Probabilidad y Estadística para Inteligencia Artificial

Dr. Ing. Pablo Briff

Laboratorio de Sistemas Embebidos - FIUBA

pbriff@fi.uba.ar

4 de Julio de 2020

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Tabla de Contenidos I

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Tabla de Contenidos II

Repaso de Octave

- Vectores y matrices, producto y exponenciación

Repaso de Octave

- Vectores y matrices, producto y exponenciación
- Funciones

Repaso de Octave

- Vectores y matrices, producto y exponenciación
- Funciones
- Operaciones matriciales, producto y exponenciación

Repaso de Octave

- Vectores y matrices, producto y exponenciación
- Funciones
- Operaciones matriciales, producto y exponenciación
- Operaciones básicas de programación, for, while, if

Repaso de Octave

- Vectores y matrices, producto y exponenciación
- Funciones
- Operaciones matriciales, producto y exponenciación
- Operaciones básicas de programación, for, while, if
- Funciones recursivas

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Repaso de Octave

- Vectores y matrices, producto y exponenciación
- Funciones
- Operaciones matriciales, producto y exponenciación
- Operaciones básicas de programación, for, while, if
- Funciones recursivas
- Histograma

Repaso de Octave

- Vectores y matrices, producto y exponenciación
- Funciones
- Operaciones matriciales, producto y exponenciación
- Operaciones básicas de programación, for, while, if
- Funciones recursivas
- Histograma
- Pdf y cdf

- Vectores y matrices, producto y exponenciación
- Funciones
- Operaciones matriciales, producto y exponenciación
- Operaciones básicas de programación, for, while, if
- Funciones recursivas
- Histograma
- Pdf y cdf
- Gráficos

Repaso de Octave

- Vectores y matrices, producto y exponenciación
- Funciones
- Operaciones matriciales, producto y exponenciación
- Operaciones básicas de programación, for, while, if
- Funciones recursivas
- Histograma
- Pdf y cdf
- Gráficos
- Guardar y recuperar datos

Producto de Convolución

- Sean $f(t), g(t)$ dos funciones reales

Producto de Convolución

- Sean $f(t), g(t)$ dos funciones reales
- Sea $y(t)$ el producto de convolución de f, g definido por

Producto de Convolución

- Sean $f(t), g(t)$ dos funciones reales
- Sea $y(t)$ el producto de convolución de f, g definido por
- $y(t) \triangleq (f * g)(t) = \int_{-\infty}^{+\infty} f(\tau)g(t-\tau)d\tau$

Producto de Convolución

- Sean $f(t), g(t)$ dos funciones reales
- Sea $y(t)$ el producto de convolución de f, g definido por
- $y(t) \triangleq (f * g)(t) = \int_{-\infty}^{+\infty} f(\tau)g(t-\tau)d\tau$
- Para funciones que son cero para argumentos negativos, la integral se reduce a:

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Producto de Convolución

- Sean $f(t), g(t)$ dos funciones reales
- Sea $y(t)$ el producto de convolución de f, g definido por
- $y(t) \triangleq (f * g)(t) = \int_{-\infty}^{+\infty} f(\tau)g(t-\tau)d\tau$
- Para funciones que son cero para argumentos negativos, la integral se reduce a:
- $y(t) = (f * g)(t) = \int_{0}^{t} f(\tau)g(t-\tau)d\tau$

Producto de Convolución

- Sean $f(t), g(t)$ dos funciones reales
- Sea $y(t)$ el producto de convolución de f, g definido por
- $y(t) \triangleq (f * g)(t) = \int_{-\infty}^{+\infty} f(\tau)g(t-\tau)d\tau$
- Para funciones que son cero para argumentos negativos, la integral se reduce a:
- $y(t) = (f * g)(t) = \int_{0}^{t} f(\tau)g(t-\tau)d\tau$
- Este se denomina producto de convolución unilateral

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Producto de Convolución

- Sean $f(t), g(t)$ dos funciones reales
- Sea $y(t)$ el producto de convolución de f, g definido por
- $y(t) \triangleq (f * g)(t) = \int_{-\infty}^{+\infty} f(\tau)g(t-\tau)d\tau$
- Para funciones que son cero para argumentos negativos, la integral se reduce a:
- $y(t) = (f * g)(t) = \int_{0}^{t} f(\tau)g(t-\tau)d\tau$
- Este se denomina producto de convolución unilateral
- El producto de convolución es lineal, asociativo y conmutativo

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Producto de Convolución

- Para funciones discretas $f(n), g(n)$ el producto de convolución es una suma dada por:

