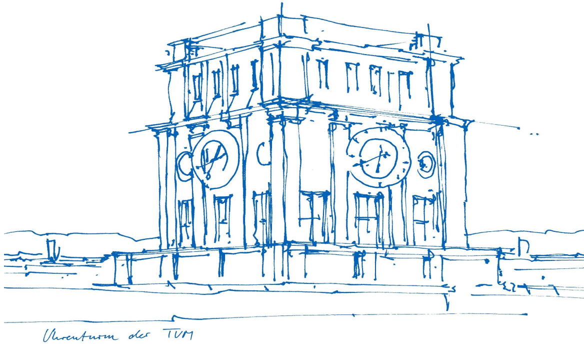
The multivariate normal distribution

Julien Gagneur

Prof. for Computational Molecular Medicine
Technical University of Munich

www.gagneurlab.in.tum.de

To understand the genetic basis of gene regulation and its implication in diseases



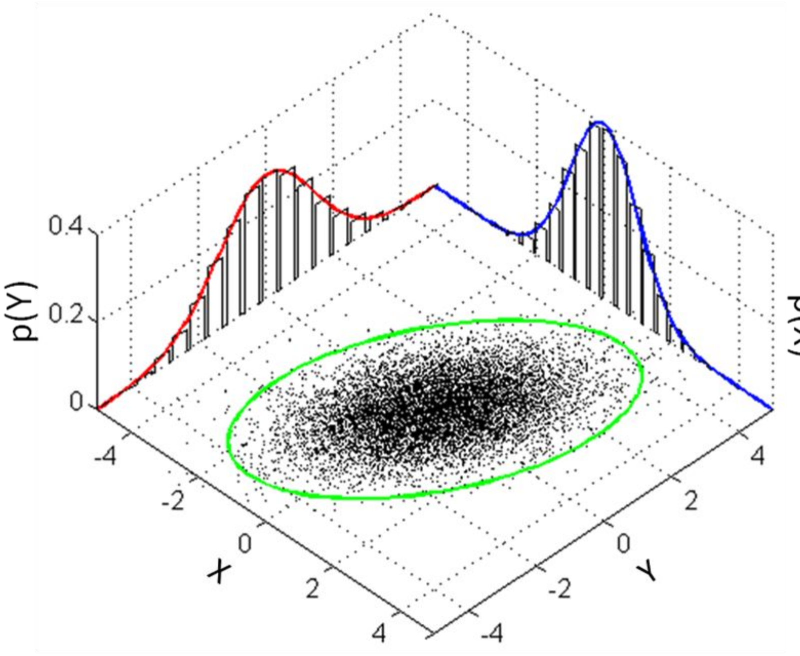
An important distribution

- The multivariate normal distribution (MVN, also called multivariate Gaussian) is the multivariate generalization of the normal distribution

- It often arises due to the central limit theorem

- It is the workhorse of a large set of statistical methods

- This lecture overviews its most important properties



Multivariate Normal Distribution (source: Wikipedia)

Definition

- We consider the joint distribution of *p* scalar random variables

$$\mathbf{x} = (x_1, ..., x_p)^\top$$

- MVN probability density

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{p/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right\}$$

- Mean vector

$$\boldsymbol{\mu} = (\mu_1, ..., \mu_p)^\top$$

- Covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & \sigma_{i,j} & \sigma_i^2 & & \\ & & & \ddots & \\ & & & & \sigma_p^2 \end{pmatrix}$$

$$\begin{aligned} \sigma_{i,j} &= \mathrm{E}[(x_i - \mu_i)(x_j - \mu_j)] \\ &= \rho_{i,j}\sigma_i \sigma j \end{aligned}$$

