Vectores Aleatorios

Cecilia Galarza

Procesos Estocásticos
Facultad de Ingeniería, Universidad de Buenos Aires

2do cuatrimestre 2024

Esperanza

El operador esperanza

Sea **X** un vector aleatorio *n*-dimensional. La media o esperanza de **X** es vector de las esperanzas de sus componentes:

$$\mathbb{E}[\mathbf{X}] = \boldsymbol{\mu}_{\mathbf{X}} = \begin{bmatrix} \mathbb{E}[X_1] \\ \vdots \\ \mathbb{E}[X_n] \end{bmatrix}.$$

El operador esperanza: Propiedades

- El operador $\mathbb{E}[]$ es un operador lineal, luego $\forall \mathbf{A} \in \mathbb{R}^{m\times n}$ y $\forall \mathbf{b} \in \mathbb{R}^m$ tenemos:
$$\mathbb{E}[\mathbf{AX}+\mathbf{b}] = \mathbf{A}\mathbb{E}[\mathbf{X}]+\mathbf{b}.$$
- En forma general, si $\boldsymbol{g}$ es una función $\boldsymbol{g}:\mathbb{R}^n \to \mathbb{R}^m$, tenemos
$$\mathbb{E}[\boldsymbol{g}(\mathbf{X})] = \begin{bmatrix} \mathbb{E}[g_1(\mathbf{X})] \\ \vdots \\ \mathbb{E}[g_m(\mathbf{X})] \end{bmatrix}.$$

La esperanza condicional

Sea un evento A. Luego, definimos el operador $\mathbb{E}[\cdot|A]$

$$\mathbb{E}[\mathbf{X}|A] = \int_{\mathbb{R}^n} \boldsymbol{x} \, f_{\mathbf{X}|A}(\boldsymbol{x}) \, d\boldsymbol{x}.$$

- En particular, si $A = \{X_1 = x_1\}$

$$\mathbb{E}[X_2|X_1 = x_1] = \int_{\mathbb{R}} x_2 \, f_{X_2}(x_2|X_1 = x_1) \, dx_2 = g(x_1).$$

  Vemos que $\mathbb{E}[X_2|X_1 = x_1]$ es una función de x_1. Si consideramos todas las posibles realizaciones de x_1 es decir, consideramos la aleatoriedad de X_1, $g(X_1)$ es una VA.

Ley de la esperanza total

Tenemos la V.A. $g(X_1) = \mathbb{E}[X_2|X_1] = \int_{\mathbb{R}} x_2 f_{X_2|X_1}(x_2|X_1)\, dx_2$. Luego,

$$
\begin{aligned}
\mathbb{E}[g(X_1)] = \mathbb{E}\Big[\mathbb{E}[X_2|X_1]\Big] &= \int_{\mathbb{R}} g(x_1)\, f_{X_1}(x_1)\, dx_1 \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} x_2 \; \underbrace{f_{X_2|X_1}(x_2|X_1 = x_1) \;\; f_{X_1}(x_1)}_{f_{X_1,X_2}(x_1,x_2)} \; dx_2 dx_1 \\
&= \int_{\mathbb{R}} x_2 \underbrace{\int_{\mathbb{R}} f_{X_1,X_2}(x_1, x_2)\, dx_1}_{f_{X_2}(x_2)} \, dx_2 = \mathbb{E}[X_2]
\end{aligned}
$$

Propiedades de la esperanza condicional

- *Ley de esperanza total.* $\mathbb{E}[\mathbb{E}[\mathbf{X}|\mathbf{Y}]] = \mathbb{E}[\mathbf{X}]$.
- *Esperanza condicional de una constante.* $\mathbb{E}[\boldsymbol{a}|\mathbf{X}] = \boldsymbol{a}$.
- *Linealidad.* $\mathbb{E}[a\mathbf{Y} + b\mathbf{Z}|\mathbf{X}] = a\mathbb{E}[\mathbf{Y}|\mathbf{X}] + b\mathbb{E}[\mathbf{Z}|\mathbf{X}]$.
- *Esperanza condicional de una función.* $\mathbb{E}[\boldsymbol{g}(\mathbf{X})\boldsymbol{h}(\mathbf{Y})|\mathbf{X}] = \boldsymbol{g}(\mathbf{X})\mathbb{E}[\boldsymbol{h}(\mathbf{Y})|\mathbf{X}]$.
- *Redundancia en condiciones.* $\mathbb{E}[\mathbf{Y}|\mathbf{X}, \boldsymbol{g}(\mathbf{X})] = \mathbb{E}[\mathbf{Y}|\mathbf{X}]$.

