
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

INTRODUCCIÓN AL PROCESAMIENTO DIGITAL DE SEÑALES ESTOCÁSTICAS

Rubén D. León V.

ELÉCTRICA Y ELECTRÓNICA

INTRODUCCIÓN AL PROCESAMIENTO DIGITAL DE SEÑALES ESTOCÁSTICAS

Rubén D. León V.

Introducción al procesamiento digital de señales estocásticas

Ing. Rubén D. León V. M.Sc.

Primera edición electrónica. Diciembre de 2014

ISBN: 978-9978-301-28-9

Par revisor: Alexis Tinoco. PhD

Universidad de las Fuerzas Armadas - ESPE

Grab. Roque Moreira Cedeño

Rector

Crnl. Francisco Armendáriz Saénz

Vicerrector Académico General

Crnl. Ricardo Urbina

Vicerrector de Investigación

Publicación autorizada por:

Comisión Editorial de la Universidad de las Fuerzas Armadas - ESPE

Edición y producción

David Andrade Aguirre

daa06@yahoo.es

Diseño

Pablo Zavala A.

Derechos reservados. Se prohíbe la reproducción de esta obra por cualquier medio impreso, reprográfico o electrónico.

El contenido, uso de fotografías, gráficos, cuadros, tablas y referencias es de **exclusiva responsabilidad** del autor.

Los derechos de esta edición electrónica son de la **Universidad de las Fuerzas Armadas - ESPE**, para consulta de profesores y estudiantes de la universidad e investigadores en: <http://www.repositorio.espe.edu.ec>.

Universidad de las Fuerzas Armadas - ESPE

Av. General Rumiñahui s/n, Sangolquí, Ecuador.

<http://www.espe.edu.ec>

Dedicatoria

A mis hijos Rubén Darío y Gelson David

Prefacio

Este texto tiene su origen en el seminario titulado “Introducción al procesamiento digital de señales estocásticas” dictado en la Facultad de Ingeniería Electrónica (FIE) de la Escuela Politécnica del Ejército en septiembre de 1993. Los objetivos del seminario fueron: (1) Actualizar los conocimientos y desarrollar la teoría matemática de los procesos estocásticos y del procesamiento digital de señales, (2) Ilustrar esta teoría con su aplicación a señales de voz, sísmica y radar, y (3) Motivar al personal docente de la FIE a incursionar en la investigación en áreas de su interés.

También, este seminario formó parte del proceso de modernización experimentado en la FIE en la actualización de los programas de estudio propuesto por el autor en octubre de 1992. El enfoque actual de materias como probabilidades y procesos estocásticos y procesamiento digital de señales han permitido que el Departamento de Eléctrica y Electrónica de la Universidad de las Fuerzas Armadas ESPE, incursione en la investigación de áreas como sísmica, voz, video y radar y cuyos trabajos han sido publicados tanto a nivel nacional como internacional.

Otro elemento que motivo la elaboración de este texto es la falta de bibliografía de buen nivel existente en nuestro país y lo costoso de la poca existente. Para su elaboración se ha guiado en textos clásicos como: [Papoulis, A., 1991] y [Oppenheim, A.V., Schafer. R., 1989] y la experiencia adquirida desde el año de 1990 que inicié la Maestría en Ciencias (Telecomunicaciones).

El texto ha sido dividido en dos partes. En la primera se realiza un estudio de las probabilidades y procesos estocásticos con cierta rigurosidad matemática mas sin descuidar el aspecto conceptual. En la segunda se estudia la teoría del procesamiento digital de señales, dando énfasis al estudio de la Fast Fourier Transform y su interpretación para su adecuada aplicación. Se incluye también la estimación espectral clásica de la densidad espectral de potencia de un proceso estocástico estacionario en el sentido amplio.

Por último, quiero agradecer a las autoridades de la Facultad de Ingeniería Electrónica y del Departamento de Eléctrica y Electrónica de la ESPE por el apoyo prestado en la elaboración de este documento. Así también quiero presentar mis especiales agradecimientos a los Profesores Dr. Fernando Toshinori Sakane, Dr. Fernando Walter (†) del Instituto Tecnológico da Aeronáutica por las observaciones y comentarios del material aquí presentado. Igualmente mis eternos agradecimientos al Prof. David Fernandes por su orientación y apoyo en el procesamiento de señales de radar.

PROBABILIDADES Y PROCESOS ESTOCÁSTICOS

Introducción

En el análisis de sistemas lineales hemos estudiado varios tipos de señales: periódicas y no periódicas. Las señales estudiadas son tales que sus valores en cualquier instante del tiempo son conocidos a partir de su descripción gráfica o una función analítica. Estas señales que pueden ser especificadas gráfica o analíticamente son llamadas *señales deterministas*.

La primera parte de este texto no será estudiada dando énfasis a estadística, ruido o cualquier otro tópico, sino como un entrenamiento básico para toda ingeniería. La mayoría de los conceptos necesarios en las aplicaciones pueden ser explicados como simples matemáticas, es decir, la *probabilidad* será vista como una estructura conceptual y sus conclusiones serán lógicas y no intuitivas.

La clase de señales, que son el resultado de eventos físicos que dependen de muchos factores y de su complejidad, tienen la característica importante de no poder ser predecibles. Sin embargo, tales eventos físicos, poseen algún tipo de "regularidad", es decir, podrían ser caracterizados mediante ciertos *promedios*. Por lo tanto, los factores y complejidad que influyen en el resultado de un evento físico son llamados *mecanismos aleatorios*.

En este contexto, señales como: voz, audio, sísmica, radar serán consideradas como siendo una realización de un proceso estocástico estacionario en el sentido amplio [03, 05, 07, 17] .

I.1 Probabilidades

I.1.1 Introducción a las probabilidades

Ha sido observado en diferentes áreas que el resultado de un experimento depende de la combinación de muchos factores impredecibles pero sin embargo ciertos *promedios* o características medias se aproximan a un "valor constante" cuando el número de observaciones incrementa. Se observa también, que este "valor constante" se mantiene el mismo si los *promedios* son nuevamente evaluados sobre cualquier conjunto de realizaciones en términos de las "*probabilidades de eventos*" o "*verosimilitud de eventos*".

Definición Clásica:

De acuerdo con la definición clásica, la probabilidad $P(A)$ de un evento A es determinada sin la realización del experimento y dada por:

$$P(A) = N_a / N \quad (1.1)$$

donde: N es el número de posibles resultados y

N_a es el número de resultados que son favorables al evento A

Sin embargo, la definición clásica adolece de ambigüedades ya que el significado de los números N y N_a no es claro. Realicemos el experimento de lanzar dos dados y queremos encontrar la probabilidad, q , de que la suma de los números resultantes sea igual a 7. Según la definición clásica, debemos determinar N_a y N .

a) N es el número de posibles sumas que se pueden obtener al lanzar dos dados : 2, 3, 4, 5, 6, ..., 12, es decir, $N = 11$. De estos eventos, suma = 7, es encontrado una vez, es decir, $N_a = 1$, entonces $q = 1/11$.

b) Todos los posibles pares, sin distinguir entre el uno y el otro dado, dando como resultado 21 sumas posibles de las cuales los pares (6,1), (4,3), (5,2) dan el evento suma = 7, es decir, $N_a = 3$, entonces, $q = 3/21$.

Se observa en los casos anteriores que cada uno de los eventos no son *igualmente probables* y por eso las soluciones son incorrectas.

c) Llevando en consideración la observación anterior, el número de posibles eventos tomando en cuenta todos los pares distinguibles entre el primero y el

segundo dado es $N = 36$ de los cuales los eventos (1-6), (6-1), (3-4), (4-3), (5-2) y (2-5) dan como resultado el evento suma = 7, es decir, $q = 6/36$.

Por lo tanto, la definición clásica es modificada a una versión mejorada como sigue:

" La probabilidad de un evento es igual a la relación de sus resultados favorables para el número total de posibles resultados permitiendo que todos los eventos sean igualmente probables."

La definición clásica fue introducida como consecuencia del *Principio de la razón insuficiente* (Bernoulli, 1713) [01,14]. En la ausencia de cualquier conocimiento a priori, debemos "asumir" que los eventos sean "igualmente probables".

Esta conclusión esta basada en la interpretación subjetiva de probabilidades como "*una medida de nuestro estado de conocimiento*" acerca de los eventos de interés. De ahí que la definición es cuestionada y mencionaremos sólo algunos:

- El término "igualmente probables" es utilizado en la versión mejorada de la definición clásica, es decir, en la definición está siendo usado el concepto a ser definido lo que lleva a los problemas de determinar N y N_a .
- La definición puede ser aplicada solamente a una limitada clase de problemas. Si en el caso de un dado cargamos una de sus caras, evidentemente que la condición de equiprobabilidad no se cumple.
- Parece que la definición clásica es una consecuencia de imperativos lógicos divorciados de la experiencia. Esto, sin embargo, no es así: Es aceptado que ciertas alternativas son igualmente probables debido a la experiencia colectiva como por ejemplo: el experimento del lanzamiento del dado, la probabilidad que un recién nacido sea niña o niño, los cuales son considerados igualmente probables después de existir una gran acumulación de información .

Interpretación Frecuencia Relativa:

La técnica denominada "*Frecuencia relativa*" intenta definir la probabilidad de un cierto evento cuantitativamente, y es dada por el límite:

$$P(A) = \lim_{N \rightarrow \infty} (N_a / N) \quad (1.2)$$

Se observa que la teoría está fundamentada en la observación a diferencia de la definición clásica; sin embargo, en un experimento físico los números N_A y N aún siendo muy grandes, ellos son finitos y, por tanto, su relación no puede ser igualada a la expresión anterior. Por lo tanto, si la definición anterior es usada para definir $P(A)$, el límite debe ser aceptado como una *hipótesis* y no como un número que puede ser determinado experimentalmente.

Se observa también que en la definición anterior (1.2) implícitamente se asume que el límite cuando N tiende a infinito, existe y que la probabilidad de un evento es una cantidad no negativa menor que la unidad.

Axiomas de la Teoría de Probabilidades

Las definiciones anteriores, desde el punto de vista matemático, dejan mucho que desear. En cualquier experimento físico, el número N podría ser muy grande, pero siempre será finito. Surge una pregunta: Cuán grande tiene que ser N para obtener el correcto valor de $P(A)$? Según algunos autores, el fundamento permitido por la definición de frecuencia relativa es demasiado imprecisa para permitirnos construir una estructura matemática precisa [01, 14].

Es necesario anotar que anteriormente ya fue mencionado que en (1.2), el límite debe ser aceptado como una hipótesis, y no como un número que puede ser determinado experimentalmente.

Desde el punto de vista matemático es necesario partir de un set de axiomas, para luego, basados en este set, solamente deducir algo más. De esta forma partimos con ciertos conceptos no definidos. Luego asumimos ciertas propiedades y las relaciones existentes entre estos conceptos. Estos son los axiomas; luego, basados en estos axiomas solamente, varias premisas son obtenidas; éstos son llamados teoremas.

Observamos que la matemática es una pura deducción lógica. Se dice que los axiomas básicos son el corazón de cualquier ramificación de las matemáticas. Estos axiomas necesitan satisfacer solamente un requerimiento, ellos deben ser consistentes con los otros, es decir, uno o más axiomas del set tomados individualmente o colectivamente no deben contradecirse con cualquier otro axioma en el set.

Sin embargo, si una rama de las matemáticas pretende ser de algún uso práctico, los axiomas deben estrictamente relacionarse a algún fenómeno

observado. Esto nos lleva directamente a la noción de un modelo matemático de un fenómeno real.

Los objetos abstractos alrededor de los cuales los axiomas son propuestos, deben ser identificados con algún fenómeno real, y los axiomas deben luego cerradamente aproximarse a las relaciones observadas del fenómeno real. Los teoremas de la teoría matemática, luego mostrarán el fenómeno real en la cuantía en que los axiomas se aproximen a las características observadas en el fenómeno.

Con este conocimiento básico, estamos listos a discutir el enfoque axiomático de las probabilidades. En este enfoque el concepto no definido, o el objeto abstracto es probabilidad; por lo tanto, debemos postular los axiomas acerca de la probabilidad de un evento en tal forma que ellos conformen a las relaciones observadas en la práctica.

Las relaciones observadas están fuertemente relacionadas a la definición de frecuencia relativa en (1.2). Es bueno indicar que el enfoque axiomático no usa la definición (1.2) para definir a la probabilidad, esto es, debemos comenzar con un concepto de probabilidad y postular los axiomas de este concepto. Luego, basados en estos axiomas, solamente nosotros investigaremos qué es verdad?. Esta es la esencia de la teoría moderna de la probabilidad.

Desde el punto de vista de frecuencia relativa de varios resultados, la siguiente terminología es a menudo utilizada:

Una simple culminación de un experimento es llamado *realización*. A cada realización observamos un simple resultado γ_i . Se dice que un evento A ocurrió durante esta realización si el evento contiene al elemento γ_i . El evento certeza ocurre cuando A o B o ambos ocurren. El evento AB ocurre cuando ambos A y B ocurren. Si los eventos A y B son mutuamente exclusivos y A ocurre, luego B no ocurre. Si $A \subset B$ y A ocurre, luego B ocurre.

En cada realización, cualquiera A o \bar{A} ocurre.

Axiomas

La probabilidad de un evento A, $P(A)$, es un número asignado a este evento. Este número es escogido para satisfacer las siguientes condiciones:

$$1.- P(A) \geq 0 \quad (1.3)$$

$$2.- P(S) = 1 \quad (1.4)$$

$$3.- \text{ Si } AB = \{ \Phi \}, \text{ luego } P(A + B) = P(A) + P(B) \quad (1.5)$$

Observemos que la técnica axiomática no dice nada de cómo asignar una probabilidad $P(A)$ a un evento A , pero postula que la función probabilidad debe obedecer los tres axiomas. En el desarrollo de la teoría, todas las conclusiones están basadas directa o indirectamente en los axiomas y sólo en los axiomas. Lo siguiente son simples consecuencias.

Propiedades

- La probabilidad del evento imposible es cero: $P\{\Phi\} = 0$.

$$A\{\Phi\} = \{\Phi\} \quad \text{y} \quad A + \{\Phi\} = A \quad \text{por ser mutuamente exclusivos}$$

$$\text{Por tanto, } P(A) = P(A + \{\Phi\})$$

$$P(A) = P(A) + P(\{\Phi\})$$

$$P\{\Phi\} = 0$$

- Para cualquier A , $P(\bar{A}) = 1 - P(A) \leq 1$

$$A + \bar{A} = S \quad \text{y} \quad A \bar{A} = \{\Phi\}$$

$$P(S) = 1 = P(A + \bar{A}) = P(A) + P(\bar{A})$$

$$P(\bar{A}) = 1 - P(A) \leq 1$$

- Para cualquier A y B , $P(A + B) = P(A) + P(B) - P(AB) \leq P(A) + P(B)$

$A + B = A + \bar{A}B$, ya que A y $\bar{A}B$ son mutuamente exclusivos, entonces:

$$P(A + B) = P(A) + P(\bar{A}B) \quad (\text{axioma-3})$$

$B = AB + \bar{A}B$ ya que también AB y $\bar{A}B$ son mutuamente excluyentes, entonces:

$$P(B) = P(AB) + P(\bar{A}B) \quad (\text{axioma-3})$$

De esta forma:

$$P(A + B) - P(A) = P(B) - P(AB)$$

$$P(A + B) = P(A) + P(B) - P(AB)$$

- Si $B \subset A$, $A = B + A\bar{B}$, luego: $P(A) = P(B) + P(A\bar{B}) \geq P(B)$

Ya que $B \subset A$. entonces son mutuamente exclusivos

$$\text{entonces } P(A) = P(B) + P(A\bar{B}) \geq P(B) \quad (\text{axioma-3})$$

Como se mencionó. la técnica axiomática "se lava las manos" en lo que tiene que ver a la asignación de la probabilidad $P(A)$ a un evento A . Ella no está interesada en la asignación de probabilidad, asume que, de alguna forma, las probabilidades son asignadas a cada evento en la base de un cierto conocimiento a priori. También fue indicado que, si el modelo matemático es conforme al fenómeno real, debemos asignar estas probabilidades de una manera consistente con la realidad física. Es posible mostrar que la definición de probabilidad, según la frecuencia relativa, satisface los axiomas:

1.- Claramente se observa que $P(A) \geq 0$ ya que $N_a \geq 0$ y $N > 0$.

2.- $P(S) = 1$ ya que S ocurre en cada realización y de ahí $N_s = N$, entonces $P(S) = 1$.

3.- Si $AB = \{\Phi\}$, luego $N_{ab} = N_a + N_b$ ya que si A y B ocurre, luego A o B ocurre pero no ambos.

$$P(A + B) = P(A) + P(B)$$

Por tanto, es evidente que la teoría axiomática use la definición de frecuencia relativa, pero también es evidente que evita tomar responsabilidad por tal definición imprecisa.

Clase F de eventos

Los eventos de S son subsets a los cuales hemos asignado probabilidades. No serán considerados como eventos todos los subsets de S sino solamente una clase F de eventos. Dependiendo de la naturaleza de la aplicación, no incluir todos los subsets de S en la clase F de eventos es por razones matemáticas.

En ciertos casos, envolviendo sets con infinito número de resultados, es imposible asignar probabilidades a todos los subsets satisfaciendo todos los axiomas, incluyendo la versión generalizada del axioma 3 como veremos más adelante.

Dado que queremos saber no solamente las probabilidades de varios eventos sino también las probabilidades de sus "sumas" y "productos", nos lleva al concepto de lo que es un campo F .

Campos

Campo F es una clase de subsets no vacío que cumple con:

$$\text{Si } A \in F \text{ luego } \bar{A} \in F \quad (\text{C.1})$$

$$\text{Si } A \in F \text{ y } B \in F \text{ luego } A + B \in F \quad (\text{C.2})$$

Estas dos propiedades dan un set mínimo de condiciones para que F sea un campo. Todas las otras propiedades son una consecuencia de las dos anteriores.

Si $A \in F$ y $B \in F$; luego $AB \in F$

De (C.1.) sigue que $A \in F$ y $B \in F$

De (C.2.) se obtiene que $A+B \in F$

De (C.1.) aplicado a $A+B$ se obtiene que:

$$A+B \in F$$

$$AB \in F \quad (C.3.)$$

Un campo F contiene el evento certeza S y el evento imposible $\{\Phi\}$

$$S \in F \text{ y } \{\Phi\} \in F \quad (C.4.)$$

Por definición F es una clase de sets no vacíos, por lo tanto F contiene al menos un elemento A , esto es: $A \in F \Rightarrow \bar{A} \in F$ de (C.1.)

De esta manera a partir de (C.3.):

$$A\bar{A} \in F \text{ y } A+A \in F$$

$$\{\Phi\} \in F \text{ y } S \in F$$

Como conclusión de lo observado anteriormente se puede decir que todos los sets que pueden ser escritos como "sumas" o "productos" de sets finitos en F , están también en F esto, sin embargo, no necesariamente es el caso para set de infinito número de resultados.

Campo Borel

Supongamos que A_1, \dots, A_n, \dots es una secuencia infinita de sets en F . Si las sumas y productos de estos sets pertenecen también a F , luego F es llamado Campo Borel.

Eventos

En teoría de las probabilidades, eventos son ciertos subsets de S formando un campo Borel. Esto nos permite asignar probabilidades no solamente a "sumas" finitas y "productos" finitos sino también a sus límites.

Para la determinación de la probabilidad de sets que pueden ser expresados como límites, aplicación repetida del axioma 3, (1.5), es necesario lo que lleva a la conclusión de que si los eventos A_1, \dots, A_n son mutuamente exclusivos, luego:

$$P(A_1 + \dots + A_n) = P(A_1) + \dots + P(A_n)$$

La extensión del anterior resultado a infinitos sets no sigue directamente desde el axioma 3, es una condición adicional conocida con el nombre **axioma de la aditividad infinita**.

3.A) Si $A_1, A_2, \dots, A_n, \dots$ son mutuamente exclusivos, luego:

$$P(A_1 + A_2 + \dots) = P(A_1) + P(A_2) + \dots \quad (1.6)$$

y asumimos que todas las probabilidades, $P(A_i)$, satisfacen los axiomas: (1.3) - (1.6)

Definición axiomática de un experimento

En teoría de probabilidades, un experimento es especificado en términos de los siguientes conceptos:

- El set S de todos los resultados del experimento. (S)
- El campo de Borel de todos los eventos del espacio S . (B)
- Probabilidad de todos estos eventos. (P)

$$\implies (S, B, P) \quad S \iff \mathbb{R}^n$$

1.1.2 Probabilidad Condicional

Es la probabilidad de un evento A dado que es conocido que el evento M ocurrió. La probabilidad condicional, $P(A/M)$, no es posible derivarla a partir de los axiomas de ahí que debemos definirla; sin embargo, la definición debe ser

razonable y conforme a la observación física. La probabilidad condicional de un evento A asumiendo que M ocurrió, $P(A/M)$, es por definición la relación:

$$P(A/M) = P(AM)/P(M) \quad (1.7)$$

donde se asume que $P(M) > 0$

A partir de la interpretación de la frecuencia relativa, es posible mostrar este resultado: Sea el experimento con dos sets resultantes, a ser realizado N veces, de las cuales N_m realizaciones son favorables al set M . De estas N_m realizaciones, N_a son favorables al evento A . Se observa que N_a es el número de realizaciones favorables al evento conjunto AM (es decir ocurren A y M). De ahí se ven directamente las siguientes propiedades:

- Si $M \subset A$ luego $P(A/M) = 1$ ya que $AM = M$
- Si $A \subset M$ luego $P(A/M) = P(A) / P(M)$ ya que $AM = A$

A continuación se muestra que para un M específico, la probabilidad condicional cumple con los axiomas:

- $P(A/M) \geq 0$
- $P(S/M) = 1$
- $P(\{A+B\}/M) = P(A/M) + P(B/M)$

I.1.3 Probabilidad Total y Teorema de Bayes

Si $C = [A_1, \dots, A_n]$ es una partición de S y B es un evento cualquiera (ver figura), luego:

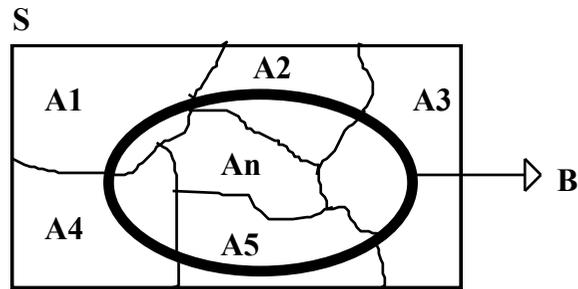


Figura 1.1 Partición de S

$$B = BS = B(A1 + \dots + An)$$

$$B = B.A1 + B.A2 + \dots + B.An$$

más, $B.A1, \dots, B.An$, son mutuamente excluyentes. Por lo tanto:

$$P(B) = P(BA1) + \dots + P(BAn)$$

Aplicando la probabilidad condicional a la expresión anterior tenemos el teorema de la probabilidad total:

$$P(B) = P(B/A1).P(A1) + \dots + P(B/An).P(An) \quad (1.8)$$

Ahora, si nuestro interés está en la probabilidad de cualquiera evento $\{Ai\}$ condicionado al $\{B\}$, tendríamos que:

$$P(Ai / B) = \frac{P(Ai.B)}{P(B)} = \frac{P(B.Ai)}{P(B)} = \frac{P(B / Ai).P(Ai)}{P(B)}$$

Substituyendo la expresión de la probabilidad total, (1.7), en la expresión anterior, resulta el Teorema de Bayes:

$$P(Ai / B) = \frac{P(B.Ai).P(Ai)}{\sum_{i=1}^n P(B / Ai).P(Ai)} \quad (1.9)$$

I.1.4. Independencia, sucesos independientes

Dos eventos A y B son llamados independientes si $P(AB) = P(A).P(B)$. El concepto de independencia es fundamental. Es este concepto que justifica el desarrollo matemático de probabilidad, no meramente como un tópico en teoría de medida, sino como una disciplina separada.