Producto de Convolución

- Para funciones discretas $f(n), g(n)$ el producto de convolución es una suma dada por:
- $y(n) \triangleq (f * g)(n) = \sum_{k=-\infty}^{+\infty} f(k)g(n-k)$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Producto de Convolución

- Para funciones discretas $f(n), g(n)$ el producto de convolución es una suma dada por:
- $y(n) \triangleq (f * g)(n) = \sum_{k=-\infty}^{+\infty} f(k)g(n-k)$
- Se reduce a $y(n) = \sum_{k=0}^{n} f(k)g(n-k)$ si $f(n) = g(n) = 0$ para $n < 0$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Producto de Convolución

- Para funciones discretas $f(n), g(n)$ el producto de convolución es una suma dada por:
- $y(n) \triangleq (f * g)(n) = \sum_{k=-\infty}^{+\infty} f(k)g(n-k)$
- Se reduce a $y(n) = \sum_{k=0}^{n} f(k)g(n-k)$ si $f(n) = g(n) = 0$ para $n < 0$
- Ejemplo: calcular el producto de convolución de dos vectores $X = [1\,2\,3\,4]^T$, $Y = X$ usando la función `conv`. ¿Cuál es la longitud del vector resultante?

Producto de Convolución

- Para funciones discretas $f(n), g(n)$ el producto de convolución es una suma dada por:
- $y(n) \triangleq (f * g)(n) = \sum_{k=-\infty}^{+\infty} f(k)g(n-k)$
- Se reduce a $y(n) = \sum_{k=0}^{n} f(k)g(n-k)$ si $f(n) = g(n) = 0$ para $n < 0$
- Ejemplo: calcular el producto de convolución de dos vectores $X = [1\,2\,3\,4]^T$, $Y = X$ usando la función `conv`. ¿Cuál es la longitud del vector resultante?
- Repetir para $X = [1\,2\,3\,1\,2\,3\,1\,2\,3]^T$, $Y = [(1/3)\,(1/3)\,(1/3)]^T$. Explicar el resultado. ¿Qué función cumple Y?

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

- Sean X, Y dos V.A. $U[0, 1]$ independientes

Suma de V.A. Uniformes Independientes

- Sean X, Y dos V.A. $U[0, 1]$ independientes
- ¿Cuál es la distribución de $Z = X + Y$?

Suma de V.A. Uniformes Independientes

- Sean X, Y dos V.A. $U[0,1]$ independientes
- ¿Cuál es la distribución de $Z = X + Y$?
- $f_X(x) = f_Y(y) = \begin{cases} 1, & \text{si} \quad 0 \leq x, y \leq 1 \\ 0, & \text{si no} \end{cases}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Suma de V.A. Uniformes Independientes

- Sean X, Y dos V.A. $U[0,1]$ independientes
- ¿Cuál es la distribución de $Z = X + Y$?
- $f_X(x) = f_Y(y) = \begin{cases} 1, & \text{si} \quad 0 \leq x, y \leq 1 \\ 0, & \text{si no} \end{cases}$
- La función de densidad de Z es:

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Suma de V.A. Uniformes Independientes

- Sean X, Y dos V.A. $U[0,1]$ independientes
- ¿Cuál es la distribución de $Z = X + Y$?
- $f_X(x) = f_Y(y) = \begin{cases} 1, & \text{si} \quad 0 \leq x, y \leq 1 \\ 0, & \text{si no} \end{cases}$
- La función de densidad de Z es:
- $f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

- Sean X, Y dos V.A. $U[0,1]$ independientes
- ¿Cuál es la distribución de $Z = X + Y$?
- $f_X(x) = f_Y(y) = \begin{cases} 1, & \text{si} \quad 0 \leq x, y \leq 1 \\ 0, & \text{si no} \end{cases}$
- La función de densidad de Z es:
- $f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$
- $f_Z(z) = (f_X * f_Y)(z) = \int_0^1 f_X(z-y) f_Y(y) dy$

Suma de V.A. Uniformes Independientes

- Sean X, Y dos V.A. $U[0,1]$ independientes
- ¿Cuál es la distribución de $Z = X + Y$?
- $f_X(x) = f_Y(y) = \begin{cases} 1, & \text{si} \quad 0 \leq x, y \leq 1 \\ 0, & \text{si no} \end{cases}$
- La función de densidad de Z es:
- $f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$
- $f_Z(z) = (f_X * f_Y)(z) = \int_0^1 f_X(z-y) f_Y(y) dy$
- Tomando la integral para $0 \leq z - y \leq 1$ queda