Esperanza condicional de VA independientes

.

Si $\mathbf{X}$ y $\mathbf{Y}$ son independientes, entonces $\mathbb{E}[\mathbf{X}|\mathbf{Y}] = \mathbb{E}[\mathbf{X}]$.

Esta propiedad es una consecuencia directa de

$$f_{\mathbf{X},\mathbf{Y}}(\boldsymbol{x},\boldsymbol{y}) = f_{\mathbf{X}}(\boldsymbol{x}) f_{\mathbf{Y}}(\boldsymbol{y}) \iff \mathbf{X}, \mathbf{Y} \text{ independientes}$$

Función Característica

Función característica de una VA

En el análisis de señales y sistemas, la transformada de Fourier es una herramienta que permite simplificar ciertos análisis. Podemos incorporar esta herramienta también al analizar VA.

Para ello, definimos la *función característica* de una variable aleatoria:

$$\Phi_X : \mathbb{R} \longrightarrow \mathbb{C}$$
$$\Phi_X(\omega) = \mathbb{E}[e^{\jmath\omega X}],$$

donde $\jmath = \sqrt{-1}$. Considerando la PDF generalizada, tenemos

$$\Phi_X(\omega) = \int_{-\infty}^{\infty} e^{\jmath\omega x} f_X(x) dx.$$

Par transformado

- $\Phi_X(\omega)$ y $f_X(x)$ forman un par Fourier transformado:

$$\Phi_X(\omega) = \int_{-\infty}^{\infty} e^{\jmath\omega x} f_X(x)dx = \mathcal{F}(f_X)(-\omega).$$

- Utilizando la transformada inversa obtenemos:

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-j\omega x}\Phi_X(\omega)d\omega = \mathcal{F}^{-1}(\Phi_X)(-x), \qquad x \in \mathbb{R}.$$

Par transformado

Si X es una VAD que toma valores en $D = \{y_1, y_2, \cdots\}$, tenemos

$$f_X(x) = \sum_{y \in D} p_X(y)\delta(x - y).$$

Luego,

$$\Phi_X(\omega) = \int_{-\infty}^{\infty} e^{\jmath\omega x} \sum_{y \in D} p_X(y)\delta(x-y)dx = \sum_{y \in D} p_X(y)e^{\jmath\omega y} = \mathcal{F}_d(p_X)(-\omega),$$

donde $\mathcal{F}_d(p_X)$ es la transformada de Fourier de tiempo discreto de la PMF p_X. Claramente,

$$p_X(x) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-j\omega x}\Phi_X(\omega)d\omega = \mathcal{F}_d^{-1}(\Phi_X)(-x), \qquad x \in D.$$

En el caso de una VAD, $\Phi_X(\omega) = \Phi_X(\omega + 2\pi)$ por propiedades de $\mathcal{F}_d$.

Función característica de un vector aleatorio

El concepto de función característica se puede ampliar a vectores aleatorios. Sea **X** un VeA *n*-dimensional. Se define la función característica del siguiente modo:

$$\Phi_{\mathbf{X}} : \mathbb{R}^n \to \mathbb{C}$$

$$\Phi_{\mathbf{X}}(\omega) = \mathbb{E}[e^{\jmath\omega^T\mathbf{X}}]$$

- Como antes, la FC y la PDF de **X** forman un par transformado:

$$\Phi_{\mathbf{X}}(\omega) = \mathcal{F}\{f_{\mathbf{X}}\}(-\omega) \qquad , \qquad f_{\mathbf{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \Phi_{\mathbf{X}}(\omega)\, e^{-j\omega^T\boldsymbol{x}}\, d\omega.$$

- En este caso, la transformada de Fourier es multidimensional. Esto queda explícito al considerar el vector de *n* dimensiones ω.

Propiedades de la función característica

Sea $\mathbf{X} \in \mathbb{R}^n$ un vector aleatorio. Si $n = 1$ es una variable aleatoria.

