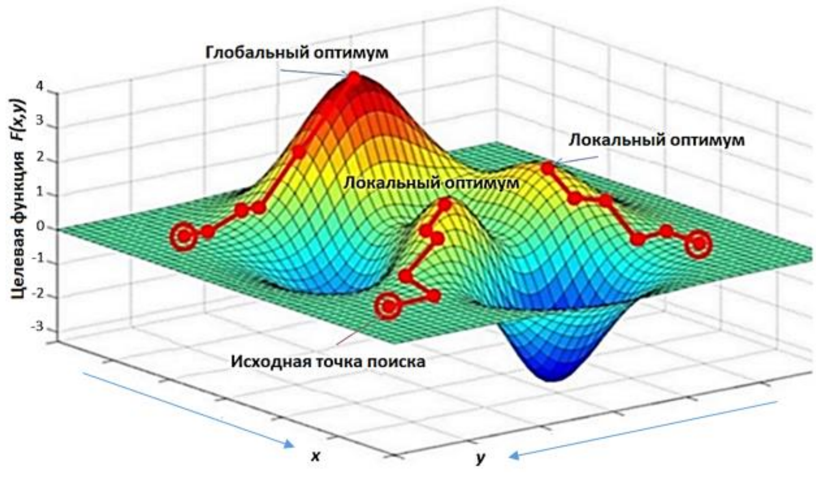
Кафедра конструирования
и производства радиоаппаратуры

Оптимизация при проектировании РЭС



Томск 2018

Кобрин Юрий Павлович
Оптимизация при проектировании РЭС. Методические указания к лабораторной работе по дисциплине «Информационные технологии проектирования электронных средств» для студентов специальности «11.03.03 «Конструирование и технология электронных средств». - Томск: Томский государственный университет систем управления и радиоэлектроники (ТУСУР), кафедра КИПР, 2018. – 36 с.

Даётся обзор основных методов оптимизации при проектировании конструкций РЭС. Приводится методика практического формирования математических моделей РЭС, используемых для решения задач оптимизации. Рассматриваются важнейшие приёмы использования системы автоматизации научно-технических расчётов MathCAD для решения задач оптимизации РЭС.

Методические указания предназначены для помощи в подготовке бакалавров и магистрантов в области разработки и моделирования РЭС различного назначения, выполнения курсовых и дипломных проектов и может быть использованы студентами других специальностей радиотехнического профиля.

Содержание

1 Цели работы

– изучение методов решения задач оптимизации при проектировании радиоэлектронных средств (РЭС);

– приобретение навыков в построении математических моделей РЭС, используемых для решения задач оптимизации;

– приобретение практических навыков в работе с системой автоматизации научно-технических расчётов ***MathCAD*** для решения задач оптимизации.

2 Порядок выполнения работы

1) Перед выполнением этой работы следует ознакомиться с теоретическим материалом, а также с примерами решения типовых задач оптимизации, представленными в разделе 5.

2) Ответить на контрольные вопросы.

3) Сформировать целевую функцию для решения задачи по своему варианту.

4) Составить программу минимизации целевой функции в системе ***MathCAD***, предусмотрев графический показ целевой функции на экране.

5) Выбрать начальное приближение и провести оптимизацию. Оценить результаты оптимизации, представив в масштабе эскиз разработанной конструкции.

6) Оформить и защитить отчёт по лабораторной работе у преподавателя.

3 Контрольные вопросы

Ответьте письменно на следующие контрольные вопросы:

1) Как формулируется задача оптимизации?
2) Какими подходами можно решить задачу оптимизации?
3) Что называется, целевой функцией?
4) Что называется, проектными параметрами?
5) Какое различие между методами условной и безусловной оптимизации?
6) В чем сущность метода штрафных функций?
7) Всегда ли оптимальные значения целевой функции совпадают со значениями глобальных экстремумов? А со значениями локальных экстремумов?
8) В чем сущность метода общего поиска?
9) В чем сущность метода золотого сечения?
10) В чем сущность методов координатного спуска (подъёма)?
11) В чем сущность методов градиентного поиска?
12) В чем сущность методов случайного поиска?

4 Защита отчёта

Отчёт должен состоять из следующих разделов:

- Тема и цель работы.
- Индивидуальное задание
- Ответы на контрольные вопросы
- Тексты программ вместе с вводимыми исходными данными и ограничениями
- Результаты выполнения индивидуального задания
- Выводы

Для получения зачёта при защите отчёта по работе студент должен:

- обосновать избранные математические модели и ограничения;
- проявить знание использованных методов оптимизации и их работоспособность на своих примерах;
- продемонстрировать навыки работы в среде ***MathCAD.***

5 Основы теории нелинейной оптимизации

5.1 Постановка задачи

Оптимальное решение при проектировании - лучшее в том или ином смысле проектное решение, допускаемое обстоятельствами.

В подавляющем большинстве случаев одна и та же техническая задача может быть решена несколькими способами, приводящими не только к различным выходным характеристикам, схемам и конструкциям, но даже и к физическим принципам, положенным в основу построения объекта. Вследствие этого на разных этапах проектирования РЭС почти всегда у проектировщиков возникает необходимость выбора из множества допустимых проектных решений наилучшего (оптимального) варианта конструкции РЭС или технологического процесса его изготовления, удовлетворяющих предъявленным требованиям.

Очевидно, что изделие или технологический процесс, выгодно отличающееся от аналогичных изделий и процессов, будет пользоваться на рынке большим спросом. В этом и состоит смысл поиска оптимальных решений.

Оптимальное решение может быть получено с помощью разных подходов.

1. Следование установленным нормативам и стандартам, в которых заложен опыт «предыдущих поколений». Эти источники создавались всевозможными путями: систематизацией опыта, экспериментальной отработкой. Конечно, не исключены случаи, когда мотивы разработчиков стандартов не очень-то и легко понять. Тем не менее, в подавляющем большинстве случаев соблюдение нормативов есть один из наиболее надёжных путей проектирования.

2. Использование конструкторами инженерной интуиции, практических навыков, опыта предыдущих разработок. Иногда этот путь даёт неплохие результаты, особенно когда решаются концептуальные вопросы. К сожалению, алгоритмы генерации новых знаний, несмотря на определённый прогресс в деталях, до сих пор не созданы.

3. Использование алгоритмов математического нелинейного проектирования, основанных на использовании математических моделей проектируемых объектов и соответственных им методов оптимизации.

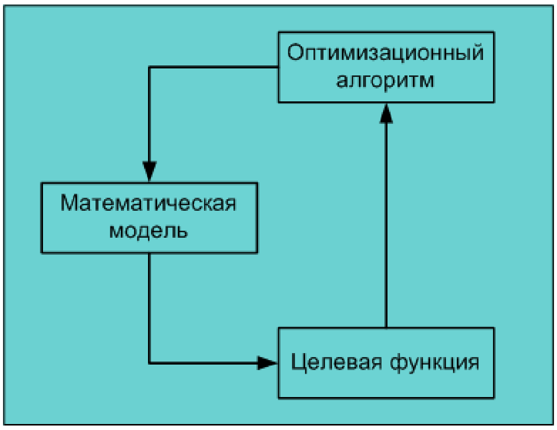
Первые два подхода требуют экспериментальной проверки (обычно вручную) большого числа вариантов, что весьма трудоёмко, требует значительных затрат времени и чаще всего экономически нецелесообразны. Нередко экспериментальное исследование объекта проектирования приводит к его разрушению. Кроме того, невозможно с уверенностью сказать, что полученное экспериментальными способами проектное решение действительно оптимально.

Наиболее продуктивным способом проектирования оптимальных РЭС являются методы оптимизации на основе *нелинейного математического программирования*, позволяющие получить за короткое время наилучший вариант конструкции из всех возможных вариантов. Безусловно, результаты будут еще лучше, если они будут разумно дополнены позитивными качествами первых двух подходов.

5.2 Основные понятия нелинейного программирования

Для проведения оптимизации необходимы математическая модель объекта, целевая функция и оптимизационный алгоритм.

При математическом моделировании на компьютерах исходный объект заменяют его «образом» — **математической моделью**. Для работы оптимизационных алгоритмов необходима математическая модель, представленная в виде *целевой функции* (*функции качества*), позволяющей количественно сравнить два альтернативных решения.



Целевая функция (критерий качества) $\boldsymbol{F}(\boldsymbol{X})$ - это выражение, количественно характеризующее качество проектируемого объекта, значение которого нужно минимизировать или максимизировать, подобрав необходимые значения совокупности (множества, вектора) n проектных параметров:

$$\boldsymbol{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n\}$$

Проектные параметры - независимые переменные величины, которые полностью и однозначно количественно описывают объект проектирования. *Число проектных параметров* характеризует степень сложности данной задачи проектирования.

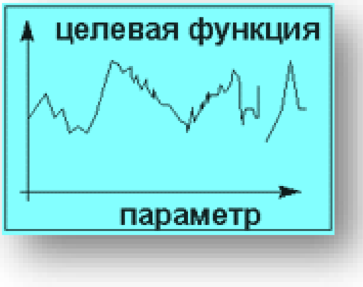
Значения проектных параметров устанавливаются в процессе оптимизации.

Целевая функция часто может приобретать самые неожиданные формы. Например, её не всегда удаётся выразить в замкнутой математической форме, в других случаях она может представлять собой кусочно-гладкую функцию. Для задания целевой функции иногда может потребоваться таблица справочных технических данных (например, зависимости некоторого параметра от времени, температуры и т.п.) или потребуется провести эксперимент.

В ряде задач оптимизации требуется введение более одной целевой функции. Часто одна из них оказывается несовместимой с другой. Примером служит проектирование бортовой радиоаппаратуры, когда одновременно требуется обеспечить максимальные прочность и надёжность, минимальный вес и минимальную стоимость. В таких случаях конструктору необходимо ввести систему приоритетов и поставить в соответствие каждой целевой функции некоторый безразмерный множитель. В результате формируется «функция компромисса», позволяющая в процессе оптимизации пользоваться одной составной целевой функцией.

В качестве проектных параметров могут служить любые основные или производные величины, служащие для количественного описания оптимизируемого объекта. Например, это могут быть неизвестные значения:

- размеров объекта,
- массы объекта;
- параметров электрорадиоэлементов,
- токов и напряжений в схеме электрической принципиальной,
- времени на проектирование или изготовление,
- температуры в корпусе РЭС,
- концентрации веществ и т.п.

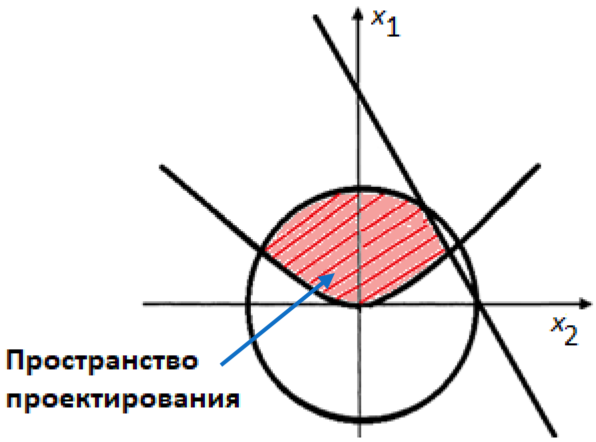


Отметим, что целевая функция и проектные параметры **обязательно** должны иметь *количественные значения*. В то же время многие показатели качества и проектные параметры (например, эстетические и эргономические характеристики, удовлетворение, которое испытывает купивший изделие покупатель и т.п.) количественно охарактеризовать очень непросто.

Проектные параметры не могут принимать произвольные значения, так как всегда есть ограничения на их физическую реализацию, условия эксплуатации и т.п.

В ряде случаев проектные параметры принимают только дискретные или целые значения. Примерами могут служить дискретность номинальных значений параметров радиокомпонентов, выпускаемых промышленностью, количество людей для выполнения какой-нибудь работы, число крепёжных болтов и т.п. Иногда проектные параметры имеют только два значения - да или нет.

Область, в которой все параметры проектирования принимают допустимые значения, называется *пространством проектирования* или *областью допустимых решений*. Пространство проектирования обычно ограничено рядом уравнений условий-ограничений, связанных с физической сущностью задачи. Ограничения могут быть настолько многообразными и взаимоисключающими, что задача оптимизации может не иметь ни одного удовлетворительного решения.



Ограничения области проектирования могут включать неравенства, равенства и требование целочисленности решения.

Под *задачей оптимизации* обычно понимают процесс или последовательность операций, позволяющих получить уточнённое значение вектора X^* из множества X допустимых решений, которому соответствует минимальное (или максимальное) значение целевой функции. И хотя конечной целью оптимизации является отыскание наилучшего, т.е. «оптимального», решения, обычно приходится довольствоваться улучшением известных решений, а не доведением их до безупречности. Поэтому под оптимизацией подразумевают скорее стремление к совершенству, которое, может быть, и не будет достигнуто.

