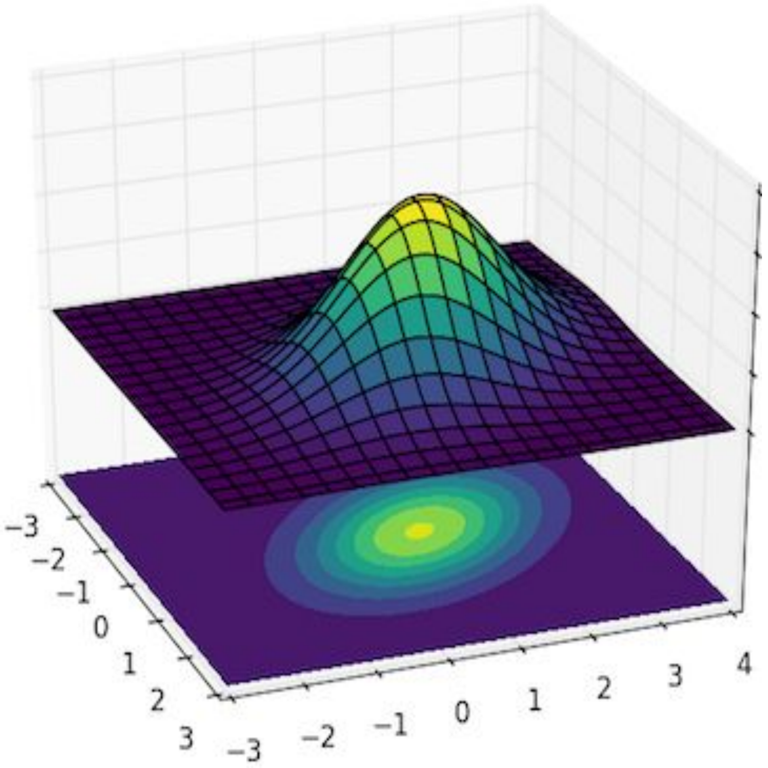
Procesos estocásticos (86.09)

- Vectores Aleatorios Gaussianos
- Transformación afín

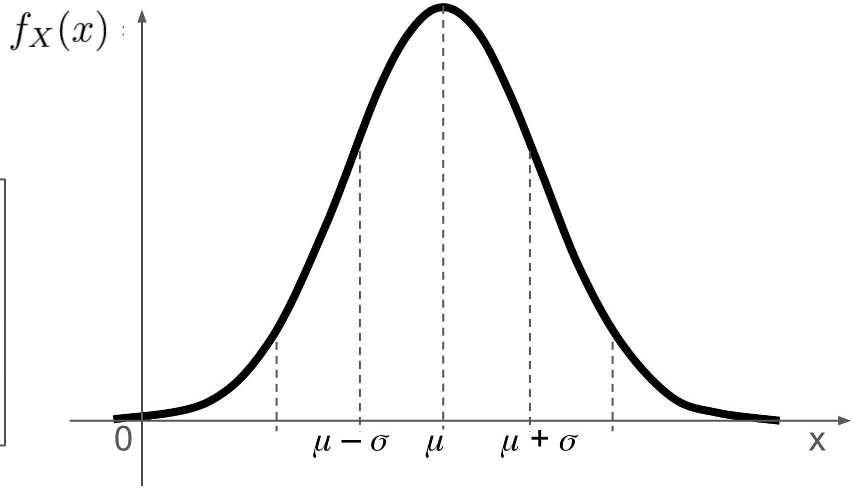


Vectores aleatorios gaussianos

Variable aleatoria con **distribución normal**

$$X \sim N(\mu, \sigma^2)$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$



Vectores aleatorios gaussianos

Vector aleatorio **gaussiano**

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \quad \begin{array}{l} \mathcal{N}(\mu_1, \sigma_1^2) \\ \mathcal{N}(\mu_2, \sigma_2^2) \\ \\ \mathcal{N}(\mu_n, \sigma_n^2) \end{array}$$

$$\boxed{\mathbf{X} \sim \mathcal{N}(\mu_{\mathbf{X}}, C_{\mathbf{X}})}$$

Cada una de sus componentes es una VA con distribución normal, cada una con su media y varianza. No necesariamente independientes.

Vectores aleatorios gaussianos

Función de densidad conjunta:

$$\boxed{f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}\,|C_X|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_X)^T C_X^{-1}(\mathbf{x}-\boldsymbol{\mu}_X)\right)}$$

Donde:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \qquad \boldsymbol{\mu}_X = \begin{bmatrix} \mu_{X_1} \\ \mu_{X_2} \\ \vdots \\ \mu_{X_N} \end{bmatrix} \qquad C_X = \begin{bmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, X_2) & \cdots & \mathrm{Cov}(X_1, X_N) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Var}(X_2) & \cdots & \mathrm{Cov}(X_2, X_N) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_N, X_1) & \mathrm{Cov}(X_N, X_2) & \cdots & \mathrm{Var}(X_N) \end{bmatrix}$$

Vectores aleatorios gaussianos (N=2)

Caso particular $\mathbf{x} = [x\ y]^T \in \mathbb{R}^2$

Sin perder generalidad, veamos el caso **normal bivariado** de media nula y covarianza C_X.

