Введение в статистическую теорию распознавания образов. Оценивание функции плотности вероятностей (ФПВ).

Тверская Е. С.
e_tverskaya@bmstu.ru

Московский государственный технический университет им. Н. Э. Баумана
(национальный исследовательский университет)

Москва, 2025

ИУ-6
Компьютерные системы и сети

Оценивание ФПВ.
Введение. О принципе максимума энтропии.

Рассмотрим **понятие**, введенное Клодом Шенноном в его работе "Математическая теория связи"в 1948г.

Определение

Рассмотрим величину $H = -\sum p_i \log_2 p_i$, где p_i – вероятность того, что система находится в состоянии i фазового простанства. Данная величина определеяет ***меру количества информации, возможности выбора и неопределенности системы.*** Данную величину называют ***энтропией*** множества вероятностей $p_1, \dots, p_n$.

Работы по теории информации и кибернетике, *Шеннон К.: Пер. с англ. –М.: Изд-во иностранной литературы, 1963г., 829с.*

Оценивание ФПВ.
Введение. О принципе максимума энтропии.

Рассмотрим **понятие**, введенное Клодом Шенноном в его работе "Математическая теория связи"в 1948г.

Определение

Рассмотрим величину $H = -\sum p_i \log_2 p_i$, где p_i – вероятность того, что система находится в состоянии i фазового простанства. Данная величина определеяет ***меру количества информации, возможности выбора и неопределенности системы.*** Данную величину называют ***энтропией*** множества вероятностей $p_1, \ldots, p_n$.

Работы по теории информации и кибернетике, *Шеннон К.: Пер. с англ. –М.: Изд-во иностранной литературы, 1963г., 829с.*

Принцип максимума энтропии.

Если плотность распределения некоторой случайной величины неизвестна, то следует выбрать такую плотность распределения, которая обеспечивает ***максимум энтропии*** с учетом всех известных ограничений.

Пусть известные ограничения представлены в форме средних оценок: математическое ожидание, дисперсия, плотность распределения и так далее.

Оценивание ФПВ.

Введение. О принципе максимума энтропии.

Обобщим понятие энтропии для непрерывной случайной величины. Также, не нарушая общности задачи, $\log_2$ заменим $\ln$.

Постановка задачи.

$$H = -\int_X f(x)\ln(f(x))\,dx \longrightarrow \max,$$

$$\int_X f(x)\,dx = 1,$$

$$\int_X b_k(x) f(x)\,dx = a_k, \quad k = \overline{1,n}.$$

Оценивание ФПВ.
Введение. О принципе максимума энтропии.

Обобщим понятие энтропии для непрерывной случайной величины. Также, не нарушая общности задачи, $\log_2$ заменим $\ln$.

Постановка задачи.

$$H = -\int\limits_X f(x)\ln(f(x))\,dx \longrightarrow \max,$$

$$\int\limits_X f(x)\,dx = 1,$$

$$\int\limits_X b_k(x)f(x)\,dx = a_k, \quad k = \overline{1, n}.$$

Получили задачу на условный экстремум. Запишем ***функцию Лагранжа***:

$$L = -\int\limits_X f(x)\ln(f(x))\,dx + \sum_{k=1}^{n}\lambda_k\left(\int\limits_X b_k(x)f(x)\,dx - a_k\right) +$$

$$+\lambda_0\left(\int\limits_X f(x)\,dx - 1\right) = -\int\limits_X f(x)\left(\ln(f(x)) - \sum_{k=0}^{n}\lambda_k b_k(x)\right)dx - \sum_{k=0}^{n}\lambda_k a_k,$$

где $b_0(x) = 1$ и $a_0 = 1$.

Оценивание ФПВ.
Введение. О принципе максимума энтропии.

Тогда

$$\frac{\partial L}{\partial f(x)} = -\int\limits_X \left(\ln f(x) - \sum_{k=0}^{n} \lambda_k b_k(x) + 1 \right) dx.$$

Оценивание ФПВ.
Введение. О принципе максимума энтропии.

Тогда

$$\frac{\partial L}{\partial f(x)} = -\int_X \left(\ln f(x) - \sum_{k=0}^{n} \lambda_k b_k(x) + 1\right) dx.$$

И можно найти

$$f(x) = \exp\left(\sum_{k=0}^{n} \lambda_k b_k(x) - 1\right). \qquad (1)$$

Параметры $\lambda_0, \lambda_1, \ldots, \lambda_n$ определяются исходя из условий (априорной информации).

