

Normal Multivariada

Graciela Boente

Definición

- Sea $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$ simétrica y definida positiva
Se dice que $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ si su densidad está dada por

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right\}$$

Definición

- Sea $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$ simétrica y definida positiva
Se dice que $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ si su densidad está dada por

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}$$

- Si $\mathbf{X} \sim N(\mathbf{0}, \text{diag}(\lambda_1, \dots, \lambda_p))$

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{\prod_{j=1}^p \lambda_j^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} \sum_{j=1}^p \frac{x_j^2}{\lambda_j} \right\}$$

Por lo tanto, X_1, \dots, X_p son independientes $X_j \sim N(0, \lambda_j)$.

Definición

- Sea $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$ simétrica y definida positiva
Se dice que $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ si su densidad está dada por

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}$$

- Si $\mathbf{X} \sim N(\mathbf{0}, \text{diag}(\lambda_1, \dots, \lambda_p))$

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{\prod_{j=1}^p \lambda_j^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} \sum_{j=1}^p \frac{x_j^2}{\lambda_j} \right\}$$

Por lo tanto, X_1, \dots, X_p son independientes $X_j \sim N(0, \lambda_j)$.

- En particular, si $\mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_p)$

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}} e^{-\frac{1}{2} \|\mathbf{x}\|^2}$$

o sea, X_1, \dots, X_p son i.i.d. $N(0, 1)$.

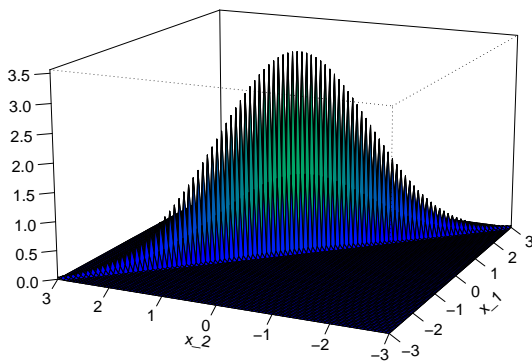
Caso $p=2$

- Sea $\boldsymbol{\mu} \in \mathbb{R}^2$ y $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ definida positiva
($|\rho| \neq 1$)

$$f(\mathbf{x}) = \frac{1}{2\pi} \frac{1}{\sigma_1\sigma_2(1-\rho^2)^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 + \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2 - 2\rho \left(\frac{x_1 - \mu_1}{\sigma_1} \right) \left(\frac{x_2 - \mu_2}{\sigma_2} \right) \right] \right\}$$

