Preliminaries

Wessel N. van Wieringen
w.n.van.wieringen@vu.nl

Department of Epidemiology and Biostatistics, VUmc
& Department of Mathematics, VU University
Amsterdam, The Netherlands

Singular value decomposition

Singular value decomposition

An (*n* x *p*)-dimensional matrix **X** is linear map:

$$f : \mathbb{R}^p \mapsto \mathbb{R}^n$$

Any element in $\mathbf{R}^n$ is formed from a linear combination of those of $\mathbf{R}^p$.

Singular value decomposition:
→ the transformation is expressible as a diagonal matrix,
→ when choosing suitable bases for $\mathbf{R}^p$ and $\mathbf{R}^n$.
Conveniently, these bases are orthonormal.

Singular value decomposition

The *singular value decomposition* of a (n x p)-dimensional matrix **X**, with $p > n$, is:

$$\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^{\top} = \sum_k d_k \mathbf{U}_{*,k} \mathbf{V}_{*,k}^{\top}$$

where:

$\mathbf{D}$ (n x n)-diagonal matrix with the nonzero singular values,

$\mathbf{U}$ (n x n)-matrix with columns containing the left singular vectors, and

$\mathbf{V}$ (p x n)-matrix with columns containing the right singular vectors.

such that:

$$\mathbf{X} \quad \mathbf{V}_{*,k} = d_k \mathbf{U}_{*,k}$$
$$\mathbf{X}^{\top}\mathbf{U}_{*,k} = d_k \mathbf{V}_{*,k}$$

Singular value decomposition

The singular vectors, both left and right, are orthonormal.

The columns of a matrix $\mathbf{X}$ are *orthonormal* if the columns are orthogonal and have a unit length. E.g.:

$$\mathbf{X} = \frac{1}{2}\begin{pmatrix} -1 & -1 \\ -1 & 1 \\ 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Then, clearly, e.g.:

$$\begin{aligned} \langle \mathbf{X}_{*,1}, \mathbf{X}_{*,1} \rangle &= \tfrac{1}{4}(-1 \times -1 + -1 \times -1 + 1 \times 1 + 1 \times 1) &= 1 \\ \langle \mathbf{X}_{*,1}, \mathbf{X}_{*,2} \rangle &= \tfrac{1}{4}(-1 \times -1 + -1 \times 1 + 1 \times -1 + 1 \times 1) &= 0 \end{aligned}$$

Singular value decomposition

In the singular value decomposition, the orthonormality of the columns of **U** and **V** implies that:

$$\mathbf{U}^\top \mathbf{U} = \mathbf{I}_{nn}$$
$$\mathbf{V}^\top \mathbf{V} = \mathbf{I}_{nn}$$

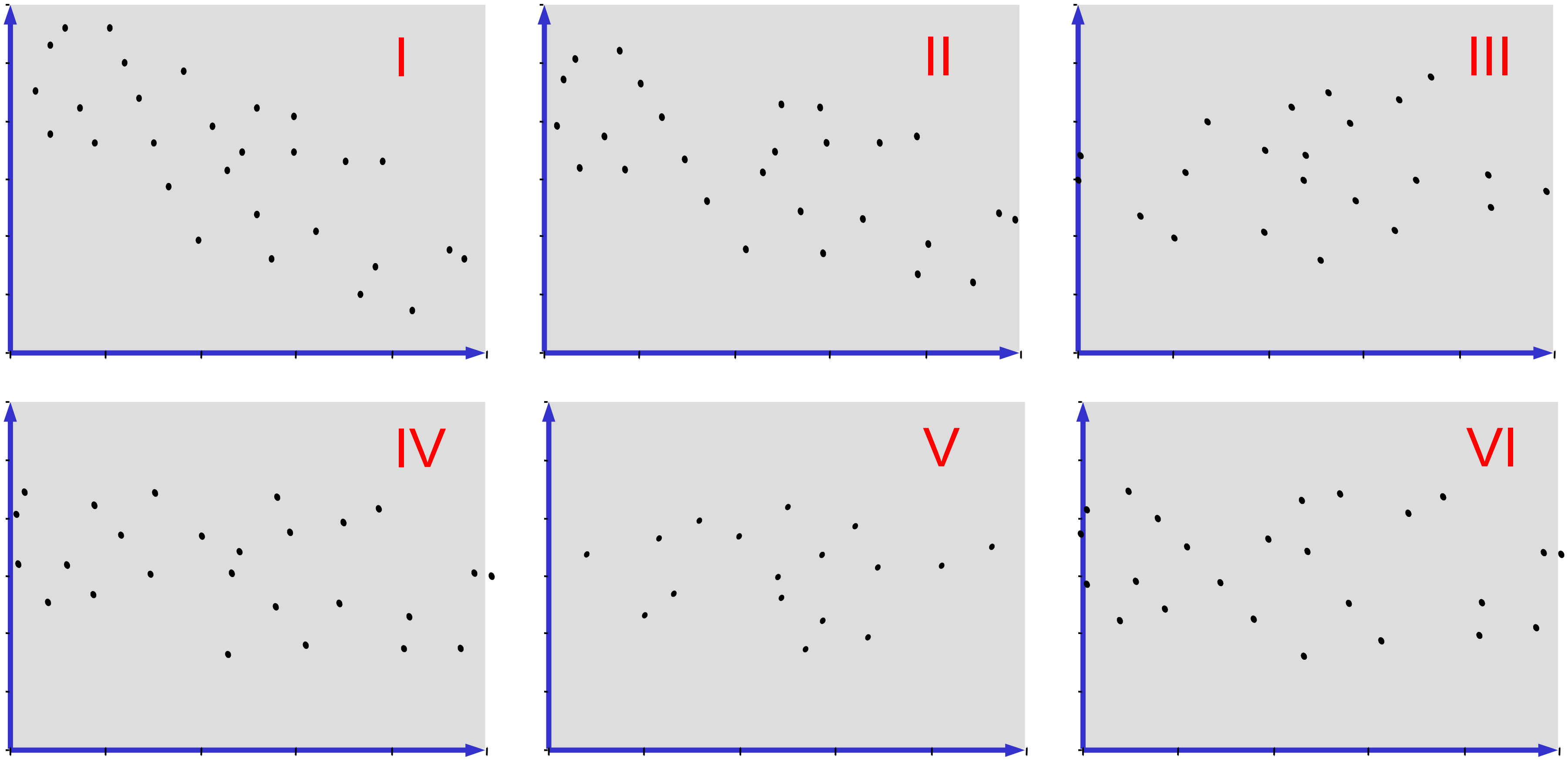
Moreover, the singular values of $\mathbf{X}$ are the square roots of the nonzero eigenvalues of $\mathbf{X}^\mathrm{T}\mathbf{X}$.

Question: show.

