

Dipartimento di Matematica

UNIVERSITÀ DI ROMA “LA SAPIENZA”

NOTE DEL CORSO

Calcolo stocastico e applicazioni

Lorenzo Bertini

Dipartimento di Matematica, Università di Roma La Sapienza

Piazzale Aldo Moro 2, 00185 Roma, Italy

E-mail: bertini@mat.uniroma1.it

Indice

1. Il moto browniano	4
1.1. <i>C</i> come spazio misurabile	4
1.2. Definizione del moto browniano	5
1.3. Costruzione del moto browniano	6
1.4. Misura di Wiener su $C(\mathbb{R}_+)$ e invarianza per cambi di scala	9
1.5. Principio di riflessione	11
1.6. (Ir)regolarità delle traiettorie browniane	12
1.7. Ulteriori proprietà delle traiettorie browniane (senza dimostrazioni)	14
1.8. Definizione di martingala e diseguaglianza di Doob	16
1.9. Definizione astratta del moto browniano e teorema di Levy	19
1.A. Cenni sulla convergenza debole di misure di probabilità	21
1.B. Diseguaglianza di Garsia, Rodemich e Rumsey	23
1.C. Cenni sulle probabilità condizionate	26
2. Equazioni differenziali stocastiche	28
2.1. Integrale di Ito	28
2.2. Formula di Ito	33
2.3. Equazioni differenziali stocastiche: esistenza e unicità per coefficienti Lipschitz	37
2.4. Formulazione di equazioni stocastiche come problema alle martingale (cenni)	41
2.5. Processo di Ornstein–Uhlenbeck	42
2.6. Approssimazione di Smoluchowski	44
2.7. Limite di McKean–Vlasov	46
2.8. Tempi locali di un browniano e rappresentazione del browniano riflesso	50
2.9. Approssimazione del browniano riflesso	54
3. Processi di Markov	58
3.1. Catene di Markov	58
3.2. Processi di Markov	67
3.3. Semigrupperi di Feller	69
3.4. Processi di Feller e proprietà di Markov forte	72
3.5. Generatori di semigrupperi di Feller	77
4. Processi di diffusione e PDE	81
4.1. Problema di Dirichlet e moto browniano	81
4.2. Formula di Fenyman–Kac	84
4.3. Legge dell'arcoseno	87
5. Misure invarianti per processi di diffusione	89
5.1. Esistenza di misure invarianti	89
5.2. Misure reversibili	91
Bibliografia	92

1. Il moto browniano

Motivazione fisica. Ci sono particelle “grandi” (qualche Angstrom) in soluzione. Nella soluzione ci sono particelle “piccole” (di massa trascurabile rispetto a quelle grandi) in agitazione termica che urtano quelle grandi. Ne risulta un moto erratico ed apparentemente imprevedibile delle particelle grandi che si osserva al microscopio (Brown era un biologo dell’ottocento). Inoltre questo moto non pare avere una velocità (le traiettorie non sono differenziabili). Vogliamo fare un modello statistico, non pretendiamo di descrivere un singolo moto, ma solo proprietà medie per il moto delle particelle grandi (moto browniano).

Motivazione matematica. Vogliamo mettere una misura su $C := C([0, 1])$, funzioni continue su $[0, 1]$ a valori in \mathbb{R} (misura di Wiener). Poi potremmo costruirne altre come perturbazione.

Cominciamo con il mostrare che in uno spazio euclideo infinito dimensionale non esiste un analogo della misura di Lebesgue. Sia H uno spazio di Hilbert infinito dimensionale con base ortonormale e_k , $k = 1, 2, \dots$. Supponiamo esista una misura λ su H invariante per traslazioni e che assegni massa finita alle palle e deduciamo un assurdo. Sia $B_r(x) := \{y \in H : \|x - y\| < r\}$ la palla di raggio r centrata in x . Abbiamo $B_2(0) \supset \bigcup_k B_{1/2}(e_k)$ con $B_{1/2}(e_k) \cap B_{1/2}(e_h) = \emptyset$ per $h \neq k$. Pertanto se λ è una misura σ -additiva dovrà risultare

$$\lambda(B_2(0)) \geq \lambda\left(\bigcup_k B_{1/2}(e_k)\right) = \sum_k \lambda(B_{1/2}(e_k))$$

se λ è invariante per traslazioni abbiamo $\lambda(B_{1/2}(e_k)) = b$ indipendentemente da k e troviamo un assurdo.

Levata la misura di Lebesgue, la misura più “semplice” che possiamo immaginare è la misura di Gauss (su \mathbb{R} definita da $\mu(dx) = (2\pi)^{-1/2} \exp\{-x^2/2\}dx$ con $\mu(\mathbb{R}) = 1$). Tale misura ha buone proprietà di tensorizzazione e passa a dimensione infinita. La misura che costruiremo su C sarà gaussiana.

1.1. C come spazio misurabile

Consideriamo C come spazio metrico rispetto alla distanza naturale della convergenza uniforme, ovvero $d(x, y) := \sup_{t \in [0, 1]} |x(t) - y(t)|$.

Dati $0 \leq t_1 < t_2 < \dots < t_n \leq 1$ e I_1, \dots, I_n intervalli aperti di \mathbb{R} definiamo il cilindro

$$\mathcal{C}(\underline{t}, \underline{I}) := \{x \in C : x(t_1) \in I_1, \dots, x(t_n) \in I_n\}$$

Volendo far diventare C uno spazio misurabile abbiamo due candidati naturali come σ -algebra.

- (i) La σ -algebra dei Boreliani, ovvero la più piccola σ -algebra che contiene tutti gli aperti di C , $\mathcal{B}_1 := \mathcal{B}(C) := \sigma\{A, A \text{ aperto in } C\}$

(ii) La σ -algebra generata dai cilindri, $\mathcal{B}_2 = \sigma\{\mathcal{C}(t, I)\}$

Il seguente risultato ci leva dall'ambiguità.

t:1=2 **Lemma 1.1.** *Risulta $\mathcal{B}_1 = \mathcal{B}_2$*

Dimostrazione. Osserviamo che $\mathcal{C}(t, I) = \{x \in C : x(t) \in I\}$ è aperto in C , quindi $\mathcal{C}(t, I) = \bigcap_{k=1}^n \mathcal{C}(t_k, I_k) \in \mathcal{B}_1$. Pertanto $\mathcal{B}_2 \subset \mathcal{B}_1$.

Siano ora \mathbb{Q} i numeri razionali in $[0, 1]$. Dato $x \in C$, abbiamo

$$\begin{aligned} \overline{B}_r(x) &:= \left\{ y \in C : \sup_{t \in [0,1]} |y(t) - x(t)| \leq r \right\} = \left\{ y \in C : \sup_{t \in \mathbb{Q}} |y(t) - x(t)| \leq r \right\} \\ &= \bigcap_{t \in \mathbb{Q}} \mathcal{C}(t, [x(t) - r, x(t) + r]) = \bigcap_{t \in \mathbb{Q}} \bigcap_{j=1}^{\infty} \mathcal{C}(t, (x(t) - r - 1/j, x(t) + r + 1/j)) \end{aligned}$$

Ciò mostra che $\overline{B}_r(x) \in \mathcal{B}_2$ e quindi anche $B_r(x) \in \mathcal{B}_2$. Per la separabilità di C questo implica che se A è un aperto di C allora $A \in \mathcal{B}_2$ e quindi $\mathcal{B}_1 \subset \mathcal{B}_2$. \square

s:dbm

1.2. Definizione del moto browniano

Terminologia. Dato uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ (ovvero uno spazio di misura con massa totale 1) diciamo che le variabili aleatorie (ovvero funzioni \mathcal{F} -misurabili su Ω a valori in \mathbb{R}) X_1, \dots, X_n sono *indipendenti* sse comunque prese f_1, \dots, f_n limitate e misurabili da \mathbb{R} a \mathbb{R} si ha

$$\mathbb{E}(f_1(X_1) \cdots f_n(X_n)) := \int d\mathbb{P}(\omega) f_1(X_1(\omega)) \cdots f_n(X_n(\omega)) = \mathbb{E}(f_1(X_1)) \cdots \mathbb{E}(f_n(X_n))$$

Data una variabile aleatoria X chiamiamo distribuzione (o legge) di X la misura di probabilità P_X su \mathbb{R} definita da $P_X(B) = \mathbb{P}(\{\omega : X(\omega) \in B\})$ per ogni $B \in \mathcal{B}(\mathbb{R})$.

Per $t > 0$ e $x \in \mathbb{R}$, indichiamo con

$$p_t(x) := \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{x^2}{2t}\right\}$$

il nucleo gaussiano. Osserviamo che $\int dx p_t(x) = 1$, $\int dx p_t(x) x = 0$, $\int dx p_t(x) x^2 = t$.

t:dbm1

Definizione 1.2. *La misura di Wiener P è la misura di probabilità su $(C, \mathcal{B}(C))$ che gode delle seguenti proprietà:*

(i) *per ogni $t \in [0, 1]$ la variabile aleatoria $x(t)$ ha distribuzione gaussiana di media 0 e varianza t , ovvero*

$$P(x(t) \in I) = \int_I dz p_t(z)$$

(ii) *$\{x(t), t \in [0, 1]\}$ è a incrementi indipendenti, ovvero comunque dati $0 \leq t_1 < t_2 < \dots < t_n \leq 1$ le variabili aleatorie $x(t_2) - x(t_1), x(t_3) - x(t_2), \dots, x(t_n) - x(t_{n-1})$ sono indipendenti.*

Nota. La condizione (i) per $t = 0$ significa che $P(x(0) = 0) = 1$ e questa è solo una scelta conveniente dell'origine. Verificare che da (i) e (ii) segue che $x(t) - x(s)$, con $t > s$ è gaussiana di media nulla e varianza $t - s$ (sugg. scrivere, con $0 < s < t$, $x(t) = x(s) + [x(t) - x(s)]$ ed utilizzare le funzioni caratteristiche delle misure).

Equivalentemente (dimostrare l'equivalenza!) la misura di Wiener è la misura P che assegna ad ogni cilindro $\mathcal{C}(\underline{t}, \underline{I})$ misura

$$P(\mathcal{C}(\underline{t}, \underline{I})) = p_n(\underline{t}; \underline{I}) := \int_{I_1} dz_1 p_{t_1}(z_1) \int_{I_2} dz_2 p_{t_2-t_1}(z_2 - z_1) \cdots \int_{I_n} dz_n p_{t_n-t_{n-1}}(z_n - z_{n-1}) \quad (1.1)$$

df dmb

L'esistenza di una misura P non è banale (l'unicità lo è?); un modo possibile per costruirla è il seguente: 1) si verifica (facilmente) che p_n verifica le condizioni di consistenza

$$p_n(t_1, \dots, t_n; I_1, \dots, I_n) = p_{n+1}(t_1, \dots, t_{k-1}, \tau, t_k, \dots, t_n; I_1, \dots, I_{k-1}, \mathbb{R}, I_k, \dots, I_n)$$

ove $\tau \in [t_{k-1}, t_k]$; 2) si evoca un teorema generale della teoria della misura (Teorema di estensione di Kolmogorov) per garantirsi esistenza di una misura \hat{P} sullo spazio enorme (e alquanto scomodo) $\mathbb{R}^{[0,1]} := \prod_{t \in [0,1]} \mathbb{R}$ con la σ -algebra prodotto; 3) Con un po' di lavoro si mostra che \hat{P} ha in realtà supporto su C .

Il punto 3 è più sottile di quello che sembra, in effetti come sopra detto è sbagliato. Il problema è che C si guarda bene dall'essere un sottoinsieme misurabile di $\mathbb{R}^{[0,1]}$. Poco male, si potrebbe pensare che completando la σ -algebra prodotto rispetto a \hat{P} (aggiungendo cioè tutti gli insiemi di \hat{P} misura nulla; stessa operazione tra i Boreliani di \mathbb{R} e gli insiemi Lebesgue misurabili) C divenga misurabile. Non è così; la \hat{P} misura interna di C è nulla (gli insiemi misurabili nel completamento sono quelli che hanno misura interna ed esterna uguale). Le precedenti affermazioni si giustificano come segue. La σ -algebra prodotto in $\mathbb{R}^{[0,1]}$ è per definizione la più piccola σ -algebra che contiene i cilindri, ora non è difficile verificare (esercizio) che in realtà tale σ -algebra è la collezione dei σ -cilindri, ove $F \subset \mathbb{R}^{[0,1]}$ è un σ -cilindro sse $F = A \times \mathbb{R}^{[0,1] \setminus T}$ per un qualche $T \subset [0, 1]$ al più numerabile e $A \in \mathcal{B}(\mathbb{R}^T)$. Sugg. mostrare che la collezione dei σ -cilindri è una σ -algebra verificando che è chiusa rispetto ad intersezioni numerabili. Gli insiemi misurabili sono pertanto quelli che dipendono dai valori di $\{x(t)\}_{t \in [0,1]}$ solo per una quantità numerabile di t . Ovviamente in tal modo non si può riconoscere la continuità.

Per superare tale difficoltà e costruire la misura di Wiener, si procede circa al modo seguente. Si considera $x(t)$ solo per t razionali e si definisce $y(t) = \lim_{s \rightarrow t, s \in \mathbb{Q}} x(s)$ dimostrando che tale limite esiste \hat{P} q.s. A questo punto si "scopre" che $t \mapsto y(t)$ è \hat{P} q.s. continua e che per ogni $t \in [0, 1]$ si ha $\hat{P}(x(t) = y(t)) = 1$; in particolare x e y hanno le stesse distribuzioni finito dimensionali. La misura di Wiener (su C , non su $\mathbb{R}^{[0,1]}$) è infine la legge di y . Per i dettagli di questa costruzione si può vedere ad esempio [9, §2.2].

Nei paragrafi successivi seguiamo una strada forse più lunga (ma che ci darà una qualche intuizione sul moto browniano) che permette di costruire la misura P direttamente su C .

s:cbm

1.3. Costruzione del moto browniano

Ritorniamo al problema fisico e pensiamo di modellare la situazione nel modo seguente.

Per semplicità pensiamo al caso di una dimensione spaziale o guardiamo solo una componente del moto. Ogni unità di tempo una pallina piccola urta quella grande, questa si muove di una unità di spazio a sinistra o a destra a seconda del risultato di una moneta non truccata. Dopo n unità di tempo la particella grande si troverà nella posizione data da $\sum_{i=1}^n \xi_i$ dove ξ_i sono variabili indipendenti tali che $\mathbb{P}(\xi_i = \pm 1) = 1/2$.

Guardiamo ora a questo processo su scala macroscopica. Pensiamo che ogni secondo (unità di tempo macroscopica) avvengono in realtà ε^{-1} urti e lo spostamento della particella a causa di un singolo urto sia in realtà pari a $\sqrt{\varepsilon}$ metri (unità di spazio macroscopica). Dopo t secondi vedremo allora la particella nella posizione $X_\varepsilon(t) = \sqrt{\varepsilon} \sum_{i=1}^{\varepsilon^{-1}t} \xi_i$. Costruiremo il moto browniano come limite di X_ε per $\varepsilon \rightarrow 0$.

In modo più formale il modello è definito come segue. Sia μ una misura di probabilità su \mathbb{R} (in cui assegnamo la σ -algebra dei Boreliani $\mathcal{B}(\mathbb{R})$) tale che $\int \mu(dx) x^4 < \infty$ (questa condizione si può indebolire) con le condizioni di normalizzazione $\int \mu(dx) x = 0$ e $\int \mu(dx) x^2 = 1$. Costruiamo lo spazio prodotto $(\Omega, \mathcal{F}, \mathbb{P}) = \bigotimes_{i=1}^{\infty} (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$. Chiamiamo ξ gli elementi di Ω e ξ_i la coordinata i -sima. Per $\varepsilon > 0$ definiamo $\varphi_\varepsilon : \Omega \rightarrow C$ come

$$\varphi_\varepsilon(\xi)(t) := \sqrt{\varepsilon} \sum_{i=1}^{[\varepsilon^{-1}t]} \xi_i + \sqrt{\varepsilon} (\varepsilon^{-1}t - [\varepsilon^{-1}t]) \xi_{[\varepsilon^{-1}t]+1} \quad (1.2) \quad \boxed{\text{feps}}$$

dichiariamo $P_\varepsilon := \mathbb{P} \circ \varphi_\varepsilon^{-1}$, ovvero $P_\varepsilon(A) := \mathbb{P}(\varphi_\varepsilon^{-1}(A))$ per ogni $A \in \mathcal{B}(C)$. Verificare che φ_ε è misurabile (sugg. dimostrare che φ_ε è continua rispetto alla topologia prodotto su Ω e verificare che la σ -algebra generata dei cilindri in Ω coincide con gli insiemi boreliani rispetto alla topologia prodotto). Vedi paragrafo 1.A per cenni sulla convergenza debole di misure di probabilità.

t:ip **Teorema 1.3.** *La successione P_ε è debolmente convergente per $\varepsilon \rightarrow 0$, ovvero esiste un'unica $P \in \mathcal{P}(C)$ per cui $P_\varepsilon \rightharpoonup P$.*

Cominciamo con il mostrare che $P_\varepsilon(\mathcal{C}(\underline{t}, \underline{I})) \rightarrow p_n(\underline{t}; \underline{I})$ e partiamo dal caso $n = 1$. In questo caso questa proprietà è nota come teorema limite centrale

t:clt **Teorema 1.4.** *Sia $\psi_N : \Omega \rightarrow \mathbb{R}$ definita da*

$$\psi_N(\xi) := \frac{1}{\sqrt{N}} \sum_{i=1}^{[Nt]} \xi_i$$

Allora $\mu_N := \mathbb{P} \circ \psi_N^{-1} \Rightarrow p_t(x) dx$ per $N \rightarrow \infty$.

