

Apunte de procesos estocásticos

Margarita Manterola

6 de junio de 2005

Resumen

Este apunte pretende ser un resumen de lo dado en clase (tanto práctica como teórica) en la materia Procesos Estocásticos (66.15) de la Facultad de Ingeniería de la UBA.

Falta: histograma, montecarlo, cambio de variable.

1. Repaso de matrices

1.1. Matriz diagonalizable

1.2. Matriz inversible

1.3. Matriz adjunta

1.4. Matriz simétrica

1.5. Matriz hermítica

1.6. Matriz unitaria

1.7. Matriz definida positiva

2. Repaso de probabilidad y estadística

2.1. Definiciones Básicas

Llamamos **señal (o variable) aleatoria** a una función que puede tomar valores aleatorios a lo largo del tiempo. Puede ser continua o discreta.

La **función de distribución de probabilidad** $F_X(x)$ es la probabilidad de que la señal aleatoria X sea menor que un determinado valor x .

$$F_X(x) = P(x \leq X) \quad (1)$$

Mientras que la **función de densidad de probabilidad** $f_X(x)$ está definida como:

$$f_X(x) = \frac{dF_X(x)}{dx} \quad (2)$$

Para la función de distribución de probabilidad no importa si la variable es continua o discreta, pero la función de densidad se define únicamente para variables continuas.

2.1.1. Ejemplos

Por ejemplo, una señal aleatoria binaria, que solamente puede tomar valores $-V$ y V , tendrá una función de probabilidad escalonada: 0 hasta $-V$, P_0 hasta V y 1 de ahí en adelante.

Mientras que su función de densidad de probabilidad estará constituida por dos impulsos, ubicados en $-V$ y V , y de peso P_0 y P_1 (siendo $P_1 = 1 - P_0$).

2.2. Momentos

El **valor medio** (o media) de una variable aleatoria está definido como:

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx \quad (3)$$

Y en general para cualquier función, se puede decir que:

$$E[g(x)] = \int_{-\infty}^{\infty} g(x) f_X(x) \quad (4)$$

Además, se define el momento de orden p :

$$m_P = E[x^P] = \int_{-\infty}^{\infty} x^P f_X(x) \quad (5)$$

De manera que $m_0 = 1$, $m_1 = E(X) = m_X$ y $m_2 = E(X^2)$.

Por otro lado, se define el momento **centrado** μ_P :

$$\mu_P = E[(x - E(X))^P] = \int_{-\infty}^{\infty} (x - m_X)^P f_X(x) \quad (6)$$

De manera que $\mu_0 = 1$, $\mu_1 = 0$ y $\mu_2 = \sigma_X^2$ (**varianza de X**).

Propiedad: $E(X^2) = m_X^2 + \sigma_X^2$.

Donde se puede equiparar la ecuación a los circuitos eléctricos y considerar $E(x^2)$ como la potencia total, m_X^2 como la potencia de corriente continua y σ_X^2 como la potencia de corriente alterna.

Demostración:

$$\sigma_X^2 = E[(X - m_X)^2] \quad (7)$$

$$= E[X^2 - 2Xm_X + m_X^2] \quad (8)$$

$$= E[X^2] - E[2Xm_X] + E[m_X^2] \quad (9)$$

$$= E[X^2] - E[2X]m_X + m_X^2 \quad (10)$$

$$= E[X^2] - 2m_Xm_X + m_X^2 \quad (11)$$

$$\sigma_X^2 = E[X^2] - m_X^2 \quad (12)$$

3. Vectores aleatorios

Puede ser necesario trabajar con varias señales aleatorias a la vez, lo que constituye un **vector aleatorio**. Por ejemplo, si se toman pares de valores de posición y velocidad de un móvil, o la altura y el peso de una persona, lo que constituiría las variables x y y .

También se puede tener un vector X de n componentes: $X = [X_1, X_2, \dots, X_n]$.

3.1. Funciones de distribución y densidad de probabilidad

En una distribución de dos variables, se puede tener una función de probabilidad conjunta:

$$F_{XY}(x, y) = P(X \leq x, Y \leq y) \quad (13)$$

Se trata de la probabilidad de que **tanto** X **como** Y sean menores a los valores x e y (es decir la intersección de las probabilidades).

De la misma manera para un vector de n variables:

$$F_{\vec{X}}(\vec{x}) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) \quad (14)$$

En ambos casos, la función de distribución de probabilidad es una función de \mathfrak{R}^N en \mathfrak{R} .

La función de densidad de probabilidades conjunta:

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y} \quad (15)$$

Propiedad: \vec{X} e \vec{Y} son variables aleatorias **independientes** si y sólo si $f_{XY}(x, y) = f_X(x)f_Y(y)$ y $P(A, B) = P(A)P(B)$.

Además, con X e Y independientes, si se tiene la función $Z = X + Y$, la función de densidad de probabilidad será:

$$f_Z(z) = f_X(z) * f_Y(z) = \int_{-\infty}^{\infty} f_Y(u)f_X(z - u)du \quad (16)$$

3.1.1. Probabilidad condicional

En los vectores aleatorios, cuando dos variables no son independientes entre sí, se puede obtener la **función de densidad de probabilidad condicional**.

$$f_{X/Y}(x/y) = \frac{f_{XY}(x, y)}{f_Y(y)} \quad (17)$$

Y también la **probabilidad condicional**

$$P(A/B) = \frac{P(A, B)}{P(B)} \quad (18)$$

Cuando las variables son independientes, $f_{X/Y}(x/y) = f_X(x)$ y $P(A/B) = P(A)$.

