Метод наименьших квадратов
и его друзья

Алексей Перегудин, 2020

Норма вектора

Норма вектора

Норма вектора – это неотрицательное число, характеризующее «размер», «массу», «значимость» этого вектора

Норма вектора

УНИВЕРСИТЕТ ИТМО

Норма вектора – это неотрицательное число, характеризующее «размер», «массу», «значимость» этого вектора

Обычная норма вектора – это его длина

Норма вектора

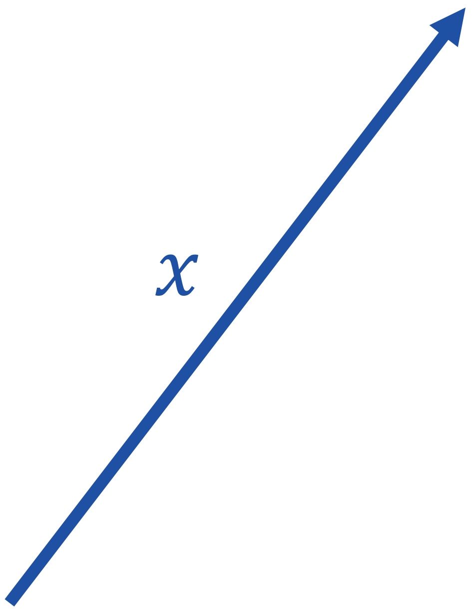
Норма вектора – это неотрицательное число, характеризующее «размер», «массу», «значимость» этого вектора

Обычная норма вектора – это его длина

Но есть и другие…

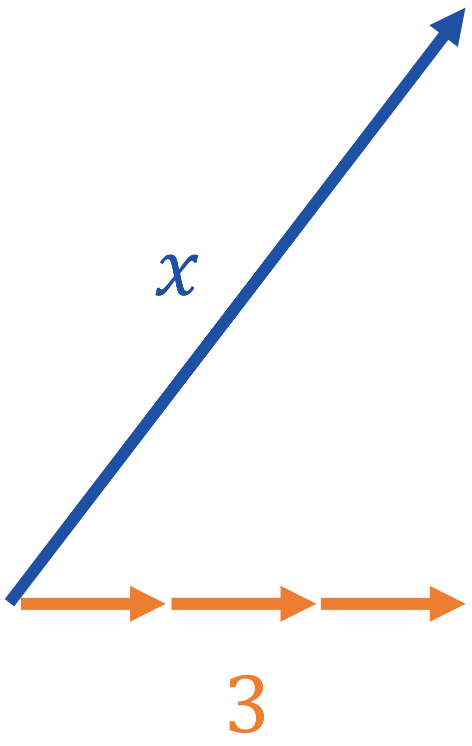
l_2-норма (Евклидова)

l_2-норма (Евклидова)



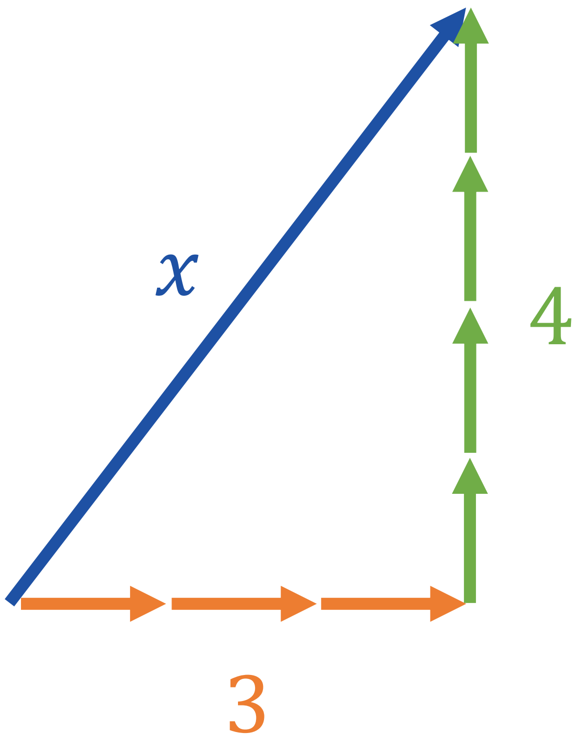
Чему равна длина вектора x?

l_2-норма (Евклидова)



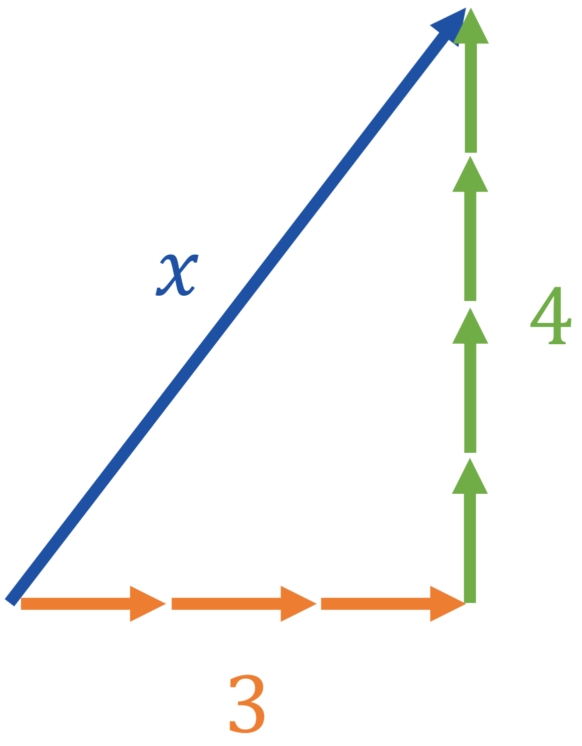
Чему равна длина вектора x?

l_2-норма (Евклидова)



Чему равна длина вектора x?

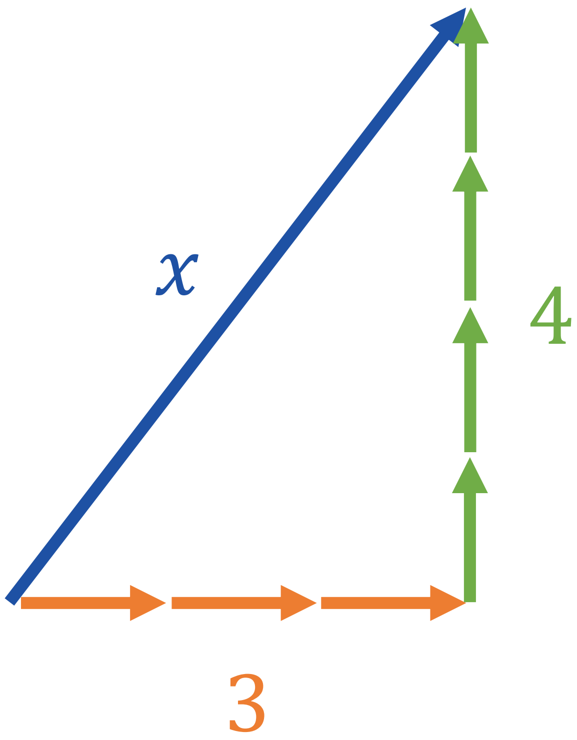
l_2-норма (Евклидова)



Чему равна длина вектора x?

$$\|x\|_2 = \sqrt{(\text{ширина})^2 + (\text{высота})^2}$$

l_2-норма (Евклидова)



Чему равна длина вектора x?

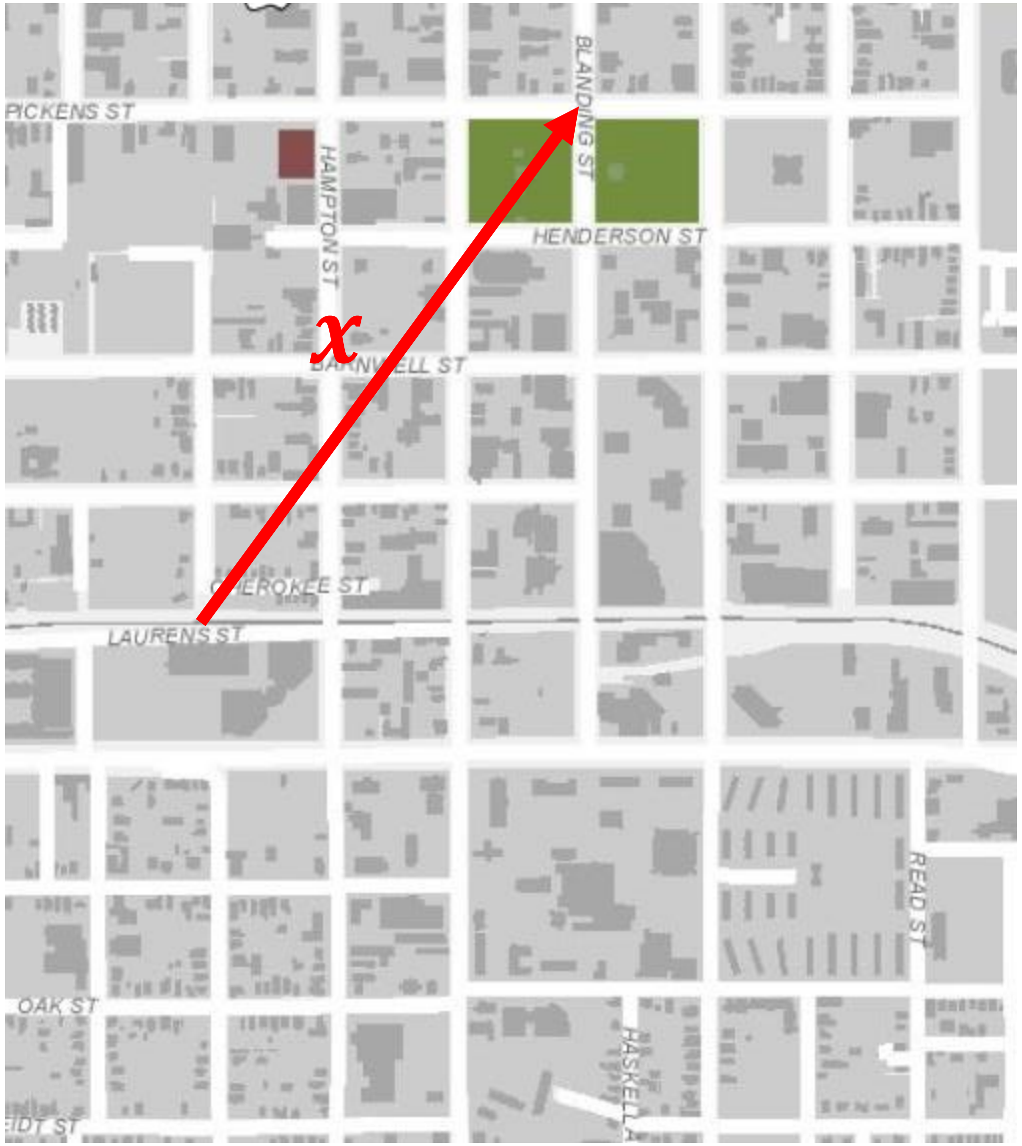
$$\|x\|_2 = \sqrt{(\text{ширина})^2 + (\text{высота})^2}$$

$$\|x\|_2 = 5$$

l_1-норма (Манхэттенская)

l_1-норма (Манхэттенская)

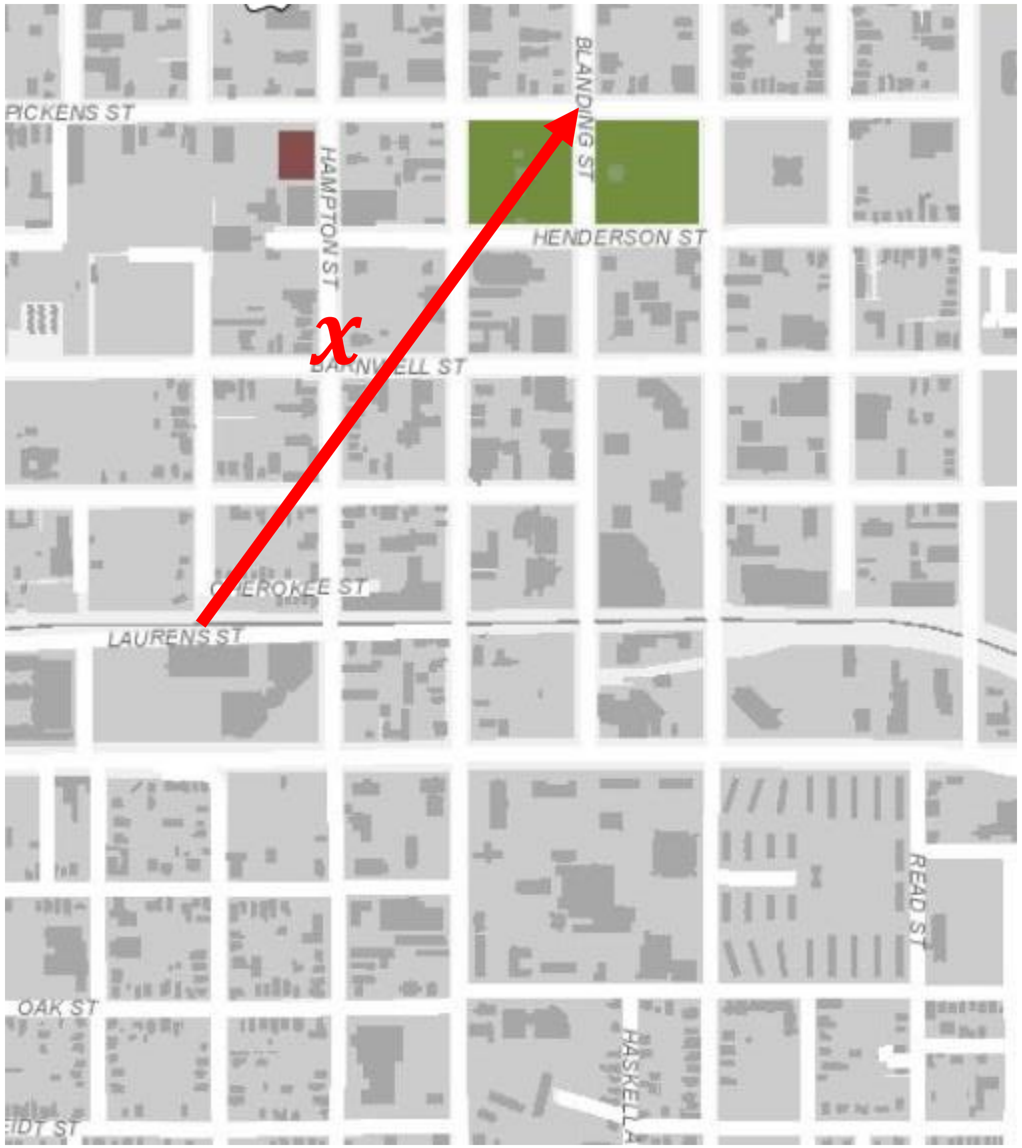
Вектор x соединяет два перекрёстка Манхэттена



l_1-норма (Манхэттенская)

Вектор x соединяет два перекрёстка Манхэттена

Какое расстояние придётся преодолеть, чтобы добраться от одного конца к другому?



l_1-норма (Манхэттенская)

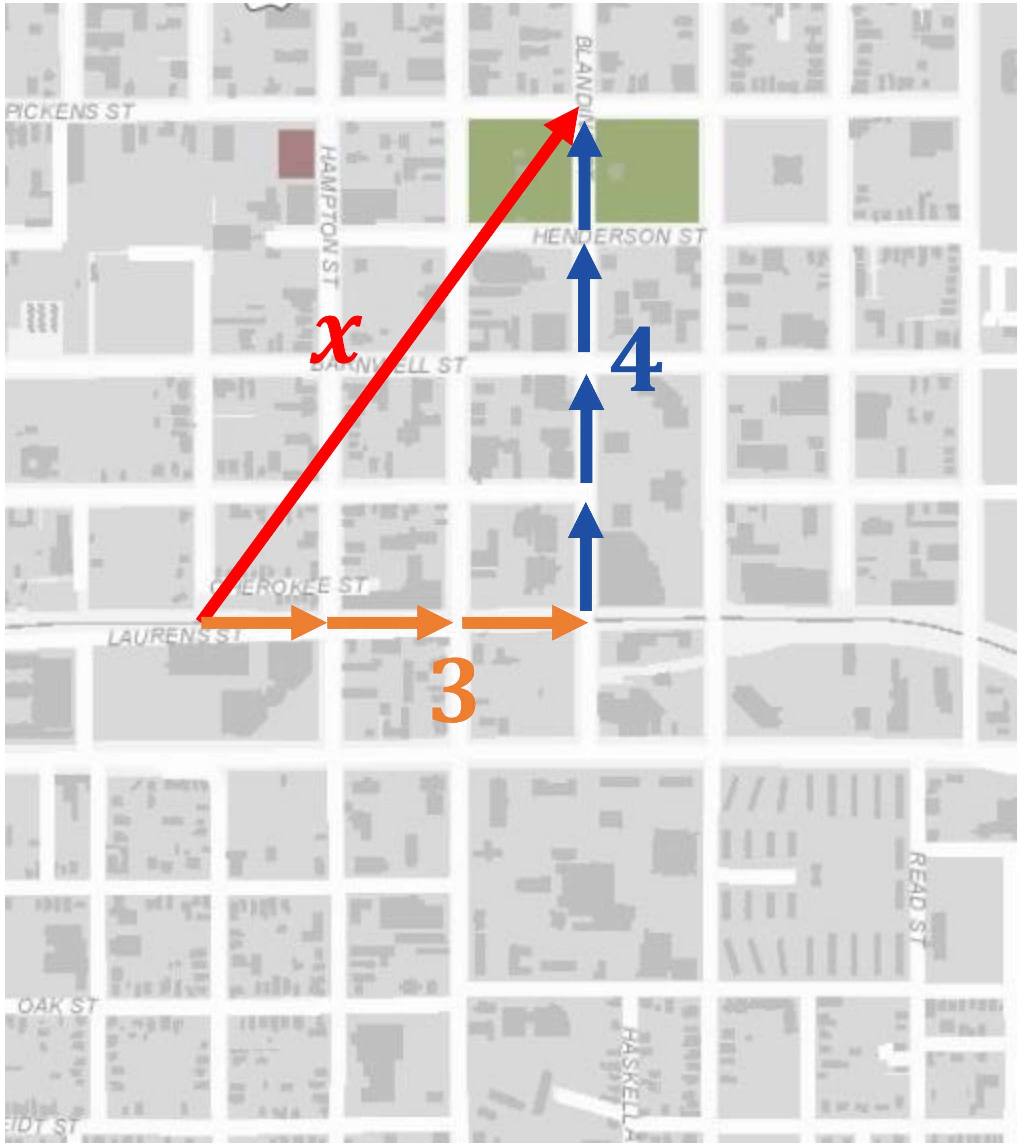
Вектор x соединяет два перекрёстка Манхэттена

Какое расстояние придётся преодолеть, чтобы добраться от одного конца к другому?

l_1-норма (Манхэттенская)

Вектор x соединяет два перекрёстка Манхэттена

Какое расстояние придётся преодолеть, чтобы добраться от одного конца к другому?

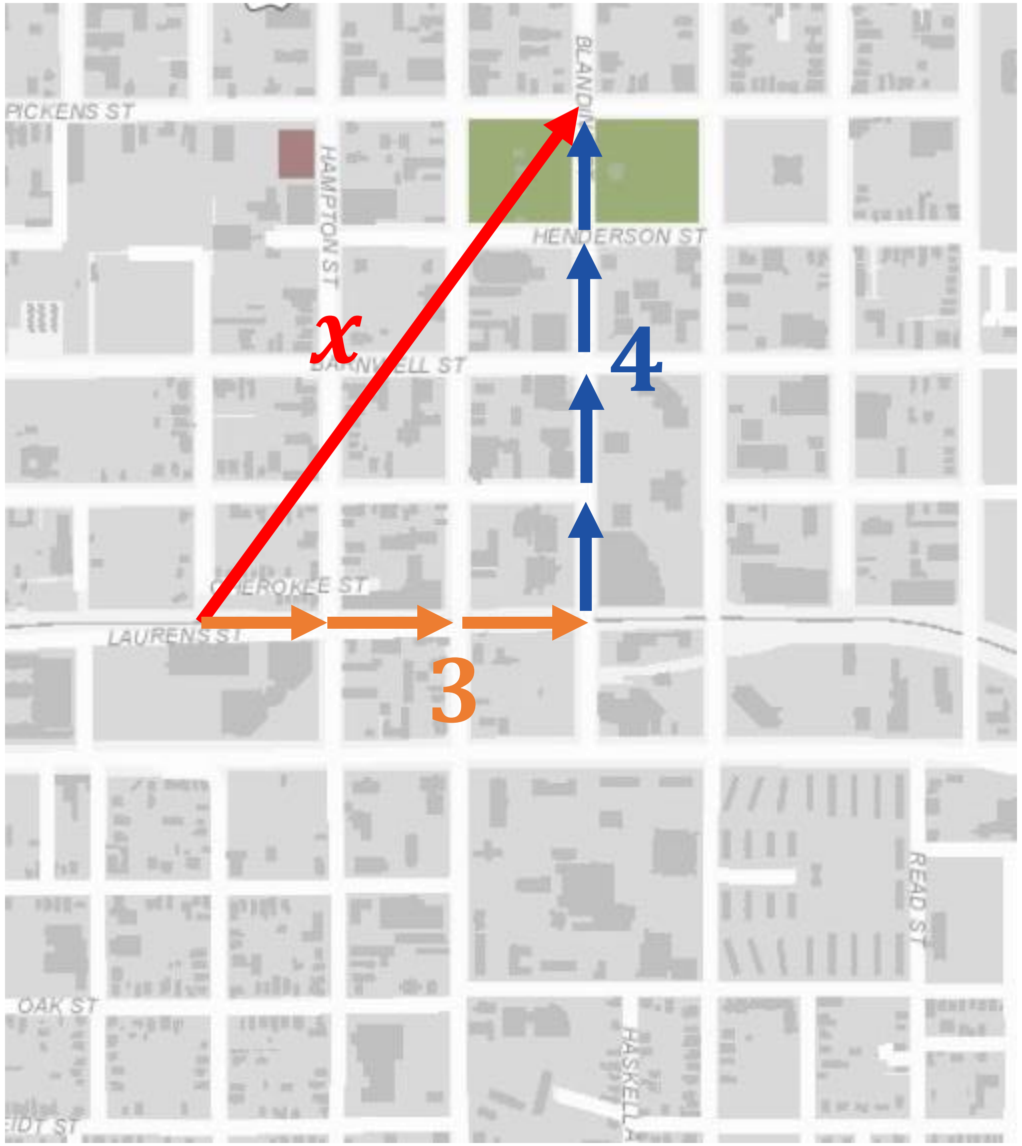


l_1-норма (Манхэттенская)

Вектор x соединяет два перекрёстка Манхэттена

Какое расстояние придётся преодолеть, чтобы добраться от одного конца к другому?

$$\|x\|_1 = |\text{ширина}| + |\text{высота}|$$



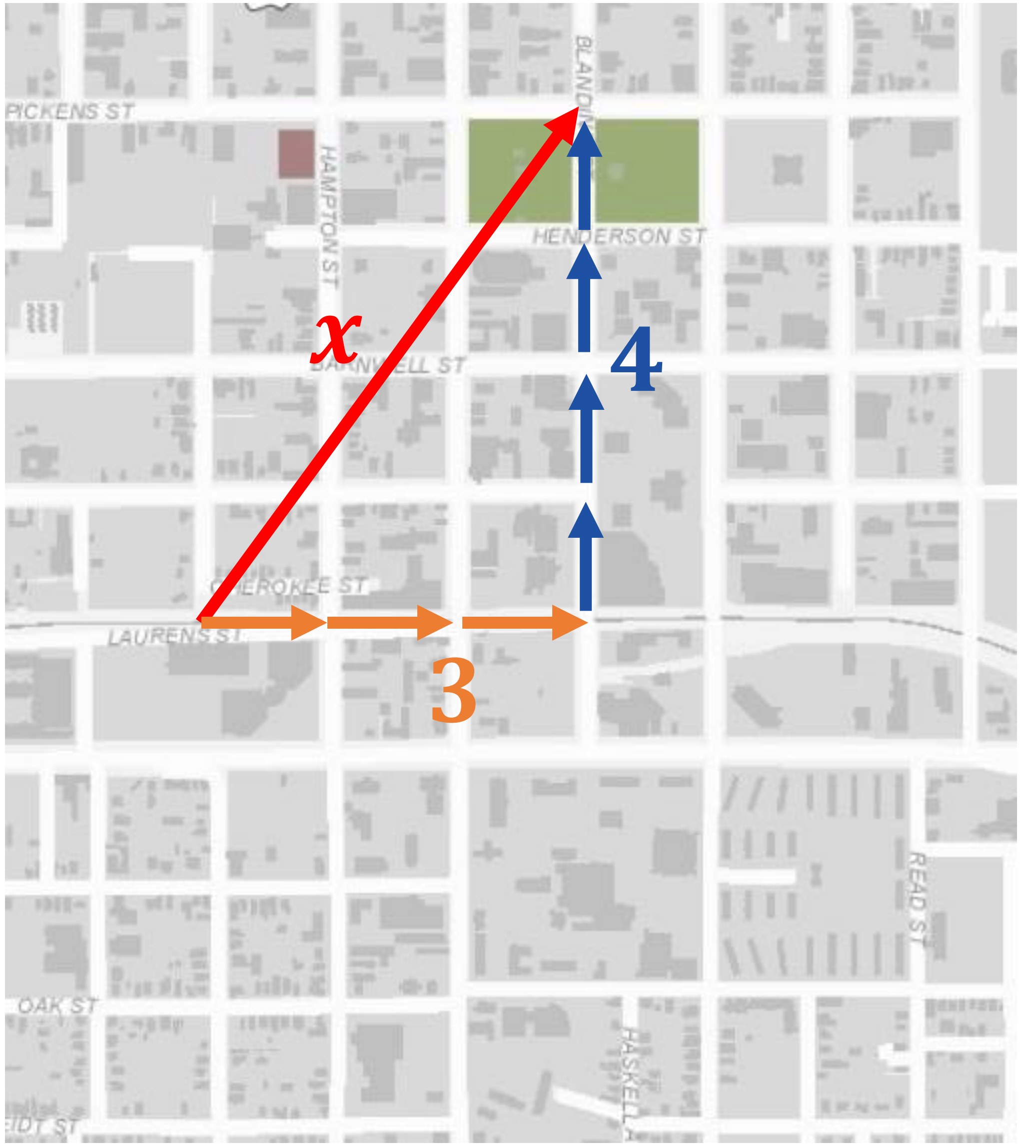
l_1-норма (Манхэттенская)

Вектор x соединяет два перекрёстка Манхэттена

Какое расстояние придётся преодолеть, чтобы добраться от одного конца к другому?

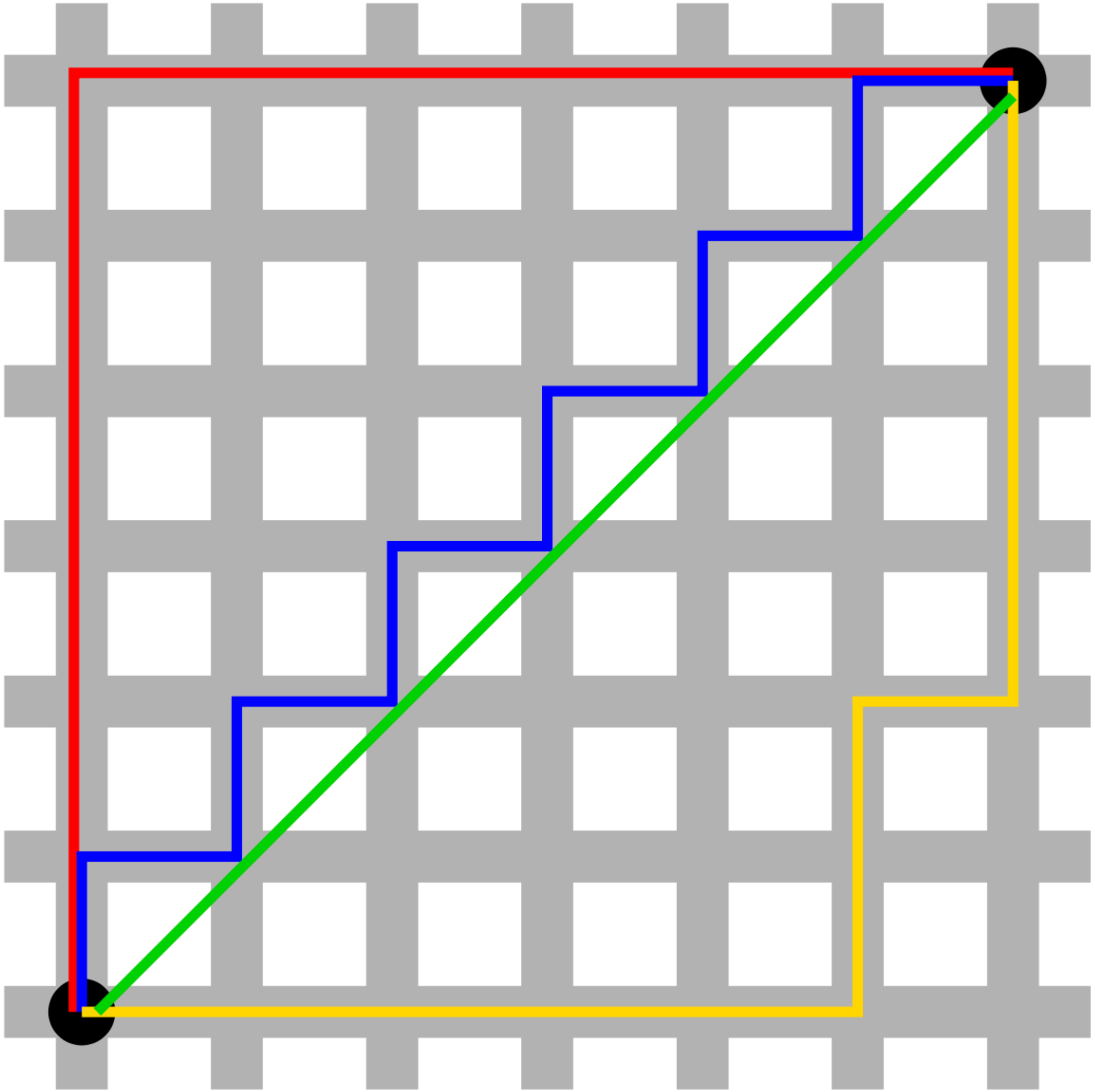
$$\|x\|_1 = |\text{ширина}| + |\text{высота}|$$

$$\|x\|_1 = 7$$



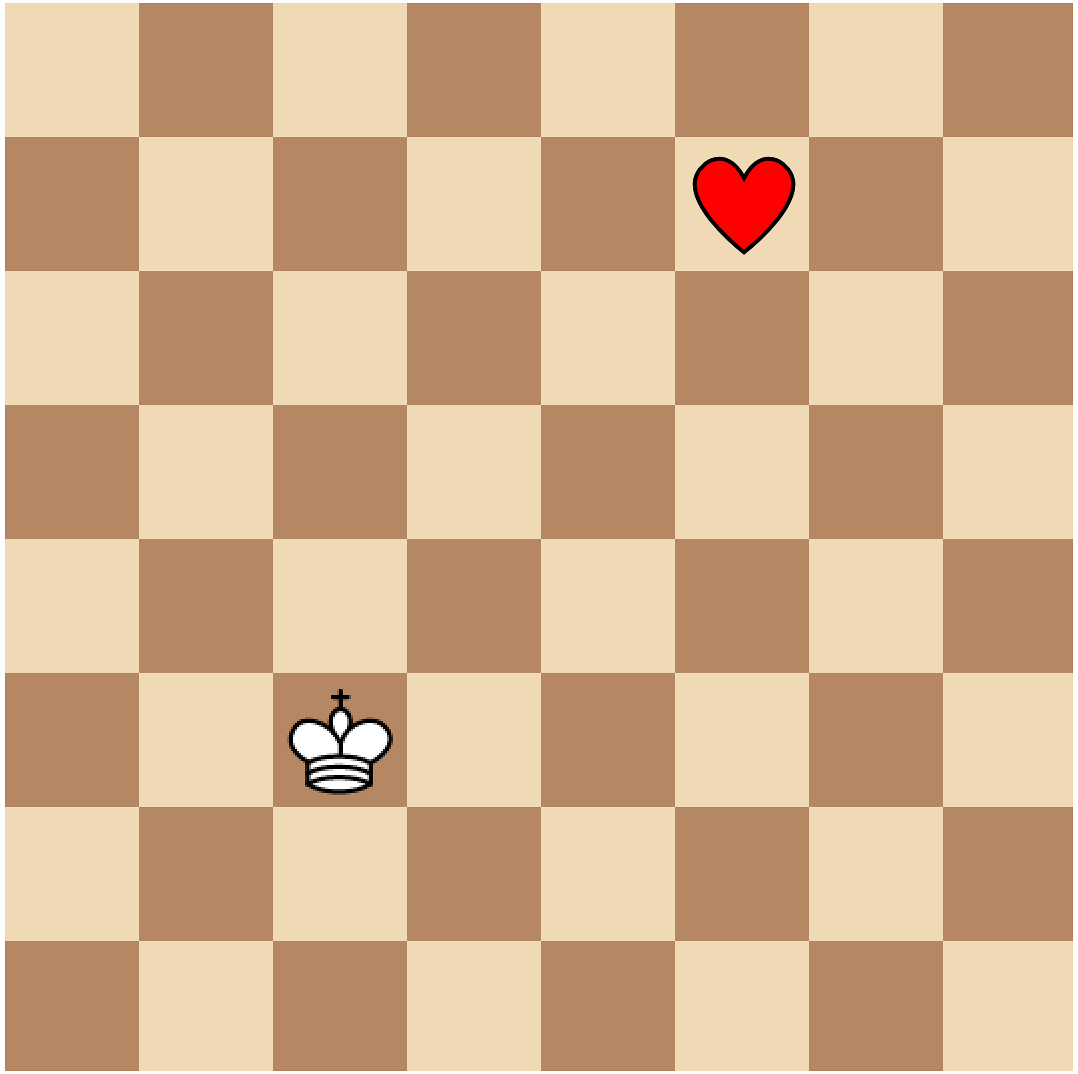
l_1-норма (Манхэттенская)

Разные пути могут дать одинаковое расстояние



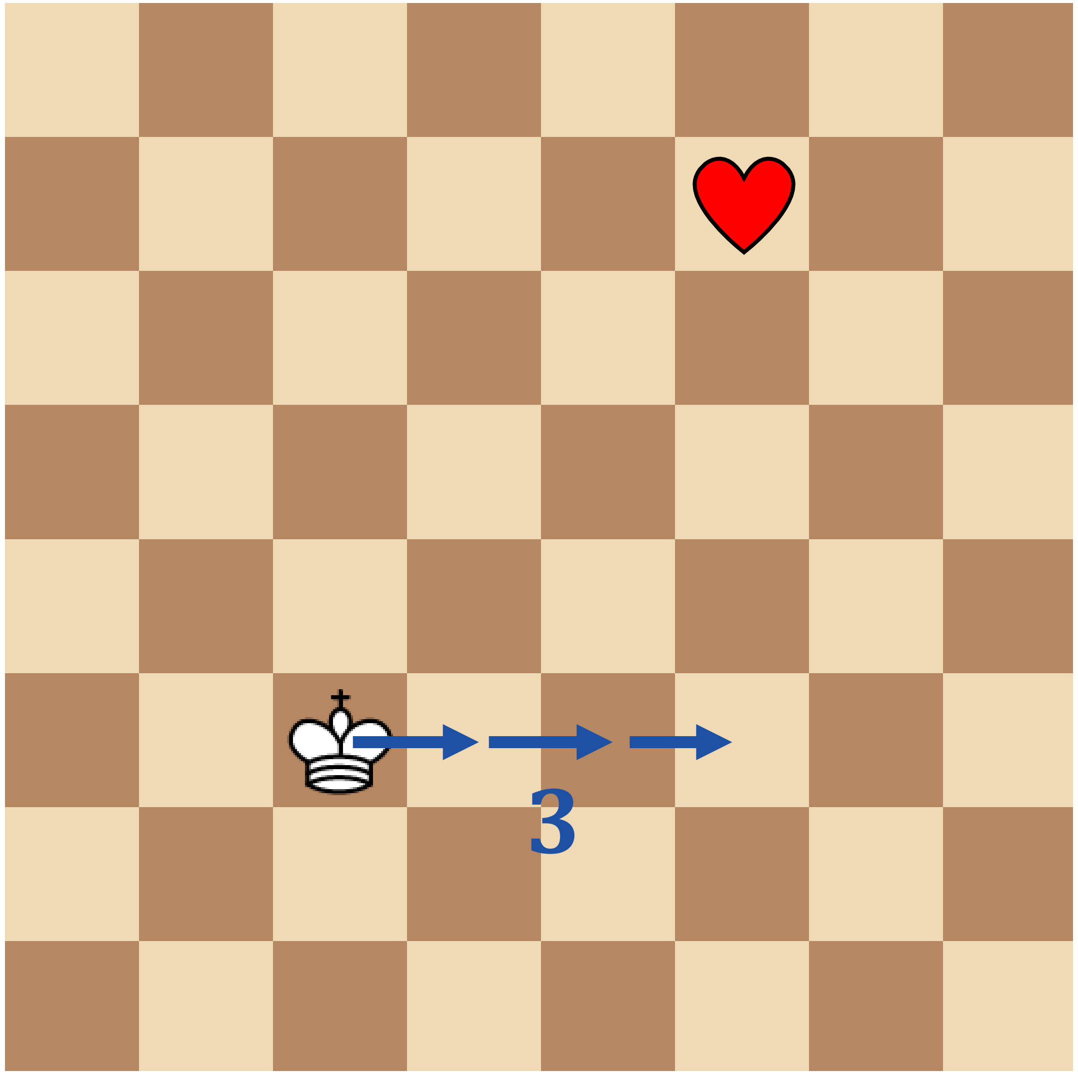
l_∞-норма (Чебышёва)

l_∞-норма (Чебышёва)



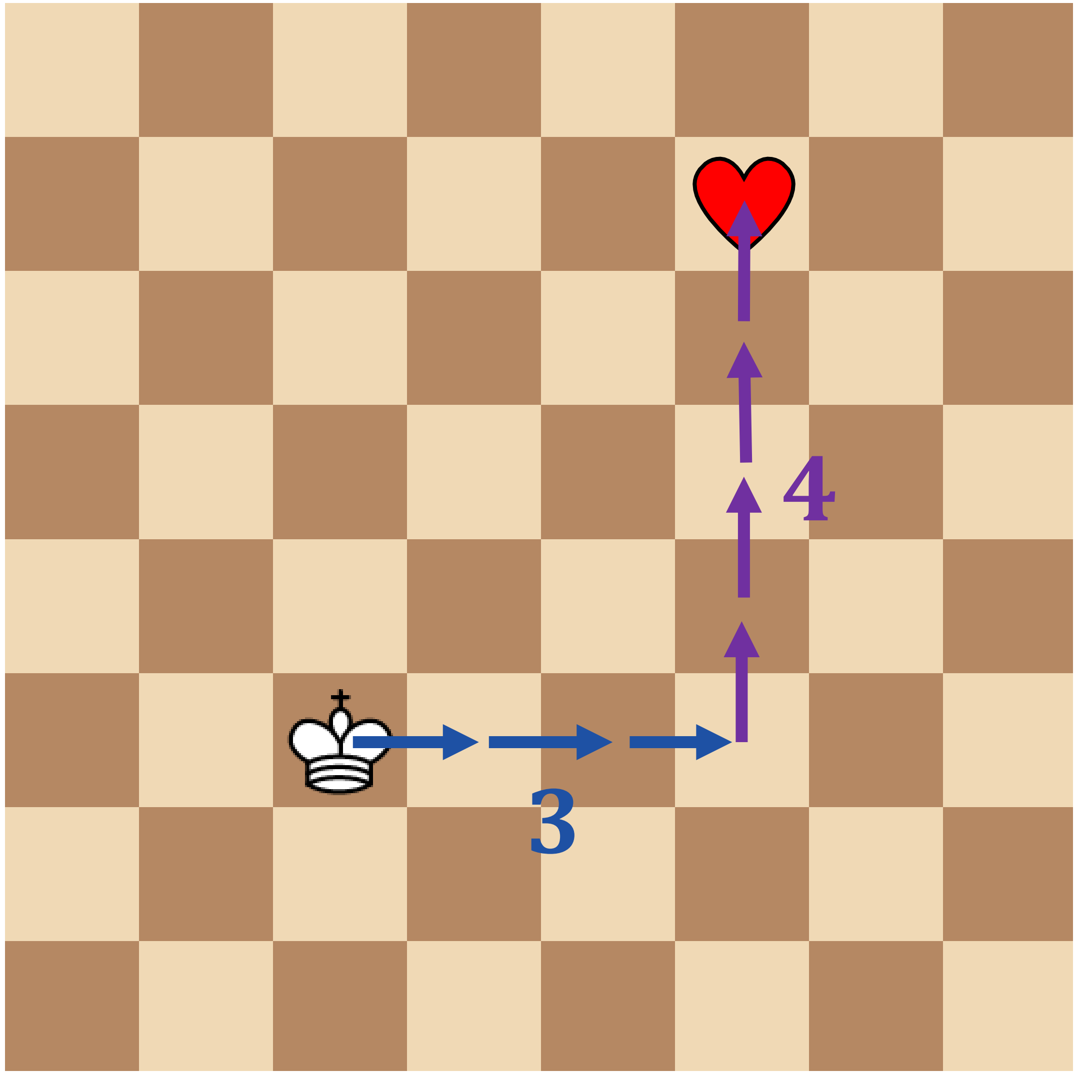
За сколько ходов король доберётся до своей любви?

l_∞-норма (Чебышёва)



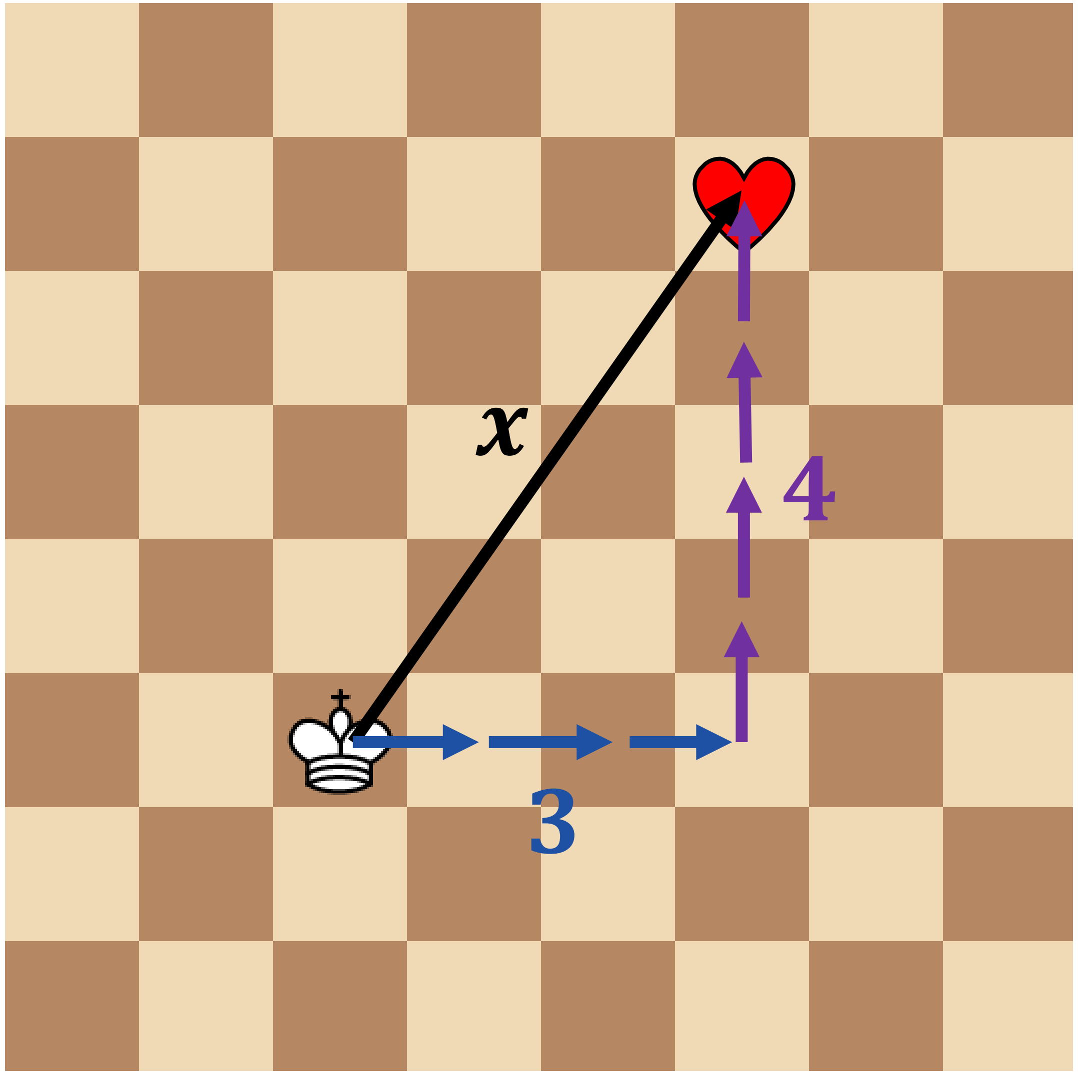
За сколько ходов король доберётся до своей любви?

l_∞-норма (Чебышёва)



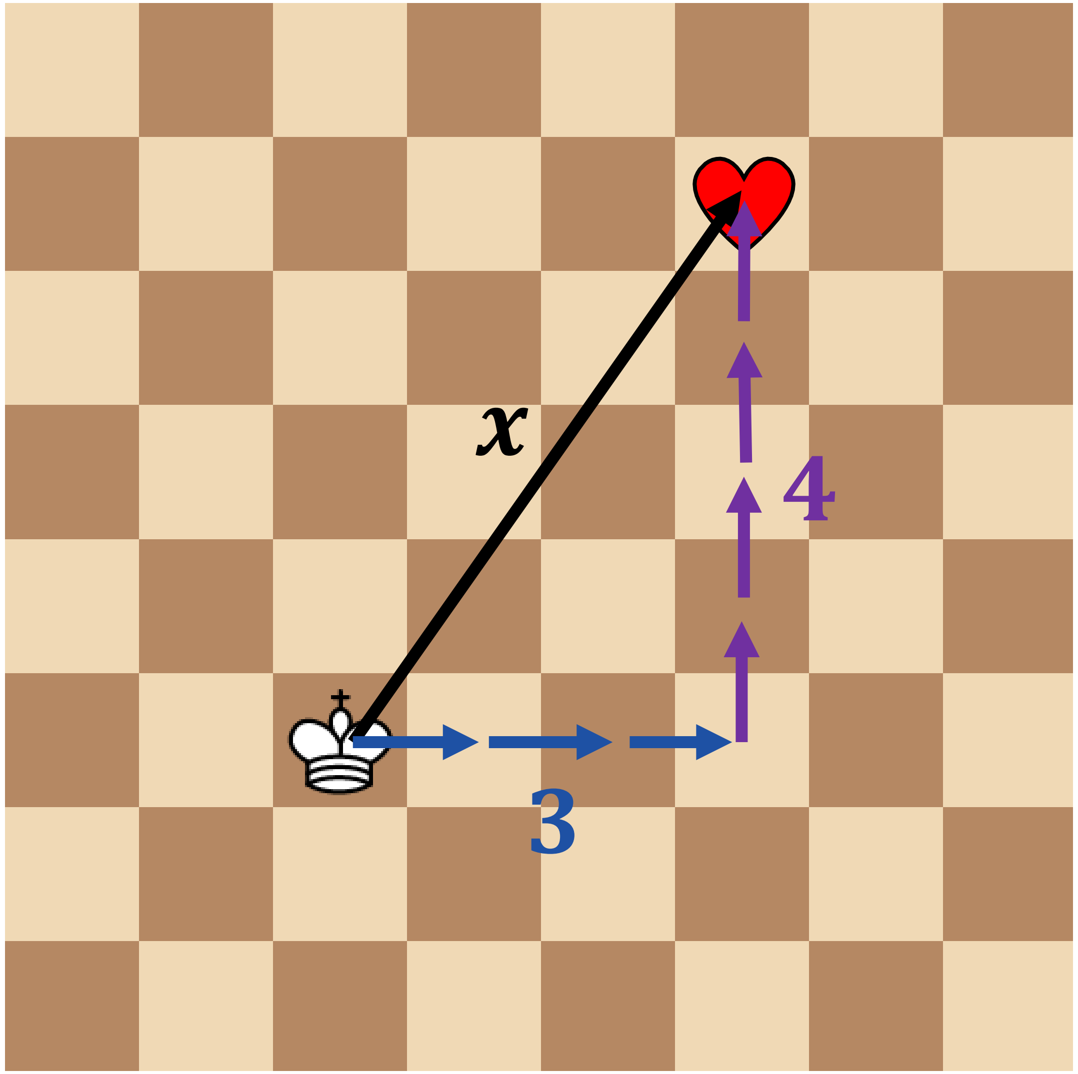
За сколько ходов король доберётся до своей любви?

l_∞-норма (Чебышёва)



За сколько ходов король доберётся до своей любви?

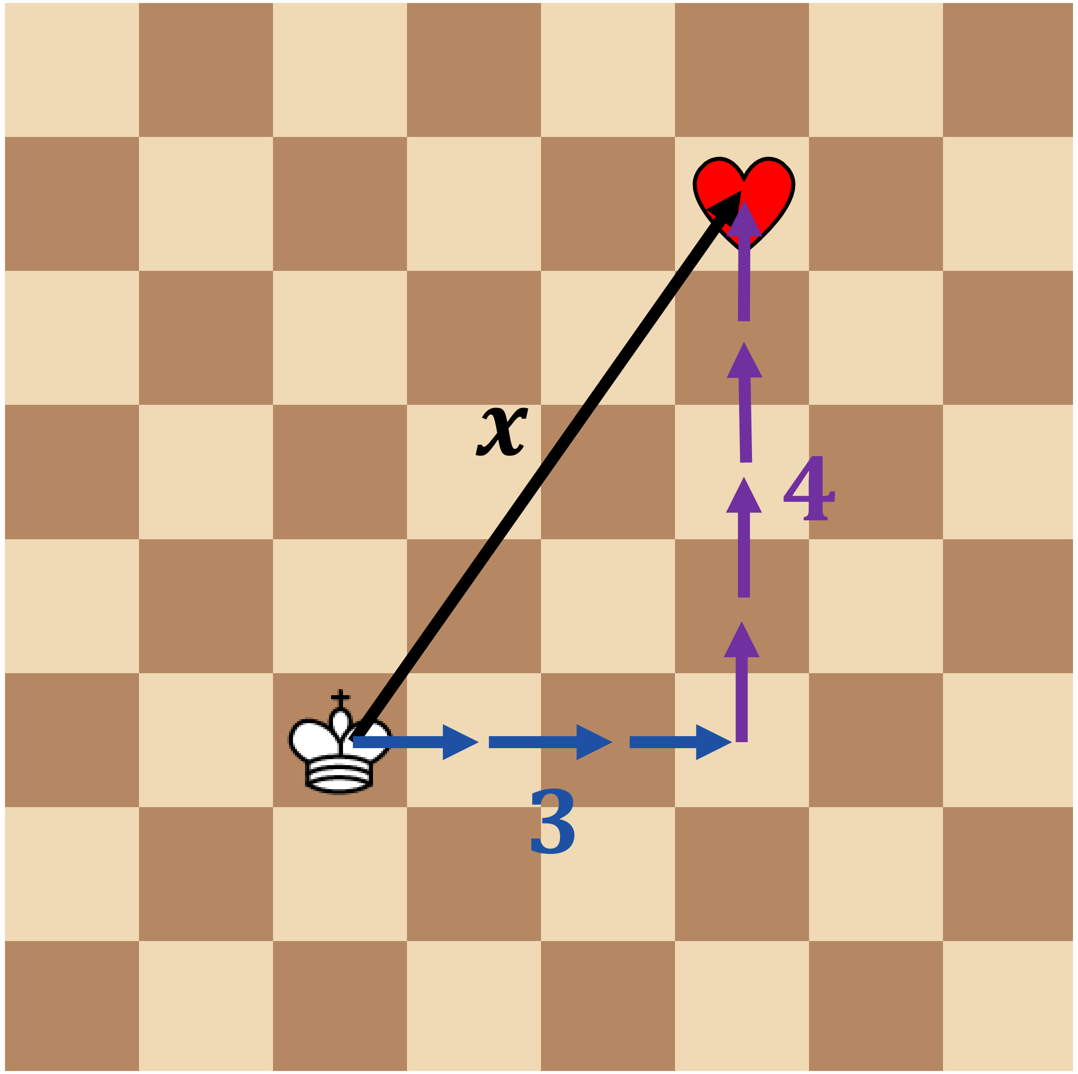
l_∞-норма (Чебышёва)



За сколько ходов король доберётся до своей любви?

$$\|x\|_\infty = \max\{\text{ширина}, \text{высота}\}$$

l_∞-норма (Чебышёва)



За сколько ходов король доберётся до своей любви?

$$\|x\|_\infty = \max\{\text{ширина}, \text{высота}\}$$

$$\|x\|_\infty = 4$$

l_p-нормы

l_p-нормы

Пусть $x = (x_1, x_2, \dots, x_n)$ – вектор с n компонентами

l_p-нормы

Пусть $x = (x_1, x_2, \dots, x_n)$ – вектор с n компонентами

l_1-норма (Манхэттенская) $\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$

l_p-нормы

Пусть $x = (x_1, x_2, \dots, x_n)$ – вектор с n компонентами

l_1-норма (Манхэттенская) $\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$

l_2-норма (Евклидова) $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$

l_p-нормы

Пусть $x = (x_1, x_2, \dots, x_n)$ – вектор с n компонентами

l_1-норма (Манхэттенская) $\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$

l_2-норма (Евклидова) $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$

l_∞-норма (Чебышёва) $\|x\|_\infty = \max\{|x_1|, |x_2|, \dots, |x_n|\}$

l_p-нормы

Пусть $x = (x_1, x_2, \ldots, x_n)$ – вектор с n компонентами

l_1-норма (Манхэттенская) $\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$

l_2-норма (Евклидова) $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$

l_∞-норма (Чебышёва) $\|x\|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$

l_p-норма (общая) $\|x\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots + |x_n|^p}$

l_p-нормы

Пусть $x = (x_1, x_2, \dots, x_n)$ – вектор с n компонентами

l_1-норма (Манхэттенская) $\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$ $p = 1$

l_2-норма (Евклидова) $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ $p = 2$

l_∞-норма (Чебышёва) $\|x\|_\infty = \max\{|x_1|, |x_2|, \dots, |x_n|\}$ $p \to \infty$

l_p-норма (общая) $\|x\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots + |x_n|^p}$

l_p-нормы

Пусть $x = (x_1, x_2, \dots, x_n)$ – вектор с n компонентами

l_1-норма (Манхэттенская) $\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$ $p = 1$

l_2-норма (Евклидова) $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ $p = 2$

l_∞-норма (Чебышёва) $\|x\|_\infty = \max\{|x_1|, |x_2|, \dots, |x_n|\}$ $p \to \infty$

l_p-норма (общая) $\|x\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots + |x_n|^p}$

(Бывают и совсем другие нормы, которые не l_p)

Пример

$$x = (3, -4)$$

Пример

$$x = (3, -4)$$

$$\|x\|_1 = |3| + |-4| = \mathbf{7}$$

Сумма модулей координат

Пример

$$x = (3, -4)$$

$$\|x\|_1 = |3| + |-4| = \mathbf{7}$$

$$\|x\|_2 = \sqrt{(3)^2 + (-4)^2} = \mathbf{5}$$

Корень из суммы
квадратов координат

Пример

$$x = (3, -4)$$

$$\|x\|_1 = |3| + |-4| = \mathbf{7}$$

$$\|x\|_2 = \sqrt{(3)^2 + (-4)^2} = \mathbf{5}$$

$$\|x\|_\infty = \max\{|3|, |-4|\} = \mathbf{4}$$

Наибольший модуль координаты

Пример

$$x = (3, -4)$$

$$\|x\|_1 = |3| + |-4| = \mathbf{7}$$

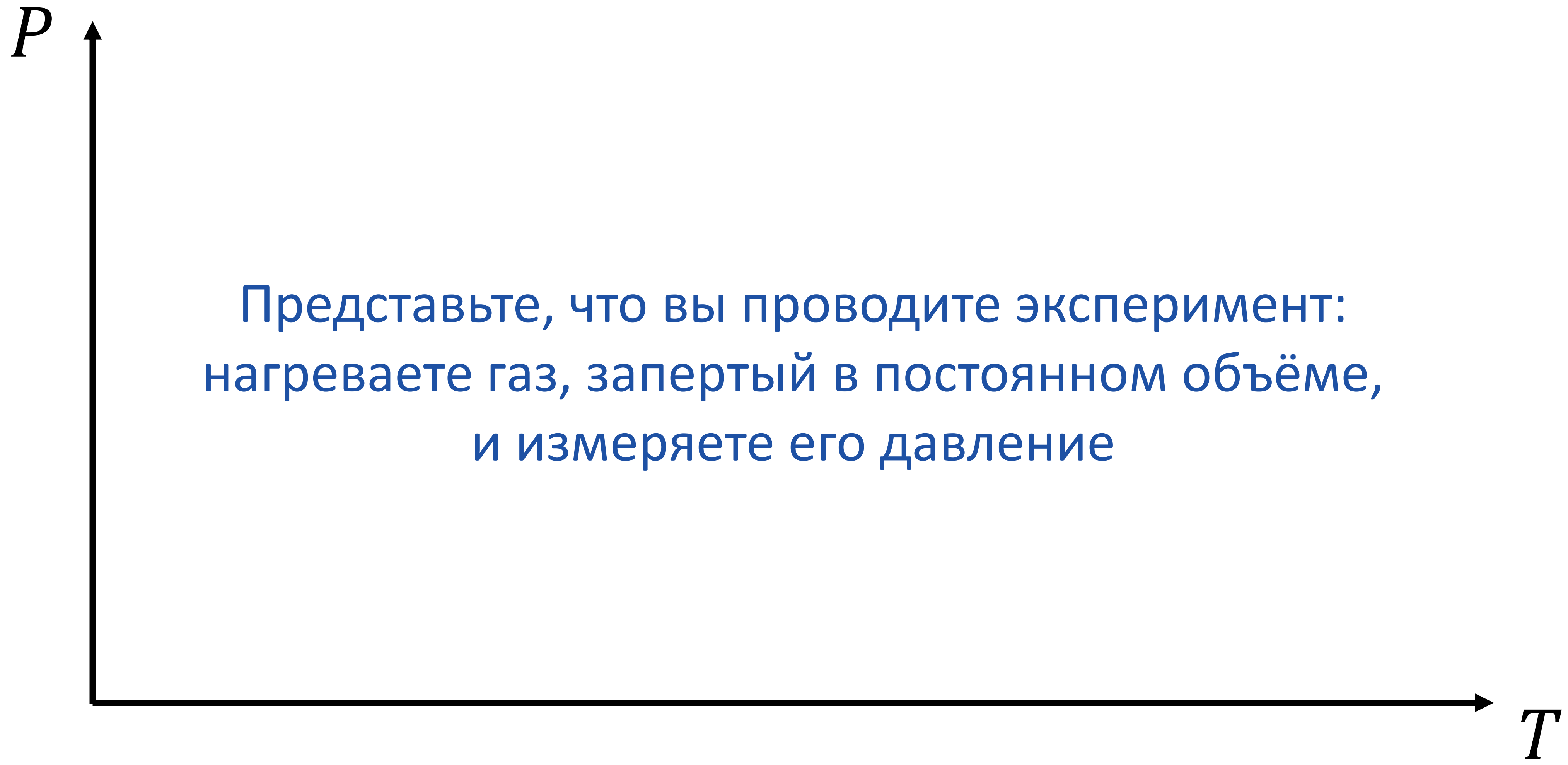
$$\|x\|_2 = \sqrt{(3)^2 + (-4)^2} = \mathbf{5}$$

$$\|x\|_\infty = \max\{|3|, |-4|\} = \mathbf{4}$$

Вообще, всегда $\|x\|_\infty \le \|x\|_2 \le \|x\|_1$

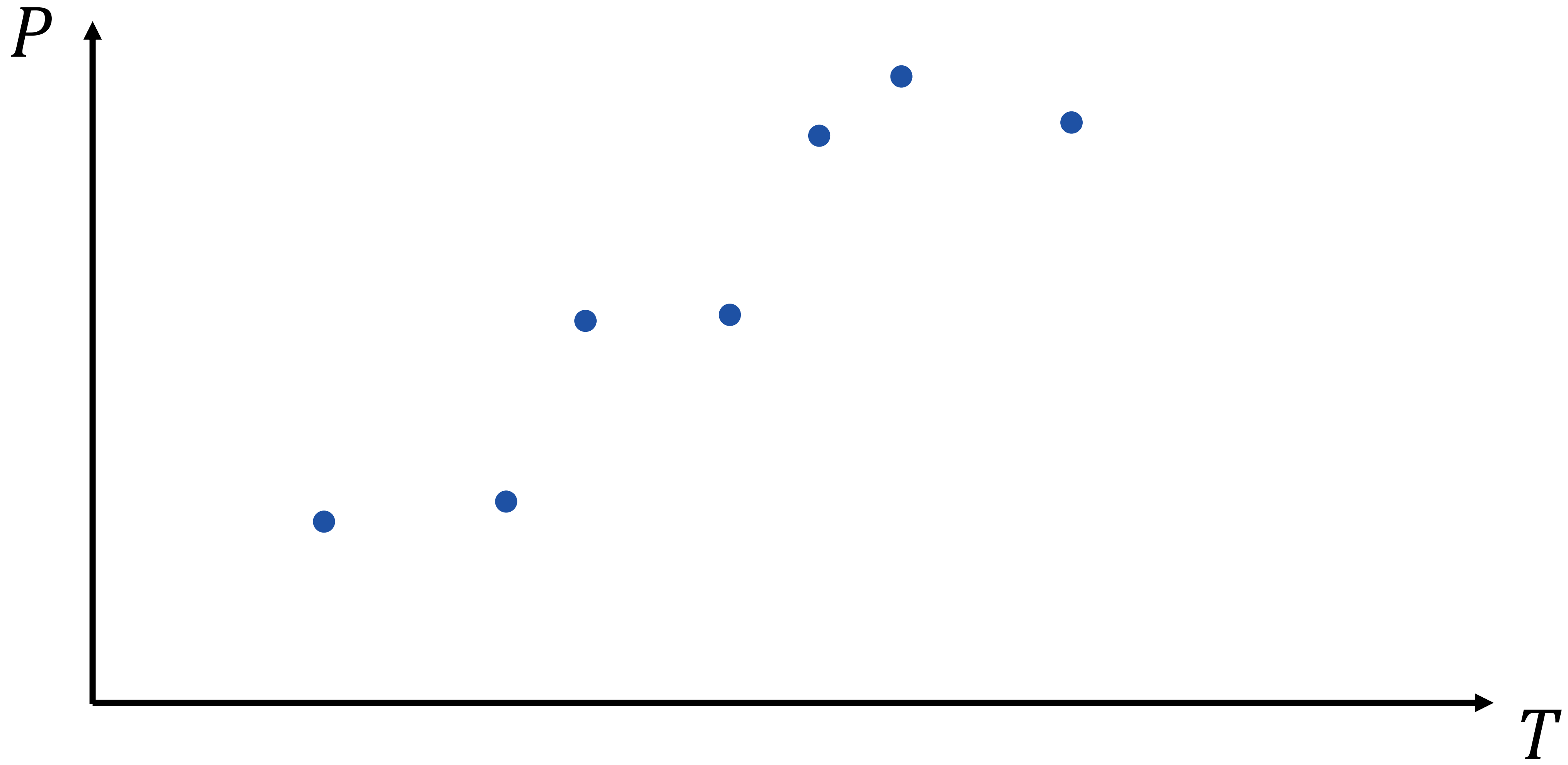
Интерполяция и аппроксимация

Интерполяция и аппроксимация

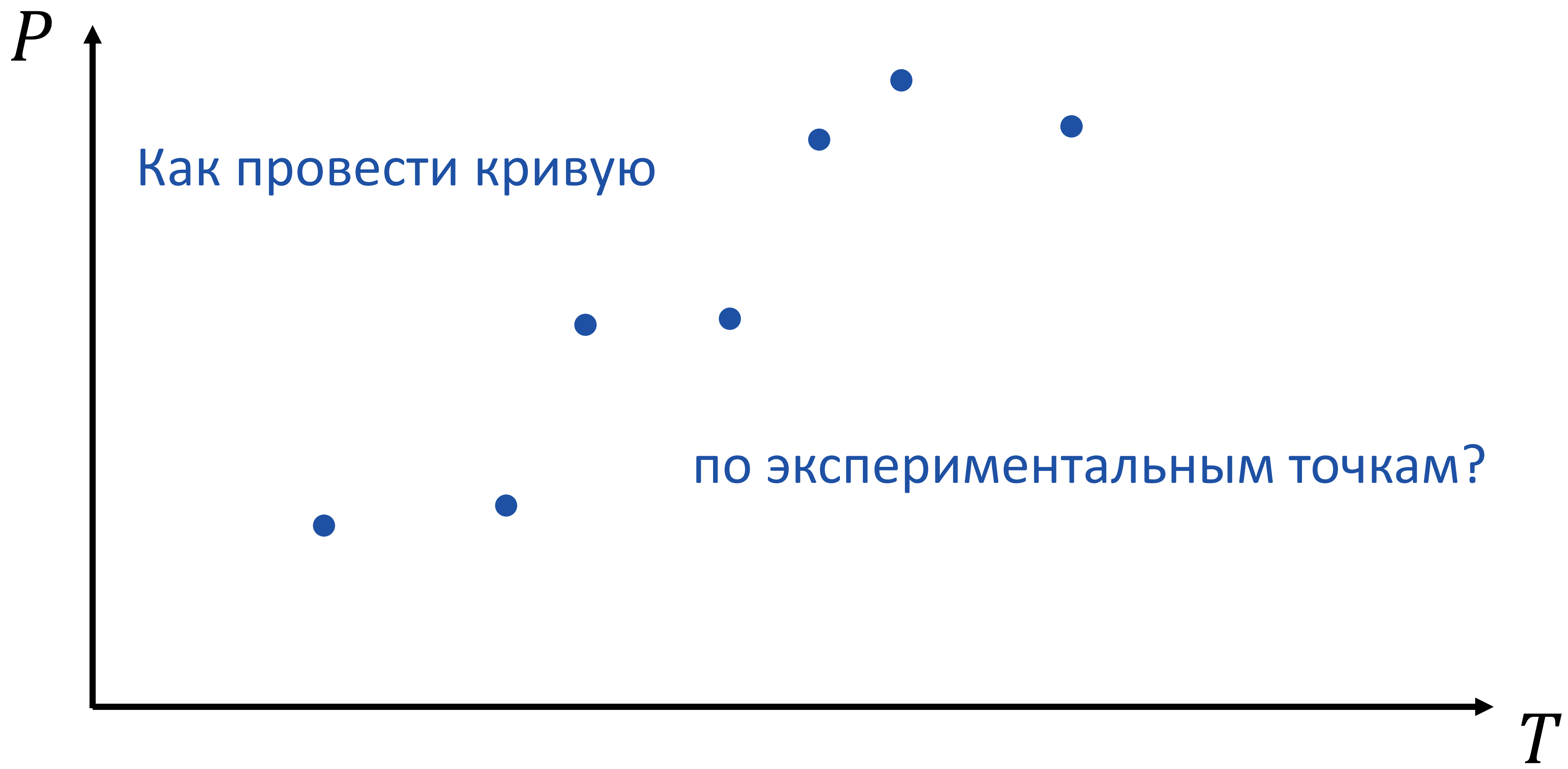


Представьте, что вы проводите эксперимент: нагреваете газ, запертый в постоянном объёме, и измеряете его давление

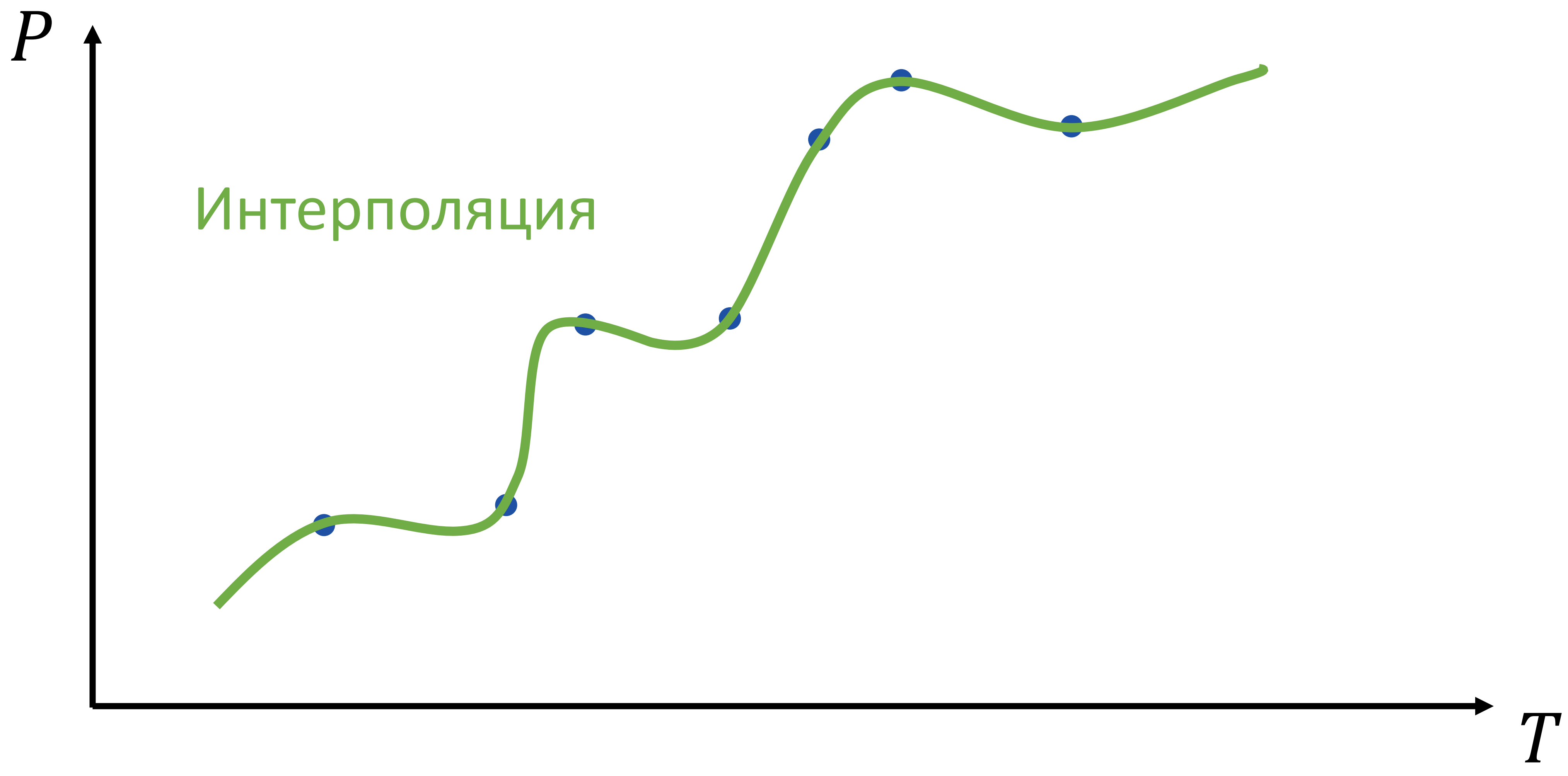
Интерполяция и аппроксимация



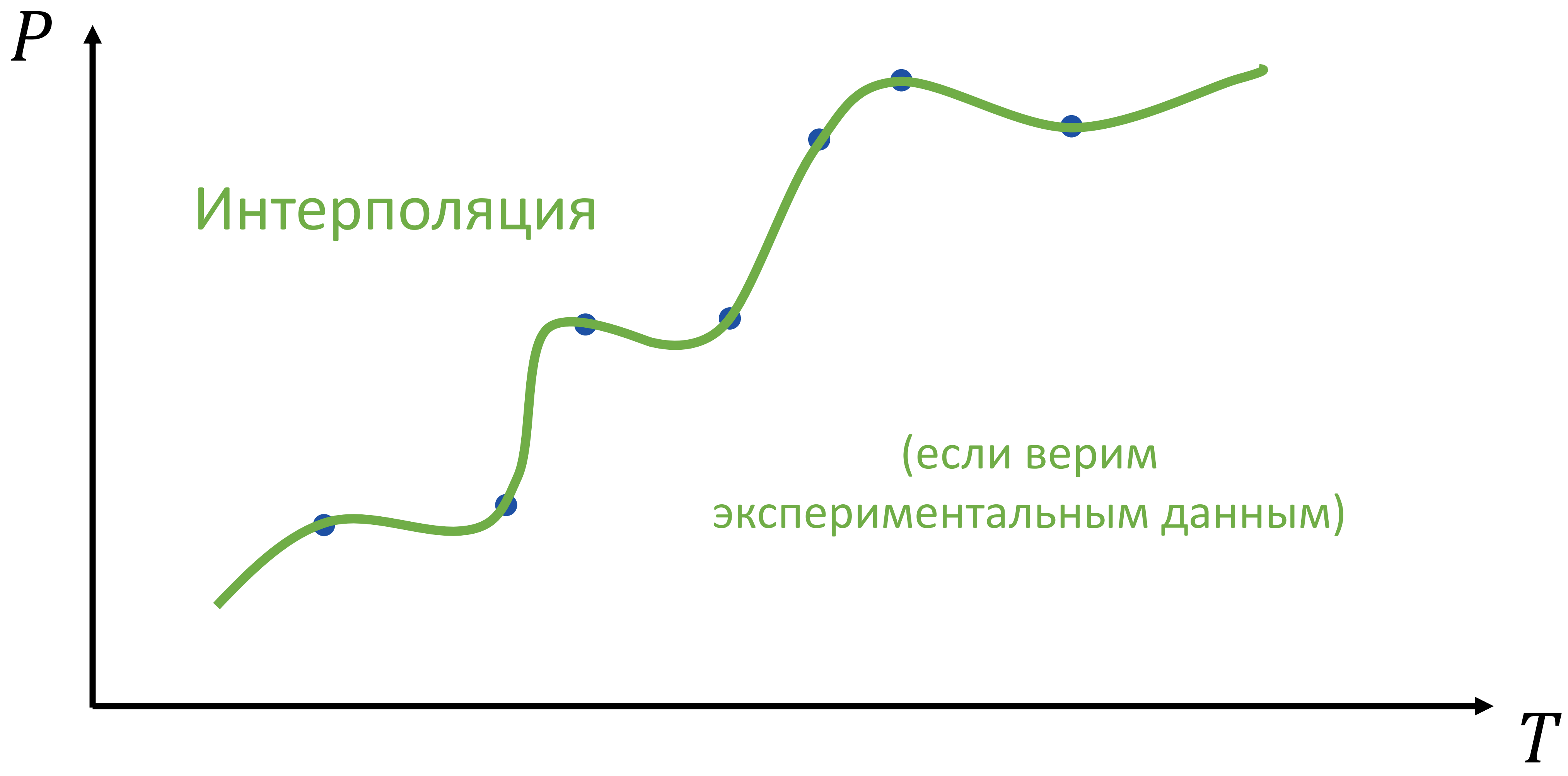
Интерполяция и аппроксимация



Интерполяция и аппроксимация



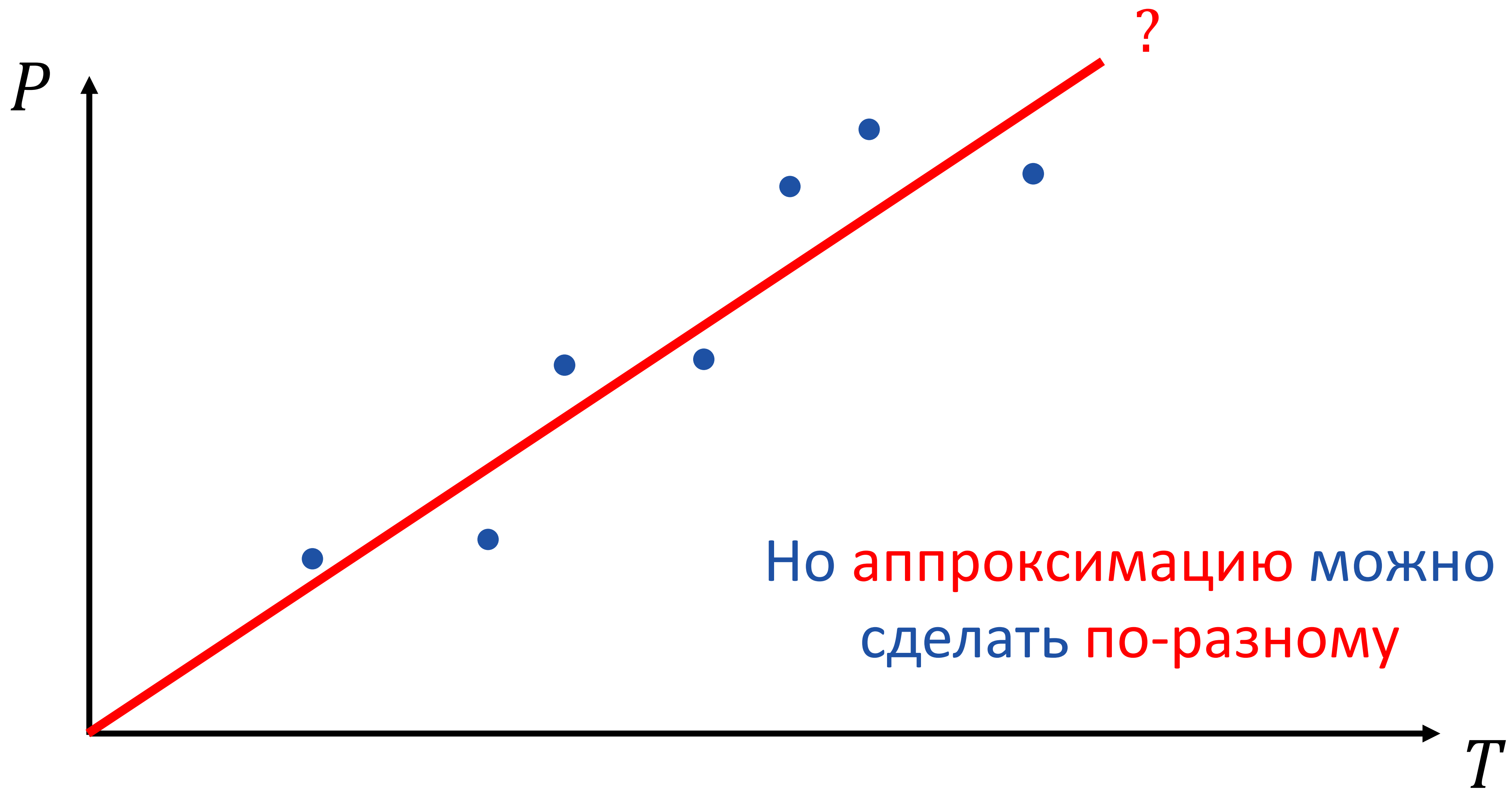
Интерполяция и аппроксимация



Интерполяция и аппроксимация

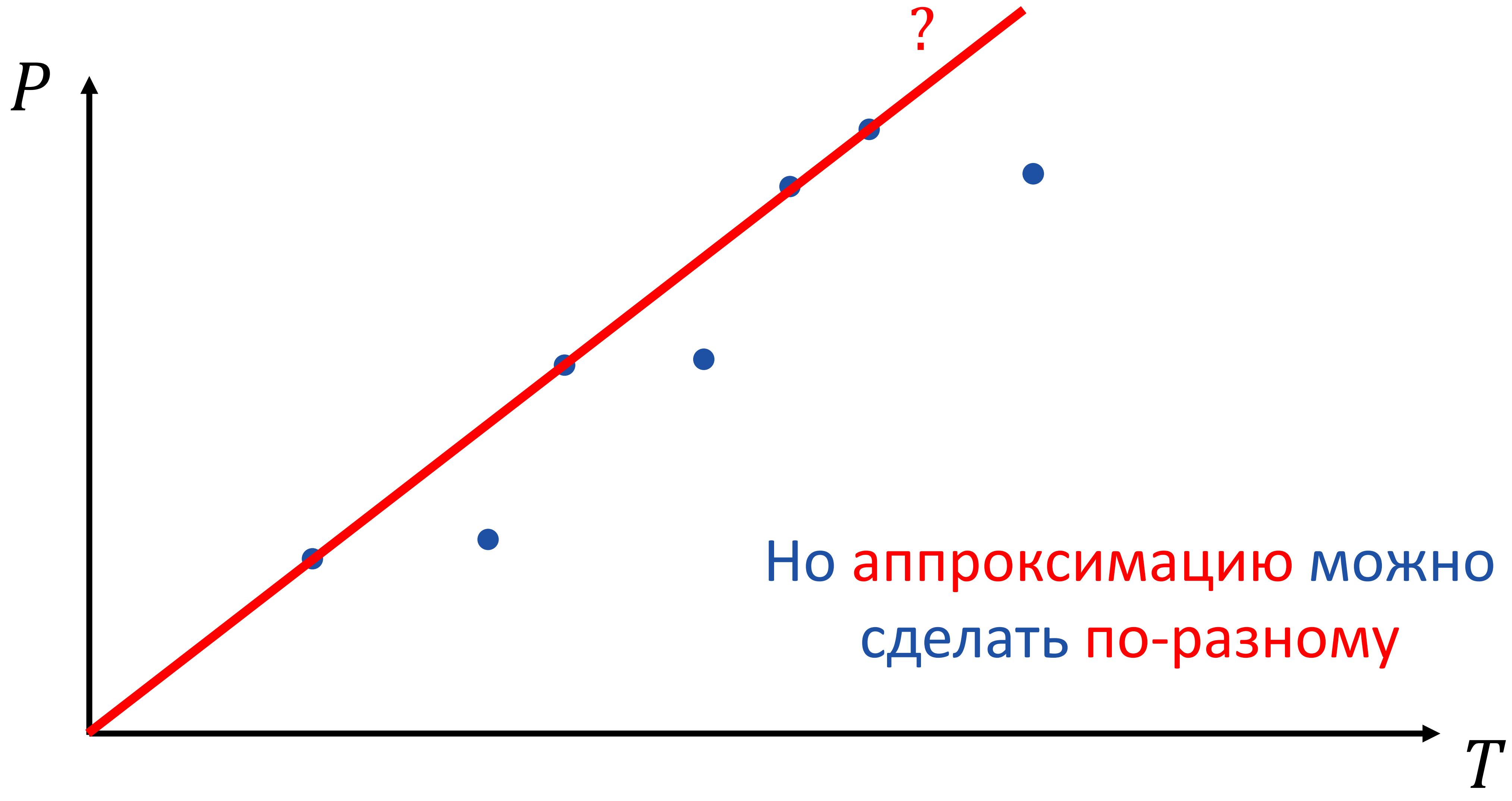
Интерполяция и аппроксимация

Интерполяция и аппроксимация



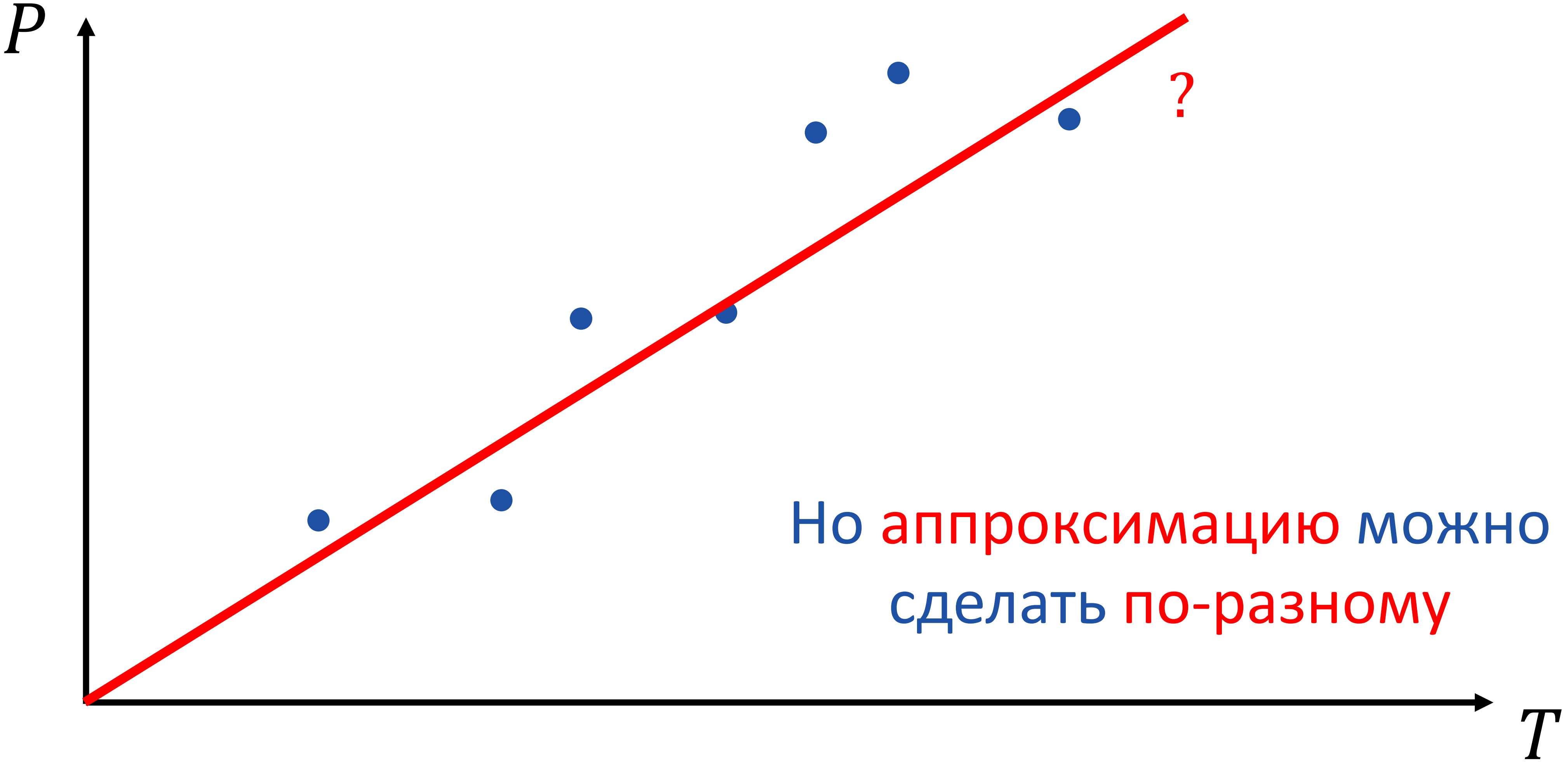
Но аппроксимацию можно сделать по-разному

