VARIANCE-COVARIANCE MATRIX

1. Introduction

- Variance-covariance matrix consists of the:
  - variances of the variables along the main diagonal
  - covariances between each pair of variables in the other matrix positions.
- It is also called just covariance matrix or dispersion or dispersion matrix.
- For two variables observed N times, variance-covariance matrix is calculated like this:

$$X = \begin{bmatrix} x_1 & \cdots & x_N \\ y_1 & \cdots & y_N \end{bmatrix} \quad , \quad \overline{X} = \begin{bmatrix} \overline{x} \\ \overline{y} \end{bmatrix} \quad , \quad X_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix}$$

$$S = \frac{1}{N} \sum_{i=1}^{N} \left(X_i - \overline{X} \right) \left(X_i - \overline{X} \right)^T \tag{1.1}$$

where: X – observation matrix containing measured values for all variables,
$\overline{X}$ – mean vector,
X_i – vector containing values of all variables at moment i,
N – number of times variables were measured,
S – variance-covariance matrix.

- For observation i=1 we get first summand of equation (1.1) like this:

$$\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} - \begin{bmatrix} \overline{x} \\ \overline{y} \end{bmatrix} \right) \left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} - \begin{bmatrix} \overline{x} \\ \overline{y} \end{bmatrix} \right)^T =$$

$$= \begin{bmatrix} x_1 - \overline{x} \\ y_1 - \overline{y} \end{bmatrix} \begin{bmatrix} x_1 - \overline{x} \\ y_1 - \overline{y} \end{bmatrix}^T =$$

$$= \begin{bmatrix} x_1 - \overline{x} \\ y_1 - \overline{y} \end{bmatrix} \begin{bmatrix} x_1 - \overline{x} & y_1 - \overline{y} \end{bmatrix} =$$

$$= \begin{bmatrix} (x_1 - \overline{x})^2 & (x_1 - \overline{x})(y_1 - \overline{y}) \\ (y_1 - \overline{y})(x_1 - \overline{x}) & (y_1 - \overline{y})^2 \end{bmatrix}$$

- Repeating this procedure for all other observations and summing resulting matrices, as defined in (1.1), we get S.

2. References

[1] http://www.itl.nist.gov/div898/handbook/pmc/section5/pmc541.htm