Оглавление

1 Аппроксимация в линейных нормированных пространствах

1.1 Существование элемента наилучшего приближения

Пусть X — произвольное линейное нормированное пространство, $L = \mathrm{lin}\{f_1, f_2, \ldots, f_n\}$ — пространство, натянутое на n линейно независимых векторов из X. Рассмотрим задачу приближения заданного элемента $x \notin L$ подпространством L:

$$d(x, L) = \inf_{y \in L} \|x - y\|_X. \tag{P}$$

Теорема 1. *Элемент наилучшего приближения в задаче* (P) *существует.*

Доказательство. Пусть вектор $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{R}^n$. Обозначим

$$\varphi(\lambda) = \|x - \lambda_1 f_1 - \lambda_2 f_2 - \cdots - \lambda_n f_n\|.$$

Доказать существование элемента наилучшего приближения, значит, доказать существование коэффициентов $\lambda_1, \lambda_2, \ldots, \lambda_n$, на которых достигается инфимум в (P).

Покажем, что $\varphi(\lambda)$ является непрерывной функцией. Действительно, в силу неравенства треугольника,

$$\varphi(\lambda') - \varphi(\lambda) = \left\| x - \sum_{k=1}^{n} \lambda'_k f_k \right\| - \left\| x - \sum_{k=1}^{n} \lambda_k f_k \right\| \leq$$

$$\leq \left\| \sum_{k=1}^{n} (\lambda'_k - \lambda_k) f_k \right\| \leq \sum_{k=1}^{n} |\lambda'_k - \lambda_k| \cdot \|f_k\| \leq \max_k |\lambda'_k - \lambda_k| \sum_{k=1}^{n} \|f_k\|.$$

Следовательно, $\varphi(\lambda') - \varphi(\lambda) \to 0$ при $\lambda' \to \lambda$.

Введем теперь вторую непрерывную функцию

$$\psi(\lambda) = \|\lambda_1 f_1 + \lambda_2 f_2 + \dots + \lambda_n f_n\|.$$

На «сфере» $|\lambda_1|^2 + |\lambda_2|^2 + \dots + |\lambda_n|^2 = 1$, которая является ограниченным замкнутым множеством (компактом) точек в конечномерном евклидовом пространстве, функция ψ по теореме Вейерштрасса достигает минимум μ. Причем, неотрицательное число μ не может равняться нулю, так как векторы $f_1, f_2, \dots, f_n$ линейно независимы и, значит, $\mu > 0$. Обозначим далее через d нижнюю грань значений функции φ. Очевидно, что $d \geq 0$.

Если

$$\sqrt{\sum_{k=1}^{n} |\lambda_k|^2} > \frac{1}{\mu}(d + 1 + \|x\|) = R,$$

то по неравенству треугольника

$$\varphi(\lambda_1, \lambda_2, \dots, \lambda_n) \geq \|\lambda_1 f_1 + \lambda_2 f_2 + \dots + \lambda_n f_n\| - \|x\| \geq$$

$$\geq \mu \sqrt{\sum_{k=1}^{n} |\lambda_k|^2} - \|x\| > d + 1.$$

Поэтому, желая найти минимум функции $\varphi(\lambda)$, мы можем ограничиться рассмотрением только таких значений λ_k, для которых $\sum\limits_{k=1}^{n} |\lambda_k|^2 \leq R^2$, т. е. рассмотрением функции $\varphi(\lambda)$ в ограниченной замкнутой области (компакте), а в такой области непрерывная функция по теореме Вейерштрасса достигает минимум.

Итак, существование линейной комбинации $\lambda_1 f_1 + \lambda_2 f_2 + \dots + \lambda_n f_n$, дающей наилучшую аппроксимацию элемента x, доказано. $\square$

1.2 Геометрическая интерпретация

Проблема существования решения, которую мы доказали выше, имеет следующую геометрическую интерпретацию. Совокупность точек вида

$$\lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_n f_n,$$

где зафиксированные векторы $f_1, f_2, \ldots, f_n$ принадлежат X и линейно независимы, а $\lambda_1, \lambda_2, \ldots, \lambda_n$ пробегают всевозможные действительные (комплексные) числа, представляет некоторое линейное многообразие $L \subset X$ в том смысле, что из $f', f'' \in L$ следует, что и $\alpha' f' + \alpha'' f'' \in L$ при произвольных действительных α', α''. Это линейное подпространство. При $n = 1$ мы получаем «прямую», при $n = 2$ — «плоскость», а вообще — «n-мерную плоскость» (при линейно независимых векторах $f_1, f_2, \ldots, f_n$). Наша задача, таким образом, состоит в нахождении точки конечномерного подпространства L пространства X, которая от заданной точки $x \in X$ находится на кратчайшем расстоянии (в метрике пространства X). Мы доказали, что такая точка в L существует.

Если само пространство X не является конечномерным, т. е. если в нем имеется сколько угодно линейно независимых между собой векторов, то X содержит бесконечномерные подпространства. Пусть L — такое подпространство.

Возникает вопрос, существует ли в L точка, наименее удаленная от заданной точки $x \in X$. Ниже мы дадим на этот вопрос положительный ответ, если X есть полное гильбертово пространство.

2 Строго нормированные пространства

Определение 1. *Линейное нормированное пространство X называется строго нормированным, если в неравенстве треугольника*

$$\|x + y\| \leq \|x\| + \|y\| \quad (x, y \neq 0)$$

равенство достигается только при $y = \alpha x$, где $\alpha > 0$.

Возникает вопрос: когда «полином» $\lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_n f_n$, дающий наилучшую аппроксимацию элемента x, будет единственным для всякого $x \in X$? Ответ на него дает следующая теорема.

2.1 Единственность элемента наилучшего приближения

Теорема 2. *В строго нормированном пространстве элемент наилучшего приближения в задаче* (P) *единственен.*

Доказательство. Если элемент $x \in L$, т. е. $x = \lambda_1 f_1 + \lambda_2 f_2 + \dots + \lambda_n f_n$, то «полином» $\lambda_1 f_1 + \lambda_2 f_2 + \dots + \lambda_n f_n$ дает наилучшую аппроксимацию элемента x и $d = d(x, L) = 0$ и полином единственен.

Пусть элемент $x \notin L$. Докажем единственность приближающего полинома в этом случае. Предположим противное, что некоторый элемент x имеет два полинома наилучшего приближения y_1 и y_2, $y_1 \neq y_2$, т. е.

$$d = d(x, L) = \inf_{y \in L} \|x - y\| = \|x - y_1\| = \|x - y_2\|.$$

Тогда

$$d \leq \left\| x - \frac{y_1 + y_2}{2} \right\| = \left\| \frac{x - y_1}{2} + \frac{x - y_2}{2} \right\| \leq \left\| \frac{x - y_1}{2} \right\| + \left\| \frac{x - y_2}{2} \right\| = d.$$

Отсюда, в вышенаписанном соотношении знак неравенства всюду можно заменить равенством, поэтому, умножая на 2, получаем, что

$$\|(x - y_1) + (x - y_2)\| = \|x - y_1\| + \|x - y_2\|.$$

Следовательно, в силу строгой нормированности пространства,

$$x - y_1 = \alpha(x - y_2), \ \alpha > 0.$$

В этом соотношении α должно равняться единице, так как в противном случае элемент x был бы линейной комбинацией элементов y_1, y_2, т. е. принадлежал подпространству L, что не так. Но если $\alpha = 1$, то $y_1 = y_2$. Получили противоречие, которое говорит о том, что наше предположение о существовании двух разных полиномов наилучшего приближения неверно. □

2.2 Строго выпуклые множества и функции

Напомним некоторые определения выпуклого анализа.

Определение 2. *Точка d выпуклого множества D называется крайней (угловой), если не существует точек $d_1, d_2 \in D$, $d_1 \neq d_2$, и числа $t \in (0,\, 1)$ таких, что $d = td_1 + (1-t)d_2$.*

Определение 3. *Замкнутое выпуклое множество называется строго выпуклым, если все его граничные точки крайние.*

Определение 4. *Функция f называется строго выпуклой, если ее надграфик — строго выпуклое множество.*

Из определения строго выпуклого множества сразу следует, что функция f строго выпукла тогда и только тогда, когда для любых точек $x_1, x_2 \in X$, $x_1 \neq x_2$, и для любого $t \in (0,\, 1)$ выполнено *строгое неравенство Йенсена*:

$$f(tx_1 + (1-t)x_2) < tf(x_1) + (1-t)f(x_2).$$

Теорема 3. *Нормированное пространство является строго нормированным тогда и только тогда, когда каждый шар в нем является строго выпуклым множеством.*[1]

[1]Замкнутое выпуклое множество называется строго выпуклым, если все его граничные точки крайние.

2.3 Неравенство Юнга

Определение. *Числа p и q называются сопряженными показателями, если они удовлетворяют условиям*

$$1 < p, q < \infty, \quad \frac{1}{p} + \frac{1}{q} = 1.$$

Теорема (неравенство Юнга)**.** *Для любых неотрицательных чисел a и b и сопряженных показателей p и q справедливо неравенство:*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Причем, равенство достигается тогда и только тогда, когда $a^p = b^q$.

Доказательство. Если хотя бы одно из чисел a или b обращается в ноль, то неравенство выполняется. Пусть теперь $a > 0$ и $b > 0$. Зафиксировав произвольное положительное число a, определим функцию $f : (0, +\infty) \to \mathbb{R}$, $f(b) = \frac{a^p}{p} + \frac{b^q}{q} - ab$. Найдем минимум функции f. Производная функции

$$f'(b) = b^{q-1} - a \iff b = a^{\frac{1}{q-1}}.$$

Поскольку $f''(b) = (q-1)b^{q-2} > 0$, то это точка глобального минимума. Значит,

$$f(b) \geq f\left(a^{\frac{1}{q-1}}\right) \quad \forall\ b > 0.$$

Посчитаем значение f в точке минимума:

$$f\left(a^{\frac{1}{q-1}}\right) = \frac{a^p}{p} + \frac{a^{\frac{q}{q-1}}}{q} - a \cdot a^{\frac{1}{q-1}} = a^p\left(\frac{1}{p} + \frac{1}{q}\right) - a^{\frac{1}{q-1}+1} = a^p - a^p = 0$$

$\Big($так как $\frac{1}{p} + \frac{1}{q} = 1 \Leftrightarrow \frac{1}{p} = \frac{q-1}{q} \Leftrightarrow p = \frac{q}{q-1}\Big)$. Следовательно,

$$f(b) \geq 0 \iff \frac{a^p}{p} + \frac{b^q}{q} - ab \geq 0 \iff ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Причем, равенство имеет место только в случае $b = a^{\frac{1}{q-1}} \Leftrightarrow b = a^{\frac{p}{q}} \Leftrightarrow a^p = b^q$. $\square$

2.4 Неравенство Гёльдера

Теорема (неравенство Гёльдера для сумм). *Пусть p и q — сопряженные показатели. Тогда*

$$\left|\sum_{k=1}^{n} x_k y_k\right| \le \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |y_k|^q\right)^{\frac{1}{q}}.$$

Доказательство. Пусть векторы $c = (c_1, \ldots, c_n), d = (d_1, \ldots, d_n) \in \mathbb{R}^n$ таковы, что выполняются равенства

$$\sum_{k=1}^{n} |c_k|^p = 1,\ \sum_{k=1}^{n} |d_k|^q = 1.$$

По неравенству Юнга $|c_k| \cdot |d_k| \le \dfrac{|c_k|^p}{p} + \dfrac{|d_k|^q}{q}$. Просуммировав по k от 1 до n, получаем

$$\sum_{k=1}^{n} |c_k| \cdot |d_k| \le \frac{\sum\limits_{k=1}^{n} |c_k|^p}{p} + \frac{\sum\limits_{k=1}^{n} |d_k|^q}{q} = \frac{1}{p} + \frac{1}{q} = 1.$$

Положим $c_k = \frac{|x_k|}{\left(\sum\limits_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}}}$, $d_k = \frac{|y_k|}{\left(\sum\limits_{j=1}^{n} |y_j|^q\right)^{\frac{1}{q}}}$. Тогда из предыдущего неравенства

$$\sum_{k=1}^{n} \frac{|x_k|}{\left(\sum\limits_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}}} \cdot \frac{|y_k|}{\left(\sum\limits_{j=1}^{n} |y_j|^p\right)^{\frac{1}{p}}} \le 1$$

$$\iff \sum_{k=1}^{n} |x_k||y_k| \le \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |y_k|^q\right)^{\frac{1}{q}}.$$

