

## 4. Обработка и анализ экспериментального материала

### 4.1 Цель обработки экспериментального материала

Обработка экспериментального материала проводится с целью:

1. Проверки адекватности (правильности) разработанных физической и математической моделей реальным процессу или системе (подсистеме).

2. Получения значений не поддающихся установлению теоретическим путем параметров или поправочных коэффициентов.

3. Установления связей между параметрами процесса или системы (подсистемы), так называемых регрессионных зависимостей.

Различают обработку:

#### 1. Прямого и непрямого экспериментального материала.

**Прямой** экспериментальный материал – это материал, полученный непосредственно из опыта, например, упругий момент в валопроводе, усилие на грани резца, скорость движения машины или тела и т.д.

**Непрямой** экспериментальный материал – это материал, полученный путем расчета по установленным соотношениям с использованием прямого экспериментального материала, например, коэффициент трения скольжения.

#### 2. Экспресс и полную обработку экспериментального материала.

**Экспресс-обработка** – получение нулевого, в лучшем случае – первого приближения значений исследуемых величин. Основан этот метод обработки на принципе “равенства площадей”, рис.4.1, суть которого сводится к приравниванию площади под кривой на продолжительности (длине)  $T$  к площади прямоугольника, одна сторона которого равна  $T$ , а вторая – искомому среднему значению  $x_0$ , т.е.

$$\int_t^{t+T} x(t) dt = x_0 T .$$

Откуда

$$x_0 = \frac{1}{T} \int_t^{t+T} x(t) dt$$

Площадь под кривой может определяться с помощью планиметра. Средний уровень (среднее значение) обрабатываемой величины может быть оценен по этому методу и визуально – прямая, параллельная оси абсцисс проводится таким образом, чтобы

заштрихованные площади под прямой и над ней были бы примерно равны, т.е.

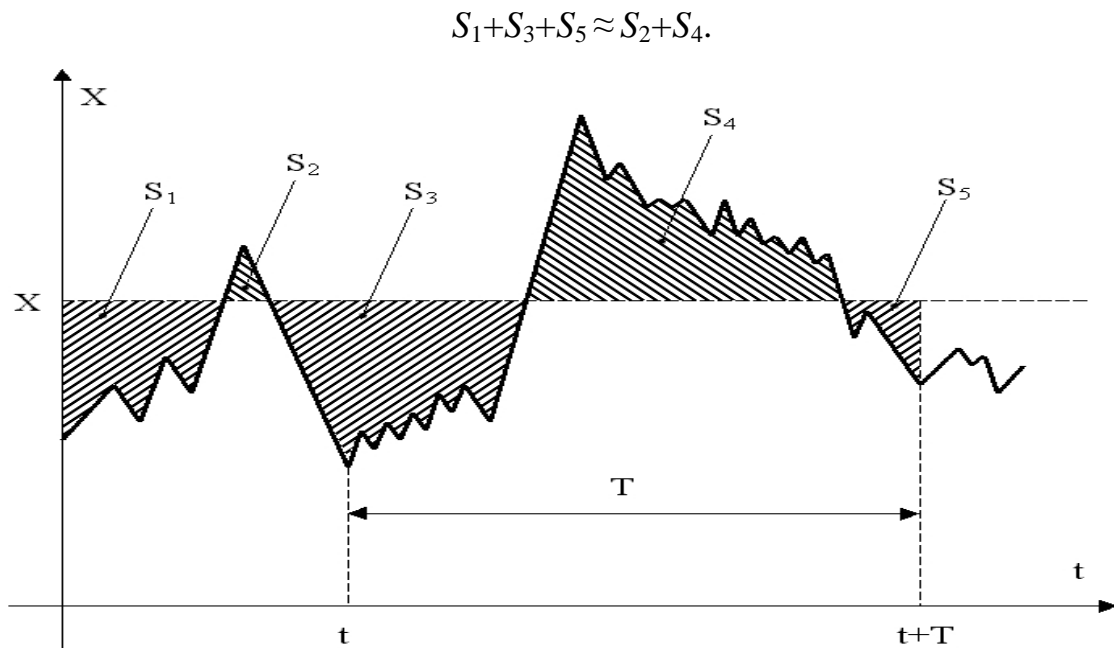


Рис.4.1 - К определению среднего значения величины по методу равенства площадей

**Экспресс-обработка** может быть использована для оценки и некоторых других параметров процесса, например, выделение низкочастотной составляющей, ее частоты (периода), амплитуды и т.д. в случае, когда процесс близок к детерминированному, см. рис. 4.2. В этом случае

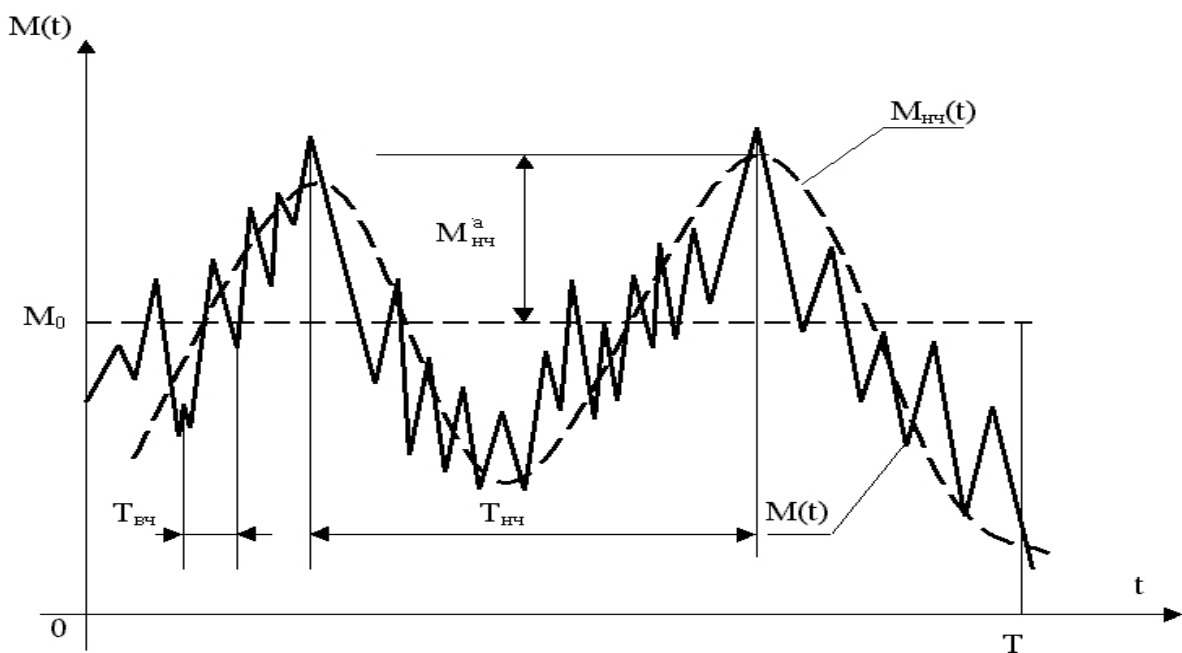


Рис.4.2-Выделение составляющих процесса при экспресс-обработке

$$M_0 = \frac{1}{T} \int_0^T M(t) dt ;$$

$$\omega_{нч} = 2\pi/T_{нч} ; \quad \omega_{вч} = 2\pi/T_{вч} .$$

**Полная обработка** экспериментального материала производится для получения доброкачественных оценок исследуемых величин, проверки адекватности моделей, установления связей между параметрами и решения ряда других задач.

Производится полная обработка материала с использованием, как правило, вычислительных машин и, поэтому, еще называется машинной обработкой с использованием известных стандартных программных средств.

#### **4.2 Задачи, решаемые при подготовке обработки результатов эксперимента**

При обработке полученного экспериментального материала, кроме достаточности экспериментального материала, решается ряд других задач, к которым в первую очередь следует отнести:

1. Подготовка экспериментального материала к вводу его в машину.
2. Ввод и автоматический контроль правильности ввода материала в машину.
3. Запись введенного экспериментального материала на носитель информации с выдачей начального адреса и длины массива.

##### **1. Подготовка экспериментального материала к вводу его в машину.**

Записанная на бумаге (магнитной ленте) физическая величина называется реализацией этой величины. Подготовка экспериментального материала к вводу его в машину включает в себя:

1. Получение копии реализации исследуемой величины (например, при помощи сканера, др. способом) с простановкой «нуля» - начало отсчета.
3. Кодирование копии реализации.

Обычно реализация представляет совокупность записей нескольких физических величин, см. рис. 4.3. Кодировующее устройство кодирует, как правило, одну реализацию. Это обуславливает необходимость в последовательном расположении реализаций на носителе, либо последовательную обработку копий реализации. При этом для каждой копии реализации физической

величины должен быть указан свой «нуль» - начало отсчета так, как показано, например, на рис.4.4.

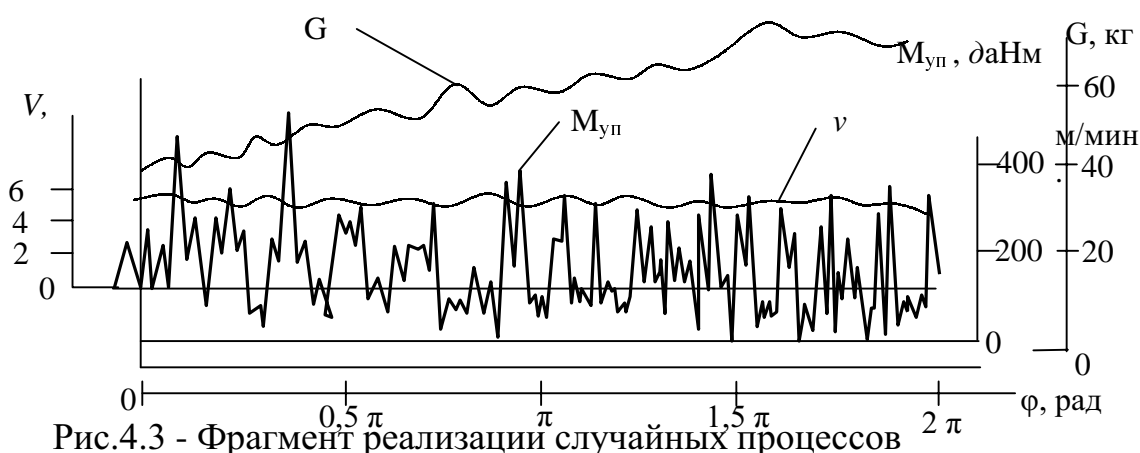


Рис.4.3 - Фрагмент реализации случайных процессов

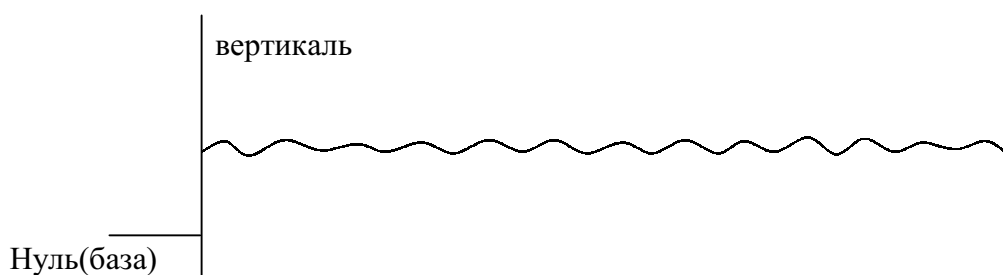


Рис.4.4 - Пример выкопировки реализации V

## 2. Ввод и автоматический контроль правильности ввода материала в машину.

Правильность ввода материала в машину проверяется за несколько этапов (некоторые программные средства проделывают эту процедуру автоматически):

1. Повторный ввод массива и сравнение его длины с первоначальной, т.е. проверка  $n_1 = n_2$ , где  $n_1, n_2$  – размерности первоначально и повторно введенного массива.

2. Контроль массива на максимум проверяется по условию  $x_i \leq C_{k,y}$ ,  $i = \overline{1, n}$ , где  $x_i$  –  $i$ -тый элемент массива;  $C_{k,y}$  – максимальное значение (константа) кодирующего устройства.

3. Контроль соседних (рядом расположенных) значений (отсчетов) на скорость изменения. При этом проверяется условие  $x_{j+1,k} - x_{j,k} \leq C_k$ ,  $j = \overline{1, n}$ ;  $k$  –  $k$ -тый массив; где  $C_k$  – константа, зависящая от типа считывающего устройства.

Если указанное условие не выполняется, то выполняется сравнение  $x_{j+2,k} - x_{j,k} \leq C_k$  и при его выполнении элементу массива  $x_{j+1,k}$  присваивается значение  $0,5(x_{j+2,k} + x_{j,k})$  – линейная интерполяция. Затем переходят к значению  $x_{j+3,k}$  и т.д.

### 3. Запись введенного экспериментального материала на носитель информации.

Запись введенного в ЭВМ материала на носитель осуществляется с обязательным указанием начального адреса и длины массива.

Полная обработка экспериментального материала зависит от вида решаемой задачи. Для большинства инженерных задач эта обработка включает оценки:

1. Числовых характеристик параметров;
2. Выделение составляющих процессов;
3. Построение гистограмм распределения;
4. Корреляционный и спектральный анализ;
5. Регрессионный анализ и другие задачи.

### 4.3 Анализ экспериментального материала

Рассмотрим вопрос выделения составляющих процессов на примере выделения низкочастотной составляющей методом “скользящей средней (площадки)”. Для обработки принимается фрагмент реализации случайного процесса продолжительностью  $T$ . На данной длине с шагом  $\Delta t$  считываются значения самой величины  $x_j$ .