Оценивание ФПВ.
Введение. О принципе максимума энтропии.

Тогда

$$\frac{\partial L}{\partial f(x)} = -\int_X \left(\ln f(x) - \sum_{k=0}^{n} \lambda_k b_k(x) + 1\right) dx.$$

И можно найти

$$f(x) = \exp\left(\sum_{k=0}^{n} \lambda_k b_k(x) - 1\right). \qquad (1)$$

Параметры $\lambda_0, \lambda_1, \ldots, \lambda_n$ определяются исходя из условий (априорной информации).

Пример №1 (равномерное распределение). Пусть априорная информация для случайной величины $x \in [\alpha, \beta]$ заключается в том, что $\int_\alpha^\beta f(x)\,dx = 1$. Тогда, согласно (1)

$$f(x) = \exp(\lambda_0 - 1).$$

Величину λ_0 вычисляем из условия $\int_\alpha^\beta \exp(\lambda_0 - 1)\,dx = 1$, тогда $\exp(\lambda_0 - 1) = {}^1/(\beta - \alpha)$ и искомая ФПВ

$$f(x) = \begin{cases} {}^1/(\beta - \alpha), & \text{если } \alpha \leqslant x \leqslant \beta, \\ 0, & \text{если } x < \alpha \text{ или } x > \beta. \end{cases}$$

Оценивание ФПВ.
Введение. О принципе максимума энтропии.

Пример №2. Пусть априорная информация для случайной величины $x \in [0, +\infty)$ заключается в том, что $\int\limits_0^\infty f(x)\,dx = 1$ и $\int\limits_0^\infty x f(x)\,dx = m$. Тогда, согласно (1)

$$f(x) = \exp(\lambda_0 - 1 + \lambda_1 x).$$

Получаем систему из 2-х уравнений относительно неизвестных λ_0 и λ_1:

$$\begin{cases} \int\limits_0^\infty \exp(\lambda_0 - 1 + \lambda_1 x)\,dx = 1, \\ \int\limits_0^\infty x\exp(\lambda_0 - 1 + \lambda_1 x)\,dx = m. \end{cases}$$

Тогда $\exp(\lambda_0 - 1) = {}^1/_m$ и $\lambda_1 = -{}^1/_m$ и искомая ФПВ

$$f(x) = \begin{cases} {}^1/_m \exp\left(-{}^x/_m\right), & \text{если } x \geqslant 0, \\ 0, & \text{если } x < 0. \end{cases}$$

Оценивание ФПВ.
Введение. О принципе максимума энтропии.

Пример №2. Пусть априорная информация для случайной величины $x \in [0, +\infty)$ заключается в том, что $\int\limits_0^\infty f(x)\,dx = 1$ и $\int\limits_0^\infty x f(x)\,dx = m$. Тогда, согласно (1)

$$f(x) = \exp(\lambda_0 - 1 + \lambda_1 x).$$

Получаем систему из 2-х уравнений относительно неизвестных λ_0 и λ_1:

$$\begin{cases} \int\limits_0^\infty \exp(\lambda_0 - 1 + \lambda_1 x)\,dx = 1, \\ \int\limits_0^\infty x\exp(\lambda_0 - 1 + \lambda_1 x)\,dx = m. \end{cases}$$

Тогда $\exp(\lambda_0 - 1) = {}^1\!/_m$ и $\lambda_1 = -{}^1\!/_m$ и искомая ФПВ

$$f(x) = \begin{cases} {}^1\!/_m \exp\left(-{}^x\!/_m\right), & \text{если } x \geqslant 0, \\ 0, & \text{если } x < 0. \end{cases}$$

Интегралы, полученные в **примере №1** и **примере №2**, являются табличными интегралами или сводятся к ним с помощью формулы интегрирования по частям, учитывая, что для сходимости несобственного интеграла в **Примере №2** должно выполняться условие $\lambda_1 < 0$.

Оценивание ФПВ.
Введение. О принципе максимума энтропии.

