

ROMA
TRE



Università degli studi di Roma Tre

Facoltà di Ingegneria

*Appunti dalle lezioni del corso di
Teoria dei Segnali*

PROCESSI ALEATORI

Prof. Alessandro Neri

Roma, novembre 2001

INDICE

<u>PROCESSI ALEATORI.....</u>	<u>3</u>
1. INTRODUZIONE.....	3
2. DEFINIZIONI	4
3. PROCESSI DETERMINISTICI.....	ERRORE. IL SEGNALIBRO NON È DEFINITO.
4. MEDIE D'INSIEME E MEDIE NEL TEMPO ...	12
5. PROCESSI STAZIONARI	15
<u>ERGODICITÀ.....</u>	<u>20</u>
<u>0</u>	
<u>TRASFORMAZIONI LINEARI DI</u>	
<u>PROCESSI.....</u>	<u>25</u>
<u>TABELLA 1.....</u>	<u>30</u>
<u>PROCESSI GAUSSIANI LIMITATI IN BANDA.....</u>	<u>31</u>
<u>INVILUPPO E FASE DI UN PROCESSO.....</u>	<u>34</u>
<u>PROCESSO DI WIENER.....</u>	<u>39</u>
<u>SERIE DI MARKOV.....</u>	<u>46</u>

Processi Aleatori

1. INTRODUZIONE

Dal punto di vista concettuale possiamo definire come segnale aleatorio un segnale il cui andamento temporale non sia noto a priori, in parte o nella sua totalità, e chiameremo sorgente aleatoria la sorgente in grado di generarlo. Ovviamente, come per i fenomeni aleatori, l'aleatorietà non rappresenta una caratteristica intrinseca del segnale, ma è associata direttamente al grado di conoscenza posseduta.

Il procedimento per la costruzione di strumenti matematici che consentono lo sviluppo e la manipolazione di modelli statistici per tali segnali si sviluppa in modo simile a quanto avviene per lo studio dei fenomeni aleatori, rappresentandone un caso limite. In particolare il singolo risultato di un esperimento si identifica con il particolare segnale aleatorio emesso da una particolare sorgente che prende il nome di realizzazione, mentre l'insieme di tutte le possibili realizzazioni costituite (da un punto di vista informale) il processo aleatorio.

Pur rappresentando concettualmente un caso limite, lo studio dei processi aleatori si distingue da quello dei fenomeni aleatori per due fattori fondamentali:

- Esplicito riferimento alla struttura dinamica contrapposto alla visione statica propria dei fenomeni aleatori;

- Necessità di considerare situazioni che coinvolgono infinità non numerabili di eventi o di variabili aleatorie che non possono essere trattate sulla sola base della Teoria delle Probabilità senza assunzioni o ipotesi aggiuntive.

2. DEFINIZIONI

Sul piano formale un processo aleatorio può essere definito come segue:

Def.1 Dato uno spazio di probabilità (Q, F, P) ed un sottoinsieme $T \subseteq \mathbb{R}$ dell'asse reale, che denota l'intervallo d'interesse, si definisce processo aleatorio e si indica con:

$$X(t; \omega), \quad (1)$$

oppure con:

$$(x(t; \omega), t \in T) \quad (2)$$

una funzione reale X definita sullo spazio prodotto $T \times \Omega$:

$$X: T \times \Omega \rightarrow \mathbb{R} \quad (3)$$

Tale che per ogni istante $t' \in T$, $X(t'; \cdot)$ sia una variabile aleatoria, cioè ogni insieme del tipo:

$$A = (\omega \in \Omega, X(t'; \omega) \leq \xi) \quad (4)$$

Sia misurabile rispetto alla σ -algebra F (cioè $A \in F$) per ogni $t' \in T$ e per ogni $\xi \in \mathbb{R}$.

Si noti che fissato $t' \in T$, $X(t; \cdot)$ diviene una variabile aleatoria e pertanto un processo aleatorio risulta essere una famiglia di variabili aleatorie, indicizzate da un parametro reale t , definite su uno spazio di probabilità (Q, F, P) comune, mentre, fissato un punto $\omega \in \Omega$ si ottiene una funzione del tempo $x(t) = X(t; \omega)$ che prende il nome di *realizzazione* (o *funzione membro*) del processo.

La precedente definizione di processo aleatorio può essere generalizzata sia considerando funzioni X che abbiano come condominio \mathbb{R}^n (processi aleatori vettoriali) sia considerando funzioni che abbiano come codominio l'insieme dei numeri complessi \mathbb{C} (processi aleatori complessi scalari e vettoriali).

Qualora T sia una sequenza $\{t_1, t_2, \dots, t_n\}$ di n istanti di tempo non necessariamente equispaziati, il processo diviene una sequenza $\{X(t_1), X(t_2), \dots, X(t_n)\}$ di variabili aleatorie e prende il nome di *processo aleatorio tempo discreto* o *serie aleatoria*. Un possibile andamento di una realizzazione di una serie aleatoria è riportato in fig. 1a

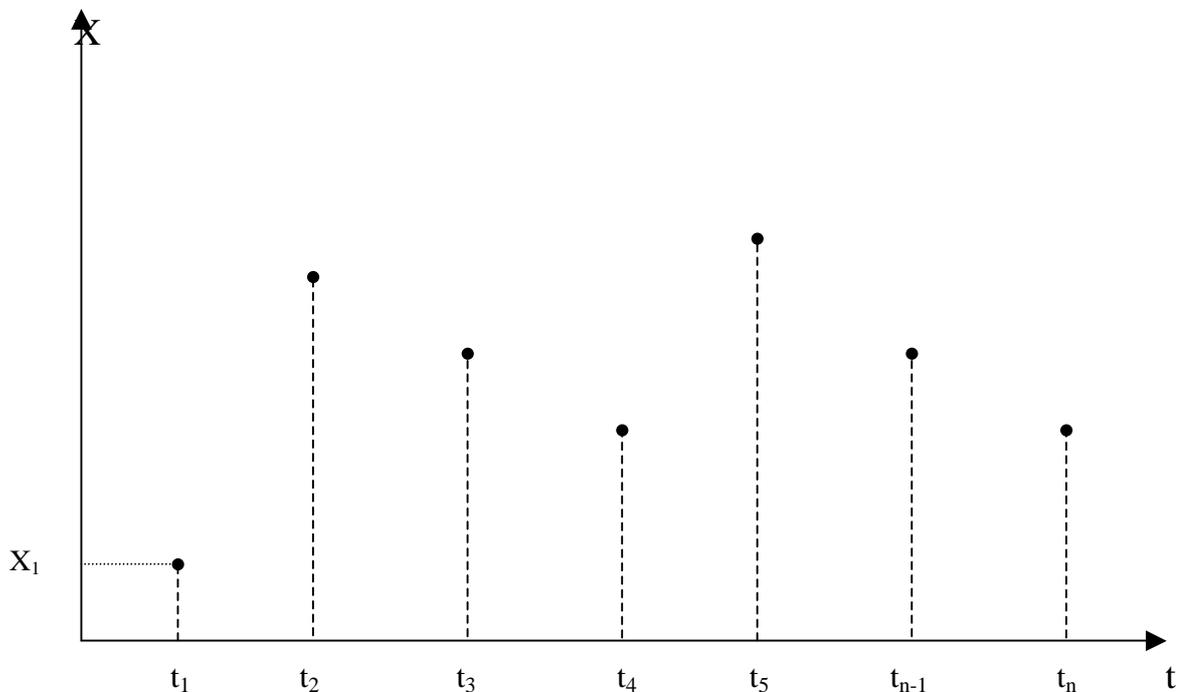


FIG. 1a

Se invece T è un intervallo di \mathbb{R} allora $X(t; \omega)$ diviene una famiglia di variabili aleatorie dipendenti da un parametro continuo ed il processo prende il nome di *processo aleatorio tempo continuo*. In tal caso le realizzazioni del processo possono assumere andamenti del tipo di quelli riportati in fig. 1b.

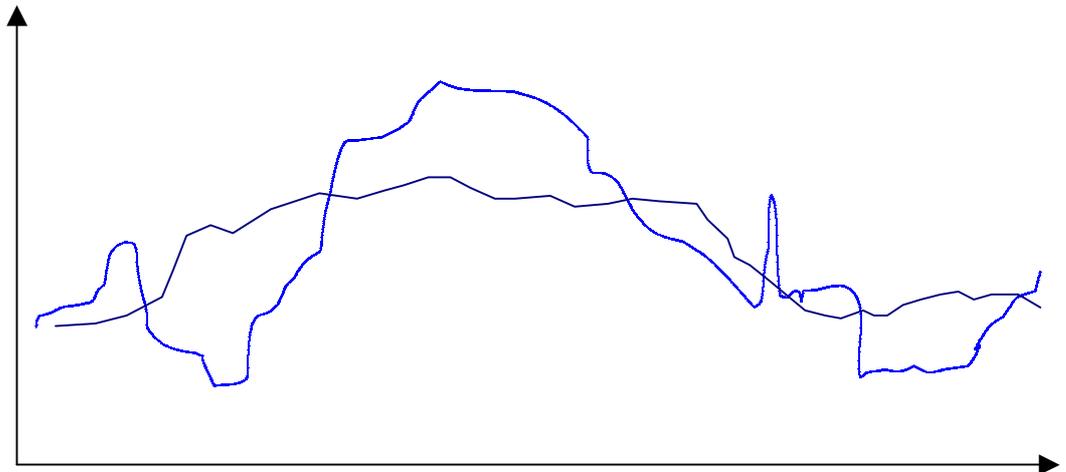


FIG. 1b

La precedente definizione di processo aleatorio richiede che, per ogni $t' \in T$, $X(t'; \cdot)$ sia una variabile aleatoria, ma non pone alcuna condizione su di esso come funzione del tempo. Allo scopo di caratterizzare completamente il processo aleatorio dal punto di vista probabilistico supponiamo inizialmente che T sia costituito da un numero discreto N di valori t_1, t_2, \dots, t_n non necessariamente equispaziati. Corrispondentemente una generica realizzazione del processo è costituita dalla determinazione della variabile aleatoria N – dimensionale :

$$X = (X_1, X_2, \dots, X_N) \tag{5a}$$

essendo:

$$X_1 = X(t_1), X_2 = X(t_2), \dots, X_N = X(t_N) \tag{5b}$$

La descrizione statistica del processo è completamente individuata, in tal caso, dalla funzione di distribuzione congiunta;

$$D_{X_1 X_2 \dots X_N} (X_1, X_2, \dots, X_N) \quad (6)$$

da cui possono essere calcolate per saturazione le funzioni di distribuzione marginali.

Se ora il numero di istanti temporali diviene infinito numerabile, la funzione di distribuzione congiunta non è più assegnabile, poiché diviene infinita la dimensione della variazione aleatoria. In questo caso si può pensare di procedere per gradi, assegnando prima le funzioni di distribuzione delle variabili aleatorie unidimensionali x_j estratte negli istanti di tempo t_j :

$$D_{X_j} (x_j; t_j) = \Pr (X(t_j) \leq x_j), \forall t_j \in T \quad (7)$$

o, nel caso in cui esse esistano, le relative f. d. p.:

$$p_{X_j} (x_j; t_j), \quad \forall t_j \in T \quad (7')$$

che chiameremo statistiche del primo ordine.

Successivamente si possono assegnare le funzioni di distribuzione congiunte relative a tutte le coppie di variabili aleatorie unidimensionali:

$$D_{X_j X_i} (X_j, X_i; t_j, t_i) \quad \forall t_i, t_j \in T \quad (8)$$

o, nel caso in cui esse esistano, le relative f. d. p. :

$$p_{X_j X_i} (X_j, X_i; t_j, t_i) \quad \forall t_i, t_j \in T \quad (8')$$

che chiameremo statistiche del secondo ordine, e così di seguito, sino ad assegnare le funzioni di distribuzione congiunte relative a tutte le n – uple di variabili aleatorie unidimensionali:

$$D_{X_{i1} X_{i2} \dots X_{in}} (X_{i1}, X_{i2}, \dots, X_{in}; t_{i1}, t_{i2}, \dots, t_{in}) \quad \forall t_{i1}, t_{i2}, \dots, t_{in} \in T \quad (9)$$

o, nel caso in cui esse esistano, le relative f. d. p. :

$$p_{X_{i1}X_{i2} \dots X_{iN}} (X_{i1}, X_{i2}, \dots, X_{iN}; t_{i1}, t_{i2}, \dots, t_{iN}), \quad \forall t_{i1}, t_{i2}, \dots, t_{iN} \in T \quad (9')$$

che chiameremo statistiche di ordine n.

Ovviamente tale collezione di distribuzione di dimensioni finita deve soddisfare delle condizioni di consistenza, in quanto dalle statistiche di ordine N è possibile derivare, per saturazione, tutte le statistiche di ordine inferiore.

Def. 2 Si definisce *gerarchia di ordine n* di un processo aleatorio l'insieme delle statistiche fino all'ordine n, cioè:

$$\begin{aligned} D_{X_j} (X_{i1}; t_{i1}) & \quad \forall t_{i1} \in T \\ D_{X_{i1} X_{i2}} (X_{i1}, X_{i2}; t_{i1}, t_{i2}) & \quad \forall t_{i1}, t_{i2} \in T \\ D_{X_{i1} X_{i2} \dots X_{iN}} (X_{i1}, X_{i2}, \dots, X_{iN}; t_{i1}, t_{i2}, \dots, t_{iN}) & \quad \forall t_{i1}, t_{i2}, \dots, t_{iN} \in T \end{aligned} \quad (10)$$

o, nel caso in cui esistano le f. d. p. :

$$\begin{aligned} p_{X_{i1}} (x_{i1}; t_{i1}), & \quad \forall t_{i1} \in T \\ p_{X_{i1}X_{i2}} (X_{i1}, X_{i2}; t_{i1}, t_{i2}) & \quad \forall t_{i1}, t_{i2} \in T \\ p_{X_{i1}X_{i2} \dots X_{iN}} (X_{i1}, X_{i2}, \dots, X_{iN}; t_{i1}, t_{i2}, \dots, t_{iN}), & \quad \forall t_{i1}, t_{i2}, \dots, t_{iN} \in T \end{aligned} \quad (10')$$

La conoscenza del processo è tanto migliore quanto più elevato è l'ordine della gerarchia assegnata. E' possibile inoltre dimostrare che [Doob, 1954, pagg. 47 – 48], indicata con F_T la σ - algebra generata dalla classe di insiemi di punti $\omega \in \Omega$ del tipo:

$$(\omega: \omega \in \Omega, X(t; \omega) \leq \xi \quad \forall t \in T, \quad \forall \xi \in R) \quad (11)$$

l'insieme delle gerarchie di ordine finito induce in modo univoco una misura di probabilità su F_T .

