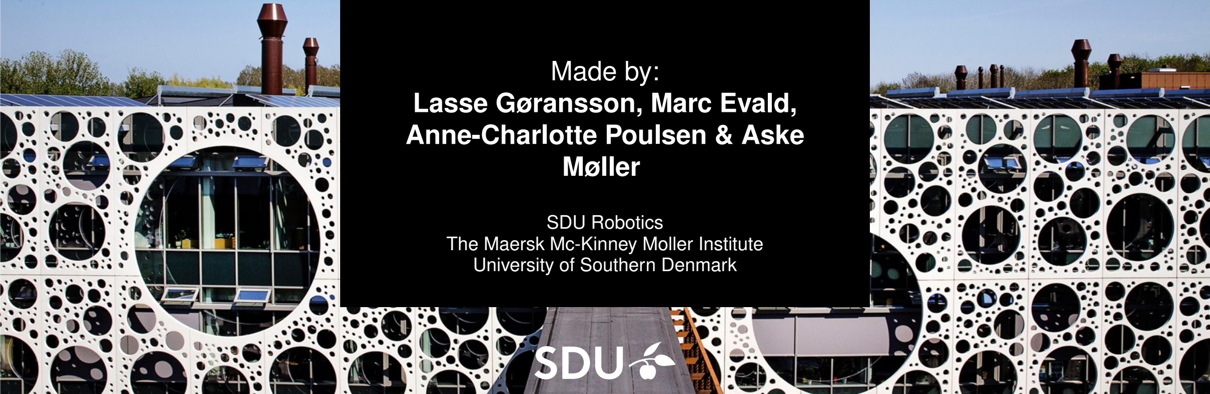
Multivariate data and multivariate normal distribution

Multivariate Statistic

Made by:

Lasse Göransson, Marc Evald, Anne-Charlotte Poulsen & Aske Møller

SDU Robotics
The Maersk Mc-Kinney Moller Institute
University of Southern Denmark



Agenda

- Univariate $\rightarrow$ Multivariate
- Multivariate theory
- Multivariate sampling
- MVN
- Model Check

Univariate Statistics

Univariate

- Stochastic Variable
- Mean
- Variance

- Stochastic Variable

$$\underset{p \times 1}{X}$$

- Mean

$$E[X] = \underset{p \times 1}{\mu} = \begin{bmatrix} E[X_1] \\ \vdots \\ E[X_p] \end{bmatrix}$$

- Variance

$$V[X] = \underset{p \times p}{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_p^2 \end{bmatrix}$$

- Mean

$$E\left[X\right] = \underset{p\times 1}{\mu} = \begin{bmatrix} E\left[X_1\right] \\ \vdots \\ E\left[X_p\right] \end{bmatrix}$$

- Variance

$$V\left[X\right] = \underset{p\times p}{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_p^2 \end{bmatrix}$$

- Correlation

$$\rho = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ & \ddots & & \rho_{2p} \\ & & & \\ & & & 1 \end{bmatrix}$$

Sampling

- Data

$$x = \begin{bmatrix} \begin{bmatrix} x_{11} & \cdots & x_{1p} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} x_{n1} & \cdots & x_{np} \end{bmatrix} \end{bmatrix}$$

- Estimated Mean

$$\hat{\mu} = \bar{x} = \begin{bmatrix} \frac{1}{n} \sum_{j=1}^{n} x_{j1} \\ \vdots \\ \frac{1}{n} \sum_{j=1}^{n} x_{jp} \end{bmatrix}$$

- Estimated Variance

$$\hat{\Sigma} = S = \frac{1}{n-1} \sum_{j=1}^{n} \left(x_j - \bar{x}\right) \left(x_j - \bar{x}\right)^T$$

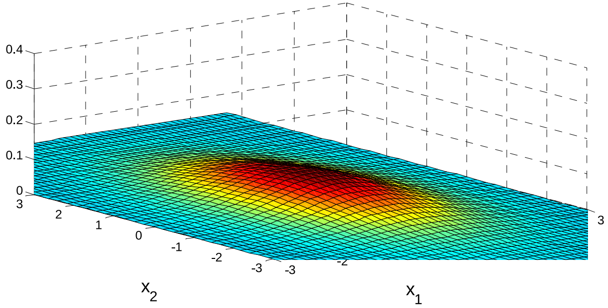
- Estimated Correlation

$$\hat{\rho} = R \rightarrow r_{ij} = \frac{s_{ij}}{\sqrt{s_{ii} s_{jj}}}$$

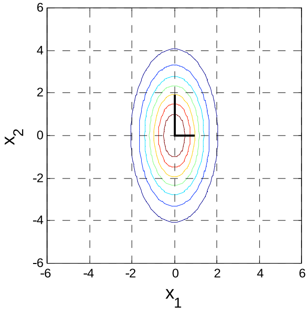
MVN

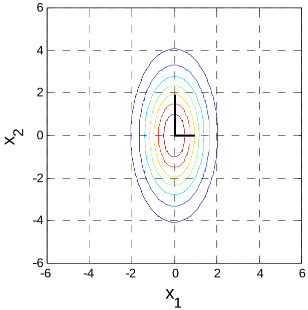
Multivariate Normal Distribution

$$\underset{p \times 1}{X} \sim N_p(\mu, \Sigma)$$



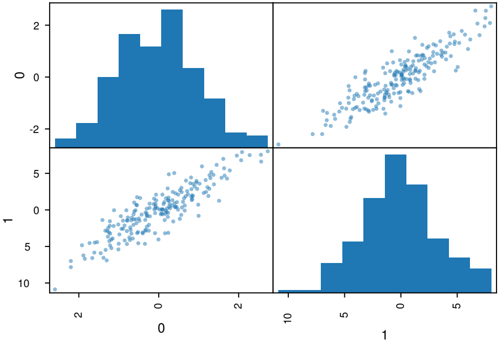
MVN
Contours



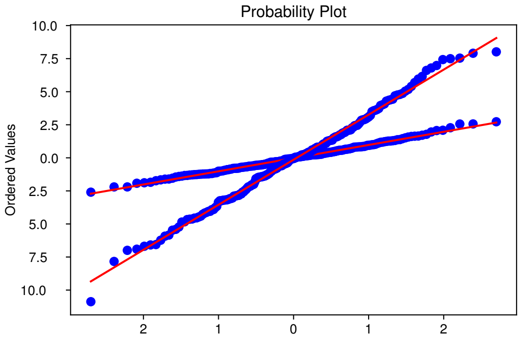


- Mahalanobis Distance $(x-\mu)^T\Sigma^{-1}(x-\mu) \sim \chi^2_p$

Model Check



0
1
2
0
-2
5
0
5
10
2
0
2
10
5
0
5



Probability Plot
Ordered Values
10.0
7.5
5.0
2.5
0.0
2.5
5.0
7.5
10.0
2
1
0
1
2