БИБЛИОТЕЧКА ФИЗИКО-МАТЕМАТИЧЕСКОЙ ШКОЛЫ

МАТЕМАТИКА

Г. И. ДРИНФЕЛЬД

ИНТЕРПОЛИРОВАНИЕ И СПОСОБ НАИМЕНЬШИХ КВАДРАТОВ

БИБЛИОТЕЧКА
ФИЗИКО-МАТЕМАТИЧЕСКОЙ
ШКОЛЫ

МАТЕМАТИКА

Г. И. ДРИНФЕЛЬД

ИНТЕРПОЛИРОВАНИЕ И СПОСОБ НАИМЕНЬШИХ КВАДРАТОВ

КИЕВ
ГОЛОВНОЕ ИЗДАТЕЛЬСТВО
ИЗДАТЕЛЬСКОГО ОБЪЕДИНЕНИЯ
«ВИЩА ШКОЛА»
1984

22.19
Д74

УДК 512.5

Интерполирование и способ наименьших квадратов. Дринфельд Г. И.—К.: Вища шк. Головное изд-во, 1984.—103 с.—(Б-чка физ.-мат. школы. Математика).

В книге изложены начальные сведения об интерполировании: линейном, полиномиальном, тригонометрическом. Рассмотрен метод наименьших квадратов и применение его к решению задач полиномиального интерполирования. Описаны операции над векторами, свойства операций, процесс ортогонализации векторов. Все понятия и приемы вычисления иллюстрируются примерами, даны задачи для самостоятельного решения.

Рассчитана на учащихся физико-математических и средних общеобразовательных школ. Она может быть использована учителями математики при проведении факультативных занятий и внеклассной работы.

Ил. 19.

Рецензенты: кандидат физико-математических наук *П. М. Моклячук* (Киевский государственный университет), учитель-методист *Р. П. Ушаков* (средняя школа № 173 г. Киева)

Редакция литературы по математике и физике
Зав. редакцией *Е. Л. Корженевич*

Д $\frac{1702050000-184}{М211(04)-84}$ 138—84

ГЛАВА 1

ВЕКТОРЫ, ОРТОНОРМИРОВАНИЕ

§ 1. Неравенство Коши — Буняковского[1]. Исследуем функцию

$$y = ax^2 + 2bx + c, \ a > 0,$$

и построим ее график.

Поскольку $y = \frac{1}{a}((ax + b)^2 + (ac - b^2))$, то:

1) при $ac - b^2 \geqslant 0$ $y \geqslant 0$ для любого действительного значения x, а равенство $y = 0$ справедливо только тогда, когда $ac - b^2 = 0$, $x = -\frac{b}{a}$;

2) если $ac - b^2 < 0$, то $y = 0$ при $x = \frac{1}{a}(-b \pm \pm \sqrt{b^2 - ac})$, $y < 0$ при условии $\frac{-b - \sqrt{ac - b^2}}{a} < x < < \frac{-b + \sqrt{ac - b^2}}{a}$ и $y > 0$ при всех остальных значениях x;

3) всегда при $x = -\frac{b}{a}$ $y = \frac{1}{a}(ac - b^2)$;

4) при $|x| \to \infty$ $y \to \infty$.

Графики этой функции при $a > 0$ изображены на рис. 1.

Сформулируем выводы, которые будут использованы в дальнейшем.

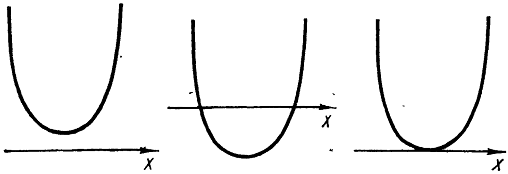

Рис. 1

[1] Буняковский В. Я. (1804—1889) — русский математик, академик. Коши О. Л. (1789—1857) — французский математик, академик.

1. Для действительных значений аргумента квадратный трехчлен $ax^2+2bx+c$ с положительным старшим коэффициентом не принимает отрицательных значений тогда и только тогда, когда $ac-b^2 \geqslant 0$.

2. Квадратный трехчлен $ax^2+2bx+c$ с положительным старшим коэффициентом при $-\infty < x < +\infty$ не имеет наибольшего значения, а наименьшего значения достигает при $x=-\frac{b}{a}$.

Теорема. *Для любых действительных* a_i, b_i $(i=1, 2, \ldots)$ *справедливо неравенство Коши — Буняковского*

$$\left|\sum_{i=1}^{n} a_i b_i\right| \leqslant \sqrt{\sum_{i=1}^{n} a_i^2}\sqrt{\sum_{i=1}^{n} b_i^2}. \qquad (1)$$

Это одно из важнейших неравенств в математике. Для его доказательства рассмотрим функцию

$$y=\sum_{i=1}^{n}(xa_i+b_i)^2=x^2\sum_{i=1}^{n}a_i^2+2x\sum_{i=1}^{n}a_ib_i+\sum_{i=1}^{n}b_i^2$$

при $-\infty < x < +\infty$. Поскольку $y \geqslant 0$, то, используя вывод 1, имеем

$$\left(\sum_{i=1}^{n}a_i^2\right)\left(\sum_{i=1}^{n}b_i^2\right)-\left(\sum_{i=1}^{n}a_ib_i\right)^2 \geqslant 0,$$

что равносильно (1).

Приведем еще одно элементарное доказательство неравенства (1). С этой целью установим тождество

$$\left(\sum_{i=1}^{n}a_i^2\right)\left(\sum_{i=1}^{n}b_i^2\right)-\left(\sum_{i=1}^{n}a_ib_i\right)^2 \equiv \sum_{r \neq s}(a_rb_s-a_sb_r)^2, \qquad (2)$$

называемое *тождеством Лагранжа*[1]. Имеем

$$\left(\sum_{i=1}^{n}a_i^2\right)\left(\sum_{i=1}^{n}b_i^2\right)-\left(\sum_{i=1}^{n}a_ib_i\right)^2=\sum_{i=1}^{n}a_i^2b_i^2+\sum_{r\neq s}a_r^2b_s^2-$$

$$-\sum_{i=1}^{n}a_i^2b_i^2-2\sum_{r\neq s}{}' a_rb_ra_sb_s=\sum_{r\neq s}(a_rb_s-a_sb_r)^2.$$

Здесь $\sum'$ означает $\sum\limits_{r=1}^{n}a_rb_r\sum\limits_{s=r+1}^{n}a_sb_s$. Если приведенное вычисление окажется затруднительным, то рекомендуем выполнить его для $n=3$ без применения знака Σ.

---

[1] Лагранж Ж. Л. (1736—1813) — французский математик.

Тождество (2) справедливо для действительных и недействительных a_i, b_i. Если a_i, b_i действительны, то из (2) следует: 1) неравенство Коши — Буняковского; 2) для того чтобы имело место равенство $\left(\sum_{i=1}^{n} a_i b_i\right)^2 = \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right)$, необходимо и достаточно выполнение условий $a_r b_s - a_s b_r = 0$ ($r, s = 1, 2, \ldots, n$). Достаточность следует непосредственно из (1).

Замечание. С помощью элементов векторной алгебры легко видеть, что при $n = 3$ неравенство (1) следует из неравенства $|\cos\alpha| \leqslant 1$. Действительно, будем считать (a_1, a_2, a_3), (b_1, b_2, b_3) соответственно проекциями векторов $\vec{u}$ и $\vec{v}$ на оси прямоугольных координат. Тогда

$$\sum_{i=1}^{3} a_i b_i = \vec{u}\cdot\vec{v} = |\vec{u}|\,|\vec{v}|\cos(\widehat{\vec{u}, \vec{v}}), \quad \sum_{i=1}^{3} a_i^2 = |\vec{u}|^2, \quad \sum_{i=1}^{3} b_i^2 = |\vec{v}|^2.$$

Следовательно, неравенство Коши — Буняковского справедливо, оно означает, что *модуль скалярного произведения двух векторов не больше произведения длин этих векторов.*

§ 2. Арифметическое n-мерное пространство.

Определение. *Арифметическим n-мерным пространством называется множество всех упорядоченных систем n действительных чисел*

$$(a_1, a_2, \ldots, a_n). \tag{1}$$

Упорядоченность означает, что, меняя местами даже два не равных между собой числа в (1), получаем другую систему. Например, если $a_1 \neq a_2$, то $(a_1, a_2, \ldots, a_n)$ и $(a_2, a_1, \ldots, a_n)$ различны. Иными словами, равенство

$$(a_1, a_2, \ldots, a_n) = (b_1, b_2, \ldots, b_n)$$

возможно тогда и только тогда, когда $a_1 = b_1$, $a_2 = b_2, \ldots, a_n = b_n$. Каждая упорядоченная система (1) называется *точкой пространства*, а числа $a_1, a_2, \ldots, a_n$ — *координатами* этой точки.

Например, при $n = 2$ имеем двойки чисел (a_1, a_2). Точку (a_1, a_2) можно рассматривать как точку плоскости, декартовы координаты которой a_1, a_2 (рис. 2). Таким образом, двумерное арифметическое пространство можно рассматривать как плоскость. Точно так же трехмерное арифметическое пространство можно рассматривать как обычное геометрическое пространство (рис. 3).

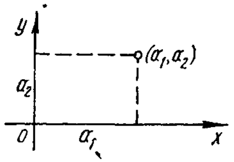

Рис. 2

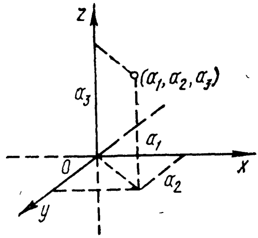

Рис. 3

§ 3. Евклидово *n*-мерное пространство.

Определение. *Арифметическое n-мерное пространство называется* м е т р и з о в а н н ы м, *если каждым двум его точкам* $A(a_1, a_2, \ldots, a_n)$ *и* $B(b_1, b_2, \ldots, b_n)$ *ставится в соответствие число* $\rho(A, B)$, *обладающее свойствами:*

1) $\rho(A, B) \geqslant 0$, $\rho(A, B) = 0$ *тогда и только тогда, когда* $A = B$;

2) $\rho(A, B) = \rho(B, A)$ (*симметричность*);

3) *для любых трех точек* A, B, C *справедливо неравенство*

$$\rho(A, C) \leqslant \rho(A, B) + \rho(B, C). \qquad (1)$$

Указанные свойства называются *аксиомами метрики*, а $\rho(A, B)$ — *расстоянием* между точками A и B.

В евклидовой плоскости длина отрезка — это расстояние между его концами. Действительно, длина отрезка очевидно удовлетворяет аксиомам 1) и 2), а также удовлетворяет аксиоме 3), так как длина любой стороны треугольника не больше суммы длин двух других его сторон (рис. 4).

По аналогии, неравенство (1) всегда называют *неравенством треугольника*.

В евклидовой плоскости длину отрезка, соединяющего точки $A(a_1, a_2)$ и $B(b_1, b_2)$, где (a_1, a_2) и (b_1, b_2) — пря-

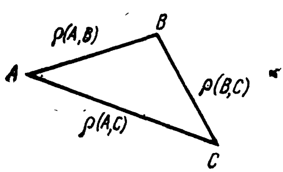

Рис. 4

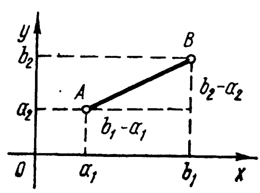

Рис. 5

моугольные декартовые координаты точек (рис. 5), можно вычислить по формуле

$$\rho(A,\ B) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2},$$

а в трехмерном евклидовом пространстве — по формуле

$$\rho(A,\ B) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + (b_3 - a_3)^2}.$$

В n-мерном арифметическом пространстве положим по определению

$$\rho(A,\ B) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + \cdots + (b_n - a_n)^2} = \sqrt{\sum_{i=1}^{n} (b_i - a_i)^2} \qquad (2)$$

и покажем, что удовлетворяются все три аксиомы метрики.

Выполнимость аксиом 1) и 2) очевидна. Проверим аксиому 3). Пусть $A(a_1, \ldots, a_n)$, $B(b_1, \ldots, b_n)$ и $C(c_1, \ldots, c_n)$ — три произвольные точки n-мерного пространства. Тогда

$$\rho^2(A,\ B) = \sum_{i=1}^{n} (b_i - a_i)^2, \quad \rho^2(A,\ C) = \sum_{i=1}^{n} (c_i - a_i)^2,$$

$$\rho^2(B,\ C) = \sum_{i=1}^{n} (c_i - b_i)^2.$$

Обозначим

$$b_i - a_i = \alpha_i,\ c_i - a_i = \gamma_i,\ c_i - b_i = \beta_i.$$

Поскольку $\gamma_i = \alpha_i + \beta_i$, то надо доказать, что

$$\sum_{i=1}^{n} \gamma_i^2 = \sum_{i=1}^{n} (\alpha_i + \beta_i)^2 \leqslant \left(\sqrt{\sum_{i=1}^{n} \alpha_i^2} + \sqrt{\sum_{i=1}^{n} \beta_i^2}\right)^2.$$

Имеем

$$\sum_{i=1}^{n} (\alpha_i + \beta_i)^2 = \sum_{i=1}^{n} \alpha_i^2 + \sum_{i=1}^{n} \beta_i^2 + 2\sum_{i=1}^{n} \alpha_i\beta_i.$$

На основании неравенства Коши — Буняковского

$$\sum_{i=1}^{n} \alpha_i\beta_i \leqslant \sqrt{\sum_{i=1}^{n} \alpha_i^2}\sqrt{\sum_{i=1}^{n} \beta_i^2}.$$

Поэтому

$$\sum_{i=1}^{n} \gamma_i^2 \leqslant \sum_{i=1}^{n} \alpha_i^2 + 2\sqrt{\sum_{i=1}^{n} \alpha_i^2}\sqrt{\sum_{i=1}^{n} \beta_i^2} + \sum_{i=1}^{n} \beta_i^2,$$

откуда

$$\sum_{i=1}^{n} \gamma_i^2 \leqslant \left(\sqrt{\sum_{i=1}^{n} \alpha_i^2} + \sqrt{\sum_{i=1}^{n} \beta_i^2}\right)^2.$$

Определение. *Арифметическое n-мерное пространство, метризованное по формуле* (2), *называется n-мерным евклидовым пространством.*

§ 4. Векторы в n-мерном арифметическом пространстве. Сложение векторов, умножение вектора на число. Упорядоченную систему n чисел $(a_1, a_2, \ldots, a_n)$ будем называть *вектором* и обозначать $\vec{a}(a_1, a_2, \ldots, a_n)$, принимая, по определению, что равенство

$$\vec{a}(a_1, a_2, \ldots, a_n) = \vec{b}(b_1, b_2, \ldots, b_n)$$

эквивалентно системе равенств

$$a_1 = b_1, \ a_2 = b_2, \ \ldots, \ a_n = b_n;$$

суммой $\vec{a} + \vec{b}$ векторов $\vec{a}(a_1, a_2, \ldots, a_n)$ и $\vec{b}(b_1, b_2, \ldots, b_n)$ является вектор $\vec{c}(a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$.

Произведением вектора $\vec{a}(a_1, a_2, \ldots, a_n)$ *на число* (скаляр) λ называется вектор $\lambda\vec{a}$ $(\lambda a_1, \lambda a_2, \ldots, \lambda a_n)$.

Вектор $\vec{0}(0, 0, \ldots, 0)$ будем называть *нулевым*.

Из приведенных определений и свойств действительных чисел следуют такие основные свойства сложения векторов и умножения вектора на число:

1) $\vec{a} + \vec{b} = \vec{b} + \vec{a}$ (переместительный закон);

2) $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$ (сочетательный закон);

3) $\vec{a} + \vec{0} = \vec{a}$ (закон поглощения нулевого вектора);

4) $\lambda(\mu\vec{a}) = (\lambda\mu)\vec{a}$ (сочетательный закон);

5) $(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$ (первый распределительный закон);

6) $\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$ (второй распределительный закон);

7) $0\vec{a} = \lambda\vec{0} = \vec{0}$ (закон поглощения нуля и нулевого вектора).

Заметим, что из указанных законов действий с векторами сумму любого числа векторов можно записывать без скобок, произвольно переставлять и группировать слагаемые. Все свойства легко проверить. Например,

$$\lambda(\vec{a}+\vec{b}) = \lambda\,\overrightarrow{(a_1+b_1,\ a_2+b_2,\ \ldots,\ a_n+b_n)} =$$

$$= \left(\overrightarrow{\lambda(a_1+b_1),\ \lambda(a_2+b_2),\ \ldots,\ \lambda(a_n+b_n)}\right) =$$

$$= \left(\overrightarrow{\lambda a_1+\lambda b_1,\ \lambda a_2+\lambda b_2,\ \ldots,\ \lambda a_n+\lambda b_n}\right) =$$

$$= \overrightarrow{(\lambda a_1,\ \lambda a_2,\ \ldots,\ \lambda a_n)} + \overrightarrow{(\lambda b_1,\ \lambda b_2,\ \ldots,\ \lambda b_n)} =$$

$$= \lambda\vec{a} + \lambda\vec{b}.$$

Вектор $(-1)\,\overrightarrow{(a_1,\ a_2,\ \ldots,\ a_n)} = \overrightarrow{(-a_1, -a_2, \ldots, -a_n)}$ будем называть противоположным вектору $\vec{a}(a_1,\ a_2,\ \ldots,\ a_n)$ и обозначать $-\vec{a}$.

По определению имеем

$$\vec{a} + (-\vec{a}) = \overrightarrow{(a_1,\ a_2,\ \ldots,\ a_n)} + \overrightarrow{(-a_1, -a_2, \ldots, -a_n)} =$$

$$= \overrightarrow{(0,\ \ldots,\ 0)} = \vec{0}.$$

Полагая $\vec{a} + (-\vec{b}) = \vec{a} - \vec{b}$, будем иметь

$$\vec{a} - \vec{b} = \vec{c}(a_1 - b_1,\ \ldots,\ a_n - b_n).$$

§ 5. Длина вектора (в евклидовом пространстве). Угол между векторами. Точки $O(0,\ 0,\ \ldots,\ 0)$ и $A(a_1,\ a_2,\ \ldots,\ a_n)$ будем называть соответственно *началом* и *концом вектора* $\overrightarrow{OA} = \vec{a}(a_1,\ a_2,\ \ldots,\ a_n)$.

Естественно поэтому следующее определение: *длиной вектора* $\vec{a}(a_1,\ a_2,\ \ldots,\ a_n)$ *называется число*

$$|\vec{a}| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2} = \sqrt{\sum_{i=1}^{n} a_i^2}\,.$$

Теорема 1. *Для любых двух векторов* $\vec{a}$ *и* $\vec{b}$ *справеоливы неравенства*

$$|\vec{a}| - |\vec{b}| \leqslant |\vec{a} \pm \vec{b}| \leqslant |\vec{a}| + |\vec{b}|.$$

Действительно, неравенство $|\vec{a} \pm \vec{b}| \leqslant |\vec{a}| + |\vec{b}|$ — это неравенство треугольника для трех точек $A(a_1, a_2, \ldots, a_n)$, $B(\pm b_1, \pm b_2, \ldots, \pm b_n)$, $C(a_1 \pm b_1, a_2 \pm b_2, \ldots, a_n \pm b_n)$, так как

$$\rho(A, B) = \sqrt{\sum_{i=1}^{n} (a_i \mp b_i)^2} = |\vec{a} \mp \vec{b}|,$$

$$\rho(A, C) = \sqrt{\sum_{i=1}^{n} b_i^2} = |\vec{b}|,$$

$$\rho(B, C) = \sqrt{\sum_{i=1}^{n} a_i^2} = |\vec{a}|.$$

Неравенство $|\vec{a}| - |\vec{b}| \leqslant |\vec{a} \pm \vec{b}|$ следует из неравенства $|\vec{a} \pm \vec{b}| \leqslant |\vec{a}| + |\vec{b}|$. Действительно,

$$|\vec{a}| = |(\vec{a} \pm \vec{b}) \mp \vec{b}| \leqslant |\vec{a} \pm \vec{b}| + |\vec{b}|,$$

$$|\vec{a}| - |\vec{b}| \leqslant |\vec{a} \pm \vec{b}|.$$

Определение. *Косинусом угла* $(\widehat{\vec{a}, \vec{b}})$ *между двумя ненулевыми векторами* $\vec{a}$ *и* $\vec{b}$ *называется число*

$$\cos(\widehat{\vec{a}, \vec{b}}) = \frac{\sum_{i=1}^{n} a_i b_i}{|\vec{a}|\,|\vec{b}|}. \tag{1}$$

Такое определение допустимо, поскольку по неравенству Коши — Буняковского

$$\left|\sum_{i=1}^{n} a_i b_i\right| \leqslant \sqrt{\sum_{i=1}^{n} a_i^2}\sqrt{\sum_{i=1}^{n} b_i^2} = |\vec{a}|\,|\vec{b}|,$$

и тогда

$$-1 \leqslant \frac{\sum_{i=1}^{n} a_i b_i}{|\vec{a}|\,|\vec{\ }|} \leqslant 1.$$

Определяя угол из равенства (1), будем полагать, что $0 \leqslant (\widehat{\vec{a}, \vec{b}}) \leqslant \pi$.

Теорема 2. *Угол* $(\widehat{\vec{a}, \vec{b}})$ *равен* 0 *или* π *тогда и только тогда, когда*

$$\vec{b} = \lambda \vec{a}. \tag{2}$$

Действительно, если справедливо равенство (2), то

$$|\vec{b}| = |\lambda \vec{a}(\lambda a_1, \lambda a_2, \ldots, \lambda a_n)| =$$

$$= \sqrt{\sum_{i=1}^{n} (\lambda a_i)^2} = |\lambda||\vec{a}|,$$

$$\sum_{i=1}^{n} a_i b_i = \sum_{i=1}^{n} a_i(\lambda a_i) = \lambda \sum_{i=1}^{n} a_i^2 = \lambda |\vec{a}|^2.$$

Следовательно,

$$\cos(\widehat{\vec{a}, \vec{b}}) = \frac{\lambda |\vec{a}|^2}{|\vec{a}| \cdot |\lambda||\vec{a}|} = \pm 1.$$

Если же $(\widehat{\vec{a}, \vec{b}}) = 0$ или $(\widehat{\vec{a}, \vec{b}}) = \pi$, то

$$\left|\sum_{i=1}^{n} a_i b_i\right| = |\vec{a}||\vec{b}| = \sqrt{\sum_{i=1}^{n} a_i^2} \sqrt{\sum_{i=1}^{n} b_i^2}.$$

Но знак равенства в неравенстве Коши — Буняковского можно получить лишь при условии, когда $b_i = \lambda a_i$ или, что то же самое, $\vec{b} = \lambda \vec{a}$.

Теорема 3. *Угол* $(\widehat{\vec{a}, \vec{b}})$ *равен* $\frac{\pi}{2}$ *тогда и только тогда, когда*

$$\sum_{i=1}^{n} a_i b_i = 0. \tag{3}$$

Эта теорема следует из определения (1).

Векторы $\vec{a}$ и $\vec{b}$ назовем *коллинеарными* («лежащими на одной линии»), если угол $(\widehat{\vec{a}, \vec{b}})$ равен 0 или π, и *ортогональными* («перпендикулярными»), если $(\widehat{\vec{a}, \vec{b}})$ равен $\frac{\pi}{2}$.

Таким образом, условие (2) является необходимым и достаточным для коллинеарности, а условие (3) — для ортогональности двух векторов $\vec{a}$ и $\vec{b}$.

§ 6. Скалярное произведение двух векторов.

Определение. *Скалярным произведением векторов* $\vec{a}(a_1, a_2, \ldots, a_n)$ *и* $\vec{b}(b_1, b_2, \ldots, b_n)$ *называется число* $a_1b_1 + a_2b_2 + \cdots + a_nb_n$.

Мы будем пользоваться обозначением $\langle\vec{a}, \vec{b}\rangle$ или $\vec{a}\cdot\vec{b}$, т. е.

$$\langle\vec{a}, \vec{b}\rangle = \vec{a}\cdot\vec{b} = \sum_{i=1}^{n} a_ib_i. \tag{1}$$

Основные свойства скалярного произведения:

1) $\langle\vec{a}, \vec{b}\rangle = \langle\vec{b}, \vec{a}\rangle$ (переместительный закон);

2) $\langle\vec{a}, (\vec{b}+\vec{c})\rangle = \langle\vec{a}, \vec{b}\rangle + \langle\vec{a}, \vec{c}\rangle$ (распределительный закон);

3) $\lambda\langle\vec{a}, \vec{b}\rangle = \langle\lambda\vec{a}, \vec{b}\rangle = \langle\vec{a}, \lambda\vec{b}\rangle$ (сочетательный закон);

4) скалярное произведение двух ненулевых векторов равно произведению их абсолютных величин на косинус угла между ними:

$$\langle\vec{a}, \vec{b}\rangle = |\vec{a}||\vec{b}|\cos(\widehat{\vec{a}, \vec{b}}). \tag{2}$$

Свойство 1) следует из (1); свойства 2) и 3) легко проверить. Например,

$$\lambda\langle\vec{a}, \vec{b}\rangle = \lambda\sum_{i=1}^{n} a_ib_i = \sum_{i=1}^{n}(\lambda a_i)b_i = \langle\lambda\vec{a}, \vec{b}\rangle,$$

Формула (2) следует из (1) и определения $\cos(\widehat{\vec{a}, \vec{b}})$.

Из 1) — 3) получаем

$$\langle(\lambda\vec{a} + \mu\vec{b}), (\alpha\vec{c} + \beta\vec{d})\rangle = \langle\lambda\vec{\alpha}, (\alpha\vec{c} + \beta\vec{d})\rangle + \langle\mu\vec{b}, (\alpha\vec{c} + \beta\vec{d})\rangle = \langle\lambda\vec{a}, \alpha\vec{c}\rangle + \langle\lambda\vec{a}, \beta\vec{d}\rangle + \langle\mu\vec{b}, \alpha\vec{c}\rangle + \langle\mu\vec{b} + \beta\vec{d}\rangle = \lambda\alpha\langle\vec{a}, \vec{c}\rangle + \lambda\beta\langle\vec{a}, \vec{d}\rangle + \mu\alpha\langle\vec{b}, \vec{c}\rangle + \mu\beta\langle\vec{b}, \vec{d}\rangle,$$

т. е. суммы векторов скалярно перемножаются как многочлены.

Теорему 3 предыдущего параграфа можно сформулировать следующим образом: *векторы* $\vec{a}$ *и* $\vec{b}$ *ортогональны тогда и только тогда, когда их скалярное произведение равно нулю.*

Замечание. Скалярное произведение $\langle \vec{a}, \vec{a} \rangle$ равно квадрату длины вектора $\vec{a}$, т. е. $\langle \vec{a}, \vec{a} \rangle = |\vec{a}|^2$.

§ 7. Линейная зависимость векторов.

Определение. *Ненулевые векторы* $\vec{a}_1(a_{11}, a_{12}, \ldots, a_{1n})$, $\vec{a}_2(a_{21}, a_{22}, \ldots, a_{2n})$, $\ldots$, $\vec{a}_r(a_{r1}, a_{r2}, \ldots, a_{rn})$ *называются* ***линейно зависимыми***, *если существуют числа* $\lambda_1, \lambda_2, \ldots, \lambda_r$, *среди которых хоть одно отлично от нуля* $\left(\sum_{i=1}^{r} \lambda_i^2 \neq 0\right)$, *такие, что*

$$\lambda_1\vec{a}_1 + \lambda_2\vec{a}_2 + \cdots + \lambda_r\vec{a}_r = \vec{0}, \qquad (1)$$

и ***линейно независимыми***, *если равенство* (1) *возможно только при условии*

$$\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0.$$

Пример 1. Покажем, что векторы

$$\vec{e}_1(1, 0, \ldots, 0), \vec{e}_2(0, 1, 0, \ldots, 0), \ldots, \vec{e}_n(0, 0, \ldots, 0, 1) \qquad (2)$$

— линейно независимы.

Действительно, если бы эти векторы были линейно зависимы, то было бы

$$\vec{r} = \lambda_1\vec{e}_1 + \lambda_2\vec{e}_2 + \cdots + \lambda_n\vec{e}_n = \vec{0}, \quad \sum_{i=1}^{n} \lambda_i^2 \neq 0, \qquad (3)$$

$$\vec{r}(\lambda_1, \lambda_2, \ldots, \lambda_n) = \vec{0}(0, 0, \ldots, 0),$$

$$\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$$

и мы пришли бы к противоречию с неравенством (3).

Совокупность векторов (2) называется *декартовым базисом* n-мерного евклидова пространства. Декартовым базисом эта совокупность векторов называется потому, что

$$\langle \vec{e}_r, \vec{e}_s \rangle = 0, \; r \neq s,$$

т. е. векторы попарно перпендикулярны, и любой вектор $\vec{u}(u_1, u_2, \ldots, u_n)$ можно представить в виде

$$\vec{u} = u_1\vec{e_1} + u_2\vec{e_2} + \cdots + u_n\vec{e_n},$$

что легко проверить. Заметим еще, что

$$|\vec{e_r}| = 1,$$

т. е. $\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}$ — единичные векторы.

Пример 2. Векторы $\vec{u_1}(1, -2, 3)$, $\vec{u_2}(2, 1, -1)$, $\vec{u_3}(4, -3, 5)$ линейно зависимы, так как

$$2\vec{u_1} + \vec{u_2} - \vec{u_3} = \vec{0}\,(0, 0, 0).$$

Вопрос о линейной зависимости векторов сводится к исследованию системы уравнений первой степени. Действительно, по определению, векторы $\vec{a_k}(a_{k1}, a_{k2}, \ldots, a_{kn})$, $k = 1, 2, \ldots, r$, линейно зависимы тогда и только тогда, когда вектор $\lambda_1\vec{a_1} + \lambda_2\vec{a_2} + \cdots + \lambda_r\vec{a_r}$ нулевой, причем не все скаляры λ_i равны нулю. Но это условие равносильно условиям

$$\begin{aligned} \lambda_1 a_{11} + \lambda_2 a_{21} + \cdots + \lambda_r a_{r1} &= 0, \\ \lambda_1 a_{12} + \lambda_2 a_{22} + \cdots + \lambda_r a_{r2} &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ \lambda_1 a_{1n} + \lambda_2 a_{2n} + \cdots + \lambda_r a_{rn} &= 0. \end{aligned} \tag{4}$$

Следовательно, для того чтобы векторы $\vec{a_k}(a_{k1}, a_{k2}, \ldots, a_{kn})$, $k = 1, 2, \ldots, r$, были линейно зависимыми, необходимо и достаточно, чтобы система уравнений (4) с неизвестными $\lambda_1, \lambda_2, \ldots, \lambda_r$ имела решение, удовлетворяющее требованию $\sum_{s=1}^{r} \lambda_s^2 \neq 0$.