3.1.2. Probabilidad marginal

$$E[X_i] = \int_{-\infty}^{\infty} x_i f_{X_i}(x_i) \quad (19)$$

3.2. Valor medio

En general, para un vector aleatorio, el valor medio estará dado por el vector de los valores medios de cada una de las variables, constituyendo un vector de $n \times 1$.

$$E[X] = [E[X_0], E[X_1], \dots, E[X_n]] \quad (20)$$

3.2.1. Momentos de orden $p + q$

Para un vector de dos variables x y y :

$$m_{pQ} = E[x^p y^q] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f_{XY}(x, y) dx dy \quad (21)$$

De manera que se pueden obtener momentos de diversos órdenes.

De **orden 0**: $m_{00} = 1$.

De **primer orden**: $m_{10} = m_X$, $m_{01} = m_Y$.

De **segundo orden**: $m_{20} = E(X^2)$, $m_{02} = E(Y^2)$ y $m_{11} = E(XY) = R_{XY}$ (la correlación entre ambas variables).

Por otro lado, se definen los momentos centrados:

$$\mu_{pQ} = E[(x - m_X)^p (y - m_Y)^q] \quad (22)$$

De **orden 0**: $\mu_{00} = 1$.

De **primer orden**: $\mu_{10} = \mu_{01} = 0$.

De **segundo orden**: $\mu_{20} = \sigma_X^2$, $\mu_{02} = \sigma_Y^2$ y $\mu_{11} = cov(x, y) = C_{XY}$ (la covarianza entre ambas variables).

3.2.2. Propiedad

$$C_{XY} = E[(X - m_X)(Y - m_Y)] \quad (23)$$

$$= E[XY + m_X m_Y - X m_Y - Y m_X] \quad (24)$$

$$= E[XY] + m_X m_Y - m_X m_Y - m_Y m_X \quad (25)$$

$$C_{XY} = E[XY] - m_Y m_X = R_{XY} - m_X m_Y \quad (26)$$

3.3. Coeficiente de correlación

El **coeficiente de correlación** para una distribución de **dos** variables está dado por:

$$\rho = \frac{C_{XY}}{\sigma_X \sigma_Y} = \frac{E[(X - m_X)(Y - m_Y)]}{\sigma_X \sigma_Y} \quad (27)$$

Este coeficiente indica qué tan relacionadas están ambas variables, y su módulo debe ser menor o igual a 1.

Cuando $\rho = 0$, las variables están **descorrelacionadas**, es decir que dado un valor de una de ellas, se puede obtener cualquier valor de la otra; si se grafica una variable en función de la otra se obtendrá un gráfico circular.

Por otro lado, si $\rho = 0,8$, las variables están bastante relacionadas y el gráfico que se obtendrá será una elipse inclinada en $\pi/4$, si es $\rho = 0,95$, la elipse será mucho más delgada, y con $\rho = 1$, será una línea recta.

Además, con $\rho = -0,8$, se obtendrá una elipse similar a la de $\rho = 0,8$, pero formando un ángulo de $3\pi/4$.

3.3.1. Matriz de correlación

Se define la matriz de correlación de un vector aleatorio de n variables como:

$$R_X = E [\vec{X} \vec{X}^T] \quad (28)$$

Es decir que la matriz está constituida por los valores medios de los productos de $[X_i X_j]$. Además, $E [\vec{X}^T \vec{X}]$ es la traza de R_X .

3.3.2. Autocorrelación

Cada instante de una misma señal aleatoria puede considerarse una señal aleatoria en sí. De manera que es posible obtener el coeficiente de correlación entre dos instantes distintos de la señal.

En estos casos, se puede establecer un función de densidad $f_X(x, t)$, y un valor medio $E[x(t)]$, ya que ambos dependen del instante elegido.

Y luego seleccionar dos o más instantes para calcular la autocorrelación de la señal.

Una señal de variación lenta, tendrá un coeficiente ρ cercano a uno, mientras que una de variación rápida tendrá un coeficiente de correlación muy cercano a cero.

3.3.3. Funciones de autocorrelación y autocovarianza

La **función de autocorrelación** se define como:

$$R_X(t_1, t_2) = E [X(t_1)X(t_2)] \quad (29)$$

Y la **función de autocovarianza** se define como:

$$C_X(t_1, t_2) = E [(X(t_1) - m_X(t_1)) (X(t_2) - m_X(t_2))] \quad (30)$$

3.3.4. Propiedades

- \vec{X} e \vec{Y} están **descorrelacionadas** ($\rho = 0$) si y sólo si:

$$E [\vec{X} \vec{Y}^T] = E [\vec{X}] E [\vec{Y}]^T \quad (31)$$

- \vec{X} e \vec{Y} son **ortogonales** si y sólo si $E[\vec{X}\vec{Y}^T] = 0$. Este principio de ortogonalidad es la base del procesamiento lineal de señales.

Por esta propiedad, se puede asimilar $E[XY]$ a un producto interno. En ese caso, la norma sería:

$$\|X\|^2 = \langle X, X \rangle = \int_{-\infty}^{\infty} x^2 f_X(x) dx \quad (32)$$

Si la norma es cero, se trata de una función de valor medio cero y varianza cero, es decir $f_X(x) = \delta(x)$.

- Si \vec{X} e \vec{Y} son independientes entonces \vec{X} e \vec{Y} están descorrelacionadas (la inversa no es válida).

Demostración:

$$R_{XY} = E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x, y) dx dy \quad (33)$$

$$= \int_{-\infty}^{\infty} x f_X(x) dx \int_{-\infty}^{\infty} y f_Y(y) dy \quad (34)$$

$$R_{XY} = m_X m_Y \quad (35)$$

Y como $R_{XY} = \rho \sigma_X \sigma_Y + m_X m_Y$, ρ debe ser cero, o lo que es lo mismo, las variables deben estar descorrelacionadas.