Nel caso in cui l'insieme T sia un intervallo di R , gli istanti di tempo costituiscono un'infinità non numerabile e quindi le variabili aleatorie marginali costituiscono un insieme non numerabile di variabili aleatorie. Ciò comporta che la descrizione statistiche dei vari ordini relative ad un istante, due istanti, tre istanti, ecc., variando la distanza tra gli istanti e la loro allocazione assoluta in tutti i modi possibili.

La gerarchia di ordine n per un processo aleatorio a tempo continuo è dunque data dall'insieme delle funzioni di distribuzione:

$$\begin{aligned}
 D_{X_1}(X_1; t_1) & \quad \forall t_1 \in T \\
 D_{X_1 X_2}(X_1, X_2; t_1, t_2) & \quad \forall t_1, t_2 \in T \\
 D_{X_1 X_2 \dots X_n}(X_1, X_2, \dots, X_n; t_1, t_2, \dots, t_n) & \quad \forall t_1, t_2, \dots, t_n \in T
 \end{aligned} \tag{12}$$

o, nel caso in cui esistano, dalle corrispondenti f. d. p., in cui t_1, t_2, \dots, t_n assumono tutti i possibili valori in T .

Anche in tal caso la conoscenza del processo è tanto migliore quanto più è elevato l'ordine della gerarchia assegnata, ma al contrario di ciò che avviene per i processi tempo discreto l'insieme delle gerarchie di ordine finito non determina in modo univoco una misura di probabilità su F_T . Si osservi ad esempio che l'insieme di F_T :

$$(\omega : \omega \in \Omega, X(t; \omega) \leq 0 \quad \forall t \in T) = \mathbf{I}_{t \in T}(\omega : \omega \in \Omega, X(t; \omega) \leq 0) \tag{13}$$

può non corrispondere ad alcun evento in quanto esso implica una intersezione di una infinità non numerabile di eventi. Il fatto che esso rappresenti o no un evento dipende dalla struttura topologica dello spazio di probabilità. In particolare è possibile definire una classe di processi detti *separabili* per i quali

la conoscenza delle gerarchie di ordine finito determina completamente le proprietà statistiche del processo. L'utilità dell'introduzione di tale classe è motivata dal fatto che dato un qualsiasi processo aleatorio $X(t; \omega)$ definito sullo stesso spazio di probabilità tale che:

$$\Pr\{X(t; \omega) = X'(t; \omega)\} = 1, \quad \forall t \in T \quad (14)$$

Tale condizione ha come ovvia conseguenza che i due processi hanno le stesse gerarchie di ordine finito. Si noti che la condizione (14) implica che l'insieme $(\omega: \omega \in \Omega, x(t; \omega) \neq x'(t; \omega))$ corrisponde all'evento nullo, ma non fornisce vincoli sull'insieme:

$$(\omega: \omega \in \Omega, x(t; \omega) = x'(t; \omega) \text{ per almeno un } t \in T) = \bigcup_{t \in T} \{\omega: \omega \in \Omega, x(t; \omega) = x'(t; \omega)\} \quad (15)$$

che potrebbe anche non corrispondere a nessun evento, in quanto implica un'infinità non numerabile di unioni.

Sul piano formale un processo separabile può essere definito come segue:

Def. 3 Un processo $X(t; \omega)$ si dice *separabile* se esiste un insieme misurabile $S \subset T$ ed un evento nullo Λ tale che, per ogni insieme chiuso $K \subset (-\infty, +\infty)$ ed ogni intervallo aperto I , i due insiemi:

$$\{\omega: \omega \in \Omega, x(t; \omega) \in K, t \in I \cap T\} \quad \{\omega: \omega \in \Omega, x(t; \omega) \in K, t \in I \cap S\}$$

differiscano per un sottoinsieme di Λ .

3. PROCESSI DETERMINISTICI

Tra tutti i processi aleatori esiste una classe la cui descrizione statistica risulta estremamente semplice. Tale situazione si verifica qualora lo spazio base ha dimensione finita ed esiste una corrispondente biunivoca tra i punti $\omega \in \Omega$ e le funzioni del tempo $x(t; \omega)$. In tal caso, in virtù della corrispondenza biunivoca, ω assume la funzione di un parametro, noto il quale risulta completamente individuata la realizzazione.

Per fare un esempio supponiamo di disporre di un generatore ai cui capi sia presente un segnale sinusoidale (oscillatore) acceso in un istante non noto, ma molto remoto; allora il segnale d'uscita del generatore potrà essere posto nella forma:

$$a \cos(2\pi f_0 t + \phi_0) \quad (16)$$

dove, non essendo noto l'istante di accensione, ϕ_0 rappresenta la determinazione di una v.a. Φ_0 definita in $[0, 2\pi)$. Noto ϕ_0 risulta, in tal caso, completamente individuata la forma d'onda. Per descrivere tale processo è quindi sufficiente conoscere la funzione di distribuzione $D_{\Phi_0}(\phi_0)$.

Tali tipi di processi prendono il nome di *processi deterministici* o ad aleatorietà parametrica. Si può asserire che la caratteristica qualificante dei processi deterministici risiede nel fatto che l'aleatorietà si manifesta in maniera statica, localizzata nel tempo all'istante iniziale dell'emissione da parte della sorgente come scelta di una determinazione particolare del parametro.

4. MEDIE D'INSIEME E MEDIE NEL TEMPO

Nel par.1 è stato mostrato come le informazioni statisticamente circa il processo possano essere espresse per mezzo delle gerarchie di probabilità. D'altronde, fissati n istanti di tempo t_1, t_2, \dots, t_n , il processo individua una variabile aleatoria n – dimensionale $X = (X_1, X_2, \dots, X_n)$ con $X_i = X(t_i)$, per cui ha senso calcolare il valore atteso di una funzione $f(X)$ di tale variabile aleatoria.

Poiché, in virtù della legge dei grandi numeri, tale valore atteso può essere calcolato tramite il limite del valore medio di tale funzione calcolata su un numero M di realizzazioni diverse $x^j(t)$ del processo, quando M tende ad infinito (vedi fig. 2) tale valore atteso viene indicato con il termine di media d'insieme (di ordine n) di f , intendendo con *insieme* l'insieme di tutte le possibili realizzazioni del processo.

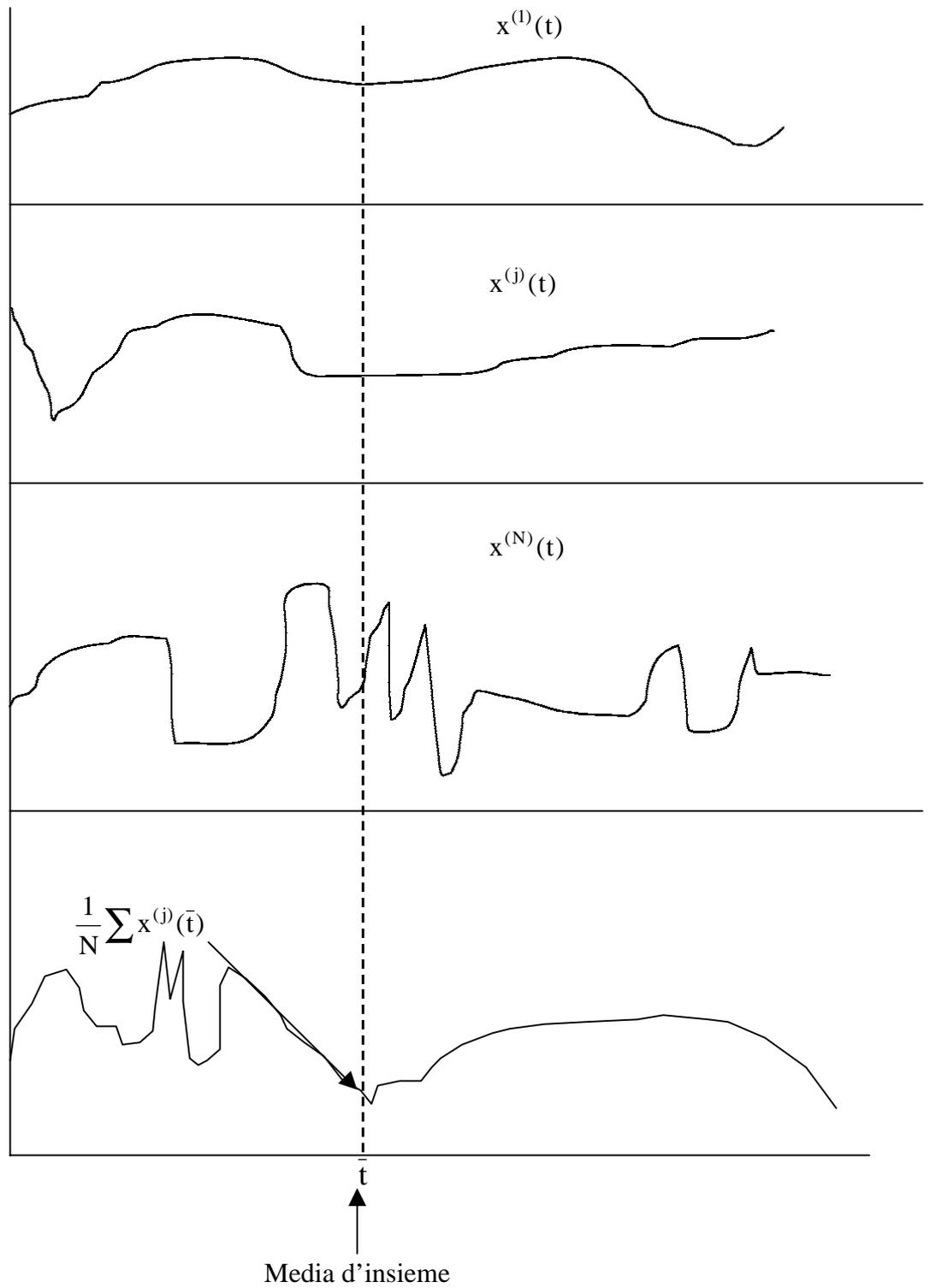


FIG. 2

Def. 4 Data una funzione $f(X_1, X_2, \dots, X_n)$ si definisce media d'insieme di ordine n di f il valore atteso:

$$E\{f(X_1, X_2, \dots, X_n)\}. \quad (17)$$

Si osservi che tale quantità dipende dalla funzione di distribuzione $D_{X_1 X_2 \dots X_n}(X_1, X_2, \dots, X_n; t_1, t_2, \dots, t_n)$ e pertanto risulta essere in generale dipendente dagli n istanti di tempo t_1, t_2, \dots, t_n .

Tra le possibili medie d'insieme risultano di particolare rilevanza nelle applicazioni le seguenti:

- Valore atteso del processo: $m_X(t_1) = E\{X(t_1)\}; \quad (18)$

- Valore quadratico atteso del processo: $m_X^{(2)}(t_1) = E\{X^2(t_1)\}; \quad (19)$

- Momento misto di ordine (1,1) del processo:
 $m_X^{(1,1)}(t_1, t_2) = E\{X(t_1) X(t_2)\} \quad (20)$

- Funzione di covarianza del processo:
 $K_X^{(1,1)}(t_1, t_2) = E\{[X(t_1) - m_X(t_1)] [X(t_2) - m_X(t_2)]\} \quad (21)$

Altre medie che possono essere considerate, oltre alle medie d'insieme, sono quelle effettuate su una singola realizzazione del processo $x^{(j)}(t)$ e che prendono il nome di *medie temporali*.