Independencia de tres eventos

Los eventos A_1, A_2, A_3 son llamados de mutuamente independientes si ellos son independientes en pares:

$$P(A_i.A_j) = P(A_i).P(A_j) \quad \text{con: } i=1,2,3 \quad j=1,2,3 \quad i \neq j$$

y
$$P(A_1A_2A_3) = P(A_1)P(A_2)P(A_3)$$

Generalización: Los eventos A_1, A_2, \dots, A_n son independientes si para cualquier $k < n$ de ellos son independientes y:

$$P(A_1A_2, \dots, A_n) = P(A_1)P(A_2) \dots P(A_n) \quad (1.10)$$

I.1.5. Experimentos repetidos

Experimentos combinados

Realicemos dos experimentos: - lanzamiento de un dado y el -lanzamiento de una moneda:

$$S_1 = \{f_1, \dots, f_6\} \quad \text{donde: } P_1\{f_i\} = 1/6$$

$$S_2 = \{c, s\} \quad \text{donde: } P_2\{c\} = P_2\{s\} = 1/2$$

Queremos encontrar la probabilidad de que obtengamos un dos (2) en el dado y una cara en la moneda. En la definición de independencia de dos eventos, estos eventos son subsets del mismo espacio S.

Para esto, debemos construir un espacio S teniendo como subsets los eventos "f2" "cara". Los dos experimentos son vistos como un simple experimento cuyos resultados son los pares $\gamma_1.\gamma_2$, donde γ_1 es una de las 6 caras y γ_2 es cara o sello. En este caso, γ_i representará un elemento arbitrario de S_i . El espacio S resultante consiste de 12 elementos.

$$S = \{f1.c, f2.c, \dots, f6.c, f1.s, f2.s, \dots, f6.s\}$$

En este espacio $\{f2\}$ y $\{c\}$ no son eventos elementales, mas si son subsets de S constituidos por:

$$\{f2\} = \{f2.s, f2.c\} \quad \{c\} = \{f1.c, \dots, f6.c\}$$

para completar el experimento debemos asignar probabilidades a los subsets de S .

$$P\{f2\} = 2/12 = 1/6 = P1\{f2\}$$

$$P\{c\} = 6/12 = 1/2 = P2\{c\}$$

Se observa que el evento $\{f2\}$ ocurre si el dado muestra un dos sin importar que muestre en la moneda. Similarmente el evento $\{c\}$. Observe que el producto de los dos eventos es el evento elemental $\{f2.c\}$ donde:

$$P\{f2.c\} = 1/12 = P1\{f2\}.P2\{c\}$$

Asumiendo que $\{f2\}$ y $\{c\}$ son independientes:

$$P\{f2.c\} = P\{f2\}.P\{c\}$$

Por lo tanto, dados dos experimentos representados por los espacios S_1 y S_2 con los elementos γ_1 y γ_2 , respectivamente, formamos todos los pares ordenados $\gamma_1\gamma_2$ donde γ_1 es cualquier elemento de S_1 y γ_2 es cualquier elemento de S_2 . El set S formado por tales pares el llamado **producto cartesiano** definido como:

$$S=S_1 \times S_2$$

- Si A es un subset de S_1 y B es un subset de S_2 , luego el set $C = A \times B$ constituido por todos los pares $\gamma_1\gamma_2$ donde γ_1 es elemento de A y γ_2 es elemento de B , es un subset de S .
- Similarmente los sets $A \times S_2$ y $S_1 \times B$. Se concluye que su producto es el set $A \times B$.

$$A \times B = (A \times S_2)(S_1 \times B)$$

Con esto concluimos que el producto cartesiano de dos experimentos S_1 y S_2 es un nuevo experimento $S = S_1 \times S_2$ cuyos eventos son todos los productos cartesianos de la forma $A \times B$, donde A es un evento del experimento S_1 y B es un evento del experimento S_2 y sus sumas y productos. Observe que en este experimento, las probabilidades de los eventos $A \times S_2$ y $S_1 \times B$ son tales que

$$P(A \times S_2) = P_1\{A\} \quad P(S_1 \times B) = P_2\{B\}$$

esto es, el evento $A \times S_2$ del experimento S ocurre si el evento A del experimento S_1 ocurre sin importar el resultado de S_2 . Así hemos determinado las probabilidades de los eventos $A \times S_2$ y $S_1 \times B$. Las probabilidades de los eventos $A \times B$ su suma y producto no pueden en forma general expresarse en términos de P_1 y P_2

Experimentos independientes

En muchas aplicaciones, los eventos $A \times S_2$ y $S_1 \times B$ del experimento combinado S son independientes. Para cualquier A y B concluimos que:

$$P(A \times B) = P[(A \times S_2)(S_1 \times B)] = P(A \times S_2)P(S_1 \times B)$$

$$P(A \times B) = P_1(A) \cdot P_2(B)$$

Esto completa la especificación del experimento S , puesto que todos sus eventos son sumas y productos de eventos de la forma $A \times B$.

Generalización

Sean n experimentos, S_1, \dots, S_n . Se define como producto cartesiano $S = S_1 \times S_2 \times \dots \times S_n$ el experimento cuyos elementos son todas las n -pletas ordenadas $\gamma_1, \dots, \gamma_n$ donde γ_i es un elemento del set S_i . Eventos en este espacio son todos los sets de la forma:

- $A_1 \times A_2 \times \dots \times A_n$ donde $A_i \subset S_i$

- Sumas de A_i

- Productos de A_i

Si los experimentos, S_i , son independientes y $P_i(A_i)$ es la probabilidad del evento A_i en el experimento S_i luego:

$$P(A_1 \times \dots \times A_n) = P_1(A_1) \times \dots \times P_n(A_n) \quad (1.11)$$

I.1.6. Experimento de Bernoulli

Es bien conocido en análisis combinatorial que si un set tiene n elementos, luego el número total de subsets constituidos de k elementos cada uno, es igual a:

$$\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{1.2\dots k} = \frac{n!}{k!(n-k)!}$$

Suceso o falla de un evento A en n realizaciones independientes

Tenemos un experimento S y un evento A con:

$$P(A) = p \quad P(\bar{A}) = q = 1-p$$

Repetimos el experimento n veces y el espacio del producto resultante denotamos por S^n

$$S^n = S \times S \times \dots \times S$$

Determinar la probabilidad, $P_n(k)$, de que el evento A ocurra k veces.

Teorema Fundamental

$$P_n(k) = P\{A \text{ ocurra } k \text{ veces en cualquier orden}\} = \binom{n}{k} p^k \cdot q^{n-k}$$

El evento $\{A \text{ ocurra } k \text{ veces en orden específico}\}$ es un producto cartesiano $B_1 \times \dots \times B_n$ donde k de los eventos B_i son iguales a A y los restantes $(n-k)$ son iguales a su complemento. Aplicando (1.10):

$$P(B_1 \times B_2 \times \dots \times B_n) = P_1(B_1) \cdot \dots \cdot P_n(B_n) = p^k \cdot q^{n-k}$$

ya que:
$$P_i(B_i) = \begin{cases} p \rightarrow \text{si } \dots B_i = A \\ q \rightarrow \text{si } \dots B_i = \bar{A} \end{cases}$$

en otras palabras $P\{A \text{ ocurra } k \text{ veces en un dado orden}\} = p^k q^{n-k}$

El evento $\{A \text{ ocurra } k \text{ veces en cualquier orden}\}$ es la unión (suma) de los $\binom{n}{k}$ eventos $\{A \text{ ocurre en un orden dado}\}$ y dado que son mutuamente exclusivos, concluimos que:

$$P_n(k) = P\{A \text{ ocurra } k \text{ veces en cualquier orden}\} = \binom{n}{k} p^k \cdot q^{n-k} \quad (1.12)$$

Generalización de los experimentos de Bernoulli [01, 14]

El experimento de realizaciones repetidas pueden ser explicadas en la siguiente forma: los eventos $A_1=A$ y $A_2=\bar{A}$ del espacio S forman una partición y sus respectivas probabilidades son iguales a $p_1=p$ y $p_2=1-p$. En el espacio S^n , la probabilidad del evento $\{A \text{ ocurre } k_1=k \text{ veces y } A_2 \text{ ocurre } k_2=n-k \text{ veces en cualquier orden}\}$ es igual a:

$$P_n(k) = P\{A \text{ ocurra } k \text{ veces en cualquier orden}\} = \binom{n}{k} p^k q^{n-k}$$

generalizando suponga que, $S = [A_1, \dots, A_r]$, es una partición de S constituido de los r eventos A_i con:

$$P(A_i) = p_i \quad \text{donde: } p_1 + \dots + p_r = 1$$

Repetimos el experimento n veces y llamamos a $P_n(k_1, \dots, k_r)$ la probabilidad del evento $\{A_1 \text{ ocurre } k_1 \text{ veces}, \dots, A_r \text{ ocurre } k_r \text{ veces en cualquier orden}\}$, donde:

$$k_1 + \dots + k_r = n$$

se mantiene que:

$$P_n(k_1, \dots, k_r) = \frac{n!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r} \quad (1.13)$$

I.2. Variables aleatorias

Introducción

Cuando un experimento es realizado, lo caracterizamos a través del espacio S y de los posibles resultados. Igualmente son definidos eventos del espacio S los que no son únicos. Estos eventos, como hemos visto, pueden ser descritos mediante bases descriptivas tales como cara y sello en el caso del lanzamiento de una moneda.

Matemáticamente, trabajar con frases es inconveniente y por lo tanto, para facilitar la manipulación podríamos asignar un número real a cada uno de los resultados mediante alguna regla prefijada. Por ejemplo: 1 si es cara y -1 si es sello.

Una VARIABLE ALEATORIA (v.a) es un número $X(\xi)$ asignando a cada resultado de un experimento. En general, asignamos un número real a cada resultado, si hay n resultados, asignamos n números $X(\xi_1) = x_1, \dots, X(\xi_n) = x_n$ a

I.2.2. Función distribución de probabilidad

Existe una probabilidad asignada a cada valor de la v.a X . Así, $P_X(x_i)$ es la probabilidad del evento al cual x_i fue asignado. Si hay un total de n resultados mutuamente exclusivos del experimento, luego de los axiomas sigue que:

$$\sum P_X(x_i) = 1$$

Ejemplo:

Lanzar tres monedas. El evento deseado es obtener un total de k caras. Hay 4 posibles resultados: 1) Obtener 0 caras, 2) una cara, 3) dos caras, 4) tres caras. Asignamos 4 números x_0, x_1, x_2, x_3 , a los eventos anteriores, respectivamente.

$$\xi_1 = \{\text{Obtener 0 caras}\} \quad X(\xi_1) = x_0 \quad \text{donde: } P_X(x_0) = 1/8$$

$$\xi_2 = \{\text{Obtener 1 cara}\} \quad X(\xi_2) = x_1 \quad \text{donde: } P_X(x_1) = 3/8$$

$$\xi_3 = \{\text{Obtener 2 caras}\} \quad X(\xi_3) = x_2 \quad \text{donde: } P_X(x_2) = 3/8$$

$$\xi_4 = \{\text{Obtener 3 caras}\} \quad X(\xi_4) = x_3 \quad \text{donde: } P_X(x_3) = 1/8$$

$$\text{Se verifica que: } \sum P_X(x_i) = 1$$

Los elementos del set S que están contenidos en el evento $\{X \leq x\}$ cambian cuando el número x toma diferentes valores. La probabilidad, $P\{X \leq x\}$, del evento $\{X \leq x\}$ es, por lo tanto, un número que depende de x . Este número es denotado por $F_X(x)$, y es llamado la FUNCIÓN DISTRIBUCIÓN de la variable aleatoria X .

$$F_X(x) = P\{X \leq x\} \quad \forall x \quad (1.14)$$

Ejemplo:

Lanzar una moneda con $P\{c\} = p$, $P\{s\} = q = 1-p$. Definimos la v.a X tal que: $X(c)=1$, $X(s)=0$. La función distribución de X es mostrada en la figura.

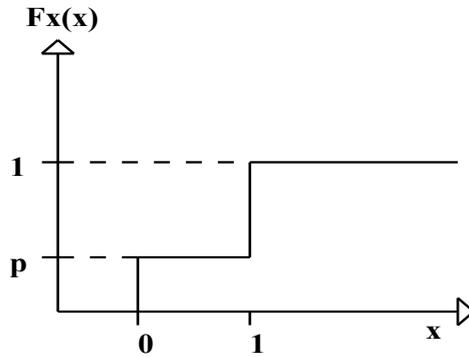


Figura 1.2 Función Distribución del lanzamiento de una moneda

Ejemplo:

En el lanzamiento de un dado, la v.a X es tal que $X(f_i) = 10i$. Donde $f_1, f_2, f_3, f_4, f_5, f_6$, representan los posibles resultados del experimento. Esto es:

$$\begin{array}{ll} X(f_1)=10 = x_1 & P_x(x_1) = P\{X = x_1\} = 1/6 \\ X(f_2)=20 = x_2 & P_x(x_2) = P\{X = x_2\} = 1/6 \\ X(f_3)=30 = x_3 & P_x(x_3) = P\{X = x_3\} = 1/6 \\ X(f_4)=40 = x_4 & P_x(x_4) = P\{X = x_4\} = 1/6 \\ X(f_5)=50 = x_5 & P_x(x_5) = P\{X = x_5\} = 1/6 \\ X(f_6)=60 = x_6 & P_x(x_6) = P\{X = x_6\} = 1/6 \end{array}$$

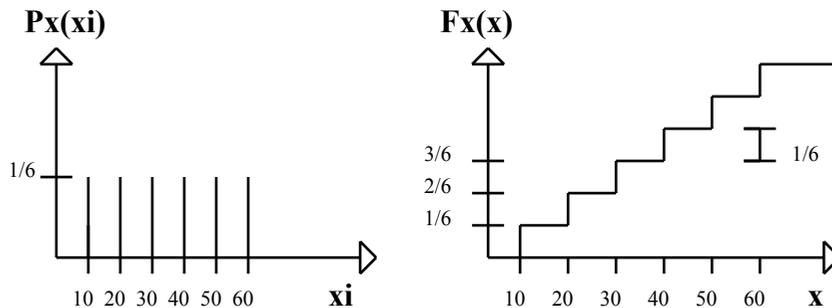


Figura 1.3 Función Distribución del lanzamiento de un dado

Propiedades de la Función Distribución

$$1) F(+\infty) = 1 \quad \text{y} \quad F(-\infty) = 0$$

$$F(+\infty) = P\{X \leq \infty\} = P\{S\} = 1$$

$$F(-\infty) = P\{X \leq -\infty\} = P\{\Phi\} = 0$$

2) Es una función no decreciente en x . Si $x_1 < x_2$ entonces, $F_X(x_1) < F_X(x_2)$.

$\{X \leq x_2\}$ es un subconjunto de $\{X \leq x_1\}$, por tanto:

$$F_X(x_1) < F_X(x_2)$$

3) Si $F_X(x_0) = 0$ luego $F_X(x) = 0$ para todo $x \leq x_0$.

Si $X(\xi) \geq x_0$ para todo ξ , entonces, $F_X(x_0) = P\{X \leq x_0\} = 0$ ya que $\{X \leq 0\}$ es el evento imposible.

4) $P\{X > x\} = 1 - F_X(x)$. $\{X \leq x\}$ y $\{X > x\}$ son mutuamente exclusivos:

$$\{X \leq x\} + \{X > x\} = S.$$

5) Utilizaremos la nomenclatura:

$$F_X(x^+) = \lim_{\varepsilon \rightarrow 0^+} [F_X(x + \varepsilon)]$$

$$F_X(x^-) = \lim_{\varepsilon \rightarrow 0^+} [F_X(x - \varepsilon)]$$

La función $F_X(x)$ es continua desde la derecha, esto es: $F_X(x^+) = F_X(x)$

6) $P\{x_1 < X \leq x_2\} = F_X(x_2) - F_X(x_1)$

Los eventos $\{X \leq x_1\}$ y $\{x_1 < X \leq x_2\}$ son mutuamente exclusivos, por tanto:

$$\{X \leq x_2\} = \{X \leq x_1\} + \{x_1 < X \leq x_2\}$$

$$P\{X \leq x_2\} = P\{X \leq x_1\} + P\{x_1 < X \leq x_2\}$$

$$P\{x_1 < X \leq x_2\} = P\{X \leq x_2\} - P\{X \leq x_1\}$$

$$P\{x_1 < X \leq x_2\} = F_X(x_2) - F_X(x_1).$$

$$7) P\{X = x\} = F_X(x) - F_X(x^-)$$

seteando $x_1 = x - \varepsilon$ y $x_2 = x$

$$P\{x - \varepsilon < X \leq x\} = F_X(x) - F_X(x - \varepsilon)$$

$$P\{X = x\} = F_X(x) - F_X(x^-)$$

$$8) P\{x_1 \leq X \leq x_2\} = F_X(x_2) - F_X(x_1^-)$$

Los eventos $\{x_1 \leq X \leq x_2\} = \{x_1 < X \leq x_2\} + \{X = x_1\}$ son mutuamente exclusivos:

$$\begin{aligned} P\{x_1 \leq X \leq x_2\} &= P\{x_1 < X \leq x_2\} + P\{X = x_1\} \\ &= F_X(x_2) - F_X(x_1) + F_X(x_1) - F_X(x_1^-) \\ &= F_X(x_2) - F_X(x_1^-) \end{aligned}$$

Una v.a X es del tipo continuo si su función distribución $F_X(x)$ es continua. Es interesante notar que si S tiene un número finito de elementos, luego cualquier v.a definido en S será del tipo discreto; sin embargo, una v.a X podría ser del tipo discreto aún si S tuviese infinito número de elementos.

I.2.3. Función Densidad de Probabilidad

Conocida también como **Función Frecuencia** es definida como:

$$f_X(x) = \frac{d}{dx}[F_X(x)] \quad (1.15)$$

Si la v.a es del tipo discreta tomando los valores x_i con probabilidades p_i , luego:

$$f_X(x) = \sum p_i \delta(x - x_i) \quad (1.16)$$

donde: $p_i = P\{X=x_i\}$ y $\delta(x)$ es la función Delta Dirac.

Propiedades de la Función Densidad de Probabilidad

Desde la monotonidad de $F_X(x)$ se obtiene que:

$$f_X(x) \geq 0$$

A partir de:

$$f_X(x) = \frac{d}{dx}[F_X(x)] \quad \text{luego:}$$

$$F_X(x) = \int_{-\infty}^x f_X(\alpha) d\alpha \quad \text{ya que } F_X(\infty) = 1$$

$$\int_{-\infty}^{+\infty} f_X(\alpha) d\alpha = 1$$

De la deducción anterior se observa que:

$$P\{X \leq x\} = F_X(x) = \int_{-\infty}^x f_X(\alpha) d\alpha \quad (1.17)$$

Teorema de la Existencia

Dado una función $f_X(x)$ o su integral, $F_X(x) = \int_{-\infty}^x f_X(\alpha) d\alpha$, podemos construir un experimento y una v.a X con distribución $F_X(x)$ o densidad $f_X(x)$, considerando como nuestro espacio S el set de todos los números reales y como sus eventos todos los intervalos en la línea real y sus uniones e intersecciones. Se define la probabilidad de un evento $\{X \leq x_1\}$ por:

$$P\{X \leq x_1\} = F_X(x_1)$$

Los resultados de nuestro experimento son los números reales. Para definir una v.a X en este experimento, debemos conocer su valor $X(x)$ para cualquier x . Defínese X tal que:

$$X(x) = x$$

Así la v.a X es el mismo resultado del experimento (caso muy utilizado en la práctica). Se dice que la función distribución de X es igual a $F_X(x)$. De hecho, el evento $\{X \leq x_1\}$ consiste de todos los resultados x tal que $X(x) \leq x_1$. De ahí:

$$P\{X \leq x_1\} = F_X(x_1)$$

Variable Aleatoria Gaussiana

Una v.a. es gaussiana si su función densidad de probabilidad es dada por la siguiente expresión:

$$f_X(x) = \frac{1}{\sigma_X \sqrt{2\pi}} \exp\left\{-\frac{(x - m_X)^2}{2\sigma_X^2}\right\}$$

donde: m_X es denominado centro de simetría.

σ_X^2 es denominado esparcimiento de la v.a.

Definamos:

$$g(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp(-x^2 / 2) \quad \Rightarrow \quad f_X(x) = \frac{1}{\sigma_X} g[(x - m_X) / \sigma_X]$$

$$\text{Si: } G(x) = \int_{-\infty}^x g(\alpha) d\alpha \quad \Rightarrow \quad G(x) = \text{erf}(x) + 1/2$$

donde la función $\text{erf}(x)$ se encuentra tabulada.