Используя то, что модуль суммы не превосходит суммы модулей, получаем неравенство Гельдера. Неравенство Гельдера обращается в равенство тогда и только тогда, когда

$$\frac{|x_k|^p}{\sum\limits_{i=1}^{n} |x_i|^p} = \frac{|y_k|^q}{\sum\limits_{i=1}^{n} |y_i|^q},\ \operatorname{sign}(x_k y_k) = \text{const},\ k = 1, \ldots n. \qquad \square$$

Теорема (неравенство Гёльдера для интегралов)**.** *Пусть функции* $x(\cdot) \in L_p([a,\, b])$*,* $y(\cdot) \in L_q([a,\, b])$*, где* p *и* q *— сопряженные показатели. Тогда*

$$\int\limits_a^b |x(t)y(t)|dt \leq \left(\int\limits_a^b |x(t)|^p dt\right)^{\frac{1}{p}} \left(\int\limits_a^b |y(t)|^q dt\right)^{\frac{1}{q}}.$$

Доказательство. Положим $a = \dfrac{|x(t)|}{\|x(\cdot)\|_p}$, $b = \dfrac{|y(t)|}{\|y(\cdot)\|_q}$. По неравенству Юнга для неотрицательных чисел a и b и сопряженных показателей p и q

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

имеем

$$\frac{|x(t)|}{\|x(\cdot)\|_p} \cdot \frac{|y(t)|}{\|y(\cdot)\|_q} \leq \frac{|x(t)|^p}{p\|x(\cdot)\|_p^p} + \frac{|y(t)|^q}{q\|y(\cdot)\|_q^q}.$$

Проинтегрируем по $t \in [a,\, b]$ обе части неравенства:

$$\frac{\int_a^b |x(t)| \cdot |y(t)|dt}{\|x(\cdot)\|_p \cdot \|y(\cdot)\|_q} \leq \frac{\int_a^b |x(t)|^p dt}{p\|x(\cdot)\|_p^p} + \frac{\int_a^b |y(t)|^q dt}{q\|y(\cdot)\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1.$$

Отсюда получаем неравенство Гельдера

$$\int\limits_a^b |x(t)y(t)|dt \leq \|x(\cdot)\|_p \cdot \|y(\cdot)\|_q.$$

Неравенство Гельдера обращается в равенство тогда и только тогда, когда

$$\frac{|x(t)|^p}{\|x(\cdot)\|_p^p} = \frac{|y(t)|^q}{\|y(\cdot)\|_q^q}, \text{ почти всюду.}$$

□

2.5 Неравенство Минковского

Теорема (неравенство Минковского для сумм)**.** *Пусть* $1 \le p < \infty$*. Тогда имеет место неравенство*

$$\left(\sum_{k=1}^{n} |x_k + y_k|^p\right)^{\frac{1}{p}} \le \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |y_k|^p\right)^{\frac{1}{p}}.$$

Доказательство. При $p = 1$ неравенство Минковского следует из неравенства для модулей

$$|x_k + y_k| \le |x_k| + |y_k|.$$

Пусть теперь $p > 1$. Используя это неравенство, получим:

$$\sum_{k=1}^{n} |x_k + y_k|^p = \sum_{k=1}^{n} |x_k + y_k|\, |x_k + y_k|^{p-1} \le$$

$$\le \sum_{k=1}^{n} |x_k|\, |x_k + y_k|^{p-1} + \sum_{k=1}^{n} |y_k|\, |x_k + y_k|^{p-1}.$$

К каждому слагаемому в правой части последнего неравенства применим неравенство Гёльдера для сумм:

$$\sum_{k=1}^{n} |x_k + y_k|^p \le \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |x_k + y_k|^{(p-1)q}\right)^{\frac{1}{q}} +$$

$$+ \left(\sum_{k=1}^{n} |y_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |x_k + y_k|^{(p-1)q}\right)^{\frac{1}{q}} =$$

$$\left(\text{так как} \frac{1}{p} + \frac{1}{q} = 1 \Leftrightarrow \frac{1}{q} = 1 - \frac{1}{p} \Leftrightarrow \frac{1}{q} = \frac{p-1}{p} \Leftrightarrow (p-1)q = p\right)$$

$$= \left(\left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |y_k|^p\right)^{\frac{1}{p}} \right) \left(\sum_{k=1}^{n} |x_k + y_k|^p\right)^{\frac{1}{q}}.$$

Отсюда

$$\left(\sum_{k=1}^{n}|x_k+y_k|^p\right)^{1-\frac{1}{q}} \le \left(\sum_{k=1}^{n}|x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n}|y_k|^p\right)^{\frac{1}{p}} \iff$$

$$\left(\sum_{k=1}^{n}|x_k+y_k|^p\right)^{\frac{1}{p}} \le \left(\sum_{k=1}^{n}|x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n}|y_k|^p\right)^{\frac{1}{p}}. \qquad \square$$

Замечание. *Неравенство Минковского для сумм при* $1 < p < \infty$ *обращается в равенство тогда и только тогда, когда* $y = \alpha x$, *где* $\alpha > 0$.

Доказательство. В процессе доказательства неравенства Минковского для сумм было использовано два неравенства Гёльдера:

$$1)\ \sum_{k=1}^{n} |x_k|\,|x_k + y_k|^{p-1} \le \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |x_k + y_k|^{(p-1)q}\right)^{\frac{1}{q}};$$

$$2)\ \sum_{k=1}^{n} |y_k|\,|x_k + y_k|^{p-1} \le \left(\sum_{k=1}^{n} |y_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |x_k + y_k|^{(p-1)q}\right)^{\frac{1}{q}}.$$

Выпишем условия обращения этих неравенств в равенства:

$$1)\ \frac{|x_k|^p}{\sum\limits_{i=1}^{n} |x_i|^p} = \frac{|x_k + y_k|^{(p-1)q}}{\sum\limits_{i=1}^{n} |x_i + y_i|^{(p-1)q}}, \quad 2)\ \frac{|y_k|^p}{\sum\limits_{i=1}^{n} |y_i|^p} = \frac{|x_k + y_k|^{(p-1)q}}{\sum\limits_{i=1}^{n} |x_i + y_i|^{(p-1)q}},$$

$k = 1, \ldots, n$. Из этих условий при $(p-1)q = p$ получается следующее равенство:

$$\frac{|x_k|^p}{\sum\limits_{i=1}^{n} |x_i|^p} = \frac{|y_k|^p}{\sum\limits_{i=1}^{n} |y_i|^p} \iff \frac{|x_k|}{\|x\|_p} = \frac{|y_k|}{\|y\|_p} \iff |y_k| = \frac{\|y\|_p}{\|x\|_p}|x_k|,$$

$k = 1, \ldots, n$. В неравенстве $|x_k + y_k| \le |x_k| + |y_k|$ знак равенства достигается, если $\operatorname{sign} x_k = \operatorname{sign} y_k$. Поэтому $y_k = \frac{\|y\|_p}{\|x\|_p} x_k$ и, значит, $y = \alpha x$, где $\alpha = \frac{\|y\|_p}{\|x\|_p} > 0$. $\square$

Теорема (неравенство Минковского для интегралов)**.** *Пусть* $1 \le p < \infty$*. Тогда имеет место неравенство:*

$$\left(\int\limits_a^b |x(t)+y(t)|^p dt\right)^{\frac{1}{p}} \le \left(\int\limits_a^b |x(t)|^p dt\right)^{\frac{1}{p}} + \left(\int\limits_a^b |y(t)|^p dt\right)^{\frac{1}{p}}.$$

Доказательство. При $p = 1$ неравенство Минковского следует из неравенства для модулей

$$|x(t)+y(t)| \le |x(t)| + |y(t)|.$$

Пусть теперь $p > 1$. Используя это неравенство, получим:

$$|x(t)+y(t)|^p = |x(t)+y(t)| \cdot |x(t)+y(t)|^{p-1} \le$$
$$\le |x(t)| \cdot |x(t)+y(t)|^{p-1} + |y(t)| \cdot |x(t)+y(t)|^{p-1}.$$

Проинтегрируем обе части последнего неравенства:

$$\int\limits_a^b |x(t)+y(t)|^p\, dt \le \int\limits_a^b |x(t)| \cdot |x(t)+y(t)|^{p-1} dt + \int\limits_a^b |y(t)| \cdot |x(t)+y(t)|^{p-1} dt.$$

К каждому слагаемому в правой части последнего неравенства применим неравенство Гёльдера для интегралов:

$$\int\limits_a^b |x(t)+y(t)|^p\, dt \le \left(\int\limits_a^b |x(t)|^p dt\right)^{\frac{1}{p}} \left(\int\limits_a^b |x(t)+y(t)|^{(p-1)q} dt\right)^{\frac{1}{q}} +$$
$$+ \left(\int\limits_a^b |y(t)|^p dt\right)^{\frac{1}{p}} \left(\int\limits_a^b |x(t)+y(t)|^{(p-1)q} dt\right)^{\frac{1}{q}} \le$$

$$\left(\text{так как} \frac{1}{p} + \frac{1}{q} = 1 \Leftrightarrow \frac{1}{q} = 1 - \frac{1}{p} \Leftrightarrow \frac{1}{q} = \frac{p-1}{p} \Leftrightarrow (p-1)q = p\right)$$

$$\le \left(\left(\int\limits_a^b |x(t)|^p dt\right)^{\frac{1}{p}} + \left(\int\limits_a^b |y(t)|^p dt\right)^{\frac{1}{p}}\right) \left(\int\limits_a^b |x(t)+y(t)|^p dt\right)^{\frac{1}{q}}.$$

Отсюда

$$\left(\int_a^b |x(t)+y(t)|^p dt\right)^{1-\frac{1}{q}} \leq \left(\int_a^b |x(t)|^p dt\right)^{\frac{1}{p}} + \left(\int_a^b |y(t)|^p dt\right)^{\frac{1}{p}} \iff$$

$$\left(\int_a^b |x(t)+y(t)|^p dt\right)^{\frac{1}{p}} \leq \left(\int_a^b |x(t)|^p dt\right)^{\frac{1}{p}} + \left(\int_a^b |y(t)|^p dt\right)^{\frac{1}{p}}. \ \square$$

Замечание. *Неравенство Минковского для интегралов при* $1 < p < \infty$ *обращается в равенство тогда и только тогда, когда* $y(t) = \alpha x(t)$ *почти всюду с константой* $\alpha > 0$.

Доказательство. В процессе доказательства неравенства Минковского для интегралов было использовано два неравенства Гёльдера:

$$1)\ \int_a^b |x(t)|\cdot|x(t)+y(t)|^{p-1}dt \le \left(\int_a^b |x(t)|^p dt\right)^{\frac{1}{p}} \left(\int_a^b |x(t)+y(t)|^{(p-1)q}dt\right)^{\frac{1}{q}};$$

$$2)\ \int_a^b |y(t)|\cdot|x(t)+y(t)|^{p-1}dt \le \left(\int_a^b |y(t)|^p t\right)^{\frac{1}{p}} \left(\int_a^b |x(t)+y(t)|^{(p-1)q}dt\right)^{\frac{1}{q}}.$$

Выпишем условия обращения этих неравенств в равенства:

$$1)\ \frac{|x(t)|^p}{\int_a^b |x|^p dt} = \frac{|x(t)+y(t)|^{(p-1)q}}{\int_a^b |x+y|^{(p-1)q}dt}; \quad 2)\ \frac{|y(t)|^p}{\int_a^b |y|^p dt} = \frac{|x(t)+y(t)|^{(p-1)q}}{\int_a^b |x+y|^{(p-1)q}dt}$$

почти всюду. Из этих условий при $(p-1)q = p$ получается следующее равенство:

$$\frac{|x(t)|^p}{\|x\|_p^p} = \frac{|y(t)|^p}{\|y\|_p^p} \iff \frac{|x(t)|}{\|x\|_p} = \frac{|y(t)|}{\|y\|_p} \iff |y(t)| = \frac{\|y\|_p}{\|x\|_p}|x(t)|$$

почти всюду. В неравенстве $|x(t)+y(t)| \le |x(t)|+|y(t)|$ знак равенства достигается, если $\operatorname{sign} x(t) = \operatorname{sign} y(t)$. Поэтому $y(t) = \frac{\|y\|_p}{\|x\|_p}x(t)$ почти всюду и, значит, $y(t) = \alpha x(t)$ почти всюду с константой $\alpha = \frac{\|y\|_p}{\|x\|_p} > 0$. $\square$

2.6 Примеры строго нормированных пространств

Пример 1. *Гильбертово пространство*[2] *H является строго нормированным пространством.*