Задача поиска минимума и максимума целевой функции $F(X)$ называется задачей *поиска экстремума*.

Локальный оптимум (экстремум) - точка пространства проектирования, в которой целевая функция имеет наибольшее (или наименьшее, если целевая функция минимизируется) значение по сравнению со значениями во всех других точках её ближайшей окрестности.

Глобальный оптимум – это оптимальное решение для всего пространства проектирования.

Если существует только один проектный параметр, который характеризует качество РЭС, то формируется однокритериальная (одномерная) целевая функция, зависящая от этого параметра. При этом другие параметры подпадают под категорию ограничений.

Однокритериальную целевую функцию можно представить кривой на плоскости (Рис. 5.1).

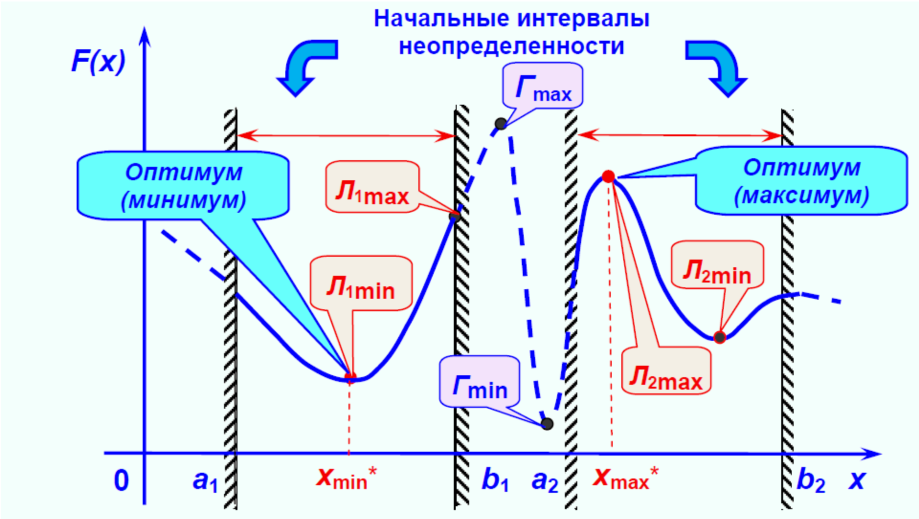


Рис. 5.1 - Одномерная целевая функция:

x^*_{max}, x^*_{min} - оптимальные значения проектного параметра целевой функции, x^*_{max}- если ищется максимум и x^*_{min}, если ищется минимум;

$Л_{max}$, $Л_{min}$- локальные максимумы и минимумы;

$Г_{max}$, $Г_{min}$ - глобальные максимумы и минимумы;

a, b - границы интервалов неопределённости.

На этом рисунке пространство проектирования определено двумя участками {a_1, b_1} и {a_2, b_2}. Заметим, что глобальные экстремумы (максимум и минимум) в данном примере вообще находятся *вне пространства проектирования* и не могут таким образом быть оптимальными решениями.

Пространство проектирования может содержать множество локальных оптимумов. Следует заметить, что нередко из-за ограничений оптимальное значение целевой функции достигается не там, где её поверхность имеет нулевой градиент и лучшее решение соответствует одной из границ области проектирования. Нужно соблюдать осторожность, чтобы не принять первый же локальный оптимум за оптимальное решение задачи.

В приведённом примере оптимальные решения являются локальными $\boldsymbol{x}^*_{\boldsymbol{max}} \Rightarrow Л^*_{\boldsymbol{2max}}$ или $\boldsymbol{x}^*_{\boldsymbol{min}} \Rightarrow Л^*_{\boldsymbol{1min}}$, так как глобальные оптимумы не допустимы, они находятся вне областей проектирования

Целевая функция *F(X)* может зависеть от нескольких проектных параметров. В этом случае ее можно представить некоторой (n+1)-мерной поверхностью (n – число проектных параметров).

Если проектных параметров два (x_1 и x_2), то целевая функция будет изображаться поверхностью в трёхмерном пространстве (Рис. 5.2). Линии на поверхности, моделирующей целевую функцию, являются её сечениями. Они показаны исключительно для наглядности и функционального значения не имеют.

При трёх и более проектных параметрах поверхности, задаваемые целевой функцией, называются гиперповерхностями и не поддаются изображению обычными средствами.

Топологические свойства поверхности целевой функции играют большую роль в процессе оптимизации, так как от них зависит выбор наиболее эффективного алгоритма.

Например, целевая функция **F** на Рис. 5.2 состоит из трёх участков. Они образуют зону, в которой функция является негладкой, а также имеет разрыв.

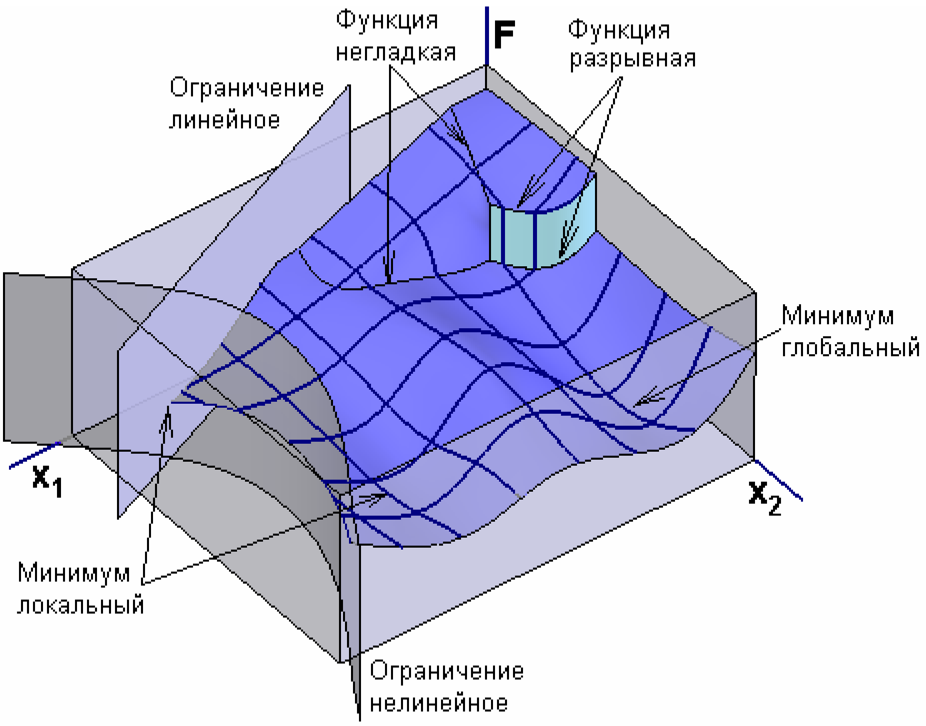


Рис. 5.2 – Поверхность, образуемая двумерной целевой функцией

На этом же рисунке показаны три типа ограничений-неравенств:

1) группа ограничений, определяющих интервалы изменения переменных x_1 и x_2 (вертикальные стенки, имеющие в основании прямоугольник);

2) группа ограничений, состоящая из единственного линейного ограничения (вертикальная неортогональная стенка);

3) группа, также представленная единственным нелинейным ограничением (вертикальная криволинейная стенка).

В пространстве проектирования может быть один или несколько экстремумов.

Унимодальная целевая функция - в пространстве проектирования существует *лишь один экстремум*, определяемый в ходе оптимизации.

Неунимодальная целевая функция - в пространстве проектирования существует *несколько экстремумов*, и в ходе оптимизации необходимо определить - какой из них оптимальный.

В нашем примере (см. Рис. 5.2) в пределах пространства проектирования целевая функция имеет несколько минимумов, следовательно, она неунимодальная. Тот, где функция имеет наименьшее значение - глобальный. Остальные минимумы являются локальными. При этом один из них расположен на границе допустимой области, в месте пересечения двух ограничений.

Чаще всего целевую функцию отображают в виде *карты линий уровня*, которые представляет собой проекцию трёхмерной поверхности на двухмерную плоскость. Карты линий уровня изображаются в виде семейства линий, получающихся в результате сечения поверхности горизонтальными плоскостями.

Если необходимо быстро представить себе общую трёхмерную картину данных, этот способ, вероятно, будет менее нагляден, чем график поверхности. Тем не менее, с помощью карт линий уровня удобно детально исследовать характерные особенности формы поверхностей целевых функций, например таких, как многоэкстремальность, разрывность, овражность[1], наличие седловых точек (Рис. 5.3) и т.п.

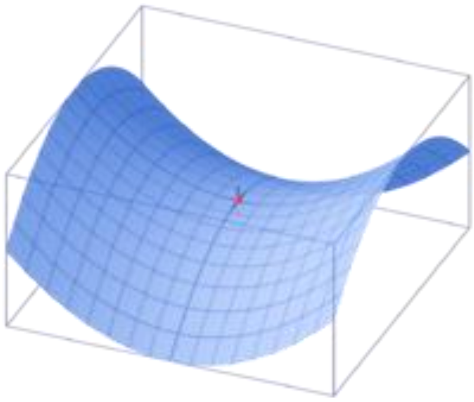


Рис. 5.3 - Седловая точка

[1] Рельеф такой целевой функции похож на овраг. Склоны оврага крутые (частные производные характеризующие их велики), а дно имеет незначительный протяжённый наклон (частные производные характеризующие его на порядок меньше)

На Рис. 5.4 та же целевая функция (см. Рис. 5.2) представлена с отображёнными линиями равного уровня[2]. Эти линии получаются при пересечении поверхности отклика функции двух переменных $F(x_1, x_2)$ плоскостью, параллельной плоскости координат (x_1, x_2). На рисунке видно, что в гладкой части функции имеется зона перегиба - это следствие того, что функция является знакопеременной по второй производной, т. е. имеет как выпуклые, так и вогнутые участки.

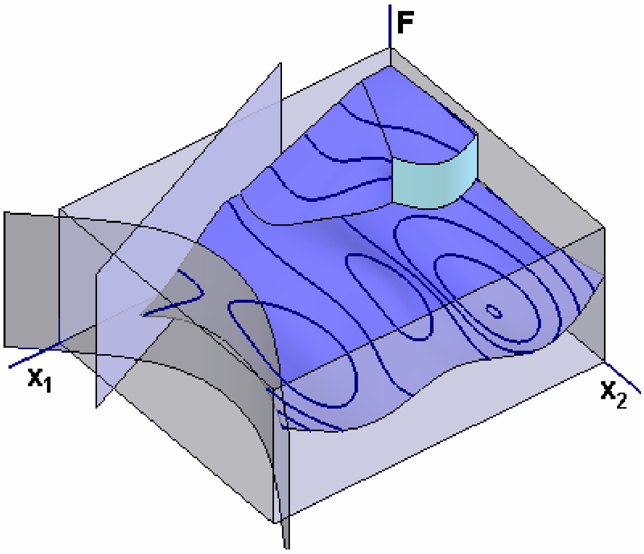


Рис. 5.4 - Линии уровня (изолинии) целевой функции

Вид целевой функции сверху в виде *карты линий одного уровня* показан на Рис. 5.5. Это традиционное представление целевой функции для задачи оптимизации.

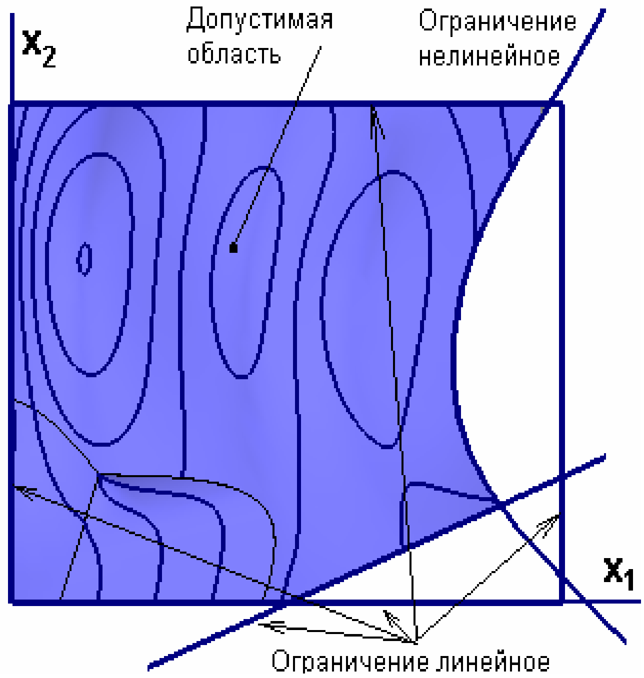


Рис. 5.5 – Представление двумерной целевой функции картой линий уровня

На приведённых рисунках совершенно намеренно не представлены *ограничения-равенства,* фигурирующие в общей постановке задачи минимизации. Попытка задать их в программе существенно усложняет как сам алгоритм, так и процесс его эксплуатации. Если какое-либо из этих ограничений-равенств можно разрешить относительно одного из проектных параметров, то стараются исключить данный параметр из процесса оптимизации. Тем самым уменьшается число измерений пространства проектирования и упрощается решение задачи.