- $\Phi_{\mathbf{X}}(0) = 1$
- La FC siempre existe y es máxima en el origen:

$$|\Phi_{\mathbf{X}}(\omega)| = \left| \int_{-\infty}^{\infty} e^{\jmath \omega^T \boldsymbol{x}} f_{\mathbf{X}}(\boldsymbol{x}) d\boldsymbol{x} \right| \leq \int_{-\infty}^{\infty} \left| e^{\jmath \omega^T \boldsymbol{x}} f_{\mathbf{X}}(\boldsymbol{x}) \right| d\boldsymbol{x}$$
$$= \int_{-\infty}^{\infty} f_{\mathbf{X}}(\boldsymbol{x}) d\boldsymbol{x} = \Phi_{\mathbf{X}}(0).$$

- La FC de $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$ donde $\mathbf{A} \in \mathbb{R}^{m \times n}$, y $\mathbf{b} \in \mathbb{R}^m$ es:

$$\Phi_{\mathbf{Y}}(\omega) = \mathbb{E}[e^{\jmath \omega^T \mathbf{Y}}] = \mathbb{E}[e^{\jmath \omega^T (\mathbf{AX}+\mathbf{b})}]$$
$$= e^{\jmath \omega^T \mathbf{b}} \mathbb{E}[e^{\jmath \omega^T \mathbf{AX}}] = e^{\jmath \omega^T \mathbf{b}} \Phi_{\mathbf{X}}(\mathbf{A}^T \omega).$$

- Si las componentes de $\mathbf{X}$ son independientes

$$\Phi_{\mathbf{X}}(\omega) = \mathbb{E}[e^{\jmath \omega^T \mathbf{X}}] = \mathbb{E}[e^{\jmath \omega_1 X_1} \dots e^{\jmath \omega_n X_n}] = \prod_{i=1}^{n} \mathbb{E}[e^{\jmath \omega_i X_i}] = \prod_{i=1}^{n} \Phi_{X_i}(\omega_i).$$

Momentos de un vector aleatorio

Momentos de una variable aleatoria

Una VA se define con su CDF. Una alternativa es definir todos sus momentos.

- El *momento* de orden n de una VA X lo denotamos m_n y se define como

$$m_n = \mathbb{E}[X^n] = \int_{-\infty}^{\infty} x^n \, f_X(x) dx.$$

Momentos de componentes de un vector aleatorio

Podemos generalizar los conceptos de momentos a las componentes de un vector aleatorio.

- Los *momentos* se definen como

$$m_{i_1,\ldots,i_n} = \mathbb{E}[X_1^{i_1} \ldots X_n^{i_n}], \qquad i_1, \ldots, i_n \in \mathbb{N}_0.$$

  Notar que esto incluye los momentos de cada una de las VAs $X_1, \ldots, X_n$ por separado. Por ejemplo:

$$m_{1,0,\ldots,0} = \mathbb{E}[X_1] = \mu_{X_1}.$$

Teorema de los momentos

Consideramos primero $n = 1$.

- Derivando k veces Φ_X con respecto a ω obtenemos

$$\Phi_X^{(k)}(\omega) = \frac{\partial^k \Phi_X(\omega)}{\partial \omega^k} = \mathbb{E}[(jX)^k e^{\jmath\omega X}].$$

- Evaluamos esta expresión en $\omega = 0$ y reordenamos:

$$\mathbb{E}[X^k] = \frac{1}{j^k} \Phi_X^{(k)}(0).$$

Este resultado nos permite obtener los momentos a partir de las derivadas de la FC, lo cual a veces resulta más sencillo que realizar la integración directa.

Teorema de los momentos

La generalización al caso vectorial es la siguiente:

$$\mathbb{E}[X_1^{i_1} \dots X_n^{i_n}] = \frac{1}{j^{i_1+\dots+i_n}} \frac{\partial^{i_1+\dots+i_n} \Phi_{\mathbf{X}}(\mathbf{0})}{\partial \omega_1^{i_1} \dots \partial \omega_n^{i_n}}.$$

La demostración es completamente análoga y queda como ejercicio.