- Es una matriz **definida no negativa** (es decir que $\forall \vec{v} \in \mathfrak{R}^N \neq 0 \Rightarrow \vec{v}^T R_X \vec{v} \geq 0$), y en general, es definida positiva (es decir que es mayor estricto).

Demostración:

$$\vec{v}^T R_X \vec{v} \geq 0 \quad \forall \vec{v} \neq 0 \quad (36)$$

$$\alpha = \vec{v}^T \vec{X} \quad \vec{x} \in \mathfrak{R} \quad (37)$$

$$\alpha^2 \geq 0 \Rightarrow E(\alpha^2) \geq \quad (38)$$

$$E[\alpha^2] = E\left[(\vec{v}^T \vec{X})^2\right] \quad (39)$$

$$= E\left[\vec{v}^T \vec{X} \vec{X}^T \vec{v}\right] \quad (40)$$

$$= \vec{v}^T E\left[\vec{X} \vec{X}^T\right] \vec{v} \quad (41)$$

El caso particular en que $\vec{v}^T R_X \vec{v} = 0$, significa que para algún $\vec{v} \neq 0$, $\vec{v}^T \vec{X} = 0$, y eso implica que al menos una de las componentes X_i es dependiente de las demás, por lo que se puede decir que no son genuinamente variables aleatorias.

3.4. Matriz de covarianza

La **matriz de covarianza** se construye a partir de:

$$C_X = E\left[(\vec{X} - \vec{m}_X)(\vec{X} - \vec{m}_X)^T\right] \quad (42)$$

Es decir que difiere de la matriz de correlación solamente para señales con media distinta de cero.

Los elementos de la diagonal principal estarán dados por:

$$\text{cov}(x_i, x_i) = E \left[(X_i - m_i)^2 \right] = \sigma_i^2 \quad (43)$$

Mientras que los demás elementos serán:

$$\text{cov}(x_i, x_j) = E \left[(X_i - m_i) (X_j - m_j) \right] \quad (44)$$

$$= E \left[X_i X_j \right] - m_i m_j = \rho_{ij} \sigma_i \sigma_j \quad (45)$$

3.4.1. Propiedades

- Se trata de una matriz **simétrica**,
- con **autovalores no negativos** (generalmente **positivos**),
- **definida no negativa**, y en general, es definida positiva.

La demostración es equivalente a la demostración para R_X , pero utilizando $(\vec{X} - m_X)$ en lugar de \vec{X} .

El caso particular en que $C_X = 0$ se da cuando $\vec{X} = m_X$ es decir que la señal permanece constantemente en su valor medio.

- Existe una matriz Q ortonormal y hermítica ($Q \cdot Q^T = I$ y $Q^T \cdot Q = I$) y una matriz diagonal Λ (que contiene los autovalores de C), tales que $C = Q \Lambda Q^T$.
- Además, $C_X = R_X - \vec{m}_X \vec{m}_X^T$
- Los autovectores son ortogonales entre sí.
- Si las componentes de \vec{X} **no están** correlacionadas, los coeficientes ρ_{ij} serán nulos, con lo que la matriz de covarianzas es **diagonal**.

3.5. Transformación lineal de vectores aleatorios

A partir de un vector aleatorio \vec{X} , con media $\vec{m}_{\vec{X}}$ y matriz de covarianza $C_{\vec{X}}$, se puede obtener el vector aleatorio \vec{Y} :

$$\vec{Y} = A\vec{X} + \vec{b} \quad (46)$$

Donde, A es una matriz de $n \times n$ y \vec{b} es un vector de dimensión n .

3.5.1. Momentos

El valor medio será:

$$E[\vec{Y}] = AE[\vec{X}] + \vec{b} \quad (47)$$

$$\vec{m}_{\vec{Y}} = A\vec{m}_{\vec{X}} + \vec{b} \quad (48)$$

La matriz de covarianza estará dada por:

$$C_{\vec{Y}} = E \left[(\vec{Y} - \vec{m}_{\vec{Y}}) (\vec{Y} - \vec{m}_{\vec{Y}})^T \right] \quad (49)$$

Donde:

$$\vec{Y} - \vec{m}_{\vec{Y}} = A\vec{X} + \vec{b} - A\vec{m}_{\vec{X}} - \vec{b} = A(\vec{X} - \vec{m}_{\vec{X}}) \quad (50)$$

Si se toma $\vec{Y} - \vec{m}_{\vec{Y}} = Y_0$ y $\vec{X} - \vec{m}_{\vec{X}} = X_0$, se tiene que la matriz de covarianza está dada por:

$$C_{\vec{Y}} = E[Y_0 Y_0^T] = E[AX_0 X_0^T A^T] \quad (51)$$

Teniendo en cuenta que $E[X_0 X_0^T] = C_{\vec{X}}$, finalmente se llega a que la matriz de covarianza de \vec{Y} será:

$$C_{\vec{Y}} = AC_{\vec{X}}A^T \quad (52)$$

Esta matriz de covarianza es análoga a $\sigma_Y^2 = a^2\sigma_X^2$ para el caso de una sola señal aleatoria.

3.6. Diagonalización de la matriz de covarianza

Teniendo una matriz de covarianza $C_{\vec{X}}$ no diagonal, es posible utilizar la transformación lineal $\vec{Y} = A\vec{X}$ para obtener una matriz $C_{\vec{Y}}$ que sí sea diagonal, es decir, obtener un vector cuyas variables no estén correlacionadas.

De manera que $C_{\vec{Y}} = AC_{\vec{X}}A^T$ sea diagonal.