[2] **Линия уровня или изолиния** — условное обозначение на карте, чертеже, схеме или графике, представляющее собой линию, в каждой точке которой измеряемая величина сохраняет одинаковое значение. Изолинии — способ представления скалярной функции от двух переменных на плоскости.

5.3 Классификация методов оптимизации

При выборе метода оптимизации более эффективным считается метод, позволяющий получить оптимальное решение с заданной максимальной погрешностью ε как можно скорее - *с минимальным числом вычислений целевой функции.*

Первостепенную роль при выборе метода оптимизации, которых к настоящему времени разработано огромное количество [1,2,3,4,5,6,7] играют *топологические свойства поверхности целевой функции*. Некоторые из этих методов приспособлены для поиска *минимума целевой функции*, другие - для поиска *максимума*. Эти различия не имеют принципиального значения, так как задача на поиск максимума легко превращается в задачу поиска минимума путём *замены знака на обратный* у целевой функции (Рис. 5.6).

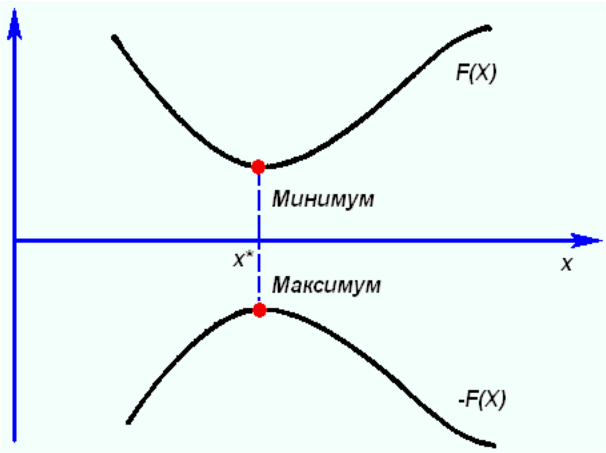


Рис. 5.6 - Изменением знака целевой функции на противоположный задача на поиск максимума превращается в задачу на поиск

Очевидно, если функция $F(x)$ имеет минимум в точке x^*, то функция $-F(x)$ имеет максимум в той же точке.

Существующие методы оптимизации классифицируют по следующим характерным признакам.

По количеству варьируемых переменных:

– методы *одномерного поиска* (один проектный параметр);

– методы *многомерного поиска* (несколько проектных параметров).

По способу изменения варьируемых переменных:

– детерминированные[3] методы (выбор параметров в соответствии с некоторым законом);

– метод *случайного поиска*, называемый также методом Монте-Карло. Основан на том, что при одном и том же числе испытаний вероятность получения решения, близкого к оптимальному, при случайном поиске больше, чем при последовательном переборе через равные интервалы изменения отдельных параметров.

По порядку используемых производных целевой функции:

– методы *нулевого порядка* (без вычисления производных);

– методы *первого порядка* (используются производные первого порядка);

– методы *второго порядка* (используются производные второго порядка).

[3] **Детерминированный** - точный, определённый, ясный, конкретный, чёткий.

По отношению к рельефу целевой функции:

– методы поиска *локальных экстремумов* сходятся к какому-нибудь локальному экстремуму целевой функции. В случае унимодальной целевой функции, этот экстремум единственен, и будет глобальным максимумом/минимумом.

– методы поиска *глобальных экстремумов*: имеют дело с многоэкстремальными целевыми функциями. При глобальном поиске основной задачей является выявление тенденций глобального поведения целевой функции;

– методы поиска оптимума при *овражном* характере целевой функции.

По отношению к ограничениям:

– методы *безусловной оптимизации* (целевая функция не ограничена);

– методы *условной оптимизации* (целевая функция ограничена).

5.4 Методы барьерных штрафных функций

Большинство численных методов оптимизации предназначены для отыскания абсолютных экстремумов, т.е. относятся к группе методов *безусловной оптимизации*. Многие из них обладают высокой эффективностью и существенно облегчают решение задачи оптимального проектирования, но получаемое с их помощью решение может лежать вне пространства проектирования ограниченного условиями

$$G_j(x_i) \geq 0, \qquad j = 1, 2, \dots, J,$$

где J – количество ограничений.

Разумно постараться использовать при решении задач условной оптимизации, более важных с практической точки зрения, простые и результативные методы безусловной оптимизации. Для этого в целевую функцию вводится добавка, характеризующая уровень нарушения ограничений.

Выражение для новой составной результирующей целевой функции $M(x_i)$ приобретает вид:

$$M(x_i) = F(x_i) + \sum_{j=0}^{n} \varphi\big[G_j(x_i)\big],$$

где новая неограниченная целевая функция $M(x_i)$ образуется сложением ограниченной целевой функции рассматриваемой задачи $F(x_i)$ и штрафной функции $\varphi[G_j(x_i)]$, учитывающей ограничения, заданные неравенствами.

Поверхность, описываемая новой штрафной функцией, должна препятствовать выходу траекторий поиска из пространства проектирования. Штрафная функция $\varphi(x_i)$ равна нулю во всех точках пространства проектирования, удовлетворяющих условиям $G_j(x_i)$, и должна стремиться к бесконечности, чем больше эти условия не удовлетворяются. Тем самым на каждое решение, попадающее в «запрещённую область», налагается «штраф».

Если все условия-ограничения удовлетворяются, то функции $M(x_i)$ и $F(x_i)$ имеют, очевидно, один и тот же минимум. Эти алгоритмы получили название методов *барьерных штрафных функций*.

Функцию штрафа можно формировать по-разному.

Первый вариант - увеличение функции происходит при приближении к ограничению изнутри допустимой области, достигая на активном ограничении бесконечной величины. Постоянная штрафная функция, является, по сути, бесконечным барьером (Рис. 5.7). При попытке попадания пробной точки в недопустимую область постоянная штрафная функция принимает «бесконечное» значение (в программных реализациях для этого некоторой переменной присваивается соответствующий признак), после чего принимается решение о том, как действовать дальше.

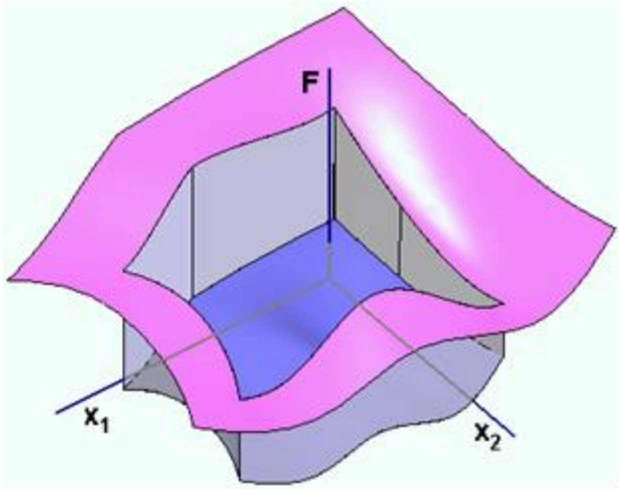


Рис. 5.7 - Постоянная штрафная функция - бесконечный барьер

Недостатком такого простейшего варианта является невозможность «участия» недопустимых точек в последующем анализе и, соответственно, склонность алгоритмов к зацикливанию (это формальное описание, реальная ситуация более сложна), если минимум лежит на границе.

Второй вариант - внутри и на границе допустимой области «добавка» равна нулю, а затем, при выходе за границу допустимой зоны, она начинает возрастать. Использование абсолютной функции штрафа показано на Рис. 5.8, где поверхность, образованная участками пирамиды и конуса общего вида и которая теоретически уходит в бесконечность, обрезана.

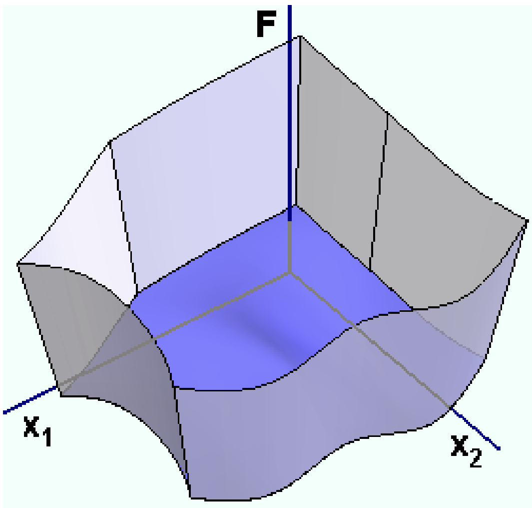


Рис. 5.8 - Абсолютная штрафная функция

По сути, данный вид штрафа - это сумма абсолютных величин невязок нарушенных ограничений. Тем самым на каждое решение, попадающее в «запрещённую область», налагается «штраф».

Теоретически параметр штрафа должен зависеть от поведения исходной целевой функции. Если, например, она «быстро» убывает в точке на линии ограничения, то «малый» штраф не сможет компенсировать это убывание, и программа продолжит поиск за пределами допустимой области. Если же штраф «слишком» большой, то возникают проблемы, присущие абсолютным штрафам. Поэтому в ситуациях, когда обнаруженный программой условный оптимум лежит на одном или нескольких ограничениях, возможен «небольшой» выход за границы допустимой области.

Преимущество метода штрафных функций в том, что он позволяет при наличии ограничений пользоваться любыми методами безусловной оптимизации, не заботясь о выполнении ограничений.

Заметим, что выбор методов оптимизации, способа формирования штрафных функций, начальных приближений и т.п. как в смысле программирования, так и при решении, является своего рода искусством и достигается опытом.

5.5 Методы одномерной оптимизации

5.5.1 Постановка задачи

Методы минимизации функций одной переменной, являясь наиболее простым типом оптимизационных задач, достаточно часто встречаются в инженерной практике. Кроме самостоятельного значения, они занимают важное место в теории оптимизации, так как часто являются важным элементом при реализации всевозможных более сложных методов многопараметрической оптимизации.

Задачу одномерной оптимизации можно представить следующим образом. Значения единственного проектного параметра x должны быть заключены в интервале, $\boldsymbol{a} \leq \boldsymbol{x} \leq \boldsymbol{b}$. Приступая к решению задачи, мы ничего не знаем о характере изменения целевой функции. Интервал значений x, в котором заключён оптимум, будем называть *интервалом неопределённости.* В начале процесса оптимизации интервал неопределённости имеет длину $\boldsymbol{b} - \boldsymbol{a}$ (см. Рис. 5.1).

Существует множество способов систематического сужения интервала неопределённости. Рассмотрим наиболее популярные методы оптимизации и отметим их главные особенности.

5.5.2 Метод общего поиска

Наиболее простым способом сужения интервала неопределённости для одномерной целевой функции является **метод общего поиска**. При этом интервал неопределённости (Рис. 5.9) делится на N равных частей с последующим вычислением значений целевой функции в полученных узлах сетки и выборе из них минимального (или максимального).

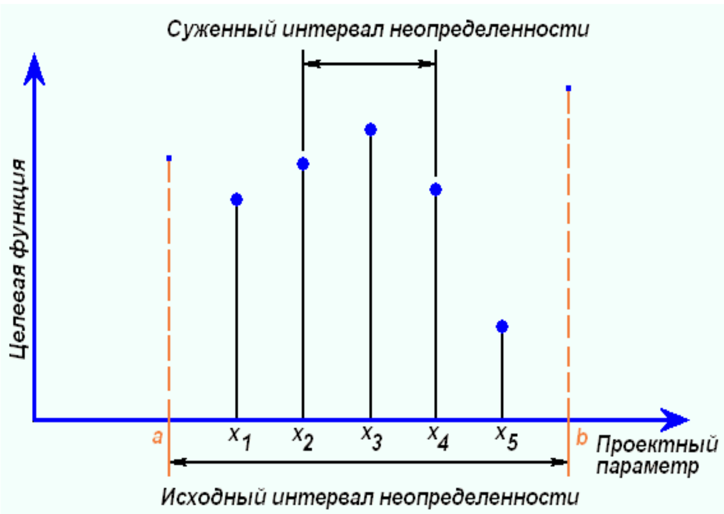


Рис. 5.9 - Метод общего поиска (ищем максимум)

В результате интервал неопределённости сужается *до двух шагов сетки*. Обычно говорят о *дроблении* интервала неопределённости, которое характеризуется коэффициентом $R(N)$. Разделив интервал неопределённости на N частей, получим $N+1$ узел, и тогда коэффициент дробления для метода общего поиска

$$R(N) = \frac{2}{N+1}$$

Таблица 5.1 – Число вычислений целевой функции N от коэффициента дробления $R(N)$

R(N)	0.1	0.01	0.001	0.0001	0.00001
N	19	199	1999	19999	199999

Таблица 5.1 показывает, что эффективность метода общего поиска при уменьшении интервала неопределённости стремительно падает.

Можно модифицировать метод общего поиска следующим образом.

