Analisi stocastica
cenni riassuntivi (in fase di stesura)

Markus Fischer, Università di Padova

Anno accademico 2014 - 2015, primo semestre
versione del 16 dicembre 2014

Notazioni

$\mathbb{N}$	insieme dei numeri naturali a partire da 1.
$\mathcal{B}(E)$	σ-algebra dei boreliani di uno spazio topologico E (cioè la σ-algebra generata dagli aperti).
$\mathbf{C}(E)$	spazio delle funzioni $E \to \mathbb{R}$ continue su uno spazio topologico E
$\mathbf{C}_b(E)$	spazio delle funzioni $E \to \mathbb{R}$ continue e limitate su uno spazio topologico E
$\vee$	massimo tra due valori numerici
$\wedge$	minimo tra due valori numerici
$\mathcal{P}(E)$	spazio delle misure di probabilità sui boreliani di uno spazio topologico E.

Convenzione: $\inf \emptyset = \infty$

Indice

Capitolo 0

Preparativi

0.1 Motivazioni

A. Problema delle diffusioni. Una motivazione, anche dal punto di visto storico, è il seguente problema delle diffusioni. Siano b, σ funzioni reali "buone", ad esempio funzioni in $\mathbf{C}_b^1(\mathbb{R})$ e $\mathbf{C}_b^2(\mathbb{R})$. Supponiamo di avere un processo stocastico $X = (X(t))_{t\geq 0}$ a valori reali e con traiettorie continue tale che, per ogni $t \geq 0$,

$$\mathbf{E}\left[X(t+h) - X(t)|\mathcal{F}_t^X\right] = b(X(t)) \cdot h + o(h), \tag{0.1}$$

dove $\mathcal{F}_t^X \doteq \sigma(X(s) : s \leq t)$ è la σ-algebra generata dal processo fino al tempo t. La (0.1) dice che la funzione di media $t \mapsto \mathbf{E}[X(t)]$ del processo varia con tasso $b(x)$ condizionatamente su $X(t) = x$. La (0.1) implica che

$$\mathbf{E}\left[(X(t+h) - X(t) - b(X(t)) \cdot h)^2 |\mathcal{F}_t^X\right] = \mathbf{E}\left[(X(t+h) - X(t))^2 |\mathcal{F}_t^X\right] + o(h).$$

Supponiamo inoltre che X sia tale che, per ogni $t \geq 0$,

$$\mathbf{E}\left[(X(t+h) - X(t))^2 |\mathcal{F}_t^X\right] = \sigma^2(X(t)) \cdot h + o(h). \tag{0.2}$$

La (0.2), insieme alla (0.1), dice che la funzione di varianza $t \mapsto \mathbf{E}[(X(t) - \mathbf{E}[X(t)])^2]$ del processo varia con tasso $\sigma^2(x)$ condizionatamente su $X(t) = x$.

Poniamo $a(x) \doteq \sigma^2(x)$. Se $a \equiv 0$, allora ci aspettiamo che X soddisfi l'equazione differenziale ordinaria (EDO)

$$dX(t) = b(X(t))dt. \tag{0.3}$$

Se la condizione iniziale $X(0)$ è deterministica, allora X sarà un processo deterministico, cioè una funzione solo del tempo. Cosa succede per $a = \sigma^2$ non identicamente zero?

Sia $f \in \mathbf{C}_c^2(\mathbb{R})$ una funzione test. Grazie a (0.1), (0.2) e la formula di Taylor, dovremmo avere, per $h > 0$ piccolo, $t \geq 0$,

$$\begin{aligned}
&\mathbf{E}\left[f(X(t+h)) - f(X(t))|\mathcal{F}_t^X\right] \\
&\approx \mathbf{E}\left[f'(X(t))\left(X(t+h)) - X(t)\right) + \frac{1}{2}f''(X(t))\left(X(t+h)) - X(t)\right)^2|\mathcal{F}_t^X\right] \\
&\approx \left(b(X(t))f'(X(t)) + \frac{1}{2}a(X(t))f''(X(t))\right)\cdot h.
\end{aligned}$$

Usando una sommatoria telescopica e prendendo valori attesi, dovremmo arrivare a

$$\mathbf{E}\left[f(X(t))\right] = \mathbf{E}\left[f(X(0))\right] + \int_0^t \mathbf{E}\left[\mathcal{A}(f)(X(s))\right] ds, \tag{0.4}$$

dove $\mathcal{A}(f)$ è l'operatore differenziale di secondo ordine dato da

$$\mathcal{A}(f)(x) \doteq b(x)f'(x) + \frac{1}{2}a(x)f''(x).$$

Poniamo $\mu_t \doteq \mathbf{P}\circ(X(t))^{-1} = \mathrm{Law}(X(t))$, cioè μ_t è la distribuzione marginale di X al tempo di t. Indicando con $\langle .,. \rangle$ l'accoppiamento di dualità tra funzioni e misure attraverso l'integrazione, la (0.4) si riscrive come

$$\langle \mu_t, f\rangle = \langle \mu_0, f\rangle + \int_0^t \langle \mu_s, \mathcal{A}(f)\rangle ds$$

oppure in forma differenziale, per ogni $f \in \mathbf{C}_c^2(\mathbb{R})$,

$$\frac{d}{dt}\langle \mu_t, f\rangle = \langle \mu_t, \mathcal{A}(f)\rangle. \tag{0.5}$$

La (0.5) è la forma debole dell'equazione di Kolmogorov forward. Se μ_t ammette una densità rispetto alla misura di Lebesgue, se cioè $\mu_t(dx) = p(t,x)dx$ per ogni $t \geq 0$, allora la (0.5) diventa

$$\frac{d}{dt}p(t,x) = \mathcal{A}^*(p(t,.))(x), \tag{0.6}$$

dove $\mathcal{A}^*$ indica l'operatore duale di $\mathcal{A}$ dato da

$$\mathcal{A}^*(f)(x) \doteq \frac{1}{2}\left(a(x)f(x)\right)'' - \left(b(x)f(x)\right)'$$

Domande / obiettivi:

1. I calcoli qui sopra possono essere resi rigorosi?

2. Qual'è la classe di processi caratterizzati dalle condizioni (0.1) e (0.2)?

3. Si può trovare una rappresentazione probabilistica di tali processi, magari analoga a quelle delle equazioni differenziali ordinarie?

B. Passeggiata aleatoria e gioco d'azzardo. Sia $(\xi_n)_{n\in\mathbb{N}}$ una successione di variabili aleatorie reali indipendenti e identicamente distribuite (i.i.d.) definita su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbf{P})$ con media $\mu \doteq \mathbf{E}[\xi_1]$ e varianza $\sigma^2 \doteq \mathbf{E}\left[(\xi_1-\mu)^2\right] < \infty$. Sia X_0 un'altra variabile aleatoria reale indipendente da $(\xi_n)_{n\in\mathbb{N}}$. Introduciamo la filtrazione a tempo discreto $\mathcal{F}_0 \doteq \sigma(X_0)$, $\mathcal{F}_\backslash \doteq \sigma(X_0, \xi_1, \ldots, \xi_n)$, $n \in \mathbb{N}$, e definiamo per ricorsione la passeggiata aleatoria

$$X_{n+1} \doteq X_n + \xi_{n+1}, \quad n \in \mathbb{N}_0.$$

Se X_0 prende valori in $\mathbb{Z}$ e le ξ_n sono a valori in $\{-1, 1\}$, allora $(X_n)_{n\in\mathbb{N}_0}$ è una passeggiata aleatoria semplice sul reticolo $\mathbb{Z}$, ed è simmetrica se $\mu = 0$. Dalla definizione di $(X_n)_{n\in\mathbb{N}_0}$ abbiamo

$$\begin{aligned} X_{n+1} = X_0 + \sum_{k=1}^{n+1} \xi_k &= X_0 + (n+1)\mu + \underbrace{\sum_{k=1}^{n+1}(\xi_k - \mu)}_{\doteq M_{n+1}} \\ &= X_n + \mu + (M_{n+1} - M_n)\,. \end{aligned} \tag{0.7}$$

Si nota che $(M_n)_{n\in\mathbb{N}_0}$ è una martingala rispetto alla filtrazione $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ con $M_0 = 0$. Esistono modelli analoghi alla (0.7) a tempo continuo?

Il termine di "deriva" $(n+1)\mu$ nella (0.7) si potrebbe riscalare come $n \approx t/h$ con $h > 0$ che tende a zero e $n = n(h)$. La difficoltà sta nel trovare un modello limite per la martingala (M_n). Osserviamo però che

$$\mathbf{E}[M_n] = 0, \qquad \operatorname{var}(M_n) = \mathbf{E}\left[\left(\sum_{k=1}^{n}(\xi_k - \mu)\right)^2\right] = n \cdot \sigma^2$$

per l'indipendenza delle ξ_k. Se $(M(t))_{t\geq 0}$ è l'analogo a tempo continuo di (M_n), ci aspetteremmo di avere

$$\operatorname{var}(M(t)) = t \cdot \sigma^2 \quad \text{per ogni } t \geq 0.$$

Poniamo, per $N \in \mathbb{N}$, $t \geq 0$,

$$M^N(t) \doteq \frac{1}{\sqrt{N}} \sum_{k=1}^{\lfloor t\cdot N\rfloor} (\xi_k - \mu).$$

Allora, per $t \geq 0$ fissato,

$$\operatorname{var}\left(M^N(t)\right) = \frac{\lfloor t \cdot N\rfloor}{N}\sigma^2 \overset{N\to\infty}{\longrightarrow} t \cdot \sigma^2$$

e, grazie al teorema del limite centrale,

$$\operatorname{Law}\left(M^N(t)\right) \doteq \mathbf{P} \circ (M^N(t))^{-1} \overset{N\to\infty}{\longrightarrow} N(0, t\cdot\sigma^2),$$

cioè $M^N(t)$ converge in distribuzione (o in legge) per $N \to \infty$ a una gaussiana di media zero e varianza $t \cdot \sigma^2$. Esiste anche un processo stocastico $(M(t))_{t\geq 0}$ e, se esiste, con quali proprietà (ad esempio, regolarità delle traiettorie, proprietà di martingala)? In termini di quel processo M, si potrà dare una caratterizzazione differenziale del modello limite per la passeggiata aleatoria (X_n) della forma

$$dX(t) = \mu dt + dM(t) \ ?$$

Consideriamo un modello per il gioco d'azzardo, simile a quello della passeggiata aleatoria. Siano X_0, ξ_n, $\mathcal{F}_n$ come sopra. Sia $(H_n)_{n\in\mathbb{N}_0}$ una successione di variabili aleatorie reali adattata alla filtrazione $(\mathcal{F}_n)$, cioè H_n è $\mathcal{F}_n$-misurabile per ogni $n \in \mathbb{N}_0$. Poniamo $V_0 \doteq X_0$ e definiamo per ricorsione

$$V_{n+1} \doteq V_n + H_n \cdot \xi_{n+1}, \quad n \in \mathbb{N}_0.$$

Le variabili H_n si possono interpretare come posta di gioco. Se le ξ_n sono a valori in $\{-1, 1\}$, V_n rappresenta il capitale di un giocatore che ha scommesso sull'esito di n lanci indipendenti di una moneta (equilibrata se $\mu = 0$) con poste di gioco $H_0, \ldots, H_{n-1}$ e capitale iniziale V_0. Dalla definizione di $(V_n)_{n\in\mathbb{N}_0}$ abbiamo

$$\begin{aligned} V_{n+1} &= V_n + H_n \cdot \mu + H_n\,(M_{n+1} - M_n) \\ &= V_0 + \mu \left(\sum_{k=0}^{n} H_k\right) + \underbrace{\sum_{k=0}^{n} H_k(M_{k+1} - M_k)}_{\doteq Y_{n+1}}. \end{aligned} \tag{0.8}$$

Poniamo $Y_0 \doteq 0$. La proprietà di martingala di $(M_n)_{n\in\mathbb{N}_0}$ insieme all'ipotesi che $(H_n)_{n\in\mathbb{N}_0}$ sia una successione adattata alla filtrazione $(\mathcal{F}_n)$ implica che $(Y_n)_{n\in\mathbb{N}_0}$ è una martingala rispetto a $(\mathcal{F}_n)$. La successione (Y_n) è l'integrale stocastico discreto della successione adattata (H_n) rispetto alla martingala (M_n). Esiste una costruzione simile di integrale stocastico a tempo continuo? Potremmo utilizzare una tale costruzione per dare senso a un'espressione differenziale della forma

$$dV(t) = \mu H(t)dt + H(t)dM(t) \ ?$$

0.2 Processi stocastici e filtrazioni

Un processo stocastico è una famiglia $(X(t))_{t\in I}$ di variabili aleatorie a valori in uno spazio misurabile comune, tutte definite sullo stesso spazio di probabilità ed indicizzate da un sottoinsieme $I \subseteq [-\infty, \infty]$. Più formalmente:

Definizione 0.1. Sia $I \subseteq [-\infty, \infty]$ non vuoto. Un *processo stocastico* è una terna $((\Omega, \mathcal{F}, \mathbf{P}), (E, \mathcal{E}), (X(t))_{t\in I})$ tale che

- $(\Omega, \mathcal{F}, \mathbf{P})$ è uno spazio di probabilità,
- $(E, \mathcal{E})$ è uno spazio misurabile,
- per ogni $t \in I$, $X(t) \colon \Omega \to E$ è una funzione $\mathcal{F}$-$\mathcal{E}$ misurabile.

L'insieme degli indici I viene detto *insieme dei tempi*. Di solito avremo $I \subseteq [0, \infty)$, in particolare $I = \mathbb{N}_0$ (tempo discreto) oppure $I = [0, \infty)$ (tempo continuo). Diremo a volte *processo* invece di processo stocastico. Lo spazio misurabile $(E, \mathcal{E})$ nella Definizione 0.1 si dice *spazio dei valori* o *spazio degli stati* del processo.

Esempio 0.1 (Passeggiata aleatoria). Siano ξ_k, $k \in \mathbb{N}$, variabili aleatorie indipendenti a valori in $\mathbb{R}^d$ definite su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbf{P})$. Poniamo $X(0) \doteq x \in \mathbb{R}^d$, $X(n) \doteq X(0) + \sum_{k=1}^{n} \xi_k$, $n \in \mathbb{N}$. Allora

$$\big((\Omega, \mathcal{F}, \mathbf{P}), (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)), (X(t))_{t\in\mathbb{N}_0}\big)$$

è un processo stocastico con insieme dei tempi $I = \mathbb{N}_0$.

Esempio 0.2 (Due classi di processi a tempo continuo). Soluzioni di equazioni differenziali ordinarie con dato iniziale aleatorio (cf. Esercizio I.1); processo di Poisson di parametro $\lambda > 0$ (cf. Esercizio I.2).

Richiami [ad esempio, Klenke, 2008]:

1. σ-algebra generata da un sistema di parti (insieme di sottoinsiemi) $\mathcal{C}$ di un insieme non-vuoto E:

   $$\sigma(\mathcal{C}) \doteq \bigcap_{\mathcal{G}\ \sigma\text{-algebra su } E : \mathcal{C} \subseteq \mathcal{G}} \mathcal{G}.$$

   Criterion di misurabilità [Teorema 1.81 in Klenke, 2008, p. 36].
   Caso particolare importante: $\mathcal{B}(E) \doteq \sigma(\mathcal{U})$ σ-algebra dei boreliani di uno spazio topologico $(E, \mathcal{U})$. Se la topologia (sistema degli aperti) $\mathcal{U}$ possiede una base $\mathcal{U}_0$ denumerabile, allora $\sigma(\mathcal{U}) = \sigma(\mathcal{U}_0)$. Esempio: $\mathcal{B}(\mathbb{R})$ σ-algebra generata dagli intervalli aperti (oppure chiusi, o semiaperti) con estremi in $\mathbb{Q}$.

2. Unicità di misure σ-finite: sistema di parti $\cap$-stabile, Teorema di Carathéodory [Teorema 1.41 e Lemma 1.42 in Klenke, 2008, pp. 19-20].

3. Spazio prodotto e σ-algebra prodotto [Sezione 14.1 in Klenke, 2008, pp. 272-275].

4. Indipendenza e misure prodotto [Klenke, 2008, pp. 275-285]. Lemma di Łomnicki e Ulam [Corollario 6.18 in Kallenberg, 2001, p. 117].

Definizione 0.2. Un processo stocastico $X = ((\Omega, \mathcal{F}, \mathbf{P}), (E, \mathcal{E}), (X(t))_{t\in I})$ si dice *misurabile* se I è un sottoinsieme boreliano di $[-\infty, \infty]$ e l'applicazione

$$I \times \Omega \ni (t, \omega) \mapsto X(t, \omega) \in E$$

è $\mathcal{B}(I) \otimes \mathcal{F} - \mathcal{E}$-misurabile.

Se X è un processo misurabile, allora, per ogni $\omega \in \Omega$ fissato, l'applicazione $I \ni t \mapsto X(.,\omega) \in E$ è misurabile. In particolare, integrali della forma $\int_I \varphi(X(t,\omega))dt$ sono ben definiti come integrali di Lebesgue per ogni funzione $\varphi\colon E \to \mathbb{R}$ misurabile e non-negativa (oppure φ misurabile tale che $\varphi(X(.,\omega))$ integrabile su I).

Definizione 0.3. Sia $I \subseteq [-\infty,\infty]$ un intervallo, E uno spazio topologico e $X = ((\Omega,\mathcal{F},\mathbf{P}),(E,\mathcal{B}(E)),(X(t))_{t\in I})$ un processo stocastico. Per $\omega \in \Omega_0$, si dice *traiettoria* del processo X l'applicazione

$$I \ni t \mapsto X(t,\omega) \in E.$$

Un processo $X = ((\Omega,\mathcal{F},\mathbf{P}),(E,\mathcal{E}),(X(t))_{t\in I})$ si dice *continuo* se le sue traiettorie sono continue con probabilità uno, cioè se esiste $\Omega_0 \in \mathcal{F}$ tale che $\mathbf{P}(\Omega_0) = 1$ e, per ogni $\omega \in \Omega$, $X(.,\omega)$ è una mappa continua $I \to E$. In analogia si definiscono processi *continui a destra*, *continui a sinistra*, *hölderiani* et cetera. A volte bisogna distinguere il caso in cui una data proprietà (come la continuità) vale per quasi tutte le traiettorie, da quello in cui vale certamente, quindi per tutte le traiettorie.

Siano $X = ((\Omega,\mathcal{F},\mathbf{P}),(E,\mathcal{E}),(X(t))_{t\in I})$, $Y = ((\Omega,\mathcal{F},\mathbf{P}),(E,\mathcal{E}),(Y(t))_{t\in I})$ due processi stocastici con lo stesso spazio degli stati $(E,\mathcal{E})$, lo stesso insieme dei tempi I e definiti sullo stesso spazio di probabilità. Si hanno le seguenti nozioni di uguaglianza.

Definizione 0.4. a) X, Y si dicono *equivalenti* o di avere *le stesse distribuzioni finito-dimensionali* se per ogni $n \in N$, ogni scelta di $t_1,\dots,t_n \in I$, ogni $A \in \otimes^n\mathcal{E}$,

$$\mathbf{P}\left((X(t_1),\dots,X(t_n)) \in A\right) = \mathbf{P}\left((Y(t_1),\dots,Y(t_n)) \in A\right).$$

b) X, Y si dicono una *modificazione* l'uno dell'altro se per ogni $t \in I$

$$X(t) = Y(t) \quad \mathbf{P}\text{-quasi certamente.}$$

c) X, Y si dicono *indistinguibili* se

$$\mathbf{P}\left(X(t) = Y(t) \text{ per ogni } t \in I\right) = 1.$$

La nozione a) di equivalenza / stesse distribuzioni finito-dimensionali si applica anche a processi definiti su spazi di probabilità diversi.

Implicazioni: c) $\Rightarrow$ b) e b) $\Rightarrow$ a).

Attenzione: In generale, per I non denumerabile, b) non implica c). Ad esempio, ponendo $(\Omega,\mathcal{F},\mathbf{P}) \doteq ([0,1],\mathcal{B}([0,1]),\lambda_{|[0,1]})$, $(E,\mathcal{E}) \doteq (\mathbb{R},\mathcal{B}(\mathbb{R}))$, $I \doteq [0,1]$ e $X(t) \doteq 0$, $Y(t,\omega) \doteq \mathbf{1}_{\{t\}}(\omega)$ per $t \in [0,1]$, $\omega \in \Omega = [0,1]$, si ha che X, Y sono una modificazione l'uno dell'altro, ma $\mathbf{P}\left(X(t) = Y(t) \text{ per ogni } t \in I\right) = 0$.

L'implicazione "b) $\Rightarrow$ c)" è valida per processi con traiettorie abbastanza regolari. Sia I un intervallo. Se X, Y sono processi con traiettorie continue a destra (oppure continue a sinistra), allora b) implica c) per X e Y.

Un processo stocastico $(X(t))_{t\in I}$ definito su $(\Omega, \mathcal{F}, \mathbf{P})$ con spazio degli stati $(E, \mathcal{E})$ può essere visto come una variabile aleatoria a valori nello spazio prodotto $(\times^I E, \otimes^I \mathcal{E})$. L'applicazione

$$X : \Omega \to (X(t,\omega))_{t\in I} \in \times^I E$$

è infatti $\mathcal{F}$-$\otimes^I\mathcal{E}$-misurabile. Di conseguenza, X induce una misura di probabilità $\mathbf{P}_X \doteq \mathbf{P} \circ (X)^{-1}$ su $\otimes^I\mathcal{E}$:

$$\mathbf{P}_X(B) \doteq \mathbf{P}\left(\{\omega \in \Omega : X(.,\omega) \in B\}\right)$$

per ogni $B \in \otimes^I\mathcal{E}$. La misura di probabilità $\mathbf{P}_X$ si dice *legge* / *distribuzione* del processo X. La legge di un processo è determinata dalle sue distribuzioni finito-dimensionali.

Proposizione 0.1. *Siano $(X(t))_{t\in I}$, $(Y(t))_{t\in I}$ processi stocastici definiti rispettivamente su $(\Omega, \mathcal{F}, \mathbf{P})$ e $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}})$ con spazio degli stati $(E, \mathcal{E})$ comune.*

Se X, Y hanno le stesse distribuzioni finito-dimensionali, allora $\mathbf{P}_X = \tilde{\mathbf{P}}_Y$.

Dimostrazione. Sia $\mathcal{C}$ il sistema degli insiemi cilindrici finito-dimensionali di $(\times^I E, \otimes^I \mathcal{E})$, cioè

$$\mathcal{C} \doteq \left\{\{f \in \times^I E : (f(t_1), \ldots, f(t_n)) \in B\} : n \in \mathbb{N},\, t_1, \ldots, t_n \in I,\, B \in \otimes^n \mathcal{E}\right\}.$$

Il sistema $\mathcal{C}$ genera la σ-algebra prodotto $\otimes^I\mathcal{E}$. In più, $\mathcal{C}$ è $\cap$-stabile (stabile per l'intersezione finita). Per il teorema di Caratheodory, due misure di probabilità su $\otimes^I\mathcal{E}$ sono quindi identiche se (e solo se) coincidono su $\mathcal{C}$. Siano $n \in \mathbb{N}$, $t_1, \ldots, t_n \in I$, $B \in \otimes^n\mathcal{E}$. Allora

$$\begin{aligned}\mathbf{P}_X\left(\{f \in \times^I E : (f(t_1), \ldots, f(t_n)) \in B\}\right) &= \mathbf{P}\left((X(t_1), \ldots, X(t_n)) \in B\right)\\ &= \mathbf{P}_{(X(t_1),\ldots,X(t_n))}(B)\end{aligned}$$

e analogamente per $\tilde{\mathbf{P}}_Y$. Per ipotesi, X, Y hanno le stesse distribuzioni finito-dimensionali, quindi $\mathbf{P}_{(X(t_1),\ldots,X(t_n))} = \tilde{\mathbf{P}}_{(Y(t_1),\ldots,Y(t_n))}$. Di conseguenza, $\mathbf{P}_X$, $\tilde{\mathbf{P}}_Y$ coincidono su $\mathcal{C}$, quindi $\mathbf{P}_X = \tilde{\mathbf{P}}_Y$. □

Al posto del sistema $\mathcal{C}$ degli insiemi cilindrici finito-dimensionali ci si può limitare al sistema dei rettangoli finito-dimensionali di $\times^I E$, cioè a

$$\mathcal{R} \doteq \left\{\{f \in \times^I E : f(t_1) \in B_1, \ldots, f(t_n) \in B_n\} : n \in \mathbb{N},\, t_1, \ldots, t_n \in I,\, B_1, \ldots, B_n \in \mathcal{E}\right\}.$$

Infatti, anche $\mathcal{R}$ genera $\otimes^I\mathcal{E}$ ed è $\cap$-stabile.

Un concetto importante, legato all'interpretazione dell'insieme degli indici I come insieme dei tempi, è quello di filtrazione.

Definizione 0.5. Sia $(\Omega, \mathcal{F})$ uno spazio misurabile, $I \subseteq [-\infty, \infty]$. Una *filtrazione* in $\mathcal{F}$ è una famiglia $(\mathcal{F}_t)_{t\in I}$ di sotto-σ-algebre di $\mathcal{F}$ tale che $\mathcal{F}_s \subseteq \mathcal{F}_t$ per ogni $s, t \in I$ con $s \leq t$.

Una filtrazione può essere interpretata come rappresentazione del flusso di informazione sulle quantità aleatorie presenti.