Пример №3. Пусть априорная информация для случайной величины $x \in \mathbb{R}$ заключается в том, что $\int\limits_0^\infty f(x)\,dx = 1$, $\int\limits_0^\infty x f(x)\,dx = m$ и $\int\limits_0^\infty x^2 f(x)\,dx = \sigma^2$ Тогда, согласно (1)

$$f(x) = \exp(\lambda_0 - 1 + \lambda_1 x + \lambda_2 x^2).$$

Получаем систему из 3-х уравнений относительно неизвестных λ_0, λ_1 и λ_2:

$$\begin{cases} \int\limits_{-\infty}^{\infty} \exp(\lambda_0 - 1 + \lambda_1 x + \lambda_2 x^2)\,dx = 1, \\ \int\limits_{-\infty}^{\infty} x \exp(\lambda_0 - 1 + \lambda_1 x + \lambda_2 x^2)\,dx = m, \\ \int\limits_{-\infty}^{\infty} x^2 \exp(\lambda_0 - 1 + \lambda_1 x + \lambda_2 x^2)\,dx = \sigma^2. \end{cases}$$

Тогда

$$\exp(\lambda_0 - 1) = \left({}^1\!/\!\sqrt{2\pi}\right)\ {}^1\!/\!\sigma \exp\left(-{}^{m^2}\!/\!{}_{2\sigma^2}\right),$$
$$\lambda_1 = {}^m\!/\!\sigma^2, \quad \lambda_2 = -{}^1\!/\!{}_{2\sigma^2}.$$

Тогда искомая ФПВ

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right).$$

Заметим, что решение полученной системы является достаточно трудоемким процессом.

Оценивание ФПВ.
Разложение по базисным функциям.

Пусть $\hat{f}(x)$ – оценка ФПВ для функции $f(x)$. Будем искать такую оценку $\hat{f}(x)$, которая обеспечивает ***минимум среднеквадратичной ошибки***:

$$\varepsilon^2 = \int\limits_{X} \rho(x) \left(f(x) - \hat{f}(x) \right)^2 dx \longrightarrow \min,$$

где $\rho(x)$ – весовая функция.
Метод аппроксимации состоит в разложении функции $f(x)$ по ***базисным функциям*** $\varphi_i(x)$:

$$f(x) = \sum_{i=1}^{\infty} c_i \varphi_i(x).$$

В этом случае оценку ФПВ можно представить в виде:

$$\hat{f}(x) = \sum_{i=1}^{m} c_i \varphi_i(x).$$

Оценивание ФПВ.
Разложение по базисным функциям.

Задача оптимизации.

Решается задача безусловной оптимизации

$$\varepsilon^2 = \int\limits_{X} \rho(x) \left(f(x) - \sum_{i=1}^{m} c_i \varphi(x) \right)^2 dx \longrightarrow \min$$

Оценивание ФПВ.

Разложение по базисным функциям.

Задача оптимизации.

Решается задача безусловной оптимизации

$$\varepsilon^2 = \int_{\mathcal{X}} \rho(x) \left(f(x) - \sum_{i=1}^{m} c_i \varphi(x) \right)^2 dx \longrightarrow \min$$

Необходимое условие экстремума: $\dfrac{\partial \varepsilon^2}{\partial c_k} = 0$, $k = \overline{1, m}$.

$$\sum_{i=1}^{m} c_i \int_{\mathcal{X}} \rho(x)\varphi_i(x)\varphi_k(x)\, dx = \int_{\mathcal{X}} \rho(x)\varphi_k(x) f(x)\, dx$$

$$= M\left[\rho(x)\varphi_k(x)\right],\ k = \overline{1, m}.$$

Оценивание ФПВ.

Разложение по базисным функциям.

Задача оптимизации.

Решается задача безусловной оптимизации

$$\varepsilon^2 = \int\limits_{\mathcal{X}} \rho(x) \left(f(x) - \sum_{i=1}^{m} c_i \varphi(x) \right)^2 dx \longrightarrow \min$$

Необходимое условие экстремума: $\dfrac{\partial \varepsilon^2}{\partial c_k} = 0$, $k = \overline{1, m}$.

$$\sum_{i=1}^{m} c_i \int\limits_{\mathcal{X}} \rho(x) \varphi_i(x) \varphi_k(x)\, dx = \int\limits_{\mathcal{X}} \rho(x) \varphi_k(x) f(x)\, dx$$

$$= M\left[\rho(x)\varphi_k(x)\right], \; k = \overline{1, m}.$$

Математическое ожидание в равенстве можно аппроксимировать выборочным средним, тогда

$$\sum_{i=1}^{m} c_i \int\limits_{\mathcal{X}} \rho(x) \varphi_i(x) \varphi_k(x)\, dx = \frac{1}{N} \sum_{j=1}^{N} \rho(x_j) \varphi_k(x_j), \quad k = \overline{1, m}.$$

Оценивание ФПВ.
Разложение по базисным функциям.