Доказательство. Предположим, что в гильбертовом пространстве выполнилось равенство:

$$\|x+y\| = \|x\| + \|y\| \quad (x, y \neq 0). \tag{1}$$

Покажем, что тогда $y = \alpha x$, где $\alpha > 0$. Поскольку норма в гильбертовом пространстве $\|x\| = \sqrt{\langle x, x\rangle}$, то равенство (1) эквивалентно соотношению

$$\sqrt{\langle x+y, x+y\rangle} = \sqrt{\langle x, x\rangle} + \sqrt{\langle y, y\rangle} \iff$$

$$\langle x+y, x+y\rangle = \left(\sqrt{\langle x, x\rangle} + \sqrt{\langle y, y\rangle}\right)^2 \iff$$

$$\langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle = \langle x, x\rangle + 2\sqrt{\langle y, y\rangle}\sqrt{\langle x, x\rangle} + \langle y, y\rangle$$

$$\langle x, y\rangle = \sqrt{\langle y, y\rangle}\sqrt{\langle x, x\rangle} \iff \langle x, y\rangle = \|x\| \cdot \|y\|. \tag{2}$$

Следовательно, для числа $\alpha \in \mathbb{R}$

$$\|-\alpha x + y\|^2 = \langle -\alpha x + y, -\alpha x + y\rangle = \alpha^2\|x\|^2 - 2\alpha\langle x, y\rangle + \|y\|^2 \overset{(2)}{=}$$

$$= \alpha^2\|x\|^2 - 2\alpha\|x\| \cdot \|y\| + \|y\|^2 = \big(-\alpha\|x\| + \|y\|\big)^2 = 0$$

при $\alpha = \frac{\|y\|}{\|x\|} > 0$. Значит, $\|-\alpha x + y\| = 0 \Leftrightarrow -\alpha x + y = 0 \Leftrightarrow$ $y = \alpha x$, где $\alpha > 0$. $\square$

[2]Гильбертово пространство — линейное (векторное) пространство (над полем вещественных или комплексных чисел), в котором для любых двух элементов пространства x и y определено скалярное произведение $\langle x, y\rangle$ и полное (значит, банахово) относительно порожденной скалярным произведением нормы $\|x\| = \sqrt{\langle x, x\rangle}$.

Пример 2. *Пространство l_p с нормой*

$$\|x\|_{l_p} := \begin{cases} \left(\sum\limits_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}}, & 1 \le p < \infty, \\ \max\{|x_1|, \ldots, |x_n|\}, & p = +\infty, \end{cases}$$

является строго нормированным при $1 < p < \infty$.

Доказательство. Пусть $1 < p < \infty$. Выпишем неравенство треугольника для норм в пространстве l_p, которое носит название неравенства Минковского для сумм:

$$\left(\sum_{k=1}^{n} |x_k + y_k|^p\right)^{\frac{1}{p}} \le \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |y_k|^p\right)^{\frac{1}{p}}$$

Предположим, что выполнилось равенство

$$\left(\sum_{k=1}^{n} |x_k + y_k|^p\right)^{\frac{1}{p}} = \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^{n} |y_k|^p\right)^{\frac{1}{p}} \quad (x, y \ne 0).$$

Согласно замечанию к неравенству Минковского для сумм неравенство при $1 < p < \infty$ обращается в равенство тогда и только тогда, когда $y = \alpha x$, где $\alpha > 0$. Значит, пространство l_p является строго нормированным при $1 < p < \infty$. $\square$

Пример 3. *Пространство $L_p([a,\, b])$ измеримых функций, для которых конечна величина*

$$\|x\|_{L_p([a,\,b])} := \begin{cases} \left(\int\limits_a^b |x(t)|^p\, dt\right)^{\frac{1}{p}}, & 1 \le p < \infty, \\ \operatorname*{ess\,max}\limits_{t\in[a,\,b]} |x(t)|, & p = +\infty, \end{cases}$$ [3]

является строго нормированным при $1 < p < \infty$.

Доказательство. Пусть $1 < p < \infty$. Выпишем неравенство треугольника для норм в пространстве $L_p([a,\, b])$, которое носит название неравенства Минковского для интегралов:

$$\left(\int\limits_a^b |x(t)+y(t)|^p dt\right)^{\frac{1}{p}} \le \left(\int\limits_a^b |x(t)|^p dt\right)^{\frac{1}{p}} + \left(\int\limits_a^b |y(t)|^p dt\right)^{\frac{1}{p}}.$$

Предположим, что выполнилось равенство

$$\left(\int\limits_a^b |x(t)+y(t)|^p dt\right)^{\frac{1}{p}} = \left(\int\limits_a^b |x(t)|^p dt\right)^{\frac{1}{p}} + \left(\int\limits_a^b |y(t)|^p dt\right)^{\frac{1}{p}}$$

для функций $x(\cdot), y(\cdot) \ne 0$ в метрике $L_p([a,\, b])$. Согласно замечанию к неравенству Минковского для интегралов неравенство при $1 < p < \infty$ обращается в равенство тогда и только тогда, когда $y(t) = \alpha x(t)$ почти всюду с константой $\alpha > 0$. Значит, пространство $L_p([a,\, b])$ является строго нормированным при $1 < p < \infty$. □

[3] ess max — essential maximum —существенный максимум означает, что максимум берется на всем отрезке, кроме быть может множества меры ноль

2.7 Упражнения

1. Доказать, что пространство l_1 не является строго нормированным пространством.

Решение. Рассмотрим для простоты пространство двумерных векторов с нормой l_1. Положим $x = (1, 0)$, $y = (0, 1)$. Тогда

$\|x\|_{l_1} = |x_1| + |x_2| = 1 + 0 = 1,$

$\|y\|_{l_1} = |y_1| + |y_2| = 0 + 1 = 1,$

$\|x + y\|_{l_1} = |x_1 + y_1| + |x_2 + y_2| = |1| + |1| = 2.$

Следовательно,

$$\|x + y\| = \|x\| + \|y\| \quad (x, y \neq 0),$$

но $y \neq \alpha x$, где $\alpha > 0$.

2. Доказать, что пространство l_∞ не является строго нормированным пространством.

Решение. Рассмотрим для простоты пространство двумерных векторов с нормой l_∞. Положим $x = (1, 0)$, $y = (2, 1)$. Тогда

$\|x\|_{l_\infty} = \max\{|x_1|, |x_2|\} = \max\{|1|, |0|\} = 1,$

$\|y\|_{l_\infty} = \max\{|y_1|, |y_2|\} = \max\{|2|, |1|\} = 2,$

$\|x + y\|_{l_\infty} = \max\{|x_1 + y_1|, |x_2 + y_2|\} = \max\{|3|, |1|\} = 3.$

Следовательно,

$$\|x + y\| = \|x\| + \|y\| \quad (x, y \neq 0),$$

но $y \neq \alpha x$, где $\alpha > 0$.

3. Доказать, что пространство $L_1([a,\, b])$ не является строго нормированным пространством.

Решение. Рассмотрим для простоты пространство $L_1([-1,\, 1])$. Положим $x(t) = \begin{cases} 0, & t < 0, \\ 1, & t \geq 0, \end{cases}$ $y(t) = \begin{cases} -1, & t < 0, \\ 0, & t \geq 0. \end{cases}$ Тогда

$$\|x(\cdot)\|_{L_1([-1,1])} = \int\limits_{-1}^{1} |x(t)|\, dt = \int\limits_{0}^{1} 1\, dt = 1,$$

$$\|y(\cdot)\|_{L_1([-1,1])} = \int\limits_{-1}^{1} |x(t)|\, dt = \int\limits_{-1}^{0} |-1|\, dt = 1,$$

$$\|x(\cdot) + y(\cdot)\|_{L_1([-1,1])} = \int\limits_{-1}^{1} |x(t) + y(t)|\, dt = \int\limits_{-1}^{1} |\mp 1|\, dt = 2.$$

Следовательно,

$$\|x + y\| = \|x\| + \|y\| \quad (x, y \neq 0),$$

но $y \neq \alpha x$, где $\alpha > 0$.

4. Доказать, что пространство $L_\infty([a,\, b])$ и $C([a,\, b])$ не являются строго нормированными пространствами.

Решение. Рассмотрим для простоты случай, когда отрезок $[a,\, b] = [-1,\, 1]$. Возьмем две линейно независимые функции $x(\cdot), y(\cdot) \in C([-1,\, 1])$, модули которых принимают свои максимальные значения в одной и той же точке τ отрезка $[-1,\, 1]$, причем $\operatorname{sign} x(\tau) = \operatorname{sign} y(\tau)$. Например, $x(t) \equiv 1$, $y(t) = t$. Тогда

$$\|x(\cdot)\|_{C([-1,\,1])} = \max_{t\in[-1,\,1]} |x(t)| = \max_{t\in[-1,\,1]} |1| = 1,$$

$$\|y(\cdot)\|_{C([-1,\,1])} = \max_{t\in[-1,\,1]} |y(t)| = \max_{t\in[-1,\,1]} |t| = y(1) = 1,$$

$$\|x(\cdot) + y(\cdot)\|_{C([-1,\,1])} = \max_{t\in[-1,\,1]} |x(t) + y(t)|\, dt = \max_{t\in[-1,\,1]} |1 + t|\, dt = 2.$$

Следовательно,

$$\|x + y\| = \|x\| + \|y\| \quad (x, y \neq 0),$$

но $y \neq \alpha x$, где $\alpha > 0$.

3 Аппроксимация в гильбертовом пространстве

3.1 Проекция вектора на подпространство

Теорема 4. *Пусть H — гильбертово пространство, $L \subset H$ — подпространство, точка $x \notin L$. Если в L существует точка $y \in L$, наименее удаленная от x $\big(d(x, L) = \|x - y\|\big)$, то вектор $x - y$ ортогонален каждому вектору f из L, т. е.*

$$\langle x - y, f\rangle = 0 \quad \forall\ f \in L \iff \langle x - y, L\rangle = 0 \iff x - y \perp L.$$

Доказательство. От противного. Предположим противное, что в подпространстве L существует вектор g, для которого

$$\langle x - y, g\rangle = \sigma \neq 0.$$

Возьмем вектор $y^* = y + \dfrac{\sigma}{\langle g, g\rangle} g$. Тогда $y^* \in L$ и

$$\|x - y^*\|^2 = \left\langle x - y - \frac{\sigma}{\langle g, g\rangle} g, x - y - \frac{\sigma}{\langle g, g\rangle} g \right\rangle =$$

$$= \|x - y\|^2 - \frac{2\sigma}{\langle g, g\rangle}\langle g, x - y\rangle + \frac{\sigma^2}{\langle g, g\rangle} \overset{\langle x-y,g\rangle=\sigma}{=} \|x - y\|^2 - \frac{\sigma^2}{\langle g, g\rangle}$$

и, значит,

$$\|x - y^*\| < \|x - y\|,$$

а это противоречит предположению, что y есть наименее удаленная от x точка подпространства L. $\square$

Вектор y из L, обладающий тем свойством, что разность $x - y$ ортогональна L, называется *проекцией* x на L.

Следствие 1. *Если расстояние* $d(x,L)=\|x\|$, *то вектор* $x \perp L$.

Доказательство. В этом случае в теореме 4 вектор $y = 0$. $\square$

Следствие 2. *Пусть* L_1, L_2 *— линейные подпространства в* H, *пространство* $L_1 \perp L_2$. *Если расстояние* $d(x, L_1 + L_2) = d(x, L_1)$, *то вектор* $x \perp L_2$.

