

Introduzione ai processi stocastici

Disciplina nata ad inizio '900 tratta fenomeni che coinvolgono un numero infinito di variabili, quindi non abbiamo più una variabile aleatoria ma un vettore infinito di esse, da qua ne deriva che i modelli probabilistici trattati sono infinito dimensionale. Ovviamente le infinità sono numerabili se trattano campi discreti e più che numerabili se trattano campi continui.

1 Definizione di processo stocastico

E' definito come un insieme di variabili causali tale che $X = X_t, t \in T$, definite su uno spazio probabilistico (Ω, A, P) , perciò X_t è un vettore di variabili casuali, dove T è insieme degli indici ed è senza alcuna restrizione, ma ha cardinalità infinità.

Essendo X_t vettore di variabili aleatorie esso assume valori su $S = R^k$ detto spazio degli stati e tale ha una sigma algebra, dove k è il numero di variabili rilevate per ogni tempo t.

Visto che X_t può essere un vettore infinito esso non ha una funzione di ripartizione ciò lo rende di difficile trattazione.

In base a S e T avremmo diversi tipi di fenomeni:

S/T	Discreto	Continuo
Discreto	Fenomeno discreto a tempo discreto (Rovina del giocatore)	Fenomeno discreto a tempo continuo (Crescita popolazione)
Continuo	Fenomeno continuo a tempo discreto (Serie storiche)	Fenomeno continuo a tempo continuo (Prezzi azionari)

Però la situazione tipica ha $(S, T) \in R$.

Un processo stocastico ha però due definizioni:

- $\forall t \in T \Rightarrow x_t$, fisso t e vedo il processo stocastico come una collezione di variabili casuali.
- $\forall \omega \in \Omega \Rightarrow f : T \Rightarrow R$, ovvero fisso la realizzazione del esperimento, ottengo una funzione al variare di t, in questo modo il processo è una collezione di traiettorie che arriva alla realizzazione. In questa seconda visione la casualità è data dalle traiettorie, esse sono infinite e si definisce un intervallo delle possibili traiettorie.

Noi sappiamo che la funzione di ripartizione identifica univocamente una variabile aleatoria o vettori aleatori. Dato A, insieme appartenente alla sigma algebra, ha probabilità $P(Y \in A)$ e tale può essere modellizzata se abbiamo la funzione di ripartizione, con un pensiero simile riusciamo a modellizzare un processo stocastico.

2 Distribuzione di un processo stocastico

Teorema di Kolmogorov

Se prendiamo un numero di marginali finite della funzione di ripartizione ottenute attraverso le traiettorie, indicate come $F_{t_1 \dots t_n}(X_1 \dots X_n) = P(X_1 < x_1 \dots X_n < x_n)$ e definiamo $F = (F_{t_1 \dots t_n}, \forall t \in T, \forall n \geq 1)$ famiglia di funzioni di ripartizioni finite dimensionali, allora se ben definita tale famiglia identifica univocamente il processo stocastico, dove tale processo è un insieme di variabili casuali $(X_t, t \in T)$ e $X_t(\omega) : \omega \Rightarrow S \in R$ definite sulla terna di probabilità.

Quindi c'è un'estensione dal finito dimensionale del vettore di cui conosciamo la funzione di ripartizione all'infinito dimensionale del processo stocastico. Tale assunzione però non è possibile ovunque infatti impone delle forti restrizioni.

Può capitare che due traiettorie molto vicine tra loro con questo metodo non vengono distinte quindi non probabilizzate.

2.1 Restrizioni di coerenza

La famiglia di funzione di ripartizione per rispettare il teorema di Kolmogorov deve rispettare le seguenti condizioni:

- **Invarianza della funzione** se $F_{t_1, t_2}(x, y)$ funzione di ripartizione marginale finita allora vale $P(X_1 < x_1, X_2 < x_2) = F_{t_1, t_2}(x, y) = F_{t_2, t_1}(y, x) = P(X_2 < x_2, X_1 < x_1)$
- **Marginalizzazione della funzione** se vale il seguente concetto:
$$\lim_{x_{n+1} \rightarrow \infty} F_{t_1 \dots t_n, t_{n+1}}(x_1 \dots x_n, x_{n+1}) = F_{t_1 \dots t_n}(x_1 \dots x_n)$$

Se non valgono queste condizioni le marginali trovate non sono estendibili al caso infinito.

2.2 Restrizioni di infinito

Il teorema di Kolmogorov vale solo in un insieme cilindrico cioè l'insieme

$C_{t_1 \dots t_n} = (q \in R^{0, T} : (q(t_1) \dots q(t_n)) \in B) \text{ e } B \in \text{borreliano}.$

Tali cilindri sono infiniti numerabili e il teorema vale sempre se la traiettoria si sviluppa in un insieme infinito al più numerabile. La famiglia di tutti gli insiemi cilindrici forma una sigma algebra e $\sigma(C)$ ed è la più piccola sigma algebra contenente tutti gli insiemi cilindrici.

Se invece l'insieme in cui si sviluppa il processo stocastico è infinito non numerabile allora è approssimabile da funzioni marginali finite solo se hanno traiettorie holderiane-continue

(<https://almostsuremath.com/2020/10/20/the-kolmogorov-continuity-theorem/>)

ed è detto estensione al continuo del teorema di Kolmogorov.

2.3 Limiti teorema di Kolmogorov

Se prendiamo due vettori uguali quasi certamente quindi che $P(X_i = Y_i) = 1 \forall i \in (1 \dots n)$ allora le loro traiettorie finite è uguale quindi che $P(X = Y)$.

Se immaginiamo che X siano le funzioni marginali finite di un processo stocastico detto PC1 e che Y lo siano per un processo stocastico PC2, e che X ed Y si hanno funzioni infinite al più numerabile, per Kolmogorov visto che $X = Y$ q.c. anche $PC1 = PC2$ e quindi $P(X_t = Y_t) = 1 \forall t \in T$.

Se la funzione è di un infinito non numerabile allora $P(X_t = Y_t) \neq 1 \forall t \in T$ visto che è unione di infinito non numerabili non è detto che la sua probabilità faccia 1.

3 Tipi di processo

3.1 Processo totalmente indipendente

Se ipotizziamo che X_t sia I.I.D. ad $N(0,1)$ allora avremmo che $P(X_1 < x_1, \dots, X_n < x_n) = \phi(X_1) * \dots * \phi(X_n)$.

Se n tende ad infinito questo processo si chiama White Noise, per approssimare tale processo possiamo prendere un numero di marginali finito e vedere se rispetta le condizioni del teorema di Kolmogorov.

Vediamo se le marginali sono invarianti; sappiamo $F_{t_1 \dots t_n}(x_1, \dots, x_n) = F_{t_i \dots t_k}(x_i, \dots, x_k)$ per la proprietà commutativa della moltiplicazione, quindi vale la condizione.

Vediamo anche che $\lim_{x_{n+1} \rightarrow \infty} \phi(x_{n+1}) = 1$ visto che è una normale quindi vale la seconda condizione tale che $\lim_{x_{n+1} \rightarrow \infty} F_{t_1 \dots t_n, t_{n+1}}(x_1 \dots x_n, x_{n+1}) = F_{t_1 \dots t_n}(x_1 \dots x_n)$ visto che si moltiplica per 1.

Visto che rispetta tali condizioni il processo può essere descritto da un numero finito di marginali, invece riguardo allo studio delle traiettorie questo processo è assolutamente non modellizzabile, le traiettorie sono totalmente casuale e irregolari.