Variable Aleatoria Uniforme

Una v.a. se dice que es uniforme entre x_1 y x_2 si su función densidad de probabilidad es:

$$f_X(x) = 1/(x_2-x_1) \quad \text{para } x_1 \leq X \leq x_2$$

$$f_X(x) = 0 \quad \text{otros } X$$

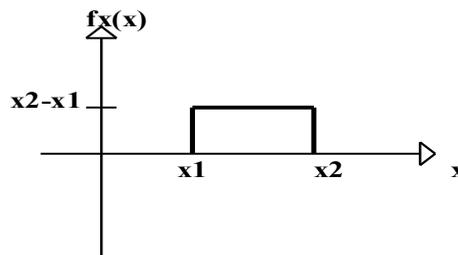


Figura 1.4 Función Densidad de Probabilidad Uniforme

Variable Aleatoria Binomial:

Se dice que una v.a. X , tiene distribución binomial de orden n si toma los valores $0, 1, 2, 3, \dots, n$ con probabilidad:

$$P\{X=k\} = \binom{n}{k} p^k \cdot q^{n-k} \quad \text{donde: } p+q=1$$

1.2.4. Distribuciones Condicionales

Recordemos que la probabilidad de un evento A condicionado a la ocurrencia de otro evento M esta dada por:

$$P(A/M) = P(AM)/P(M) \quad P(M) \neq 0$$

La *distribución condicional* $F_X(x/M)$ de una v.a. X condicionada al evento M es definida como la probabilidad condicional del evento $\{X \leq x/M\}$

$$F_X(x/M) = P(X \leq x / M) = P\{(X \leq x).M\} / P(M) \quad (1.18)$$

En la expresión anterior, $\{(X \leq x, M)\}$ es el producto de los eventos $\{X \leq x\}$ y M , esto es, el evento constituido por todos los resultados ξ tal que $X(\xi) \leq x$ y $\xi \in M$. Es por tanto, que $F_x(x/M)$ tiene las mismas propiedades que $F_x(x)$:

$$F_x(\infty/M) = 1$$

$$F_x(-\infty/M) = 0$$

$$P\{x_1 < x \leq x_2 / M\} = F_x(x_2/M) - F_x(x_1/M) = P\{(x_1 < x \leq x_2), M\} / P(M)$$

La *función densidad condicional*, $f_x(x/M)$ es dada por:

$$f_x(x/M) = dF_x(x/M) / dx \quad (1.19)$$

1.2.5. Probabilidad Total y Teorema de Bayes

Extenderemos los resultados ya obtenidos en ítems anteriores. Es bien conocido que:

$$P(B) = P(A_1B) + P(A_2B) + \dots + P(A_nB)$$

$$P(B) = P(B/A_1)P(A_1) + \dots + P(B/A_n)P(A_n)$$

Supongamos la v.a X asignada a los resultados ξ de un dado experimento, tal que:

$$B = \{X \leq x\}$$

$$P\{X \leq x\} = P\{X \leq x/A_1\}P\{A_1\} + \dots + P\{X \leq x/A_n\}P\{A_n\}$$

$$F_x(x) = F_x(x/A_1)P\{A_1\} + \dots + F_x(x/A_n)P\{A_n\} \quad (1.19)$$

$$f_x(x) = f_x(x/A_1)P\{A_1\} + \dots + f_x(x/A_n)P\{A_n\} \quad (1.20)$$

Igualmente de:

$$P(A/M) = P(AB)/P(B) = P(BA)/P(B) = P(B/A)P(A)/P(B)$$

Con $B = \{X \leq x\}$ donde $X(\xi)$ es una v.a con $\xi \in B$

$$P(A/\{X \leq x\}) = P(X \leq x / A)P(A)/P(B)$$

$$P(A/X \leq x) = F_x(X \leq x/A)P(A)/P(B) = F_x(X \leq x/A)P(A)/F_x(x) \quad (1.21)$$

Para la función densidad de probabilidad, consideremos el siguiente evento B:

$$B = \{x_1 < X \leq x_2\}$$

$$P\{A/x_1 < X \leq x_2\} = P\{x_1 < X \leq x_2/A\}P(A)/P\{x_1 < X \leq x_2\}$$

$$P\{A/x_1 < X \leq x_2\} = (F_x(x_2/A) - F_x(x_1/A))P(A)/(F_x(x_2) - F_x(x_1))$$

La $P(A/X=x)$ del evento A asumiendo que $X=x$ no puede ser determinada directamente ya que $P\{X=x\} = 0$. Para sobrellevar este inconveniente, definimos:

$$P\{A/X=x\} = \lim_{\Delta x \rightarrow 0} P\{A/x_1 \leq X \leq x_1 + \Delta x\}$$

Con $x_1 = x$ y $x_2 = x_1 + \Delta x$

$$P\{A/X=x\} = P(A/x_1 \leq X \leq x_1)$$

$$P\{A/X=x\} = (F_x(x_2/A) - F_x(x_1/A))P(A)/(F_x(x_2) - F_x(x_1))$$

$$P\{A/X=x\} = \lim_{\Delta x \rightarrow 0} (F_x(x_1 + \Delta x/A) - F_x(x_1/A))P(A)/(F_x(x_1 + \Delta x) - F_x(x_1))$$

$$P\{A/X=x\} = f_x(x/A)P(A)/f_x(x) \quad (1.22)$$

I.2.6. Funciones de una Variable Aleatoria

Supongamos que X es una v.a y que $g(X)$ es una función de la variable aleatoria X . La expresión, $Y=g(X)$, es una v.a definida de la siguiente manera: Para un dado resultado ξ del experimento, $X(\xi)$ es un número y $g[X(\xi)]$ es otro número especificado en términos de $X(\xi)$ y $g(X)$.

Este número es el valor $Y(\xi) = g[X(\xi)] = y$. Así:

$$F_Y(y) = P\{Y \leq y\} = P\{g(X) \leq y\}$$

Para un valor específico, y , los valores de X tales que $g(X) = y$ forman un set en el eje x denotado por S_y . Claramente $g[X(\xi)] = y$ si $X(\xi)$ es un número en el set S_y . De ahí que:

$$F_Y(y) = P\{X \in S_y\}$$

Determinación de $f_Y(y)$:

Deseamos determinar la densidad de $Y = g(X)$ en función de la densidad de X , y de $g(X)$. Supongamos primero que el set S del eje y no está en el rango de la función $g(X)$, esto es, que $g(X)$ no es un punto de S para cualquier x . En este caso, la probabilidad de que $g(X)$ este en S es igual a 0. De ahí, $f_Y(y) = 0$ para $y \in S$. Es suficiente por tanto, considerar los valores de y tales que para algún x , $g(x) = y$.

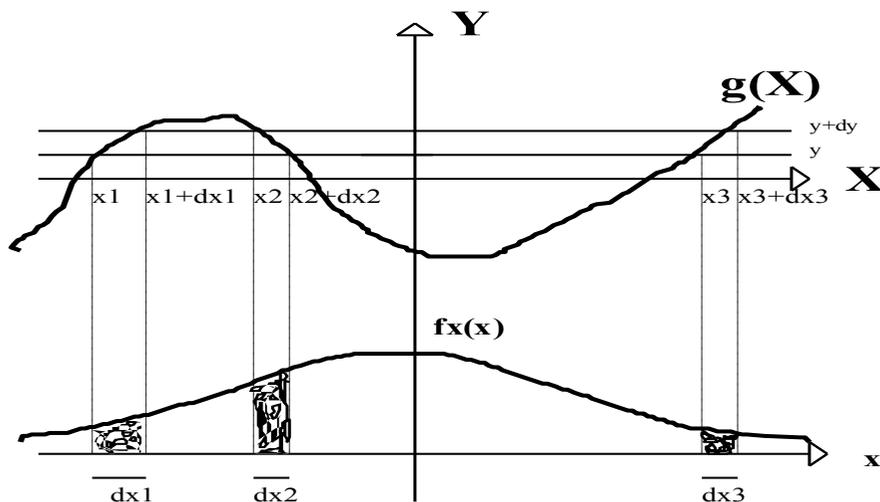


Figura 1.5 Relación entre X y Y

Para encontrar $f_y(y)$, resolveremos la ecuación $y = g(x)$. Denotando sus raíces reales por: x_1, x_2, \dots, x_n , esto es:

$$y = g(x_1) = \dots = g(x_n)$$

Entonces:

$$f_y(y) = f_x(x_1)/g'(x_1) + \dots + f_x(x_n)/g'(x_n) \quad (1.23)$$

donde: $g'(x) = dy/dx$

Sabemos que:

$$P\{y < Y \leq y+dy\} = f_y(y)dy$$

Es suficiente, por tanto, encontrar el set de valores de x tal que $y < g(x) \leq y+dy$ y la probabilidad de que x se encuentre en este set. Como vemos en los gráficos, este set consiste de los intervalos:

$$x_1 < X \leq x_1+dx_1 \quad \text{donde: } dx_1 > 0$$

$$x_2+dx_2 < X \leq x_2 \quad \text{donde: } dx_2 < 0$$

$$x_3 < X \leq x_3+dx_3 \quad \text{donde: } dx_3 > 0$$

$$P\{y < Y \leq y+dy\} = P\{x_1 < X \leq x_1+dx_1\} + P\{x_2+dx_2 < X \leq x_2\} + P\{x_3 < X \leq x_3+dx_3\}$$

$$P\{x_1 < X \leq x_1+dx_1\} = f_x(x_1)dx_1 \quad \text{donde: } dx_1 = dy/g'(x_1)$$

$$P\{x_2+dx_2 < X \leq x_2\} = f_x(x_2)dx_2 \quad \text{donde: } dx_2 = dy/g'(x_2)$$

$$P\{x_3 < X \leq x_3+dx_3\} = f_x(x_3)dx_3 \quad \text{donde: } dx_3 = dy/g'(x_3)$$

Por lo tanto:

$$f_y(y) = f_x(x_1)/g'(x_1) + \dots + f_x(x_n)/g'(x_n)$$

$$\text{donde: } g'(x) = dy/dx$$

Media de una Variable Aleatoria

El valor esperado o media de una variable aleatoria X , $E[X]$, es por definición:

$$E[X] = \int_{-\infty}^{+\infty} x \cdot f_x(x) dx = mx \quad (1.24)$$

Este número será notado por mx .

La definición propone que la variable aleatoria X es continua. En el caso de X sea una variable aleatoria discreta, el integral se torna en un sumatorio. Supongamos que X toma los valores x_i con probabilidades p_i , en este caso:

$$f_x(x) = \sum p_i \cdot \delta(x - x_i)$$

$$\text{donde: } p_i = P\{X=x_i\}$$

$$E[X] = \int x \cdot [\sum p_i \cdot \delta(x - x_i)] \cdot dx$$

$$E[X] = \sum [p_i] \cdot \int x \cdot \delta(x - x_i) \cdot dx$$

$$E[X] = \sum x_i \cdot p_i \quad (1.25)$$

Ahora, la media condicional de una variable aleatoria X asumiendo que está condicionada a un evento M , está dado por:

$$E[X/M] = \int x \cdot f_x(x/M) dx \quad (1.26)$$

Para v. a. discretas

$$E\{X/M\} = \sum x_i \cdot P\{X = x_i / M\} \quad (1.27)$$

Dada una v. a. X y una función $Y = g(X)$, tenemos que el valor medio de $g(X)$ es:

$$E[Y] = E[g(X)] = \int g(x) \cdot f_X(x) dx \quad (1.28)$$

Pero cuando X es una v. a. discreta:

$$E[g(X)] = \sum g(x_i) \cdot P\{X = x_i\} \quad (1.28)$$

Si $Y = a_1 \cdot g_1(X) + \dots + a_n \cdot g_n(X)$

$$E[Y] = a_1 \cdot E[g_1(X)] + \dots + a_n \cdot E[g_n(X)]$$

Si $Y = K$ siendo K una constante

$$E[Y] = \int K \cdot f_Y(y) dy$$

$$E[Y] = K \cdot \int f_Y(y) dy$$

$$E[Y] = K$$

Si $X = X_i + jX_q$ es una v. a. de valor complejo, luego:

$$E[X] = E[X_i] - jE[X_q]$$

y si $g(X) = g(X_i) + jg(X_q)$

$$E[g(X)] = E[g(X_i)] + jE[g(X_q)]$$

Varianza de una Variable Aleatoria

La varianza de una v. a. X es por definición:

$$\text{var}(X) = \sigma^2 = E[(X - mx)^2] \quad (1.30)$$

$$\text{var}(X) = \sigma^2 = \int (x - mx)^2 \cdot f_X(x) dx$$

σ^2 es siempre ≥ 0 y el valor σ es conocido como la DESVIACION STANDARD. Si la variable aleatoria X es del tipo discreta, luego:

$$\sigma^2 = \sum p_i (x_i - mx)^2 \quad \text{donde: } p_i = P\{X=x_i\} \quad (1.31)$$

Una medida de la concentración de una v.a. alrededor de su valor medio es la varianza. De hecho, la probabilidad que X esté fuera de un intervalo arbitrario $(mx - \epsilon, mx + \epsilon)$ es despreciable si la relación σ/ϵ es suficientemente pequeña. Este resultado es conocido como DESIGUALDAD DE CHEBYCHEFF. Es posible mostrar que:

$$\forall \epsilon > 0 \quad P\{|X - mx| \geq \epsilon\} \leq (\sigma/\epsilon)^2 \quad (1.32)$$

Ejemplo:

Si tenemos una variable aleatoria X con fdp gaussiana:

$$f_X(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left\{-\frac{(x - mx)^2}{2\sigma^2}\right\}$$

Es posible mostrar que:

$$E[X] = mx \quad \text{y} \quad \text{var}(X) = \sigma x^2$$

Ejemplo:

Una variable aleatoria distribuída POISSON de parámetro α toma valores $0, 1, \dots$, con probabilidad dada por:

$$P\{X = k\} = \frac{\exp(-\alpha) \cdot \alpha^k}{k!}$$

$$E[X] = mx = \alpha$$

$$\text{var}(X) = \sigma x^2 = \alpha$$

Momentos

Los momentos son estadísticas de las variables aleatorias de orden general n definidos como:

$$E[x^n] = \int x^n \cdot f_x(x) dx \quad (1.33)$$

De gran interés es el momento de segundo orden, denominado **Valor Cuadrático Medio** e igual a $E[X^2]$

De igual manera, se definen los llamados momentos centrales dados por:

$$E[(X - mx)^n] = \int (x - mx)^n \cdot f_x(x) dx \quad (1.34)$$

Por último, se define la función característica de una variable aleatoria mediante la integral de Fourier:

$$\Phi(\omega) = \int f_x(x) \cdot \exp(j\omega x) dx = \mathfrak{F}[f_x(x)] \quad (1.35)$$

I.2.7. Estadísticas Conjuntas de dos Variables Aleatorias

Se considera dos variables aleatorias X y Y , siendo el objetivo determinar sus estadísticas conjuntas, esto es, la probabilidad que el punto (x,y) esté en una región D en el plano xy .

Las funciones distribución $F_X(x)$ y $F_Y(y)$ de las variables aleatorias X y Y , respectivamente, determinan sus estadísticas separadas o marginales, pero no sus estadísticas conjuntas. Por lo tanto, la probabilidad del evento conjunto: $\{X \leq x, Y \leq y\}$ no puede ser expresado en términos de $F_X(x)$ y $F_Y(y)$.

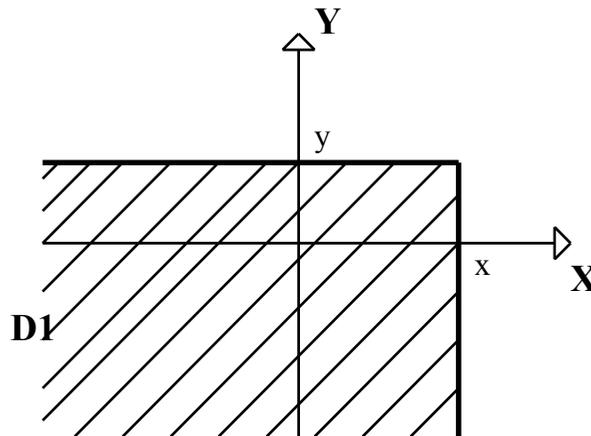


Figura 1.6 Región de las variables aleatorias $X \leq x, Y \leq y$

La distribución conjunta, $F_{XY}(x,y)$, de las dos variables aleatorias X y Y es la probabilidad del evento :

$$\{X \leq x; Y \leq y\} = \{(x, y) \in D1\}$$

donde x, y son dos números reales arbitrarios y $D1$ es el cuadrante mostrado en la figura. Entonces:

$$F_{XY}(x,y) = P\{X \leq x, Y \leq y\} \quad (1.36)$$

Propiedades:

$$F(-\infty, y) = 0$$

$$F(x, -\infty) = 0$$

$$F(\infty, \infty) = 1$$

Sean los eventos:

$$\{x_1 < X \leq x_2, Y \leq y\} \quad \{X \leq x, y_1 < Y \leq y_2\}$$

Entonces: $P\{x_1 < X \leq x_2, Y \leq y\} = F_{xy}(x_2, y) - F_{xy}(x_1, y)$

$$P\{X \leq x, y_1 < Y \leq y_2\} = F_{xy}(x, y_2) - F_{xy}(x, y_1)$$

Igualmente la función densidad de probabilidad conjunta $f_{xy}(x, y)$ es definida como:

$$f_{xy}(x, y) = \frac{\partial^2 [F_{xy}(x, y)]}{\partial x \partial y} \quad (1.37)$$

A partir de esta última relación, tenemos que:

$$P\{X \leq x, Y \leq y\} = F_{xy}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{xy}(\alpha, \beta) d\alpha d\beta \quad (1.38)$$

esto es, que la probabilidad de que el punto (x, y) esté en la región D del plano xy es igual al integral de $f_{xy}(x, y)$ en D . Dicho en otras palabras:

$$P\{(x, y) \in D\} = \iint_D f_{xy}(x, y) dx dy$$

donde, $\{(x, y) \in D\}$, es el evento constituido por todos los resultados ξ tal que los puntos $[X(\xi), Y(\xi)]$ esten en la region D .

Hemos dicho que $F_x(x)$ es denominada distribución marginal y $f_x(x)$ es la densidad marginal. Es posible determinar las distribuciones marginales, a partir de la distribución conjunta, esto es:

$$F_x(x) = F_{xy}(x, \infty) \quad F_y(y) = F_{xy}(\infty, y)$$

$$\begin{aligned} f_x(x) &= \int_{-\infty}^{+\infty} f_{xy}(x, y) dy \\ f_y(y) &= \int_{-\infty}^{+\infty} f_{xy}(x, y) dx \end{aligned} \tag{1.39}$$

Independencia

Dos variables aleatorias, X y Y , son estadísticamente independientes si los eventos $\{X \in A\}$ y $\{Y \in B\}$ son independientes. Esto es, si se da que:

$$P\{X \in A, Y \in B\} = P\{X \in A\} \cdot P\{Y \in B\}$$

donde A y B son dos sets arbitrarios en los ejes X y Y , respectivamente. Aplicando lo anterior a los eventos $\{X \leq x\}$ y $\{Y \leq y\}$ se concluye que, si X , Y son independientes entonces será:

$$F_{xy}(xy) = F_x(x) F_y(y) \tag{1.40}$$

$$f_{xy}(xy) = f_x(x) f_y(y) \tag{1.41}$$

Momentos Conjuntos

Sean X y Y dos variables aleatorias, su momento conjunto es definido como:

$$E[X^j \cdot Y^k] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^j \cdot y^k \cdot f_{xy}(x, y) \cdot dx \cdot dy \tag{1.42}$$

Correlación:

Un momento conjunto de particular interés es el definido como correlación, dado cuando $j = k = 1$.

$$\text{Corr}(XY) = E[XY] = \iint x.y.f_{xy}(x,y).dx dy \quad (1.43)$$

Covarianza:

La correlación de las variables aleatorias X y Y centradas en torno de su valor medio es denominado de covarianza de X y Y.

$$\text{Cov}(XY) = E[\{X - m_x\} \cdot \{Y - m_y\}] = E[XY - m_y \cdot X - m_x \cdot Y + m_x \cdot m_y] \quad (1.44)$$

$$\text{Cov}(XY) = E[XY] - m_x \cdot m_y$$

$$\text{Cov}(XY) = \text{Corr}(XY) - m_x \cdot m_y$$

Coefficiente de Correlación

Siendo σ_x^2 y σ_y^2 las varianzas de X y Y, respectivamente; la covarianza de X y Y normalizada con respecto a $\sigma_x \cdot \sigma_y$ es el denominado *coeficiente de correlación*

$$\rho_{xy} = \frac{\text{Cov}(XY)}{\sigma_x \cdot \sigma_y} \quad (1.44.a)$$

Dos v.a. X y Y son descorrelacionadas si y sólo si su covarianza es cero:

$$\text{Cov}(XY) = 0$$

Dos v.a. X y Y son ortogonales si y sólo si su correlación es cero:

$$E[xy] = 0$$

Si una o ambas v.a. tienen media cero y si son ortogonales, luego X y Y son descorrelacionadas y viceversa.

I.3 Procesos Estocásticos

I.3.1. Introducción

En los sistemas de comunicación una caracterización de *señales aleatorias* tales como señales de voz, señales de televisión, datos de un computador digital, ruido del canal, ruido de receptor, es importante.

Estas señales aleatorias tienen dos propiedades: 1) las señales son funciones del tiempo definidas para algún intervalo de observación y 2) Las señales son aleatorias en el sentido en que antes de realizar el experimento, no es posible describir la forma de onda que será observada, esto es, que el ensamble del espacio constituido por funciones temporales es llamado de proceso estocástico o aleatorio.

Podemos definir un PE como un conjunto (ensamble) de funciones temporales junto con una regla de probabilidades que asigna una probabilidad a cualquier evento significativo asociado con una observación de una de estas funciones.

Consideremos un proceso estocástico $X(t)$ representado por el ensamble $\{X_i(t), i = 1, 2, \dots\}$ como se ilustra en la figura. La función muestra $X_i(t)$ con probabilidad de ocurrencia $P(s_i)$ corresponde al punto muestra s_i en el espacio S .

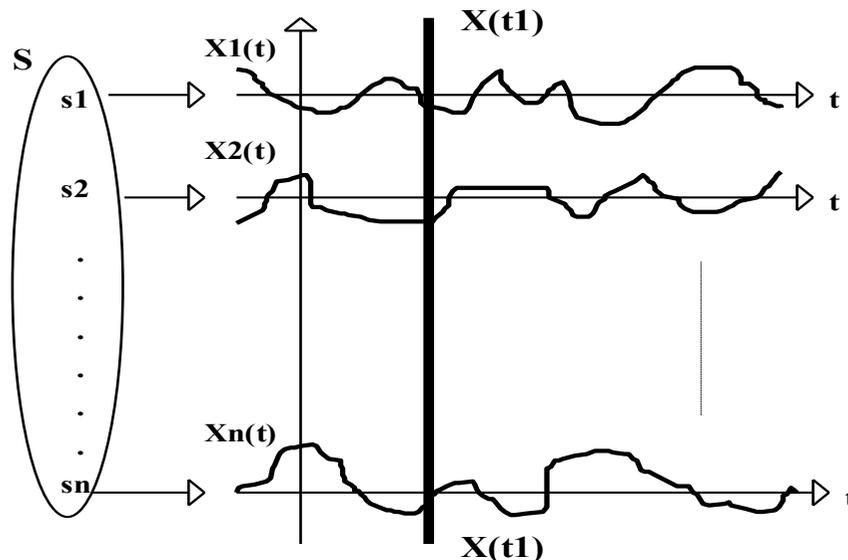


Figura 1.7 Ilustración de varias realizaciones de un Proceso Estocástico

Observemos el ensamble de las formas de onda a un instante $t = t_1$. Dado que cada uno de los eventos s_j del espacio S tiene asociado con ello un número $X_j(t_1)$ y una probabilidad $P(s_j)$, podemos decir que $\{X_j(t_1), j = 1, \dots\}$ forma una variable aleatoria $X(t_1)$. Por tanto, la diferencia entre Proceso Estocástico (PE) y v.a. es que para una v.a. el resultado de un experimento es mapeado en un número, mientras que para un PE el resultado es mapeado en una forma de onda que es función del tiempo.