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{2\pi\,|C_X|^{1/2}} \exp\left(-\frac{1}{2}\mathbf{x}^T C_X^{-1}\mathbf{x}\right)$$

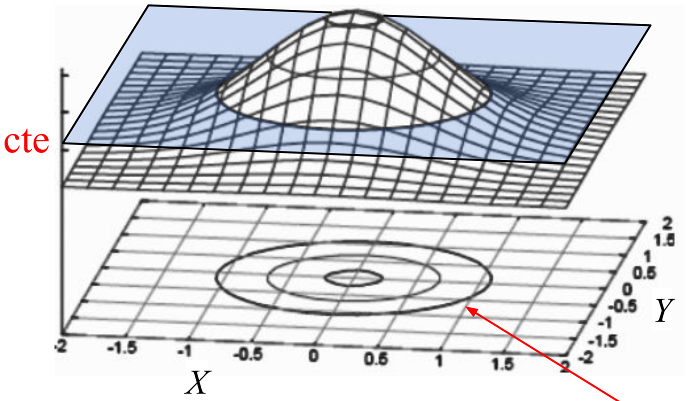
Podemos expresar la matriz de covarianza $C_X \in \mathbb{R}^{2x2}$ y su inversa C_X^{-1} en términos de las varianzas y coeficiente de correlación:

$$C_X = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix} \qquad C_X^{-1} = \frac{1}{\sigma_X^2\sigma_Y^2(1-\rho^2)} \begin{bmatrix} \sigma_Y^2 & -\rho\sigma_X\sigma_Y \\ -\rho\sigma_X\sigma_Y & \sigma_X^2 \end{bmatrix}$$

Veamos qué pasa con el exponente: $\mathbf{x}^T C_X^{-1}\mathbf{x} = \frac{1}{1-\rho^2}\left(\frac{x^2}{\sigma_X^2} - \frac{2\rho\,xy}{\sigma_X\sigma_Y} + \frac{y^2}{\sigma_Y^2}\right)$

Vectores aleatorios gaussianos (N=2)

Caso particular $\mathbf{x} = [x\ y]^T \in \mathbb{R}^2$



Elipses

Evaluando los puntos de la superficie que tienen una $f_X(\mathbf{x})$ = cte, obtenemos las curvas de nivel (Elipses para la normal bivariada)

$$\frac{1}{2\pi\,|C_X|^{1/2}}\exp\left(-\frac{1}{2}\mathbf{x}^T C_X^{-1}\mathbf{x}\right) = \text{cte}$$

$$\exp\left(-\frac{1}{2}\mathbf{x}^T C_X^{-1}\mathbf{x}\right) = \text{cte}\,2\pi|C_X|^{1/2}$$

$$\mathbf{x}^T C_X^{-1}\mathbf{x} = -2\ln\left(\text{cte}\,2\pi|C_X|^{1/2}\right) = \text{cte'}$$

$$\frac{x^2}{\sigma_X^2} - \frac{2\rho\,xy}{\sigma_X\sigma_Y} + \frac{y^2}{\sigma_Y^2} = \text{cte'}(1-\rho^2) = \text{cte''}$$

Curvas de nivel – Gaussiana multivariada

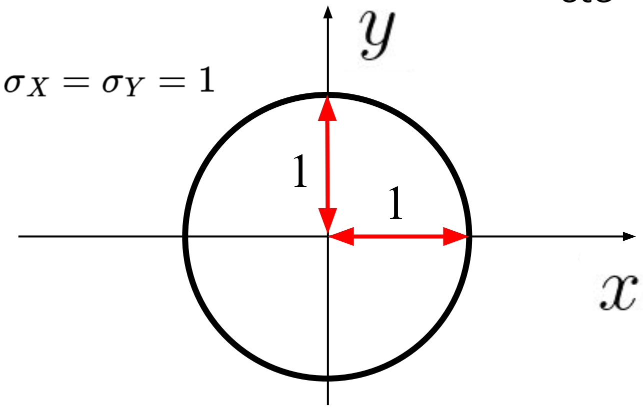
Vector descorrelacionado y normalizado

$$C_X = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$x^2 + y^2 = 1$$

Suponer *cte* = 1

$\sigma_X = \sigma_Y = 1$



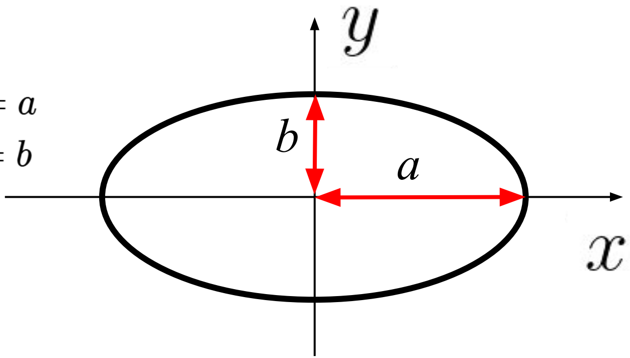
Vector descorrelacionado

$$C_X = \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_Y^2 \end{bmatrix}$$

$$\frac{x^2}{\sigma_X^2} + \frac{y^2}{\sigma_Y^2} = 1$$

$\sigma_X = a$

$\sigma_Y = b$



Curvas de nivel – Gaussiana multivariada

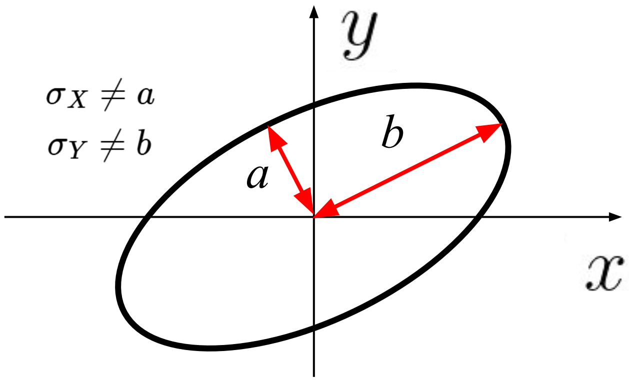
Vector correlacionado

$$C_X = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$$

$$\frac{x^2}{\sigma_X^2} - \frac{2\rho\, xy}{\sigma_X\sigma_Y} + \frac{y^2}{\sigma_Y^2} = 1$$

