

Лекция №5

Параметрическая оптимизация

5.1. Введение

Проектирование – это сложный и трудноформализуемый процесс, объединяющий такие важные процедуры, как синтез структуры, выбор параметров элементов, анализ и принятие решения.

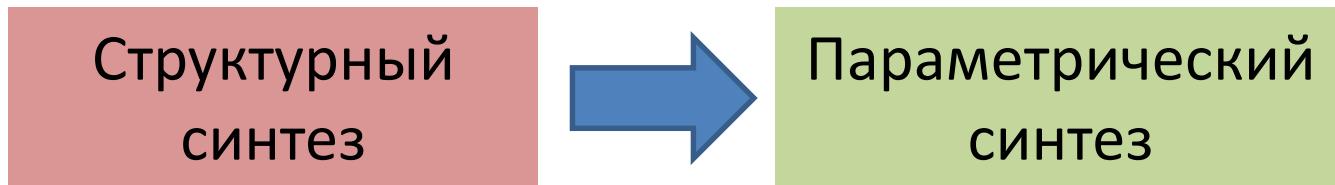
Поиск рационального технического решения при выбранном физическом принципе действия осуществляется методами **структурного синтеза**. Определение оптимальных значений параметров элементов технической системы известной структуры – задача **параметрического синтеза** или **параметрической оптимизации**.

Структурный
синтез

Параметрический
синтез

5.1. Введение

Проектирование начинается со **структурного синтеза**, при котором генерируется принципиальное решение. Но конструкции, получаемые на этом этапе, представляются в параметрическом виде, т.е. без указания числовых значений параметров элементов. Поэтому прежде чем приступить к верификации проектного решения, нужно задать или рассчитать значения этих параметров, т.е. выполнить **параметрический синтез**. Примерами результатов параметрического синтеза могут служить геометрические размеры деталей в механическом узле или в оптическом приборе и т.п.



Типовая блок-схема процесса автоматизированного проектирования

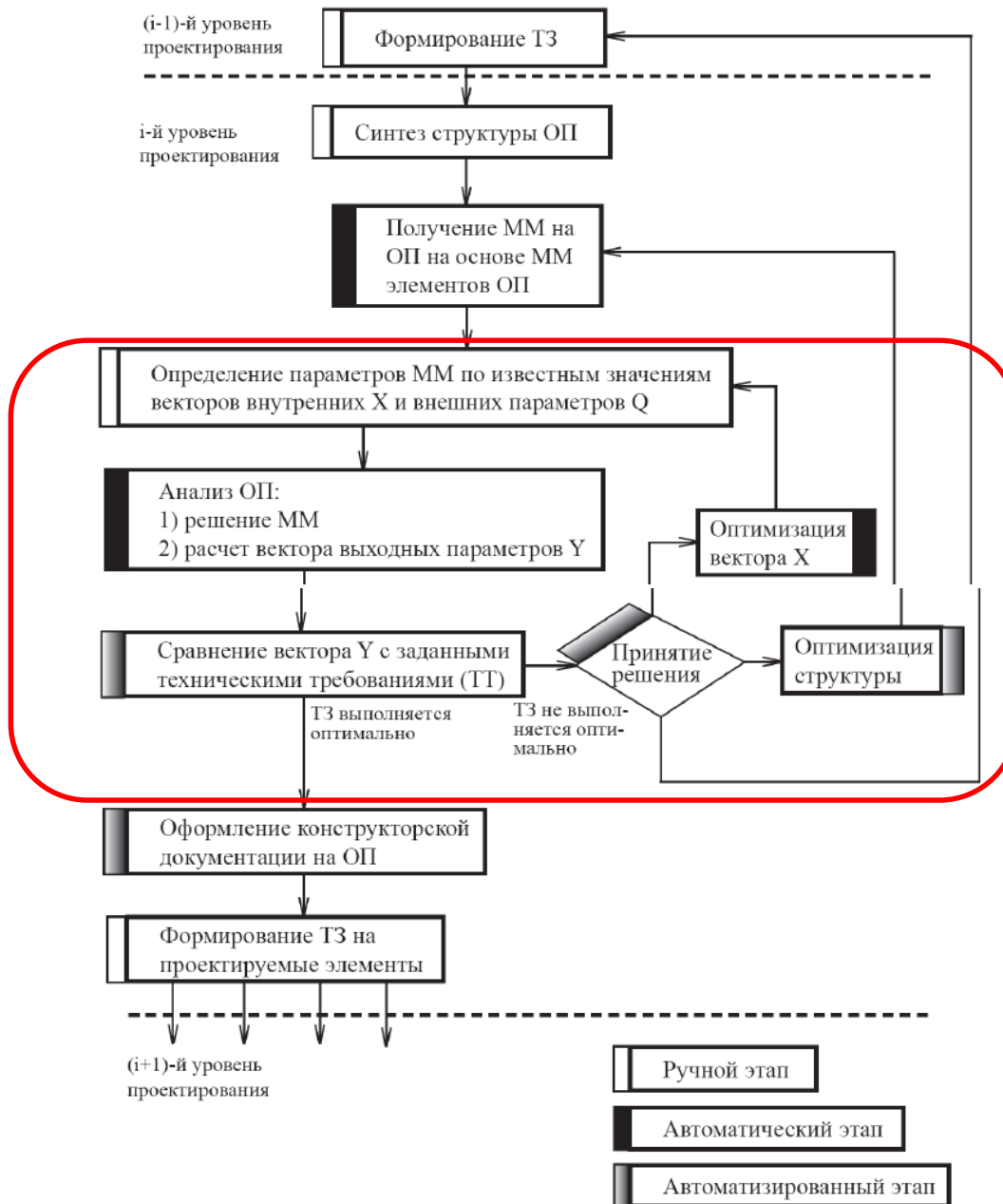
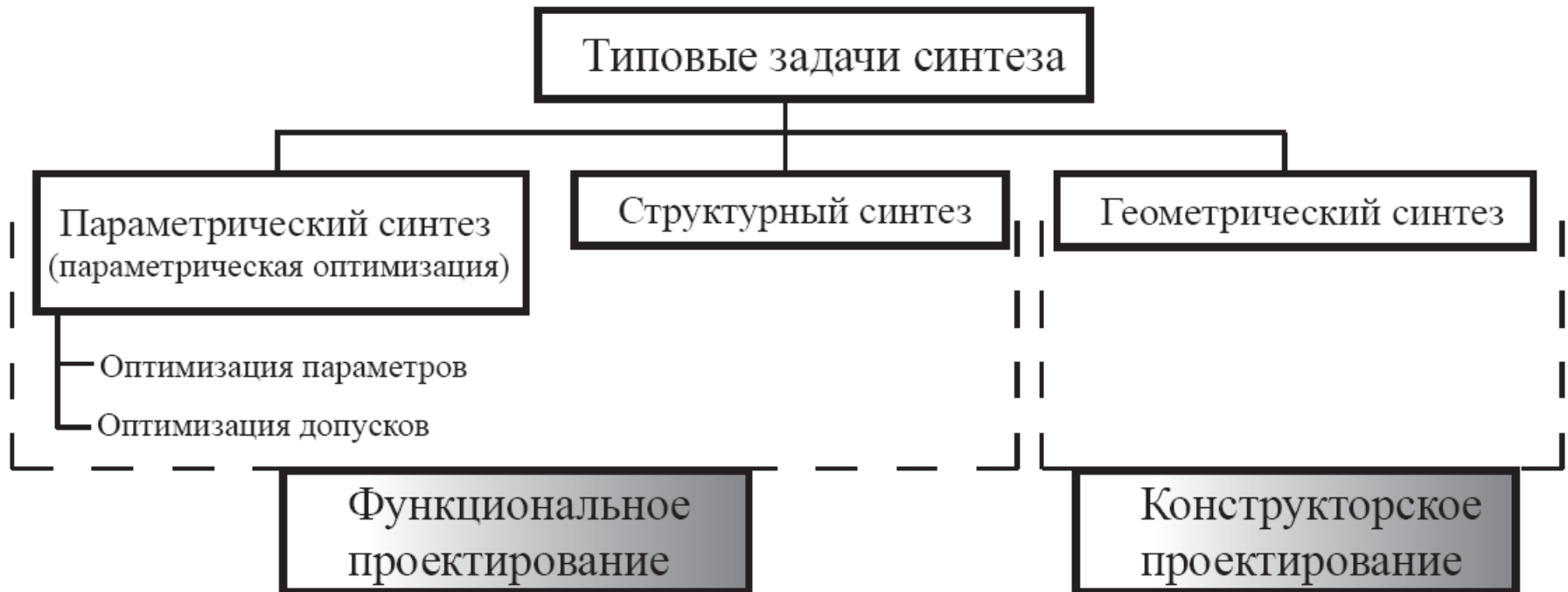