Imponiendo la condición adicional de que $\|\vec{Y}\|^2 = \|\vec{X}\|^2$:

$$\vec{Y}^T \vec{Y} = \vec{X}^T A^T A \vec{X} = \vec{X}^T \vec{X} \quad (53)$$

Ya que se debe cumplir que las normas sean iguales, es necesario que $A^T A = I$, es decir que $A^T = A^{-1}$: se trata de una matriz ortogonal (en el caso real) o unitaria (en el caso complejo); representa una rotación.

$$C_{\vec{Y}} = AC_{\vec{X}}A^{-1} \quad (54)$$

$$C_{\vec{Y}}A = AC_{\vec{X}} \quad (55)$$

$$A^T C_{\vec{Y}} = C_{\vec{X}} A^T \quad (56)$$

$$BC_{\vec{Y}} = C_{\vec{X}}B \quad (57)$$

Donde $B = A^T$ es una matriz cuyas columnas son los autovectores de $C_{\vec{X}}$, y los autovalores son $b_i \sigma_{Y_i}^2$.

3.7. Distribución gaussiana multivariable

Una variable aleatoria gaussiana de media m y varianza σ^2 (denominada $N(m, \sigma^2)$) tiene una función de densidad de probabilidad:

$$f_X(x) = \frac{e^{-\frac{(x-m)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \quad (58)$$

Utilizando el cambio de variable anterior, un vector aleatorio cuyas componentes sean gaussianas de media \vec{m} y covarianza C , denominado $N(\vec{m}, C)$, tiene una función de densidad de probabilidad:

$$f_{\vec{X}}(\vec{x}) = \frac{e^{-1/2(\vec{x}-\vec{m})^T C^{-1}(\vec{x}-\vec{m})}}{\sqrt{2\pi} \det(C)^{1/2}} \quad (59)$$

Esto se puede dar, por ejemplo, en la distribución de los tiros al blanco en el eje x y y .

3.7.1. Cómo construir una gaussiana

Es posible construir una distribución gaussiana usando una distribución uniforme y una distribución de Rayleigh:

$$x = r \cos \phi \quad (60)$$

$$y = r \sin \phi \quad (61)$$

$$(62)$$

Donde ϕ es una uniforme de valor $1/2\pi$ que va de 0 a 2π , y r cumple con la distribución de Rayleigh.

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}} = \frac{1}{2\pi\sigma} e^{-\frac{r}{2\sigma^2}} \quad (63)$$

$$f_{R\Phi}(r, \phi) = \frac{f_{XY}(x, y)}{\left| \det \frac{\partial(r, \phi)}{\partial(x, y)} \right|} = f_{XY}(x, y) \left| \det \frac{\partial(x, y)}{\partial(r, \phi)} \right| \quad (64)$$

$$= r \frac{e^{-\frac{r}{2\sigma^2}}}{2\pi\sigma} \quad (65)$$

$$f_{\Phi}(\phi) = \int_0^{\infty} f_{R\Phi}(r, \phi) dr = \int_0^{\infty} r \frac{e^{-\frac{r}{2\sigma^2}}}{2\pi\sigma} dr = \frac{1}{2\pi} \quad (66)$$

$$f_R(r) = \int_0^{2\pi} f_{R\Phi}(r, \phi) d\phi = \int_0^{2\pi} r \frac{e^{-\frac{r}{2\sigma^2}}}{2\pi\sigma} d\phi = \frac{e^{-\frac{r}{2\sigma^2}}}{\sigma} \quad (67)$$

4. Procesos Estocásticos

4.1. Concepto de proceso estocástico

Un proceso estocástico es una secuencia de experimentos aleatorios. Generalmente esta secuencia es en el tiempo pero no necesariamente.

Para visualizar este concepto:

- $X(t, \omega)$ constituye una familia de funciones, un proceso estocástico, donde t es un instante de realización y ω es el parámetro de la señal.
- $X(t, \omega_s)$ es una función de t .
- $X(t_0, \omega)$ es una variable aleatoria.

- $X(t_0, \omega_8)$ es un resultado experimental, un número real.

Por ejemplo, se puede tener un proceso en que cual $P(\omega_0) = P(\omega_1) = 1/2$ y $X(t, \omega_0) = 8t + 1$, mientras que $X(t, \omega_1) = \cos(2\pi 1000t)$.

Uno de los problemas importantes a resolver es cómo representar este proceso, ya que cada una de las realizaciones es una señal aleatoria que tiene asociada una forma de onda a lo largo del tiempo.

Lo que se puede hacer es dejar fija una de las dimensiones y caracterizar las otras. Por ejemplo, si se deja fija la realización ξ (es decir, se elige una realización en particular), pueden calcularse los momentos correspondientes:

$$m_X(\xi) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_T x(t, \xi) dt \quad (68)$$

$$R_{XX}(\tau, \xi) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_T x(t, \xi) x(t + \tau, \xi) dt \quad (69)$$

$$P_X(\xi) = R_{XX}(0, \xi) \quad (70)$$

Si, en cambio, se deja fijo el instante del tiempo t_0 , se pueden obtener otros momentos:

$$\mu_X(t) = \int_{-\infty}^{\infty} x f(x; t) dx \quad (71)$$

$$\sigma_X^2(t) = \int_{-\infty}^{\infty} (x - \mu_X(t))^2 f(x; t) dx \quad (72)$$

4.2. Clasificación: tiempo continuo y discreto, amplitud continua y discreta

4.3. Funciones de distribución y densidad de probabilidad de orden n

4.3.1. Orden 1

$$F_X(x, t) = P[X(t) \leq x] \quad (73)$$

$$f_X(x, t) = \frac{dF(x, t)}{dx} \quad (74)$$

Al comparar las realizaciones de un mismo proceso en distintos instantes de tiempo, se obtienen las funciones de densidad $f(x_1, x_2; t_1, t_2)$.