Исходя из теории определителей, легко доказать теорему, которую приводим для $n = 3$: *для того чтобы три вектора*

$$\vec{a_1}(a_{11}, a_{12}, a_{13}), \quad \vec{a_2}(a_{21}, a_{22}, a_{23}), \quad \vec{a_3}(a_{31}, a_{32}, a_{33})$$

были линейно зависимыми, необходимо и достаточно, чтобы выполнялось условие

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 0.$$

§ 8. Процесс Сонина — Шмидта. Русский математик Н. Я. Сонин[1] и немецкий математик Э. Шмидт[2] независимо друг от друга решили важную задачу, которую сформулируем следующим образом:

Пусть заданные векторы $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$ *линейно независимы, требуется найти векторы* $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m$, *удовлетворяющие условиям:*

1) $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m$ *должны линейно выражаться через* $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m$, *т. е.*

$$\vec{b}_1 = \alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \cdots + \alpha_m\vec{a}_m,$$
$$\vec{b}_2 = \beta_1\vec{a}_1 + \beta_2\vec{a}_2 + \cdots + \beta_m\vec{a}_m,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

где α_i, β_i — *числа (скаляры);*

2) $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m$ *должны быть линейно независимыми;*

3) $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m$ *должны быть ортонормированными, т. е.*

$$\langle\vec{b}_r, \vec{b}_s\rangle = 0, \ r \neq s,$$
$$\langle\vec{b}_r, \vec{b}_r\rangle = |\vec{b}_r|^2 = 1.$$

Рассмотрим решение задачи по Сонину — Шмидту. Положим

$$\vec{b}_1 = \frac{\vec{a}_1}{|\vec{a}_1|}$$

и заметим, что

$$\langle\vec{b}_1, \vec{b}_1\rangle = \left\langle \frac{\vec{a}_1}{|\vec{a}_1|}, \frac{\vec{a}_1}{|\vec{a}_1|} \right\rangle = \frac{1}{|\vec{a}_1|^2}\langle\vec{a}_1, \vec{a}_1\rangle = 1.$$

Затем рассмотрим вектор $\vec{c}_2 = \vec{a}_2 - \langle\vec{a}_2, \vec{b}_1\rangle\vec{b}_1$. Поскольку векторы $\vec{a}_1$, $\vec{a}_2$, а значит, и векторы $\vec{a}_2$, $\vec{b}_1$ линейно независимы, то вектор $\vec{c}_2$ ненулевой и линейно выражается через $\vec{a}_1$, $\vec{a}_2$.

[1] Сонин Н. Я. (1849—1915) — русский математик, академик.
[2] Шмидт Э. (1876—1959) — немецкий математик, академик.

Кроме того, $\langle\vec{c}_2, \vec{b}_1\rangle = \langle\vec{a}_2, \vec{b}_1\rangle - \langle\vec{a}_2, \vec{b}_1\rangle\langle\vec{b}_1, \vec{b}_1\rangle = 0$. Теперь положим

$$\vec{b}_2 = \frac{\vec{c}_2}{|\vec{c}_2|}.$$

Таким образом, нашли два вектора $\vec{b}_1$, $\vec{b}_2$ (читатель может самостоятельно проверить их линейную независимость), удовлетворяющие всем условиям задачи.

Теперь положим

$$\vec{c}_3 = \vec{a}_3 - \langle\vec{a}_3, \vec{b}_2\rangle\vec{b}_2 - \langle\vec{a}_3, \vec{b}_1\rangle\vec{b}_1.$$

Этот вектор ненулевой и линейно выражается через $\vec{b}_1$, $\vec{b}_2$, $\vec{a}_3$ и тем самым через $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$ (почему?). Кроме того,

$$\langle\vec{c}_3, \vec{b}_1\rangle = \langle\vec{a}_3, \vec{b}_1\rangle - \langle\vec{a}_3, \vec{b}_2\rangle\langle\vec{b}_2, \vec{b}_1\rangle - \langle\vec{a}_3, \vec{b}_1\rangle\langle\vec{b}_1, \vec{b}_1\rangle =$$
$$= \langle\vec{a}_3, \vec{b}_1\rangle - \langle\vec{a}_3, \vec{b}_2\rangle \cdot 0 - \langle\vec{a}_3, \vec{b}_1\rangle \cdot 1 = 0.$$

Точно так же проверяется равенство $\langle\vec{c}_3, \vec{b}_2\rangle = 0$. Положим

$$\vec{b}_3 = \frac{\vec{c}_3}{|\vec{c}_3|}.$$

Векторы $\vec{b}_1$, $\vec{b}_2$, $\vec{b}_3$ удовлетворяют всем условиям задачи. Далее полагаем

$$\vec{c}_4 = \vec{a}_4 - \langle\vec{a}_4, \vec{b}_3\rangle\vec{b}_3 - \langle\vec{a}_4, \vec{b}_2\rangle\vec{b}_2 - \langle\vec{a}_4, \vec{b}_1\rangle\vec{b}_1,$$
$$\vec{b}_4 = \frac{\vec{c}_4}{|\vec{c}_4|}$$

и т. д.

Замечание 1. Уже отмечалось, что $\vec{c}_2$, $\vec{c}_3$, ... — ненулевые векторы. Это дает возможность делить их на $|\vec{c}_2|$, $|\vec{c}_3|$, ...

Замечание 2. На каждом шаге решения задачи обращалось внимание, что $\vec{a}_1$, $\vec{a}_2$, ... можно линейно выразить через $\vec{b}_1$, $\vec{b}_2$, ...

Пример. Применим процесс Сонина — Шмидта к векторам

$$\vec{a}_1\,(2,\ 1,\ 2),\quad \vec{a}_2\,(1,\ 2,\ 1),\quad \vec{a}_3\,(1,\ 1,\ 2).$$

Полагаем, что

$$\vec{b}_1 = \frac{\vec{a}_1}{|\vec{a}_1|} = \vec{b}_1\left(\frac{2}{3},\ \frac{1}{3},\ \frac{2}{3}\right).$$

Затем

$$\vec{c_2}=\vec{a_2}-\langle\vec{a_2},\vec{b_1}\rangle\vec{b_1}=\vec{a_2}(1,\ 2,\ 1)-\left(\frac{2}{3}\cdot 1+\frac{1}{3}\cdot 2+\right.$$

$$\left.+\frac{2}{3}\cdot 1\right)\vec{b_1}\left(\frac{2}{3},\ \frac{1}{3},\ \frac{2}{3}\right)=\vec{c_2}\left(-\frac{1}{3},\ \frac{4}{3},\ -\frac{1}{3}\right),$$

$$|\vec{c_2}|=\sqrt{2},$$

$$\vec{b_2}=\vec{b_2}\left(-\frac{1}{3\sqrt{2}},\ \frac{4}{3\sqrt{2}},\ -\frac{1}{3\sqrt{2}}\right).$$

Далее находим

$$\vec{c_3}=\vec{a_3}-\langle\vec{a_3},\vec{b_2}\rangle\vec{b_2}-\langle\vec{a_3},\vec{b_1}\rangle\vec{b_1}=\vec{a_3}(1,\ 1,\ 2)-$$

$$-\frac{1}{3\sqrt{2}}\vec{b_2}\left(-\frac{1}{3\sqrt{2}},\ \frac{4}{3\sqrt{2}},\ -\frac{1}{3\sqrt{2}}\right)-\frac{7}{3}\vec{b_1}\left(\frac{2}{3},\ \frac{1}{3},\ \frac{2}{3}\right)=$$

$$=\vec{c_3}\left(-\frac{1}{2},\ 0,\ \frac{1}{2}\right),$$

$$|\vec{c_3}|=\frac{\sqrt{2}}{2},$$

$$\vec{b_3}=\vec{b_3}\left(-\frac{1}{\sqrt{2}},\ 0,\ \frac{1}{\sqrt{2}}\right).$$

Рекомендуем проверить самостоятельно, что совокупность векторов $\vec{b_1}$, $\vec{b_2}$, $\vec{b_3}$ ортонормированная.

Заметим еще, что

$$\vec{b_1}=\frac{1}{3}\vec{a_1},\ \vec{b_2}=\frac{1}{\sqrt{2}}\vec{a_2}-\frac{2}{3}\vec{a_1},\ \vec{b_3}=\sqrt{2}\,\vec{a_3}-\frac{1}{6}\vec{a_2}+\frac{2-7\sqrt{2}}{9\sqrt{2}}\vec{a_1},$$

$$\vec{a_1}=3\vec{b_1},\ \vec{a_2}=\sqrt{2}\,\vec{b_2}+2\sqrt{2}\,\vec{b_1},\ \vec{a_3}=$$

$$=\frac{1}{\sqrt{2}}\vec{b_3}+\frac{1}{6}\vec{b_2}+\frac{7}{3\sqrt{2}}\vec{b_1}.$$

§ 9. Геометрические истолкования. Мы могли бы, шаг за шагом, показать, что приведенные в предыдущих параграфах определения, теоремы и формулы при $n=2$ и $n=3$ аналогичны тем определениям, теоремам, правилам, которые знакомы читателю из школьных курсов элементарной математики и физики. Ограничимся здесь случаем $n=2$ и наиболее важными фактами.

1. Пусть имеем вектор $\vec{a}(a_1,\ a_2)$. Возьмем в плоскости декартову систему координат и найдем точку M с координатами $(a_1,\ a_2)$. Соединим начало координат с этой точкой M, снабдим отрезок OM стрелкой (рис. 6) и, таким образом, получим геометрическое представление о векторе.

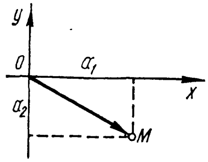

Рис. 6

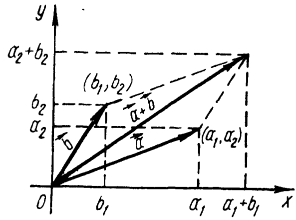

Рис. 7

2. Рассмотрим векторы $\vec{a}(a_1, a_2)$, $\vec{b}(b_1, b_2)$. Тогда $\vec{a}+\vec{b}=\vec{c}(a_1+b_1,\ a_2+b_2)$, $\vec{a}-\vec{b}=\vec{c}\ (a_1-b_1,\ a_2-b_2)$. Из рис. 7 видно, что введенное нами правило сложения это не что иное, как известное правило параллелограмма.

3. Рассмотрим скалярное произведение

$$\langle\vec{a},\ \vec{b}\rangle = a_1b_1 + a_2b_2.$$

Имеем (рис. 8)

$$\langle\vec{a},\ \vec{b}\rangle = |\vec{a}|\,|\vec{b}|\left(\frac{a_1}{|\vec{a}|}\cdot\frac{b_1}{|\vec{b}|}+\frac{a_2}{|\vec{a}|}\cdot\frac{b_2}{|\vec{b}|}\right)=$$

$$=|\vec{a}|\,|\vec{b}|(\cos\alpha\cos\beta+\sin\alpha\sin\beta)=|\vec{a}|\,|\vec{b}|\cos(\beta-\alpha).$$

Мы получили известное правило вычисления работы.

4. Попытаемся понять геометрический смысл процесса Сонина — Шмидта (рис. 9). Построение вектора $\vec{b}_1 = \dfrac{\vec{a}_1}{|\vec{a}_1|}$ естественно и понятно. Построение вектора $\vec{c}_2$ тоже естественно и мы получаем:

$$\vec{b}_2=\lambda\vec{c}_2,\ |\vec{b}_2|=\lambda|\vec{c}_2|;\ 1=\lambda|\vec{c}_2|,\ \lambda=\frac{1}{|\vec{c}_2|},$$

$$\vec{c}_2=\vec{a}_2+\alpha\vec{b}_1.$$

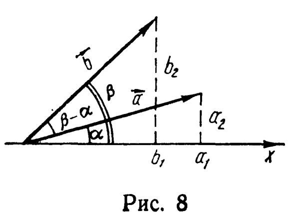

Рис. 8

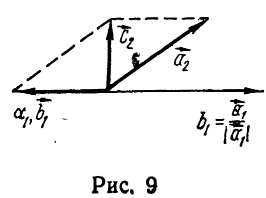

Рис. 9

Чтобы найти α, умножим скалярно последнее равенство на $\vec{b}_1$ и, так как $\langle \vec{c}_2, \vec{b}_1 \rangle = 0$, получим

$$0 = \langle \vec{a}_2, \vec{b}_1 \rangle + \alpha \langle \vec{b}_1, \vec{b}_1 \rangle = \langle \vec{a}_2, \vec{b}_1 \rangle + \alpha,$$
$$\alpha = -\langle \vec{a}_2, \vec{b}_1 \rangle,$$
$$\vec{c}_2 = \vec{a}_2 - \langle \vec{a}_2, \vec{b}_1 \rangle \vec{b}_1.$$

Когда сделаны два первых шага по методу Сонина—Шмидта, нетрудно сделать третий и последующие шаги аналогично и без рисунков.

§ 10. Геометрический смысл линейного уравнения. Выясним геометрический смысл уравнения первой степени

$$Ax + By + C = 0, \qquad (1)$$

где A, B, C — числа. Будем считать x и y декартовыми координатами точки. Выражение $Ax + By$ имеет вид скалярного произведения. Поэтому рассмотрим векторы $\vec{r}(x, y)$, $\vec{n}(A, B)$ и запишем уравнение (1) в виде

$$\langle \vec{n}, \vec{r} \rangle + C = 0. \qquad (2)$$

Пусть $N(x_0, y_0)$ — точка, координаты которой удовлетворяют уравнению (1). Введем вектор $\vec{r}_0(x_0, y_0)$. Тогда

$$\langle \vec{n}, \vec{r}_0 \rangle + C = 0.$$

Найдя из последнего равенства C и подставив его значение в (2), получим уравнение

$$\langle \vec{n}, (\vec{r} - \vec{r}_0) \rangle = 0, \qquad (3)$$

равносильное уравнению (1). Уравнение (3) легко истолковать (рис. 10): при изменении x и y, связанных уравнением

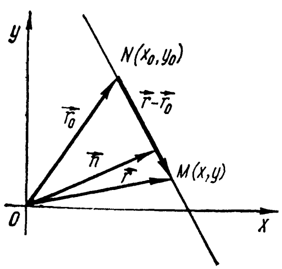

Рис. 10

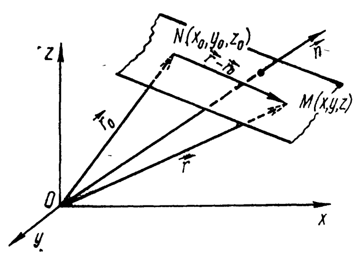

Рис. 11

(1), следовательно, при изменении вектора $\vec{r}(x, y)$ соответствующая точка M — конец вектора $\vec{r}$ — перемещается так, что вектор $\vec{r}-\vec{r}_0$ остается перпендикулярным к фиксированному направлению $\vec{n}$. Поскольку вектор (его конец) $\vec{r}_0$ фиксирован, то конец вектора $\vec{r}$ (точка $M(x, y)$) перемещается по прямой, проходящей через точку $N(x_0, y_0)$, перпендикулярно к $\vec{n}$.

Итак, уравнение (1) геометрически (в декартовых координатах) на плоскости представляет собой прямую, перпендикулярную к вектору $\vec{n}(A, B)$ и проходящую через точку (x_0, y_0), абсциссу x_0 которой можно задать, а ординату y_0 найти из равенства $Ax_0+By_0+C=0$.

Уравнение (1) называется *линейным*, а $\vec{n}$ — *нормальным вектором* прямой. Если уравнение (1) определяет прямую, то для ее построения необходимо найти две точки.

Теперь рассмотрим уравнение

$$Ax+By+Cz+D=0, \tag{4}$$

которое также называем *линейным*. Введем в рассмотрение векторы

$$\vec{r}(x, y, z), \ \vec{n}(A, B, C), \ \vec{r}_0(x_0, y_0, z_0),$$

где x_0, y_0, z_0 удовлетворяют единственному условию Ax_0+ $+By_0+Cz_0+D=0$. Уравнение (4) можно заменить равносильным ему векторным уравнением

$$\langle\vec{n}, (\vec{r}-\vec{r}_0)\rangle=0, \tag{5}$$

которое по виду ничем не отличается от уравнения (3). Однако в этом случае имеем трехмерное пространство, в нем: фиксированную точку $N(x_0, y_0, z_0)$ — вектор $\vec{r}_0$; фиксированное направление — вектор $\vec{n}$; переменные векторы $\vec{r}$, $\vec{r}-\vec{r}_0$. Вектор $\vec{r}-\vec{r}_0$ всегда остается перпендикулярным к $\vec{n}$, а так как один из его концов фиксирован, то этот вектор находится в плоскости, проходящей через точку $N(x_0, y_0, z_0)$ перпендикулярно к $\vec{n}$ (рис. 11).

Итак, уравнение (4) определяет плоскость в трехмерном пространстве, если x, y, z считать декартовыми координатами точки. Эта плоскость перпендикулярна к вектору $\vec{n}$, который называется нормальным вектором плоскости, и проходит через точку $N(x_0, y_0, z_0)$, две координаты которой можно задать, а третью координату найти из уравнения $Ax_0 + By_0 + Cz_0 + D = 0$.

Обратив внимание на однотипность уравнений (1), (4) и на одинаковость перехода к уравнениям (3), (5), не отличающимся по виду, запишем уравнение

$$A_1x_1 + A_2x_2 + \cdots + A_nx_n + C = 0 \tag{6}$$

в виде

$$\langle \vec{n}, (\vec{r} - \vec{r}_0)\rangle = 0,$$

где $\vec{r}(x_1, x_2, \ldots, x_n)$, $\vec{n}(A_1, A_2, \ldots, A_n)$, $\vec{r}_0(x_1^0, x_2^0, \ldots, x_n^0)$ — векторы и $A_1x_1^0 + A_2x_2^0 + \cdots + A_nx_n^0 + C = 0$. Поэтому справедливо следующее определение: *совокупность точек n-мерного евклидова пространства, координаты которых удовлетворяют линейному уравнению* (6), *называется* *гиперплоскостью.*

§ 11. Решение задачи на отыскание минимума. Выясним геометрический смысл следующей задачи.

Пусть в трехмерном евклидовом пространстве даны два неколлинеарных вектора $\vec{u}_1$, $\vec{u}_2$ и третий вектор $\vec{v}$, не выражающийся линейно через $\vec{u}_1$, $\vec{u}_2$, т. е. не компланарный с ними:

$$\vec{v} \neq b_1\vec{u}_1 + b_2\vec{u}_2, \tag{1}$$

где b_1, b_2 — числа. Требуется найти коэффициенты a_1, a_2 такие, чтобы величина $\langle \vec{w}, \vec{w}\rangle$, где

$$\vec{w} = \vec{v} - (a_1\vec{u}_1 + a_2\vec{u}_2),$$

стала минимальной, и вычислить $\min\langle \vec{w}, \vec{w}\rangle$.

Выяснить геометрический смысл задачи не трудно. В самом деле, если векторы $\vec{u}_1$, $\vec{u}_2$ не коллинеарны, то они определяют плоскость p (рис. 12), в которой, ввиду формулы (1), вектор $\vec{v}$ не лежит. Вектор $\vec{u} = a_1\vec{u}_1 + a_2\vec{u}_2$ при

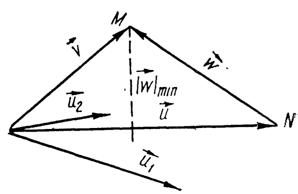

Рис. 12

любых a_1, a_2 лежит в плоскости p, вектор $\vec{w}$ соединяет концы векторов $\vec{v}$, $\vec{u}$ (заданную точку M пространства и некоторую точку N плоскости p) и $\langle\vec{w}, \vec{w}\rangle$ является квадратом его длины. Требуется подобрать вектор $\vec{u}$ так, чтобы длина $\overline{MN}$ был минимальной. Иными словами: рассматриваются отклонения точки M пространства от различных точек плоскости p, требуется найти на p такую точку N, для которой расстояние $\overline{MN}$ было бы минимальным.

Геометрически искомая точка N является проекцией точки M на плоскость p. Чтобы найти N, надо построить проекцию вектора $\vec{v}$ на плоскость p. Для этого заменим пару векторов $\vec{u}_1$, $\vec{u}_2$ ортонормированной парой векторов $\vec{v}_1$, $\vec{v}_2$, расположенных в той же плоскости p, спроектируем $\vec{v}$ на направлении векторов $\vec{v}_1$, $\vec{v}_2$ и полученные векторы-проекции сложим. Замена $\vec{u}_1$, $\vec{u}_2$ на $\vec{v}_1$, $\vec{v}_2$ всегда возможна (процесс Сонина — Шмидта).

Решим данную задачу аналитически (с помощью вычислений). Положим

$$\langle\vec{u}_1, \vec{u}_1\rangle = \langle\vec{u}_2\ \vec{u}_2\rangle = 1, \quad \langle\vec{u}_1\ \vec{u}_2\rangle = 0 \tag{2}$$

и вычислим $\langle\vec{w}, \vec{w}\rangle$. На основании свойств скалярного произведения получим

$$\begin{aligned}\langle\vec{w}, \vec{w}\rangle &= \langle\vec{v} - a_1\vec{u}_1 - a_2\vec{u}_2\ \ \vec{v} - a_1\vec{u}_1 - a_2\vec{u}_2\rangle = \\ &= \langle\vec{v}, \vec{v}\rangle + a_1^2\langle\vec{u}_1, \vec{u}_1\rangle + a_2^2\langle\vec{u}_2, \vec{u}_2\rangle - \\ &- 2a_1\langle\vec{v}, \vec{u}_1\rangle - 2a_2\langle\vec{v}, \vec{u}_2\rangle + 2a_1a_2\langle\vec{u}_1, \vec{u}_2\rangle.\end{aligned}$$

Тогда, используя равенства (2) и вводя обозначения

$$\alpha_1 = \langle\vec{v}, \vec{u}_1\rangle, \ \alpha_2 = \langle\vec{v}, \vec{u}_2\rangle, \tag{3}$$

имеем

$$\langle\vec{w}, \vec{w}\rangle = \langle\vec{v}, \vec{v}\rangle + a_1^2 + a_2^2 - 2a_1\alpha_1 - 2a_2\alpha_2,$$

$$\langle\overleftarrow{w}, \overleftarrow{w}\rangle = \langle\vec{v}, \vec{v}\rangle + (a_1 - \alpha_1)^2 + (a_2 - \alpha_2)^2 - (\alpha_1^2 + \alpha_2^2).$$

Отсюда ясно, что значение $\langle\vec{w}, \vec{w}\rangle$ станет минимальным, если положим

$$a_1 = \alpha_1 = \langle\vec{v}, \vec{u}_1\rangle, \ a_2 = \alpha_2 = \langle\vec{v}, \vec{u}_2\rangle \tag{4}$$

и только в этом случае. Кроме того,

$$\min|\vec{w}|^2 = \min\langle\vec{w}, \vec{w}\rangle = \langle\vec{v}, \vec{v}\rangle - (\alpha_1^2 + \alpha_2^2) = \\ = |\vec{v}|^2 - (\langle\vec{v}, \vec{u}_1\rangle)^2 - (\langle\vec{v}, \vec{u}_2\rangle)^2.$$

Заметим также, что

$$\alpha_1 = |\vec{v}||\vec{u}_1|\cos(\widehat{\vec{v}, \vec{u}_1}) = |\vec{v}|\cdot 1\cdot\cos(\widehat{\vec{v}, \vec{u}_1}),$$

$$\langle\vec{w}, \vec{u}_1\rangle = \langle\vec{v}, \vec{u}_1\rangle - \alpha_1\langle\vec{u}_1, \vec{u}_1\rangle - \alpha_2\langle\vec{u}_1, \vec{u}_2\rangle = \langle\vec{v}, \vec{u}_1\rangle - \\ - \alpha_1 = 0, \ \langle\vec{w}, \vec{u}_2\rangle = 0$$

и поэтому α_1 — это величина проекции вектора $\vec{v}$ на направление вектора $\vec{u}_1$ и точно также α_2 — величина проекции $\vec{v}$ на $\vec{u}_2$, вектор $\vec{w}$ перпендикулярен к плоскости векторов $\vec{u}_1$, $\vec{u}_2$.

Мы получили:

1. Если в трехмерном евклидовом пространстве векторы $\vec{u}_1$, $\vec{u}_2$ образуют ортонормированную пару, а вектор $\vec{v}$ удовлетворяет условию

$$\vec{v} \neq b_1\vec{u}_1 + b_2\vec{u}_2,$$

то величина

$$\langle\vec{w}, \vec{w}\rangle = \langle\vec{v} - a_1\vec{u}_1 - a_2\vec{u}_2, \ \vec{v} - a_1\vec{u}_1 - a_2\vec{u}_2\rangle$$

принимает минимальное значение тогда и только тогда, когда $a_1 = \langle\vec{v}, \vec{u}_1\rangle$, $a_2 = \langle\vec{v}, \vec{u}_2\rangle$. Кроме того,

$$\min\langle\vec{w}, \vec{w}\rangle = \langle\vec{v}, \vec{v}\rangle - (\langle\vec{v}, \vec{u}_1\rangle)^2 - (\langle\vec{v}, \vec{u}_2\rangle)^2.$$

2. Задача отыскания $\min\langle\vec{w}, \vec{w}\rangle$ с геометрической точки зрения означает отыскание проекции точки M (конца вектора $\vec{v}$), не лежащей в плоскости p ортонормированной пары векторов $\vec{u}_1$, $\vec{u}_2$, на плоскость p. Найденное решение означает, что проектируем $\vec{v}$ на направления $\vec{u}_1$, $\vec{u}_2$ и полученные векторы складываем. Конец вектора-суммы и есть искомая проекция точки M. Наконец, $\min\langle\vec{w}, \vec{w}\rangle$ — это квадрат расстояния точки M от своей проекции на плоскость векторов $\vec{u}_1$, $\vec{u}_2$.

Обобщим полученные результаты.

Теорема. *Пусть дана ортонормированная последовательность (линейно независимых) векторов* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$ $(m<n)$ *в* n*-мерном евклидовом пространстве* E_n *и вектор* $\vec{v}$, *не выражающийся линейно через* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$. *Пусть*

$$\vec{w}=\vec{v}-(a_1\vec{u}_1+a_2\vec{u}_2+\cdots+a_m\vec{u}_m), \tag{5}$$

где $a_1, \ldots, a_m$ — *числа. Для того чтобы величина* $\langle\vec{w}, \vec{w}\rangle=|\vec{w}|^2$ *приняла наименьшее значение, необходимо и достаточно положить*

$$a_1=\langle\vec{v}, \vec{u}_1\rangle,\ a_2=\langle\vec{v}, \vec{u}_2\rangle, \ldots, \ a_m=\langle\vec{v}, \vec{u}_m\rangle. \tag{6}$$

Кроме того,

$$\min\langle\vec{w}, \vec{w}\rangle=\langle\vec{v}, \vec{v}\rangle-\{(\langle\vec{v}, \vec{u}_1\rangle)^2+\cdots+(\langle\vec{v}, \vec{u}_m\rangle)^2\}. \tag{7}$$

Доказательство. Имеем

$$\langle\vec{w}, \vec{w}\rangle=\langle\vec{v}-\sum_{i=1}^{m} a_i\vec{u}_i, \ \vec{v}-\sum_{i=1}^{m} a_i\vec{u}_i\rangle=$$

$$=\langle\vec{v}, \vec{v}\rangle+\sum_{i=1}^{m} a_i^2\langle\vec{u}_i, \vec{u}_i\rangle-2\sum_{i=1}^{m} a_i\langle\vec{v}, \vec{u}_i\rangle+$$

$$+2\sum_{(r<s)} a_r a_s\langle\vec{u}_r, \vec{u}_s\rangle=\langle\vec{v}, \vec{v}\rangle+\sum_{i=1}^{m} a_i^2-2\sum_{i=1}^{m} a_i\alpha_i=$$

$$=\langle\vec{v}, \vec{v}\rangle+\sum_{i=1}^{m}(a_i-\alpha_i)^2-\sum_{i=1}^{m}\alpha_i^2, \ \alpha_i=\langle\vec{v}, \vec{u}_i\rangle.$$

Следовательно,

$$\min\langle\vec{w}, \vec{w}\rangle=\langle\vec{v}, \vec{v}\rangle-\sum_{i=1}^{m}(\langle\vec{v}, \vec{u}_i\rangle)^2$$

и достигается при $a_i=\alpha_i=\langle\vec{v}, \vec{u}_i\rangle$. Кроме того,

$$\langle\vec{w}, \vec{u}_r\rangle=\langle\vec{v}-\sum_{i=1}^{m}\alpha_i\vec{u}_i, \vec{u}_r\rangle=\langle\vec{v}, \vec{u}_r\rangle-\alpha_r=0.$$

Исходя из геометрических соображений, сформулируем теорему следующим образом: *расстояние между концом вектора* $\vec{v}$ *в* n*-мерном евклидовом пространстве* E_n *и точками* m*-мерного подпространства пространства* E_n —

концами векторов $a_1\vec{u}_1 + a_2\vec{u}_2 + \cdots + a_m\vec{u}_m$, *где последовательность векторов* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$ *ортонормирована, становится минимальным для той точки подпространства, которая находится как конец вектора, являющегося суммой векторов-проекций* $\vec{v}$ *на направления векторов* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$.