3.2 Processo totalmente dipendente

Se ipotizziamo che tutte le variabili casuali sono totalmente dipendenti nella formula $x_t = y = N(0, 1)$ quindi qualsiasi x è uguale all'osservazione normale y .

Questo processo è quasi totalmente deterministico infatti dipende da un'unica osservazione aleatoria, può essere descritto come $P(X_1 < x_1, \dots, X_n < x_n)$ ma visto che $X_1 = \dots = X_n$ può essere scritto come $P(X_1 < \min(x_1 \dots x_n)) = \phi(\min(x_1 \dots x_n))$ per essere un processo descrivibile con marginali finite valutiamo le condizioni del teorema di Kolmogorov.

L'invarianza della funzione è rispettata infatti $\phi(\min(x_1 \dots x_n)) = \phi(\min(x_i \dots x_k))$, anche la condizione di marginalizzazione di funzione è rispettata infatti $\lim_{x_{n+1} \rightarrow \infty} \phi(\min(x_1 \dots x_{n+1})) = \phi(\min(x_1 \dots x_n))$ visto che esisterà almeno un numero minore di infinito tale da rendere il minimo invariato.

3.3 Processo gaussiano

Quelli analizzati in precedenza sono processi gaussiani che ha **definizione**: è un processo stocastico che prendendo un qualsiasi numero finito di variabili aleatorie che compongono il processo hanno la distribuzione congiunta gaussiana.

Si nota che è l'estensione nel caso infinito di una normale multivariata, e grazie al teorema di Kolmogorov questa estensione è corretta. Per determinare univocamente un processo stocastico gaussiano abbiamo bisogno del vettore delle medie e della matrice di varianza covarianza, ma essendo ad infiniti momenti il processo dobbiamo definire:

- $E(x_t) = \mu(t)$ cioè una funzione che indica il livello medio del processo al tempo t .
- $Cov(x_t, x_s) = C(s, t)$ Una funzione che descriva la covarianza tra qualsiasi coppia in ogni istante di tempo. Questa funzione per essere giusta deve rispettare $C(s, t) = C(t, s)$, inoltre $C(t, t) = Var(t)$. C deve essere semi definita positiva.

Quindi si forma $F_{t_1 \dots t_n} = N_n(\mu, \Sigma)$, questa descrizione rispetta le condizioni del teorema di Kolmogorov, banale dato le proprietà della normale, quindi esiste un processo gaussiano descrivibile con delle marginali finite.

3.4 Processo markoviano

Un processo stocastico di Markov è un processo nel quale la probabilità di transazione che determina il passaggio ad uno stato dipende esclusivamente dello stato di sistema immediatamente precedente, e non è quindi condizionato dal passato. Quindi ogni nuova osservazione da tale processo è indipendente dalla traiettoria precedente del processo.

E' il processo più semplice su cui modellare la dipendenza e indipendenza all'interno di un processo stocastico.

E' possibile modellizzarlo nel caso discreto come $P(X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_1 = x_1) = P(X_{n+1} = x_{n+1} | X_n = x_n)$ invece nel continuo essendo un processo con n che va ad infinito non numerabile, per rappresentarlo attraverso il teorema di Kolomogorov, dovremmo supporre che dati un'estrazione di tempi k ordine le $t_k > \dots > t_i$ allora posso scrivere come $P(X_t = x_t | X_{t_k} = x_{t_k}, \dots, X_{t_1} = x_{t_1}) = P(X_t = x_t | X_{t_k} = x_{t_k})$.

4 Catena di Markov

E' il caso di un processo markoviano a tempo discreto con spazio degli stati discreto esso è caratterizzato dalla matrice di transizione. Si può scrivere come $P(X_{n+1} = j | X_n = i, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i) = P_n(i, j)$ la matrice di transazione è l'insieme di tutte le $P_n(i, j)$ probabilità.

Si dice che una catena è **omogenea** se il valore è indipendente da n quindi la posizione della traiettoria e si scrive che $P(X_{n+1} = j | X_n = i) = P(i, j)$. Notiamo che lo spazio degli stati ha la cardinalità di n quindi numerabile.

La catena di Markov può essere descritta da due elementi:

- Stato iniziale della catena che è il valore $P(X_0 = i) = \mu_i$ dove $i \in S$ (spazio degli stati) dove $\mu_i \geq 0$ la sua sommatoria su S da ovviamente 1, essendo funzione di probabilità.
- Matrice di transizione cioè una matrice che contiene tutti i valori di $P_n(i, j)$, essendo matrice di probabilità ha le righe che sommano ad 1, infatti è detta anche matrice stocastica.

Ora verifichiamo se tale processo è descrivibile da un numero finito di variabili attraverso il teorema di Kolmogorov, scriviamo tale processo come $P(X_0 = i_0, \dots, X_n = i_n) = P(X_0 = i_0) * P(X_1 = i_1 | X_0 = i_0) * \dots * P(X_n = i_n | X_{n-1} = i_{n-1}, \dots, X_0 = i_0)$, visto che dipende solo dal presente è possiamo vederlo scritto come $\mu_0 * P(i_1, i_0) * \dots * P(i_n, i_{n-1})$.

Notiamo subito che essendo una produttoria la prima restrizione del teorema è rispettata, verifichiamo la seconda e marginalizziamo rispetto a $n + 1$, essendo una variabile discreta usiamo la sommatoria e otteniamo quindi $\sum_{i_{n+1} \in S} P(X_0 = i_0, \dots, X_n = i_n) = \mu_0 * P(i_1, i_0) * \dots * P(i_n, i_{n-1}) * \sum_{i_{n+1} \in S} P(i_{n+1}, i_n)$ dove la sommatoria essendo sul dominio di probabilità fa 1 e quindi è rispettata anche la seconda condizione di Kolmogorov. L'esempio del white noise è un caso di catena di markow omogenea.

4.1 Struttura e proprietà delle catene di Markov omogenee discrete

Sia X_0 il punto di partenza della catena sia Y_i variabile I.I.D e indipendente rispetto ad X allora un processo markoviano può essere scritto come $X_{n+1} = f(x_n, y_n)$ abbiamo una componente determinista data da x_n e una aleatoria data da y_n , se il processo ha tale struttura è un processo markoviano omogeneo e vale anche il viceversa. Se mancasse la componente y_n avremmo un processo totalmente deterministico.

Dimostriamo che tale scrittura è corretta sappiamo che $P(X_{n+1} = j | X_n = i_n, \dots, X_0 = i_0) = P(f(x_n, y_n) = j | X_n = i_n, \dots, X_0 = i_0)$ sappiamo che le Y_n sono I.I.D. quindi $P(f(i_n, y_n) = j) = P(X_{n+1} = j | X_n = i_n)$ cioè la marginale dato X_n quindi vale la proprietà del processo markoviano che $X_{n+1} | X_n, \dots, X_0 = X_{n+1} | X_n$.

Ora dobbiamo dimostrare che l'intero futuro non dipende dal passato cioè $X_{n+m}|X_n, \dots, X_0 = X_{n+m}|X_n$ questo perchè $P(X_{n+m} = i_{n+m}, \dots, X_{n+1} = i_{n+1} | X_n = i_n, \dots, X_0 = i_0) = \frac{P(X_{n+m}=i_{n+m}, \dots, X_0=i_0)}{P(X_n=i_n, \dots, X_0=i_0)}$ i termini della moltiplicazione si elidono ed otteniamo $P(i_n, i_{n+1}) * \dots * P(i_{n+m}, i_{n+m}) = P(X_{n+m} = i_{n+m}, \dots, X_{n+1} = i_{n+1} | X_n = i_n)$ cioè la tesi.