La autocorrelación entre dos instantes t_1 y t_2 será:

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] \quad (75)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2; t_1, t_2) dx_1 dx_2 \quad (76)$$

Y la probabilidad estará dada por:

$$f_{X_1 X_2}(x_1, x_2; t_1, t_2) = P[X(t_1) \leq x, X(t_2) \leq x] \quad (77)$$

4.4. Procesos estacionarios

Un proceso estacionario es aquél que es invariante ante una traslación en el origen de tiempos.

4.4.1. Proceso estacionario de orden 1

Debe cumplirse que para todo $\varepsilon \in \mathfrak{R}$, sus funciones de probabilidad y de distribución no dependan del tiempo, sino sólo de x :

$$F_X(x, t) = F_X(x, t + \varepsilon) \quad (78)$$

$$f_X(x, t) = f_X(x, t + \varepsilon) \quad (79)$$

Es decir que la probabilidad no cambia ante una traslación en el origen de tiempos. Por esto mismo, el valor medio y la varianza serán constantes para todo t : $m_X(t) = m_X$ y $\sigma_X(t) = \sigma_X$.

4.4.2. Proceso estacionario de orden 2

Debe cumplirse que para todo $\varepsilon \in \mathfrak{R}$:

$$F_{X_1 X_2}(x_1, x_2, t_1, t_2) = F_{X_1 X_2}(x_1, x_2, t_1 + \varepsilon, t_2 + \varepsilon) \quad (80)$$

Por lo que se puede decir que $F_{X_1 X_2}(x_1, x_2, t_1, t_2) = F_{X_1 X_2}(x_1, x_2, t_2 - t_1)$, ya que solamente depende de la diferencia de tiempos entre t_1 y t_2 .

Que un proceso sea estacionario de orden 2 implica que es estacionario de orden 1, es decir que su función de densidad de probabilidad es independiente del tiempo y que el valor medio y la varianza son constantes en el tiempo.

Como puede verse en la ecuación (76, en un proceso estacionario de orden 2 es claro que la autocorrelación sólo depende de la diferencia de tiempos: $R_X(t_1, t_2) = R_X(t_2 - t_1) = R_X(\tau)$. Y de la misma manera con $C_X(\tau)$

4.4.3. Proceso estacionario en sentido amplio

Se denomina proceso **estacionario en sentido amplio** (ESA o *WSS*) o también proceso *débilmente estacionario* a aquellos procesos que tengan m_X y σ_X^2 constantes y R_X dependiente únicamente de τ .

No interesa cómo es la función en sí, sino los datos sobre el valor medio, la varianza y la autocorrelación. Ya que si la entrada a una sistema LTI tiene estas propiedades, la salida también las tendrá.

No existen en la práctica los procesos estacionarios ideales, sin embargo, en un gran número de casos, se puede tomar un proceso como estacionario ya que durante el tiempo pertinente los parámetros no varían.

4.5. Valor medio, varianza y autocorrelación

Para una variable en dos instantes t_1 y t_2 :

$$C_X(t_1, t_2) = E[(X(t_1) - m_X(t_1))(X(t_2) - m_X(t_2))] \quad (81)$$

$$C_X(t_1, t_2) = \rho(t_1, t_2)\sigma_X(t_1)\sigma_X(t_2) \quad (82)$$

$$R_X(t_1, t_2) = C_X(t_1, t_2) + m_X(t_1)m_X(t_2) \quad (83)$$

$$(84)$$

4.6. Interpretación de la función de autocorrelación

A partir de la función de autocorrelación se puede distinguir una señal cuyos valores varían muy rápidamente de una señal cuyos valores varían lentamente. En el primer caso, la función $R_X(\tau)$ se extenderá mucho más sobre la recta τ que en el segundo caso.

En otras palabras, cuando la autocorrelación es alta, entre dos instantes cercanos tendremos un valor similar; pero cuando es baja, podremos tener cualquier valor.

Propiedades

- Tanto $R_X(\tau)$ como $C_X(\tau)$ son funciones pares

▪

$$R_X(\tau) = C_X(\tau) + m_X^2 = \rho(\tau)\sigma_x^2 + m_X^2 \quad (85)$$

$$C_X(0) = \sigma_X^2 \geq |C_X(\tau)| \quad \forall \tau \quad (86)$$

$$(87)$$

▪

$$\begin{aligned} E[(X(t) - X(t + \tau))^2] &= E[X^2(t)] + E[X^2(t + \tau)] - 2E[X(t)X(t + \tau)] \quad (88) \\ &= 2R_X(0) - 2R_X(\tau) \quad (89) \end{aligned}$$

$$= 2C_X(0) - 2C_X(\tau) \quad (90)$$

$$= 2\sigma_X^2(1 - \rho(\tau)) > 0 \quad (91)$$

De manera que el coeficiente de autocorrelación para $\tau = 0$ será $\rho(0) = 1$ siempre, ya que se trata de la misma medición; y $R_X(0) = \sigma_X^2 + m_X^2$ puede ser interpretado como la potencia del proceso.

4.7. Teorema de Wiener-Khinchin

El teorema de Wiener-Khinchin establece que la transformada de Fourier de la función de autocorrelación constituye la densidad espectral de potencia de esa función.