$\sigma_X \neq a$

$\sigma_Y \neq b$



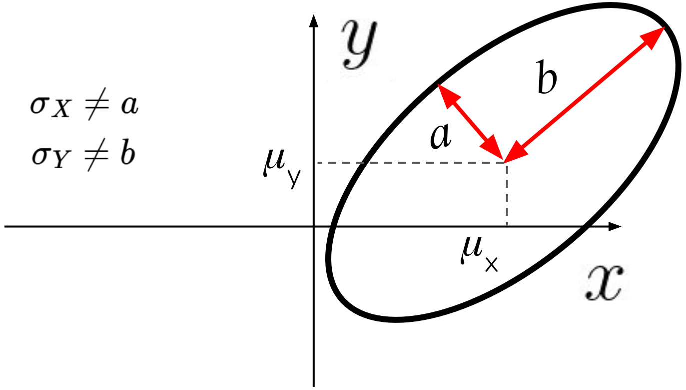
Vector correlacionado y desplazado

$$C_X = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix} \qquad \boldsymbol{\mu} = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}$$

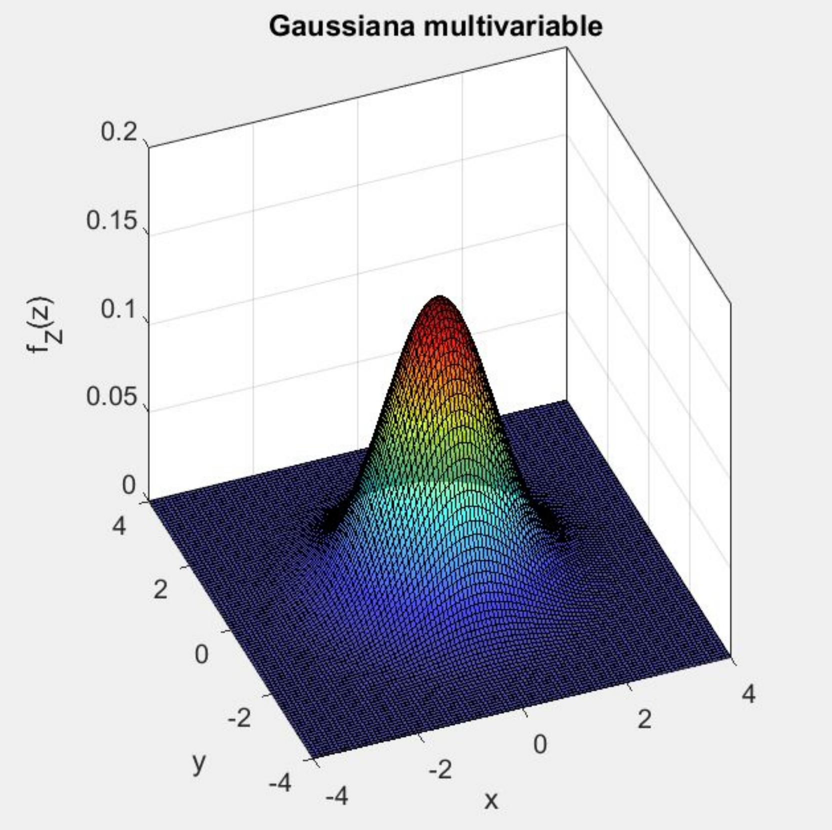
$$\frac{(x-\mu_x)^2}{\sigma_X^2} - \frac{2\rho\,(x-\mu_x)(y-\mu_y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_y)^2}{\sigma_Y^2} = 1$$

