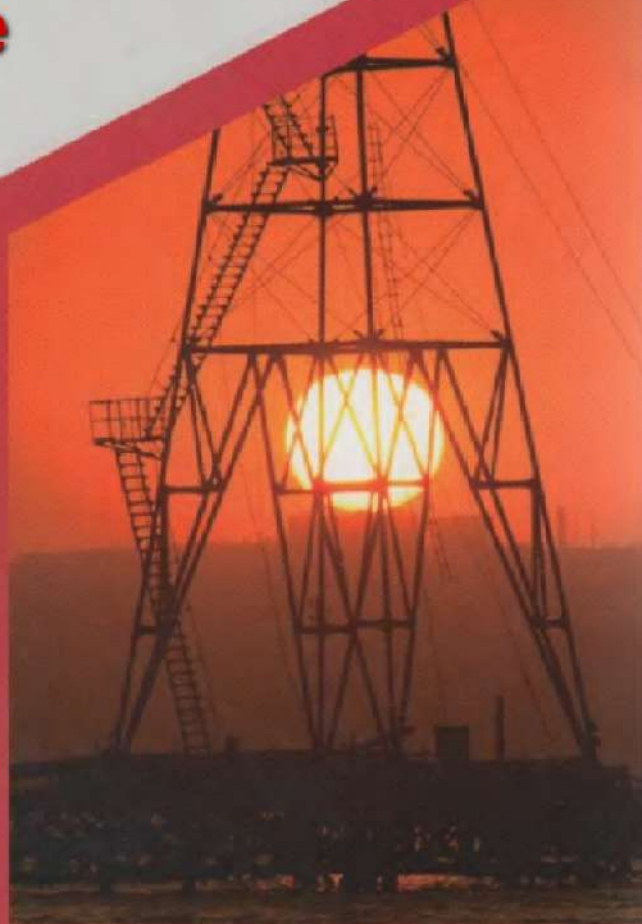


В.Г. Сазыкин, А.Г. Кудряков

ОПТИМИЗАЦИЯ СИСТЕМ ЭНЕРГОСНАБЖЕНИЯ

Учебное пособие



В.Г. Сазыкин, А.Г. Кудряков

ОПТИМИЗАЦИИ СИСТЕМ ЭНЕРГОСНАБЖЕНИЯ

Учебное пособие

(для обучающихся по направлению подготовки «Агроинженерия»
профиль подготовки «Информационные технологии и
автоматизированные системы управления»).

Краснодар
КубГАУ
2017

УДК 621.311. 1(075)
ББК 31.19
С 14

Р е ц е н з е н т ы :

И.И. Надтока – д.т.н., профессор (Южно-Российский государственный технический университет – Новочеркасский политехнический институт);
В.А. Атрощенко – д.т.н., профессор (Кубанский государственный технологический университет).

С14 В.Г. Сазыкин, А.Г. Кудряков. Оптимизации систем энергоснабжения: учебное пособие для вузов. – Краснодар: КубГАУ. – 2017. – 210 с.

ISBN 978-5-9907812-5-2

Учебное пособие выполнено в соответствии с компетентностной концепцией федерального государственного образовательного стандарта подготовки магистрантов, обучающихся по направлению 35.04.06 «Агроинженерия» профиль подготовки «Информационные технологии и автоматизированные системы управления» по дисциплине «Оптимизации систем энергоснабжения».

Учебное пособие необходимо при формировании профессиональных компетенций в процессе аудиторных занятий и индивидуальной работы при изучении отдельных глав, тем или вопросов лекционного курса и нормативных документов; подготовке к очередным лекциям и практическим занятиям; выполнении индивидуальных заданий, контрольных, курсовых и выпускных квалификационных работ и проектов; подготовке к индивидуальному собеседованию с преподавателем, к зачету, тестированию; участию в выполнении программы научно-исследовательской работы.

УДК 621.31: 1(075)
ББК 31.19

ISBN 978-5-9907812-5-2

© Сазыкин В.Г., Кудряков А.Г., 2017
© ФГБОУ Кубанский государственный аграрный университет имени И.Т. Трубилина, 2017

ВВЕДЕНИЕ

Качественное и надежное энергоснабжение с его последующим рациональным и экономичным использованием – залог успешной работы любого предприятия. Одним из главных факторов в этом является организованная, продуманная и эффективная система энергоснабжения. Современный этап развития энергетического комплекса в России предопределяет необходимость разработки механизмов и инструментов, использование которых позволит обеспечить максимально эффективное функционирование энергосистем. Все большее значение отводится решению задач, связанных с оптимизацией функционирования систем энергоснабжения, которые должны обеспечивать высокий уровень технологической безопасности, качественное решение социальных задач, высокую экономическую эффективность.

При проектировании и эксплуатации систем энерго- и электроснабжения приходится иметь дело с многовариантными задачами, в которых из некоторого множества допустимых по техническим условиям решения нужно выбрать одно, которое является лучшим по какому либо критерию.

Оптимальное управление развивающимися и саморазвивающимися энергетическими системами, применительно к энергетике, требует изучения закономерности их оптимального управления, развития и функционирования.

Важной составной частью управления энергосистемой является ведение оптимальных режимов. Применение методов оптимизации состоит, прежде всего, в отыскании допустимых, удовлетворяющих условиям надежности энергоснабжения и надлежащего качества энергии, режимов. Не меньшее значение имеет выбор из числа допустимых режимов наиболее экономичного, что позволяет практически без дополнительных затрат обеспечить экономию за счет снижения расхода топлива и потерь в электрических и тепловых сетях.

Оптимизация режимов работы энергосистемы – сложная многоуровневая иерархическая задача. Оптимальный режим энергосистемы – это такой режим из допустимых, удовлетворяющих условиям надежности и качества электроэнергии, при котором обеспечивается минимум затрат при заданной в каждый момент времени нагрузке потребителей. Исключительная сложность оптимального управления режимами определяется не только большим количеством управляемых элементов, но и тем, что разные регулируемые и настраиваемые параметры требуется поддерживать оптимальными на большой территории.

Целью учебного пособия является помощь магистрантам в освоении системы знаний о фундаментальных и прикладных возможностях методов оптимизации систем энергоснабжения, в том числе и агропромышленного направления, ознакомление их с математическими методами описания, исследования и оптимизации электроэнергетических процессов и энергетического оборудования.

1. СЛОЖНЫЕ ЭЛЕКТРОЭНЕРГЕТИЧЕСКИЕ СИСТЕМЫ

1.1. Понятие системы. Общие определения

Система представляет собой совокупность элементов, являющаяся объектом исследования, изучения или наблюдения. Элементами могут быть физические объекты (оборудование, машины, приборы, здания и т.п.), явления (нагревание, охлаждение, свечение, электромагнетизм), процессы, в том числе и технологические (регулировка, ремонт и т. п.) [3, 5, 7, 12, 17].

Элемент системы – ее неделимая часть в рамках конкретного исследования, реализующая конкретные функции. Элемент системы описывается множеством различных характеристик, параметров, связями с соседними элементами. Связи между элементами делают систему единым целым. Элементы отличаются друг от друга выполняемыми функциями, состояниями, входами и выходами. Любой элемент может рассматриваться как более мелкая система.

Термин «система» появился в научной литературе давно и является таким же неопределенным, как термины «множество» или «совокупность». Наиболее широко этот термин первоначально использовался в механике, где обозначал материальную систему, т.е. совокупность материальных точек, подчиненных определенным связям. В дальнейшем понятие системы было распространено на биологические, экономические, технологические и другие объекты.

Система – понятие относительное. Некоторая совокупность элементов может быть частью более крупной системы, небольшой ее частью или рассматриваться самостоятельно, не зависимо от окружающего мира. Это зависит от цели исследования. Для установления системы, сферы ее действия необходимо выявить ее границы и состав. При установлении границ системы выявляются причинно-следственные взаимосвязи между ее элементами.

Для выделения системы требуется определить:

- цель, для достижения которой формируется система;
- объект исследования, состоящий из множества элементов, связанных с точки зрения цели в единое целое системными признаками;
- субъект исследования, наблюдения, заказчика, формирующего систему;
- характеристики внешней среды по отношению к системе и отражение их взаимосвязей с системой.

Цель функционирования определяет системные признаки, с помощью которых описываются элементы системы. Система с точки зрения цели есть упорядоченное представление об объекте (существующем или проектируемом). Разные субъекты, в зависимости от цели, могут иметь свои представления об элементах системы, их взаимосвязях и связях с внешней средой.

Цель – это субъективный образ, абстрактная модель несуществующего, но желаемого состояния производства, которое решило бы возникшую проблему.

Цели, которые ставит перед собой человек, редко достижимы только за счет его собственных возможностей, или возможностей производства, к которому он причастен.

Стечение обстоятельств, характеризующееся различием между необходимым (желаемым) и существующим, называется проблемой, или проблемной ситуацией. Проблема существующего положения, в частности с производством продукции, осознается в несколько стадий: от смутного ощущения, что «что-то не так», к осознанию потребности, затем выявлению проблемы и, наконец, к формулировке цели.

Вся последующая деятельность, способствующая решению этой проблемы, направлена на достижение поставленной цели. Эта деятельность направлена на отбор из окружающей среды элементов, свойства которых можно использовать на достижение поставленной цели, и на объединение этих элементов надлежащим образом, т.е. как работу по созданию того, что мы называем системой.

Упорядоченность представления субъекта есть целенаправленное выделение элементов системы, установление их признаков, взаимосвязей между собой и с внешней средой. При выделении системы учитывают наиболее существенные признаки, все второстепенное, несущественное – исключается.

Решение проблемы есть то, что заполняет промежуток между существующей и желаемой системами. Значение для человека имеют наглядные, образные, визуальные модели.

Для наглядного представления системы ее изображают в виде «черного ящика», выделенного из окружающей среды и имеющего входы и выходы, рис. 1.1. Название «черный ящик» образно подчеркивает полное отсутствие сведений о внутреннем содержании ящика: задаются, фиксируются, перечисляются только входные и выходные связи системы со средой. Такой подход, несмотря на его простоту и на отсутствие сведений о внутренней структуре системы, часто оказывается полезным.

Сопоставляя входы и выходы за ряд моментов времени, находят такие входные параметры X , при которых рассчитанные значения выходных параметров Y лучше всего аппроксимируют фактические значения выходов.

Сущность метода «черного ящика» состоит в том, что при исследовании объектов они рассматриваются как недоступный для наблюдения, изучения и описания «черный ящик», имеющий определенные входы и выходы. Вследствие сложности устройства «черного ящика», т.е. изучаемого объекта, возможно лишь наблюдать состояние входов в него и соответствующих им выходов, т.е. изучать поведение, не зная его внутреннего устройства.

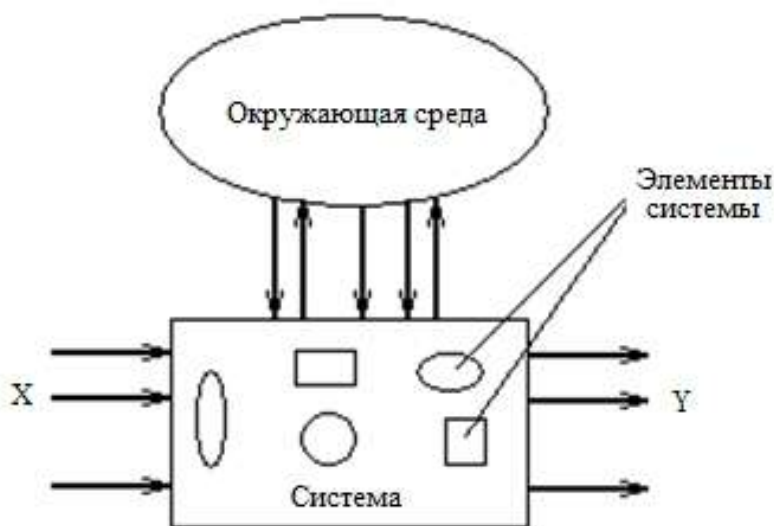


Рис. 1.1. Понятие «черного ящика»

Однако, как бы детально ни изучалось поведение «черного ящика», нельзя вывести обоснованного суждения о его внутреннем устройстве, ибо одним и тем же поведением могут обладать различные объекты, а одно и то же соотношение между входами и выходами может в пределах имеющихся статистических данных удовлетворительно описываться несколькими различными математическими выражениями. С увеличением числа факторов регрессионной модели обычно падает ее достоверность. Как показывает практика, удовлетворительные модели получаются при описании ситуации, в которой выходной фактор существенно связан не более чем с пятью-шестью входными факторами.

Во многих случаях достаточно содержательного словесного описания входов и выходов.

Различают *детерминированные* и *стохастические* системы.

В детерминированных системах цель исследования полностью определена, сами элементы и отношения между ними и внешней средой известны.

Системы со стохастической структурой не имеют либо ясно выраженной цели исследования, либо выраженных существенных элементов и отношений между ними (признаков). Подобные системы выделяются на этапах разработки, проектирования сложных производств, технологических процессов и оборудования.

Системы разделяются на *управляемые* и *неуправляемые*. Управление можно определить как организацию различных действий, процессов для достижения намеченной цели.

Управляемые системы обеспечивают целенаправленное функционирование при изменяющихся внутренних или внешних условиях. Управление осуществляется человеком или специальным устройством (для технических систем). К управляемым системам относятся, например, движение автотранспорта, работа технологической линии или предприятия в целом.

Неуправляемые системы не обеспечивают целенаправленного функционирования. К неуправляемым относятся стихийные явления природы, работа оборудования после отказа, движение ветра.

При рассмотрении, анализе и синтезе систем существуют два подхода: *индуктивный* (классический) и *системный*.

Индуктивный подход предполагает изучение системы путем перехода от частного к общему и дальнейший системы за счет слияния ее компонентов.

Системный подход предполагает переход от общего к частному при выделении исследуемого объекта из окружающей среды при единой цели.

Структуру системы можно изучать исходя из состава отдельных подсистем (структурный подход) или путем анализа функционирования отдельных свойств, позволяющих достичь заданной цели (функциональный подход).

Структурный подход позволяет выделить состав элементов системы и связи между ними. Наиболее общее описание структуры – топологическое описание на базе теории сетей и графов.

Структура системы – совокупность связей между элементами системы, отражающая их взаимодействие. Структура системы может изучаться с разных позиций – извне (состава отдельных элементов системы и отношений между ними) и изнутри (при анализе свойств системы, приводящих к намеченной цели). Связи между элементами, определяющие систему, могут быть устойчивые, неустойчивые, статистически устойчивые.

Устойчивые связи существуют постоянно в течение рассматриваемого промежутка времени или возникают регулярно.

Неустойчивые связи возникают редко, от случая к случаю.

Статистически устойчивые связи с течением времени стремятся к определенным значениям.

Связи могут определяться экономическими отношениями, физическими или социальными законами, отношениями родства, подчинения и т.д. Они могут быть функциональными, информационными, причинными, логическими и т.д.

Функциональный подход рассматривает отдельные функции, алгоритмы, приводящие к достижению цели.

Характеристики системы могут быть количественные и качественные. Количественно система характеризуется числами, выражающими отношение между заданной величиной (эталоном) и исследуемой величиной. Качественные характеристики выражаются описанием типа хороший, плохой, больше, меньше или с помощью различных шкал, например методами экспертных оценок.

Функционирование системы – проявление функций системы во времени, переход от одного состояния к другому (движение в пространстве состояний). При использовании системы важно качество ее функционирования. Один и тот же закон функционирования может быть реализован с помощью различных алгоритмов. Процесс функционирования можно рассматривать как последова-

тельную смену состояний, Совокупность всех возможных значений состояний системы называют пространством состояний системы.

Внешняя среда – множество существующих вне системы элементов любой природы, оказывающих влияние на систему или находящихся под ее воздействием. Внешняя среда определяет условия функционирования системы посредством воздействия *внешних факторов*, являющихся движущей силой процесса и определяющих характеристики этого процесса. В зависимости от цели внешние факторы могут быть стимулирующими, регулируемыми, ограничивающими, возмущающими и разрушающими.

Стимулирующие факторы стимулируют развитие процесса.

Регулирующие, управляющие факторы приводят к изменению целей, режимов и алгоритмов функционирования системы.

Ограничивающими факторами являются различные нормативно-правовые акты, законы, нормы поведения, технические условия, регламенты и стандарты функционирования технологических процессов и технических систем.

Возмущающие факторы – это отрицательные факторы, негативно влияющие на работу системы, достижение ее цели. Эти факторы можно спрогнозировать и компенсировать.

Разрушающие факторы – это отрицательные факторы, которые сложно спрогнозировать, а значит, и предотвратить. Они приводят к частичному или полному уничтожению системы.

Отношения между элементами системы и системой определяются их иерархией.

Иерархия – это упорядоченная по старшинству совокупность элементов и подсистем, входящих в данную систему, например, завод – цех – участок – линия – аппарат.

В *иерархической системе* объект расчленяется на уровни согласно принципу подчинения низших уровней высшим. Степень декомпозиции будет определяться как спецификой решаемой задачи, так и имеющейся информацией об объекте.

Иерархическая организация, конечно, не является исключительной особенностью сельского хозяйства – такой подход к структурированию приложим к самым разнообразным системам – коммерческим предприятиям, комплектам компьютерных программ, социальному устройству, электронному оборудованию и т. п.

Объекты, принадлежащие каждому структурному уровню, могут рассматриваться и как системы, образованные из подсистем (объекты более низких уровней), и как подсистемы, входящие в состав некоторой системы (объект более высокого уровня).

Для иерархических систем характерны три важных свойства:

1) каждый уровень иерархии имеет свой собственный язык, свою систему концепций или принципов;

2) на каждом уровне иерархии происходит обобщение свойств объектов более низких уровней. Закономерности, обнаруженные и описанные для по-

следних, могут быть включены в объясняющую (функциональную) схему, обратая при этом связь с объектом высшего уровня. Таким образом, описание на уровне i способствует объяснению (пониманию) явлений, имеющих место на уровне $i-1$;

3) взаимосвязи между уровнями не симметричны. Для нормального функционирования объектов высшего уровня необходимо, чтобы успешно «работали» объекты более низкого уровня, но не наоборот.

Однако главная задача при этом – выбрать компоненты системы таким образом, чтобы каждому из них была присуща относительная автономия, то есть, чтобы внутренние связи в пределах каждой подсистемы были сильными, а взаимодействия между подсистемами – слабыми. Обычно решающим оказывается то обстоятельство, что подсистемы, подлежащие рассмотрению, должны быть хорошо изучены и описаны.

1.2. Иерархия энергетических систем

В настоящее время энергетика существует в виде большой и сложной целенаправленной управляемой системы. Обычно под *системой* понимают то, что определенным образом взаимосвязано, подчинено общим целям и законам, может быть описано в рамках структуры, которая отражает устойчивые отношения, связи и взаимодействия между элементами и частями, объединяющими несколько элементов или частей – *подсистем*.

Система – целостная совокупность взаимосвязанных элементов и частей, средство достижения поставленной цели [5, 17]. Понятие *система* является многогранным и развивающимся. Элемент и часть, часть и подсистема, подсистема и система образуют диалектические противоположности категории часть и целое. Связи в системе также являются элементами, частями или подсистемами, назначение которых объединять другие элементы, части или подсистемы в единое целое, в систему. *Структура системы* – совокупность всех элементов, частей и подсистем и связей между ними обеспечивает целостность и устойчивость при различных внешних и внутренних изменениях.

Для возникновения системы необходимо выполнение трех условий:

1) объединение частей или объектов, между которыми существуют *каузальные* – причинно-следственные связи;

2) наличие методов, способов и средств, т.е. подходящих связей для объединения частей или объектов в систему;

3) появление у системы новых эмерджентных свойств, которые отдельным элементам, рассматриваемым вне системы, не присущи. *Эмерджентность* характеризует эффект организации, который является результатом возникновения между элементами усиливающих связей, проявлением закона перехода количества в качество.

Большая система энергетики состоит из отдельных взаимосвязанных подсистем, функционирующих как единое целое, но при решении самостоятельных

задач подсистемы сами становятся системами и рассматриваются отдельно [2, 3, 5, 12, 17].

Составной частью энергетики является энергетическая система. Все элементы системы функционально связаны единством генерирования, передачи и потребления электрической энергии. *Энергетическая система* (энергосистема) – совокупность электростанций, электрических и тепловых сетей, соединенных между собой и связанных общностью режима в непрерывном, процессе производства, преобразования и распределения электрической энергии и тепла при общем управлении этим режимом. В свою очередь и эта подсистема является иерархичной и включает электроэнергетическую систему.

Электроэнергетическая система (ЭЭС) – та часть энергетической системы, в которой теплота и различные виды энергии преобразуются в электрическую энергию, передаваемую на расстояние, распределяемую по потребителям, где она вновь преобразуется в другие виды энергии. Одну из основных функциональных составляющих энергосистемы выполняет ее электрическая часть. Электрическая часть энергосистемы играет роль одной из подсистем большой энергетики, которая также обладает системными признаками. *Электрическая часть энергосистемы* – совокупность электрического оборудования объектов энергосистемы [9].

Для оптимального функционирования сложной системы необходимо осуществлять управление. *Управление* – это совокупность воздействий на систему, ее части или объекты с целью достижения поставленной цели. В сложной системе управление реализуется через *принятие решений*.

Итак, ЭЭС – это находящееся в работе электрооборудование энергосистемы и приемников электрической энергии, объединенное общим режимом и рассматриваемое как единое целое в отношении протекающих в нем физических процессов. ЭЭС представляет собой совокупность *силовых элементов* – вырабатывающих, преобразующих, передающих, распределяющих и потребляющих электрическую энергию и *элементов управления* – регулирующих и изменяющих состояние системы. ЭЭС является сложной автоматизированной системой, совокупностью *управляемого объекта* и *субъекта управления* – автоматических управляющих устройств, где часть функций управления (наблюдение за работой автоматических устройств, изменение при необходимости программы их работы, принятие решений по управлению в изменившихся или сложных ситуациях) выполняет человек. Обязательное присутствие человека в системе управления, характеризует ее как *эргатическую*.

С целью выбора управляющего воздействия на систему предварительно необходимо проанализировать возможности и последствия разных действий. Такая возможность предоставляется при использовании моделей ситуации. Осуществление моделирования требует затрат определенных ресурсов, как правило, ограниченных. В зависимости от различия в удовлетворении потребности моделирования в ресурсах конкретизируются понятия *больших* и *сложных систем*. При моделировании, например на ЭВМ, ограничены ресурсы объема памяти и скорость решения задач большой размерности в реальном време-

ни. Системы, моделирование которых затруднительно вследствие их размерности, называются *большими* [17].

Имеющаяся о системе информация предстает в виде той самой модели, об использовании которой идет речь. Признаком простоты системы является успешность управления, т.е. *достаточность* информации для управления. Но если полученные с помощью модели управления результаты приводят к неожиданным, непредвиденным или нежелательным последствиям, отличающимся от предсказанных моделью, это характеризуется как *сложность* системы и объясняется недостаточностью информации для управления.

Обычно система состоит из элементов с внутренними и внешними связями. Если учитывать только двусторонние внутренние одномерные связи между элементами, то общее число связей выражается квадратичной зависимостью числа управляемых переменных от количества элементов системы. Обеспечение требований по качеству связей предполагает возможность выбора номинальных значений для каждой управляемой переменной, поэтому число управляемых переменных с учетом возможностей выбора эталонов еще более возрастает. Число способов управления объемом информации, переносимой сигналами, растет как квадратичная функция количества используемых сигналов. В сложных системах такой объем информации в процессе управления переработать практически невозможно.

Сложной системой называется система, в модели которой не хватает информации для эффективного управления. Из-за влияния неопределенных в полной мере и случайных факторов на сложную систему модель ее управления неадекватна заданной цели, поэтому большую систему электроэнергетики определяют как сложную систему.

В сложных системах выделяются специфические контуры управления, по которым циркулируют потоки информации: исходной – от элементов системы к управляющим устройствам и управляющей – от управляющих устройств к элементам системы.

Сложным системам свойственны в той или другой степени черты самоорганизации. Система называется *самоорганизующейся*, если она способна на основании оценки воздействий внешней среды, путем последовательного изменения своих свойств прийти к некоторому устойчивому состоянию, когда воздействия внешней среды окажутся в допустимых пределах. Самоорганизация системы связана с ее *адаптацией*, приспособлением к условиям окружающей среды. Развитие системы, изменение во времени ее состояния свидетельствует о ее *динамичности*.

Следовательно, основными отличительными признаками сложных систем электроэнергетики являются:

- 1) большое количество организованных и управляемых частей и элементов, эргатичность;
- 2) сложность и многовариантность функций, выполняемых системой и направленных на достижение общей цели функционирования;
- 3) возможность разбиения системы на подсистемы – цели, функционирование которых подчинено общей цели функционирования всей системы;

- 4) иерархическая структура управления, разветвленная информационная сеть и потоки информации;
- 5) взаимодействие с внешней средой и функционирование в условиях воздействия случайных факторов и неполноты информации;
- 6) самоорганизация, адаптивность и динамичность.

1.3. Принятие решений при управлении сложными системами

В основе управления сложными системами заложено постоянное решение множества вопросов, задач, проблем. Выработка и принятие решения относится к информационным процессам, связанным с переработкой субъектом управления информации об окружающей среде, состоянии объекта управляемой системы, накопленного опыта и знаний, которые заканчиваются волевым, творческим действием лица, принимающего решение (ЛПР) – принятием решения [2–4, 10, 18].

Присутствие в процессе управления субъекта – ЛПР вносит в процесс формирования и выбора рациональных решений обогащенный применением объективных математических методов элемент субъективизма. Понимание субъективной природы решений способствует поиску путей обогащения возможностей и способностей ЛПР. К этому привлекаются математические и экспертные модели и методы, позволяющие достичь в решении оптимального сочетания субъективного и объективного.

Принятие решения характеризуется следующей ситуацией:

- 1) наличием проблемы – цели, необходимой для достижения;
- 2) наличием участников процесса принятия решения, несущих ответственность за принятое решение;
- 3) наличием альтернатив выбора линий поведения. Если выбора нет, то решение не требуется и predetermined действия могут быть выполнены автоматически;
- 4) риском и неопределенностью использования различных альтернатив, с которыми связана вероятность достижения цели. Риск и неопределенность являются частыми, но не обязательными чертами ситуации принятия решения;
- 5) присутствием ограничений на значения показателей, которыми описывается альтернатива, один из вариантов решения. Наличие ограничений, как правило, затрудняет достижение поставленной цели.

Выявление всех элементов ситуации должно предшествовать непосредственному принятию решения.

Процесс принятия решения – это уникальное многоплановое явление с присущими ему закономерностями и особенностями. Во-первых, это информационный процесс, связанный с переработкой информации субъектом управления. ЛПР на основе знания основных закономерностей функционирования объекта управления, анализа и оценки информации об окружающей среде, состоянии управляемой системы и предшествующего опыта вырабатывает новую ин-

формацию в виде решения. Во-вторых, это волевое, творческое действие ЛПР, которое завершает психологический процесс отражения действительности интеллектом субъекта управления, где доводы находятся на одном уровне с эмоциями, а логика с интуицией.

Рациональное решение зависит от накопленного ЛПР опыта и обосновывается объективным анализом с применением количественных и качественных математических методов. При этом в описании вариантов решения используется структурированное множество показателей, а для выбора оптимального решения – набор критериев.

Реализация всевозможных вариантов решения обычно связана с различными последствиями. Для оценки вариантов решения используются критерии эффективности. Вариант считается оптимальным, если обеспечивается максимизирующее значение функции при заданных ограничениях или минимизируются ресурсы для достижения заданного значения целевой функции.

Технология принятия решения представляет собой структуру выбора направления и способов действия для достижения цели и состоит из следующих основных стадий:

- 1) формирование подмножества вариантов решения из всего множества, универсума возможных решений;
- 2) прогнозирование развития событий в окружающей среде и результатов функционирования системы в произвольный момент времени в будущем;
- 3) построение модели функционирования системы в условиях окружающей среды и использование ее для прогноза последствий решения;
- 4) сравнение ожидаемой полезности разных вариантов многокритериального решения;
- 5) выбор наилучшего варианта решения из сформированного подмножества при наличии противоречивых оценок по нескольким частным критериям;
- 6) составление плана конкретных действий на основе прогноза развития обстановки, в которой план будет исполняться.

Процесс выработки и принятия решения заключен в нескольких типичных этапах:

- 1) уяснение задачи, оценка оперативной обстановки, выполнение расчетов, обоснование замысла и решения;
- 2) принятие решения, его оформление, доведение до сведения непосредственных исполнителей;
- 3) разработка плана и необходимых документов, регламентирующих выполнение принятого решения, утверждение плана;
- 4) контроль исполнения решения, анализ и обобщение опыта.

Процесс выработки и принятия решения может быть циклическим, итерационным и содержать повторяющиеся действия. Итерации, как правило, необходимы для получения новой или уточненной информации.

1.4. Особенности системы электроснабжения

Анализ свойств объекта управления сложной электрической системой осуществим на одной из ее подсистем – системе электроснабжения (СЭС) крупного предприятия (рис. 1.2), также являющейся сложной системой. Для этого воспользуемся методологией, терминами и определениями системного анализа [2, 3, 12, 17], представляя СЭС как проблемосодержащую систему предметной области.

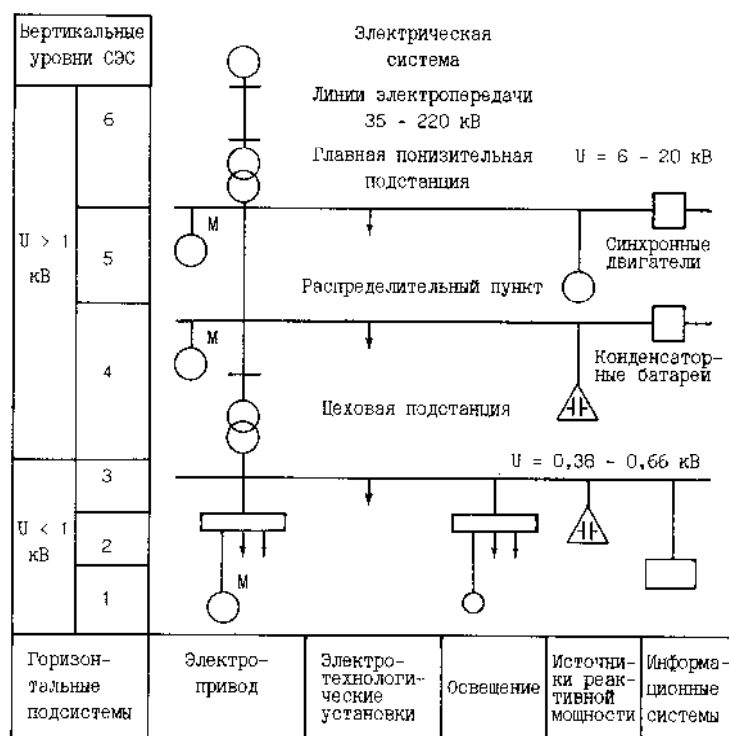


Рис. 1.2. Иерархия уровней-подсистем системы электроснабжения:
1–6 – уровни электрических нагрузок

В СЭС района или крупного предприятия первоначально выделяются две подсистемы:

1) объект управления (подсистема исполнения решений – совокупность технологически взаимосвязанных производственных объектов и агрегатов, электроустановок, материальных ресурсов и производственного персонала);

2) субъект управления (подсистема принятия решений – АСУЭ: управление отделом главного энергетика производственно-хозяйственной деятельностью своих подразделений; подразделений оперативно-диспетчерского управления; средств автоматики и телемеханики). С точки зрения субъекта предметом управления являются информационные процессы, происходящие в контролируемой области.

СЭС района или современного предприятия представляет собой совокупность иерархически организованных различных по характеру подсистем. Ядром

является технология или технологический процесс. Подчинены технологии электроприемники потребителей. Связывающее звено в иерархии от источников питания и электрической системы до электропотребителей образуют линии электропередачи, трансформаторные подстанции, преобразователи, источники реактивной мощности. Подсистемы СЭС связаны между собой как технологией, так и единством передачи, распределения, преобразования и потребления электрической энергии. Поэтому СЭС является подсистемой всего промышленного комплекса.

Цель рассматриваемой системы установлена в определении СЭС как совокупности электроустановок, предназначенных для обеспечения потребителей электрической энергией, осуществляющих производство, преобразование, передачу, распределение, потребление и превращение электрической энергии в другие виды энергии.

Уровни-подсистемы СЭС имеют вертикальную и горизонтальную структуру (рис. 1.1).

Вертикальные уровни. В соответствии с Правилами устройства электроустановок (ПУЭ) [19] все электроустановки принято делить на два уровня напряжением U до 1 кВ и выше. При определении электрических нагрузок СЭС представляется иерархической системой: 1 – отдельные электроприемники; 2 – группы электроприемников; 3 – шины цеховой подстанции напряжением 0,38–0,66 кВ; 4 – шины распределительного пункта напряжением 6–20 кВ; 5 – шины главной понизительной подстанции напряжением 6–20 кВ; 6 – граница раздела промышленного предприятия и энергосистемы.

Горизонтальные подсистемы (назначение электроприемников) включают электропривод, электротехнологические установки, электрическое освещение, источники реактивной мощности, устройства для сбора и обработки информации, а также устройства управления.

Практически все используемые уровни напряжений до и выше 1 кВ переменного и постоянного тока опасны для жизни людей. Опасность электрического тока в отличие от прочих опасностей усугубляется быстротечностью поражения и тем, что человек не в состоянии без специальных приборов дистанционно обнаружить напряжение. Промышленная частота переменного тока 50 Гц является одной из наиболее опасных частот для поражения людей. Действующие электроустановки являются источниками шума, ультразвука, инфразвука, вибрации, окружены электрическими, магнитными и электромагнитными полями, представляют опасность возникновения пожара и взрыва.

СЭС обеспечивает бесперебойное электроснабжение в пределах допустимых показателей качества электроэнергии (ПКЭ) и исключения опасных для жизни людей и окружающей среды ситуаций при соблюдении условий безотказности, долговечности, устойчивости, управляемости, безопасности, ремонтпригодности и избыточности входящих элементов. На практике надежность СЭС обеспечивается техническими и организационными мероприятиями (дублированием элементов, их функций, подсистем, использованием при управлении избыточной информации, совершенствованием конструкций, материалов и

операций технического обслуживания, снижением продолжительности аварийных ремонтов, совершенствованием систем контроля и управления).

СЭС ухудшает состояние окружающей среды за счет повышения концентрации вредных веществ в атмосфере, воде, почве, нарушении теплового баланса, природного ландшафта, увеличении уровней шума, вибрации, напряженности электромагнитного поля и т.п. При аварийных ситуациях в СЭС перечисленные факторы многократно возрастают. Особую роль приобретают социальные последствия экологических нарушений.

В последние годы часть экологических проблем в рамках Международной конференции по большим энергетическим системам – CIGRE сформировалась в отдельное направление исследований – электромагнитную экологию. Электромагнитная экология связана с теми явлениями в электроэнергетических системах, которые оказывают вредное и опасное влияние на другие системы – связь, электронное оборудование, трубопроводы, живые организмы людей, животных и растений. Учитываются как сами электромагнитные процессы в электроустановках, так и последствия их влияния на окружающую среду. Основные вопросы этого направления относятся к проблемам электромагнитной совместимости (ЭМС) [8]: изучение влияния короны высоковольтных линий электропередачи и электромагнитных полей на вспомогательное электронное оборудование, линии связи, протяженные металлические сооружения и низковольтные линии электропередачи.

Экономичность функционирования СЭС (минимум затрат и максимум эффекта) достигается путем оптимального соотношения между затратами на оборудование и технико-экономическими последствиями нарушения требований к обеспечению электроснабжения. Оптимальное соотношение находится в результате сопоставления нескольких возможных технических решений и выбора наиболее целесообразного варианта воплощения.

Взаимное влияние электрооборудования, связанного электрической сетью в СЭС, породило комплекс проблем ЭМС. Нарушение ЭМС проявляется в экономическом ущербе, ухудшении ПКЭ и коэффициента мощности, увеличении нагрева элементов и потерь мощности в сети и самом электрооборудовании, возникновении резонансных явлений, ухудшении диэлектрических свойств изоляционных материалов, сокращении срока службы и надежности элементов СЭС, уменьшении полезной мощности электроприемников, сбоях в работе, нарушениях нормальных режимов электро- и технологического оборудования, влиянии на автоматику, линии связи, информационные потоки. Главным техническим аспектом нарушения ЭМС является связанность всех элементов СЭС. Искажения и помехи со стороны низшего напряжения проникают в сети среднего и высшего напряжения, сети других потребителей и влияют на работу всего электрооборудования. Внедряется принципиально новое электрооборудование, которое можно разделить на помехочувствительное (электронная и цифровая техника, являющаяся главным элементов средств автоматики, вычислительных машин и информационных систем) и производящее помехи (силовая полупроводниковая техника, электроустановки с элегазовым оборудованием, лазерные принтеры и персональные компьютеры, потребители, имеющие бес-

трансформаторные схемы выпрямителей с емкостным сглаживанием пульсаций и другие установки). Неожиданные физические эффекты оказывают отрицательное воздействие, например, выпрямление переменного тока окислами проводников в местах электрического контакта. Недостаточна приборная база для выявления, контроля и диагностики искажений и помех.

СЭС обладают рядом свойств (присущих большим и сложным искусственным человеко-машинным системам кибернетического типа), которые можно разделить на общие, структурные и режимные.

Общие свойства СЭС:

- иерархичность системы показателей;
- неожиданность, необъяснимость, уникальность и случайность поведения, т.е. нехватка информации для ее управления по единой модели, адекватной заданной цели. Поэтому для каждой подсистемы или группы ее элементов используется собственная модель;
- эмерджентность, целостность обладания качествами, отсутствующими у отдельно взятых ее частей;
- состояние СЭС в любой момент времени не определено системой показателей достаточно точно. Возникает парадокс: чем больше детерминированных параметров и точнее каждый из них, тем менее точен образ-характеристика описываемой системы;
- управляемость параметров режима и параметров системы, с помощью изменения которых СЭС способна изменять внутренние параметры;
- наблюдаемость, заключающаяся в том, что внутренние параметры системы связаны, по крайней мере, с одной, позволяющей их анализировать, выходной величиной;
- идентифицируемость состояний СЭС, способствующая выявлению нужной информации для контроля и управления как отдельными элементами и группами, так и всей СЭС;
- вероятностная предсказуемость (неопределенность, нечеткость) позволяет, используя коэффициент определенности или функцию принадлежности [11, 14], представить описание поведения СЭС на некотором временном интервале будущего;
- прогнозируемость как составная часть планирования дает возможность судить о состоянии СЭС в будущем или о путях и сроках достижения планируемых состояний;
- кибернетическая моделируемость, подразумевающая рассмотрение СЭС в виде преобразователя, имеющего входы и выходы. Например, при моделировании СЭС ее моделью может быть любая другая система, в которой совершается подобное преобразование состояний входов в состояния выходов;
- целенаправленность (адаптивность, саморегулирование), проявляющаяся в способности СЭС устранять последствия внешних и внутренних воздействий для достижения необходимого состояния благодаря наличию средств резервирования, управления и автоматики.

Структурные свойства СЭС:

- многомерность, определяемая наличием и необходимостью измерения параметров и показателей;
- многосвязанность элементов, имеющих прямые, обратные, рекурсивные (необратимые, причинно-следственные), синергические (усиливающие), циклические, жесткие и гибкие связи;
- неоднозначность (нечеткость) структуры как различных функций каждого уровня СЭС, так и назначения каждого из элементов; дискретность физической неделимости элементов СЭС (генераторов, трансформаторов, электродвигателей и проч.) и показателей системы в целом, которые носят дискретный характер.

Режимные свойства СЭС:

- одновременность процессов производства (для промышленных предприятий, имеющих собственные электростанции), передачи, распределения и потребления электрической энергии;
- непрерывность всего функционального процесса в СЭС от выработки до электропотребления;
- взаимосвязь между производством и потреблением с учетом потерь в элементах СЭС;
- случайность и многообразие возмущений, что характеризует СЭС как открытую систему, связанную с внешней средой, технологическим процессом, атмосферными воздействиями и проч.;
- динамичность изменения во времени параметров режима и системы;
- наблюдаемое соответствие между физическими и информационными процессами.

Режимы СЭС подразделяются на *установившиеся и переходные*. Внутри этих групп различаются нормальные, утяжеленные, аварийные и послеаварийные виды режимов.

В *нормальном* режиме работают все необходимые элементы СЭС, ПКЭ не должны выходить за пределы максимальных значений, указанных в ГОСТ 13109-97 [8], при этом в течение не менее 95 % времени каждых суток ПКЭ должны находиться в пределах нормальных значений. Разброс параметров в результате «дыхания» системы – нечеткость (размытость) параметров находится в пределах, установленных [5].

В *утяжеленном* режиме появляется повышенная опасность дальнейших ухудшений параметров режима и системы.

При *аварийных* нарушениях электроснабжения допускается кратковременный выход значений ПКЭ за установленные максимальные границы с последующим восстановлением до границ послеаварийного режима.

В *послеаварийном режиме* СЭС происходит изменение параметров системы, а параметры режима характеризуются значениями ПКЭ, не выходящими за максимальные границы, установленные ГОСТом [8].

Особую опасность для СЭС представляют аварийные режимы [3, 5, 12], последствия которых отражаются на других промышленных подсистемах и окружающей среде. При высокой скорости протекающих процессов ликвида-

ция аварий возможна только с помощью быстродействующих устройств системной автоматики.

Последствиями аварий в зависимости от тяжести являются: ухудшение условий электробезопасности; понижение надежности электроснабжения; нарушение технологии производства; загрязнение окружающей среды; тяжелые социальные последствия экологических нарушений; перегрузка неповрежденных элементов; нарушение устойчивости работы и, как следствие, возможность каскадного характера развития аварийного процесса; материальный ущерб от простоев, недовыпуска продукции, увеличения потерь, повреждения агрегатов, ухудшения параметров СЭС и ее элементов; изменение конфигурации схемы электроснабжения.

Причинами аварийных режимов являются: случайные факторы; нарушения нормальных режимов, работоспособности и разрушения элементов СЭС; развивающиеся дефекты, которые не могут быть выявлены в промежутках между плановыми ремонтами; отказы исполнительных механизмов; сложность и территориальная распределенность СЭС; инерционность и неадекватность управления; поступление ложной информации и отсутствие информации о реально происходящих процессах. Для оптимального осуществления нормальных режимов, предупреждения аварийных ситуаций, а также быстрого устранения их последствий нужны сложные системы автоматического и автоматизированного управления, что требует системного рассмотрения и анализа субъекта управления.

1.5. Особенности автоматизированного управления в СЭС

Наибольшее развитие в энергетике получили автоматизированные системы управления – АСУ производственными энергетическими объединениями – предприятиями для диспетчерского, производственно-технического и организационно-экономического управления [2–4, 6, 12]. Входящие в состав АСУ производственных энергетических объединений структурные единицы: электростанции, электрические и тепловые сети, СЭС крупных промышленных предприятий, имеют собственные АСУ, реализующие информационные, управляющие и вычислительные функции. Крупные промышленные предприятия, потребляющие в большом количестве различные виды энергии, имеющие развитое и сложное энергетическое оборудование, создают автоматизированные системы управления энергоснабжением – АСУЭнерго (АСУЭ) [2, 6].

АСУЭ состоит из двух основных подсистем – автоматизации управления СЭС и автоматизации управления основным энергооборудованием промышленного предприятия. Вторая подсистема подчиняется технологии производства – АСУТП, но согласовывает с энергодиспетчером основные управляющие решения, связанные с эксплуатацией присутствующего электрооборудования. В связи с этим обстоятельством, а также учитывая, что подсистема АСУЭ СЭС по сравнению с системами, предназначенными для автоматизации обеспечения другими видами энергии, наиболее полно отвечает предъявляемым к автомати-

зированной системы управления требованиям, в дальнейшем изложении остановимся на ней.

АСУЭ характеризуется [2, 4, 6] наличием системы автоматизированного сбора и обработки информации; необходимостью выработки и использования в управлении информации, получаемой расчетным путем; возможностью частичного автоматического управления; применением математических методов управления на базе управляющих ЭВМ; автономностью входящих в нее подсистем, имеющих самостоятельные цели управления и общую цель, единую для всей АСУЭ; наличием внутренних и внешних связей у каждой подсистемы; уплотнением информации при восходящем иерархическом движении; необходимостью согласования целей, критериев управления, процедур кодирования и обработки информации между отдельными подсистемами как в пределах самой АСУЭ, так и с подсистемами производства – АСУП и технологических процессов – АСУТП. Система ЛПР – ЭВМ обеспечивает достаточно эффективное функционирование объекта, когда сбор и переработка информации, необходимой для реализации функций управления, осуществляется с применением методов и средств автоматизации и вычислительной техники. Все АСУ имеют техническое, математическое, программное, лингвистическое, информационное, организационное и правовое виды обеспечения.

В энергосистеме и ее подсистемах, ЭЭС и СЭС субъектом в управлении сложными процессами является диспетчер – ЛПР [4, 6]. ЛПР обычно работает в условиях ограниченности времени для анализа большого объема поступающей оперативной информации, часть из которой может быть неполной. В процессе решения задач диспетчерского управления ЛПР выполняет функции: планирования и контроля, выработки регулировочных воздействий и устранения нарушений контролируемого процесса, мониторинга и диагностики системы.

Нередко возникают проблемные ситуации, когда проявляется противоречие между запланированным и реальным ходом контролируемого процесса, а у ЛПР отсутствует четкое представление о том, что нужно делать для ликвидации возмущений в системе. Возникновению подобных ситуаций способствуют и специфичные для диспетчерского труда условия: эмоциональный фон, порожденный огромной ответственностью за результаты деятельности, а также недостаток времени на принятие решения. Под влиянием отрицательного эмоционального состояния и высокого темпа принятия решений у ЛПР иногда выпадают из памяти нужные ему инструкции решения подобных задач, поэтому он вынужден конструировать их вновь.

Структура деятельности диспетчера-ЛПР основана на оперативном мышлении. Прежде чем отреагировать на проблемную ситуацию, ЛПР мысленно представляет элементы, из которых складывается ситуация, затем приводит в движение образы этих элементов и на основе такого перемещения представляет совокупности будущих действий. Диспетчер учитывает большое количество как взаимосвязанных, так и противоречивых факторов. Он должен принимать решения в условиях отсутствия четких критериев их оценки.

Принятие и реализация управляющих воздействий решаются тем сложнее, чем сложнее объект управления, чем больше управляемых параметров и связей между ними. В АСУЭ связи между параметрами значительно усложнились, они стали менее определенными. В этих условиях возникает альтернативность возможных решений со своими положительными и отрицательными сторонами, что затрудняет однозначность окончательного выбора. Жесткость предъявляемых к СЭС требований, заключающихся в обеспечении безопасности, надежности, экологичности, экономичности и ЭМС, вызывает необходимость автоматизированного управления жизненным циклом всей системы. С позиции субъекта управления СЭС имеет непрерывный характер управления, осуществляемый как с помощью средств системной автоматики, так и АСУЭ.

Задачами системной автоматики является управление при электромагнитных и электромеханических переходных процессах с помощью устройств релейной защиты и автоматики.

СЭС связана с АСУЭ информационной подсистемой, которая обеспечивает сбор и регистрацию значений параметров режима и системы для последующего адекватного управления. Управление в автоматизированном режиме осуществляется дистанционно периодически с помощью команд. Постоянно производится передача измеренной информации к АСУЭ для оценки функционирования СЭС. Кроме набора показаний контрольно-измерительных приборов выделяется визуальная информация, которая не фиксируется приборами, а поступает от специалистов, наблюдающих за объектом. В общем случае вся объективная и субъективная информация характеризует определенные параметры режима и системы. В результате оценки информации по мере необходимости ЛПР принимает решения о характере и силе управляющих воздействий, которые оперативно передаются для реализации в нужном месте.

Имеется ряд задач оперативного управления параметрами режима и системы, интерпретации, мониторинга и диагностики состояния СЭС и ее элементов, проектирования, а также планирования, в принятии решений, по которым необходимо присутствие человека. Решение перечисленных выше задач осуществляется в диалоге ЛПР – ЭВМ и не может быть пока полностью автоматизировано по следующим причинам [1, 4, 13, 20]:

- неопределенность полной совокупности параметров при оценке состояния объекта и, как следствие, отсутствие четкого интегрального критерия;
- большое количество факторов, влияющих на состояние объекта, в том числе таких, которые не имеют количественного выражения;
- сложная связь между влияющими на работу факторами, часто неявная и неоднозначная, трудно формализуемая, плохо структурируемая;
- недостаточная информация о параметрах СЭС и влияющих на них факторах;
- построение математической модели СЭС затрудняет высокая связанность процессов, протекающих в различных подсистемах, группах и отдельных элементах. По этой причине модель СЭС должна включать не только модели отдельных групп и элементов, но и модели, описывающие их взаимодей-

ствие между собой. В результате математическая модель оказывается очень громоздкой и трудно реализуемой;

- реализация многих моделей требует большого ресурса вычислительных средств;

- реальный процесс диагностирования состояния СЭС является прерогативой специалистов высокого класса – экспертов, обладающих высокой степенью компетентности в соответствующей узкоспециальной области. ЛПР может исходить в своих рассуждениях не из самих результатов измерений, а из тенденций, качественных оценок или структуры этих измерений.

Эффективное функционирование АСУЭ основано на предъявляемых к ней требованиях [1, 2, 6]. Это наличие оперативного управления в реальном времени, комплексность, многоуровневость и четкость иерархии управления, высокая надежность и программная совместимость с другими АСУ. Задачей управления АСУЭ является необходимость воздействовать как на СЭС, так и взаимодействовать с внешними системами, чтобы эффективно обеспечить основные и вспомогательные производства промышленного предприятия необходимым количеством электроэнергии надлежащего качества. В АСУЭ управление смещается из области стабилизации основных электроэнергетических параметров и автоматизации повторяющихся операций в область решения задач оптимального управления и эксплуатации.

Практика эксплуатации промышленных СЭС показывает, что при утяжеленных, аварийных, послеаварийных или динамически развивающихся переходных режимах работы особенно проявляется нечеткость и неполнота в информации, необходимой для автоматизированного принятия адекватного текущей ситуации управляющего решения. Существующие автоматизированные системы АСУЭ резко снижают свою работоспособность в подобных условиях. Возникает опасность для жизни людей и предприятия.

Действенным средством автоматизации обеспечения жизненного цикла промышленного электроснабжения стали ЭВМ новых поколений. Но наличие даже самых современных ЭВМ не является достаточным условием для успешного функционирования АСУЭ. Для этого требуются более совершенные методы организации и соответствующие им средства управления.

1.6. Анализ причин неопределенности информации в ЭЭС

Проблема автоматизированного принятия решений в СЭС и ЭЭС связана с выбором альтернатив в условиях неопределенности имеющейся информации [1, 10, 13–17, 18, 20]. Неопределенность информации проявляется как неопределенность составляющих – знаний и данных. Причины неопределенности выразим в виде дерева (рис. 1.3).

На первом иерархическом уровне неопределенность классифицируем триадой: неопределенностью природы явлений в предметной области, выражаемой через неизвестные факторы; неопределенностью характера поведения предметной области, результаты действия которой нельзя полностью учесть и предсказать; неоднозначностью собственных целей управления, вследствие то-

го, что перед ЛПР стоит сразу несколько целей, которые одним, даже интегральным критерием выразить невозможно.

На верхнем уровне дерева неопределенность образована тремя главными причинами, характеризующими количество недостающей информации о составляющих элементах ЭЭС. Во-первых, сложностью ЭЭС – неполнотой и недостаточностью знаний. Неопределенность общих знаний обнаруживается в неполноте имеющихся знаний о состоянии как ЭЭС, так и всей предметной области, входящей в энергосистему, недостаточной изученности сложных процессов, отсутствии полного осмысления структуры сложной системы и всех механизмов взаимодействия между подсистемами и отдельными элементами. Неопределенность конкретных знаний характеризуется во-первых: многосвязанностью данных, когда одни функциональные отклонения проявляются в симптомах других изменений; недостатком статистического материала о неопределенных параметрах; случайным характером изменения параметров; зашумленностью данных.

Во-вторых, нечеткостью ситуации, которая вызвана недетерминированным характером нагрузки и показателей качества электроэнергии, наличием качественной и субъективной информации, характеризующей параметры ЭЭС. Источниками неопределенности такого рода могут быть: невозможность точного измерения параметров режима и параметров системы; неточность исполнительских действий; невозможность полного и четкого описания многих элементов ЭЭС и ситуаций; недостаточность учета влияющих факторов, несвоевременность поступления информации. В-третьих, присутствующей в информации нечеткости, которая связана с отсутствием точных границ областей определения и свойственна большинству понятий.

Типы источников возникновения неопределенной информации в информационных видах обеспечения АСУЭ также довольно разнообразны. Первый тип связан с надежностью информации – неопределенность может присутствовать в фактическом знании. Второй тип обусловлен неточностью языка представления правил принятия решений, так как, если правило не выражено на формальном языке, его значение не может быть выражено точно. Третий тип возникает, когда принятие решения основано либо на неполной информации, либо при синтезе правил принятия решений, полученных из различных источников или от разных экспертов.



Рис. 1.3. Классификация причин неопределенности информации, возникающих при решении задач предметной области

Анализ информационной базы ЭЭС и особенностей АСУЭ показывают, что, переходя к разработке систем, осуществляющих автоматизированное управление, приходится развивать формальный механизм преобразования информации, приближающийся по своим возможностям к процессу понимания явлений человеком. Углубляя исследования в этом направлении, становится все более очевидным, что необходимая для разрабатываемых систем информация все больше отходит от объектов, которыми оперирует традиционная математика. Используются не только числовые, но и лингвистические модели рассуждений специалистов. Если традиционные методы предъявляют требования приближения к реальности, согласования со средой, обеспеченности вычислительными ресурсами, конечности, адекватности, однозначности, точности, полноты, замкнутости, непротиворечивости и др., то во многих утяжеленных режимах ЭЭС или близких к ним экстраординарных условиях большинство из этих требований выполнить невозможно.

Общим термином, характеризующим неопределенность информации, было предложено понятие НЕ-фактора [13]. Основной чертой различных по природе факторов информации, является то, что каждый из них отражает компонент модели понимания, проявляясь в виде содержательного и лексического отрицания классических свойств формальных систем: полноты, определенно-

сти, достоверности и т.п. НЕ-факторы представляют собой явления одного порядка и отражают различия между сложной реальной системой и ее моделью.

Второй уровень дерева указывает общие причины неопределенности информации в предметной области ЭЭС: неизвестность, неадекватность, недостоверность. В ситуации неизвестности, например, на первых стадиях формулирования проблемы имеются математическая неопределенность и трудности описания, когда еще не выработаны требования к критериям оценки – многовариантность, а также неизвестны требования заказчика или возможности исполнителя – неопределенность предпочтений, информация о задаче практически отсутствует. При сборе данных на первом этапе, временно прекращенном в связи с нехваткой нужных для сбора информации ресурсов, может оказаться, что собранная информация обладает недостоверностью, поскольку имеются еще не все возможные сведения. В этом случае проявляется неполнота, если получена не вся необходимая информация, то – недостаточность. Когда для некоторых элементов ЭЭС определены не их однозначные описания, а лишь множества, которым эти описания принадлежат, это – недоопределенность. Ряд элементов задачи в какой-то период времени может быть описан лишь приблизительно – неадекватно по грубой аналогии с уже решавшимися типовыми задачами. Но при продолжении исследований эти составляющие недостоверности могут быть устранены. Дальнейшая обработка информации, возможно, приведет либо к однозначному описанию элементов, либо неоднозначному, когда собрана возможная информация, но полного описания нет, и оно не может быть получено этим путем.

Следующий третий уровень дерева раскрывает конкретные причины возможной неполноты описания, оказывающие влияние на субъект управления со стороны внешней среды: физической и математической неопределенности, а также лингвистической неопределенности профессионального языка. Математическую интерпретацию исходной и обрабатываемой информации классифицируем по виду ее существующего представления: детерминированная, вероятностная и нечеткая (размытая).

Детерминированными являются параметры, значения которых либо точно известны, либо вероятностные величины обладают пренебрежительно малой дисперсией. Но в процессе принятия решения детерминированная величина, представляемая точкой на числовой оси, может быть также неопределенной по причинам отсутствия контекста, в котором она существует. Это возможно при отсутствии качественных сторон описания. Результат приводит к нестабильности существования даже при нормальных режимах, например, применяемое в проектировании номинальное значение напряжения в условиях эксплуатации не является стабильной величиной питающего напряжения из-за постоянного корреляционного изменения других параметров режима.

Для более полного представления параметров режима в процессе функционирования ЭЭС вводятся интервалы параметров режима в виде допустимых областей определения, например, показатели качества электроэнергии для нормального и послеаварийного режимов. Однако интервал также является достаточно грубой моделью представления информации.

С целью повышения гибкости интервальной оценки модели могут, например, использоваться функции распределения случайной величины в рассматриваемом интервале. Принятое решение в этом случае должно учитывать фактор риска. Другой путь уточнения модели проходит через представление параметров режима в виде множеств. При этом все элементы множества обладают общими свойствами, но предлагаемые варианты принимаемых решений являются равноценными, поэтому для выбора оптимального решения необходима дополнительная информация.

Более точный учет взглядов и суждений о моделях ЭЭС дают нечеткие множества (НМ) – элементы множества обладают общим свойством, но в различной степени. Функция принадлежности НМ является способом формализации контекста существования множества показателей; изображает гипотезу субъективного представления ЛПР об особенностях процесса управления, о характере ограничений и целей решения; описывает способ оценки альтернатив; расширяет входную информацию, переходя к ее представлению от числовой оси к плоскости. Подобная информация в большинстве экстраординарных случаев остается единственной, характеризующей процессы в ЭЭС.

Физическая неопределенность связана как со стохастической неопределенностью – случайностью, так и с погрешностью измерений физическими приборами – неточностью. Однако физическая неопределенность предполагает знание законов распределения вероятностей в обоих случаях. Здесь также присутствует и нечеткость, связанная с качественными характеристиками самой информации.

Лингвистическая неопределенность профессионального языка порождается множественностью и неопределенностью значений слов – полисемией и неоднозначностью смысла фраз. В полисемии обозначаемые одним и тем же словом объекты различны, например, в ПУЭ [19] в ст. 1.2.8 говорится об электроприемниках первой категории как «... с особо сложным непрерывным технологическим процессом, требующим *длительного* времени на восстановление рабочего режима ...», а в ст. 1.3.2 говорится, что «допустимые *длительные* токи ... выбираются в зависимости от ...». Если аналогичные описания не сходны, то это – омонимия, но если они сходны, то это – нечеткость. Рассматриваемый термин *длительный* для обозначения продолжительности процесса весьма приблизителен, поскольку в первом случае время восстановления рабочего режима зависит от продолжительности ремонта поврежденного оборудования, а во втором случае длительность связана с теплофизическими и режимными характеристиками проводников электрооборудования и электроустановок.

Источником многозначности смысла фраз является синтаксическая, семантическая и прагматическая неоднозначность. При синтаксической неоднозначности уточнение синтаксиса предложения позволяет понять смысл фразы, например, «*Уменьшить потери*, слегка изменяя напряжение» и «*Уменьшить потери слегка*, изменяя напряжение». При семантической неопределенности отдельные слова из предложения понятны, но неясен смысл всей фразы, например, постороннему мало понятен смысл происходящего процесса при передаче словесных управляющих воздействий: «еще, чуть-чуть, еще немного,

достаточно». Прагматическая неопределенность связана с неоднозначностью понимания описания текста решаемой задачи, например, «оптимизировать режим электропотребления».

Традиционные способы решения задач автоматизированного управления предъявляют к используемой информации требования неопременного приближения к реальности, четкости, однозначности, точности, полноты, непротиворечивости и др., но в экстраординарных и экстремальных условиях многие из этих требований выполнить невозможно. Если информация о СЭС имеет неопределенный характер, то в полной мере нельзя воспользоваться моделями этой системы, опираясь на теорию подобия. Однако управление высококвалифицированными специалистами осуществляется даже при такой информации, следовательно, необходимо разрабатывать и использовать модели рассуждений экспертов-ЛПР и стремиться к уменьшению неопределенности информации путем применения новых интеллектуальных методов.

Формализация нечетких данных и качественных знаний, связанная с использованием методов нечеткой математики, позволяет переходить от нечеткой символьной формы представления знаний к числовым аналогам, пригодным для интерпретации в интеллектуальных системах поддержки принятия решений, основанных на эвристическом подходе в решении сложных задач с нечеткой и неопределенной входной информацией.

1.7. Компоненты сложной системы

По мере постижения проблемы функционирования таких сложных систем, как ЭЭС и СЭС, становится понятной неотъемлемость для жизненного цикла решения задач их управления. Эволюция методов и средств управления этих сложных системам была связана с трансформацией развития составляющих общей проблемы функционирования энергосистемы: надежностью, оптимизацией, устойчивостью, электробезопасностью и др. Объединяющим фактором был и остается информационный аспект проблемы, поскольку все подсистемы, входящие в сложную систему, связаны единством информационного обмена. Основным решением информационного аспекта задачи управления для ЛПР является преодоление «информационного кризиса». В одних присутствующих подсистемах автоматизированного управления, например, АСУП, это избыток циркулирующей информации, в других (АСУТП, АСУЭ) – ее недостаток и неопределенность. Особенно явно это проявляется в экстраординарных режимах. В нормальных же режимах работы для ЛПР необходима дополнительная информация о работе изношенного оборудования, объем которого в ЭЭС и СЭС постоянно увеличивается, опережая темпы совершенствования АСУЭ.

Проблема информации, несмотря на нематериальность последней существует объективно, и исследователи сначала обратили на нее внимание, а затем начали познавать ее в зависимости от своих возможностей, способностей и потребностей. Проблема возникла в ходе развития используемых в энергосистеме разделов науки и техники – это наша апперцепция, осознанное восприятие не-

знания. Поскольку проблемы не рождаются на пустом месте, но вырастают из обобщения полученных ранее результатов, то существующего знания уже достаточно, чтобы указать на некоторое неизвестное из системы знания. Установление наличия неизвестного звена в проблемной области, его выделение и осмысление являются целью этого раздела.

Объектом управления рассматриваемых сложных систем (в дальнейшем систем) является электрооборудование со всеми связями. Во многих случаях имеется естественное стремление к возникновению неисправностей, старению, неустойчивой работе при внешних и внутренних возмущениях. Объект, как правило, увеличивает информационную энтропию системы.

Субъект управления представляет собой ЛПР в АСУ и персонал, который непосредственно связан с функционированием объекта. Субъект управления стремится к сохранению устойчивой, безопасной, оптимальной работы объекта без аварий и уменьшает энтропию системы.

В системе изначально присутствует и третий трансцендентальный компонент, находящийся за пределами рассмотренных определений в первом разделе. О наличии третьего системного компонента неявно говорится во многих исследованиях, например в [2, 3, 15]. Он существует, но его пока не выделяют. Он проявляется в эффективности преобразования и использования информации, циркулирующей в системе.

Рассмотрев источники неопределенности информации в предыдущем разделе, можно сделать вывод о том, что главной причиной феномена возникновения неопределенности информации является субъективная системная ограниченность. Количественная и качественная мера определенности информации зависит в конечном итоге от рассудка, мыслительных и умственных способностей субъекта управления системы, его возможности воспринимать и обрабатывать информацию. Такая характеристика третьего системного компонента, функционирующего до сих пор «по умолчанию», позволяет сделать вывод о том, что в энциклопедическом смысле рассматриваемая часть системы является ее *интеллектом*.

К третьему системному компоненту следует относить опыт, интуицию и навыки работы персонала, модели и методы, базы данных и знаний, алгоритмы и программы систем автоматики и АСУ. Интеллектуальная системная составляющая выступает в роли «органов чувств и ума» субъекта управления. Это способность субъекта преобразовывать энергетический поток в отклик, сопровождающий взаимодействие элементов системы, в информацию. Основная задача состоит не только в этом преобразовании, но и выделении полезной информации, доведении ее до понимания субъекта. Интеллектуальная составляющая в организационном плане является арбитром, выполняющим рассуждения и исследования, средством соизмерения деятельности элементов и частей системы. Новый взгляд на интересующую нас сущность с позиций системного анализа дает возможность представить сложную систему в виде трех компонентов:

$$\text{Система} = \text{Объект} + \text{Субъект} + \text{Интеллект}.$$

Очевидно, что энергосистема, ЭЭС и СЭС как сложные технико-энергетические системы – это не только совокупность взаимосвязанных составляющих объекта и субъекта, но и обязательное присутствие наличия интеллекта, который помогает субъекту уменьшать энтропию системы и снижать неопределенность информации. Целью всей сложной системы (рис. 1.4) является функционирование (выполнение основных функций по обеспечению потребителей электроэнергией) и забота о собственном «здоровье» (поддержание необходимой степени организованности связей между ее компонентами, частями и элементами).



Рис. 1.4. Взаимодействие компонентов сложной системы

Интеллект в системе подразделяется на естественный (свойственный ЛПР и обслуживающему персоналу) и искусственный (выполнение с помощью ЭВМ функций моделирования человеческой деятельности).

Для ответа на вопрос о причинах «информационного кризиса», возникающего в процессе управления сложной системой, необходимо хотя бы кратко остановиться на феномене информации.

Информация определяется как отражение сущности, явления, процесса через результаты взаимодействия элементов системы [17]. Информация в электрических цепях ЭЭС обнаруживается по изменению поля (полей) в результате действия, преобразования одной формы поля в другую. Наведение ЭДС в индуктивности является результатом изменения тока в этой или другой магнитосвязанной цепи. Мощность есть результат скорости изменения энергии. Информация проявляется также через толкование происшедших изменений, последствий в элементах. Поэтому результат отражения откликов взаимодействия элементов с помощью интеллекта и есть информация согласно ее определению.

Отклик на результат взаимодействия элементов системы находится в пространстве, в котором совершается работа поля вокруг явления, ситуации, первопричины.

Напряжение, ток, мощность предстают в виде информации о функционировании электрооборудования ЭЭС и становятся параметрами режима. Сопротивления, проводимости, схемные факторы – информация о физических свойствах электрооборудования – параметры системы. Параметры режима и системы являются истолкованием, интерпретацией состояния сложной системы. Интерпретированная информация может передаваться, храниться, накапливаться в различных видах. Ее нужно только выделить, ограничить и правильно объяснить.

Поскольку функционирование объекта вызывает изменение электрических, магнитных и электромагнитных полей, а изменение полей преобразуется в информацию о функционировании объекта, то наиболее информативно объект проявляет себя в работе, в динамике, во взаимодействии с другими объектами. Все действия нарушают состояние баланса, равновесия и приводят к противодействию. В механике это описывает третий закон Ньютона – действие равно противодействию. В электротехнике это выражено в уравнениях Максвелла самоподдерживающееся распространение электромагнитных волн в вакууме. В электромеханике – это действие обратного момента в двигателе, получение в генераторе ЭДС и т.п. По противодействию мы судим о действии. Поэтому, например, авария объекта есть противодействие неправильным действиям в системе.

Одно из главных свойств информации заключено в возможности ее интерпретации, истолкования, объяснения, описания. Интерпретация предстает как вид отображения информации. Информация многогранна, многопланова, многообразна в видах ее отражения. Она присутствует постоянно, не исчезает, а переходит из одного вида в другой. При правильной, неискаженной интерпретации сути действия соответствующая информация об объекте не должна изменяться. Таким образом, информация о сложной системе предстает в следующем виде:

$$\text{Информация} = \text{Полезная информация} + \text{Информационный шум} + \text{Метаинформация.}$$

Полезная информация – информация, которая необходима (с точки зрения субъекта) для функционирования системы. Информация, которую может воспринимать субъект – функция интеллекта, совершаемая с информацией в границах восприятия субъекта.

$$\text{Полезная информация} = \text{Данные} + \text{Знания.}$$

Данные – сведения, представленные в знаковой системе на определенном носителе для обеспечения возможностей их хранения, передачи, приема и об-

работки. Абстрактная субстанция, несущая некоторую информацию. Данные, как правило, безотносительны к содержанию информации.

Знания – проверенные общественной практикой полезные сведения, которые могут многократно использоваться людьми для решения тех или иных задач.

Информационный шум – ненужная с точки зрения субъекта информация. Эта информация в текущий момент времени не может быть использована для функционирования системы. Например, функционирование силовых трансформаторов сопровождается гудением. В недавнем прошлом это воспринималось субъектом просто как вредный и бесполезный гул. В настоящее время разработаны акустические методы и средства для диагностики функционирования силовых трансформаторов. Информационный шум благодаря развитию интеллекта превратился в полезную для субъекта информацию.

Информационный шум возрастает гораздо быстрее объема полезной информации, так как интеллект развивается гораздо медленнее этого роста. Создается ложное представление о *лавинообразном* росте *всей* информации, поскольку информационный шум стал преобладать над полезной информацией. Сейчас нужно формулировать проблему иначе: не лавинообразное увеличение информации, но *катастрофическое* замедление развития интеллекта. В этом нужно искать стратегическую цель системных научных исследований.

