Recall:

1) The least square solutions of a matrix equation $A\mathbf{x} = \mathbf{b}$ are the solutions of the equation

$$A\mathbf{x} = \text{proj}_{\text{Col}(A)}\mathbf{b}$$

2) If $A\mathbf{x} = \mathbf{b}$ is a consistent equation, then $\mathbf{b} \in \text{Col}(A)$, and $\text{proj}_{\text{Col}(A)}\mathbf{b} = \mathbf{b}$. In such case the least square solutions of $A\mathbf{x} = \mathbf{b}$ are just the ordinary solutions.

3) If $A\mathbf{x} = \mathbf{b}$ is inconsistent, then the least square solutions are the best substitute for the (nonexistent) ordinary solutions.

4) If $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an orthogonal basis of a subspace V of $\mathbb{R}^n$ then

$$\text{proj}_V\mathbf{w} = \left(\frac{\mathbf{w} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 + \ldots + \left(\frac{\mathbf{w} \cdot \mathbf{v}_k}{\mathbf{v}_k \cdot \mathbf{v}_k}\right)\mathbf{v}_k$$

5) If $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an arbitrary basis of V then we can use the Gram-Schmidt process to obtain an orthogonal basis of V.

How to compute least square solutions of $A\mathbf{x} = \mathbf{b}$
(version 1.0)

1) Compute a basis of $\mathrm{Col}(A)$.

2) Use the Gram-Schmidt process to get an orthogonal basis of $\mathrm{Col}(A)$.

3) Use the orthogonal basis to compute $\mathrm{proj}_{\mathrm{Col}(A)}\mathbf{b}$.

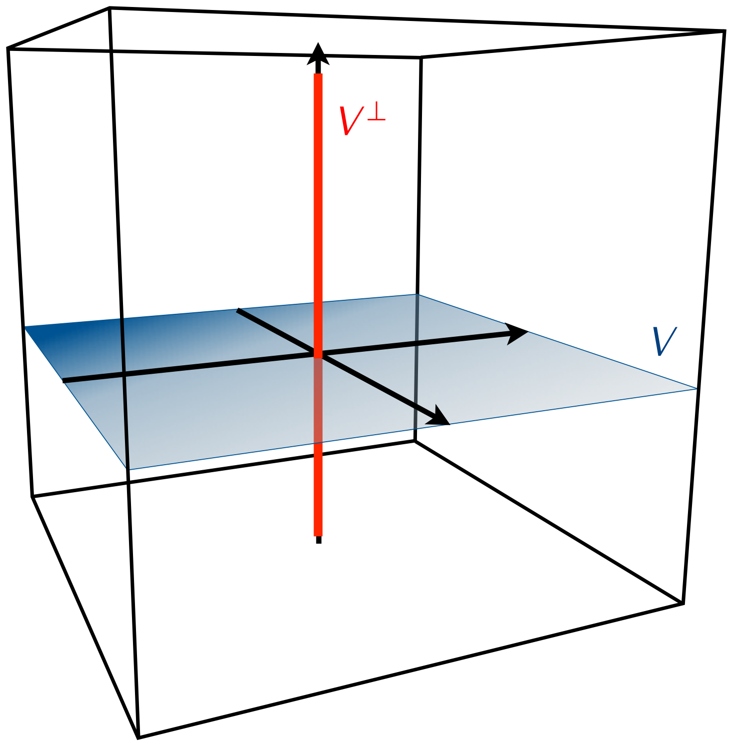
4) Compute solutions of the equation $A\mathbf{x} = \mathrm{proj}_{\mathrm{Col}(A)}\mathbf{b}$.

Next goal: Simplify this.

Definition

If V is a subspace of $\mathbb{R}^n$ then the *orthogonal complement* of V is the set $V^\perp$ of all vectors orthogonal to V:

$$V^\perp = \{\mathbf{w} \in \mathbb{R}^n \mid \mathbf{w} \cdot \mathbf{v} = 0\}$$



Propositon

If V is a subspace of $\mathbb{R}^n$ then:

1) $V^\perp$ is also a subspace of $\mathbb{R}^n$.
2) For each vector $\mathbf{w} \in \mathbb{R}^n$ there exist unique vectors $\mathbf{v} \in V$ and $\mathbf{z} \in V^\perp$ such that $\mathbf{w} = \mathbf{v} + \mathbf{z}$.

Definition

If A is an $m \times n$ matrix then the *row space* of A is the subspace $\mathrm{Row}(A)$ of $\mathbb{R}^n$ spanned by rows of A.

Example

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

Proposition

If A is a matrix then

$$\mathrm{Row}(A)^{\perp} = \mathrm{Nul}(A)$$

Corollary

If A is a matrix then

$$\mathrm{Col}(A)^{\perp} = \mathrm{Nul}(A^T)$$

Back to least square solutions

Theorem

A vector $\hat{\mathbf{x}}$ is a least square solution of a matrix equation

$$A\mathbf{x} = \mathbf{b}$$

if and only if $\hat{\mathbf{x}}$ is an ordinary solution of the equation

$$(A^T A)\mathbf{x} = A^T\mathbf{b}$$

Theorem

The equation

$$(A^T A)\mathbf{x} = A^T\mathbf{b}$$

is called the *normal equation* of $A\mathbf{x} = \mathbf{b}$.