Рис. 1.2

1.6.6. Типовые задачи проектирования

По приведенной блок-схеме процесса проектирования, можно отметить, что его основу составляет последовательность решения задач синтеза, задач анализа и задач принятия решений.

На рис. 1.3 и 1.4 представлены типовые задачи синтеза и анализа, встречающиеся при проектировании изделий в автомобилестроении. Отметим, что наиболее сложными, с точки зрения формализации и автоматизации, являются задачи синтеза (структурного и геометрического).



5.1. Введение

В случае если по результатам анализа проектное решение признается неокончательным, то начинается процесс последовательных приближений к приемлемому варианту проекта. Во многих приложениях для улучшения проекта удобнее варьировать значения параметров элементов, т.е. использовать параметрический синтез на базе многовариантного анализа. При этом **задача параметрического синтеза может быть сформулирована как задача определения значений параметров элементов, наилучших с позиций удовлетворения требований технического задания при неизменной структуре проектируемого объекта.** Тогда параметрический синтез называют *параметрической оптимизацией*. Если параметрический синтез не приводит к успеху, то повторяют процедуры структурного синтеза, т.е. на очередных итерациях корректируют или перевыбирают структуру объекта.

5.1.Введение

Окончательный выбор варианта технического объекта (принятие решения) проводится с учетом выработанных правил предпочтения на основании установленных критериев. В основе правила предпочтения лежит **целевая функция**, количественно выражающая качество объекта и потому называемая также **функцией качества** или **критерием оптимальности**.

5.1. Введение

Содержание типичного ТЗ включает в себя конкретные числовые требования к основным выходным параметрам (технические требования); и качественное описание требований, ограничений и условий, непосредственно не поддающихся количественной оценке. Далее будем рассматривать только количественные параметры.

Требуемые соотношения между выходными параметрами и техническими требованиями (ТТ) называют **условиями работоспособности** и записывают в виде:

$$y_i < T_i, \quad y_j > T_j, \quad y_k = T_k \pm \Delta y_k \quad (5.1)$$

где Δy_k – допустимое отклонение k -го выходного параметра от указанного в ТЗ значения T_k

5.1. Введение

Условия работоспособности имеют определяющее значение в разработке технического устройства, так как задачей проектирования становится выбор проектного решения, в котором наилучшим образом выполняются все условия работоспособности во всем диапазоне изменения внешних параметров и при выполнении всех качественных требований ТЗ.

Во время процесса проектирования на этапе параметрического синтеза значения некоторых внутренних параметров назначаются и не подлежат изменению. К таким параметрам, например, можно отнести параметры унифицированных элементов или те из них, значения которых оговорены в ТЗ. Остальные параметры нужно выбрать оптимальным образом.

5.1. Введение

Внутренние параметры, значения которых могут меняться в процессе оптимизации и которые являются аргументами целевой функции, называют **управляемыми параметрами**.

Пусть в проектируемом объекте имеется n управляемых параметров, образующих вектор $\mathbf{X} = (x_1, x_2, \dots, x_n)$. Обозначим целевую функцию через $F(\mathbf{X})$, а область ее определения – через \mathbf{XO} . Вектор \mathbf{X} определяет координаты точки в области определения \mathbf{XO} . Если элементы вектора \mathbf{X} принимают только дискретные значения, то \mathbf{XO} является **дискретным множеством точек** и задача оптимизации относится к области **дискретного** (в частном случае целочисленного) **программирования**.

5.1. Введение

Большинство задач параметрической оптимизации технических объектов формулируется в терминах непрерывных параметров. Если экстремум целевой функции ищется в неограниченной области **ХО**, то его называют **безусловным**, а методы поиска – методами **безусловной оптимизации**.

Чаще на область определения **ХО** наложены некоторые ограничения, ограничивающие ее. В этом случае искомый экстремум становится **условным**, а методы поиска – методами **условной оптимизации**.

5.2. Постановка задач оптимизации

Проблема оптимизации имеет два основных аспекта:

- 1) нужно **поставить задачу**, формализовав понятие «оптимальный»;
- 2) нужно **решить задачу**, уже имеющую математическую формулировку.

Решение задачи строгими математическими методами может быть выполнено после того, как задача поставлена.

Процедура постановки задачи носит неформальный характер и включает следующие этапы:

- **Выбор целевой функции**
- **Выбор управляемых параметров**
- **Назначение ограничений**
- **Нормирование управляемых и выходных параметров.**