FACULTAD DE INGENIERIA Universidad de Buenos Aires

- Sean X, Y dos V.A. $U[0,1]$ independientes
- ¿Cuál es la distribución de $Z = X + Y$?
- $f_X(x) = f_Y(y) = \begin{cases} 1, & \text{si} \quad 0 \leq x, y \leq 1 \\ 0, & \text{si no} \end{cases}$
- La función de densidad de Z es:
- $f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$
- $f_Z(z) = (f_X * f_Y)(z) = \int_0^1 f_X(z-y) f_Y(y) dy$
- Tomando la integral para $0 \leq z - y \leq 1$ queda
- $f_Z(z) = \begin{cases} z, & \text{si} \quad 0 \leq z \leq 1 \\ 2 - z, & \text{si} \quad 1 \leq z \leq 2 \\ 0, & \text{si no} \end{cases}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Suma de V.A. Uniformes Independientes

- Simular la distribución de Z en Octave usando la función `conv`

Suma de V.A. Uniformes Independientes

- Simular la distribución de Z en Octave usando la función `conv`
- Extender a $Z = X_1 + X_2 + X_3 + X_4$
  con $X_i \sim U[0, 1]$ (agrupar de a pares) y simular

- Sean X, Y dos V.A. $\mathcal{N}(0,1)$ independientes

Suma de V.A. Gaussianas Independientes

- Sean X, Y dos V.A. $\mathcal{N}(0, 1)$ independientes
- ¿Cuál es la distribución de $Z = X + Y$?

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Suma de V.A. Gaussianas Independientes

- Sean X, Y dos V.A. $\mathcal{N}(0,1)$ independientes
- ¿Cuál es la distribución de $Z = X + Y$?
- $f_Z(z) = (f_X * f_Y)(z) = \frac{1}{\sqrt{4\pi}} e^{-z^2/4}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Suma de V.A. Gaussianas Independientes

- Sean X, Y dos V.A. $\mathcal{N}(0,1)$ independientes
- ¿Cuál es la distribución de $Z = X + Y$?
- $f_Z(z) = (f_X * f_Y)(z) = \frac{1}{\sqrt{4\pi}} e^{-z^2/4}$
- $\mu_Z = \mu_X + \mu_Y = 0$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Suma de V.A. Gaussianas Independientes

- Sean X, Y dos V.A. $\mathcal{N}(0,1)$ independientes
- ¿Cuál es la distribución de $Z = X + Y$?
- $f_Z(z) = (f_X * f_Y)(z) = \frac{1}{\sqrt{4\pi}} e^{-z^2/4}$
- $\mu_Z = \mu_X + \mu_Y = 0$
- Se puede demostrar la $f_Z(z)$ a partir del producto de convolución

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Suma de V.A. Gaussianas Independientes

- Sean X, Y dos V.A. $\mathcal{N}(0,1)$ independientes
- ¿Cuál es la distribución de $Z = X + Y$?
- $f_Z(z) = (f_X * f_Y)(z) = \frac{1}{\sqrt{4\pi}} e^{-z^2/4}$
- $\mu_Z = \mu_X + \mu_Y = 0$
- Se puede demostrar la $f_Z(z)$ a partir del producto de convolución
- $\mathrm{var}[Z] = \mathrm{var}[X] + \mathrm{var}[Y] = 2$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Suma de V.A. Gaussianas Independientes

- Sean X, Y dos V.A. $\mathcal{N}(0,1)$ independientes
- ¿Cuál es la distribución de $Z = X + Y$?
- $f_Z(z) = (f_X * f_Y)(z) = \frac{1}{\sqrt{4\pi}} e^{-z^2/4}$
- $\mu_Z = \mu_X + \mu_Y = 0$
- Se puede demostrar la $f_Z(z)$ a partir del producto de convolución
- $\text{var}[Z] = \text{var}[X] + \text{var}[Y] = 2$
- Simular en Octave usando dos alternativas: i) la función `conv` y ii) la función `hist`

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

- Sea $Y = g(X)$ una función de una v.a. continua

Funciones de una V.A.

- Sea $Y = g(X)$ una función de una v.a. continua
- La esperanza de Y se puede encontrar sin calcular $f_Y(y)$, como $E[Y] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Funciones de una V.A.

- Sea $Y = g(X)$ una función de una v.a. continua
- La esperanza de Y se puede encontrar sin calcular $f_Y(y)$, como $E[Y] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$
- En general, $F_Y(y) = P(g(X) \leq y) = \int_{\{x|g(x) \leq y\}} f_X(x) dx$

Funciones de una V.A.

- Sea $Y = g(X)$ una función de una v.a. continua
- La esperanza de Y se puede encontrar sin calcular $f_Y(y)$, como $E[Y] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$
- En general, $F_Y(y) = P(g(X) \leq y) = \int_{\{x | g(x) \leq y\}} f_X(x) dx$
- La derivada de la cdf es la pdf, entonces:

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Funciones de una V.A.