Special case: Independent Gaussian variables

- For MVNs, Pearson correlation = 0 is equivalent to independence

$$\rho_{i,j} = 0 \iff x_i \perp\!\!\!\perp x_j$$

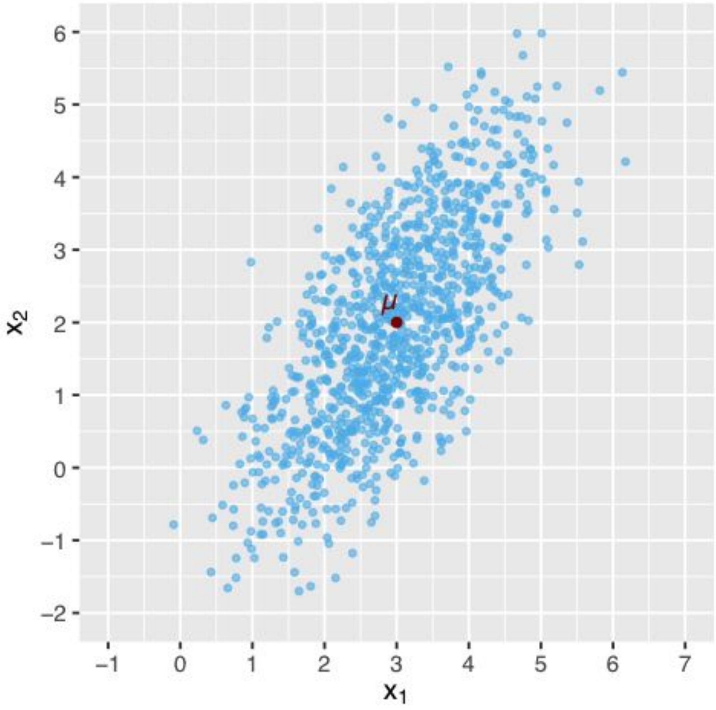
- Independent univariate Gaussian can be jointly modeled as:

$$p(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{pmatrix}$$

Density

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{p/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}$$



1,000 random draws with

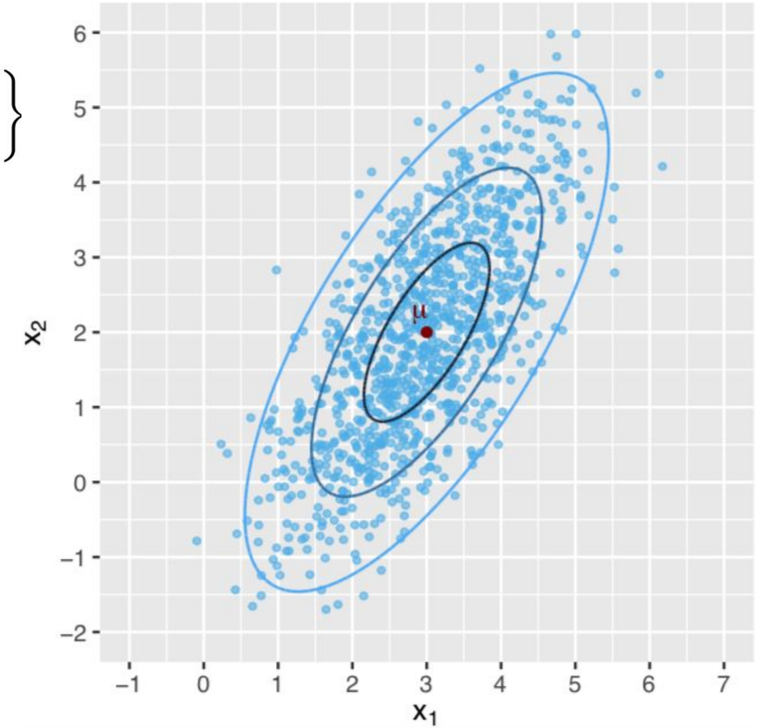
$$\boldsymbol{\mu} = (3, 2)^\top \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

Surfaces of equi-likelihood are ellipsoids

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{p/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}$$

- Squared Mahalonobis distance

$$\Delta^2 = (\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})$$



1,000 random draws with

$$\boldsymbol{\mu} = (3, 2)^\top \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