Def. 5 Data una funzione $f(X_1, X_2, \dots, X_n)$ si definisce media temporale di ordine n di f il limite :

$$\begin{aligned} \langle f[x^{(j)}(t_1), x^{(j)}(t_2), \dots, x^{(j)}(t_n)] \rangle &= \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} f[x^{(j)}(t+t_1), x^{(j)}(t+t_2), \dots, x^{(j)}(t+t_n)] dt \end{aligned} \quad (22)$$

Tra le possibili medie temporali risultano di particolare rilevanza nelle applicazioni le seguenti:

- Valore medio:

$$\langle x^{(j)} \rangle = \langle x^{(j)}(t_1) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x^{(j)}(t+t_1) dt \quad (23)$$

- Potenza:

$$P = \langle [x^{(j)}(t_1)]^2 \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} [x^{(j)}(t+t_1)]^2 dt \quad (24)$$

- Funzione di autocorrelazione:

$$R_x(t_2 - t_1) = \langle X^{(j)}(t_1) X^{(j)}(t_2) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} [X^{(j)}(t+t_1) X^{(j)}(t+t_2)] dt \quad (25)$$

5. PROCESSI STAZIONARI

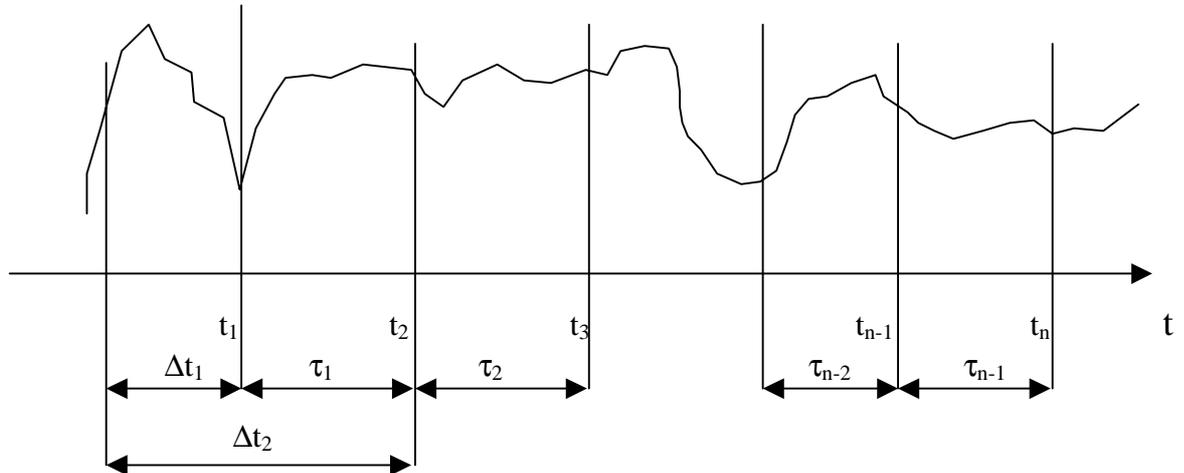
Def. 6 Un processo aleatorio si dice stazionario in senso stretto se le sue proprietà statistiche sono invarianti rispetto ad una traslazione temporale

Ciò implica che fissati n intervalli di tempo $\Delta t_1, \Delta t_2, \dots, \Delta t_n$, la f.d.p. congiunta

$$p_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n; t_0 + \Delta t_1, t_0 + \Delta t_2, \dots, t_0 + \Delta t_n) \quad (26)$$

Relativa alle v. a. X_i estratte negli istanti di tempo $t_i = t_0 + \Delta t_i$ non dipende da t_0 , ma solo dalla allocazione relativa degli istanti di tempo t_1, t_2, \dots, t_n , completamente individuata dai tempi Δt_i con $i = 1, 2, \dots, n-1$.

In alternativa agli intervalli Δt_i è usuale nella letteratura considerare gli intervalli $\tau_i = t_{i+1} - t_i$ (vedi fig.)



In particolare, come conseguenza diretta della definizione si hanno le seguenti proprietà di particolare rilevanza nelle applicazioni:

- La gerarchia del primo ordine non dipende dall'istante di tempo considerato:

$$p_{X_1}(x_1; t_1) = p_{X_1}(x_1) \quad \forall t_1 \in T; \quad (27)$$

- Le medie d'insieme di ordine 1 sono indipendenti dal tempo. In particolare si ha:

$$m_X(t_1) = m_X \quad (28.a) \quad m_X^{(2)}(t_1) = m_X^{(2)} \quad (28.b)$$

- La gerarchia del secondo ordine dipende solo dalle differenze $\tau = t_2 - t_1$ tra gli istanti di tempo considerati:

$$p_{X_1 X_2}(x_1, x_2; t_1, t_2) = p_{X_1 X_2}(x_1, x_2; \tau) \quad \forall t_1, t_2 \in T \quad (29)$$

- Le medie d'insieme del secondo ordine non dipendono che dalla differenza $\tau = t_2 - t_1$ tra gli istanti di tempo considerati. In particolare si ha:

$$m_X^{(1,1)}(\tau) = \iint_{X_1 X_2} x_1 x_2 p_{X_1 X_2}(x_1, x_2; \tau) dx_1 dx_2 \quad (30.a)$$

$$K_X(\tau) = \iint_{x_1 x_2} (x_1 - m_{x_1})(x_2 - m_{x_2}) p_{x_1 x_2}(x_1, x_2; \tau) dx_1 dx_2 \quad (30.b)$$

Per illustrare i concetti precedenti consideriamo il processo armonico nel par.3 allo scopo di individuare per quale tipo di distribuzione di ϕ_0 tale processo risulti stazionario. A tale scopo iniziamo con il derivare la gerarchia di ordine 1 del processo.

$$\text{Poiché} \quad x(t) = a \cos(2\pi f_0 t + \phi_0) \quad (31)$$

la f. d. p. $p_{x_1}(x_1; t_1)$ può essere calcolata a partire da quella di ϕ_0 osservando che, fissato t_1 , ad una determinazione x_1 corrispondono due valori di ϕ_0 : dati da

$$\phi_0 = -2\pi f_0 t_1 \pm \arccos\left(\frac{x_1}{a}\right), \text{ mod. } 2\pi, \quad (32)$$

e pertanto applicando la regola per le funzioni di v. a. e osservando che per ambedue le funzioni inverse si ha:

$$\left| \frac{d\phi_0}{dx_1} \right| = \frac{1}{\sqrt{a^2 - x_1^2}}, \quad \forall x_1 \in [-a, a] \quad (33)$$

segue che:

$$p_{x_1}(x_1; t_1) = \frac{1}{\sqrt{a^2 - x_1^2}} \left\{ p_{\phi_0} \left[\left(-2\pi f_0 t_1 + \arccos\left(\frac{x_1}{a}\right) \right)_{\text{mod. } 2\pi} \right] + p_{\phi_0} \left[\left(-2\pi f_0 t_1 - \arccos\left(\frac{x_1}{a}\right) \right)_{\text{mod. } 2\pi} \right] \right\}, \quad \forall x_1 \in [-a, a] \quad (34)$$

Dalla precedente relazione si ha immediatamente che, affinché $p_{x_1}(x_1; t_1)$ non dipenda da t_1 , la $p_{\phi_0}(\phi_0)$ deve risultare costante su tutto l'intervallo $[0, 2\pi)$ e quindi la v. a. ϕ_0 deve risultare a distribuzione uniforme su tale intervallo, ottenendo

$$p_{X_1}(x_1) = \begin{cases} 0 & x_1 < -a \text{ o } x_1 > a \\ \frac{1}{\sqrt{a^2 - x_1^2}} & -a \leq x_1 \leq a \end{cases} \quad (35)$$

Tale condizione risulta essere, quindi, necessaria per la stazionarietà del processo armonico. Per dimostrare l'eventuale sufficienza occorrerebbe considerare le gerarchie di ordine n , con n qualsiasi.

In tal caso si può trarre vantaggio dal fatto che, per $n=2$

$$\begin{aligned} x_2 = x(t_2) &= a \cos(2\pi f_0 t_2 + \phi_0) = a \cos[(2\pi f_0 t_1 + \phi_0) + 2\pi f_0 \tau] = \\ &= x_1 \cos(2\pi f_0 \tau) \pm \sqrt{a^2 - x_1^2} \sin(2\pi f_0 \tau) \end{aligned} \quad (36)$$

e quindi la f. d. p. condizionata $p_{X_2/X_1}(x_2/x_1; t_1, t_2)$ dipende solo da τ e di conseguenza la f. d. p. congiunta

$$p_{X_1 X_2}(x_2, x_1; t_1, t_2) = p_{X_2/X_1}(x_2/x_1; \tau) p_{X_1}(x_1) \quad (37)$$

dipende solo da τ . Per $n > 2$ inoltre le v. a. X_3, X_4, \dots, X_n sono legate in modo biunivoco a X_1 e X_2 una volta fissati $\tau_1, \tau_2, \dots, \tau_{n-1}$.

In effetti essendo il processo deterministico, si può procedere in alternativa come segue. Considerando una traslazione Δt si ha:

$$x(t_1 + \Delta t) = a \cos[(2\pi f_0 t_1 + 2\pi f_0 \Delta t + \phi_0)] = a \cos(2\pi f_0 t_1 + \phi'_0) \quad (38)$$

con

$$\phi'_0 = \phi_0 + 2\pi f_0 \Delta t \pmod{2\pi} \quad (39)$$

pertanto il processo può essere pensato come dipendente dalla nuova v. a. ϕ'_0 . Affinché le sue proprietà statistiche risultino inalterate rispetto ad una traslazione arbitraria, le due variabili aleatorie ϕ_0 e ϕ'_0 devono avere la stessa f. d. p., ma poiché, in base alla (39), si ha:

$$p_{\phi'_0}(\phi'_0) = p_{\phi_0}[(\phi'_0 - 2\pi \Delta t)_{\text{mod. } 2\pi}] \quad (40)$$

ed inoltre l'eguaglianza deve valere per ogni Δt , ne consegue che tali condizioni sono verificate se e solo se $p_{\phi_0}(\phi_0)$ risulta essere costante in $[0, 2\pi)$.

Risulta dimostrato pertanto il seguente teorema:

Teorema 1 Il processo armonico $\{a \cos(2\pi f_0 t + \phi_0)\}$ risulta stazionario in senso stretto se e solo se la v. a. ϕ_0 è uniformemente distribuita nell'intervallo $[0, 2\pi)$.

Poiché la proprietà di stazionarietà in senso stretto risulta molto restrittiva nonché difficile da verificare, è utile introdurre una seconda forma di stazionarietà più debole della precedente.

Def. 7 Un processo aleatorio $X(t; \omega)$ si dice stazionario in senso lato se:

A. $m_X^{(2)}(t_1) < +\infty$ (41.a)

B. Il valore atteso è indipendente dal tempo $m_X(t_1) = m_X = \text{cost}$ (41.b)

C. La funzione di covarianza dipende solo dall'intervallo $\tau = t_2 - t_1$:

$$K_X(t_1, t_2) = K_X(t_2 - t_1) = K_X(\tau) \quad (41.c)$$

Ovviamente un processo aleatorio stazionario in senso stretto lo è anche in senso lato, mentre in generale non è vero il viceversa.



ERGODICITA'

La definizione di processo aleatorio come famiglia di variabili aleatorie indicizzate da un parametro implica che per poterne derivare le proprietà statistiche su base sperimentale occorre fare ricorso ad un numero sufficientemente elevato di realizzazioni ottenute da sorgenti di costituzione e condizioni di funzionamento identiche. Ciò non di meno è lecito chiedersi se esistano delle classi di processi per le quali tali informazioni possono essere dedotte a partire da una singola realizzazione del processo, ovvero in quali condizioni le medie temporali e le medie d'insieme coincidano. I processi per i quali, ciò accade vengono detti ergodici. La nozione di ergodicità ha fatto la sua prima apparizione nella teoria cinetica e nella meccanica classica statistica ad opera soprattutto di Maxwell, Boltzmann, Clausius e Gibbs, mentre le prime dimostrazioni matematiche di teoremi ergodici inerenti l'intercambiabilità tra medie temporali e medie d'insieme sono state date da Birkhoff (1931) e Von Neumann (1932). Ovviamente una definizione di "ergodicità" basata sull'intercambiabilità tra tali operazioni, anche se chiara dal punto di vista concettuale, rischia di non essere operativa (per verificare, infatti, l'ergodicità di un processo occorrerebbe disporre delle medie d'insieme) e di non mettere in luce la struttura interna del processo. L'approccio qui seguito è pertanto quello di definire l'ergodicità in modo formale in base a proprietà statistiche del processo e derivare successivamente dei teoremi ergodici che stabiliscono

l'intercambiabilità tra medie temporali e medie d'insieme. Il punto di partenza sul piano concettuale è costituito dall'osservazione che, dato un fenomeno aleatorio, ed un evento ammissibile E con probabilità $\Pr\{E\}$, la legge forte dei grandi numeri ci assicura che:

$$\lim_{N \rightarrow \infty} \frac{N_E}{N} = \Pr\{E\} \quad (\text{Pr ob.1}) \quad (42)$$

cioè che si ha la coincidenza tra il limite del rapporto di frequenza N_E/N e la probabilità $\Pr\{E\}$ dell'evento, a meno di un insieme di esperimenti di misura nulla, essendo N_E il numero di prove in cui si verifica l'evento E e N il numero di prove totali. Ora elemento essenziale nella derivazione della legge forte dei grandi numeri è costituito dal requisito che le prove siano statisticamente indipendenti. Se quindi tale condizione appare come un requisito naturale nel costruire il fenomeno aleatorio combinato costituito da N ripetizioni del fenomeno di partenza, essa non è in genere soddisfatta quando si passa ai processi.

Supposto infatti che $x(t)$ sia un processo stazionario, se si considerano le determinazioni x_1, x_2, \dots, x_n assunte dalle variabili aleatorie estratte negli n istanti di tempo t_1, t_2, \dots, t_n ; poiché la f. d. p. ad essa associata è la stessa in virtù della stazionarietà, esse potrebbero essere pensate come il risultato di N ripetizioni del fenomeno aleatorio associato alla singola variabile aleatoria X_i e si potrebbe pensare di applicare la legge dei grandi numeri. In tal caso però richiedere l'indipendenza statistica tra le prove condurrebbe a considerare solo la classe estremamente ristretta di processi per i quali le singole v. a. X_i siano mutualmente statisticamente indipendenti.

In effetti, come mostrato in [Doob, 1954, pagg 465 e seg.], affinché la legge dei grandi numeri possa essere applicata ciò che occorre è che, preso l'insieme di tutte le realizzazioni del processo, non appena si toglie da tale insieme un sottoinsieme di realizzazioni di misura di probabilità non nulla (e ovviamente diversa da 1) il processo che così si ottiene non sia stazionario.

