Fundamentos de Análisis de Datos
Estadística Descriptiva (Caso Multivariado)

Dra. Andrea Alejandra Rey

Especialización en Ciencia de Datos - ITBA

Empleo de arreglos: vectores y matrices

Extensión de la estadística descriptiva unidimensional

Arreglos

Un vector n-dimensional es un arreglo $X \in \mathbb{R}$, que puede ser:

Vector columna

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Vector fila

$$X = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}$$

Arreglos

Una matriz de tamaño $n \times m$ es un arreglo $X \in \mathbb{R}^{n \times m}$ de la forma:

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix}.$$

Observación

- Si $X \in \mathbb{R}^{n \times m}$, entonces X tiene n filas y m columnas.
- Si $n = m$, decimos que X es una matriz **cuadrada** de orden n.

Arreglos

Definición

Dada $X \in \mathbb{R}^{n\times m}$, se define su **matriz traspuesta** como $X' \in \mathbb{R}^{m\times n}$ donde las filas de X' son las columnas de X o, equivalentemente, las columnas de X' son las filas de X.
La matriz traspuesta también se denota por X^t o X^T.

Ejemplo

$$X = \begin{pmatrix} 1 & 2 & 0 & -1 \\ 2 & 3 & 1 & 0 \\ 0 & 0 & -2 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 4} \Longrightarrow X' = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 1 & -2 \\ -1 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{4\times 3}$$

Operaciones de matrices

Definición

Sean $X, Y \in \mathbb{R}^{n\times m}$, se define la **suma** como:

$$X + Y = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix} + \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1m} \\ y_{21} & y_{22} & \cdots & y_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nm} \end{pmatrix}$$

$$= \begin{pmatrix} x_{11}+y_{11} & x_{12}+y_{12} & \cdots & x_{1m}+y_{1m} \\ x_{21}+y_{21} & x_{22}+y_{22} & \cdots & x_{2m}+y_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1}+y_{n1} & x_{n2}+y_{n2} & \cdots & x_{nm}+y_{nm} \end{pmatrix}.$$

Operaciones de matrices

Propiedades de la suma

- La suma está definida sólo para matrices del mismo tamaño.
- La suma de dos matrices da por resultado una matriz del mismo tamaño.
- La suma es matrices es conmutativa; es decir, $X + Y = Y + X$.
- La suma de matrices es asociativa; es decir, si X, Y, Z son tres matrices del mismo tamaño entonces $X + (Y + Z) = (X + Y) + Z$.

Operaciones de matrices

Definición

Sean $X \in \mathbb{R}^{n\times m}$ y $\lambda \in \mathbb{R}$, se define el **producto por un escalar** como:

$$\lambda X = \lambda \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix} = \begin{pmatrix} \lambda x_{11} & \lambda x_{12} & \cdots & \lambda x_{1m} \\ \lambda x_{21} & \lambda x_{22} & \cdots & \lambda x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda x_{n1} & \lambda x_{n2} & \cdots & \lambda x_{nm} \end{pmatrix}.$$

Operaciones de matrices

Propiedades del producto por un escalar

- Entendemos por escalar a cualquier número real.
- El producto de una matriz por un escalar da por resultado una matriz del mismo tamaño.
- El producto por un escalar es distributivo respecto de la suma; es decir, $\lambda(X+Y)=\lambda X+\lambda Y$.
- Dados dos escalares $\lambda,\mu\in\mathbb{R}$, se verifica lo siguiente:

$$
\begin{aligned}
(\lambda+\mu)X &= \lambda X+\mu X,\\
(\lambda\mu)X &= \lambda(\mu X)=\mu(\lambda X).
\end{aligned}
$$

- El 1 es el elemento neutro; es decir, $1X=X$.