Multivariate normal distribution

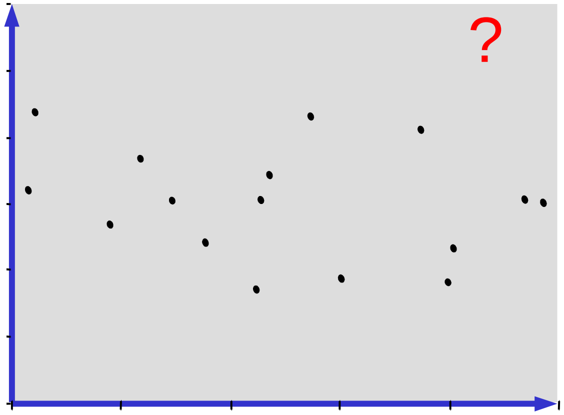
Covariance matrix

Scatterplots of data on two random variables.
Which show association?



Covariance matrix

Association between two random variables may be assessed graphically. This is not very exact and in boundary cases difficult to reach consensus.



Ideally, a measure of interrelatedness of the two variables.

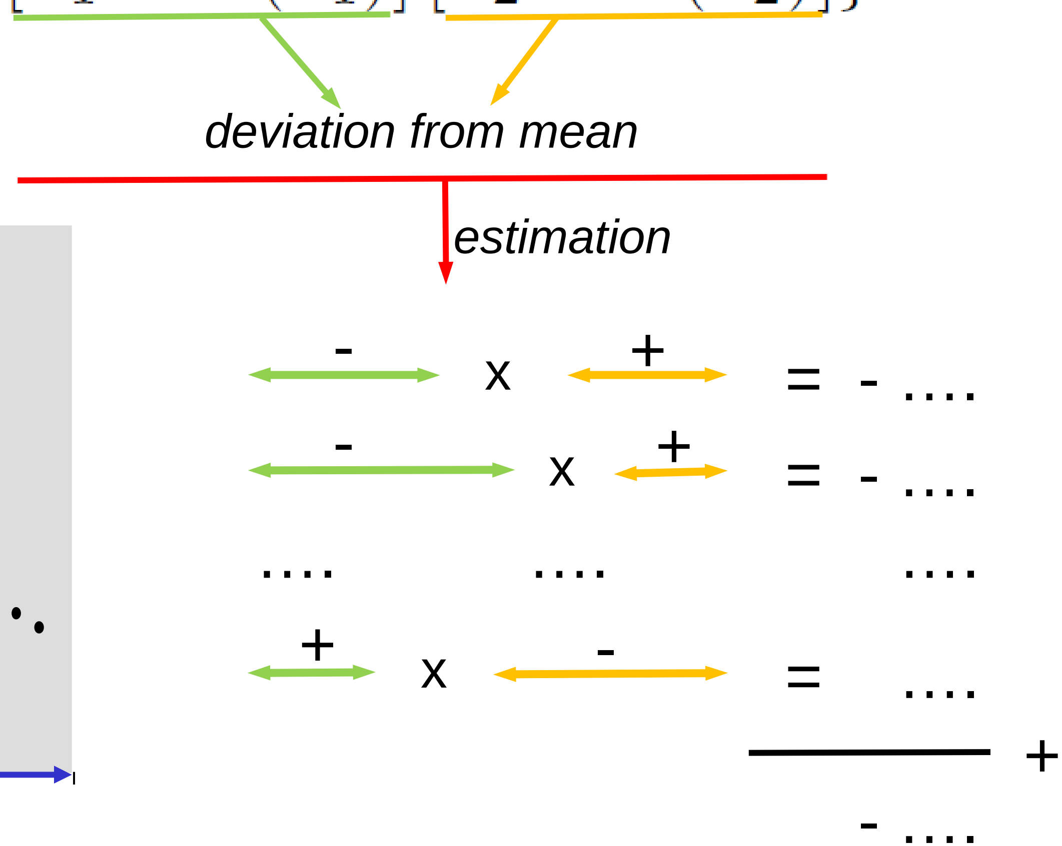
Covariance is such a measure. It measures whether a positive deviation from the mean in one variable systemically coincides with a positive (or negative) deviation from the mean in another variable.

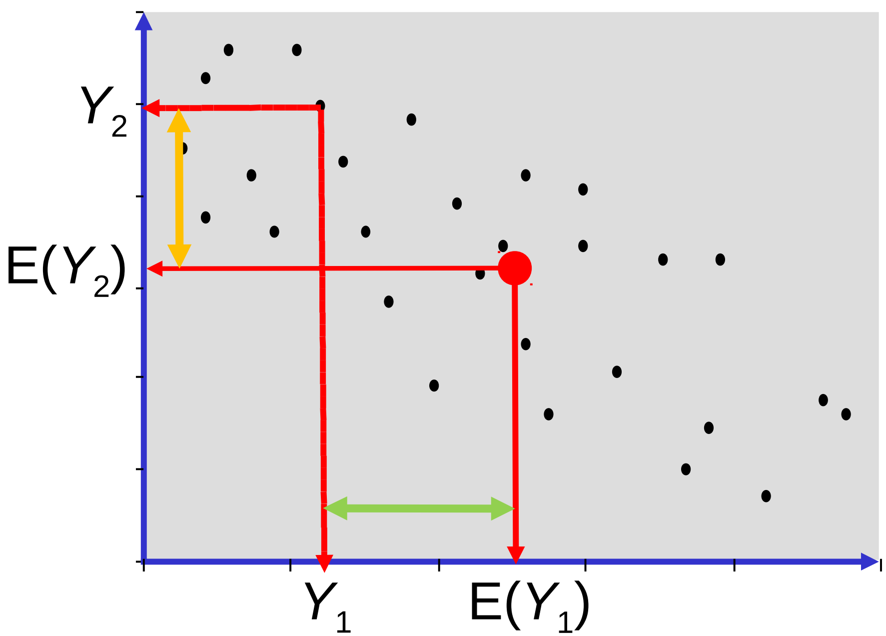
Covariance matrix

Covariance measures the linear dependence between two random variables.