Conseguentemente con tale impostazione si ha la seguente definizione di processo ergodico:

Def. 8 Un processo è ergodico se:
A. E' stazionario in senso stretto
B. Non contiene sottoinsiemi stazionari in senso stretto con probabilità diversa da 0 o da 1

La coincidenza tra medie d'insieme e medie temporali è poi assicurata dal seguente teorema, noto anche come *legge forte dei grandi numeri per processi stazionari in senso stretto*, per la cui dimostrazione si rimanda a [Doob, 1954, pagg 465 e seg.]:

Teorema 2 In un processo ergodico la media temporale di ordine n di una funzione di una realizzazione è uguale per tutte le realizzazioni (a meno di un sottoinsieme di probabilità nulla) e coincide con la media d'insieme della stessa funzione.

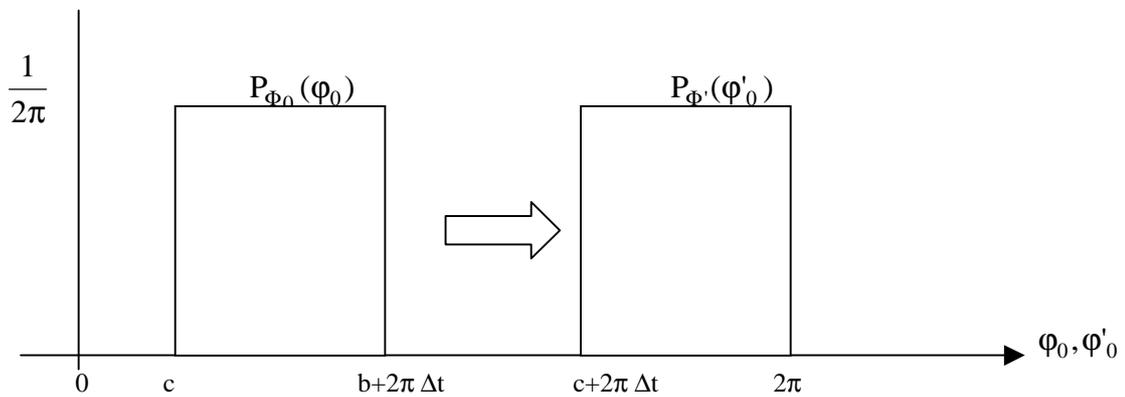
Commento: In essenza il teorema implica che ogni realizzazione di un processo ergodico contiene in sé tutta l'informazione possibile sul processo in quanto una sorgente ergodica produce, nel corso di una realizzazione, tutte le situazioni ed i casi possibili per il processo con una frequenza pari alla probabilità di detti eventi. Quindi prese M sorgenti aleatorie identiche, all'interno di ogni realizzazione si troveranno ovviamente con istanti di presentazione diversi, tutti i casi possibili.

Esempio: Allo scopo di illustrare l'operatività della definizione si consideri il processo armonico: $X(t; \omega)$ di cui $x(t) = a \cos(2\pi f_0 t + \phi_0)$ costituisce una realizzazione. E' stato già dimostrato nel par. 4 come tale processo sia stazionario in senso stretto. Se ora si

considera, tra tutte le realizzazioni, il sottoinsieme costituito dalle realizzazioni per cui $\phi_0 \in [b, c) \subset [0, 2\pi)$ si ha immediatamente che tale sottoinsieme a cui compete una misura di probabilità pari a:

$$\int_b^c p_{\phi_0}(\phi_0) d\phi = \frac{c-b}{2\pi} \quad (43)$$

non risulta stazionario in senso stretto, infatti con riferimento al par. 3 la v. a. ϕ'_0 presenta una f. d. p. data dalla eq. (40) che pertanto si modifica come rappresentato in fig.



Conseguentemente il processo è ergodico.

Qualora a sia invece la determinazione di una v. a. A per cui il processo risulta dipendente dalla coppia (A, ϕ_0) ogni sottoinsieme di realizzazione per cui $a \in [a_1, a_2]$ e $\phi_0 \in [0, 2\pi)$ risulta stazionario in senso stretto. Infatti si noti che la f. d. p. congiunta può essere posta nella forma

$$\begin{aligned} P_{X_1 X_2 \dots X_n} (x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \\ = P_{X_1 X_2 \dots X_n / A} (x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n / A) p_A(a) \end{aligned} \quad (44)$$

e che la f. d. p. condizionata è esattamente del tipo analizzato nel par.3 nel caso di a noto. Poiché tale sottoinsieme ha una misura di probabilità pari a

$$\int_{a_1}^{a_2} dD_A(a) \quad (45)$$

in generale diverso da 0 e 1; tale processo non è ergodico.

Si noti che il Teorema 2 impone la coincidenza tra medie temporali e medie d'insieme comunque si scelga la funzione $f(x_1, x_2, \dots, x_n)$. In effetti tale coincidenza potrebbe essere "investita" anche da un punto di vista differente, ovvero dato un processo aleatorio $X(t; \omega)$ ed una ben precisa funzione f (ad esempio $f(x_1) = x_1^2$) ci si potrebbe chiedere sotto quali condizioni risulti:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} f[x(t)] dt = E_{X_1} \{f(x_1)\} \quad (45)$$

[ovvero per definizione di limite in media quadratica

$$\lim_{T \rightarrow \infty} E \left\{ \left[\frac{1}{T} \int_{-T/2}^{T/2} f[x(t)] dt - E_{X_1} \{f(x_1)\} \right]^2 \right\} = 0 \quad (45')$$

Come mostrato in [Wong e Haick, 1985, pagg. 69 – 70] in tal caso una condizione sufficiente affinché la (45) sia valida è che risulti

$$\int_{-\infty}^{+\infty} |R_f(\tau)| d\tau < +\infty \quad (46)$$

in cui $R_f(\tau)$ è la funzione di covarianza di $f(x_1)$. Ovviamente in tal caso la (46) va verificata per ogni funzione d'interesse.

Dato un processo stazionario non ergodico, l'insieme di tutte le possibili realizzazioni può essere suddiviso in sottoinsiemi ergodici.

TRASFORMAZIONI LINEARI DI PROCESSI

1. INTRODUZIONE

Dato un processo aleatorio $X(t)$ ed un filtro con risposta impulsiva $h(t)$, sia $Y(t; \omega)$ il processo le cui realizzazioni $v(t)$ siano costituite dalle uscite del filtro quando in ingresso vengano applicate le realizzazioni $x(t)$ di $X(t; \omega)$:

$$\{y(t; \omega), t \in T\} = \{x(t; \omega) * h(t), t \in T\} \quad (1)$$

Allo scopo di analizzare il legame esistente tra le proprietà statistiche di $X(t; \omega)$ e $Y(t; \omega)$, si osservi che il filtro costituisce una trasformazione permanente. Di conseguenza se il processo d'ingresso è stazionario in senso stretto, anche il processo d'uscita risulterà stazionario in senso stretto, in quanto sia le proprietà statistiche dell'ingresso che il comportamento del filtro risulteranno invarianti rispetto alle traslazioni temporali. Inoltre se il processo d'ingresso è ergodico, anche il processo d'uscita è ergodico.

Il legame esistente tra le medie d'insieme del 1° e del 2° ordine di $X(t; \omega)$ e $Y(t; \omega)$ possono essere derivate in modo semplice dalle regole relative al transito di un segnale certo attraverso un filtro e ricordando che l'ipotesi di ergodicità

implica la coincidenza tra medie temporali e le medie d'insieme. In particolare si esaminerà ora il caso del valore atteso e della funzione di covarianza:

- Valore Atteso:

Per il calcolo del valore atteso $m_Y(t)$ di $Y(t; \omega)$ si osservi che presa una realizzazione $x(t)$ di $X(t; \omega)$ per il valor medio di $y(t)$ vale la seguente relazione:

$$\begin{aligned} \langle y(t) \rangle = \overline{y(t)}^t &= \int_{-\infty}^{+\infty} h(\tau) \overline{x(t-\tau)}^t d\tau = \int_{-\infty}^{+\infty} h(\tau) \overline{x(t-\tau)}^t d\tau = \\ &= \int_{-\infty}^{+\infty} h(\tau) \overline{x(t-\tau)}^t d\tau = \overline{x(t)}^t \cdot \int_{-\infty}^{+\infty} h(\tau) d\tau \end{aligned} \quad (2)$$

Se quindi il processo $X(t; \omega)$ è ergodico si ha:

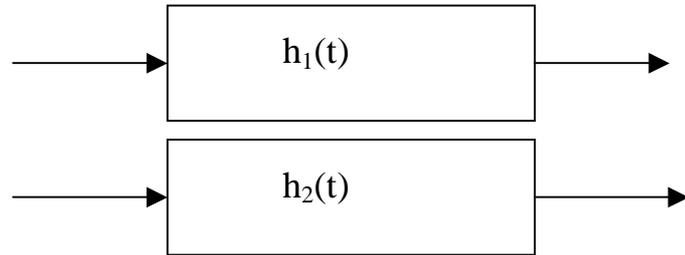
$$m_Y = E\{y(t)\} = \overline{y(t)}^t = \overline{x(t)}^t \int_{-\infty}^{+\infty} h(\tau) d\tau = m_X \cdot \int_{-\infty}^{+\infty} h(\tau) d\tau \quad (3)$$

Si noti che l'integrale della risposta impulsiva del filtro è uguale al valore della funzione di trasferimento $H(f)$ calcolato in $f=0$. Pertanto la precedente relazione può anche essere posta nella forma:

$$m_Y = m_X \cdot H(0) \quad (4)$$

- Funzione di covarianza:

Per il calcolo del legame tra la funzione di covarianza $K_{XX}(t)$ dell'ingresso e quella dell'uscita $K_{YY}(t)$ conviene esaminare il caso più generale di una coppia di processi $X_1(t; \omega)$ $X_2(t; \omega)$ le cui realizzazioni transitano in due filtri con risposte impulsive $h_1(t)$ e $h_2(t)$.



In tal caso la funzione di intercorrelazione $P_{Y_1Y_2}(\tau)$ tra le uscite $y_1(t)$ e $y_2(t)$ dei due filtri, per definizione è pari a:

$$P_{Y_1Y_2}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} y_1^*(t) y_2(t + \tau) dt \quad (5)$$

Poiché inoltre $y_1(t)$ e $y_2(t)$ sono legati a $x_1(t)$ e a $x_2(t)$ dalle note relazioni:

$$y_1(t) = \int_{-\infty}^{+\infty} h_1(\vartheta) x_1(t - \vartheta) d\vartheta \quad (6.a)$$

$$y_2(t) = \int_{-\infty}^{+\infty} h_2(\lambda) x_2(t - \lambda) d\lambda \quad (6.b)$$

si ha:

$$\begin{aligned} P_{Y_1Y_2}(\tau) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-\infty}^{+\infty} h_1^*(\vartheta) x_1^*(t - \vartheta) d\vartheta \int_{-\infty}^{+\infty} h_2(\lambda) x_2(t + \tau - \lambda) d\lambda dt = \\ &= \int_{-\infty}^{+\infty} h_1^*(\vartheta) \int_{-\infty}^{+\infty} h_2(\lambda) \left[\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x_1^*(t - \vartheta) x_2(t + \tau - \lambda) dt \right] d\lambda d\vartheta = \\ &= \int_{-\infty}^{+\infty} h_1^*(\vartheta) \int_{-\infty}^{+\infty} h_2(\lambda) \cdot P_{X_1X_2}(\tau - \lambda + \vartheta) d\lambda d\vartheta \end{aligned} \quad (7)$$

posto, quindi, $\mu = \lambda - \vartheta$ si ha:

$$\begin{aligned} P_{Y_1Y_2}(\tau) &= \int_{-\infty}^{+\infty} P_{X_1X_2}(\tau - \mu) \int_{-\infty}^{+\infty} h_1^*(\vartheta) h_2(\vartheta + \mu) d\vartheta d\mu = \\ &= \int_{-\infty}^{+\infty} P_{X_1X_2}(\tau - \mu) E_{h_1h_2}(\mu) d\mu = P_{X_1X_2}(\tau) * E_{h_1h_2}(\tau) \end{aligned} \quad (8)$$

Si noti che ponendo $h_1(t) = u_0(t)$ si ha:

$$y_1(t) = x_1(t) \quad (9)$$

$$E_{h_1 h_2}(t) = u_0(t) \otimes h_2(t) = h_2(t) \quad (10)$$

e quindi

$$P_{Y_1 Y_2}(\tau) = P_{X_1 X_2}(\tau) * h_2(t) \quad (11)$$

Analogamente posto $h_2(t) = u_0(t)$ si ha:

$$y_2(t) = x_2(t) \quad (12)$$

$$E_{h_1 h_2}(t) = h_1(t) \otimes u_0(t) = h_1^*(-t) \quad (13)$$

$$\begin{aligned} P_{Y_1 Y_2}(\tau) &= P_{X_1 X_2}(\tau) * h_1^*(-t) = h_1^*(-t) * P_{X_1 X_2}(\tau) = \\ &= h_1(t) \otimes P_{X_1 X_2}(\tau) \end{aligned} \quad (14)$$

Le regole relative alla trasformazione di un solo segnale possono essere dedotte immediatamente da quelle relative ad una coppia di segnali. Infatti posto $x_1(t) = x_2(t) = x(t)$ e $h_1(t) = h_2(t) = h(t)$ dalla (8) si ha:

$$P_{Y Y}(\tau) = P_{X X}(\tau) * E_{hh}(t) \quad (15)$$

Inoltre dalle (11) e (14) si ha:

$$P_{X Y}(\tau) = P_{X X}(\tau) * h(t) \quad (16)$$

$$P_{Y X}(\tau) = P_{X X}(\tau) \otimes h(t) \quad (17)$$

Le precedenti relazioni sono riassunte in Tabella 1 insieme alle corrispondenti relazioni nel dominio della frequenza.