Доказательство. Предположим, что $\hat{l}_1$ — наименее удаленная от x точка подпространства L_1. Тогда по условию следствия

$$\|x - \hat{l}_1\| \overset{\mathrm{def}\,\hat{l}_1}{=} d(x, L_1) = d(x, L_1 + L_2) \overset{\mathrm{def}}{=}$$

$$= \inf_{\substack{l_1 \in L_1 \\ l_2 \in L_2}} \|x - l_1 - l_2\| = \inf_{\substack{l_1 \in L_1 \\ l_2 \in L_2}} \|x - \hat{l}_1 - l_1 - l_2\|.$$

Поскольку гильбертово пространство является строго нормированным пространством, то приближающий элемент единственен, значит, $l_1 = l_2 = 0$. Тем самым имеем равенство:

$$\inf_{\substack{l_1 \in L_1 \\ l_2 \in L_2}} \|x - \hat{l}_1 - l_1 - l_2\| = \inf_{l_2 \in L_2} \|x - \hat{l}_1 - l_2\| = \|x - \hat{l}_1\|.$$

$$d(x - \hat{l}_1, L_1 + L_2) = d(x - \hat{l}_1, L_2) = \|x - \hat{l}_1\|.$$

Поэтому по следствию 1 вектор $x - \hat{l}_1 \perp L_2$.
Поскольку $\hat{l}_1 \perp L_2$, то отсюда следует, что вектор $x \perp L_2$. $\square$

3.2 Определитель Грама

В том случае, когда подпространство конечномерно и образовано линейно независимыми векторами $f_1, f_2, \dots, f_n$ мы можем, пользуясь доказанным предложением, фактически найти вектор $y = \lambda_1 f_1 + \lambda_2 f_2 + \dots + \lambda_n f_n$, наименее уклоняющийся от вектора x (существование этого вектора доказано выше). Действительно, вектор y есть проекция x на L, и значит, он должен удовлетворять уравнениям

$$\langle x - y, f_k\rangle = 0 \quad \big(\Leftrightarrow \langle y, f_k\rangle = \langle x, f_k\rangle\big), \quad k = 1, 2, \dots, n, \tag{1}$$

которые в подробной записи имеют вид

$$\begin{array}{c} \lambda_1\langle f_1, f_1\rangle + \lambda_2\langle f_2, f_1\rangle + \ldots + \lambda_n\langle f_n, f_1\rangle = \langle x, f_1\rangle, \\ \lambda_1\langle f_1, f_2\rangle + \lambda_2\langle f_2, f_2\rangle + \ldots + \lambda_n\langle f_n, f_2\rangle = \langle x, f_2\rangle, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \lambda_1\langle f_1, f_n\rangle + \lambda_2\langle f_2, f_n\rangle + \ldots + \lambda_n\langle f_n, f_n\rangle = \langle x, f_n\rangle, \end{array} \tag{1$'$}$$

и представляют систему линейных уравнений для нахождения коэффициентов λ_k.

Детерминант этой системы, т. е.

$$\begin{vmatrix} \langle f_1, f_1\rangle & \langle f_2, f_1\rangle & \dots & \langle f_n, f_1\rangle \\ \langle f_1, f_2\rangle & \langle f_2, f_2\rangle & \dots & \langle f_n, f_2\rangle \\ \dots & \dots & \dots & \dots \\ \langle f_1, f_n\rangle & \langle f_2, f_n\rangle & \dots & \langle f_n, f_n\rangle \end{vmatrix} = G(f_1, f_2, \dots, f_n),$$

носит название *детерминанта Грама* системы векторов $f_1, \dots, f_n$.

Теорема 5. *Пусть H — гильбертово пространство, подпространство $L = \mathrm{lin}\{f_1, f_2, \dots, f_n\}$ натянуто на n линейно независимых векторов из H, точка $x \in H \setminus L$. Тогда*

$$d^2(x, L) = \frac{G(x, f_1, \dots, f_n)}{G(f_1, \dots, f_n)},$$

и детерминант Грама $G(f_1, \dots, f_n) > 0$.

Доказательство. Поскольку элемент наилучшего приближения конечномерным подпространством существует, то система (2) всегда имеет решение. Так как гильбертово пространство H строго нормировано, то при любом векторе x это решение единственно. Отсюда вытекает, что детерминант Грама не равен нулю.

Найдем выражение для квадрата погрешности, с которой вектор y аппроксимирует вектор x. Имеем

$$d^2 = \langle x-y, x-y\rangle = \langle x-y, x\rangle - \langle x-y, y\rangle \overset{(1)}{=} \langle x-y, x\rangle = \langle x, x\rangle - \langle y, x\rangle$$

$$\Longleftrightarrow \lambda_1\langle f_1, x\rangle + \lambda_2\langle f_2, x\rangle + \ldots + \lambda_n\langle f_n, x\rangle = \langle x, x\rangle - d^2.$$

Присоединяя это уравнение к системе (2), состоящей из n линейных уравнений с n неизвестными $\lambda_1, \lambda_2, \dots, \lambda_n$, получим систему из $(n+1)$-го линейного уравнения с n неизвестными, которая имеет то же решение. Значит, ранг расширенной матрицы не увеличился. Поэтому

$$\begin{vmatrix} \langle f_1, x\rangle & \dots & \langle f_n, x\rangle & d^2 - \langle x, x\rangle \\ \langle f_1, f_1\rangle & \dots & \langle f_n, f_1\rangle & -\langle x, f_1\rangle \\ \dots & \dots & \dots & \dots \\ \langle f_1, f_n\rangle & \dots & \langle f_n, f_n\rangle & -\langle x, f_n\rangle \end{vmatrix} = 0 \iff$$

$$\begin{vmatrix} d^2 - \langle x, x\rangle & \langle f_1, x\rangle & \dots & \langle f_n, x\rangle \\ -\langle x, f_1\rangle & \langle f_1, f_1\rangle & \dots & \langle f_n, f_1\rangle \\ \dots & \dots & \dots & \dots \\ -\langle x, f_n\rangle & \langle f_1, f_n\rangle & \dots & \langle f_n, f_n\rangle \end{vmatrix} = 0 \iff$$

$$\begin{vmatrix} d^2 & \langle f_1, x\rangle & \dots & \langle f_n, x\rangle \\ 0 & \langle f_1, f_1\rangle & \dots & \langle f_n, f_1\rangle \\ \hdotsfor{4} \\ 0 & \langle f_1, f_n\rangle & \dots & \langle f_n, f_n\rangle \end{vmatrix} = \begin{vmatrix} \langle x, x\rangle & \langle f_1, x\rangle & \dots & \langle f_n, x\rangle \\ \langle x, f_1\rangle & \langle f_1, f_1\rangle & \dots & \langle f_n, f_1\rangle \\ \hdotsfor{4} \\ \langle x, f_n\rangle & \langle f_1, f_n\rangle & \dots & \langle f_n, f_n\rangle \end{vmatrix}$$

$$\iff d^2 \cdot G(f_1, \dots, f_n) = G(x, f_1, \dots, f_n).$$

Таким образом,

$$d^2 = \min_{\alpha_k} \|x - \alpha_1 f_1 - \dots - \alpha_n f_n\|^2 = \frac{G(x, f_1, \dots, f_n)}{G(f_1, \dots, f_n)}. \qquad (3)$$

Из этого соотношения и из того, что $G(f) = \langle f, f\rangle > 0$ $(f \neq 0)$, последовательно получаем, что $\min\limits_{\alpha_2} \|f_1 - \alpha_2 f_2\|^2 = \frac{G(f_1, f_2)}{G(f_2)} > 0 \Rightarrow G(f_1, f_2) > 0, \dots, G(f_1, \dots, f_n) > 0.$ $\square$

3.3 Свойства детерминанта Грама

Свойство 1. *Детерминант Грама системы линейно зависимых векторов равен нулю.*

Доказательство. Свойство следует непосредственно из системы $(1')$. □

Свойство 2. *Неравенство* $G(f_1, f_2) \geq 0$ *эквивалентно неравенству Коши-Буняковского.*

Доказательство. Действительно,

$$G(f_1, f_2) \geq 0 \iff \begin{vmatrix} \langle f_1, f_1 \rangle & \langle f_2, f_1 \rangle \\ \langle f_1, f_2 \rangle & \langle f_2, f_2 \rangle \end{vmatrix} \geq 0 \iff$$
$$\langle f_1, f_1 \rangle \langle f_2, f_2 \rangle - \langle f_1, f_2 \rangle \langle f_2, f_1 \rangle \geq 0 \iff$$
$$\langle f_2, f_2 \rangle \geq \langle f_1, f_2 \rangle \langle f_2, f_1 \rangle \iff \|f_1\| \cdot \|f_2\| \geq |\langle f_1, f_2 \rangle|.$$

Значит, полученный результат о неотрицательности детерминанта Грама можно рассматривать как обобщение неравенства Коши-Буняковского. □

Свойство 3. *Имеет место неравенство*

$$G(f_1, \dots, f_n) \leq G(f_1, \dots, f_m) G(f_{m+1}, \dots, f_n), \tag{$*$}$$

$m = 1, \dots, n-1$. *Причем, знак* $=$ *в соотношении* $(*)$ *имеет место тогда и только тогда, когда каждый из векторов* $f_1, \dots, f_m$ *ортогонален каждому из векторов* $f_{m+1}, f_{m+2}, \dots, f_n$.

Доказательство. При $k = 1, 2, \dots, m-1$ для расстояний имеем очевидные неравенства:

$$d(f_k, \operatorname{lin}\{f_{k+1}, \dots, f_m\} + \operatorname{lin}\{f_{m+1}, \dots, f_n\}) \leq d(f_k, \operatorname{lin}\{f_{k+1}, \dots, f_m\}), \tag{4_1}$$

и

$$d(f_m, \operatorname{lin}\{f_{m+1}, \dots, f_n\}) \leq \|f_m\|. \tag{4_2}$$

По теореме 5 эти неравенства переписываются в виде

$$\frac{G(f_k, f_{k+1}, \dots, f_n)}{G(f_{k+1}, \dots, f_n)} \le \frac{G(f_k, f_{k+1}, \dots, f_m)}{G(f_{k+1}, \dots, f_m)} \iff$$

$$\frac{G(f_k, f_{k+1}, \dots, f_n)}{G(f_k, f_{k+1}, \dots, f_m)} \le \frac{G(f_{k+1}, \dots, f_n)}{G(f_{k+1}, \dots, f_m)}, \quad k = 1, 2, \dots, m-1,$$

и

$$\frac{G(f_m, f_{m+1}, \dots, f_n)}{G(f_{m+1}, \dots, f_n)} \le G(f_m) \iff$$

$$\frac{G(f_m, f_{m+1}, \dots, f_n)}{G(f_m)} \le G(f_{m+1}, f_{m+2}, \dots, f_n),$$

или

$$\frac{G(f_1, \dots, f_n)}{G(f_1, \dots, f_m)} \le \frac{G(f_2, \dots, f_n)}{G(f_2, \dots, f_m)} \le \dots \le \frac{G(f_m, f_{m+1}, \dots, f_n)}{G(f_m)} \le$$

$$\le G(f_{m+1}, f_{m+2}, \dots, f_n).$$

Отсюда следует доказываемое неравенство

$$G(f_1, \dots, f_n) \le G(f_1, \dots, f_m) G(f_{m+1}, f_{m+2}, \dots, f_n).$$

В силу наших рассмотрений, знак $=$ в соотношении $(*)$ имеет место тогда и только тогда, когда все неравенства (4_1), (4_2) превращаются в равенства, которые перепишем в виде:

$$d(f_k, \mathrm{lin}\{f_{k+1}, \dots, f_m\} + L) = d(f_k, \mathrm{lin}\{f_{k+1}, \dots, f_m\}),\ k = 1, \dots, m-1.$$

и

$$d(f_m, L) = \|f_m\|,$$

где $L := \mathrm{lin}\{f_{m+1}, \dots, f_n\}$. Из последнего соотношения в силу следствия 1 теоремы 4 вытекает, что $f_m \perp L$, а предыдущие соотношения для $k = m-1, m-2, \dots, 1$ последовательно дают в силу следствия 2 теоремы 4

$$\begin{array}{ll}
d(f_{m-1}, \mathrm{lin}\{f_m\} + L) = d(f_{m-1}, \mathrm{lin}\{f_m\}) & \Rightarrow f_{m-1} \perp L, \\
d(f_{m-2}, \mathrm{lin}\{f_{m-1}, f_m\} + L) = d(f_{m-2}, \mathrm{lin}\{f_{m-1}, f_m\}) & \Rightarrow f_{m-2} \perp L, \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\
d(f_1, \mathrm{lin}\{f_2, \dots, f_m\} + L) = d(f_1, \mathrm{lin}\{f_2, \dots, f_m\}) & \Rightarrow f_1 \perp L
\end{array}$$

Таким образом, свойство 3 доказано. $\square$

3.4 Неравенство Адамара

Из неравенства $(*)$ вытекает далее, что

$$G(f_1, \ldots, f_n) \leq \langle f_1, f_1 \rangle \cdot \ldots \cdot \langle f_n, f_n \rangle \tag{5}$$

и знак $=$ здесь достигается тогда и только тогда, когда

$$\langle f_i, f_k \rangle = 0, \quad i \neq k;\ i, k = 1, \ldots, n.$$

Неравенство $(*)$ содержит как частный случай классическую теорему Адамара о детерминантах.

Теорема (неравенство Адамара)**.** *Пусть матрица* $A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$, *где* a_{ik} *— любые действительные числа. Тогда имеет место неравенство:* $(\det A)^2 \leq \prod\limits_{i=1}^{n} \left(\sum\limits_{j=1}^{n} a_{ij}^2 \right)$.