Метаинформация – вспомогательная информация, позволяющая отделять полезную информацию от информационного шума.

НЕ-факторы [13] и энтропия являются функцией двух переменных – интеллекта и восприятия субъекта. Энтропия в этом случае может трактоваться как мера не отсутствия полезной информации, а как мера ограниченности, недостатка интеллекта, который не может сам воспринять или восполнить недостающую часть имеющейся информации. В действии, поведении объекта нет случайного и неопределенного. Здесь все закономерно, а случайность и неопределенность являются недостаточностью знаний о закономерном действии в системе.

Объект управления не должен подвергаться необоснованной критике за аварии и отказы, поскольку он не наделен интеллектом. Приданные искусственные системы автоматики контролируют ограниченное количество его степеней деятельности (свободы). Субъект управляет объектом, используя свойства объекта. Если специфика объекта будет применена с толком, то цель будет достигнута, субъект окажется на высоте и объект будет удобен в функционировании. Неправильное управление субъектом приводит к неоднократным сбоям, быстрому изнашиванию оборудования, опасности разрушения и авариям.

Объект обладает естественными степенями свободы деятельности, которые не являются отрицательными и непригодными параметрами, но их можно и нужно употребить с пользой. Не имея приданного интеллекта, объект может выйти за допустимые границы функционирования. Если субъект правильно использует степени свободы деятельности объекта, то и объект правильно функционирует и субъект находится вне опасности.

В исходное состояние систему могут возвращать внешние и внутренние воздействия. Компоненты системы обладают ограниченным набором возможностей, чтобы противостоять внутренним и внешним воздействиям и характеризуются параметрами режима и системы, а также параметрами состояния. Субъекту, отягощенному разными функциями, трудно достичь успеха в главном из-за проявления ограниченности возможностей «человеческого фактора». Выход из этого положения состоит в совершенствовании интеллектуальной составляющей, применении систем и технологий искусственного интеллекта, освобождающих естественный интеллект от рутинных, дополнительных и сложных функций. В этих условиях создается надстройка управления – система поддержки принятия решений: модель – метод – средство (рис. 1.5). Модель служит эквивалентом объекта управления. Метод выражает в алгоритмической форме свойства объекта, передаваемые моделью. Средства реализуют методы принятия решений с помощью интеллектуальных компьютерных программ.

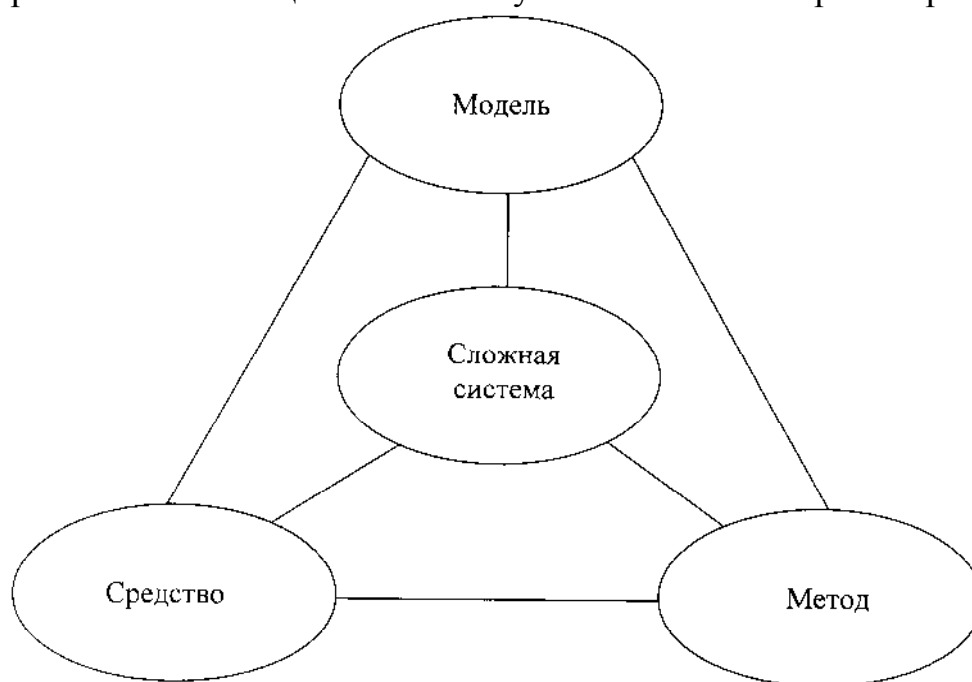


Рис. 1.5. Структура поддержки принятия решений при управлении сложной системой

Чем более развит системный интеллект, тем более гибок процесс возможного функционирования при имеющихся ограничениях. Свобода действий в естественной замкнутой системе характеризуется увеличением ее энтропии, однако, в искусственной системе это, наоборот, поддержание низкой энтропии системы, т.е. более высокой степени организованности системы, переход от наблюдения внешних проявлений к исследованию и устранению внутренних первопричин неустойчивых и аварийных состояний.

1.8. Понятие модели и моделирования. Классификация моделей

Научные знания можно разделить на две категории: фундаментальные и прикладные [3, 7, 17].

Фундаментальные знания описывают наиболее общие законы природы и техники.

Прикладные знания представляют собой разновидность фундаментальных знаний и находят применение при организации производства товаров и в сфере услуг. Какая-то часть этих товаров и услуг используется в процессе исследований, что, в свою очередь, повышает уровень фундаментальных и прикладных знаний.

Для согласования результатов «смежных» исследовательских программ и выработки единого убедительного для практики заключения – хорошим средством оказывается модель.

Модель – материальный или мысленно представляемый объект, который в процессе изучения замещает объект-оригинал, сохраняя некоторые важные для данного исследования типичные его черты.

Моделирование можно рассматривать как замещение исследуемого объекта (оригинала) его условным образом, описанием или другим объектом, именуемым моделью и обеспечивающим адекватное с оригиналом поведение в рамках заданных допущений. Моделирование обычно выполняется с целью познания свойств оригинала путем исследования его модели, а не самого объекта. Моделирование оправдано в том случае, когда оно проще создания самого оригинала или когда последний по каким-то причинам лучше вообще не создавать. Реальная польза от моделирования может быть получена при условии, что модель адекватна оригиналу в том смысле, что должна с достаточной точностью отображать интересующие исследователя характеристики оригинала.

В большинстве случаев моделирование вовсе не заменяет реальный объект и не отменяет необходимости в его разработке и натурном испытании. Оно просто значительно уменьшает объем работ по проектированию и исследованию объектов. В тех же случаях, когда это не так, стоимость моделирования может оказаться вполне сравнимой со стоимостью разработок и натурных испытаний изделий.

Дадим классификацию моделей, отражающую в первую очередь методологические вопросы процедуры построения математических моделей и нахождения их решения с помощью ЭВМ.

Если исходить из целевого направления информационных потоков, циркулирующих между объектами и окружающим миром, модели можно разделить на модели для исследования и модели для управления.

Модели для исследования являются формой организации и представления знаний, средством соединения новых знаний с имеющимися. При расхождении модели с реальностью это несоответствие ликвидируется путем изменения модели.

Модели для управления являются средством организации практических действий, способом представления эталонных действий или их результата, т.е. являются рабочим представлением целей. Модели для управления используются для того, чтобы при обнаружении расхождения между моделью и реальным

процессом направить усилия на изменение реальности так, чтобы приблизить ее к модели. Они носят нормативный характер, играют роль стандарта, под который подгоняются как сама деятельность, так и ее результат. Примерами моделей управления служат планы и программы, уставы организаций, законы, алгоритмы, рабочие чертежи и шаблоны, параметры отбора, технологические допуски, технические требования и т.д.

Основное различие между исследовательскими моделями и моделями для управления состоит в том, что модели для исследований отражают существующее, а модели для управления – не существующее, но желаемое и возможно осуществимое.

По форме представления модели делят на физические, символические и смешанные.

Физические модели подразделяются на модели подобия и аналоговые.

Модели подобия характеризуются некоторыми масштабными изменениями, выбираемыми в соответствии с критериями подобия (например, глобус – модель земного шара). Природа процесса и его физическая сущность одинаковы как для модели, так и для исследуемого оригинала.

Аналоговые модели основаны на известных аналогиях между протеканием процессов в механических, тепловых, электрических, пневматических, гидравлических и других динамических системах и предназначены для исследования статических и динамических свойств объекта.

Символические модели характеризуются тем, что параметры реального объекта и отношения между ними представлены символами:

- семантическими (словами);
- математическими;
- логическими.

Класс символических моделей весьма широк. Наряду со словесными описаниями функционирования объектов – сценариями, сюда также относятся схематические модели: чертежи, графики и блок-схемы, логические блок-схемы (например, алгоритмы программ) и таблицы решений, таблицы и номограммы, а также математические описания – математические модели.

Математическая модель представляет собой набор формальных соотношений, которые отображают поведение исследуемой системы и состоящее из совокупности связанных между собой математическими зависимостями (формулами, уравнениями, неравенствами, логическими условиями) величин – факторов. По своей роли эти факторы целесообразно подразделить на параметры и характеристики.

Модели функционирования включают широкий спектр символических моделей, например:

- модель жизненного цикла системы, описывающая процессы существования системы от зарождения до прекращения функционирования;
- модели операций, выполняемых объектом, представляют описание взаимосвязанной совокупности процессов функционирования отдельных элементов объекта. Так, в состав моделей операций могут входить модели надежности,

характеризующие выход элементов системы из строя под влиянием эксплуатационных факторов;

- информационные модели, отображающие во взаимосвязи источников и потребителей информации, виды информации, характер ее преобразования, а также их временные и количественные характеристики;

- процедурные модели, описывающие порядок взаимодействия элементов исследуемого объекта при выполнении различных операций, например обработки материалов, деятельности персонала, использования информации, в том числе и реализации процедур принятия управленческих решений;

- временные модели, описывающие процедуру функционирования объекта во времени и распределение ресурса «время» по отдельным компонентам объекта.

Параметрами объекта называются *факторы*, характеризующие свойства объекта или составляющих его элементов (рис. 1.6). В процессе исследования объекта ряд параметров может изменяться, поэтому они называются переменными, которые в свою очередь подразделяются на переменные состояния и переменные управления.

Переменные состояния объекта являются функцией переменных управления и воздействий внешней среды.

Характеристиками (выходными характеристиками) называются интересующие исследователя непосредственные конечные результаты функционирования объекта (естественно, что выходные характеристики являются переменными состояниями).

Характеристики внешней среды описывают свойства внешней среды, которые сказываются на процессе и результата функционирования объекта. Значения ряда факторов, определяющие начальное состояние объекта или внешней среды, называются начальными условиями.

При описании математической модели оперируют следующими понятиями:

- критерий оптимальности;
- целевая функция;
- система ограничений;
- уравнение связи;
- решение модели.

Критерием оптимальности называется некоторый показатель, служащий формализацией конкретной цели управления и выражаемый при помощи целевой функции через факторы модели. Критерий оптимальности определяет смысловое содержание целевой функции. В ряде случаев в качестве критерия оптимальности может выступать одна из выходных характеристик объекта.

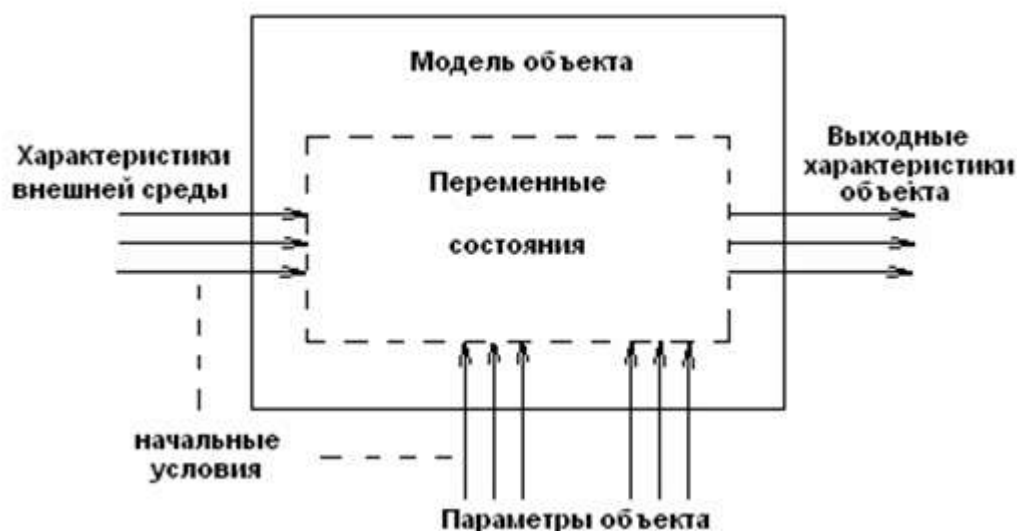


Рис.1.6. Классификация факторов по их роли в модели.

Целевая функция математически связывает между собой факторы модели, и ее значение определяется значениями этих величин. Содержательный смысл целевой функции придает только критерии оптимальности.

Система ограничения определяет пределы, сужающие область осуществимых, приемлемых или допустимых решений и фиксирующие внешние и внутренние свойства объекта. Ограничения определяют область протекания процесса, пределы изменения параметров и характеристик объекта.

Уравнения связи являются математической формализацией системы ограничений.

Критерии оптимальности и система ограничений определяют концепцию построения будущей математической модели, т.е. концептуальную модель, а их формализация, т.е. целевая функция и уравнения связи, представляет собой математическую модель.

Решением математической модели называется такой набор (совокупность) значений переменных, который удовлетворяет ее уравнениям связи.

Модели, имеющие много решений, называются вариантными в отличие от безвариантных, имеющих одно решение. Среди допустимых решений вариантной модели, как правило, находится одно решение, при котором целевая функция, в зависимости от смысла модели, имеет наибольшее или наименьшее значение. Такое решение, как и соответствующее значение целевой функции, называется *оптимальным*.

В зависимости от степени формализации связей между факторами различают аналитические и алгоритмические модели.

Аналитической называется модель в виде уравнений или неравенств, не имеющих разветвлений вычислительного процесса при определении значений любых переменных состояния модели, целевой функции и уравнений связи.

Если в математических моделях единственная целевая функция и ограничения заданы аналитически, то подобные модели относятся к классу моделей *математического программирования*.

Характер функциональных зависимостей может быть линейным и нелинейным. Соответственно этому математические модели делятся на линейные и нелинейные.

В сложной системе зачастую гораздо легче построить ее модель в виде алгоритма, показывающего отношения между элементами системы в процессе ее функционирования, задаваемые обычно в виде логических условий – разветвлений хода процесса.

К *алгоритмическим моделям* относятся и имитационные модели – моделирующие алгоритмы, имитирующие поведение элементов изучаемого объекта и взаимодействие между ними в процессе функционирования.

При *имитационном моделировании* процесс функционирования подсистем, выраженный в виде правил и уравнений, связывающих переменные, имитируется на компьютере. Для имитации используются специальные среды имитационного моделирования, позволяющие строить модели, имитирующие работу моделируемой системы, с любой степенью достоверности без проведения подробных аналитических преобразований.

В зависимости от того, содержит ли математическая модель случайные факторы, она может быть отнесена к классу стохастических или детерминированных.

В *детерминированных моделях* ни целевая функция, ни уравнения связи не содержат случайных факторов. Следовательно, для данного множества входных значений модели на выходе может быть получен только один единственный результат. Главная особенность детерминированной модели заключается в том, что любой прогноз получается в виде числа, а не в виде распределения вероятностей. Это в ряде случаев приемлемо, однако когда приходится иметь дело с величинами, значение которых предсказать трудно (количество осадков, ветровая нагрузка), такой подход оказывается совершенно неудовлетворительным.

Стохастические математические модели имеют факторы с вероятностной природой и характеризуются какими-либо законами распределения. Значения выходных характеристик в таких моделях могут быть предсказаны только в вероятностном смысле. Это даёт возможность оценивать не только среднее значение прогнозируемого параметра, но и его дисперсию.

Следующим признаком, по которому можно различать математические модели, является связь с фактором времени.

Статическая модель – это математическая функция, в которую не включена переменная времени. Все особенности поведения системы, имеющие выраженную зависимость от времени, при этом игнорируют. А поскольку все в мире быстро ли, медленно ли, но меняется, то любая статическая модель условна.

Статическими моделями пользуются, когда в рамках поставленной задачи инерционностью и «памятью» реальной системы можно пренебречь. Это возможно при выполнении ряда условий, в число которых входят следующие:

- система устойчива, т.е. переходные процессы после скачкообразного изменения входов затухают;
- входы меняются медленно;
- выходы изменяются редко.

Динамические модели позволяют учесть наличие «памяти», инерционности системы. Математическим аппаратом описания динамических систем являются дифференциальные, разностные уравнения, конечные автоматы, случайные процессы. Динамические модели, имеющие практическую ценность, обычно строятся на основе дифференциальных уравнений, не поддающихся прямому интегрированию, и решение их нельзя получить в виде простых аналитических выражений. В этом случае прибегают к численным методам решений на компьютере с помощью специального программного обеспечения.

Система может быть дискретной или непрерывной по входам, выходам и по времени. Под дискретным понимается конечное или счетное множество – один, два, три и т.д. Под непрерывным множеством понимается отрезок, луч или прямая линия, т.е. связное числовое множество, количество элементов которого стремится к бесконечности. Как правило, дискретность входа влечет за собой дискретность выхода объекта. Кроме того, для статических систем исчезает разница между непрерывным и дискретным временем.

Смешанные модели могут содержать как физические, так и символические элементы.

Эмпирические модели описывают связи между параметрами элементов одного уровня. Разработчик эмпирической модели всегда остается в пределах одного единственного уровня организационной иерархии, где он и строит уравнения, связывающие между собой параметры, свойственные подсистеме только данного уровня.

Функциональная модель объясняет связи между элементами как одного уровня иерархии, так и между различными уровнями. Разработчик функциональной модели стремится описать поведение системы с фундаментальных позиций, затрагивающих основу работы объекта, учитывающих наиболее общие закономерности его работы.

Всегда можно построить такую эмпирическую модель, которая была бы согласована с массивом опытных данных лучше, чем функциональная, т.к. эмпирическая модель практически свободна от ограничений, в то время как возможности функциональной модели ограничиваются положенными в ее основу допущениями, идеями и гипотезами.

Получение данных и обработка данных для моделирования. Исследование реального объекта и его математической модели связано с использованием исходной информации, получаемой в процессе непосредственного измерения на объекте. Получение данных осуществляют путем:

- 1) всеобщего контроля;
- 2) выборочного исследования;
- 3) планирования эксперимента.

При всеобщем контроле осуществляют измерения со всех объектов, по всем параметрам на всех временных интервалах. Это предполагает большие материальные и временные затраты на осуществление исследования.

Выборочное исследование – это метод исследования, при котором параметры изучаемого явления, происходящего на объекте, устанавливаются по определенной части этого объекта на основе положений случайного отбора-выборки. Результаты исследования части объекта распространяются на весь объект – генеральную совокупность. В ряде исследований этот метод является единственно возможным, например, при контроле качества продукции, проводимом путем уничтожения или разложения на составляющие изучаемого продукта.

Особенность выборочного исследования состоит в том, что выбор единиц для обследования происходит по принципу равных возможностей попадания в выборку каждой единицы исследуемого параметра- считается, что клейковина в массе зерна постоянна для всего элеватора- генеральной совокупности (для одной партии или потока). При распространении результатов выборки на всю генеральную совокупность возникают ошибки, зависящие от разных факторов: степени вариации изучаемого явления, численности выборки, методов отбора единиц для исследования, принятого уровня достоверности результатов. Для снижения ошибки применяют случайные (рандомизированные) выборки.

Рандомизация – это случайный выбор объекта исследования, его уровня или варианта.

Исходные экспериментальные данные с объекта, например для двух величин x и y , формируются в виде таблиц измерений зависимой (выходной) величины y от независимой (входной) величины x .

Исходные данные об объекте или его модели могут быть представлены в виде:

- 1) отдельных чисел;
- 2) векторов и матриц чисел;
- 3) временного (динамического) ряда.

Экспертные оценки применяются, когда нет надлежащей теоретической или экспериментальной информации об объекте исследования. Исходя из полученной в результате анализа модели объекта исходной информации, определяются направления, специальности, по которым необходимо привлечь экспертов. В оценке эксперта будут интегрированы его знания, интуиция и опыт, относящиеся к конкретному явлению.

Например, один из методов экспертной оценки – метод Дельфи, состоит в последовательном анкетировании мнений экспертов различных направлений деятельности по интересующим вопросам, основанных на логическом анализе, интуиции и опыте. Метод предполагает использование серии анкет, в каждой из которых содержится информация и мнения, полученные из предыдущих анкет. Степень достоверности экспертизы устанавливается по погрешности, с которой оценка эксперта в итоге подтверждается последующими событиями.

Свертывание векторов (скаляризация). В случаях, когда выходная информация представлена в виде вектора, для упрощения анализа применяют его свертывание.

Линейная свертка применяется в случае необходимости иметь один выходной параметр или в случае разных по своей физической природе частных параметров u , с разными шкалами и размерностями.

Планирование эксперимента – это метод исследования, при котором параметры изучаемого явления устанавливаются с помощью специальных планов.

1.9. Задачи оптимизации и основные понятия

В практической деятельности управления в сложной системе специалисты часто сталкиваются с ситуацией, когда им из некоторой совокупности возможных вариантов своего поведения или принятия решения необходимо выбрать один вариант [1, 3, 7, 17]. Наилучший вариант поведения, принятие наилучшего решения можно выбирать по-разному. Если такой выбор предусматривает проведение количественного анализа ситуации путем сравнения различных вариантов с помощью какой-либо количественной оценки этих вариантов, то говорят о необходимости решения задачи оптимизации (лат. *optimus* – наилучший).

При этом задача оптимизации имеет смысл, если есть несколько возможных вариантов ее решения. Эти варианты обычно называют альтернативами.

По содержанию задачи оптимизации весьма разнообразны. Они могут быть связаны с проектированием технических устройств и технологических процессов, с распределением ограниченных ресурсов и планированием работы предприятий, наконец, с решением проблем, возникающих в повседневной жизни человека. Всевозможные устройства, процессы и ситуации, применительно к которым предстоит решать задачу оптимизации, объединяются общим названием объекта оптимизации.

Основные понятия. Для решения задачи оптимизации нужно, прежде всего, найти ответы на следующие вопросы:

- 1) что значит «лучше»?
- 2) что конкретно нужно улучшить?
- 3) за счет чего можно добиться улучшения, что можно изменить?
- 4) в каких пределах можно производить изменения?

Отвечая на первый вопрос, необходимо сформулировать критерий оптимальности, т.е. определить те признаки и предпочтения, по которым следует провести сравнительную оценку альтернатив и выбрать среди них наилучшую с точки зрения поставленной цели оптимизации. Именно с этой точки зрения можно ответить на второй вопрос: что конкретно нужно улучшить? Это может быть повышение производительности оборудования или срока службы технического устройства, снижение массы конструкции аппарата или затрат на его производство и т.п.

Для ответа на два последних вопроса необходимо располагать математической моделью объекта оптимизации. Эта модель описывает объект при помощи соотношений между величинами, характеризующими его свойства. Обычно хотя бы часть этих величин можно изменять в некоторых пределах, что и порождает множество альтернатив, среди которых и предстоит выбрать наилучшую. Изменяемые при оптимизации величины, входящие в математическую модель объекта оптимизации, называют параметрами оптимизации, а соотношения, устанавливающие пределы возможного изменения этих параметров, – ограничениями. Эти ограничения могут быть заданы в форме равенств или неравенств. Их называют соответственно ограничениями типа равенства или ограничениями типа неравенства.

Если множество параметров оптимизации является подмножеством конечномерного линейного пространства, то говорят о конечномерной задаче оптимизации в отличие от бесконечномерных задач, которые рассматривают в вариационном исчислении и оптимальном управлении. При этом критерием оптимальности может быть требование достижения наибольшего или наименьшего значения одной или несколькими действительными (скалярными) функциями параметров оптимизации, выражающими количественно меру достижения цели оптимизации рассматриваемого объекта. Каждую из таких функций принято называть целевой.

Если целевая функция единственная, то задачу конечномерной оптимизации называют задачей математического программирования, а в противном случае – задачей многокритериальной (векторной) оптимизации.

Если целевая функция и ограничения являются линейными относительно параметров оптимизации, то говорят о задаче линейного программирования. Одну из первых таких задач сформулировал и решил Л.В. Канторович – отечественный математик и экономист, лауреат Нобелевской премии 1975 г. Задача Канторовича была связана с выбором оптимальной производственной программы, что и объясняет появление в названии этого класса задач слова «программирование».

При нелинейной зависимости целевой функции или ограничений от параметров оптимизации говорят о задаче нелинейного программирования.

Теория оптимизации находит эффективное применение во всех направлениях инженерной деятельности, и в первую очередь в случаях четырех ее областях:

- 1) проектирование систем и их составных частей;
- 2) планирование и анализ функционирования существующих систем;
- 3) инженерный анализ и обработка информации;
- 4) управление динамическими системами.

Теория оптимизации представляет собой совокупность фундаментальных математических результатов и численных методов, ориентированных на нахождение и идентификацию наилучших вариантов из множества альтернатив и позволяющих избежать полного перебора и оценивания возможных вариантов.

Этапы постановки задачи. Для того, чтобы использовать математические результаты и численные методы теории оптимизации для решения конкретных инженерных задач, необходимо:

1) установить границы подлежащей оптимизации инженерной системы. Система представляется как некоторая изолированная часть реального мира. Границы системы задаются пределами, отделяющими систему от высшей среды, и служат для выделения системы из ее окружения. При проведении анализа обычно предполагается, что взаимосвязи между системой и внешней средой зафиксированы на некотором выбранном уровне представления;

2) определить количественный критерий, на основе которого можно произвести анализ вариантов, с целью выявления «наилучшего». Под критерием оптимальности подразумевается мера, которая дает возможность выбрать наилучший вариант из имеющейся совокупности. В одних приложениях выбираются критерии экономического характера; в других - критерии, которые основываются на некоторых технологических факторах и т.п. Независимо от того, какой критерий выбирается при оптимизации, «наилучшему» варианту всегда соответствует минимальное или максимальное значение характеристического критерия;

3) осуществить выбор внутрисистемных переменных, которые используются для определения характеристик идентификации вариантов. Это переменные, которые должны адекватно описывать допустимые проекты или условия функционирования системы. При выборе независимых переменных целесообразно руководствоваться правилом, согласно которому следует рассматривать только те переменные, которые оказывают существенное влияние на характеристический критерий, выбранный для анализа сложной системы;

4) построить модель системы. Модель описывает взаимосвязи между переменными задачи и отражает влияние независимых переменных на степень достижения цели, определяемой характеристическим критерием. Структура модели может включать основные уравнения материальных и энергетических балансов, соотношения, связанные с проектными решениями, уравнения, описывающие физические процессы, протекающие в системе. Эти уравнения обычно дополняются неравенствами, которые определяют область допустимых значений независимых переменных, позволяют определить требования, накладываемые на верхние или нижние границы изменения характеристик функционирования системы, и установить лимиты имеющихся ресурсов.

Таким образом, задача в виде, пригодном для применения оптимизационных методов, объединяет:

- 1) характеристическую меру;
- 2) множество независимых переменных;
- 3) модель, отражающую взаимосвязь переменных.

Структура оптимизационных задач. Оптимальные задачи можно классифицировать как задачи минимизации вещественнозначной функции $f(x)$ N -мерного векторного аргумента $x = (x_1, x_2, \dots, x_N)$, компоненты которого удовле-

творяют системе уравнений $h_k(x) = 0$; набору неравенств $g_j(x) \geq 0$, а также ограничены сверху и снизу, т.е. $x_i(u) \geq x_i \geq x_i(l)$.

Функция $f(x)$ называется целевой функцией, уравнения $h_u(x) = 0$ – ограничениями типа равенств, неравенства $g_j(x) \geq 0$ – ограничениями типа неравенств. При этом предполагается, что все функции являются вещественнозначными, а число ограничений – конечно.

Классификация оптимизационных задач. Задачи оптимизации можно классифицировать в соответствии с видами функций $f(x)$, h_k , g_i и размерностью вектора X .

Задачи без ограничений, в которых X представляет собой одномерный вектор, называются задачами с одной переменной и составляют простейший, но весьма важный подкласс оптимизационных задач.

Задачи условной оптимизации, в которых функция h_x и g_j являются линейными, называются задачами с линейными ограничениями. В таких задачах функция $f(x)$ может быть линейной и нелинейной.

Задачи, в которых целевая функция $f(x)$ – линейная функция вектора непрерывных переменных, называются задачами линейного программирования.

Задачи, в которых целевая функция $f(x)$ – линейная функция целочисленных переменных, называются задачами целочисленного линейного программирования.

Задачи с нелинейной целевой функцией $f(x)$ и линейными ограничениями h_u , g_j называются задачами нелинейного программирования с линейными ограничениями. Оптимизация задачи такого рода можно классифицировать на основе структурных особенностей нелинейных целевых функций.

Если $f(x)$ – квадратичная функция, это задача квадратичного программирования.

Если $f(x)$ – есть отношение линейных функций – это задача дробно-линейного программирования. Деление оптимальных задач на эти классы важно тем, что специфические особенности тех или иных задач играют важную роль при разработке методов их решения.

Список литературы к 1 главе

1. Алиев Р.А., Церковный А.Э., Мамедова Г.А. Управление производством при нечеткой исходной информации. – М.: Энергоатомиздат, 2001.
2. Арзамасцев Д.А., Бартоломей П.И., Холян А.М. АСУ и оптимизация режимов энергосистем. – М.: Высшая школа, 1993.
3. Бусленко Н.П. Моделирование сложных систем. – М.: Наука, 1998.
4. Башлыков А.А., Еремеев А.П. Экспертные системы поддержки принятия решений в энергетике. – М.: Изд-во МЭИ, 2004.
5. Веников В.А. Переходные электромеханические процессы в электрических системах. – М.: Высшая школа, 1985.

6. Гельман Г.А. Автоматизированные системы управления энергоснабжением промышленных предприятий.– М.: Энергоатомиздат, 1994.
7. Гордеев А.С. Моделирование в агроинженерии. – Мичуринск: Мичуринский государственный аграрный университет, 2007.
8. ГОСТ 13109-97. Электрическая энергия. Нормы качества электрической энергии в системах электроснабжения общего назначения.– Минск: Изд-во стандартов.
9. ГОСТ 21027-75. Системы энергетические. Термины и определения. – М.: Изд-во стандартов.
10. Заде Л.А. Понятие лингвистической переменной и его приложение к принятию приближенных решений.– М.: Мир, 1996.
11. Кофман А. Введение в теорию нечетких множеств.– М.: Радио и связь, 2002.
12. Кудрин Б.И. Электроснабжение промышленных предприятий.– М.: Интермет Инжиниринг, 2006.
13. Нариньяни А.И. Недоопределенность в системе представления и обработки знаний // Изв. АН СССР. Техническая кибернетика. 1986. № 5.
14. Нечеткие множества в моделях управления и искусственного интеллекта / Под ред. Д.А. Поспелова.– М.: Наука, 1996.
15. Обработка нечеткой информации в системах принятия решений/ А.Н. Борисов, А.В. Алексеев, Г.В. Меркурьева и др. – М.: Радио и связь, 1999.
16. Орловский С.А. Проблемы принятия решений при нечеткой исходной информации.– М.: Наука, 1998.
17. Перегудов Ф.И., Тарасенко Ф.П. Введение в системный анализ. – М.: Высш. шк., 1989.
18. Поспелов Д.А. Ситуационное управление: теория и практика.– М.: Наука, 2006.
19. Правила устройства электроустановок. – М.: НЦ ЭНАС, 2006.
20. Прикладные нечеткие системы / Под ред. Т. Тэрано, К. Асаи, М. Сугэно.– М.: Мир, 2003.

2. ОСНОВЫ ТЕОРИИ ОПТИМИЗАЦИИ

2.1. Характеристика задач оптимизации

В области вычислительной математики сложилось самостоятельное направление по численным методам оптимизации, ориентированное на исследование новых классов задач, разработку методов и построение единого аппарата для анализа сходимости численных процедур. В настоящее время разработано множество численных методов для всех основных теоретически обоснованных классов задач: безусловной минимизации гладких и негладких функций в различных пространствах; условной статической и динамической оптимизации при ограничениях различного типа.

Для большинства методов имеется строгое доказательство, выявлена скорость сходимости и обоснована область применения.

Однако такие проблемы как построение эффективных методов для некоторых специальных типов задач, выбор оптимальных методов с учетом специфики проблемы, создание доступных и работоспособных программ для ЭВМ пока ещё не решены.

Оптимизация технологических процессов долгое время не являлась задачей первостепенной важности. Возможности для практического применения методов в течение длительного времени не существовало в связи с отсутствием и низкой адекватностью математических моделей, недостаточной прикладной пригодностью известных методов оптимизации и низкой вычислительной мощностью ЭВМ.

Вопрос о целесообразности расходования средств на внедрение методов оптимизации оставался спорным. Первыми областями применения современных методов оптимизации принято считать авиацию и космос, где раньше, чем где-либо было признано, что применение методов позволит достичь лучших и более совершенных технических результатов. Большую роль при этом сыграли точные знания о процессах, подлежащих оптимизации, которые обеспечивали получение реальных результатов в короткие сроки, что в свою очередь способствовало проведению систематических исследований и развитию уже известных методов с целью их более широкого практического применения.

Становится очевидным, что решение практических задач требует неформальных действий разработчика, возможности участия в диалоговом режиме в процессе поиска и необходимости проведения предварительных вычислительных экспериментов с целью получения начального приближения.

В наиболее общем смысле теория оптимизации представляет собой совокупность фундаментальных математических результатов и численных методов, ориентированных на нахождение и идентификацию наилучших вариантов из множества альтернатив и позволяющих избежать полного перебора и оценивания возможных вариантов. Процесс оптимизации лежит в основе всей инженерной деятельности, поскольку классические функции инженера заключаются в повышении качества функционирования существующих и в проектировании

новых более эффективных и менее дорогостоящих технических систем и, с другой стороны, разрабатывать методы систем.

Эффективность оптимизационных методов, позволяющих осуществить выбор наилучшего варианта, тесно связана с широким использованием достижений в области математики путем реализации на базе применения вычислительной техники интерактивных вычислительных схем, опирающихся на строго обоснованные логические процедуры и алгоритмы. Поэтому для понимания основ оптимизации требуется привлечение важнейших результатов теории матриц, элементов линейной алгебры и дифференциального исчисления, а также положений математического анализа. Математические понятия и конструкции используются не только для того, чтобы повысить уровень строгости представления материала, но и потому, что они составляют терминологическую базу изложения, которая позволяет упростить описание и определение структурных элементов рассматриваемых вычислительных процедур и облегчить их понимание.

Поскольку размерность инженерных задач, как правило, достаточно велика, а расчеты в соответствии с алгоритмами оптимизации требуют значительных затрат времени, оптимизационные методы ориентированы главным образом на реализацию с помощью ЭВМ.

Проблема оптимизации технологий процессов в настоящее время как никогда актуальна в связи с низкими технико-экономическими показателями существующих агрегатов, большими затратами материальных и энергетических ресурсов на производство единицы продукции. Необходимо создание новых экологически чистых, ресурсо- и энергосберегающих технологий. Много задач связано с реконструкцией и совершенствованием существующих технологий, рациональным использованием имеющихся ресурсов в рамках функционирования нового хозяйственного механизма. Отмеченное предполагает формулировку и решение различных по сложности задач оптимизации.

2.2. Обозначения и терминология

Чтобы выделить определенные аспекты используемой в пособии терминологии, в этом разделе сделаны некоторые предварительные замечания относительно ряда терминов, довольно часто встречающихся в литературе по оптимизации.

Оптимальное решение. Вектор $X^* = [x_1^*, \dots, x_n^*]^T$, удовлетворяющий условию $f(X^*) = \max(\min) f(X)$ называется оптимальной точкой, а соответствующее значение $f(X^*)$ – оптимальным значением целевой функции. Здесь индекс T обозначает вектор-строку. Пара X^* и $f(X^*)$ составляет оптимальное решение. Как показано на рис. 2.1, могут существовать различные типы оптимальных решений, если целевая функция не является унимодальной (т.е. имеющей один экстремум). Глобальное оптимальное решение представляет собой наименьшее значение $f(X)^*$, тогда как локальное или относительное оптимальное решение

представляет собой наименьшее значение $f(X)$ в окрестности некоторого вектора X . Как для глобального, так и для локального минимума

$$f(X^*) \leq f(X)$$

но для глобального оптимального решения это соотношение выполняется для всех $X \in E^n$ либо допустимой области, тогда как для локального оптимального решения это имеет место только для малой области ξ , где $\|X - X^*\| < \xi$. Если принимается во внимание и точность решения, то условие оптимальности можно представить в виде

$$f(X^*) \leq f(X) - \gamma,$$

где γ - некоторая малая величина.

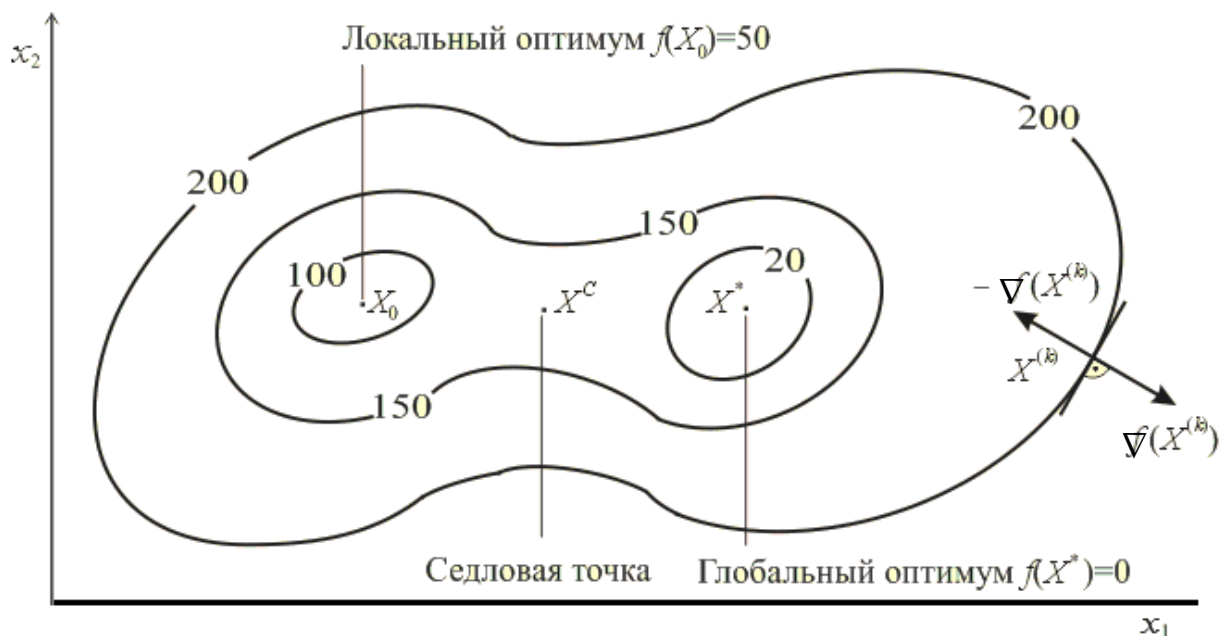


Рис. 2.1. Виды оптимальных решений

Существуют области, в которых целевая функция изменяется незначительно. Такие окрестности получили название седловых точек.

В основном все алгоритмы, описываемые в последующих главах, дают лишь локально оптимальные решения. На практике предположение о том, что локальный экстремум является глобальным, может быть проверено путем использования нескольких начальных векторов. Однако, если найдено только одно решение, ещё недостаточно полагать, что оно является обязательно для данной системы или процесса глобальным.

Допустимость решений. В практических задачах область изменения факторов всегда конечна. Это объясняется наличием конструктивных, техноло-

гических ограничений и определенным запасом ресурсов. Задачи, в которых $X \in E^n$, называют безусловной оптимизацией, а задачи с ограничениями – условной. Выделяют ограничения типа равенств $h(X) = 0$ и неравенств $g(X) \leq 0$.

Любой вектор X , удовлетворяющий ограничениям в виде равенств и неравенств, называется допустимой точкой или допустимым вектором. Множество всех точек, удовлетворяющих ограничениям, образуют допустимую область R функции $f(X)$. Любая точка вне R называется недопустимой. Условный оптимум представляет собой локальный оптимум, лежащий на границе или внутри допустимой области. Если ограничения имеют вид равенств, то допустимый вектор X должен лежать на пересечении всех гиперповерхностей, соответствующих $h_j(X) = 0$. При ограничениях в виде неравенств точка X может быть либо внутренней точкой (допустимой точкой), либо граничной точкой (тоже допустимой точкой), либо внешней точкой (недопустимой точкой). Область допустимых значений переменных может быть односвязной и не односвязной. В последнем случае может оказаться, что алгоритм осуществит обследование одной из допустимых областей.

Методы решения задач безусловной оптимизации рассматриваются во второй главе, а условной в третьей.

Поверхность отклика. Область $f(X)$, в которой осуществляется нахождение оптимального решения, представляет некоторую поверхность в многомерном пространстве. Поверхность отклика наглядно можно представить для одномерного и двухмерного случаев. Для графического представления функций двух переменных используют линии уровня. Это множество точек, для которых целевая функция имеет постоянное значение. Изображение функции с помощью линий уровня показано на рис. 2.1. В топологии поверхности оптимума выделяют несколько характерных видов: круговой холм, эллиптический холм, симметричное или вытянутое седло, стационарный или спадающий овраг и др. Вопросы геометрической интерпретации поверхностей отклика подробно изложены в [1, 13].

Градиент. Если целевая функция непрерывна и дифференцируема, то существует градиент $f(X)$, определяемый как вектор-столбец из первых частных производных $f(X)$ по X , значения которых вычисляются в данной точке $X^{(k)}$. Верхний индекс k , $k = 1, 2, \dots$ используется для обозначения точки в E^n , в которой определяется значение градиента. Таким образом, градиент в точке $X^{(k)}$ равен

$$\nabla f(X^{(k)}) = \begin{bmatrix} \frac{\partial f(X^{(k)})}{\partial x_1} \\ \frac{\partial f(X^{(k)})}{\partial x_n} \end{bmatrix}. \quad (2.1)$$

Выражение $\nabla^T f(X^{(k)})$ обозначает вектор-строку. Доказано, что градиент скалярной функции направлен в сторону наискорейшего увеличения функции,

т.е. наискорейшего подъёма, и ортогонален линии уровня $f(X)$, проходящей через данную точку $X^{(k)}$. Вектор, противоположный этому градиенту (отрицательный градиент), направлен в сторону наискорейшего спуска. На рис. 2.1. для точки $X^{(k)}$ показаны два этих направления.

Норма вектора градиента обозначается через $\|\nabla f(X)\|$ и определяется следующим образом $\|\nabla f(X)\| = [\nabla^T f(X) \nabla f(X)]^{1/3}$.

Вектор единичного направления \hat{S} , совпадающий с направлением градиента рассчитывается по формуле $\hat{S} = \nabla f(X) / \|\nabla f(X)\|$. Следует отметить, что градиент не является направлением наискорейшего увеличения $f(X)$, если рассмотрение ведется не в евклидовой, а в какой-то другой метрике. Например, если определять длину вектора X не по формуле $\|X\| = (X^T X)^{1/2}$, а по формуле $\|X\| = \sum_j \|x_j\|$, то направление наискорейшего увеличения $f(X)$ можно получить,

находя ту компоненту градиента, которая имеет наибольшее абсолютное значение, и полагая соответствующую компоненту x равной либо +1, либо -1 (в зависимости от знака компоненты градиента), а остальные компоненты приравнять к нулю, как, например, в симплекс-методе при линейном программировании.

Составляющие градиента в точке $X^{(k)}$ могут быть определены аналитически, в соответствии с (2.1), а также численно на основе разностных схем или методов регрессионного анализа и планирования эксперимента.

Аппроксимация функций. Для некоторых процедур оптимизации необходимо осуществлять линейную или квадратичную аппроксимацию функций $f(X)$, $g_j(X)$ и $h_j(X)$. Например, линейная или первого порядка, аппроксимация целевой функции $f(X)$ может быть выполнена с помощью усеченного ряда Тейлора в окрестности $X^{(k)}$:

$$f(X) = f(X^{(k)}) + \nabla^T f(X^{(k)})(X - X^{(k)}). \quad (2.2)$$

Квадратичную аппроксимацию $f(X)$ можно получить, отбрасывая в рядах Тейлора члены третьего и более высокого порядков:

$$f(X) = f(X^{(k)}) + \nabla^T f(X^{(k)})(X - X^{(k)}) + \frac{1}{2} (X - X^{(k)})^T \nabla^2 f(X^{(k)})(X - X^{(k)}), \quad (2.3)$$

где $\nabla^2 f(X^{(k)})$ – матрица Гессе $f(X)$, представляющая квадратичную матрицу вторых частных производных $f(X)$ взятых в точке $X^{(k)}$

$$\nabla^2 f(X^{(k)}) = H(X^{(k)}) = \begin{bmatrix} \frac{\partial^2 f(X^{(k)})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(X^{(k)})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(X^{(k)})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(X^{(k)})}{\partial x_n^2} \end{bmatrix}. \quad (2.4)$$

Матрица Гессе, кроме того, используется в методах оптимизации для ускорения процесса поиска за счет более правильного выбора направления движения.

Точность определения экстремальных точек. В настоящее время, в связи с применением ЭВМ, разработано большое множество численных методов нахождения оптимального решения. Любой численный метод предполагает реализацию последовательности многократно повторяющихся действий (итераций), которые должны привести к цели. С каждой итерацией мы должны приближаться к решению задачи. В этом случае говорят, что метод сходится. Всегда возникает вопрос: до каких пор следует выполнять итерации, тем более, что каждая из них может занимать довольно значительное время. Для ответа на этот вопрос вводят понятие точности определения экстремума. В каждом конкретном случае точность задается самим исследователем в зависимости от поставленной задачи и размерности величин.

Можно использовать следующие критерии оценки точности:

- заданная точность определения значения переменных;
- заданная точность определения значения функции в экстремальной точке;
- оценка значений составляющих вектора градиента или его нормы;
- комбинация перечисленных условий или их одновременное выполнение.

Первое условие имеет вид $|X^{(k+1)} - X^{(k)}| \leq \varepsilon$, где $X^{(k+1)}$ и $X^{(k)}$ – значения переменных на $k+1$ и k -ой итерациях, а ε – некоторая константа, определяющая точность поиска. Это условие вполне универсально за исключением случаев с «острым» пиком поверхностей отклика.

Условие второго вида $|f(X^{(k+1)}) - f(X^{(k)})| \leq \varepsilon$ не пригодно для функций с «пологим» экстремумом или «оврагом». В этом случае найденное решение может находиться довольно далеко от точки оптимума.

Критерий оценки точности определения оптимума третьего вида вытекает из необходимых условий существования экстремума ($\nabla f(X^*) = 0$ и $\|\nabla f(X^*)\| = 0$), поэтому проверяется одно из условий: $\nabla f(X^{(k)}) \leq \varepsilon$ или $\|\nabla f(X^{(k)})\| \leq \varepsilon$.

В ряде задач, когда переменные имеют различные диапазоны изменения, необходимо задавать набор ε . Так, например, если x_1 – температура металла, а x_2 – концентрация углерода в металле, то следует задать, например, $\varepsilon_1 = 1$, а $\varepsilon_2 = 0,002$.

Необходимые и достаточные условия существования решений для некоторых случаев рассматриваются ниже в соответствующих разделах.

2.3. Основные этапы решения задач оптимизации

Прежде чем использовать математические процедуры и численные методы теории оптимизации для решения конкретных инженерных задач, необходимо:

- обосновать критерий, на основе которого можно произвести количественный анализ вариантов с целью выявления «наилучшего»;
- установить границы подлежащей оптимизации системы или объекта;
- осуществить выбор внутрисистемных переменных и построить модель, отражающую взаимосвязи между переменными.

Эта последовательность действий составляет первую часть процесса решения задач оптимизации. Корректная постановка задачи служит ключом к успеху и ассоциируется в большей степени с искусством, нежели с точной наукой. Искусство постановки задач постигается в практической деятельности на примерах успешно реализованных разработок и должно быть взаимосвязано с четким представлением преимуществ, недостатков и специфических особенностей различных методов теории оптимизации.

Вторая часть этапов решения содержит:

- выбор способа оптимизации;
- формализацию задачи и выбор метода;
- реализацию задачи;
- анализ полученных результатов.

Кратко остановимся на характеристике некоторых этапов.

Формулировка цели и обоснование критерия. На первом этапе производится выбор критерия, на основе которого можно выявить «наилучший» или множество «наилучших» условий функционирования системы, являющихся целью решаемой задачи. Обычно используют критерии двух видов: экономического плана (капитальные вложения, энергетические и материальные затраты, себестоимость продукции, прибыль в единицу времени, отношение затрат к прибыли) и технологического характера (продолжительность технологических операций, надежность, скорость протекания процессов и др.).

Цель имеет, как правило, содержательный смысл, а критерий должен быть:

- количественно определяемым;
- иметь конечное значение в области изменения факторов;
- обладать чувствительностью к входящим в него переменным.

Формально критерием может быть функция линейного

$$f(X) = c_1x_1 + c_2x_2 + \dots + c_nx_n$$

или произвольного вида, а также функционал, характеризующий, например, суммарные затраты за некоторый промежуток времени

$$Q(U, \tau) = \int_0^{\tau} \varphi[U(\tau)] d\tau.$$

В первом случае имеет место задача статической оптимизации, а во втором – динамической или задача оптимального управления.

При решении практических задач обычно формулируется несколько критериев, чаще всего противоречивых. Такую ситуацию называют неопределенностью цели. Раздел, в котором решаются такие задачи, называют векторной оптимизацией. Для преодоления трудностей можно использовать следующие подходы [12]: линейную свертку; контрольные показатели; выделение главной цели; введение метрики в пространстве целевых функций; построение компромиссов Паретто и др.

В настоящем пособии рассматриваются методы решения задач статической скалярной оптимизации.

Выделение объекта. Этот этап необходим:

- для установления границ изучаемой системы с внешней средой;
- формирования вида допустимой области и характера ограничений.

Под системой понимается совокупность взаимосоединяющихся элементов, действие которых направлено на достижение поставленной цели. Границы системы носят условный характер и задаются условиями, отделяющими систему от внешней среды.

В результате выполнения этого этапа определяют размерность задачи, границы изменения независимых переменных. Вид и характер ограничений зависит от диапазонов изменения переменных, их взаимосвязей и наличия запасов ресурсов на управление. Для установления и формализации взаимосвязей необходим этап математического описания объекта оптимизации.

Выбор способа оптимизации. Применяются два направления для решения задач оптимизации:

- прямая оптимизация на объекте;
- оптимизация на основе математической модели.

В задачах проектирования и создания новых технологий второй подход является основным. Каждое из отмеченных направлений требует решения комплекса вопросов и применения соответствующих методов, имеет свои преимущества и недостатки.

В задачах прямой оптимизации могут быть использованы поисковые методы и методы планирования эксперимента при поиске оптимальных условий [13–15]. Это направление имеет свои специфические особенности (учет влияния помех, нестационарности процессов и т.д.), которые не рассматриваются в данной работе.

Последующие этапы, которые будем рассматривать, относятся ко второму направлению.

Формализация задачи. Решается в два этапа. На первом этапе осуществляется построение математической модели объекта с учетом сформулированного критерия и границ изменения независимых переменных. Затем про-

изводится математическая запись задачи с выделением критерия и системы ограничений.

В задачах статической оптимизации математическая модель преобразуется и является составной частью системы ограничений. В задачах оптимизации управления модель присутствует в явном виде. Результат этого этапа является основанием для выбора метода оптимизации.

Выбор метода. Аналитическое решение задачи оптимизации может быть получено лишь в исключительных случаях, а графическое представление – при размерности не более двух переменных. Поэтому основным способом является численное решение на ЭВМ. Выбор численного метода определяется:

- видом критерия;
- наличием и характером ограничений;
- размерностью задачи.

Общая схема классификации методов скалярной оптимизации показана на рис. 2.2. В данной работе рассматриваются только методы статической оптимизации, более детальная классификация и сущность которых излагается в соответствующих разделах пособия.

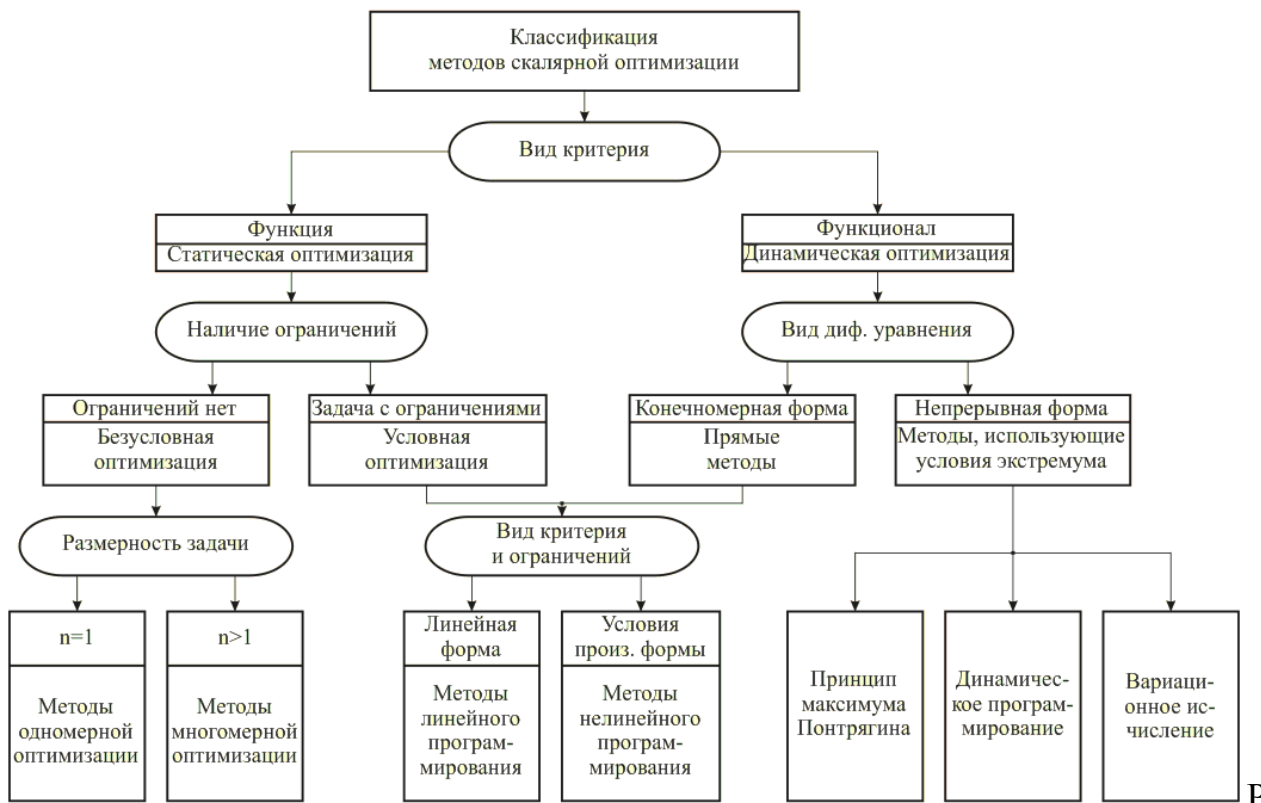
При выборе метода необходимо учитывать скорость сходимости, объем вычислений на каждой итерации, устойчивость к разного рода помехам, сложность реализации и требуемый объем памяти ЭВМ. После выбора метода необходима разработка алгоритмов и программного обеспечения.

Реализация задачи на ЭВМ. Переход от метода на том уровне, который рассматривается в литературе, к алгоритму непростая задача. Необходимо выбрать схемы решения вспомогательных задач, обосновать параметры, определяющие точность решения и критерии прерывания счета и т.д. От этих мало-значущих на идейном уровне подробностей в очень большой мере зависит эффективность алгоритма. Похожие проблемы возникают при реализации на ЭВМ. Это выбор языка программирования, организация массивов данных, процедур ввода и вывода информации.

Работоспособность алгоритма проверяется на нескольких характерных (типовых, контрольных) задачах, в процессе решения которых настраиваются параметры, заменяются некоторые процедуры, устанавливается требуемая точность решения и т.д. Удобным является диалоговый режим, который имеет для таких задач значительные преимущества.

Анализ результатов. Полученные результаты расчета на ЭВМ требуют глубокого анализа по точности, содержательной интерпретации и практической применимости. Точность найденного решения для некоторых задач можно оценить теоретически или экспериментально, осуществляя решение из различных начальных координат, а также с использованием различных критериев останова. Содержательный анализ и проверка соответствия результатов практическим условиям необходимы по той причине, что математическая модель всегда является отражением реальной действительности. Поэтому возможно получение

противоречивых результатов. В таких случаях возникает необходимость корректировки модели и проведение повторного решения поставленной задачи.



ис. 2.2. Общая схема классификации методов скалярной оптимизации

3. МЕТОДЫ РЕШЕНИЯ ЗАДАЧ ИНВАРИАНТНОЙ ОПТИМИЗАЦИИ

Задача безусловной оптимизации состоит в нахождении минимума или максимума функции при отсутствии каких-либо ограничений. Несмотря на то, что практические задачи содержат ограничения, изучение методов безусловной оптимизации важно с нескольких точек зрения. Во-первых, многие алгоритмы решения задач условной оптимизации предполагают сведение ее к последовательности задач без ограничений и, во-вторых, линейный поиск по направлению, применяемый в задачах условной оптимизации, эквивалентен минимизации функции одной переменной. Наконец, рассмотрение особенностей методов безусловной оптимизации является полезным при обосновании и разработке процедур решения задач с ограничениями.

В этой главе рассматриваются численные методы определения экстремума функции одной и нескольких переменных, которые принято называть методами безусловной одномерной и многомерной оптимизации.

3.1. Методы безусловной одномерной оптимизации

Постановка задачи и её анализ. Одномерная оптимизация заключается в нахождении точки x^* , в которой целевая функция $f(x)$ принимает максимальное или минимальное значение. При этом $x \in E^n$ и $n = 2$. Часто в постановках задачи может быть задан интервал $[a, b]$, в котором находится оптимальное значение. Этот интервал можно рассматривать в качестве начального приближения задачи.

Функция $f(x)$ имеет локальный минимум в точке x_0 , если при $\delta > 0$ существует окрестность $[x_0 - \delta, x_0 + \delta]$, такая, что для всех значений x в этой окрестности выполняется условие $f(x_0) \leq f(x)$. Функция $f(x)$ имеет глобальный минимум в точке x^* , если для всех x , принадлежащих E^n или $[a, b]$, справедливо неравенство $f(x^*) \leq f(x)$.

На рисунке 3.1 дано графическое представление функции $f(x)$, которая имеет локальный минимум в точке x_0 и глобальный минимум в точке x^* .

Классический подход к задаче нахождения значений x_0 и x^* состоит в поиске условий, которым они должны удовлетворять. Представленная на рис. 3.1 функция и её производные непрерывны. Известно, что в точках x_0 и x^* производная $f'(x)$ равна нулю. Следовательно, x_0 и x^* будут решениями уравнения

$$f'(x_0) = 0 \text{ или } f'(x^*) = 0. \quad (3.1)$$

Точка x_m , в которой достигается локальный максимум, и точка x_c в которой происходит горизонтальный перегиб функции, также удовлетворяют этому уравнению. Следовательно, уравнение (3.1) является только необходимым условием, но не достаточным.

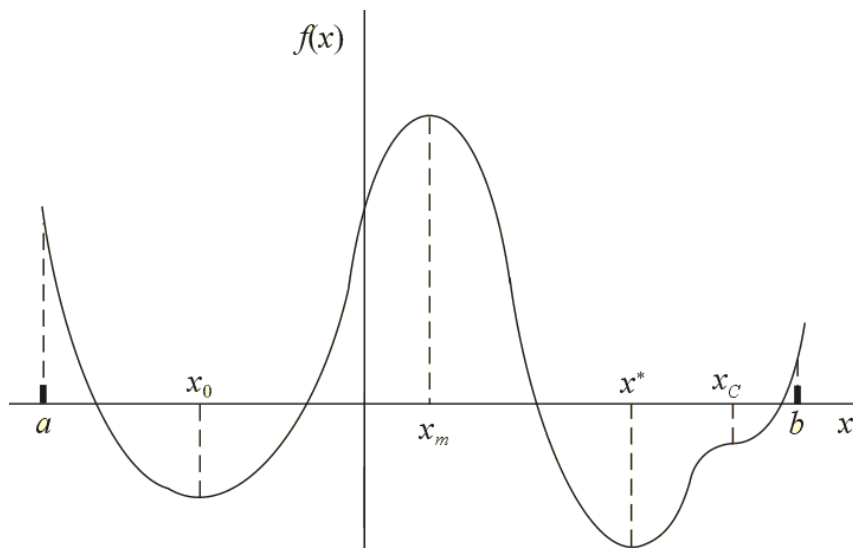


Рис. 3.1. Графическое представление функции, имеющей различные экстремумы

Заметим, что в точках x_0, x^* производная $f'(x)$ меняет знак с отрицательного на положительный. В точке x_m знак меняется с положительного на отрицательный, в то время как в точке x_c он не меняется. Достаточным условием является существование второй производной, знак которой характеризует вид экстремума. В точке \min вторая производная должна быть положительной, а в точке \max – отрицательной.

Однако, если вторая производная равна нулю, ситуация остается неопределенной.

Рассмотренные условия можно пояснить теоретически, если рассмотреть разложение функции $f(x)$ в ряд Тейлора в окрестности, например, точки x_0 , что требует непрерывности функции $f(x)$ и ее производных:

$$f(x_0+h) - f(x_0) = hf'(x_0) + \frac{h^2}{2!} f''(x_0) + \dots \quad (3.2)$$

Если в точке x_0 достигается минимум, то левая часть (3.2) будет неотрицательной для любого достаточно малого h ($|h| < \delta$). Следовательно, первая производная $f'(x_0)$ должна быть равна нулю, и это является достаточным условием (3.1), потому что достаточно малое отрицательное или положительное значение h делало бы правую часть отрицательной.

Так как в следующем члене (3.2) всегда $h^2 > 0$, то если $f''(x_0) > 0$, то в точке x_0 достигается минимум. Если $f'(x_m) = 0$ и $f''(x_m) < 0$, то из аналогичных соображений в точке x_m достигается максимум. Для определения различия между локальным и глобальным минимумами необходимо сравнить значения функций $f(x_0)$ и $f(x^*)$.

Неоднозначность, возникающую при $f''(x) = 0$ можно разрешить, увеличив количество членов в формуле разложения в ряд Тейлора:

$$f(x_0+h) - f(x) = hf'(x_0) + \frac{h^2}{2!} f''(x_0) + \frac{h^3}{3!} f'''(x_0) + \frac{h^4}{4!} f^{(4)}(x_0) + \dots$$

При этом можно сформулировать следующее правило: если функция $f(x)$ и ее производные непрерывны, то точка x_0 является точкой экстремума (максимума или минимума) тогда и только тогда, когда n четное, где n – порядок первой не обращающейся в нуль в точке x_0 производной. Если $f^{(n)}(x_0) < 0$, то в точке x_0 достигается максимум, если $f^{(n)}(x_0) > 0$, то в точке x_0 достигается минимум.

Отмеченные положения носят в большей степени теоретический характер и полезны лишь для понимания сущности задачи. Кроме того, они могут быть использованы при аналитическом нахождении решения. В численных методах используются приближенные условия, на что рекомендуется обратить особое внимание при анализе условий окончания работы алгоритмов.

Задача оптимизации, в которой критерий задан функцией одной переменной, относится к наиболее простому типу. Тем не менее, анализ таких задач занимает одно из центральных мест в оптимизационных исследованиях как

теоретической, так и практической направленности, что предопределило разработку большого числа методов и алгоритмов их решения.

Классификация численных методов. Для определения экстремума функции одной переменной можно применять поисковые методы и методы с использованием производных. Среди методов поиска выделяют методы последовательного сокращения интервала (сканирования, локализации оптимума, половинного деления, золотого сечения, Фибоначчи) и методы точечного оценивания (Пауэлла, квадратичной аппроксимации, обратного переменного шага). На основе использования производных построены методы: средней точки, секущих, кубичной аппроксимации, Ньютона и др.

Теоретические основы указанных методов подробно изложены [1, 11–13, 17, 18], алгоритмы рассматриваются в работах [11, 13, 17], программная реализация некоторых методов приведена в монографии [11].

Для понимания особенностей методов, предпосылок, практической направленности рассмотрим вопросы построения алгоритмов и примеры решения простейших задач.

3.2. Поисковые методы одномерной минимизации

Задача одномерной минимизации на основе поисковых методов последовательного сокращения интервала реализуется в два этапа:

- установление границ исходного интервала;
- уменьшение границ интервала путем выполнения итераций до заданной величины.

Установление границ интервала. На этом этапе на основе априорных данных или правил исключения строится относительно широкий интервал $[a, b]$, содержащий точку оптимума x^* . Обычно поиск граничных точек такого интервала проводится с помощью эвристических методов поиска, хотя в ряде случаев можно также использовать методы экстраполяции. В соответствии с одним из эвристических методов [17], $(k+1)$ -я пробная точка определяется по рекуррентной формуле:

$$x_{k+1} = x_k + 2^k \Delta; k = 0, 1, 2, \dots$$

где x_0 – произвольно выбранная точка;

Δ – подбираемая некоторым способом величина шага.

Знак Δ определяется путем сравнения значений $f(x_0)$, $f(x_0 + |\Delta|)$ и $f(x_0 - |\Delta|)$. Если

$$f(x_0 - |\Delta|) \geq f(x_0) \geq f(x_0 + |\Delta|),$$

то, согласно предположению об унимодальности, точка минимума должна располагаться правее точки x_0 и величина Δ выбирается положительной. Если из-

менить знаки неравенств на противоположные, то Δ следует выбирать отрицательной. Если же

$$f(x_0 - |\Delta|) \geq f(x_0) \leq f(x_0 + |\Delta|),$$

то точка минимума лежит в интервале $[x_0 - |\Delta|, x_0 + |\Delta|]$ и поиск граничных точек завершен. Случай, когда

$$f(x_0 - |\Delta|) \leq f(x_0) \geq f(x_0 + |\Delta|),$$

противоречит предположению об унимодальности. Выполнение этого условия свидетельствует о том, что функция не является унимодальной.

После определения знака Δ производят расчет координат пробных точек и значений функций до тех пор, пока, например, при поиске минимума, не будет выполнено условие

$$f(x_{k-1}) \geq f(x_k) \leq f(x_{k+1}).$$

Полученные при этом значения x_{k-1} и x_{k+1} будут соответствовать границам интервала, внутри которого находится x^* . Схему рассмотренного метода поясним на примере 3.1.

Пример 3.1. Установление границ исходного интервала.

Рассмотрим задачу минимизации функции $f(x) = (100 - x)^2$ при заданной начальной точке $x_0 = 30$ и величине шага $|\Delta| = 5$.

Знак Δ определяется на основе сравнения значений

$$f(x_0) = f(30) = 4900,$$

$$f(x_0 + |\Delta|) = f(35) = 4225,$$

$$f(x_0 - |\Delta|) = f(25) = 5625.$$

Так как $f(x_0 - |\Delta|) \geq f(x_0) \geq f(x_0 + |\Delta|)$, то величина Δ должна быть положительной, а координата точки минимума x^* должна быть больше 30. Имеем $x_1 = x_0 + \Delta = 35$. Далее $x_2 = x_1 + 2\Delta = 45$; $f(45) = 3025 < f(x_1)$

$$x^* > 35. \quad x_3 = x_2 + 2^2\Delta = 65 \quad f(65) = 1225 < f(x_2)$$

$$x^* > 45. \quad x_4 = x_3 + 2^3\Delta = 105 \quad f(105) = 25 < f(x_3)$$

$$x^* > 65. \quad x_5 = x_4 + 2^4\Delta = 185 \quad f(185) = 7225 > f(x_4)$$

Следовательно, $x^* < 185$. Таким образом, шесть шагов вычислений позволили выявить интервал $[65, 185]$, в котором расположена точка x^* . Заметим, что эффективность поиска граничных точек непосредственно зависит от величины шага Δ . Если Δ велико, то получаем грубые оценки координат граничных точек, и построенный интервал оказывается весьма широким. С другой стороны, если величина шага мала, для определения граничных точек может потребоваться достаточно большой объем вычислений.

