Example.

$$P = \begin{array}{r c} & \begin{array}{ccc} \text{height} & \text{weight} & \text{age} \end{array} \\ \begin{array}{r} \text{Aly} \\ \text{Bob} \\ \text{Chen} \\ \text{Deb} \end{array} & \begin{bmatrix} 62 & 141 & 19 \\ 82 & 164 & 21 \\ 79 & 154 & 19 \\ 70 & 135 & 25 \end{bmatrix} \end{array}$$

<u>**General form**</u>

$$A = \begin{bmatrix} \boxed{X_1} & \boxed{X_2} & \cdots & \boxed{X_M} \end{bmatrix} = \begin{bmatrix} \boxed{S_1} \\ \boxed{S_2} \\ \vdots \\ \boxed{S_N} \end{bmatrix}$$

Notation. Given a data matrix $A = \begin{bmatrix} X_1 & X_2 & \dots & X_M \end{bmatrix}$ we will denote

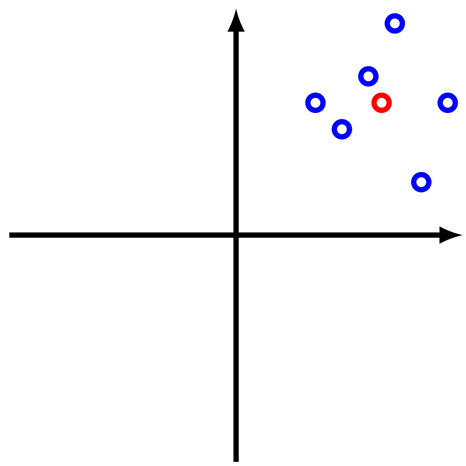
$$\widetilde{A} = \begin{bmatrix} \widetilde{X}_1 & \widetilde{X}_2 & \dots & \widetilde{X}_M \end{bmatrix}$$

where $\widetilde{X}_i$ is the demeaning of the vector X_i.

Example.

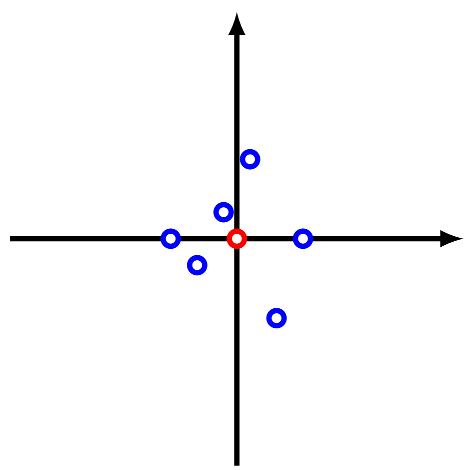
$$A = \begin{bmatrix} 5.0 & 6.0 \\ 7.0 & 2.0 \\ 4.0 & 4.0 \\ 3.0 & 5.0 \\ 6.0 & 8.0 \\ 8.0 & 5.0 \end{bmatrix}$$

$$\text{mean} = \begin{bmatrix} 5.5 & 5.0 \end{bmatrix}$$



$$\widetilde{A} = \begin{bmatrix} -0.5 & 1.0 \\ 1.5 & -3.0 \\ -1.5 & -1.0 \\ -2.5 & 0.0 \\ 0.5 & 3.0 \\ 2.5 & 0.0 \end{bmatrix}$$

$$\text{mean} = \begin{bmatrix} 0.0 & 0.0 \end{bmatrix}$$



Definition

The *covariance matrix* of a data matrix A is the matrix

$$C_A = \frac{1}{N}\,\widetilde{A}^T\widetilde{A}$$

Proposition

If $A = \begin{bmatrix} X_1 & \dots & X_M \end{bmatrix}$ is a data matrix then

$$C_A = \begin{bmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, X_2) & \dots & \mathrm{Cov}(X_1, X_M) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Var}(X_2) & \dots & \mathrm{Cov}(X_2, X_M) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_M, X_1) & \mathrm{Cov}(X_M, X_2) & \dots & \mathrm{Var}(X_M) \end{bmatrix}$$

Note. For any matrix A the matrix C_A is

- symmetric
- positive semidefinite

Total variance and trace

Definition

If $A = \begin{bmatrix} X_1 & \dots & X_M \end{bmatrix}$ is a data matrix then the *total variance* of A is the number

$$\mathrm{Var}(A) = \mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_M)$$

Definition

For a square matrix

$$B = \begin{bmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{bmatrix}$$

the *trace* of B is the number

$$\mathrm{tr}\, B = b_{11} + b_{22} + \ldots + b_{nn}$$

Note. If A is a data matrix and C_A is its covariance matrix then

$$\mathrm{Var}(A) = \mathrm{tr}\, C_A$$

Proposition

If A, B are $n \times n$ matrices then

1) If A, B are $n \times n$ matrices then $\operatorname{tr} AB = \operatorname{tr} BA$.
2) If A, P, B are $n \times n$ matrices such that $A = PBP^{-1}$ then $\operatorname{tr} A = \operatorname{tr} B$.

Corollary

If a matrix A is diagonalizable,

$$A = PDP^{-1}$$

for some invertible matrix P and a diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

then $\operatorname{tr} A = \operatorname{tr} D = \lambda_1 + \lambda_2 + \ldots + \lambda_n$.