Se il processo $X(T)$ è ergodico dalle (15) e (16) segue che

$$\boxed{m_{YY}^{(1,1)}(\tau) = m_{XX}^{(1,1)} * E_{hh}(\tau)} \quad (18)$$

$$\mathbf{m}_{XY}^{(1,1)}(\tau) = \mathbf{m}_{XX}^{(1,1)} * \mathbf{h}(\tau) \quad (19)$$

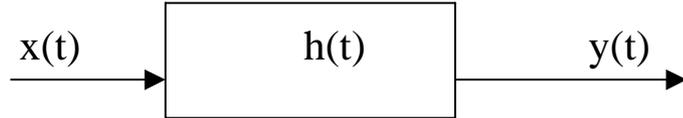
Più in generale se $X_1(t; \omega)$ e $X_2(t; \omega)$ sono congiuntamente ergodici, dalle (8) e (11) si ha

$$\mathbf{m}_{Y_1 Y_2}^{(1,1)}(\tau) = E \{ y_1^*(t) y_2(t + \tau) \} = P_{Y_1 Y_2}(\tau) = P_{X_1 X_2}(\tau) * E_{h_1 h_2}(\tau) = \mathbf{m}_{X_1 X_2}^{(1,1)}(\tau) * E_{h_1 h_2}(\tau) \quad (20)$$

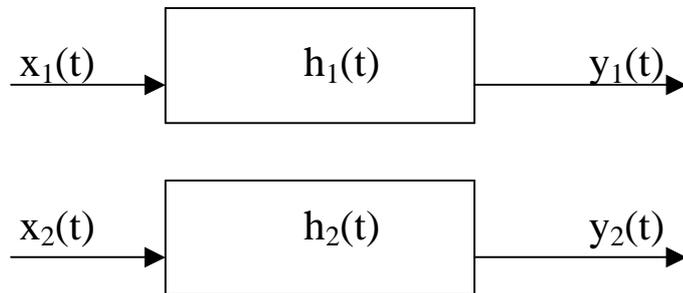
$$\mathbf{m}_{X_1 Y_2}^{(1,1)}(\tau) = E \{ x_1^*(t) y_2(t + \tau) \} = P_{X_1 Y_2}(\tau) = P_{X_1 X_2}(\tau) * h_2(\tau) = \mathbf{m}_{X_1 X_2}^{(1,1)}(\tau) * h_2(\tau) \quad (21)$$

$$\mathbf{m}_{Y_1 X_2}^{(1,1)}(\tau) = E \{ y_1^*(t) x_2(t + \tau) \} = h_1(\tau) \otimes P_{X_1 X_2}(\tau) = h_1(\tau) \otimes \mathbf{m}_{X_1 X_2}^{(1,1)}(\tau) \quad (22)$$

TABELLA 1



- | | | |
|----|--|--|
| 1. | $\overline{y(t)}^t = \overline{x(t)}^t \cdot \int_{-\infty}^{+\infty} h(t) dt$ | $\overline{y(t)}^t = \overline{x(t)}^t \cdot H(0)$ |
| 2. | $P_{YY}(\tau) = P_{XX}(\tau) * E_{hh}(\tau)$ | $P_{YY}(f) = P_{XX}(f) \cdot H(f) ^2$ |
| 3. | $P_{XY}(\tau) = h(\tau) * P_{XX}(\tau)$ | $P_{XY}(f) = P_{XX}(f) \cdot H(f)$ |
| 4. | $P_{YX}(\tau) = h(\tau) \otimes P_{XX}(\tau)$ | $P_{YX}(f) = P_{XX}(f) \cdot H^*(f)$ |



- | | | |
|----|---|---|
| 5. | $P_{Y_1 Y_2}(\tau) = P_{X_1 X_2}(\tau) * E_{h_1 h_2}(\tau)$ | $P_{Y_1 Y_2}(f) = P_{X_1 X_2}(f) \cdot H_1^*(f) \cdot H_2(f)$ |
| 6. | $P_{X_2 Y_2}(\tau) = P_{X_1 X_2}(\tau) * h_2(\tau)$ | $P_{X_1 Y_2}(f) = P_{X_1 X_2}(f) \cdot H_2(f)$ |
| 7. | $P_{Y_1 X_2}(\tau) = h_1(\tau) \otimes P_{X_1 X_2}(\tau)$ | $P_{Y_1 X_2}(f) = P_{X_1 X_2}(f) \cdot H_1^*(f)$ |

PROCESSI GAUSSIANI LIMITATI IN BANDA CON BANDA $2w$ CENTRATA INTORNO A f_0 [$f_0 > W$]

In analogia a quanto effettuato per i segnali certi si definisce:

Def. 1 Processo ergodico limitato in banda, con occupazione di banda $2w$ centrata intorno a f_0 , un processo la cui generica realizzazione non venga alterata nel transito attraverso un filtro passa banda ideale con funzione di trasferimento

$$H(f) = \begin{cases} 1 & \text{per } f_0 - w < |f| < f_0 + w \\ 0 & \text{altrove} \end{cases}$$

Ciò posto consideriamo un processo $X(t)$, gaussiano, ergodico, a valor atteso nullo, limitato in banda con occupazione di banda $2w$ centrata intorno a f_0 . In questo caso la singola realizzazione $x(t)$ può essere espressa in termini di componenti analogiche di bassa frequenza rispetto a f_0 , come segue:

$$x(t) = x_c(t) \cos(2\pi f_0 t) - x_s(t) \sin(2\pi f_0 t)$$

D'altro canto, indicata con $\hat{X}(t)$ la trasformata di Hilbert di $x(t)$, le componenti analogiche di bassa frequenza sono legate a $x(t)$ e $\hat{X}(t)$ dalla trasformazione lineare

$$\begin{bmatrix} x_c(t) \\ x_s(t) \end{bmatrix} = \begin{bmatrix} \cos(2\pi f_0 t) & \sin(2\pi f_0 t) \\ -\sin(2\pi f_0 t) & \cos(2\pi f_0 t) \end{bmatrix} \begin{bmatrix} x(t) \\ \hat{x}(t) \end{bmatrix}$$

Conseguentemente le proprietà statistiche dei due processi $X_c(t)$ e $X_s(t)$ possono essere dedotte, in modo semplice, da quelle di $X(t)$ e della sua trasformata di Hilbert.

A tale scopo si osservi che essendo $X(t)$, gaussiano, stazionario, ergodico, a valor medio nullo, anche il processo $\hat{X}(t)$ ottenuto come uscita del filtro di Hilbert, quando venga applicato in ingresso $X(t)$ è ancora gaussiano, stazionario ed ergodico ed inoltre poiché la funzione di trasferimento del filtro di Hilbert $H_H(f)$ vale

$$H_H(f) = -j \operatorname{sign}(f)$$

si ha:

$$P_{\hat{X}\hat{X}}(f) = P_{XX}(f) \cdot |H_H(f)|^2 = P_{XX}(f)$$

Ovvero ricordando il Teorema di Wiener – Khintchine ed osservando che il valore atteso è nullo si ha:

$$K_{\hat{X}\hat{X}}(\tau) = F^{-1} \{ P_{\hat{X}\hat{X}}(f) \} = F^{-1} \{ P_{XX}(f) \} = K_{XX}(\tau)$$

Poiché quindi $\hat{X}(t)$ ha la stessa funzione di covarianza di $X(t)$, ne consegue che $\hat{X}(t)$ ha la stessa gerarchia di probabilità di $X(t)$.

Si noti inoltre che $x(t)$ e $\hat{X}(t)$ risultano ortogonali tra loro, infatti applicando le regole del transito di un segnale attraverso un filtro si ha:

$$P_{X\hat{X}}(\tau) = h_H(\tau) \otimes P_{XX}(\tau) = \frac{1}{\pi\tau} * P_{XX}(\tau) = \hat{P}_{XX}(\tau)$$

Poiché $P_{XX}(\tau)$ è una funzione pari e $1/\pi\tau$ è una funzione dispari si ha:

$$P_{x\hat{x}}(0) = \lim_{\Delta T \rightarrow \infty} \frac{1}{\Delta T} \int_{-\Delta T/2}^{\Delta T/2} x(t) \hat{x}(t) dt = \int_{-\infty}^{+\infty} P_{XX}(\tau) \frac{1}{\pi \tau} d\tau = 0$$

Conseguentemente essendo, $m_X = m_{\hat{X}} = 0$, per l'ergodicit  si ha:

$$K_{X\hat{X}}(0) = P_{X\hat{X}}(0) - m_X m_{\hat{X}} = 0$$

e quindi le determinazioni $x(t_1)$ e $\hat{x}(t_1)$ estratte da $X(t)$ e $\hat{X}(t)$ nello stesso istante di tempo sono statisticamente indipendenti.

Ci  posto, poich  $x_c(t)$ e $x_s(t)$ sono ottenute da $x(t)$ e $\hat{x}(t)$ tramite una trasformazione lineare, anche le v. a. X_c e X_s estratte da $X_c(t)$ e $X_s(t)$ nell'istante t_1 saranno ancora congiuntamente gaussiane, incorrelate, a valor atteso nullo e varianza $\sigma_{X_c}^2 = \sigma_{X_s}^2 = \sigma_X^2$

Infatti:

$$\begin{bmatrix} m_{X_c} \\ m_{X_s} \end{bmatrix} = \begin{bmatrix} \cos(2\pi f_0 t) & \sin(2\pi f_0 t) \\ -\sin(2\pi f_0 t) & \cos(2\pi f_0 t) \end{bmatrix} \cdot \begin{bmatrix} m_X \\ m_{\hat{X}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \sigma_{X_c}^2 & \sigma_{X_c X_s} \\ \sigma_{X_c X_s} & \sigma_{X_s}^2 \end{bmatrix} = \begin{bmatrix} \cos(2\pi f_0 t) & \sin(2\pi f_0 t) \\ -\sin(2\pi f_0 t) & \cos(2\pi f_0 t) \end{bmatrix} \cdot \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_X^2 \end{bmatrix} \cdot \begin{bmatrix} \cos(2\pi f_0 t) & -\sin(2\pi f_0 t) \\ \sin(2\pi f_0 t) & \cos(2\pi f_0 t) \end{bmatrix} = \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_X^2 \end{bmatrix}$$

c. d. d.

Pi  in generale allo stesso modo si pu  dimostrare che $X_c(t)$ e $X_s(t)$ sono due processi congiuntamente gaussiani con valor atteso nullo e funzioni di covarianza:

$$K_{X_c X_c}(\tau) = K_{X_s X_s}(\tau) = K_X^c(\tau)$$

$$K_{X_c X_s}(\tau) = K_{X_s X_c}(\tau) = K_X^s(\tau)$$

In cui $K_X^c(\tau)$ e $K_X^s(\tau)$ sono le componenti analogiche di bassa frequenza della funzione di covarianza di $K_X(\tau)$ rispetto a f_0 :

$$K_X(\tau) = K_X^c(\tau) \cos(2\pi f_0 \tau) - K_X^s(\tau) \sin(2\pi f_0 \tau)$$

INVILUPPO E FASE DI UN PROCESSO GAUSSIANO

Dato un processo gaussiano $X(t; \omega)$, ergodico, a valor medio nullo e limitato in banda, con banda $2w$ centrato intorno a f_0 , si definisce inviluppo e fase di $X(t; \omega)$ rispetto a f_0 il modulo e la fase dell'inviluppo complesso, ovvero

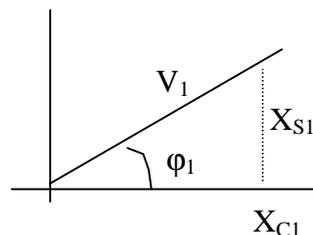
$$\begin{cases} V(t) = \sqrt{x_c^2(t) + x_s^2(t)} \\ \Phi(t) = \text{tag}^{-1} \frac{x_s(t)}{x_c(t)} \end{cases}$$

Pertanto per la singola realizzazione $x(t)$ si ha la seguente rappresentazione:

$$x(t) = v(t) \cos[2\pi f_0 t + \varphi(t)]$$

Per il calcolo della gerarchia di ordine 1 di $v(t)$ e di $\varphi(t)$ si osservi che

$$\begin{cases} x_c(t) = v(t) \cos[\varphi(t)] \\ x_s(t) = v(t) \sin[\varphi(t)] \end{cases}$$



E quindi poiché la gerarchia di ordine 1 di $x_c(t)$ e $x_s(t)$ vale:

$$P_{X_{c_1} X_{s_1}}(x_{c_1}, x_{s_1}) = \frac{1}{2\pi\sigma_x^2} e^{-\frac{x_{c_1}^2 + x_{s_1}^2}{2\sigma_x^2}}$$

e lo Jacobiano della trasformazione, che equivale ad un cambiamento di variabile da cartesiano a polare, vale:

$$J(v_1, \varphi_1) = \begin{vmatrix} \cos \varphi_1 & \sin \varphi_1 \\ -v_1 \sin \varphi_1 & v_1 \cos \varphi_1 \end{vmatrix} = v_1$$

in base alla regola del cambiamento di variabile si ha:

$$P_{V_1 \Phi_1}(v_1, \varphi_1) = \frac{v_1}{2\pi\sigma_x^2} e^{-\frac{v_1^2}{2\sigma_x^2}}, \quad v_1 > 0, \varphi_1 \in [0, 2\pi)$$

La gerarchia di ordine 1 dell'involuppo può essere calcolata dalla precedente relazione per saturazione, ottenendo:

$$P_{V_1}(v_1) = \int_0^{2\pi} \frac{v_1}{2\pi\sigma_x^2} e^{-\frac{v_1^2}{2\sigma_x^2}} d\varphi_1 = \frac{v_1}{\sigma_x^2} e^{-\frac{v_1^2}{2\sigma_x^2}}, \quad v_1 \geq 0$$

Tale f. d. p. è nota come f. d. p. di Rayleigh. La funzione di distribuzione di V_1 si ottiene immediatamente per integrazione della f. d. p. :

$$D_{V_1}(v_1) = \int_0^{v_1} \frac{r}{\sigma_x^2} e^{-\frac{r^2}{2\sigma_x^2}} dr = 1 - e^{-\frac{v_1^2}{2\sigma_x^2}}$$

Inoltre per il valore atteso, per il valore quadratico atteso e la varianza dell'involuppo si ha:

$$m_V = E\{v\} = \int_0^{\infty} \frac{v}{\sigma_x^2} e^{-\frac{v^2}{2\sigma_x^2}} dv = \int_0^{\infty} v d\left(e^{-\frac{v^2}{2\sigma_x^2}}\right) = \left[-v e^{-\frac{v^2}{2\sigma_x^2}} \right]_0^{\infty} + \int_0^{\infty} e^{-\frac{v^2}{2\sigma_x^2}} dv = \sigma_x \sqrt{\frac{\pi}{2}}$$

$$m_V^{(2)} = E_V \{v^2\} = E_{X_c, X_s} \{x_c^2 + x_s^2\} = 2\sigma_x^2$$

$$\sigma_V^2 = \left(2 - \frac{\pi}{2}\right) \sigma_x^2$$

Si noti che nel derivare il valore quadratico atteso, si è fatto uso dell'incorrelazione tra $x_c(t)$ e $x_s(t)$, allo stesso istante di tempo.