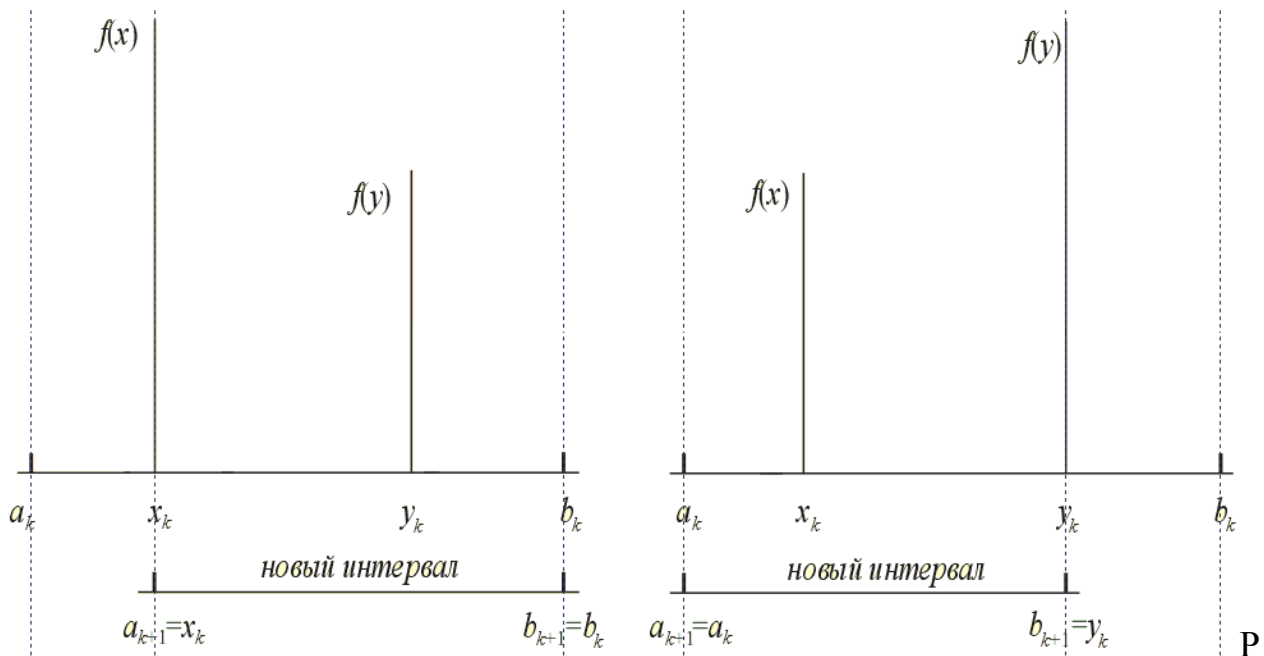
Для установления границ начального интервала, а также предварительного анализа вида функции можно использовать метод сканирования, выбрав при этом рациональную величину шага. Определение границ исходного интер-

вала является как бы заданием начального состояния для последующей оптимизации, поэтому общий успех будет существенно зависеть от качества решения этой задачи.

Реализация этапа уменьшения границ интервала. Так как минимум x^* на $[a, b]$ неизвестен, то этот интервал называется интервалом неопределенности.

Для унимодальной функции $f(x)$ доказано [17], что интервал неопределенности может быть сокращен с помощью вычисления значений функции в двух точках $x, y \in [a, b]$.

На рис. 3.2. показаны два возможных случая. Если $f(x) > f(y)$, то новым интервалом неопределенности является $[x, b]$, поэтому переприсвоение границ интервала на итерациях для этого случая осуществляется по правилу $a_{k+1} = x_k$, а $b_{k+1} = b_k$. При выполнении условия $f(x) \leq f(y)$, новым интервалом будет $[a, y]$ и, соответственно, $a_{k+1} = a_k$, $b_{k+1} = y_k$.



ис. 3.2. Варианты возможного сокращения интервала для унимодальной функции

Рассмотрим несколько процедур минимизации унимодальной функции на замкнутом ограниченном интервале посредством итеративного сокращения интервала неопределенности.

Метод сканирования. Сканирование или равномерный поиск является примером одновременного поиска, когда точки, в которых вычисляется значение функции, выбираются заранее. Начальный интервал неопределенности $[a_1, b_1]$ делится на отрезки величиной δ сеткой из точек $a_1 + k\delta$ для $k = 1, \dots, n$. Тогда $b_1 = a_1 + (n - 1)\delta$. Затем функция $f(x)$ вычисляется в каждой из n точек сетки и определяется \bar{x} , в которой она имеет наименьшее значение. Если $f(x)$ унимодальна, то минимум принадлежит интервалу $[\bar{x} - \delta, \bar{x} + \delta]$.

После n вычислений функции интервал $[a_1, b_1]$ сокращен до длины 2δ . Так как $n = [(b_1 - a_1)/\delta] - 1$, то для достижения более высокой точности нахождения минимума необходимо большое число вычислений функции (табл. 3.1).

Метод локализации оптимума. С целью повышения точности и уменьшения числа расчетов $f(x)$ можно проводить двойное сканирование. В начале определить $[\bar{x} - \delta, \bar{x} + \delta]$, а затем вновь повторить сканирование. Этапы повторяются до достижения заданной точности. Этот метод называется локализацией оптимума. Он работает наиболее эффективно, если текущий интервал неопределенности делится на четыре части.

Метод половинного деления. Является процедурой последовательного поиска. Используется стратегия выбора точек x_1 и y_1 на интервале $[a_1, b_1]$, направленная на минимизацию максимумов из $x_1 - a_1$ и $b_1 - y_1$. Это может быть достигнуто выбором в качестве x_1 и y_1 середины интервала $[a_1, b_1]$. Однако в этом случае будем иметь одну точку и не сможем сократить далее интервал неопределенности. Поэтому, как показано на рис. 3.3, x_1 и y_1 выбраны симметрично на расстоянии $\varepsilon > 0$ от середины интервала каждая. Здесь число $\varepsilon > 0$ настолько мало, чтобы длина нового интервала $\varepsilon + (b_1 - a_1)/2$ являлась достаточно близкой к значению $(b_1 - a_1)/2$ и в то же время, чтобы значения функций $f(x_1)$, $f(y_1)$ были различны [2].

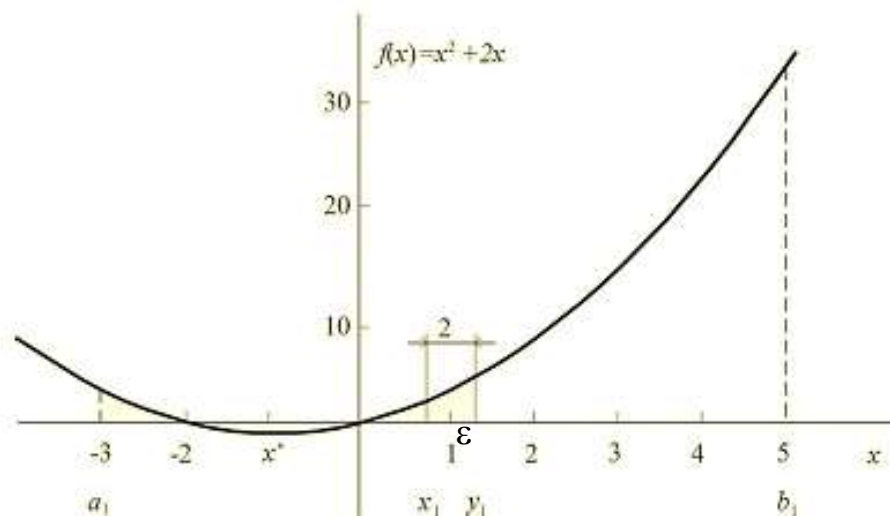


Рис. 3.3. Схема метода половинного деления

Алгоритм метода половинного деления.

Начальный этап. Выбрать константу $2\varepsilon > 0$, допустимую конечную длину интервала $l > 0$, начальный интервал $[a_1, b_1]$. Задать $k = 1$ и перейти к основному этапу.

Основной этап. Шаг 1. Если $b_k - a_k < l$ то остановиться; точка минимума принадлежит интервалу $[a_k, b_k]$, иначе вычислить

$$x_k = \frac{a_k + b_k}{2} - \varepsilon; y_k = \frac{a_k + b_k}{2} + \varepsilon$$

и перейти к шагу 2.

Шаг 2. Если $f(x_k) < f(y_k)$ то $a_{k+1} = a_k$ и $b_{k+1} = y_k$. В противном случае положить $a_{k+1} = y_k$ и $b_{k+1} = b_k$. Заменить $k = k + 1$ и перейти к шагу 1.

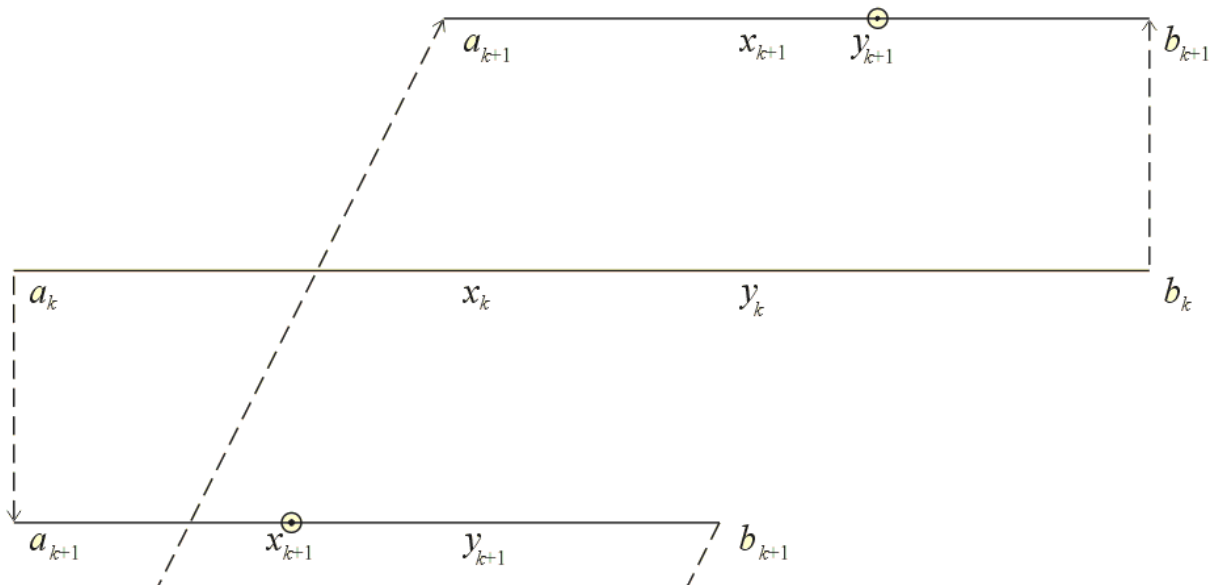
Необходимо отметить, что длина интервала неопределенности в начале $(k + 1)$ -й итерации равна

$$(b_{k+1} - a_{k+1}) = \frac{1}{2^k} (b_1 - a_1) + 2\varepsilon \left(1 - \frac{1}{2^k}\right).$$

Это выражение может быть использовано для определения числа итераций, необходимых для достижения желаемой точности и количества расчетов функции, так как каждая итерация требует двух дополнительных расчетов целевой функции.

Метод золотого сечения. Представляет дальнейшее развитие метода половинного деления, когда при выполнении каждой итерации осуществляется только один раз расчет значения целевой функции.

Основные соотношения метода вытекают из следующих предпосылок [11, 2, 18]. Пусть на k -ой итерации метода золотого сечения как показано на рис. 3.4 интервал неопределенности равен $[a_k, b_k]$. Точки x_k, y_k выбираются на основе следующих условий.



Р

ис. 3.4. Схема метода золотого сечения

1. Длина нового интервала неопределенности

$$b_{k+1} - a_{k+1} = \alpha (b_k - a_k),$$

где $\alpha \in (0, 1)$ называется коэффициентом сжатия.

Кроме того, должны выполняться равенства

$$b_k - x_k = y_k - a_k; x_k = a_k + (1 - \alpha)(b_k - a_k); y_k = a_k + \alpha(b_k - a_k).$$

2. Для новой итерации x_{k+1} и y_{k+1} выбираются так, что x_{k+1} совпадает с y_k , или y_{k+1} совпадает с x_k . Это условие обеспечивает возможность того, что на $(k+1)$ -й итерации потребуется только одно новое вычисление функции.

Чтобы пояснить это, рассмотрим рис. 3.4 и следующие два случая.

Случай 1: Если $f(x_k) > f(y_k)$, то в этом случае $a_{k+1} = x_k$ и $b_{k+1} = b_k$. При $x_{k+1} = y_k$ имеем $y_{k+1} = a_{k+1} + \alpha(b_{k+1} - a_{k+1}) = x_k + \alpha(b_{k+1} - a_{k+1})$. Подставляя в это выражение для x_k и y_k уравнения из предыдущего первого условия, получим, что $\alpha^2 + \alpha - 1 = 0$.

Случай 2: Если $f(x_k) \leq f(y_k)$, то в этом случае $a_{k+1} = a_k$ и $b_{k+1} = y_k$. При $y_{k+1} = x_k$ имеем $x_{k+1} = a_{k+1} + (1 - \alpha)(b_{k+1} - a_{k+1}) = a_k + (1 - \alpha)(y_k - a_k)$. Подставляем в это выражение для x_k и y_k уравнения из предыдущего первого условия также получим $\alpha^2 + \alpha - 1 = 0$.

Корнями уравнения $\alpha^2 + \alpha - 1 = 0$ являются $\alpha = 0,618$ и $\alpha = -2,618$. Так как α должно быть из интервала $(0, 1)$, то принимаем $\alpha = 0,618$.

На первой итерации метода необходимы два вычисления функции в точках x_1 и y_1 , но на каждой последующей требуется только одно вычисление, так как либо $x_{k+1} = y_k$, либо $y_{k+1} = x_k$.

Алгоритм метода золотого сечения.

Начальный этап. Задается конечная длина интервала неопределенности $l > 0$, начальный интервал $[a_1, b_1]$. Вычислить $x_1 = a_1 + (1 - \alpha)(b_1 - a_1)$ и $y_1 = a_1 + \alpha(b_1 - a_1)$, приняв $\alpha = 0,618$. Найти $f(x_1)$, и $f(y_1)$. Задать $k = 1$ и перейти к основному этапу.

Основной этап. Шаг 1. Если $b_k - a_k < 1$, то остановиться; оптимальная точка принадлежит интервалу $[a_k, b_k]$. При $f(x_k) > f(y_k)$, перейти к шагу 2, если $f(x_k) \leq f(y_k)$, то к шагу 3.

Шаг 2. Положить $a_{k+1} = x_k$; $b_{k+1} = b_k$; $x_{k+1} = y_k$; $f(x_{k+1}) = f(y_k)$. Вычислить $y_{k+1} = a_{k+1} + \alpha(b_{k+1} - a_{k+1})$.

Вычислить $f(y_{k+1})$ и перейти к шагу 4.

Шаг 3. Положить $a_{k+1} = a_k$, $b_{k+1} = y_k$, $y_{k+1} = x_k$, $f(y_{k+1}) = f(x_k)$. Вычислить $x_{k+1} = a_{k+1} + (1 - \alpha)(b_{k+1} - a_{k+1})$. Вычислить $f(x_{k+1})$ и перейти к шагу 4.

Шаг 4. Заменить $k = k + 1$ и перейти к шагу 1.

Пример 3.2. Метод золотого сечения.

Минимизировать $f(x) = x^2 + 2x$ на интервале $x \in [-3, 5]$ сократив его до величины $l \leq 0.2$.

Итерация 1.

Начальный этап. $k = 1$, $a_1 = -3$, $b_1 = 5$, $l = 0,2$, $\alpha = 0,618$.

$x_1 = -3 + (1 - 0,618) \cdot (5 - (-3)) = 0,056$, $y_1 = -3 + 0,618(5 - (-3)) = 1,944$;
 $f(x_1) = (0,056) + 2 \cdot 0,056 = 0,115$, $f(y_1) = (1,944) + 2 \cdot 1,944 = 7,667$.

Основной этап. Шаг 1.

$b_1 - a_1 = 5 - (-3) = 8 > 0,2$;

$f(x_1) = 0,115 < f(y_1) = 7,667$, перейти к шагу 2.

Шаг 2.

$$a_2 = a_1 = -3; b_2 = y_1 = 2,944; y_2 = x_1 = 0,056$$

$$f(y_2) = f(x_1) = 0,115; x_2 = -3 + (1-0,618) \cdot (2,944 - (-3)) = -1,112; f(x_2) = -0,987.$$

Шаг 3.

$k = 1 + 1 = 2$, перейти к шагу 1.

Итерация 2. Шаг 1.

$$b_2 - a_2 = 2,944 < f(y_2) = 0,115, \text{ перейти к шагу 3.}$$

Шаг 3.

$$a_3 = a_2 = -3; b_3 = y_2 = 0,056; y_3 = x_2 = -1,112; f(y_3) = f(x_2) = -0,987;$$

$$x_3 = -3 + (1-0,618) \cdot (0,056 - (-3)) = -1,832; f(x_3) = -0,208.$$

Шаг 4. $k = 2 + 1 = 3$, перейти к шагу 1.

После восьми итераций, содержащих девять вычислений функций, интервал неопределенности составил $[-1,112, -0,936]$. При этом $b_9 - a_9 = 0,176$, что меньше заданной величины 0,2. В качестве точки минимума может быть взята середина интервала -1,024. Следует отметить, что искомый минимум равен -1,0.

Метод Фибоначчи. В отличие от рассмотренных поисковых процедур в методе Фибоначчи используется числовая последовательность, поэтому общее число n вычислений функции заранее задается. Это объясняется тем, что точки, в которых производятся вычисления, определяются по зависящим от n формулам

$$x_k = a_k + \frac{F_{n-k-1}}{F_{n-k+1}} (b_k - a_k), \quad k = 1, \dots, n-1;$$

$$y_k = a_k + \frac{F_{n-k}}{F_{n-k+1}} (b_k - a_k), \quad k = 1, \dots, n-2.$$

где $F_j = F_{j-1} + F_{j-2}, j = 3, 4, \dots; F_0 = F_1 = 1$ называется последовательностью чисел Фибоначчи (1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, ...).

В этом методе длина интервала неопределенности на k -й итерации сократится от $b_1 - a_1$ до $(b_1 - a_1)/F_n$.

Поэтому для требуемой точности можно заранее рассчитать величину F_n и количество вычислений функции [11, 2, 18].

Алгоритм метода Фибоначчи.

Начальный этап. Задается конечная длина интервала $l > 0$ и константа различимости $\varepsilon > 0$, начальный интервал $[a_1, b_1]$. Выбрать число вычислений функции n так, чтобы $F_n > (b_1 - a_1)/l$. Вычислить $x_1 = a_1 + (F_{n-2}/F_n)(b_1 - a_1)$, $y_1 = a_1 + (F_{n-1}/F_n)(b_1 - a_1)$; $f(x_1)$; $f(y_1)$. Принять $k = 1$ и перейти к основному этапу.

Основной этап. Шаг 1. Если $f(x_k) > f(y_k)$, то перейти к шагу 2, если $f(x_k) \leq f(y_k)$, то к шагу 3.

Шаг 2. Принять $a_{k+1} = x_k$; $b_{k+1} = b_k$. Затем вычислить $x_{k+1} = y_k$, $y_{k+1} = a_{k+1} + (F_{n-k-1}/F_{n-k})(b_{k+1} - a_{k+1})$. Если $k = n - 2$, то перейти к шагу 3.

Шаг 3. Принять $a_{k+1} = a_k$; $b_{k+1} = y_k$; $x_{k+1} = a_{k+1} + (F_{n-k-2}/F_{n-k}) \cdot (b_{k+1} - a_{k+1})$. Если $k = n - 2$, то перейти к шагу 5, иначе вычислить $f(x_{k+1})$ и перейти к шагу 4.

Шаг 4. Присвоить $k = k + 1$ и перейти к шагу 1.

Шаг 5. Принять $x_n = x_{n-1}$ и $y_n = x_n + \varepsilon$. Если $f(x_n) > f(y_n)$, то $a_n = x_n$ и $b_n = b_{n-2}$. При $f(x_n) \leq f(y_n)$ вычислить $a_n = a_{n-1}$ и $b_n = x_n$ и остановиться.

Оптимальное решение содержится в интервале $[a_n, b_n]$.

Пример 3.3. Метод Фибоначчи.

Рассмотрим задачу примера 3.1.

Минимизировать $f(x) = x^2 + 2x$ на интервале $[-3, 5]$.

Итерация 1.

Начальный этап. $k = 1$, $\varepsilon = 0,01$, $a_1 = -3$, $b_1 = 5$.

$F_n > (5 - (-3))/0,2 = 40$, следовательно, $n=9$; $F_9 = 34$; $F_8 = 21$; $F_7 = 13$.

$x_1 = -3 + 13/34 \cdot (5 - (-3)) = 0,054$; $y_1 = -3 + 21/31 \cdot (5 - (-3)) = 2,945$;

$f(x_1) = 0,112$; $f(y_1) = 7,675$.

Основной этап. Шаг 1. $f(x_1) < f(y_1)$, перейти к шагу 3.

Шаг 3. $a_2 = a_1 = -3$; $b_2 = y_1 = 2,945$; $y_2 = x_1$; $f(y_2) = f(x_1)$;

$x_2 = -3 + (8/21) \cdot (2,945 - (-3)) = -1,109$; $f(x_2) = -0,988$; $k = 1 \neq 9 - 2$, перейти к шагу 4.

Шаг 4. $k = 1 + 1 = 2$, переходим к шагу 1.

После первой итерации интервал неопределенности сократился до $[-3, 2,945]$. При $k = 9$ получается $x_9 = x_8 = -0,963$ и $y_9 = x_9 + \varepsilon = -0,954$.

Конечный интервал неопределенности равен $[-1,109, -0,963]$, длина которого $l = 0,146$.

Середина интервала равна $-1,03$.

Заметим, что длина интервала неопределенности в предыдущем примере была равна $0,176$.

Сравним рассмотренные выше поисковые методы, основанные на последовательном сокращении интервала.

Каждый из рассмотренных в этом параграфе поисковых методов исключения интервала дает за конечное число шагов такую точку \hat{x} , что $|\hat{x} - x^*| \leq l$, где l – длина конечного интервала, а x^* – точка минимума на заданном интервале $[a_1, b_1]$. Необходимое число вычислений функции n может быть определено как наименьшее положительное целое, удовлетворяющее следующим соотношениям [2]:

$$\text{метод равномерного поиска } n \geq \frac{2(b_1 - a_1)}{l} - 1;$$

$$\text{половинного деления } \left(\frac{1}{2}\right)^{n/2} \geq \frac{l}{b_1 - a_1};$$

метод золотого сечения $(0,618)^{n-1} \geq \frac{l}{b_1 - a_1}$;

метод Фибоначчи $\frac{1}{F_n} \geq \frac{l}{b_1 - a_1}$.

В таблице 3.1 приведены данные о количествах вычислений значений функции, необходимых для определения координаты x^* с заданной точностью $\varepsilon = l/(b_1 - a_1)$.

Видно, что наиболее эффективными являются методы Фибоначчи и золотого сечения, которые при достаточно больших n становятся идентичными по производительности, поскольку значение $1/F_n$ стремится к $(0,618)^{n-2}$.

Рассмотрим несколько методов поиска, основанных на точечном оценивании значений функции.

Таблица 3.1

Метод поиска	Заданная точность ε				
	0,1	0,05	0,01	0,001	0,001
Равномерный поиск	19	39	199	1999	19999
Половинного деления	7	9	14	20	27
Золотого сечения	6	8	11	16	20
Фибоначчи	5	7	10	15	19

Метод обратного переменного шага. Предполагает задание начальной координаты x_1 и приращения Δx . Работа алгоритма заключается в следующем.

Рассчитывается координата $x_{k+1} = x_k + \Delta x$.

Если $f(x_{k+1}) < f(x_k)$, то шаг считается «удачным» и его значение увеличивается $\Delta x = \alpha \Delta x$, $\alpha > 2$.

Если $f(x_{k+1}) \geq f(x_k)$, то шаг считается «неудачным». Поэтому направление изменяется на противоположное, а значение шага уменьшается на величину $\beta < 1$ и будет равно $\Delta x = -\beta \Delta x$.

Затем проверяется условие окончания $|x_{k+1} - x_k| \leq \varepsilon$ и поиск продолжается.

Алгоритмически метод очень прост, поэтому часто применяется в многомерном поиске, например, прямом поиске, в методе Розенброка.

Метод квадратичной аппроксимации. Основан на предположении о том, что в ограниченном интервале можно аппроксимировать функцию квадратичным полиномом, который используется для оценивания координаты оптимума [18].

Оценка оптимального значения рассчитывается по формуле

$$\bar{x} = (x_2 + x_1)/2 - (a_1/2a_2).$$

Предполагается, что заданы x_1, x_2, x_3 , и известны значения функции в этих точках f_1, f_2, f_3 , а аппроксимирующая функция

$$g(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2)$$

совпадает с $f(x)$ в трех указанных точках.

Коэффициенты полинома определяются уравнениями

$$a_0 = f_1; a_1 = (f_2 - f_1)/(x_2 - x_1); a_2 = 1/(x_3 - x_2) \cdot [(f_3 - f_1)/(x_3 - x_1) - (f_2 - f_1)/(x_2 - x_1)].$$

Для унимодальных функций \bar{x} оказывается приемлемой для оценки оптимума x^* .

Метод Пауэлла. Заключается в последовательном применении процедуры оценивания с использованием квадратичной аппроксимации. Схему алгоритма можно описать следующим образом [11, 17]. Пусть x_1 – начальная точка, Δx – выбранная величина шага по оси x .

Шаг 1. Вычислить $x_2 = x_1 + \Delta x$.

Шаг 2. Вычислить $f(x_1), f(x_2)$.

Шаг 3. Если $f(x_1) > f(x_2)$, положить $x_3 = x_1 + 2\Delta x$. Если $f(x_1) \leq f(x_2)$, положить $x_3 = x_1 - \Delta x$.

Шаг 4. Вычислить $f(x_3)$ и найти $F_{\min} = \min\{f_1, f_2, f_3\}$.

$X_{\min} = (\cdot) x_i$, которая соответствует F_{\min} .

Шаг 5. По трем точкам x_1, x_2, x_3 вычислить \bar{x} используя формулу для оценивания с помощью квадратичной аппроксимации.

Шаг 6. Проверка на окончание поиска:

- является ли разность $F_{\min} - f(\bar{x})$ достаточно малой?
- является ли разность $X_{\min} - \bar{x}$ достаточно малой?

Если оба условия выполняются, закончить поиск. В противном случае перейти к шагу 7.

Шаг 7. Выбрать «наилучшую» точку (X_{\min} или \bar{x}) и две точки по обе стороны от нее. Обозначить эти точки в естественном порядке и перейти к шагу 4.

Необходимо отметить, что за счет последовательных приближений, совмещенных с квадратичной аппроксимацией, метод имеет высокую эффективность.

3.3. Методы с использованием производных

Рассмотренные в предыдущем разделе методы поиска основываются на предположениях об унимодальности и в ряде случаев о непрерывности исследуемой целевой функции. Целесообразно предположить, что если в дополнении

к условию непрерывности ввести требование дифференцируемости функции, то эффективность поисковых процедур можно существенно повысить [17].

Метод Ньютона. В рамках схемы Ньютона предполагается, что функция f дважды дифференцируема. Работа алгоритма начинается в точках x_1 , которая представляет начальное приближение и строится по рекуррентному соотношению $x_{k+1} = x_k - [f'(x_k)/f''(x_k)]$, где присутствуют первая и вторая производные. Однако, в зависимости от выбора начальной точки и вида функции алгоритм может как сходиться к истинной стационарной точке, так и расходиться. Несмотря на этот недостаток метод Ньютона, требует наименьшего количества расчетов функции.

Метод средней точки. Если функция $f(x)$ унимодальна в заданном интервале поиска, то точкой оптимума является точка, в которой $f'(x) = 0$. Если при этом имеется возможность вычислить как значение функции, так и её производной, то для нахождения корня уравнения $f'(x)=0$ можно воспользоваться эффективным алгоритмом исключения интервала, на каждой итерации которого рассматривается лишь одна пробная точка.

Метод поиска с использованием **кубической аппроксимации**. В соответствии с рассматриваемым методом подлежащая минимизации функция f аппроксимируется полиномом третьего порядка. Логическая схема метода аналогична схеме методов с использованием квадратичной аппроксимации. Однако, в данном случае построение аппроксимирующего полинома проводится на основе меньшего числа точек, поскольку в каждой точке можно вычислить значения как функции, так и её производной.

Сравнение одномерных методов. С помощью теоретических выкладок показано, что такие методы точечного оценивания, как метод Пауэлла или метод поиска с использованием кубической аппроксимации и производных, существенно эффективнее методов последовательного сокращения интервала, среди которых выделяется метод золотого сечения.

Результаты численных экспериментов, представленные в литературе, не подтверждают преимущества метод с использованием производных и квадратичной аппроксимации или метода исключения интервалов над остальными методами. Если для вычисления значений целевой функции требуется значительное машинное время, то, по мнению авторов работы [17], предпочтительнее использовать стратегию поиска, основанную на модификации метода Пауэлла. Это мнение подтверждается результатами вычисленных экспериментов, проведенных Химмельблау [1], который сравнил указанную стратегию поиска со схемой поиска по методу золотого сечения и показал, что модифицированный метод Пауэлла требует меньшего количества вычислений значения функции для достижения заданной точности при оценивании координаты точки минимума. Если необходимо получить решение с очень высокой степенью точности, то лучшими оказываются методы поиска на основе полиномиальной аппроксимации.

Для повышения надежности целесообразно использовать совместно два или три метода.

В заключении необходимо отметить, что достаточно подробное рассмотрение методов одномерной оптимизации не является собственной целью, поскольку в практических задачах критерии с одним фактором встречаются очень редко. Более важным представляются следующие обстоятельства. С методической стороны физическую сущность предмета можно более наглядно раскрыть при рассмотрении именно этих методов. Второй причиной выступает тот факт, что многие методы многомерной оптимизации с ограничениями и без ограничений базируются на линейном поиске вдоль заданного направления, который является одномерным. В качестве заданного направления могут быть: направление градиента, сопряженные или ортонормированные направления и др.

3.4. Методы многомерной безусловной оптимизации

Постановка задачи и её анализ. Для функции n действительных переменных

$$f(x_1, x_2, x_3, \dots, x_n) = f(X), \quad x \in E^n.$$

Требуется найти компоненты вектора X^* , которые дают условие

$$f(X^*) = \min(\max)f(X).$$

Рассматривая локальный X_0 и глобальный X^* экстремумы функции можно отметить их особенности.

Функция $f(X)$ имеет локальный минимум в точке X_0 , если существует окрестность, такая, что $f(X)$ больше $f(X_0)$ во всех точках этой окрестности.

В случае глобального минимума в точке X^* для всех X справедливо неравенство $f(X) \geq f(X^*)$ (см. рис. 2.1).

При таких определениях и очевидных предположениях относительно дифференцируемости можно обобщить уравнение (3.2) и получить

$$\begin{aligned} f(X_0 + h) - f(X_0) &= \sum_{i=1}^n h_i \frac{\partial f(x_1, \dots, x_n)}{\partial x_i} + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n h_i h_j \frac{\partial^2 f(x_1, \dots, x_n)}{\partial x_i \partial x_j} + \dots = \\ &= \nabla^T f(X_0) h + \frac{1}{2} h^T H(X_0) h + \dots \end{aligned} \quad (3.3)$$

Тогда, если x_0 является точкой минимума функции $f(x)$, то каждая первая частная производная $\partial f / \partial x_i$ ($i = 1, \dots, n$) должна обращаться в нуль в точке X_0 .

Если это не так, то соответствующим выбором h_i можно добиться того, что разность $f(X_0 + h) - f(X_0)$ будет отрицательна [11].

Следовательно, необходимым условием минимума в точке X_0 является уравнение $\nabla f(X_0) = 0$, т.е. $\partial f(X_0)/\partial x_i = 0$ ($i = 1, \dots, n$). Тогда знак разности $f(X_0 + h) - f(X_0)$ определяется членом $1/2 \cdot h^T H(X_0) h$, поэтому если матрица $H(X_0)$ положительно определена, то этот член положителен для всех h . Следовательно, необходимыми и достаточными условиями минимума являются: $\nabla f(X_0) = 0$; $H(X_0)$ положительно определена.

Необходимыми и достаточными условиями максимума являются: $\nabla f(X_m) = 0$; $H(X_m)$ – отрицательно определена.

Классификация методов. Выделяют два класса методов: поисковые или методы нулевого порядка и методы с использованием производных.

Поисковые методы используют только значение функции, поэтому их ещё называют методами нулевого порядка. По принципу построения направлений движения поисковые методы можно разделить на две группы. Первая группа основана на одномерном линейном поиске вдоль направлений. К этой группе относятся следующие методы: сканирования, покоординатного и прямого поиска, Пауэлла, Розенброка, случайного поиска и др. В методах покоординатного и прямого поиска осуществляются одномерные движения вдоль координатных осей. Метод Розенброка основан на последовательных одномерных поисках вдоль ортонормированных направлений. В методе Пауэлла одномерный поиск производится вдоль системы сопряженных направлений. Случайный поиск осуществляется по направлениям, заданным случайным образом. Вторая группа движения к оптимуму предполагает расчет значений функции в вершинах многогранника: симплекса или комплекса. Здесь следует выделить методы: обычный симплекс, Нелдера и Мида, комплекс-методы.

Методы с использованием производных делятся также на две группы: первого и второго порядка. К методам первого порядка относятся методы: градиентный, Бокса-Уилсона, наискорейшего спуска, методы сопряженных направлений и сопряженного градиента, переменной метрики и др. К методам второго порядка, которые используют вторые частные производные, относится метод Ньютона и его модификации.

Ниже рассматривается идеология построения методов и работа некоторых алгоритмов при нахождении минимального значения целевой функции.

3.5. Поисковые методы безусловной оптимизации

Простейшим методом является *сканирование*, которое строится аналогично одномерному поиску. Рассчитывается значение целевой функции в узлах сетки, затем определяется координата, в которой функция принимает минимальное значение. Можно проводить локализацию оптимума. Метод очень прост в реализации. Позволяет найти все локальные и глобальные решения поставленной задачи в заданной области определения критерия, а результаты поиска не зависят от вида целевой функции.

Однако у этого метода есть один решающий недостаток. Это большое количество вычислений целевой функции, которое необходимо произвести для достижения требуемой точности. Так, для девяти интервалов по каждой переменной понадобится десять вычислений целевой функции. Пусть время расчета целевой функции на ЭВМ равно одной секунде. Тогда для девяти переменных потребуется 10^9 с для выполнения всего одной итерации, т.е. результат этой итерации будет известен примерно через 31 год непрерывной работы ЭВМ. Этот недостаток особенно резко проявляется при применении метода сканирования с постоянным шагом, когда число разбиений для каждой переменной подбирается таким образом, чтобы шаг был соизмерим с точностью.

Идея метода *поочередного изменения переменных* заключается в следующем: из начальной точки делается шаг по первой переменной, если он «удачен», то переходят к следующей переменной. Если шаг оказался «неудачным», то делается шаг в противоположном направлении. Эта процедура повторяется до тех пор, пока по всем направлениям не будет получаться одни «неудачные» шаги. Поиск продолжается до тех пор, пока абсолютное значение величины шага оказывается меньше заданной точности.

Метод Гауса-Зайделя или циклического координатного поиска заключается в том, что на итерациях по каждой переменной определяется минимум целевой функции вдоль направления координатных осей. После чего осуществляется поиск вдоль другой координаты. Поиск по направлениям, совпадающим с координатными осями, можно проводить любым известным методом одномерной оптимизации, например, методом золотого сечения или обратного переменного шага. Таким образом, задача сводится к многократному использованию метода одномерной оптимизации. Очередность варьирования переменных обычно устанавливается произвольно и в процессе поиска не меняется.

Точность поиска проверяется по сравнению значений переменных и функции на $(k + 1)$ и (k) итерациях. Траектория работы метода для двух итераций показана на рис. 3.5.

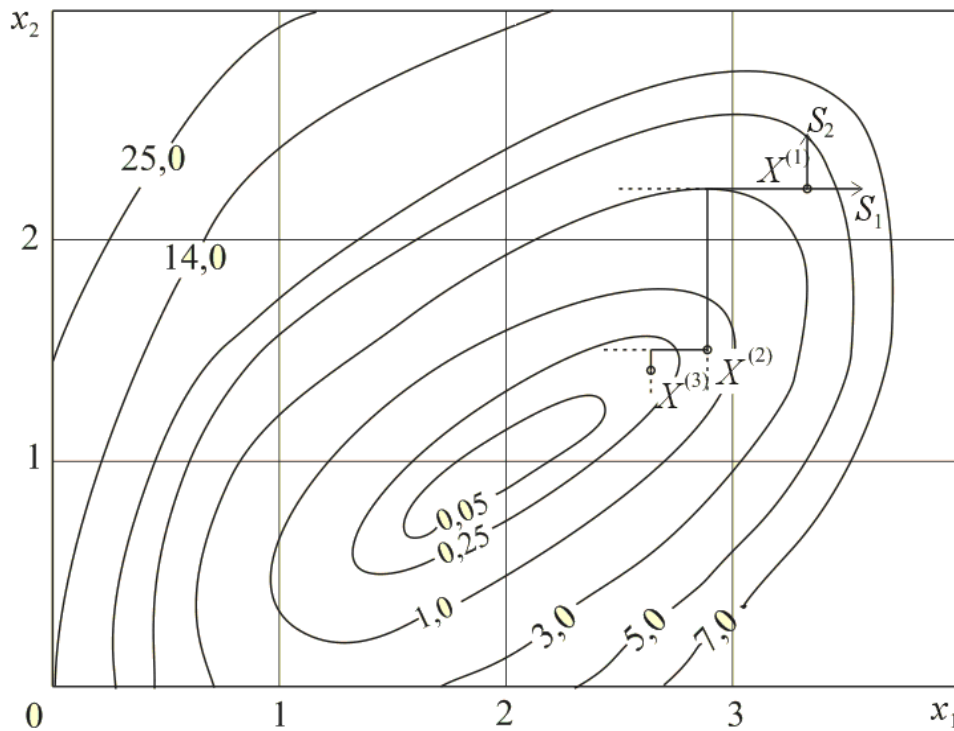


Рис. 3.5. Траектория метода покоординатного поиска

Метод прямого поиска. Хук и Дживс [1] предложили логически простую стратегию поиска, использующую априорные сведения и в то же время отвергающую устаревшую информацию относительно характера топологии целевой функции. Метод включает два этапа: «исследующий поиск» вокруг базисной точки и «поиск по образцу» в направлении, выбранном для минимизации. На рис. 3.6 представлена упрощенная траектория этого метода.

Рассматриваемый метод состоит из следующих операций. Исследующий поиск. Задается начальное приближение $X^{(1)}$ и приращения по координатам ΔX . Рассчитывается значение $f(X^{(1)})$ в базисной точке. Затем в циклическом порядке совершаются пробные шаги. Если приращение улучшает целевую функцию, то шаг считается «удачным». По этой переменной значение изменяется на величину шага и дается приращение по другой переменной (рис. 3.6). Иначе – «неудачным» и делается шаг в противоположном направлении. И если он тоже оказался «неудачным», то значение этой переменной оставляют без изменения, и дается приращение по другой переменной и т.д. пока не будут изменены все независимые переменные. На этом завершается первый исследующий поиск, найдена точка $X^{(2)}$.

Поиск по образцу осуществляется вдоль направления, соединяющего $X^{(2)}$ и $X^{(1)}$. Совершается один или несколько шагов до тех пор, пока шаги являются «удачными».

Шаги исследующего поиска:

—— успех
 неудача

Поиск по образцу:

..... успех
 + + + + + неудача

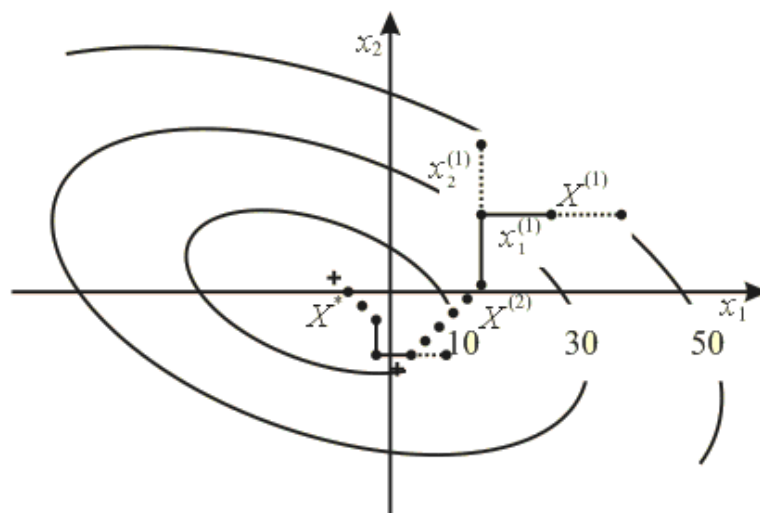


Рис. 3.6. Траектория метода прямого поиска с дискретным шагом

Применяют две модификации метода прямого поиска:

- в исследующем поиске используется одномерная минимизация вдоль координатных направлений (рис. 3.7);
- исследующий поиск осуществляется на основе дискретных шагов по направлениям (рис. 3.6).

Поиск по образцу:

..... успех
 + + + + + неудача

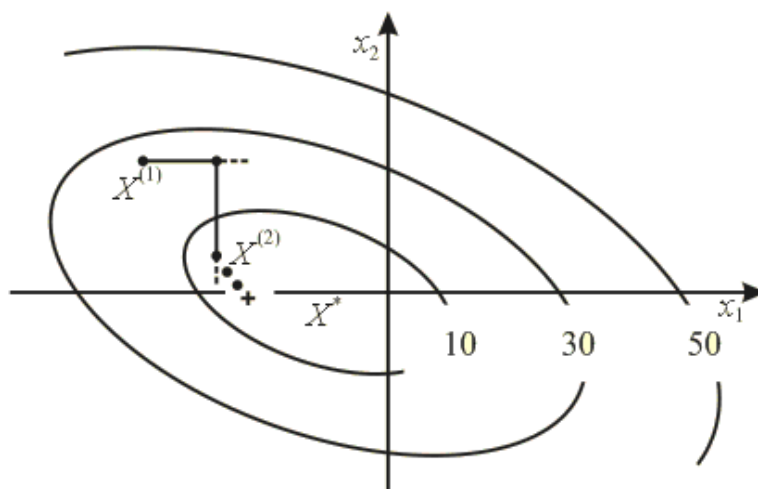


рис. 3.7. Траектория метода прямого поиска
с минимизацией по направлениям

Р

Метод Розенброка является итерационной процедурой, имеющей некоторое сходство с исследующим поиском Хука и Дживса. Отличие состоит в том, что с помощью дискретных шагов (рис. 3.8) или одномерной оптимизации (рис. 3.9) поиски осуществляются вдоль системы ортонормированных направлений $S_1^{(k)}, \dots, S_n^{(k)}$, полученных при помощи процедуры Грама-Шмидта, теоретические вопросы построения которых подробно изложены в [1, 2].

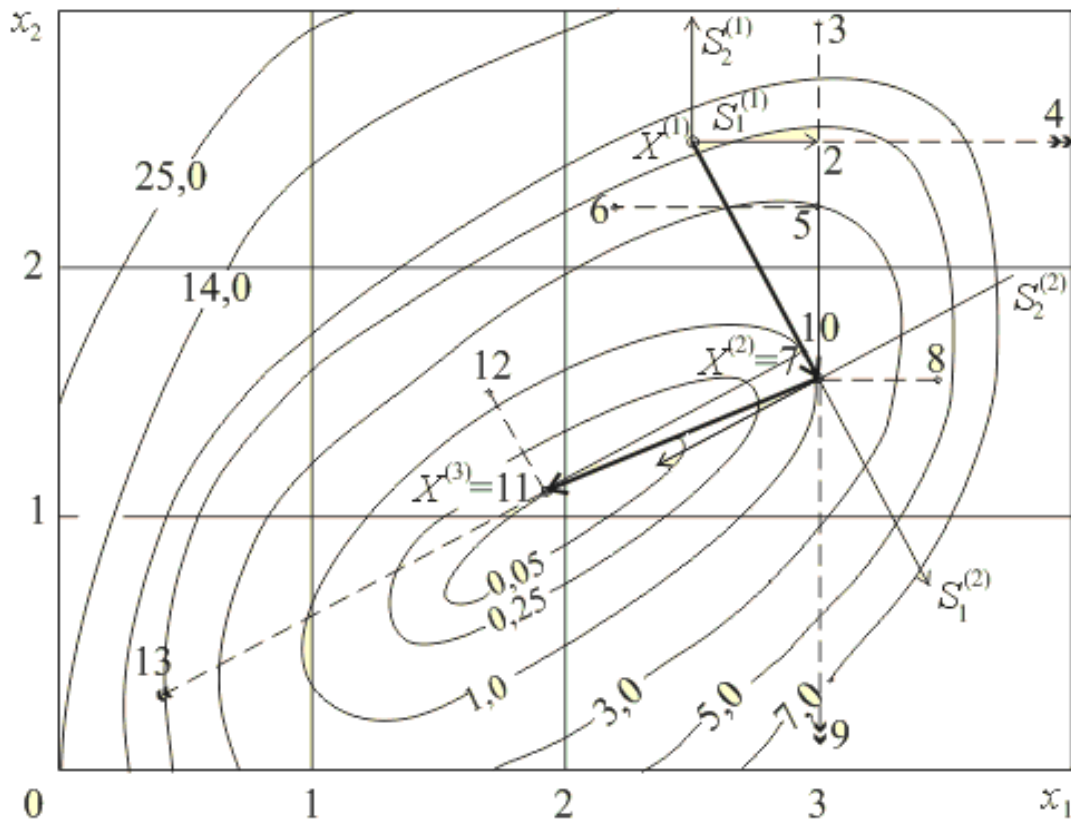


Рис. 3.8. Траектория метода Розенброка с дискретным шагом

Рассмотрим схему метода и алгоритмы его работы.

На первой итерации процесс поиска из начального приближения $X^{(1)}$ осуществляется вдоль координатных осей. Результатом этого процесса будет точка $X^{(2)}$. После этого происходит смена направлений. Причем, как показано на рис. 3.8, единичный вектор направления $\hat{S}_1^{(2)}$ совпадает с вектором перехода из $X^{(1)}$ в $X^{(2)}$, а $\hat{S}_2^{(2)}$ достраивается ортогональное к $\hat{S}_1^{(2)}$.

В общем случае набор ортонормированных векторов $\hat{S}_1^{(k+1)}, \dots, \hat{S}_n^{(k+1)}$ на $(k+1)$ – М этапе вычисляется при помощи следующих соотношений

$$\hat{S}_1^{(k+1)} = \frac{a_1^{(k)}}{\|a_1^{(k)}\|}; \quad (3.4)$$

$$b_2^{(k)} = a_2^{(k)} - \left[(a_2^{(k)})^T \cdot \hat{S}_1^{(k+1)} \right] \cdot \hat{S}_1^{(k+1)};$$

$$\hat{S}_2^{(k+1)} = \frac{b_2^{(k)}}{\|b_2^{(k)}\|}; \quad (3.5)$$

...

$$b_n^{(k)} = a_n^{(k)} - \sum_{i=1}^{n-1} \left[\left(a_n^{(k)} \right)^T \cdot \hat{S}_i^{(k+1)} \right] \cdot \hat{S}_i^{(k+1)} ;$$

$$\hat{S}_n^{(k+1)} = \frac{b_n^{(k)}}{\|b_n^{(k)}\|},$$

где $\|a_i\|$ - норма a_i , являющаяся вектором перехода из $X^{(k)}$ в $X^{(k+1)}$ по направлениям.

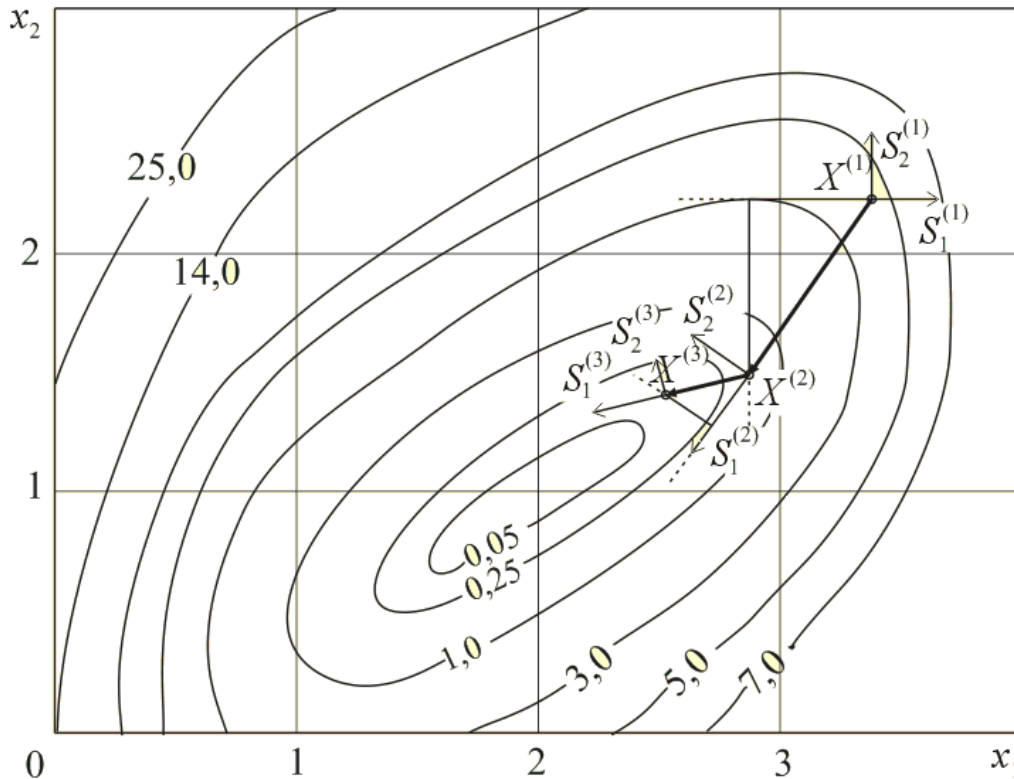


Рис. 3.9. Траектория метода Розенброка с минимизацией по направлению

Векторы a_i рассчитываются по формуле:

$$a_1^{(k)} = \bar{\lambda}_1^{(k)} \cdot \hat{S}_1^{(k)} + \bar{\lambda}_2^{(k)} \cdot \hat{S}_2^{(k)} + \dots + \bar{\lambda}_n^{(k)} \cdot \hat{S}_n^{(k)}$$

$$a_2^{(k)} = \bar{\lambda}_2^{(k)} \cdot \hat{S}_2^{(k)} + \dots + \bar{\lambda}_n^{(k)} \cdot \hat{S}_n^{(k)} \quad (3.6)$$

.....

$$a_n^{(k)} = \bar{\lambda}_n^{(k)} \cdot \hat{S}_n^{(k)}$$

где $\bar{\lambda}_i^{(k)}$ – величина шага, равная переходу из $X^{(k)}$ в $X^{(k+1)}$ по направлениям.

Алгоритм метода Розенброка с минимизацией по направлению. Приведем теперь алгоритм Розенброка, использующий линейный поиск по направлению для минимизации функции нескольких переменных [2].

Начальный этап. Пусть $\varepsilon > 0$ – скаляр, используемый в критерии останова. Выбрать в качестве $\hat{S}_1^{(1)}, \dots, \hat{S}_n^{(1)}$ координатные направления, начальную точку $X^{(1)}$, положить $Y^{(1)} = X^{(1)}$, $k = j = 1$ и перейти к основному этапу.

Основной этап. Шаг 1. Найти λ_j^* – оптимальное решение задачи минимизации $f(Y^{(j)} + \lambda_j \hat{S}_j^{(k)})$ при условии $\lambda \in \varepsilon$ и положить $Y^{(j+1)} = Y^{(j)} + \lambda_j^* \hat{S}_j^{(k)}$. Если $j < n$, то заменить j на $j + 1$ и вернуться к шагу 2. В противном случае перейти к шагу 2.

Шаг 2. Положить $X^{(k+1)} = Y^{(n+1)}$. Если $\|X^{(k+1)} - X^{(k)}\| < \varepsilon$, то остановиться; в противном случае положить $Y^{(1)} = X^{(k+1)}$, заменить k на $k + 1$, положить $j = 1$ и перейти к шагу 3.

Шаг 3. Построить новое множество линейно независимых и взаимно ортогональных направлений в соответствии с (3.5) и (3.6). Обозначить новые направления через $\hat{S}_1^{(k)}, \dots, \hat{S}_n^{(k)}$ и вернуться к шагу 1.

Траектория движения по алгоритму для функции $f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$ показана на рис. 3.9.

Алгоритм метода Розенброка с дискретным шагом.

Начальный этап. Выбрать число $\varepsilon > 0$ для останова алгоритма, коэффициент растяжения $\alpha > 1$ и коэффициент сжатия $\beta \in (-1, 0)$. Взять в качестве $\hat{S}_1^{(1)}, \dots, \hat{S}_n^{(1)}$ координатные направления и выбрать $\lambda_1, \dots, \lambda_n > 0$ начальную длину шага вдоль каждого из направлений. Выбрать начальную точку $X^{(1)}$, положить $Y^{(1)} = X^{(1)}$, $k = j = 1$. При этом $X^{(k)}$ – координаты минимальной точки на k -той итерации, $Y^{(j)}$ – координаты точки на j -том шаге каждой итерации. Перейти к основному этапу.

Основной этап. Шаг 1. Если $f(Y^{(j)} + \lambda_j \hat{S}_j^{(k)}) < f(Y^{(j)})$, то шаг по j -му направлению считается «успешным». Положить $Y^{(j+1)} = Y^{(j)} + \lambda_j \hat{S}_j^{(k)}$ и заменить λ_j на $\alpha \lambda_j$. Если же $f(Y^{(j)} + \lambda_j \hat{S}_j^{(k)}) \geq f(Y^{(j)})$, то шаг считается «неудачным». Положить $Y^{(j+1)} = Y^{(j)}$ и заменить λ_j на $\beta \lambda_j$. Если $j < n$, то заменить j на $j + 1$ и вернуться к шагу 1. В противном случае, т.е. при $j = n$ перейти к шагу 2.

Шаг 2. Если $f(Y^{(n+1)}) < f(Y^{(1)})$, т.е. если хотя бы один спуск по направлению при шаге 1 оказался успешным, положить $Y^{(1)} = Y^{(n+1)}$, $j = 1$ и повторить шаг 1. Пусть $f(Y^{(n+1)}) = f(Y^{(1)})$, т.е. каждый из n последних спусков по направлению на шаге 1 был неудачным. Если $f(Y^{(n+1)}) < f(X^{(k)})$, т.е. по крайней мере один удачный спуск встретился в течении k -й итерации на шаге 1, то перейти к шагу 3. Если $f(Y^{(n+1)}) = f(X^{(k)})$, т.е. не было не одного удачного спуска по направлению, то остановиться. При этом $X^{(k)}$ является приближенным оптимальным решением, если $|\lambda_j| < \varepsilon$ для всех j . В противном случае положить $Y^{(1)} = Y^{(n+1)}$, $j = 1$ и перейти к шагу 1.

Шаг 3. Положить $X^{(k+1)} = Y^{(n+1)}$. Если $\|X^{(k+1)} - X^{(k)}\| < \varepsilon$, то остановиться; $X^{(k+1)}$ – приближенное оптимальное решение. В противном случае вычислить $\lambda_1, \dots, \lambda_n$ из соотношения $X^{(k+1)} - X^{(k)} = \sum_{j=1}^n \lambda_j \hat{S}_j^{(k)}$ построить новые направления в соответствии с (3.5), обозначить их через $\hat{S}_1^{(k+1)}, \dots, \hat{S}_n^{(k+1)}$ положить $Y^{(1)} = X^{(k+1)}$, заменить k на $k + 1$ положить $j = 1$ и перейти к шагу 1.

Дискретные шаги выбираются вдоль n направлений поиска на шаге 1. Если движение вдоль S_j оказалось успешным, то λ_j заменяется на $\alpha\lambda_j$, если же на этом направлении постигла неудача, то λ_j заменяется на $\beta\lambda_j$. Так как $\beta < 0$, то неудача приводит к сдвигу в обратном направлении вдоль j -го вектора на следующей реализации шага 1. Шаг 1 повторяется до тех пор, пока неудача будет иметь место при спуске по каждому направлению поиска. В этом случае строятся новые направления поиска в соответствии с процедурой Грама-Шмидта. Траектория алгоритма показана на рис. 3.8, а принцип работы поясняется ниже в примере.

Пример 3.4. Метод Розенброка с дискретным шагом.

Минимизировать $f(X) = (x_1 - 2)^4 + (x_1 - 2x_2)^3$. Эта функция имеет $X^* = [2 \ 1]^T$, где $f(X) = 0$. Возьмем начальную точку $X^{(1)} = [3.5 \ 3.5]^T$, где $f(X^{(1)}) = 6,4$. $Y^{(1)} = X^{(1)}$.

Параметры алгоритма: $\lambda_1 = \lambda_2 = 0,5$; $\alpha = 3$; $\beta = 0,5$; направления начального поиска совпадают с координатными осями $\hat{S}_1^{(1)} = [1 \ 0]^T$, $\hat{S}_2^{(1)} = [0 \ 1]^T$.

Рассмотрим процесс выполнения первого исследующего поиска и построения новых ортонормированных направлений $S^{(2)}$. Все необходимые построения изображены на рис. 3.8.

Исследующий поиск.

Вычислим $f(Y^{(2)})$ в точке 2

$$Y^{(2)} = \begin{bmatrix} 2,5 + 0,5 \cdot 1 \\ 2,5 + 0,5 \cdot 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 2,5 \end{bmatrix}.$$

Здесь $f(Y^{(2)}) = 5,0$, т.е. имеет место «удача», поэтому шаг по x_1 увеличивается $\lambda_1 = 3 \cdot 0,5 = 1,5$. Затем вычисляем $f(Y^{(3)})$ в точке 3

$$Y^{(3)} = \begin{bmatrix} 3 + 0,5 \cdot 0 \\ 2,5 + 0,5 \cdot 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}.$$

Здесь $f(Y^{(3)}) = 10,0$, т.е. «неудача», поэтому шаг по x_2 уменьшается и направление изменяется на противоположное $\lambda_2 = -0,5 \cdot 0,5 = -0,25$.

При наличии одной «удачи» поиск продолжаем. Считаем $Y^{(1)} = Y^{(3)}$.

Вычисляем $f(Y^{(2)})$ в точке 4

$$Y^{(2)} = \begin{bmatrix} 3 + 1,5 \cdot 1 \\ 2,5 + 1,5 \cdot 0 \end{bmatrix} = \begin{bmatrix} 4,5 \\ 2,5 \end{bmatrix}.$$

Здесь $f(Y^{(2)}) = 39,31$, что говорит о «неудаче», поэтому величина шага уменьшается $\lambda_1 = -1,5 \cdot 0,5 = -0,75$.

Рассчитываем $f(Y^{(3)})$ в точке 5

$$Y^{(3)} = \begin{bmatrix} 3 - 0,25 \cdot 0 \\ 2,5 - 0,25 \cdot 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2,25 \end{bmatrix}.$$

Здесь $f(Y^{(5)}) = 3,25$. Следовательно, шаг является успешным, $\lambda_2 = 3 \cdot 0,25 = 0,75$

Поиск продолжается аналогичным образом до 9 точки. На этом первый исследующий поиск заканчивается, т. к. два последних расчета 8 и 9 неудачны. В качестве результата принимается координата $[3 \ 1,5]$ 7 точки, которая обозначается за $X^{(2)} = Y^{(1)}$.

Построение новых направлений поиска. Новое направление $\hat{S}_1^{(2)}$ совпадает с вектором перехода из $X^{(1)}$ в $X^{(2)}$, а $\hat{S}_2^{(2)}$ достраивается ортогонально к $\hat{S}_1^{(2)}$. Рассчитаем единичные направления $\hat{S}_j^{(2)}$ по формуле (3.5), а векторы $a_1^{(1)}$ и $a_2^{(1)}$ по формуле (3.6).

Определяем составляющие шага, где $\bar{\lambda}^{(1)}$ перехода из $X^{(1)} = [3.5 \ 3.5]^T$ в $X^{(2)} = [3 \ 1,5]^T$. $\bar{\lambda}_1^{(1)} = 3 - 2,5 = 0,5$, $\bar{\lambda}_2^{(1)} = 1,5 - 2,5 = -1$

Находим компоненты векторов

$$a_1^{(1)} = 0,5 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 1 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0,5 \\ -1 \end{bmatrix}; \quad a_2^{(1)} = -1 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix};$$

Рассчитываем направления $\hat{S}_1^{(2)}$ и $\hat{S}_2^{(2)}$

$$\begin{aligned} \hat{S}_1^{(2)} &= \frac{[0,5 - 1]^T}{[0,5^2 + (-1)^2]^{1/2}} = [0,45 - 0,89]^T \\ b_2^{(1)} &= \begin{bmatrix} 0 \\ -1 \end{bmatrix} - [0 - 1] \begin{bmatrix} 0,45 \\ -0,89 \end{bmatrix} = \begin{bmatrix} 0,45 \\ -0,89 \end{bmatrix} \\ \hat{S}_2^{(2)} &= \frac{[-0,4 - 0,21]}{[(-0,4)^2 + (-0,21)^2]^{1/2}} = \begin{bmatrix} -0,89 \\ -0,47 \end{bmatrix}. \end{aligned}$$

На этом завершена первая итерация метода. Точка с 10 по 13 соответствуют результатам исследующего поиска на второй итерации. После 16 итераций получается следующий результат: $X^{(*)} = [1,99959 \ 0,99979]$; $f(X^{(*)}) = 0,285 \cdot 10^{-13}$, что указывает на высокую эффективность метода.

Симплекс-метод. Одним из эффективных методов поиска оптимума функции многих переменных является симплекс метод. Симплексом в пространстве n переменных называется выпуклый многогранник, имеющий $n + 1$ вершину. В пространстве двух переменных – это треугольник, в пространстве трех переменных – тетраэдр. В первоначальном симплексном методе использовался правильный симплекс, т.е. симплекс, все ребра которого равны между собой (например, равносторонний треугольник).

Размещение правильного симплекса в пространстве n переменных может быть осуществлено различными путями. Часто используются два следующих способа.

1. Одна вершина симплекса помещается в начало координат, а остальные вершины располагаются так, чтобы ребра, выходящие из первой вершины, образовали одинаковые углы с соответствующими координатными осями (рис. 3.10). Координаты вершин симплекса в этом случае могут быть заданы следующей матрицей:

Номера вершин	x_1	x_2	x_3	...	x_n
1	0	0	0	...	0
2	P	Q	Q	...	Q
3	Q	P	Q	...	Q
...
$n + 1$	Q	Q	Q	...	P

В этой матрице $P = (1/n \cdot \sqrt{2}) \cdot (\sqrt{n+1} + n - 1)$; $Q = (1/n \cdot \sqrt{2}) (\sqrt{n+1} - 1)$.

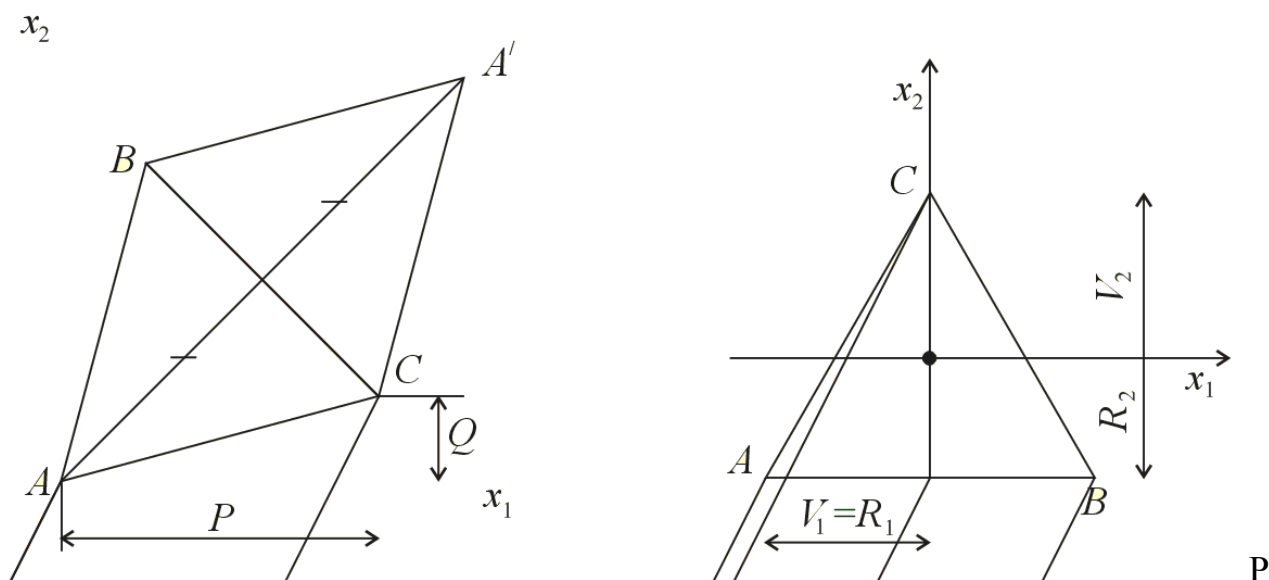
2. Центр симплекса помещается в начало координат, а $(n+1)$ вершина на ось x_n . Остальные вершины располагаются симметрично относительно координатных осей (рис. 3.10). В этом случае координаты вершин симплекса определяются матрицей:

Номера вершин	x_1	x_2	x_3	...	x_n
1	$-R_1$	$-R_2$	$-R_3$...	$-R_n$
2	V_1	$-R_2$	$-R_3$...	$-R_n$
3	0	V_2	$-R_3$...	$-R_n$
...
$n + 1$	0	0	0	...	V_n

где $R_i = 1/\sqrt{2 \cdot i(i+1)}$; $V_i = 1/\sqrt{2 \cdot (i+1)}$.

И в первом, и во втором случаях формулы получены для симплекса, длина ребра которого равна единице. Для произвольной длины каждую формулу надо умножить на длину ребра l .

Идея симплекс метода заключается в следующем. Выбирается начальный симплекс $(A-B-C)$ и в его вершинах рассчитываются значения целевой функции $f(A)$, $f(B)$, $f(C)$. Из этих значений находится «наихудшая» точка: при поиске минимума функции эта та точка, в которой функция принимает наибольшее значение. Например, точка A на рис. 3.10. Через центр противоположной грани (в данном случае это сторона $B-C$) строится новая вершина симплекса A' , симметричная A . В результате получается симплекс $A-B-A'$, причем значения целевой функции в двух его вершинах уже известны. Поэтому вычисляется значение функции в точке A' и среди всех вершин ищется вершина с «наихудшим» значением. Эта вершина (например, C) снова отображается через центр противоположной грани (сторона $B-A'$) и вся процедура повторяется.



ис. 3.10. Варианты построения исходного симплекса

Возникает вопрос: как заканчивать поиск? В случае симплексного метода признаком того, что мы находимся в окрестности экстремума является характерная ситуация, которая называется заикливанием. Эта ситуация возникает, когда значения целевой функции в полученной новой вершине снова оказывается «наихудшим». В этом случае, отображая эту вершину через центр противоположащей грани, мы снова попадаем в отброшенную на предыдущем шаге вершину, и этот процесс будет повторяться бесконечное количество раз.

В простом симплексном методе обнаружение заикливания является признаком того, что поиск экстремума необходимо закончить. Однако в этом случае точность нахождения экстремума будет целиком определяться размером симплекса. Если исходный симплекс имеет большие размеры, то, как правило, маловероятно, что в ходе поиска одна из вершин поиска попадет в экстремум. Если же взять исходный симплекс с малой длиной ребра, то в этом случае резко возрастает количество вычислений целевой функции. Поэтому в применяемых модификациях данного метода при обнаружении заикливания уменьшают размеры исходного симплекса. В окрестности экстремума процедура уменьшения размеров симплекса будет многократно повторяться и в результате симплекс будет постоянно уменьшаться, стягиваясь в точку. Как только размер симплекса станет меньше выбранной точности поиска, процесс оптимизации закончится.

Метод Нелдера и Муда. Этот метод называют еще поиском по деформируемому многограннику. Отмеченные выше недостатки регулярного симплекса, а также отсутствие ускорения поиска и трудности при проведении поиска на искривленных «оврагах» и «хребтах» явились причиной разработки Нелдером и Мидом метода, в котором симплекс меняет свою форму, т. е. представляет деформируемый многогранник. Изменение формы многогранника происходит за счет операций растяжения, сжатия и редукции.

1. *Построение начального многогранника.* Выбирается произвольной формы многогранник с координатами вершин

$$X_i^{(k)} = [x_{i1}^{(k)}, \dots, x_{ij}^{(k)}, \dots, x_{in}^{(k)}]^T, i = 1, \dots, n+1.$$

Индекс (k) будет обозначать k -й этап поиска. Значение целевой функции в $X_i^{(k)}$ равно $f(X_i^{(k)})$.