Замечание. Слова «линейно независимых», заключенные в скобки в условии теоремы, можно опустить. В самом деле, пусть векторы $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$ взаимно ортогональны и пусть

$$\lambda_1\vec{u}_1 + \lambda_2\vec{u}_2 + \ldots + \lambda_m\vec{u}_m = 0. \tag{8}$$

Умножив последнее равенство на $\vec{u}_1$, получим

$$\lambda_1\langle\vec{u}_1, \vec{u}_1\rangle + \lambda_2\langle\vec{u}_2, \vec{u}_1\rangle + \ldots + \lambda_m\langle\vec{u}_m, \vec{u}_1\rangle = 0.$$

Отсюда

$$\lambda_1\langle\vec{u}_1, \vec{u}_1\rangle = 0.$$

Но $\langle\vec{u}_1, \vec{u}_1\rangle = |\vec{u}_1|^2 \neq 0$, поэтому $\lambda_1 = 0$. Точно так же можно показать, что $\lambda_2 = \ldots = \lambda_m = 0$. Итак, равенство (8) возможно только при условии $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$. Следовательно, векторы $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$ линейно независимы.

Упражнения

1. Докажите, что угол α между двумя плоскостями

$$\begin{aligned} A_1x + B_1y + C_1z + D_1 &= 0, \\ A_2x + B_2y + C_2z + D_2 &= 0 \end{aligned} \tag{1}$$

определяется формулой

$$\cos\alpha = \frac{A_1A_2 + B_1B_2 + C_1C_2}{\sqrt{A_1^2 + A_2^2 + A_3^2}\sqrt{B_1^2 + B_2^2 + B_3^2}}. \tag{2}$$

2. Докажите, что для того чтобы плоскости (1) были перпендикулярными (параллельными), необходимо и достаточно выполнение условия $A_1A_2 + B_1B_2 + C_1C_2 = 0$ $\left(\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}\right)$. Чем объяснить, что в приведенных условиях и в формуле (2) отсутствуют D_1, D_2?

3. Сформулируйте две задачи, подобные задачам 1 и 2 для двух прямых на плоскости.

4. Докажите, что система уравнений (сколько их ?)

$$\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}$$

определяет прямую в пространстве. Каков смысл каждого из уравнений системы?

5. Докажите, что квадрат площади параллелограма, построенного на векторах $\vec{a}(a_1, a_2, a_3)$, $\vec{b}(b_1, b_2, b_3)$ как на сторонах, равен $(a_1b_2 - a_2b_1)^2 + (a_1b_3 - a_3b_1)^2 + (a_2b_3 - a_3b_2)^2$. Укажите геометрический смысл тождества Лагранжа.

6. Пусть $\langle \vec{u_i}, \vec{u_k} \rangle = \begin{cases} 0, & i \neq k, \\ 1, & i = k, \end{cases}$ $\vec{v} = a_1\vec{u_1} + a_2\vec{u_2} + \ldots + a_m\vec{u_m}$.

Докажите, что $a_i = \langle \vec{v}, \vec{u_i} \rangle$, $i = 1, 2, \ldots, m$.

7. Назовем *векторным произведением* векторов $\vec{a}(a_1, a_2, a_3)$, $\vec{b}(b_1, b_2, b_3)$ вектор $\vec{c}(a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1)$. Докажите, что этот вектор перпендикулярен к плоскости перемножаемых векторов и что длина вектора $\vec{c}$ численно равна площади параллелограма, построенного на векторах $\vec{a}, \vec{b}$. Векторное произведение обозначают двумя способами: $\vec{a} \times \vec{b}$, $[\vec{a}, \vec{b}]$. Докажите, что $[\vec{a}, \vec{b}] = -[\vec{b}, \vec{a}]$.

8. Примените процесс Сонина — Шмидта к векторам $\vec{a}(a_1, 0, 0)$, $\vec{b}(b_1, b_2, 0)$, $\vec{c}(c_1, c_2, c_3)$. Объясните результат геометрически и обобщите его на четырехмерное пространство.

9. Даны три вектора: $\vec{i}(1, 0, 0)$, $\vec{j}(0, 1, 0)$, $\vec{k}(0, 0, 1)$. Вычислите $\vec{i} \times \vec{j}$, $\vec{j} \times \vec{k}$, $\vec{k} \times \vec{i}$.

10. *Векторно-скалярным* или *смешанным произведением* трех векторов $\vec{a}, \vec{b}, \vec{c}$ в трехмерном пространстве называется произведение

$$\vec{a} \cdot (\vec{b} \times \vec{c}) = \langle \vec{a}, \vec{b} \times \vec{c} \rangle.$$

Докажите с точностью до знака, что $\langle \vec{a}, \vec{b} \times \vec{c} \rangle$ равно объему параллелепипеда, построенного на векторах $\vec{a}, \vec{b}, \vec{c}$ как на ребрах.

11. $\langle \vec{a}, \vec{b} \times \vec{c} \rangle = 0$. Что можно сказать о расположении векторов $\vec{a}, \vec{b}, \vec{c}$?

12. Пусть $\vec{c} = \vec{a} - \vec{b}$. Вычислите $\langle \vec{c}, \vec{c} \rangle$ и выясните геометрический смысл полученного результата.

13. Примените процесс Сонина — Шмидта к векторам $\vec{a}(1, -1, 2)$, $\vec{b}(-1, 2, -1)$, $\vec{c}(-1, 7, 4)$. Укажите геометрический смысл результата.

14. Векторы $\vec{a}, \vec{b}, \vec{c}$ взаимно перпендикулярны и образуют с осями координат соответственно углы $\alpha_1, \alpha_2, \alpha_3$; $\beta_1, \beta_2, \beta_3$; $\gamma_1, \gamma_2, \gamma_3$. Докажите, что:

$$\cos^2\alpha_1 + \cos^2\alpha_2 + \cos^2\alpha_3 = 1, \quad \cos^2\beta_1 + \cos^2\beta_2 + \cos^2\beta_3 = 1,$$

$$\cos^2\gamma_1 + \cos^2\gamma_2 + \cos^2\gamma_3 = 1;$$
$$\cos\alpha_1\cos\beta_1 + \cos\alpha_2\cos\beta_2 + \cos\alpha_3\cos\beta_3 = 0,$$
$$\cos\alpha_1\cos\gamma_1 + \cos\alpha_2\cos\gamma_2 + \cos\alpha_3\cos\gamma_3 = 0,$$
$$\cos\beta_1\cos\gamma_1 + \cos\beta_2\cos\gamma_2 + \cos\beta_3\cos\gamma_3 = 0.$$

15. Произведение $\vec{a} \times (\vec{b} \times \vec{c}) = [\vec{a}, [\vec{b}, \vec{c}]]$ называется *векторно-векторным* или *двойным векторным произведением*. Вычислите $\vec{a} \times (\vec{b} \times \vec{c})$, если даны векторы $\vec{a}(1, 0, 0)$, $\vec{b}(0, 1, 0)$ $\vec{c}(0, 0, 1)$.

ГЛАВА II

ПОСЛЕДОВАТЕЛЬНОСТИ ОРТОНОРМИРОВАННЫЕ НА КОНЕЧНОМ МНОЖЕСТВЕ

§ 1. Общие замечания. В предыдущей главе (см. § 4) рассматривались векторы $\vec{a}(a_1, a_2, \ldots, a_n)$. Для них были определены действия сложения и умножения на число и отмечены свойства этих действий. Заметим, что возможны множества объектов, отличных от $\vec{a}(a_1, a_2, \ldots, a_n)$, для которых таким же образом определяются действия сложения и умножения на число, причем сохраняются основные свойства этих действий, перечисленные в гл. I, § 4.

Рассмотрим пример. Пусть дана совокупность E всех четных функций с общей областью определения $-1 \leqslant x \leqslant 1$. Слова «сумма», «умножение на число» будем понимать в том смысле, что, если значениям функций $\varphi(x)$, $\psi(x)$ в точке $x = \alpha$ соответствуют значения $\varphi(\alpha)$, $\psi(\alpha)$, то значению $\varphi(x) + \psi(x)$ соответствует значение $\varphi(\alpha) + \psi(\alpha)$, а значением функции $a\varphi(x)$ является $a\varphi(\alpha)$. Поскольку четность означает, что E состоит из функций, удовлетворяющих условию

$$f(x) \equiv f(-x),$$

то четными будут и функции $\varphi(x) + \psi(x)$, $a\varphi(x)$. Следовательно, применяя действия сложения и умножения на число к четным функциям из E, получим функции из того же E.

Очевидно, имеют место свойства:

1) $f(x) + \varphi(x) = \varphi(x) + f(x)$;

2) $(f(x) + \varphi(x)) + \psi(x) = f(x) + (\varphi(x) + \psi(x))$;

3) $\begin{cases} \lambda(f(x) + \varphi(x)) = \lambda f(x) + \lambda\varphi(x), \\ (\lambda + \mu) f(x) = \lambda f(x) + \mu f(x); \end{cases}$

4) $\lambda\mu f(x) = \lambda(\mu f(x)) = \mu(\lambda f(x))$,

которые аналогичны свойствам, рассмотренным в гл. I, § 4.

В главе I было введено определение скалярного произведения двух векторов, указаны свойства скалярного произведения, а затем установлены некоторые результаты. При этом не использовалось непосредственно определение скалярного произведения, а лишь свойства этого произведения. Отсюда следует: если для некоторых величин введены понятия сложения, умножения на число и понятие скалярного произведения так, что сохраняются свойства этих действий, установленные для векторов, то можно автоматически распространить на рассмотренные величины те понятия, рассуждения и результаты, которые были приведены в главе I.

Пример 1. Рассмотрим множество всех бесконечных числовых последовательностей $\{a_1, a_2, \ldots, a_n, \ldots\}$, обладающих свойством: для каждой последовательности существует $\lim\limits_{n\to\infty}\sum\limits_{i=1}^{n} a_i^2$. (Тогда для каждых двух последовательностей $a=\{a_1, \ldots a_n, \ldots\}$, $b=\{b_1, \ldots b_n, \ldots\}$ существует $\lim\limits_{n\to\infty}\sum\limits_{i=1}^{n} a_i b_i$. Это следует из неравенства Коши — Буняковского и теории пределов). Если принять, что

$$a+b=\{a_1+b_1,\ a_2+b_2, \ldots, a_n+b_n, \ldots\},$$
$$\lambda a=\{\lambda a_1,\ \lambda a_2, \ldots, \lambda a_n, \ldots\},$$
$$\langle a,\ b\rangle=\lim_{n\to\infty}\sum_{i=1}^{n} a_i b_i,$$

то установленные для векторов в n-мерном пространстве свойства сложения, умножения на скаляр и скалярного произведения сохранятся. Например,

$$\lambda\langle a,\ b\rangle=\lambda\lim_{n\to\infty}\sum_{i=1}^{n} a_i b_i=\lim_{n\to\infty}\lambda\sum_{i=1}^{n} a_i b_i=\lim_{n\to\infty}\sum_{i=}^{n}\lambda a_i b_i=$$
$$=\lim_{n\to\infty}\sum_{i=1}^{n}(\lambda a_i)\, b_i=\langle\lambda a,\ b\rangle,$$

но чтобы можно было обобщить и понятие о длине вектора, мы предположили существование предела $\lim\limits_{n\to\infty}\sum\limits_{i=1}^{n} a_i^2$, который и будем считать квадратом длины.

Пример 2. Рассмотрим множество всех непрерывных на отрезке $[a,\ b]$ функций.

Свойство непрерывности функции $f(x)$ при $x=\alpha$ можно определить равенством

$$\lim_{x\to\alpha} f(x)=f(\alpha).$$

Поэтому, если определить сложение двух функций и умножение на число, как это было сделано в начале параграфа для четных функций, то снова будут выполняться все свойства сложения и умножения на число.

Напомним, что функция непрерывна на интервале (отрезке), если она непрерывна в каждой точке интервала (отрезка).

Каждой непрерывной функции $f(x)$ соответствует некоторое число — ее определенный интеграл $\int\limits_a^b f(x)\,dx$ (или площадь фигуры, ограниченной кривой $y = f(x)$ и прямыми $x = a$, $x = b$, $y = 0$, при условии, что площадь, соответствующая участку кривой, лежащей под осью Ox, взята со знаком минус). Эти интегралы (площади) обладают следующими свойствами:

$$\int\limits_a^b (f(x) + \varphi(x))\,dx = \int\limits_a^b f(x)\,dx + \int\limits_a^b \varphi(x)\,dx, \tag{1}$$

$$\int\limits_a^b cf(x)\,dx = c\int\limits_a^b f(x)\,dx, \quad c \text{ — число.} \tag{2}$$

Определим скалярное произведение формулой

$$\langle f(x),\ \varphi(x)\rangle = \int\limits_a^b f(x)\,\varphi(x)\,dx. \tag{3}$$

Тогда, используя свойства (1), (2), легко доказать, что:

1) $\langle f(x),\ \varphi(x)\rangle = \langle \varphi(x),\ f(x)\rangle$;

2) $\langle f(x),\ \varphi(x) + \psi(x)\rangle = \langle f(x),\ \varphi(x)\rangle + \langle f(x),\ \psi(x)\rangle$;

3) $\lambda\langle f(x),\ \varphi(x)\rangle = \langle \lambda f(x),\ \varphi(x)\rangle = \langle f(x),\ \lambda\varphi(x)\rangle$.

Например,

$$\langle f(x),\ \varphi(x) + \psi(x)\rangle = \int\limits_a^b f(x)\,(\varphi(x) + \psi(x))\,dx =$$
$$\int\limits_a^b f(x)\,\varphi(x)\,dx + \int\limits_a^b f(x)\,\psi(x)\,dx = \langle f(x),\ \varphi(x)\rangle + \langle f(x),\ \psi(x)\rangle.$$

Заметим также, что:

4) $\langle g(x),\ g(x)\rangle = \int\limits_a^b g^2(x)\,dx \geqslant 0$;

5) $\langle g(x),\ g(x)\rangle = 0 \leftrightarrow g(x) = 0$.

§ 2. Ортонормированные последовательности функций, заданных на конечном числовом множестве.

Определение. *Если для последовательности (конечной или бесконечной) функций* $f_1(x)$, $f_2(x)$, ... *с общей об-*

ластью определения введено скалярное произведение, то последовательность называется ортонормированной при выполнении условий

$$\langle f_r(x),\ f_s(x)\rangle = 0,\ r \neq s \text{ (ортогональность)}, \tag{4}$$

$$\langle f_r(x),\ f_r(x)\rangle = 1 \text{ (нормированность)}. \tag{5}$$

Нормой функции называется

$$\| f_r(x)\| = \sqrt{\langle f_r(x),\ f_r(x)\rangle},$$

что аналогично длине вектора.

Полезно заметить, что если условие (5) не выполнено, то, положив

$$\varphi_r(x) = \frac{f_r(x)}{\| f_r(x)\|}, \tag{6}$$

получим

$$\langle \varphi_r(x),\ \varphi_r(x)\rangle = \left\langle \frac{f_r(x)}{\| f_r(x)\|},\ \frac{f_r(x)}{\| f_r(x)\|}\right\rangle = \frac{1}{\| f_r(x)\|} \times$$
$$\times \frac{1}{\| f_r(x)\|}\langle f_r(x),\ f_r(x)\rangle = 1.$$

Таким образом, нормирование по формуле (6) дает возможность весьма просто, не нарушая ортогональности, заменить ненормированную последовательность нормированной.

Нас, главным образом, будут интересовать последовательности функций, заданных на некотором конечном множестве значений аргумента. Это связано с тем, что на практике очень часто встречаются функции, заданные графически, и ряд значений функции находим по графику, или функции значения которых находим с помощью непосредственных измерений. Иногда функция задана так, что в принципе можно вычислить (или измерить) значения функции для любого значения аргумента, но в этом нет необходимости или это чрезмерно сложно (а то и дорого), и поэтому находим только несколько значений функции.

Возьмем последовательность

$$\varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_m(x)\ \ldots$$

функций, заданных на конечном множестве $\{x_i\}_{i=1}^n$ значений аргумента, т. е. пусть задана такая таблица:

$$\varphi_1(x_1),\ \varphi_1(x_2),\ \ldots,\ \varphi_1(x_n),$$
$$\varphi_2(x_1),\ \varphi_2(x_2),\ \ldots,\ \varphi_2(x_n),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$

Поскольку каждую строку этой таблицы можно рассматривать как вектор в n-мерном евклидовом пространстве, то естественно определить обычным способом сумму $\varphi_r(x) + \varphi_s(x)$, произведение $\lambda\varphi_r(x)$ и ввести скалярное произведение по формуле

$$\langle \varphi_r(x),\ \varphi_s(x)\rangle = \sum_{i=1}^{n} \varphi_r(x_i)\,\varphi_s(x_i),$$

а норму — по формуле

$$\|\varphi_r(x)\| = \sqrt{\langle \varphi_r(x),\ \varphi_r(x)\rangle}.$$

Тогда условие ортонормированности выразится формулой

$$\sum_{i=1}^{n} \varphi_r(x_i)\,\varphi_s(x_i) = \begin{cases} 0, & r \neq s, \\ 1, & r = s. \end{cases}$$

Пример 1. Пусть $\varphi_1(x) = \frac{x}{\sqrt{14}}$, $\varphi_2(x) = \frac{7}{\sqrt{266}}\left(x^2 - \frac{18}{7}x\right)$ при $x = 0,\ 1,\ 2,\ 3$. Покажем, что имеет место ортонормированность.

Имеем

$$\langle \varphi_1(x),\ \varphi_2(x)\rangle = 0\cdot 0 + \frac{1}{\sqrt{14}}\cdot\frac{7}{\sqrt{266}}\left(1\cdot\left(1 - \frac{18}{7}\right) + 2\left(4 - \frac{36}{7}\right) + 3\left(9 - \frac{18}{7}\cdot 3\right)\right) = \frac{7}{\sqrt{14}\cdot\sqrt{266}}\cdot\frac{7 - 18 + 56 - 72 + 189 - 162}{7} = 0,$$

$$\langle \varphi_1(x),\ \varphi_1(x)\rangle = \frac{1}{14}(0\cdot 0 + 1\cdot 1 + 2\cdot 2 + 3\cdot 3) = 1,$$

$$\langle \varphi_2(x),\ \varphi_2(x)\rangle = \frac{49}{266}\left(0\cdot 0 + \frac{121}{49} + \frac{64}{49} + \frac{81}{49}\right) = 1.$$

Таким образом, ортонормированность имеет место.

Пример 2. Пусть $\varphi_1(x) = \frac{1}{\sqrt{2}}\cos x$, $\varphi_2(x) = \sin x$ при $x = 0,\ \frac{\pi}{2},\ \pi$.

Имеем

$$\langle \varphi_1(x),\ \varphi_2(x)\rangle = \frac{1}{\sqrt{2}}\cdot 0 + 0\cdot 1 + \left(-\frac{1}{\sqrt{2}}\right)\cdot 0,$$

$$\langle \varphi_1(x),\ \varphi_1(x)\rangle = \langle \varphi_2(x),\ \varphi_2(x)\rangle = 1.$$

§ 3. Две ортонормированные последовательности тригонометрических функций.

1. Рассмотрим последовательность функций

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \ldots,\ \cos mx,\ \sin mx,\ \ldots \qquad (1)$$

на множестве $\left\{0, \frac{2\pi}{2n+1}, \frac{4\pi}{2n+1}, \ldots, \frac{4n\pi}{2n+1}\right\}$ и докажем что

$$\langle 1, \cos rx\rangle = \langle 1, \sin rx\rangle = 0. \qquad (2)$$

Проще всего это сделать на основании формулы Эйлера[1] и ее следствия:

$$\cos\alpha + i\sin\alpha = e^{i\alpha}, \ e^{2i\pi} = 1 \ (i^2 = -1). \qquad (3)$$

Действительно, на основании формул (3) получаем

$$\langle 1, \cos rx\rangle + i\langle 1, \sin rx\rangle = \left(1 + \cos\frac{2}{2n+1}\pi r + \right.$$

$$\left. + \cos\frac{4}{2n+1}\pi r + \ldots + \cos\frac{4n}{2n+1}\pi r\right) + i\left(\sin\frac{2}{2n+1}\pi r + \right.$$

$$\left. + \sin\frac{4}{2n+1}\pi r + \ldots + \sin\frac{4n}{2n+1}\pi r\right) =$$

$$= 1 + e^{i\frac{2}{2n+1}\pi r} + e^{i\frac{4}{2n+1}\pi r} + \ldots + e^{i\frac{4n}{2n+1}\pi r} =$$

$$= \frac{1 - e^{i\frac{4n+2}{2n+1}\pi r}}{1 - e^{i\frac{2}{2n+1}\pi r}} = \frac{1 - e^{i\,2\pi}}{1 - e^{i\frac{2}{2n+1}\pi r}} = 0.$$

Тем самым равенства (2) доказаны. Их можно также доказать, используя известные формулы тригонометрии. В самом деле,

$$2\sin\frac{2}{2n+1}\pi r\left(1 + \cos\frac{2}{2n+1}\pi r + \cos\frac{4}{2n+1}\pi r + \ldots + \right.$$

$$\left. + \cos\frac{4n-2}{2n+1}\pi r + \cos\frac{4n}{2n+1}\pi r\right) = 2\sin\frac{2}{2n+1}\pi r +$$

$$+ \sin\frac{4}{2n+1}\pi r + \left(\sin\frac{6}{2n+1}\pi r - \sin\frac{2}{2n+1}\pi r\right) +$$

$$+ \left(\sin\frac{8}{2n+1}\pi r - \sin\frac{4}{2n+1}\pi r\right) + \ldots + \left(\sin\frac{4n}{2n+1}\pi r - \right.$$

$$\left. - \sin\frac{4n-4}{2n+1}\pi r\right) + \left(\sin\frac{4n+2}{2n+1}\pi r - \sin\frac{4n-2}{2n+1}\pi r\right) =$$

$$= \sin\frac{2}{2n+1}\pi r + \sin\frac{4n}{2n+1}\pi r + \sin 2\pi r =$$

---

[1] Эйлер Л. П. (1707—1783) — великий математик, академик Петербургской академии наук.

Доказательство формулы Эйлера можно найти во многих учебниках по курсу высшей математики, а также в книге: Дринфельд Г. И. Квадратура круга и трансцендентность числа π. К.: Вища школа. Головное изд-во, 1976. 84 с.

$$= 2\sin \pi r \cdot \cos \frac{2n-1}{2n+1}\pi r = 0.$$

Следовательно, $\langle 1, \cos rx \rangle = 0$. Аналогично доказывается второе из равенств (2).

Используя равенства (2), получаем

$$\langle \cos rx, \cos sx \rangle = \langle \sin rx, \sin sx \rangle = 0, \quad r \neq s,$$
$$\langle \cos rx, \sin sx \rangle = 0. \tag{4}$$

Действительно, обозначая $\frac{2k}{2n+1}\pi t = x_t$, имеем

$$\langle \cos rx, \cos sx \rangle = \sum_{t=0}^{2n} \cos rx_t \cos sx_t = \frac{1}{2}\sum_{t=0}^{2n} \cos(r+s)x_t +$$
$$+ \frac{1}{2}\sum_{t=0}^{2n} \cos(r-s)x_t = \frac{1}{2}\langle 1, \cos(r+s)x\rangle +$$
$$+ \frac{1}{2}\langle 1, \cos(r-s)x\rangle = 0.$$

Аналогично доказываются и остальные из равенств (4). Наконец, найдем

$$\langle 1, 1 \rangle = 2n+1, \ \langle \cos mx, \cos mx \rangle =$$
$$= \langle \sin mx, \sin mx \rangle = \frac{2n+1}{2}. \tag{5}$$

Действительно, первое из равенств (5) очевидно (множество состоит из $2n+1$ чисел). Второе и третье следуют из равенств (2), так как

$$\langle \cos mx, \cos mx \rangle = \langle 1, \cos^2 mx \rangle = \left\langle 1, \frac{1}{2} + \frac{1}{2}\cos 2mx \right\rangle =$$
$$= \left\langle 1, \frac{1}{2} \right\rangle + \frac{1}{2}\langle 1, \cos 2mx \rangle = \frac{1}{2}\langle 1, 1 \rangle = \frac{2n+1}{2}.$$

Равенства (2), (4), (5) означают, что на множестве $\left\{0, \frac{2\pi}{2n+1}, \frac{4\pi}{2n+1}, \ldots, \frac{4n\pi}{2n+1}\right\}$ последовательность

$$\sqrt{\frac{2}{2n+1}} \cdot \frac{1}{\sqrt{2}}, \ \sqrt{\frac{2}{2n+1}}\cos x, \ \sqrt{\frac{2}{2n+1}}\sin x, \ \ldots,$$
$$\sqrt{\frac{2}{2n+1}}\cos mx, \ \sqrt{\frac{2}{2n+1}}\sin mx, \ \ldots$$

— ортонормирована.

2. Точно так же доказывается ортонормированность последовательности

$$\sqrt{\frac{2}{n}}\cos x, \sqrt{\frac{2}{n}}\sin x, \sqrt{\frac{2}{n}}\cos 2x, \ldots,$$
$$\sqrt{\frac{2}{n}}\cos mx, \sqrt{\frac{2}{n}}\sin mx, \ldots$$

на множестве $\left\{\frac{2\pi}{n}, \frac{4\pi}{n}, \ldots, \frac{2\pi}{n}n = 2\pi\right\}$.

§ 4. Примеры применения процесса Сонина — Шмидта. Изложенный в гл. I, § 8, процесс Сонина — Шмидта вполне применим к последовательности функций на множестве $\{x_1, x_2, \ldots, x_n\}$, поскольку в § 8 мы пользовались только свойствами скалярного произведения. Для рассматриваемой последовательности функций необходимо только ввести скалярное произведение так, чтобы сохранились нужные свойства. Это было сделано в § 2, положив

$$\langle \varphi_r(x), \varphi_s(x)\rangle = \sum_{i=1}^{n} \varphi_r(x_i)\varphi_s(x_i).$$

Итак, не будем здесь излагать метод Сонина — Шмидта, а перейдем к рассмотрению примеров.

Пример 1. Найти три многочлена $\varphi_0(x)$, $\varphi_1(x)$, $\varphi_2(x)$ степеней соответственно 0, 1, 2 так, чтобы они образовали ортонормированную систему на множестве $\{0, 1, 2, 3\}$.

Для решения задачи достаточно применить процесс ортонормирования Сонина — Шмидта к простейшим многочленам 1, x, x^2. Получим:

1) $\varphi_0(x) = \dfrac{1}{\sqrt{\langle 1, 1\rangle}} = \dfrac{1}{\sqrt{1+1+1+1}} = 0{,}5;$

2) $u_1(x) = x - \left\langle x, \dfrac{1}{2}\right\rangle \varphi_0(x) = x - 1{,}5,$

$$\varphi_1(x) = \frac{x-1{,}5}{\sqrt{2{,}25+0{,}25+0{,}25+2{,}25}} = \frac{1}{\sqrt{5}}(x-1{,}5);$$

3) $u_2(x) = x^2 - \left\langle x^2, \dfrac{x-1{,}5}{\sqrt{5}}\right\rangle \dfrac{1}{\sqrt{5}}(x-1{,}5) - \left\langle x^2, \dfrac{1}{2}\right\rangle \times$
$\times \dfrac{1}{2} = x^2 - \left(\langle x^2, x\rangle - \left\langle x^2, \dfrac{3}{2}\right\rangle\right)\dfrac{x-1{,}5}{\sqrt{5}\cdot\sqrt{5}} - \dfrac{1}{4}\langle x^2, 1\rangle = x^2 - 3x +$
$+ 1,$

$$\varphi_2(x) = \frac{x^2-3x+1}{\sqrt{1+1+1+1}} = \frac{1}{2}(x^2-3x+1).$$

Приведенный пример должен был напомнить процесс Сонина — Шмидта и дать некоторое представление об объеме вычислений. Назначение следующего примера другое.

Пример 2. Найти три многочлена степеней 0, 1, 2, которые образуют ортонормированную систему на множестве $\{1, 2\}$.

Применим метод Сонина — Шмидта к многочленам $1, x, x^2$. Имеем:

1) $\varphi_0(x) = \dfrac{1}{\sqrt{1+1}} = \dfrac{1}{\sqrt{2}}$;

2) $u_1(x) = x - \left\langle x, \dfrac{1}{\sqrt{2}} \right\rangle \cdot \dfrac{1}{\sqrt{2}} = x - \dfrac{3}{2}$,

$$\varphi_1(x) = \frac{x - \frac{3}{2}}{\sqrt{\frac{1}{4} + \frac{1}{4}}} = \sqrt{2}\left(x - \frac{3}{2}\right);$$

3) $u_2(x) = x^2 - \left\langle x^2, \sqrt{2}\left(x - \dfrac{3}{2}\right)\right\rangle \cdot \sqrt{2}\left(x - \dfrac{3}{2}\right) - \left\langle x^2, \dfrac{1}{\sqrt{2}}\right\rangle \cdot \dfrac{1}{\sqrt{2}} = x^2 - 2\left(1 + 8 - \dfrac{3}{2} - 6\right)\left(x - \dfrac{3}{2}\right) - \dfrac{1}{2}(1+4) = x^2 - 3x + 2$.

Но $u_2(x)$ уже нельзя нормировать, так как $\langle u_2(x), u_2(x)\rangle = 0$ ибо $u_2(1) = u_2(2) = 0$. Это не случайно, поскольку фактически применили процесс Сонина — Шмидта к векторам $\vec{a}(1, 1)$, $\vec{b}(1, 2)$, $\vec{c}(1, 4)$, среди которых только два первых линейно независимы (больше и не могло быть, так как пространство двумерное). И в самом деле

$$\vec{c} = 3\vec{b} - 2\vec{a}.$$

Процесс Сонина — Шмидта должен был прерваться, ведь он предполагает линейно независимыми данные векторы.

§ 5. Коэффициенты Эйлера — Фурье[1], свойство минимума. Повторим то, что было изложено в гл. 1, § 11, но для случая, когда векторы порождены функциями, заданными на некотором конечном множестве.