The covariance between random variables Y_1 and Y_2 is:

$$\mathrm{Cov}(Y_1, Y_2) \quad = \quad E\{[Y_1 - E(Y_1)]\,[Y_2 - E(Y_2)]\}$$



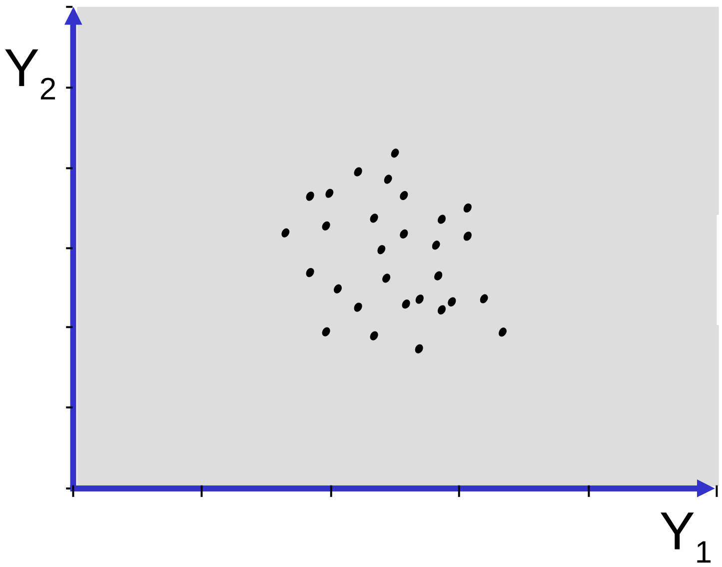


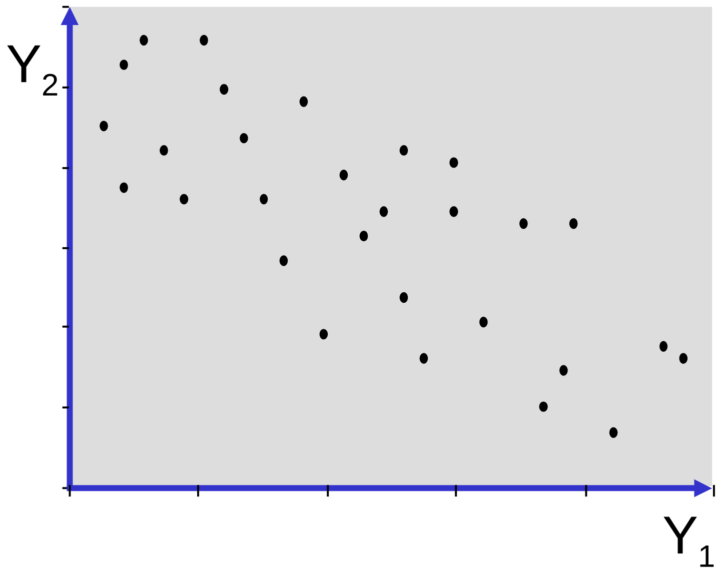
Covariance matrix

Example

$$\begin{cases} \varepsilon_1, \varepsilon_2 & \sim & \mathcal{N}(0, \sigma_\varepsilon^2) \text{ i.i.d.} \\ Y_1 & = & \varepsilon_1 \\ Y_2 & = & \varepsilon_2 \end{cases}$$

$$\begin{cases} \varepsilon_1, \varepsilon_2 & \sim & \mathcal{N}(0, \sigma_\varepsilon^2) \text{ i.i.d.} \\ Y_1 & = & \varepsilon_1 \\ Y_2 & = & \beta Y_1 + \varepsilon_2 \end{cases}$$





$$\mathrm{Cov}(Y_1, Y_2) = \mathrm{Cov}(\varepsilon_1, \varepsilon_2) = 0$$

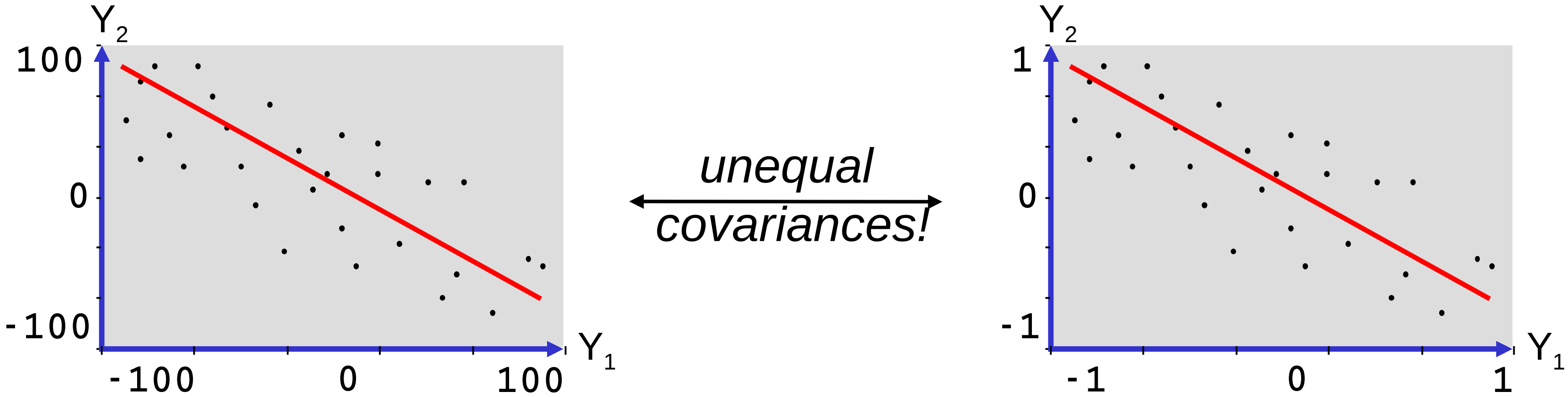
$$\mathrm{Cov}(Y_1, Y_2) = \ ???$$

Covariance matrix

Example

$$\begin{cases} \varepsilon_1, \varepsilon_2 & \sim & \mathcal{N}(0, \sigma_\varepsilon^2) \text{ i.i.d.} \\ Y_1 & = & \varepsilon_1 \\ Y_2 & = & \beta Y_1 + \varepsilon_2 \end{cases} \longrightarrow \quad \mathrm{Cov}(Y_1, Y_2) = \beta \sigma_\varepsilon^2$$

Covariance thus depends on variance of Y_1, while linear relation (β) between Y_1 and Y_2 is unchanged.



Undesirable property for a measure of linear dependence.

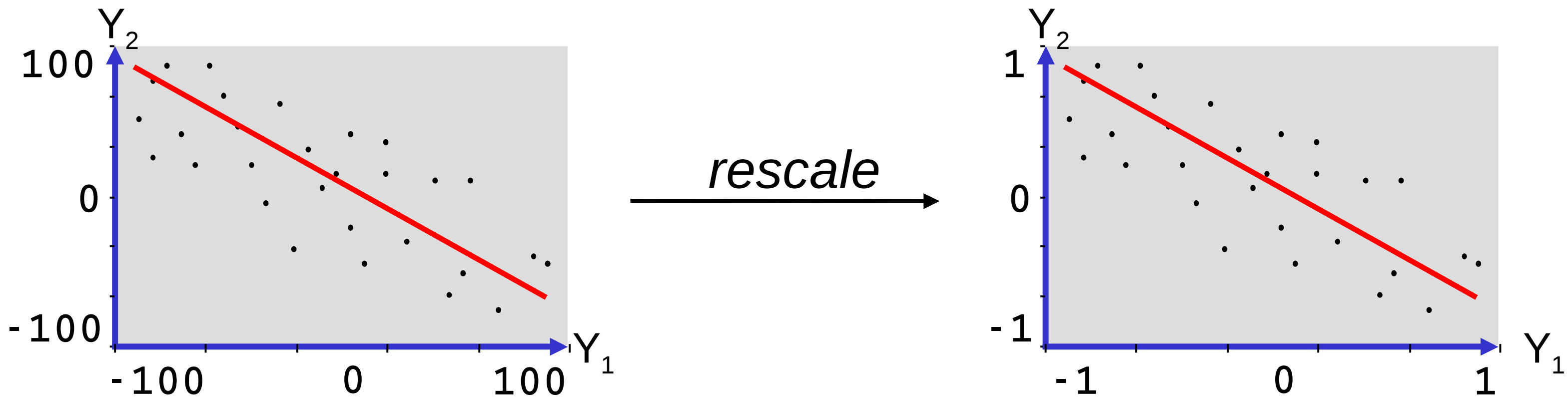
Covariance matrix

Solution

Standardization of Y_1 and Y_2

$$\mathrm{Cov}(\tilde{Y}_1, \tilde{Y}_2) = \mathrm{Cov}(a_1 Y_1, a_2 Y_2) = a_1 a_2 \mathrm{Cov}(Y_1, Y_2)$$