Dimostrazione. Sia $\phi(\lambda) := \int \mu(dx) e^{i\lambda x}$ la funzione caratteristica di μ . Per il teorema di Taylor si ha (dimostrare che $\int \mu(dx) x^2 < \infty$ implica $\phi(\lambda) \in C^2$)

$$\begin{aligned} \phi(\lambda) &= \phi(0) + \phi'(0)\lambda + \frac{1}{2}\phi''(0)\lambda^2 + o(\lambda^2) \\ &= 1 + i\lambda \int \mu(dx) x - \frac{1}{2}\lambda^2 \int \mu(dx) x^2 + o(\lambda^2) = 1 - \frac{1}{2}\lambda^2 + o(\lambda^2) \end{aligned}$$

pertanto, per ogni $\lambda \in \mathbb{R}$ si ha, per l'indipendenza delle ξ_j ,

$$\begin{aligned}\phi_N(\lambda) &= \int \mu_N(dx) e^{i\lambda x} = \int d\mathbb{P}(\xi) \exp \left\{ i\lambda \frac{1}{\sqrt{N}} \sum_{j=1}^{[Nt]} \xi_j \right\} \\ &= \phi \left(\frac{\lambda}{\sqrt{N}} \right)^{[Nt]} = \left[1 - \frac{1}{2} \frac{\lambda^2}{N} + o \left(\frac{1}{N} \right) \right]^{[Nt]} \longrightarrow \exp \left\{ -\frac{1}{2} \lambda^2 t \right\}\end{aligned}$$

per $N \rightarrow \infty$. Osservando che per misure su \mathbb{R} la convergenza debole equivale alla convergenza puntuale della funzione caratteristica (vedi teorema 1.24) e che $\exp\{-\frac{1}{2}\lambda^2 t\}$ è la funzione caratteristica di $p_t(x)dx$, il teorema è dimostrato. \square

Possiamo ora facilmente dimostrare la *convergenza delle distribuzioni finito dimensionali*.

t:cdfd **Lemma 1.5.**

$$\lim_{\varepsilon \rightarrow 0} P_\varepsilon(\mathcal{C}(\underline{t}, \underline{I})) = p_n(\underline{t}; \underline{I})$$

Dimostrazione. Poichè $\sqrt{\varepsilon}(\varepsilon^{-1}t - [\varepsilon^{-1}t]) \leq \sqrt{\varepsilon}$ abbiamo che l'ultimo termine nel lato destro della (1.2) converge a 0 (addirittura in $L_2(d\mathbb{P})$). Possiamo quindi dimenticarlo. Consideriamo, per semplicità il caso $n = 2$ con $t_1 < t_2$. Poichè \mathbb{P} è una misura prodotto, abbiamo

$$\begin{aligned}& \mathbb{P} \left(\sqrt{\varepsilon} \sum_{i=1}^{[\varepsilon^{-1}t_1]} \xi_i \in I_1, \sqrt{\varepsilon} \sum_{i=[\varepsilon^{-1}t_1]+1}^{[\varepsilon^{-1}t_2]} \xi_i \in I_2 \right) \\ &= \mathbb{P} \left(\sqrt{\varepsilon} \sum_{i=1}^{[\varepsilon^{-1}t_1]} \xi_i \in I_1 \right) \cdot \mathbb{P} \left(\sqrt{\varepsilon} \sum_{i=[\varepsilon^{-1}t_1]+1}^{[\varepsilon^{-1}t_2]} \xi_i \in I_2 \right) \longrightarrow \int_{I_1} dx_1 p_{t_1}(x_1) \cdot \int_{I_2} dx_2 p_{t_2-t_1}(x_2)\end{aligned}$$

da cui segue

$$P_\varepsilon(x(t_1) \in I_1, x(t_2) - x(t_1) \in I_2) \longrightarrow \int_{I_1} dx_1 p_{t_1}(x_1) \cdot \int_{I_2} dx_2 p_{t_2-t_1}(x_2)$$

che equivale a

$$P_\varepsilon(x(t_1) \in I_1, x(t_2) \in I_2) \rightarrow p_2(t_1, t_2; I_1, I_2)$$

Il caso $n > 2$ si dimostra con lo stesso argomento. \square

Abbiamo fatto vedere che sui cilindri $\mathcal{C}(\underline{t}, \underline{I})$ la misura P_ε converge a P . D'altra parte è sicuramente vero se che conosciamo una misura $P \in \mathcal{P}(C)$ sui cilindri la conosciamo ovunque. Il seguente esempio mostra che questo non basta a concludere che $P_\varepsilon \implies P$.

Sia P la probabilità concentrata su $x = 0$ e P_n concentrata sulla funzione $x_n(t)$ definita da

$$x_n(t) := \begin{cases} 2nt & \text{se } t \in [0, 1/(2n)) \\ 2 - 2nt & \text{se } t \in [1/(2n), 1/n) \\ 0 & \text{se } t \in [1/n, 1] \end{cases}$$

è ovvio che le distribuzioni finito dimensionali di P_n convergono a quelle di P , d'altra parte sia $f(x) := \sup_{t \in [0,1]} |x(t)|$ che è una funzione continua da C a \mathbb{R} . Poichè $f(x_n) = 1$ mentre $f(0) = 0$ non è possibile che $P_n \implies P$.

Questo esempio che mostra come i cilindri non determinino la convergenza debole è tipico di C . Su \mathbb{R}^n (o anche $\mathbb{R}^{\mathbb{N}}$) è invece vero che se una probabilità converge sui cilindri allora converge debolmente (provare a dimostrarlo).

Pe poter concludere la dimostrazione del Teorema 1.3 ci serve una proprietà di compattezza (relativa) della famiglia $\{P_\varepsilon\}$ (vedi appendice, teoremi 1.26 e 1.27).

t:tpe **Lemma 1.6.** *La famiglia $\{P_\varepsilon\}$ è tight.*

Dimostrazione. Poichè $P_\varepsilon(x(0) = 0) = 1$, la condizione (i) in teorema 1.27 è soddisfatta gratuitamente. Per dimostrare che anche la condizione (ii) è soddisfatta, usiamo la diseguaglianza di Garsia (vedi 1.28). Ricordando che ξ_i sono v.a. indipendenti e che $\mathbb{E}(\xi_i) = 0$, per $0 \leq s < t \leq 1$ tali che $\varepsilon^{-1}t$ e $\varepsilon^{-1}s$ sono interi abbiamo

$$\begin{aligned} \int dP_\varepsilon(x) |x(t) - x(s)|^4 &= \mathbb{E} \left(\sqrt{\varepsilon} \sum_{i=\varepsilon^{-1}s+1}^{\varepsilon^{-1}t} \xi_i \right)^4 \\ &= \varepsilon^2 \sum_{i=\varepsilon^{-1}s+1}^{\varepsilon^{-1}t} \mathbb{E}(\xi_i^4) + 6\varepsilon^2 \sum_{\varepsilon^{-1}s+1 \leq i < j \leq \varepsilon^{-1}t} \mathbb{E}(\xi_i^2 \xi_j^2) \\ &\leq C[\varepsilon(t-s) + (t-s)^2] \leq 2C(t-s)^2 \end{aligned}$$

con un po' di pazienza si verifica che la stessa stima vale se $\varepsilon^{-1}t$ e $\varepsilon^{-1}s$ non sono necessariamente interi. Utilizzando il corollario 1.29 con $r = 4$ e $\alpha = 1$ abbiamo, per $\beta \in (0, 1/4)$,

$$\begin{aligned} P_\varepsilon(x : w_x(\delta) \geq \eta) &\leq P_\varepsilon \left(\delta^\beta \sup_{|t-s| \leq \delta} \frac{|x(t) - x(s)|}{|t-s|^\beta} \geq \eta \right) \\ &\leq \left(\frac{\delta^\beta}{\eta} \right)^r E_\varepsilon \left(\sup_{0 \leq s < t \leq 1} \frac{|x(t) - x(s)|}{|t-s|^\beta} \right)^r \leq C \left(\frac{\delta^\beta}{\eta} \right)^r \end{aligned}$$

in cui si è utilizzata la diseguaglianza di Chebshev.

Da ciò segue

$$\lim_{\delta \rightarrow 0} \sup_{\varepsilon > 0} P_\varepsilon(x : w_x(\delta) \geq \eta) = 0$$

mostra che anche la condizione (ii) del teorema 1.27 è soddisfatta. □

s:ics Il teorema 1.3 segue facilmente dai lemmi 1.5 e 1.6.

1.4. Misura di Wiener su $C(\mathbb{R}_+)$ e invarianza per cambi di scala

Abbiamo costruito la misura di Wiener su $C([0, 1])$, ovviamente nulla cambia se $T > 0$ e consideriamo $C([0, T])$. Vogliamo però definirla direttamente su $C(\mathbb{R}_+) = C([0, \infty))$. La definizione 1.2 della misura di Wiener fa intervenire solo le distribuzioni finito dimensionali e quindi è la stessa sia in $C([0, T])$ che in $C(\mathbb{R}_+)$. Vogliamo però affermare che si può

ottenere come limite debole della passeggiata aleatoria in $C(\mathbb{R}_+)$. Per far ciò dobbiamo dotare $C(\mathbb{R}_+)$ di una struttura metrica definita come segue.

Consideriamo $C(\mathbb{R}_+)$ con la topologia della convergenza uniforme sui compatti, ovvero $x_n(t) \rightarrow x(t)$ sse per ogni $T > 0$ si ha che $x_n(t) \rightarrow x(t)$ converge uniformemente per $t \in [0, T]$. Verificare che questa topologia è metrizzabile (sugg. una possibile metrica è $d(x, y) = \sum_{n=1}^{\infty} 2^{-n} \min \{1, \sup_{t \in [0, n]} |x(t) - y(t)|\}$).

N.B. Tutto quello appena scritto è giusto, ma solo un volgare trucco: ci siamo organizzati la topologia di $C(\mathbb{R}_+)$ in modo che l'affermazione che P_ε (la misura indotta dalla passeggiata aleatoria) converge debolmente a P (la misura di Wiener) è equivalente a dire che per ogni $T > 0$ si ha che P_ε ristretta a $C([0, T])$ converge debolmente a P ristretta a $C([0, T])$. Non abbiamo alcun controllo su quello che succede per $T \rightarrow \infty$!

Una proprietà fondamentale della misura di Wiener è l'invarianza rispetto a cambi di scala diffusivi. Dato $\lambda > 0$ sia $\psi_\lambda : C(\mathbb{R}_+) \rightarrow C(\mathbb{R}_+)$ definita da $(\psi_\lambda x)(t) := \sqrt{\lambda} x(\lambda^{-1}t)$ e P la misura di Wiener su $C(\mathbb{R}_+)$ allora $P = P \circ \psi_\lambda^{-1}$. È importante sottolineare che questo non dice che la funzione $\psi_\lambda x$ è uguale a x (cosa ovviamente falsa), ma solo che hanno la stessa distribuzione. A volte questa invarianza si esprime dicendo che se $x(\cdot)$ è un browniano allora anche $\sqrt{\lambda} x(\lambda^{-1}\cdot)$ lo è (ma non è lo stesso browniano!). Questa proprietà segue facilmente dalla Definizione 1.2 (l'invarianza per cambi di scala diffusivi è scritta nel nucleo p_t), ma è ancora più ovvia dal Teorema 1.3: mandate ε in $\lambda \varepsilon$.

Utilizziamo l'invarianza per cambi di scala per dimostrare che il moto browniano è *ricorrente*, ovvero ripassa infinite volte per ogni punto $a \in \mathbb{R}$.

t:ric **Proposizione 1.7.** *Sia P la misura di Wiener su $C(\mathbb{R}_+)$ allora*

$$P\left(\sup_{t \geq 0} x(t) = +\infty, \inf_{t \geq 0} x(t) = -\infty\right) = 1$$

Dimostrazione. Poiché $-x(t)$ è un moto browniano tanto quanto $x(t)$ basta dimostrare che

$$P\left(\sup_{t \geq 0} x(t) = +\infty\right) = 1$$

Sia $X^* := \sup_{t \geq 0} x(t)$ (è una variabile aleatoria a valori $\mathbb{R}_+ \cup \{+\infty\}$); dall'invarianza per cambi di scala diffusivi segue che per $\lambda > 0$ le variabili aleatorie X^* e λX^* hanno la stessa legge e quindi (perché?) X^* è concentrata su $\{0, \infty\}$. Basta quindi dimostrare che $P(X^* = 0) = 0$. L'idea è ora di usare l'indipendenza degli incrementi per chiudere una disuguaglianza.

Evidentemente

$$P(X^* = 0) \leq P\left(x(1) \leq 0, x(t) \leq 0 \forall t \geq 1\right)$$

Ora per $s \in \mathbb{R}_+$, l'incremento $x(1+s) - x(1)$ è ancora un moto browniano, pertanto $\sup_{s \geq 0} [x(1+s) - x(1)] \in \{0, \infty\}$ per quello visto sopra. Dato che $x(1)$ è finito troviamo

allora

$$\begin{aligned}
P(X^* = 0) &\leq P\left(x(1) \leq 0, x(t) \leq 0 \forall t \geq 1, \sup_{s \geq 0} [x(1+s) - x(1)] = 0\right) \\
&\quad + P\left(x(1) \leq 0, x(t) \leq 0 \forall t \geq 1, \sup_{s \geq 0} [x(1+s) - x(1)] = +\infty\right) \\
&= P\left(x(1) \leq 0, \sup_{s \geq 0} [x(1+s) - x(1)] = 0\right) \\
&= P(x(1) \leq 0) P(X^* = 0) = \frac{1}{2} P(X^* = 0)
\end{aligned}$$

dove abbiamo usato l'indipendenza di $x(1+s) - x(1)$, $s \geq 0$ rispetto a $x(1)$. Dalla precedentemente diseuguaglianza concludiamo infine che $P(X^* = 0) = 0$. \square

s:rp

1.5. Principio di riflessione

Come applicazione della costruzione precedente della misura di Wiener, dimostriamo il seguente risultato

t:rp

Teorema 1.8. *Sia x il moto browniano e $t > 0$; allora $\sup_{s \leq t} x(s) \stackrel{Law}{=} |x(t)|$ ovvero per ogni $\alpha \in \mathbb{R}_+$*

$$P\left(\sup_{s \in [0, t]} x(s) < \alpha\right) = P(|x(t)| < \alpha) = 2 \int_0^\alpha dx p_t(x) \quad (1.3) \quad \mathbf{rp}$$

Osservazione 1. Se, dato $\alpha > 0$ introduciamo la variabile aleatoria $\tau_\alpha := \inf \{t \geq 0 : x(t) = \alpha\}$ (ovvero il primo tempo in cui il browniano $x(t)$ raggiunge α), il principio di riflessione (1.3) si può riformulare come

$$P(\tau_\alpha \leq t) = P\left(\sup_{s \in [0, t]} x(s) \geq \alpha\right) = 2 \int_\alpha^\infty dx p_t(x)$$

da cui segue (verificare. Sugg. integrare per parti)

$$\frac{d}{dt} P(\tau_\alpha \leq t) = \frac{1}{\sqrt{2\pi}} \frac{\alpha}{t^{3/2}} \exp\left\{-\frac{\alpha^2}{2t}\right\}$$

in particolare $P(\tau_\alpha < \infty) = 1$ (verificare), ma $E(\tau_\alpha) = \infty$. Riconciare la propria intuizione sul moto browniano con questa proprietà.

Osservazione 2. Siano ξ_i indipendenti ed identicamente distribuite come in sezione 1.3, $S_n = \sum_{i=1}^n \xi_i$. Dal teorema 1.3 e (1.3) segue allora

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{1}{\sqrt{n}} \sup_{i \leq n} S_i < \alpha\right) = 2 \int_0^\alpha dx \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

che indentifica il comportamento asintotico del massimo di somme di variabili indipendenti.

Dimostrazione. L'idea è la seguente. Abbiamo

$$\begin{aligned} P\left(\sup_{s \in [0, t]} x(s) > \alpha\right) &= P\left(\sup_{s \in [0, t]} x(s) > \alpha, x(t) > \alpha\right) + P\left(\sup_{s \in [0, t]} x(s) > \alpha, x(t) < \alpha\right) \\ &= P(x(t) > \alpha) + P\left(\sup_{s \in [0, t]} x(s) > \alpha, x(t) < \alpha\right) \end{aligned}$$

Sia $\tau_\alpha := \inf\{t > 0 : x(t) = \alpha\}$, come in osservazione 1, il primo tempo in cui il moto browniano arriva in α . Ora, poichè gli incrementi del moto browniano sono indipendenti possiamo riflettere intorno a α la traiettoria $x(t)$ dopo τ_α , ovvero definire

$$\tilde{x}(t) := \begin{cases} x(t) & \text{se } t \in [0, \tau_\alpha] \\ 2\alpha - x(t) & \text{se } t \geq \tau_\alpha \end{cases}$$

che dovrebbe avere la stessa probabilità di $x(t)$. Otteniamo quindi

$$P\left(\sup_{s \in [0, t]} x(s) > \alpha, x(t) < \alpha\right) = P\left(\sup_{s \in [0, t]} x(s) > \alpha, x(t) > \alpha\right)$$

da cui l'enunciato segue facilmente.

Questa idea si può formalizzare, vedi [9, §2.6]. D'altra parte a livello discreto (per la passeggiata aleatoria) è già corretta e dimostriamo (1.3) come limite di una formula analoga per la passeggiata aleatoria semplice. Siano ξ_i i.i.d. (indipendenti e identicamente distribuite) con $\mathbb{P}(\xi_i = \pm 1) = 1/2$, $S_n = \sum_{i=1}^n \xi_i$ e a intero positivo. Allora

$$\mathbb{P}\left(\sup_{i \leq n} S_i \geq a\right) = 2\mathbb{P}(S_n > a) + \mathbb{P}(S_n = a) \quad (1.4) \quad \boxed{\text{rprw}}$$

da questa identità, via scaling e teorema 1.3, la formula (1.3) segue (notate che il secondo termine a secondo membro della (1.4) non dà contributo nel limite).