Пусть последовательность $\varphi_1(x), \varphi_2(x), \ldots, \varphi_m(x)$ ортонормирована на множестве $\{x_1, x_2, \ldots, x_n\}$, $m \leqslant n$. Коэффициентами Эйлера — Фурье функции $f(x)$, заданной на том же множестве, назовем числа

$$\alpha_i = \langle f(x), \varphi_i(x)\rangle = \sum_{k=1}^{n} f(x_k)\varphi_i(x_k). \qquad (1)$$

Частично роль этих коэффициентов выясняется утверждением:

---

[1] Фурье Ж. (1768—1830) — французский математик.

Теорема 1. Пусть на множестве $\{x_1, x_2, \ldots, x_n\}$ *справедливо равенство*

$$f(x) = A_1\varphi_1(x) + A_2\varphi_2(x) + \cdots + A_m\varphi_m(x). \qquad (2)$$

Тогда

$$A_r = \langle f(x), \varphi_r(x)\rangle = \alpha_r, \quad r = 1, 2, \ldots, m.$$

Действительно, имеем

$$\langle f(x), \varphi_r(x)\rangle = A_1\langle\varphi_1(x), \varphi_r(x)\rangle + \\ + A_2\langle\varphi_2(x), \varphi_r(x)\rangle + \cdots + A_m\langle\varphi_m(x), \varphi_r(x)\rangle.$$

Поскольку

$$\langle\varphi_s(x), \varphi_r(x)\rangle = \begin{cases} 0, & r \neq s, \\ 1, & r = s, \end{cases}$$

то

$$\langle f(x), \varphi_r(x)\rangle = A_r.$$

Особо важна следующая теорема.

Теорема 2 (свойство минимума коэффициентов Эйлера—Фурье). *Пусть последовательность* $\varphi_1(x), \varphi_2(x), \ldots, \varphi_m(x)$ *ортонормирована на множестве* $\{x_1, x_2, \ldots, x_n\}$. *Величина*

$$R_m^2 = \langle f(x) - \sum_{r=1}^{m} A_r\varphi_r(x),\ f(x) - \sum_{r=1}^{m} A_r\varphi_r(x)\rangle = \\ = \sum_{i=1}^{n} [f(x_i) - \sum_{r=1}^{m} A_r\varphi_r(x)]^2 \qquad (3)$$

принимает минимальное значение тогда и только тогда, когда

$$A_r = \alpha_r = \langle f(x), \varphi_r(x)\rangle.$$

Действительно,

$$R_m^2 = \langle f(x), f(x)\rangle + \sum_{r=1}^{m} A_r^2\langle\varphi_r(x), \varphi_r(x)\rangle - \\ - 2\sum_{r=1}^{m} A_r\langle f(x), \varphi_r(x)\rangle + \sum_{r\neq s} A_rA_s\langle\varphi_r(x), \varphi_s(x)\rangle.$$

Поскольку

$$\langle\varphi_r(x), \varphi_r(x)\rangle = 1,\ \langle\varphi_r(x)\ \varphi_s(x)\rangle = 0,\ r \neq s; \\ \langle f(x), \varphi_r(x)\rangle = \alpha_r,$$

то

$$R_m^2 = \langle f(x), f(x)\rangle + \sum_{r=1}^{m} A_r^2 - 2\sum_{r=1}^{m} A_r\alpha_r =$$

$$= \langle f(x),\ f(x)\rangle + \sum_{r=1}^{m} (A_r - \alpha_r)^2 - \sum_{r=1}^{m} \alpha_r^2.$$

Очевидно, R_m^2 будет минимальным тогда и только тогда, когда выполняется условие

$$\sum_{r=1}^{m} (A_r - \alpha_r)^2 = 0,$$

эквивалентное условиям $A_r = \alpha_r$, $r = 1, 2, \ldots, m$. Теорема доказана.

Замечание 1. Минимальное значение R_m^2 равно

$$\rho_m^2 = \langle f(x, f(x)\rangle - \sum_{r=1}^{m} \alpha_r^2 = \sum_{i=1}^{n} f^2 x_i -$$

$$- \sum_{r=1}^{m} (\sum_{i=1}^{n} f(x_i)\, \varphi_r(x_i))^2. \tag{4}$$

Замечание 2. Очевидно, при любых A_r $R_m^2 \geqslant 0$. Следовательно, и $\rho_m^2 \geqslant 0$. Поэтому

$$\sum_{r=1}^{m} \alpha_r^2 \leqslant \langle f(x), f(x)\rangle = \sum_{i=1}^{n} f^2(x_i).$$

Это *неравенство Бесселя*[1].

Замечание 3. Выражение (3) (пренебрегая отсутствием множителя $\frac{1}{n}$) естественно назвать *квадратом среднего квадратического отклонения* величины $A_1\varphi_1(x) + A_2\varphi_2(x) + \cdots + A_m\varphi_m(x)$ от $f(x)$ на множестве $\{x_1, x_2, \ldots, x_n\}$. Поэтому теорему 2 можно сформулировать и так: *среднее квадратическое отклонение линейной комбинации* $\sum_{r=1}^{m} A_r\varphi_r(x)$ *от* $f(x)$, *в случае когда последовательность* $\varphi_1(x)$, $\varphi_2(x), \ldots$ *ортонормирована на множестве* $\{x_1, x_2, \ldots, x_n\}$, *минимально тогда и только тогда, когда в качестве коэффициентов* A_r *взяты коэффициенты Эйлера — Фурье.* Слова «среднее квадратическое отклонение» можно заменить словами «средняя квадратическая погрешность».

Замечание 4. Вообще то, средним квадратическим отклонением величины $u(x)$ от $v(x)$ на $\{x_1, x_2, \ldots, x_n\}$ называют число

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n} [v(x_i) - u(x_i)]^2}.$$

[1] Бессель Ф. В. (1784—1846) — немецкий математик и астроном.

Тогда, совершенно очевидно, σ содержится между наименьшей и наибольшей из величин $|v(x_i)-u(x_i)|$, чем и оправдывается слово «среднее».

Замечание 5. П. Л. Чебышев[1] доказал одно важное неравенство в теории вероятностей. Арифметическая трактовка этого неравенства такова: пусть σ — среднее квадратическое чисел $\Delta_1, \Delta_2, \ldots, \Delta_m$, т. е.

$$\sigma=\sqrt{\frac{\Delta_1^2+\Delta_2^2+\cdots+\Delta_m^2}{m}}.$$

Тогда число k величин Δ_i, больших σt, меньше чем $\frac{m}{t^2}$.

Предположим, что первые k величин Δ_i больше σt. Тогда имеем

$$\sigma>\sqrt{\frac{\Delta_1^2+\cdots+\Delta_k^2}{m}}>\sqrt{\frac{\sigma^2t^2+\cdots+\sigma^2t^2}{m}}=$$
$$=\sigma t\sqrt{\frac{k}{m}}. \tag{5}$$

Значит,

$$\sigma^2>\sigma^2t^2\frac{k}{m}, \quad k<\frac{m}{t^2}.$$

Например, величин Δ_i, больших 3σ, всегда меньше $\frac{m}{9}$. Если учесть, что неравенства (5) получены с «запасом», то можно быть вполне уверенным, что величин Δ_i, больших 3σ, меньше 10 %.

Все сказанное убеждает в том, что среднее квадратическое может служить одной из важных характеристик приближений (измерений).

Следует отметить, что задача о минимизации величины R_m^2 получила геометрическую трактовку (см. гл. I, § 11).

§ 6. Тригонометрическое интерполирование по способу наименьших квадратов. Рассмотрим частный случай изложенной выше (см. § 5, теорема 2) задачи, поскольку она широко применяется на практике.

Пусть заданы значения функции $f(x)$ для значений аргумента $\frac{2\pi}{n}, \frac{4\pi}{n}, \cdots, \frac{2\pi}{n}n=2\pi$. Требуется найти коэффициенты A_i, B_i так, чтобы приближённая формула

$$f(x)\approx(A_1\cos x+B_1\sin x)+(A_2\cos 2x+B_2\sin 2x)+$$
$$+\cdots+(A_m\cos mx+B_m\sin mx) \tag{1}$$

[1] Чебышев П. Л. (1821—1894) — академик, основатель Петербургской математической школы.

была наилучшей в том смысле, что минимальным является среднее квадратическое отклонение левой части формулы от правой на множестве $\left\{\frac{2\pi}{n}, \frac{4\pi}{n}, \ldots, \frac{2\pi}{n}\cdot n\right\}$.

Поскольку на рассматриваемом множестве последовательность функций

$$\sqrt{\frac{2}{n}}\cos x, \ \sqrt{\frac{2}{n}}\sin x, \ \sqrt{\frac{2}{n}}\cos 2x, \ \sqrt{\frac{2}{n}}\sin 2x, \ \ldots$$
$$\ldots, \ \sqrt{\frac{2}{n}}\cos mx, \ \sqrt{\frac{2}{n}}\sin mx \qquad (2)$$

ортонормирована (см. § 3), то квадрат среднего квадратического отклонения

$$\sum_{i=1}^{n}\left[f\left(\frac{2\pi i}{n}\right)-\sum_{k=1}^{m}\left(A_i\cos\frac{2\pi i}{n}+B_i\sin\frac{2\pi i}{n}\right)\right]^2$$

будет минимальным, если $A_i\sqrt{\frac{n}{2}}$ и $B_i\sqrt{\frac{n}{2}}$ бу ут коэффициентами Эйлера — Фурье функции $f(x)$ относительно функций (2), т. е. если

$$A_i\sqrt{\frac{n}{2}}=\sum_{i=1}^{n}f\left(\frac{2\pi i}{n}\right)\sqrt{\frac{2}{n}}\cos\frac{2\pi i}{n},$$
$$B_i\sqrt{\frac{n}{2}}=\sum_{i=1}^{n}f\left(\frac{2\pi i}{n}\right)\sqrt{\frac{2}{n}}\sin\frac{2\pi i}{n}$$

или, что то же самое,

$$A_i=\frac{2}{n}\sum_{i=1}^{n}f\left(\frac{2\pi i}{n}\right)\cos\frac{2\pi i}{n},$$
$$B_i=\frac{2}{n}\sum_{i=1}^{n}f\left(\frac{2\pi i}{n}\right)\sin\frac{2\pi i}{n}. \qquad (3)$$

Приближенной формулой (1) с коэффициентами (3) можно пользоваться при вычислении значений, принимаемых функцией $f(x)$ для значений x, отличных от $\frac{2\pi i}{n}$ $(i=1, 2, \ldots, n)$, но содержащихся на отрезке $[0, 2\pi]$. Формулу (1) назовем *интерполирующей по способу наименьших* *квадратов*, а значения аргумента $\frac{2\pi i}{n}$ — *узлами интерполирования.*

Точность приближения характеризуется средним квадратическим отклонением, равным (см. § 5, равенство (4))

$$\rho_m^2 = \sum_{i=1}^{n} f^2\left(\frac{2\pi i}{n}\right) - \frac{n}{2}\sum_{i=1}^{n}(A_i^2 + B_i^2).$$

Замечание 1. Совершенно так же можно рассмотреть случай задания функции $f(x)$ на множестве $\left\{0, \frac{2\pi}{2n+1}, \frac{4\pi}{2n+1}, \ldots, \frac{4\pi n}{2n+1}\right\}$ и функций

$$\sqrt{\frac{2}{2n+1}}\cdot\frac{1}{\sqrt{2}},\ \sqrt{\frac{2}{2n+1}}\cos x,\ \sqrt{\frac{2}{2n+1}}\sin x, \ldots$$
$$\ldots, \sqrt{\frac{2}{2n+1}}\cos mx,\ \sqrt{\frac{2}{2n+1}}\sin mx, \ldots,$$

ортонормированность которых доказана ранее (см. § 3).

Замечание 2. Тригонометрическое интерполирование по способу наименьших квадратов в случае неравностоящих узлов практически не приводит к столь простым (в записи) формулам, как формулы (3), хотя вычисления по формулам (3) тоже громоздки (если не пользоваться электронными вычислительными машинами).

Формулы (3) рекомендуем запомнить.

§ 7. Обобщения. Под символом $\int\limits_a^b f(x)\,dx$, как указывалось в § 1, можно понимать площадь фигуры, ограниченной отрезком $[a, b]$ оси абсцисс, прямыми $x = a$, $x = b$ и кривой $y = f(x)$, но приписывать знак минус той части площади, которая находится под осью абсцисс.

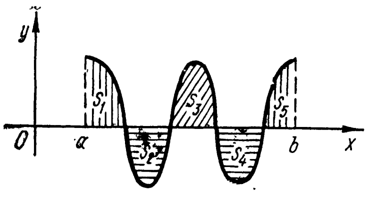

Рис. 13

Так, для фигуры, изображенной на рис. 13, получим

$$\int\limits_a^b f(x)\,dx = S_1 - S_2 + S_3 - S_4 + S_5.$$

Кривая $y = f(x)$ предполагается непрерывной.

На практике часто применяют следующие свойства определенного интеграла:

$$\int\limits_a^b (f(x) + \varphi(x))\,dx = \int\limits_a^b f(x)\,dx + \int\limits_a^b \varphi(x)\,dx, \qquad (1)$$

$$\int_a^b (cf(x))\,dx = c\int_a^b f(x)\,dx, \quad c \text{ — постоянная},$$

которые легко доказываются в интегральном исчислении, а геометрически достаточно очевидны.

Введем обозначение

$$\langle f(x),\ \varphi(x)\rangle = \int_a^b p(x) f(x) \varphi(x)\,dx,$$

где $p(x)$ — положительная функция, называемая *весом*. На основании свойств (1) имеем

$$\langle f(x),\ \varphi(x)\rangle = \langle \varphi(x),\ f(x)\rangle,$$
$$\langle f(x),\ \varphi(x)+\psi(x)\rangle = \langle f(x),\ \varphi(x)\rangle + \langle f(x),\ \psi(x)\rangle,$$
$$\lambda\langle f(x),\ \varphi(x)\rangle = \langle \lambda f(x),\ \varphi(x)\rangle = \langle f(x),\ \lambda\varphi(x)\rangle.$$

Это означает, что $\langle f(x),\ \varphi(x)\rangle$ обладает известными нам свойствами скалярного произведения. Назовем $\langle f(x),\ \varphi(x)\rangle$ скалярным произведением функций $f(x)$ и $\varphi(x)$ с весом $p(x)$. Заметим также, что $\langle f(x),\ f(x)\rangle \geqslant 0$, $\langle f(x),\ f(x)\rangle = 0 \leftrightarrow f(x) \equiv 0$.

Число

$$\sqrt{\langle f(x),\ f(x)\rangle} = \left\{\int_a^b p(x) f^2(x)\,dx\right\}^{\frac{1}{2}}$$

назовем нормой функции $f(x)$ и обозначим $\|f(x)\|$. Поскольку легко можно доказать неравенство Коши — Буняковского

$$\left|\int_a^b p(x) f(x) \varphi(x)\,dx\right|^2 \leqslant \int_a^b p(x) f^2(x)\,dx + \int_a^b p(x) \varphi^2(x)\,dx,$$

то, аналогично доказательству для случая пространства Евклида (см. гл. I, § 3), можно доказать неравенство треугольника

$$\|f(x)+\varphi(x)\| \leqslant \|f(x)\| + \|\varphi(x)\|.$$

Последовательность функций

$$\varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_n(x),\ \ldots,$$

заданных на отрезке $[a,\ b]$, назовем *ортонормированной* относительно веса $p(x)$, $p(x) > 0$, на $[a,\ b]$, если

$$\langle \varphi_r(x), \varphi_s(x)\rangle = \begin{cases} 0, & r \neq s, \\ 1, & r = s. \end{cases}$$

Возможность применения процесса Сонина — Шмидта очевидна.

Ортонормирование последовательности 1, x, x^2, ..., в зависимости от выбора a, b, $p(x)$, приводит к тем или иным классическим многочленам. Приведем примеры (ортогональность сохранена).

1. Полиномы Чебышева: $a = -1$, $b = 1$, $p(x) = \frac{1}{\sqrt{1-x^2}}$

$$T_0(x) = 1,$$
$$T_1(x) = x,$$
$$T_2(x) = 2x^2 - 1,$$
$$T_3(x) = 4x^3 - 3x,$$
$$T_4(x) = 8x^4 - 8x^2 + 1,$$
$$\cdots\cdots\cdots$$

Полиномы Чебышева играют большую роль в некоторых разделах математики и ее приложений, они обладают многими важными свойствами. Приведем некоторые из них:

а) $T_n(x) = \cos(n \arccos x)$;

б) $T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x)$, $k = 2, 3, \ldots$;

в) $\frac{1-tx}{1-2tx+t^2} = T_0(x) + T_1(x)\,t + T_2(x)\,t^2 + \cdots + T_k(x) \times$ $\times\, t^k + \cdots$.

2. Полиномы Лежандра[1]: $a = -1$, $b = 1$, $p(x) = 1$

$$X_0(x) = 1,$$
$$X_1(x) = x,$$
$$X_2(x) = \frac{1}{3}(3x^2 - 1),$$
$$X_3(x) = \frac{1}{5}(5x^3 - 3x),$$
$$\cdots\cdots\cdots$$

Упражнения

1. Пусть на $\{x_1, x_2, \ldots, x_n\}$ функции $\varphi(x)$, $\psi(x)$ ортогональны. Докажите, что хотя бы при одном i $\varphi(x_i)$, $\psi(x_i)$ имеют разные знаки.

2. Пусть функции 1, $\varphi_1(x)$, $\varphi_2(x)$, ..., $\varphi_m(x)$ — ортогональны на $\{x_1, x_2, \ldots, x_n\}$ и непрерывны на отрезке, содержащем $x_1, x_2, \ldots$ ..., x_n. Докажите, что для каждой функции $\varphi_r(x)$ найдутся среди

---

[1] Лежандр А. М. (1752—1833) — французский математик.

заданных значений аргумента два соседних значения x_i, x_{i+1} таких, что $\varphi_r(x_i)\,\varphi_r(x_{i+1})<0$. Сделайте рисунок.

3. Пусть $p(x)>0$ на множестве $\{x_1, x_2, \ldots, x_n\}$ и $\varphi_1(x)$, $\varphi_2(x), \ldots, \varphi_m(x)$ заданы на этом множестве. Проверьте, можно ли определить скалярное произведение формулой

$$\langle \varphi_r(x), \varphi_s(x)\rangle = \sum_{i=1}^{n} p(x_i)\,\varphi_r(x_i)\,\varphi_s(x_i), \qquad (1)$$

сохранив нужные свойства скалярного произведения (функцию $p(x)$ называют весом).

4. Назовем $\varphi_i(x)$, $\varphi_k(x)$ ортонормированными на $\{x_i\}_1^n$, относительно веса $p(x)$, $p(x)>0$, на $\{x_i\}_1^n$, если

$$\langle \varphi_i(x), \varphi_k(x)\rangle = \begin{cases} 0, & i\neq k, \\ 1, & i=k \end{cases}$$

при определении (1), упражнение 3. Сформулируйте определение коэффициентов Эйлера — Фурье, докажите свойства минимума, получите неравенство Бесселя.

5. На множестве $\left\{0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}\right\}$ заданы функции 1, $\sin \pi x$, $\cos \pi x$. Выполните ортонормирование.

6. Найдите коэффициенты Эйлера — Фурье a_0, a_1, a_2 функции $\sin \pi x$, заданной на множестве $\left\{0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}\right\}$, для алгебраических многочленов $\varphi_0(x)$, $\varphi_1(x)$, $\varphi_2(x)$ степеней 0, 1, 2, ортонормированных на этом множестве. Вычислите среднее квадратическое отклонение $\sin \pi x$ от

$$a_0\varphi_0(x)+a_1\varphi_1(x)+a_2\varphi_2(x).$$

7. Докажите, что полином Чебышева $T_n(x)$ имеет n различных действительных корней

$$\cos\frac{\pi}{2n},\ \cos\frac{3\pi}{2n},\ \cos\frac{5\pi}{2n},\ \ldots,\ \cos\frac{2n-1}{2n}\pi.$$

8. Найдите $T_5(x)$, $T_6(x)$, $T_7(x)$.

9. На одном рисунке постройте графики $T_0(x)$, $T_1(x)$, $T_2(x)$, $T_3(x)$. Обратите внимание на расположение точек пересечения этих графиков с осью абсцисс.

10. На одном рисунке постройте графики полиномов Лежандра $X_0(x)$, $X_1(x)$, $X_2(x)$, $X_3(x)$.

11. Докажите второе из равенств (2), § 3.

12. Докажите ортонормированность последовательности, указанной в конце § 3.

13. Воспользовавшись интегралом

$$\int_a^b p(x)(\lambda f(x)+\varphi(x))^2\,dx,$$

докажите неравенство Коши — Буняковского:

$$\left|\int_a^b p(x)\,f(x)\,\varphi(x)\,dx\right| \leqslant \left\{\int_a^b p(x)\,f^2(x)\,dx\right\}^{\frac{1}{2}} \left\{\int_a^b p(x)\,\varphi^2(x)\,dx\right\}^{\frac{1}{2}}.$$

ГЛАВА III

ЭЛЕМЕНТЫ ГАРМОНИЧЕСКОГО АНАЛИЗА И ТРИГОНОМЕТРИЧЕСКОЕ ИНТЕРПОЛИРОВАНИЕ

§ 1. Гармоники, амплитуда, частота, начальная фаза.

Выражения

$$A_1\cos x+B_1\sin x,\ A_2\cos 2x+B_2\sin 2x,\ \ldots,\ A_k\cos kx+B_k\sin kx,\ \ldots$$

называются соответственно гармониками 1, 2, ..., k-го, ... порядков, или короче, первой, второй, ..., k-й гармониками.

В приложениях удобнее записывать гармоники в другой форме, а именно:

$$A_k\cos kx+B_k\sin kx=R_k\sin(kx+\varphi_k). \tag{1}$$

Равенство (1) легко получить. Для этого определим φ_k из уравнений

$$\frac{A_k}{\sqrt{A_k^2+B_k^2}}=\sin\varphi_k,\quad \frac{B_k}{\sqrt{A_k^2+B_k^2}}=\cos\varphi_k$$

(что возможно, так как выполняется условие $\sin^2\varphi_k+\cos^2\varphi_k=1$) и обозначим

$$R_k=\sqrt{A_k^2+B_k^2}.$$

Тогда

$$A_k\cos kx+B_k\sin kx=R_k(\sin\varphi_k\cos kx+\cos\varphi_k\sin kx)=R_k\sin(kx+\varphi_k).$$

Выясним смысл величин R_k, k, φ_k. Для этого прежде всего сравним график k-й гармоники со стандартной синусоидой. Синусоида $y=\sin x$ хорошо известна (рис. 14, a). На отрезке $[0, 2\pi]$ она пересекает ось абсцисс в точках $0, \pi, 2\pi$; ее минимальная ордината равна -1, максимальная $+1$, появляются они по одному разу. Так как $\sin(x\pm 2\pi)\equiv\sin x$, то вне отрезка $[0, 2\pi]$ график получаем сдвигом

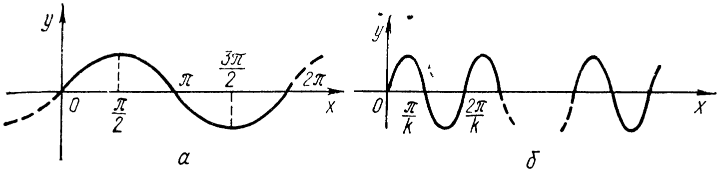

Рис. 14

показанной на рисунке кривой на 2π, 4π, ..., $n\pi$ вправо и влево от начала.

График функции $y = R_k \sin kx$ пересекает ось абсцисс на отрезке $[0, 2\pi]$ в точках $x = 0,\ \frac{\pi}{k},\ \frac{2\pi}{k},\ \ldots\ \pi,\ \frac{k+1}{k}\pi, \ldots$ $\ldots,\ \frac{2\pi k}{k} = 2\pi$, максимальная и минимальная ординаты появляются по k раз, т. е. в k раз чаще, чем у стандартной синусоиды. Максимальная ордината теперь равна R_k, минимальная $-R_k$. График функции $y = R_k \sin kx$ на отрезке $[0, 2\pi]$ имеет вид, указанный на рис. 14, *б*. С помощью сдвига можно получить график вне отрезка $[0, 2\pi]$.

Наконец, чтобы построить график k-й гармоники $y = R_k \sin(kx + \varphi_k)$ достаточно сдвинуть кривую $y = R_k \sin kx$ вдоль оси Ox на φ_k влево, если $\varphi_k > 0$, или вправо, если $\varphi_k < 0$.

Если интерпретировать x как время, y — как отклонение точки, движущейся вдоль Ox, от начала координат, то из рис. 14, *а*, *б* видим, что точка движется колебательно — она отклоняется то вправо (на 1, на R_k), то влево (на -1, на $-R_k$) от положения, занятого в некоторый начальный момент. Если двигаясь по закону $y = \sin x$ точка за время 2π дважды возвращается в начальное положение (совершает одно колебание), то, двигаясь по закону $y = \sin kx$, точка за то же время 2π совершит k колебаний. Наконец, движение по закону $y = R_k \sin(kx + \varphi_k)$ означает: наибольшее удаление от исходного положения R_k, число колебаний (частота) k, начало движения сдвинуто во времени на φ_k (назад или вперед). Вполне естественна общепринятая терминология: R_k — амплитуда, k — частота, φ_k — начальная фаза. С большой точностью описанное движение можно наблюдать, следя за маятником настенных часов или за движением груза, подвешенного на пружине к балке.

§ 2. Понятие о гармоническом анализе функций (ряд Фурье). Многие вопросы математики, электротехники, механики и других прикладных наук приводят к следующей задаче: *представить заданную периодическую, с периодом* 2π, *функцию* $f(x)$ *в виде*

$$f(x)=\frac{A_0}{2}+(A_1\cos x+B_1\sin x)+(A_2\cos 2x+B_2\sin 2x)+\\+\cdots+(A_k\cos kx+B_k\sin kx)+\cdots,\quad 0\leqslant x\leqslant 2\pi, \tag{1}$$

т. е. требуется представить $f(x)$ *в виде суммы гармоник.*

Если число слагаемых в правой части (1) конечно, т. е.

$$f(x)=\frac{A_0}{2}+(A_1\cos x+B_1\sin x)+\cdots+\\+(A_n\cos nx+B_n\sin nx), \tag{2}$$

то $f(x)$ — *тригонометрический многочлен n-го порядка.*

Если число слагаемых в правой части (1) не конечно, то говорят, что функция $f(x)$ *разложена в тригонометрический ряд*, а равенство (2) становится приближенным (тем более точным, чем больше n).

В равенстве (2) коэффициенты A_k, B_k могут быть вычислены по формулам Эйлера — Фурье

$$A_0=\frac{1}{\pi}\int_0^{2\pi} f(x)\,dx;\quad A_k=\frac{1}{\pi}\int_0^{2\pi} f(x)\cos kx\,dx,$$

$$B_k=\frac{1}{\pi}\int_0^{2\pi} f(x)\sin kx\,dx \tag{3}$$

(формула для A_0 получается из формулы для A_k при $k=0$. Для этого мы и писали в (2) не A_0, а $\frac{A_0}{2}$).

Для доказательства формул (3) достаточно воспользоваться элементарными равенствами

$$\int_0^{2\pi} 1\cdot dx=2\pi,\quad \int_0^{2\pi}\cos kx\,dx=\int_0^{2\pi}\sin kx\,dx=0, \tag{4}$$

$$\int_0^{2\pi}\cos kx\sin mx\,dx=0,$$

$$\int_0^{2\pi}\cos kx\cos mx\,dx=\begin{cases}0, & k\neq m,\\ \pi, & k=m,\end{cases} \tag{5}$$

$$\int_0^{2\pi}\sin kx\sin mx\,dx=\begin{cases}0, & k\neq m,\\ \pi, & k=m.\end{cases}$$

Действительно, первую из формул (3) получим из равенства (2) на основании (4); остальные формулы (3) получим на основании (5), предварительно умножив (2) на $\cos kx$ (на $\sin kx$). Формулы (4) легко получить, если читатель знаком с элементами интегрального исчисления. В противном случае следует обратиться к рис. 13 и вспомнить (гл. II, § 7) связь интеграла $\int_a^b F(x)\,dx$ с площадью фигуры (если $F(x) = 1$, то имеем прямоугольник). Формулы (5) — это следствия формул (4), так как, например,

$$\cos^2 kx = \frac{1}{2}(1 + \cos 2kx), \quad \cos kx \cos mx =$$
$$= \frac{1}{2}(\cos(k+m)k + \cos(k-m)x).$$

Формулы (3) называются *формулами Эйлера — Фурье*. За редким исключением, если возможно разложение функции $f(x)$ в тригонометрический ряд, то обязательно коэффициентами ряда являются коэффициенты Эйлера—Фурье, вычисляемые по формулам (3). Тогда говорят, что *функция разложена в ряд Фурье*.

Класс функций, которые можно разложить в ряд Фурье, очень широк. Достаточно, например, чтобы $f(x)$ была непрерывной.

С точки зрения механики разложение функции $f(x)$, $0 \leqslant x \leqslant 2\pi$, в тригонометрический ряд истолковывается так: x — время, $f(x)$ — путь прямолинейно движущейся точки, пройденный за время x, разложение в ряд — представление о движении как о результате наложения (сложения) колебаний.

Приведем два примера, опустив вычисления

Пример 1. Рассмотрим функцию, график которой изображен на рис. 15.

Это периодическая функция, совпадающая с функцией $y = x^2$ на отрезке $[-\pi, \pi]$. Разложение в ряд Фурье можно записать следующим образом:

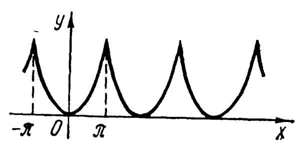

Рис. 15

$$x^2 = \frac{\pi^2}{3} - 4\left(\frac{\cos x}{1^2} - \frac{\cos 2x}{2^2} + \frac{\cos 3x}{3^2} - \dots \pm \frac{\cos mx}{m^2} \mp \dots\right), \quad (6)$$
$$-\pi \leqslant x \leqslant \pi.$$

Отсутствие синусов объясняется четностью функции (симметричностью графика относительно оси ординат).

