

Appunti del corso Processi stocastici

Marco Frego,
Marco Pizzato,
Luca Tasin,
Luciano Tubaro

anno accademico 2007-2008

Indice

1 Introduzione ai Processi Stocastici	5
1 Nota Introduttiva	5
2 Processi Stocastici	5
3 Classi Monotone	7
4 Probabilità Condizionata	8
5 Teoremi di Kolmogorov	13
2 Processi Stocastici	17
1 Processi Gaussiani	17
2 Processi di Markov	18
3 Relazione di Chapman-Kolmogorov per Processi di Markov	19
4 Processo di Wiener	21
4.1 Distribuzione del Processo di Wiener	22
4.2 Wiener è un Processo di Markov	23
4.3 Costruzione del Processo di Wiener	25
5 Brownian Bridge	26
5.1 Transizione nel Brownian Bridge	26
5.2 Costruzione del Brownian Bridge	28
5.3 Moto Browniano	32
6 Martingale	32
3 Equazioni Differenziali Stocastiche	37
1 Brownian Bridge da un'Equazione Differenziale Stocastica	37
2 Brownian Bridge Condizionando un Processo di Wiener	38
3 Alcune Equazioni Differenziali Stocastiche	39
4 Semigruppo e Integrale	41
1 Semigruppo di Operatori Lineari	41
2 Richiami di Alcuni Tipi di Integrazione	45
2.1 Integrale di Riemann-Stieltjes	45
2.2 Integrale di Lebesgue-Stieltjes	47
3 Integrazione Rispetto ad un Processo di Wiener	47
3.1 Il Caso più Semplice	47
3.2 Un Caso Particolare di Integrale di Itô	48
4 Un Caso Particolare di Meyer	50

4 Appunti del corso

Capitolo 1

Introduzione ai Processi Stocastici

1 Nota Introduttiva

Πάντα ḥεῖ, οὐδὲν μένει

Tutto scorre, nulla sta fermo
Ἡράκλειτος ὁ σκοτεινός (ca. 535 - 475 a.C.)

La Probabilità si occupa di studiare quei fenomeni che avvengono spesso in tempi molto brevi e coinvolgono un così grande numero di eventi che non si possono investigare singolarmente. Fin dall'antichità questa attenzione è stata rivolta ai fenomeni naturali, che hanno portato poi all'astrazione e ai modelli matematici successivi che meglio si prestavano alla loro descrizione. Solo verso la fine del XIX secolo il matematico russo Andrey Markov (1856-1922) cominciò a dare rigore formale alle osservazioni dei mutamenti casuali che venivano osservati. L'esempio più famoso a riguardo è senz'altro rappresentato dal *moto browniano*, fenomeno che coinvolge il movimento delle particelle microscopiche del polline all'interno di un liquido: benché esso porti il nome del botanico Robert Brown (1773-1858) che per primo lo descrisse, fu modellizzato solo successivamente da Albert Einstein e altri nel 1905. Tuttavia solo grazie alla teoria assiomatica, introdotta da Andrey Kolmogorov (1903-1987) intorno al 1930 si può guardare oggi al moto browniano come ad un processo stocastico a tempi continui: il processo di Wiener. Lo studio dei processi stocastici è alla base di molte applicazioni e coinvolge molte branche della matematica come l'Analisi, la Topologia, la Teoria della Misura.

L'aggettivo *stocastico* deriva dall'aggettivo *στοχαστικός* \prec *στοχάζομαι* che gli antichi greci usavano col significato di *sono abile a mirare nel lancio* e quindi più figuratamente *ho in vista qualcosa, congetturo*.

2 Processi Stocastici

Definizione 1.1. *Un processo stocastico è una famiglia di variabili casuali $\{X_t, t \in [0, T] \subset \mathbb{R}\}$ definite tutte su uno stesso spazio di probabilità $(\Omega, \mathcal{E}, \mathbb{P})$.*

6 Appunti del corso

Ad ogni processo stocastico possiamo associare la famiglia \mathcal{M} di tutte le distribuzioni di probabilità congiunte (finite)

$$\mu_{t_1, t_2, \dots, t_n}(B) = \mathbb{P}((X_{t_1}, X_{t_2}, \dots, X_{t_n}) \in B)$$

al variare di t_1, t_2, \dots, t_n e di n .

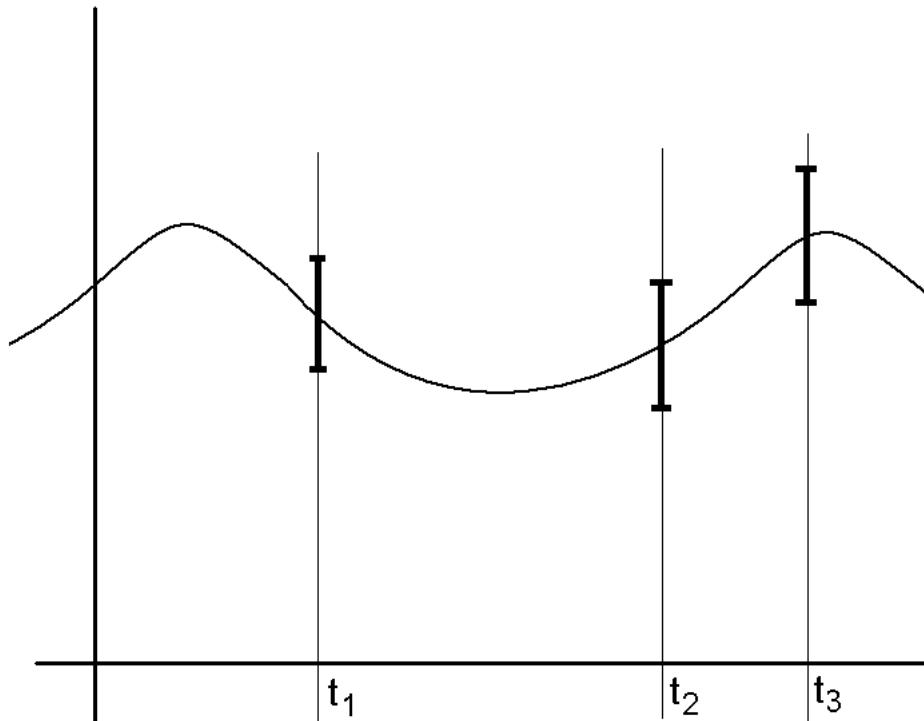


Figura 1.1: Insiemi cilindrici

Introduciamo lo spazio vettoriale reale $\mathbb{R}^{[0,T]}$, formato da tutte le funzioni $q: [0, T] \rightarrow \mathbb{R}$.

Definizione 1.2. *Fissati n valori $t_1 < t_2 < \dots < t_n$ in $[0, T]$ e fissato inoltre un boreiano $B \subset \mathbb{R}^n$, chiameremo insieme cilindrico il sottoinsieme di $\mathbb{R}^{[0,T]}$ definito da*

$$C_{t_1, t_2, \dots, t_n, B} = \{q \in \mathbb{R}^{[0,T]} : (q(t_1), q(t_2), \dots, q(t_n)) \in B\}.$$

Lemma 1.3. *La famiglia di tutti gli insiemi cilindrici, \mathcal{C} , forma un'algebra. Denotiamo con $\sigma(\mathcal{C})$ la più piccola σ -algebra contenente l'algebra \mathcal{C} .*

Ripetiamo ora la stessa operazione sostituendo lo spazio $\mathbb{R}^{[0,T]}$ con lo spazio $C([0, T], \mathbb{R})$ (che è un sottospazio di $\mathbb{R}^{[0,T]}$). In altre parole, introduciamo lo spazio (vettoriale su \mathbb{R}) $C([0, T], \mathbb{R})$, formato da tutte le funzioni continue $q: [0, T] \rightarrow \mathbb{R}$.

Definizione 1.4. Fissati n valori $t_1 < t_2 < \dots < t_n$ in $[0, T]$ e fissato inoltre un boreliano $B \subset \mathbb{R}^n$, chiameremo insieme cilindrico il sottoinsieme di $C([0, T], \mathbb{R})$ definito da

$$C_{t_1, t_2, \dots, t_n, B}^{(\mathcal{C})} = \{q \in C([0, T], \mathbb{R}) : (q(t_1), q(t_2), \dots, q(t_n)) \in B\}.$$

Lemma 1.5. La famiglia di tutti gli insiemi cilindrici, \mathcal{C}_c , forma un'algebra. Denotiamo con $\sigma(\mathcal{C}_c)$ la più piccola σ -algebra contenente l'algebra \mathcal{C}_c .

Abbiamo il seguente risultato: se pensiamo $C([0, T], \mathbb{R})$ con la topologia della convergenza uniforme, allora

$$\sigma(\mathcal{C}_c) \equiv \mathcal{B}$$

dove \mathcal{B} è la σ -algebra di Borel di $C([0, T], \mathbb{R})$.

Definizione 1.6. Per ogni $\omega \in \Omega$ possiamo considerare la traiettoria $t \in X_t(\omega)$ come una funzione $q \in \mathbb{R}^{[0, T]}$, cioè $q(t) = X_t(\omega)$. Al variare di $\omega \in \Omega$, abbiamo una famiglia di funzioni $q \in \mathbb{R}^{[0, T]}$. Quindi possiamo definire una mappa, che chiameremo X , definita da

$$X : \Omega \rightarrow \mathbb{R}^{[0, T]}, \quad X(\omega) = q$$

dove q è data da $q(t) = X_t(\omega)$.

Ebbene risulta che la mappa X è misurabile come mappa da $(\Omega, \mathcal{E}) \rightarrow (\mathbb{R}^{[0, T]}, \sigma(\mathcal{C}))$.

Possiamo allora considerare la misura immagine di \mathbb{P} , e la denoteremo con \mathbb{Q}_X , definita al solito come

$$\mathbb{Q}_X(B) = \mathbb{P}(X^{-1}(B)), \quad B \in \sigma(\mathcal{C})$$

Ad ogni processo stocastico $\{X_t, t \in [0, T]\}$ abbiamo associato allora un nuovo spazio di probabilità

$$(\mathbb{R}^{[0, T]}, \sigma(\mathcal{C}), \mathbb{Q}_X)$$

Abbiamo il risultato che se due processi stocastici X ed X' , eventualmente definiti anche su spazi di probabilità diversi, con l'ipotesi che ammettano la stessa famiglia \mathcal{M} , cioè

$$\mathbb{P}((X_{t_1}, X_{t_2}, \dots, X_{t_n}) \in B) = \mathbb{P}'((X'_{t_1}, X'_{t_2}, \dots, X'_{t_n}) \in B),$$

allora abbiamo anche che $\mathbb{Q}_X \equiv \mathbb{Q}_{X'}$. Quindi la \mathbb{Q}_X in ultima analisi è determinata solo da \mathcal{M} : potremo scrivere $\mathbb{Q}_{\mathcal{M}}$.

Potremo allora considerare un altro processo stocastico \dot{X} , definito su $(\mathbb{R}^{[0, T]}, \sigma(\mathcal{C}), \mathbb{Q}_{\mathcal{M}})$, da

$$\dot{X}_t(q) = q(t)$$

che ha esattamente la famiglia \mathcal{M} come distribuzioni congiunte.

3 Classi Monotone

Definizione 1.7. Definiamo π -sistema una famiglia \mathcal{P} di sottoinsiemi di Ω tale che

$$A, B \in \mathcal{P} \Rightarrow A \cap B \in \mathcal{P}.$$

8 Appunti del corso

Definizione 1.8. Definiamo *l-sistema* una famiglia \mathcal{L} di sottoinsiemi di Ω tale che

- 1) Se $A_1 \subset A_2 \subset A_3 \subset \dots$ con ogni $A_i \in \mathcal{L}$ allora $\bigcup A_i \in \mathcal{L}$
- 2) $A \subset B$ entrambi $\in \mathcal{L} \Rightarrow B \setminus A \in \mathcal{L}$

Data una famiglia \mathcal{A} di sottoinsiemi di Ω , denoteremo con $\sigma(\mathcal{A})$ la σ -algebra generata da \mathcal{A} .

Proposizione 1.9. Un *l-sistema* \mathcal{L} , tale che $\Omega \in \mathcal{L}$ e che per ogni $A, B \in \mathcal{L}$ anche $A \cap B \in \mathcal{L}$, è una σ -algebra.

Proposizione 1.10. Dati un π -sistema \mathcal{P} e un *l-sistema* \mathcal{L} , tali che $\Omega \in \mathcal{L}$ e $\mathcal{P} \subset \mathcal{L}$, allora $\mathcal{P} \subset \sigma(\mathcal{P}) \subset \mathcal{L}$.

Dati un π -sistema \mathcal{P} e un *l-sistema* \mathcal{L} tali che $\mathcal{P} \subset \mathcal{L}$ e $\Omega \in \mathcal{L}$, definiamo

$$\mathcal{L}_1 = \{E \in \mathcal{L} : \forall C \in \mathcal{P}, \quad E \cap C \in \mathcal{L}\}$$

$$\mathcal{L}_2 = \{E \in \mathcal{L}_1 : \forall A \in \mathcal{L}_1, \quad E \cap A \in \mathcal{L}\}$$

Lemma 1.11. Risulta $\mathcal{P} \subset \mathcal{L}_2 \subset \mathcal{L}_1 \subset \mathcal{L}$; inoltre \mathcal{L}_1 ed \mathcal{L}_2 sono *l-sistemi*.

Corollario 1.12. Sia $\tilde{\mathcal{L}}$ il più piccolo *l-sistema* tale che $\Omega \in \tilde{\mathcal{L}}$; allora $\tilde{\mathcal{L}}_2 \equiv \tilde{\mathcal{L}}_1 \equiv \tilde{\mathcal{L}}$. Risulta in altre parole che $A \cap B \in \tilde{\mathcal{L}}$ per ogni $A, B \in \tilde{\mathcal{L}}$, e quindi, per la proposizione 1.9, $\tilde{\mathcal{L}}$ risulta una σ -algebra.

Quindi la proposizione 1.10 è dimostrata in quanto $\mathcal{P} \subset \sigma(\mathcal{P}) \subset \tilde{\mathcal{L}} \subset \mathcal{L}$.

Teorema 1.13. Se due misure di probabilità μ_1 e μ_2 coincidono su un π -sistema \mathcal{P} allora coincidono sulla σ -algebra $\sigma(\mathcal{P})$ generata da \mathcal{P} .

4 Probabilità Condizionata

Teorema 1.14 (teorema di Radon-Nikodym). Sia \mathcal{A} una σ -algebra su X e siano $\alpha, \nu : \mathcal{A} \rightarrow \mathbb{R}$ misure σ -additive a valori reali. Se α è positiva ed assolutamente continua rispetto a ν , allora esiste una funzione $g \in L^1(\alpha)$ detta densità o derivata di ν rispetto ad α tale che

$$\nu(A) = \int_A g \, d\alpha$$

per ogni $A \in \mathcal{A}$. Inoltre ogni altra funzione di densità è quasi ovunque uguale a g .

Teorema 1.15 (teorema di Beppo Levi o convergenza dominata). Sia $\{f_k\}$ una successione di funzioni misurabili non negative, $f_k : \mathbb{R}^n \rightarrow [0, +\infty]$ e supponiamo che la successione sia anche crescente, allora se $f(x) = \lim_{k \rightarrow \infty} f_k(x)$ si ha

$$\int_{\mathbb{R}^n} f(x) \, d\mu(x) = \lim_{k \rightarrow \infty} \int_{\mathbb{R}^n} f_k(x) \, d\mu(x)$$

Sia $(\Omega, \mathcal{E}, \mathbb{P})$ il nostro spazio di probabilità.

Lemma 1.16. *Siano $\mathcal{F} \subseteq \mathcal{E}$ una σ -algebra e $A \subseteq \Omega$. Esiste un'unica funzione $Y : \Omega \rightarrow \mathbb{R}$ \mathcal{F} -misurabile tale che per ogni $F \subseteq \mathcal{F}$*

$$\mathbb{P}(A \cap F) = \int_F Y(\omega) \mathbb{P}(d\omega).$$

Dimostrazione. La chiave di volta è che la funzione $F \mapsto \mathbb{P}(A \cap F)$ è una misura, come è facile controllare. Inoltre è assolutamente continua rispetto a \mathbb{P} , dunque per Radon-Nikodym esiste Y tale che

$$\mathbb{P}(A \cap F) = \int_F Y d\mathbb{P}.$$

Infine

$$\int_F Y d\mathbb{P} = \int_F Y' d\mathbb{P} \Rightarrow \int_F (Y - Y') d\mathbb{P} = 0 \quad \text{per ogni } F \subseteq \Omega$$

da cui segue che $Y = Y'$ quasi ovunque rispetto a \mathbb{P} . Qui è decisivo che Y e Y' siano \mathcal{F} -misurabili. ■

Definizione 1.17. *Chiamiamo probabilità condizionata di A rispetto ad \mathcal{F} , indicandola con $\mathbb{P}(A|\mathcal{F})$, l'unica Y tale che*

$$\mathbb{P}(A \cap F) = \int_F Y(\omega) \mathbb{P}(d\omega).$$

Analogamente

Lemma 1.18. *Data una variabile casuale $X \in L^1(\Omega)$ esiste un'unica funzione Y \mathcal{F} -misurabile tale che per ogni $F \subseteq \mathcal{F}$*

$$\int_F X(\omega) \mathbb{P}(d\omega) = \int_F Y(\omega) \mathbb{P}(d\omega).$$

Definizione 1.19 (aspettazione condizionata). *Sia X una variabile casuale in $L^1(\Omega)$. Sia Y una variabile casuale \mathcal{F} -misurabile tale che*

$$\int_F X(\omega) \mathbb{P}(d\omega) = \int_F Y(\omega) \mathbb{P}(d\omega), \quad \text{per ogni } F \in \mathcal{F}$$

L'esistenza di una tale funzione si può stabilire tramite il teorema di Radon-Nikodym. Inoltre essa è unica a meno di insiemi di misura nulla.