Per dimostrare (1.4) osserviamo che

$$\mathbb{P}\left(\sup_{i \leq n} S_i \geq a\right) = \mathbb{P}(S_n = a) + \mathbb{P}\left(\sup_{i \leq n} S_i \geq a, S_n > a\right) + \mathbb{P}\left(\sup_{i \leq n} S_i \geq a, S_n < a\right)$$

e (1.4) segue da

$$\mathbb{P}\left(\sup_{i \leq n} S_i \geq a, S_n > a\right) = \mathbb{P}\left(\sup_{i \leq n} S_i \geq a, S_n < a\right) \quad (1.5) \quad \boxed{\text{rprw1}}$$

Avendo scelto la passeggiata aleatoria semplice, ci sono 2^n possibili traiettorie S_1, \dots, S_n che hanno tutte la medesima probabilità. La riflessione intorno a a (dopo che S_i raggiunge a) stabilisce ora una corrispondenza biunivoca tra le traiettorie che contribuiscono all'evento a primo membro di (1.5) con quelle che contribuiscono all'evento a secondo membro e quindi (1.5) vale. \square

s:ptbm

1.6. (Ir)regolarità delle traiettorie browniane

t:hbm

Proposizione 1.9. *Con probabilità 1 rispetto a P , $t \rightarrow x(t)$ è Hölder continua di esponente α per ogni $\alpha < 1/2$.*

Dimostrazione. Basta utilizzare il corollario 1.29. Un calcolo gaussiano mostra infatti che

$$\int dx p_t(x) x^{2n} = (2n - 1)!! t^n$$

pertanto, ricordando che l'incremento $x(t) - x(s)$ ha distribuzione $p_{t-s}(x)dx$, abbiamo $E(|x(t) - x(s)|^{2n}) \leq C_n |t - s|^n$. Prendendo n grande abbastanza la proposizione segue. \square

Sia Δ una partizione di $[0, t]$ con $0 = t_0 < t_1 < \dots < t_n = t$ e $|\Delta| := \max_i |t_i - t_{i-1}|$. Definiamo la *variazione quadratica* rispetto a Δ di $x \in C([0, t])$ come

$$V_{\Delta}^2(x) := \sum_{i=1}^n |x(t_i) - x(t_{i-1})|^2 \tag{1.6} \quad \text{dqvd}$$

Osserviamo che non è vero, in generale, che $V_{\Delta}^2(x) \leq V_{\Delta'}^2(x)$ se Δ' è più fine di Δ (come avviene per la variazione *tout court*).

t:qvbm

Teorema 1.10. *Sia P la misura di Wiener. Allora*

$$\lim_{|\Delta| \rightarrow 0} V_{\Delta}^2(x) = t \quad \text{in } L_2(P) \tag{1.7} \quad \text{qvbm}$$

Osservazione Da (1.7) segue che esiste una sottosuccessione Δ_n per cui $V_{\Delta_n}^2(x) \rightarrow t$ P -a.s. In realtà dalla dimostrazione seguente segue (esercizio) che la convergenza con probabilità uno è garantita per una successione di partizioni Δ_n tale che $\sum_n |\Delta_n| < \infty$, ad esempio la partizione fatta con i diadici, ovvero $\Delta_n = \{0, t 2^{-n}, \dots, t i 2^{-n}, \dots, t\}$. Si può dimostrare che vale la convergenza P -a.s. anche per partizioni che si raffinano, $\Delta_{n+1} \supset \Delta_n$, vedi [12, Prop. II.2.12]. D'altra parte per la "vera" variazione quadratica si ha $\sup_{\Delta} V_{\Delta}^2(x) = \infty$ P -a.s.

Dimostrazione. Poichè gli incrementi di x sono indipendenti e con distribuzione gaussiana, abbiamo

$$\begin{aligned} E(V_{\Delta}^2(x) - t)^2 &= E\left(\sum_{i=1}^n [|x(t_i) - x(t_{i-1})|^2 - (t_i - t_{i-1})]\right)^2 \\ &= \sum_{i=1}^n E[|x(t_i) - x(t_{i-1})|^2 - (t_i - t_{i-1})]^2 = \sum_{i=1}^n 2(t_i - t_{i-1})^2 \end{aligned}$$

che converge a 0 per $|\Delta| \rightarrow 0$. \square

t:ivbm

Corollario 1.11. *Con probabilità 1 rispetto a P , $t \rightarrow x(t)$ ha variazione infinita in ogni intervallo.*

Dimostrazione. Per il teorema precedente, esiste un insieme $\bar{\Omega} \subset C$ con $P(\bar{\Omega}) = 1$ per cui vale la seguente proprietà. Data comunque una coppia di razionali $p < q$ esiste una successione di partizioni Δ_n , $|\Delta_n| \rightarrow 0$, di $[p, q]$ per cui

$$\lim_{n \rightarrow \infty} \sum_{t_i \in \Delta_n} [x(t_i) - x(t_{i-1})]^2 = q - p$$

per ogni $x \in \bar{\Omega}$.

Sia ora $V(x)$ la variazione di x su $[p, q]$, abbiamo

$$\sum_{t_i \in \Delta_n} [x(t_i) - x(t_{i-1})]^2 \leq V(x) \sup_{t_i \in \Delta_n} [x(t_i) - x(t_{i-1})]$$

se $V(x) < \infty$ il secondo membro converge a zero poichè x è continuo; troviamo quindi un assurdo. \square

t:ndbm

Corollario 1.12. *Con probabilità 1 rispetto a P , $t \rightarrow x(t)$ non è Hölder continuo con esponente $\alpha > 1/2$ in alcun intervallo.*

Osservazione L'affermazione di Corollario 1.12 non è equivalente a dire che con probabilità 1 rispetto a P , $t \rightarrow x(t)$ non è Hölder continuo con esponente $\alpha > 1/2$ in alcun punto. Infatti, come sarà chiaro dalla dimostrazione, l'affermazione in Corollario 1.12 si riduce immediatamente ad un'affermazione su di un insieme numerabile di punti, l'affermazione sopra fatta richiede un minimo di accortezza in più, vedi [9, Rem. 2.9.15] per un diverso esempio di questo fenomeno. Ciò non di meno, l'affermazione sopra fatta è vera e nemmeno troppo difficile da dimostrare, vedi [9, Ex. 2.9.21]

Dimostrazione. La dimostrazione è analoga al corollario precedente. Se esistesse un intervallo $[p, q]$ per cui $|x(t) - x(s)| \leq C|t - s|^\alpha$ con $\alpha > 1/2$ troveremmo

$$\sum_{t_i \in \Delta_n} [x(t_i) - x(t_{i-1})]^2 \leq C^2(q - p) \sup_{t_i \in \Delta_n} |t_i - t_{i-1}|^{2\alpha-1}$$

che conduce ad un assurdo. \square

s:upbm

1.7. Ulteriori proprietà delle traiettorie browniane (senza dimostrazioni)

Proposizione 1.9 e corollario 1.12 lasciano un certo margine su quale sia il corretto modulo di continuità del moto browniano. Per la dimostrazione del seguente teorema vedi ad esempio [9, Thm. 2.9.25]. Un risultato un po' più debole si può trovare via una scelta astuta di q e Ψ nella disegualianza di Garsia, vedi [15, Ex. 2.4.8].

t:lmcbm

Teorema 1.13. *Sia P la misura di Wiener, allora*

$$P \left(\overline{\lim}_{\delta \rightarrow 0} \sup_{|t-s| \leq \delta} \frac{|x(t) - x(s)|}{\sqrt{2\delta \log(1/\delta)}} = 1 \right) = 1$$

Il comportamento di una traiettoria browniana in punto fissato, ad esempio $t = 0$ è invece lievemente diverso, cambia di un logaritmo. Vale infatti il seguente risultato, noto come *legge dei logaritmi iterati*.

t:11i **Teorema 1.14.** *Sia P la misura di Wiener, allora*

$$P \left(\overline{\lim}_{t \downarrow 0} \frac{x(t)}{\sqrt{2t \log \log(1/t)}} = 1 \right) = 1$$

Evidentemente, per la simmetria della misura di Wiener, vale pure

$$P \left(\underline{\lim}_{t \downarrow 0} \frac{x(t)}{\sqrt{2t \log \log(1/t)}} = -1 \right) = 1$$

In particolare segue che P -a.s. esiste una successione $t_n \downarrow 0$ per cui $x(t_n) = 0$: prima di lasciare lo zero il moto browniano ci ripassa infinite volte.

Si può utilizzare un'ulteriore proprietà di invarianza della misura di Wiener per ottenere il comportamento delle traiettorie per $t \rightarrow \infty$ da teorema 1.14. Più precisamente se $x(t)$, $t \geq 0$ è un moto browniano allora anche $\tilde{x}(t) := tx(1/t)$ è un moto browniano. Esercizio: verificare la precedente affermazione (in particolare si deve dimostrare che $\lim_{t \downarrow 0} \tilde{x}(t) = 0$ con probabilità uno). Da teorema 1.14 segue allora

$$P \left(\overline{\lim}_{t \rightarrow \infty} \frac{x(t)}{\sqrt{2t \log \log t}} = 1, \underline{\lim}_{t \rightarrow \infty} \frac{x(t)}{\sqrt{2t \log \log t}} = -1 \right) = 1$$

Notiamo infine che dalla legge dei logaritmi iterati per il moto browniano si può ricavare l'analoga affermazione per una passeggiata aleatoria; con le solite notazioni

$$\mathbb{P} \left(\overline{\lim}_{n \rightarrow \infty} \frac{S_n}{\sqrt{2n \log \log n}} = 1 \right) = 1$$

La dimostrazione di teorema 1.14 è basata sulla proprietà di martingala del moto browniano e la diseguaglianza di Doob, si veda [12, Thm. II.1.9] o [9, Thm. 2.9.23].

Sia $x(t)$ una traiettoria browniana, ci chiediamo quanto tempo passi nel semiasse positivo. Per $t > 0$ definiamo pertanto la variabile aleatoria

$$A_t^+(x) := \int_0^t ds \mathbb{1}_{(0, \infty)}(x(s))$$

Per ovvio motivi di simmetria il browniano ha tanta probabilità di essere positivo quanto negativo, ma presa una traiettoria $x(\cdot)$ è molto più probabile che questa abbia passato un frazione significativa del tempo in un semiasse piuttosto che sia stata (circa) positiva per metà del tempo e negativa per l'altra metà. Ricordiamo che per la ricorrenza del moto browniano $x(\cdot)$ ripassa comunque per zero infinite volte.

La precedente discussione qualitativa acquista un senso in virtù del risultato seguente, noto come *legge dell'arcoseno*.

t:las0 **Teorema 1.15.** *Sia P la misura di Wiener, allora per $\theta \in [0, 1]$*

$$P\left(\frac{1}{t}A_t^+(x) \leq \theta\right) = \int_0^\theta du \frac{1}{\pi\sqrt{u(1-u)}} = \frac{2}{\pi} \arcsin \sqrt{\theta}$$

Vi sono varie dimostrazioni del precedente teorema. Forse la più “elementare” è basata sui dei calcoli espliciti per la passeggiata aleatoria ed utilizza il principio di invarianza, vedi [2, §11] in cui la legge dell’arccoseno è ottenuta come caso particolare. Altre dimostrazioni utilizzano strumenti più avanzati come la teoria dei tempi locali per il moto browniano, vedi [12, Thm. VI.2.7], oppure la connessione con equazioni alle derivate parziali (formula di Feynman–Kac), vedi nel seguito Teorema 4.8.

s:mart

1.8. Definizione di martingala e diseguaglianza di Doob

La teoria delle martingale occupa ormai diversi tomi dei manuali di probabilità, ed raggiunge livelli di eleganza sopraffini. È comunque bene avere in mente il seguente esempio elementare.

Consideriamo una successione di risultati del colore uscito nei giri di una roulette (senza zero). Questa situazione è modellata da una successione di v.a. ξ_k , $k \in \mathbb{N}$ i.i.d. con $\mathbb{P}(\xi_i = \pm 1) = 1/2$, qui abbiamo codificato “rosso” con -1 e “nero” con 1 . Ora noi siamo al banco della roulette e possiamo scegliere la strategia di gioco che preferiamo, ovvero dopo n giri possiamo decidere di puntare, per l’ $n + 1$ -mo, quello che vogliamo sul rosso o sul nero. Però non possiamo barare: dobbiamo puntare prima del faditico “rien ne va plus”; mentre sappiamo tutto dei risultati dei primi n giri, nulla ci è noto di quelli successivi. Una strategia di gioco è allora definita dichiarando che all’ $n + 1$ -mo giro puntiamo $f_n(\xi_1, \dots, \xi_n)$ euro sul rosso. Qui $f(\cdot)$ è una funzione reale arbitraria di n , ξ_1, \dots, ξ_n . Se $f_n(\xi_1, \dots, \xi_n) < 0$ vuol dire che puntiamo $-f_n(\xi_1, \dots, \xi_n)$ sul nero. Ora costruiamo il bilancio dei soldi vinti (o persi) nei primi n giri, cioè definiamo

$$M_n = \sum_{k=1}^n f_{k-1}(\xi_1, \dots, \xi_{k-1}) \xi_k, \quad n = 1, 2, \dots \quad (1.8) \quad \text{em}$$

e decidiamo anche $M_0 = 0$. La successione M_n è il prototipo di una martingala ed anche -in qualche senso- la generica martingala. È immediato verificare che $\mathbb{E}(M_n) = 0$ per ogni n , ovvero per una qualsiasi strategia di gioco la vincita media è zero (ma va!).

Passiamo ora alla definizione formale considerando solo martingale a tempo continuo, sarà evidente come modificare la definizione nel caso (come nell’esempio precedente) di tempo discreto. Come al solito ci diamo uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ e vogliamo aggiungere una struttura in più detta *filtrazione* \mathcal{F}_t , $t \in \mathbb{R}_+$. La *filtrazione* \mathcal{F}_t è semplicemente una famiglia crescente di sotto σ -algebre di \mathcal{F} , ovvero $\mathcal{F}_t \subset \mathcal{F}$, $\mathcal{F}_s \subset \mathcal{F}_t$ per $0 \leq s \leq t$.

Convieni inoltre assumere le seguenti condizioni sulla filtrazione \mathcal{F}_t (*condizione usuale*): \mathcal{F}_0 contiene tutti gli insiemi \mathbb{P} nulli per $\mathcal{F}_\infty := \bigvee_{t \geq 0} \mathcal{F}_t := \sigma\left\{\bigcup_{t \geq 0} \mathcal{F}_t\right\}$ ed \mathcal{F}_t è continua da destra, ovvero $\mathcal{F}_{s+} := \bigcap_{t > s} \mathcal{F}_t = \mathcal{F}_s$.

Vogliamo descrivere un processo stocastico (= famiglia di v.a. indicizzate da $t \in \mathbb{R}_+$) ed il senso intuitivo della filtrazione \mathcal{F}_t è quello di un sacco dove conserviamo l'informazione accumulata fino al tempo t . Nell'esempio precedente la filtrazione è invece discreta: $\mathcal{F}_k = \sigma\{\xi_1, \dots, \xi_k\}$. Possiamo ora dare la definizione formale di martingala. Vedi sezione 1.C per una brevissima introduzione delle probabilità condizionate.

d:mart

Definizione 1.16. *Dato uno spazio di probabilità filtrato $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ una martingala M è una famiglia di v.a. $M_t = M_t(\omega)$, $t \in \mathbb{R}_+$ tale che*

1. *Per ogni $t \in \mathbb{R}_+$ la v.a. M_t è \mathcal{F}_t misurabile (questa condizione si abbrevia dicendo che M è adattata alla filtrazione \mathcal{F}_t);*
2. *Per ogni $s, t \in \mathbb{R}_+$ con $s < t$ si ha $\mathbb{E}(M_t | \mathcal{F}_s) = M_s$.*

Diciamo inoltre che M è una martingala continua se $\mathbb{P}(\{\omega : t \mapsto M_t(\omega) \text{ è continua}\}) = 1$.

Osserviamo che la definizione di martingala richiede (implicitamente) che per ogni $t \in \mathbb{R}_+$ M_t abbia attesa finita, $\mathbb{E}|M_t| < \infty$. La successione (1.8) è un esempio di martingala a tempo discreto (verificare) e, come vedremo, il moto browniano è un esempio di martingala continua.

Un'esercizio (facile) di teoria della misura, vedi [9, Prop. 1.1.13], mostra che nel caso di processi continui (o solo continui da destra) la condizione 1 equivale in realtà alla seguente condizione più forte (si esprime dicendo che M è *progressivamente misurabile* rispetto a \mathcal{F}_t). Per ogni $t \in \mathbb{R}_+$ l'applicazione $[0, t] \times \Omega \ni (s, \omega) \mapsto M_s(\omega) \in \mathbb{R}$ è $\mathcal{B}([0, t]) \times \mathcal{F}_t$ misurabile.

Il seguente teorema di convergenza illustra l'utilità della nozione di martingala, che svolge un ruolo analogo a quello di successione monotona.

t:mc

Teorema 1.17. *Sia M una martingala continua (o solo continua da destra) tale che $\sup_{t>0} \mathbb{E}(|M_t|) < \infty$ allora $\lim_{t \rightarrow \infty} M_t(\omega)$ esiste con \mathbb{P} probabilità 1.*

Vedi [9, Thm. 1.3.15] per la dimostrazione, è però altamente consigliata la lettura di [17, Cap. 10,11] (ove si considera il caso di tempo discreto) per la nozione di martingala ed il teorema di convergenza.

Un'altro risultato molto utile per le martingale è la seguente *diseguaglianza di Doob*, di cui faremo uso ripetuto nella teoria delle equazioni stocastiche.

t:dlp

Teorema 1.18. *Sia M_t una martingala continua (o solo continua da destra) allora: per ogni $p \in (1, \infty)$, $T \in \mathbb{R}_+$ vale*

$$\mathbb{E}\left(\sup_{t \in [0, T]} |M_t|^p\right) \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}|M_T|^p \quad (1.9) \quad \text{dlp}$$

Inoltre per ogni $\lambda > 0$, $p \in [1, \infty)$, $T \in \mathbb{R}_+$

$$\mathbb{P}\left(\sup_{t \in [0, T]} |M_t| \geq \lambda\right) \leq \frac{1}{\lambda^p} \mathbb{E}|M_T|^p \quad (1.10) \quad \text{dp12}$$

Si sottolinea come (1.9) valga solo per $p > 1$ (per $p = 1$ esistono controesempi) mentre (1.10) vale anche per $p = 1$, anzi nelle applicazioni è spesso usata in tale caso.