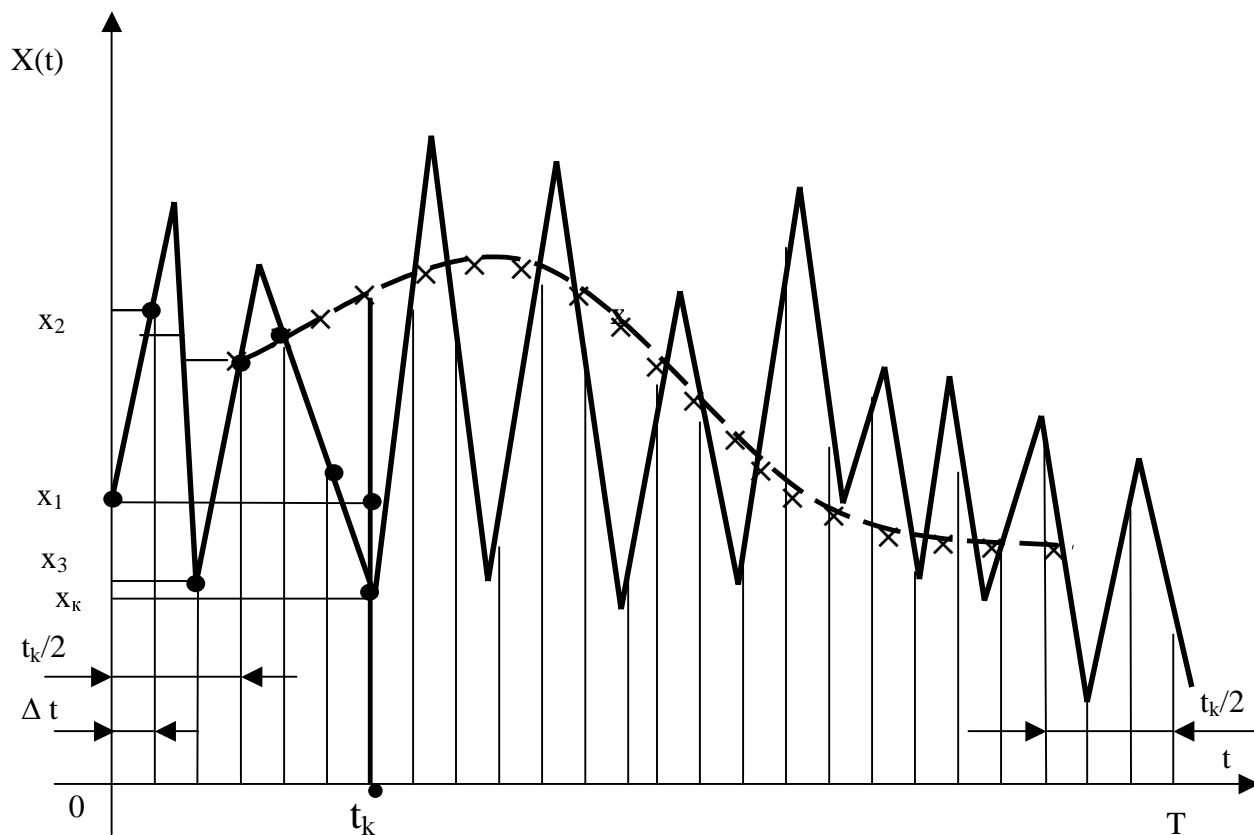


Рис.4.5 - Графическое представление выделения низкочастотной составляющей процесса методом скользящей средней

Затем на реализацию величины  $x(t)$ , начиная с ее первого значения  $x_1$ , накладывается “площадка” длиной  $t_k$ , рис.4.5. В пределах этой площадки рассчитывается среднее значение величины  $x(t)$  по зависимости

$$\bar{x}_1(t_k) = N_k^{-1} \sum_{j=1}^{N_k} x_j,$$

где  $N_k$  - число отчетов величины  $x(t)$ , приходящихся на площадку длиной  $t_k$ .

Полученное значение  $\bar{x}_1(t_k)$  приводится к середине площадки, т.е. к  $t_{k/2}$ . Далее площадка сдвигается вправо, т.е. в сторону приращения координаты  $t$ , на величину  $\Delta t$  и рассчитывается среднее значение

$$\bar{x}_2(t_k) = N_k^{-1} \sum_{j=1}^{N_k} x_{j+1},$$

которое приводится к середине площадки, т.е. к  $t_{k/2} + \Delta t$ . При  $n$ -ном сдвиге площадки среднее значение величины  $x(t)$  рассчитывается по зависимости

$$\bar{x}_n(t_k) = N_k^{-1} \sum_{j=1}^{N_k} x_{j+(n-1)\Delta t}$$

и приводится к координате  $t_{k/2} + (n-1)\Delta t$ .

Таким образом, из реализации процесса можно выделить низкочастотную составляющую. Недостатком этого способа является потеря информации, которая заключается в отбрасывании значений величины  $x(t)$  на интервале времени  $t_{k/2} + t_{k/2} = t_k$ .

Как было указано выше, специалисту-исследователю приходится изучать связи между явлениями, которые часто являются статистическими. При этом данному значению входной величины соответствует множество значений выходной. Установить степень тесноты связи входной величины с выходной можно с помощью **коэффициента корреляции**. Однако при получении информации данные необходимо проверять по критериям грубых ошибок. Это касается данных, существенно отличающихся от остальных. Существует несколько таких критериев, одним из которых является сравнение выборочного значения с величиной доверительного интервала для данной выборки.

Предположим, что мы располагаем данными нескольких параллельных опытов. Пусть какой-либо результат, с нашей точки зрения, существенно отличается от остальных. Как выявить, допущена или нет грубая ошибка?

Сделаем следующее. По уравнению

$$\bar{x} - t_{\alpha} S_x \leq x_i \leq \bar{x} + t_{\alpha} S_x$$

для имеющейся выборки с определенной доверительной вероятностью (пусть это будет  $P=0,95$ ) определяется доверительный интервал. Здесь  $t_\alpha$  - критерий Стьюдента;  $S_x$  – среднее квадратичное отклонение выборки. Если сомнительный результат выходит за пределы интервала, то он является грубой ошибкой, его следует исключить и оценку всех параметров провести заново (некоторые программные средства позволяют выполнить эти процедуры автоматически).

**Пример:** Имеются значения выборки производительности роторной линии: 105; 95; 100; 103; 99 и 98 шт/мин., а  $S_x = 3,76$  шт/мин. Из справочной таблицы для числа степеней свободы  $f = n - 1 = 6 - 1 = 5$ , определим  $t_\alpha = 2,571$ . Тогда доверительный интервал для отдельных размеров

$$\bar{x} \pm t_\alpha S_x = (100 \pm 2,571 \cdot 3,76) = 100 \pm 9,66 \text{ шт/мин,}$$

следовательно, с вероятностью  $P=0,95$  все данные приведенной выборки однородны. Замеры 95 и 105 шт/мин не являются грубыми ошибками.

### Проверка однородности дисперсий

Допустим, требуется экспериментально установить связь между параметром объекта и фактором. Для этого на каждом контрольном уровне фактора для повышения точности результатов будем проводить несколько параллельных замеров параметра. В результате эксперимента получим ряд выборок. Каждая из них характеризуется своим математическим ожиданием и своей дисперсией. В общем случае для дисперсий можно записать:

$$S_1^2 \neq S_2^2 \neq S_3^2 \neq \dots \neq S_n^2.$$

Наиболее распространенным методом является проверка с помощью **критерия Фишера**. Этот критерий ( $F$  – критерий) представляет отношение большей дисперсии к меньшей

$$F = \frac{S_1^2}{S_2^2}, \text{ где } S_1^2 > S_2^2.$$

Расчетное значение критерия сравнивают с критерием табличным, определяемым для принятого уровня значимости и соответствующих дисперсиям  $S_1^2$  и  $S_2^2$  степеней свободы  $f_1$  и  $f_2$ . Значение  $F$ -критерия

$$F_\alpha(f_1; f_2) \neq F_\alpha(f_2; f_1).$$

Если расчетное значение  $F < F_\alpha(f_1; f_2)$ , то дисперсии однородны и вместо  $S_1^2$  и  $S_2^2$  можно пользоваться средневзвешенным значением:

$$S_{CB}^2 = \frac{S_1^2 F_1 + S_2^2 F_2}{F_1 + F_2}.$$

В случае, когда мы располагаем не 2, а набором  $u$  дисперсий, то с помощью  $F$ -критерия проверяется однородность наименьшей и наибольшей. Если различие между ними окажется незначительным, то другие дисперсии тем более будут однородными. В этом случае для всей совокупности средневзвешенная дисперсия определяется из выражения

$$S_{CB}^2 = \frac{\sum_{u=1}^N S_u^2 F_u}{\sum_{u=1}^N F_u},$$

где  $N$  – число дисперсий.

### Сравнение двух выборочных средних

Пусть усовершенствована конструкция специального приспособления для выполнения определенной технологической операции. Для оценки полезности сделанного поставлен сравнительный эксперимент: проведен ряд замеров производительности до и после усовершенствования. В результате эксперимента получены две выборки. Первая – число замеров  $n_1$ ; математическое ожидание  $\bar{x}_1$ ; дисперсия  $S_1^2$ . Вторая – число замеров  $n_2$ ; математическое ожидание  $\bar{x}_2$ ; дисперсия  $S_2^2$ . При этом  $\bar{x}_1 \neq \bar{x}_2$ . Ставится вопрос: можно ли утверждать, что различие между  $\bar{x}_1$  и  $\bar{x}_2$  значимо? Иными словами можно ли утверждать, что реконструкция дала положительные результаты.

Обозначим  $\bar{x}_1 - \bar{x}_2 = Z$ .

$Z$  будет значимой величиной, если выйдет за пределы доверительного интервала

$$\mu_z - t_\alpha S_z \leq Z \leq \mu_z + t_\alpha S_z.$$

Предположим, что различие между  $\bar{x}_1$  и  $\bar{x}_2$  незначимо, т.е.  $\mu_z = 0$ . Тогда доверительный интервал  $Z = \pm t_\alpha S_z$ . Но

$$S_z^2 = \left( \frac{\partial Z}{\partial x_1} \right)^2 \cdot S_{x_1}^2 + \left( \frac{\partial Z}{\partial x_2} \right)^2 \cdot S_{x_2}^2,$$



тогда

$$S_z^2 = S_{x_1}^2 + S_{x_2}^2;$$

так как  $S_x^2 = \frac{S_x^2}{n}$ , то  $S_z^2 = \frac{S_{x_1}^2}{n_1} + \frac{S_{x_2}^2}{n_2}$ .

Если  $S_{x_1}^2$  и  $S_{x_2}^2$  однородны, а только в этом случае можно сравнить  $\overline{x_1}$  и  $\overline{x_2}$ , то

$$S_z^2 = S_{CB}^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right) \text{ или } S_z = S_{CB} \sqrt{\frac{n_1 + n_2}{n_1 n_2}}.$$

Таким образом, доверительный интервал

$$Z = \pm t_\alpha S_{CB} \sqrt{\frac{n_1 + n_2}{n_1 n_2}}.$$

Критерий  $t_\alpha$  берется для соответствующего уровня значимости и числа степеней свободы  $f = n_1 + n_2 - 2$  [23].

**Пример.** В одинаковых условиях проведения испытаний двух конструкций специального приспособления для фрезерно-центровального станка МР-74. Получены следующие равновероятностные данные для производительности:

- конструкция № 1: 39; 34; 58; 45; 34; 38 шт/час.
- конструкция № 2: 40; 47; 50; 53; 48; 45 шт/час.

Найдем оценки математических ожиданий

$$\begin{aligned} \overline{x_1} &= \frac{39 + 38 + 51 + 45 + 34 + 38}{6} = 41,3 \text{ шт/час;} \\ \overline{x_2} &= \frac{40 + 47 + 50 + 53 + 49 + 45}{6} = 47,3 \text{ шт/час.} \end{aligned}$$

На первый взгляд кажется, что реконструкция привела к существенному увеличению производительности приспособления № 2 – на 14,5%. Проверим достоверность этого.

Определим оценки дисперсий -  $S_1^2 = 83,07$  (шт/час)<sup>2</sup>,  $S_2^2 = 20,27$  (шт/час)<sup>2</sup>. Определим

$$S_x = M_k S ; \quad S_{x_1} = 1,051\sqrt{83,07} = 9,58 \text{ шт/час};$$

$$S_{x_2} = 1,051\sqrt{20,27} = 4,73 \text{ шт/час}.$$

Для проверки однородности параллельных опытов в выборках определим доверительный интервал

$$\bar{x}_1 \pm t_\alpha S_{x_1} = 41,3 \pm 2,571 \cdot 9,58 = (41,3 \pm 24,64) \text{ шт/час};$$

$$\bar{x}_2 \pm t_\alpha S_{x_2} = 47,3 \pm 2,571 \cdot 4,73 = (47,3 \pm 12,16) \text{ шт/час}.$$

Таким образом, каждая из выборок состоит из однородных результатов. Проверим однородность дисперсий. Определим расчетное значение критерия Фишера.

$$F = \frac{S_1^2}{S_2^2} = \frac{83,07}{20,27} = 4,1.$$

Табличное значение F критерия при уровне значимости  $\alpha=0,05$  и  $f_1=f_2=5$  составляет 5,05. Следовательно, дисперсии однородны.

Определим средневзвешенную дисперсию

$$S_{CB}^2 = \frac{23,07 \cdot 5 + 20,27 \cdot 5}{5 + 5} = 51,67 \text{ (шт/час)}^2.$$

Тогда доверительный интервал

$$Z = \pm 2,228\sqrt{51,67} \cdot \sqrt{\frac{6+6}{6 \cdot 6}} = \pm 9,25 \text{ шт/час}.$$

Так как интервал включает различие между математическими ожиданиями

$$\bar{x}_2 - \bar{x}_1 = 47,3 - 41,3 = 6 \text{ шт/час},$$

то с вероятностью  $P=0,95$  экспериментальный материал не дает оснований утверждать, что конструкция приспособления № 2 привела к значимому увеличению производительности.