- Sea $Y = g(X)$ una función de una v.a. continua
- La esperanza de Y se puede encontrar sin calcular $f_Y(y)$, como $E[Y] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$
- En general, $F_Y(y) = P(g(X) \leq y) = \int_{\{x | g(x) \leq y\}} f_X(x) dx$
- La derivada de la cdf es la pdf, entonces:
- $f_Y(y) = \frac{dF_Y}{dy}(y)$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Funciones de una V.A.

- Consideremos ahora el caso de funciones monótonas

Funciones de una V.A.

- Consideremos ahora el caso de funciones monótonas
- $g(t)$ es monótona y $h(t)$ es su inversa, asumamos que existe dh/dt

Funciones de una V.A.

- Consideremos ahora el caso de funciones monótonas
- $g(t)$ es monótona y $h(t)$ es su inversa, asumamos que existe dh/dt
- Sea X una variable aleatoria contenida en un intervalo I (es decir, $f_X(x) = 0$ fuera de I)

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Funciones de una V.A.

- Consideremos ahora el caso de funciones monótonas
- $g(t)$ es monótona y $h(t)$ es su inversa, asumamos que existe dh/dt
- Sea X una variable aleatoria contenida en un intervalo I (es decir, $f_X(x) = 0$ fuera de I)
- Consideremos $Y = g(X), X = h(Y)$

Funciones de una V.A.

- Consideremos ahora el caso de funciones monótonas
- $g(t)$ es monótona y $h(t)$ es su inversa, asumamos que existe dh/dt
- Sea X una variable aleatoria contenida en un intervalo I (es decir, $f_X(x) = 0$ fuera de I)
- Consideremos $Y = g(X), X = h(Y)$
- $f_Y(y) = f_X(h(y))|\frac{dh}{dy}(y)|$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

- Consideremos ahora el caso de funciones monótonas
- $g(t)$ es monótona y $h(t)$ es su inversa, asumamos que existe dh/dt
- Sea X una variable aleatoria contenida en un intervalo I (es decir, $f_X(x) = 0$ fuera de I)
- Consideremos $Y = g(X), X = h(Y)$
- $f_Y(y) = f_X(h(y))|\frac{dh}{dy}(y)|$
- La expresión anterior se verifica a partir de la definición de cdf de Y, y diferenciando usando la regla de la cadena

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Funciones de una V.A.

- Consideremos ahora el caso de funciones monótonas
- $g(t)$ es monótona y $h(t)$ es su inversa, asumamos que existe dh/dt
- Sea X una variable aleatoria contenida en un intervalo I (es decir, $f_X(x) = 0$ fuera de I)
- Consideremos $Y = g(X), X = h(Y)$
- $f_Y(y) = f_X(h(y))|\frac{dh}{dy}(y)|$
- La expresión anterior se verifica a partir de la definición de cdf de Y, y diferenciando usando la regla de la cadena
- Para funciones monótonamente crecientes $g(x)$, planteamos

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Funciones de una V.A.

- Consideremos ahora el caso de funciones monótonas
- $g(t)$ es monótona y $h(t)$ es su inversa, asumamos que existe dh/dt
- Sea X una variable aleatoria contenida en un intervalo I (es decir, $f_X(x) = 0$ fuera de I)
- Consideremos $Y = g(X), X = h(Y)$
- $f_Y(y) = f_X(h(y))|\frac{dh}{dy}(y)|$
- La expresión anterior se verifica a partir de la definición de cdf de Y, y diferenciando usando la regla de la cadena
- Para funciones monótonamente crecientes $g(x)$, planteamos
- $F_Y(y) = P(g(X) \leq y) = P(X \leq h(y)) = F_X(h(y))$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Funciones de una V.A.

- Consideremos ahora el caso de funciones monótonas
- $g(t)$ es monótona y $h(t)$ es su inversa, asumamos que existe dh/dt
- Sea X una variable aleatoria contenida en un intervalo I (es decir, $f_X(x) = 0$ fuera de I)
- Consideremos $Y = g(X), X = h(Y)$
- $f_Y(y) = f_X(h(y))|\frac{dh}{dy}(y)|$
- La expresión anterior se verifica a partir de la definición de cdf de Y, y diferenciando usando la regla de la cadena
- Para funciones monótonamente crecientes $g(x)$, planteamos
- $F_Y(y) = P(g(X) \leq y) = P(X \leq h(y)) = F_X(h(y))$
- Para funciones monótonamente crecientes planteamos: $F_Y(y) = P(X \geq h(y)) = 1 - F_X(h(y))$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Funciones de una V.A.