Confrontare la diseguaglianza di Doob con il principio di riflessione (Teo. 1.8) che ci dice la legge di $\sup_{t \in [0, T]} M_t$ quando M_t è il moto browniano.

Si può utilizzare la diseguaglianza di Doob (invece della diseguaglianza di Garsia) per dimostrare facilmente la compattezza della misura su $C(\mathbb{R}_+)$ indotta dalla passeggiata aleatoria (Lemma 1.6); è infatti immediato vedere che una passeggiata aleatoria simmetrica è una martingala. Qualche dettaglio in più. È facile verificare che (ii) in teorema 1.27 è implicata dalla condizione seguente. Per ogni $\eta > 0$ e $t \in [0, 1]$

$$\lim_{\delta \downarrow 0} \overline{\lim}_{\varepsilon \rightarrow 0} \frac{1}{\delta} P_\varepsilon \left(\sup_{s \in [t, t+\delta]} |x(s) - x(t)| > \eta \right) = 0 \quad (1.11) \quad \boxed{\text{bpc}}$$

A questo punto l'applicazione della diseguaglianza di Doob (1.10) con $p = 4$ alla martingala $S_{[\varepsilon^{-1}t]} - S_{[\varepsilon^{-1}s]}$, $t \in [s, s + \delta]$ permette di concludere la dimostrazione. Con un argomento un po' più sofisticato, vedi [2, §10], si può anche dimostrare teorema 1.3 senza la condizione $\mathbb{E}(\xi_i^4) < \infty$.

Esercizio. Verificare che

$$\sup_{t \in [0, T]} \mathbb{E}|M_t|^p = \mathbb{E}|M_T|^p$$

Sugg. usare la proprietà di martingala e la diseguaglianza di Jensen.

La stima (1.9) segue facilmente dal lemma seguente. Consideriamo un numero finito di tempi $t_1 < \dots < t_n$ in $[0, T]$ e scriviamo M_i invece di M_{t_i} , $i = 1, \dots, n$

$\boxed{\text{t:ldlp}}$

Lemma 1.19. *Siano $M_*^+ := \sup\{0, M_1, \dots, M_n\}$ e $M_n^+ := [M_n + |M_n|]/2$ allora, per ogni $p > 1$*

$$\mathbb{E}((M_*^+)^p) \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}(M_n^+)^p \quad (1.12) \quad \boxed{\text{ldlp}}$$

Dimostrazione. Per $a > 0$ sia

$$\chi_k(a) := \begin{cases} 1 & \text{se } M_1 < a, \dots, M_{k-1} < a, M_k \geq a \\ 0 & \text{altrimenti} \end{cases}$$

allora

$$\sum_{k=1}^n \chi_k(a) = \begin{cases} 1 & \text{se } M_*^+ \geq a \\ 0 & \text{altrimenti} \end{cases}$$

Poiché $x^p = \int_0^\infty da pa^{p-1} \mathbb{1}_{x \geq a}$ e $a\chi_k(a) \leq M_k \chi_k(a)$ abbiamo

$$(M_*^+)^p = \int_0^\infty da pa^{p-1} \sum_{k=1}^n \chi_k(a) \leq p \int_0^\infty da a^{p-2} \sum_{k=1}^n M_k \chi_k(a)$$

Dato che M è adattata, si ha che $\chi_k(a)$ è misurabile rispetto a $\mathcal{F}_k := \mathcal{F}_{t_k}$. Dalla condizione di martingalità segue allora

$$\mathbb{E}([M_n - M_k] \chi_k(a)) = \mathbb{E}(\chi_k(a) \mathbb{E}(M_n - M_k | \mathcal{F}_k)) = 0$$

pertanto

$$\begin{aligned} \frac{1}{p} \mathbb{E}(M_*^+)^p &\leq \sum_{k=1}^n \int_0^\infty da a^{p-2} \mathbb{E}(M_k \chi_k(a)) \\ &\leq \sum_{k=1}^n \int_0^\infty da a^{p-2} \mathbb{E}(M_n^+ \chi_k(a)) = \frac{1}{p-1} \mathbb{E}(M_n^+ (M_*^+)^{p-1}) \end{aligned}$$

Utilizzando Hölder troviamo

$$\mathbb{E}(M_*^+)^p \leq \frac{p}{p-1} \mathbb{E}(M_n^+ (M_*^+)^{p-1}) \leq \frac{p}{p-1} \left\{ \mathbb{E}((M_*^+)^p) \right\}^{(p-1)/p} \left\{ \mathbb{E}((M_n^+)^p) \right\}^{1/p}$$

che conclude la dimostrazione. \square

La dimostrazione della diseguaglianza (1.10) è più facile, nello stesso contesto di Lemma 1.19 definiamo ora

$$\chi_k(a) := \begin{cases} 1 & \text{se } |M_1| < a, \dots, |M_{k-1}| < a, |M_k| \geq a \\ 0 & \text{altrimenti} \end{cases}$$

allora

$$\mathbb{P}\left(\sup_{1 \leq k \leq n} |M_k| \geq a\right) = \mathbb{E} \sum_{k=1}^n \chi_k(a) \leq \frac{1}{a^p} \sum_{k=1}^n \mathbb{E}(|M_k|^p \chi_k(a))$$

Osserviamo ora che per la convessità di $|x|^p$, $p \geq 1$, la diseguaglianza di Jensen (convicersi che si può usare Jensen anche per attese condizionate) e la proprietà di martingala si ha $\mathbb{E}(|M_n|^p | \mathcal{F}_k) \geq |\mathbb{E}(M_n | \mathcal{F}_k)|^p = |M_k|^p$ per $k < n$. È ora immediato concludere la dimostrazione di (1.10).

s:damb1

1.9. Definizione astratta del moto browniano e teorema di Levy

A volte conviene dare una definizione astratta di moto browniano. Ad esempio vogliamo considerare come punto iniziale non l'origine (o un qualsiasi punto fissato), ma un punto aleatorio. Con la sola misura di Wiener non c'è "abbastanza stocasticità" per modellare tale situazione, ma è chiaro come procedere: prediamo \mathbb{R} con una probabilità μ (la legge della condizione iniziale) e facciamo il prodotto con $C(\mathbb{R}_+)$ con la misura di Wiener. A questo il moto browniano con dato iniziale di legge μ sarà definito sullo spazio prodotto (a valori in $C(\mathbb{R}_+)$) da $B(t, \omega) = \omega_1 + \omega_2(t)$ ove $(\omega_1, \omega_2) \in \mathbb{R} \times C(\mathbb{R}_+)$.

t:damb

Definizione 1.20. Dato uno spazio di probabilità filtrato $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, il moto browniano standard $\{w_t\}_{t \in \mathbb{R}_+}$ (la scelta della lettera w è in omaggio a Wiener) è una collezione di variabili aleatorie $w_t : \Omega \rightarrow \mathbb{R}$ tali che:

1. per ogni $t \in \mathbb{R}_+$ la variabile aleatoria w_t è \mathcal{F}_t -misurabile con legge data da una gaussiana di media zero e varianza uno;
2. per ogni $0 \leq s < t$ l'incremento $w_t - w_s$ è indipendente dalla σ -algebra \mathcal{F}_s ;

3. le traiettorie $t \mapsto w_t$ sono continue con \mathbb{P} probabilità 1.

Verificare che dalla definizione precedente segue che il moto browniano è ad incrementi indipendenti come abbiamo richiesto nella Definizione 1.2.

Rispetto alla costruzione della misura di Wiener fatta in precedenza, si può prendere $\Omega = C(\mathbb{R}_+)$, \mathcal{F} la σ -algebra dei boreliani, \mathbb{P} la misura di Wiener e definire $w_t(\omega) := \omega(t)$, ovvero w_t è la coordinata “canonica” su $C(\mathbb{R}_+)$. Detta $\mathcal{A}_t := \sigma\{w_s, s \leq t\}$ la *filtrazione canonica* in $C(\mathbb{R}_+)$ una filtrazione che soddisfa la condizione usuale è data da $\mathcal{F}_t = \mathcal{A}_t \vee \mathcal{N} := \sigma\{\mathcal{A}_t \cup \mathcal{N}\}$ dove \mathcal{N} è la collezione degli insiemi \mathbb{P} -nulli per \mathcal{A}_∞ ; ovvero $N \in \mathcal{N}$ sse esiste $M \in \mathcal{A}_\infty$ tale che $N \subset M$ e $\mathbb{P}(M) = 0$.

Come accennato precedentemente, il vantaggio (minimo) di questa definizione più astratta di moto browniano è che lo spazio di probabilità $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ può essere “più grande” della particolare realizzazione scelta nella costruzione della misura di Wiener; questo può essere utile se si vuole modellare un fenomeno che coinvolga altri aspetti aleatori oltre al moto browniano. Ricordiamo anche che il senso intuitivo della filtrazione \mathcal{F}_t è quello di un sacco dove conserviamo l’informazione accumulata fino al tempo t : un evento è in \mathcal{F}_t se possiamo decidere se è capitato oppure no guardando solo alla traiettoria browniana fino al tempo t ; anche in questo caso \mathcal{F}_t può essere “più grande”, richiediamo però che - noto \mathcal{F}_t - sappiamo tutto della porzione di traiettoria $\{w_s\}_{s \in [0, t]}$ e che l’incremento $w_{t+h} - w_t$, $h > 0$ sia indipendente da quello che succede prima di t . Questo è importante! Se vogliamo mettere una condizione iniziale aleatoria al browniano la dobbiamo prendere indipendente dal resto; non a caso nell’esempio precedente ci siamo organizzati la misura prodotto. Se gli avessimo dato una condizione iniziale che dipendeva dalla storia futura sarebbe venuto un qualche processo complicato, non un moto browniano. La condizione 2 in definizione 1.20 richiede che l’incremento sia indipendente dalla storia pregressa e questo (se ci sono altri oggetti aleatori) è più che chiedere la mera indipendenza degli incrementi.

La definizione 1.20 caratterizza univocamente il moto browniano nel senso che se pensiamo a w_t , $t \in \mathbb{R}_+$ come una mappa da Ω a $C(\mathbb{R}_+)$ (cosa che possiamo fare con probabilità uno in virtù della richiesta della continuità delle traiettorie), allora la misura $P := \mathbb{P} \circ (w.)^{-1}$ è quella di Wiener.

Dalla definizione 1.20 segue immediatamente (verificare) che

$$(i) \mathbb{E}(w_t - w_s | \mathcal{F}_s) = 0$$

$$(ii) \mathbb{E}([w_t - w_s]^2 | \mathcal{F}_s) = t - s$$

In altri termini (verificare) w_t e $w_t^2 - t$ sono martingale su $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$.

Poiché dalla definizione di moto browniano segue che l’incremento $w_t - w_s$ è indipendente dalla σ -algebra \mathcal{F}_s ed ha distribuzione gaussiana di varianza $t - s$ segue anche (verificare) che per ogni $\lambda \in \mathbb{R}$ (anzi per ogni $\lambda \in \mathbb{C}$) il processo

$$\mathcal{M}_t^\lambda := \exp \left\{ \lambda w_t - \frac{1}{2} \lambda^2 t \right\}$$

è una martingala continua. Questa famiglia di *martingale esponenziali* risulta cruciale in diverse circostanze.

Si può dimostrare (teorema di Levy) le proprietà (i) e (ii) sopra, unite alla continuità delle traiettorie, caratterizzano completamente il moto browniano. Più precisamente vale il seguente risultato.

t:levy

Teorema 1.21. *Sia $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ uno spazio di probabilità filtrato e M una martingala continua tale che $M_t^2 - t$ è una martingala e $M_0 = 0$ con \mathbb{P} probabilità 1; allora M è il moto browniano standard.*

L'ipotesi di continuità è essenziale: se N_t è il processo di Poisson allora $N_t - t$ è una martingala tale che $(N_t - t)^2 - t$ è ancora una martingala. Le dimostrazioni “moderne” del teorema di Levy, si veda [9, Thm. 3.3.16], [12, Thm. IV.3.6] utilizzano il calcolo stocastico (che noi svilupperemo in seguito), una dimostrazione con meno prerequisiti si trova in [3].

Il punto cruciale consiste nel riconoscere, per ogni $\theta \in \mathbb{R}$, nel processo

$$\mathcal{M}_t^\lambda := \exp \left\{ i\theta M_t + \frac{1}{2}\theta^2 t \right\}$$

una martingala. Infatti ciò equivale (è la stessa verifica richiesta sopra) a dire che

$$\mathbb{E} \left(e^{i\theta(M_t - M_s)} | \mathcal{F}_s \right) = e^{-\frac{1}{2}\theta^2(t-s)}$$

e questo implica che $M_t - M_s$ è una variabile gaussiana di varianza $t - s$ indipendente dalla σ -algebra \mathcal{F}_s , ovvero l'incremento di un browniano.

Nel dimostrare che \mathcal{M}_t^λ è una martingala possiamo prima ridurci ad intervalli di tempo piccoli. Ora uno sviluppo di Taylor al secondo ordine è quello che riconosce in \mathcal{M}_t^λ una martingala a partire dalla martingalità di M_t e $M_t^2 - t$. Tutto il lavoro consiste nel mostrare che i termini di ordine superiore non contano. Per fare questo abbiamo a disposizione “solo” l'ipotesi di continuità delle traiettorie.

Il calcolo della variazione quadratica per il moto browniano (Teo. 1.10) è in realtà valido per tutte le martingale continue. Più precisamente data una martingala continua M_t (al quadrato integrabile) esiste un unico processo crescente che si annulla in zero, detto *variazione quadratica* della martingala M_t e denotato con $\langle M \rangle_t$ tale che $M_t^2 - \langle M \rangle_t$ è ancora una martingala. Il processo $\langle M \rangle_t$ si può ottenere anche come limite (in probabilità) della variazione quadratica $V_{\Delta}^2(M)$ sull'intervallo $[0, t]$ quando $|\Delta| \rightarrow 0$. Vedi [9, §1.1.4].

s:wc

1.A. Cenni sulla convergenza debole di misure di probabilità

Consideriamo (S, d) spazio metrico equipaggiato con la sua σ -algebra dei Boreliani $\mathcal{B}(S)$, ovvero la minima σ -algebra che contiene tutti gli aperti. Indichiamo con $C(S)$ le funzioni continue da S a \mathbb{R} e con $\mathcal{P}(S)$ l'insieme delle misure di probabilità su $(S, \mathcal{B}(S))$.

t:wc

Definizione 1.22. *Data una successione $P_n \in \mathcal{P}(S)$ e $P \in \mathcal{P}(S)$, diciamo che P_n converge debolmente a P , $P_n \Rightarrow P$, sse*

$$\lim_{n \rightarrow \infty} \int f dP_n = \int f dP$$

per ogni $f \in C(S)$ limitata.

Osserviamo che se $\varphi : S \rightarrow S'$ (dove S' è un altro spazio metrico) è un'applicazione continua e $P_n \Rightarrow P$ si ha che $P_n \circ \varphi^{-1} \Rightarrow P \circ \varphi^{-1}$.

Il seguente teorema segue facilmente dalla definizione, vedi [2, Thm. 2.1].

t:wce **Teorema 1.23.** *Le seguenti affermazioni sono equivalenti:*

(i) $P_n \Rightarrow P$

(ii) $\int dP_n f \rightarrow \int dP f$ per ogni f uniformemente continua e limitata

(iii) $\overline{\lim}_n P_n(C) \leq P(C)$ per ogni C chiuso

(iv) $\underline{\lim}_n P_n(A) \geq P(A)$ per ogni A aperto

(v) $\lim_n P_n(B) = P(B)$ per ogni B tale che $P(\partial B) = 0$ (ove ∂B è la frontiera di B)

Nel caso in cui $S = \mathbb{R}$ possiamo caratterizzare la convergenza debole via la *funzione caratteristica* (= trasformata di Fourier della misura). Data μ , misura di probabilità su \mathbb{R} introduciamo

$$\phi(\lambda) := \int \mu(dx) e^{i\lambda x}$$

t:wcf **Teorema 1.24.** *Siano $\{\mu_n\}$ e μ misure di probabilità su \mathbb{R} con funzioni caratteristiche $\phi_n(\lambda)$ e $\phi(\lambda)$. Allora $\mu_n \Rightarrow \mu$ sse $\phi_n(\lambda) \rightarrow \phi(\lambda)$ puntualmente in λ .*

Vedi e.g. [2, Thm. 7.6] per la dimostrazione.

t:cet **Definizione 1.25.** *Sia $\Pi \subset \mathcal{P}(S)$ una famiglia di probabilità su S . Diciamo che Π è (sequenzialmente) relativamente compatta sse per ogni successione in Π possiamo estrarre una sottosuccessione debolmente convergente. Diciamo inoltre che Π è tight sse per ogni $\varepsilon > 0$ esiste un compatto K tale che $P(K) > 1 - \varepsilon$ per ogni $P \in \Pi$.*

Il seguente teorema di Prohorov fornisce una condizione necessaria e sufficiente per la compattezza di Π , vedi [2, Thm. 6.1,6.2]

t:prohorov **Teorema 1.26.** *Se Π è tight allora Π è relativamente compatta. Se inoltre S è completo e separabile vale anche il viceversa, ovvero se Π è relativamente compatta allora Π è tight.*

Tightness in C .