Per ciò che concerne la fase $\varphi(t)$, la sua gerarchia di ordine 1 può essere ottenuta dalla f. d. p. congiunta $P_{v_1, \varphi_1}(v_1, \varphi_1)$ per saturazione rispetto a v_1 , ottenendo

$$P_{\varphi_1}(\varphi_1) = \int_0^{\infty} \frac{v_1}{2\pi\sigma_x^2} e^{-\frac{v_1^2}{2\sigma_x^2}} dv_1 = \frac{1}{2\pi} \int_0^{\infty} d\left(e^{-\frac{v_1^2}{2\sigma_x^2}}\right) = \frac{1}{2\pi}$$

Cioè la fase $\varphi(t)$ è una v. a. a distribuzione uniforme nell'intervallo $[0, 2\pi)$. Si noti inoltre che, poiché:

$$P_{v_1, \varphi_1}(v_1, \varphi_1) = P_{v_1}(v_1) P_{\varphi_1}(\varphi_1)$$

l'involuppo e la fase del processo $X(t; \omega)$, considerati allo stesso istante di tempo, sono statisticamente indipendenti. Infine, in talune applicazioni può essere utile considerare il quadrato $B(t) = V^2(t)$ dell'involuppo. Per tale processo si ha:

$$P_{B_1}(b_1) = \begin{cases} \frac{1}{2\sigma_X^2} e^{-\frac{|b|}{2\sigma_X^2}} & b \geq 0 \\ 0 & b < 0 \end{cases}$$

$$m_V = E\{B(t)\} = E\{V^2(t)\} = 2\sigma_X^2$$

$$\begin{aligned} m_B^{(2)} = E\{V^4(t)\} &= E\{[x_c^2(t) + x_s^2(t)]^2\} = E\{x_c^4(t) + 2x_c^2(t)x_s^2(t) + x_s^4(t)\} = \\ &= 3\sigma_X^4 + 2\sigma_X^2\sigma_X^2 + 3\sigma_X^4 = 8\sigma_X^4 \end{aligned}$$

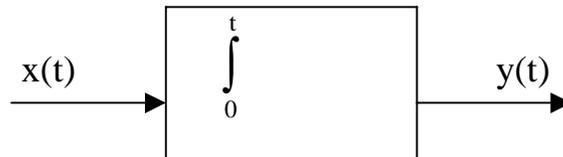
$$\sigma_B^2 = m_B^{(2)} - m_B^2 = 4\sigma_X^4$$

Infine per la funzione di autocorrelazione di tale processo si ha:

$$\begin{aligned} P_{BB}(\tau) = m_{BB}^{(1,1)}(\tau) &= E\{[x_c^2(t) + x_s^2(t)][x_c^2(t+\tau) + x_s^2(t+\tau)]\} = \\ &= E\{x_c^2(t)x_c^2(t+\tau)\} + E\{x_c^2(t)x_s^2(t+\tau)\} + E\{x_s^2(t)x_c^2(t+\tau)\} + \\ &+ E\{x_s^2(t)x_s^2(t+\tau)\} = \sigma_X^2 + 2[E\{x_c(t)x_c(t+\tau)\}]^2 + \sigma_X^2 + 2[E\{x_c(t)x_s(t+\tau)\}]^2 + \\ &+ \sigma_X^2 + 2[E\{x_s(t)x_c(t+\tau)\}]^2 + \sigma_X^2 + 2[E\{x_s(t)x_s(t+\tau)\}]^2 = \\ &= 4\sigma_X^2 + 4[K_X^c(\tau)]^2 + 4[K_X^s(\tau)]^2 \end{aligned}$$

PROCESSO DI WIENER

Nel definire il processo di Wiener appare utile partire dall'analisi di un possibile meccanismo di generazione di tale processo. Pertanto, con riferimento allo schema di fig.



sia $X(t;\omega)$ un processo gaussiano, ergodico, a valor atteso nullo e spettro di densità di potenza uniforme in $(-\infty, +\infty)$:

$$P_{XX}(f) = N_0$$

conseguentemente, la funzione di covarianza di $X(t;\omega)$, per il teorema di Wiener – Khintchine, varrà:

$$K_{XX}(\tau) = N_0 u_0(\tau)$$

Sia inoltre $y(t)$ l'uscita dell'integratore quando in ingresso viene applicata una generica realizzazione $x(t)$ di $X(t; \omega)$:

$$y(t) = \int_0^t x(\vartheta) d\vartheta$$

Si noti che, poiché l'integrale viene attivato dall'istante $t=0$, l'uscita $y(t)$ sarà identicamente nulla per $t < 0$, quindi il processo d'uscita non potrà essere stazionario.

Poiché la trasformata, anche se variante nel tempo, è lineare il processo di uscita $Y(t; \omega)$ risulta ancora gaussiano.

Il valore atteso $m_Y(t_1)$ può essere calcolato come segue:

$$m_Y(t_1) = E_Y \{y(t_1)\} = E_X \left\{ \int_0^{t_1} x(\vartheta) d\vartheta \right\} = \int_0^{t_1} E\{x(\vartheta)\} d\vartheta = \int_0^{t_1} m_X d\vartheta = m_X \int_0^{t_1} d\vartheta = 0$$

Allo stesso modo per il valore quadratico atteso $m_Y^{(2)}(t_1)$ si ha:

$$m_Y^{(2)}(t_1) = E_Y \{[y(t_1)]^2\} = E_X \left\{ \int_0^{t_1} x(\vartheta_1) d\vartheta_1 \cdot \int_0^{t_1} x(\vartheta_2) d\vartheta_2 \right\} = \int_0^{t_1} \int_0^{t_1} E\{x(\vartheta_1)x(\vartheta_2)\} d\vartheta_1 d\vartheta_2 = \int_0^{t_1} \int_0^{t_1} N_0 u_0(\vartheta_1 - \vartheta_2) d\vartheta_1 d\vartheta_2 = N_0 \int_0^{t_1} d\vartheta = N_0 t_1$$

Quindi il processo $Y(t; \omega)$ è ancora a valor atteso nullo, ma il suo valore quadratico atteso cresce linearmente nel tempo, cosa che conferma la non-stazionarietà di $Y(t; \omega)$.

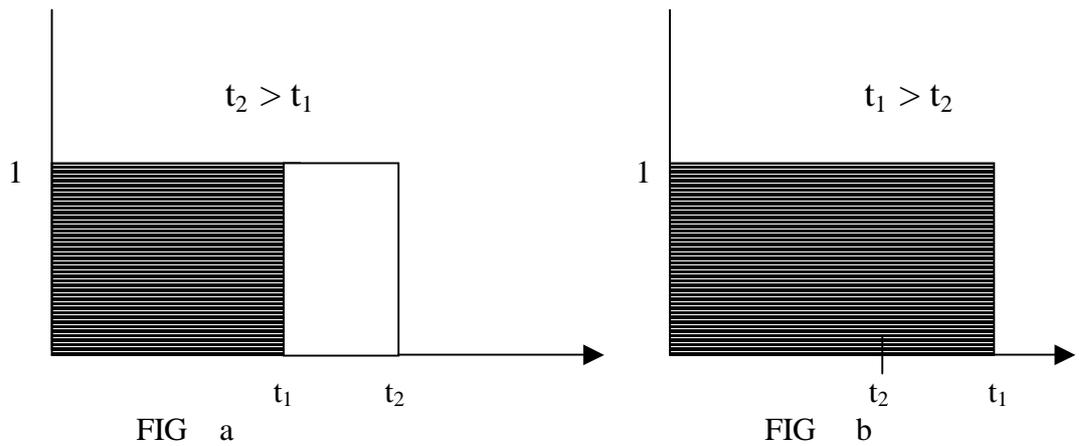
Essendo $Y(t; \omega)$ gaussiano, a valor atteso nullo, per poterne scrivere la generica gerarchia di ordine n , e quindi per poterlo descrivere in modo completo dal punto di vista statistico, occorre calcolare la funzione di covarianza $K_{YY}(t_1, t_2)$.

L'espressione di tale funzione può essere dedotta seguendo il processo già delineato nel calcolo del valore quadratico atteso. In particolare, osservando che il valore atteso di $Y(t; \omega)$ è nullo si ha:

$$\begin{aligned}
 K_{YY}(t_1, t_2) &= E_Y \{ [y(t_1) - m_Y(t_1)] [y(t_2) - m_Y(t_2)] \} = \\
 &= E_X \left\{ \int_0^{t_1} x(\vartheta_1) d\vartheta_1 \cdot \int_0^{t_2} x(\vartheta_2) d\vartheta_2 \right\} = \int_0^{t_1} \int_0^{t_2} E \{ x(\vartheta_1) x(\vartheta_2) \} d\vartheta_1 d\vartheta_2 = \\
 &\quad \int_0^{t_1} \int_0^{t_2} N_0 u_0(\vartheta_1 - \vartheta_2) d\vartheta_1 d\vartheta_2
 \end{aligned}$$

Per il calcolo dell'integrale doppio che compare nella precedente relazione si osservi che in virtù delle proprietà dell'impulso matematico si ha:

$$\int_0^{t_2} u_0(\vartheta_1 - \vartheta_2) d\vartheta_2 = \begin{cases} 1 & \text{se } \vartheta_1 \leq t_2 \\ 0 & \text{se } \vartheta_1 > t_2 \end{cases}$$



e quindi, come appare evidente dalle figg. a e b, l'integrale rispetto a ϑ_1 di tale funzione sull'intervallo $[0, t_1]$ vale t_1 se $t_1 < t_2$ e t_2 se $t_1 > t_2$. Di conseguenza:

$$K_{YY}(t_1, t_2) = N_0 \min(t_1, t_2)$$

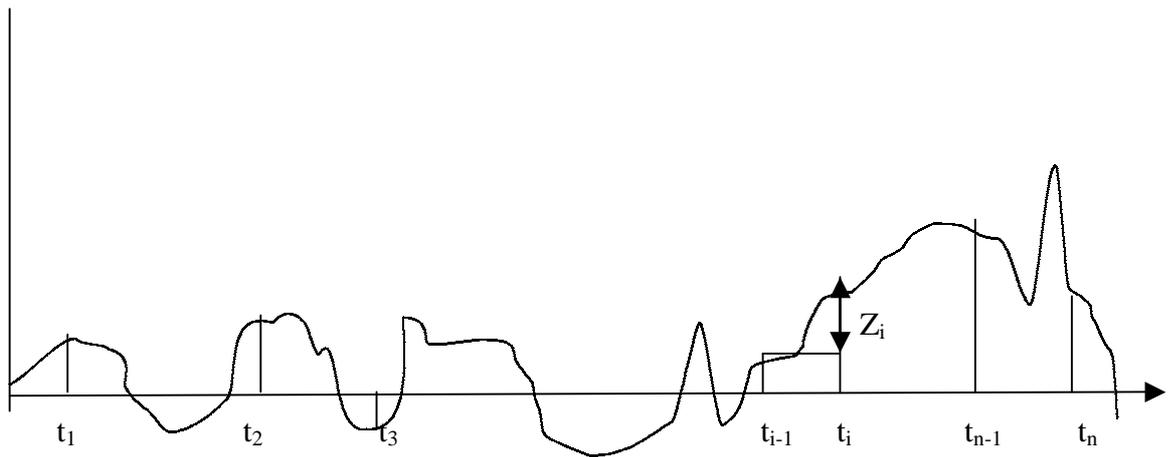
Ovviamente il valore quadratico atteso può anche essere calcolato direttamente da tale relazione ponendo $t_1 = t_2$.

Ciò posto può essere data la seguente definizione formale di processo di Wiener

Def. Si definisce processo di Wiener, un processo gaussiano, definito sull'intervallo $T = [0, \infty)$,

- a valor atteso nullo
- funzione di covarianza del tipo

$$K_{YY}(t_1, t_2) = N_0 \min(t_1, t_2)$$



Il processo di Wiener gode dell'importante proprietà che comunque si fissino $n + 1$ istanti di tempo

$$0 \leq t_0 < t_1 < t_2 < \dots < t_{i-1} < t_i < \dots < t_n$$

gli incrementi Z_i :

$$Z_i = Y(t_i) - Y(t_{i-1}) , \quad i=1, \dots, n$$

che il processo $Y(t; \omega)$ subisce negli n intervalli di tempo $[t_{i-1}, t_i]$ sono congiuntamente gaussiani, a valor atteso nullo, varianza $\sigma_{Z_i}^2$ pari a

$$\sigma_{Z_i}^2 = N_0 (t_i - t_{i-1}) = N_0 \Delta t_i \quad \text{con } \Delta t_i = t_i - t_{i-1}$$

e tra loro staticamente indipendenti.