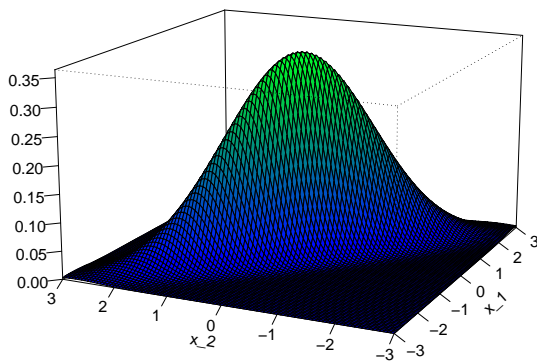
Caso $p=2$

$\rho = -0.999$



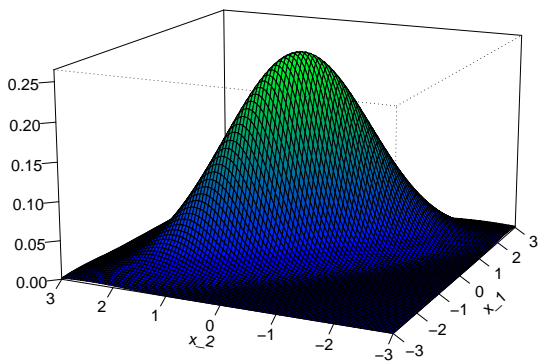
Caso $p=2$

$\rho = -0.8991$



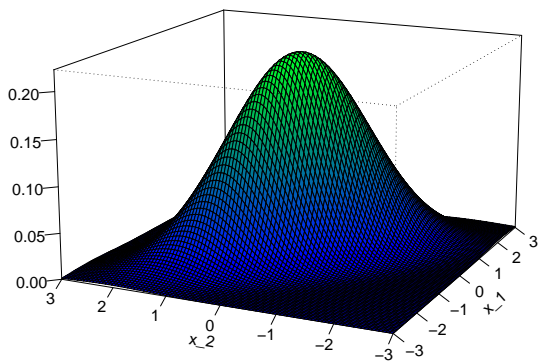
Caso $p=2$

$\rho = -0.7992$



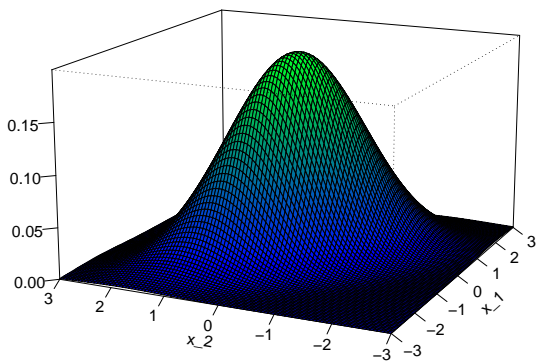
Caso $p=2$

$\rho = -0.6993$



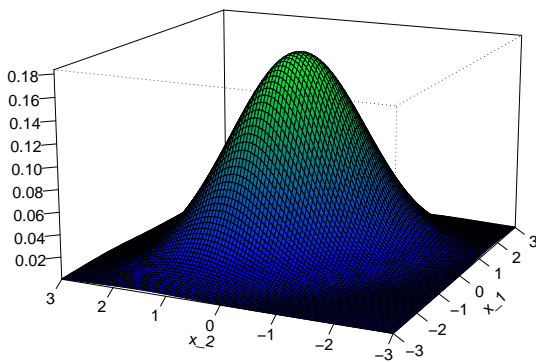
Caso $p=2$

$\rho = -0.5994$



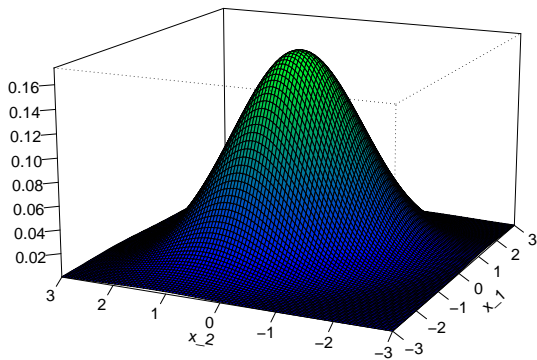
Caso $\rho=2$

$\rho = -0.4995$



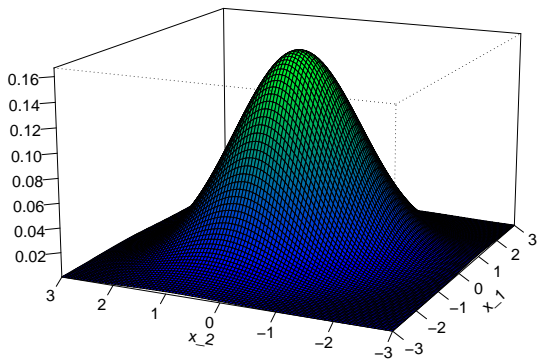
Caso $p=2$

$\rho = -0.3996$



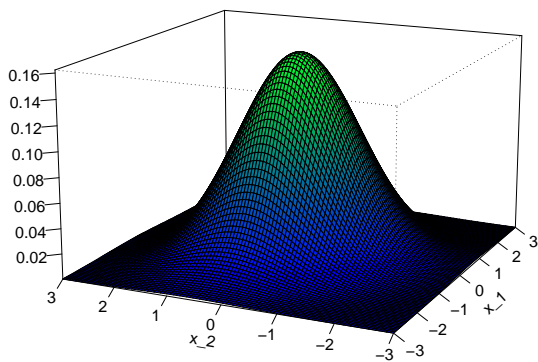
Caso $\rho=2$

$\rho = -0.2997$



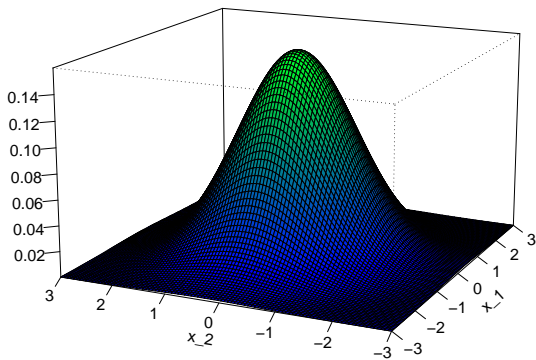
Caso $p=2$

$\rho = -0.1998$



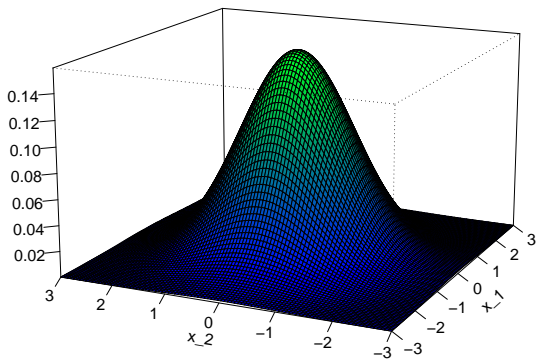
Caso $p=2$

$\rho = -0.0999$



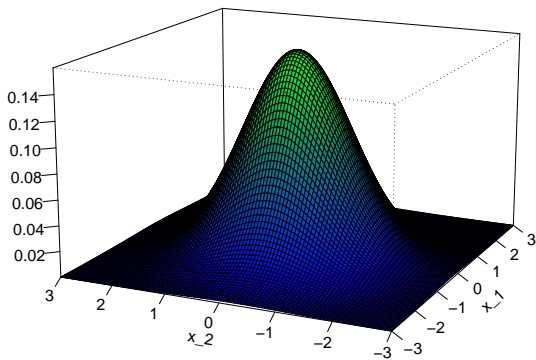
Caso $p=2$

$\rho = 0$



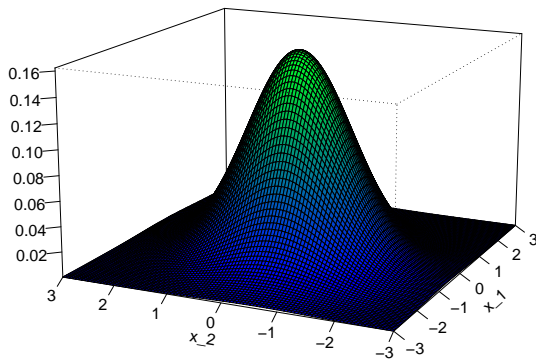
Caso $p=2$

$\rho = 0.0999$



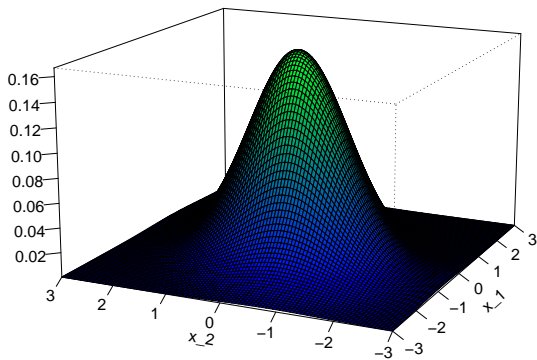
Caso $p=2$

$\rho = 0.1998$



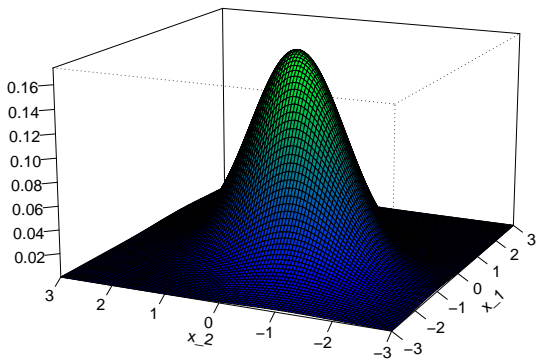
Caso $p=2$

$\rho = 0.2997$



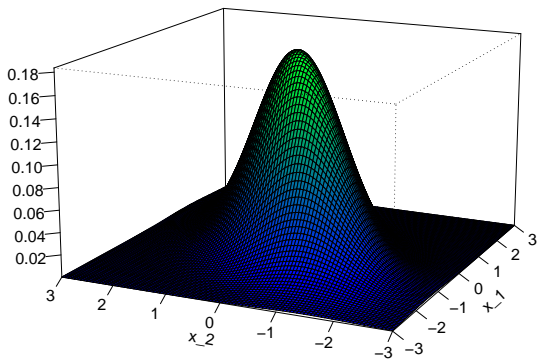
Caso $p=2$

$\rho = 0.3996$



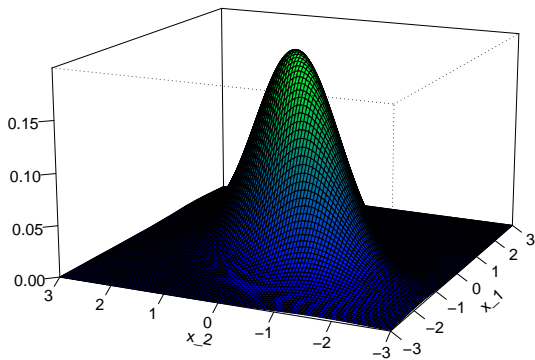
Caso $p=2$

$\rho = 0.4995$



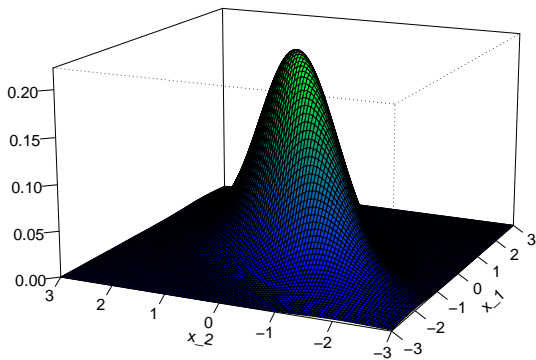
Caso $p=2$

$\rho = 0.5994$



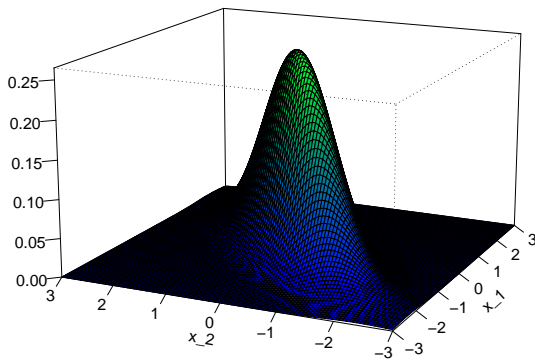
Caso $p=2$

$\rho = 0.6993$



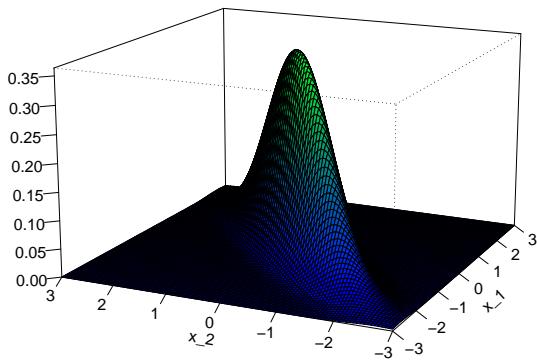
Caso $p=2$

$\rho = 0.7992$



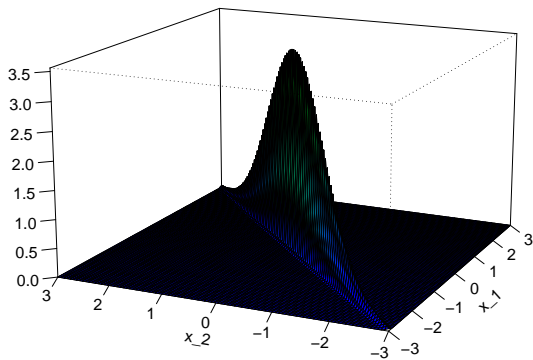
Caso $p=2$

$\rho = 0.8991$



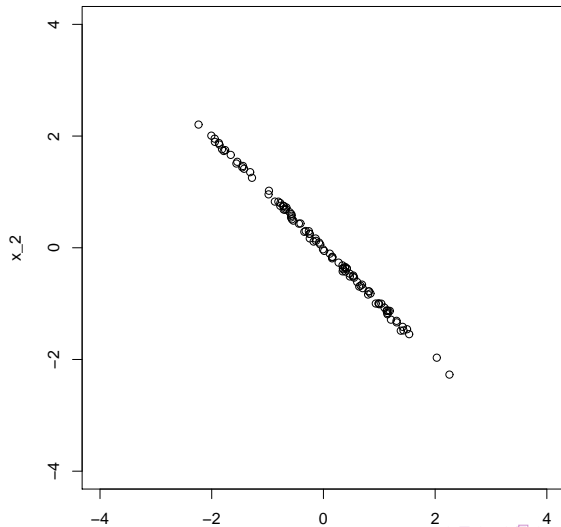
Caso $p=2$

$\rho = 0.999$



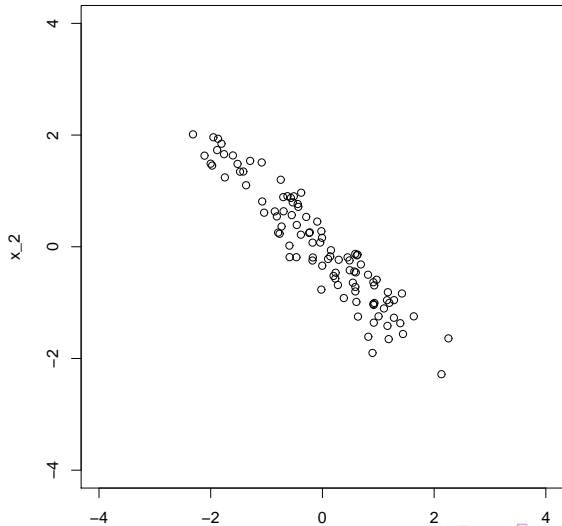
Caso $p=2$

$\rho = -0.999$



