Distribuciones Multivariadas

Alan Reyes-Figueroa

Métodos Numéricos II (Aula 05a) 18.julio.2023

Distribuciones multivariadas

Sea $X = (X_1, X_2, \ldots, X_d) \in \mathbb{R}^d$ un vector aleatorio (esto es, cada componente X_i es una variable aleatoria $X_i : \Omega \to \mathbb{R}$).

Definición

*Definimos el **valor esperado** de X como el vector $\mu \in \mathbb{R}^d$ dado por*

$$\mathbb{E}(X) = \mu = (\mu_1, \mu_2, \ldots, \mu_d)^T \in \mathbb{R}^d,$$

donde $\mu_i = \mathbb{E}(X_i)$, para $i = 1, 2, \ldots, d$.

Definición

*Definimos la **varianza** de X como la matriz $\Sigma \in \mathbb{R}^{d \times d}$ dada por*

$$\mathrm{Cov}(X) = \Sigma = \big(\mathrm{Cov}(X_i, X_j)\big)_{i,j}.$$

*La entrada (i, j) de esta matriz corresponde a la covarianza de las variables X_i y X_j. A Σ también se le conoce como la **matriz de covarianza** de X.*

Distribuciones multivariadas

Para cualquier vector aleatorio $X \in \mathbb{R}^d$, la matriz de covarianzas $\Sigma = \mathrm{Cov}(X)$ satisface

1. Σ es simétrica (como consecuencia, tiene autovalores reales).
2. Σ es semi-definida positiva (todos sus autovalores son no-negativos).
   En particular, para todo vector $\mathbf{x} \in \mathbb{R}^d$, se cumple que $\mathbf{x}^T\Sigma\mathbf{x} \geq 0$.
3. La diagonal de Σ contiene a las varianzas $\sigma_i^2 = \mathrm{Cov}(X_i)$, para $i = 1, 2, \ldots, d$.
4. Si $\mathbb{X} \in \mathbb{R}^{n\times d}$ es la matriz de datos. Entonces $\Sigma = \mathbb{X}^T\mathbb{X}$.

Distribuciones multivariadas

Sea $X_1, X_2, \ldots, X_n \in \mathbb{R}^d$ es una muestra aleatoria de vectores i.i.d (independientes e identicamente distribuidos), todos con distribución X.

Podemos codificar esta muestra dentro de una matriz, $\mathbb{X} \in \ltimes \times$, llamada la **matriz de datos** (cada dato de la muestra es un renglón de $\mathbb{X}$).

$$\mathbb{X} = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1d} \\ x_{21} & x_{22} & \ldots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{nd} \end{pmatrix},$$

donde

$$X_i = (x_{i1}, x_{i2}, \ldots, x_{id}) \in \mathbb{R}^d, \quad i = 1, 2, \ldots, n.$$

Distribuciones multivariadas

Observe que la $i-$ésima columna de $\mathbb{X}$ corresponde a una muestra (de tamaño n) de la variable aleatoria X_i. Podemos entonces restar a cada columna su correspondiente media $\mu_i = \mathbb{E}(X_i)$. Así, obtenemos una versión centrada de la matriz de datos:

$$\mathbb{X}_c = \mathbb{X} - \mu = \begin{pmatrix} x_{11} - \mu_1 & x_{12} - \mu_2 & \dots & x_{1d} - \mu_d \\ x_{21} - \mu_1 & x_{22} - \mu_2 & \dots & x_{2d} - \mu_d \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} - \mu_1 & x_{n2} - \mu_2 & \dots & x_{nd} - \mu_d \end{pmatrix}.$$

Es posible mostrar (con las propiedades de la página siguiente) que la matriz de covarianzas empírica (muestral) se puede escribir como

$$\Sigma = \mathrm{Cov}(X) = \mathbb{X}_c^T \mathbb{X}_c.$$

Distribuciones multivariadas

Sea $X, Y \in \mathbb{R}^d$ vectores aleatorios, $a, b \in \mathbb{R}$, $c \in \mathbb{R}^d$ constantes, $A \in \mathbb{R}^{p\times d}$ una matriz constante. Entonces

1. $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$,
2. $\mathbb{E}(c) = c$,
3. $\mathbb{E}(\mathbb{E}(X)) = \mathbb{E}(X)$,
4. $\operatorname{Cov}(aX) = a^2 \operatorname{Cov}(X)$,
5. $\operatorname{Cov}(\mathbf{AX}) = \mathbf{A}^{\mathbf{T}} \operatorname{Cov}(\mathbf{X})\mathbf{A}$,
6. $\operatorname{Cov}(aX + bY) = a^2 \operatorname{Cov}(X) + b^2 \operatorname{Cov}(Y) + 2abCov(X, Y)$,
7. Si $X \perp Y$, entonces $\operatorname{Cov}(X, Y) = 0$,
8. Si $X \perp Y$, entonces $\operatorname{Cov}(aX + bY) = a^2 \operatorname{Cov}(X) + b^2 \operatorname{Cov}(Y)$.