with $a_j = [\mathrm{Var}(Y_j)]^{-1/2}$



$$\mathrm{Cov}(Y_1, Y_2) = \mathbf{8483.662}$$
$$\mathrm{Cov}(\tilde{Y}_1, \tilde{Y}_2) = \mathbf{0.675}$$

$$\mathrm{Cov}(Y_1, Y_2) = \mathbf{1.007}$$
$$\mathrm{Cov}(\tilde{Y}_1, \tilde{Y}_2) = \mathbf{0.697}$$

Covariance matrix

Pearson's correlation coefficient

Normalized covariance between Y_1 and Y_2:

$$\rho(Y_1, Y_2) = \mathrm{Cor}(Y_1, Y_2) = \frac{\mathrm{Cov}(Y_1, Y_2)}{\sqrt{\mathrm{Var}(Y_1)}\sqrt{\mathrm{Var}(Y_2)}}$$

It measures the degree of linear dependence between the two random variables Y_1 and Y_2.

$\rho(Y_1, Y_2)$ in [-1, 1], with

→ $\rho = 1$: perfect positive linear relationship.

→ $\rho = 0$: absence of linear dependency.

→ $\rho = -1$: perfect negative linear relationship.

Closer $|\rho|$ to one: stronger linear dependency.

Covariance matrix

Covariance matrix

The definition of covariance extends to random vectors:

$$\begin{aligned}
&\mathrm{Cov}(\mathbf{X}, \mathbf{Y}) \\
&= E\{[\mathbf{X} - E(\mathbf{X})][\mathbf{Y} - E(\mathbf{Y})]^\top\} \\
&= \int_{\mathbb{R}^p}\int_{\mathbb{R}^p} [\mathbf{x} - E(\mathbf{X})][\mathbf{y} - E(\mathbf{Y})]^\top f_{(\mathbf{X},\mathbf{Y})}(\mathbf{x}, \mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}
\end{aligned}$$

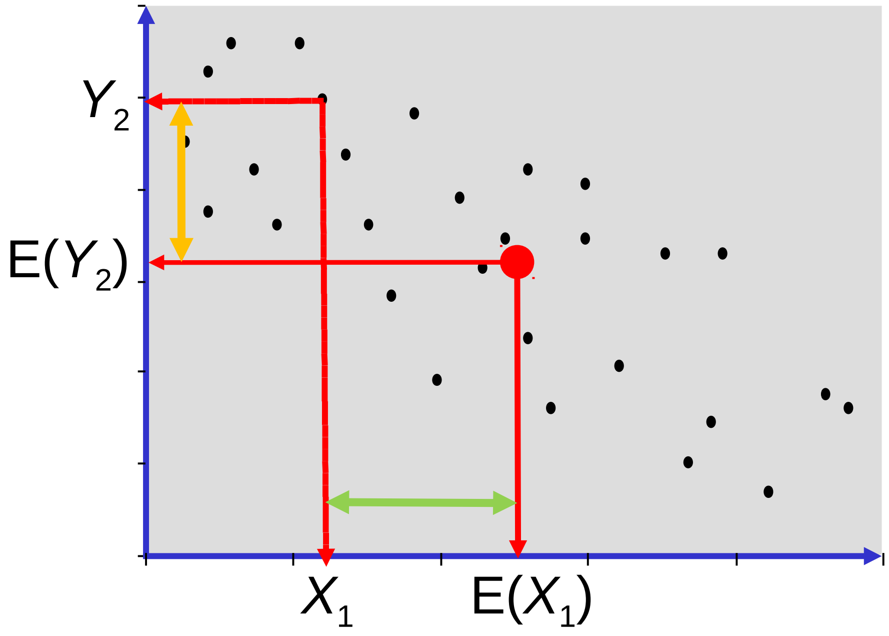
No longer a scalar, covariance is now a *pxp* matrix:

$$\begin{pmatrix}
\mathrm{Cov}[(\mathbf{X})_1, (\mathbf{Y})_1] & \dots & \mathrm{Cov}[(\mathbf{X})_1, (\mathbf{Y})_p] \\
\vdots & \ddots & \vdots \\
\mathrm{Cov}[(\mathbf{X})_p, (\mathbf{Y})_1] & \dots & \mathrm{Cov}[(\mathbf{X})_p, (\mathbf{Y})_p]
\end{pmatrix}$$

Covariance matrix

Covariance matrix

The elements of a covariance matrix are the pairwise covariances of the elements of random vectors **X** and **Y**:



$$
\begin{aligned}
[\mathrm{Cov}(\mathbf{X}, \mathbf{Y})]_{1,2} \\
&= \mathrm{Cov}[(\mathbf{X})_1, (\mathbf{Y})_2] \\
&= \mathrm{Cov}(X_1, Y_2) \\
&= E\{[X_1 - E(X_1)] \\
&\qquad [Y_2 - E(Y_2)]\}
\end{aligned}
$$

Covariance matrix

Question

Consider the random vector $\mathbf{Y} = (Y_1, Y_2, Y_3)^\top$ with covariance matrix:

$$\mathrm{Cov}(\mathbf{Y}, \mathbf{Y}) = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 3 & 1 \\ -1 & 1 & 1 \end{pmatrix}$$

→ What is the meaning of the diagonal elements?
→ Why is the above matrix symmetric?
→ What does the value of (1,2) element imply?