- Consideremos ahora el caso de funciones monótonas
- $g(t)$ es monótona y $h(t)$ es su inversa, asumamos que existe dh/dt
- Sea X una variable aleatoria contenida en un intervalo I (es decir, $f_X(x) = 0$ fuera de I)
- Consideremos $Y = g(X), X = h(Y)$
- $f_Y(y) = f_X(h(y))|\frac{dh}{dy}(y)|$
- La expresión anterior se verifica a partir de la definición de cdf de Y, y diferenciando usando la regla de la cadena
- Para funciones monótonamente crecientes $g(x)$, planteamos
- $F_Y(y) = P(g(X) \leq y) = P(X \leq h(y)) = F_X(h(y))$
- Para funciones monótonamente crecientes planteamos: $F_Y(y) = P(X \geq h(y)) = 1 - F_X(h(y))$
- $f_Y(y) = \frac{dF_Y}{dy}(y)$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Cuadrado de una V. A.

- Sea $X \sim U[0,1]$, encontrar la pdf de $Y = X^2$

- El objetivo es generar una cdf a partir de una variable uniforme cero-uno

Método de la Transformada Inversa

- El objetivo es generar una cdf a partir de una variable uniforme cero-uno
- Sea $X \sim U[0,1]$. Sea $F(y)$ una función monótona biyectiva, la cual usaremos como cdf objetivo

Método de la Transformada Inversa

- El objetivo es generar una cdf a partir de una variable uniforme cero-uno
- Sea $X \sim U[0,1]$. Sea $F(y)$ una función monótona biyectiva, la cual usaremos como cdf objetivo
- $F^{-1}(x) = \inf\{y|F(y) \geq x\} \quad (0 < x < 1)$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Método de la Transformada Inversa

- El objetivo es generar una cdf a partir de una variable uniforme cero-uno
- Sea $X \sim U[0,1]$. Sea $F(y)$ una función monótona biyectiva, la cual usaremos como cdf objetivo
- $F^{-1}(x) = \inf\{y | F(y) \geq x\} \quad (0 < x < 1)$
- Si $Y = F^{-1}(X)$, entonces la cdf de Y es $F(y)$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Método de la Transformada Inversa

- El objetivo es generar una cdf a partir de una variable uniforme cero-uno
- Sea $X \sim U[0,1]$. Sea $F(y)$ una función monótona biyectiva, la cual usaremos como cdf objetivo
- $F^{-1}(x) = \inf\{y | F(y) \geq x\} \quad (0 < x < 1)$
- Si $Y = F^{-1}(X)$, entonces la cdf de Y es $F(y)$
- Demostración:
  $F_Y(y) = P(Y \leq y) = P(F^{-1}(X) \leq y) = P(X \leq F(y)) = F(y)$

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Método de la Transformada Inversa

- El objetivo es generar una cdf a partir de una variable uniforme cero-uno
- Sea $X \sim U[0,1]$. Sea $F(y)$ una función monótona biyectiva, la cual usaremos como cdf objetivo
- $F^{-1}(x) = \inf\{y | F(y) \geq x\} \quad (0 < x < 1)$
- Si $Y = F^{-1}(X)$, entonces la cdf de Y es $F(y)$
- Demostración: $F_Y(y) = P(Y \leq y) = P(F^{-1}(X) \leq y) = P(X \leq F(y)) = F(y)$
- También se puede demostrar lo anterior usando la regla de la cadena: $f_Y(y) = f_X(h(y)) |\frac{dh}{dy}(y)|$

Distribución Gaussiana Bivariable

- Sean U, V dos v.a. normales independientes

Distribución Gaussiana Bivariable

- Sean U, V dos v.a. normales independientes
- Sean $X = aU + bV$, $Y = cU + dV$

FACULTAD
DE INGENIERIA
Universidad de Buenos Aires

Distribución Gaussiana Bivariable

- Sean U, V dos v.a. normales independientes
- Sean $X = aU + bV$, $Y = cU + dV$
- Entonces, la distribución conjunta de X, Y es una distribución normal bivariable

Distribución Gaussiana Bivariable

- Sean U, V dos v.a. normales independientes
- Sean $X = aU + bV$, $Y = cU + dV$
- Entonces, la distribución conjunta de X, Y es una distribución normal bivariable
- Agrupamos las v.a. en el vector $\mathrm{x} = [X\ Y]^T$