Esempio 0.3. Sia $(X(t))_{t\in I}$ un processo stocastico definito su $(\Omega, \mathcal{F}, \mathbf{P})$. Allora

$$\mathcal{F}_t^X \doteq \sigma(X(s) : s \leq t), \quad t \in I.$$

definisce una filtrazione in $\mathcal{F}$, la *filtrazione naturale* del processo X.

Definizione 0.6. Sia $(X(t))_{t\in I}$ un processo stocastico definito su $(\Omega, \mathcal{F}, \mathbf{P})$, sia $(\mathcal{F}_t)_{t\in I}$ una filtrazione in $\mathcal{F}$. Il processo X si dice *adattato* a $(\mathcal{F}_t)_{t\in I}$ se $X(t)$ è $\mathcal{F}_t$-misurabile per ogni $t \in I$.

Un processo X è sempre adattato alla propria filtrazione naturale.

Definizione 0.7. Siano $(\Omega, \mathcal{F})$ uno spazio misurabile, $(\mathcal{F}_t)_{t\in I}$ una filtrazione in $\mathcal{F}$. Per $t \in I$, si definiscono

$$\mathcal{F}_{t+} \doteq \bigcap_{s>t} \mathcal{F}_s, \qquad \mathcal{F}_{t-} \doteq \sigma\left(\bigcup_{s<t} \mathcal{F}_s\right),$$

e si dice che $\mathcal{F}_{t+}$ è la *σ-algebra degli eventi immediatamente dopo t* e $\mathcal{F}_{t-}$ la *σ-algebra degli eventi prima di t*.

Una filtrazione $(\mathcal{F}_t)_{t\in I}$ si dice *continua a destra* (oppure *continua a sinistra*) se, per ogni $t \in I$, $\mathcal{F}_t = \mathcal{F}_{t+}$ (oppure $\mathcal{F}_t = \mathcal{F}_{t-}$).

Definizione 0.8. Sia $I \subseteq [0, \infty]$ un boreliano. Siano $X = ((\Omega, \mathcal{F}, \mathbf{P}), (E, \mathcal{E}), (X(t))_{t\in I})$ un processo stocastico e $(\mathcal{F}_t)_{t\in I}$ una filtrazione in $\mathcal{F}$. Il processo X si dice *progressivamente misurabile* rispetto a $(\mathcal{F}_t)_{t\in I}$ se, per ogni $t \in I$, l'applicazione

$$I \cap [0, t] \times \Omega \ni (s, \omega) \mapsto X(s, \omega) \in E$$

è $\mathcal{B}(I \cap [0, t]) \otimes \mathcal{F}_t$-$\mathcal{E}$-misurabile.

Un processo progressivamente misurabile rispetto a una filtrazione è misurabile ed è adattato a quella filtrazione. Se le traiettorie del processo sono sufficientemente regolari, allora vale anche l'implicazione conversa [cf. Karatzas and Shreve, 1991, p. 5]. Per semplicità supponiamo $I = [0, T]$ oppure $I = [0, \infty)$.

Proposizione 0.2. *Siano $(\Omega, \mathcal{F}, \mathbf{P})$ uno spazio di probabilità, $(\mathcal{F}_t)_{t\in I}$ una filtrazione in $\mathcal{F}$. Sia $(X(t))_{t\in I}$ un processo stocastico definito su $(\Omega, \mathcal{F}, \mathbf{P})$ a valori in uno spazio topologico E con tutte le traiettorie continue a destra (oppure tutte continue a sinistra). Se X è adattato a $(\mathcal{F}_t)$, è anche progressivamente misurabile rispetto a $(\mathcal{F}_t)$.*

Dimostrazione. Supponiamo che le traiettorie di X siano tutte continue a destra (il caso di traiettorie continue a sinistra è analogo). Sia $t_0 \in I$ fissato. Dobbiamo mostrare che la mappa $[0, t_0] \times \Omega \ni (t, \omega) \mapsto X(t, \omega) \in E$ è $\mathcal{B}([0, t_0]) \otimes \mathcal{F}_{t_0}$-$\mathcal{E}$-misurabile. Per $n \in \mathbb{N}$ poniamo

$$X^{(n)}(t, \omega) \doteq \begin{cases} X\left(\frac{k+1}{2^n} t_0, \omega\right) & \text{if } t \in (\frac{k}{2^n} t_0, \frac{k+1}{2^n} t_0] \text{ per un } k \in \{0, \dots, 2^n - 1\}, \\ X(0, \omega) & \text{if } t = 0. \end{cases}$$

Allora la mappa $[0,t_0]\times\Omega \ni (t,\omega)\mapsto X^{(n)}(t,\omega)\in E$ è $\mathcal{B}([0,t_0])\otimes\mathcal{F}_{t_0}$-$\mathcal{E}$-misurabile poiché, per $A\in\mathcal{E}$,

$$\begin{aligned}
&\left\{(t,\omega)\in[0,t_0]\times\Omega : X^{(n)}(t,\omega)\in A\right\}\\
&\qquad = \{0\}\times\underbrace{\{\omega\in\Omega : X(t_0,\omega)\in A\}}_{\in\mathcal{F}_{t_0}}\cup\\
&\qquad\quad \bigcup_{k=0}^{2^n-1}\left(\tfrac{k}{2^n}t_0,\tfrac{k+1}{2^n}t_0\right]\times\underbrace{\left\{\omega\in\Omega : X\left(\tfrac{k+1}{2^n}t_0,\omega\right)\in A\right\}}_{\in\mathcal{F}_{2^{-n}(k+1)t_0}\subseteq\mathcal{F}_{t_0}}\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \in\mathcal{B}([0,t_0])\otimes\mathcal{F}_{t_0}.
\end{aligned}$$

Come processo, $X^{(n)}$ ha traiettorie costanti a tratti e continue a sinistra. Grazie a questa continuità e quella a destra delle traiettorie di X, per ogni $(t,\omega)\in[0,t_0]\times\Omega$,

$$X^{(n)}(t,\omega)\overset{n\to\infty}{\longrightarrow}X(t,\omega).$$

Ne segue che X come limite puntuale di funzioni $\mathcal{B}([0,t_0])\otimes\mathcal{F}_{t_0}$-$\mathcal{E}$-misurabili gode della stessa misurabilità. □

Non-esistenza del rumore bianco come processo stocastico misurabile [cf. Kallianpur, 1980, p. 10]

Proposizione 0.3. *Sia $(X(t))_{t\in[0,1]}$ un processo stocastico a valori reali definito su $(\Omega,\mathcal{F},\mathbf{P})$. Sia $\sigma\in[0,\infty)$. Se X è misurabile e tale che, per ogni $t\in[0,1]$, $\mathbf{E}[X(t)]=0$, $\mathbf{E}\left[X(t)^2\right]=\sigma^2$, $X(t)$ è indipendente da $X(s)$ per ogni $s\neq t$, allora $\sigma=0$ e X è una modificazione del processo costante uguale a zero.*

Dimostrazione. Denotiamo con λ la misura di Lebesgue su $\mathcal{B}([0,1])$. Poiché X è misurabile, X^2 è una funzione non-negativa definita su $[0,1]\times\Omega$ e misurabile rispetto a $\mathcal{B}([0,1])\otimes\mathcal{F}$. Grazie al teorema di Fubini-Tonelli ed all'ipotesi $\mathbf{E}\left[X(t)^2\right]=\sigma^2$ per ogni $t\in[0,1]$,

$$\mathbf{E}\left[\int_0^1 X^2(t)dt\right]=\int_0^1\mathbf{E}\left[X^2(t)\right]dt=\sigma^2<\infty.$$

Questo implica in particolare che $\int_0^1|X(t)|^2dt<\infty$ $\mathbf{P}$-quasi certamente; quindi per la disuguaglianza di Hölder e la monotonia di $t\mapsto\int_0^t|X(s)|ds$, abbiamo che $\int_0^t|X(s)|ds<\infty$ per ogni $t\in[0,1]$ con probabilità uno. Ne deduciamo, usando due volte il teorema di Fubini-Tonelli, che per ogni $t\in[0,1]$

$$\mathbf{E}\left[\left(\int_0^t X(s)ds\right)^2\right]=\mathbf{E}\left[\left(\int_0^t\int_0^t X(\tilde{s})X(s)ds\mathrm{d}\tilde{s}\right)\right]=\int_0^t\int_0^t\mathbf{E}\left[X(\tilde{s})X(s)\right]ds\,d\tilde{s}.$$

Per l'ipotesi di indipendenza, $\mathbf{E}[X(\tilde{s})X(s)]=\mathbf{1}_{\{s\}}(\tilde{s})\sigma^2$ per $s,\tilde{s}\in[0,1]$; di conseguenza,

$$\mathbf{E}\left[\left(\int_0^t X(s)ds\right)^2\right]=0\quad\text{per ogni } t\in[0,1].$$

Questo implica che $X(t,\omega) = 0$ per $\lambda \otimes \mathbf{P}$-quasi ogni $(t,\omega) \in [0,1] \times \Omega$. Infatti, per ogni $t \in [0,1]$ esiste $\Omega_t \in \mathcal{F}$ tale che $\mathbf{P}(\Omega_t) = 1$ e

$$\int_0^t X(s,\omega)ds = 0 \quad \text{per ogni } \omega \in \Omega_t.$$

Poniamo $\tilde{\Omega} \doteq \bigcap_{t\in[0,1]\cap\mathbb{Q}} \Omega_t$. Allora $\mathbf{P}(\tilde{\Omega}) = 1$ e, per ogni $\omega \in \tilde{\Omega}$, $X(.,\omega)$ è integrabile con $\int_0^t X(s,\omega)ds = 0$ per ogni $t \in [0,1] \cap \mathbb{Q}$. Ne segue, grazie alla continuità dell'integrale di Lebesgue, che la funzione F_ω definita per $\omega \in \tilde{\Omega}$ mediante

$$F_\omega(t) \doteq \int_0^t X(s,\omega)ds, \quad t \in [0,1],$$

è identicamente uguale a zero. Dall'altra parte, per il teorema fondamentale del calcolo nella versione di Lebesgue-Vitali,

$$\frac{d}{dt}F_\omega(t) = X(t,\omega) \quad \lambda\text{-quasi ovunque},$$

ma, siccome $F_\omega \equiv 0$, $\frac{d}{dt}F_\omega(t) = 0$ per ogni $t \in [0,1]$. Ne segue che per ogni $\omega \in \tilde{\Omega}$ esiste $B_\omega \in \mathcal{B}([0,1])$ tale che $\lambda(B_\omega) = 1$ e $X(t,\omega) = 0$ per ogni $t \in B_\omega$. Usando di nuovo il teorema di Fubini-Tonelli, ne deduciamo che

$$\begin{aligned}
\lambda \otimes \mathbf{P}\left(\{(t,\omega) : X(t,\omega) = 0\}\right) &= \int_\Omega \int_0^1 \mathbf{1}_{\{0\}}(X(t,\omega))dt\,\mathbf{P}(d\omega) \\
&= \int_{\tilde{\Omega}} \int_0^1 \mathbf{1}_{B_\omega}(t) \cdot \mathbf{1}_{\{0\}}(X(t,\omega))dt\,\mathbf{P}(d\omega) \\
&= \int_{\tilde{\Omega}} \lambda(B_\omega)\,\mathbf{P}(d\omega) \\
&= \mathbf{P}(\tilde{\Omega}) \qquad = 1.
\end{aligned}$$

Di conseguenza, $X = 0$ $\lambda \otimes \mathbf{P}$-quasi certamente. Questo implica

$$\mathbf{E}\left[\int_0^1 X^2(t)dt\right] = 0,$$

quindi $\sigma^2 = 0$. □

Osservazione 0.1. Nella Proposizione 0.3, si è scelto l'intervallo $[0,1]$ per l'insieme dei tempi I solo per semplicità di notazioni. L'ipotesi di varianza σ^2 finita non è essenziale per il risultato "negativo." Infatti, se esistesse un processo reale misurabile con componenti indipendenti a due a due ed identicamente distribuite con varianza infinita, i valori di quel processo si potrebbero troncare e centrare generando in quel modo un processo a media zero e varianza finita, che però per la Proposizione 0.3 sarebbe una versione del processo costante uguale a zero.

0.3 Nozioni di convergenza

Sono rilevanti varie nozioni di convergenza per successioni di variabili aleatorie. Sia (E, d) uno spazio metrico, e siano X_n, $n \in \mathbb{N}$, X variabili aleatorie a valori in $(E, \mathcal{B}(E))$ definite su $(\Omega, \mathcal{F}, \mathbf{P})$.

Definizione 0.9. Si dice che la successione $(X_n)_{n\in\mathbb{N}}$ converge a X per $n \to \infty$

– *quasi certamente* se esiste $\Omega_0 \in \mathcal{F}$ tale che $\mathbf{P}(\Omega_0) = 1$ e per ogni $\omega \in \Omega_0$

$$X_n(\omega) \overset{n\to\infty}{\longrightarrow} X(\omega);$$

– *in probabilità* se per ogni $\varepsilon > 0$

$$\mathbf{P}\left(d(X_n, X) > \varepsilon\right) \overset{n\to\infty}{\longrightarrow} 0;$$

– *in L^p* con $p \in [1, \infty)$ se

$$\mathbf{E}\left[d(X_n, X)^p\right] \overset{n\to\infty}{\longrightarrow} 0;$$

– *in legge* (o *in distribuzione*) se per ogni $f \in \mathbf{C}_b(E)$

$$\mathbf{E}\left[f(X_n)\right] \overset{n\to\infty}{\longrightarrow} \mathbf{E}\left[f(X)\right].$$

Le nozioni di convergenza in Def. 0.9 dipendono chiaramente dalla misura di probabilità. La convergenza in legge ha senso anche per variabili aleatorie definite su spazi di probabilità diversi. La variabile limite X nel caso di convergenza quasi certa, in probabilità o in L^p è determinata con probabilità uno. La convergenza in legge dipende invece, come suggerisce il nome, solo dalle leggi delle variabili aleatorie; essa è infatti equivalente alla seguente nozione di convergenza per misure di probabilità.

Definizione 0.10. Siano μ_n, $n \in \mathbb{N}$, μ misure di probabilità su $\mathcal{B}(E)$. Si dice che la successione $(\mu_n)_{n\in\mathbb{N}}$ *converge debolmente* a μ per $n \to \infty$ se per ogni $f \in \mathbf{C}_b(E)$

$$\int_E f(x)\mu_n(dx) \overset{n\to\infty}{\longrightarrow} \int_E f(x)\mu_n(dx).$$

Una misura di probabilità μ su $\mathcal{B}(E)$ con (E, d) uno spazio metrico è univocamente determinata dagli integrali $\int_E f(x)\mu(dx)$ su tutte le funzioni $f\colon E \to \mathbb{R}$ limitate e uniformemente continue; di conseguenza, μ è determinata dagli integrali $\int_E f(x)\mu(dx)$, $f \in \mathbf{C}_b(E)$.

Confronto di nozioni di convergenza. Supponiamo che lo spazio metrico (E, d) dei valori sia *separabile* (i.e., esiste un sottoinsieme denumerabile e denso in E). Allora

convergenza quasi-certa $\Rightarrow$ convergenza in probabilità $\Rightarrow$ convergenza in legge.

Convergenza in legge implica la convergenza in probabilità se la misura limite è una misura di Dirac (la variabile limite è quindi quasi-certamente costante). Se $E = \mathbb{R}^d$ (oppure uno spazio di Banach), allora

$$\text{convergenza in } L^p,\ p \geq 1 \Rightarrow \text{convergenza in } L^1 \Rightarrow \text{convergenza in probabilità.}$$

La convergenza dominata fornisce condizioni in cui la convergenza in probabilità implica la convergenza in L^1 (in L^p).

Il seguente lemma è spesso utile nelle dimostrazioni di convergenza.

Lemma 0.1 (Borel-Cantelli)**.** *Sia $(\Omega, \mathcal{F}, \mathbf{P})$ uno spazio di probabilità e sia $(A_n)_{n\in\mathbb{N}}$ una successione di eventi in $\mathcal{F}$. Poniamo*

$$A \doteq \limsup_{n\to\infty} A_n \doteq \{\omega \in \Omega : \omega \in A_n \text{ per infiniti indici } n\}.$$

(i) Se $\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty$, allora $\mathbf{P}(A) = 0$.

(ii) Se $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty$ e gli eventi A_n, $n \in \mathbb{N}$, sono a due a due indipendenti, allora $\mathbf{P}(A) = 1$.

Dimostrazione. (i). Supponiamo che $\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty$. Sia $\varepsilon > 0$. Per la sommabilità, esiste $n_0(\varepsilon) \in \mathbb{N}$ tale che $\sum_{n=n_0(\varepsilon)}^{\infty} \mathbf{P}(A_n) \leq \varepsilon$. Per la sua definizione, $A \subseteq \bigcup_{n=k}^{\infty} A_n$ per ogni $k \in \mathbb{N}$. Di conseguenza,

$$\mathbf{P}(A) \leq \mathbf{P}\left(\bigcup_{n=n_0(\varepsilon)}^{\infty} A_n\right) \leq \sum_{n=n_0(\varepsilon)}^{\infty} \mathbf{P}(A_n) \leq \varepsilon.$$

Siccome $\varepsilon > 0$ era arbitrario, si ha $\mathbf{P}(A) = 0$.

(ii). Supponiamo che $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty$ e che gli eventi A_k, $k \in \mathbb{N}$, siano indipendenti a due a due. Poniamo $S_n \doteq \sum_{k=1}^{n} \mathbf{1}_{A_k}$, $S \doteq \sum_{k=1}^{\infty} \mathbf{1}_{A_k}$. Dobbiamo mostrare che $\mathbf{P}(S = \infty) = 1$. Per l'indipendenza a due a due, abbiamo per ogni $n \in \mathbb{N}$ che

$$\operatorname{var}(S_n) = \sum_{k=1}^{n} \operatorname{var}(\mathbf{1}_{A_k}) = \sum_{k=1}^{n} \left(\mathbf{E}[\mathbf{1}_{A_k}] - \mathbf{E}[\mathbf{1}_{A_k}]^2\right) \leq \sum_{k=1}^{n} \mathbf{E}[\mathbf{1}_{A_k}] = \mathbf{E}[S_n].$$

Utilizzando la disuguaglianza di Chebyshev, otteniamo

$$\begin{aligned}
\mathbf{P}\left(S_n < \frac{1}{2}\mathbf{E}[S_n]\right) &= \mathbf{P}\left(\mathbf{E}[S_n] - S_n > \frac{1}{2}\mathbf{E}[S_n]\right) \\
&\leq \mathbf{P}\left(|\mathbf{E}[S_n] - S_n| \geq \frac{1}{2}\mathbf{E}[S_n]\right) \\
&\leq \frac{4\operatorname{var}(S_n)}{\mathbf{E}[S_n]^2} \\
&\leq \frac{4}{\mathbf{E}[S_n]}.
\end{aligned}$$

Per l'ipotesi di non-sommabilità, $\mathbf{E}[S_n] = \sum_{k=1}^{n} \mathbf{P}(A_k) \nearrow \infty$ per $n \to \infty$. Di conseguenza, $\mathbf{P}\left(S_n < \frac{1}{2}\mathbf{E}[S_n]\right) \to 0$ e $\mathbf{P}\left(S_n \geq \frac{1}{2}\mathbf{E}[S_n]\right) \to 1$ per $n \to \infty$. Siccome $S_n \nearrow S$ per $n \to \infty$, ne segue che

$$\mathbf{P}\left(S \geq \frac{1}{2}\mathbf{E}[S_n]\right) \geq \mathbf{P}\left(S_n \geq \frac{1}{2}\mathbf{E}[S_n]\right) \stackrel{n\to\infty}{\longrightarrow} 1.$$

Poiché $\mathbf{E}[S_n] \nearrow \infty$, questo implica $\mathbf{P}(S = \infty) = 1$, quindi $\mathbf{P}(A) = 1$. □

Caratterizzazione e convergenza in legge di variabili aleatorie a valori in $\mathbb{R}^d$: *funzioni caratteristiche* [ad esempio, Klenke, 2008, pp. 297-316].

0.4 Leggi normali e famiglie gaussiane

Definizione 0.11. Una misura di probabilità $\mathcal{B}(\mathbb{R})$ si dice *legge normale* (o *legge gaussiana*) di media $b \in \mathbb{R}$ e varianza $\sigma^2 > 0$ e si indica con $N(b, \sigma^2)$ se essa è assolutamente continua rispetto alla misura di Lebesgue con densità

$$\mathbb{R} \ni y \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-b)^2}{2\sigma^2}}.$$

La legge normale $N(b, \sigma^2)$ ha funzione caratteristica

$$\varphi_{N(b,\sigma^2)}(y) = e^{\sqrt{-1}b\cdot y} e^{-\frac{\sigma^2}{2}y^2}, \quad y \in \mathbb{R}.$$

Alcune proprietà elementari delle leggi normali univariate:

- $N(b, \sigma^2)$ è simmetrica rispetto a b.
- Leggi normali possiedono momenti di ogni ordine $p \in (0, \infty)$.
- Convoluzione: $N(b, \sigma^2) * N(\tilde{b}, \tilde{\sigma}^2) = N(b + \tilde{b}, \sigma^2 + \tilde{\sigma}^2)$; equivalentemente, se $X \sim N(b, \sigma^2)$, $Y \sim N(\tilde{b}, \tilde{\sigma}^2)$ sono due variabili gaussiane indipendenti, allora $X + Y$ ha legge $N(b + \tilde{b}, \sigma^2 + \tilde{\sigma}^2)$.
- Trasformazioni lineari-affini: Se $X \sim N(b, \sigma^2)$, $\tilde{b} \in R$, $c \in R$, allora $c\cdot X + \tilde{b}$ ha legge $N(b + \tilde{b}, c^2 \cdot \sigma^2)$.
- Convergenza per $\sigma^2 \to 0$: $N(b, \sigma^2)$ converge debolmente alla di misura di Dirac δ_b per $\sigma^2 \to 0$. Perciò si definisce $N(b, 0) \doteq \delta_b$.

Leggi normali d-dimensionali si definiscono attraverso la funzione caratteristica, la quale è ben definita anche nel caso "degenere" quando la legge non è assolutamente continua rispetto alla misura di Lebesgue in $\mathbb{R}^d$.

Definizione 0.12. Una misura di probabilità $\mathcal{B}(\mathbb{R}^d)$ si dice *legge normale* (o *legge gaussiana*) di media $b \in \mathbb{R}^d$ e matrice di covarianza $\Gamma \in \mathbb{R}^{d\times d}$ semidefinita positiva e si indica con $N(b, \Gamma)$ se la sua funzione caratteristica è data da

$$\mathbb{R}^d \ni y \mapsto e^{\sqrt{-1}\langle b,y\rangle} \cdot e^{-\frac{1}{2}\langle \Gamma y, y\rangle} \in \mathbb{C}.$$

Proprietà delle leggi normali multivariate. Siano $b \in \mathbb{R}^d$ un vettore, $\Gamma \in \mathbb{R}^{d\times d}$ una matrice semidefinita positiva, e sia X una variabile aleatoria di legge $N(b,\Gamma)$.

1) Trasformazioni lineari-affini. Siano $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times d}$. Allora $Y \doteq AX + c$ ha legge $N(Ab + c, A\Gamma A^{\mathsf{T}})$.

2) Riduzione a normali standard. Esiste Z di legge normale standard d-dimensionale (cioè $Z \sim N(0, \mathrm{Id}_d)$) tale che $X = b + \sqrt{\Gamma}Z$.

3) Densità. La legge di X è assolutamente continua rispetto alla misura di Lebesgue in $\mathbb{R}^d$ se e solo se Γ è invertibile, e in quel caso la densità è data da

$$f_X(x) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(\Gamma)}} \exp\left(-\frac{1}{2}\langle \Gamma^{-1}(x-b), x-b\rangle\right), \quad x \in \mathbb{R}^d.$$

4) Leggi marginali. Una variabile aleatorie Z a valori in $\mathbb{R}^d$ è gaussiana se e solo se $\langle Z, y\rangle$ è una gaussiana uno-dimensionale per ogni $y \in \mathbb{R}^d$.

5) Indipendenza:

   a) Somme. Siano X, Y gaussiane a valori in $\mathbb{R}^d$ con $X \sim N(b,\Gamma)$, $Y \sim N(\tilde{b},\tilde{\Gamma})$. Se X, Y sono indipendenti, allora $X+Y$ ha legge $N(b+\tilde{b}, \Gamma+\tilde{\Gamma})$.

   b) Correlazione / covarianza. Siano X, Y con legge congiunta gaussiana, cioè tali che (X,Y) è gaussiana. Allora X, Y sono indipendenti se e solo se X, Y non sono correlate.

Attenzione: Né la legge congiunta né la legge della somma di due gaussiane è necessariamente gaussiana. (Contro-)esempio: Siano X, Z variabili aleatorie indipendenti tali che X è normale standard, Z di Rademacher, cioè $\mathbf{P}(Z = -1) = \frac{1}{2} = \mathbf{P}(Z = 1)$. Allora anche $Y \doteq Z \cdot X$ è una gaussiana standard, ma né $X + Y$ né (X,Y) sono gaussiane, e X, Y non sono indipendenti.