Operaciones de matrices

Definición

Sean $X \in \mathbb{R}^{n\times m}$ e $Y \in \mathbb{R}^{m\times p}$, se define el **producto** como:

$$XY = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix} \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1m} \\ y_{21} & y_{22} & \cdots & y_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mp} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=1}^{m} x_{1k}y_{k1} & \sum_{k=1}^{m} x_{1k}y_{k2} & \cdots & \sum_{k=1}^{m} x_{1k}y_{kp} \\ \sum_{k=1}^{m} x_{2k}y_{k1} & \sum_{k=1}^{m} x_{2k}y_{k2} & \cdots & \sum_{k=1}^{m} x_{2k}y_{kp} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{m} x_{nk}y_{k1} & \sum_{k=1}^{m} x_{nk}y_{k2} & \cdots & \sum_{k=1}^{m} x_{nk}y_{kp} \end{pmatrix}.$$

Operaciones de matrices

Veamos el siguiente ejemplo:

$$\underbrace{\begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix}}_{X \in \mathbb{R}^{2\times 2}} \underbrace{\begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \end{pmatrix}}_{Y \in \mathbb{R}^{2\times 3}} \Longrightarrow XY \in \mathbb{R}^{2\times 3}$$

$$(XY)_{11} = \begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \end{pmatrix} = 2 \cdot 1 + 1 \cdot 0 = \boxed{2}$$

$$(XY)_{12} = \begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \end{pmatrix} = 2 \cdot 2 + 1 \cdot 3 = \boxed{7}$$

$$(XY)_{13} = \begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \end{pmatrix} = 2 \cdot 1 + 1 \cdot 2 = \boxed{4}$$

Operaciones de matrices

$$(XY)_{21} = \begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \end{pmatrix} = 2 \cdot 1 + 2 \cdot 0 = \boxed{2}$$

$$(XY)_{22} = \begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \end{pmatrix} = 2 \cdot 2 + 2 \cdot 3 = \boxed{10}$$

$$(XY)_{23} = \begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \end{pmatrix} = 2 \cdot 1 + 2 \cdot 2 = \boxed{6}$$

Luego,

$$XY = \begin{pmatrix} 2 & 7 & 4 \\ 2 & 10 & 6 \end{pmatrix}.$$

Operaciones de matrices

Propiedades del producto

- El producto matricial está definido sólo cuando la cantidad de columnas de la matriz de la izquierda coincide con la cantidad de filas de la matriz de la derecha.
- El producto de dos matrices da por resultado una matriz con la misma cantidad de filas que la matriz de la izquierda y la misma cantidad de columnas que la matriz de la derecha.
- El producto de matrices **no** es conmutativo.
- El producto de matrices es asociativo; es decir, si X, Y, Z son tres matrices con los tamaños adecuados $X(YZ) = (XY)Z$.
- La matriz identidad I, que es una matriz cuadrada con unos en su diagonal principal y ceros fuera de ella, es el elemento neutro del producto matricial.

Operaciones de matrices

Inversa de una matriz

Sea $A \in \mathbb{R}^{n\times n}$. Se dice que A es **inversible** si existe $B \in \mathbb{R}^{n\times n}$ tal que:

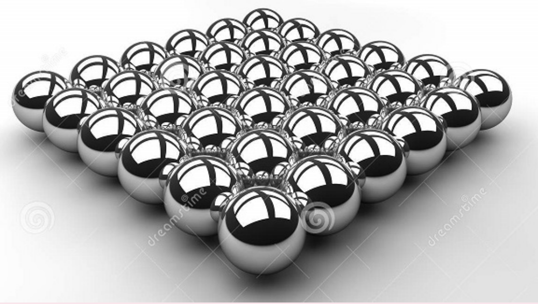
$$AB = BA = I,$$

donde $I \in \mathbb{R}^{n\times n}$ es la matriz identidad.
En este caso, se dice que B es la **matriz inversa** de A y se denota por A^{-1}.