Чтобы получить коэффициент дробления $R(N) = 0.01$, вычислим вначале на исходном интервале неопределённости функцию в 19 точках, и получим коэффициент дробления $R(N) = 0.1$. Затем, вычислив ещё 19 значений функции *на сокращённом интервале неопределённости*, получим $R(N) = 0.01$. Таким образом, чтобы в 100 раз уменьшить интервал неопределённости потребуется всего 38, а не 199 вычислений.

Несмотря на значительные затраты времени из-за необходимости многократного вычисления целевой функции, метод общего поиска имеет важное достоинство – **он универсален**, так как используемая целевая функция может быть *неунимодальной* и не критична к разрывности.

5.5.3 Метод золотого сечения

В методе золотого сечения целевая функция вычисляется в точках интервала неопределённости, расположенных в пропорциях золотого сечения таким образом, чтобы каждое вычисленное значение целевой функции давало новую полезную информацию.

Золотое сечение — деление величины на две части в таком отношении, при котором *меньшая часть так относится к большей*, как *большая ко всей величине*. Отношение большей части к меньшей в этой пропорции выражается квадратичной иррациональностью

$$\varphi = \frac{\sqrt{5}+1}{2} \approx 1{,}6180339887\ldots$$

и, наоборот, отношение меньшей части к большей

$$\frac{1}{\varphi} = \frac{\sqrt{5}-1}{2} \approx 0{,}6180339887\ldots$$

При использовании метода золотого сечения целевая функция обязательно должна быть *унимодальной*.

Сущность метода золотого сечения состоит в следующем. Интервал неопределённости делится на две неравные части так, что отношение длины большого отрезка к длине всего интервала равно отношению длины меньшего отрезка к длине большего отрезка. *На* Рис. 5.10 показан интервал неопределённости Z, состоящий из отрезков z_1 и z_2, отношение длин которых определяется правилом золотого сечения $\frac{z_1}{Z} = \frac{z_2}{z_1}$, и кроме того $z_1 + z_2 = Z$.

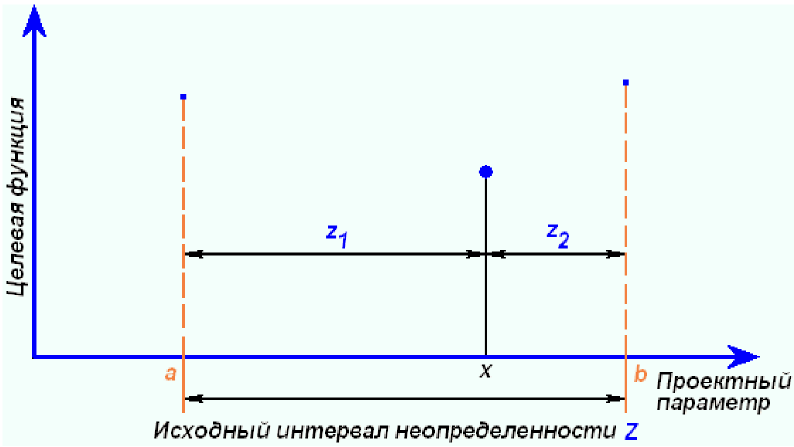


Рис. 5.10 - Обозначения, используемые в методе золотого сечения

В соответствии с методом золотого сечения (Рис. 5.11) внутри отрезка [*a, b*] выделяют две промежуточные симметрично расположенные точки x_1и x_2 в «золотом» отношении»:

$$x_1 = a + (1-r)(b-a) \approx a + 0.382 \cdot (b-a) = a + 0.382 \cdot Z;$$
$$x_2 = a + r \cdot (b-a) \approx a + 0.618 \cdot (b-a) = a + 0.618 \cdot Z.$$

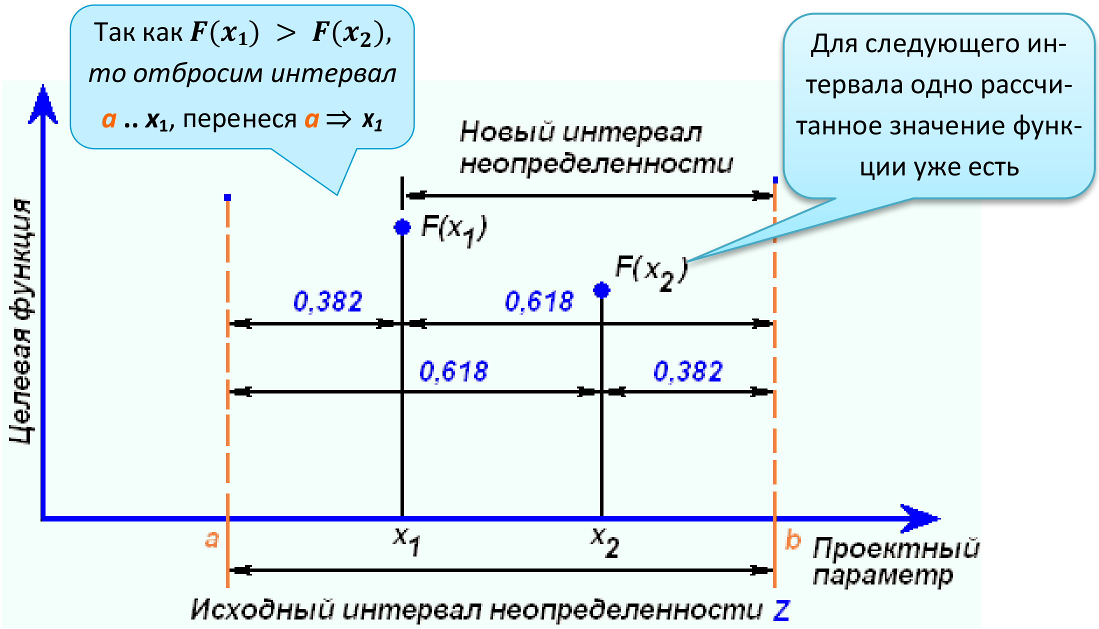


Рис. 5.11 - Метод золотого сечения (ищем минимум)

В точках x_1 и x_2 вычисляют значения целевой функции $F(x_1)$ и $F(x_2)$. Сравнение этих значений (ещё раз напомним, что функция унимодальна – минимум лишь один!) даёт возможность отбросить либо кусок интервала [a .. x_1], либо кусок интервала [x_2 .. b].

Но в том и другом случае одна из точек x_1или x_2 обязательно окажется внутренней для нового интервала, делящей его две части, длины которых снова относятся в «золотом» отношении». Таким образом, начиная со второго шага, следует вычислять *лишь одно новое значение целевой функции* в недостающей «золотой» точке нового интервала неопределённости. Положение этой точки определяется правилом золотого сечения по приведённым выше формулам.

Например, если у целевой функции на Рис. 5.11 мы ищем минимум, и выполнилось условие $F(x_1) > F(x_2)$, то следует отбросить интервал $[a .. x_1]$, а левую границу a перенести в точку x_1 . При каждом таком отбрасывании исходный интервал неопределённости уменьшается в $1 / 0.618$ раза.

После вычисления *N* значений целевой функции коэффициент дробления интервала неопределённости составляет $R(N) = 0.618^{N-1}$.

Используя эту формулу нетрудно показать, что для достижения той же величины коэффициента дробления $R(N)$ = 0.01 по методу золотого сечения достаточно сделать только лишь 11 вычислений целевой функции.

5.6 Методы многомерной оптимизации

5.6.1 Общие соображения

На первый взгляд может показаться, что различие между методами одномерного и многомерного поиска состоит лишь в том, что первые требуют меньшего объёма вычислений и что в принципе методы, пригодные для функций одной переменной, можно применять и для функций многих переменных. Однако это не совсем так, поскольку многомерное пространство качественно отличается от одномерного.

Прежде всего, с увеличением числа измерений *уменьшается вероятность унимодальности целевой функции* (Рис. 5.12). Кроме того, множество элементов, образующих многомерное пространство, гораздо мощнее множества элементов одномерного пространства. Объем вычислений, необходимых для сужения интервала неопределённости в многомерном пространстве, является *степенной функцией,* показатель которой равен размерности пространства.

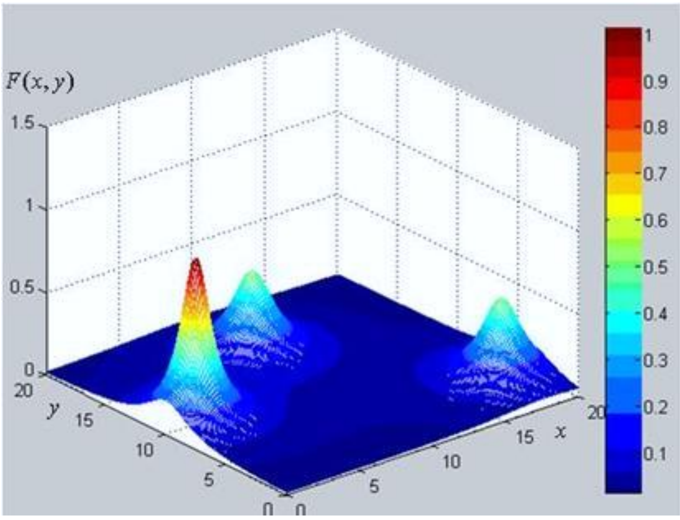


Рис. 5.12- Неунимодальная целевая функция

Например, если в случае одномерного пространства для достижения $R(N)$= 0.1 методом общего поиска требуется вычислить 19 значений целевой функции, то в случае двумерного пространства это число составляет 361, трёхмерного - 6859, четырёхмерного - 130321, а пятимерного - 2476099! Поскольку при выборе оптимальном проектировании нередко приходится иметь дело с пятью и более проектными параметрами, серьёзность трудностей, обусловленных многомерностью, становится бесспорной.

Методы оптимизации в многомерном пространстве делятся на две большие группы - *прямые* и *косвенные*. **Прямые методы** основаны на сравнении вычисляемых значений целевой функции в различных точках. **Косвенные методы** - на использовании необходимых и достаточных условий математического определения максимума и минимума функции.

Стратегия прямых методов - постепенное приближение к оптимуму; при использовании косвенных методов стремятся найти решение, не исследуя неоптимальные точки.

5.6.2 Метод покоординатного спуска (подъёма)

Одним из простейших прямых методов поиска оптимума считается **метод покоординатного спуска (подъёма, если ищется максимум)**, который заключается в сведении многомерной задачи к последовательным одномерным задачам, которые решаются методами минимизации функции одной переменной, например, методом золотого сечения. В этом методе в качестве очередного направления спуска выбирается направление одной из координатных осей.

Логическим развитием рассмотренной выше методов одномерного поиска было бы последовательное изменение каждого проектного параметра до тех пор, пока не будет достигнут минимум (максимум) целевой функции. По завершении этой процедуры для всех проектных параметров можно вернуться к первому параметру и посмотреть, нельзя ли ещё более усовершенствовать решение. Направление поиска выбирают поочерёдно вдоль координатных осей каждого проектного параметра до тех пор, пока не будет достигнут максимум (минимум) целевой функции.

На Рис. 5.13 показана двумерная целевая функция, заданная линиями уровня.

Пусть значения всех проектных параметров фиксированы, кроме последнего. Принимая в качестве нулевого приближения исходную точку x_1^0 ищем минимум или максимум (точка $M1$) вдоль координаты $\boldsymbol{x_2}$ одним из методов одномерной оптимизации. Затем, сохраняя найденное значение координаты x_2^1 постоянным, ищут оптимум вдоль координаты $\boldsymbol{x_1}$ (точка $M2$). Вновь возвращаемся к параметру $\boldsymbol{x_2}$ и т.д. Выход из этого итерационного процесса осуществляется по достижению точки оптимума с координатами x_1^*, x_2^* (приращения координат точки меняются не более чем на величину заданной абсолютной погрешности ε).

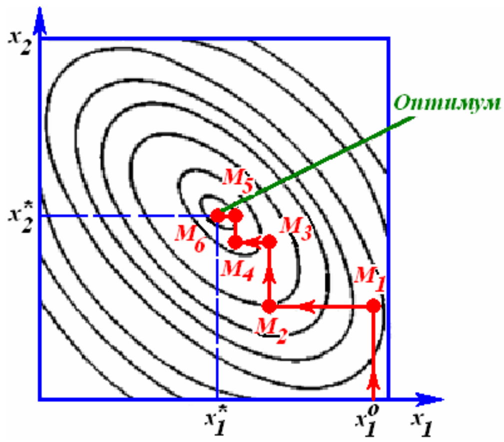


Рис. 5.13 - Поиск оптимума целевой функции методом покоординатного спуска

Особенно эффективен метод, если линии уровня близки по форме к окружностям или эллипсам, оси которых параллельны осям координат (Рис. 5.14). Физически это означает, что проектные параметры практически независимы друг от друга. Если же эти оси наклонены к осям координат (как на Рис. 5.13), то приходится много раз изменять направление поиска, и эффективность алгоритма снижается, так как для нахождения оптимума приходится вычислять гораздо больше значений целевой функции.