5.2. Постановка задач оптимизации

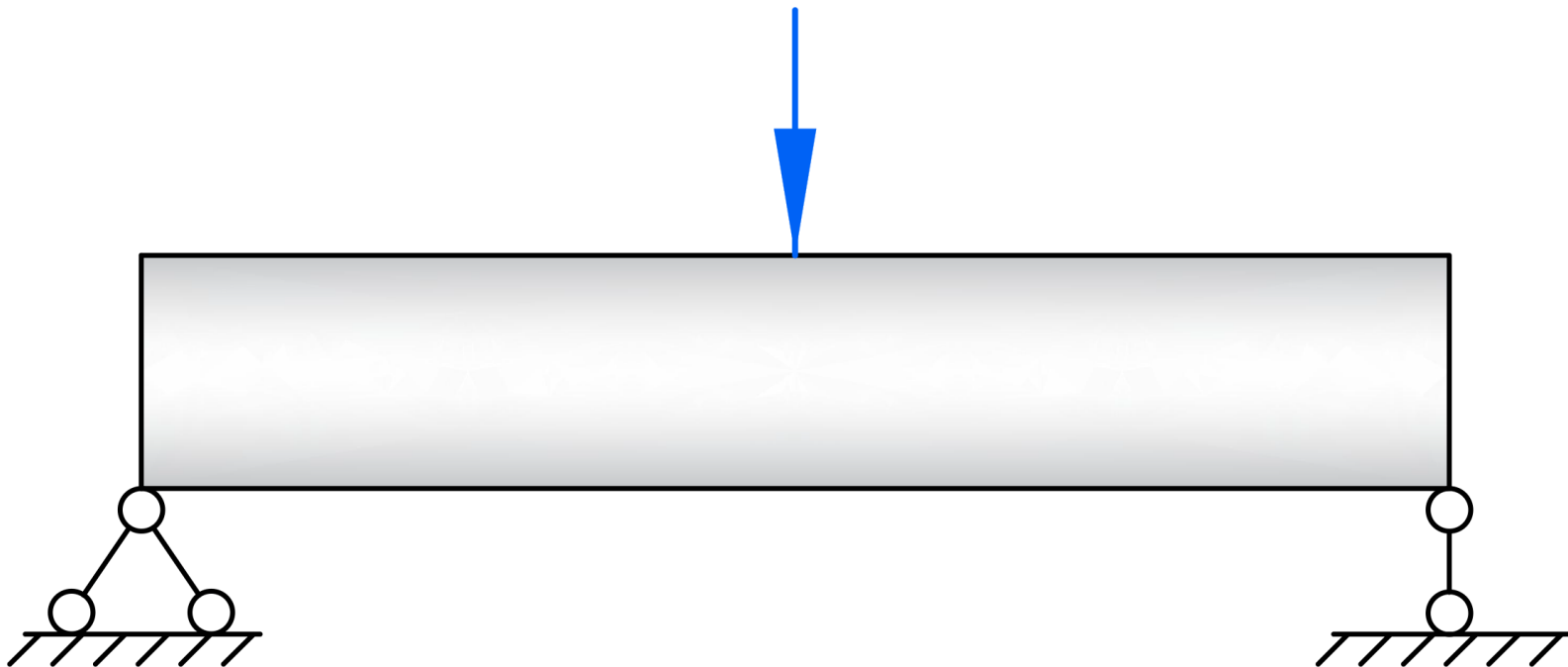
Пример: Необходимо провести оптимизацию поперечного сечения балки, работающей на изгиб.

Целевая функция: минимизировать массу

Управляемые параметры: размеры поперечного сечения

Ограничения: прогиб в центре балки

Контролируемые параметры: масса, прогиб



5.2. Постановка задач оптимизации

Пример: Оптимизация передаточных чисел коробки передач, для сокращения времени разгона до 100 км/ч.

Целевая функция: минимизировать время

Управляемые параметры: передаточные числа

Ограничения: ограничения на передаточные числа из условий компоновки

Контролируемые параметры: Время



5.2.1. Выбор целевой функции

Основная проблема постановки экстремальных задач заключается в формулировке целевой функции. Сложность выбора целевой функции (функции качества) состоит в том, что любой технический объект первоначально имеет **векторный характер критериев оптимальности, причем улучшение одного из выходных параметров, как правило, приводит к ухудшению другого**, так как все выходные параметры являются функциями одних и тех же управляемых параметров и не могут изменяться независимо друг от друга. Такие выходные параметры называют ***конфликтными параметрами***.

Целевая функция должна быть одна. Сведение многокритериальной задачи к однокритериальной называют ***сверткой векторного критерия***.

В зависимости от того, каким образом выбираются и объединяются выходные параметры в скалярной функции качества, различают ***частные, аддитивные, мультипликативные, минимаксные, статистические критерии*** и т. д.

5.2.1. Выбор целевой функции

Частные критерии могут применяться в случаях, когда среди выходных параметров можно выделить один основной параметр $y_i(X)$, наиболее полно отражающий эффективность проектируемого объекта.

Этот параметр принимают за целевую функцию. Примерами таких параметров являются: для энергетического объекта – мощность, для технологического автомата – производительность, для транспортного средства – грузоподъемность. Для многих технических объектов таким параметром служит стоимость. Условия работоспособности всех остальных выходных параметров объекта относят при этом к функциональным ограничениям. Оптимизация на основе такой постановки называется оптимизацией по частному критерию.

Достоинство такого подхода – простота, существенный недостаток – то, что большой запас работоспособности можно получить только по основному параметру, который принят в качестве целевой функции; а другие выходные параметры вообще не будут иметь запасов.

5.2.1. Выбор целевой функции

На рис. 5.1 представлено двумерное пространство выходных параметров y_1 и y_2 , для которых заданы условия работоспособности $y_1 < T_1$ и $y_2 < T_2$. Допустим, что повышению качества объекта соответствует уменьшение этих параметров. **Кривая АВ** является границей достижимых значений выходных параметров. Это ограничение объективное и связано с существующими физическими и технологическими условиями производства, называемыми **условиями реализуемости**.



Рис. 5.1. Области работоспособности и Паретто

5.2.1. Выбор целевой функции

Область, в пределах которой выполняются все условия реализуемости и работоспособности, называют **областью работоспособности**. Множество точек пространства выходных параметров, из которых невозможно перемещение, приводящее к улучшению всех выходных параметров, называют **областью компромиссов**, или **областью Паретто**. Участок кривой AB (см. рис. 5.1) относится к области Паретто.



рис. 5.1. Области работоспособности и Паретто

5.2.1. Выбор целевой функции

Если в качестве целевой функции в ситуации рис. 5.1. выбрать параметр y_1 , то результатом оптимизации будут параметры X , соответствующие точке **B**. Но это граница области работоспособности и, следовательно, при нестабильности внутренних и внешних параметров велика вероятность выхода за пределы области работоспособности.



рис. 5.1. Области работоспособности и Паретто

5.2.1. Выбор целевой функции

Пример: Изгиб балки

Выходные параметры: масса и прогиб

Целевыми функциями можно выбрать и минимум массы и минимум прогиба.

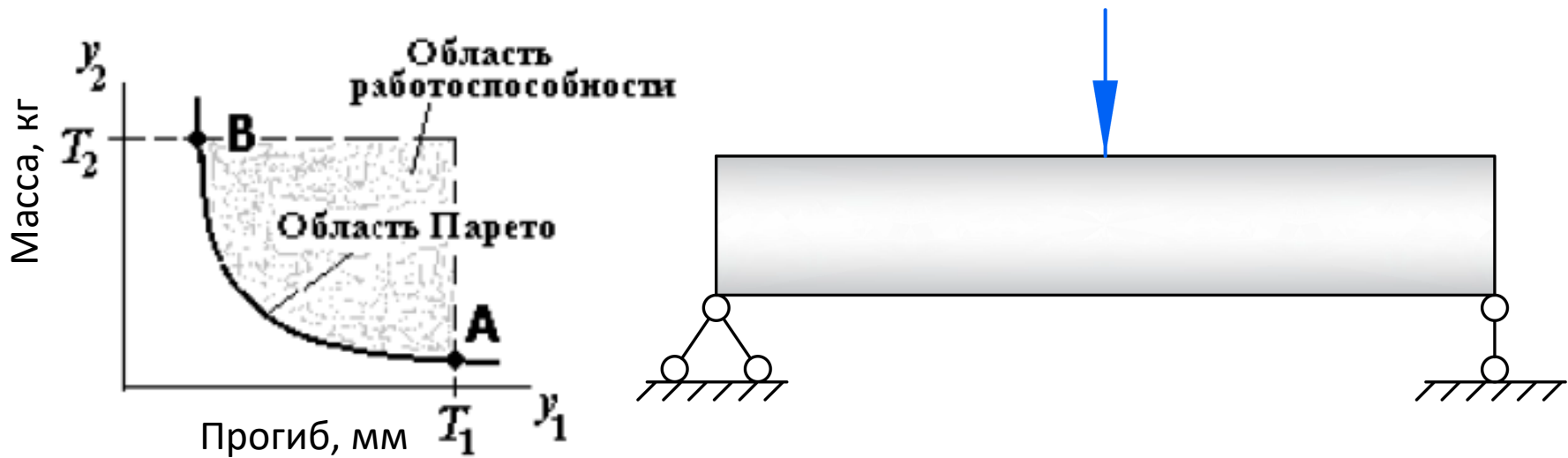


рис. 5.1. Области работоспособности и Паретто

5.2.1. Выбор целевой функции

Взвешенный аддитивный критерий объединяет (свертывает) все выходные параметры $y_j(\mathbf{X})$ (частные критерии) в одну целевую функцию, представляющую собой взвешенную сумму частных критериев:

$$F(\mathbf{X}) = \sum_{j=1}^m \omega_j y_j(\mathbf{X}), \quad (5.2)$$

где ω_j – весовой коэффициент, m – число выходных параметров. Для случая минимизации функции (5.1) $\omega_j < 0$, если выходной параметр $y_j(\mathbf{X})$ в процессе оптимизации нужно увеличивать; $\omega_j > 0$, если выходной параметр $y_j(\mathbf{X})$ в процессе оптимизации нужно уменьшать.

Объединение нескольких выходных параметров, имеющих в общем случае различную физическую размерность, в одной скалярной целевой функции требует предварительного нормирования этих параметров. Способы нормирования параметров будут рассмотрены ниже.

5.2.1. Выбор целевой функции

Мультипликативный критерий может применяться в случаях, когда выходные параметры не могут принимать нулевые значения. В этом случае целевая функция имеет вид:

$$F(\mathbf{X}) = \prod_{j=1}^m (y_j(\mathbf{X}))^{\omega_j} \quad (5.3)$$

Для случая минимизации функции (5.2) $\omega_j < 0$, если выходной параметр $y_j(\mathbf{X})$ в процессе оптимизации нужно увеличивать; $\omega_j > 0$, если выходной параметр $y_j(\mathbf{X})$ в процессе оптимизации нужно уменьшать.

Наиболее существенными недостатками как аддитивного, так и мультипликативного критерия является субъективный подход к выбору весовых коэффициентов и неучет в постановке задачи технических требований, предъявляемых к выходным параметрам.

5.2.1. Выбор целевой функции

Максиминные (минимаксные) критерии позволяют достичь одной из целей оптимального проектирования – наилучшего удовлетворения условий работоспособности.

Введем количественную оценку степени выполнения j -го условия работоспособности, обозначим ее через z_j , и будем называть *запасом работоспособности параметра y_j* . Расчет запаса по j -му выходному параметру можно выполнить различными способами, например,

$$z_j = a_j \left(\frac{T_j - y_{j\text{ном}}}{\delta_j} - 1 \right), \quad (5.4)$$

где a_j – весовой коэффициент; $y_{j\text{ном}}$ – номинальное значение j -го выходного параметра; δ_j – величина, характеризующая разброс j -го выходного параметра.

5.2.1. Выбор целевой функции

Здесь предполагается, что все соотношения сведены к виду $uj < Tj$. Следует принимать $aj > 1$ (рекомендуемые значения $5 < aj < 20$), если желательно достичь выполнения j -го технического требования с заданным допуском; $aj = 1$, если необходимо получить максимально возможную оценку zj .

Естественно теперь поставить задачу о выборе такой стратегии поиска вектора внутренних параметров \mathbf{X} , которая максимизировала бы минимальный из запасов, т.е.:

$$F(\mathbf{X}) = \max_{\mathbf{X} \in \mathbf{XD}} \min_{1 \leq j \leq m} z_j(\mathbf{X}), \quad (5.4)$$

где m – количество запасов работоспособности; \mathbf{XD} – допустимая для поиска область.

Критерий оптимизации с целевой функцией (5.4) называют максиминным критерием.

5.2.2. Назначение ограничений

Назначение ограничений. Ограничения объективно появляются при проектировании технических объектов и вытекают из конкретной физической и технологической реализуемости внутренних параметров элементов, ограниченности ресурсов и т. п. Ограничения суживают область **XO** и искомый экстремум становится условным.

Различают *прямые* и *функциональные* ограничения.

5.2.2. Назначение ограничений

Прямые ограничения имеют вид

$$x_{ni} \leq x_i \leq x_{ei} \text{ при } i \in [1:n],$$

где x_{ni} , x_{ei} – минимально и максимально допустимые значения i -го управляемого параметра; n – размерность пространства управляемых параметров. Например, для многих объектов параметры элементов не могут быть отрицательными: $x_{ni} > 0$ (геометрические размеры, электрические сопротивления, массы и т. п.).

5.2.2. Назначение ограничений

Функциональные ограничения, как правило, представляют собой условия работоспособности выходных параметров, не вошедших в целевую функцию. Функциональные ограничения могут быть:

1) типа равенств

$$\psi(\mathbf{X}) = 0 ;$$

2) типа неравенств

$$\varphi(\mathbf{X}) > 0 ,$$

где $\psi(\mathbf{X})$ и $\varphi(\mathbf{X})$ – вектор-функции.

5.2.3. Нормирование управляемых и выходных параметров

Пространство управляемых параметров – метрическое. Поэтому при выборе направлений и величин шагов поиска необходимо вводить ту или иную норму, отождествляемую с расстоянием между двумя точками. Последнее предполагает, что все управляемые параметры имеют одинаковую размерность или являются безразмерными.

Возможны различные способы нормирования. В качестве примера рассмотрим способ логарифмического нормирования, достоинством которого является переход от абсолютных приращений параметров к относительным. В этом случае i -й управляемый параметр u_i преобразуется в безразмерный x_i , следующим образом:

$$x_i = \ln \left(\frac{u_i}{\zeta_i} \right),$$

где ζ_i — коэффициент, численно равный единице измерения параметра u_i .

5.2.4. Математическая формулировка задач оптимизации

Задачу параметрического синтеза технических объектов в некоторых случаях можно сформулировать как задачу безусловной оптимизации:

$$\min_{\mathbf{X} \in \mathbf{X}_0} F(\mathbf{X}) .$$

Наиболее типичным случаем параметрической оптимизации технических объектов является поиск значений вектора управляемых параметров, доставляющих экстремум целевой функции при наличии ограничений. Формулировка задачи в этом случае запишется так:

$$\min_{\mathbf{X} \in \mathbf{X}_D} F(\mathbf{X})$$

$$\psi(\mathbf{X}) = 0; \quad \varphi(\mathbf{X}) > 0;$$

$$x_{ni} \leq x_i \leq x_{ei}, \quad \text{при } i \in [1:n].$$

5.3. Обзор методов оптимизации

Решение задач оптимизации в САПР ведется с помощью *поисковых методов*, использующих предшествующую информацию для построения улучшенного решения задачи, т.е. эти методы являются итерационными. Поисковые методы основаны на пошаговом изменении управляемых параметров:

$$\mathbf{X}_{k+1} = \mathbf{X}_k + \Delta \mathbf{X}_k ,$$

где в большинстве методов приращение $\Delta \mathbf{X}_k$ вектора управляемых параметров вычисляется по формуле:

$$\Delta \mathbf{X}_k = h \cdot \mathbf{g}(\mathbf{X}_k)$$

Здесь \mathbf{X}_k – значение вектора управляемых параметров на k -м шаге, h – шаг, а $\mathbf{g}(\mathbf{X}_k)$ – направление поиска. Следовательно, если выполняются условия сходимости, то реализуется пошаговое (итерационное) приближение к экстремуму.