Covariance matrix

Covariance matrix properties (I)

Let **X** and **Y** be two independent multivariate random variables. Then:

$$\mathrm{Cov}(\mathbf{X}, \mathbf{Y}) = 0$$

Let **X** be a multivariate random variable and **c** a vector with constants. Then:

$$\mathrm{Cov}(\mathbf{c}, \mathbf{X}) = 0$$

Let **Y** be a multivariate random variable. Then:

$$\mathrm{Cov}(\mathbf{Y}, \mathbf{Y}) = \mathrm{Var}(\mathbf{Y})$$

Covariance matrix

Covariance matrix properties (II)

Let **W**, **X, Y** and **Z** be multivariate random variables. Then:

$$\begin{aligned}\mathrm{Cov}(\mathbf{W}+\mathbf{X},\mathbf{Y}+\mathbf{Z}) \\ &= \mathrm{Cov}(\mathbf{W},\mathbf{Y})+\mathrm{Cov}(\mathbf{W},\mathbf{Z}) \\ &\quad +\mathrm{Cov}(\mathbf{X},\mathbf{Y})+\mathrm{Cov}(\mathbf{X},\mathbf{Z})\end{aligned}$$

Let **X** and **Y** be two multivariate random variables and **A** and **B** coefficient matrices. Then:

$$\mathrm{Cov}(\mathbf{A}\mathbf{X},\mathbf{B}\mathbf{Y}) = \mathbf{A}\,\mathrm{Cov}(\mathbf{X},\mathbf{Y})\mathbf{B}^{\mathrm{T}}$$

Covariance matrix

Correlation matrix

Similarly, the correlation between two random vectors is:

$$\mathrm{Cor}(\mathbf{X}, \mathbf{Y}) = \begin{pmatrix} \mathrm{Cor}[(\mathbf{X})_1, (\mathbf{Y})_1] & \dots & \mathrm{Cor}[(\mathbf{X})_1, (\mathbf{Y})_p] \\ \vdots & \ddots & \vdots \\ \mathrm{Cor}[(\mathbf{X})_p, (\mathbf{Y})_1] & \dots & \mathrm{Cor}[(\mathbf{X})_p, (\mathbf{Y})_p] \end{pmatrix}$$

with e.g.:

$$\mathrm{Cor}[(\mathbf{X})_1, (\mathbf{Y})_2] = \mathrm{Cor}(X_1, Y_2) = \frac{\mathrm{Cov}(X_1, Y_2)}{\sqrt{\mathrm{Var}(X_1)}\sqrt{\mathrm{Var}(Y_2)}}$$

The correlation matrix contains the pairwise correlations.

Multivariate normal distribution

Denote a *p*-dimensional $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_p)^{\mathrm{T}}$ random variable following a *multivariate normal distribution* by:

$$\mathbf{Y}_i \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

with a *mean* parameter:

$$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)^{\mathrm{T}} \in \mathbb{R}^p$$

and a *covariance* parameter $\boldsymbol{\Sigma} \in \mathcal{S}_{++}^p$:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix}$$

Multivariate normal distribution

Density

The *p*-variate normal distribution has density f(**Y**$_i$) equal to:

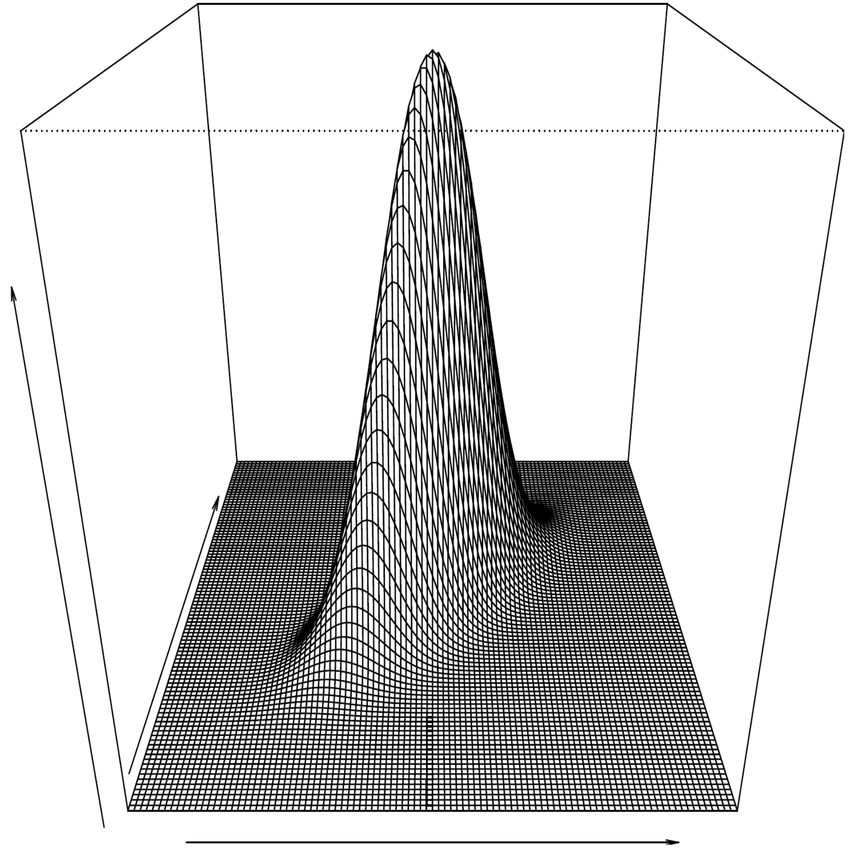
$$\frac{1}{(2\pi)^{p/2}\,|\boldsymbol{\Sigma}|^{1/2}}\,\exp\left[-\frac{1}{2}(\mathbf{Y}_i-\boldsymbol{\mu})^T\,\boldsymbol{\Sigma}^{-1}\,(\mathbf{Y}_i-\boldsymbol{\mu})\right]$$

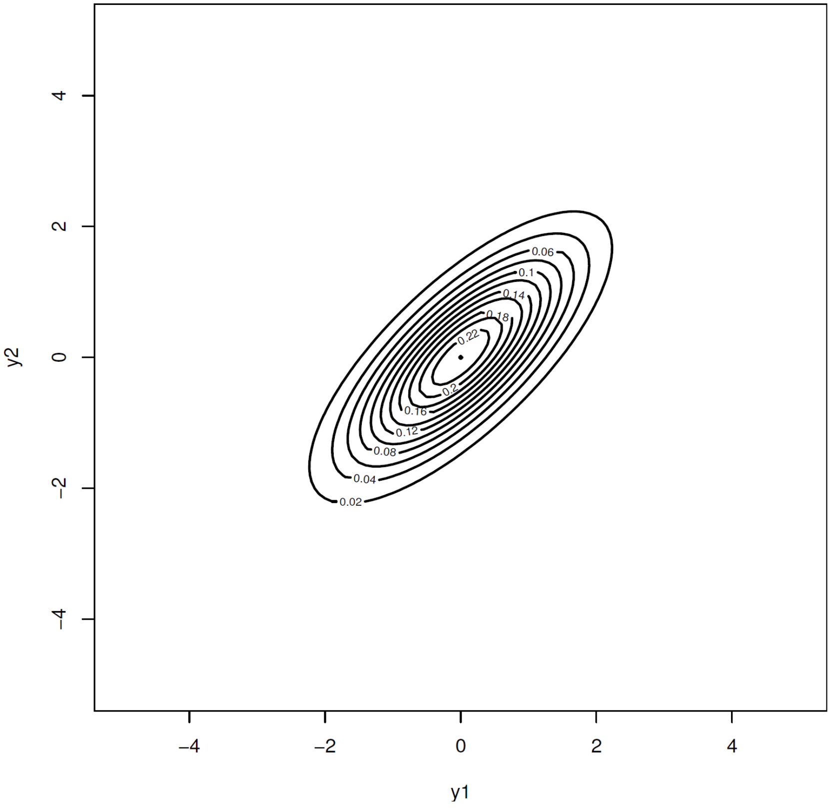
Recall the univariate normal distribution density:

$$\begin{aligned} f(Y_i) &= \frac{1}{\sqrt{2\pi\sigma^2}}\,\exp\left[-\frac{1}{2}(Y_i-\mu)^2/\sigma^2\right] \\ &= \frac{1}{(2\pi)^{1/2}\,\sigma}\,\exp\left[-\frac{1}{2}(Y_i-\mu)\,\sigma^{-2}\,(Y_i-\mu)\right] \end{aligned}$$