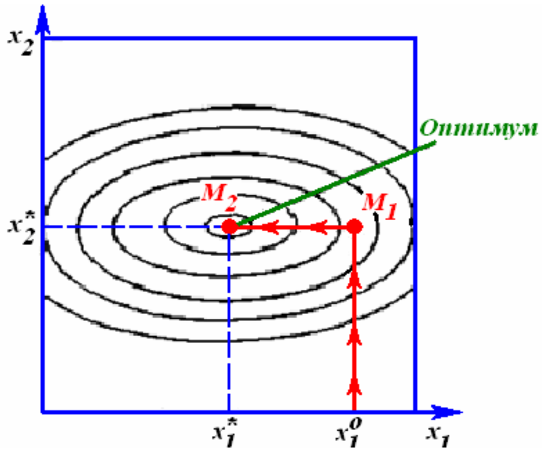


Рис. 5.14 - Линии уровня целевой функции по форме близки к окружностям или эллипсам, оси которых параллельны осям координат

Метод покоординатного подъёма практически неприменим, если линии уровня имеют точки излома (Рис. 5.15). Поскольку линии уровня такого типа очень часто встречаются в инженерной практике, то прежде, чем воспользоваться указанным методом, следует убедиться, что решаемая задача не имеет подобного недостатка. Можно однако попытаться выбрать начальное приближение так, чтобы линия излома на траектории поиска не встретилась.

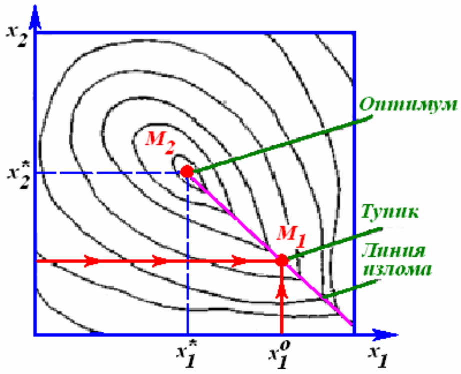


Рис. 5.15 - Целевая функция имеет линию излома

Несмотря на перечисленные недостатки, метод покоординатного подъёма (спуска) часто используют на первой стадии решения задачи, применяя затем более сложные методы. К достоинствам метода покоординатного подъёма следует отнести возможность использования простых алгоритмов одномерного поиска, таких, как метод золотого сечения.

5.6.1 Градиентные методы

Во многих алгоритмах многомерной оптимизации так или иначе используется информация о градиентах.

Градиент - вектор, своим направлением указывающий направление наискорейшего возрастания некоторой величины F, значение которой меняется от одной точки пространства к другой, а по величине (модулю) равный быстроте роста этой величины в этом направлении.

Чтобы наглядно представить идею метода, рассмотрим следующий простой пример. Представим себе, что альпинисту завязали глаза и предложили добраться до вершины «унимодальной» горы. Даже ничего не видя, он может это сделать, если все время будет двигаться вверх. Хотя любая ведущая вверх тропа, в конечном счёте, приведёт его к вершине, кратчайшей из них будет самая крутая, если, правда, альпинист не натолкнётся на вертикальный обрыв[4], который придётся обходить.

Метод оптимизации, в основу которого положена идея движения по самой крутой тропе, называется методом наискорейшего подъёма или наискорейшего спуска. Вектор градиента перпендикулярен линии уровня и указывает направление к новой точке в пространстве проектирования (Рис. 5.16).

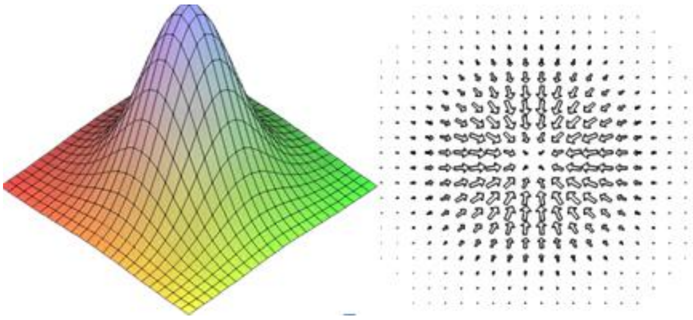


Рис. 5.16 - Операция градиента преобразует холм (слева), если смотреть на него сверху, в поле векторов (справа). Видно, что векторы направлены «в горку» и тем длиннее, чем круче наклон

Чтобы лучше понять идею градиентных методов, подробнее остановимся на свойствах градиентов. Рассмотрим систему независимых единичных векторов $\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3}, \ldots, \overrightarrow{e_N}$, направленных вдоль осей координат $x_1, x_2, x_3, \ldots, x_N$, являющихся в то же время проектными параметрами. Вектор градиента произвольной целевой функции $F(x_1, x_2, x_3, \ldots, x_N)$ имеет вид

$$\mathbf{grad}\, F = \nabla F = \frac{\partial F}{\partial x_1} \cdot \overrightarrow{e_1} + \frac{\partial F}{\partial x_2} \cdot \overrightarrow{e_2} + \cdots + \frac{\partial F}{\partial x_N} \cdot \overrightarrow{e_N}$$

где частные производные вычисляются в рассматриваемой точке. Этот вектор направлен вверх, в направлении подъёма; обратный ему вектор (антиградиент) указывает направление спуска. Единичный вектор градиента часто представляют в виде

$$v_1\overrightarrow{e_1} + v_2\overrightarrow{e_2} + v_3\overrightarrow{e_3} + \cdots + v_N\overrightarrow{e_N},$$

где

$$v_i = \frac{\dfrac{\partial F}{\partial x_i}}{\sum_{i=1}^{N} \left[\left(\dfrac{\partial F}{\partial x_j} \right)^2 \right]^{1/2}}.$$

[4] Математическим эквивалентом обрыва на поверхности, образуемой целевой функцией, являются те её места, где поставлены условные ограничения.

Иногда характер целевой функции бывает достаточно хорошо известен, чтобы можно было вычислить компоненты вектора градиента путём непосредственного дифференцирования. Если таким способом частные производные получить не удаётся, то можно найти их приближенные значения в непосредственной окрестности рассматриваемой точки:

$$\frac{\partial F}{\partial x_i} = \frac{F(x_1, x_2, \dots, x_i + \Delta, \dots, x_N) - F(x_1, x_2, \dots, x_i, \dots, x_N)}{\Delta}$$

Здесь Δ - небольшое смещение в направлении x_i. Эту формулу часто называют «приближением секущей». Полученную информацию о направлении градиента можно использовать различным образом для построения алгоритма поиска.

Пошаговые градиентные методы при отсутствии ограничений сводятся к следующему.

Этап 1. Пусть нам известны начальные значения проектных параметров $x_1^0, x_2^0, \dots, x_i^0, \dots, x_N^0$

Этап 2. Определяется направление наискорейшего изменения целевой функции $F(x_1^0, x_2^0, \dots, x_i^0, \dots, x_N^0)$, т.е. ее градиента $grad\ F(x_1^0, x_2^0, \dots, x_i^0, \dots, x_N^0)$ в исходной точке.

Этап 3. Осуществляется перемещение из точки $X^0 = \{x_1^0, x_2^0, \dots, x_i^0, \dots, x_N^0\}$ в точку $X^1 = \{x_1^1, x_2^1, \dots, x_i^1, \dots, x_N^1\}$ по направлению противоположному градиенту.

Этап 4. В точке $X^1 = \{x_1^1, x_2^1, \dots, x_i^1, \dots, x_N^1\}$ определяется новое направление градиента целевой функции $grad\ F(x_1^1, x_2^1, \dots, x_i^1, \dots, x_N^1)$ и осуществляется перемещение в точку $X^2 = \{x_1^2, x_2^2, \dots, x_i^2, \dots, x_N^2\}$ и т.д. Этот этап повторяется до тех пор, пока не будет выполнено условие окончания поиска (изменение проектных параметров или целевой функции станут меньше заданных погрешностей).

Ряд вариантов методов градиентного поиска основан на смещении на постоянный шаг в направлении градиента с последующим вычислением целевой функции. Если её величина оказывается больше предыдущей, вычисляется градиент в новой точке, и вся процедура повторяется, причём часто при этом шаг увеличивают. Если же величина целевой функции не изменяется или убывает, то шаг смещения от предыдущей точки уменьшают и повторяют всю процедуру вычислений. Так поступают до тех пор, пока дальнейшее уменьшение шага уже не приводит к улучшению результата.

Особенностью *метода наискорейшего спуска* (***Ошибка! Источник ссылки не найден.***) является движение с оптимальным шагом *h*, рассчитанным с помощью одномерной минимизации целевой функции по *h* вдоль антиградиентного направления. Действительно, если в какой-либо точке направление поиска определено, то целевая функция может считаться функцией переменного параметра *h,* характеризующего положение новой точки на заданной прямой.

Поэтому алгоритм метода наискорейшего спуска содержит следующие этапы.

Этап 1. Определяется направление градиента целевой функции $F(x_1, x_2, \dots, x_i, \dots, x_N)$ в исходной точке.

Этап 2. Нахождение одним из методов одномерного поиска оптимального шага вдоль антиградиентного направления. Величина шага *h* определяется из условия минимума функции *dF.*

Этап 3. Осуществляется движение с этим шагом по лучу, направленному по антиградиенту целевой функции $grad\ F(x_1^0, x_2^0, \dots, x_i^0, \dots, x_N^0)$ в точке $X^0 = \{x_1^0, x_2^0, \dots, x_i^0, \dots, x_N^0\}$ до тех пор, пока функция F не достигнет минимума на этой прямой (точка $X^1 = \{x_1^1, x_2^1, \dots, x_i^1, \dots, x_N^1\}$).

Этап 4. В точке $X^1 = \{x_1^1, x_2^1, \dots, x_i^1, \dots, x_N^1\}$ определяется новое направление градиента целевой функции $grad\ F(x_1^1, x_2^1, \dots, x_i^1, \dots, x_N^1)$ и осуществляется перемещение в точку $X^2 = \{x_1^2, x_2^2, \dots, x_i^2, \dots, x_N^2\}$ и т. д. до тех пор, не будет выполнено условие прекращения поиска.

Пример траектории поиска этим методом показана на Рис. 5.17. Из рисунка видно, что движение вдоль одного направления прекращается, когда линия направления поиска становится касательной к какой-либо линии равного уровня. Каждое новое направление движения к экстремуму ортогонально предшествующему.

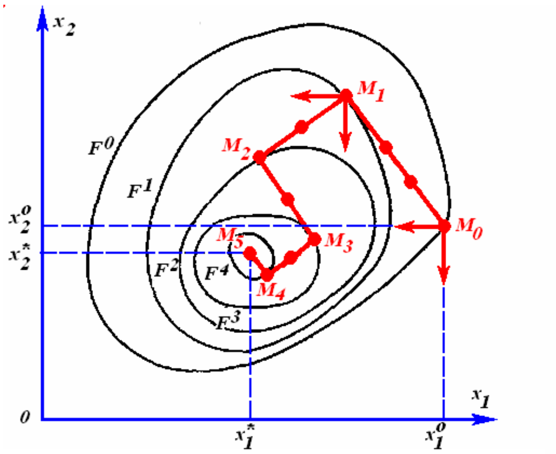


Рис. 5.17 - Траектория поиска методом наискорейшего подъёма (спуска)

Как видно на Рис. 5.18, для мультимодальных функций градиентные методы позволяют найти лишь локальные оптимумы. Поэтому, если характер поверхности недостаточно хорошо известен, следует испробовать несколько исходных точек и убедиться, что во всех случаях получается одно и то же оптимальное решение.

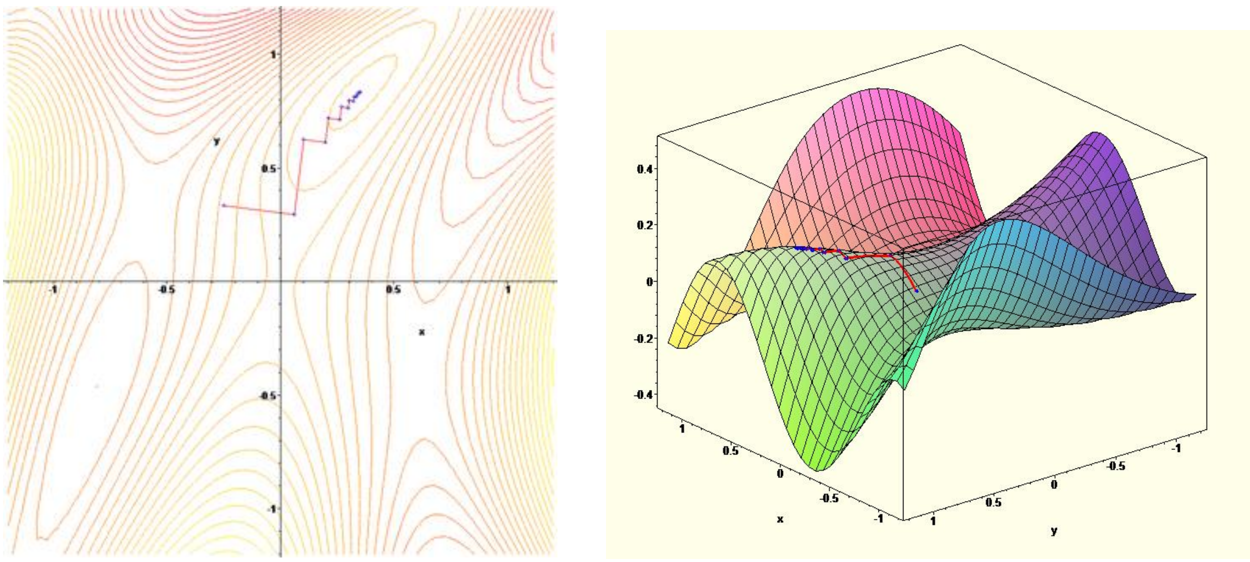


Рис. 5.18 - Метод градиентного спуска оказывается очень медленным при движении вдоль оврага.