После установления тесноты связи между случайными величинами корреляционный анализ следует продолжить еще на один шаг – установить аналитическую зависимость между выходной и входной величинами, т.е. выполнить регрессионный анализ. **Задача получения аналитической зависимости решается за три этапа:**

1. Выбор вида уравнения регрессии;
2. Определение значений коэффициентов уравнения;
3. Проверка соответствия (адекватности) установленной зависимости экспериментальному материалу.

Первый этап представляет неформализованную процедуру. Для облегчения выбора вида аналитической зависимости в прямоугольной системе координат с равномерными шкалами на осях наносят точки, полученные экспериментально. Координаты точек – значения аргумента и функции. При выборе вида зависимости следует стремиться к тому, чтобы выбранная кривая проходила бы как можно ближе к точкам. Не следует ориентироваться на сложные формулы, так как ценность формулы определяется не сложностью, а той ошибкой, которую допускаем мы при их использовании. Следует также заметить, что для сложных формул возрастает объем вычислений.

Определение значений коэффициентов уравнения регрессии в настоящее время, как правило, проводится с применением ЭВМ по стандартным программам.

В случае, когда вид уравнения предполагаем, то наиболее достоверные результаты будут получены при определении коэффициентов уравнения регрессии с использованием метода наименьших квадратов. В этом случае коэффициенты уравнения регрессии ищутся такими, чтобы сумма квадратов отклонений экспериментальных данных от значений, вычисленных по формуле, была минимальной.

Рассмотрим методику определения коэффициентов для простейшей зависимости

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 .$$

Так как во всех опытах  $x_0=1$ , то это уравнение можно представить в виде

$$\hat{y} = b_0x_0 + b_1x_1 + b_2x_2 .$$

Для определения коэффициентов необходимо, чтобы сумма отклонений стремилась к минимуму, т.е.

$$\sum_{u=1}^n [y_u - (b_0 x_0 + b_1 x_{1u} + b_2 x_{2u})]^2 = \min ,$$

где  $y_u$  – значение функции в эксперименте;

$b_0 x_0 + b_1 x_{1u} + b_2 x_{2u}$  – значение функции при расчете.

Сумма квадратов отклонений будет минимальной, если ее частные производные по  $b_0, b_1, b_2$  будут равны нулю, т.е.

$$\frac{\partial \sum}{\partial b_0} = 0; \quad \frac{\partial \sum}{\partial b_1} = 0; \quad \frac{\partial \sum}{\partial b_2} = 0;$$

откуда, сократив двойку:

$$\sum_{u=1}^n [y_u - (b_0 x_0 + b_1 x_{1u} + b_2 x_{2u})] x_0 = 0;$$

$$\sum_{u=1}^n [y_u - (b_0 x_0 + b_1 x_{1u} + b_2 x_{2u})] x_{1u} = 0;$$

$$\sum_{u=1}^n [y_u - (b_0 x_0 + b_1 x_{1u} + b_2 x_{2u})] x_{2u} = 0.$$

После преобразования получим

$$\sum_{u=1}^n y_u x_0 - b_0 \sum_{u=1}^n x_0^2 - b_1 \sum_{u=1}^n x_{1u} x_0 - b_2 \sum_{u=1}^n x_{2u} x_0 = 0;$$

$$\sum_{u=1}^n y_u x_{1u} - b_0 \sum_{u=1}^n x_0 x_{1u} - b_1 \sum_{u=1}^n x_{1u}^2 - b_2 \sum_{u=1}^n x_{2u} x_{1u} = 0;$$

$$\sum_{u=1}^n y_u x_{2u} - b_0 \sum_{u=1}^n x_0 x_{2u} - b_1 \sum_{u=1}^n x_{1u} x_{2u} - b_2 \sum_{u=1}^n x_{2u}^2 = 0.$$

Однако вследствие нормировки

$$\sum_{u=1}^n x_{iu}^2 = n;$$

вследствие ортогональности

$$\sum_{u=1}^n x_{1u} x_{ju} = 0,$$

получим

$$\sum_{u=1}^n y_u x_{0u} - b_{0u} = 0;$$

$$\sum_{u=1}^n y_u x_{1u} - b_{1u} = 0;$$

$$\sum_{u=1}^n y_u x_{2u} - b_{2u} = 0.$$

В случае, когда у нас имеется уравнение с 3 факторами,

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3,$$

то к полученным уравнениям добавится еще одно:

$$\sum_{u=1}^n y_u x_{3u} - b_{3u} = 0.$$

Таким образом, в общем случае коэффициент регрессии определится из уравнения

$$\sum_{u=1}^n y_u x_{iu} - b_{iu} = 0.$$

Отсюда следует, что при любом числе факторов

$$b_i = \frac{\sum_{u=1}^n y_u x_{iu}}{n}.$$

В уравнении регрессии, например,

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2$$

все коэффициенты имеют размерность параметра  $y$ , т.к.  $x_1$  и  $x_2$  – безразмерные величины. Кроме этого факторы  $x_1$  и  $x_2$  имеют одинаковый порядок. Поэтому коэффициенты  $b_1$  и  $b_2$  отражают силу влияния соответствующего фактора.

Величина коэффициента соответствует вкладу данного фактора в параметр при переходе с нулевого на верхний или нижний уровень. По уравнению можно вычислить значение параметра при любом дробном значении факторов.

Уравнение можно привести к виду, в котором факторы будут иметь свое первоначальное физическое значение. Для этого необходимо в него подставить

$$x_i = \frac{x_i - x_{i0}}{\Delta x_i}.$$

### Оценка воспроизводимости опытов

Предположим, что при построении модели учтены все факторы, влияющие на параметр. Предположим также, что между  $y$  и  $x$  существует функциональная связь, и что никакие случайные величины не повлияют на объект. В этом случае каждый раз при постановке опыта с данным набором значений факторов будем получать наугад одно и тоже значение параметра. Воспроизводимость опытов будет полная. В реальных условиях учесть влияние абсолютно всех факторов нельзя, да и нет необходимости. Поэтому повторные (параллельные) опыты не дают абсолютно одинаковых результатов, т.е. существует ошибка опыта [23].

Ошибки оцениваются дисперсией. В рассматриваемом случае дисперсия называется **дисперсией воспроизводимости опытов**. Для ее получения каждую строку матрицы планирования повторяют несколько раз. С целью исключения систематических ошибок опыты повторяют в случайной последовательности. Эта процедура и называется **рандомизацией**. Практически поступают следующим образом. Предположим, что фактора 2. ПФЭ требует постановки 4 опытов. Для оценки воспроизводимости решили каждую строку повторить дважды, т.е. поставить восемь опытов. Нумеруем эти опыты, а затем, пользуясь таблицей случайных чисел, устанавливаем порядок их проведения.

После проведения опытов определяются построчные дисперсии

$$S_u^2 = \frac{1}{m-1} \sum_{j=1}^m (y_{uj} - \bar{y}_u)^2;$$

где  $m$  – число параллельных опытов;

$y_{ij}$  – значение параметра в опыте  $j$  данной  $i$  строки;  
 $\overline{y_u}$  – математическое ожидание параметра в строке.

$$S_{Bi}^2 = \sum_{u=1}^n \left( \frac{db_i}{dy_u} \right)^2 S_{y_u}^2.$$

Например, при  $n=4$  (2 фактора на двух уровнях)

$$b_0 = \frac{\overline{y_1} + \overline{y_2} + \overline{y_3} + \overline{y_4}}{4};$$

$$S_{B0}^2 = \frac{1}{16} S_{y_1}^2 + \frac{1}{16} S_{y_2}^2 + \frac{1}{16} S_{y_3}^2 + \frac{1}{16} S_{y_4}^2;$$

Так как дисперсии  $S_{y_1}^2$ ,  $S_{y_2}^2$  и т.д. однородны, то можно считать

$$S_{y_1}^2 \approx S_{y_2}^2 \approx S_{y_3}^2 \approx S_{y_4}^2 \approx S_{CB}^2 \text{ и}$$

$$S_{B0}^2 = \frac{S_{CB}^2}{4};$$

В общем случае

$$S_{B0}^2 = \frac{S_{CB}^2}{f}.$$

Число степеней свободы  $S_{Bi}^2$  равно  $f=n(m-1)$ .

Доверительный интервал для коэффициента  $b_i$

$$|\overline{b_i}| - t_\alpha S_{Bi} \leq |b_i| \leq |\overline{b_i}| + t_\alpha S_{Bi},$$

где  $t_\alpha$  – критерий Стьюдента.

Предположим, что коэффициент  $b_i$  незначим. В этом случае

$$|\overline{b_i}| = 0.$$

Для значимости коэффициентов должно выполняться условие:

$$|\overline{b_i}| > t_\alpha S_{Bi}.$$

Критерий  $t_\alpha$  принимается из таблицы для соответствующего уровня значимости  $\alpha$  и числа степеней свободы  $f=n(m-1)$ , но  $\alpha=1-p$ , где  $p$  – доверительная вероятность.

Однородность дисперсии  $S_u^2$  определяется с помощью критерия Фишера (отношение большей дисперсии к меньшей):

$$F = \frac{S_1^2}{S_2^2} < F_\alpha(f_1; f_2), \text{ где } S_1^2 > S_2^2.$$

Если дисперсии окажутся однородными, а только в этом случае по экспериментальным данным можно определить коэффициенты уравнения регрессии, они определяются по выражению:

$$S_{CB}^2 = \frac{\sum_{u=1}^N S_u^2 f_u}{\sum_{u=1}^N f_u},$$

где  $f_u$  – число степеней свободы;  
 $S_u^2$  - дисперсия величины;  
 $N$  - число дисперсий.

### Проверка значимости коэффициентов уравнения регрессии

Случайная величина является значимой, если с принятой вероятностью выходит за пределы соответствующего доверительного интервала.

Коэффициенты уравнения регрессии определяются с учетом параллельных опытов по выражению

$$b_i = \frac{\sum_{u=1}^n y_u x_{iu}}{n}.$$

В этом выражении  $n$  – величина точная,  $x_{iu}$  – считаем точной, т.к. предполагаем, что ошибка определения фактора на порядок меньше ошибки определения параметра. Тогда

$$b_i = f(\sum \bar{y}_u).$$

### Проверка адекватности уравнений регрессии

На первом этапе ищется дисперсия адекватности

$$S_{ад}^2 = \frac{1}{n - \ell} \sum_{u=1}^n (\bar{y}_u - \hat{y}_u)^2,$$

где  $\ell$  – число коэффициентов в уравнении регрессии;  
 $\hat{y}_u$  – рассчитанное по уравнению значение параметра.

Затем с использованием F-критерия проверяется однородность дисперсий адекватности и воспроизводимости.



**Пример.** Исследовать влияние режимов резания ( $v$ ,  $s$ ,  $t$ ) на силу резания при торцовом фрезеровании высокопрочного чугуна ВЧ 50-1,5 с пределами факторов  $v$  (90...150 м/мин);  $S_Z$  (2...4 мм/зуб) на станке 6Н13П с силоизмерительной трехкомпонентной головкой (учитывает тангенциальную  $P_T$ , радиальную  $P_R$  и осевую  $P_A$  силу).

Результаты реализации ПФЭ  $2^3$  по одной составляющей силы резания  $P_T$  приведены в таблице 4.1.

Здесь 
$$\bar{y}_u = \frac{y_{u1} + y_{u2} + y_{u3}}{3}.$$

Определим коэффициенты уравнения регрессии

$$b_0 = \frac{925 + 803 + 1358 + 1245 + 1568 + 1477 + 2021 + 1918}{8};$$

$$b_0 = \frac{\sum_{u=1}^n \bar{y}_u \cdot x_0}{n} \approx 1114,4 \text{ Н};$$

$$b_1 = \frac{\sum_{u=1}^n \bar{y}_u \cdot x_1}{n} = \frac{-925 + 803 - 1358 + 1245 - 1568 + 1477 - 2021 + 1918}{8} = -53,6 \text{ Н};$$

Таблица 4.1 – Результаты реализации ПФЭ  $2^3$

Номер серии опытов	Матрица планирования			Результаты экспериментов, Н			$\bar{y}_u$ , Н	$S_u^2$ , Н <sup>2</sup>
	$x_1$	$x_2$	$x_3$	$y_{u1}$	$y_{u2}$	$y_{u3}$		
1	-1	-1	-1	910	930	935	925	175
2	+1	-1	-1	797	821	791	803	252
3	-1	+1	-1	1390	1338	1346	1358	784
4	+1	+1	-1	1220	1251	1264	1245	511
5	-1	-1	+1	1540	1603	1561	1568	1029
6	+1	-1	+1	1504	1440	1487	1477	1099
7	-1	+1	+1	2030	2041	1992	2021	661
8	+1	+1	+1	1926	1900	1928	1918	244
сумма							4755	

$$b_2 = \frac{\sum_{u=1}^n \bar{y}_u \cdot x_2}{n} = \frac{-925 - 803 + 1358 + 1275 - 1568 - 1477 + 2021 + 1918}{8} = 221,1 \text{ Н};$$

$$b_3 = \frac{\sum_{u=1}^n y_u \cdot x_3}{n} = \frac{-925-803-1358-1245+1568+1477+2021+1918}{8} = 3316 H$$

Произведем оценку воспроизводимости опытов. Рассчитаем построчные дисперсии:

$$S_1^2 = \frac{1}{3-1} [(910-925)^2 + (930-925)^2 + (935-925)^2] = 175 (H^2);$$

$$S_2^2 = \frac{1}{3-1} [(797-803)^2 + (821-803)^2 + (791-803)^2] = 252 (H^2);$$

$$S_3^2 = 784 (H^2); S_4^2 = 511 (H^2); S_5^2 = 1029 (H^2);$$

$$S_6^2 = 1099 (H^2); S_7^2 = 661 (H^2); S_8^2 = 244 (H^2).$$

Результаты расчетов занесем в таблицу 4.1.

Наименьшее значение  $S_1^2 = 175 (H^2)$ , наибольшее -  $S_6^2 = 1099 (H^2)$ .