Написанное разложение является основанием для приближенного равенства

$$x^2 \approx \frac{\pi^2}{3} - 4\left(\cos x - \frac{\cos 2x}{2^2} + \frac{\cos 3x}{3^2}\right) \tag{7}$$

и более точных (это равенство дает грубое приближение).

Пример 2.

$$|\sin x| = \frac{2}{\pi} - \frac{4}{\pi}\left(\frac{\cos 2x}{3} + \frac{\cos 4x}{15} + \cdots + \frac{\cos 2mx}{4m^2-1} + \cdots\right). \tag{8}$$

Замечание. Формулы (4), (5) означают, что последовательность функций

$$\frac{1}{\sqrt{2\pi}},\ \frac{1}{\sqrt{\pi}}\cos x,\ \frac{1}{\sqrt{\pi}}\sin x,\ \ldots,\ \frac{1}{\sqrt{\pi}}\cos kx,\ \frac{1}{\sqrt{\pi}}\sin kx,\ \ldots$$

ортонормирована на отрезке $[0, 2\pi]$, а числа

$$\frac{1}{\sqrt{2\pi}}\int_0^{2\pi} f(x)\,dx,\ \frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\cos kx\,dx,\ \frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\sin kx\,dx,$$

$$k = 1, 2, \ldots,$$

являются коэффициентами Эйлера—Фурье.

Тригонометрический многочлен (и ряд) мы получим в виде

$$\left(\frac{1}{\sqrt{2\pi}}\int_0^{2\pi} f(x)\,dx\right)\frac{1}{\sqrt{2\pi}} + \left(\frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\cos x\,dx\right)\frac{1}{\sqrt{\pi}}\cos x +$$

$$+\left(\frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\sin x\,dx\right)\frac{1}{\sqrt{\pi}}\sin x + \cdots + \left(\frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\cos kx\,dx\right)\times$$

$$\times\frac{1}{\sqrt{\pi}}\cos kx + \left(\frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\sin kx\,dx\right)\frac{1}{\sqrt{\pi}}\sin kx + \cdots.$$

Но это и есть правая часть формул (1), (2), в которых A_i, B_i вычислены по формулам (3).

Формулы (3) определяют, строго говоря, коэффициенты Эйлера—Фурье функции $\frac{1}{\sqrt{\pi}}f(x)$. Поэтому для разложения (1) с коэффициентами (3) выражение для среднего квадратического отклонения и неравенство Бесселя имеют вид

$$\rho_n^2 = \frac{1}{\pi}\int_0^{2\pi} f^2(x)\,dx - \left(\frac{A_0^2}{2} + \sum_{m=1}^{n}(A_m^2 + B_m^2)\right),$$

$$\frac{A_0}{2} + \sum_{m=1}^{n}(A_m^2 + B_m^2) \leqslant \frac{1}{\pi}\int_0^{2\pi} f^2(x)\,dx.$$

Доказательство отличается от приведенного в гл. II, § 5, только тем, что знак $\sum$ меняется на $\int\limits_0^{2\pi}$.

§ 3. О колебании струны. Задача о колебании струны — это одна из задач, решение которой привело к возникновению теории рядов Фурье (а также созданию гармонического анализа). Рассмотрим эту задачу, не приводя доказательств и вычислений. Дадим только некоторое представление о возможности и характере применений гармонического анализа.

Струной называется тонкая и свободно изгибающаяся нить, закрепленная в точках $x = 0$, $x = \pi$ оси Ox, находящаяся в равновесии под действием сильного натяжения, в состоянии равновесия расположенная на оси Ox. Если в начальный момент времени ($t = 0$) оттянуть струну и сообщить каждой ее точке некоторую скорость $\psi(x)$, то струна прийдет в движение. Предположим, что струна движется только в одной плоскости и точки струны движутся перпендикулярно к Ox (внешняя сила пусть отсутствует). Ясно, что струна будет колебаться. Таким образом, имеем свободные (отсутствие внешней силы) поперечные (в одной плоскости) колебания. Точка струны, которая при равновесии находилась на оси Ox на расстоянии x от начала координат в момент t, будет отклонена от Ox на величину $u(t, x)$. Задача заключается в нахождении и изучении $u(t, x)$.

Сведения о форме оттянутой (при $t = 0$) струны и начальных скоростях составляют так называемые начальные условия:

$$u(0, x) = \varphi(x), \quad u'(0, x) = \psi(x) \tag{1}$$

(штрих означает производную по времени t). То, что струна закреплена, дает так называемые граничные условия:

$$u(t, 0) = 0, \quad u(t, \pi) = 0.$$

Сила натяжения и плотность струны влияют на то, что $u(t, x)$ зависит еще от некоторой, постоянной относительно x, t величины, называемой параметром. Для простоты будем считать этот параметр равным единице. С помощью основных законов механики можно составить уравнение, содержащее производные от $u(t, x)$, — так называемое дифференциальное уравнение. Решение этого уравнения с учетом начальных и граничных условий следующее:

$$u(t,\ x)=\sum_{k=1}^{\infty}(A_k\cos kt+B_k\sin kt)\sin kx, \qquad (2)$$

где

$$A_k=\frac{2}{\pi}\int_0^{\pi}\varphi(y)\sin ky\,dy,\quad B_k=\frac{2}{k\pi}\int_0^{\pi}\psi(y)\sin ky\,dy. \qquad (3)$$

Перепишем (2) в виде

$$u(t,\ x)=\sum_{k=1}^{\infty}(R_k\sin kx)\sin(kt+\varphi_k), \qquad (4)$$

где $R_k\sin kx$ — амплитуды, φ_k — фазы.

Формулу (4) можно истолковать так: отклонение (колебание) струны $u(t,\ x)$ является суммой отклонений $((R_k\sin kx)\sin(kt+\varphi_k))$-гармоник. Первая из этих гармоник (у нее наибольшая амплитуда $R_1\sin x$ и наименьшая частота $k=1$) определяет основной тон струны, остальные гармоники определяют *обертоны*. Так как числа R_k убывают достаточно быстро, то влияние обертонов особо ощущается только при малых k, остальные обертоны создают тембр звука.

Дальнейший анализ решения (4), объясняющий почему музыкант прижимает струны пальцами, опускаем.

§ 4. Практический гармонический анализ. В приложениях теории рядов Фурье часто приходится вычислять коэффициенты Эйлера—Фурье:

$$A_0=\frac{1}{\pi}\int_0^{2\pi}f(x)\,dx,\quad A_k=\frac{1}{\pi}\int_0^{2\pi}f(x)\cos kx\,dx,$$

$$B_k=\frac{1}{\pi}\int_0^{2\pi}f(x)\sin kx\,dx. \qquad (1)$$

Выполнить точно это вычисление удается очень редко. Дело в том, что чаще всего встречаются такие случаи: 1) $f(x)$ слишком сложна; 2) $f(x)$ задана графически (график $f(x)$ изображает прибор, например осцилограф); 3) $f(x)$ задана таблично (для некоторых значений аргумента x значения $f(x)$ найдены из наблюдений, с помощью измерений).

Все указанные случаи приводят к одной и той же задаче: приближенно вычислить коэффициенты Эйлера—Фурье. Это и есть задача практического гармонического анализа.

Уточним ее. Прежде всего будем полагать, что на отрезке $[0, 2\pi]$ заданы значения аргумента $x_1, x_2, \ldots, x_n$ и соответствующие значения функции $f(x_1), f(x_2), \ldots, f(x_n)$ (они вычисляются в случае 1), снимаются с графика в случае 2), находятся с помощью измерений в случае 3)).

Существует много формул и правил приближенного вычисления интегралов (1) (геометрически — площадей фигур) — простых и сложных, менее точных и очень точных.

Чем же следует руководствоваться в выборе формулы или правила приближенного вычисления коэффициентов Эйлера—Фурье?

Естественно принять следующий общий принцип: заменяя выражение некоторой величины приближенным выражением, последнее выбирают так, чтобы оно сохранило характерные свойства точной величины. Воспользуемся этим принципом. Интегралы (площади), как известно, обладают свойством линейности:

$$\int_a^b (F(x) + \Phi(x))\,dx = \int_a^b F(x)\,dx + \int_a^b \Phi(x)\,dx,$$

$$\int_a^b cF(x)\,dx = c\int_a^b F(x)\,dx.$$

Следовательно, надо сохранить это свойство, когда вычисляем интеграл $\int_a^b F(x)\,dx$ по значениям $F(x_1), F(x_2), \ldots$ $\ldots, F(x_n)$ функции $F(x)$. Свойством линейности обладает также и сумма вида

$$a_1F(x_1) + a_2F(x_2) + \cdots + a_nF(x_n).$$

Следовательно, надо искать приближенную формулу вида

$$\int_a^b F(x)\,dx \approx \sum_{i=1}^{n} a_iF(x_i).$$

Для коэффициентов (1) Эйлера—Фурье это означает

$$A_0 \approx \frac{1}{\pi}\sum_{i=1}^{n} a_if(x_i),$$

$$A_k \approx \frac{1}{\pi}\sum_{i=1}^{n} \alpha_if(x_i)\cos kx_i, \quad B_k \approx \frac{1}{\pi}\sum_{i=1}^{n} \beta_if(x_i)\sin kx_i. \quad (2)$$

Но коэффициенты Эйлера—Фурье обладают еще *свойством минимума*. Следуя принятому принципу, нужно взять для A_0, A_1, A_2, ..., A_m, B_1, B_2, ..., B_m те значения, при которых величина

$$\sum_{i=1}^{n}\left(f(x_i)-\left(A_0+\sum_{k=1}^{m}\left(A_k\cos kx_i+B_k\sin kx_i\right)\right)\right)^2 \qquad (3)$$

принимает минимальное значение. Задачу минимизации выражения (3) мы решаем для таких наборов значений x_i, для которых функции 1, $\cos kx$, $\sin kx$ ($k=1, 2, \ldots, m$) образуют ортонормированную систему. Если это так, то, как известно (см. гл. II, § 5), выражение (2) принимает минимальное значение, когда в качестве A_0, A_k, B_k ($k = 1, 2, \ldots, m$) берем коэффициенты Эйлера—Фурье функции $f(x)$ на множестве $\{x_1, x_2, \ldots, x_n\}$ значений аргумента.

Функции 1, $\cos kx$, $\sin kx$ ортонормированы на множествах $\left\{0, \frac{2\pi}{2n+1}, \frac{4\pi}{2n+1}, \ldots, \frac{4n\pi}{2n+1}\right\}$, $\left\{\frac{2\pi}{n}, \frac{4\pi}{n}, \ldots, \frac{2\pi}{n}\cdot n\right\}$ (см. гл. II, § 3), если их соответственно умножить на $\sqrt{\frac{2}{2n+1}}$, $\sqrt{\frac{2}{n}}$.

Пусть заданы значения $f\left(\frac{2\pi}{n}\right)$, $f\left(\frac{4\pi}{n}\right)$, ..., $f\left(\frac{2\pi}{n}\cdot n\right)$ функции $f(x)$. Тогда коэффициенты Эйлера—Фурье функции $f(x)$ относительно ортонормированной на множестве $\left\{\frac{2\pi}{n}, \frac{4\pi}{n}, \ldots\right\}$ системы функций $\sqrt{\frac{2}{n}}\cos kx$, $\sqrt{\frac{2}{n}}\sin kx$ будут иметь вид

$$\sqrt{\frac{2}{n}}\sum_{i=1}^{n} f\left(\frac{2\pi i}{n}\right)\cos\frac{2\pi i}{n}, \quad \sqrt{\frac{2}{n}}\sum_{i=1}^{n} f\left(\frac{2\pi i}{n}\right)\sin\frac{2\pi i}{n}$$

и поэтому окончательно получим, что выражение (3) минимально, когда

$$A_0=\frac{2}{n}\sum_{i=1}^{n} f\left(\frac{2\pi i}{n}\right),$$

$$A_k=\frac{2}{n}\sum_{i=1}^{n} f\left(\frac{2\pi i}{n}\right)\cos\frac{2\pi i}{n}k, \quad B_k=\frac{2}{n}\sum_{i=1}^{n} f\left(\frac{2\pi i}{n}\right)\sin\frac{2\pi i}{n}k, \qquad (4)$$

а так как полученные выражения имеют вид (2), то все требования к выбору приближенных значений коэффициентов Эйлера—Фурье будут выполнены, если вычислим их по формулам (4), впервые указанным Бесселем.

Легко выяснить геометрический смысл этих формул. Обозначим

$$\frac{2\pi i}{n} = x_i,\ f(x)\cos kx = \varphi(x),\ x_{i+1} - x_i = \frac{2\pi}{n} = \Delta x_i.$$

Тогда A_k запишется в виде

$$A_k = \frac{1}{\pi}\sum_{i=1}^{n} \varphi(x_i)\,\Delta x_i. \tag{5}$$

Произведение $\varphi(x_i)\,\Delta x_i$ геометрически означает площадь прямоугольника с основанием Δx_i и высотой $\varphi(x_i)$. Таким образом, формула (5) означает, что заменена площадь фигуры,

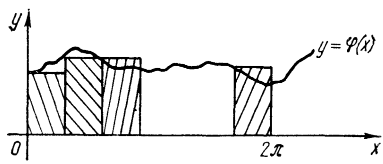

Рис. 16

ограниченной линиями $x = 0$, $x = 2\pi$, $y = 0$, $y = \varphi(x)$, суммой площадей прямоугольников (рис. 16). В интегральном исчислении приближенная формула вида

$$\int_a^b F(x)\,dx \approx \sum_{i=1}^{n} F(x_i)\,\Delta x_i$$

является простейшей и называется *формулой прямоугольников.*

Замечание 1. Для простоты предполагаем, что имеем дело с отрезком $[0, 2\pi]$. Простые преобразования

$$x_1 = 2\pi\,\frac{x-a}{b-a},\quad x_1 = 2\pi\,\frac{x - \frac{a+b}{2}}{b-a} \tag{6}$$

переводят отрезок $[a, b]$ в отрезки $[0, 2\pi]$, $[-\pi, \pi]$.

Замечание 2. До широкого распространения ЭВМ прямое применение формул (4) было весьма нежелательным вследствие большого объема громоздких вычислений. По этой причине математиками и инженерами были разработаны приемы, облегчающие (в значительной мере автоматизирующие) вычисления. Создавались также приборы (гармонические анализаторы), выполняющие гармонический анализ. Много приемов практического гармонического анализа рассмотрено в книге Серебренникова М. Г. Гармонический анализ (М., 1948).

Вычисления по формулам (4) для ЭВМ не представляются сложными — программа проста и затрата машинного времени не велика.

§ 5. **Тригонометрическое интерполирование.** Задача тригонометрического интерполирования формулируется следующим образом.

Пусть известны значения функции $f(x)$ в $2n+1$ точках $x_0, x_1, x_2, \ldots, x_{2n}$, требуется найти тригонометрический многочлен n-го порядка

$$T_n(x) = A_0 + (A_1 \cos x + B_1 \sin x) + \cdots + (A_n \cos nx + B_n \sin nx), \tag{1}$$

удовлетворяющий условиям

$$T_n(x_i) = f(x_i), \ i = 0, 1, \ldots, 2n. \tag{2}$$

Предполагаем, что никакие два заданных значения x_i, x_j не отличаются на число, кратное 2π:

$$x_i - x_j \neq m \cdot 2\pi. \tag{3}$$

В самом деле, если условие (3) не выполнено, то (ввиду периодичности косинуса и синуса) $T_n(x_i) = T_n(x_j)$, и тогда два из условий (2) или совпадают, если $f(x_i) = f(x_j)$, или противоречат друг другу, если $f(x_i) \neq f(x_j)$. Нам же надо иметь точно $2n+1$ различных и непротиворечивых условий, так как определению подлежат ровно $2n+1$ коэффициентов многочлена $T_n(x)$. Если условий меньше чем $2n+1$, то часть коэффициентов останутся неопределенными. Если условий больше чем $2n+1$, то, записав подробно условия $T_n(x_i) = f(x_i)$, получим для определения $2n+1$ величин A_0, A_1, $B_1, \ldots, A_n, B_n$ уравнения

$$f(x_i) = A_0 + (A_1 \cos x_i + B_1 \sin x_i) + \cdots + (A_n \cos nx_i + B_n \sin nx_i), \tag{4}$$

которые могут оказаться несовместимыми (противоречивыми), поскольку уравнений больше, чем неизвестных.

Графически условия (2) означают: графики функции $f(x)$ и многочлена $T_n(x)$ совпадают в точках с заданными абсциссами $x_0, x_1, \ldots, x_{2n}$.

Абсциссы $x_0, x_1, \ldots, x_{2n}$ называются *узлами интерполяции*, $T_n(x)$ — *интерполяционным многочленом*, приближенное равенство

$$f(x) \approx T_n(x)$$

(точное в узлах) называется *интерполяционной формулой*.

Условия (2), записанные в виде (4), представляют собой систему уравнений первой степени с неизвестными A_0, A_1, B_1, ..., A_n, B_n. Решение этой системы не совсем простое.

Пусть $n = 1$. Найдем коэффициенты A_0, A_1, B_1 многочлена

$$T_1(x) = A_0 + A_1 \cos x + B_1 \sin x \tag{4}$$

из уравнений

$$\begin{aligned} f(x_0) &= A_0 + A_1 \cos x_0 + B_1 \sin x_0, \\ f(x_1) &= A_0 + A_1 \cos x_1 + B_1 \sin x_1, \\ f(x_2) &= A_0 + A_1 \cos x_2 + B_1 \sin x_2. \end{aligned} \tag{5}$$

Вычтем первое из уравнений (5) поочередно из равенства (4) и оставшихся двух уравнений (5) и воспользуемся формулами разности косинусов и синусов. Получим

$$T_1(x) - f(x_0) = 2\sin\frac{x-x_0}{2}\left(-A_1 \sin\frac{x+x_0}{2} + B_1 \cos\frac{x+x_0}{2}\right),$$

$$f(x_1) - f(x_0) = 2\sin\frac{x_1-x_0}{2}\left(-A_1 \sin\frac{x_1+x_0}{2} + B_1 \cos\frac{x_1+x_0}{2}\right),$$

$$f(x_2) - f(x_0) = 2\sin\frac{x_2-x_0}{2}\left(-A_1 \sin\frac{x_2+x_0}{2} + B_1 \cos\frac{x_2+x_0}{2}\right).$$

Если из последних двух равенств найдем A_1, B_1 и подставим их значения в первое равенство, то получим

$$T_1(x) = f(x_0)\frac{\sin\frac{x-x_1}{2}\sin\frac{x-x_2}{2}}{\sin\frac{x_0-x_1}{2}\sin\frac{x_0-x_2}{2}} +$$

$$+ f(x_1)\frac{\sin\frac{x-x_0}{2}\sin\frac{x-x_2}{2}}{\sin\frac{x_1-x_0}{2}\sin\frac{x_1-x_2}{2}} + f(x_2)\frac{\sin\frac{x-x_0}{2}\sin\frac{x-x_1}{2}}{\sin\frac{x_2-x_0}{2}\sin\frac{x_2-x_1}{2}}. \tag{6}$$

Проверка: положив в последнем равенстве $x = x_0$, x_1, x_2, соответственно получим

$$T_1(x_0) = f(x_0),\ T_1(x_1) = f(x_1),\ T_1(x_2) = f(x_2).$$

По аналогии можно написать общую формулу

$$T_n(x) = \sum_{i=0}^{2n} \frac{\sin\frac{x-x_0}{2}\ldots\sin\frac{x-x_{i-1}}{2}}{\sin\frac{x_i-x_0}{2}\ldots\sin\frac{x_i-x_{i-1}}{2}} \times$$

$$\times \frac{\sin\frac{x-x_{i+1}}{2}\ldots\sin\frac{x-x_{2n}}{2}}{\sin\frac{x_i-x_{i+1}}{2}\ldots\sin\frac{x_i-x_{2n}}{2}} f(x). \tag{7}$$

Выполнение условий

$$T_n(x_k) = f(x_k),\ k = 0,\ 1,\ \ldots,\ 2n,$$

легко проверяется.

Приведенная формула не единственная (по форме). Так, имеем следующую формулу:

$$T_n^1(x) = \sum_{i=1}^{n+1} \frac{(\cos x - \cos x_1)\ \ldots\ (\cos x - \cos x_{i-1})}{(\cos x_i - \cos x_1)\ \ldots\ (\cos x_i - \cos x_{i-1})} \times$$

$$\times \frac{(\cos x - \cos x_{i+1})\ \ldots\ (\cos x - \cos x_{n+1})}{(\cos x_i - \cos x_{i+1})\ \ldots\ (\cos x_i - \cos x_{n+1})} f(x_i). \tag{8}$$

Очевидно, $f(x_i) = T_n(x_i)$, $i = 1,\ 2,\ \ldots,\ n+1$. В правой части формулы имеем тригонометрический многочлен n-го порядка, так как

$$\cos^2 x = \frac{1}{2}(1 + \cos 2x),$$
$$\cos^3 x = \frac{1}{4}(\cos 3x + 3\cos x), \tag{9}$$
$$\ldots\ldots\ldots\ldots$$

Этот многочлен является четной функцией. Формула

$$f(x) \approx T_n^1(x)$$

применяется, если $f(x)$ — четная функция.

Если $f(x)$ — нечетная функция, то в формуле (8) умножим стоящее под знаком суммы i-е слагаемое на $\frac{\sin x}{\sin x_i}$ и получим многочлен $(n+1)$-го порядка:

$$T_n^2(x) = \sum_{i=1}^{n+1} \frac{\sin x}{\sin x_i}\cdot\frac{(\cos x - \cos x_1)\ \ldots\ (\cos x - \cos x_{i-1})}{(\cos x_i - \cos x_1)\ \ldots\ (\cos x_i - \cos x_{i-1})} \times$$

$$\times \frac{(\cos x - \cos x_{i+1})\ \ldots\ (\cos x - \cos x_{n+1})}{(\cos x_i - \cos x_{i+1})\ \ldots\ (\cos x_i - \cos x_{n+1})} f(x_i), \tag{10}$$

так как

$$\sin x\cos x = \frac{1}{2}\sin 2x,$$

$$\sin x\cos^2 x = \frac{1}{4}\sin 3x + \frac{1}{4}\sin x, \qquad (11)$$

$$\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$$

Наконец, имеем также

$$T_n^3(x) = \sum_{i=1}^{n+1} \frac{(\sin x - \sin x_1)\ \ldots\ (\sin x - \sin x_{i-1})}{(\sin x_i - \sin x_1)\ \ldots\ (\sin x_i - \sin x_{i-1})} \times$$

$$\times \frac{(\sin x - \sin x_{i+1})\ \ldots\ (\sin x - \sin x_{n+1})}{(\sin x_i - \sin x_{i+1})\ \ldots\ (\sin x_i - \sin x_{n+1})} f(x_i). \qquad (12)$$

Формулы (7), (8), (10), (12) легко запоминаются и характеризуются общностью — на узлы не накладываются никакие ограничения. Недостатком этих формул является то, что они не имеют вида

$$T_n(x) = A_0 + A_1\cos x + B_1\sin x + \cdots + A_n\cos nx + \\ + B_n\sin nx = A_0 + R_1\sin(x + \varphi_1) + R_2\sin(2x + \varphi_2) + \cdots \\ \cdots + R_n\sin(nx + \varphi_n), \qquad (13)$$

удобного для приложений.

Если узлы интерполирования выбраны так, что на множестве этих узлов функции 1, $\cos x$, $\sin x$, ..., $\cos nx$, $\sin nx$ взаимно ортогональны, то получить формулу вида (13) несложно. Пусть узлами будут $x_1, x_2, \ldots, x_{2n+1}$. Тогда, поскольку существование интерполяционного многочлена установлено (хотя бы формулой (7)), то найти его коэффициенты можно следующим образом:

$$\langle T_n(x),\ \cos kx\rangle = \langle f(x),\ \cos kx\rangle = \sum_{i=1}^{2n+1} f(x_i)\cos kx_i,$$

кроме того, из равенства (13) имеем

$$\langle T_n(x),\ \cos kx\rangle = A_k\langle \cos kx,\ \cos kx\rangle.$$

Следовательно,

$$A_k = \frac{\langle f(x),\ \cos kx\rangle}{\langle \cos kx,\ \cos kx\rangle}.$$

Аналогично получим

$$B_k = \frac{\langle f(x),\ \sin kx\rangle}{\langle \sin kx,\ \sin kx\rangle}.$$

В частности, если $x_k = \frac{2\pi k}{2n+1}$, $k = 0, 1, \ldots, 2n$, то (см. гл. II, § 3)

$$\langle \cos 0 \cdot x, \cos 0 \cdot x \rangle = 2n+1, \ \langle \cos kx, \cos kx \rangle = \frac{2n+1}{2},$$

$$\langle \sin kx, \sin kx \rangle = \frac{2n+1}{2}, \ k = 1, 2, \ldots, m.$$

В рассматриваемом частном случае

$$A_0 = \frac{1}{2n+1} \sum_{k=0}^{2n} f\left(\frac{2\pi k}{2n+1}\right), \ A_k = \frac{2}{2n+1} \sum_{i=0}^{2n} f\left(\frac{2\pi i}{2n+1}\right) \cos \frac{2\pi i}{2n+1} k,$$

$$B_k = \frac{2}{2n+1} \sum_{i=0}^{2n} f\left(\frac{2\pi i}{2n+1}\right) \sin \frac{2\pi i}{2n+1} k. \tag{14}$$

Упражнения

1. Воспользовавшись равенством (6), § 2, докажите, что

$$\frac{\pi^2}{12} = 1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \cdots + (-1)^{m-1} \frac{1}{m^2} + \cdots$$

Затем найдите π с точностью до 0,01.

2. Докажите, что

$$\frac{\pi^2}{6} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{m^2} + \cdots,$$

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots + \frac{1}{(2k+1)^2} + \cdots.$$

Воспользовавшись последним равенством, вычислите π.

3. Воспользовавшись равенством (8), § 2, докажите, что

$$\frac{1}{2} = \frac{1}{3} + \frac{1}{15} + \cdots + \frac{1}{4m^2-1} + \cdots.$$

Это же равенство проверьте, исходя из того, что

$$\frac{1}{4k^2-1} = \frac{1}{2}\left(\frac{1}{2k-1} - \frac{1}{2k+1}\right).$$

4. Найдите амплитуду и начальную фазу гармоник

$$\cos x + \sqrt{3} \sin x, \quad \sqrt{3} \cos x + \sin x.$$

5. Разбив отрезок $[0, 2\pi]$ на 12 равных частей, примените формулу прямоугольников к вычислению $\int\limits_0^{2\pi} \cos^2 x \, dx$. Сравните с точным ответом, приведенным в § 2, формула 5 (при $k = m = 1$).

6. На одном чертеже постройте графики функций

$$x^2,\ \frac{\pi^2}{3} - 4\cos x,\ \frac{\pi^2}{3} - 4\cos x + \cos 2x,$$

$$\frac{\pi^2}{3} - 4\cos x + \cos 2x - \frac{4}{9}\cos 3x,\ 0 \leqslant x \leqslant 2\pi.$$

7. Воспользовавшись формулами Эйлера

$$\cos x = \frac{e^{ix} + e^{-ix}}{2},\quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}$$

и формулой для $(a+b)^n$, докажите равенства (9), (11), § 5, найдите общие выражения для $\cos^n x$. $\sin^n x$.

8. Пусть $x_k = \frac{2\pi k}{5}$, $k = 0, 1, 2. 3, 4$. Составьте интерполяционные многочлены по формулам (13), (14) и (17), § 5. Проверьте, совпадают ли, в конечном счете, результаты.

9. Положив $x_k = \frac{\pi k}{4}$, $k = 0, 1, 2, 3$, составьте формулу тригонометрического интерполирования для функции $f(x) = x$. На одном чертеже постройте графики функции $f(x)$ и найденного тригонометрического многочлена.

ГЛАВА IV

ИНТЕРПОЛИРОВАНИЕ И СПОСОБ НАИМЕНЬШИХ КВАДРАТОВ

§ 1. Вспомогательные сведения.

I. Программа по математике средней общеобразовательной школы содержит понятие производной, напомним его.

Производной функции $f(x)$ при $x = a$ называется предел отношения приращения функции $f(x)$ в точке a к приращению h аргумента, когда это последнее стремится к нулю:

$$f'(a) = \lim_{|h|\to 0} \frac{f(a+h) - f(a)}{h}.$$

Если предел не существует, то говорят, что функция не имеет производной при $x = a$.

Производная

$$f'(x) = \lim_{|h|\to 0} \frac{f(x+h) - f(x)}{h}$$

сама есть функция аргумента x и называется производной (без добавления слов «при $x = a$»).

Из свойств пределов легко вывести свойства производной. Приведем только два свойства:

$$(f(x)+\varphi(x))'=f'(x)+\varphi'(x),$$

$$(cf(x))'=cf'(x), \text{ где } c \text{ — постоянная.}$$

Легко доказать формулу

$$((x-a)^k)'=k(x-a)^{k-1}.$$

Рассмотрим случай, когда $k=3$ (формула верна и для дробных показателей, но мы ограничимся только целыми). Имеем

$$((x-a)^3)'=\lim_{|h|\to 0}\frac{(x-a+h)^3-(x-a)^3}{h}=$$

$$=\lim_{|h|\to 0}\frac{3(x-a)^2h+3(x-a)h^2+h^3}{h}=$$

$$=\lim_{|h|\to 0}(3(x-a)^2+3h(x-a)+h^2)=3(x-a)^2.$$

Производная от $f'(x)$ называется *второй производной* и обозначается $f''(x)$, третья производная $f'''(x)$ или $f^{(3)}(x)$ — это производная от $f''(x)$ и т. д.

II. Число a наз вается корнем многочлена

$$P(x)=a_0x^n+a_1x^{n-1}+\cdots+a_{n-2}x^2+a_{n-1}x+a_n, \qquad (1)$$

если $P(a)=0$.