Интерполяция и аппроксимация



Но аппроксимацию можно сделать по-разному

Интерполяция и аппроксимация



Но аппроксимацию можно сделать по-разному

Интерполяция и аппроксимация

Отклонение функции от набора точек

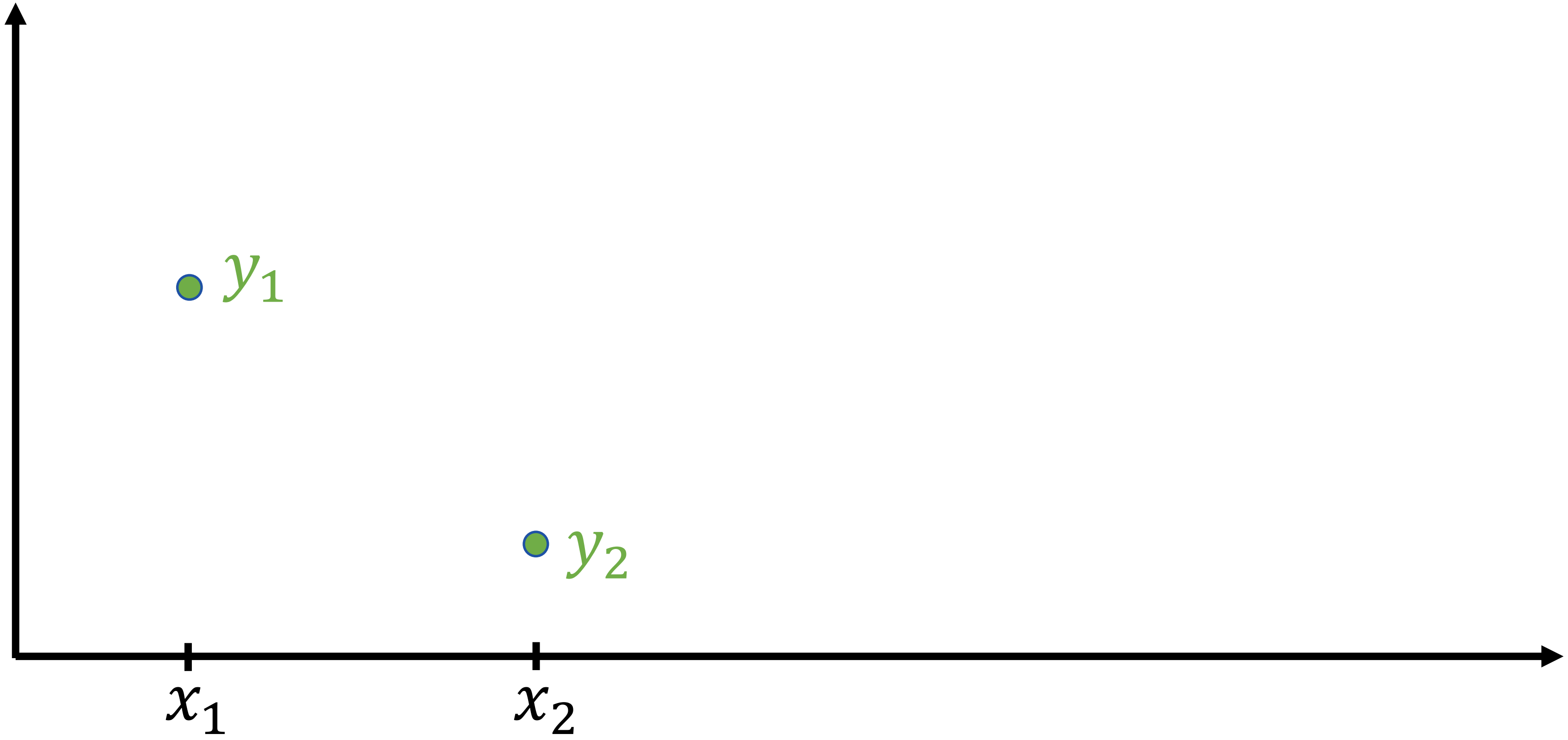
Отклонение функции от набора точек

Отклонение функции от набора точек

Точки

(x_1, y_1)

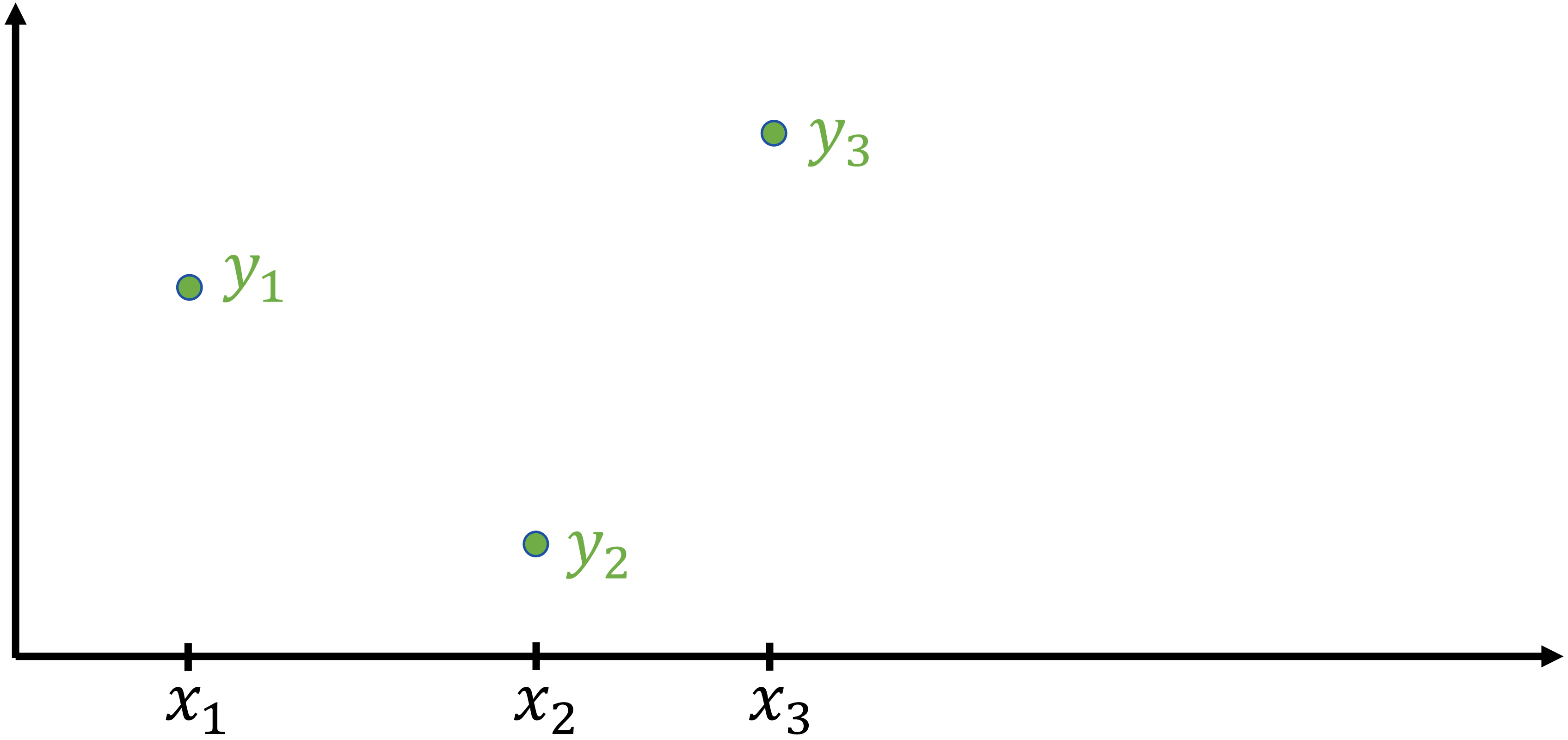
Отклонение функции от набора точек



Точки

(x_2, y_2)

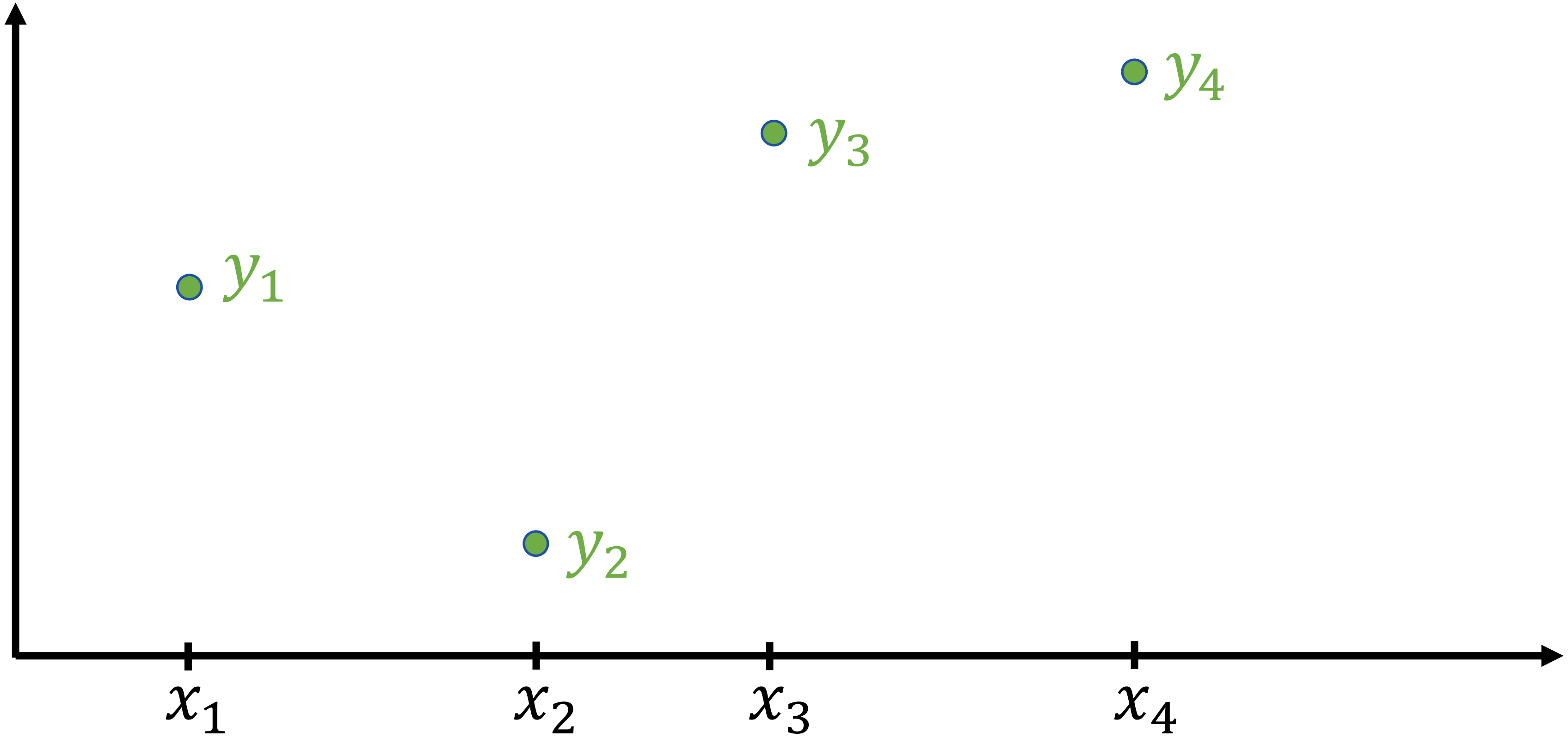
Отклонение функции от набора точек



Точки

(x_3, y_3)

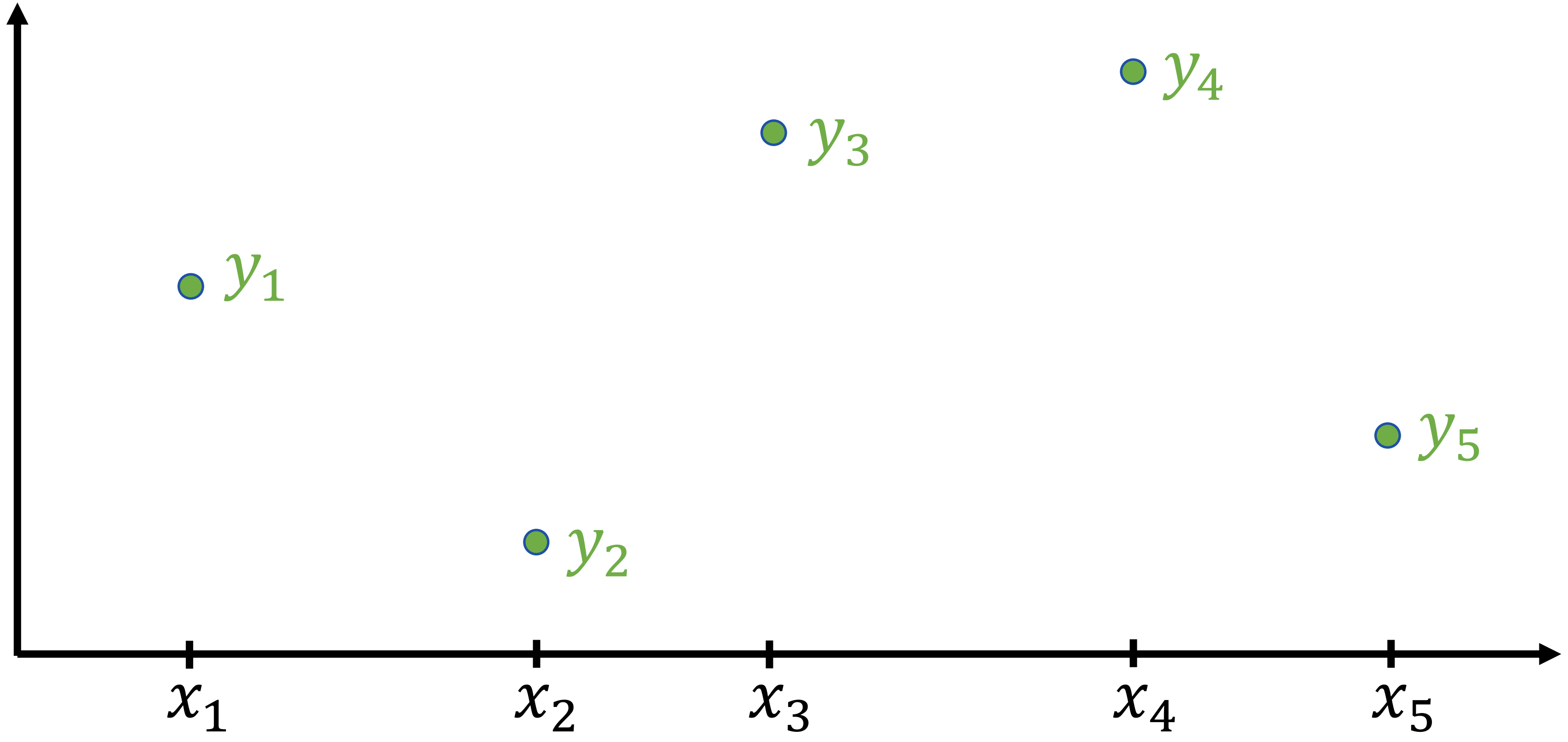
Отклонение функции от набора точек



Точки

(x_4, y_4)

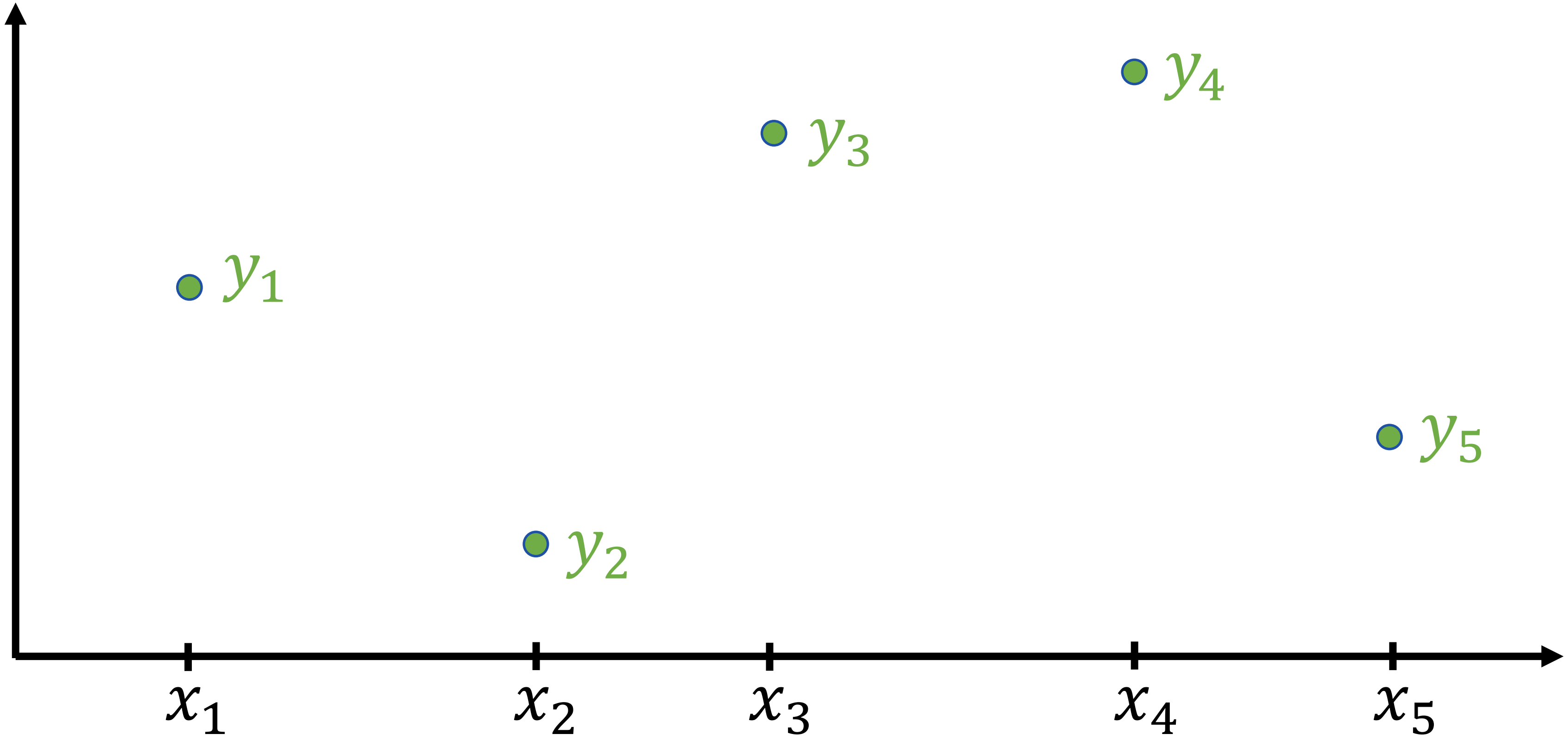
Отклонение функции от набора точек



Точки

(x_5, y_5)

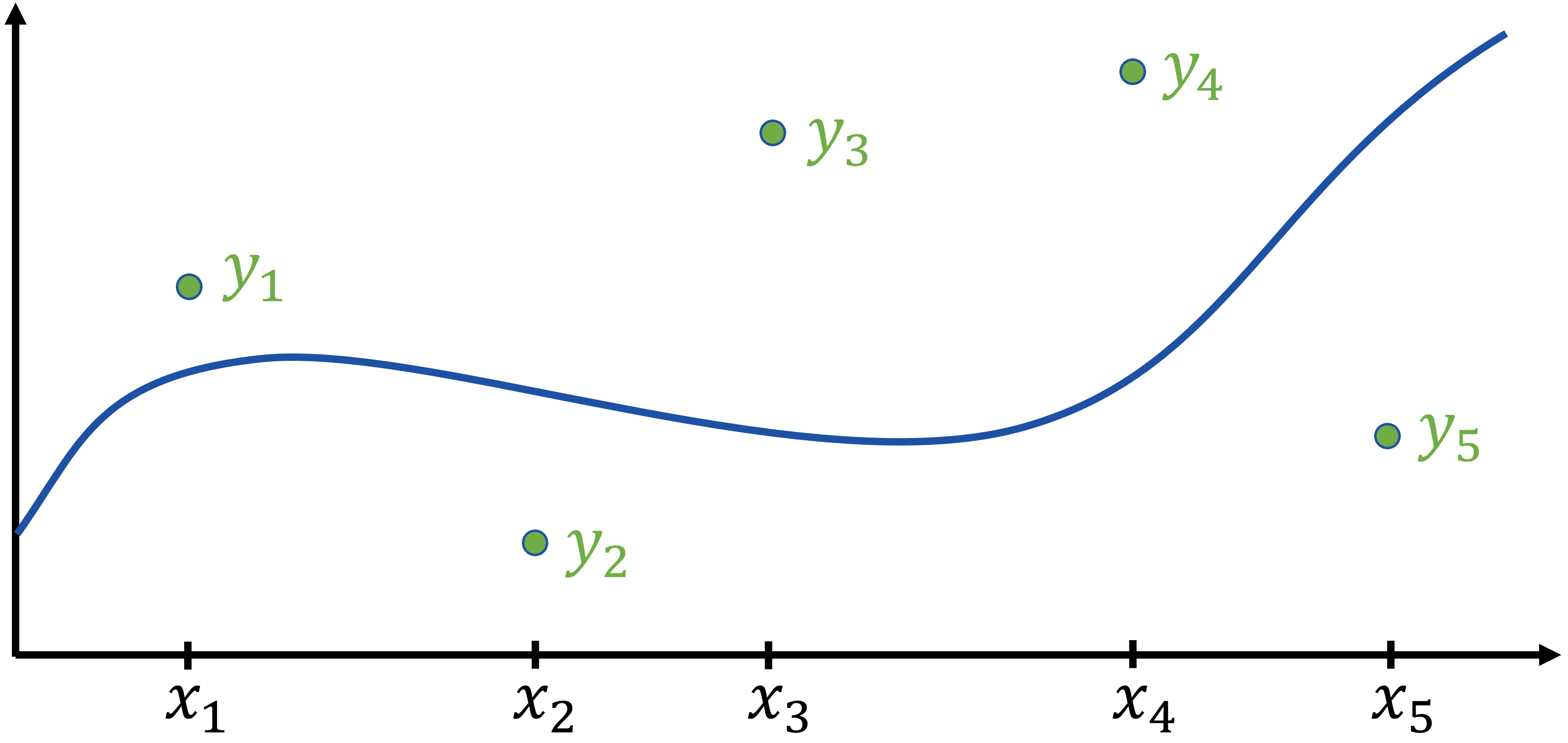
Отклонение функции от набора точек



Точки

(x_i, y_i)

Отклонение функции от набора точек



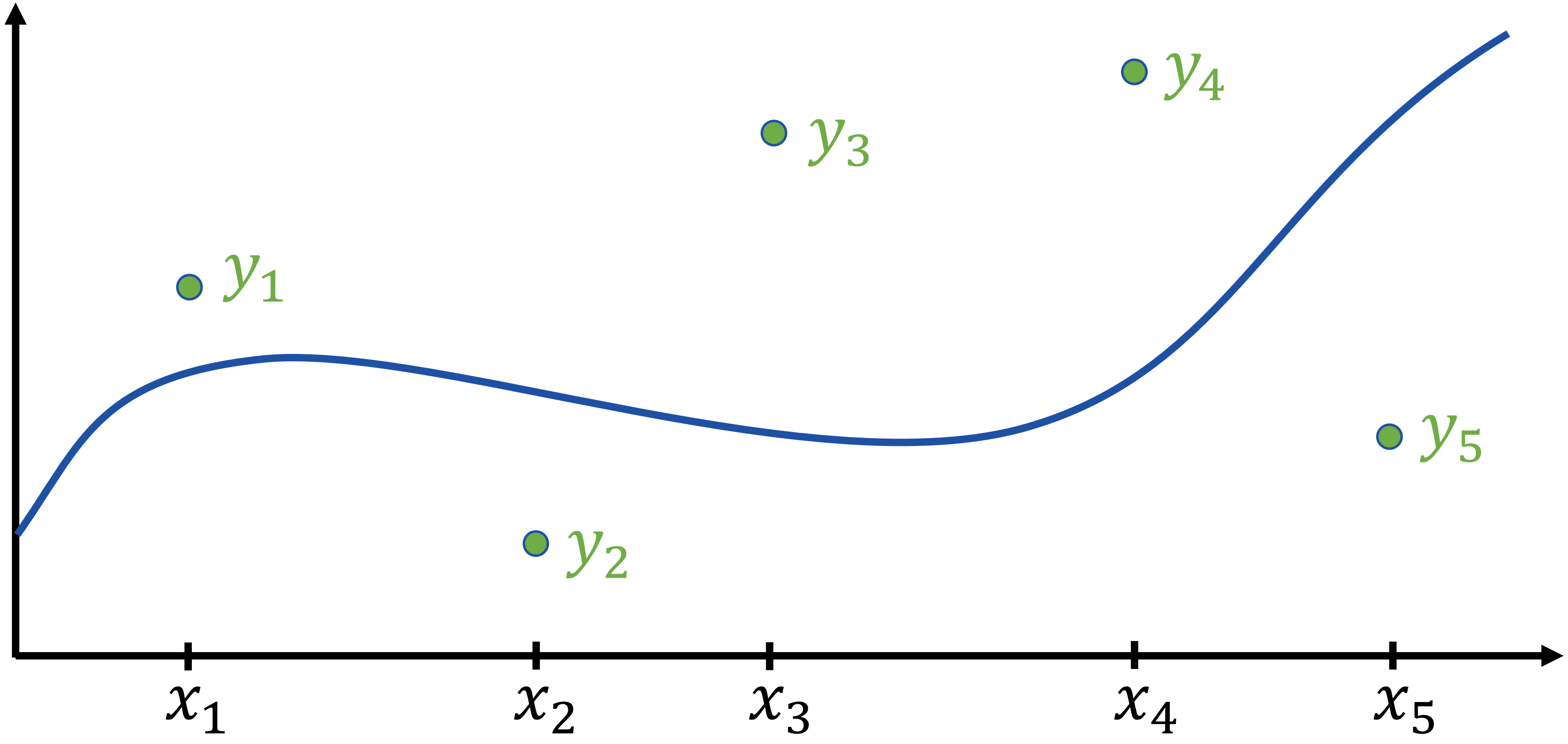
Точки

(x_i, y_i)

Функция

$\hat{y} = f(x)$

Отклонение функции от набора точек



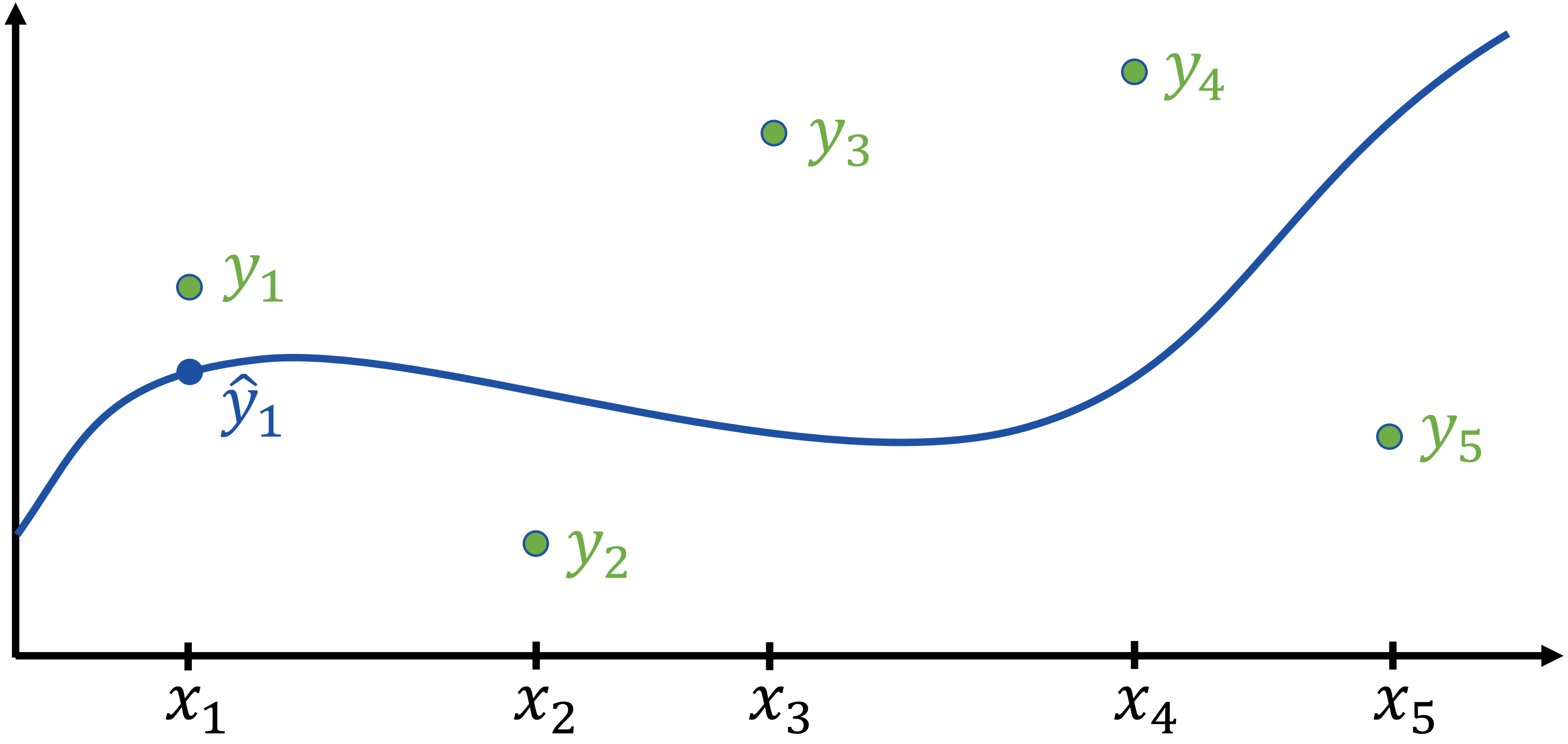
Точки

(x_i, y_i)

Значения функции

$\hat{y}_i = f(x_i)$

Отклонение функции от набора точек



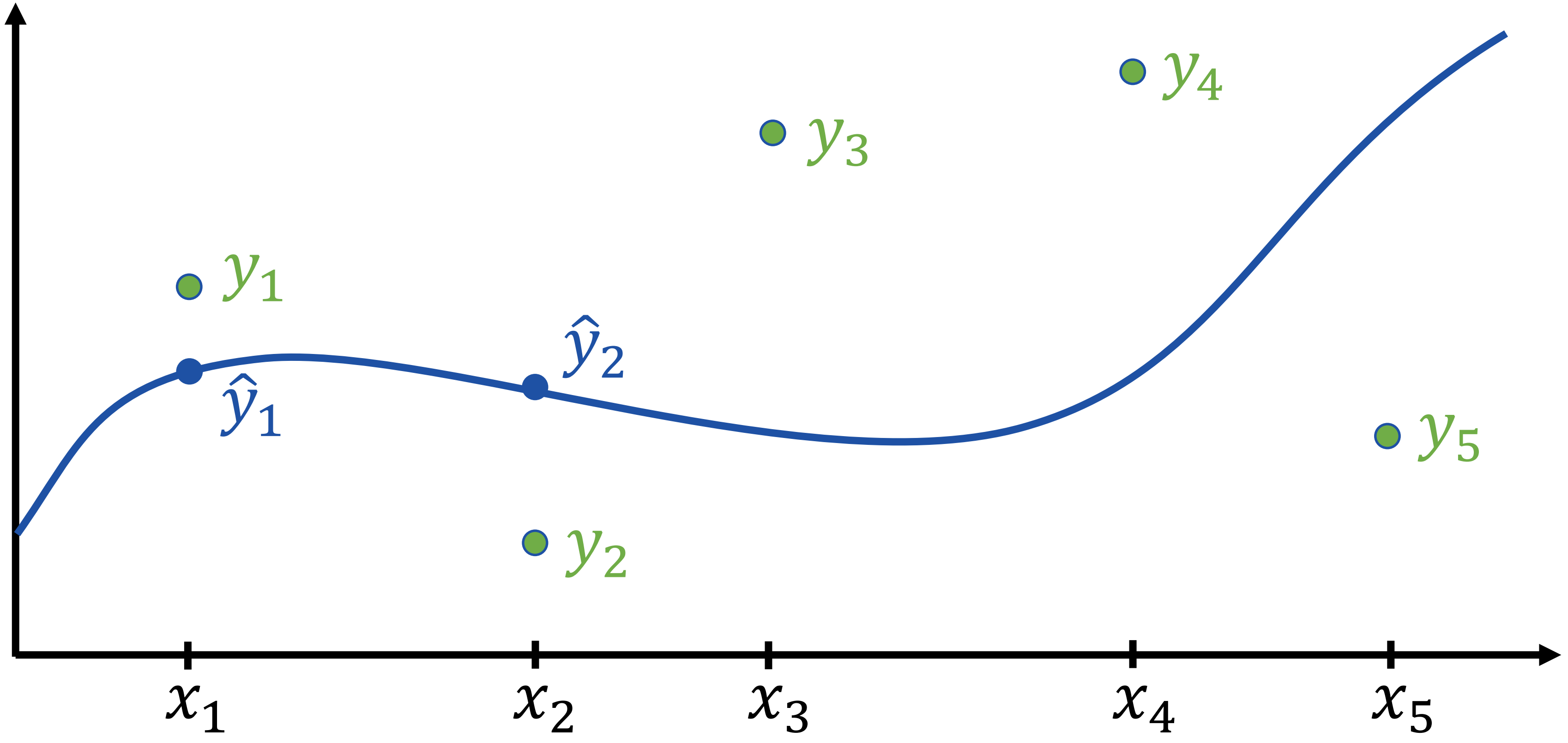
Точки

(x_i, y_i)

Значения функции

$\hat{y}_1 = f(x_1)$

Отклонение функции от набора точек



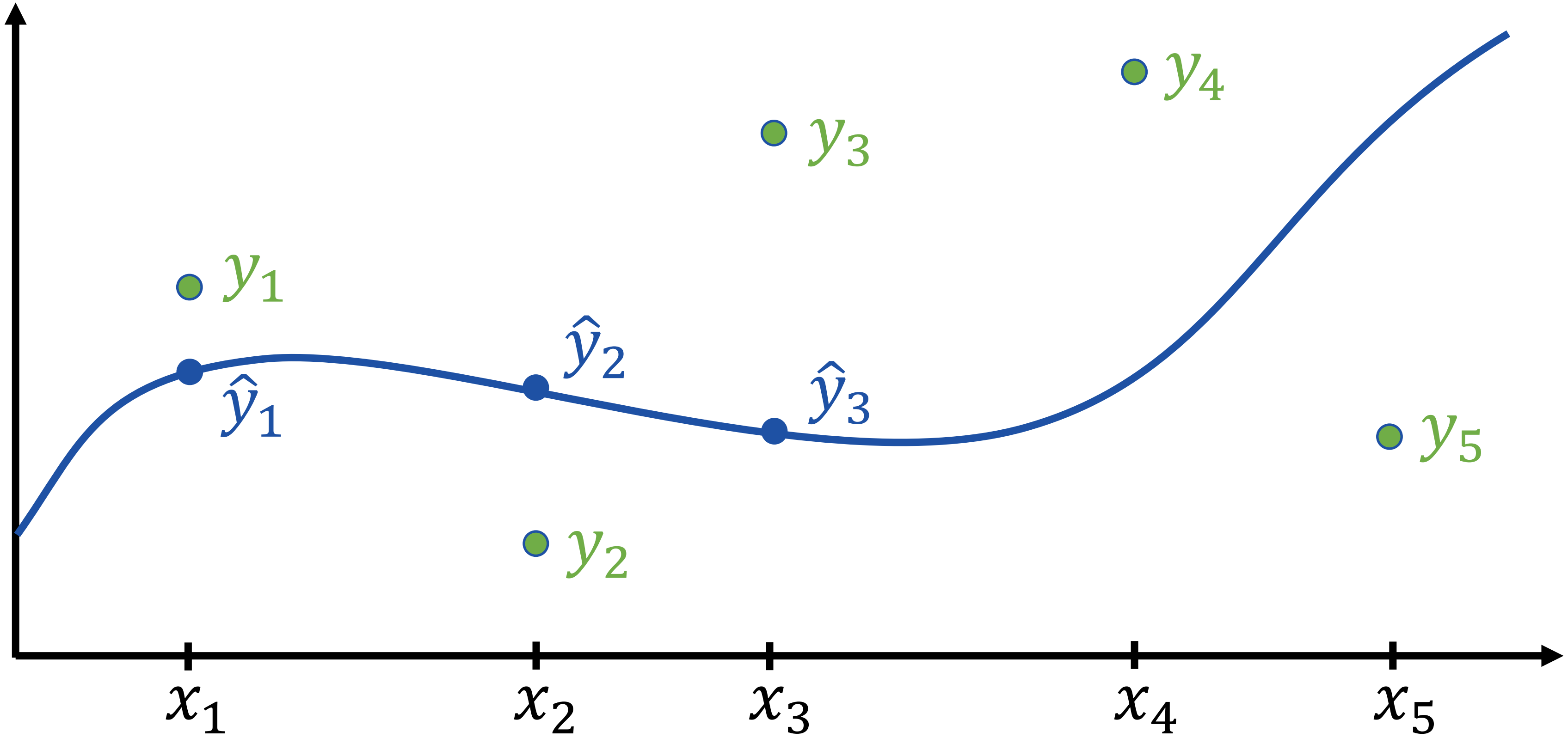
Точки

(x_i, y_i)

Значения функции

$\hat{y}_2 = f(x_2)$

Отклонение функции от набора точек



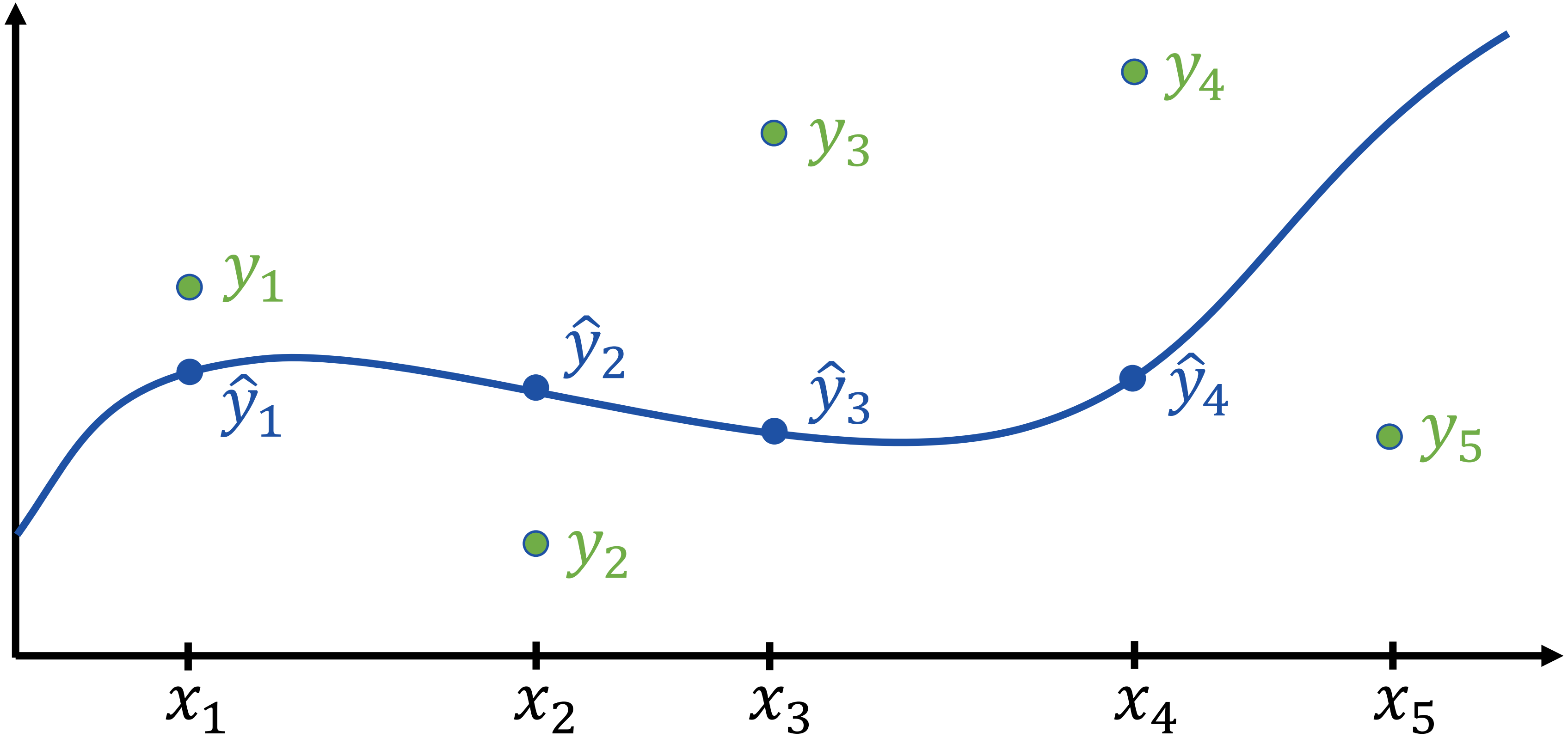
Точки

(x_i, y_i)

Значения функции

$\hat{y}_3 = f(x_3)$

Отклонение функции от набора точек



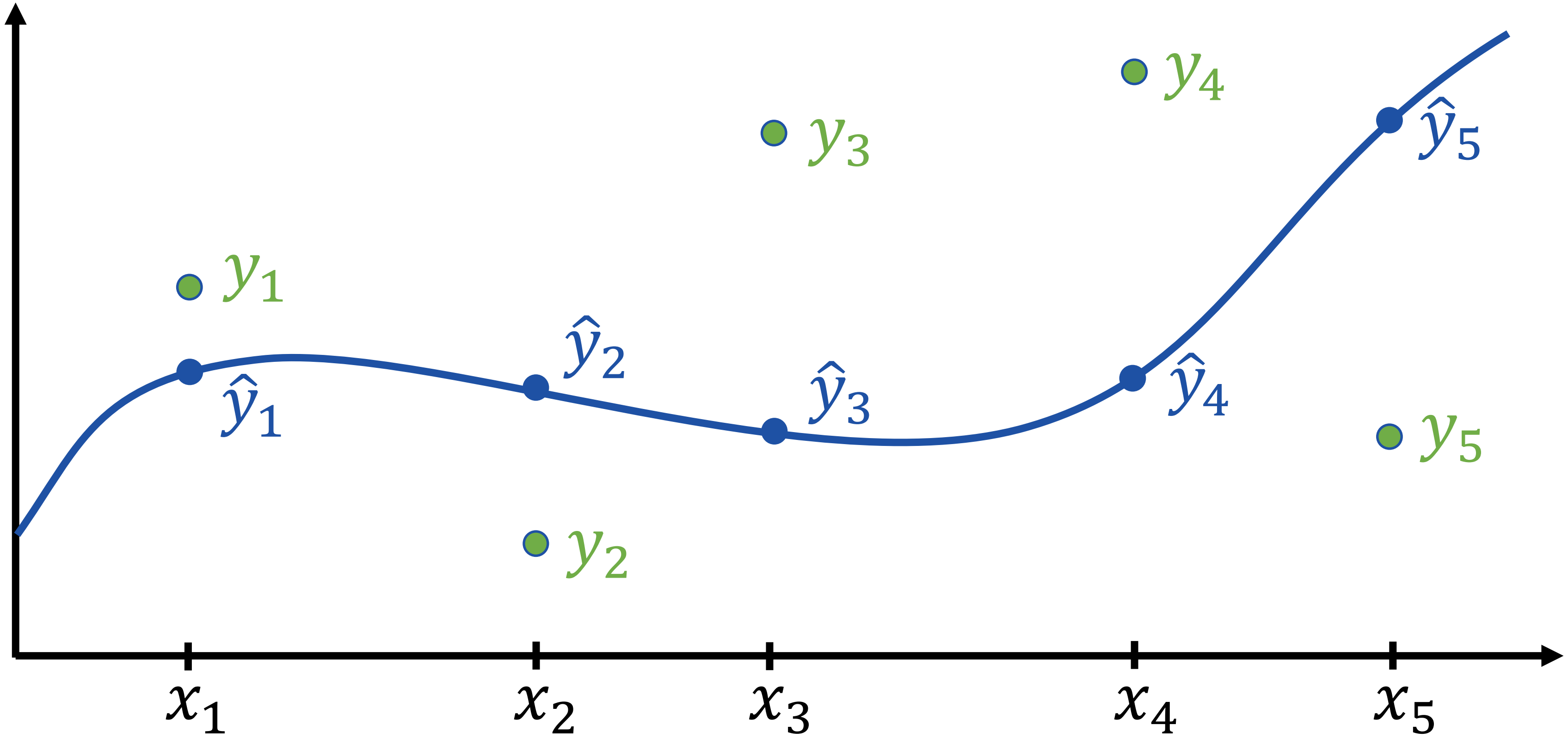
Точки

(x_i, y_i)

Значения функции

$\hat{y}_4 = f(x_4)$

Отклонение функции от набора точек



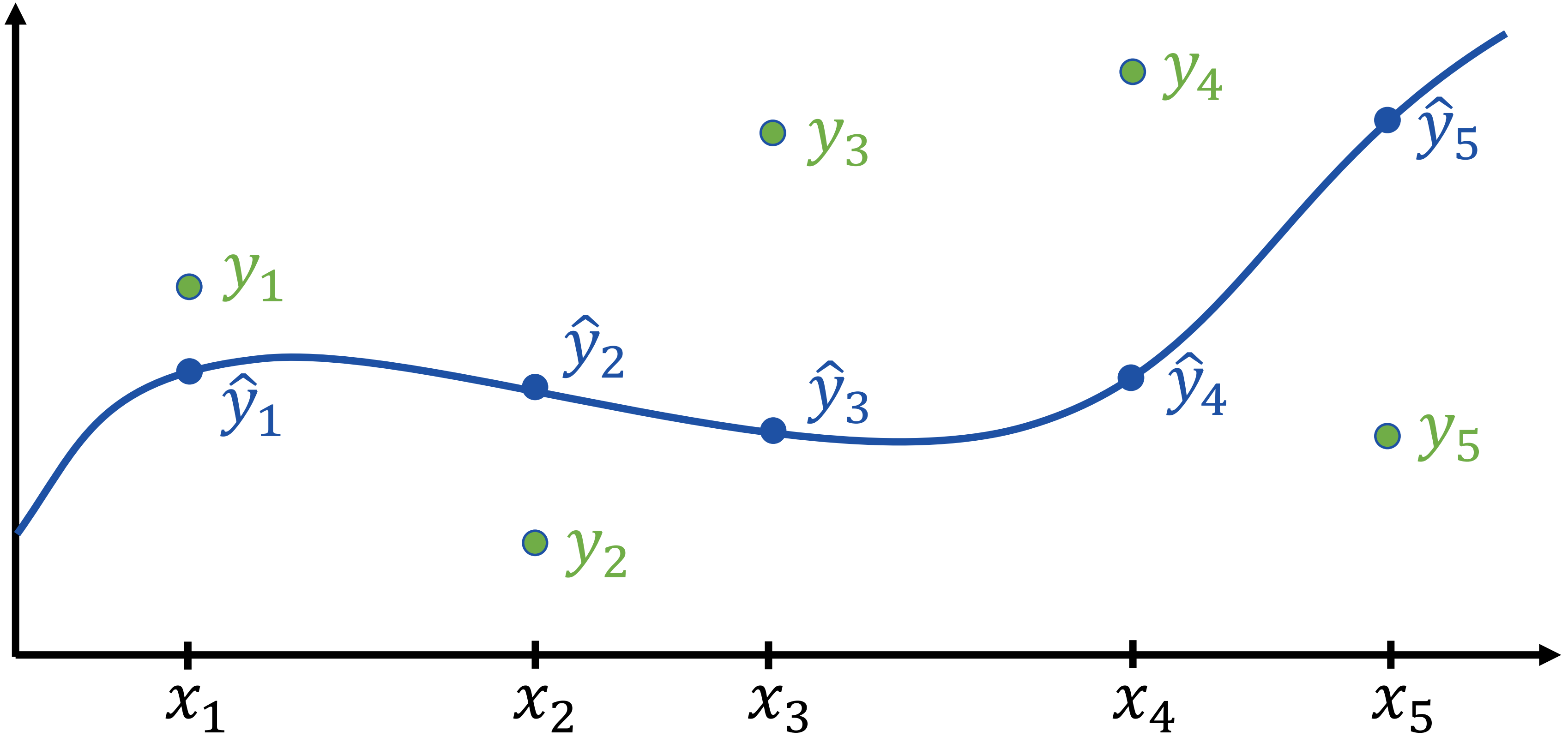
Точки

(x_i, y_i)

Значения функции

$\hat{y}_5 = f(x_5)$

Отклонение функции от набора точек



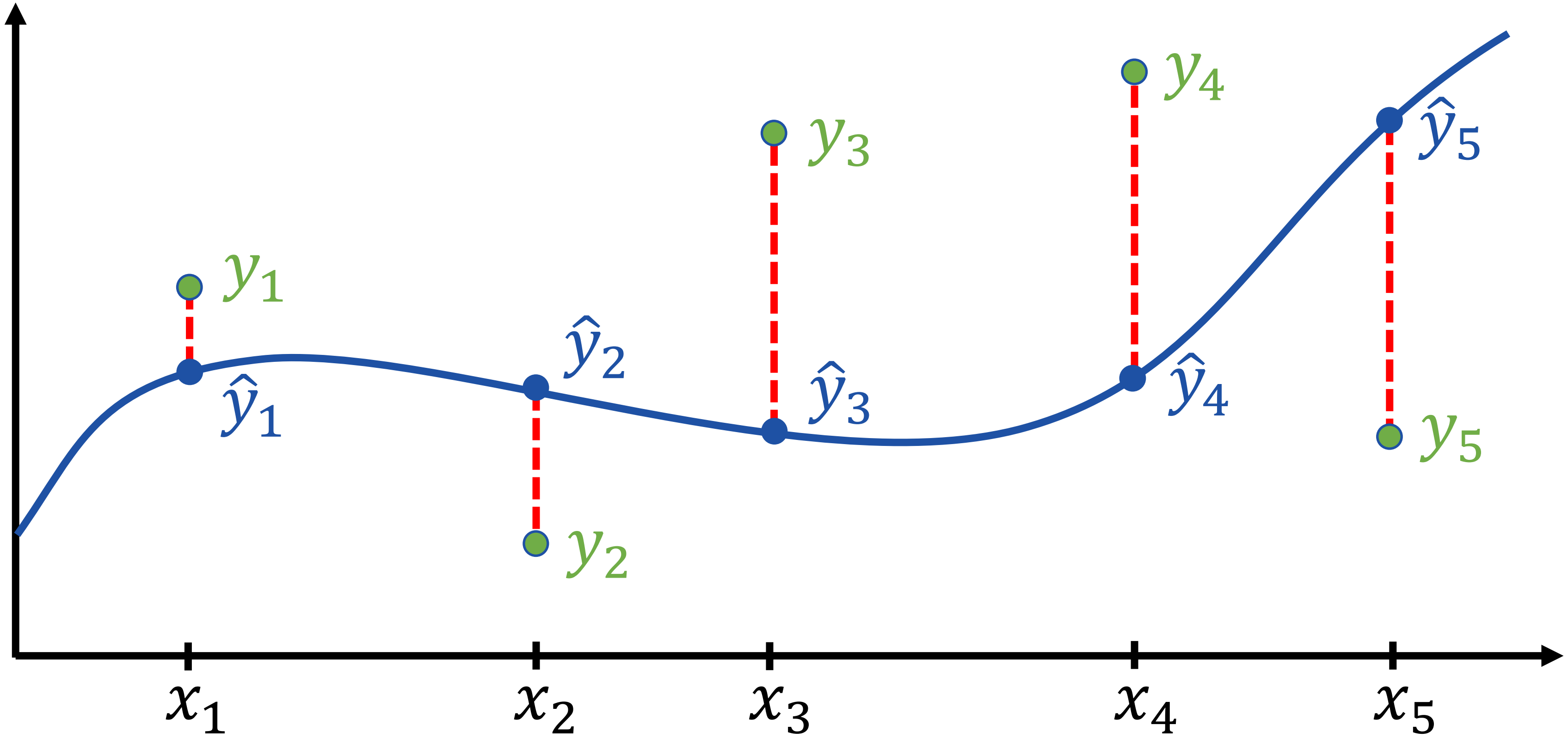
Точки

$$(x_i, y_i)$$

Значения функции

$$\hat{y}_i = f(x_i)$$

Отклонение функции от набора точек



Точки

(x_i, y_i)

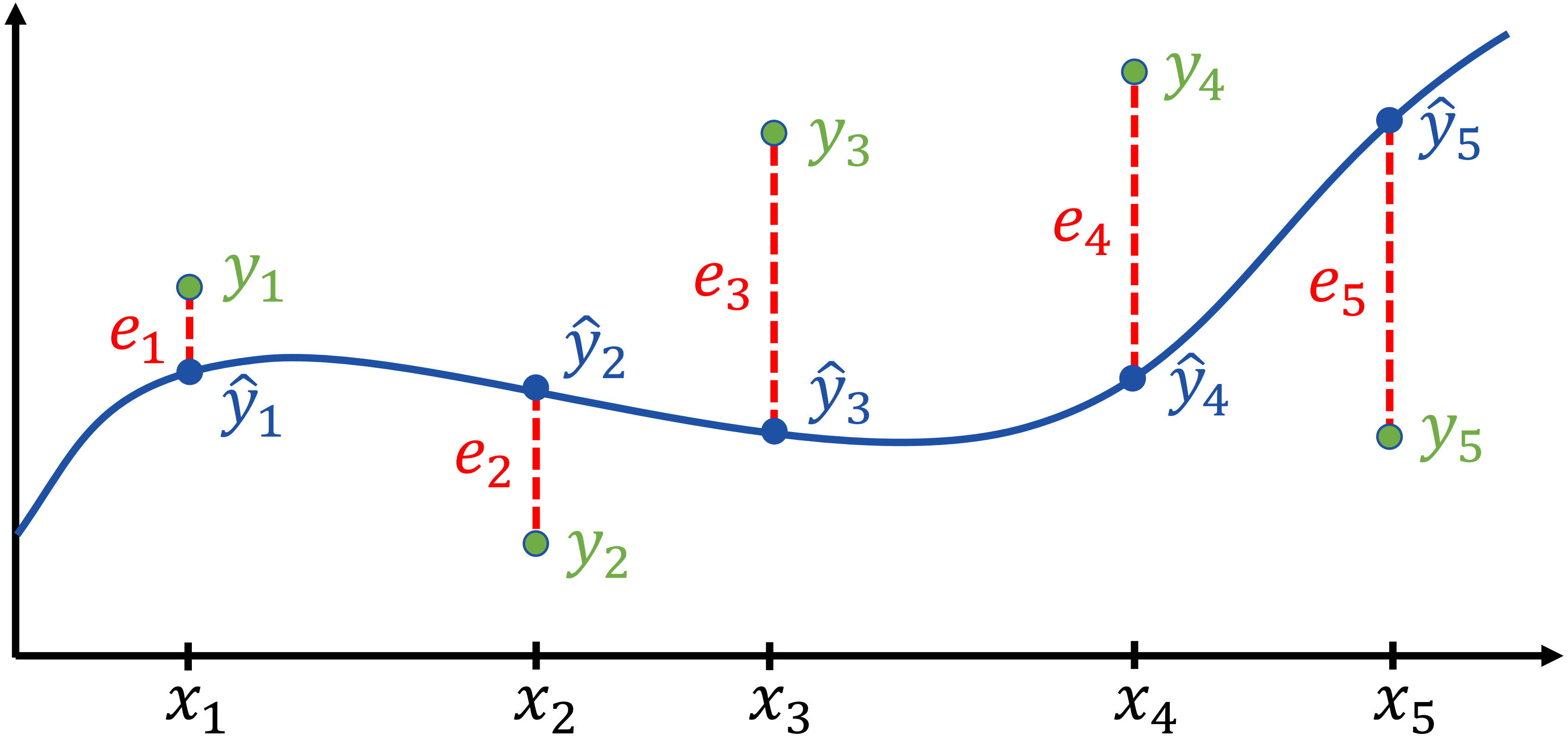
Значения функции

$\hat{y}_i = f(x_i)$

Отклонения

$y_i - \hat{y}_i$

Отклонение функции от набора точек



Точки

$$(x_i, y_i)$$

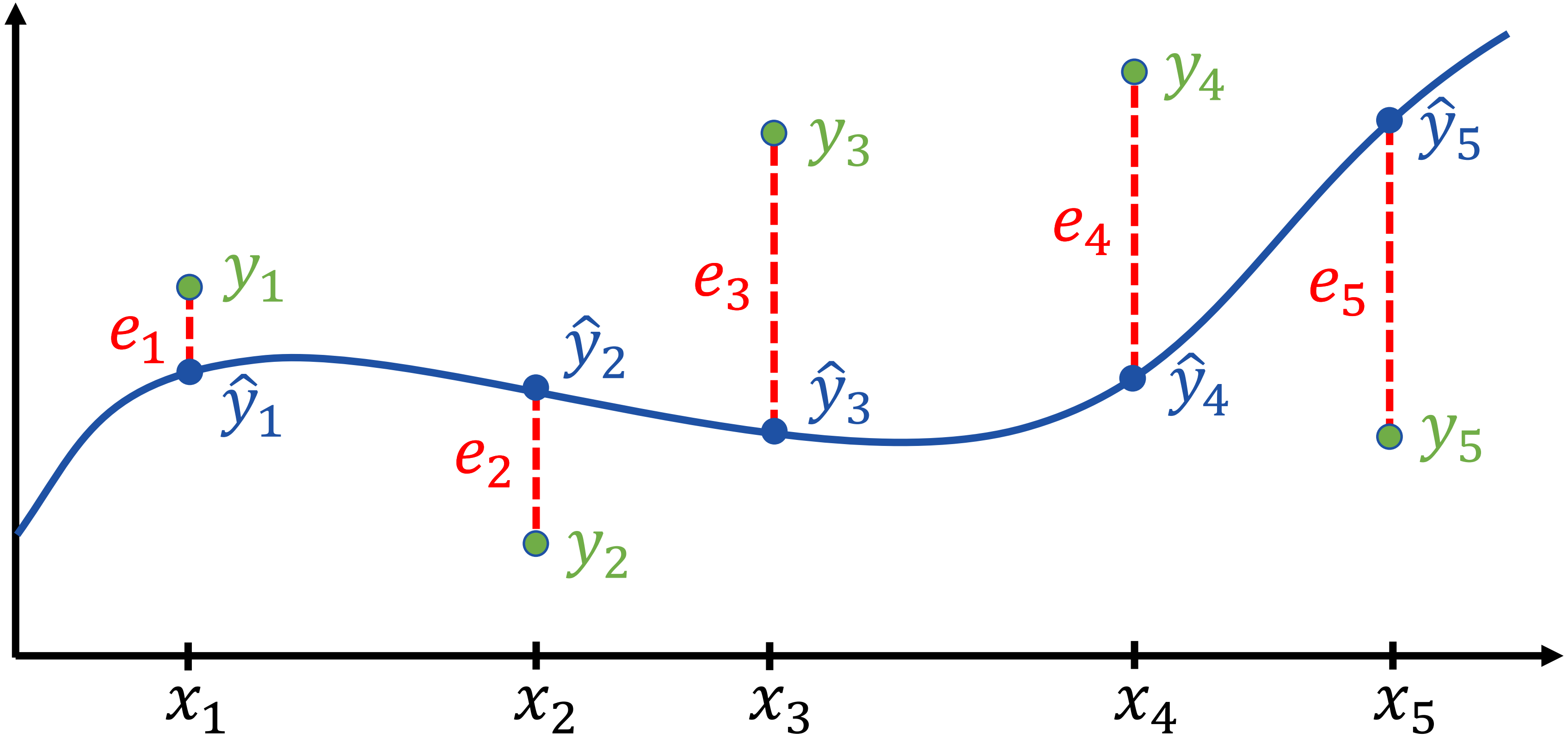
Значения функции

$$\hat{y}_i = f(x_i)$$

Отклонения

$$e_i = y_i - \hat{y}_i$$

Отклонение функции от набора точек



Точки

$$(x_i, y_i)$$

Значения функции

$$\hat{y}_i = f(x_i)$$

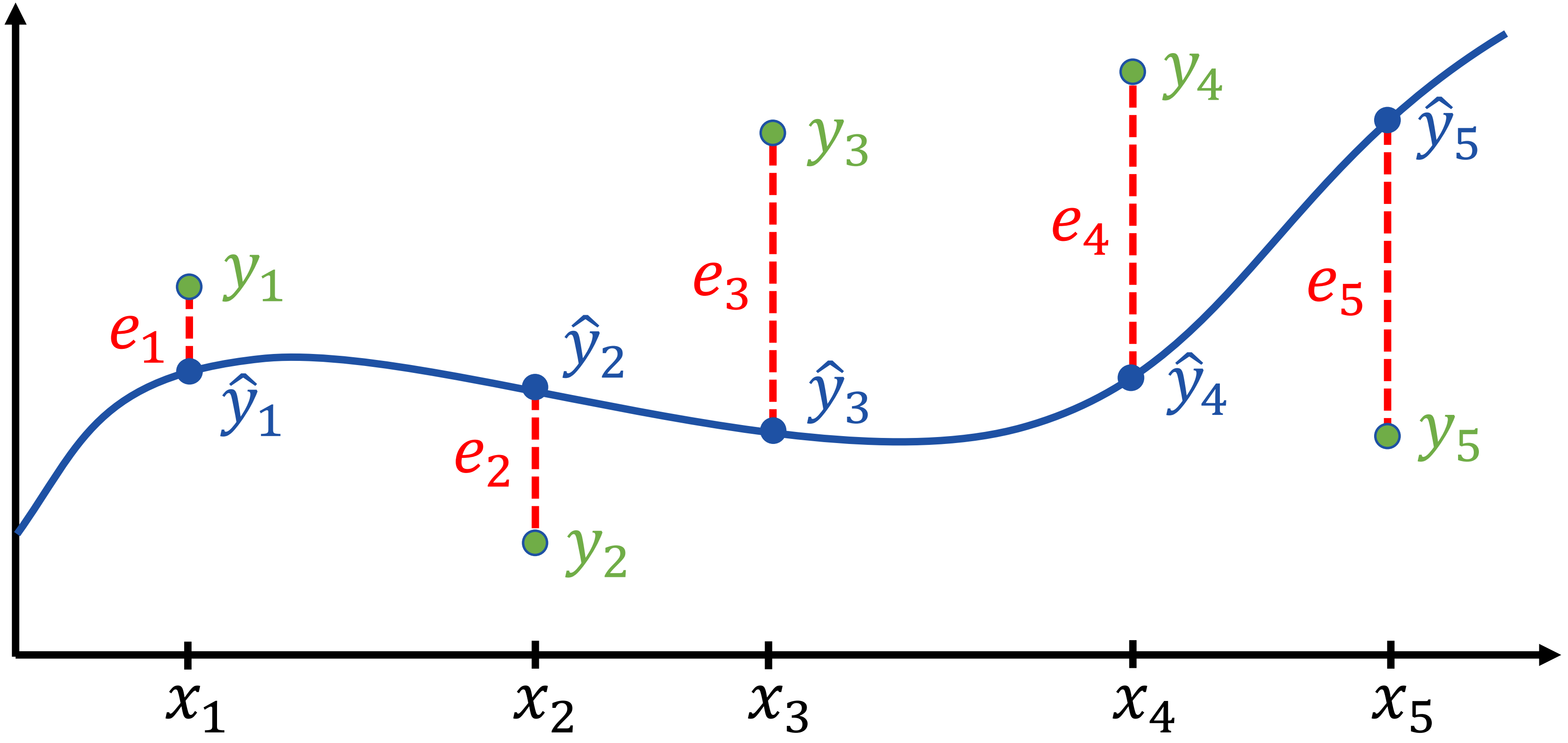
Отклонения

$$e_i = y_i - \hat{y}_i$$

Вектор отклонений

$$e = (\quad , \quad , \quad , \quad , \quad)$$

Отклонение функции от набора точек



Точки

(x_i, y_i)

Значения функции

$\hat{y}_i = f(x_i)$

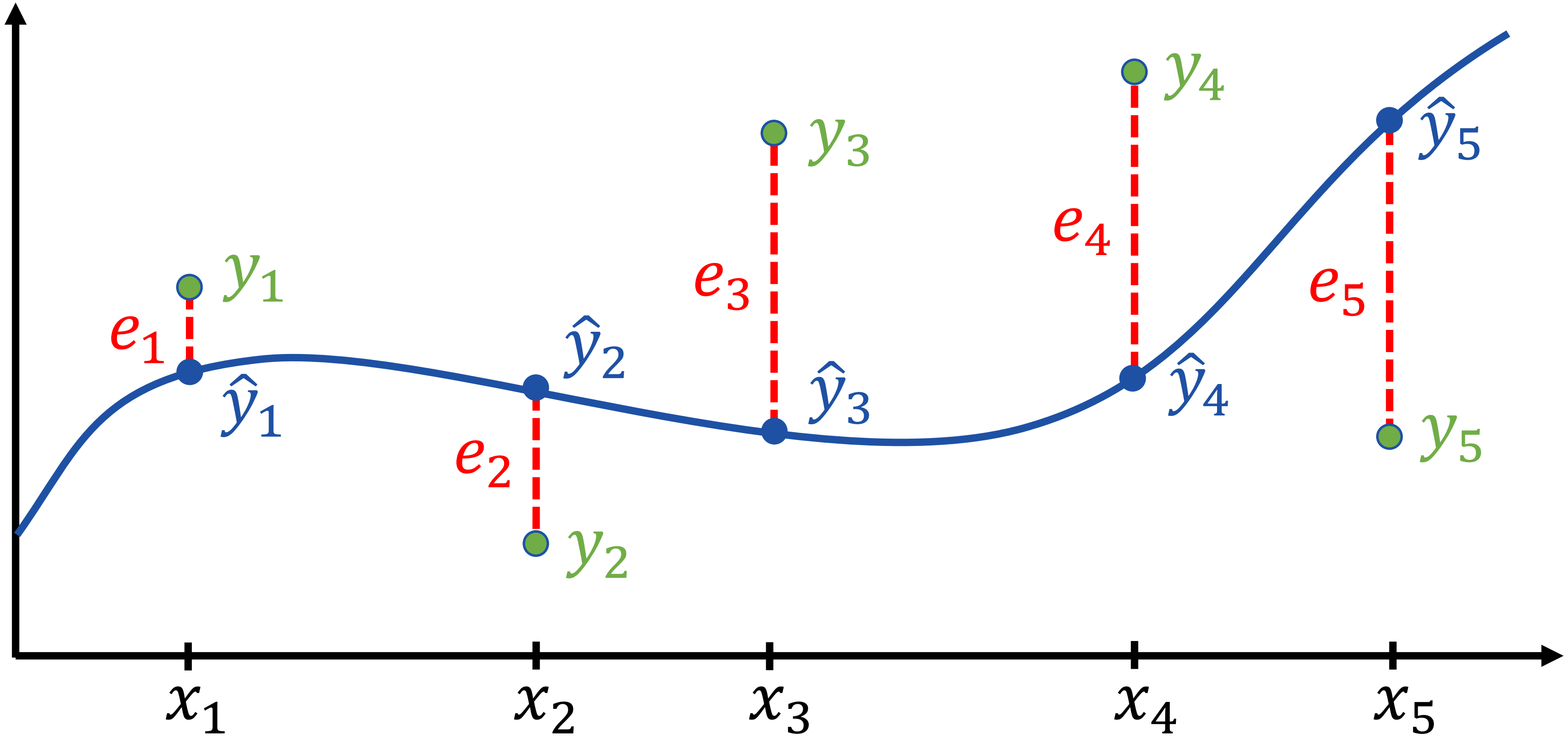
Отклонения

$e_i = y_i - \hat{y}_i$

Вектор отклонений

$e = (e_1, e_2, e_3, e_4, e_5)$

Отклонение функции от набора точек



Точки

$$(x_i, y_i)$$

Значения функции

$$\hat{y}_i = f(x_i)$$

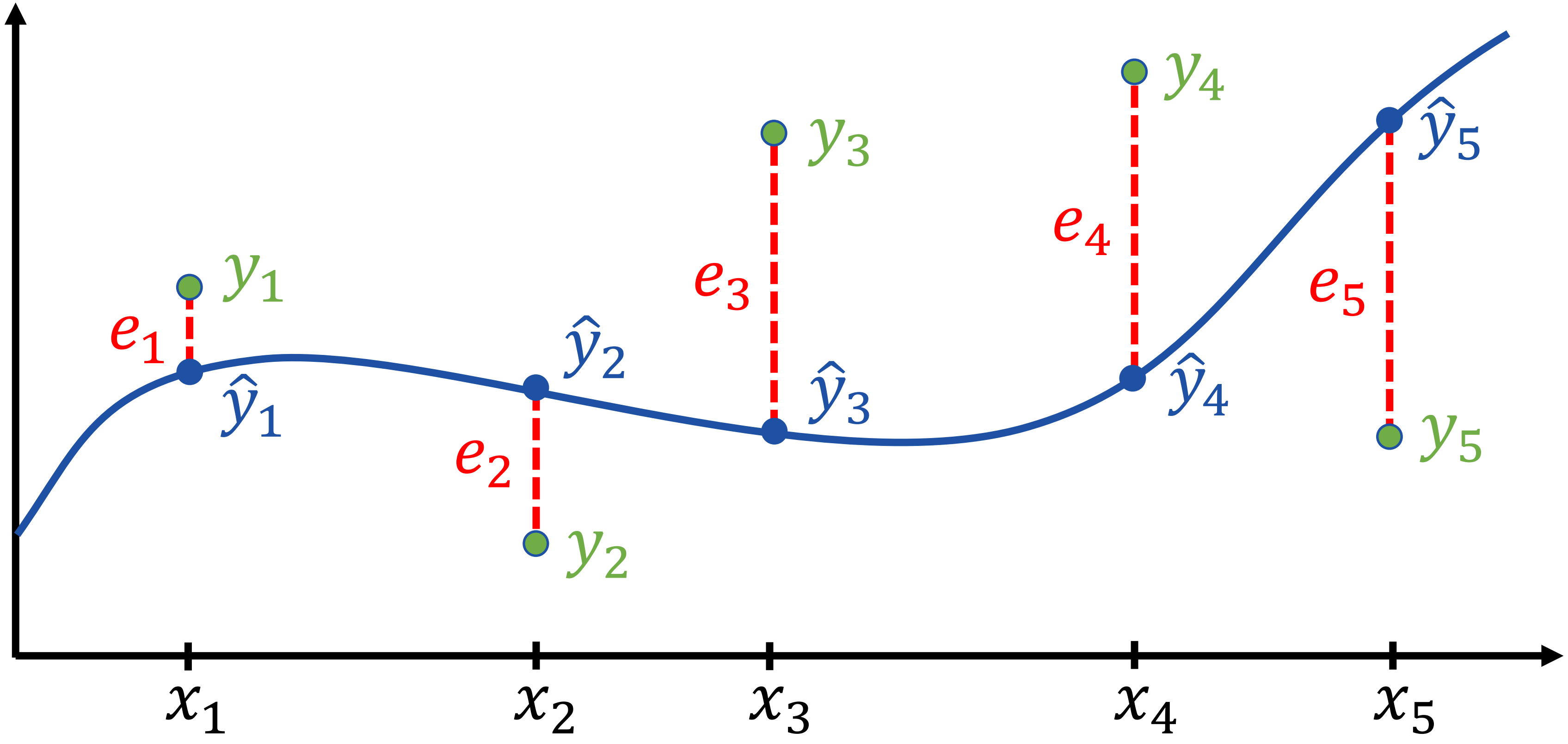
Отклонения

$$e_i = y_i - \hat{y}_i$$

Вектор отклонений

$$e = (e_1, e_2, e_3, e_4, e_5)$$

Отклонение функции от набора точек



Точки

(x_i, y_i)

Значения функции

$\hat{y}_i = f(x_i)$

Отклонения

$e_i = y_i - \hat{y}_i$

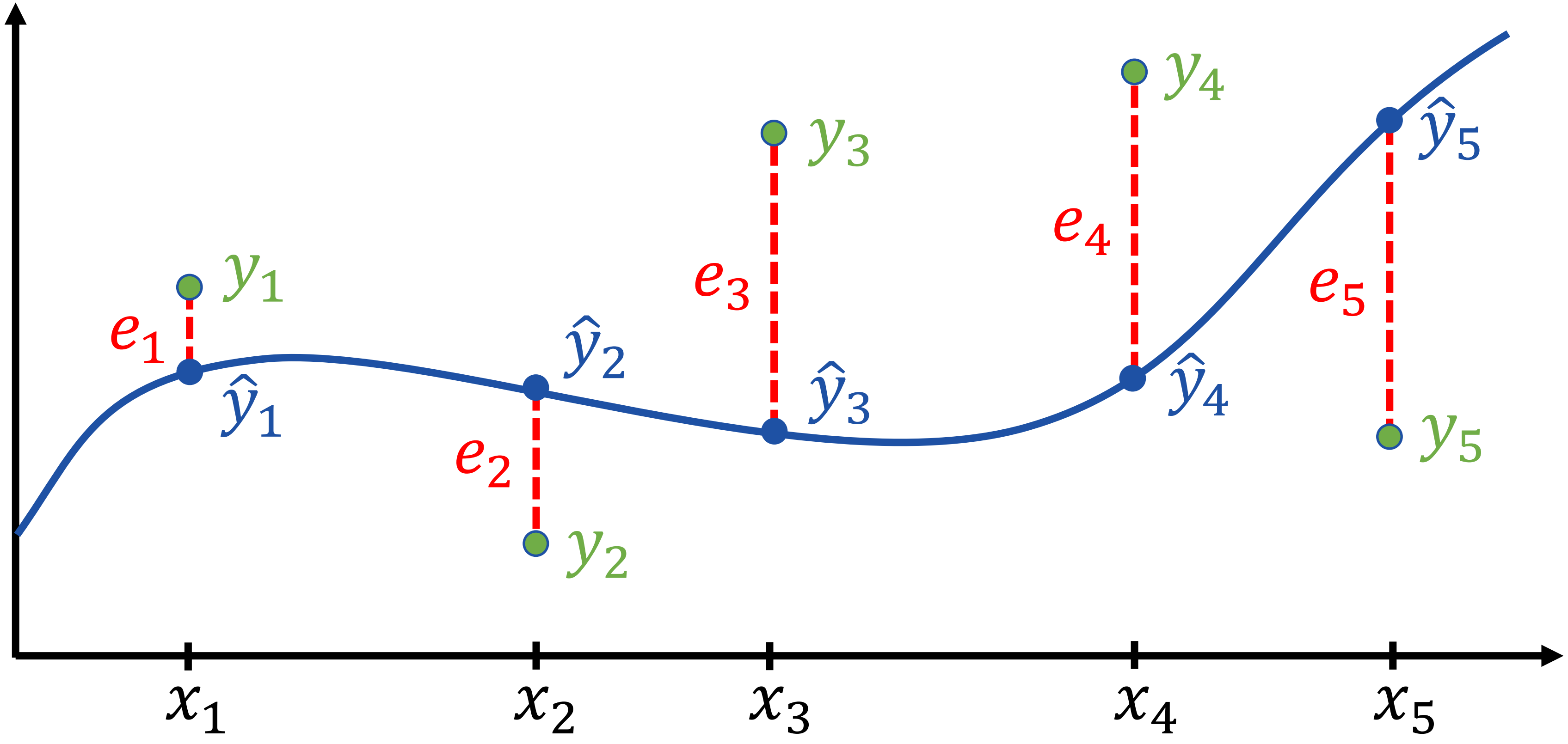
Вектор отклонений

$e = (e_1, e_2, e_3, e_4, e_5)$

Его норма

$\|e\|$

Отклонение функции от набора точек



Точки

$$(x_i, y_i)$$

Значения функции

$$\hat{y}_i = f(x_i)$$

Отклонения

$$e_i = y_i - \hat{y}_i$$

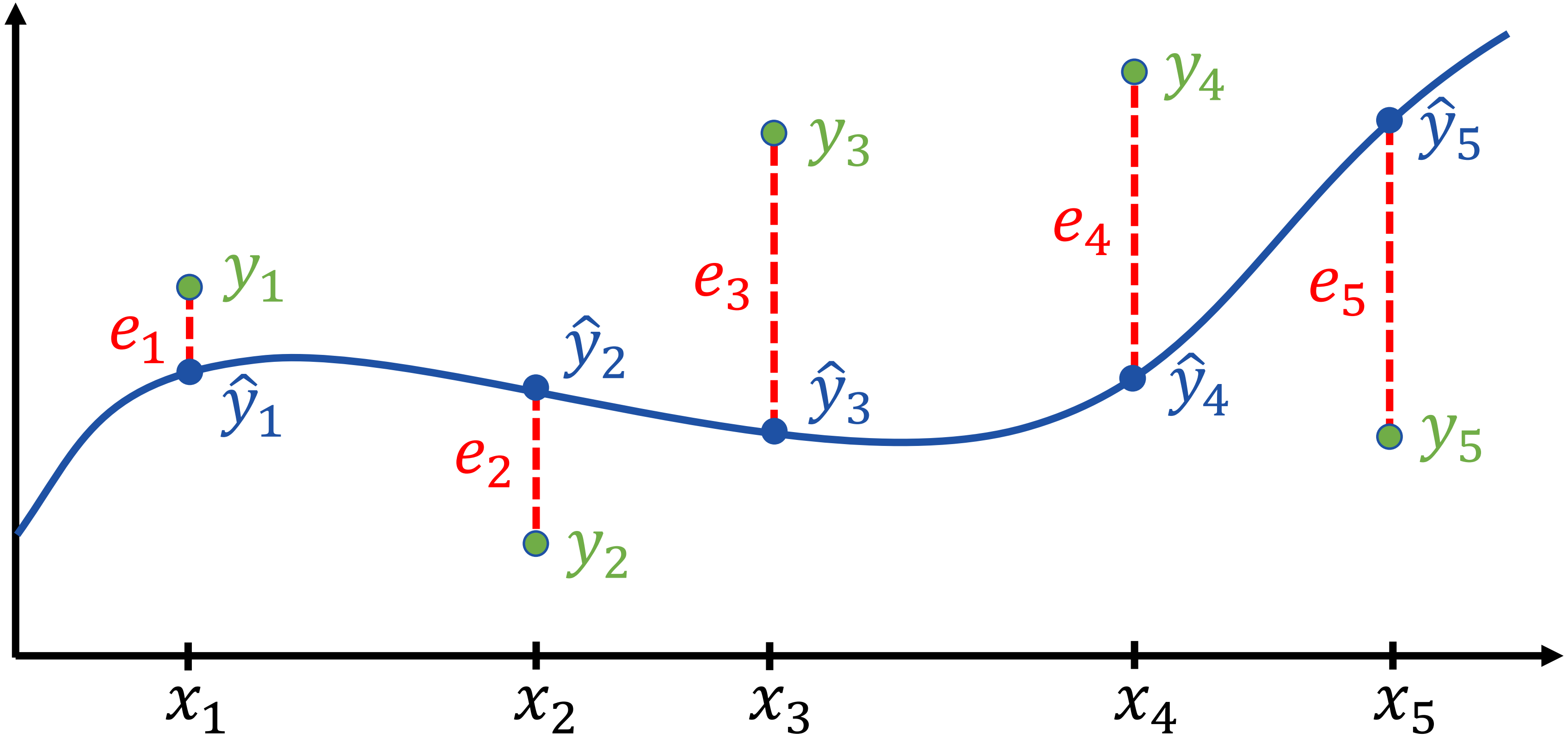
Вектор отклонений

$$e = (e_1, e_2, e_3, e_4, e_5)$$

Его норма (например, такая)

$$\|e\|_1 = |e_1| + |e_2| + |e_3| + |e_4| + |e_5|$$

Отклонение функции от набора точек



Точки

(x_i, y_i)

Значения функции

$\hat{y}_i = f(x_i)$

Отклонения

$e_i = y_i - \hat{y}_i$

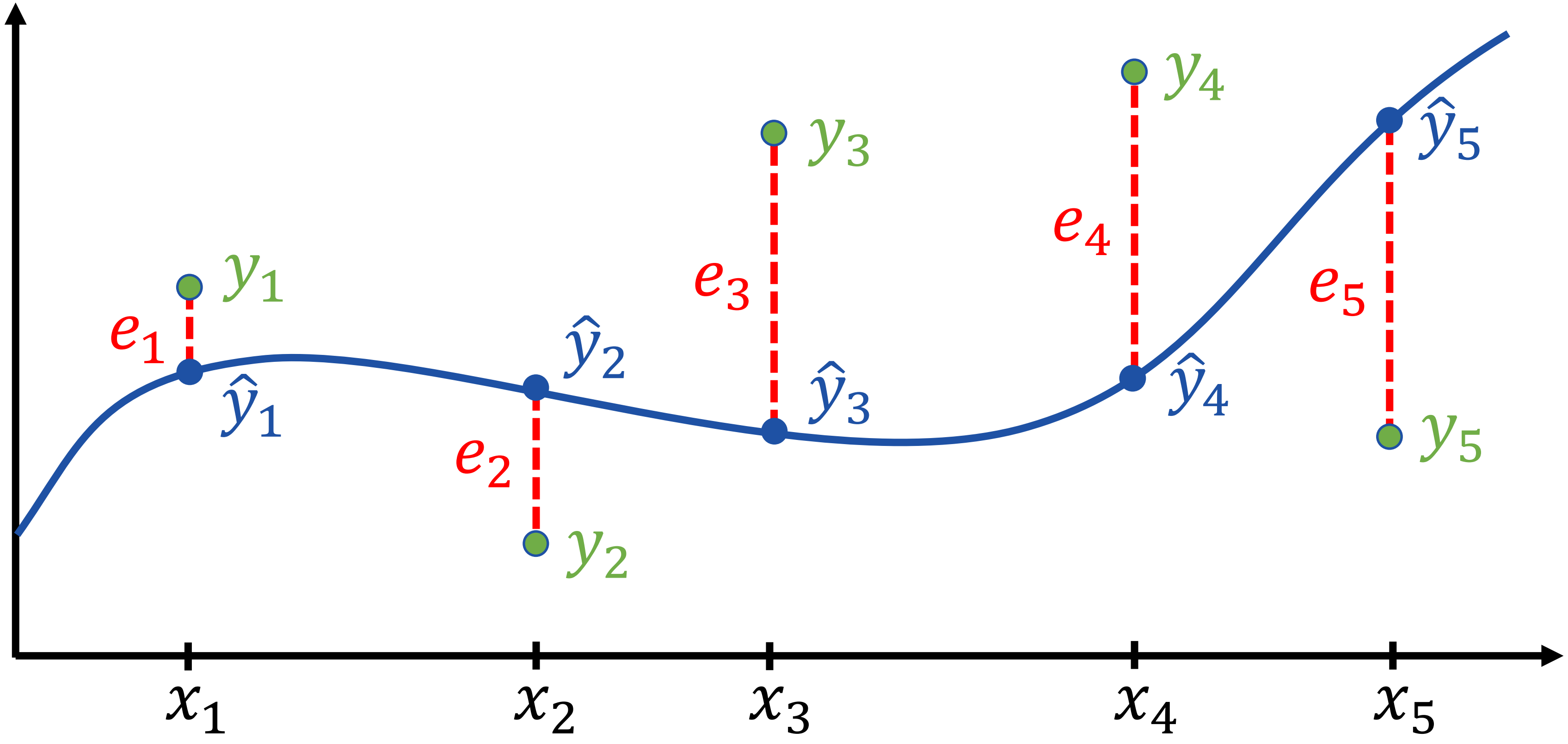
Вектор отклонений

$e = (e_1, e_2, e_3, e_4, e_5)$

Его норма (или такая)

$$\|e\|_2 = \sqrt{e_1^2 + e_2^2 + e_3^2 + e_4^2 + e_5^2}$$

Отклонение функции от набора точек



Точки

(x_i, y_i)

Значения функции

$\hat{y}_i = f(x_i)$

Отклонения

$e_i = y_i - \hat{y}_i$

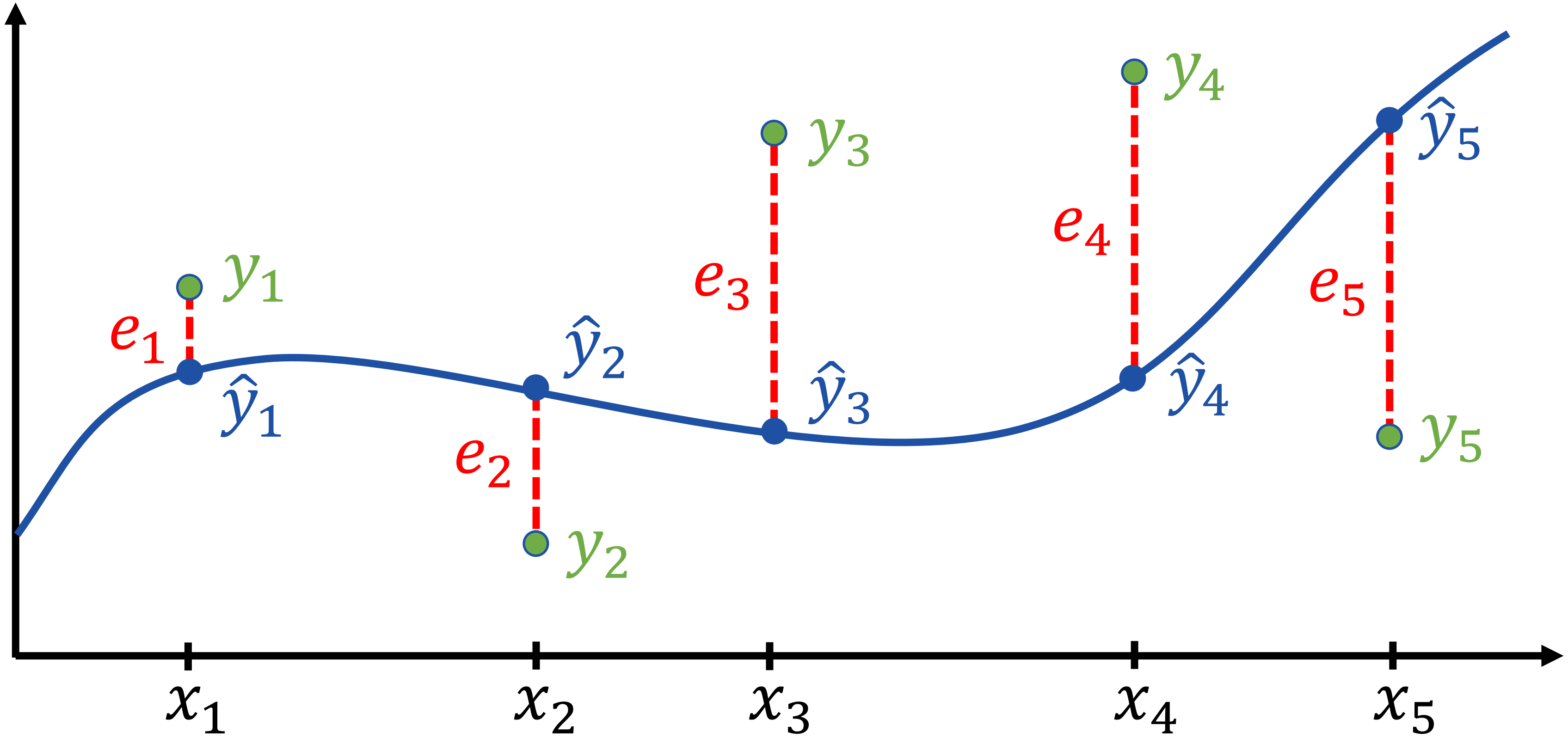
Вектор отклонений

$e = (e_1, e_2, e_3, e_4, e_5)$

Его норма (или даже такая)