5.3. Обзор методов оптимизации

Методы оптимизации **классифицируют** по ряду признаков.

В зависимости от числа управляемых параметров различают методы **одномерной** и **многомерной** оптимизации, в первых из них управляемый параметр единственный, во вторых размер вектора **X** не менее двух. Реальные задачи в САПР многомерны, методы одномерной оптимизации играют вспомогательную роль на отдельных этапах многомерного поиска.

5.3. Обзор методов оптимизации

Различают методы *условной* и *безусловной*; оптимизации по наличию или отсутствию ограничений. Для реальных задач характерно наличие ограничений, однако методы безусловной оптимизации также представляют интерес, поскольку задачи условной оптимизации с помощью специальных методов могут быть сведены к задачам без ограничений.

5.3. Обзор методов оптимизации

В зависимости от числа экстремумов различают задачи одно- и многоэкстремальные. Если метод ориентирован на определение какого-либо локального экстремума, то такой метод относится к **локальным**. Если же результатом является глобальный экстремум, то метод называют **глобальным**. Удовлетворительные по вычислительной эффективности методы глобального поиска для общего случая отсутствуют и потому на практике в САПР используют методы поиска локальных экстремумов.

5.3. Обзор методов оптимизации

Наконец, в зависимости от того, используются при поиске производные целевой функции по управляемым параметрам или нет, различают методы нескольких порядков. Если производные не используются, то имеет место метод **нулевого порядка**, если используются первые или вторые производные, то соответственно метод **первого** или **второго порядка**. Методы первого порядка называют также градиентными, поскольку вектор первых производных $F(\mathbf{X})$ по \mathbf{X} есть градиент целевой функции:

$$\mathbf{grad} F(\mathbf{X}) = \left(\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \dots, \frac{\partial F}{\partial x_n} \right)$$

Конкретные методы определяются следующими факторами:

- 1) способом вычисления направления поиска $\mathbf{g}(\mathbf{X}^k)$ в формуле (5.13);
- 2) способом выбора шага h ;
- 3) способом определения окончания поиска.

5.3. Обзор методов оптимизации

Определяющим фактором является первый из перечисленных в этом списке, он подробно описан ниже.

Шаг может быть или постоянным, или выбираться исходя из одномерной оптимизации – поиска минимума целевой функции в выбранном направлении $\mathbf{g}(\mathbf{X}_k)$. В последнем случае шаг называется оптимальным.

Окончание поиска обычно осуществляют по правилу: если на протяжении r подряд идущих шагов траектория поиска остается в малой ε -окрестности текущей точки поиска \mathbf{X}_k , то поиск следует прекратить, следовательно, условие окончания поиска имеет вид:

$$\left| \mathbf{X}_k - \mathbf{X}_{k-r} \right| < \varepsilon$$

5.3. Обзор методов оптимизации

Классификацию методов оптимизации можно представить в виде следующей схемы:



5.4 Методы одномерной оптимизации

К методам одномерной оптимизации относятся методы **дихотомии**, **золотого сечения**, **чисел Фибоначчи** (методы последовательного поиска), **полиномиальной аппроксимации** и ряд их модификаций.

В методах последовательного поиска, исходя из свойств унимодальности целевой функции, строится такая стратегия последовательного поиска экстремума x^* , при которой любая пара вычислений $F(x)$ позволяет сузить область поиска. Вычисляя $F(x)$ в точках x_1 и x_2 таких, что $a < x_1 < x_2 < b$, можно локализовать интервал неопределенности путем анализа полученных значений функции (рис. 5.2):

если $F(x_1) < F(x_2)$, $x^* \in [a, x_2]$;

если $F(x_1) = F(x_2)$, $x^* \in [x_1, x_2]$;

если $F(x_1) > F(x_2)$, $x^* \in [x_1, b]$.

5.4 Методы одномерной оптимизации

Стратегия выбора значений x_1 и x_2 для проведения опытов с учетом предыдущих результатов определяет сущность различных методов последовательного поиска.

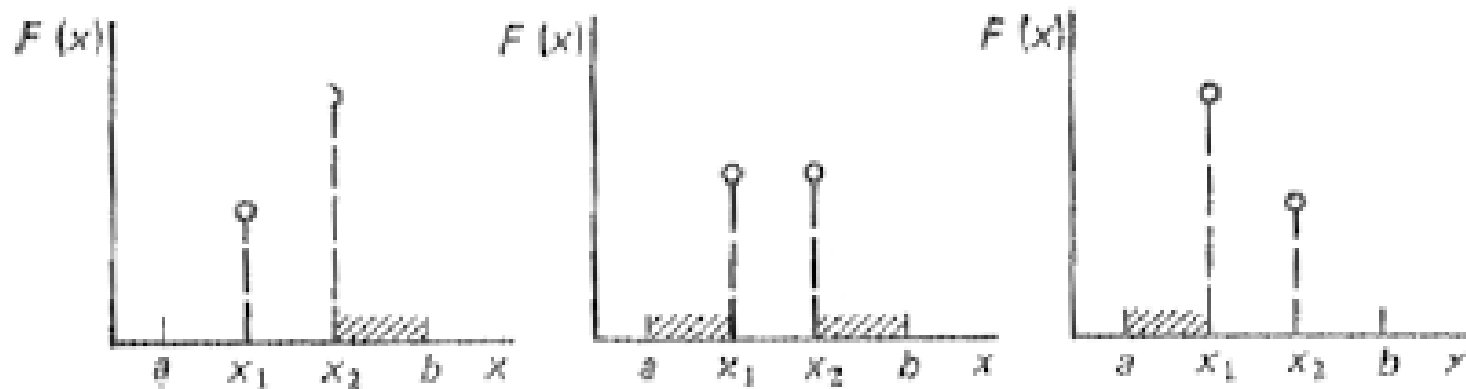


Рис. 5.2. Определение области поиска

5.4 Методы одномерной оптимизации

В **методе дихотомии (половинного деления)** на k -ом шаге проводится два опыта, при которых аргументы $x_{1,k}$ и $x_{2,k}$ выбираются на расстоянии $\delta/2$ справа и слева от середины интервала:

$$x_{1,k} = \frac{a_k + b_k}{2} - \frac{\delta}{2}; \quad x_{2,k} = \frac{a_k + b_k}{2} + \frac{\delta}{2}$$

Вычислив $F(x_{1,k})$ и $F(x_{2,k})$ и сравнив полученные значения можно найти новый интервал неопределенности (как показано выше) и продолжить поиск до тех пор, пока длина интервала не уменьшится до величины погрешности δ .

Для этого потребуется N шагов, где N – ближайшее к $\log_2 \left(\frac{2(b-a)}{\delta} \right)$ целое значение, но на каждом шаге целевую функцию приходится вычислять дважды.

5.4 Методы одномерной оптимизации

В *методе золотого сечения* точки $x_{1,k}$ и $x_{2,k}$ выбираются симметрично относительно центра интервала неопределенности с учетом коэффициента пропорциональности τ следующим образом:

$$x_{1,k} = a_k + \tau(b_k - a_k); \quad x_{2,k} = b_k - \tau(b_k - a_k)$$

Затем в этих точках вычисляются значения целевой функции, анализ которых позволяющие сократить интервал неопределенности, и поиск продолжается. Если подобрать значение τ так, что в полученном отрезке меньшей длины одна из промежуточных точек совпадает с промежуточной точкой от предыдущего шага, то это позволит сократить число вычислений целевой функции на всех шагах (кроме первого) в два раза. Условие получения такого значения τ формулируется следующим образом:

В результате имеем значение

$$\tau = \frac{3 - \sqrt{5}}{2} \approx 0,382$$

, которое называется *золотым сечением*.