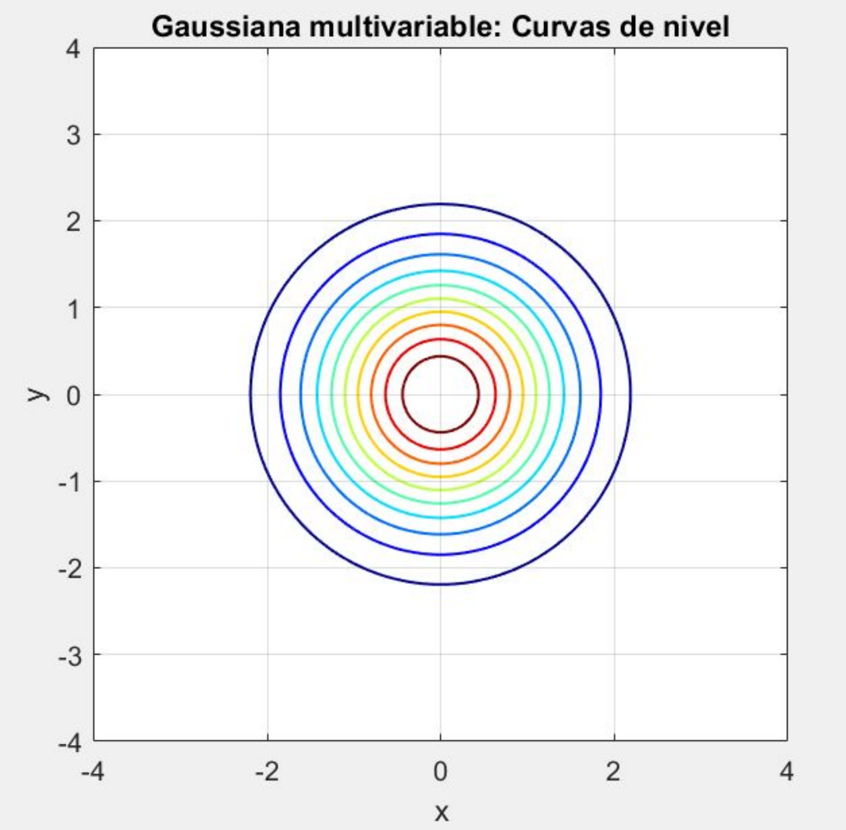
$\sigma_X \neq a$

$\sigma_Y \neq b$

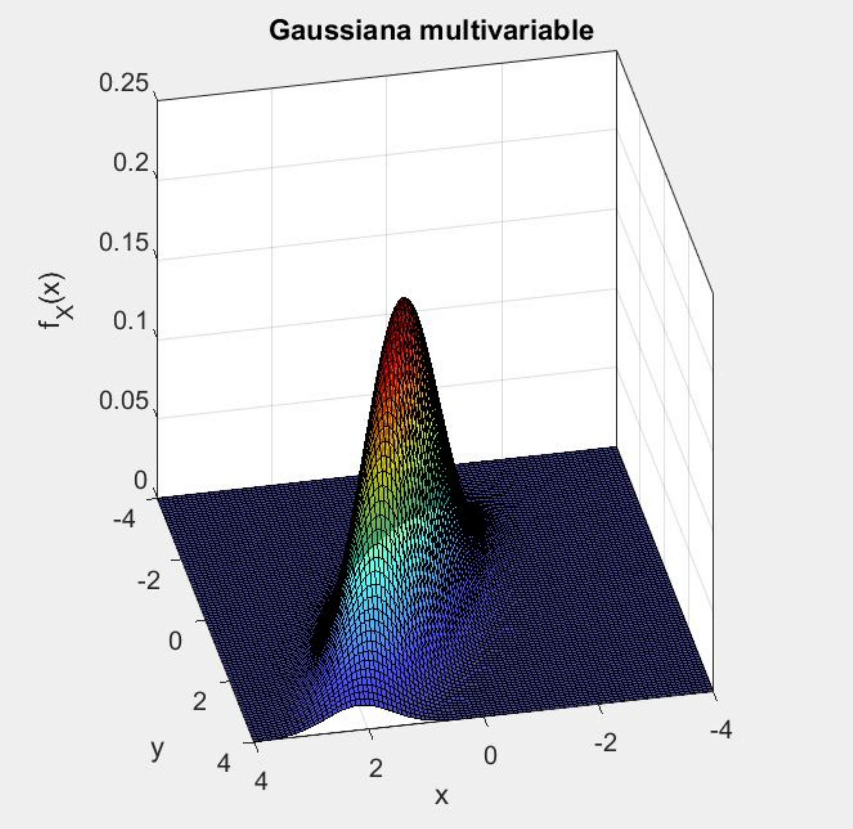


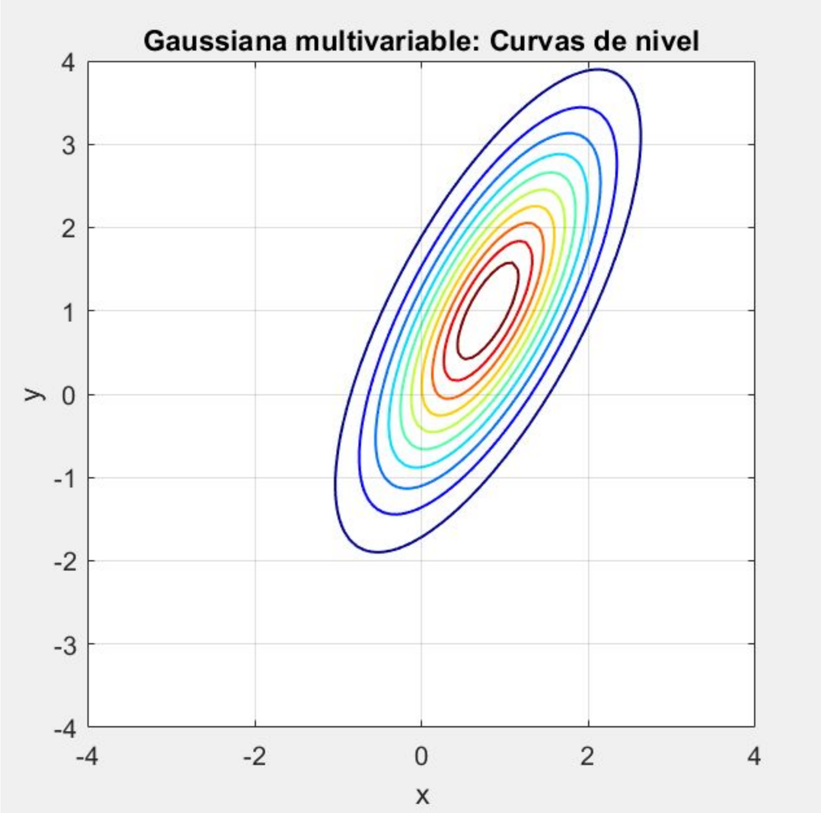
Vectores aleatorios gaussianos (N=2)





Vectores aleatorios gaussianos (N=2)





Vector aleatorio normal estándar

Vector aleatorio

$$\mathbf{Z} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_N \end{bmatrix} \in \mathbb{R}^N$$

VeA normal estándar

$$\mathbf{Z} \sim N(\mathbf{0}, I)$$

$$\mathbf{0} \in \mathbb{R}^N$$

$$I \in \mathbb{R}^{N\times N}$$

VAs normales estándar

$$Z_i \sim N(0,1)\ ;\ i = 1, \ldots, N$$

$$\mathrm{Cov}(Z_i Z_j) = 0 \quad ; \forall\, i \neq j$$

Función de densidad normal estándar multivariada

$$\boxed{f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{(2\pi)^{N/2}} \exp\left(-\frac{1}{2}\mathbf{z}^T\mathbf{z}\right)}$$

Vector aleatorio normal estándar

$$\mu_{\mathbf{Z}} = \mathbb{E}\left[\mathbf{Z}\right] = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$C_Z = \begin{bmatrix} \mathrm{Var}(Z_1) & \mathrm{Cov}(Z_1, Z_2) & \cdots & \mathrm{Cov}(Z_1, Z_N) \\ \mathrm{Cov}(Z_2, Z_1) & \mathrm{Var}(Z_2) & \cdots & \mathrm{Cov}(Z_2, Z_N) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(Z_N, Z_1) & \mathrm{Cov}(Z_N, Z_2) & \cdots & \mathrm{Var}(Z_N) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Transformación Afín

Transformación Afín

Una transformación afín es un caso más general que una transformación lineal, ya que incluye también una traslación de la media en un vector constante **b**.