Infine notiamo la proprietà di **assenza di memoria** nella catena ovvero che il futuro non dipende dalla traiettorie del processo questo è immediato dalla scrittura precedente infatti se $X_n = X_0$ allora $X_{n+m}|X_n = X_{n+m}|X_0$.

4.2 Trasformazioni in catene markoviane

Un processo può non essere markoviano ma possiamo renderlo tale, infatti possiamo creare un nuovo processo in cui l'osservazione sia indipendente dal passato.

Per esempio se una variabile dipende anche dall'istante precedente oltre che quello presente posso creare tutte le combinazioni dei risultati dei due istanti e numerarle, così facendo accoppio i valori dipendenti man mano creando un processo dipendente solo dal presente.

4.2.1 Esempio rovina del giocatore

Se due giocatori hanno capitale iniziale C_a e C_b e ogni volta scommettono un'euro avranno come spazio degli stati $(0, \dots, C_a + C_b)$ e se ci mettiamo nella posizione di A, esso ha p come probabilità di vincita e $1-p$ come probabilità di perdita.

Tale processo formerà una catena di Markov dove l'osservazione $X_n = I_n$ avrà come evoluzione $X_{n+1} = I_n + 1$ o $X_{n+1} = I_n - 1$, notiamo quindi che X_{n+1} dipende sì da X_n ma esclusivamente della posizione in cui si trova, infatti avrà $+1$ o -1 indifferentemente dal percorso fatto in precedenza per raggiungere X_n .

Tale processo markoviano avrà come matrice di transizione:

$$A = \begin{bmatrix} & 0 & 1 & 2 & C_a + C_b \\ 0 & 1 & 0 & \dots & 0 \\ 1 & 1-p & 0 & p & 0 \\ 2 & 0 & 1-p & 0 & p \\ C_a + C_b & 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

Vediamo che in 0 e in $C_a + C_b$ la probabilità è uno ed è detta barriera assorbente cioè se il processo arriva in quel punto il gioco finisce.

4.3 Equazione di Champan-Kolomogorv

Se prendo una catena di Markov omogenea con un numero finito di stati ed ho μ e la matrice di transizione $P^{(1)}$, allora $P_{ij}^{(1)} = P(X_{n+1} = j | X_n = i)$ e la matrice di transizione m -esima è indicata come $P_{ij}^{(m)} = P(X_{n+m} = j | X_n = i)$.

Se $X_0 = i$ allora possiamo scrivere $P_{ij}^{(m)} = P(X_{n+m} = j | X_0 = i)$ visto che essendo processo markoviano omogeneo il processo non dipende dal tempo dello stato ma solo dallo stato in se.

Se il processo è omogeneo operativamente posso anche moltiplicare la matrice delle transizione n -volte e prendere P^n nella cella (i,j) .

Se voglio trovare $P_{ij}^{(2)}$ esso è uguale a $P(X_2 = j | X_0 = i) = \sum_{h \in S} P(X_2 = j, X_1 = h | X_0 = i)$ dove S spazio degli stati, che è uguale a $\sum_{h \in S} P(X_2 = j | X_1 = h, X_0 = i) * P(X_1 = h | X_0 = i)$
 $= \sum_{h \in S} P(X_2 = j | X_1 = h) * P(X_1 = h | X_0 = i) = \sum_{h \in S} P(i, h) * P(h, j)$. Quindi in questo modo passo da i a j e ne calcolo la probabilità sommando le probabilità di ogni traiettoria.

4.3.1 Prima equazione

Si ha che $P^{(m+n)} = \sum_{h \in S} P^{(m)}(i, h) * P^{(n)}(h, j)$ dove h è il valore in m . In questo modo calcolo la probabilità di tutte le traiettorie da 0 a m che partono da i , e le moltiplico per tutte le probabilità da m a n che partono da un generico h e arrivano a j , mentre i e j sono fissi h varia in tutto S .

Vale la seguente: $P(X_{n+m} = j | X_0 = i) = \sum_{h \in S} P(X_{n+m} = j | X_m = h) * P(X_m = h | X_0 = i)$.

Tale calcolo si può scrivere in formula matriciale come $P^{(m+n)} = P^{(m)} * P^{(n)}$ quindi è semplice ma ci sono moltissime moltiplicazioni matriciali.

Per far quadrare le formule precedenti è necessario dire che $P^{(0)} = I$ cioè indice che la transazione dallo stato 0 non è possibile infatti come visto precedentemente a 0 si ferma il gioco.

4.3.2 Seconda equazione:

Se conosco $P(X_m = j | X_0 = i)$ posso ottenere $P(X_m = j)$ marginalizzando cioè:

$P(X_m = j) = \sum_{i \in S} P(X_m = j | X_0 = i) * P(X_0 = i) = \sum_{i \in S} P^{(m)}(i, j) * \mu(i)$ tale scrittura può essere trasformata in maniera matriciale come $P(X_m = j) = \mu' * P^m$. E' facile notare che più m è grande più il potere previsivo sulla posizione del punto si abbassa.

4.4 Classificazione degli stati

Gli stati possono essere:

- **Stati transitori:** uno stato destinato a non essere più percorso con il crescere del tempo
- **Stati ricorrenti:** uno stato percorso infinite volte con il passare del tempo.

4.4.1 Proprietà degli stati

- **Accessibilità:** Si dice che lo stato j è accessibile da i se esiste almeno un valore di n tale che $P^{(n)}(i, j) > 0$ e si indica come $i \Rightarrow j$
- **Comunicabilità:** Se $i \Rightarrow j$ e $j \Rightarrow i$ allora si ha una relazione biunivoca della comunicabilità e si scrive come $i < - - > j$.
- **Riflessività:** Se i è comunicabile con se stesso.
- **Simmetria:** Cioè se $i < - - > j$ vale anche il vice versa $j < - - > i$
- **Transitività:** se $i < - - > j$ e $j < - - > k$ allora $i < - - > k$.
- **Assorbente:** Se ha solo accessibilità da uno stato ad esso ed è riflessivo.

Date queste proprietà si possono creare delle ripartizioni spaziali.

4.4.2 Partizioni degli stati

- **Chiusura** un sotto insieme di stati C si dice chiuso se dato $i \in C$ e $j \notin C$ allora non esiste alcuna $P(i, j) > 0$
- **Irriducibilità** Un insieme formato solo da stati comunicanti

4.5 Stati transitori e ricorrenti

Se scriviamo che $f_i = P(\text{tornare in } i \text{ prima o poi} | X_n = i)$, che è uguale a $P^n(i, i)$, e $f_i(n) = P(\text{tornare in } i \text{ prima per la prima volta al tempo } n | X_n = i) = P(X_n = i, X_n \neq i, \dots, X_1 \neq i | X_0 = i)$ questa probabilità descrive tutto il percorso del processo ed è diversa da quella precedente.

Allora per trovare $f_i = \sum_{n=1}^{\infty} f_i(n)$ è una sommatoria perchè la probabilità sono disgiunte visto che essendoci assenza di memoria una volta toccato i si azzerà tutto.

Se $f_i = 1$ allora lo stato i è uno stato ricorrente quindi si ritornerà sempre a quello stato, e visto che il processo è infinito si passerà infinite volte allo stato i .

Se indichiamo il passaggio dal punto i come una funzione indicatrice, quindi se passa in i è 1, avremmo che se $N = \sum I_n = \infty$ quindi è una variabile casuale degenera.