Доказательство. Поскольку $\det A = \det A^*$, то

$$(\det A)^2 = \det(AA^*) = \begin{vmatrix} \sum\limits_{k=1}^{n} |a_{1k}|^2 & \sum\limits_{k=1}^{n} a_{1k}a_{2k} & \ldots & \sum\limits_{k=1}^{n} a_{1k}a_{nk} \\ \sum\limits_{k=1}^{n} a_{2k}a_{1k} & \sum\limits_{k=1}^{n} |a_{2k}|^2 & \ldots & \sum\limits_{k=1}^{n} a_{2k}a_{nk} \\ \ldots & \ldots & \ldots & \ldots \\ \sum\limits_{k=1}^{n} a_{nk}a_{1k} & \sum\limits_{k=1}^{n} a_{nk}a_{2k} & \ldots & \sum\limits_{k=1}^{n} |a_{nk}|^2 \end{vmatrix} =$$

$$= \begin{vmatrix} \langle f_1, f_1 \rangle & \langle f_1, f_2 \rangle & \ldots & \langle f_1, f_n \rangle \\ \langle f_2, f_1 \rangle & \langle f_2, f_2 \rangle & \ldots & \langle f_2, f_n \rangle \\ \ldots & \ldots & \ldots & \ldots \\ \langle f_n, f_1 \rangle & \langle f_n, f_2 \rangle & \ldots & \langle f_n, f_n \rangle \end{vmatrix} = G(f_1, \ldots, f_n),$$

где $f_k = (a_{k1}, a_{k2}, \ldots, a_{kn})$.

На основании неравенства (5) следует неравенство Адамара:

$$(\det A)^2 = G(f_1, \ldots, f_n) \leq \langle f_1, f_1 \rangle \cdot \ldots \cdot \langle f_n, f_n \rangle =$$

$$= \sum_{k=1}^{n} |a_{1k}|^2 \sum_{k=1}^{n} |a_{2k}|^2 \cdot \ldots \cdot \sum_{k=1}^{n} |a_{nk}|^2,$$

причем знак равенства будет иметь место только в том случае, когда

$$\langle f_\lambda, f_\mu \rangle = 0 \iff \sum_{1}^{n} a_{\lambda k} a_{\mu k} = 0, \quad \lambda \neq \mu. \qquad \square$$

Из неравенства Адамара вытекает следующая оценка определителя матрицы.

Следствие. *Пусть матрица* $A = (a_{ij})_{i,j=1}^{n}$, $|a_{ik}| \leq M$. *Тогда* $|\det A| \leq n^{\frac{n}{2}} M^n$.

Дадим еще одно доказательство теоремы Адамара с помощью теории экстремальных задач.

Теорема (неравенство Адамара)**.** *Пусть матрица* $X = \left(x_i^j\right)_{i,j=1}^n = \begin{pmatrix} x_1^1 & \dots & x_1^n \\ \multicolumn{3}{c}{\dotfill} \\ x_n^1 & \dots & x_n^n \end{pmatrix}$ *— произвольная квадратная матрица со строками* $x_i = (x_i^1, \dots, x_i^n)$, $i = 1, \dots, n$. *Тогда имеет место неравенство:* $(\det X)^2 \le \prod\limits_{i=1}^n |x_i|^2$.

Доказательство. Рассмотрим следующую экстремальную задачу: найти наибольший объем n-мерного параллелепипеда, у которого все ребра имеют единичную длину. Эта задача допускает такую формализацию:

$$D(x_1, \dots, x_n) = \det X \to \sup; \ |x_i|^2 = 1, \quad i = 1, \dots, n, \qquad (P)$$

здесь $|x_i|^2 = \sum\limits_{j=1}^n \left(x_i^j\right)^2$. Имеем гладкую экстремальную задачу с ограничениями типа равенств. Она решается с помощью принципа Лагранжа. Согласно нему существует вектор множителей Лагранжа $\lambda = (\lambda_0, \dots, \lambda_n) \ne 0$ такой, что для функция Лагранжа задачи (P) $\mathscr{L} = \lambda_0 D + \sum\limits_{j=1}^n \lambda_i |x_i|^2$ выполняются условия стационарности:

$$\widehat{\mathscr{L}_{x_i}} = 0 \iff \lambda_0 \widehat{D}_{x_i} + 2\lambda_i \hat{x}_i = 0, \quad i = 1, \dots, n. \qquad (*)$$

Если $\lambda_0 = 0$, то $\lambda_i = 0$, т. к. $\hat{x}_i \ne 0$ (напомним, $|\hat{x}_i| = 1$). Поскольку вектор λ ненулевой, то $\lambda_0 \ne 0$.

Положим для удобства $\lambda_0 = -2$. Тогда условие стационарности, деленное на два, перепишется в виде:

$$\widehat{D}_{x_i} = \lambda_i \hat{x}_i, \quad i = 1, \dots, n. \qquad (*)$$

Поскольку определитель матрицы раскладывается по любой строке как сумма произведений элементов строки матрицы на из алгебраические дополнения $\widehat{D} = \langle \widehat{A}_i, \hat{x}_i \rangle$, то производная определителя по строке — это вектор, составленный из алгебраических дополнений к элементам этой строки: $\widehat{D}_{x_i} = \widehat{A}_i = (\widehat{A}_i^1, \ldots, \widehat{A}_i^n)$. Умножая уравнения в условии стационарности $(*)$ на $\hat{x}_i$, получаем:

$$\langle \widehat{D}_{x_i}, \hat{x}_i \rangle = \lambda_i \langle \hat{x}_i, \hat{x}_i \rangle = \lambda_i, \ i = 1, \ldots, n.$$

Откуда $\lambda_i = \langle \widehat{D}_{x_i}, \hat{x}_i \rangle = \langle \widehat{A}_i, \hat{x}_i \rangle = \widehat{D}$. Соответственно, из $(*)$

$$\hat{x}_i = \frac{\widehat{D}_{x_i}}{\lambda_i} = \frac{\widehat{A}_i}{\widehat{D}}, \ i = 1, \ldots, n, \iff \widehat{X} = \left(\frac{\widehat{A}_i^j}{\widehat{D}} \right)_{i,j=1}^n.$$

Из алгебры известно, что матрица $\left(\frac{\widehat{A}_i^j}{\widehat{D}} \right)_{i,j=1}^n$ является матрицей обратной к транспонированной матрице $\widehat{X}^*$, т. е. $\widehat{X} = (\widehat{X}^*)^{-1} \Leftrightarrow \widehat{X}\widehat{X}^* = I$. Таким образом, $\widehat{X}$ — ортогональная матрица и, следовательно, $|\det X| = 1$.

Из доказанного и однородности определителя и умножения строки матрицы на число следует неравенство Адамара. $\square$

4 Полиномы, наименее уклоняющиеся от нуля в пространстве $L_p([-\pi,\, \pi])$

4.1 Определение

Пусть $L = L_{2n-1} = \mathrm{lin}\{1, \cos t, \sin t, \ldots, \cos(n-1)t, \sin(n-1)t\}$ — $2n-1$-мерное пространство, тригонометрических полиномов степени не выше $n-1$. Рассмотрим задачу приближения заданного элемента $x(t) = \alpha \cos nt + \beta \sin nt$ подпространством L в пространстве $L_p([-\pi,\, \pi])$, $1 \le p \le \infty$:

$$d(x, L) = \inf_{y \in L} \|x - y\|_{L_p} \iff \int\limits_{-\pi}^{\pi} |x(t) - y(t)|^p dt \to \min, \qquad (P')$$

т. е. среди полиномов вида

$$P_n(t) = \alpha \cos nt + \beta \sin nt +$$
$$+ a_{n-1} \cos(n-1)t + b_{n-1} \sin(n-1)t + \ldots + a_1 \cos t + b_1 \sin t + a_0,$$

ищем полином наименее уклоняющийся от нуля в метрике пространства L_p, т. е. доставляющий минимум интегралу: $\int\limits_{-\pi}^{\pi} |P_n(t)|^p dt$.

4.2 Теорема Бернштейна

Теорема 6 (Бернштейн С.Н.). *Полиномом, наименее уклоняющимся от нуля в пространстве $L_p([-\pi, \pi])$, $1 \le p \le \infty$, является полином $\widehat{P}_n(t) = \alpha \cos nt + \beta \sin nt$.*

Доказательство. Примем вначале, что $1 < p < \infty$, и воспользуемся тем, что при этом условии пространство L_p есть строго нормированное пространство. Предположим, что полиномом, наименее уклоняющимся от нуля является полином $P_n(t) = \alpha \cos nt + \beta \sin nt +$
$+ a_{n-1} \cos(n-1)t + b_{n-1} \sin(n-1)t + \ldots + a_1 \cos t + b_1 \sin t + a_0$. Так как, в силу периодичности,

$$\int\limits_{-\pi}^{\pi} |P_n(t)|^p dt = \int\limits_{-\pi+\varphi}^{\pi+\varphi} |P_n(t)|^p dt = \int\limits_{-\pi}^{\pi} |P_n(t+\varphi)|^p dt \quad \forall\, \varphi \in \mathbb{R},$$

то полином $P_n(t+\varphi)$ для любого $\varphi \in \mathbb{R}$ также является полиномом наименее уклоняющимся от нуля. Поскольку

$$\begin{aligned} P_n\big(t + \tfrac{2\pi}{n}\big) &= \alpha \cos n\big(t + \tfrac{2\pi}{n}\big) + \beta \sin n\big(t + \tfrac{2\pi}{n}\big) + \\ &+ a_{n-1} \cos(n-1)\big(t + \tfrac{2\pi}{n}\big) + b_{n-1} \sin(n-1)\big(t + \tfrac{2\pi}{n}\big) + \ldots \\ &+ a_1 \cos\big(t + \tfrac{2\pi}{n}\big) + b_1 \sin\big(t + \tfrac{2\pi}{n}\big) + a_0 = \end{aligned}$$

$\big($по тригонометрическим формулам $a_k \cos k(t+\varphi) + b_k \sin k(t+\varphi) =$
$= a_k(\cos kt \cos k\varphi - \sin kt \sin k\varphi) + b_k(\sin kt \cos k\varphi + \cos kt \sin k\varphi) =$
$= (a_k \cos k\varphi + b_k \sin k\varphi) \cos kt + (-a_k \sin k\varphi + b_k \cos k\varphi) \sin kt\big)$

$$\begin{aligned} &= \alpha \cos nt + \beta \sin nt + \\ &+ \big(a_{n-1} \cos(n-1)\tfrac{2\pi}{n} + b_{n-1} \sin(n-1)\tfrac{2\pi}{n}\big) \cos(n-1)t + \\ &+ \big(-a_{n-1} \sin(n-1)\tfrac{2\pi}{n} + b_{n-1} \cos(n-1)\tfrac{2\pi}{n}\big) \sin(n-1)t + \ldots \\ &+ \big(a_1 \cos\tfrac{2\pi}{n} + b_1 \sin\tfrac{2\pi}{n}\big) \cos t + \big(-a_1 \sin\tfrac{2\pi}{n} + b_1 \cos\tfrac{2\pi}{n}\big) \sin t + a_0, \end{aligned}$$

то $P_n\bigl(t+\frac{2\pi}{n}\bigr) \equiv P_n(t)$ в силу единственности полинома, наименее уклоняющегося от нуля в строго нормированном пространстве. Следовательно, для $k = 1, \dots, n-1$

$$\begin{cases} a_k \cos k\frac{2\pi}{n} + b_k \sin k\frac{2\pi}{n} = a_k, \\ -a_k \sin k\frac{2\pi}{n} + b_k \cos k\frac{2\pi}{n} = b_k, \end{cases} \iff$$

$$\begin{cases} a_k\bigl(\cos k\frac{2\pi}{n} - 1\bigr) + b_k \sin k\frac{2\pi}{n} = 0, \\ -a_k \sin k\frac{2\pi}{n} + b_k\bigl(\cos k\frac{2\pi}{n} - 1\bigr) = 0, \end{cases} \iff \begin{cases} a_k = 0, \\ b_k = 0. \end{cases}$$

Аналогично, поскольку

$$P_n\bigl(t+\tfrac{\pi}{n}\bigr) = \alpha \cos n\bigl(t+\tfrac{\pi}{n}\bigr) + \beta \sin n\bigl(t+\tfrac{\pi}{n}\bigr) + a_0 = \\ -\alpha \cos nt - \beta \sin nt + a_0,$$

то $P_n\bigl(t+\frac{\pi}{n}\bigr) \equiv -P_n(t)$ в силу единственности полинома, наименее уклоняющегося от нуля в строго нормированном пространстве, то $a_0 = 0$. Таким образом, единственным решением данной задачи при $1 < p < \infty$ является полином $\widehat{P}_n(t)$.