Vector Aleatorio obtenido a partir de Procesos Estocásticos

Un PE $X(t)$ implica un número infinito de v.a. una para cada t ($-\infty < t < \infty$). Podemos hallar una $F_X(x_1)$ de la v.a. $X(t_1)$ obtenida al observar $X(t)$ en $t=t_1$. En general, para k instantes t_1, t_2, \dots, t_k , definimos el evento conjunto:

$$\{X(t_1) \leq x_1, \dots, X(t_k) \leq x_k\}$$

donde: $X(t_i)$ es una v.a. tomada del PE $X(t)$ al instante t_i

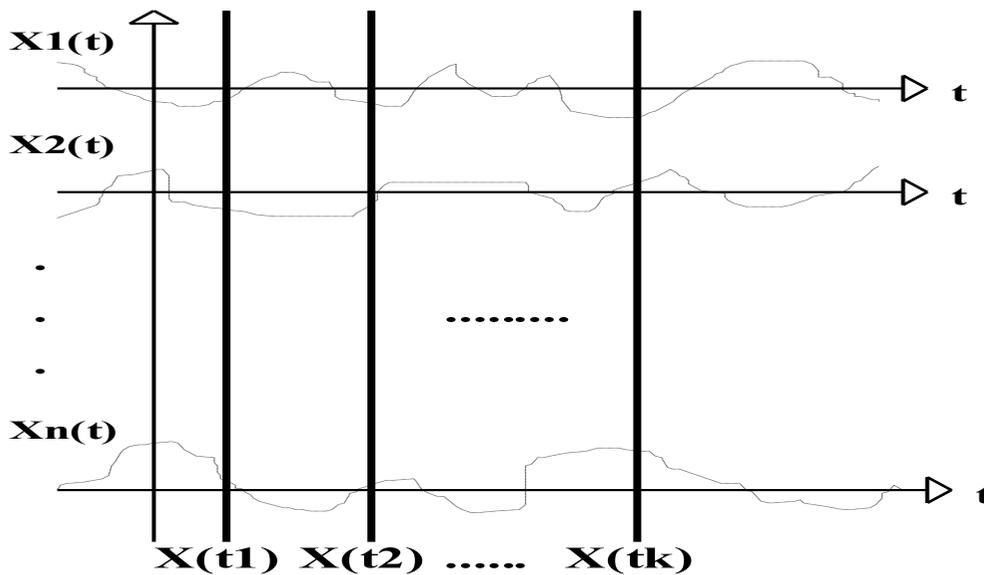


Figura 1.8 Proceso Estocástico caracterizado por k variables aleatorias

La probabilidad de este evento define la función distribución conjunta

$$F_{\underline{X}(t)}(\underline{x}) = P\{X(t_1) \leq x_1, \dots, X(t_k) \leq x_k\} \tag{1.44}$$

donde:

$$\underline{X}(t) = [X(t_1)X(t_2)\dots X(t_k)]^T$$

$$\underline{x} = [x_1, x_2, \dots, x_k]^T$$

La función densidad de probabilidad del vector aleatorio $X(t)$ es por tanto:

$$f_{\underline{X}(t)}(\underline{x}) = \frac{\partial^k}{\partial x_1 \partial x_2 \dots \partial x_k} F_{\underline{X}(t)}(\underline{x}) \quad (1.45)$$

I.3.2. Procesos Estocásticos Estacionarios

Considere un set de instantes t_1, t_2, \dots, t_k , en el cual un PE es definido. Una caracterización completa del PE $X(t)$ nos habilita a especificar la función densidad de probabilidad conjunta $f_{\underline{X}(t)}(\underline{x})$. $X(t)$ es ESTRICTAMENTE ESTACIONARIO si su fdp conjunta, $f_{\underline{X}(t)}(\underline{x})$, es invariante a los desplazamientos del origen del tiempo. Esto es:

$$f_{\underline{X}(t)}(\underline{x}) = f_{\underline{X}(t-T_1)}(\underline{x}) \quad (1.46)$$

PE Estacionarios (PEE) son importantes por dos razones:

1. Son frecuentemente encontrados en la práctica o aproximados con un alto grado de precisión. Observe que no es necesario que un PE sea estacionario todo el tiempo sino para un dado intervalo de observación.
2. Muchas de las propiedades importantes de los PEE encontrados son descritas por momentos de primer y segundo orden; consecuentemente, es más fácil desarrollar una simple pero útil teoría que describa otros PE.

I.3.3. Media y funciones de Correlación y Covarianza

En la práctica la mayoría de veces no es posible determinar la función densidad conjunta del PE, más una descripción parcial de éste lo dan los parámetros como la media y funciones de autocorrelación y autocovarianza. Consideremos $X(t)$ un PEE. Sea $X(t_k)$ una v.a. observada en el PE $X(t)$ al instante $t=t_k$.

La media del PE es definida como:

$$m_x(t_k) = E[X(t_k)] \quad \text{para cualquier } t_k \quad (1.47)$$

En general, la media del PE no es una constante, mas si el PE es Estacionario entonces:

$$m_x(t_k) = m_x \quad \text{para cualquier } t_k$$

Igualmente, la función de autocorrelación del PE. $X(t)$ es definida como:

$$R_x(t_j, t_k) = E[X(t_j) \cdot X(t_k)] \quad \text{Para todo } t_j, t_k \quad (1.48)$$

Si el PE, $X(t)$, es Estacionario, es posible mostrar que $R_x(t_j, t_k)$ es apenas función de la diferencia de los tiempos $(t_k - t_j)$.

$$R_x(t_j, t_k) = R_x(t_k - t_j) = R_x(\tau) = E[X(t) \cdot X(t - \tau)] \quad (1.49)$$

La función autocovarianza del PE $X(t)$ es definida como:

$$K_x(t_j, t_k) = E[\{X(t_j) - m_x(t_j)\} \cdot \{X(t_k) - m_x(t_k)\}] \quad (1.50)$$

Si el PE, $X(t)$, es estacionario, luego $m_x(t_j) = m_x(t_k) = m_x$, y su función de autocorrelación es función apenas de la diferencia de los tiempos $(t_k - t_j)$. Así:

$$K_x(t_j, t_k) = E[X(t_k)X(t_j) - m_x \cdot X(t_j) - m_x \cdot X(t_k) + m_x^2] = E[X(t_j) \cdot X(t_k)] - m_x^2$$

$$K_x(t_j, t_k) = K_x(t_j - t_k) = R_x(t_k - t_j) - m_x^2$$

$$\mathbf{Kx}(\tau) = \mathbf{Rx}(\tau) - \mathbf{mx}^2 \quad (1.51)$$

De lo anterior, se observa que **media y función de autocorrelación** describen parcialmente un PEE , esto es:

- la media es constante.
- $R_x(\tau)$ es independiente de un desplazamiento en el tiempo.
- $R_x(0)$ es finito.

Estas 3 condiciones no son suficientes para garantizar que el PE sea estrictamente estacionario. Un PE que no es estrictamente estacionario, pero que mantiene estas 3 condiciones es dicho **Proceso Estocástico Estacionario en el Sentido Amplio (PEESA)**.

Propiedades de $R_x(\tau)$ de un PEESA

- 1.) La función de autocorrelación (fac) de un PEESA es una función par:

$$R_x(\tau) = R_x(-\tau)$$

- 2.) El valor cuadrático medio (potencia total del proceso) de un PEESA es igual a:

$$E[X^2(t)] = E[X(t).X(t)] = R_x(0)$$

- 3.) Si $Z(t) = X(t) + Y(t)$ luego su fac es igual a:

$$R_z(\tau) = E[Z(t).Z(t-\tau)] = E[\{X(t)+Y(t)\} . \{X(t-\tau)+Y(t-\tau)\}]$$

$$R_z(\tau) = E[X(t).X(t-\tau)] + E[Y(t).Y(t-\tau)] + E[X(t).Y(t-\tau)] + E[Y(t).X(t-\tau)]$$

$$R_z(\tau) = R_x(\tau) + R_y(\tau) + R_{xy}(\tau) + R_{yx}(\tau)$$

- 4.) Si $X(t)$ tiene una componente periódica de período T , luego la fac del proceso también tiene una componente periódica de período T .
- 5.) $R_x(\tau)$ de un PEESA tiene su magnitud máxima en $\tau = 0$.
- 6.) $\lim_{\tau \rightarrow \infty} R_x(\tau) = [E(x)]^2 = m_x^2$
- 7.) Si el valor medio del proceso $X(t)$ es diferente de 0, luego $R_x(\tau)$ tendrá una componente DC.

La fac, $R_x(\tau)$, nos permite un medio de describir la interdependencia de dos v.a. obtenidas observando el PE, $X(t)$, a instantes separados τ segundos, como se ilustra en la figura:

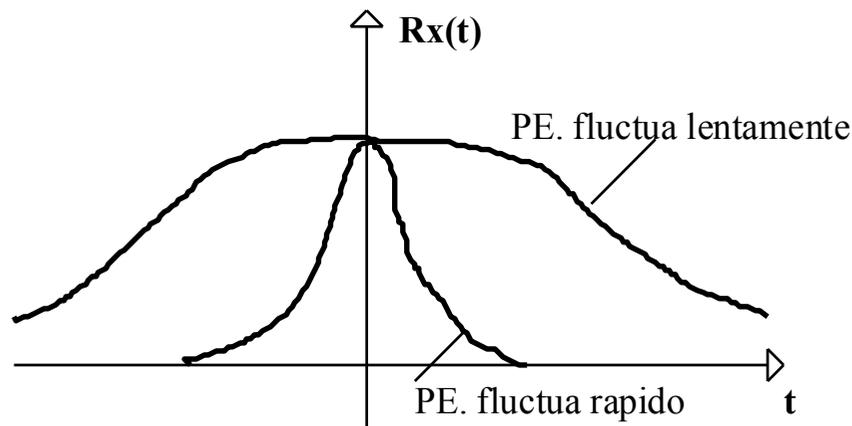


Figura 1.9 Ilustración de la función de auto correlación con fluctuación lenta y rápida

I.3.4. Ergodicidad y Medias Temporales

Si la teoría de PE es un método para describir señales de interés, debemos ser capaces de estimar a partir de observaciones de $X(t)$ tales cantidades probabilísticas como la media y la función de autocorrelación.

Para determinar $m_x(t_k)$ y $R_x(t_j, t_k)$ de un PE, es preciso disponer de todas las realizaciones. En particular, se requiere conocimiento completo de las estadísticas de 1ero. y 2do. orden del PE $X(t)$.

En la práctica, lo único disponible es, apenas, una simple realización del PE $X(t)$. Es natural entonces considerar medias temporales de simples realizaciones

Media Temporal

La media temporal de una realización, $x(t)$, de un PE. $X(t)$ es definida como:

$$\langle X(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} x(t) \cdot dt \quad (1.52)$$

Función Autocorrelación Temporal

La fac temporal de una realización, $x(t)$, de un PE, $X(t)$, es definida como:

$$\langle X(t) \cdot X(t - \tau) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} x(t) \cdot x(t - \tau) \cdot dt \quad (1.53)$$

Estas dos medias temporales son v.a. ya que sus valores dependen de la realización $x(t)$.

En general, MEDIAS ESTADÍSTICAS son iguales a las MEDIAS TEMPORALES para el caso en que el PEE sea ERGÓDICO (Boltzman 1887) [14]. Un proceso $X(t)$ es dicho ser ergódico si todas sus propiedades estadísticas pueden ser determinadas a partir de una simple realización del proceso $x(t)$. Esto es:

$$E[X(t)] = \langle X(t) \rangle \quad E[X(t) \cdot X(t - \tau)] = \langle X(t) \cdot X(t - \tau) \rangle \quad (1.54)$$

I.3.5. Densidad Espectral de Potencia de un PEESA

La Densidad Espectral de Potencia (DSP), $S_x(f)$, y función de autocorrelación, $R_x(\tau)$, de un PEESA, $X(t)$, forman un par de transformadas de Fourier dadas por:

$$S_x(f) = \int_{-\infty}^{+\infty} R_x(\tau) \cdot \exp(-j2\pi f\tau) \cdot d\tau \quad (1.55)$$

$$R_x(\tau) = \int_{-\infty}^{+\infty} S_x(f) \cdot \exp(+j2\pi f\tau) \cdot df$$

Este par de ecuaciones constituyen las relaciones de **Einstein-Wiener-Khintchine** para un PEESA.

Propiedades de la DSP

- 1.) Si el PEESA es de valor real la DSP es una función par, esto es:

$$S_x(f) = S_x(-f)$$

- 2.) La potencia total del proceso $X(t)$ es el área bajo la curva de la DSP:

$$E[X^2(t)] = R_x(0) = \int_{-\infty}^{+\infty} S_x(f) \cdot df$$

- 3.) $S_x(f) > 0$ para todo f

- 4.) $S_x(0) = \int_{-\infty}^{+\infty} R_x(\tau) \cdot d\tau$

I.3.6. Procesos Aleatorios a través de Filtros

Suponga que un PE $X(t)$ es aplicado a la entrada de un filtro lineal invariante con el tiempo de respuesta impulsiva $h(t)$, produciendo un PE $Y(t)$ de salida, como se ilustra en la figura.

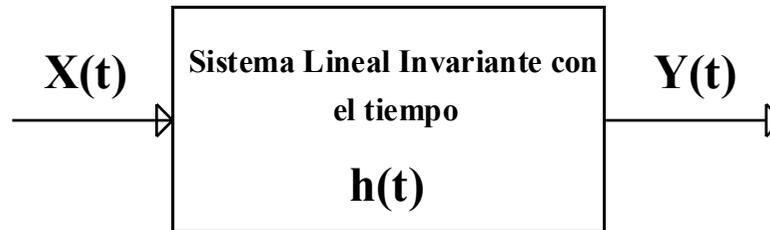


Figura 1.10 Sistema LTI alimentado con un Proceso Estocástico $X(t)$

Es difícil determinar la distribución de probabilidad de $Y(t)$ aún cuando la de $X(t)$ sea perfectamente conocida en el intervalo $-\infty < t < \infty$. En este punto, el objetivo es determinar la media y la función de autocorrelación del proceso de salida $Y(t)$ en función de la media y función autocorrelación del proceso de entrada, $X(t)$ considerado PEESA [10].

Media del Proceso de Salida

Por definición:

$$m_Y(t) = E[Y(t)]$$

ya que el sistema es lineal, la integral de convolución es posible mostrar que:

$$m_Y(t) = \int_{-\infty}^{+\infty} h(\tau) \cdot m_X(t - \tau) \cdot d\tau$$

Si $X(t)$ es Estacionario en el sentido amplio, $m_X(t - \tau) = m_X = \text{constante}$, luego:

$$m_y(t) = m_y = m_x \cdot \int h(\tau) \cdot d\tau$$

$$m_y(t) = m_y = m_x \cdot H(0) \tag{1.56}$$

donde $\rightarrow H(f) = \mathfrak{F}[h(t)]$

Función de Autocorrelación del proceso de salida:

Aplicando la definición y la integral de convolución, obtenemos:

$$R_y(t, u) = E[Y(t) \cdot Y(u)]$$

$$R_y(t, u) = E \left[\int h(\tau_1) \cdot X(t - \tau_1) \cdot d\tau_1 \int h(\tau_2) \cdot X(t - \tau_2) \cdot d\tau_2 \right]$$

$$R_y(t, u) = \int h(\tau_1) \cdot d\tau_1 \int h(\tau_2) \cdot R_x(t - \tau_1, u - \tau_2) \cdot d\tau_2$$

Si X(t) es un PEESA, luego $R_x(t - \tau_1, u - \tau_2) = R_x(t - u - \tau_1 + \tau_2)$

$$R_y(t, u) = R_y(t - u) = R_y(\tau) = \iint h(\tau_1) \cdot h(\tau_2) \cdot R_x(\tau - \tau_1 + \tau_2) \cdot d\tau_1 d\tau_2$$

con $\rightarrow \tau = t - u$

$$R_y(\tau) = h(\tau) * h(-\tau) * R_x(\tau) \text{ donde: } * \text{ denota la operación convolución.} \tag{1.57}$$

Se observa entonces que si X(t) es estacionario en el sentido amplio, Y(t) también lo es.

Densidad Espectral de Potencia del Proceso de Salida [04, 11]:

Para determinar $S_y(f)$ encontramos la transformada de Fourier de $R_y(\tau)$:

$$S_y(f) = \int R_y(\tau) \cdot \exp(-j2\pi f\tau) \cdot d\tau = \iiint h(\tau_1) \cdot h(\tau_2) \cdot R_x(\tau - \tau_1 + \tau_2) \cdot \exp(-j2\pi f\tau) \cdot d\tau_1 d\tau_2 d\tau$$

$$\text{Con: } \tau_0 = \tau - \tau_1 + \tau_2 \quad \Rightarrow \quad \tau = \tau_0 + \tau_1 - \tau_2$$

$$S_y(f) = \int h(\tau_1) \cdot \exp(-j2\pi f \cdot \tau_1) \cdot d\tau_1 \int h(\tau_2) \cdot \exp(j2\pi f \cdot \tau_2) \cdot d\tau_2 \int R_x(\tau_0) \cdot \exp(-j2\pi f \cdot \tau_0) \cdot d\tau_0$$

$$S_y(f) = H(f) \cdot H^*(f) \cdot S_x(f) = |H(f)|^2 \cdot S_x(f) \quad (1.58)$$

I.3.7. Procesos Estocásticos Discretos en el Tiempo

Un PEESA Discreto en el Tiempo (PEESADT) puede ser obtenido a partir del muestreo de un PEESA continuo en el tiempo, $X(t)$. Si la DSP de $X(t)$ es limitada en frecuencia en el intervalo ($|f| \leq B$), luego la mínima frecuencia de muestreo utilizada, según el teorema de Nyquist, es de $2B$ con el objetivo de recuperar $X(t)$ a partir de sus muestras $X(nT)$, donde T es el período de muestreo [05, 09, 18].

Para efectos de simplificación en el análisis matemático, vamos a normalizar el período de muestreo a $T=1$, y una simple realización muestreada del proceso $X(t)$ la simbolizaremos de la siguiente forma: $\{x(n); n = \dots -1, 0, 1, \dots\}$.

Es posible mostrar que la fac del PEESADT corresponde a muestras de la fac del PEESA continuo en el tiempo, $X(t)$. Esto es:

$$R_x(k) = E[x(n) \cdot x(n-k)] = R_x(\tau) \Big|_{\tau=kT} \quad (1.59)$$

Y la DSP del PEESADT es una repetición periódica de la DSP del PEESA continuo en el tiempo $X(t)$, centradas en torno de múltiplos enteros de la frecuencia de muestreo. En nuestro caso como $T=1$ es suficiente analizar el intervalo $-\frac{1}{2} < f < \frac{1}{2}$.

Matriz de Autocorrelación del PEESADT [08]:

Supongamos un intervalo de observación donde el proceso estocástico discreto en el tiempo pueda ser considerado estacionario y representado por el conjunto de muestras $\{x(n); n = 0, \dots, N-1\}$. Se define la matriz de autocorrelación (NxN) del PEESADT como:

$$\underline{R_x} = E\{\underline{X} \cdot \underline{X}^T\} \tag{1.60}$$

donde:

$$\underline{X} = [x(0) \ x(1) \ x(2) \ \dots \ x(N-1)]^T$$

Con esto:

$$\underline{R_x} = \begin{bmatrix} E[x(0)x(0)] & E[x(0)x(1)] & E[x(0)x(2)] & \dots & E[x(0)x(N-1)] \\ E[x(1)x(0)] & E[x(1)x(1)] & E[x(1)x(2)] & \dots & E[x(1)x(N-1)] \\ E[x(2)x(0)] & E[x(2)x(1)] & E[x(2)x(2)] & \dots & E[x(2)x(N-1)] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ E[x(N-1)x(0)] & E[x(N-1)x(1)] & E[x(N-1)x(2)] & \dots & E[x(N-1)x(N-1)] \end{bmatrix}$$

Substituyendo $E[x(i) \cdot x(j)] = R_x(i - j)$

$$\underline{R_x} = \begin{bmatrix} R_x(0) & R_x(-1) & R_x(-2) & \dots & R_x(-N+1) \\ R_x(1) & R_x(0) & R_x(-1) & \dots & R_x(-N+2) \\ R_x(2) & R_x(1) & R_x(0) & \dots & R_x(-N+3) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ R_x(N-1) & R_x(N-2) & R_x(N-3) & \dots & R_x(0) \end{bmatrix}$$

Si el PEESADT es de valor real, entonces: $R_x(k) = R_x(-k)$

Si el PEESADT es de valor complejo entonces $R_x(k) = R_x^*(-k)$

I.3.8. Estimación Cuadrática Media [03, 10]

Vamos a suponer que deseamos hacer una estimativa de una variable de interés, $d(n)$, a partir de observaciones ruidosas de la misma $x(n)$. En el siguiente gráfico se ilustra lo dicho:

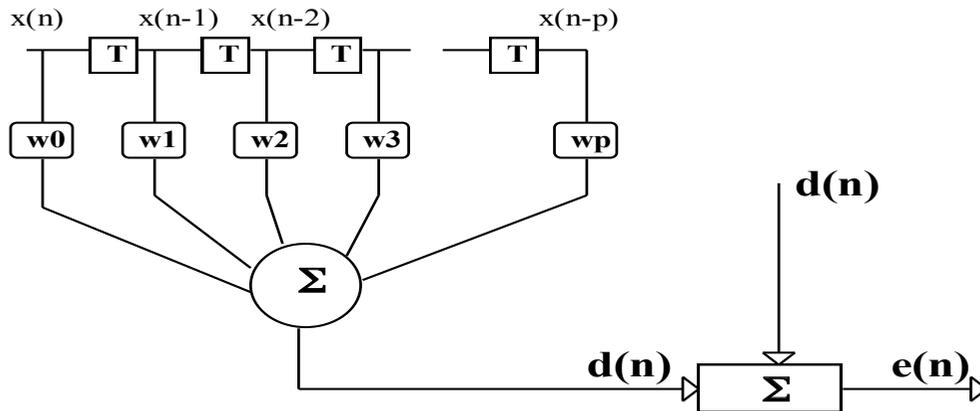


Figura 1.12 Filtro de Wiener

El objetivo es realizar una *estimativa lineal* de $d(n)$ a partir de $(p+1)$ muestras del proceso $\{x(n)\}$, esto es:

$$\hat{d}(n) = w_0 \cdot x(n) + w_1 \cdot x(n-1) + \dots + w_p \cdot x(n-p)$$

$$\hat{d}(n) = \sum_{i=0}^p w_i \cdot x(n-i)$$

Las constantes, w_i , son obtenidas a partir de *algún criterio de optimización*. En Procesamiento de Señales Estocásticas la función costo que normalmente es utilizada es el *valor cuadrático medio* de algún parámetro.