Другой причиной, снижающей эффективность градиентных методов, являются изломы линий уровня целевой функции. Так как такие точки соответствуют разрыву в наклоне линий контура, то здесь возможны ошибки в определении направления дальнейшего поиска. Поэтому поиск может замедлиться и идти зигзагами поперёк линии излома, а время, необходимое для получения решения, будет столь велико, что расчёт придётся прекратить. В действительности большинство исследуемых поверхностей имеет одну или более линий излома, которые нередко проходят через точку оптимума. Поэтому, наткнувшись на линию излома, следует в дальнейшем двигаться вдоль неё.

Из всех методов локальной оптимизации методы градиентного спуска наиболее просты в реализации. Однако они характеризуются довольно слабыми условиями сходимости. Шаг градиентного метода часто используется как часть других, более совершенных методов оптимизации, например, в методе Флетчера-Ривса [4,5]. Заметим, что метод градиентного спуска оказывается очень медленным при движении вдоль оврага (см. Рис. 5.18). Более эффективным считается метод сопряжённых градиентов [6].

5.6.1 Метод случайного поиска

Если целевая функция неунимодальна, то использовать методы покоординатного и градиентного спуска нельзя, так как они ищут только локальные экстремумы целевой функции. Какой из множества локальных экстремумов будет найден - зависит от координат начальной точки поиска.

Можно воспользоваться методом общего поиска, однако даже в случае многомерного пространства как было показано в примере (см. Таблица 5.1), число вычислений целевой функции для нахождения оптимума с заданной точностью очень велико. Было показано, что для многомерного пространства это число растёт как степенная функция, показатель степени которой равен размерности вектора проектных параметров, что непомерно много даже для быстродействующих компьютеров.

Оригинальный подход, позволяющий обойти эту трудность, основан на использовании методов случайного поиска (методов Монте-Карло[5]).

Сущность метода Монте-Карло заключается в том, что на каждом k-цикле с помощью специальной программы формируется последовательность псевдослучайных координат для N проектных параметров $X^{(k)} = \left\{x_1^{(k)}, x_2^{(k)}, \dots, x_N^{(k)}\right\}$ с равномерным законом распределения. Перебирая таким образом координаты случайных точек (Рис. 5.19), компьютер отбрасывает те из них, которые оказались вне области допустимых значений переменных (вне проектного пространства). Для каждой точки, попавшей в пространство проектирования, вычисляется значение целевой функции. Если значение целевой функции меньше (или больше, если ищется максимум) предыдущего минимума, то оно запоминается как точка с наименьшим значением.

---

[5] **Монте-Карло** — столица княжества Монако, славится своими казино с различными карточными играми, рулетками и др., в которых выигрыши зависят от случая. **Методы Монте-Карло** —группа численных методов для изучения случайных процессов. В этих методах процесс оптимизации моделируется при помощи программы-генератора псевдослучайных величин, с помощью которых вычисляются вероятностные характеристики решаемой задачи.

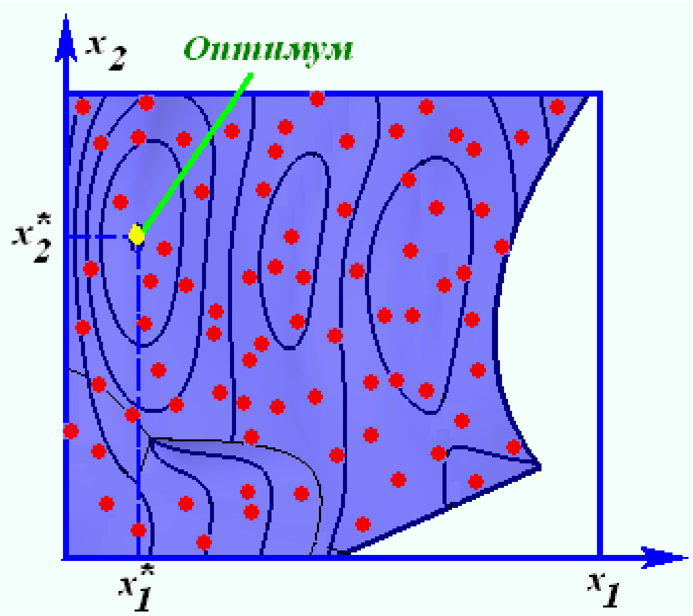


Рис. 5.19 - Метод случайного поиска (метод Монте-Карло)

Результат, полученный с помощью метода Монте-Карло характеризуется вероятностью P того, что при данном числе случайных проб K, расположение точки оптимума будет определено с точностью $R(N)$ где $R(N)$ - объем N-мерного куба, выраженный в долях от общего объёма поиска. Если делается K случайных проб, то вероятность того, что одна из этих проб попадёт в заданную область, может быть определена по формуле

Для нахождения числа случайных проб K, необходимых для того, чтобы с заданной вероятностью P расположение точки оптимума определялось с точностью $R(N)$, можно воспользоваться формулой

$$K = \frac{log(1-P)}{log\big(1-R(N)\big)}$$

Таблица 5.2 показывает, сколько ячеек K надо выбрать случайным образом, чтобы обеспечить заданную вероятность P при заданной выборке наиболее перспективных ячеек. Из таблицы видно, что при случайной выборке 44 ячеек вероятность достижения $R(N)$ = 0.1 составит 99%. Вспомним, что для 100%-ного обеспечения целевую функцию в случае пяти переменных пришлось бы вычислить 2 476 099 раз!

Таблица 5.2 – Зависимость необходимого числа проб в случайной выборке для обеспечения вероятности P достижения заданного коэффициента дробления $R(N)$

Коэффициент дробления $R(N)$	**Необходимое число проб для достижения заданного коэффициента дробления при заданной вероятности**			
	0.80	**0.90**	**0.95**	**0.99**
0.100	16	22	29	44
0.050	32	25	59	90
0.010	161	230	299	459
0.005	322	460	598	919

Существует громадное многообразие алгоритмов случайного поиска, что обусловлено их простотой, устойчивой работой, отсутствием необходимости вычисления производных, наглядностью и удовлетворительной и хорошей сходимостью, особенно на задачах большой размерности.

Достоинства методов случайного поиска:

– они пригодны для любой целевой функции независимо от того, является она унимодальной или нет.

– вероятность успеха при попытках не зависит от размерности рассматриваемого пространства.

Основные недостатки метода:

– большая трудоёмкость и длительность поиска экстремума;

– возможность ошибки при попадании в область локального экстремума.

Хотя этот метод случайного поиска часто не позволяет непосредственно найти оптимальное решение, он создаёт подходящие предпосылки для применения в дальнейшем других методов поиска. Поэтому его часто применяют в сочетании с одним или несколькими методами других типов.

5.7 Численные методы оптимизации на основе MathCAD

5.7.1 Введение

В MathCAD задачи оптимизации решаются с помощью встроенных функций [8,9,10]. Причём решаются только задачи поиска локального экстремума. Чтобы найти глобальный максимум (или минимум), требуется либо сначала вычислить все их локальные значения и потом выбрать из них наибольший (наименьший), либо предварительно просканировать с некоторым шагом рассматриваемую область, чтобы выделить из неё подобласть наибольших (наименьших) значений функции и осуществить поиск глобального экстремума, уже находясь в его окрестности.

5.7.2 Поиск локального экстремума функции одной переменной без ограничений

Построим график функции $f(x) = x^4 + 5x^3 - 10x$ на интервале (-5, 2). Как видно из Рис. 6.2, она имеет глобальный максимум на левой границе интервала, глобальный минимум, локальный максимум, локальный минимум и локальный максимум на правой границе интервала (в порядке слева направо).

Для решения задач поиска максимума и минимума в MathCAD имеются встроенные функции *Minerr, Minimize и Maximize*. Все они используют градиентные численные методы.

Если никаких дополнительных условий при их использовании не вводится, поиск экстремумов выполняется для любых значений x от -∞ до ∞.

Всем аргументам целевой функции f предварительно следует присвоить начальные значения, которые будут восприниматься как начальные приближения для градиентного метода.

$Minimize(f, var1, var2, \ldots)$ возвращает вектор значений переменных $var1, var2, \ldots$, при которых функция f достигает минимума, при условии выполнения ограничений в блоке решения. Возвращает скаляр, если указана одна переменная, иначе возвращает вектор решений (Рис. 6.2).

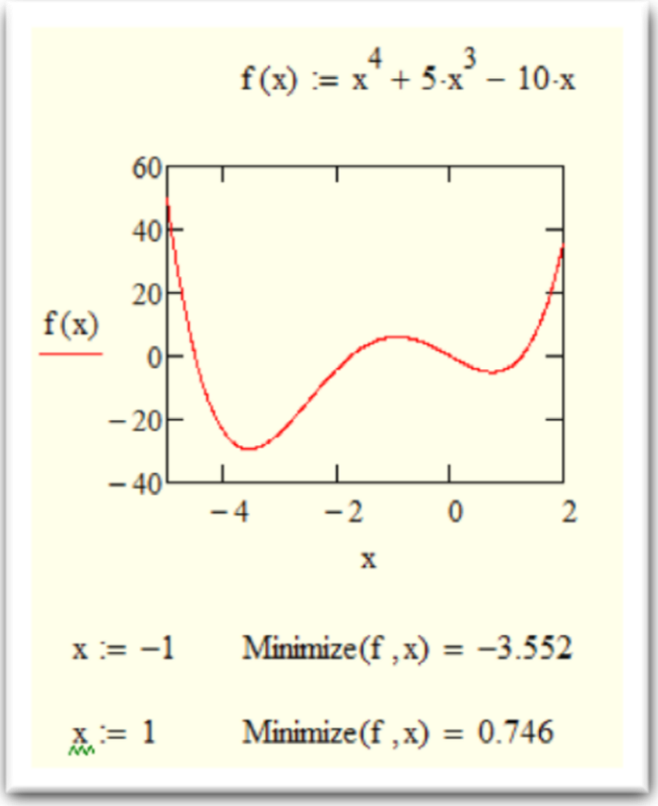


Рис. 5.20 – Нахождение локальных минимумов функции $f(x) = x^4 + 5x^3 - 10x$

$\mathrm{Maximize(f, var1, var2, \ldots)}$ - возвращает вектор значений переменных *var1, var2, ...,* при которых функция *f* достигает максимума, при условии выполнения ограничений в блоке решения (Рис. 6.3).

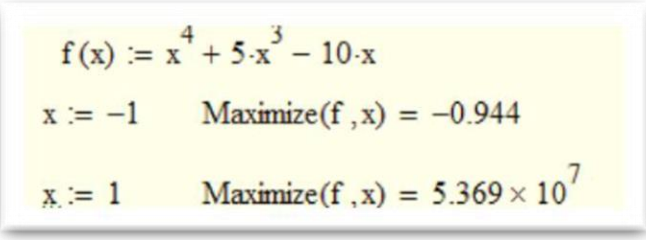


Рис. 5.21 – Нахождение локальных максимумов функции $f(x) = x^4 + 5x^3 - 10x$

Как видно из рисунков, существенное влияние на результат оказывает выбор начального приближения, в зависимости от чего в качестве ответа выдаются различные локальные экстремумы. Численный метод (см.) вообще не справляется с задачей, поскольку начальное приближение $x = 1$ выбрано далеко от области локального максимума, и поиск решения уходит в сторону увеличения $f(\mathrm{x})$.

5.7.3 Поиск экстремума функции одной переменной при ограничениях

В задачах на условный экстремум функции минимизации и максимизации должны быть включены в вычислительный блок, т.е. им должно предшествовать ключевое слово *Given* (*Учитывая*).

В промежутке между *Given* и функцией поиска экстремума с помощью условных операторов записываются логические выражения (неравенства, уравнения), задающие ограничения на значения аргументов минимизируемой функции. В листинге Рис. 6.4 показаны пример поиска условных экстремумов на различных интервалах, границы которых определены соответствующими неравенствами. Сравните результаты работы этого листинга с двумя предыдущими.

