

Definition

Given a vector

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \in \mathbb{R}^N$$

the *mean* of X is the number

$$m_X = \frac{x_1 + \dots + x_N}{N}$$

Notation. For a vector $X \in \mathbb{R}^N$ as above, by \tilde{X} we will denote the vector

$$\tilde{X} = \begin{bmatrix} x_1 - m_X \\ \vdots \\ x_N - m_X \end{bmatrix}$$

We will say that \tilde{X} is the *demeaning* of X .

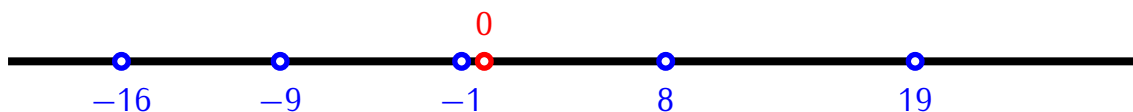
Note. $m_{\tilde{X}} = 0$.

Example.

$$X = [62 \ 90 \ 79 \ 70 \ 55]^T, m_X = 71$$



$$\tilde{X} = [-9 \ 19 \ 8 \ -1 \ -16]^T, m_{\tilde{X}} = 0$$



Definition

Given a vector

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \in \mathbb{R}^N$$

the *variance* of X is the number

$$\text{Var}(X) = \frac{1}{N} \sum_{i=1}^N (x_i - m_X)^2$$

Note. $\text{Var}(X) = \frac{1}{N} \tilde{X}^T \tilde{X}$.

Proposition

For a vector

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \in \mathbb{R}^N$$

we have

$$\text{Var}(X) = \frac{1}{2N^2} \sum_{i,j} (x_i - x_j)^2$$

Definition

Given vectors

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \in \mathbb{R}^N$$

the *covariance* of X and Y is the number

$$\text{Cov}(X, Y) = \frac{1}{N} \sum_{i=1}^N (x_i - m_X)(y_i - m_Y)$$

Note.

- 1) $\text{Cov}(X, Y) = \frac{1}{N} \tilde{X}^T \tilde{Y}$
- 2) $\text{Var}(X) = \text{Cov}(X, X)$

Proposition

For vectors

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \in \mathbb{R}^N$$

we have

$$\text{Cov}(X, Y) = \frac{1}{2N^2} \sum_{i,j} (x_i - x_j)(y_i - y_j)$$

Example.