2. *Расчет координат центра тяжести.* Определяют те векторы x многогранника, которые дают максимальное и минимальное значение $f(X)$, а именно

$$f(X_h^{(k)}) = \max\{f(X_1^{(k)}), \dots, f(X_{n+1}^{(k)})\}; \quad f(X_l^{(k)}) = \min\{f(X_1^{(k)}), \dots, f(X_{n+1}^{(k)})\}.$$

Тогда координаты центра тяжести рассчитываются по формуле

$$x_{n+2,j}^{(k)} = \frac{1}{n} \left[\left(\sum_{i=1}^{n+1} x_{ij}^{(k)} \right) - x_{hj}^{(k)} \right], j = 1, \dots, n,$$

где индекс j обозначает координатное направление.

3. *Отражение.* Представляет проектирование $X_h^{(k)}$ через центр тяжести в соответствии с соотношением

$$X_{n+3}^{(k)} = X_{n+2}^{(k)} + \alpha(X_h^{(k)} - X_{n+2}^{(k)}),$$

где $\alpha > 0$ - коэффициент отражения;

$X_h^{(k)}$ - вершина, в которой функция $f(X)$ имеет наибольшее значение.

В обычном симплексе $\alpha = 1$.

4. *Разряжение.* Происходит в случае, если $f(X_{n+3}^{(k)}) \leq f(X_l^{(k)})$. Вектор $(X_{n+3}^{(k)} - X_{n+2}^{(k)})$ увеличивается в соответствии с соотношением

$$X_{n+4}^{(k)} = X_{n+2}^{(k)} + \gamma(X_{n+3}^{(k)} - X_{n+2}^{(k)}),$$

где $\gamma > 1$ - коэффициент растяжения.

Если $f(X_{n+4}^{(k)}) < f(X_{n+3}^{(k)})$, то $X_h^{(k)}$ заменяется на $X_{n+4}^{(k)}$ и процедура продолжается с этапа 2 при $k = k + 1$. Иначе $X_h^{(k)}$ заменяется на $X_{n+3}^{(k)}$ и также осуществляется переход к этапу 2 при $k = k + 1$.

5. *Сжатие.* Если $f(X_{n+3}^{(k)}) > f(X_i^{(k)})$ для всех $i \neq h$, то вектор $(X_h^{(k)} - X_{n+2}^{(k)})$ уменьшается в соответствии с формулой

$$X_{n+5}^{(k)} = X_{n+2}^{(k)} + \beta(X_h^{(k)} - X_{n+2}^{(k)}),$$

где $0 < \beta < 1$ - коэффициентом сжатия.

Затем $X_h^{(k)}$ заменить на $X_{n+5}^{(k)}$ и происходит возврат к операции 2 при $k = k + 1$.

6. *Редукция*. Происходит при условии, если $f(X_{n+3}^{(k)}) > f(X_l^{(k)})$. Все векторы $(X_i^{(k)} - X_l^{(k)})$ уменьшаются в 2 раза с отсчетом от $X_l^{(k)}$ в соответствии с формулой

$$X_i^{(k)} = X_l^{(k)} + 0,5(X_i^{(k)} - X_l^{(k)}), i=1, \dots, n+1.$$

Затем возвращаемся к операции 2 для продолжения поиска на $(k+1)$ – М шаге.

Для окончания поиска Нилдер и Мид используют критерий вида

$$\left\{ \frac{1}{n+1} \sum_{i=1}^{n+1} [f(X_i^{(k)}) - f(X_{n+2}^{(k)})]^2 \right\}^{1/2} \leq \varepsilon,$$

где ε – заданная точность поиска;

$f(X_i^{(k)})$ – значение функции в центре тяжести.

Поясним работу метода на конкретном примере.

Пример 3.5. Поиск методом деформируемого многогранника.

Рассматриваем задачу минимизации функции $f(X) = (x_1 - 2)^4 + (x_1 - 2x_2)^3$.

Графические построения изображены на рис. 3.11.

В качестве начальных вершин многогранника возьмем следующие:

$$X_1^{(1)} = [0,0 \ 1,0]^T; X_2^{(1)} = [0,5 \ 1,5]^T; X_3^{(1)} = [1,0 \ 1,0]^T.$$

Зададим параметры алгоритма

$$\alpha = 1; \beta = 0,5; \gamma = 2.$$

На первой итерации при $k = 2$. Вычислим значение функции в вершинах исходного многогранника

$$f(X_1^{(1)}) = 20,0; f(X_2^{(1)}) = 11,3; f(X_3^{(1)}) = 2,0.$$

Значит, наихудшей вершиной является $X_h^{(1)} = X_1^{(1)}$, а наилучшей $X_l^{(1)} = X_2^{(1)}$.

Определяем центр тяжести

$$x_{4,1}^{(1)} = \frac{1}{2} [(0,0 + 0,5 + 1,0) - 0,0] = 0,75,$$

$$x_{4,2}^{(1)} = \frac{1}{2} [(1,0 + 1,5 + 1,0) - 1,0] = 1,25.$$

и рассчитываем координату отражения

$$x_{5,1}^{(1)} = 0,75 + 1(0,75 - 0,0) = 1,5,$$

$$x_{5,2}^{(1)} = 1,25 + 1(1,25 - 1) = 1,5.$$

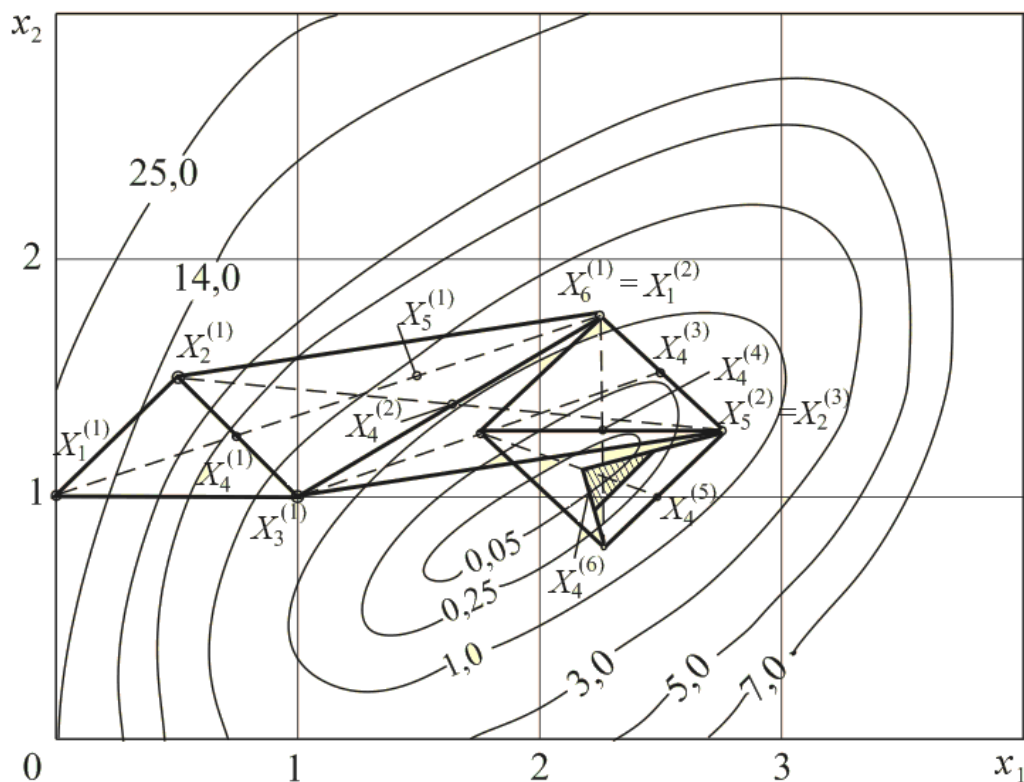


Рис. 3.11. Траектория поиска методом Нельдера-Мида

Значение функции в этой точке $f(X_5^{(1)}) = 1,98$, что меньше чем в $X_1^{(1)}$. Поэтому перейдем к операции растяжения

$$x_{6,1}^{(1)} = 0,75 + 2(1,5 - 0,75) = 2,25,$$

$$x_{6,2}^{(1)} = 1,25 + 2(1,5 - 1,25) = 1,75.$$

Значение функции $f(2,25 \ 1,75) = 1,56$.

Поскольку $f(X_6^{(1)}) < f(X_5^{(1)})$, заменяем $X_1^{(1)}$ на $X_6^{(1)}$ и полагаем $X_6^{(1)} = X_1^{(2)}$ на следующем этапе поиска.

На рис. 3.11 показано шесть итераций работы алгоритма. На втором этапе произведено отражение, затем сжатие отражения, и снова сжатие многогранника. На шестом этапе редукция. Полученный в результате многогранник выделен заштрихованной областью.

После выполнения 43 этапов достигнут результат: $X_4^{(43)} = [2,007 \ 1,003]$, $f(X_4^{(43)}) = 0,31 \cdot 10^{-5}$, который характеризует достаточно высокую эффективность метода.

Методы случайного поиска. Существует специальная группа методов многомерной оптимизации, принципиально отличающихся от рассмотренных выше, в которых используется алгоритм случайного поиска. В ряде случаев методы детерминированного поиска, когда направление и шаг определяются по заранее выбранному правилу, оказываются не в состоянии продолжить поиск или становятся малоэффективными, почему-бы не выбрать направление случайным [10].

К этим методам относятся: комплексный, слепого случайного поиска, случайных направлений и другие.

Комплексный метод возник из симплексного метода и предполагает использование случайных направлений за счет вычеркивания и добавления вершин. В этом методе используется многогранник более чем с $(n + 1)$ вершиной, который называется комплексом. На начальном шаге выбирается произвольная точка, а затем дополнительные вершины достраиваются с помощью псевдослучайных чисел. Затем вычисляется значение функции в вершинах, и вершина, в которой $f(X)$ имеет наихудшее значение заменяется новой, находящейся на прямой, проходящей через отброшенную точку и центр оставшихся точек на расстоянии, равном или большем, чем расстояние от отброшенной точки до центра тяжести. Используются процедуры изменения формы многогранника.

В других методах используется случайный выбор направлений и радиуса поиска.

Применение современных быстродействующих ЭВМ позволяет с помощью таких методов достичь решения даже быстрее, чем каким-либо другим способом, правда, если повезет.

3.6. Методы с использованием производных

Градиентные методы. Стратегия обычного градиентного метода оптимизации без ограничений использует только первые производные целевой функции. На k -М этапе переход из точки $X^{(k)}$ в точку $X^{(k+1)}$ описывается следующим соотношением [1]:

$$X^{(k+1)} = X^{(k)} + \Delta X^{(k)} = X^{(k)} + \lambda^{(k)} \cdot \hat{S}^{(k)}, \quad (3.7)$$

где $\Delta X^{(k)}$ - вектор перехода из точки $X^{(k)}$ в точку $X^{(k+1)}$;

$\hat{S}^{(k)}$ - единичный вектор в направлении $\Delta X^{(k)}$;

$+\lambda^{(k)}$ - скаляр, равный величине шага.

Величина шага $\lambda^{(k)}$ в процессе движения остается постоянной. В ряде случаев предусматривается адаптация к топологии поверхности.

Применение метода градиента при решении задач минимизации было рассмотрено еще известным французским математиком Коши. Градиент целевой функции $f(X)$ в любой точке X есть вектор в направлении наибольшего локального увеличения $f(X)$. Следовательно, нужно двигаться в направлении, противоположном градиенту $f(X)$, поскольку отрицательный градиент $f(X)$ в точке $X^{(k)}$ направлен в сторону наибольшего уменьшения $f(X)$ по всем компонентам X и ортогонален линии уровня $f(X)$ в точке $X^{(k)}$. Введение направления, противоположного нормированному (единичному) градиенту $f(X)$, определяемого в точке $X^{(k)}$ определяется по формуле

$$\hat{S}^{(k)} = -\frac{\nabla f(X^{(k)})}{\|\nabla f(X^{(k)})\|}, \quad (3.8)$$

Подстановка (3.8) в (3.7) следующую формулу перехода из $X^{(k)}$ в $X^{(k+1)}$,

$$X^{(k+1)} = X^{(k)} - \frac{\lambda^{(k)} \nabla f(X^{(k)})}{\|\nabla f(X^{(k)})\|}, \quad (3.9)$$

которая используется в градиентном методе с постоянным шагом.

Траектория градиентного метода показана на рис. 3.13. Здесь отмечен характерный случай, возникающий при переходе к минимуму или в овражных ситуациях, который заключается в зигзагообразном движении. Поэтому величину шага необходимо уменьшать. Одним из возможных подходов к адаптации является расчет угла φ между последовательными векторами шагов. При малых φ величину шага следует делать больше, а при больших – соответственно уменьшать. Это позволит сократить число шагов и повысит работоспособность метода.

Пример 3.6. Обычный градиентный метод.

Минимизировать $f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$

$$\frac{\partial f(X)}{\partial x_1} = 4(x_1 - 2)^3 + 2(x_1 - 2x_2);$$

$$\frac{\partial f(X)}{\partial x_2} = 2(x_1 - 2x_2);$$

$$\|\nabla f(X)\| = \sqrt{\left(\frac{\partial f(X)}{\partial x_1}\right)^2 + \left(\frac{\partial f(X)}{\partial x_2}\right)^2}.$$

В качестве начальной возьмем точку $X^{(1)} = [2,5 \ 2,5]^T$. Величина шага $\lambda^{(k)} = 0,5$ и остается постоянной. Результаты выполнения 5 этапов приведены в таблице 3.2, а траектория на рис. 3.12.

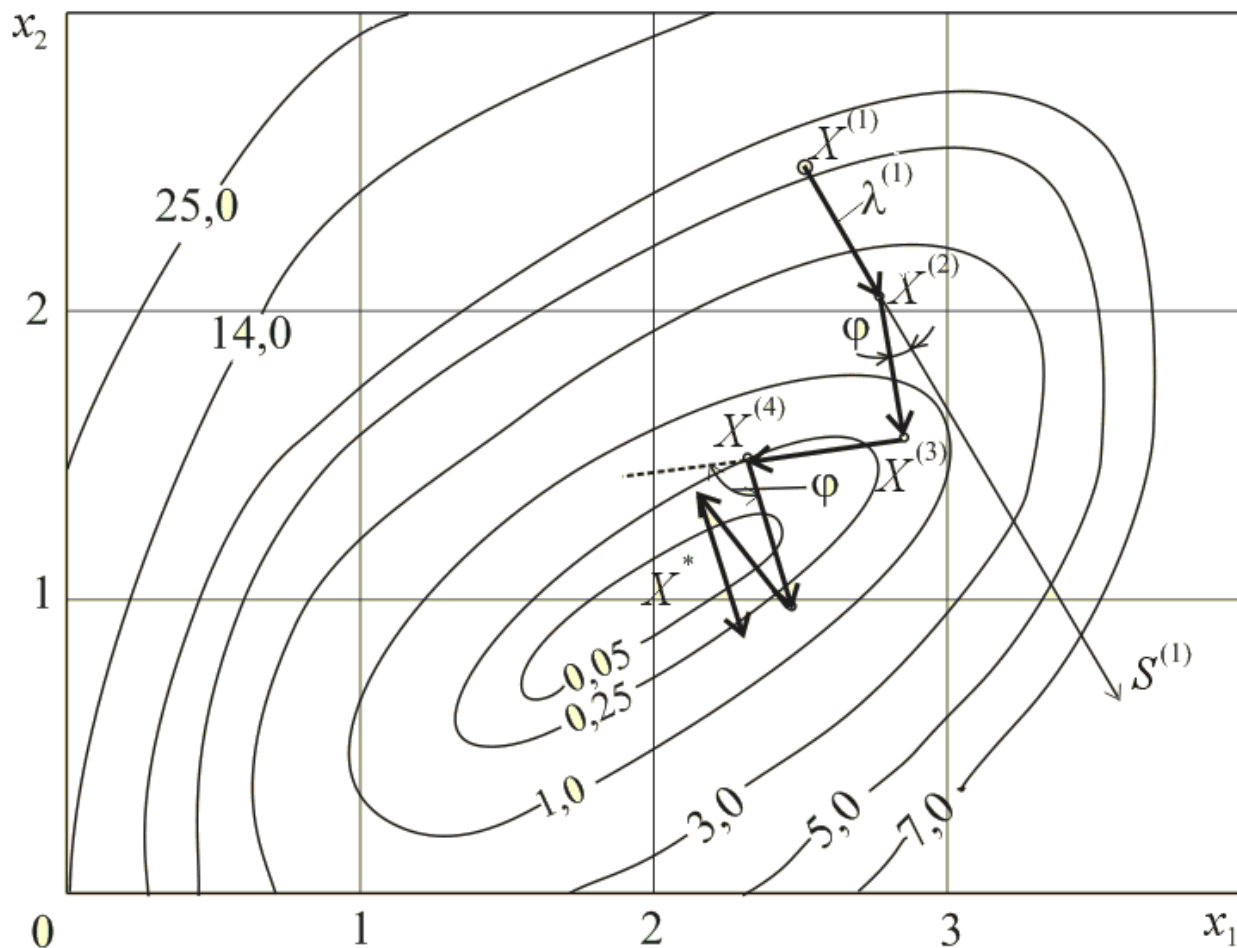


Рис. 3.12. Градиентный метод с постоянным шагом

Таблица 3.2

Этап	x_1	x_2	$\frac{\partial f(X)}{\partial x_1}$	$\frac{\partial f(X)}{\partial x_2}$	$\ \nabla f(X)\ $	Величина шага	
						Δx_1	Δx_2
1	2,5	2,5	-4,5	10	10,96	0,21	-0,46
2	2,71	2,04	-0,59	5,16	5,19	0,06	-0,49
3	2,77	1,55	2,32	0,28	2,33	-0,49	-0,06
4	2,28	1,49	-0,74	1,76	1,91	0,19	-0,46
5	2,47	1,03	2,0	-2,68	3,34	-0,29	0,40
6	2,18	1,43					

Видно, что, начиная с 6 этапа, возникает зигзагообразное движение, которое при $\lambda = 0,5$ не приведет к оптимуму. Необходимо уменьшить величину шага, однако при этом движение в направлении X^* остается очень медленным.

Метод наискорейшего спуска. Отличается тем, что вдоль направления указанного градиентом из $X^{(k)}$ в $X^{(k+1)}$ осуществляется одномерный поиск, тогда вектор перехода может быть определен выражением

$$X^{(k+1)} = X^{(k)} + \lambda^{*(k)} \cdot \hat{S}^{(k)},$$

где $\lambda^{*(k)}$ – оптимальная величина шага.

В точке $X^{(k+1)}$ выбирается новое направление по градиенту и опять совершается одномерный поиск для направления $\lambda^{*(k)}$. Траектория метода наискорейшего спуска показана на рис. 3.14. Величина $\lambda^{*(k)}$ может быть определена любым методом одномерного поиска или рассчитана на основе следующего уравнения:

$$\frac{df(X^{(k)} + \lambda \hat{S}^{(k)})}{d\lambda} = 0. \quad (3.10)$$

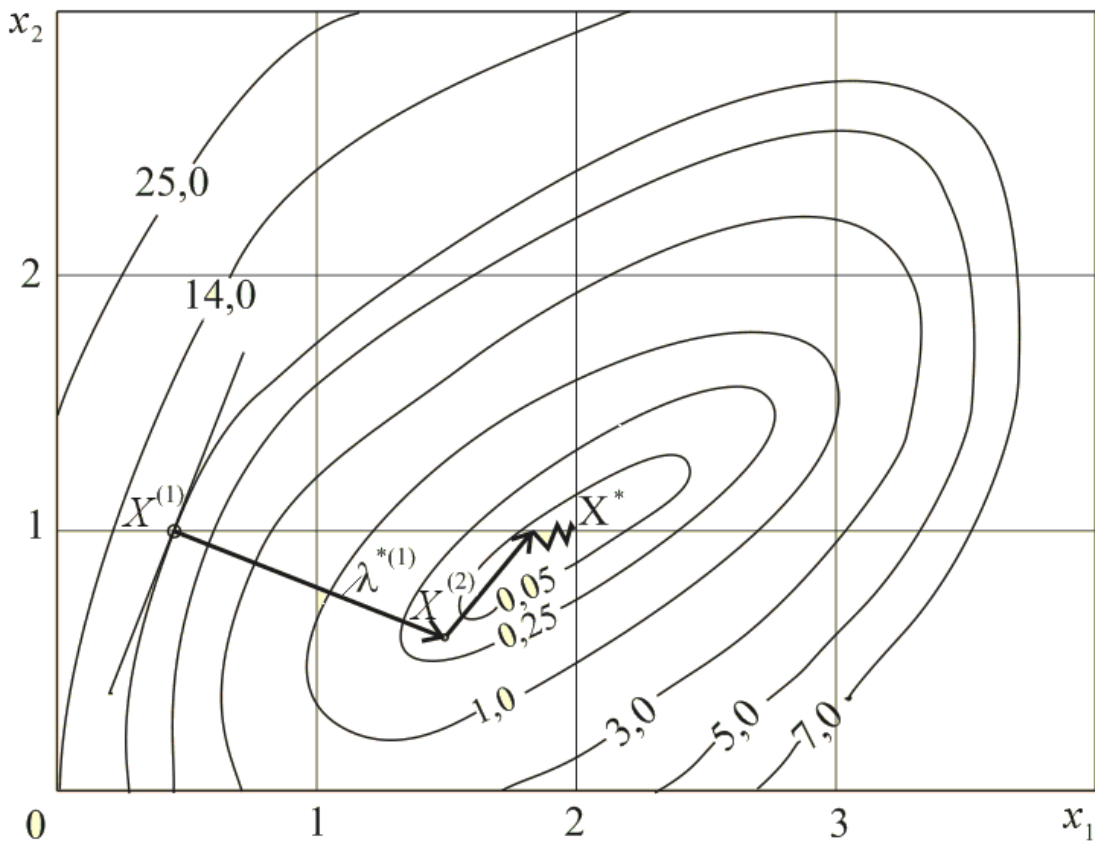


Рис. 3.14. Метод наискорейшего спуска

В уравнение (3.3) вместо (h) подставим $\lambda \hat{S}^{(k)}$. Тогда

$$\frac{df(X^{(k)} + \lambda \hat{S}^{(k)})}{d\lambda} = 0 = \nabla^T f(X^{(k)}) \cdot \hat{S}^{(k)} + (\hat{S}^{(k)})^T \cdot H(X^{(k)}) \cdot \hat{S}^{(k)} \cdot \lambda^{(k)},$$

что дает следующее выражение для $\lambda^{*(k)}$

$$\lambda^{*(k)} = \frac{\nabla^T f(X^{(k)}) \cdot \hat{S}^{(k)}}{(\hat{S}^{(k)})^T \cdot H(X^{(k)}) \cdot \hat{S}^{(k)}}. \quad (3.11)$$

Необходимо отметить, что уравнение (3.11) можно использовать лишь в случае, когда удастся определить матрицу Гессе.

Величину $\lambda^{*(k)}$ можно еще в некоторых случаях найти аналитически. Этот подход рассматривается ниже наиболее подробно в примере.

Пример 3.7. Метод наискорейшего спуска.

В этом примере рассмотрим решение предыдущей задачи. В качестве начальной точки возьмем $X^{(1)} = [0,5 \ 1,0]^T$. Вместо постоянного λ будем определять $\lambda^{*(k)}$ из условия минимума функции в направлении, совпадающем с направлением градиента.

Рассчитаем составляющие градиента в точке $X^{(1)} = [0,5 \ 1,0]^T$. Они равны $\nabla f(X^{(1)}) = [-16,5 \ 6]^T$, а $\|\nabla f(X^{(1)})\| = 17,55$.

Находим единичные направления $\hat{S}_1^{(1)} = 16,5/17,55 = -0,94$, $\hat{S}_2^{(1)} = 6/17,55 = 0,34$.

Чтобы найти новую координату $X^{(2)}$ необходимо выбрать величину шага $\lambda^{*(1)}$, т.е. найти то значение λ , при котором $f(X)$ достигает минимума в направлении, заданном единичным вектором $\hat{S}_1^{(1)}$.

В точке $X^{(1)} = [0,5 \ 1,0]^T$ уравнение (3.9) имеет вид

$$X^{(2)} = \begin{bmatrix} 0,5 \\ 1,0 \end{bmatrix} - \lambda^{(1)} \cdot \begin{bmatrix} -0,94 \\ 0,34 \end{bmatrix}.$$

Отсюда $f(X^{(2)}) = f(\lambda^{(1)}) = (0,5 + 0,92 \cdot \lambda^{(1)} - 2)^4 + [0,5 + 0,94 \cdot \lambda^{(1)} - 2 \cdot (1 - 0,34 \cdot \lambda^{(1)})]^2 = (-1,5 + 0,94 \cdot \lambda^{(1)})^4 + (-1,5 + 1,62 \cdot \lambda^{(1)})^2$

$$\frac{\partial f(\lambda)}{\partial \lambda} = 3,76 \cdot (0,94 \cdot \lambda^{(1)} - 1,5)^3 + 2,56 \cdot (1,62 \cdot \lambda^{(1)} - 1,5).$$

Из условия $\nabla f(\lambda) = 0$ находим $\lambda^{*(1)} = 1,049$ и рассчитываем

$$x_1^{(2)} = 0,5 + 1,049 \cdot 0,94 = 1,49;$$

$$x_2^{(2)} = 1 - 1,049 \cdot 0,34 = 0,64.$$

Аналогичным образом осуществляется выполнение следующих шагов.

Выбор из условия одномерного поиска необходимой величины шага обеспечивает значительно лучшую сходимость этого метода по сравнению с градиентным.

Метод крутого восхождения. Этот метод называют еще методом Бокса и Уильсона [33]. Представляет пошаговую процедуру по поверхности отклика, в которой для оценки составляющих градиента $\hat{\nabla}f(X) = [b_1^{(k)}, b_2^{(k)}, \dots, b_n^{(k)}]$ используется линейное уравнение регрессии в кодированных переменных $\hat{\nabla}f(X) = b_0^{(k)} + b_1^{(k)}x_1^{(k)} + b_2^{(k)}x_2^{(k)} + \dots + b_n^{(k)}x_n^{(k)}$, полученное в результате планирования эксперимента в окрестности точки $X^{(k)}$. Затем совершается движение (крутое восхождение) по направлению $\hat{\nabla}f(X^{(k)})$, которое определяется произведением коэффициента $b_j^{(k)}$ на интервал варьирования переменных ΔZ_j при планировании эксперимента в окрестности $X^{(k)}$. Это движение можно рассматривать как этап, аналогичный методу наискорейшего спуска, в котором решается зада-

ча одномерного поиска. В полученной точке $X^{(k+1)}$ снова производится планирование эксперимента и выбор нового направления движения.

Необходимо отметить, что данный метод по принципу используемой информации (только значение функции) можно отнести к поисковым методам, а по способу движения – к градиентным. Кроме того, появляется возможность свертки информации о поверхности отклика в виде совокупности линейных уравнений регрессии. Отмеченные особенности характеризуют положительные достоинства этого метода.

Методы, использующие сопряженные направления отличаются от метода наискорейшего спуска тем, что одномерный поиск осуществляется в сопряженных направлениях. Важной особенностью является быстрая сходимость за счет того, что при выборе направления используется матрица Гессе, которая описывает область топологии поверхности отклика в отличие от градиента, характеризующего направление только для данной точки. В частности, если целевая функция квадратична, то можно получить точку минимума не более чем за количество шагов, равных размерности задачи [1, 2].

Предположим, что процесс поиска начинается в $X^{(1)}$ с произвольным единичным направлением $\hat{S}^{(1)}$. Тогда результатом первого одномерного поиска будет

$$X^{(2)} = X^{(1)} + \lambda^{*(1)} \cdot \hat{S}^{(1)},$$

где $\lambda^{*(1)}$ – результат минимизации $f(X^{(1)} + \lambda \cdot \hat{S}^{(1)})$ по λ .

Затем новое направление $\hat{S}^{(2)}$ выбирается сопряженным с $\hat{S}^{(1)}$ из условия

$$(\hat{S}^{(2)})^T \cdot [\nabla^2 f(X^{(1)})] \cdot \hat{S}^{(1)} = 0. \quad (3.12)$$

В общем случае система n – линейно независимых направлений поиска $S^{(1)}, S^{(2)}, \dots, S^{(n-1)}$ называется сопряженной по отношению к некоторой положительно определенной (квадратичной) матрице Q , если

$$(\hat{S}^{(2)})^T \cdot Q \cdot S^{(j)} = 0, \quad 0 \leq i \neq j \leq n - 1.$$

Сопряженность – понятие, аналогичное ортогональности, когда $H = 1$.

Для квадратичной функции, преобразованной к каноническому виду, $f(X) = \tilde{b}_{11}x_1^2 + \tilde{b}_{22}x_2^2$, собственные значения $1/2 \cdot \tilde{H}$ находятся на диагонали

$$\tilde{H} = \begin{bmatrix} \tilde{b}_{11} & 0 \\ 0 & \tilde{b}_{22} \end{bmatrix}.$$

Значения и знак коэффициентов канонической формы характеризует топологию поверхности [1], поэтому использование матрицы Гессе и обеспечивает высокую эффективность методов сопряженных направлений.

Пример 3.8. Метод сопряженных направлений.

Рассмотрим задачу минимизации функции из начальной точки $X^{(1)} = [2 \ 10]^T$. Траектория показана на рис. 3.14.

$$F(X) = 4(x_1 - 5)^2 + (x_2 - 6)^2;$$

$$\nabla f(X) = \begin{bmatrix} 8(x_1 - 5) \\ 2(x_2 - 6) \end{bmatrix};$$

$$H(X) = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}.$$

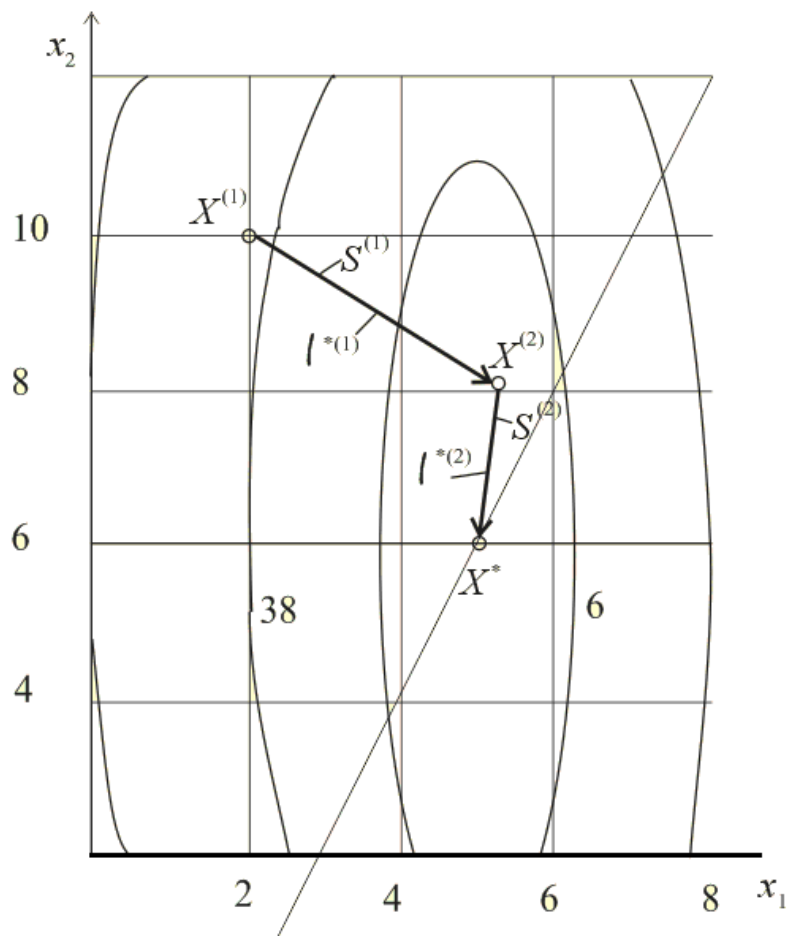


Рис. 3.14. Метод сопряженных направлений

Возьмем произвольное начальное единичное направление $\hat{S}^{(1)} = [0,87 \ -0,50]$. По формуле (3.11) определим длину шага

$$\lambda^{*(1)} = - \frac{[-24 \ 8] \cdot \begin{bmatrix} 0,87 \\ -0,5 \end{bmatrix}}{[0,87 \ -0,5] \cdot \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0,87 \\ -0,5 \end{bmatrix}} = \frac{24,88}{6,55} = 3,79.$$

$$\text{Тогда } X^{(2)} = \begin{bmatrix} 2 \\ 10 \end{bmatrix} - 3,39 \cdot \begin{bmatrix} 0,87 \\ -0,50 \end{bmatrix} = \begin{bmatrix} 5,3 \\ 8,1 \end{bmatrix}.$$

Следующее направление $\hat{S}^{(2)}$ выбираем сопряженным к $\hat{S}^{(1)}$. Для этого используем два уравнения: сопряженности (3.12) и единичной нормировки.

$$\begin{aligned} [\hat{S}_1^{(2)} \cdot \hat{S}_2^{(2)}] \cdot \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0,87 \\ -0,5 \end{bmatrix} &= 0; \\ [\hat{S}_1^{(2)} \cdot \hat{S}_2^{(2)}] \cdot \begin{bmatrix} \hat{S}_1^{(2)} \\ \hat{S}_2^{(2)} \end{bmatrix} &= 1. \end{aligned}$$

Из этих уравнений следует, что

$$\left. \begin{aligned} 6,96 \cdot \hat{S}_1^{(2)} - \hat{S}_2^{(2)} &= 0 \\ (\hat{S}_1^{(2)})^2 + (\hat{S}_2^{(2)})^2 &= 1 \end{aligned} \right\}, \hat{S}_1^{(1)} = \begin{bmatrix} 0,14 \\ 0,98 \end{bmatrix}.$$

Далее определяем $\lambda^{*(2)}$

$$\lambda^{*(2)} = - \frac{\begin{bmatrix} -24 & 4,2 \end{bmatrix} \cdot \begin{bmatrix} 0,14 \\ 0,99 \end{bmatrix}}{\begin{bmatrix} 0,14 & 0,99 \end{bmatrix} \cdot \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0,14 \\ 0,99 \end{bmatrix}} = -2,12$$

и находим конечный результат поиска

$$X^{(3)} = \begin{bmatrix} 5,3 \\ 8,1 \end{bmatrix} - 2,12 \cdot \begin{bmatrix} 0,14 \\ 0,99 \end{bmatrix} = \begin{bmatrix} 5,003 \\ 6,001 \end{bmatrix}.$$

Дальнейшим развитием методов сопряженных направлений являются методы переменной метрики и сопряженных градиентов, в которых используется аппроксимация обратной матрицы Гессе и предыстории направлений поиска [1, 17].

Метод Ньютона. В случае функции нескольких переменных метод Ньютона является процедурой, которая отклоняет направление наискорейшего спуска умножением его на матрицу, обратную к матрице Гессе. Эта операция объясняется нахождением подходящего направления для квадратичной аппроксимации функции, в то время как в градиентном поиске выбор направления связан с минимизацией линейной аппроксимации функции.

Квадратичная аппроксимация функции $f(X)$, при замене в уравнении (2.4) $X - X^{(k)}$ на величину $\Delta X^{(k)} = X^{(k+1)} - X^{(k)}$, имеет следующий вид:

$$f(X^{(k+1)}) = f(X^{(k)}) + \nabla^T f(X^{(k)}) \Delta X^{(k)} + \frac{1}{2} (\Delta X^{(k)})^T \nabla^2 f(X^{(k)}) \Delta X^{(k)}.$$

Минимум этой функции в направлении $\Delta X^{(k)}$ определяется дифференцированием $f(X)$ по каждой из компонент ΔX и приравниванием нулю полученных выражений. Последнее приводит к

$$\Delta X^{(k)} = - [\nabla^2 f(X^{(k)})]^{-1} \nabla f(X^{(k)}),$$

где $-\nabla^2 f(X^{(k)})^{-1}$ - матрица, обратная матрице Гессе.

Подстановка полученного уравнения в (3.7) дает формулу перехода из $X^{(k)}$ в $X^{(k+1)}$ по методу Ньютона:

$$X^{(k+1)} = X^{(k)} - [\nabla^2 f(X^{(k)})]^{-1} \nabla f(X^{(k)}).$$

В отличие от рассмотренных выше способов оптимизации, данный метод не использует параметров, обладает высокой сходимостью и является одной из наиболее эффективных процедур. Однако, применение данного метода сдерживается необходимостью наличия аналитического описания критерия или численной оценки $\hat{\nabla}^2 f(X^{(k)})$ и $\hat{\nabla} f(X^{(k)})$, что является сложной задачей в практической реализации.

4. МЕТОДЫ РЕШЕНИЯ ЗАДАЧ СТАТИЧЕСКОЙ УСЛОВНОЙ ОПТИМИЗАЦИИ

Задача условной оптимизации заключается в нахождении экстремума целевой функции при заданных ограничениях произвольного вида.

В этой главе рассмотрим методы статической условной оптимизации, в которых исключаются ограничения, где присутствует параметр время, и не используются дифференциальные уравнения. Это область оптимального управления и динамической оптимизации.

В этом классе задач в зависимости от вида критерия и ограничений выделяют линейное и нелинейное программирование. Особенности постановок задач и классификация методов этих направлений рассмотрена ниже.

4.1. Линейное программирование

Постановка задачи. Линейное программирование состоит в оптимизации линейной целевой функции на многогранном множестве. Многие практические задачи могут быть сформулированы как задачи линейного программирования.

Постановка задачи в стандартной форме с m ограничениями и n переменными имеет вид:

Минимизировать

$$f(X) = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \tag{4.1}$$

при ограничениях $a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1;$

$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2; \quad (4.2)$$

$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m;$$

или для произвольной строки

$$\sum_{j=1}^n a_{ij}x_j = b_i.$$

При этом $x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0$.

Постановку задачи можно привести к каноническому виду и записать в матричной форме: минимизировать

$$f(C, X) = C^T X \quad (4.3)$$

при ограничениях

$$AX = b, X \geq 0, \quad (4.4)$$

где A – матрица коэффициентов; X – вектор переменных; C – вектор оценок задачи; b – вектор ресурсов.

Каноническую форму всегда можно получить путем известных в математике преобразований.

Например, неравенство вида $\sum_{j=1}^n a_{ij}x_j \leq b_i$, можно заменить равенством,

вводя неотрицательную дополнительную переменную x_{j+1} , так что

$$\sum_{j=1}^n a_{ij}x_j + x_{j+1} = b_i.$$

Выделяют целочисленное линейное программирование, когда переменные принимают лишь целые значения, и непрерывное, когда переменные имеют вещественные значения.

Геометрическая интерпретация задачи линейного программирования рассматривается ниже при решении конкретного примера.

Классификация методов. Наибольшее распространение получили следующие методы: симплекс, последовательного улучшения плана, последовательного уточнения оценок, ветвей и границ, Гомори и др. Рассмотрим один из методов получивших широкое распространение.

Симплекс-метод. В основе этого метода лежит возможность представления точек многогранного множества через экстремальные точки (вершины) и

экстремальные направления (грани). Представляет систематическую процедуру решения, состоящую в движении от одной экстремальной точки \bar{X} допустимой области к другой с лучшими значениями целевой функции с помощью ведущего преобразования. Процесс продолжается до тех пор, пока не будет найдена оптимальная экстремальная точка X^* , либо экспериментальное направление S^* для которого $c^T S^* < 0$.

Экстремальные точки. Преобразуем матрицу A так, что $A = [B, N]$, где B – матрица порядка $m = (n - m)$. Компоненты векторов X_B, X_N отвечают столбцам соответствующих матриц и называются базисными и небазисными. Тогда ограничения $AX = b, X \geq 0$ можно преобразовать следующим образом

$$BX_B + NX_N = b, X_B \geq 0, X_N \geq 0. \quad (4.5)$$

Существует теорема [2], в которой доказано, что точка \bar{X} является экстремальной точкой множества $S = \{X: AX = b, X \geq 0\}$, когда перестановкой столбцов матрица A может быть представлена в виде $[B, N]$ так, что

$$\bar{X} = \begin{bmatrix} X_B \\ X_N \end{bmatrix} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}, \quad (4.6)$$

где B – невырожденная матрица порядка $m \times m$, удовлетворяющая неравенству $B^{-1}b \geq 0$. Число экстремальных точек не превосходит

$$N = \frac{n!}{m!(n-m)!}.$$

Экстремальные направления. Вектор \bar{S} задает экстремальное направление множества S если матрица A перестановкой столбцов может быть представлена в виде $[B, N]$ так, что $B^{-1}a_j \leq 0$ для некоторого столбца a_j из N и \bar{S} положительное кратное вектора $S^T = \left[(-B^{-1}a_j)^T, e_j^T \right]$. Здесь e_j есть $(n - m)$ - Мерный вектор, j -я компонента которого равна 1, а остальные – нулю.

Анализ условий оптимальности. Рассматриваем задачу минимизации целевой функции. Возьмем экстремальнойю точку \bar{X} и сопоставим с произвольной точкой X , для которой используя (4.5) получим

$$X_B = B^{-1}b - B^{-1}NX_N. \quad (4.7)$$

Преобразуем уравнение для целевой функции

$$f(C, X) = C_B^T X_B + C_N^T X_N = C_B^T B^{-1}b + (C_N^T - C_B^T B^{-1}N)X_N. \quad (4.8)$$

В это уравнение подставим условие (4.6)

$$f(C, X) = C_B^T \bar{X} + (C_N^T - C_B^T B^{-1}N)X_N = C^T \bar{X} + dX_N. \quad (4.9)$$

Проведем анализ полученного уравнения. Если $d \geq 0$, то в силу условия $X_N \geq 0$ справедливо неравенство $C^T X \geq C^T \bar{X}$, следовательно \bar{X} – оптимальная экстремальная точка, которая является результатом решения.

При $d < 0$, в частности для некоторого j компонента $C_j - C_B^T B^{-1} a_j$, необходимо дальнейшее изменение $X = \bar{X} + \lambda S_j$.

Направление S_j определяется по формуле

$$S_j = \begin{bmatrix} -B^{-1} a_j \\ e_j \end{bmatrix},$$

где e_j – $(n - m)$ - Мерный единичный вектор с единицей j -й позиции.

Тогда из уравнения (4.9) следует, что

$$f(C, X) = C_B^T X + \lambda (C_j - C_B^T B^{-1} a_j). \quad (4.10)$$

При $\lambda > 0$ рассмотрим следующие два случая.

1. Если $B^{-1} a_j \leq 0$, то из (4.10) следует, что значения целевой функции неограниченны. В этом случае найдено экстремальное направление S_j .

2. При $B^{-1} a_j > 0$ необходим расчет координат новой экстремальной точки. Величина шага λ определяется следующим образом

$$\lambda = \min_{i \leq j \leq m} \left\{ \frac{\bar{b}_i}{y_{ij}} : y_{ij} > 0 \right\} = \frac{\bar{b}_r}{y_{rj}} \geq 0, \quad (4.11)$$

где $\bar{b} = B^{-1} b$; y_{ij} – j -я компонента вектора Y_j .

Тогда компоненты вектора рассчитываются по формулам

$$x_{bj} = b_j - \frac{\bar{b}_r}{y_{rj}}, \quad i = 1, \dots, m, \quad (4.12)$$

$$x_{bj} = \frac{\bar{b}_r}{y_{rj}}, \text{ а остальные равны нулю.}$$

Положительными могут быть только $x_{b1}, \dots, x_{br-1}, x_{br+1}, x_{bm}, x_j$, т.е. не более m компонент.

Базисная переменная x_{br} , соответствующая минимальному отношению в (4.11), выводится из базиса, а небазисная x_j вводится в базис на ее место.

Таким образом, если задана произвольная экстремальная точка, то можно установить, что она оптимальна, либо найти экстремальное направление, приводящее к неограниченному значению целевой функции, либо найти экс-

тремальную точку с лучшими значениями целевой функции. В последнем случае процесс повторяется.

Табличное представление симплекс-метода. Рассмотрим для пояснения процедуры преобразования задачи при переходе от одной экстремальной точки к другой. Преобразуем постановку задачи к виду:

- строка – целевая функция $f - C_B^T X_B - C_N^T X_N = 0$,

- строка – ограничений $BX_B + NX_N = b$.

Эти равенства можно свести в симплекс-таблицу, в которой ПЧ соответствует правым частям:

f	X_N^T	X_B^T	ПЧ
1	$-C_N^T$	$-C_B^T$	0
0	N	B	b

В соответствии с рассмотренными выше условиями преобразуем эту таблицу. Строки-ограничения умножим на B^{-2} . К строке целевой функции прибавим новые строки ограничения, умноженные на C_B^T . При этом получается следующая таблица:

	f	X_N^T	X_B^T	ПЧ
f	1	$C_B^T B^{-1} - C_N^T$	0^T	$C_B^T \bar{b}$
X_B	0	$B^{-1}N$	1	\bar{b}

Базисные переменные отмечены в таблице слева. Значения базисных переменных в целевой функции записаны в правой части таблицы. Кроме того, вектор $C_B^T B^{-1} - C_N^T$ и матрицу $B^{-1}N$ удобно хранить под внебазисными переменными.

Последняя таблица содержит всю информацию, необходимую для завершения первого шага симплекс-метода. Если $C_B^T B^{-1}N - C_N^T \leq 0$, то процесс прекращается – последняя (текущая) экстремальная точка является оптимальной. В противном случае при просмотре строки целевой функции можно отобрать внебазисную переменную с отрицательным значением $C_B^T B^{-1}a_j - C_j$. Если $B^{-1}a_j \leq 0$, то процесс прекращается – оптимальное значение целевой функции неограниченно. Предположим, что $Y_j = B^{-1}a_j \leq 0$. Так как \bar{b} и Y_j записаны под ПЧ и x_j соответственно, то, следуя (4.11), по таблице легко вычислить λ . Базисная переменная x_{br} , соответствующая минимальному отношению в (4.11), выводится из базиса, а x_j вводится в базис.

Ведущее преобразование симплекс-таблицы осуществляется при переходе от одной экстремальной точки к другой. Используя условие (4.11) определяется ведущий элемент y_{rj} . Строка отвечающая выводимой из базиса переменной x_{br} , называется ведущей строкой, а столбец, отвечающей вводимой в базис x_j – ведущим столбцом. Основные этапы ведущего преобразования заключаются в следующем:

1. Преобразуется ведущая строка (r – строка, соответствующая x_{br}) путем деления на ведущий элемент y_{rj} .

2. Производится умножение полученной r – строки на y_{ij} и результат вычитается из i -ой строки ограниченной для всех $i = 1, \dots, m, i \neq r$.

3. Осуществляется умножение полученной r -строки на $C_B^T B^{-1} a_j - C_j$ и результат вычитается из строки целевой функции.

Алгоритм симплекс-метода рассматривается для задачи минимизации.

Начальный этап. Найти начальную экстремальную точку \bar{X} с базисом B . Если это сделать трудно, то использовать метод искусственных переменных, который рассматривается в работах [6–21].

Основной этап. Шаг 1. Пусть \bar{X} – экстремальная точка с базисом B . Вычислить $C_B^T B^{-1} N - C_N^T$. Если этот вектор неположительный, то остановиться; \bar{X} – оптимальная экстремальная точка.

В противном случае выбрать наибольшую положительную компоненту $C_B^T B^{-1} a_j - C_j$. Если $Y_j = B^{-1} a_j \leq 0$, то остановиться: оптимальное значение целевой функции неограниченно вдоль луча.

$$\left\{ X + \lambda \begin{pmatrix} -Y_j \\ e_j \end{pmatrix} : \lambda \geq 0 \right\},$$

где e_j – вектор, все компоненты которого, кроме j -й, равны нулю, а j -я – единица.

Если $Y_j > 0$, то перейти к шагу 2.

Шаг 2. Вычислить номер r в соответствии с (4.11) и построить новую экстремальную точку по формулам (4.12). Сформировать новый базис, выходя из B столбец a_{br} и вводя вместо него a_j . Повторить шаг 1.

Алгоритм обеспечивает сходимость за конечное число шагов. Это объясняется тем, что если на каждой итерации, т. е. при прохождении всего основного этапа, $\bar{b} = B^{-1} b > 0$, то λ , определяемое по формуле (4.11) будет строго положительным и значение целевой функции в текущей экстремальной точке будет строго меньше, чем во всех экстремальных точках на предыдущих итерациях. Отсюда следует, что текущая точка отличается от всех, построенных на предыдущих итерациях. Так как число экстремальных точек конечно, то симплексный алгоритм останавливается через конечное число шагов.

Работу симплекс-метода поясним на следующем примере.

Пример 4.1. Симплекс-метод линейного программирования

Минимизировать $f(C, X) = x_1 - 2x_2$ при ограничениях

$$4x_1 + 5x_2 \leq 40;$$

$$-x_1 + x_2 \leq 4;$$

$$x_2 \leq 5;$$

$$x_1, x_2 \geq 0.$$

Геометрическая иллюстрация этой задачи показана на рис. 4.1, где изображена линия для $f(C, X) = 3$ с направлением, в котором целевая функция убывает. Ограничения образуют замкнутый многогранник, значения целевой функ-

ции в вершинах которого соответственно равны 0, -8, -9, -6, 10. Координаты вершин указаны на рисунке 4.1. Видно, что если параллельно перемещать прямую $f(C, X)$, (т. е. уменьшать ее значения), то последней точкой, которая будет принадлежать допустимой области, является $\bar{X}^{(3)}$, что и будет решением задачи.

Исходную задачу путем введения дополнительных переменных x_3, x_4, x_5 преобразуем к каноническому виду:

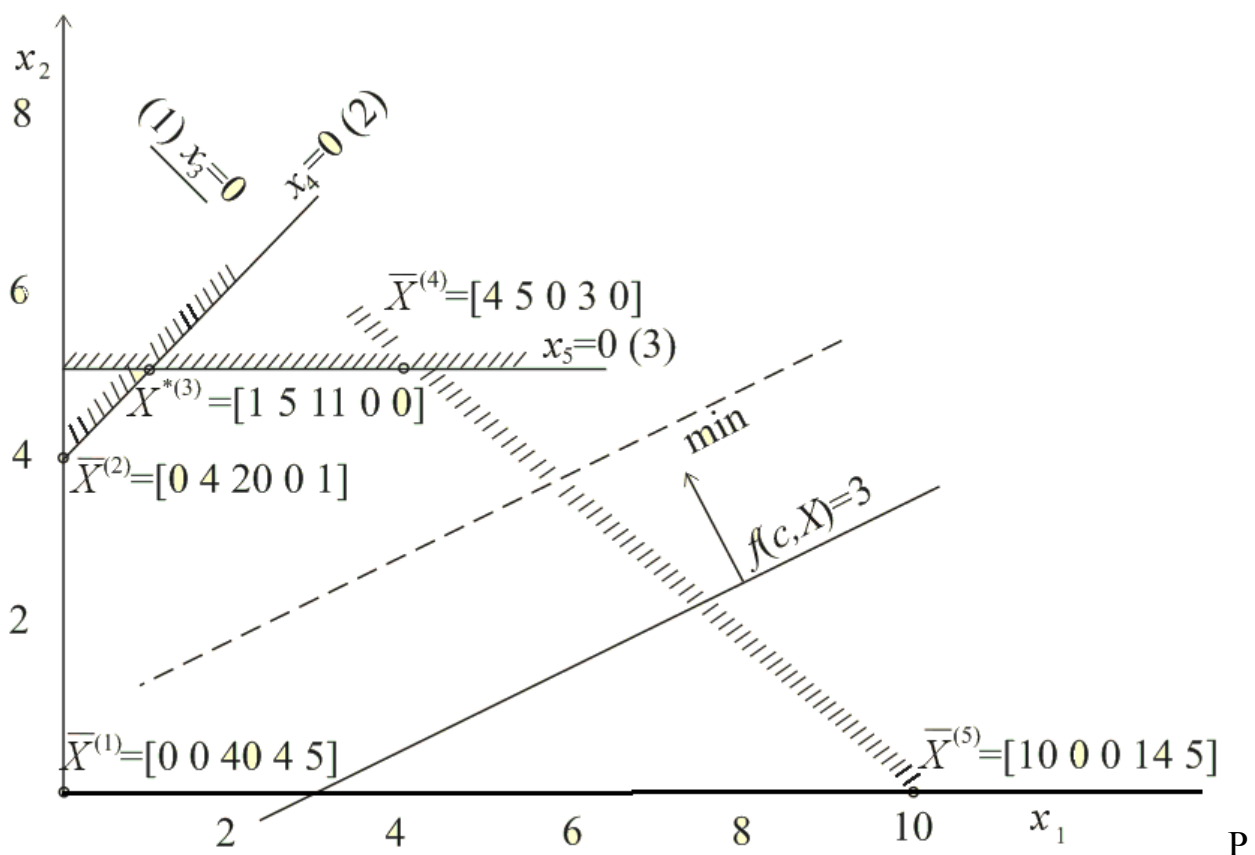
Минимизировать $f(C, X) = x_1 - 2x_2$ при ограничениях

$$4x_1 + 5x_2 + x_3 = 40;$$

$$-x_1 + x_2 + x_4 = 4;$$

$$x_2 + x_5 = 5;$$

$$x_1, x_2, x_3, x_4, x_5 \geq 0.$$



ис. 4.1. Симплекс-метод линейного программирования

Здесь $C = [1 \ -2 \ 0 \ 0 \ 0]^T$, $b = [40 \ 4 \ 5]^T$, $A = \begin{bmatrix} 4 & 5 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}$. Выберем в

качестве B матрицу $[a_3 \ a_4 \ a_5] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Так как $B^{-1}b = \bar{b} > 0$, то найдена

начальная экстремальная точка, которой является вершина $\bar{X}^{(1)}$.

Соответствующая ей симплекс-таблица:

	f	x_1	x_2	x_3	x_4	x_5	ПЧ
f	1	-1	2	0	0	0	0
x_3	0	4	5	1	0	0	40
x_4	0	-1	①	0	1	0	4
x_5	0	0	1	0	0	1	5

Анализируем условие оптимальности $C_B^T B^{-1}N - C_N^T < 0$, компоненты вектора равны [1-2], поэтому переходим к новой экстремальной точке. Находим ведущий элемент, используя (4.11). Минимальное отношение $b_j/y_{ij} = 4$ для положительных y_{ij} соответствует y_{23} . Следовательно x_4 выводится из базиса, а x_2 вводится. Новый базис $B = [a_3 \ a_2 \ a_5]$. Строим новую симплекс-таблицу.

	f	x_1	x_2	x_3	x_4	x_5	ПЧ
f	1	1	0	0	-2	0	-8
x_3	0	9	0	1	-5	0	20
$*x_2$	0	-1	1	0	1	0	4
x_5	0	①	0	0	-1	1	1

Преобразованная ведущая строка $*$ для базисной переменной x_2 получается путем деления исходной на ведущий элемент $y_{22} = 1$. Определяем значение $C_B^T B^{-1}a_2 - C_3$. Оно равно 2, что $> C$. Преобразуем строку для целевой функции, а именно умножим преобразованную строку на 2 и результат вычитаем из строки целевой функции.

При преобразовании первой и третьей строк-ограничений новую ведущую строку умножаем соответственно на $y_{12} = 5$ и $y_{32} = 1$ и результат вычитаем из строк-ограничений. Полученная таблица характеризует экстремальную точку $\bar{X}^{(2)} = [0 \ 4 \ 20 \ 2]$. Здесь также проверяем условие оптимальности. Имеем один положительный компонент для x_1 , поэтому процесс преобразования продолжаем.

Строим новую симплекс-таблицу. Ведущий элемент $y_{13} = 1$, следовательно x_5 выводится из базиса, а x_1 вводится. Новый базис $B = [a_3 \ a_2 \ a_1]$. Получаем следующую таблицу.

	f	x_1	x_2	x_3	x_4	x_5	ПЧ
f	1	0	0	0	-1	-1	-9
x_3	0	0	0	1	4	-9	11
x_2	0	0	1	0	0	-1	5
x_1	0	1	0	0	-1	1	1

Так как обе компоненты вектора $(C_B^T B^{-1}N - C_N^T) = [-1 \ -1] < 0$, то получено оптимальное решение $\bar{X}^* = [1 \ 5 \ 11 \ 0 \ 0]$, $f(\bar{X}^*) = -9$.

Рассмотренные выше аспекты симплекс-Метода характерны для других модификаций симплексных процедур. Особенности состоят в способах перехода от одной экстремальной точки к другой. Так в методах последовательного

улучшения плана и уточнения оценок используются модифицированные жордановы исключения [21].

4.2. Нелинейное программирование

Постановка задачи. Нелинейное программирование состоит в оптимизации функции произвольного вида при ограничениях в виде равенств и неравенств. Формальная постановка имеет вид:

Геометрическая интерпретация задачи: Минимизировать $f(X)$, $X \in E^n$ при ограничениях:

$$h_i(X) = 0, i = 1, \dots, \quad (4.13)$$

$$g_i(X) \leq 0, i = m + 1, \dots, p.$$

Вид задачи показан на рис. 4.2. Здесь решением условной задачи является \bar{X}^* , а без условной X^* .

Условия оптимальности. Структура задачи нелинейного программирования такова, что универсальный критерий оптимальности еще не разработан. Доказаны необходимые и достаточные условия оптимальности лишь для частных случаев, которые используют аналитическое описание целевой функции и ограничений. В таких задачах можно применять условия Ф. Джона и Куна-Таккера [2]. Поэтому при численном решении необходима многократная проверка сходимости при различных начальных условиях и параметрах используемых методов, а также физическая интерпретация полученных результатов.

Классификация методов. Для решения задач нелинейного программирования используются следующие методы: множителей Лагранжа, линейной аппроксимации, штрафных функций, барьеров, последовательной безусловной минимизации, возможных направлений, скользящего допуска и др. [1, 2, 17, 18, 25]. Рассмотрим предпосылки построения и особенности некоторых методов.

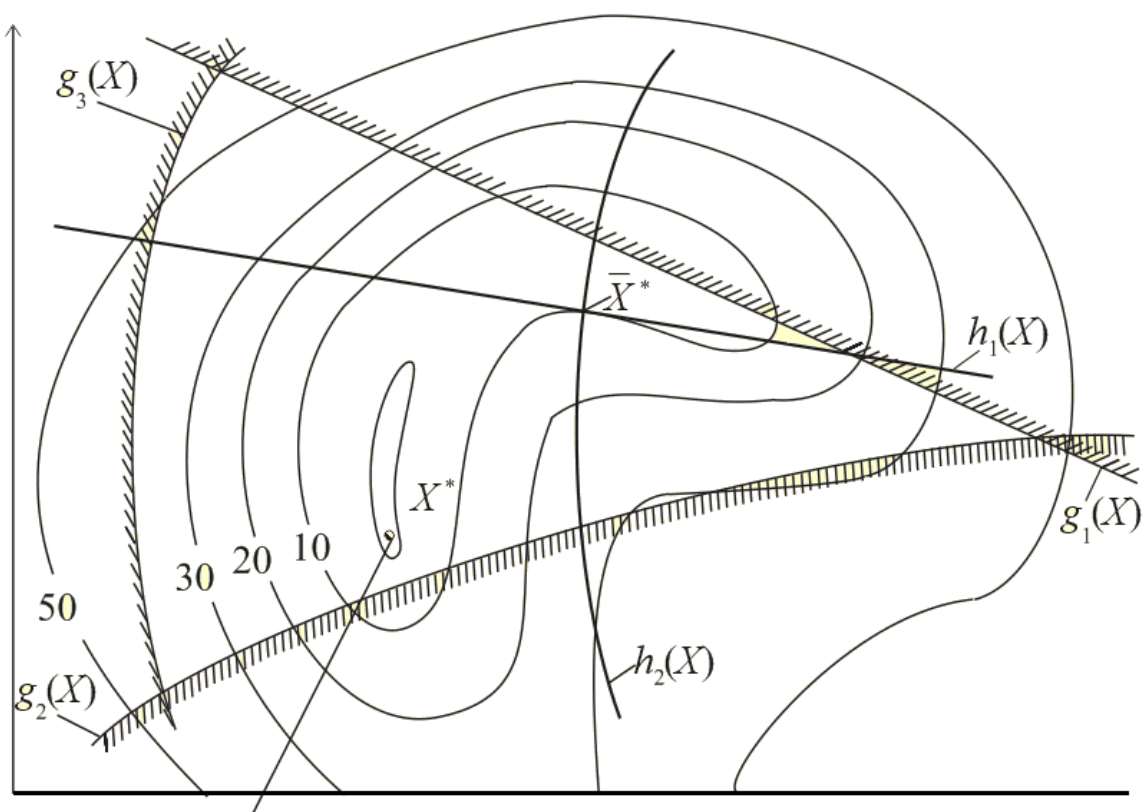


Рис. 4.2. Геометрическая интерпретация задачи нелинейного программирования

Методы линейной аппроксимации. В этих методах на каждой k -ой итерации функция и ограничения линейризуются. В результате вместо (4.13) получается задача минимизации $f(X^{(k)})$ при ограничениях

$$h_i(X^{(k)}) + \nabla^T h_i(X^{(k)})(X - X^{(k)}) = 0, i = 1, \dots, m;$$

$$g_i(X^{(k)}) + \nabla^T g_i(X^{(k)})(X - X^{(k)}) \leq 0, i = m + 1, \dots, p. \quad (4.14)$$

Поскольку $f(X^{(k)})$, $\nabla f(X^{(k)})$, $h_i(X^{(k)})$, $\nabla h_i(X^{(k)})$, $g_i(X^{(k)})$, являются постоянными векторами, либо скалярными константами (значения вычисляются в точке $X^{(k)}$), задача (4.14) представляет задачу линейного программирования. Реализация пошаговых процедур задач (4.14) обеспечивает сходимость к решению задачи (4.13) нелинейного программирования [1]. К методам линейной аппроксимации относятся: аппроксимирующее линейное программирование, проективные методы, метод приведенного градиента и др. [1, 3].

Метод множителей Лагранжа. Заключается в постановке и решении модифицированной задачи, которая преобразуется из исходной (4.13) к задаче безусловной оптимизации (4.15).

Сначала все ограничения в виде неравенств заменяются эквивалентными ограничениями типа равенств путем введения ослабляющих переменных $g_i(X) - U_i^2 = 0, i = m + 1, \dots, p$. Затем формулируется задача безусловной оптимизации

$$L(X, \lambda) = f(X) + \sum_{i=1}^m \lambda_i h_i(X) + \sum_{i=m+1}^p \lambda_i [g_i(X) - U_i^2]. \quad (4.15)$$

Функцию $L(X, \lambda)$ называют функцией Лагранжа, а величина λ_i – множителями Лагранжа. Решение сводится к нахождению стационарных точек (4.15) путем решения системы уравнений, полученных приравниванием нулю частных производных по λ, X, U [1, 3, 21].

Методы штрафных и барьерных функций основаны на преобразовании исходной задачи (4.13) к задаче безусловной оптимизации. С помощью функций, задающих ограничения, формируется штраф, который действует таким образом, что нарушение какого-либо ограничения становится невыгодным для преобразованной задачи.

В методе штрафных функций преобразованная задача имеет вид

$$F(X, \mu) = f(X) + \mu \left\{ \sum_{i=1}^m |h_i(X)|^l + \sum_{i=m+1}^p [\max(0, g_i(X))]^l \right\}, \quad (4.16)$$

где μ – положительное большое число, коэффициент штрафа;

l – целое положительное число.

Метод барьерных функций приемлем для задачи с ограничениями типа неравенств. Здесь задача преобразуется следующим образом [1]:

$$F(X, r) = f(X) + r \sum_{i=m+1}^p \frac{-1}{g_i(X)}, \quad (4.17)$$

где r – положительная константа.

При приближении к границе изнутри, $\{X | g(X) \leq 0\}$ добавка к $f(X)$ становится большой, что невыгодно для задачи (4.17). Решение задачи (4.17, 4.16) может быть осуществлено любым методом безусловной оптимизации. При этом следует помнить, что точность решения зависит от правильности выбора параметров «штрафа» μ и r в процессе поиска.

Так, например, в методе штрафных функций при очень больших μ возникает «овражность», а малые значения не дают результата.

Метод последовательно безусловной минимизации является обобщением метода барьерных и штрафных функций [1, 8], сводится к решению последовательности задач без ограничений, которая в пределе сходится к задаче нелинейного программирования.

На каждом этапе используется функция

$$P(X^{(k)}, r^{(k)}) = f(X^{(k)}) + (r^{(k)})^{-1/2} \sum_{i=1}^m h_i^2 X^{(k)} + r^{(k)} \sum_{i=m+1}^p \frac{-1}{g_i(X^{(k)})}, \quad (4.18)$$

где значения весовых коэффициентов r положительны и образуют монотонно убывающую последовательность $\{r \mid r^{(1)} > r^{(2)} > \dots > 0\}$. Для задания этой последовательности используются различные подходы [1, 8]. В качестве простого соотношения можно использовать зависимость $r^{(k)} = r^{(k-1)}$, где $r^{(1)} = 2$.

Метод скользящего допуска основан на совместном использовании информации, получаемой как в допустимой области, так и в точках являющихся близкими к допустимым (почти допустимые). Интервалы, в пределах которых точки можно считать почти допустимыми, в ходе оптимизационного поиска постепенно сокращаются. Это делается для того, чтобы сразу исключить возможность возникновения «оврага». С этой целью вводится критерий скользящего допуска $\Phi^{(k)}$, который при приближении к точке оптимума исключает возможность нарушения допустимых границ.

Задача нелинейного программирования (4.13) сводится к задаче с единственным укрупненным ограничением, которое требуется.

Минимизировать $f(X)$, $X \in E^n$
при ограничении $\Phi^{(k)} = T(X) \geq 0$

$$T(X) = \left[\sum_{i=1}^m h_i^2(X) + \sum_{i=m+1}^p U_i(X) g_i^2(X) \right]^{\frac{1}{2}}, \quad (4.19)$$

где U_i – оператор Хевисайда, обладающий свойством, обладающий свойством: $U_i = 0$ при $g_i(X) \leq 0$ и $U_i = 1$ при $g_i(X) > 0$.

Произвольная точка $X^{(k)}$ является:

- 1) допустимой, если $T(X^{(k)}) = 0$;
- 2) почти допустимой, при $0 \leq T(X^{(k)}) \leq \Phi^{(k)}$;
- 3) недопустимой, если $T(X^{(k)}) > \Phi^{(k)}$.

Задачу (4.19) целесообразно решать одним из методов поиска. При этом задается своя стратегия выбора $\Phi^{(k)}$ в виде убывающей функции, примеры построения которой рассмотрены в [1].

По оценкам, приведенным в работе [1] данный метод является одним из наиболее эффективных.

4.3. Практическое приложение задач оптимизации

В заключении кратко остановимся на практическом приложении описанных выше методов оптимизации. Рассмотрим содержательные формулиров-

ки наиболее часто встречающихся задач оптимизации. При этом сохраним принятые обозначения и терминологию.

Задачи идентификации. Очень часто необходимость оптимизации возникает при подстройке параметров модели по экспериментальным данным. Задачи идентификации отличаются большим разнообразием в зависимости от вида модели (статическая, динамическая, линейная, нелинейная) и данных (прямые, косвенные и др.)

При настройке по «пассивным» данным параметров модели заданной структуры $Y=F(U,X)$ требуется подобрать параметры X^* так, чтобы минимизировать отклонения предсказываемых значений выхода от фактических, т. е. найти

$$\min f(X) = \sum_{i=1}^m \varphi [y_i - F(U_i, X)], \quad (4.20)$$

где $\varphi(Z)$ – мера рассогласования ($\varphi(Z) = |Z|$; $\varphi(Z) = Z^2$ и др.);

U_i, y_i – результат измерений.

Эта задача безусловной оптимизации с неявно заданным критерием, которая для простейших структур может быть решена аналитически. В других случаях возникает необходимость использования методов оптимизации.

В теории **планирования эксперимента** возникают задачи оптимизации трех типов. Первый связан с выбором оптимального плана (A -оптимальные, D -оптимальные и др. планы). Здесь характерны детерминированные многоэкстремальные задачи при наличии ограничений на входные воздействия. Второй тип задач связан с отработкой результатов эксперимента при фиксированном плане. Наконец, во многих случаях целью исследования является не локальное описание зависимости некоторого показателя от других факторов, а нахождение экстремума этого показателя.

Оптимальное проектирование осуществляется после разработки математической модели проектируемого объекта. Предполагается, что задан некоторый показатель (целевая функция), который подлежит оптимизации, и сформулированы ограничения на допустимые значения параметров и других характеристик объекта. Здесь возможны задачи, связанные с наилучшим выбором параметров отдельного изделия (элементы конструкции агрегата, химический состав и свойства металла), технологического процесса (оптимизация режимов, создание новых процессов), предприятия (оптимальное размещение и выбор совокупности агрегатов, цехов, образующих технологическое производство). При этом типичен случай нелинейной целевой функции и нелинейных ограничений, а также необходимость выбора нескольких альтернативных вариантов.

Оптимальное планирование. Исторически именно экономические задачи привели к возникновению и развитию линейного программирования. Это

объясняется тем, что в экономике проблема планирования является центральной, а большинство экономических моделей являются линейными. К линейному программированию сводятся многие задачи оперативного и долгосрочного планирования, задачи оптимизации снабжения и перевозок, задачи управления запасами, рационального выбора шихтовых материалов, составления расписаний и др.

Экстремальное регулирование заключается в выборе входных воздействий на объект, дающих экстремум показателя качества. Эта проблема с математической точки зрения рассматривается как задача минимизации функции, единственной информацией о которой является ее значение, вычисленное со случайными ошибками.