Distribución Gaussiana Bivariable

- Sean U, V dos v.a. normales independientes
- Sean $X = aU + bV$, $Y = cU + dV$
- Entonces, la distribución conjunta de X, Y es una distribución normal bivariable
- Agrupamos las v.a. en el vector $\mathrm{x} = [X\ Y]^T$
- El vector de medias es $\mu = [\mu_X\ \mu_Y]^T$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Gaussiana Bivariable

- Sean U, V dos v.a. normales independientes
- Sean $X = aU + bV$, $Y = cU + dV$
- Entonces, la distribución conjunta de X, Y es una distribución normal bivariable
- Agrupamos las v.a. en el vector $\mathrm{x} = [X\ Y]^T$
- El vector de medias es $\mu = [\mu_X\ \mu_Y]^T$
- La matriz de covarianza es $\Sigma = \begin{bmatrix} \sigma_X^2 & \rho_{XY}\sigma_X\sigma_Y \\ \rho_{XY}\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Gaussiana Bivariable

- Sean U, V dos v.a. normales independientes
- Sean $X = aU + bV$, $Y = cU + dV$
- Entonces, la distribución conjunta de X, Y es una distribución normal bivariable
- Agrupamos las v.a. en el vector $\mathrm{x} = [X\ Y]^T$
- El vector de medias es $\mu = [\mu_X\ \mu_Y]^T$
- La matriz de covarianza es $\Sigma = \begin{bmatrix} \sigma_X^2 & \rho_{XY}\sigma_X\sigma_Y \\ \rho_{XY}\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$
- $f_{XY}(\mathrm{x}) = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\mathrm{x}-\mu)^T \cdot \Sigma^{-1} \cdot (\mathrm{x}-\mu)\right]$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Gaussiana Bivariable

- Sean U, V dos v.a. normales independientes
- Sean $X = aU + bV$, $Y = cU + dV$
- Entonces, la distribución conjunta de X, Y es una distribución normal bivariable
- Agrupamos las v.a. en el vector $\mathrm{x} = [X\ Y]^T$
- El vector de medias es $\mu = [\mu_X\ \mu_Y]^T$
- La matriz de covarianza es $\Sigma = \begin{bmatrix} \sigma_X^2 & \rho_{XY}\sigma_X\sigma_Y \\ \rho_{XY}\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$
- $f_{XY}(\mathrm{x}) = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\mathrm{x}-\mu)^T \cdot \Sigma^{-1} \cdot (\mathrm{x}-\mu)\right]$
- Si σ_X^2, σ_Y^2 son no nulas y $|\rho_{XY}| < 1$ entonces Σ es no-singular, simétrica y definida positiva

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Gaussiana Bivariable

- Sean U, V dos v.a. normales independientes
- Sean $X = aU + bV$, $Y = cU + dV$
- Entonces, la distribución conjunta de X, Y es una distribución normal bivariable
- Agrupamos las v.a. en el vector $\mathrm{x} = [X\ Y]^T$
- El vector de medias es $\mu = [\mu_X\ \mu_Y]^T$
- La matriz de covarianza es $\Sigma = \begin{bmatrix} \sigma_X^2 & \rho_{XY}\sigma_X\sigma_Y \\ \rho_{XY}\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$
- $f_{XY}(\mathrm{x}) = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\mathrm{x}-\mu)^T \cdot \Sigma^{-1} \cdot (\mathrm{x}-\mu)\right]$
- Si σ_X^2, σ_Y^2 son no nulas y $|\rho_{XY}| < 1$ entonces Σ es no-singular, simétrica y definida positiva
- Es decir, $x^T \Sigma x > 0$, $\mathrm{eig}(\Sigma) > 0$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Gaussiana Bivariable

- La pdf conjunta de la distribución normal bivariable se puede expresar como:

Distribución Gaussiana Bivariable

- La pdf conjunta de la distribución normal bivariable se puede expresar como:
- $f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \exp\left[-\frac{1}{2(1-\rho_{XY}^2)}(z_X^2 - 2\rho_{XY}x_X z_Y + z_Y^2)\right]$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Gaussiana Bivariable

- La pdf conjunta de la distribución normal bivariable se puede expresar como:
- $f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \exp\left[-\frac{1}{2(1-\rho_{XY}^2)}(z_X^2 - 2\rho_{XY} x_X z_Y + z_Y^2)\right]$
- $z_X = (x - \mu_X)/\sigma_X$, $z_Y = (y - \mu_Y)/\sigma_Y$ son variables normalizadas

Distribución Gaussiana Bivariable

- La pdf conjunta de la distribución normal bivariable se puede expresar como:
- $f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \exp\left[-\frac{1}{2(1-\rho_{XY}^2)}(z_X^2 - 2\rho_{XY}x_X z_Y + z_Y^2)\right]$
- $z_X = (x-\mu_X)/\sigma_X$, $z_Y = (y-\mu_Y)/\sigma_Y$ son variables normalizadas
- Sea el conjunto de nivel $z_X^2 - 2\rho_{XY}x_X z_Y + z_Y^2 = \text{const.}$ ¿Qué forma toma en el plano $z_X z_Y$ esta ecuación?