Se invece $f_i < 1$ allora lo stato i è transitorio, ciò significa che non tutti le traiettorie passano in i e che con il crescere dei passi nessuna traiettoria attraverserà più lo stato i .

Se indichiamo il passaggio dal punto i come una funzione indicatrice, quindi se passa in i è 1, avremmo che se $N = \sum I_n = N$ quindi è una variabile casuale.

La variabile in questione è la geometrica infatti la probabilità che il percorso non passi più nello stato i è scrivibile come $Z(n) = f_i^{n-1} * (1 - f_i)$ e quindi ha valore atteso $\frac{1}{1-f_i}$.

Ne consegue che valgono le seguenti proprietà:

- Un insieme di stati chiuso è formato da almeno uno stato ricorrente o da un numero infinito di spazi transitori.
- Se i è ricorrente e $i < \dots > j$ anche j è ricorrente e vice versa.
- Se i è transitorio e $i < \dots > j$ anche j è transitorio e vice versa.
- Se un insieme chiuso è formato da stati comunicanti allora sono tutti ricorrenti, vale anche il vice versa.
- **Teorema di decomposizione degli stati:** se S spazio degli stati allora $S = c_1 \cup \dots \cup T$ dove c_i sono gli insieme chiusi e T tutti i sotto insiemi aperti cioè l'insieme di tutti gli stati transitori, in ogni spazio esiste almeno un c_1 ma non sempre un T .
- Se $i \Rightarrow j$ allora lo stato i è uno stato transitorio, non vale il vice versa.

Possiamo riconoscere uno stato se ricorrente o transitorio valutando se $f1 = \sum_{n>0} f1(n) = 1$ operativamente possiamo vederlo analizzando il grafo e vedendo le probabilità $P(1,1) + P(1,2) * P(2,1) + \dots$ la difficoltà qua sta nel saper risolvere la serie che uscirà, con questa sommatoria calcoliamo la probabilità che il processo parta in 1 e ritorni in 1.

Possiamo anche calcolare la probabilità di tornare lo stato i per la prima volta.

Possiamo anche calcolare il tempo medio di ritorno in i ed è calcolabile come $\sum_{n>0} n * f1(n)$.

4.6 Distribuzioni limite: Limite delle catene di Markov

Visto che nel lungo periodo gli stati transitori vengono abbandonati allora se i è uno stato transitorio, allora $\lim_{n \rightarrow \infty} P_n(i, i) = 0$ ovvero che lo stato non tornerà più in se stesso, inoltre ciò è indipendente dal punto di partenza infatti se j è uno stato transitorio, allora $\lim_{n \rightarrow \infty} P_n(j, i) = 0$.

Per capire se uno stato è transitorio oppure no decomponiamo gli spazii in $c_1 \cup \dots \cup T$, dove c_i sono detti insiemi irriducibili.

Su T è interessante calcolare il tempo medio di abbandono e invece su c_i è interessante sapere la probabilità di raggiungere c_i se parto da c_j .

Questa seconda considerazione possiamo calcolarla. Prendiamo per esempio un insieme chiuso formato da due stati che ha come matrice di transizione

$$A = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix} \quad (2)$$

Allora se voglio calcolare la probabilità di raggiungere c_i partendo da c_i calcolo il $\lim_{n \rightarrow \infty} P_n(i, i)$ che in questo caso è $\frac{\beta}{\alpha + \beta}$ ma posso calcolare anche $\lim_{n \rightarrow \infty} P_n(j, i)$ che è $\frac{\alpha}{\alpha + \beta}$. Posso anche creare la matrice di transizione al limite che è:

$$A = \begin{bmatrix} \frac{\beta}{\alpha + \beta} & \frac{\alpha}{\alpha + \beta} \\ \frac{\beta}{\alpha + \beta} & \frac{\alpha}{\alpha + \beta} \end{bmatrix} \quad (3)$$

Notiamo che il punto di partenza con $n \rightarrow \infty$ non è più importante, notiamo anche che il limite che troviamo è finito, però non è sempre così.

4.7 Catene periodiche e aperiodiche

Se le catene sono senza limite possono essere periodiche cioè che la $P_{k*n}(1,1) > 0$, ma se m/kn , con k periodo del ciclo, non è un numero naturale, allora $P_m(1,1) = 0$ perciò il limite non è definibile.

Esistono anche le catene senza limite definite aperiodiche, questo indica che tra tutti i passi per ritornare ad i il massimo comune divisore è 1.

Inoltre con queste definizioni sappiamo che:

- Se tutti gli stati sono aperiodici allora anche la catena che formano è aperiodica.
- Se i e j comunicano allora hanno la stessa periodicità, ciò intende che se ho uno insieme chiuso allora la periodicità di tale insieme è uguale per tutti.
- Se esiste un $P_1(i,i) > 0$ allora si ha uno stato riflessivo ciò crea uno stato aperiodico che rende aperiodica tutta la catena.

4.8 Distribuzione di lungo periodo

Se consideriamo ora solo catene aperiodiche con tutti gli stati comunicanti, sappiamo che la probabilità di lungo periodo cioè $\lim_{n \rightarrow \infty} P_n(i,j) = \pi(j)$ notiamo che dipende solo dal punto di arrivo e non quello di partenza, cioè la probabilità che una catena finisca in un certo stato.

La distribuzione per una catena di Markov è una funzione di probabilità che assegna una probabilità ad ognuno degli stati della catena cioè $\pi = (\pi(1), \dots, \pi(n))$, vettore colonna.

Tale vettore si definisce **stazionario** se vale l'uguaglianza $\pi'P = \pi'$ dove P matrice di transizione. Se vale tale proprietà vuol dire che il vettore π rimane costante per ogni X_i , quindi non cambia nel tempo.

Supponiamo che $\mu = x_0 \sim \pi$ allora $P(X_n = j) = \sum_{i=1}^k P(X_{n-1} = i) * P(X_n = j | X_{n-1} = i) = \pi(i) * P(i,j) = \pi(j)$ da qui ne consegue che ogni $X_i \sim \pi$. Notiamo che questa proprietà indica anche che il processo è identicamente distribuito.

Teorema: Se X_n è una catena irriducibile e aperiodica allora esiste un'unica soluzione all'equazione $\pi'P = \pi'$ e tale soluzione $\lim_{n \rightarrow \infty} P_n(i,j) = \pi(j)$, ciò garantisce esistenza e unicità di π inoltre il tempo medio di ritorno allo stato i è $\frac{1}{\pi(i)}$.

Legge dei grandi numeri Sappiamo che se Y_i sono I.I.D. ed hanno un certo valore finito $E(Y_i) = \mu$ allora $\frac{\sum Y_i}{n} \rightarrow \mu$ quindi ne converge in probabilità, estendendo questo concetto alle catene markoviane irriducibili e aperiodiche otteniamo che $\frac{\sum f(x_i)}{n} \rightarrow \sum f(j) * \pi(j)$.

Da qua ne consegue che posso trovare la condizionata $\pi(j) = \lim_{n \rightarrow \infty} P^{(n)}(i,j) = \lim_{n \rightarrow \infty} P(X_n = j | x_0 = i)$ che essendo I.I.D è uguale alla marginale $\lim_{n \rightarrow \infty} P(X_n = j) = \pi(j)$.

Inoltre l'equazione $\pi'P = \pi'$ è un sistema a k equazione con $k-1$ incognite quindi ha rango $k-1$ e ciò ci dice che esiste $\sum \pi(i) = 1$.