Проверим однородность дисперсий:

$$F_P = \frac{S_6^2}{S_1^2} = \frac{1099}{175} = 6,28.$$

Для уровня значимости  $\alpha=0,05$  и степеней свободы  $f_1=2$  и  $f_2=2$  ( $n-1=3-1=2$ ), табличное значение критерия Фишера  $F_\alpha = 19,0$ .

Так как  $F_P=6,28 < F_\alpha=19,0$ , то дисперсии однородны.

Дисперсия воспроизводимости

$$S_B^2 = S_{CB}^2 = \frac{175 \cdot 2 + 252 \cdot 2 + 784 \cdot 2 + 511 \cdot 2 + 1029 \cdot 2 + 1099 \cdot 2 + 661 \cdot 2 + 244 \cdot 2}{2+2+2+2+2+2+2+2};$$

$$S_B^2 = 594,4 (H^2).$$

Проведем проверку значимости коэффициентов уравнения регрессии. Определим дисперсию коэффициентов по формуле

$$S_{Bi}^2 = \frac{S_B^2}{n} = \frac{594,4}{8} = 74,3 (H^2).$$

Доверительный интервал для коэффициентов  $|b_i| = t_\alpha S_B$ .

Табличное значение критерия  $t_\alpha$  для уровня значимости 0,05 и числа степеней свободы  $f = n(m - 1) = 16$  составляет 2,143. Таким образом доверительный интервал

$$|b_i| = 2,143 \cdot \sqrt{74,3} = 18,5.$$

Следовательно, все коэффициенты значимы. Тогда уравнение регрессии будет иметь вид

$$\hat{y}_T = 1414,4 - 53,6x_1 + 221,1x_2 + 331,6x_3.$$

У нас  $x_1 \rightarrow V$ ;  $x_2 \rightarrow S$ ;  $x_3 \rightarrow t$ .

Анализ этой модели показывает, что увеличение подачи  $S_z$  и глубины резания  $t$  приводит к росту силы резания из-за увеличения сечения снимаемой стружки; при увеличении же скорости резания  $V$  наблюдается снижение силы резания (однако увеличивается тепловыделение). Таким образом, данное уравнение не только надежно описывает физическую сущность влияния режимов резания на процессы в зоне резания, но и вскрывает физическую сущность влияния режимов резания на процессы в зоне резания.

Чаще в инженерных задачах определяют критерии согласия, суть которых сводится к установлению меры случайности в несовпадении (отклонении) экспериментального и теоретического распределения вероятностей. Рассмотрим критерий Пирсона, который устанавливает меру случайностей между статистической  $p^*$  и теоретической  $p$  вероятностями, приняв в качестве расхождения

$$u = \sum_{i=1}^k C_i (p_i^* - p_i)^2,$$

где  $C_i$  – весовой коэффициент;  
 $k$  – число разрядов.

Установлено, что при  $C_i = np_i^{-1}$  закон распределения вероятностей для  $u$  зависит только от числа разрядов  $k$  и при увеличении  $n$  стремится к распределению  $\chi^2$ :

$$\chi^2 = n \sum_{i=1}^k C_i (p_i^* - p_i)^2 p_i^{-1},$$

или

$$\chi^2 = \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i}; \quad m_i = n p_i^*,$$

где  $m_i$  – число значений случайной величины  $\chi$  в  $i$ -том разряде.

Распределение  $\chi^2$  зависит от параметра  $r$ , называемого числом степеней свободы распределения:  $r = k - n_{CB}$ , где  $n_{CB}$  – число связей. В качестве связей принято принимать независимые условия

$$1. \sum_{i=1}^k p_i^* = 1; \quad 2. \sum_{i=1}^k \tilde{x}_i p_i^* = m_x; \quad 3. \sum_{i=1}^k (\tilde{x}_i - m_x)^2 p_i^* = D_x.$$

Величина  $\chi^2$  табулирована, входами в таблицу которой являются число степеней свободы распределения и доверительная вероятность.

## 5 ОСНОВНЫЕ ПОНЯТИЯ ОПТИМИЗАЦИИ МОДЕЛЕЙ

При определении параметров проектируемого изделия или процесса инженеру часто приходится решать задачи выбора этих параметров из множества вариантов по некоторому критерию. Такой выбор возможен при использовании методов оптимизации. В инженерном деле под оптимизацией (*optimum* – лат. наилучший) получают наилучшее из множества решение удовлетворяющее всем требованиям, предъявляемых к проектируемому объекту. Теория оптимизации включает в себя совокупность фундаментальных математических положений и численных методов нахождения наилучших вариантов из множества альтернатив и позволяющих избежать полностью перебора и оценки возможных вариантов [24].

### 5.1 Процедура оптимизации модели. Критерий оптимизации.

#### Методы линейного программирования. Симплекс-метод

При исследовании операций выделяют три главных направления:

1. Построение модели, т.е. формализация изучаемого процесса или явления.
2. Описание операции - постановка задачи.
3. Решение оптимизационной задачи.

Первый этап мы рассмотрели довольно подробно в первой главе. Третий этап требует больше участие математика, а не исследователя. А на втором этапе мы должны сформулировать оптимизационную задачу. Все условия записываются в виде соответствующей системы ограничений (уравнений или неравенств). Любое решение этой системы называется допустимым. Критерий оптимизации записывается в виде функции, которую называют целевой. А решение задачи оптимизации состоит в отыскании на множестве решений системы ограничений максимального или минимального значения целевой функции.

При решении оптимизации, возникающие на различных этапах технологического проектирования используют различные виды математических моделей и методы оптимизации, которые могут быть классифицированы по следующим признакам.

1. По структуре математические модели делят на модели с ограничениями и без них. Ограничения делят в свою очередь на линейные и нелинейные. Целевые функции подразделяются на линейные, квадратичные, разомкнутые и нелинейные.

2. По виду переменных математических моделей различают такие группы: с непрерывными значениями переменных; с дискретными значениями переменных. Во второй группе различают целочисленные и нецелочисленные переменные.

Для отыскания оптимума целевой функции могут быть использованы 4 группы методов: аналитические, рекурсивные, итерационные и стохастические. Аналитические методы применяют при решении классических задач с помощью исследования производной целевой функции.

Рекурсивные методы относят к методам, позволяющим определить одну переменную за одну расчетную операцию. Наиболее распространено среди таких методов динамическое программирование.

Стохастические методы основаны на методах поиска случайных решений.

Наибольшее же распространение получили итерационные методы, к которым относятся способы расчета функций цели в одной или нескольких вероятностных точках для определения “лучшей” точки. Различают два итерационных метода: линейное и нелинейное программирование.

Линейное программирование применяют для решения линейных задач, когда функция цели и ограничения являются линейными, а все переменные непрерывными функциями. Для методов нелинейного программирования характерно непосредственное отыскание оптимума. Они подразделяются на методы, базирующиеся на определении градиента, и методы, в которых не требуется расчет градиента.

С помощью численных методов мы непосредственно ищем минимум функции  $f(x)$  в некотором интервале  $a < x < b$ , в котором, как предполагается, лежит минимум, вычисляя значения функции в выбранных точках данного интервала. Иногда это единственно возможная стратегия поиска. Например, себестоимость проведения технологического процесса может зависеть от режимных параметров. Инженер знает, что стоимость является функцией от этих параметров, хотя может и не знать явного вида функции. Однако он может поставить эксперимент и провести процесс при различных режимных параметрах и, следовательно, найти себестоимости для этих параметров и надеяться определить минимальную себестоимость и лучшие режимы проведения процесса, при которых она достигается.

Можно попытаться найти положение минимума в точке, аппроксимирующей его с нужной точностью, или определить малый интервал, в котором находится минимум. Попытаемся достичь поставленной цели как можно более эффективным способом, т. е.

осуществляя наименьшее количество вычислений функции. Однако, поскольку проведение эксперимента требует определенных затрат, инженер захочет добиться этой точности, проведя как можно меньше экспериментов.

Предположим, что точки  $a$  и  $b$  определяют (возможно, очень грубо) интервал, который содержит истинную точку минимума, и внутри этого интервала функция унимодальна, т. е. имеет один минимум в точке  $x^*$ . Следовательно, данная функция имеет форму, близкую к той, что приведена на рис. 5.1.

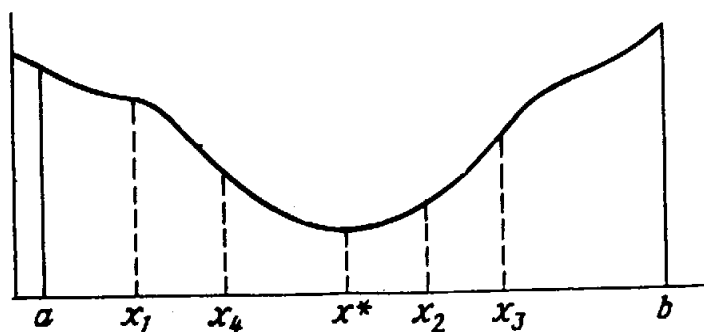


Рис. 5.1- Истинная точка минимума.

Если известны значения функции такого вида в трех точках  $x_1, x_2, x_3$ , таких, что  $a < x_1 < x_2 < x_3 < b$ , а  $f(x_2) < f(x_1)$  и  $f(x_2) < f(x_3)$ , то  $x_1 < x^* < x_3$ .

Тогда точка  $x^*$  будет лежать внутри интервала  $(x_1; x_3)$ , меньшего по размеру, чем интервал  $(a; b)$ .

В зависимости от вида оптимизации технологических процессов можно выделить следующие группы критериев оптимизации, или целевых функций:

1. Экономические, например, себестоимость.
2. Техничко - экономические, например, максимальная производительность.
3. Технологические, например, точность изготовления детали.
4. Эксплуатационные, например, усталостная прочность.
5. Прочие – психологические, эстетические, экологические др.

Рассмотрим методы линейного программирования. Наиболее известный из этих методов – симплекс-метод, автором которого считается американский математик Дж. Данциг. С помощью симплекс-метода решаются задачи оптимизации режимов резания металлов, а также транспортная задача, когда требуется составить план перевозок (перемещение заготовок по цеху) однородного груза так, чтобы общая стоимость перевозок была минимальной.

Рассмотрим конкретный пример. Но прежде следует сказать, что если число переменных системы ограничений и целевой функции

в математической модели задачи линейного программирования равно 2 или 3, то такую задачу можно решать графически. Итак, пример.

Процесс изготовления 2 видов изделий заводом требует:

- последовательной обработки заготовок на токарно – винторезных и вертикально – фрезерных станках;
- затрат 2 вида сырья: стали и цветных металлов.

Данные о потребности ресурсов приведены в таблице 5.1.

Таблица 5.1 - Потребность ресурсов

		Затраты на 1 изделие		Ресурсы
		А	В	
Материалы	Сталь, кг	10	70	320
	Цвет. мет., кг	20	50	420
Оборудование	Токарные ст., ст-ч	300	400	6200
	Фрезерные ст, ст-ч	200	100	34000
Прибыль на изделие, грн.		3,0	8,0	

Определим такой план выпуска продукции, который бы обеспечивал максимальную прибыль при условии, что время работы фрезерных станков должно быть использовано полностью.

Решение.

Обозначим через  $X$  число изделий А, а через  $Y$  - число изделий В. Из стали необходимо  $10 + 70Y \leq 320$ , а цветного металла  $20X + 50Y \leq 420$ . Время отработки на токарных станках  $300X + 400Y \leq 6200$ , а время обработки на фрезерных станках  $200X + 10Y = 3400$ . Кроме того,  $X \geq 0; Y \geq 0$ . Следовательно, система ограничений примет вид

$$\begin{cases} X + 7Y \leq 32 \\ 2X + 5Y \leq 42 \\ 3X + 4Y \leq 62 \\ 2X + Y = 34 \\ X \geq 0 \\ Y \geq 0. \end{cases}$$

Общая прибыль (целевая функция)  $F = 3X + 8Y$ .

Надо на множестве ограничений найти наибольшее значение  $F$ .



Найдем координаты множество точек плоскости, координаты которых удовлетворяют системе ограничений. Так как  $X \geq 0$  и  $Y \geq 0$ , то решение будем искать в I квадранте. Графическая интерпретации неравенств ограничений дает фигуру OACKL, рис. 5.2. Прямая  $2X + Y = 34$  в этой фигуре выделяет множество дополнительных планов, которые графически представляют собой точки отрезка EN. Тогда среди точек отрезка EN выберем такую, в которой  $F \rightarrow F_{\max}$ .

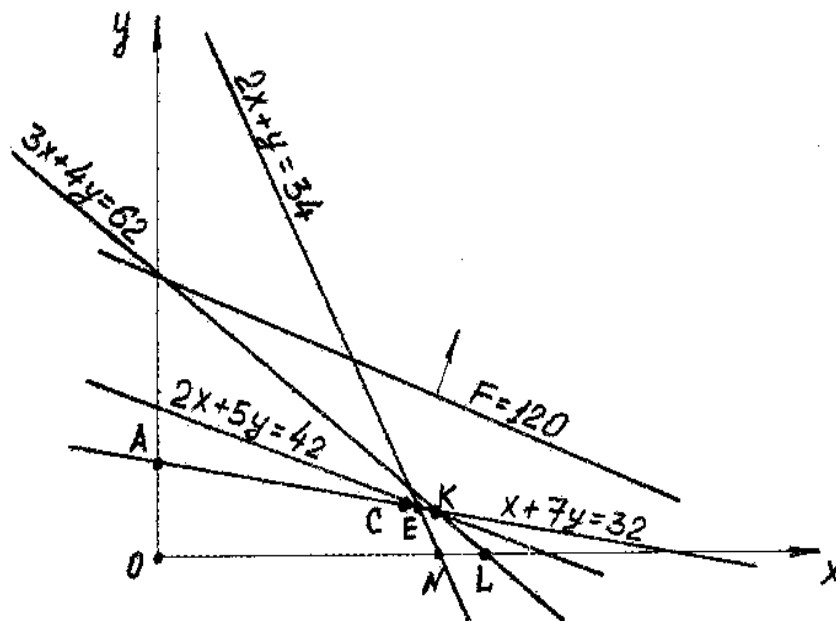


Рис. 5.2 - Графическое представление симплекс – метода

По уравнению  $3X + 8Y = C$  строим несколько прямых, придавая  $C$  произвольные значения. Например, 120, 240 и т.д. Следовательно, получаем семейство параллельных прямых. Причем, с увеличением  $C$  прямая  $F$  движется вверх - вправо, рис. 5.2. Последней точкой касания отрезка EN прямой  $F$  будет точка E. Найдём её координаты из совместного решения уравнений, описывающих прямые, на пересечении которых лежит искомая точка:

$$\begin{cases} 2X + Y = 34 \\ 2X + 5Y = 42. \end{cases}$$

Решением этой системы алгебраических уравнений будет  $X=16$  и  $Y=2$ .