Пусть базисные функции удовлетворяют условию:

$$\int\limits_{\boldsymbol{x}} \boldsymbol{\rho(x)\varphi_i(x)\varphi_k(x)\,dx = \lambda_k\delta_{ik}} = \begin{cases} \boldsymbol{\lambda_k}, & \text{если } \boldsymbol{i = k}, \\ \boldsymbol{0}, & \text{если } \boldsymbol{i \neq k}. \end{cases}$$

Оценивание ФПВ.

Разложение по базисным функциям.

Пусть базисные функции удовлетворяют условию:

$$\int\limits_{\mathcal{X}} \rho(x)\varphi_i(x)\varphi_k(x)\,dx = \lambda_k\delta_{ik} = \begin{cases} \lambda_k, & \text{если } i = k, \\ 0, & \text{если } i \neq k. \end{cases}$$

Тогда

$$c_k\lambda_k = \frac{1}{N}\sum_{j=1}^{N} \rho(x_j)\varphi_k(x_j), \quad k = \overline{1, m}.$$

Оценивание ФПВ.
Разложение по базисным функциям.

Пусть базисные функции удовлетворяют условию:

$$\int\limits_{\mathcal{X}} \rho(x)\varphi_i(x)\varphi_k(x)\,dx = \lambda_k\delta_{ik} = \begin{cases}\lambda_k, & \text{если } i = k,\\ 0, & \text{если } i \neq k.\end{cases}$$

Тогда

$$c_k\lambda_k = \frac{1}{N}\sum_{j=1}^{N}\rho(x_j)\varphi_k(x_j), \quad k = \overline{1, m}.$$

Предположим, что система используемых базисных функций являются ортонормированной, тогда

$$c_k = \frac{1}{N}\sum_{j=1}^{N}\rho(x_j)\varphi_k(x_j), \quad k = \overline{1, m}. \tag{2}$$

Так как в оценке $\hat{f}(x)$ ограничились только первыми m членами, то среднеквадратичная ошибка

$$\varepsilon^2 = \int\limits_{\mathcal{X}} \rho(x)\left\{f(x) - \sum_{i=1}^{m} c_i\varphi_i(x)\right\}^2 dx =$$

$$= \int\limits_{\mathcal{X}} \rho(x)\left\{\sum_{i=m+1}^{\infty} c_i\varphi_i(x)\right\}^2 dx = \sum_{i=m+1}^{\infty} \lambda_i\,|c_i|^2$$

Оценивание ФПВ.
Разложение по базисным функциям.

$\lambda_i \left|c_i\right|^2$ – ошибка, обусловленная исключением i-го члена разложения.

Вывод!

Если удастся найти систему базисных функций таких, что $\boldsymbol{\lambda_i \left|c_i\right|^2}$ быстро убывает с ростом $\boldsymbol{i}$, то такая система даст удовлетворительный результат представления ФПВ.

Оценивание ФПВ.
Разложение по базисным функциям.

$\lambda_i \left|c_i\right|^2$ – ошибка, обусловленная исключением i-го члена разложения.

Вывод!

Если удастся найти систему базисных функций таких, что $\lambda_i \left|c_i\right|^2$ быстро убывает с ростом i, то такая система даст удовлетворительный результат представления ФПВ.

Примеры базисных функций.

1. Полиномы Лежандра, ортогональны на $[-1, 1]$
2. Полиномы Эрмита, ортогональны на $(-\infty, \infty)$.
3. Полиномы Лагера, ортогональны на $[0, \infty)$.
4. Полиномы Гегенбауэра, отогональны на $[-1, 1]$.

Оценивание ФПВ.
Разложение по базисным функциям.

$\lambda_i |c_i|^2$ – ошибка, обусловленная исключением i-го члена разложения.

Вывод!

Если удастся найти систему базисных функций таких, что $\lambda_i |c_i|^2$ быстро убывает с ростом i, то такая система даст удовлетворительный результат представления ФПВ.

Примеры базисных функций.

1. Полиномы Лежандра, ортогональны на $[-1, 1]$
2. Полиномы Эрмита, ортогональны на $(-\infty, \infty)$.
3. Полиномы Лагера, ортогональны на $[0, \infty)$.
4. Полиномы Гегенбауэра, отогональны на $[-1, 1]$.