Por ejemplo, $\begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$ y $\begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}$ son matrices inversas pues:

$$\begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix} = \begin{pmatrix} 6-5 & -3+3 \\ 10-10 & -5+6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.$$

Ejemplo



Vectores aleatorios

Dadas $X_1, X_2, \ldots, X_m$ variables aletorias, podemos definir el **vector aleatorio** $\boldsymbol{X} = (X_1, X_2, \ldots, X_m)$.
Una **muestra** de $\boldsymbol{X}$ es un vector $\boldsymbol{x}_1 = (x_{11}, x_{12}, \ldots, x_{1m})$.
Si tenemos n muestras, podemos armar el arreglo:

$$\boldsymbol{x} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix}.$$

Vectores aleatorios

Definición

Si $\boldsymbol{x}$ es una muestra de tamaño n del vector aletorio $\boldsymbol{X}$, se define:

- el **vector de medias muestral** como $\bar{\boldsymbol{x}} = \begin{pmatrix} \bar{\boldsymbol{x}}_1 \\ \bar{\boldsymbol{x}}_2 \\ \vdots \\ \bar{\boldsymbol{x}}_m \end{pmatrix}$.
- el **vector de varianzas muestral** como $s^2_{\boldsymbol{x}} = \begin{pmatrix} s^2_{\boldsymbol{x}_1} \\ s^2_{\boldsymbol{x}_2} \\ \vdots \\ s^2_{\boldsymbol{x}_m} \end{pmatrix}$.

Ejemplo de base

País	Calidad de vida	Seguridad	Costo de vida	Polución*
Argentina	111.77	35.89	30.78	51.32
Brasil	105.13	33.36	33.77	53.68
México	127.36	45.55	36.99	58.52
Estados Unidos	173.79	50.97	71.60	36.06
Francia	154.59	44.73	67.55	42.83
Italia	141.49	53.14	60.39	54.73
India	118.31	55.36	22.15	73.00
Nueva Zelanda	177.61	54.16	71.60	24.19
Nepal	84.51	62.40	24.67	84.14
Finlandia	191.31	73.71	66.47	11.99

*Descargado desde `https://www.numbeo.com/cost-of-living/` (enero, 2023).

Ejemplo de base

Podemos pensar el ejemplo como 10 muestras de un vector aleatorio $\boldsymbol{X} = (X_1, X_2, X_3, X_4)$, donde:

- X_1 indica el índice de la calidad de vida de un país,
- X_2 indica el índice de la seguridad en un país,
- X_3 indica el índice del costo de vida de un país,
- X_4 indica el índice de polución en un país.

En este caso,

$$\bar{\boldsymbol{x}} = \begin{pmatrix} 138.587 \\ 50.927 \\ 48.597 \\ 49.046 \end{pmatrix} \quad \text{y} \quad s^2_{\boldsymbol{x}} = \begin{pmatrix} 1234.04 \\ 143.08 \\ 424.44 \\ 461.37 \end{pmatrix}$$

Covarianza

Mide la existencia de asociación lineal entre dos variables; es decir, si el comportamiento de las variables se mueve en la misma dirección o en direcciones opuestas.

Definición

Dadas dos muestras $x = (x_1, x_2, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_n)$ de las variables X e Y, respectivamente, se definen:

- la **covarianza poblacional** entre X e Y como:

$$\mathrm{Cov}(X, Y) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}),$$

- la **covarianza muestral** entre X e Y como:

$$\mathrm{Cov}(X, Y) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}).$$

Covarianza

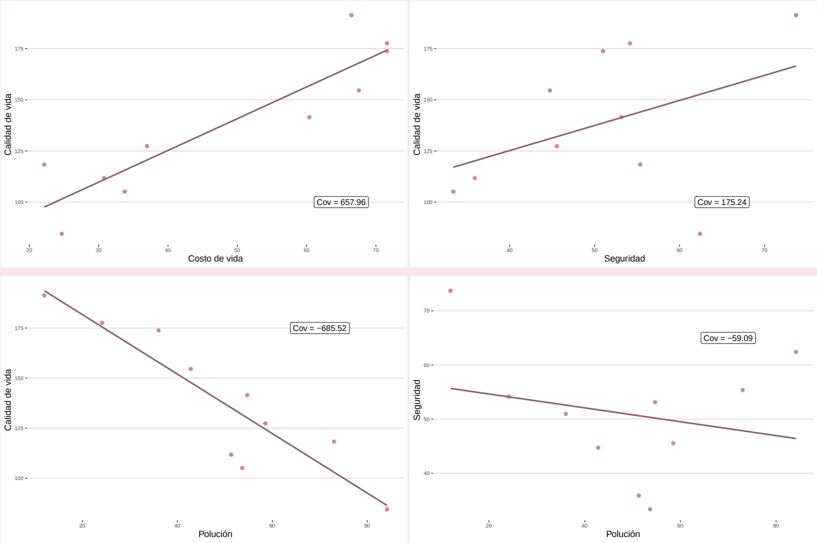
Observación

La **covarianza muestral** entre X y X, $\mathrm{Cov}(X,X)$, es en realidad la varianza muestral de X.

Interpretación

- Una covarianza positiva indica que ambas variables se mueven la misma dirección (ambas crecen o ambas decrecen).
- Una covarianza negativa indica que las variables se mueven en direcciones opuestas (cuando una crece, la otra decrece).
- Una covarianza nula indica que no hay una relación lineal entre las variables. En este caso se dice que las variables son **no correlacionadas**.

Covarianza


Calidad de vida
Costo de vida
Cov = 657.96
Calidad de vida
Seguridad
Cov = 175.24
Calidad de vida
Polución
Cov = −685.52
Seguridad
Polución
Cov = −59.09

Grado de asociación lineal

Pregunta

- ¿La covarianza mide si la relación entre las variables es fuerte o débil?
- ¿Se puede conseguir una relación directa o inversa perfecta entre dos variables?

Grado de asociación lineal

Mismas direcciones	Direcciones perpendiculares	Direcciones opuestas

$\cos(\theta) = 1$	$\cos(\theta) = 0$	$\cos(\theta) = -1$

Grado de asociación lineal

Pregunta

¿Cómo se calcula el ángulo entre dos vectores?

Definición

Sean $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ e $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ dos vectores en $\mathbb{R}^n$.
Si θ es el ángulo formado por ambos vectores, entonces:

$$\cos(\theta) = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}},$$

donde θ es un ángulo entre 0° y 180°.

Grado de asociación lineal

Supongamos ahora que tenemos dos muestras $x = (x_1, x_2, \ldots, x_n)$ e $y = (y_1, y_2, \ldots, y_n)$ de las variables X e Y, respectivamente.
Consideramos los vectores centrados en la media $x - \bar{x}$ e $y - \bar{y}$.
El ángulo formado por estos vectores satisface:

$$\begin{aligned}
\cos(\theta) &= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \\
&= \frac{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}} \\
&= \frac{\mathrm{Cov}(X, Y)}{s_x s_y}.
\end{aligned}$$

Correlación muestral

Mide el grado de asociación lineal entre dos variables.

Definición

Dadas dos muestras $x = (x_1, x_2, \ldots, x_n)$ e $y = (y_1, y_2, \ldots, y_n)$ de las variables X e Y, respectivamente, se define la **correlación muestral** entre X e Y como:

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{s_x s_y}.$$

Correlación muestral

Interpretación

- La correlación toma valores en el intervalo $[-1, 1]$.
- $\rho(X, Y) = 1$ indica una relación lineal directa perfecta, donde los datos están sobre una recta de pendiente positiva.
- $\rho(X, Y) = -1$ indica una relación lineal inversa perfecta, donde los datos están sobre una recta de pendiente negativa.
- $\rho(X, Y) = 0$ indica que no existe una relación lineal entre los datos.
- En particular:

$$\begin{aligned}
0 \leq |\rho(X, Y)| \leq 0.25 &\Longrightarrow \text{No hay relación lineal} \\
0.25 < |\rho(X, Y)| \leq 0.50 &\Longrightarrow \text{La relación lineal es débil} \\
0.50 < |\rho(X, Y)| \leq 0.75 &\Longrightarrow \text{La relación lineal es moderada} \\
0.75 < |\rho(X, Y)| \leq 1 &\Longrightarrow \text{La relación lineal es fuerte}
\end{aligned}$$

Correlación muestral

En nuestro ejemplo:

- El coeficiente de correlación entre el costo y la calidad de vida es 0.9091.
- El coeficiente de correlación entre la seguridad y la calidad de vida es 0.4170.
- El coeficiente de correlación entre la polución y la calidad de vida es -0.9085.
- El coeficiente de correlación entre la polución y la seguridad es -0.2280.