Процесс оптимизации производится на реальном объекте, поэтому такие задачи имеют свои особенности. Нельзя делать большие шаги, необходимо учитывать нестационарность и инерционность объекта, запаздывание в каналах измерения и воздействия. Это приводит к тому, что применяемые на практике способы экстремального регулирования заметно отличаются от рассмотренных стандартных алгоритмов безусловной оптимизации.

Оптимальное управление. В этих задачах требуется найти закон изменения управляющих воздействий для динамической системы так, чтобы минимизировать критерий оптимальности. При этом предполагается, что система описывается дифференциальными уравнениями. Результатом решения таких задач является выбор режима управления технологическим процессом, электрической машиной и другими объектами. Конечномерный вариант таких задач может быть решен методами нелинейного программирования. Условия оптимальности и другие методы решения подробно рассматриваются в литературе по теории оптимального управления.

4.4. Практические методы инвариантной оптимизации

1. Процесс оптимизации. Имеется задача. Для решения задачи нужно формализовать объект и представить его в виде математической модели.

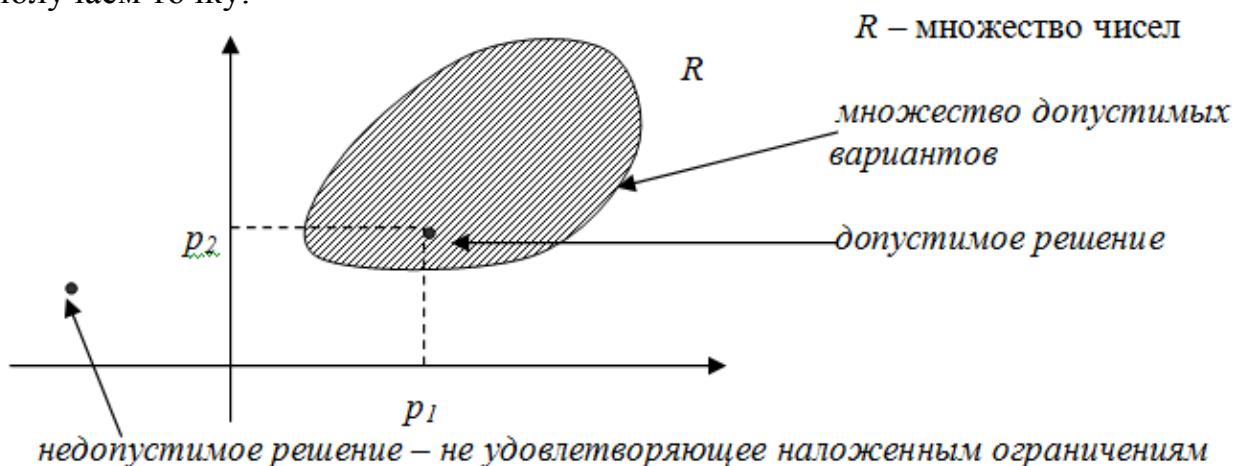
Модели могут быть:

- физические;
- геометрические (схема, фотография, рисунок);
- математические.

Математическая модель, такая, которая определена с помощью математических формализмов. Математическая модель не является точной, а является идеализацией. Модель характеризуется параметрами, которые могут быть и числовыми p_1, p_2, \dots, p_n . Их часть может характеризовать состояние объекта – параметры состояния (p_1, p_2) , а другие могут относиться к процессу проектирования – переменные проектирования (p_n) .

Определение параметров состояния – задача моделирования. Определение переменных проектирования – задачи проектирования или задачи оптимизации.

Допустим, имеются 2 переменные p_1, p_2 . Задавая конкретные значения, получаем точку.



Плоскость множества возможных вариантов, на нее могут быть наложены ограничения.

Отображение множества $F : G \rightarrow R$ – целевая функция, которая позволяет формировать критерий для сравнения различных решений.

2 вида задач оптимизации:

- максимизации;
- минимизации.

Для оптимизационного решения задачи требуется:

1. Сформулировать задачу.
2. Построить математическую модель (определить множество переменных).
3. Определить ограничения на возможные решения.
4. Определить целевую функцию.

Далее применим формальные математические методы, позволяющие найти решения.

2. Методы одномерной оптимизации. Постановка задачи: требуется оптимизировать x (формальная постановка)

$$\begin{cases} Z = f(x) \longrightarrow opt \\ a \leq x \leq b \end{cases}$$

$f(x)$ – функция одной переменной $a, b, x, f(x) \in R$

$f(x)$ – целевая функция.

Решение: найти x , при котором $f(x)$ принимает оптимальное значение.

2 варианта:

- минимизировать – задача минимизации;
- максимизировать – задача максимизации.

Рассмотрим случай минимизации

$$\begin{cases} Z = f(x) \longrightarrow opt \\ a \leq x \leq b \end{cases} \quad a, b, x, f(x) \in R.$$

2 способа:

- аналитический;
- численный.

В аналитическом способе $f(x)$ задается в виде формулы, в численном способе $f(x)$ задается в виде черного ящика, на входе подается x , на выходе – значение целевой функции в этой точке.

Пусть функция определена в некоторой области S ($x \in S$), в случае одномерной оптимизации S – интервал $S = \{x \mid a \leq x \leq b\}$:

1. точка x^* называется глобальным минимумом, если для $\forall x \in S, f(x^*) \leq f(x)$.

2. точка x^* называется строгим глобальным минимумом, если для $\forall x \in S, f(x^*) < f(x)$.

3. точка x^* называется локальным минимумом, если для $\forall x \in E(x^*), f(x^*) \leq f(x)$.

4. точка x^* называется строгим локальным минимумом, если для $\forall x \in E(x^*), f(x^*) < f(x)$.

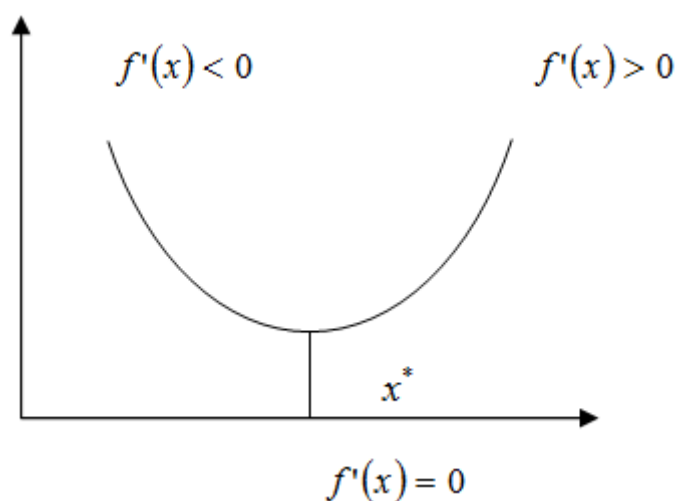
Следствие: любая точка глобального минимума является локальным минимумом, обратное – не верно.

3. Аналитический способ нахождения локального минимума. Задана функция, которая:

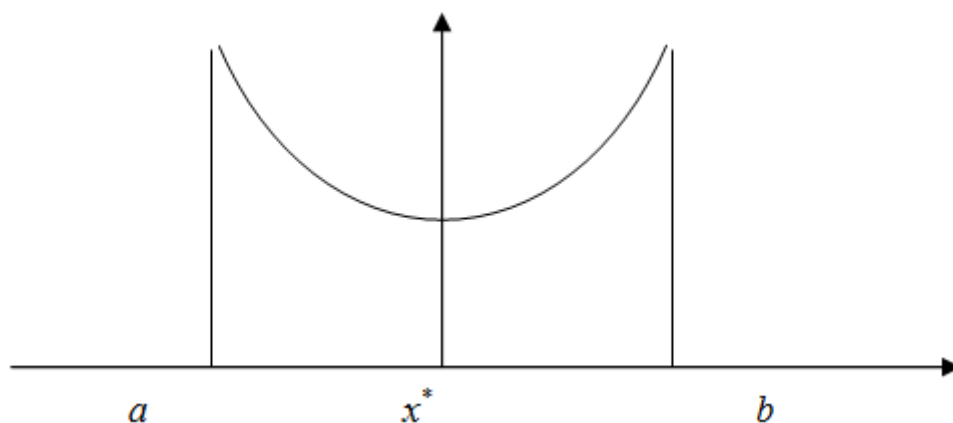
- $f(x)$ – дифференцируема;

- $f'(x) = 0$ – необходимое условие точки локального минимума.

Графическая интерпретация решения

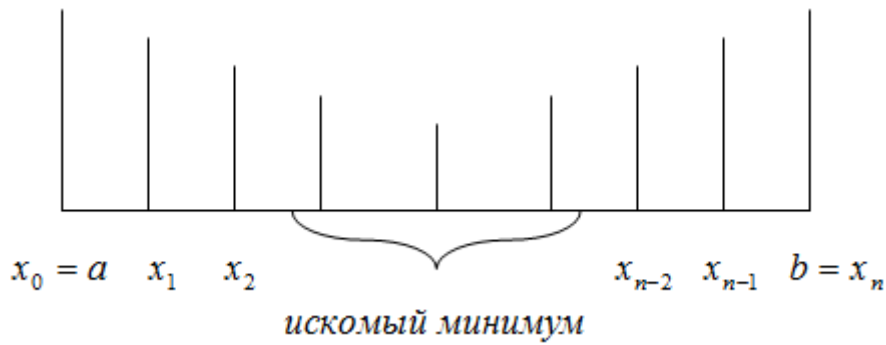


4. Численные методы. Пусть функция $f(x)$ задана на интервале (a, b) , при этом существует такая точка x^* , которая на $[a, x^*]$ – монотонно убывает, а на $[x^*, b]$ – монотонно возрастает, следовательно, – функция *унимодальная* (одна мода, один экстремум).



Если из того, что $x_1 \leq x_2$ следует, что $f(x_1) \leq f(x_2)$, то функция называется *монотонно возрастающей*. Если из того, что $x_2 \leq x_1$ следует, что $f(x_2) \leq f(x_1)$, то функция называется *монотонно убывающей*.

5. Методы одномерного поиска. Разобьем $[a, b]$ и вычислим значение функции в каждой точке.



В результате остается интервал меньшего размера, к которому применяется тот же метод; находим еще один интервал, и т. д.; в конце находим интервал с заведомо нужной точкой.

Интервал неопределенности – интервал, в котором заведомо находится точка минимума. Наиболее эффективное разбиение – двумя точками на 3 равных отрезка.

- 1) $f(x_1) < f(x_2) \Rightarrow (a, x_2)$;
- 2) $f(x_1) > f(x_2) \Rightarrow (x_1, b)$;

$\left(\frac{2}{3}\right)^n$ – после выполнения n шагов происходит сокращение исходного интервала;

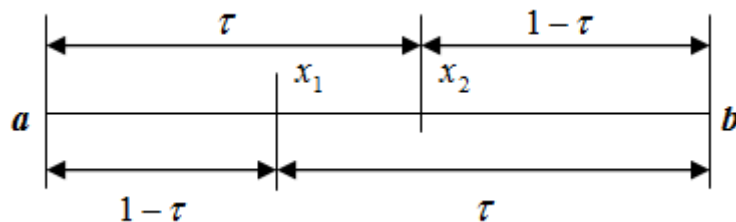
$$\left(\frac{2}{3}\right)^n (b - a) = \delta,$$

где δ – точность, с которой надо найти решение задачи.

$$\left(\frac{2}{3}\right)^n = \frac{\delta}{b - a} \quad n \ln \frac{2}{3} = \ln \frac{\delta}{b - a},$$

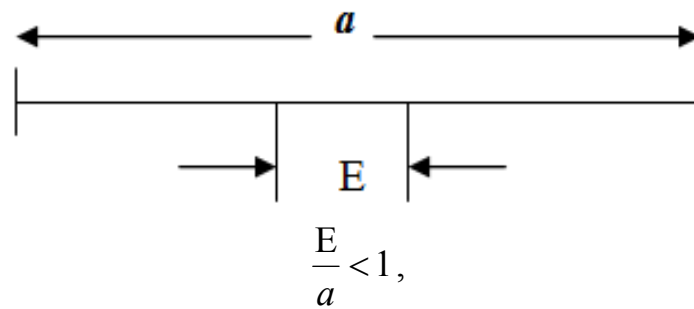
где $N = 2n$; n – число шагов, N – число вычислений (мера эффективности данного решения).

6. Метод золотого сечения. Точки должны быть расположены на равном расстоянии.



$$\frac{\tau}{1} = \frac{1 - \tau}{\tau}; \quad \tau^2 = 1 - \tau; \quad \tau^2 + \tau - 1 = 0;$$

$$\tau_{1,2} = \frac{-1 \pm \sqrt{5}}{2} \Rightarrow \tau = \frac{\sqrt{5} - 1}{2} \approx 0,618; \quad \tau - \text{золотое сечение.}$$

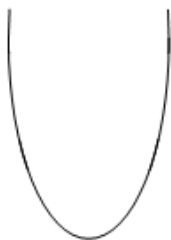


λ - величина сокращения на каждом шаге.
Число итераций растет как логарифм функции

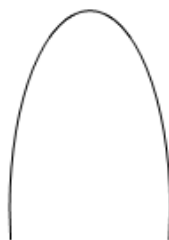
$$(\lambda)^N = \frac{E}{a}; N = \ln \frac{E}{a\lambda}.$$

7. Одномерная оптимизация с использованием производных
 $f(x) \rightarrow \min$. Пусть целевая функция дифференцируема $f'(x) = 0$.

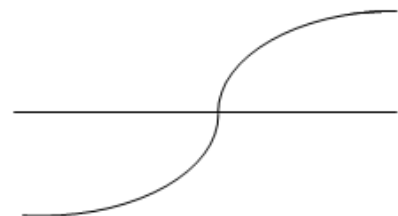
Графическая интерпретация



точка локального
минимума

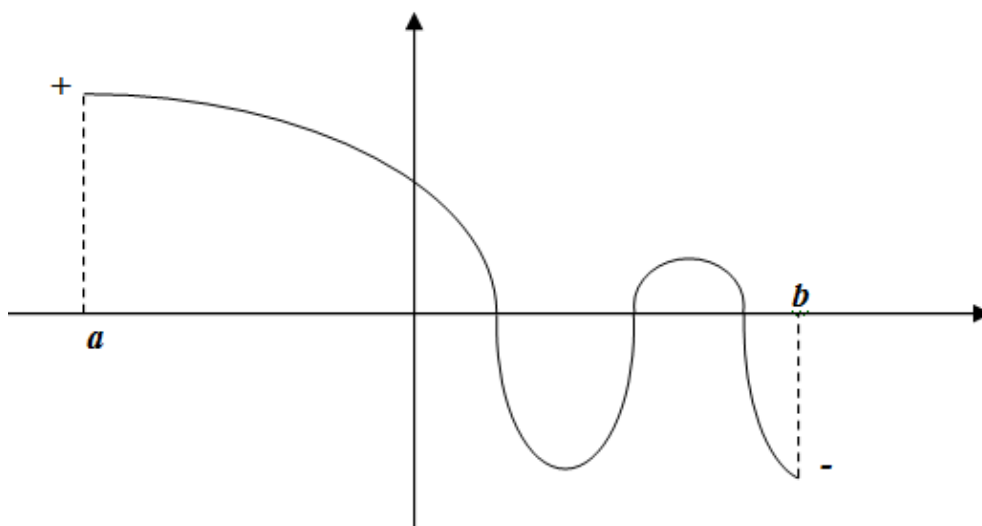


точка локального
максимума

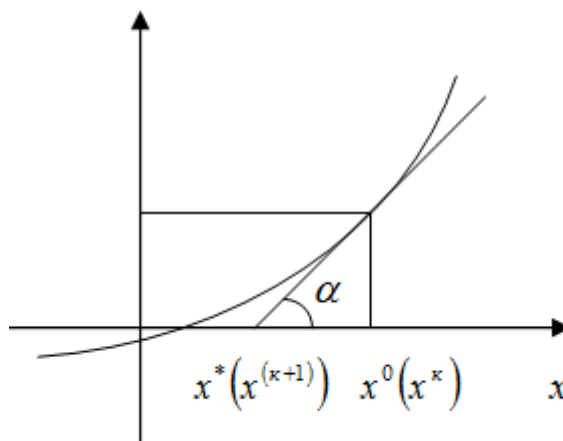


точка перегиба

8. Методы для нахождения корня уравнения функции одной переменной. Деление пополам. Имеется, по крайней мере, один корень. Выбираем любую точку и смотрим, какой знак она имеет, такой знак нам и искать. Выбираем точку приблизительно в середине интервала, исследуя значения, можно отбросить половину интервала.



Метод Ньютона (метод касательной). В случае если известна производная, то выбираем x^0 – начальное приближение.



Допустим, что точка x^0 достаточно близка к корню функции и примерно ведет себя линейно, не отклоняется. Проведем касательную и находим точку ближе чем x^0 , последовательно повторяем до x^* .

Для метода Ньютона необходимо, чтобы:

- функция имела производную;
- точка должна быть взята близко к корню;
- функция изменяется близко к линейной функции.

$$\Delta x^{(k)} = x^{(k+1)} - x^k;$$

$$\underbrace{f(x^k + \Delta x^k)}_0 = f(x^k) + f'(x^k) \Delta x^k + O(\dots) \text{ – уравнение касательной};$$

$$\begin{cases} \Delta x^k = -\frac{h(x^k)}{h'(x^k)} \\ x^{k+1} = x^k + \Delta x^k \end{cases}$$

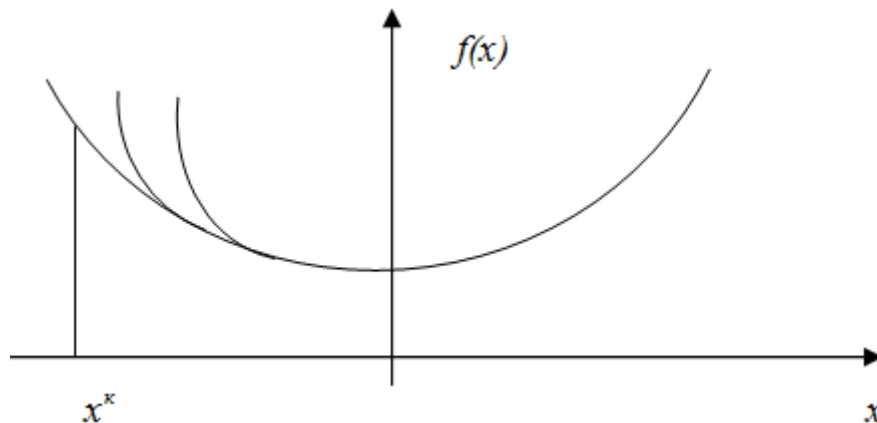
Если $|\Delta x^k| < \epsilon$, то вычисления можно прекратить и считать, что нужный нам корень – условие прекращения поиска. (ϵ – значение корня с некоторой точностью).

В методе Ньютона каждая его итерация удваивает количество значащих цифр. Если все условия выполнены, то эти методы удваивают (ускоряют) количество значащих цифр:

$$\begin{cases} h(x) = f'(x); \\ \Delta x^k = -\frac{h(x^k)}{h'(x^k)}; \\ x^{k+1} = x^k + \Delta x^k \end{cases}$$

Представим что $h(x)$ линейная функция, то метод Ньютона позволяет найти ее корень за одну итерацию. Целевая функция представляет собой квадратичную зависимость, следовательно, метод Ньютона позволяет найти минимум или максимум квадратичной функции за одну итерацию.

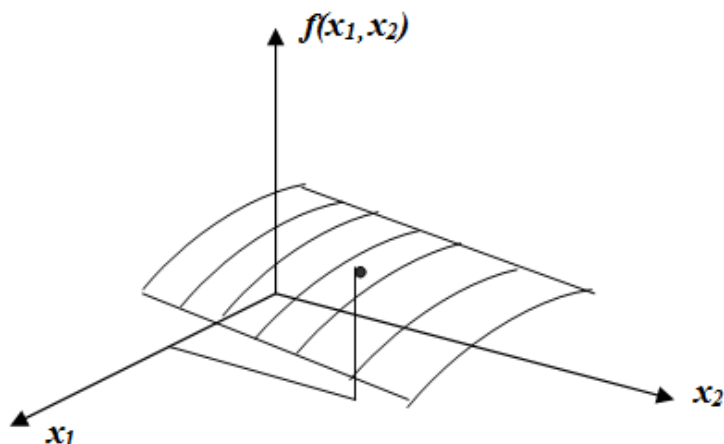
Замена функции на касательную, называется линейной аппроксимацией, и ее применение к целевой функции – парабола в точке приближения.



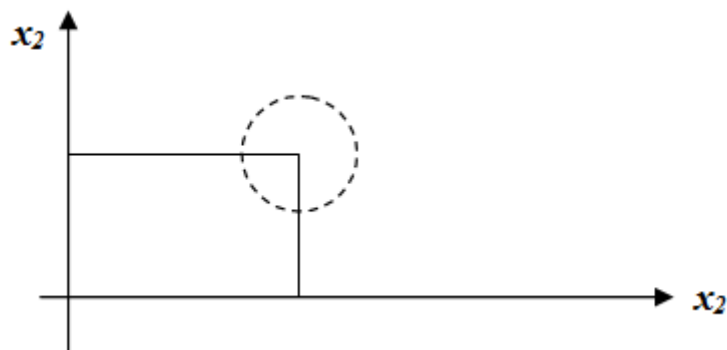
Замена заданной зависимости квадратичной зависимостью, называется – квадратичной аппроксимацией. Метод Ньютона основан на замене заданной зависимости более простой зависимостью.

9. Безусловная оптимизация. Целевая функция зависит от нескольких переменных $f(x_1, x_2, \dots, x_n) \rightarrow \min$. Т.к. нет дополнительных условий, накладывающихся на переменные, то – безусловная оптимизация.

Функции 2-х переменных



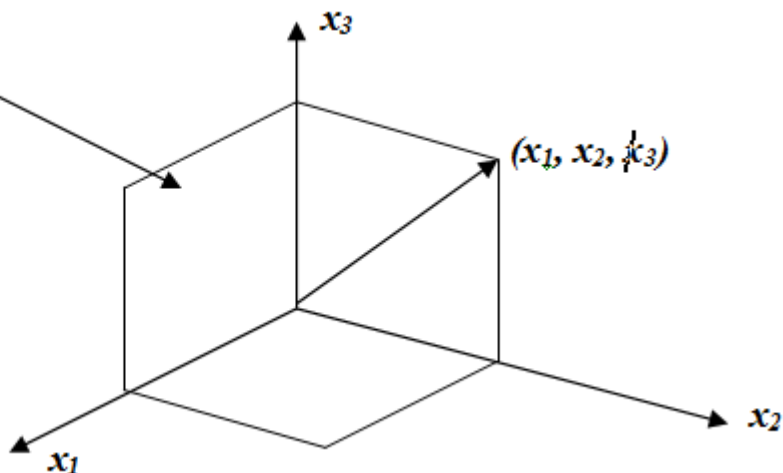
Условия, определяющие точку минимума, – необходимо проанализировать поведение функции в некоторой точке. Часто под окрестностью подразумевают шар.



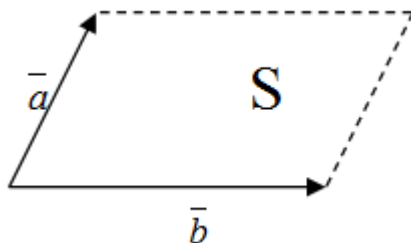
Рассмотрим вспомогательное построение:

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2) = f(x_1, x_2) + \left[\frac{\partial f}{\partial x_1} \Delta x_1 + \frac{\partial f}{\partial x_2} \Delta x_2 \right] + O(\dots).$$

линейное векторное пространство



Скалярное произведение векторов $(\bar{a}, \bar{b}) = \|\bar{a}\| \|\bar{b}\| \cos \bar{a} \wedge \bar{b}$, где $\|\cdot\|$ - длина вектора (норма вектора), $\bar{a} \wedge \bar{b}$ - угол между векторами.

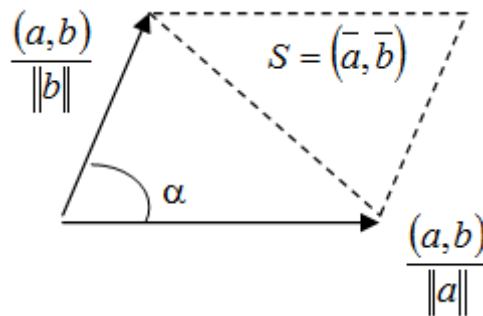


$$\|a\| = (a, a)^{1/2}.$$

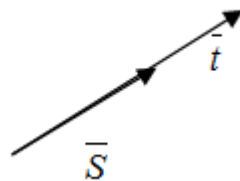
Допустим, что: $a = (a_1, a_2, \dots, a_n)$, $b = (b_1, b_2, \dots, b_n)$.

Тогда: $(a, b) = \sum_{i=1}^n a_i b_i$; $\|a\| = \sqrt{\sum_{i=1}^n a_i^2}$.

Допустим, что имеется 2 вектора:

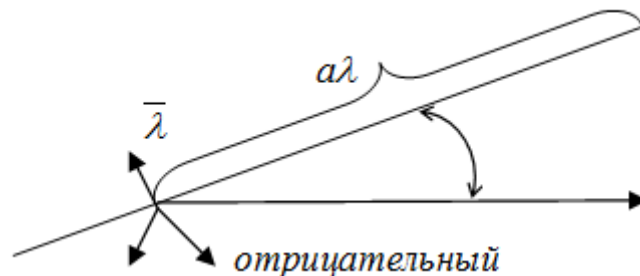


Чтобы задать направление, мы задаем вектор



Нормируем вектор $\bar{\lambda} = \frac{\bar{S}}{\|\bar{S}\|}$.

Нормированный вектор имеет то же самое направление, но $\|\bar{\lambda}\| = 1$ — длина. Допустим, что задан нормированный вектор $\|\bar{\lambda}\| = 1$

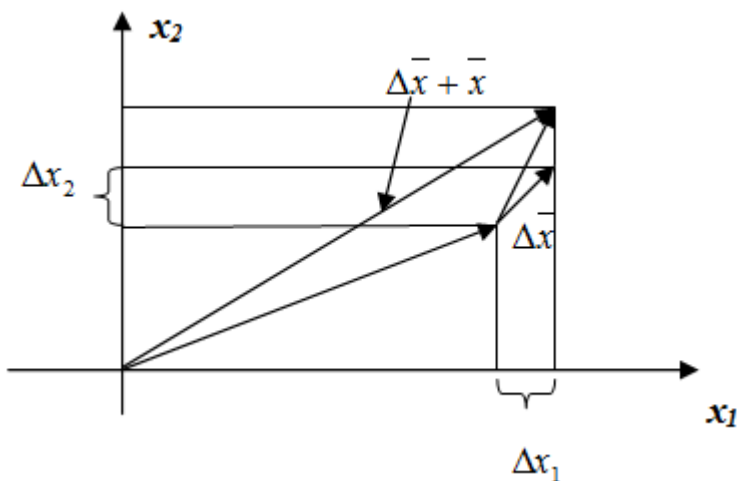


$$a\lambda = (\lambda, a) = \|a\| \cos a \wedge \lambda.$$

Скалярное произведение равно 0, когда $\cos a \wedge \lambda$ (прямой угол). Возвращаемся к функции 2-х переменных:

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2) = f(x_1, x_2) + \underbrace{\frac{\partial f}{\partial x_1} \Delta x_1 + \frac{\partial f}{\partial x_2} \Delta x_2}_{grad f} + O(\dots).$$

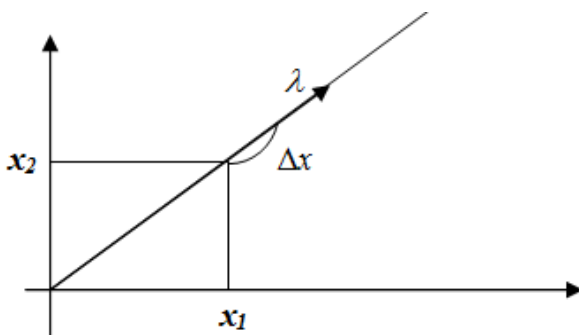
Отбрасываем члены $O(\dots)$, приращение будет более точным.



$$x = (x_1, x_2)$$

$$\Delta x = (\Delta x_1, \Delta x_2)$$

Вектор $\text{grad} f = \left(\frac{\partial f}{\partial x_1}; \frac{\partial f}{\partial x_2} \right) \Rightarrow f(x + \Delta x) = f(x) + (\text{grad} f, \Delta x) + O(\dots)$ – формула Тейлора.



$$\|\bar{\lambda}\| = 1.$$

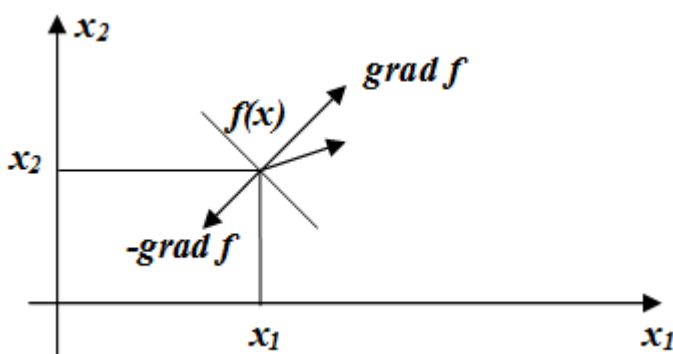
Мы рассматриваем, как изменяется точка вдоль данного направления. Функция становится функцией одной переменной, Δx – скалярная величина.

$$f(x + \Delta x) = f(x) + (\text{grad} f, \Delta x) + O(\dots),$$

где $\frac{\partial f}{\partial x} = (\text{grad} f, x)$ – производная по направлению (вдоль данного направления);

$\max \frac{\partial f}{\partial x}$ – направление ряда, равное направлению grad (≤ 0); grad – вектор, в сторону которого функция изменяется более быстро.

Антиградиент – grad направленный в другую сторону ($-\text{grad}$).



Необходимое условие:

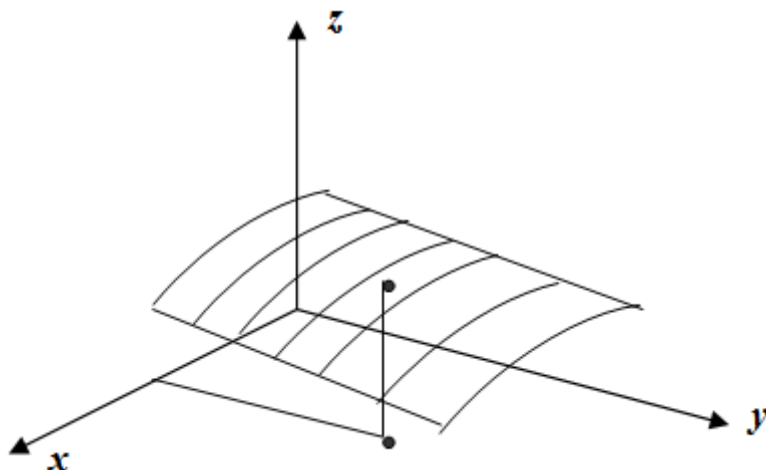
$\text{grad} f = 0$ – локальный минимум (или максимум). Точки локального экстремума.

Допустим, что мы совершаем малое перемещение Δx . В каком случае (в точке) будет: * больше, чем заданная: тогда, когда угол – острый $\Rightarrow f(x + \Delta x) = f(x) + (\text{grad} f, \Delta x) + O(\dots) = 0$.

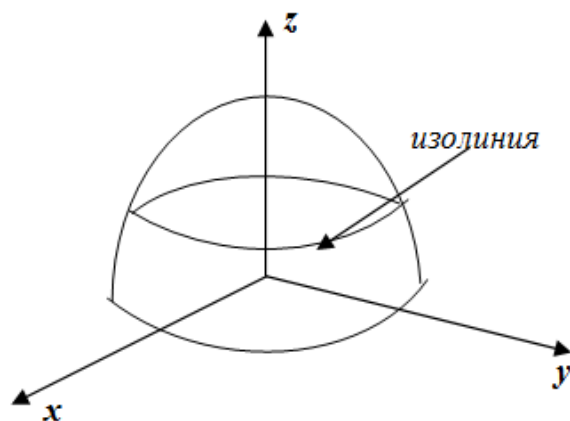
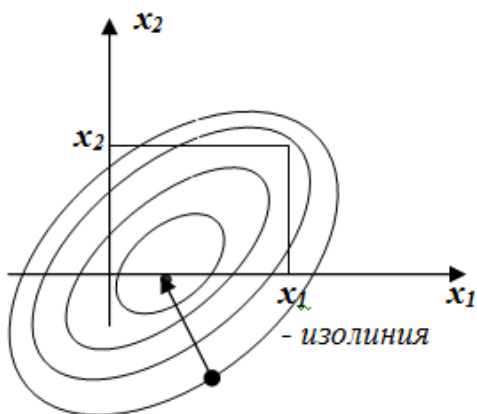
* – если под прямым углом, то не изменяется;

* – если под тупым углом, то приводит к уменьшению функции.

1. $y = f(x_1, x_2)$ строим поверхности



2. Идет построение в плоскости x_1 и x_2 . Выбираем точку, определяющую значение аргумента. Находим точку, в которой функция имеет то же самое значение, в результате получаем линию, в которой функция имеет постоянное значение – изолиния (линия уровня).



Вектор *grad* составляет прямой угол с изолинией.

Вернемся к формуле:

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2) = \underbrace{f(x_1, x_2)}_{\text{постоянный член}} + \frac{\partial f}{\partial x_1} \Delta x_1 + \frac{\partial f}{\partial x_2} \Delta x_2 + \\ + \frac{\partial^2 f}{\partial x_1^2} (\Delta x_1)^2 + \frac{\partial^2 f}{\partial x_2^2} (\Delta x_2)^2 + O(\dots).$$

10. Квадратичная аппроксимация (или квадратичное приращение).

Линейное отображение:

$$L: R^m \rightarrow R^n;$$

$$x \in R^m;$$

$$y \in R^n;$$

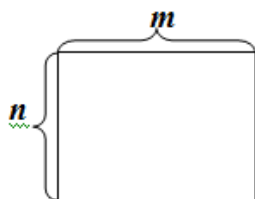
$$\bar{y} = y(\bar{x}) - \text{линейное отображение, если:}$$

1) свойство аддитивности – $\varphi(x + y) = \varphi(x) + \varphi(y)$;

2) свойство однородности – $\varphi(kx) = k\varphi(x)$.

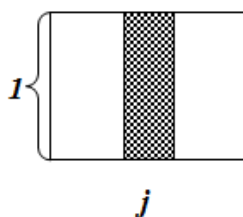
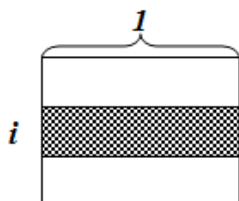
Линейное отображение можно задать матрицей:

$$y = Ax$$



$$y = A(Bx); C = AB;$$

$$a_j = \sum_k a_{ik} b_{kj} \text{ - основная формула}$$



$$A: R^n \rightarrow R^n \text{ отображение} \quad y = Ax \quad y_i = \sum_j a_{ij} x_j \quad Z = By = (BA)x$$

Две задачи:

1) решение системы уравнений $y = Ax$ – *линейный вектор* (x) ;

2) обратное отображение – найти x ;

A^{-1} – обратное отображение;

$AA^{-1} = I = A^{-1}A$ следовательно, строки матрицы ортогональны столбцам другой матрицы.

Нахождение собственных значений.

Используя матрицу можно найти более сложную функцию: $Z = (Ax, x)$ – квадратичная форма.

$Z = Z(x) = Z(x_1, x_2, \dots, x_n)$ – функция нескольких переменных

$$Z = \sum_i \sum_j a_{ij} x_i x_j.$$

Рассмотрим подробнее.

$$\text{Есть матрица: } A = \begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix};$$

$Z = 2x_1x_1 + 3x_1x_2 - 1x_2x_1 + 4x_2x_2$ – квадратичная форма;

$$Z = 2x_1^2 + 2x_1x_2 + 4x_2^2;$$

$$A' = \begin{pmatrix} 2 & 2 \\ 0 & 4 \end{pmatrix}.$$

A и A' определяют одну и ту же квадратичную форму следовательно значения этой формы не неоднозначно. Если по заданной квадратичной форме найдем симметрию, то она будет однозначная.

$$\bar{A} = \frac{1}{2}(A + A^T);$$

$$A^T = \begin{pmatrix} 2 & -1 \\ 3 & 4 \end{pmatrix}; \quad \frac{1}{2}(A + \bar{A}) = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}.$$

Без ограничения общности можно считать, что матрица, определяющая квадратичную форму, является симметричной.

Вернемся к квадратичной форме:

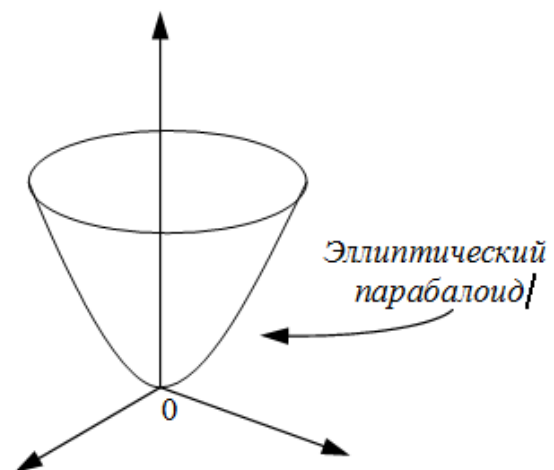
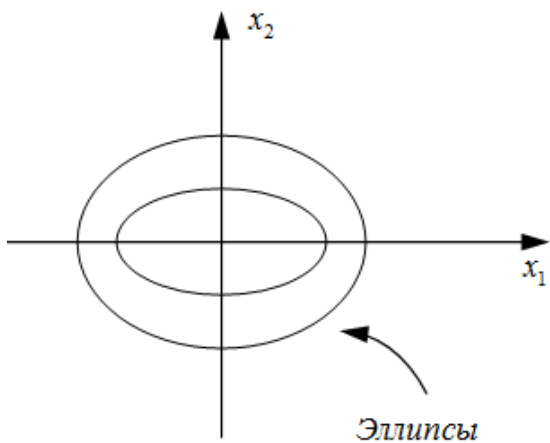
$$Z = (Ax, x) = \sum_i \sum_j a_{ij} x_i x_j.$$

Рассмотрим функцию 2-го порядка:

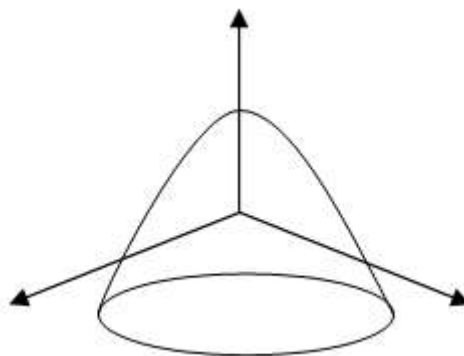
$$Z = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

Допустим, что $a_{12} = 0$, матрица диагональная.

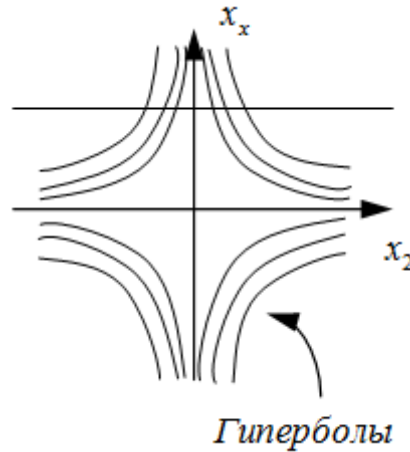
1. $a_{11}, a_{22} > 0$



2. $a_{11}, a_{22} < 0$



3. $a_{11} > 0, a_{22} < 0$



Допустим, что $a_{12} \neq 0$. Тогда вся картина просто повернется на некоторый угол по оси Z .

Рассмотрим n -мерный случай.

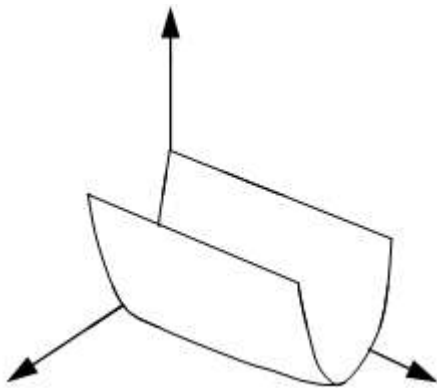
Квадратичная форма называется положительно определенной областью, если она не отрицательная.

1. $(Ax, x) \geq 0$, причем обращается в ноль, в том случае если $x = 0$ ($(Ax, x) = 0 \Leftrightarrow x = 0$). Этот случай соответствует эллиптическому параболоиду.

$$(Ax, x) \leq 0, (Ax, x) = 0 \Leftrightarrow x = 0.$$

2. Знаконеопределенность $(Ax, x) \begin{matrix} < \\ > \end{matrix} 0$ соответствует n -мерному эллиптическому гиперболоиду (n -мерное седло).

Рассмотрим 2-мерное пространство: $Z = 0x_1^2 + x_2^2$.



Если квадратная матрица называется положительно определенной, то и матрица положительно определена.

Рассмотрим разложение функции 2-х переменных в ряд Тейлора:

$$\begin{aligned} f(x_1 + \Delta x_1, x_2 + \Delta x_2) = & f(x_1, x_2) + \frac{\partial f}{\partial x_1} \Delta x_1 + \frac{\partial f}{\partial x_2} \Delta x_2 + \frac{\partial^2 f}{\partial x_1^2} (\Delta x_1)^2 + \\ & + \frac{\partial^2 f}{\partial x_1 \partial x_2} \Delta x_1 \Delta x_2 + \frac{\partial^2 f}{\partial x_2^2} (\Delta x_2)^2 + O(\dots) \end{aligned}$$

квадратичная матрица задается матрицей H

$$f(x + \Delta x) = f(x) + (\text{grad} f, \Delta x) + (H\Delta x, \Delta x) + O(\dots) = 0,$$

матрица, составленная из членов 2-го порядка

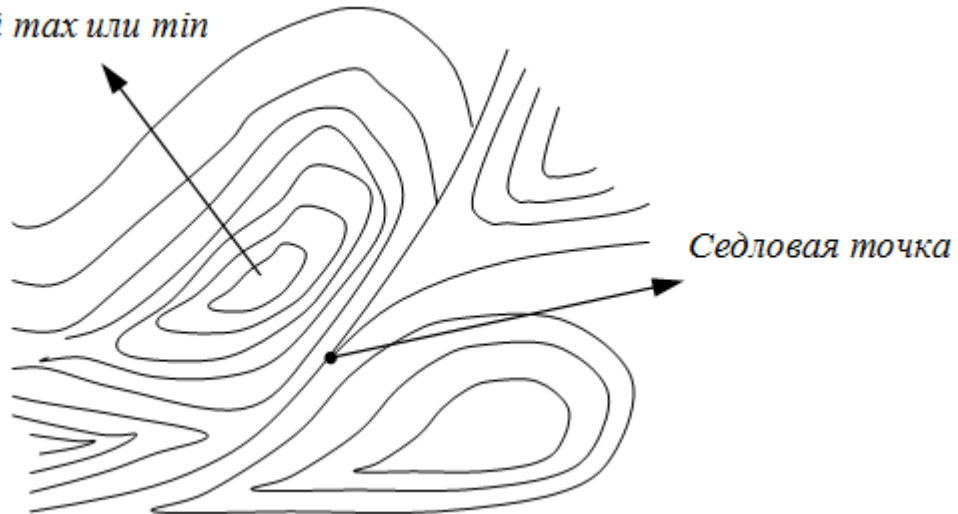
$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} - \text{матрица симметрична.}$$

Матрица H – матрица Гессе.

$$h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j} - \text{определение матрицы Гессе.}$$

Если матрица (матрица Гессе) в точке локального экстремума положительно определена, то это – точка локального минимума, если матрица отрицательно определена, то это – точка локального максимума, а если не определена – седловые точки.

Локальный max или min



Минимизируем:

$$f(x_1, x_2) = 2x_1^2 + \cos x_1 x_2 + 3x_2^2 \rightarrow \min.$$

Найти частные производные:

1. $\frac{\partial f}{\partial x_1} = 4x_1 - x_2 \sin x_1 x_2 = 0 \text{ (grad} = 0\text{)};$
2. $\frac{\partial f}{\partial x_2} = 6x_2 - x_1 \sin x_1 x_2 = 0.$

Эта система позволяет найти все точки экстремума:

$$\begin{cases} \frac{\partial f}{\partial x_1} = 4x_1 - x_2 \sin x_1 x_2 \\ \frac{\partial f}{\partial x_2} = 6x_2 - x_1 \sin x_1 x_2 \end{cases}$$

те x_1 и x_2 которые удовлетворяют уравнениям и будет точками экстремума.

Допустим, что $x_1^* = 3,5$ $x_2^* = 2,7$. Надо составить функцию второго порядка и подставить x_1^* x_2^* и посмотреть их.

Необходимые условия – помогают охарактеризовать искомую точку:

1. $\text{grad } f = 0$
2. $H = \frac{\partial^2 f}{\partial x_i \partial x_j}$

$H \geq 0$ – локальный минимум;

$H \leq 0$ – локальный максимум;

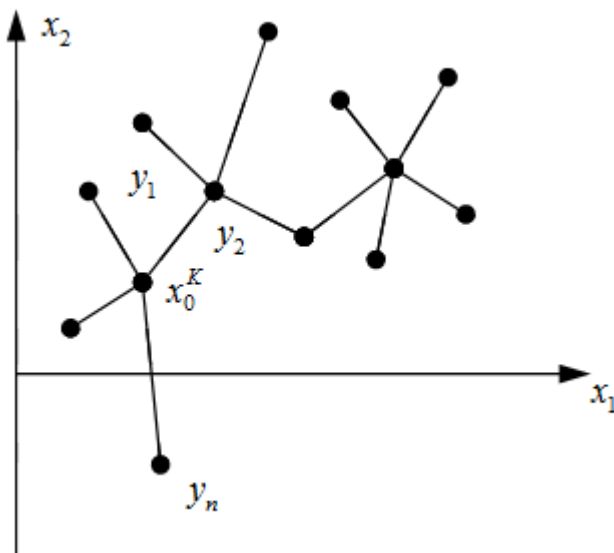
H – не определена – седловая точка.

Для поиска используют численные методы.

Постановка задачи. Требуется $f(x) \rightarrow \min$, где x – вектор $x = (x_1, \dots, x_n)$ – т.к. нет ограничений задача безусловной оптимизации.

Есть черный ящик, который по заданным значениям x позволяет вычислить значение функции.

12. Методы прямого поиска



Должны задать начальное приближение точки x_0 ;

x^k – некоторое приближение полученное после k – итераций; вычислить значение точки в окрестности точки x^k .

Из данных точек выбрать точку, в которой функция принимает наименьшее значение, выбираем ее и строим вокруг нее окрестность.

Выбираем точку. В окрестности существующей точки выбираем точку с меньшим значением, опять в ее окрестности есть точки с меньшим значением и т.д.

В таком виде этот метод не эффективен.

Пример. Шаг по x_1 берем больше, а по x_2 – сохраняем. Поскольку мы свободны в выборе точек, то можно менять шаг и направление.

Методы:

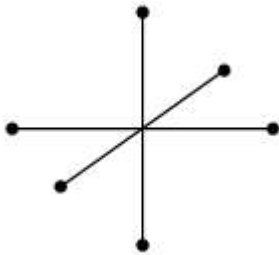
- 1) Хука-Дживса.
- 2) Нелдера-Мида (используется п-1 угольник).

Преимущества метода прямого поиска:

- простота;
- не нужны производные.

Недостатки:

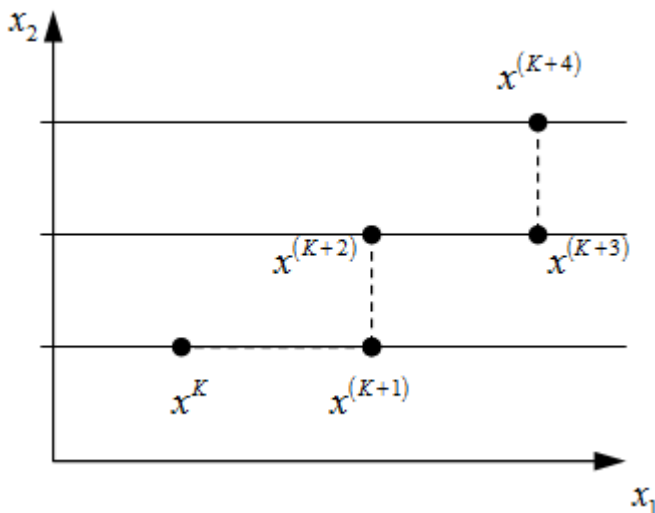
- плохая сходимость;
- применим для небольшого числа переменных $n \leq 10 \div 20$.



$2n$ точек:
в случае 2-х переменных – 4 точки;
в случае 3-х переменных – 6 точек.

Этот метод применим в простых случаях, когда эти недостатки себя не проявляют.

13. Метод координатного спуска



Существует приближенная точка минимума. Минимизируя функцию по направлению к x_1 на прямой, используется любой метод одномерной минимизации, x_2 – фиксируют. Далее выполняют одномерную оптимизацию по x_2 , фиксируя x_1 .

Этот процесс повторяют до тех пор, пока следующая точка не окажется близка к точке приближения.

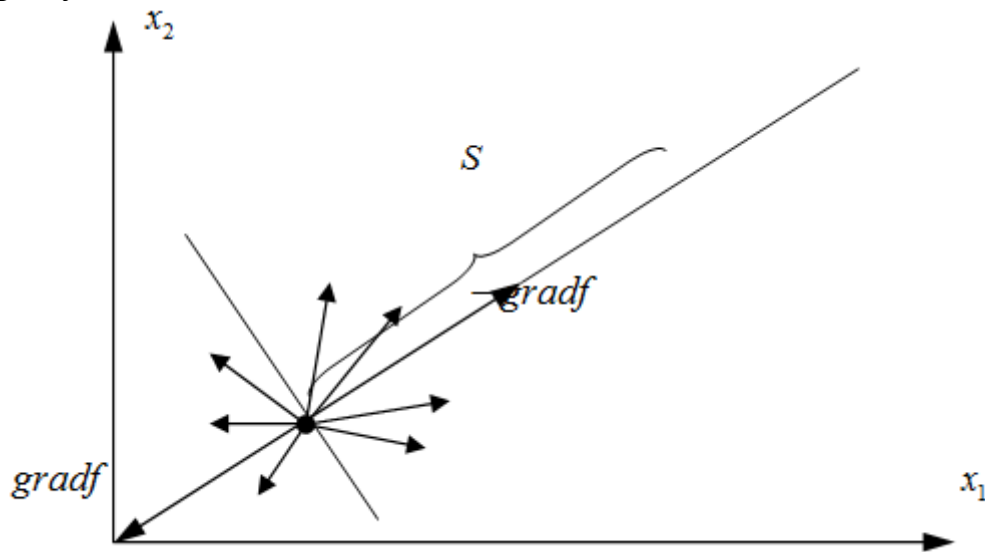
14. Градиентные методы

Метод наискорейшего спуска

Рассмотрим grad целевой функции.

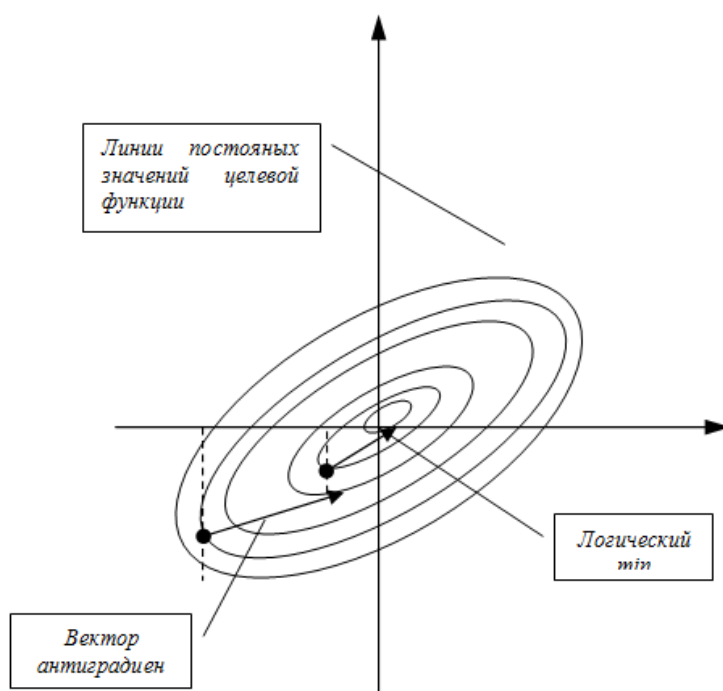
Движение по направлению вектора под острым углом будет приводить к уменьшению целевой функции, а движение против направления функции к

увеличению целевой функции. Разумно за направление движения принять сам вектор – $\text{grad } f$.



Для выбора расстояния нужно применить метод одномерной оптимизации. Прекратить поиск, когда величина $\text{grad } f$ станет достаточно малой. Этот метод гарантирует, что найдена либо точка локального минимума, либо седловая точка.

Анализ метода

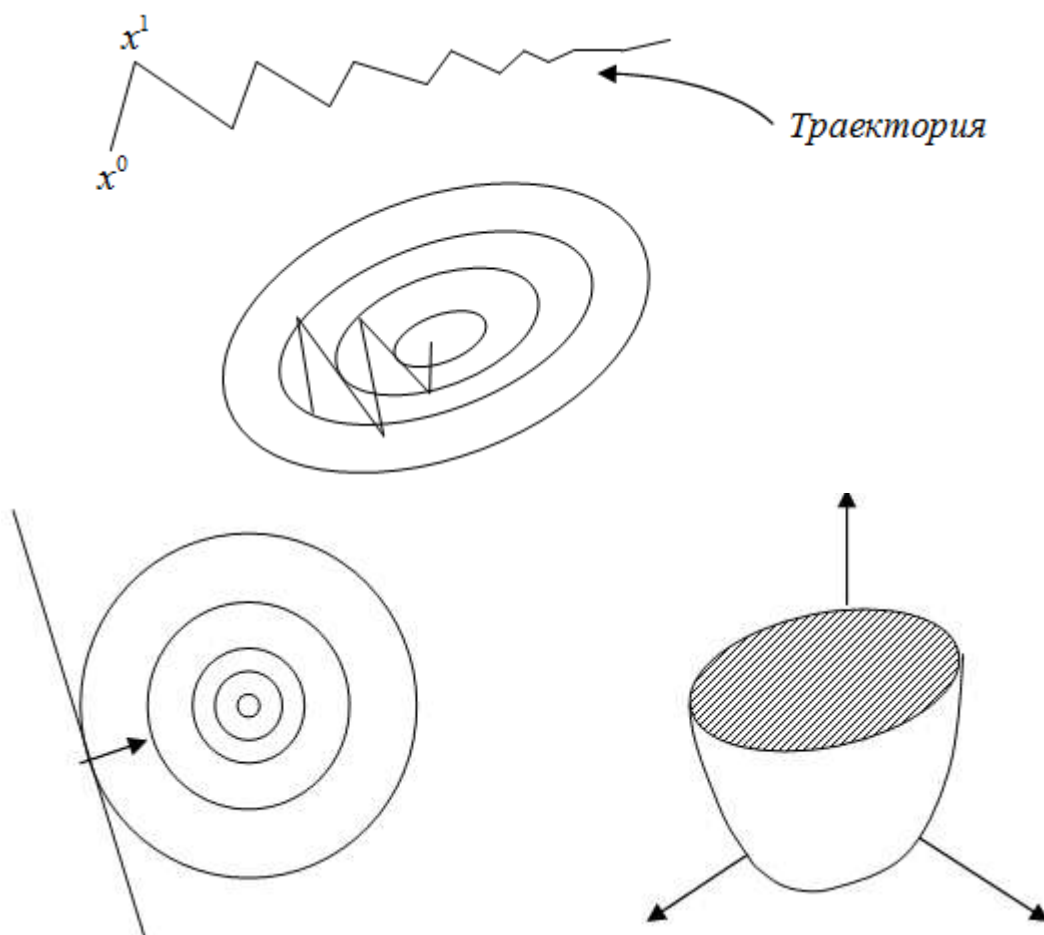


Рассмотрим целевую функцию, которая является квадратичной функцией, точка локального минимума совпадает с точкой начала координат.

Пусть мы выбрали начальное приближение.

Отыскивая наименьшее значение по направлению траектории (наименьшее значение там, где происходит касание $\text{grad } f$ линии уровня).

В случае, когда масштаб выбирается следующим образом (линии уровня вытянуты).



Если линии уровня $Z = x_1^2 + x_2^2$ – окружности, то приходим сразу в точку локального минимума.

14. Метод Ньютона

- $f(x) = \text{const}$ (одночлен) – один постоянный член любой точки данной функции является оптимальным – тривиальный случай;

- линейная функция (двучлен)

$$f(x) = C_1 x_1 + C_2 x_2 + \dots + C_n x_n + C_0 = (C, x) + C_0;$$

- $\text{grad} f = \nabla f = C$ (возможно бесконечное уменьшение и увеличение).

1 и 2 не подходят для оптимизации.

Рассмотрим трехчлен $f(x) = (C_1 x_1 + C_2 x_2 + \dots + C_n x_n) + C_0 + \sum_i \sum_j S_{ij} x_i x_j$;

$$f(x) = \frac{1}{2} (Qx, x) + (C, x) + C_0;$$

$$(Qx, x) = \sum_j y_j x_j = \sum_i \sum_j q_{ij} x_i x_j$$

без ограничения общности можно положить, что матрица q – симметричная $q_{ij} = q_{ji}$.

Разложим функцию в ряд Тейлора (должно быть 3 члена). Чтобы найти линейный член квадратичной функции, надо взять $grad$.

$$\begin{aligned}\frac{\partial}{\partial x_k} \left(\sum_i \sum_j q_{ij} x_i x_j \right) &= \sum_i \sum_j q_{ij} \frac{\partial}{\partial x_k} (x_i x_j) = \sum_i \sum_j q_{ij} \left(\frac{\partial}{\partial x_k} x_i + \frac{\partial}{\partial x_k} x_j \right) = \sum_i q_{ki} x_i + \\ &+ \sum_j q_{kj} x_j = \sum_j q_{kj} x_j + \sum_j q_{jk} x_j = 2 \sum_j q_{kj} x_j = 2Qx; \\ grad(Qx, x) &= 2Qx; \quad gradf = Qx + C; \quad C = 0.\end{aligned}$$

Найдем матрицу Гессе (матрица вторых частных производных)

$$H = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right) = Q,$$

элемент матрицы Гессе является элементом функции Q . $h_{ij} = q_{ij}$ (все частные производные высших порядков равны 0).

Функция экстремальна, если $grad$ в данной точке равен 0, $gradf = Qx + C$, следовательно, условие экстремальности $Qx = -C$ – система.

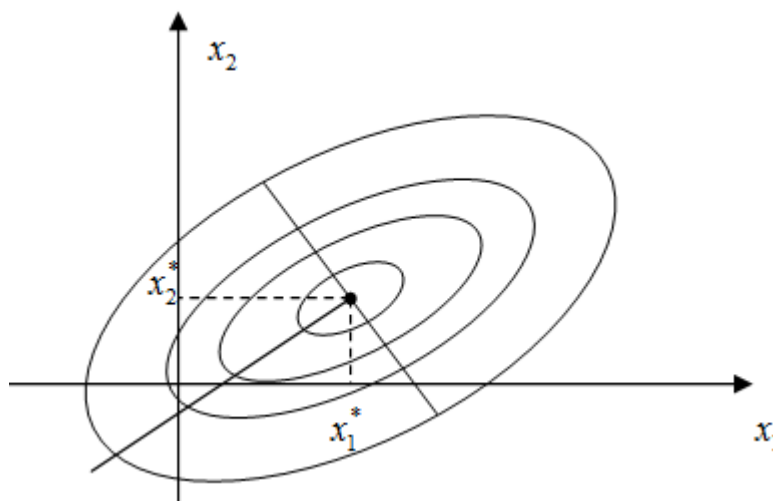
Необходимое условие оптимальности:

1. Если $Q > 0$, то решение данной системы существует и оно единственное (совместная система).

2. Если $Q < 0$, то решение данной системы существует и оно единственное, т.е. если Q знакоопределена, то существует решение и оно единственное.

3. Если имеем квадратичную функцию и матрица положительно определена, то линии уровня – эллипсы.

Собственные значения определяют оси эллипсов.



Чтобы определить координаты точки локального минимума, нужно решить систему $Qx^2 = -C$.

Пусть $f(x)$ – произвольная функция и надо найти точку локального минимума. Разложим функцию в ряд Тейлора в окрестности точки.

$$f(x^K + \Delta x^K) = f(x^K + \nabla f \Delta x^K) + (H \Delta x^K, \Delta x^K) + O(\dots)$$

Пусть функция не квадратичная, эллипсы примерно отражают кривизну линий уровня и находятся в окрестности точки x^K . В окрестности точки x^K находим приближение и заменяем эту функцию квадратичной функцией, которая получается из разложения в ряд Тейлора. Далее решаем задачу минимизации.

Находим точку минимума и рассматриваем эту точку как следующее приближение и т.д.

Для нахождения точки минимума квадратичной функции (зависит от Δx) необходимо решить систему:

$$H\Delta x^K = -\nabla f(x^K).$$

Окончательно следующее приближение

$$x^{K+1} = x^K + \Delta x^K.$$

где $x^{K+1} = x^K - H^{-1}\nabla f(x^K)$ – формула Ньютона (обобщение формулы минимизации одной переменной).

Выполнение метода останавливается когда $\nabla \approx 0$, т.е. когда Δx^K очень мало. Для получения практической точности достаточно выполнить 4 итерации метода Ньютона.

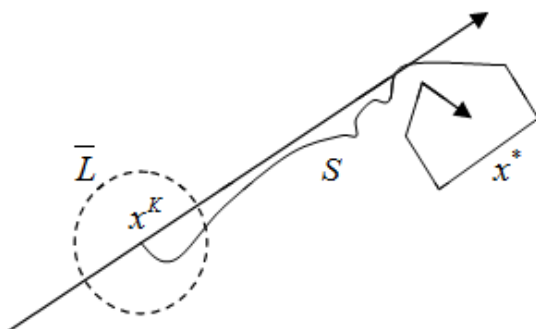
Если f – хороша, то метод Ньютона подходит, если f – квадратичная функция, то метод Ньютона приводит к минимальной точке за 1 шаг, из любой точки.

Недостатки:

- на каждом шаге итерации надо находить решение системы $H\Delta x^K = -\nabla f(x^K)$;
- с ростом числа итераций H – разрежается, т.е. большое число членов становится равными 0.

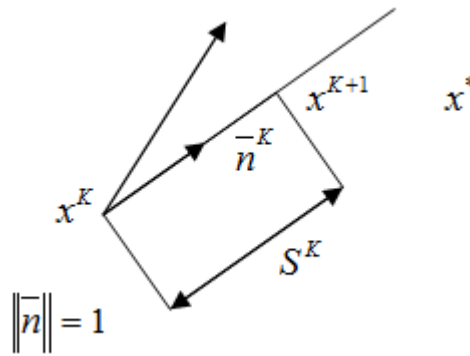
Все формулы безусловной минимизации можно записать в общую схему:

- выбор направления;
- выбор шага.



x^K – приближение в точке локального минимума, чтобы приблизиться к искомой точке x^* . Мы должны выбрать направление, в конце получим локальную линию.

Допустим, требуется $f(x) \rightarrow \min$; x^0 - начальное приближение; x^K – текущее приближение:



а) выбор направления $(grad f, \bar{n}) < 0$;

б) движение вдоль выбранного направления $\overline{X^{K+1}} = \overline{X^K} + S^K * n^K$.

15. Задачи оптимизации с ограничениями – разностями (ЗОР)

$$\begin{cases} f(x_1, x_2, \dots, x_n) \rightarrow \min \\ g_1(x_1, x_2, \dots, x_n) = 0 \\ \dots \dots \dots \\ g_m(x_1, x_2, \dots, x_n) = 0 \end{cases} \quad \begin{matrix} f(x) \rightarrow \min \\ g(x) = 0 \end{matrix}.$$

Пример:

$$\begin{cases} f(x_1, x_2, x_3, x_4) \rightarrow \min \\ g_1(x_1, x_2, x_3, x_4) = 0 \\ g_m(x_1, x_2, x_3, x_4) = 0 \end{cases} \quad \begin{matrix} \text{Функции заданы аналитическим выражением} \Rightarrow \\ q_1 \text{ можно разрешить относительно одной из перемен-} \\ \text{ных } x_1 = h_1(x_2, x_3, x_4) \Rightarrow x_1 - \text{можно исключить из } f \text{ и} \\ q_2, \text{ подставив вместо нее } L_1: \end{matrix}$$

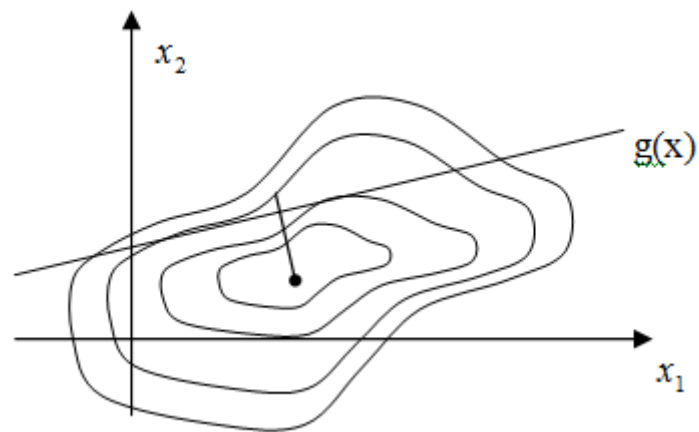
$$\begin{cases} \overline{f}(x_2, x_3, x_4) \\ \overline{g}(x_2, x_3, x_4) = 0 \end{cases} \Rightarrow x_2 = h_2(x_3, x_4).$$

Тогда, $\overline{\overline{f}}(x_3, x_4)$ – задача безусловной оптимизации.

Находим $x_3, x_4 \Rightarrow$ вычисляем $x_2 \Rightarrow x_1$

Метод исключения. Численное решение:

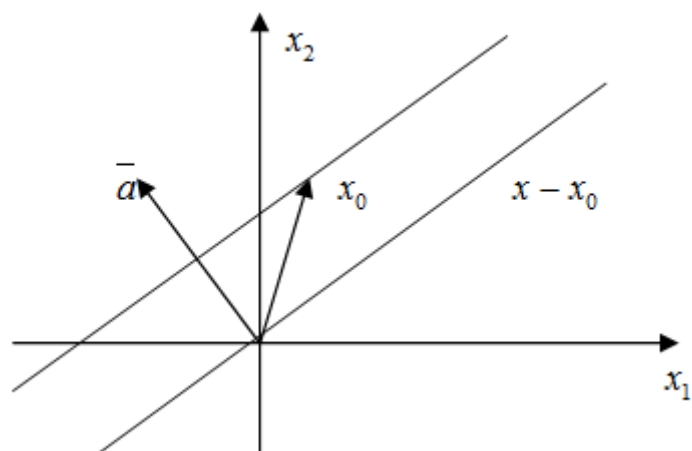
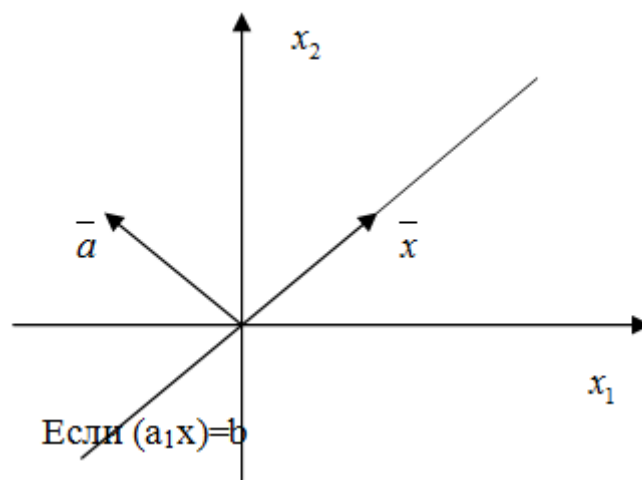
$$\begin{cases} f(x_1, x_2) \rightarrow \min \\ a_1 x_1 + a_2 x_2 = b \end{cases} \\ [g(x_1, x_2) = a_1 x_1 + a_2 x_2 - b] \Rightarrow \text{точка } \min \text{ должна лежать на прямой.}$$



В каждый момент линия уровня будет касаться прямой \Rightarrow эта точка и является точкой условного локального min. Если в окрестности заданной точки, удовлетворяющей всем значениям равенства, значение функции больше, чем в точке, то эта точка – есть точка условного локального min.