En este caso de estimación es evidente que existirá un *error en la estimativa* dado por:

$$e(n) = d(n) - \hat{d}(n)$$

$$e(n) = d(n) - \sum_{i=0}^p w_i \cdot x(n-i)$$

Y la función costo entonces será el *valor cuadrático medio* (vcm) *del error de estimación*

$$J(n) = E[e(n)^2]$$

Por lo tanto, los coeficientes w_0, w_1, \dots, w_n son obtenidos para minimizar la función costo y se dicen que son óptimos en el sentido que minimizan el vcm del error de estimación. Para determinar w_i , derivamos $J(n)$ en relación a cada w_i e igualamos a cero. Después de las respectivas manipulaciones matemáticas se llega al siguiente conjunto de ecuaciones, llamadas ECUACIONES DE WIENER-HOPF [10]:

$$E[d(n).x(n-j)] = \sum_{i=0}^p w_i E[x(n-i).x(n-j)] \quad \text{con } j=0,..,p \quad (1.61)$$

Tenemos $(p+1)$ ecuaciones, con $(p+1)$ incógnitas. En estas ecuaciones se verifican que:

$$E[x(n-i).x(n-j)] = R_x(j-i) \quad (j-i)\text{-ésimo atraso de la fac del proceso } X(nT)$$

$$E[d(n).x(n-j)] = R_{xd}(j) \quad (j)\text{-ésimo atraso de la fac cruzada de } X(n) \text{ y } d(n)$$

Las mismas pueden expresarse en forma matricial como sigue:

$$\begin{bmatrix} R_x(0) & R_x(-1) & R_x(-2) & \dots & R_x(-p) \\ R_x(1) & R_x(0) & R_x(-1) & \dots & R_x(-p+1) \\ R_x(2) & R_x(1) & R_x(0) & \dots & R_x(-p+2) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ R_x(p) & R_x(p-1) & R_x(p-2) & \dots & R_x(0) \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_p \end{bmatrix} = \begin{bmatrix} R_{dx}(0) \\ R_{dx}(1) \\ R_{dx}(2) \\ \vdots \\ R_{dx}(p) \end{bmatrix}$$

$$\underline{R_x} \cdot \underline{w} = \underline{R_{dx}} \quad (1.62)$$

donde:

\underline{R}_X = matriz de autocorrelación $(p + 1, p + 1)$ del proceso $X(nT)$

\underline{w} = vector de los coeficientes óptimos del estimador lineal

\underline{R}_{dx} = vector de correlación cruzada entre los procesos $X(nT)$ y $d(nT)$

I.3.9. Predicción Lineal [02, 12]

En muchos casos prácticos, la respuesta deseada, $d(n)$, no está disponible, por tanto, aplicación de la ecuaciones de Wiener-Hopf "no es posible". En el problema de predicción lineal se desea estimar la muestra actual, $x(n)$, de un PEESADT, $X(nT)$, en función de un conjunto de p muestras anteriores $\{x(n-1), x(n-2), \dots, x(n-p)\}$, de tal forma que minimice el vcm del denominado error de predicción, $e(n)$, como se ilustra en la figura

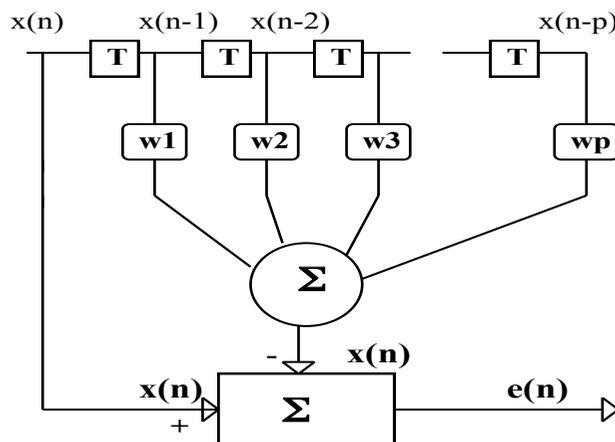


Figura 1.13 Filtro Predictor

La señal $e(n)$ es el denominado error de predicción lineal y dado por:

$$e(n) = x(n) - \hat{x}(n)$$

$$e(n) = x(n) - \sum_{i=1}^p w_i \cdot x(n-i)$$

La función costo a minimizar es el vcm del error de predicción:

$$J(n) = E[e^2(n)]$$

Luego de derivar la función costo en relación a cada w_i se llega a las denominadas **Ecuaciones Normales o de Yule-Walker**:

BIAS: El bias de un estimador es definido como el valor verdadero del parámetro, menos el valor esperado de la estimativa, es decir:

$$bias = B = \alpha - E[\hat{\alpha}]$$

donde: α es el parámetro a estimar y $\hat{\alpha}$ es la estimativa de α

Un estimador es dicho no bisiado si su bias es igual a cero, esto es, si la función densidad de probabilidad del estimador es simétrica y que su centro de simetria es el valor verdadero del parámetro a estimar.

Varianza: La varianza del estimador mide el ancho de la densidad de probabilidad del estimador y es definida como:

$$var[\hat{\alpha}] = E[\{\hat{\alpha} - E(\hat{\alpha})\}^2]$$

Una varianza pequeña del estimador sugiere que la densidad de probabilidad del estimador está concentrada en torno de su valor medio, lo que si el estimador es no bisiado, será el valor verdadero del parámetro a estimar.

Un estimador es dicho ser **consistente** si cuando el número de muestras utilizadas en la estimativa tiende al infinito, el bias y la varianza tienden a cero.

Dos estimadores de la fac de un PEESADT serán consideradas:

$$\tilde{R}_x(m) = \frac{1}{N-m} \sum_{n=0}^{N-m+1} x(n) \cdot x(n+m) = \frac{1}{N-m} \sum_{n=m}^{N-1} x(n) \cdot x(n-m) \quad (1.65)$$

Es posible mostrar que este estimador es un estimador no bisiado y que su varianza es proporcional a $1/N$. Dado que el bias y la varianza de este estimador tienden a cero cuando $N \rightarrow \infty$, luego es un estimador consistente.

$$\hat{R}_x(k) = \frac{1}{N} \sum_{n=0}^{N-k-1} x(n) \cdot x(n+k) = \frac{1}{N} \sum_{n=k}^{N-1} x(n) \cdot x(n-k) \quad (1.66)$$

Es posible mostrar que este estimador es un estimador bisiado más asintóticamente no bisiado y que su varianza es proporcional a $1/N$. Luego, el estimador es un estimador consistente.

BIAS: El bias de un estimador es definido como el valor verdadero del parámetro, menos el valor esperado de la estimativa, es decir:

$$bias = B = \alpha - E[\hat{\alpha}]$$

donde: α es el parámetro a estimar y $\hat{\alpha}$ es la estimativa de α

Un estimador es dicho no bisiado si su bias es igual a cero, esto es, si la función densidad de probabilidad del estimador es simétrica y que su centro de simetria es el valor verdadero del parámetro a estimar.

Varianza: La varianza del estimador mide el ancho de la densidad de probabilidad del estimador y es definida como:

$$var[\hat{\alpha}] = E[\{\hat{\alpha} - E(\hat{\alpha})\}^2]$$

Una varianza pequeña del estimador sugiere que la densidad de probabilidad del estimador está concentrada en torno de su valor medio, lo que si el estimador es no bisiado, será el valor verdadero del parámetro a estimar.

Un estimador es dicho ser **consistente** si cuando el número de muestras utilizadas en la estimativa tiende al infinito, el bias y la varianza tienden a cero.

Dos estimadores de la fac de un PEESADT serán consideradas:

$$\tilde{R}x(m) = \frac{1}{N-m} \sum_{n=0}^{N-m+1} x(n) \cdot x(n+m) = \frac{1}{N-m} \sum_{n=m}^{N-1} x(n) \cdot x(n-m) \quad (1.65)$$

Es posible mostrar que este estimador es un estimador no bisiado y que su varianza es proporcional a $1/N$. Dado que el bias y la varianza de este estimador tienden a cero cuando $N \rightarrow \infty$, luego es un estimador consistente.

$$\hat{R}x(k) = \frac{1}{N} \sum_{n=0}^{N-k-1} x(n) \cdot x(n+k) = \frac{1}{N} \sum_{n=k}^{N-1} x(n) \cdot x(n-k) \quad (1.66)$$

Es posible mostrar que este estimador es un estimador bisiado más asintóticamente no bisiado y que su varianza es proporcional a $1/N$. Luego, el estimador es un estimador consistente.

Puede observarse que, en los dos estimadores, cuando m o k tiende a N , el número de muestras disponibles para hacer las estimativas disminuyen dando una estimativa no confiable. Es posible utilizar el siguiente criterio insertado por YULE-WALKER: El número de atrasos que pueden estimarse con buen grado de confiabilidad es del orden del 10% de los datos disponibles para la estimativa, N [.

Varios investigadores utilizan el valor cuadrático medio de la estimativa para cuantificar un estimador:

$$E[\hat{\alpha}^2] = \text{var}(\hat{\alpha}) + (\text{bias})^2$$

Basados en este criterio, es posible mostrar que para muchas aplicaciones el estimador bisiado $\hat{R}_x(k)$ posee un error cuadrático medio menor que el del estimador no bisiado $\tilde{R}_x(m)$.

Parte II

Procesamiento Digital de Señales

Introducción

Señales son representaciones de experimentos que, en muchos casos, no son fácilmente observables. El procesamiento de las mismas tiene a ver con la *representación, transformación y manipulación* de señales y de la información contenida en ellas. Las técnicas y aplicaciones del procesamiento de señales han jugado un papel importante en diferentes campos, siendo éstas tan antiguas como Newton y Gauss (XVII,XVIII) y tan modernas como los circuitos integrados de muy alta escala de integración (VLSI).

El procesamiento de señales en general posee una rica historia y su importancia es evidente por el amplio espectro de su aplicación en áreas como: radar, biomédica, sonar, sísmica, voz, video, datos, comunicación de datos-voz-video, robótica, instrumentación y muchas otras.

Antes de 1960, el procesamiento de señales era exclusivo de la tecnología analógica continua en el tiempo. La rápida evolución de los computadores digitales, así como de la tecnología de alta integración, hizo que exista un desplazamiento a la tecnología digital.

De esta forma, el término Procesamiento Digital de Señales debe ser entendido como la manipulación de secuencias discretas en el tiempo (si la señal es analógica debe ser discretizada) representadas por números de precisión finita donde el procesamiento es implementado utilizando un computador digital o hardware dedicado.

El procesamiento de señales no está restringido a sistemas o señales unidimensionales. La mayoría de las técnicas de procesamiento unidimensional tienen su contraparte en sistemas multidimensionales tal como el procesamiento de imágenes.

En este caso, áreas como codificación de video, imágenes médicas, análisis y mejoramiento de fotografías aéreas, imágenes de Radares de Apertura Sintética (SAR) y ótras, requieren técnicas de procesamiento de señales bidimensionales y hasta tridimensionales.

Los fundamentos del procesamiento de señales pueden ser aplicados a muchos sectores avanzados y especializados. En **análisis espectral digital** por ejemplo, el uso de la Transformada Discreta de Fourier (DFT) es de suma importancia para determinar el denominado periodograma clásico y otros métodos como los de Daniel, Bartlett y Welch.

En modelamiento, el procesamiento de señales juega un rol importante en la codificación y compresión de datos a partir de la técnica de **codificación por predicción lineal (LPC)**, conjuntamente con la teoría de **procesos estocásticos** [02, 04, 05, 09].

Un tópico avanzado de mucha importancia es el procesamiento adaptativo de señales estocásticas el cual encuentra sus fundamentos en: *procesamiento de señales y procesos estocásticos*. Señales como voz, video, datos, radar, sísmica y otras son aleatorias por naturaleza, esto es, que pueden ser modeladas como siendo una realización de un Proceso Estocástico No Estacionario. Es evidente entonces que las técnicas de procesamiento de señales sean adaptativas, las que serán objeto de un futuro texto.

Finalmente, el procesamiento de señales se ha visto beneficiado por el acople existente entre la teoría y aplicaciones y principalmente por las tecnologías para su implementación a partir de los denominados Procesadores Digitales de Señales (DSP).

Estos DSP's tienen la característica de ser proyectados para realizar un gran volumen de cálculo en el menor tiempo posible, dando la posibilidad de implementar algoritmos complejos (radar por ejemplo) de procesamiento de señales en tiempo real, área en la cual hay mucho camino por recorrer, lo que será posible con la continua auto-actualización y especialización.

II.1. Señales y Sistemas Discretos en el tiempo

II.1.1. Relación entre Formas de Representación de Sistemas o Señales

Así como los sistemas y señales continuos en el tiempo disponen de herramientas matemáticas como la Transformada de Fourier (FT) y Transformada de Laplace para su representación, su contraparte para los sistemas y señales discretas en el tiempo lo son la Transformada Discreta de Fourier (DTFT) y la Transformada Discreta de Fourier (DFT) [13, 15, 16].

Las interrelaciones existentes entre estas diversas formas de representación son ilustradas en el siguiente gráfico:

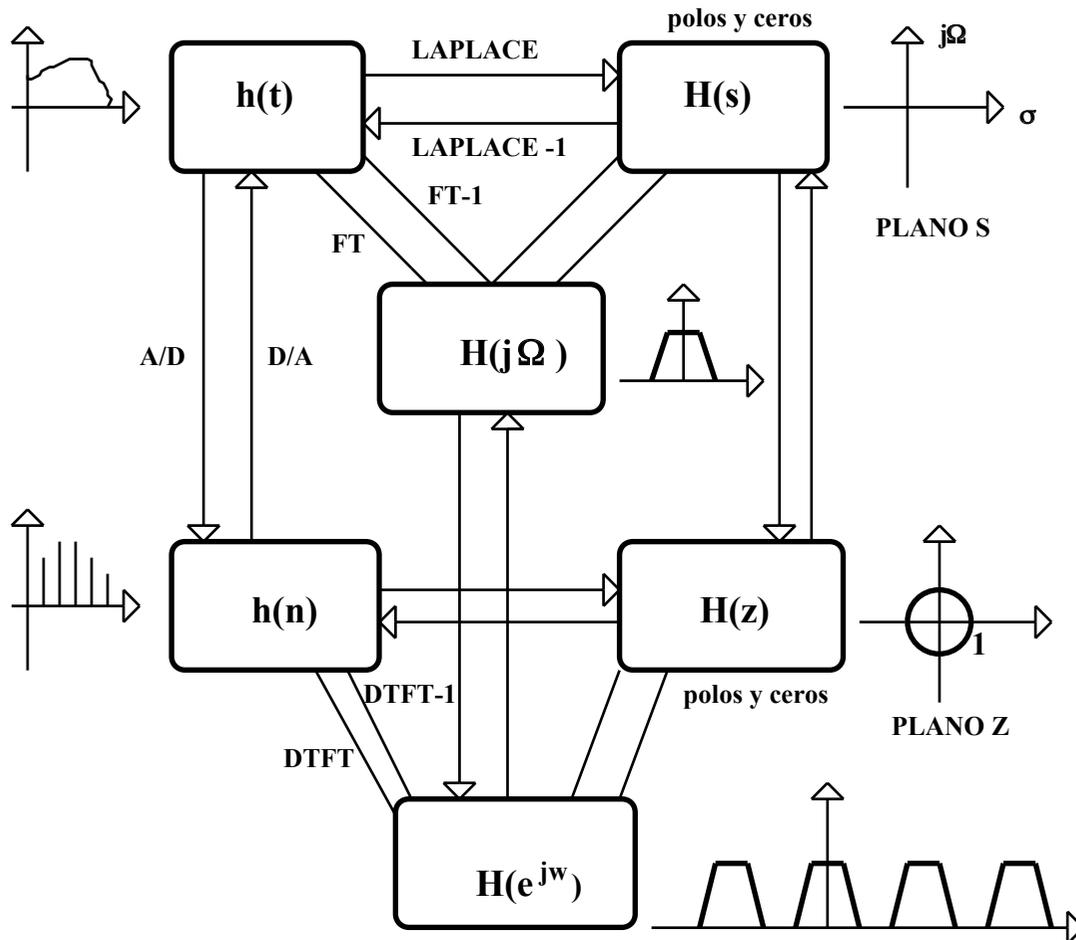


Figura 2.1 Relaciones entre diversas transformadas

II.1.2. Esquema Básico de Procesamiento Digital de Señales

El esquema básico de un sistema de Procesamiento Digital de Señales originariamente analógicas es ilustrado a continuación:

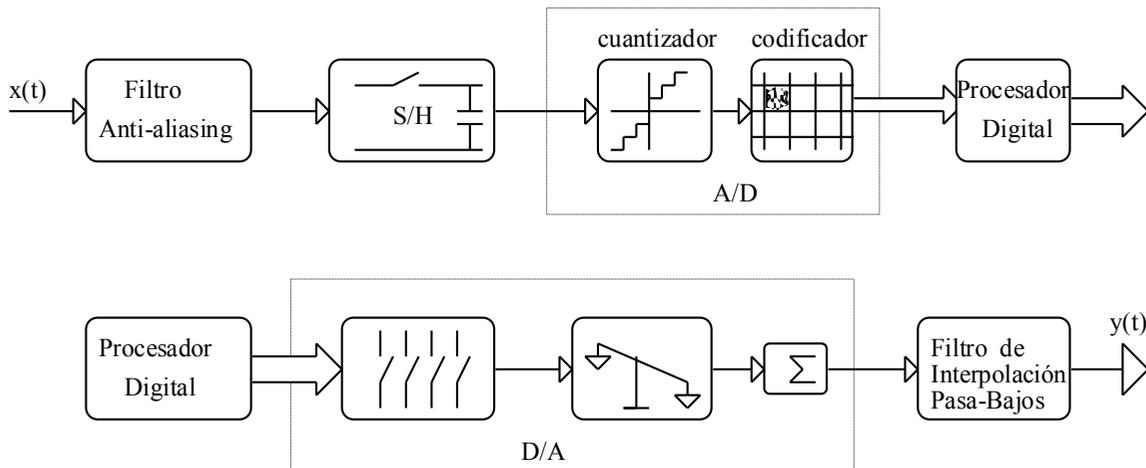


Figura 2.2 Sistema de Procesamiento Digital

El filtro pasa-bajos analógico tiene la función de limitar el espectro de la señal de entrada para mitigar el efecto de "aliasing". Otra función que desempeña este filtro es la de eliminar el posible ruido existente fuera de la banda de interés, tópico que será analizado en detalle.

El Sample/Hold (S/H) tiene la función de realizar el proceso de discretización en el tiempo de la señal analógica de entrada y de mantener estable la muestra durante el proceso de conversión análogo-digital.

El Conversor Análogo-Digital (A/D) está constituido por: 1) Cuantizador (Q) el que realiza el proceso de discretización en amplitudes de la señal y 2) Codificador, el que realiza el proceso de asociación de números (binarios u otros).

El Procesador Digital que puede ser un computador, microcomputador o hardware dedicado, en el que se implementa el procesamiento deseado.

El Conversor Digital-Analógico (D/A) que convierte la señal digital entregada por el procesador, en una señal continua en el tiempo más discreta en amplitudes.

Por último, el Filtro de Interpolación que, en muchos casos, es un filtro pasa-bajos que tiende a suavizar la señal entregada por el D/A.

II.1.3. Ventajas y puntos que se deben observar sobre el Procesamiento Digital en relación al Analógico

- **Robustez de los Sistemas Digitales:**

Reproductividad.- La reproducción de un Circuito Integrado (IC) es mucho más confiable que la reproducción de elementos como Resistores (R), Condensadores (C) y otros, los que, en muchos casos, tienen tolerancias grandes.

Almacenamiento y Transmisión.- La degradación de la señal al ser almacenada o transmitida en forma digital es mucho menor que en el caso analógico.

Regenerabilidad de señales.- La regeneración de una señal digital se limita al reconocimiento de dos niveles de tensión (en el caso binario).

- **Flexibilidad:**

Procesadores Programables y/o Adaptables.- Si se desea modificar alguna característica del sistema que va a proyectarse, será necesario apenas variar el contenido de la memoria.

Controlados por Software.- Las características de ciertos IC's (AD7111), tales como ganancia del amplificador, números de bits de cuantización y otros parámetros, pueden ser controlados por software.

Implementación de Procesamiento por Software.- Esto es que los algoritmos de procesamiento pueden ser optimizados numéricamente.

- **Posibilidad de Implementación en Tiempo no Real:**

Un intervalo de interés de la salida del A/D puede ser almacenado en memoria para su posterior análisis y procesamiento (Osciloscopio con Almacenamiento Digital).

- **Capacidad de Operación con Números:**

Algoritmos Complejos.- La capacidad de operación del procesador nos permite la implementación de algoritmos complejos.

Resultados Deseados.- La repetibilidad de los cálculos numéricos hace con el sistema sea exacto y su precisión dependerá del número de bits de cuantización y del tipo de aritmética que debe utilizarse.

- **Simulación/Emulación:**

CAD/CAE.- Se abre la posibilidad del Diseño Asistido por Computador.

Dentro de los puntos que deben ser observados de los sistemas de procesamiento digital de señales, podemos mencionar los siguientes:

Son exactos, más no precisos, más dependiendo de la aplicación se puede llegar a resultados deseados.

El mantenimiento y el diagnóstico son relativamente complicados.

La velocidad de procesamiento: el proceso de suma es mucho más lento que en una analógica.

II.1.4. Señales

1) Señales continuas en el tiempo/espacio y en amplitudes:

Una señal es llamada analógica, si alguno de sus parámetros (amplitud por ejemplo) varía **análogamente** a algún parámetro del fenómeno de interés. Por ejemplo: El voltaje de salida de un micrófono varía análogamente a la señal acústica producida por una persona. El brillo en una fotografía es un ejemplo de una señal analógica bidimensional. Evidentemente, estas señales son continuas en el tiempo/espacio y en amplitudes.

2) Señales discretas en el tiempo/espacio y continuas en amplitud:

La salida de un muestreador ideal es una señal de este tipo. Sólo dispone valores a ciertos instantes de tiempo. Varios **cuadros** de una señal de video son un ejemplo de una señal discreta en el tiempo y continua en amplitudes, bidimensional.

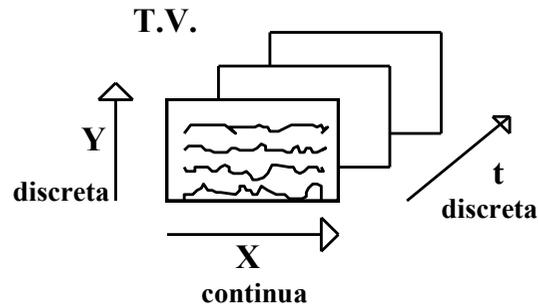
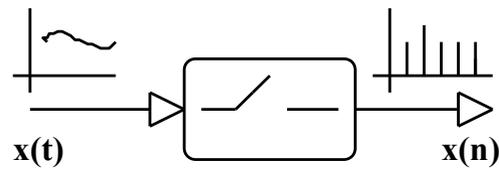


Figura 2.3 Muestreo en el tiempo de señales de voz y vídeo

3.) Señales continuas en el tiempo/espacio y discretas en amplitud:

Las señales que son procesadas digitalmente más sin el proceso de discretización en el tiempo. Un ejemplo es la salida del conversor D/A.

4.) Señal discreta en el tiempo/espacio y en amplitud, con estas codificadas numéricamente. Estas son las denominadas señales digitales.

II.1.5 Sistemas

Un sistema discreto es un **algoritmo, mapeamiento, operación o transformación** que actúa sobre una secuencia de muestras de entrada, $x(n)$, para generar una secuencia de muestras de salida, $y(n)$.