Covarianza versus correlación

- La covarianza sólo mide la dirección del movimiento de las variables, mientras que la correlación mide además el grado de la relación.
- El valor de la covarianza no está acotado, mientras que el valor de la correlación está acotado entre -1 y 1.
- La correlación es invariante por escalamiento de las variables. Sin embargo, la covarianza depende de la escala elegida.
- La covarianza tiene unidad de medida que hereda de las variables, mientras que la correlación es un coeficiente que no tiene unidad de medida.

Matriz de covarianzas muestral

Es una matriz cuadrada que muestra las covarianzas de todas las variables de un conjunto de datos, resultando una herramienta muy utilizada en el análisis multivariado.

Definición

Sea $\boldsymbol{X} = (X_1, X_2, \ldots, X_m)$ un vector aletorio.
Se define la **matriz de covarianzas** de $\boldsymbol{X}$ como:

$$\Sigma = \begin{pmatrix} \mathrm{Cov}(X_1, X_1) & \mathrm{Cov}(X_1, X_2) & \cdots & \mathrm{Cov}(X_1, X_m) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Cov}(X_2, X_2) & \cdots & \mathrm{Cov}(X_2, X_m) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_m, X_1) & \mathrm{Cov}(X_m, X_2) & \cdots & \mathrm{Cov}(X_m, X_m) \end{pmatrix} \in \mathbb{R}^{m \times m}.$$

Matriz de covarianzas muestral

Propiedades

- Como $\mathrm{Cov}(X_i, X_i) = \mathrm{Var}(X_i)$ para $i = 1, 2, \ldots, m$, la diagonal de la matriz de covarianzas está formada por las varianzas de las componentes de X. Es por ello que esta matriz también es conocida como **matriz de varianzas y covarianzas**.
- Como $\mathrm{Cov}(X_i, X_j) = \mathrm{Cov}(X_j, X_i)$ para todo para $i, j = 1, 2, \ldots, m$, la matriz de covarianzas es **simétrica**; es decir, $\Sigma = \Sigma'$.
- Si las variables son no correlacionadas, la matriz de covarianzas es diagonal; es decir, los elementos fuera de la diagonal son nulos.
- La matriz de covarianzas es semidefinida positiva; es decir, $x'\Sigma x \geq 0$ para cualquier vector columna $x \in \mathbb{R}^m$.

Matriz de covarianzas muestral

La matriz de covarianzas de nuestro ejemplo de base es:

$$
\begin{array}{r c}
 & \begin{array}{rrrr} \text{Calidad de vida} & \text{Seguridad} & \text{Costo de vida} & \text{Polución} \end{array} \\
\begin{array}{r} \text{Calidad de vida} \\ \text{Seguridad} \\ \text{Costo de vida} \\ \text{Polución} \end{array} &
\left(\begin{array}{rrrr}
1234.04 & 175.24 & 657.96 & -685.52 \\
175.24 & 143.08 & 65.58 & -59.09 \\
657.96 & 65.58 & 424.44 & -364.00 \\
-685.52 & -59.09 & -364 & 461.37
\end{array} \right)
\end{array}
$$

Ejemplo*

*Base de datos `trees` de R

Distribución Normal multivariada

Importancia

- Tiene una fórmula matemática relativamente sencilla por lo que resulta simple obtener modelos multivariados basados en esta distribución.
- Permite generalizar el Teorema Central del Límite para el caso univariado, el cual asegura que el vector de medias de un conjunto de vectores aleatorios independientes e idénticamente distribuidos, tiene una distribución Normal multivariada aproximada para un número grande de muestras.
- Al igual que en el caso univariado, muchos problemas naturales pueden ser modelados usando esta distribución.