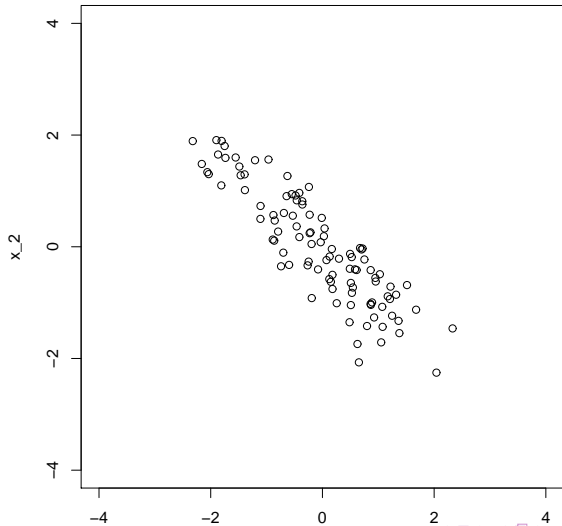
Caso $p=2$

$\rho = -0.8991$



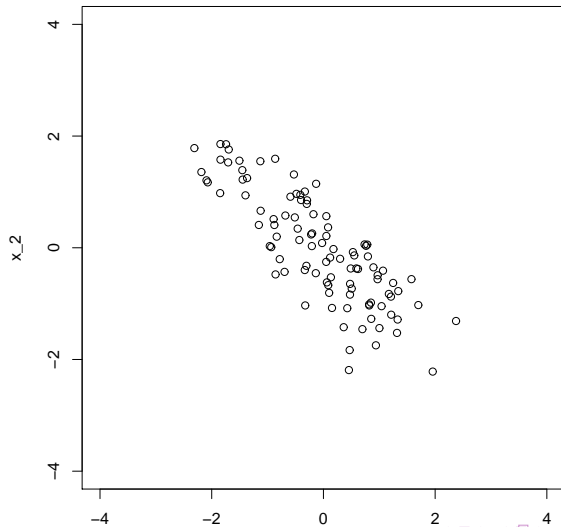
Caso $p=2$

$\rho = -0.7992$



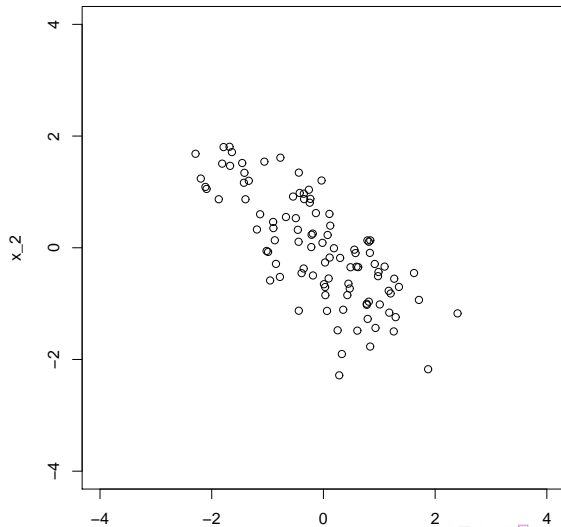
Caso $p=2$

$\rho = -0.6993$



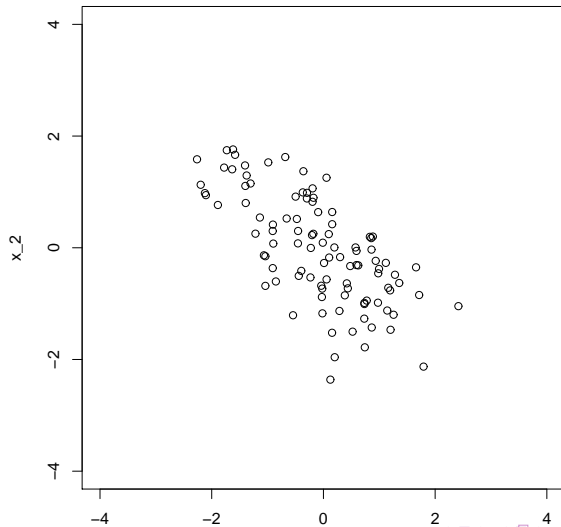
Caso $p=2$

$\rho = -0.5994$



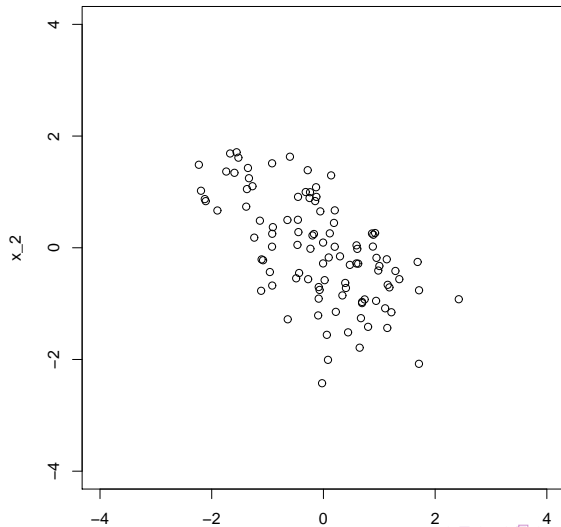
Caso $p=2$

$\rho = -0.4995$



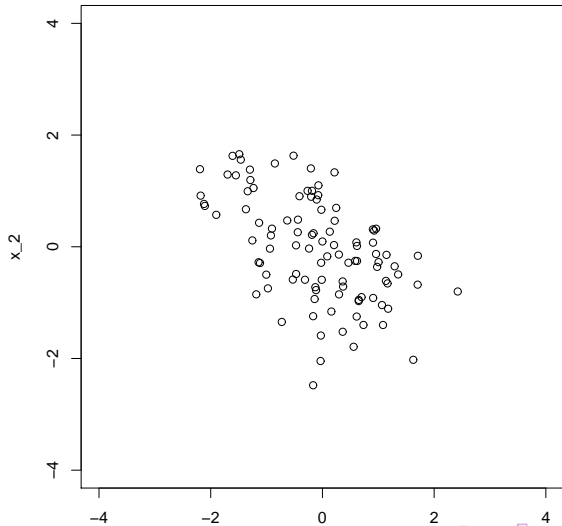
Caso $p=2$

$\rho = -0.3996$



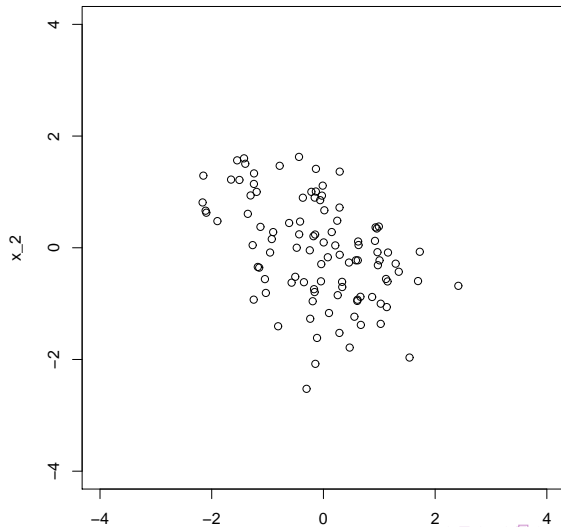
Caso $p=2$

$\rho = -0.2997$



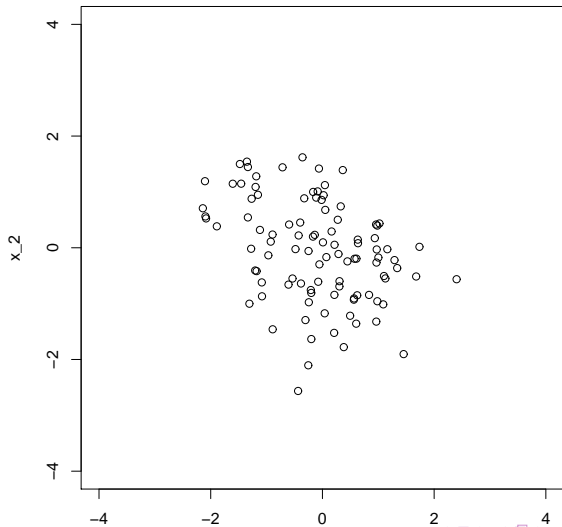
Caso $p=2$

$\rho = -0.1998$



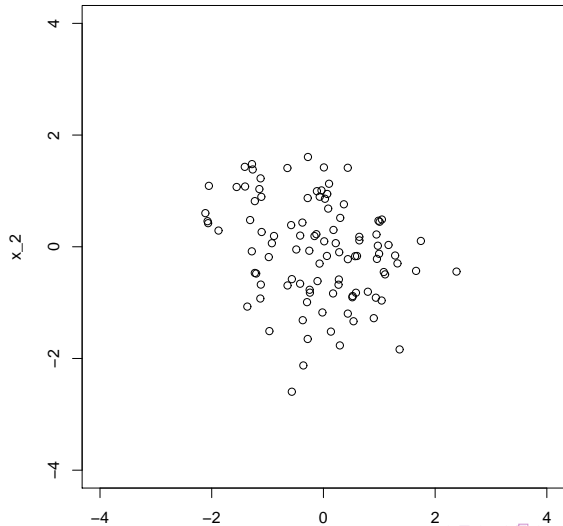
Caso $p=2$

$\rho = -0.0999$



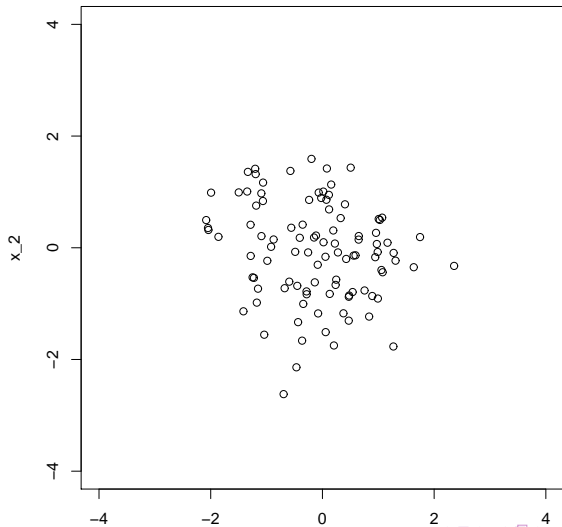
Caso $\rho=2$

$\rho = 0$



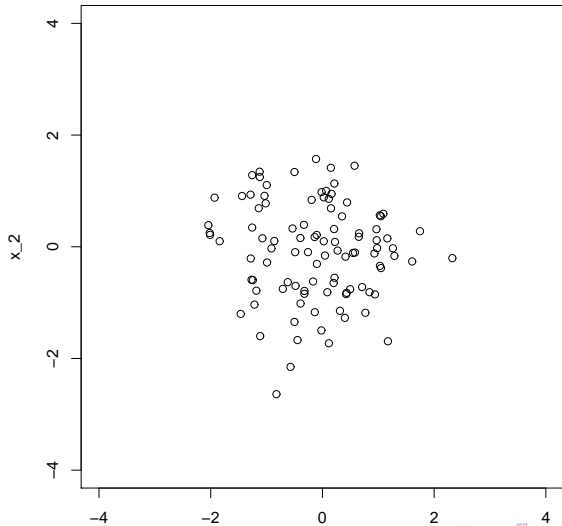
Caso $p=2$

$\rho = 0.0999$



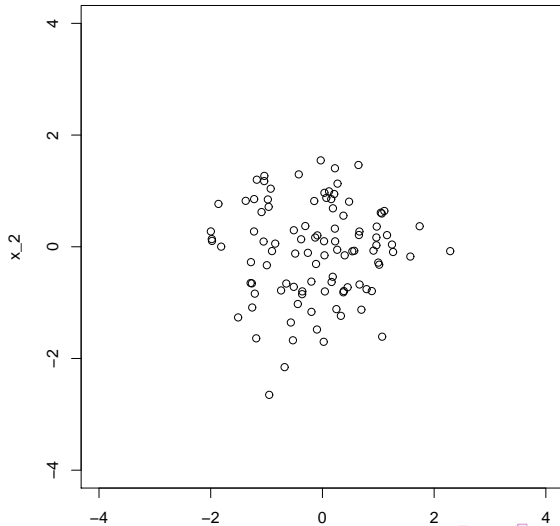
Caso $p=2$

$\rho = 0.1998$



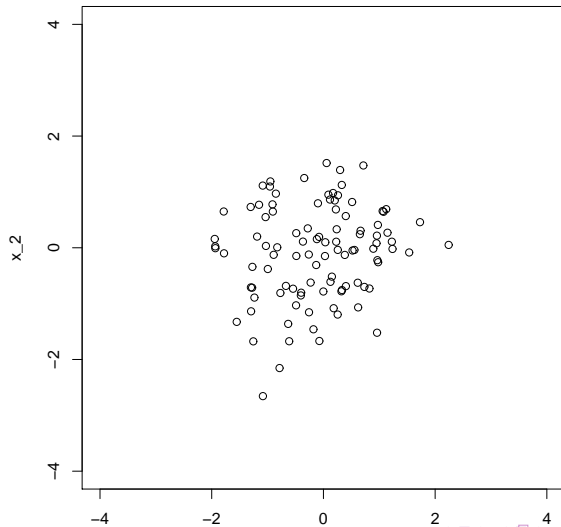
Caso $p=2$

$\rho = 0.2997$



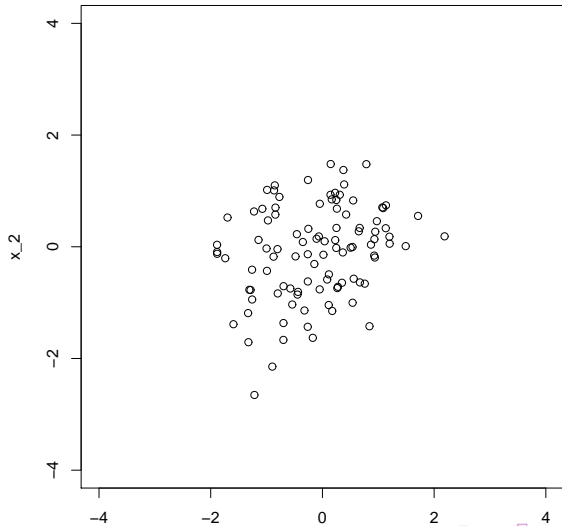
Caso $p=2$

$\rho = 0.3996$



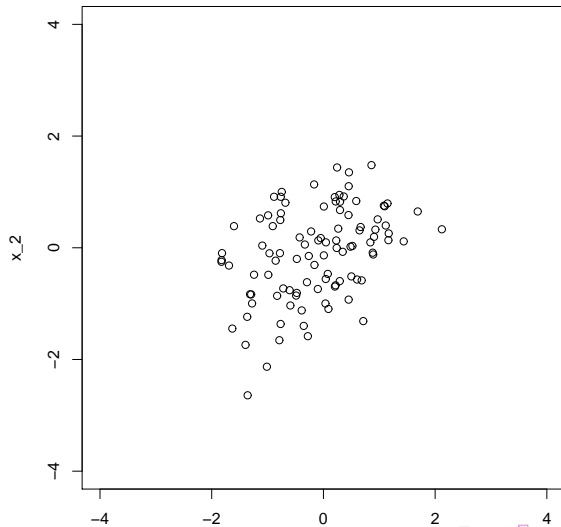
Caso $p=2$

$\rho = 0.4995$



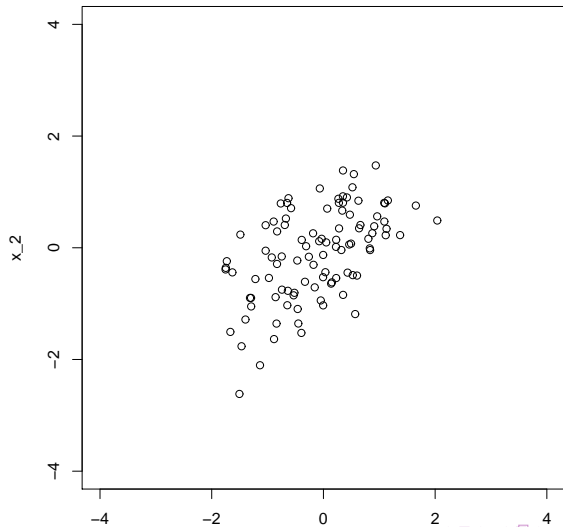
Caso $p=2$

$\rho = 0.5994$



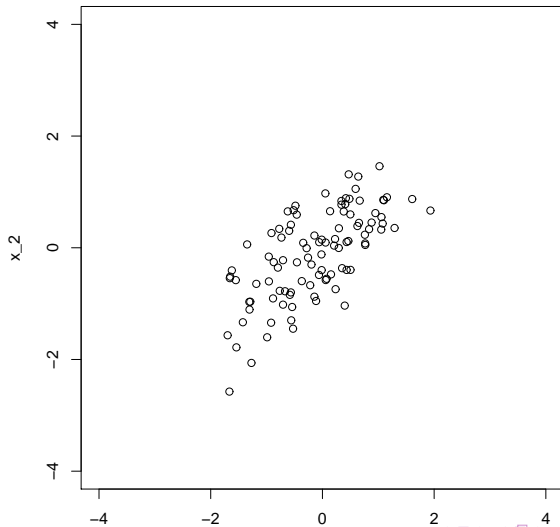
Caso $p=2$

$\rho = 0.6993$



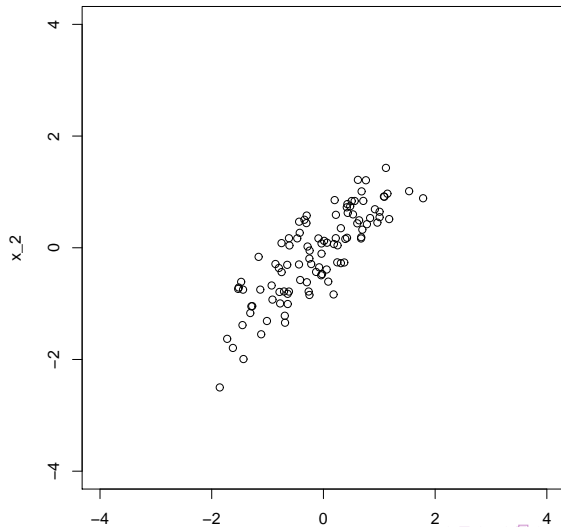
Caso $p=2$

$\rho = 0.7992$



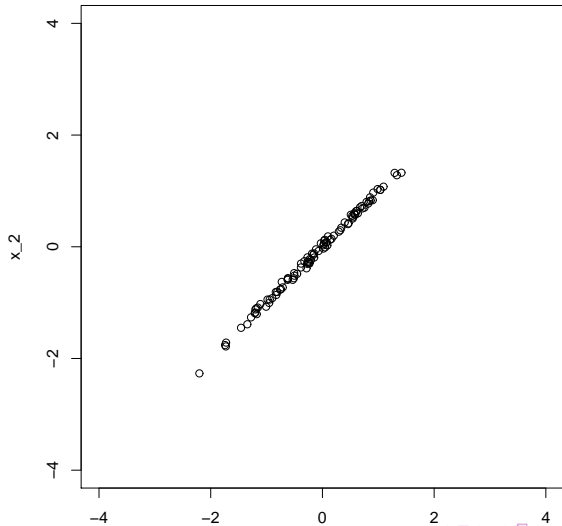
Caso $p=2$

$\rho = 0.8991$



Caso $p=2$

$\rho = 0.999$



Propiedades

- Si $\mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_p) \implies \varphi_{\mathbf{X}}(\mathbf{t}) = \exp\{-\frac{1}{2}\|\mathbf{t}\|^2\}$

Propiedades