Transformaciones de vectores aleatorios

Transformaciones de vectores aleatorios

Sea $\mathbf{X}, \mathbf{Y}$ dos VeAs tal que $\mathbf{Y} = \boldsymbol{g}(\mathbf{X})$ donde

$$\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^m$$

.
En general, la CDF de $\mathbf{Y}$ es

$$F_{\mathbf{Y}}(\boldsymbol{y}) = \mathbb{P}(\mathbf{Y} \leq \boldsymbol{y}) = \mathbb{P}(\boldsymbol{g}(\mathbf{X}) \leq \boldsymbol{y}) = \int_{R_{\boldsymbol{y}}} f_{\mathbf{X}}(\boldsymbol{x}) d\boldsymbol{x},$$

donde $R_{\boldsymbol{y}} = \{\boldsymbol{x} : \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{y}\}$. Podemos obtener la PDF de $\mathbf{Y}$ (posiblemente generalizada) como

$$f_{\mathbf{Y}}(\boldsymbol{y}) = \frac{\partial^m F_{\mathbf{Y}}(\boldsymbol{y})}{\partial y_1 \dots \partial y_m}.$$

Transformaciones de vectores aleatorios

Una alternativa al procedimiento anterior, es hallar la FC de **Y** y luego tomar la antitransformada de Fourier para hallar $f_{\mathbf{Y}}(\boldsymbol{y})$. Analizaremos este procedimiento en el caso de una función afín.

Distribución de una combinación lineal de VA

Sea $Y = \sum_{i=1}^{n} a_i X_i = \boldsymbol{a}^T \mathbf{X}$.

- Usando las propiedades de la FC, tenemos que

$$f_Y(y) = \mathcal{F}^{-1}\left[\Phi_Y(\omega)\right](-y) = \mathcal{F}^{-1}\left[\Phi_{\mathbf{X}}(\boldsymbol{a}\omega)\right]$$

- Si las componentes de **X** son independientes tenemos que

$$\Phi_{\mathbf{X}}(\boldsymbol{a}\omega) = \prod_{i=1}^{n} \Phi_{X_i}(a_i\omega).$$

Distribución de una combinación lineal de VA

- Por la propiedad de escalado de Fourier,

$$\mathcal{F}^{-1}\left(\Phi_X(a\omega)\right)(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)$$

- Luego, por el teorema de convolución de Fourier,

$$\begin{aligned} f_Y(y) &= \mathcal{F}^{-1}\left[\prod_{i=1}^{n} \Phi_{X_i}(a_i\omega)\right] \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \left(\prod_{i=1}^{n} \Phi_{X_i}(a_i\omega)\right) e^{-j\omega y}\, d\omega \\ &= \frac{1}{|a_1|} f_{X_1}\left(\frac{y}{a_1}\right) * \ldots * \frac{1}{|a_n|} f_{X_n}\left(\frac{y}{a_n}\right). \end{aligned}$$

Suma de V.A independientes

Sean $X_1, \cdots X_n$ independientes

Si $Y = \sum_{i=1}^{n} X_i$

$$f_Y(y) = f_{X_1}(y) * \ldots * f_{X_n}(y).$$

Transformación lineal o afín (caso biyectivo)

Sea $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$, con $\mathbf{A}$ una matriz cuadrada y no singular. Vimos que

$$\Phi_{\mathbf{Y}}(\omega) = \Phi_{\mathbf{X}}(\mathbf{A}^T\omega)e^{\jmath\omega^T\mathbf{b}}$$

Considerando el par transformado entre la FC y la PDF,

$$\begin{aligned} f_{\mathbf{Y}}(\boldsymbol{y}) &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \Phi_{\mathbf{X}}(\mathbf{A}^T\omega)\ e^{\jmath\omega^T(\mathbf{b}-\boldsymbol{y})}\ d\omega \\ &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \Phi_{\mathbf{X}}(\mathbf{A}^T\omega)\ e^{-j\omega^T\mathbf{A}\mathbf{A}^{-1}(\boldsymbol{y}-\mathbf{b})}\ d\omega \\ &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \Phi_{\mathbf{X}}(\mathbf{A}^T\omega)\ e^{-j(\mathbf{A}^T\omega)^T\mathbf{A}^{-1}(\boldsymbol{y}-\mathbf{b})}\ d\omega \end{aligned}$$

Transformación lineal o afín (caso biyectivo)