Per dimostrare tali proprietà è sufficiente osservare che gli incrementi Z_i possono essere in termini del processo $X(t; \omega)$ come segue

$$Z_i = \int_{t_{i-1}}^{t_i} x(\vartheta) d\vartheta$$

e procedere come per il calcolo del valore atteso e della funzione di covarianza del processo di Wiener, In particolare si ha;

$$E_{Z_i} \{z_i\} = E_X \left\{ \int_{t_{i-1}}^{t_i} x(\vartheta) d\vartheta \right\} = \int_{t_{i-1}}^{t_i} E_X \{x(\vartheta)\} d\vartheta = 0$$

$$\begin{aligned} \sigma_{Z_i}^2 &= E_{Z_i} \{z_i - m_{Z_i}\} = E_X \left\{ \int_{t_{i-1}}^{t_i} x(\vartheta_1) d\vartheta_1 \int_{t_{i-1}}^{t_i} x(\vartheta_2) d\vartheta_2 \right\} = \\ &= \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} E_X \{x(\vartheta_1) x(\vartheta_2)\} d\vartheta_1 d\vartheta_2 = \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} N_0 u_0(\vartheta_1 - \vartheta_2) d\vartheta_1 d\vartheta_2 = N_0 \Delta t_i \end{aligned}$$

$$\begin{aligned} E_{Z_i Z_j} \{z_i, z_j\} &= E_X \left\{ \int_{t_i}^{t_{i-1}} x(\vartheta_1) d\vartheta_1 \int_{t_j}^{t_{j-1}} x(\vartheta_2) d\vartheta_2 \right\} = \\ &= \int_{t_i}^{t_{i-1}} \int_{t_j}^{t_{j-1}} E_X \{x(\vartheta_1) x(\vartheta_2)\} d\vartheta_1 d\vartheta_2 = N_0 \int_{t_i}^{t_{i-1}} \int_{t_j}^{t_{j-1}} u_0(\vartheta_1 - \vartheta_2) d\vartheta_1 d\vartheta_2 = \\ &= N_0 \int_{t_i}^{t_{i-1}} \text{rect}_{\Delta t_j} \left(\vartheta_1 - \frac{t_j - t_{j-1}}{2} \right) d\vartheta_1 = 0 \end{aligned}$$

Si noti che nell'ultimo integrale, l'integrando è diverso da zero solo nell'intervallo $[t_{j-1}, t_j]$ che non ha parti in comune con l'intervallo d'integrazione.

Poiché gli incrementi Z_i sono incorrelati e gaussiani, essi risultano anche statisticamente indipendenti.

La f. d. p. $p_{\underline{z}}(\underline{z})$ relativa agli incrementi vale:

$$p_{\underline{z}}(\underline{z}) = \frac{1}{\prod_{i=1}^n (2\pi\Delta t_i)^{1/2}} e^{-\sum_{i=1}^n \frac{z_i^2}{2N_0\Delta t_i}}$$

Utilizzando tale f. d. p. è agevole dimostrare che il processo di Wiener è un processo di Markov semplice. Infatti osservando che, preso $t_0 = 0$ ed essendo $Y(0) = 0$, si ha

$$\begin{bmatrix} z_1 \\ z_2 \\ \dots \\ z_n \end{bmatrix} = \begin{bmatrix} 1 & \dots & \dots & 0 \\ -1 & 1 & \dots & \dots \\ \dots & -1 & 1 & \dots \\ 0 & \dots & -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}$$

e poiché la matrice A è triangolare, il determinante di tale matrice è pari al prodotto degli elementi della diagonale principale, e quindi il determinante Jacobiano della trasformazione è pari a 1, per la nota regola del cambiamento di variabile si ha:

$$P_{Y_1, \dots, Y_n}(y_1, y_2, \dots, y_n; t_1, t_2, \dots, t_n) = \frac{1}{\prod_{i=1}^n (2\pi N_0 \Delta t_i)^{1/2}} e^{-\sum_{i=1}^n \frac{(y_i - y_{i-1})^2}{2N_0 \Delta t_i}}$$

Tale relazione fornisce la gerarchia di ordine n del processo di Wiener.

Per dimostrare che il processo di Wiener è un processo di Markov semplice si osservi che in base alla relazione precedente, si ha:

$$\begin{aligned}
 P_{Y_n / Y_{n-1} \dots Y_1} (y_n / y_{n-1}, \dots, y_1; t_1, t_2, \dots, t_n) &= \frac{P_{Y_1 \dots Y_n} (y_1, \dots, y_n; t_1, \dots, t_n)}{P_{Y_1 \dots Y_{n-1}} (y_1, \dots, y_{n-1}; t_1, \dots, t_{n-1})} = \\
 &= \frac{\prod_{i=1}^{n-1} (2\pi N_0 \Delta t_i)^{1/2} e^{-\sum_{i=1}^{n-1} \frac{(y_i - y_{i-1})^2}{2N_0 \Delta t_i}}}{\prod_{i=1}^n (2\pi N_0 \Delta t_i)^{1/2} e^{-\sum_{i=1}^{n-1} \frac{(y_i - y_{i-1})^2}{2N_0 \Delta t_i}}} = \frac{e^{-\frac{(y_n - y_{n-1})^2}{2N_0 \Delta t_n}}}{\sqrt{2\pi N_0 \Delta t_n}}
 \end{aligned}$$

e quindi, poiché la f. d. p. condizionata dipende solo da Y_{n-1} e come è facile verificare, coincide con la $P_{Y_n / Y_{n-1} \dots Y_1} (y_n / y_{n-1}, \dots, y_1; t_1, t_2, \dots, t_n)$, il processo è di Markov semplice.

In particolare si osservi che condizionatamente a Y_{n-1} , Y_n è una v. a. con valor atteso pari a Y_{n-1} e varianza pari a $N_0 \Delta t_n$.

Ciò risulta evidente se si osserva che Y_n è pari a :

$$Y_n = Y_{n-1} + Z_n$$

La proprietà di Markovianità può essere spiegata dal punto di vista concettuale, considerando che una volta noto il valore dell'integratore nell'istante t_{i-1} , il valore assunto dall'uscita dell'integratore all'istante t_i dipenderà solo dalla storia dell'ingresso nell'intervallo $[t_{i-1}, t_i]$, poiché per i valori assunti dall'ingresso in tale intervallo non dipendono statisticamente da quelli assunti nell'intervallo $[0, t_{j-1}]$, tutta l'informazione statistica relativa all'intervallo $[0, t_{j-1}]$ è assunta dallo stato dell'integratore all'istante t_{j-1} .

SERIE DI MARKOV

Presa una serie aleatoria $X(t; \omega)$ essa si dice discreta se le variabili aleatorie $X(t_1), X(t_2), \dots, X(t_n)$ estratte negli n istanti t_1, t_2, \dots, t_n sono variabili aleatorie discrete che possono assumere una fra L possibili determinazioni appartenenti all'insieme A :

$$A = \{x_1, x_2, \dots, x_L\}$$

con L eventualmente infinito. L'insieme A prende il nome di Alfabeto di sorgente.

Analogamente la serie aleatoria si dice continua se le predette variabili aleatorie risultano essere v. a. continue.

Ciò posto, è possibile formulare la seguente definizione di sorgente, ovvero di serie di Markov discreta.

Def. Una serie aleatoria discreta (e la sorgente che la emette) si dice di Markov di ordine M se detti

$$t_{k_1}, t_{k_2}, \dots, t_{k_N}$$

N istanti di tempo in ordine crescente con $N > M+1$, e indicate con

$$x[k_1]=x(t_{k_1}), x[k_2]=x(t_{k_2}), \dots, x[k_n]=x(t_{k_n})$$

le generiche determinazioni delle v. a. estratte in detti istanti, le probabilità delle determinazioni della v. a. $X[k_n]$ condizionate alle determinazioni $x[k_1], x[k_2], \dots, x[k_{n-1}]$ dipendono dalle determinazioni delle M variabili aleatorie che precedono $X[k_n]$. In formule

$$\begin{aligned} & \Pr \{X[k_n]=x_{i_{k_n}} / X[k_{n-1}]=x_{i_{k_{n-1}}}, \dots, X[k_1]=x_{i_{k_1}}\} = \\ & = \Pr \left\{ X[k_n]=x_{i_{k_n}} / X[k_{n-1}]=x_{i_{k_{n-1}}}, \dots, X[k_{n-M}]=x_{i_{k_{n-M}}} \right\} \quad (1) \end{aligned}$$

Per $M = 1$ la serie viene detta di Markov semplice.

Si osservi che come per i processi di Markov, anche per le serie di Markov la gerarchia di qualsiasi ordine è ottenibile a partire dalla conoscenza della gerarchia di probabilità di ordine $M+1$.

Infatti in base alla (1) per $n > M+1$ si ha:

$$\begin{aligned} & \Pr \{X[n]=x_{i_n} / X[n-1]=x_{i_{n-1}}, \dots, X[1]=x_{i_1}\} = \\ & = \Pr \{X[n]=x_{i_n} / X[n-1]=x_{i_{n-1}}, \dots, X[n-M]=x_{i_{n-M}}\} \\ & \Pr \{X[n-1]=x_{i_{n-1}} / X[n-2]=x_{i_{n-2}}, \dots, X[n-1-M]=x_{i_{n-1-M}}\} \cdot \\ & \dots \cdot \Pr \{X[M+1]=x_{i_{M+1}} / X[M]=x_{i_M}, \dots, X[1]=x_{i_1}\} \\ & \cdot \Pr \{X[M]=x_{i_M}, \dots, X[1]=x_{i_1}\} \quad (2) \end{aligned}$$

ed inoltre:

$$\begin{aligned} & \Pr \{X[n]=x_{i_n} / X[n-1]=x_{i_{n-1}}, \dots, X[n-M]=x_{i_{n-M}}\} = \\ & = \frac{\Pr \{X[n]=x_{i_n}, X[n-1]=x_{i_{n-1}}, \dots, X[n-M]=x_{i_{n-M}}\}}{\Pr \{X[n-1]=x_{i_{n-1}}, \dots, X[n-M]=x_{i_{n-M}}\}} \quad (3) \end{aligned}$$

Ovviamente le gerarchie di ordine inferiore a $M+1$ possono essere ottenute per saturazione della gerarchia di ordine $M+1$.

Anche nel caso delle serie di Markov, la gerarchia di probabilità di ordine $M+1$ dovrà soddisfare la relazione di consistenza rappresentata dalla relazione di

CHAPMAN – KOLMOGOROFF. Per derivare le implicazioni di tale relazione, si cominci con il considerare una serie di Markov semplice.

A tale scopo si osservi che nel caso di serie discreta, la gerarchia di probabilità del 1° ordine può essere riscritta in forma più compatta introducendo il vettore $P[k_n]$ i cui elementi sono costituiti dalle probabilità che la v. a. $X[k_n]$ assuma le singole determinazioni x_1, \dots, x_L . In formule:

$$P[k_n] = \begin{bmatrix} \Pr\{X[k_n]=x_1\} \\ \Pr\{X[k_n]=x_2\} \\ \dots \\ \Pr\{X[k_n]=x_L\} \end{bmatrix} \quad (4)$$

D'altronde, considerando due istanti di tempo t_{k_n} e t_{k_1} , per l'elemento i -esimo di tale vettore si ha

$$\Pr\{X[k_n]=x_i\} = \sum_{j=1}^L \Pr\{X[k_n]=x_i, X[k_1]=x_j\} \quad i = 1, \dots, L \quad (5)$$

ovvero, esprimendo la probabilità congiunta tramite le probabilità condizionate si ha:

$$\Pr\{X[k_n]=x_i\} = \sum_{j=1}^L \Pr\{X[k_n]=x_i/X[k_1]=x_j\} \cdot \Pr\{X[k_1]=x_j\} \quad i = 1, \dots, L \quad (6)$$

Tale Relazione che vale per il singolo elemento del vettore può essere riscritta in modo più compatto, in forma matriciale, come segue:

$$P[k_n] = \prod [k_n, k_1] \cdot P[k_1] \quad (7)$$

in cui $\prod [k_n, k_1]$ è la matrice $L \times L$ il cui generico elemento (i, j) è costituito dalla probabilità condizionata $\Pr\{X[k_n]=x_i/X[k_1]=x_j\}$:

$$\prod [k_n, k_1] = \begin{bmatrix} \Pr\{X[k_n]=x_1/X[k_1]=x_1\} & \dots & \Pr\{X[k_n]=x_1/X[k_1]=x_L\} \\ \Pr\{X[k_n]=x_2/X[k_1]=x_1\} & \dots & \Pr\{X[k_n]=x_2/X[k_1]=x_L\} \\ \dots & \dots & \dots \\ \Pr\{X[k_n]=x_L/X[k_1]=x_j\} & \dots & \Pr\{X[k_n]=x_L/X[k_1]=x_L\} \end{bmatrix} \quad (8)$$

Si noti che la somma degli elementi di ciascuna colonna della matrice $\prod [k_n, k_1]$ è pari a 1, in quanto la funzione di normalizzazione

$$\sum_{i=1}^L \Pr\{X[k_n]=x_i/X[k_1]=x_j\} = 1 \quad (9)$$

Si noti che la reazione (7) vale per qualsiasi serie discreta.