- Si $\mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_p) \implies \varphi_{\mathbf{X}}(\mathbf{t}) = \exp\{-\frac{1}{2}\|\mathbf{t}\|^2\}$
- Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \varphi_{\mathbf{X}}(\mathbf{t}) = \exp\{i\mathbf{t}^T \boldsymbol{\mu}\} \exp\{-\frac{1}{2}\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}\}$

Propiedades

- Si $\mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_p) \implies \varphi_{\mathbf{X}}(\mathbf{t}) = \exp\{-\frac{1}{2}\|\mathbf{t}\|^2\}$
- Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \varphi_{\mathbf{X}}(\mathbf{t}) = \exp\{i\mathbf{t}^T \boldsymbol{\mu}\} \exp\{-\frac{1}{2}\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}\}$
- $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \iff \mathbf{a}^T \mathbf{X} \sim N(\mathbf{a}^T \boldsymbol{\mu}, \mathbf{a}^T \boldsymbol{\Sigma} \mathbf{a}), \forall \mathbf{a} \in \mathbb{R}^p$

Propiedades

- Si $\mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_p) \implies \varphi_{\mathbf{X}}(\mathbf{t}) = \exp\{-\frac{1}{2}\|\mathbf{t}\|^2\}$
- Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \varphi_{\mathbf{X}}(\mathbf{t}) = \exp\{i\mathbf{t}^T \boldsymbol{\mu}\} \exp\{-\frac{1}{2}\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}\}$
- $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \iff \mathbf{a}^T \mathbf{X} \sim N(\mathbf{a}^T \boldsymbol{\mu}, \mathbf{a}^T \boldsymbol{\Sigma} \mathbf{a}), \forall \mathbf{a} \in \mathbb{R}^p$
- $\mathbf{X} = (X_1, \dots, X_p)^T \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \mathbb{E}(\mathbf{X}) = \boldsymbol{\mu}$ y