Multivariate normal distribution

The density of a bivariate (p=2) normal distribution.

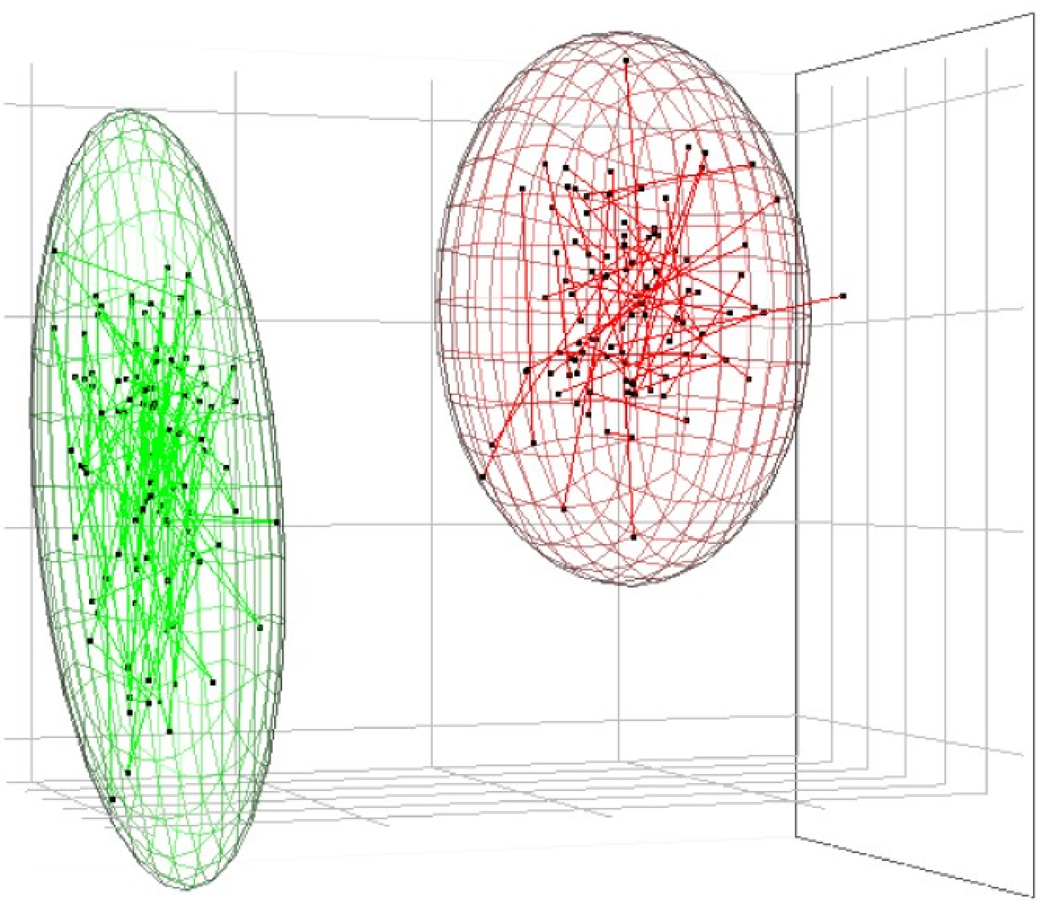


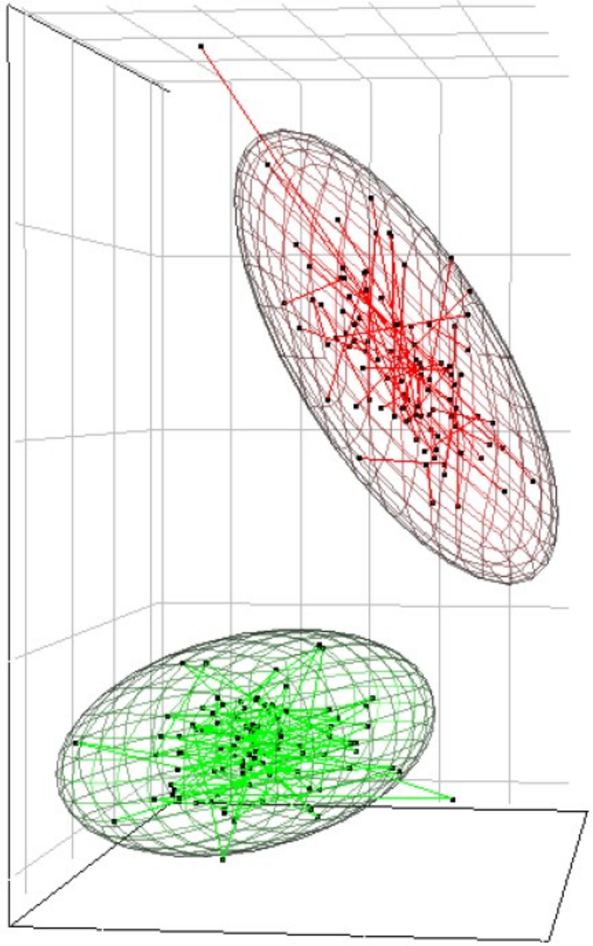


Density represented by level sets: {**Y** : f(**Y**) = c}. Observations with equal likelihood.

Multivariate normal distribution

Data distribution of trivariate (*p*=3) normal distributions.





