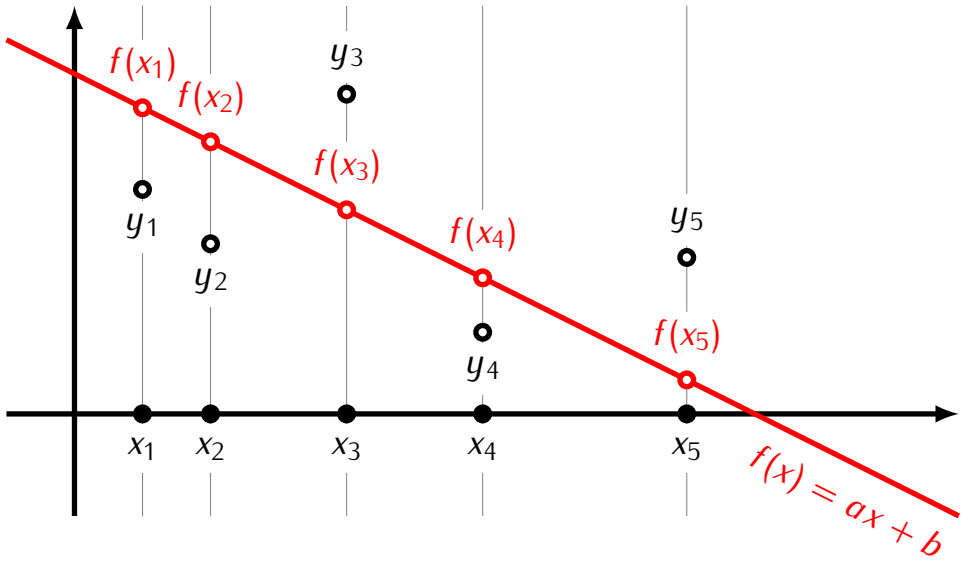
L_1 regression

Problem. Given points with coordinates (x_i, y_i) for $i = 1, \ldots, n$ find a function $f(x) = ax + b$ such that the sum

$$\sum_{i=1}^{n} |f(x_i) - y_i|$$

is a as small as possible.



Note. Compare with L_2 regression (least squares): we want to minimize

$$\sqrt{\sum_{i=1}^{n} |f(x_i) - y_i|^2}$$

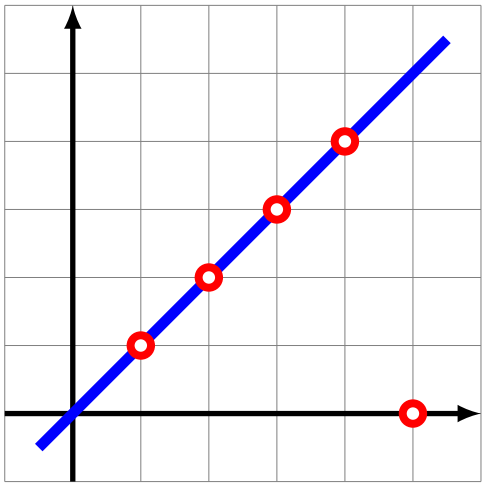
Problem. Given points with coordinates (x_i, y_i) for $i = 1, \ldots, n$ find a function $f(x) = ax + b$ such that the sum

$$\sum_{i=1}^{n} |f(x_i) - y_i|$$

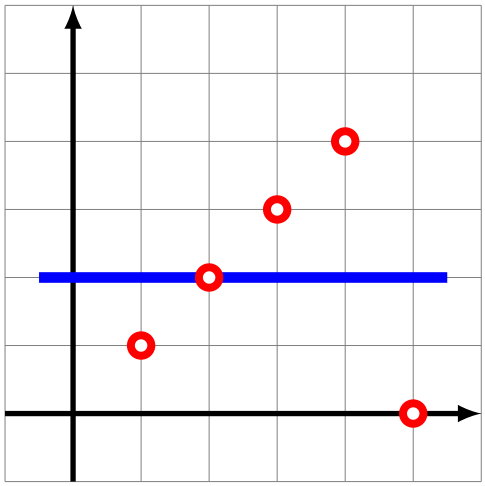
is a as small as possible.

L_1 regression vs L_2 regression

- L_1 regression is less sensitive that L_2 if we change the value of a single point.



L_1 regression



L_2 regression

- L_2 regression gives a uniquely defined line if there are at least two points with different x-coordinates. L_1 regression can have infinitely many solutions.