$\|e\|_\infty = \max\{|e_1|, |e_2|, |e_3|, |e_4|, |e_5|\}$

Отклонение функции от набора точек



Точки

(x_i, y_i)

Значения функции

$\hat{y}_i = f(x_i)$

Отклонения

$e_i = y_i - \hat{y}_i$

Вектор отклонений

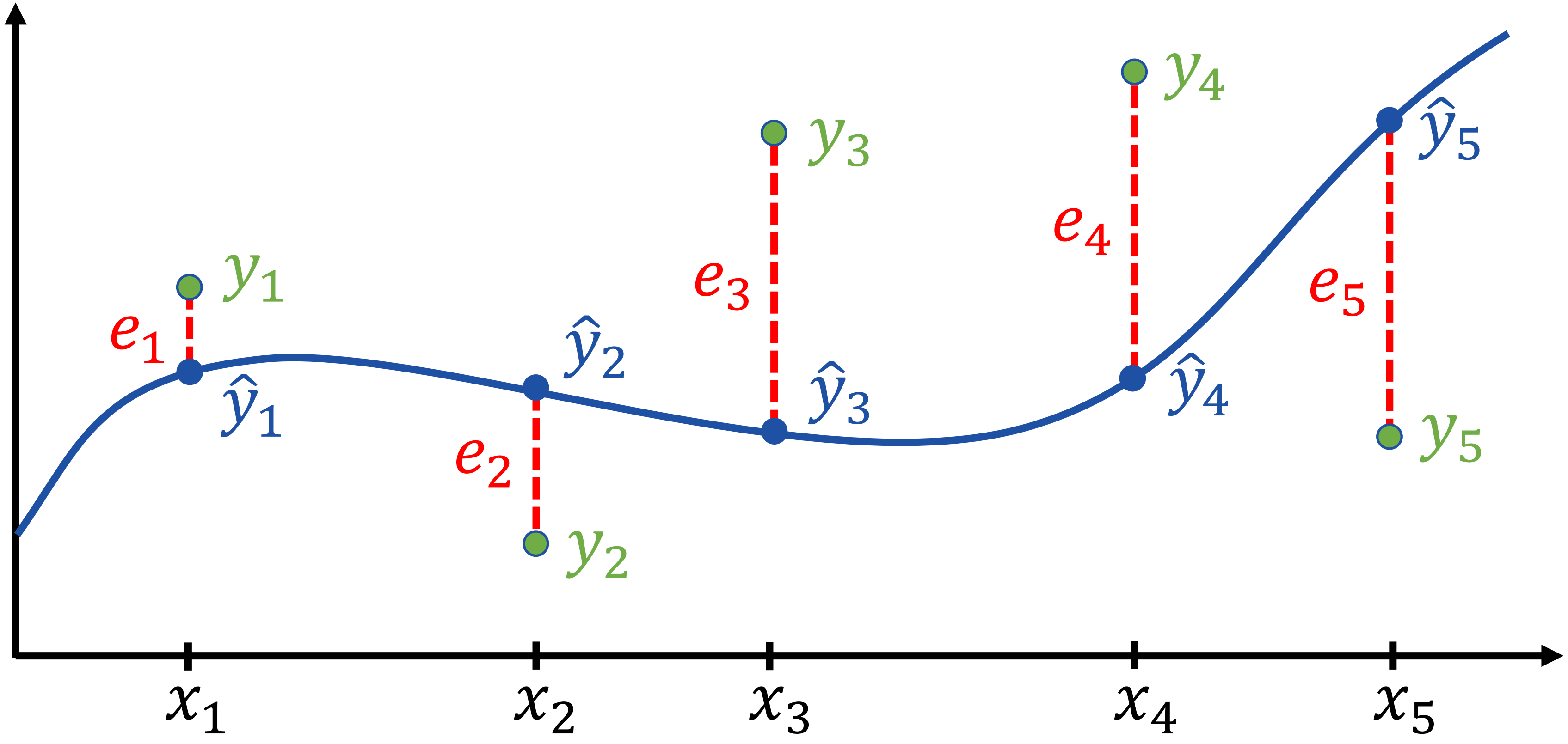
$e = (e_1, e_2, e_3, e_4, e_5)$

Его норма

$\|e\|$

(в общем, какая-нибудь)

Отклонение функции от набора точек



Точки

$$(x_i, y_i)$$

Значения функции

$$\hat{y}_i = f(x_i)$$

Отклонения

$$e_i = y_i - \hat{y}_i$$

Вектор отклонений

$$e = (e_1, e_2, e_3, e_4, e_5)$$

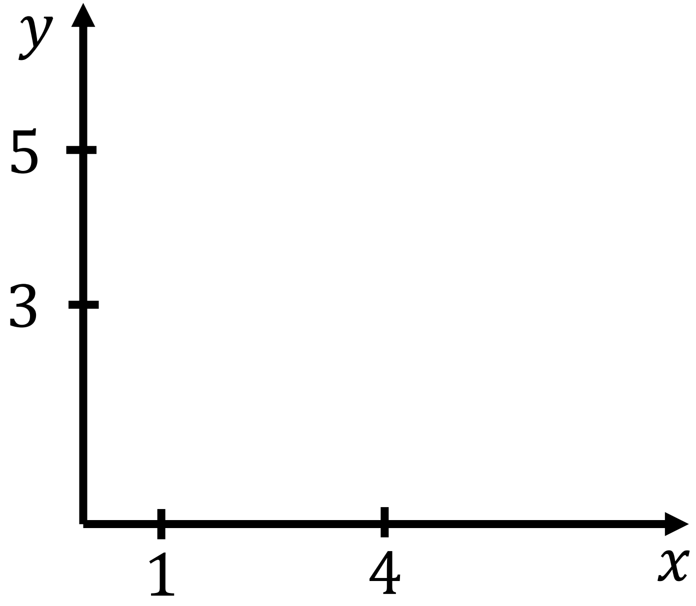
Его норма

$$\|e\|$$

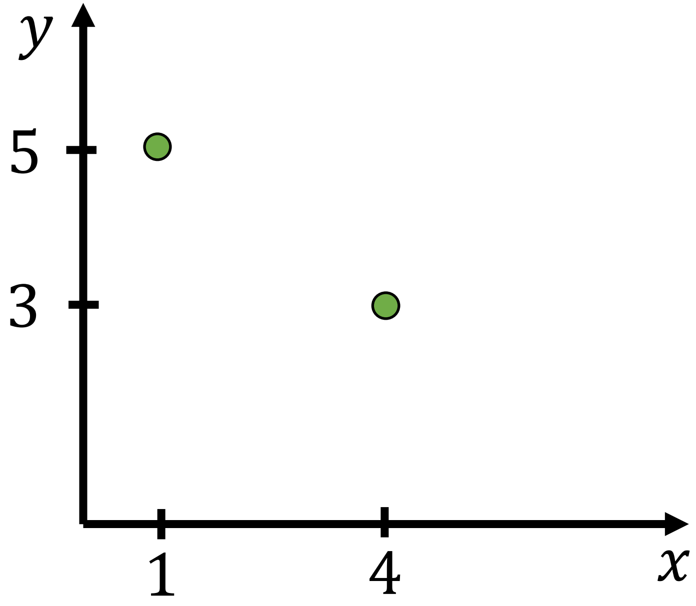
Число, характеризующее «промах» функции

Простой пример аппроксимации

Простой пример аппроксимации



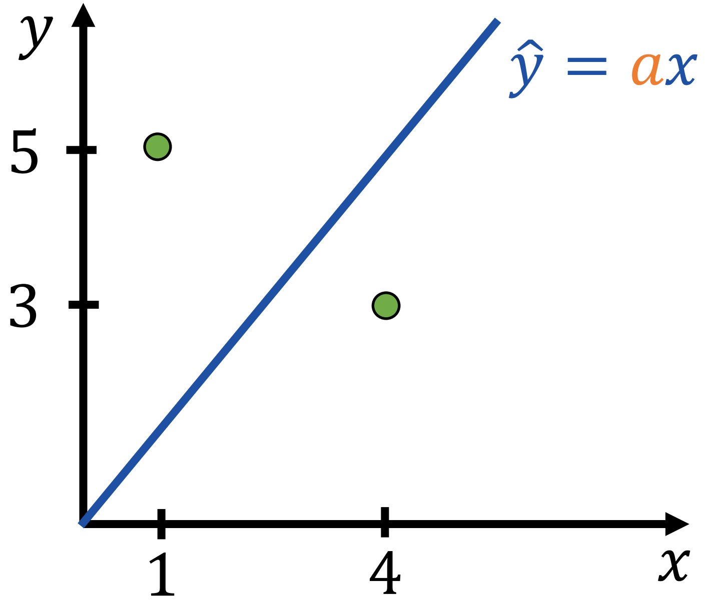
Простой пример аппроксимации



Точки

$(1, 5)$ и $(4, 3)$

Простой пример аппроксимации



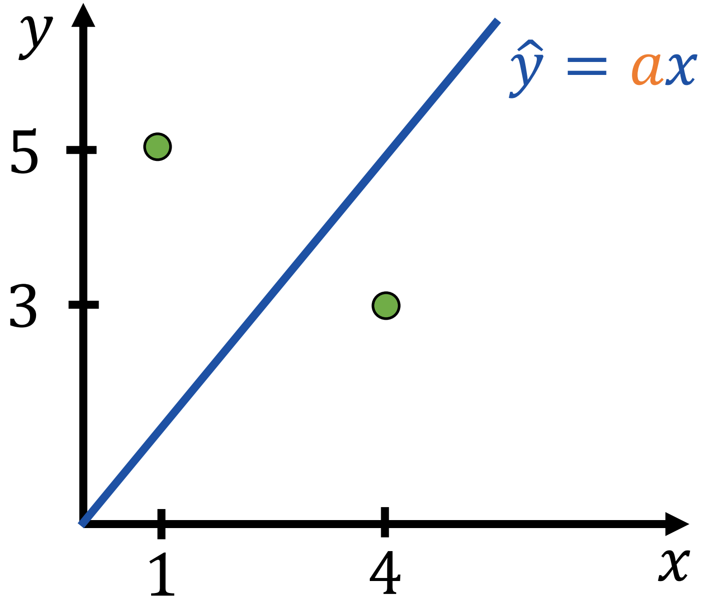
Точки

$(1, 5)$ и $(4, 3)$

Общий вид функции

$\hat{y} = ax$

Простой пример аппроксимации



Точки

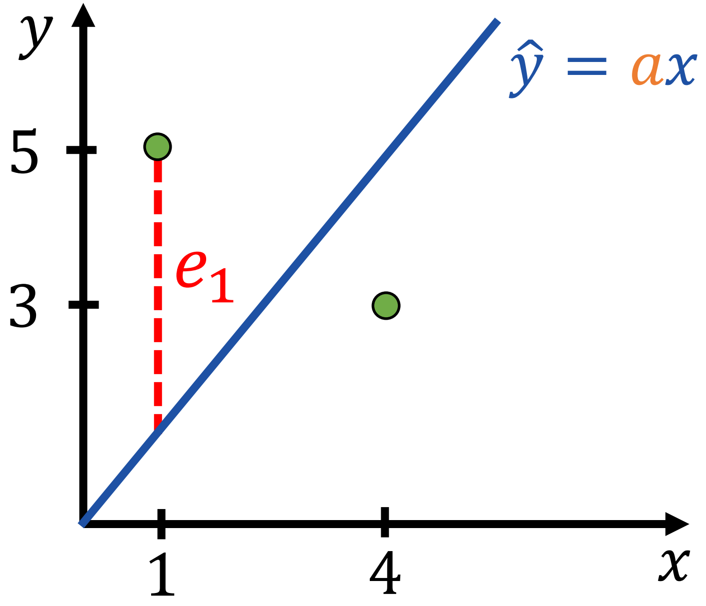
$(1, 5)$ и $(4, 3)$

Общий вид функции

$\hat{y} = ax$

Требуется найти наилучший параметр a

Простой пример аппроксимации



Точки

$(1, 5)$ и $(4, 3)$

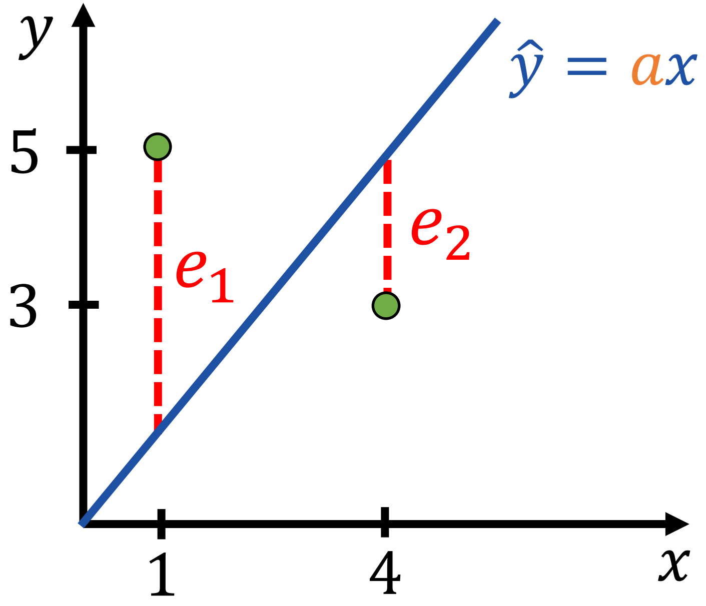
Общий вид функции

$\hat{y} = ax$

Требуется найти наилучший параметр a

$$e_1 = y_1 - ax_1 = 5 - 1a$$

Простой пример аппроксимации



Точки

$(1, 5)$ и $(4, 3)$

Общий вид функции

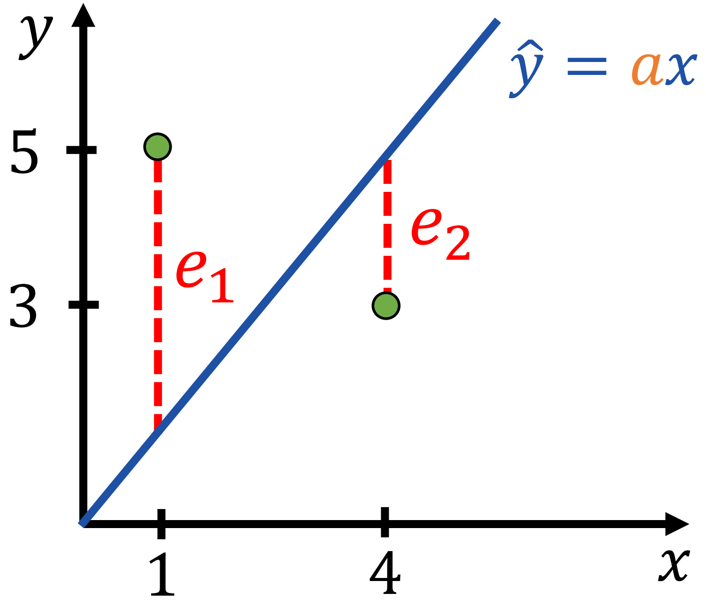
$$\hat{y} = ax$$

Требуется найти наилучший параметр a

$$e_1 = y_1 - ax_1 = 5 - 1a$$

$$e_2 = y_2 - ax_2 = 3 - 4a$$

Простой пример аппроксимации



Точки

$(1, 5)$ и $(4, 3)$

Общий вид функции

$\hat{y} = ax$

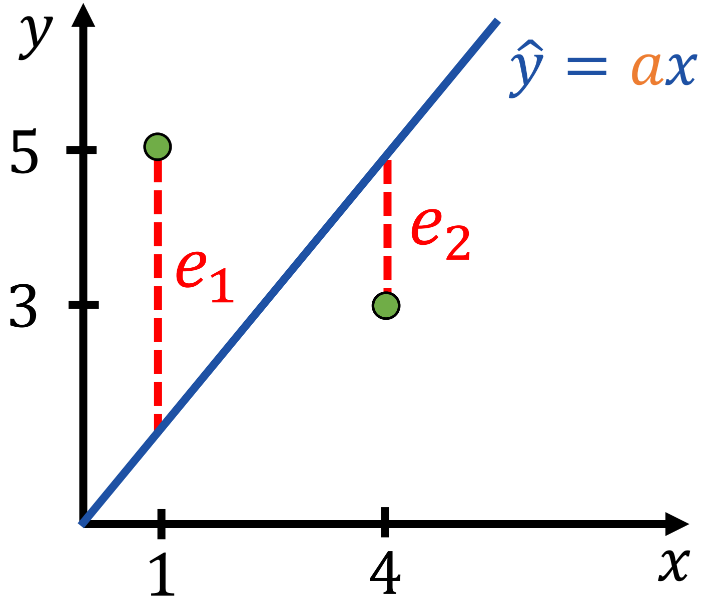
Требуется найти наилучший параметр a

$$e_1 = y_1 - ax_1 = 5 - 1a$$

$$e_2 = y_2 - ax_2 = 3 - 4a$$

$$\Rightarrow \quad e = (5 - 1a, 3 - 4a)$$

Простой пример аппроксимации



Точки

$(1, 5)$ и $(4, 3)$

Общий вид функции

$\hat{y} = ax$

Требуется найти наилучший параметр a

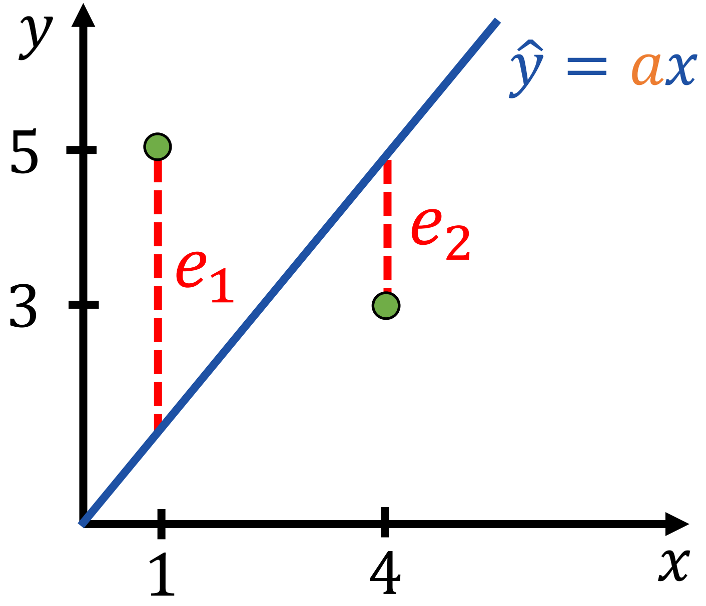
$$e_1 = y_1 - ax_1 = 5 - 1a$$

$$e_2 = y_2 - ax_2 = 3 - 4a$$

$$\Rightarrow \quad e = (5 - 1a, 3 - 4a)$$

Вектор отклонений зависит от выбора a

Простой пример аппроксимации



Точки

$(1, 5)$ и $(4, 3)$

Общий вид функции

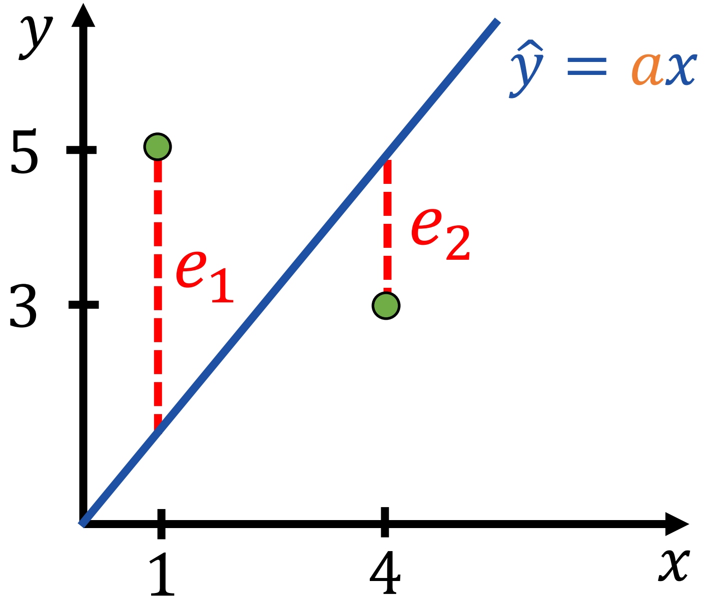
$$\hat{y} = ax$$

Требуется найти наилучший параметр a

Вектор отклонений

$$e = (5 - 1a, 3 - 4a)$$

Простой пример аппроксимации



Точки

$(1, 5)$ и $(4, 3)$

Общий вид функции

$\hat{y} = ax$

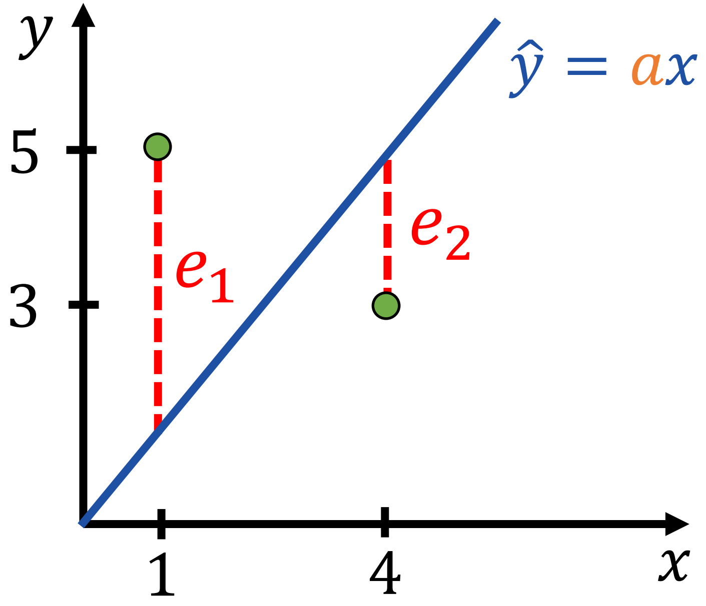
Требуется найти наилучший параметр a

Вектор отклонений

$e = (5 - 1a, 3 - 4a)$

$\Rightarrow$ Разные варианты его нормы

Простой пример аппроксимации



Точки

$(1, 5)$ и $(4, 3)$

Общий вид функции

$\hat{y} = ax$

Требуется найти наилучший параметр a

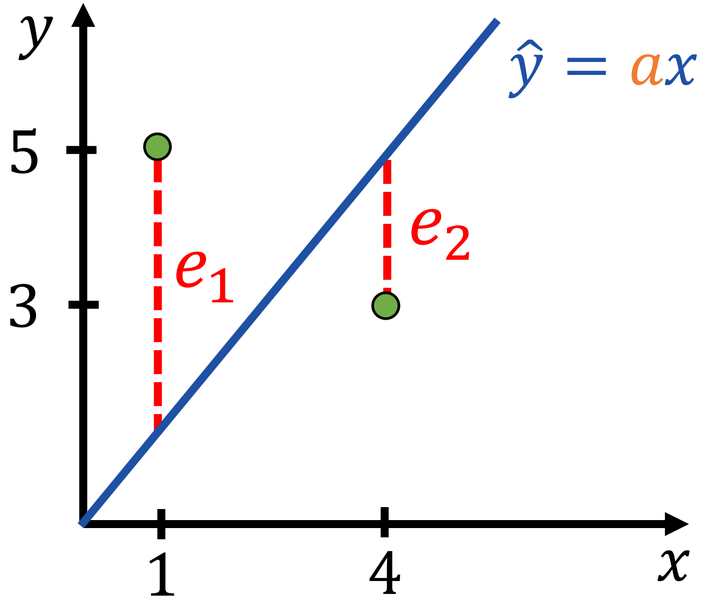
Вектор отклонений

$e = (5 - 1a, 3 - 4a)$

Разные варианты его нормы

$\Rightarrow \quad \|e\|_1 = |5 - 1a| + |3 - 4a|$

Простой пример аппроксимации



Точки

$(1, 5)$ и $(4, 3)$

Общий вид функции

$\hat{y} = ax$

Требуется найти наилучший параметр a

Вектор отклонений

$e = (5 - 1a, 3 - 4a)$

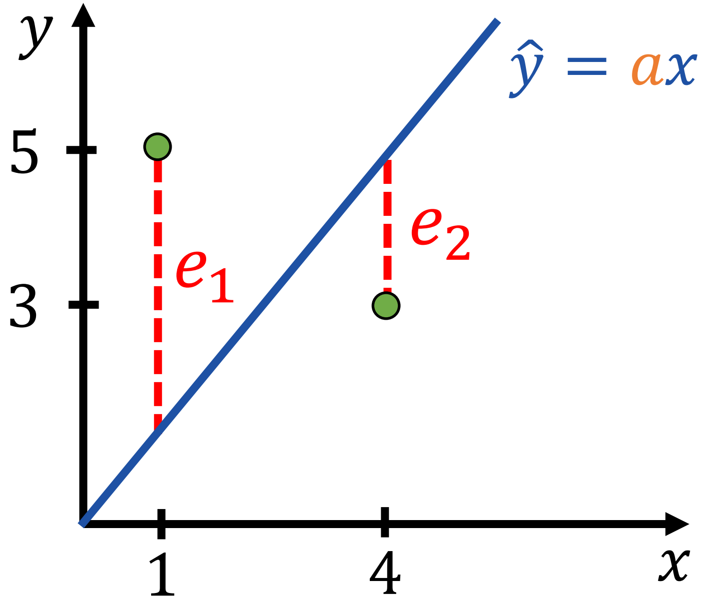
$\Rightarrow$

Разные варианты его нормы

$$\|e\|_1 = |5 - 1a| + |3 - 4a|$$

$$\|e\|_2 = \sqrt{(5 - 1a)^2 + (3 - 4a)^2}$$

Простой пример аппроксимации



Точки

$(1,5)$ и $(4,3)$

Общий вид функции

$$\hat{y} = ax$$

Требуется найти наилучший параметр a

Вектор отклонений

$$e = (5 - 1a, 3 - 4a)$$

$\Rightarrow$

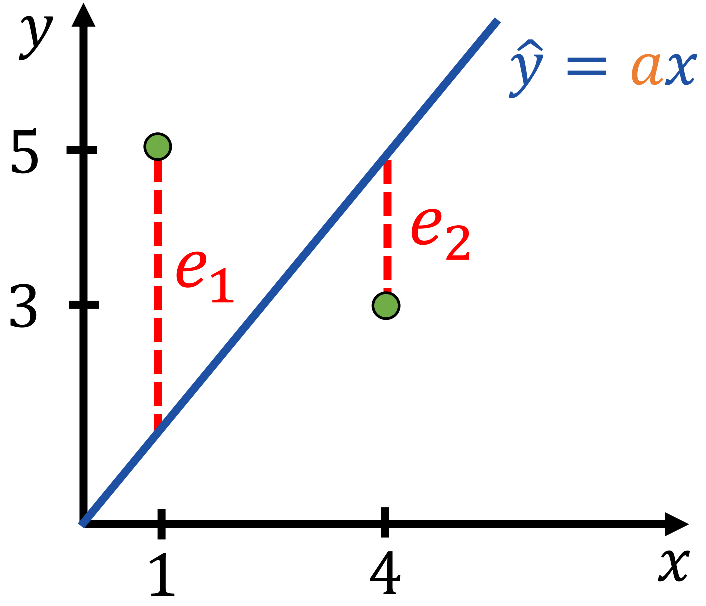
Разные варианты его нормы

$$\|e\|_1 = |5 - 1a| + |3 - 4a|$$

$$\|e\|_2 = \sqrt{(5 - 1a)^2 + (3 - 4a)^2}$$

$$\|e\|_\infty = \max\{|5 - 1a|, |3 - 4a|\}$$

Простой пример аппроксимации



Точки

$(1, 5)$ и $(4, 3)$

Общий вид функции

$$\hat{y} = ax$$

Требуется найти наилучший параметр a

Разные варианты его нормы

Каждая норма является функцией от параметра a

$$\|e\|_1 = |5 - 1a| + |3 - 4a|$$

$$\|e\|_2 = \sqrt{(5 - 1a)^2 + (3 - 4a)^2}$$

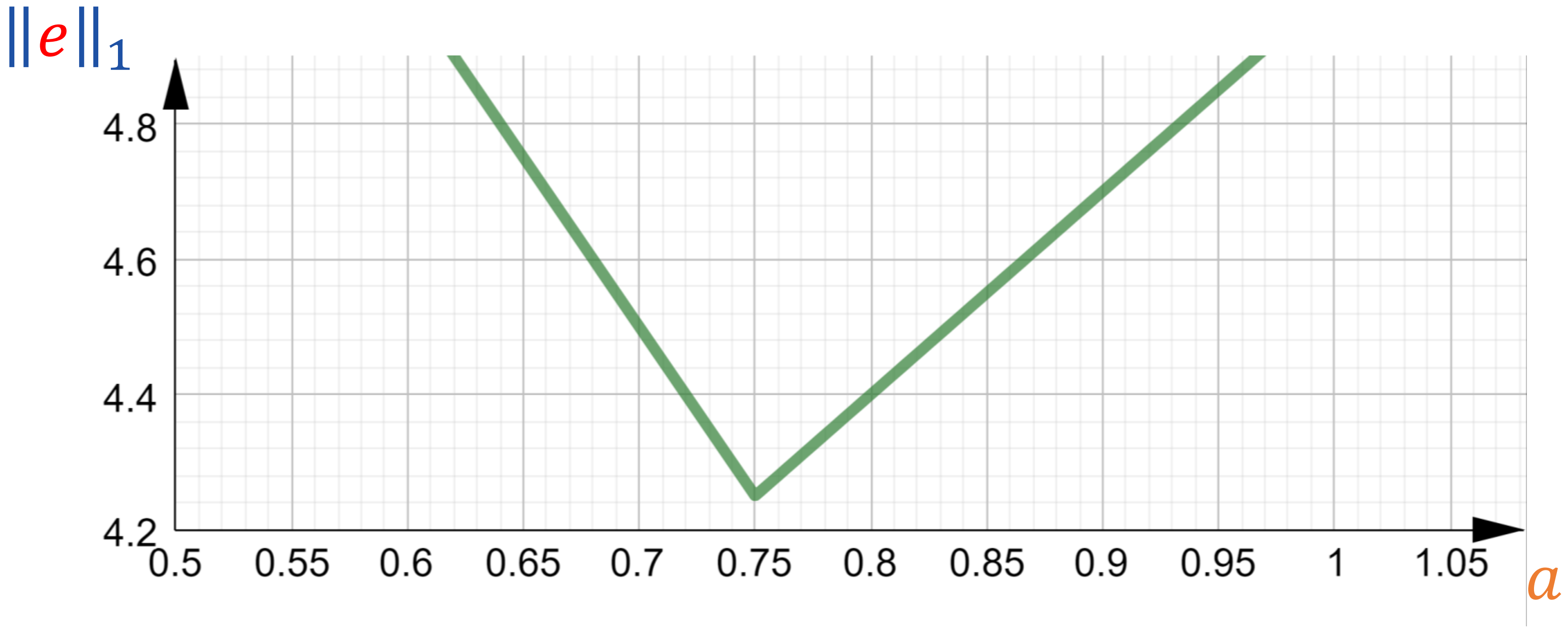
$$\|e\|_\infty = \max\{|5 - 1a|, |3 - 4a|\}$$

Простой пример аппроксимации

$$\|e\|_1 = |5 - 1a| + |3 - 4a|$$

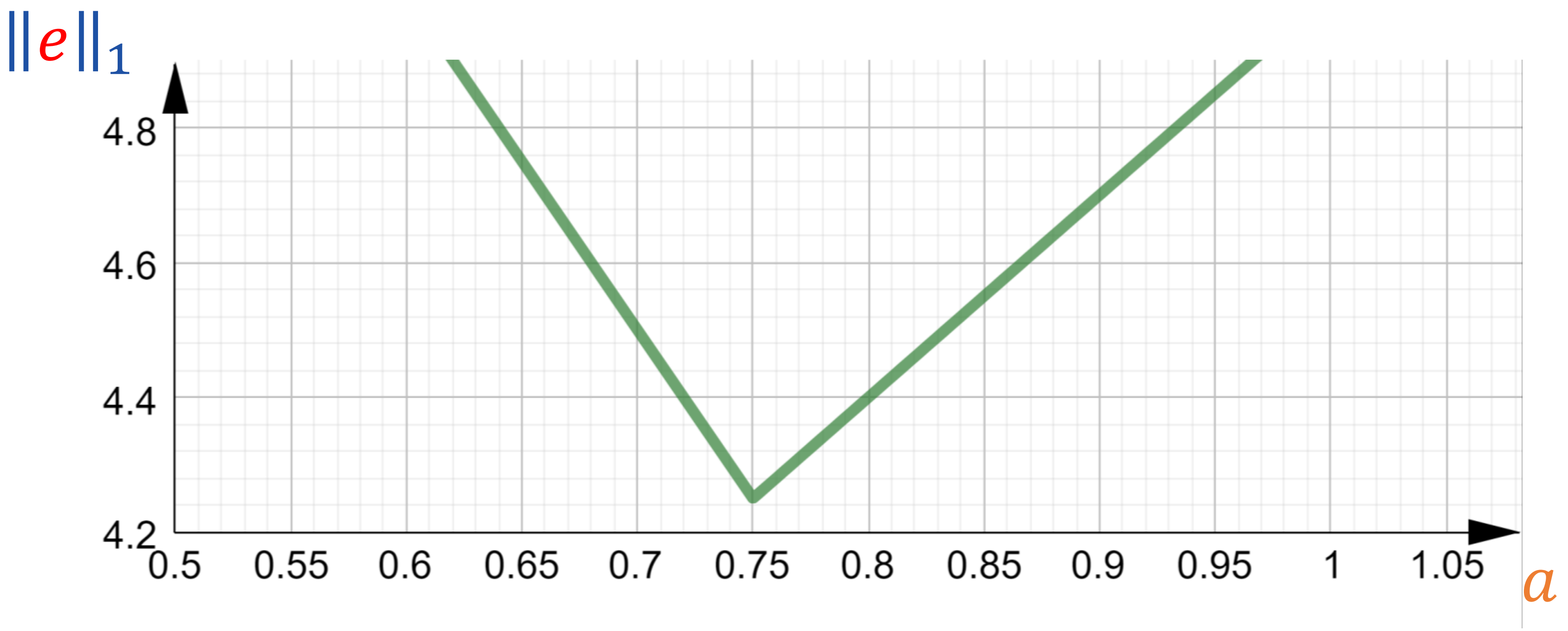
Простой пример аппроксимации

$$\|e\|_1 = |5 - 1a| + |3 - 4a|$$



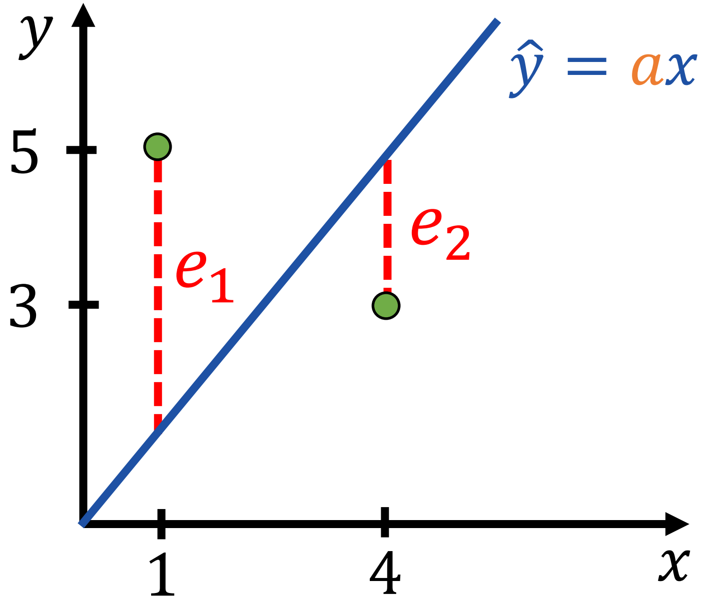
Простой пример аппроксимации

$$\|e\|_1 = |5 - 1a| + |3 - 4a|$$



При $a = 0.75$ норма вектора отклонений оказывается наименьшей, то есть попадание – «наилучшее»

Простой пример аппроксимации



Точки

$(1, 5)$ и $(4, 3)$

Общий вид функции

$\hat{y} = ax$

Требуется найти наилучший параметр a

Разные варианты его нормы

Каждая норма является функцией от параметра a

$$\|e\|_1 = |5 - 1a| + |3 - 4a|$$

$$\|e\|_2 = \sqrt{(5 - 1a)^2 + (3 - 4a)^2}$$

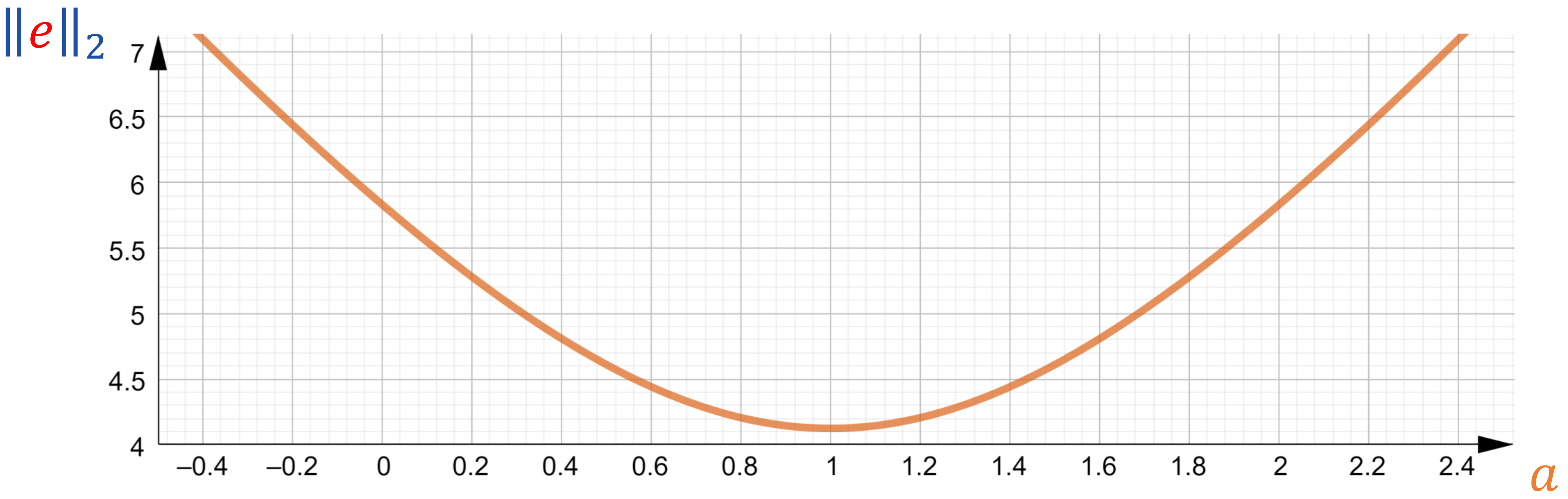
$$\|e\|_\infty = \max\{|5 - 1a|, |3 - 4a|\}$$

Простой пример аппроксимации

$$\|e\|_2 = \sqrt{(5 - 1a)^2 + (3 - 4a)^2}$$

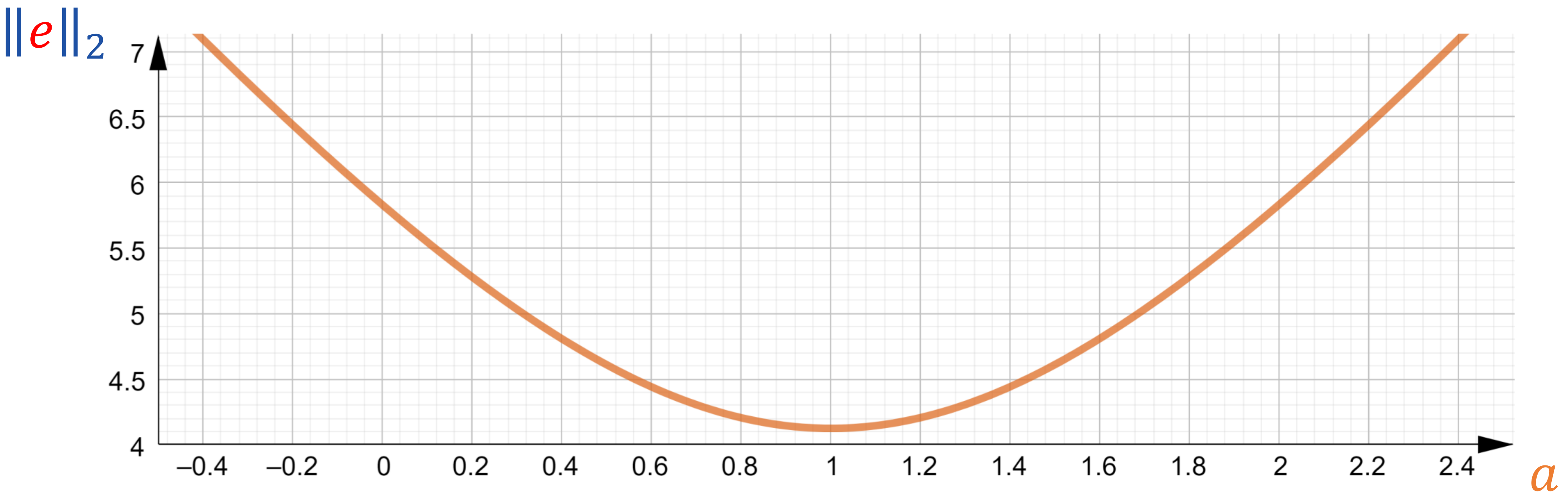
Простой пример аппроксимации

$$\|e\|_2 = \sqrt{(5 - 1a)^2 + (3 - 4a)^2}$$



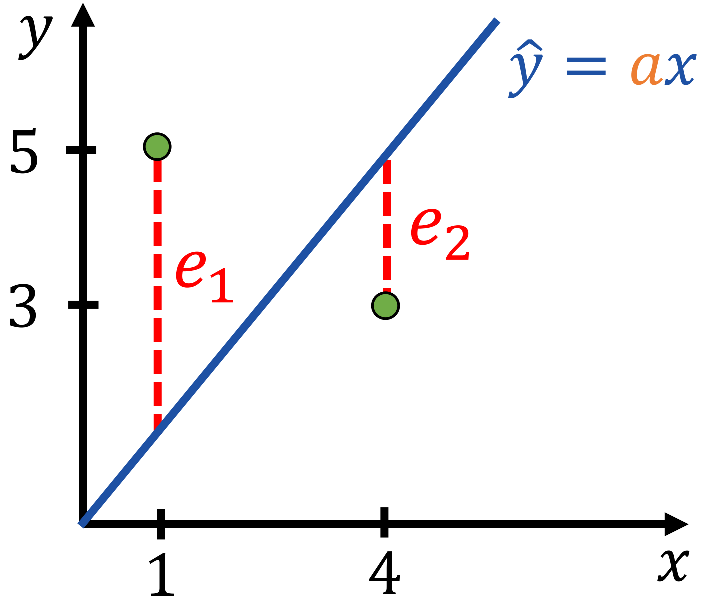
Простой пример аппроксимации

$$\|e\|_2 = \sqrt{(5 - 1a)^2 + (3 - 4a)^2}$$



При $a = 1$ норма вектора отклонений оказывается наименьшей, то есть попадание – «наилучшее»

Простой пример аппроксимации



Точки

$(1, 5)$ и $(4, 3)$

Общий вид функции

$\hat{y} = ax$

Требуется найти наилучший параметр a

Разные варианты его нормы

Каждая норма является функцией от параметра a

$$\|e\|_1 = |5 - 1a| + |3 - 4a|$$

$$\|e\|_2 = \sqrt{(5 - 1a)^2 + (3 - 4a)^2}$$

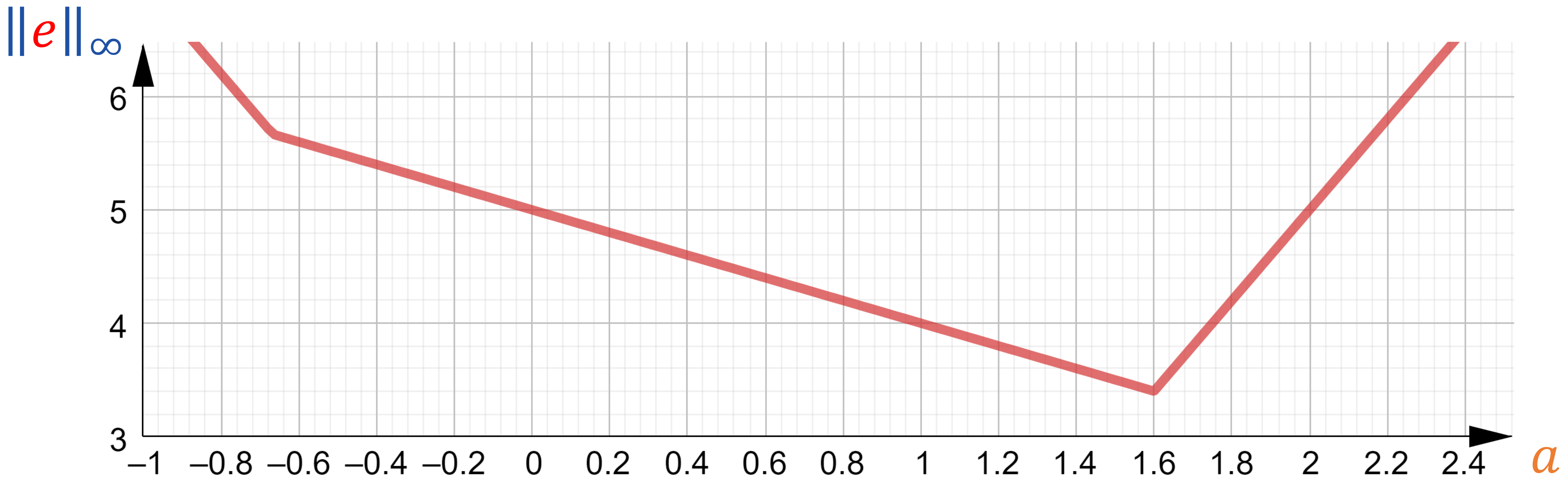
$$\|e\|_\infty = \max\{|5 - 1a|, |3 - 4a|\}$$

Простой пример аппроксимации

$$\|e\|_\infty = \max\{|5 - 1a|, |3 - 4a|\}$$

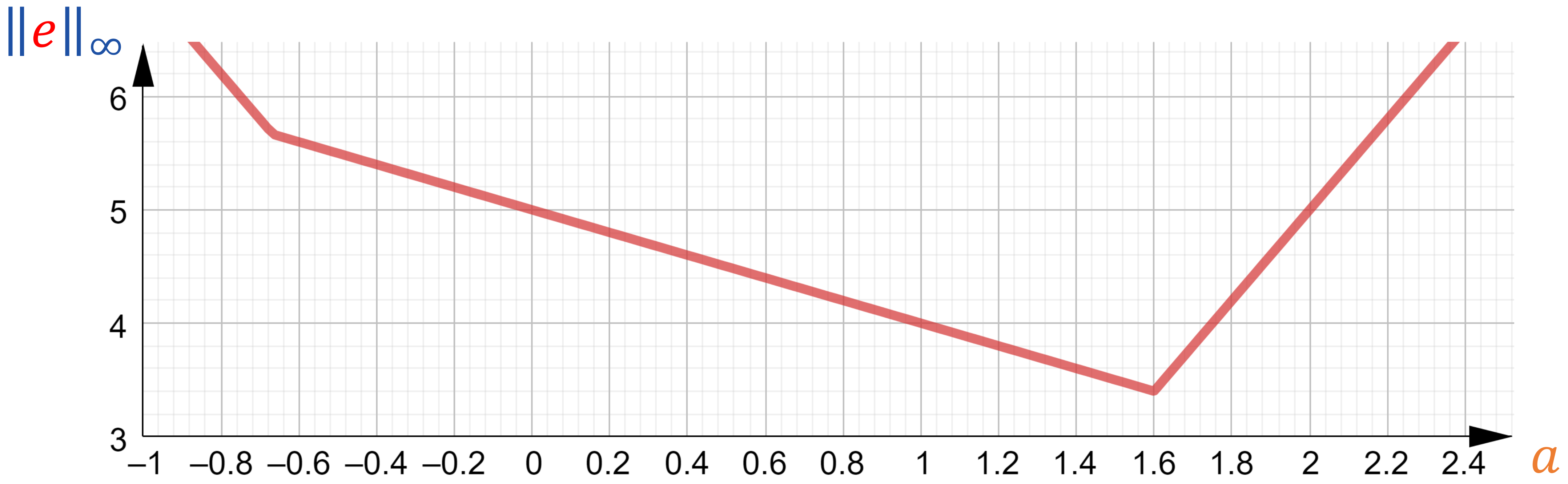
Простой пример аппроксимации

$$\|e\|_\infty = \max\{|5 - 1a|, |3 - 4a|\}$$



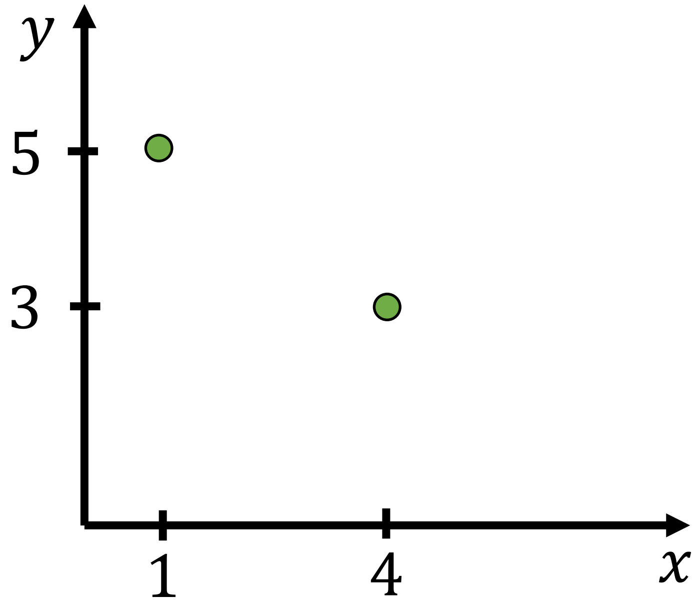
Простой пример аппроксимации

$$\|e\|_\infty = \max\{|5 - 1a|, |3 - 4a|\}$$

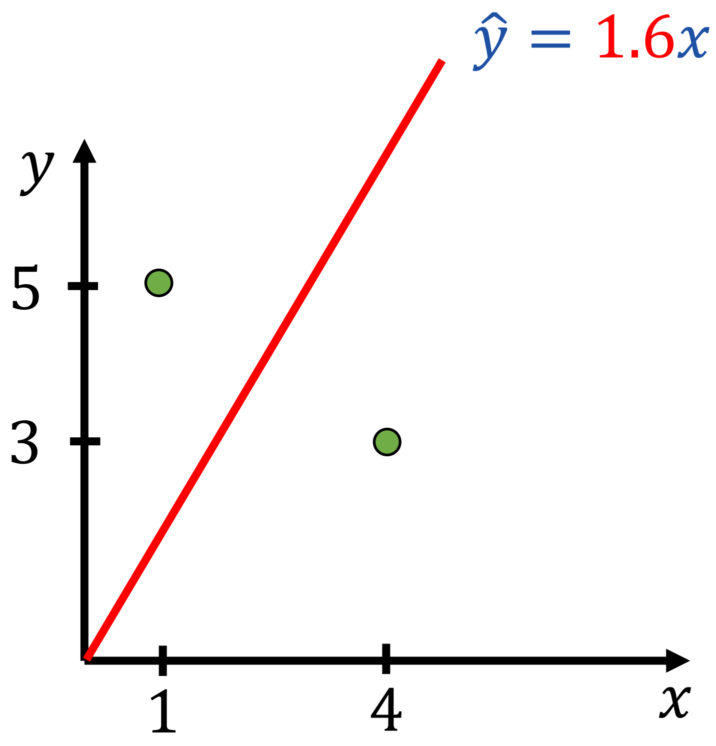


При $a = 1.6$ норма вектора отклонений оказывается наименьшей, то есть попадание – «наилучшее»

Простой пример аппроксимации

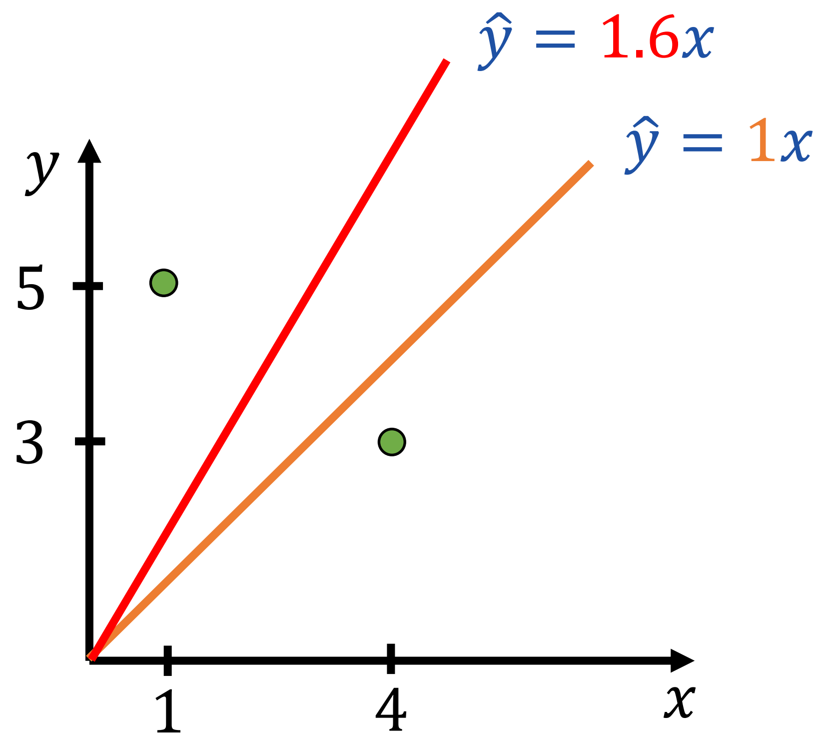


Простой пример аппроксимации



Наилучшая аппроксимация относительно l_∞-нормы (l_∞-минимизация)

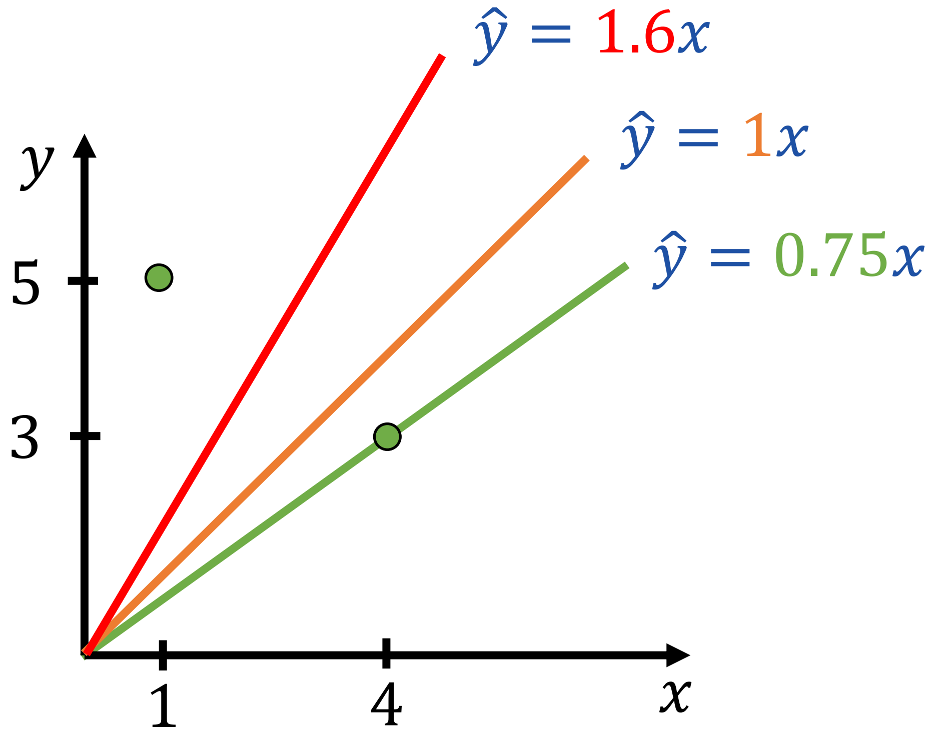
Простой пример аппроксимации



Наилучшая аппроксимация относительно l_∞-нормы (l_∞-минимизация)

Наилучшая аппроксимация относительно l_2-нормы (Метод наименьших квадратов)

Простой пример аппроксимации

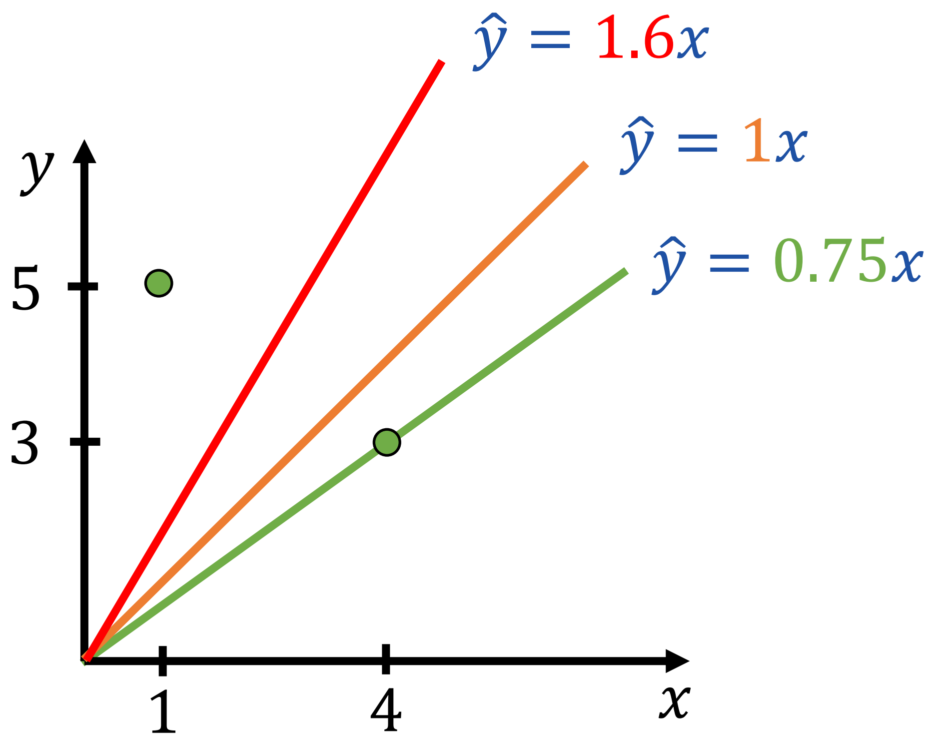


Наилучшая аппроксимация относительно l_∞-нормы (l_∞-минимизация)

Наилучшая аппроксимация относительно l_2-нормы (Метод наименьших квадратов)

Наилучшая аппроксимация относительно l_1-нормы (Метод наименьших модулей)

Простой пример аппроксимации



Наилучшая аппроксимация
относительно l_∞-нормы
(l_∞-минимизация)

Наилучшая аппроксимация
относительно l_2-нормы
(Метод наименьших квадратов)

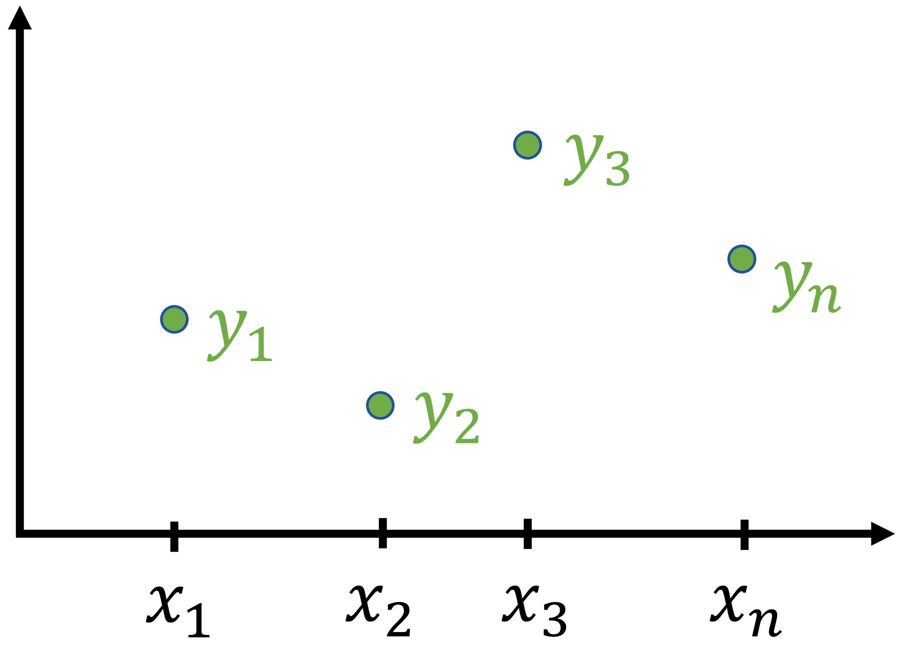
Наилучшая аппроксимация
относительно l_1-нормы
(Метод наименьших модулей)

Чаще всего используется метод наименьших квадратов.
Познакомимся с ним поближе!

Метод наименьших квадратов

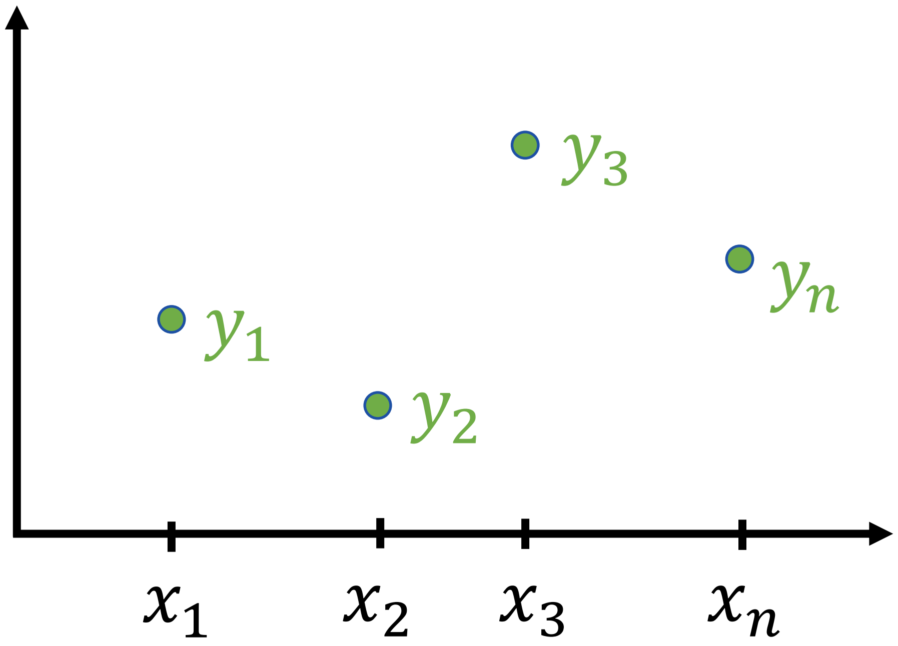
Метод наименьших квадратов

Даны точки, которые надо аппроксимировать



Метод наименьших квадратов

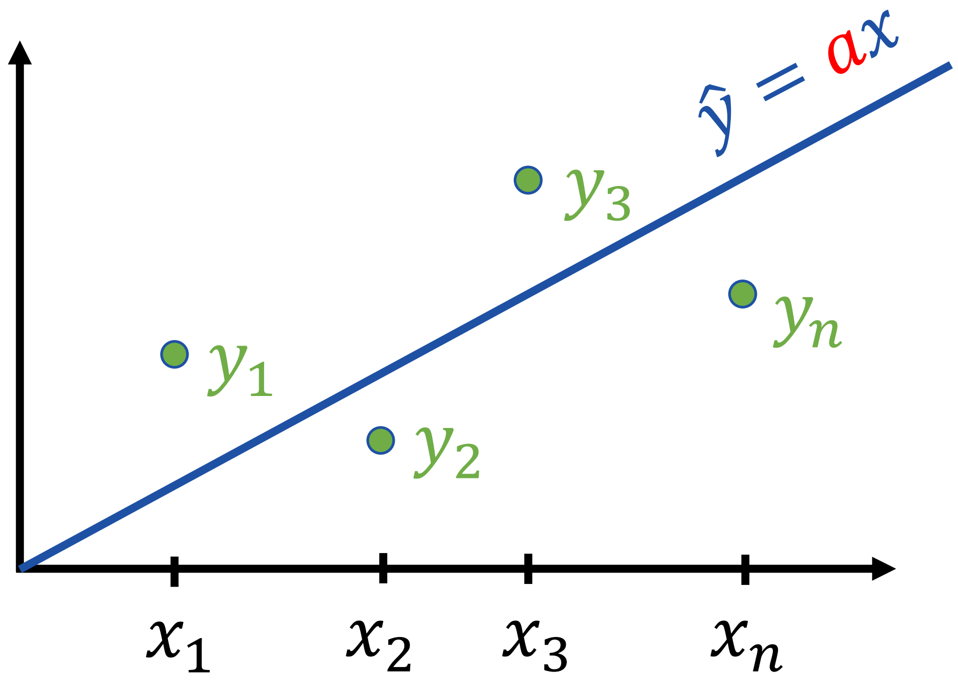
Даны точки, которые надо аппроксимировать



Шаг 1. Выбрать общий вид аппроксимирующей функции

Метод наименьших квадратов

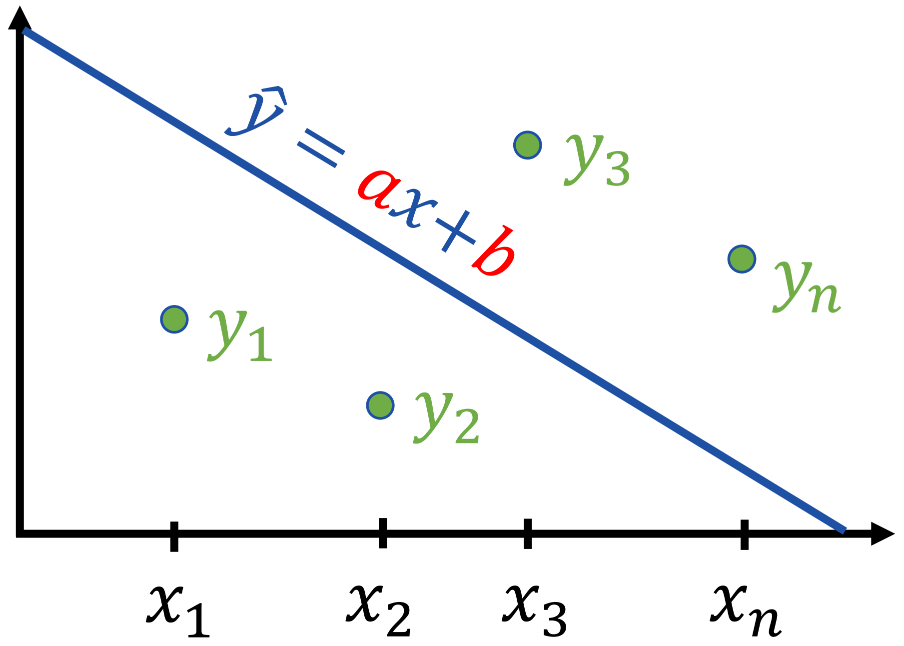
Даны точки, которые надо аппроксимировать



Шаг 1. Выбрать общий вид аппроксимирующей функции

Метод наименьших квадратов

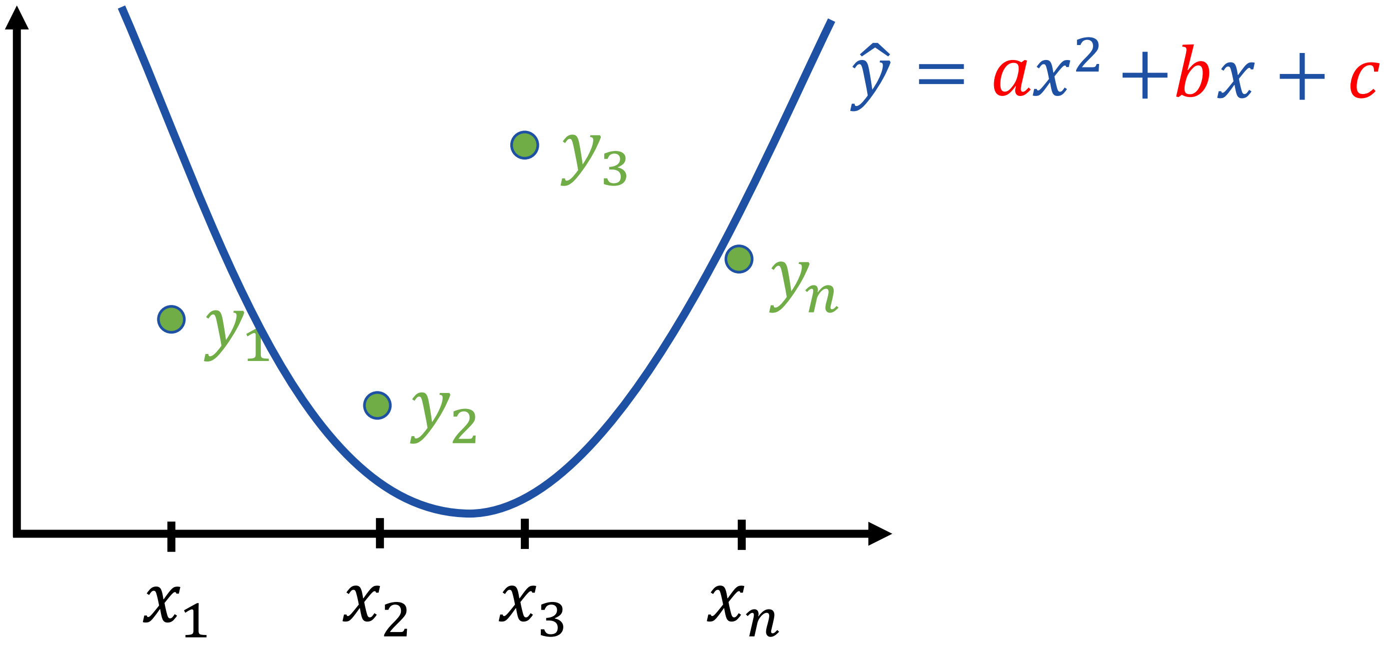
Даны точки, которые надо аппроксимировать



Шаг 1. Выбрать общий вид аппроксимирующей функции

Метод наименьших квадратов

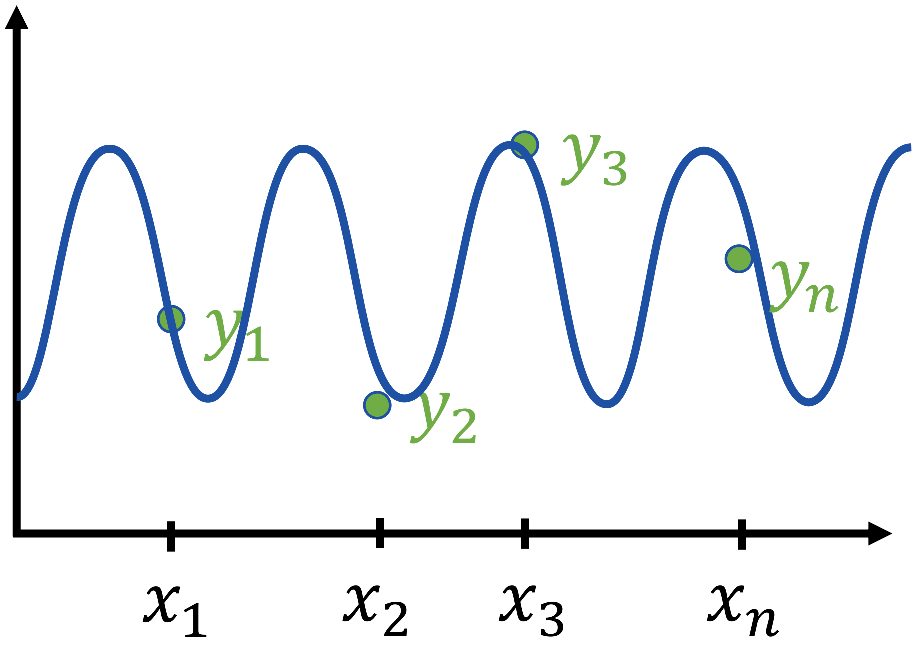
Даны точки, которые надо аппроксимировать



Шаг 1. Выбрать общий вид аппроксимирующей функции

Метод наименьших квадратов

Даны точки, которые надо аппроксимировать

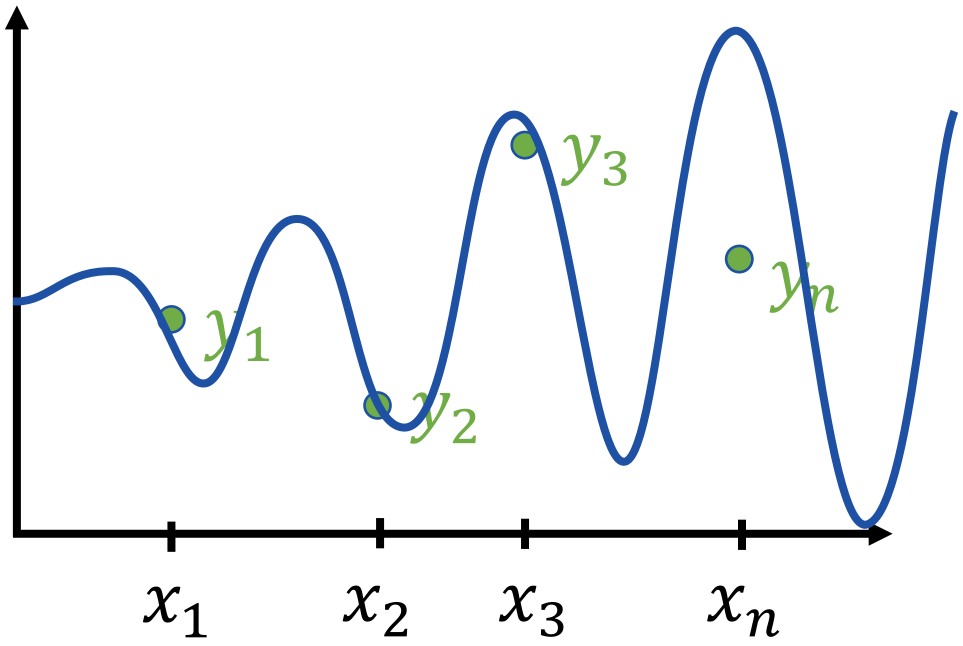


$$\hat{y} = a\sin(bx + c) + d$$

Шаг 1. Выбрать общий вид аппроксимирующей функции

Метод наименьших квадратов

Даны точки, которые надо аппроксимировать

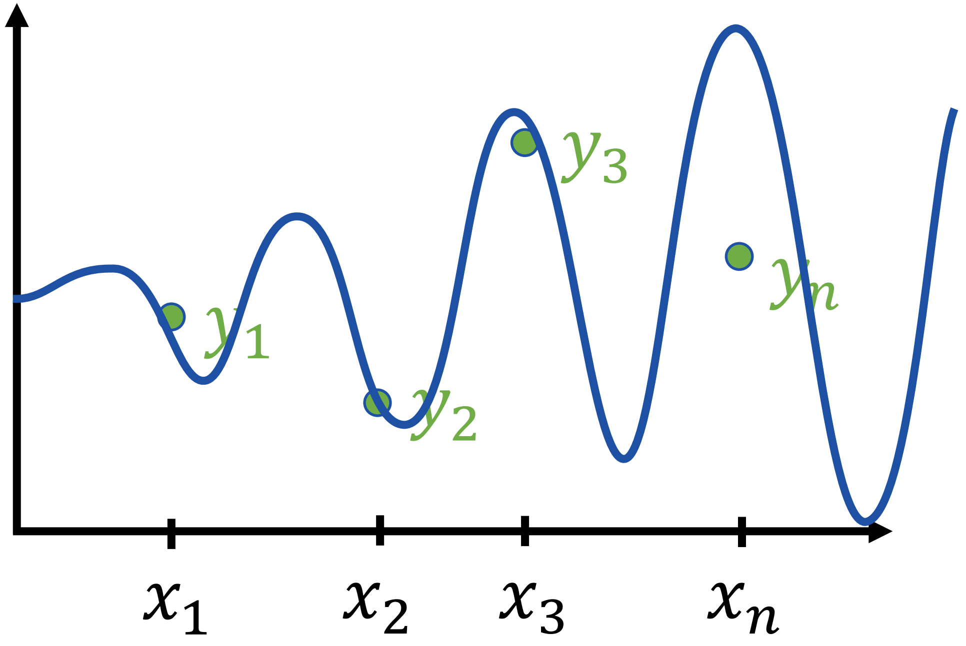


$$\hat{y} = ae^{bx}\sin(cx) + d$$

Шаг 1. Выбрать общий вид аппроксимирующей функции

Метод наименьших квадратов

Даны точки, которые надо аппроксимировать



$$\hat{y} = ae^{bx}\sin(cx) + d$$

Шаг 2. Найти значения коэффициентов, при которых l_2-норма вектора отклонений наименьшая

Метод наименьших квадратов

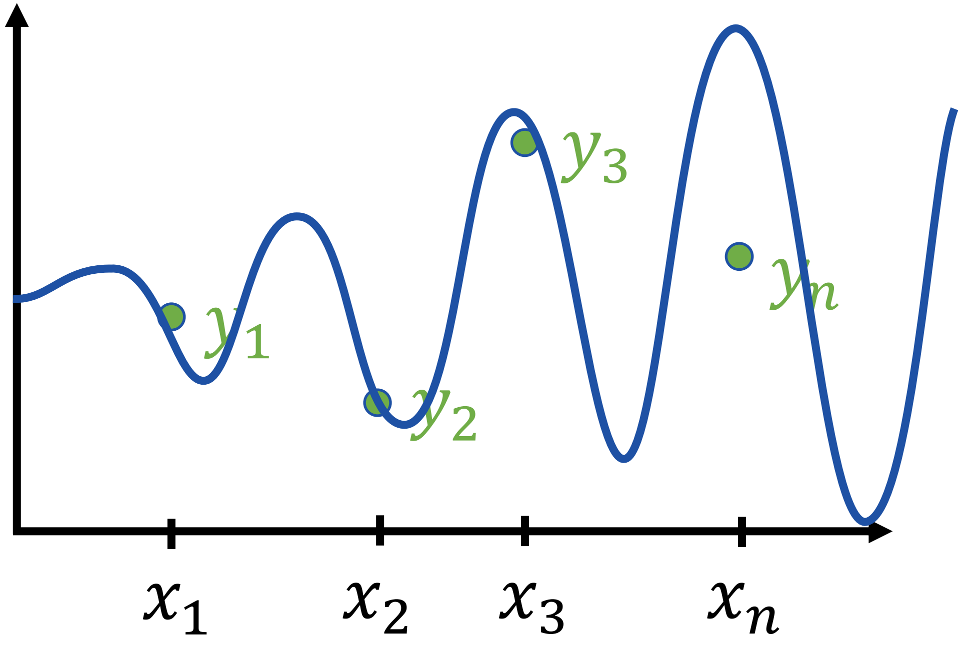
Даны точки, которые надо аппроксимировать

$a = ?\quad b = ?$

$c = ?\quad d = ?$

$\Downarrow$

$\|e\|_2 \to \min$

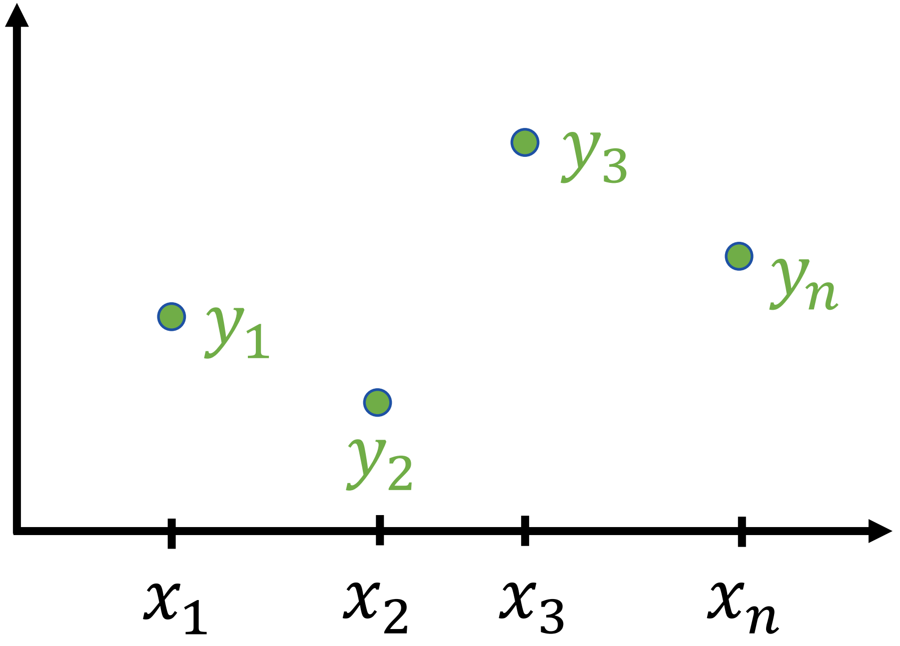


$\hat{y} = ae^{bx}\sin(cx) + d$

Шаг 2. Найти значения коэффициентов, при которых l_2-норма вектора отклонений наименьшая

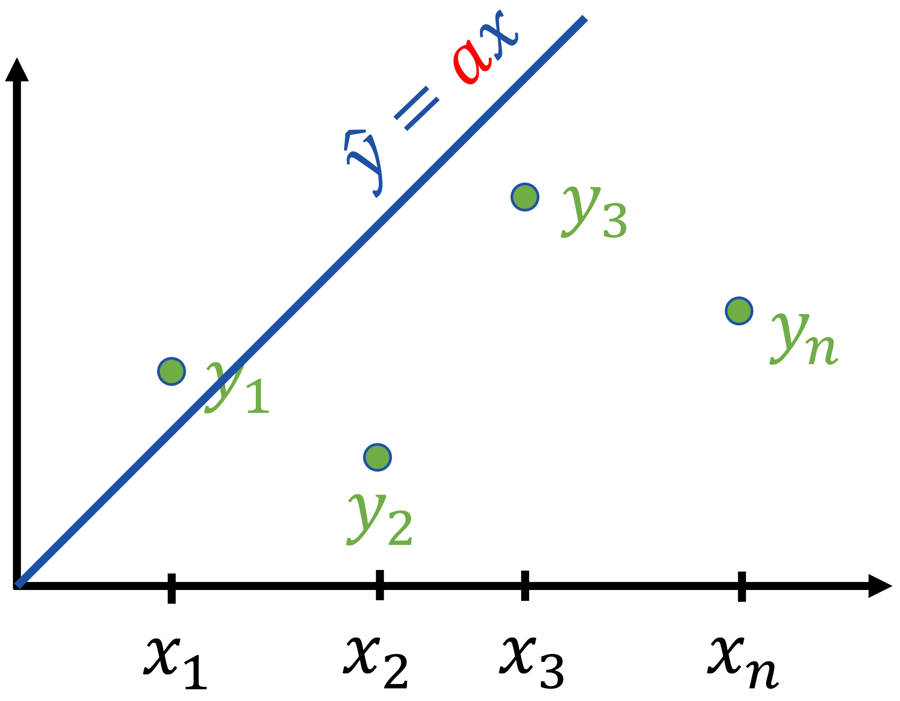
МНК для $\hat{y} = ax$

МНК для $\hat{y} = ax$



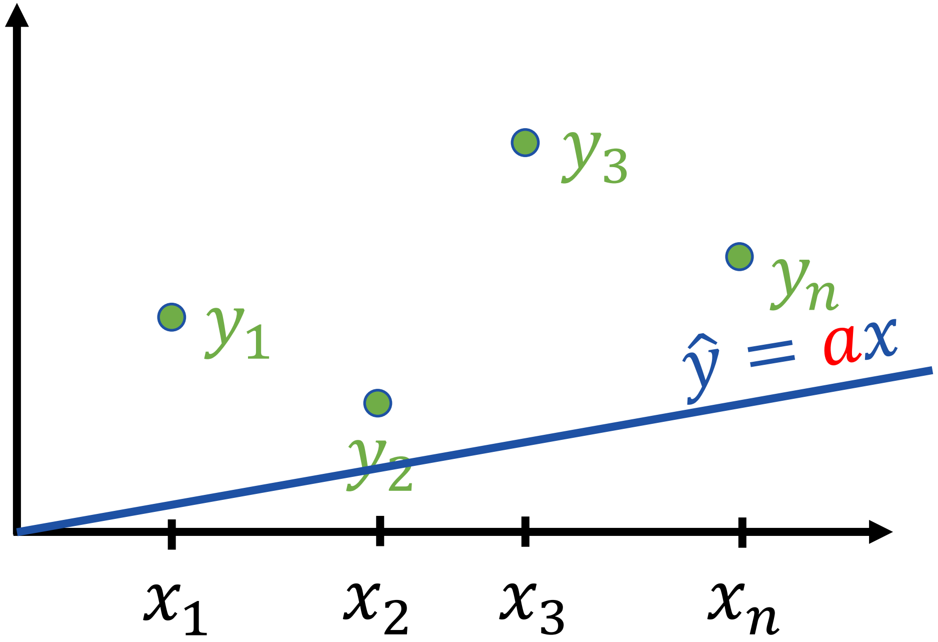
Даны точки, которые надо аппроксимировать прямой $\hat{y} = ax$

МНК для $\hat{y} = ax$



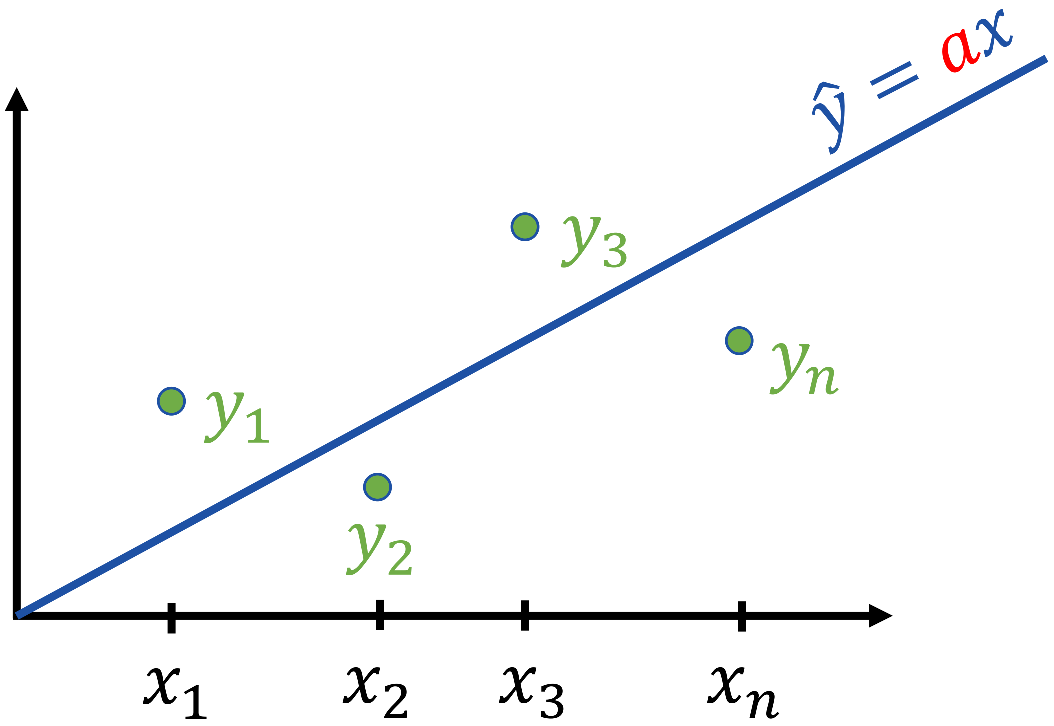
Предположим, мы проведём прямую с каким-то коэффициентом a.