Denoteremo tale funzione con $\mathbb{E}(X|\mathcal{F})(\omega)$, spesso omettendo di scrivere esplicitamente la variabile ω .

Lemma 1.20. *Proprietà dell'aspettazione condizionata.*

1. Se X è essa stessa \mathcal{F} -misurabile allora $\mathbb{E}(X|\mathcal{F}) = X$ quasi ovunque.
2. $\mathbb{E}(cX_1 + X_2|\mathcal{F}) = c \mathbb{E}(X_1|\mathcal{F}) + \mathbb{E}(X_2|\mathcal{F})$ quasi ovunque.

Nel caso in cui sceglieremo $X = \mathbf{1}_A$ useremo anche la notazione

$$\mathbb{P}(A|\mathcal{F}) := \mathbb{E}(\mathbf{1}_A|\mathcal{F})$$

10 Appunti del corso

Proposizione 1. *Se risulta*

$$\int_F X(\omega) \mathbb{P}(d\omega) = \int_F Y(\omega) \mathbb{P}(d\omega), \quad \text{per ogni } F \in \mathcal{P}$$

dove \mathcal{P} è un qualunque π -sistema tale che $\sigma(\mathcal{P}) = \mathcal{F}$. Allora l'eguaglianza è vera anche per tutti gli $F \in \mathcal{F}$.

Sia Y un vettore casuale n -dimensionale, $Y = (Y_1, Y_2, \dots, Y_n)$. Consideriamo come σ -algebra \mathcal{F} la σ -algebra generata da Y , cioè

$$\mathcal{F} = \sigma(Y) = \sigma(Y_1, Y_2, \dots, Y_n)$$

oppure la σ -algebra generata dal π -sistema \mathcal{P} formato da tutte le intersezioni

$$Y_1^{-1}(I_1) \cap Y_2^{-1}(I_2) \cap Y_3^{-1}(I_3) \cap \dots \cap Y_n^{-1}(I_n).$$

Se la variabile casuale X è \mathcal{F} -misurabile allora esiste una funzione boreiana $\varphi(y) = \varphi(y_1, y_2, \dots, y_n)$ tale che

$$X = \varphi(Y) = \varphi(Y_1, Y_2, \dots, Y_n), \quad \text{quasi ovunque}$$

e questa funzione è unica a meno di insiemi di misura nulla rispetto alla misura, distribuzione di probabilità di Y ,

$$\mu(B) = \mathbb{P}(Y \in B).$$

Con tale σ -algebra \mathcal{F} useremo la notazione

$$\mathbb{E}(X | Y) := \mathbb{E}(X | \mathcal{F})$$

alla quale possiamo associare una φ tale che $\mathbb{E}(X | Y) = \varphi(Y)$. Useremo anche l'ulteriore notazione

$$\mathbb{E}(X | Y = y) = \mathbb{E}(X | Y_1 = y_1, Y_2 = y_2, \dots, Y_n = y_n) := \varphi(y) = \varphi(y_1, y_2, \dots, y_n)$$

Nel caso particolare in cui come variabile X scegliamo la variabile

$$\mathbf{1}_{X \in I} \equiv \mathbf{1}_I(X)$$

scrivremo anche

$$\mathbb{P}(X \in I | Y = y) := \mathbb{E}(\mathbf{1}_I(X) | Y = y)$$

con l'identità

$$\mathbb{P}(X \in I, Y \in B) = \int_B \mathbb{P}(X \in I | Y = y) \mathbb{P}(Y \in dy).$$

Definizione 1.21. Chiamiamo valore di aspettazione condizionato di X rispetto ad \mathcal{F} , indicandolo con $\mathbb{E}(X | \mathcal{F})$, l'unica Y tale che

$$\int_F X(w) \mathbb{P}(dw) = \int_F Y(w) \mathbb{P}(dw).$$

Dimostrazione. Del tutto simile a quella del lemma precedente, considerando la misura $F \mapsto \int_F X d\mathbb{P}$. ■

Vogliamo ora comprendere meglio gli oggetti appena definiti, per cominciare notiamo che $\mathbb{E}(\mathbf{1}_A | \mathcal{F}) = \mathbb{P}(A | \mathcal{F})$, cioè l'aspettazione condizionata di una funzione indicatrice è la probabilità condizionata del suo insieme, dunque la seconda definizione generalizza la prima.

Esempio 1. Sia $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Ovviamente c'è solo una condizione da controllare, in quanto integrare sull'insieme vuoto non porta a nulla:

$$\int_{\Omega} \mathbb{E}(X | \mathcal{F}_0) d\mathbb{P} = \int_{\Omega} X d\mathbb{P} = \mathbb{E}(X).$$

Ma $\mathbb{E}(X | \mathcal{F}_0)$ deve essere \mathcal{F}_0 -misurabile, quindi costante, portandola fuori dal segno di integrale concludiamo che $\mathbb{E}(X | \mathcal{F}_0) = \mathbb{E}(X)$, cioè il buon vecchio valore di aspettazione.

Esempio 2. Consideriamo ora $\mathcal{F}_1 = \{\emptyset, E, E^c, \Omega\}$ con $E \subseteq \Omega$. Dire che una variabile casuale è \mathcal{F}_1 -misurabile equivale a dire che è costante su E e su E^c . Supponiamo $0 < \mathbb{P}(E) < 1$.

$$\int_E X d\mathbb{P} = c \int_E d\mathbb{P} = c\mathbb{P}(E).$$

Dunque

$$\mathbb{E}(X | \mathcal{F}_1)(\omega) = \frac{1}{\mathbb{P}(E)} \int_E X d\mathbb{P} \quad \text{se } \omega \in E$$

e simmetricamente per $\omega \in E^c$

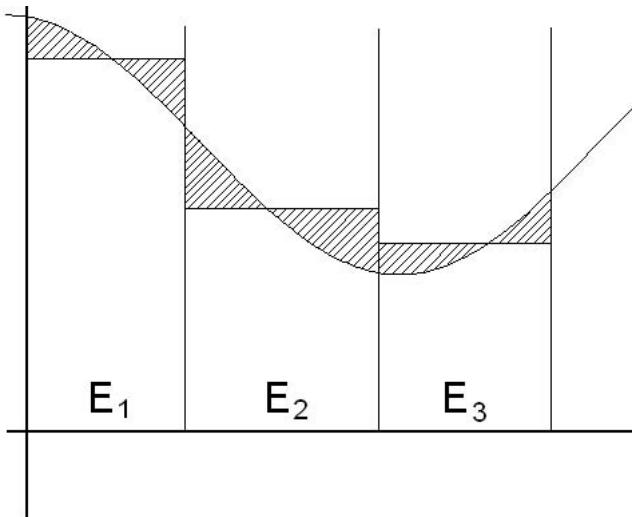


Figura 1.2: Generalizzazione dell'esempio

L'esempio si generalizza senza colpo ferire al caso di una σ -algebra $\mathcal{F} = \{\emptyset, E_1, E_2, \dots, E_n, \Omega, \dots\}$. Anche in questo caso, il risultato sarà di rendere costante la X su ognuno degli E_i , prendendone la media.

12 Appunti del corso

Data una variabile casuale X possiamo considerare la σ -algebra degli insiemi misurabili secondo X , poniamo $\mathcal{F}_X = \{X^{-1}(I) : I \in \mathcal{B}\}$. A questo punto diremo che X e \mathcal{F} sono indipendenti se lo sono \mathcal{F}_X e \mathcal{F} , cioè se per ogni boreiano I di \mathbb{R} ed ogni $F \in \mathcal{F}$, $\mathbb{P}(X^{-1}(I) \cap F) = \mathbb{P}(X^{-1}(I))\mathbb{P}(F)$.

La prossima proposizione raccoglie alcune utili proprietà dell'aspettazione condizionata che in seguito utilizzeremo.

Proposizione 1.22. *Siano $X, Y \in L^1(\Omega)$ e $a \in \mathbb{R}$.*

- i) *X è \mathcal{F} -misurabile se e solo se $\mathbb{E}(X|\mathcal{F}) = X$.*
- ii) *$\mathbb{E}(aX + Y|\mathcal{F}) = a\mathbb{E}(X|\mathcal{F}) + \mathbb{E}(Y|\mathcal{F})$.*
- iii) *Se X e \mathcal{F} sono indipendenti allora $\mathbb{E}(X|\mathcal{F}) = \mathbb{E}(X)$.*

Dimostrazione. Le prime due affermazioni sono immediate. Mostriamo iii). Procediamo in modo standard, analizzando per prima cosa il caso di una funzione indicatrice $X = \mathbf{1}_A$. Abbiamo $\mathcal{F} = \{\Omega, A, A^c, \emptyset\}$ e l'essere indipendente è equivalente ad avere $\mathbb{P}(A \cap F) = \mathbb{P}(A)\mathbb{P}(F)$ per ogni $F \subseteq \mathcal{F}$.

$$\int_F \mathbb{E}(X|\mathcal{F}) d\mathbb{P} = \int_F \mathbf{1}_A d\mathbb{P} = \mathbb{P}(A \cap F) = \mathbb{P}(A)\mathbb{P}(F) = \int_F \mathbb{P}(A) d\mathbb{P}.$$

Chiaramente $\mathbb{P}(A)$ come variabile casuale costante è sempre misurabile, dunque $\mathbb{E}(X|\mathcal{F}) = \mathbb{P}(A)$ quasi ovunque rispetto a \mathbb{P} e ovviamente $\mathbb{E}(X) = \mathbb{P}(A)$.

Il caso in cui $X = \sum_i c_i \mathbf{1}_{A_i}$ è una funzione semplice, è subito vinto per linearità, poi si prosegue approssimando da sotto con una successione di funzioni semplici positive (indipendenti rispetto a \mathcal{F} !) passando al limite con Beppo-Levi e si conclude scomponendo X qualsiasi in parte positiva e parte negativa. ■

Sia $Y = (Y_1, \dots, Y_n)$ un vettore aleatorio, di nuovo possiamo considerare la più piccola σ -algebra di insiemi misurabili secondo Y , abbiamo

$$\mathcal{F}_Y = \sigma\{Y_1^{-1}(I_1) \cap Y_2^{-1}(I_2) \cap \dots \cap Y_n^{-1}(I_n)\} \subseteq \mathcal{E}.$$

Per proseguire abbiamo bisogno del seguente importante

Teorema 1.23. *Siano $Y = (Y_1, \dots, Y_n)$ e X variabili casuali. Allora X è \mathcal{F}_Y -misurabile se e solo se esiste $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ boreiana tale che $X = \phi(Y)$. Tale ϕ è unica a meno di insiemi di misura nulla rispetto alla misura indotta da Y .*

Dimostrazione. La struttura della dimostrazione è la stessa di quella della proposizione: esercizio! ■

Possiamo scrivere, più suggestivamente, $\mathbb{E}(X|Y) := \mathbb{E}(X|\mathcal{F}_Y)$ la quale per definizione è \mathcal{F}_Y -misurabile.

Definizione 1.24. *Siano $Y = (Y_1, \dots, Y_n)$ e X variabili casuali. Sia ϕ tale che $X = \phi(Y)$, allora poniamo*

$$\mathbb{E}(X|Y = y) := \phi(y).$$

Notiamo che $\mathbb{P}(X \in I|Y) = \mathbb{E}(\mathbf{1}_I(X)|Y)$, dunque esiste ψ boreiana tale che $\mathbb{P}(X \in I|Y) = \psi(Y)$ e quindi

$$\begin{aligned}\mathbb{P}(X \in I, Y \in B) &= \int_{Y^{-1}(B)} \mathbb{P}(X \in I|Y) d\mathbb{P} = \int_{Y^{-1}(B)} \psi(Y) d\mathbb{P} = \\ &= \int_B \psi(y) \mathbb{P}(Y \in dy) = \int_B \mathbb{P}(X \in I|Y = y) \mathbb{P}(Y \in dy).\end{aligned}$$

Rigore è così fatto nella fitta giungla della probabilità condizionata; forse.

Data una successione di eventi A_1, A_2, A_3, \dots , si definisce $\limsup A_k$ l'evento

$$\limsup A_k = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

Lemma 1.25 (Borel-Cantelli). *Sia A_1, A_2, A_3, \dots una successione di eventi, tali che*

$$\sum_{k=1}^{\infty} \mathbb{P}(A_k) < \infty,$$

allora l'evento $A = \limsup A_k$ ha probabilità nulla, $\mathbb{P}(A) = 0$. Naturalmente si ha che $\mathbb{P}(A^c) = 1$ e $A^c = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c$.

Dimostrazione. Dalla definizione abbiamo che $A \subset \bigcup_{k=n}^{\infty} A_k$. Quindi

$$\mathbb{P}(A) \leq \sum_{k=n}^{\infty} \mathbb{P}(A_k)$$

che per ipotesi tende a zero quando $n \rightarrow \infty$. ■

5 Teoremi di Kolmogorov

Teorema 1.26 (I Teorema di Kolmogorov). *Data una famiglia \mathcal{M} , esiste un'unica misura immagine \mathbb{Q}_{μ} su $(\mathbb{R}^{[0,T]}, \sigma(\mathcal{C}), \mathbb{Q}_{\mu})$.*

Benchè questo teorema garantisca l'esistenza di quello che stiamo cercando, non ci dice nulla su come effettuare la nostra ricerca, cioè non è un teorema costruttivo. Per questo faremo riferimento al prossimo, quando parleremo genericamente di teorema di Kolmogorov.

Teorema 1.27 (II Teorema di Kolmogorov). *Sia $(X_t)_t, t \in [0, T]$ un processo stocastico tale che valga la seguente ipotesi: $\mathbb{E}(|X_t - X_s|^{\alpha}) \leq c|t - s|^{1+\beta}$, per qualche $\alpha, \beta, c > 0$. Allora esiste una versione X'_t di X_t , tale che le sue traiettorie sono quasi ovunque hölderiane di esponente $\gamma < \frac{\beta}{\alpha}$.*

14 Appunti del corso

Dimostrazione. Usando la diseguaglianza di Chebychev otteniamo che

$$\mathbb{P}(|X_t - X_s| > \varepsilon) \leq \frac{1}{\varepsilon^\alpha} \mathbb{E}(|X_t - X_s|^\alpha) \leq \frac{c}{\varepsilon^\alpha} |t - s|^{1+\beta}.$$

Introduciamo il concetto di numeri *diadi* che ci permetteranno di continuare la nostra dimostrazione.

Sia $D_m := \{\frac{k}{2^m} : k = 0, \dots, 2^m\}$. Sia quindi l'insieme dei diadi $D = \bigcup_{m=0}^{\infty} D_m$.

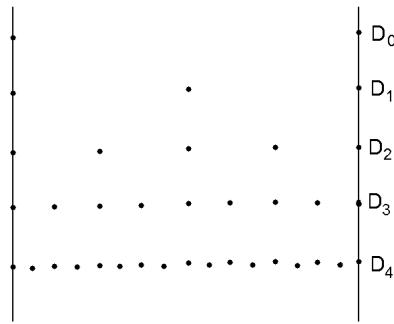


Figura 1.3: Costruzione dei diadi

Cerchiamo ora di trovare qualche stima utile valutando la funzione sui diadi.

$$\mathbb{P}(|X_{t_{i+1}} - X_{t_i}| > \varepsilon_m) \leq \frac{c}{\varepsilon_m^\alpha} \frac{1}{2^{m(1+\beta)}}, t_i \in D_m.$$

Consideriamo ora l'evento $A = (\max_i |X_{t_{i+1}} - X_{t_i}| > \varepsilon_m)$. Appare quindi evidente (esercizio) che

$$A \subset \bigcup_{i=0}^{2^m-1} A_i \text{ ove } A_i = (|X_{t_{i+1}} - X_{t_i}| > \varepsilon_m).$$

Quindi sfruttando le proprietà della misura di probabilità e le stime ottenute precedentemente otteniamo che

$$\mathbb{P}(\max_i |X_{t_{i+1}} - X_{t_i}| > \varepsilon_m) \leq 2^m \frac{c}{\varepsilon_m^\alpha} \frac{1}{2^{m(1+\beta)}} = \frac{c}{\varepsilon_m^\alpha 2^{m\beta}}.$$

Sia ora $A_m = (\max_i |X_{t_{i+1}} - X_{t_i}| > \varepsilon_m)$. E sia inoltre $\varepsilon_m = \frac{1}{2^{\gamma m}}$. Con queste scelte otteniamo

$$\sum_{m=0}^{\infty} \mathbb{P}(A_m) \leq c \sum_{m=0}^{\infty} \frac{2^{m\gamma\alpha}}{2^{m\beta}} = c \sum_{m=0}^{\infty} \frac{1}{2^{m(\beta-\gamma\alpha)}}.$$

Scegliendo $\gamma < \frac{\beta}{\alpha}$ la serie converge.