Nel caso in cui la serie discreta sia di Markov semplice, la relazione di CHAPMAN – KOLMOGOROFF può essere scritta in forma matriciale come fattorizzazione della matrice $\prod [k_n, k_1]$. Ovvero presi 3 istanti $t_{k_1} < t_{k_m} < t_{k_n}$ si ha:

$$\prod [k_n, k_1] = \prod [k_n, k_m] \cdot \prod [k_m, k_1] \quad (10)$$

Per la dimostrazione della (10) si osservi che per ciascun elemento della matrice $\prod [k_n, k_1]$ si ha:

$$\begin{aligned} \Pr\{X[k_n]=x_i/X[k_1]=x_j\} &= \sum_{h=1}^L \Pr\{X[k_n]=x_i, X[k_m]=x_h/X[k_1]=x_j\} \\ &= \sum_{h=1}^L \Pr\{X[k_n]=x_i/X[k_m]=x_h, X[k_1]=x_j\} \cdot \Pr\{X[k_m]=x_h/X[k_1]=x_j\} \end{aligned} \quad (11)$$

ed essendo la serie di Markov semplice ne consegue che:

$$\Pr\{X[k_n]=x_i/X[k_1]=x_j\} = \sum_{h=1}^L \Pr\{X[k_n]=x_i/X[k_m]=x_h\} \cdot \Pr\{X[k_m]=x_h/X[k_1]=x_j\} \quad (10')$$

Come è facile verificare la (10'), altro non è che la (10) riscritta elemento per elemento.

Ciò posto, applicando iterativamente la (10) nella (7) e ponendo per semplicità di notazione

$$\prod_1 [k] \stackrel{\Delta}{=} \prod [k+1, k] \quad (12)$$

si ha:

$$P[k_n] = \prod_1 [k_{n-1}] \prod_1 [k_{n-2}] \dots \prod_1 [k_1] P[k_1] \quad (13)$$

che fornisce la funzione di probabilità all'istante t_{k_n} a partire dalla conoscenza della funzione di probabilità all'istante t_{k_1} e delle matrici delle probabilità condizionate. Di conseguenza si può affermare che una serie di Markov semplice è completamente descritta sul piano statistico non appena si conoscono le matrici delle probabilità condizionate a passo 1, e la funzione di probabilità relativa all'istante iniziale $P[0] = P_0$ che rappresenta la condizione iniziale dell'equazione alle differenze:

$$\boxed{P[k+1] = \prod_1 [k] \cdot P[k]} \quad (14)$$

soddisfatto dalla gerarchia di ordine 1.

Nel caso in cui la serie di Markov semplice sia omogenea le matrici delle probabilità condizionate non dipenderanno più dai 2 indici relativi agli istanti di tempo, ma solo dalla loro differenza. Ciò implica che la matrice $\Pi_1[k]$ diviene costante rispetto al tempo e le relazioni (14) e (13) si semplificano come segue:

$$P[k+1] = \prod_1 \cdot P[k] \quad (15)$$

$$P[k+1] = \left[\prod_1 \right]^n \cdot P[k] \quad (16)$$

Nel caso in cui la serie sia stazionaria oltre alle (15) e (16) si ha:

$$P[k+1] = P[k] = P = \text{cost.} \quad (17)$$

Dalle (15) e (17) ne consegue che la gerarchia di ordine 1 deve soddisfare la relazione:

$$(I - \Pi_1) P = 0 \quad (18)$$

Si osservi che poiché la somma della matrice degli elementi di ciascuna colonna della matrice Π_1 è pari a 1, e che quindi la somma degli elementi di ciascuna colonna della matrice $(I - \Pi_1)$ è zero, il determinante della matrice $I - \Pi_1$ è nullo. Quindi il sistema omogeneo (18) ammette almeno ∞^1 soluzioni. Affinché però il vettore P rappresenti una funzione di probabilità deve essere soddisfatta la condizione di normalizzazione :

$$\sum_{i=1}^L P_i = 1 \quad (19)$$

In effetti, si potrebbe dimostrare che qualora la serie sia anche ergodica, il rango della matrice $(I - \Pi_1)$ è proprio pari a $L-1$ e quindi il sistema omogeneo (18) ammette ∞^1 soluzioni, ovvero la funzione di probabilità individuata dalle (18) e (19) è unica.

In definitiva, se la serie è ergodica, la matrice delle probabilità, condizionate a passo 1, descrive completamente il comportamento statistico della serie, in quando da essa può essere calcolata la gerarchia di qualsiasi ordine.

Quanto precede può essere esteso agevolmente al caso di sorgente di Markov di ordine M attraverso l'introduzione del concetto di stato.

Per definizione lo stato della sorgente $S[k]$ all'istante t_k è rappresentato dalle determinazioni degli M termini precedenti:

$$S[k] \triangleq (x[k-1], x[k-2], \dots, x[k-M])$$

Pertanto il numero di stati possibili è pari a L^M , ovvero pari al numero di combinazioni di M termini consecutivi delle L possibili determinazioni. Lo stato della serie di Markov potrà assumere valori sull'insieme

$$S = \{s_1, s_2, \dots, s_{L^M}\}$$

i cui elementi sono i punti di uno spazio prodotto $L \times L \times L \times \dots \times L$ (M volte).

Da ciò ne consegue che una volta noto lo stato all'istante t_{k+1} , risulta nota anche la determinazione assunta dalla serie nell'istante t_k . Inoltre, per come è definito una volta note le proprietà statistiche dello stato $S[k+1]$ da esse possono essere dedotte (tramite operazioni di somma) le proprietà statistiche delle v. a. $X[k]$, \dots , $X[k-M+1]$.

L'impiego dello stato, nello studio di una sorgente di Markov è motivato dal seguente teorema:

Teorema Data una serie di Markov discreta di ordine M , la serie (discreta) degli stati è una serie di Markov semplice

Dim. Dalla definizione di stato della serie di Markov ne consegue che:

$$\begin{aligned}
 & \Pr \{S[k]=s_k/S[k-1]=s_{k-1}, \dots, S[1]=s_1\} = \\
 = & \Pr \{X[k-1]=x_{k-1}, X[k-2]=x_{k-2}, \dots, X[k-M]=x_{k-M}/X[k-2]=x_{k-2}, \dots, X[1]=x_1\} = \\
 & = \Pr \{X[k-1]=x_{k-1}/X[k-2]=x_{k-2}, \dots, X[1]=x_1\} \cdot \\
 & \Pr \{X[k-2]=x_{k-2}, \dots, X[k-M]=x_{k-M}/X[k-2]=x_{k-2}, \dots, X[1]=x_1\} = \\
 & = \Pr \{X[k-1]=x_{k-1}/X[k-2]=x_{k-2}, \dots, X[k-M-1]=x_{k-M-1}\} = \\
 & \Pr \{X[k-1]=x_{k-1}/S[k-1]=s_{k-1}\}
 \end{aligned} \tag{20}$$

in quanto la probabilità condizionata di una v. a. rispetto a sé stessa e altre v. a. è pari a 1 e la serie è di Markov di ordine M .

Analogamente si ha:

$$\begin{aligned}
 & \Pr \{S[k]=s_k/S[k-1]=s_{k-1}\} = \\
 = & \Pr \{X[k-1]=x_{k-1}, X[k-2]=x_{k-2}, \dots, X[k-M]=x_{k-M}/X[k-2]=x_{k-2}, \dots, X[k-M-1]=x_{k-M-1}\} = \\
 & = \Pr \{X[k-1]=x_{k-1}/X[k-2]=x_{k-2}, \dots, X[k-M-1]=x_{k-M-1}\} \cdot \\
 & \Pr \{X[k-2]=x_{k-2}, \dots, X[k-M]=x_{k-M}/X[k-2]=x_{k-2}, \dots, X[k-M-1]=x_{k-M-1}\} = \\
 & = \Pr \{X[k-1]=x_{k-1}/X[k-2]=x_{k-2}, \dots, X[k-M-1]=x_{k-M-1}\} = \\
 & \Pr \{X[k-1]=x_{k-1}/S[k-1]=s_{k-1}\}
 \end{aligned} \tag{21}$$

dall'eguaglianza tra la (20) e la (21) ne discende che la serie degli stati è di Markov semplice. C. d. d.

In virtù del precedente teorema per l'analisi del comportamento statistico della serie degli stati sono immediatamente applicabili i risultati derivanti in precedenza per le serie di Markov semplici.

A tale scopo giova osservare che una serie di Markov stazionaria di ordine M con il cosiddetto diagramma di stato.

Il diagramma di stato è un grafo orientato i cui nodi rappresentano i possibili stati della serie e gli archi le emissioni dei successivi termini della serie che si verificano nelle transizioni da uno stato all'altro. Gli archi sono etichettati con i valori delle probabilità condizionate di uno stato rispetto al suo predecessore, ovvero del termine della serie emesso in funzione dello stato di partenza. Ad esempio, con riferimento alla fig.1

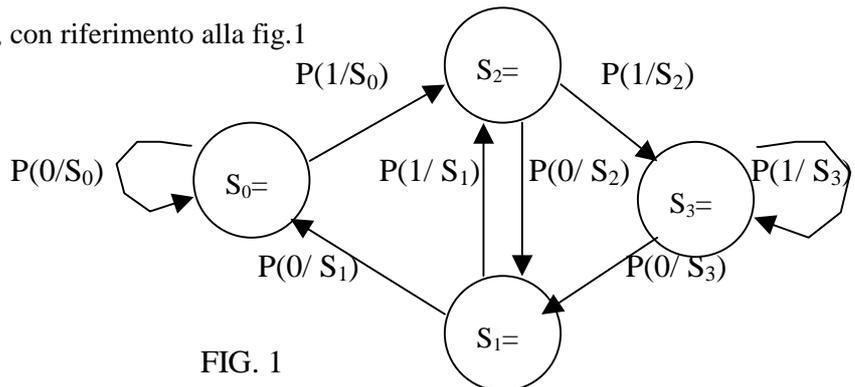


FIG. 1

Si consideri una serie di Markov di ordine $M=2$ che possa assumere determinazioni su un alfabeto binario $\{0, 1\}$. In tal caso gli stati possibili saranno:

- $S_0 \equiv (0,0)$
- $S_1 \equiv (0,1)$
- $S_2 \equiv (1,0)$
- $S_3 \equiv (1,1)$

Nella costruzione del diagramma di stato, occorrerà considerare che, per come è definito lo stato, solo alcune transizioni sono possibili. In particolare,

supponendo $S[k] = s_0$, lo stato $S[k+1]$ all'istante $k+1$ varrà ancora s_0 se la sorgente avrà emesso il simbolo 0 e varrà s_2 se la sorgente avrà emesso il simbolo 1. Si noti che la relazione (19) di normalizzazione delle probabilità condizionate si traduce nella probabilità che la somma dei pesi dei rami uscenti da un nodo sia pari a 1.

Un esempio numerico è riportato in fig. 2.

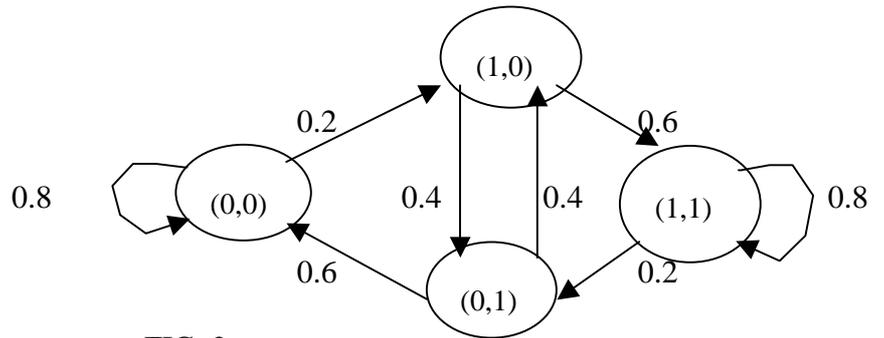


FIG. 2

Da tale figura è immediatamente deducibile la matrice Π_1 che nel caso in cui si riferisce agli stati prende anche il nome di matrice delle probabilità di transizione tra gli stati. In tal caso dalla definizione e dai dati riportati in fig. 2 si ha:

$$\Pi_1 = \begin{bmatrix} 0.8 & 0.6 & 0 & 0 \\ 0 & 0 & 0.4 & 0.2 \\ 0.2 & 0.4 & 0 & 0 \\ 0 & 0 & 0.6 & 0.8 \end{bmatrix}$$

La gerarchia di ordine 1 può essere calcolata tramite le (18) e (19) ottenendo

$$\begin{bmatrix} 0.2 & -0.4 & 0 & 0 \\ 0 & 1 & -0.4 & -0.2 \\ -0.2 & -0.4 & 1 & 0 \\ 0 & 0 & 0.4 & -0.8 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

da cui si ha

$$\begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \end{bmatrix} = \begin{bmatrix} 0.375 \\ 0.125 \\ 0.125 \\ 0.375 \end{bmatrix}$$

Si noti infine che le gerarchie di ordine 1 della serie è calcolabile in base alla gerarchia degli stati, infatti (in base alla fig. 3):

$$\begin{aligned} \Pr\{X[k]=0\} &= \Pr\{S[k+1]=s_0\} + \Pr\{S[k+1]=s_1\} = 0.375 + 0.125 = 0.5 \\ \Pr\{X[k]=1\} &= \Pr\{S[k+1]=s_2\} + \Pr\{S[k+1]=s_3\} = 0.375 + 0.125 = 0.5 \end{aligned}$$

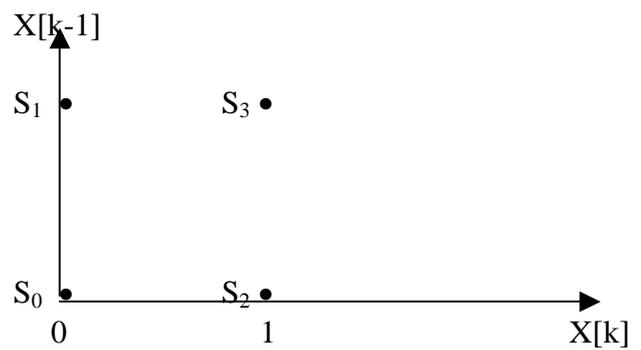


FIG. 3 – Spazio base per lo stato $S[k+1]$