МНК для $\hat{y} = ax$



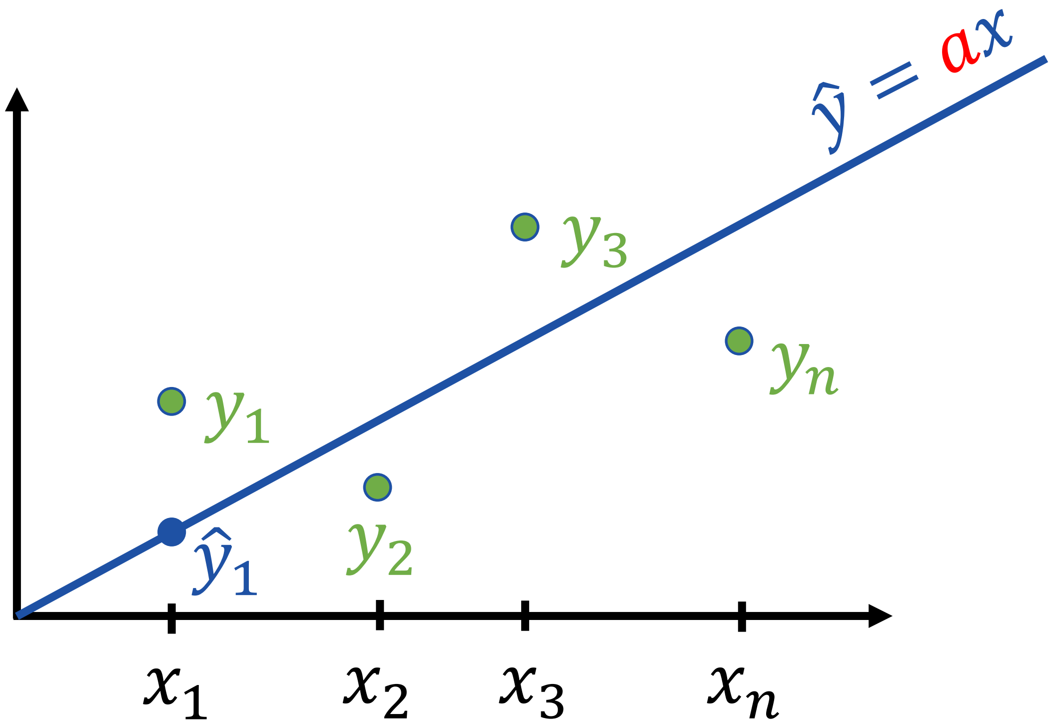
Предположим, мы проведём прямую с каким-то коэффициентом a.

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

МНК для $\hat{y} = ax$

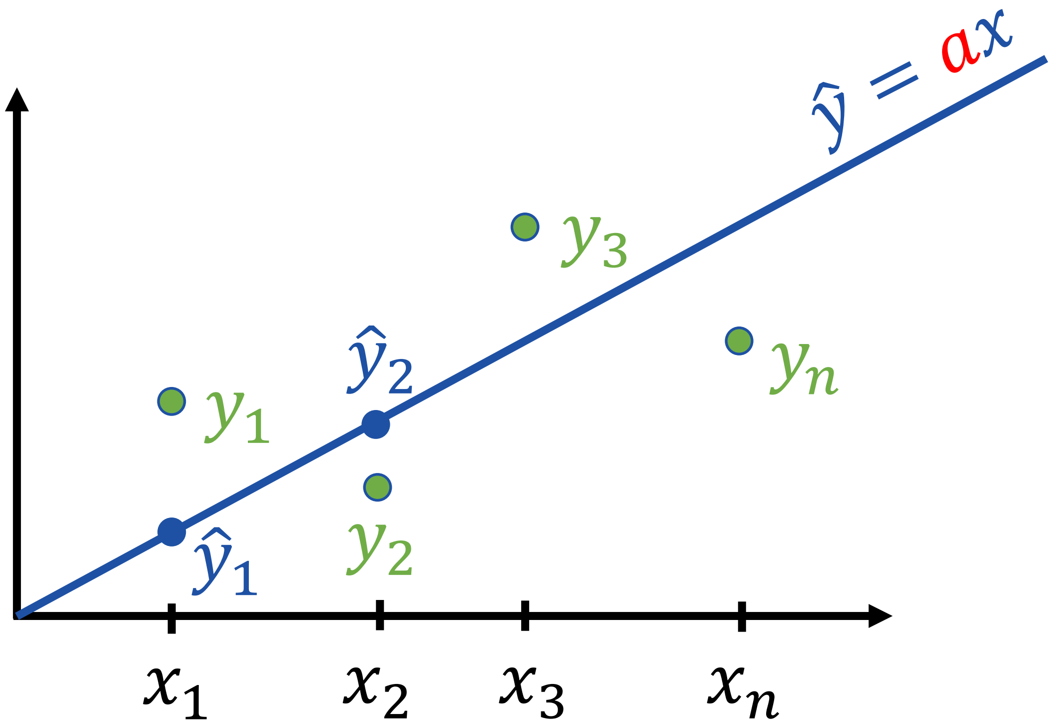


Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

$$\hat{y}_1 = ax_1$$

МНК для $\hat{y} = ax$



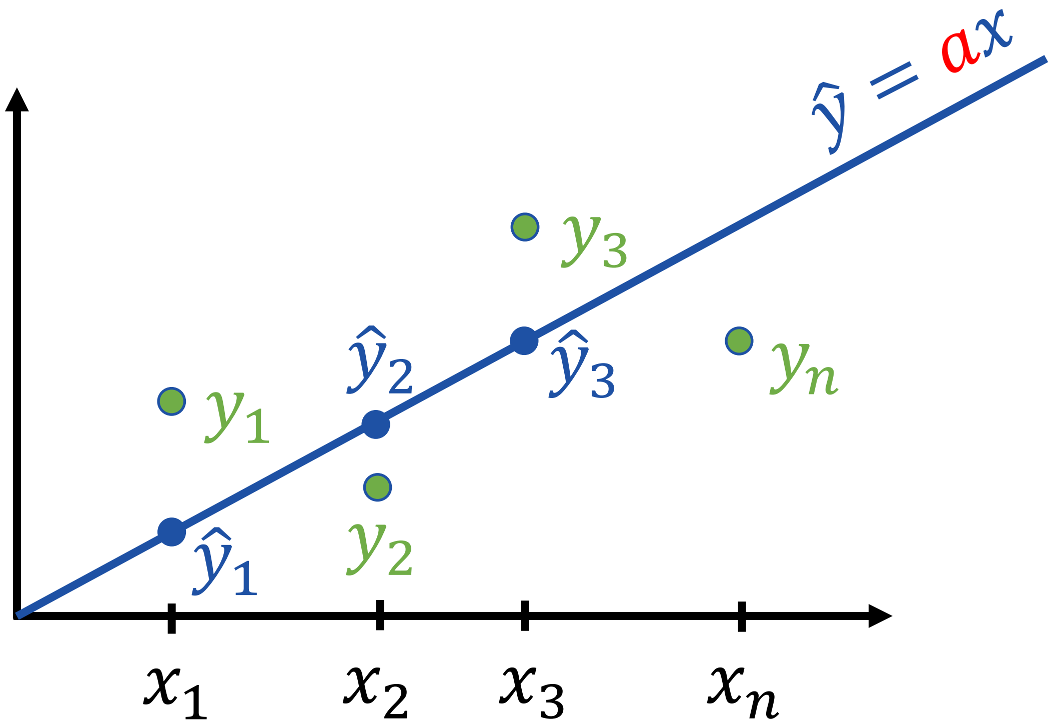
Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

$$\hat{y}_1 = ax_1$$

$$\hat{y}_2 = ax_2$$

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

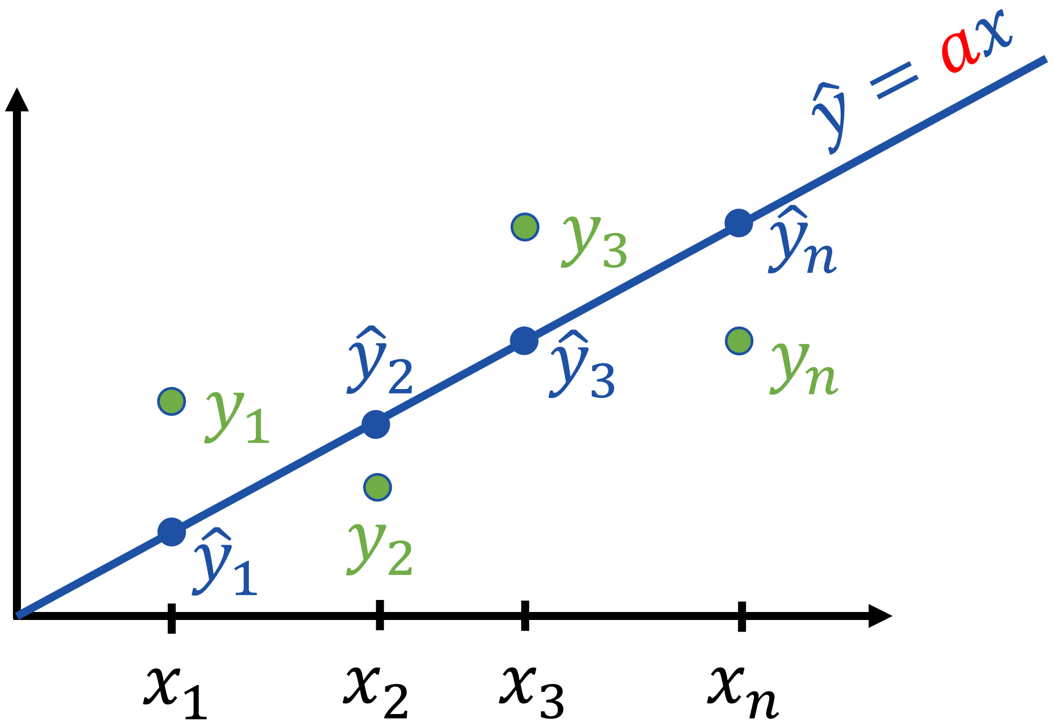
Тогда

$$\hat{y}_1 = ax_1$$

$$\hat{y}_2 = ax_2$$

$$\hat{y}_3 = ax_3$$

МНК для $\hat{y} = ax$

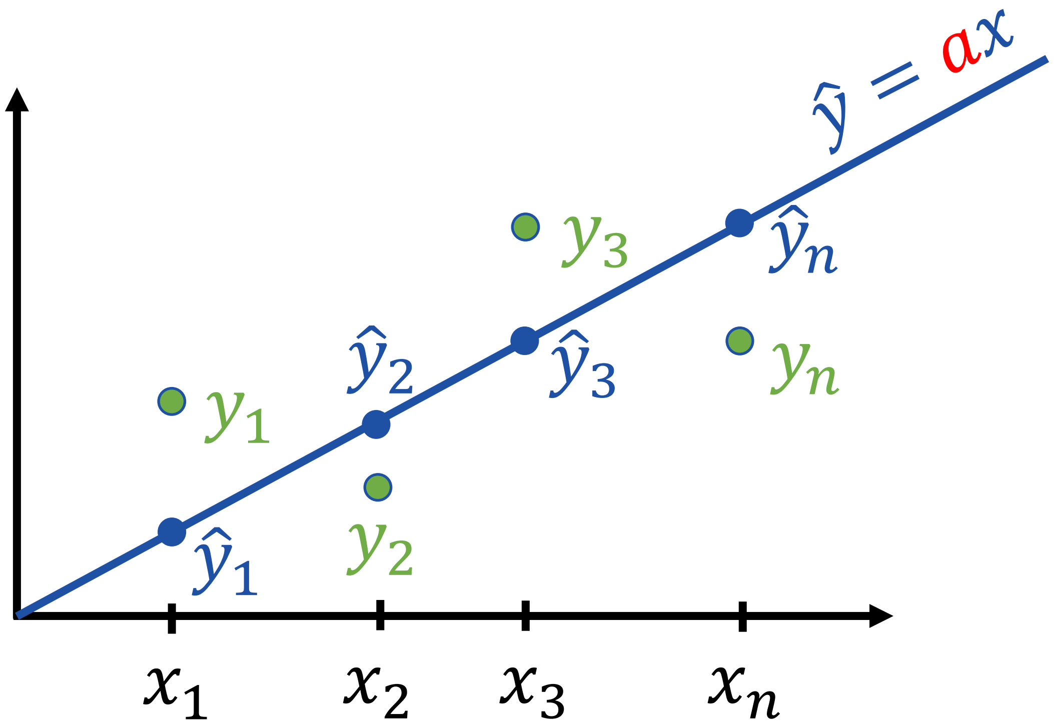


Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

$$\hat{y}_1 = ax_1$$
$$\hat{y}_2 = ax_2$$
$$\hat{y}_3 = ax_3$$
$$\vdots$$
$$\hat{y}_n = ax_n$$

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

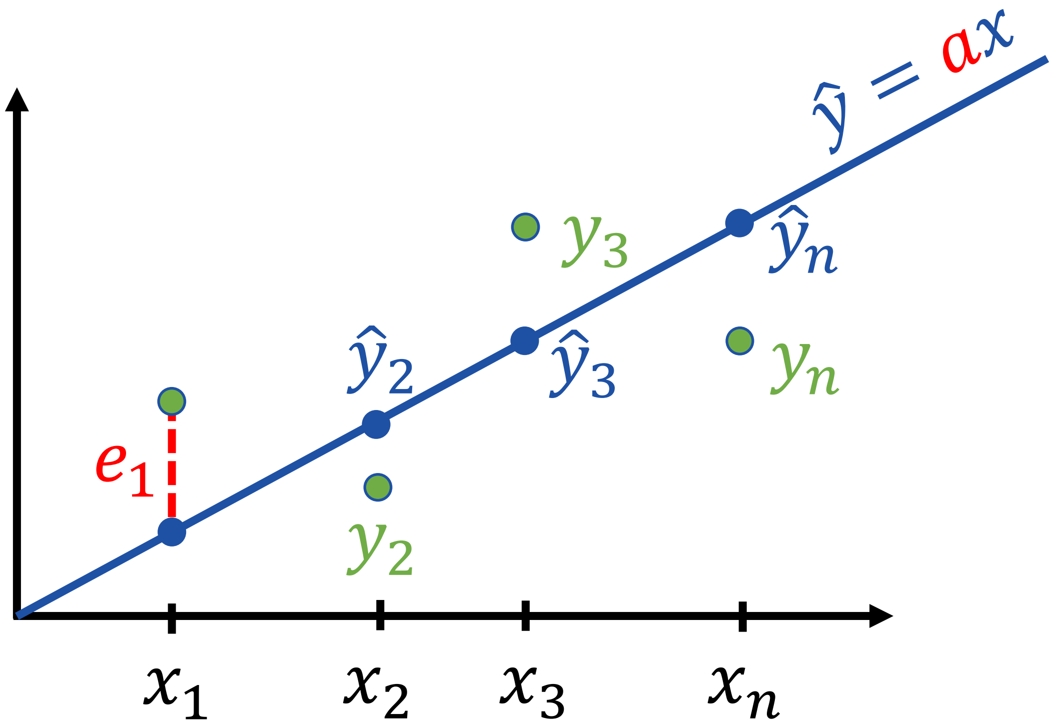
Тогда

$$\hat{y}_1 = ax_1$$
$$\hat{y}_2 = ax_2$$
$$\hat{y}_3 = ax_3$$
$$\vdots$$
$$\hat{y}_n = ax_n$$

$\Rightarrow$

Отклонения

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

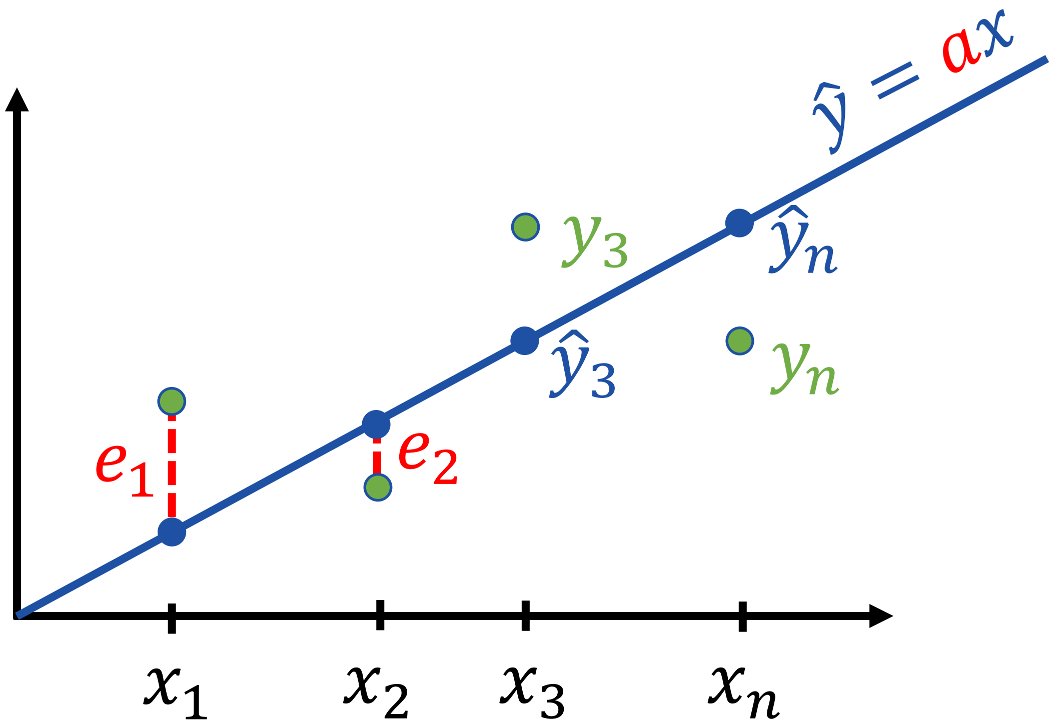
$$\hat{y}_1 = ax_1$$
$$\hat{y}_2 = ax_2$$
$$\hat{y}_3 = ax_3$$
$$\vdots$$
$$\hat{y}_n = ax_n$$

$$\Rightarrow$$

Отклонения

$$e_1 = y_1 - ax_1$$

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

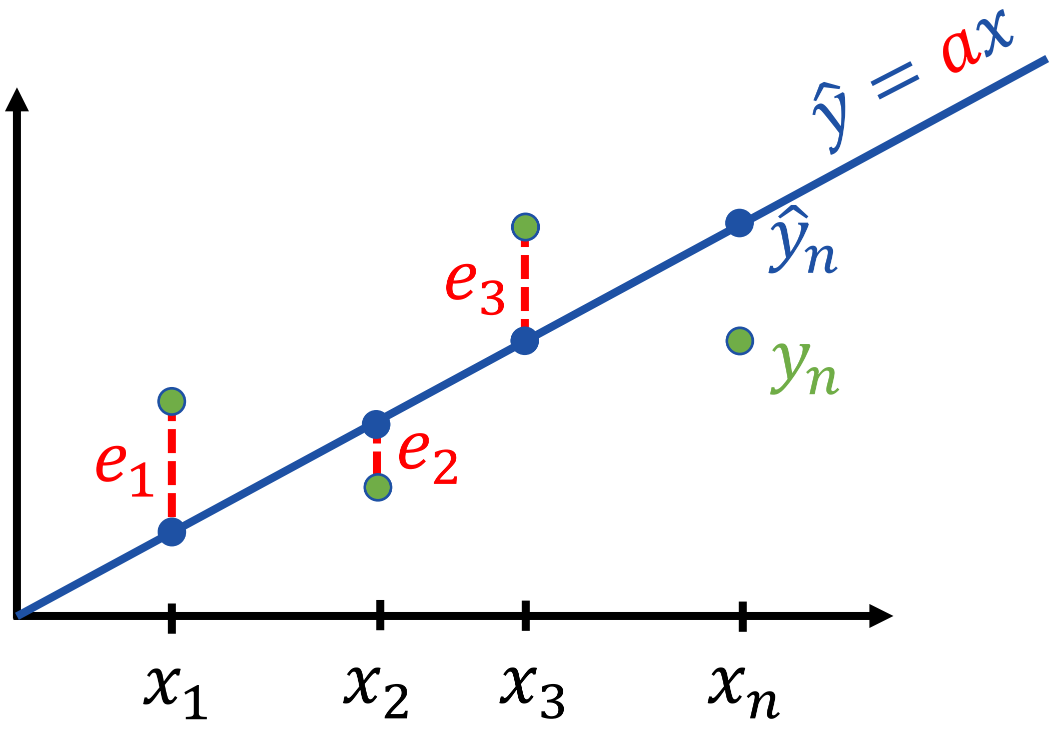
$$\hat{y}_1 = ax_1$$
$$\hat{y}_2 = ax_2$$
$$\hat{y}_3 = ax_3$$
$$\vdots$$
$$\hat{y}_n = ax_n$$

$\Rightarrow$

Отклонения

$$e_1 = y_1 - ax_1$$
$$e_2 = y_2 - ax_2$$

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

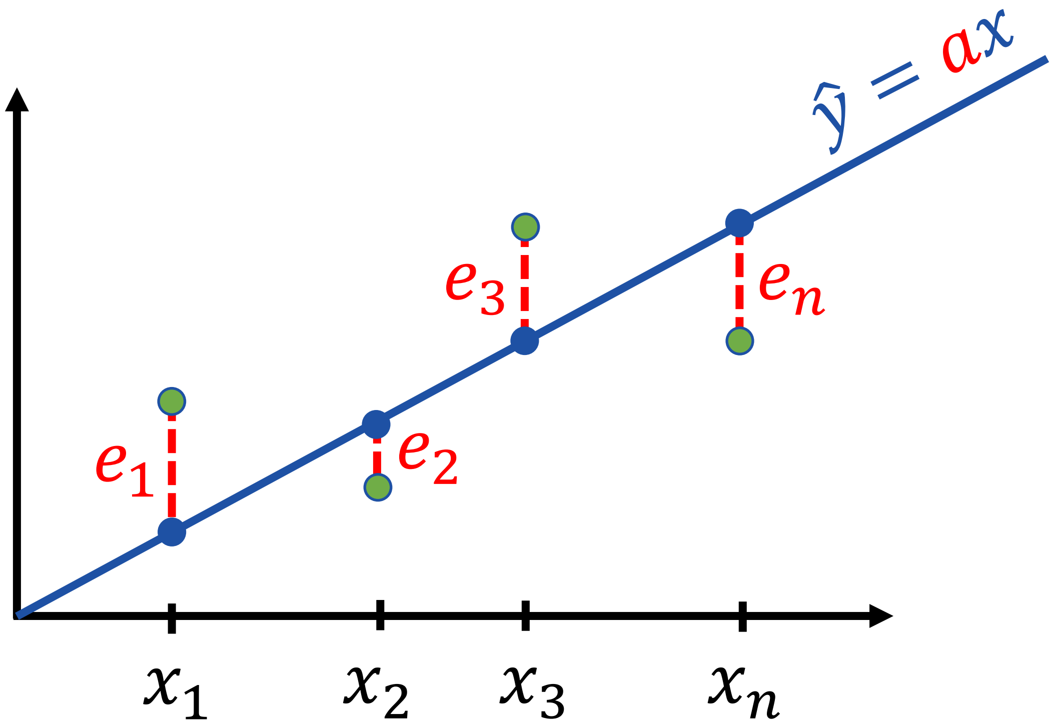
$$\hat{y}_1 = ax_1$$
$$\hat{y}_2 = ax_2$$
$$\hat{y}_3 = ax_3$$
$$\vdots$$
$$\hat{y}_n = ax_n$$

$\Rightarrow$

Отклонения

$$e_1 = y_1 - ax_1$$
$$e_2 = y_2 - ax_2$$
$$e_3 = y_3 - ax_3$$

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

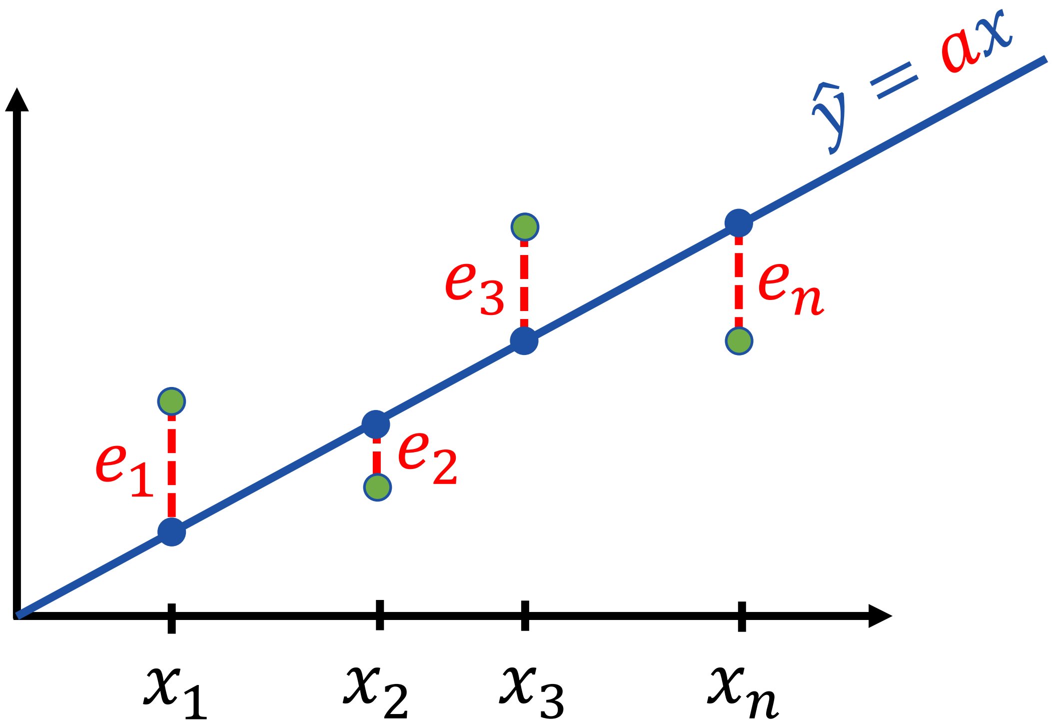
$$\hat{y}_1 = ax_1$$
$$\hat{y}_2 = ax_2$$
$$\hat{y}_3 = ax_3$$
$$\vdots$$
$$\hat{y}_n = ax_n$$

$\Rightarrow$

Отклонения

$$e_1 = y_1 - ax_1$$
$$e_2 = y_2 - ax_2$$
$$e_3 = y_3 - ax_3$$
$$\vdots$$
$$e_n = y_n - ax_n$$

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

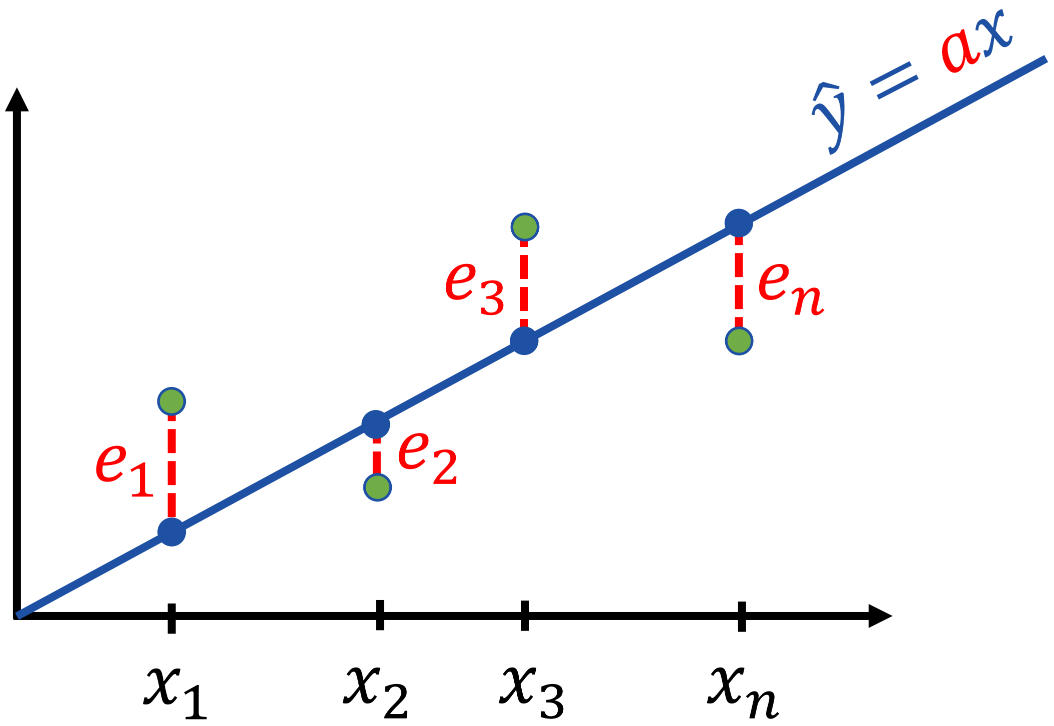
Тогда

$$\hat{y}_i = ax_i$$

Отклонения

$$e_i = y_i - ax_i$$

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

$$\hat{y}_i = ax_i$$

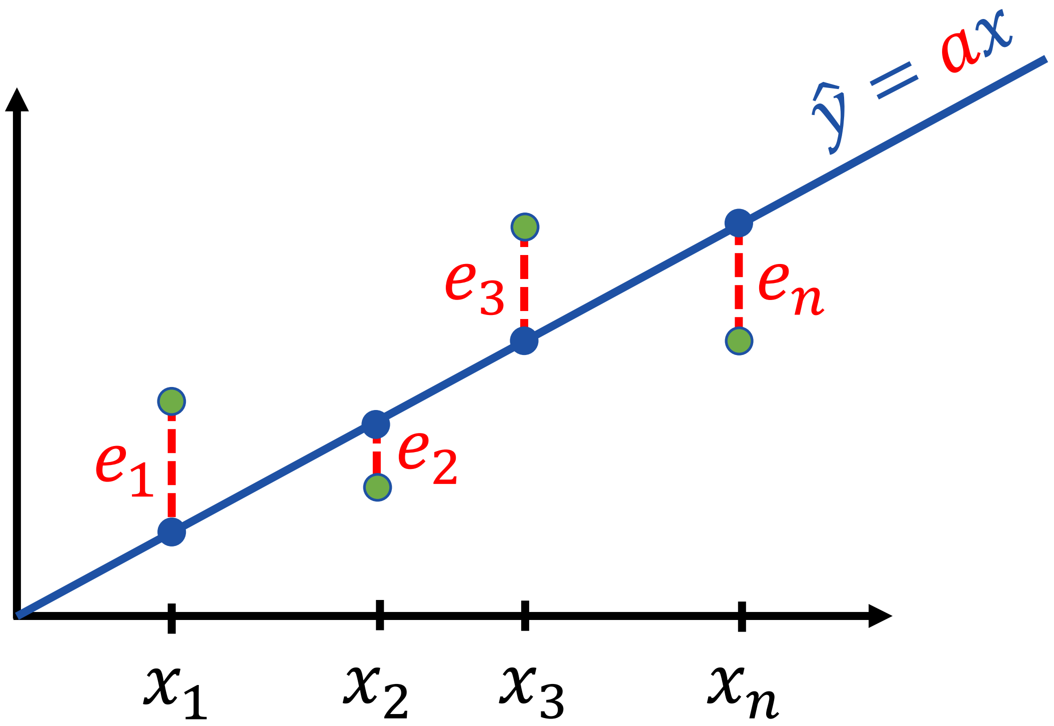
Отклонения

$$e_i = y_i - ax_i$$

Норма вектора отклонений

$$\|e\| = \sqrt{\sum_{i=1}^{n} e_i^2}$$

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

$$\hat{y}_i = ax_i$$

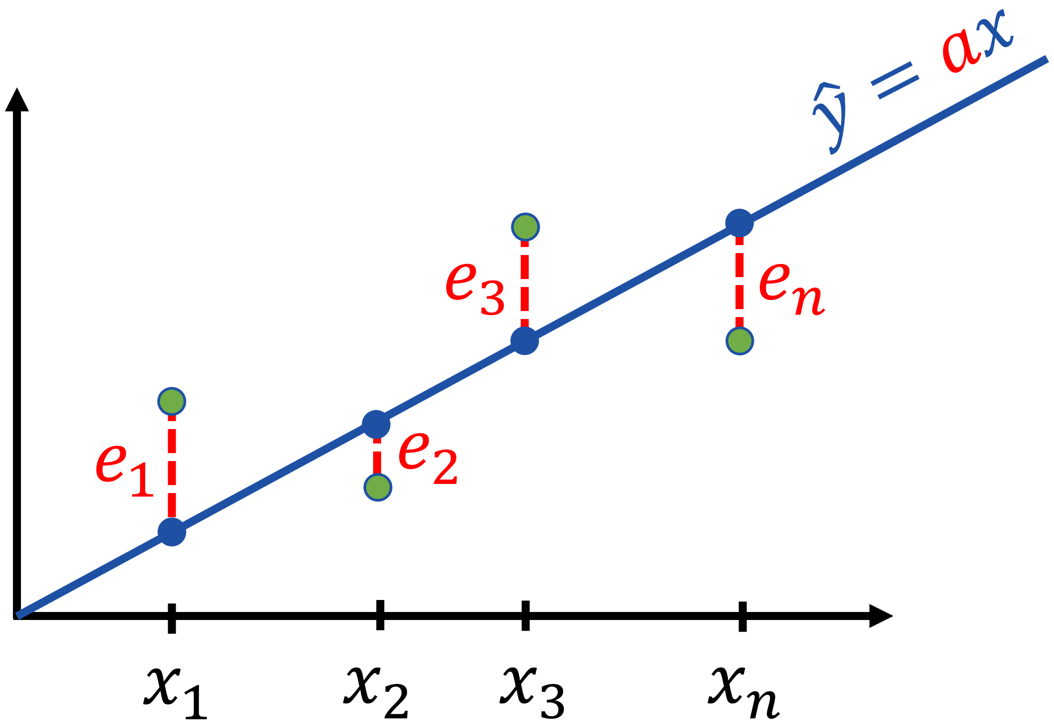
Отклонения

$$e_i = y_i - ax_i$$

Норма вектора отклонений

$$\|e\| = \sqrt{\sum_{i=1}^{n} e_i^2} = \sqrt{\sum_{i=1}^{n} (y_i - ax_i)^2}$$

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

$$\hat{y}_i = ax_i$$

Отклонения

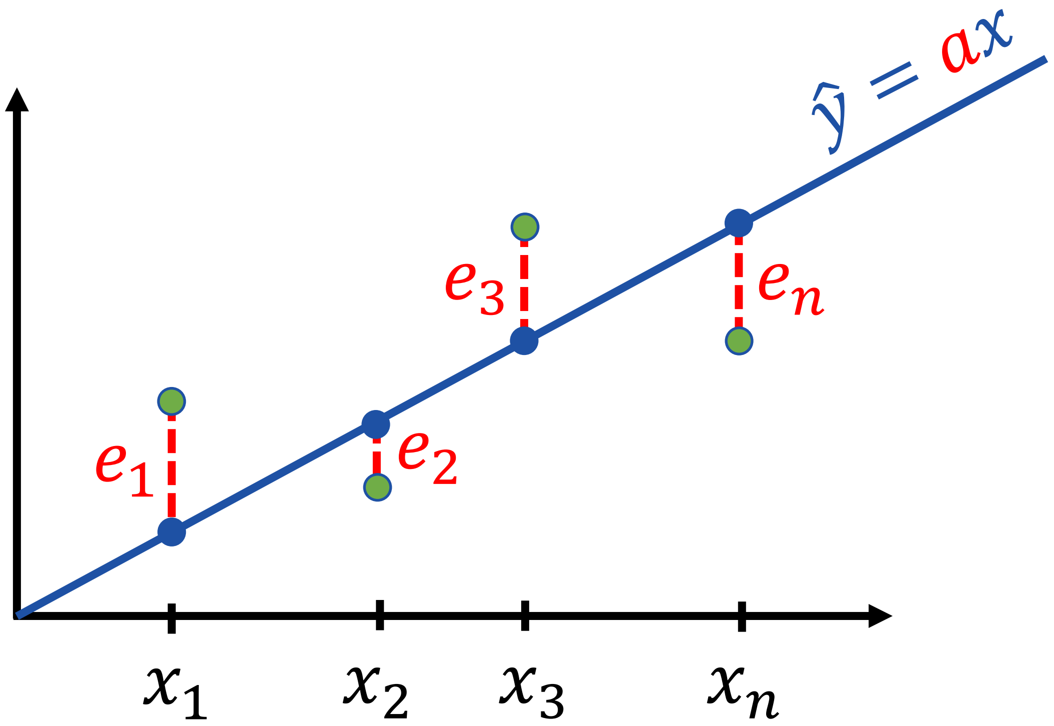
$$e_i = y_i - ax_i$$

Норма вектора отклонений

$$\|e\| = \sqrt{\sum_{i=1}^{n} e_i^2} = \sqrt{\sum_{i=1}^{n} (y_i - ax_i)^2}$$

Надо найти число a, при котором норма наименьшая

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

$$\hat{y}_i = ax_i$$

Отклонения

$$e_i = y_i - ax_i$$

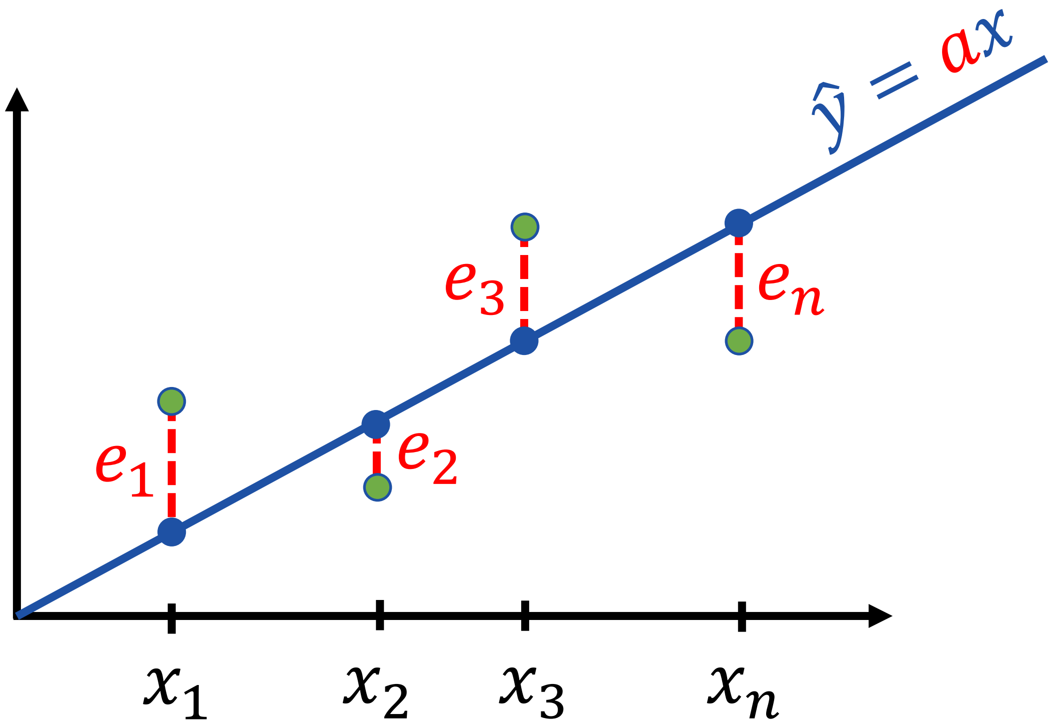
Норма вектора отклонений

$$\|e\| = \sqrt{\sum_{i=1}^{n} e_i^2} = \sqrt{\sum_{i=1}^{n} (y_i - ax_i)^2}$$

Надо найти число a, при котором норма наименьшая

Но если норма наименьшая, значит и квадрат нормы наименьший

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

$$\hat{y}_i = ax_i$$

Отклонения

$$e_i = y_i - ax_i$$

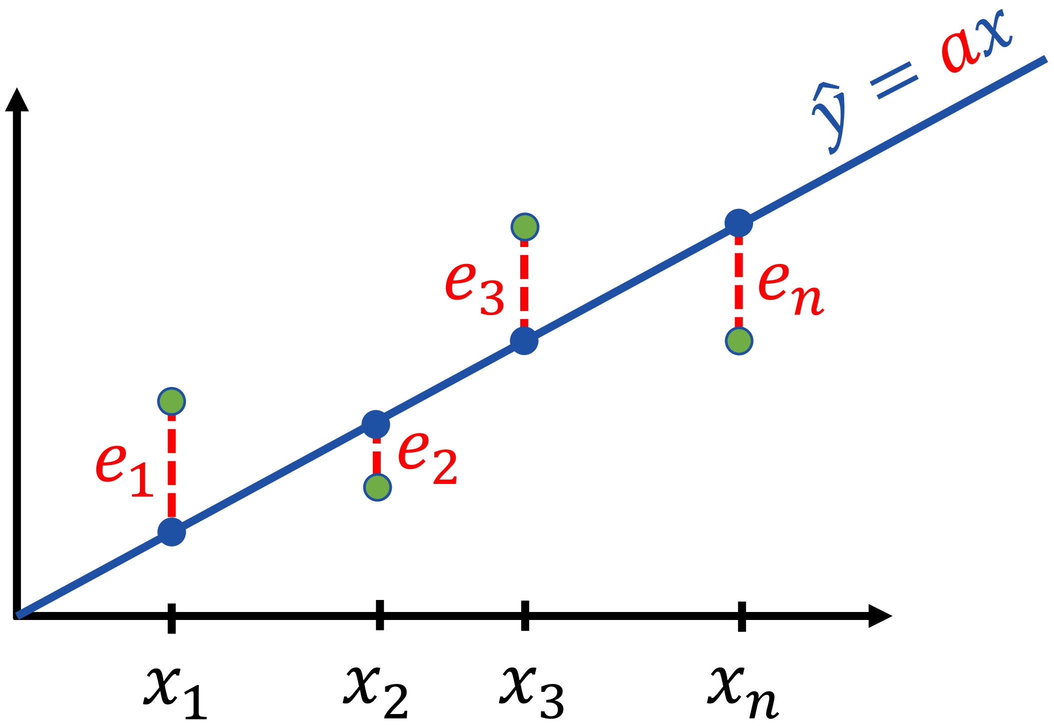
Квадрат нормы вектора отклонений

$$\|e\|^2 = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - ax_i)^2$$

Надо найти число a, при котором норма наименьшая

Но если норма наименьшая, значит и квадрат нормы наименьший

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

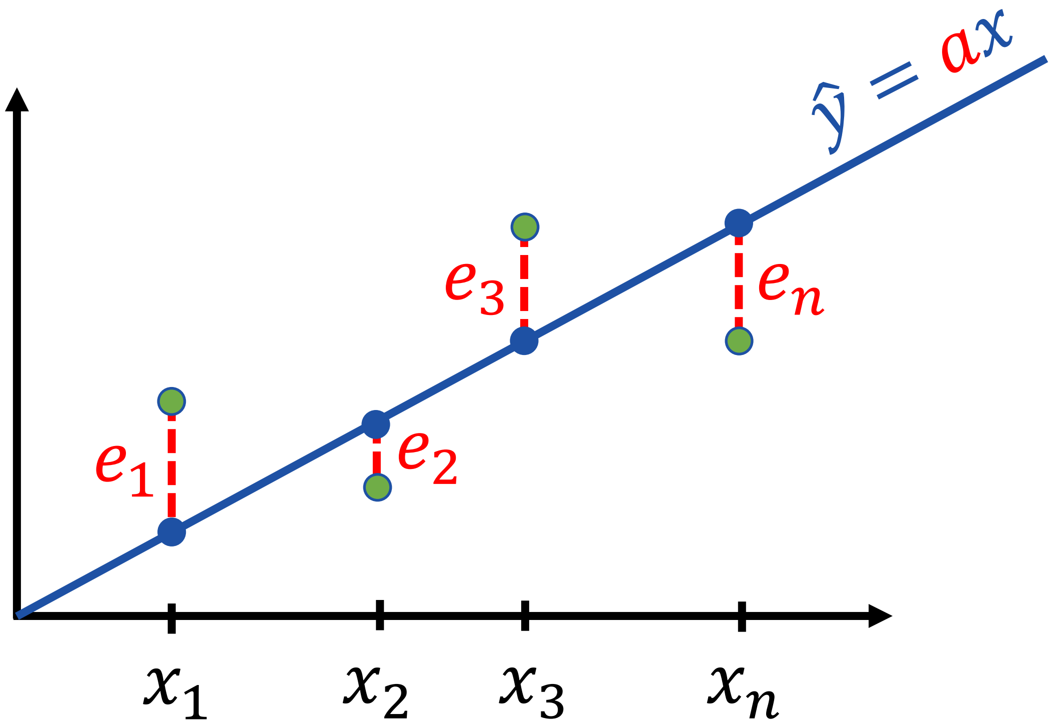
$$\hat{y}_i = ax_i$$

Отклонения

$$e_i = y_i - ax_i$$

$$\sum_{i=1}^{n} (y_i - ax_i)^2$$

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

$$\hat{y}_i = ax_i$$

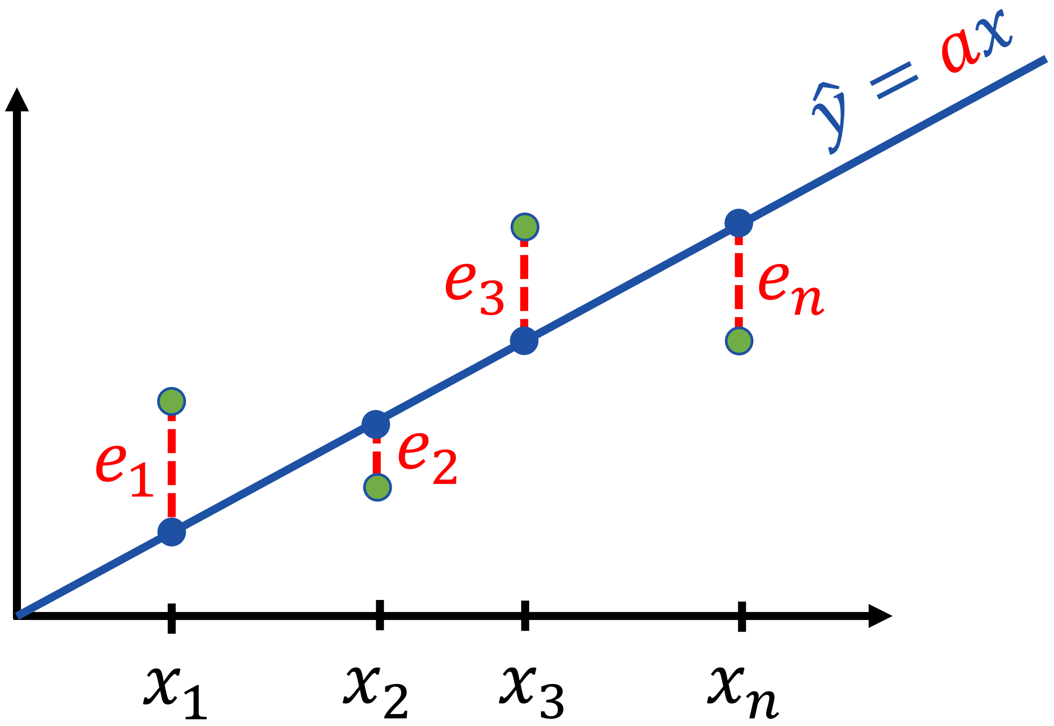
Отклонения

$$e_i = y_i - ax_i$$

Надо минимизировать это число

$$\sum_{i=1}^{n} (y_i - ax_i)^2$$

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

$$\hat{y}_i = ax_i$$

Отклонения

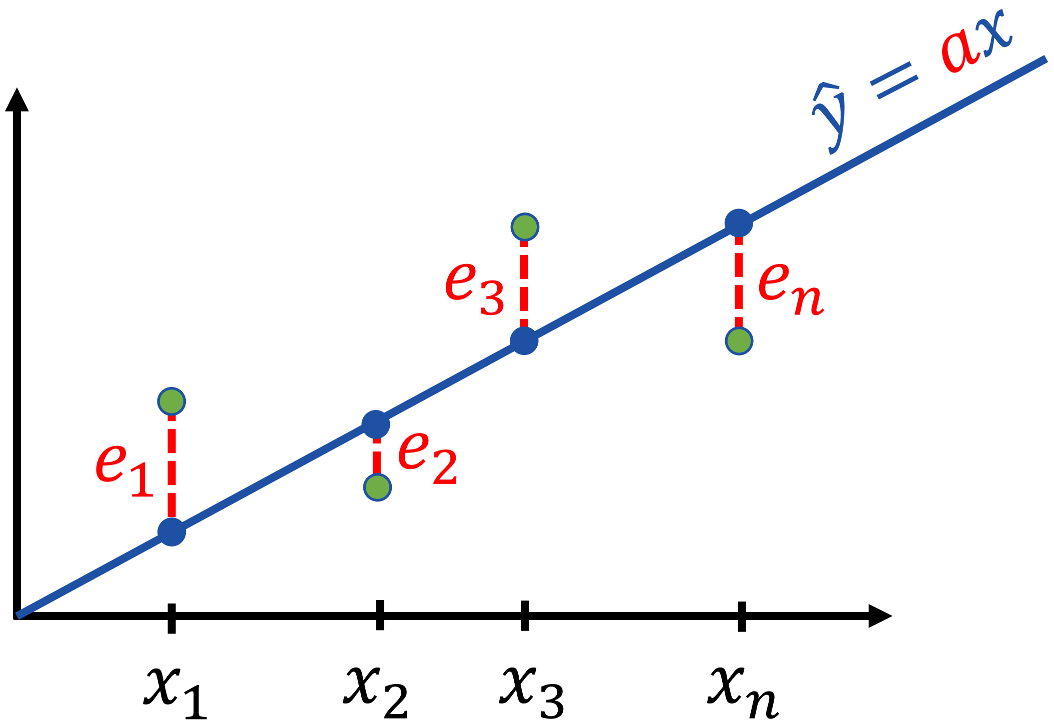
$$e_i = y_i - ax_i$$

Надо минимизировать это число

$$\sum_{i=1}^{n} (y_i - ax_i)^2$$

Оно является функцией от числа a

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

$$\hat{y}_i = ax_i$$

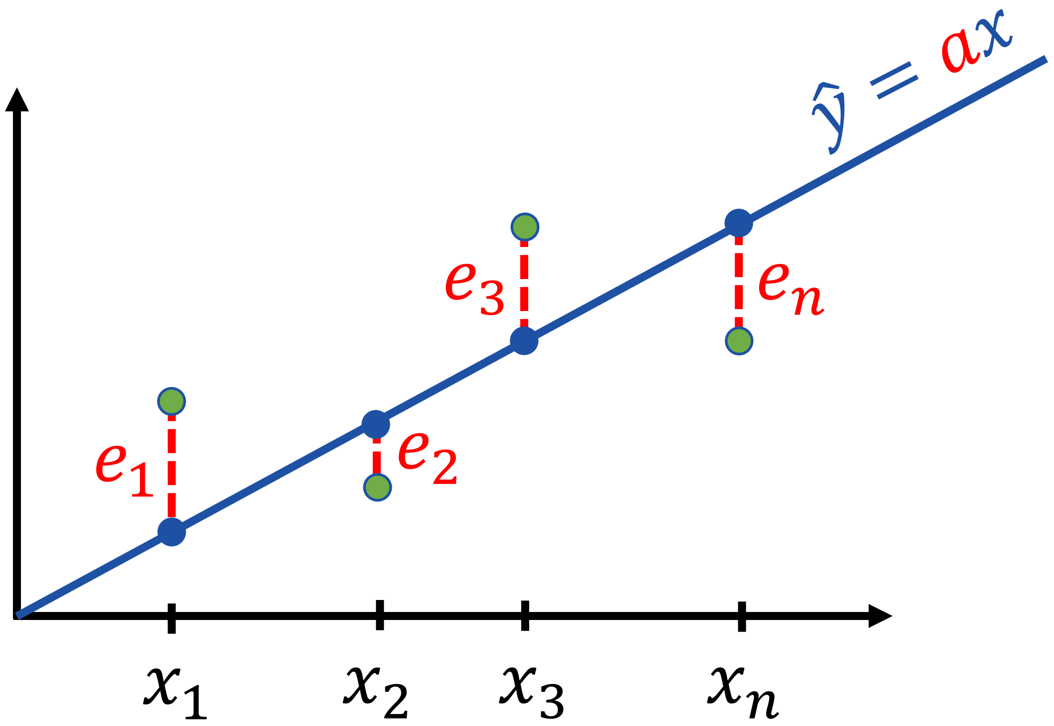
Отклонения

$$e_i = y_i - ax_i$$

$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2$$

Оно является функцией от числа a

МНК для $\hat{y} = ax$



Предположим, мы проведём прямую с каким-то коэффициентом a.

Тогда

$$\hat{y}_i = ax_i$$

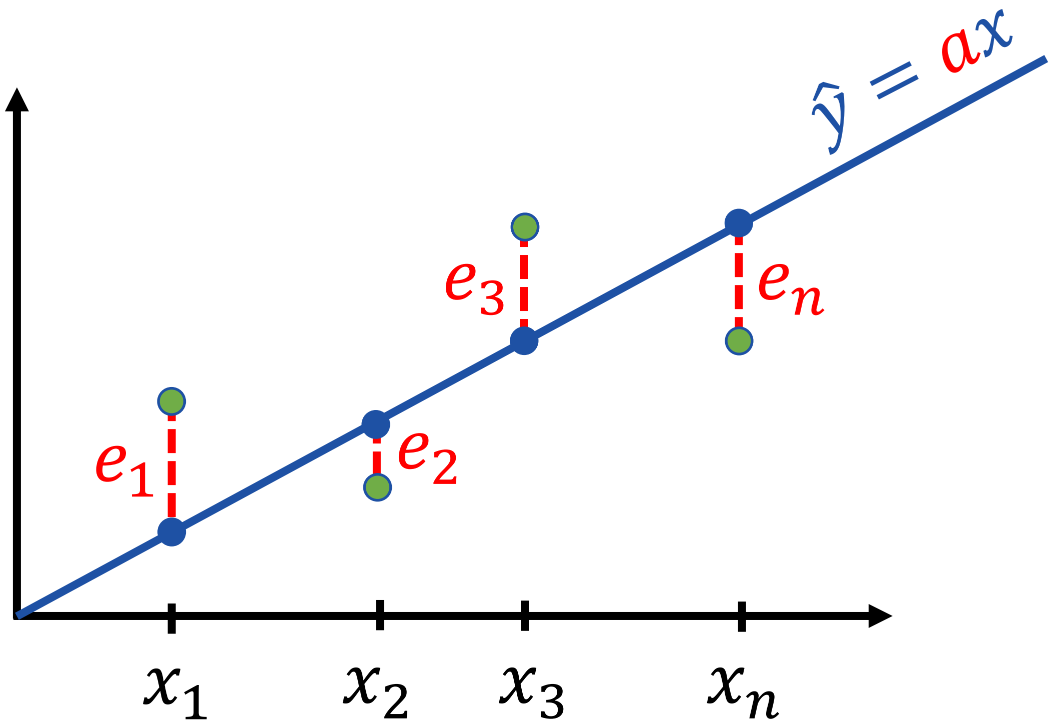
Отклонения

$$e_i = y_i - ax_i$$

Надо найти точку минимума этой функции

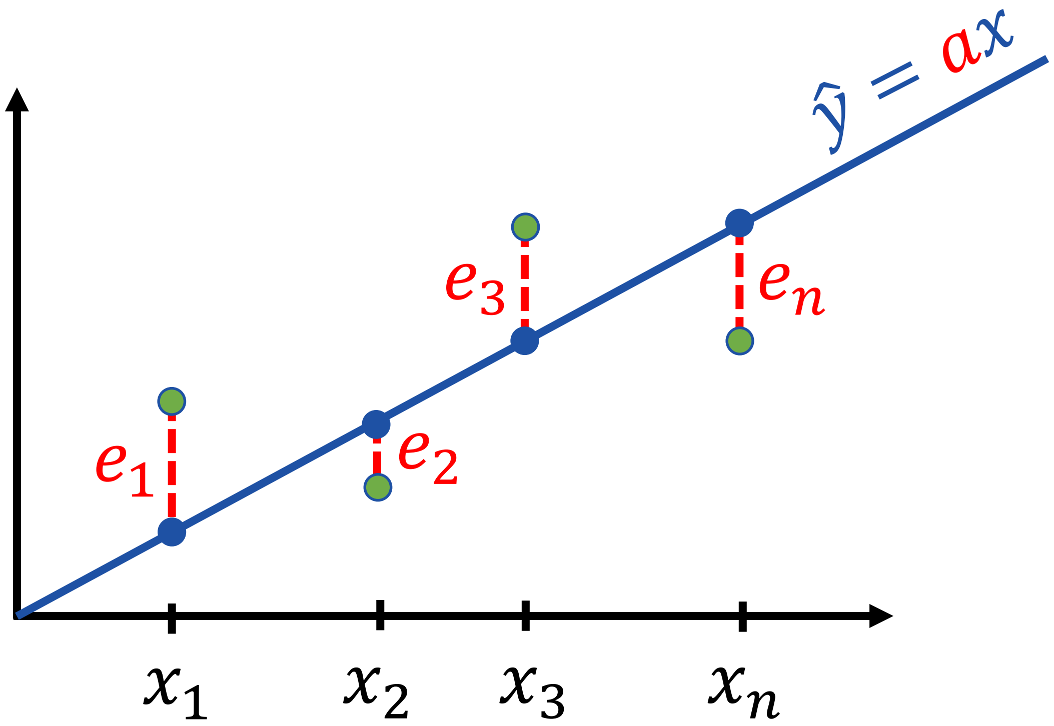
$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2$$

МНК для $\hat{y} = ax$



$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2 \to \min$$

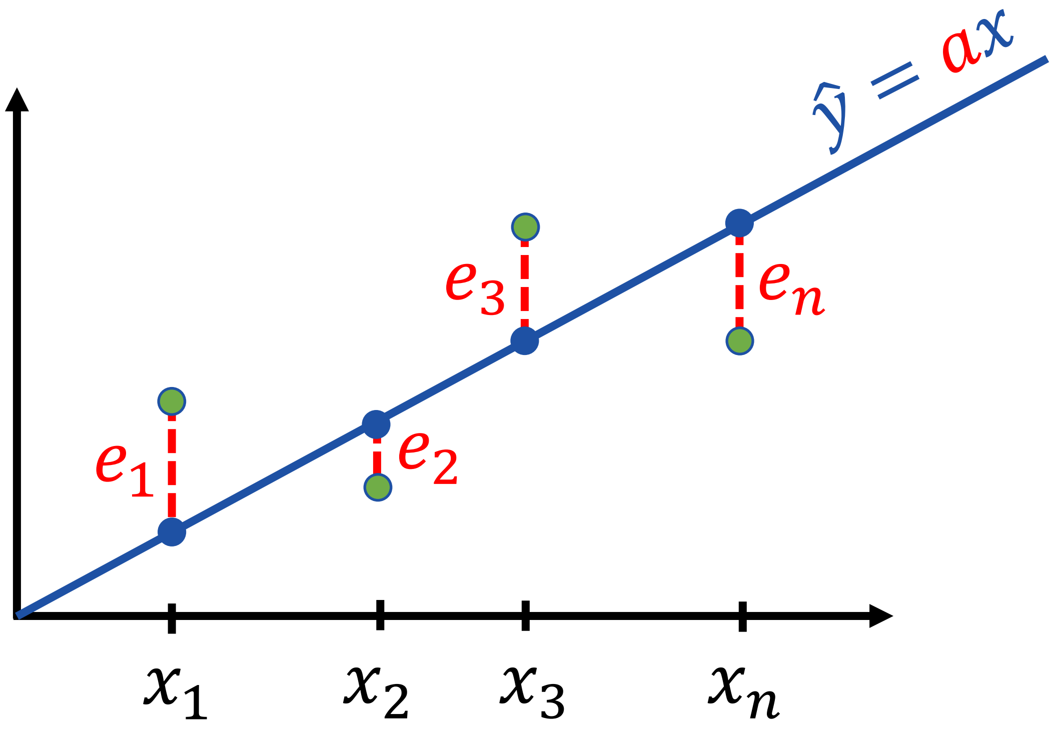
МНК для $\hat{y} = ax$



$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2 \rightarrow \min$$

Как найти точку минимума функции?

МНК для $\hat{y} = ax$

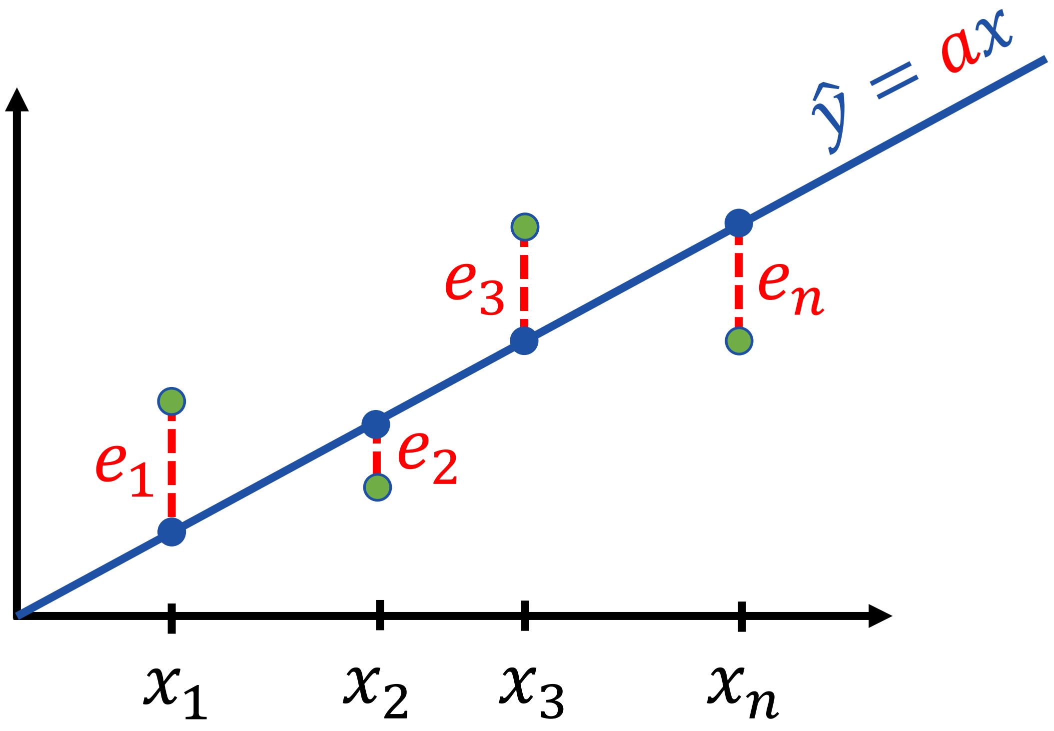


$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2 \to \min$$

Как найти точку минимума функции?

Взять производную и приравнять к нулю!

МНК для $\hat{y} = ax$



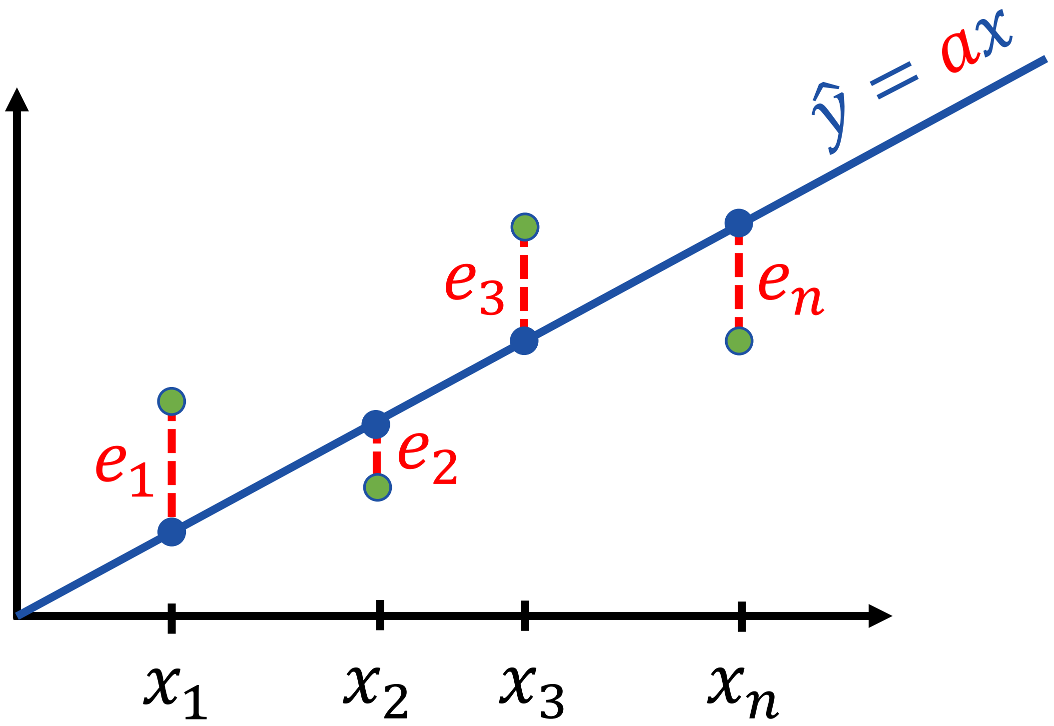
$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2 \to \min$$

Как найти точку минимума функции?

Взять производную и приравнять к нулю!

$$\frac{dS(a)}{da} = ?$$

МНК для $\hat{y} = ax$



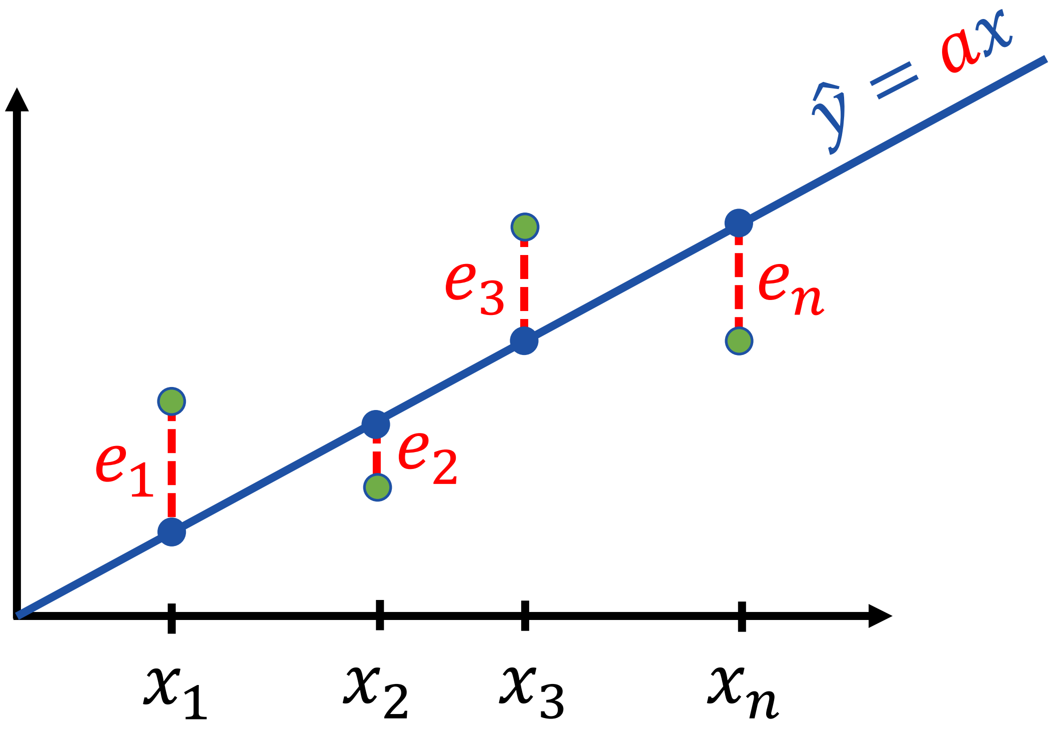
$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2 \ \to \min$$

Как найти точку минимума функции?

Взять производную и приравнять к нулю!

$$\frac{dS(a)}{da} = -2\sum_{i=1}^{n} x_i(y_i - ax_i)$$

МНК для $\hat{y} = ax$



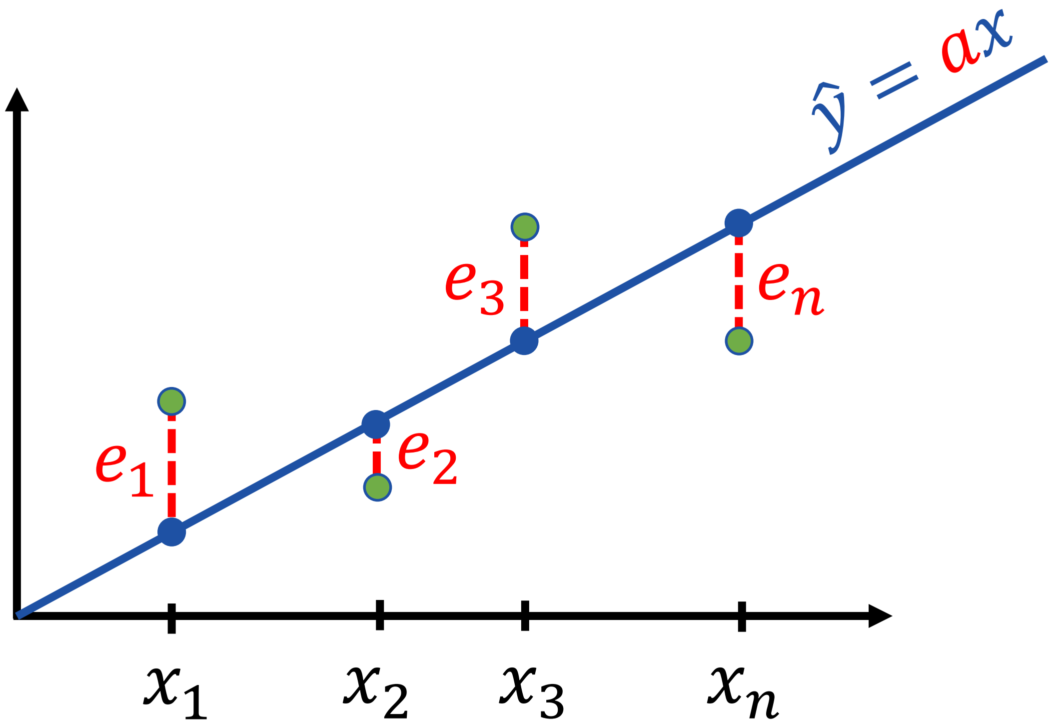
$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2 \;\to \min$$

Как найти точку минимума функции?

Взять производную и приравнять к нулю!

$$\frac{dS(a)}{da} = -2\sum_{i=1}^{n} x_i(y_i - ax_i) = 0$$

МНК для $\hat{y} = ax$



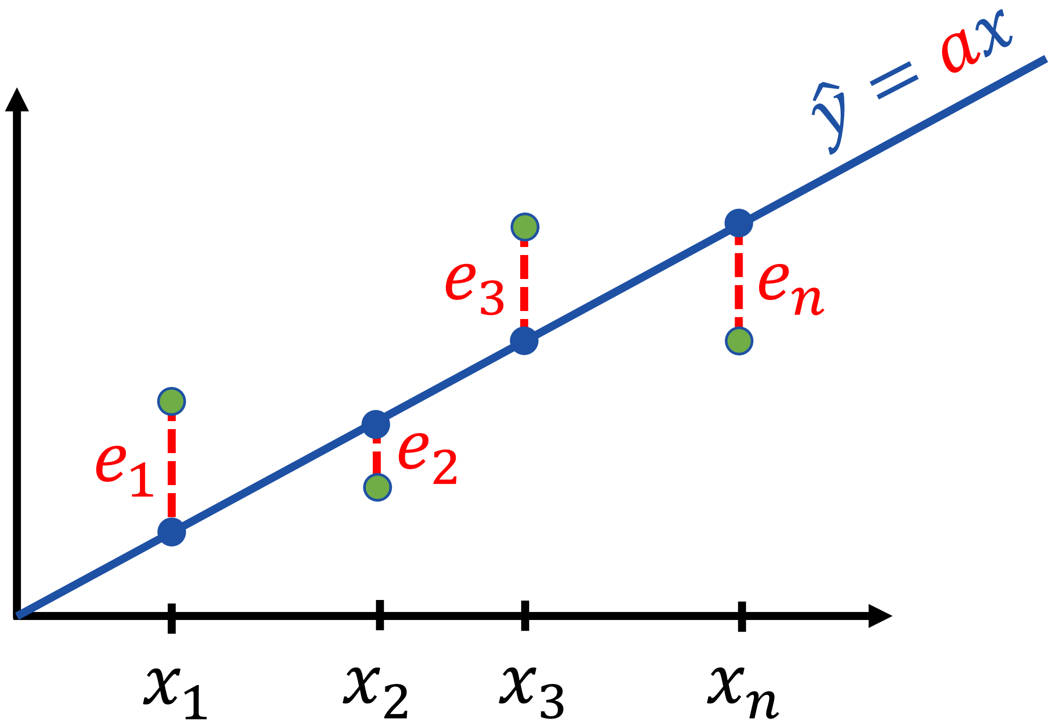
$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2 \to \min$$

Как найти точку минимума функции?

Взять производную и приравнять к нулю!

$$-2\sum_{i=1}^{n} x_i(y_i - ax_i) = 0$$

МНК для $\hat{y} = ax$



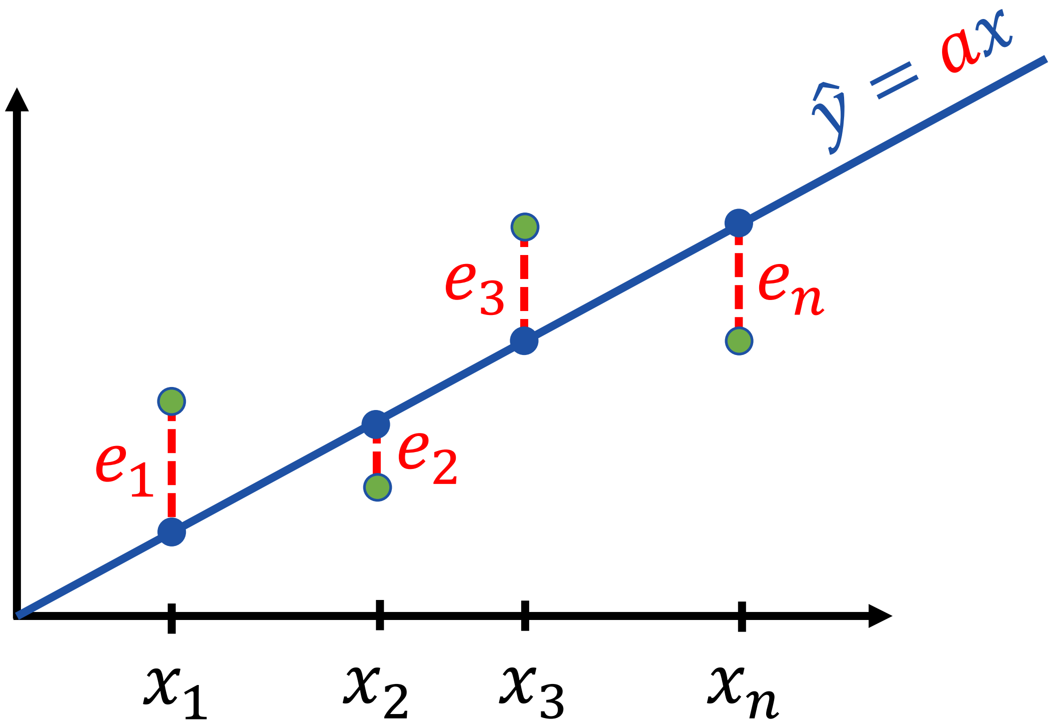
$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2 \rightarrow \min$$

Как найти точку минимума функции?

Взять производную и приравнять к нулю!

$$\sum_{i=1}^{n} x_i(y_i - ax_i) = 0$$

МНК для $\hat{y} = ax$



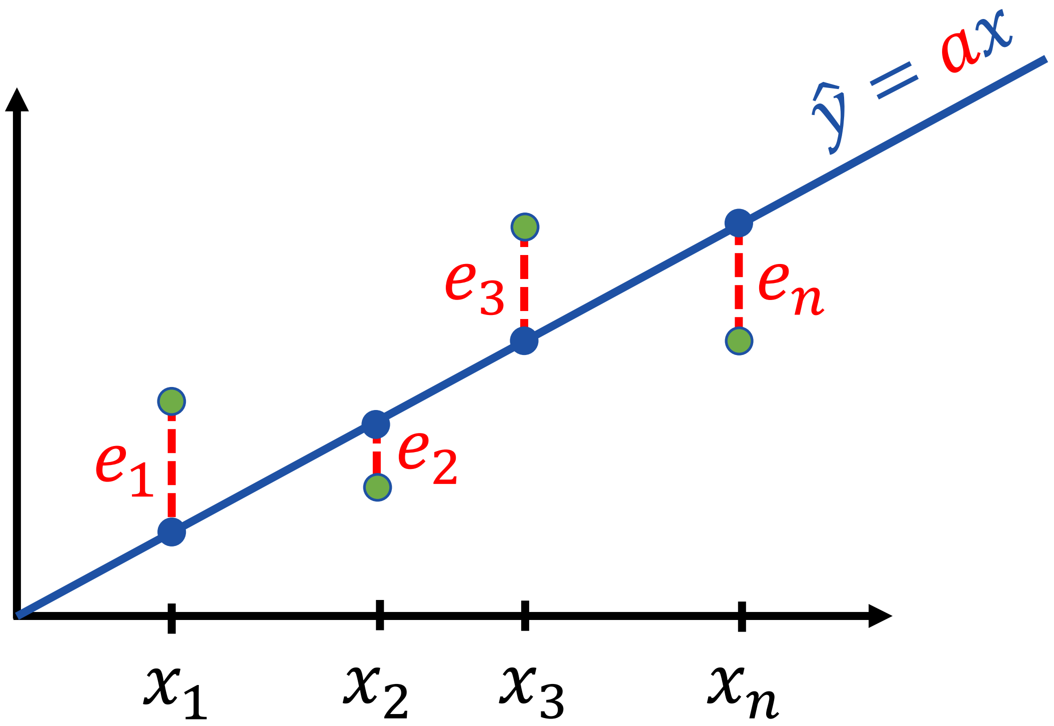
$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2 \to \min$$

Как найти точку минимума функции?

Взять производную и приравнять к нулю!

$$\sum_{i=1}^{n} x_i(y_i - ax_i) = 0 \quad \Rightarrow \quad \sum_{i=1}^{n} ax_i^2 = \sum_{i=1}^{n} x_i y_i$$

МНК для $\hat{y} = ax$



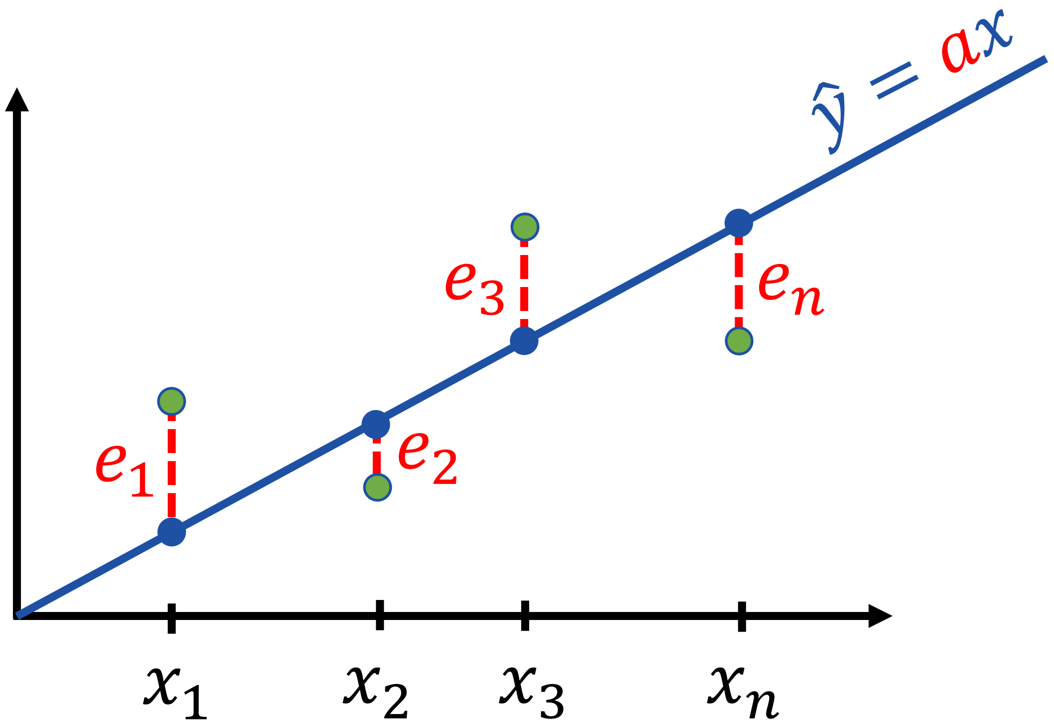
$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2 \to \min$$

Как найти точку минимума функции?

Взять производную и приравнять к нулю!

$$\sum_{i=1}^{n} x_i(y_i - ax_i) = 0 \quad \Rightarrow \quad a\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i$$

МНК для $\hat{y} = ax$



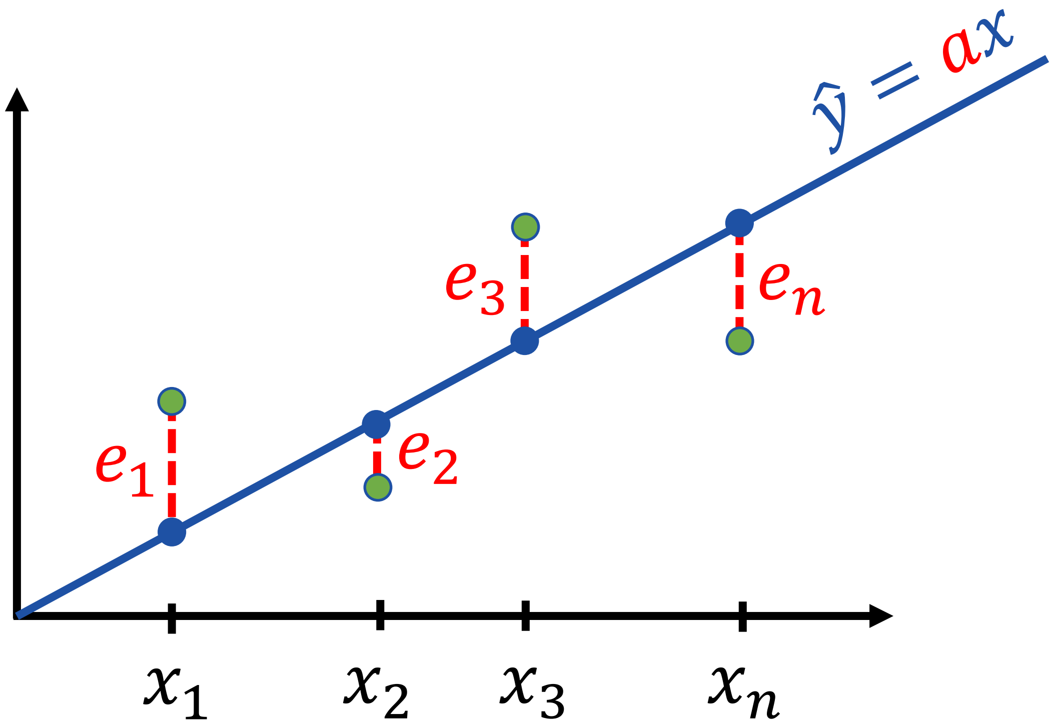
$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2 \to \min$$

Как найти точку минимума функции?