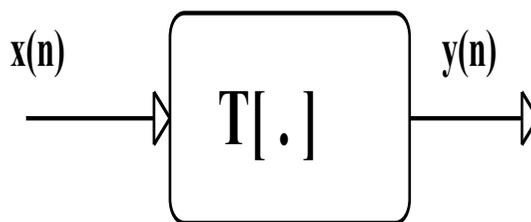


Figura 2.4 Sistema discreto en el tiempo

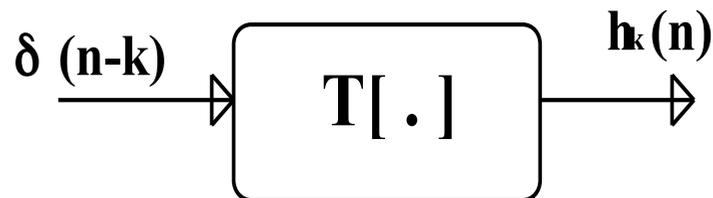
Propiedades:**Linealidad**

$$\text{Si: } y_1(n) = T[x_1(n)] \quad y_2(n) = T[x_2(n)]$$

El sistema es lineal si:

$$T[a_1 \cdot x_1(n) + a_2 \cdot x_2(n)] = a_1 \cdot T[x_1(n)] + a_2 \cdot T[x_2(n)]$$

Para caracterizar un sistema discreto, disponemos de la denominada respuesta a la muestra unitaria $\delta(n-k)$. Entonces, dado un sistema, $T[\cdot]$, notaremos como $h_k(n)$ a la respuesta del sistema a una entrada $\delta(n-k)$.



Esto es:

$$h_k(n) = T[\delta(n-k)] \quad (2.1)$$

Para una entrada $x(n)$ obtenemos una salida $y(n)$ y dado que $x(n)$ puede ser expresada como:

$$x(n) = \sum_{k=-\infty}^{+\infty} x(k) \cdot \delta(n-k)$$

entonces, $y(n)$ puede expresarse como:

$$y(n) = T \left[\sum_{k=-\infty}^{+\infty} x(k) \cdot \delta(n-k) \right]$$

$$y(n) = \sum_{k=-\infty}^{+\infty} x(k) \cdot T[\delta(n-k)]$$

Sustituyendo (2.1) en la última expresión, tenemos:

$$y(n) = \sum_{k=-\infty}^{+\infty} x(k) \cdot h_k(n) \quad (2.2)$$

Invariante con el Desplazamiento del Tiempo/Espacio

$$\text{Si } y(n) = T[x(n)]$$

Luego, el sistema es invariante con el desplazamiento si:

$$T[x(n-n_0)] == y(n-n_0)$$

Sumatoria de Convolución

Si el sistema es lineal e invariante con el desplazamiento del tiempo/espacio (LTI), luego:

$$h(n) = T[(n)] \qquad h(n-k) = T[(n-k)]$$

Sustituyendo lo anterior en la expresión (2.2), tenemos:

$$y(n) = \sum_{k=-\infty}^{+\infty} x(k) \cdot h(n-k) \quad (2.3)$$

La expresión anterior es denominada **SUMATORIA DE CONVOLUCIÓN**

Causalidad

Un sistema causal es aquél cuya respuesta a cada instante, n_0 , depende apenas de las entradas para $n \leq n_0$. En el caso particular de sistemas LTI, el sistema es causal si y solamente si:

$$h(n) = 0 \quad \text{para } n < 0$$

Estabilidad

Un sistema es estable, si para cualquier entrada limitada en amplitud, la salida también es limitada en amplitud. En el caso particular de sistemas LTI una condición necesaria y suficiente para que el sistema sea estable es que:

$$S = \sum_{k=-\infty}^{k=+\infty} |h(k)| < \infty$$

II.1.6. Respuesta en Frecuencia

Vamos a considerar un sistema LTI caracterizado por una respuesta a la muestra unitaria, $h(n)$, y una entrada, $x(n)$, exponencial de valor complejo dada por :

$$x(n) = k \cdot \exp(j \cdot \omega \cdot n)$$

La salida, $y(n)$, puede ser determinada a partir de (2.3) ya que el sistema es LTI:

$$y(n) = \sum_{k=-\infty}^{k=+\infty} k \cdot \exp[j\omega(n - k)] \cdot h(k)$$

Arreglados los términos de la expresión anterior, tenemos:

$$y(n) = \underbrace{k \cdot \exp(j\omega \cdot n)}_{\text{Señal de entrada}} \cdot \underbrace{\sum_{k=-\infty}^{k=+\infty} h(k) \cdot \exp(-j\omega k)}_{\text{Respuesta en frecuencia del sistema}}$$

La expresión:

$$H(e^{j\omega}) = \sum_{k=-\infty}^{k=+\infty} h(k) \cdot \exp(-j\omega k) \quad (2.4)$$

es un número complejo que puede ser expresado como:

$$H(e^{j\omega}) = |H(e^{j\omega})| \cdot \exp[\arg\{H(e^{j\omega})\}]$$

Es decir, que la magnitud y fase de la exponencial compleja de entrada son modificadas por la magnitud y fase, respectivamente, de $H(e^{j\omega})$. Por tanto, la expresión (2.4) es denominada Respuesta en Frecuencia de un sistema LTI.

Es posible mostrar que $H(e^{j\omega})$ es una función continua de ω y periódica en ω con período igual a 2π . Observando la expresión (2.4) vemos que si $H(e^{j\omega})$ es una función continua de ω y periódica, es posible representarla como una serie de Fourier, donde los $h(k)$ son los coeficientes de la serie y dados por:

$$h(k) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} H(e^{j\omega}) \cdot \exp(j\omega k) \cdot d\omega \quad (2.5)$$

El par de ecuaciones (2.4) y (2.5) es el par denominado: Discrete Time Fourier Transform (DTFT) y es válida para cualquier secuencia $h(n)$, siempre que:

$$\sum_{k=-\infty}^{k=+\infty} |h(k)| < \infty$$

Es fácil mostrar que si en el dominio del tiempo la salida de un sistema LTI es dado por:

$$y(n) = x(n) * h(n)$$

donde, el símbolo $*$ denota sumatoria de convolución, en el dominio de la frecuencia tenemos:

$$Y(e^{j\omega}) = H(e^{j\omega}) \cdot X(e^{j\omega})$$

II.1.7. Relación entre Señales Discretas y Continuas en el Tiempo y la Frecuencia

Consideremos una señal analógica $x_a(t)$ que está siendo muestreada con un período de muestreo igual a T . La señal resultante la notaremos $x_a(nT) = x(n)$. La representación de Fourier, $X_a(j\Omega)$, de $x_a(t)$ es dada por el par de transformadas:

$$x_a(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X_a(j\Omega) \cdot \exp(j\Omega t) \cdot d\Omega \quad (2.6.a)$$

$$X_a(j\Omega) = \int_{-\infty}^{+\infty} x_a(t) \cdot \exp(-j\Omega t) \cdot dt \quad (2.6.b)$$

donde: $\Omega = 2\pi f$ es la frecuencia angular analógica.

De (2.6.a), tenemos que:

$$x(n) = x_a(nT) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X_a(j\Omega) \cdot \exp(j\Omega nT) \cdot d\Omega \quad (2.7)$$

A partir de la DTFT de $x(n)$, expresión (2.5), tenemos:

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} X(e^{j\omega}) \cdot \exp(j\omega n) \cdot d\omega \quad (2.8)$$

El objetivo es relacionar las ecuaciones (2.7) y (2.8). Para esto, es conveniente expresar (2.7) de la siguiente manera:

$$x(n) = x_a(nT) = \frac{1}{2\pi} \sum_{r=-\infty}^{\infty} \left[\int_{(2r-1)\pi/T}^{(2r+1)\pi/T} X_a(j\Omega) \cdot \exp(j\Omega nT) \cdot d\Omega \right]$$

Cada término en la sumatoria puede reducirse a un integral sobre el intervalo $-\pi/T$ a $+\pi/T$, mediante un cambio de variables para obtener:

$$x(n) = xa(nT) = \frac{1}{2\pi} \sum_{r=-\infty}^{\infty} \int_{-\pi/T}^{+\pi/T} Xa\left(j\Omega + j\frac{2\pi r}{T}\right) \cdot \exp(j\Omega nT) \cdot \exp(j2\pi rn) \cdot d\Omega$$

Intercambiando el orden de la integral y la sumatoria y observando que $\exp(j2\pi rn)=1$ para cualquier valor entero de r y n , obtenemos:

$$x(n) = xa(nT) = \frac{1}{2\pi} \int_{-\pi/T}^{+\pi/T} \left[\sum_{r=-\infty}^{r=+\infty} Xa\left(j\Omega + j\frac{2\pi r}{T}\right) \right] \cdot \exp\{j(\Omega T)n\} \cdot d\Omega \quad (2.9)$$

Comparando las ecuaciones (2.8) y (2.9), se concluye que:

$$\Omega = \omega/T \quad (2.10)$$

$$X(e^{j\omega}) = X(e^{j\Omega T}) = \frac{1}{T} \sum_{r=-\infty}^{\infty} Xa\left(j\Omega + j\frac{2\pi r}{T}\right) \quad (2.11)$$

Se observa que ω es adimensional y denominada frecuencia angular digital. La expresión (2.11) nos dice que el espectro de la señal muestreada, $xa(nT)=x(n)$, es una versión repetida del espectro de la señal continua en torno de múltiplos de la frecuencia de muestreo. Este efecto es ilustrado en el siguiente diagrama:

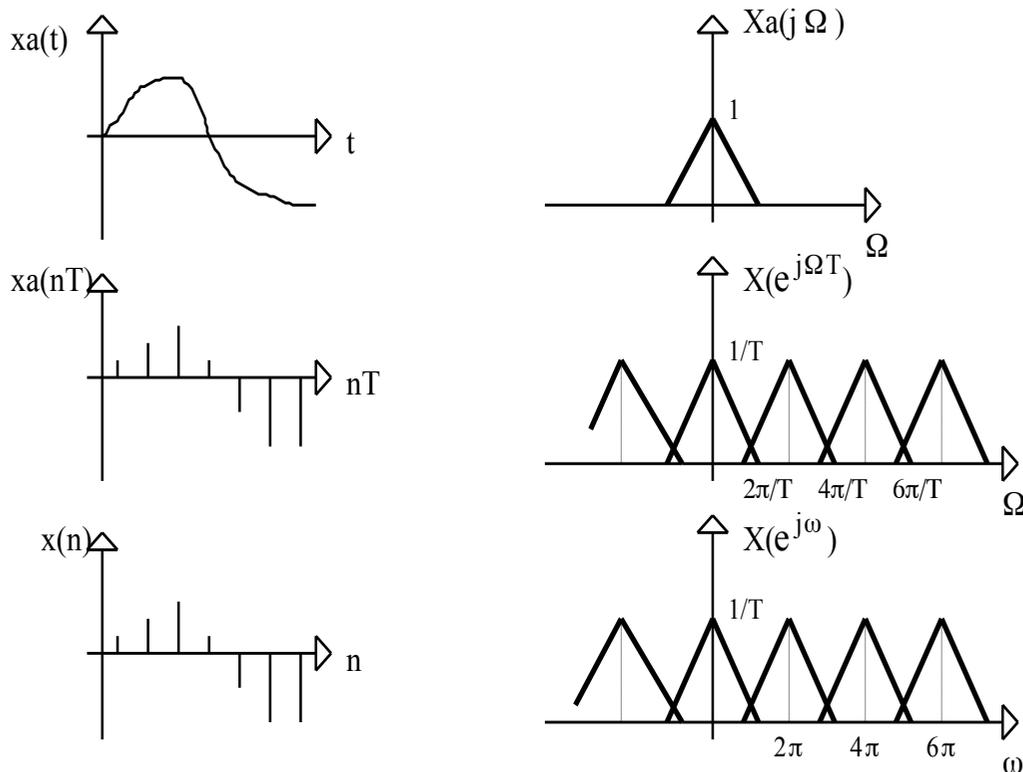


Figura 2.5 Relación entre el espectro de una señal continua y la discreta en el tiempo

Es posible mostrar a partir de las expresiones anteriores que la señal continua en el tiempo, $x_a(t)$, puede ser recuperada a partir de sus muestras (obtenidas con la mínima frecuencia de muestreo = Nyquist) mediante la denominada fórmula de Whittaker:

$$x_a(t) = \sum_{k=-\infty}^{\infty} x_a(kT) \cdot \frac{\text{Sen}\left[\frac{\pi}{T}(t-kT)\right]}{\frac{\pi}{T}(t-kT)} \quad (2.12)$$

La expresión (2.12) nos dice que para recuperar $x_a(t)$ sin distorsión alguna, es necesario disponer de un filtro cuya respuesta impulsiva sea una $\text{Sinc}(\cdot)$.

En la práctica, los conversores D/A utilizan un filtro con una respuesta impulsiva $u(t)-u(t-T)$, donde $u(\cdot)$ es la función paso unitaria.

II.2. Transformada Z

Dada una secuencia $x(n)$ cualquiera, su transformada Z es definida como:

$$X(z) = \sum_{n=-\infty}^{\infty} x(n) \cdot z^{-n} \quad (2.13)$$

donde: $z = r \cdot e^{j\omega}$

Con esto:

$$X(z) = X(r \cdot e^{j\omega}) = \sum_{n=-\infty}^{\infty} [x(n) \cdot r^{-n}] \cdot \exp(-j\omega n)$$

Esto es, la transformada Z, $\text{Tr}[Z]$, de $x(n)$ corresponde a la DTFT del producto de $x(n)$ por una secuencia exponencial de valor real. La existencia de la DTFT de una secuencia cualquiera estaba restringida a que la energía de la secuencia debía ser finita.

En el caso de la $\text{Tr}[Z]$, mismo que la secuencia no tenga DTFT, ésta puede existir dependiendo de la secuencia exponencial r^{-n} .

II.2.1. Secuencias y Regiones de Convergencia de Su Tr[Z]

Dada una secuencia $x(n)$ y su transformada Z , $X(z)$, con z de valor complejo. Su región de convergencia (RC) puede expresarse en forma general como:

$$0 \leq R_{x-} \leq |z| \leq R_{x+} \leq \infty$$

a.-) Secuencias de duración finita

Una secuencia de duración finita está definida dentro de un intervalo $[n_1, n_2]$.

Si $n_1 < 0 \Rightarrow |X(z)| \rightarrow \infty$ cuando $|z| \rightarrow \infty \Rightarrow$ RC no incluye $|z| = \infty$

Si $n_2 > 0 \Rightarrow |X(z)| \rightarrow \infty$ cuando $|z| \rightarrow 0 \Rightarrow$ RC no incluye $|z| = 0$

b.-) Secuencias a la Derecha

Una secuencia a derecha es tal que $x(n) = 0$ para $n < n_1$. En este caso:

$$X(z) = \sum_{n=n_1}^{\infty} x(n) \cdot z^{-n}$$

Si $n_1 \geq 0$, la RC no incluye el origen más si el infinito, esto es, la RC es el exterior de un círculo de radio R_{x-}

Si $n_1 < 0$, la RC no incluye el origen y el infinito, esto es, la RC es el exterior de un círculo de radio R_{x-} que no incluya al infinito.

c.-) Secuencias a la Izquierda

Una secuencia a la izquierda es tal que: $x(n) = 0$ para $n > n_2$. En este caso:

$$X(z) = \sum_{n=-\infty}^{n_2} x(n) \cdot z^{-n}$$

Si $n_2 \leq 0$, la RC no incluye al infinito, esto es, la RC es el interior de un círculo de radio R_{x+}

Si $n_2 > 0$, la RC no incluye el origen y al infinito, esto es, la RC es el interior de un círculo de radio R_{x+} que no incluya al origen.

d.-) Secuencia Bilateral

En este caso:

$$X(z) = \sum x(n).z^{-n}$$

$$X(z) = \underbrace{\sum_{n=-\infty}^{-1} x(n).z^{-n}}_{\text{Converge para } |z| < R_{x+}} + \underbrace{\sum_{n=0}^{\infty} x(n).z^{-n}}_{\text{Converge para } |z| > R_{x-}}$$

Por tanto, la RC es una region anular de radio externo R_{x+} y radio interno R_{x-} . Si $R_{x+} < R_{x-}$ luego la $Tr[Z]$ no existe.

PROPIEDADES:

Linealidad

Dadas 2 secuencias $x(n)$ $y(n)$ con $Tr[Z]$ $X(z)$ y $Y(z)$, respectivamente:

$$Z[a.x(n) + b.y(n)] = a.X[z] + b.Y[z]$$

Desplazamiento Temporal:

$$Z[x(n-n_0)] = z^{-n_0}.X(z)$$

II.2.2. Función del Sistema [06]

Ahora, veamos porque la $Tr[Z]$ es utilizada para el análisis de sistemas y señales discretas en el tiempo. Es posible representar las relaciones de entrada-salida de un sistema LTI a través de la denominada **Ecuación de Diferencias con Coeficientes Constantes** de la siguiente forma:

$$\sum_{k=0}^N a_k \cdot y(n-k) = \sum_{k=0}^M b_k \cdot x(n-k)$$

En el dominio de z , tenemos

$$\sum_{k=0}^N a_k \cdot z^{-k} \cdot Y(z) = \sum_{k=0}^M b_k \cdot z^{-k} \cdot X(z)$$

$$Y(z) = \frac{\sum_{k=0}^M b_k \cdot z^{-k}}{\sum_{k=0}^N a_k \cdot z^{-k}} \cdot X(z)$$

Observe que si $x(n) = \delta(n)$ $X(z) = 1$ y por tanto:

$$Y(z) \Big|_{X(z)=1} = H(z) = \frac{\sum_{k=0}^M b_k \cdot z^{-k}}{\sum_{k=0}^N a_k \cdot z^{-k}}$$

$$\text{Para Sistemas LTI: } Y(z) = H(z) \cdot X(z) \quad (2.14)$$

donde: $H(z)$ es la respuesta del sistema a la muestra unitaria

Se observa que, como $H(z)$ es una relación de polinomios en z^{-1} , podemos expresarla como:

$$H(z) = A \cdot \frac{\prod_{k=1}^M (1 - c_k \cdot z^{-1})}{\prod_{k=1}^N (1 - d_k \cdot z^{-1})}$$

El numerador contribuye con ceros en $z = c_k$ y polos en $z = 0$, en tanto, que el denominador contribuye con polos en $z = d_k$ y ceros en $z = 0$, por lo tanto, a menos de una constante A , $H(z)$ es caracterizado por un diagrama de polos y ceros. En el siguiente gráfico se ilustra la posición de polos para varios tipos de filtros.

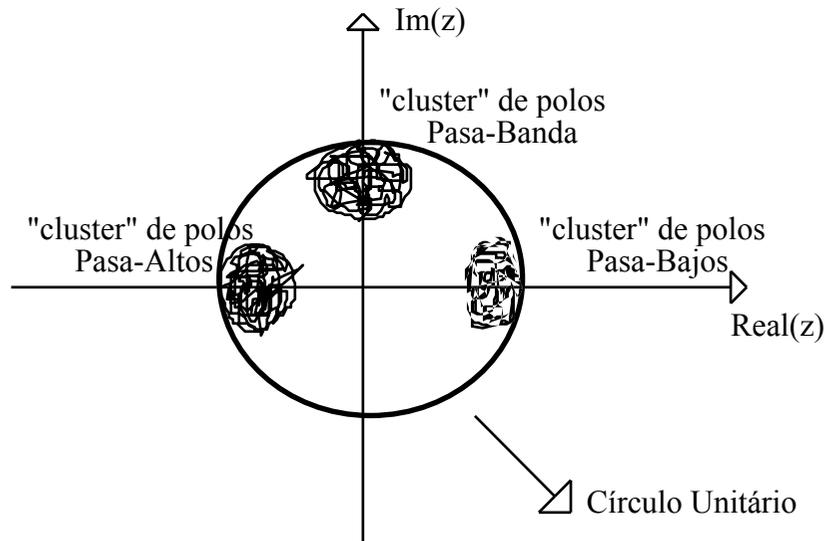


Figura 2.6 Ubicación de polos y su relación con el tipo del filtro

Por último, en la ecuación a diferencias de un sistemas LTI, podemos hacer $a_0 = 1$ sin pérdida de generalización, obteniendo:

$$y(n) = \sum_{k=0}^M b_k \cdot x(n-k) - \sum_{k=1}^N a_k \cdot y(n-k)$$

Con esto $H(z)$ puede expresarse como:

$$H(z) = \frac{\sum_{k=0}^M b_k \cdot z^{-k}}{1 + \sum_{k=1}^N a_k \cdot z^{-k}} \quad (2.15)$$

Si $\{a_k\} = 0$, la función del sistema tiene $(M+1)$ ceros y $(M+1)$ polos en el origen por lo que el sistema es estable ya que la RC es todo el plano z excepto el origen y por lo tanto el círculo unitario está en la RC. Estos sistemas o filtros son llamados de Respuesta Impulsiva Finita (FIR).

Si por lo menos uno de los $a_k \neq 0$, la estructura se torna recursiva y su estabilidad está determinada por la ubicación de los polos. Son llamados de Respuesta Impulsiva Infinita (IIR).

II.3. Transformada Discreta de Fourier

La Transformada Discreta de Fourier (DFT) es una herramienta para análisis y representación de sistemas y señales discretas en el tiempo de duración finita.

II.3.1. Serie Discreta de Fourier (DFS)

Como sabemos, una secuencia periódica posee energía infinita, esto es, su DTFT no existe. La Serie Discreta de Fourier (DFS) es una herramienta para representar secuencias discretas en el tiempo periódicas. Sea $\tilde{x}(n)$ una secuencia periódica con período N , esto es:

$$\tilde{x}(n) = \tilde{x}(n + kN) \quad \text{para todo } k \text{ entero}$$

Siendo $\tilde{x}(n)$ periódica, es posible representarla mediante una serie de Fourier, dada por:

$$\tilde{x}(n) = \sum_k \tilde{X}(k) \cdot \exp(j2\pi kn / N) \quad (2.16)$$

siendo una combinación lineal de exponenciales complejas armónicamente relacionadas. Comparando con el caso continuo, sea la señal continua $\tilde{x}(t)$, expresada como una serie de Fourier, esto es:

$$\tilde{x}(t) = \sum_{k=-\infty}^{\infty} \tilde{X}(k) \cdot \exp(jk \cdot \omega_0 \cdot t) \quad (2.17)$$

donde: $\omega_0 = 2\pi/T_0$ es la frecuencia fundamental de $\tilde{x}(t)$ y T_0 es la periodicidad de $\tilde{x}(t)$.