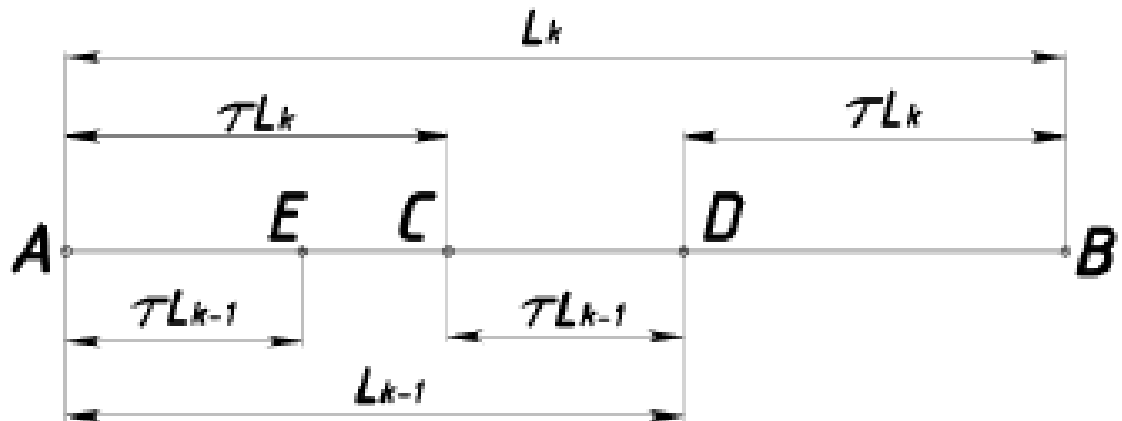
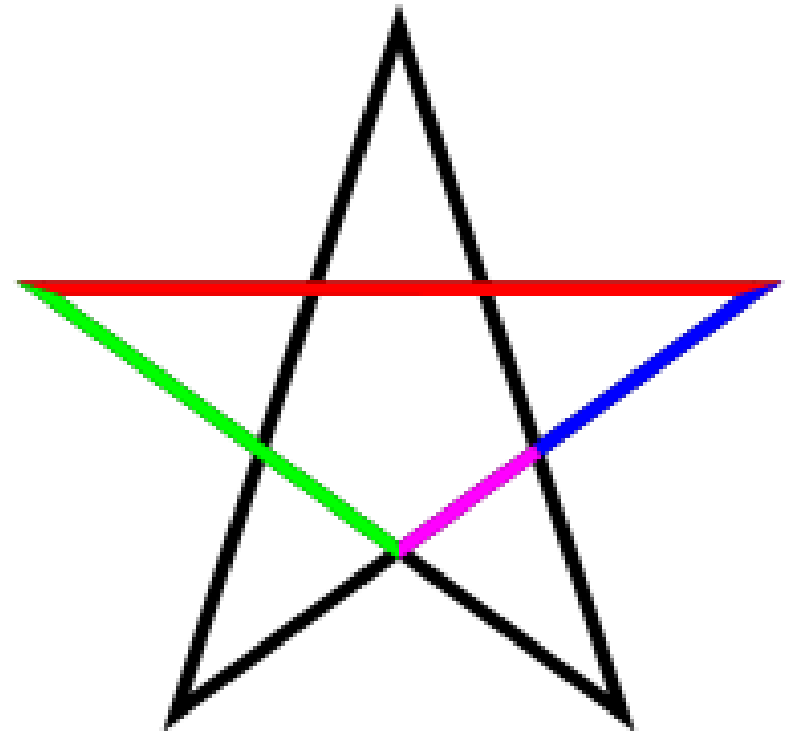
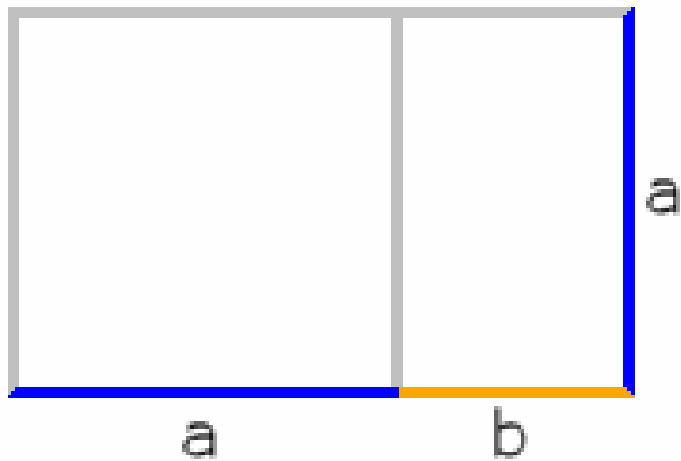


Рис. 5.3 Золотое сечение

5.4 Методы одномерной оптимизации

Золотое сечение (золотая пропорция, деление в крайнем и среднем отношении, **гармоническое деление**) — соотношение двух величин a и b , $b > a$, когда справедливо $b/a = (a+b)/b$. Число, равное отношению b/a , обычно обозначается греческой буквой φ , реже — греческой буквой τ . Из исходного равенства нетрудно получить, что число

$$\varphi = \frac{1 + \sqrt{5}}{2}$$



5.4 Методы одномерной оптимизации

Краткое описание метода чисел Фибоначчи

Метод Фибоначчи является процедурой линейного поиска минимума унимодальной функции $f(x)$ на замкнутом интервале $[a, b]$, отличающейся от процедуры золотого сечения тем, что **очередная пробная точка делит интервал локализации в отношении двух последовательных чисел Фибоначчи**. Последовательность чисел Фибоначчи задается условиями $F_0 = F_1 = 1$, $F_{k+1} = F_k + F_{k-1}$, $k = 1, 2, \dots$. Начальными членами последовательности будут $1, 1, 2, 3, 5, 8, 13, \dots$.

Стратегия поиска Фибоначчи требует заранее указать n – число вычислений минимизируемой функции и ε – константу различимости двух значений $f(x)$.

5.4 Методы одномерной оптимизации

Числа Фибона́ччи — элементы числовой последовательности 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, 4181, 6765, 10946, ... в которой каждое последующее число равно сумме двух предыдущих чисел. Название по имени средневекового математика Леонардо Пизанского (известного как Фибоначчи). Иногда число 0 не рассматривается как член последовательности.

Иногда числа Фибоначчи рассматривают и для отрицательных номеров n как двусторонне бесконечную последовательность, удовлетворяющую тому же рекуррентному соотношению. При этом члены с отрицательными индексами легко получить с помощью эквивалентной формулы «назад»:

$$F_n = F_{n+2} - F_{n+1}$$

n	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10
F_n	-55	34	-21	13	-8	5	-3	2	-1	1	0	1	1	2	3	5	8	13	21	34	55

Легко заметить, что

$$F_{-n} = (-1)^{n+1} F_n$$

5.4 Методы одномерной оптимизации

По *методу полиномиальной аппроксимации* целевая функция $F(x)$ заменяется квадратичным полиномом

$$P(x) = a_0 + a_1x + a_2x^2$$

вычисляют значения целевой функции на граничных точках интервала и в одной промежуточной точке. Далее решают систему из трех алгебраических уравнений, полученных подстановкой в (5.20) значений координат указанных точек вместо x и вычисленных значений функции вместо $P(x)$. В результате становятся известными значения коэффициентов k и a в (5.20) и, исходя из условия $dP(x)/dx = 0$, определяют экстремальную точку x^* .