Теорема 1. *Для того чтобы a было корнем многочлена* (1), *необходимо и достаточно, чтобы многочлен делился без остатка на $x-a$.*

Доказательство. *Достаточность.* Если $P(x)$ делится без остатка на $x-a$, то

$$P(x)=P_1(x)(x-a),$$

где $P_1(x)$ — частное (многочлен). Ясно, что $P(a)=0$.

Необходимость. Пусть $P(a)=0$, т. е.

$$a_0a^n+a_1a^{n-1}+\cdots+a_{n-1}a+a_n=0. \qquad (2)$$

Вычитая (2) из (1), получаем

$$P(x)=a_0(x^n-a^n)+a_1(x^{n-1}-a^{n-1})+\cdots+a_{n-1}(x-a).$$

Так как разность целых степеней двух чисел делится на разность первых степеней этих чисел, то каждое слагаемое в правой части последнего равенства делится на $x-a$, и потому $P(x)$ делится на $x-a$.

Теорема 2. *Для того чтобы многочлен* (1) *имел n различных корней $a_1, a_2, \ldots, a_n$, необходимо и достаточно, чтобы имело место разложение на множители*

$$P(x) = a_0 (x - a_1)(x - a_2) \ldots (x - a_n).$$

Доказательство. *Достаточность* очевидна.

Необходимость. Если $a_1, a_2, \ldots, a_n$ корни многочлена, то по теореме 1 он должен делиться без остатка на $(x-a_1)$, $(x-a_2), \ldots, (x-a_n)$. Следовательно, он должен делиться без остатка и на произведение

$$(x - a_1)(x - a_2) \ldots (x - a_n), \qquad (3)$$

являющееся многочленом той же степени n. Поэтому, поделив $P(x)$ на произведение (3), обязательно получим постоянное частное. Значит,

$$P(x) = K(x - a_1)(x - a_2) \ldots (x - a_n).$$

Если бы мы выполнили умножение, то коэффициентом при x^n было бы K, но в (1) коэффициент при x^n равен a_0. Отсюда $K = a_0$.

Теорема 3. *Любой многочлен*

$$P(x) = a_0 + a_1 x + a_2 x^2 + \cdots$$

можно расположить по степеням $x - a$, т. е. представить в виде

$$P(x) = A_0 + A_1 (x - a) + A_2 (x - a)^2 + \cdots$$

при любом a.

В самом деле,

$$\begin{aligned} P(x) &= a_0 + a_1((x-a)+a) + a_2((x-a)+a)^2 + \cdots = \\ &= (a_0 + a_1 a + a_2 a^2 + \cdots) + (x-a)(a_1 + 2aa_2 + \cdots) + \\ &+ (x-a)^2 (a_2 + 3aa_3 + \cdots) = A_0 + A_1(x-a) + \\ &+ A_2 (x-a)^2 + \cdots. \end{aligned}$$

Теорема 4 (Ролля)[1]. *Если функция $f(x)$ обладает свойствами:* 1) $f(a) = f(b)$; 2) *для каждого значения x, $a \leqslant x \leqslant b$, $f(x)$ — непрерывна;* 3) *в каждой точке интервала $a < x < b$ существует производная $f'(x)$, то по меньшей мере в одной точке c интервала $f'(c) = 0$.*

Строгое аналитическое доказательство теоремы привести не можем. Геометрически теорема кажется очевидной, она

[1] Ролль М. (1652—1719) — французский математик.

означает существование на графике функции $f(x)$ хотя бы одной точки, в которой касательная к графику параллельна оси абсцисс. Рекомендуем читателю самостоятельно сделать рисунок, начав его с учета условия 1).

Замечание. Геометрическое доказательство нельзя считать строгим хотя бы уже потому, что изображаем кривые, обладающие свойством: при движении точки касания вдоль кривой угол наклона касательной к оси абсцисс изменяется непрерывно. Это означает непрерывность производной, а эта непрерывность не предполагается.

§ 2. Тейлорово разложение многочлена и общая формула Тейлора. Для некоторой функции значение функции и ряда ее производных могут быть заданы, найдены экспериментально или вычислены при $x=a$. Например, для точки, движущейся прямолинейно, в момент времени $x=a$ могут быть заданы путь $S(a)$, пройденный за время a, скорость $S'(a)$ и ускорение $S''(a)$.

В том случае, когда известно, что функция $P(x)$ является многочленом n-й степени и заданы при $x=a$ значение функции и ее производные до n-го порядка включительно, то многочлен $P(x)$ может быть найден при любом x.

Теорема. *Если $P(x)$ многочлен n-й степени, то*

$$P(x)=P(a)+P'(a)(x-a)+\frac{P''(a)}{2!}(x-a)^2+$$
$$+\frac{P'''(a)}{3!}(x-a)^3+\cdots+\frac{P^{(n)}(a)}{n!}(x-a)^n. \qquad (1)$$

Доказательство. Запишем $P(x)$ в виде

$$P(x)=A_0+A_1(x-a)+A_2(x-a)^2+\cdots+A_n(x-a)^n, \qquad (2)$$

что всегда возможно, и последовательно найдем

$$\begin{aligned}
P'(x)&=A_1+2A_2(x-a)+3A_3(x-a)^2+\cdots\\
&\qquad\qquad+nA_n(x-a)^{n-1},\\
P''(x)&=\quad 2A_2 \quad +3\cdot 2A_3(x-a)+\cdots\\
&\qquad\qquad\cdots+n(n-1)A_n(x-a)^{n-2},\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
P^{(n)}(x)&=\qquad\qquad n!\,A_n
\end{aligned} \qquad (3)$$

Положив в формулах (2) и (3) $x=a$, получим

$$A_0=P(a),\ A_1=P'(a),\ A_2=\frac{P''(a)}{2},\ \ldots,\ A_n=\frac{P^{(n)}(a)}{n!}. \qquad (4)$$

Из (2) и (4) следует формула (1), которую называют *формулой Тейлора*[1] *для многочлена.*

[1] Тейлор Б. (1685—1731) — английский математик.

Теперь предположим, что $f(x)$ не многочлен, но известно, что на некотором интервале, содержащем $x=a$, $f(x)$ имеет производные до $(n+1)$-го порядка включительно и пусть заданы

$$f(a),\ f'(a),\ f''(a),\ \ldots,\ f^{(n)}(a). \tag{5}$$

По формуле (1) можем построить многочлен $P(x)$ n-й степени, принимающий, вместе со своими производными, при $x=a$ те же значения (5), т. е.

$$P(a)=f(a),\ P'(a)=f'(a),\ \ldots,\ P^{(n)}(a)=f^{(n)}(a).$$

Тогда получим

$$P(x)=f(a)+f'(a)(x-a)+\frac{f''(a)}{2}(x-a)^2+\cdots$$
$$\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n. \tag{6}$$

Естественно положить

$$f(x)=P(x)+R_n(x)=f(a)+f'(a)(x-a)+$$
$$+\frac{f''(a)}{2}(x-a)^2+\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n+R_n(x). \tag{7}$$

Эта формула называется *формулой Тейлора*, в ней $R_n(x)$ является остаточным членом. Ясно, что $R_n(x)$ — это погрешность приближенной формулы

$$f(x)\approx f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots$$
$$\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n. \tag{8}$$

Пока не найден остаточный член или же оценка для него формула (7) окончательного вида не имеет, а применение формулы (8) может привести к большой погрешности.

В курсе дифференциального исчисления для $R_n(x)$ указаны различные по форме выражения. Самым простым и чаще всего применяемым является остаточный член в форме Лагранжа

$$R_n(x)=\frac{(x-a)^{n+1}}{(n+1)!}f^{(n+1)}(t),\ a<t<x\ \text{(или } x<t<a).$$

Поскольку значение t неизвестно, то остается неизвестным и точное значение выражения для $R_n(x)$. Чаще всего достаточным является очевидная оценка

$$|R_n(x)| \leqslant \frac{|x-a|^{n+1}}{(n+1)!} \cdot \max |f^{(n+1)}(x)|.$$

Например, пусть $f(x) = \sin x$, $-\frac{\pi}{4} \leqslant x \leqslant \frac{\pi}{4}$.

Последовательными производными функции $\sin x$ являются

$$\cos x, \ -\sin x, \ -\cos x, \ \sin x, \ \cos x, \ \ldots . \qquad (9)$$

Поэтому $\max |(\sin x)^{(n+1)}| \leqslant 1$. Положим $a = 0$, тогда для $R_n(x)$ будем иметь оценку

$$|R_n(x)| \leqslant \frac{\pi^{n+1}}{4^{n+1}(n+1)!}.$$

Если возьмем $n = 3$, то

$$|R_3(x)| \leqslant \frac{\pi^4}{4^4 \cdot 24} < \frac{1}{60}.$$

Подставив значения производных (9) при $x = 0$ в (8), получим

$$\sin x \approx x - \frac{x^3}{3!}$$

$\left(\text{погрешность меньше } \frac{1}{60}\right)$.

Замечание 1. Приближенное представление $f(x)$ с помощью многочлена имеет важное значение в теории и приложениях.

Замечание 2. Хотя класс функций, имеющих производные, не очень узок, требование существования производных надо считать жестким.

§ 3. Интерполяционная формула Лагранжа. Предположим известными $n+1$ значений функции $f(x)$ при $n+1$ различных значениях аргумента. Такое предположение осуществляется весьма часто: в определенные моменты времени с большой точностью фиксируются расстояния прямолинейно движущейся точки от некоторого начала; измеряется длина стержня при различных температурах; прибор вычерчивает график некоторой функции, а с помощью этого графика находятся значения функции; функция задана сложной формулой, и потому вычисляются только ее значения для ряда значений аргумента.

Требуется по данным значениям

$$f(a_1), \ f(a_2), \ \ldots, \ f(a_n), \ f(a_{n+1})$$

найти приближенное выражение для $f(x)$ в виде многочлена.

Определение. *Многочлен $P(x)$ не выше n-й степени, обладающий свойством*

$$P(a_1) = f(a_1),\ P(a_2) = f(a_2),\ \ldots,\ P(a_{n+1}) = f(a_{n+1}), \quad (1)$$

называется ***интерполяционным многочленом*** *для $f(x)$, а формула*

$$f(x) \approx P(x), \quad (2)$$

используемая для приближенного нахождения значений $f(x)$ при значениях аргумента, содержащихся между наименьшим и наибольшим из $a_1, a_2, \ldots, a_{n+1}$, называется ***интерполяционной формулой.***

Геометрически условия (1) означают, что графики многочлена $P(x)$ и функции $f(x)$ заведомо имеют не менее чем $n+1$ общих точек.

Требование, чтобы многочлен $P(x)$ был не выше n-й степени, естественно — если взять многочлен более высокой степени, то он будет содержать более чем $n+1$ коэффициентов, для нахождения которых имеется только $n+1$ условий (1), и нельзя ожидать, что коэффициенты определятся однозначно. Если потребовать, чтобы степень многочлена была меньше n, то число условий (1) будет больше числа искомых коэффициентов и, как правило, уравнения для определения неизвестных окажутся несовместными. Итак, надо найти многочлен

$$P(x) = A_0 + A_1 x + \cdots + A_n x^n, \quad (3)$$

удовлетворяющий условиям (1). Эти условия подробно записываются следующим образом:

$$\begin{aligned} A_0 + A_1 a_1 + A_2 a_1^2 + \cdots + A_n a_1^n &= f(a_1), \\ A_0 + A_1 a_2 + A_2 a_2^2 + \cdots + A_n a_2^n &= f(a_2), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ A_0 + A_1 a_{n+1} + A_2 a_{n+1}^2 + \cdots + A_n a_{n+1}^n &= f(a_{n+1}) \end{aligned} \quad (4)$$

и представляют собой систему $n+1$ уравнений первой степени с $n+1$ неизвестными $A_0, A_1, \ldots, A_n$.

Докажем, что если существует решение системы (4), то только одно. Это означает, что интерполяционный многочлен, если он существует, может быть записан в различных формах, но, будучи приведен к виду (3), всегда будет одним и тем же.

Пусть система уравнений (4) имеет два решения: 1) A_0, $A_1, \ldots, A_n$; 2) $B_0, B_1, \ldots, B_n$. Тогда будут тождествами равенства (4) и равенства

$$B_0 + B_1 a_1 + B_2 a_1^2 + \cdots + B_n a_1^n = f(a_1),$$
$$B_0 + B_1 a_2 + B_2 a_2^2 + \cdots + B_n a_2^n = f(a_2),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$B_0 + B_1 a_{n+1} + B_2 a_{n+1}^2 + \cdots + B_n a_{n+1}^n = f(a_{n+1}),$$

Вычитая эти равенства из соответствующих равенств (4), получим

$$(A_0 - B_0) + (A_1 - B_1) a_i + (A_2 - B_2) a_i^2 + \cdots$$
$$\cdots + (A_n - B_n) a_i^n = 0, \ i = 1, 2, \ldots, n+1.$$

Это означает, что многочлен n-й степени

$$(A_0 - B_0) + (A_1 - B_1) x + \cdots + (A_n - B_n) x^n$$

имеет $n+1$ (больше n) различных корней, что невозможно. Мы пришли к противоречию. Значит, больше одного решения система (4) иметь не может.

Ниже найдем многочлен $P(x)$ и тем самым вопрос о существовании решения системы (4) будет решен положительно.

Итак, требуется найти многочлен $P(x)$, удовлетворяющий условиям (1), не решая систему уравнений (4).

Рассмотрим частный случай, когда

$$P(a_1) = f(a_1), \ P(a_2) = 0, \ \ldots, \ P(a_{n+1}) = 0. \tag{9}$$

На основании теоремы 2, § 1, имеем

$$P(x) = a_0 (x - a_2)(x - a_3) \ldots (x - a_{n+1}).$$

Найдем a_0. Для этого воспользуемся первым из условий (9), получим

$$f(a_1) = a_0 (a_1 - a_2)(a_1 - a_3) \ldots (a_1 - a_{n+1}),$$
$$a_0 = \frac{1}{(a_1 - a_2)(a_1 - a_3) \ldots (a_1 - a_{n+1})} f(a_1).$$

Окончательно, многочлен n-й степени

$$P_1(x) = \frac{(x - a_2)(x - a_3) \ldots (x - a_{n+1})}{(a_1 - a_2)(a_1 - a_3) \ldots (a_1 - a_{n+1})} f(a_1)$$

удовлетворяет условиям

$$P_1(a_1) = f(a_1),\ P_1(a_i) = 0 \text{ при } i \neq 1.$$

Аналогично находим многочлен n-й степени

$$P_2(x) = \frac{(x-a_1)(x-a_3)\ \dots\ (x-a_{n+1})}{(a_2-a_1)(a_2-a_3)\ \dots\ (a_2-a_{n+1})} f(a_2),$$

удовлетворяющий условиям

$$P_2(a_1) = 0,\ P_2(a_2) = f(a_2),\ P_2(a_i) = 0 \text{ при } i \neq 2.$$

Затем находим $P_3(x),\ \dots\ ,\ P_{n+1}(x)$. Тогда

$$P(x) = P_1(x) + P_2(x) + \cdots + P_{n+1}(x)$$

является многочленом не выше n-й степени, удовлетворяющим всем условиям (1). Найденный многочлен

$$P(x) = \frac{(x-a_2)(x-a_3)\ \dots\ (x-a_{n+1})}{(a_1-a_2)(a_1-a_3)\ \dots\ (a_1-a_{n+1})} f(a_1) + \\ + \frac{(x-a_1)(x-a_3)\ \dots\ (x-a_{n+1})}{(a_2-a_1)(a_2-a_3)\ \dots\ (a_2-a_{n+1})} f(a_2) + \qquad (10) \\ \dots\dots\dots\dots\dots\dots \\ + \frac{(x-a_1)(x-a_2)\ \dots\ (x-a_n)}{(a_{n+1}-a_1)(a_{n+1}-a_2)\ \dots\ (a_{n+1}-a_n)} f(a_{n+1})$$

называется *интерполяционным многочленом Лагранжа*, а формула $f(x) \approx P(x)$ — *интерполяционной формулой Лагранжа.*

Если предположить существование у $f(x)$ производных до $(n+1)$-го порядка включительно, то можно найти выражение для погрешности интерполяционной формулы. Ниже (см. § 4) мы найдем это выражение при $n = 1$.

Пример 1. Пусть $f(-1) = 6,\ f(0) = -2,\ f(1) = 4,\ f(2) = 12.$ Имеем

$$f(x) \approx \frac{x(x-1)(x-2)}{(-1)\cdot(-2)\cdot(-3)}\cdot 6 + \frac{(x+1)(x-1)(x-2)}{1\cdot(-1)\cdot(-2)}\cdot(-2) + \\ + \frac{(x+1)\,x\,(x-2)}{2\cdot 1\cdot(-1)}\cdot 4 + \frac{(x+1)\,x\,(x-1)}{3\cdot 2\cdot 1}\cdot 12 = -2x^3 + 7x^2 + x - 2.$$

Пример 2. Пусть $f(-1) = 2,\ f(0) = 1,\ f(1) = 4,\ f(2) = 11.$ Имеем

$$f(x) \approx \frac{x(x-1)(x-2)}{(-1)\cdot(-2)\cdot(-3)}\cdot 2 + \frac{(x+1)(x-1)(x-2)}{1\cdot(-1)\cdot(-2)}\cdot 1 +$$

$$+\frac{(x+1)\,x\,(x-2)}{2\cdot 1\cdot(-1)}\cdot 4+\frac{(x+1)\,x\,(x-1)}{3\cdot 2\cdot 1}\cdot 11=2x^2+x+1.$$

Как видим, степень интерполяционного многочлена может быть меньше n (в данном случае $n=3$).

§ 4. Линейное интерполирование. Пусть заданы только два значения x_1, x_2 аргумента и соответственные значения функции $f(x)$. Интерполяционный многочлен будет тогда многочленом первой степени, его графиком будет прямая, проходящая через точки $(x_1,\ f(x_1))$, $(x_2,\ f(x_2))$. Отсюда название — *линейное интерполирование.*

К линейному интерполированию прибегают довольно часто.

Существуют таблицы численных значений многих функций (например, таблицы логарифмов, таблицы значений тригонометрических функций). В этих таблицах последовательные значения аргумента отличаются на некоторую величину h. Например, в таблицах десятичных логарифмов даны логарифмы целых чисел, $h=1$. В таблицах тригонометрических величин $h=30'\left(=\frac{\pi}{360}\right)$ $\left(\text{или}\quad h=\frac{\pi}{720}\right)$. Между тем, иногда необходимо вычислить значение функции для значения аргумента, содержащегося между двумя рядом стоящими аргументами. Скажем, надо найти $\sin 32°42'$, а в таблицах есть синусы углов $32°30'$ и $32°45'$.

В подобной ситуации можно успешно применять линейное интерполирование.

Допустим, что некоторый стержень вмонтирован в установке так, что экспериментатору доступны только концы этого стержня.

Измерив величины некоторой физической характеристики, например температуры на концах стержня, экспериментатор в качестве первого приближения для нахождения значения характеристики в некоторой недоступной точке стержня вынужден воспользоваться линейным интерполированием.

При этом возникает вопрос о величине погрешности с которой находится значение функции по линейной интер поляционной формуле. В самом деле, если в таблицах даны логарифмы с точностью до 0,0001, то было бы бессмысленным вычислять логарифм с помощью интерполяционной формулы, дающей результат с точностью до 0,01.

Докажем, что погрешность линейной интерполяционной формулы для $f(x)$ допускает оценку

$$|\text{погр.}| < \frac{(x_2 - x_1)^2}{8} \max |f''(x)|. \tag{1}$$

Пусть многочлен

$$P_1(x) = K + Lx$$

удовлетворяет условиям интерполирования

$$P_1(x_1) = f(x_1), \ P_1(x_2) = f(x_2)$$

и пусть $\bar{x}$ — любое, но фиксированное, значение аргумента такое, что $x_1 < \bar{x} < x_2$. При $x = \bar{x}$, вообще говоря, $f(x) \neq \neq P_1(x)$, и поэтому

$$f(\bar{x}) = P_1(\bar{x}) + \text{погр.}$$

Погрешность будем записывать в виде

$$\text{погр.} = N(\bar{x} - x_1)(\bar{x} - x_2),$$

причем число N еще не определено, но если $\bar{x}$ окажется равным x_1 или x_2, то погрешность будет равна нулю.

Рассмотрим функцию

$$\varphi(x) = f(x) - P_1(x) - N(x - x_1)(x - x_2).$$

Ясно, что

$$\varphi(x_1) = 0, \ \varphi(\bar{x}) = 0, \ \varphi(x_2) = 0.$$

По теореме Ролля, получаем

$$\varphi'(c_1) = 0, \ x_1 < c_1 < \bar{x}; \ \varphi'(c_2) = 0, \ \bar{x} < c_2 < x_2.$$

Снова применяя теорему Ролля, получим

$$\varphi''(\bar{c}) = 0, \ c_1 < \bar{c} < c_2 \Rightarrow x_1 < \bar{c} < x_2.$$

Но

$$\varphi''(x) = f''(x) - 2N,$$

так как $x' = 1$, $x'' = (1)' = 0$. Значит,

$$0 = \varphi''(\bar{c}) = f''(\bar{c}) - 2N,$$

$$N = \frac{f''(\bar{c})}{2}, \ x_1 < \bar{c} < x_2.$$

Итак

$$\text{погр.} = \frac{f''(\bar{c})}{2}(x - x_1)(x - x_2).$$

Отсюда
$$|\text{погр.}| \leqslant \frac{1}{2}\max|f''(x)|\max(x-x_1)(x_2-x).$$
Так как
$$(x-x_1)(x_2-x) = -x^2+(x_2+x_1)x-x_1x_2 =$$
$$= -\left(x-\frac{x_1+x_2}{2}\right)^2+\left(\frac{x_2-x_1}{2}\right)^2,$$
то
$$\max(x-x_1)(x_2-x) = \frac{(x_2-x_1)^2}{4}.$$
Окончательно
$$|\text{погр.}| < \frac{(x_2-x_1)^2}{8}\max|f''(x)|.$$

Рассмотрим два важных частных случая.

1. $f(x)=\sin x,\ 0 \leqslant x_1,\ x_2 \leqslant \frac{\pi}{4},\ x_2-x_1=\frac{\pi}{720}$ $(=15')$.

Так как $(\sin x)''=-\sin x$, то получим
$$|\text{погр.}| < \frac{\pi^2}{8\cdot(720)^2} < \frac{10{,}24}{8\cdot 518400} < 0{,}00001.$$
Таким образом, с помощью линейной интерполяционной формулы можно вычислять (с точностью до 0,00001) значения синусов углов, содержащихся между углами, которые отличаются на 15′. Это обычно вполне достаточная точность.

Очевидно, такой же результат будем иметь для $\cos x$. Мы предполагали, что $x_1,\ x_2 \in \left[0,\ \frac{\pi}{4}\right]$ по той причине, что $\sin x = \cos\left(\frac{\pi}{2}-x\right)$, и поэтому в таблицах даются значения тригонометрических функций лишь для углов, содержащихся между 0 и 45°.

2. $f(x)=\lg x,\ x_1,\ x_2 > 100,\ x_2-x_1=1$.

В этом случае из (1) получим (так как $(\lg x)''=-\frac{1}{x^2}\lg e$, где $e=2{,}71\ldots$)
$$|\text{погр.}| < \frac{1}{8\cdot 100^2}\cdot\lg 2{,}8 < 0{,}00001.$$

Запишем теперь формулу линейного интерполирования и приведем ее к виду, удобному для пользования. Общая формула (10) предыдущего параграфа при $n=2$ будет иметь вид

$$f(x) \approx \frac{x-x_2}{x_1-x_2} f(x_1) + \frac{x-x_1}{x_2-x_1} f(x_2).$$

Отсюда

$$f(x) \approx \frac{(x-x_1)+(x_1-x_2)}{-(x_2-x_1)} f(x_1) + \frac{x-x_1}{x_2-x_1} f(x_2),$$

$$f(x) \approx f(x_1) + \frac{f(x_2)-f(x_1)}{x_2-x_1}(x-x_1).$$

Это и есть формула линейного интерполирования, имеющая простой геометрический смысл (рис. 17) и удобная для вычислений.

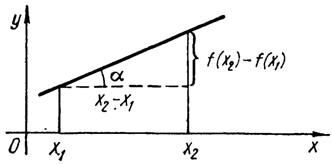

Рис. 17

§ 5. Интерполяционная формула Ньютона. Рассмотренная в § 3 формула Лагранжа пригодна для любого набора узлов интерполирования, но имеет два существенных недостатка. Первый из них заключается в том, что формула громоздка — каждое слагаемое формулы является многочленом n-й степени. Второй недостаток заключается в следующем. Случается, что сначала задается, допустим, m узлов и вычисляется интерполяционный многочлен Лагранжа. Затем по какой-то причине добавляются дополнительные узлы интерполирования (потому, например, что полученная интерполяционная формула оказалась недостаточно точной). Приходится тогда все вычисления выполнять снова — ни одно из слагаемых формулы Лагранжа не сохраняется.

Ньютон[1] нашел такую форму для интерполяционного многочлена, которая лишена указанных недостатков формулы Лагранжа, но эта форма проста лишь в том случае, когда заданные значения аргумента образуют арифметическую прогрессию. Чтобы написать формулу Ньютона, введем понятие о так называемых *конечных разностях функции*.

Пусть даны значения аргумента a, $a+h$, $a+2h$, $a+3h$, ... и соответствующие значения функции. Выпишем значения аргумента и функции двумя колонками. Затем

[1] Ньютон И. (1643—1727) — английский физик, механик, астроном и математик.

каждое число второй колонки, начиная с первого, вычтем из последующего числа. Получим первые разности функции и первую из них обозначим через $\Delta f(a)$. Найденные разности записываем в третьей колонке. Аналогично по первым разностям составляем вторые. Первую из вторых разностей обозначаем символом $\Delta^2 f(a)$. Подобно составляем третьи разности и т. д. Интерполяционная формула Ньютона имеет вид

$$f(x) \approx P(x) = f(a) + \frac{x-a}{h}\Delta f(a) + \frac{(x-a)(x-a-h)}{2!\,h^2} \times$$
$$\times \Delta^2 f(a) + \frac{(x-a)(x-a-h)(x-a-2h)}{3!\,h^3}\Delta^3 f(a) + \ldots .$$

Доказательство весьма элементарно и приводить его не будем.

Пример 1. Использовав данные примера 1, § 3, написать формулу Ньютона.

Имеем

x	$f(x)$	Δf	$\Delta^2 f$	$\Delta^3 f$
−1	6	−8	14	−12
0	−2	6	2	
1	4	8		
2	12			

Откуда

$$f(x) \approx 6 + \frac{x+1}{1}\cdot(-8) + \frac{(x+1)\,x}{2}\cdot 14 +$$
$$+ \frac{(x+1)\,x\,(x-1)}{6}\cdot(-12) = -2 + x + 7x^2 - 2x^3.$$

Пример 2. Использовав данные примера 2, § 3, написать формулу Ньютона.

Имеем

x	$f(x)$	Δf	$\Delta^2 f$	$\Delta^3 f$
−1	2	−1	4	0
0	1	3	4	
1	4	7		
2	11			

Откуда

$$f(x) \approx 2 + \frac{x+1}{1}\cdot(-1) + \frac{(x+1)\,x}{2}\cdot 4 = 1 + x + 2x^2.$$

Замечание. Формула линейного интерполирования в конечной ее форме представляет собой формулу Ньютона при $x_1 = a$, $x_2 = a + h$.

В предыдущих параграфах были рассмотрены две возможности:

1) даны $f(a)$, $f'(a)$, $f''(a)$, ... ;

2) даны $f(x_1)$, $f(x_2)$, $f(x_3)$, ...

Существуют интерполяционные формулы и для «смешанного» случая, когда даны, например, $f(x_1)$; $f(x_2)$, $f'(x_2)$, $f''(x_2)$; $f(x_3)$.

Такой случай встречается редко, поэтому его рассматривать не будем. В последующих параграфах рассмотрим очень важный для практики случай, когда даны $n+1$ значений функции, а мы хотим выразить приближенно нашу функцию в виде многочлена не степени n, а степени m, причем обязательно $m < n$.

§ 6. Избыточные системы линейных уравнений, решение по способу наименьших квадратов. Непосредственное измерение или вычисление величин x_1, x_2, ... , x_m во многих случаях затруднительно или даже невозможно, а допускающими непосредственные измерения оказываются некоторые выражения (функции), содержащие величины x_1, x_2, ... , x_m.

Простейшим случаем, к которому сводятся многие другие, является тот, когда в процессе измерения получаем результаты вида

$$\begin{aligned} a_1x_1 + b_1x_2 + \cdots + l_1x_m &= h_1, \\ a_2x_1 + b_2x_2 + \cdots + l_2x_m &= h_2, \\ \cdots\cdots\cdots\cdots\cdots&\cdots\cdots \\ a_nx_1 + b_nx_2 + \cdots + l_nx_m &= h_n, \end{aligned} \qquad (1)$$

т. е. для определения m неизвестных x_1, x_2, ... , x_m имеем систему n линейных уравнений. Число уравнений обычно больше числа неизвестных, так как экспериментатор, даже измеряя одну величину, не может ограничиться одним измерением. Дело в том, что отдельные измерения неизбежно содержат случайные ошибки и лишь на основании некоторой совокупности измерений экспериментатор может сделать более или менее обоснованные выводы. Например, если измеряя некоторую величину n раз, экспериментатор получит в результате величины α_1, α_2, ... , α_n, то полагается, что достаточно близким к истинному значению величины является среднее арифметическое:

$$\frac{\alpha_1 + \alpha_2 + \cdots + \alpha_n}{n}. \qquad (2)$$

Во многих случаях (но не всегда), что доказано в теории вероятностей и проверено на практике, гипотеза о достаточной близости величины (2) к истинному значению измеряемой величины верна.

Система (1), в которой число уравнений больше числа неизвестных, называется *избыточной*. Она противоречива, несовместна. Это означает, что нельзя найти такие значения неизвестных, которые точно удовлетворяют всем уравнениям системы. Следует иметь в виду, что коэффициенты уравнений, и прежде всего свободные члены, обычно не точны, и поэтому даже точное решение системы (1), если бы оно существовало, не давало бы точных значений интересующих нас величин.