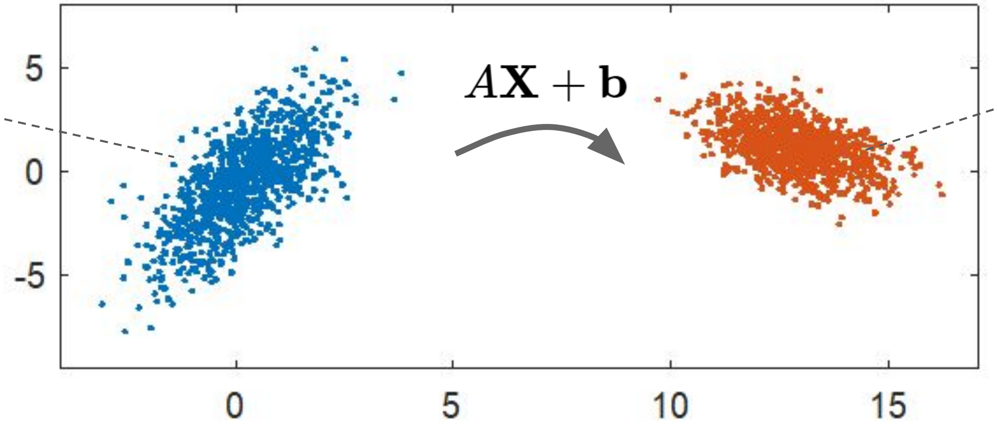
$$\mathbf{Y} = A\mathbf{X} + \mathbf{b}$$

Constantes determinísticas: $A \in \mathbb{R}^{M \times N}$ $\quad \mathbf{b} \in \mathbb{R}^{M}$

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

$$\mathbb{E}[\mathbf{X}] \in \mathbb{R}^{N}$$

$$C_X \in \mathbb{R}^{N \times N}$$



$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

$$\mathbb{E}[\mathbf{Y}] \in \mathbb{R}^{M}$$

$$C_Y \in \mathbb{R}^{M \times M}$$

Transformación Afín

Resultado 1

Sea un VeA $\mathbf{X} \in \mathbb{R}^{\mathrm{N}}$ de media $\boldsymbol{\mu}_{\mathrm{X}} \in \mathbb{R}^{\mathrm{N}}$ y covarianza C_X (conocidas) y una transformación afín $\mathbf{Y} = \mathrm{A}\mathbf{X}+\mathbf{b}$. La media $\boldsymbol{\mu}_{\mathrm{Y}}$ y covarianza C_Y resultan:

$$\boxed{\boldsymbol{\mu}_Y = A\boldsymbol{\mu}_X + \mathbf{b}}$$

$$\boxed{C_Y = AC_XA^T}$$

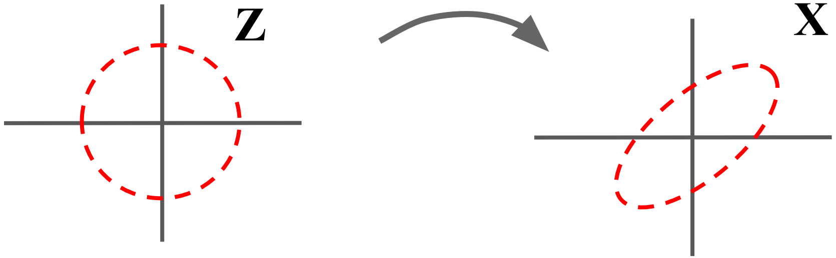
Transformación lineal “coloreado”

Resultado 2

A partir de un vector descorrelacionado y normalizado se puede generar otro con una matriz de covarianza arbitraria C_X.

Interpretación geométrica

$$\mathbf{X} = \mathbf{A}\mathbf{Z}$$



$$\boxed{A = P_X \Lambda_X^{1/2}}$$

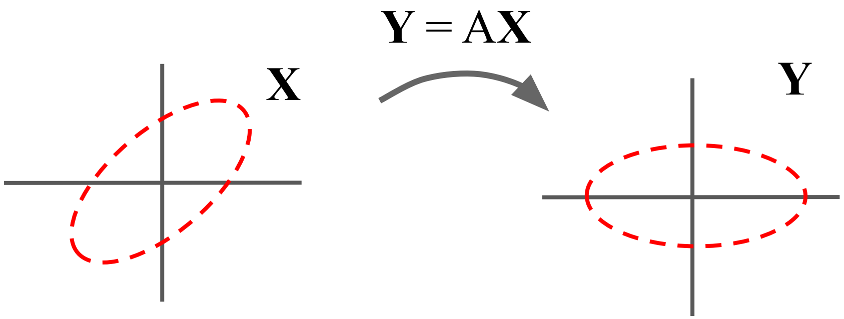
Se suele denominar “coloreado” el proceso de correlacionar un vector descorrelacionado.

Transformación lineal “blanqueo”

Resultado 3

Proyectamos el vector correlacionado $\mathbf{X}$ en el espacio de direcciones principales (autovectores) para generar el vector $\mathbf{Y}$ (descorrelacionado).