Consideriamo ora l'insieme $A = \limsup A_m$. Per il lemma di Borel-Cantelli $\mathbb{P}(A) = 0$.

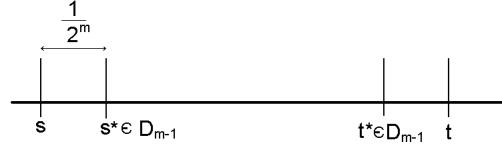


Figura 1.4: Suddivisione degli intervalli

Sia ora $\omega \notin A$. Questo implica che $\omega \in \bigcup_{m=1}^{\infty} \bigcap_{k=m}^{\infty} A_k^c$. Segue che $\exists m^*$ tale che $\omega \in \bigcap_{k=m^*}^{\infty} A_k^c$. E quindi $\omega \in A_k^c \quad \forall k \geq m^*$.

$$\text{Ne segue che } \max_i |X_{t_{i+1}}(\omega) - X_{t_i}(\omega)| \leq \frac{1}{2^{m\gamma}} \quad \forall m \geq m^*.$$

Riassumendo abbiamo costruito un insieme A di misura nulla tale che

$$\omega \notin A \Rightarrow \exists m^* \text{ con } \max_i |X_{t_{i+1}}(\omega) - X_{t_i}(\omega)| \leq \frac{1}{2^{m\gamma}} \quad \forall m \geq m^*.$$

Quello che vorremmo vedere è che per $t, s \in D$ vale la stima

$$|X_t - X_s| \leq c|t - s|^\gamma.$$

Supponiamo quindi $|t - s| < \frac{1}{2^n} = \delta$. Questo implica che $t, s \in D_m, m > n$. Affermiamo che

$$|X_t - X_s| \leq 2\left(\frac{1}{2^{(n+1)\gamma}} + \frac{1}{2^{(n+2)\gamma}} + \cdots + \frac{1}{2^{m\gamma}}\right) \leq \left(\frac{1}{2^\gamma}\right)^{n+1} \frac{2}{1 - \frac{1}{2^\gamma}}.$$

Dimostriamo l'affermazione per induzione. Supponiamo infatti che valga per differenza di diadii appartenenti a D_{m-1} . Prendiamo $t, s \in D_m$ e altri punti $s^*, t^* \in D_{m-1}$ tali che $s \leq s^*, t^* \leq t$. Quindi

$$\begin{aligned} |X_t - X_s| &\leq |X_t - X_{t^*}| + |X_{t^*} - X_{s^*}| + |X_{s^*} - X_s| \leq \\ &\leq \frac{1}{2^{m\gamma}} + \frac{1}{2^{m\gamma}} + 2\left(\frac{1}{2^{(n+1)\gamma}} + \cdots + \frac{1}{2^{(m-1)\gamma}}\right) = 2\left(\frac{1}{2^{(n+1)\gamma}} + \cdots + \frac{1}{2^{m\gamma}}\right). \end{aligned}$$

L'affermazione vale evidentemente per il passo iniziale, cioè $n+1$, in quanto $|X_t - X_s| \leq \frac{1}{2^{(n+1)\gamma}} \leq \frac{2}{2^{(n+1)\gamma}}$. E quindi l'affermazione è provata.

Poniamo ora $\delta = \frac{1}{2^{m^*}}$, $c = 2\left(\frac{1}{1 - \frac{1}{2^\gamma}}\right)$. Consideriamo due istanti $t, s \in D$. Allora $\exists n$ tale che $\frac{1}{2^{n+1}} < |t - s| < \frac{1}{2^n}$. Sia inoltre $|t - s| < \delta$. Allora

$$|X_t(\omega) - X_s(\omega)| \leq c \left(\frac{1}{2^{n+1}}\right)^\gamma \leq c|t - s|^\gamma.$$

16 Appunti del corso

Abbiamo ottenuto la stima che cercavamo quando gli istanti appartengono ai diadi. Ovviamente per altri tempi non abbiamo nessun controllo del processo, che può allegramente fare quello che gli pare e piace senza preoccuparsi di quello che piacerebbe a noi. Per questo andiamo a definire un nuovo processo $X'_t(\omega)$ nella maniera seguente.

1. $X'_t(\omega) = X_t(\omega)$, se $t \in D$
2. $X'_t(\omega) = \lim_n X_{t_n}(\omega)$, se $t \notin D$, ove $t_n \rightarrow t, t_n \in D$
3. $X'_t(\omega) = 17$, se $\omega \in A$

Per le stime viste in precedenza il nuovo processo è evidentemente a traiettorie continue.

Ci resta da dimostrare che effettivamente i due processi sono uno versione dell'altro cioè che $\mathbb{P}(X_t = X'_t) = 1$.

Se $t \in D$ il fatto è vero per definizione. Altrimenti, se $t \notin D$, consideriamo una successione $t_n \rightarrow t$ tale che $X_{t_n} \rightarrow X_t$ quasi ovunque. Questo implica che $X_{t_n} \rightarrow X_t$ in probabilità, e questo basta a farci vincere.

Mostriamo ora che le misure finito dimensionali dei due processi coincidono.

Infatti, consideriamo l'insieme $\Omega_t = \{\omega : X_t(\omega) = X'_t(\omega)\}$. Ovviamente $\mathbb{P}(\Omega_t) = 1$.

Inoltre $\mathbb{P}(\bigcup_{t_i} \Omega_{t_i}^c) \leq \sum_{t_i} \mathbb{P}(\Omega_{t_i}^c) = 0$.

E quindi

$$\begin{aligned}\mathbb{P}((X_{t_1}, \dots, X_{t_n}) \in B) &= \mathbb{P}((X_{t_1}, \dots, X_{t_n}) \in B, \bigcap_{t_i} \Omega_{t_i}) = \\ &= \mathbb{P}((X'_{t_1}, \dots, X'_{t_n}) \in B, \bigcap_{t_i} \Omega_{t_i}) = \mathbb{P}((X'_{t_1}, \dots, X'_{t_n}) \in B).\end{aligned}$$

E con ciò concludiamo la dimostrazione. ■

Capitolo 2

Processi Stocastici

1 Processi Gaussiani

I *Processi Stocastici Gaussiani* sono caratterizzati dalla proprietà che tutte le misure finito dimensionali $\mu \in \mathcal{M}$ sono distribuzioni normali, in cui è tradizione indicare la media con la funzione $m(t) := \mathbb{E}(X_t)$ e la funzione di covarianza con $\varphi(s, t) := \mathbb{E}((X_t - m(t))(X_s - m(s)))$. Dunque secondo questa definizione, l' n -esima misura finito dimensionale si scrive così:

$$\mu_{t_1, \dots, t_n}(B) := \int_B \left(\frac{1}{\sqrt{2\pi}} \right)^n \frac{1}{\sqrt{\det(A)}} e^{-\frac{1}{2}\langle A^{-1}(x - m_{t_1, \dots, t_n}), x - m_{t_1, \dots, t_n} \rangle} dx$$

dove $x \in \mathbb{R}^n$, $m_{t_1, \dots, t_n} = (m(t_1), \dots, m(t_n))$, $A = a_{ij} = \varphi(t_i, t_j) \forall i, j = 1 \dots n$.

La matrice di covarianza A deve avere rango massimo in questa notazione, altrimenti non risulterebbe invertibile. Per evitare questo problema si preferisce usare la funzione caratteristica

$$\phi_{t_1, \dots, t_n}(z) = \int_{\mathbb{R}^n} e^{i\langle z, x \rangle \mu_{t_1, \dots, t_n}} dx = e^{i\langle m_{t_1, \dots, t_n}, z \rangle - \frac{1}{2}\langle Az, z \rangle}$$

In questa maniera non si ha più la dipendenza da A^{-1} e non serve dunque richiedere che sia non degenere.

Esempio 3. Calcolare $\mathbb{E}((X_t - X_s)^2)$.

$$(X_t - X_s)^2 = (X_t - m(t) + m(t) - (X_s - m(s)) - m(s))^2$$

chiamando $Y_t = X_t - m(t)$, $Y_s = X_s - m(s)$ e sostituendo si ricava

$$(Y_t - Y_s + m(t) - m(s))^2 = (Y_t - Y_s)^2 + 2(m(t) - m(s))(Y_t - Y_s) + (m(t) - m(s))^2$$

Ora, poichè

$$\mathbb{E}(Y_t - Y_s) = \mathbb{E}(Y_t) - \mathbb{E}(Y_s) = m(t) - m(t) - m(s) + m(s) = 0$$

rimane solamente

$$\begin{aligned} \mathbb{E}((X_t - X_s)^2) &= \mathbb{E}((Y_t - Y_s)^2) + 0 + (m(t) - m(s))^2 = \\ \mathbb{E}(Y_t^2 - 2Y_t Y_s + Y_s^2) + (m(t) - m(s))^2 &= \varphi(t, t) - 2\varphi(s, t) + (m(t) - m(s))^2. \end{aligned}$$

2 Processi di Markov

Definizione 2.1. Un processo di Markov è il dato di un processo (X_t) e di una famiglia di probabilità di transizione $p(s, x; t, I)$ tali che siano verificate le seguenti condizioni:

1. $(s, x, t) \rightarrow p(s, x; t, I)$ sia una funzione misurabile
2. $I \rightarrow p(s, x; t, I)$ sia una misura di probabilità
3. Sia soddisfatta la relazione di Chapman - Kolmogorov,

$$p(s, x; t, I) = \int_{\mathbb{R}} p(s, x; r, dz)p(r, x; t, I)$$
4. $p(s, x; s, I) = \delta_x(I)$
5. $\mathbb{E}(\mathbf{1}_I(X_t) | \mathcal{F}_s) = \mathbb{E}(\mathbf{1}_I(X_t) | X_s) = p(s, X_s; t, I),$
 ove $\mathcal{F}_t = \sigma\{X_u : u \leq t\}$

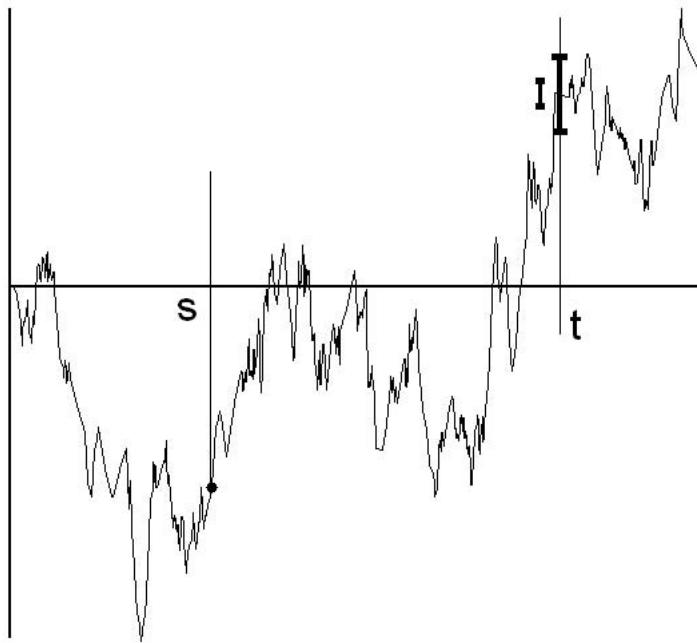


Figura 2.1: Proprietà di transizione

L'ultima proprietà della definizione è sostanzialmente equivalente alla nota condizione $\mathbb{P}(X_t \in I | X_s = x, X_{s_1} = x_1, \dots, X_{s_n} = x_n) = \mathbb{P}(X_t \in I | X_s = x)$ che il lettore troverà più consueta.

Si può inoltre notare come nella definizione di processo di Markov non sia considerata la famiglia \mathcal{M} di misure finito-dimensionali, che caratterizza (per il primo teorema di Kolmogorov) il processo stocastico. Vediamo quindi, a titolo esplicativo, come sia possibile sfruttare la conoscenza delle probabilità di transizione per ricavare \mathcal{M} .

$$\begin{aligned}
 & \mu_{t_1, t_2, \dots, t_n}(I_1 \times \cdots \times I_n) = \\
 &= \mathbb{P}((X_{t_1}, \dots, X_{t_n}) \in I_1 \times \cdots \times I_n) \\
 &= \mathbb{P}((X_{t_1}, \dots, X_{t_{n-1}}) \in I_1 \times \cdots \times I_{n-1}, X_{t_n} \in I_n) \\
 &= \int_{I_1 \times \cdots \times I_{n-1}} \mathbb{P}(X_{t_n} \in I_n | X_{t_{n-1}} = x_{n-1}, \dots, X_{t_1} = x_1) \mathbb{P}(X_{t_1} \in dx_1, \dots, X_{t_{n-1}} \in dx_{n-1}) \\
 &= \int_{I_1 \times \cdots \times I_{n-1}} p(t_{n-1}, x_{n-1}; t_n, I_n) \mathbb{P}(X_{t_1} \in dx_1, \dots, X_{t_{n-1}} \in dx_{n-1}) \\
 &= \int_{I_1 \times \cdots \times I_{n-1}} p(t_{n-1}, x_{n-1}; t_n, I_n) p(t_{n-2}, x_{n-2}; t_{n-1}, dx_{n-1}) \mathbb{P}(X_{t_1} \in dx_1, \dots, X_{t_{n-2}} \in dx_{n-2})
 \end{aligned}$$

Reiterando lo stesso procedimento otteniamo alla fine la seguente espressione:

$$\begin{aligned}
 & \mu_{t_1, t_2, \dots, t_n}(I_1 \times \cdots \times I_n) = \\
 &= \int_I \mathbb{P}(X_0 \in dx_0) p(0, x_0; t_1, dx_1) p(t_1, x_1; t_2, dx_2) \dots p(t_{n-1}, x_{n-1}; t_n, dx_n)
 \end{aligned}$$

ove $I = I_1 \times \cdots \times I_{n-1}$ e $\mathbb{P}(X_0 \in dx_0)$ è la distribuzione iniziale.

3 Relazione di Chapman-Kolmogorov per Processi di Markov

Sia X_t un processo di Markov, cioè valga la seguente proprietà:

$$\mathbb{P}(X_t \in I | X_s = x, X_{s_1} = x_1, \dots, X_{s_n} = x_n) = \mathbb{P}(X_t \in I | X_s = x).$$

Indichiamo inoltre con $p(s, x; t, I)$ la probabilità $\mathbb{P}(X_t \in I | X_s = x)$. Allora

Definizione 2.2. La relazione di Chapman-Kolmogorov può essere enunciata nel modo seguente:

$$p(s, x; t, I) = \int_{\mathbb{R}} p(s, x; r, dz) p(r, x; t, I),$$

ove $s < t$.

Vediamo ora come può essere scritta la relazione facendo alcune ipotesi.

Supponiamo per prima cosa che $p(s, x; t, I) = \int_I p(s, x; t, y) dy$, cioè che la probabilità di transizione ammetta densità.

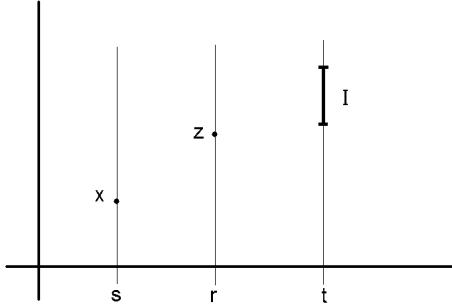


Figura 2.2: Relazione di Chapman - Kolmogorov

Vediamo a cosa si riduce la relazione in questo caso:

$$\int_I p(s, x; t, y) dy = \int_{\mathbb{R}} p(s, x; r, z) dz \int_I p(r, z; t, y) dy = \\ \int_I \int_{\mathbb{R}} p(s, x; r, z) p(r, z; t, y) dz dy.$$

E questo implica che

$$p(s, x; t, y) = \int_{\mathbb{R}} p(s, x; r, z) p(r, z; t, y) dz$$

quasi ovunque.

Supponiamo ora che $p(s, x; t, I) = p(s+\delta, x; t+\delta; I)$. Allora vale evidentemente $p(s, x; t, I) = p(0, x; t-s; I) =: p(t-s, x, I) =: p_{t-s}(x, I)$.

In questo caso la relazione diventa semplicemente

$$p_{s+t}(x, I) = \int_{\mathbb{R}} p_s(x, dy) p_t(y, I).$$

Se ammette densità si può anche scrivere in questa forma

$$p_{s+t}(x, y) = \int_{\mathbb{R}} p_s(x, z) p_t(z, y) dz.$$

Se oltre all'omogeneità temporale vale anche l'ipotesi di omogeneità spaziale, cioè

$$p(s, x + \xi; t, y + \xi) = p(s, x; t, y) = p(s, 0; t, y - x) =: p_{t-s}(y - x) =: p_t(x)$$

la relazione diventa semplicemente

$$p_{s+t}(x) = \int_{\mathbb{R}} p_s(x - y) p_t(y) dy,$$

cioè $p_{s+t} = p_s * p_t$, ove $*$ indica il prodotto di convoluzione.

4 Processo di Wiener

Definizione 2.3. Il processo di Wiener ha media nulla $m(t) := 0 \forall t$, funzione di covarianza $\varphi(s, t) := \min\{s, t\} = s \wedge t$.