Примером разложения по базисным функциям являются ряды Фурье.

Оценивание ФПВ.

Классические ортогональные полиномы.

Более подробную информацию можно прочитать в книге: **Свешников А. Г., Боголюбов А. Н., Кравцов В. В.** Лекции по математической физике: Учеб. пособие. –М.: Изд-во МГУ, 1993. –352 с.

Определение

Будем называть систему $\{p_n(x)\}$ полиномов всех степеней, заданных на отрезке $[a, b]$ ***системой классических ортогональных полиномов*** если они ортогональны на отрезке $[a, b]$ с весом $\rho(x)$, удовлетворяющим на интервале (a, b) дифференциальному уравнению Пуассона

$$\frac{d}{dx}\left(\sigma(x)\rho(x)\right) = \tau(x)\rho(x), \tag{3}$$

где $\sigma(x)$ и $\tau(x)$ – заданные функции, удовлетворяющие условию

$$x^m\sigma(x)\rho(x)|_a^b = 0, \quad m = 0, 1, \ldots \tag{4}$$

Оценивание ФПВ.
Классические ортогональные полиномы.

Более подробную информацию можно прочитать в книге: **Свешников А. Г., Боголюбов А. Н., Кравцов В. В.** Лекции по математической физике: Учеб. пособие. –М.: Изд-во МГУ, 1993. –352 с.

Определение

Будем называть систему $\{p_n(x)\}$ полиномов всех степеней, заданных на отрезке $[a, b]$ ***системой классических ортогональных полиномов*** если они ортогональны на отрезке $[a, b]$ с весом $\rho(x)$, удовлетворяющим на интервале (a, b) дифференциальному уравнению Пуассона

$$\frac{d}{dx}\left(\sigma(x)\rho(x)\right) = \tau(x)\rho(x), \tag{3}$$

где $\sigma(x)$ и $\tau(x)$ – заданные функции, удовлетворяющие условию

$$x^m \sigma(x)\rho(x)|_a^b = 0, \quad m = 0, 1, \ldots \tag{4}$$

Замечание

Граничные точки a и b отрезка $[a, b]$ могут соответсвенно принимать значения $-\infty$ и $+\infty$.

Оценивание ФПВ.
Классические ортогональные полиномы.

$\tau(x) = Ax + B$ – линейная функция.
Функция $\sigma(x)$ имеет вид

$$\sigma(x) = \begin{cases} (x-a)(b-x), & \text{при } a \neq -\infty,\ b \neq \infty, \\ x-a, & \text{при } a \neq -\infty,\ b = \infty, \\ b-x, & \text{при } a = -\infty,\ b \neq \infty, \\ 1, & \text{при } a = -\infty,\ b = \infty. \end{cases} \tag{5}$$

Общее решение уравнения Пуассона имеет вид

$$\rho(x) = \frac{1}{\sigma(x)} \exp\left(\int \frac{\tau(x)}{\sigma(x)} dx\right). \tag{6}$$

Выражения (3)-(6) определяю целый класс классических ортогональных полиномов.
Для решения поставленной задачи будем исследовать классические ортогональные полиномы – полиномы Эрмита.

Оценивание ФПВ.

Классические ортогональные полиномы. Полиномы Эрмита.

Пусть $a = -\infty$ и $b = \infty$. Тогда при $\sigma(x) = 1$ и $\tau(x) = -2x$ получим

$$\rho(x) = \exp\left(\int \tau(x)\, dx\right) = e^{-x^2}$$

Оценивание ФПВ.
Классические ортогональные полиномы. Полиномы Эрмита.

Пусть $a = -\infty$ и $b = \infty$. Тогда при $\sigma(x) = 1$ и $\tau(x) = -2x$ получим

$$\rho(x) = \exp\left(\int \tau(x)\,dx\right) = e^{-x^2}$$

Определение

Классические ортогональные полиномы, заданные на прямой $(-\infty, \infty)$ и ортогональные на ней с весом $\rho(x) = e^{-x^2}$, называются ***полиномами Эрмита***.

Можно получить следующее выражение для полиномов Эрмита:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n}\left(e^{-x^2}\right)$$

Выпишем явно несколько первых полиномов

$$H_0(x) = 1,$$
$$H_1(x) = 2x,$$
$$H_2(x) = 4x^2 - 2,$$
$$H_3(x) = 8x^3 - 12x.$$

Запишем квадрат нормы полиномов Эрмита

$$||H_n||^2 = 2^n n! \int_{-\infty}^{\infty} e^{-x^2} dx = 2^n n! \sqrt{\pi}.$$

Оценивание ФПВ.