Interpretación geométrica



$$\boxed{A = P_X^T}$$

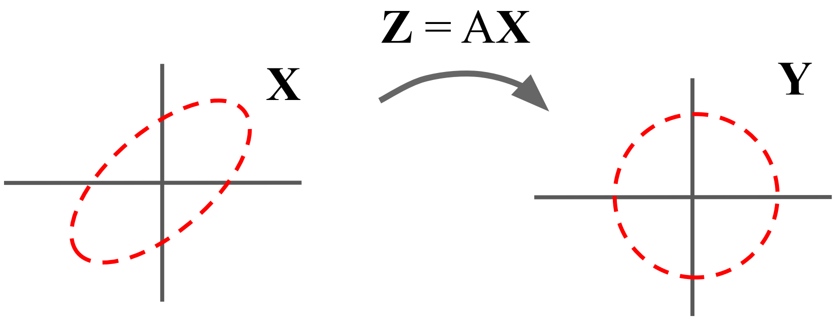
Se suele denominar “blanqueo” el proceso de descorrelacionar un vector aleatorio..

Transformación lineal “blanqueo y normalización”

Resultado 4

Si además de descorrelacionar un vector $\mathbf{X}$, necesitamos normalizar para que todas sus componentes tengan varianza unitaria, la transformación resulta:

Interpretación geométrica



$$\boxed{A = \Lambda_X^{-1/2} P_X^T}$$

blanqueamos y normalizamos

Resumen

$\boldsymbol{\mu}_{\mathrm{X}}$ $\quad \mathrm{C_X} = P_X \Lambda_{\mathrm{X}} P_X^{\mathrm{T}}$

$$\mathbf{X}$$

Descorrelación, centrado y normalización

$$\mathrm{A} = \Lambda_{\mathrm{X}}^{-1/2} P_X^{\mathrm{T}}$$

$$\mathbf{b} = -\mathrm{A}\boldsymbol{\mu}_{\mathrm{X}}$$

$\boldsymbol{\mu}_{\mathrm{Z}} = \mathbf{0}$ $\quad \mathrm{C_Z} = I$

$$\mathbf{Z}$$

Coloreado y traslación

$$\mathrm{A} = P_Y \Lambda_{\mathrm{Y}}^{1/2}$$

$$\mathbf{b} = \boldsymbol{\mu}_{\mathrm{Y}}$$

$\boldsymbol{\mu}_{\mathrm{Y}}$ $\quad \mathrm{C_Y} = P_Y \Lambda_{\mathrm{Y}} P_Y^{\mathrm{T}}$

$$\mathbf{Y}$$

Actividad 1

Se quiere utilizar una transformación afín $\mathbf{Y} = A\,\mathbf{Z} + \mathbf{b}$ que permita convertir un vector aleatorio con parámetros C_Z y $\boldsymbol{\mu}_Z$=0 en otro vector con parámetros C_Y y $\boldsymbol{\mu}_Y$.

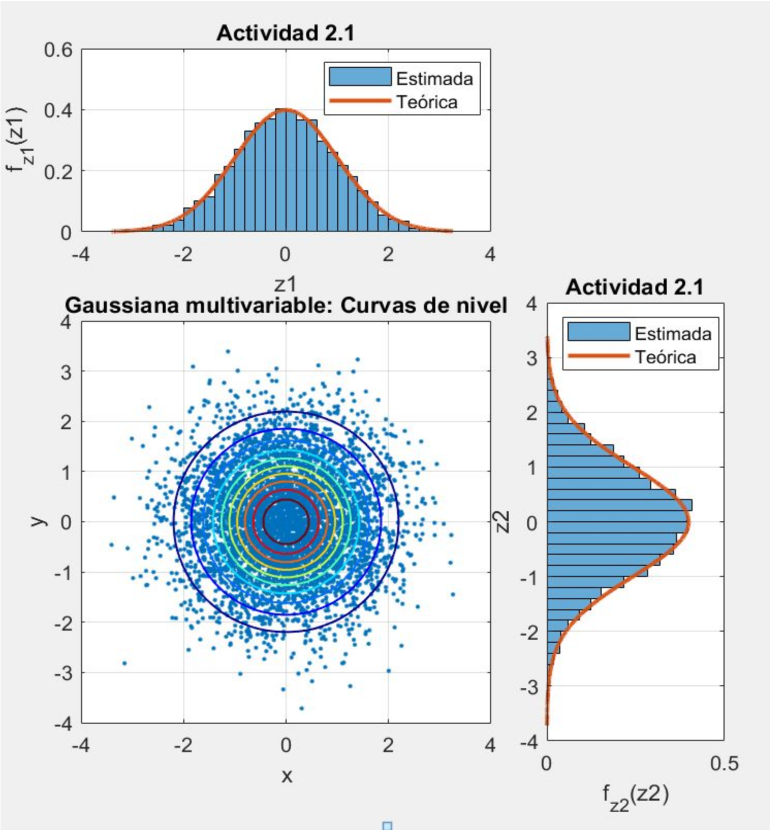
1. Genere un vector normal estándar $\mathbf{Z} = [Z_1\ Z_2]^T$ de 5000 realizaciones. Grafique el histograma de cada componente y el gráfico de dispersión de **Z**.