Пример:

$$\begin{aligned} a_1x_1 + a_2x_2 &= 0; \\ (a, x) &= 0; \\ a &\perp x. \end{aligned}$$



Допустим, $b = (a, x_0)$;

$$(a, x) - (a, x_0) = 0;$$

$$(a, x - x_0).$$

Прямая будет проходить через некоторую точку, удовлетворяющую условию и $\perp \bar{a}$.

Для n переменных $f(\bar{x}) \rightarrow \min$, $Ax = b$.

Рассмотрим i -ое ограничение:

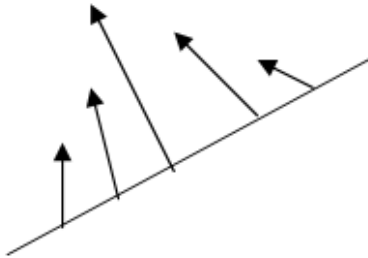
$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = 0, \quad (a, x) = 0,$$

где \bar{a} – вектор задан $\rightarrow x$ – все вектора, лежащие $\perp \bar{a}$. Они и составляют гиперплоскость.

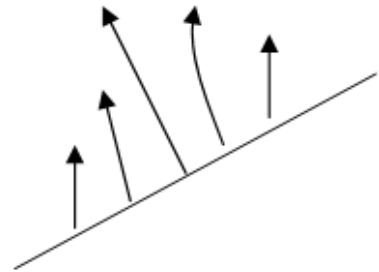
При добавлении еще одного условия, уменьшаются размерности. В конечном итоге получится пространство $n-m$.

Для двух переменных возможно 2 случая:

1.

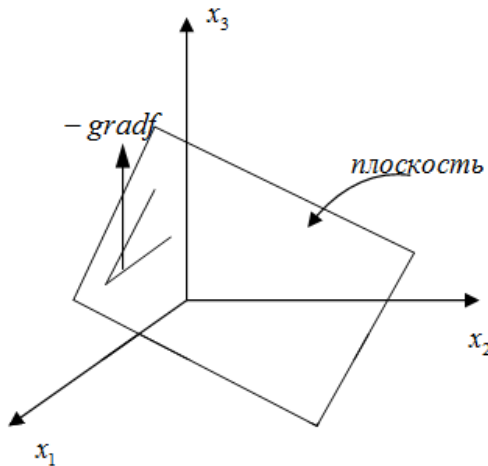


2.



В случае 2 это не точка минимума, а седловая точка.

Рассмотрим точку 3-х переменных:



Ограничение – плоскость, следовательно, все допустимые точки на плоскости. Если угол $grad$ не равен 90 градусам, следовательно, можно двигаться дальше. На плоскости существует направление, которое будет составлять острый угол с $grad$ и, двигаясь в этом направлении, можно уменьшить значение f . Если $-grad f$ перпендикулярен плоскости, то эта точка может быть точкой минимума.

Пусть существует ограничения:

$$\begin{cases} f(x) \rightarrow \min; \\ a_{11}x_1 + a_{12}x_2 + \dots + a_{1n} = b_1; \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n} = b_2. \end{cases}$$

Рассмотрим опять случай 3-х переменных.

Точка минимума должна принадлежать пересечению плоскостей.

Необходимое условие – вектор антиградиента должен составлять угол 90 градусов с прямой пересечения плоскостей.

Для n -мерного случая имеется n переменных, следовательно, рассматривая каждое ограничение; получаем $n-1$ гиперплоскость, следовательно, рассмотрев m ограничений, получим $n-m$ гиперплоскость ($m < n$).

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n} = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n} = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn} = b_m \end{array} \right. \quad \begin{array}{l} Ax = b \\ \text{все ограничения} \\ \text{независимы} \end{array}$$

Если вектор $grad$ (n -мерный) будет ортогонален $n-m$ – пространству.

Допустим, имеется $n-1$ пространство, n -мерный вектор может принадлежать ему или нет. Пусть вектор не принадлежит данному подпространству, следовательно, его можно разложить на 2 вектора – один который принадлежит подпространству, и второй который ортогонален данному. Ортогональное дополнение – вектора, которые ортогональны данному подпространству.

В 3D – пространстве, если подпространство равно 1, следовательно, ортогональное дополнение равно 3.

В $n-m$ -мерном подпространстве ортогональное дополнение имеет размерность m .

Необходимое условие. Если мы находим точку, где вектор градиента принадлежит ортогональному дополнению к пространству, заданному ограничениями – равенствами, то эта точка может быть точкой локального минимума.

Пусть есть 2 плоскости. Если записать систему ограничений равенств следующим образом:

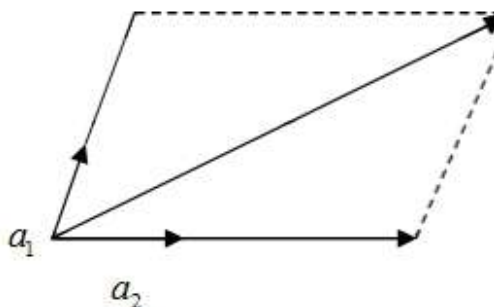
$$\left\{ \begin{array}{l} (a_1, x) = b_1 \\ (a_2, x) = b_2 \\ \vdots \\ (a_m, x) = b_m \end{array} \right. ,$$

где $a_i = a_{i1}, a_{i2}, \dots, a_{in}$.

Таким образом, вектора a_1, a_2, \dots, a_m порождают ортогональное дополнение. Существующие могут быть выбраны в качестве базиса ортогонального дополнения, следовательно, градиент принадлежит ортогональному дополнению:

$$-gradf = \lambda_1 \overline{a_1} + \lambda_2 \overline{a_2} + \dots + \lambda_m \overline{a_m} ,$$

т.е. линейная комбинация базисных векторов.



$$\bar{V} = \lambda_1 \bar{a}_1 + \lambda_2 \bar{a}_2,$$

где $\lambda_1 + \lambda_2 + \dots + \lambda_m$ – множители Лагранжа.

Рассмотрим матрицу A^T , в ней a_1, \dots, a_m – столбцы.

$$-gradf = A^T \lambda,$$

это условие может быть использовано для численного решения задачи оптимизации с ограничивающими уравнениями.

Пример:
$$\begin{cases} f(x) \rightarrow \min; \\ Ax = b \end{cases}$$

$$Ax = b$$

$$A^T \lambda = -gradf$$

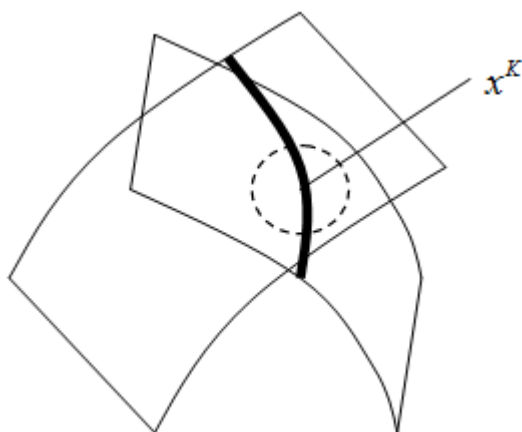
Если найдем такие вектора x и λ , для которых эти условия выполняются, то точка может быть точкой локального минимума.

Рассмотрим случай, когда система ограничений равенств нелинейная:

$$\begin{cases} f(x_1, \dots, x_n) \rightarrow \min \\ g_1(x_1, \dots, x_n) = 0 \\ \vdots \\ g_m(x_1, \dots, x_n) = 0 \end{cases};$$

$$\begin{cases} f(\bar{x}) \rightarrow \min \\ g(\bar{x}) = 0 \end{cases}.$$

Если функции дифференцируемы, то в окрестности точки минимума они будут вести себя как линейные.



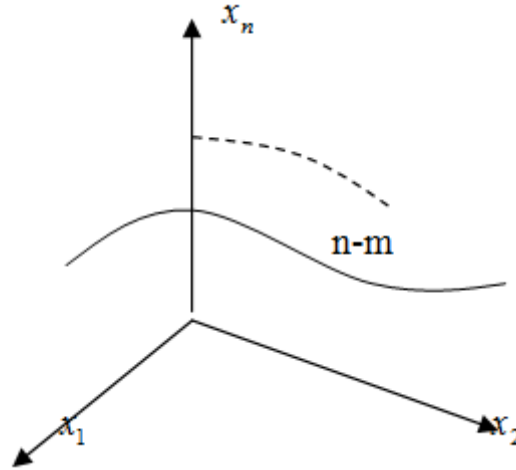
$$\bar{g}_i(x^* + \Delta x) = g_i(x^*) + \frac{\partial g_i}{\partial x_j} \Delta x + O(\dots),$$

следовательно, в окрестности точки локального минимума эта зависимость линейная, следовательно, получается система вида:

$$A\Delta x = g_i(x^*) - \bar{g}_i(x^* + \Delta x) + O(\dots),$$

$$\text{где } A = \frac{\partial g_i}{\partial x_j},$$

следовательно, необходимое условие локального минимума:



$$-gradf = \lambda_1 gradg_1 + \lambda_2 gradg_2 + \dots + \lambda_m gradg_m,$$

где $\lambda_1 + \lambda_2 + \dots + \lambda_m$ – множители Лагранжа.

$$\begin{cases} f(\bar{x}) \rightarrow \min \\ g(\bar{x}) = 0 \end{cases}$$

$$-gradf = \lambda_1 gradg_1 + \lambda_2 gradg_2 + \dots + \lambda_m gradg_m;$$

x^* – точка может быть искомой в задаче;

$\lambda_1 + \lambda_2 + \dots + \lambda_m$ – множители Лагранжа;

$$\begin{cases} n \\ m \end{cases} \begin{cases} -gradf = \lambda_1 gradg_1 + \lambda_2 gradg_2 + \dots + \lambda_m gradg_m \\ g(x) = 0 \end{cases}^m.$$

Обозначения для скалярного произведения $(\bar{x}, \bar{y}) = x^T y$;

$$(\bar{x}, \bar{y}) = (x_1, \dots, x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_i x_i y_i;$$

$$L(x, y) = f(x) + \lambda^T g(x) = f(x) + \lambda_1 g_1(x) + \dots + \lambda_m g_m(x).$$

Необходимое условие точки локального минимума – исходное задание с ограничениями представляет собой необходимое условие точки локального экстремума для функции Лагранжа.

$$\frac{\partial L}{\partial x_j} = \dots = f(x) + \lambda_1 g_1(x) + \dots + \lambda_m g_m(x) = \frac{\partial f}{\partial x_j} + \lambda_1 \frac{\partial g_1}{\partial x_j} + \dots + \lambda_m \frac{\partial g_m}{\partial x_j} = 0;$$

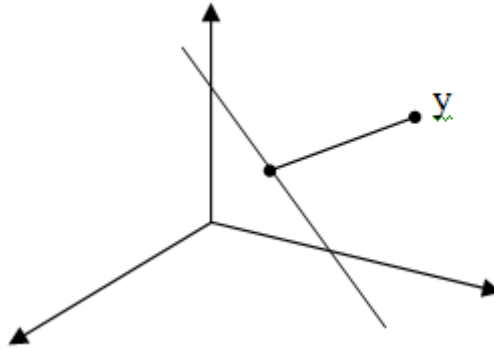
$$\frac{\partial L}{\partial x_j} = g_i(x) = 0, i = 1, \dots, m.$$

16. Метод множителей Лагранжа. Применяется для нахождения точки локального минимума для точек исходной задачи $h(x, \lambda) = f(x) + \lambda^T g(x)$. Экстремальными точками локального минимума являются седловые.

Пример:

Найти расстояние от точки до прямой в 3-х мерном пространстве.

Плоскость:
$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n} &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n} &= b_2 \end{aligned}$$



Пересечение плоскостей – линия

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2} \rightarrow \min;$$

$$Z = (x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 \rightarrow \min;$$

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n} - b_1 &= 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n} - b_2 &= 0 \end{aligned};$$

$$L(x_1, x_2, x_3, \lambda_1, \lambda_2) = \sum_{j=1}^3 (x_j - y_j)^2 + \lambda_1 \left(\sum_{j=1}^3 a_{1j}x_j - b_1 \right) + \lambda_2 \left(\sum_{j=1}^3 a_{2j}x_j - b_2 \right);$$

$$\begin{cases} \frac{\partial L}{\partial x_1} = 0 \\ \frac{\partial L}{\partial x_2} = 0 \\ \vdots \\ \frac{\partial L}{\partial x_3} = 0 \\ \frac{\partial L}{\partial \lambda_1} = 0 \\ \frac{\partial L}{\partial \lambda_2} = 0 \end{cases}$$

5 условий дают систему линейных уравнений.

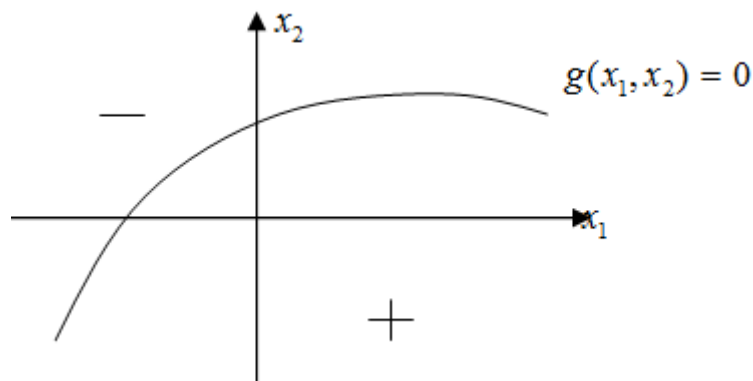
17. Нелинейное программирование (НЛП)

$$\begin{cases} f(x_1, \dots, x_n) \rightarrow \min \\ g_i(x_1, \dots, x_n) = 0, i = 1, \dots, m; \\ h_k(x_1, \dots, x_n) \geq 0 \end{cases}$$

f, g_i, h_k – заданные функции нелинейные.

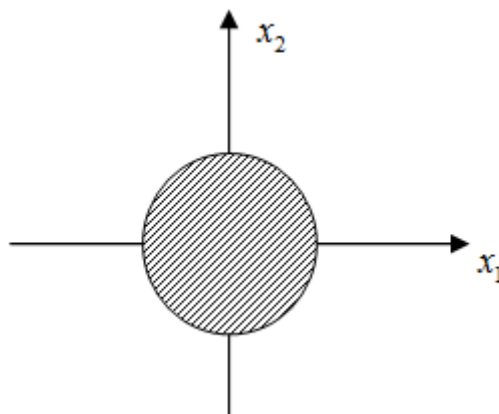
$$\text{НЛП} \begin{cases} f(\bar{x}) \rightarrow \min \\ \bar{g}(\bar{x}) = 0; \\ \bar{h}(\bar{x}) \geq 0 \end{cases}$$

Рассмотрим $g(x_1, x_2) \geq 0$.



Пример:

$$\begin{aligned} x_1^2 + x_2^2 &\leq 4; \\ g(x_1, x_2) &= 4 - x_1^2 - x_2^2 \geq 0. \end{aligned}$$

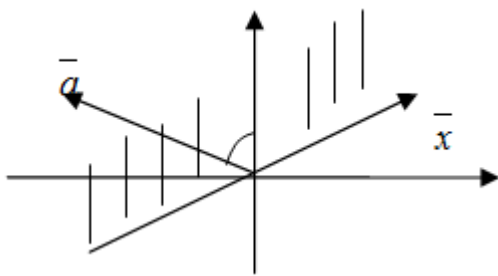


В случае системы неравенств пересечение всех областей. Если $g > 0$, то ограничение неравенства – неактивно (точку можно смещать).

Если точка находится точно на границе, то говорят, что ограничение активно.

Рассмотрим случай: $a^T x = a_1 x_1 + a_2 x_2 \geq 0$.

Е



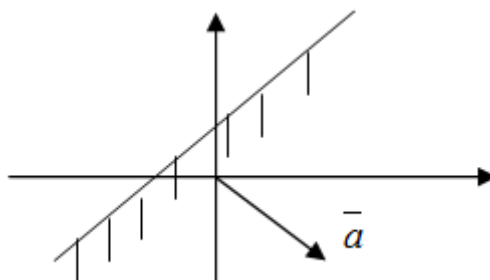
Если задано линейное ограничивающее неравенство, то вектор \bar{a} направлен внутрь допустимой области. Если $a^T x \leq 0$, то вектор \bar{a} будет направлен из допустимой области.

если $a^T x \geq b$, то граница проходит

дит не через начало координат.

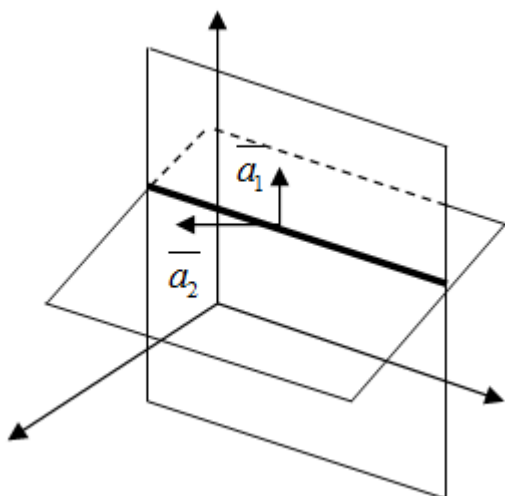
Необходимые условия:

$$\begin{aligned} f(x_1, x_2) &\rightarrow \min; \\ a_1 x_1 + a_2 x_2 &\geq b \end{aligned}$$



1. Если локальный минимум внутри допустимой области, то $gradf = 0$;

2. Если точка локального минимума точно на границе, то $gradf = \lambda \bar{a}$, точка является точкой локального, если $gradf = \lambda \bar{a}$ и $\lambda \geq 0$.



\bar{a}_1, \bar{a}_2 – вектора нормали к соответствующей плоскости.
 $\lambda_1, \lambda_2 \geq 0$

В общем случае:

$$\left\{ \begin{array}{l} f(\bar{x}) \rightarrow \min \\ g_1(\bar{x}) = 0 \\ \vdots \\ g_m(\bar{x}) = 0 \\ h_1(\bar{x}) \geq 0 \\ \vdots \\ h_m(\bar{x}) \geq 0 \end{array} \right.$$

$$\text{а) } \text{grad} f(\bar{x}) = \bar{h}_1 \text{grad} g_1(\bar{x}) + \dots + \bar{h}_m \text{grad} g_m(\bar{x}) + \lambda_1 \text{grad} h_1(\bar{x}) + \dots + \lambda_p \text{grad} h_p(\bar{x});$$

$$\begin{array}{l} \lambda_1 \geq 0; \\ \text{б) } \lambda_2 \geq 0; \\ \quad \vdots \\ \lambda_p \geq 0 \end{array}$$

в) Если $h_k = 0$, то $\lambda_k > 0$. Если $h_k > 0$, то $\lambda_k = 0$. Т.е. $\lambda_k + h_k = 0, k = 1, \dots, p$. Условие дополняющей нежесткости.

Все три условия в совокупности называются условиями Куна-Таккера (условия оптимальности первого порядка).

$$\left\{ \begin{array}{l} f(\bar{x}) \rightarrow \min \\ \bar{g}(\bar{x}) = 0 \\ \bar{h}(\bar{x}) \geq 0 \end{array} \right.$$

Ограничения неравенства

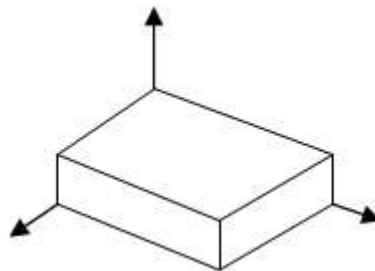
$$t_i = h_i(\bar{x}) \geq 0$$

\Downarrow

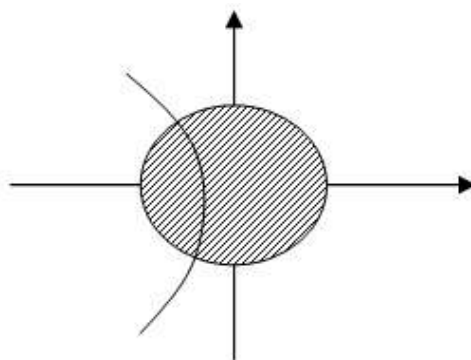
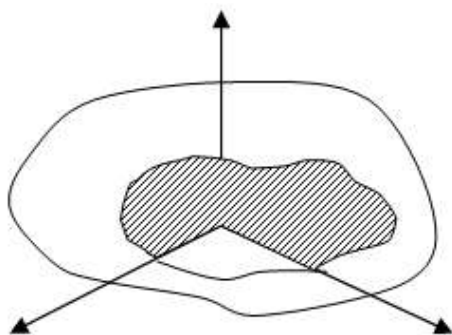
$$t_i - h_i(\bar{x}) = 0, t_i \geq 0$$

Можно записать и так:

$$\left\{ \begin{array}{l} f(\bar{x}) \rightarrow \min \\ \bar{g}(\bar{x}) = 0 \\ \bar{x} \geq 0 \end{array} \right. .$$



Поскольку постановка задачи $\bar{h}(\bar{x}) \geq 0$, то:



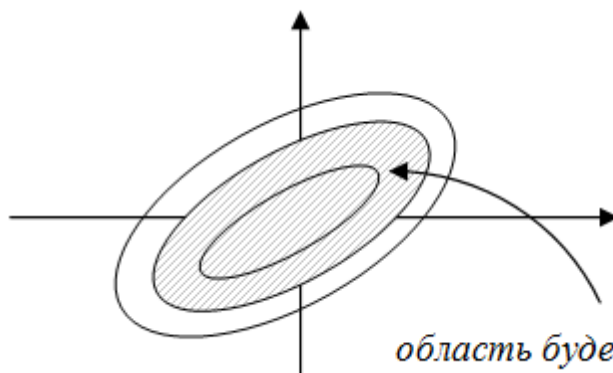
Основные результаты: Область n -мерного пространства называется выпуклой, если вместе с двумя точками, она содержит весь отрезок, соединяющий эти 2 точки.

Пример:

Функция нескольких переменных $f(\bar{x})$ называется выпуклой, если ее матрица Гессе положительно определена.

$$H = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right); \quad x^T H x \geq 0.$$

Если мы рассматриваем неравенство $f(x) \leq C$, то данное неравенство определяет выпуклую область.



область будет выпуклой

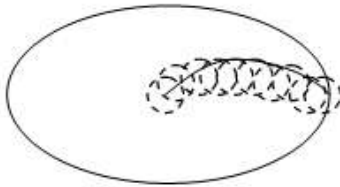
Пусть дана задача НЛП, если целевая функция этой задачи – выпуклая, и область целевых решений так же выпукла, то локальный оптимум совпадает с ее глобальным оптимумом задачи (задачи выпуклого программирования):

- случай – когда все ограничительные неравенства являются не активными;
- случай – когда точка лежит на границе.

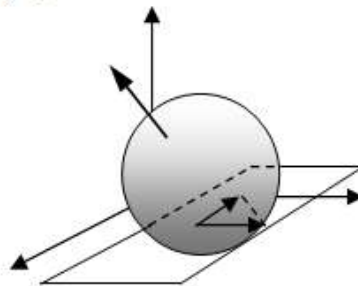
Методы решения НЛП

<p><i>Нулевого порядка</i> – поисковые методы (безусловные ориентиры похожи на это). Используется только значение целевой функции (Z).</p>	<p><i>Первого порядка</i> – аналогичны градиентным методам. Условно градиентные методы. Используется и Z и вектор градиента ($grad\ Z$).</p>	<p><i>Второго порядка.</i> Ньютоновские методы. Они являются специальными вариантами методов Ньютона для оптимизации. Используется Z, $grad\ Z$ и матрица Гесса (H)</p>
---	--	--

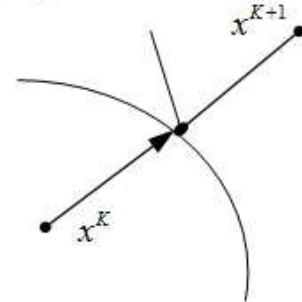
(*)



(**)



(***)



(**)

Первый случай – вектор $grad$ направлен по нормали; второй случай – идет под углом (надо спроецировать поверхность, следовательно, она будет показывать направление).

Если мы находимся внутри, то двигаемся как в (*), а далее (**). Это более эффективный метод.

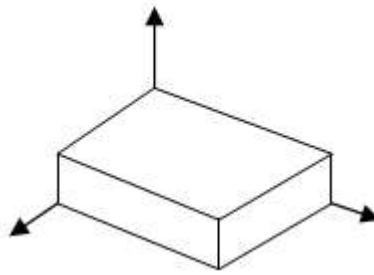
(***)

Рассмотреть отрезок, это может дать нам еще один отрезок.

18. Задачи линейного программирования (ЛП)

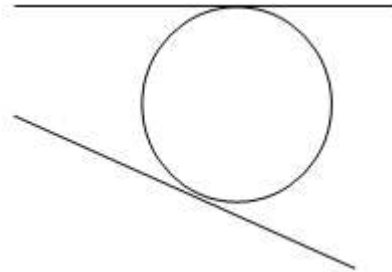
$$\begin{cases} Z = C^T x + C_0 \rightarrow \min \\ Ax = b \\ x \geq 0 \end{cases} \quad \begin{array}{l} \text{- стандартная форма} \\ \text{задачи ЛП} \end{array}$$

В общем случае, если $Ax \geq b$, то допустимая область представляет собой многогранник в пространстве.



В случае $Ax = b, x \geq 0$ – многогранник, имеющий неполную размерность.

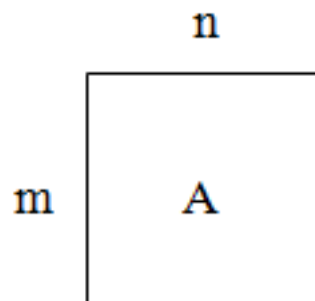
Допустим, имеется некоторое выпуклое множество. Тогда в любой граничной точке этого множества всегда можно провести опорную гиперплоскость, т.е. такую гиперплоскость, которая имеет с множеством только одну общую точку, и все множество находится по одну сторону от гиперплоскости.



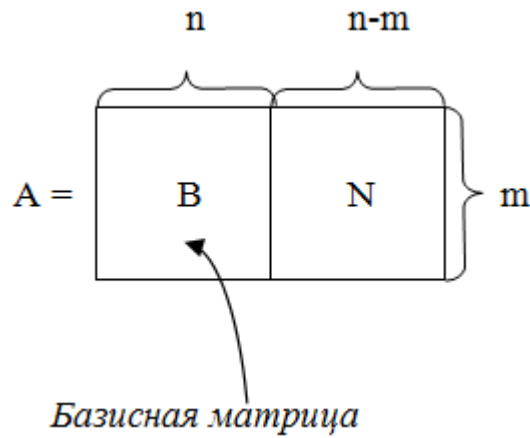
Если $-grad$ является нормалью к гиперплоскости и плоскость не опорная, то можно двигаться под острым углом к $grad$, тем самым улучшая значение функции. Такое движение невозможно, если антиградиент определяет опорную плоскость. Следовательно, в этом случае это точка локального минимума, который является и глобальным минимумом.

Геометрически, чтобы найти точку локального минимума, необходимо найти такую вершину глобального множества, что плоскость, которая является нормальной к антиградиенту, является опорной.

Рассмотрим $Ax = b, x \geq 0$, m – ограничений равенств, n – число переменных.



Первые m столбцов линейно независимы. $rank(B) = m$, $\det(B) \neq 0$.



$Ax = A_1x_1 + A_2x_2 + \dots + A_nx_n$, A_1, A_2, \dots, A_n – столбцы матрицы,

$$A = \frac{B}{N} \Rightarrow x = \frac{x_B}{x_N} \text{ – базисные переменные.}$$

Тогда систему ограничений равенств можно записать

$$\frac{B}{N} \frac{x_B}{x_N} + b;$$

$$Bx_B + Nx_N = b \quad Bx_B = -Nx_N + b;$$

Для B существует обратная матрица $B^{-1}x_B = -B^{-1}Nx_N + B^{-1}b$;

$$x_B = -B^{-1}Nx_N + B^{-1}b.$$

Если для данного базиса зафиксируем не базисные переменные в нуле, то получим точку, которая является вершиной многогранника.

Вершины многогранника множества, характеризующиеся тем, что небазисные переменные равны 0.

Исследование в случае, когда вершина не является точкой оптимума.

Рассмотрим целевую функцию:

$$\begin{aligned} Z = C^T x &= \begin{pmatrix} C_B \\ C_N \end{pmatrix}^T \begin{pmatrix} x_B \\ x_N \end{pmatrix} = C_B^T x_B + C_N^T x_N = C_B^T (-B^{-1}Nx_N + B^{-1}b) + C_N^T x_N = \\ &= -C_B^T B^{-1}Nx_N + C_B^T B^{-1}b + C_N^T x_N = (C_N^T - C_B^T B^{-1}N)x_N + C_B^T B^{-1}b = \\ &= \underbrace{\left(C_N - N^T B^{-T} C_B \right)}_d^T x_N + \underbrace{b^T B^{-T} C_B}_{d_0} = B^{-T} = (B^T)^{-1}. \end{aligned}$$

где d – показывает суммарное влияние небазисных переменных на целевую функцию;

d_0 – множители Лагранжа или относительные оценки небазисных переменных.

$$Z = C^T x = d^T x_N + Z_0$$



Точка будет точкой оптимума, если все $d \geq 0$.

Если имеется один отрицательный коэффициент

$$Z = d_1(x_N)_1 + \dots + d_j(x_N)_j + \dots + d_n(x_N)_n + Z_0, \text{ то:}$$

- $d_j < 0$, следовательно, можно увеличить x_N , тогда целевая функция начнет улучшаться;

- $(x_B)_i \geq 0$, если $(x_B)_j = 0$, то дальше x_N увеличивать нельзя, x_{Nj} и x_{Bj} меняются местами.

Список литературы к 2–4 главам

1. Химмельблау Д. Прикладное нелинейное программирование. – М.: Мир, 1999.
2. Базара М., Шетти К. Нелинейное программирование. Теория и алгоритмы. – М.: Мир, 1996.
3. Поляк Б. Т. Введение в оптимизацию. – М.: Наука, 1998.
4. Моисеев Н. Н., Иванюков Ю. П., Столярова Е. М. Методы оптимизации. – М.: Наука, 1997.
5. Пшеничный Б. Н., Данилин Ю. М. Численные методы в экстремальных задачах. – М.: Наука, 2000.
6. Юдин Д. Б., Гольштейн Е. Г. Линейное программирование. Теория, методы и приложения. – М.: Наука, 2001.
7. Муртаф Ф. Современное линейное программирование. – М.: Мир, 1995.
8. Фиакко А., Мак-Кормик Дж. Нелинейное программирование: методы последовательной безусловной минимизации. – М.: Мир, 1992.
9. Васильев Ф. П. Численные методы решения экстремальных задач. – М.: Наука, 2001.
10. Растринин Л. А. Системы экспериментального управления. – М.: Наука, 1994.
11. Банди. Методы оптимизации. – М.: Радио и связь, 1998.
12. Моисеев Н. Н. Математические задачи системного анализа. – М.: Наука, 1997.
13. Ахназарова С.А., Кафаров В.В. Оптимизация эксперимента в химии и химической технологии. – М.: Высшая школа, 1998.
14. Денисов В.И., Попов А.А. Пакет программ оптимального планирования эксперимента. – М.: Финансы и статистика, 2001.

15. Бродский В.З. Введение в факторное планирование эксперимента. – М.: Наука, 2000.
16. Гордеев Л. С., Кафаров В. В., Бояринов А. И. Оптимизация процессов химической технологии. – М.: МХТИ, 1999.
17. Реклейтис Г., Рейвиндран А., Рэгсделк. Оптимизация в техники. В 2-х кн. – М.: Мир, 1998.
18. Уайлд Д. Д. Методы поиска оптимума. – М.: Мир, 1997.
19. Дамбраускас А. П. Симплексный поиск. – М.: Энергия, 2000.
20. Федоров В. В. Численные методы максимина. – М.: Наука, 1999.
21. Карманов В. Г. Математическое программирование. – М.: Наука, 1995.
22. Зангвилл У. Нелинейное программирование. Единый подход. – М.: Советское радио, 1993.
23. Федоренко Р. П. Приближенные решения задач оптимального управления. – М.: Наука, 1998.
24. Данилин Ю. М. Скорость сходимости методов сопряженных направлений. – Кибернетика, 1977, № 6.
25. Зойтендейк Г. Методы возможных направлений. – М.: Мир, 1993.
26. Белов Е. Н. Алгоритмы и программы решения задач квадратичного и линейного программирования. -Экономические и математические методы, 1998, т. 16, № 1.
27. Белов Е. Н., Поляк Б. Т., Скоков В. А. Комплекс программ оптимизации. – Экономические и математические методы. 1998, т. 14, № 4.
28. Федоров В. В. Теория оптимального эксперимента. – М.: Наука, 1991.
29. Романовский Н. В. Алгоритмы решения экспериментальных задач. – М.: Наука, 1997.
30. Мочалов С.П. Методы оптимизации металлургических процессов: учебное пособие. – Новокузнецк: Сибирская государственная горно- Металлургическая академия, 1997.
31. Вычислительные методы выбора оптимальных проектных решений / Под редакцией В. С. Михалевича. – Киев.: Наукова думка, 1997.
32. Евтушенко Ю. Г. Методы решения экстремальных задач и их применение в системах оптимизации. – М.: Наука, 1992.
33. Евтушенко Ю. Г., Жадан В. Г. К вопросу о систематизации численных методов нелинейного программирования. – М.: ВЦ АН РФ, 1998.
34. Веселов Е. Н., Евтушенко Ю. Г., Мазурик В. П. Диалоговый комплекс ДИСО – вычислительная станция прикладного математика // Вычислительные системы и программное обеспечение. - Н.: ВЦ СО АН РФ, 1996.
35. Мазурик В. П. Пакет прикладных программ «Администратор полей» // Пакеты прикладных программ. Программное обеспечение оптимизационных задач. – М.: Наука, 1997.
36. Рыков А.С. Поисковая оптимизация. Методы деформируемых конфигураций. – М.: Наука, 2003.

5. ВЕРОЯТНОСТНЫЕ МОДЕЛИ ОПТИМИЗАЦИИ РЕЖИМОВ ЭЛЕКТРОЭНЕРГЕТИЧЕСКИХ СИСТЕМ

5.1. Характеристика информации для расчетов установившихся режимов ЭЭС

Общая характеристика факторов, воздействующих на ЭЭС. ЭЭС непрерывно подвергается воздействию случайных факторов, которые можно разделить на следующие две группы [1, 3, 5, 8, 11]:

1) факторы, приводящие к аномальным ситуациям в ЭЭС:

- отказы оборудования;
- неправильная работа автоматики и защиты;
- стихийные бедствия (ветер, гроза, гололед);
- короткие замыкания;
- ошибки персонала;

2) факторы, не приводящие к нарушению нормального режима работы ЭЭС:

- случайный характер нагрузки;
- метеорологические факторы (температура, влажность, ветер);
- погрешности измерения исходных данных;
- погрешности математических моделей и ошибки значений параметров этих моделей;
- погрешности методов расчета (незавершенность итерационных процессов и ошибки округления);
- ошибки реализации решений при управлении режимами.

Традиционные расчеты режимов игнорировали случайные факторы и использовали в качестве исходных данных одну из возможных реализаций, которая с инженерной точки зрения казалась наиболее правдоподобной. При этом обнаруживалась проблема оценки погрешности результатов. Приходилось делать определенные запасы, допуски и резервы, что, как правило, определялось на основе опыта, интуиции и с перестраховкой.

Знание погрешностей результатов позволяет решить проблему запасов более обоснованно.

Учет случайных факторов нужен не только при проектировании или планировании работы ЭЭС, но и при оперативном диспетчерском управлении (в темпе процесса или в режиме реального времени).

Режимными параметрами ЭЭС являются мощности в узлах и ветвях схемы сети, модули и фазы напряжений и токов и др. Эти режимные параметры могут быть как исходными данными, так и подлежать расчету.

Параметры ЭЭС – это фактически параметры модели ЭЭС (сопротивления, проводимости, характеристики состояния и топология схемы сети).

Режимные параметры, являющиеся исходными данными, измеряются измерительными приборами, а параметры системы обычно не могут быть прямо измерены и подлежат вычислению через справочные и каталожные данные или идентификации.

Режимные параметры имеют погрешности измерения и передачи данных. Управляющие воздействия имеют погрешности прогноза, модели и ошибки вычислений. Управляющие воздействия отрабатываются с погрешностью, вызванной люфтами и ошибками персонала.

Погрешности нагрузок потребителей. Электроприемники потребителей включаются, изменяют свою мощность и отключаются случайным образом. Вместе с тем, цикл суточной деятельности большинства людей однотипен – примерно в одно и то же время начинают работать учреждения, промышленные предприятия и с наступлением сумерек включается освещение, и к полуночи бытовая нагрузка снижается. Все это обуславливает закономерность суточного графика нагрузки и позволяет прогнозировать ее с относительно высокой точностью [5, 6].

График нагрузки любого потребителя можно считать случайным процессом, т. е. функцией времени, значения которой являются случайными величинами, а параметры распределения зависят от времени.

Для получения характеристик случайного процесса необходимо иметь ряд его реализаций. Для отдельного потребителя получить ряд его реализаций за один и тот же период времени в принципе невозможно. Поэтому принимают гипотезу о том, что условия работы потребителя практически повторяются в каждом периоде (сутки, год).

Таким образом, выполняя наблюдения n периодов T , получаем n реализаций случайного процесса (ансамбль реализаций).

$$\begin{aligned} &P(t), \\ &P(t + T), \\ &P(t + 2T), \\ &\dots \\ &P[t + (n - 1)T] \end{aligned} \tag{5.1}$$

Период T выбирается исходя из таких закономерностей как:

- суточное вращение Земли;
- смена времен года;
- семидневная неделя с двумя выходными;
- восьми (семи) часовой рабочий день;

- метеорологические факторы.

Обычно в качестве периода выбирается цикл человеческой деятельности: год, месяц, неделя, сутки. Однако, вследствие наличия общей тенденции изменения нагрузки – тренда, ансамбль реализаций для нагрузки можно принять лишь условно. С большим основанием можно говорить о цикличности графика i -ой нагрузки $P_i(t)$, если из него вычесть среднюю за период нагрузку, рис. 5.1 а) и тренд, рис. 5.2 б), таким образом, получим график, изображенный на рис. 5.2 в).

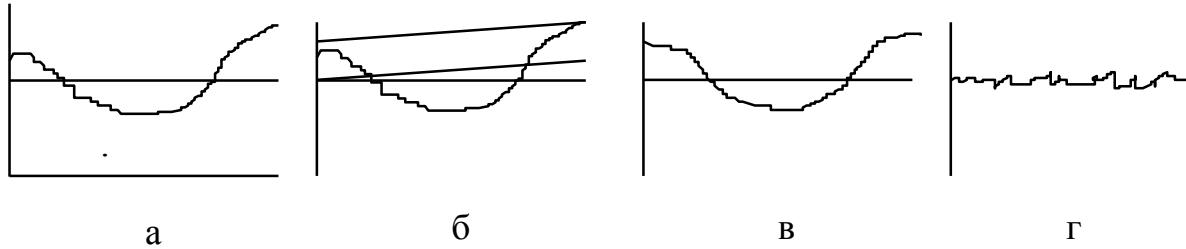


Рис. 5.1. Годовой график нагрузки и его составляющие

Часто линия тренда является прямой линией

$$P_{\text{icp}} + a_i t, \quad (5.2)$$

однако возможно наличие квадратичной составляющей, тогда

$$P_{\text{icp}} + a_i t + b_i t^2. \quad (5.3)$$

Если теперь из полученного графика вычесть периодическую составляющую $C_i \cos\left(\frac{1}{T}t - \theta\right)$, то в результате получим остаток, который изменяется случайным образом – чисто случайный процесс $\xi_i(t)$ с математическим ожиданием, равным нулю.

$$\xi_i(t) = P_i(t) - P_{\text{icp}} - a_i t - C_i \cos\left(\frac{1}{T}t - \theta\right). \quad (5.4)$$

То же самое можно проделать для T , равного суткам, неделе, месяцу.

Характеристики регулярных и случайной составляющих процесса изменения нагрузки используются для его прогнозирования.

Все регулярные составляющие процесса легко экстраполируются. Функция $\xi_i(t)$ используется для оценки ошибки прогноза, а также может содержать информацию для прогнозирования. Для этого может использоваться ковариационная функция случайного процесса $X(t)$

$$R_X(t, t + \tau) = \lim_{T \rightarrow \infty} \int_0^T \overset{\circ}{X}(t) \overset{\circ}{X}(t + \tau) dt, \quad (5.5)$$

где $\overset{\circ}{X}(t) = X(t) - m_X(t)$ – центрированный случайный процесс; $m_X(t)$ – математическое ожидание случайного процесса.

Корреляционная функция

$$r_X(t, t + \tau) = \frac{R_X(t, t + \tau)}{\sigma_X(t)\sigma_X(t + \tau)}. \quad (5.6)$$

Если математическое ожидание $m_X(t)$ и ковариационная функция $R_X(t, \tau)$ не зависят от времени, то случайный процесс $X(t)$ является стационарным. Использование же фактически одной реализации случайного процесса, наблюдаемой в течение большого промежутка времени, разбитой на периоды и представленной как ансамбль реализаций за выбранный период T , основано на свойстве эргодичности случайного процесса.

На множестве n реализаций случайного процесса

$$R_X^*(\tau) = \frac{1}{n} \sum_{i=1}^n x_i(t) x_i(t + \tau). \quad (5.7)$$

и

$$r_X^*(\tau) = \frac{R_X^*(\tau)}{D_X^*}, \quad (5.8)$$

где $D_X^* = R_X^*(0)$ – оценка дисперсия случайного процесса $X(t)$.

Прогнозируемое значение случайного процесса на интервал упреждения θ может быть получено по формуле

$$\hat{X}(t_0 + \theta) = m_X^* + r_X^*(\theta)[X(t_0) - m_X^*], \quad (5.9)$$

где $\hat{X}(t + \theta)$ – прогнозируемое значение случайного процесса; m_X^* – оценка математического ожидания случайного процесса.

С увеличением интервала упреждения θ значение корреляционной функции приближается к нулю, а прогнозируемое значение случайного процесса (5.9) приближается к математическому ожиданию.

Ошибки прогнозирования зависят от:

- характера нагрузок;
- времени упреждения;
- количества составляющих отдельных электроприемников в одной нагрузке (чем больше нагрузок входит в состав суммарной нагрузки, тем меньше относительная погрешность).

Если число потребителей, образующих суммарную нагрузку велико, то в соответствие с центральной предельной теоремой теории вероятностей распределение суммы случайных величин стремится к нормальному закону распределения.

Следует, однако, оговорить тот факт, данная теорема применима для случайных величин, среди которых нет таких, которые имеют слишком большую по отношению к другим дисперсию.

Ошибки измерения. Измерения неизбежно содержат ошибки. Измерению в ЭЭС подлежат:

- напряжения;
- активные и реактивные мощности;
- токи и др.

Ошибки возникают по всему тракту передачи измеряемой информации от трансформатора тока и напряжения до ЭВМ, куда эти данные вводятся для последующей обработки.

При считывании со щитовых приборов имеет место дополнительный источник ошибок – ошибки считывания, передачи по телефонной линии связи и ручного ввода в ЭВМ.

Существует три канала поступления данных:

- 1) телеизмерения;
- 2) диспетчерские отчетные документы (ведомости), формируемые в основном на основе сообщений по телефону от персонала энергообъектов, считывающего показания щитовых приборов;
- 3) контрольные замеры – массовые одновременные считывания и фиксация в документах показаний приборов в максимально возможном числе точек электрической сети, проводимые несколько раз в год.

Все ошибки измерений можно разделить на *статические* и *динамические*.

Статические ошибки обусловлены:

- ошибками измерительных приборов;
- квантованием сигналов (непрерывные преобразуются в дискретные);
- возникновением помех в каналах и т. п.

Обычно принимают *закон распределения ошибок нормальным*. При этом исходят из подтверждающегося практикой предположения о том, что малые ошибки появляются намного чаще, чем большие. Положительные и отрицательные ошибки (после устранения систематической составляющей) равновероятны – эти условия характерны для нормального закона.

Округленные данные погрешностей измерительных щитовых приборов приведены в табл. 5.1. Щитовые приборы используются персоналом для контрольных замеров.

Таблица 5.1

Статические ошибки измерений по щитовым приборам

Измеряемый параметр	Математическое ожидание ошибки, %	Среднеквадратичная ошибка, %
Напряжение 110 кВ	-1,0 ... 1,5	0,7 ... 1,2
Напряжение 220–500 кВ	-0,5 ... 1,5	0,7 ... 1,2
Активная мощность генераторов	-0,5	1,0 ... 12*
Реактивная мощность генераторов	-0,5	2,0 ... 15*
Переток активной мощности	-1,0	1,5 ... 70*
Переток реактивной мощности	-1,0	2,5 ... 100*
Ток	-1,0	0,8 ... 17*
Активная нагрузка	-1,0	1,5 ... 40*
Реактивная нагрузка	-6,0	15 ... 80*

* Большие погрешности появляются при измерении малых перетоков, когда приборы работают в начале шкалы.

При телеизмерениях, вводимых в ЭВМ, человек как передаточное звено исключается. В настоящее время энергосистемы оборудованы телемеханическими устройствами дискретного типа. В них аналоговый непрерывный сигнал

на выходе измерительного устройства преобразуется аналого-цифровым преобразователем (АЦП) в дискретный вид, передача которого позволяет обеспечить надежную защиту от случайных помех. При этом точность измерения лимитируется числом квантов, на которые делится диапазон измеряемой величины.

Так, если переток мощности измеряется датчиком с диапазоном ± 500 МВт и АЦП допускает преобразование в 100 квантов, то цена одного кванта есть $1000/100 = 10$ МВт. Возникает погрешность ± 10 МВт, например сигналы 325 и 334,7 будут иметь одинаковый цифровой код, соответствующий 330 МВт.

Поскольку появление значения измеряемого параметра в диапазоне кванта можно считать равновероятным, ошибки измерений, связанные с преобразованием аналогового сигнала в дискретный можно считать распределенными по *равномерному закону*.

Используемые в энергетике телеизмерительные системы имеют класс точности не ниже 1,0, причем здесь уже учтена погрешность измерительных трансформаторов, датчиков и АЦП. Эта погрешность отнесена к максимальному значению измеряемой величины (рис. 5.2), и поэтому измеренные значения внутри диапазона измерений могут иметь большую погрешность [7, 9].



Говоря о статических ошибках измерений, следует иметь в виду, что между ошибками отдельных измерений может иметь место корреляционная связь (используются одни и те же измерительные трансформаторы).

Динамические погрешности. Между моментом времени измерения и моментом времени использования данных проходит время τ . Режим ЭЭС все время меняется и измеренное значение $x(t)$ и текущее значение измеряемого параметра режима $x(t + \tau)$ уже будут разными

$$\Delta x = x(t + \tau) - x(t). \quad (5.10)$$

Динамическая ошибка Δx зависит от:

- степени изменчивости измеряемого параметра x (для медленно изменяющихся параметров она будет меньше, чем для быстроизменяющихся при одном и том же τ);
- времени запаздывания τ , которое определяется неодновременностью проведения измерений и ввода информации в ЭВМ, а также временем между

моментом получения измеряемых данных и моментом реализации решения, принятого на основе этих данных.

Для анализа динамической погрешности удобно использовать ковариационную функцию. В самом деле, чем более статистически взаимосвязаны значения x в момент измерения и в момент принятия решения, тем меньше динамическая погрешность.

Если считать, что за время τ математическое ожидание x не изменилось, то дисперсия динамической ошибки Δx_d .

$$\begin{aligned} D_{\Delta x} &= M[(\Delta x_d)^2] = M\left\{[\dot{X}(t+\tau) - \dot{X}(t)]^2\right\} = \\ &= M[\dot{X}(t+\tau)^2] - 2M[\dot{X}(t+\tau)\dot{X}(t)] + \\ &+ M[\dot{X}(t)^2] = 2\sigma_x^2 - 2K_x(\tau). \end{aligned} \quad (5.11)$$

Для современных систем телеизмерений, у которых цикл опроса датчиков составляет 1–10 с, динамическая среднеквадратическая ошибка составляет:

- для напряжений 0,2–0,7 %;
- для перетоков активной мощности 0,5–30 %;
- для перетоков реактивной мощности 1–60 %.

Большие значения соответствуют быстроизменяющимся величинам, когда измеряемое значение находится в начальной части диапазона проводимых измерений.

По данным некоторых исследований средние динамические погрешности телеизмерений в 2–5 раз выше статических ошибок.

В последнее время в пунктах сбора информации стала производиться первичная обработка данных по более сложным алгоритмам (с учетом производных $x(t)$) и таким образом снижать данную погрешность.

Грубые ошибки измерений. Помимо ошибок измерений, связанных с нормальным функционированием системы сбора данных, возникают ошибки, вызванные отказами аппаратуры, редко встречающимися помехами. Такие ошибки, как правило, значительно выше описанных шумов, но появляются гораздо реже. Измерения, содержащие такие ошибки, часто называют сбоями, выбросами, грубыми ошибками, плохими данными. Распределение вероятностей появления сбоев $P(m)$ можно описать законом Пуассона

$$P(m) = \frac{a^m e^{-a}}{m!}, \quad (5.12)$$

а распределение амплитуд этих грубых ошибок $\Delta_{сб}$ обычно принимают нормальным с дисперсией, большей, чем у обычных ошибок.

Грубые ошибки измерений, как правило, не коррелированы между собой.

Полная ошибка измерений

$$\Delta_{\Sigma} = \Delta x(1 - p_{сб}) + \Delta_{сб} p_{сб}.$$

Погрешности моделей ЭЭС. Погрешности моделей ЭЭС, используемых при расчетах режимов, могут возникать в силу двух причин:

- 1) расчетная схема содержит эквиваленты, например узел является эквивалентом нескольких нагрузок вместе со связывающей их сетью;
- 2) параметры элементов расчетной схемы имеют погрешности из-за:
 - погрешностей паспортных данных;
 - погрешностей экспериментальных данных;
 - ошибок в расчете;
 - старения и модернизации оборудования, а также неучета некоторых физических свойств объектов, таких как:
 - распределенности параметров ЛЭП;
 - поверхностного эффекта в проводниках;
 - влияния земли и заземленных тросов;
 - неоднородности поверхности земли и снежного покрова.

Кроме того, имеет место влияние метеорологических факторов (температура, скорость ветра, влажность и т. п.).

Так, например, для воздушных линий (ВЛ) погрешность в определении активного сопротивления R может составить:

- ± 1 % из-за погрешности в расчете длины трасы;
- $\pm 1,5 \dots 3,0$ % из-за неучета поверхностного эффекта;
- $-20 \dots +16$ % из-за неучета влияния температуры.

Для индуктивного сопротивления ВЛ:

- ± 1 % из-за погрешности в расчете длины трасы;
- $0,25$ % из-за неточности и неоднородности вдоль трассы геометрии проводов;
- $1 \dots 3$ % из-за неучета влияния заземленных тросов;
- $-(4 \dots 8)$ из-за неучета параллельных цепей.

Для трансформаторов погрешности в определении их сопротивлений и проводимостей холостого хода вследствие работы РПН может достигать ± 10 %, ± 15 %, ± 30 %. Погрешность в задании коэффициента трансформации может достигать величины, соответствующей одной ступени регулирования (2 %).

Значения погрешностей параметров математических моделей ВЛ и трансформаторов приведены в табл. 5.2.

Данные о распределении вероятностей ошибок параметров схем замещения в литературе отсутствуют.

Таблица 5.2

**Значения погрешностей параметров математических моделей ЛЭП
и трансформаторов, %**

Погрешности	Линии электропередачи				Трансформаторы			
	R	X	G	B	R	X	G	B
Диапазон изменения	-16–20	-6–3	-20–40	-4–20	-16–20	-15–20	-12–24	-15–45
Систематическая ошибка	2	-1,5	10	8	2	0	6	6
Среднеквадратическое отклонение	6	1,5	10	4	6	5	5	10

При выполнении оптимизационных расчетов весьма важна информация о расходных характеристиках электростанций $B = f(P)$ – зависимостей расхода топлив в единицу времени B от генерирующей мощности P и характеристик относительных приростов (ХОП) $\frac{\partial B(P)}{\partial P}$. Исследования в этой области показали, что погрешности ХОП не выше 5–6 %.

Погрешности реализации управления. Конечной целью расчета режима ЭЭС является принятие тех или иных решений – задание суточных графиков выработки электростанций, графиков поддержания уровней напряжения в контрольных точках, планов и графиков ремонта оборудования, выбор уставок автоматики и релейной защиты. Однако реализация решений всегда будет отличаться от плановых в силу ряда причин, таких, как:

- «люфты» в регуляторах скорости и возбуждения, отрабатывающие запланированные графики параметров режима;
- дискретность возможных значений параметров (РПН, уставки);
- неточность контроля за реализацией графиков из-за погрешностей самих регулирующих приборов и погрешностей считывания их показаний;
- недостаточная дисциплина диспетчерского персонала.

Указанные причины приводят к отклонению реальных параметров режима от заданных.

Известны результаты анализа отклонений сумм генерируемых активных мощностей в энергосистемах (табл. 5.5).

Таблица 5.3

Параметры распределения отклонений суммарных активных мощностей

Исследуемый параметр	m , %	σ , %	s	ex
Отклонение мощности ЕЭС	-1,5–0,5	1,5–3,0	-1,0–2,0	0,5–1,5
Отклонение генерируемых мощностей ОЭС	-2,0–1,0	2,0–5,0	-2,0–2,0	-1,0–5,0
Отклонение генерируемых мощностей ГЭС	-2,0–2,0	10–20	-1,0–2,0	0–6,0

Ни для одной из величин гипотеза о нормальном или равномерном законе не подтвердилась.

Кроме того, как правило, при плановых значениях, близких к минимальным, электростанции работают с большей мощностью, чем получено в расчетах, а при значениях, близких к максимальным, — с меньшей, т. е. имеется тенденция удержания мощностей станции подальше от границ допустимой области.

В приведенных результатах отклонение реализованных управлений от плановых включало и отклонения вследствие отличия нагрузок от прогнозируемых значений.

5.2. Расчет установившихся режимов ЭЭС в вероятностной постановке

Модели установившихся режимов в представлены детерминированной постановке. Схема ЭЭС состоит из узлов и ветвей. Ветви разделяются на продольные и поперечные [2, 6, 10, 11].

Модель узла. Каждый узел связан с хотя бы одним другим узлом схемы сети посредством продольных ветвей. узел описывается мощностью, проводимостью поперечной ветви и напряжением, рис. 5.3.

Для узла i использованы следующие обозначения: P_{gi} — активная мощность генерации; Q_{gi} — реактивная мощность генерации; P_{li} — активная мощность нагрузки; Q_{li} — реактивная мощность нагрузки; Y_{Ni} — проводимость поперечной ветви; U_i, δ_i — модуль и фаза напряжения; ω_i — множество номеров узлов, смежных i -у.

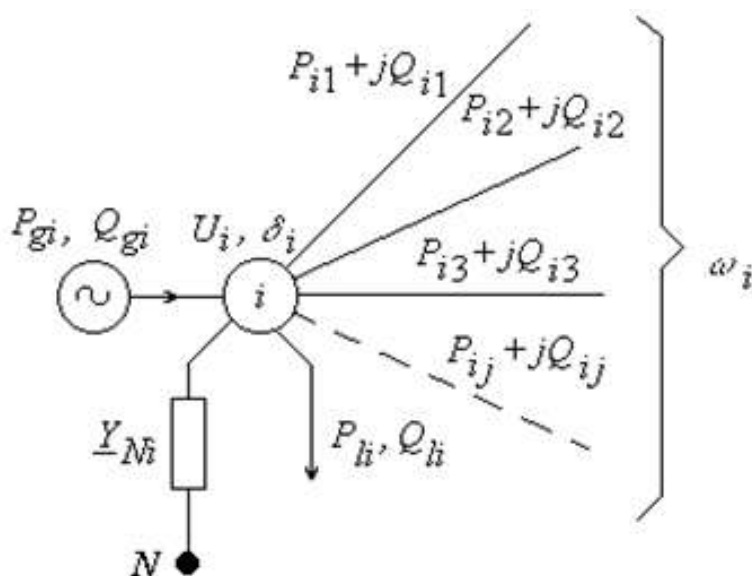


Рис. 5.3. Модель узла схемы сети

$$\begin{aligned}
\underline{U}_i &= U'_i + jU''_i, \\
U'_i &= U_i \cos \delta_i, \\
U''_i &= U_i \sin \delta_i, \\
U_i &= \sqrt{U'^2_i + U''^2_i}, \\
\delta_i &= \operatorname{arctg} \frac{U''_i}{U'_i}.
\end{aligned} \tag{5.13}$$

Мощность нагрузки задается постоянной величиной или статическими характеристиками:

$$\begin{aligned}
P_{li} &= P_{0i}(U_i, f), \\
Q_{li} &= Q_{0i}(U_i, f).
\end{aligned} \tag{5.14}$$

Обычно статические характеристики представляются в виде полиномов отдельно по напряжению и по частоте:

$$\begin{aligned}
P_l &= P_{l0} \left(1 + a_1 \frac{U - U_0}{U_0} + a_2 \left(\frac{U - U_0}{U_0} \right)^2 \right), \\
Q_l &= Q_{l0} \left(1 + b_1 \frac{U - U_0}{U_0} + b_2 \left(\frac{U - U_0}{U_0} \right)^2 \right).
\end{aligned} \tag{5.15}$$

где P_{l0} и Q_{l0} – нагрузки при номинальном напряжении U_0 ; a_1, a_2, b_1, b_2 – коэффициенты, определяемые экспериментальным путем.

В ряде случаев получают обобщенные типовые статические характеристики. Зависимость $P_l = P(U)$ почти линейна, при этом $\frac{\partial P}{\partial U} \approx a_1 \approx 0,7 \dots 0,9$. $Q_l = Q(U)$ – имеет явно нелинейный характер.

Статические характеристики нагрузки по частоте считаются линейными:

$$\begin{aligned}
P_l &= P_0 \left(1 + c \left(\frac{f - f_0}{f_0} \right) \right), \\
Q_l &= Q_0 \left(1 + d \left(\frac{f - f_0}{f_0} \right) \right).
\end{aligned} \tag{5.16}$$

где $c = 1,5$ для промышленных потребителей; $c = 1,2$ для смешанных потребителей.

Генерируемая мощность также может быть задана статическими характеристиками:

$$\begin{aligned}
P_g &= P_{go} \left(1 + k \left(\frac{f - f_0}{f_0} \right) \right) \left(1 + g \left(\frac{U - U_0}{U_0} \right) \right), \\
Q_g &= Q_{go} \left(1 + m \left(\frac{f - f_0}{f_0} \right) \right) \left(1 + l \left(\frac{U - U_0}{U_0} \right) \right).
\end{aligned} \tag{5.17}$$

Коэффициенты k, g, m, l – устанавливаются при настройке сетевых регуляторов.

Параметры режима в узле связаны между собой по первому закону Кирхгофа:

$$\begin{aligned}
\sum_{j \in \omega_i} P_{i,j} - P_{gi} + P_{li} + U_i^2 g_i &= 0, \\
\sum_{j \in \omega_i} Q_{i,j} - Q_{gi} + Q_{li} + U_i^2 b_i &= 0.
\end{aligned} \tag{5.18}$$

или

$$\sum_{j \in \omega_i} \underline{S}_{ij} - \underline{S}_{gi} + \underline{S}_{li} + U_i^2 \underline{Y}_i = 0. \tag{5.19}$$

Модель ветви. Продольные ветви являются элементами схем замещения ЛЭП или трансформатора. ЛЭП – моделируется симметричной П-образной схемой замещения, а трансформаторы – несимметричной, рис. 5.4.

Модель ЛЭП:

$$\underline{Z} = \underline{Z}_C \text{sh}(\underline{\gamma}_0 l), \tag{5.20}$$

$$\underline{Y}_1 = \underline{Y}_2 + \frac{\text{ch}(\underline{\gamma}_0 l) - 1}{\underline{Z}_C \text{sh}(\underline{\gamma}_0 l)} = \frac{1}{\underline{Z}_C} \text{th} \left(\frac{\underline{\gamma}_0 l}{2} \right). \tag{5.21}$$

Обычно $\underline{Z} = r_0 l + jx_0 l$ и $\underline{Y}_1 = \underline{Y}_2 = \frac{(g_0 + jb_0)l}{2}$.

$$\begin{aligned}
\underline{Z} &= \frac{\underline{Z}_T}{a}, \\
\underline{Y}_1 &= \frac{1}{\underline{Z}_T} (1 - a) + \underline{Y}_\mu, \\
\underline{Y}_2 &= \frac{a}{\underline{Z}_T} (a - 1).
\end{aligned} \tag{5.22}$$

при $a > 5$.

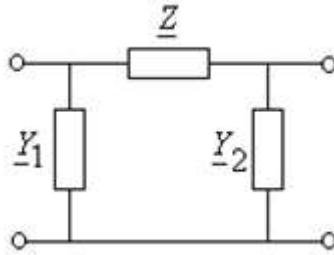


Рис. 5.4. Модель ветви

Модель сети представляет собой узловые уравнения вида (5.18) или (5.19), в которых потоки мощности по ветвям выражены через узловые напряжения и элементы матрицы узловых проводимостей, а алгебраическая сумма мощности нагрузки и мощности генерации есть суммарная мощность узла $\underline{S} = P + jQ$,

$$\underline{U}_i \left(\sum_{j \in \omega_i} \underline{Y}_{i,j}^* \underline{U}_j^* - \underline{Y}_{i0}^* \underline{U}_0^* \right) = \underline{S}_i. \quad (5.23)$$

при $i = 1 \dots n$, где n – число независимых узлов схемы сети.

В некоторых задачах кроме уравнений узловых напряжений используются уравнения для токов узлов или ветвей, модулей и фаз напряжений, а также потоков мощности по ветвям схемы сети.

Вектор состояния и базис установившегося режима. Совокупность всех режимных параметров ЭЭС, а также мощности нагрузок и генерации называется *вектором параметров режима*. Обозначим его через \mathbf{z} .

Совокупность уравнений, связывающих все параметры режима, образует общую систему уравнений установившегося режима

$$\mathbf{g}(\mathbf{z}) = 0. \quad (5.24)$$

Некоторые из этих уравнений могут быть взаимосвязанными, т. е. одни уравнения получаются из других. Поэтому некоторые уравнения могут быть исключены – выделяется *система независимых уравнений*. Число независимых уравнений равно удвоенному числу независимых узлов сети, которое равно числу всех узлов без единицы $N - 1$.

В традиционной постановке заданы мощности всех узлов кроме одного – базисного балансирующего узла $[(N - 1) \text{ узла}]$.

Введем понятие *вектора состояния*. Под вектором состояния \mathbf{x} будем понимать такой набор параметров режима, знание которого обеспечивает возможность дорассчитать все остальные параметры режима. Рассмотрим схему электрической сети, состоящую из шести узлов, рис. 5.5.

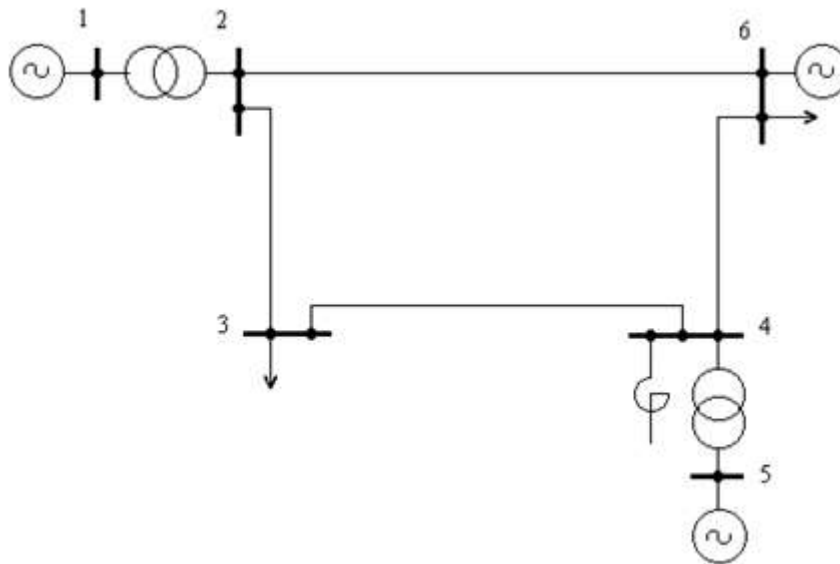


Рис. 5.5. Шестиузловая схема электрической сети

В этой схеме общее число узлов $N = 6$, а число независимых узлов $N - 1 = 5$.

В традиционной постановке задаются модуль и фаза базисного балансирующего узла – узел 1 и мощности остальных пяти узлов $\mathbf{x} = \{U_1 \delta_1 P_2 Q_2 P_3 Q_3 P_4 Q_4 P_5 Q_5 P_6 Q_6\}$ – 12 параметров. Всего же можно указать 24 узловых параметра: 6 модулей напряжений, 6 фаз напряжений, 6 активных и 6 реактивных мощностей узлов. Для расчета в традиционной постановке остальных 12 параметров необходимо решить систему из 10 уравнений узловых напряжений $2(N - 1) = 10$, отыскав таким образом 5 модулей и 5 фаз напряжений узлов и из баланса мощности в схеме сети, найти мощность балансирующего узла – оставшиеся два параметра P_1 и Q_5 .

Если базисный и балансирующий узлы не совпадают, то, например, в балансирующем узле вообще может быть ничего не известно: ни мощности ни напряжение и его фаза. В то же время в каком-либо другом узле могут быть заданы все узловые параметры: U, δ, P, Q .

В табл. 5.4 приведены несколько вариантов вектора состояния для шестиузловой схемы, включая традиционный вариант.

Таблица 5.4

Примеры вектора состояния	1	2	3	4	5	6	Требуется решение уравнений?
Традиционный	$U_1 \delta_1$	$P_2 Q_2$	$P_3 Q_3$	$P_4 Q_4$	$P_5 Q_5$	$P_6 Q_6$	Да
Вариант 2	$U_1 \delta_1$	$U_2 \delta_2$	$U_3 \delta_3$	$U_4 \delta_4$	$U_5 \delta_5$	$U_6 \delta_6$	Нет
Вариант 3	$U_2 \delta_2$	$P_2 Q_2$	$U_4 \delta_4$	$P_{2,3} Q_{2,3}$	$P_{4,6} Q_{4,6}$	$P_4 Q_4$	Нет

Задание вектора состояния набором модулей и фаз напряжений – не единственный способ обойтись без решения системы уравнений при расчете режима. Вариант 3 вектора состояния для шестиузловой схемы также позволяет

обойтись без решения системы уравнений путем последовательного расчета одного неизвестного параметра через заданные и уже найденные.

В случае, когда мощности некоторых узлов не заданы, а подлежат вычислению (подбору) так, чтобы достичь предельного или заданного значения некоторой функции, называемой целевой, существуют некоторые степени свободы в расчете режима. Иными словами решений (без целевой функции) бесконечно много. Целевая функция вводит условия, которые добавляют некоторое количество уравнений в систему независимых уравнений. Вид уравнений зависит от метода решения.

Число степеней свободы:

$$p = n - q, \quad (5.25)$$

где n – число независимых уравнений $n = 2(N - 1)$; q – число заданных параметров режима (меньшее по сравнению с вектором состояния).

Обозначим через \mathbf{d} – вектор заданных параметров режима, через \mathbf{y} – вектор переменных, образующих степени свободы, а через \mathbf{s} – параметры режима, которые могут быть вычислены по \mathbf{d} и \mathbf{y} из уравнений установившегося режима:

$$\mathbf{w}(\mathbf{s}, \mathbf{y}, \mathbf{d}) = 0. \quad (5.26)$$

Здесь \mathbf{w} вектор-функция уравнений установившегося режима.

Компоненты вектора состояния \mathbf{x} могут быть компонентами вектора \mathbf{d} , компонентами \mathbf{y} , а также компонентами вектора \mathbf{s} , т. е. дорасчитываться по заданным \mathbf{d} и \mathbf{y} . Например, модули и фазу напряжений в некоторых узлах могут быть заданы или дорасчитываться и могут входить в вектор состояния.

В состав вектора \mathbf{y} можно ввести и коэффициенты трансформации, емкости регулируемых батарей конденсаторов, сопротивления отключаемых цепей, реакторов и т. п.

В качестве критерия L может выступать мощность потерь в сети, а на вектор \mathbf{z} накладываются ограничения

$$\mathbf{z}_{\min} < \mathbf{z} < \mathbf{z}_{\max}. \quad (5.27)$$

На рис. 5.6 приведена схема сети из трех узлов, для которой можно дать несколько вариантов состава векторов \mathbf{d} , \mathbf{x} , \mathbf{y} .