Взять производную и приравнять к нулю!

$$\sum_{i=1}^{n} x_i(y_i - ax_i) = 0 \quad \Rightarrow \quad a\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \quad \Rightarrow \quad a = \frac{\sum x_i y_i}{\sum x_i^2}$$

МНК для $\hat{y} = ax$

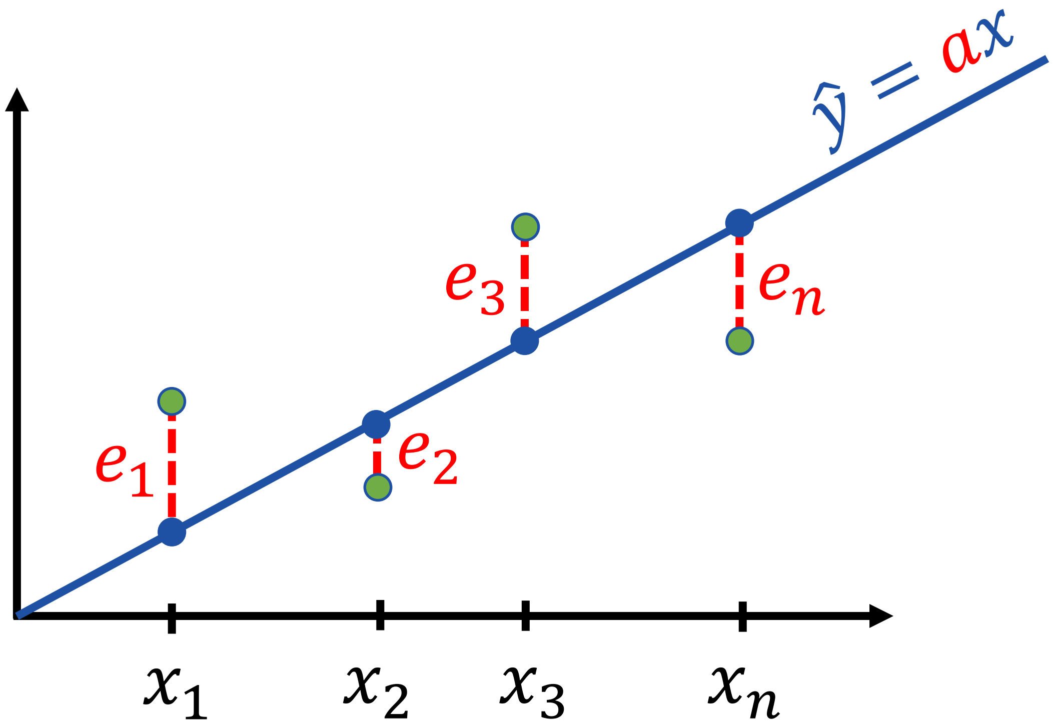


$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2 \to \min$$

Точка минимума функции $S(a)$

$$a = \frac{\sum x_i y_i}{\sum x_i^2}$$

МНК для $\hat{y} = ax$



$$S(a) = \sum_{i=1}^{n} (y_i - ax_i)^2 \to \min$$

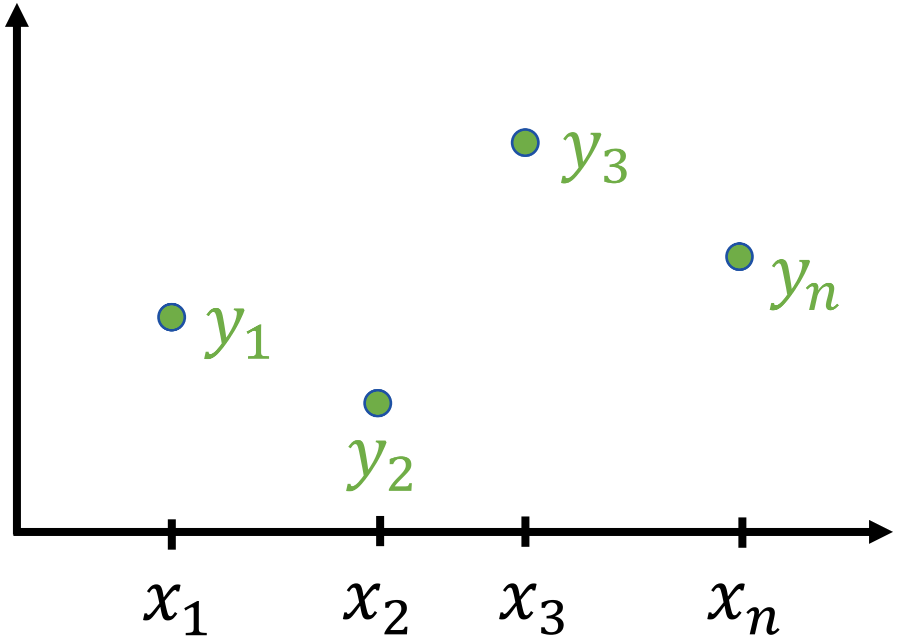
Точка минимума функции $S(a)$

$$a = \frac{\sum x_i y_i}{\sum x_i^2}$$

При этом значении a вектор отклонений будет иметь наименьшую l_2-норму

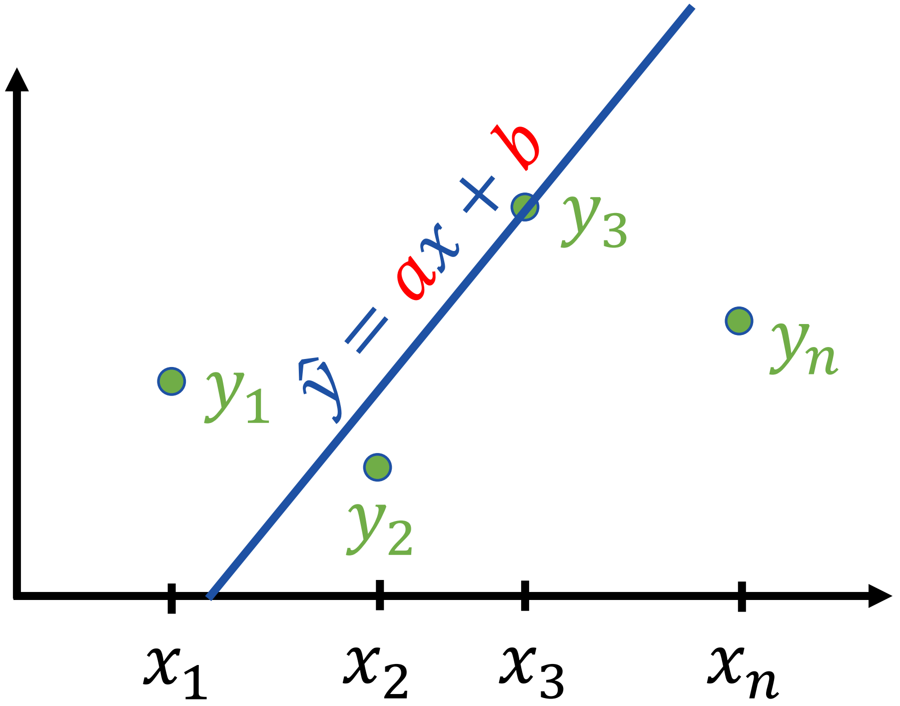
МНК для $\hat{y} = ax + b$

МНК для $\hat{y} = ax + b$



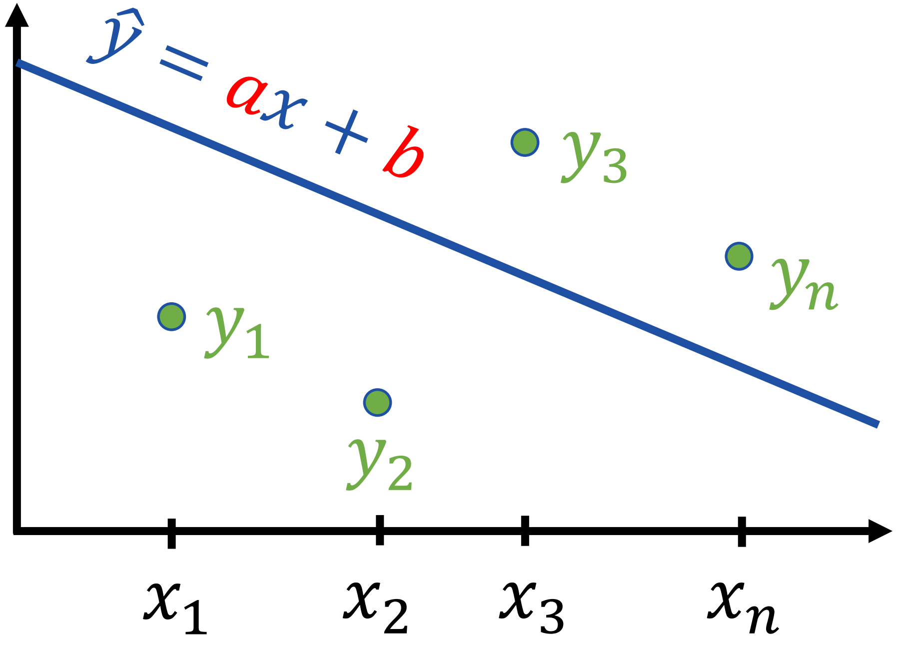
Даны точки, которые надо аппроксимировать прямой $\hat{y} = ax + b$

МНК для $\hat{y} = ax + b$



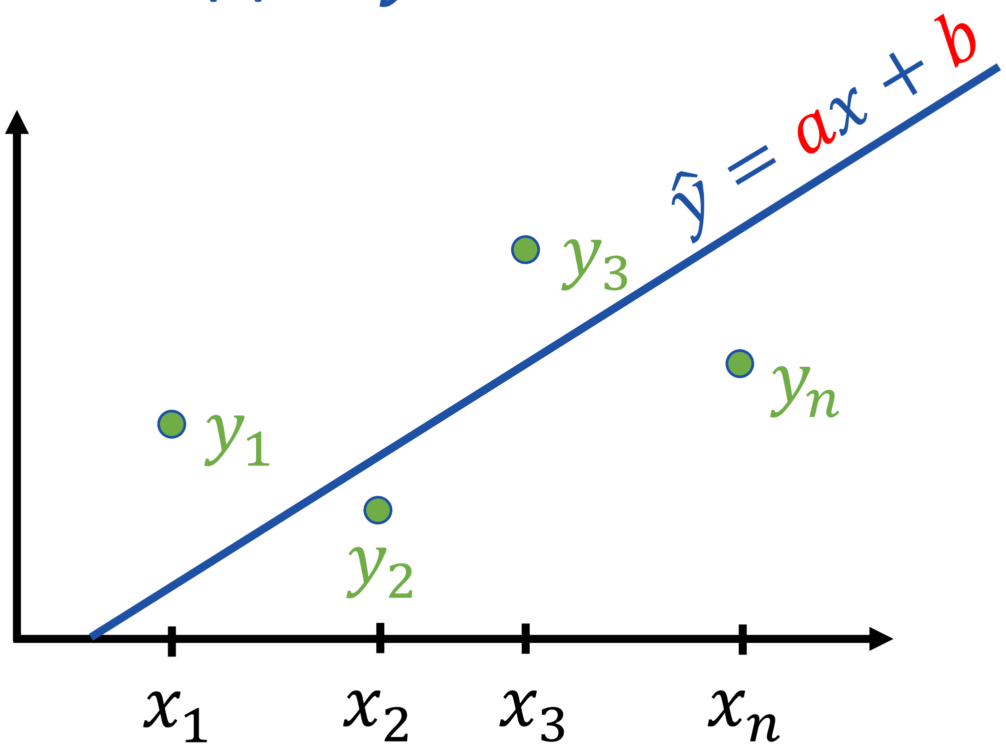
Предположим, мы проведём прямую с какими-то коэффициентами a и b

МНК для $\hat{y} = ax + b$



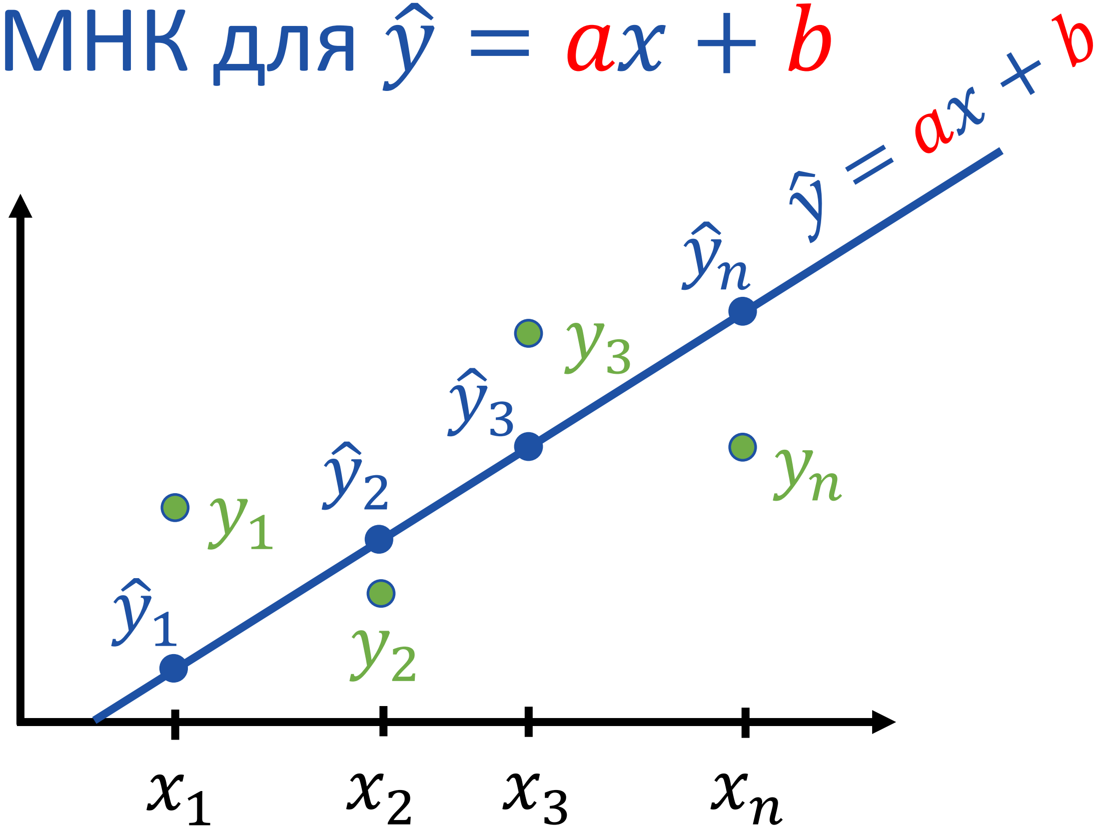
Предположим, мы проведём прямую с какими-то коэффициентами a и b

МНК для $\hat{y} = ax + b$



Предположим, мы проведём прямую с какими-то коэффициентами a и b

МНК для $\hat{y} = ax + b$

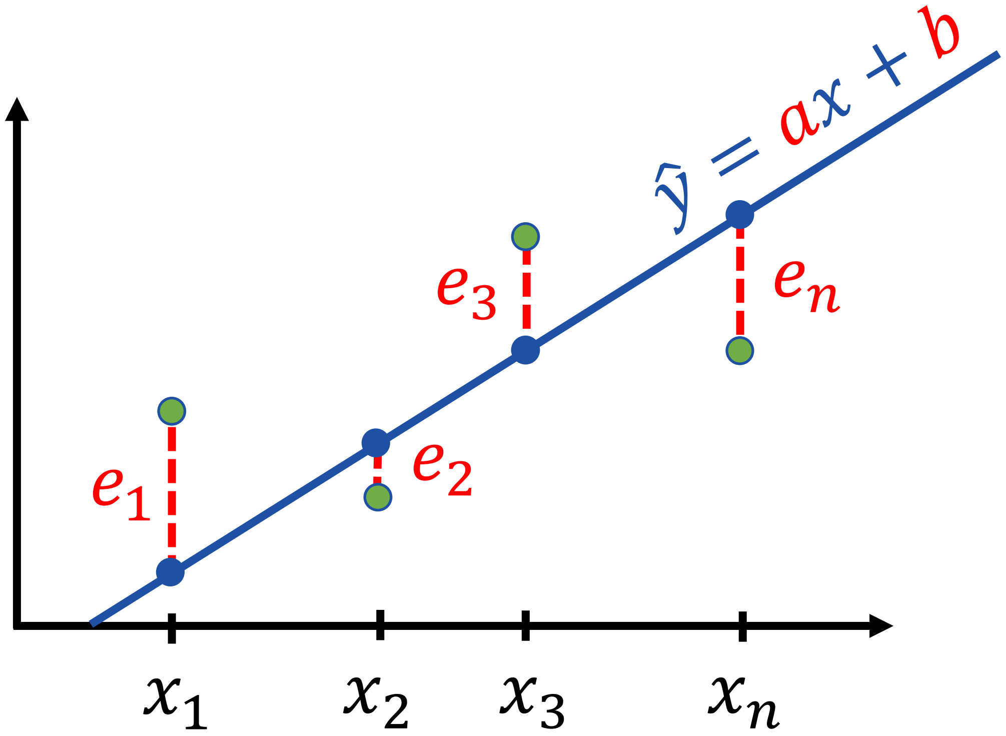


Предположим, мы проведём прямую с какими-то коэффициентами a и b

Тогда

$$\hat{y}_i = ax_i + b$$

МНК для $\hat{y} = ax + b$



Предположим, мы проведём прямую с какими-то коэффициентами a и b

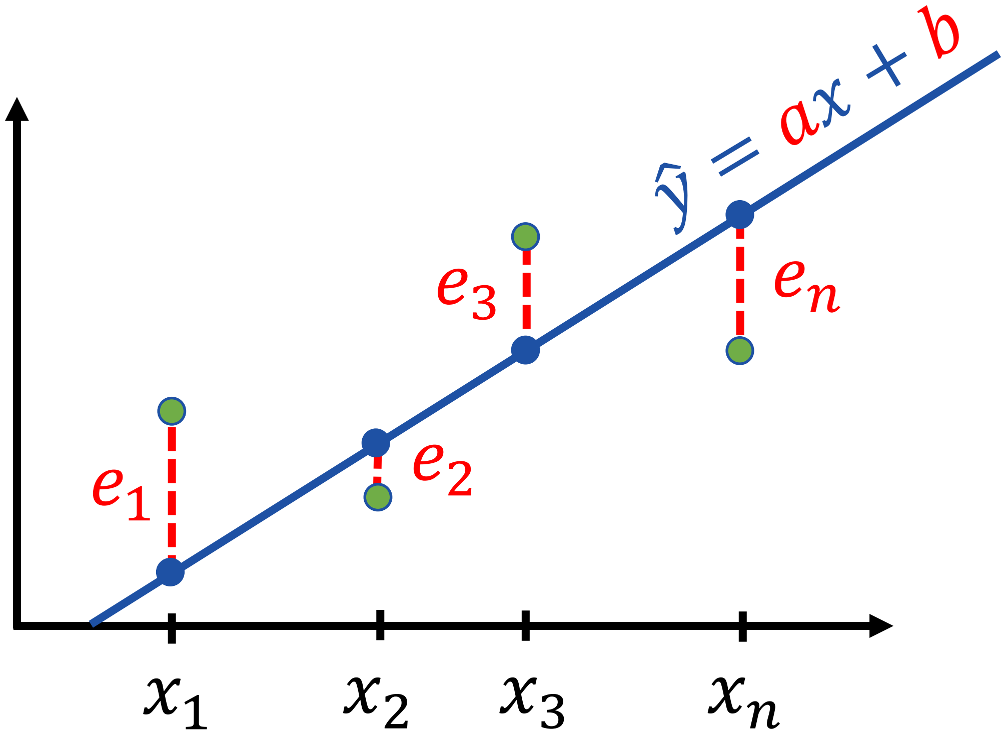
Тогда

$$\hat{y}_i = ax_i + b$$

Отклонения

$$e_i = y_i - (ax_i + b)$$

МНК для $\hat{y} = ax + b$



Предположим, мы проведём прямую с какими-то коэффициентами a и b

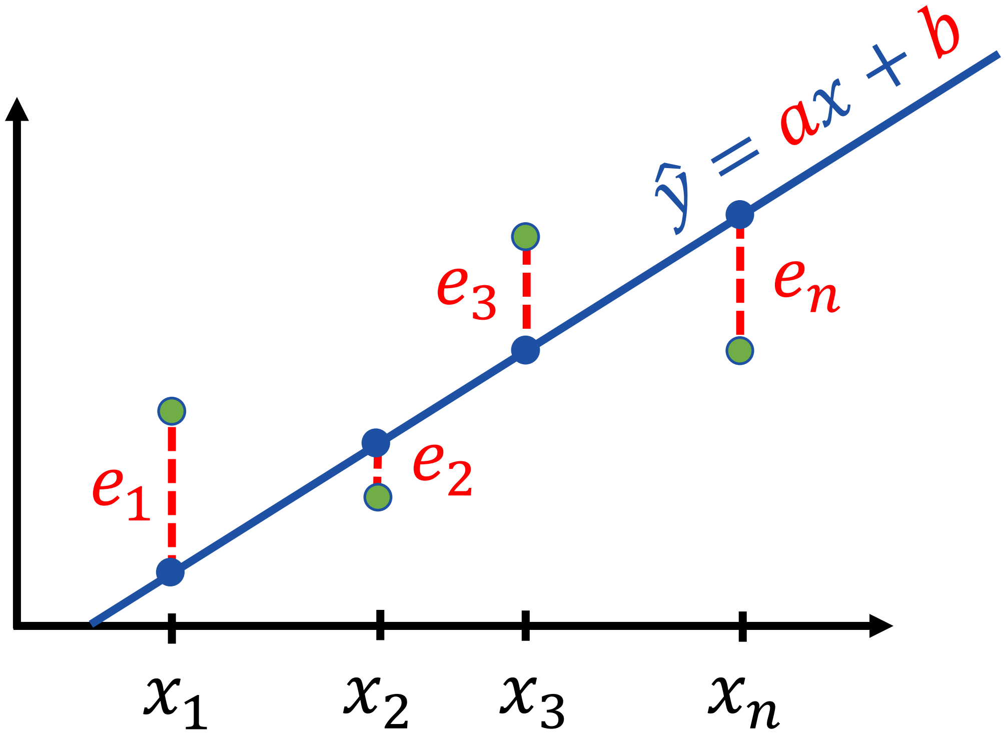
Тогда

$$\hat{y}_i = ax_i + b$$

Отклонения

$$e_i = y_i - ax_i - b$$

МНК для $\hat{y} = ax + b$



Предположим, мы проведём прямую с какими-то коэффициентами a и b

Тогда

$$\hat{y}_i = ax_i + b$$

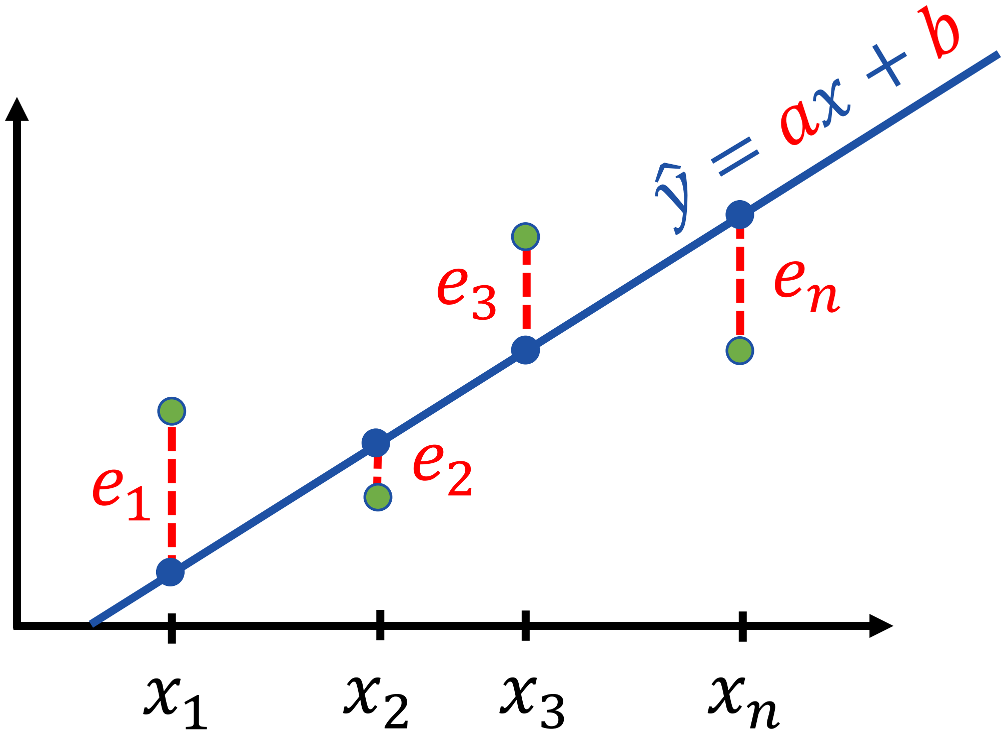
Отклонения

$$e_i = y_i - ax_i - b$$

Норма вектора отклонений

$$\|e\| = \sqrt{\sum_{i=1}^{n} (y_i - ax_i - b)^2}$$

МНК для $\hat{y} = ax + b$



Предположим, мы проведём прямую с какими-то коэффициентами a и b

Тогда

$$\hat{y}_i = ax_i + b$$

Отклонения

$$e_i = y_i - ax_i - b$$

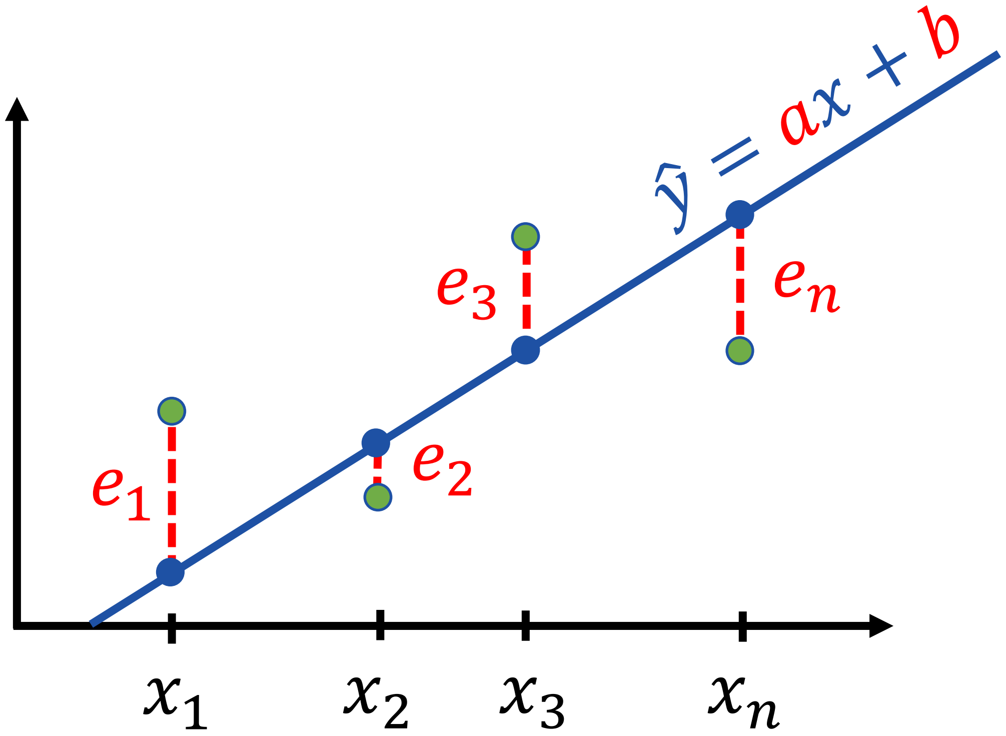
Норма вектора отклонений

$$\|e\| = \sqrt{\sum_{i=1}^{n} (y_i - ax_i - b)^2}$$

Квадрат нормы вектора отклонений

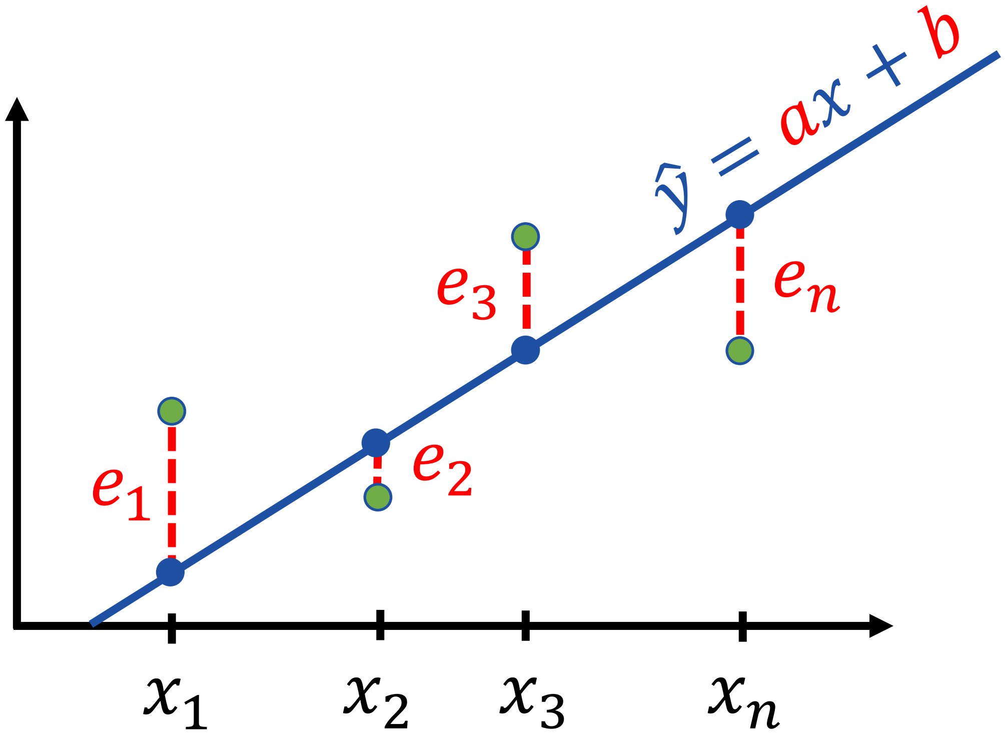
$$S(a, b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

МНК для $\hat{y} = ax + b$



$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \to \min$$

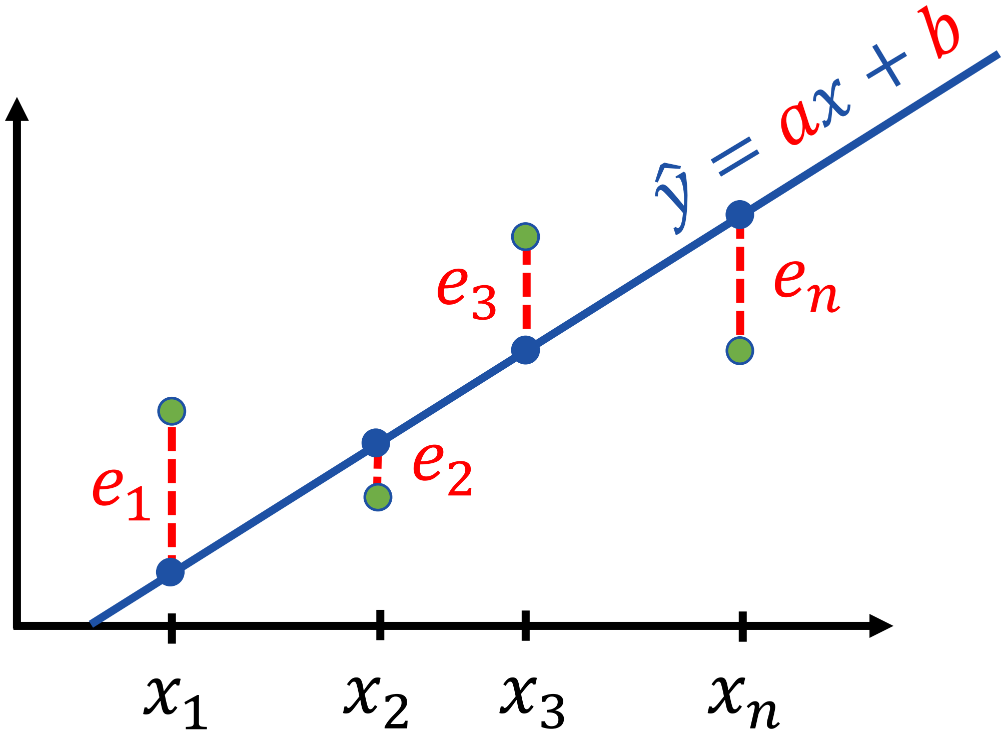
МНК для $\hat{y} = ax + b$



$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \to \min$$

Как найти точку минимума функции двух аргументов?

МНК для $\hat{y} = ax + b$

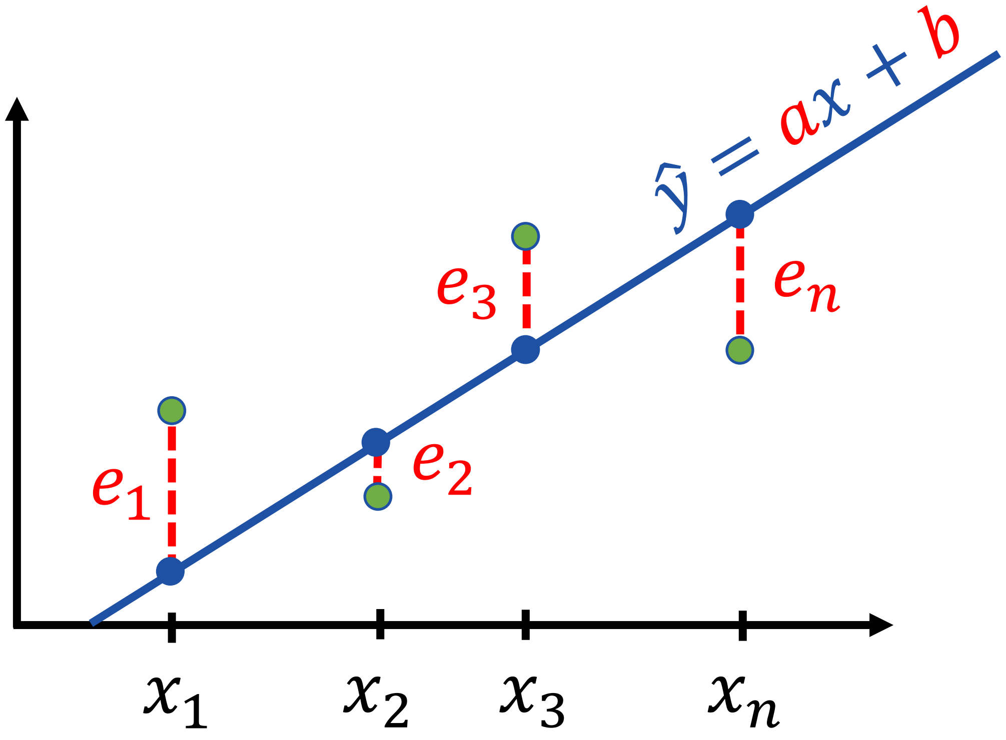


$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \rightarrow \min$$

Как найти точку минимума функции двух аргументов?

Взять две частные производные и приравнять к нулю!

МНК для $\hat{y} = ax + b$



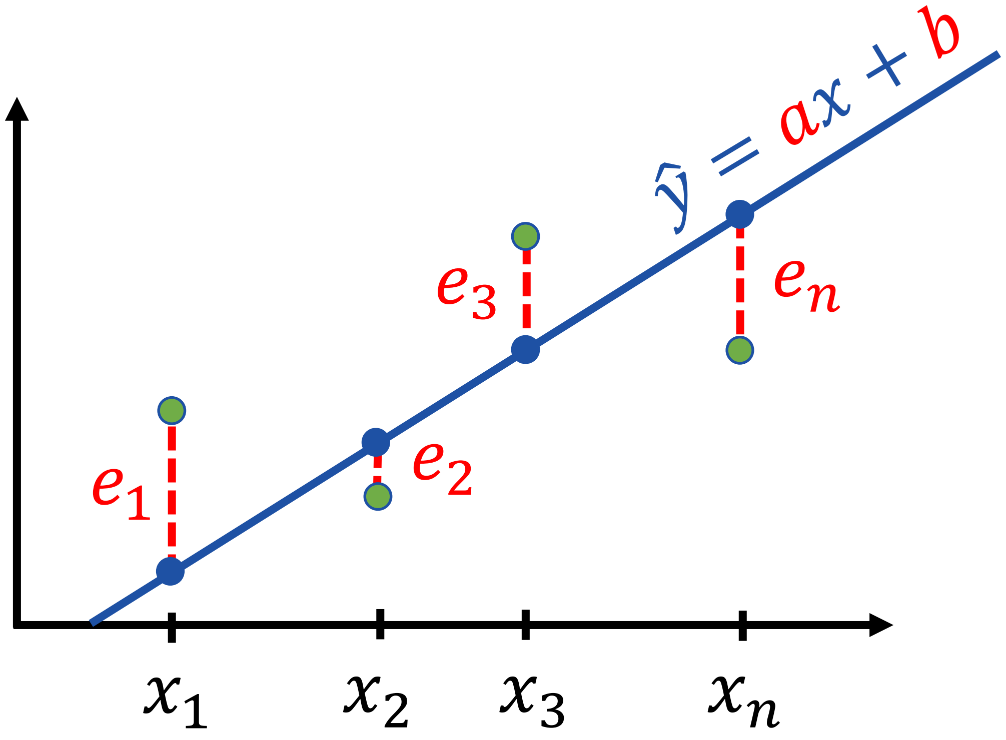
$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \to \min$$

Как найти точку минимума функции двух аргументов?

Взять две частные производные и приравнять к нулю!

$$\frac{\partial S(a,b)}{\partial a} = 0 \qquad \frac{\partial S(a,b)}{\partial b} = 0$$

МНК для $\hat{y} = ax + b$



$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \to \min$$

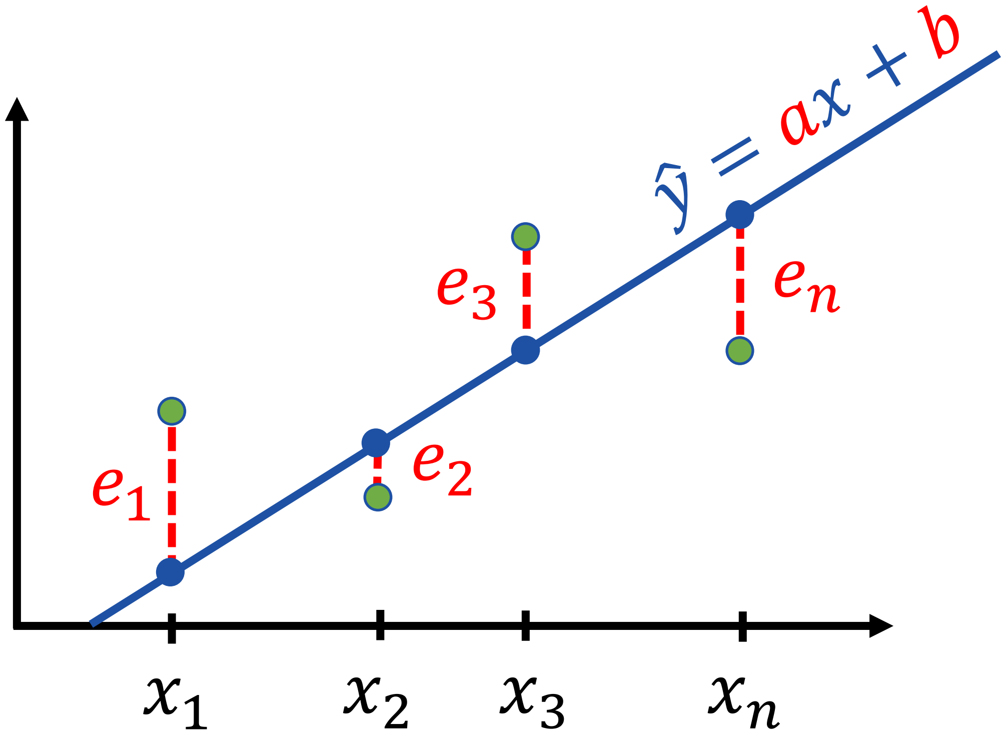
Как найти точку минимума функции двух аргументов?

Взять две частные производные и приравнять к нулю!

$$\frac{\partial S(a,b)}{\partial a} = \ ?$$

$$\frac{\partial S(a,b)}{\partial b} = \ ?$$

МНК для $\hat{y} = ax + b$



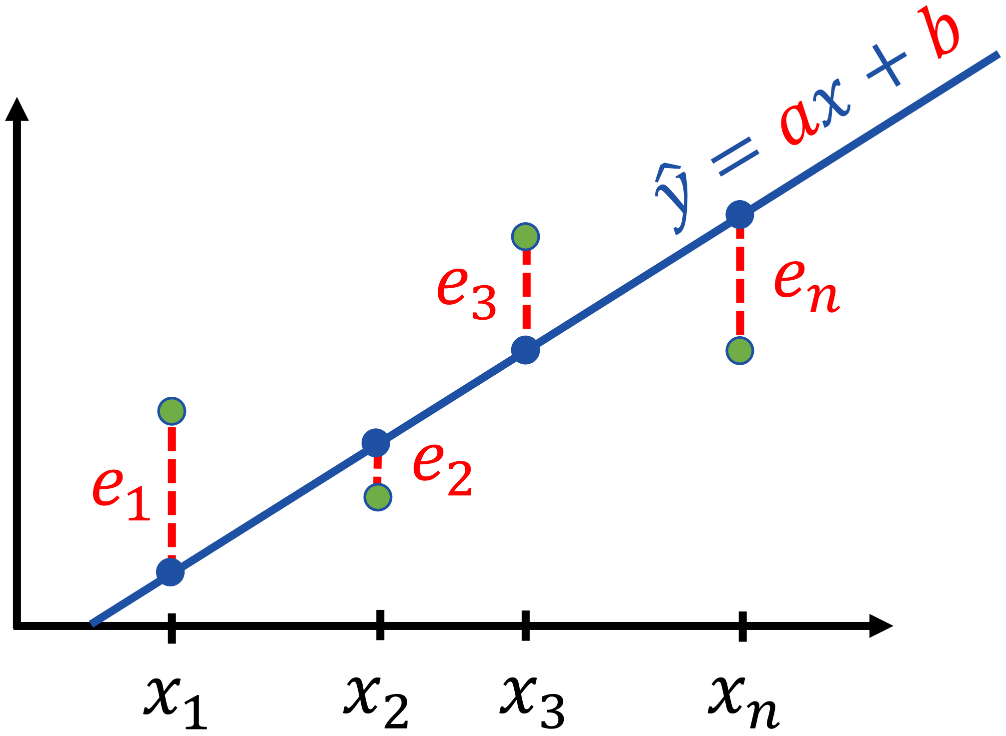
$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \;\to \min$$

Как найти точку минимума функции двух аргументов?

Взять две частные производные и приравнять к нулю!

$$\frac{\partial S(a,b)}{\partial a} = -2\sum_{i=1}^{n} x_i(y_i - ax_i - b) \qquad \frac{\partial S(a,b)}{\partial b} = \;?$$

МНК для $\hat{y} = ax + b$



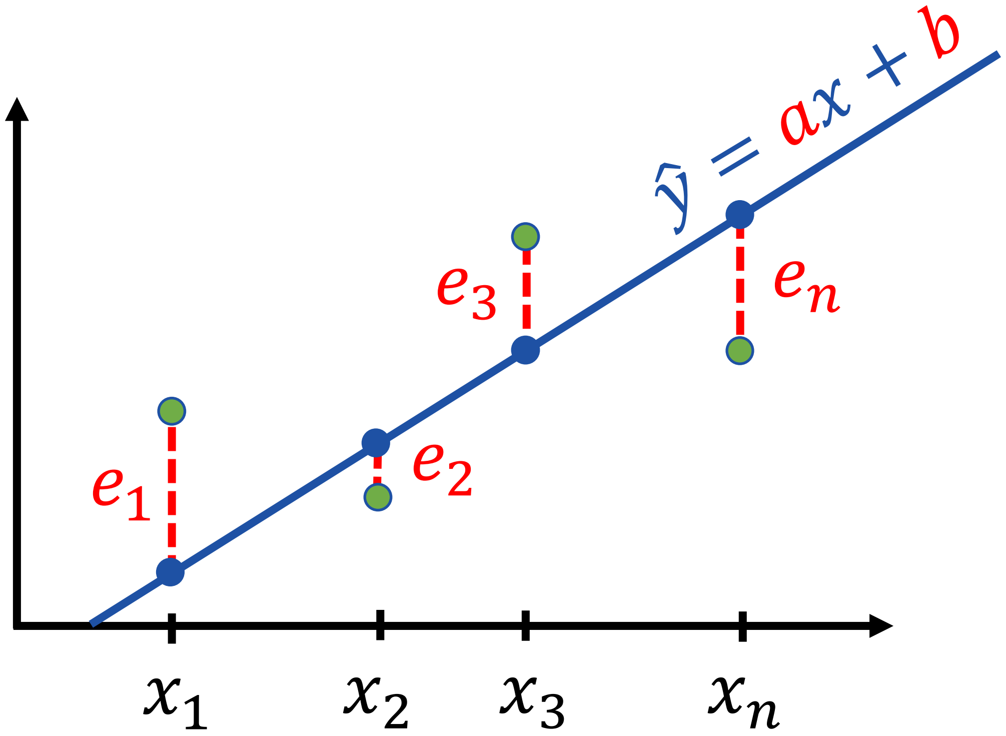
$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \ \to \min$$

Как найти точку минимума функции двух аргументов?

Взять две частные производные и приравнять к нулю!

$$\frac{\partial S(a,b)}{\partial a} = -2\sum_{i=1}^{n} x_i(y_i - ax_i - b) \qquad \frac{\partial S(a,b)}{\partial b} = -2\sum_{i=1}^{n} (y_i - ax_i - b)$$

МНК для $\hat{y} = ax + b$



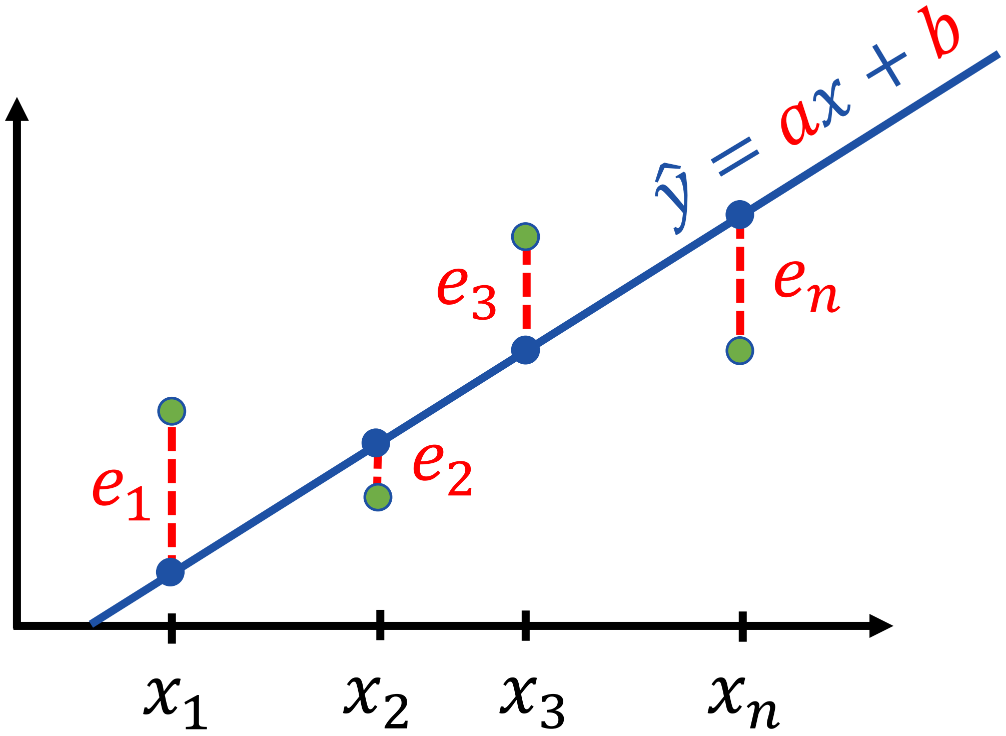
$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \;\to \min$$

Взять две частные производные и приравнять к нулю!

$$-2\sum_{i=1}^{n} x_i(y_i - ax_i - b) = 0$$

$$-2\sum_{i=1}^{n} (y_i - ax_i - b) = 0$$

МНК для $\hat{y} = ax + b$



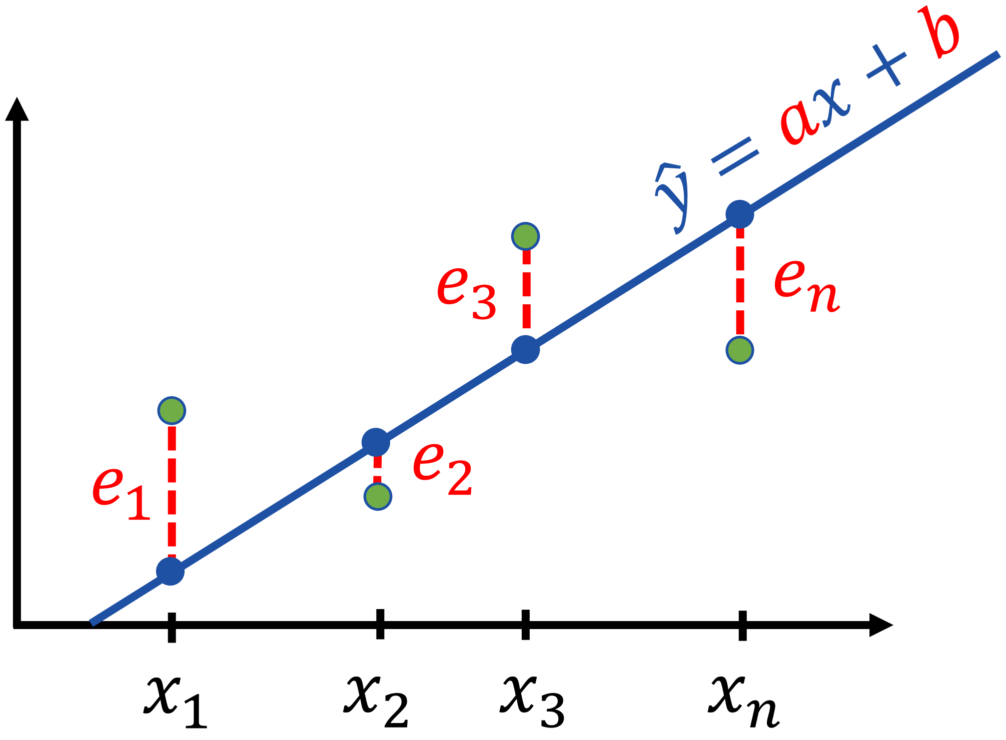
$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \to \min$$

Взять две частные производные и приравнять к нулю!

$$\sum_{i=1}^{n} x_i(y_i - ax_i - b) = 0$$

$$\sum_{i=1}^{n} (y_i - ax_i - b) = 0$$

МНК для $\hat{y} = ax + b$



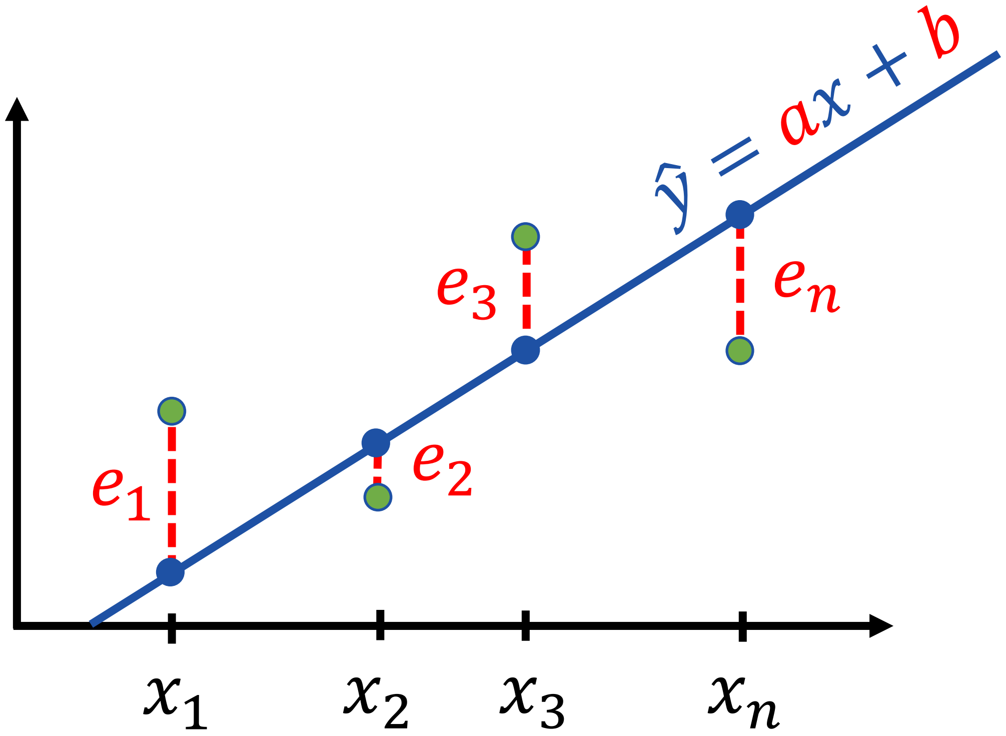
$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \rightarrow \min$$

Взять две частные производные и приравнять к нулю!

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i$$

$$\sum_{i=1}^{n} (y_i - ax_i - b) = 0$$

МНК для $\hat{y} = ax + b$



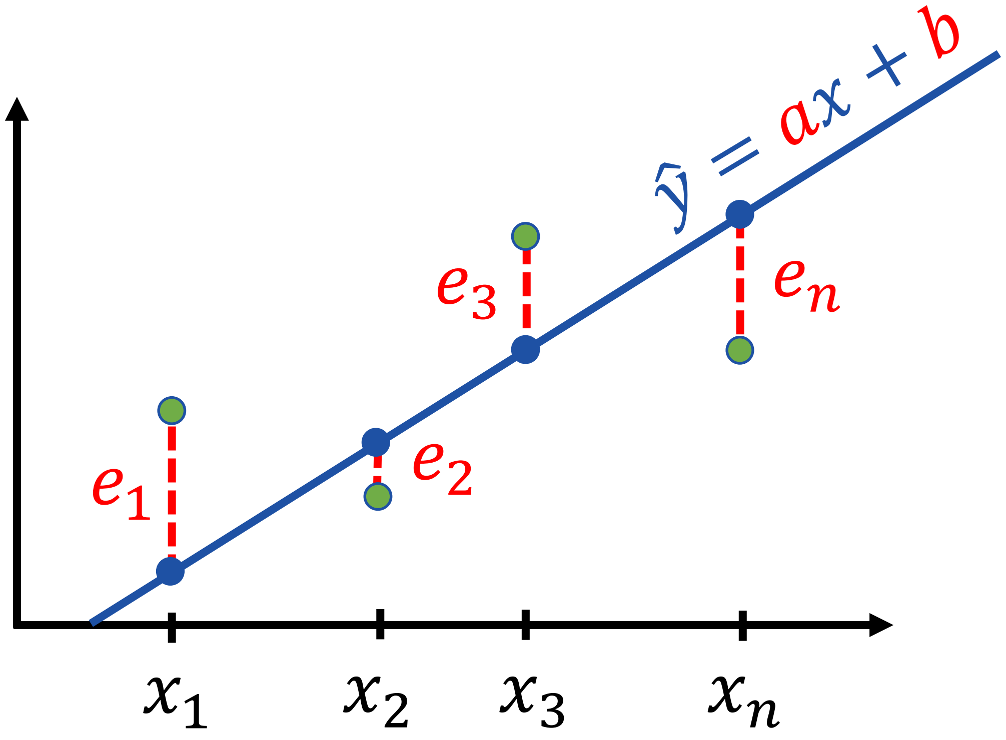
$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \to \min$$

Взять две частные производные и приравнять к нулю!

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i$$

$$a\sum_{i=1}^{n} x_i + bn = \sum_{i=1}^{n} y_i$$

МНК для $\hat{y} = ax + b$



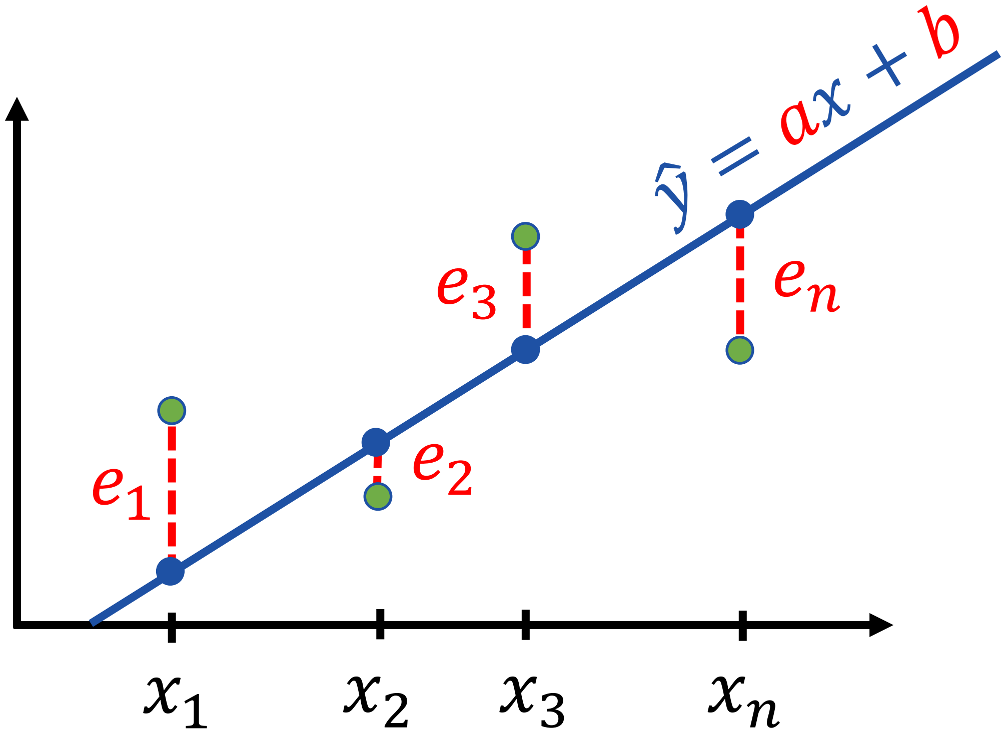
$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \to \min$$

Взять две частные производные и приравнять к нулю!

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i$$

$$a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} 1 = \sum_{i=1}^{n} y_i$$

МНК для $\hat{y} = ax + b$



$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \to \min$$

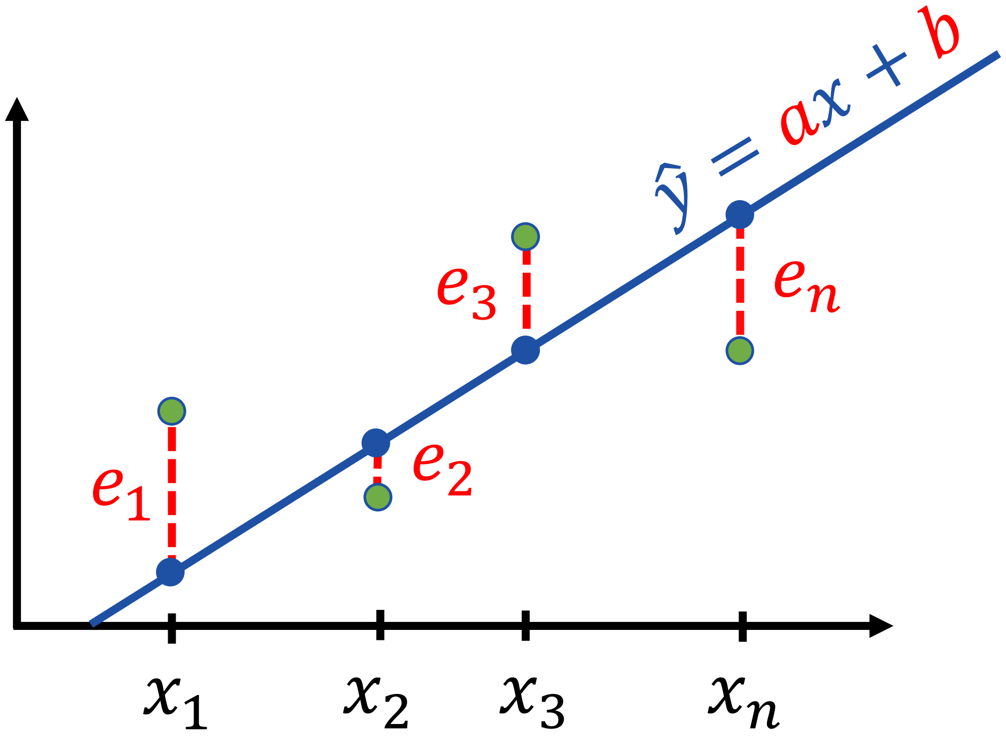
Взять две частные производные и приравнять к нулю!

Система из двух уравнений с двумя неизвестными

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i$$

$$a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} 1 = \sum_{i=1}^{n} y_i$$

МНК для $\hat{y} = ax + b$



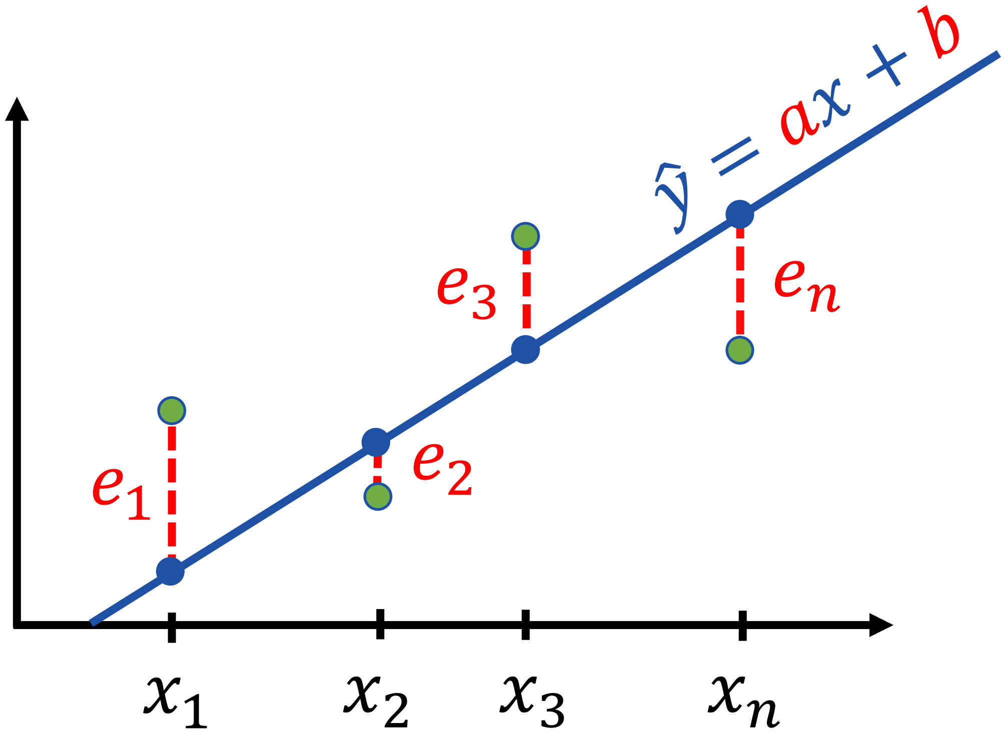
$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \rightarrow \min$$

Взять две частные производные и приравнять к нулю!

Матричное уравнение
с неизвестным вектором

$$\begin{bmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & \sum 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum x_i y_i \\ \sum y_i \end{bmatrix}$$

МНК для $\hat{y} = ax + b$



$$S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \to \min$$

Взять две частные производные и приравнять к нулю!

Матричное уравнение с неизвестным вектором

$$\begin{bmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & \sum 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum x_i y_i \\ \sum y_i \end{bmatrix}$$

Можно написать явную формулу решения, но мы не будем этого делать

А если много параметров?

А если много параметров?

$$\hat{y} = f(x, a_1, a_2, \dots, a_m)$$

А если много параметров?

Общий вид аппроксимирующей функции

$$\hat{y} = f(x, a_1, a_2, \dots, a_m)$$

А если много параметров?

Общий вид аппроксимирующей функции

$$\hat{y} = f(x, a_1, a_2, \dots, a_m)$$

Квадрат нормы вектора отклонений

$$S(a_1, a_2, \dots, a_m) = \sum_{i=1}^{n} (y_i - f(x_i, a_1, a_2, \dots, a_m))^2$$

А если много параметров?

Общий вид аппроксимирующей функции

$$\hat{y} = f(x, a_1, a_2, \dots, a_m)$$

Квадрат нормы вектора отклонений

$$S(a_1, a_2, \dots, a_m) = \sum_{i=1}^{n} (y_i - f(x_i, a_1, a_2, \dots, a_m))^2$$

Необходимое условие минимума

$$\frac{\partial S}{\partial a_1} = 0 \qquad \frac{\partial S}{\partial a_2} = 0 \qquad \dots \qquad \frac{\partial S}{\partial a_m} = 0$$

В каких случаях
точка минимума функции $S(a_1, a_2, \ldots, a_m)$
может быть найдена аналитически?

В каких случаях
точка минимума функции $S(a_1, a_2, \dots, a_m)$
может быть найдена аналитически?

При каких условиях
компьютер наверняка сможет найти её
(подобрать наилучшие параметры $a_1, a_2, \dots, a_m$)?

Про разрешимость МНК

Аналитическое решение возможно

Аналитическое решение почти наверняка невозможно

Про разрешимость МНК

Аналитическое решение возможно

Аналитическое решение почти наверняка невозможно

$$\hat{y} = ax + b$$

$$\hat{y} = ax^2 + bx + c$$

$$\hat{y} = ae^x + bx^3 + c\sin(7x)$$

Про разрешимость МНК

УНИВЕРСИТЕТ ИТМО

Аналитическое решение возможно

$$\hat{y} = ax + b$$

$$\hat{y} = ax^2 + bx + c$$

$$\hat{y} = ae^x + bx^3 + c\sin(7x)$$

Аналитическое решение почти наверняка невозможно

$$\hat{y} = \sin(ax + b)$$

$$\hat{y} = x^a + bx + c$$

$$\hat{y} = ax + b + e^{cx}$$

Про разрешимость МНК

Аналитическое решение возможно

$$\hat{y} = ax + b$$

$$\hat{y} = ax^2 + bx + c$$

$$\hat{y} = ae^x + bx^3 + c\sin(7x)$$

Аналитическое решение почти наверняка невозможно

$$\hat{y} = \sin(ax + b)$$

$$\hat{y} = x^a + bx + c$$

$$\hat{y} = ax + b + e^{cx}$$

В чём между ними разница?

Про разрешимость МНК

Аналитическое решение возможно

$$\hat{y} = ax + b$$

$$\hat{y} = ax^2 + bx + c$$

$$\hat{y} = ae^x + bx^3 + c\sin(7x)$$

Функция $\hat{y}$ линейна относительно параметров

Аналитическое решение почти наверняка невозможно

$$\hat{y} = \sin(ax + b)$$

$$\hat{y} = x^a + bx + c$$

$$\hat{y} = ax + b + e^{cx}$$

Функция $\hat{y}$ не линейна относительно параметров

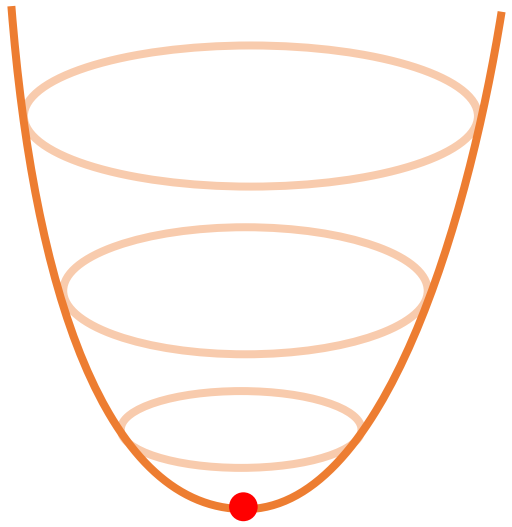
Про разрешимость МНК

Функция $\hat{y}$ линейна относительно параметров

Функция $\hat{y}$ не линейна относительно параметров

Про разрешимость МНК

Функция $\hat{y}$ линейна
относительно параметров

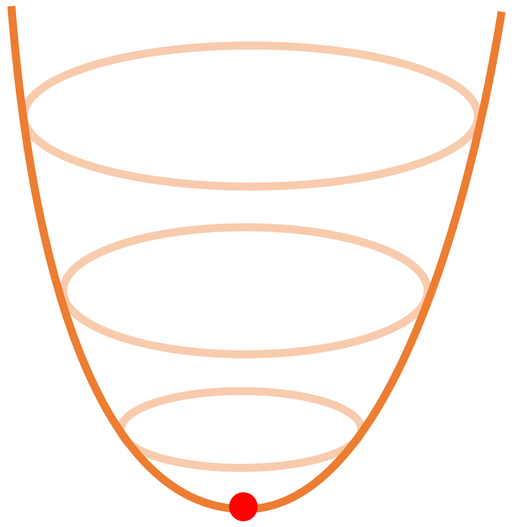


Функция S имеет
одну точку минимума

Функция $\hat{y}$ не линейна
относительно параметров

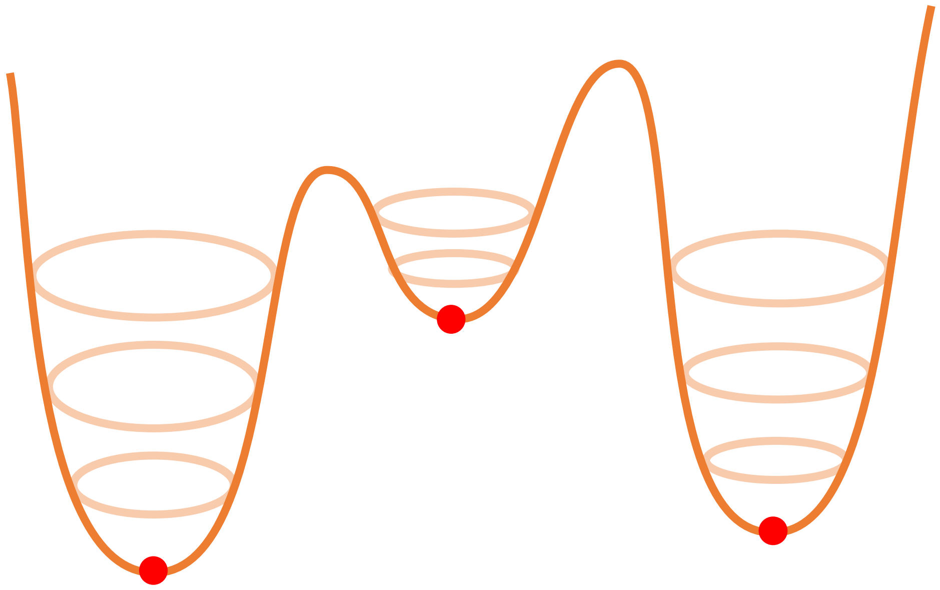
Про разрешимость МНК

Функция $\hat{y}$ линейна относительно параметров



Функция S имеет одну точку минимума

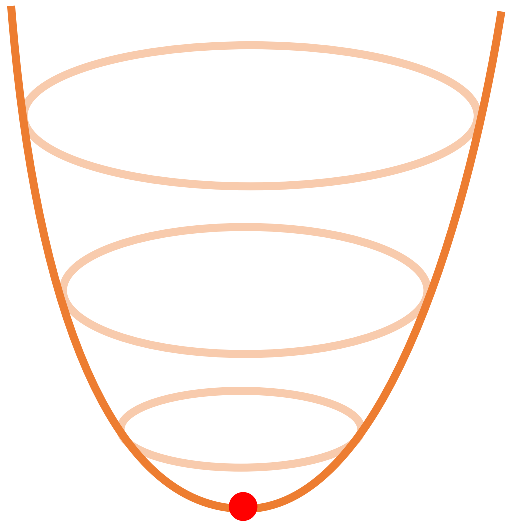
Функция $\hat{y}$ не линейна относительно параметров



Функция S может иметь много точек минимума

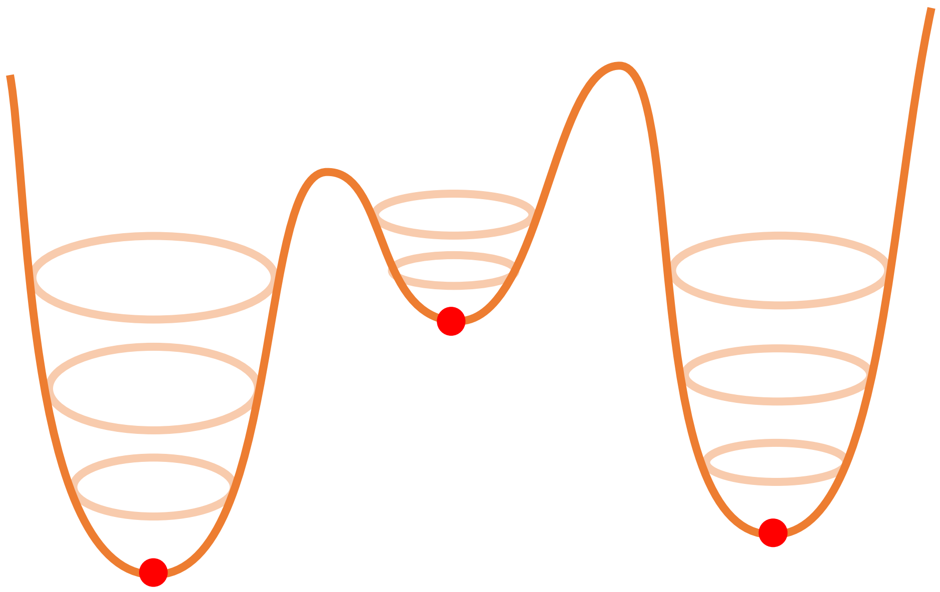
Про разрешимость МНК

Функция $\hat{y}$ линейна
относительно параметров



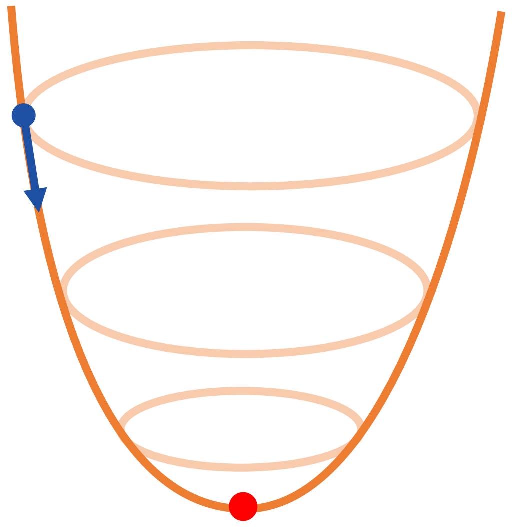
Компьютер точно найдёт
глобальный минимум

Функция $\hat{y}$ не линейна
относительно параметров



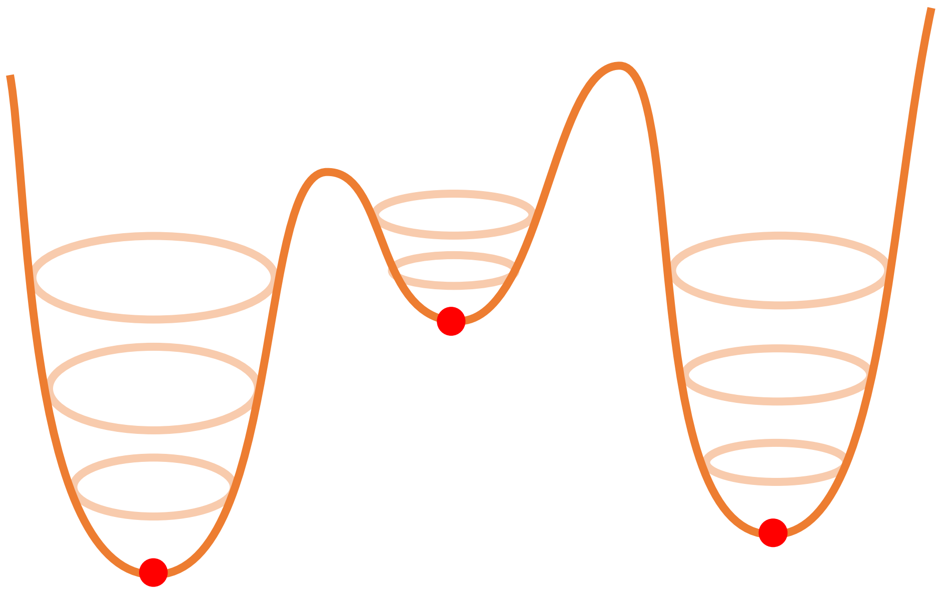
Про разрешимость МНК

Функция $\hat{y}$ линейна относительно параметров



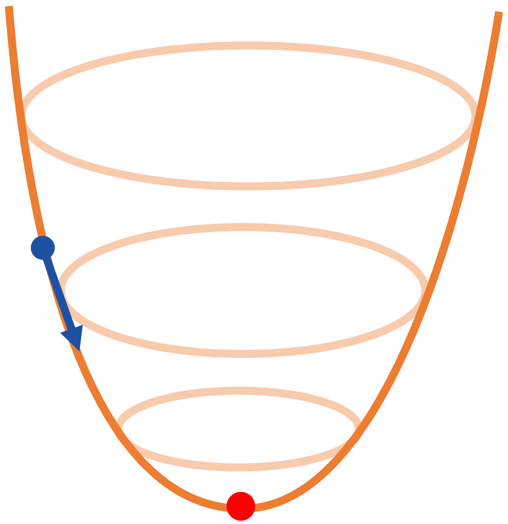
Компьютер точно найдёт глобальный минимум

Функция $\hat{y}$ не линейна относительно параметров



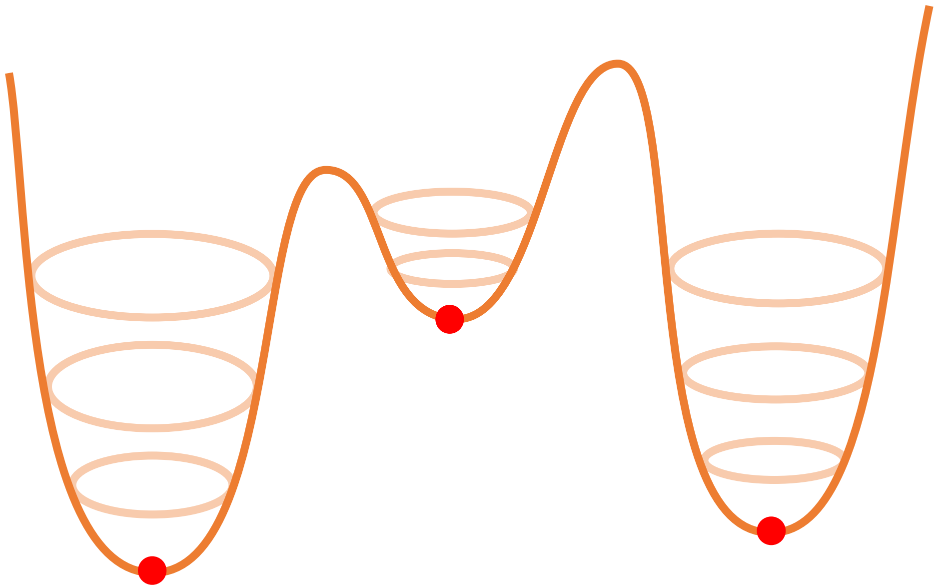
Про разрешимость МНК

Функция $\hat{y}$ линейна относительно параметров



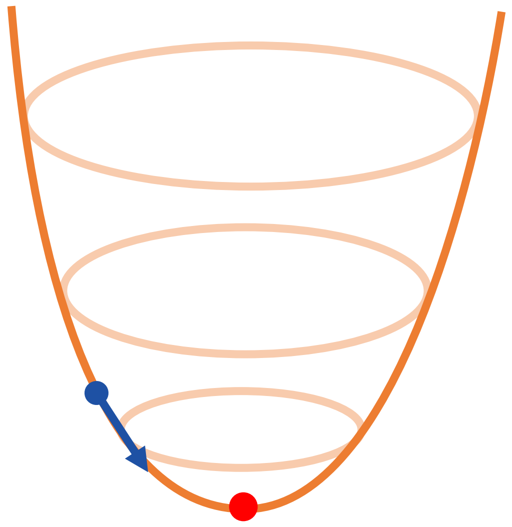
Компьютер точно найдёт глобальный минимум

Функция $\hat{y}$ не линейна относительно параметров



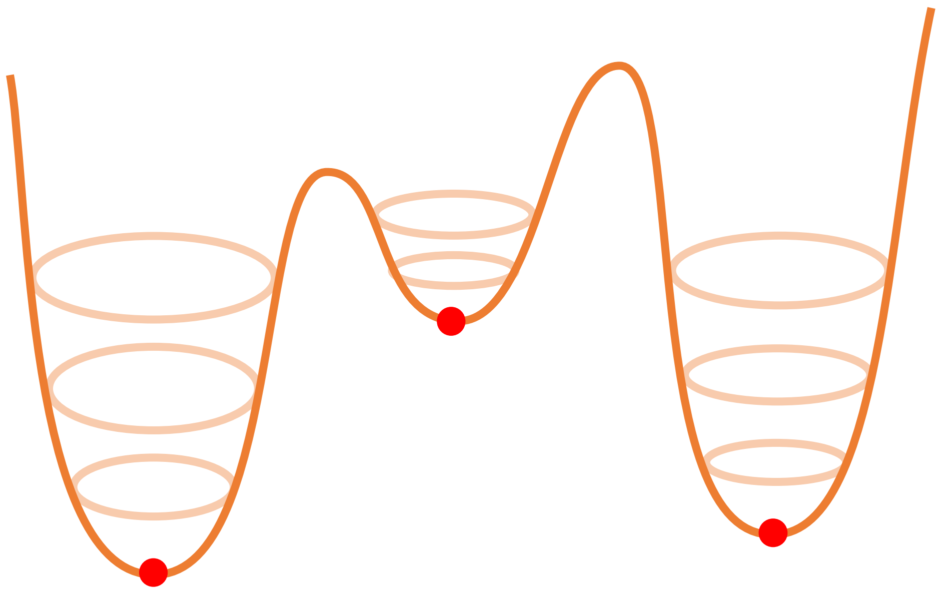
Про разрешимость МНК

Функция $\hat{y}$ линейна относительно параметров



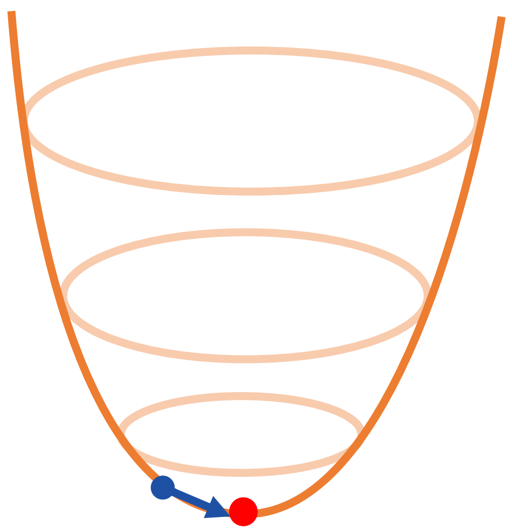
Компьютер точно найдёт глобальный минимум

Функция $\hat{y}$ не линейна относительно параметров



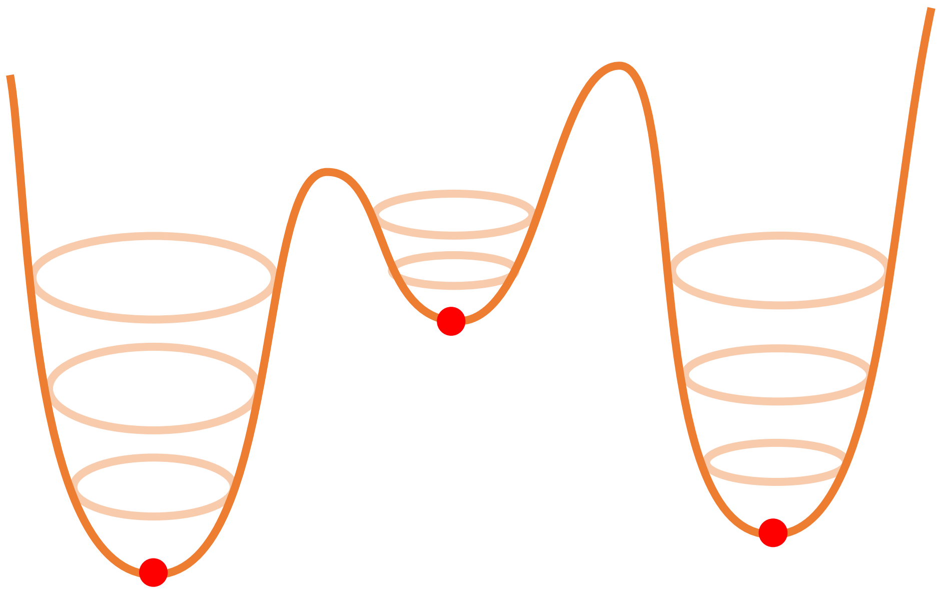
Про разрешимость МНК

Функция $\hat{y}$ линейна относительно параметров



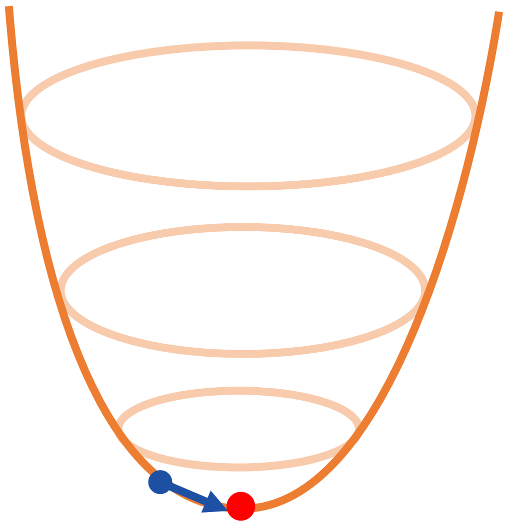
Компьютер точно найдёт глобальный минимум

Функция $\hat{y}$ не линейна относительно параметров



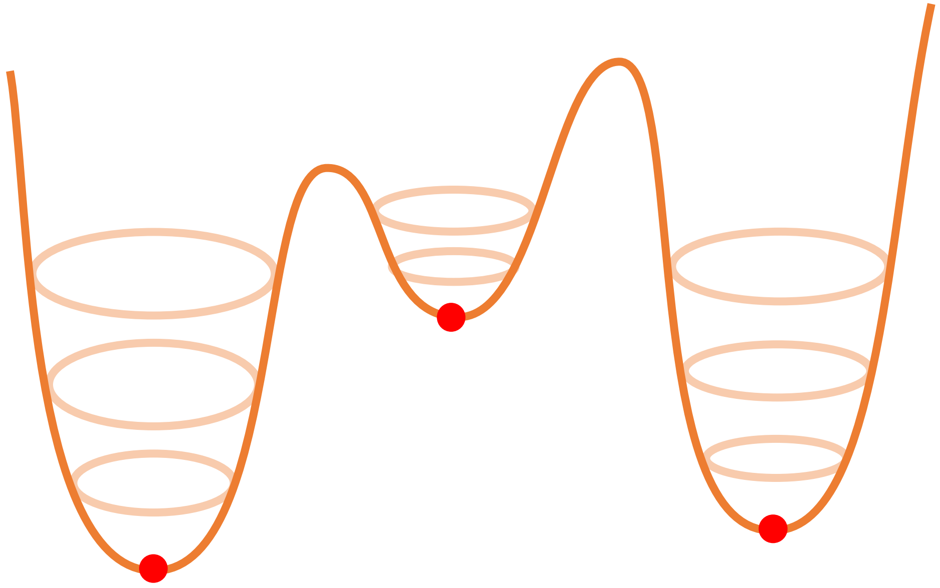
Про разрешимость МНК

Функция $\hat{y}$ линейна относительно параметров



Компьютер точно найдёт глобальный минимум

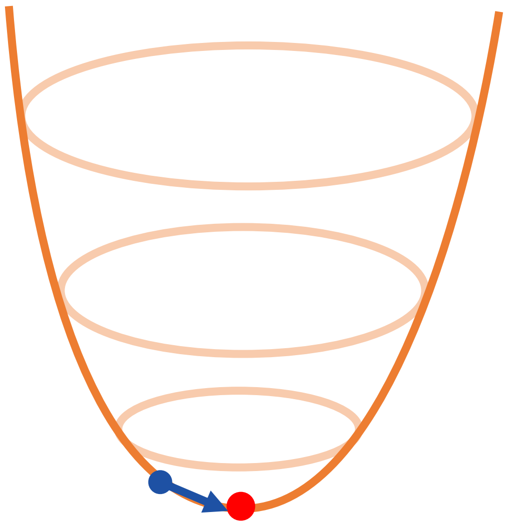
Функция $\hat{y}$ не линейна относительно параметров