Multivariate normal distribution

Standard multivariate normal

The random variable $\mathbf{Y} = (Y_1, Y_2, Y_3)^\top$ is standard normally distributed if:

$$\boldsymbol{\mu} = \mathbf{0}_{p\times 1} \quad \text{and} \quad \boldsymbol{\Sigma} = \mathbf{I}_{p\times p}$$

Thus:

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{pmatrix} \right)$$

Put differently:

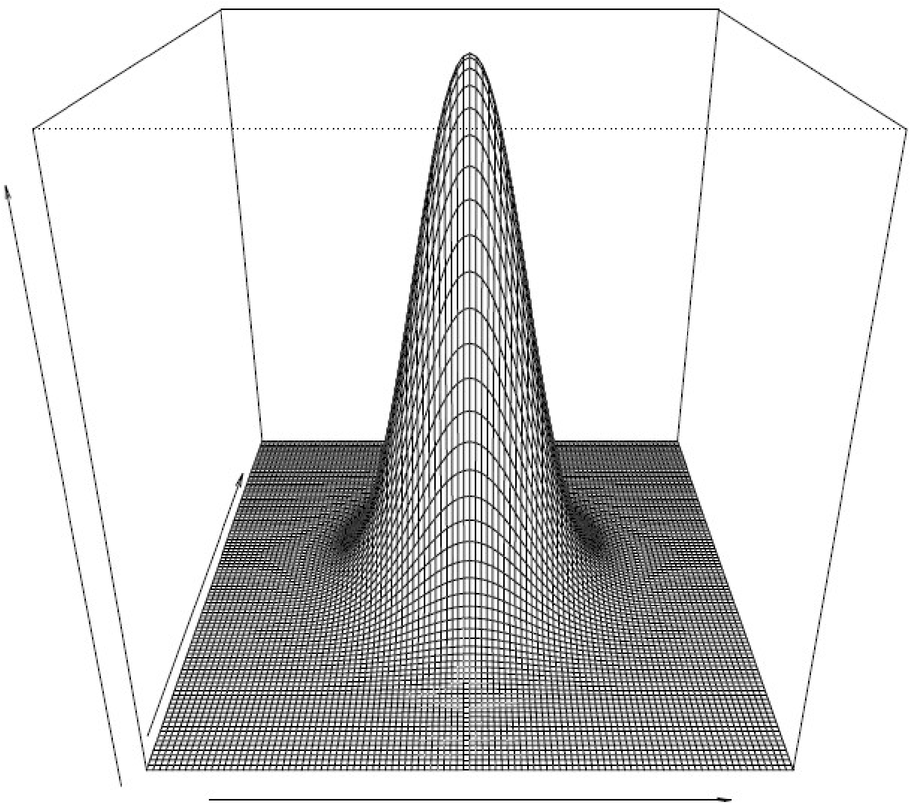
$$Y_j \text{ i.i.d. with } Y_j \sim \mathcal{N}(0,1)$$

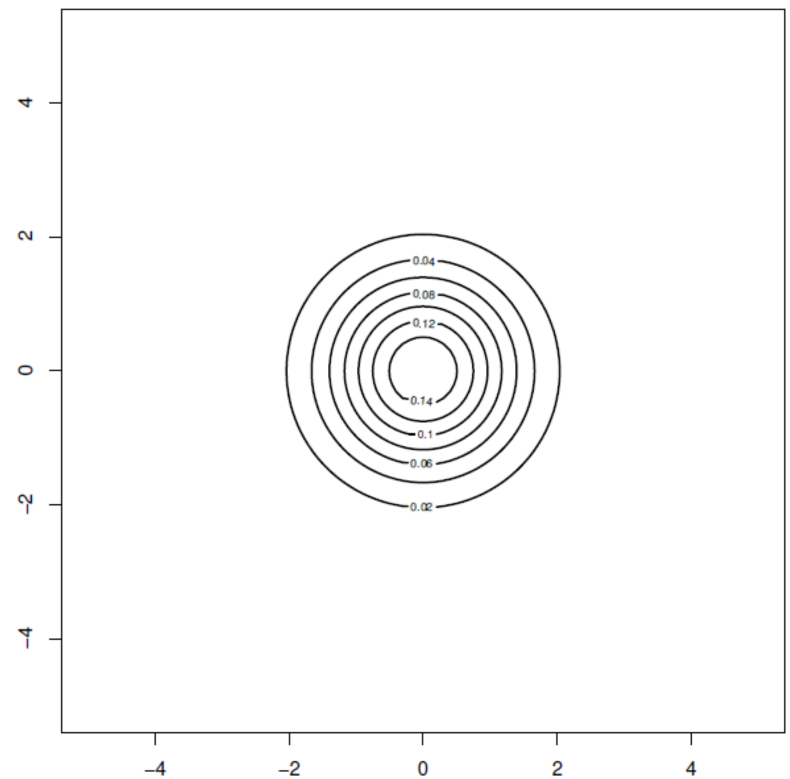
Multivariate normal distribution

Standard bivariate normal

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

$$\begin{cases} Y_1 \sim \mathcal{N}(0,1), \\ Y_2 \sim \mathcal{N}(0,1), \\ Y_1 \perp\!\!\!\perp Y_2 \end{cases}$$





Multivariate normal distribution

Any multivariate normal random variable can be derived from the standard normal one.

Let:

$$\mathbf{Z} \sim \mathcal{N}(\mathbf{0}_{p\times 1}, \mathbf{I}_{p\times p})$$

$$\boldsymbol{\mu} \in \mathbb{R}^p$$

$\mathbf{L} \in \mathcal{M}^p$ such that $\operatorname{rank}(\mathbf{L}) = p$

Now define:

$$\mathbf{Y} = \boldsymbol{\mu} + \mathbf{L}\mathbf{Z}$$

Then:

$$\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{L}\mathbf{L}^\top)$$

Multivariate normal distribution

Question

Let the random variable $\mathbf{Y}$ be defined as on the previous slide. Verify:

$$\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^\top$$

and

$$\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^\top \in \mathbf{S}_{++}^p ?$$

Hint (for part 2)

Use the singular value decomposition of $\mathbf{L}$:

$$\mathbf{L} = \mathbf{U}_\ell \mathbf{D}_\ell \mathbf{V}_\ell^\top$$

Multivariate normal distribution

Bivariate normal distribution.

Recall model:

$$\begin{cases} \varepsilon_1, \varepsilon_2 & \sim & \mathcal{N}(0, \sigma_\varepsilon^2) \text{ i.i.d.} \\ Y_1 & = & \varepsilon_1 \\ Y_2 & = & \beta Y_1 + \varepsilon_2 \end{cases}$$

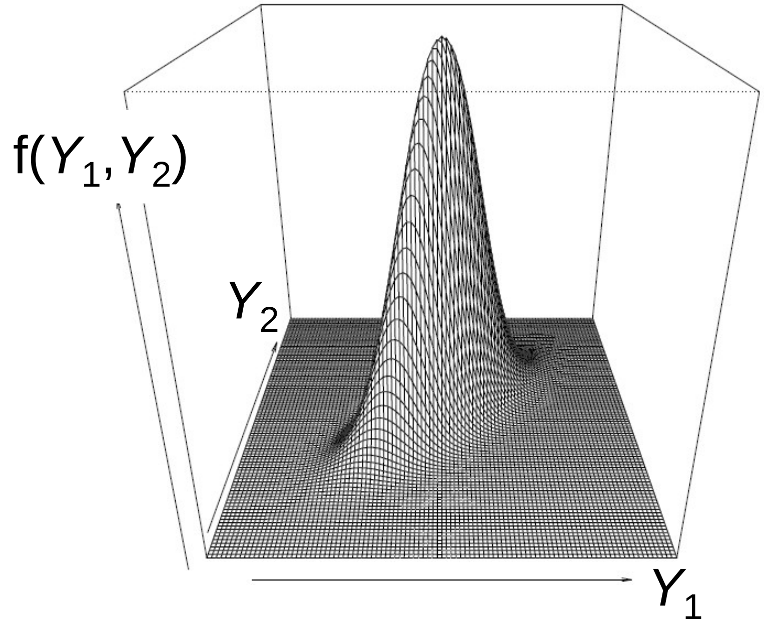
Then:

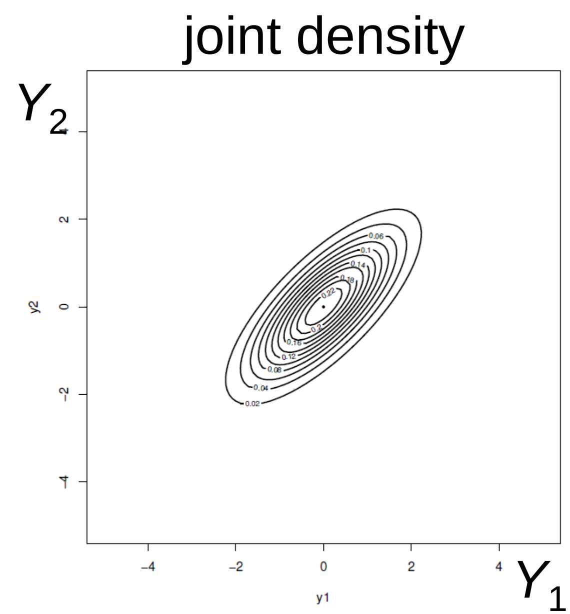
$$\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

with:

$$\boldsymbol{\mu} = (0, 0)^\top$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_\varepsilon^2 & \beta\sigma_\varepsilon^2 \\ \beta\sigma_\varepsilon^2 & (1+\beta^2)\sigma_\varepsilon \end{pmatrix}$$





Multivariate normal distribution

The matrix $\mathbf{\Sigma}$ is often parameterized as:

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} & \cdots & \sigma_1\sigma_p\rho_{1p} \\ \sigma_1\sigma_2\rho_{12} & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & \vdots \\ \sigma_1\sigma_p\rho_{1p} & \cdots & \cdots & \sigma_p^2 \end{pmatrix}$$

where:

$$\sigma_j^2 = \mathrm{Var}(Y_{ij})$$

$$\rho_{j_1,j_2} = \frac{\mathrm{Cov}(Y_{ij_1}, Y_{ij_2})}{\sqrt{\mathrm{Var}(Y_{ij_1})}\sqrt{\mathrm{Var}(Y_{ij_2})}}$$

The latter is the *correlation* between Y_{ij1} and Y_{ij2}.

Multivariate normal distribution

The parameterization in matrix form:

$$\mathbf{\Sigma} = \tilde{\mathbf{\Sigma}}_d \mathbf{R} \tilde{\mathbf{\Sigma}}_d$$

where:

$$\tilde{\mathbf{\Sigma}}_d = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \sigma_p \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{12} & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1p} & \cdots & \cdots & 1 \end{pmatrix}$$

correlation matrix

Multivariate normal distribution

From covariance to correlation matrix:

$$\mathbf{R} = \tilde{\mathbf{\Sigma}}_d^{-1} \mathbf{\Sigma} \tilde{\mathbf{\Sigma}}_d^{-1}$$

where:

$$\tilde{\mathbf{\Sigma}}_d^{-1} = \begin{pmatrix} \sigma_1^{-1} & 0 & \cdots & 0 \\ 0 & \sigma_2^{-1} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \sigma_p^{-1} \end{pmatrix}$$

Question

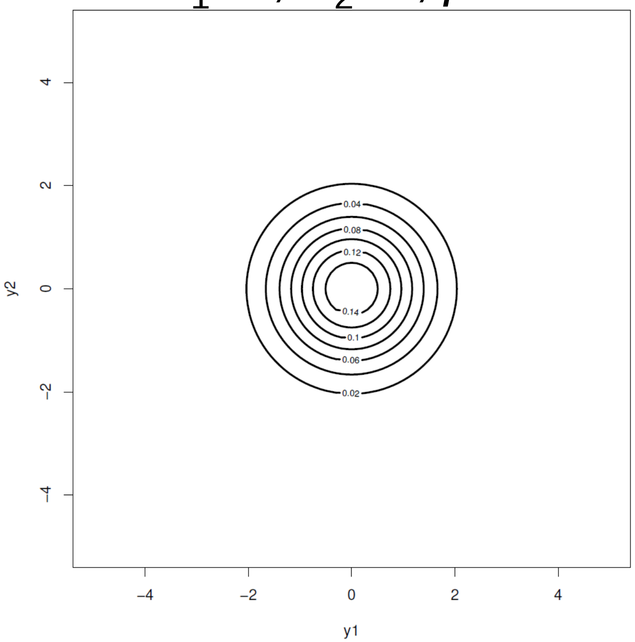
→ Verify for p=2.

→ How to go from correlation to covariance matrix?

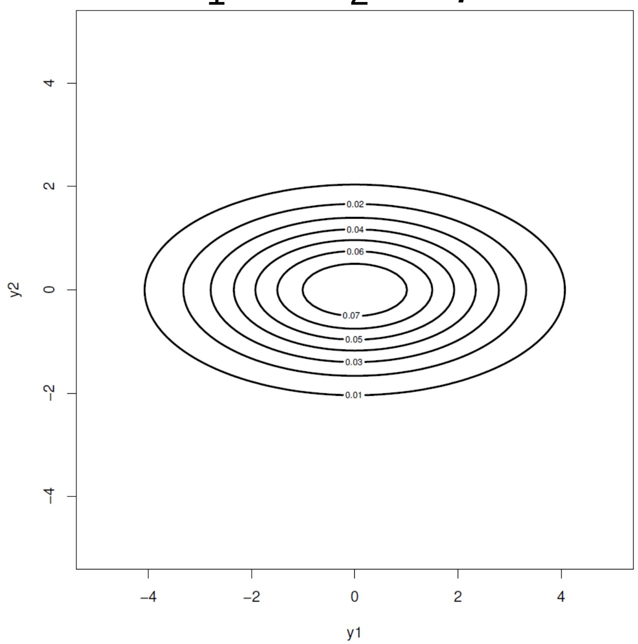
Multivariate normal distribution

Effect of σ_1, σ_2, ρ in the bivariate normal distribution.

$\sigma_1=1, \sigma_2=1, \rho=0$



$\sigma_1=2, \sigma_2=1, \rho=0$



$\sigma_1=1, \sigma_2=1, \rho=3/4$