Haciendo un cambio de variable $\nu = \mathbf{A}^T\omega$, tenemos

$$\begin{aligned} f_{\mathbf{Y}}(\boldsymbol{y}) &= \frac{1}{(2\pi)^n}\frac{1}{|\det(\mathbf{A}^T)|}\int_{\mathbb{R}^n}\Phi_{\mathbf{X}}(\nu)\, e^{-j\nu^T\mathbf{A}^{-1}(\boldsymbol{y}-\mathbf{b})}\, d\nu \\ &= \frac{1}{|\det(\mathbf{A})|}\underbrace{\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\Phi_{\mathbf{X}}(\nu)\, e^{-j\nu^T\mathbf{A}^{-1}(\boldsymbol{y}-\mathbf{b})}\, d\nu}_{\mathcal{F}^{-1}(\Phi_X)(-\mathbf{A}^{-1}(\boldsymbol{y}-\mathbf{b}))} \\ &= \frac{f_{\mathbf{X}}(\mathbf{A}^{-1}(\boldsymbol{y}-\mathbf{b}))}{|\det(\mathbf{A})|}, \qquad \boldsymbol{y}\in\mathbb{R}^n. \end{aligned}$$

$\mathbf{Y} = \mathbf{AX} + \mathbf{b}$

$$f_{\mathbf{Y}}(\boldsymbol{y}) = \frac{f_{\mathbf{X}}(\mathbf{A}^{-1}(\boldsymbol{y}-\mathbf{b}))}{|\det(\mathbf{A})|}$$

Transformación biyectiva y suave

- Sea $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^n$ una función inversible
- Luego, $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{y}$ tiene una única solución.
- Sea $E(\boldsymbol{x}_0)$ un entorno alrededor de $\boldsymbol{x}_0$ y $F(\boldsymbol{x}_0) = \{\boldsymbol{g}(\boldsymbol{x}) : \boldsymbol{x} \in E(\boldsymbol{x}_0)\}$ el entorno alrededor de $\boldsymbol{g}(\boldsymbol{x}_0)$. Luego,

$$\mathbb{P}(\mathbf{X} \in E(\boldsymbol{x}_0)) = \mathbb{P}(\mathbf{Y} \in F(\boldsymbol{x}_0)) \Rightarrow \int_{E(\boldsymbol{x}_0)} f_{\mathbf{X}}(\boldsymbol{x}) d\boldsymbol{x} = \int_{F(\boldsymbol{x}_0)} f_{\mathbf{Y}}(\boldsymbol{y}) d\boldsymbol{y}.$$

Transformación biyectiva y suave

- Considerando un cambio de variable en la integral, obtenemos la siguiente relación entre las PDFs conjuntas:

$$f_{\mathbf{Y}}(\boldsymbol{y}) = \frac{f_{\mathbf{X}}(\boldsymbol{g}^{-1}(\boldsymbol{y}))}{|\det(J(\boldsymbol{g}^{-1}(\boldsymbol{y})))|},$$

donde J es el Jacobiano de la transformación dado por

$$J(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1}(\boldsymbol{x}) & \dots & \frac{\partial g_1}{\partial x_n}(\boldsymbol{x}) \\ \vdots & \vdots & \vdots \\ \frac{\partial g_n}{\partial x_1}(\boldsymbol{x}) & \dots & \frac{\partial g_n}{\partial x_n}(\boldsymbol{x}) \end{bmatrix}.$$

Transformación general suave

- Si $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^n$ es una función suave pero no inversible, dado un $\boldsymbol{y}$ la ecuación $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{y}$ tiene en general varias soluciones $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k$.
- En este caso, hay que sumar las contribuciones de cada una de las soluciones. Luego,

$$f_{\mathbf{Y}}(\boldsymbol{y}) = \sum_{i=1}^{k} \frac{f_{\mathbf{X}}(\boldsymbol{x}_i)}{|\det(J(\boldsymbol{x}_i))|},$$

donde $J(\boldsymbol{x}_i)$ es el Jacobiano de la transformación dado por

$$J(\boldsymbol{x}_i) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1}(\boldsymbol{x}_i) & \dots & \frac{\partial g_1}{\partial x_n}(\boldsymbol{x}_i) \\ \vdots & \vdots & \vdots \\ \frac{\partial g_n}{\partial x_1}(\boldsymbol{x}_i) & \dots & \frac{\partial g_n}{\partial x_n}(\boldsymbol{x}_i) \end{bmatrix}.$$

Transformación general suave

El método anterior también se puede aplicar a una transformación $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^m$ con $m \neq n$ agregando variables auxiliares X_i o Y_i de forma tal que se obtenga una nueva transformación $\boldsymbol{h} : \mathbb{R}^p \to \mathbb{R}^p$ donde $p = \max(n, m)$.