Proposizione 0.4 (Convergenza di variabili gaussiane)**.** *Sia $(X_n)_{n\in\mathbb{N}}$ una successione di variabili aleatorie gaussiane a valori in $\mathbb{R}^d$ tale che, per $n \to \infty$, $X_n \to X$ in legge per una variabile aleatoria X. Allora anche X è gaussiana.*

Dimostrazione. Sia φ la funzione caratteristica (della legge) di X, e sia φ_n la funzione caratteristica (della legge) di X_n, $n \in \mathbb{N}$. Per l'ipotesi di convergenza in legge e grazie al teorema di continuità di Lévy [cf. Teorema 15.23 in Klenke, 2008, p. 309], abbiamo $\varphi_n \to \varphi$ per $n \to \infty$ uniformemente sui compatti di $\mathbb{R}^d$. Per ipotesi, X_n ha distribuzione gaussiana; quindi esistono $b_n \in \mathbb{R}^d$, $\Gamma_n \in \mathbb{R}^d \times \mathbb{R}^d$ semidefinita positiva tali che

$$\varphi_n(y) = e^{\sqrt{-1}\langle b_n, y\rangle} \cdot e^{-\frac{1}{2}\langle \Gamma_n y, y\rangle}, \quad y \in \mathbb{R}^d.$$

Poiché le funzioni caratteristiche sono limitate in modulo da uno, basta mostrare che esistono una sotto-successione $(n_j)_{j\in\mathbb{N}}$, un vettore $b \in \mathbb{R}^d$ e una matrice $\Gamma_n \in \mathbb{R}^d \times \mathbb{R}^d$ tali che $n_j \nearrow \infty$, $b_{n_j} \to b$ e $\Gamma_{n_j} \to \Gamma$ per $j \to \infty$ (la matrice Γ sarà semidefinita positiva come limite di matrici semidefinite positive). E' quindi sufficiente verificare che le successioni $(b_n)_{n\in\mathbb{N}}$ e $(\Gamma_n)_{n\in\mathbb{N}}$ sono limitate. Supponiamo per assurdo che $(\Gamma_n)_{n\in\mathbb{N}}$ sia illimitata. Allora esisterebbero una sotto-successione $(n_i)_{i\in\mathbb{N}}$ e un vettore $v \in \mathbb{R}^d \setminus \{0\}$ tali che $n_i \nearrow \infty$ e $\sigma_i^2 \doteq \mathrm{var}(\langle X_{n_i}, v\rangle) \nearrow \infty$ per $i \to \infty$. Poiché $\langle X_{n_i}, v\rangle$ è una gaussiana uno-dimensionale di media $\langle b_{n_i}, v\rangle$ e varianza σ_i^2, si avrebbe, per ogni $c \geq 0$,

$$\begin{aligned}
\mathbf{P}\left(\langle X_{n_i}, v\rangle \in [-c,c]\right) &\leq \mathbf{P}\left(\langle X_{n_i}, v\rangle - \langle b_{n_i}, v\rangle \in [-c,c]\right)\\
&= \frac{1}{\sqrt{2\pi\sigma_i^2}}\int_{-c}^{c} e^{-\frac{x^2}{2\sigma_i^2}}\,dx\\
&\leq \frac{2c}{\sqrt{2\pi\sigma_i^2}}\\
&\overset{i\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Ma questo implicherebbe, per la convergenza in legge[1], $\mathbf{P}\left(\langle X, v\rangle \in [-c,c]\right) = 0$ per ogni $c \geq 0$ — in contraddizione col fatto che $\mathbf{P} \circ \langle X, v\rangle^{-1}$ sia una misura di probabilità sui boreliani di $\mathbb{R}$. Ne segue che $(\Gamma_n)_{n\in\mathbb{N}}$ è una successione limitata in $\mathbb{R}^d \times \mathbb{R}^d$. Supponiamo ora per assurdo che $(b_n)_{n\in\mathbb{N}}$ sia illimitata. Allora esisterebbero una sotto-successione $(n_i)_{i\in\mathbb{N}}$, una costante $K \in (0,\infty)$ e un vettore $v \in \mathbb{R}^d \setminus \{0\}$ tali che $n_i \nearrow \infty$ e $m_i \doteq \langle b_{n_i}, v\rangle \nearrow \infty$ per $i \to \infty$, mentre $\sup_{i\in\mathbb{N}} \sigma_i^2 \leq K$, dove $\sigma_i^2 \doteq \mathrm{var}(\langle X_{n_i}, v\rangle)$. Ma allora si avrebbe, per ogni $c \geq 0$,

$$\begin{aligned}
\mathbf{P}\left(\langle X_{n_i}, v\rangle \in [-c,c]\right) &= \frac{1}{\sqrt{2\pi\sigma_i^2}}\int_{-c}^{c} e^{-\frac{(x-m_i)^2}{2\sigma_i^2}}\,dx\\
&\leq \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{c-m_i}{\sigma_i}} e^{-\frac{y^2}{2}}\,dy\\
&\overset{i\to\infty}{\longrightarrow} 0
\end{aligned}$$

perché $(c - m_i)/\sigma_i \to -\infty$ per $i \to \infty$. Ma questo porterebbe alla stessa contraddizione di prima. Di conseguenza, le successioni $(b_n)_{n\in\mathbb{N}}$ e $(\Gamma_n)_{n\in\mathbb{N}}$ sono ambedue limitate. □

Di conseguenza, $\{\mu \in \mathcal{P}(\mathbb{R}^d) : \mu \text{ gaussiana}\}$ è un sottoinsieme chiuso dello spazio $\mathcal{P}(\mathbb{R}^d)$ delle misure di probabilità su $\mathcal{B}(\mathbb{R}^d)$ munito della topologia della convergenza debole. L'insieme $\{X \in L^2((\Omega,\mathcal{F},\mathbf{P});\mathbb{R}^d) : X \text{ gaussiana}\}$ è un sottoinsieme chiuso di $L^2((\Omega,\mathcal{F},\mathbf{P});\mathbb{R}^d)$ (rispetto alla convergenza in L^2), ma non è un sottospazio lineare.

Ricordiamo la seguente versione del teorema del limite centrale; ad esempio, Teorema 15.56 in Klenke [2008, p. 326].

[1] La convergenza in legge di (X_n) a X implica la convergenza in legge di $(\langle X_n, v\rangle)$ a $\langle X, v\rangle$, il che implica $\mathbf{P}\left(\langle X, v\rangle \in [a,b]\right) \geq \limsup_{n\to\infty} \mathbf{P}\left(\langle X_n, v\rangle \in [a,b]\right)$ e $\mathbf{P}\left(\langle X, v\rangle \in (a,b)\right) \leq \liminf_{n\to\infty} \mathbf{P}\left(\langle X_n, v\rangle \in (a,b)\right)$ per ogni $-\infty \leq a < b \leq \infty$.

Teorema 0.1 (Teorema del limite centrale). *Sia $(X_n)_{n\in\mathbb{N}}$ una successione i.i.d. di variabili aleatorie a valori in $\mathbb{R}^d$ di quadrato integrabili. Sia Γ la matrice di covarianza comune (finita per ipotesi). Allora la successione di variabili aleatorie definita da*

$$Y_n \doteq \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(X_i - \mathbf{E}\left[X_i\right]\right), \quad n \in \mathbb{N},$$

converge in legge a una gaussiana di media zero e matrice di covarianza Γ.

Definizione 0.13. Un processo stocastico $(X(t))_{t\geq 0}$ a valori in $\mathbb{R}^d$ si dice *processo gaussiano* se per ogni scelta di $n \in \mathbb{N}$, $t_1, \ldots, t_n \in [0, \infty)$ le variabili aleatorie $X(t_1), \ldots, X(t_n)$ hanno legge congiunta gaussiana.

La legge di un processo gaussiano $(X(t))_{t\geq 0}$ è determinata dalle funzioni

$$m_X(t) \doteq \mathbf{E}\left[X(t)\right] \quad \text{e} \quad \rho_X(s,t) \doteq \mathbf{E}\left[(X(s) - m(s))(X(t) - m(t))^{\mathsf{T}}\right],$$

dette funzioni rispettivamente di media e di covarianza.

Lemma 0.2. *Sia $\mathcal{V}$ una famiglia gaussiana reale (cioè una collezione non vuota di variabili aleatorie reali definite sullo stesso spazio di probabilità tale che la distribuzione congiunta di ogni sottoinsieme finito di $\mathcal{V}$ è gaussiana). Siano $\mathcal{V}_1, \mathcal{V}_2 \subset \mathcal{V}$ tali che $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$. Allora $\mathcal{V}_1$, $\mathcal{V}_2$ sono indipendenti se e solo se X, Y sono indipendenti per ogni scelta di $X \in \mathcal{V}_1$ e $Y \in \mathcal{V}_2$.*

Dimostrazione. L'implicazione non ovvia è "$\Leftarrow$". Supponiamo quindi che X, Y siano indipendenti (oppure scorrelate) per ogni $X \in \mathcal{V}_1$, $Y \in \mathcal{V}_2$. Per mostrare che ciò implica l'indipendenza di $\mathcal{V}_1$, $\mathcal{V}_2$ basta far vedere che le variabili aleatorie $\bar{X} \doteq (X_1, \ldots, X_m)$ e $\bar{Y} \doteq (Y_1, \ldots, Y_n)$ a valori rispettivamente in $\mathbb{R}^m$ e $\mathbb{R}^n$ sono indipendenti per ogni scelta di $X_i \in \mathcal{V}_1$, $i \in \{1, \ldots, m\}$, $Y_j \in \mathcal{V}_2$, $j \in \{1, \ldots, n\}$. Siano $\bar{X}$, $\bar{Y}$ di tale forma. Poiché $\mathcal{V}_1 \cup \mathcal{V}_2 \subset \mathcal{V}$ è una famiglia gaussiana, $(\bar{X}, \bar{Y})$ ha legge gaussiana. Per l'ipotesi, $\mathrm{cov}(X_i, Y_j) = 0$ per ogni $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$. La matrice di covarianza di $(\bar{X}, \bar{Y})$ ha quindi la forma

$$\begin{pmatrix} \mathrm{cov}(\bar{X}) & 0 \\ 0 & \mathrm{cov}(\bar{Y}) \end{pmatrix}.$$

Ne discende che la funzione caratteristica $\varphi_{(\bar{X},\bar{Y})}$ di $(\bar{X}, \bar{Y})$, che è la funzione caratteristica di una gaussiana $(m+n)$-dimensionale, ha forma prodotto

$$\varphi_{(\bar{X},\bar{Y})}(x,y) = \varphi_{\bar{X}}(x)\varphi_{\bar{Y}}(y), \quad x \in \mathbb{R}^m,\ y \in \mathbb{R}^n,$$

il che implica l'indipendenza delle gaussiane $\bar{X}$, $\bar{Y}$. □

0.5 Speranza condizionale

Capitolo 8 in Klenke [2008, pp. 169-187], Sezioni 3.1-3.3 in Baldi [2000, pp. 49-58]; Esercizio I.7.

I seguenti tre risultati ci saranno utili. Sia $(\Omega, \mathcal{F}, \mathbf{P})$ uno spazio di probabilità, e sia $\mathcal{G} \subseteq \mathcal{F}$ una sotto-σ-algebra.

Lemma 0.3 (Fattorizzazione)**.** *Siano X, Y variabili aleatorie su $(\Omega, \mathcal{F}, \mathbf{P})$ a valori rispettivamente in $(E_1, \mathcal{E}_1)$ e $(E_2, \mathcal{E}_2)$. Sia φ una mappa $E_1 \times E_2 \to \mathbb{R}$ misurabile e limitata. Se X è indipendente da $\mathcal{G}$ e Y è $\mathcal{G}$-misurabile, allora*

$$\mathbf{E}\left[\varphi(X,Y)|\mathcal{G}\right] = \mathbf{E}\left[\varphi(X,y)\right]_{y=Y} \quad \mathbf{P}\text{-}q.c., \tag{0.9}$$

cioè $\mathbf{E}\left[\varphi(X,Y)|\mathcal{G}\right] = f(Y)$ $\mathbf{P}$-quasi certamente, dove la funzione $f\colon E_2 \to \mathbb{R}$ è data da $f(y) \doteq \mathbf{E}\left[\varphi(X,y)\right]$, $y \in E_2$.

Dimostrazione. Mostriamo l'affermazione in tre passi. Chiaramente, se l'equazione (0.9) è valida per φ, allora è valida anche per $c + \varphi$ con $c \in \mathbb{R}$ una costante.

Primo passo. Supponiamo che φ sia della forma $\varphi(x,y) = \varphi_1(x) \cdot \varphi_2(y)$ per funzioni $\varphi_1\colon E_1 \to \mathbb{R}$, $\varphi_2\colon E_2 \to \mathbb{R}$ limitate e misurabili. Allora con probabilità uno

$$\begin{aligned}
\mathbf{E}\left[\varphi(X,Y)|\mathcal{G}\right] &= \mathbf{E}\left[\varphi_1(X)\,\varphi_2(Y)|\mathcal{G}\right] \\
&= \varphi_2(Y)\;\mathbf{E}\left[\varphi_1(X)|\mathcal{G}\right] \\
&= \varphi_2(Y)\;\mathbf{E}\left[\varphi_1(X)\right] \\
&= \mathbf{E}\left[\varphi_1(X)\,\varphi_2(y)\right]_{y=Y} = \mathbf{E}\left[\varphi(X,y)\right]_{y=Y},
\end{aligned}$$

dove abbiamo utilizzato l'ipotesi che Y sia $\mathcal{G}$-misurabile e X indipendente da $\mathcal{G}$.

Secondo passo. Dal primo passo e per la linearità della speranza condizionale segue che la (0.9) è valida per funzioni della forma $\varphi(x,y) = \sum_{i=1}^{n} \varphi_1^{(i)}(x) \cdot \varphi_2^{(i)}(y)$ con $n \in \mathbb{N}$, $\varphi_1^{(1)}, \ldots, \varphi_1^{(n)}$, $\varphi_2^{(1)}, \ldots, \varphi_2^{(n)}$ misurabili e limitate (combinazioni lineari di funzioni di forma prodotto limitate e misurabili).

Terzo passo. La σ-algebra $\mathcal{E}_1 \otimes \mathcal{E}_2$ è generata dal sistema degli insiemi della forma $A \times B$ con $A \in \mathcal{E}_1$, $B \in \mathcal{E}_2$ ("rettangoli"). Ne segue che, data una mappa $\varphi\colon E_1 \times E_2 \to [0,\infty]$ misurabile, esiste una successione di funzioni ψ_n tale che, per ogni $n \in \mathbb{N}$, ψ_n è una combinazione lineare di funzioni di forma prodotto limitate e misurabili come al secondo passo e $\psi_n \nearrow \varphi$ per $n \to \infty$. La validità della (0.9) si ottiene ora usando la limitatezza di φ (esiste una costante $c \geq 0$ tale che $c + \varphi$ è non-negativa), il risultato del secondo passo e il teorema della convergenza monotona. □

Lemma 0.4. *Sia X una variabile aleatoria a valori in $\mathbb{R}^d$. Se esiste una funzione $\varphi\colon \mathbb{R}^d \to \mathbb{C}$ tale che per ogni $x \in \mathbb{R}^d$*

$$\mathbf{E}\left[e^{\sqrt{-1}\langle X,x\rangle} \,\middle|\, \mathcal{G}\right] = \varphi(x) \quad \mathbf{P}\text{-quasi certamente},$$

allora X ha funzione caratteristica φ ed è indipendente da $\mathcal{G}$.

Dimostrazione. Per la proprietà di raffinamento della speranza condizionale, l'ipotesi implica che

$$\mathbf{E}\left[e^{\sqrt{-1}\langle X,x\rangle} \right] = \varphi(x) \quad \text{per ogni } x \in \mathbb{R}^d,$$

il che significa che X ha funzione caratteristica φ. Per verificare l'indipendenza mostriamo che X, Y sono indipendenti ogniqualvolta Y è una variabile aleatoria a valori in $\mathbb{R}$ e $\mathcal{G}$-misurabile. Sia dunque Y $\mathcal{G}$-misurabile. Indichiamo con $\psi_{(X,Y)}$ la funzione caratteristica della legge congiunta di X e Y. Allora per ogni $x \in \mathbb{R}^d$, ogni $y \in \mathbb{R}$,

$$\begin{aligned}
\psi_{(X,Y)}(x,y) &= \mathbf{E}\left[\exp\left(\sqrt{-1}\left(\langle X, x\rangle + Y \cdot y\right)\right)\right] \\
&= \mathbf{E}\left[\mathbf{E}\left[e^{\sqrt{-1}\langle X,x\rangle} \cdot e^{\sqrt{-1}Y\cdot y} \middle| \mathcal{G}\right]\right] \\
&= \mathbf{E}\left[\mathbf{E}\left[e^{\sqrt{-1}\langle X,x\rangle} \middle| \mathcal{G}\right] e^{\sqrt{-1}Y\cdot y}\right] \\
&= \mathbf{E}\left[\varphi(x) e^{\sqrt{-1}Y\cdot y}\right] \\
&= \varphi(x) \cdot \psi_Y(y),
\end{aligned}$$

ove $\varphi = \psi_X$ è la funzione caratteristica di X e ψ_Y quella di Y. Ne segue, per le proprietà delle funzioni caratteristiche, che X, Y sono indipendenti. □

Lemma 0.5 (Bayes)**.** *Sia X una variabile aleatoria non-negativa con $\mathbf{E}[X] = \mathbf{E}_{\mathbf{P}}[X] = 1$. Definiamo una misura di probabilità $\mathbf{Q}$ su $\mathcal{F}$ mediante*

$$\mathbf{Q}(A) \doteq \mathbf{E}_{\mathbf{P}}\left[\mathbf{1}_A \cdot X\right], \quad A \in \mathcal{F}.$$

Se $\mathbf{E}_{\mathbf{P}}\left[X | \mathcal{G}\right] > 0$ $\mathbf{P}$-quasi certamente, allora per ogni variabile aleatoria Y a valori in $\mathbb{R}$ limitata

$$\mathbf{E}_{\mathbf{Q}}\left[Y | \mathcal{G}\right] = \frac{\mathbf{E}_{\mathbf{P}}\left[Y \cdot X | \mathcal{G}\right]}{\mathbf{E}_{\mathbf{P}}\left[X | \mathcal{G}\right]} \quad \mathbf{Q}\text{-}\textit{quasi certamente.}$$

Dimostrazione. Per costruzione, $\mathbf{Q}$ is assolutamente continua rispetto a $\mathbf{P}$. Per ogni $A \in \mathcal{G}$,

$$\begin{aligned}
\mathbf{E}_{\mathbf{Q}}\left[Y \cdot \mathbf{1}_A\right] &= \mathbf{E}_{\mathbf{P}}\left[X \cdot Y \cdot \mathbf{1}_A\right] \\
&= \mathbf{E}_{\mathbf{P}}\left[\mathbf{E}_{\mathbf{P}}\left[X \cdot Y | \mathcal{G}\right] \mathbf{1}_A\right] \\
&= \mathbf{E}_{\mathbf{P}}\left[\mathbf{E}_{\mathbf{P}}\left[X | \mathcal{G}\right] \frac{\mathbf{E}_{\mathbf{P}}\left[X \cdot Y | \mathcal{G}\right]}{\mathbf{E}_{\mathbf{P}}\left[X | \mathcal{G}\right]} \mathbf{1}_A\right] \\
&= \mathbf{E}_{\mathbf{P}}\left[\mathbf{E}_{\mathbf{P}}\left[X \cdot \frac{\mathbf{E}_{\mathbf{P}}\left[X \cdot Y | \mathcal{G}\right]}{\mathbf{E}_{\mathbf{P}}\left[X | \mathcal{G}\right]} \mathbf{1}_A \middle| \mathcal{G}\right]\right] \\
&= \mathbf{E}_{\mathbf{P}}\left[X \cdot \frac{\mathbf{E}_{\mathbf{P}}\left[X \cdot Y | \mathcal{G}\right]}{\mathbf{E}_{\mathbf{P}}\left[X | \mathcal{G}\right]} \mathbf{1}_A\right] \\
&= \mathbf{E}_{\mathbf{Q}}\left[\frac{\mathbf{E}_{\mathbf{P}}\left[X \cdot Y | \mathcal{G}\right]}{\mathbf{E}_{\mathbf{P}}\left[X | \mathcal{G}\right]} \mathbf{1}_A\right],
\end{aligned}$$

da cui la tesi. □

Capitolo 1

Moto browniano

Motivazione: moto browniano come processo limite di passeggiate aleatorie interpolate e riscalate secondo il teorema del limite centrale:

$$Y_n(t) \doteq \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor t\cdot n \rfloor} \xi_k^{(n)}, \quad t \geq 0,\ n \in \mathbb{N},$$

dove $\xi_k^{(n)}$, $k \in \mathbb{N}$, sono variabili aleatorie i.i.d. a valori in $\mathbb{R}$ con media zero e varianza uno. Il processo limite Y dei processi Y_n, se esiste, per il teorema del limite centrale multidimensionale, deve essere gaussiano con media zero e funzione di covarianza $\rho(s,t) = \min\{s,t\}$; ne segue che gli incrementi devono essere indipendenti e gaussiani con $Y(t) - Y(s) \sim N(0, t-s)$ per ogni $t \geq s$.

Esistenza e regolarità delle traiettorie del processo limite?

Varie costruzioni:

1. Teorema di esistenza di Kolmogorov più teorema di continuità di Kolmogorov-Chentsov (non a lezione); cf. Sezione 2.2 in Karatzas and Shreve [1991, pp. 49-56].

2. Principio di invarianza di Donsker: processo limite (convergenza in distribuzione) di passeggiate aleatorie interpolate e riscalate (Sezione 1.5).

3. Costruzione di Lévy-Ciesielski: processo limite (convergenza quasi certa delle traiettorie) di certi processi continui definiti per ricorrenza (Sezione 1.1).

1.1 Costruzione di Lévy-Ciesielski

Siano $\xi_k^{(n)}$, $n \in \mathbb{N}_0$, $k \in \mathbb{N}$, variabili aleatorie indipendenti e normali standard definite su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbf{P})$. Definiamo processi $(B^{(n)}(t))_{t\in[0,1]}$ per ricorrenza: Poniamo $B^{(0)}(t,\omega) \doteq t \cdot \xi_1^{(0)}$, $(t,\omega) \in [0,1] \times \Omega$, e se $B^{(n-1)}$ è stato definito per un $n \in \mathbb{N}$,

$$B^{(n)}(t,\omega) \doteq \begin{cases} B^{(n-1)}(t,\omega) \text{ se } t = \frac{k}{2^n} \text{ per un } k \in \{0,\ldots,2^n\} \text{ pari,} \\ B^{(n-1)}(t,\omega) + \frac{1}{\sqrt{2^{n+1}}}\xi_k^{(n)}(\omega) \text{ se } t = \frac{k}{2^n} \text{ per un } k \in \{0,\ldots,2^n\} \text{ dispari,} \\ \text{interpolazione lineare a tratti altrimenti.} \end{cases}$$

Per l'interpolazione lineare a tratti, se $t = \frac{k}{2^n}$ per un $k \in \{0,\ldots,2^n\}$ dispari, allora $B^{(n-1)}(t,\omega) = B^{(n-1)}(\frac{k}{2^n},\omega) = \frac{1}{2}\left(B^{(n-1)}(\frac{k-1}{2^n},\omega) + B^{(n-1)}(\frac{k+1}{2^n},\omega)\right)$. Spiegazione della scelta dei valori di $B^{(n)}(\frac{k}{2^n},\omega)$ per $k \in \{0,\ldots,2^n\}$ dispari: Esercizio II.1.

Costruzione in tre passi:

1. **Rappresentazione:**

$$B^{(n)}(t,\omega) = \sum_{m=0}^{n} \sum_{k\in J(m)} S_k^{(m)}(t)\, \xi_k^{(m)}(\omega), \tag{1.1}$$

dove $J(m) \doteq \{k \in \{1,\ldots,2^m\} : k \text{ dispari}\}$ e $S_k^{(m)} : [0,1] \to \mathbb{R}$, $m \in \mathbb{N}_0$, $k \in J(m)$, sono le *funzioni di Schauder* definite da $S_1^{(0)}(t) \doteq t$ e, per $m \in \mathbb{N}$,

$$S_k^{(m)}(t) \doteq \begin{cases} 0 & \text{se } 0 \le t \le \frac{k-1}{2^m}, \\ \sqrt{2^{m-1}}(t - \frac{k-1}{2^m}) & \text{se } \frac{k-1}{2^m} \le t \le \frac{k}{2^m}, \\ \sqrt{2^{m-1}}(\frac{k+1}{2^m} - t) & \text{se } \frac{k}{2^m} \le t \le \frac{k+1}{2^m}, \\ 0 & \text{se } \frac{k+1}{2^m} \le t \le 1. \end{cases}$$

Si potrebbe partire dalla rappresentazione (1.1) prendendola come definizione dei processi $B^{(n)}$.

2. **Convergenza:** Esiste $\Omega_0 \in \mathcal{F}$ tale che $\mathbf{P}(\Omega_0) = 1$ e, per ogni $\omega \in \Omega_0$, la successione $(B^{(n)}(.,\omega))_{n\in\mathbb{N}}$ è di Cauchy nello spazio di Banach $(\mathbf{C}([0,1]), \|.\|_\infty)$.