Компьютер может попасть в локальный минимум

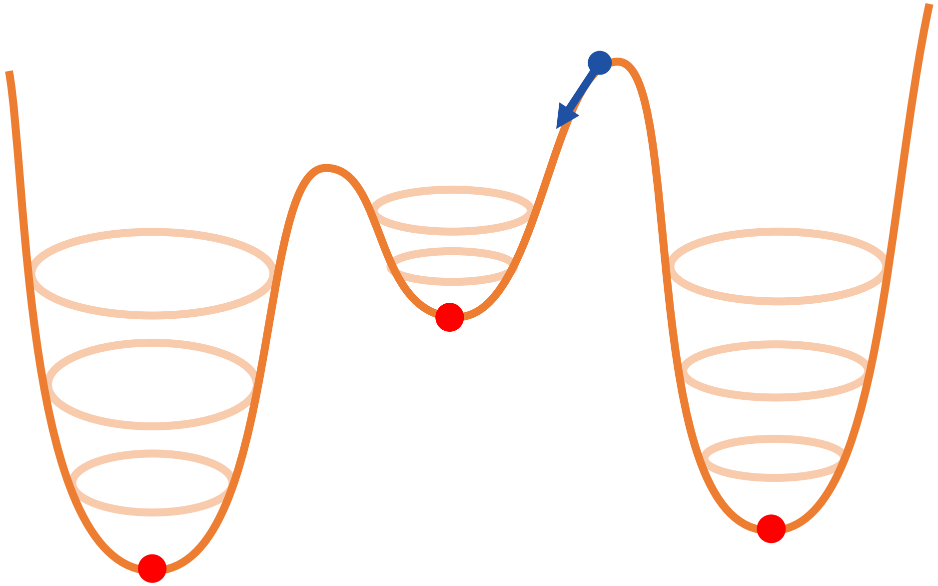
Про разрешимость МНК

Функция $\hat{y}$ линейна относительно параметров



Компьютер точно найдёт глобальный минимум

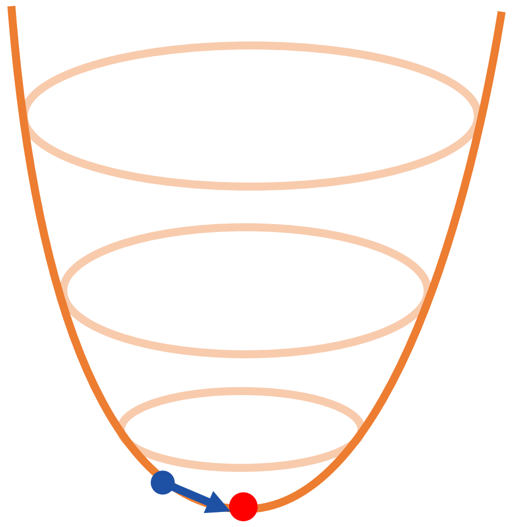
Функция $\hat{y}$ не линейна относительно параметров



Компьютер может попасть в локальный минимум

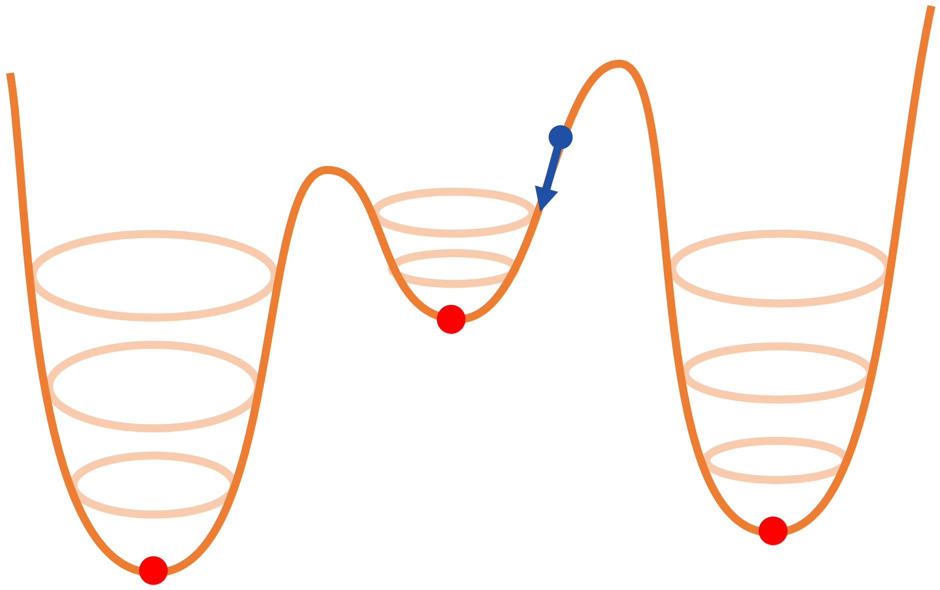
Про разрешимость МНК

Функция $\hat{y}$ линейна
относительно параметров



Компьютер точно найдёт
глобальный минимум

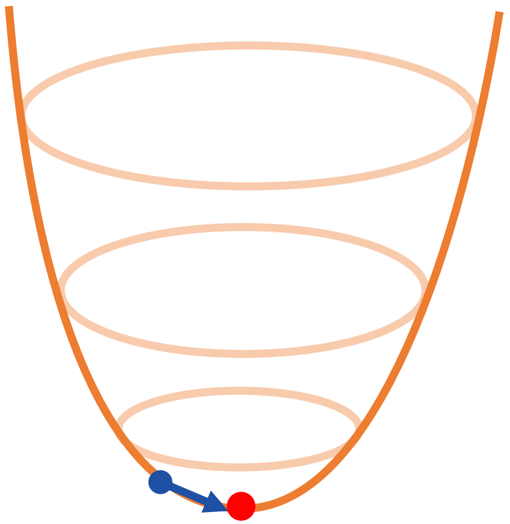
Функция $\hat{y}$ не линейна
относительно параметров



Компьютер может попасть
в локальный минимум

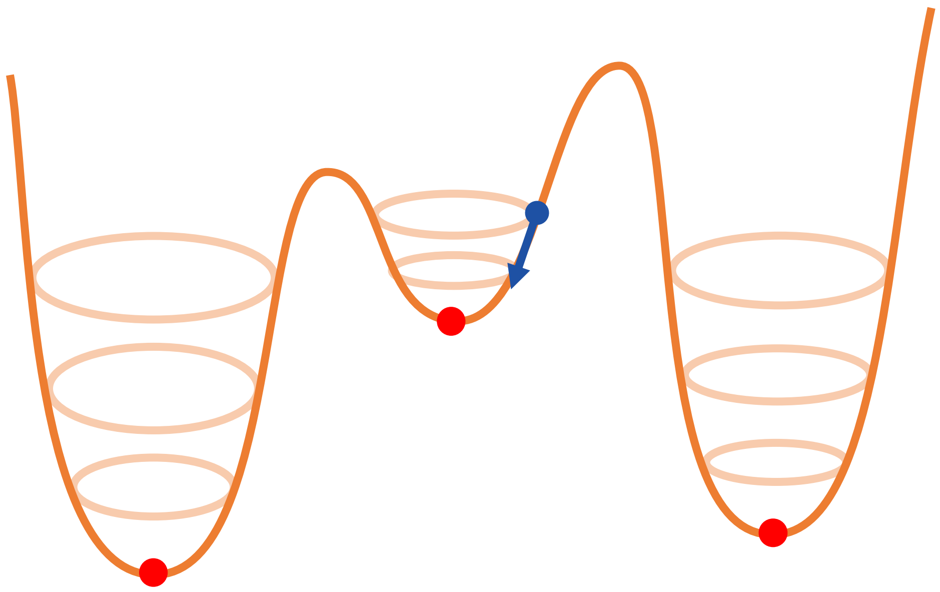
Про разрешимость МНК

Функция $\hat{y}$ линейна относительно параметров



Компьютер точно найдёт глобальный минимум

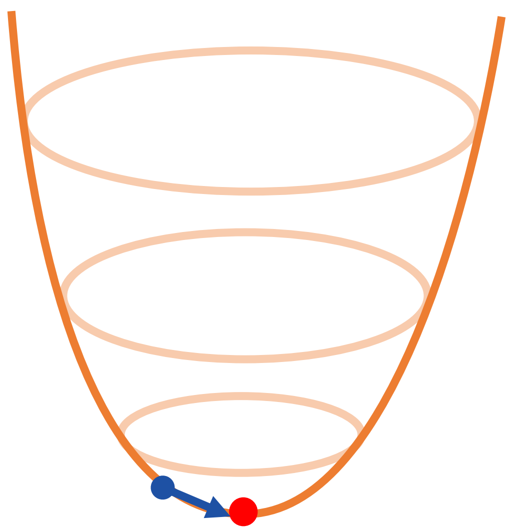
Функция $\hat{y}$ не линейна относительно параметров



Компьютер может попасть в локальный минимум

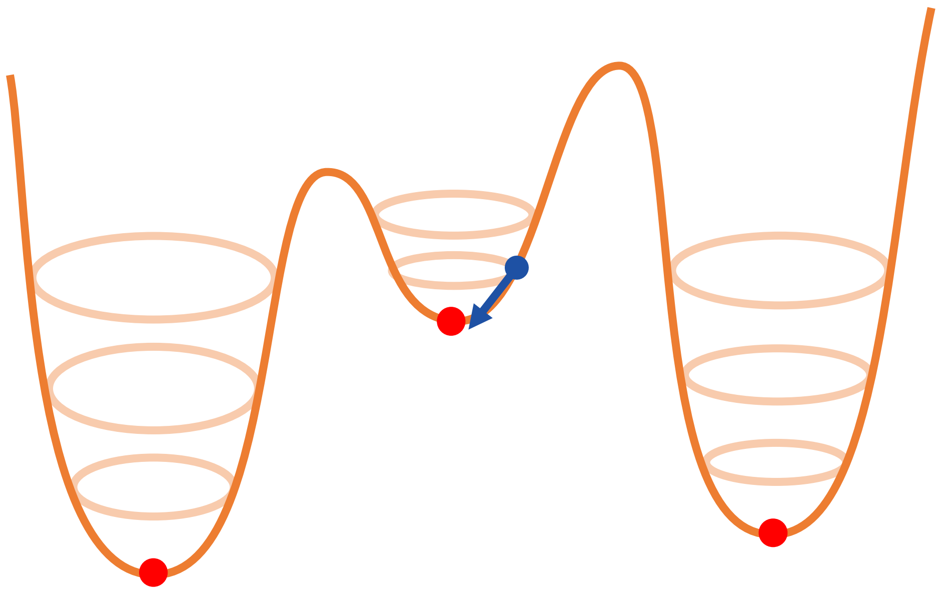
Про разрешимость МНК

Функция $\hat{y}$ линейна относительно параметров



Компьютер точно найдёт глобальный минимум

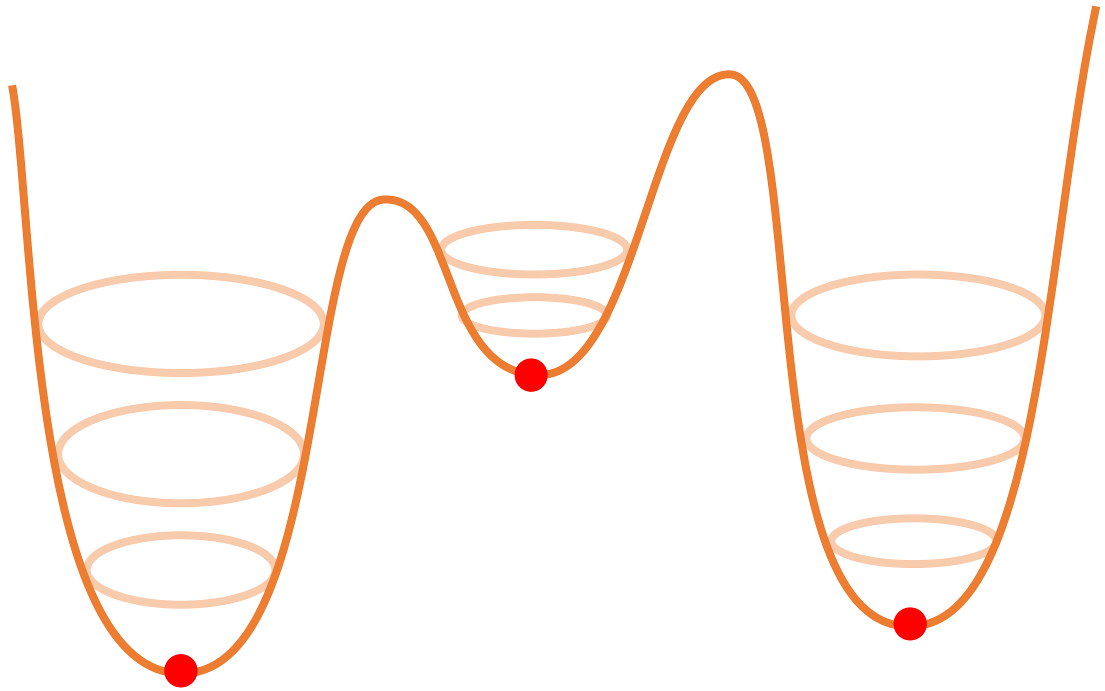
Функция $\hat{y}$ не линейна относительно параметров



Компьютер может попасть в локальный минимум

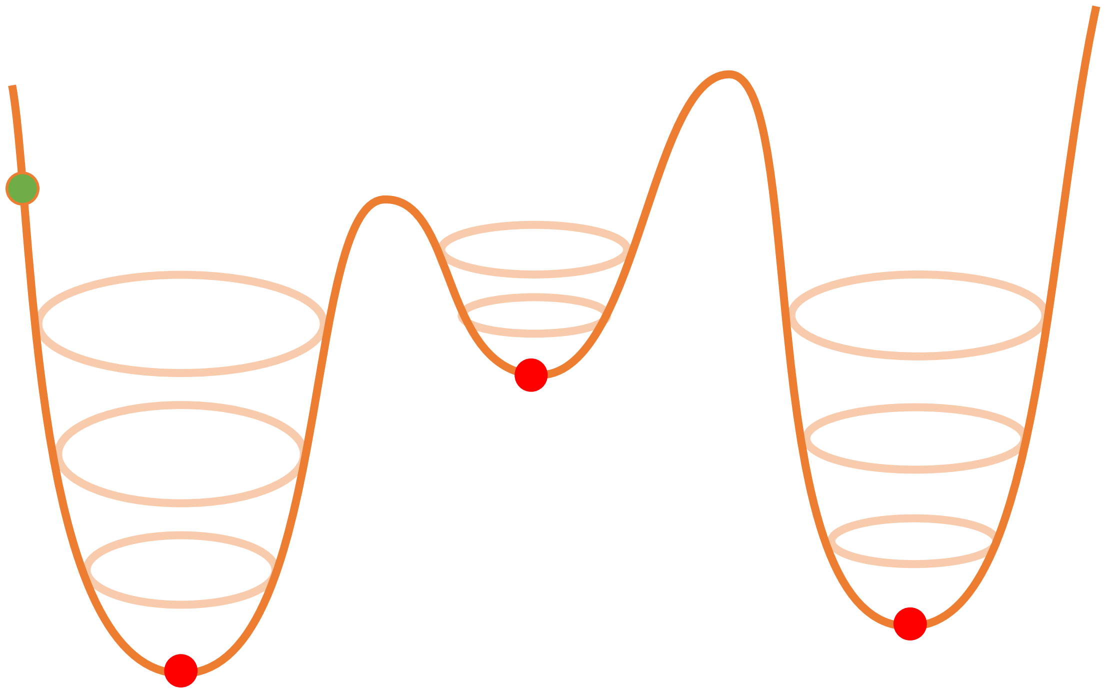
Про разрешимость МНК

Для успешного компьютерного поиска, важно задать хорошую начальную точку, близкую к точке глобального минимума



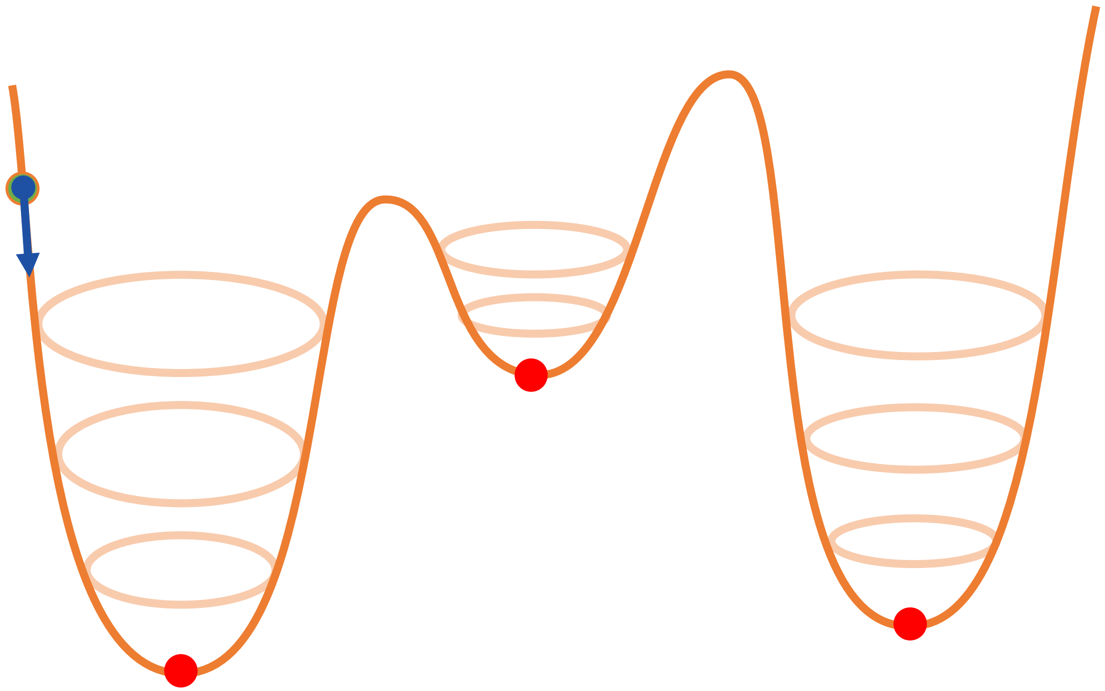
Про разрешимость МНК

Для успешного компьютерного поиска, важно задать хорошую начальную точку, близкую к точке глобального минимума



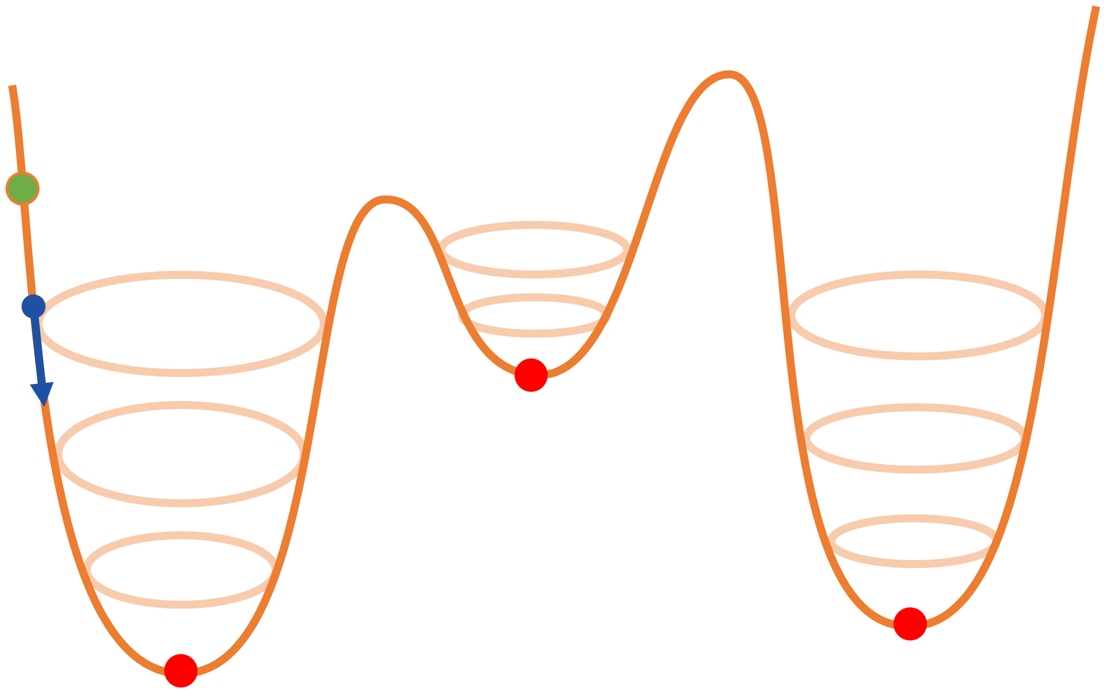
Про разрешимость МНК

Для успешного компьютерного поиска, важно задать хорошую начальную точку, близкую к точке глобального минимума



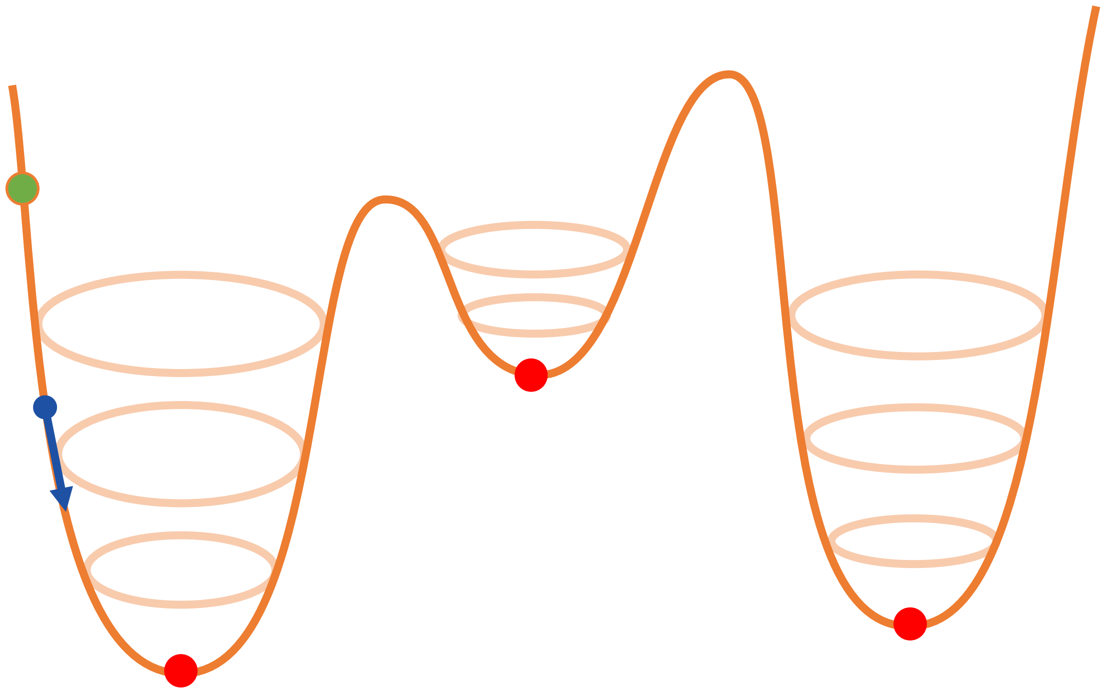
Про разрешимость МНК

Для успешного компьютерного поиска, важно задать хорошую начальную точку, близкую к точке глобального минимума



Про разрешимость МНК

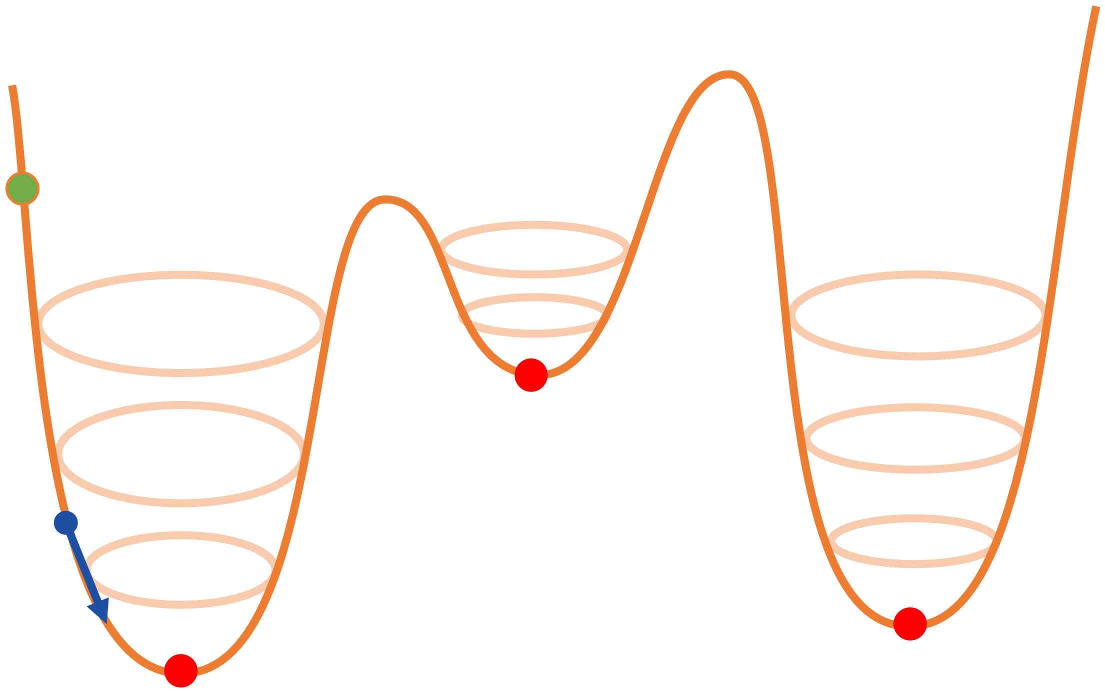
Для успешного компьютерного поиска, важно задать хорошую начальную точку, близкую к точке глобального минимума



Про разрешимость МНК

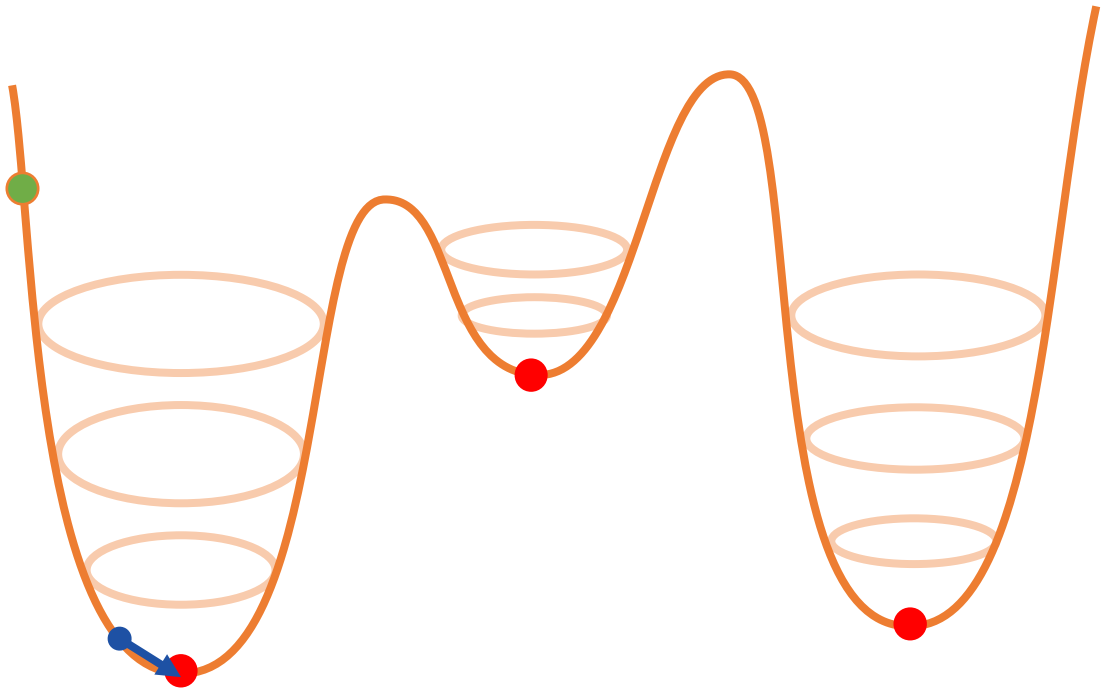
УНИВЕРСИТЕТ ИТМО

Для успешного компьютерного поиска, важно задать хорошую начальную точку, близкую к точке глобального минимума



Про разрешимость МНК

Для успешного компьютерного поиска, важно задать хорошую начальную точку, близкую к точке глобального минимума



МНК в первой лабораторной

МНК в первой лабораторной

$$\theta(t) = \omega_{nls}\left(t - T_m\left(1 - e^{-\frac{t}{T_m}}\right)\right)$$

МНК в первой лабораторной

Параметры входят нелинейно

$$\theta(t) = \omega_{nls}\left(t - T_m\left(1 - e^{-\frac{t}{T_m}}\right)\right)$$

МНК в первой лабораторной

Параметры входят нелинейно

$$\theta(t) = \omega_{nls}\left(t - T_m\left(1 - e^{-\frac{t}{T_m}}\right)\right)$$

Важно выбрать начальную точку (ω_{nls}, T_m) недалеко от глобального минимума

МНК в первой лабораторной

Параметры входят нелинейно

$$\theta(t) = \omega_{nls}\left(t - T_m\left(1 - e^{-\frac{t}{T_m}}\right)\right)$$

Важно выбрать начальную точку (ω_{nls}, T_m) недалеко от глобального минимума

$$\mathrm{att} = [15; 0.06]$$

МНК в первой лабораторной

Параметры входят нелинейно

$$\theta(t) = \omega_{nls}\left(t - T_m\left(1 - e^{-\frac{t}{T_m}}\right)\right)$$

Важно выбрать начальную точку (ω_{nls}, T_m) недалеко от глобального минимума

$$\text{att} = [15; 0.06]$$

Если вы получали плохой результат, и вам приходилось менять эти числа, значит вы попадали в локальный минимум

МНК в первой лабораторной

Фрагмент из методички

2.9.4 Задайте одностолбцовую (нельзя одностроковую) матрицу att из двух элементов, где надо разместить те значения параметров ω_{nls} и T_m, которые, как вы думаете, получатся. Это помогает лишь ускорить процесс, поэтому содержание данной матрицы может быть любым, например `att=[15;0.06]`.

МНК в первой лабораторной

Фрагмент из методички

2.9.4 Задайте одностолбцовую (нельзя однострококовую) матрицу att из двух элементов, где надо разместить те значения параметров ω_{nls} и T_m, которые, как вы думаете, получатся. Это помогает лишь ускорить процесс, поэтому содержание данной матрицы может быть любым, например `att=[15;0.06]`.

Неправда

МНК в первой лабораторной

Фрагмент из методички

2.9.4 Задайте одностолбцовую (нельзя одностроковую) матрицу att из двух элементов, где надо разместить те значения параметров ω_{nls} и T_m, которые, как вы думаете, получатся. Это помогает лишь ускорить процесс, поэтому содержание данной матрицы может быть любым, например `att=[15;0.06]`.

Неправда

Если аппроксимирующая функция не линейна относительно параметров, то может повлиять на результат

Почему МНК?

Почему МНК?

Почему надо использовать именно
метод наименьших квадратов?

Почему МНК?

Почему надо использовать именно
метод наименьших квадратов?

Почему не метод наименьших модулей?

Почему МНК?

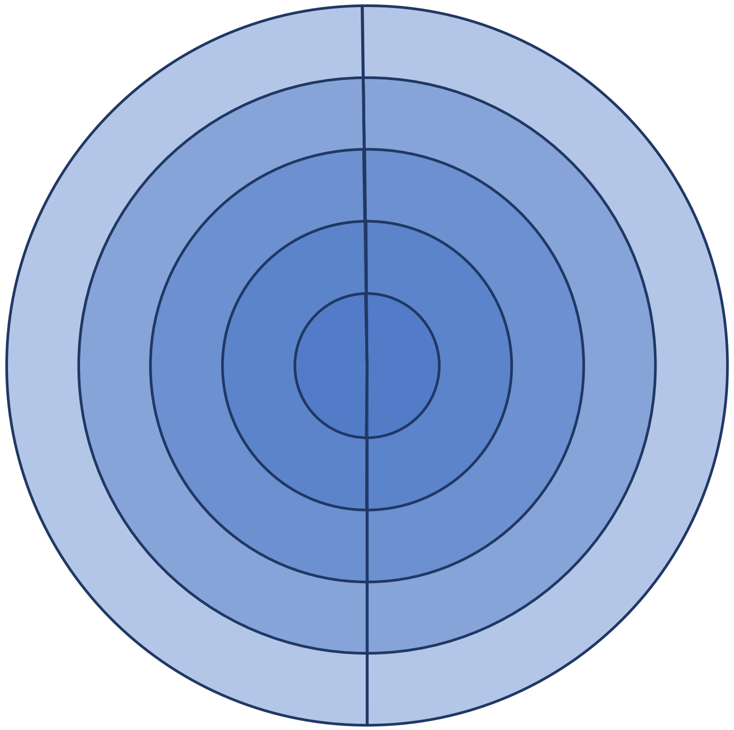
Почему надо использовать именно
метод наименьших квадратов?

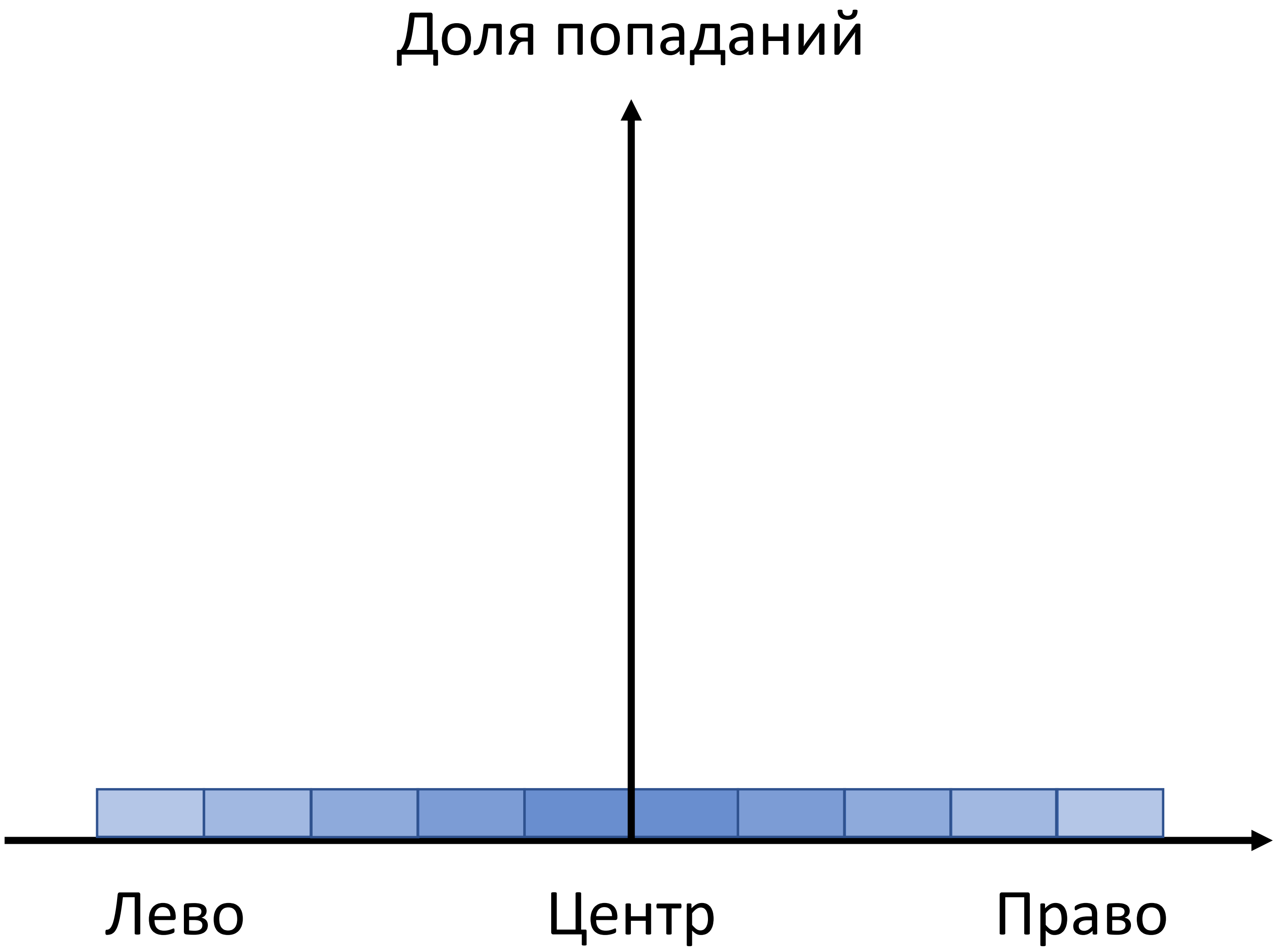
Почему не метод наименьших модулей?

Отчего мы минимизируем именно l_2-норму,
а не какую-то другую?