Distribución normal multivariada

Sea $X = (X_1, X_2, \ldots, X_d) \in \mathbb{R}^d$ un vector aleatorio. Decimos que X sigue una **distribución normal multivariada** $\mathcal{N}_d(\mu, \Sigma)$ si su densidad está dada por

$$f_X(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}\det(\Sigma)^{1/2}} \exp\Big(-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)\Big)\, d\mathbf{x}.$$

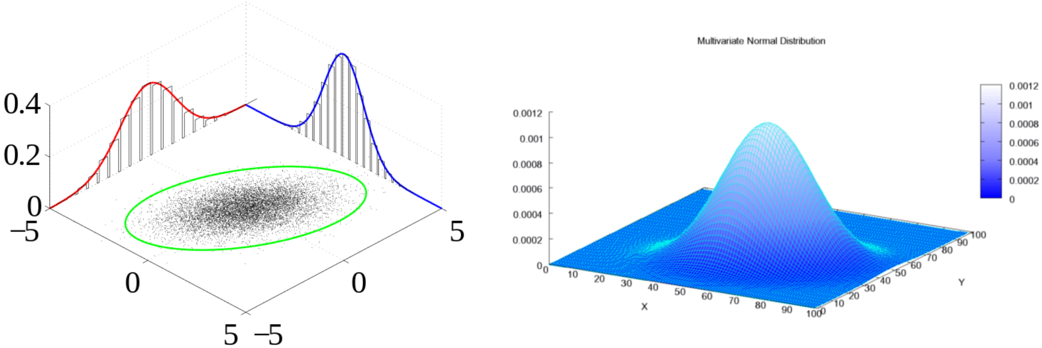
Aquí,

$$\mathbb{E}(X) = \mu = (\mu_1, \mu_2, \ldots, \mu_d)^T,$$

y

$$\text{Cov}(X) = \Sigma = \begin{pmatrix} Cov(X_1,X_1) & Cov(X_1,X_2) & \ldots & Cov(X_1,X_d) \\ Cov(X_2,X_1) & Cov(X_2,X_2) & \ldots & Cov(X_2,X_d) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_d,X_1) & Cov(X_d,X_2) & \ldots & Cov(X_d,X_d) \end{pmatrix}.$$

Distribución normal multivariada



Densidad de una normal bivariada: (a) como nube de puntos, (b) como función.

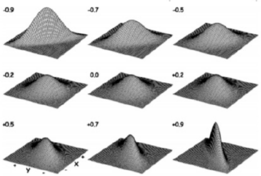
Típicamente la matriz Σ proporciona información sobre la relación entre las variables componentes.

$Cov(X) = [Cov(X_i, X_j)]$ es una matriz simétrica y pos. definida.

Caso $d = 2$:

$$\begin{pmatrix} Var(X_1) & Cov(X_1, X_2) \\ Cov(X_1, X_2) & Var(X_2) \end{pmatrix} = \begin{pmatrix} \sigma_{X_1}^2 & Cor(X_1, X_2)\sigma_{X_1}\sigma_{X_2} \\ Cor(X_1, X_2)\sigma_{X_1}\sigma_{X_2} & \sigma_{X_2}^2 \end{pmatrix}$$

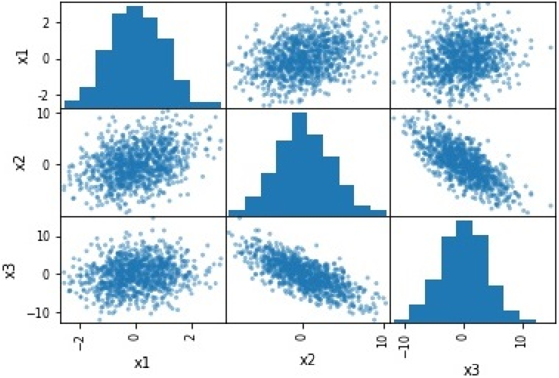
Cambiar $\rho = Cor(X_1, X_2)$:



http://personal.kenyon.edu/hartlaub/MellonProject/images/Bivariate52.gif.

Distribución normal multivariada

Una forma práctica de ver esta información de covarianza o correlación entre las componentes es a través de *pair-plots*.



Pairplot de una muestra para una normal 3-variada.

Distribución normal multivariada

¿Cómo generar una muestra de una distribución normal d-variada con μ y Σ específicas?

Algoritmo (o receta):

1. Generar d muestras (de tamaño n), independientes, de distribuciones normales estándar $Z_1, Z_2, \ldots, Z_d \in \mathbb{R}^n$, y construir una matriz de datos $\mathbb{Z}$ con las muestras Z_i como columnas.
Como son independientes y estándar el vector $Z = (Z_1, \ldots, Z_d)$ sigue una distribución normal estándar $\mathcal{N}_d(\mathbf{0}, I_d)$.

2. Asegurarse que la matriz Σ es simétrica y positiva definida. Luego, construir descomposición de Cholesky $L^T L = \Sigma$, (veremos este algoritmo más adelante).

3. Construir la variable aleatoria $X = LZ + \mu$, la cual tiene una matriz de datos dada por $\mathbb{X} = L\mathbb{Z} + \mu$ (la muestra que queremos). De las propiedades anteriores, tenemos que $\mathbb{E}(Z) = \mu$ y $\text{Cov}(X) = L^T I_d L = L^T L = \Sigma$.