Таким образом, получили оптимальный план:

$$F_{\max} = 3 \times 16 + 8 \times 2 = 64.$$

Таким образом, из графического решения задачи линейного программирования следует, что минимальное или максимальное значение целевой функции (если оно существует) достигается в

вершинах многогранника допустимых планов. Следовательно, основу симплекс - метода составляет такой алгоритм: найти значение целевой функции в какой либо вершине, затем, переходя к соседней вершине, попытаться уменьшить (увеличить) его и т.д. В общем виде это звучит так:

1. Система уравнений ограничений приводится к виду

$$\left\{ \begin{array}{l} x_1 = b_1 - (ax_{1,r+1} + \dots + a_{1n}x_n) \\ x_2 = b_2 - (a_{2,r+1}x_{r+1} + \dots + a_{2n}x_n) \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ x_r = b_r - (a_{r,r+1}x_{r+1} + \dots + a_{rn}x_n) \\ x_i \geq 0, \quad i = \overline{1, n} \end{array} \right. \quad (5.1)$$

где  $b_1 \dots b_r$  - неотрицательны;  $x_1 \dots x_r$  - базисные переменные базиса Б;  $x_{r+1} \dots x_n$  - свободные (небазисные) переменные.

Пусть все базисные переменные равны 0. Найдём из системы (5.1) значения базисных переменных:

$$x_1 = b_1; \quad x_2 = b_2; \quad \dots \quad x_r = b_r.$$

В результате получаем базисные решения  $(b_1, b_2, \dots, b_r, 0, \dots, 0)$ , соответствующие базису Б.

Целевая функция  $F(x_1, \dots, x_n)$  также выражается через небазисные переменные

$$F = C_0 + C'_{r+1} X_{r+1} + \dots + C'_n X_n.$$

2. Проверка достижения оптимума. Если оптимум не достигнут, то из базиса Б удаляется одна из переменных в небазисные, а вместо нее из числа прежних небазисных переменных вводится новая. Получаем новый базис  $B'$ . Далее – повторяется тоже, что и в 1 шаге. Если опять оптимум не достигнут, - все шаги повторяются снова, причем каждый новый шаг заключается в переходе от базиса Б к новому базису  $B'$ , такому, что значение  $F_{B'}$  уменьшается (увеличивается) или, по крайней мере, не увеличивается (не уменьшается), например,

$$F_{B'} \leq F_B.$$

Этот процесс оканчивается одним из 3 случаев: находится оптимум; доказываемая противоречивость ограничений; доказываемая неограниченность целевой функции при базисных решениях, т.е. когда задача решений не имеет.

## 5.2 Метод Нелдера – Мида

Метод Нелдера - Мида является развитием симплексного метода Спендли, Хекста и Химсворта. Множество  $(n + 1)$ -й равноудаленной точки в  $n$ -мерном пространстве называется регулярным симплексом. Эта конфигурация рассматривается в методе Спендли, Хекста и Химсворта. Следовательно, в двумерном пространстве симплексом является равносторонний треугольник, а в трехмерном пространстве - правильный тетраэдр. Идея метода состоит в сравнении значений функции в  $(n + 1)$  вершинах симплекса и перемещении симплекса в направлении оптимальной точки с помощью итерационной процедуры. В симплексном методе, предложенном первоначально, регулярный симплекс использовался на каждом этапе. Нелдер и Мид предложили несколько модификаций этого метода, допускающих, чтобы симплексы были неправильными. В результате получился очень надежный метод прямого поиска, являющийся одним из самых эффективных, если  $n < 6$ .

В методе Спендли, Хекста и Химсворта симплекс перемещается с помощью трех основных операций: отражения, растяжения и сжатия. Смысл этих операций станет понятным при рассмотрении шагов процедуры:

А. Найдем значения функции в вершинах симплекса.

Б. Найдем наибольшее значение функции  $f_h$ , следующее за наибольшим значением функции  $f_g$ , наименьшее значение функции  $f_l$  и соответствующие им точки  $x_h, x_g$  и  $x_l$ .

В. Найдем центр тяжести всех точек, за исключением точки  $x_h$ . Пусть центром тяжести будет

$$x_0 = \frac{1}{n} \sum_{i=h} x_i$$

и вычислим  $f(x_0) = f_0$ .

Г. Удобнее всего начать перемещение от точки  $x_h$ . Отразив точку  $x_h$  относительно точки  $x_0$ , получим точку  $x_r$  и найдем  $f(x_r) = f_r$ .

Д. Сравним значения функций  $f_r$  и  $f_l$ .

Е. Сравним значения функций  $f_r$  и  $f_h$ .

Ж. Сравним значения функций  $f_c$  и  $f_h$ .

З. На этом шаге мы уменьшаем размерность симплекса делением пополам расстояния от каждой точки симплекса до  $x_l$  — точки,

определяющей наименьшее значение функции.

### 5.3 Методы нелинейного программирования

#### Метод Фибоначчи

Метод Фибоначчи относится к группе методов нелинейного программирования. Причем, по этому методу оптимум ищется вдоль произвольно выбранного направления и не требуется расчета градиента.

Метод Фибоначчи может быть эффективно использован в случае, если переменная определена только в дискретных точках. Такая оптимизация имеет место при поиске, например, оптимальной частоты вращения шпинделя станка, если задан конкретный ряд чисел его оборотов.

Минимизация функции одного переменного является, как правило, необходимым этапом также многих методов минимизации многомерных функций. На первый, взгляд это просто: если функция  $\varphi(x)$ , которую надо минимизировать на отрезке  $[a, b]$ , дифференцируема, то надо найти  $\varphi'(x)=0$ , присоединить к этим нулям концы отрезка, выделить из этих точек локальные минимумы и найти точку, в которой имеется абсолютный (глобальный) минимум.

Однако задача взятия производной функции и решения уравнения  $\varphi'(x)=0$  бывает весьма сложной. Следовательно, важное значение приобретают методы минимизации, не требующие вычисления производной. Одним из таких методов является как раз, метод Фибоначчи, заключающийся в построении отрезков  $[a_k, b_k]$ , стягивающихся к точке  $x^*$  минимума функции  $\varphi(x)$  на исходном отрезке  $[a_0, b_0]$ .

Свойства этого метода:

1. Для любого  $k=0, 1 \dots k$  - м приближением  $x_k$  точки  $x^*$  является одна из 2 точек:  $y_k$  либо  $z_k$  отстоящих на одинаковых расстояниях

$$y_k - a_{k-1} = b_{k-1} - z_k$$

от соответствующих концов отрезка  $[a_k, b_k]$ .

Для этого должны выполняться следующие соотношения:

$$y_k = a_{k+1} + c_{k-1}(b_{k-1} - a_{k-1}); \quad (5.2)$$

$$z_k = a_{k-1} + (1 - c_{k-1})(b_{k-1} - a_{k-1}), \quad (5.3)$$

$$\text{где число } c_{k-1} \text{ выбирается так, чтобы } 0 \leq c_{k-1} \leq 0,5. \quad (5.4)$$

2. На каждом шаге (за исключением первого) допускается вычисление одного значения функции  $\varphi(x)$ .

Из выражений (5.2 – 5.4) будем находить отрезок  $[a_k, b_k]$  следующим образом:

$$\left. \begin{array}{l} \text{если } \varphi(y_k) \leq \varphi(z_k), \text{ то } a_k \leq x^* \leq z_k, \\ \text{и выбираем } a_k = a_{k-1}, b_k = z_k, x_k = y_k \end{array} \right\} \quad (5.5)$$

$$\left. \begin{array}{l} \text{если } \varphi(y_k) > \varphi(z_k), \text{ то } y_k \leq x^* \leq b_{k-1}, \\ \text{и выбираем } a_k = y_k, b_k = b_{k-1}, x_k = z_k \end{array} \right\} \quad (5.6)$$

Анализ процедуры (5.2 – 5.6). Из (5.2) и (5.3) следует, что

$$y_k - a_{k-1} = b_{k-1} - z_k. \quad (5.7)$$

Следовательно, из (5.5) и (5.6) получаем

$$b_k - a_k = z_k - a_{k-1} = b_{k-1} - y_k = (1 - c_{k-1})(b_{k-1} - a_{k-1}) \quad (5.8)$$

Обращаясь к (5.5) и (5.6) видим, что

$$x_k \in [a_k, b_k] \text{ и } \varphi(x_k) = \min\{\varphi(y_k), \varphi(z_k)\}.$$

Теперь в соответствии со вторым свойством, потребуем, чтобы на каждом шаге (за исключением первого, когда необходимы 2 вычисления) вычислялось одно значение функции  $\varphi(x)$ . Возникает соотношение:

$$x_k = \begin{cases} z_{k+1}, & \text{если } \varphi(y_k) \leq \varphi(z_k) \\ y_{k+1}, & \text{если } \varphi(y_k) > \varphi(z_k). \end{cases} \quad (5.9)$$

Чтобы обеспечить выполнение условий (5.9), необходимо подобрать числа  $c_k$  соответствующим образом.

Пусть  $\varphi(y_k) \leq \varphi(z_k)$ . Тогда из (5.2), (5.5) и (5.8) следует:

$$x_k = y_k = a_{k-1} + c_{k-1}(b_{k-1} - a_{k-1}) = a_k + \frac{c_{k-1}}{1 - c_{k-1}}(b_k - a_k).$$

$$\text{Если } 1 - c_k = \frac{c_{k-1}}{1 - c_k}, \quad k=1, 2, \dots, \quad (5.10)$$

$$\text{т. е. } c_k = \frac{1-2c_{k-1}}{1-c_{k-1}}, \quad k=1,2,\dots,$$

то из (5.3) следует; что  $x_k=z_{k+1}$ .

Если  $\varphi(y_k) > \varphi(z_k)$ , то (смотри выражения (5.6), (5.3), (5.2), (5.10))

$$\begin{aligned} x_k = z_k &= a_{k-1} + (1+c_{k-1})(b_{k-1} - a_{k-1}) = a_{k-1} + c_{k-1}(b_{k-1} - a_{k-1}) + (1-2c_{k-1})(b_{k-1} - a_{k-1}) = \\ &= y_k + (1-2c_{k-1})(b_{k-1} - a_{k-1}) = a_k + \frac{1-2c_{k-1}}{1-c_{k-1}}(b_k - a_k) = y_{k+1}. \end{aligned}$$

Таким образом, соотношения (5.10) обеспечивают выполнение условия (5.9).

Из соотношения (5.10) получаем

$$c_k = \frac{-1+3c_{k-2}}{c_{k-2}} \quad (5.11)$$

и далее

$$c_k = \frac{F_i - F_{i+2}c_{k-i}}{F_{i-2} - F_i c_{k-i}}, \quad k=3,4,\dots, \quad i=3,4,\dots,k; \quad (5.12)$$

где

$$F_j = F_{j-1} + F_{j-2}, \quad j=3,4,\dots \quad (5.13)$$

Числа  $F_j$ , удовлетворяющие этому соотношению и условию  $F_1=F_2=1$ , называются числами Фибоначчи. Некоторые из этих чисел:

$$F_3=2; F_4=3; F_5=5; F_6=8; F_7=13; F_8=21; F_9=34; F_{10}=55 \text{ и т.д.}$$

Сущность метода Фибоначчи.

Будем выбирать числа  $c_k$  из следующих соображений: при заранее заданном числе шагов  $n$ , т.е.  $k=1,2,\dots,n$ , числа  $c_k$  выберем такими, чтобы длина отрезка  $[a_n, b_n]$  была наименьшей.

Так как из (5.8) следует

$$b_n - a_n = \prod_{j=0}^{n-1} (1-c_j)(b_0 - a_0), \quad (5.14)$$

то возникает задача

$$\min_{c_j} \prod_{j=0}^{n-1} (1-c_j) \quad (5.15)$$

при условиях  $0 \leq c_j \leq 0,5$ ;

$$1 - c_1 = \frac{c_0}{1 - c_0}; \quad (5.16)$$

$$1 - c_2 = \frac{1 - 2c_0}{c_0}; \quad (5.17)$$

$$1 - c_j = -\frac{F_{j-1} - F_{j+1}c_0}{F_{j-2} - F_jc_0}, \quad j = 3, 4, \dots, n-1 \quad (5.18)$$

Условия (5.16 – 5.18) с очевидностью вытекают из условий (5.10 – 5.12).