In letteratura questo particolare processo si indica spesso con W_t in onore di Wiener. Supponendo $s < t$ si ottengono le seguenti caratteristiche, proprie di questo processo:

$$\mathbb{E}(W_t) = m(t) = 0, \quad \text{Var}(W_t) = \min\{t, t\} = t$$

$$\mathbb{E}(W_t - W_s) = 0, \quad \mathbb{E}((W_t - W_s)^2) = t - 2s + s + 0 = t - s$$

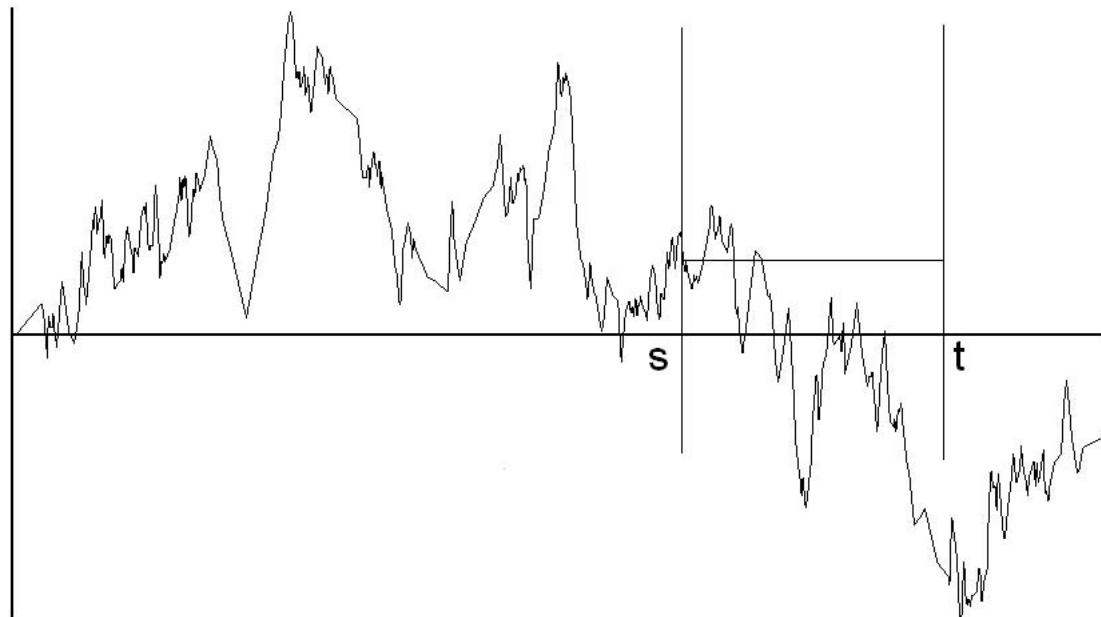


Figura 2.3: Una realizzazione del Processo di Wiener

Ciò mostrerà che è un processo a incrementi indipendenti, inoltre come si vedrà nel seguito, è un processo che si può considerare continuo, ma non differenziabile (in nessun punto). Vediamo se W_s e $W_t - W_s$ sono indipendenti. Iniziamo a verificare che sono gaussiani: li possiamo scrivere come combinazione lineare.

$$(W_s, W_t - W_s) = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} W_s \\ W_t \end{pmatrix}$$

Da ciò si evince che sono gaussiani, se dimostriamo che sono anche scorrelati, allora otteniamo l'indipendenza.

$$a_{12} = a_{21} = \mathbb{E}(W_s(W_t - W_s)) = \mathbb{E}(W_s W_t) - \mathbb{E}(W_s^2) = s - s = 0$$

Nel caso generale per n variabili, il ragionamento è analogo. In quel caso:

$$a_{ij} = \mathbb{E}((W_{t_i} - W_{t_{i-1}})(W_{t_j} - W_{t_{j-1}})) =$$

$$= \mathbb{E}(W_{t_i}W_{t_j} - W_{t_{i-1}}W_{t_j} + W_{t_{i-1}}W_{t_{j-1}} - W_{t_{j-1}}W_{t_i}) = t_i - t_{i-1} + t_{i-1} - t_i = 0$$

In questo modo possiamo riscrivere il processo di Wiener:

$$W_t = (W_{\frac{t}{N}} - W_0) + (W_{\frac{2t}{N}} - W_{\frac{t}{N}}) + \cdots + (W_{\frac{tN}{N}} - W_{\frac{t(N-1)}{N}})$$

e i singoli incrementi sono indipendenti, come abbiamo appena dimostrato. Allora si ha che

$$W_{\frac{kt}{N}} - W_{\frac{(k-1)t}{N}} \sim N(0, \frac{t}{N}) \Rightarrow \frac{W_{\frac{kt}{N}} - W_{\frac{(k-1)t}{N}}}{\sqrt{\frac{t}{N}}} \sim N(0, 1)$$

ovvero

$$W_t = \sqrt{\frac{t}{N}}(X_1 + \cdots + X_N)$$

4.1 Distribuzione del Processo di Wiener

Teorema 2.4. *Le distribuzioni finito dimensionali del processo di Wiener sono date da*

$$\mu_{t_1, \dots, t_n}(B) = \int_B g_{t_1}(x_1)g_{t_2-t_1}(x_2-x_1)\cdots g_{t_n-t_{n-1}}(x_n-x_{n-1})dx_1\cdots dx_n$$

Dimostrazione. Consideriamo il processo W_t , cioè il processo di Wiener che è un processo gaussiano con $m(t) = 0$ e $\phi(s, t) = \min(s, t)$. Vogliamo calcolare $\mu_{t_1, \dots, t_n}(B)$.

Sappiamo già che

$$\mu_{t_1, \dots, t_n}(B) := \int_B \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\sqrt{\det A}} e^{-\frac{1}{2}\langle A^{-1}x, x\rangle} dx,$$

ove

$$A = \begin{pmatrix} t_1 & t_1 & \cdots & \cdots & t_1 \\ t_1 & t_2 & \cdots & \cdots & t_2 \\ \vdots & t_2 & t_3 & \cdots & t_3 \\ \vdots & \cdots & \cdots & \cdots & \cdots \\ t_1 & t_2 & \cdots & \cdots & t_n \end{pmatrix}$$

è la matrice di covarianza.

Vorremmo rendere esplicito il determinante della matrice e il prodotto scalare risultante.

Compiendo operazioni sulle righe, otteniamo facilmente che

$$\det(A) = t_1(t_2 - t_1) \cdots (t_n - t_{n-1}).$$

Passiamo ora a calcolare il prodotto scalare.

Sia $A^{-1}x = y$, cioè $x = Ay$. Questo implica che $\langle A^{-1}x, x \rangle = \langle y, x \rangle$. Scrivendo esplicitamente otteniamo

$$\begin{cases} x_1 = t_1(y_1 + \cdots + y_n) \\ x_2 = t_1y_1 + t_2(y_2 + \cdots + y_n) \\ \vdots \\ x_n = t_1y_1 + t_2y_2 + \cdots + t_ny_n \end{cases}$$

Manipolando il sistema otteniamo questa semplice espressione per il prodotto scalare:

$$\langle A^{-1}x, x \rangle = \frac{x_1^2}{t_1} + \frac{(x_2 - x_1)^2}{t_2 - t_1} + \cdots + \frac{(x_n - x_{n-1})^2}{t_n - t_{n-1}}$$

Quindi, introducendo la funzione $g_t(\alpha) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{\alpha^2}{t}}$, otteniamo la seguente espressione

$$\mu_{t_1, \dots, t_n}(B) = \int_B g_{t_1}(x_1)g_{t_2-t_1}(x_2 - x_1) \cdots g_{t_n-t_{n-1}}(x_n - x_{n-1}) dx_1 \cdots dx_n$$

■

4.2 Wiener è un Processo di Markov

Consideriamo il processo di Wiener caratterizzato nel solito modo:

$$m(t) = 0 \quad e \quad \phi(s, t) = \min(s, t).$$

Teorema 2.5. *Il processo di Wiener rispetta la condizione di Markov.*

$$\mathbb{P}(W_t \in I | W_s = x, W_{s_1} = x_1, \dots, W_{s_n} = x_{s_n}) = \mathbb{P}(W_t \in I | W_s = x)$$

Dimostrazione. Calcoliamo distintamente le due probabilità, cominciando da

$$\mathbb{P}(W_t \in I | W_s = x).$$

Per definizione vale il seguente fatto:

$$\mathbb{P}(W_t \in I, W_s \in B) = \int_B \mathbb{P}(W_t \in I | W_s = x) \mathbb{P}(W_s \in dx)$$

Ma d'altro canto

$$\begin{aligned} \mathbb{P}(W_t \in I, W_s \in B) &= \int_{B \times I} g_s(x_1) g_{t-s}(x_2 - x_1) dx_1 dx_2 = \\ &\quad \int_B \left\{ \int_I g_{t-s}(x_2 - x_1) dx_2 \right\} g_s(x_1) dx_1 \end{aligned}$$

E quindi, confrontando le due espressioni e notando che

$$\mathbb{P}(W_s \in dx) = g_s(x)dx$$

24 Appunti del corso

otteniamo

$$\mathbb{P}(W_t \in I | W_s = x) = \int_I g_{t-s}(y - x) dy,$$

ove l'uguaglianza è da intendersi quasi ovunque.

Calcoliamo quindi il valore di

$$\mathbb{P}(W_t \in I | W_s = x, W_{s_1} = x_1, \dots, W_{s_n} = x_{s_n})$$

e vediamo che in effetti coincide con quello calcolato in precedenza.

Consideriamo quindi una variabile aleatoria $n + 1$ dimensionale, Y , e facciamo il conto per $B = J_1 \times \dots \times J_n \times J$ (il lettore può verificare che una tale scelta è sufficiente).

Come in precedenza possiamo scrivere le seguenti uguaglianze

$$\begin{aligned} \int_B \mathbb{P}(W_t \in I | Y = y) \mathbb{P}(Y \in dy) &= \mathbb{P}(W_t \in I, Y \in B) = \\ &= \mathbb{P}(W_t \in I, W_{s_1} \in J_1, \dots, W_{s_n} \in J_n, W_s \in J) \end{aligned}$$

D'altronde l'ultima probabilità può essere espressa nel seguente modo

$$\int_{B \times I} g_{s_1}(y_1) g_{s_2 - s_1}(y_2 - y_1) \cdots g_{s - s_n}(y_{n+1} - y_{s_n}) g_{t-s}(y - y_{n+1}) dy_1 \cdots dy_{n+1} dy$$

Riconoscendo che

$$\mathbb{P}(Y \in dy) = g_{s_1}(y_1) \cdots g_{s - s_n}(y_{n+1} - y_n) dy_1 \cdots dy_{n+1},$$

dove per il primo dy si intende in effetti $dy_1 \cdots dy_{n+1}$ otteniamo

$$\int_B \mathbb{P}(W_t \in I | Y = y) \mathbb{P}(Y \in dy) = \int_B \left\{ \int_I g_{t-s}(y - y_{n+1}) dy \right\} \mathbb{P}(Y \in dy)$$

La qual cosa implica senza tema di dubbio

$$\mathbb{P}(W_t \in I | Y = y) = \int_I g_{t-s}(y - y_{n+1}) dy,$$

quasi ovunque. Notando che nel nostro caso $y_{n+1} = x$, possiamo ritenere conclusa la dimostrazione.

Nella formula della misura di Wiener

$$\mu_{t_1, \dots, t_n}(B) = \int_B g_{t_1}(x_1) g_{t_2 - t_1}(x_2 - x_1) \cdots g_{t_n - t_{n-1}}(x_n - x_{n-1}) dx_1 \cdots dx_n$$

si riconosce in $g_t(\xi) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{\xi^2}{2t}}$ l'equazione del calore (Wiener ~ 1923).

Abbiamo quindi trovato un'espressione del tipo $\int_{\Omega} F(q) W(dq)$ con $F(q)$ un funzionale e $W(dq)$ la misura. A noi interessa un caso particolare, sui cilindri (che sono misurabili), scegliendo $F(q) = \phi(q(t_1), \dots, q(t_n))$. Dunque questa scelta comporta che:

$$\int_{\Omega} F(q) W(dq) =$$

$$= \int_{\mathbb{R}^n} \phi(x_1, \dots, x_n) g_{t_1}(x_1) g_{t_2-t_1}(x_2 - x_1) \dots g_{t_n-t_{n-1}}(x_n - x_{n-1}) dx_1 \dots dx_n$$

nel caso in cui $\phi = 1_B$ ci si riduce al noto $\mu(B)$. Ci interessa ora verificare che $\mathbb{E}((W_t - W_s)^2) = t - s$, per poterlo fare ricorriamo all'espressione trovata in precedenza:

$$(W_t - W_s)^2 = \phi(q(s), q(t)) \Rightarrow \mathbb{E}((W_t - W_s)^2) = \int_{\Omega} \phi(q(s), q(t)) W(dq) =$$

$$\int_{\mathbb{R}^2} (x_2 - x_1)^2 g_s(x_1) g_{t-s}(x_2 - x_1) dx_1 dx_2 = t - s$$

Inoltre, dal fatto che $X_t - X_s \sim N(0, t - s) \Rightarrow \mathbb{E}((W_t - W_s)^2) =$

$$= \int_{\mathbb{R}} x^2 \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{x^2}{2(t-s)}} dx =$$

indicando con $y = \frac{x}{\sqrt{t-s}} \Rightarrow x = y\sqrt{t-s}$ si ottiene (nel caso generale di potenze pari):

$$(t-s)^m \int_{\mathbb{R}} y^{2m} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = (t-s)^m c_m$$

Rendiamo esplicito c_m :

$$c_m := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} y^{2m} e^{-\frac{y^2}{2}} dy = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} y^{2m-1} (ye^{-\frac{y^2}{2}}) dy$$

osservando che $\frac{d}{dy} \left(e^{-\frac{y^2}{2}} \right) = -ye^{-\frac{y^2}{2}}$ si semplifica in:

$$-\frac{1}{\sqrt{2\pi}} \left[y^{2m-1} e^{-\frac{y^2}{2}} \right]_{-\infty}^{+\infty} + \frac{(2m-1)}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{y^2}{2}} y^{2m-2} dy = (2m-1)c_{m-1}$$

dunque per ricorsione si trova $c_m = (2m-1)(2m-3)(2m-5) \dots 5 \cdot 3 \cdot 1 = (2m-1)!! = \frac{(2m)!}{2^m m!}$. Come vedremo più avanti, questo fatto implica che le traiettorie del processo di Wiener siano *hölderiane* per $\alpha < \frac{1}{2}$, cioè $|f(t) - f(s)| \leq L|t - s|^{\alpha}$ dove L è la costante di Lipschitz. E ciò conclude la dimostrazione. ■

4.3 Costruzione del Processo di Wiener

Consideriamo il seguente operatore su $C([0, 1])$

$$Qf(t) = \int_0^1 (s \wedge t) f(s) ds$$

che ha come autovalori ed autofunzioni ($Qf_k = \lambda_k f_k$)

$$\frac{1}{\sqrt{\lambda_k}} = \frac{2k+1}{2}, \pi \quad f_k(t) = \sin\left(\frac{2k+1}{2}\pi t\right)$$

Inoltre

$$\int_0^1 \sin^2 \left(\frac{2k+1}{2} \pi t \right) dt = \frac{1}{2}$$

ed il processo di Wiener ha il seguente sviluppo

$$W(t) = \sum_{k=0}^{\infty} a_k X_k \sin \left(\frac{2k+1}{2} \pi t \right)$$

con

$$a_k = \sqrt{2\lambda_k}$$

Infatti

$$\frac{1}{2} a_k X_k = \int_0^1 W(t) \sin \left(\frac{2k+1}{2} \pi t \right) dt$$

da cui

$$\frac{1}{4} a_k^2 = \int_0^1 \int_0^1 \mathbb{E}[W(t)W(s)] \sin \left(\frac{2k+1}{2} \pi t \right) \sin \left(\frac{2k+1}{2} \pi s \right) ds dt$$

cioè

$$\frac{1}{4} a_k^2 = \int_0^1 \int_0^1 (s \wedge t) \sin \left(\frac{2k+1}{2} \pi t \right) \sin \left(\frac{2k+1}{2} \pi s \right) ds dt,$$

e quindi

$$\begin{aligned} \frac{1}{4} a_k^2 &= \int_0^1 f_k(t) dt \int_0^1 s \wedge t f_k(s) ds = \int_0^1 f_k(t) Qf_k(t) dt = \\ &= \lambda_k \int_0^1 f_k(t)^2 dt = \frac{\lambda_k}{2} \end{aligned}$$

In conclusione abbiamo

$$W(t) = \sqrt{2} \sum_{k=0}^{\infty} \sqrt{\lambda_k} X_k \sin \left(\frac{2k+1}{2} \pi t \right) = \sqrt{2} \sum_{k=0}^{\infty} X_k \frac{\sin \left(\frac{2k+1}{2} \pi t \right)}{\frac{2k+1}{2} \pi}$$

5 Brownian Bridge

5.1 Transizione nel Brownian Bridge

Siamo interessati a calcolare la probabilità di transizione quando il processo è il Brownian Bridge.