Distribución Normal multivariada

Definición

Sea $\boldsymbol{X} = (X_1, X_2, \ldots, X_m)$ un vector aleatorio.
Decimos que $\boldsymbol{X}$ sigue una distribución Normal multivariada con vector de medias $\boldsymbol{\mu}$ y matriz de covarianzas Σ, si la función de densidad de probabilidad conjunta es de la forma:

$$f(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^m|\Sigma|}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right],$$

donde $|\Sigma|$ denota el determinante de la matriz de covarianzas y Σ^{-1} su matriz inversa, y $\boldsymbol{x} \in \mathbb{R}^m$.

En este caso, notamos $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$.

Distribución Normal multivariada

Propiedades

Supongamos que $\boldsymbol{X} = (X_1, X_2, \ldots, X_m) \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$.

- La variable X_i sigue una distribución Normal univariada con media μ_i y varianza Σ_{ii}, para $i = 1, 2, \ldots, m$.
- Un subconjunto de variables de $\boldsymbol{X}$ tiene una distribución Normal multivariada.
- Cualquier combinación lineal $\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_m X_m$, con $\alpha_i \in \mathbb{R}$ para $i = 1, 2, \ldots, m$, tiene una distribución Normal univariada.
- Las componentes $X_1, X_2, \ldots, X_m$ son independientes si y sólo si son no correlacionadas; es decir, $\Sigma_{ij} = 0$ para todo $i \neq j$.
- Las distribuciones condicionales derivadas de las componentes son normales.

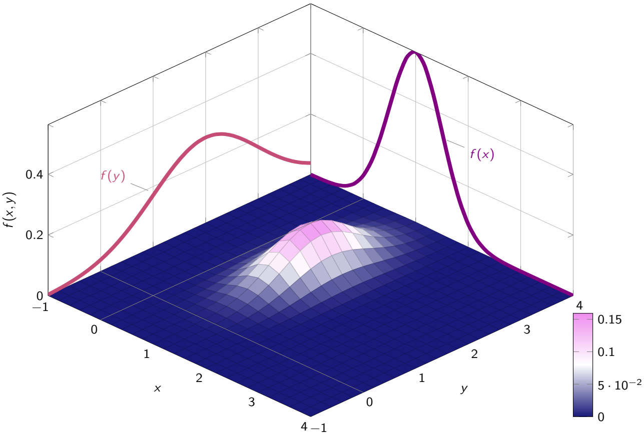
Distribución Normal multivariada

Observación

- Si $\boldsymbol{X} = (X_1, X_2, \ldots, X_m)$ tiene distribución Normal multivariada, todas las variables X_i siguen una distribución Normal.
- Sin embargo, la recíproca no es cierta. Es decir, la distribución conjunta de variables normales no necesariamente sigue una distribución Normal bivariada.
- Si $X_1, X_2, \ldots, X_m$ son variables aleatorias **independientes** con distribución Normal, entonces el vector $\boldsymbol{X} = (X_1, X_2, \ldots, X_m)$ tiene distribución Normal multivariada.

Distribución Normal bivariada

Gráfico



Distribución Normal bivariada

En el caso particular en que el vector aleatorio tenga dos componentes (X, Y), tenemos que:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} \qquad \text{y} \qquad \Sigma = \begin{pmatrix} \sigma_X^2 & C \\ C & \sigma_Y^2 \end{pmatrix},$$

donde $C = \mathrm{Cov}(X, Y)$.
La matriz inversa de la matriz de covarianzas está dada por:

$$\Sigma^{-1} = \frac{1}{\sigma_X^2 \sigma_Y^2 - C^2} \begin{pmatrix} \sigma_Y^2 & -C \\ -C & \sigma_X^2 \end{pmatrix}.$$