 $\text{COV}(\mathbf{X}) = (\text{COV}(X_i, X_j))_{1 \leq i, j \leq p} = \boldsymbol{\Sigma}$

Propiedades

- Si $\mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_p) \implies \varphi_{\mathbf{X}}(\mathbf{t}) = \exp\{-\frac{1}{2}\|\mathbf{t}\|^2\}$
- Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \varphi_{\mathbf{X}}(\mathbf{t}) = \exp\{i\mathbf{t}^T \boldsymbol{\mu}\} \exp\{-\frac{1}{2}\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}\}$
- $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \iff \mathbf{a}^T \mathbf{X} \sim N(\mathbf{a}^T \boldsymbol{\mu}, \mathbf{a}^T \boldsymbol{\Sigma} \mathbf{a}), \forall \mathbf{a} \in \mathbb{R}^p$
- $\mathbf{X} = (X_1, \dots, X_p)^T \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \mathbb{E}(\mathbf{X}) = \boldsymbol{\mu}$ y
 $\text{COV}(\mathbf{X}) = (\text{COV}(X_i, X_j))_{1 \leq i, j \leq p} = \boldsymbol{\Sigma}$
- Sea $\mathbf{X} = (X_1, \dots, X_p)^T \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, entonces,
 X_1, \dots, X_p son independientes $\iff \boldsymbol{\Sigma}$ es diagonal.

Propiedades

- Sea $\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}$, $\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}$ y $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$. Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \mathbf{X}_1 \sim N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$ y $\mathbf{X}_2 \sim N(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$. Más aún, \mathbf{X}_1 y \mathbf{X}_2 son independientes $\iff \boldsymbol{\Sigma}_{21} = \mathbf{0}$.

Propiedades