Можно выбрать из (1) часть уравнений (m уравнений), решить их точно, но тогда, почти наверняка, подстановка найденных решений в оставшиеся $n-m$ уравнений приведет к совершенно несуразному результату.

Какие приближенные значения $x_1, x_2, \ldots, x_m$ надо считать наилучшими? По-видимому такие, чтобы одновременно, более или менее точно, удовлетворялись все уравнения системы (1).

Еще в начале XIX в. Лежандр предложил следующий принцип:

Из всех возможных значений неизвестных $x_1, x_2, \ldots$ $\ldots, x_m$ надо взять те, для которых сумма квадратов ошибок — отклонений левых частей уравнений от правых — является наименьшей.

Таким образом, требуется найти значения $x_1, x_2, \ldots$ $\ldots, x_m$, при которых величина

$$(a_1x_1+b_1x_2+\cdots+l_1x_m-h_1)^2+(a_2x_1+b_2x_2+\cdots+$$
$$+l_2x_m-h_2)^2+\cdots+(a_nx_1+b_nx_2+\cdots+l_nx_m-h_n)^2$$

имеет наименьшее значение (минимум).

Принцип Лежандра называется *принципом наименьших квадратов*, а решение системы (1) на основании этого принципа называется *решением по способу наименьших квадратов*. Естественно возникает вопрос обоснования принципа наименьших квадратов. Теоретически доказано и практически проверено, что в большинстве случаев принципу наименьших квадратов можно (и нужно) следовать. Однако считать его универсально пригодным нельзя.

Вопрос обоснования принципа наименьших квадратов по существу относится к теории вероятностей. Здесь рассмотрим только геометрический смысл принципа наименьших квадратов. Пусть дана система уравнений

$$
\begin{gathered}
a_1x + b_1y + c_1z = d_1, \\
a_2x + b_2y + c_2z = d_2, \\
\cdots\cdots\cdots\cdots \\
a_nx + b_ny + c_nz = d_n
\end{gathered} \tag{2}
$$

с тремя неизвестными. Кроме того, предположим, что выполнены условия

$$
a_i^2 + b_i^2 + c_i^2 = 1, \quad i = 1, 2, \ldots, n. \tag{3}
$$

Как известно (см. гл. I, § 10), каждое линейное уравнение (2) является уравнением плоскости. В аналитической геометрии весьма просто доказывается, что числа $|a_i\bar{x} + b_i\bar{y} + c_i\bar{z} - d_i|$ при условиях (3) равны расстояниям точки с декартовыми координатами $(\bar{x}, \bar{y}, \bar{z})$ от плоскостей (2). Следовательно, отыскать значения $\bar{x}$, $\bar{y}$, $\bar{z}$ из системы (1) при условиях (3) по способу наименьших квадратов означает найти точку, для которой минимально среднее квадратическое расстояние ее от заданных плоскостей (2).

Теперь докажем основное утверждение для $m = 3$ (для $m > 3$ рассуждения не сложнее).

Теорема. *Для того чтобы приближенное решение $\bar{x}$, $\bar{y}$, $\bar{z}$ избыточной системы уравнений*

$$
\begin{gathered}
a_1x + b_1y + c_1z = h_1, \\
a_2x + b_2y + c_2z = h_2, \\
\cdots\cdots\cdots\cdots \\
a_nx + b_ny + c_nz = h_n
\end{gathered} \tag{4}
$$

было наилучшим в смысле принципа наименьших квадратов, необходимо и достаточно, чтобы $\bar{x}$, $\bar{y}$, $\bar{z}$ было решением системы трех уравнений

$$
\begin{gathered}
\sum_{i=1}^{n} a_i^2x + \sum_{i=1}^{n} a_ib_iy + \sum_{i=1}^{n} a_ic_iz = \sum_{i=1}^{n} a_ih_i, \\
\sum_{i=1}^{n} a_ib_ix + \sum_{i=1}^{n} b_i^2y + \sum_{i=1}^{n} b_ic_iz = \sum_{i=1}^{n} b_ih_i, \\
\sum_{i=1}^{n} c_ia_ix + \sum_{i=1}^{n} c_ib_iy + \sum_{i=1}^{n} c_i^2z = \sum_{i=1}^{n} c_ih_i.
\end{gathered} \tag{5}
$$

Система (5) называется *нормальной системой* уравнений, соответствующей избыточной системе (4).

Надо доказать, что величина

$$R = (a_1x + b_1y + c_1z - h_1)^2 + (a_2x + b_2y + c_2z - \\ - h_2)^2 + \cdots + (a_nx + b_ny + c_nz - h_n)^2 \qquad (6)$$

принимает наименьшее значение при тех и только тех x, y, z, которые образуют решение системы (5). Легко понять, что наибольшего значения R иметь не может, так как эта величина бесконечно возрастает, например, при $x \to \infty$.

Существование минимума R не вызывает сомнений, так как $R \geqslant 0$. Требуется только найти те значения $\bar{x}$, $\bar{y}$, $\bar{z}$ аргументов, при которых R достигает минимума. Записав, что

$$R_{\min} = (a_1\bar{x} + b_1\bar{y} + c_1\bar{z} - h_1)^2 + \cdots + (a_n\bar{x} + \\ + b_n\bar{y} + c_n\bar{z} - h_n)^2, \qquad (7)$$

рассмотрим функцию аргумента t:

$$\varphi(t) = (a_1(\bar{x} + t) + b_1\bar{y} + c_1\bar{z} - h_1)^2 + (a_2(\bar{x} + t) + \\ + b_2\bar{y} + c_2\bar{z} - h_2)^2 + (a_3(\bar{x} + t) + b_3\bar{y} + c_3\bar{z} - h_3)^2 + \ldots.$$

Эта функция, согласно равенству (7), должна иметь минимум при $t = 0$. Однако легко подсчитать, что

$$\varphi(t) = t^2 \sum_{i=1}^{n} a_i^2 + 2t \sum_{i=1}^{n} a_i(a_i\bar{x} + b_i\bar{y} + c_i\bar{z} - h_i) + \\ + \sum_{i=1}^{n} (a_i\bar{x} + b_i\bar{y} + c_i\bar{z} - h_i)^2,$$

т. е. это квадратный трехчлен относительно t. Как уже было установлено (см. гл. I, § 1), квадратный трехчлен, $at^2 + 2bt + c$, $a > 0$, имеет минимум при $t = -\frac{b}{a}$. В данном случае квадратный трехчлен имеет минимум при $t = 0$. Значит,

$$b = \sum_{i=1}^{n} a_i(a_i\bar{x} + b_i\bar{y} + c_i\bar{z} - h_i) = 0,$$

а это первое из уравнений (5). Второе из этих уравнений получим, рассмотрев

$$\psi(t) = (a_1\bar{x} + b_1(\bar{y} + t) + c_1\bar{z} - h_1)^2 + \cdots + (a_n\bar{x} + \\ + b_n(\bar{y} + t) + c_n\bar{z} - h_n)^2.$$

Аналогично находим и третье из уравнений (5).

Доказав теорему, заметим, что нормальную систему (5) можно составить по избыточной системе (4) следующим образом: умножая уравнения (4) соответственно на коэффициенты при x и складывая, получаем первое уравнение нормальной системы (5); умножая на коэффициенты при y и складывая, получаем второе уравнение системы (5) и т. д.

Пример 1. Решить систему уравнений

$$2x+y=1,$$
$$x-2y=2,$$
$$3x+2y=3.$$

Умножая уравнения соответственно один раз на 2, 1, 3, а второй — на 1, —2, 2 и складывая результаты, получаем нормальную систему

$$14x+6y=13,$$
$$6x+9y=3,$$

решением которой является $x=1{,}1$; $y=-0{,}4$.

Пример 2. Решить систему уравнений

$$x-y=1,$$
$$2x+y=2,$$
$$x+3y=1.$$

Нормальная система

$$6x+4y=6,$$
$$4x+11y=4$$

имеет решение $x=1$, $y=0$, которое является точным решением исходной системы. Значит, данная система совместна, а это означает, что одно из уравнений является следствием остальных (если умножить первое уравнение на $-5/3$, второе на $4/3$ и сложить результаты, то получится третье уравнение).

Пример 3. Решить систему

$$x-y=1,$$
$$2x+y=2,$$
$$2x+y=3.$$

Здесь третье уравнение явно противоречит второму, но вполне может быть, что, измеряя величину $u=2x+y$, получим два резко отличных результата, но нет оснований предпочесть один результат другому. Нормальная система

$$9x+3y=11,$$
$$3x+3y=4$$

имеет решение $x=7/6$, $y=1/6$. Подстановка в заданную систему показывает, что решение является точным для системы уравнений

$$x - y = 1,$$
$$2x + y = 2{,}5$$

(2,5 это среднее арифметическое чисел 2 и 3).

В общем виде, если перейти от системы уравнений

$$ax + by = h,$$
$$Ax + By = H,$$
$$Ax + By = K$$

к нормальной системе, получим

$$(a^2 + 2A^2)\,x + (ab + 2AB)\,y = ah + A(H + K),$$
$$(ab + 2AB)\,x + (b^2 + 2B)\,y = bh + B\,(H + K).$$

Решение этой системы

$$x = \frac{2Bh - b\,(H + K)}{2\,(aB - Ab)}, \quad y = \frac{a\,(H + K) - 2Ah}{2\,(aB - Ab)}$$

является точным решением системы уравнений

$$ax + by = h,$$
$$Ax + By = \frac{H + K}{2}.$$

Полученный результат можно обобщить (см. упражнение 9 в конце главы).

§ 7. Решение нормальной системы.

§ 7. Решение нормальной системы. В примерах предыдущего параграфа рассматривались системы с двумя неизвестными и коэффициенты при неизвестных были небольшими целыми числами. В практических задачах число неизвестных в основном больше двух, а коэффициенты записываются с помощью нескольких значущих цифр. Для решения таких систем достаточно воспользоваться методом подстановки (это заметил еще Гаусс). Надо лишь удачно ввести обозначения и подметить некоторое простое правило. Так, рассмотрим систему двух уравнений

$$\begin{aligned} au + bv + cw + \cdots &= \alpha, \\ Au + Bv + Cw + \cdots &= \beta. \end{aligned} \qquad (1)$$

Разделив обе части первого из уравнений на a, получим

$$u + \frac{b}{a}\,v + \frac{c}{a}\,w + \cdots = \frac{\alpha}{a}. \qquad (2)$$

Из этого и второго из уравнений (1) исключим u, для чего умножим уравнение (2) на A и результат вычтем из второго уравнения системы:

$$\left(B - \frac{bA}{a}\right)v + \left(C - \frac{cA}{a}\right)w + \cdots = \left(\beta - \frac{\alpha A}{a}\right).$$

Умножив обе части полученного уравнения на a и разделив на $Ba - bA$, находим

$$v + \frac{aC - cA}{aB - bA} w + \cdots = \frac{\beta a - \alpha A}{aB - bA}.$$

Та им образом, систему (1) заменяем эквивалентной системой

$$\begin{aligned} u + \frac{b}{a} v + \frac{c}{a} w + \cdots &= \frac{\alpha}{a}, \\ v + \frac{aC - cA}{aB - bA} w + \cdots &= \frac{\beta a - \alpha A}{aB - bA}. \end{aligned} \tag{3}$$

Чтобы автоматизировать процесс перехода от системы (1) к системе (3), достаточно запомнить правило составления общего знаменателя по двум первым столбцам коэффициентов системы уравнений (1) и заметить, что числители получаем по тому же правилу, если, по очереди, заменить второй столбец коэффициентов последующими столбцами.

Вернемся теперь к рассмотрению системы уравнений

$$\begin{aligned} a_1 x + b_1 y + c_1 z &= h_1, \\ a_2 x + b_2 y + c_2 z &= h_2, \\ &\cdots\cdots\cdots \\ a_n x + b_n y + c_n z &= h_n \end{aligned} \tag{4}$$

и соответствующей нормальной системы уравнений

$$\begin{aligned} \sum_{i=1}^{n} a_i^2 x + \sum_{i=1}^{n} a_i b_i y + \sum_{i=1}^{n} a_i c_i z &= \sum_{i=1}^{n} a_i h_i, \\ \sum_{i=1}^{n} b_i a_i x + \sum_{i=1}^{n} b_i^2 y + \sum_{i=1}^{n} b_i c_i z &= \sum_{i=1}^{n} b_i h_i, \\ \sum_{i=1}^{n} c_i a_i x + \sum_{i=1}^{n} c_i b_i y + \sum_{i=1}^{n} c_i^2 z &= \sum_{i=1}^{n} c_i h_i. \end{aligned} \tag{5}$$

Очевидно, коэффициенты системы (5) получаем как скалярное произведение столбцов коэффициентов системы (4), воспользовавшись при этом известными обозначениями.

Воспользуемся обозначением Гаусса

$$\sum_{i=1}^{n} \alpha_i \beta_i = [\alpha\beta]$$

и запишем систему (5) в таком виде:

$$
\begin{aligned}
&[aa]x + [ab]y + [ac]z = [ah],\\
&[ba]x + [bb]y + [bc]z = [bh],\\
&[ca]x + [cb]y + [cc]z = [ch].
\end{aligned}
\tag{6}
$$

Разделим первое уравнение на $[aa]$ и исключим x из остальных двух уравнений. Воспользовавшись правилом, полученным выше (см. систему (3)), найдем

$$
\begin{gathered}
x + \frac{[ab]}{[aa]}y + \frac{[ac]}{[aa]}z = \frac{[ah]}{[aa]},\\
([aa][bb] - [ab]^2)\,y + ([aa][bc] - [ab][ac])\,z =\\
= [aa][bh] - [ab][ah],\\
([aa][bc] - [ab][ac])\,y + ([aa][cc] - [ac]^2)\,z =\\
= [aa][ch] - [ah][ac].
\end{gathered}
\tag{7}
$$

Если введем обозначения

$$[aa][bb] - [ab]^2 = [bb1],\quad [aa][bc] - [ab][ac] = [bc1],\ \ldots,$$

то систему (7) перепишем в виде

$$
\begin{aligned}
x + \frac{[ab]}{[aa]}y + \frac{[ac]}{[aa]}z &= \frac{[ah]}{[aa]},\\
[bb1]\,y + [bc1]\,z &= [bh1],\\
[bc1]\,y + [cc1]\,z &= [ch1].
\end{aligned}
\tag{8}
$$

Разделив второе уравнение на $[bb1]$ и исключив y из третьего уравнения, получим

$$
\begin{aligned}
x + \frac{[ab]}{[aa]}y + \frac{[ac]}{[aa]}z &= \frac{[ah]}{[aa]},\\
y + \frac{[bc1]}{[bb1]}z &= \frac{[bh1]}{[bb1]},\\
z &= \frac{[ch2]}{[cc2]},
\end{aligned}
\tag{9}
$$

где $[cc2] = [cc1][bb1] - [bc1]^2$, $[ch2] = [bb1][ch1] - [bc1][bh1]$.

Нормальная система, приведенная к виду (9), называется *элиминационной* (элиминирование — исключение). Из нее последовательно находим x, y, z.

Замечание 1. Можно воспользоваться тем, что таблицы коэффициентов нормальной системы (6) и промежуточной, состоящей из двух последних уравнений (8), симметричны относительно диагонали таблицы. Опытный вычислитель выписывает каждый раз только таблицы коэффициентов и лишь элиминационную систему.

Замечание 2. Если иметь в виду только решение нормальной системы, то вовсе не обязательно приводить ее к виду (9). Иногда удобней, например, преобразовать нормальную систему к виду

$$\begin{aligned} z + \cdots \quad &= \cdots, \\ x + \cdots &= \cdots, \\ y &= \cdots, \end{aligned}$$

т. е. исключать неизвестные можно в любом порядке. Строго придерживаться схемы Гаусса следует при решении нескольких систем уравнений, если при этом используется ЭВМ.

Замечание 3. Пусть $\bar{x}$, $\bar{y}$, $\bar{z}$ — решение нормальной системы. Введем обозначения

$$\varepsilon_1 = a_1\bar{x} + b_1\bar{y} + c_1\bar{z} - h_1,$$
$$\dots\dots\dots\dots\dots$$
$$\varepsilon_n = a_n\bar{x} + b_n\bar{y} + c_n\bar{z} - h_n$$

для погрешностей — отклонений левых частей уравнений от правых — при подстановке решения нормальной системы в исходную. Точность решения исходной системы естественно характеризовать величиной

$$[\varepsilon\varepsilon] = \sum_{i=1}^{n} (a_i\bar{x} + b_i\bar{y} + c_i\bar{z} - h_i)^2.$$

Достаточно простые вычисления приводят к следующей формуле Гаусса:

$$[\varepsilon\varepsilon] = [hh] - \left(\frac{[ah]^2}{[aa]} + \frac{[bh1]^2}{[bb1]} + \frac{[ch2]^2}{[cc2]}\right). \tag{10}$$

Рекомендуем читателю сопоставить формулу Гаусса с элиминационной системой (9).

§ 8. Параболическое интерполирование по способу наименьших квадратов. Формулы П. Л. Чебышева.

Графики многочленов

$$y = A_0 + A_1x + A_2x^2 + \cdots + A_mx^m$$

называются *параболическими кривыми* или точнее *параболами* m-й степени. При $m = 2$ имеем известную параболу.

Формы парабол, в зависимости от подбора коэффициентов A_0, A_1, ..., A_m, весьма разнообразны. Поэтому, имея значения $f(x_1)$, $f(x_2)$, ..., $f(x_{n+1})$ некоторой функции $f(x)$, естественно воспользоваться приближенной формулой вида

$$f(x) \approx P_m(x) = A_0 + A_1x + A_2x^2 + \cdots + A_mx^m. \tag{1}$$

Естественно найти коэффициенты A_0, A_1, ..., A_m так, чтобы выполнялись равенства

$$P_m(x_i) = f(x_i), \quad i = 1, 2, \dots, n+1. \tag{2}$$

Случай $m=n$ уже рассматривался в главе III, где для интерполяционного многочлена $P_n(x)$ были приведены формулы Лагранжа и Ньютона. Теперь рассмотрим случай, когда $m<n$. Этот случай особо важен, поскольку число значений функции $f(x)$, вычисляемых или измеряемых, лучше иметь большим (информация о функции тем надежней, чем она полней), а интерполирующий многочлен должен, по возможности, иметь не слишком высокую степень.

Запишем условия (2) в виде

$$\begin{aligned} &A_0+A_1x_1+A_2x_1^2+\cdots+A_mx_1^m=f(x_1),\\ &A_0+A_1x_2+A_2x_2^2+\cdots+A_mx_2^m=f(x_2),\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ &A_0+A_1x_{n+1}+A_2x_{n+1}^2+\cdots+A_mx_{n+1}^m=f(x_{n+1}). \end{aligned} \tag{3}$$

Для определения коэффициентов A_0, A_1, A_2, ... имеем избыточную систему линейных уравнений, поэтому можно говорить лишь о приближенном ее решении.

К системе линейных уравнений (3) применим принцип наименьших квадратов, т. е. коэффициенты A_0, A_1, A_2, ... будем искать такими, чтобы величина

$$\sum_{i=1}^{n+1}\left(f(x_i)-(A_0+A_1x_i+A_2x_i^2+\cdots+A_mx_i^m)\right)^2 \tag{4}$$

приняла свое минимальное значение. Для решения этой задачи можно применить схему Гаусса, изложенную в предыдущем параграфе, но возможен и другой подход. Допустим, что даны многочлены $\psi_0(x)$, $\psi_1(x)$, $\psi_2(x)$, ... степеней соответственно 0, 1, 2, ..., образующие ортонормированную систему на множестве $\{x_1, x_2, \ldots, x_{n+1}\}$. Сгруппировав соответственно слагаемые в многочлене $P_m(x)$ или подставив вместо 1, x, x^2, ... их выражения через $\psi_0(x)$, $\psi_1(x)$, $\psi_2(x)$, ..., запишем $P_m(x)$ в виде

$$P_m(x)=B_0\psi_0(x)+B_1\psi_1(x)+B_2\psi_2(x)+\cdots+B_m\psi_m(x),$$

а условие минимизации выражения (4) станет условием минимизации выражения

$$\begin{aligned} &\langle f(x)-P_m(x),\ f(x)-P_m(x)\rangle=\\ &=\sum_{i=1}^{n+1}\left(f(x_i)-\sum_{r=0}^{m}B_r\psi_r(x)\right)^2. \end{aligned} \tag{5}$$

Такая задача была рассмотрена в гл. II, § 5, и сейчас воспользуемся указанным там решением.

Как было показано, выражение (5) принимает минимальное значение, когда коэффициенты B_r являются коэффициентами Эйлера — Фурье, т. е.

$$B_r = \langle f(x),\ \psi_r(x)\rangle = \sum_{i=1}^{n+1} f(x_i)\,\psi_r(x_i). \tag{6}$$

Таким образом, формула интерполирования имеет вид

$$f(x) \approx P_m(x) = \langle f(x),\ \psi_0(x)\rangle\,\psi_0(x) + \\ + \langle f(x),\ \psi_1(x)\rangle\,\psi_1(x) + \cdots + \langle f(x),\ \psi_m(x)\rangle\,\psi_m(x). \tag{7}$$

Кроме того, минимальным значением величины (5), т. е. квадратом среднего квадратического отклонения $f(x)$ от $P_m(x)$, будет (см. гл. II)

$$\rho_m^2 = \langle f(x),\ f(x)\rangle - \sum_{i=0}^{m} (\langle f_m(x),\ \psi_i(x)\rangle)^2. \tag{8}$$

Если имеем не ортонормированную систему многочленов $\psi_0(x)$, $\psi_1(x)$, ..., $\psi_m(x)$, а ортогональную $\varphi_0(x)$, $\varphi_1(x)$, ..., $\varphi_m(x)$, то равенства (7) и (8) будут иметь другой вид. В самом деле, нормируя, получим

$$\psi_r(x) = \frac{\varphi_r(x)}{\sqrt{\langle \varphi_r(x),\ \varphi_r(x)\rangle}}.$$

Далее

$$B_r = \left\langle f(x),\ \frac{\varphi_r(x)}{\sqrt{\langle \varphi_r(x),\ \varphi_r(x)\rangle}}\right\rangle = \\ = \frac{1}{\sqrt{\langle \varphi_r(x),\ \varphi_r(x)\rangle}}\langle f(x),\ \varphi_r(x)\rangle.$$

Поэтому равенство (7) перепишем в виде

$$f(x) \approx P_m(x) = \frac{\langle f(x),\ \varphi_0(x)\rangle}{\langle \varphi_0(x),\ \varphi_0(x)\rangle}\varphi_0(x) + \\ + \frac{\langle f(x),\ \varphi_1(x)\rangle}{\langle \varphi_1(x),\ \varphi_1(x)\rangle}\varphi_1(x) + \cdots + \frac{\langle f(x),\ \varphi_m(x)\rangle}{\langle \varphi_m(x),\ \varphi_m(x)\rangle}\varphi_m(x), \tag{7'}$$

а равенство (8) — в виде

$$\rho_m^2 = \langle f(x),\ f(x)\rangle - \sum_{i=1}^{m} \frac{(\langle f(x),\ \varphi_i(x)\rangle)^2}{\langle \varphi_i(x),\ \varphi_i(x)\rangle}. \tag{8'}$$

Замечание 1. Обычно на практике приходится сначала искать интерполяционный многочлен невысокой степени. Если требуется увеличить степень многочлена, то дописывают одно или больше слагаемых в формулах (7), (8) или в формулах (7′) и (8′).

Замечание 2. Если не указана ортонормированная или только ортогональная система, то (см. гл. II, § 4) строим такую последовательность методом Сонина — Шмидта, используя последовательность 1, x, x^2, ... При этом многочлены находим последовательно один за другим. Например, для построения $P_3(x)$ ($m = 3$) заканчиваем процесс Сонина — Шмидта, как только получаем $\varphi_0(x)$, $\varphi_1(x)$, $\varphi_2(x)$, $\varphi_3(x)$.

Замечание 3. Произвести ортогонализацию последовательности 1, x, x^2, ... не сложно, но иногда утомительно (см., например, гл. II, § 4, пример 1).

Для выявления основной идеи П. Л. Чебышева и получения некоторого удобного способа построения ортогональных многочленов воспользуемся схемой Гаусса, изложенной в предыдущем параграфе.

Для удобства положим, что

$$a = 1,\ \ a_i = 1;\ \ b = x,\ \ b_i = x_i;\ \ c = x^2,\ \ c_i = x_i^2;$$
$$A_0 = X,\ \ A_1 = Y,\ \ A_2 = Z,\ \ f(x_i) = h,$$

и примем $m = 2$. Тогда найдем многочлен

$$P(x) = aX + bY + cZ,$$

удовлетворяющий условиям

$$\begin{aligned} a_1X + b_1Y + c_1Z &= h_1, \\ a_2X + b_2Y + c_2Z &= h_2, \\ &\ldots\ldots\ldots \\ a_nX + b_nY + c_nZ &= h_n. \end{aligned}$$

Эта система аналогична системе (4) предыдущего параграфа и поэтому, применяя принцип наименьших квадратов, мы можем сразу написать элиминационную систему уравнений:

$$\begin{aligned} X + \frac{[ab]}{[aa]}Y + \frac{[ac]}{[aa]}Z &= \frac{[ah]}{[aa]}, \\ Y + \frac{[bc1]}{[bb1]}Z &= \frac{[bh1]}{[bb1]}, \\ Z &= \frac{[ch2]}{[cc2]}. \end{aligned} \tag{9}$$

П. Л. Чебышев, никогда не забывавший о практике, конечно не мог не заметить большое неудобство системы (9). Эта система предполагает обязательно, что каждое ее уравнение выписывается полностью, так как решать систему надо с конца. Составление коэффициентов системы (9) достаточно сложное, требует использования сразу всех a_i, b_i, ... Если бы мы были уверены, что не придется увеличивать степень интерполяционного многочлена, то отмеченные недостатки системы (9) нас бы не беспокоили. Однако на практике естественно сначала искать интерполяционный многочлен невысокой степени, затем, в случае необходимости, увеличить степень на единицу и т. д. В случае увеличения степени многочлена всю систему (9) надо составлять заново. Вот этого, следуя П. Л. Чебышеву, мы хотим избежать.

Уравнения (9) определяют неизвестные X, Y, Z как линейные функции правых частей этих уравнений. Значит, если обозначим правые части уравнений (9) соответственно через K_0, K_1, K_2, найдем X, Y, Z и полученные значения подставим в $P(x)$, то получим

$$P(x) = K_0\varphi_0(x) + K_1\varphi_1(x) + K_2\varphi_2(x), \tag{10}$$

где $\varphi_0(x)$, $\varphi_1(x)$, $\varphi_2(x)$ требуется определить.

Умножив соответственно уравнения (9) на $\varphi_0(x)$, $\varphi_1(x)$, $\varphi_2(x)$ и сложив результаты, получим в правой части многочлен (10). Следовательно,

$$X\varphi_0(x) + Y\left(\frac{[ab]}{[aa]}\varphi_0(x) + \varphi_1(x)\right) + Z\left(\frac{[ac]}{[aa]}\varphi_0(x) + \right.$$
$$\left. + \frac{[bc1]}{[bb1]}\varphi_1(x) + \varphi_2(x)\right) = P(x).$$

Но $P(x) = aX + bY + cZ$. Так как X, Y, Z могут определяться только однозначно, то

$$\begin{gathered} \varphi_0(x) = a = 1, \\ \frac{[ab]}{[aa]}\varphi_0(x) + \varphi_1(x) = b = x, \\ \frac{[ac]}{[aa]}\varphi_0(x) + \frac{[bc1]}{[bb1]}\varphi_1(x) + \varphi_2(x) = c = x^2, \end{gathered} \tag{11}$$

откуда и найдем $\varphi_0(x)$, $\varphi_1(x)$, $\varphi_2(x)$. Они, очевидно, будут многочленами степени 0, 1, 2.

Итак, если написать только первое уравнение системы (11), то найдем $\varphi_0(x)$ и многочлен

$$P_0(x) = K_0\varphi_0(x) = \frac{[ah]}{[aa]}\varphi_0(x). \tag{12}$$

Написав первых два уравнения (11), найдем $\varphi_0(x)$, $\varphi_1(x)$ и

$$P_1(x) = K_0\varphi_0(x) + K_1\varphi_1(x) = \frac{[ah]}{[aa]}\varphi_0(x) + \frac{[bh1]}{[bb1]}\varphi_1(x). \tag{13}$$

Используя все три уравнения (11), получим

$$P_2(x) = K_0\varphi_0(x) + K_1\varphi_1(x) + K_2\varphi_2(x) =$$
$$= P_1(x) + \frac{[ch2]}{[cc2]}\varphi_2(x). \qquad (14)$$

Равенства (11) и (12) — (14) называются *формулами Чебышева.*

Следовательно, имея первых два столбца элиминационной системы и столбец правых ее частей, можем написать первые два уравнения (11), а имея три столбца элиминационной системы и столбец свободных членов, можем написать три уравнения (11).

Пример. Пусть даны значения аргумента: 0, 1, 2, 3 и соответствующие значения функции: 1, —1, 3, 4. Требуется найти интерполяционные многочлены первой и второй степени.