2. Encuentre los parámetros de la transformación *A* y **b** para que cumpla lo pedido.

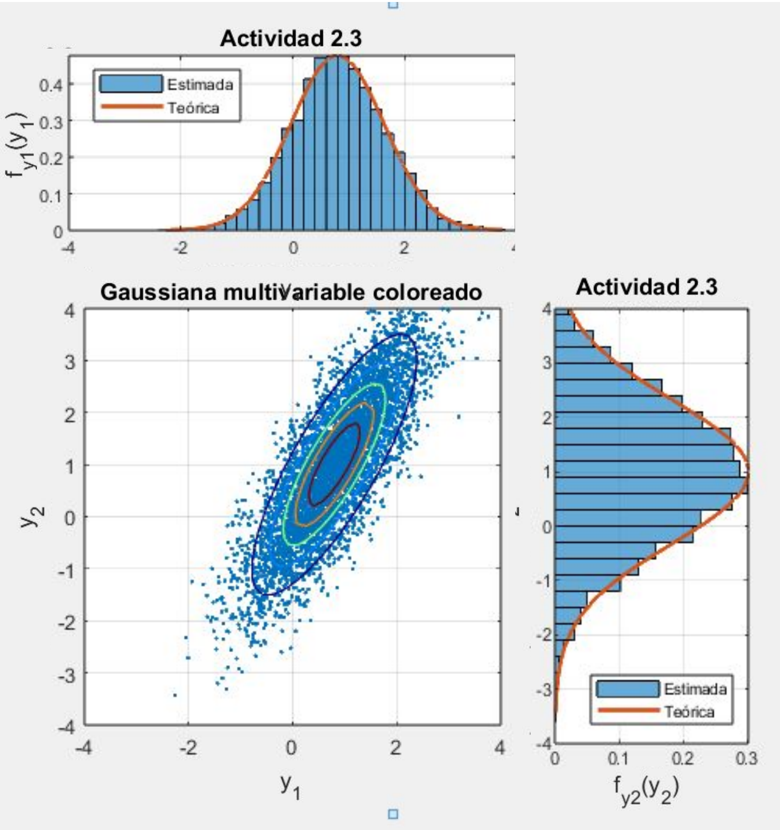
3. Con la transformación definida, genere 5000 realizaciones de la variable **Y** transformando las muestras del vector **Z**. Para cada componente del vector **Y** generado grafique el histograma.. También haga un gráfico de dispersión.

$$C_X = \begin{bmatrix} 0.70 & 0.90 \\ 0.90 & 1.75 \end{bmatrix} \qquad \boldsymbol{\mu}_X = \begin{bmatrix} 0.80 \\ 1.90 \end{bmatrix}$$

Actividad 1



Actividad 2.1
Estimada
Teórica
$f_{z1}(z1)$
z1
Gaussiana multivariable: Curvas de nivel
x
y
Actividad 2.1
z2
$f_{z2}(z2)$



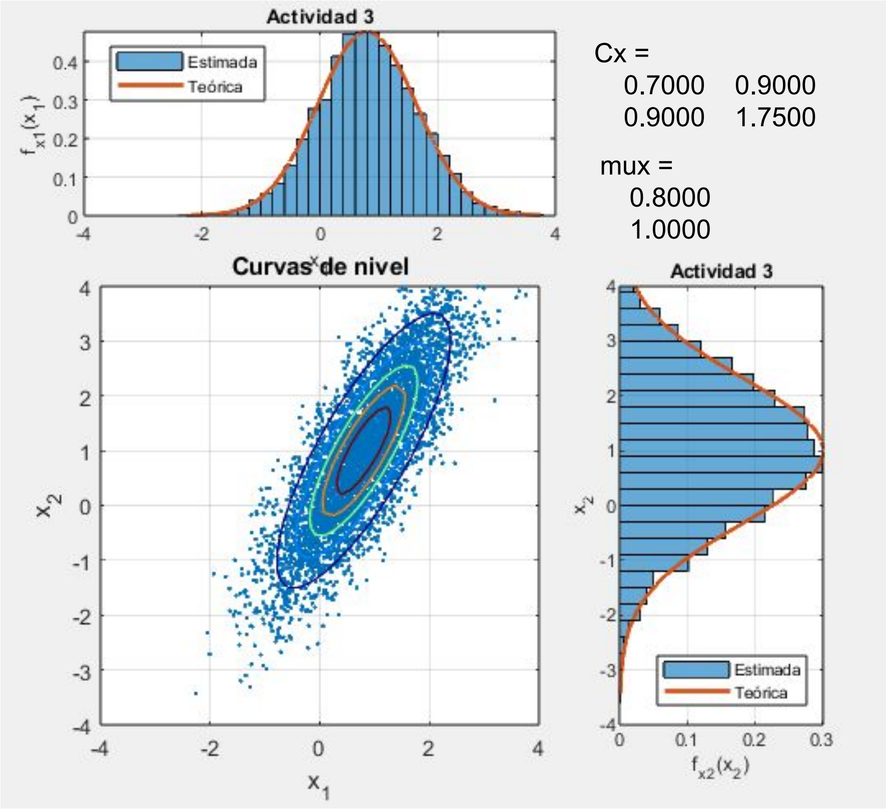
Actividad 2.3
Estimada
Teórica
$f_{y1}(y_1)$
Gaussiana multivariable coloreado
y_1
y_2
Actividad 2.3
$f_{y2}(y_2)$

Actividad 2

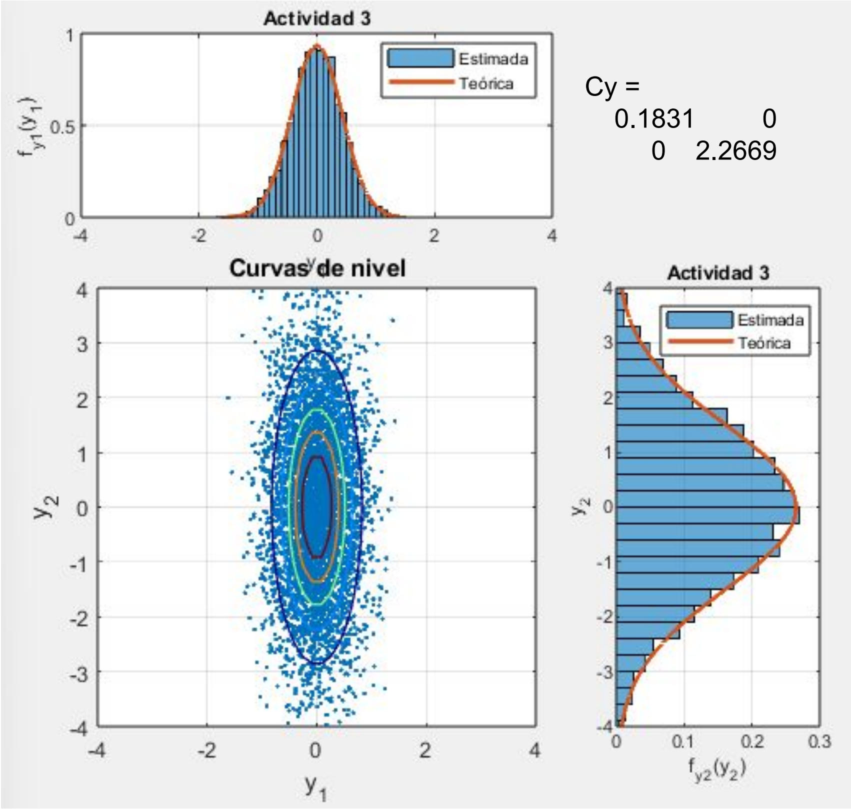
Genere un vector aleatorio normal de media $\boldsymbol{\mu}_X$ y covarianza C_X. Luego aplique una transformación para generar un nuevo vector **Y** descorrelacionado (si normalizar) y de media nula. Haga los gráficos de dispersión de ambos vectores (**X** e **Y**) y sus histogramas de cada componente. También grafique la superficie de la función de densidad teórica para **Y**.