Poiché i compatti di C sono caratterizzati dal teorema di Ascoli–Arzelà, l'applicazione del teorema di Prohorov al caso di C è quasi immediata. Per $\delta \in (0, 1]$, il modulo di continuità di $x \in C$ è definito da

$$w_x(\delta) := \sup_{|s-t|<\delta} |x(t) - x(s)| \tag{1.13} \quad \boxed{\text{mc}}$$

Il teorema di Ascoli–Arzelà stabilisce che un insieme $\mathcal{A} \subset C$ è a chiusura compatta sse $\sup_{x \in \mathcal{A}} |x(0)| < \infty$ e $\lim_{\delta \rightarrow 0} \sup_{x \in \mathcal{A}} w_x(\delta) = 0$. Basta infatti osservare che, prendendo k abbastanza grande in modo che $\sup_{x \in \mathcal{A}} w_x(1/k) < \infty$ risulta

$$|x(t)| \leq |x(0)| + \sum_{i=1}^k \left| x\left(\frac{i}{k}t\right) - x\left(\frac{i-1}{k}t\right) \right| \leq |x(0)| + kw_x(1/k)$$

gli elementi di \mathcal{A} sono pertanto equilimitati e equicontinui.

t:cC **Teorema 1.27.** *La famiglia $\Pi \subset \mathcal{P}(C)$ è tight sse valgono le seguenti condizioni*

$$(i) \lim_{a \rightarrow \infty} \sup_{P \in \Pi} P(x : |x(0)| > a) = 0$$

$$(ii) \text{ per ogni } \varepsilon > 0 \quad \lim_{\delta \downarrow 0} \sup_{P \in \Pi} P(x : w_x(\delta) \geq \varepsilon) = 0$$

Dimostrazione. Dimostriamo solo che (i) e (ii) implicano che Π è tight. Dato $\eta > 0$ scegliamo a tale che, detto $A_0 := \{x : |x(0)| < a\}$, si abbia $P(A_0) \geq 1 - \eta/2$ per ogni $P \in \Pi$. Per $k \geq 1$, scegliamo inoltre δ_k tale che, detto $A_k := \{x : w_x(\delta_k) < 1/k\}$ si abbia $P(A_k) \geq 1 - 2^{-(k+1)}\eta$ per ogni $P \in \Pi$. Sia K la chiusura di $\bigcap_{k=0}^{\infty} A_k$, che risulta compatto per il teorema di Ascoli–Arzelà. Poichè $P(x : x \notin K) \leq \eta \sum_{k=0}^{\infty} 2^{-(k+1)} = \eta$ la famiglia Π è tight. \square

s:grr

1.B. Diseguaglianza di Garsia, Rodemich e Rumsey

Il seguente risultato (funzioni di una variabile reale, nulla di probabilità) va inteso nello stesso senso delle immersioni di Sobolev: controlla una norma puntuale in termini di una integrale.

t:grr

Teorema 1.28. *Siano q e Ψ funzioni continue su $[0, \infty)$ strettamente crescenti tali che $q(0) = \Psi(0) = 0$ e $\lim_{t \rightarrow \infty} \Psi(t) = \infty$. Se*

$$\int_0^1 dt \int_0^1 ds \Psi\left(\frac{|x(t) - x(s)|}{q(|t-s|)}\right) \leq B$$

allora, per ogni $0 \leq s < t \leq 1$

$$|x(t) - x(s)| \leq 8 \int_0^{t-s} q(du) \Psi^{-1}\left(\frac{4B}{u^2}\right) \tag{1.14}$$

grr

Noi useremo la questa diseguaglianza scegliendo per q e Ψ delle potenze. La maggiore flessibilità dell’enunciato generale è comunque utile in qualche applicazione.

Dimostrazione. Vediamo prima che è sufficiente dimostrare (1.14) per $s = 0$ e $t = 1$. Ponendo infatti

$$\bar{x}(u) = x(s + (t-s)u) \quad \bar{q}(u) = q((t-s)u)$$

abbiamo

$$\int_0^1 du \int_0^1 dv \Psi\left(\frac{|\bar{x}(u) - \bar{x}(v)|}{\bar{q}(|u-v|)}\right) = \frac{1}{(t-s)^2} \int_s^t du \int_s^t dv \Psi\left(\frac{|x(u) - x(v)|}{q(|u-v|)}\right) \leq \frac{B}{(t-s)^2} =: \bar{B}$$

Utilizzando (1.14) (per $s = 0$ e $t = 1$) per \bar{x} e \bar{q} troviamo

$$|x(t) - x(s)| = |\bar{x}(1) - \bar{x}(0)| \leq \int_0^1 \bar{q}(du) \Psi^{-1}\left(\frac{4\bar{B}}{u^2}\right) = \int_0^{t-s} q(du) \Psi^{-1}\left(\frac{4B}{u^2}\right)$$

ovvero (1.14).

Sia

$$I(t) := \int_0^1 ds \Psi\left(\frac{|x(t) - x(s)|}{q(|t - s|)}\right) \quad (1.15) \quad \boxed{\text{g0}}$$

poiché $\int_0^1 dt I(t) \leq B$ esiste $t_0 \in (0, 1)$ tale che $I(t_0) \leq B$. Definamo ora ricorsivamente una successione decrescente $t_n \in [0, t_0]$, $t_n \downarrow 0$ al modo seguente.

Sia d_0 tale che $q(d_0) = q(t_0)/2$ (notiamo che $d_0 < t_0$), affermiamo che esiste $t_1 \in (0, d_0)$ per cui

$$\begin{aligned} I(t_1) &\leq \frac{2B}{d_0} \\ \Psi\left(\frac{|x(t_1) - x(t_0)|}{q(|t_1 - t_0|)}\right) &\leq \frac{2I(t_0)}{d_0} \end{aligned} \quad (1.16)$$

Sia λ_{d_0} la misura di Lebesgue su $(0, d_0)$; abbiamo che

$$\lambda_{d_0}\left(t : I(t) > \frac{2B}{d_0}\right) < \frac{d_0}{2} \quad (1.17)$$

$$\lambda_{d_0}\left(t : \Psi\left(\frac{|x(t) - x(t_0)|}{q(|t - t_0|)}\right) > \frac{2I(t_0)}{d_0}\right) < \frac{d_0}{2} \quad (1.18)$$

se infatti (1.17) fosse falsa avremmo $\int_0^{d_0} dt I(t) > B$ e se fosse falsa (1.18) avremmo

$$\int_0^{d_0} ds \Psi\left(\frac{|x(s) - x(t_0)|}{q(|s - t_0|)}\right) > I(t_0)$$

che invaliderebbe $I(t_0) \leq B$. Da (1.17)–(1.18) segue che esiste $t_1 \in (0, d_0)$ per cui vale (1.16).

In generale dato t_{n-1} definiamo d_{n-1} tale che $q(d_{n-1}) = q(t_{n-1})/2$ e scegliamo $t_n \in (0, d_{n-1})$ in modo che

$$\begin{aligned} I(t_n) &\leq \frac{2B}{d_{n-1}} \\ \Psi\left(\frac{|x(t_n) - x(t_{n-1})|}{q(|t_n - t_{n-1}|)}\right) &\leq \frac{2I(t_{n-1})}{d_{n-1}} \end{aligned} \quad (1.19)$$

l'esistenza di un tale t_n si mostra come per t_1 . Poiché

$$2q(d_{n+1}) = q(t_{n+1}) \leq q(d_n) = q(t_n)/2$$

segue che $t_n \downarrow 0$.

Abbiamo ora

$$q(t_n - t_{n+1}) \leq q(t_n) = 2q(d_n) = 4[q(d_n) - q(d_n)/2] \leq 4[q(d_n) - q(d_{n+1})] \quad (1.20) \quad \boxed{\text{g4}}$$

che, insieme a (1.19), implica

$$\begin{aligned} |x(t_n) - x(t_{n-1})| &\leq \Psi^{-1}\left(\frac{2I(t_{n-1})}{d_{n-1}}\right) q(|t_n - t_{n+1}|) \leq \Psi^{-1}\left(\frac{4B}{d_n d_{n-1}}\right) 4[q(d_n) - q(d_{n+1})] \\ &\leq 4\Psi^{-1}\left(\frac{4B}{d_n^2}\right) [q(d_n) - q(d_{n+1})] \leq 4 \int_{d_{n+1}}^{d_n} q(du) \Psi^{-1}\left(\frac{4B}{u^2}\right) \end{aligned}$$

Sommando su n otteniamo

$$|x(t_0) - x(0)| \leq 4 \int_0^1 q(du) \Psi^{-1}\left(\frac{4B}{u^2}\right)$$

Ripetendo l'argomento per $x(1-t)$ troviamo anche

$$|x(1) - x(t_0)| \leq 4 \int_0^1 q(du) \Psi^{-1}\left(\frac{4B}{u^2}\right)$$

che sommata alla precedente conclude la dimostrazione. \square

$\boxed{\text{t:cgr}}$

Corollario 1.29. *Sia \mathbb{P} una probabilità du C per cui esistono $\alpha, r > 0$ e $C_0 < \infty$ tali che per ogni $s, t \in [0, 1]$*

$$\mathbb{E}(|x(t) - x(s)|^r) \leq C_0 |t - s|^{1+\alpha}$$

Allora per ogni $\beta \in (0, \alpha/r)$ esiste una costante $C = C(C_0, \beta, \alpha, r) < \infty$ (indipendente dalla probabilità \mathbb{P}) per cui

$$\mathbb{E}\left(\sup_{0 \leq s < t \leq 1} \frac{|x(t) - x(s)|}{|t - s|^\beta}\right)^r \leq C$$

Dimostrazione. Scegliamo $\Psi(u) = u^r$ e $q(u) = u^{\gamma/r}$, con $\gamma \in (2, 2 + \alpha)$ nel Teorema 1.28 ponendo

$$B(x) := \int_0^1 dt \int_0^1 ds \left(\frac{|x(t) - x(s)|}{|t - s|^{\gamma/r}}\right)^r$$

Notiamo infatti che in virtù dell'ipotesi risulta

$$\mathbb{E}(B(x)) \leq \int_0^1 dt \int_0^1 ds C_0 |t - s|^{1+\alpha-\gamma} =: C_1$$

purché $\gamma < 2 + \alpha$.

Troviamo

$$|x(t) - x(s)| \leq 8 \int_0^{|t-s|} \left(\frac{4B(x)}{u^2}\right)^{1/r} du^{\gamma/r} = \frac{8\gamma}{\gamma-2} (4B(x))^{1/r} |t-s|^\beta =: C_2 (B(x))^{1/r} |t-s|^\beta$$

s:cp purché $\beta = (\gamma - 2)/r > 0$. Prendendo il valore di attesa rispetto a \mathbb{P} il risultato segue. \square

1.C. Cenni sulle probabilità condizionate

Dato uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, una sotto σ -algebra $\mathcal{G} \subset \mathcal{F}$ e una variabile aleatoria $X \in L_1(\mathbb{P})$, l'attesa condizionata di X data \mathcal{G} , che indichiamo con $\mathbb{E}(X|\mathcal{G})$, è la variabile aleatoria \mathcal{G} misurabile (definita \mathbb{P} -a.s.) tale che

$$\int_G X(\omega) d\mathbb{P}(\omega) = \int_G \mathbb{E}(X|\mathcal{G})(\omega) d\mathbb{P}(\omega) \quad \text{per ogni } G \in \mathcal{G}$$

L'esistenza dell'attesa condizionata segue dal teorema di Radon–Nikodym (perché?). Nel caso in cui $X \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ si verifica immediatamente (farlo!) che $E(X|\mathcal{G})$ è la proiezione di X sul sottospazio (chiuso) $L_2(\Omega, \mathcal{G}, \mathbb{P}) \subset L_2(\Omega, \mathcal{F}, \mathbb{P})$.

Se Ω è un insieme finito ad ogni σ -algebra è associata una partizione $\mathcal{D} = \{D_i\}$ di Ω ed, in tal caso, assumendo inoltre $X = \mathbb{1}_B$ per un qualche $B \in \mathcal{F}$, si ha

$$\mathbb{E}(\mathbb{1}_B|\mathcal{D})(\omega) = \sum_i \mathbb{P}(B|D_i) I_{D_i}(\omega)$$

dove, se $\mathbb{P}(D_i) > 0$, $\mathbb{P}(B|D_i) := \mathbb{P}(B \cap D_i) / \mathbb{P}(D_i)$ è l'usuale probabilità condizionata. Vediamo quindi che l'attesa condizionata è costante sugli atomi di \mathcal{D} (condizione equivalente alla \mathcal{D} -misurabilità).

Nel caso in cui $\mathcal{G} = \sigma\{Y\}$ per una qualche v.a. Y , scriviamo $\mathbb{E}(X|Y)$ per $\mathbb{E}(X|\sigma\{Y\})$. In tal caso si dimostra esistere una funzione boreliana $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ tale che

$$\mathbb{E}(X|Y)(\omega) = \varphi(Y(\omega)) \quad \mathbb{P}\text{-a.s.}$$

La seguente proposizione riassume alcune utili proprietà dell'attesa condizionata, la dimostrazione segue direttamente dalla definizione data.

t:cex **Proposizione 1.30.** *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità, $\mathcal{G} \subset \mathcal{F}$ uno sotto σ -algebra di \mathcal{F} e X una v.a. in $L_1(\Omega, \mathcal{F}, \mathbb{P})$; allora:*

1. $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$;
2. Se X è \mathcal{G} misurabile allora $\mathbb{E}(X|\mathcal{G}) = X$;
3. Se X è \mathcal{G} misurabile e Y è una v.a. arbitraria (con le necessarie condizioni di integrabilità) allora $\mathbb{E}(XY|\mathcal{G}) = X \mathbb{E}(Y|\mathcal{G})$ (funzioni \mathcal{G} misurabili escono gratis dall'attesa condizionata);
4. Se X è indipendente dalla σ -algebra \mathcal{G} (ovvero $\mathbb{E}(X\mathbb{1}_G) = \mathbb{E}(X)\mathbb{P}(G)$ per ogni $G \in \mathcal{G}$) allora $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$;
5. Se $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$ allora

$$\mathbb{E}(X|\mathcal{G}_1) = \mathbb{E}(\mathbb{E}(X|\mathcal{G}_2)|\mathcal{G}_1)$$

possiamo cioè inserire condizionamenti rispetto a σ -algrebre più fini.

Note bibliografiche

Sulla “teoria fisica” del moto browniano si consigliano gli articoli originali di Einstein, vedi [5] per la traduzione inglese. Come motivazione è anche molto consigliata la lettura di [11]. La costruzione della misura di Wiener a partire da una passeggiata aleatoria si può trovare, per esempio, in [9, §2.2.4] o in [12, Thm. XIII.1.9]. È però consigliato [2, Thm. 10.1]; in tale libro si può anche trovare la dimostrazione del teorema di Prohorov e una discussione dettagliata della convergenza debole e criteri di compatezza. Applicazioni del teorema 1.3 per il calcolo di distribuzioni di funzioni del moto browniano (nello stesso spirito di teorema 1.8) sono discusse in [2, §11]. La dimostrazione di Proposizione 1.7 è tratta da [13], si consiglia la lettura del capitolo “introduttivo” per vari problemi non banali (quasi tutti oltre gli argomenti trattati in questo corso) collegati al moto browniano. Per la costruzione del valore di attesa condizionato a partire da un proiettore in L_2 (senza utilizzare il teorema di Radon–Nikodym) vedi [17, Ch. 9]. Come detto prima, [17] è anche altamente consigliato per un’introduzione alla teoria delle martingale (a tempo discreto). Per una teoria generale delle martingale a tempo continuo, vedi [9, Ch. 1] o [12, Ch. II]. La diseguaglianza di Garsia, Rodemich e Rumsey (con alcune applicazioni) si può trovare in [15, Thm. 2.1.3].

s:1

s:ii

2. Equazioni differenziali stocastiche

2.1. Integrale di Ito

Vogliamo definire un integrale rispetto al moto browniano. Poichè il moto browniano ha variazione infinita, non è l'usuale integrale di Lebesgue–Stieltjes. Vediamo in un esempio quale sia il problema. Supponiamo di voler definire $\int_0^t w_s dw_s$ (come funzione di t e ω , ma ommettiamo scrivere esplicitamente la dipendenza da ω). Presa $0 = t_0 < t_1 < \dots < t_n = t$ una partizione di $[0, t]$ e $\vartheta \in [0, 1]$ guardiamo

$$\sum_i w(t_i + \vartheta[t_{i+1} - t_i]) [w(t_{i+1}) - w(t_i)]$$

e proviamo a calcolare il valore di attesa. Poiché $\mathbb{E}(w_t w_s) = \min\{t, s\} =: t \wedge s$ troviamo

$$\mathbb{E} \sum_i w(t_i + \vartheta[t_{i+1} - t_i]) [w(t_{i+1}) - w(t_i)] = \sum_i [t_i + \vartheta[t_{i+1} - t_i] - t_i] = \vartheta t$$

Il risultato dipende da ϑ . Come (sperabilmente) sarà chiaro nel seguito la scelta più naturale è di prendere $\vartheta = 0$ che dà luogo all'integrale di Ito che risulta una martingala. Altre definizioni (meno convenienti) sono però possibili.

Dato il solito spazio di probabilità filtrato $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, in cui supponiamo che la filtrazione \mathcal{F}_t soddisfi la condizione usuale (continua da destra e completa) ed un moto browniano w_t , cominciamo con il definire la classe degli “integrandi” $\varphi_t := \varphi(t, \omega)$ per $\int_0^t \varphi_s dw_s$. Ricordiamo che il processo (= famiglia di variabili aleatorie parametrizzate da t) φ_t è *progressivamente misurabile* rispetto a \mathcal{F}_t sse per ogni $t \in \mathbb{R}_+$ l'applicazione $[0, t] \times \Omega \ni (s, \omega) \mapsto \varphi_s(\omega)$ è $\mathcal{B}([0, t]) \times \mathcal{F}_t$ misurabile. Intuitivamente richiediamo che la traiettoria $\{\varphi_s, s \in [0, t]\}$ sia determinata dalla conoscenza di $\{w_s, s \in [0, t]\}$ (pensiamo che la filtrazione \mathcal{F}_t sia quella generata dal browniano). Dato $T > 0$ introduciamo inoltre la norma

$$\|\varphi\|_T^2 := \int_0^T dt \mathbb{E}(\varphi_t^2) \tag{2.1} \quad \text{tn}$$

la classe degli integrandi φ sarà data da

$$\mathcal{H}_T := \{\varphi : [0, T] \times \Omega \rightarrow \mathbb{R} : \varphi \text{ è progr. mis. rispetto a } \mathcal{F}_t, \|\varphi\|_T < \infty\} \tag{2.2} \quad \text{hT}$$

che risulta un sottospazio chiuso di $L_2([0, T] \times \Omega, \mathcal{B}([0, T]) \times \mathcal{F}_\infty, dt \times d\mathbb{P})$.