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Gaussiana Bivariable

- La pdf conjunta de la distribución normal bivariable se puede expresar como:
- $f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \exp\left[-\frac{1}{2(1-\rho_{XY}^2)}(z_X^2 - 2\rho_{XY}x_X z_Y + z_Y^2)\right]$
- $z_X = (x-\mu_X)/\sigma_X$, $z_Y = (y-\mu_Y)/\sigma_Y$ son variables normalizadas
- Sea el conjunto de nivel $z_X^2 - 2\rho_{XY}x_X z_Y + z_Y^2 = \text{const.}$ ¿Qué forma toma en el plano $z_X z_Y$ esta ecuación?
- ¿Cómo varía la elipse anterior en función de $|\rho_{XY}|$?

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

- La pdf conjunta de la distribución normal bivariable se puede expresar como:
- $f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \exp\left[-\frac{1}{2(1-\rho_{XY}^2)}(z_X^2 - 2\rho_{XY}x_X z_Y + z_Y^2)\right]$
- $z_X = (x-\mu_X)/\sigma_X$, $z_Y = (y-\mu_Y)/\sigma_Y$ son variables normalizadas
- Sea el conjunto de nivel $z_X^2 - 2\rho_{XY}x_X z_Y + z_Y^2 = \text{const.}$ ¿Qué forma toma en el plano $z_X z_Y$ esta ecuación?
- ¿Cómo varía la elipse anterior en función de $|\rho_{XY}|$?
- ¿Qué significa que $|\rho_{XY}| = 1$?

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

- En general, para un vector $\mathrm{x} = [X_1 \, X_2 \ldots X_n]$ de v.a. multivariable normal de dimensión n tiene una pdf conjunta dada por

Distribución Gaussiana Multivariable

- En general, para un vector $\mathrm{x} = [X_1\, X_2 \ldots X_n]$ de v.a. multivariable normal de dimensión n tiene una pdf conjunta dada por
- $f_{XY}(\mathrm{x}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\mathrm{x}-\mu)^T \cdot \Sigma^{-1} \cdot (\mathrm{x}-\mu)\right]$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Gaussiana Multivariable

- En general, para un vector $\mathrm{x} = [X_1 \, X_2 \ldots X_n]$ de v.a. multivariable normal de dimensión n tiene una pdf conjunta dada por
- $f_{XY}(\mathrm{x}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\mathrm{x}-\mu)^T \cdot \Sigma^{-1} \cdot (\mathrm{x}-\mu)\right]$
- El vector de medias es ahora $\mu = [\mu_{X1} \, \mu_{X2} \ldots \mu_{Xn}]^T$

Distribución Gaussiana Multivariable

- En general, para un vector $\mathrm{x} = [X_1\, X_2 \ldots X_n]$ de v.a. multivariable normal de dimensión n tiene una pdf conjunta dada por
- $f_{XY}(\mathrm{x}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\mathrm{x}-\mu)^T \cdot \Sigma^{-1} \cdot (\mathrm{x}-\mu)\right]$
- El vector de medias es ahora $\mu = [\mu_{X1}\, \mu_{X2} \ldots \mu_{Xn}]^T$
- La matriz de covarianza es $\Sigma = \begin{bmatrix} \mathrm{cov}[X_1, X_1] & \ldots & \mathrm{cov}[X_1, X_n] \\ \vdots & \ddots & \vdots \\ \mathrm{cov}[X_n, X_1] & \ldots & \mathrm{cov}[X_n, X_n] \end{bmatrix}$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Distribución Gaussiana Multivariable

- En general, para un vector $\mathrm{x} = [X_1\, X_2 \dots X_n]$ de v.a. multivariable normal de dimensión n tiene una pdf conjunta dada por
- $f_{XY}(\mathrm{x}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\mathrm{x}-\mu)^T \cdot \Sigma^{-1} \cdot (\mathrm{x}-\mu)\right]$
- El vector de medias es ahora $\mu = [\mu_{X1}\, \mu_{X2} \dots \mu_{Xn}]^T$
- La matriz de covarianza es $\Sigma = \begin{bmatrix} \mathrm{cov}[X_1, X_1] & \dots & \mathrm{cov}[X_1, X_n] \\ \vdots & \ddots & \vdots \\ \mathrm{cov}[X_n, X_1] & \dots & \mathrm{cov}[X_n, X_n] \end{bmatrix}$
- ¿Cómo queda la expresión de $f_{XY}(\mathrm{x})$ cuando todas las variables son independientes?