Положив

$$P(x) = X \cdot 1 + Yx + Zx^2,$$

получим для определения X, Y, Z избыточную систему уравнений

$$X = 1,$$
$$X + Y + Z = -1,$$
$$X + 2Y + 4Z = 3,$$
$$X + 3Y + 9Z = 4.$$

Ей соответствует нормальная система

$$\begin{aligned} 4X + 6Y + 14Z &= 7,\\ 6X + 14Y + 36Z &= 17,\\ 14X + 36Y + 98Z &= 47. \end{aligned} \qquad (15)$$

Используя нормальную систему (15), запишем

$$X + \frac{6}{4}Y + \ldots = \frac{7}{4}, \qquad (16)$$
$$Y + \ldots = \frac{26}{20}.$$

Второе равенство получаем в результате исключения X из второго уравнения (15). Это уже дает возможность найти

$$\varphi_0(x) = 1,$$
$$\frac{6}{4} \cdot 1 + \varphi_1(x) = x, \quad \varphi_1(x) = x - 1{,}5, \qquad (17)$$

$$P_1(x) = \frac{7}{4} + \frac{26}{20}(x - 1{,}5) = \frac{13}{10}x - \frac{1}{5}.$$

Если понадобится повышение степени многочлена, то дополняем равенства (16):

$$X+\frac{6}{4}Y+\frac{14}{4}Z=\frac{7}{4},$$
$$Y+\frac{60}{20}Z=\frac{26}{20},$$
$$Z=\frac{3}{4}$$

и составляем уравнение для нахождения $\varphi_2(x)$:

$$\frac{14}{4}\cdot 1+\frac{60}{20}(x-1{,}5)+\varphi_2(x)=x^2.$$

Отсюда,

$$\varphi_2(x)=x^2-3x+1,$$

$$P_2(x)=\frac{7}{4}+\frac{26}{20}(x-1{,}5)+\frac{3}{4}(x^2-3x+1)=$$
$$=\frac{3}{4}x^2-\frac{19}{20}x+\frac{11}{20}. \qquad (18)$$

Замечание 1. В данном случае умышленно взяты те же значения аргумента, что и в гл. II, § 4, пример 1. Найденные многочлены те же самые, что и в упомянутом примере (только обозначения изменены).

Замечание 2. П. Л. Чебышев получил формулы вида (12) — (14), пользуясь так называемыми непрерывными дробями.

§ 9. Эмпирические линии регрессии. В статистических исследованиях часто возникает следующий вопрос: наблюдения (измерения) дают n пар

$$(x_1,\ y_1),\ (x_2,\ y_2),\ \ldots,\ (x_n,\ y_n)$$

значений величин x, y. Эти величины могут быть случайными, но связаны они между собой некоторой зависимостью. Например, количество ягнят y связано с количеством овцематок x, хотя строгой зависимости между x и y нет. Наличие случайных факторов видно уже из того, что у одного и того же чабана на сотню овцематок каждый год приходится разное количество ягнят.

Простейшее предположение заключается в том, что

$$y=kx+b. \qquad (1)$$

График уравнения (1) называется *линией регрессии*.

Для определения коэффициентов k, b имеем избыточную систему уравнений

$$y_i = kx_i + b, \quad i = 1, 2, \ldots, n,$$

которую решаем по способу наименьших квадратов с применением ортогональных многочленов. Легко проверить, что ортогональны многочлены 1, $x - M_x$, где

$$M_x = \frac{x_1 + x_2 + \cdots + x_n}{n}. \tag{2}$$

Действительно,

$$\langle 1, x - M_x \rangle = \sum_{i=1}^{n} (x_i - M_x) = \sum_{i=1}^{n} x_i - nM_x = 0.$$

Для применения формулы (7'), § 8, находим

$$\langle 1, 1 \rangle = n, \quad \langle y, 1 \rangle = \sum_{i=1}^{n} y_i = nM_y,$$

$$\langle x - M_x, x - M_x \rangle = \sum_{i=1}^{n} (x_i - M_x)^2 = n\sigma_x^2,$$

где

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n} (x_i - M_x)^2}{n}}, \tag{3}$$

$$\langle y, x - M_x \rangle = \sum_{i=1}^{n} y_i (x_i - M_x) = \sum_{i=1}^{n} (x_i - M_x)(y_i - M_y),$$

так как

$$\sum_{i=1}^{n} M_y (x_i - M_x) = M_y \sum_{i=1}^{n} (x_i - M_x) = 0.$$

По формуле (7'), § 8, имеем

$$y = \frac{nM_y}{n} \cdot 1 + \frac{\sum_{i=1}^{n} (x_i - M_x)(y_i - M_y)}{n\sigma_x^2} (x - M_x),$$

$$y - M_y = \frac{\sum_{i=1}^{n} (x_i - M_x)(y_i - M_y)}{n\sigma_x^2} (x - M_x). \tag{4}$$

Величины M_x, M_y называются *выборочными средними*, величина σ_x^2 — *выборочной дисперсией*, угловой коэффициент

$$m = \frac{\sum_{i=1}^{n} (x_i - M_x)(y_i - M_y)}{n\sigma_x^2}$$

называется *выборочным коэффициентом регрессии*, прямая (4) — *эмпирической линией регрессии*.

Обратим внимание на то, что прямая (4) проходит через точку (M_x, M_y).

Конечно, предположение (1) не единственно возможное. Можно, например, ввести квадратичную (параболическую) регрессию

$$y = ax^2 + bx + c.$$

Если точки (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) расположены достаточно близко к эмпирической линии регрессии, то можно полагать, что предположение (1) оправдано.

§ 10. О приближенном решении нормальной системы. Способ Гаусса решения нормальной системы уравнения применяется широко. Он используется при решении любой системы n уравнений первой степени с n неизвестными. С помощью этого способа получаем формулы П. Л. Чебышева и приходим к уравнениям для нахождения многочленов ортогональных на данном конечном множестве значений аргумента. Наконец, получаем полезную формулу Гаусса (см. § 6, формула (10)). Следует заметить, однако, что способ Гаусса не использует важную особенность таблицы (матрицы) коэффициентов нормальной системы уравнений — ее симметричность. Эта симметричность упрощает решение нормальной системы, особенно при решении системы с помощью определителей. Если число уравнений велико, то вычисления и с помощью определителей очень сложны.

В том случае, когда коэффициенты $[aa]$, $[bb]$, ..., находящиеся в таблице коэффициентов нормальной системы на диагонали, больше остальных коэффициентов, а это бывает очень часто, удобен такой приближенный способ решения нормальной системы.

В качестве первого приближенного решения нормальной системы берется то, которое получаем, если в левых частях уравнений оставить только члены, стоящие по диагонали.

Имеем

$$x_1 = \frac{[ah]}{[aa]}, \quad y_1 = \frac{[bh]}{[bb]}, \ \ldots .$$

Подставив эти значения в первое уравнение системы, находим его погрешность $[aa]x_1 + [ab]y_1 + \cdots - [ah] = \alpha_1$, и полагаем

$$x_2 = x_1 - \frac{\alpha_1}{[aa]}.$$

Подставив во второе уравнение системы $x_2, y_1, z_1, \ldots$, находим

$$\alpha_2 = [ba]x_2 + [bb]y_1 + [bc]z_1 + \cdots - [bh].$$

Затем полагаем

$$y_2 = y_1 - \frac{\alpha_2}{[bb]}.$$

Дойдя таким путем до второго приближения последнего из неизвестных, повторяем весь процесс, исходя уже из приближенного решения $x_2, y_2, z_2, \ldots$. В случае необходимости процесс повторяем.

Пример. Решить систему уравнений

$$\begin{aligned} 27x + 6y &= 88, \\ 6x + 15y + z &= 70, \\ y + 54z &= 107. \end{aligned} \qquad (1)$$

Находим первое приближенное решение нормальной системы:

$$x_1 = \frac{88}{27} \approx 3, \quad y_1 = \frac{70}{15} \approx 5, \quad z_1 = \frac{107}{54} \approx 2.$$

Далее

$$\alpha_1 = 27 \cdot 3 + 6 \cdot 5 - 88 = 23,$$

$$x_2 = 3 - \frac{23}{27} \approx 2{,}15;$$

$$\alpha_2 = 6 \cdot 2{,}15 + 15 \cdot 5 + 2 - 70 = 19{,}9,$$

$$y_2 = 5 - \frac{19{,}9}{15} \approx 3{,}67;$$

$$\alpha_3 = 3{,}67 + 54 \cdot 2 - 107 = 4{,}67$$

$$z_2 = 2 - \frac{4{,}67}{54} = 1{,}914.$$

Итак,

$$x_2 = 2{,}15, \quad y_2 = 3{,}67, \quad z_2 = 1{,}914.$$

Теперь процесс повторяем

$$\beta_1 = 27 \cdot 2{,}15 + 6 \cdot 3{,}67 - 88 = -7{,}93,$$

$$x_3 = 2{,}15 + \frac{7{,}93}{27} = 2{,}446.$$

Легко проверить, что

$$y_3 = 3{,}561, \quad z_3 = 1{,}915.$$

Точное решение системы (1) следующее:

$$x = 2{,}470, \quad y = 3{,}551, \quad z = 1{,}916.$$

§ 11. Об избыточной системе уравнений не первой степени. Было бы ошибкой полагать, что всегда можно измерить значения линейных функций искомых величин, которые непосредственно не допускают измерения. Поэтому возникает вопрос о приближенном решении избыточной системы n уравнений не первой степени с m неизвестными ($n > m$). Прежде всего попытаемся найти какое-нибудь приближенное решение. Его можно подобрать, исходя из «физических» соображений, путем нескольких попыток. Возьмем некоторый набор $a_1, a_2, \ldots$ приближенных значений неизвестных. Обозначим поправки к ним через $h_1, h_2, \ldots$. Если эти поправки относительно малы, то для них тоже приближенно можно получить избыточную систему линейных уравнений, и эту последнюю уже можно решать по способу наименьших квадратов. После этого исходим уже из значений

$$a_1 + h_1, \quad a_2 + h_2, \ \ldots$$

для неизвестных и повторяем процедуру и т. д.

Покажем это на примерах.

Пример 1. Найти приближенное решение избыточной системы уравнений

$$x^2 + y^2 = 4{,}5,$$
$$x^3 - xy = 5{,}7,$$
$$x^2 - 3y^2 + x = 3{,}8.$$

В качестве первого приближения берем $x = 2$, $y = 1$, поскольку x, y входят в уравнения далеко неравноправно и мало вероятно, что они приблизительно равны, а из второго и третьего уравнений можно сделать вывод, что $x > y$.

Теперь положим $x = 2 + h$, $y = 1 + k$, тогда

$$(2 + h)^2 + (1 + k)^2 = 4{,}5,$$
$$(2 + h)^3 - (2 + h)(1 + k) = 5{,}7,$$
$$(2 + h)^2 - 3(1 + k)^2 + (1 + h) = 3{,}8.$$

Раскрыв скобки и отбросив слагаемые, содержащие h и k в степени, выше первой (h, k предполагаем малыми), находим

$$4h + 2k = -0{,}5,$$
$$11h - 2k = -0{,}3,$$
$$5h - 6k = 0{,}8.$$

Это избыточная система уравнений первой степени. Составим нормальную систему

$$162h - 44k = -1{,}3,$$
$$-44h + 44k = -5{,}2.$$

Отсюда

$$h = -0{,}06, \quad k = -0{,}18.$$

Таким образом, новые значения для x и y соответственно равны 1,94 и 0,82.

При подстановке первого приближения в заданные уравнения находим погрешности $-0{,}5$, $0{,}3$, $0{,}8$. Сумма квадратов погрешностей равна $0{,}98 \approx 1$.

Подстановка второго приближения в заданные уравнения дает погрешности 0,16, 0,16, 0,2, сумма квадратов которых равна $0{,}09 \approx 0{,}1$.

Пример 2. Найти решение системы уравнений

$$\sqrt[3]{x+y} + \sqrt{y} = 3{,}8,$$
$$\sqrt{2y - x} + \sqrt[3]{y^2} = 3{,}1,$$
$$x^2 - y^2 = 3{,}9,$$

исходя из приближения $x = 5$, $y = 3$ (цель задачи показать элементарный способ сведения данной системы к системе уравнений первой степени).

Пусть α величина малая по сравнению с единицей, тогда

$$\sqrt[n]{1+\alpha} = \sqrt[n]{1 + n\frac{\alpha}{n}} = \left(\left(1 + \frac{\alpha}{n}\right)^n - \frac{n(n-1)}{2}\left(\frac{\alpha}{n}\right)^2 - \right.$$
$$\left. - \frac{n(n-1)(n-2)}{1\cdot 2\cdot 3}\left(\frac{\alpha}{n}\right)^3 - \ldots\right)^{\frac{1}{n}}$$

(если читатель не знаком с этой формулой, то можно проверить, взяв $n = 2$, $n = 3$). Отбрасывая слагаемые, содержащие $\frac{\alpha}{n}$ в степени, большей 1, получим

$$\sqrt[n]{1+\alpha} \approx \sqrt[n]{\left(1 + \frac{\alpha}{n}\right)^n} = 1 + \frac{\alpha}{n}. \tag{1}$$

Предполагая малыми h и k, положим в заданных уравнениях $x = 5 + h$, $y = 3 + k$, тогда

$$\sqrt[3]{8 + (h+k)} + \sqrt{3 + k} = 3{,}8,$$
$$\sqrt{1 + (2k - h)} + \sqrt[3]{(3+k)^2} = 3{,}1,$$
$$(5+h)^2 - (3+k)^2 = 3{,}9.$$

Переписав эти уравнения в виде

$$2\sqrt[3]{1+\frac{h+k}{8}}+\sqrt{3}\cdot\sqrt{1+\frac{k}{3}}=3{,}8,$$

$$\sqrt{1+(2k-h)}+\sqrt[3]{9}\cdot\sqrt[3]{1+\frac{6k}{9}}=3{,}1,$$

$$5^2+10h-3^2-6k=3{,}9$$

и воспользовавшись приближенным равенством (1), находим

$$2\left(1+\frac{h+k}{24}\right)+\sqrt{3}\left(1+\frac{k}{6}\right)=3{,}8,$$

$$1+\frac{2k-h}{2}+\sqrt[3]{9}\left(1+\frac{6k}{27}\right)=3{,}1,$$

$$10h-6k=-12{,}1.$$

Это система уравнений первой степени.

Пользуясь одной формулой дифференциального исчисления (формулой Тейлора для функций нескольких переменных), практически всегда можно свести задачу решения избыточной системы уравнений не первой степени к решению системы уравнений первой степени.

§ 12. Общие замечания относительно эмпирических формул. Формулы, основанные на данных, полученных в результате проведения опытов, называются эмпирическими. Но это не означает, что только на основе опыта определяется эмпирическая формула. Вид формулы может подсказать теория, аналогия или просто интуиция, но формула содержит некоторые параметры (коэффициенты, показатели), нахождение которых основано на результатах опытов. Само собою разумеется, что любая эмпирическая формула подлежит проверке, которая может привести к уточнению этой формулы, к ограничению области ее применения или полной ее отмене. В данной работе речь может идти лишь о составлении эмпирических формул на основании числовых значений одной величины (аргумента) и числовых значений другой величины (функции), предполагаемой однозначно зависящей от первой. Как правило, и те и другие числовые значения являются приближенными, поэтому из каких бы соображений мы не выбрали б вид формулы и значения входящих в нее параметров, ожидать нахождение точной формулы не приходится. Если обеспечена удовлетворительная точность вычислений, это обстоятельство не должно смущать. Ведь мы же пользуемся формулами, содержащими число π, логарифмы, косинусы, ..., хотя π, логарифмы, ... точно вычислить невозможно.

Следовательно, эмпирические формулы всегда будут приближенными. Нам уже известны способы нахождения некоторых приближенных формул в основном полиномиального типа: тригонометрические и алгебраические многочлены. Однако этого не достаточно. Вот простейший пример. При известных условиях объем газа и давление связаны формулой $V = \frac{c}{p}$, заменить эту формулу полиномиальной нет возможности. Следующий пример показ вает, что интерполяционная формула

$$f(x) \approx 1 - \frac{x^2}{2} \tag{1}$$

для некоторого интервала изменений аргумента может быть достаточно точной и вполне пригодной для приближенного вычисления значений функции, однако эта формула может дать совершенно искаженное представление об истинной функции $f(x)$. Действительно, имеют место равенства

$$\cos x = 1 - \frac{x^2}{2} + \cdots,$$

$$e^{-\frac{x^2}{2}} = 1 - \frac{x^2}{2} + \cdots$$

и ряд других. Была бы допущена грубая ошибка, если бы мы характеризовали некоторые явления функцией $e^{-\frac{x^2}{2}}$ в то время, когда функция $1 - \frac{x^2}{2}$ фактически появилась как приближение $\cos x$.

Между тем физическое чутье и внимательное рассмотрение полученных числовых данных может подсказать, что вообще не следует искать приближение вида (1). Не трудно понять, что составление эмпирических формул это дело не простое и очень ответственное.

Заметим, что если число измерений (наблюдений) не очень мало (больше пяти), то следует найти в декартовой системе координат точки, соответствующие результатам измерений (наблюдений). Часто расположение точек подсказывает возможный вид эмпирической формулы.

§ 13. Примеры нахождения эмпирических формул.

I. Если значения аргумента образуют точно, или почти точно, арифметическую прогрессию, а соответствующие значения функции образуют почти геометрическую прогрессию, то искомую зависимость между аргументом и функцией следует искать в виде

$$y = a^{bx+c}, \quad a > 0. \tag{1}$$

Поскольку $a = m^{\log_m a}$, $a^{bx+c} = m^{(bx+c)\log_m a}$, то в качестве основания a выберем любое положительное число (обычно берем основание 10, что удобно, если понадобится пользоваться логарифмическими таблицами, или основание натуральных логарифмов $e \approx 2{,}71828\ldots$). Итак, пусть даны значения аргумента $x_1, x_2, \ldots, x_n$ и соответствующие значения функции $f(x_1), f(x_2), \ldots, f(x_n)$, которые будем предполагать положительными, и пусть

$$f(x) = 10^{bx+c}. \tag{2}$$

Прологарифмировав равенство (2) и воспользовавшись данными значениями x и $f(x)$, получим избыточную систему уравнений

$$bx_i + c = \lg f(x_i), \quad i = 1, 2, \ldots, n, \tag{3}$$

которую решаем по способу наименьших квадратов.

Замечание 1. Мы не пишем

$$f(x) = A \cdot 10^{bx+c},$$

так как A можно считать положительным (в противном случае вместо $f(x)$ возьмем $-f(x)$) и тогда

$$A = 10^{\lg A}, \quad f(x) = 10^{bx+(c+\lg A)}$$

Замечание 2. Если среди значений $f(x)$ есть часть отрицательных, то зависимость (2) заменим зависимостью вида

$$f(x) + k = 10^{bx+c},$$

где k не меньше наибольшего из модулей отрицательных $f(x_i)$.

Пример 1. Пусть значения аргумента равны 0,1, 0,2, 0,3, 0,4, а значения функции соответственно 2,5, 5,1, 9,8, 20,0. Значения аргумента образуют арифметическую прогрессию, а значения функции «почти геометрическую»—каждое значение функции почти вдвое больше предшествующего. Найдем формулу вида

$$y = 10^{ax+b}.$$

Имеем

$$\begin{gathered} ax + b = \lg y, \\ 0{,}1a + b = \lg 2{,}5 = 0{,}398, \\ 0{,}2a + b = \lg 5{,}1 = 0{,}707, \\ 0{,}3a + b = \lg 9{,}8 = 0{,}991, \\ 0{,}4a + b = \lg 20 = 1{,}301 \end{gathered}$$

и составляем нормальную систему

$$0{,}3a + b = 0{,}999\ (\approx 1),$$
$$a + 4b = 3{,}397\ (\approx 3{,}4).$$

Отсюда $a = 3$, $b = 0{,}1$ и $f(x) \approx 10^{3x+0,1}$.

Замечание. Само собою разумеется, что четырех точек (измерений) мало для обоснованных выводов. Пример носит чисто иллюстративный характер.

Пример 2. Пусть значения аргумента равны 0,5, —0,3, —0,1, 0,1, 0,3, а значения функции соответственно 1,8, 3,1, 6,5, 16,4, 49,1. Здесь значения аргумента снова образуют арифметическую прогрессию, а значения функции возрастают быстро:

$$\frac{3{,}1}{1{,}8} < 1{,}8, \quad \frac{49{,}1}{16{,}4} \approx 3.$$

Найдем формулу вида

$$y = 10^{ax^2+bx+c}.$$

Имеем

$$ax^2 + bx + c = \lg y,$$
$$0{,}25a - 0{,}5b + c = \lg 1{,}8 = 0{,}255,$$
$$0{,}09a - 0{,}3b + c = \lg 3{,}1 = 0{,}491,$$
$$0{,}01a - 0{,}1b + c = \lg 6{,}5 = 0{,}812,$$
$$0{,}01a + 0{,}1b + c = \lg 16{,}4 = 1{,}215,$$
$$0{,}09a + 0{,}3b + c = \lg 49{,}1 = 1{,}691.$$

Нормальная система:

$$0{,}079a - 0{,}125b + 0{,}45c = 0{,}284,$$
$$-0{,}125a + 0{,}45b - 0{,}5c = 0{,}273,$$
$$0{,}450a - 0{,}500b + 5c = 4{,}464.$$

Отсюда $a = 0{,}763$; $b = 1{,}783$; $c = 1{,}02$ и

$$f(x) = e^{0,763x^2+1,783x+1,020}.$$

II. На рис. 18 изображены кривые, характерной особенностью которых является убывание в одну сторону и

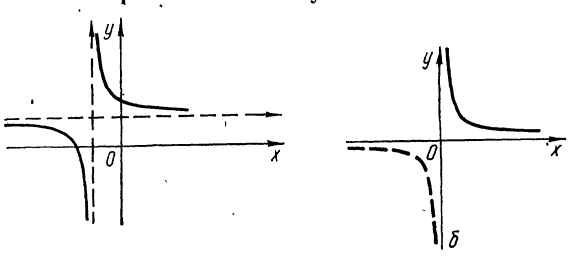

Рис. 18

возрастание в другую с замедлением, в результате чего точки кривой как бы стремятся расположиться на двух взаимно перпендикулярных прямых (асимптотах).

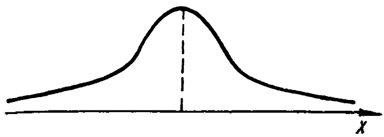

Рис. 19

На рис. 19 изображена кривая (ось Oy не показана) уравнение которой имеет вид

$$y = \frac{1}{(x-a)^2 + b}.$$

Эта кривая имеет одну асимптоту и один максимум.

Если значения аргумента не скучены и их не слишком мало, то иногда можно заметить, что заданные точки располагаются, примерно, на кривых указанных видов. Таким образом появляются эмпирические формулы вида

$$y = \frac{1}{ax+b}, \quad y = \frac{1}{ax^2+bx+c}.$$

Рассмотрим один пример.

Пример. Значения аргумента равны 1, 2, 3, 4, 5, а соответствующие значения функции 0,30, 0,20, 0,13, 0,11, 0,08. Найдем формулу вида

$$y = \frac{1}{ax+b}.$$

Имеем

$$ax + b = \frac{1}{y},$$

$$a + b = 3{,}10,$$
$$2a + b = 4{,}90,$$
$$3a + b = 7{,}15,$$
$$4a + b = 8{,}85,$$
$$5a + b = 11{,}15.$$

Составляем нормальную систему

$$55a + 15b = 125{,}50,$$
$$15a + 5b = 35{,}15$$

и находим, что $a = 2{,}005$, $b = 1{,}001$. Полагаем $a = 2$, $b = 1$, тогда

$$y = \frac{1}{2x+1}.$$

Не трудно догадаться, что возможно появление формулы вида

$$y = \frac{1}{\alpha x + \beta} + \gamma.$$

которую заменяем следующей:

$$y = \frac{a}{x+b} + c.$$

Соответствующую кривую получаем параллельным сдвигом кривой $y = \frac{1}{\alpha x + \beta}$ вверх или вниз на $|\gamma|$.

Если при увеличении значений x точки графика размещаются приблизительно на прямой $y = \alpha$, параллельной Ox, то полагаем $\gamma = \alpha$ и приходим к известному по существу случаю

$$y - \alpha = \frac{1}{\alpha x + \beta},$$
$$\alpha x + \beta = \frac{1}{y - \alpha},$$
$$\alpha x_1 + \beta = \frac{1}{y_1 - \alpha},$$
$$\alpha x_2 + \beta = \frac{1}{y_2 - \alpha},$$
$$\cdots\cdots\cdots$$

Полагая, что

$$y = \frac{a}{x+b} + c,$$

получим

$$xy + by - cx - bc - a = 0.$$

Отсюда, обозначая $bc + a = a_1$, получим

$$a_1 + cx - by = xy,$$
$$a_1 + cx_1 - by_1 = x_1y_1,$$
$$a_1 + cx_2 - by_2 = x_2y_2,$$
$$\cdots\cdots\cdots$$

Найдя из соответствующей нормальной системы b, c, a_1, найдем и a.

III. Кривые вида

$$y = (x - a_1)^{\alpha_1} (a_2 - x)^{\alpha_2},$$

если известны a_1, a_2 и не известны α_1, α_2, находим следующим образом.

Логарифмирование приведет к уравнению

$$\alpha_1 \lg (x - a_1) + \alpha_2 \lg (a_2 - x) = \lg y,$$

линейному относительно искомых α_1, α_2. Значит получим систему уравнений

$$\alpha_1 \lg(x_i - a_1) + \alpha_2 \lg(a_2 - x_i) = \lg y_i, \; i = 1, 2, \ldots, m,$$

и сможем применить способ наименьших квадратов.

Упражнения

1. Найдите многочлен третьей степени, корни которого равны 1, —1, 2.

2. С помощью теоремы Ролля докажите, что многочлен $x^n - a^n$ (n — целое) имеет не более двух действительных корней.

3. Не вычисляя производных $f'(x)$, $f''(x)$, ..., найдите $f'(a)$, $f''(a)$, $f'''(a)$, ..., если

$$f(x) = x^5 + 7x^4 + 6x^3 + 4x^2 + 5.$$

4. Пусть $f(x)$ многочлен m-й степени и известны значения $f(x_1)$ $f(x_2)$, ..., $f(x_n)$, $n > m$. Докажите, что интерполяционный многочлен Лагранжа $P(x)$ будет многочленом m-й степени. Более того, $P(x) \equiv f(x)$.

5. Пусть $f(a)$, $f(a+h)$, $f(a+2h)$, ... — значения многочлена m-й степени. Докажите, что конечные разности $\Delta^r f(a)$ равны нулю при $r > m$.

6. Докажите, что

$$f(a+h) = f(a) + \Delta f(a),$$
$$f(a+2h) = f(a) + 2\Delta f(a) + \Delta^2 f(a),$$
$$f(a+3h) = f(a) + 3\Delta f(a) + 3\Delta^2 f(a) + \Delta^3 f(a).$$

На основании этих формул проверьте интерполяционную формулу Ньютона.

7. Составьте формулу линейного интерполирования для $\sqrt{x}$. Исходя из равенств $\sqrt{4} = 2$, $\sqrt{9} = 3$, найдите с помощью линейного интерполирования $\sqrt{7}$, $\sqrt{6}$ и проверьте, что результаты верны с точностью до 0,1.

8. Пусть

$$\begin{vmatrix} a & b \\ A & B \end{vmatrix} = aB - bA.$$

Проверьте, что решение системы уравнений

$$ax + by = c,$$
$$Ax + By = C$$

можно записать в виде

$$x = \frac{\begin{vmatrix} c & b \\ C & B \end{vmatrix}}{\begin{vmatrix} a & b \\ A & B \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} a & c \\ A & C \end{vmatrix}}{\begin{vmatrix} a & b \\ A & B \end{vmatrix}} \tag{1}$$

или, что то же самое,

$$x = \frac{P}{Q}, \quad P = \begin{vmatrix} c & b \\ C & B \end{vmatrix}, \quad Q = \begin{vmatrix} a & b \\ A & B \end{vmatrix}. \tag{2}$$

Пусть дана избыточная система

$$
\begin{aligned}
x - y &= 1,\\
2x + y &= 2,\\
-x + 3y &= -1,\\
3x + 4y &= 2.
\end{aligned}
$$

Сгруппируйте эти уравнения парами (6 пар!) и найдите для каждой пары уравнений

$$x_1 = \frac{P_1}{Q_1}, \quad x_2 = \frac{P_2}{Q_2}, \quad \ldots$$

по формуле (1). Проверьте, что решение нормальной системы для x совпадает со значением дроби

$$\frac{P_1Q_1 + P_2Q_2 + \cdots + P_6Q_6}{Q_1^2 + Q_2^2 + \cdots + Q_6^2}.$$

9. Пусть дана система уравнений

$$
\begin{aligned}
ax + by + cz &= h,\\
\alpha x + \beta y + \gamma z &= k,\\
Ax + By + Cz &= l_1,\\
Ax + By + Cz &= l_2,\\
Ax + By + Cz &= l_3.
\end{aligned}
$$

Докажите, отыскивая минимум выражения

$$(ax + by + cz - h)^2 + \cdots + (Ax + By + Cz - l_3)^2,$$

что решение по способу наименьших квадратов будет точным решением системы уравнений

$$
\begin{aligned}
ax + by + cz &= h,\\
\alpha x + \beta y + \gamma z &= k,\\
Ax + By + Cz &= \frac{l_1 + l_2 + l_3}{3}.
\end{aligned}
$$

ОГЛАВЛЕНИЕ

БИБЛИОТЕЧКА
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ШКОЛЫ

МАТЕМАТИКА

Гершон Ихелевич Дринфельд

ИНТЕРПОЛИРОВАНИЕ И СПОСОБ НАИМЕНЬШИХ КВАДРАТОВ

Редактор *Л. П. Онищенко*
Литредактор *Л. П. Никитина*
Художественный редактор *Е. В. Чурий*
Технический редактор *А. И. Омоховская*
Корректор *Ф. И. Слободская*

Информ. бланк № 8215

Сдано в набор 28.03.84. Подп. в печать 06.08.84. БФ 30316 Формат 84×108/32. Бумага типогр. № 3 Лит. гарн. Выс. печ. 5,46 усл. печ. л. 5,77 усл. кр.-отт. 4,58 уч.-изд. л. Тираж 6000 экз. Изд. № 6365. Зак. 4-664. Цена 15 коп.

Головное издательство издательского объединения «Вища школа», 252054, Киев-54, ул. Гоголевская, 7

Отпечатано с матриц книжной фабрики им. М. В. Фрунзе, 310057, Харьков-57, ул. Донец-Захаржевского, 6/8 в харьковской типографии № 16, г. Харьков, ул. Университетская, 16. Зак. 1335.