Se voglio sapere nei primi passi quante volte una catena è passata da un certo stato, di fatto devo calcolare una percentuale, ovvero la percentuale di volte che ci si trova in un certo stato in un certo periodo: $\frac{N^n}{n}$ essa però se $n \rightarrow \infty$ converge in probabilità a $\pi(j)$.

5 Catene di markov a tempo continuo

Se consideriamo il processo $(X_t, t \geq 0)$ esso è una catena markoviana a tempo continuo se per ogni istante di tempo ordinato $0 < S_0 < \dots < S_n < S < \dots < S + t$ e per ogni stato discreto i vale che: $P(X_{t+s} = j | X_s = i_n, \dots, X_0 = i_0) = P(X_{t+s} = j | X_s = i_n)$ e se $i_n = i_0 = j$ allora $= P(X_t = j | X_0 = j) = P_t(i, j)$, quindi come nella catena di markov gli stati passati non influenzano gli stati futuri, si dice omogenea se P non dipende da t .

La differenza nel continuo delle equazione di chanpan si ha che non avremo più dei punti come valore del processo ma dei segmenti discontinui, quindi invece che usare le sommatorie uso gli integrali. Per descrivere tale percorso ci servono quindi due informazioni:

- **I salti** essi sono una variabile aleatoria che come nella catena markoviana descriveremo attraverso una matrice P .
- **I tempi in cui avvengono i salti** Sono una variabile aleatoria che calcola il tempo tra un salto e un altro, esso non è influenzato del tempo che è in quella posizione, il tempo di intra-salto si dispone come una $Exp(\lambda(j))$ notiamo quindi che se il tasso di abbandono è basso rimaniamo molto a lungo in quello stato.

Per modellizzare la catena o sappiamo quando avvengono i salti oppure il tempo che trascorre tra uno e l'altro ci è più facile calcolare la seconda.

5.1 Altri tipi di processo markoviano continui

5.1.1 Processo a incrementi indipendenti

Se abbiamo un processo a tempo continuo $X = (X_t, t \geq 0)$. Tale processo è detto a incrementi indipendenti se vi è correlazione tra i livelli dei vari stati ma l'incremento tra uno stato e un altro è totalmente indipendente dal presente e dal passato. Quindi anche se so quello che è successo tra 0 e S chiamato Σ_s allora $X_t | \Sigma_s \rightarrow X_t | X_s$.

Si parla di incrementi stazionari se gli incrementi dipendo solo dal Δ_t e non dalla posizione di esso cioè $\Delta_t = X_t - X_0 = X_{t+s} - X_{t-s}$.

5.1.2 Processo di conteggio

Un processo costante a tratti su \mathbb{R} che ha dei salti positivi, inoltre tale funzione ha $N_0 = 0$ cioè il punto di partenza uguale a 0, e all'aumentare di t il numero N aumenta o rimane costante.

6 Processo di Poisson

E' un processo di conteggio, ed è una speciale catena markoviana che ha come unico salto ammesso il passaggio di stato in $+1$, e gli incrementi sono tra loro indipendenti. Dove N_t è il valore dello stato in cui siamo, tale valore si dispone come una poissoniana di media λ_t .

Per creare tale processo si sono fatte due importanti assunzioni:

- Il numero di arrivi è limitato quindi $0 < E(N_t) < \infty$
- Assenza di arrivi multipli, cioè in ogni istante temporale può aumentare di solo un unità il processo.

Inoltre il processo ha le seguenti proprietà:

- $N_0 = 0$
- incrementi indipendenti
- incrementi stazionari ovvero $N_{s+t} - N_t$ e si dispongono come una $Po(\lambda * t)$ dove λ è un tasso costante che corrisponde alla probabilità di arrivo in un determinato stato.

Queste proprietà sono sufficienti quindi se valgono tali condizioni il processo è di Poisson. E' dimostrabile che il miglior modo per trattare il conteggio è quello poissoniano infatti, se abbiamo k intervalli di lunghezza t/k che presentano un numero di arrivi I.I.D. se $k \rightarrow \infty$ in questi minuscoli intervalli è dimostrabile che il numero con probabilità quasi certa o è 1 o 0, quindi per vedere se nell'istante t_i/k c'è stato un arrivo lo si parametrizza con una bernulliana $Ber(\lambda * t/k)$ per valutare il numero di arrivi su t , useremo la sua somma cioè una binomiale $Bi(k, \lambda * t/k)$ però come abbiamo detto $k \rightarrow \infty$ quindi si trasforma in una poissoniana $Po(\lambda * t)$ e si nota la stazionarietà.

Possiamo modellare anche quanto tempo passa da un alto all'altro questo ha forma esponenziale di parametro λ ($\text{Exp}(\lambda)$). Quindi possiamo calcolare la probabilità di "per quanto tempo non c'è un arrivo" che equivale a dire "non ci sono stati arrivi tra s e t " $P(T_i > t) = P(N_t = 0) = \exp -\lambda * t$. Inoltre visto che è modellato da un esponenziale il tempo in cui rimane in i non è influenzato dal tempo in cui c'è stato fin'ora, ciò porta ad un assenza di memoria, infatti $P(X > s + t | X > t) = P(X > s)$, posso vedere anche il tempo di inter-arrivo come $P(N_{t+s} - N_t = 0 | N_t = 1) = P(N_{t+s} - N_t = 0)$. Da qua possiamo ricavare che $= P(N_{t+s} - N_s = 0 | N_s = N_0) = P(N_{t+s} - N_s = 0) = Po(\lambda * t) = \exp -\lambda * t$ e così facendo possiamo dire che i tempi di inter-arrivo cioè $T_i = s_i - s_{i-1}$ si dispongono come un esponenziale ($\text{Exp}(\lambda)$), e sono tra di loro I.I.D.

Dunque per simulare un processo di poisson invece che simulare gli N_t che sono dipendenti gli uni dagli altri si preferisce utilizzare T_i .

Ora vogliamo modellare la $P(N_t = j | N_s = i)$ cioè la probabilità di arrivare j al tempo t dopo che sono passato per i al tempo s , si può trasformare come $\frac{P(N_t=j, N_s=i)}{P(N_s=i)} = \frac{P(N_t=j, N_t-N_s=j-i)}{P(N_s=i)}$ che essendo indipendenti ho $P(N_t - N_s = j - i)$ e quindi si dispongono come una $\text{Poisson}(\lambda(t - s))$.

Se modelliamo invece $P(T_1 < s | N_t = 1)$ cioè la probabilità che l'arrivo arrivi nel segmento di tempo s sapendo in un tempo t c'è stato un arrivo. Può essere scritto come $\frac{P(T_1 < s, N_t=1)}{P(N_t=1)} = \frac{P(N_s - N_0=1, N_t - N_s=0)}{P(N_t=1)}$ essendo intervalli slegati $= \frac{P(N_s - N_0=1) * P(N_t - N_s=0)}{P(N_t=1)} = \frac{s}{t}$ notiamo quindi che si dispone come una uniforme in $U(0, t)$.

Ora immaginiamo che ci siano molti arrivi nel tempo t , per modellizzare tali arrivi possiamo generare da un uniforme $U(0, t)$ valori quanti gli arrivi, e poi ordinarli dal più grande al più piccolo e quelli saranno i tempi d'arrivo.

Se invece voglio sapere il modello che ha $P(N_s = i | N_t = j) = \frac{P(N_s=i, N_t-N_s=j-i)}{P(N_t=j)} = \text{Bin}(j, \frac{s}{t})$. Ed è la forma generale di quanto visto nel paragrafo prima.