Sia $\mathcal{S} \subset \mathcal{H}_T$ il sottoinsieme dei processi (progressivamente misurabili) semplici, ovvero $\varphi \in \mathcal{S}$ se esiste una partizione $0 = t_0 < t_1 < \dots < t_n = T$ di $[0, T]$ e variabili aleatorie limitate $\xi_k \in \mathcal{F}_{t_k}$, $k = 0, \dots, n-1$ per cui

$$\varphi_t(\omega) = \xi_0 \mathbb{I}_{\{0\}}(t) + \sum_{k=0}^{n-1} \xi_k(\omega) \mathbb{I}_{(t_k, t_{k+1}]}(t)$$

con un po' di pazienza si mostra che \mathcal{S} è denso in \mathcal{H}_T , vedi [9, §3.2.A]. Anzi nel presente contesto (integrale rispetto al moto browniano) qualche sforzo in più mostra come sia sufficiente richiedere che φ sia adattato (ovvero $\varphi_t \in \mathcal{F}_t$ per ogni $t \in \mathbb{R}_+$) invece che progressivamente misurabile.

Dato $\varphi \in \mathcal{S}$ definiamo l'integrale di Ito (rispetto al moto browniano w_t) come il processo $I_t(\varphi)$, $t \in [0, T]$ dato da

$$\begin{aligned} I_t(\varphi) &:= \sum_{i=0}^{i_*(t)-1} \xi_i [w_{t_{i+1}} - w_{t_i}] + \xi_{i_*(t)} [w_t - w_{t_{i_*(t)}}] \\ &= \sum_{i=0}^{i_*(t)-1} \varphi_{t_i} [w_{t_{i+1}} - w_{t_i}] + \varphi_{t_{i_*(t)}} [w_t - w_{t_{i_*(t)}}] \end{aligned} \quad (2.3)$$

dove $i_*(t)$ è tale che $t_{i_*(t)} \leq t < t_{i_*(t)+1}$. Si può trarre giovamento nel confrontare questa definizione con l'esempio di martingala della strategia di gioco alla roulette, eq.ne (1.8).

La seguente Proposizione riassume le proprietà di $I_t(\varphi)$.

t:iis **Proposizione 2.1.** *Il processo $I_t(\varphi)$ soddisfa le seguenti proprietà:*

- (i) $I_t(\varphi)$ è progressivamente misurabile e $t \rightarrow I_t(\varphi)$ è continuo \mathbb{P} -q.s.;
- (ii) $I_t(\varphi)$ è lineare, i.e. per ogni $\alpha, \beta \in \mathbb{R}$ e $\varphi, \psi \in \mathcal{S}$ si ha $I_t(\alpha\varphi + \beta\psi) = \alpha I_t(\varphi) + \beta I_t(\psi)$;
- (iii) per $0 \leq s \leq t \leq T$ abbiamo

$$\begin{aligned} \mathbb{E}(I_t(\varphi) | \mathcal{F}_s) &= I_s(\varphi) \\ \mathbb{E}\left(I_t(\varphi)^2 - \int_0^t du \varphi_u^2 \middle| \mathcal{F}_s\right) &= I_s(\varphi)^2 - \int_0^s du \varphi_u^2 \end{aligned} \quad (2.4)$$

ovvero $I_t(\varphi)$ è una martingala con variazione quadratica $\int_0^t ds \varphi_s^2$;

- (iv) per ogni $\lambda \in \mathbb{R}_+$

$$\mathbb{P}\left(\sup_{t \in [0, T]} |I_t(\varphi)| \geq \lambda\right) \leq \frac{1}{\lambda^2} \|\varphi\|_T^2 \quad (2.5) \quad \text{diis}$$

Osservazione Da (iii) segue in particolare che

$$\mathbb{E}(I_t(\varphi)^2) = \int_0^t ds \mathbb{E}(\varphi_s^2) \quad t \in [0, T]$$

che mostra come l'applicazione $\varphi \mapsto I_T(\varphi)$ sia una isometria (definita su un insieme denso) dallo spazio (di Hilbert) \mathcal{H}_T a $L_2(\Omega, \mathbb{P})$. Grazie a tale osservazione possiamo immediatamente definire $I_T(\varphi)$ per ogni $\varphi \in \mathcal{H}_T$. Vogliamo però considerare $I_t(\varphi)$ come un processo (i.e. come funzione di ω e t) e dire, ad esempio, che $t \mapsto I_t(\varphi)$ è \mathbb{P} -a.s. continuo per ogni $\varphi \in \mathcal{H}_T$. Nel fare questo utilizzeremo anche la proprietà (iv).

Dimostrazione. Le proprietà (i) e (ii) seguono immediatamente dalla definizione; notiamo infatti che $t \mapsto I_t(\varphi)$ è continuo per gli $\omega \in \Omega$ per cui $t \mapsto w_t$ è continuo. Dimostriamo (iii). Dalla definizione (2.3) segue

$$I_t(\varphi) - I_s(\varphi) = \varphi_{t_{i_*(s)}} \left[w_{t_{i_*(s)+1}} - w_s \right] + \sum_{i=i_*(s)+1}^{i_*(t)-1} \varphi_{t_i} \left[w_{t_{i+1}} - w_{t_i} \right] + \varphi_{t_{i_*(t)}} \left[w_t - w_{t_{i_*(t)}} \right]$$

Poichè $\varphi_{t_i} \in \mathcal{F}_{t_i}$ e per $u \geq s$ $\mathbb{E}(f|\mathcal{F}_s) = \mathbb{E}[\mathbb{E}(f|\mathcal{F}_u)|\mathcal{F}_s]$ troviamo

$$\begin{aligned} \mathbb{E}(I_t(\varphi) - I_s(\varphi) | \mathcal{F}_s) &= \mathbb{E} \left(\varphi_{t_{i_*(s)}} \left[w_{t_{i_*(s)+1}} - w_s \right] \middle| \mathcal{F}_s \right) \\ &+ \sum_{i=i_*(s)+1}^{i_*(t)-1} \mathbb{E} \left(\varphi_{t_i} \mathbb{E} \left([w_{t_{i+1}} - w_{t_i}] \middle| \mathcal{F}_{t_{i-1}} \right) \middle| \mathcal{F}_s \right) \\ &+ \mathbb{E} \left(\varphi_{t_{i_*(t)}} \mathbb{E} \left([w_t - w_{t_{i_*(t)}}] \middle| \mathcal{F}_{t_{i_*(t)}} \right) \middle| \mathcal{F}_s \right) = 0 \end{aligned}$$

La seconda proprietà di martingala in (2.4) si dimostra, con un po' di algebra in più, allo stesso modo. La stima (2.5) segue applicando la diseguaglianza di Doob (teorema 1.18) alla martingala $I_t(\varphi)$. \square

Siamo ora in grado di definire $I_t(\varphi)$ per ogni $\varphi \in \mathcal{H}_T$ in modo che valgano ancora le proprietà della proposizione 2.1. Ricordiamo il seguente risultato (lemma di Borel–Cantelli) che segue immediatamente dalle proprietà di continuità delle misure di probabilità.

t:bc **Lemma 2.2.** *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità e $A_n \in \mathcal{F}$, $n = 1, \dots$ una successione di eventi tali che*

$$\sum_n \mathbb{P}(A_n) < \infty$$

allora

$$\mathbb{P}(A_n \text{ i.o.}) = \mathbb{P} \left(\bigcap_n \bigcup_{k \geq n} A_k \right) = 0$$

Definizione integrale di Ito.

Data $\varphi \in \mathcal{H}_T$ scegliamo una successione $\varphi^{(n)} \in \mathcal{S}$ tale che $\|\varphi - \varphi^{(n)}\|_T \leq n^{-3}$ allora per (ii) e (iv) in Proposizione 2.1 abbiamo

$$\mathbb{P} \left(\sup_{t \in [0, T]} |I_t(\varphi^{(n)}) - I_t(\varphi^{(n+1)})| \geq \frac{1}{n^2} \right) \leq n^4 \left(\frac{2}{n^3} \right)^2$$

applicando il lemma di Borel–Cantelli, detto A_n l'evento a primo membro sopra, concludiamo che posto $N = \{A_n \text{ i.o.}\}$ vale $\mathbb{P}(N) = 0$. D'altra parte per $\omega \in \Omega \setminus N$ abbiamo che esiste un $n_0 = n_0(\omega)$ per cui $\sup_{t \in [0, T]} |I_t(\varphi^{(n+1)}) - I_t(\varphi^{(n)})| < 1/n^2$ per ogni $n > n_0$. In particolare se $\omega \in \Omega \setminus N$ si ha che $I_t(\varphi^{(n)})(\omega)$ converge uniformemente per $t \in [0, T]$. Possiamo quindi definire

$$I_t(\varphi)(\omega) := \begin{cases} \lim_n I_t(\varphi^{(n)})(\omega) & \text{se } \omega \in \Omega \setminus N \\ 0 & \text{se } \omega \in N \end{cases}$$

poiché \mathcal{F}_t contiene gli eventi \mathbb{P} -nulli per \mathcal{F}_∞ , $I_t(\varphi)$ è progressivamente misurabile. Per costruzione risulta inoltre che $t \rightarrow I_t(\varphi)$ è \mathbb{P} -a.s. continua (come limite uniforme di funzioni continue). È anche facile vedere che $I_t(\varphi)$ non dipende dalla successione scelta per approssimare φ .

Osserviamo infine che, per costruzione, per ogni $t \in [0, T]$ fissato $I_t(\varphi^{(n)}) \rightarrow I_t(\varphi)$ in $L_2(\mathbb{P})$, in particolare ciò ci garantisce che $\mathbb{E}(I_t(\varphi)^2) < \infty$. La convergenza in L_2 ci permette di passare al limite nelle relazioni di martingala (2.4) ed affermare che $I_t(\varphi)$ è una martingala con variazione quadratica $\int_0^t ds \varphi_s^2$. Vediamo i dettagli.

Abbiamo la successione $\varphi^{(n)}$ che approssima φ nella norma $\|\cdot\|_T$. La prima relazione di martingala in (2.4) equivale a dire che per ogni $0 \leq s < t \leq T$ ed ogni evento $G \in \mathcal{F}_s$ si ha

$$\mathbb{E}\left(\mathbb{1}_G [I_t(\varphi^{(n)}) - I_s(\varphi^{(n)})]\right) = 0 \quad (2.6) \quad \boxed{\text{1Im}}$$

Vogliamo ora passare al limite per $n \rightarrow \infty$ nella precedente identità e concludere che $I_t(\varphi)$ è una martingala. Poiché $I_t(\varphi^{(n)})$ (con $t \in [0, T]$ fissato) converge a $I_t(\varphi)$ in $L_2(\Omega, d\mathbb{P})$ converge anche in $L_1(\Omega, d\mathbb{P})$ e possiamo passare al limite in (2.6).

Allo stesso modo la seconda relazione di martingala in (2.4) equivale a dire che per ogni $0 \leq s < t \leq T$ ed ogni evento $G \in \mathcal{F}_s$ si ha

$$\mathbb{E}\left(\mathbb{1}_G \left\{ [I_t(\varphi^{(n)}) - I_s(\varphi^{(n)})]^2 - \int_s^t du [\varphi_u^{(n)}]^2 \right\}\right) = 0 \quad (2.7) \quad \boxed{\text{1IIm}}$$

Poiché $I_t(\varphi^{(n)})$ converge a $I_t(\varphi)$ in $L_2(\Omega, d\mathbb{P})$ e $\mathbb{E} \int_s^t du [\varphi_u^{(n)} - \varphi_u]^2 \leq \|\varphi^{(n)} - \varphi\|_T^2$ che converge a zero per $n \rightarrow \infty$, è immediato passare al limite in (2.7) e concludere che la variazione quadratica di $I_t(\varphi)$ è $\int_0^t ds \varphi_s^2$.

La costruzione precedente conclude la costruzione dell'integrale di Ito per processi in \mathcal{H}_T . Questo non è completamente soddisfacente poiché la condizione $\|\varphi\|_T < \infty$ impone delle condizioni di integrabilità (rispetto a ω) non naturali; in particolare il processo $\varphi_t = \exp\{w_t^4\}$ non è in \mathcal{H}_T e non si capisce perché non possiamo integrarlo. Vediamo quindi come si possa allargare la classe degli integrandi ai processi φ_t progressivamente misurabili tali che

$$\mathbb{P}\left(\int_0^T dt \varphi_t^2 < \infty\right) = 1 \quad (2.8) \quad \boxed{\text{ipg}}$$

Osserviamo che per $\varphi \in \mathcal{H}_T$ vale la seguente stima. Per ogni $\lambda, K \in (0, \infty)$

$$\mathbb{P}\left(\sup_{t \in [0, T]} \left| \int_0^t \varphi_s dw_s \right| > \lambda\right) \leq \frac{K}{\lambda^2} + \mathbb{P}\left(\int_0^T dt \varphi_t^2 > K\right) \quad (2.9) \quad \boxed{\text{dd1}}$$

Definendo

$$\varphi_t^{(K)} := \varphi_t \mathbb{1}_{[0, K]} \left(\int_0^t ds \varphi_s^2 \right)$$

abbiamo $\varphi^{(K)} \in \mathcal{H}_T$ e $\|\varphi^{(K)}\|_T^2 \leq K$. Esercizio: verificare tale affermazione (sugg. sia $\tau := \inf\{t \geq 0 : \int_0^t ds \varphi_s^2 > K\}$). Osserviamo che se $\omega \in A_K := \{\tilde{\omega} : \int_0^T dt \varphi_t^2(\tilde{\omega}) \leq K\}$ si

ha $\varphi_t^{(K)}(\omega) = \varphi_t(\omega)$, $\forall t \in [0, T]$. Pertanto

$$\begin{aligned} \mathbb{P}\left(\sup_{t \in [0, T]} \left| \int_0^t \varphi_s dw_s \right| > \lambda\right) &\leq \mathbb{P}\left(\sup_{t \in [0, T]} \left| \int_0^t \varphi_s dw_s \right| > \lambda, A_K\right) + \mathbb{P}(A_K^c) \\ &\leq \mathbb{P}\left(\sup_{t \in [0, T]} \left| \int_0^t \varphi_s^{(K)} dw_s \right| > \lambda\right) + \mathbb{P}(A_K^c) \leq \frac{\|\varphi^{(K)}\|_T^2}{\lambda^2} + \mathbb{P}(A_K^c) \end{aligned}$$

che conclude la dimostrazione di (2.9).

Per costruire l'integrale di Ito di un processo φ_t progressivamente misurabile che soddisfa (2.8), costruiamo la successione

$$\varphi_t^{(n)} := \varphi_t \mathbb{1}_{[0, n]} \left(\int_0^t ds \varphi_s^2 \right)$$

Ovviamente $\varphi \in \mathcal{H}_T$ e da (2.8) segue immediatamente che per ogni $\delta > 0$

$$\lim_n \mathbb{P}\left(\int_0^T dt [\varphi_t - \varphi_t^{(n)}]^2 > \delta\right) = 0 \quad (2.10) \quad \boxed{\text{dd2}}$$

Possiamo quindi definire la successione di integrali stocastici $I_t(\varphi^{(n)}) = \int_0^t \varphi_t^{(n)} dw_t$ che dichiariamo essere, uniformemente per $t \in [0, T]$, di Cauchy rispetto alla convergenza in probabilità, ovvero per ogni $\varepsilon > 0$ si ha

$$\lim_{n, m \rightarrow \infty} \mathbb{P}\left(\sup_{t \in [0, T]} |I_t(\varphi^{(n)}) - I_t(\varphi^{(m)})| > \varepsilon\right) = 0$$

come si verifica immediatamente utilizzando la disuguaglianza (2.9) e (2.10). L'integrale stocastico $I_t(\varphi)$ è allora definito come il limite in probabilità, uniforme per $t \in [0, T]$, di $I_t(\varphi^{(n)})$. Tale limite risulta un processo continuo \mathbb{P} -a.s. (è la stessa dimostrazione fatta prima) ma non è necessariamente una martingala poiché nessuno ci garantisce la condizione di integrabilità $\mathbb{E} \left| \int_0^t \varphi_s dw_s \right| < \infty$.

D'altra parte, in qualche senso, $I_t(\varphi)$ soddisfa le condizione di martingalità anche nella sola ipotesi (2.8). Viene introdotto il concetto di *martingala locale* come segue.

Dato il solito spazio di probabilità filtrato $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ una variabile aleatoria $\tau : \Omega \rightarrow [0, +\infty]$ viene detta *tempo di arresto* sse per ogni $t \in \mathbb{R}_+$ si ha che l'evento $\{\tau \leq t\}$ è nella σ -algebra \mathcal{F}_t . In soldoni richiediamo che per sapere se τ è "scattato" prima di t ci basta guardare la porzione di traiettoria browniana in $[0, t]$ (pensando, come al solito, che \mathcal{F}_t è la filtrazione generata dal browniano). L'esempio tipico di tempo di arresto è $\tau := \inf\{t \geq 0 : X_t \in A\}$ per un A aperto in \mathbb{R} ed un processo X con traiettorie \mathbb{P} -q.s. continue adattato a \mathcal{F}_t (che soddisfa la condizione usuale). Vedi [9, § 1.2] per un prontuario sui tempi di arresto. Una *martingala locale* M_t è allora un processo \mathcal{F}_t adattato per cui esiste una successione di tempi di arresto $\tau_n \uparrow +\infty$ \mathbb{P} -q.s. per cui, detto

$$M_t^n(\omega) := M_{t \wedge \tau_n(\omega)}(\omega) = \begin{cases} M_t(\omega) & \text{se } t < \tau_n(\omega) \\ M_{\tau_n(\omega)}(\omega) & \text{se } t \geq \tau_n(\omega) \end{cases}$$

il *processo arrestato*, si ha che M_t^n risulta una martingala per ogni $n \in \mathbb{N}$.