Dimostrazione. Per $m \in \mathbb{N}_0$, $\omega \in \Omega$, $t \in [0,1]$, poniamo

$$g^{(m)}(t,\omega) \doteq \sum_{k\in J(m)} S_k^{(m)}(t)\, \xi_k^{(m)}(\omega), \qquad \Xi_m(\omega) \doteq \max_{k\in J(m)} |\xi_k^{(m)}(\omega)|.$$

Basta mostrare che esiste $\Omega_0 \in \mathcal{F}$ tale che $\mathbf{P}(\Omega_0) = 1$ e, per ogni $\omega \in \Omega_0$, esiste $n_0(\omega) \in \mathbb{N}$ tale che $\sum_{m=n_0(\omega)}^{\infty} \|g^{(m)}(.,\omega)\|_\infty < \infty$. Siccome le $\xi_k^{(m)}$ hanno distribuzione normale standard, abbiamo

$$\mathbf{P}(\Xi_m > x) \le \sum_{k\in J(m)} \mathbf{P}\left(|\xi_k^{(m)}| > x\right) \le 2^m e^{-x^2/2} \quad \text{per ogni } x \ge 1.$$

Ne segue che $\sum_{m=1}^{\infty} \mathbf{P}(\Xi_m > m) < \infty$ e, con $A \doteq \limsup_{m\to\infty} \{\Xi_m > m\}$, $\mathbf{P}(A) = 0$ grazie al lemma di Borel-Cantelli (parte prima). Scegliamo ora $\Omega_0 \doteq \Omega \setminus A$. Allora $\mathbf{P}(\Omega_0) = 1$ e per ogni $\omega \in \Omega_0$ esiste $n_0(\omega) \in \mathbb{N}$ tale che $\Xi_m(\omega) \le m$

per ogni $m \geq n_0(\omega)$. Per costruzione, $\| \sum_{k \in J(m)} S_k^{(m)} \|_\infty \leq 1/\sqrt{2^{m+1}}$. Di conseguenza, per ogni $\omega \in \Omega_0$,

$$\sum_{m=n_0(\omega)}^{\infty} \|g^{(m)}(.,\omega)\|_\infty \leq \sum_{m=n_0(\omega)}^{\infty} \Xi_m(\omega) \cdot \| \sum_{k \in J(m)} S_k^{(m)} \|_\infty \leq \sum_{m=n_0(\omega)}^{\infty} \frac{m}{\sqrt{2^{m+1}}},$$

ma $\sum_{m=1}^{\infty} m/\sqrt{2^{m+1}} < \infty$. $\square$

3. **Identificazione:** Il processo B definito da $B(t,\omega) \doteq \lim_{n\to\infty} B^{(n)}(t,\omega)$ per $\omega \in \Omega_0$, $B(t,\omega) \doteq 0$ altrimenti, è continuo ed ha incrementi indipendenti con $B(t) - B(s) \sim N(0, t-s)$ per $t > s$.

Costruzione fornisce un processo continuo $(B(t))_{t\in[0,1]}$. Costruzione di un moto browniano $(W(t))_{t\geq 0}$ come processo stocastico definito per tempi $t \geq 0$ (invece di $t \in [0,1]$) mediante "copia ed incolla".

Definizione 1.1. Un processo stocastico $W = (W(t))_{t\geq 0}$ a valori in $\mathbb{R}$ definito su $(\Omega, \mathcal{F}, \mathbf{P})$ si dice *moto browniano standard* se

(i) $\mathbf{P}(W(0) = 0) = 1$,

(ii) per ogni $0 \leq s < t$, $W(t) - W(s)$ ha distribuzione $N(0, t-s)$,

(iii) per ogni scelta di $n \in \mathbb{N}$, $0 \leq t_0 < t_1 < \ldots < t_n$, le variabili aleatorie $W(t_n) - W(t_{n-1}, \ldots, W(t_1) - W(t_0)$ sono indipendenti,

(iv) le traiettorie di W sono continue con probabilità uno.

Un moto browniano standard è quindi un processo stocastico reale continuo che parte da zero ed ha incrementi indipendenti gaussiani stazionari (stazionari poiché la distribuzione di $W(t) - W(s)$ dipende solo dalla differenza $t - s$ dei tempi).

Fonti: Sezione 2.3 in Karatzas and Shreve [1991, pp. 56-59] (utilizzando l'esercizio 3.3 al posto dell'identità di Parseval per le funzioni di Haar, cf. anche Esercizio II.4); Sezione 2.3 in Caravenna [2011, pp. 32-37].

1.2 Proprietà di base

Proposizione 1.1 (Caratterizzazioni equivalenti). *Sia $W = (W(t))_{t\geq 0}$ un processo stocastico reale definito su $(\Omega, \mathcal{F}, \mathbf{P})$ tale che, con probabilità uno, $W(0) = 0$ e W è continuo. Le seguenti condizioni sono equivalenti:*

a) W è un moto browniano standard;

b) W è un processo gaussiano di media zero e funzione di covarianza $(s,t) \mapsto s \wedge t$;

c) per ogni $t > s \geq 0$, $W(t) - W(s)$ ha distribuzione $N(0, t-s)$ ed è indipendente dalla σ-algebra generata da $W(\tilde{s})$, $\tilde{s} \in [0, s]$.

Dimostrazione. **a) ⇒ b).** La famiglia $\{W(t) : t \geq 0\}$ di variabili aleatorie reali è gaussiana poiché la legge congiunta di $W(t_n), \ldots, W(t_1)$ è gaussiana per ogni scelta di $n \in \mathbb{N}$, $t_1, \ldots, t_n \in [0, \infty)$ (senza perdità di generalità $t_1 < \ldots < t_n$) essendo il risultato di una trasformazione lineare del vettore gaussiano $(W(t_n) - W(t_{n-1}), \ldots, W(t_1) - W(0))$. Le variabili $W(t)$, $t \geq 0$, sono tutte centrate; la funzione di media è quindi costante zero. La covarianza per $t \geq s$ è data da

$$\mathbf{E}\left[W(s) \cdot W(t)\right] = \mathbf{E}\left[W(s)\left(W(t) - W(s)\right)\right] + \mathbf{E}\left[(W(s))^2\right] = 0 \cdot 0 + s = s,$$

quindi $\mathbf{E}\left[W(s) \cdot W(t)\right] = \min\{s, t\}$ per ogni $s, t \geq 0$.

b) ⇒ c). Per $t > s$, $(W(t), W(s))^{\mathsf{T}}$ ha legge congiunta gaussiana di media zero e matrice di covarianza $\begin{pmatrix} t & s \\ s & s \end{pmatrix}$. Ne segue che

$$W(t) - W(s) = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} W(t) \\ W(s) \end{pmatrix}$$

ha distribuzione $N(0, t-s)$. Essendo $\{W(t) : t \geq 0\}$ una famiglia gaussiana, per mostrare che $W(t) - W(s)$ è indipendente da $\sigma(W(\tilde{s}) : \tilde{s} \leq s)$, basta verificare (cf. Esercizio II.3) che, per ogni $\tilde{s} \in [0, s]$, $W(t) - W(s)$ e $W(\tilde{s})$ sono indipendenti; grazie alla gaussianietà congiunta, l'indipendenza è equivalente a covarianza zero, ma

$$\operatorname{cov}(W(t) - W(s), W(\tilde{s})) = \mathbf{E}\left[\left(W(t) - W(s)\right) W(\tilde{s})\right] = t \wedge \tilde{s} - s \wedge \tilde{s} = 0$$

per la forma della funzione di covarianza di W, che è data da $(s, t) \mapsto s \wedge t$.

c) ⇒ a). Sia $n \in \mathbb{N}$, siano $0 \leq t_0 < t_1 < \ldots < t_n$. Si dimostra per induzione che la legge congiunta di $W(t_n) - W(t_{n-1}), \ldots, W(t_1) - W(t_0)$ coincide con la misura prodotto $N(0, t_n - t_{n-1}) \otimes \ldots \otimes N(0, t_1 - t_0)$. □

La proprietà di un processo di essere un moto browniano è preservata dalle seguenti trasformazioni (cambia ovviamente il processo, ma non la sua legge).

Proposizione 1.2 (Trasformazioni)**.** *Sia W un moto browniano standard. Le seguenti trasformazioni di W sono ancora moti bowniani standard:*

a) $-W$ (simmetria rispetto all'origine);

b) $(W(t + t_0) - W(t_0))_{t \geq 0}$ per ogni $t_0 \geq 0$ fissato (traslazione temporale);

c) $(\frac{1}{\sqrt{c}} \cdot W(c \cdot t))_{t \geq 0}$ per ogni $c > 0$ fissato (riscalamento diffusivo);

d) $(t \cdot W(1/t))_{t \geq 0}$ (inversione del tempo – valore in $t = 0$ per continuità).

Inoltre, per $T > 0$ fissato, il processo $(W(T - t) - W(T))_{t \in [0,T]}$ è un moto browniano su $[0, T]$ (riflessione temporale).

Dimostrazione. Le affermazioni **a)** e **b)** seguono direttamente dalla Definizione 1.2. La **c)** si verifica facilmente usando la caratterizzazione del moto browniano come processo gaussiano; cf. Proposizione 1.1 b).

Per la **d)**, si osserva che $(t \cdot W(1/t))_{t>0}$ è un processo gaussiano centrato con funzione di covarianza $s \wedge t$ e traiettorie continue (per tempi strettamente maggiore di zero) $\mathbf{P}$-quasi certamente. Resta da mostrare che $\lim_{t\to 0+} t \cdot W(1/t) = 0$ con probabilità uno. Ma

$$\mathbf{P}\left(\left\{\omega \in \Omega : \lim_{t\to 0+, t\in\mathbb{Q}} t \cdot W(1/t,\omega) = 0\right\}\right) = 1$$

poiché l'evento $\{\omega \in \Omega : \lim_{t\to 0+, t\in\mathbb{Q}} t \cdot W(1/t,\omega) = 0\}$ coincide con

$$\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \underbrace{\bigcap_{t\in(0,1/n]\cap\mathbb{Q}} \left\{\omega \in \Omega : |t \cdot W(1/t,\omega)| \leq \frac{1}{m}\right\}}_{\doteq A_{m,n}}$$

e, per ogni $m \in \mathbb{N}$ fissato, $A_{m,n} \subseteq A_{m,n+1}$ e

$$\mathbf{P}\left(A_{m,n}\right) = \mathbf{P}\left(\bigcap_{t\in(0,1/n]\cap\mathbb{Q}} \left\{\omega \in \Omega : |W(t,\omega)| \leq \frac{1}{m}\right\}\right) \overset{n\to\infty}{\nearrow} 1,$$

dove abbiamo usato che i processi $(t \cdot W(1/t))_{t>0}$ e $(W(t))_{t>0}$ hanno la stessa legge. Infine, poiché $(t \cdot W(1/t))_{t>0}$ ha traiettorie continue quasi certamente, gli eventi $\{\omega \in \Omega : \lim_{t\to 0+} t \cdot W(1/t,\omega) = 0\}$ e $\{\omega \in \Omega : \lim_{t\to 0+, t\in\mathbb{Q}} t \cdot W(1/t,\omega) = 0\}$ coincidono a meno di un insieme di $\mathbf{P}$-misura zero. □

La Proposizione 1.2 d) implica la seguente "legge dei grandi numeri" per il moto browniano:

$$\lim_{t\to\infty} \frac{W(t)}{t} = 0 \quad \mathbf{P}\text{-q.c.}$$

Sia $(\mathcal{F}_t)_{t\geq 0}$ una filtrazione in $\mathcal{F}$, $(\Omega, \mathcal{F}, \mathbf{P})$ uno spazio di probabilità.

Definizione 1.2. Un processo stocastico $W = (W(t))_{t\geq 0}$ a valori in $\mathbb{R}$ definito su $(\Omega, \mathcal{F}, \mathbf{P})$ si dice *moto browniano* rispetto a $(\mathcal{F}_t)_{t\geq 0}$ se

(i) $\mathbf{P}(W(0) = 0) = 1$,

(ii) W è adattato rispetto a $(\mathcal{F}_t)$,

(iii) per ogni $0 \leq s < t$, $W(t) - W(s)$ ha distribuzione $N(0, t-s)$ ed è indipendente da $\mathcal{F}_s$,

(iv) le traiettorie di W sono continue con probabilità uno.

Fonti: Sezione 2.2 in Baldi [2000, pp. 32-35]; Sezione 2.2.1 in Caravenna [2011, pp. 27-31].

1.3 Struttura delle traiettorie

In questa sezione, indichiamo con W un moto browniano standard definito su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbf{P})$ che supponiamo essere completo (per evitare problemi di misurabilità).

Teorema 1.1 (Wiener-Paley-Zygmund)**.** *Le traiettorie del moto browniano, con probabilità uno, non sono differenziabili in nessun punto:*

$$\mathbf{P}\left(\{\omega \in \Omega : \exists t \geq 0 : W(.,\omega) \text{ è derivabile in } t\}\right) = 0.$$

Dimostrazione. Sia $f : [0,\infty) \to \mathbb{R}$ una funzione, e sia $t \geq 0$. Una condizione necessaria per la differenziabilità di f in t è che esistono costanti $h_0 \in (0,1]$, $C \in \mathbb{N}$ tali che

$$\frac{1}{h}\left|f(t+h) - f(t)\right| \leq C \quad \text{per ogni } h \in (0, h_0].$$

Questa condizione necessaria è equivalente a: esistono numeri $n_0 \in \mathbb{N}$, $C \in \mathbb{N}$ tali che per ogni numero $n \geq n_0$, ogni $k \in \{i+1, i+2, i+3\}$, ove $i \in \{0,\ldots,n-1\}$ e $j \in \mathbb{N}_0$ sono univocamente determinati da $j + i/n \leq t < j + (i+1)/n$,

$$|f(j + (k+1)/n) - f(j + k/n)| \leq \frac{7C}{n}.$$

Ponendo, per $n, c \in \mathbb{N}$, $j \in \mathbb{N}_0$,

$$A_{n,c,j} \doteq \bigcup_{i=0}^{n-1} \bigcap_{k=i+1}^{i+3} \left\{\omega \in \Omega : \left|W\left(j + \tfrac{k+1}{n}, \omega\right) - W\left(j + \tfrac{k}{n}, \omega\right)\right| \leq \frac{c}{n}\right\},$$

abbiamo l'inclusione

$$\{\omega \in \Omega : \exists t \geq 0 : W(.,\omega) \text{ è derivabile in } t\} \subseteq \bigcup_{j=0}^{\infty} \bigcup_{c=1}^{\infty} \bigcup_{n_0=1}^{\infty} \bigcap_{n=n_0}^{\infty} A_{n,c,j}.$$

Per verificare la tesi, è quindi sufficiente mostrare che $\mathbf{P}\left(\bigcap_{n=n_0}^{\infty} A_{n,c,j}\right) = 0$ per ogni $n_0, c \in \mathbb{N}$, $j \in \mathbb{N}_0$. Siano $n, c \in \mathbb{N}$, $j \in \mathbb{N}_0$. Allora, per l'indipendenza e la distribuzione degli incrementi di W,

$$\mathbf{P}\left(A_{n,c,j}\right) \leq \sum_{i=0}^{n-1} \mathbf{P}\left(|W(1/n)| \leq \frac{c}{n}\right)^3 = n \cdot \mathbf{P}\left(|W(1)| \leq \frac{c}{\sqrt{n}}\right)^3 \leq n \cdot \left(\frac{c}{\sqrt{n}}\right)^3,$$

dove in particolare è stato usato il fatto che $W(1) \sim N(0,1)$. Per monotonia, ne segue che

$$\mathbf{P}\left(\bigcap_{n=n_0}^{\infty} A_{n,c,j}\right) \leq \lim_{n\to\infty} \mathbf{P}\left(A_{n,c,j}\right) \leq \lim_{n\to\infty} \frac{c^3}{\sqrt{n}} = 0$$

per ogni $n_0, c \in \mathbb{N}$, $j \in \mathbb{N}_0$. □

Definizioni: funzione *a variazione totale finita*, funzione *a variazione quadratica finita* (attenzione al ruolo delle partizioni).

Teorema 1.2 (Lévy)**.** *Sia $T > 0$, e sia (t_i^n) una successione di partizioni di $[0,T]$ con passo tendente a zero. Allora per ogni $t \in [0,T]$,*

$$\sum_{i=1}^{N_n} \left| W(t\wedge t_i^n) - W(t\wedge t_{i-1}^n) \right|^2 \overset{n\to\infty}{\longrightarrow} t \quad \textit{in } L^2(\mathbf{P}).$$

Dimostrazione. Chiaramente, per ogni $n \in \mathbb{N}$, $t = \sum_{i=1}^{N_n} t\wedge t_i^n - t\wedge t_{i-1}^n$. Di conseguenza e grazie all'indipendenza degli incrementi di W,

$$\begin{aligned}
&\mathbf{E}\left[\left| t - \sum_{i=1}^{N_n} \left| W(t\wedge t_i^n) - W(t\wedge t_{i-1}^n) \right|^2 \right|^2\right] \\
&= \mathbf{E}\left[\left| \sum_{i=1}^{N_n} \left| W(t\wedge t_i^n) - W(t\wedge t_{i-1}^n) \right|^2 - \left(t\wedge t_i^n - t\wedge t_{i-1}^n\right) \right|^2\right] \\
&= \sum_{i=1}^{N_n} \mathbf{E}\left[\left| \left| W(t\wedge t_i^n) - W(t\wedge t_{i-1}^n) \right|^2 - \left(t\wedge t_i^n - t\wedge t_{i-1}^n\right) \right|^2\right] \\
&= \sum_{i=1}^{N_n} 2\left| t\wedge t_i^n - t\wedge t_{i-1}^n \right|^2 \\
&\leq 2t \cdot \max_{i\in\{1,\ldots,N_n\}} \left| t\wedge t_i^n - t\wedge t_{i-1}^n \right| \\
&\overset{n\to\infty}{\longrightarrow} 0,
\end{aligned}$$

poiché per ipotesi il passo delle partizioni tende a zero. Abbiamo usato inoltre il fatto che $W(t\wedge t_i^n) - W(t\wedge t_{i-1}^n)$ sia una variabile aleatoria gaussiana di media zero e varianza $t\wedge t_i^n - t\wedge t_{i-1}^n$, quindi $\left(W(t\wedge t_i^n) - W(t\wedge t_{i-1}^n)\right)^2$ una variabile aleatoria di media $t\wedge t_i^n - t\wedge t_{i-1}^n$ e varianza $2\left(t\wedge t_i^n - t\wedge t_{i-1}^n\right)^2$. □

Osservazioni:

- Dal Teorema 1.2 discende che con probabilità uno le traiettorie del moto browniano *non* sono a variazione totale finita (questo fatto si può dedurre anche dal Teorema 1.1). Di conseguenza, non si può definire l'integrale rispetto a una traiettoria tipica del moto browniano nel senso di Lebesgue-Stieltjes.

- Il Teorema 1.2 afferma la convergenza delle sommatorie degli incrementi quadratici del moto browniano in L^2, quindi anche in probabilità. In realtà vale convergenza con probabilità uno (quasi certa); cf. Teorema 21.64 in Klenke [2008, pp. 486-487].

Lemma 1.1 (Stime per la gaussiana standard)**.** *Per ogni $x > 0$,*

$$\frac{x}{1+x^2}\cdot e^{-x^2/2} \leq \int_x^\infty e^{-z^2/2}dz \leq \frac{1}{x}\cdot e^{-x^2/2}.$$

In particolare, per $x \geq 1$, $\int_x^\infty e^{-z^2/2}dz \leq e^{-x^2/2}$.

Dimostrazione. Sia $x > 0$. Allora

$$\int_x^{\infty} e^{-z^2/2} dz \leq \frac{1}{x} \int_x^{\infty} z \cdot e^{-z^2/2} dz = \frac{1}{x} e^{-x^2/2}.$$

Dall'altra parte, siccome $\frac{d}{dx}\left(\frac{1}{x}e^{-x^2/2}\right) = -\frac{x^2+1}{x^2}e^{-x^2/2}$,

$$\frac{1}{x}e^{-x^2/2} = \int_x^{\infty} \frac{z^2+1}{z^2} e^{-z^2/2} dz \leq \frac{x^2+1}{x^2} \int_x^{\infty} e^{-z^2/2} dz,$$

poiché $\frac{z^2+1}{z^2} \leq \frac{x^2+1}{x^2}$ per $z \geq x$. □

Lemma 1.2 (Massimo corrente del moto browniano)**.** *Per ogni* $T > 0$*, ogni* $x > 0$*,*

$$\mathbf{P}\left(\sup_{t\in[0,T]} W(t) > x\right) \leq 2\,\mathbf{P}\left(W(T) > x\right).$$

Dimostrazione. Siano $T > 0$, $x > 0$ fissati. Per $n \in \mathbb{N}$, $i \in \{0, \ldots, 2^n\}$, poniamo $t_i^n \doteq \frac{i}{2^n} \cdot T$. Possiamo supporre che *tutte* le traiettorie di W siano continue. Allora per la continuità delle traiettorie,

$$\begin{aligned}\left\{\omega \in \Omega : \sup_{t\in[0,T]} W(t,\omega) > x\right\} &= \left\{\omega \in \Omega : \max_{t\in[0,T]} W(t,\omega) > x\right\} \\ &= \bigcup_{n=1}^{\infty} \left\{\omega \in \Omega : \max_{i\in\{0,\ldots,2^n\}} W(t_i^n,\omega) > x\right\}.\end{aligned}$$

Sia $n \in \mathbb{N}$. Poniamo $\tau_n(\omega) \doteq \inf\{i \in \{0,\ldots,2^n\} : W(t_i^n,\omega) > x\}$, $\omega \in \Omega$, con la convenzione $\inf \emptyset = \infty$. Allora per l'indipendenza degli incrementi del moto browniano e la simmetria rispetto a zero delle loro distribuzioni,

$$\begin{aligned}&\mathbf{P}\left(\max_{i\in\{0,\ldots,2^n\}} W(t_i^n) > x\right) \\ &= \mathbf{P}\left(\tau_n < \infty\right) \\ &= \sum_{i=0}^{2^n} \mathbf{P}\left(\tau_n = i\right) \\ &= \sum_{i=0}^{2^n} \mathbf{P}\left(\{\tau_n = i\} \cap \{W(t_i^n) > x\}\right) \\ &= \sum_{i=0}^{2^n} \mathbf{P}\left(\{\tau_n = i\} \cap \{W(t_i^n) > x\} \cap \{W(t_n^n) - W(t_i^n) \geq 0\}\right) \\ &\quad + \sum_{i=0}^{2^n} \mathbf{P}\left(\{\tau_n = i\} \cap \{W(t_i^n) > x\} \cap \{W(t_n^n) - W(t_i^n) < 0\}\right) \\ &= 2\sum_{i=0}^{2^n} \mathbf{P}\left(\{\tau_n = i\} \cap \{W(t_i^n) > x\} \cap \{W(T) - W(t_i^n) \geq 0\}\right)\end{aligned}$$

$$\leq 2 \sum_{i=0}^{2^n} \mathbf{P}\left(\{\tau_n = i\} \cap \{W(T) > x\}\right)$$
$$= 2\,\mathbf{P}\left(W(T) > x\right).$$

Ne segue che

$$\mathbf{P}\left(\sup_{t\in[0,T]} W(t) > x\right) = \lim_{n\to\infty} \mathbf{P}\left(\max_{i\in\{0,\dots,2^n\}} W(t_i^n) > x\right) \leq 2\,\mathbf{P}\left(W(T) > x\right)$$

poiché $\{\max_{i\in\{0,\dots,2^m\}} W(t_i^m) > x\} \subseteq \{\max_{i\in\{0,\dots,2^n\}} W(t_i^n) > x\}$ per $m \leq n$ e $\{\sup_{t\in[0,T]} W(t) > x\} = \bigcup_{n=1}^{\infty}\{\max_{i\in\{0,\dots,2^n\}} W(t_i^n) > x\}$. □

Nel Lemma 1.2 vale in realtà uguaglianza; cf. Teorema 1.6 (principio di riflessione).

Teorema 1.3 (Legge del logaritmo iterato)**.** *Si ha*

$$\mathbf{P}\left(\limsup_{t\to 0+} \frac{W(t)}{\sqrt{2t\log(\log(1/t))}}\right) = 1.$$

Dimostrazione. Poniamo $\varphi(t) \doteq \sqrt{2t\log(\log(1/t))}$, $t \in (0, 1/e]$. Notiamo che $\lim_{t\to 0+}\varphi(t) = 0$, $\varphi(t) \geq \sqrt{2t}$ per $t \in (0, e^{-e}]$ e φ è strettamente crescente su $(0, e^{-e}]$.

Prima parte: $\limsup_{t\to 0+} \frac{W(t)}{\varphi(t)} \leq 1$ **P**-quasi certamente. Sia $\delta > 0$ fissato. Sia $(t_n) \subset (0, e^{-e}]$ una successione decrescente (da specificare) tale che $t_n \searrow 0$ per $n \to \infty$. Poniamo

$$A_n = A_{n,\delta} \doteq \{\omega \in \Omega : \exists t \in [t_{n+1}, t_n] : W(t,\omega) > (1+\delta)\varphi(t)\}.$$

Poiché $\delta > 0$ è arbitrario, basta mostrare che $\mathbf{P}(\limsup_{n\to\infty} A_n) = 0$. Grazie al lemma di Borel-Cantelli, è sufficiente mostrare che

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty.$$

Siccome φ è crescente su $(0, e^{-e}]$ e (t_n) decrescente, abbiamo per ogni $n \in \mathbb{N}$

$$A_n \subseteq \left\{\omega \in \Omega : \sup_{t\in[t_{n+1},t_n]} W(t,\omega) > (1+\delta)\varphi(t_{n+1})\right\}$$
$$\subseteq \left\{\omega \in \Omega : \sup_{t\in[0,t_n]} W(t,\omega) > (1+\delta)\varphi(t_{n+1})\right\}.$$

Poniamo $x_n \doteq \frac{1+\delta}{\sqrt{t_n}}\varphi(t_{n+1})$. Allora per i Lemmata 1.1 e 1.2,

$$\begin{aligned}\mathbf{P}(A_n) &\leq \mathbf{P}\left(\sup_{t\in[0,t_n]} W(t) > \sqrt{t_n}x_n\right)\\ &\leq 2\,\mathbf{P}\left(W(t_n) > \sqrt{t_n}x_n\right)\\ &\leq \frac{2}{\sqrt{2\pi}}\int_{x_n}^{\infty} e^{-z^2/2}dz\\ &\leq \frac{1}{x_n}e^{-x_n^2/2}.\end{aligned}$$

Scegliendo la successione (t_n) mediante $t_n \doteq (1+\delta)^{-n}$, $n\in\mathbb{N}$, otteniamo

$$x_n = \frac{1+\delta}{(1+\delta)^{-n/2}}\varphi\big((1+\delta)^{-(n+1)}\big) = \sqrt{2(1+\delta)\log\big((n+1)\log(1+\delta)\big)}.$$

In particolare, $x_n \geq 1$ se $n \geq 2/\log(1+\delta)$. Di conseguenza, per $n \geq 2/\log(1+\delta)$,

$$\mathbf{P}(A_n) \leq e^{-x_n^2/2} = (\log(1+\delta))^{-(1+\delta)}\cdot\frac{1}{(n+1)^{1+\delta}},$$

quindi $\sum_{n=1}^{\infty}\mathbf{P}(A_n) < \infty$.