- Sea $q \leq p$, $\mathbf{A} \in \mathbb{R}^{q \times p}$.

$$\text{Si } \mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \sim N(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T)$$

Propiedades

- Sea $q \leq p$, $\mathbf{A} \in \mathbb{R}^{q \times p}$.
Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \sim N(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T)$
- En particular, si $\mathbf{H} = (\mathbf{h}_1, \dots, \mathbf{h}_q) \in \mathbb{R}^{p \times q}$, es ortogonal incompleta, o sea, $\mathbf{H}^T \mathbf{H} = \mathbf{I}_q$, entonces si $\mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_p)$, $\mathbf{Y} = \mathbf{H}^T \mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_q)$.

Propiedades

- Sea $q \leq p$, $\mathbf{A} \in \mathbb{R}^{q \times p}$.
Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \sim N(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T)$
- En particular, si $\mathbf{H} = (\mathbf{h}_1, \dots, \mathbf{h}_q) \in \mathbb{R}^{p \times q}$, es ortogonal incompleta, o sea, $\mathbf{H}^T \mathbf{H} = \mathbf{I}_q$, entonces si $\mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_p)$, $\mathbf{Y} = \mathbf{H}^T \mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_q)$.
- Sea $\boldsymbol{\Sigma} = \mathbf{H}\boldsymbol{\Lambda}\mathbf{H}^T$, con \mathbf{H} ortogonal y $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_p)$, $\lambda_1 \geq \dots, \lambda_p$.
Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \mathbf{H}^T(\mathbf{X} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \boldsymbol{\Lambda})$.