Definizione 2.6. *Brownian Bridge è un processo gaussiano, caratterizzato da media nulla, $m(t) = 0 \forall t$ e ha funzione di covarianza data da $\varphi(s, t) = s \wedge t - st$.*

Proposizione 2. *Cerchiamo la seguente probabilità:*

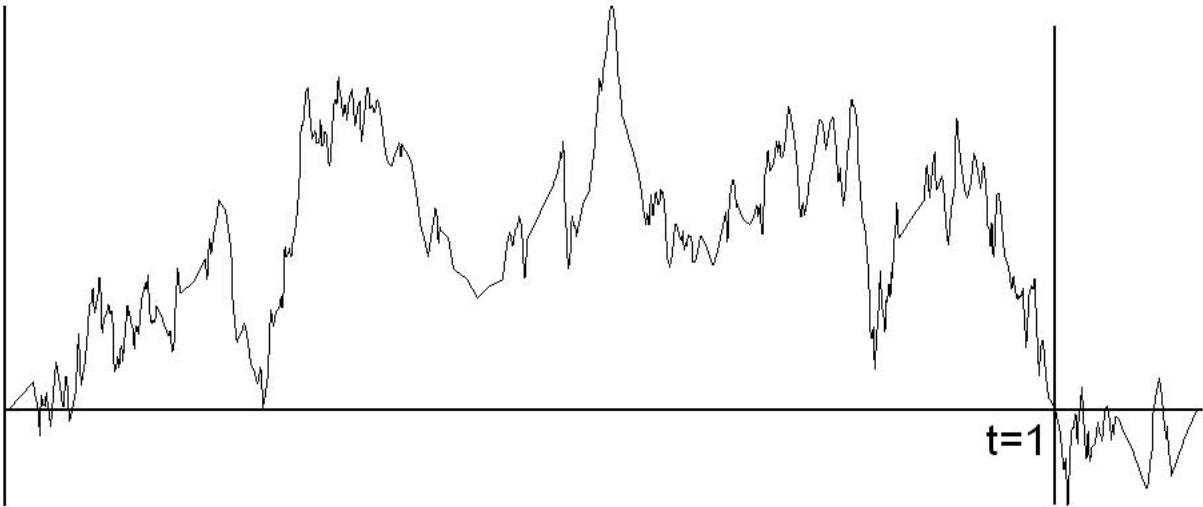


Figura 2.4: Una realizzazione del Brownian Bridge

Dimostrazione.

$$\mathbb{P}(X_t \in I | X_s = x)$$

Per poterla calcolare conviene passare per la formula di Bayes e scrivere

$$\mathbb{P}(X_s \in J, X_t \in I) = \mathbb{P}(X_t \in I | X_s = x) \mathbb{P}(X_s \in dx)$$

Dobbiamo quindi risolvere l'integrale al primo membro e riconoscere in esso l'integrale del secondo fattore al secondo membro per poterlo isolare e ottenere la probabilità condizionata che ci serve.

$$\mathbb{P}(X_s \in J, X_t \in I) = \int_{J \times I} \left(\frac{1}{\sqrt{2\pi}} \right)^2 \frac{1}{\sqrt{\det(A)}} e^{-\frac{1}{2}\langle A^{-1}\xi, \xi \rangle} d\xi$$

con $\xi = (x, y) \in \mathbb{R}^2$. La matrice di covarianza del Brownian Bridge si ricava dalla sua φ , ed è:

$$A = \begin{pmatrix} s(1-s) & s(1-t) \\ s(1-t) & t(1-t) \end{pmatrix} \Rightarrow A^{-1} = \begin{pmatrix} \frac{t}{s(1-s)} & \frac{1}{(s-t)} \\ \frac{1}{(s-t)} & \frac{1-s}{(s-t)(t-1)} \end{pmatrix}$$

segue facilmente che $\det(A) = s(s-t)(t-1)$ e $\det(A^{-1}) = 1/\det(A)$. Dobbiamo ora eseguire il prodotto scalare, cosa che richiederà parecchi (semplici) conti, in breve si ha:

$$\begin{aligned} (x \ y) \begin{pmatrix} \frac{t}{s(1-s)} & \frac{1}{(s-t)} \\ \frac{1}{(s-t)} & \frac{1-s}{(s-t)(t-1)} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \\ = \frac{x^2 t(1-t) + y^2 s(1-s) + 2xys(t-1)}{s(s-t)(t-1)} \end{aligned}$$

Allora il nostro integrale si traduce in:

$$\int_{J \times I} \left(\frac{1}{\sqrt{2\pi}} \right)^2 \frac{1}{\sqrt{s(s-t)(t-1)}} e^{-\frac{1}{2}\frac{x^2 t(1-t) + y^2 s(1-s) + 2xys(t-1)}{s(s-t)(t-1)}} dx dy$$

Ora spezziamolo nelle due parti e isoliamo il termine corrispondente a $\mathbb{P}(X_s \in dx)$.

$$\begin{aligned} & \int_J \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{s(1-s)}} e^{\frac{-x^2}{2s(1-s)}} dx \int_I \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\frac{(1-t)(t-s)}{1-s}}} e^{-\frac{1}{2} \frac{x^2 s(t-1)^2 + y^2 s(s-1)^2 + 2xys(t-1)(1-s)}{s(1-s)(s-t)(t-1)}} dy = \\ & = \mathbb{P}(X_s \in dx) \int_I \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\frac{(1-t)(t-s)}{1-s}}} e^{-\frac{1}{2} \frac{(y - \frac{1-t}{1-s}x)^2}{\frac{1-t}{1-s}(t-s)}} dy \end{aligned}$$

Quindi si vede che la probabilità di transizione del Brownian Bridge è una normale di media $\frac{1-t}{1-s}x$ e varianza $\frac{(1-t)(t-s)}{1-s}$, riassumendo:

$$\mathbb{P}(X_t \in I | X_s = x) \sim N\left(\frac{1-t}{1-s}x, \frac{(1-t)(t-s)}{1-s}\right).$$

■

5.2 Costruzione del Brownian Bridge

Nel caso del brownian bridge per calcolare i corrispondenti autovalori e autofunzioni abbiamo che

$$\begin{aligned} Qf(t) = g(t) &= \int_0^1 (s \wedge t - st) f(s) ds = \\ &= (1-t) \int_0^t sf(s) ds + t \int_t^1 (1-s)f(s) ds \end{aligned}$$

da cui

$$g(0) = 0, \quad g(1) = 0.$$

Inoltre

$$g'(t) = - \int_0^t sf(s) ds + \int_t^1 (1-s)f(s) ds = - \int_0^1 sf(s) ds + \int_t^1 f(s) ds,$$

e

$$g''(t) = -f(t).$$

Quindi da

$$Qf = g = \lambda f$$

otteniamo

$$\lambda f'' = -f, \quad f(0) = f(1) = 0$$

che risulta dà

$$\frac{1}{\sqrt{\lambda_k}} = k\pi, \quad k = 1, 2, 3, \dots, \quad f_k(t) = \sin(k\pi t)$$

Quindi volendo lo sviluppo del brownian bridge

$$Y(t) = \sum_{k=1}^{\infty} a_k X_k \sin(k\pi t)$$

abbiamo che

$$\frac{1}{2}a_k X_k = \int_0^1 Y(t) \sin(k\pi t) dt$$

da cui

$$\frac{1}{4}a_k^2 = \int_0^1 \int_0^1 \mathbb{E}[Y(t)Y(s)] \sin(k\pi t) \sin(k\pi s) ds dt$$

cioè

$$\frac{1}{4}a_k^2 = \int_0^1 f_k(t) Q f_k(t) dt = \lambda_k \int_0^1 f_k(t)^2 dt = \frac{\lambda_k}{2}$$

e quindi

$$a_k = \sqrt{2\lambda_k}.$$

Concludendo abbiamo

$$Y(t) = \sqrt{2} \sum_{k=1}^{\infty} X_k \frac{\sin(k\pi t)}{k\pi}$$

Siccome il brownian bridge si ottiene anche da un processo di Wiener

$$B(t) - tB(1) = Y(t)$$

possiamo anche scrivere, visto che $X_0 = B(1)$ è indipendente da $Y(\cdot)$,

$$B(t) = tX_0 + \sqrt{2} \sum_{k=1}^{\infty} X_k \frac{\sin(k\pi t)}{k\pi}$$

Teorema 2.7. *Il Brownian Bridge, scritto come serie, dà luogo ad un processo continuo.*

Dimostrazione. La dimostrazione si articola in più passi:

- Partiamo da una successione di variabili casuali $X_k, k = 1, 2, 3, \dots$, indipendenti ed equidistribuite come normali standard $N(0, 1)$.
- Consideriamo la serie per $0 \leq t \leq 1$

$$Y(t) = \sqrt{2} \sum_{k=1}^{\infty} X_k \frac{\sin(k\pi t)}{k\pi}$$

- Mostriamo che la serie è assolutamente uniformemente convergente.

Per questo, usiamo le seguenti notazioni, con $m < n$ fissati.

Consideriamo le somme finite

$$V_{m,n}(t) = \sqrt{2} \sum_{k=m+1}^n X_k \frac{\sin(k\pi t)}{k\pi}, \quad U_{m,n}(t) = \sqrt{2} \sum_{k=m+1}^n X_k \frac{\cos(k\pi t)}{k\pi}$$

e la somma a valori complessi

$$Z_{m,n}(t) = U_{m,n}(t) + i V_{m,n}(t) = \sqrt{2} \sum_{k=m+1}^n X_k \frac{e^{ik\pi t}}{k\pi}.$$

Definiamo inoltre

$$S_{m,n} = \sup_{0 \leq t \leq 1} |V_{m,n}(t)|.$$

Primo passo

Evidentemente abbiamo

$$|V_{m,n}(t)| \leq |Z_{m,n}(t)|$$

Secondo passo

$$\begin{aligned} |Z_{m,n}(t)|^2 &= Z_{m,n}(t)\bar{Z}_{m,n}(t) = 2 \sum_{k=m+1}^n X_k \frac{e^{ik\pi t}}{k\pi} \sum_{k=m+1}^n X_k \frac{e^{-ik\pi t}}{k\pi} = \\ &= \frac{2}{\pi^2} \left\{ \sum_{k=m+1}^n X_k^2 \frac{1}{k^2} + \sum_{h \neq k} X_h X_h \frac{e^{i(k-h)\pi t}}{kh} \right\} \end{aligned}$$

Poniamo $p = k - h$, allora potremo scrivere

$$\begin{aligned} |Z_{m,n}(t)|^2 &= \frac{2}{\pi^2} \left\{ \sum_{k=m+1}^n \frac{X_k^2}{k^2} + \sum_{p=1}^{n-m+1} e^{ip\pi t} \sum_{h=m+1}^{n-p} \frac{X_{h+p} X_h}{(h+p)h} + \right. \\ &\quad \left. + \sum_{p=1}^{n-m+1} e^{-ip\pi t} \sum_{h=m+1}^{n-p} \frac{X_{h+p} X_h}{(h+p)h} \right\}, \end{aligned}$$

da cui

$$|Z_{m,n}(t)|^2 \leq \frac{2}{\pi^2} \left\{ \sum_{k=m+1}^n \frac{X_k^2}{k^2} + 2 \sum_{p=1}^{n-m+1} \left| \sum_{h=m+1}^{n-p} \frac{X_{h+p} X_h}{(h+p)h} \right| \right\},$$

e

$$|V_{m,n}(t)|^2 \leq \frac{2}{\pi^2} \left\{ \sum_{k=m+1}^n \frac{X_k^2}{k^2} + 2 \sum_{p=1}^{n-m+1} \left| \sum_{h=m+1}^{n-p} \frac{X_{h+p} X_h}{(h+p)h} \right| \right\},$$

e quindi abbiamo anche che

$$|S_{m,n}|^2 \leq \frac{2}{\pi^2} \left\{ \sum_{k=m+1}^n \frac{X_k^2}{k^2} + 2 \sum_{p=1}^{n-m+1} \left| \sum_{h=m+1}^{n-p} \frac{X_{h+p} X_h}{(h+p)h} \right| \right\}.$$

Terzo passo

Prendendo i valori di aspettazione si ha

$$\mathbb{E}(S_{m,n}^2) \leq \frac{2}{\pi^2} \left\{ \sum_{k=m+1}^n \frac{1}{k^2} + 2 \sum_{p=1}^{n-m+1} \mathbb{E} \left(\left| \sum_{h=m+1}^{n-p} \frac{X_{h+p} X_h}{(h+p)h} \right| \right) \right\}$$

Per l'ultimo termine abbiamo

$$\left(\mathbb{E} \left| \sum_{h=m+1}^{n-p} \frac{X_{h+p} X_h}{(h+p)h} \right| \right)^2 \leq \mathbb{E} \left[\left(\sum_{h=m+1}^{n-p} \frac{X_{h+p} X_h}{(h+p)h} \right)^2 \right]$$

e sviluppando

$$\left(\sum_{h=m+1}^{n-p} \frac{X_{h+p} X_h}{(h+p)h} \right)^2 = \sum_{h=m+1}^{n-p} \frac{X_{h+p}^2 X_h^2}{(h+p)^2 h^2} + 2 \sum_{m+1 \leq h < k \leq n-p} \frac{X_{h+p} X_h X_{k+p} X_k}{(h+p)h(k+p)k},$$

da cui, prendendo i valori di aspettazione

$$\mathbb{E} \left[\left(\sum_{h=m+1}^{n-p} \frac{X_{h+p} X_h}{(h+p)h} \right)^2 \right] = \sum_{h=m+1}^{n-p} \frac{1}{(h+p)^2 h^2},$$

e quindi

$$\mathbb{E} \left| \sum_{h=m+1}^{n-p} \frac{X_{h+p} X_h}{(h+p)h} \right| \leq \left[\sum_{h=m+1}^{n-p} \frac{1}{(h+p)^2 h^2} \right]^{\frac{1}{2}}$$

Quarto passo

Ricapitolando abbiamo

$$\begin{aligned} \mathbb{E}(S_{m,n}^2) &\leq \frac{2}{\pi^2} \left\{ \sum_{k=m+1}^n \frac{1}{k^2} + 2 \sum_{p=1}^{n-m+1} \left[\sum_{h=m+1}^{n-p} \frac{1}{(h+p)^2 h^2} \right]^{\frac{1}{2}} \right\} \\ &\leq \frac{2}{\pi^2} \left\{ \frac{n-m}{m^2} + 2(n-m) \sqrt{\frac{n-m}{m^4}} \right\}. \end{aligned}$$

In particolare, con $n = 2m$ abbiamo

$$\mathbb{E}(S_{m,2m}^2) \leq \frac{2}{\pi^2} \left\{ \frac{1}{m} + 2 \sqrt{\frac{1}{m}} \right\} \leq \frac{6}{\pi^2} \sqrt{\frac{1}{m}}.$$

Quindi

$$\mathbb{E}(S_{m,2m}) \leq \left(\mathbb{E}(S_{m,2m}^2) \right)^{\frac{1}{2}} \leq \frac{\sqrt{6}}{\pi} m^{-\frac{1}{4}}$$

ed in particolare

$$\sum_{m=1}^{\infty} \mathbb{E}(S_{2^m, 2^{m+1}}) \leq \frac{\sqrt{6}}{\pi} \sum_{m=1}^{\infty} 2^{-\frac{m}{4}} < \infty$$

Quinto passo

Scriviamo la serie di partenza come

$$\sum_{k=1}^{\infty} X_k \frac{\sin(k\pi t)}{k\pi} = X_1 \frac{\sin(\pi t)}{\pi} + \sum_{m=0}^{\infty} \sum_{k=2^m+1}^{2^{m+1}} X_k \frac{\sin(k\pi t)}{k\pi}$$

ed applicare il risultato del quarto passo per avere la convergenza assoluta ed uniforme.
Quindi il processo $Y(t)$

$$Y(t) = \sqrt{2} \sum_{k=1}^{\infty} X_k \frac{\sin(k\pi t)}{k\pi}$$

è a traiettorie continue.

Inoltre è un processo gaussiano di media nulla e funzione di covarianza $\mathbb{E}(Y(t)Y(s)) = s \wedge t - st$: è un Brownian Bridge. ■

5.3 Moto Browniano

Dato un moto browniano $B(t)$, il processo $Y(t) = B(t) - tB(1)$ è un Brownian Bridge. Inoltre $Y(t)$ e $B(1)$ sono indipendenti! (Basta calcolare la covarianza).

Partendo ora da un Brownian Bridge $Y(t)$ come precedentemente definito, e considerando una ulteriore variabile aleatoria gaussiana X_0 ($N(0, 1)$), che prende il ruolo di $B(1)$, definiamo

$$B(t) = tX_0 + \sqrt{2} \sum_{k=1}^{\infty} X_k \frac{\sin(k\pi t)}{k\pi},$$

che risulta essere un processo stocastico a traiettorie continue ed è un moto browniano.

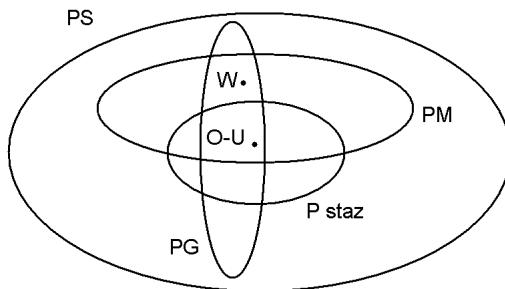


Figura 2.5: Divisione fondamentale dei processi stocastici

6 Martingale

La parola martingala deriva dal francese *martingale* e lo spagnolo *martingala*, termine che proviene a sua volta da *martingaux*, abitante di Martigues, cittadina della Provenza. Si dicono “alla martingala” certi calzoni che si abbottonano da dietro, poi si è passati ad intendere certi nastri che per bellezza ricadevano dalla cintura dei pantaloni. Per similitudine, si dice ora martingala la correggia che, attaccata da uno dei capi, alla cinghia che stringe il corpo, dall’altro al muso del cavallo, obbliga quest’ultimo a camminare tenendo la testa in una posizione fissa. In matematica, l’etimologia deriva da un tipo di scommesse che prevedevano il raddoppio della posta per ogni giocata successiva.