Seconda parte: $\limsup_{t\to 0+}\frac{W(t)}{\varphi(t)} \geq 1$ **P**-quasi certamente. Sia $\delta\in(0,1/2]$ fissato. Basta trovare una successione $(t_n)\subset(0,e^{-e}]$ decrescente tale che $t_n \searrow 0$ per $n\to\infty$ e

$$\mathbf{P}\left(\limsup_{n\to\infty}\{\omega\in\Omega : W(t_n,\omega) > (1-\delta)\varphi(t_n)\}\right) = 1.$$

Sia $\varepsilon\in(0,\delta)$. Scegliamo q maggiore di zero e minore di $\min\{1-(1-\varepsilon)^2, \frac{(\delta-\varepsilon)^2}{(1+\varepsilon)^2}, \frac{1}{e}\}$. Poniamo, per $n\in\mathbb{N}$,

$$t_n \doteq q^n, \qquad x_n \doteq \frac{1-\varepsilon}{\sqrt{t_n - t_{n+1}}}\varphi(t_n),$$
$$B_n \doteq \{\omega\in\Omega : W(t_n,\omega) - W(t_{n+1},\omega) > (1-\varepsilon)\varphi(t_n)\}.$$

Allora

$$x_n = \frac{1-\varepsilon}{\sqrt{1-q}}\sqrt{2\log\big(n\log(1/q)\big)} > 0$$

e $x_n \nearrow \infty$ per $n\to\infty$. In particolare, $x_n \geq 1$ per $n \geq 8$. Grazie al Lemma 1.1, per $n\geq 8$,

$$\begin{aligned}\mathbf{P}(B_n) &= \mathbf{P}\left(\frac{W(t_n)-W(t_{n+1})}{\sqrt{t_n - t_{n+1}}} > x_n\right)\\ &= \frac{1}{\sqrt{2\pi}}\int_{x_n}^{\infty} e^{-z^2/2}dz\\ &\geq \frac{1}{\sqrt{2\pi}}\frac{x_n}{x_n^2+1}e^{-x_n^2/2}\\ &\geq \frac{1}{6x_n}e^{-x_n^2/2}\\ &\geq \frac{\sqrt{1-q}}{6(1-\varepsilon)}\cdot\frac{1}{\sqrt{2\log(n\log(1/q))}}\cdot\left(\frac{1}{n\log(1/q)}\right)^{\frac{(1-\varepsilon)^2}{1-q}}.\end{aligned}$$

Poiché q è tale che $(1-\varepsilon)^2 < 1-q$, questo implica che

$$\sum_{n=1}^{\infty} \mathbf{P}(B_n) = \infty.$$

Ma per l'indipendenza degli incrementi del moto browniano, B_n, $n \in \mathbb{N}$, sono eventi indipendenti. Di conseguenza, per la seconda parte del lemma di Borel-Cantelli,

$$\mathbf{P}\left(\limsup_{n\to\infty} B_n\right) = 1.$$

Esiste dunque un evento $\Omega_0 \in \mathcal{F}$ tale che $\mathbf{P}(\Omega_0) = 1$ e per ogni $\omega \in \Omega_0$, $W(t_n,\omega) - W(t_{n+1},\omega) > (1-\varepsilon)\varphi(t_n)$ per infiniti indici $n \in \mathbb{N}$. Dalla prima parte della dimostrazione applicata al moto browniano $-W$ si ha che esiste $\Omega_1 \in \mathcal{F}$ tale che $\mathbf{P}(\Omega_1) = 1$ e per ogni $\omega \in \Omega_1$ esiste $n_0(\omega) \in \mathbb{N}$ tale che $W(t_n,\omega) > -(1+\varepsilon)\varphi(t_n)$ per ogni $n \geq n_0(\omega)$. Poniamo $\tilde{\Omega} \doteq \Omega_0 \cap \Omega_1$. Allora $\mathbf{P}(\tilde{\Omega}) = 1$ e per ogni $\omega \in \tilde{\Omega}$ si ha che

$$W(t_n,\omega) = W(t_{n+1},\omega) + W(t_n,\omega) - W(t_{n+1},\omega) > \varphi(t_n)\left(1-\varepsilon-(1+\varepsilon)\frac{\varphi(t_{n+1})}{\varphi(t_n)}\right)$$

per infiniti indici $n \in \mathbb{N}$. Poiché $\lim_{n\to\infty} \frac{\varphi(t_{n+1})}{\varphi(t_n)} = \sqrt{q}$ e $1-\varepsilon-(1+\varepsilon)\sqrt{q}) > 1-\delta$ per la scelta di q, si ottiene che con probabilità uno $W(t_n) > (1-\delta)\varphi(t_n)$ per infiniti indici n. Di conseguenza, per ogni $\delta \in (0,1/2]$, $\limsup_{t\to 0+} \frac{W(t)}{\varphi(t)} \geq 1-\delta$ $\mathbf{P}$-quasi certamente. □

Corollario 1.1.

$$\mathbf{P}\left(\limsup_{t\to 0+} \frac{W(t)}{\sqrt{2t\log(\log(1/t))}}\right) = 1, \qquad \mathbf{P}\left(\limsup_{t\to \infty} \frac{W(t)}{\sqrt{2t\log(\log(t))}}\right) = 1,$$

$$\mathbf{P}\left(\liminf_{t\to 0+} \frac{W(t)}{\sqrt{2t\log(\log(1/t))}}\right) = -1, \quad \mathbf{P}\left(\liminf_{t\to \infty} \frac{W(t)}{\sqrt{2t\log(\log(t))}}\right) = -1.$$

Dimostrazione. Basta applicare il Teorema 1.3 ai moti browniani standard W, $-W$, $(t\cdot W(1/t))_{t\geq 0}$, $(-t\cdot W(1/t))_{t\geq 0}$; cf. Proposizione 1.2. Attenzione all'argomento del logaritmo iterato (cambia da $1/t$ a t). □

Definizione 1.3. Sia $f\colon \mathbb{R} \to \mathbb{R}^d$ una funzione. Il *modulo di continuità* di f su un intervallo $[a,b]$ è la mappa

$$(0,\infty) \ni \delta \mapsto w_f(\delta;[a,b]) \doteq \sup_{t,s\in[a,b]:|t-s|\leq\delta} |f(t)-f(s)| \in [0,\infty).$$

Una funzione f è continua nell'intervallo $[a,b]$ se e solo se $w_f(\delta;[a,b]) \to 0$ per $\delta \searrow 0$.

Teorema 1.4 (Lévy). *Per* $\mathbf{P}$*-quasi ogni* $\omega \in \Omega$*,*

$$\limsup_{h\to 0+} \frac{w_{W(.,\omega)}(h;[0,1])}{\sqrt{2h\log(1/h)}} = 1.$$

A lezione senza dimostrazione; vedi Sezione 2.9.F in Karatzas and Shreve [1991, pp. 114-116].

Fonti: Sezione 2.4 in Baldi [2000, pp. 39-42], Sezioni 2.9.D e 2.9.E in Karatzas and Shreve [1991, pp. 109-114].

1.4 Tempi d'arresto e teorema d'arresto per il moto browniano

Sia $(\Omega, \mathcal{F}, \mathbf{P})$ uno spazio di probabilità, sia $(\mathcal{F}_t)_{t\in I}$ una filtrazione in $\mathcal{F}$, I un insieme di tempi.

Definizione 1.4. Un *tempo d'arresto* rispetto alla filtrazione $(\mathcal{F}_t)_{t\in I}$ è una mappa $\tau\colon \Omega \to I \cup \{\infty\}$ tale che per ogni $t \in I$

$$\{\omega \in \Omega : \tau(\omega) \leq t\} \in \mathcal{F}_t.$$

Ogni tempo costante $t \in I \cup \{\infty\}$ è un tempo d'arresto. Sia τ un tempo d'arresto rispetto a $(\mathcal{F}_t)$. Allora il sistema di parti

$$\mathcal{F}_\tau \doteq \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ per ogni } t \in I\}$$

è una σ-algebra, la σ-algebra degli eventi determinati fino al tempo τ. Ricordiamo le seguenti proprietà dei tempi d'arresto:

Proposizione 1.3. *Siano σ, τ tempi d'arresto rispetto a $(\mathcal{F}_t)_{t\in I}$. Allora*

a) τ è $\mathcal{F}_\tau$-misurabile;

b) $\sigma \vee \tau$, $\sigma \wedge \tau$ sono tempi d'arresto rispetto a $(\mathcal{F}_t)$;

c) $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$ se $\sigma \leq \tau$;

d) $\mathcal{F}_{\sigma\wedge\tau} = \mathcal{F}_\sigma \cap \mathcal{F}_\tau$ e gli eventi $\{\tau < \sigma\}$, $\{\tau \leq \sigma\}$, $\{\tau = \sigma\}$ sono in $\mathcal{F}_{\sigma\wedge\tau}$;

e) $\tau + \sigma$ è un tempo d'arresto rispetto a $(\mathcal{F}_t)$ se $t_1, t_2 \in I$ implica $t_1 + t_2 \in I$.

Proposizione 1.4. *Sia $(X(t))_{t\in I}$ un processo stocastico a valori in $(E, \mathcal{E})$ definito su $(\Omega, \mathcal{F}, \mathbf{P})$, e sia τ un tempo aleatorio a valori in I, cioè una funzione misurabile $\Omega \to I$.*

1. *Se X è un processo misurabile, allora $X(\tau)$ è una variabile aleatoria.*

2. *Se X è un processo progressivamente misurabile rispetto a $(\mathcal{F}_t)$ e τ un tempo d'arresto rispetto a $(\mathcal{F}_t)$, allora $X(\tau)$ è $\mathcal{F}_\tau$-misurabile.*

Una classe particolarmente importante di tempi aleatori sono i tempi di prima uscita e prima entrata di un processo stocastico da/in una parte dello spazio di stati. Sia I tale che $\inf I \in I$, e sia $(X(t))_{t\in I}$ un processo stocastico su $(\Omega, \mathcal{F}, \mathbf{P})$ a valori in uno spazio metrico (E, d). Per $D \in \mathcal{B}(E)$, il tempo aleatorio $\tau_D \colon \Omega \to I \cup \{\infty\}$ definito da

$$\tau_D(\omega) \doteq \inf\{t \in I : X(t,\omega) \in D\}$$

si dice *tempo di prima entrata* di X in D. Analogamente si definisce il *tempo di prima uscita* di X da D come

$$\tilde{\tau}_D(\omega) \doteq \inf\{t \in I : X(t,\omega) \notin D\}.$$

Chiaramente, $\tilde{\tau}_D = \tau_{D^c}$ per ogni $D \in \mathcal{B}(E)$.

Proposizione 1.5. *Sia $(X(t))_{t\in I}$ un processo stocastico a valori in uno spazio metrico (E, d), adattato alla filtrazione $(\mathcal{F}_t)_{t\in I}$ e con traiettorie continue. Sia $D \in \mathcal{B}(E)$, e sia τ_D il tempo di prima entrata di X in D. Allora τ_D è un tempo d'arresto rispetto a $(\mathcal{F}_t)$*

– *se D è un chiuso, oppure*

– *se D è un aperto e la filtrazione $(\mathcal{F}_t)$ è continua a destra.*

Dimostrazione. Primo caso: D un chiuso. Allora per ogni $t \in I$

$$\{\omega \in \Omega : \tau_D(\omega) > t\} = \bigcup_{n=1}^{\infty} \bigcap_{s\in\{t\}\cup[0,t]\cap\mathbb{Q}} \left\{\inf_{x\in D} d(x, X(s,\omega)) \geq 1/n\right\}$$

grazie alla continuità delle traiettorie di X e l'ipotesi che D sia chiuso. Poiché $\{\inf_{x\in D} d(x, X(s)) \geq 1/n\} \in \mathcal{F}_s$ per ogni $s \in [0,t]$ e si usano solo intersezioni e unioni denumerabili, abbiamo $\{\tau_D > t\} \in \mathcal{F}_t$, quindi anche $\{\tau_D \geq t\} \in \mathcal{F}_t$.

Secondo caso: D un aperto. Allora per ogni $t \in I$

$$\{\omega \in \Omega : \tau_D(\omega) \geq t\} = \bigcap_{s\in\{t\}\cup[0,t]\cap\mathbb{Q}} \{X(s,\omega) \neq D\}.$$

Ne segue che $\{\tau_D \geq t\} \in \mathcal{F}_t$, quindi anche $\{\tau_D < t\} \in \mathcal{F}_t$ per ogni $t \in I$. Per concludere che $\{\tau_D \leq t\} \in \mathcal{F}_t$ si osserva che

$$\{\tau_D \leq t\} = \bigcap_{\varepsilon>0} \{\tau_D < t + \varepsilon\}$$

è in $\mathcal{F}_{t+}$, ma $\mathcal{F}_{t+} = \mathcal{F}_t$ per l'ipotesi che la filtrazione sia continua a destra. □

Lemma 1.3 (Approssimazione di tempi d'arresto)**.** *Sia $I = [0,\infty)$, e sia τ un tempo d'arresto rispetto a $(\mathcal{F}_t)_{t\geq 0}$. Allora esiste una successione $(\tau_n)_{n\in\mathbb{N}}$ di $(\mathcal{F}_t)$-tempi d'arresto tale che, per ogni $n \in \mathbb{N}$, τ_n prende valori in un insieme numerabile, $\{\tau_n = \infty\} = \{\tau = \infty\}$ e, per ogni $\omega \in \Omega$, $\tau_n(\omega) \searrow \tau(\omega)$ per $n \to \infty$.*

Dimostrazione. Poniamo per $n \in \mathbb{N}$, $\omega \in \Omega$,

$$\tau_n(\omega) \doteq \begin{cases} 0 & \text{se } \tau(\omega) = 0, \\ \frac{k}{2^n} & \text{se } \tau(\omega) \in (\frac{k-1}{2^n}, \frac{k}{2^n}] \text{ per un } k \in \mathbb{N}, \\ \infty & \text{se } \tau(\omega) = \infty. \end{cases}$$

Allora τ_n prende valori nell'insieme denumerabile $\{k/2^n : k \in \mathbb{N}_0\} \cup \{\infty\}$, $\{\tau_n = \infty\} = \{\tau = \infty\}$ e, per ogni $\omega \in \Omega$, $\tau_n(\omega) \searrow \tau(\omega)$ per $n \to \infty$. Poiché

$$\{\tau_n \leq t\} = \bigcup_{k\in\mathbb{N}_0:k\leq 2^n t} \left\{\tau_n = \frac{k}{2^n}\right\} = \bigcup_{k\in\mathbb{N}_0:k\leq 2^n t} \left\{\frac{k-1}{2^n} < \tau \leq \frac{k}{2^n}\right\} \in \mathcal{F}_t$$

per ogni $t \geq 0$, si ha che τ_n è un tempo d'arresto rispetto a $(\mathcal{F}_t)$. □

Il seguente risultato viene a volte detto "teorema d'arresto per il moto browniano"; esso generalizza l'invarianza temporale del moto browniano (punto b) della Proposizione 1.2) ed è legato alla proprietà di Markov forte.

Teorema 1.5. *Sia W un moto browniano rispetto a $(\mathcal{F}_t)_{t\geq 0}$. Sia τ un tempo d'arresto finito (cioè $\tau(\omega) < \infty$ per ogni $\omega \in \Omega$). Allora*

$$Y(t) \doteq W(\tau + t) - W(\tau), \quad t \geq 0,$$

è un moto browniano rispetto alla filtrazione $(\mathcal{F}_{\tau+t})_{t\geq 0}$.

Dimostrazione. Poniamo $\mathcal{G}_t \doteq \mathcal{F}_{\tau+t}$, $t \geq 0$. Il processo Y ha traiettorie continue $\mathbf{P}$-quasi certamente, $Y(0) = 0$, e Y è adattato alla filtrazione $(\mathcal{G}_t)$ poiché $W(\tau+t)$ è $\mathcal{F}_{\tau+t}$-misurabile per ogni $t \geq 0$. Resta mostrare che, per ogni scelta di $s \leq t$, $Y(t) - Y(s)$ è indipendente da $\mathcal{G}_s$ con distribuzione $N(0, t - s)$.

Siano $s \leq t$ fissati. Mostriamo che $Y(t) - Y(s)$ è indipendente da $\mathcal{G}_s$ con distribuzione $N(0, t - s)$ prima per τ un tempo d'arresto a valori in un insieme denumerabile, poi nel caso generale.

Primo passo. Supponiamo che τ prenda valori in $\{t_j : j \in \{1, \ldots, N\}\}$ per un $N \in \mathbb{N}$ oppure in $\{t_j : j \in \mathbb{N}\}$, ove $0 \leq t_j < t_{j+1}$ per ogni j. Allora, grazie alla Proposizione 1.2, per ogni $B \in \mathcal{B}(\mathbb{R})$, ogni $A \in \mathcal{G}_s$,

$$\begin{aligned}
&\mathbf{P}\left(\{Y(t) - Y(s) \in B\} \cap A\right) \\
&= \sum_j \mathbf{P}\left(\{W(t_j + t) - W(t_j + s) \in B\} \cap A \cap \{\tau = t_j\}\right) \\
&= \sum_j \mathbf{P}\left(W(t_j + t) - W(t_j + s) \in B\right) \cdot \mathbf{P}\left(A \cap \{\tau = t_j\}\right) \\
&= \mathbf{P}\left(W(t) - W(s) \in B\right) \cdot \sum_j \mathbf{P}\left(A \cap \{\tau = t_j\}\right) \\
&= \mathbf{P}\left(W(t) - W(s) \in B\right) \cdot \mathbf{P}(A).
\end{aligned}$$

Secondo passo. Sia τ un tempo d'arresto finito, ma del resto generale. Scegliamo una successione di tempi d'arresto (τ_n) che approssimi τ secondo il Lemma 1.3. Allora per ogni $\varphi \in \mathbf{C}_b(\mathbb{R})$, $A \in \mathcal{G}_s$, $n \in \mathbb{N}$,

$$\mathbf{E}\left[\varphi\left(W(\tau_n+t)-W(\tau_n+s)\right)\cdot \mathbf{1}_A\right] = \mathbf{E}\left[\varphi\left(W(t)-W(s)\right)\cdot \mathbf{1}_A\right]$$

grazie al primo passo. Per la continuità delle traiettorie di W, la scelta dei tempi d'arresto τ_n e la convergenza dominata abbiamo

$$\mathbf{E}\left[\varphi\left(W(\tau_n+t)-W(\tau_n+s)\right)\cdot \mathbf{1}_A\right] \overset{n\to\infty}{\longrightarrow} \mathbf{E}\left[\varphi\left(W(\tau+t)-W(\tau+s)\right)\cdot \mathbf{1}_A\right].$$

Ne segue che

$$\mathbf{E}\left[\varphi\left(Y(t)-Y(s)\right)\cdot \mathbf{1}_A\right] = \mathbf{E}\left[\varphi\left(W(t)-W(s)\right)\cdot \mathbf{1}_A\right]$$

per ogni $\varphi \in \mathbf{C}_b(\mathbb{R})$, ogni $A \in \mathcal{G}_s$. Questo implica che $Y(t)-Y(s)$ è indipendente da $\mathcal{G}_s$ con distribuzione $N(0, t-s)$. □

Teorema 1.6 (Principio di riflessione)**.** *Sia $(W(t))_{t\geq 0}$ un moto browniano su $(\Omega, \mathcal{F}, \mathbf{P})$. Per $a > 0$ poniamo $\tau_a(\omega) \doteq \inf\{t \geq 0 : W(t,\omega) \geq a\}$, $\omega \in \Omega$. Allora per ogni $t \geq 0$*

$$\mathbf{P}\left(\sup_{s\in[0,t]} W(s) \geq a\right) = \mathbf{P}\left(\tau_a \leq t\right) = 2\,\mathbf{P}\left(W(t) \geq a\right).$$

Dimostrazione. Siano $t \geq 0$, $a > 0$ fissati. Possiamo supporre che tutte le traiettorie di W siano continue. Allora

$$\left\{\omega \in \Omega : \sup_{s\in[0,t]} W(s,\omega) \geq a\right\} = \left\{\omega \in \Omega : \tau_a(\omega) \leq t\right\}.$$

E' quindi sufficiente mostrare che $\mathbf{P}(\tau_a \leq t) = 2\,\mathbf{P}(W(t) \geq a)$. A questo scopo, basta mostrare che le trasformate di Laplace delle due funzioni

$$[0,\infty) \ni t \mapsto \mathbf{P}\left(W(t) \geq a\right) \qquad e \qquad [0,\infty) \ni t \mapsto \frac{1}{2}\,\mathbf{P}\left(\tau_a \leq t\right)$$

coincidono. La trasformata di Laplace di una funzione $f\colon [0,\infty) \to \mathbb{R}$ misurabile e limitata è definita come

$$\mathcal{L}(f)(\lambda) \doteq \int_0^\infty e^{-\lambda t} f(t)dt, \quad \lambda > 0.$$

[L'ipotesi di limitatezza è più forte del necessario.] Se f, g sono due funzioni $[0,\infty) \to \mathbb{R}$ continue a destra (oppure continue a sinistra) e limitate tali che $\mathcal{L}(f)(\lambda) = \mathcal{L}(g)(\lambda)$ per ogni $\lambda > 0$ allora $f = g$.

Le due funzioni

$$[0,\infty) \ni t \mapsto \mathbf{P}\left(W(t) \geq a\right) \in [0,1] \quad e \quad [0,\infty) \ni t \mapsto \frac{1}{2}\,\mathbf{P}\left(\tau_a \leq t\right) \in [0, \tfrac{1}{2}]$$

sono limitate, non decrescenti e continue a destra (in verità continue). Per verificare la continuità a destra si può osservare che per $t \geq 0$, $h \geq 0$

$$\begin{aligned}\mathbf{P}\left(W(t+h) \geq a\right) - \mathbf{P}\left(W(t) \geq a\right) &= \frac{1}{\sqrt{2\pi t}} \int_a^\infty e^{-\frac{x^2}{2t}} \left(\frac{\sqrt{t}}{\sqrt{t+h}} e^{\frac{x^2 h}{2(t+h)t}} - 1 \right) dx \\ &\overset{h\to 0}{\longrightarrow} 0\end{aligned}$$

grazie alla convergenza dominata, mentre

$$\begin{aligned}\mathbf{P}\left(\tau_a \leq t+h\right) - \mathbf{P}\left(\tau_a \leq t\right) &= \mathbf{P}\left(\sup_{s\in[0,t]} W(s) < a, \sup_{s\in[t,t+h]} W(s) \geq a \right) \\ &\overset{h\to 0+}{\longrightarrow} 0,\end{aligned}$$

poiché le traiettorie di W sono continue quasi certamente e per ogni funzione $\varphi\colon [0,\infty) \to \mathbb{R}$ continua $\sup_{s\in[0,t]} \varphi(s) < a$ implica che esiste $\delta = \delta(\varphi) > 0$ tale che $\varphi(s) < a$ per ogni $s \in [t, t+\delta]$.