4.7.1. Caso continuo

4.7.2. Caso discreto

Sea $x(n)$ una señal aleatoria **ESA**, cuya función de autocorrelación está dada por:

$$r_{XX}(m) = E [x(n+m)x^*(n)] \quad (92)$$

La densidad de potencia espectral será la transformada discreta de Fourier de r_{XX} :

$$S_{XX}(\Omega) = T \sum_{n=-\infty}^{\infty} r_{XX}(m) e^{i2\pi\Omega m T} \quad (93)$$

Donde T es el intervalo de muestreo, y se asume que la señal es de banda limitada a $\pm 1/2T$, y es periódica, con frecuencia $1/T$.

4.8. Densidad espectral de potencia

Se define $S_X(\omega)$, la **densidad espectral de potencia** del vector \vec{X} como la transformada de Fourier de la función de autocorrelación R_X .

$$S_X(\omega) = \mathcal{F} \{R_X\} = \int R_X(\tau) e^{-j\omega\tau} d\tau \quad (94)$$

La densidad espectral de potencia representa el promedio de la energía que está transmitiendo un proceso en una determinada frecuencia.

Para un proceso estocástico no estacionario, la densidad espectral de potencia estará dada por:

$$S_X(t, \omega) = \mathcal{F} \{R_X(t, \tau)\} = \int R_X(t, \tau) e^{-j\omega\tau} d\tau \quad (95)$$

En este caso, se verifica que su área es la potencia del proceso o su autocorrelación en $\tau = 0$.

Además, de esta misma manera se puede calcular la densidad espectral de potencia para la correlación entre dos procesos ($r_{XY}(t, \tau) = E [x(t)y(t+\tau)]$), teniendo en cuenta que cuando la correlación entre ambos procesos es estacionaria, r_{XY} dependerá únicamente de τ .

$$S_{XY}(\omega) = \mathcal{F} \{R_{XY}(\tau)\} = \int R_{XY}(\tau) e^{-j\omega\tau} d\tau \quad (96)$$

4.9. Comparación de formulaciones en tiempo continuo y discreto

4.10. Expansión de Karhunen-Loève

4.11. Procesos ergódicos

Se dice que un proceso estacionario es ergódico cuando las funciones que entrañan valores esperados a lo largo de las realizaciones pueden obtenerse también a partir de una sola realización.

Es decir que una sola realización es representativa de toda la familia.

Un ejemplo de proceso ergódico es un promedio temporal:

$$\vec{X} = \frac{1}{N} \sum_{i=1}^N x_i \quad (97)$$

5. Sistemas lineales con excitaciones aleatorias

Si un proceso estocástico se aplica a un sistema lineal, a la salida se obtiene otro proceso estocástico. La relación entre las variables que caracterizan al de la entrada y las que caracterizan al de la salida, se obtienen a partir de la respuesta impulsiva $h(t)$.

5.1. Valor medio, autocorrelación y densidad espectral de potencia de la salida

Teniendo en cuenta que:

$$y(t) = \int h(\tau)x(t - \tau)d\tau \quad (98)$$

Se puede obtener el valor medio de $y(t)$ como:

$$E[Y] = \int h(\tau)E[x(t - \tau)]d\tau \quad (99)$$

Y, si X es estacionario:

$$E[Y] = E[X] \int h(\tau)d\tau = E[X]H(0) \quad (100)$$

Por otro lado, la autocorrelación de la salida será:

$$R_Y(t) = h(t)h^T(-t)R_X(t) \quad (101)$$

Y la densidad espectral de potencia:

$$S_{XY}(\omega) = H(\omega)S_X(\omega) \quad (102)$$

$$S_Y(\omega) = |H(\omega)|^2 S_X(\omega) \quad (103)$$

Estas relaciones son importantes, ya que constituyen una manera real y práctica de obtener la respuesta impulsiva de un sistema lineal.

Es decir que utilizando la ecuación:

$$H(\omega) = \frac{S_{XY}(\omega)}{S_X(\omega)} \quad (104)$$

Se puede obtener la respuesta impulsiva a partir de las densidades espectrales de la entrada y de la salida en función de la entrada.

5.2. Sistemas de múltiples entradas y salidas

5.3. Filtros

5.4. Formulación en tiempo continuo y discreto

6. Procesos particulares

6.1. Ruido blanco

Se denomina **ruido blanco** a una señal aleatoria con ancho de banda infinito, es decir que contiene a todas las frecuencias.

Hay dos tipos de ruido blanco: uniforme o gaussiano.

La densidad espectral del ruido blanco es una uniforme que va desde $-pi/T$ hasta pi/T , con valor $N_0/2$.

El ruido blanco real no existe, ya que no se puede tener ancho de banda infinito. En la práctica se utilizan señales que tienen un ancho de banda mayor al del sistema estudiado, por lo que se lo considera infinito.

6.2. Modelos lineales en tiempo discreto: procesos AR, MA y ARMA

6.2.1. Proceso AR1

Un proceso **autoregresivo de primer orden** estará dado por la siguiente ecuación:

$$X_n = aX_{n-1} + W_n \quad (105)$$

Donde W_n es ruido blanco, con varianza σ_W^2 y valor medio nulo. En las ecuaciones que se incluyen a continuación se supone que la entrada es determinística, ya que de no ser así no es posible efectuar la transformada z .

Si X_n es un proceso **ESA**, se puede obtener el valor medio:

$$E[X_n] = aE[X_{n-1}] + E[W_n] \quad (106)$$

$$E[X_n] = aE[X_n] + 0 \quad (107)$$

$$E[X_n] = 0 \quad (108)$$

Y la varianza:

$$\sigma_X^2 = E[X_n^2] = E[a^2X_{n-1}^2 + W_n^2 + 2aX_{n-1}W_n] \quad (109)$$

$$E[X_n^2] = a^2E[X_{n-1}^2] + E[W_n^2] + 2aE[X_{n-1}W_n] \quad (110)$$

$$E[X_n^2] = a^2E[X_n^2] + \sigma_W^2 + 2a0 \quad (111)$$

$$\sigma_X^2 = a^2\sigma_X^2 + \sigma_W^2 \quad (112)$$

$$\sigma_X^2 = \frac{\sigma_W^2}{1-a} \quad (113)$$

Donde $|a| < 1$.