Это же выражение, как легко видеть из соображений непрерывности, будет решением также при $p = 1$ (и притом единственным, в силу общих предложений, доказываемых ниже). $\square$

5 Теорема Чебышева об альтернансе

5.1 Теоремы об очистке

Пусть f_1 и f_2 — выпуклые непрерывные функции. Тогда функция $\max\{f_1, f_2\}(x)$ является выпуклой функцией. Рассмотрим следующую выпуклую задачу без ограничений

$$\max\{f_1, f_2\}(\hat{x}) \to \min .$$

Из курса выпуклого анализа известно, что необходимым и достаточным условием минимума в точке $\hat{x}$ является условие $0 \in \partial \max\{f_1, f_2\}(x)$.

Субдифференциал максимума описывается с помощью теоремы Дубовицкого–Милютина.

Теорема (Дубовицкий–Милютин)**.** *Пусть f_1 и f_2 — выпуклые функции, непрерывные в точке $\hat{x}$, $f_1(\hat{x}) = f_2(\hat{x})$. Тогда*

$$\partial \max\{f_1, f_2\}(\hat{x}) = \operatorname{conv}(\partial f_1(\hat{x}) \cup \partial f_2(\hat{x})).$$

Тогда если $f_1(\hat{x}) = f_2(\hat{x})$, то по теореме Дубовицкого–Милютина условие $0 \in \partial \max\{f_1, f_2\}(\hat{x})$ равносильно тому, что существуют неотрицательные числа α_1, α_2, $\alpha_1 + \alpha_2 = 1$ такие, что $\alpha_1 y_1 + \alpha_2 y_2 = 0$, где $y_1 \in \partial f_1(\hat{x})$, $y_2 \in \partial f_2(\hat{x})$.

Если $f_1(\hat{x}) \neq f_2(\hat{x})$ (возьмем, для определенности, $f_1(\hat{x}) > f_2(\hat{x})$), то $\max\{f_1, f_2\}(\hat{x}) = f_1(\hat{x})$ и условие $0 \in \partial \max\{f_1, f_2\}(\hat{x})$ равносильно тому, что $0 \in \partial f_1(\hat{x})$. Этот случай тоже укладывается в схему предыдущего абзаца при $\alpha_1 = 1$, $\alpha_2 = 0$, то есть $y_1 = 0$, $y_2 \in \partial f_2(\hat{x})$ — любое.

Для решения более сложных задач вида

$$\max_{t \in T} f(t,x) \to \min$$

вместо теоремы Дубовицкого–Милютина используется теорема об очистке.

Теорема (об очистке)**.** *Пусть T — компакт, $f : T \times \mathbb{R}^n \to \mathbb{R}$ — функция, обладающая следующим свойствами: a) $f(t,\cdot)$ выпукла и непрерывна на $\mathbb{R}^n$ для любого $t \in T$, b) $f(\cdot,x)$ полунепрерывна сверху для любого $x \in \mathbb{R}^n$, c) $\inf\limits_{x \in \mathbb{R}^n} \max\limits_{t \in T} f(t,x) =: M > -\infty$. Тогда существуют число $s \in \mathbb{N}$, $s \le n+1$ и набор точек $\tau_1, \ldots, \tau_s \in T$, таких, что $M = \inf\limits_{x \in \mathbb{R}^n} \max\limits_{1 \le i \le s} f(\tau_i, x)$.*

Таким образом, компакт T можно “очистить” до s точек ($s \le n+1$), что минимакс по всему семейству $\{f(t,\cdot)\}_{t \in T}$ окажется равным минимаксу лишь семейства $\{f(t_i,\cdot)\}_{i=1}^s$, состоящего лишь из конечного числа функций.

Теорема (об очистке (субдифференциальная форма))**.** *Пусть T — компакт, $f : T \times \mathbb{R}^n \to \mathbb{R}$ — функция, обладающая следующим свойствами: a) $f(t,\cdot)$ выпукла и непрерывна на $\mathbb{R}^n$ для любого $t \in T$, b) $f(\cdot,x)$ полунепрерывна сверху для любого $x \in \mathbb{R}^n$. Положим $f(x) = \max\limits_{t \in T} f(t,x)$ и возьмем точку $\hat{x} \in R^n$. Тогда существуют число $s \in \mathbb{N}$, $s \le n+1$, и набор точек $\tau_1, \ldots, \tau_s \in T$, таких, что $f(\tau_i, \hat{x}) = f(\hat{x})$, $i = 1, \ldots, s$, и*

$$\partial f(\hat{x}) = \operatorname{conv}\{\partial_x f(\tau_1, \hat{x}), \ldots, \partial_x f(\tau_s, \hat{x})\}.$$

5.2 Альтернанс. Теорема об альтернансе

Рассмотрим задачу о приближении заданной функции $h(\cdot) \in C([t_0,\, t_1])$ пространством $\mathcal{P}_n = \left\{ p(t) = \sum\limits_{k=1}^{n} x_k t^{k-1} \right\}$ полиномов степени не выше $n-1$ в пространстве $C([t_0,\, t_1])$. Полином наименьшего уклонения $\widehat{p}(\cdot)$ определяется соотношением

$$\|h(\cdot) - \widehat{p}(\cdot)\|_{C([t_0,\, t_1])} = \min_{p(\cdot)\in\mathcal{P}_n} \|h(\cdot) - p(\cdot)\|_{C([t_0,\, t_1])}.$$

Теорема (Чебышева об альтернансе)**.** *Для того чтобы полином $\widehat{p}(\cdot)$ степени $\le n-1$ наименее уклонялся от непрерывной функции $h(\cdot)$ в метрике пространства $C([t_0,\, t_1])$ необходимо и достаточно, чтобы нашлись $s \ge n+1$ точек, в которых разность $h(\cdot) - \widehat{p}(\cdot)$ принимает, чередуясь, свои максимальное и минимальное значения.*

Доказательство. Необходимость. Естественно предположить, что $h(\cdot) \notin \mathcal{P}_n$. Обозначим для вектора $x \in \mathbb{R}^n$

$$f(t,x) = |h(t) - p(t)| = \left| h(t) - \sum_{k=1}^{n} x_k t^{k-1} \right|,$$

$$f(x) = \max_{t\in[t_0,\, t_1]} f(t,x) = \|h(\cdot) - p(\cdot)\|_{C([t_0,\, t_1])}.$$

Тогда задача об отыскании полинома, наименее уклоняющегося от функции $h(\cdot)$, перепишется в виде:

$$f(x) \to \min.$$

Легко доказать, что функция f выпукла, непрерывна и растет на бесконечности. Значит, по следствию из теоремы Вейерштрасса решение $\hat{x}$ задачи существует и тогда $0 \in \partial f(\hat{x})$. Существование элемента наилучшего приближения конечномерным подпространством в линейном нормированном пространстве следует также по теореме 1. Выписанная задача является выпуклой задачей без ограничений. По необходимому и достаточному условию минимума в выпуклой задаче без ограничений $0 \in \partial f(\hat{x})$.

Применяем теорему об очистке. Согласно этой теореме, если $0 \in \partial f(\hat{x})$, то найдутся $s \le n+1$ точек $t_0 \le \tau_1 < \tau_2 < \ldots < \tau_s \le t_1$, s неотрицательных чисел $\alpha_1, \alpha_2, \ldots, \alpha_s$, $\sum\limits_{i=1}^{s} \alpha_i = 1$ и s векторов $y_i \in \partial_x f(\tau_i, \hat{x})$, таких, что

$$f(\tau_i, \hat{x}) = \|h(\cdot) - \widehat{p}(\cdot)\|_{C([t_0, t_1])}, \quad \sum_{i=1}^{s} \alpha_i y_i = 0. \tag{1}$$

Вследствие того, что функция $f(\tau_i, x)$ достигает свой положительный максимум в точке $\hat{x}$, то $f(\tau_i, \hat{x}) = M \neq 0$, и, следовательно, функция $f(\tau_i, x)$ (модуль в точке отличной от нуля) дифференцируема по x. Значит, ее субдифференциал $\partial_x f(\tau_i, \hat{x})$ совпадает с производной. Дифференцируя функцию $f(\tau_i, x) = \left| h(\tau_i) - \sum\limits_{k=1}^{n} x_k \tau_i^{k-1} \right|$ по вектору x в точке $\hat{x}$, получаем вектор

$$y_i = -\operatorname{sign}\left(h(\tau_i) - \widehat{p}(\tau_i)\right)(1, \tau_i, \ldots, \tau_i^{n-1}). \tag{2}$$

Из (1) и (2) вытекает, что система

$$\sum_{i=1}^{s} \alpha_i y_i = 0 \iff \sum_{i=1}^{s} \alpha_i \operatorname{sign}\left(h(\tau_i) - \widehat{p}(\tau_i)\right) \tau_i^{k-1} = 0, \quad k = 1, \ldots, n, \tag{3}$$

имеет нетривиальное решение $\left(\sum\limits_{i=1}^{s} \alpha_i = 1\right)$. При $s \le n$ эта система имеет лишь тривиальное решение, так как определитель системы является определителем Вандермонда. Значит, $s = n+1$.

Умножая k-тое уравнение системы (3) на x_k и складывая, получаем

$$\sum_{i=1}^{n+1} \alpha_i \operatorname{sign}\left(h(\tau_i) - \widehat{p}(\tau_i)\right) p(\tau_i) = 0 \quad \forall\, p(\cdot) \in \mathcal{P}_n. \tag{4}$$

Пусть $p_i(\cdot) \in \mathcal{P}_n$, $i = 1, \ldots, n$, такой полином, что $p_i(\tau_j) = \delta_{ij}$, $j = 1, \ldots, n$. Получим $p_i(\cdot)$ — полином степени $n-1$ с фиксированными $n-1$ нулем, причем все нули кратности 1, других нулей у него нет и полином меняет свои знаки, проходя через точки τ_j, $j \neq i$. Поэтому $\operatorname{sign} p_n(\tau_{n+1}) = 1$, $\operatorname{sign} p_{n-1}(\tau_{n+1}) = -1$, $\ldots$. т. е.

$$\operatorname{sign} p_i(\tau_{n+1}) = (-1)^{i-n}. \tag{5}$$

Поскольку $p_i(\tau_j) = \delta_{ij}$, $j = 1, \ldots, n$, то, подставляя в соотношение (4) вместо функции p функции p_i, получаем только по два слагаемых, отличных от нуля

$$\alpha_i \operatorname{sign}\big(h(\tau_i) - \widehat{p}(\tau_i)\big) + \alpha_{n+1}\operatorname{sign}\big(h(\tau_{n+1}) - \widehat{p}(\tau_{n+1})\big)p_i(\tau_{n+1}) = 0,$$

т. е.

$$\operatorname{sign}\big(h(\tau_i) - \widehat{p}(\tau_i)\big) = -\operatorname{sign}\big(h(\tau_{n+1}) - \widehat{p}(\tau_{n+1})\big)\operatorname{sign} p_i(\tau_{n+1}) =$$

$$= (-1)^{i-n+1}\operatorname{sign}\big(h(\tau_{n+1}) - \widehat{p}(\tau_{n+1})\big).$$

Это означает, что функция $h(\cdot) - \widehat{p}(\cdot)$, чередуясь, принимает свои максимальное и минимальное значения.

Достаточность. Пусть в точках $t_0 \leq \tau_1 < \tau_2 < \ldots < \tau_s \leq t_1$, $s \geq n+1$, функция $h(\cdot) - \widehat{p}(\cdot)$ принимает, чередуясь, максимальное и минимальное значения. Если допустить, что для некоторого $\overline{p}(\cdot) \in \mathcal{P}_n$

$$\|h(\cdot) - \overline{p}(\cdot)\|_{C([t_0,\, t_1])} < \|h(\cdot) - \widehat{p}(\cdot)\|_{C([t_0,\, t_1])},$$

то окажется, что числа $\widehat{p}(\tau_i) - \overline{p}(\tau_i)$ являются попеременно то отрицательными, то положительными $\geq n + 1$ раз. Но тогда функция $\widehat{p}(\cdot) - \overline{p}(\cdot)$ должна иметь $\geq n$ нулей. Это невозможно, так как функция $\widehat{p}(\cdot) - \overline{p}(\cdot)$ — полином степени $\leq n-1$. Получили противоречие, из которого следует требуемое. $\square$

Следствие. *Полином* $T_{n\infty}(t) = 2^{-(n-1)}\cos(n\arccos t) = t^n + \sum\limits_{k=0}^{n-1}\hat{x}_k t^k$ *является полиномом, наименее уклоняющимся от нуля в метрике* $C([-1,\,1])$ *среди всех полиномов степени* n *со старшим коэффициентом равным единице.*

Полином $T_{n\infty}(t)$ называется *полиномом Чебышева первого рода.* Для того, чтобы коэффициент при t^n был равен единице, коэффициент при $\cos(n\arccos t)$ надо взять $2^{-(n-1)}$. Этот факт будет доказан позже.