Из условий (5.15 – 5.18) следует

$$\prod_{j=0}^{n-1} (1 - c_j) = (1 - c_0) \frac{c_0}{1 - c_0} \times \frac{1 - 2c_0}{1 - c_0} \times \frac{1 - 2c_0}{c_0} \times \frac{F_2 - F_4c_0}{F_1 - F_3c_0} \dots \frac{F_{n-3} - F_{n-1}c_0}{F_{n-4} - F_{n-2}c_0} \times \frac{F_{n-2} - F_n c_0}{F_{n-3} - F_{n-1}c_0} (-1)^{n-1}.$$

Таким образом,

$$\prod_{j=0}^{n-1} (1 - c_j) = \begin{cases} F_{n-2} - F_n c_0, & \text{если } n - \text{нечетное;} \\ -F_{n-2} + F_n c_0, & \text{если } n - \text{четное.} \end{cases} \quad (5.19)$$

Легко проверить, что из условия  $0 \leq c_k \leq 0,5$ ,  $k = 0, 1, \dots$  и соотношения

$$c_k = \frac{F_k - F_{k+2}c_0}{F_{k-2} - F_k c_0}, \quad k = 3, 4, \dots, n,$$

следует для  $k \geq 3$

$$\frac{F_k}{F_{k+2}} \leq c_{n-k} \leq \frac{F_{k-1}}{F_{k+1}}, \quad \text{если } k - \text{четное;}$$

$$\frac{F_{k-1}}{F_{k+1}} \leq c_{n-k} \leq \frac{F_k}{F_{k+1}}, \quad \text{если } k - \text{нечетное,}$$

$$\text{и } 0 \leq c_{n-1} \leq 0,5, \quad \frac{1}{3} \leq c_{n-2} \leq \frac{1}{2}.$$

Тогда задача (5.15), (5.4), (5.16), (5.17), (5.18) сведется к двум.

Если  $n$  - нечетное, то для  $n \geq 3$

$$\min_{c_0} (F_{n-2} - F_n c_0),$$

$$\frac{F_{n-1}}{F_{n+1}} \leq c_0 \leq \frac{F_n}{F_{n+2}}.$$

$$\text{Откуда } c_0 = \frac{F_n}{F_{n+1}}. \quad (5.20)$$

Если  $n$  - четное, то для  $n \geq 4$

$$\min_{c_0} (F_n c_0 - F_{n-2})$$

$$\frac{F_n}{F_{n+2}} \leq c_0 \leq \frac{F_{n-1}}{F_{n+1}}.$$

$$\text{Откуда } c_0 = \frac{F_n}{F_{n+1}}$$

и кроме того,  $c_0 = 0,5$  при  $n=1$ ,  $c_0 = 1/3$  при  $n=2$ .

Наконец, вычислим минимальное значение величины

$$\prod_{j=1}^{n-1} (1 - c_j).$$

Для этого используем соотношение

$$F_n^2 - F_{n-2} F_{n+2} = (-1)^n.$$

Тогда из (5.19) и (5.20) следует, что  $\prod_{j=0}^{n-1} (1 - c_j) = 1/(F_{n+2})$ .

$$\text{Таким образом, } b_n - a_n = F_{n+2}^{-1} (b_0 + a_0). \quad (5.21)$$

Это соотношение позволяет выбрать число  $n$  при заданной точности  $\varepsilon > 0$  вычисления точки  $x^*$ .

$$\text{Так как } |x^* - x_n| \leq b_n - a_n,$$

то величина  $n$  должна обеспечивать выполнение неравенств

$$b_n - a_n \leq \varepsilon < b_{n-1} - a_{n-1},$$

$$\text{или } \frac{b_0 - a_0}{F_{n+2}} \leq \varepsilon \leq \frac{b_0 - a_0}{F_{n+1}}.$$

Следовательно, метод Фибоначчи при заданном числе шагов  $n$  обеспечивает минимальную длину отрезка  $[a_n, b_n]$ , содержащего точку  $x^*$ .

В заключение приведем вычислительную схему этого метода, обеспечивающую минимальную погрешность вычислений:

I. Выбираем в качестве  $n$  наименьшее натуральное число,



удовлетворяющее неравенствам:

$$F_{n+1} < \frac{1}{\varepsilon} (b_0 - a_0) \leq F_{n+2}.$$

2. Первый шаг:

$$y_1 = a_0 + \frac{F_n}{F_{n+2}} (b_0 - a_0);$$

$$z_1 = a_0 + \frac{F_{n+1}}{F_{n+2}} (b_0 - a_0).$$

Вычисляем  $\varphi(y_1)$  и  $\varphi(z_1)$ .

Числа  $a_1$  и  $b_1$  выбираем так:

$$a_1 = \begin{cases} a_0, & \text{если } \varphi(y_1) \leq \varphi(z_1); \\ y_1, & \text{если } \varphi(y_1) > \varphi(z_1); \end{cases}$$

$$b_1 = \begin{cases} z_1, & \text{если } \varphi(y_1) \leq \varphi(z_1); \\ b_0, & \text{если } \varphi(y_1) > \varphi(z_1); \end{cases}$$

3.  $k$ -й шаг ( $2 \leq k \leq n$ ). Пусть известны:  $a_{k-1}, b_{k-1}, \varphi(y_{k-1}), \varphi(z_{k-1})$ . Если  $\varphi(y_{k-1}) \leq \varphi(z_{k-1})$ , то вычисляем

$$y_k = a_{k-1} + \frac{F_{n-k+1}}{F_{n+2}} (b_0 - a_0) \text{ и } \varphi(y_k)$$

и полагаем  $z_k = y_{k-1}; \varphi(z_k) = \varphi(y_{k-1})$ .

Если  $\varphi(y_{k-1}) > \varphi(z_{k-1})$ , то полагаем  $y_k = z_{k-1}; \varphi(y_k) = \varphi(z_{k-1})$

и вычисляем  $z_k = a_{k+1} + \frac{F_{n-k+2}}{F_{n+2}} (b_0 - a_0)$ .

Числа  $a_k$  и  $b_k$  выбираем так:

$$a_k = \begin{cases} a_{k-1}, & \text{если } \varphi(y_k) \leq \varphi(z_k); \\ b_0, & \text{если } \varphi(y_k) > \varphi(z_k); \end{cases}$$

$$b_k = \begin{cases} z_k, & \text{если } \varphi(y_k) \leq \varphi(z_k); \\ b_{k-1}, & \text{если } \varphi(y_k) > \varphi(z_k); \end{cases}$$

## 5.4 Поиск методом "золотого сечения"

Не всегда можно заранее определить, сколько раз придется вычислять функцию. В методе Фибоначчи это нужно знать для

определения  $L_2$ , т. е. положения начальной точки от данного из концов начального интервала.

Метод "золотого сечения" почти столь же эффективен, как и метод Фибоначчи, однако при этом не требуется знать  $n$  — количество вычислений функции, определяемое вначале. После того как выполнено  $j$  вычислений, исходя из тех же соображений, что и ранее, записываем

$$L_{j-1} = L_j + L_{j-1}$$

Здесь  $L$  — интервал неопределенности, содержащий оптимальное значение функции.

Однако если  $n$  неизвестно, то мы не можем использовать условие  $L_{n-1} = 2L_n - \varepsilon$ . Если отношение последующих интервалов будет постоянным, т. е.

$$\frac{L_{j-1}}{L_j} = \frac{L_j}{L_{j+1}} = \frac{L_{j+1}}{L_{j+2}} = \dots = \tau$$

то

$$\frac{L_{j-1}}{L_j} = 1 + \frac{L_{j+1}}{L_j}$$

Таким образом,  $\tau^2 - \tau - 1 = 0$ , откуда  $\tau = (1 + \sqrt{5})/2 \approx 1,618033989$ .

Тогда

$$\frac{L_{j-1}}{L_{j+1}} = \tau^2, \frac{L_{j-2}}{L_{j+1}} = \tau^3$$

Следовательно,  $\frac{L_1}{L_n} = \tau^{n-1}$

$$L_n = \frac{L_1}{\tau^{n-1}}$$

В результате анализа двух рассмотренных значений функции будет определен тот интервал, который должен исследоваться в дальнейшем. Этот интервал будет содержать одну из предыдущих точек и следующую точку, помещаемую симметрично ей. Первая точка находится на расстоянии  $L_1/\tau$  от одного конца интервала, вторая — на таком же расстоянии от другого. Поскольку  $\lim_{n \rightarrow \infty} F_{n-2}/F_n = 1/n$ , то видно, что поиск методом "золотого сечения"

является предельной формой поиска методом Фибоначчи. Название "золотое сечение" произошло от названия отношения в уравнении . Видно, что  $\tilde{y}_{i+1}$  делится на две части так, что отношение целого к большей части равно отношению большей части к меньшей, т. е. равно так называемому "золотому отношению".

## 5.5 Аппроксимация кривыми

В этом разделе применяется иной подход - найти малый интервал, в котором находится минимум функции. Используется несколько значений функции в определенных точках для аппроксимации функции обычным полиномом по крайней мере в небольшой области значений. Затем положение минимума функции аппроксимируется положением минимума полинома, поскольку последний вычислить проще.

### 5.5.1 Квадратичная интерполяция

Если известны значения функции  $f(x)$  в трех различных точках  $\alpha, \beta, \gamma$ , равные соответственно  $f_\alpha, f_\beta, f_\gamma$ , то функция  $f(x)$  может быть аппроксимирована квадратичной функцией

$$\varphi(x) = Ax^2 + Bx + C,$$

где  $A, B, C$  – определяются из уравнений

$$A\alpha^2 + B\alpha + C = f_\alpha$$

$$A\beta^2 + B\beta + C = f_\beta$$

$$A\gamma^2 + B\gamma + C = f_\gamma$$

После преобразований этих уравнений получаем

$$A = [(\gamma - \beta) f_\alpha + (\alpha - \gamma) f_\beta + (\beta - \alpha) f_\gamma] / \Delta,$$

$$B = [(\beta^2 - \gamma^2) f_\alpha + (\gamma^2 - \alpha^2) f_\beta + (\alpha^2 - \beta^2) f_\gamma] / \Delta,$$

$$C = [\beta\gamma(\gamma - \beta) f_\alpha + \gamma\alpha(\alpha - \gamma) f_\beta + \alpha\beta(\beta - \alpha) f_\gamma] / \Delta,$$

где  $\Delta = (\alpha - \beta)(\beta - \gamma)(\gamma - \alpha)$ . Ясно, что  $\varphi(x)$  будет иметь минимум в точке  $x = -B/2A$ , если  $A > 0$ . Следовательно, можно аппроксимировать точку минимума функции  $f(x)$  значением

Этот метод может непосредственно применяться к функциям одной переменной. Он может быть очень полезен для выполнения линейного поиска в процедурах градиентных методов. В этих процедурах требуется найти минимум функции  $f(x)$  в точках прямой  $x_0 + \lambda d$ , где  $x_0$  - заданная точка, а  $d$  определяет заданное направление. Значения функции  $f(x_0 + \lambda d)$  на этой прямой являются значениями функции одной переменной  $\lambda$ :

$$\varphi(\lambda) = f(x_0 + \lambda d)$$

### 5.5.2 Кубическая интерполяция

Квадратичная интерполяция часто называется методом Пауэлла и аппроксимирует функцию квадратичным трехчленом. Метод Давидона обеспечивает большую точность и аппроксимирует функцию кубическим полиномом. Для кубической интерполяции в этом методе используются значения функции и ее производной, вычисленные в двух точках. Этот метод широко используется в процедурах линейного поиска, и именно с этой точки зрения мы и будем его рассматривать.

Необходимо минимизировать функцию  $f(x)$  на прямой  $x_0 + hd$ , т.е. минимизировать функцию

$$\varphi(h) = f(x_0 + hd) = f(x_{01} + hd_1, x_{02} + hd_2, \dots, x_{0n} + hd_n).$$

Тогда, сократив промежуточные выкладки,

$$\frac{d\varphi}{dh} = \nabla f(x_0 + hd)^T d = g(x_0 + hd)^T d.$$

$$\delta = 0.5 \left[ \frac{(\beta^2 - \gamma^2)f_\alpha + (\gamma^2 - \alpha^2)f_\beta + (\alpha^2 - \beta^2)f_\gamma}{(\beta - \gamma)f_\alpha + (\gamma - \alpha)f_\beta + (\alpha - \beta)f_\gamma} \right]$$

Предполагаем, что известны следующие значения

$$\varphi(p) = \varphi_p; \quad \varphi(q) = \varphi_q; \quad \frac{d\varphi}{dh}(p) = G_p; \quad \frac{d\varphi}{dh}(q) = G_q.$$

Эту информацию можно использовать для построения кубического полинома

$$a + bh + ch^2 + dh^3,$$

который будет аппроксимировать функцию  $\varphi(h)$ . Если  $p=0$ , то уравнения, определяющие  $a, b, c, d$ , выглядят так:

$$\begin{aligned} a &= \varphi_p; \\ a + bq + cq^2 + dq^3 &= \varphi_q; \\ b &= G_p; \\ b + 2cq + 3dq^2 &= G_q. \end{aligned}$$

Положим, что точка  $r$  является точкой минимума кубического полинома. Тогда:

$$\begin{aligned} \frac{r}{q} &= \frac{G_p + z \pm \omega}{G_p + G_q + 2z}, \\ \text{где} \quad z &= \frac{3(\varphi_p - \varphi_q)}{q} + G_p + G_q; \\ \omega &= (z^2 - G_p G_q)^{0,5}. \end{aligned}$$

Эта итерационная процедура имеет следующие шаги:

1. Найти  $\varphi_p=f(x_0)$  и  $G_p=[g(x_0)]^T d$ .
2. Проверить, выполняется ли условие  $G_p < 0$ , и если оно не выполняется, производить поиск вдоль направления  $-d$ .
3. Вычислить  $\varphi_q=f(x_0+qd)$  и  $G_q=[g(x_0+qd)]^T d$ .
4. Если  $G_q > 0$  или  $\varphi_q > \varphi_p$ , то интервал, содержащий минимум, найден. В противном случае заменить  $q$  на  $2q$  и вернуться к шагу 3.
5. Использовать уравнение для аппроксимации точки минимума на интервале  $(0, q)$  значением  $r$ .
6. Если  $|d\varphi/dh| = |[g(x_0+rd)]^T d| = |G_r| < \varepsilon$ , где  $\varepsilon$  - заданная точность, то остановиться.
7. Вернуться на шаг 5, используя интервал  $(0, r)$ , если  $G_r < 0$ , либо используя интервал  $(r, q)$ , если  $G_r \leq 0$ .

На шаге 6 производится проверка значения производной. Предшествующие проверки приводят к остановке тогда, когда положение минимума не изменяется. Следует отметить, что в общем случае проще найти минимум функции, чем положение точки минимума. Последнее определяется с меньшей точностью.

