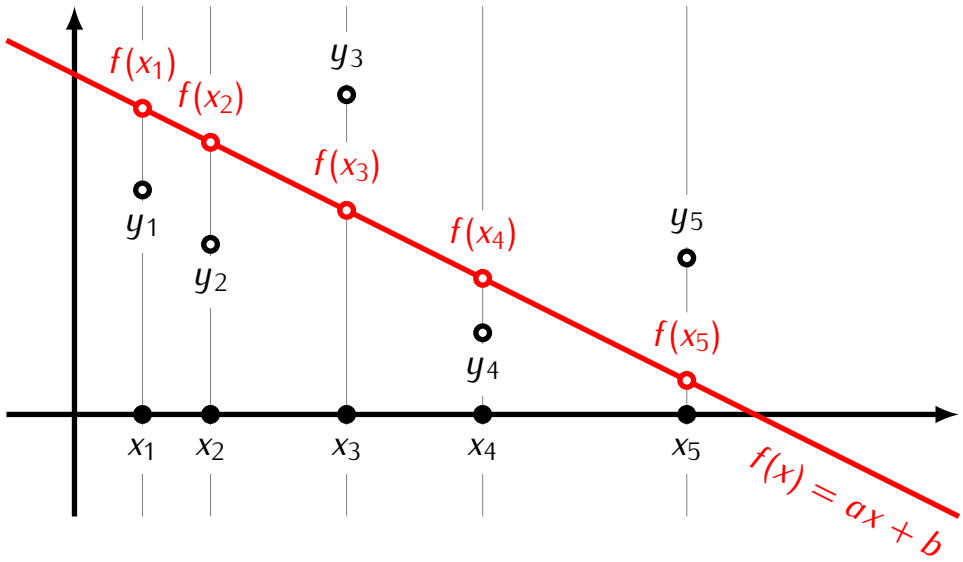
Application: Least square lines



Definition

If $(x_1, y_1), \ldots, (x_p, y_p)$ are points on the plane then the *least square line* for these points is the line given by an equation $f(x) = ax + b$ such that the number

$$\operatorname{dist}\left(\begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}, \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_p) \end{bmatrix}\right) = \sqrt{(y_1 - f(x_1))^2 + \ldots + (y_p - f(x_p))^2}$$

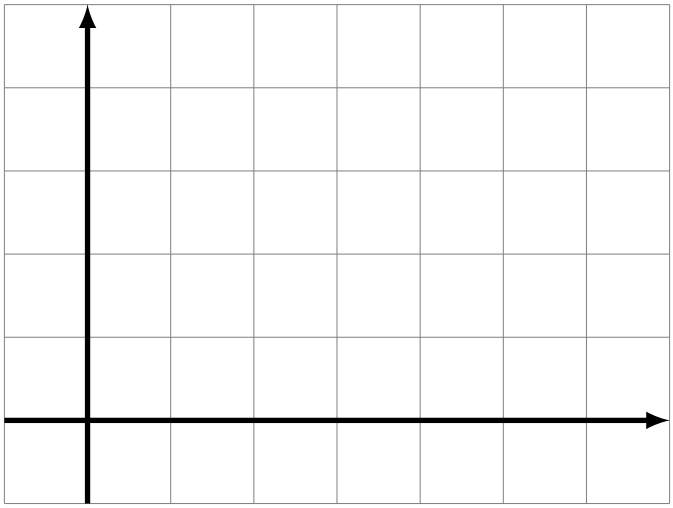
is the smallest possible.

Proposition

The line $f(x) = ax + b$ is the least square line for points $(x_1, y_1), \ldots, (x_p, y_p)$ if the vector $\begin{bmatrix} a \\ b \end{bmatrix}$ is the least square solution of the equation

$$\begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_p & 1 \end{bmatrix} \cdot \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}$$

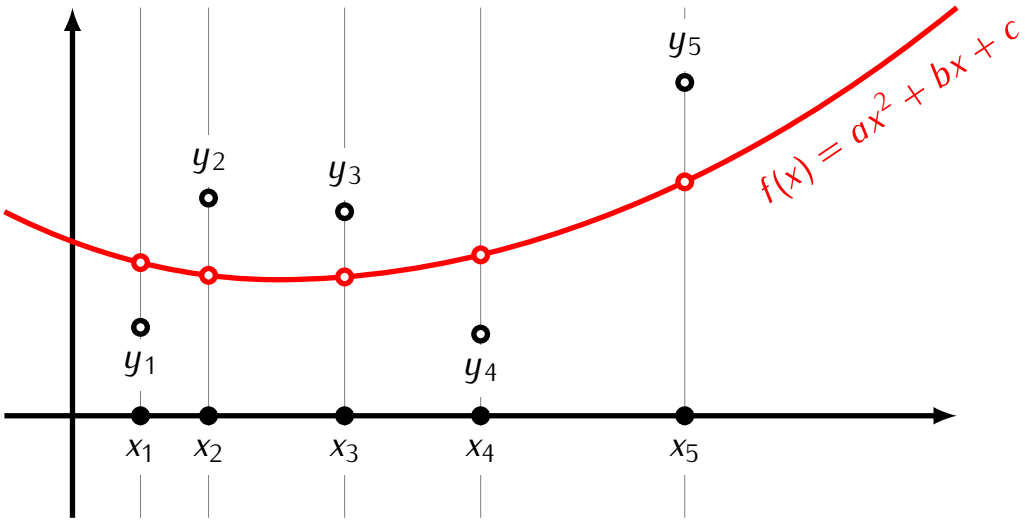
Example. Find the equation of the least square line for the points $(0, 0)$, $(1, 1)$, $(3, 1)$, $(5, 3)$.



Application: Least square curves

The above procedure can be used to determine curves other than lines that fit a set of points in the least square sense.

Example: Least square parabolas



Definition

If $(x_1, y_1), \ldots, (x_p, y_p)$ are points on the plane then the *least square parabola* for these points is the parabola given by an equation $f(x) = ax^2 + bx + c$ such that the number

$$\operatorname{dist}\left(\begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}, \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_p) \end{bmatrix}\right) = \sqrt{(y_1 - f(x_1))^2 + \ldots + (y_p - f(x_p))^2}$$

is the smallest possible.

Proposition

The parabola $f(x) = ax^2 + bx + c$ is the least square parabola for points $(x_1, y_1), \ldots, (x_p, y_p)$ if the vector $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ is the least square solution of the equation

$$\begin{bmatrix} x_1^2 & x_1 & 1 \\ \vdots & \vdots & \\ x_p^2 & x_p & 1 \end{bmatrix} \cdot \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}$$

Example. Find the equation of the least square parabola for the points $(-2, 2)$, $(0, 0)$, $(1, 1)$, $(2, 3)$.