$f(x) := x^4 + 5 \cdot x^3 - 10 \cdot x$

$x := -1$

Given $-5 < x < -2$ Minimize$(f, x) = -3.552$

$x := 1$

Given $x > 0$ Minimize$(f, x) = 0.746$

$x := -10$

Given $-3 < x < 0$ Maximize$(f, x) = -0.944$

Рис. 5.22 - Пример поиска условных экстремумов функции

Не забывайте о важности выбора правильного начального приближения и в случае задач на условный экстремум.

5.7.4 Нахождение экстремума функции многих переменных

Вычисление экстремума функции многих переменных не несёт принципиальных особенностей по сравнению с функциями одной переменной. Поэтому ограничимся примером (Рис. 6.5) нахождения максимума и минимума функции, показанной в виде графиков трёхмерной поверхности и линий уровня.

$f(x,y) := 2(x - 5.07)^2 + (y - 10.03)^2 - 0.2(x - 5.07)^3$

Координаты исходной точки поиска

$x := 1$ $y := 1$

Given

$0 < x < 15$ $0 < y < 20$ Ограничения

$\text{Minimize}(f, x, y) = \begin{pmatrix} 5.07 \\ 10.03 \end{pmatrix}$

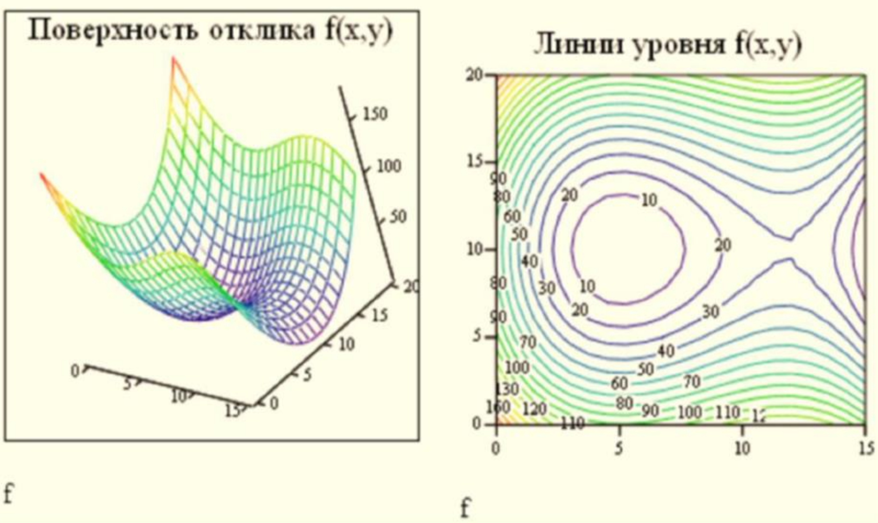


Рис. 5.23 – Нахождение экстремума функции двух переменных

Обратите внимание, как с помощью неравенств, введённых логическими операторами, задаётся пространство проектирования (область определения целевой функции) на плоскости (x, y).

6 Практические задания

6.1 Пример формирования целевой функции

Необходимо определить оптимальные размеры x_1, x_2, x_3 прямоугольного блока РЭС (Рис. 6.1) объёмом V = 1000 см^3, для которого корпус должен иметь минимальную массу m. Корпус выполняется из листового материала с постоянной толщиной стенок h = 0.2 см. Толщина блока x_2 не должна превышать $x_2 \le 8$ см.

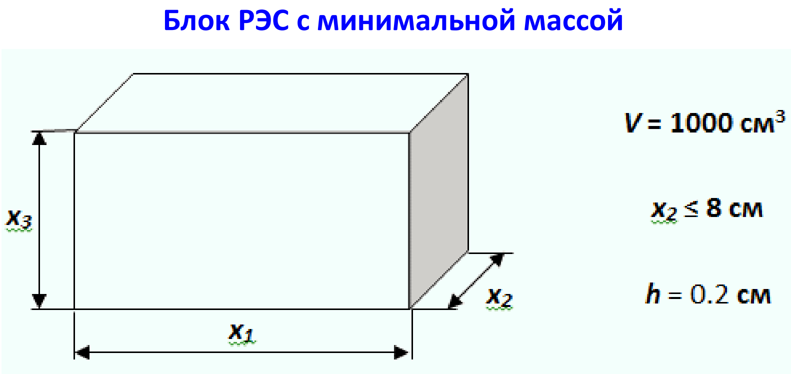


Рис. 6.1 - Оптимизация размеров блока РЭС

Решение.

Проектные параметры: размеры x_1, x_2, x_3.

Масса блока: $V = \gamma \cdot h \cdot S$

Учитывая, что удельный вес γ *и* толщина стенок материала корпуса блока РЭС h постоянные коэффициенты, достаточно минимизировать площадь *поверхности* корпуса S. Целевая функция имеет вид:

$$S(x1,\ x2,\ x3) = 2 \cdot (x1 \cdot x2 + x2 \cdot x3 + x1 \cdot x3) \quad (6.1)$$

Ограничение-неравенство на размер x_1:

$$0 < x2 \le 8. \quad (6.2)$$

Ограничение-равенство:

$$V = x1 \cdot x2 \cdot x3 = 1000. \quad (6.3)$$

Из последнего ограничения-равенства легко выразить один из размеров, лучше всего, *x3*:

$$x3 = V/(x1 \cdot x2) \quad (6.4)$$

и исключить его из проектный параметров, упрощая целевую функцию и вводя вместо *x3* в неё заданной объём (а это константа, а не переменная!):

$$S(x1,\ x2) = 2 \cdot (x1 \cdot x2 + \frac{V}{x_1} + \frac{V}{x_2}). \quad (6.5)$$

При использовании классических *методов математического анализа*, оптимальные значения проектных параметров определяют приравниванием нулю полученных в аналитической форме выражений для первых производных целевой функции:

$$\begin{aligned} \frac{\partial S(x_1,x_2)}{\partial x_1} &= 2 \cdot x_2 - 2\frac{V}{x_1^2} = 0; \\ \frac{\partial S(x_1,x_2)}{\partial x_2} &= 2 \cdot x_1 - 2\frac{V}{x_2^2} = 0. \end{aligned} \quad (6.6)$$

Решая систему алгебраических уравнений (6.6) получим, что экстремум (глобальный минимум) получается при $x_1 = x_2 = x_3 = 10$. Однако при подобном решении не выполняется ограничение-неравенство (6.2) и, следовательно, оно непригодно. Результат подтверждает, что не всегда глобальный минимум является оптимальным решением.

6.2 Индивидуальные задания

Определить оптимальные размеры x_1, x_1, x_1 блока РЭС заданной формы объёмом ***V = 1000 см³*** для которого корпус должен иметь минимальную массу ***m***. Корпус выполняется из листового материала с постоянной толщиной стенок ***h* = 0.2** см. Глубина блока не должна превышать $x_1 \le 8$ см (Рис. 6.2).

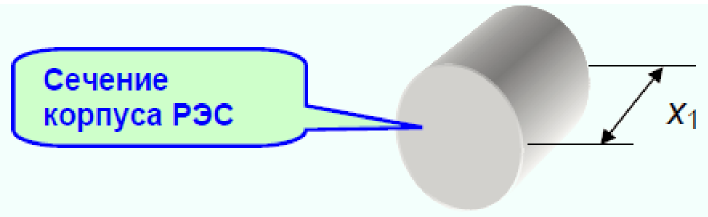
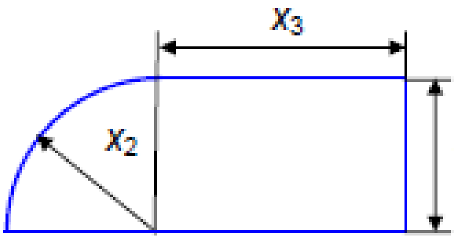


Рис. 6.2 - Характерные параметры оптимизируемого корпуса РЭС

Чтобы избежать ошибок при формировании целевой функции целесообразно воспользоваться чрезвычайно полезной способностью MathCAD преобразовать символьные выражения к простому виду [8,9,10]. Если при вычислении выражения в числовом виде MathCAD возвращает одно или несколько чисел, то результатом аналитического вычисления выражения становится другое выражение.

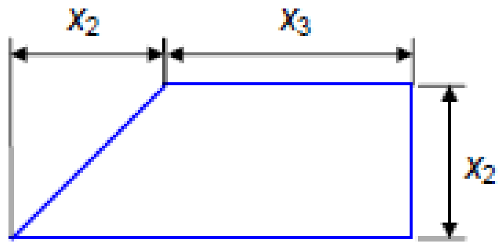
Варианты сечения корпуса РЭС приведены ниже.