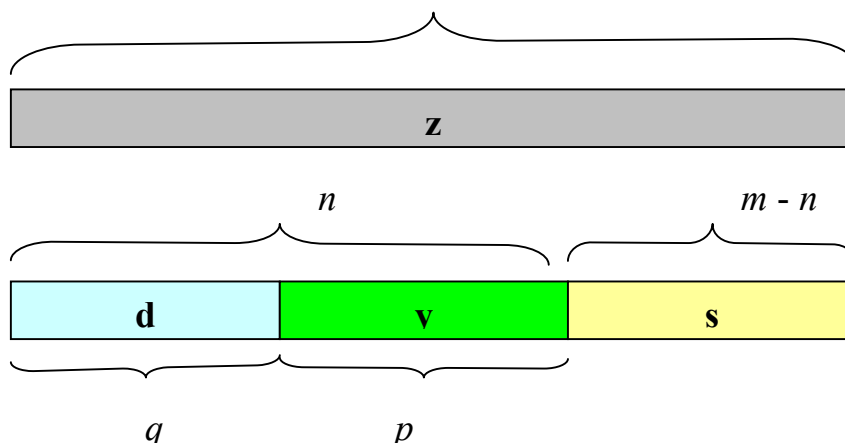


Рис. 5.6. Общая структура разделения переменных по векторам

Вариант 5. $\mathbf{d} = \{P_2, Q_2, P_3, Q_3\}$; $\mathbf{x} = \{U_2, \delta_2, U_3, \delta_3\}$; $\mathbf{y} = \{U_1\}$.

Составьте два других варианта указанных векторов.

Вариант 5. _____

Вариант 5. _____

Возможна ситуация, когда незаданность параметров режима определяется не наличием степеней свободы, а отсутствием информации, допустим, просто нет измерений нагрузок. В этом случае часть параметров не может быть дорассчитана по известным данным, и сеть становится ненаблюдаемой. Так в трехузловой схеме сети, рис. 5.7, если заданы только $U_2, \delta_2, U_3, \delta_3$, то невозможно определить P_3, Q_3, U_3, δ_5 .

Отсутствие возможности рассчитать параметры означает, что теряет смысл сохранение в расчетной схеме соответствующей части сети. В примере трехузловой схемы сети схема будет состоять из двух узлов, а в узле 2 расчетная мощность узла равна сумме нагрузки узла 2 и перетока мощности узлов 2 – 5.

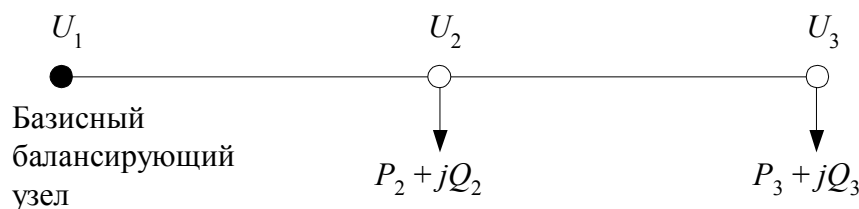


Рис. 5.7. Схема сети из трех узлов

Определение. Минимальное количество данных о параметрах режима, необходимое для однозначного определения всех его остальных параметров, называется в теории наблюдаемости *базисом*.

Любой базис может быть вектором состояния.

Любой вектор состояния является базисом.

Имеет смысл различать эти понятия, т. к. понятие вектора состояния связано с координатами решения задачи, т. е. вектор состояния является чаще вектором неизвестных, который вычисляется по исходным данным. Базис – это фиксированные параметры режима, задаваемые обычно в качестве исходных данных

Поскольку все параметры режима, в том числе и базис, являются функциями вектора состояния, уравнения установившегося режима можно представить в виде:

$$\bar{\mathbf{z}}_b - \mathbf{z}_b(\mathbf{x}_2) = 0, \quad (5.28)$$

где $\bar{\mathbf{z}}_b$ – заданное значение базиса; $\mathbf{z}_b(\mathbf{x}_2)$ – известные функции вектора базисных переменных от неизвестных компонент вектора состояния \mathbf{x}_2 ; $\mathbf{x}_2 \in \mathbf{x}$.

Например, при традиционной постановке

$$\begin{aligned}
\bar{P}_i - P_i(\mathbf{U}, \boldsymbol{\delta}) &= 0, \\
\bar{Q}_i - Q_i(\mathbf{U}, \boldsymbol{\delta}) &= 0, \\
P_i(\mathbf{U}, \boldsymbol{\delta}) &= \sum_{j \in \omega_i} P_{i,j}, \\
Q_i(\mathbf{U}, \boldsymbol{\delta}) &= \sum_{j \in \omega_i} Q_{i,j}.
\end{aligned} \tag{5.29}$$

В случае, если известны перетоки мощности или напряжения, то используются уравнения:

$$\begin{aligned}
\bar{P}_{i,j} - P_{i,j}(U_i, \delta_i, U_j, \delta_j) &= 0, \\
\bar{Q}_{i,j} - Q_{i,j}(U_i, \delta_i, U_j, \delta_j) &= 0, \\
\bar{U}_i - U_i &= 0.
\end{aligned} \tag{5.30}$$

Если в качестве исходных данных выступает набор параметров, число компонент которого больше базисного, то возникающая при этом система уравнений вида (5.16) является переопределенной, т. е. число уравнений больше числа неизвестных. Если выделить из всего вектора \mathbf{z} базисный состав \mathbf{z}_b , то

$$\begin{aligned}
\bar{\mathbf{z}}_b - \mathbf{z}_b(\mathbf{x}_2) &= 0, \\
\bar{\mathbf{z}}_l - \mathbf{z}_l(\mathbf{x}_2) &= 0,
\end{aligned} \tag{5.31}$$

Так как \mathbf{x}_2 вычисляется из $\bar{\mathbf{z}}_b - \mathbf{z}_b(\mathbf{x}_2) = 0$, то $\bar{\mathbf{z}}_l - \mathbf{z}_l(\mathbf{x}_2) \neq 0$ — имеет место несогласованность задания исходных данных. Систему уравнений $\bar{\mathbf{z}}_l - \mathbf{z}_l(\mathbf{x}_2) = 0$ называют подсистемой контрольных уравнений, по которой контролируется степень согласованности задания исходных данных. Эти уравнения имеют большое значение для достоверизации исходных данных.

Модели установившихся режимов в вероятностной постановке. Неопределенность режима обусловлена случайным характером исходных данных.

Пусть заданы вектор математических ожиданий $M[\mathbf{d}] = \mathbf{m}_d$ и матрица ковариаций $\text{cov}(\mathbf{d}) = M[(\mathbf{d} - \mathbf{m}_d)(\mathbf{d} - \mathbf{m}_d)^T]$ вектора исходных данных. Требуется найти математические ожидания и ковариационные матрицы параметров режима электрической сети.

В случае линейной задачи матричная система уравнений

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{5.32}$$

в вероятностной постановке связывает вектор известных \mathbf{y} и вектор неизвестных \mathbf{x} случайных величин. Для этой системы уравнений искомые числовые характеристики вектора \mathbf{y} могут быть получены по уравнениям:

$$\begin{aligned}
\mathbf{m}_y &= \mathbf{A}\mathbf{m}_x, \\
\text{cov}(\mathbf{y}) &= \mathbf{A} \text{cov}(\mathbf{x}) \mathbf{A}^T.
\end{aligned} \tag{5.33}$$

Решение этой системы уравнений имеет вид:

$$\begin{aligned}
\mathbf{m}_x &= \mathbf{A}^{-1} \mathbf{m}_y, \\
\text{cov}(\mathbf{x}) &= \mathbf{A}^{-1} \text{cov}(\mathbf{y}) (\mathbf{A}^T)^{-1}
\end{aligned} \tag{5.34}$$

Ситуация осложняется, если вместо линейного соотношения имеет место нелинейное $Y = \varphi(X)$, рис 5.8.

Математическое ожидание случайной величины X , будет преобразовано в моду случайной величины Y , а математическое ожидание для Y будет смещено вверх (рис. 5.8). Чем больше среднеквадратическое отклонение Y , тем сильнее разница m_y и моды Y .

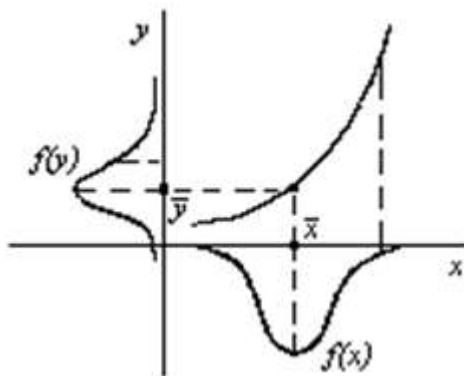


Рис. 5.8. Нелинейное преобразование случайных величин

Уравнения установившегося режима являются нелинейными и поэтому необходимы специальные методы получения законов распределения или числовых характеристик неизвестных параметров режима.

Методы расчета установившихся режимов в вероятностной постановке. Для расчета числовых характеристик параметров установившихся режимов нашли применения два метода [1–3, 6, 9]:

- 1) метод статистических испытаний (метод Монте-Карло);
- 2) метод статистической линеаризации.

Метод статистических испытаний. Метод статистических испытаний является универсальным методом, позволяющим моделировать на ЭВМ любой процесс. С помощью датчика (генератора) случайных чисел (ДСЧ) получают серии значений случайных величин (псевдослучайные числа), подчиненных заданному закону распределения. Каждое из полученных значений участвует в одном расчете режима – испытании. Таких испытаний проводится достаточно много и после этого проводится статистическая обработка полученных результатов.

Пусть необходимо определить функцию распределения случайной величины C , ее математическое ожидание μ и дисперсию σ^2 .

$$C = \varphi(x_1, x_2, \dots, x_n). \quad (5.35)$$

Для каждой из случайных величин x_1, x_2, \dots, x_n проводится получение случайного значения при соответствующем законе распределения с параметрами μ_i, σ_i^2 и находится $x_i^{(1)}, (i = 1, 2, \dots, n)$.

Вычисляется частное значение $C^{(1)}$.

Пункты 1 и 2 повторяются до тех пор, пока не будет выполнено N испытаний, т. е. получено множество $\{C^{(1)}, C^{(2)}, \dots, C^{(N)}\}$.

Производится статистическая обработка и находятся плотность (функция) распределения и числовые характеристики случайной величины C .

Таким образом, требуется N раз рассчитать установившийся режим. Обычно $N = 100-1000$. При этом расчет больших схем электрических сетей идет достаточно большое время.

Метод статистической линеаризации. Система нелинейных уравнений установившегося режима записывается для математических ожиданий случайных величин, а для вычисления ковариационных матриц линеаризуется в точке математических ожиданий случайных величин и задача сводится к линейной

$$\mathbf{w}(\mathbf{s}, \mathbf{d}) \approx \mathbf{w}_1(\mathbf{s}, \mathbf{d}) = \mathbf{w}_0 + \mathbf{C}_s(\mathbf{s} - \mathbf{m}_s) + \mathbf{C}_d(\mathbf{d} - \mathbf{m}_d) = 0 \quad (5.36)$$

или с обозначениями для центрированных случайных величин

$$\overset{\circ}{\mathbf{w}}_1(\mathbf{s}, \mathbf{d}) = \mathbf{C}_s \overset{\circ}{\mathbf{s}} + \mathbf{C}_d \overset{\circ}{\mathbf{d}}, \quad (5.37)$$

откуда

$$\overset{\circ}{\mathbf{s}} = \mathbf{C}_s^{-1} \left[\overset{\circ}{\mathbf{w}}_1 - \mathbf{C}_d \overset{\circ}{\mathbf{d}} \right]. \quad (5.38)$$

Линейное выражение (5.38) используется для расчета ковариационной матрицы $\text{cov}(\mathbf{s})$.

$$\begin{aligned} \text{cov}(\mathbf{s}) &= M \left[\overset{\circ}{\mathbf{s}} \cdot \overset{\circ}{\mathbf{s}}^T \right] = M \left\{ \mathbf{C}_s^{-1} \left[\overset{\circ}{\mathbf{w}}_1 - \mathbf{C}_d \overset{\circ}{\mathbf{d}} \right] \cdot \left[\overset{\circ}{\mathbf{w}}_1 - \mathbf{C}_d \overset{\circ}{\mathbf{d}} \right]^T (\mathbf{C}_s^{-1})^T \right\} = \\ &= \mathbf{C}_s^{-1} \left\{ \text{cov}(\mathbf{w}) + \mathbf{C}_d \text{cov}(\mathbf{d}) \mathbf{C}_d^T - \mathbf{C}_d \text{cov}(\mathbf{d}, \mathbf{w}) - \text{cov}(\mathbf{w}, \mathbf{d}) \mathbf{C}_d^T \right\} (\mathbf{C}_s^{-1})^T. \end{aligned} \quad (5.39)$$

Совместно с системой (5.39) решается система уравнений, записанная для математических ожиданий

$$\mathbf{w}(\mathbf{m}_s, \mathbf{m}_d, \text{cov}(\mathbf{s}), \text{cov}(\mathbf{d}), \text{cov}(\mathbf{s}, \mathbf{d})) = 0. \quad (5.40)$$

В некоторых случаях упрощенно можно использовать уравнения:

$$\mathbf{w}(\mathbf{m}_s, \mathbf{m}_d) = 0. \quad (5.41)$$

Вычисления числовых характеристик некоторых параметров режимов.

Пример 1. Дисперсия модуля напряжения при известных случайных продольной и поперечной составляющих напряжения:

$$\begin{aligned}
U &= \sqrt{U'^2 + U''^2}, \\
D[U] &= \left(\frac{\partial U}{\partial U'} \right)^2 D[U'] + \left(\frac{\partial U}{\partial U''} \right)^2 D[U''] + 2 \left(\frac{\partial U}{\partial U'} \right) \left(\frac{\partial U}{\partial U''} \right) \text{cov}(U', U''), \\
\left(\frac{\partial U}{\partial U'} \right) &= -\frac{U'}{\sqrt{U'^2 + U''^2}}, \\
\left(\frac{\partial U}{\partial U''} \right) &= -\frac{U''}{\sqrt{U'^2 + U''^2}}, \\
D[U] &= \frac{1}{U'^2 + U''^2} \{ U'^2 D[U'] + U''^2 D[U''] + 2U'U'' \text{cov}[U', U''] \}.
\end{aligned} \tag{5.42}$$

$$D[U] = \frac{1}{U'^2 + U''^2} \{ U'^2 D[U'] + U''^2 D[U''] + 2U'U'' \text{cov}[U', U''] \}. \tag{5.43}$$

где под U' и U'' понимаются их математические ожидания.

Выполним расчет режима простой сети, рис. 5.9, и найдем числовые характеристики модуля напряжения на шинах нагрузки. Будем считать, что случайными величинами являются только активная и реактивная мощности нагрузки, а параметры сети являются неслучайными величинами. Для простоты пренебрежем мощностью холостого хода трансформаторов и зарядной мощностью ЛЭП. Расчет выполним в системе Mathcad [4].

Одноцепная линия 110 кВ длиной 30 км, рис. 5.9, выполнена проводами марки АС-120/19. Трансформатор типа ТДЦН-40000/110.

Все параметры ЛЭП и трансформатора, а также известные параметры режима, приведены ниже.



Рис. 5.9. Схема электрической сети для расчета

Расчетные соотношения:

1) Математические ожидания модуля напряжения и его вещественной и мнимой составляющих находятся обычным способом, как при детерминированных данных.

2) Дисперсии напряжений и его составляющих находятся по линеаризованной зависимости с использованием матрицы частных производных первого порядка – матрицы Якоби (5.44). По соображениям, вытекающим из обеспечения сходимости итерационного процесса расчета установившегося режима, уравнения баланса мощности располагаются таким образом, что для каждого узла вначале записывается уравнение для реактивной мощности, а затем для активной мощности. В таком же порядке записаны ниже и производные этих функций.

Матрица Якоби

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} -2U'B_{11} + U_0B_{10} & -2U''B_{11} - U_0G_{10} \\ 2U'G_{11} - U_0G_{10} & 2U''G_{11} - U_0B_{10} \end{pmatrix}, \tag{5.44}$$

где
$$G_{10} = \frac{R}{R^2 + X^2}, \quad B_{10} = \frac{X}{R^2 + X^2},$$

$$G_{11} = G_{10}, \quad B_{11} = \frac{B_C}{2}.$$

Ковариационная матрица напряжений находится по выражению:

$$\mathbf{R}_U = \mathbf{A}^{-1} \mathbf{R}_{PQ} (\mathbf{A}^{-1})^T. \quad (5.45)$$

Ковариационная матрица мощностей имеет вид

$$\mathbf{R}_{PQ} = \begin{pmatrix} D_Q & \text{cov}(Q, P) \\ \text{cov}(P, Q) & D_P \end{pmatrix}. \quad (5.46)$$

Матрица \mathbf{R}_U является симметричной и имеет следующую структуру

$$\mathbf{R}_U = \begin{pmatrix} D_{U'} & \text{cov}(U', U'') \\ \text{cov}(U'', U') & D_{U''} \end{pmatrix}. \quad (5.47)$$

Дисперсия модуля напряжения вычисляется по формуле (5.43).

Единицы измерения kvolt \equiv 1000·volt MVA \equiv 10³·kvolt·amp

kamp := 1000·amp Mwatt \equiv MVA Mvar \equiv MVA

Исходные данные

ORIGIN := 1

1. Погонные параметры и длина ЛЭП

$$r_0 := 0.249 \frac{\text{ohm}}{\text{km}} \quad x_0 := 0.427 \frac{\text{ohm}}{\text{km}} \quad l := 30 \cdot \text{km}$$

2. Параметры трансформатора

$$R_T := 1.3 \cdot \text{ohm} \quad X_T := 34.7 \cdot \text{ohm} \quad Z_T := R_T + i \cdot X_T$$

3. Мощность нагрузки

$$P_{\text{load}} := 30 \cdot \text{Mwatt} \quad Q_{\text{load}} := 25 \cdot \text{Mvar} \quad S_{\text{load}} := P_{\text{load}} + i \cdot Q_{\text{load}}$$

$$S := -S_{\text{load}} \quad \text{Задающие мощности узлов}$$

4. Напряжения

$$U_{\text{nom}} := 110 \cdot \text{kvolt} \quad \text{Номинальное напряжение сети.}$$

$$U_0 := 115 \cdot \text{kvolt} \quad \text{Напряжение базисного узла.}$$

Расчетные данные

5. Расчетные параметры ЛЭП

$$R_L := r_0 \cdot l \quad X_L := x_0 \cdot l \quad Z_L := R_L + i \cdot X_L$$

$$Z_L = 7.47 + 12.81i \cdot \text{ohm}$$

1. Получение узловых проводимостей

$$G := \text{Re}(Y) \quad B := \text{Im}(Y) \quad G_0 := \text{Re}(Y_0) \quad B_0 := \text{Im}(Y_0)$$

$$Y = 3.757 \cdot 10^{-3} - 0.02i \cdot \text{siemens} \quad Y_0 = 3.757 \cdot 10^{-3} - 0.02i \cdot \text{siemens}$$

$$Y := \frac{1}{Z_L + Z_T} \quad Y0 := \frac{1}{Z_L + Z_T}$$

3. Определение ковариационной матрицы мощностей

$$K_{qp} := \begin{bmatrix} 60 & 48 \\ 48 & 60 \end{bmatrix} \cdot \text{MVA}^2$$

Вычисления

1. Определение математического ожидания напряжения

Решение нелинейного уравнения установившегося режима.

Начальное приближение напряжения:

$$U1 := (1.0 - 1i \cdot 0.15) \cdot U_{nom}$$

Решающий блок - приближенное решение

Given

$$U1 \cdot (\overline{Y \cdot U1} - \overline{Y0 \cdot U0}) = S$$

$$U1 := \text{MinErr}(U1)$$

$$U1 = 99.278 - 10.487i \cdot \text{kvolt}$$

Результат решения - узловые напряжения -
в экспоненциальной форме записи:

$$|U1| = 99.831 \cdot \text{kvolt} \quad \arg(U1) = -6.03 \cdot \text{deg}$$

$$U1' := \text{Re}(U1) \quad U1' = 99.278 \cdot \text{kvolt}$$

$$U1'' := \text{Im}(U1) \quad U1'' = -10.487 \cdot \text{kvolt}$$

2. Вычисление ковариационной матрицы напряжений

2.1. Матрица Якоби

$$A := \begin{bmatrix} -2 \cdot U1' \cdot B + U0 \cdot B0 & -2 \cdot U1'' \cdot B - U0 \cdot G0 \\ 2 \cdot U1' \cdot G - U0 \cdot G0 & 2 \cdot U1'' \cdot G - U0 \cdot B0 \end{bmatrix} \quad A = \begin{bmatrix} 1.701 & -0.859 \\ 0.314 & 2.262 \end{bmatrix} \cdot \text{kamp}$$

2.2. Ковариационная матрица напряжений

$$K_{u'u''} := A^{-1} \cdot K_{qp} \cdot (A^T)^{-1} \quad K_{u'u''} = \begin{bmatrix} 31.733 & 12.79 \\ 12.79 & 7.565 \end{bmatrix} \cdot \text{kvolt}^2$$

2.3. Дисперсия модуля напряжения

$$DU := \frac{1}{(|U1|)^2} \cdot \left(U1'^2 \cdot K_{u'u''_{1,1}} + U1''^2 \cdot K_{u'u''_{2,2}} + 2 \cdot U1' \cdot U1'' \cdot K_{u'u''_{1,2}} \right)$$

$$DU = 28.794 \cdot \text{kvolt}^2$$

3. Оценка вероятности допустимости уровня напряжения

Числовые характеристики напряжения, приведенного к напряжению первичной обмотки

$$m_u := |U_1| \quad m_u = 99.831 \cdot \text{kvolt} \quad \text{Математическое ожидание}$$

$$\sigma_u := \sqrt{DU} \quad \sigma_u = 5.366 \cdot \text{kvolt} \quad \text{Среднеквадратическое отклонение}$$

3.2. Пусть диапазон приемлемых значений напряжений составляет:

$$U_{\min} := 96.5 \cdot \text{kvolt} \quad U_{\max} := 133.42 \cdot \text{kvolt}$$

$$\text{pnorm}\left(\frac{U_{\min}}{\text{kvolt}}, \frac{m_u}{\text{kvolt}}, \frac{\sigma_u}{\text{kvolt}}\right) = 0.267 \quad \text{Вероятность выхода напряжения за допустимые пределы}$$

Пример 2. Дисперсия напряжения при известных случайных параметрах продольного сопротивления ветви.

$$U = \sqrt{\left(U_0 - \frac{PR + QX}{U}\right)^2 + \left(\frac{PX - QR}{U}\right)^2} = \sqrt{(U_0 - \Delta U')^2 + \Delta U''^2}, \quad (5.48)$$

$$D[U] = \frac{1}{U^4} [(\Delta U'P + \Delta U''Q)^2 D[R] + (\Delta U'P - \Delta U''Q)^2 D[X]]$$

Дисперсия потерь мощности при случайном характере продольного сопротивления ветви.

$$\Delta P = 3I^2 R, \quad (5.49)$$

$$D[\Delta P] = (3I^2)^2 D[R].$$

Найдем числовые характеристики модуля напряжения и потерь активной мощности для сети из примера 1, но теперь случайными величинами будем считать только параметры сети, а мощности нагрузок будут неслучайными величинами. Значения погрешностей параметров сети возьмем из табл. 5.3.

Дисперсия активного и реактивного сопротивления сети складывается из дисперсий сопротивления ЛЭП и трансформатора

$$D_R := (0.06 \cdot R_L)^2 + (0.06 \cdot R_T)^2 \quad D_R = 0.207 \cdot \text{ohm}^2$$

$$D_X := (0.015 \cdot X_L)^2 + (0.05 \cdot X_T)^2 \quad D_X = 3.047 \cdot \text{ohm}^2$$

Введем обозначения для падения напряжения:

$$dU := U_0 - U_1 \quad dU' := \text{Re}(dU) \quad dU'' := \text{Im}(dU)$$

$$dU' = 15.722 \cdot \text{kvolt} \quad dU'' = 10.487 \cdot \text{kvolt}$$

$$D_U := \frac{\left[(dU' \cdot P_{\text{load}} + dU'' \cdot Q_{\text{load}})^2 \cdot D_R + (dU'' \cdot P_{\text{load}} - dU' \cdot Q_{\text{load}})^2 \cdot D_X\right]}{(|U_1|)^4}$$

$$D_U = 1.311 \cdot 10^{-3} \text{ } \circ \text{kvolt}^2 \quad \sqrt{D_U} = 0.036 \text{ } \circ \text{kvolt}$$

Ток по линии:

$$I := \frac{S_{\text{load}}}{\sqrt{3} \cdot U_1} \quad |I| = 0.226 \text{ } \circ \text{kamp}$$

Математическое ожидание и дисперсия потерь мощности:

$$m_{\Delta P} := 3 \cdot (|I|)^2 \cdot (R_L + R_T) \quad m_{\Delta P} = 1.342 \text{ } \circ \text{Mwatt}$$

$$D_{\Delta P} := [3 \cdot (|I|)^2]^2 \cdot D_R \quad D_{\Delta P} = 4.846 \cdot 10^{-3} \text{ } \circ \text{Mwatt}^2$$

$$\sqrt{D_{\Delta P}} = 0.07 \text{ } \circ \text{Mwatt}$$

Из расчета следует, что влияние погрешностей параметров схемы замещения ЛЭП очень мало, и при решении большинства задач анализа режимов ЭЭС их можно не учитывать.

5.3. Оценивание состояния ЭЭС

Схема формирования модели текущего состояния ЭЭС. Модель текущего состояния ЭЭС необходима для контроля допустимости текущего режима, соблюдения требований по качеству, надежности и экономичности режима [6, 9–11]. Эта модель включает в себя:

- 1) топологию сети – множество узлов и связей между ними на текущий момент времени (в виде графа сети и матриц инцидентий);
- 2) параметры элементов расчетной схемы (сопротивления, проводимости, коэффициенты трансформации, коэффициенты статических характеристик нагрузок и генераторов, параметры электростанций и т. д.);
- 3) параметры режима расчетной схемы (напряжения и фазы напряжений узлов, перетоки активной и реактивной мощности и токи в ветвях (связях), мощности генерации и потребления (токи) в узлах).

Данная модель текущего состояния обычно задается в табличной форме (таблица узловых параметров и таблица параметров ветвей).

Исходными данными для формирования расчетной модели, используемой при оперативном управлении, должны быть:

- данные о положении телекоммуникационной аппаратуры;
- паспортные, экспериментальные и другие данные такого же рода о характеристиках элементов схемы, имеющиеся в базе данных, с которыми работает оперативный комплекс управления режимом ЭЭС;

- технологические пределы, не зависящие от текущего режима, в том числе предельные токи, уровни напряжения;
- измерения параметров режима;
- статистические сведения об измерениях параметров режима, наложенные в процессе эксплуатации;
- сведения о составе и состоянии оборудования;
- метеорологические факторы;
- информация, поступающая от смежных отраслей промышленности и потребителей;
- характеристики водохранилищ;
- сведения о располагаемых запасах и качестве топлива;
- метрологические характеристики поступающих измерений;
- состав и настройка автоматики и релейной защиты.

Соответственно можно выделить четыре основных задачи формирования модели состояния ЭЭС, решаемые в цикле реального времени:

- формирование топологии электрической сети;
- формирование расчетной схемы сети;
- оценивание состояния электрической сети;
- получение ограничений на параметры режима.

Формирование топологии электрической сети. Задача формулируется следующим образом.

По данным телесигналов и состоянии коммутационной аппаратуры подстанций сформировать топологию (способ соединения) электрической сети. Анализ топологии сети состоит из двух частей:

1) анализ подстанций, на которых произошло изменение состояния коммутационной аппаратуры; здесь проверяется, изменился ли при этой коммутации состав связей и узлов расчетной схемы. На каждой подстанции составляется список всех ветвей, инцидентных данному узлу;

2) проверка – не произошло ли разбивки расчетной схемы на несвязанные между собой части – потери связности сети.

Оценивание состояния ЭЭС. Процедура оценивания состояния ЭЭС заключается в том, что находятся такие значения параметров режима (называемые в дальнейшем оценками и обозначаемые знаками \wedge сверху), которые удовлетворяют условиям установившихся и наиболее близки (в смысле некоторого критерия) к измеренным значениям.

Назначение оценивания состояния состоит в следующем:

- сбалансировать измеренные значения параметров режима так, чтобы выполнялись законы Ома и Кирхгофа, описывающие режим электрической сети;
- дорассчитать неизмеренные параметры;

- отфильтровать ошибки измерений;
- дать прогнозы изменения параметров режима;
- оценить статические характеристики ошибок полученных результатов.

Различают статическое и динамическое оценивание состояния.

При статическом оценивании состояния используют измерения, полученные в один и тот же момент времени (хотя бы теоретически), а результаты предшествующих моментов времени не учитываются.

Динамическое оценивание состояния учитывает данные предшествующих моментов времени с помощью так называемых моделей динамики.

Здесь рассматриваются только вопросы статического оценивания состояния.

Статическое оценивание состояния. Измерения, поступающие по каналам связи в компьютер, образуют вектор

$$\bar{\mathbf{v}} = \mathbf{v}(\mathbf{x}) + \boldsymbol{\xi}, \quad (5.50)$$

где $\mathbf{v}(\mathbf{x})$ – истинные значения измеренных параметров; \mathbf{x} – вектор состояния; $\bar{\mathbf{v}}$ – вектор измеренных значений параметров; $\boldsymbol{\xi}$ – вектор ошибок измерений, имеющий, в общем случае, математическое ожидание отличное от нуля (систематическую ошибку) и ковариационную матрицу $\mathbf{R} = \text{cov}(\boldsymbol{\xi}) = M[(\boldsymbol{\xi} - \Delta)(\boldsymbol{\xi} - \Delta)^T]$, характеризующие случайную составляющую ошибки измерения.

Обычно считают, что случайные составляющие ошибок измерений некоррелированы между собой, т. е. элементы матрицы \mathbf{R} равны

$$r_{i,j} = \begin{cases} \sigma_i^2 & \text{при } i = j; \\ 0 & \text{при } i \neq j. \end{cases} \quad (5.51)$$

Иногда считают, что и $\Delta = 0$.

Зависимости $\mathbf{v}(\mathbf{x})$ известны, они представляют собой выражения для потоков мощности, токов, узловых мощностей в зависимости от модулей и фаз напряжений в узлах.

Поскольку $\boldsymbol{\xi}$ априори неизвестны, то полагая их равными нулю, получаем систему нелинейных уравнений относительно вектора \mathbf{x}

$$\bar{\mathbf{v}} - \mathbf{v}(\mathbf{x}) = \mathbf{0}. \quad (5.52)$$

Если число измеряемых величин равно n (порядку вектора \mathbf{x}) и в качестве компонент вектора \mathbf{v} выступают узловые мощности, то имеем традиционную систему уравнений установившегося режима. При этом две компоненты вектора \mathbf{v} есть модуль и фаза (обычно принимается равной нулю) базисного узла.

Если же число измерений больше минимально необходимого, т. е. $m > n = 2(N - 1)$, то система уравнений (5.52) из-за ошибок измерений, как правило,

несовместная, т. е. равенство нулю невязок $2(N - 1)$ уравнений из m оставит ненулевыми невязки остальных $m - 2(N - 1)$ уравнений. В этом случае в качестве решения берется такое значение вектора \mathbf{x} , которое минимизирует взвешенную сумму квадратов невязок, причем в качестве весов выступают величины, обратные дисперсиям:

$$\varphi = \sum_{i=1}^n \frac{[\bar{v}_i - v_i(\mathbf{x})]^2}{\sigma_{vi}^2} \quad (5.53)$$

или в матричном виде

$$\varphi = [\bar{\mathbf{v}} - \mathbf{v}(\mathbf{x})]^T \mathbf{R}_v^{-1} [\bar{\mathbf{v}} - \mathbf{v}(\mathbf{x})]. \quad (5.54)$$

Минимум функции φ находится приравниванием нулю производных φ по компонентам вектора \mathbf{x}

$$\frac{\partial \varphi}{\partial \mathbf{x}} = -2 \sum_{i=1}^m [\bar{v}_i - v_i(\mathbf{x})] \frac{1}{\sigma_{vi}^2} \frac{\partial v_i}{\partial \mathbf{x}} = 0 \quad (5.55)$$

или в матричной записи

$$\frac{\partial \varphi}{\partial \mathbf{x}} = -2 [\bar{\mathbf{v}} - \mathbf{v}(\mathbf{x})]^T \mathbf{R}_v^{-1} \frac{\partial \mathbf{v}}{\partial \mathbf{x}} = 0. \quad (5.56)$$

После упрощения будем иметь

$$\left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right)^T \mathbf{R}_v^{-1} [\bar{\mathbf{v}} - \mathbf{v}(\mathbf{x})] = 0. \quad (5.57)$$

Эта система уравнений является нелинейной и ее решение $\mathbf{x} = \hat{\mathbf{x}}$ является оценкой вектора состояния.

Дисперсии составляющих оценок вектора \mathbf{x} можно найти по линеаризованной зависимости в точке $\hat{\mathbf{x}}$:

$$\left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right)^T \mathbf{R}_v^{-1} [\bar{\mathbf{v}} - \mathbf{v}(\mathbf{x})] \approx \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right)^T \bigg|_{\mathbf{x}=\hat{\mathbf{x}}} \mathbf{R}_v^{-1} \left[\bar{\mathbf{v}} - \mathbf{v}(\hat{\mathbf{x}}) + \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right) \bigg|_{\mathbf{x}=\hat{\mathbf{x}}} \Delta \mathbf{x} \right] = 0. \quad (5.58)$$

Отсюда следует, что

$$\left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right)^T \bigg|_{\mathbf{x}=\hat{\mathbf{x}}} \mathbf{R}_v^{-1} \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right) \bigg|_{\mathbf{x}=\hat{\mathbf{x}}} \Delta \mathbf{x} = \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right)^T \bigg|_{\mathbf{x}=\hat{\mathbf{x}}} \mathbf{R}_v^{-1} \Delta \mathbf{v}. \quad (5.59)$$

Обозначим $\mathbf{A} = \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right)^T \bigg|_{\mathbf{x}=\hat{\mathbf{x}}} \mathbf{R}_v^{-1} \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right) \bigg|_{\mathbf{x}=\hat{\mathbf{x}}}$ $\mathbf{b} = \left(\frac{\partial \mathbf{v}}{\partial \mathbf{x}} \right)^T \bigg|_{\mathbf{x}=\hat{\mathbf{x}}} \mathbf{R}_v^{-1}$,

тогда (5.10) есть система линейных уравнений вида $\mathbf{A} \Delta \mathbf{x} = \mathbf{b} \Delta \mathbf{v}$.

Это линейное соотношение между $\Delta \mathbf{x} = \mathbf{x} - \hat{\mathbf{x}}$ и $\Delta \mathbf{v} = \mathbf{v} - \hat{\mathbf{v}}$ позволяет вычислить ковариационную матрицу оценок вектора \mathbf{x}

$$\mathbf{R}_x = \mathbf{A}^{-1} \mathbf{b} \mathbf{R}_v \mathbf{b}^T (\mathbf{A}^{-1})^T. \quad (5.60)$$

Обнаружение «плохих» данных. Под плохими понимаются данные, ошибки которых значительно больше априори предполагаемых.

Для обнаружения плохих данных используются:

- сравнение текущего замеренного значения с априори известным диапазоном изменения данной величины;
- сравнение с измерением в предыдущий момент времени – в этом случае считается априори известным диапазон изменения скорости параметра;
- анализ распределения ошибок измерения;
- использование специальных неквадратичных критериев;
- анализ невязок так называемых контрольных уравнений – уравнений, в которые входят только измеренные величины.

Все эти меры и их комбинации обеспечивают помехоустойчивость (робастность) алгоритмов оценивания и повышает надежность их работы.

На рис. 5.10–5.14 показаны эти основные приемы повышения помехоустойчивости

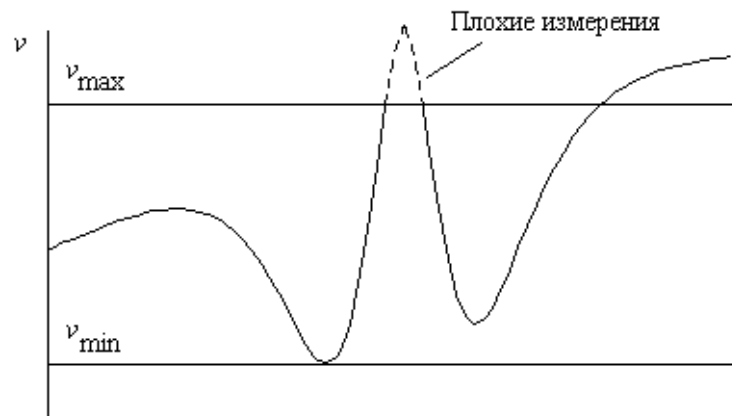


Рис.5.10. Обнаружение плохих данных по диапазону изменения параметра

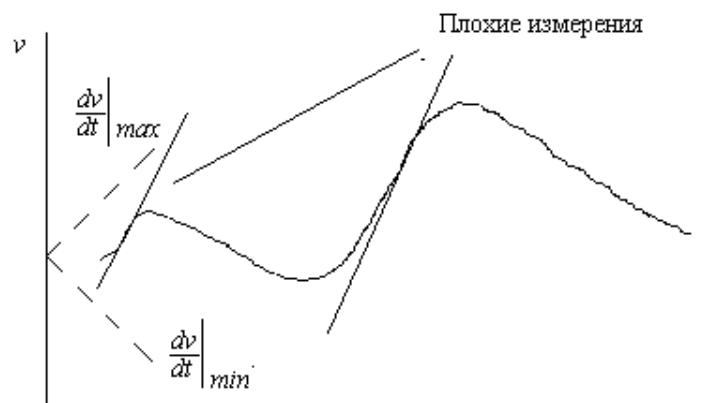


Рис. 5.11. Обнаружение плохих данных по диапазонам скоростей изменения параметра

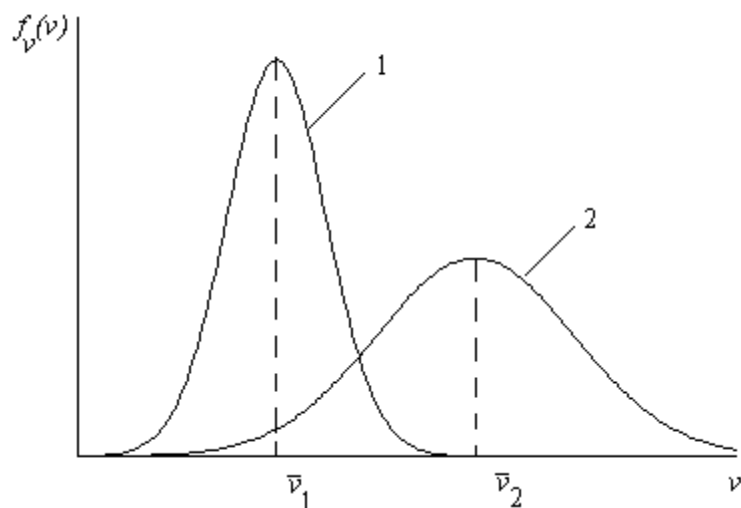


Рис. 5.12. Обнаружение плохих данных по распределениям ошибок: 1 – распределение «хороших» и 2 – распределение «грубых» измерений

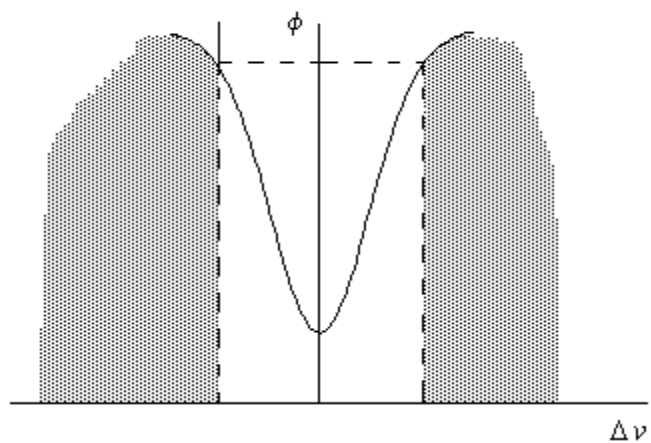


Рис. 5.13. Обнаружение плохих данных по неквадратичным критериям (заштрихованная область «плохих» измерений)

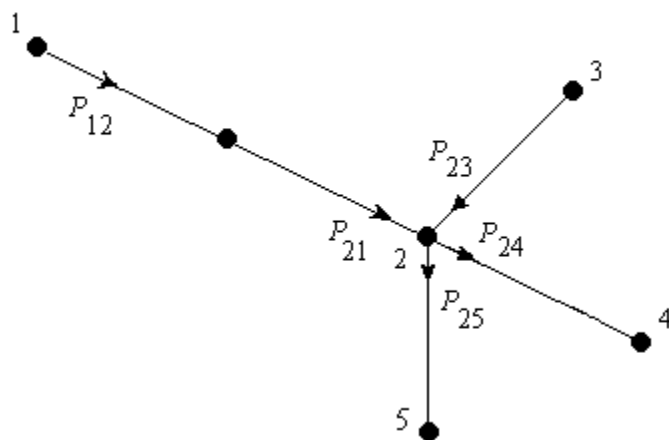


Рис. 5.14. Обнаружение плохих данных с помощью контрольных уравнений:

$$\bar{P}_{1,2} - \bar{P}_{2,1} = w_1 > \varepsilon; \quad -\bar{P}_{2,1} - \bar{P}_{2,3} + \bar{P}_{2,4} + \bar{P}_{2,5} = w_2 < \varepsilon.$$

Алгоритм обнаружения «плохих» данных с помощью аппарата контрольных уравнений. Исходное положение метода контрольных уравнений состоит в том, что, если невязка контрольного уравнения мала, то в него не входит ни одно «плохое» измерение. Это означает, что событие взаимной компенсации ошибок нескольких «плохих» измерений в одном контрольном уравнении считается практически невозможным.

1) Проверка, есть ли среди контрольных уравнений такие, которые имеют невязку

$$|w_{k,i}(\mathbf{v})| \leq b_i, \quad (5.61)$$

где величина b_i определяется через дисперсии измерений, входящих в i -е контрольное уравнение.

Дисперсия невязки контрольного уравнения w_i определяется по линеаризованной зависимости

$$\sigma_{wi}^2 = \sum_{j \in \omega_i} \left(\frac{\partial w_i}{\partial v_j} \right)^2 \sigma_{vj}^2. \quad (5.62)$$

и

$$b_i = t_i \sigma_{wi}, \quad (5.63)$$

где ω_i – множество индексов, входящих в i -е контрольное уравнение переменных; t_i – коэффициент, определяемый вероятностью появления невязки контрольного уравнения больше b_i .

Так как предполагается, что все измерения имеют нормальное распределение, то в силу линейной зависимости величина w_i также распределена нормально.

Все измерения, входящие в контрольные уравнения, для которых выполняется условие (5.61), объявляются достоверными. Они помечаются индексом d , т. е. \mathbf{v}_d . Контрольные уравнения с малыми невязками исключаются из дальнейшего рассмотрения.

2) Если нет контрольных уравнений, для которых условие (5.61) не выполняется, то алгоритм завершен.

3) Из контрольных уравнений, для которых (5.60) не выполняются, составляются такие подсистемы уравнений, чтобы из них вычислялись не вошедшие в \mathbf{v}_d измерения. Если это невозможно, то процесс заканчивается, и переходим на п. 5. В противном случае для каждой такой подсистемы

$w_{s,r}(\mathbf{v}_s, \bar{\mathbf{v}}_d) = 0$ находятся эти измеряемые параметры \mathbf{v}_s . Обозначим их $\hat{\mathbf{v}}_s$. Они

заменяют измеренные параметры $\bar{\mathbf{v}}_s$. Измерения $\hat{\mathbf{v}}_s$ вводятся в вектор проверенных переменных \mathbf{v}_d , их дисперсии принимаются в соответствии с соотношением:

$$\mathbf{R}_{vs} = \frac{\partial \mathbf{v}_s}{\partial \mathbf{v}_d} \mathbf{R}_{vd} \left(\frac{\partial \mathbf{v}_s}{\partial \mathbf{v}_d} \right)^T, \text{ где } \frac{\partial \mathbf{v}_s}{\partial \mathbf{v}_d} = - \left(\frac{\partial \mathbf{w}_{sk}}{\partial \mathbf{v}_s} \right)^{-1} \frac{\partial \mathbf{w}_{sk}}{\partial \mathbf{v}_d}. \quad (5.64)$$

В матрице \mathbf{R}_s оставляются только диагональные элементы.

После этого определяются новые невязки всех тех контрольных уравнений, которые не были использованы для вычисления \mathbf{v}_s , но в которые \mathbf{v}_s входят. Переход к п. 5.

4) Все непроверенные измерения, оставшиеся в неисключенных контрольных уравнениях, не могут быть теперь проверены и объявляются сомнительными, так как входят в контрольные уравнения с относительно большими невязками.

Чтобы придать им меньший вес при оценивании состояния для них вычисляется новая ковариационная матрица, которая как бы «размазывает» по ним имеющуюся невязку контрольного уравнения

$$w_{ki}^2 = t^2 \sum_{j \in \omega_{iid}} \left(\frac{\partial w_{ki}}{\partial v_i} \right)^2 \sigma_j^2 + t^2 \sum_{j \in \omega_{iiq}} \left(\frac{\partial w_{ki}}{\partial v_i} \right)^2 \sigma_{js}^2 \quad (5.65)$$

Откуда

$$\frac{w_{ki}^2}{t^2} - \sum_{j \in \omega_{iid}} \left(\frac{\partial w_{ki}}{\partial v_i} \right)^2 \sigma_j^2 = \sum_{j \in \omega_{iiq}} \left(\frac{\partial w_{ki}}{\partial v_i} \right)^2 \sigma_{js}^2, \quad (5.66)$$

где ω_{id} – множество индексов достоверных измерений; ω_{iq} – множество индексов сомнительных измерений, входящих в i -е контрольное уравнение.

Выбор σ_{js}^2 из (5.66) неоднозначен. Для определенности можно в одно и то же число раз α увеличить все дисперсии сомнительных измерений

$$\alpha = \frac{\left[\frac{w_{ki}^2}{c^2} - \sum_{j \in \omega_{id}} \left(\frac{\partial w_{ki}}{\partial v_i} \right)^2 \sigma_j^2 \right]}{\sum_{j \in \omega_{iq}} \left(\frac{\partial w_{jk}}{\partial v_j} \right)^2 \sigma_{js}^2}. \quad (5.67)$$

5.4. Регулирование напряжение в распределительных сетях методом характеристического узла

Регулирование напряжения в распределительных и питающих сетях можно рассматривать отдельно, поскольку они независимы друг от друга [7, 11]. РПН – используется для компенсации потерь напряжений, появляющихся в распределительных сетях, рис. 5.15.

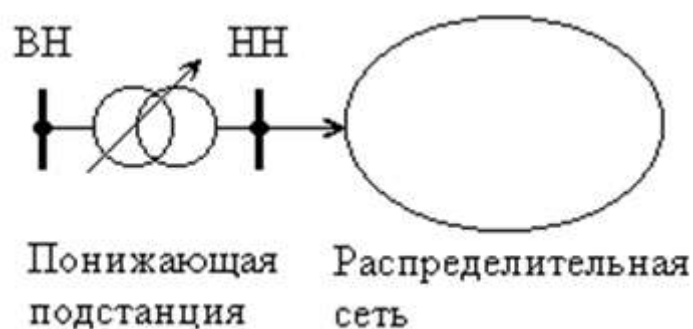


Рис. 5.15. РПН на понижающей подстанции

Для этой цели нужны сведения о потерях напряжений в распределительной сети, для чего можно:

- использовать телеизмерения в распределительной сети;
- применить модель полного сопротивления распределительной сети;
- выполнить разовые расчеты максимальных потерь напряжения.

Первый способ является очень дорогим для такого рода сетей, второй нашел применения за рубежом, в частности во Франции.

В России регулирование напряжения на понижающих подстанциях ведется по принципу встречного регулирования.

Рассмотрим метод регулирования напряжения, в основе которого лежит модель полного сопротивления. На подстанции, где размещено устройство РПН, располагают только информацией о U , P , Q , поставляемыми в сеть, а также токе I .

Модель полного сопротивления, по которому протекает ток, рис. 5.16, соответствует отдаваемой в сеть мощности. Получается одно напряжение, которое подлежит регулированию, но потребители многочисленны и распределены по всей сети.

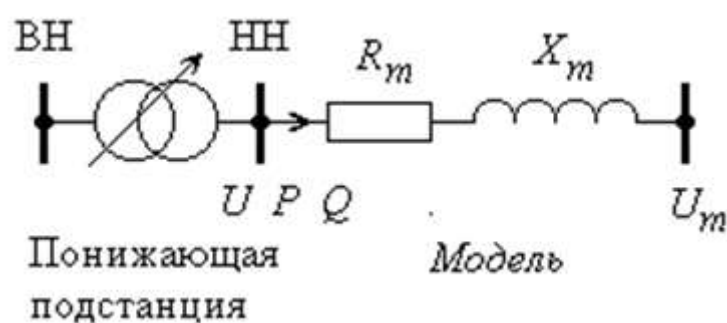


Рис. 5.15. Модель полного сопротивления

Следовательно, схема позволяет регулировать напряжение только в одной точке отходящей сети, которую выбирают оптимально расположенной среди нагрузок; такую точку называют *характеристической или производящей точкой*.

Характеристическая точка должна быть выбрана таким образом, чтобы при хорошем регулировании напряжения общий ущерб (для всей сети) был ми-

нимальным, при этом всякое перемещение регулируемой точки вызывает увеличение общего ущерба.

Если отходящая линия является неразветвленной, то, следовательно, в характеристической точке среднее квадратическое отклонение напряжения должно быть равным средневзвешенной величине (по потребляемым энергиям) от квадратов отклонений в различных точках линии. Если линия имеет длину L и равномерную нагрузку, то характеристическая точка расположена на расстоянии $0,4L$ от подстанции. В разветвленной сети расчет сложнее, однако опыт показывает, что характеристическая точка близка к точке, расстояние до которой $0,4L$.

Положение характеристической точки меняется во времени, поскольку изменения нагрузок происходят неодновременно. Следовательно, нельзя осуществить хорошее регулирование, если изменения нагрузок не будут скорректированы. На практике, как правило, потребители имеют одинаковую природу, например бытовые потребители в жилых квартирах, или сельскохозяйственные потребители, или же промышленные потребители; тогда значения коэффициента корреляции оказывается равным $0,7 \dots 0,9$ (идеальная однородность была бы при $1,0$). Для нагрузок с различной природой коэффициент корреляции незначителен (менее $0,3$) и даже отрицателен (он равен $-1,0$ для двух нагрузок, постоянно меняющихся в противоположных направлениях). В этом случае необходимо разделить сеть на подсети, имеющие независимое регулирование напряжения.

На практике не обязательно уточнять местоположение характеристической точки; достаточно знать статистическое распределение напряжения в этой точке, называемое фиктивным напряжением U_f . Его измеряют статистическим вольтметром в различных точках сети, связанных с нагрузками, потребляемыми вблизи каждой из этих точек. Нагрузки могут быть определены с помощью годовых энергий W_i , потребляемых всем множеством нагрузок, сгруппированных вокруг каждой точки. Итак имеем

$$U_f(t) = \frac{\sum_{i=1}^n W_i U_i(t)}{\sum_{i=1}^n W_i}. \quad (5.68)$$

Из последнего выражения можно найти относительное отклонение фиктивного напряжения V_f , которое позволяет характеризовать качество напряжения в электрической сети

$$V_f = \frac{(U_f - U_{\text{ном}})}{U_{\text{ном}}}. \quad (5.69)$$

Математическая модель, представленная полным сопротивлением, должна воспроизводить напряжение U_f .

Пусть R_m и X_m – значения активного и реактивного сопротивлений модели; U_0 – параметр для настройки системы регулирования; P и Q – активная и

реактивная мощности полной нагрузки регулируемой сети; U – напряжение на шинах подстанции, питающей сеть; U_m – напряжение модели.

При этом имеем

$$U_m = U - \frac{PR_m + QX_m}{U_{\text{НОМ}}}. \quad (5.70)$$

Регулятор должен поддерживать напряжение U_m фиксированным и равным по величине $U_{\text{НОМ}} + U_0$. Выразим из (5.70) напряжение, которое необходимо поддерживать на шинах подстанции

$$U = U_m + \frac{PR_m + QX_m}{U_{\text{НОМ}}}, \quad (5.71)$$

и подставляя вместо модельного напряжения выражение $U_m = U_{\text{НОМ}} + U_0$, имеем

$$U = U_{\text{НОМ}} + U_0 + \frac{PR_m + QX_m}{U_{\text{НОМ}}}, \quad (5.72)$$

и в относительных единицах

$$V = \frac{U - U_{\text{НОМ}}}{U_{\text{НОМ}}} = \frac{U_0}{U_{\text{НОМ}}} + \frac{PR_m + QX_m}{U_{\text{НОМ}}^2} \quad (5.73)$$

или

$$V = V_0 + \frac{PR_m + QX_m}{U_{\text{НОМ}}^2}, \quad (5.74)$$

где $V_0 = \frac{U_0}{U_{\text{НОМ}}}$.

Фиктивное напряжение U_f равно

$$U_f = U - \Delta U, \quad (5.75)$$

откуда

$$\Delta U = U - U_f \quad (5.76)$$

или в относительных единицах

$$\Delta V = \frac{(U - U_f)}{U_{\text{НОМ}}}. \quad (5.77)$$

Относительное отклонение фиктивного напряжения

$$V_f = V - \Delta V \quad (5.78)$$

или

$$V_f = V_0 + \frac{PR_m + QX_m}{U_{\text{НОМ}}^2} - \Delta V. \quad (5.79)$$

Член V_0 – постоянная величина, а остальные – случайные величины (P , Q , ΔV). Эти три величины можно определить с помощью средних величин (математических ожиданий) \bar{P} , \bar{Q} , $\bar{\Delta V}$, среднеквадратических отклонений $\sigma[P]$, $\sigma[Q]$, $\sigma[\Delta V]$ и коэффициентов корреляции $r_{P\Delta V}$, $r_{Q\Delta V}$, r_{PQ} .

Регулирование напряжения оптимально, если параметры R_m , X_m и V_0 выбраны такими, что среднеквадратическое отклонение V_f минимально. В этом случае экономический ущерб, испытываемый всеми потребителями, минимален.

Определим оптимальные значения R_m , X_m и V_0 . Для этого дисперсию отклонения V_f по R_m , X_m

$$\sigma_{V_f}^2 = \frac{R_m^2}{U_{\text{НОМ}}^4} \sigma_P^2 + \frac{X_m^2}{U_{\text{НОМ}}^4} \sigma_Q^2 + \sigma_{\Delta V}^2 +$$

$$+ 2 \left(\frac{R_m X_m}{U_{\text{НОМ}}^4} r_{PQ} \sigma_P \sigma_Q - \frac{R_m}{U_{\text{НОМ}}^2} r_{P\Delta V} \sigma_P \sigma_{\Delta V} - \frac{X_m}{U_{\text{НОМ}}^2} r_{Q\Delta V} \sigma_Q \sigma_{\Delta V} \right) \quad (5.80)$$

продифференцируем и приравняем полученные производные нулю:

$$\frac{\partial \sigma_{V_f}^2}{\partial R_m} = 2R_m \frac{\sigma_P^2}{U_{\text{НОМ}}^4} + 2 \frac{X_m}{U_{\text{НОМ}}^4} r_{PQ} \sigma_P \sigma_Q - \frac{1}{U_{\text{НОМ}}^2} r_{P\Delta V} \sigma_P \sigma_{\Delta V} = 0, \quad (5.81)$$

$$\frac{\partial \sigma_{V_f}^2}{\partial X_m} = 2X_m \frac{\sigma_Q^2}{U_{\text{НОМ}}^4} + 2 \frac{R_m}{U_{\text{НОМ}}^4} r_{PQ} \sigma_P \sigma_Q - \frac{1}{U_{\text{НОМ}}^2} r_{Q\Delta V} \sigma_Q \sigma_{\Delta V} = 0.$$

В результате имеем систему линейных уравнений:

$$R_m \frac{\sigma_P}{U_{\text{НОМ}}^2} + X_m \frac{\sigma_Q}{U_{\text{НОМ}}^2} r_{PQ} - \sigma_{\Delta V} r_{P\Delta V} = 0, \quad (5.82)$$

$$R_m \frac{\sigma_P}{U_{\text{НОМ}}^2} r_{PQ} + X_m \frac{\sigma_Q}{U_{\text{НОМ}}^2} - \sigma_{\Delta V} r_{Q\Delta V} = 0,$$

решение которой можно записать в матричном виде

$$\begin{pmatrix} R_m \\ X_m \end{pmatrix} = \frac{U_{\text{НОМ}}^2}{\Delta} \begin{pmatrix} \sigma_Q & -\sigma_Q r_{PQ} \\ -\sigma_P r_{PQ} & \sigma_P \end{pmatrix} \begin{pmatrix} \sigma_{\Delta V} r_{P\Delta V} \\ \sigma_{\Delta V} r_{Q\Delta V} \end{pmatrix}, \quad (5.83)$$

где

$$\Delta = \sigma_P \sigma_Q - \sigma_P \sigma_Q r_{PQ}^2 = \sigma_P \sigma_Q (1 - r_{PQ}^2). \quad (5.84)$$

Таким образом, имеем выражения для определения оптимальных значений R_m , X_m , а выражение для V_0 определится из (5.79), если считать, что математическое ожидание V_f равно нулю $\bar{V}_f = 0$:

$$R_m = U_{\text{НОМ}}^2 \frac{\sigma_{\Delta V}}{\sigma_P} \frac{r_{P\Delta V} - r_{PQ} r_{Q\Delta V}}{1 - r_{PQ}^2},$$

$$X_m = U_{\text{НОМ}}^2 \frac{\sigma_{\Delta V}}{\sigma_Q} \frac{r_{Q\Delta V} - r_{PQ} r_{P\Delta V}}{1 - r_{PQ}^2}, \quad (5.85)$$

$$V_0 = \Delta \bar{V} - \frac{R_m}{U_{\text{НОМ}}^2} \bar{P} - \frac{X_m}{U_{\text{НОМ}}^2} \bar{Q}.$$

Следовательно, для оптимального регулирования необходимо иметь следующие числовые характеристики режимных параметров:

- математические ожидания и среднеквадратические отклонения активной и реактивной мощности, отдаваемой в сеть;
- математическое ожидание и среднеквадратическое отклонение относительной величины потери напряжения ΔV ;
- коэффициенты корреляции: r_{PQ} , $r_{P\Delta V}$, $r_{Q\Delta V}$.

Первые две группы характеристик определяются путем обработки статистического материала, полученного из натурного или вычислительного экспериментов. Что касается коэффициентов корреляции, то вследствие того, что в большинстве случаев в распределительных сетях активная и реактивная нагрузки меняются одновременно, коэффициент корреляции $r_{PQ} = 1$, а $r_{P\Delta V} = 0,7-0,95$ в большинстве случаев и $r_{P\Delta V} = r_{Q\Delta V}$. С учетом указанных значений формулы (5.85) приобретают вид:

$$\begin{aligned} R_m &= U_{\text{ном}}^2 \frac{\sigma_{\Delta V}}{\sigma_P} \frac{r_{P\Delta V}}{2}, \\ X_m &= U_{\text{ном}}^2 \frac{\sigma_{\Delta V}}{\sigma_Q} \frac{r_{Q\Delta V}}{2}, \\ V_0 &= \Delta \bar{V} - \frac{R_m}{U_{\text{ном}}^2} \bar{P} - \frac{X_m}{U_{\text{ном}}^2} \bar{Q}. \end{aligned} \quad (5.86)$$

Регулирование напряжения заключается в поддержании на шинах подстанции напряжения:

$$U = U_{\text{ном}} + U_{\text{ном}} V_0 + \frac{PR_m + QX_m}{U_{\text{ном}}}. \quad (5.87)$$

Пример 3. Получим закон регулирования напряжения на шинах центра питания распределительной сети 10 кВ, приведенной на рис. 5.16, по методу характеристического узла. В табл. 5.5 приведены параметры схемы сети.

Таблица 5.5

Сопротивления ветвей схемы сети (провод марки АС-50)

Имя ветви	R , Ом	X , Ом
1 - 3	0,5	0,36
2 - 3	0,5	0,36
3 - 5	0,5	0,36
4 - 5	0,5	0,36
5 - 6	1,04	0,72

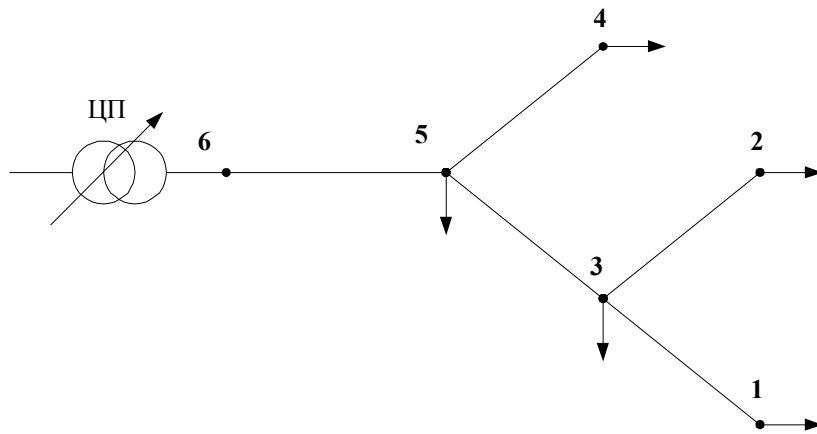


Рис. 5.16. Схема распределительной сети
Расчет приведен в системе Mathcad [4].

ORIGIN := 1 t := 1..6 U_{nom} := 10

Графики нагрузки узлов схемы сети по активной и реактивной мощности
шесть ступеней по 4 часа.

Мощности даны в киловаттах (P) и киловарах (Q).

$$\begin{aligned}
 P_1 &:= \begin{bmatrix} 210 \\ 235 \\ 511 \\ 420 \\ 430 \\ 402 \end{bmatrix} & P_2 &:= \begin{bmatrix} 210 \\ 235 \\ 511 \\ 420 \\ 430 \\ 402 \end{bmatrix} & P_3 &:= \begin{bmatrix} 126 \\ 141 \\ 280 \\ 250 \\ 270 \\ 241 \end{bmatrix} & P_4 &:= \begin{bmatrix} 315 \\ 352 \\ 725 \\ 528 \\ 550 \\ 601 \end{bmatrix} & P_5 &:= \begin{bmatrix} 419 \\ 468 \\ 832 \\ 713 \\ 720 \\ 801 \end{bmatrix} \\
 Q_1 &:= \begin{bmatrix} 289 \\ 315 \\ 250 \\ 220 \\ 300 \\ 420 \end{bmatrix} & Q_2 &:= \begin{bmatrix} 289 \\ 315 \\ 250 \\ 220 \\ 300 \\ 420 \end{bmatrix} & Q_3 &:= \begin{bmatrix} 116 \\ 126 \\ 110 \\ 120 \\ 160 \\ 168 \end{bmatrix} & Q_4 &:= \begin{bmatrix} 289 \\ 315 \\ 320 \\ 330 \\ 380 \\ 420 \end{bmatrix} & Q_5 &:= \begin{bmatrix} 289 \\ 315 \\ 310 \\ 300 \\ 359 \\ 420 \end{bmatrix}
 \end{aligned}$$

Графики напряжений в узлах схемы сети, полученные путем расчета
режимов для каждой ступени графиков мощностей, кВ.

$$\begin{aligned}
 U_1 &:= \begin{bmatrix} 10.21 \\ 10.18 \\ 9.99 \\ 10.07 \\ 10.03 \\ 9.99 \end{bmatrix} & U_2 &:= \begin{bmatrix} 10.21 \\ 10.18 \\ 9.99 \\ 10.07 \\ 10.03 \\ 9.99 \end{bmatrix} & U_3 &:= \begin{bmatrix} 10.23 \\ 10.20 \\ 10.03 \\ 10.01 \\ 10.03 \\ 10.03 \end{bmatrix} & U_4 &:= \begin{bmatrix} 10.25 \\ 10.23 \\ 10.07 \\ 10.13 \\ 10.1 \\ 10.07 \end{bmatrix} & U_5 &:= \begin{bmatrix} 10.28 \\ 10.26 \\ 10.11 \\ 10.17 \\ 10.15 \\ 10.11 \end{bmatrix}
 \end{aligned}$$

График мощностей 5-го узла

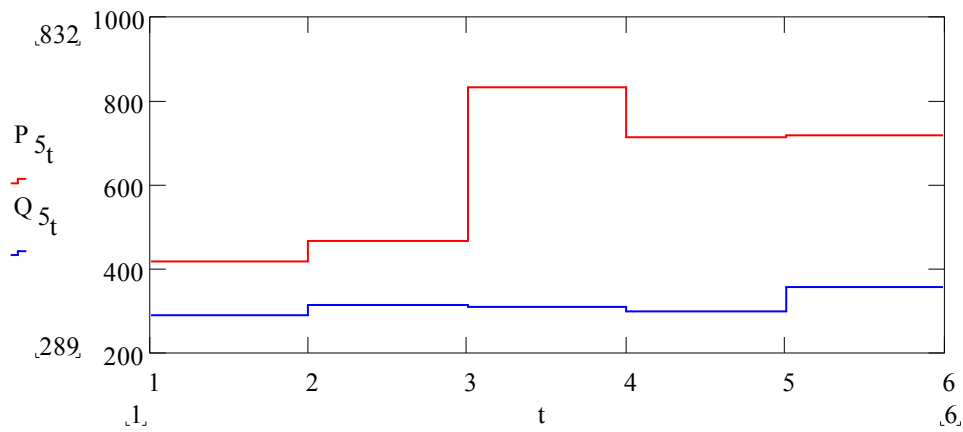
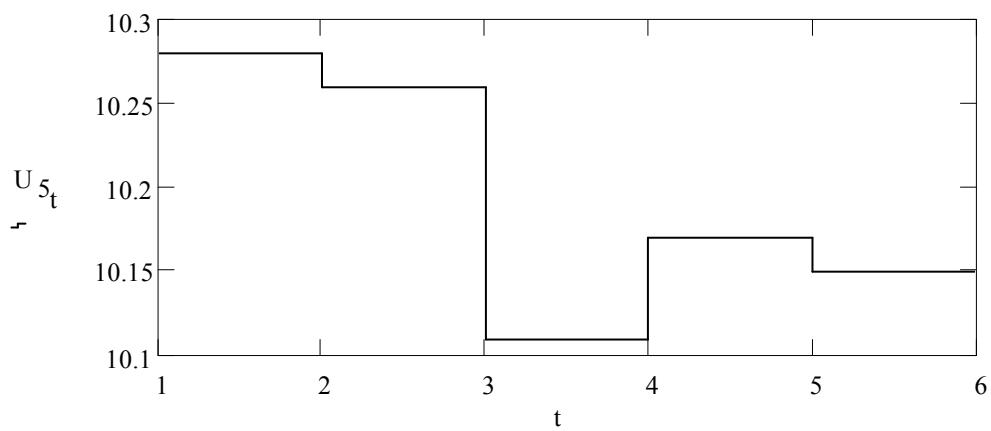


График напряжения в 5-м узле



Суммарный график нагрузки сети

$$P_s := P_1 + P_2 + P_3 + P_4 + P_5 \quad Q_s := Q_1 + Q_2 + Q_3 + Q_4 + Q_5$$

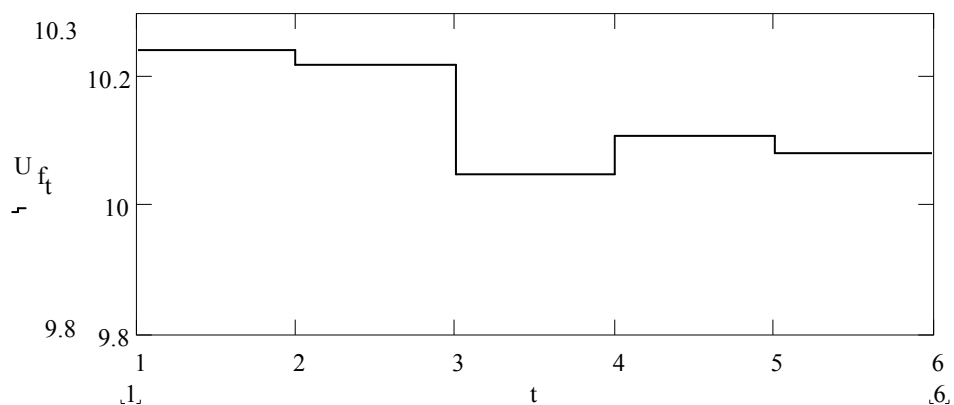
Суточное потребление энергии в каждом узле нагрузки и по сети в целом

$$W_1 := \sum P_1 \cdot 4 \quad W_2 := \sum P_2 \cdot 4 \quad W_3 := \sum P_3 \cdot 4 \quad W_4 := \sum P_4 \cdot 4 \quad W_5 := \sum P_5 \cdot 4$$

$$W := \sum P_s \cdot 4 \quad W = 5.099 \cdot 10^4$$

График напряжения фиктивного узла

$$U_f := \frac{W_1 \cdot U_1 + W_2 \cdot U_2 + W_3 \cdot U_3 + W_4 \cdot U_4 + W_5 \cdot U_5}{W}$$



Расчет статистических характеристик параметров модели

$$V_f := \frac{U_f - U_{nom}}{U_{nom}} \quad \Delta V := \text{mean}(V_f) \quad \Delta V = 0.013$$

$$\sigma_{\Delta V} := \text{stdev}(V_f) \quad \sigma_{\Delta V} = 7.727 \cdot 10^{-3}$$

Примем коэффициенты корреляции:

$$r_{p\Delta V} := 0.8 \quad r_{q\Delta V} := 0.8$$

$$m_p := \text{mean}(P_s) \quad \sigma_p := \text{stdev}(P_s) \quad m_p = 2.125 \cdot 10^3 \quad \sigma_p = 570.988$$

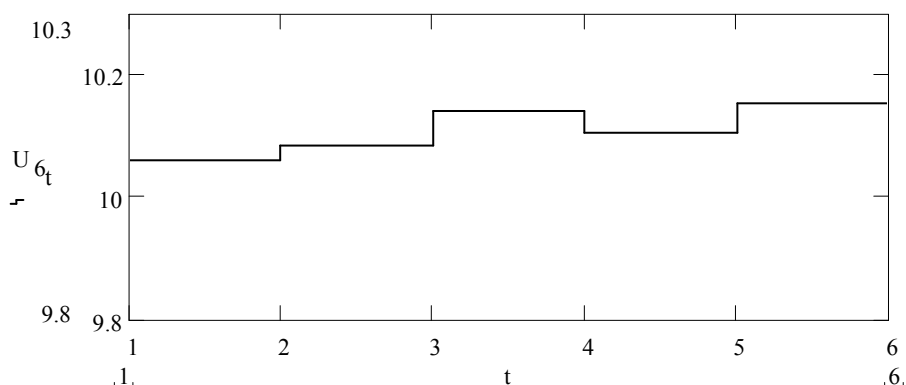


График напряжения в центре питания, который должна поддерживать система регулирования

$$\Delta U := \frac{(P_s \cdot R + Q_s \cdot X) \cdot 10^{-3}}{U_{nom}}$$

$$U_6 := U_{nom} + V_0 \cdot U_{nom} + \Delta U$$

$$U_6 = \begin{bmatrix} 10.062 \\ 10.086 \\ 10.143 \\ 10.107 \\ 10.154 \\ 10.205 \end{bmatrix}$$

$$m_q := \text{mean}(Q_s) \quad \sigma_q := \text{stdev}(Q_s) \quad m_q = 1.406 \cdot 10^3 \quad \sigma_q = 222.329$$

Сопротивления модели и параметр настройки

$$R := \frac{U_{\text{nom}}^2 \cdot \sigma_{\Delta V}}{2 \cdot \sigma_p} \cdot r_{p \Delta V} \cdot 10^3 \quad R = 0.541$$

$$X := \frac{U_{\text{nom}}^2 \cdot \sigma_{\Delta V}}{2 \cdot \sigma_q} \cdot r_{q \Delta V} \cdot 10^3 \quad X = 1.39$$

$$V_0 := \Delta V - \left(\frac{R}{U_{\text{nom}}^2} \cdot m_p + \frac{X}{U_{\text{nom}}^2} \cdot m_q \right) \cdot 10^{-3} \quad V_0 = -0.018$$

Список литературы к 5 главе

1. Гамм А. З. Вероятностные модели режимов электроэнергетических систем. – Новосибирск: ВО «Наука», 1997.
2. Гамм А. З. Статистические методы оценивания состояния электроэнергетических систем. – М.: Наука, 1999.
3. Манусов В. З. Вероятностные задачи в электроэнергетике. – Новосиб. электротехн. ин-т, Новосибирск, 1995.
4. Лыкин А. В. Mathcad в задачах электроэнергетики. Уч. пособие. – Новосибирск: изд-во НГТУ, 1998.
5. Лыкин А. В. Режимы электроэнергетических систем. Вероятностные модели режимов электроэнергетических систем. – Изд. Новосиб. гос. техн. ун-та, Новосибирск, 2000.
6. Методы решения задач реального времени в электроэнергетике / А. З. Гамм, Ю. Н. Кучеров, С. И. Паламарчук и др. – Новосибирск: Наука, 1998.
7. Автоматизированные системы оперативно-диспетчерского управления электроэнергетическими системами / О. Н. Войтов, В. П. Воронин, А. З. Гамм, и др. – Новосибирск: Наука, 2006.
8. Оценивание состояния в электроэнергетике / А. З. Гамм, Л. Н. Герасимов, И. И. Голуб и др. – М.: Наука, 1995.
9. Гамм А. З., Голуб И. И. Наблюдаемость электроэнергетических систем. – М.: Наука, 1999.
10. Пелисье Р. Энергетические системы / Пер. с франц. – М.: Высшая школа, 1995.
11. Холмский В.Г. Расчет и оптимизация режимов электрических сетей. – М.: Высшая школа, 2005.