5.5 Методы многомерной оптимизации

Среди методов нулевого порядка в САПР находят применение методы Розенброка, конфигураций (Хука-Дживса), деформируемого многогранника (Нелдера-Мида), случайного поиска. К методам с использованием производных относятся методы наискорейшего спуска, сопряженных градиентов, переменной метрики.

5.5 Методы многомерной оптимизации

Метод покоординатного спуска характеризуется выбором направлений поиска поочередно вдоль всех n координатных осей, шаг рассчитывается на основе одномерной оптимизации, критерий окончания поиска $|\mathbf{X}_k - \mathbf{X}_{k+n}| < \varepsilon$, где ε – заданная точность определения локального экстремума, n – размерность пространства управляемых параметров. Траектория покоординатного спуска для примера двумерного пространства управляемых параметров показана на рис. 5.4, где \mathbf{X}^k – точки на траектории поиска, x_i – управляемые параметры. Целевая функция представлена своими линиями равного уровня, около каждой линии записано соответствующее ей значение $F(\mathbf{X})$. Очевидно, что \mathcal{E} есть точка минимума.

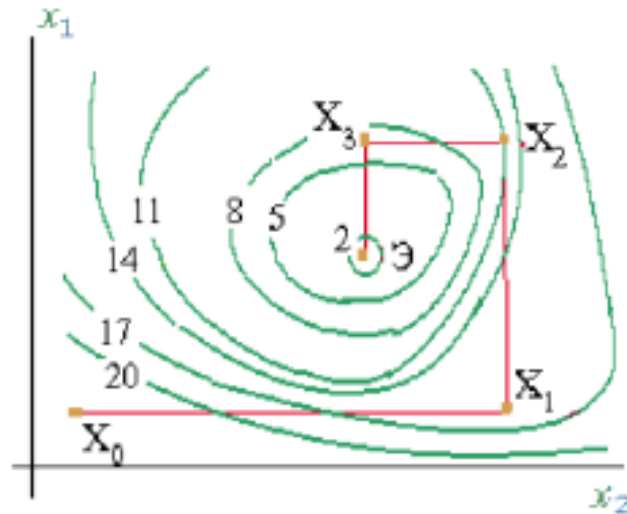


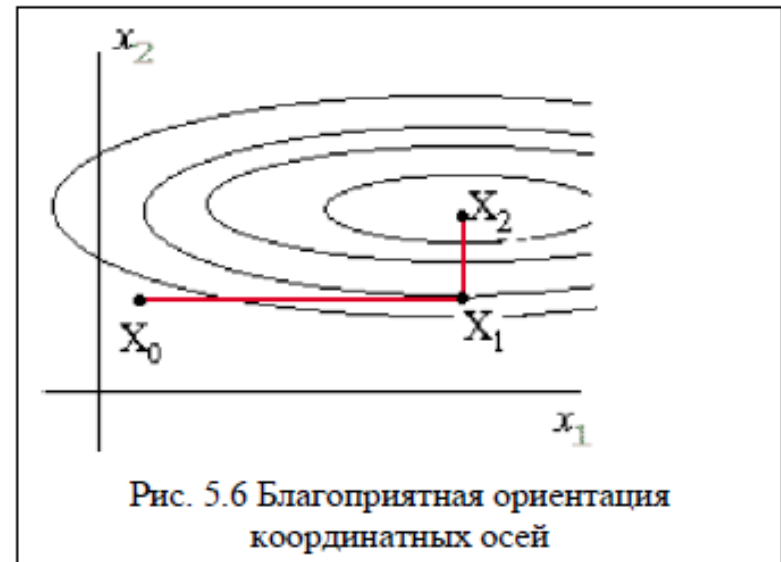
Рис. 5.4 Траектория покоординатного спуска

5.5 Методы многомерной оптимизации

При использовании метода покоординатного спуска велика вероятность “застревания” поиска на дне оврага вдали от точки экстремума. На рис. 5.5 видно, что после попадания в точку C , расположенную на дне оврага, дальнейшие шаги возможны лишь в направлениях aa или bb , но они приводят к ухудшению целевой функции. Следовательно, поиск прекращается в точке C .

Оврагом называют часть пространства управляемых параметров, в которой наблюдаются слабые изменения производных целевой функции по одним направлениям и значительные изменения с переменной знака – по некоторым другим направлениям. Знак производной меняется в точках, принадлежащих дну оврага.

В то же время при благоприятной ориентации дна оврага, а именно при положении одной из координатных осей, близком к параллельности с дном оврага, поиск оказывается весьма быстрым. Эта ситуация показана на рис. 5.6.



5.5 Методы многомерной оптимизации

Метод Розенброка заключается в таком повороте координатных осей, чтобы одна из них оказалась квазипараллельной дну оврага. Такой поворот осуществляют на основе данных, полученных после серии из n шагов покоординатного спуска. Положение новых осей s_i может быть получено линейным преобразованием прежних осей x_i : ось s_1 совпадает по направлению с вектором $X_{k+n} - X_k$ остальные оси выбирают из условия ортогональности к X_1 и друг к другу.

5.5 Методы многомерной оптимизации

Другой удачной модификацией покоординатного спуска является **метод конфигураций**. В соответствии с этим методом вначале выполняют обычную серию из n шагов покоординатного спуска, затем делают дополнительный шаг в направлении вектора $\mathbf{X}_k - \mathbf{X}_{k-n}$

как показано на рис. 5.7, где дополнительный шаг выполняют в направлении вектора $\mathbf{X}_3 - \mathbf{X}_1$, что и приводит в точку \mathbf{X}_4 .

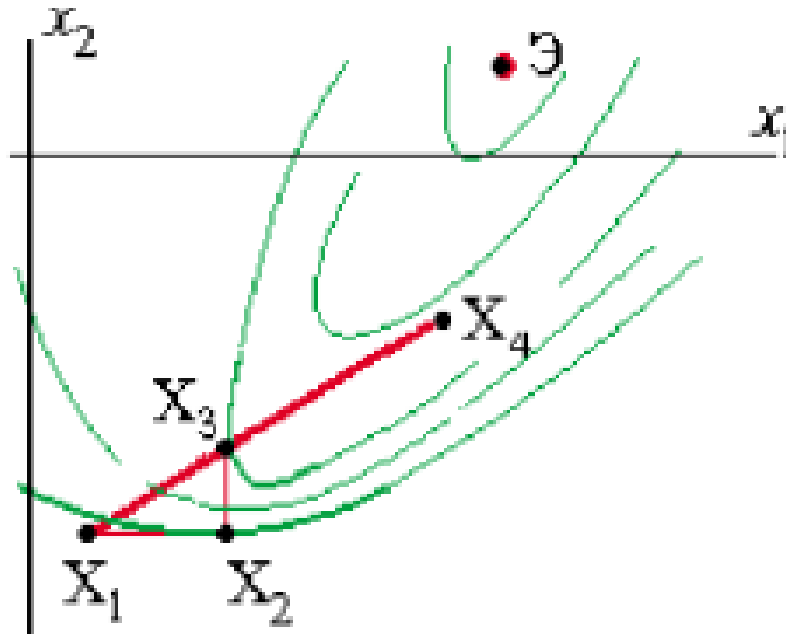


Рис. 5.7 Метод конфигураций

5.5 Методы многомерной оптимизации

Поиск экстремума *методом деформируемого многогранника* основан на построении многогранника с $(n + 1)$ вершинами на каждом шаге поиска, где n – размерность пространства управляемых параметров. В начале поиска эти вершины выбирают произвольно, на последующих шагах выбор подчинен правилам метода.