Доказательство. Полином $\cos(n\arccos t)$ достигает своего минимального и максимального значения равного ± 1 в точках $t_k = \cos\dfrac{k\pi}{n}$, $k = 0, 1, \ldots, n$. Всего $n+1$ точка. Значит, он имеет $n+1$-альтернанс и по теореме Чебышева об альтернансе является наилучшим приближением старшего члена t^n полиномами степени $n-1$. $\square$

6 Неравенство Бернштейна

Теорема 7. *Для любого тригонометрического полинома $x(\cdot)$ степени n имеет место неравенство*

$$\|\dot{x}(\cdot)\|_{C([-\pi,\,\pi])} \le n\|x(\cdot)\|_{C([-\pi,\,\pi])}.$$

Это неравенство является точным, и оно достигается на функции $t \to A\sin(nt+\gamma)$.

Доказательство. Возьмем произвольный степени n тригонометрический полином $x(t) = x_0 + \sum\limits_{k=1}^{n}(x_k \sin kt + x_{k+n}\cos kt)$. Рассмотрим следующую экстремальную задачу в пространстве $\mathbb{R}^{2n+1}$ векторов $x = (x_0, x_1, \dots, x_{2n})$:

$$f_0(x) \to \inf; \quad f_1(x) \le 1, \tag{P}$$

где $f_0(x) := \dot{x}(0) = \sum\limits_{k=1}^{n} kx_k \cos kt|_{t=0} = \sum\limits_{k=1}^{n} kx_k$, $f(t,x) = |x(t)| = \left|x_0 + \sum\limits_{k=1}^{n}(x_k \sin kt + x_{k+n}\cos kt)\right|$, $f_1(x) = \max\limits_{t\in[-\pi,\,\pi]} f(t,x)$. Это задача выпуклого программирования, так как функции f_0, f_1 выпуклые. Из теоремы Вейерштрасса и того обстоятельства, что множество допустимых элементов $D(P) = \{x \in \mathbb{R}^{2n+1} \mid f_1(x) \le 1\}$ компактно, вытекает, что решение задачи (P) существует. Обозначим его $\hat{x}$, функцию $x(\cdot)$, соответствующую вектору $\hat{x}$, обозначим $\hat{x}(\cdot)$.

Применим теорему Куна–Таккера, учитывая, что выполнено условие Слейтера ($\overline{x} = 0 \Rightarrow f_1(\overline{x}) < 1$). По теореме существуют множители Лагранжа $\lambda_0 = 1$, $\lambda \ge 0$, при которых функция Лагранжа задачи (P) $\mathscr{L}(x) = f_0(x) + \lambda(f_1(x) - 1)$ имеет минимум в точке $\hat{x}$: $\min_{x\in D(P)} \mathscr{L}(x) = \mathscr{L}(\hat{x})$. Следовательно, по теореме о необходимых и достаточных условиях минимума выпуклой функции $0 \in \partial\mathscr{L}(\hat{x})$. Кроме того выполняется условие дополняющей нежесткости: $\lambda(f_1(\hat{x}) - 1) = 0$.

Множитель $\lambda \ne 0$, так как в противном случае условие $0 \in \partial_x\mathscr{L}(\hat{x}) \Leftrightarrow 0 \in \partial f_0(\hat{x}) \Leftrightarrow 0 = (0,1,2,\dots,n,0,...,0)$ не выполняется.

Поскольку $0 \in \partial \mathscr{L}(\hat{x})$, то по теореме об очистке существует $s \le 2n+2$ точек $\tau_1 < \tau_2 < \ldots < \tau_s$, где $|\hat{x}(\tau_i)| = 1$, и s чисел $\alpha_1, \ldots, \alpha_s$, $\alpha_i > 0$, $\sum\limits_{i=1}^{s} \alpha_i = 1$, таких, что (более подробно это объяснялось в предыдущем пункте) выполняется система из $2n+1$ уравнения:

$$(0, 1, 2, \ldots, n, 0, ..., 0) + \lambda \sum_{i=1}^{s} \alpha_i \mathrm{sign}\, \hat{x}(\tau_i)(1, \sin \tau_i, ..., \cos n\tau_i) = 0.$$

откуда, умножая уравнения соответственно на $x_0, \ldots, x_{2n}$, приходим к равенству

$$\lambda \sum_{i=1}^{s} \alpha_i \mathrm{sign}\, \hat{x}(\tau_i) x(\tau_i) = -\dot{x}(0) \; \forall \, x(t) = x_0 + \sum_{k=1}^{n} (x_k \sin kt + x_{k+n} \cos kt). \tag{1}$$

Это тождество называется *тождеством Рисса*. Ясно, что число $s \le 2n$, ибо в точках τ_i функция $\hat{x}(\cdot)$ достигает максимума и, значит, $\dot{\hat{x}}(\tau_i) = 0$, а производная $\dot{\hat{x}}(\cdot)$ полинома степени n не может иметь больше $2n$ нулей. Если допустить, что $s < 2n$, то мы присоединим к $\{\tau_i\}_{i=1}^{s}$ точку нуль (а ее, очевидно, нет среди $\{\tau_i\}$, ибо иначе значение задачи (P) было бы равно $\dot{\hat{x}}(0) = 0$, что абсурдно) и еще быть может, несколько других точек, чтобы получилась система $\{\tau_i\}_{i=1}^{2n}$ из $2n$ разных точек. Существует полином $\overline{x}(\cdot)$ степени n, имеющий точки $\{\tau_i\}_{i=1}^{2n}$ своими нулями. Тогда $\dot{\overline{x}}(0) \ne 0$, ибо нуль полинома степени n с $2n$ нулями не может быть кратным. Получилось противоречие: с одной стороны $\dot{\overline{x}}(0) \ne 0$, а с другой (из (1)) $\dot{\overline{x}}(0) = 0$. Итак, $s = 2n$ и, значит, в каждой точке τ_i полином $1 - \hat{x}^2(\cdot)$ степени $2n$ имеет двукратный корень. Полином $\dot{\hat{x}}^2(\cdot)$ степени $2n$ в этих же точках также имеет двукратные корни. Значит, эти полиномы пропорциональны. Приравнивая старшие коэффициенты, приходим к уравнению $\dot{\hat{x}}^2(t) = n^2(1 - \hat{x}^2(t)$. Интегрируя его, получаем $x(t) = \sin nt$. Доказано, что $|\dot{x}(0)| \le n\|x(\cdot)\|_{C([-\pi,\,\pi])}$. Отсюда

$$|\dot{x}(\tau)| = |\dot{x}(t+\tau)|_{t=0}| \le n\|x(\cdot + \tau)\|_{C([-\pi,\,\pi])} = n\|x(\cdot)\|_{C([-\pi,\,\pi])}. \;\Box$$

7 Полиномы, наименее уклоняющиеся от нуля

7.1 Постановка задачи. Лемма о нулях

Общая постановка задачи о полиномах, наименее уклоняющихся от нуля, такова. Среди полиномов $p_{n-1}(\cdot) \in P_n$, т. е. среди полиномов вида $p_{n-1}(t) = \sum\limits_{k=0}^{n-1} x_k t^k$ найти тот, который доставляет минимум функции

$$f_{nq}(x) = \|t^n + p_{n-1}(t)\|_{L_q([-1,\,1])} = \left\| t^n + \sum_{k=0}^{n-1} x_k t^k \right\|_{L_q([-1,\,1])}. \qquad (P)$$

Функция $f_{nq}(\cdot)$ является выпуклой и непрерывной по x, следовательно, поставленная задача относится к числу простейших задач выпуклого программирования (задача без ограничений). Функция $f_{nq}(\cdot)$, очевидно, неограниченно возрастает при стремлении x к бесконечности. Отсюда следует, что минимум в задаче (P) существует. При $1 < q < \infty$ функция $f_{nq}(\cdot)$ является строго выпуклой по x и, значит, решение задачи (P) единственно. Можно показать, что и при $q = 1, +\infty$ решение будет также единственным. Обозначим минималь в задаче (P) через $\hat{x}$, значение задачи через γ_{nq}, а соответствующий полином через $T_{nq}(\cdot)$. Таким образом,

$$T_{nq}(t) = t^n + \sum_{k=1}^{n} \hat{x}_k t^{k-1}, \quad \gamma_{nq} = \|T_{nq}(\cdot)\|_{L_q([-1,\,1])}.$$

Полином $T_{nq}(\cdot)$ называется *полиномом наименьшего уклонения* в метрике $L_q([-1,\,1])$.

В дальнейшем нам понадобится следующая лемма.

Лемма (о нулях). *Полином* $T_{nq}(\cdot)$ *имеет в интервале* $(-1,\,1)$ *ровно* n *нулей.*

Доказательство. Не будем пока исследовать случаи $q = 1, +\infty$, так как далее будут даны явные выражения для полиномов $T_{n1}(\cdot)$ и $T_{n\infty}(\cdot)$ и окажется, что у них n нулей. Итак, пусть $1 < q < \infty$. Тогда функция $f_{nq}(\cdot)$ дифференцируема и вектор $\hat{x}$ является единственным решением уравнения

$$(f_{nq}(x))' = 0. \tag{2}$$

Производя дифференцирование, получаем

$$\frac{\partial f_{nq}(\hat{x})}{\partial x_i} = q\int\limits_{-1}^{1} |T_{nq}(t)|^{q-1}\operatorname{sign} T_{nq}(t)\cdot t^k\,dt = 0,\ k = 0,\ldots,n-1. \quad (*)$$

Следовательно, функция $|T_{nq}(\cdot)|^{q-1}\operatorname{sign} T_{nq}(\cdot)$, имеющая те же перемены знака, что и $T_{nq}(\cdot)$, ортогональна на $[-1,\,1]$ всем полиномам из P_n. Отсюда следует, что число ее перемен знака в интервале $(-1,\,1)$ не меньше n. Действительно, если полином $T_{nq}(t)$ меняет знак в s точках $-1 < t_1 < \ldots < t_s < 1$ и $s \le n-1$, то для полинома $p(t) = (t-t_1)\cdot\ldots\cdot(t-t_s)$ $\int\limits_{-1}^{1} |T_{nq}(t)|^{q-1}\operatorname{sign} T_{nq}(t)\cdot p(t)\,dt \ne 0$, что и противоречит соотношению $(*)$. $\square$

В трех случаях, когда $q = 1$, 2 и ∞ удается в конечном виде найти и величины γ_{nq} и соответствующие им полиномы наилучшего приближения. Далее мы приводим такие решения этих задач, которые послужат некоторой опорой для дальнейшего.

7.2 Полиномы Чебышева первого рода

Теорема 8. *Полином, наименее уклоняющийся от нуля в пространстве $C([-1,\,1])$, имеет вид*

$$T_{n\infty}(t) = 2^{-(n-1)}\cos(n\arccos t).$$

Полином $T_{n\infty}(t)$ называется *полиномом Чебышева* (полиномом Чебышева первого рода).

Доказательство. Рассмотрим задачу

$$f(x) = \max_{t\in[-1,\,1]} f(t,x) = \max_{t\in[-1,\,1]} \left| t^n + \sum_{k=0}^{n-1} x_k t^k \right| \to \inf .$$

Функция f — выпуклая, непрерывная и растет на бесконечности. Значит, по следствию из теоремы Вейерштрасса решение $\hat{x}$ задачи существует.

Необходимое и достаточное условие: $0 \in \partial f(\hat{x})$.

Если $0 \in \partial f(\hat{x})$, то по теореме об очистке найдутся $s \le n+1$ точек $t_0 \le \tau_1 < \tau_2 < \ldots < \tau_s \le t_1$, неотрицательных чисел $\lambda_1, \lambda_2, \ldots, \lambda_s$, $\sum\limits_{i=1}^{s} \lambda_i = 1$ и векторов $y_i \in \partial_x f(\tau_i, \hat{x})$, таких, что

$$f(\tau_i, \hat{x}) = \|T_{n\infty}\|_\infty, \quad \sum_{i=1}^{s} \lambda_i y_i = 0. \tag{1}$$

Вследствие того, что функция $f(\tau_i, x)$ достигает свой положительный максимум в точке $\hat{x}$, то $f(\tau_i, \hat{x}) = M \neq 0$, и, следовательно, функция $f(\tau_i, x)$ (модуль в точке отличной от нуля) дифференцируема по x. Значит, ее субдифференциал $\partial_x f(\tau_i, \hat{x})$ совпадает с производной. Дифференцируя функцию $f(\tau_i, x) = \left| \tau_i^n - \sum\limits_{k=0}^{n-1} x_k \tau_i^k \right|$ по вектору x в точке $\hat{x}$, получаем вектор

$$y_i = -\operatorname{sign} T_{n\infty}(\tau_i) \cdot (1, \tau_i, \ldots, \tau_i^{n-1}). \tag{2}$$

Из соотношений (1) и (2) вытекает, что система

$$\sum_{i=1}^{s} \alpha_i y_i = 0 \iff \sum_{i=1}^{s} \alpha_i \operatorname{sign} T_{n\infty}(\tau_i)\tau_i^k = 0, \quad k = 0, \dots, n-1, \tag{3}$$

имеет нетривиальное решение $\left(\sum\limits_{i=1}^{s} \alpha_i = 1\right)$. При $s \le n$ эта система имеет лишь тривиальное решение, так как определитель системы является определителем Вандермонда. Значит, $s = n+1$.