Y, por último, la autocorrelación:

$$R_X(k) = a^{|k|} \sigma_X^2 = \frac{a^{|k|}}{1-a} \sigma_W^2 \quad (114)$$

$$R_X(0) = a^0 \sigma_X^2 = \sigma_X^2 \quad (115)$$

$$R_X(1) = C_X(1) = a^1 \sigma_X^2 = \rho \sigma_{X_n} \sigma_{X_{n-1}} \quad (116)$$

Es claro que en este último caso, $\rho = a$, y para el caso general, $\rho(k) = a^k$, es el coeficiente de correlación entre X_n y X_{n+k} .

Si se obtiene $S_X(\omega)$ como la transformada de Fourier de la correlación, se puede ver la *coloración* del ruido.

Ejemplos de valores de a :

- $a = 0,9$, constituye un pasabajos, donde cada valor está muy relacionado con el anterior, la función $C_X(k)$ tendrá una numerosa cantidad de valores y el trazado de X_{n+1} en función de X_n es una elipse muy cerrada, inclinada $\pi/4$.
- $a = 0,5$, se trata de un pasabajos más leve, donde cada valor está un poco relacionado con el anterior, la función $C_X(k)$ tiene menor cantidad de valores y el trazado de X_{n+1} en función de X_n es una elipse levemente cerrada.
- $a = -0,9$, se trata de un pasaltos, donde los valores suelen alternar el signo, la función $C_X(k)$ también alternará el signo de sus componentes y el trazado de X_{n+1} en función de X_n es una elipse muy cerrada, inclinada a $3\pi/4$.

6.2.2. Proceso AR2

Un **proceso autoregresivo de orden 2** se describe mediante la siguiente ecuación:

$$X_n = aX_{n-1} + bX_{n-2} + W_n \quad (117)$$

Dentro del plano a, b es necesario buscar las regiones para las que el comportamiento del sistema es estable. Y dentro de esa región hay que buscar las zonas donde los polos de la transferencia sean reales y donde sean complejos.

Cuando los polos son complejos conjugados, X tiende a oscilar y es mucho más *rico*.

6.2.3. Proceso AR-q

Un **proceso autoregresivo de orden q** se describe mediante la siguiente ecuación:

$$X_n = b_1 X_{n-1} + b_2 X_{n-2} + \dots + b_q X_{n-q} + W_n \quad (118)$$

$$X_n = W_n + \sum_{i=1}^q b_i X_{n-i} \quad (119)$$

Haciendo la transformada Z :

$$X(z) = W(z) + \sum_{i=1}^q b_i z^{-1} X(z) \quad (120)$$

Teniendo en cuenta que la entrada del sistema es $W(z)$ y la salida es $X(z)$, se puede obtener una expresión para la respuesta impulsiva $H(z)$.

$$H(z) = \frac{1}{1 - \sum_{i=1}^q b_i z^{-i}} \quad (121)$$

Se trata de una respuesta que tiene q polos y ningún cero, es decir que se trata de un filtro **IIR**.

6.2.4. Proceso MA

Se denomina *Moving Average* o promedio móvil, o filtro transversal o ventana deslizante, a un sistema descrito por la siguiente ecuación:

$$Y_n = \sum_{i=0}^P b_i W_{n-i} \quad (122)$$

Es decir que en este sistema se suman los valores previos de W_n , multiplicados por un factor.

Teniendo en cuenta que W_n es la entrada y Y_n la salida, se puede obtener la respuesta impulsiva:

$$h(i) = b_i \quad (123)$$

$$H(z) = \sum_{i=0}^P b_i z^{-i} \quad (124)$$

Es decir que se trata de un filtro cuya transferencia tiene ceros y no tiene ningún polo, es un filtro **FIR**.

Por otro lado, es posible calcular el valor medio de la salida:

$$E[Y_n] = \vec{h}^T E[W_n] \quad (125)$$

Dado que W_n es ruido blanco, con media cero, la media de la salida será también cero.

También se puede calcular la varianza:

$$E[Y_n^2] = E[\vec{h}^T W_n W_n^T \vec{h}] \quad (126)$$

$$E[Y_n^2] = \vec{h}^T E[W_n W_n^T] \vec{h} = \vec{h}^T C_X \vec{h} \quad (127)$$

6.2.5. Procesos ARMA

Combinando la ecuación general de los procesos autoregresivos con la del promedio móvil, se puede obtener la ecuación de un sistema en que se realimentan tanto la entrada como la salida.

De esta manera es posible obtener procesos con una gran variedad de características espectrales, a la vez que una gran cantidad de procesos pueden pensarse como si hubieran sido generados a partir de ruido blanco y un sistema AR, MA o ARMA.

6.3. Procesos gaussianos

6.4. Ruido de banda angosta: representación de Rice, distribución de envolvente, densidad espectral de potencia

6.5. Proceso de Poisson

6.6. Procesos de Markov: cadenas de Markov en tiempo discreto y continuo

6.7. Ecuación de Chapman-Kolmogoroff

6.8. Ecuaciones de Kolmogoroff y ecuaciones de balance global

7. Elementos de teoría de filtros óptimos

Cuando se tiene una señal que pasa por un filtro, normalmente también se tiene un determinado ruido que se suma a la señal, y cuya salida por el filtro se sumará a la salida de la señal.

$$X(t) + N(t) \rightarrow Y(t) + V(t) \quad (128)$$

Teniendo esto en cuenta, el objetivo de los filtros debe ser **maximizar la relación señal-ruido**, de manera que se preserve la mayor cantidad de información.