Per la trasformata di Laplace di $t \mapsto \mathbf{P}\left(W(t) \geq a\right)$, utilizzando il teorema di Fubini, il teorema d'arresto per il moto browniano (Teorema 1.5) ed il fatto che $W(\tau_a) = a$ $\mathbf{P}$-q.c., troviamo

$$\begin{aligned}&\mathcal{L}\left(\mathbf{P}\left(W(.) \geq a\right)\right)(\lambda) \\ &= \int_0^\infty e^{-\lambda t}\, \mathbf{P}\left(W(t) \geq a\right) dt \\ &= \mathbf{E}\left[\int_0^\infty e^{-\lambda t} \mathbf{1}_{[a,\infty)}\left(W(t)\right) dt\right] \\ &= \mathbf{E}\left[\int_{\tau_a}^\infty e^{-\lambda t} \mathbf{1}_{[a,\infty)}\left(W(t)\right) dt\right] \\ &= \mathbf{E}\left[\int_0^\infty e^{-\lambda(t+\tau_a)} \mathbf{1}_{[a,\infty)}\left(W(t+\tau_a) - W(\tau_a) + W(\tau_a)\right) dt\right] \\ &= \mathbf{E}\left[\int_0^\infty e^{-\lambda(t+\tau_a)} \mathbf{1}_{[0,\infty)}\left(W(t+\tau_a) - W(\tau_a)\right) dt\right] \\ &= \mathbf{E}\left[e^{-\lambda \tau_a}\right] \int_0^\infty e^{-\lambda t}\, \mathbf{P}\left(W(t+\tau_a) - W(\tau_a) \geq 0\right) dt \\ &= \mathbf{E}\left[e^{-\lambda \tau_a}\right] \frac{1}{2} \int_0^\infty e^{-\lambda t} dt \\ &= \frac{1}{2\lambda}\, \mathbf{E}\left[e^{-\lambda \tau_a}\right].\end{aligned}$$

La trasformata di Laplace di $t \mapsto \frac{1}{2}\, \mathbf{P}\left(\tau_a \leq t\right)$, dall'altra parte, si può riscrivere come

$$\begin{aligned}\mathcal{L}\left(\frac{1}{2}\, \mathbf{P}\left(\tau_a \leq .\right)\right)(\lambda) &= \frac{1}{2} \int_0^\infty e^{-\lambda t}\, \mathbf{P}\left(\tau_a \leq t\right) dt \\ &= \frac{1}{2}\, \mathbf{E}\left[\int_0^\infty e^{-\lambda t} \mathbf{1}_{[0,t]}\left(\tau_a\right) dt\right] \\ &= \frac{1}{2}\, \mathbf{E}\left[\int_{\tau_a}^\infty e^{-\lambda t} dt\right]\end{aligned}$$

$$= \frac{1}{2\lambda}\mathbf{E}\left[e^{-\lambda\tau_a}\right].$$

Le due trasformate, di conseguenza le due funzioni, coincidono. □

Fonti: Sezioni 1.2 e 2.5 in Baldi [2000, pp. 26-27 e 42-44];

1.5 Spazio canonico e principio di invarianza

Sezione 2.4 in Karatzas and Shreve [1991, pp. 59-71]; a lezione senza dimostrazioni. Versione breve nella Sezione 2.7 in Caravenna [2011, pp. 46-48].

1.6 Moto browniano multidimensionale

Sia $(\Omega, \mathcal{F}, \mathbf{P})$ uno spazio di probabilità, $(\mathcal{F}_t)_{t\geq 0}$ una filtrazione in $\mathcal{F}$.

Definizione 1.5. Un processo stocastico $(W(t))_{t\geq 0}$ a valori in $\mathbb{R}^d$ definito su $(\Omega, \mathcal{F}, \mathbf{P})$ si dice *moto browniano d-dimensionale* rispetto a $(\mathcal{F}_t)$ se

(i) $W(0) = 0 \in \mathbb{R}^d$ **P**-quasi certamente,

(ii) W è $(\mathcal{F}_t)$-adattato,

(iii) per ogni $t > s \geq 0$, $W(t) - W(s)$ ha distribuzione $N(0, (t-s)\,\mathrm{Id}_d)$ ed è indipendente da $\mathcal{F}_s$,

(iv) W ha traiettorie continue **P**-quasi certamente.

Quando la filtrazione $(\mathcal{F}_t)$ non viene specificata si intende per $(\mathcal{F}_t)$, come al solito, la filtrazione naturale del processo.

Proposizione 1.6. *Sia $(W(t))_{t\geq 0}$ un moto browniano d-dimensionale rispetto alla filtrazione $(\mathcal{F}_t)_{t\geq 0}$, $W = (W_1, \ldots, W_d)^{\mathsf{T}}$.*

a) I processi $W_1, \ldots, W_d$ a valori in $\mathbb{R}$ sono dei moti browniani rispetto a $(\mathcal{F}_t)$ indipendenti (cioè le σ-algebre $\sigma(W_i(t) : t \geq 0)$, $i \in \{1, \ldots, d\}$), formano una famiglia indipendente). La σ-algebra degli incrementi $\sigma(W(t) - W(s) : t \geq s)$ è indipendente da $\mathcal{F}_s$ per ogni $s \geq 0$.

b) Sia $z \in \mathbb{R}^d$ con $|z| = 1$. Allora $\langle z, W\rangle$ è un moto browniano rispetto a $(\mathcal{F}_t)$.

c) Sia A una matrice $d \times d$. Il processo X dato da $X(t,\omega) \doteq AW(t,\omega)$, $t \geq 0$, $\omega \in \Omega$, X è un moto browniano d-dimensionale se e solo se A è una matrice ortogonale.

d) Sia $B \in \mathcal{B}(\mathbb{R}^d)$ un insieme di misura di Lebesgue finita e strettamente positiva. Allora

$$\mathbf{E}\left[\int_0^\infty \mathbf{1}_B(W(t))dt\right] = \begin{cases} \infty & \text{se } d \in \{1,2\} \\ \frac{1}{2}\pi^{-d/2}\Gamma\left(\frac{d}{2}-1\right)\int_B |x|^{2-d}dx & \text{se } d \geq 3, \end{cases}$$

dove $\Gamma(.)$ indica la funzione Gamma.

Dimostrazione. **a)** Il fatto che, per $i \in \{1, \ldots, d\}$, W_i sia un moto browniano rispetto a $(\mathcal{F}_t)$ segue direttamente dalla definizione di un moto browniano (uno-dimensionale) rispetto a una filtrazione. Per mostrare l'indipendenza di $W_1, \ldots, W_d$ si osserva che $(W_i(t))_{i \in \{1,\ldots,d\}, t \geq 0}$ è una famiglia gaussiana. Per il Lemma 0.2, la gaussianità congiunta e il fatto che le $W_i(t)$ siano centrate, basta verificare che, per $i \neq j$, $0 \leq s \leq t$,

$$\mathbf{E}\left[W_i(t)W_j(s)\right] = \mathbf{E}\left[\left(W_i(t) - W_i(s)\right)W_j(s)\right] + \mathbf{E}\left[W_i(s)W_j(s)\right] = 0;$$

il che è vero perché $(W_i(t) - W_i(s))$ è indipendente da $\mathcal{F}_s$ e $W_j(s)$ è $\mathcal{F}_s$-misurabile, mentre $\mathbf{E}\left[W_i(s)W_j(s)\right] = 0$ poiché $(W_1, \ldots, W_d)$ ha media zero e matrice di covarianza $s\,\mathrm{Id}_d$.

b) Sia $z \in \mathbb{R}^d$. Allora $\langle z, W\rangle$ è un processo gaussiano a valori in $\mathbb{R}$ centrato con traiettorie continue e $\langle z, W(0)\rangle = 0$ $\mathbf{P}$-quasi certamente. Inoltre $\langle z, W\rangle$ è $(\mathcal{F}_t)$-adattato. Per $t \geq s$, abbiamo che $\langle z, W(t)\rangle - \langle z, W(s)\rangle = \langle z, W(t) - W(s)\rangle$ è indipendente da $\mathcal{F}_s$, di conseguenza e per il punto a)

$$\begin{aligned}
\mathbf{E}\left[\langle z, W(t)\rangle\langle z, W(s)\rangle\right] &= \mathbf{E}\left[\langle z, W(t) - W(s)\rangle\langle z, W(s)\rangle\right] + \mathbf{E}\left[\left|\langle z, W(s)\rangle\right|^2\right] \\
&= 0 + \mathbf{E}\left[\sum_{i,j=1}^{d} z_i z_j W_i(s) W_j(s)\right] \\
&= \mathbf{E}\left[\sum_{i=1}^{d} z_i^2 W_i^2(s)\right] \\
&= s|z|^2,
\end{aligned}$$

che dà la funzione di covarianza di un moto browniano standard se (e solo se) $|z| = 1$. In alternativa al calcolo di sopra si possono usare le trasformazioni lineari di un processo gaussiano.

c) Il processo X è gaussiano centrato con traiettorie continue e $X(0) = 0$ $\mathbf{P}$-quasi certamente. Inoltre X è $(\mathcal{F}_t)$-adattato e, per $t \geq s$, l'incremento $X(t) - X(s) = A(W(t) - W(s))$ è indipendente da $\mathcal{F}_s$ con legge gaussiana di media zero e matrice di covarianza $A(t-s)\,\mathrm{Id}_d\, A^{\mathsf{T}} = (t-s)AA^{\mathsf{T}}$ (trasformazione lineare). Affinché X sia un moto browniano d-dimensionale dobbiamo avere $AA^{\mathsf{T}} = \mathrm{Id}_d$, cioè A una matrice ortogonale.

d) Per il teorema di Fubini e usando la densità di $N(0, t\,\mathrm{Id}_d)$,

$$\begin{aligned}
\mathbf{E}\left[\int_0^\infty \mathbf{1}_B(W(t))dt\right] &= \int_0^\infty \mathbf{P}\left(W(t) \in B\right) dt \\
&= \int_0^\infty \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} \mathbf{1}_B(x) e^{-\frac{|x|^2}{2t}} dx\, dt \\
&= \int_B \int_0^\infty \frac{1}{(2\pi t)^{d/2}} e^{-\frac{|x|^2}{2t}} dt\, dx.
\end{aligned}$$

Se $d = 1$, allora per ogni $x \in \mathbb{R}$

$$\int_0^\infty \frac{1}{(2\pi t)^{d/2}} e^{-\frac{|x|^2}{2t}} dt \geq \frac{e^{-1/2}}{\sqrt{2\pi}} \int_{|x|^2}^\infty \frac{1}{\sqrt{t}} dt = \infty,$$

in maniera simile, se $d = 2$, per ogni $x \in \mathbb{R}^2$

$$\int_0^\infty \frac{1}{(2\pi t)^{d/2}} e^{-\frac{|x|^2}{2t}} dt \geq \frac{e^{-1/2}}{2\pi} \int_{|x|^2}^\infty \frac{1}{t} dt = \infty.$$

Se $d \geq 3$, allora per ogni $x \in \mathbb{R}^2 \setminus \{0\}$,

$$\begin{aligned}
&\int_0^\infty \frac{1}{(2\pi t)^{d/2}} e^{-\frac{|x|^2}{2t}} dt && \mid \text{cambio di variabili } s \doteq \frac{1}{t} \\
&= \int_0^\infty \frac{1}{(2\pi)^{d/2}} s^{d/2-2} e^{-\frac{|x|^2}{2}s} ds && \mid \text{cambio di variabili } u \doteq \frac{|x|^2}{2} s \\
&= \frac{1}{(2\pi)^{d/2}} \int_0^\infty \left(\frac{2}{|x|^2} u\right)^{d/2-2} e^{-u} \frac{2}{|x|^2} du \\
&= \frac{1}{2} \pi^{-d/2} |x|^{2-d} \int_0^\infty u^{d/2-2} e^{-u} du \\
&= \frac{1}{2} \pi^{-d/2} |x|^{2-d} \Gamma\left(d/2 - 1\right).
\end{aligned}$$

□

Proposizione 1.7. *Sia $(W(t))_{t\geq 0}$ un moto browniano d-dimensionale e B la palla aperta di raggio uno e centro l'origine in $\mathbb{R}^d$. Sia τ il tempo di prima uscita di W da B, cioè $\tau(\omega) \doteq \inf\{t \geq 0 : W(t, \omega) \notin B\}$, $\omega \in \Omega$. Allora $W(\tau)$ può essere considerata una variabile aleatoria a valori in ∂B (ovvero nella sfera unitaria $(d-1)$-dimensionale). La legge di $W(\tau)$ è la misura di Haar con massa uno su ∂B, e le variabili aleatorie τ e $W(\tau)$ sono indipendenti.*

Dimostrazione. Basta applicare la legge del logaritmo iterato a una delle componenti di W (che sono moti browniani unodimensionali) per vedere che W lascia ogni insieme limitato con probabilità uno. Di conseguenza, $\tau < \infty$ $\mathbf{P}$-quasi certamente e $W(\tau)$ può essere considerata una variabile aleatoria a valori in ∂B.

La misura di Haar su ∂B è la misura di probabilità sui boreliani di ∂B che è invariante rispetto al gruppo di isometrie della sfera, cioè rispetto al gruppo delle trasformazioni ortogonali. Sia Q la legge di $W(\tau)$, cioè $Q(U) \doteq \mathbf{P}\left(W(\tau) \in U\right)$ per $U \in \mathcal{B}(\partial B)$. Dobbiamo mostrare che $Q(U) = Q(AU)$ per ogni $U \in \mathcal{B}(\partial B)$, ogni matrice $d \times d$ A ortogonale. Sia A ortogonale e poniamo $Y \doteq A^{-1}W$. Per il punto c) della Proposizione 1.6 abbiamo che Y è un moto browniano d-dimensionale (con A anche $A^{-1} = A^\mathsf{T}$ è ortogonale). Poiché $|Y| = |AW| = |W|$, τ è anche il tempo di prima uscita di Y dalla palla B. Ne segue che

$$\mathbf{P}\left(Y(\tau) \in U\right) = Q(U) \text{ per ogni } U \in \mathcal{B}(\partial B).$$

Dall'altra parte, per la definizione di Y,

$$\mathbf{P}\left(Y(\tau) \in U\right) = \mathbf{P}\left(W(\tau) \in AU\right) = Q(AU) \text{ per ogni } U \in \mathcal{B}(\partial B).$$

Per quanto riguarda l'indipendenza di τ e $W(\tau)$, dobbiamo mostrare che $\mathbf{P}\left(W(\tau) \in U, \tau \in J\right) = \mathbf{P}\left(W(\tau) \in U\right) \mathbf{P}\left(\tau \in J\right)$ per ogni $U \in \mathcal{B}(\partial B)$, ogni

$J \in \mathcal{B}([0,\infty))$. Sia $J \in \mathcal{B}([0,\infty))$ e poniamo $\mu_J(U) \doteq \mathbf{P}\left(W(\tau) \in U, \tau \in J\right)$, $U \in \mathcal{B}(\partial B)$. Ragionando come sopra si vede che μ_J è invariante rispetto alle trasformazioni ortogonali. Per l'unicità della misura di Haar, esiste una costante $c(J) \geq 0$ tale che $\mu_J = c(J) \cdot Q$ con Q la legge di $W(\tau)$). In particolare $\mu_J(\partial B) = c(J) \cdot Q(\partial B) = c(J)$ poiché $Q(\partial B) = 1$. Di conseguenza, per ogni $U \in \mathcal{B}(\partial B)$,

$$\begin{aligned}
\mathbf{P}\left(W(\tau) \in U, \tau \in J\right) &= \mu_J(U) \\
&= \mu_J(\partial B) \cdot Q(U) \\
&= \mathbf{P}\left(W(\tau) \in \partial B, \tau \in J\right) \cdot \mathbf{P}\left(W(\tau) \in U\right) \\
&= \mathbf{P}\left(\tau \in J\right) \cdot \mathbf{P}\left(W(\tau) \in U\right).
\end{aligned}$$

$\square$

Capitolo 2

Martingale

2.1 Martingale (sub- e super-)

Sia $(\Omega, \mathbf{P}, \mathcal{F})$ uno spazio di probabilità, $(\mathcal{F}_t)_{t\in I}$ una filtrazione in $\mathcal{F}$, $I \subset [-\infty, \infty]$ l'insieme dei tempi (in particolare $I = \mathbb{N}_0$ oppure $I = [0, \infty)$).

Definizione 2.1. Sia $(X(t))_{t\in I}$ un processo stocastico reale definito su $(\Omega, \mathcal{F}, \mathbf{P})$. Il processo X si dice una *martingala* rispetto alla filtrazione $(\mathcal{F}_t)_{t\in I}$ se

(i) $X(t) \in L^1(\Omega, \mathcal{F}_t, \mathbf{P})$ per ogni $t \in I$;

(ii) $\mathbf{E}\left[X(t)|\mathcal{F}_s\right] = X(s)$ $\mathbf{P}$-q.c. per ogni $s, t \in I$ con $s \leq t$.

Il processo X si dice una *submartingala* rispetto a $(\mathcal{F}_t)_{t\in I}$ se vale la condizione (i) e $\mathbf{E}\left[X(t)|\mathcal{F}_s\right] \leq X(s)$ $\mathbf{P}$-q.c. per ogni $s \leq t$. Il processo X si dice una *supermartingala* rispetto a $(\mathcal{F}_t)_{t\in I}$ se vale la (i) e $\mathbf{E}\left[X(t)|\mathcal{F}_s\right] \geq X(s)$ $\mathbf{P}$-q.c. per ogni $s \leq t$.

Per la condizione (i), una (sub-/super-)martingala è un processo adattato ed integrabile. Quando la filtrazione non viene specificata e non risulta dal contesto si intende quella naturale del processo.

- Un processo reale X è una submartingala se e solo se $-X$ è una supermartingala; X è una martingala se e solo se X è una submartingala e una supermartingala.
- La funzione di media $I \ni t \mapsto \mathbf{E}[X(t)]$ di un processo reale X è
  - costante se X è una martingala,
  - crescente se X è una submartingala,
  - decrescente se X è una supermartingala.
- Una submartingala (oppure supermartingala) con funzione di media costante è una martingala.
- Combinazioni lineari di martingale sono martingale.

- Combinazioni lineari a coefficienti non-negativi di submartingale sono submartingale; analogamente per le supermartingale.

- Sia $\bar{X} \in L^1(\Omega, \mathcal{F}, \mathbf{P})$. Poniamo, per $t \in I$, $X(t) \doteq \mathbf{E}\left[\bar{X}|\mathcal{F}_t\right]$. Allora $(X(t))_{t\in I}$ è una martingala rispetto a $(\mathcal{F}_t)_{t\in I}$.

- Sia $(X(t))_{t\in I}$ una submartingala rispetto a $(\mathcal{F}_t)_{t\in I}$. Sia $T \in I$. Allora $X(t) \leq \mathbf{E}[X(T)|\mathcal{F}_t]$ per ogni $t \leq T$. Se I possiede un elemento massimale T_* (cioè $T_* \geq t$ per ogni $t \in I$), allora $X(t) \leq \mathbf{E}[X(T_*)|\mathcal{F}_t] \doteq Y(t)$ per ogni $t \in I$ e Y è una martingala. "Una submartingala è sotto la sua martingala."

Proposizione 2.1. *Sia $(X(t))_{t\in I}$ un processo stocastico reale su $(\Omega, \mathcal{F}, \mathbf{P})$.*

a) Sia $\varphi\colon \mathbb{R} \to \mathbb{R}$ una funzione convessa tale che $\varphi(X(t)) \in L^1(\mathbf{P})$ per ogni $t \in I$. Se X è una martingala rispetto a $(\mathcal{F}_t)$, allora $\varphi(X)$ è una submartingala rispetto a $(\mathcal{F}_t)$.

b) Sia $\varphi\colon \mathbb{R} \to \mathbb{R}$ una funzione convessa e crescente tale che $\varphi(X(t)) \in L^1(\mathbf{P})$ per ogni $t \in I$. Se X è una submartingala rispetto a $(\mathcal{F}_t)$, allora $\varphi(X)$ è una submartingala rispetto a $(\mathcal{F}_t)$.

Dimostrazione. L'enunciato è una conseguenza dalla disuguaglianza di Jensen per la speranza condizionale. □

Dalla Proposizione 2.1 segue in particolare:

- Se $(X(t))_{t\in I}$ è una martingala in L^p per un $p \geq 1$, allora $(|X(t)|^p)_{t\in I}$ è una submartingala (rispetto alla stessa filtrazione).

- Se $(X(t))_{t\in I}$ è una submartingala, allora $(X^+(t))_{t\in I}$ è una submartingala.

- Se $(X(t))_{t\in I}$ è una submartingala in L^p per un $p \geq 1$, allora $((X^+(t))^p)_{t\in I}$ è una submartingala.

Esempio 2.1 (Passeggiata aleatoria).

Esempio 2.2 ((Sub-/super-)Martingale del moto browniano).

2.2 Arresto e stime sul massimo

Teorema 2.1 (Arresto di submartingale)**.** *Sia $(X(t))_{t\geq 0}$ una submartingala continua a destra rispetto alla filtrazione $(\mathcal{F}_t)$, e sia τ un tempo d'arresto rispetto a $(\mathcal{F}_t)$ tale che $\tau < \infty$ $\mathbf{P}$-quasi certamente.*

a) Il processo $(X(\tau \wedge t))_{t\geq 0}$ è una submartingala sia rispetto a $(\mathcal{F}_t)_{t\geq 0}$ che a $(\mathcal{F}_{\tau\wedge t})_{t\geq 0}$.

b) Se τ è quasi certamente limitato (cioè se esiste una costante determinista $C \in [0,\infty)$ tale che $\tau \leq C$ $\mathbf{P}$-q.c.) oppure $|X(\tau \wedge t)| \leq Y$ per ogni $t \geq 0$ e una variabile aleatoria $Y \in L^1(\mathbf{P})$, allora

$$\mathbf{E}\left[X(\tau)\right] \geq \mathbf{E}\left[X(0)\right].$$

Se X è una martingala, allora $(X(\tau \wedge t))_{t\geq 0}$ è una martingala e si ha, sotto le ipotesi aggiuntive su τ, $\mathbf{E}\left[X(\tau)\right] = \mathbf{E}\left[X(0)\right]$.

Teorema 2.2 (disuguaglianze massimali e di Doob)**.** *Sia $(X(t))_{t\geq 0}$ una submartingala continua a destra rispetto alla filtrazione $(\mathcal{F}_t)$.*

1. *Per ogni $T \geq 0$, ogni $\lambda > 0$,*

$$\mathbf{P}\left(\sup_{t\in[0,T]} X(t) \geq \lambda\right) \leq \frac{1}{\lambda}\,\mathbf{E}\left[0 \vee X(T)\right],$$
$$\mathbf{P}\left(\inf_{t\in[0,T]} X(t) \leq -\lambda\right) \leq \frac{1}{\lambda}\left(\mathbf{E}\left[0 \vee X(T)\right] - \mathbf{E}\left[X(0)\right]\right) \leq \frac{1}{\lambda}\,\mathbf{E}\left[|X(T)|\right].$$

2. *Se X è una martingala oppure una submartingala non-negativa, allora per ogni $T \geq 0$, ogni $p > 1$,*

$$\mathbf{E}\left[\sup_{t\in[0,T]} |X(t)|^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbf{E}\left[|X(T)|^p\right].$$

In particolare, con $p = 2$, $\mathbf{E}\left[\sup_{t\in[0,T]} |X(t)|^2\right] \leq 4\,\mathbf{E}\left[|X(T)|^2\right]$.

2.3 Convergenza

2.4 Compensazione e variazione quadratica

[appunti]

Def. variazione quadratica / processo crescente associato per una martingala locale (continua).

Variazione quadratica del processo arrestato: $\langle Y(\tau \wedge \cdot)\rangle(t) = \langle Y\rangle(\tau \wedge t)$.

Teorema 2.3 (Caratterizzazione del moto browniano, P. Lévy)**.** *Sia X una martingala locale rispetto a $(\mathcal{F}_t)$ con $X(0) = 0$ e traiettorie continue $\mathbf{P}$-quasi certamente. Supponiamo inoltre che per la variazione quadratica di X valga con probabilità uno*

$$\langle X\rangle(t) = t \text{ per ogni } t \geq 0.$$

Allora X è un moto browniano uno-dimensionale rispetto a $(\mathcal{F}_t)$.

Dimostrazione. Per le ipotesi, esiste una successione $(\tau_n)_{n\in\mathbb{N}}$ di tempi d'arresto con $\tau_n \nearrow \infty$ **P**-quasi certamente tale che, per ogni $n \in \mathbb{N}$, $X(\tau_n \wedge \cdot)$ è una $(\mathcal{F}_t)$-martingala in L^2 di media zero e $(X^2(\tau_n \wedge t) - \tau_n \wedge t)_{t\geq 0}$ è una $(\mathcal{F}_t)$-martingala. Grazie a queste proprietà di martingala e la disuguaglianza di Doob,

$$\mathbf{E}\left[\sup_{s\in[0,t]} X^2(\tau_n \wedge s)\right] \leq 4\,\mathbf{E}\left[X^2(\tau_n \wedge t)\right] = 4\,\mathbf{E}\left[\tau_n \wedge t\right] \leq 4t.$$

Questo implica, per il lemma di Fatou e per la convergenza **P**-quasi certa $\sup_{s\in[0,t]} X^2(\tau_n \wedge s) \to \sup_{s\in[0,t]} X^2(s)$, che $\sup_{s\in[0,t]} X^2(s) \in L^1(\mathbf{P})$ per ogni $t \geq 0$. In particolare, X è un processo di quadrato integrabile. Inoltre, per $t \geq s \geq 0$, $n \in \mathbb{N}$,

$$\begin{aligned}
&\mathbf{E}\left[X^2(t) - t \middle| \mathcal{F}_s\right] \\
&\quad = \mathbf{E}\left[X^2(t) - X^2(\tau_n \wedge t) \middle| \mathcal{F}_s\right] + X^2(\tau_n \wedge s) - \tau_n \wedge s + \mathbf{E}\left[\tau_n \wedge t - t \middle| \mathcal{F}_s\right] \\
&\qquad\qquad \overset{n\to\infty}{\longrightarrow} X^2(s) - s,
\end{aligned}$$

poiché $\mathbf{E}\left[X^2(t) - X^2(\tau_n \wedge t) \middle| \mathcal{F}_s\right] \to 0$ per $n \to \infty$ grazie al teorema della convergenza dominata con $2\sup_{s\in[0,t]} X^2(s)$ come variabile integrabile dominante e $\mathbf{E}\left[\tau_n \wedge t - t \middle| \mathcal{F}_s\right] \to 0$ per convergenza dominata o convergenza monotona, mentre $X^2(\tau_n \wedge s) - \tau_n \wedge s \to X^2(s) - s$ poiché $\tau_n \nearrow \infty$ **P**-quasi certamente. Di conseguenza, $(X^2(t) - t)_{t\geq 0}$ è una $(\mathcal{F}_t)$-martingala. Analogamente si verifica che X è una $(\mathcal{F}_t)$-martingala.