6. ПРАКТИКА РЕШЕНИЯ ЗАДАЧ ОПТИМИЗАЦИИ

6.1. Общие положения

При проектировании и эксплуатации систем энерго- и электроснабжения часто приходится иметь дело с многовариантными задачами, т.е. с задачами в которых из некоторого множества допустимых по техническим условиям решения нужно выбрать одно, которое является лучшим по какому либо критерию. Такое решение принято называть *оптимальным*, а задачи, в которых производится поиск такого решения, получили название *оптимизационных задач*.

Применительно к системам электроснабжения оптимизационные задачи приходится решать при выборе напряжения электрических сетей, выборе числа и мощности источников питания, выборе оптимальной конфигурации электрической сети, выборе сечений проводников, определении рационального распределения источников реактивной мощности, выборе мест размещения источников питания и т.д.

В качестве *критериев оптимальности* в большинстве практических задач электроснабжения используются экономические показатели (себестоимость, прибыль, финансовые затраты и т.п.), хотя в некоторых случаях могут быть использованы и другие: минимум потерь напряжения (энергии), надежность электроснабжения, качество электроэнергии и т.л.

Таким образом, *критерием оптимальности является количественная оценка оптимизируемого качества объекта*.

На основании выбранного критерия оптимальности составляется целевая функция, представляющая собой зависимость критерия оптимальности от параметров, влияющих на ее значение. Вид критерия оптимальности или целевой функции определяется конкретной задачей оптимизации.

Соответственно, *задача оптимизации сводится к нахождению экстремума целевой функции*.

При решении оптимизационных задач после выбора параметров оптимизации (сечения проводников, количество трансформаторов и т.д.), необходимо определить ограничения на эти параметры. При этом ограничения могут накладываться как по техническим, так и по экономическим соображениям.

В общем случае математическая модель оптимизационной задачи содержит три базовых компонента: целевая функция, ограничения, граничные условия.

Целевая функция представляет собой математическую запись критерия оптимальности:

$$Z(x_1, x_2, \dots, x_n) \rightarrow \text{extr},$$

где x_1, x_2, \dots, x_n – искомые переменные, значения которых необходимо определить в процессе решения задачи.

Ограничения представляют собой различные технические, экономические и другие условия, которые необходимо учесть при решении задачи:

$$f_j(x_1, x_2, \dots, x_n) (\leq \text{или} =) b_j \text{ где } j = 1, 2, \dots, m.$$

Граничные условия определяют диапазон изменения искоемых переменных:

$$d_i \leq x_i \leq D_i, \quad i = 1, 2, \dots, n,$$

где d_i, D_i – нижняя и верхняя граница диапазона изменения переменной x_i соответственно.

Наиболее распространенным случаем граничных условий искоемых переменных в реальных технических задачах является их неотрицательность – $x_i \geq 0$.

Для решения оптимизационных задач используют специальные математические приемы и методы, которые получили название *методов математического программирования*.

В соответствии с характером зависимости между переменными в выражении целевой функции оптимизационные задачи классифицируются на *задачи линейного программирования* и *задачи нелинейного программирования*.

Кроме того, по характеру изменения искоемые переменные могут иметь *непрерывный, целочисленный* или *дискретный* характер. Соответственно, задачи оптимизации, в которых имеются целочисленные или дискретные переменные, подразделяются на задачи *целочисленного* или *дискретного программирования*.

Примерами непрерывных переменных являются значения тока и мощности в линии электропередачи; целочисленными переменными являются количество силовых трансформаторов и компенсирующих устройств; дискретными переменными являются сечения проводников и мощности трансформаторов, которые выбираются из стандартного ряда.

Важное влияние на вид оптимизационной задачи накладывает характер исходной информации. Если исходная информация однозначно определена, то она называется *детерминированной*; если же она носит случайный характер и подчиняется законам теории вероятностей, то она называется *случайной*. Исходная информация, которая носит неопределенный характер и не подчиняется теории вероятностей, называется *недетерминированной*.

Оптимизационные задачи, в которых исходная информация носит случайный характер, классифицируются как задачи *стохастического программирования*, а задачи, в которых исходная информация не определена, могут быть решены с помощью *теории игр*.

Существуют и другие виды классификации задач оптимизации, основной целью которых является выявление специфических особенностей тех или иных задач, играющих важную роль при разработке методов их решения.

6.2. Решение задач линейного программирования

Общая задача линейного программирования состоит в минимизации (максимизации) линейной функции

$$Z = c_1 x_1 + c_2 x_2 + \dots + c_n x_n$$

от n переменных x_1, x_2, \dots, x_n , удовлетворяющих условиям неотрицательности

$$x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0$$

и m линейным ограничениям

$$\begin{aligned} a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n &\leq (=, \geq) b_1, \\ a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n &\leq (=, \geq) b_2, \\ &\dots \dots \dots \\ a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n &\leq (=, \geq) b_m. \end{aligned}$$

Решение подобных задач может быть выполнено с помощью электронных таблиц MS Excel, которые содержат в своем составе специализированные средства, позволяющие решать большинство типовых практических задач оптимизации.

Для того чтобы решить задачу линейного программирования в табличном редакторе Microsoft Excel, необходимо придерживаться следующего плана действий.

1. Ввести условие задачи:
 - а) *создать экранную форму для ввода условия задачи:*
 - переменных,
 - целевой функции (ЦФ),
 - ограничений,
 - граничных условий;
 - б) *ввести исходные данные в экранную форму*
 - коэффициенты целевой функции,
 - коэффициенты при переменных в ограничениях,
 - правые части ограничений;
 - в) *ввести зависимости из математической модели в экранную форму*
 - формулу для расчета целевой функции;
 - формулы для расчета значений левых частей ограничений;
 - г) *задать целевую функцию (в окне «Поиск решения»):*
 - целевую ячейку,
 - направление оптимизации ЦФ;
 - д) *ввести ограничения и граничные условия (в окне «Поиск решения»):*
 - ячейки со значениями переменных,
 - граничные условия для допустимых значений переменных,
 - соотношения между правыми и левыми частями ограничений.

6. Решить задачу:

а) установить параметры решения задачи (в окне «Поиск решения»),

б) запустить задачу на решение (в окне «Поиск решения»;

в) выбрать формат вывода решения (в окне «Результаты поиска решения»)

Задача 1. Агропромышленное мебельное предприятие выпускает три вида изделий: И1, И2 и И3. На изготовление каждого изделия требуется три вида сырья: элемент 1 (Э1), элемент 2 (Э2) и элемент 3 (Э3). Для упрощения задачи будем считать, что расход энергетических, трудовых и других ресурсов на изготовления каждого вида продукции одинаков. Нормы расхода каждого вида сырья на изготовление одного изделия и прибыль от его реализации приведены в табл. 6.1. При планировании производственной программы необходимо учесть ограничения на каждый вид сырьевого ресурса, которые составляют 80, 120, 60 для Э1, Э2 и Э3 соответственно. Требуется определить, в каком количестве нужно выпускать каждый вид изделия (И), чтобы получить максимальную прибыль (маркетинговую ситуацию на рынке не учитываем – все что произвели, сразу продали).

Таблица 6.1

Исходные данные к задаче 1

Параметры	Изделие 1	Изделие 2	Изделие 3
Элемент 1	1	3	2
Элемент 2	2	4	8
Элемент 3	1	2	4
Прибыль	25	70	120

Решение. Математическая модель решаемой задачи будет иметь следующий вид.

Целевая функция

$$Z = 25 \cdot x_1 + 70 \cdot x_2 + 120 \cdot x_3 \rightarrow \max. \quad (6.1)$$

Система ограничений

$$\left. \begin{aligned} 1 \cdot x_1 + 3 \cdot x_2 + 2 \cdot x_3 &\leq 80; \\ 2 \cdot x_1 + 4 \cdot x_2 + 8 \cdot x_3 &\leq 120; \\ 1 \cdot x_1 + 2 \cdot x_2 + 4 \cdot x_3 &\leq 60. \end{aligned} \right\}$$

Граничные условия

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0. \quad (6.3)$$

Из математической модели задачи видно, что целевая функция имеет линейную зависимость от переменных, значит, решаемая задача является задачей линейного программирования.

На рабочем листе книги MS Excel создадим экранную форму для ввода условий задачи, рис. 6.1

	A	B	C	D	E	F	G
1		Переменные					
2	Вид	Табурет	Стул	Стол			
3	Значение	0	0	0			
4	Нижн. граница	>=0	>=0	>=0			
5	Верх. граница				Целевая функция	Тип	
6	Козфф. ЦФ	25	70	120	=СУММПРОИЗВ(B3:D3;B6:D6)	max	
7							
8		Ограничения					
9	Вид				Левая часть	Знак	Пр. часть
10	Э1	1	3	2	=СУММПРОИЗВ(B3:D3;B10:D10)	<=	80
11	Э2	2	4	8	=СУММПРОИЗВ(B3:D3;B11:D11)	<=	120
12	Э3	1	2	4	=СУММПРОИЗВ(B3:D3;B12:D12)	<=	60
13							
14							

Рис. 6.1. Экранная форма задачи 1

В экранной форме на рис. 6.1 каждой переменной и каждому коэффициенту задачи поставлена в соответствие конкретная ячейка в Excel. Имя ячейки состоит из буквы, обозначающей столбец, и цифры, обозначающей строку, на пересечении которых находится объект задачи линейного программирования. Переменным задачи соответствуют ячейки B3:D3, коэффициентам целевой функции соответствуют ячейки B6:D6, правым частям ограничений соответствуют ячейки G10:G16.

В ячейку E6, в которой будет отображаться значение ЦФ, необходимо ввести формулу, по которой это значение будет рассчитано. В экранной форме значение целевой функции можно вычислить по формуле:

$$E6 = \text{СУММПРОИЗВ}(B3:D3;B6:D6) . \quad (6.4)$$

Формулы для расчета левых частей системы ограничений (6.2) разместим в ячейках E10:E16. Для экранной формы, представленной на рис. 6.1, они могут быть рассчитаны по следующим формулам

$$\left. \begin{aligned} E10 &= \text{СУММПРОИЗВ}(B3:D3;B10:D10); \\ E11 &= \text{СУММПРОИЗВ}(B3:D3;B11:D11); \\ E12 &= \text{СУММПРОИЗВ}(B3:D3;B12:D12). \end{aligned} \right\}$$

Экранная форма готова для решения задачи линейного программирования.

В качестве начального приближения значений искомых переменных примем нулевые значения, для чего введем в ячейки B3:D3 нули.

Для решения задачи линейного программирования средствами MS Excel в меню СЕРВИС нужно выбрать команду ПОИСК РЕШЕНИЯ.

Примечание. При первоначальном обращении к инструменту поиска решения в главном меню СЕРВИС можно не обнаружить операции ПОИСК РЕШЕНИЯ. Это означает, что компонент поиска решения в программе MS Excel не установлен. Для установки компонента необходимо выбрать в меню СЕР-

ВИС команду НАДСТРОЙКИ и в открывшемся диалоговом окне установить галочку в строке ПОИСК РЕШЕНИЯ, рис. 6.6.

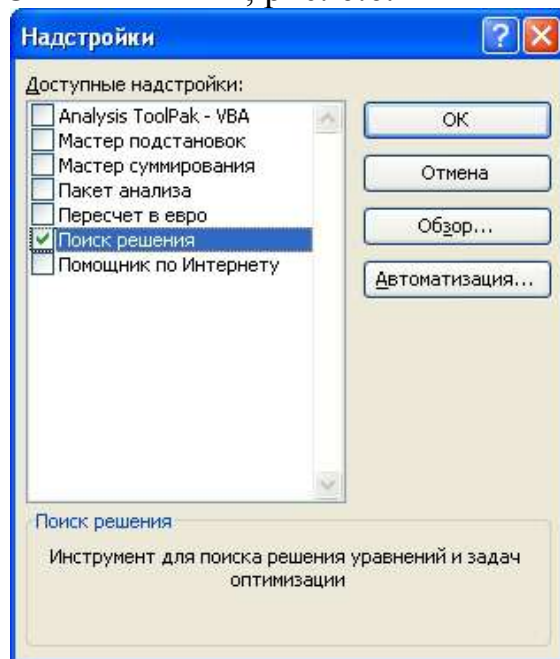


Рис. 6.6. Диалоговое окно установки дополнительных компонент MS Excel

При установке дополнительных компонент программы MS Excel может потребоваться установочный диск пакета MS Office.

Параметры решения задачи устанавливаются в диалоговом окне «Поиск решения», представленном на рис. 6.3

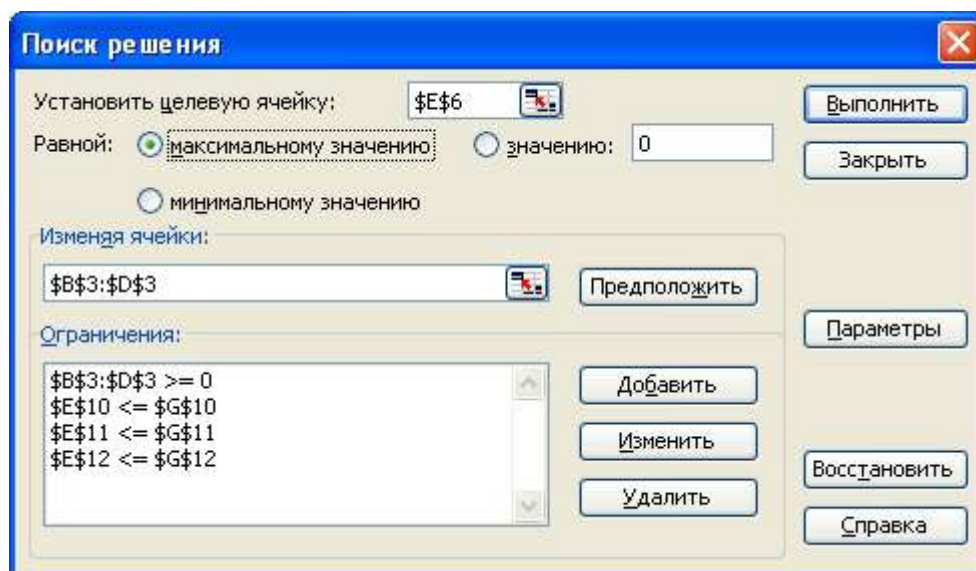


Рис. 6.3. Диалоговое окно «Поиск решения» задачи 1

В открывшемся диалоговом окне необходимо выбрать целевую ячейку (в нашей задаче E6), установить переключатель характера решаемой задачи в требуемое положение (в нашем случае «максимальное значение»), в поле «Изменяя ячейки» указать ячейки, которые должны изменяться в процессе поиска

решения задачи (в нашей задаче B3:D3). После этого в поле «Ограничения» нужно ввести ограничения и граничные условия решаемой задачи оптимизации.

Для того, чтобы добавить ограничение, нужно нажать кнопку ДОБАВИТЬ, и в отрывшемся диалоговом окне «Добавление ограничения» установить ссылку на требуемую ячейку, вид и значение ограничения (рис.6.4).

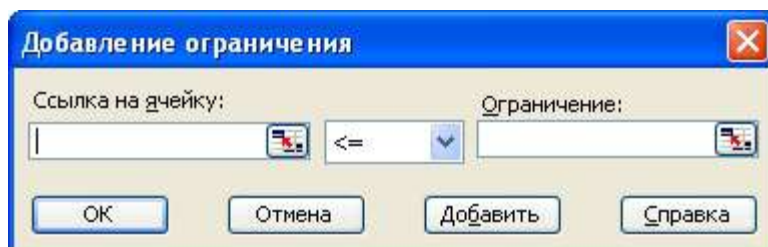


Рис. 6.4. Диалоговое окно «Добавление ограничения»

В соответствии с заданными условиями и экранной формой решаемой нами задачи в нашем случае требуется ввести 4 ограничения:

- значения ячеек E10:E12 должны быть меньше или равны значениям ячеек G10:G12 соответственно (система ограничений 6.2);
- значения ячеек B3:D3 должны быть больше или равны нулю (граничные условия 6.3).

Для решения задач оптимизации определенного класса необходима установка конкретных параметров поиска решения. Для этого нужно нажать кнопку ПАРАМЕТРЫ и заполнить некоторые поля окна «Параметры поиска решения» (рис. 6.5).

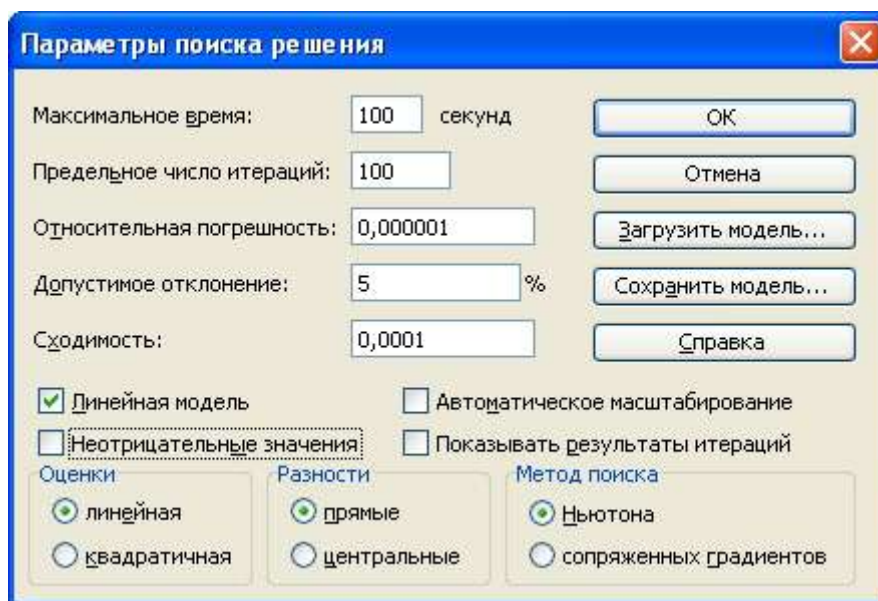


Рис. 6.5. Диалоговое окно «Параметры поиска решения»

Параметр «Максимальное время» служит для назначения времени (в секундах), выделяемого на решение задачи. В поле можно ввести время, не превышающее 32 767 секунд (более 9 часов).

Параметр «Предельное число итераций» служит для управления временем решения задачи путем ограничения числа промежуточных вычислений. В поле можно ввести количество итераций, не превышающее 32 767.

Параметр «Относительная погрешность» служит для задания точности, с которой определяется соответствие ячейки целевому значению или приближение к указанным границам. Поле должно содержать число из интервала от 0 до 1. Чем меньше количество десятичных знаков во введенном числе, тем ниже точность. Высокая точность увеличит время, которое требуется для того, чтобы сошелся процесс оптимизации.

Параметр «Допустимое отклонение» служит для задания допуска на отклонение от оптимального решения в целочисленных задачах. При указании большего допуска поиск решения заканчивается быстрее.

Параметр «Сходимость» применяется только при решении нелинейных задач.

Установка флажка «Линейная модель» обеспечивает ускорение поиска решения линейной задачи за счет применения симплекс-метода.

Установленные по умолчанию значения параметров подходят для большинства решаемых задач. Оставим их без изменения, отметив галочкой «Линейная модель», так как наша задача линейная.

Подтвердим установленные параметры нажатием кнопки ОК.

Запуск задачи на решение производится из окна «Поиск решения» путем нажатия кнопки ВЫПОЛНИТЬ.

На экране появляется окно «Результаты поиска решения», представленное на рис. 6.6.

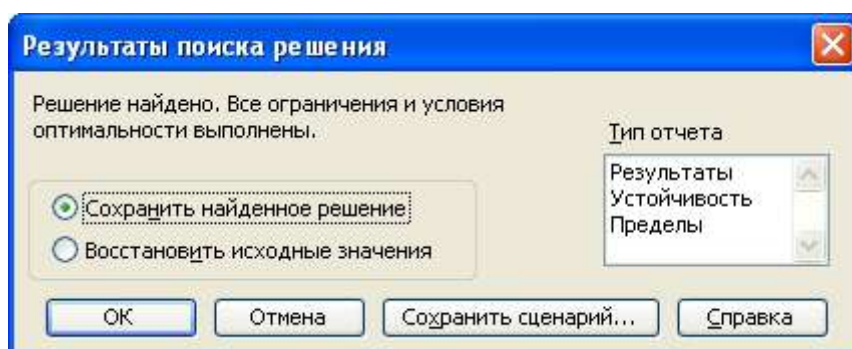


Рис. 6.6. Диалоговое окно «Результаты поиска решения»

Примечание. Иногда при решении оптимизационных задач характер выводимых в диалоговом окне «Результаты поиска решения» сообщений может быть другим: «Поиск не может найти подходящего решения» или «Значения целевой ячейки не сходятся». Вывод таких сообщений не всегда свидетельствуют о характере оптимального решения задачи, а чаще о том, что при вводе условий задачи в Excel были допущены ошибки, не позволяющие Excel найти

оптимальное решение, которое в действительности существует.

Если решение оптимизационной задачи найдено, мы можем сохранить его в ячейках экранной формы и просмотреть три типа отчетов «Результаты», «Устойчивость», «Пределы», которые предлагает Excel для анализа полученного решения. Нажмем в диалоговом окне «Результаты поиска решения» кнопку ОК и сохраним результаты решения в экранной форме.

Экранная форма с результатами решения задачи представлена на рис. 6.7.

	А	В	С	Д	Е	Ф	Г	Н
1		Переменные						
2	Вид	И1	И2	И3				
3	Значение	0	25	2,5				
4	Нижн. граница	≥ 0	≥ 0	≥ 0				
5	Верх. граница				Целевая функция	Тип		
6	Козфф. ЦФ	25	70	120	2050	max		
7								
8		Ограничения						
9	Вид				Левая часть	Знак	Пр. часть	
10	Э1	1	3	2	80	\leq	80	
11	Э2	2	4	8	120	\leq	120	
12	Э3	1	2	4	60	\leq	60	
13								
14								

Готово

Рис. 6.7. Экранная форма задачи 1 с результатом решения

Из результатов решения задачи, представленных на рис.1.7, следует, что для получения максимальной прибыли в планируемом производственном периоде в размере 2050 у.е. мы должны произвести 25 стульев и 2,5 стола. При этом запасы всех наших ресурсов (Э1, Э2 и Э3) будут израсходованы полностью без остатка.

С практической точки зрения полученное решение не имеет смысла, так как производство 0,5 изделия если теоретически и возможно, то получение прибыли от его реализации крайне затруднительно.

С математической точки зрения при решении задачи мы допустили ошибку, выполнив решение без учета того, что наша задача относится к классу задач целочисленного программирования – переменные: табуретки, столы и стулья могут быть только целыми числами. Для исправления ошибки выберем в меню СЕРВИС команду ПОИСК РЕШЕНИЯ и в поле «Ограничения» добавим еще одно ограничение – ячейки В3:D3 должны быть целыми числами (рис. 6.8).

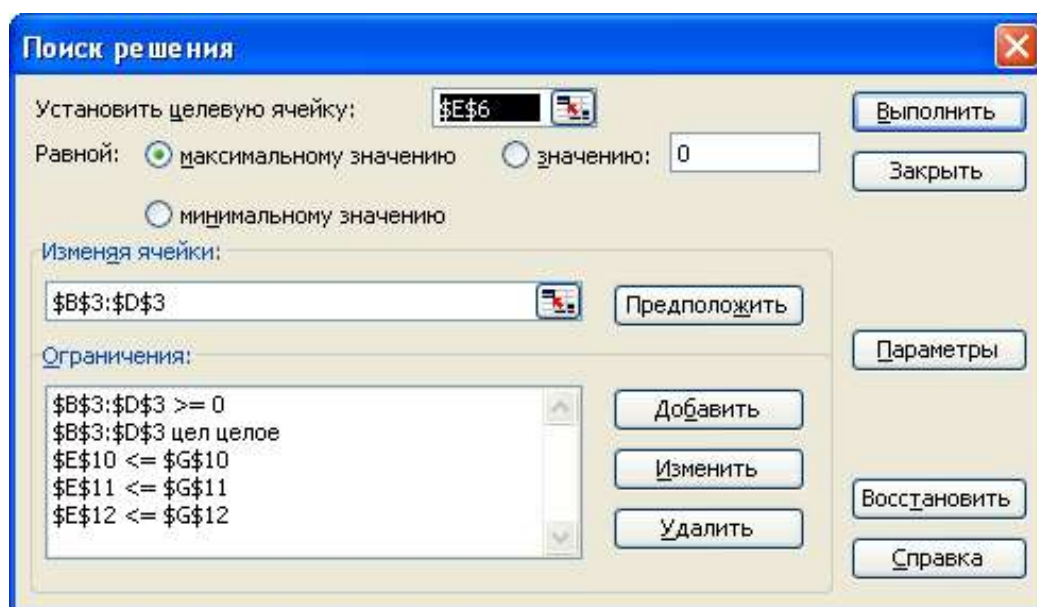


Рис. 6.8. Диалоговое окно «Поиск решения» целочисленной задачи 1

Экранная форма с результатами решения задачи линейного целочисленного программирования представлена на рис. 6.9.

	A	B	C	D	E	F	G	H
1		Переменные						
2	Вид	I1	I2	I3				
3	Значение	0	24	3				
4	Нижн. граница	>=0	>=0	>=0				
5	Верх. граница							
6	Козфф. ЦФ	25	70	120	Целевая функция	Тип		
7					2040	max		
8		Ограничения						
9	Вид				Левая часть	Знак	Пр. часть	
10	Э1	1	3	2	78	<=	80	
11	Э2	2	4	8	120	<=	120	
12	Э3	1	2	4	60	<=	60	
13								
14								

Рис. 6.9. Экранная форма задачи 1 с результатом решения

Выберем в диалоговом окне «Результаты поиска решения» (рис. 6.6) тип отчета «Результаты» и нажмем кнопку ОК.

MS Excel автоматически создаст в рабочей книге новый лист «Отчет по результатам 1», представленный на рис. 6.10.

1	Microsoft Excel 11.0 Отчет по результатам						
2	Рабочий лист: [Пример1.xls]Задача 1			И1 - изделие 1			
3	Отчет создан: 04.03.2007 14:52:46			И2 - изделие 2			
4				И3 - изделие 3			
5							
6	Целевая ячейка (Максимум)						
7	Ячейка		Имя	Исходное значение		Результат	
8	\$E\$6	Козфф. ЦФ	Целевая функция	0		2040	
9							
10							
11	Изменяемые ячейки						
12	Ячейка		Имя	Исходное значение		Результат	
13	\$B\$3	Значение	И1	0		0	
14	\$C\$3	Значение	И2	0		24	
15	\$D\$3	Значение	И3	0		3	
16							
17							
18	Ограничения						
19	Ячейка		Имя	Значение	Формула	Статус	Разница
20	\$E\$10	Э1	Левая часть	78	\$E\$10<=\$G\$10	не связан.	2
21	\$E\$11	Э2	Левая часть	120	\$E\$11<=\$G\$11	связанное	0
22	\$E\$12	Э3	Левая часть	60	\$E\$12<=\$G\$12	связанное	0
23	\$B\$3	Значение	И1	0	\$B\$3>=0	связанное	0
24	\$C\$3	Значение	И2	24	\$C\$3>=0	не связан.	24
25	\$D\$3	Значение	И3	3	\$D\$3>=0	не связан.	3
26	\$B\$3	Значение	И1	0	\$B\$3=целое	связанное	0
27	\$C\$3	Значение	И2	24	\$C\$3=целое	связанное	0
28	\$D\$3	Значение	И3	3	\$D\$3=целое	связанное	0

Готово

NUM

Рис. 6.10. Отчет по результатам решения задачи

Отчет состоит из трех таблиц.

В первой таблице указана ячейка целевой функции, ее исходное (начальное) значение и полученный оптимальный результат – 2040 у.е.

Во второй таблице приведены номера ячеек, наименование, исходное и полученные в результате решения задачи значения переменных: И1, И2 и И3. Из таблицы видно, что для получения максимальной прибыли мы должны произвести 24 И2 и 3 И2. Изготавливать И1 в данной постановке задачи экономически не целесообразно.

В третьей таблице представлены данные по ограничениям решаемой задачи. Из таблицы видно, что при планируемом оптимальном объеме производства Э2 и Э3 будут израсходованы полностью, а Э1 останется 2 у.е.

6.3. Решение задач нелинейного программирования

Решение задач нелинейного программирования принципиально ничем не отличается от решения задач линейного и целочисленного программирования. Единственное отличие заключается в том, что при установке параметров поиска решения в диалоговом окне «Параметры поиска решения» (рис. 6.5), необходимо снять галочку в строке «Линейная модель».

Кроме того, процедура поиска решения задач нелинейного программирования более критична к исходным начальным данным.

Для решения задач нелинейного программирования в Excel реализовано два метода: метод Ньютона и метод сопряженных градиентов Флетчера-Ривса. Выбор метода решения производится в диалоговом окне «Параметры поиска решения». В качестве критерия останова поиска решения в Excel используется следующее условие:

$$\Delta f_k = \left| \frac{f(x^{k-1}) - f(x^k)}{f(x^k)} \right| \leq \varepsilon \quad (6.6)$$

Значение ε вводится в окне «Параметры поиска решения» в строке «Относительная погрешность».

В соответствии с выражением (6.6) начальные значения переменных желательно назначать близкими к оптимальным значениям, что значительно ускорит процесс решения задачи. Обязательным условием является требование неравенства целевой функции в начальной точке нулю, иначе при вычислении погрешности по выражению (6.6) возможно деление на ноль.

6.4. Решение транспортной задачи

Транспортная задача (ТЗ) является частным типом задачи линейного программирования и определяется как задача разработки наиболее экономичного плана перевозки продукции одного вида из нескольких пунктов отправления в пункты назначения. При этом величина транспортных расходов прямо пропорциональна объему перевозимой продукции и задается с помощью тарифов на перевозку единицы продукции.

Задачи транспортного типа широко распространены в практике. К ним сводятся многие задачи линейного программирования – задачи о назначениях, сетевые, календарного планирования и др.

И хотя ТЗ может быть решена одним из методов решения любой задачи линейного программирования, ее математическая модель и структура ограничений имеют ряд специфических особенностей.

Стандартная ТЗ формулируется следующим образом. Имеется m пунктов отправления (или пунктов производства) A_1, \dots, A_m , в которых сосредоточены запасы однородных продуктов в количестве a_1, \dots, a_m единиц. Имеется n пунктов назначения (или пунктов потребления) B_1, \dots, B_n , потребность которых в указанных продуктах составляет b_1, \dots, b_n единиц. Известны также транспортные расходы C_{ij} , связанные с перевозкой единицы продукта из пункта A_i в пункт B_j , $i = 1, 2, \dots, m; j = 1, 2, \dots, n$.

Требуется составить такой план перевозок (откуда, куда и сколько единиц продукта везти), чтобы удовлетворить спрос всех пунктов потребления за счет реализации всего продукта, произведенного всеми пунктами производства, при минимальной общей стоимости всех перевозок.

Пусть x_{ij} – количество единиц продукта, поставляемого из пункта A_i в пункт B_j . Подлежащие минимизации суммарные затраты на перевозку продуктов из всех пунктов производства во все пункты потребления выражаются формулой:

$$\sum_{i=1}^m \sum_{j=1}^n C_{ij} \cdot x_{ij} \rightarrow \min. \quad (6.7)$$

Таким образом, целевая функция ТЗ представляет собой транспортные расходы на осуществление всех перевозок в целом.

Математическая модель ТЗ содержит также две группы ограничений.

Первая группа ограничений указывает, что запас продукции в любом пункте отправления должен быть равен суммарному объему перевозок продукции из этого пункта

$$\sum_{j=1}^n x_{ij} = a_i, \text{ где } i = 1, \dots, m. \quad (6.8)$$

Вторая группа ограничений указывает, что суммарные перевозки продукции в некоторый пункт потребления должны полностью удовлетворить спрос на продукцию в этом пункте.

$$\sum_{i=1}^m x_{ij} = b_j, \text{ где } j = 1, \dots, n. \quad (6.9)$$

Объемы перевозок – неотрицательные числа, так как перевозки из пунктов потребления в пункты производства исключены:

$$x_{ij} \geq 0, i = 1, \dots, m; j = 1, \dots, n. \quad (6.10)$$

Из (6.8), (6.9) следует, что сумма запасов продукции во всех пунктах отправления должна равняться суммарной потребности во всех пунктах потребления, то есть

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j. \quad (6.11)$$

Если условие (6.11) выполняется, то ТЗ называется *сбалансированной* (замкнутая модель), в противном случае – *несбалансированной* (открытая модель). Поскольку ограничения модели ТЗ (6.8), (6.9) могут быть выполнены только при сбалансированной ТЗ, то при построении транспортной модели необходимо проверять условие баланса (6.11). В случае, когда суммарные запасы превышают суммарные потребности, необходим дополнительный фиктивный пункт потребления, который будет формально потреблять существующий излишек запасов, то есть

$$b_{\Phi} = \sum_{i=1}^m a_i - \sum_{j=1}^n b_j. \quad (6.12)$$

Если суммарные потребности превышают суммарные запасы, то необходим дополнительный фиктивный пункт отправления, формально восполняющий существующий недостаток продукции в пунктах отправления:

$$a_{\phi} = \sum_{j=1}^n b_j - \sum_{i=1}^m a_i. \quad (6.13)$$

Введение фиктивного потребителя или отправителя повлечет необходимость формального задания фиктивных тарифов C_{ij}^{ϕ} (реально не существующих) для фиктивных перевозок.

Так как нас интересует суммарная стоимость всех перевозок, можно принять величину фиктивного тарифа равной нулю $C_{ij}^{\phi}=0$, что не изменит значение искомой целевой функции.

Рассмотрим процедуру решения транспортной задачи на конкретном примере.

Задача 2. Оптовая база занимается поставкой электрооборудования (ЭО) заказчикам. ЭО поставляется из трех складов, месячные запасы которых составляют 1500, 1300 и 1600 единиц товара соответственно. ЭО нужно развести по трем заказчикам, месячные потребности которых равны 2100, 1600 и 1000 единиц товара соответственно. Транспортные расходы по доставке единицы ЭО из соответствующего склада соответствующему заказчику приведены в табл. 6.6. Необходимо определить оптимальные по транспортным расходам способы доставки ЭО со складов заказчикам.

Таблица 6.2

Транспортные расходы по доставке товара, руб/шт.

Склады	Заказчик 1	Заказчик 2	Заказчик 3
Склад 1	80	200	70
Склад 2	100	105	120
Склад 3	120	70	90

Составим в MS Excel транспортную матрицу для решения задачи, рис. 6.11.

	A	B	C	D	E
1	Склады	Заказчики			Запасы ЭО , шт
2		Заказчик 1	Заказчик 2	Заказчик 3	
3		Стоимость перевозки товара, руб/шт.			
4	Склад 1	80	200	70	1500
5	Склад 2	100	105	120	1300
6	Склад 3	120	70	90	1600
7	Потребность в ЭО , шт.	2100	1600	1000	

Рис. 6.11 Транспортная матрица задачи в MS Excel

Проверяем баланс ТЗ.

Суммарные запасы составляют $\sum_{i=1}^m a_i = 4400$ шт. ЭО, суммарная потребность $\sum_{j=1}^n b_j = 4700$ шт.

Транспортная задача не сбалансирована – спрос превышает предложение. Приведем ТЗ к сбалансированной, добавив еще один фиктивный склад, стоимость перевозки ЭО в магазины заказчика в который будет равна нулю.

Определим количество ЭО на фиктивном складе $a_{\phi} = \sum_{j=1}^n b_j - \sum_{i=1}^m a_i = 4700 - 4400 = 300$ шт. и внесем изменения в лист книги MS Excel.

	A	B	C	D	E
1	Склады	Магазины			Запасы товара, шт
2		Магазин 1	Магазин 2	Магазин 3	
3		Стоимость перевозки товара, руб/шт.			
4	Склад 1	80	200	70	1500
5	Склад 2	100	105	120	1300
6	Склад 3	120	70	90	1600
7	Фиктивн. Склад	0	0	0	300
8	Потребность в товаре, шт.	2100	1600	1000	

Рис. 6.12. Сбалансированная транспортная матрица

Сбалансированность транспортной матрицы легко проверить средствами Excel, последовательно вычислив суммы ячеек В8:D8 и Е4:Е7. Суммы должны быть одинаковыми.

Создадим экранную форму решения ТЗ, для чего выделим на листе книги все ячейки транспортной матрицы, создадим ее копию и внесем в копию необходимые исправления. За начальные условия перевозимого количества товара примем нули, рис. 6.13.

	A	B	C	D	E
1	Склады	Магазины			Запасы товара, шт
2		Магазин 1	Магазин 2	Магазин 3	
3		Стоимость перевозки товара, руб/шт.			
4		Склад 1	80	200	
5	Склад 2	100	105	120	1300
6	Склад 3	120	70	90	1600
7	Фиктивн. Склад	0	0	0	300
8	Потребность в товаре, шт.	2100	1600	1000	
9					
10					
11	Склады	Магазины			Запасы товара, шт
12		Магазин 1	Магазин 2	Магазин 3	
13		Количество перевозимого товара, шт.			
14		Склад 1	0	0	
15	Склад 2	0	0	0	
16	Склад 3	0	0	0	
17	Фиктивн. Склад	0	0	0	
18	Потребность в товаре, шт.				
19	Стоимость перевозок, руб				

Рис. 6.13. Транспортная матрица и экранная форма решения ТЗ

Заполним полученную экранную форму необходимыми формулами. Для этого удобно воспользоваться встроенными в Excel математическими формулами СУММ() и СУММПРОИЗВ(), рис. 6.14.

	A	B	C	D	E
1	Склады	Магазины			Запасы товара, шт
2		Магазин 1	Магазин 2	Магазин 3	
3		Стоимость перевозки товара, руб/шт.			
4	Склад 1	80	200	70	1500
5	Склад 2	100	105	120	1300
6	Склад 3	120	70	90	1600
7	Фиктивн. Склад	0	0	0	300
8	Потребность в товаре, шт.	2100	1600	1000	
9					
10					
11	Склады	Магазины			Запасы товара, шт
12		Магазин 1	Магазин 2	Магазин 3	
13		Количество перевозимого товара, шт.			
14	Склад 1	0	0	0	=СУММ(B14:D14)
15	Склад 2	0	0	0	=СУММ(B15:D15)
16	Склад 3	0	0	0	=СУММ(B16:D16)
17	Фиктивн. Склад	0	0	0	=СУММ(B17:D17)
18	Потребность в товаре, шт.	=СУММ(B14:B17)	=СУММ(C14:C17)	=СУММ(D14:D17)	
19	Стоимость перевозок, руб	=СУММПРОИЗВ(B4:B7;B14:B17)	=СУММПРОИЗВ(C4:C7;C14:C17)	=СУММПРОИЗВ(D4:D7;D14:D17)	=СУММ(B19:D19)
20					

Рис. 6.14. Экранная форма решения ТЗ с введенными формулами

Ячейка E19 содержит формулу для вычисления значения целевой функции – суммарной стоимости всех транспортных перевозок.

Выберем в меню СЕРВИС команду ПОИСК РЕШЕНИЯ и установим в диалоговом окне формы «Поиск решения» целевую ячейку, требуемый вид за-

висимости целевой функции, диапазон ячеек переменных и ограничения в соответствии с условиями решаемой ТЗ аналогично примеру Задачи 1, рис. 6.15.

Рис. 6.15. Окно форма «Поиск решения» транспортной задачи

Нажмем кнопку «Параметры» и в открывшемся диалоговом окне формы «Параметры поиска решения» отметим галочкой «Линейная модель», так как решаемая нами ТЗ является задачей линейного программирования. После выполнения процедуры поиска решения на рабочем листе книги получим решение ТЗ, рис. 6.16.

	А	В	С	Д	Е
11	Склады	Магазины			Запасы товара, шт
12		Магазин 1	Магазин 2	Магазин 3	
13		Количество перевозимого товара, шт.			
14	Склад 1	500	0	1000	1500
15	Склад 2	1300	0	0	1300
16	Склад 3	0	1600	0	1600
17	Фиктивн. Склад	300	0	0	300
18	Потребность в товаре, шт.	2100	1600	1000	
19	Стоимость перевозок, руб	170000	112000	70000	352000

Рис. 6.16. Результат решения ТЗ на листе книги MS Excel

6.5. Определение оптимального количества трансформаторов цеховых подстанций

Задача 3. От шин 10 кВ главной понизительной подстанции (ГПП) предприятия осуществляется электроснабжение цехов с суммарными расчетными нагрузками P_p и Q_p . (рис. 6.17). Определить оптимальное

количество цеховых трансформаторов напряжением 10/0,4 кВ с заданными номинальной мощностью $S_{тр}$ и коэффициентом загрузки k_3 при условии, что со стороны питания потребляемая реактивная мощность не должна превышать значения Q_c . Устройства для компенсации реактивной мощности могут быть установлены как на шинах 10 кВ ГПП Q_{10} , так и на шинах 0,4 кВ цеховых трансформаторов Q_{04} .

Исходные данные для решения задачи приведены в табл. 6.3. Затраты на единицу мощности трансформаторов и компенсирующих устройств обозначены через Z . Необходимо составить математическую модель задачи, создать экранную форму и решить задачу в программе MS Excel.

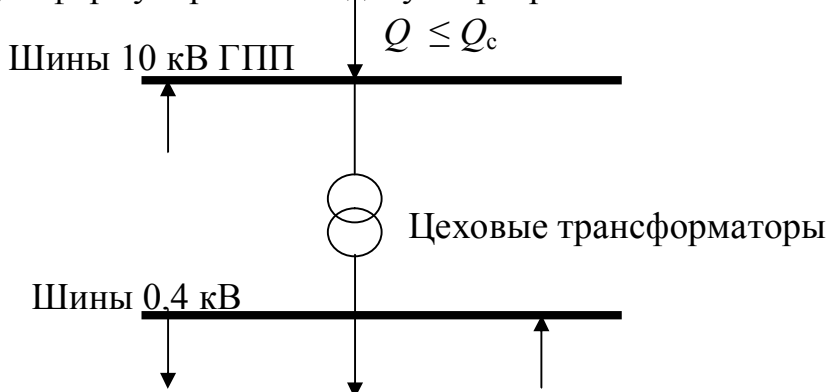


Рис. 6.17. Расчетная схема электроснабжения предприятия

Таблица 6.3

Исходные данные

№ вар	P_p , МВА	Q_p , Мвар	k_3	$S_{тр}$, кВА	$Z_{тр}$, у.е./кВА	Z_{04} , у.е./квар	Z_{10} , у.е./квар
1	20	18	0,7	1000	9	10	4
2	25	20	0,75	1600	10	9	4,5
3	30	25	0,8	2500	9	8	4
4	33	29	0,85	1000	10	8,5	5
5	35	30	0,9	1600	9	10	5
6	28	24	0,75	2500	11	9	4,5
7	33	25	0,8	1000	10	8	3,8
8	30	28	0,7	1600	12	9	5
9	32	27	0,75	2500	11	10	4
10	26	22	0,8	630	12	9	3,8
11	28	25	0,7	1000	10	10	5
12	24	20	0,75	630	10	9	4
13	30	25	0,8	1600	9	8	3,5
14	25	20	0,7	630	10	9	5
15	23	19	0,8	1000	9	9	5

Методические указания. По заданным расчетным активной P_p и реактивной Q_p нагрузкам определяется полная расчетная нагрузка

$$S_p = \sqrt{P_p^2 + Q_p^2}. \quad (6.14)$$

По величине S_p и заданному коэффициенту загрузки k_3 определяется максимальное количество цеховых трансформаторов с заданной номинальной мощностью $S_{тр}$

$$N_{\max} = \frac{S_p}{S_{тр} \cdot k_3}. \quad (6.15)$$

При полной компенсации реактивной мощности на шинах 0,4 кВ ($Q_p = Q_{04}$) количество цеховых трансформаторов будет минимальным

$$N_{\min} = \frac{P_p}{S_{тр} \cdot k_3}. \quad (6.16)$$

Значения N_{\max} и N_{\min} округляются до ближайших больших целых чисел.

Оптимальное количество цеховых трансформаторов N , подлежащее определению, будет лежать в пределах $N_{\max} \geq N \geq N_{\min}$.

Величина мощности компенсирующих устройств на шинах 0,4 кВ, позволяющая сократить количество трансформаторов на единицу составит

$$Q_{04}^1 = \frac{Q_p}{N_{\max} - N_{\min}}. \quad (6.17)$$

Для определения оптимального количества трансформаторов необходимо найти минимум целевой функции

$$Z = Z_{тр} \cdot S_{тр} \cdot N + Z_{04} \cdot Q_{04} + Z_{10} \cdot Q_{10}, \quad (6.18)$$

представляющей собой суммарные затраты на цеховые трансформаторы и компенсирующие устройства на 0,4 и 10 кВ.

Минимум целевой функции (6.18) ищется при следующих ограничениях:

1. Суммарная величина мощности компенсирующих устройств на шинах 0,4 и 10 кВ должна быть равна расчетной реактивной нагрузке (перекомпенсация не допускается)

$$Q_{04} + Q_{10} = Q_p. \quad (6.19)$$

6. Искомое количество трансформаторов, уменьшаемое за счет установки компенсирующих устройств на шинах 0,4 кВ, определится условием

$$N \geq N_{\max} - \frac{Q_{04}}{Q_{04}^1}. \quad (6.20)$$

Решение задачи должно выполняться при граничных условиях

$$N \geq 0, Q_{04} \geq 0, Q_{10} \geq 0. \quad (6.21)$$

Задача приведена к формализованному виду для решения в MS Excel.

6.6. Оптимальное распределение компенсирующих устройств в радиальной схеме электроснабжения

Задача 4. Питание цеховых трансформаторных подстанций промышленного предприятия осуществляется от шин $U = 10$ кВ главной понизительной подстанции (ГПП) кабельными линиями по радиальной схеме (рис. 6.18).

Значения активных сопротивлений кабельных линий R_i , реактивные нагрузки цехов Q_i и суммарная мощность компенсирующих устройств Q_k приведены в табл. 6.4.

Технические данные нерегулируемых конденсаторных установок, используемых на предприятии, приведены в табл. 6.5.

Требуется найти оптимальный вариант распределения компенсирующих устройств заданной суммарной мощности Q_k между цеховыми подстанциями по условию минимума потерь активной мощности в линиях. Найти теоретически возможный и практический минимум потерь активной мощности в системе электроснабжения.

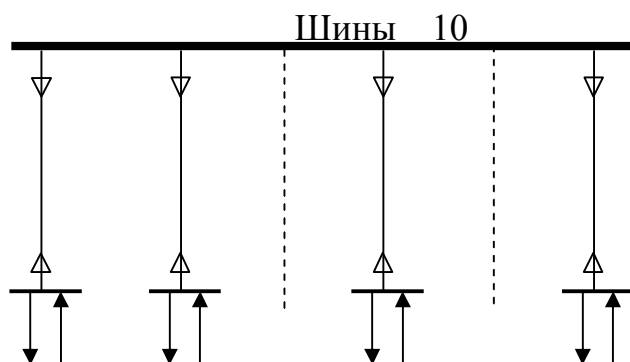


Рис. 6.18. Расчетная схема электроснабжения предприятия

Методические указания. Потери активной мощности в линии при передаче по ней реактивной мощности определяются выражением:

$$\Delta P = \frac{Q^2 \cdot R}{U^2}, \text{ Вт}, \quad (6.24)$$

где Q – значение передаваемой по линии реактивной мощности, квар; R – активное сопротивление линии, Ом; U – номинальное напряжение, кВ.

Суммарные потери активной мощности в радиальной схеме электроснабжения от реактивных нагрузок Q_i при установке у каждой нагрузки компенсирующего устройства мощностью Q_{ki} можно определить по выражению

$$\Delta P = \sum_{i=1}^n \frac{(Q_i - Q_{ki})^2 \cdot R_i}{U^2}, \quad (6.25)$$

где n – количество узлов нагрузки (цеховых трансформаторных подстанций).

Таблица 6.4

Исходные данные

№ вар	R_1 , Ом	R_2 , Ом	R_3 , Ом	R_4 , Ом	Q_1 , квар	Q_2 , квар	Q_3 , квар	Q_4 , квар	Q_k , квар
1	0,1	0,2	0,3	0,4	4000	3000	1000	2000	7500
2	0,15	0,1	0,3	0,25	1000	2000	3000	4000	7500
3	0,35	0,25	0,2	0,1	2500	3500	1000	2000	6000
4	0,4	0,3	0,25	0,2	1000	2000	2500	3500	6000
5	0,25	0,25	0,1	0,1	1500	3500	3500	1500	7500
6	0,1	0,1	0,25	0,25	2500	2000	2000	2500	6600
7	0,25	0,1	0,25	0,1	3000	1500	1500	3000	6600
8	0,15	0,25	0,15	0,25	1500	2500	2500	2500	6000
9	0,2	0,4	0,3	0,1	2500	2500	2500	2500	7500
10	0,4	0,3	0,2	0,1	1500	1500	3000	3000	6600
11	0,4	0,4	0,2	0,15	1000	2000	3000	4000	7500
12	0,25	0,15	0,1	0,4	2000	3000	2500	2500	7500
13	0,25	0,25	0,4	0,1	1500	3000	3000	1500	6600
14	0,2	0,3	0,35	0,25	2000	2500	2000	2500	6600
15	0,35	0,45	0,2	0,25	1000	4000	1500	2500	6000

Таблица 6.5

Технические данные нерегулируемых конденсаторных установок

Тип	Номинальное напряжение, кВ	Шкала номинальных мощностей, квар
УК-10-Q УЗ	10,5	300, 600, 900, 1200, 1500, 1800

Выражение (6.25) является целевой функцией решаемой задачи. Требуется найти минимум целевой функции при следующих ограничениях

1. Мощность устанавливаемых компенсирующих устройств должна быть положительной

$$Q_{ki} \geq 0, i = 1, 2 \dots n. \quad (6.26)$$

6. Суммарная мощность устанавливаемых компенсирующих устройств должна быть равна заданной Q_k

$$\sum_{i=1}^n Q_{ki} = Q_k. \quad (6.27)$$

Выражения (6.25)–(6.27) являются математической моделью решаемой задачи с непрерывными переменными. Решение данной задачи позволяет определить теоретически возможный минимум потерь активной мощности в системе электроснабжения, достижение которого возможно только при использовании в каждом узле нагрузки регулируемых источников реактивной

мощности, например синхронных компенсаторов. Однако такое решение требует значительных капитальных затрат, поэтому на практике для компенсации реактивной мощности на промышленных предприятиях обычно используют нерегулируемые статические конденсаторные батареи. В этом случае мощность компенсирующих устройств может изменяться только с дискретным шагом, величина которого зависит от типа и технических характеристик используемых конденсаторных установок.

Для конденсаторных батарей, технические характеристики которых приведены в табл.1.5, шаг дискретизации составляет 300 квар.

Для нахождения практического минимума потерь активной мощности в системе электроснабжения требуется формализовать математическую модель, описываемую выражениями (6.25)–(6.27), к математической модели с дискретными переменными. В общем случае, способы формализации математических уравнений могут быть различными, выбор же конкретного варианта определяется исследователем самостоятельно.

6.7. Оптимальное распределение компенсирующих устройств в магистральной схеме электроснабжения

Задача 5. Питание группы потребителей промышленного предприятия осуществляется от шин $U = 10$ кВ главной понизительной подстанции (ГПП) кабельными линиями по магистральной схеме (рис. 6.19).

Значения активных сопротивлений кабельных линий R_i , реактивные нагрузки потребителей Q_i и суммарную мощность компенсирующих устройств Q_k возьмем из условий предыдущей задачи (табл. 6.4).

Технические данные нерегулируемых конденсаторных установок, используемых на предприятии, приведены в табл. 6.5.

Требуется найти оптимальный вариант распределения компенсирующих устройств заданной суммарной мощности Q_k между узлами нагрузки по условию минимума потерь активной мощности в линиях. Найти теоретически возможный и практический минимум потерь активной мощности в системе электроснабжения.

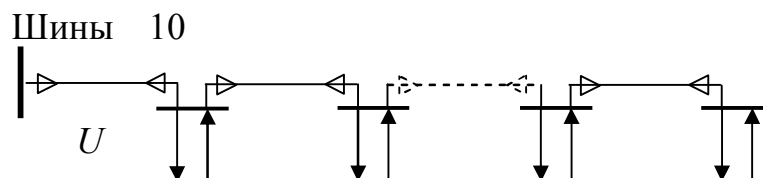


Рис. 6.19. Расчетная схема электроснабжения группы потребителей предприятия

Методические указания. Суммарные потери активной мощности в магистральной линии при передаче по ней реактивной мощности определяются выражением:

$$\Delta P = \frac{1}{U^2} \left(R_1 \cdot \left(\sum_1^n Q_i - \sum_1^n Q_{ki} \right)^2 + R_2 \cdot \left(\sum_2^n Q_i - \sum_2^n Q_{ki} \right)^2 + \dots + \right. \\ \left. + R_i \cdot \left(\sum_i^n Q_i - \sum_i^n Q_{ki} \right)^2 + R_n \cdot (Q_n - Q_{kn})^2 \right), \text{ Вт}, \quad (6.28)$$

где Q_i , Q_{kii} – значения потребляемой и компенсирующей реактивной мощности в узле нагрузки соответственно, квар; R_i – активное сопротивление участка линии, Ом; U – номинальное напряжение, кВ, n – количество узлов нагрузки.

Выражение (6.28) представляет собой целевую функцию решаемой задачи. Требуется найти минимум целевой функции при следующих ограничениях:

1) мощность устанавливаемых компенсирующих устройств должна быть положительной

$$Q_{ki} \geq 0, i = 1, 2 \dots n; \quad (6.29)$$

2) суммарная мощность устанавливаемых компенсирующих устройств должна быть равна заданной Q_k

$$\sum_{i=1}^n Q_{ki} = Q_k. \quad (6.30)$$

Выражения (6.28)–(6.30) являются математической моделью решаемой задачи с непрерывными переменными.

Разработайте экранную форму математической модели задачи и найдите ее решение средствами MS Excel.

Последовательно произвольно измените сопротивления участков линий R_i схемы электроснабжения и значения потребляемой реактивной мощности Q_i в узлах нагрузки, так чтобы суммарные значения сопротивлений и потребляемой реактивной мощности остались неизменными ($\sum R_i = \text{const}$, $\sum Q_i = \text{const}$). Найдите решения и проанализируйте полученные результаты.

Преобразуйте математическую модель задачи в модель с целочисленными дискретными переменными и найдите ее решение. Проведите анализ полученных результатов. Сформулируйте условия размещения компенсирующих устройств в магистральных схемах электроснабжения.

6.8. Оптимальное проектирование схемы электроснабжения

Задача 6. Производится проектирование схемы электроснабжения сетевого района. Определены требуемая мощность производственных объектов и располагаемый резерв мощности близлежащих подстанций (ПС). Известны также расстояния от производственных объектов до каждой подстанции (табл. 6.6).

Требуется определить оптимальную с точки зрения минимума приведенных затрат схему электроснабжения сетевого района.

Методические указания. Приведенные годовые затраты на сооружение и эксплуатацию электрической сети определяются по выражению:

$$Z_{\text{пр}} = p_n \cdot K + C_3, \text{ руб.}, \quad (6.31)$$

где p_n – нормативный коэффициент капитальных вложений; K – капитальные вложения; C_3 – стоимость потерь электроэнергии в проводах ЛЭП.

Величина капитальных вложений на сооружение ЛЭП зависит от ее сечения и длины:

$$K = (a + b \cdot S) \cdot l, \quad (6.32)$$

где a, b – расчетные коэффициенты; S, l – сечение и длина ЛЭП соответственно.

Таблица 6.6

Исходные данные на проектирование

№ вар	ПС	Резерв мощности, МВт	Мощность производственных объектов, МВт					
			Объект 1	Объект 2	Объект 3	Объект 4	Объект 5	Объект 6
			25	40	30	10	20	15
			Расстояние от объекта до подстанции, км					
1	ПС 1	50	15	30	18	42	16	19
	ПС 2	80	25	12	21	30	20	40
	ПС 3	40	30	16	25	25	32	42
2	ПС 1	50	25	25	10	15	15	20
	ПС 2	50	25	20	25	20	25	20
	ПС 3	60	25	15	30	35	20	10
3	ПС 1	50	10	20	25	25	20	15
	ПС 2	40	30	20	20	25	10	20
	ПС 3	80	30	25	25	20	15	15
4	ПС 1	60	10	10	10	20	20	20
	ПС 2	60	30	30	20	15	15	30
	ПС 3	40	35	15	35	30	25	20
5	ПС 1	80	30	20	10	10	20	30
	ПС 2	60	15	15	35	35	35	15
	ПС 3	40	20	10	30	30	20	10

Стоимость потерь электрической энергии в проводах ЛЭП определяется законом Джоуля-Ленца:

$$C_3 = 3 \cdot I^2 \cdot R \cdot \text{Ц} \cdot t, \quad (6.33)$$

где I, R – ток в фазе линии и ее активное сопротивление; Ц – отпускная цена одного кВт·ч электрической энергии, руб.; t – число часов работы линии в год.

Активное сопротивление линии можно определить по выражению:

$$R = \rho \frac{l}{S}, \quad (6.34)$$

где ρ – удельное сопротивление материала провода.

Сечения проводников в сетях высокого напряжения определяются по экономической плотности тока:

$$S = \frac{I}{J_{\text{ЭК}}}. \quad (6.35)$$

Преобразуя выражение (6.31) с учетом (6.32)-(6.35), получим:

$$Z_{\text{пр}} = a \cdot l + \left(\frac{b}{J_{\text{ЭК}}} + 3 J_{\text{ЭК}} \cdot \rho \cdot \Pi \cdot t \right) \cdot I \cdot l. \quad (6.36)$$

Для n участков электрической сети суммарные приведенные затраты определяются выражением:

$$Z_{\text{пр}} = \sum_{i=1}^m \sum_{j=1}^n k_1 \cdot l_{ij} + \sum_{i=1}^m \sum_{j=1}^n k_2 \cdot l_{ij} \cdot I_{ij}, \quad (6.37)$$

где I_{ij} – ток, потребляемый j -м потребителем с i -й подстанции; l_{ij} – расстояние от j -го потребителя до i -й подстанции; k_1, k_2 – постоянные коэффициенты.

Для достижения минимальных приведенных затрат достаточно минимизировать второй член уравнения (6.37), при этом значение коэффициента k_2 можно не учитывать. С учетом того, что ток в линии прямо пропорционален передаваемой по ней мощности, получим выражение целевой функции решаемой задачи:

$$Z = \sum_{i=1}^m \sum_{j=1}^n l_{ij} \cdot P_{ij} \rightarrow \min. \quad (6.38)$$

Требуется найти минимум целевой функции при следующих ограничениях

Суммарная мощность, потребляемая всеми потребителями с одной подстанции должна быть равна располагаемой мощности подстанции

$$\sum_{j=1}^n P_{ij} = P_i, \quad i = 1, 2 \dots m. \quad (6.39)$$

Суммарная мощность, передаваемая всеми подстанциями одному потребителю должна быть равна требуемой мощности этого объекта

$$\sum_{i=1}^m P_{ij} = P_j, \quad j = 1, 2 \dots n. \quad (6.40)$$

Величина мощности, передаваемой по линии должна быть положительной

$$P_{ij} \geq 0. \quad (6.41)$$

Выражения (6.38)–(6.41) являются математической моделью решаемой задачи.

Перед решением задачи необходимо проверить баланс располагаемой и требуемой мощности и при необходимости привести задачу к сбалансированной. Разработайте экранную форму математической модели задачи и найдите ее решение средствами MS Excel.

Список литературы к 6 главе

1. Лещинская Т.Б., Наумов И.В. Электроснабжение сельского хозяйства: Учебник. – М.: КолосС, 2008.
6. Лыкин А.В. Электрические системы и сети: Учебное пособие. – М.: Университетская книга; Логос, 2008.
3. Методы оптимизации: Учебник / Под ред. В.С. Зарубина, А.П. Крищенко. – 2-е изд., стереотип. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2003.
4. Методы оптимизации режимов энергосистем / В.М. Горнштейн, Б.П. Мирошниченко, А.В. Пономарев. – М.: Энергоатомиздат, 2001.
5. Математическое моделирование в системах электроснабжения. Методические указания к выполнению лабораторных работ по курсу «Математическое моделирование в системах электроснабжения» для студентов 3 курса, обучающихся по направлению 140200 «Электроэнергетика» / С.Г.Обухов. – Изд-во Томского политехнического университета, 2011.
6. Савоськин Н.Е. Надежность электрических систем: Учебное пособие. – Издательство Пензенского государственного университета, Пенза, 2004.
7. Холмский В.Г. Расчет и оптимизация режимов электрических сетей (специальные вопросы): Учебное пособие. – М.: Высшая школа, 2005.
8. Шапкин А.С., Мазаева Н.П. Математические методы и модели исследования операций: Учебник. 2-е изд. – М.: Издательско-торговая корпорация «Дашков и К», 2005.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1. СЛОЖНЫЕ ЭЛЕКТРОЭНЕРГЕТИЧЕСКИЕ СИСТЕМЫ	4
1.1. Понятие системы. Общие определения	4
1.2. Иерархия энергетических систем	9
1.3. Принятие решений при управлении сложными системами	12
1.4. Особенности системы электроснабжения	14
1.5. Особенности автоматизированного управления в СЭС	19
1.6. Анализ причин неопределенности информации в ЭЭС	23
1.7. Компоненты сложной системы	27
1.8. Понятие модели и моделирования. Классификация моделей	33
1.9. Задачи оптимизации и основные понятия	40
Список литературы к 1 главе	44
2. ОСНОВЫ ТЕОРИИ ОПТИМИЗАЦИИ	45
2.1. Характеристика задач оптимизации	45
2.2. Обозначения и терминология	46
2.3. Основные этапы решения задач оптимизации	51
3. МЕТОДЫ РЕШЕНИЯ ЗАДАЧ ИНВАРИАНТНОЙ ОПТИМИЗАЦИИ	54
3.1. Методы безусловной одномерной оптимизации	55
3.2. Поисковые методы одномерной минимизации	57
3.3. Методы с использованием производных	66
3.4. Методы многомерной безусловной оптимизации	68
3.5. Поисковые методы безусловной оптимизации	70
3.6. Методы с использованием производных	83
4. МЕТОДЫ РЕШЕНИЯ ЗАДАЧ СТАТИЧЕСКОЙ УСЛОВНОЙ ОПТИМИЗАЦИИ	91
4.1. Линейное программирование	91
4.2. Нелинейное программирование	99
4.3. Практическое приложение задач оптимизации	102
4.4. Практические методы инвариантной оптимизации	105
Список литературы к 2–4 главам	140
5. ВЕРОЯТНОСТНЫЕ МОДЕЛИ ОПТИМИЗАЦИИ РЕЖИМОВ ЭЛЕКТРОЭНЕРГЕТИЧЕСКИХ СИСТЕМ	142
5.1. Характеристика информации для расчетов установившихся режимов ЭЭС	142
5.2. Расчет установившихся режимов ЭЭС в вероятностной постановке ..	151
5.3. Оценивание состояния ЭЭС	165
5.4. Регулирование напряжения в распределительных сетях методом характеристического узла	172
Список литературы к 5 главе	181

6. ПРАКТИКА РЕШЕНИЯ ЗАДАЧ ОПТИМИЗАЦИИ	182
6.1. Общие положения	182
6.2. Решение задач линейного программирования	184
6.3. Решение задач нелинейного программирования	192
6.4. Решение транспортной задачи	193
6.5. Определение оптимального количества трансформаторов цеховых подстанций	198
6.6. Оптимальное распределение компенсирующих устройств в радиальной схеме электроснабжения	201
6.7. Оптимальное распределение компенсирующих устройств в магистральной схеме электроснабжения	203
6.8. Оптимальное проектирование схемы электроснабжения	205
Список литературы к 6 главе	207

Учебное издание

Сазыкин Василий Георгиевич
Кудряков Александр Георгиевич

ОПТИМИЗАЦИИ СИСТЕМ ЭНЕРГОСНАБЖЕНИЯ

Учебное пособие

Компьютерная верстка
Багметов А.А, Кучеренко Д.Е.

Подписано в печать 01.12.2016. Формат 60×84/8.
Усл. печ.л. -24,0, уч изд.л. -14,5
Тираж 100 экз. Заказ №

Типография ООО «КРОН»
350044, г.Краснодар, ул.Алма-Атинская, 57, офис 3.