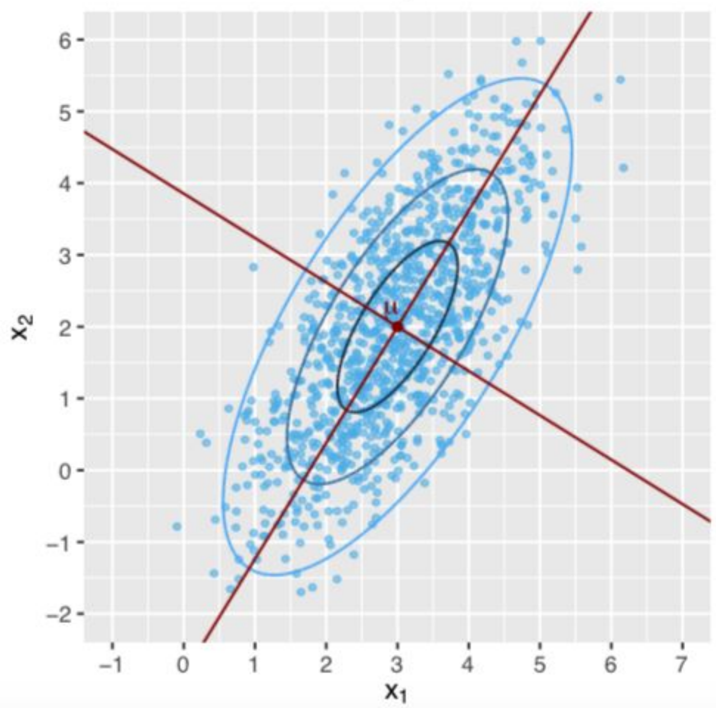
Eigendirections of the covariance

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{p/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}$$

- Squared Mahalonobis distance

$$\Delta^2 = (\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})$$

- The covariance matrix is symmetric positive definite
- The axes are the eigendirections of the covariance



1,000 random draws with

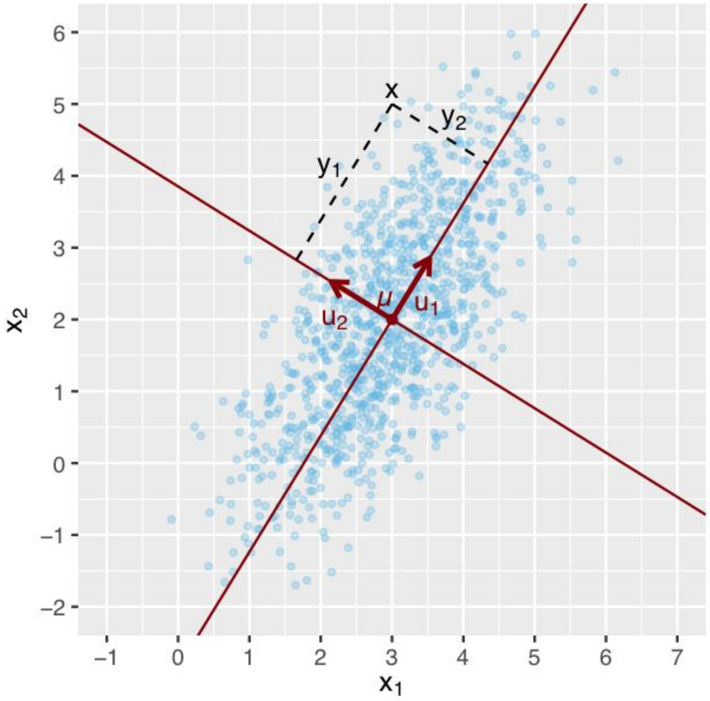
$$\boldsymbol{\mu} = (3, 2)^\top \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

Independence of transformed coordinates

- Eigenvalues $\lambda_1 > ... > \lambda_p$
- Eigenvectors $\mathbf{u}_1, ..., \mathbf{u}_p$
- Transformed coordinates $y_i = \mathbf{u}_i^\top(\mathbf{x} - \boldsymbol{\mu})$
- MVN and independence

$$p(\mathbf{y}) = \mathcal{N}(\mathbf{y}|\mathbf{0}, \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_p \end{pmatrix})$$

- This transformation is used in data analysis (Principal Component Analysis), even for non-Gaussian data



Conclusion

- MVN probability density $$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{p/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}$$

- Geometry

- Transformed variables

- Details in Bishop, CM. 2007 (Springer)
  Pattern Recognition and Machine Learning