Классические ортогональные полиномы. Полиномы Эрмита.

ИУ-6 Компьютерные системы и сети

Следует отметить, что при аппроксимации функций целесообразно использовать следующую ортонормированную систему

$$\psi(x) = \sqrt{\frac{\rho(x)}{\lambda_i}}\varphi_i(x).$$

Оценивание ФПВ.

Классические ортогональные полиномы. Полиномы Эрмита.

ИУ-6 Компьютерные системы и сети

Следует отметить, что при аппроксимации функций целесообразно использовать следующую ортонормированную систему

$$\psi(x) = \sqrt{\frac{\rho(x)}{\lambda_i}}\varphi_i(x).$$

В этом случае коэффициенты

$$c_k = \frac{1}{N}\sum_{j=1}^{N}\psi_k(x_j), \quad k = \overline{1, m}.$$

Оценивание ФПВ.

Классические ортогональные полиномы. Полиномы Эрмита.

Следует отметить, что при аппроксимации функций целесообразно использовать следующую ортонормированную систему

$$\psi(x) = \sqrt{\frac{\rho(x)}{\lambda_i}}\varphi_i(x).$$

В этом случае коэффициенты

$$c_k = \frac{1}{N}\sum_{j=1}^{N}\psi_k(x_j), \quad k = \overline{1,m}.$$

Возвращаясь к полиномам Эрмита

$$\psi_n(x) = \frac{e^{-x^2/2}}{\sqrt{2^n n!\sqrt{\pi}}}H_n(x).$$

Оценивание ФПВ.
Классические ортогональные полиномы. Полиномы Эрмита.

ИУ-6 Компьютерные системы и сети

Следует отметить, что при аппроксимации функций целесообразно использовать следующую ортонормированную систему

$$\psi(x) = \sqrt{\frac{\rho(x)}{\lambda_i}}\varphi_i(x).$$

В этом случае коэффициенты

$$c_k = \frac{1}{N}\sum_{j=1}^{N}\psi_k(x_j), \quad k = \overline{1, m}.$$

Возвращаясь к полиномам Эрмита

$$\psi_n(x) = \frac{e^{-x^2/2}}{\sqrt{2^n n!\sqrt{\pi}}}H_n(x).$$

Замечание

Следует отметить, что полиномы Эрмита используются для аппроксимации плотностей вероятности не «очень сильно» отличающихся от нормальной. Также их использование может вызывать вычислительные трудности из-за вычисления коэффицентов при $H_n(x)$.

Оценивание ФПВ.

Классические ортогональные полиномы. Полиномы Эрмита.

Так как в выражении (2) члены $\rho(x_j)$ не зависят от k и, как следствие, во всех коэффициентах одинаковы, то их можно исключить из аппроксимирующего выражения. Тогда

$$c_k = \frac{1}{N}\sum_{j=1}^{N}\varphi_k(x_j), \quad k = \overline{1, m}$$

Оценивание ФПВ.
Классические ортогональные полиномы. Полиномы Эрмита.

ИУ-6 Компьютерные системы и сети

Так как в выражении (2) члены $\rho(x_j)$ не зависят от k и, как следствие, во всех коэффициентах одинаковы, то их можно исключить из аппроксимирующего выражения. Тогда

$$c_k = \frac{1}{N}\sum_{j=1}^{N}\varphi_k(x_j), \quad k = \overline{1,m}$$

Замечание

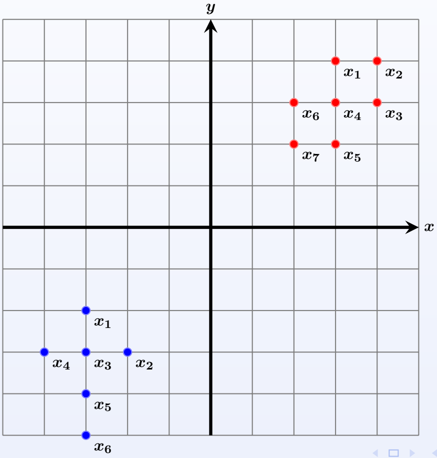
В иллюстативных целях, для демонстрации предложенного метода, будем обращаться с ортогональными функциями так, как если бы они были ортонормированными.