## 5.6 Метод Хука – Дживса

Этот метод был разработан в 1961 году, но до сих пор является

весьма эффективным и оригинальным. Поиск состоит из последовательности шагов последующего поиска вокруг базисной точки, за которой в случае успеха следует поиск по образцу.

Описание этой процедуры представлено ниже:

А. Выбрать начальную базисную точку  $b_1$  и шаг длиной  $h_j$  для каждой переменной  $x_j$ ,  $j = 1, 2, \dots, n$ . Для каждой переменной используется шаг  $h$ , однако указанная выше модификация тоже может оказаться полезной.

Б. Вычислить  $f(x)$  в базисной точке  $b_1$  с целью получения сведений о локальном поведении функции  $f(x)$ . Эти сведения будут использоваться для нахождения подходящего направления поиска по образцу, с помощью которого можно надеяться достичь большего убывания значения функции.

В. При поиске по образцу используется информация, полученная в процессе исследования, и минимизация функции завершается поиском в направлении, заданном образцом.

Г. Завершить этот процесс, когда длина шага (длины шагов) будет уменьшена до заданного малого значения.

## 5.7 Градиентные методы. Метод наискорейшего спуска

Рассмотрим методы поиска, в которых наряду со значениями функции используется и ее градиент. С помощью метода покоординатного спуска осуществляется поиск из заданной точки в направлении, параллельном одной из осей, до точки минимума в данном направлении. Затем поиск производится в направлении, параллельном другой оси, и т. д. Направления, конечно, фиксированы. Кажется разумным попытаться модифицировать этот метод таким образом, чтобы на каждом этапе поиск точки минимума производился вдоль "наилучшего" направления. Не ясно, какое направление является "наилучшим", но известно, что направление градиента является направлением наискорейшего возрастания функции. Следовательно, противоположное направление является направлением наискорейшего убывания функции.

Здесь возникает задача максимизации с ограничением.

В методе наискорейшего спуска желательно использовать свойство направления градиента. Это свойство заключается в том, что направления градиента перпендикулярно в любой точке линии постоянного уровня, поскольку вдоль этой линии функция постоянна. Таким образом, если  $(d_1, d_2, \dots, d_n)$  – малый шаг вдоль линии уровня, то  $f(x_1+d_1, x_2+d_2, \dots, x_n+d_n) = f(x_1, x_2, \dots, x_n)$  и следовательно,

$$df = \sum_{j=1}^n \frac{\partial f}{\partial x_j} d_j = |\nabla f(x)|^T d = 0.$$

Поэтому, если мы находимся в точке  $x_i$ , на некотором шаге процесса оптимизации, то поиск минимума функции осуществляется вдоль направления  $-\nabla f(x_i)$ . Данный метод является итерационным. На шаге  $i$  точка минимума аппроксимируется точкой  $x_i$ .

### 5.7.1 Квадратичные функции

Квадратичная функция

$$F(x) = a + x^T b + \frac{1}{2} x^T G x$$

где  $a$  — константа;  $b$  — постоянный вектор и  $G$  — положительно определенная симметрическая матрица — имеет минимум в точке  $x^*$ , причем  $x^*$  определяется следующим образом:

$$\nabla F(x^*) = b + Gx^* = 0$$

откуда  $x^* = -G^{-1}b$ . Следует, что при выполнении условия непрерывности любую функцию можно аппроксимировать в окрестности точки  $x_0$  функцией

$$\varphi(x) = f(x_0) + (x - x_0)^T \nabla f(x_0) + \frac{1}{2} (x - x_0)^T G(x_0) (x - x_0),$$

где  $G(x_0)$  — матрица Гессе, вычисленная в точке  $x_0$ .

Разумной аппроксимацией минимума функции  $f(x)$  может быть минимум функции  $\varphi(x)$ . Если последний находится в точке  $x_m$ , то

$$\nabla f(x_0) + G(x_0)(x_m - x_0) = 0,$$

откуда

$$x_m = x_0 - G^{-1}(x_0) \nabla f(x_0),$$

или

$$x_m = x_0 - G^{-1}(x_0) g(x_0)$$

Таким образом следует модифицировать итерационное уравнение, и точкой  $x$ , следующей аппроксимации минимума будет

$$x_{i+1} = x_i - G^{-1}(x_i) g(x_i)$$

или, и в более удобном виде

$$x_{i+1} = x_i - \lambda_i G^{-1}(x_i)g(x_i)$$

где длина шага  $\lambda_i$ , определяется одномерным поиском в направлении  $G^{-1}(x_i)g(x_i)$ .

Метод Ньютона — Рафсона основан на последнем уравнении. Не будем рассматривать его во всех подробностях, упомянем лишь некоторые его особенности. Если точка  $x_i$ , близко расположена к точке  $x^*$ , то сходимость будет быстрой, поскольку в общем случае функция  $\varphi(x)$  будет хорошо аппроксимировать функцию  $f(x)$  в этой окрестности. Как норму градиента  $|g(x_{i+1})|$ , так и расстояние между точками  $|x_{i+1}-x_i|$  следует проверить на выполнение критерия завершения. Интересно отметить, что по сравнению с простым методом наискорейшего спуска направлением спуска в данном случае будет не  $-g(x_i)$ , а  $-G^{-1}(x_i)g(x_i)$ , где учитываются и вторые производные. Методом Давидона — Флетчера — Пауэлла можно получить наилучший результат, производя поиск на  $i$ -м этапе в направлении  $-H_i g(x_i)$ , где  $H_i$  - положительно определенная симметрическая матрица, которая в конечном счете становится равной  $-G^{-1}(x^*)$ . Таким образом, этот метод обходит как вычисление, так и обращение матрицы  $G(x_i)$  при каждой итерации.

Следовательно, направление поиска при каждой итерации является решающим фактором, если говорить об эффективности итерационных методов поиска. При каждой итерации желательно произвести одномерный поиск в наилучшем направлении. Для квадратичной функции  $n$  переменных наилучшим направлением является направление, сопряженное с предыдущим направлением поиска.

Ясно, что не следует сразу принимать решение о направлениях поиска, а лучше накапливать информацию, полученную на предыдущих этапах поиска, для того чтобы определять дальнейшие направления поиска.

Следовательно, если поиск производится по взаимно сопряженным направлениям, то минимум квадратичной функции  $n$  переменных будет найден не более чем за  $n$  шагов.

### 5.7.2. Метод Давидона - Флетчера - Пауэлла

Метод Давидона — Флетчера — Пауэлла (ДФП) основан на использовании соотношений

$$\nabla f(x_0) + G(x_0)(x_m - x_0) = 0,$$

откуда

$$x_m = x_0 - G^{-1}(x_0)\nabla f(x_0),$$

или



$$x_m = x_0 - G^{-1}(x_0)g(x_0)$$

и

$$x_{i+1} = x_i - G^{-1}(x_i)g(x_i)$$

или, и в более удобном виде

$$x_{i+1} = x_i - \lambda_i G^{-1}(x_i)g(x_i)$$

Но в нем не требуется на каждом шаге вычислять обратный гессиан  $G^{-1}(x_i)$ , так как направление поиска на шаге  $i$  является направлением  $-H_i g(x_i)$ , где  $H_i$  - положительно определенная симметрическая матрица, которая обновляется на каждом шаге. В пределе матрица  $H$  становится равной обратному гессиану.

Начнем поиск из начальной точки  $X_0$ , взяв в качестве начальной матрицу  $H_0$

1. На шаге  $i$  имеются точка  $x_i$  и положительно определенная симметрическая матрица  $H_i$ .

2. В качестве направления поиска взять направление

$$d_i = -H_i g_i$$

3. Чтобы найти функцию  $\lambda_i$ , минимизирующую функцию  $f(x_i + \lambda d_i)$ , произвести одномерный поиск вдоль прямой  $x_i + \lambda d_i$ .

4. Положить

$$v_i = \lambda_i d_i$$

5. Положить

$$X_{i+1} = X_i + V_i$$

6. Найти  $f(x_{i+1})$  и  $g_{i+1}$ . Завершить процедуру, если величины  $|g_{i+1}|$  достаточно малы. В противном случае продолжить.

7. Положить

$$u_i = g_{i+1} - g_i$$

8. Обновить матрицу  $H$  следующим образом:

$$H_{i+1} = H_i + A_i + B_i,$$

где

$$A_i = v_i v_i^T / (v_i^T u_i),$$

$$B_i = -H_i u_i u_i^T H_i / (u_i^T H_i u_i)$$

9. Увеличить на единицу и вернуться на шаг 2.

Этим завершается теоретическое изложение метода ДФП, который использует как идеи метода Ньютона — Рафсона, так и свойство сопряженных направлений, и при применении для минимизации квадратичной функции  $n$  переменных он сходится не более чем за  $n$  итераций. Это весьма мощная оптимизационная процедура, очень эффективная при оптимизации большинства функций независимо от того, квадратичны они или нет.

### 5.7.3 Метод Флетчера – Ривса

Метод флетчера — Ривса основан на том, что для квадратичной функции  $n$  переменных  $n$  одномерных поисков вдоль взаимно сопряженных направлений позволяют найти минимум.

Направления поиска в методе Флетчера — Ривса являются взаимно сопряженными и в данном методе минимум квадратичной функции  $n$  переменных можно найти не более чем за  $n$  шагов. Это означает, что одномерный поиск производится с нужной точностью и устраняются любые ошибки округления, которые могут возникнуть.

Метод применим и к неквадратичным функциям, так как если поиск осуществляется вблизи минимума, то можно надеяться на достижение квадратичной сходимости, когда имеет место квадратичная аппроксимация. Флетчер и Ривс полагают, что в этой ситуации каждое  $n$ -е направление поиска должно быть направлением наискорейшего спуска и при построении сопряженных направлений должен быть произведен рестарт.

## 5.8 Штрафные функции

Основная идея штрафной функции состоит в преобразовании задачи минимизации функции

$$z = f(x)$$

с соответствующими ограничениями, наложенными на  $x$ , в задачу поиска минимума без ограничений функции

$$z = f(x) + P(x)$$

Функция  $P(x)$  является штрафной. Необходимо, чтобы при нарушении ограничений она "штрафовала" функцию  $z$ , т. е. увеличивала ее значение. В этом случае минимум  $Z$  будет находиться внутри области ограничений. Функция  $P(x)$ , удовлетворяющая этому условию, может быть не единственной.

Задачу минимизации можно сформулировать следующим образом:

минимизировать функцию  $z = f(x)$  при ограничениях  $c_j(x) > 0$ ,  $j = 1, 2, \dots, m$ .

Функцию  $P(x)$  удобно записать следующим образом:

$$P(x) = r \sum_{j=1}^m \frac{1}{c_j(x)},$$

где  $r$  – положительная величина. Тогда

$$z = \varphi(x, r) = f(x) + r \sum_{j=1}^m \frac{1}{c_j(x)}.$$

Если  $x$  принимает допустимые значения, т. е. значения, для которых  $c_j(x) \geq 0$ , то  $z$  принимает значения, которые больше соответствующих значений  $f(x)$ , и разность можно уменьшить за счет того, что  $r$  может быть очень малой величиной. Но если  $x$  принимает значения, которые хотя и являются допустимыми, но близки к границе области ограничений, и по крайней мере одна из функций  $c_j(x)$  близка к нулю, тогда значения функции  $P(x)$ , и, следовательно, значения функции  $z$  станут очень велики. Таким образом, влияние функции  $P(x)$  состоит в создании «гребня с крутыми краями» вдоль каждой границы области ограничений. Значит, если поиск начать из допустимой точки и осуществляется поиск минимума функции  $\varphi(x, r)$  без ограничений, то минимум будет достигаться внутри допустимой области для задачи с ограничениями. Полагая  $r$  достаточно малой величиной, для того чтобы влияние  $P(x)$  было малым в точке минимума, мы можем сделать точку минимума функции  $\varphi(x, r)$  без ограничений совпадающей с точкой минимума функции  $f(x)$  с ограничениями.

Очевидно, что если функция  $f(x)$  выпукла, а функция  $c_j(x)$  при  $j = 1, \dots, n$  вогнута, то функция  $f(x)$  при наличии ограничений имеет единственный минимум.

## 5.9 Метод SUMT Фиакко и Маккормика

Метод SUMT (sequential unconstrained minimisation technique) был впервые предложен Кэрролом в 1961 году. Его идеи были исследованы Фиакко и Маккормиком, которые не только рассмотрели теоретические вопросы и сходимость метода, но создали практическую систему для его реализации.

Редко можно будет использовать данный метод, поскольку далеко не всегда можно найти оптимальную точку для функций  $\varphi(x,r)$  в виде функции  $x^*(r)$ , предел которой при  $r \rightarrow 0$  можно исследовать.

Поэтому для того, чтобы можно было применить настоящий метод на практике, необходимо построить вычислительный метод, использующий теоретическое свойство сходимости, указанное в предыдущем разделе. Теоретически здесь не возникает трудностей.

Предполагается, что в начале процедуры имеется допустимая точка. Важно, чтобы в процессе последующих вычислений получаемые точки принадлежали допустимой области.

Выбор начального значения  $r$  может оказаться важным с точки зрения сокращения числа итераций при минимизации функции  $\varphi(x,r)$ . Если сначала  $r$  выбрано очень малым, для того чтобы функция  $\varphi(x,r)$  мало отличалась от функции  $f(x)$ , то метод будет сходиться очень быстро. Однако такой выбор может привести к серьезным осложнениям при вычислениях.

## 5.10 Неопределенность целей. Методы построения обобщенных критериев. Оптимизация по Парето

Исследования тех операций, которые мы рассматривали в двух предыдущих параграфах, сводились к четко поставленным математическим задачам оптимизации, т.е. эти задачи не содержали неопределенностей. Но такие задачи являются скорее исключением, чем правилом.