Propiedades

- Sea $q \leq p$, $\mathbf{A} \in \mathbb{R}^{q \times p}$.
Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \sim N(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T)$
- En particular, si $\mathbf{H} = (\mathbf{h}_1, \dots, \mathbf{h}_q) \in \mathbb{R}^{p \times q}$, es ortogonal incompleta, o sea, $\mathbf{H}^T \mathbf{H} = \mathbf{I}_q$, entonces si $\mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_p)$, $\mathbf{Y} = \mathbf{H}^T \mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_q)$.
- Sea $\boldsymbol{\Sigma} = \mathbf{H}\boldsymbol{\Lambda}\mathbf{H}^T$, con \mathbf{H} ortogonal y $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_p)$, $\lambda_1 \geq \dots, \lambda_p$.
Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \mathbf{H}^T(\mathbf{X} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \boldsymbol{\Lambda})$.
- Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \iff \mathbf{X} = \mathbf{A}\mathbf{Z} + \boldsymbol{\mu}$ con $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_p)$ y $\mathbf{A}\mathbf{A}^T = \boldsymbol{\Sigma}$.

Propiedades

- Sea $q \leq p$, $\mathbf{A} \in \mathbb{R}^{q \times p}$.
Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \sim N(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T)$
- En particular, si $\mathbf{H} = (\mathbf{h}_1, \dots, \mathbf{h}_q) \in \mathbb{R}^{p \times q}$, es ortogonal incompleta, o sea, $\mathbf{H}^T \mathbf{H} = \mathbf{I}_q$, entonces si $\mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_p)$, $\mathbf{Y} = \mathbf{H}^T \mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_q)$.
- Sea $\boldsymbol{\Sigma} = \mathbf{H}\boldsymbol{\Lambda}\mathbf{H}^T$, con \mathbf{H} ortogonal y $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_p)$, $\lambda_1 \geq \dots, \lambda_p$.
Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \mathbf{H}^T(\mathbf{X} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \boldsymbol{\Lambda})$.
- Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \iff \mathbf{X} = \mathbf{A}\mathbf{Z} + \boldsymbol{\mu}$ con $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_p)$ y $\mathbf{A}\mathbf{A}^T = \boldsymbol{\Sigma}$.
- Si $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies (\mathbf{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_p^2$

Propiedades

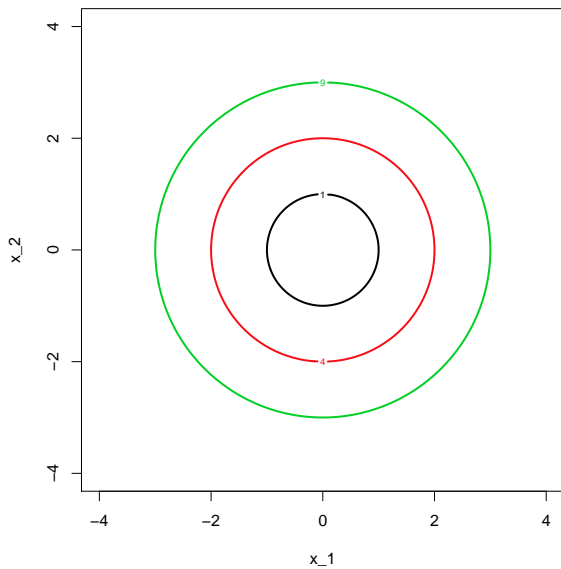
- Las regiones de densidad constante son los elipsoides

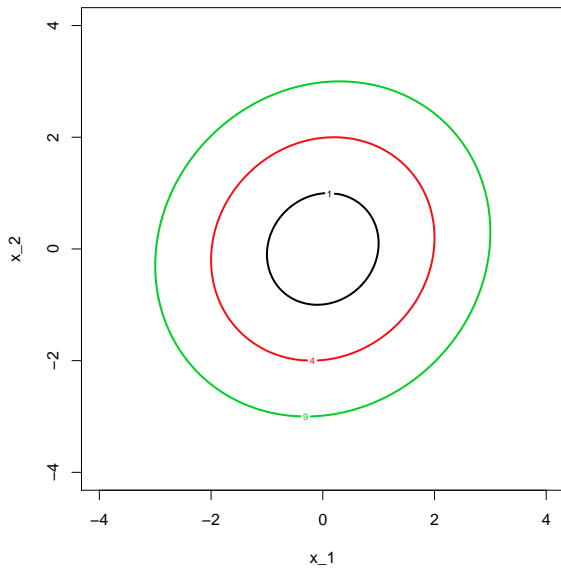
$$(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) = c^2$$

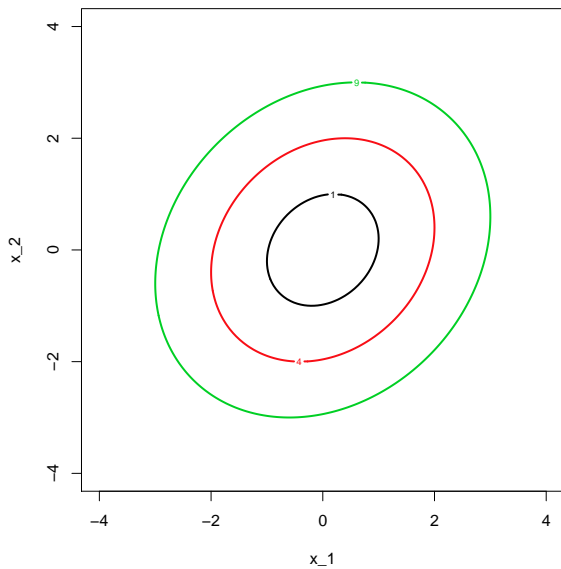
- Si $\boldsymbol{\Sigma} = \mathbf{H}\boldsymbol{\Lambda}\mathbf{H}^T$, con \mathbf{H} ortogonal y $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_p)$, definamos $\mathbf{y} = \mathbf{H}^T \mathbf{x}$ y $\boldsymbol{\nu} = \mathbf{H}^T \boldsymbol{\mu}$.

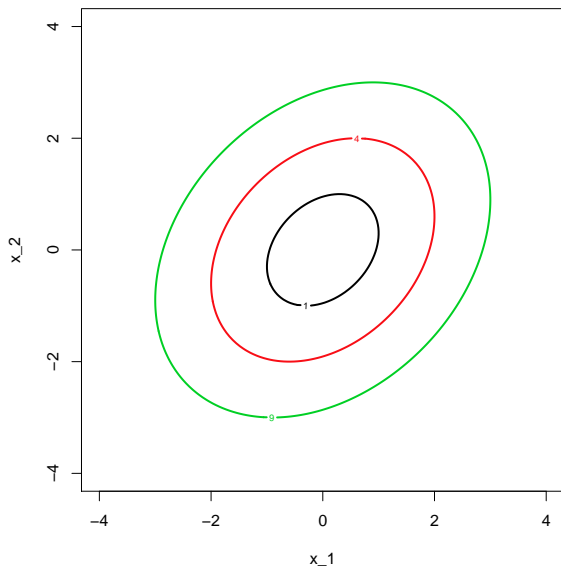
Los contornos de densidad constante son los elipsoides centrados en $\boldsymbol{\nu}$ con ejes principales de longitud $2c\lambda_j^{\frac{1}{2}}$ soportados en los autovectores, $\mathbf{h}_1, \dots, \mathbf{h}_p$, o sea,