6.1 Processo di poisson eterogeneo

Fin'ora abbiamo trattato processi omogenei, se venisse rimossa questa condizione avremmo che, il parametro della poisson non sarebbe più costante ad λ , ma dipende dall'integrale sull'intervallo preso in considerazione. In particolare se l'intervallo va da s a $s+t$ allora avremmo che $\lambda = \int_s^{s+t} \lambda(x) dx$, dove la funzione λ varia.

In questo caso però $S_1 | N_t$ non si dispone come una $U(0, 1)$, ma come $f(x) = \frac{\lambda(x)}{\int_0^t \lambda(x) dx}$.

7 Moto browniano

E' un processo definito nel continuo che si muove con movimenti casuali in direzioni casuali. Tale processo deriva da una camminata aleatoria simmetrica nel discreto ma trattato al limite.

Se definisco Y_n come una variabile dicotomica che assume i valori ± 1 e sono I.I.D, e costruiamo $X_n = Y_1 + \dots + Y_n$ in questo modo si avranno incrementi indipendenti, ora studiamo al limite tali salti spazio-temporali, così da analizzare in continuo tali processi.

In particolare se chiamiamo τ il salto temporale e γ il salto spaziale entrambi tendenti a 0. In questo modo posso definire il salto spaziale come $\gamma * Y_n$ e il salto temporale come t/τ dove $t < \tau$, così facendo posso costruire il processo come $X_t = \gamma(Y_1 + \dots + Y_{\frac{t}{\tau}})$ ed è il processo nel continuo.

Ora analizziamo i momenti di tale processo vediamo facilmente che $E(Y_n) = 0$ e $Var(Y_n) = 1$. Quindi essendo somma di I.I.D il processo X_n avremo $E(X_n) = 0$ e $Var(X_n) = \gamma^2 \frac{t}{\tau}$ ciò ci fa notare che γ perchè Var non sia degenerare deve avere velocità di caduta $\sqrt{\tau}$, nel nostro caso porremo $\gamma = \sigma * \sqrt{\tau}$, così facendo ottengo $Var(X_n) = \sigma^2 * t$.

E' facile dedurre che essendo somma di I.I.D allora varrà il teorema centrale del limite, quindi $X_t \rightarrow N(0, \sigma^2 * t)$, si note che la varianza cresce linearmente rispetto al tempo passato.

Da qua possiamo dare allora una definizione di moto browniano, un processo con stati infiniti e ricorrenti. Dato B_t moto browniano uni-dimensionale con varianza σ^2 sarà tale se:

- $B_0=0$ per convenzione.
- Incrementi indipendenti cioè $B_t - B_s$ è indipendente da B_s , essendo un particolare processo markoviano.
- Incrementi stazionari cioè che $B_t - B_s$ dipende solo da $t - s$.
- B_t è distribuito come $N(0, \sigma^2 * t)$.
- t nel processo è continua.

Questo processo è continuo ma non differenziabile, quindi completamente indipendente. Infine passa infinite volte per le 0 o qualsiasi valore fissato.

7.1 Distribuzione di B_t

Vediamo se riusciamo ad identificare il processo come un insieme di distribuzioni marginali finite, in questo caso non si può usare il teorema di Kolmogorov visto che il tempo è in un insieme più grande di quelli al più numerabile, almeno che la funzione non sia regolare.

Sappiamo che le differenze $X_{t_1} - X_0, \dots, X_n - X_{n_1}$ si dispongono come una normale multivariata di media 0, matrice varianza covarianza una matrice diagonali di σ^2 , questo perchè gli incrementi sono indipendenti.

Sappiamo che X_t è la somma di tutti questi incrementi ed essendo gaussiano otteniamo a nostra volta un processo gaussiano con media 0, varianza $\sigma^2 * t$. Con covarianza calcolabile come $Cov(X_t, X_s) = E(X_t X_s) - E(X_t)E(X_s) = E(X_t X_s) = E(((X_t - X_s) + (X_s - X_0))(X_s - X_0))$ che essendo tutti intervalli disgiunti sono anche indipendenti ed ottengo $\sigma^2 * s$, perciò la covarianza è proporzionale al tratto di traiettoria in comune tra i due processi, vediamo quindi che vale indipendenza e stazionarietà del processo.

7.2 Moto browniano come trasformato di poisson

Può essere visto come una trasformato di una poisson (X_t) come $Y_t = X_t + u * t$ dove $u * t$ rappresenta il drift, in questa forma però non attraverserà lo 0 infinite volte ma abbandonerà pian piano le condizioni iniziali nel verso del drift.

7.3 Assenza di memoria

Il moto browniano ha come proprietà l'assenza di memoria infatti se X_t è browniano allora anche $Y_t = X_{t+h} - X_h$ è browniano, il processo in questione ha come valore atteso 0 essendo trasformazione lineare di normali con media 0, ha come varianza $\sigma^2 t$ dove $t = t + h - h$. Ha come covarianza $Cov(Y_s, Y_t) = E(Y_s Y_t) = E((X_{s+h} - X_h)(X_{t+h} - X_h))$ essendo non disgiunti devo dividerli, perciò $E((X_{s+h} - X_h)(X_{t+h} - X_{s+h} + X_{s+h} - X_h))$ ora sono disgiunti quindi indipendenti da qui ottengo $\sigma^2 * s$, ed ecco dimostrato che è un moto browniano.

In generale vale che $X_t | X_s = x \rightarrow X_t - X_s + x | X_s = x \rightarrow N(x, \sigma^2(t - s))$.

7.4 Distribuzioni condizionate del moto browniano

Sappiamo che data una soglia a il processo raggiungerà tale soglia con probabilità 1.

Ci si può anche chiedere: Se T_a è il tempo di arrivo in a come è distribuita questa variabile casuale.

Modelliamo la probabilità che x raggiunga a in T_a come:

$$P(X_t > a) = P(X_t > a | T_a < t) * P(T_a < t) + P(X_t > a | T_a > t) * P(T_a > t) =$$

$P(X_t > a | T_a < t) * P(T_a < t)$ partendo da 0 il modello ed avendo assenza di memoria sappiamo che la probabilità di salire cioè è $P(X_t > a | T_a < t) = 1/2$ perciò abbiamo $P(X_t > a) = 1/2 * P(T_a < t)$.

Da qua posso trovare $P(T_a < t) = 2 * P(X_t > a)$ sapendo che X_t si dispone come una normale di varianza t la indico come $2 * \int_a^\infty \frac{e^{-x^2/2t}}{\sqrt{2\pi * t}} dx$ se cambio variabile e dico che $y = x/\sqrt{t}$ ottenendo

$2 * \int_{a/\sqrt{t}}^\infty \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy$ posso derivare rispetto a t usando la formula $\frac{d}{dt} \int_{a/\sqrt{t}}^\infty f(y) dy = f(g(t)) * (-\frac{dg(t)}{dt})$ così da ottenere la funzione di distribuzione di T_a cioè $\frac{a}{\sqrt{2\pi}} * t^{-3/2} * e^{-\frac{a^2}{2t}}$.

Posso chiedermi anche se preso un momento finito $(0, t)$ quale sia il massimo del processo e che distribuzione ha.

Come considerazione iniziale sappiamo che tale variabile casuale $W_t = \max(X_s) \geq 0$ visto che il moto parte da 0. Si può modellare come $P(W_t > w) = P(T_w < t)$, possiamo vederlo come il tempo di arrivo in tale valore, se il tempo d'arrivo non esiste allora il valore massimo è insensato.