$$C_X = \begin{bmatrix} 0.70 & 0.90 \\ 0.90 & 1.75 \end{bmatrix} \qquad \boldsymbol{\mu}_X = \begin{bmatrix} 0.80 \\ 1.90 \end{bmatrix}$$

Actividad 2



Actividad 3
Estimada
Teórica
$f_{x1}(x_1)$
x
Curvas de nivel
x_2
x_1
Actividad 3
x_2
Estimada
Teórica
$f_{x2}(x_2)$
Cx =
0.7000 0.9000
0.9000 1.7500
mux =
0.8000
1.0000



Actividad 3
Estimada
Teórica
$f_{y1}(y_1)$
y
Curvas de nivel
y_2
y_1
Actividad 3
y_2
Estimada
Teórica
$f_{y2}(y_2)$
Cy =
0.1831 0
0 2.2669

¿Preguntas?

Transformación Afín

Resultado 1

Calculamos la media:

$$\mathbf{Y} = A\,\mathbf{X} + \mathbf{b} \quad \Longrightarrow \quad \mathbb{E}[\mathbf{Y}] = A\,\mathbb{E}[\mathbf{X}] + \mathbf{b} \quad \Longrightarrow \quad \boxed{\boldsymbol{\mu}_Y = A\boldsymbol{\mu}_X + \mathbf{b}}$$

Calculamos la covarianza:

$$\mathbf{Y} - \mathbb{E}[\mathbf{Y}] = A\,(\mathbf{X} - \mathbb{E}[\mathbf{X}])$$

$$C_{\mathbf{Y}} = \mathbb{E}\left[(\mathbf{Y} - \mathbb{E}[\mathbf{Y}])(\mathbf{Y} - \mathbb{E}[\mathbf{Y}])^T\right] = \mathbb{E}\left[A\,(\mathbf{X} - \mathbb{E}[\mathbf{X}])\{A(\mathbf{X} - \mathbb{E}[\mathbf{X}])\}^T\right] =$$
$$= \mathbb{E}\left[A\,(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^T A^T\right] = A\,\mathbb{E}\left[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^T\right] A^T = A\,C_{\mathbf{X}} A^T$$

$$\boxed{C_Y = A C_X A^T}$$

Transformación lineal “coloreado”

Resultado 2

$$C_X = \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^T]$$

De la transformación $\mathbf{X} = \mathrm{A}\mathbf{Z}$ sabemos que C_X resulta:

$$C_X = \mathbb{E}[(A\mathbf{Z})(A\mathbf{Z})^T] = AC_Z A^T = AI A^T$$

Podemos hacer la descomposición espectral de C_X y ver que se puede expresar como el producto de una matriz A por su traspuesta (AA^{T}):

$$C_X = P_X \Lambda_X P_X^T = P_X \Lambda_X^{1/2} \Lambda_X^{1/2} P_X^T = P_X \Lambda_X^{1/2} (\Lambda_X^{1/2})^T P_X^T = \underbrace{(P_X \Lambda_X^{1/2})}_{A} (P_X \Lambda_X^{1/2})^T = AA^T$$

Transformación lineal "blanqueo"

Resultado 3

Sabemos de resultados anteriores que:

$$C_Y = AC_XA^T = AP_X\Lambda_XP_X^TA^T$$

Viendo la diagonalización de la ecuación anterior, si elegimos que $AP_X = I$, entonces $A = P_X^T$ y C_Y será igual a una matriz diagonal.

$$C_Y = \underbrace{P_X^TP_X}_{I}\Lambda_X\underbrace{P_X^TP_X}_{I} = \Lambda_X$$

Esto implica que la matriz A que descorrelaciona el vector $\mathbf{X}$ resulta P_X^T.

Transformación lineal “blanqueo” y normalización

Resultado 4

Buscamos que $C_Z = I$:

$$C_Z = AC_XA^T = AP_X\Lambda_X P_X^T A^T = AP_X\Lambda_X^{1/2}(\Lambda_X^{1/2})^T P_X^T A^T = \underbrace{AP_X\Lambda_X^{1/2}}_{I}(AP_X\Lambda_X^{1/2})^T = I$$

Imponemos esta condición

Entonces:

$$AP_X\Lambda_X^{1/2} = I$$

$$AP_X\Lambda_X^{1/2}\Lambda_X^{-1/2} = \Lambda_X^{-1/2} \quad \rightarrow \quad AP_X = \Lambda_X^{-1/2}$$

$$AP_XP_X^T = \Lambda_X^{-1/2}P_X^T \quad \rightarrow \quad \boxed{A = \Lambda_X^{-1/2}P_X^T}$$