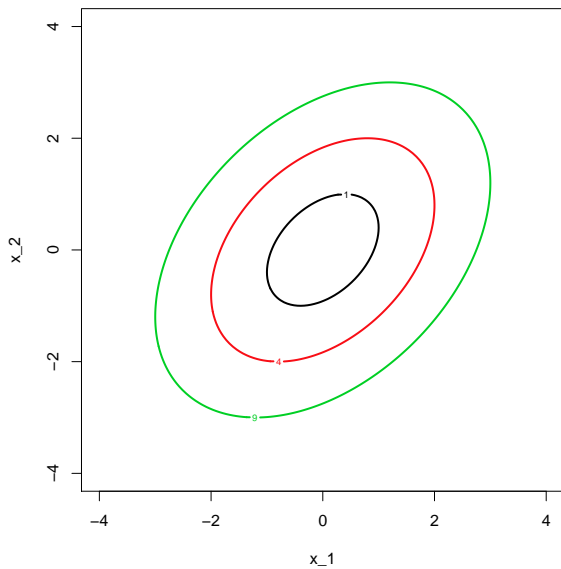
$$\sum_{j=1}^p \frac{(y_j - \nu_j)^2}{\lambda_j} = c^2$$

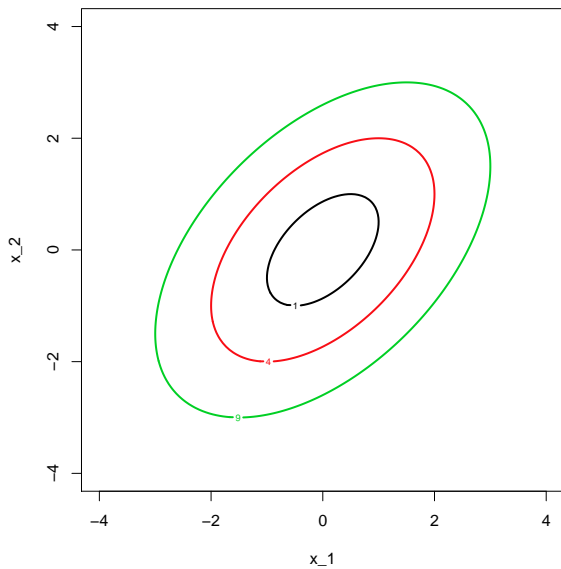
$\rho = 0$ 

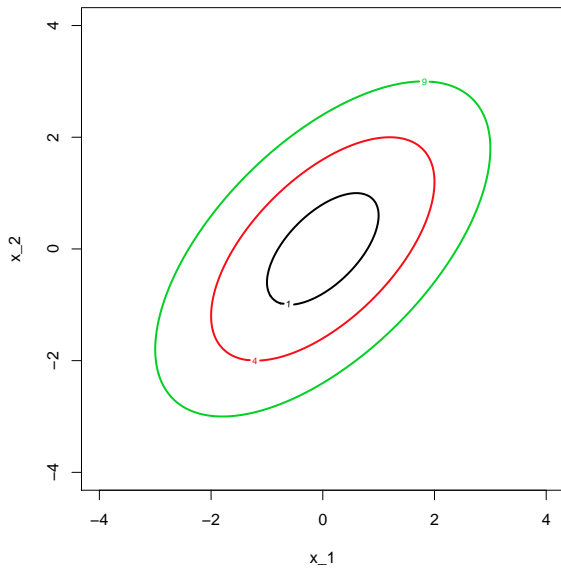
$\rho = 0.0999$ 

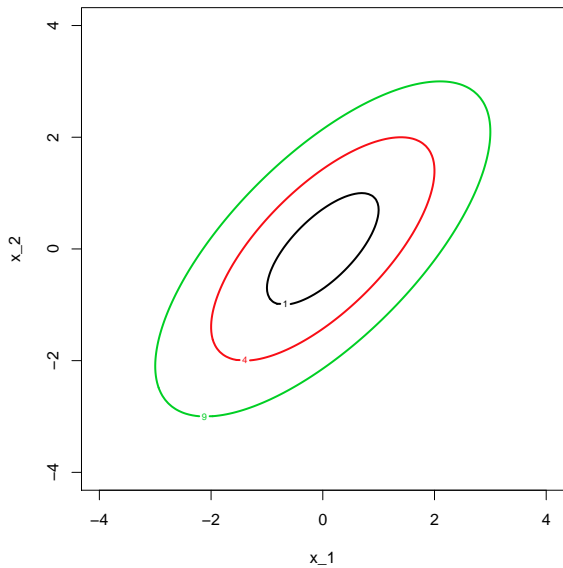
$\rho = 0.1998$ 

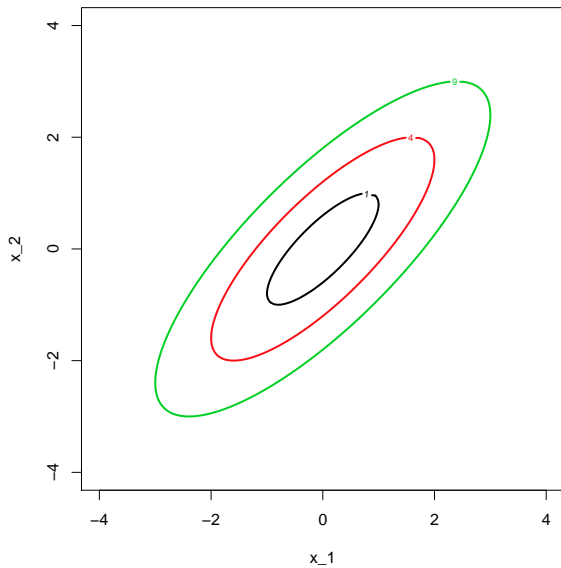
$\rho = 0.2997$ 

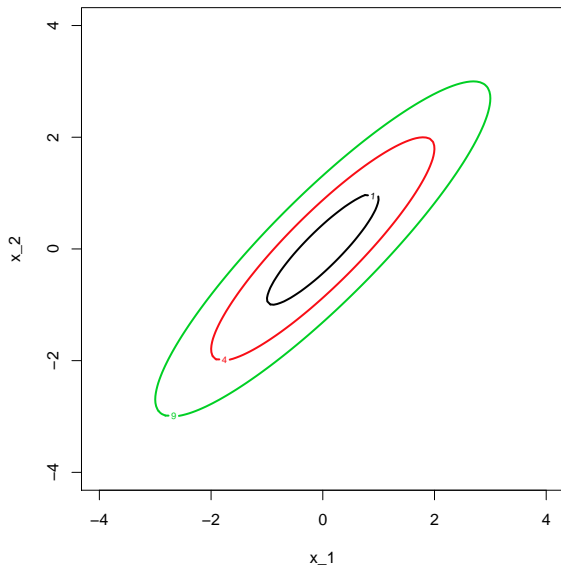
$\rho = 0.3996$ 

$\rho = 0.4995$ 

$\rho = 0.5994$ 

$\rho = 0.6993$ 

$\rho = 0.7992$ 

$\rho = 0.8991$ 

$\rho = 0.999$ 