How to compute least square solutions of $A\mathbf{x} = \mathbf{b}$

(version 2.0)

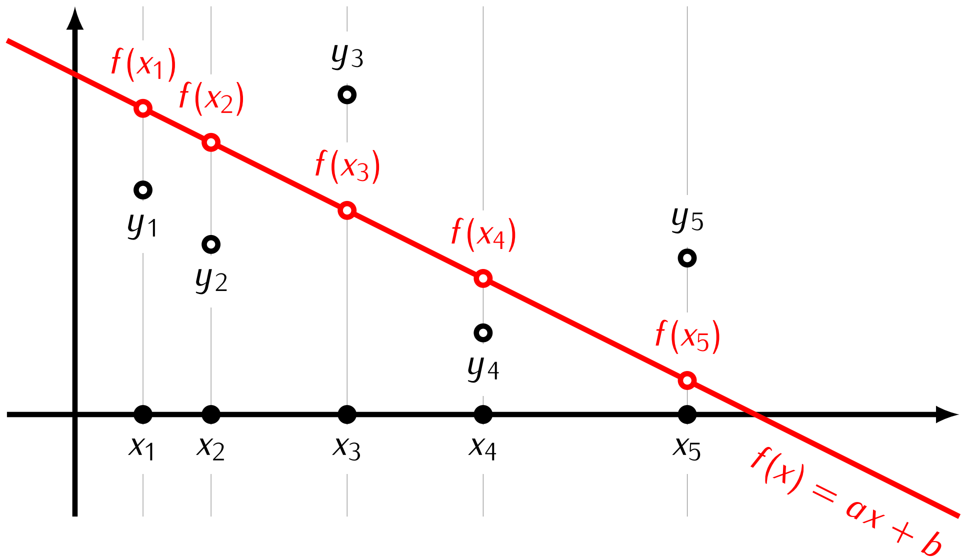
1) Compute $A^T A$, $A^T\mathbf{b}$.

2) Solve the normal equation $(A^T A)\mathbf{x} = A^T\mathbf{b}$.

Example. Compute least square solutions of the following equation:

$$\begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

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

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

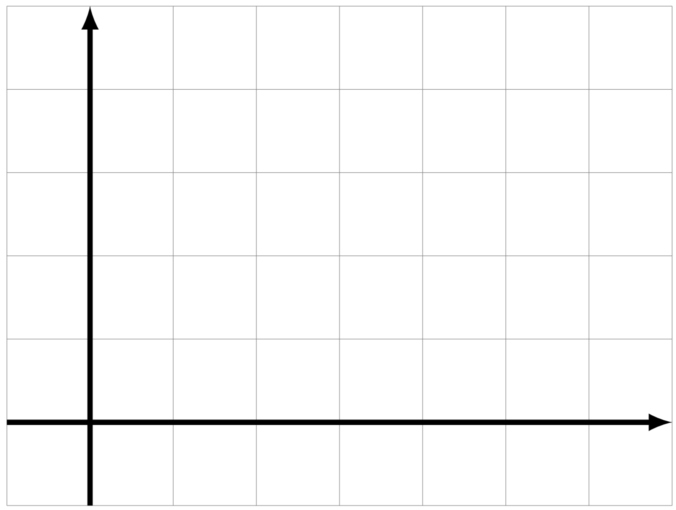
is the smallest possible.

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

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

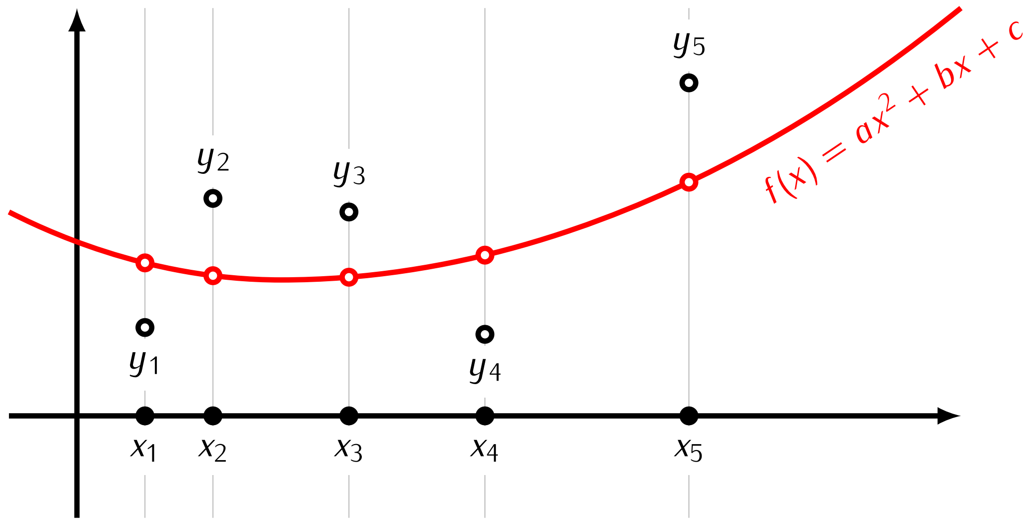
Example. Find the equation of the least square line for the points $(0, 0)$, $(1, 1)$, $(3, 1)$, $(5, 3)$.



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 $(-2, 2)$, $(0, 0)$, $(1, 1)$, $(2, 3)$.