Sappiamo da prima la distribuzione di T_w , la derivo rispetto a w e trovo la distribuzione di W_t che è $\sqrt{\frac{2}{\pi * t}} * e^{-\frac{w^2}{2t}}$. Questa distribuzione assomiglia molto a una normale anche se tutta via è il suo modulo.

8 Processi di punto nello spazio

In questa sezione si tratteranno processi a temporali quindi dove l'ascissa non è più il tempo. In particolare un processo di punto è un realizzazione atemporale di punti. Dato un insieme A contenuto in D sottoinsieme di R^d vogliamo sapere il numero di punti che cadono in A $N(A)$, ed a tale numero si potrà associare una variabile casuale.

Tali processi soddisfano la proprietà di additività quindi se A e B disgiunti allora $N(A \cup B) = N(A) + N(B)$ e ciò vale anche se è un unione infinita di insiemi.

8.1 Processo di poisson omogeneo

E' detto omogeneo se dati due insiemi $N(a_i, b_i)$ si dispone come $Po(\lambda(b_i - a_i))$ o più precisamente come $Po(\lambda * v(A_i))$ dove v è la misura di Lebesgue, quindi il numero di punti rimane costante in ogni punti del sottoinsieme.

Tale processo è una catena di Markov e perciò soddisfa le condizioni di Kolmogorov.

Ha le seguenti proprietà:

- Indipendenza: siano B_1, B_2 insieme disgiunti allora $N(B_1) \perp N(B_2)$ ne consegue che $N_{t_2} - N_{t_1} = N(t_1, t_2)$ e vale $N(t_1, t_2) \perp N(t_2, t_3)$
- Stazionarietà: $N(B) = Po(\lambda * v(B))$ notiamo quindi che dipende solo dall'area analizzata, infatti è invariante rispetto alle traslazioni di area, visto che l'area spostata non si allarga ne restringe. E' detto invariante rispetto ai momenti rigidi cioè rotazione, traslazione e riflessione.

Vediamo che se $A_n \downarrow \emptyset$ avrò che $N(A_n) = 0$ quindi abbiamo continuità, questo perchè $Po(0)=0$ quindi è ovvio che se ho area 0 ho anche 0 punti, questa proprietà rende N una misura su sigma algebra.

Da qua ne consegue che se B_1, B_2 insieme disgiunti allora $N(B_1) + N(B_2) = Po(\lambda * v(B_1 \cup B_2)) = Po(\lambda * v(B_1) + \lambda * v(B_2))$.

Il problema sorge se gli insiemi non sono disgiunti risulta più difficile applicare il teorema di Kolmogorov. Ecco perchè se ho A e B non disgiunti mi ricavo tre insiemi disgiunti da essi $C_1 = A - B$ $C_2 = B - A$ $C_3 = A \cap B$ allora trovo che le sue marginali sono $P(N(A) = i, N(B) = j) = P(N(C_1) + N(C_3) = i, N(C_2) + N(C_3) = j)$ che possiamo condizionarlo a $C_3 = n$ visto che dipendono da essi, e sommare tutti i valori condizionati ottenendo così $P(N(C_1) = i - n, N(C_2) = j - n)$ che è indipendente dall'intersezione.

8.1.1 Generazione di un processo

- Individuo un generico sottoinsieme I e ne trovo il numero di punti, dove tale insieme è generato come una $Po(\lambda * v(I))$.
- Ogni singolo punto è generato in maniera I.I.D da un uniforme $U(A)$, che ha come valore A cioè l'area di I . Da notare che i punti sono casuali non perfettamente equispaziati, infatti se si ha un equidistanza potrebbe esserci correlazione negativa.

8.1.2 Momenti di un processo

- Il momento primo μ è una misura, perciò non ha valori negativi ed è additiva, essa rappresenta il numero medio di punti che cadono in B . $\mu(B) = \lambda * v(B)$.
- Il momento secondo μ è definito come $\mu(B_1, B_2) = \mu(N(B_1) * n(B_2)) - \mu(B_1)\mu(B_2)$ che nel caso del processo di poisson $\mu(N(B_1) * n(B_2)) = \lambda^2 * v(B_1) * v(B_2) + \lambda * v(B_1 \cap B_2)$

8.2 Processo di poisson non omogeneo

Tale processo non ha uniformità di probabilità nell'area dei sottoinsiemi creati, mantiene però la proprietà degli incrementi indipendenti.

Questa volta però il numero di punti in B $N(B)$ si dispone come $Po(\mu(B))$ dove $\mu(B)$ è l'integrale di f_{xy} funzione di probabilità sull'area, se tale funzione fosse costante ho un processo omogeneo.

Ne consegue che i punti non saranno generati da un uniforme ma da una funzione $\frac{f_{xy}}{\mu(B)}$ quindi dipende dall'altezza della funzione in quel punto.

9 Processi spaziali

E' un processo a valori reali $S_x, x \in A$ ed $A \in R^d$, notiamo quindi che il processo non ha un'unica direzione di movimento.

Si definisce processo stocastico spaziale una collezione di variabili casuali definite su uno spazio di probabilità e indicizzate da un parametro X che varia in maniera continua su $A \in R^d$.

Un processo stocastico spaziale è detto **Gaussiano** se le sue distribuzioni marginali finito-dimensionali sono normali, da qua si può intuire che un processo è definito completamente se si identificano:

- $E(S_x) = \mu(X)$ chiamato drift del processo
- $Cov(S_x, S_y) = C(X, Y)$ funzione semi definita positiva

9.1 Stazionarietà e Isotropia

Si ha stazionarietà in senso:

- **Forte** se data una traslazione h la distribuzione di $S_{x_1} \dots S_{x_n} = S_{x_1+h} \dots S_{x_n+h}$.
- **Debole** se vi è invarianza dei primi due momenti rispetto ad una traslazione h cioè $E(S_{x_1}) = E(x_{1+h})$ e $Var(S_{x_1}) = Var(x_{1+h})$ con $Cov(S_{x_1}, S_{x_2}) = Cov(S_{x_1+h}, S_{x_2+h})$ che equivale a dire $= C(X_1 - X_2)$.

Nei processi la stazionarietà forte implica quella debole in quello gaussiano vale anche il viceversa.

Si ha **isotropia** se ho invarianza dei primi due momenti rispetto a rotazioni, traslazioni o compressioni dell'insieme. In questo caso per la covarianza avrò $Cov(S_{x_1}, S_{x_2}) = Cov(S_{x_1+h}, S_{x_2+h}) = C(|X_1 - X_2|)$.

9.2 Metodo Covariogramma e Correlogramma

Se voglio calcolare i momenti di un processo ma non ho abbastanza replicazione posso assumere che tale processo è stazionario debolmente e calcolare:

- Media come $E(S_x) = \mu(X)$ dove ipotizziamo che il processo ruoti attorno alla media come $a + b * X + \delta$, un processo parzialmente stazionario.
- Covarianza so che essendo stazionario debolmente ho $Cov(S_x, S_y) = C(X - Y)$, in questo caso esiste una funzione $C(h)$ ovvero la covarianza tra il vettore $Cov(S_{x_1+h}, S_{x_1})$. La sua rappresentazione grafica è detta **covariogramma** notiamo che è condizione necessaria affinché vi sia stazionarietà debole. Ha come proprietà:
 - $C(0) = Var(S_x) = Var(S_{x+h})$ quindi dalla disuguaglianza di Cauchy-Schwarz sappiamo che $C(h) = C(S_{x_1+h}, S_{x_1}) < C(0)$ quindi $C(0)$ è un punto di massimo.
 - Funzione è simmetrica rispetto all'origine cioè $C(h) = C(-h)$
 - C è definita positiva.
- Correlazione possiamo calcolare tale funzione come $\frac{C(h)}{C(0)}$ e la sua rappresentazione grafica è detta correlogramma, spesso i due grafici si inter-scambiano e viene usato di più quest'ultimo perchè normalizzato.