Distribución Normal bivariada

Observemos que en este caso, la expresión $(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$ es igual a:

$$\begin{pmatrix} x-\mu_X & y-\mu_Y \end{pmatrix} \frac{1}{\sigma_X^2\sigma_Y^2-C^2} \begin{pmatrix} \sigma_Y^2 & -C \\ -C & \sigma_X^2 \end{pmatrix} \begin{pmatrix} x-\mu_X \\ y-\mu_Y \end{pmatrix} =$$

$$\frac{1}{\sigma_X^2\sigma_Y^2-C^2} \begin{pmatrix} x-\mu_X & y-\mu_Y \end{pmatrix} \left[\begin{pmatrix} \sigma_Y^2 & -C \\ -C & \sigma_X^2 \end{pmatrix} \begin{pmatrix} x-\mu_X \\ y-\mu_Y \end{pmatrix} \right] =$$

$$\frac{1}{\sigma_X^2\sigma_Y^2-C^2} \begin{pmatrix} x-\mu_X & y-\mu_Y \end{pmatrix} \begin{pmatrix} \sigma_Y^2(x-\mu_X)-C(y-\mu_Y) \\ -C(x-\mu_X)+\sigma_X^2(y-\mu_Y) \end{pmatrix} =$$

$$\frac{(x-\mu_X)[\sigma_Y^2(x-\mu_X)-C(y-\mu_Y)]+(y-\mu_Y)[-C(x-\mu_X)+\sigma_X^2(y-\mu_Y)]}{\sigma_X^2\sigma_Y^2-C^2} =$$

$$\frac{\sigma_Y^2(x-\mu_X)^2-2C(x-\mu_X)(y-\mu_Y)+\sigma_X^2(y-\mu_Y)^2}{\sigma_X^2\sigma_Y^2-C^2}.$$

Distribución Normal bivariada

Recordemos que si ρ denota el coeficiente de correlación entre X e Y, entonces $\rho = C/(\sigma_X\sigma_Y)$ o, equivalentemente, $C = \rho\sigma_X\sigma_Y$.
Reemplazando en lo anterior:

$$\frac{\sigma_Y^2(x-\mu_X)^2 - 2\rho\sigma_X\sigma_Y(x-\mu_X)(y-\mu_Y) + \sigma_X^2(y-\mu_Y)^2}{\sigma_X^2\sigma_Y^2 - \rho^2\sigma_X^2\sigma_Y^2} =$$

$$\frac{\sigma_X^2\sigma_Y^2\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\sigma_X^2\sigma_Y^2\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \sigma_X^2\sigma_Y^2\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2}{\sigma_X^2\sigma_Y^2(1-\rho^2)} =$$

$$\frac{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2}{1-\rho^2}.$$

Además,

$$|\Sigma| = \sigma_X^2\sigma_Y^2 - C^2 = \sigma_X^2\sigma_Y^2 - \rho^2\sigma_X^2\sigma_Y^2 = \sigma_X^2\sigma_Y^2(1-\rho^2).$$

Distribución Normal bivariada

Reemplazando en la fórmula de la función de densidad de probabilidad, en el caso $m = 2$, concluimos que la expresión para el caso bivariado es:

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

$$\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right\}.$$

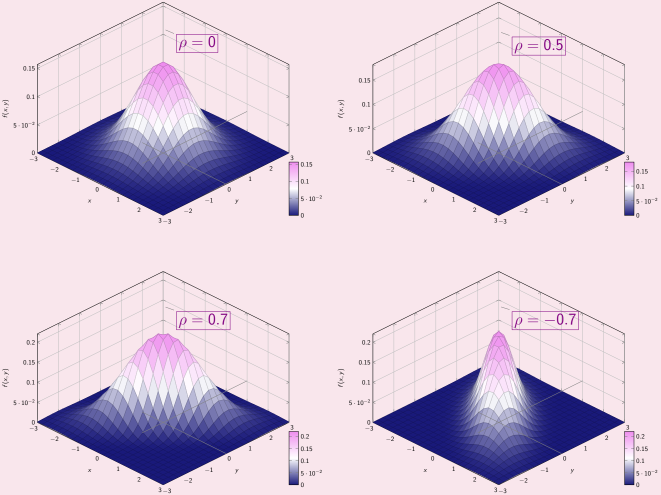
Distribución Normal bivariada

Pregunta

¿Cómo influye el valor de ρ en la gráfica?

Supongamos que $(X, Y) \sim \mathcal{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right)$.

Distribución Normal bivariada



Ejemplo