7.1. Criterios de optimización de un filtro

El coeficiente $\rho = S/N$ se puede maximizar de distintas maneras según el criterio adoptado. Un criterio puede ser maximizar la relación:

$$\rho = \frac{E[Y(t)]}{E[V(t)]} \quad (129)$$

Otra relación posible:

$$\rho = \frac{|Y(t)|^2}{E[V(t)]} \quad (130)$$

En este caso, $|Y(t)|^2$ representa la potencia instantánea de la señal, mientras que $E[V(t)]$ representa la potencia media del ruido.

Este segundo criterio se utiliza cuando se quiere discriminar entre unos y ceros, positivo y negativo, presencia o ausencia de pulsos (ej: radar).

7.2. Filtro adaptado: formulaciones en tiempo continuo y tiempo discreto (FIR e IIR)

Para obtener la función de transferencia $h(t)$ que maximiza la relación señal-ruido, se utiliza la siguiente deducción:

$$V(t) = h(t) * N(t) \quad (131)$$

$$E[V^2(t)] = \sigma_V^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_V(\omega) d\omega \quad (132)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(\omega)|^2 S_N(\omega) d\omega \quad (133)$$

$$y(t) = h(t) * x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(\omega) e^{j\omega t} d\omega \quad (134)$$

Con esto es posible encontrar el coeficiente *rho*, para ello se utiliza el instante t_0 en el que la salida es máxima.

$$\rho = \frac{\left(\frac{1}{2\pi} \int_{-\infty}^{\infty} Y(\omega) e^{j\omega t_0} d\omega \right)^2}{\frac{1}{2\pi} \int_{-\infty}^{\infty} |H(\omega)|^2 S_N(\omega) d\omega} \quad (135)$$

$$= \frac{1}{2\pi} \frac{\left| \int_{-\infty}^{\infty} H(\omega) X(\omega) e^{j\omega t_0} d\omega \right|^2}{\int_{-\infty}^{\infty} |H(\omega)|^2 S_N(\omega) d\omega} \quad (136)$$

El objetivo es encontrar $H(\omega)$ tal que este cociente sea máximo, pero no se puede hacer derivando e igualando a cero, es necesario utilizar la desigualdad de Cauchy-Schwartz

Desigualdad de Cauchy-Schwartz:

$$|\langle u, v \rangle|^2 \leq \|u\|^2 \cdot \|v\|^2 \quad (137)$$

Esta desigualdad es válida en los *espacios de Banach*, en los que $\langle u, u \rangle = \|u\|^2$. La igualdad solamente se da cuando $u = kv$.

Para el espacio de las variables aleatorias se define:

$$\langle u, v \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} U(\omega) V^*(\omega) d\omega \quad (138)$$

$$\|u\|^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} U(\omega)^2 d\omega = \int_{-\infty}^{\infty} U(t)^2 dt \quad (139)$$

Es decir que se están tomando a las señales como vectores (se utiliza mucho en comunicaciones digitales).

De manera que para las ecuaciones anteriores, podemos definir:

$$U(\omega) = H(\omega) \sqrt{S_N(\omega)} \quad (140)$$

$$V(\omega) = \frac{X(\omega)}{\sqrt{S_N(\omega)}} e^{j\omega t_0} \quad (141)$$

$$\rho = \frac{2\pi |\langle u, v \rangle|^2}{(2\pi)^2 \|u\|^2} \quad (142)$$

$$= \frac{1}{2\pi} \frac{|\langle u, v \rangle|^2}{\|u\|^2} \leq \frac{1}{2\pi} \|v\|^2 \quad (143)$$

$$\frac{1}{2\pi} \|v\|^2 = \left(\frac{1}{2\pi} \right)^2 \int_{-\infty}^{\infty} \frac{|X(\omega)|^2}{S_N(\omega)} d\omega \quad (144)$$

Se cuenta, entonces con una cota superior para la relación señal-ruido, de manera que el valor de esta cota es $\rho_{\text{máx}}$. Ese valor se alcanzará cuando valga el $=$ del \leq , es decir cuando $u = kv$:

$$H(\omega)\sqrt{S_N(\omega)} = k \frac{X^*(\omega)}{\sqrt{S_N(\omega)}} \quad (145)$$

- 7.3. Fundamentos de estimación lineal en medida cuadrática: espacios de Hilbert de variables aleatorias de 2º orden, teorema de la proyección, principio de ortogonalidad**
- 7.4. Aplicaciones: filtrado, predicción y alisado de datos**
- 7.5. Ecuación de Wiener-Hopf**
- 7.6. Filtro de Wiener no causal y causal**
- 7.7. Ecuación de Yule-Walker**
- 8. Elementos de teoría de decisión**
 - 8.1. Decisión entre hipótesis binarias**
 - 8.2. Relación de verosimilitud**
 - 8.3. Reglas de decisión de Bayes y de Neyman-Pearson**
 - 8.4. Comportamiento del clasificador: probabilidad de error, de pérdida y de falsa alarma**
 - 8.5. Decisión entre hipótesis múltiples**
 - 8.6. Detección binaria con observaciones múltiples y ruido gaussiano; relación con el filtro adaptado**
- 9. Elementos de teoría de colas**
 - 9.1. Teorema de Little**
 - 9.2. Colas M/M/1, M/M/1/K, M/M/c, M/M/c/c**
 - 9.3. Ejemplos de aplicación**