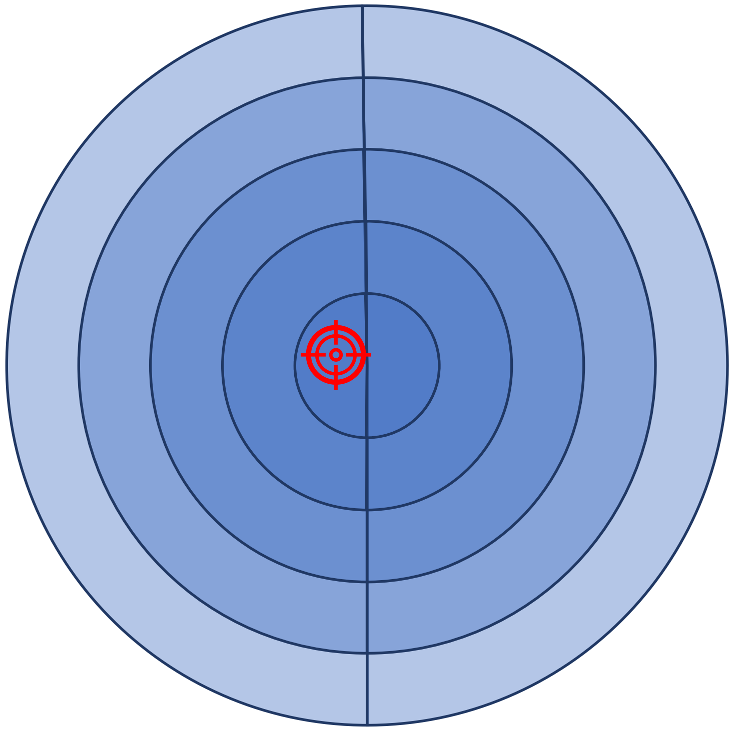
Распределения вероятностей

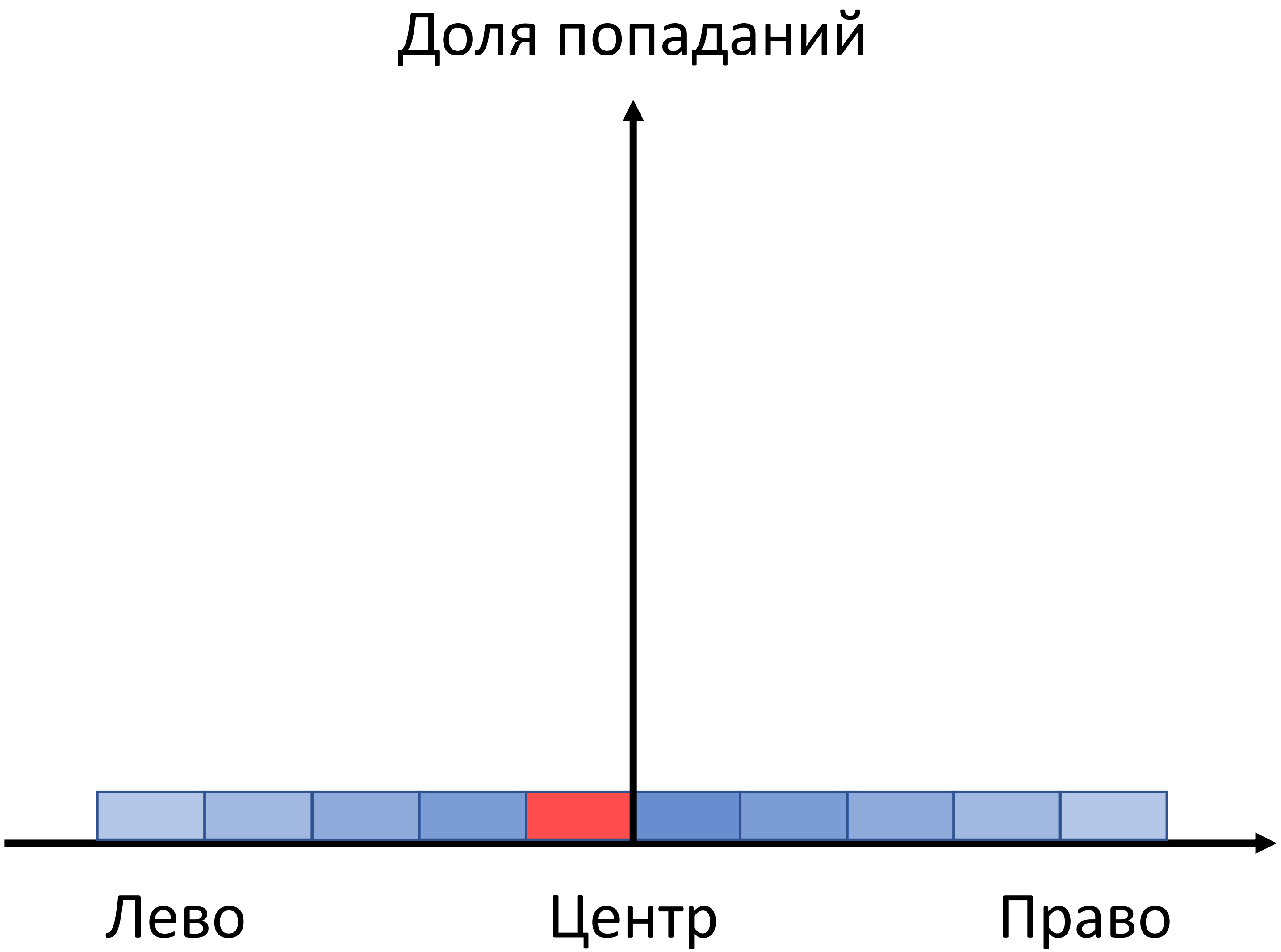
Распределения вероятностей



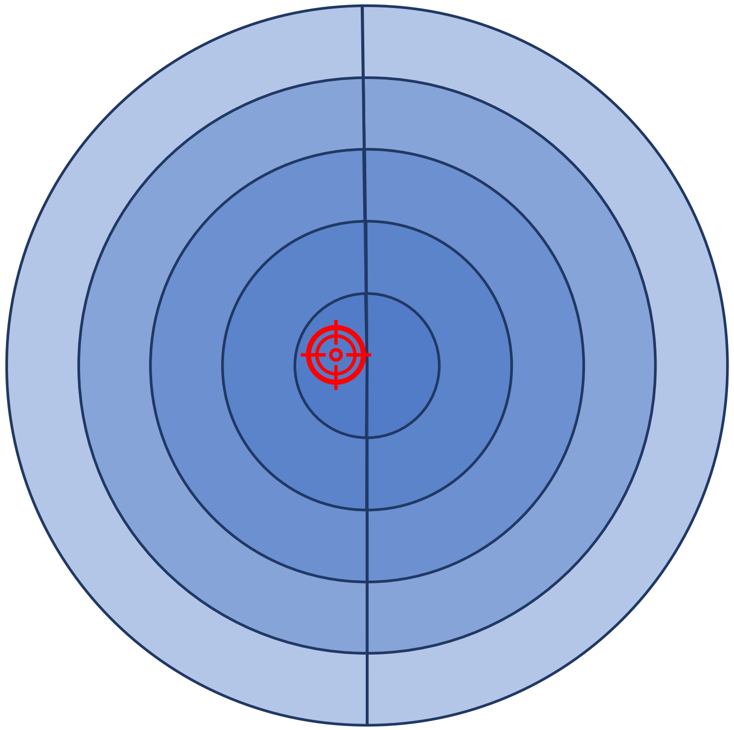


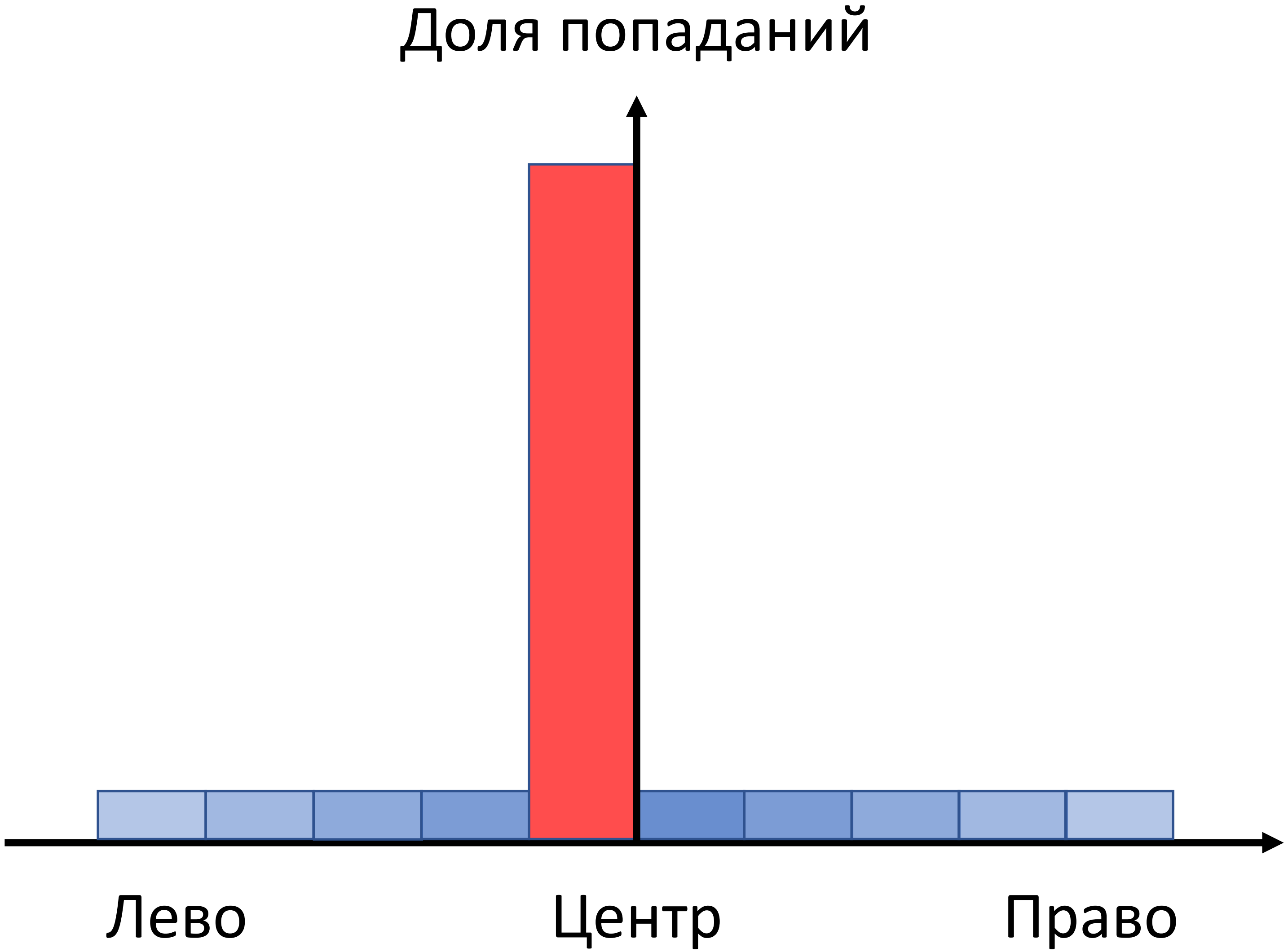
Распределения вероятностей



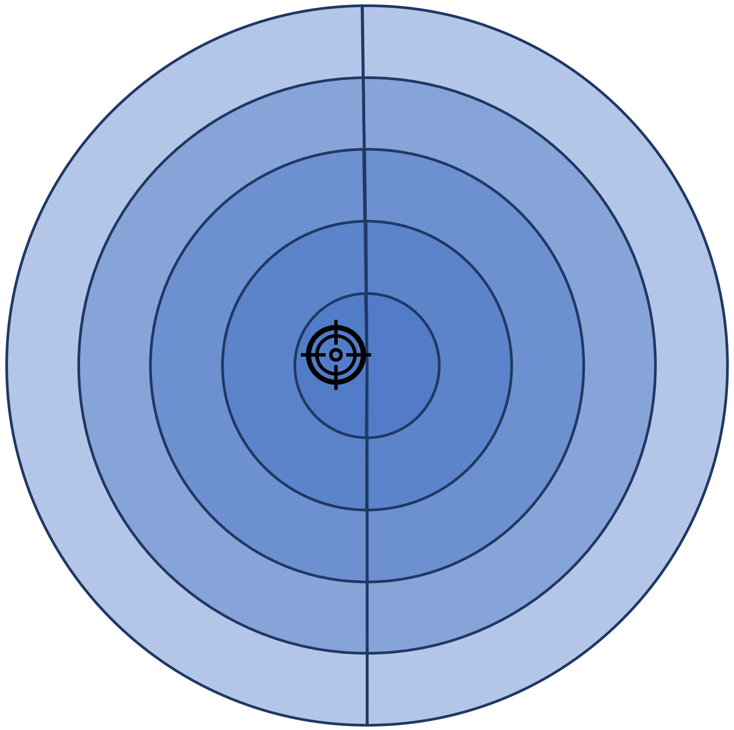


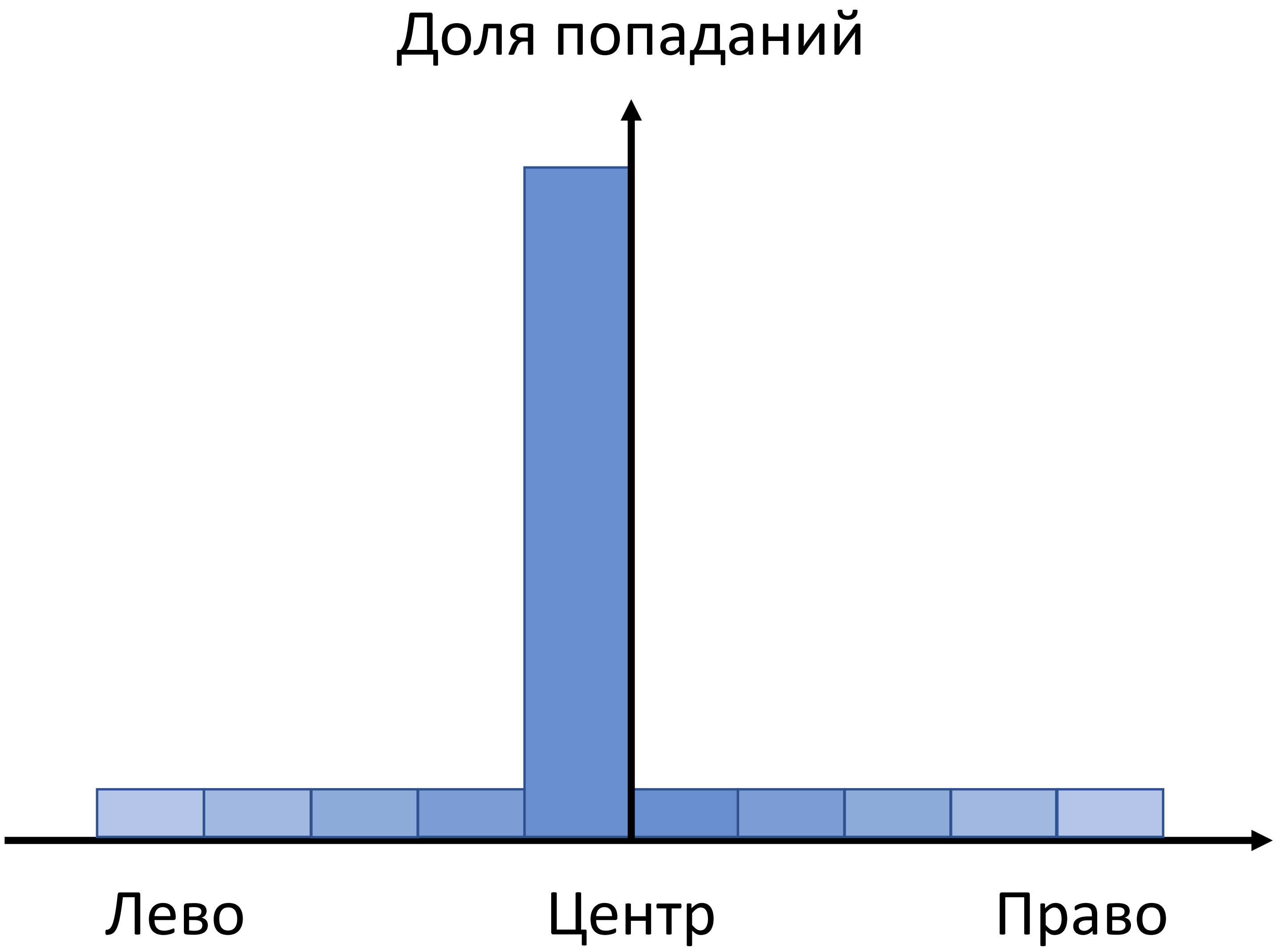
Распределения вероятностей



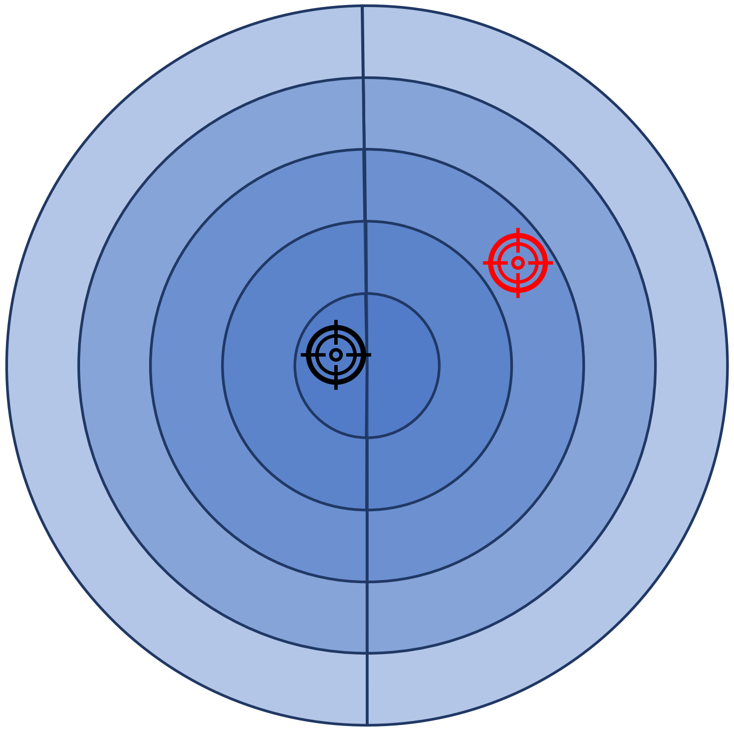


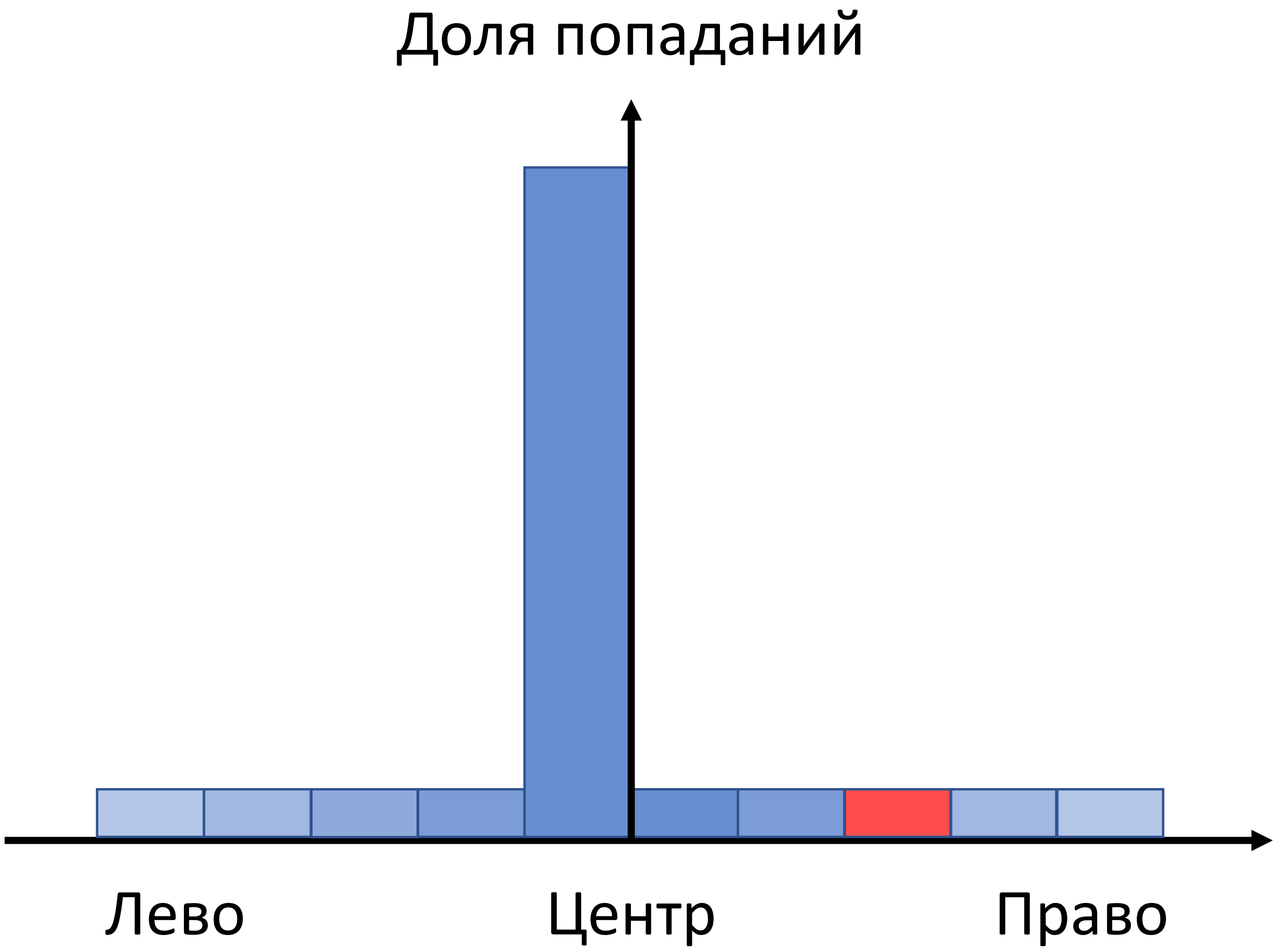
Распределения вероятностей



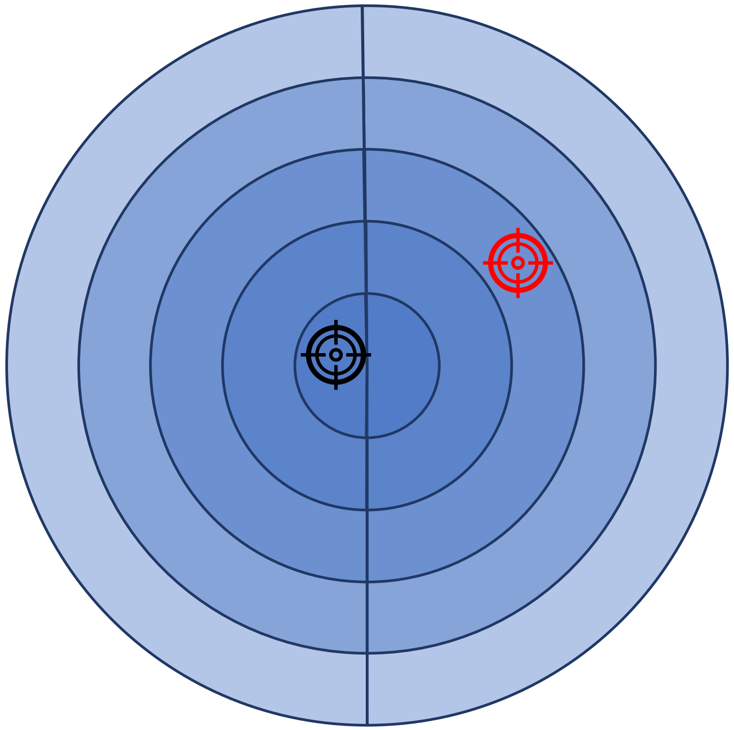


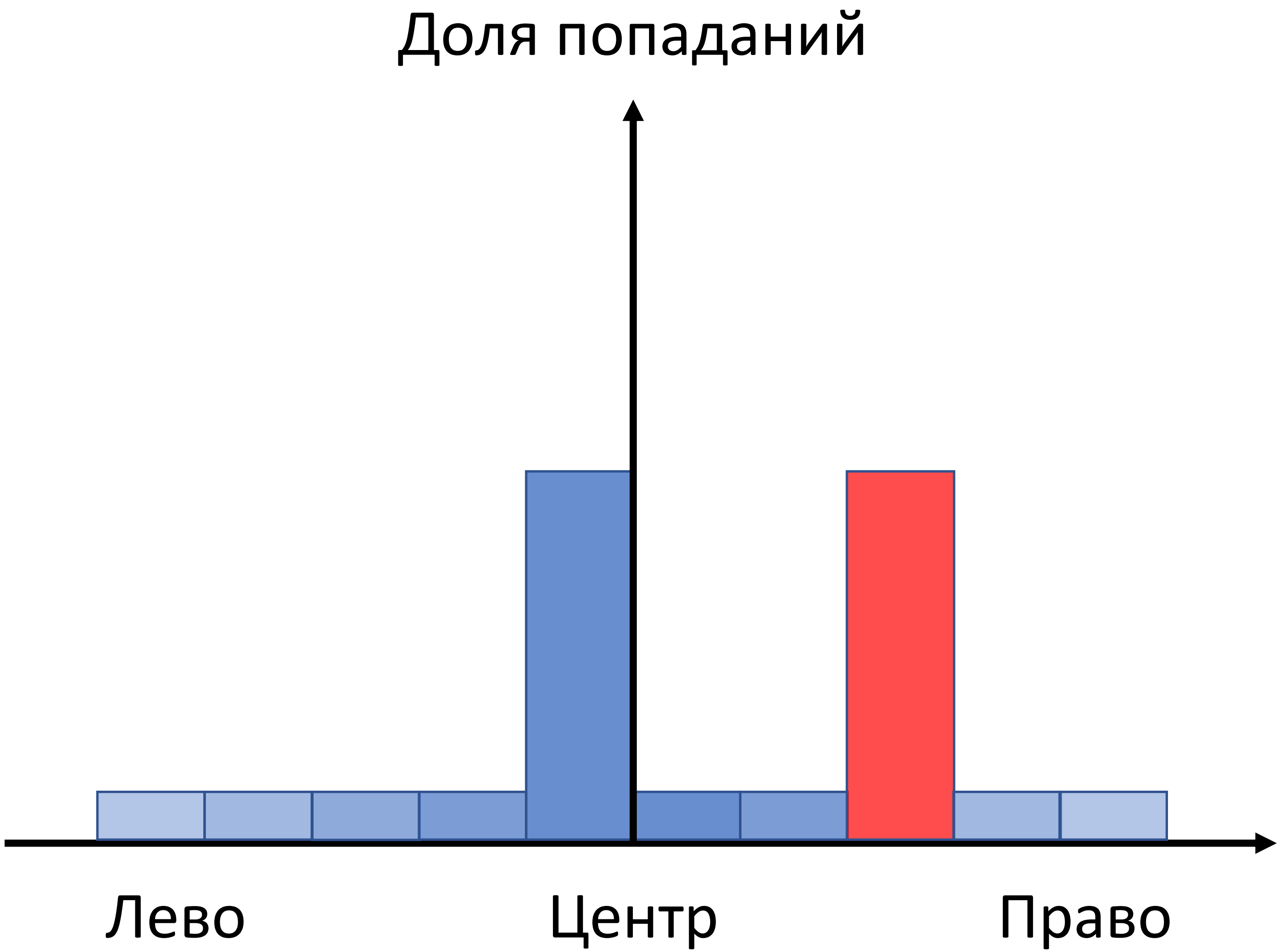
Распределения вероятностей



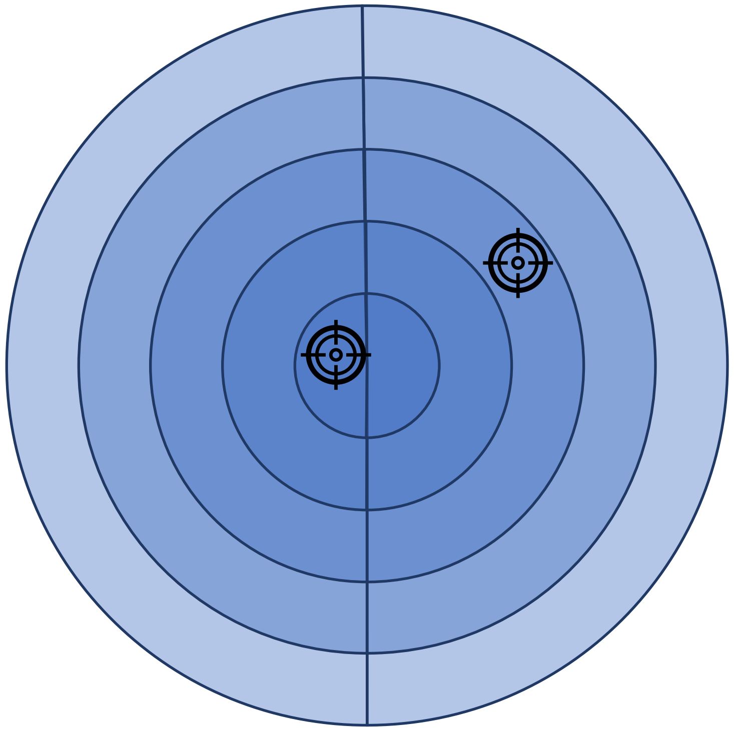


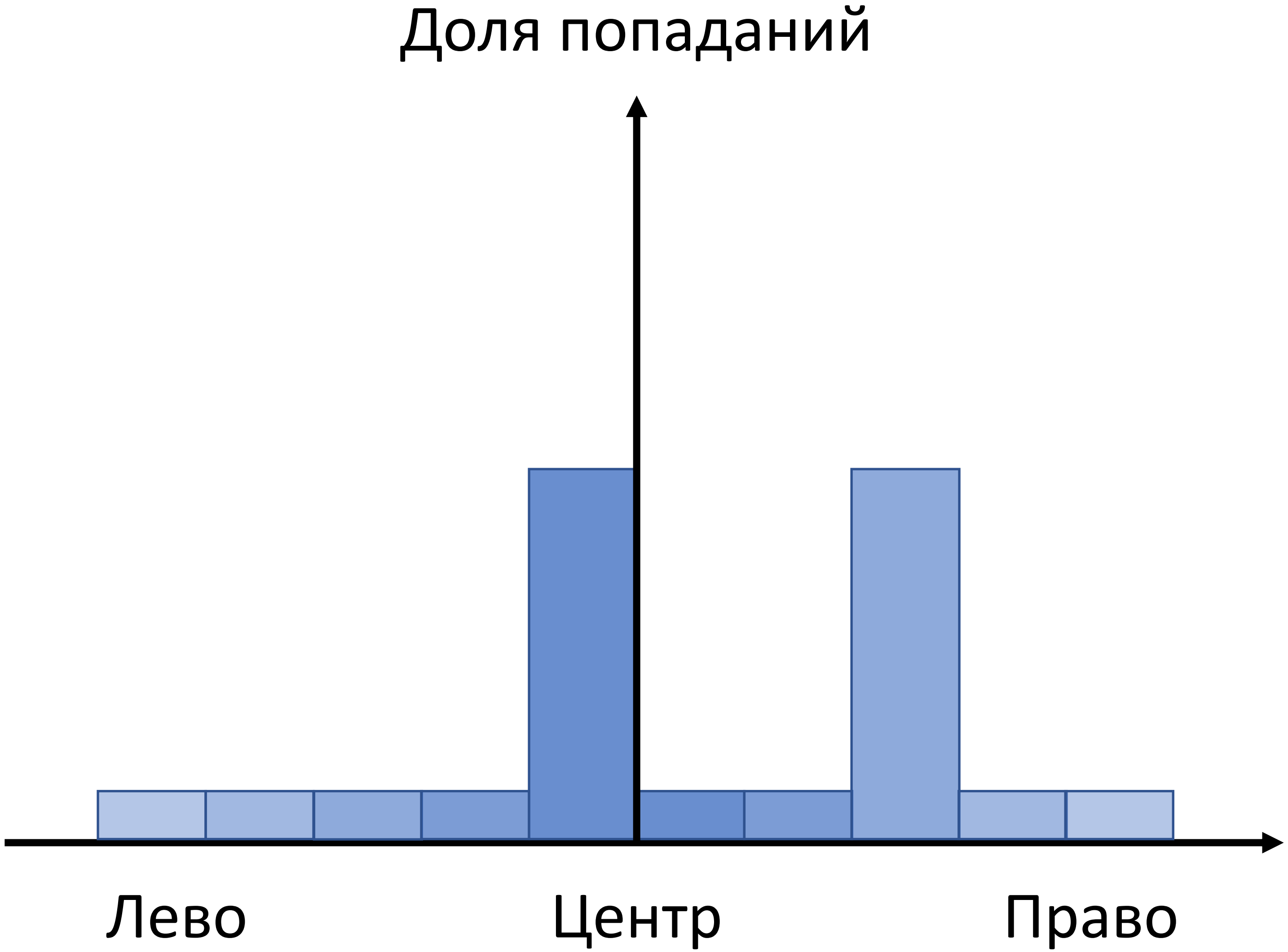
Распределения вероятностей





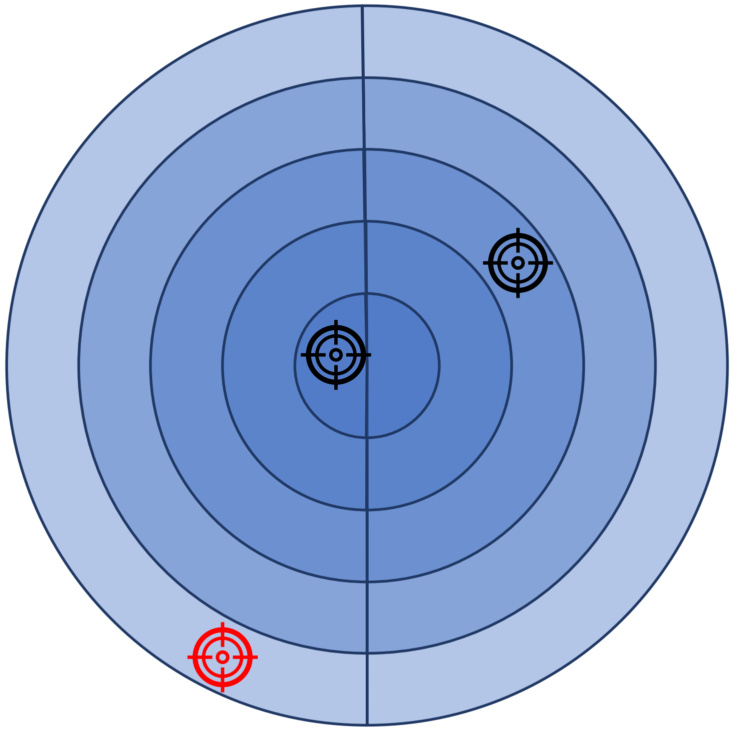
Распределения вероятностей

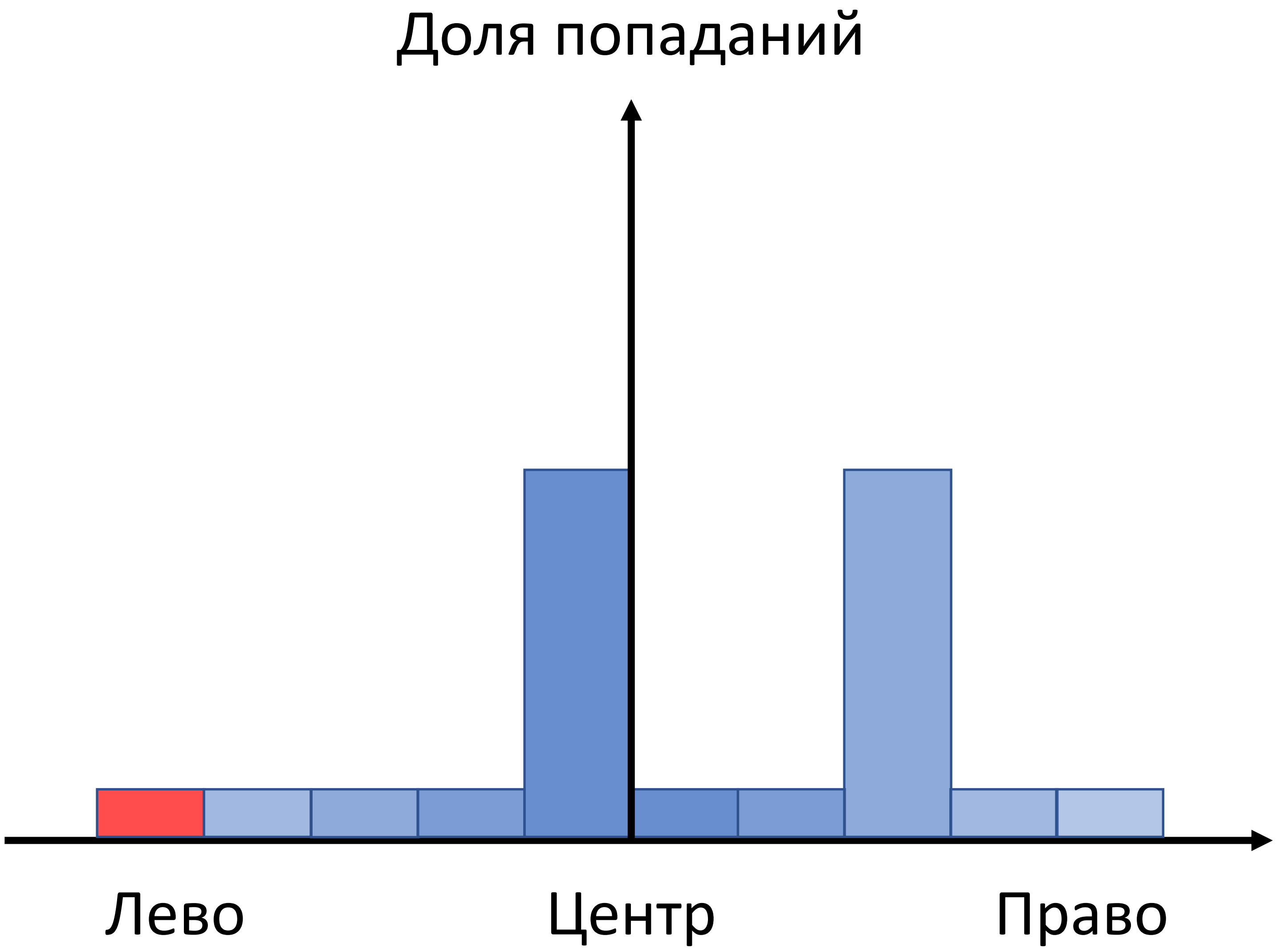




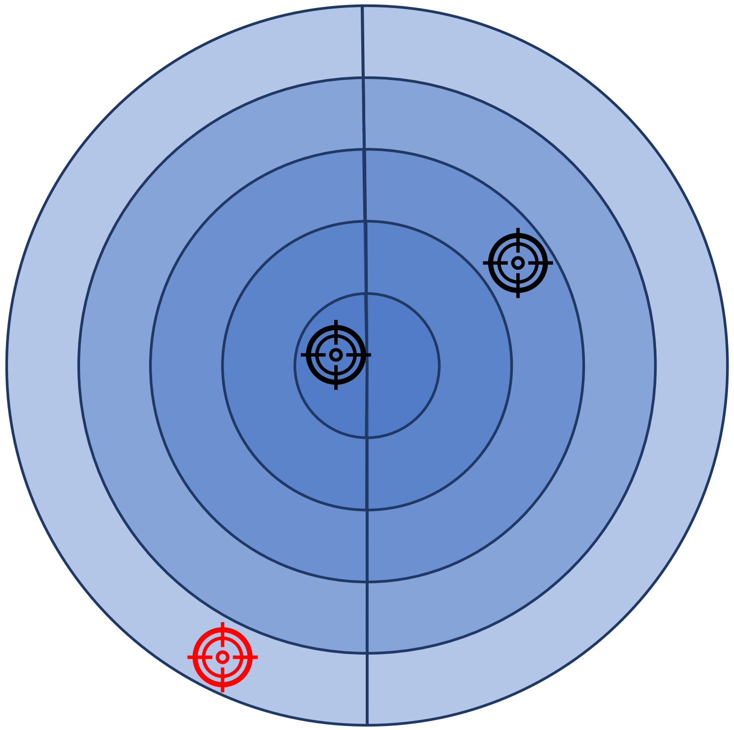
Распределения вероятностей

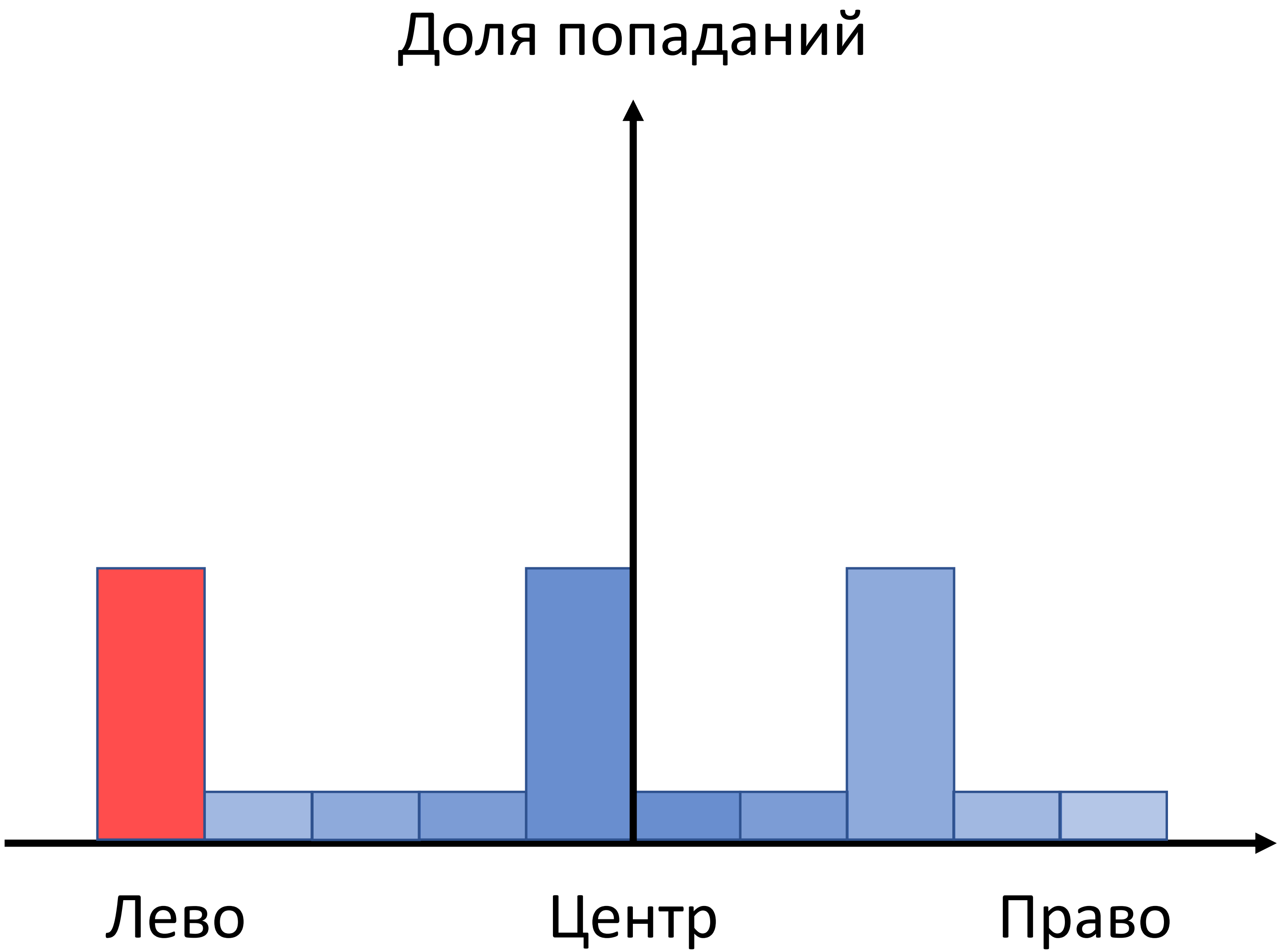
УНИВЕРСИТЕТ ИТМО



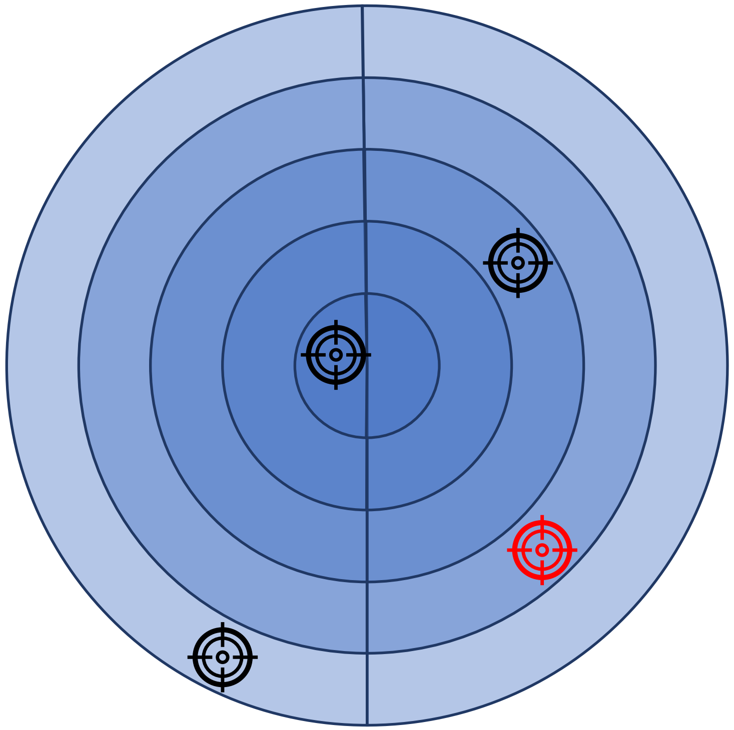


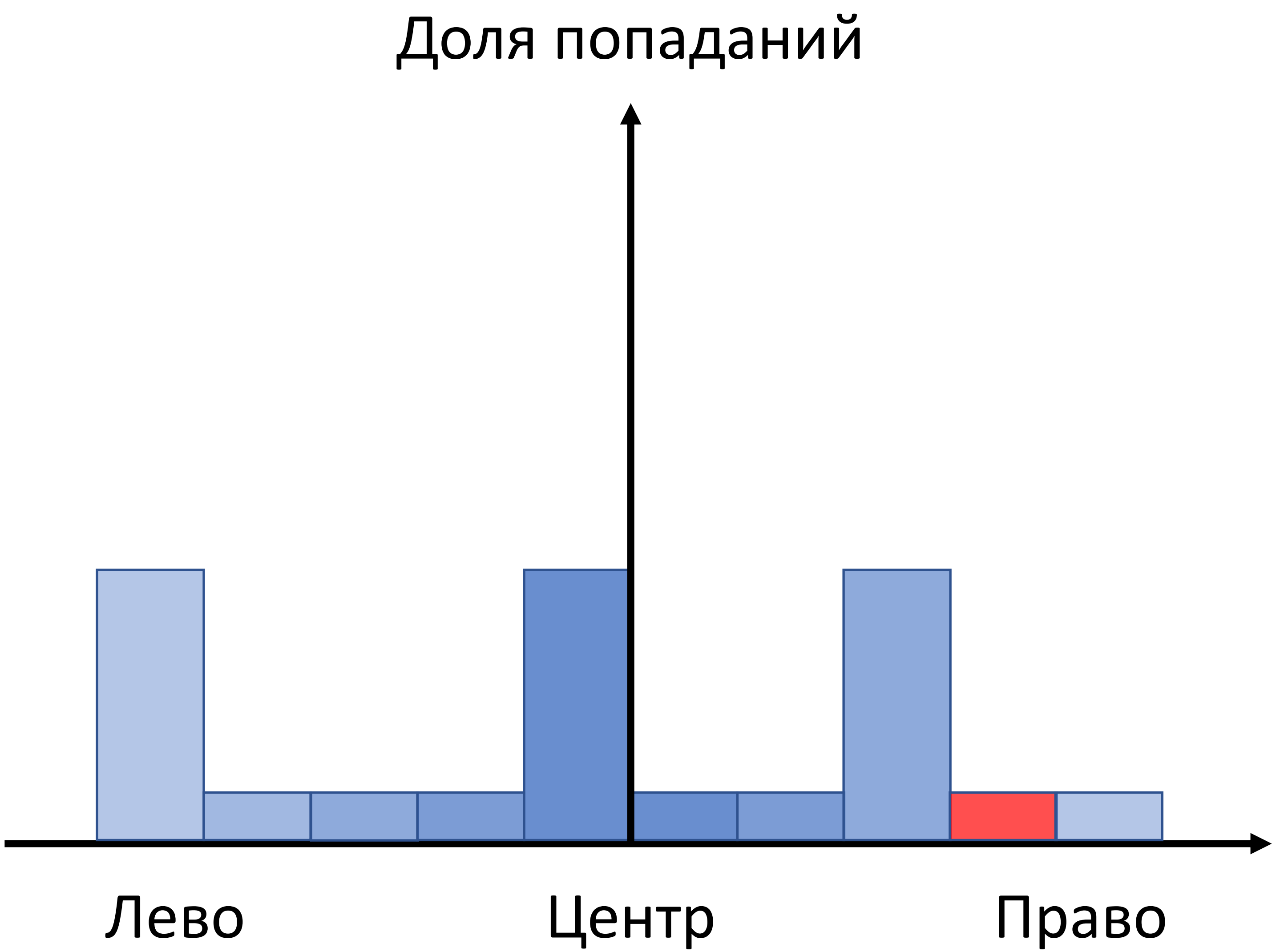
Распределения вероятностей



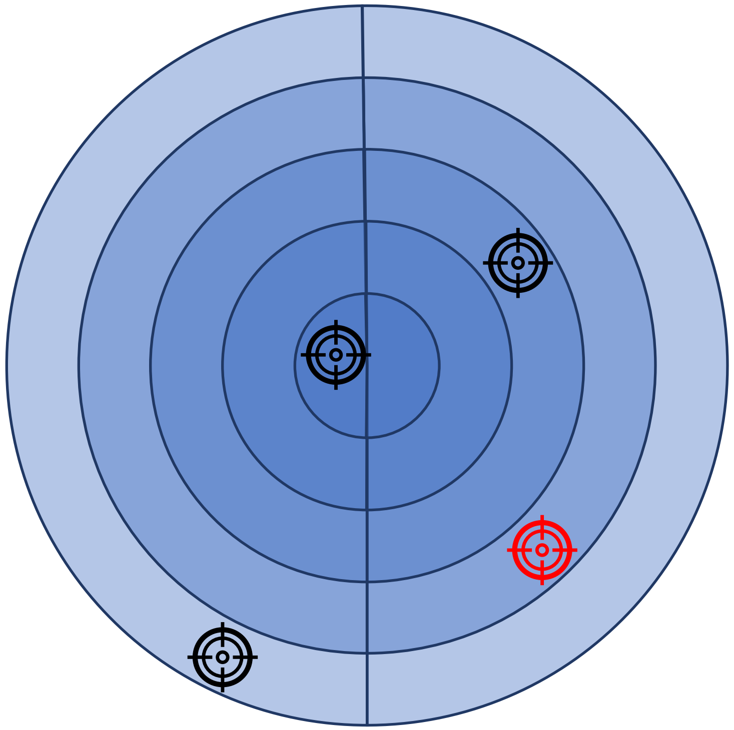


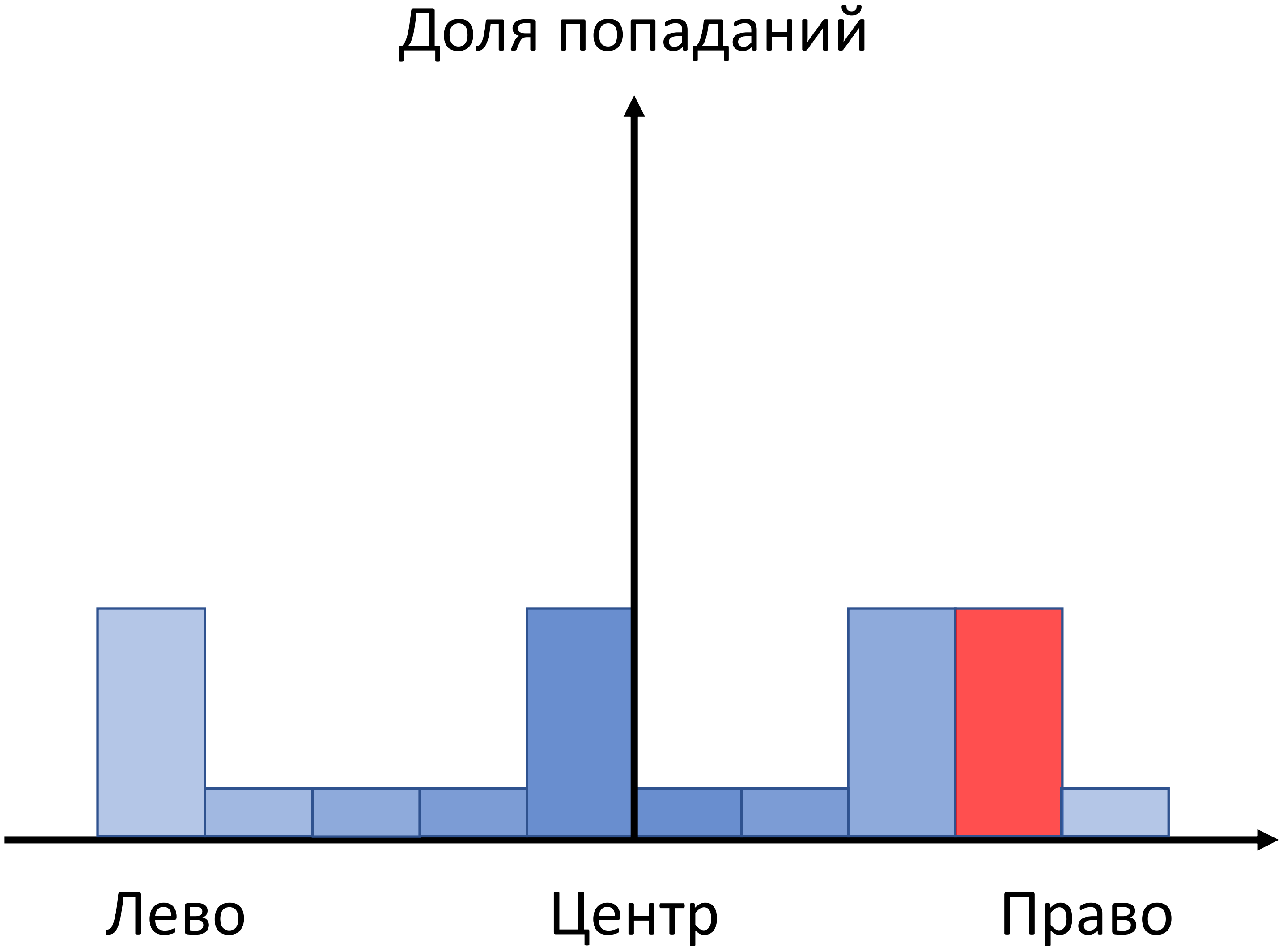
Распределения вероятностей



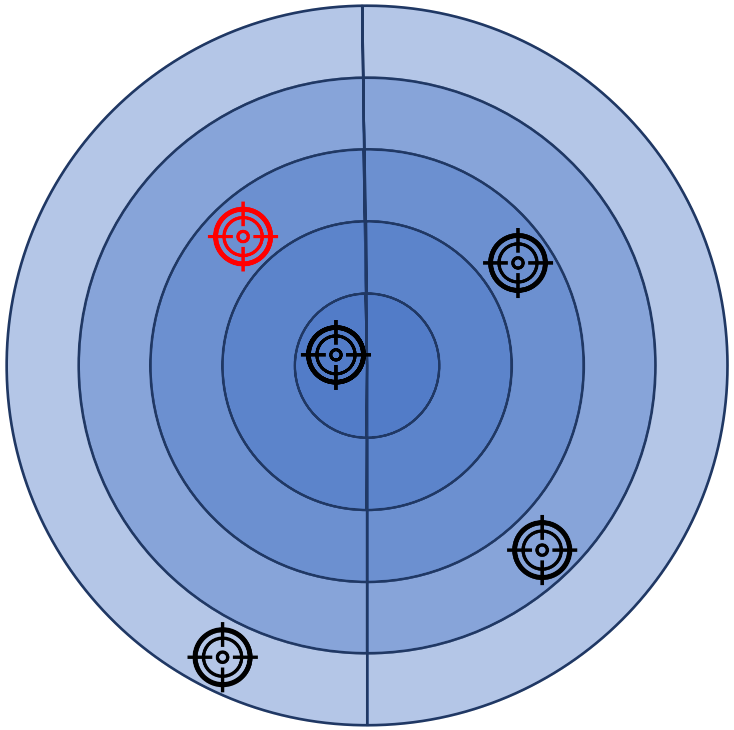


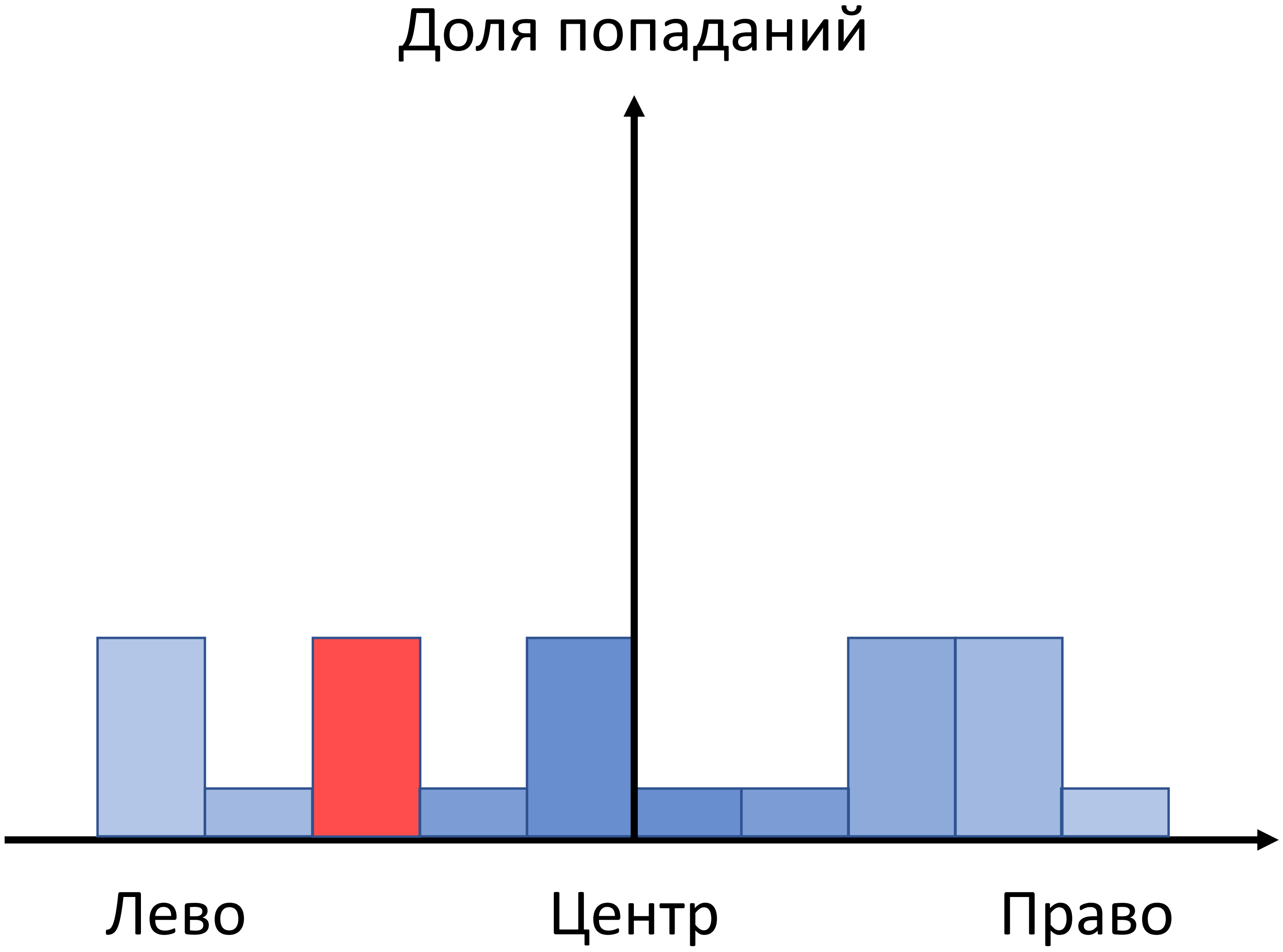
Распределения вероятностей



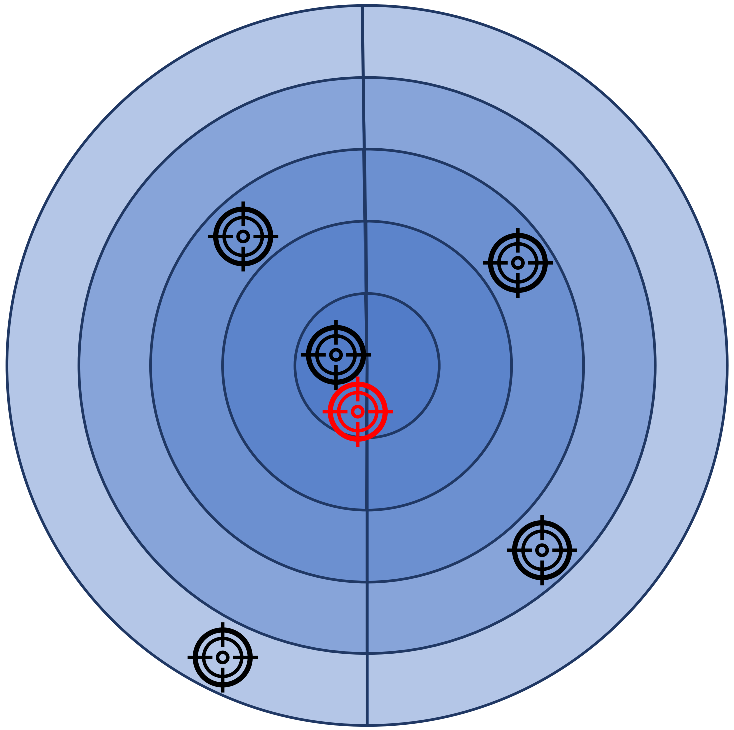


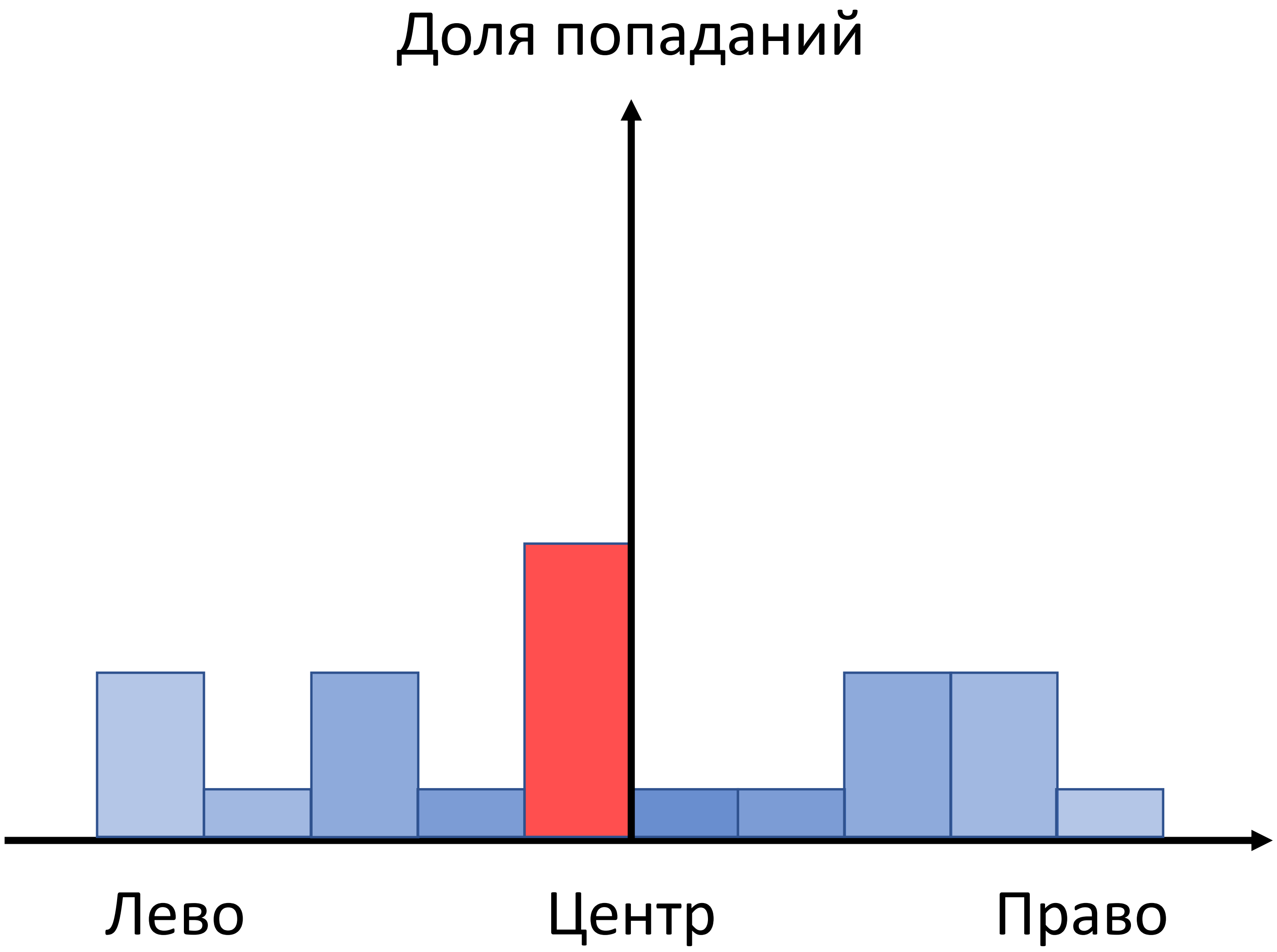
Распределения вероятностей



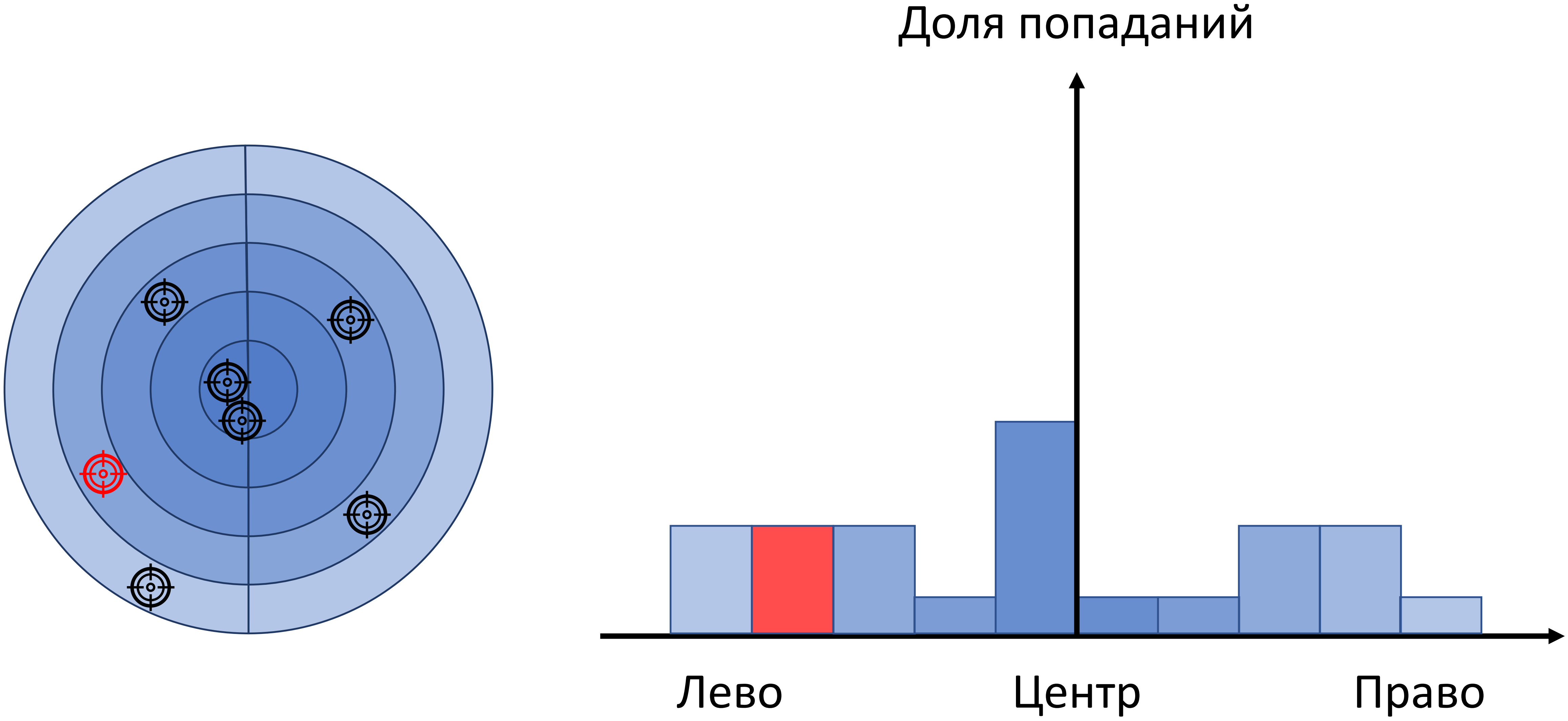


Распределения вероятностей

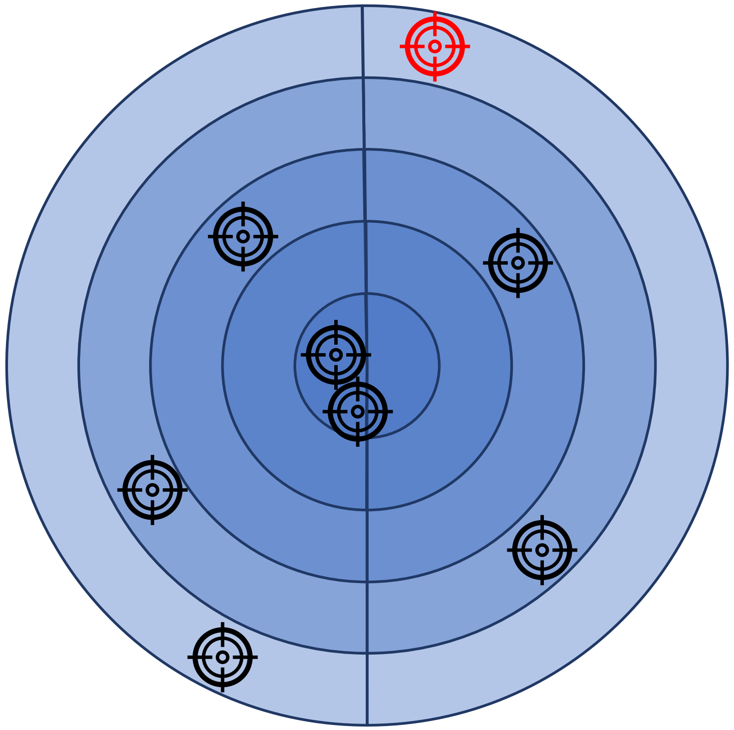


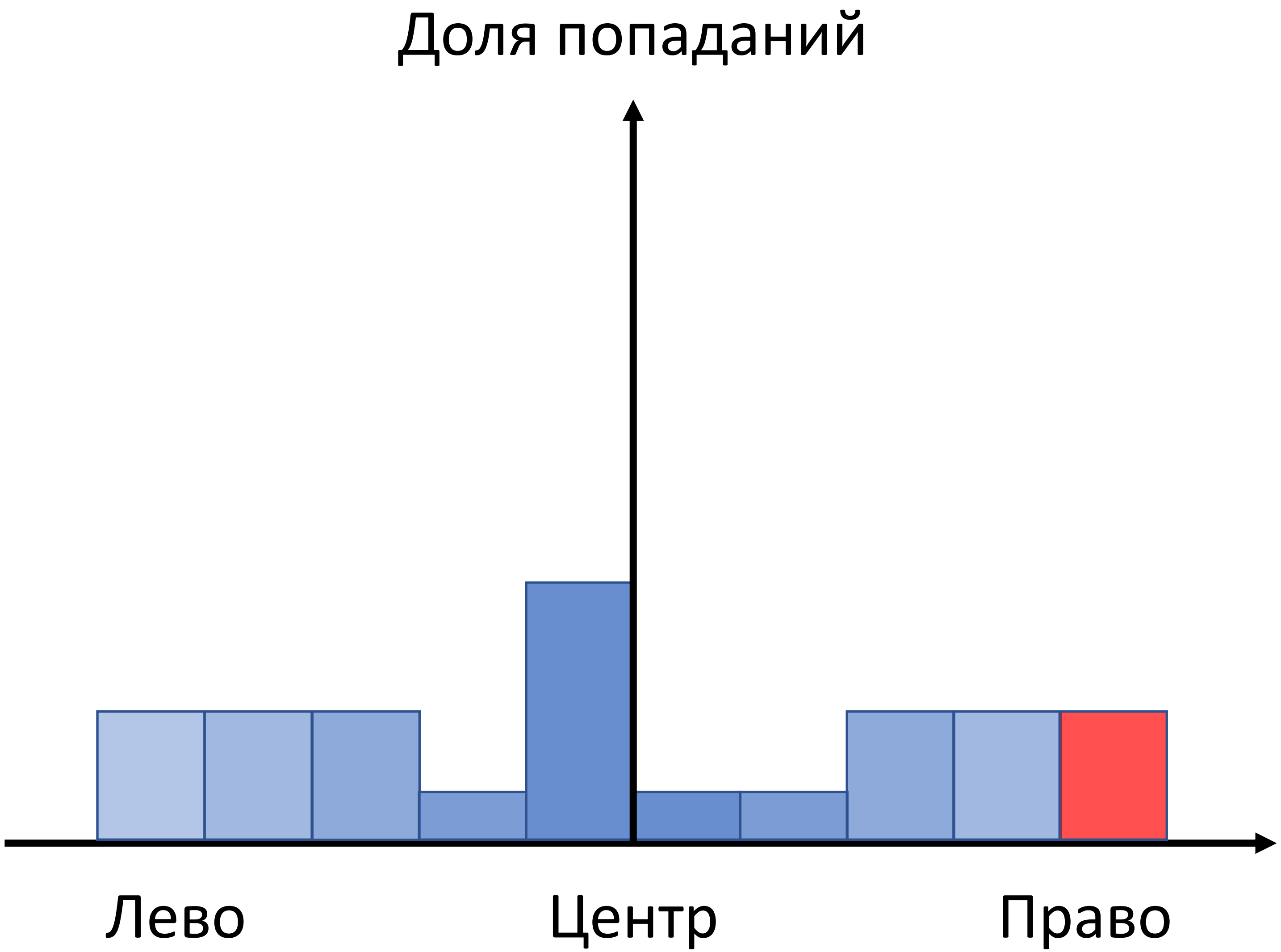


Распределения вероятностей

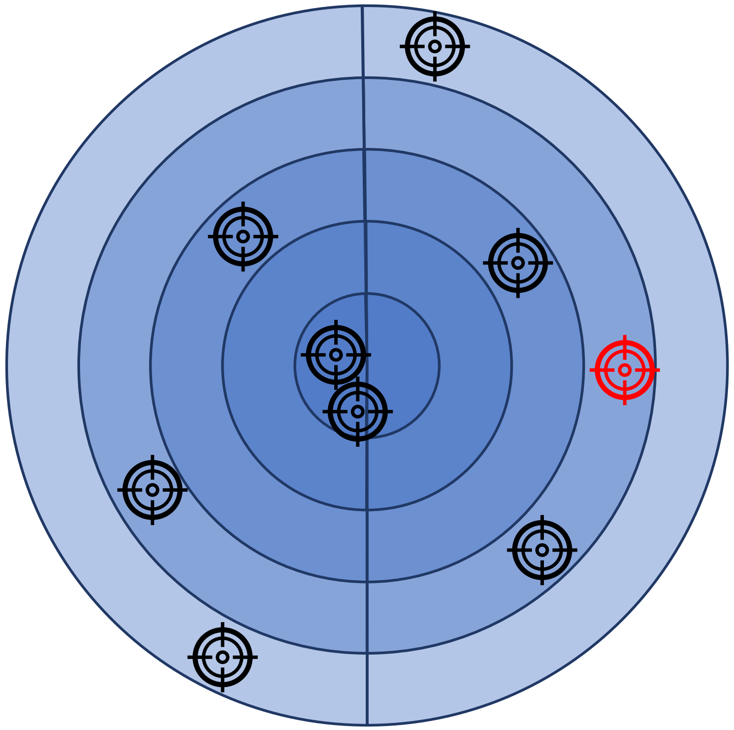


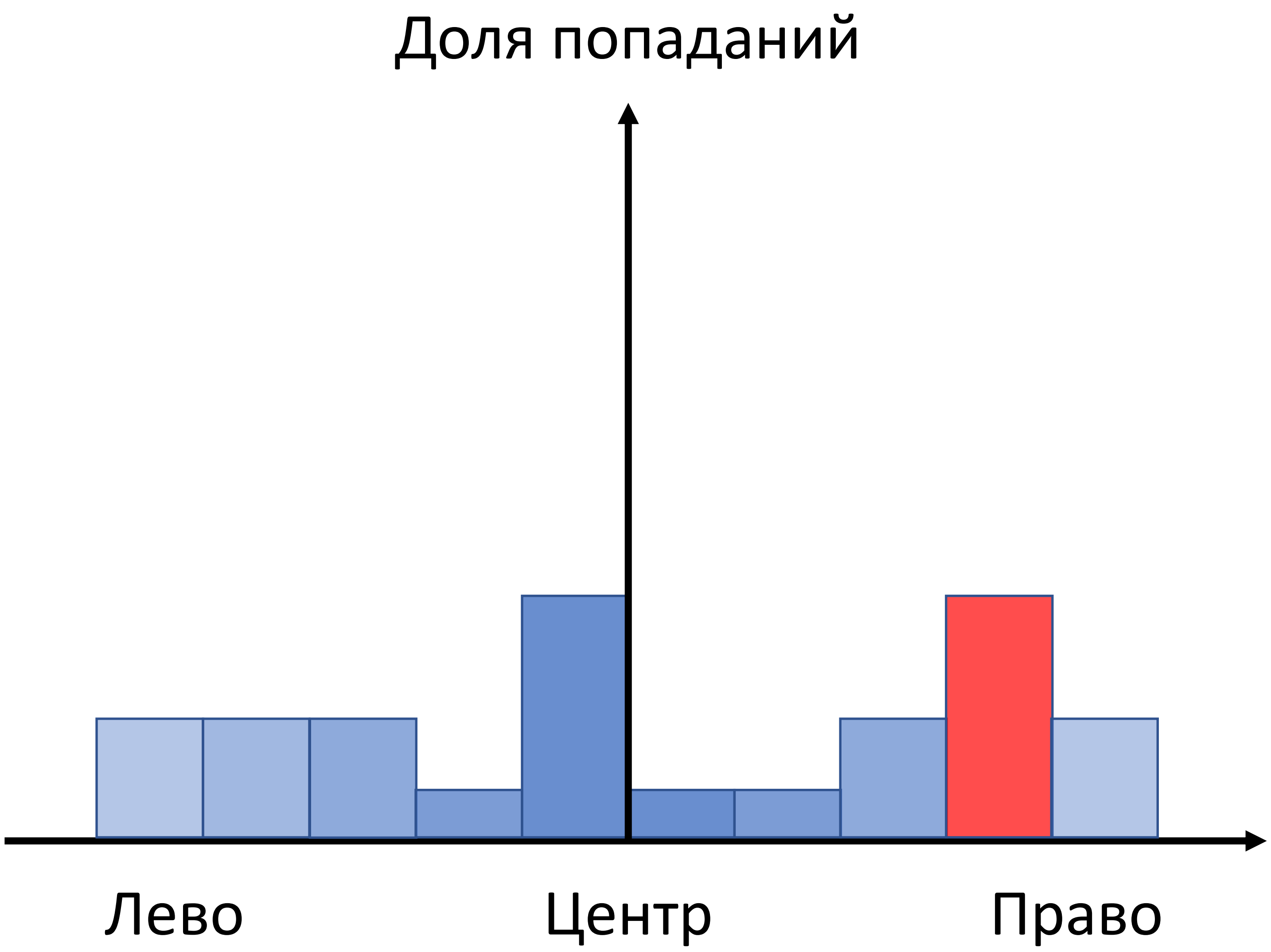
Распределения вероятностей



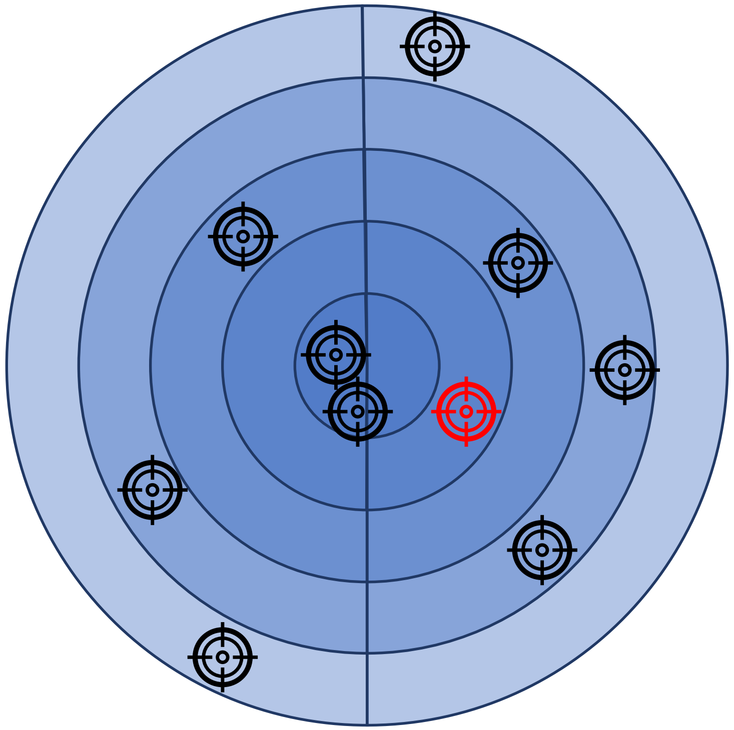


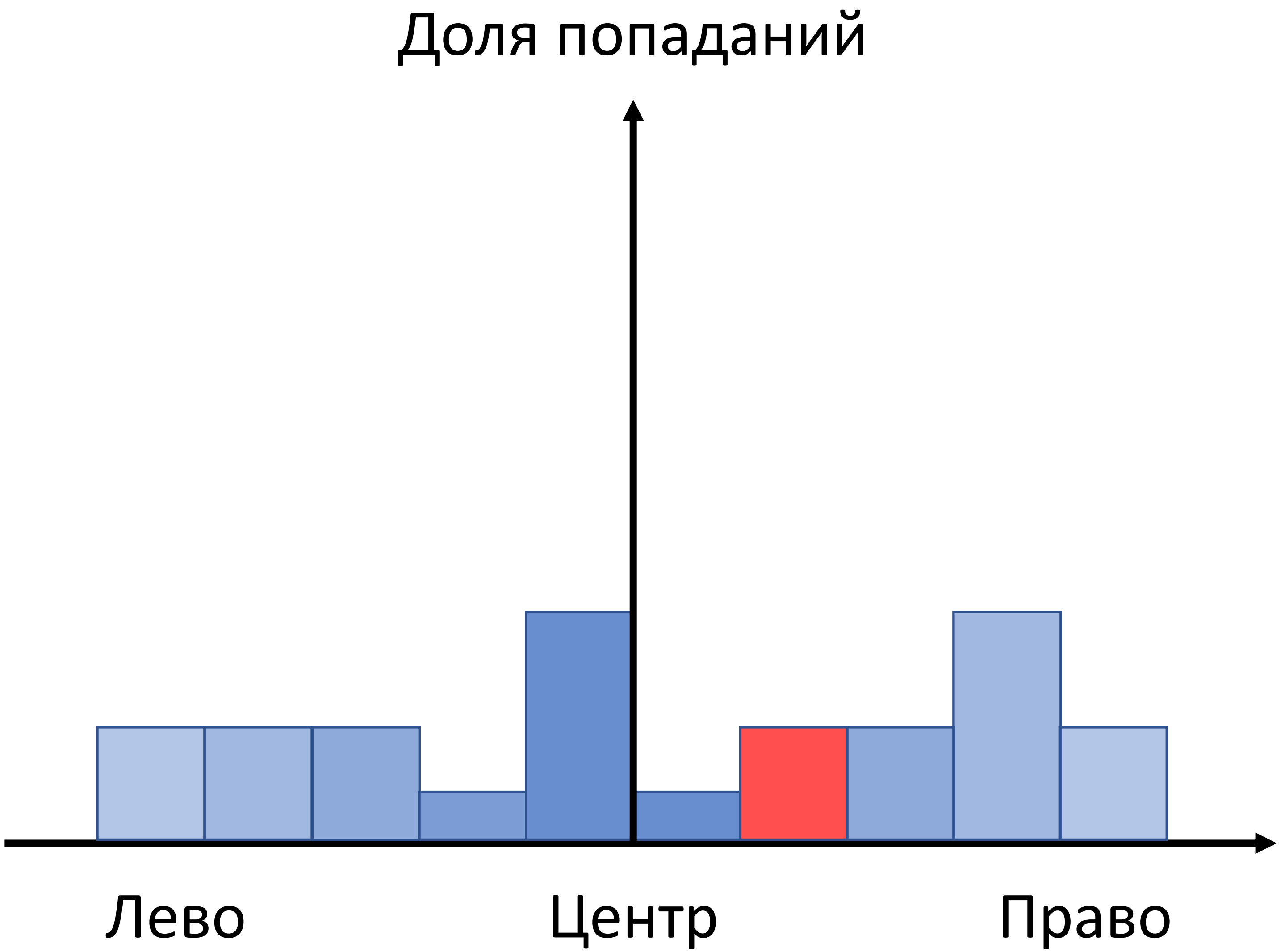
Распределения вероятностей



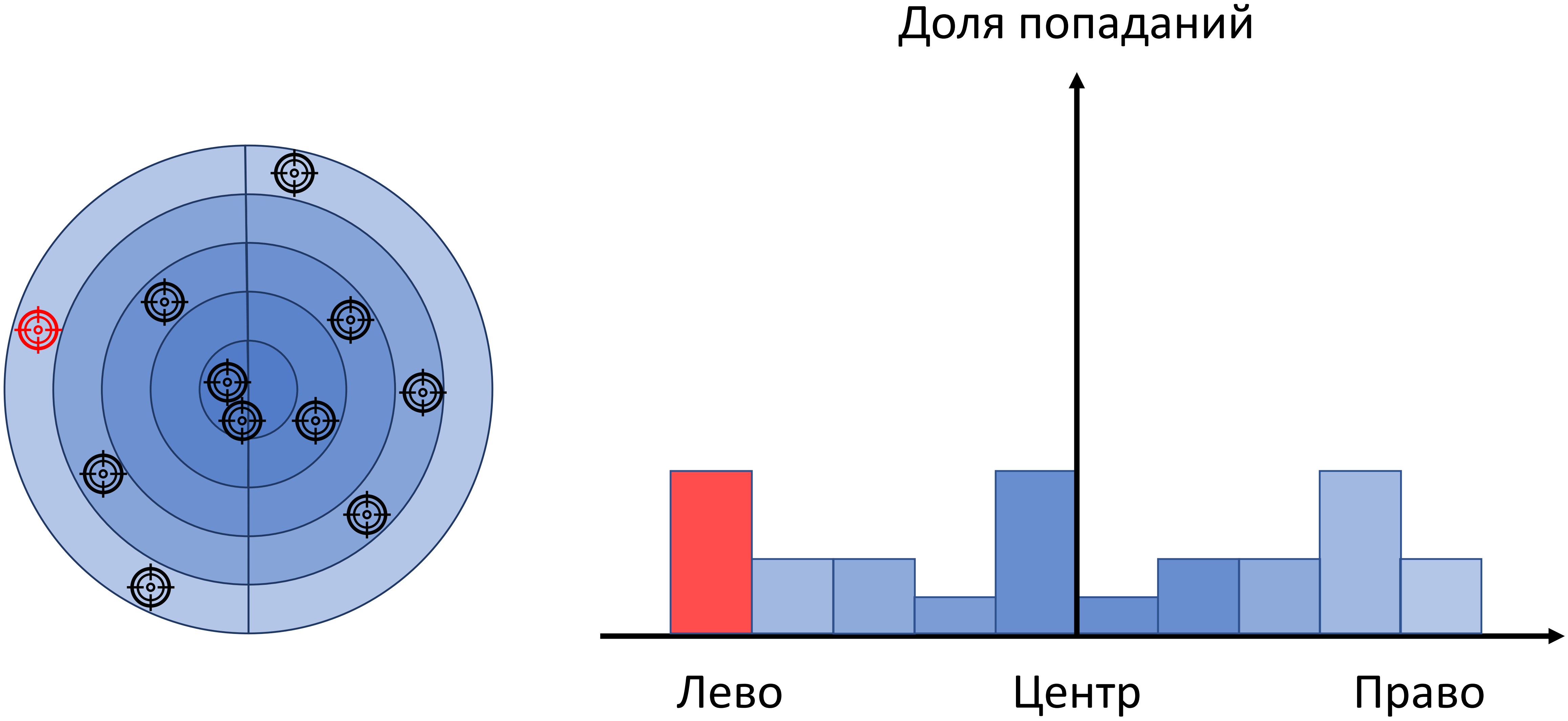


Распределения вероятностей

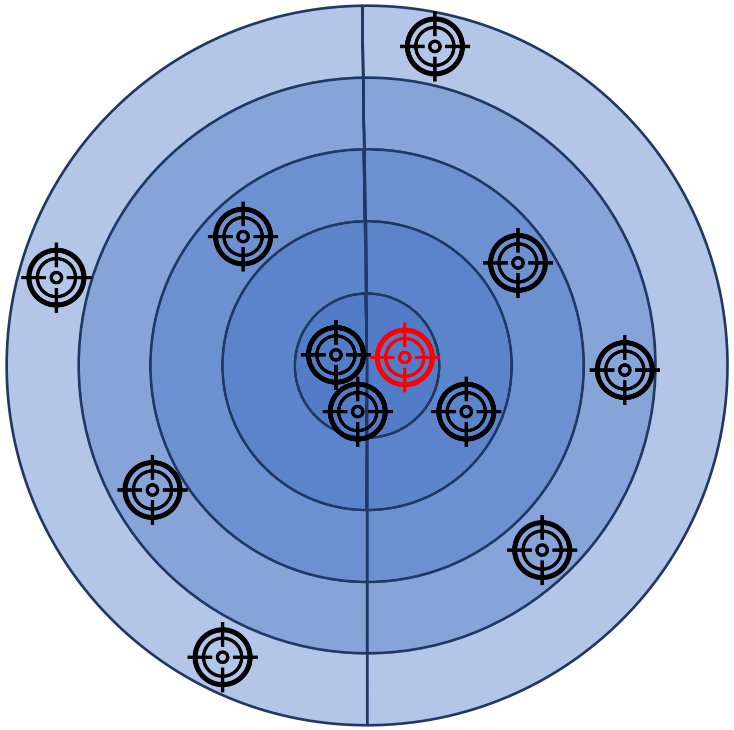


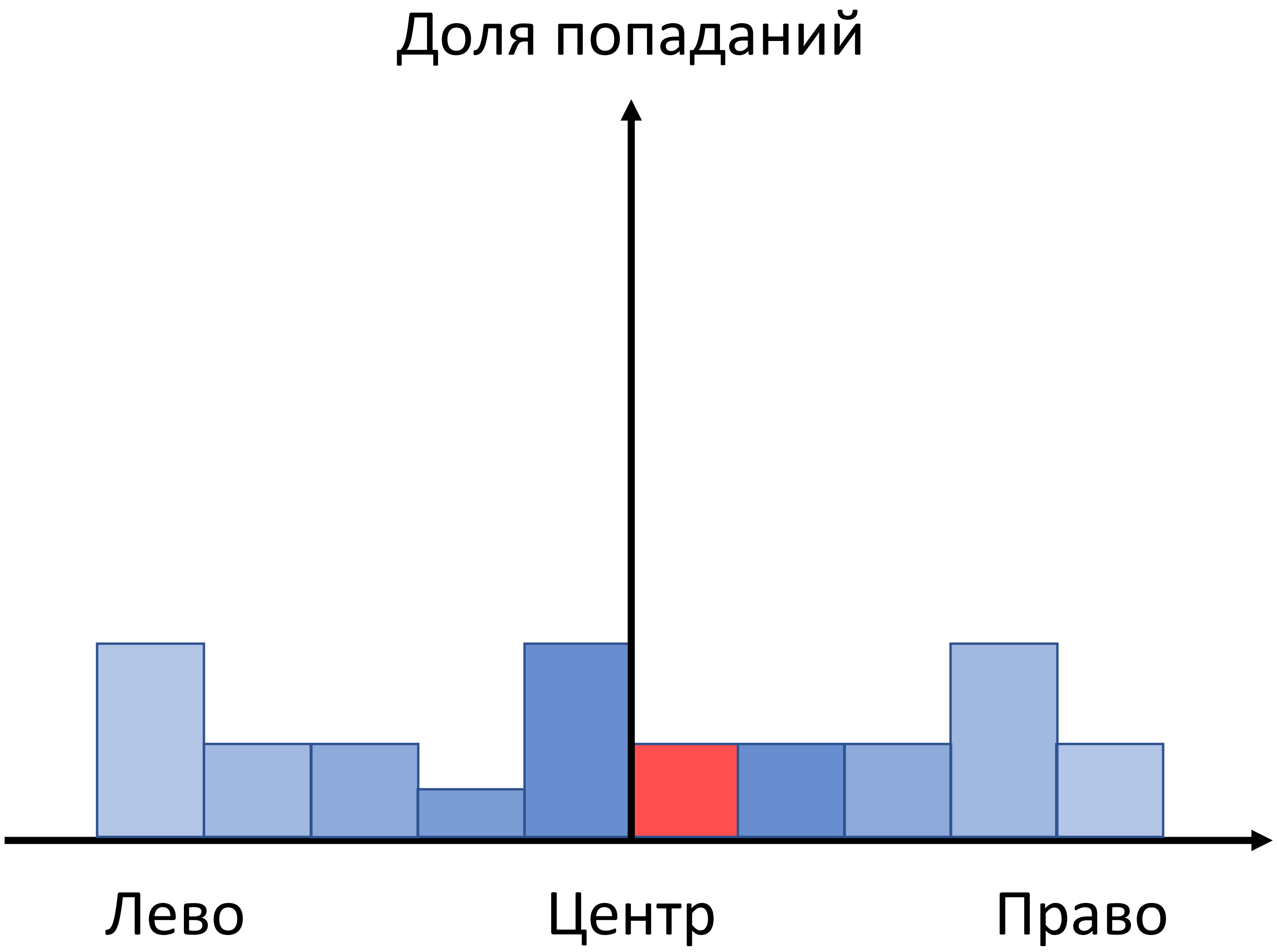


Распределения вероятностей

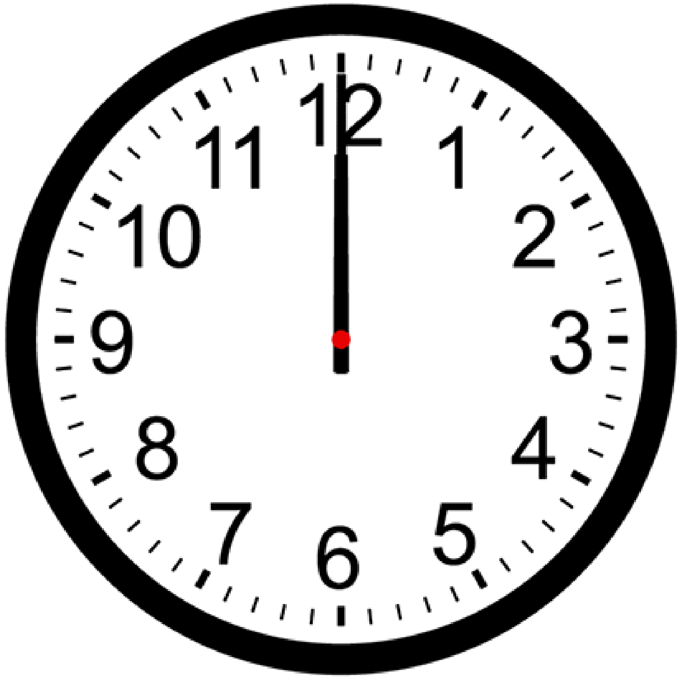


Распределения вероятностей

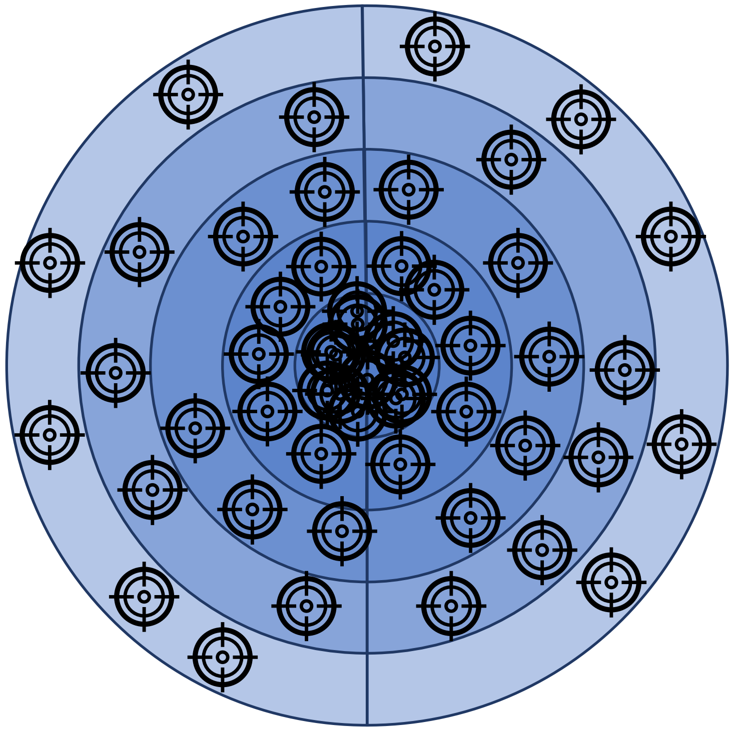


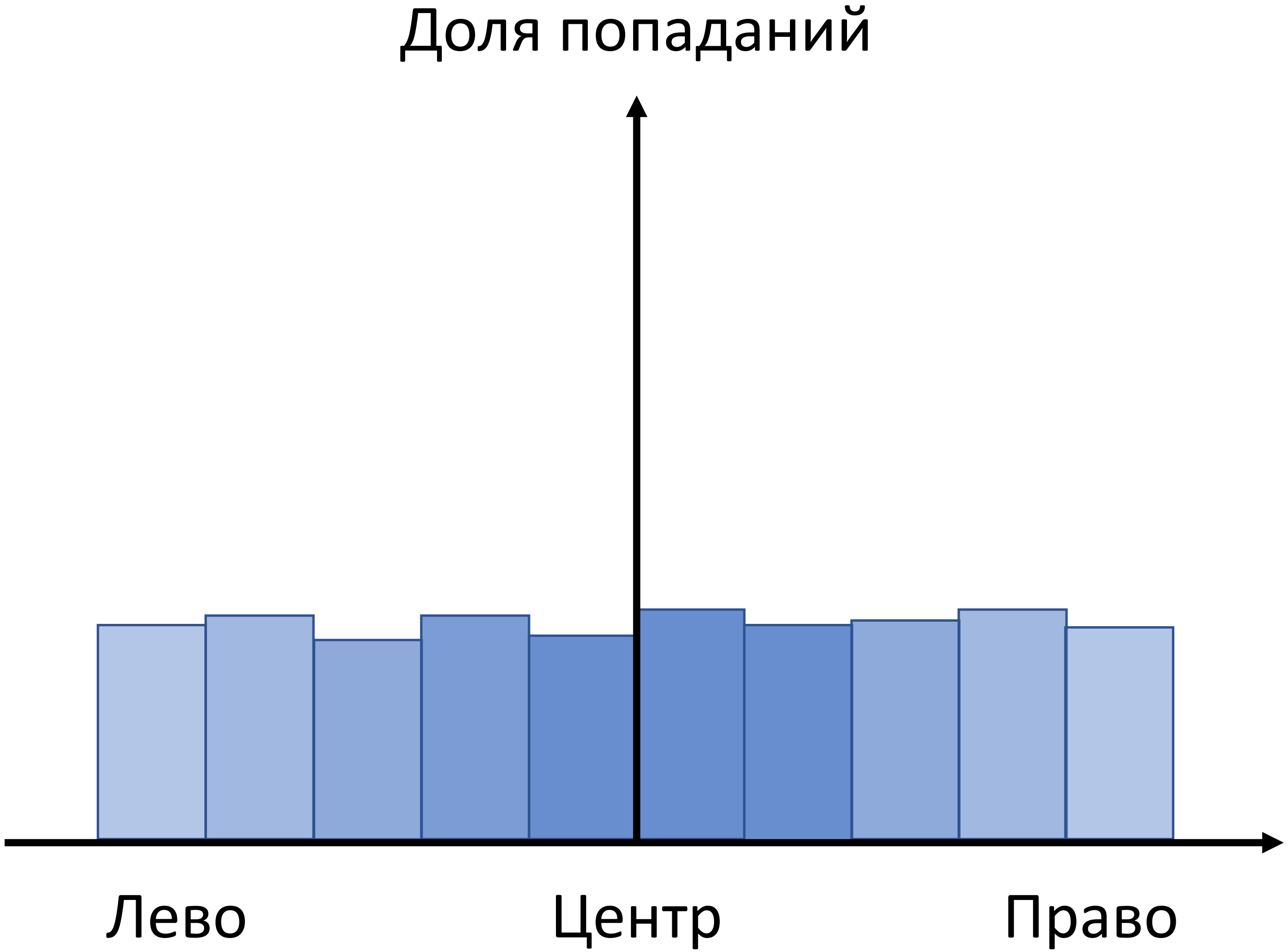


Распределения вероятностей

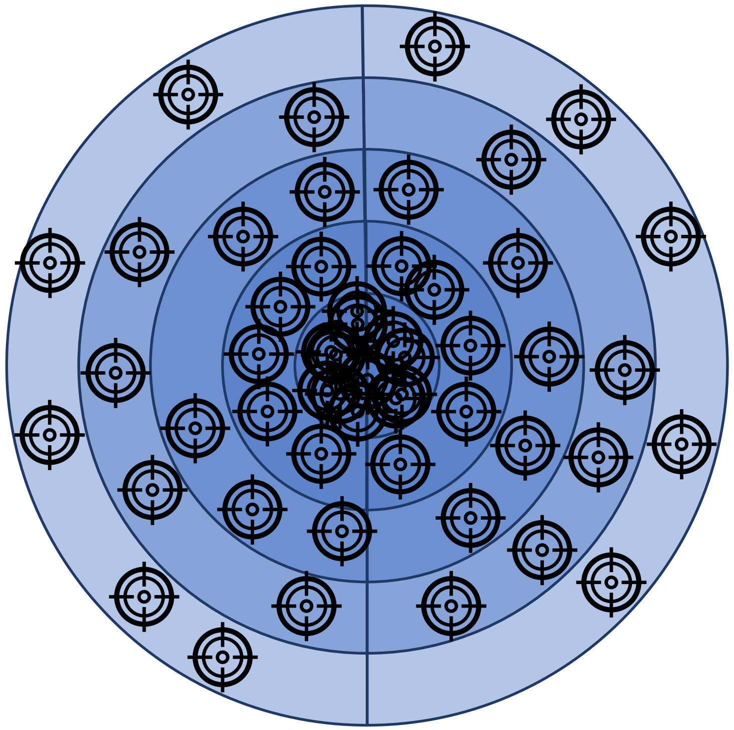


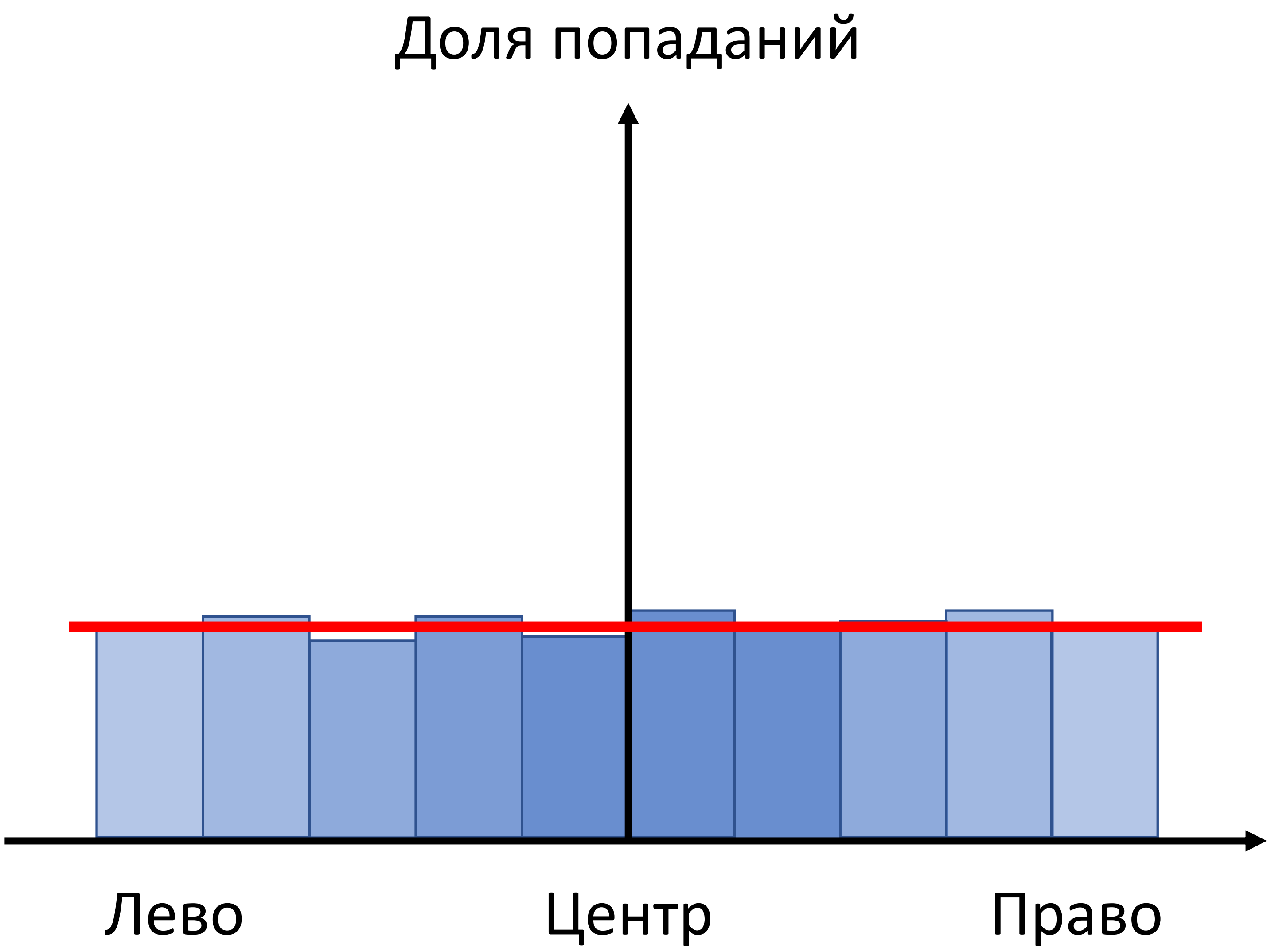
Распределения вероятностей



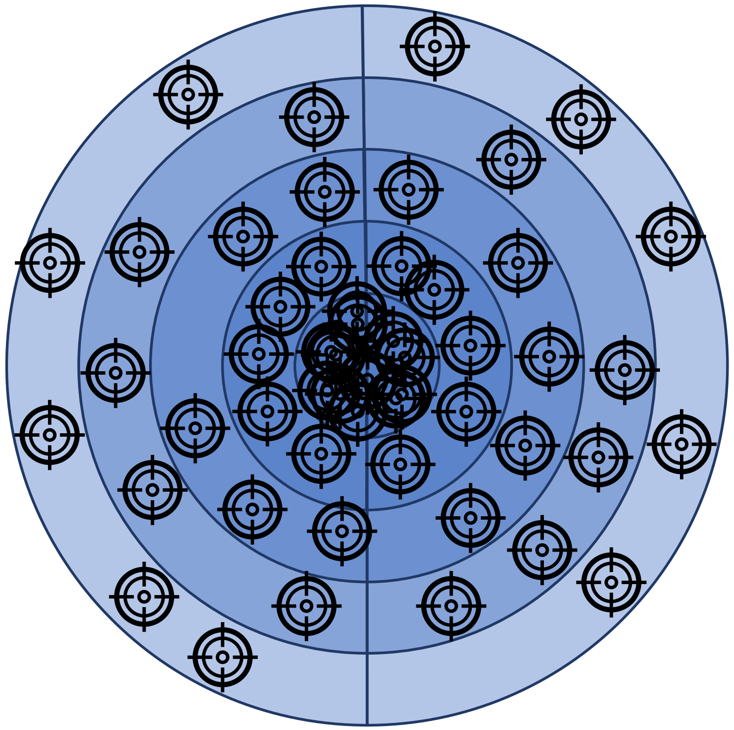


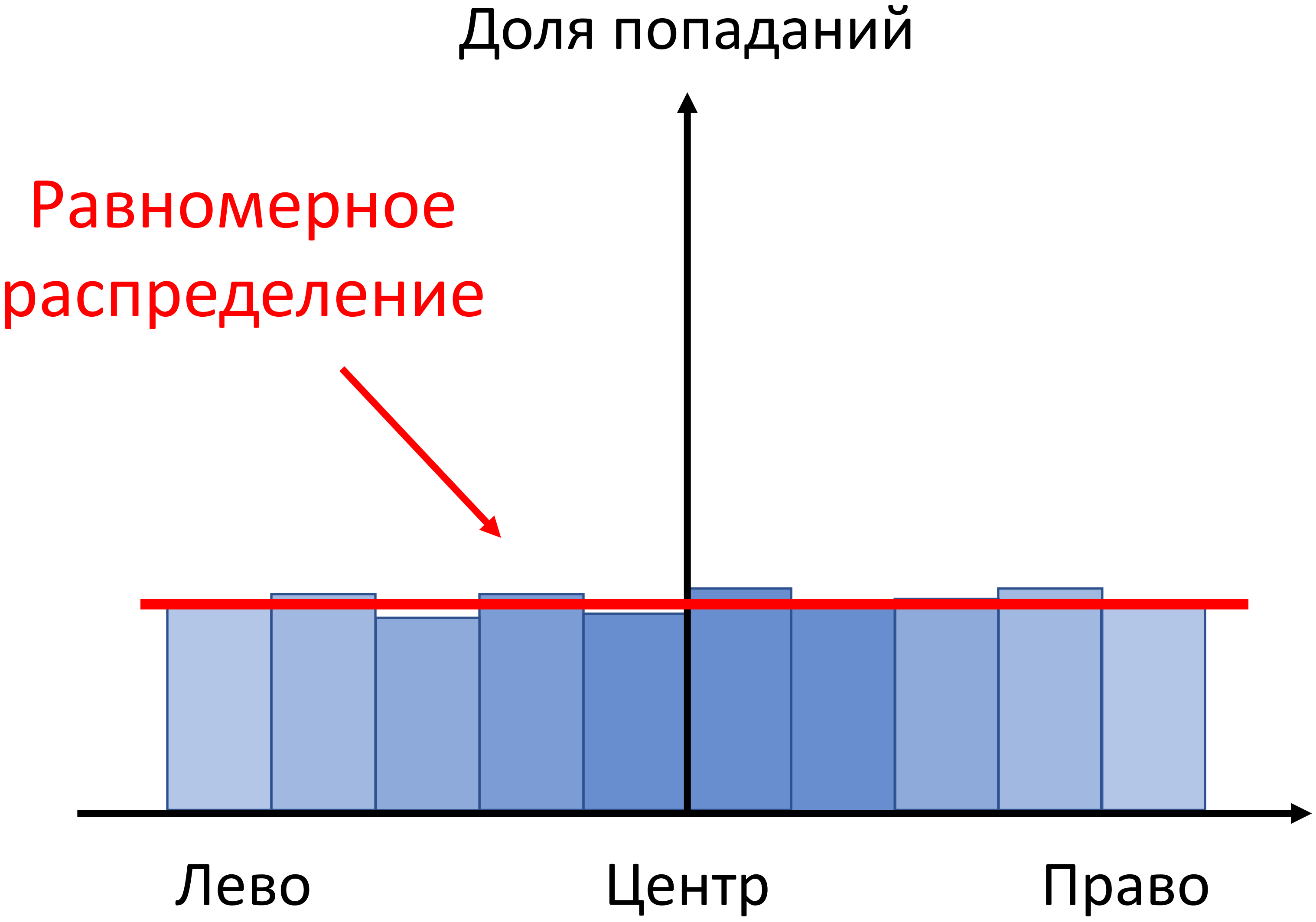
Распределения вероятностей



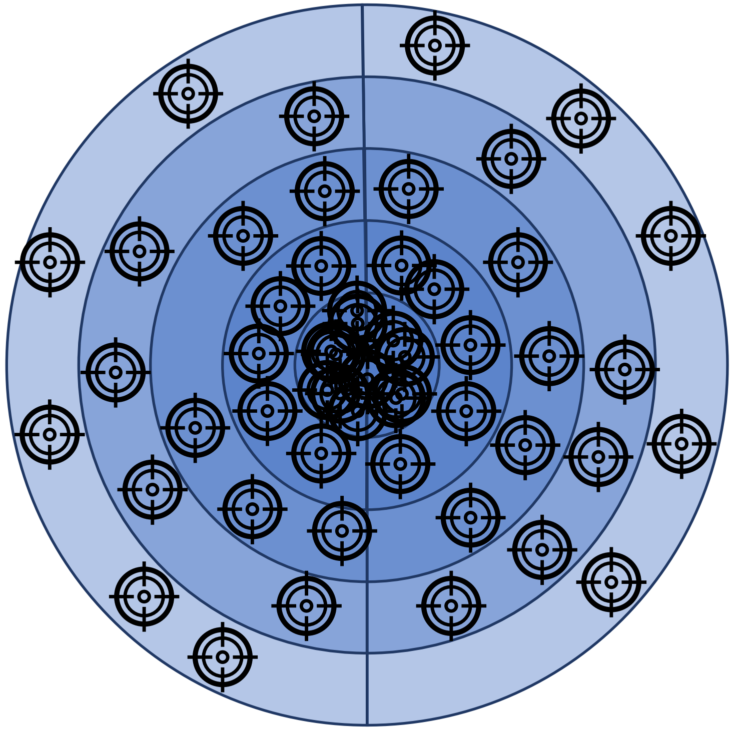


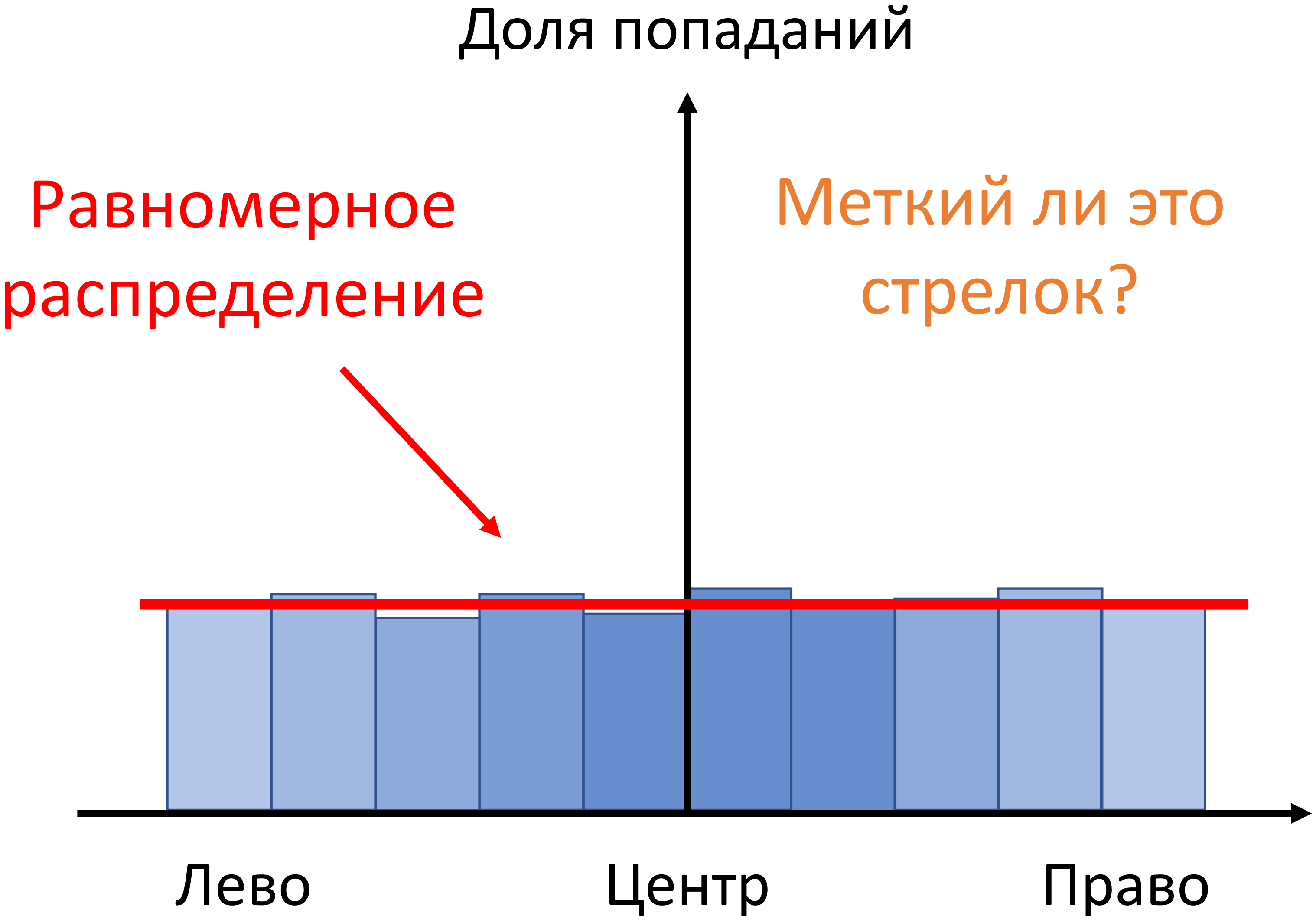
Распределения вероятностей



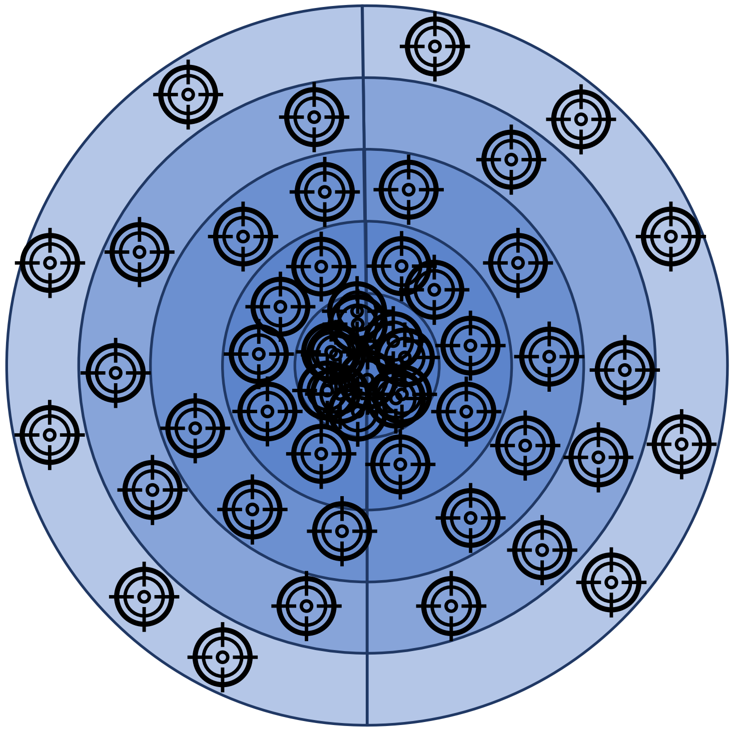


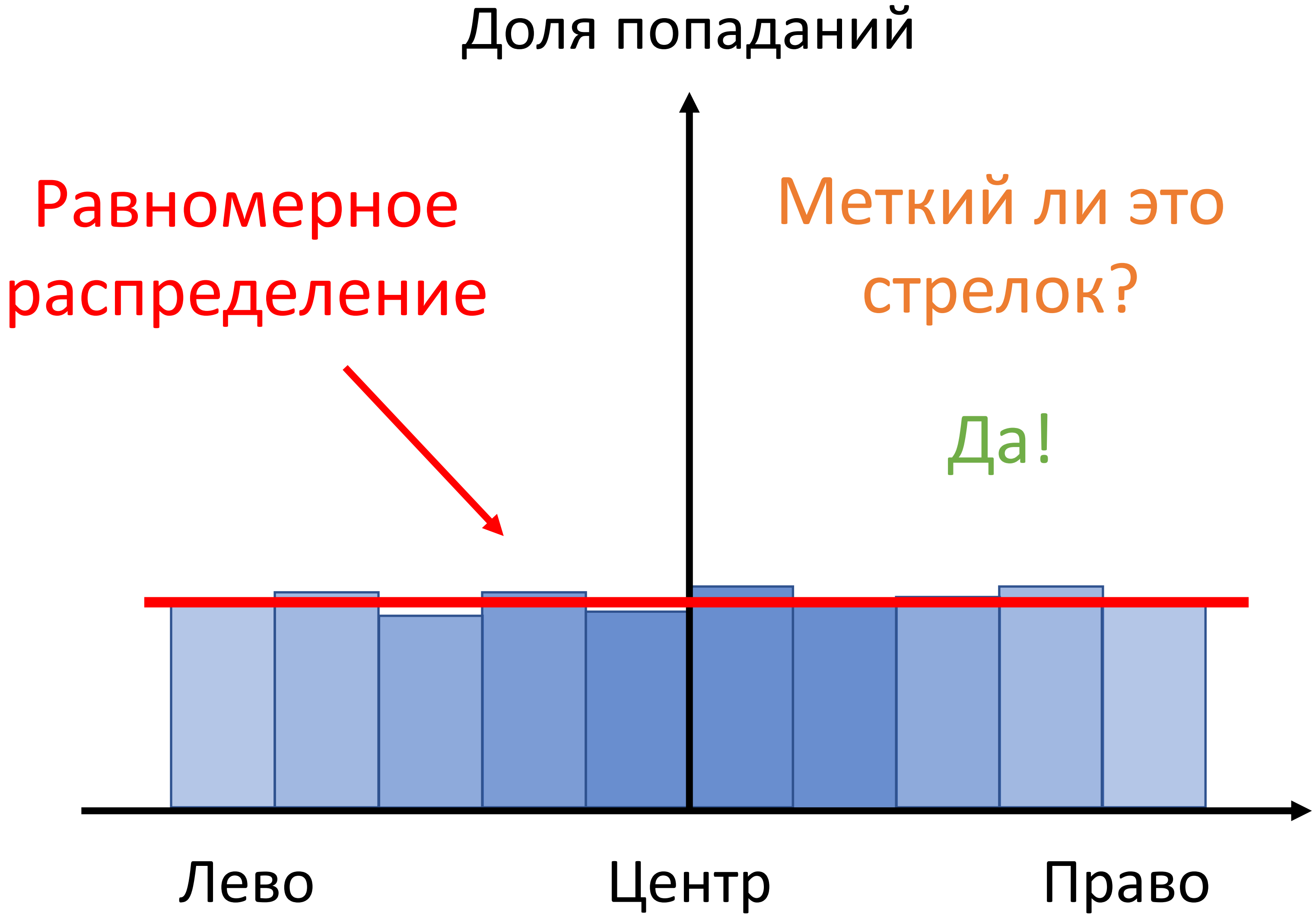
Распределения вероятностей



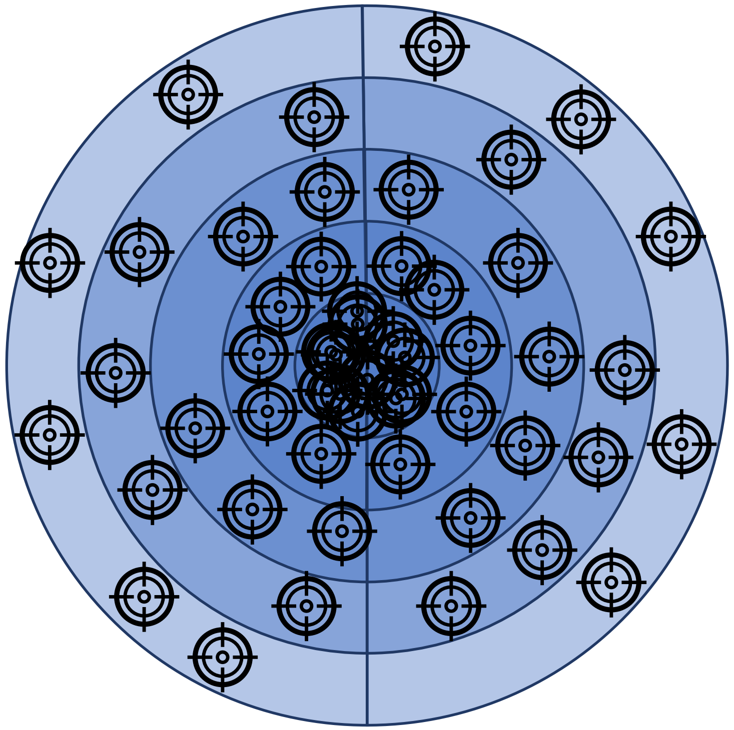


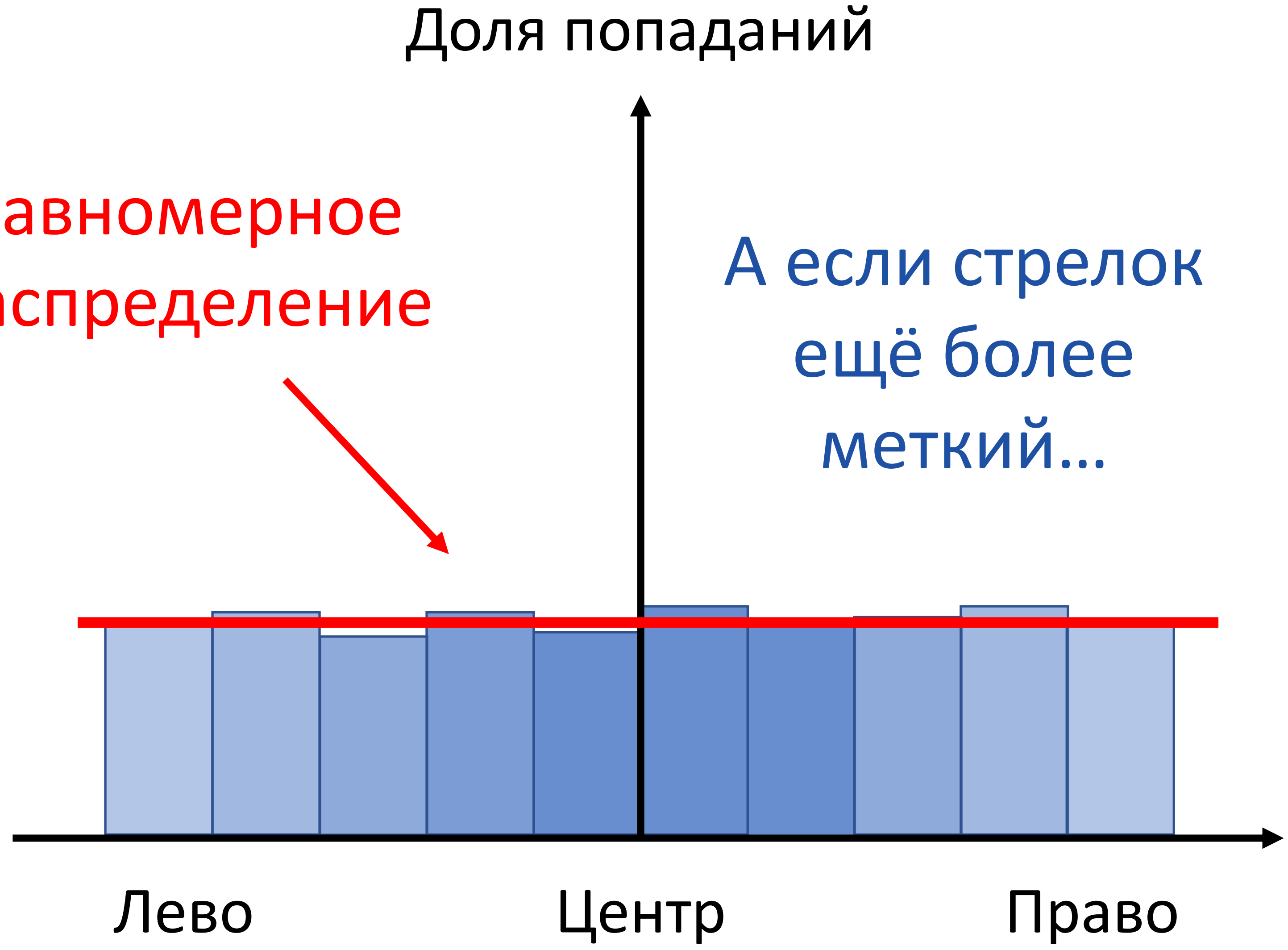
Распределения вероятностей





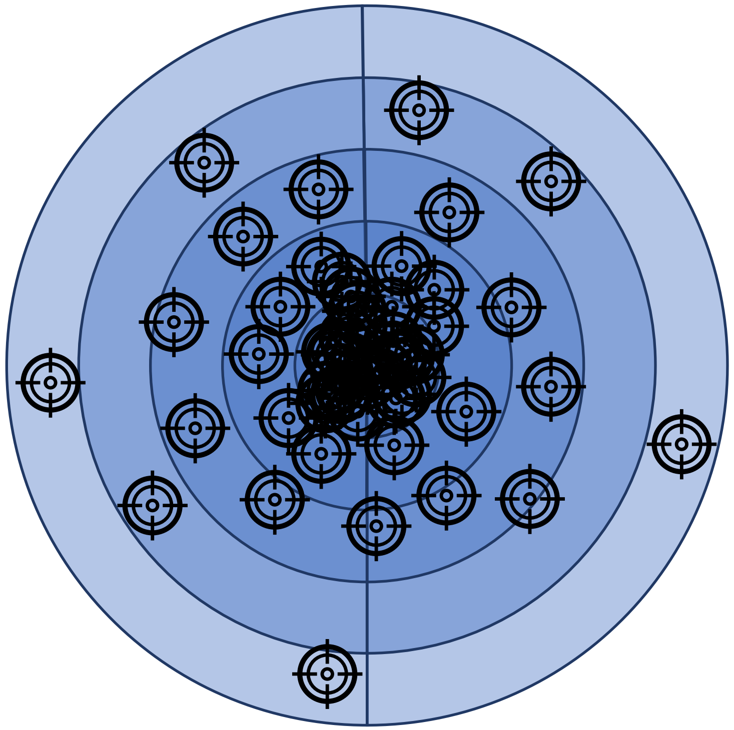
Распределения вероятностей

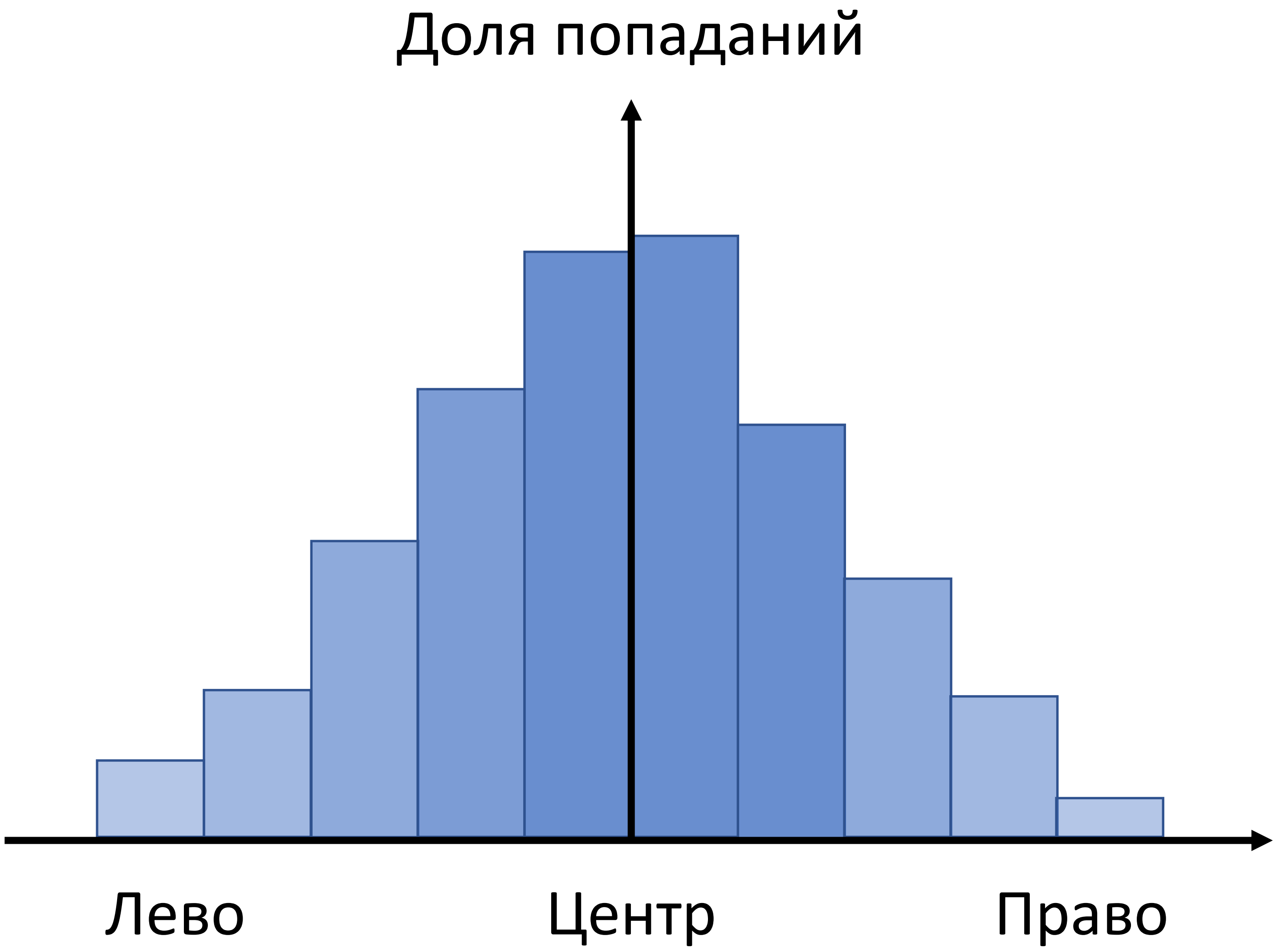




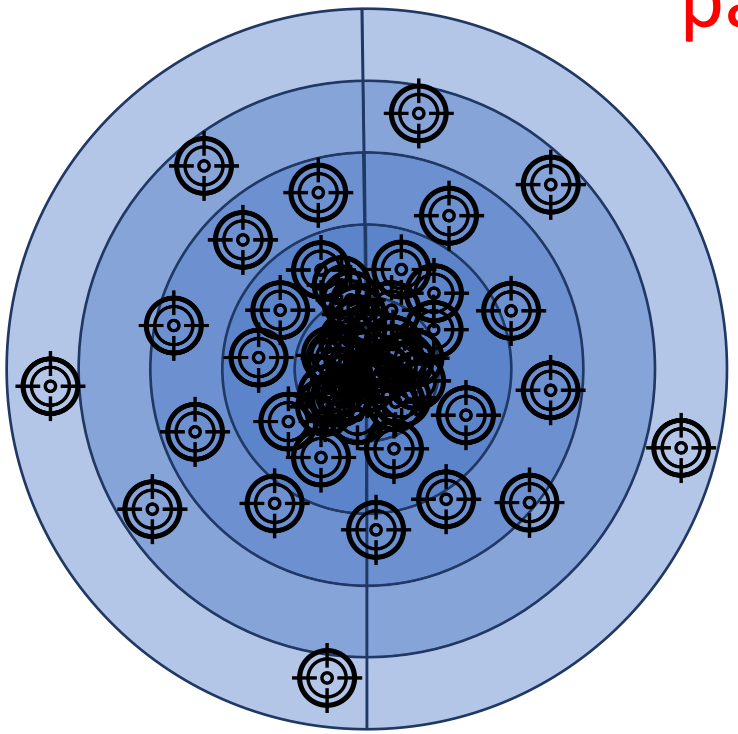
А если стрелок ещё более меткий…

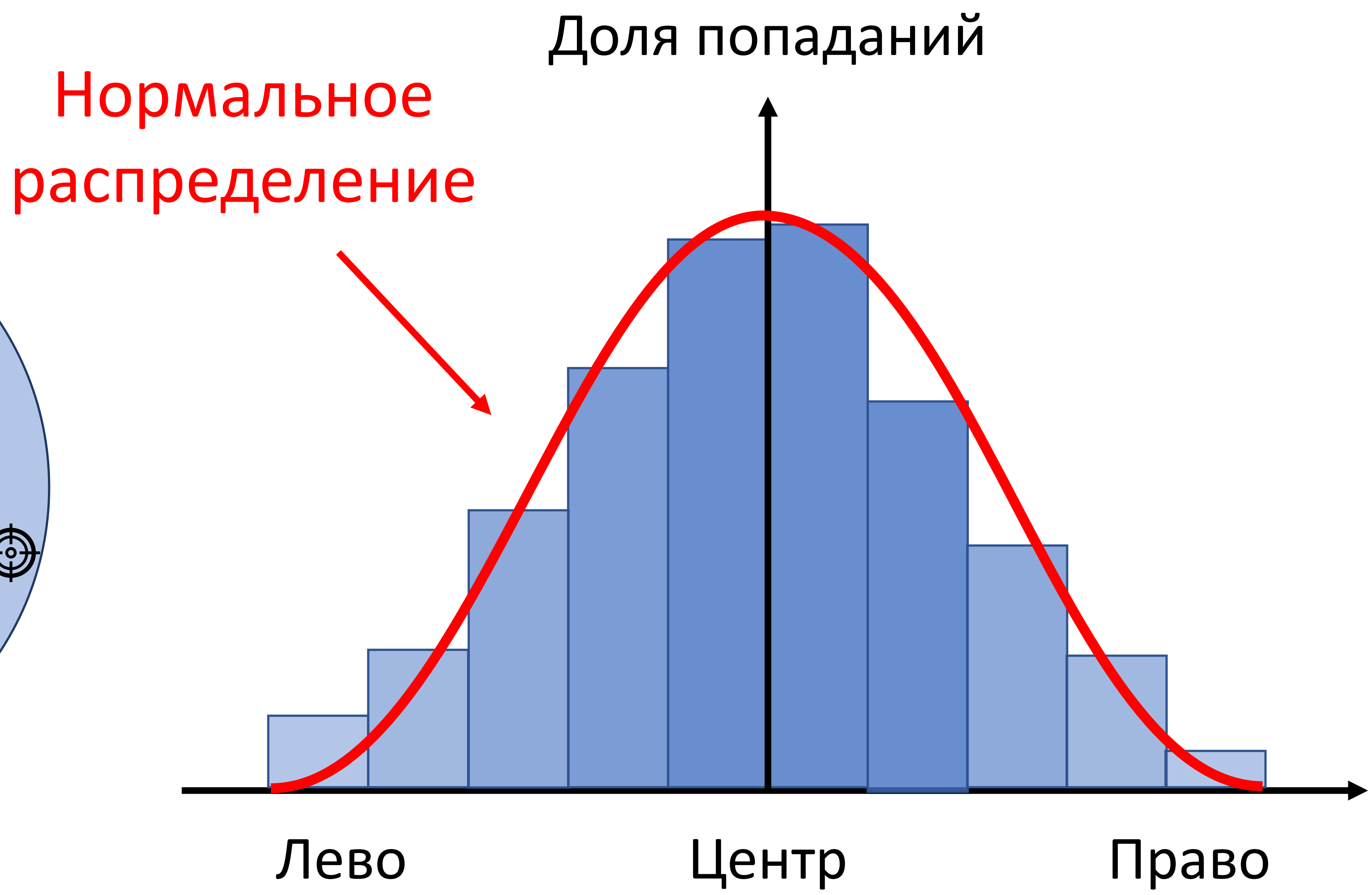
Распределения вероятностей





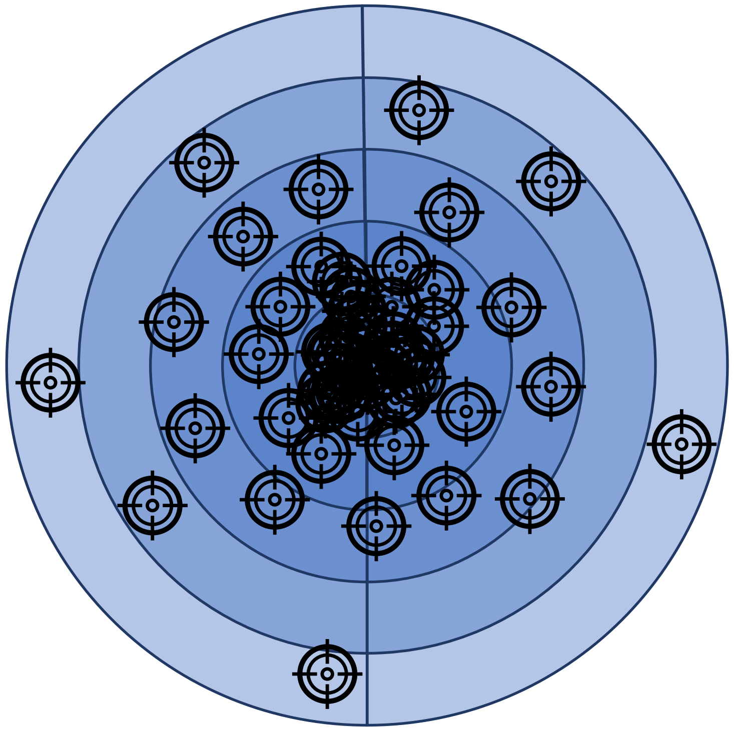
Распределения вероятностей

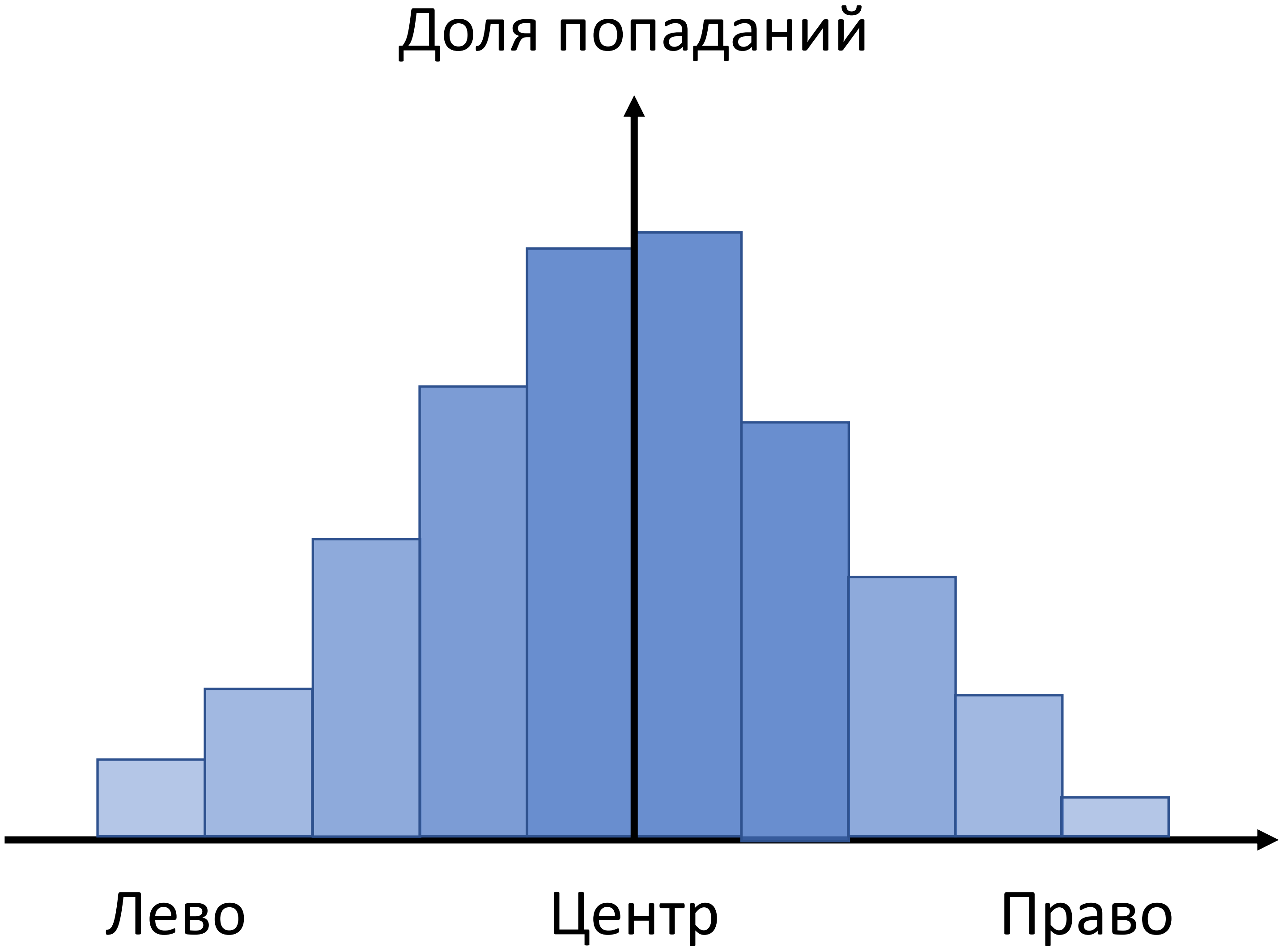




Распределения вероятностей

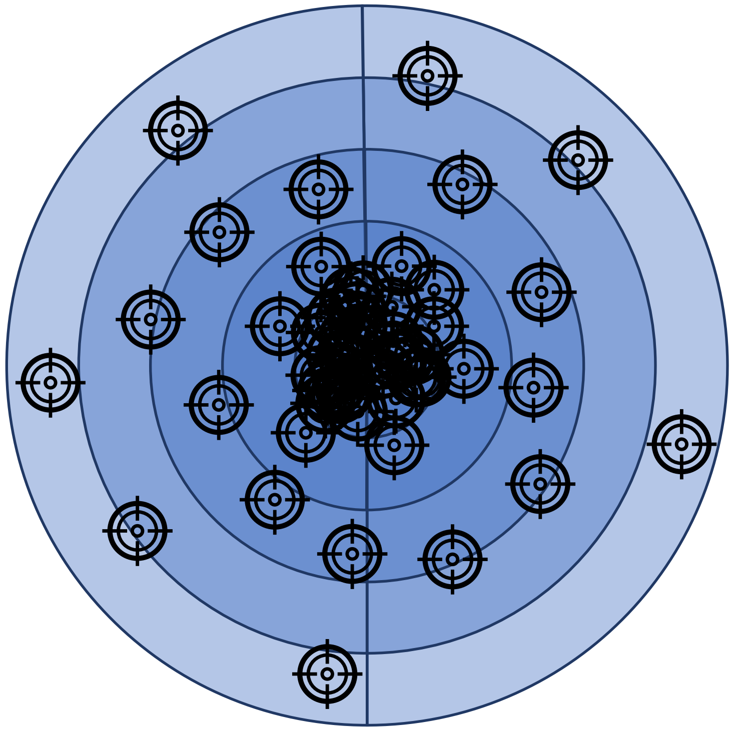
Или чуть иначе…

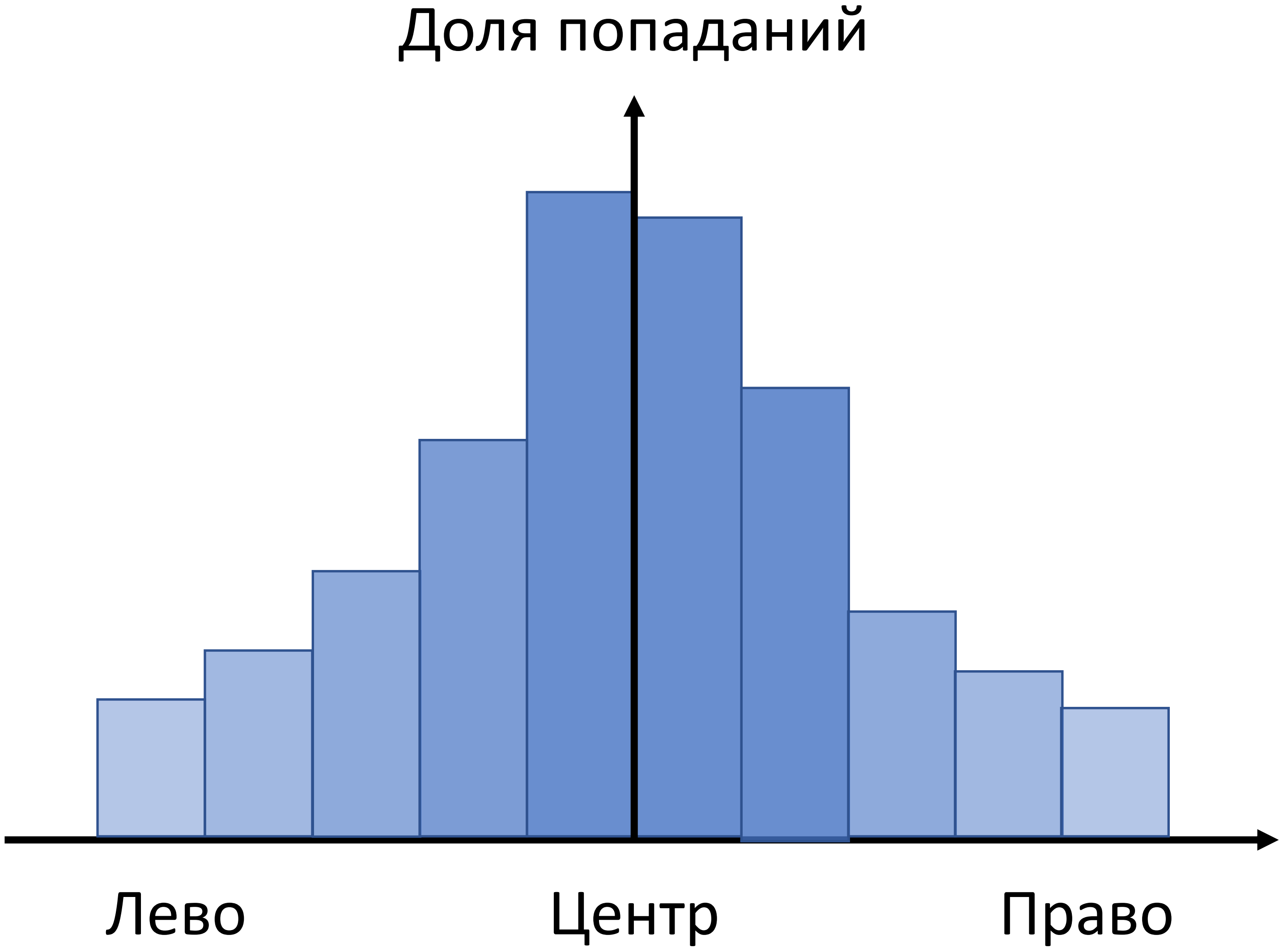




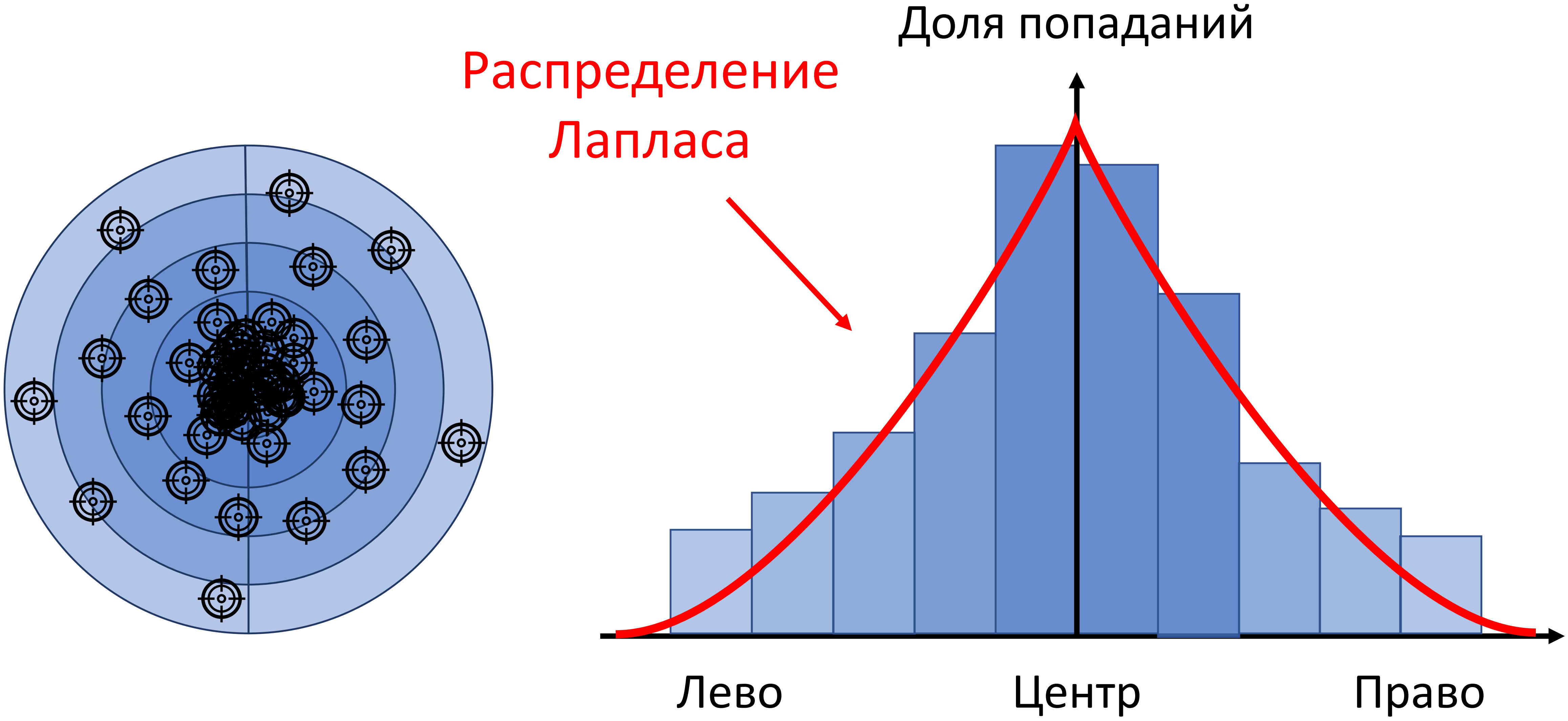
Распределения вероятностей

Или чуть иначе…





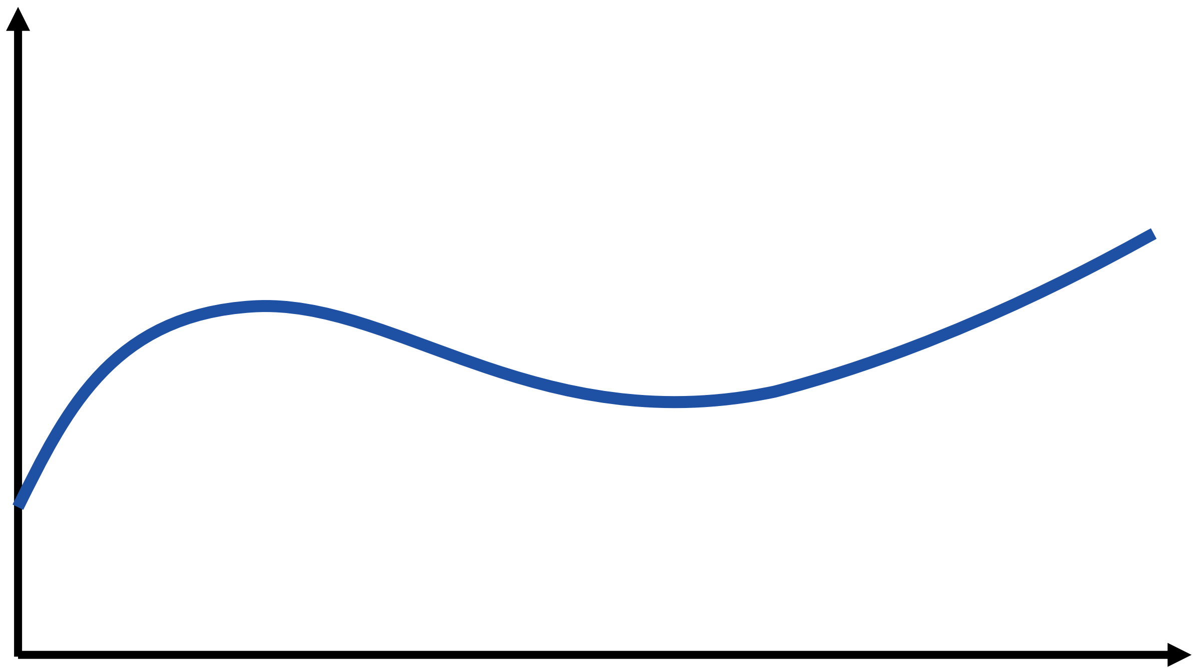
Распределения вероятностей



Распределения отклонений

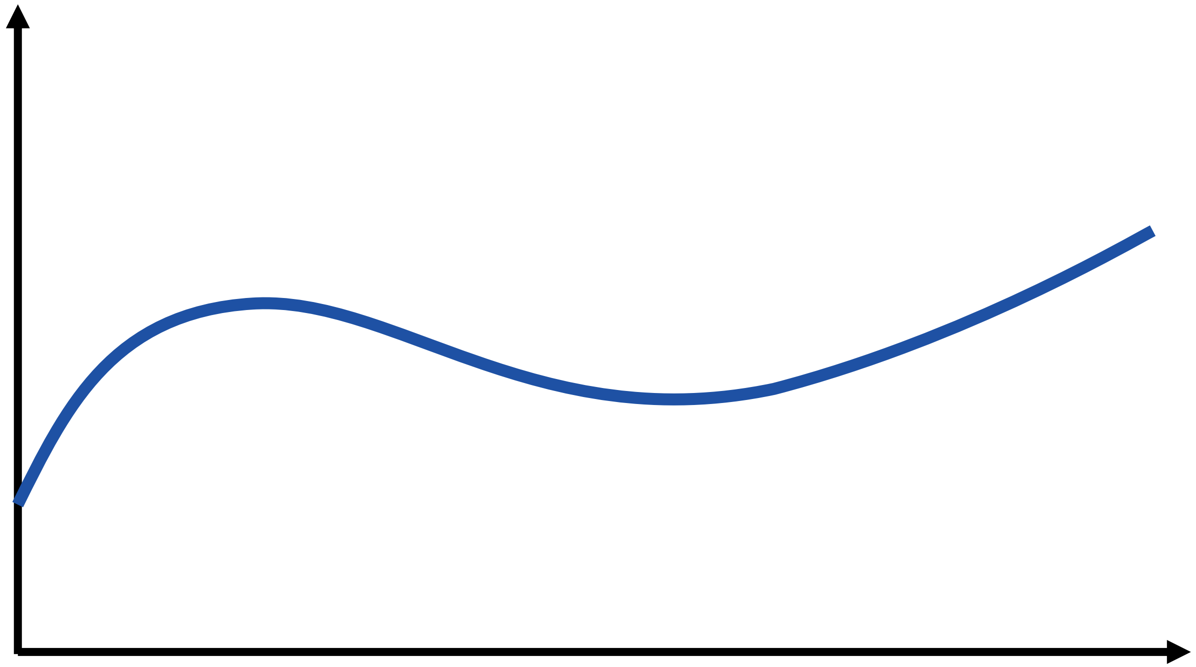
Распределения отклонений

Есть истинная функция

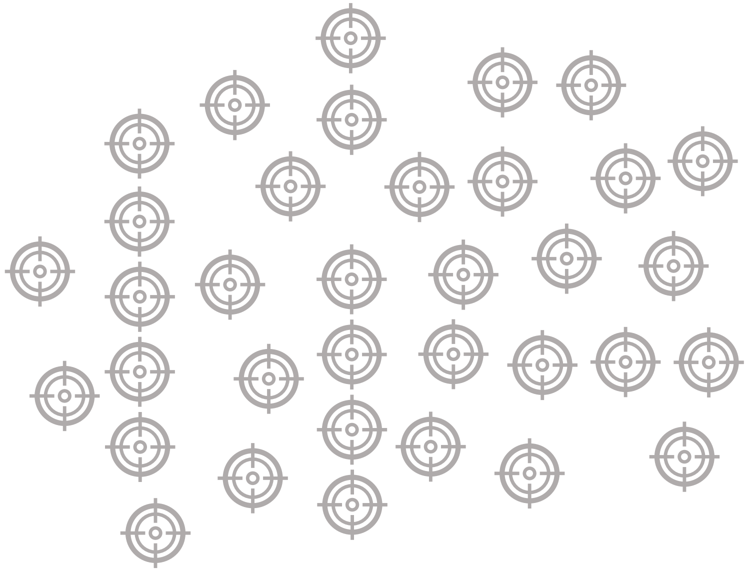


Распределения отклонений

Есть истинная функция

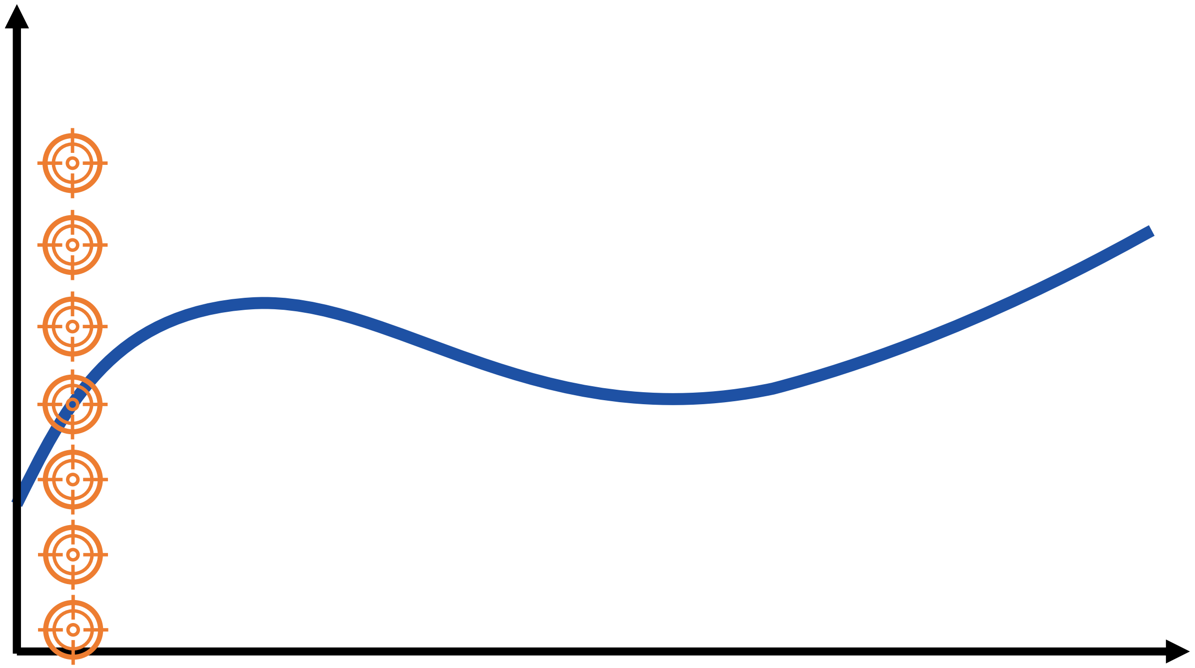


Делаем много неточных измерений – «выстрелов»

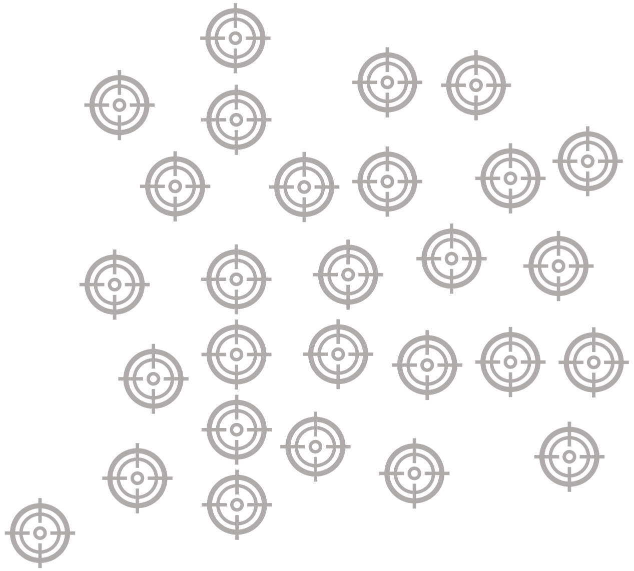


Распределения отклонений

Есть истинная функция

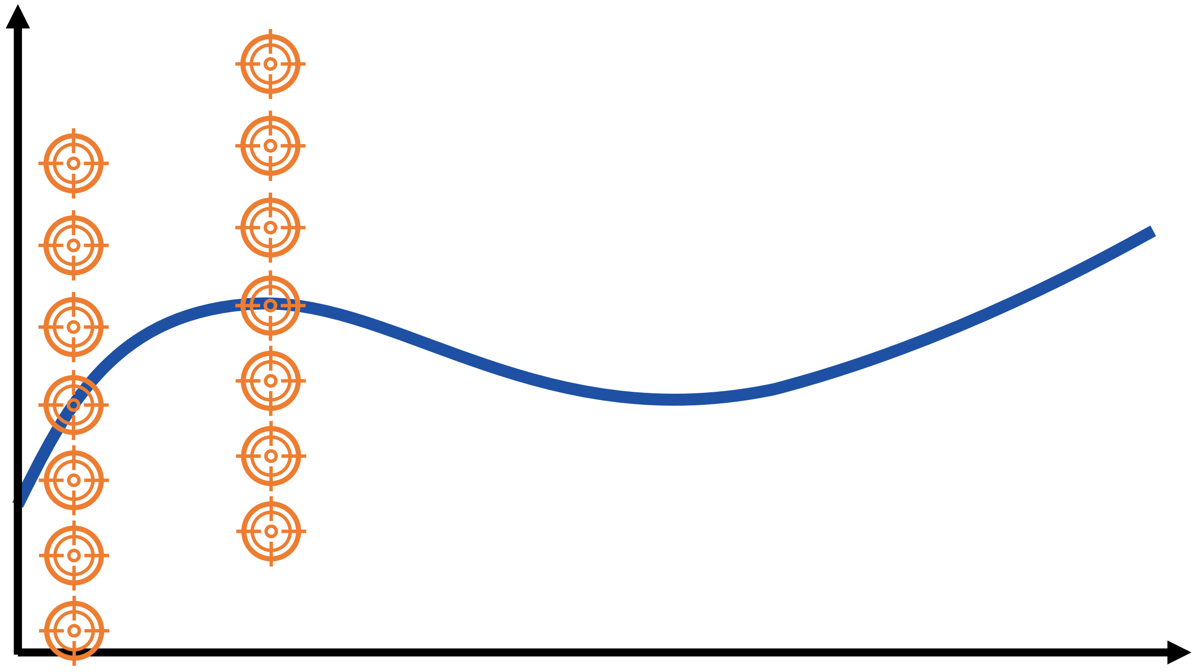


Делаем много неточных измерений – «выстрелов»

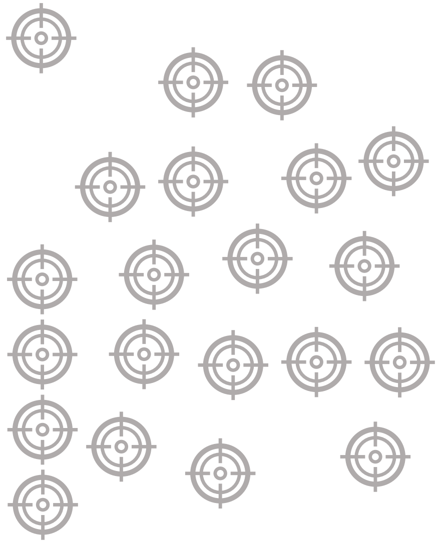


Распределения отклонений

Есть истинная функция

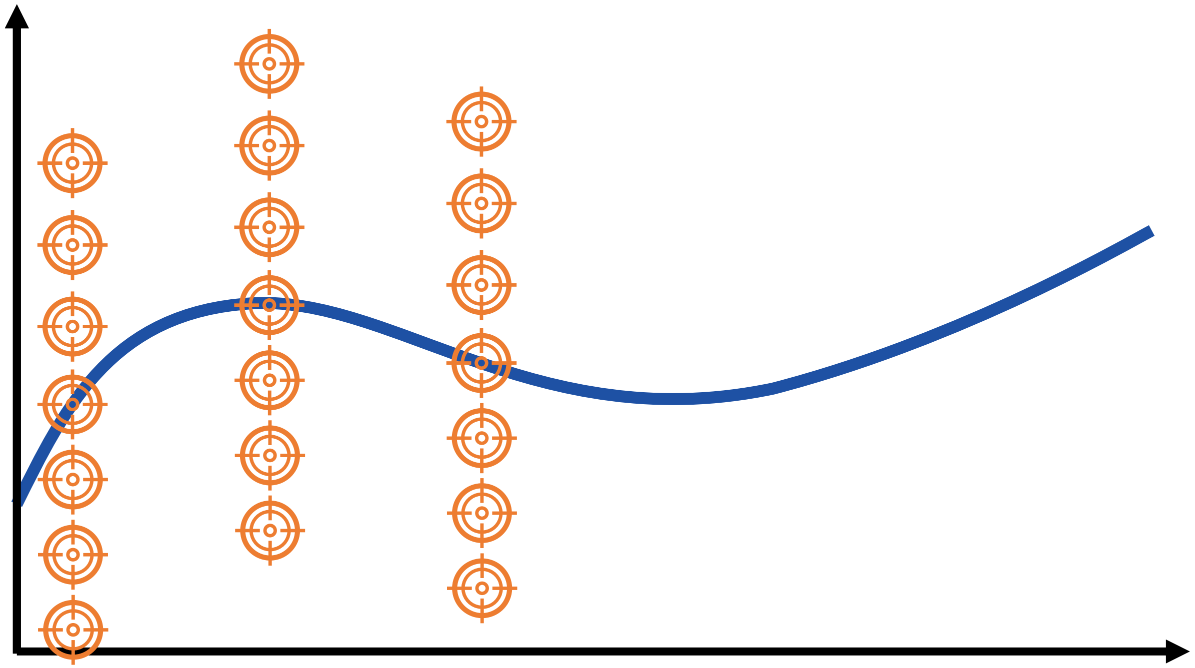


Делаем много неточных измерений – «выстрелов»

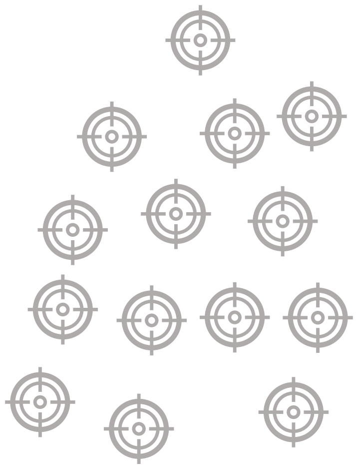


Распределения отклонений

Есть истинная функция

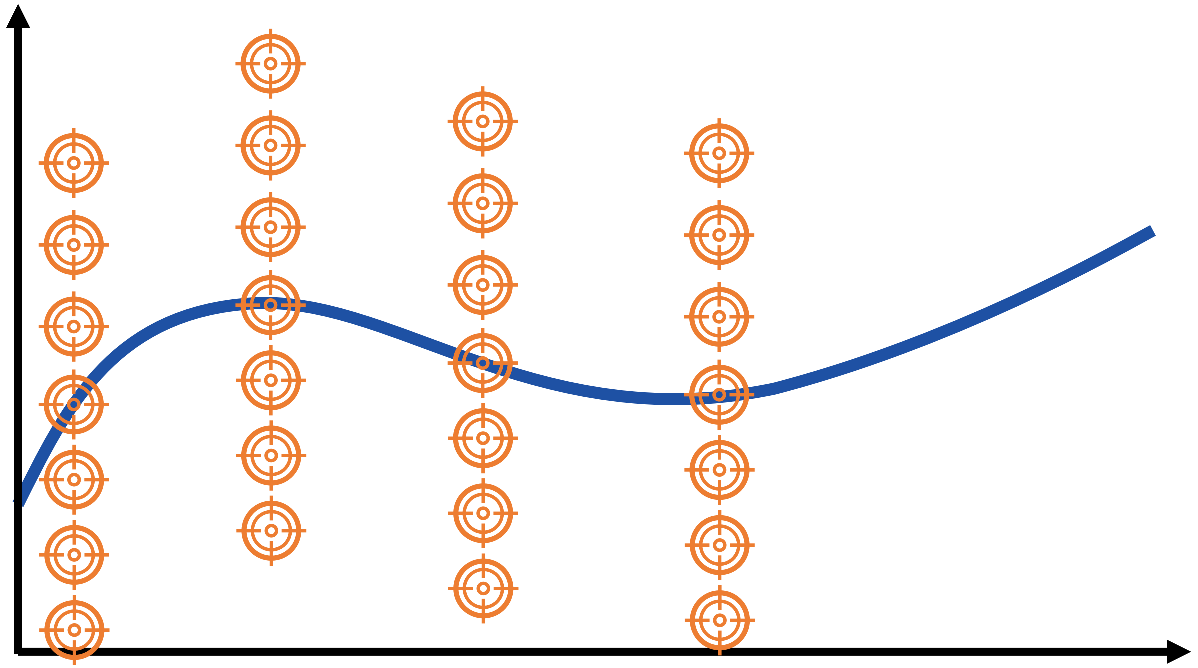


Делаем много неточных измерений – «выстрелов»

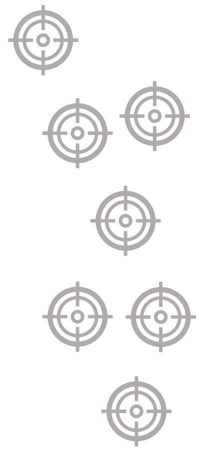


Распределения отклонений

Есть истинная функция

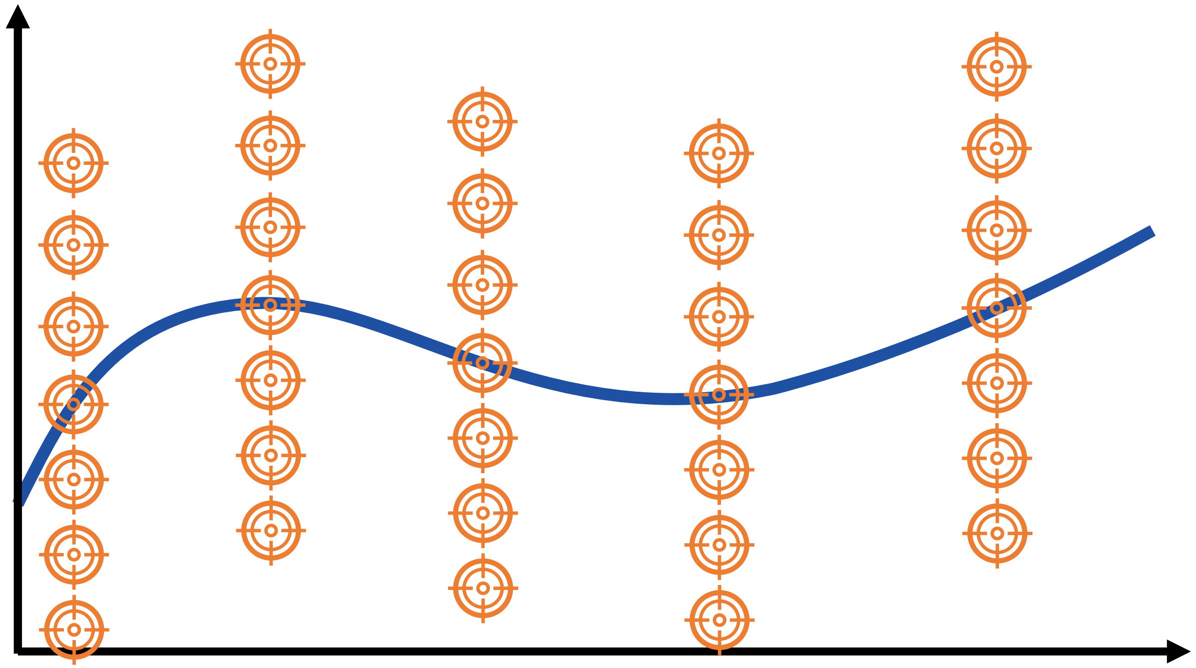


Делаем много неточных измерений – «выстрелов»



Распределения отклонений

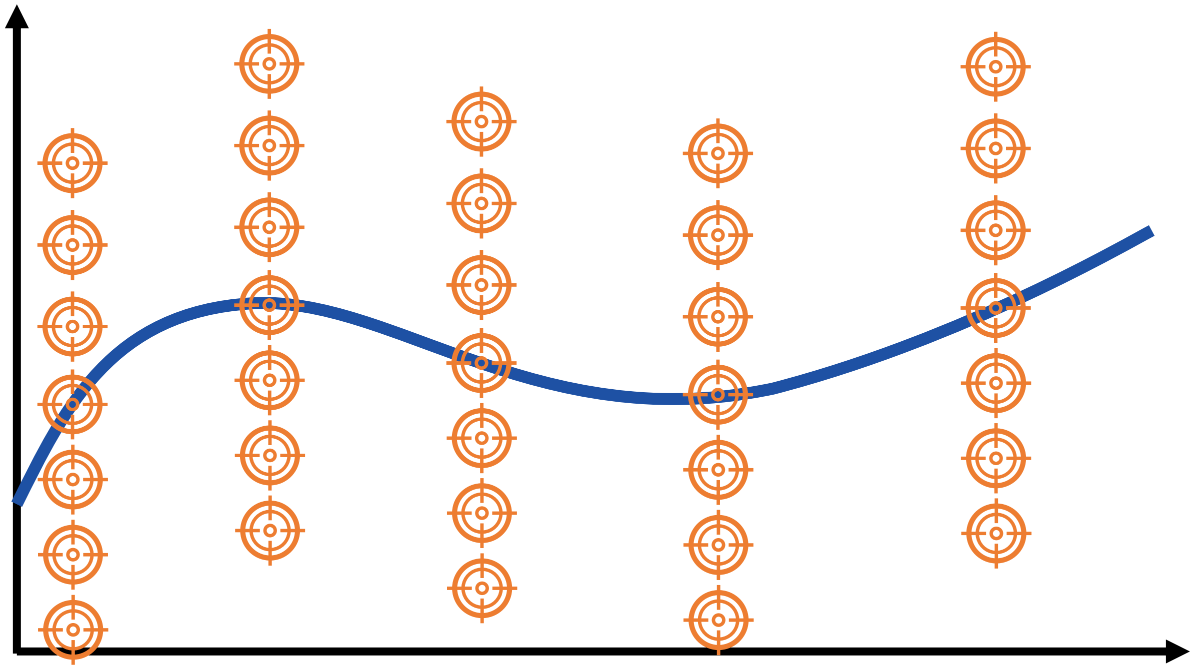
Есть истинная функция



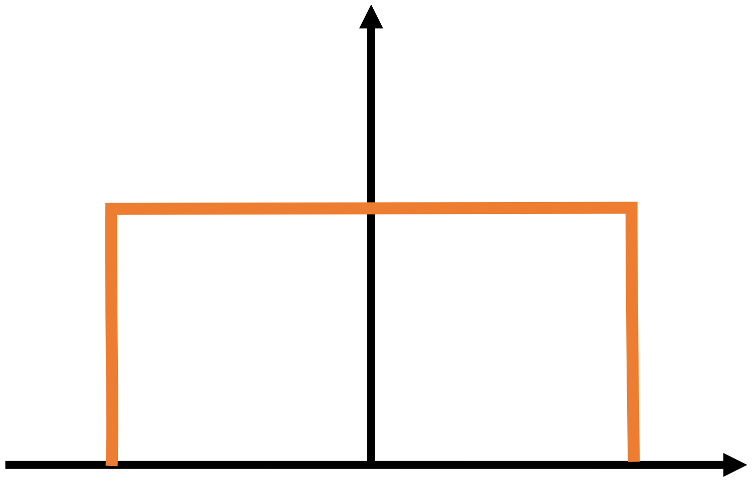
Делаем много неточных измерений – «выстрелов»

Распределения отклонений

Потенциальные результаты измерений

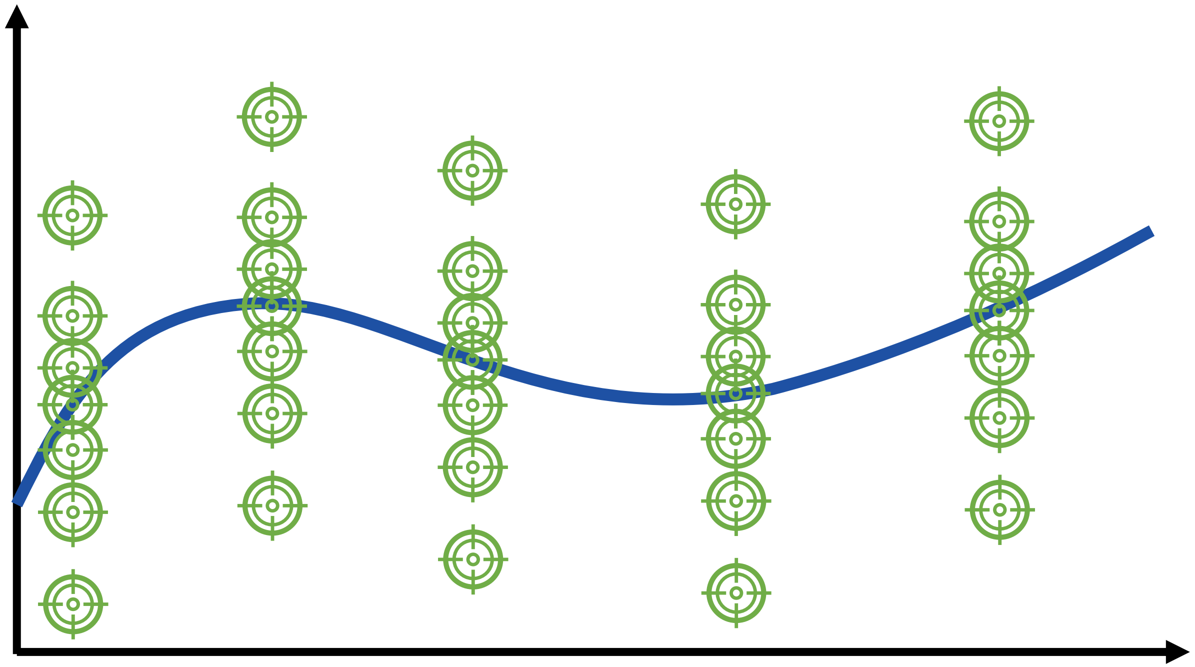


При равномерном распределении отклонений

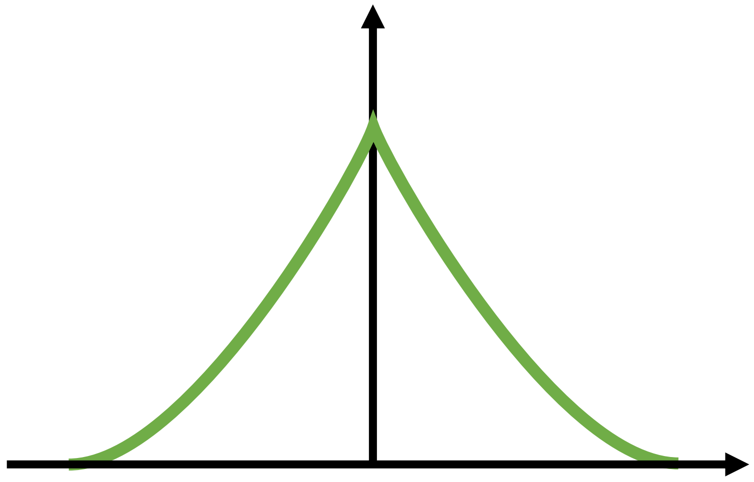


Распределения отклонений

Потенциальные результаты измерений

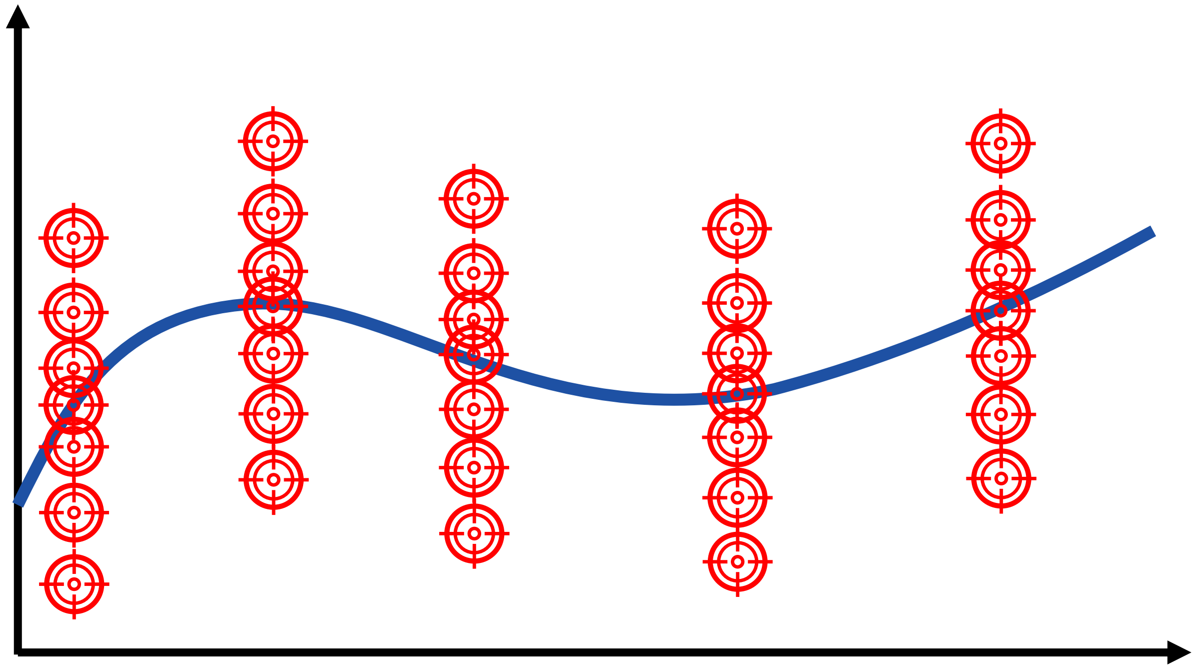


При распределении отклонений по Лапласу

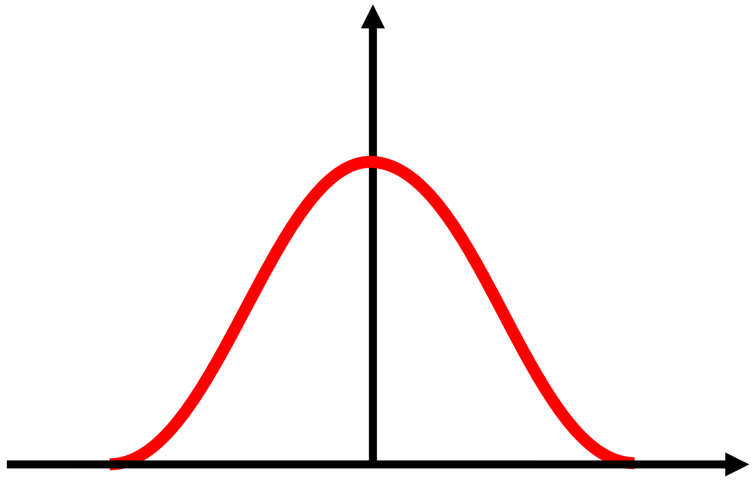


Распределения отклонений

Потенциальные результаты измерений