Basta quindi dimostrare l'enunciato supponendo che X sia una martingala vera di quadrato integrabile con $X(0) = 0$, traiettorie continue e variazione quadratica $\langle X\rangle(t) = t$, $t \geq 0$. In questa situazione, resta da mostrare che, per ogni $t \geq s \geq 0$, $X(t) - X(s)$ abbia distribuzione $N(0, t-s)$ e sia indipendente da $\mathcal{F}_s$. E' equivalente verificare che, per ogni $t \geq s \geq 0$, ogni $y \in \mathbb{R}$,

$$\mathbf{E}\left[e^{\sqrt{-1}y(X(t)-X(s))} \middle| \mathcal{F}_s\right] = e^{-\frac{1}{2}y^2(t-s)} \quad \mathbf{P}\text{-q.c.}$$

Siano $t \geq s \geq 0$, $y \in \mathbb{R}$. Per $n \in \mathbb{N}$, $i \in \{1, \ldots, n\}$ poniamo

$$t_0^n \doteq s, \qquad t_i^n \doteq s + \tfrac{i}{n}(t-s), \qquad \Delta_i^n(X) \doteq X\left(t_i^n\right) - X\left(t_{i-1}^n\right).$$

Allora

$$\mathbf{E}\left[e^{\sqrt{-1}y(X(t)-X(s))} \middle| \mathcal{F}_s\right] = \mathbf{E}\left[\prod_{i=1}^{n} \exp\left(\sqrt{-1}y\Delta_i^n(X)\right) \middle| \mathcal{F}_s\right].$$

Usiamo ora la formula di Taylor applicata alla mappa $x \mapsto \exp(\sqrt{-1}x)$:

$$e^{\sqrt{-1}x} = 1 + \sqrt{-1}x - \frac{1}{2}x^2 + \mathcal{O}(x^3), \quad x \in \mathbb{R}.$$

Indicando con R_n la variabile aleatoria a valori in $\mathbb{C}$ che risulta dagli errori di approssimazione di secondo ordine, possiamo scrivere

$$\begin{aligned}
&\mathbf{E}\left[\prod_{i=1}^{n} \exp\left(\sqrt{-1}y\Delta_i^n(X)\right)\middle|\,\mathcal{F}_s\right] \\
&\qquad = \mathbf{E}\left[\prod_{i=1}^{n}\left(1+\sqrt{-1}y\Delta_i^n(X)-\frac{1}{2}y^2\Delta_i^n(X)^2\right)\middle|\,\mathcal{F}_s\right] + \mathbf{E}\left[R_n\middle|\,\mathcal{F}_s\right].
\end{aligned}$$

Senza dare i dettagli, notiamo che $|\mathbf{E}\left[R_n|\,\mathcal{F}_s\right]| \to 0$ in probabilità per $n\to\infty$ grazie all'integrabilità e alla continuità delle traiettorie di X. Dall'altra parte,

$$\begin{aligned}
&\mathbf{E}\left[\prod_{i=1}^{n}\left(1+\sqrt{-1}y\Delta_i^n(X)-\frac{1}{2}y^2\Delta_i^n(X)^2\right)\middle|\,\mathcal{F}_s\right] \\
&= \mathbf{E}\left[\mathbf{E}\left[\prod_{i=1}^{n}\left(1+\sqrt{-1}y\Delta_i^n(X)-\frac{1}{2}y^2\Delta_i^n(X)^2\right)\middle|\,\mathcal{F}_{t_{n-1}^n}\right]\middle|\,\mathcal{F}_s\right] \\
&= \mathbf{E}\left[\mathbf{E}\left[\left(1+\sqrt{-1}y\Delta_n^n(X)-\frac{1}{2}y^2\Delta_n^n(X)^2\right)\middle|\,\mathcal{F}_{t_{n-1}^n}\right]\prod_{i=1}^{n-1}\left(1+\sqrt{-1}y\Delta_i^n(X)-\frac{1}{2}y^2\Delta_i^n(X)^2\right)\middle|\,\mathcal{F}_s\right].
\end{aligned}$$

Ora, poiché i processi X, $X^2-\langle X\rangle$ sono delle $(\mathcal{F}_t)$-martingale e $\langle X\rangle(r)=r$, $r\geq 0$, per ipotesi,

$$\begin{aligned}
&\mathbf{E}\left[\left(1+\sqrt{-1}y\Delta_n^n(X)-\frac{1}{2}y^2\Delta_n^n(X)^2\right)\middle|\,\mathcal{F}_{t_{n-1}^n}\right] \\
&= 1+\sqrt{-1}y\underbrace{\mathbf{E}\left[X\left(t_n^n\right)-X\left(t_{n-1}^n\right)\middle|\,\mathcal{F}_{t_{n-1}^n}\right]}_{=0}-\frac{1}{2}y^2\,\mathbf{E}\left[\left(X\left(t_n^n\right)-X\left(t_{n-1}^n\right)\right)^2\middle|\,\mathcal{F}_{t_{n-1}^n}\right] \\
&= 1-\frac{1}{2}y^2\,\mathbf{E}\left[X^2\left(t_n^n\right)-X^2\left(t_{n-1}^n\right)\middle|\,\mathcal{F}_{t_{n-1}^n}\right] \\
&= 1-\frac{1}{2}y^2\left(t_n^n-t_{n-1}^n\right) \\
&= 1-\frac{y^2(t-s)}{2n}.
\end{aligned}$$

Ne segue che

$$\begin{aligned}
&\mathbf{E}\left[\prod_{i=1}^{n}\left(1+\sqrt{-1}y\Delta_i^n(X)-\frac{1}{2}y^2\Delta_i^n(X)^2\right)\middle|\,\mathcal{F}_s\right] \\
&\qquad = \left(1-\frac{y^2(t-s)}{2n}\right)\mathbf{E}\left[\prod_{i=1}^{n-1}\left(1+\sqrt{-1}y\Delta_i^n(X)-\frac{1}{2}y^2\Delta_i^n(X)^2\right)\middle|\,\mathcal{F}_s\right].
\end{aligned}$$

Proseguendo per iterazione, si ottiene che

$$\mathbf{E}\left[\prod_{i=1}^{n}\left(1+\sqrt{-1}y\Delta_i^n(X)-\frac{1}{2}y^2\Delta_i^n(X)^2\right)\middle|\,\mathcal{F}_s\right] = \left(1-\frac{y^2(t-s)}{2n}\right)^n.$$

Di conseguenza, per una sotto-successione $(n_j)_{j\in\mathbb{N}}$ tale che $\left|\mathbf{E}\left[R_{n_j}\middle|\mathcal{F}_s\right]\right| \to 0$ $\mathbf{P}$-quasi certamente per $j \to \infty$, si ha che con probabilità uno

$$\mathbf{E}\left[e^{\sqrt{-1}y(X(t)-X(s))}\middle|\mathcal{F}_s\right] = \lim_{j\to\infty} \mathbf{E}\left[\prod_{i=1}^{n_j}\left(1+\sqrt{-1}y\Delta_i^{n_j}(X)-\frac{1}{2}y^2\Delta_i^{n_j}(X)^2\right)\middle|\mathcal{F}_s\right] = e^{-\frac{1}{2}y^2(t-s)}.$$

□

Versione multi-dimensionale:

Capitolo 3

Integrale stocastico

Sezioni 3.2.A e 3.2.B in Karatzas and Shreve [1991, pp. 129-141]

3.1 Costruzione per processi semplici

3.2 Estensione a processi adattati di quadrato integrabili

3.3 L'integrale stocastico come processo

Sezione 6.3 in Baldi [2000, pp. 118-122]

3.4 Localizzazione

Siano $(\Omega, \mathcal{F}, \mathbf{P})$ uno spazio di probabilità completo, $(\mathcal{F}_t)_{t\geq 0}$ una filtrazione in $\mathcal{F}$ che soddisfa alle ipotesi standard, e W un moto browniano uno-dimensionale rispetto a $(\mathcal{F}_t)$.

$$\Lambda^2 \doteq \big\{u\colon [0,\infty)\times\Omega \to \mathbb{R} : u \text{ progressivamente misurabile rispetto a } (\mathcal{F}_t) \text{ e} \\ \int_0^T |u(t)|^2 dt < \infty \;\; \mathbf{P}\text{-q.c. per ogni } T > 0\big\}.$$

Teorema 3.1 (Integrale di Itô su Λ^2)**.** *Sia $T > 0$.*

1. *L'integrale di Itô è un operatore lineare su Λ^2_T e Λ^2.*
2. *Processo. Per ogni $u \in \Lambda^2$, il processo $(\int_0^t u(s)dW(s))_{t\geq 0}$ è una martingala locale continua rispetto a $(\mathcal{F}_t)$.*
3. *Disuguaglianza massimale. Per ogni $u \in \Lambda^2$, ogni $\varepsilon, \delta > 0$,*

$$\mathbf{P}\left(\sup_{t\in[0,T]}\left|\int_0^t u(s)dW(s)\right| > \varepsilon\right) \leq \frac{\delta}{\varepsilon^2} + \mathbf{P}\left(\int_0^T |u(t)|^2 dt > \delta\right).$$

4. *Continuità. Sia* $u \in \Lambda^2$*, e sia* $(u_n)_{n\in\mathbb{N}} \subset \Lambda^2$ *tale che* $\int_0^T |u(t)-u_n(t)|^2 dt \to 0$ *in probabilità per* $n \to \infty$*. Allora*

$$\int_0^t u_n(s)dW(s) \overset{n\to\infty}{\longrightarrow} \int_0^t u(s)dW(s)$$

in probabilità uniformemente in $t \in [0,T]$*.*

5. *Località rispetto a tempi d'arresto. Sia* $u \in \Lambda^2$*, e sia* τ *un tempo d'arresto rispetto a* $(\mathcal{F}_t)$ *tale che* $\tau \le T$ $\mathbf{P}$*-quasi certamente. Allora*

$$\int_0^\tau u(t)dW(t) = \int_0^T u(t)\cdot \mathbf{1}_{[0,\tau]}(t)dW(t) \quad \mathbf{P}\text{-}q.c.$$

6. *Località rispetto a eventi. Siano* $u, v \in \Lambda^2$*, e sia* $A \in \mathcal{F}$*. Se* $u(t,\omega) = v(t,\omega)$ *per* $\lambda \otimes \mathbf{P}$*-quasi ogni* $(t,\omega) \in [0,T] \times \mathbf{P}$*, allora per* $\mathbf{P}$*-quasi ogni* $\omega \in A$*, ogni* $t \in [0,T]$*,*

$$\left(\int_0^t u(s)dW(s)\right)(\omega) = \left(\int_0^t v(s)dW(s)\right)(\omega).$$

7. *Somme di Riemann-Itô. Sia* u *un processo reale* $(\mathcal{F}_t)$*-adattato e continuo. Allora* $u \in \Lambda^2$ *e, per ogni successione* (t_i^n) *di partizioni di* $[0,T]$ *con passo tendente a zero,*

$$\sum_i u(t_{i-1}^n)\left(W(t_i^n \wedge t) - W(t_{i-1}^n \wedge t)\right) \overset{n\to\infty}{\longrightarrow} \int_0^t u(s)dW(s)$$

in probabilità uniformemente in $t \in [0,T]$*.*

Sezione 6.4 in Baldi [2000, pp. 118-122]

Capitolo 4

Calcolo di Itô

4.1 Formula di Itô e processi di Itô

Sia $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t)$ una base stocastica standard, e sia W un moto browniano uno-dimensionale rispetto a $(\mathcal{F}_t)$.

Teorema 4.1 (Formula di Itô I)**.** *Sia* $f \in \mathbf{C}^2(\mathbb{R})$. *Allora con probabilità uno per ogni* $t \geq 0$,

$$f\big(W(t)\big) = f\big(W(0)\big) + \int_0^t f'\big(W(s)\big)dW(s) + \frac{1}{2}\int_0^t f''\big(W(s)\big)ds.$$

Dimostrazione. Per la continuità del processo $f(W)$ e dei processi di integrale, basta verificare l'equazione per $t \geq 0$ fissato. Sia (t_i^n) una successione di partizioni di $[0, t]$ con passo tendente a zero. Grazie alla formula di Taylor, per ogni $n \in N$,

$$\begin{aligned}
&f\big(W(t)\big) - f\big(W(0)\big)\\
&= \sum_{i=1}^{n} \Big(f\big(W(t_i^n)\big) - f\big(W(t_{i-1}^n)\big)\Big)\\
&= \sum_{i=1}^{n} f'\big(W(t_{i-1}^n)\big)\big(W(t_i^n) - W(t_{i-1}^n)\big) + \frac{1}{2}f''\big(W(\theta_i^n)\big)\big(W(t_i^n) - W(t_{i-1}^n)\big)^2,
\end{aligned}$$

dove $\theta_i^n \in [t_{i-1}^n, t_i^n]$, $i \in \{1, \ldots, n\}$. Poiché il processo $f'(W)$ è continuo e $(\mathcal{F}_t)$-adattato, abbiamo grazie al punto 7 del Teorema 3.1 (somme di Riemann-Itô) che

$$\sum_{i=1}^{n} f'\big(W(t_{i-1}^n)\big)\big(W(t_i^n) - W(t_{i-1}^n)\big) \overset{n\to\infty}{\longrightarrow} \int_0^t f'\big(W(s)\big)dW(s) \text{ in probabilità.}$$

Poniamo

$$\begin{aligned}
X_n &\doteq \sum_{i=1}^{n} f''\big(W(\theta_i^n)\big)\left(W(t_i^n) - W(t_{i-1}^n)\right)^2,\\
Y_n &\doteq \sum_{i=1}^{n} f''\big(W(t_{i-1}^n)\big)\left(W(t_i^n) - W(t_{i-1}^n)\right)^2,\\
Z_n &\doteq \sum_{i=1}^{n} f''\big(W(t_{i-1}^n)\big)\left(t_i^n - t_{i-1}^n\right).
\end{aligned}$$

Chiaramente,

$$Z_n \stackrel{n\to\infty}{\longrightarrow} \int_0^t f''\big(W(s)\big)ds \quad \mathbf{P}\text{-quasi certamente.}$$

Dall'altro canto, indicando con δ_n la lunghezza del passo della partizione (t_i^n),

$$\begin{aligned}
|X_n - Y_n| &\leq \sum_{i=1}^{n} \left|f''\big(W(\theta_i^n)\big) - f''\big(W(t_{i-1}^n)\big)\right| \cdot \left|W(t_i^n) - W(t_{i-1}^n)\right|^2\\
&\leq \max_{s,\tilde{s}\in[0,t]:|s-\tilde{s}|\leq\delta_n} \left|f''\big(W(s)\big) - f''\big(W(\tilde{s})\big)\right| \cdot \sum_{i=1}^{n} \left|W(t_i^n) - W(t_{i-1}^n)\right|^2\\
&\stackrel{n\to\infty}{\longrightarrow} 0 \text{ in probabilità,}
\end{aligned}$$

poiché $f''(W)$ è un processo continuo e, per il Teorema 1.2,

$$\sum_{i=1}^{n} \left|W(t_i^n) - W(t_{i-1}^n)\right|^2 \stackrel{n\to\infty}{\longrightarrow} t \text{ in } L^2(\mathbf{P}),$$

quindi in probabilità. Resta dunque da mostrare che $|Y_n - Z_n| \to 0$ in probabilità. Per $R > 0$ poniamo

$$\begin{aligned}
Y_n^{(R)} &\doteq \sum_{i=1}^{n} \mathbf{1}_{[-R,R]}\big(W(t_{i-1}^n)\big) \cdot f''\big(W(t_{i-1}^n)\big)\left(W(t_i^n) - W(t_{i-1}^n)\right)^2,\\
Z_n^{(R)} &\doteq \sum_{i=1}^{n} \mathbf{1}_{[-R,R]}\big(W(t_{i-1}^n)\big) \cdot f''\big(W(t_{i-1}^n)\big)\left(t_i^n - t_{i-1}^n\right).
\end{aligned}$$

Allora, per $R > 0$ fissato, grazie all'indipendenza degli incrementi del moto browniano ed il fatto che f'' sia limitata sui compatti,

$$\begin{aligned}
&\mathbf{E}\left[|Y_n^{(R)} - Y_n^{(R)}|^2\right]\\
&\leq \max_{x\in[-R,R]} |f''(x)| \cdot \mathbf{E}\left[\left|\sum_{i=1}^{n} \left|W(t_i^n) - W(t_{i-1}^n)\right|^2 - \left(t_i^n - t_{i-1}^n\right)\right|^2\right]\\
&= \max_{x\in[-R,R]} |f''(x)| \cdot \sum_{i=1}^{n} \mathbf{E}\left[\left(\left|W(t_i^n) - W(t_{i-1}^n)\right|^2 - \left(t_i^n - t_{i-1}^n\right)\right)^2\right]\\
&= \max_{x\in[-R,R]} |f''(x)| \cdot \sum_{i=1}^{n} 2\left(t_i^n - t_{i-1}^n\right)^2\\
&\stackrel{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Di conseguenza, $|Y_n^{(R)} - Z_n^{(R)}| \to 0$ in probabilità per ogni $R > 0$ fissato. Ora, grazie alla continuità di W, per ogni $\varepsilon > 0$ esiste $R = R(\varepsilon) > 0$ tale che

$$\mathbf{P}\left(\sup_{s\in[0,t]} |W(s)| > R(\varepsilon)\right) \leq \varepsilon.$$

Ne segue la convergenza di $(|Y_n - Z_n|)_{n\in\mathbb{N}}$ a zero in probabilità. □

La formula di Itô del Teorema 4.1 si esprime in notazione differenziale come

$$f\big(W(t)\big) = f'\big(W(t)\big)dW(t) + \frac{1}{2}f''\big(W(t)\big)dt.$$

Si noti che il differenziale dt corrisponde a $d\langle W\rangle(t)$, il differenziale della variazione quadratica del moto browniano.

La formula di Itô vale per processi scalari più generali di quelli del Teorema 4.1, cioè processi della forma $f(W)$ con $f \in \mathbf{C}^2(\mathbb{R})$, W un moto browniano uno-dimensionale.

Definizione 4.1. Un processo $(X(t))_{t\geq 0}$ a valori in $\mathbb{R}$ si dice *processo di Itô* rispetto a $(\mathcal{F}_t)$ e W se X è un processo continuo $(\mathcal{F}_t)$-adattato ed esistono processi φ, ψ a valori in $\mathbb{R}$ tali che $\varphi \in \Lambda^1((\mathcal{F}_t))$, $\psi \in \Lambda^2((\mathcal{F}_t))$ e $\mathbf{P}$-quasi certamente per ogni $t \geq 0$

$$X(t) = X(0) + \int_0^t \varphi(s)ds + \int_0^t \psi(s)dW(s).$$

Se X è un processo di Itô rispetto a $(\mathcal{F}_t)$ e W, allora il differenziale stocastico di X ammette la decomposizione $dX(t) = \varphi(t)dt+\psi(t)dW(t)$, dove i processi φ, ψ sono univocamente determinati a meno di un insieme di $\mathrm{Leb}_{[0,\infty)} \otimes \mathbf{P}$-misura zero.

Teorema 4.2 (Formula di Leibniz-Itô)**.** *Siano X, Y processi di Itô rispetto a $(\mathcal{F}_t)$ e W a valori in $\mathbb{R}$. Allora $\mathbf{P}$-quasi certamente per ogni $t \geq 0$*

$$X(t)\cdot Y(t) = X(0)\cdot Y(0) + \int_0^t X(s)dY(s) + \int_0^t Y(s)dX(s) + \int_0^t d\langle X, Y\rangle(s).$$

Dimostrazione. □

La formula di Leibniz-Itô del Teorema 4.2 si esprime in notazione differenziale come

$$d(X\cdot Y)(t) = X(t)dY(t) + Y(t)dX(t) + d\langle X, Y\rangle(t).$$

Si noti la comparsa del differenziale della covariazione quadratica tra X e Y, che nel calcolo classico sarebbe uguale a zero.

Teorema 4.3 (Formula di Itô II)**.** *Sia $f \in \mathbf{C}^{1,2}([0,\infty)\times\mathbb{R})$, e sia X un processo di Itô rispetto a $(\mathcal{F}_t)$ e W con decomposizione $dX(t) = \varphi(t)dt + \psi(t)dW(t)$. Allora $\mathbf{P}$-quasi certamente per ogni $t \geq 0$,*

$$\begin{aligned}
&f\big(t, X(t)\big) - f\big(0, X(0)\big) \\
&= \int_0^t \frac{\partial f}{\partial t}\big(s, X(s)\big)ds + \int_0^t \frac{\partial f}{\partial x}\big(s, X(s)\big)dX(s) + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}\big(s, X(s)\big)d\langle X\rangle(s) \\
&= \int_0^t \frac{\partial f}{\partial x}\big(s, X(s)\big)\psi(s)dW(s) \\
&\quad + \int_0^t \left(\frac{\partial f}{\partial t}\big(s, X(s)\big) + \frac{\partial f}{\partial x}\big(s, X(s)\big)\varphi(s) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\big(s, X(s)\big)|\psi(s)|^2\right) ds
\end{aligned}$$

Dimostrazione. □

La formula di Itô del Teorema 4.3 mostra in particolare che la classe dei processi di Itô è invariante rispetto alle trasformazioni $X \hookrightarrow f(.,X(.))$ per $f \in \mathbf{C}^{1,2}([0,\infty)\times\mathbb{R})$.

4.2 Moto browniano, funzioni armoniche e problema di Dirichlet

4.3 Rappresentazione di martingale

Qui: Rappresentazione di martingale della filtrazione naturale aumentato di un moto browniano come integrale stocastico.

Sia W un moto browniano d-dimensionale su uno spazio di probabilità completo $(\Omega, \mathcal{F}, \mathbf{P})$, e sia $(\mathcal{G}_t)$ la filtrazione naturale di W aumentata di tutti gli eventi $\mathbf{P}$-trascurabili, cioè

$$\mathcal{G}_t = \sigma(W(s) : s \leq t) \wedge \sigma(A \in \mathcal{F} : \mathbf{P}(A) = 0).$$

Primo passo: rappresentazione per variabili aleatorie $\xi \in L^2(\Omega, \mathcal{G}_T, \mathbf{P})$ della forma

$$\xi = c + \int_0^T u(t)\cdot dW(t) = c + \sum_{j=1}^d \int_0^T u_j(t)dW_j(t)$$

per un vettore costante $c \in \mathbb{R}^d$ e un processo Itô integrabile u (cioè $u \in \mathcal{H}^2$). Necessariamente si avrà $c = \mathbf{E}[\xi]$.

Teorema 4.4. *Sia $T > 0$, e sia $\xi \in L^2(\Omega, \mathcal{G}_T, \mathbf{P})$. Allora esiste un processo u in $\mathcal{H}^2_T = \mathcal{H}^2_T(\mathcal{G}_t)$ a valori in $\mathbb{R}^d$ tale che*

$$\xi = \mathbf{E}\left[\xi\right] + \int_0^T u(t)\cdot dW(t) \quad \mathbf{P}\text{-quasi certamente.}$$

Il processo u è unico $\lambda_{|[0,T]} \otimes \mathbf{P}$-quasi certamente.

Dimostrazione. □

Teorema 4.5. *Sia X una martingala rispetto a $(\mathcal{G}_t)$ di quadrato integrabile (cioè $X(t) \in L^2(\Omega, \mathcal{G}_t, \mathbf{P})$ per ogni $t \geq 0$). Allora esiste un processo u in $\mathcal{H}^2 = \mathcal{H}^2(\mathcal{G}_t)$ a valori in $\mathbb{R}^d$ tale che, per ogni $t \geq 0$,*

$$X(t) = \mathbf{E}[X(0)] + \int_0^t u(s) \cdot dW(s) \quad \mathbf{P}\text{-}\textit{quasi certamente.}$$

Dimostrazione. Sia $T > 0$. Per il Teorema 4.4, esiste $u \in \mathcal{H}^2_T \subset \mathcal{H}^2$ tale che

$$X(T) = \mathbf{E}[X(T)] + \int_0^T u(s) \cdot dW(s) \quad \mathbf{P}\text{-q.c.}$$

Per la proprietà di martingala di X abbiamo $\mathbf{E}[X(T)] = \mathbf{E}[X(0)]$ e per ogni $t \in [0, T]$, $\mathbf{P}$-quasi certamente,

$$\begin{aligned} X(t) = \mathbf{E}[X(T)|\mathcal{G}_t] &= \mathbf{E}[X(0)] + \mathbf{E}\left[\int_0^T u(s) \cdot dW(s)\big|\mathcal{G}_t\right] \\ &= \mathbf{E}[X(0)] + \int_0^t u(s) \cdot dW(s), \end{aligned}$$

dove l'ultima uguaglianza segue grazie alla proprietà di martingala dell'integrale stocastico per processi in $\mathcal{H}^2$.