Definizione 2.8. Un processo stocastico $\{X_t, t \geq 0\}$ e una filtrazione $\{\mathcal{F}_t\}_{t \geq 0}$ si dicono una martingala se:

1. X_t è \mathcal{F}_t -misurabile (si dice che il processo è adattato alla filtrazione) cioè $\mathcal{F}_t^X \subset \mathcal{F}_t$, ove $\mathcal{F}_t^X = \sigma\{X_u \text{ t.c. } u \leq t\}$

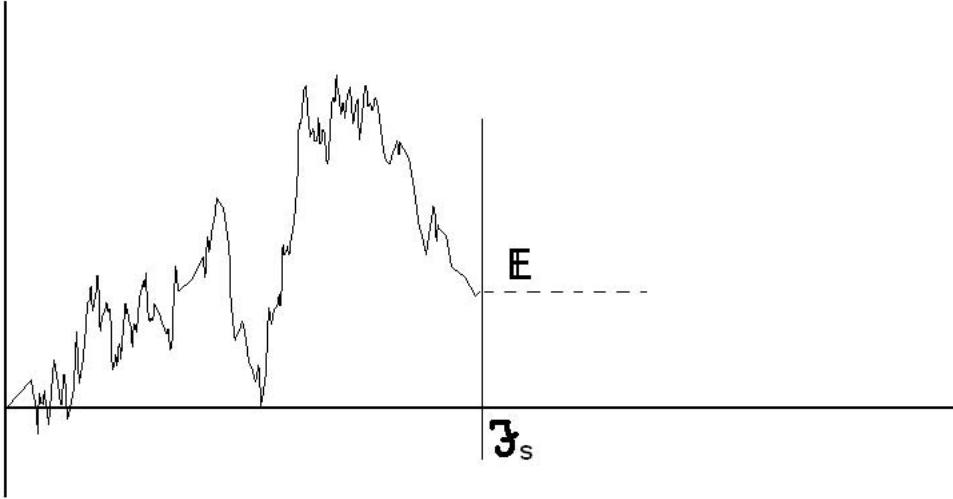


Figura 2.6: Martingala

$$2. \mathbb{E}(X_t | \mathcal{F}_s) = X_s$$

Esempio 4. Calcoliamo $\mathbb{E}(\mathbb{E}(X_t | \mathcal{F}_s))$.

Se definiamo $\mathcal{F}_0 := \{\emptyset, \Omega\}$ abbiamo naturalmente che la media è costante, cioè $\mathbb{E}(X_t | \mathcal{F}_0) = \mathbb{E}(X)$, inoltre poichè risulta $\mathcal{F}_0 \subset \mathcal{F}_s$ si possono scrivere le seguenti uguaglianze.

$$\mathbb{E}(\mathbb{E}(X_t | \mathcal{F}_s)) = \mathbb{E}(\mathbb{E}(X_t | \mathcal{F}_s) | \mathcal{F}_0) = \mathbb{E}(X_t | \mathcal{F}_0) = \mathbb{E}(X_t) = \mathbb{E}(X_s) = \mathbb{E}(X_0)$$

Sia ora $X_\infty \in \mathcal{L}^1$, definiamo un nuovo processo $X_t := \mathbb{E}(X_\infty | \mathcal{F}_t)$ al variare di t . Ci chiediamo se sia ancora una martingala. Facciamo le dovute verifiche: innanzitutto è \mathcal{F}_t -misurabile e risulta $\mathbb{E}(X_t | \mathcal{F}_s) = \mathbb{E}(\mathbb{E}(X_\infty | \mathcal{F}_t) | \mathcal{F}_s) = \mathbb{E}(X_\infty | \mathcal{F}_s) = X_s$. Dunque è effettivamente una martingala.

Sotto opportune ipotesi, si può dimostrare che esiste il limite

$$\lim_{t \rightarrow \infty} X_t = X_\infty$$

che viene chiamato *legge dei grandi numeri in forma generalizzata*.

Per le martingale vale questa disegualanza di Chebychev, che può risultare utile in alcuni casi perchè permette di confrontare la norma del sup con una norma integrale (che può risultare computazionalmente vantaggiosa), ad esempio per mostrare la continuità per limitatezza:

$$\mathbb{P}\left(\sup_{0 \leq s \leq t} |X_s| > \varepsilon\right) \leq \frac{1}{\varepsilon^p} \mathbb{E}(X_t^p)$$

Essa vale anche nel caso di *sottomartingale*, ovvero le martingale in cui vale (punto 2) $\mathbb{E}(X_t | \mathcal{F}_s) \geq X_s$; invece per *supermartingale* (caso col \geq) la disegualanza non vale più. Consideriamo ora il processo di Wiener adattato alla filtrazione \mathcal{F}_t , cioè il caso in cui $\mathcal{F}_t^W \subset \mathcal{F}_t$, \mathcal{F}_t e $W_{t+h} - W_h$ sono indipendenti. In queste ipotesi abbiamo, per il processo di Wiener che:

$$\mathbb{E}(W_t | \mathcal{F}_s) = \mathbb{E}(W_t - W_s + W_s | \mathcal{F}_s) = \mathbb{E}(W_t - W_s | \mathcal{F}_s) + \mathbb{E}(W_s | \mathcal{F}_s) =$$

34 Appunti del corso

$$= \mathbb{E}(W_t - W_s) + W_s = 0 + W_s = W_s$$

Questo ragionamento si può applicare al processo W_t^n nella maniera seguente:

$$\begin{aligned} \mathbb{E}(W_t^n | \mathcal{F}_s) &= \mathbb{E}((W_t - W_s + W_s)^n | \mathcal{F}_s) = \\ &= \mathbb{E}\left(\sum_{k=0}^n \binom{n}{k} (W_t - W_s)^k W_s^{n-k} | \mathcal{F}_s\right) = \sum_{k=0}^n \binom{n}{k} \mathbb{E}((W_t - W_s)^k W_s^{n-k} | \mathcal{F}_s) = \\ &= \sum_{k=0}^n \binom{n}{k} W_s^{n-k} \mathbb{E}((W_t - W_s)^k | \mathcal{F}_s) = \sum_{k=0}^n \binom{n}{k} W_s^{n-k} \mathbb{E}((W_t - W_s)^k) \end{aligned}$$

Si vede come per gli indici k dispari il termine generale si annulli, restano quindi solo i pari, allora, posto $b = \frac{n}{2}$ (o alla sua parte intera se n è dispari):

$$\sum_{k=0}^b \binom{n}{2k} W_s^{n-2k} \mathbb{E}((W_t - W_s)^{2k}) = \sum_{k=0}^b \binom{n}{2k} \frac{(2k)!}{2^k k!} W_s^{n-2k} (t-s)^k$$

Per $n = 2$ si ha $\mathbb{E}(W_t^2 | \mathcal{F}_s) = W_s^2 - s + t$.

Proviamo adesso a fare un conto simile al precedente ma per una funzione esponenziale, scegliamo $\vartheta \in \mathbb{R}$ e vediamo che cosa dà $\mathbb{E}(e^{\vartheta W_t} | \mathcal{F}_s)$. Per poter calcolare questo valore di aspettazione condizionata conviene scrivere l'esponenziale nel suo sviluppo di Taylor, $e^{\vartheta W_t} = \sum_{n=0}^{\infty} \frac{\vartheta^n W_t^n}{n!}$. Quindi, se sostituiamo all'interno del valore di aspettazione, otteniamo:

$$\begin{aligned} \mathbb{E}(e^{\vartheta W_t} | \mathcal{F}_s) &= \sum_{n=0}^{\infty} \frac{\vartheta^n}{n!} \sum_{k=0}^b \binom{n}{2k} \frac{(2k)!}{2^k k!} W_s^{n-2k} (t-s)^k = \\ &= \sum_{n=0}^{\infty} \vartheta^n \sum_{k=0}^b \frac{1}{(n-2k)! 2^k k!} W_s^{n-2k} (t-s)^k = \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^b \frac{1}{(n-2k)! k!} (\vartheta W_s)^{n-2k} \left[\frac{\vartheta^2}{2} (t-s)\right]^k = \\ &= \sum_{k=0}^{\infty} \frac{\left[\frac{\vartheta^2}{2} (t-s)\right]^k}{k!} \sum_{n=2k}^{\infty} \frac{(\vartheta W_s)^{n-2k}}{(n-2k)!} = \\ &= \sum_{k=0}^{\infty} \frac{\left[\frac{\vartheta^2}{2} (t-s)\right]^k}{k!} e^{\vartheta W_s} = e^{\frac{\vartheta^2}{2} (t-s) + \vartheta W_s} \end{aligned}$$

Ciò mostra dunque che $\mathbb{E}(e^{\vartheta W_t} | \mathcal{F}_s) = e^{\frac{\vartheta^2}{2} (t-s) + \vartheta W_s}$ e quindi che $\mathbb{E}(e^{\vartheta W_t - \frac{\vartheta^2 t}{2}} | \mathcal{F}_s) = e^{\vartheta W_s - \frac{\vartheta^2 s}{2}}$ cioè che è una martingala.

Nei casi successivi si ha che

$$e^{\vartheta x - \frac{\vartheta^2}{2}} = \sum_{n=0}^{\infty} H_n(x) \frac{\vartheta^n}{n!}$$

dove gli $H_n(x)$ sono polinomi, che sono chiamati *polinomi di Hermite*. Essi risultano dallo sviluppo in serie di Taylor dell'espressione precedente (funzione generatrice). Si possono dunque esplicitare in

$$H_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}$$

i primi termini sono i seguenti:

$$\begin{aligned} H_0 &= 1 \\ H_1 &= x \\ H_2 &= x^2 - 1 \\ H_3 &= x^3 - 3x \\ H_4 &= x^4 - 6x^2 + 3 \end{aligned}$$

Essi hanno varie proprietà, vediamone alcune. Posto δ_{ij} il delta di Kronecker, valgono queste relazioni notevoli:

$$\int_{\mathbb{R}} H_n(x) H_m(x) e^{-\frac{x^2}{2}} dx = n! \sqrt{2\pi} \delta_{nm}$$

Questo fatto mostra come essi siano tra loro ortogonali rispetto alla densità di probabilità normale standard. Ci sono ulteriori identità significative, ad esempio soddisfano l'equazione differenziale di Hermite oppure si possono trovare per ricorsione.

$$\begin{aligned} H_n''(x) - xH_n'(x) + nH_n(x) &= 0 \\ H_{n+1}(x) &= xH_n(x) - H_n'(x) \\ H_n'(x) &= nH_{n-1}(x) \end{aligned}$$

A questo punto possiamo normalizzarli mediante $H_n(\frac{W_t}{\sqrt{t}})\sqrt{t}^n$ e quindi ricaviamo che $\mathbb{E}(H_n(\frac{W_t}{\sqrt{t}})\sqrt{t}^n | \mathcal{F}_s) = H_n(\frac{W_s}{\sqrt{s}})\sqrt{s}^n$ che come si vede, rappresenta una martingala. Facciamo la verifica per $n = 2$, $H_2(x) = x^2 - 1$, dunque

$$H_2\left(\frac{W_t}{\sqrt{t}}\right)\sqrt{t}^2 = \left(\frac{W_t^2}{t} - 1\right)t = W_t^2 - t.$$

Torniamo a rivolgere la nostra attenzione alle martingale. Come abbiamo visto per le passeggiate casuali, anche con questi processi si possono calcolare le probabilità di raggiungere un certo stato ad un certo tempo. Iniziamo con $\mathbb{P}(\sup_{0 \leq s \leq t} W_s \geq c)$, ovvero la prob-

abilità di toccare o superare la quota c entro il tempo t . Come nel caso delle passeggiate casuali, applichiamo il principio di riflessione ad ogni traiettoria che una volta toccata quota c torna più bassa. In questo modo possiamo decomporre l'evento $(\sup_{0 \leq s \leq t} W_s \geq c)$

nell'unione di $(W_t \geq c) \bigcup (\sup_{0 \leq s \leq t} W_s \geq c, W_t < c)$. Si può notare inoltre che i due eventi sono gli stessi e di conseguenza

$$\mathbb{P}(\sup_{0 \leq s \leq t} W_s \geq c) = 2\mathbb{P}(W_t \geq c) = 2 \int_c^\infty \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} dx = 2 \int_c^\infty g_t(x) dx$$

36 Appunti del corso

Passiamo adesso alla probabilità di arrivare a c la prima volta ad un certo tempo, $\mathbb{P}(\tau_c \leq t)$. Anche in questo caso osserviamo che $(\tau_c \leq t) = (\sup_{0 \leq s \leq t} W_s \geq c)$. Ne possiamo ora provare a vedere la densità:

$$\frac{d}{dt} \int_c^\infty g_t(x) dx = \int_c^\infty \frac{\partial g_t}{\partial t}(x) dx$$

In quest'ultimo integrale riconosciamo il nucleo di Gauss che soddisfa l'equazione del calore, quindi possiamo proseguire con

$$\frac{1}{2} \int_c^\infty \frac{\partial^2 g_t}{\partial x^2}(x) dx = \frac{1}{2} \left[\frac{\partial g_t}{\partial x} \right]_c^\infty$$

In conclusione $\mathbb{P}(\tau_c \in dt) = \frac{c}{t} \frac{1}{\sqrt{2\pi t}} e^{-\frac{c^2}{2t}} dt$.

Capitolo 3

Equazioni Differenziali Stocastiche

1 Brownian Bridge da un'Equazione Differenziale Stocastica

Consideriamo l'equazione differenziale stocastica

$$dX_t = \frac{b - X_t}{T - t} dt + dW_t, \quad X_0 = a$$

dove W_t è un processo di Wiener standard. Possiamo scrivere una soluzione “esplicita” osservando che

$$d\left(\frac{X_t - b}{T - t}\right) = \frac{dX_t}{T - t} + \frac{X_t - b}{(T - t)^2} dt = \frac{dW_t}{T - t},$$

da cui

$$\frac{X_t - b}{T - t} - \frac{X_s - b}{T - s} = \int_s^t \frac{dW_r}{T - r}$$

e anche

$$X_t = \frac{T - t}{T - s} X_s + \frac{t - s}{T - s} b + \int_s^t \frac{T - t}{T - r} dW_r.$$

Dall'ultima si può verificare subito che $X_T = b$. Inoltre le probabilità di transizione sono date da

$$p(s, x; t, I) := \mathbb{P}(X_t \in I \mid X_s = x)$$

e sono gaussiane con valor medio e varianza dati da

$$\frac{(T - t)x + (t - s)b}{T - s}, \quad \int_s^t \left(\frac{T - t}{T - r}\right)^2 dr = \frac{T - t}{T - s}(t - s)$$

o in altre notazioni

$$p(s, x; t, I) = \mathcal{N}\left(\frac{(T - t)x + (t - s)b}{T - s}, \frac{T - t}{T - s}(t - s)\right)(I)$$

oppure con la notazione

$$g_t(z) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{z^2}{t}},$$

$p(s, x; t, I)$ ha la seguente densità di probabilità

$$g_{\frac{T-t}{T-s}(t-s)}\left(y - \frac{(T-t)x + (t-s)b}{T-s}\right).$$

Lemma 3.1. *Si ha l'identità*

$$g_{\frac{T-t}{T-s}(t-s)}\left(y - \frac{(T-t)x + (t-s)b}{T-s}\right) = g_{t-s}(y-x) \frac{g_{T-t}(b-y)}{g_{T-s}(b-x)}$$

che può essere dimostrata con facili e noiosi calcoli.

2 Brownian Bridge Condizionando un Processo di Wiener

Consideriamo

$$\mathbb{P}(W_t \in I \mid W_s \in [x - \varepsilon, x + \varepsilon], W_T \in [b - \varepsilon, b + \varepsilon]).$$

o in altre parole

$$\mathbb{P}(W_t \in I \mid |W_s - x| \leq \varepsilon, |W_T - b| \leq \varepsilon) = \frac{\mathbb{P}(W_t \in I, |W_s - x| \leq \varepsilon, |W_T - b| \leq \varepsilon)}{\mathbb{P}(|W_s - x| \leq \varepsilon, |W_T - b| \leq \varepsilon)}.$$

È facile mostrare che esiste il limite per $\varepsilon \rightarrow 0$ ed è

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathbb{P}(W_t \in I, |W_s - x| \leq \varepsilon, |W_T - b| \leq \varepsilon)}{\mathbb{P}(|W_s - x| \leq \varepsilon, |W_T - b| \leq \varepsilon)} = \int_I g_{t-s}(y-x) \frac{g_{T-t}(b-y)}{g_{T-s}(b-x)} dy.$$

Consideriamo ora la seguente probabilità

$$\mathbb{P}(W_t \in I \mid W_s = x, W_T = b);$$

è un facile esercizio mostrare che

$$\mathbb{P}(W_t \in I \mid W_s = x, W_T = b) = \int_I g_{t-s}(y-x) \frac{g_{T-t}(b-y)}{g_{T-s}(b-x)} dy,$$

cioè, arriviamo alla stessa conclusione ottenuta tramite limite.