1



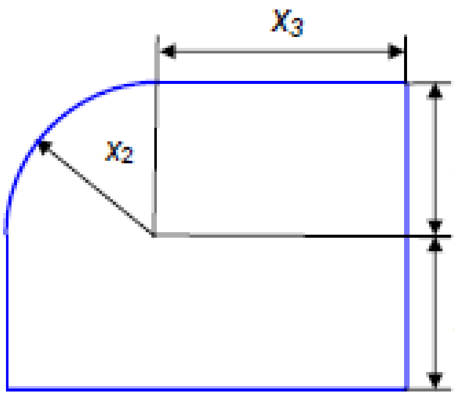
X_3
X_2
X_2

2



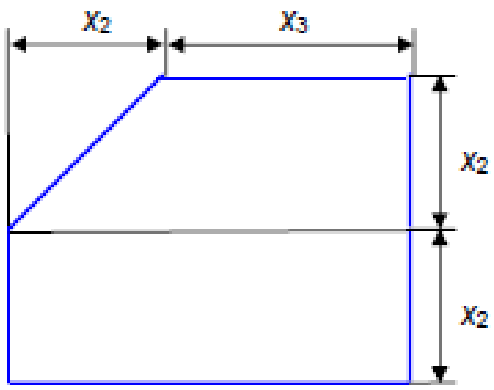
X_2
X_3
X_2

3



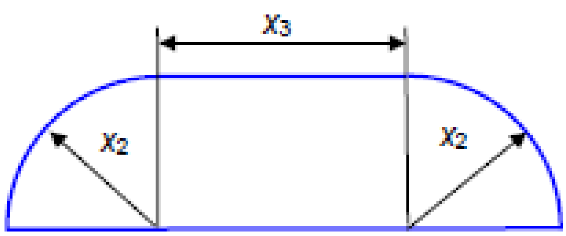
X_3
X_2
X_2
X_2

4



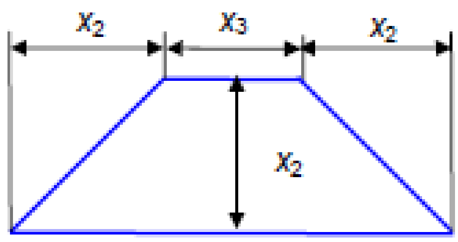
X_2
X_3
X_2
X_2

5



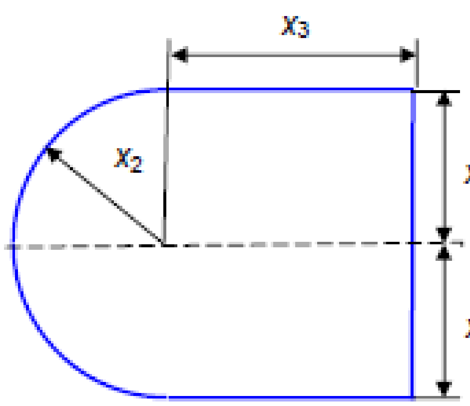
X_3
X_2
X_2

6



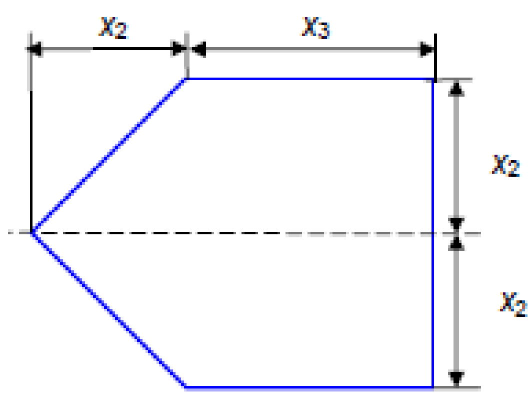
X_2
X_3
X_2
X_2

7



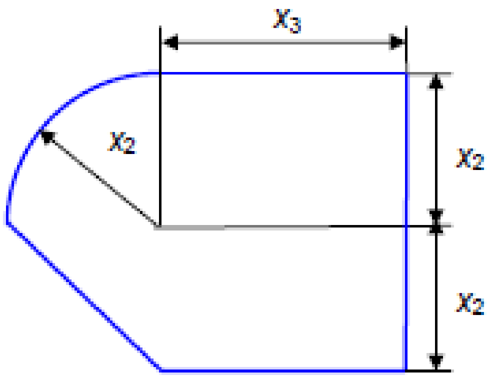
X_3
X_2
X_2
X_2

8

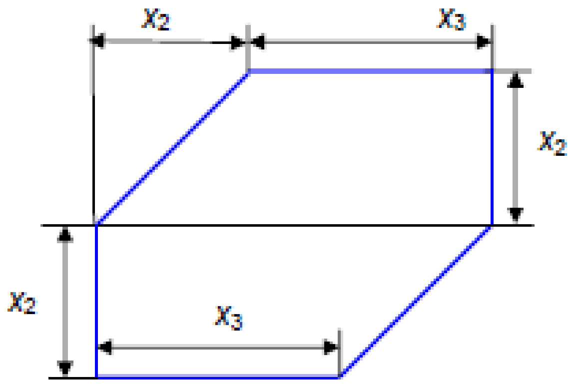


X_2
X_3
X_2
X_2

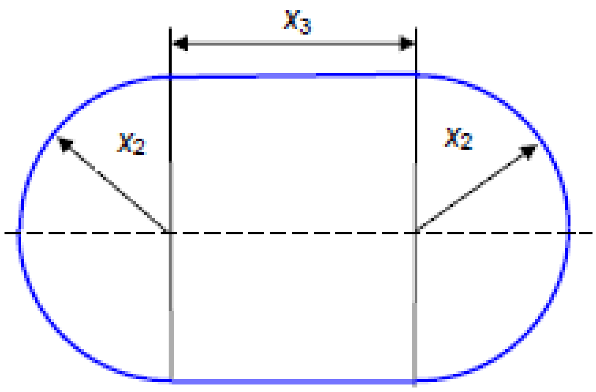
9



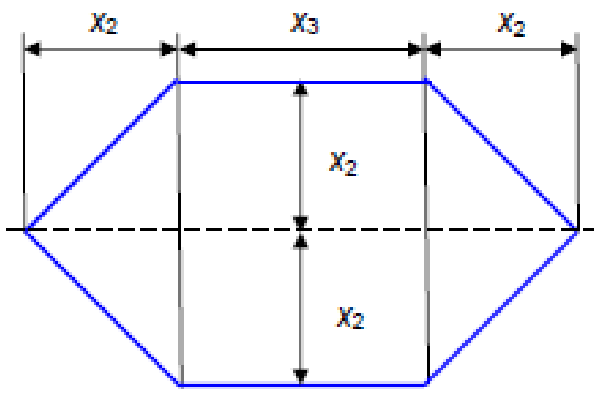
10



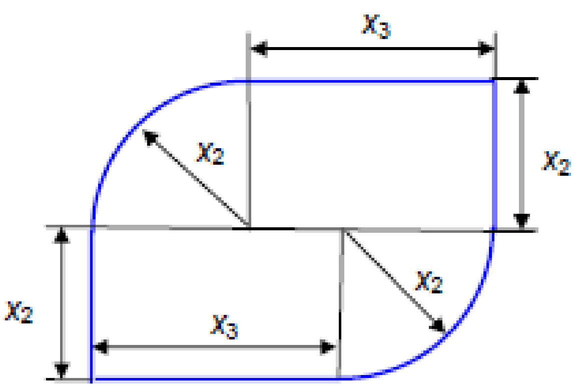
11



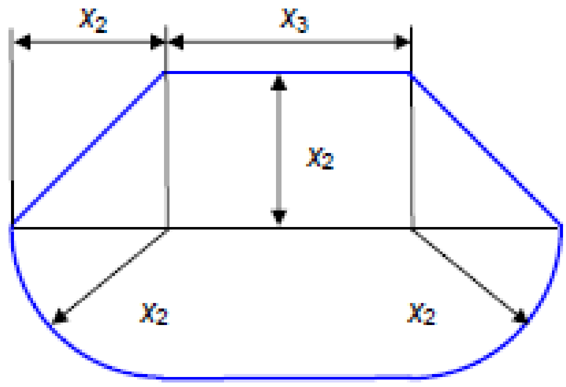
12



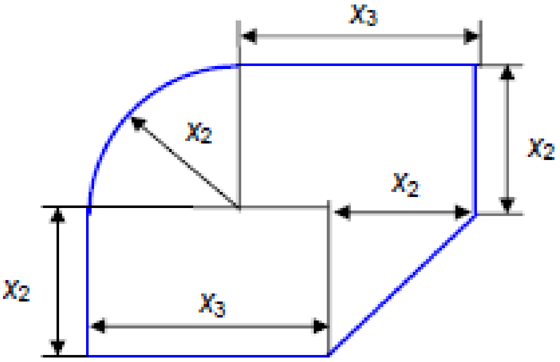
13



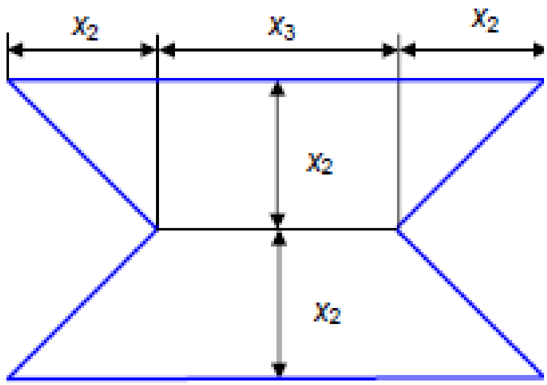
14



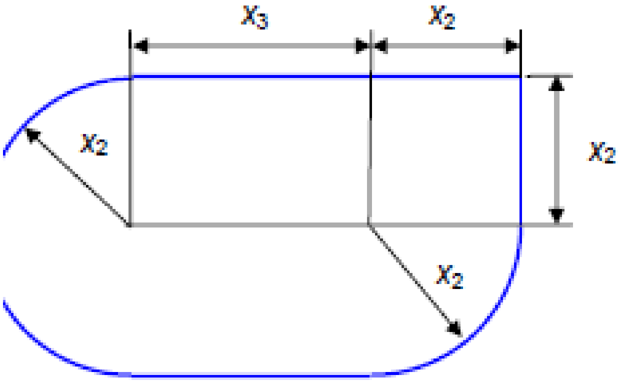
15



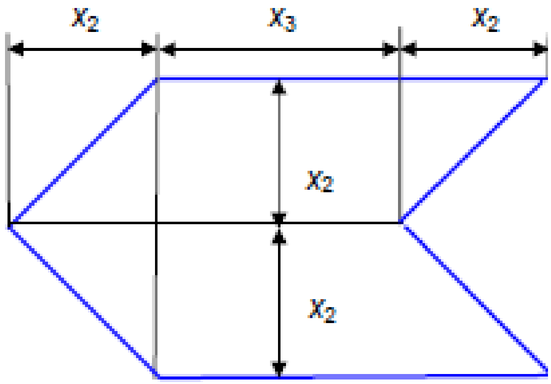
16



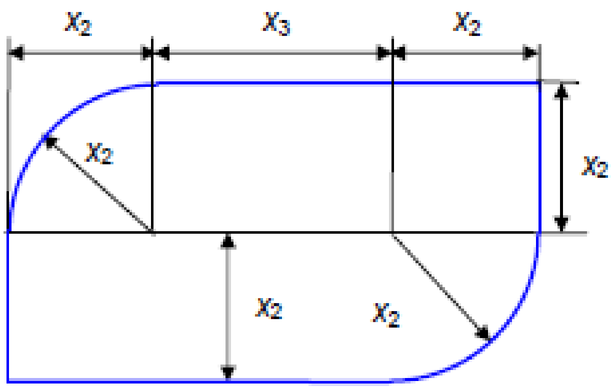
17



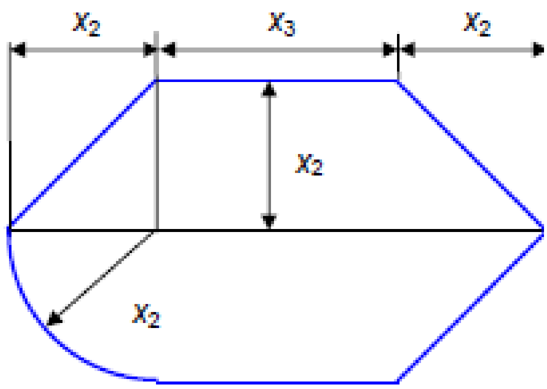
18



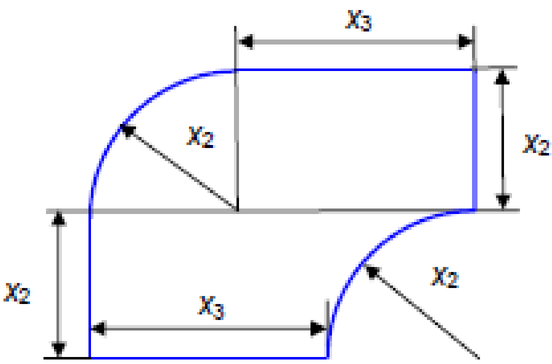
19



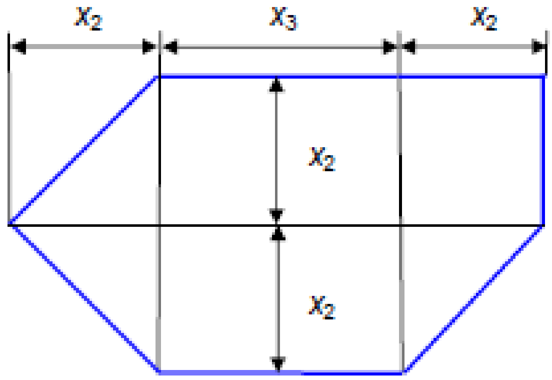
20



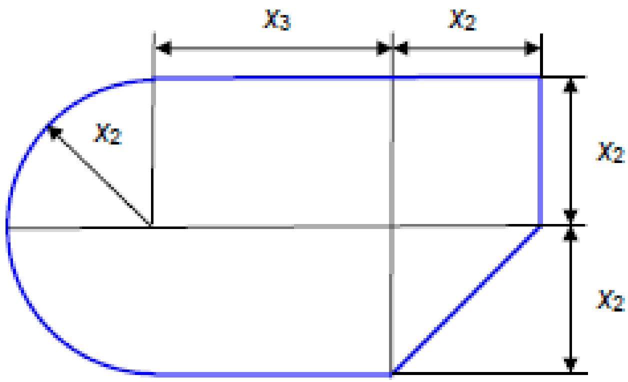
21



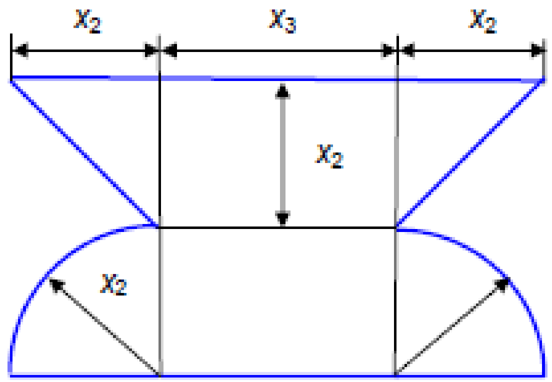
22



23



24



7 Список литературы

1. Бененсон З.М., Елистратов М.Р., Ильин Л.К. и др. Моделирование и оптимизация на ЭВМ радиоэлектронных устройств / Под ред. З.М. Бененсона. – М.: Радио и связь, 1981. – 272 с.
2. Гилл Ф., Мюррей У., Райт М. Практическая оптимизация. Пер. с англ. — М.: Мир, 1985.— 509 с.
3. Д. Мак-Кракен, У. Дорн. Численные методы и программирование на ФОРТРАНЕ. – М.: МИР, 1977. – 583 с.

4. Пантелеев А.В., Летова Т.А. Методы оптимизации в примерах и задачах. - М.: Высшая школа, 2005. - 544 с.
5. Пашкеев С.Д., Минязов Р.И., Могилевский В.Л. Машинные методы оптимизации в технике связи. Под ред. С. Д. Пашкеева. Учеб. пособие для вузов. - М.: Связь, 1976. - 272 с.
6. Химмельблау Д.. Прикладное нелинейное программирование. Пер. с англ. - М.: Мир, 1975. - 536 с.
7. Шуп Т. Решение инженерных задач на ЭВМ: Практическое руководство. - М.: Мир, 1962. - 238 с.
8. Очков В.Ф. Mathcad 14 для студентов. инженеров и конструкторов. - СПб.: БХВ-Петербург, 2007. - 368 с.
9. Рыжиков Ю.И. Вычислительные методы. - СПб.: БХВ-Петербург, 2007. - 400 с.
10. Кобрин Ю.П. Применение системы автоматизации научно-технических расчетов MathCAD при проектировании РЭС : Методические указания к лабораторной работе. [Электронный ресурс] / Ю. П. Кобрин. - Томск: ТУСУР, 2012. — 53 с.