Per $N \in \mathbb{N}$ scegliamo $u^N \in \mathcal{H}^2_N$ tale che

$$X(N) = \mathbf{E}[X(0)] + \int_0^N u^N(s) \cdot dW(s) \quad \mathbf{P}\text{-q.c.}$$

Poniamo $u(t, \omega) \doteq u^N(t, \omega)$ se $(t, \omega) \in [N-1, N) \times \Omega$. Per la parte di unicità del Teorema 4.4, u^N è unico $\lambda_{|[0,N]} \otimes \mathbf{P}$-quasi certamente. Si verifica allora per induzione che per ogni $N \in \mathbb{N}$ $u_{|[0,N]} = u^N$ $\lambda_{|[0,N]} \otimes \mathbf{P}$-quasi certamente. Ne segue che $u \in \mathcal{H}^2$ e, per ogni $t \geq 0$,

$$X(t) = \mathbf{E}[X(0)] + \int_0^t u(s) \cdot dW(s) \quad \mathbf{P}\text{-quasi certamente.}$$

□

Una rappresentazione analoga a quella del Teorema 4.5 vale anche per X una martingala locale rispetto alla filtrazione browniana $(\mathcal{G}_t)$. In questo caso il processo di rappresentazione u è solo in $\Lambda^2 = \Lambda^2(\mathcal{G}_t)$.

4.4 Teorema di Girsanov

Sia $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t))$ una base stocastica con W un moto browniano d-dimensionale rispetto a $(\mathcal{F}_t)$. Sia u un processo in Λ^2 (rispetto a $(\mathcal{F}_t)$) a valori in $\mathbb{R}^d$. Poniamo

$$Z(t) \doteq \exp\left(\int_0^t u(s) \cdot dW(s) - \frac{1}{2}\int_0^t |u(s)|^2 ds\right), \quad t \geq 0. \tag{4.1}$$

Sappiamo che Z è una martingala locale e anche supermartingala non-negativa rispetto a $(\mathcal{F}_t)$. Di conseguenza, Z è una martingala vera e propria se e solo se $\mathbf{E}[Z(t)] = 1$ per ogni $t \geq 0$.

Sia $T > 0$ fissato. Il processo Z è una martingala sull'intervallo $[0, T]$ se (e solo se) $\mathbf{E}[Z(T)] = 1$. Definiamo una misura positiva $\mathbf{Q}_T$ su $\mathcal{F}_T$ tramite

$$\mathbf{Q}_T(A) \doteq \mathbf{E}\left[\mathbf{1}_A \cdot Z(T)\right], \quad A \in \mathcal{F}_T. \tag{4.2}$$

La misura $\mathbf{Q}_T$ è una misura di probabilità se e solo se $(Z(t))_{t\in[0,T]}$ è una martingala. In questo caso $\mathbf{Q}_T$ è assolutamente continua rispetto a $\mathbf{P}_{|\mathcal{F}_T}$ (la restrizione di $\mathbf{P}$ su $\mathcal{F}_T$) con densità

$$\frac{d\mathbf{Q}_T}{d\,\mathbf{P}_{|\mathcal{F}_T}} = Z(T) \quad \mathbf{P}\text{-quasi certamente.}$$

Teorema 4.6 (Girsanov)**.** *Nella situazione di cui sopra, supponiamo che u sia tale che $\mathbf{E}\left[Z(T)\right] = 1$, dove Z è il processo definito da* (4.1)*. Sia $\mathbf{Q}_T$ la misura di probabilità definita da* (4.2)*. Poniamo*

$$\tilde{W}(t) \doteq W(t) - \int_0^t u(s)ds, \quad t \in [0, T].$$

Allora $\tilde{W}$ è un moto browniano rispetto a $\mathbf{Q}_T$, cioè $(\tilde{W}(t))_{t\in[0,T]}$ è un moto browniano d-dimensionale rispetto a $(\mathcal{F}_t)_{t\in[0,T]}$ definito su $(\Omega, \mathcal{F}_T, Q_T)$.

Dimostrazione. Poiché $\mathbf{Q}_T$ è assolutamente continua rispetto a $\mathbf{P}_{|\mathcal{F}_T}$, si ha con $\mathbf{Q}_T$-probabilità uno che le traiettorie di $\tilde{W}$ sono continue e $\tilde{W}(0) = 0$. Per $\lambda \in \mathbb{R}^d$ poniamo

$$Y_\lambda(t) \doteq \exp\left(\sqrt{-1}\langle\lambda, \tilde{W}(t)\rangle + \frac{1}{2}|\lambda|^2 t\right), \quad t \in [0, T],$$

e mostriamo che Y_λ è una martingala (a valori in $\mathbb{C}$) rispetto a $(\mathcal{F}_t)_{t\in[0,T]}$ e $\mathbf{Q}_T$. Questo implicherà che, per ogni $0 \leq s \leq t \leq T$, ogni $\lambda \in \mathbb{R}^d$,

$$\mathbf{E}_{\mathbf{Q}_T}\left[\exp\left(\sqrt{-1}\langle\lambda, \tilde{W}(t) - \tilde{W}(s)\rangle\right) \| \mathcal{F}_s\right] = \exp\left(-\frac{1}{2}|\lambda|^2 t\right) \quad \mathbf{Q}_T\text{-q.c.},$$

il che implicherà a sua volta che, rispetto a $\mathbf{Q}_T$, $\tilde{W}(t) - \tilde{W}(s)$ ha distribuzione $N(0, (t-s)\,\mathrm{Id}_d)$ ed è indipendente da $\mathcal{F}_s$.

Per mostrare che Y_λ è una martingala rispetto a $\mathbf{Q}_T$ è sufficiente mostrare che $Y_\lambda \cdot Z$ sia una martingala rispetto a $\mathbf{P}$.

□

Esempio 4.1 (Traiettorie browniane al di sotto di una funzione differenziabile). Sia W un moto browniano uno-dimensionale su $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t))$, sia $T > 0$, e sia $\varphi : [0, T] \to \mathbb{R}$ una funzione deterministica tale che $\varphi \in \mathbf{C}^1([0, T])$ e $b \doteq \varphi(0) > 0$. Domanda: $\mathbf{P}\left(W(t) < \varphi(t) \text{ per ogni } t \in [0, T]\right) > 0$?

Poniamo $\psi(t) \doteq \varphi(t) - b$ e $\tilde{W}(t) \doteq W(t) - \psi(t)$, $t \in [0,T]$. Allora

$$A \doteq \left\{W(t) < \varphi(t) \text{ per ogni } t \in [0,T]\right\} = \left\{\sup_{t\in[0,T]} \tilde{W}(t) < b\right\}.$$

Definiamo una misura Q_T su $\mathcal{F}_T$ mediante $Q_T(A) \doteq \mathbf{E}\left[\mathbf{1}_A \cdot Z(T)\right]$, $A \in \mathcal{F}_T$, ove

$$Z(t) \doteq \exp\left(\int_0^t \psi'(s)dW(s) - \frac{1}{2}\int_0^t |\psi'(s)|^2 ds\right), \quad t \in [0,T].$$

Poiché $\psi' = \varphi'$ è una funzione deterministica e limitata, abbiamo $\mathbf{E}\left[Z(T)\right] = 1$. Il Teorema 4.6 implica quindi che $\tilde{W}$ é un moto browniano fino al tempo T rispetto a Q_T. Ne segue (ricordando che $b > 0$)

$$Q_T(A) = 1 - Q_T\left(\sup_{t\in[0,T]} \tilde{W}(t) \geq b\right) = 1 - 2Q_T\left(\tilde{W}(T) \geq b\right) > 0.$$

Osserviamo che $Z(T) \in (0,\infty)$ certamente rispetto sia a $\mathbf{P}$ che a Q_T; abbiamo quindi che $\mathbf{P}_{|\mathcal{F}_T}$ è assolutamente continua rispetto a Q_T con densità $1/Z(T) > 0$. Il fatto che $Q_T(A) > 0$ implica dunque

$$\mathbf{P}(A) = \int_A \frac{1}{Z(T,\omega)} Q_T(d\omega) > 0.$$

Capitolo 5

Equazioni differenziali stocastiche

Un'equazione differenziale stocastica (EDS) serve a descrivere l'evoluzione nel tempo di una quantità aleatoria $X(t)$. Supponendo che $X(t)$ prenda valori in $\mathbb{R}^n$, ci occupiamo di equazioni della forma

$$dX(t) = b\big(t, X(t)\big)dt + \sigma\big(t, X(t)\big)dW(t), \tag{5.1}$$

dove W è un moto browniano d-dimensionale e b, σ sono funzioni deterministiche definite su $[0,\infty) \times \mathbb{R}^n$ ("prodotto tempo - spazio degli stati") a valori in $\mathbb{R}^n$ e $\mathbb{R}^{n\times d}$ rispettivamente. Studieremo in particolare il problema di Cauchy per l'EDS (5.1), cioè cercheremo una soluzione della (5.1) con condizione iniziale $X(0)$ prescritta.

I differenziali che appaiono in (5.1) sono differenziali stocastici; l'EDS sarà quindi interpretata come equazione integrale in cui l'integrale rispetto a W sarà l'integrale stocastico di Itô. Le funzioni b, σ in (5.1) si dicono rispettivamente *vettore di deriva (drift)* e *matrice di dispersione* (spesso σ viene chiamato matrice o coefficiente di diffusione, anche se quest'ultima denominazione si applicherebbe piuttosto a $\sigma\sigma^\mathsf{T}$ oppure a $\frac{1}{2}\sigma\sigma^\mathsf{T}$). Il vettore $b(t,x)$ dà il tasso di cambiamento infinitesimale della media di X in t se $X(t) = x$, mentre la matrice $\sigma\sigma^\mathsf{T}(t,x)$ dà il tasso di cambiamento infinitesimale della matrice di covarianza di X in t se $X(t) = x$.

5.1 Esistenza ed unicità

Siano $n, d \in \mathbb{N}$ fissati, e siano $b\colon [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$ funzioni misurabili.

Definizione 5.1. Una soluzione dell'equazione differenziale stocastica associata a (b, σ) è una quaterna $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t), W, X)$ tale che

(i) $(\Omega, \mathcal{F}, \mathbf{P})$ è uno spazio di probabilità completo;

(ii) $(\mathcal{F}_t)_{t\geq 0}$ è una filtrazione in $\mathcal{F}$ che sodisfa alle ipotesi standard (cioè continua a destra e tale che $\mathcal{F}_t$ contiene tutti gli eventi $\mathbf{P}$-trascurabili);

(iii) W è un moto browniano d-dimensionale rispetto a $(\mathcal{F}_t)$;

(iv) X è un processo a valori in $\mathbb{R}^n$ adattato a $(\mathcal{F}_t)$ e $\mathbf{P}$-quasi certamente continuo tale che $(b(t, X(t)))_{t\geq 0}$ è un processo in Λ^1, $(\sigma(t, X(t)))_{t\geq 0}$ è un processo in Λ^2 e $\mathbf{P}$-quasi certamente per ogni $t \geq 0$

$$X(t) = X(0) + \int_0^t b\big(s, X(s)\big)ds + \int_0^t \sigma\big(s, X(s)\big)dW(s).$$

Una soluzione $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t), W, X)$ si dice *soluzione forte* se il processo X è adattato alla filtrazione data da $\mathcal{G}_t \doteq \sigma(X(0)) \vee \mathcal{F}_t^W \vee \mathcal{N}_{\mathbf{P}}$, $t \geq 0$, ove $(\mathcal{F}_t^W)$ è la filtrazione naturale di W e $\mathcal{N}_{\mathbf{P}}$ il sistema degli insiemi $\mathbf{P}$-trascurabili.

Le due condizioni su $(\Omega, \mathcal{F}, \mathbf{P})$ e $(\mathcal{F}_t)$ nella Definizione 5.1 si possono riassumere dicendo che $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t))$ è una base stocastica standard.

Le soluzioni della Definizione 5.1 sono soluzioni globali; il moto browniano W e il processo di soluzione X sono quindi definiti per tutti i tempi non-negativi. A volte ci si deve accontentare di soluzioni definite solo fino a un tempo terminale finito, che può essere deterministico oppure aleatorio (tempo d'arresto) e nel quale la soluzione può anche essere non definita. In alcune applicazioni i coefficienti b, σ sono funzioni più generali; in particolare possono dipendere dalla traiettoria del processo di soluzione o del moto browniano fino al tempo corrente.

Ci sono due nozioni di unicità e due nozioni di esistenza per le equazioni differenziali stocastiche.

Definizione 5.2. Si dice che vale *unicità traiettoriale* per l'EDS associata a (b, σ) se quando $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t), W, X)$, $((\Omega, \mathcal{F}, \mathbf{P}), (\tilde{\mathcal{F}}_t), W, \tilde{X})$ sono due soluzioni dell'equazione tali che $\mathbf{P}(X(0) = \tilde{X}(0)) = 1$, allora

$$\mathbf{P}\left(X(t) = \tilde{X}(t) \text{ per ogni } t \geq 0\right) = 1.$$

Si dice che vale *unicità in legge* per l'EDS associata a (b, σ) se quando $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t), W, X)$, $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}}), (\tilde{\mathcal{F}}_t), \tilde{W}, \tilde{X})$ sono due soluzioni dell'equazione tali che $\mathbf{P} \circ (X(0))^{-1} = \tilde{\mathbf{P}} \circ (\tilde{X}(0))^{-1}$, allora $\mathbf{P} \circ X^{-1} = \tilde{\mathbf{P}} \circ \tilde{X}^{-1}$.

Per stabilire unicità traiettoriale si confrontano soluzioni con lo stesso spazio di probabilità, lo stesso processo di Wiener e condizioni iniziali che coincidono con probabilità uno. Per stabilire unicità in legge, invece, si confrontano soluzioni con condizioni iniziali identicamente distribuite, ma processi che possono essere definiti su basi stocastiche diverse. In particolare, il moto browniano che muove la dinamica non è necessariamente lo stesso.

L'unicità traiettoriale è una proprietà più forte di quella in legge nel senso che l'unicità traiettoriale implica l'unicità legge. Per una dimostrazione di questo risultato dovuto a Yamada e Watanabe si veda, ad esempio, Sezione 5.3.D in Karatzas and Shreve [1991, pp. 308-311].

Definizione 5.3. Si dice che vale *esistenza forte* per l'EDS associata a (b, σ) se quando $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t))$ è una base stocastica standard con W un moto browniano d-dimensionale rispetto a $(\mathcal{F}_t)$ e ξ una variabile aleatoria $\mathcal{F}_0$-misurabile a valori in $\mathbb{R}^n$, allora esiste un processo X a valori in $\mathbb{R}^n$ tale che $X(0) = \xi$ $\mathbf{P}$-quasi certamente e $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t), W, X)$ è una soluzione forte dell'equazione.

Si dice che vale *esistenza debole* per l'EDS associata a (b, σ) se quando μ è una misura di probabilità su $\mathcal{B}(\mathbb{R}^n)$, allora esiste una soluzione $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t), W, X)$ dell'equazione tale che $\mathbf{P} \circ (X(0))^{-1} = \mu$.

Le nozioni di unicità ed esistenza possono essere ristrette prescrivendo la distribuzione iniziale oppure una classe di distribuzioni iniziali. Ad esempio, si può richiedere unicità traiettoriale ed esistenza debole solo per condizioni iniziali deterministiche.

Se i coefficienti b, σ sono funzioni lipschitziane nella variabile di spazio a crescita sublineare uniformemente nel tempo, si hanno esistenza di soluzioni forti e unicità traiettoriale per l'EDS associata a (b, σ). Per il risultato preciso introduciamo le seguenti condizioni.

(Lip) Esiste una costante $L > 0$ tale che per ogni $t \geq 0$, ogni $x, y \in \mathbb{R}^n$

$$|b(t,x) - b(t,y)| \leq L|x-y|, \qquad |\sigma(t,x) - \sigma(t,y)| \leq L|x-y|.$$

(Cre) Esiste una costante $K > 0$ tale che per ogni $t \geq 0$, ogni $x \in \mathbb{R}^n$

$$|b(t,x)|^2 \leq K\left(1+|x|\right)^2, \qquad |\sigma(t,x)|^2 \leq K\left(1+|x|\right)^2.$$

Se le funzioni b, σ sono globalmente lipschitziane e non dipendono dalla variabile temporale, allora soddisfanno non solo alla condizione (Lip), ma in automatico anche alla condizione (Cre).

Teorema 5.1. *Supponiamo che b, σ soddisfino a (Lip) e (Cre). Allora valgono unicità traiettoriale ed esistenza forte per l'EDS associata a (b, σ).*

Dimostrazione. □

5.2 Esempi di equazioni differenziali stocastiche

Esempio 5.1 (Equazioni lineari con rumore additivo). Siano $A\colon [0,\infty) \to \mathbb{R}^{n\times n}$, $\tilde{b}\colon [0,\infty) \to \mathbb{R}^n$, $\tilde{\sigma}\colon [0,\infty) \to \mathbb{R}^{n\times d}$ funzioni deterministiche misurabili e localmente limitate. Allora valgono unicità traiettoriale ed esistenza forte di soluzioni per l'equazione lineare con rumore additivo

$$dX(t) = \left(A(t)X(t) + \tilde{b}(t)\right) + \tilde{\sigma}(t)dW(t), \tag{5.2}$$

cioè per l'EDS associata ai coefficienti b, σ dati da

$$b(t,x) \doteq A(t)x + \tilde{b}(t), \qquad \sigma(t,x) \doteq \tilde{\sigma}(t), \qquad (t,x) \in [0,\infty) \times \mathbb{R}^n.$$

La soluzione si può rappresentare in termini della soluzione fondamentale dell'equazione differenziale lineare ordinaria associata a $A(.)$ usando il metodo di variazione delle costanti; cf. Esercizio V.6.

Esempio 5.2 (Processo di Ornstein-Uhlenbeck). Siano $\alpha > 0$, $\tilde{\sigma} > 0$ costanti positive. Consideriamo l'EDS lineare scalare

$$dX(t) = -\alpha X(t)dt + \tilde{\sigma} dW(t), \tag{5.3}$$

che è un caso particolare dell'EDS (5.2) con $n = d = 1$, $A(.) \equiv -\alpha$, $\tilde{b}(.) \equiv 0$, $\tilde{\sigma}(.) \equiv \tilde{\sigma}$. Sia W è un moto browniano uno-dimensionale definito su una base stocastica standard $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t))$, e sia ξ una variabile aleatoria reale $\mathcal{F}_0$-misurabile. Il processo di soluzione X dell'EDS (5.3) con condizione iniziale ξ si può allora scrivere come (cf. Esercizio V.6)

$$X(t) = \xi e^{-\alpha t} + \tilde{\sigma} \int_0^t e^{-\alpha(t-s)} dW(s), \quad t \geq 0.$$

Se $\xi \in L^2(\mathbf{P})$, allora le funzioni di media e di covarianza di X sono date da

$$\begin{aligned} \mathbf{E}\left[X(t)\right] &= \mathbf{E}\left[\xi\right] e^{-\alpha t}, \\ \operatorname{cov}\left(X(s), X(t)\right) &= \left(\operatorname{var}(\xi) + \frac{\tilde{\sigma}^2}{2\alpha}(e^{2\alpha(t\wedge s)} - 1)\right) e^{-\alpha(t+s)}. \end{aligned}$$

Se ξ ha distribuzione gaussiana, allora X è un processo gaussiano. Se la distribuzione di ξ è gaussiana di media zero e varianza $\frac{\tilde{\sigma}^2}{2\alpha}$, allora X diventa un processo gaussiano stazionario con funzione di covarianza

$$\operatorname{cov}\left(X(s), X(t)\right) = \frac{\tilde{\sigma}^2}{2\alpha} e^{2\alpha|t-s|}.$$

Esempio 5.3 (Moto browniano geometrico). Siano $\mu \in \mathbb{R}$, $\tilde{\sigma} \geq 0$. Consideriamo l'equazione scalare

$$dX(t) = \mu X(t)dt + \tilde{\sigma} X(t) dW(t), \tag{5.4}$$

cioè l'EDS associata ai coefficienti b, σ dati da

$$b(t,x) \doteq \mu \cdot x, \qquad \sigma(t,x) \doteq \tilde{\sigma} \cdot x, \qquad (t,x) \in [0,\infty) \times \mathbb{R}.$$

La (5.4) è lineare sia nel coefficiente di deriva b che in quello di dispersione σ, e b, σ soddisfanno alle ipotesi del Teorema 5.1. Data una base stocastica standard $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t))$ con W un moto browniano uno-dimensionale rispetto a $(\mathcal{F}_t)$) e una variabile aleatoria reale ξ $\mathcal{F}_0$-misurabile, il processo di soluzione forte della (5.4) si scrive esplicitamente come

$$X(t) = \xi \exp\left(\tilde{\sigma} W(t) + \left(\mu - \frac{\tilde{\sigma}^2}{2}\right) t\right), \quad t \geq 0.$$

Supponiamo $\xi > 0$ **P**-quasi certamente. Allora $X(t) > 0$ per ogni $t \geq 0$ con probabilità uno. A seconda che $\mu - \frac{\tilde{\sigma}^2}{2}$ sia positivo, negativo o uguale a zero, cambia il comportamento asintotico del processo di soluzione X:

$$\begin{aligned}
&\mu - \frac{\tilde{\sigma}^2}{2} > 0, \quad \text{allora } \mathbf{P}\text{-q.c.} \quad \lim_{t\to\infty} X(t) = \infty, \quad \inf_{t\geq 0} X(t) > 0;\\
&\mu - \frac{\tilde{\sigma}^2}{2} < 0, \quad \text{allora } \mathbf{P}\text{-q.c.} \quad \sup_{t\geq 0} X(t) < \infty, \quad \lim_{t\to\infty} X(t) = 0;\\
&\mu - \frac{\tilde{\sigma}^2}{2} = 0, \quad \text{allora } \mathbf{P}\text{-q.c.} \quad \limsup_{t\to\infty} X(t) = \infty, \quad \liminf_{t\to\infty} X(t) = 0.
\end{aligned}$$

Esempio 5.4 (Esplosione). Consideriamo l'equazione stocastica scalare

$$X(t) = x + \int_0^t (X(s))^3 ds + \int_0^t (X(s))^2 dW(s), \tag{5.5}$$

dove $x \in \mathbb{R}$ (condizione iniziale deterministica) e W è un moto browniano unodimensionale definito su una base stocastica standard $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t))$. Poniamo $\tau_x \doteq \inf\{t \geq 0 : W(t) = 1/x\}$ se $x \neq 0$, $\tau_0 \doteq \infty$ se $x = 0$. Il processo X definito sull'intervallo (aleatorio) $[0, \tau_x)$ mediante

$$X(t,\omega) \doteq \frac{x}{1 - x \cdot W(t,\omega)}, \quad t \in [0, \tau_x(\omega)),$$

soddisfa l'equazione integrale (5.5) con $X(0) = x$ **P**-quasi certamente. Se $x \neq 0$, allora per **P**-quasi ogni $\omega \in \Omega$

$$\tau_x(\omega) < \infty, \qquad \limsup_{t \nearrow \tau_x(\omega)} |X(t,\omega)| = \infty,$$

cioè X esplode in tempo finito con probabilità uno.

5.3 Proprietà di Markov e diffusioni

5.4 Formula di Feynman-Kac

Siano $b : \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^d$ funzioni globalmente lipschitziane. Sia $(\Omega, \mathcal{F}, \mathbf{P})$ uno spazio di probabilità completo con una filtrazione $(\mathcal{F}_t)$ in $\mathcal{F}$ che soddisfi alle ipotesi standard e con W un moto browniano d-dimensionale rispetto a $(\mathcal{F}_t)$. Per $(t,x) \in [0,\infty) \times \mathbb{R}^n$, sia $X^{(t,x)}$ la soluzione dell'equazione

$$X^{(t,x)}(s) = \begin{cases} x & \text{se } s \in [0,t], \\ x + \int_t^s b(X(r))dr + \int_t^s \sigma(X(r))dW(r) & \text{se } s > t. \end{cases} \tag{5.6}$$

Per il Teorema 5.1, i processi $X^{(t,x)}$ sono ben definiti.

Teorema 5.2.

$$v(t,x) = \mathbf{E}\left[\int_t^T g\big(s, X^{(t,x)}(s)\big)\exp\left(-\int_t^s k\big(r, X^{(t,x)}(r)\big)dr\right)ds \right.$$
$$\left. + F\big(X^{(t,x)}(T)\big)\exp\left(-\int_t^T k\big(r, X^{(t,x)}(r)\big)dr\right)\right]$$

Dimostrazione. Sia $(t,x) \in [0,T] \times \mathbb{R}^n$ fissato Poniamo

$$Y(s) \doteq$$

$\square$

Letteratura

Le fonti principali per il corso sono il Baldi [2000] e le dispense di Caravenna [2011]. Un testo di riferimento, piuttosto tecnico ma molto utile, è Karatzas and Shreve [1991]. Potrà essere utile, soprattutto per richiami di calcolo delle probabilità, il Klenke [2008]. Qualche esempio e risultato sarà tratto da Øksendal [1998] o Rogers and Williams [2000a,b].

Bibliografia

P. Baldi. *Equazioni differenziali stocastiche e applicazioni.* Pitagora Editrice, Bologna, 2000.

F. Caravenna. *Moto browniano e analisi stocastica.* 2011. URL `http://www.matapp.unimib.it/~fcaraven/`. Dispense rielaborate, versione 3.3.

O. Kallenberg. *Foundations of Modern Probability.* Probability and Its Applications. Springer, New York, 2nd edition, 2001.

G. Kallianpur. *Stochastic Filtering Theory*, volume 13 of *Applications of Mathematics*. Springer, New York, 1980.

I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer, New York, 2nd edition, 1991.

A. Klenke. *Probability Theory. A Comprehensive Course.* Springer, London, 2008.

B. Øksendal. *Stochastic Differential Equations. An Introduction with Applications*. Springer, Berlin, 5th edition, 1998.

L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales. Volume 1: Foundations.* Cambridge University Press, Cambridge, 2000a.

L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales. Volume 2: Itô Calculus.* Cambridge University Press, Cambridge, 2000b.