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Proyección de una V.A. Gaussiana Multivariable

- Sea $\mathrm{x} = [X_1\, X_2 \dots X_n]$ un vector de v.a. multivariable normal de dimensión n

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Proyección de una V.A. Gaussiana Multivariable

- Sea $\mathrm{x} = [X_1\, X_2 \dots X_n]$ un vector de v.a. multivariable normal de dimensión n
- Sea $\mathrm{w} \in \mathcal{R}^n$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Proyección de una V.A. Gaussiana Multivariable

- Sea $\mathrm{x} = [X_1\, X_2 \ldots X_n]$ un vector de v.a. multivariable normal de dimensión n
- Sea $\mathrm{w} \in \mathcal{R}^n$
- $z = \mathrm{w}^T \cdot \mathrm{x}$ es la proyección de x sobre w

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Proyección de una V.A. Gaussiana Multivariable

- Sea $\mathrm{x} = [X_1\, X_2 \ldots X_n]$ un vector de v.a. multivariable normal de dimensión n
- Sea $\mathrm{w} \in \mathcal{R}^n$
- $z = \mathrm{w}^T \cdot \mathrm{x}$ es la proyección de x sobre w
- La pdf de z es $\mathcal{N}(\mathrm{w}^T \cdot \mu, \mathrm{w}^T \cdot \Sigma \cdot \mathrm{w})$

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Ejercicios de Correlación

- Dada $X = aU + bV$, $Y = cU + dV$, con $U, V \sim \mathcal{N}(\mu, \sigma^2)$ generar diferentes valores (en particular $b = 0, c = 0$ y $a = d = 1$) para variar $|\rho_{XY}|$ en Octave

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Ejercicios de Correlación

- Dada $X = aU + bV$, $Y = cU + dV$, con $U, V \sim \mathcal{N}(\mu, \sigma^2)$ generar diferentes valores (en particular $b = 0, c = 0$ y $a = d = 1$) para variar $|\rho_{XY}|$ en Octave
- Graficar X vs. Y y analizar

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Ejercicio 1

(a) Sean X, Y dos v.a. distribuidas normalmente, con distribucion conjunta normal bivariable f_{XY}. Demostrar que X, Y son independientes si y sólo si son v.a. descorrelacionadas (pista: usar $\rho = 0$ en f_{XY})

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Ejercicio 2

(a) Sean X, Y dos v.a. normales de media cero correlacionadas con $\rho_{XY} = 0.5$. Se tiene que $\mathrm{var}[X] = 2\mathrm{var}[Y] = 1$.

(b) Encontrar la expresión de la pdf conjunta de X, Y

(c) Proyectar $\mathrm{x} = [X\ Y]^T$ sobre $\mathrm{w} = [1\,1]^T$ y encontrar la pdf de la proyección.

FACULTAD DE INGENIERIA Universidad de Buenos Aires

Ejercicio 3

(a) Usando el método de la transformada inversa, calcular la pdf de la distribución exponencial $f_Y(y) = \lambda e^{-\lambda y}$ para $y \geq 0$, o $f_Y(y) = 0$ si $y < 0$. Usar un valor arbitrario de $\lambda > 0$

(b) Simular en Octave

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Ejercicio 4

1. Sea $\mathrm{x} = (X_1, X_2) \in \mathcal{R}^2$ un vector de v.a.
2. Definamos $\mathrm{y} = \mathrm{T}(\mathrm{x})$, con $\mathrm{T} = (g_1(x_1, x_2), g_2(x_1, x_2)) \in \mathcal{R}^2$, con g_1, g_2 diferenciables
3. Encontrar la expresión de la pdf del vector transformado y en función de $f_{X_1 X_2}$
4. Resolver la pdf de y para el caso del vector x del Ejercicio 2 con $T = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ y simular en Octave

FACULTAD DE INGENIERIA
Universidad de Buenos Aires

Bibliografía I

E. Alpaydin, *Introduction to machine learning*.
MIT press, 2020.

"MIT - Convolution, 18.031, Haynes Miller and Jeremy Orloff."
http://math.mit.edu/~hrm/18.031/convolution.pdf.
Accessed: 2020-05-15.

"MIT - Lecture Notes Course 6.041-6.431, Fall 2000."
https://vfu.bg/en/e-Learning/Math--Bertsekas_Tsitsiklis_Introduction_to_probability.pdf.
Accessed: 2020-05-15.

FACULTAD DE INGENIERIA
Universidad de Buenos Aires