Possiamo ora verificare mediante calcolo diretto che

$$p(s, x; t, I) := \mathbb{P}(W_t \in I \mid W_s = x, W_T = b)$$

verifica la relazione di Chapman-Kolmogorov; quindi possiamo costruire il processo di Markov corrispondente X_t , che è il brownian bridge.

3 Alcune Equazioni Differenziali Stocastiche

Teorema 3.2. *L'equazione che definisce il processo di Ornstein - Uhlenbeck*

$$\dot{X}_t = -\lambda X_t + \sigma \dot{W}_t$$

Dimostrazione. Moltiplicando entrambi i membri per $e^{\lambda t}$ otteniamo

$$(e^{\lambda t} X_t)' = \sigma e^{\lambda t} \dot{W}_t$$

Da cui integrando otteniamo

$$X_t = e^{-\lambda t} X_0 + \sigma \int_0^t e^{-\lambda(t-s)} dW_s$$

Inoltre da fatti noti sull'integrale di *Stieltjes* sappiamo che

$$\sigma \int_0^t e^{-\lambda(t-s)} dW_s \sim \mathbf{N} \left(0, \sigma^2 \int_0^t e^{-2\lambda(t-s)} ds = \frac{1 - e^{-2\lambda t}}{2\lambda} \right)$$

Quindi supponendo X_0 indipendente dalla parte integrale e $X_0 \sim \mathbf{N}(m, v^2)$ otteniamo che

$$X_t \sim \mathbf{N} \left(e^{-\lambda t} m, \frac{\sigma^2}{2\lambda} + e^{-2\lambda t} \left(v^2 - \frac{\sigma^2}{2\lambda} \right) \right)$$

Se poniamo $X_0 \sim \mathbf{N} \left(0, \frac{\sigma^2}{2\lambda} \right)$ otteniamo che anche X_t è distribuito nello stesso modo. Abbiamo trovato cioè la distribuzione uniforme.

Volendo obiettare che \dot{W}_t non esiste, si può risolvere l'equazione integrale

$$X_t = X_0 - \lambda \int_0^t X_s ds + \sigma W_t$$

e risolvendola si ottiene la medesima soluzione. ■

Teorema 3.3. *L'equazione che definisce il Brownian Bridge è:*

$$\dot{X}_t = -\frac{X_t}{1-t} + \dot{W}_t$$

Dimostrazione. Come in precedenza risolviamola senza tenere conto di cosa voglia dire \dot{W}_t .

Moltiplicando entrambi i membri per $\frac{1}{1-t}$ otteniamo

$$\left(\frac{X_t}{1-t} \right)' = \frac{\dot{W}_t}{1-t}$$

40 Appunti del corso

Integrando otteniamo

$$X_t = (1-t)X_0 + \int_0^t \frac{1-t}{1-s} dW_s$$

Ponendo inoltre $X_0 = 0$ otteniamo

$$X_t = \int_0^t \frac{1-t}{1-s} dW_s \sim \mathbf{N}(0, t(1-t))$$

che è proprio il *brownian bridge*.

Se ora integriamo l'equazione tra s e t otteniamo la seguente espressione

$$X_t = \frac{1-t}{1-s} X_s + \int_s^t \frac{1-t}{1-r} dW_r$$

da cui si può vedere facilmente (notando che il termine integrale è indipendente da X_s perché il processo di *Wiener* è a incrementi indipendenti) che

$$\mathbb{P}(X_t \in I | X_s = x) \sim \mathbf{N}\left(\frac{1-t}{1-s}x, \frac{1-t}{1-s}(t-s)\right)$$

come già sapevamo. ■

Capitolo 4

Semigruppo e Integrale

1 Semigruppo di Operatori Lineari

In questa prima parte consideriamo processi di Markov omogenei nel tempo.

Introduciamo una funzione $u(t, x)$ nel seguente modo

$$u(t, x) = \mathbb{E}(f(X_t) | X_0 = x) = \int_{\mathbb{R}} f(y) \mathbb{P}(X_t \in dy | X_0 = x) = \int_{\mathbb{R}} f(y) p_t(x, dy)$$

Teorema 4.1. Se scriviamo $u(t, x) = U_t f(x)$ vediamo che U è un operatore lineare (su \mathcal{L}^∞) e inoltre vale la seguente relazione di semigruppo

$$U_{t+s} = U_t U_s = U_s U_t$$

Dimostrazione. Sia $g(y) = U_s f(y) = \int_{\mathbb{R}} f(z) p_s(y, dz)$. Allora

$$\begin{aligned} U_t g(x) &= \int_{\mathbb{R}} g(y) p_t(x, dy) = \int_{\mathbb{R}} p_t(x, dy) \int_{\mathbb{R}} f(z) p_s(y, dz) = \\ &= \int_{\mathbb{R}} f(z) \int_{\mathbb{R}} p_s(y, dz) p_t(x, dy) = \int_{\mathbb{R}} f(z) p_{s+t}(x, dz) = U_{s+t} f(x) \end{aligned}$$

■

Teorema 4.2 (Equazione di Kolmogorov). $\frac{\partial u}{\partial t} = A u(t, x)$

Dimostrazione. Proviamo ora a derivare la funzione $u(t, x)$.

$$u(t+h, x) - u(t, x) = U_{t+h} f(x) - U_t f(x) = (U_h - I) U_t f(x)$$

e quindi

$$\frac{u(t+h, x) - u(t, x)}{h} = \frac{U_h - I}{h} U_t f(x) = \frac{U_h g(x) - g(x)}{h}$$

dove abbiamo posto $g(x) = U_t f(x)$.

42 Appunti del corso

Consideriamo ora il seguente insieme $D_A = \{g \in \mathcal{L}^\infty : \frac{U_h g - g}{h} \longrightarrow Ag\}$, ove A è un operatore lineare.

Prendendo $f(x) \in D_A$, otteniamo la seguente equazione differenziale alle derivate parziali

$$\begin{cases} \frac{\partial u}{\partial t} = Au(t, x) \\ u(0, x) = f(x) \end{cases}$$

Riassumendo abbiamo visto che partendo da un processo di Markov omogeneo nel tempo e definendo una funzione in un certo modo otteniamo un'equazione differenziale, detta anche equazione di *Kolmogorov*. In generale si può vedere che queste equazioni sono di tipo parabolico e l'operatore A è un operatore del secondo ordine del tipo

$$Au(x) = \frac{1}{2}a(x)\frac{\partial^2 u}{\partial x^2} + b(x)\frac{\partial u}{\partial x}$$

■

Teorema 4.3 (Equazione di Wiener). $dX_t = b(X_t)dt + \sigma(X_t)dW_t$

Dimostrazione. Ad esempio partendo dal processo di *Wiener* W_t , per cui vale

$$p_t(x, y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}}$$

si ottiene l'operatore $Au = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}$.

Si può vedere inoltre che vale la seguente equazione differenziale stocastica

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t$$

ove $b(x)$ viene chiamata deriva (drift) e $\sigma(x) = \sqrt{a(x)}$ diffusione. ■

Vediamo ora cosa succede se consideriamo processi di Markov non omogenei nel tempo.

Per tali processi definiamo come in precedenza un operatore nel modo seguente

$$U_{s,t}f(x) = \int_{\mathbb{R}} f(y)p(s, x; t, dy)$$

ove $p(s, x; t, I) = \mathbb{P}(X_t \in I | X_s = x)$.

In questo caso la relazione di Chapman-Kolmogorov ci dice che vale

$$U_{s,t} = U_{s,r}U_{r,t}$$

Ponendo $u(s, x, t) = U_{s,t}f(x)$ otteniamo come in precedenza un sistema

$$\begin{cases} \frac{\partial u}{\partial s} + A(s)u(s) = 0 \\ u(t, x, t) = f(x) \end{cases}$$

ove A è un operatore lineare del tipo definito in precedenza.

L'operatore $U_{s,t}$ viene anche detto operatore di evoluzione all'indietro (backward).

Introduciamo ora un operatore sullo spazio delle misure, nel seguente modo

$$V_{t,s}\mu(I) = \int_{\mathbb{R}} \mu(dx)p(s, x; t, I)$$

Come in precedenza si può verificare che vale $V_{t,s} = V_{t,r}V_{r,s}$. Inoltre vale l'importante relazione $U_{s,t} = V_{t,s}^*$ (dove per V^* si intende l'operatore duale di V).

Definizione 4.4. Consideriamo ora lo spazio $\mathcal{L}^\infty \times \mathcal{M}$ e introduciamo una forma bilineare nel seguente modo

$$\langle f, \mu \rangle := \int_{\mathbb{R}} f(z)\mu(dz)$$

Per l'osservazione precedente si può vedere che vale

$$\langle U_{s,t}f, \mu \rangle = \langle f, V_{t,s}\mu \rangle$$

Consideriamo ora lo spazio $D(s) = \left\{ \phi \in \mathcal{L}^\infty : \frac{U_{s,s+h}-I}{h}(\phi) \longrightarrow A(s)\phi \right\}$, ove A è un operatore lineare. Ripetendo i conti già visti in precedenza vediamo che, per $\phi \in D(s)$, vale

$$\frac{\partial}{\partial s}U_{s,t} = -A(s)U_{s,t}$$

Nello stesso modo, considerando $\phi \in D(t)$ si vede che vale

$$\frac{\partial}{\partial t}U_{s,t} = U_{s,t}A(t)$$

Proposizione 3. Per l'insieme $D_V(t) = \left\{ \mu \in \mathcal{M} : \frac{V_{t+h,t}-I}{h}\mu \longrightarrow B(t)\mu \right\}$, ove $B(t)$ è lineare, valgono le seguenti relazioni

$$\frac{\partial}{\partial s}V_{t,s} = -V_{t,s}B(s)$$

$$\frac{\partial}{\partial t}V_{t,s} = B(t)V_{t,s}$$

prendendo μ rispettivamente in $D_V(t)$ e in $D_V(s)$.

Teorema 4.5 (Equazione di Fokker-Planck). *Sia $v(t, I) = V_{t,s}\mu(I)$. Otteniamo quindi l'equazione*

$$\frac{\partial v(t, I)}{\partial t} = B(t)v(t, I)$$

che unitamente ad un dato iniziale in s costituisce l'equazione forward di Kolmogorov, o equazione di Fokker-Planck.

Si può inoltre vedere facilmente che vale

$$A(s)^* = B(s)$$

e quindi le equazioni forward e backward sono strettamente collegate una all'altra.

Se ad esempio

$$A(s)\phi = \frac{1}{2}a(x)\phi_{xx} + b(x)\phi_x$$

si può vedere che

$$A(s)^*\psi = \frac{1}{2}(a\psi)_{xx} - (b\psi)_x$$

Questo fatto segue direttamente da un'integrazione per parti.

Se ora torniamo ai processi di Markov omogenei nel tempo possiamo definire $U_{s,t}f(x) = U_{t-s}f(x)$, come in precedenza. Così facendo otteniamo

$$\frac{\partial U_{s-t}}{\partial(s-t)} = AU_{s-t}$$

e

$$B = A^*$$

cioè i calcoli fatti per processi di Markov generali vengono decisamente semplificati.

Vediamo ora qualche esempio di processi noti e di operatori associati.

Esempio 5. *Processo di Wiener. Come visto in precedenza al processo di Wiener corrisponde l'operatore*

$$A\phi = \frac{1}{2}\phi_{xx}$$

Inoltre si vede che $A^ = A$ e quindi l'operatore è autoaggiunto.*

Esempio 6. *Processo di Ornstein-Uhlenbeck. Consideriamo il processo definito secondo la solita equazione differenziale stocastica $dX_t = -\lambda X_t dt + \sigma dW_t$. In questo caso l'operatore assume la seguente forma*

$$A\phi = \frac{1}{2}\sigma^2\phi_{xx} - \lambda x\phi_x$$

e il suo aggiunto

$$A^* \phi = \frac{1}{2} \sigma^2 \phi_{xx} + \lambda(x\phi)_x$$

Esempio 7. *Brownian Bridge.* Prendiamo il processo definito mediante l'equazione differenziale stocastica $dX_t = -\frac{X_t}{1-t} dt + dW_t$. Allora

$$A(t)\phi = \frac{1}{2}\phi_{xx} - \frac{x}{1-t}\phi_x$$

e possiamo notare che in questo caso l'operatore dipende dal tempo.

2 Richiami di Alcuni Tipi di Integrazione

2.1 Integrale di Riemann-Stieltjes

Rievociamo velocemente i principali risultati che riguardano l'integrazione alla Riemann-Stieltjes. Siano $F : [a, b] \rightarrow \mathbb{R}$ una funzione non decrescente, $P = \{a = x_0, x_1, \dots, x_n = b\}$ una partizione di $[a, b]$ tale che $|x_{i+1} - x_i| = (b - a)/n$ e $C = \{c_0, c_1, \dots, c_{n-1}\}$ con $c_i \in [x_i, x_{i+1}]$.

Consideriamo ora $g : \mathbb{R} \rightarrow \mathbb{R}$ e poniamo:

$$S(g, P, C) = \sum_{i=0}^{n-1} g(c_i)[F(x_{i+1}) - F(x_i)]$$

Se esiste unico (indipendente da P e C) il limite di $S(g)$ per n tendente all'infinito allora diciamo che g è integrabile secondo Riemann-Stieltjes e chiamiamo integrale rispetto a F di g su $[a, b]$ tale limite:

$$\int_a^b g(x)dF := \lim_{n \rightarrow \infty} \sum_{i=0}^n g(c_i)[F(x_{i+1}) - F(x_i)].$$

L'integrale così definito gode delle solite proprietà dell'integrale di Riemann, elenchi-amo ora i risultati più importanti.

Teorema 4.6. Se g è continua in $[a, b]$, allora g è integrabile rispetto ad una $F : [a, b] \rightarrow \mathbb{R}$ non decrescente.

Proposizione 4.7. Siano g_1, g_2 integrabili rispetto a F su $[a, b]$, allora

1. $g_1 + g_2$ e cg_1 per ogni costante c sono integrabili e

$$\int_a^b (g_1 + g_2)dF = \int_a^b g_1 dF + \int_a^b g_2 dF \quad \text{e} \quad \int_a^b cg_1 dF = c \int_a^b g_1 dF.$$

46 Appunti del corso

2. Se $g_1(x) \leq g_2(x)$ su $[a, b]$ allora

$$\int_a^b g_1 dF \leq \int_a^b g_2 dF.$$

3. Se $a < c < b$, g_1 è integrabile su $[a, c]$ e $[c, b]$ e

$$\int_a^c g_1 dF + \int_c^b g_1 dF = \int_a^b g_1 dF.$$

4. Se g_1 è integrabile rispetto a F e ad H su $[a, b]$ allora

$$\int_a^b g_1 d(F+H) = \int_a^b g_1 dF + \int_a^b g_1 dH.$$

5. $g_1 g_2$ è integrabile.

6. $|g_1|$ è integrabile e

$$\left| \int_a^b g_1 dF \right| \leq \int_a^b |g_1| dF.$$

Il teorema che segue formalizza un'idea che abbiamo sempre avuto:

Teorema 4.8. Supponiamo che F sia derivabile e che F' e g siano integrabili rispetto a F . Allora

$$\int_a^b g dF = \int_a^b g F'(x) dx.$$

Il prossimo teorema è di importanza fondamentale ed è noto come teorema del cambiamento di variabile.

Teorema 4.9. Sia $\phi : [A, B] \rightarrow [a, b]$ continua e strettamente crescente. Sia g integrabile rispetto a F su $[a, b]$. Allora $g(\phi)$ è integrabile su $[A, B]$ rispetto a $g(\phi)$ e

$$\int_A^B g(\phi) d(F(\phi)) = \int_a^b g dF$$

Ricordiamo che F è detta a variazione limitata se

$$\sup \left\{ \sum_{i=0}^m |f(x_i) - f(x_{i+1})| \right\} < \infty$$

Inoltre abbiamo la seguente caratterizzazione:

Proposizione 4.10. $F : [a, b] \rightarrow \mathbb{R}$ è a variazione limitata se e solo se esistono $G, H : [a, b] \rightarrow \mathbb{R}$ non decrescenti tali che $F = G - H$.

Risulta chiaro che possiamo integrare rispetto a funzioni a variazione limitata.

Nota 4.11. Per $F(x) = x$ otteniamo l'integrale di Riemann classico.

2.2 Integrale di Lebesgue-Stieltjes

Vogliamo ora affrontare il problema in un'ottica più alla Lebesgue.

Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ una funzione non decrescente. Vogliamo definire attraverso f una misura su \mathbb{R} . Cominciamo col porre $\alpha_f((a, b]) := f(b) - f(a)$ per poi definire la misura di *Lebesgue-Stieltjes* di un sottoinsieme $e \subseteq \mathbb{R}$ come il numero

$$\lambda_f(e) = \inf \left\{ \sum_{i=1}^{\infty} \alpha_f((a_i, b_i]) \text{ t.c. } \bigcup_{i=1}^{\infty} (a_i, b_i] \supseteq e \right\}$$

Risulta così definita una misura esterna $\lambda_f : \mathcal{P} \mapsto [a, \infty]$, in realtà molto di più:

Teorema 4.12. *Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ non decrescente, allora λ_f è una misura esterna di Radon su \mathbb{R} .*

Si potrebbe ingenuamente pensare che sia sempre $\lambda_f((a, b]) = \alpha_f((a, b])$, in realtà abbiamo bisogno di qualche ipotesi di regolarità.