Эти правила поясняются рис. 5.8 на примере двумерной задачи оптимизации. Выбраны вершины исходного треугольника: X_1 , X_2 , X_3 . Новая вершина X_4 находится на луче, проведенном из худшей вершины X_1 (из вершины с наибольшим значением целевой функции) через центр тяжести $ЦТ$ многогранника, причем рекомендуется X_4 выбирать на расстоянии d от $ЦТ$, равном $|ЦТ - X_1|$. Новая вершина X_4 заменяет худшую вершину X_1 . Если оказывается, что X_4 имеет лучшее значение целевой функции среди вершин многогранника, то расстояние d увеличивают. На рисунке именно эта ситуация имеет место и увеличение d дает точку X_5 . В новом многограннике с вершинами X_2 , X_3 , X_5 худшей является вершина X_2 , аналогично получают вершину X_6 , затем вершину X_7 и т.д. Если новая вершина окажется худшей, то в многограннике нужно сохранить лучшую вершину, а длины всех ребер уменьшить, например, вдвое (стягивание многогранника к лучшей вершине). Поиск прекращается при выполнении условия уменьшения размеров многогранника до некоторого предела.

5.5 Методы многомерной оптимизации

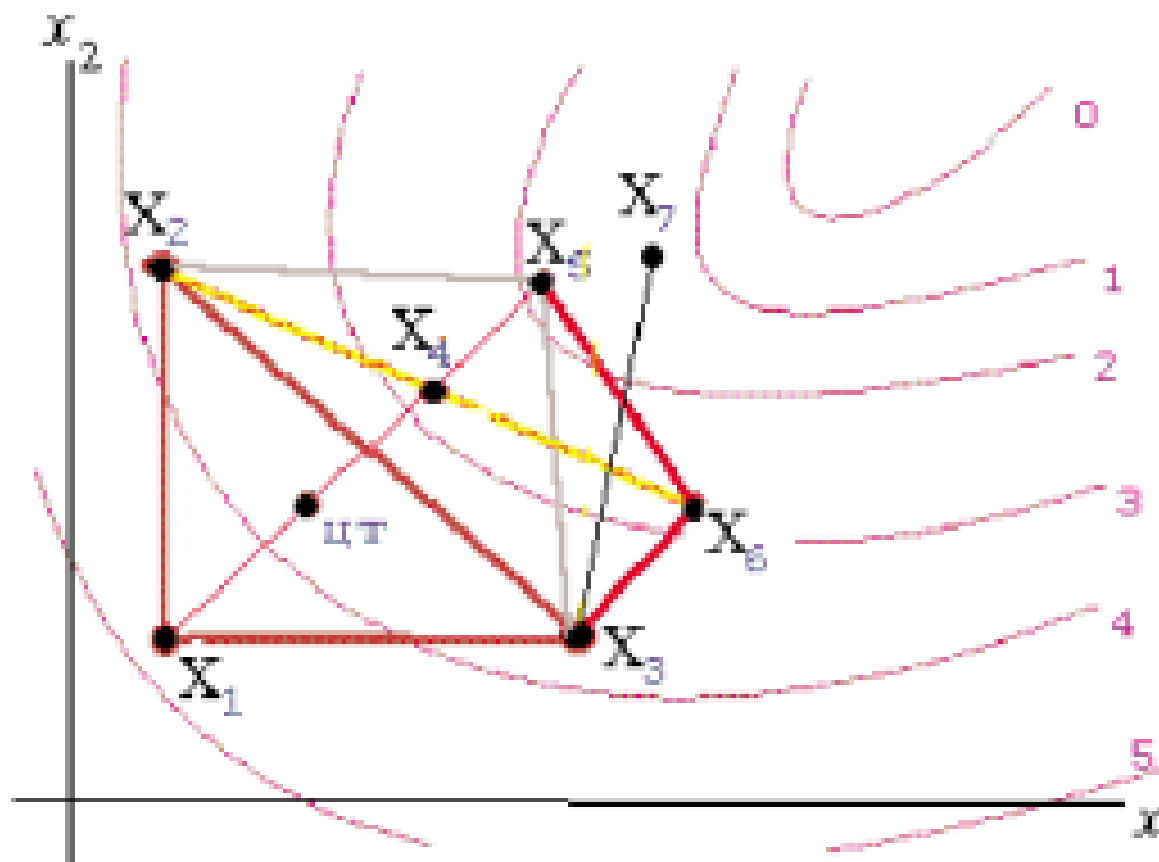


Рис. 5.8. Метод деформируемого многогранника

5.5 Методы многомерной оптимизации

Случайные методы поиска характеризуются тем, что направления поиска g выбирают случайным образом.

Особенностью **метода наискорейшего спуска** является выполнение шагов поиска в антиградиентном направлении:

$$\mathbf{X}_{k+1} = \mathbf{X}_k + h \cdot (-\text{grad}F(\mathbf{X})) / |\text{grad}F(\mathbf{X})|, \quad (5.21)$$

шаг h выбирается оптимальным с помощью одномерной оптимизации. Элементы вектор-градиента могут вычисляться по аналитическим зависимостям или численными методами. В последнем случае говорят о **квазиградиентных методах**.

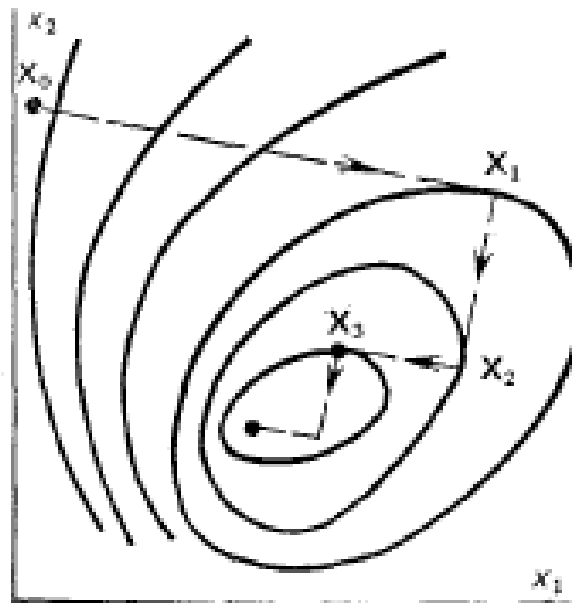


Рис. 5.9 Метод наискорейшего спуска

5.5 Методы многомерной оптимизации

При использовании метода наискорейшего спуска, как и большинства других методов, эффективность поиска существенно снижается в овражных ситуациях. Траектория поиска приобретает зигзагообразный вид с медленным продвижением вдоль дна оврага в сторону экстремума. Чтобы повысить эффективность градиентных методов, используют несколько приемов.

5.5 Методы многомерной оптимизации

Спасибо за внимание!