Адекватное реальности описание проблемы практически всегда содержит различного рода неопределенности, отражающие то естественное положение, в котором находится исследователь: любое его знание относительно и неточно. В исследовании операций различают три вида неопределенностей: неопределенность наших знаний об окружающей обстановке (неопределенность природы); неопределенность целей; неопределенность действий реального противника или партнера, например, нетождественность целей партнеров в игре. Для инженеров - механиков представляет интерес второй вид неопределенности.

Так, при рассмотрении вопроса оптимизации режимов резания

в качестве целевой функции можно принимать либо максимум производительности, либо минимум себестоимости обработки. Однако производителя, скорее всего, будет интересовать и то, и другое одновременно, несмотря на противоречивость этих целей друг другу (в самом деле, рост производительности обязательно приведет к повышенному расходу инструмента, электроэнергии, износу оборудования и др., т.е. к росту себестоимости продукции). Этот факт нашел свое выражение в распространенной фразе: добиться максимума производства с минимумом затрат. Фраза не имеет строго научного смысла, т.к. минимум затрат - нуль, а с нулевыми затратами произвести какую либо полезную работу нельзя. Но эта, на первый взгляд бессмысленная, фраза отражает тенденцию, интересы оперирующей стороны. Рассмотренная ситуация типична: она показывает, что, даже зная цели оперирующей стороны, исследователь операции еще не может решить оптимизационную задачу. Для того, чтобы свести задачу исследования операции к стандартной задаче оптимизации, необходимо сформулировать еще дополнительные гипотезы, не вытекающие из постановки задачи.

Рассмотрим некоторые способы преодоления неопределенности целей. Итак, перед исследователем стоит задача выбора способа действия (вектора  $X$ ) обеспечивающего максимальное значение функциям

$$f_1(x), f_2(x), \dots, f_n(x)$$

одновременно.

Линейная свертка. Вместо  $n$  частных критериев  $f_i$  предлагается рассматривать один критерий вида

$$F(x) = \sum_{i=1}^n c_i f_i(x), \quad (5.22)$$

где  $c_i$  - некоторые положительные числа, тем или иным способом нормированные, например,  $\sum_{i=1}^n c_i = 1$ .

Такой способ свертки вводит отношение эквивалентности различных критериев (целевых функций), т.к. величины  $c_i$  показывают, насколько изменяется целевая функция  $F(x)$  при изменении критерия на единицу:

$$c_i = \frac{\partial F}{\partial f_i}.$$

Коэффициенты  $c_i$  - результат экспертизы; они отражают представление оперирующей стороны о содержании компромисса, т.е.

о ранжировании целей, которое вместе с назначением весовых коэффициентов и является той дополнительной гипотезой, которая позволяет свести задачу со многими критериями к задаче с единственным критерием, определяемым формулой (5.22).

Компромиссы Парето. Смысл сведения частных критериев к одному общему очевиден: одну задачу заменяют другой, причем в правомочности такой замены состоит содержание новых гипотез.

Но к анализу многокритериальных задач можно подойти и с других позиций: попытаться сократить множество исходных вариантов, т.е., исключить из неформального анализа те варианты решений, которые заведомо будут плохи. Один из подобных путей предложил в 1904 г. итальянский математик В. Парето.

Предположим, что мы сделали некоторый выбор. Обозначим его через  $x_1$  и предположим, что существует некоторый другой выбор  $x_2$  такой, что для всех критериев  $f_i(x)$  имеют место неравенства

$$f_i(x_2) \geq f_i(x_1), \quad i = \overline{1, n}, \quad (5.23)$$

причем хотя бы одно из неравенств — строгое.

Очевидно, что выбор  $x_2$  предпочтительнее  $x_1$ . Поэтому все векторы  $x_1$ , удовлетворяющие (5.23), следует сразу исключить из рассмотрения. Имеет смысл заниматься сопоставлением, подвергать неформальному анализу только векторы  $x_1$ , для которых не существует  $x_2$  такого, что для всех критериев удовлетворяются неравенства. Множество всех таких значений  $x_1$  называют множеством Парето, а вектор  $x_1$  называют не улучшаемым вектором результатов (вектором Парето), если из  $f_i(x_2) \geq f_i(x_1)$  для любого  $i$  следует  $f_i(x_2) = f_i(x_1)$ .

Предположим, что цели субъекта определяются двумя однозначными функциями:

$$f_1(x) \rightarrow \max; \quad f_2(x) \rightarrow \max.$$

Тогда каждому допустимому значению переменной  $x$  отвечает одна точка на плоскости  $(f_1, f_2)$ , рис.5.3, и равенства

$$f_1 = f_1(x), \quad f_2 = f_2(x)$$

определяют параметрическое задание некоторой кривой  $abcd$  в этой плоскости. Но к множеству Парето можно отнести далеко не всю

кривую. Так, участок  $bc$ , очевидно, не принадлежит множеству Парето, поскольку вместе с ростом  $f_1$  происходит и рост  $f_2$ .

Таким образом, на этом участке изменению переменной  $x$  отвечает одновременное увеличение обеих целевых функций, и, следовательно, такие варианты решений должны быть сразу исключены из дальнейшего рассмотрения.

Из тех же соображений должен быть исключен участок  $a'b$ , поскольку для каждой его точки  $e$  найдется точка  $g$ , принадлежащая участку  $cd$ , в которой значения обеих функций  $f_1$  и  $f_2$  больше, чем в точке  $e$ . Значит, претендовать на принадлежность к множеству Парето могут только участки  $aa'$  и  $cd$ , причем точка  $a'$  также должна быть исключена.

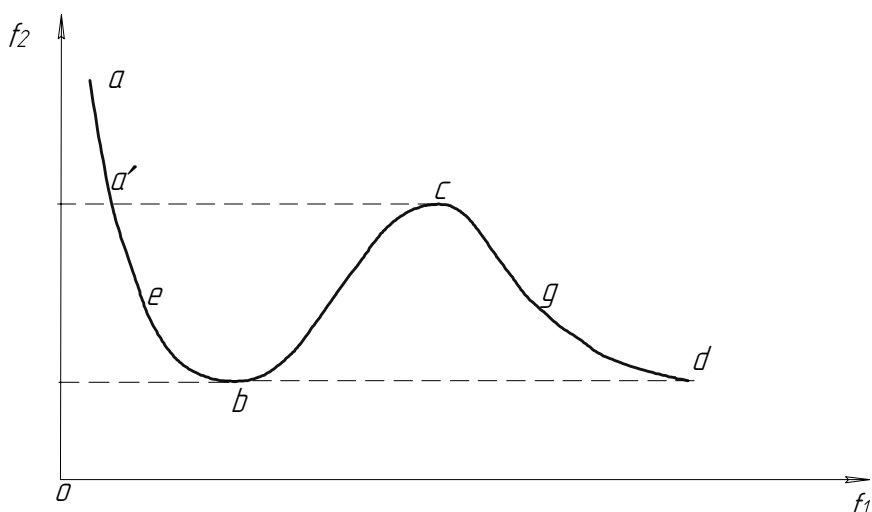


Рис. 5.2 - Компромиссы Парето.

Таким образом, принцип Парето заключается в том, что выбирать в качестве решения следует только тот вектор  $x$ , который принадлежит множеству Парето. Принцип Парето не выделяет единственного решения, он только сужает множество альтернатив. Окончательный выбор остается за лицом, принимающим решение. Но исследователь, построив множество Парето, конечно, облегчает процедуру выбора решения.

Принцип Парето играет очень важную роль в автоматизации проектирования. Очень часто проектировщик оказывается перед необходимостью искать компромисс. Одним из путей отыскания этого компромисса является построение множества Парето, изучение которого дает большую информацию. Лицо, принимающее решение, видит, в частности, сколько "стоит" увеличение одного из показателей, как оно сказывается на остальных показателях, значения которых непременно ухудшаются. Это множество оказывается, как правило, весьма сложной природы. Его анализ интуитивными методами вряд ли возможен. Но, помимо частных критериев  $f_i(x)$ , достаточно часто в

распоряжении проектировщика есть еще некоторый общий критерий. Введение общего критерия  $F(x)$  и максимизация его значений на множестве Парето также является некоторой гипотезой, поскольку из совокупности критериев  $f_1, \dots, f_n, F$  один из критериев мы специальным образом выделяем.

В заключении следует указать, что принцип Парето относится к методам последовательного сжатия множества альтернатив, которые занимают одно из центральных мест в системном анализе.



## ПОСЛЕСЛОВИЕ

Создание сложных технических систем, конструкций, устройств, технологий, с которыми приходится иметь дело современным специалистам, потребовало организации проектирования, совмещенного с исследованиями, которое бы носило нетрадиционный характер. Сам процесс проектирования стал близок к процессу исследования, поэтому в данной работе рассмотрены основные направления инженерных исследований. Предложенный подход базируется на понятии «системный анализ», предполагающем рассмотрение инженерных исследований как целостной системы. Рассмотрены этапы исследований, которые совпадают по смыслу с основными этапами принятия решения в системном анализе:

1) Построение качественной модели рассматриваемой проблемы. Предложены подходы к построению моделей. Показано, как определить параметры моделей.

2) Построение математической модели рассматриваемой проблемы. Приведены математические модели некоторых механических систем. Показано, как необходимо выбирать функцию цели.

3) Исследование влияния переменных на значение целевой функции. Показано, каким образом необходимо проводить аналитическое решение модели, а если это невозможно – то численными методами на ЭВМ. Приведены способы оптимизации параметров моделей.

4) Проверка адекватности модели. Приведены основные подходы к планированию эксперимента, обработке экспериментального материала, а также собственно проверке адекватности модели реальному объекту. При этом необходимо учитывать современный уровень и возможности существующих технологических средств проведения эксперимента – первичных датчиков, контрольно-измерительных приборов, исполнительных устройств и др.

Надеемся на то, что работа, с которой Вы познакомились, окажется полезной в Вашей работе.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Советов Б.Я., Яковлев С.А. Моделирование систем. - М.:Высшая школа, 1985.- 271 с.
2. Тимошенко С.П. Колебания в инженерном деле. - М.: Наука, 1967.- 444 с.
3. Седов Л.Н. Методы подобия и размерностей в механике. - М.: Наука, 1981.- 448 с.
4. Шенк Х. Теория инженерного эксперимента. - М.: Мир, 1972.- 381 с.
5. Адлер Ю.П., Марков Е.В., Грановский Ю.В. Планирование эксперимента при поиске оптимальных решений. - М.: Наука, 1976.- 279 с.
6. Джонсон Н., Лион Ф. Статистика и планирование эксперимента в технике. Методы планирования эксперимента. - М.: Мир, 1981.- 520 с.
7. Бендат Дж., Пирсол А. Применения корреляционного и спектрального анализа. - М.: Мир, 1983.- 312 с.
8. Смирнов Н.В., Дудин-Барковский И.В. Курс теории вероятностей и математической статистики для технических приложений: учебное пособие для вузов. - М.: Наука, 1969.- 505 с.
9. Вентцель Е.С., Овчаров Л.А. Прикладные задачи теории вероятности. - М.: Радио и связь, 1983.- 416 с.
10. Пальчевський Б.О. Дослідження технологічних систем (моделювання, проектування, оптимізація): Навчальний посібник.- Львів: Світ, 2001.- 232с.
11. Бондарь А.Г., Статюха Г.А. Планирование эксперимента в химической технологии.- К.: Вища школа, 1976.-184с.
12. Баранов Е.Г. Основы научных исследований.-К.: Вища школа, 1984.-176с.
13. Проектирование датчиков для измерения механических величин/ Под ред. А.П.Осадчего. - М.: Машиностроение, 1979.- 480 с.
14. Тензометрия в машиностроении / Под общ. ред. Р.А.Маркова.- М.: Машиностроение, 1975.- 288 с.
15. Тензометрическое устройство для измерения усилий в зубчатом зацеплении: Проспект ВДНХ УССР/ Н.Г.Бойко, И.А.Горобец, В.Г.Нечепав и др. – Донецк, 1988.-2с.
16. А.с. № 1399462 СССР, МКИ<sup>4</sup> E21C 29/02. Устройство для измерения усилий зацепления движителей очистных комбайнов/ Н.Г.Бойко, И.А.Горобец, В.Г.Нечепав и др. (СССР).- Опубл.30.05.88. Бюл.№ 20.-С.136.
17. Gorobez I, Golubov N. Adaptronsystem in der Werkzeugmaschine für die spanabhebende Formung – Entwicklungsmethoden und

Entwicklungsprozesse im Maschinenbau. 5 Magdeburger Maschinenbau-Tage.- Berlin, Logos-Verl.,2001,-S.45-51.

18Gorobez I., Navka I., Lapajeva I., Schaban K. Die Parameter der Adaptronsysteme der Drehmaschinen – Modern Technologies, Quality and Restructuring International Conference N.C.M.R – Bulletin of the Politechnic institute of Jassy, Iassy, Romania 23-25 of May 2002 , Vol. XLVIII - S.100-104.

19. Горобец И.А. , Шабан К. Параметры регулирования адаптронных систем токарных станков – Материалы Третьего Международного научно-практического семинара “Практика и перспективы развития институционального партнерства”. В 2-х кн. – Таганрог, ТРТУ, кн..2, 2002, №2 – с.47-52.

20. Вицко А.В., Горобец И.А., Разработка конструкции силового датчика адаптивной системы управления токарного станка с ЧПУ – Инженер. Студенческий научно-технический журнал.- Донецк: ДонНТУ, 2003. №4.- С.59-63.

21. А.С. СССР №1604989 Горобец И.А., Бойко Н.Г. Устройство для исследования параметров бесцепной системы подачи угольного комбайна. - Оpubл. БИ №41, 1990.- с.36.

22. Molitor M. Messtechnik- Magdeburg :TU. – 1999.-120s.

23. Теория инженерного эксперимента: Учебное пособие/ Г.М. Тимошенко, П.Ф.Зима.- Донецк: ДПИ, 1984.- 60с.

24. Банди Б. Методы оптимизации. Вводн. курс: Пер. с англ..- М.: Радио и связь, 1988. – 128 с.