При нормальном распределении отклонений



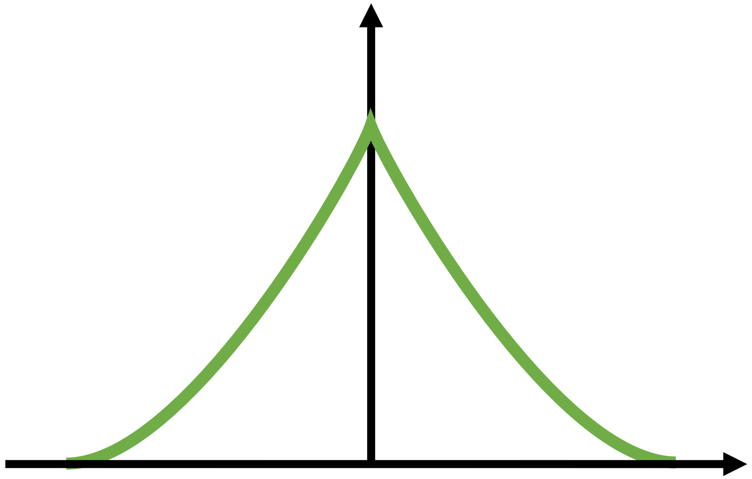
Выбор метода аппроксимации

Выбор метода аппроксимации

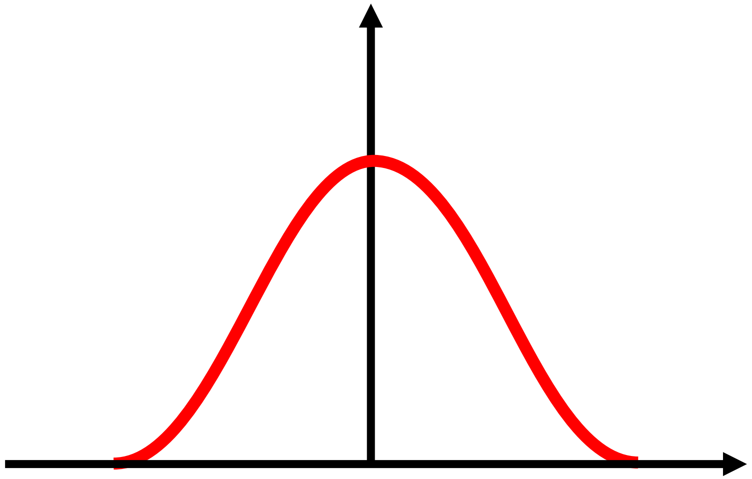
Как распределены отклонения?

Выбор метода аппроксимации

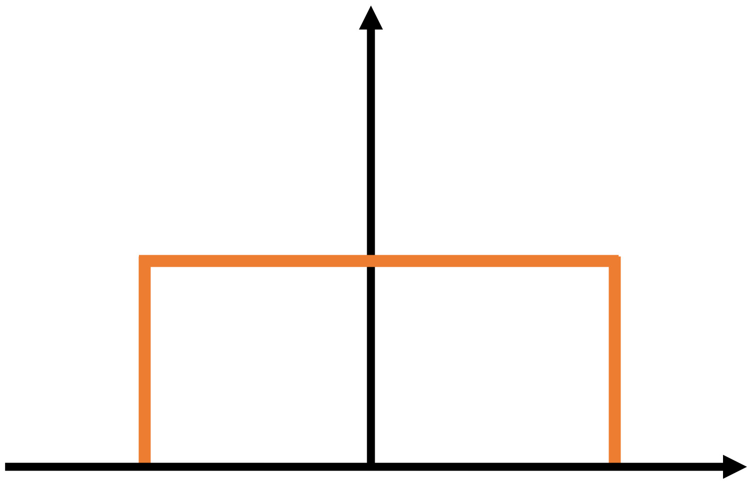
Как распределены отклонения?



Распределение Лапласа



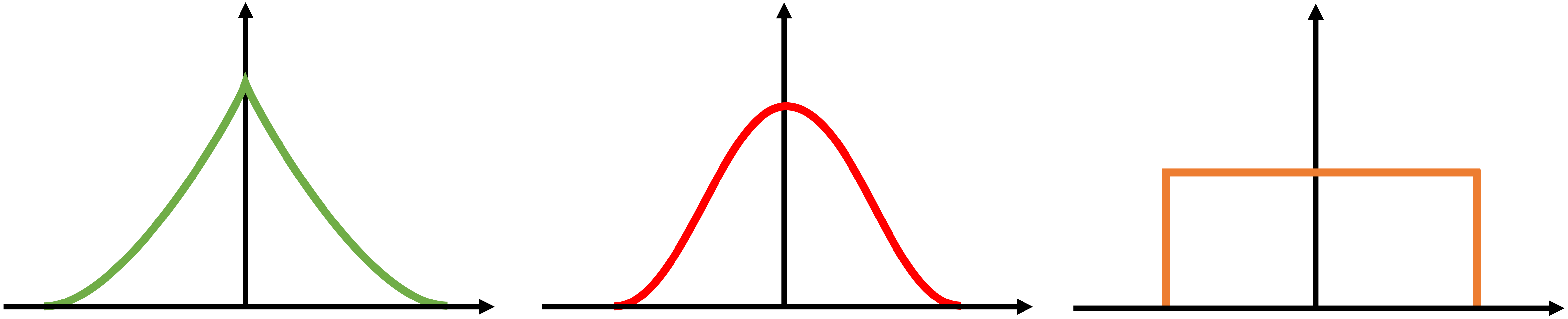
Нормальное распределение



Равномерное распределение

Выбор метода аппроксимации

Как распределены отклонения?



Распределение Лапласа

$\Downarrow$

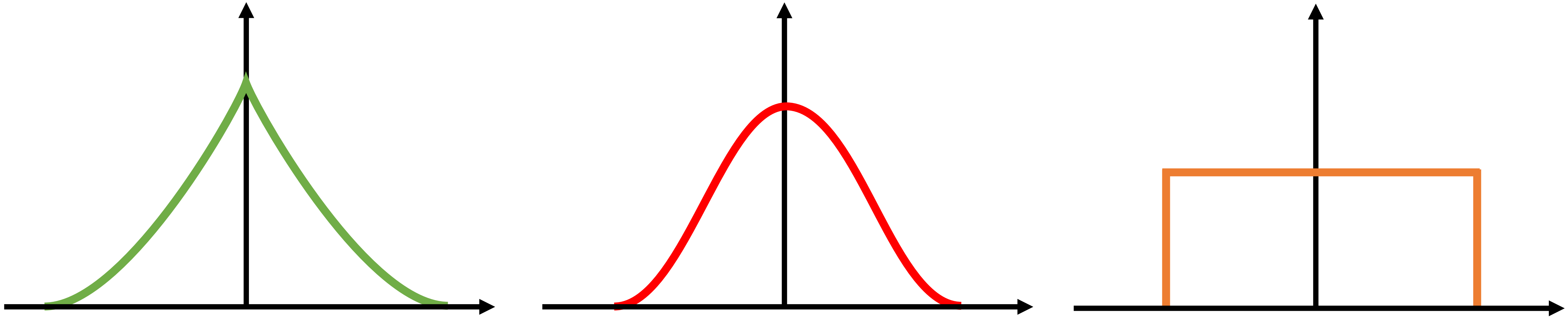
Минимизация l_1-нормы вектора отклонений

Нормальное распределение

Равномерное распределение

Выбор метода аппроксимации

Как распределены отклонения?



Распределение Лапласа

$\Downarrow$

Минимизация l_1-нормы вектора отклонений

Нормальное распределение

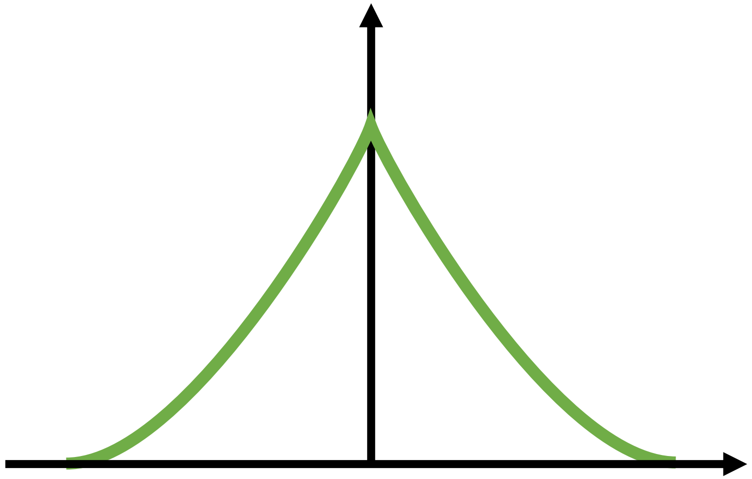
$\Downarrow$

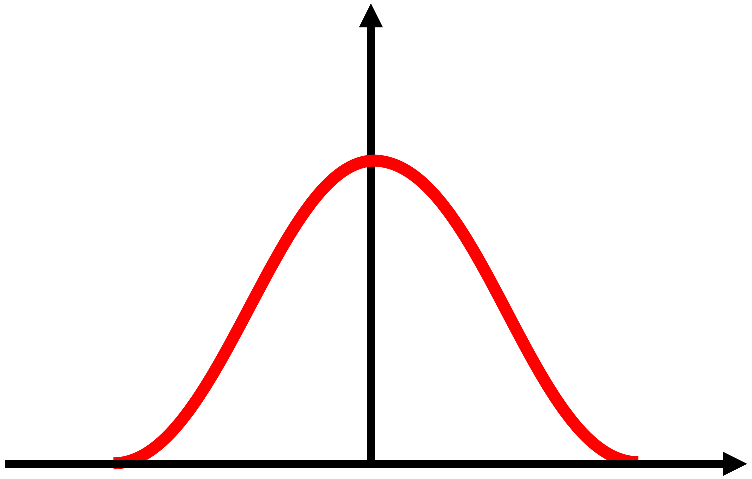
Минимизация l_2-нормы вектора отклонений

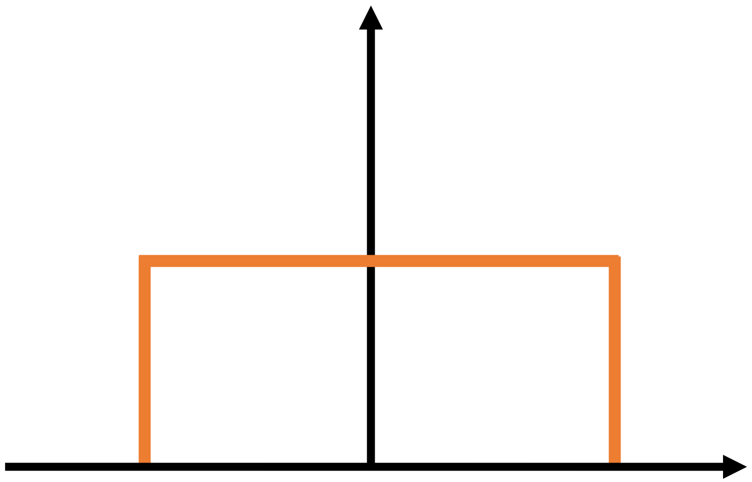
Равномерное распределение

Выбор метода аппроксимации

Как распределены отклонения?







Распределение Лапласа

$\Downarrow$

Минимизация l_1-нормы вектора отклонений

Нормальное распределение

$\Downarrow$

Минимизация l_2-нормы вектора отклонений

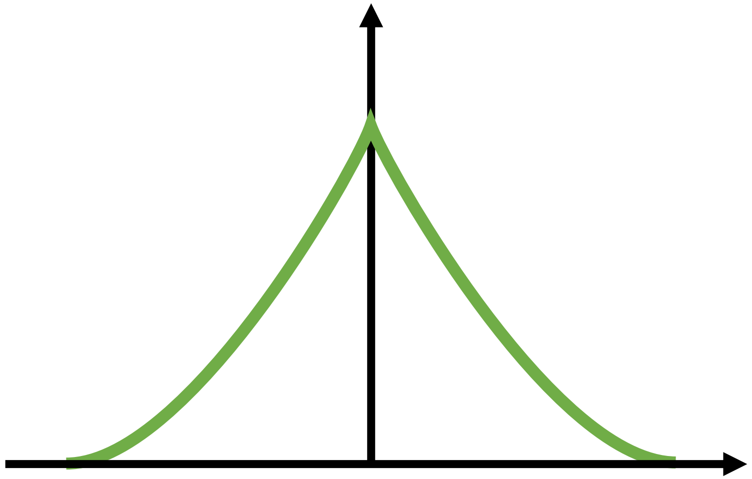
Равномерное распределение

$\Downarrow$

Минимизация l_∞-нормы вектора отклонений

Выбор метода аппроксимации

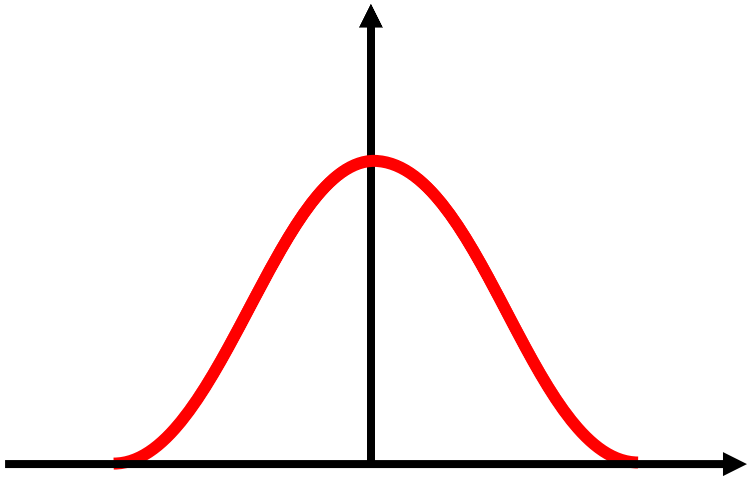
Как распределены отклонения?



Распределение Лапласа

$\Downarrow$

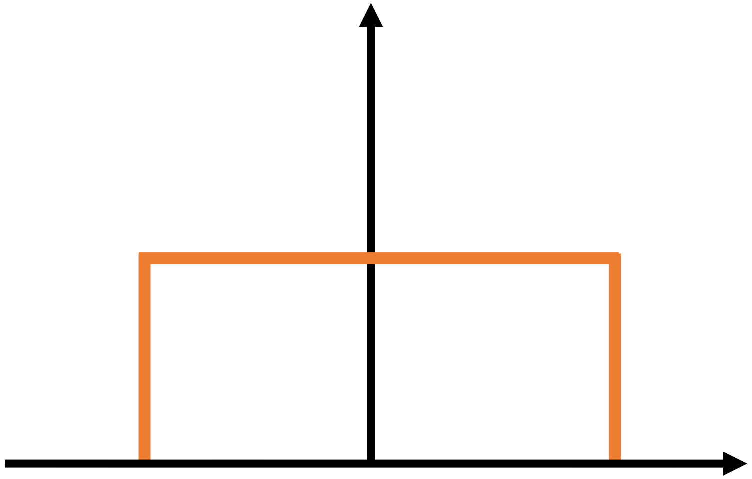
$$\sum |e_i| \to \min$$



Нормальное распределение

$\Downarrow$

Минимизация l_2-нормы вектора отклонений



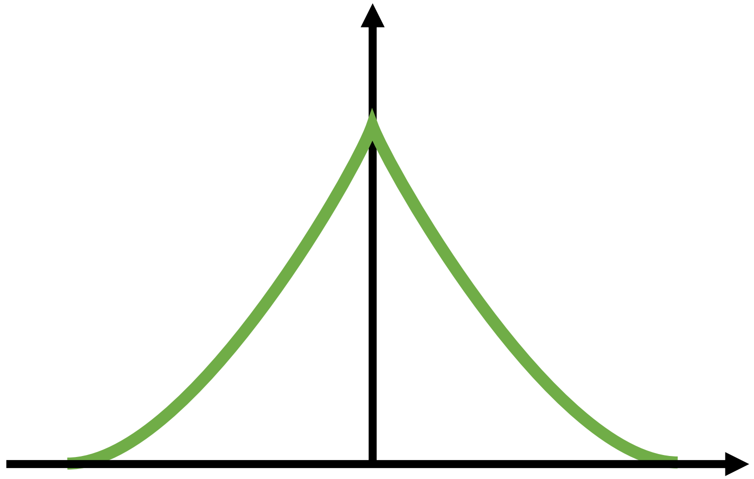
Равномерное распределение

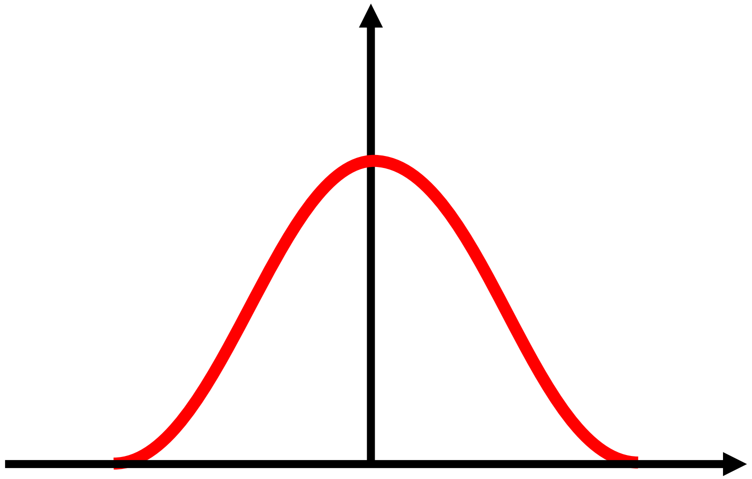
$\Downarrow$

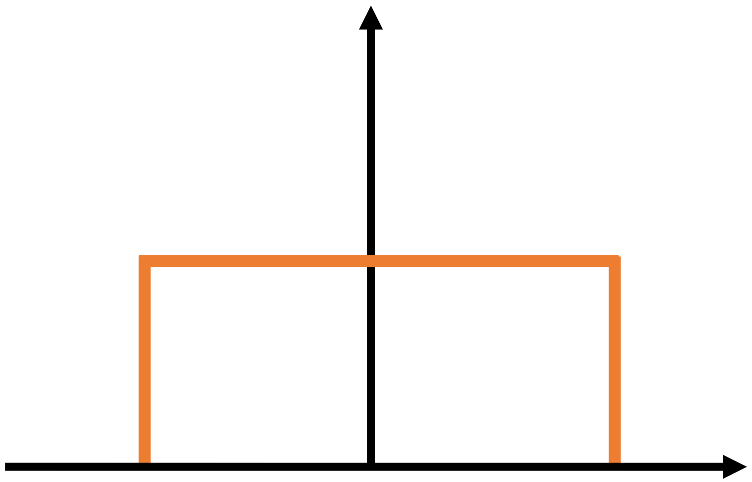
Минимизация l_∞-нормы вектора отклонений

Выбор метода аппроксимации

Как распределены отклонения?



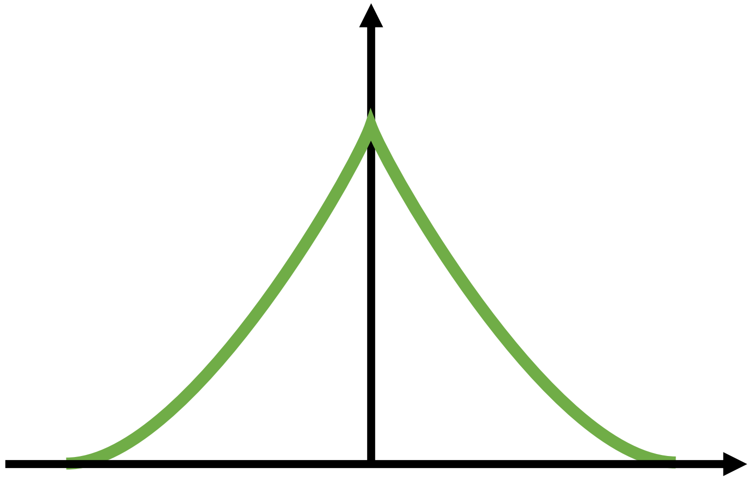




Распределение Лапласа	Нормальное распределение	Равномерное распределение
$\Downarrow$	$\Downarrow$	$\Downarrow$
$\sum \lvert e_i \rvert \to \min$	$\sum e_i^2 \to \min$	Минимизация l_∞-нормы вектора отклонений

Выбор метода аппроксимации

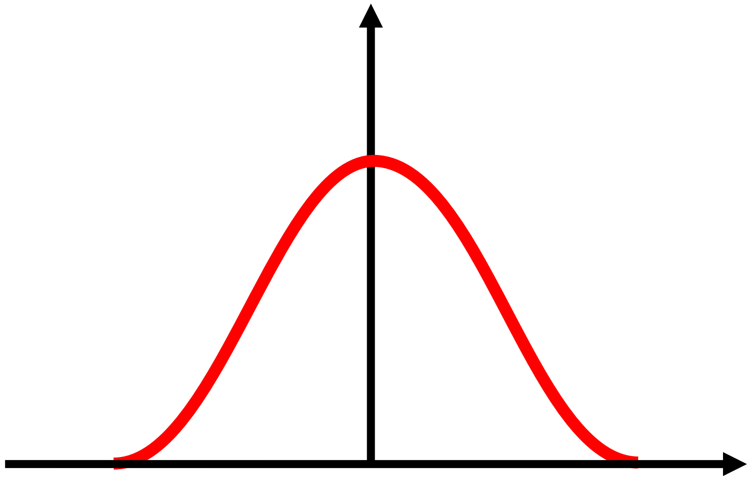
Как распределены отклонения?



Распределение Лапласа

$\Downarrow$

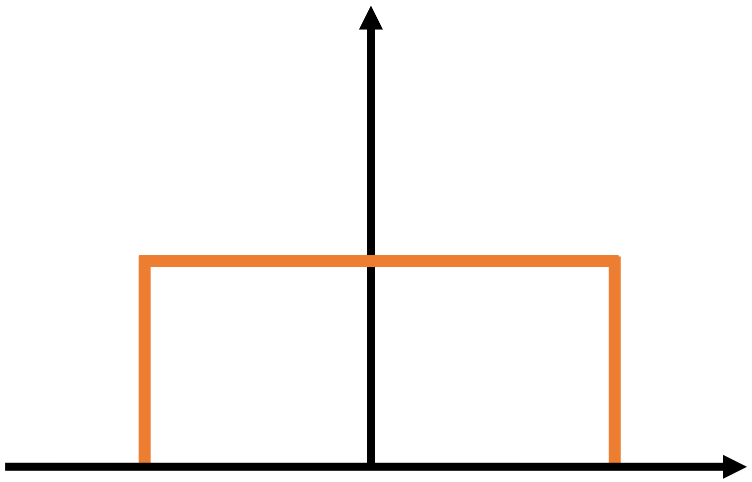
$$\sum |e_i| \to \min$$



Нормальное распределение

$\Downarrow$

$$\sum e_i^2 \to \min$$



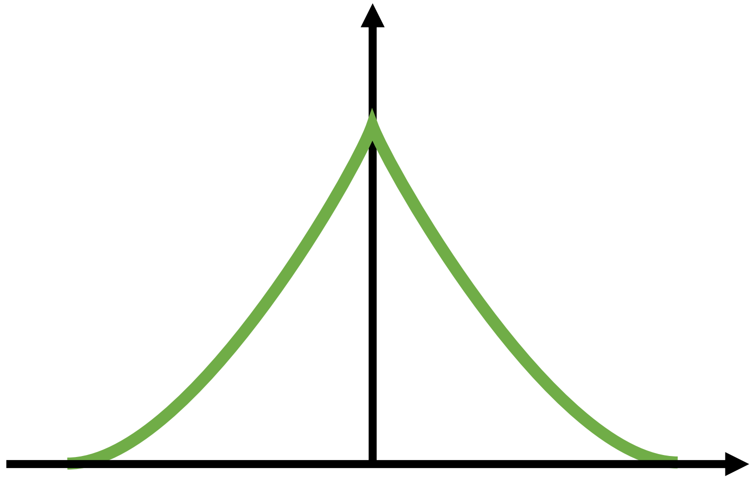
Равномерное распределение

$\Downarrow$

$$\max |e_i| \to \min$$

Выбор метода аппроксимации

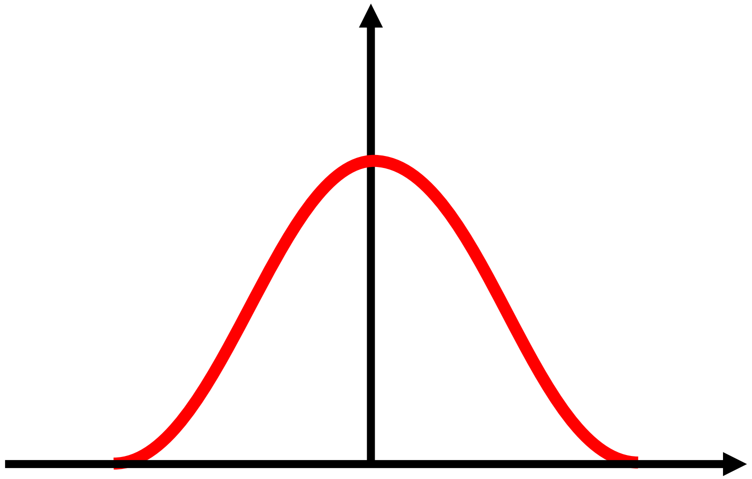
Как распределены отклонения?



Распределение Лапласа

$\Downarrow$

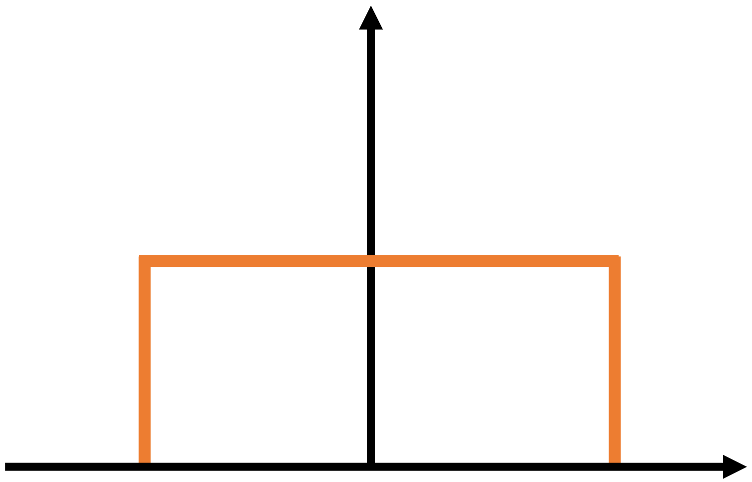
Метод наименьших модулей



Нормальное распределение

$\Downarrow$

$$\sum e_i^2 \to \min$$



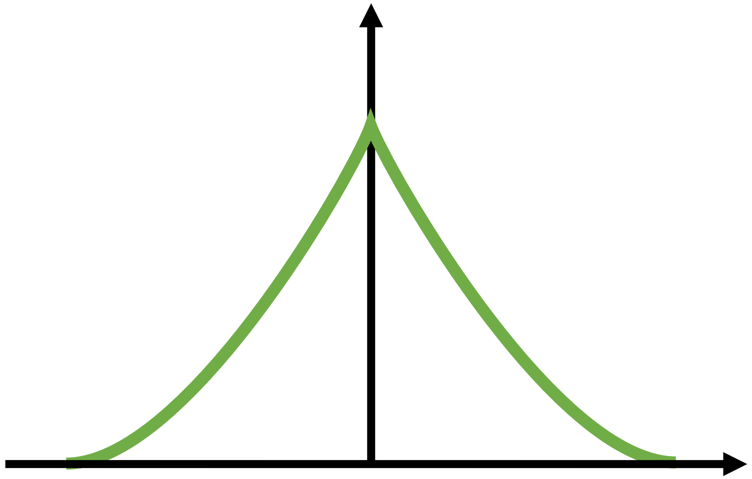
Равномерное распределение

$\Downarrow$

$$\max |e_i| \to \min$$

Выбор метода аппроксимации

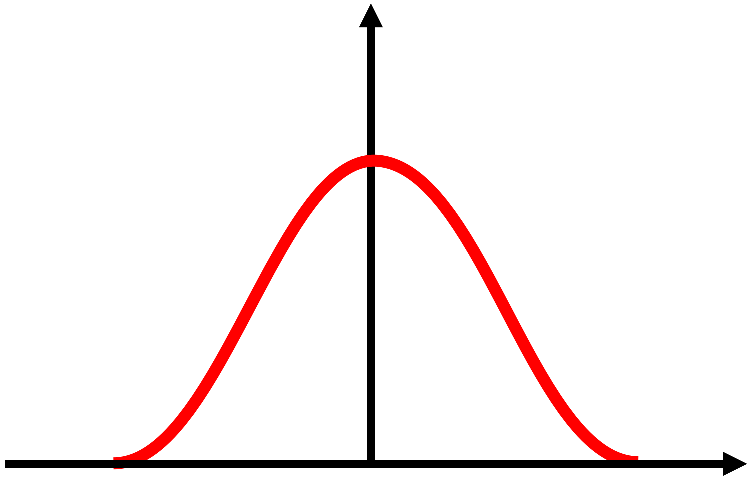
Как распределены отклонения?



Распределение Лапласа

$\Downarrow$

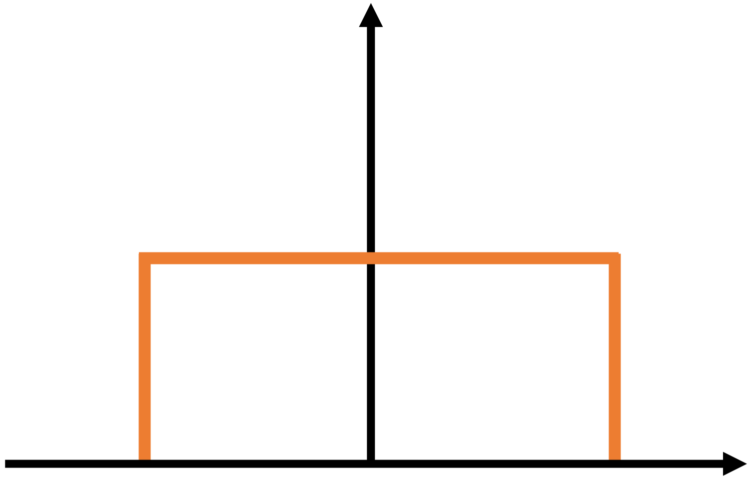
Метод наименьших модулей



Нормальное распределение

$\Downarrow$

Метод наименьших квадратов



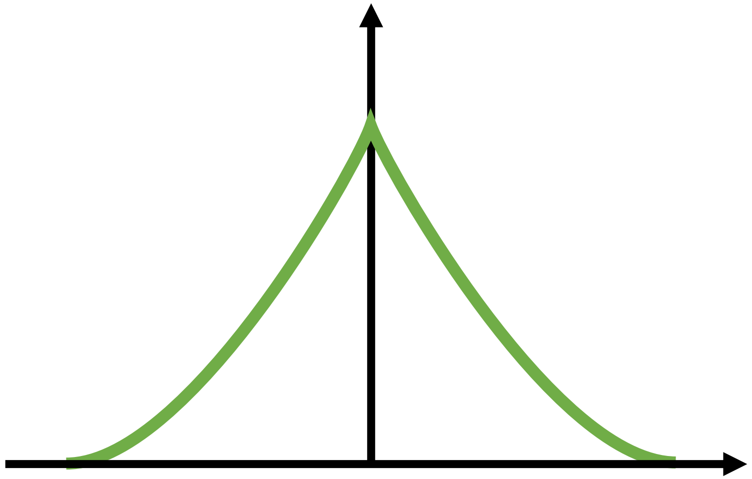
Равномерное распределение

$\Downarrow$

$\max |e_i| \to \min$

Выбор метода аппроксимации

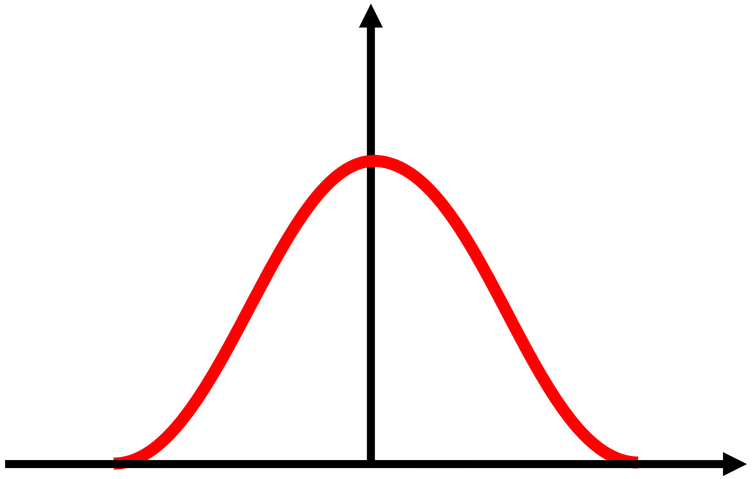
Как распределены отклонения?



Распределение Лапласа

⇓

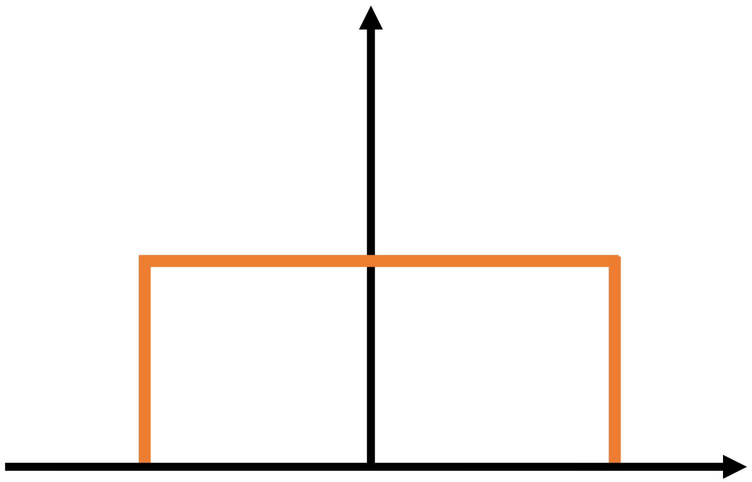
Метод наименьших модулей



Нормальное распределение

⇓

Метод наименьших квадратов



Равномерное распределение

⇓

Нет красивого названия

Выбор метода аппроксимации

Почему чаще всего используется метод наименьших квадратов?

Выбор метода аппроксимации

Почему чаще всего используется метод наименьших квадратов?

Чем он лучше других?

Центральная предельная теорема

Центральная предельная теорема

Если отклонения вызваны большим числом независимых причин,
то – какими бы ни были эти причины –
– итоговое распределение отклонений будет близким к нормальному

Центральная предельная теорема

Центральная предельная теорема

УНИВЕРСИТЕТ ИТМО

Если мы не можем проконтролировать причины отклонений

Центральная предельная теорема

Если мы не можем проконтролировать причины отклонений

$\Downarrow$

То логично предположить, что их много разных и независимых

Центральная предельная теорема

Если мы не можем проконтролировать причины отклонений

$\Downarrow$

То логично предположить, что их много разных и независимых

$\Downarrow$

Если это так, то распределение отклонений оказывается нормальным

Центральная предельная теорема

Если мы не можем проконтролировать причины отклонений

$\Downarrow$

То логично предположить, что их много разных и независимых

$\Downarrow$

Если это так, то распределение отклонений оказывается нормальным

$\Downarrow$

И метод наименьших квадратов даёт наилучшую аппроксимацию!

Центральная предельная теорема

НО ЭТО

Центральная предельная теорема

УНИВЕРСИТЕТ ИТМО

НО ЭТО

НЕ ВСЕГДА

Центральная предельная теорема

НО ЭТО

НЕ ВСЕГДА

ТАК

Спасибо!