Haciendo $t = nT$, donde T es el período de muestreo tal que: $T_0 = NT$, la expresión (2.17) es la (2.16). Observe que en (2.16), los límites de la sumatoria no fueron definidos, para ello definamos:

$$WN = \exp(-j2\pi/N) \quad (2.18)$$

La expresión (2.18) significa que hemos dividido al círculo unitario uniformemente en N espacios. Así:

$$W^{N-nk} = \exp(j2\pi kn/N)$$

es una secuencia de valor complejo de período N en n y k . Por lo tanto, en la expresión (2.16) hay apenas N exponenciales complejas diferentes. Con esto los límites de la sumatoria de (2.16) serán:

$$\tilde{x}(n) = \sum_{k=0}^{N-1} \tilde{X}(k) \cdot \exp(j2\pi kn / N) \quad (2.19)$$

Para obtener $\tilde{X}(k)$, multiplicamos ambos miembros de (2.19) por $\exp(-j2\pi kr/N)$ y luego realizamos el sumatorio desde $n=0$ hasta $n=N-1$, obteniéndose:

$$\tilde{X}(k) = \frac{1}{N} \sum_{n=0}^{N-1} \tilde{x}(n) \cdot \exp(-j2\pi kn / N) \quad (2.20)$$

El par (2.19) y (2.20) es conocido como Serie Discreta de Fourier (DFS), más por convención es utilizado el par:

$$\tilde{X}(k) = \sum_{n=0}^{N-1} \tilde{x}(n) \cdot \exp(-j2\pi kn / N) \quad (2.21)$$

$$\tilde{x}(n) = \frac{1}{N} \cdot \sum_{k=0}^{N-1} \tilde{X}(k) \cdot \exp(j2\pi kn / N) \quad (2.22)$$

La razón por la que se adopta la convención (2.21) y (2.22) es analizada a seguir.

II.3.2. Relación entre la Tr[Z] y la DFS

Sea $x(n)$ una secuencia de longitud finita obtenida a partir de una secuencia periódica $\tilde{x}(n)$ de período N de la siguiente manera:

$$x(n) = \tilde{x}(n) \quad 0 \leq n \leq N-1$$

$$x(n) = 0 \quad \text{otros } n$$

Con esto:

$$X(z) = \sum_{n=0}^{N-1} x(n) \cdot z^{-n}$$

Se observa que si $X(z)$ es evaluada en N puntos uniformemente espaciados sobre el círculo unitario tenemos:

$$\text{Para: } z_k = \exp(j2\pi k/N)$$

se llega a:

$$X(z) \Big|_{z=z_k} = \sum_{n=0}^{N-1} x(n) \cdot \exp(-j2\pi kn / N) = \tilde{X}(k) \quad (2.23)$$

Comparando (2.23) con (2.21), concluimos que la evaluación de la Tr[Z] de un período de la secuencia $\tilde{x}(n)$, en N puntos uniformemente espaciados sobre el círculo unitario, genera los coeficientes $\tilde{X}(k)$ de la DFS de la secuencia $\tilde{x}(n)$.

II.3.3. Propiedades de la DFS

Linealidad:

$$\tilde{x}_1(n) \xleftrightarrow{DFS} \tilde{X}_1(k)$$

$$\tilde{x}_2(n) \xleftrightarrow{DFS} \tilde{X}_2(k)$$

ambas con el mismo período

Si $\tilde{x}_3(n) = a.\tilde{x}_1(n) + b.\tilde{x}_2(n)$, entonces:

$$\tilde{X}_3(k) = a.\tilde{X}_1(k) + b.\tilde{X}_2(k)$$

Desplazamiento

$$\tilde{x}(n) \xleftrightarrow{DFS} \tilde{X}(k)$$

$$\tilde{x}(n+m) \xleftrightarrow{DFS} \exp(j2\pi km/N).\tilde{X}(k)$$

$$\tilde{X}(k+m) \xleftrightarrow{DFS} \exp(-j2\pi km/N).\tilde{x}(n)$$

Convolución

$$\text{Si } \tilde{x}_1(n) \xleftrightarrow{DFS} \tilde{X}_1(k)$$

$$\tilde{x}_2(n) \xleftrightarrow{DFS} \tilde{X}_2(k)$$

ambas con período N

$$\tilde{x}_3(n) = ? \xleftrightarrow{DFS} \tilde{X}_3(k) = \tilde{X}_1(k).\tilde{X}_2(k)$$

Es posible llegar a mostrar que:

$$\tilde{x}_3(n) = \sum_{m=0}^{N-1} \tilde{x}_1(n).\tilde{x}_2(n-m)$$

Ejemplo:

$\tilde{x}_1(n) = \tilde{x}_2(n)$ como se ilustra en el gráfico:

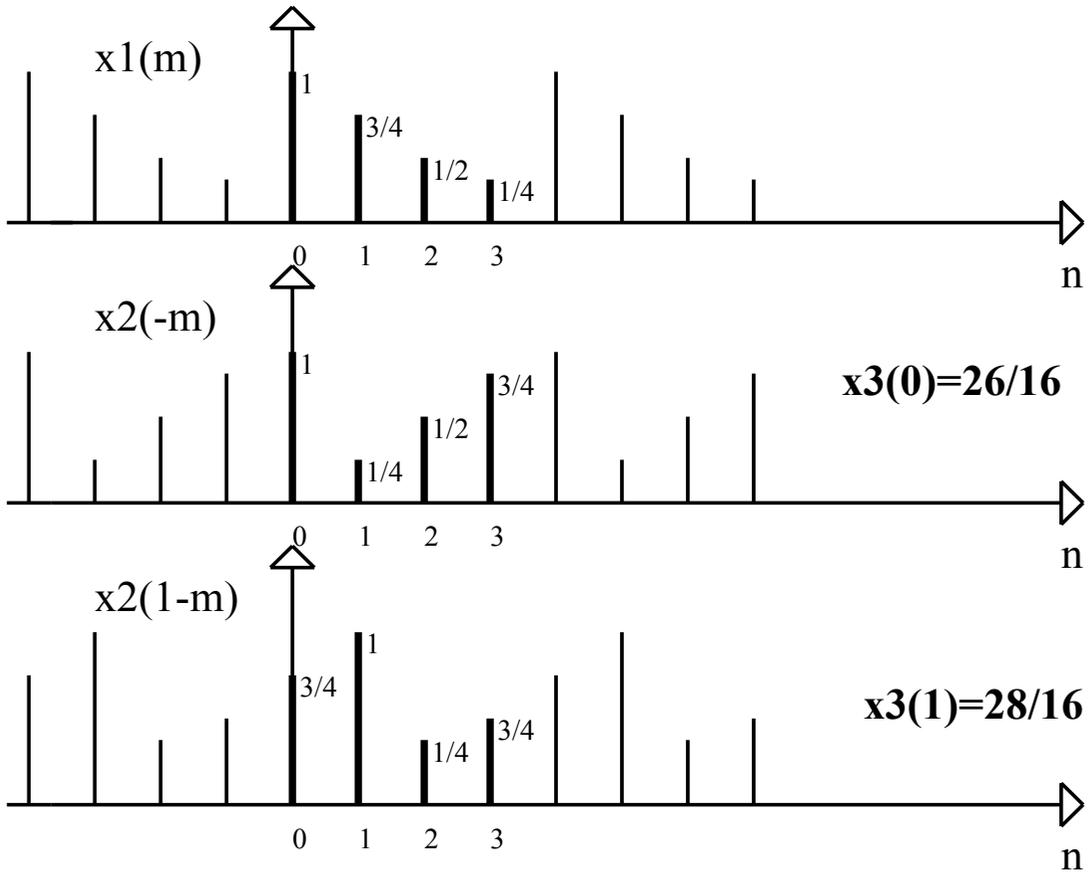
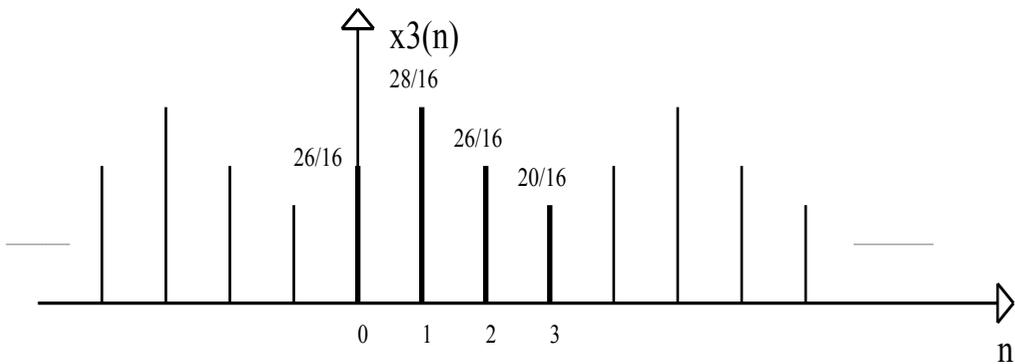


Figura 2.7 Ilustración de la convolución

Resultando:

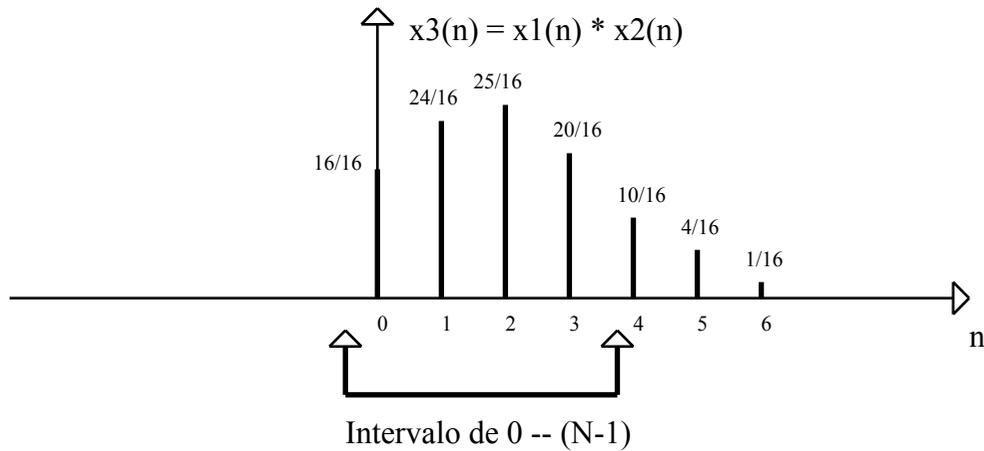


Comparando con el resultado de la convolución lineal de las secuencias:

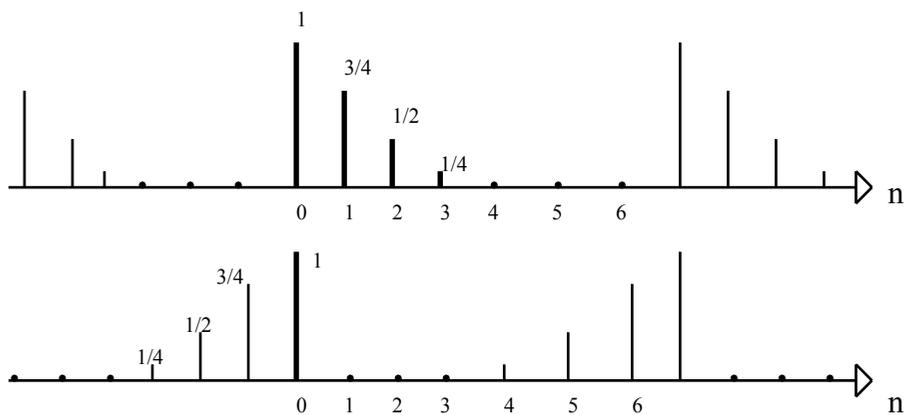
$$x_1(n) = \tilde{x}_1(n) \text{ en el intervalo } 0 \leq n \leq N-1$$

$$x_2(n) = \tilde{x}_2(n) \text{ en el intervalo } 0 \leq n \leq N-1$$

y 0 para otros n



Observando los gráficos de $x_3(n)$ y $\tilde{x}_3(n)$ en el intervalo $[0, N-1]$ concluimos que no son el mismo resultado. Mas si se inserta $(N-1)$ ceros al final de cada período en las dos secuencias $\tilde{x}_1(n)$ y $\tilde{x}_2(n)$, el resultado de la convolución circular es igual al resultado de la convolución lineal de $x_1(n)$ y $x_2(n)$ en el intervalo $[0, 2(N-1)]$, como ilustra el siguiente gráfico.



Es evidente que el resultado de la convolución circular de las secuencias anteriores en el intervalo $[0, 2(N-1)]$ es igual a la convolución lineal de $x_1(n)$ con $x_2(n)$. Al aumentar ceros en las secuencias, la $TR[Z]$ de $x_1(n)$ y $x_2(n)$ no varía, más el número de puntos de evaluación de la $Tr[Z]$ ha aumentado de N a $2N-1$.

II.3.4. Transformada Discreta de Fourier (DFT)

a.-) Definición:

Considere una secuencia finita $x(n)$: $0 \leq n \leq N-1$. A partir de $x(n)$ finita creamos una secuencia periódica $\tilde{x}(n)$ con período N , de la siguiente forma:

$$\tilde{x}(n) = \sum_{k=-\infty}^{\infty} x(n+k.N) = x(n)_N$$

Se observa que:

$$x(n) = \tilde{x}(n) \cdot RN(n)$$

donde: $RN(n) = 1 \quad 0 \leq n \leq N-1$
 $RN(n) = 0 \quad \text{otros } n$

Recordando que para $\tilde{x}(n)$ existe un par DFS, periódicas con período N . Reescribiendo (2.21) y (2.22):

$$\tilde{X}(k) = \sum_{n=0}^{N-1} \tilde{x}(n) \cdot \exp(-j2\pi kn / N)$$

$$\tilde{x}(n) = \frac{1}{N} \cdot \sum_{k=0}^{N-1} \tilde{X}(k) \cdot \exp(j2\pi kn / N)$$

Entonces, se define la DFT de $x(n)$ como:

$$X(k) = \tilde{X}(k) \cdot R_N(k)$$

$$x(n) = \tilde{x}(n) \cdot R_N(n)$$

Por lo tanto:

$$\begin{aligned} X(k) &= \sum_{n=0}^{N-1} x(n) \cdot \exp(-j2\pi kn / N) & 0 \leq k \leq N-1 \\ X(k) &= 0 & \text{otros } k \end{aligned} \quad (2.24)$$

$$\begin{aligned} x(n) &= \sum_{k=0}^{N-1} X(k) \cdot \exp(j2\pi kn / N) & 0 \leq n \leq N-1 \\ x(n) &= 0 & \text{otros } n \end{aligned} \quad (2.25)$$

Se observa la alta similaridad con la DFS. De hecho la DFT es definida como siendo igual a la DFS en una ventana rectangular de longitud N, razón por la cual en algunos libros especializados no llevan en cuenta la DFS.

Nota: A pesar que la DFT es aplicada a secuencias no periódicas, en el cálculo de la misma se **asume** implícitamente que la secuencia se torna periódica.

b.-) Propiedades

Apenas por un aspecto formal, las propiedades de la DFT serán enumeradas, mas como ya fue indicado, éstas serán similares a las de la DFS, dentro de una ventana rectangular de duración N.

i.) Linealidad:

Sean: $x_1(n)$ de longitud N_1 $x_2(n)$ de longitud N_2

Si: $x_3(n) = a \cdot x_1(n) + b \cdot x_2(n)$ luego $X_3(k) = a \cdot X_1(k) + b \cdot X_2(k)$

Apenas una observación: En la DFS se supuso que las secuencias eran de igual longitud. En este caso, para determinar $X_1(k)$ y $X_2(k)$, se debe obtener la $\text{Tr}[Z]$ de $x_1(n)$ y $x_2(n)$ y muestrearlas en un número $N_3 = \max(N_1, N_2)$ de muestras, o, en su defecto, incrementamos ceros en la secuencia de menor longitud.

ii.) Desplazamiento Circular:

$$x_1(n) = x[(n+m)]_N \cdot R_N(n)$$

$$X_1(k) = \exp(j2\pi km/N) \cdot X(k)$$

Analogamente:

$$X_1(k) = X[(k+l)]_N \cdot R_N(k)$$

$$x_1(n) = \exp(-j2\pi lm/N) \cdot x(n)$$

iii.) Convolución Circular:

Dadas:

$$x_1(n) \xleftrightarrow{DFT} X_1(k) \quad \text{de longitud } N$$

$$x_2(n) \xleftrightarrow{DFT} X_2(k) \quad \text{de longitud } N$$

$$x_3(n) = ? \xleftrightarrow{DFT} X_1(k) \cdot X_2(k) = X_3(k)$$

Consideremos la propiedad del desplazamiento circular. Supongamos que $x_3(n)$ puede expresarse como:

$$x_3(n) = \left[\sum_{m=0}^{N-1} x_1[m]_N \cdot x_2[n-m]_N \right] \cdot R_N(n) \quad (2.26)$$

$$x_3(n) = \left[\sum_{m=0}^{N-1} \tilde{x}_1(m) \cdot \tilde{x}_2(n-m) \right] \cdot R_N(n)$$

Observando en la ecuación anterior, la expresión en corchetes representa la convolución circular de dos secuencias periódicas. Aplicando la propiedad de la convolución circular de la DFS tenemos que:

$$\tilde{X}_1(k).\tilde{X}_2(k) \xleftrightarrow{DFS} \sum_{m=0}^{N-1} \tilde{x}_1(m).\tilde{x}_2(n-m)$$

Multiplicando a ambos lados de la ecuación por una ventana de duración N:

$$[\tilde{X}_1(k).\tilde{X}_2(k)].R_N(k) \xleftrightarrow{DFS} \left[\sum_{m=0}^{N-1} \tilde{x}_1(m).\tilde{x}_2(n-m) \right].R_N(n)$$

Mas $\tilde{X}_1(k).R_N(k) = X_1(k)$ y $\tilde{X}_2(k).R_N(k) = X_2(k)$

Con esto, obtenemos:

$$X_1(k).X_2(k) \longleftrightarrow \left[\sum_{m=0}^{N-1} \tilde{x}_1(m).\tilde{x}_2(n-m) \right].R_N(n) \quad (2.27)$$

Comparando (2.26) y (2.27), concluimos que:

$$X_3(k) = X_1(k).X_2(k)$$

Asi, esta propiedad puede ser resumida de la siguiente manera:

Si $x_3(n) = x_1(n) \otimes x_2(n)$ donde \otimes denota convolución circular, luego:

$$x_3(n) = \text{DFT}^{-1}[X_1(k).X_2(k)]$$

iv.) Convolución Lineal:

Es evidente que el análisis que se realizó para obtener la convolución lineal a partir de la convolución circular en la DFS sirve para la DFT. Lo que haremos aqui es generalizar el resultado obtenido para cuando las secuencias sean de longitud diferente.

Si $x_1(n)$ es una secuencia de longitud N_1 y $x_2(n)$ es otra de longitud N_2 , el resultado de la convolución de $x_1(n)$ con $x_2(n)$ es una secuencia $x_3(n)$ de longitud $N_3=N_1+N_2-1$.

La obtención de la convolución lineal, a partir de convolución circular, es ilustrada en el diagrama siguiente:

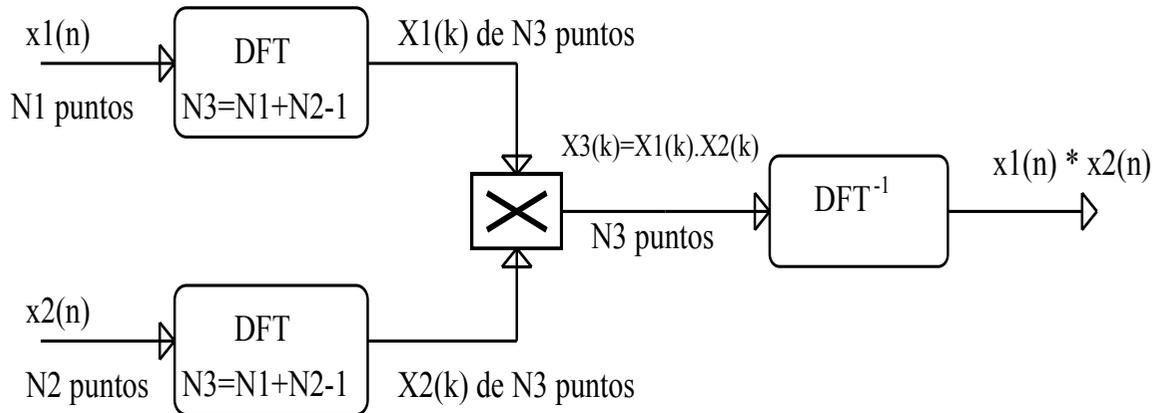


Figura 2.8 Implementación de la convolución usando la DFT

II.3.5. Interpretaciones de la DFT

Los coeficientes de la DFT, $X(k)$, son muestras de la $\text{Tr}[Z]$ de $x(n)$, tomadas a intervalos $[\exp(j2\pi/N)]$ uniformemente espaciados en el círculo unitario. La DFT es una combinación lineal de exponenciales complejas muestreadas y armónicamente relacionadas donde los $\{X(k)\}$ son los coeficientes usados en la representación.

Recordando las propiedades de ortogonalidad y ortonormalidad de las exponenciales complejas, entonces, la DFT es una representación en un espacio ortonormal donde las exponenciales complejas muestreadas son las bases del espacio.

A partir de la expresión de $X(k)$, es posible expresarse como:

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot \exp(j2\pi k(m-n)/N) \cdot \exp(-j2\pi km/N)$$

$$X(k) = \underbrace{\exp(-j2\pi km/N)}_{\text{apenas un desfase}} \cdot \underbrace{\sum_{n=0}^{N-1} x(n) \cdot \exp[j2\pi k(m-n)/N]}_{\text{sumatoria de convolución: } x(m) * h_k(m)}$$

donde: $h_k(n) = \exp(j2\pi kn/N)$, es decir que la DFT puede esquematizarse como:

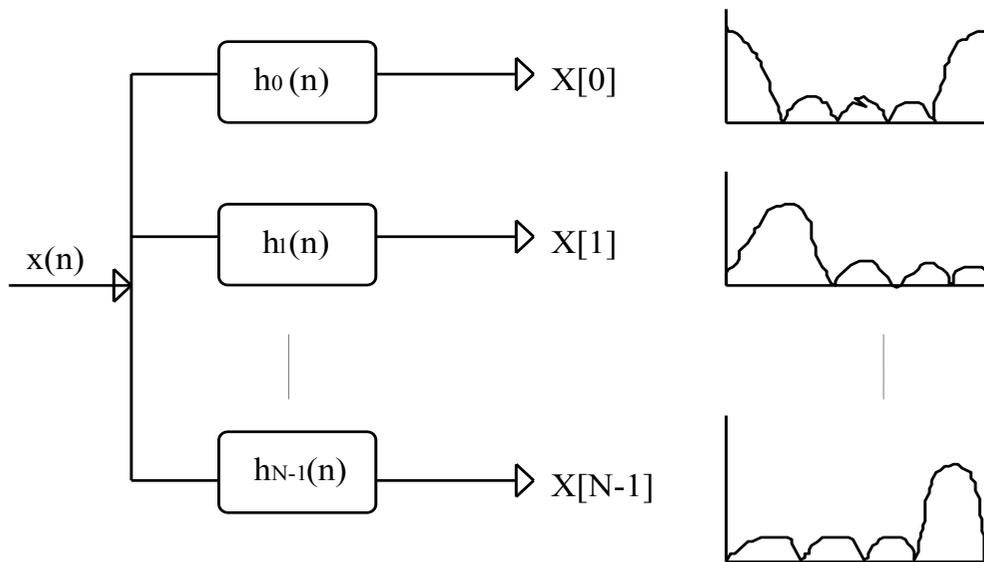


Figura 2.9 Interpretación de la DFT como un banco de filtros

Técnica Zero-Padding

En muchos casos, en la práctica, se dispone apenas de una secuencia $x(n)$ de corta duración de un fenómeno de interés. La técnica zero-padding intenta evitar posibles ambigüedades en la determinación del espectro de la secuencia.