Итак, доказано, что $T_{n,\infty}$ в $n+1$ точке $-1 \le \tau_1 < \ldots < \tau_{n+1} \le 1$ достигает своего максимального и минимального значения. Отсюда ясно, что $\hat{\tau}_1 = -1$, $\hat{\tau}_{n+1} = 1$, ибо иначе производная $\dot{T}_{n\infty}$, будучи полиномом степени $n-1$, имела бы $\ge n$ нулей; $T_{n\infty}(\pm 1) = M$. Но тогда полиномы $M^2 - T_{n\infty}^2$ и $(1-t^2)\dot{T}_{n\infty}^2$, где $M = \|T_{n\infty}\|_\infty$, имеют однократные нули $\hat{\tau}_1$ и $\hat{\tau}_{n+1}$ и двукратные нули в точках $\hat{\tau}_2, \dots, \hat{\tau}_n$, причем оба полинома имеют степень $2n$. Приравнивая коэффициенты при t^{2n} приходим к равенству:

$$n^2(M^2 - T_{n\infty}^2) = (1-t^2)\dot{T}_{n\infty}^2 \qquad (*).$$

Решим это дифференциальное уравнение, извлекая квадратный корень:

$$\pm n\sqrt{M^2 - T_{n\infty}^2} = \sqrt{1-t^2} \cdot \dot{T}_{n\infty} = \sqrt{1-t^2} \cdot \frac{dT_{n\infty}}{dt}$$

$$\iff \pm\frac{dT_{n\infty}}{\sqrt{M^2 - T_{n\infty}^2}} = \frac{n\,dt}{\sqrt{1-t^2}}.$$

Возьмем интеграл от обеих частей последнего уравнения:

$$\int \pm\frac{dT_{n\infty}}{\sqrt{M^2 - T_{n\infty}^2}} = \int \frac{n\,dt}{\sqrt{1-t^2}}$$

$$\iff \arccos\frac{T_{n\infty}}{M} = \pm n \arccos t + C.$$

Если $t = 1$, то $\frac{T_{n\infty}}{M} = 1$ и, значит, $C = 0$. Поэтому

$$T_{n\infty} = M\cos(\pm n \arccos t) = M\cos(n \arccos t).$$

Осталось доказать, что $T_{n\infty} = 2^{-(n-1)}\cos(n\arccos t)$. $\square$

7.3 Полиномы Чебышева второго рода

Теорема 9. *Полином, наименее уклоняющийся от нуля в пространстве* $L_1([-1,\,1])$*, имеет вид*

$$T_{n1}(t) = \frac{\dot{T}_{n+1\,\infty}(t)}{n+1} = \frac{\sin\bigl(\bigl((n+1)\arccos t\bigr)}{2^n\sqrt{1-t^2}}.$$

Полином $T_{n+1\,\infty}(t)$ называется *полиномом Чебышева* (полиномом Чебышева второго рода).

Доказательство. □

7.4 Полиномы Лежандра

Теорема 10. *Полином, наименее уклоняющийся от нуля в пространстве* $L_2([-1,\,1])$*, имеет вид*

$$T_{n2}(t) = \frac{n!}{(2n)!}\Big(\frac{d}{dt}\Big)^n (t^2-1)^n.$$

При этом $\|T_{n2}\|_{L_2([-1,\,1])} = \dfrac{2^n(n!)^2}{(2n)!}\sqrt{\dfrac{2}{2n+1}}$.

Полиномы $P_n(t) = \dfrac{1}{2^n n!}\Big(\dfrac{d}{dt}\Big)^n (t^2-1)^n$ называются *полиномами Лежандра*, $\|P_n\|_{L_2([-1,\,1])} = \sqrt{\dfrac{2}{2n+1}}$.

Полиномы $\sqrt{\dfrac{2n+1}{2}}P_n$ образуют ортонормированный базис в пространстве $L_2([-1,\,1])$, получающийся процессом ортогонализации полиномов $1, t, \ldots, t^n, \ldots$.

Доказательство. Запишем нашу экстремальную задачу в виде:

$$\int\limits_{-1}^{1} x^2(t)dt \to \inf; \quad x^{(n)}(t) = n!$$

Сведем задачу к задаче Лагранжа, вводя вместо функции x вектор-функцию $x = (x_1, \ldots, x_n)$, и обозначения: $x_1 = x$, $x_2 = \dot{x}$, $\ldots$, $x_n = x^{(n-1)}$. Тогда переменные $x_1, \ldots, x_n$ связаны дифференциальными уравнениями и получаем задачу:

$$\int\limits_{-1}^{1} x_1^2(t)dt \to \inf; \quad \dot{x}_1 = x_2,\ \dot{x}_2 = x_3,\ \ldots,\ \dot{x}_{n-1} = x_n,\ \dot{x}_n = n!.$$

Функция Лагранжа:

$$\mathscr{L} = \int\limits_{-1}^{1} \big(\underbrace{\lambda_0 x_1^2 + p_1(\dot{x}_1 - x_2) + \ldots + p_{n-1}(\dot{x}_{n-1} - x_n) + p_n(\dot{x}_n - n!)}_{L}\big)dt.$$

Необходимые условия:

a) система уравнений Эйлера для лагранжиана L

$$\begin{cases} -\dfrac{d}{dt}L_{\dot{x}_i} + L_{x_i} = 0, \\ i = 1, \dots, n, \end{cases} \iff \begin{cases} -\dot{p}_1 + 2\lambda_0 x_1 = 0, \\ -\dot{p}_2 - p_1 = 0, \\ \dots\dots\dots\dots \\ -\dot{p}_n - p_{n-1} = 0; \end{cases}$$

b) трансверсальность по x для терминанта $l = 0$:

$$\begin{cases} L_{\dot{x}_k}(-1) = l_{x_k(-1)}, \\ L_{\dot{x}_k}(1) = -l_{x_k(1)}, \end{cases} \iff p_k(\pm 1) = 0,\ k = 1, \dots, n.$$

Если $\lambda_0 = 0$, то из (a) следует, что $\dot{p}_1 = 0$. Тогда $p_1 = \mathrm{const}$. В силу условия трансверсальности $p_1 = 0$. Аналогично, $p_2 = \ldots = p_n = 0$. Все множители Лагранжа — нули. Этого не может быть. Значит, $\lambda_0 \neq 0$.

Обозначим $p(t) := p_n(t)$. Из системы (a) следует, что

$$p^{(n)}(t) = 2(-1)^{n-1}\lambda_0 x(t),$$

а условия трансверсальности перепишутся в виде:

$$p^{(k-1)}(\pm 1) = 0,\ k = 1, \dots, n.$$

Положим $\lambda_0 = \dfrac{(-1)^{n-1}}{2}$. Получим, что $\begin{cases} p^{(n)}(t) = x(t), \\ x^{(n)}(t) = n!. \end{cases}$

Таким образом, для функции p выполняются условия: $p^{(2n)}(t) = n!$, $p^{(k-1)}(\pm 1) = 0$, $k = 1, \dots, n$. Следовательно, p — многочлен степени $2n$, имеющий n-кратные нули в точках ± 1, т. е. $p(t) = C(t^2 - 1)^n$. Константа C находится из того условия, что $p^{(2n)}(t) = n!$. Поэтому $p(t) = \dfrac{n!}{(2n)!}(t^2 - 1)^n$.

Отсюда получаем явное выражение для полинома Лежандра

$$T_{n2}(t) = \frac{n!}{(2n)!}\left(\frac{d}{dt}\right)^n (t^2 - 1)^n.$$

Найдем выражение для $\|T_{n2}\|_{L_2([-1,\,1])}$. Поскольку в силу соотношения $(*)$ многочлен T_{n2} ортогонален t^k, $k=0,\dots,n-1$, то

$$\|T_{n2}\|^2_{L_2([-1,\,1])} = \int\limits_{-1}^{1} T_{n2}(t)\Bigl(t^n + \sum_{k=0}^{n-1}\hat{x}_k t^k\Bigr)dt = \int\limits_{-1}^{1} T_{n2}(t)t^n\,dt =$$

$$= \frac{n!}{(2n)!}\int\limits_{-1}^{1}\left(\frac{d}{dt}\right)^n (t^2-1)^n\cdot t^n\,dt = \frac{n!}{(2n)!}\int\limits_{-1}^{1} t^n\,d\Bigl(\frac{d}{dt}\Bigr)^{n-1}(t^2-1)^n =$$

$$= \frac{n!}{(2n)!}\left(t^n\,d\Bigl(\frac{d}{dt}\Bigr)^{n-1}(t^2-1)^n\Big|_{-1}^{1} - \int\limits_{-1}^{1}\left(\frac{d}{dt}\right)^{n-1}(t^2-1)^n\cdot nt^{n-1}\,dt\right) =$$

$$= -\frac{n!n}{(2n)!}\int\limits_{-1}^{1}\left(\frac{d}{dt}\right)^{n-1}(t^2-1)^n\cdot t^{n-1}\,dt = \ldots = \frac{(n!)^2}{(2n)!}(-1)^n\int\limits_{-1}^{1}(t^2-1)^n\,dt$$

(берем n раз интеграл по частям).
Известно, что $\int\limits_{-1}^{1}(t^2-1)^n\,dt = (-1)^n\dfrac{2^{2n+1}(n!)^2}{(2n+1)(2n)!}$. Поэтому

$$\|T_{n2}\|^2_{L_2([-1,\,1])} = \frac{2^{2n+1}(n!)^4}{(2n+1)((2n)!)^2}.$$

Извлекая квадратный корень, получаем доказываемую норму полинома T_{n2}. $\square$

Приведем еще одно доказательство теоремы.

Доказательство. 2. Рассмотрим задачу

$$f(x) = \int_{-1}^{1} \left(t^n + \sum_{k=0}^{n-1} x_k t^k \right)^2 dt \to \min .$$

Функция f является гладкой выпуклой функцией (проверьте!). Необходимое и достаточное условие минимума в такой задаче — $0 \in \partial f(\hat{x})$ — равносильно условию $f'(\hat{x}) = 0$, т. е.

$$\int_{-1}^{1} \left(t^n + \sum_{k=0}^{n-1} \hat{x}_k t^k \right) t^k \, dt = 0 \iff \int_{-1}^{1} T_{n2}(t) t^k \, dt = 0, \; k = 0, \ldots, n-1. \tag{$*$}$$

Положим $p(t) = \dfrac{1}{(n-1)!} \int_{-1}^{t} (t-\tau)^{n-1} T_{n2}(\tau) \, d\tau$. Тогда $\dot{p}(t) = \dfrac{1}{(n-2)!} \int_{-1}^{t} (t-\tau)^{n-2} T_{n2}(\tau) \, d\tau, \; \ldots, \; p^{(n-1)}(t) = \int_{-1}^{t} T_{n2}(\tau) \, d\tau$, $p^{(n)}(t) = T_{n2}(t)$. Поэтому $p^{(k)}(-1) = 0, \; k = 0, \ldots, n-1$.

В силу соотношений $(*)$

$$p^{(k)}(1) = \frac{1}{(n-1-k)!} \int_{-1}^{1} (t-\tau)^{n-1-k} T_{n2}(\tau) \, d\tau = 0, \; k = 0, \ldots, n-1.$$

Значит, p — полином степени $2n$ такой, что $p^{(k)}(\pm 1) = 0$, $k = 0, \ldots, n-1$, т. е. $p(t) = C(t^2-1)^n$. Константа C находится из того условия, что $p^{(n)}(t) = T_{n2}(t) = t^n + \ldots \Leftrightarrow p^{(2n)}(t) = n!$. Поэтому $p(t) = \dfrac{n!}{(2n)!}(t^2-1)^n$.

Отсюда получаем явное выражение для полинома Лежандра

$$T_{n2}(t) = p^{(n)}(t) = \frac{n!}{(2n)!} \left(\frac{d}{dt} \right)^n (t^2-1)^n. \qquad \square$$