Teorema 4.13. *Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ non decrescente e continua a destra, allora $\lambda_f((a, b]) = \alpha_f((a, b])$.*

Per completezza enunciamo il viceversa del penultimo teorema.

Teorema 4.14. *Sia μ una misura esterna di Radon su \mathbb{R} e $f(x) := \mu((-\infty, x])$ per $x \in \mathbb{R}$. Allora $\lambda_f(B) = \mu(B)$ per ogni $B \in \mathcal{B}$.*

Naturalmente avendo a disposizione una misura esterna di Radon su \mathbb{R} possiamo costruire la teoria dell'integrazione astratta nella consueta maniera.

Nota 4.15. *Se prendiamo $f(x) = x$ otteniamo la misura di Lebesgue sulla retta reale.*

Esempio 8. *Integrale rispetto alla funzione di ripartizione di una variabile casuale, cioè rispetto alla sua misura immagine.*

3 Integrazione Rispetto ad un Processo di Wiener

3.1 Il Caso più Semplice

Supponiamo ora di voler definire l'integrale alla Riemann rispetto ad una funzione F che non sia a variazione limitata. Possiamo scaricare le difficoltà sulla funzione integranda con la formula per parti, precisamente:

$$\sum_{i=0}^{n-1} g(x_i)[F(x_{i+1}) - F(x_i)] = F(b)g(b) - F(a)g(a) - \sum_{i=0}^{n-1} F(x_{i+1})[g(x_{i+1}) - g(x_i)]$$

e se g è a variazione limitata abbiamo risolto il problema. Così ad esempio, possiamo scrivere

$$\int_0^t f(s)df = f(t)^2 - f(0)^2 - \int_0^t f(s)df$$

da cui

$$\int_0^t f(s)df = \frac{f(t)^2 - f(0)^2}{2}$$

come auspicavamo.

3.2 Un Caso Particolare di Integrale di Itô

Un grosso guaio sorge quando si vuole integrare il processo di Wiener rispetto al processo di Wiener, che è tutt'altro che a variazione limitata. Sicuramente vale

$$\sum_{i=0}^{n-1} W(t_i)[W(t_{i+1}) - W(t_i)] = W(t_n)^2 - W(t_0)^2 - \sum_{i=0}^{n-1} W(t_{i+1})[W(t_{i+1}) - W(t_i)]$$

Da inguaribili ottimisti quali siamo potremmo sperare che

$$S = \sum_{i=0}^{n-1} W(t_i)[W(t_{i+1}) - W(t_i)]$$

e

$$S' = \sum_{i=0}^{n-1} W(t_{i+1})[W(t_{i+1}) - W(t_i)]$$

convergano allo stesso limite, ma così non è!

Notiamo subito che

$$S' - S = \sum_{i=0}^{n-1} (W(t_{i+1}) - W(t_i))^2 \quad e \quad S' + S = W(t_n)^2 - W(t_0)^2$$

Allora abbiamo

$$S' = \frac{W(t_n)^2 - W(t_0)^2}{2} + \frac{1}{2} \sum_{i=0}^{n-1} (W(t_{i+1}) - W(t_i))^2$$

e

$$S = \frac{W(t_n)^2 - W(t_0)^2}{2} - \frac{1}{2} \sum_{i=0}^{n-1} (W(t_{i+1}) - W(t_i))^2$$

Per comodità poniamo $t = t_n$ e $t_0 = 0$ e ricordiamo che $W(0) = 0$.

Teorema 4.16. *La somma $V = \sum_{i=0}^{n-1} (W(t_{i+1}) - W(t_i))^2$ converge in $L^2(\Omega)$ a t.*

Dimostrazione. Per le proprietà del processo di Wiener $W(t_{i+1}) - W(t_i) \sim N(0, t_{i+1} - t_i)$ e quindi

$$W(t_{i+1}) - W(t_i) = \sqrt{t_{i+1} - t_i} \frac{(W(t_{i+1}) - W(t_i))}{\sqrt{t_{i+1} - t_i}} = \sqrt{t_{i+1} - t_i} \xi_i$$

con $\xi_i \sim N(0, 1)$.

Ovviamente stiamo studiando la somma V al variare della partizione, nel nostro caso al variare di n , sottointendendo i dovuti indici. Dal fatto che $\mathbb{E}(\xi_i^2) = 1$ segue $\mathbb{E}(V) = t$. A questo punto ci basta far vedere che la varianza di V tende a 0, infatti così facendo potremmo concludere che $V \xrightarrow{L^2} t$.

Espandiamo il quadrato:

$$(V - t)^2 = \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 (\xi_i^2 - 1)^2 + \sum_{i \neq j} (t_{i+1} - t_i)(t_{j+1} - t_j)(\xi_i^2 - 1)(\xi_j^2 - 1)$$

La media della seconda sommatoria è zero per l'indipendenza degli incrementi e poiché $\mathbb{E}((\xi_i^2 - 1)^2) = \mathbb{E}(\xi_i^4) - 1 = 2$ deduciamo che

$$\mathbb{E}((V - t)^2) = 2 \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 \leq 2 \max_k |t_{k+1} - t_k| t \rightarrow 0$$

e abbiamo finito. ■

Possiamo allora scrivere

$$S = \sum_{i=0}^{n-1} W(t_i)(W(t_{i+1}) - W(t_i)) \xrightarrow{L^2} \frac{W(t)^2 + t}{2}$$

e

$$S' = \sum_{i=0}^{n-1} W(t_{i+1})(W(t_{i+1}) - W(t_i)) \xrightarrow{L^2} \frac{W(t)^2 - t}{2}$$

da cui si evince che non esiste l'integrale di Riemann-Stieltjes di W rispetto a W in quanto il limite di $\sum_{i=0}^{n-1} W(c_i)(W(t_{i+1}) - W(t_i))$ dipende dalla scelta dei $c_i \in [t_i, t_{i+1}]$. Ci troviamo dunque innanzi all'ambiguità scelta dei c_i con cui calcolare $\int_0^t W(s)dW$.

Impulsivamente si potrebbe pensare che la decisione migliore sia per $c_i = (t_i + t_{i+1})/2$ con la quale otteniamo l'integrale detto di Stratonovic (che è anche quello che si otterrebbe se W fosse a variazione limitata):

$$\int_0^t W(s)dW := \frac{W(t)^2}{2}.$$

In realtà si ottiene qualcosa di più interessante per $c_i = t_i$ definendo così l'integrale di Itô:

$$\int_0^t W := \lim_{L^2} \sum_{i=0}^{n-1} W(t_i)(W(t_{i+1}) - W(t_i)) = \frac{W(t)^2 - t}{2}.$$

Ad esempio con questa definizione è valido l'importante

Teorema 4.17. *L'integrale di Itô è una martingala rispetto a $\mathcal{F}_t = \sigma\{W_u, u \leq t\}$.*

Dimostrazione. Osserviamo che $\mathbb{E}(W_s|\mathcal{F}_s) = W_s$ in quanto W_s è \mathcal{F}_s -misurabile e che

$$\mathbb{E}(W_t|\mathcal{F}_s) - \mathbb{E}(W_s|\mathcal{F}_s) = \mathbb{E}(W_s - W_t|\mathcal{F}_s) = \mathbb{E}(W_t - W_s) = 0.$$

Inoltre

$$\begin{aligned}\mathbb{E}(W_t^2 - W_s^2|\mathcal{F}_s) &= \mathbb{E}((W_t - W_s)^2|\mathcal{F}_s) + 2\mathbb{E}(W_s(W_t - W_s)|\mathcal{F}_s) = \\ &= \mathbb{E}((W_t - W_s)^2) + 2W_s\mathbb{E}(W_t - W_s|\mathcal{F}_s) = t - s\end{aligned}$$

e quindi, dato che $W_s^2 - s$ è \mathcal{F}_s -misurabile

$$\mathbb{E}(W_t^2 - t - W_s^2 + s|\mathcal{F}_s) = \mathbb{E}(W_t^2 - W_s^2) - (t - s) = 0$$

come volevamo. ■

4 Un Caso Particolare di Meyer

Teorema 4.18 (Teorema di Meyer(1963-)). *Se M_t è a traiettorie continue ed è una martingala allora*

$$\sum_{i=0}^{n-1} (M_{t_{i+1}} - M_{t_i})^2 \rightarrow A(t)$$

e inoltre

$$M_t^2 - A(t)$$

è una martingala.

Nel seguito consideriamo il solito processo di Wiener adattato alla filtrazione $(\mathcal{F}_t)_{t \geq 0}$ e inoltre tale che $W_{t+h} - W_t$ sia indipendente da \mathcal{F}_t per ogni incremento h .

Cerchiamo di vedere cosa succede nel caso in cui $M_t = \int_0^t W(s)dW(s)$.

$$M_{t_{i+1}} - M_{t_i} = \int_{t_i}^{t_{i+1}} W(s)dW(s) \cong W(t_i)(W_{t_{i+1}} - W_{t_i})$$

E questo implica

$$\sum (M_{t_{i+1}} - M_{t_i})^2 \cong \sum W(t_i)^2 (W_{t_{i+1}} - W_{t_i})^2 \cong \sum W(t_i)^2 (t_{i+1} - t_i) \rightarrow \int_0^t W_s^2 ds$$

Quindi, supponendo che $A_t = \int_0^t W_s^2 ds$, vediamo che $M_t^2 - A_t$ è una martingala.

Dobbiamo cioè mostrare che

$$\mathbb{E} \left(\left(\int_0^t W(r)dW_r \right)^2 - \int_0^t W(r)^2 dr | \mathcal{F}_s \right) = \left(\int_0^s W(r)dW_r \right)^2 - \int_0^s W(r)^2 dr$$

Dato che l'aspettazione è lineare e inoltre $\int_0^s W(r)^2 dr$ è \mathcal{F}_s misurabile mostriamo che

$$\mathbb{E} \left(\left(\int_0^t W(r) dW_r \right)^2 - \left(\int_0^s W_r dW_r \right)^2 \mid \mathcal{F}_s \right) = \mathbb{E} \left(\int_s^t W_r^2 dr \mid \mathcal{F}_s \right)$$

Riscriviamo il primo membro in altra maniera.

$$\begin{aligned} & \mathbb{E} \left(\left(\int_0^t W_r dW_r - \int_0^s W_r dW_r \right)^2 + 2 \int_0^t W_r dW_r \int_0^s W_r dW_r - 2 \left(\int_0^s W_r dW_r \right)^2 \mid \mathcal{F}_s \right) \\ &= \mathbb{E} \left(\left(\int_s^t W_r dW_r \right)^2 + 2 \int_0^s W_r dW_r \left(\int_s^t W_r dW_r \right) \mid \mathcal{F}_s \right) = \\ &= \mathbb{E} \left(\left(\int_s^t W_r dW_r \right)^2 \mid \mathcal{F}_s \right) + 2 \int_0^s W_r dW_r \mathbb{E} \left(\int_s^t W_r dW_r \mid \mathcal{F}_s \right) \end{aligned}$$

dove l'integrale è stato portato fuori dall'aspettazione perchè \mathcal{F}_s misurabile.

Concentriamo per ora la nostra attenzione su $\mathbb{E} \left(\int_s^t W_r dW_r \mid \mathcal{F}_s \right)$.

Sfruttando varie proprietà del valore di aspettazione condizionato possiamo scrivere

$$\begin{aligned} \mathbb{E} \left(\int_s^t W_r dW_r \mid \mathcal{F}_s \right) &= \int_s^t \mathbb{E} (W_r dW_r \mid \mathcal{F}_s) = \int_s^t \mathbb{E} (\mathbb{E} (W_r dW_r \mid \mathcal{F}_r) \mid \mathcal{F}_s) = \\ &= \int_s^t \mathbb{E} (W_r \mathbb{E} (dW_r \mid \mathcal{F}_r) \mid \mathcal{F}_s) = \int_s^t \mathbb{E} (W_r \mathbb{E} (dW_r) \mid \mathcal{F}_s) = 0 \end{aligned}$$

E questo è vero perchè $dW_r = W_{r+dr} - W_r$ è indipendente da \mathcal{F}_r e a media nulla.

Quindi ci rimane solo da considerare $\mathbb{E} \left(\left(\int_s^t W_r dW_r \right)^2 \mid \mathcal{F}_s \right)$ che possiamo riscrivere per linearità e scambiando gli integrali così

$$\int_s^t \int_s^t \mathbb{E} (W_r W_{r'} dW_r dW_{r'} \mid \mathcal{F}_s)$$

52 Appunti del corso

Consideriamo il caso in cui $r < r'$. Del tutto analogo è il caso opposto.

$$\begin{aligned} \int_s^t \int_s^t \mathbb{E}(W_r W_{r'} dW_r dW_{r'} | \mathcal{F}_s) &= \int_s^t \int_s^t \mathbb{E}(\mathbb{E}(W_r W_{r'} dW_r dW_{r'} | \mathcal{F}_{r'}) | \mathcal{F}_s) = \\ &= \int_s^t \int_s^t \mathbb{E}(W_r W_{r'} dW_r \mathbb{E}(dW_{r'} | \mathcal{F}_{r'}) | \mathcal{F}_s) = \int_s^t \int_s^t \mathbb{E}(W_r W_{r'} dW_r \mathbb{E}(dW_{r'}) | \mathcal{F}_s) = 0 \end{aligned}$$

per lo stesso motivo precedente.

Ci rimane solo il caso in cui $r = r'$. In questo caso il doppio integrale si riduce ad uno singolo e l'espressione si riduce a

$$\begin{aligned} \int_s^t \mathbb{E}(W_r^2 dW_r^2 | \mathcal{F}_s) &= \int_s^t \mathbb{E}(\mathbb{E}(W_r^2 dW_r^2 | \mathcal{F}_r) | \mathcal{F}_s) = \int_s^t \mathbb{E}(W_r^2 \mathbb{E}(dW_r^2 | \mathcal{F}_r) | \mathcal{F}_s) = \\ &= \int_s^t \mathbb{E}(W_r^2 \mathbb{E}(dW_r^2) | \mathcal{F}_s) = \int_s^t \mathbb{E}(W_r^2 dr | \mathcal{F}_s) = \mathbb{E}\left(\int_s^t W_r^2 dr | \mathcal{F}_s\right) \end{aligned}$$

e questo conclude la dimostrazione.

Il lettore noterà che in vari punti la faccenda è abbastanza sporca, ma può essere facilmente compresa considerando il seguente fatto

$$\int_0^t W_s dW_s = \lim \sum_{i=0}^{n-1} W_{s_i} (W_{s_{i+1}} - W_{s_i})$$

Se ora al posto degli integrali consideriamo le somme tutti i passaggi risultano maggiormente comprensibili e formali. Se il lettore è interessato, come esercizio rifaccia il conto usando le somme e poi passando al limite.

Bibliografia

- [1] **P. Baldi**, EQUAZIONI DIFFERENZIALI STOCASTICHE E APPLICAZIONI, Pitagora Editrice
- [2] **I. Karatzas, S.E. Shreve**, BROWNIAN MOTION AND STOCHASTIC CALCULUS, Springer Verlag
- [3] **W. Feller**, INTRODUCTION TO PROBABILITY THEORY AND ITS APPLICATIONS, VOL.1, Wiley
- [4] **W. Feller**, INTRODUCTION TO PROBABILITY THEORY AND ITS APPLICATIONS, VOL.2, Wiley

Indice analitico

- l -sistema, 8
 - π -sistema, 7
- aspettazione condizionata, 9, 10
- Chebychev, 14, 33
- covarianza
 - nel Brownian Bridge, 27
- delta di Kronecker, 35
- deriva(drift), 42
- diadici, 14
- diffusione, 42
- distribuzioni congiunte, 6, 7
- equazione
 - di Kolmogorov, 41
 - del calore, 24, 36
 - di Fokker-Planck, 44
 - di Kolmogorov, 42
 - di Wiener, 42
 - differenziale di Hermite, 35
- formula di Bayes, 27
- funzione di covarianza, 17
- funzione generatrice, 35
- indipendenza, 12
 - incrementi, 22
- insieme cilindrico, 6
- integrale
 - di Stieltjes, 39
 - di Itô, 49
 - di Lebesgue-Stieltjes, 47
 - di Riemann, 45, 46
 - di Riemann-Stieltjes, 49
 - di Stratonovic, 49
- legge grandi numeri generalizzata, 33
- lemma
- Borel-Cantelli, 13, 14
- Lipschitz, 25
- martingale, 32
 - sottomartingale, 33
 - supermartingale, 33
- misura esterna di Radon, 47
- misura immagine, 7
- moto browniano, 32
- nucleo di Gauss, 36
- operatore di evoluzione, 43
- polinomi di Hermite, 35
- probabilità condizionata ad una σ -algebra, 9
- processo stocastico, 5
 - Brownian Bridge, 26, 45
 - di Markov, 18, 19
 - di Ornstein - Uhlenbeck, 39
 - di Ornstein-Uhlenbeck, 44
 - di Wiener, 21, 44
 - gaussiano, 17
- relazione di Chapman - Kolmogorov, 18, 19, 38, 42
- teorema
 - Beppo-Levi, 12
 - cambio di variabili, 46
 - di Beppo Levi, 8
 - di Radon-Nikodym, 8
 - Kolmogorov I, 13
 - Kolmogorov II, 13
 - Meyer, 50
 - Radon-Nikodym, 9
- traiettoria, 7
- variazione limitata, 46