Si nosotros incrementamos con N_2 ceros la secuencia $x(n)$ de N_1 puntos y extraemos la DFT (N_1+N_2 puntos) de la secuencia aumentada, es equivalente a extraer la $\text{Tr}[Z]$ de la secuencia original y muestrearla a intervalos uniformes en N_1+N_2 puntos.

El hecho de incrementar ceros a la secuencia original no implica en un aumento de la resolución del estimador del espectro, ya que si la $\text{Tr}[Z]$ se mantiene, por lo tanto $X(e^{j\omega})$ que es la $\text{Tr}[Z]$ evaluada en el círculo unitario, también se mantiene. Este efecto será ilustrado de mejor manera con los ejemplos.

II.4. Fast Fourier Transform (FFT)

La Fast Fourier Transform (FFT), es un algoritmo para computar la DFT de una secuencia. El par DFT (2.24) y (2.25) son reproducidas por comodidad:

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot \exp(-j2\pi kn / N) \quad 0 \leq k \leq N-1$$

$$X(k) = 0 \quad \text{otros } k$$

$$x(n) = \sum_{k=0}^{N-1} X(k) \cdot \exp(j2\pi kn / N) \quad 0 \leq n \leq N-1$$

$$x(n) = 0 \quad \text{otros } n$$

Para su cálculo, son necesarias N multiplicaciones complejas y $(N-1)$ sumas complejas por cada coeficiente $X(k)$. Ya que son N coeficientes de la DFT o DFT⁻¹, el número de operaciones para realizar una DFT de N puntos es:

$$N^2 \quad \text{multiplicaciones complejas} \quad (2.28.a)$$

$$N(N-1) \quad \text{sumas complejas} \quad (2.28.b)$$

Básicamente, la FFT aprovecha las propiedades de periodicidad y simetría del operador:

$$W_N^{nk} = \exp(-j2\pi nk / N) \quad (2.29)$$

Así, es posible mostrar que:

$$W_N^{nk} \text{ es periódica en } n \text{ y } k \text{ con período igual a } N \quad (2.30)$$

$$W_N^{(n+mN)(k+lN)} = W_N^{nk} \quad \text{con } n, k, l, m \text{ enteros}$$

y

$$W_N^{k(N-m)} = \left(W_N^{nk} \right)^* \quad (2.31)$$

Algoritmo de Decimación en el Tiempo: FFT de raíz 2

Vamos a suponer que el número N de muestras de la secuencia que se analiza es una potencia de 2, esto es, $N=2^M$. Dada una secuencia $\{x(n); n=0, \dots, N-1\}$, separamos en dos secuencias:

$x_e(n)$ formada por las muestras de $x(n)$ de índice par

$x_o(n)$ formada por las muestras de $x(n)$ de índice impar

De esta forma:

$$x_e(n) = x(2n) \quad 0 \leq n \leq N/2 - 1$$

$$x_o(n) = x(2n+1) \quad 0 \leq n \leq N/2 - 1$$

Por lo tanto, $X(k)$ puede expresarse como:

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot W_N^{nk}$$

$$X(k) = \sum_{n=par}^{N-1} x(n) \cdot W_N^{nk} + \sum_{n=impar}^{N-1} x(n) \cdot W_N^{nk} \quad \text{para } k=0, \dots, N-1$$

$$X(k) = \sum_{n=0}^{\frac{N}{2}-1} x_e(n) \cdot W_N^{2nk} + \sum_{n=0}^{\frac{N}{2}-1} x_o(n) \cdot W_N^{(2n+1)k}$$

La última ecuación puede ser escrita como:

$$X(k) = \underbrace{\sum_{n=0}^{\frac{N}{2}-1} x_e(n) \cdot W_{N/2}^{nk}}_{\text{DFT de } N/2 \text{ puntos}} + W_N^k \cdot \underbrace{\sum_{n=0}^{\frac{N}{2}-1} x_o(n) \cdot W_{N/2}^{nk}}_{\text{DFT de } N/2 \text{ puntos}}$$

$$X(k) = X_e(k) + W_N^k \cdot X_o(k) \quad k = 0, \dots, N-1 \quad (2.32)$$

Aprovechando que $\tilde{X}_e(k)$ y $\tilde{X}_o(k)$ son periódicas con período igual a $N/2$, entonces (2.32) puede expresarse:

$$X(k) = X_e(k) + W_N^k \cdot X_o(k) \quad 0 \leq k \leq N/2 - 1 \quad (2.33)$$

$$X(k) = X_e(k - N/2) + W_N^{k-N/2} \cdot X_o(k - N/2) \quad \frac{N}{2} \leq k \leq N-1$$

En la última expresión, aprovechamos la periodicidad de W_N , esto es: $W_N^{k-N/2} = -W_N^k$, con lo que obtenemos:

$$X(k) = X_e(k - N/2) - W_N^k \cdot X_o(k - N/2) \quad N/2 \leq k \leq N-1 \quad (2.34)$$

Las expresiones (2.33) y (2.34) son ilustradas en el siguiente gráfico para $N=8$.

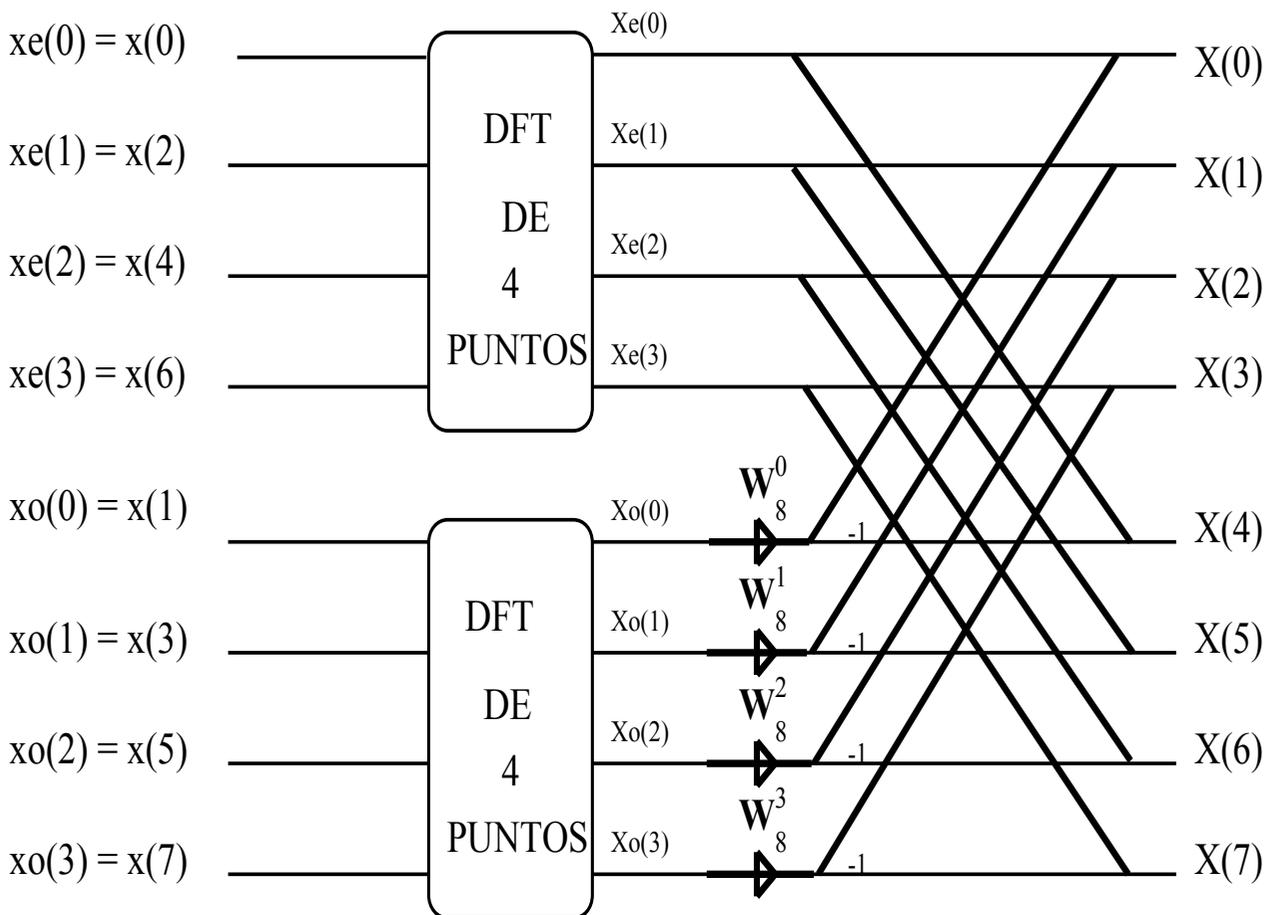


Figura 2.10 Implementación parcial de la FFT

Aplicando el mismo raciocinio a las secuencias $x_e(n)$ y $x_o(n)$, esto es, coger las muestras de índice par e impar de cada una, las DFT's de $N/2$ puntos pueden ser implementadas a través de 2 DFT's de $N/4$ puntos, llegando a mostrar que $X_e(k)$ y $X_o(k)$ pueden expresarse como:

$$Xe(k) = A(k) + W_N^{2k} \cdot B(k) \quad 0 \leq k \leq N/4 - 1$$

$$Xe(k) = A(k - N/4) - W_N^{2k} \cdot B(k - N/4) \quad N/4 \leq k \leq N/2 - 1$$

y

$$Xo(k) = C(k) + W_N^{2k} \cdot D(k) \quad 0 \leq k \leq N/4 - 1$$

$$Xo(k) = C(k - N/4) - W_N^{2k} \cdot D(k - N/4) \quad N/4 \leq k \leq N/2 - 1$$

Este procedimiento sigue hasta llegar a la DFT de dos puntos. En el siguiente gráfico es ilustrada la FFT completa con N=8.

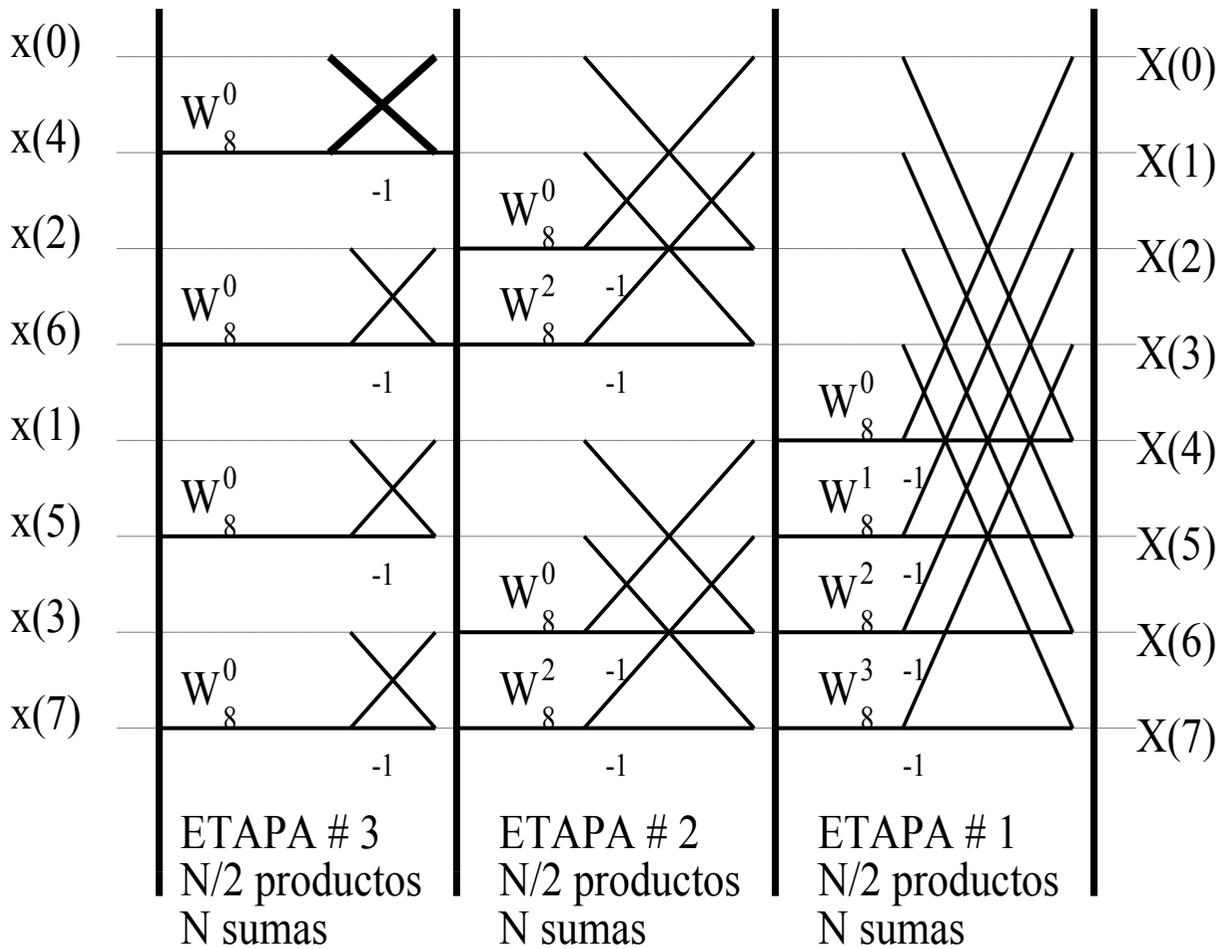


Figura 2.11 Implementación total de la FFT

Observaciones:

- Las entradas son reordenadas en tanto que las salidas están ordenadas
- La mariposa marcada con negro intenso es denominada computación "IN PLACE".
- El factor que multiplica una de las entradas de la mariposa de raíz-2 es llamado FACTOR DE ROTACION o "Twiddle Factor". Para cada etapa de la FFT son necesarios $N/2$ factores de rotación. El número de etapas de la FFT es $\log_2 N$. El número de multiplicaciones y sumas complejas por etapa es de $(N/2)$ y N , respectivamente. Esto implica que el número total de multiplicaciones y sumas complejas es:

$(N/2) \cdot \log_2 N$ multiplicaciones complejas

$N \cdot \log_2 N$ sumas complejas

Hacemos a continuación una comparación del costo computacional, considerando apenas las multiplicaciones complejas, de la DFT y la FFT:

N	DFT (N^2)	FFT ($(N/2) \cdot \log_2(N)$)	
16	256	32	8:1
64	4096	192	21:1
256	65536	1024	64:1

A continuación veremos cómo es que los datos de entrada son reordenados. Para esto, refirámonos a la ilustración siguiente:

$$S_x(e^{j\omega}) = \sum_{n=-\infty}^{\infty} R_x(n) \cdot \exp(-j\omega n) \quad (2.35)$$

Por tanto, toda la teoría de la DTFT, DFS, DFT (FFT), que fue desarrollada para señales y sistemas determinísticos tienen su contraparte en procesos estocásticos a través del análisis de la función de autocorrelación del proceso.

Como conocemos, un PEESADT está formado por un "ensamble" de realizaciones. Una sola realización está formada por un conjunto de variables aleatorias. Por tanto, si intentamos representarla en el dominio de Fourier a esta realización, lo que se va a obtener es un conjunto de variables aleatorias que no representan la PSD del proceso.

Normalmente no se dispone de toda la información de $R_x(n)$ sino apenas de una ventana de corta duración $[n=-M, \dots, M]$, con esto, (2.35) se expresa como:

$$\hat{S}_x(e^{j\omega}) = \sum_{n=-M}^M R_x(n) \cdot \exp(-j\omega n) \quad (2.36)$$

El espectro resultante es la convolución de la PSD verdadera $S_x(e^{j\omega})$, y la DTFT de una secuencia $\{w(n)=1, n=-M, \dots, M\}$.

Ahora, en la práctica solo se dispone de un pequeño bloque de una realización del proceso, digamos: $\{x(n), n=0, \dots, N-1\}$. Por lo tanto, debemos estimar $R_x(n)$ utilizando cualquiera de los estimadores estudiados en la primera parte de este seminario. Si utilizáramos el estimador biased, $\hat{R}_x(n)$, la estimativa del espectro resultante es conocido con el nombre de PERIODOGRAMA, $P_N(\omega)$.

$$P_N(\omega) = \sum_{m=-(N-1)}^{N-1} \hat{R}_x(n) \cdot \exp(-j\omega n) \quad (2.37)$$

Substituyendo el valor de $\hat{R}_x(n)$ en la expresión anterior y llevando en consideración que $X(e^{j\omega}) = \sum_{n=0}^{N-1} x(n) \cdot \exp(-j\omega n)$, (2.37) se reduce a:

$$P_N(\omega) = \frac{1}{N} |X(e^{j\omega})|^2 \quad (2.38)$$

(2.38) es también conocida como PERIODOGRAMA DE SCHUSTER (1890-1900) [11]. Es posible mostrar que este estimador de la PSD del PEESADT, $X(nT)$, incrementa su varianza a medida que el número N de datos disponibles aumenta. Esto quiere decir que el Periodograma Clásico, PC, no es un estimador consistente.

La popularidad del PC nace del ahorro computacional que representa la determinación de $X(ej)$ a partir del algoritmo de la FFT.

Referencias Bibliográficas

- [01] Akaike, H. "A new look at the statistical model identification", *IEEE Trans. on Automatic Control*, vol. AC-19, pp. 255-265, 1974.
- [02] Atal, B.S., and Hanauer, S.L., "Speech analysis and synthesis by linear prediction of the speech wave", *J. Acoust. Soc. Amer.*, vol. 50, pp. 637-655, 1971.
- [03] Box, G.E.P., Jenkins, G.M., *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, 1976.
- [04] Burg, J.P., "Maximum Entropy Spectral Analysis", Ph.D. dissertation, Stanford University, Stanford Calif., 1975.
- [05] Flanagan, J.L., and others, "Speech coding", *IEEE Trans. on Communications*, vol. 27, pp. 710-737, 1979.
- [06] Gray, A.H. and Markel, J.D., "A Normalized digital filter structure", *IEEE Trans. on ASSP*, vol. ASSP-23, pp. 268-277, 1975.
- [07] Haykin, S., "Radar Signal Processing", *IEEE ASSP Magazine*, vol. 2, # 2, 1985.
- [08] Haykin, S., *Adaptive Filter Theory*, Prentice-Hall, Englewood Cliffs, N.J., 1991.
- [09] Jayant, N.S., Noll, P., *Digital Coding of Waveforms*, Prentice-Hall, Englewood Cliffs, N.J., 1984.
- [10] Kailath, T., "A view of three decades of linear filtering theory", *IEEE Trans. on Information Theory*, vol. IT-20, pp 146-181, 1974.
- [11] Kay, S.M., *Modern Spectral Estimation – Theory & Application*, Prentice Hall, 1999
- [12] Makhoul, J., "Linear Prediction: A tutorial review", *Proc. IEEE*, vol. 63, pp. 561-580, 1975. Popular Articles Nov 2014
- [13] Oppenheim, A.V., and Schaffer, R.W., *Discrete-Time Signal Processing*, 3rd. Edition Prentice-Hall Signal Processing, 2009.
- [14] Papoulis, A., Unnikrishna, P., *Probability, Random Variables, and Stochastic Processes*, 4th ed., International Edition, 2002.
- [15] Rabiner, L.R., and Gold, B., *Theory and Application of Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, N.J., 1975.
- [16] Rabiner, L.R., and Schaffer, R.W., *Digital Processing of Speech Signals*, Prentice-Hall, Englewood Cliffs, N.J., 1978.
- [17] Robinson, E.A., and Treitel, S., *Geophysical Signal Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1980.

[18] **Rabiner, L.**, " A tutorial on hidden Markov models and selected applications in speech recognition", *Proc. IEEE*, vol. 63, pp. 561-580, 1975. Popular Articles Nov 2014

Contenido

Prefacio	4
Parte I: Probabilidades y Procesos Estocásticos	6
I.1 PROBABILIDADES	7
I.1.1 Introducción a las Probabilidades	
I.1.2 Probabilidad Condicional	
I.1.3 Probabilidad Total y Teorema de Bayes	
I.1.4 Independencia, Sucesos Independientes	
I.1.5 Experimentos Repetidos	
I.1.6 Experimento de Bernoulli	
I.2 VARIABLES ALEATORIAS	23
I.2.1 Definición	
I.2.2 Función Distribución de Probabilidad	
I.2.3 Función Densidad de Probabilidad	
I.2.4 Distribuciones Condicionales	
I.2.5 Probabilidad Total y Teorema de Bayes	
I.2.6 Funciones de una Variable Aleatoria	
Media de una variable aleatoria	
Varianza de una variable aleatoria	
Momentos	
I.2.7 Estadísticas Conjuntas de dos variables aleatorias	
Independencia	
Momentos Conjuntos	
Correlación	
Covarianza	
Coeficiente de Correlación	
I.3 PROCESOS ALEATORIOS	44
I.3.1 Introducción	
I.3.2 Procesos Estocásticos Estacionarios	
I.3.3 Media y Funciones de Correlación y Covarianza	
I.3.4 Ergodicidad y Medias Temporales	
I.3.5 Densidad Espectral de Potencia de un PEESA	
I.3.6 Procesos Aleatorios a través de Filtros	
I.3.7 Procesos Estocásticos Discretos en el Tiempo	
I.3.8 Estimación Cuadrática Media	
I.3.9 Predicción Lineal	

Parte II: Procesamiento Digital de Señales	62
II.1 SEÑALES Y SISTEMAS DISCRETOS EN EL TIEMPO	66
II.1.1 Relación entre formas de representación de sistemas o señales	
II.1.2 Esquema Básico de Procesamiento Digital de Señales	
II.1.3 Ventajas y puntos que deben observarse sobre el procesamiento digital en relación al analógico	
II.1.4 Señales	
II.1.5 Sistemas	
II.1.6 Respuesta en Frecuencia	
II.1.7 Relación entre señal discreta y continua en el tiempo y la frecuencia	
II.2 TRANSFORMADA Z	75
II.2.1 Secuencias y Regiones de Convergencia de su $Tr[Z]$	
II.2.2 Función del Sistema	
II.3 TRANSFORMADA DISCRETA DE FOURIER	80
II.3.1 Serie Discreta de Fourier (DFS)	
II.3.2 Relación entre la $Tr[Z]$ y la DFS	
II.3.3 Propiedades de la DFS	
II.3.4 Transformada Discreta de Fourier (DFT)	
II.3.5 Interpretaciones de la DFT	
II.4 FAST FOURIER TRANSFORM (FFT)	91
II.5 ESTIMACION DE LA DSP DE UN PEESADT	97
Referencias bibliográficas	100



ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA



ISBN: 978-9978-301-28-9



9 789978 301289