Dalla precedente costruzione risulta ovvia che $I_t(\varphi)$ è una martingala locale. Basta dichiarare $\tau_n := \inf\{t \geq 0 : \int_0^t ds \varphi_s^2 > n\}$. La precedente definizione è più sottile di quello che sembra. Affinché una martingala locale sia una “vera” martingala non basta una condizione di integrabilità; vi sono infatti esempi di martingale locali uniformemente integrabili che non sono però delle martingale.

s:if

2.2. Formula di Ito

La formula di Ito è null’altro che il teorema fondamentale del calcolo per integrali stocastici ed è alla base dell’analisi stocastica. Ci diamo il solito spazio di probabilità filtrato $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ ed un moto browniano w su tale spazio. Supponiamo inoltre di avere un processo X_t dato da

$$X_t = X_0 + \int_0^t ds \psi_s + \int_0^t \varphi_s dw_s \quad (2.11) \quad \text{xi}$$

con $X_0 \in \mathbb{R}$, ψ_s e φ_s progressivamente misurabili tali che φ e $|\psi|^{1/2}$ soddisfano (2.8) per ogni $T > 0$. Un processo siffatto viene detto *semimartingala*, essenzialmente è un processo a variazione finita q.s. più una martingala. Simbolicamente scriviamo (2.11) nella forma differenziale (senza altro significato che una notazione)

$$dX_t = \psi_t dt + \varphi_t dw_t \quad (2.12) \quad \text{xid}$$

dalla costruzione precedente dell’integrale di Ito (per l’integrale di Lebesgue non ci sono problemi) è immediato verificare che X_t è progressivamente misurabile ed ha traiettorie continue \mathbb{P} -q.s. Il senso di (2.12) è il seguente se abbiamo un processo f_t progressivamente misurabile tale che per ogni $T > 0$

$$\mathbb{P}\left(\int_0^T dt [|f_t| |\psi_t| + f_t^2 \varphi_t^2] < \infty\right) = 1$$

allora definiamo l’integrale di f rispetto alla semimartingala X come

$$\int_0^t f_s dX_s := \int_0^t ds f_s \psi_s + \int_0^t f_s \varphi_s dw_s$$

ove il secondo termine è l’integrale di Ito (ben definito in virtù delle ipotesi fatte).

Data una funzione regolare f su \mathbb{R} vogliamo scoprire qual’è il “differenziale” di $f(X_t)$.

t:if

Teorema 2.3. Sia $f(t, x) \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$ (funzione C^1 in t , C^2 in x) e X dato in (2.11).

Allora

$$df(t, X_t) = \left[\partial_t f(t, X_t) + \partial_x f(t, X_t) \psi_t + \frac{1}{2} \partial_x^2 f(t, X_t) \varphi_t^2 \right] dt + \partial_x f(t, X_t) \varphi_t dw_t \quad (2.13) \quad \text{ifd}$$

con cui intendiamo semplicemente che, \mathbb{P} -q.s., vale

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t ds \left\{ \partial_s f(s, X_s) + \partial_x f(s, X_s) \psi_s + \frac{1}{2} \partial_x^2 f(s, X_s) \varphi_s^2 \right\} \\ &\quad + \int_0^t \partial_x f(s, X_s) \varphi_s dw_s \end{aligned} \quad (2.14) \quad \text{ifdint}$$

In altri termini, utilizzando la notazione (2.12)

$$df(t, X_t) = \partial_t f(t, X_t) dt + \partial_x f(t, X_t) dX_t + \frac{1}{2} \partial_x^2 f(t, X_t) d\langle X \rangle_t$$

in cui $\langle X \rangle_t := \int_0^t ds \varphi_s^2$ denota la variazione quadratica della martingala locale in (2.11).

Osserviamo come l'integrale stocastico in (2.14) sia ben definito in quanto il processo $\partial_x f(s, X_t) \varphi_t$ è progressivamente misurabile e soddisfa (2.8). Se invece avessimo costruito l'integrale stocastico solo per processi in \mathcal{H}_T , per poter scrivere la formula di Ito avremmo dovuto imporre delle condizioni di crescita a $f(t, x)$ per $x \rightarrow \infty$.

Il passo fondamentale per consiste nel dimostrare (2.13) quando ψ e φ in (2.12) sono costanti. Da ciò segue il risultato per $\psi, \varphi \in \mathcal{S}$ e, via approssimazione, il caso generale. Per semplicità di notazione consideriamo solo il caso f indipendente da t , il caso generale si ottiene comunque considerando combinazione lineare di funzioni prodotto. Cominciamo con un esempio particolare, che sarà l'ingrediente cruciale.

t:dwt **Lemma 2.4.** *Sia w_t il moto browniano, allora $d[w_t]^2 = 2w_t dw_t + dt$ ovvero*

$$\int_0^t w_s dw_s = \frac{1}{2} [w_t^2 - t]$$

Dimostrazione. Dato $t > 0$ e $0 = t_0 < t_1 < \dots < t_n = t$ una partizione di $[0, t]$, utilizzando l'identità $a(b - a) = (b^2 - a^2)/2 - (b - a)^2/2$ possiamo scrivere

$$\sum_{i=0}^{n-1} w_{t_i} [w_{t_{i+1}} - w_{t_i}] = \frac{1}{2} \sum_{i=0}^{n-1} [w_{t_{i+1}}^2 - w_{t_i}^2 - (w_{t_{i+1}} - w_{t_i})^2] \xrightarrow{n \rightarrow \infty} \frac{1}{2} [w_t^2 - t] \quad \text{in } L_2(\mathbb{P})$$

per la convergenza della variazione quadratica del moto browniano a t , vedi teorema 1.10. □

t:dwt **Lemma 2.5.** *Siano $X_t^i = b_i t + a_i w_t$, $i = 1, 2$ con a_i, b_i v.a. limitate e \mathcal{F}_0 misurabili. Allora*

$$d(X_t^1 X_t^2) = X_t^1 dX_t^2 + X_t^2 dX_t^1 + a_1 a_2 dt \tag{2.15} \quad \text{dpr}$$

con cui intendiamo semplicemente che

$$X_t^1 X_t^2 = X_0^1 X_0^2 + \int_0^t [X_s^1 dX_s^2 + X_s^2 dX_s^1 + a_1 a_2 ds]$$

Osservazione Per additività dell'integrale è immediato estendere (2.15) a processi semplici, $X^i \in \mathcal{S}$, ottenendo

$$d(X_t^1 X_t^2) = X_t^1 dX_t^2 + X_t^2 dX_t^1 + \int_0^t ds \varphi_s^1 \varphi_s^2$$

Per densità di \mathcal{S} in \mathcal{H}_T ed approssimazione di processi soddisfacenti (2.8), si ottiene infine il caso generale.

Dimostrazione. Per linearità è sufficiente dimostrare $dw_t^2 = 2w_t dw_t + dt$ (vedi lemma 2.4) e $d[tw_t] = w_t dt + t dw_t$. Considerando la solita partizione di $[0, t]$ ricaviamo subito

$$\begin{aligned} \int_0^t (sdw_s + w_s ds) &= \lim_n \sum_{i=0}^{n-1} \left[t_i (w_{t_{i+1}} - w_{t_i}) + w_{t_{i+1}} (t_{i+1} - t_i) \right] \\ &= \lim_n \sum_{i=0}^{n-1} [t_{i+1} w_{t_{i+1}} - t_i w_{t_i}] = tw_t \end{aligned}$$

che conclude la dimostrazione. □

Procedendo per induzione (esercizio) troviamo allora il seguente lemma.

Lemma 2.6. *Se $f = f(x)$ è un polinomio, allora l'identità (2.14) è vera.*

Dimostrazione teorema 2.3.

Supponiamo, per un attimo, che esista un compatto K di \mathbb{R} tale che $X_t(\omega) \in K$ per ogni $t \in [0, T]$ e \mathbb{P} -q.s. in ω . Grazie al teorema di Weierstrass sull'approssimazione, uniforme in compatti, di funzioni continue mediante polinomi, la dimostrazione dell'identità (2.14) segue subito dal Lemma 2.6 (basta approssimare, uniformemente in K , f'' con un polinomio). Per concludere la dimostrazione, vediamo come ci si può ricondurre, mediante *arresto*, al caso in cui X_t assume valori in un compatto.

Dato X_t come in (2.11) ed $n > 0$ introduciamo la variabile aleatoria

$$\tau_n := \inf\{t \geq 0 : |X_t| > n\}$$

ovvero il primo tempo in cui X_t esce dalla palla di raggio n (intendiamo $\tau_n = +\infty$ se $\sup_{t \geq 0} |X_t| \leq n$). Poichè X è progressivamente misurabile abbiamo che τ_n è un tempo di arresto. Inoltre, per la continuità \mathbb{P} -q.s. di X_t , abbiamo che $\tau_n \uparrow +\infty$ \mathbb{P} -q.s. per $n \rightarrow \infty$. Verifichiamo: poiché $\tau_{n+1} \geq \tau_n$ abbiamo

$$\{\omega : \tau_n(\omega) \uparrow +\infty\} = \bigcap_{k>0} \bigcup_{n_0>0} \bigcap_{n \geq n_0} \{\omega : \tau_n(\omega) > k\} = \bigcap_{k>0} \bigcup_{n>0} \{\omega : \tau_n(\omega) > k\}$$

Pertanto $\tau_n \uparrow +\infty$ \mathbb{P} -q.s. sse

$$\mathbb{P}\left(\bigcup_{k>0} \bigcap_{n \geq 0} \{\omega : \tau_n(\omega) \leq k\}\right) = 0$$

è quindi sufficiente verificare che

$$\lim_n \mathbb{P}\left(\bigcap_{n \geq 0} \{\omega : \tau_n(\omega) \leq k\}\right) = \lim_n \mathbb{P}\left(\{\omega : \sup_{t \in [0, k]} |X_t(\omega)| > n\}\right) = 0$$

e questo segue immediatamente dalla continuità \mathbb{P} -q.s. di $t \mapsto X_t$.

Introduciamo adesso il processo arrestato al tempo τ_n , come al solito $X_t^n := X_{t \wedge \tau_n}$ in cui ricordiamo che X^n è ancora progressivamente misurabile poichè τ_n è un tempo di arresto. Scriviamo ora la formula di Ito per il processo arrestato. Ci è concesso in quanto

X^n rimane dentro la palla di raggio n (diciamo anche $|X_0| < n$) e possiamo approssimare, uniformemente nella palla di raggio n , $f \in C^2(\mathbb{R})$ e le sue prime due derivate con un polinomio

$$f(X_t^n) = f(X_0) + \int_0^{t \wedge \tau_n} ds \left[\partial_x f(X_s) \psi_s + \frac{1}{2} \partial_x^2 f(X_s) \varphi_s^2 \right] + \int_0^{t \wedge \tau_n} \partial_x f(X_s) \varphi_s dw_s \quad (2.16)$$

itostop

in cui l'integrale stocastico è ben definito (per quello rispetto a Lebesgue non ci sono problemi) poiché τ_n è un tempo di arresto. Più precisamente

$$\int_0^{t \wedge \tau_n} \partial_x f(X_s) \varphi_s dw_s = \int_0^t \mathbb{1}_{[0, \tau_n]}(s) \partial_x f(X_s) \varphi_s dw_s$$

e $s \mapsto \mathbb{1}_{[0, \tau_n]}(s)$ è progressivamente misurabile. Poiché $\tau_n \uparrow +\infty$ \mathbb{P} -q.s. si può (verificare) passare al limite $n \rightarrow \infty$ in (2.16) completando la dimostrazione della formula di Ito. \square

La teoria dell'integrale di Ito ha come ambito naturale le semimartingale (continue) piuttosto che il solo moto browniano, vedi [9, §3.3.2]; in [12, Thm. IV.2.2] $I_t(\varphi)$ è addirittura definito da una proprietà di martingala. Qui è stata svolta invece, come in [4, Ch. I.1] una teoria un po' meno generale. È infine possibile sviluppare tale teoria, con interessanti applicazioni, anche per martingale "con salti" (traiettorie continue da destra).

Assumendo tale teoria, vediamo ora come si possa dimostrare la caratterizzazione martingalesca del moto browniano, teorema 1.21 (teorema di Levy).

Dimostrazione del teorema di Levy.

Come già osservato è sufficiente mostrare che se M è una martingala (anzi basta che sia una martingala locale) continua con variazione quadratica t e $M_0 = 0$, allora per ogni $\theta \in \mathbb{R}$

$$\mathbb{E} \left(e^{i\theta(M_t - M_s)} \middle| \mathcal{F}_s \right) = e^{-\frac{1}{2}\theta^2(t-s)} \quad (2.17)$$

cfmb

applicando la formula di Ito alla funzione $f(x) = e^{i\theta x}$ troviamo (qui stiamo barando perché non abbiamo sviluppato l'integrazione stocastica in questa generalità)

$$e^{i\theta M_t} = e^{i\theta M_s} + i\theta \int_s^t e^{i\theta M_u} dM_u - \frac{1}{2}\theta^2 \int_s^t e^{i\theta M_u} d\langle M \rangle_u$$

poiché $e^{i\theta x}$ è una funzione limitata e $\langle M \rangle_u = u$ abbiamo che l'integrale stocastico sopra è una vera martingala e non solo una martingala locale, in particolare il suo valore di attesa condizionato a \mathcal{F}_s si annulla. Per $A \in \mathcal{F}_s$, troviamo quindi

$$\mathbb{E} \left(e^{i\theta(M_t - M_s)} \mathbb{1}_A \right) = \mathbb{P}(A) - \frac{1}{2}\theta^2 \int_s^t du \mathbb{E} \left(e^{i\theta(M_u - M_s)} \mathbb{1}_A \right)$$

Risolvendo tale equazione integrale per la funzione (continua) $t \mapsto \mathbb{E} \left(e^{i\theta(M_t - M_s)} \mathbb{1}_A \right)$ troviamo

$$\mathbb{E} \left(e^{i\theta(M_t - M_s)} \mathbb{1}_A \right) = \mathbb{P}(A) e^{-\frac{1}{2}\theta^2(t-s)}$$

che equivale (2.17). □

Si osserva che la precedente dimostrazione funziona anche nel caso multidimensionale, vedi [9, Thm. 3.3.16].

Enunciamo infine un teorema di rappresentazione, per la dimostrazione si veda [9, Thm. 3.4.2], che dice che una “generica” martinagala locale continua si può rappresentare come integrale di Ito rispetto al moto browniano. Questo risultato giustifica l’affermazione precedente che l’esempio di martingala dato dalla strategai di gioco alla roulette è quasi la generica martingala. Ovviamente l’ipotesi di traiettorie q.s. continue è essenziale, non c’è modo di rappresentare il processo di Poisson *compensato* $N_t - t$ come un’integrale di Ito.

t:rmbm

Teorema 2.7. *Sia M_t una martingala locale continua sullo spazio di probabilità filtrato $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ tale che la sua variazione quadratica $t \mapsto \langle M \rangle_t(\omega)$ è assolutamente continua \mathbb{P} -q.s. Allora esiste un ingradimento dello spazio di probabilità filtrato $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathcal{F}}_t, \tilde{\mathbb{P}})$ dotato di un moto browniano w_t e di un processo φ_t adattato a $\tilde{\mathcal{F}}_t$ con $\tilde{\mathbb{P}}\left(\int_0^T dt \varphi_t^2 < \infty\right)$ per ogni $T > 0$ tale che*

$$M_t = \int_0^t \varphi_s dw_s \quad \tilde{\mathbb{P}}\text{-q.s.}$$

Non c’è alcun motivo per cui sia possibile costruire un moto browniano nello spazio di probabilità originario $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$; al fine di rappresentare la martingala M bisogna “arricchirlo” opportunamente (semplicemente si prende il prodotto con un altro spazio di probabilità).

Idea della dimostrazione. La prima cosa da fare è trovarsi il browniano “giusto” per tale rappresentazione; cerchiamo di usare il teorema di Levy e la formula di Ito per costruire un browniano da M_t . Dichiariamo inoltre $q - t = q_t(\omega)$ la funzione tale che $\langle M \rangle_t = \int_0^t ds q_s^2$. Per Ito e Levy il brownaino putativo sembra essere la martingala

$$N_t = \int_0^t \frac{1}{q_s} dM_s$$

Il baco è che q potrebbe essere zero e non possiamo veramente definire N come sopra. Ci prendiamo allora un browniano B indipendente da M (per questo dobbiamo ingrandire lo spazio di probabilità) e definiamo

$$N_t = \int_0^t \mathbb{1}_{\{q_s \neq 0\}} \frac{1}{q_s} dM_s + \int_0^t \mathbb{1}_{\{q_s = 0\}} dB_s$$

Grazie alla formula di Ito nella versione per martingale continue troviamo allora $\langle N \rangle_t = t$ e quindi, per Levy, N è un browniano. Per concludere la dimostrazione basta ora convincersi che $M_t = \int_0^t q_s dN_s$.

s:eul

2.3. Equazioni differenziali stocastiche: esistenza e unicità per coefficienti Lipschitz

Finora abbiamo considerato solo il moto browniano unidimensionale, ma non ci sono difficoltà a definire (ed a costruire) il moto browniano in \mathbb{R}^m . Dato il solito spazio di

probabilità filtrato $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ il moto browniano m -dimensionale è il processo a valori in \mathbb{R}^m , $w_t = (w_t^1, \dots, w_t^m)$ tale che le coordinate sono moti browniani unidimensionali indipendenti. Equivalentemente w_t è il processo gaussiano di media nulla e funzione di covarianza data da $\mathbb{E}(w_t^i w_s^j) = \delta_{i,j} t \wedge s$, con $s, t \in \mathbb{R}_+$ e $i, j = 1, \dots, m$. Dato un campo vettoriale $b : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ ed una funzione $\sigma : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{M}_{n,m}$ (a valori matrici $n \times m$) vogliamo considerare l'equazione stocastica

$$\begin{cases} dX_t &= b(t, X_t)dt + \sigma(t, X_t)dw_t \\ X_0 &= x \end{cases} \quad (2.18) \quad \boxed{\text{sde}}$$

dove $x \in \mathbb{R}^n$ è il dato iniziale. Scritta esplicitamente in coordinate l'equazione (2.18) si legge

$$\begin{cases} dX_t^i &= b_i(t, X_t)dt + \sum_{j=1}^m \sigma_{i,j}(t, X_t)dw_t^j \\ X_0^i &= x^i \end{cases} \quad (2.19) \quad \boxed{\text{sdeco}}$$

dove $i = 1, \dots, n$. Come nella sezione precedente non attribuiamo alcun significato ai differenziali, è solo una notazione conveniente per non scrivere ogni volta tutti gli integrali. Il senso intuitivo di (2.18) è che in un intervallo dt l'incremento dX_t è dato da un termine proporzionale a dt più un termine di “rumore” proporzionale a \sqrt{dt} , ma a media nulla. Per questo ci si riferisce a b come al termine di *deriva* ed a $a_{ij} := (\sigma\sigma^*)_{ij} = \sum_{k=1}^m \sigma_{ik}\sigma_{jk}$ come al termine di *diffusione*. Per poter scrivere l'equazione (2.18) richiediamo che b e σ siano funzioni boreliane, per avere un teorema di esistenza e unicità serviranno, ovviamente, condizioni più forti.