Один из возможных вариантов решения данной задачи в многомерном случае предложен в работе: ***Курант Р., Гильберт Д.*** *Методы математической физики. 1945. в 2-х томах.* Следует отметить, что данный подход в многомерном случае не является единственным и отыскание универсальной системы базисных функций и коэффициентов разложения является достаточно сложной задачей.

$$\begin{aligned}
\varphi_1(x) &= \varphi_1(x_1, x_2) = H_0(x_1)H_0(x_2) = 1,\\
\varphi_2(x) &= \varphi_2(x_1, x_2) = H_1(x_1)H_0(x_2) = 2x_1,\\
\varphi_3(x) &= \varphi_3(x_1, x_2) = H_0(x_1)H_1(x_2) = 2x_2,\\
\varphi_4(x) &= \varphi_4(x_1, x_2) = H_1(x_1)H_1(x_2) = 4x_1x_2.
\end{aligned}$$

Оценивание ФПВ.
Разложение по базисным функциям. Пример.

Рассмотрим набор точек на плоскости, которые представляются двумя классами: $\boldsymbol{\omega_1}$ (красный цвет) и $\boldsymbol{\omega_2}$ (синий цвет).



Оценивание ФПВ.
Разложение по базисным функциям. Пример.

Сделаем допущение, что $\mathbf{P}(\omega_1) = \mathbf{P}(\omega_2) = 0.5$.

$$c_{1k} = \frac{1}{N_1}\sum_{j=1}^{N_1}\varphi_k(x_{1j}), \quad k = \overline{1, m},$$

где $N_1 = 7$ – количество элементов в классе ω_1, $m = 4$ – количество базисных функций.

$$c_{11} = \frac{1}{7}\sum_{j=1}^{7}\varphi_1(x_{1j}) = \frac{1}{7}(1 + \cdots + 1) = 1,$$

$$c_{12} = \frac{1}{7}\sum_{j=1}^{7}\varphi_2(x_{1j}) = \frac{1}{7}(4 + 4 + 6 + 6 + 6 + 8 + 8) = 6,$$

$$c_{13} = \frac{1}{7}\sum_{j=1}^{7}\varphi_3(x_{1j}) = \frac{1}{7}(4 + 4 + 6 + 6 + 6 + 8 + 8) = 6,$$

$$c_{13} = \frac{1}{7}\sum_{j=1}^{7}\varphi_4(x_{1j}) = \frac{1}{7}(16 + 24 + 24 + 36 + 48 + 48 + 64) \simeq 37.1.$$

Тогда

$$\hat{f}\,(x|\omega_1) = 1 + 12x_1 + 12x_2 + 148.4x_1x_2.$$

Оценивание ФПВ.

Разложение по базисным функциям. Пример.

$$c_{2k} = \frac{1}{N_2}\sum_{j=1}^{N_2}\varphi_k(x_{2j}), \quad k = \overline{1, m}$$

где $N_2 = 6$ – количество элементов в классе ω_2, $m = 4$ – количество базисных функций.

$$c_{21} = \frac{1}{6}\sum_{j=1}^{6}\varphi_1(x_{2j}) = \frac{1}{6}(1 + \cdots + 1) = 1,$$

$$c_{22} = \frac{1}{6}\sum_{j=1}^{6}\varphi_2(x_{2j}) = \frac{1}{6}(-4 - 8 - 4 \cdot 8) = -6,$$

$$c_{23} = \frac{1}{6}\sum_{j=1}^{6}\varphi_3(x_{2j}) = \frac{1}{6}(-4 - 6 \cdot 3 - 8 - 10) \simeq -6.7,$$

$$c_{23} = \frac{1}{6}\sum_{j=1}^{6}\varphi_4(x_{2j}) = \frac{1}{6}(-24 - 24 - 36 - 48 - 48 - 60) = 40.$$

Тогда

$$\hat{f}\,(x|\omega_2) = 1 - 12x_1 - 13.4x_2 + 160x_1x_2.$$

Тогда уравнение разделяющей границы будет:

$$\hat{f}(x|\omega_1)\mathrm{P}(\omega_1) = \hat{f}(x|\omega_2)\mathrm{P}(\omega_2).$$

Оценивание ФПВ.
Разложение по базисным функциям. Пример.

Уравнение разделяющей границе изображено на рисунке, она разделяет наше множество на два класса.