9.3 Metodo Variogramma

Un processo S ammette variogramma solo se è intrinsecamente stazionario, cioè la variabile casuale varianza delle differenze dipende solo della differenza dei punti, cioè $Var(S_y - S_x) = f(Y - X)$. Esso ha come variogramma $2 * \gamma(h) = Var(S_{x+h} - S_x)$ dove $\gamma(h)$ è detto semi-variogramma, più basso è tale valore più c'è un alta dipendenza tra le due variabili.

Affinché $2\gamma(h)$ sia variogramma deve essere:

- $2\gamma(0) = 0$ cioè ha un minimo nell'origine.
- $\gamma(h) = \gamma(-h)$
- è sempre definita negativa

Un esempio di moto intrinsecamente stazionario è il moto browniano definito come $S_x = N(0, \sigma^2 * x)$ vedo che non è stazionario in senso debole però $Var(S_x - S_y) = f(X - Y)$ cioè dipende solo dalla distanza dei due punti.

Visto che ha media 0 posso dire che $2 * \gamma(h) = E((S_{x+h} - S_x)^2)$.

9.4 Relazione tra variogramma e covariogramma

La nozione di stazionarietà intrinseca è contenuta in quella di stazionarietà debole, ciò implica che se esiste il covariogramma allora esiste anche il variogramma, ma non vale il vice-versa.

Infatti $2 * \gamma(h) = Var(S_{x+h} - S_x) = Var(S_{x+h}) + Var(S_x) - 2Cov(S_{x+h}, S_x) = 2C(0) - 2C(h)$, notiamo quindi che se ho il covariogramma posso derivare il variogramma, vice-versa non è possibile.

E' possibile dimostrare che se $\lim_{h \rightarrow \infty} C(h) = 0$ quindi $\lim_{h \rightarrow \infty} \gamma(h) = C(0)$ da qua trovo che $C(h) = \lim_{h \rightarrow \infty} \gamma(h) - \gamma(h)$.

Se ho una processo isotropico vale $\gamma(h) = \gamma(|h|)$, in questo caso si dice che il variogramma è omogeneo. C'è una relazione inversa tra variogramma e covariogramma.

9.5 Caratteristiche del semi-variogramma e covariogramma/correlogramma

Il semi-variogramma è una funzione crescente di h per cui il covariogramma dovrà essere decrescente. So che hanno rispettivamente minimo e massimo in 0.

Il semi-variogramma può crescere avendo un asintoto orizzontale, e in questo caso il covariogramma decrescerà con asintoto orizzontale lo 0. Ne consegue che c'è sempre correlazione tra due elementi ma che tende a 0, in maniera operativa prendiamo un livello α dove diciamo che non c'è più correlazione. Se invece il semi-variogramma cresce e poi si trasforma in una retta ho che il covariogramma tocca lo 0 e si ferma.

9.6 Processo spaziale gaussiano isotropico

Tale processo può avere:

- Continuità: è una continuità in media quadratica, un processo è continuo in X_0 se $\lim_{x \rightarrow x_0} E((S_x - S_{x_0})^2) = 0$ cioè se $S_x \rightarrow S_{x_0}$ in media quadratica. Se il processo stocastico è limitato allora la continuità normale include la continuità in media quadratica.

Se il processo è stazionario posso dire che $\lim_{x \rightarrow x_0} E((S_x - S_{x_0})^2) = 2\gamma(h) = 2(C(0) - C(h))$, dunque si ha continuità in media quadratica se e solo se il variogramma è continuo nell'origine, dunque l'effetto pepita crea traiettorie discontinue.

- Differenziabilità: cioè se è vera la relazione $\lim_{h \rightarrow 0} (\frac{S_{x+h} - S_x}{h} - S'_x)^2 = 0$. Se S è stazionario esiste derivata seconda in $C(0)$ allora S è differenziabile una volta. S è differenziabile n volte se e solo se esiste derivata $2n$ nell'origine del covariogramma.

9.6.1 Effetto pepita

Se $\lim_{h \rightarrow 0} \gamma(h) = \alpha$ che è diverso da 0 allora ho l'effetto pepita, cioè una discontinuità nell'origine del semivariogramma.

Se presente tale effetto ho discontinuità nelle osservazioni, tali nella realtà sono definiti errori ed è necessario modellarli. Avremmo quindi che il nostro fenomeno è modellato da:

- Modello S_x quello ufficiale che modella i dati con covariogramma C_s variogramma γ_s e correlogramma ρ_s esso è continuo.
- Modello W_x che modella gli errori ha momento prima uguale a 0, una variabilità finita $Var(W_x) = \tau^2$ e la covarianza tra gli errori nulla. Vediamo che tale modello è stazionario debolmente.

Con questi due processi posso creare quello che descrive la realtà come $R=S+W$ esso ha:

- Valore atteso uguale a quello di S.
- Varianza uguale alla somma delle due varianze.
- Covarianza uguale alla covarianza di S.
- Correlazione 1 se $h = 0$ e se $h \neq 0$ è $\frac{C_s(h)}{\tau^2 + C(0)}$, vediamo quindi che c'è discontinuità.
- Variogramma è scritto come $2\gamma_s(h) + \tau^2$ noto quindi che è discontinuo nell'origine e l'effetto pepita è dato da τ^2 .

Notiamo quindi la facilità di costruire un modello discontinuo partendo da uno continuo.

9.7 Principali modelli isotropici

- Modello esponenziale ha come correlogramma la funzione $\rho(|h|) = e^{-\frac{|h|}{r}}$ quindi ha asintoto orizzontale lo 0; variogramma la funzione $\gamma(|h|) = C(0) * (1 - e^{-\frac{|h|}{r}})$ ed ha asintoto orizzontale $C(0)$. Il parametro r indica l'inclinazione della curva ed è detto parametro di scala. Le traiettorie sono continue ma non derivabili essendoci il modulo.
- Modello gaussiano ha funzione di correlazione $\rho(|h|) = e^{-\frac{|h|^2}{r}}$ ciò lo rende continuo e derivabile n volte.
- Modello esponenziale potenziato $\rho(|h|) = e^{-\frac{|h|^k}{r}}$ dove k sta tra 0 e 2 ed è parametro di forma.
- Modello sferico $p(|h|) = 0$ se $|h| > r$ se no $p(|h|) = 1 - \frac{3|h|}{4r} + \frac{|h|^3}{2r^3}$, vediamo quindi che è limitata cioè arriva veramente a 0 e non asintoticamente. E' continuo ma non derivabile.
- Modllo di Matérn ha funzione di correlazione $\rho(|h|) = \frac{(\frac{|h|}{r})^k}{2^{k-1}\Gamma(k)} * B_k(\frac{|h|}{r})$. Se il parametro $K=1/2$ otteniamo il modello esponenziale se $K = \infty$ otteniamo quello gaussiano. P è sempre continua ed è differenziabile in media quadratica m volte se $K > m$, perciò più K è grande più è regolare la traiettoria.