Nello sviluppare la teoria dell'equazioni stocastiche dobbiamo cominciare con il decidere che cosa dichiarare soluzione di (2.18). Questo può essere precisato in vari modi, sviluppiamo in questa sezione la teoria di *soluzioni forti* (o alla Ito). Osserviamo che la nozione di integrale stocastico è definita solo \mathbb{P} -a.s. non possiamo quindi pretendere di risolvere (2.18) per ogni fissato $\omega \in \Omega$.

s:ssde **Definizione 2.8.**

Una soluzione di (2.18) è un processo (a valori in \mathbb{R}^n) X_t su $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ tale che:

- (i) $t \mapsto X_t$ è continuo \mathbb{P} -q.s.;
- (ii) X_t è adattato alla filtrazione \mathcal{F}_t ;
- (iii) per ogni $t \geq 0$, $i = 1, \dots, n$ e $j = 1, \dots, m$

$$\mathbb{P} \left(\int_0^t ds |b_i(s, X_s)| + \int_0^t ds \sigma_{i,j}(s, X_s)^2 < \infty \right) = 1$$

- (iv) Per ogni $t \geq 0$ e $i = 1, \dots, n$ l'identità

$$X_t^i = x^i + \int_0^t ds b_i(s, X_s)dt + \sum_{j=1}^m \int_0^t \sigma_{i,j}(s, X_s)dw_s^j \quad (2.20) \quad \boxed{\text{sdeint}}$$

è soddisfatta \mathbb{P} -q.s.

La condizione (ii) merita un commento: possiamo pensare che \mathcal{F}_t sia la filtrazione generata dal browniano, richiediamo quindi di poter determinare la soluzione nell'intervallo $[0, s]$ conoscendo solo la parte della traiettoria browniana in tale intervallo. Questa condizione è quindi una richiesta tipo “principio di causalità” ed è altresì necessaria per la buona definizione (nel senso di Ito) dell'integrale stocastico in (2.20).

Nel caso in cui σ sia la matrice identità (con $n = m$), l'equazione (2.20) si riduce a

$$X_t = x + \int_0^t ds b(s, X_s)dt + w_t \quad (2.21) \quad \boxed{\text{sdes1}}$$

che possiamo invece (con opportune ipotesi su b) risolvere per ogni fissato ω . Si può infatti leggere (2.21) come la ricerca di un punto fisso in $C(\mathbb{R}_+; \mathbb{R}^n)$ per un opportuna applicazione continua...

t:eusde

Teorema 2.9. *Siano b e σ globalmente Lipschitz, più precisamente esiste una costante L tale che per ogni $t \in \mathbb{R}_+$, $x, y \in \mathbb{R}^n$ (qui $|\cdot|$ è la norma euclidea in \mathbb{R}^n , per lo spazio delle matrici si può usare una norma qualunque, tanto sono tutte equivalenti)*

$$\begin{aligned} |b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| &\leq L|x - y| \\ |b(t, x)|^2 + |\sigma(t, x)|^2 &\leq L(1 + |x|^2) \end{aligned} \quad (2.22) \quad \boxed{\text{lip}}$$

allora esiste una soluzione, nel senso della definizione 2.8, all'equazione (2.18). Inoltre se X_t e \tilde{X}_t sono due soluzioni si ha

$$\mathbb{P}(X_t = \tilde{X}_t, \quad t \geq 0) = 1$$

La dimostrazione di questo teorema è sostanzialmente identica all'analogo risultato per equazioni differenziali ordinarie: l'unicità segue dal lemma di Gronwall, l'esistenza dalla convergenza delle iterate di Picard. L'unico ingrediente in più è la disuguaglianza di Doob che ci consente di stimare il superiore di integrali stocastici (vietatissimo portare i moduli dentro integrali stocastici!). Vediamo comunque i dettagli.

Dimostrazione.

Unicità. Per la continuità di X_t e \tilde{X}_t è sufficiente dimostrare che per ogni $t \geq 0$ si ha $\mathbb{P}(X_t = \tilde{X}_t) = 1$. Assumiamo che per ogni $t \geq 0$ si abbia $\mathbb{E}(|X_t|^2) < \infty$, $\mathbb{E}(|\tilde{X}_t|^2) < \infty$, il caso generale si riconduce a questo mediante *arresto*. Scrivendo (2.20) per X_t , \tilde{X}_t e sottraendo membro a membro troviamo

$$X_t - \tilde{X}_t = \int_0^t ds [b(s, X_s) - b(s, \tilde{X}_s)] + \int_0^t [\sigma(s, X_s) - \sigma(s, \tilde{X}_s)] dw_s$$

da cui troviamo, per Cauchy-Swartz nel primo integrale e le proprietà degli integrali stocastici,

$$\mathbb{E}|X_t - \tilde{X}_t|^2 \leq 2t \int_0^t ds \mathbb{E}|b(s, X_s) - b(s, \tilde{X}_s)|^2 + 2C \int_0^t ds \mathbb{E}|\sigma(s, X_s) - \sigma(s, \tilde{X}_s)|^2$$

per un'opportuna costante $C = C(n, m) > 0$ (il cui valore numerico dipende dalla norma che abbiamo deciso di usare nello spazio delle matrici). Utilizzando la condizione di Lipschitz (2.22) ed il lemma di Gronwall il risultato segue.

Esistenza. Definiamo ricorsivamente la successione di processi $X^{(k)}$ come

$$X_t^{(k+1)} = x + \int_0^t ds b(s, X_s^{(k)}) + \int_0^t \sigma(s, X_s^{(k)}) dw_s \quad (2.23) \quad \boxed{\text{picard}}$$

ed $X_t^{(0)} = x$. Procedendo per induzione ed utilizzando la seconda riga in (2.22) si verifica che per ogni $T > 0$ esiste una costante $C = C(T, L)$ tale che per ogni $t \in [0, T]$ e $k \geq 0$ si ha

$$\mathbb{E}|X_t^{(k)}|^2 \leq C(1 + |x|^2)e^{Ct} \quad (2.24) \quad \boxed{\text{stimaxk}}$$

che ci garantisce l'integrale stocastico in (2.23) essere una martingala. Abbiamo

$$X_t^{(k+1)} - X_t^{(k)} = \int_0^t ds [b(s, X_s^{(k)}) - b(s, X_s^{(k-1)})] + \int_0^t [\sigma(s, X_s^{(k)}) - \sigma(s, X_s^{(k-1)})] dw_s$$

Per la condizione (2.22), via Cauchy–Swartz e la disuguaglianza di Doob (teorema 1.9), otteniamo che per ogni $T > 0$ esiste una costante $C = C(T, L)$ tale che per ogni $k \geq 0$ e $t \in [0, T]$ si ha

$$\mathbb{E} \left(\sup_{s \in [0, t]} |X_s^{(k+1)} - X_s^{(k)}|^2 \right) \leq C \int_0^t ds \mathbb{E} \left(|X_s^{(k)} - X_s^{(k-1)}|^2 \right)$$

e quindi, per ricorsione,

$$\mathbb{E} \left(\sup_{s \in [0, t]} |X_s^{(k+1)} - X_s^{(k)}|^2 \right) \leq \frac{(Ct)^k}{k!} \sup_{t \in [0, T]} \mathbb{E} \left(|X_t^{(1)} - x|^2 \right) = \bar{C} \frac{(Ct)^k}{k!} \quad (2.25) \quad \boxed{\text{stimaxk2}}$$

In particolare ne segue che

$$\sum_{k=0}^{\infty} \mathbb{P} \left(\sup_{t \in [0, T]} |X_t^{(k+1)} - X_t^{(k)}| > \frac{1}{k^2} \right) < \infty$$

Utilizzando il lemma di Borel–Cantelli (lemma 2.2) troviamo un insieme $\bar{\Omega}$ di \mathbb{P} misura uno per cui vale la seguente affermazione. Per ogni $\omega \in \bar{\Omega}$ esiste $k_0 = k_0(\omega)$ per cui se $k \geq k_0$ e si ha

$$\sup_{t \in [0, T]} |X_t^{(k+1)}(\omega) - X_t^{(k)}(\omega)| \leq \frac{1}{k^2}$$

cosicché $X_t^{(k)}(\omega)$ è convergente, uniformemente per t in compatti, se $\omega \in \bar{\Omega}$. Considerando il limite $k \rightarrow \infty$ costruiamo un processo X_t continuo \mathbb{P} -q.s. Si tratta ora di mostrare che X_t soddisfa le altre richieste in definizione 2.8. Dalla disuguaglianza (2.25) segue che, per t fissato $X_t^{(k)}$ è una successione di Cauchy in $L_2(\mathbb{P})$; per la convergenza \mathbb{P} -q.s. di $X_t^{(k)}$ si ha allora che $X_t^{(k)}$ converge a X_t in $L_2(\mathbb{P})$. In particolare, grazie alla stima (2.24),

$\mathbb{E}|X_t|^2 \leq C(1 + |x|^2)e^{Ct}$. Utilizzando (2.22) è ora facile passare al limite $k \rightarrow \infty$ in (2.23) e dimostrare che X soddisfa (2.20). \square

s:mp

2.4. Formulazione di equazioni stocastiche come problema alle martingale (cenni)

Consideriamo l'equazione stocastica (2.18) quando b e σ non dipendono esplicitamente dal tempo. Ricordando che la matrice $(n \times n)$ di diffusione è definita da $a(x) := \sigma(x)\sigma(x)^*$, ovvero $a_{i,j}(x) = \sum_{k=1}^m \sigma_{i,k}(x)\sigma_{j,k}(x)$, introduciamo l'operatore differenziale (su \mathbb{R}^n)

$$Lu(x) := \frac{1}{2} \sum_{1 \leq i,j \leq n} a(x)_{i,j} \partial_i \partial_j u(x) + \sum_{i=1}^n b_i(x) \partial_i u(x) \quad (2.26)$$

Lmp

Se abbiamo una soluzione, nel senso della definizione 2.8, di (2.18) utilizzando la formula di Ito (nella versione multidimensionale) troviamo che per ogni $f \in C_K^2(\mathbb{R}^n)$ (C^2 a supporto compatto) il processo

$$M_t^f := f(X_t) - f(x) - \int_0^t ds Lf(X_s) \quad (2.27)$$

Mmp

è una martingala sullo spazio filtrato $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. D'altra parte il teorema di Levy (teorema 1.21) ci suggerisce che tale proprietà di martingala possa caratterizzare univocamente il processo X .

L'idea è quindi di formulare (2.18) come un problema per una misura di probabilità su $C(\mathbb{R}_+; \mathbb{R}^n)$ (dotato della σ -algebra dei boreliani e della filtrazione canonica) facendo sparire ogni riferimento al browniano w .

t:mp

Definizione 2.10. *Una misura di probabilità P_x su $C(\mathbb{R}_+; \mathbb{R}^n)$ risolve il problema alle martingale per L in (2.26) con dato iniziale $x \in \mathbb{R}^n$ se $P(X_0 = x) = 1$ (qui X_t è la coordinata canonica su $C(\mathbb{R}_+; \mathbb{R}^n)$) e per ogni $f \in C_K^2(\mathbb{R}^n)$ il processo M_t^f in (2.27) è una P_x -martingala.*

Come detto prima è sparita ogni traccia del browniano w , di conseguenza anche la condizione (fondamentale nella teoria di Ito) che X fosse adattato a w non c'è più. La formula di Ito ci consente di affermare che se il processo X^x è una soluzione forte di (2.18) (ove evidenziamo nella notazione la dipendenza dal dato iniziale) allora la sua legge $P_x := \mathbb{P} \circ (X^x)^{-1}$ risolve il corrispondente problema alle martingale; anzi la formula di Ito ci dice proprio chi sia (in termini di w) la martingala M^f . Meno ovvio, ma ancora vero, è che l'unicità nel senso di teorema 2.9 implica l'unicità di P_x .

Questo esempio può aiutare a capire la differenza tra unicità nel senso di Ito e nel senso del problema alle martingale. Consideriamo $n = 1$, $b = 0$, $\sigma(x) = \text{sgn}(x)$ e $X_0 = 0$. Supponiamo che $w_t = \int_0^t \text{sgn}(\beta_s) d\beta_s$ per un altro browniano β (notiamo che w è ancora un browniano per il teorema di Levy); allora troviamo subito due soluzioni a (2.18): $X_t = \pm \beta_t$. D'altra parte queste due soluzioni hanno la stessa legge e non contraddicono l'unicità nel senso del problema alle martingale.

Il vantaggio della formulazione 2.10 è che possiamo dimostrare l'esistenza in condizioni assai più deboli. L'unicità è più difficile (sorprendentemente si può tradurre in un

problema di esistenza per equazioni alle derivate parziali), ma si riesce a fare meglio di quanto si possa fare per equazioni ordinarie: pensare agli esempi classici di non unicità per equazioni differenziali ordinarie, magari un po' di rumore può aiutare.

Il seguente risultato, vedi [9, Thm. 5.4.22 e Rem. 5.4.30] fornisce delle condizioni sufficienti per l'esistenza e unicità del problema alle martingale.

t:eump

Teorema 2.11. *Se $b(x)$ e $a(x)$ sono limitati e continui allora, per ogni $x \in \mathbb{R}^n$ esiste una soluzione al problema alle martingale 2.10. Se inoltre $b(x)$ e $a(x)$ sono Hölder continui e $a(x) \geq c\mathbb{1}$, con $c > 0$ (ovvero L è uniformemente ellittico) allora la soluzione è unica.*

s:ou

2.5. Processo di Ornstein–Uhlenbeck

In questa sezione consideriamo l'equazione stocastica unidimensionale con termine di deriva lineare e termine di diffusione costante. Ovvero

$$\begin{cases} dX_t &= -\lambda X_t dt + dw_t \\ X_0 &= x \end{cases} \quad (2.28) \quad \text{oueq}$$

La cui soluzione si scrive immediatamente via variazione delle costanti, i.e.

$$X_t = e^{-\lambda t} x + \int_0^t e^{-\lambda(t-s)} dw_s = e^{-\lambda t} x + e^{-\lambda t} \int_0^t e^{\lambda s} dw_s \quad (2.29) \quad \text{oues}$$

la verifica è immediata:

$$dX_t = -\lambda \left[e^{-\lambda t} x + e^{-\lambda t} \int_0^t e^{\lambda s} dw_s \right] dt + e^{-\lambda t} e^{\lambda t} dw_t$$

Come appare chiaro dall'espressione (2.29) X_t si può ottenere come limite (in $L_2(d\mathbb{P})$) di combinazione lineare di processi gaussiani ed è quindi un processo gaussiano. La media e funzione di covarianza del processo X_t caratterizzano univocamente la sue distribuzioni finito dimensionali e pertanto il processo di Ornstein–Uhlenbeck stesso. Effettuiamo il calcolo, grazie alle proprietà dell'integrale stocastico troviamo $\mathbb{E}(X_t) = e^{-\lambda t} x$ e

$$\begin{aligned} \mathbb{E}\left((X_t - e^{-\lambda t} x)(X_s - e^{-\lambda s} x)\right) &= \mathbb{E}\left(\int_0^{s \wedge t} e^{-\lambda(t-u)} dw_u \int_0^{s \wedge t} e^{-\lambda(s-u)} dw_u\right) \\ &= \frac{1}{2\lambda} e^{-\lambda(t+s)} [e^{2\lambda t \wedge s} - 1] = \frac{1}{2\lambda} [e^{-\lambda|t-s|} - e^{-\lambda(t+s)}] \end{aligned} \quad (2.30) \quad \text{covou}$$

Si consiglia di eseguire lo stesso calcolo quando consideriamo il dato iniziale non fisso, ma aleatorio anch'esso. In particolare se prendiamo $\lambda > 0$ e x una gaussiana a media nulla e varianza $(2\lambda)^{-1}$ indipendente dal moto browniano w scopriamo che $\mathbb{E}(X_t) = 0$ mentre $\mathbb{E}(X_t^2) = (2\lambda)^{-1}$ è indipendente da t . Questo vuol dire, per la gaussianità di X , che la distribuzione ad un tempo fissato rimane sempre la stessa. Diremo allora, in modo naturale, che la distribuzione gaussiana di media nulla e varianza $(2\lambda)^{-1}$ è una misura invariante per il processo di Ornstein–Uhlenbeck. Poichè per v.a. gaussiane la convergenza della media e della varianza implica la convergenza in legge (dimostrare

questa affermazione. Sugg. scrivere le funzioni caratteristiche) facendo il limite $t = s \rightarrow \infty$ in (2.30) troviamo che la v.a. X_t converge in legge alla variabile gaussiana di media nulla e varianza $(2\lambda)^{-1}$ per $t \rightarrow \infty$. La misura invariante descrive quindi - almeno in questo caso - il comportamento asintotico del processo.

s:outcbm

Proposizione 2.12. *Sia $S : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ la funzione strettamente crescente definita da $S(t) := \int_0^t ds e^{2\lambda s} = (2\lambda)^{-1}[e^{2\lambda t} - 1]$. Siano inoltre X_t^x il processo di Ornstein–Uhlenbeck (in cui evidenziamo nella notazione la dipendenza dal dato iniziale $x \in \mathbb{R}$) e W un moto browniano standard. Allora, per ogni $x \in \mathbb{R}$,*

$$X_t^x \stackrel{Law}{=} e^{-\lambda t} [x + W_{S(t)}] \quad t \geq 0 \quad (2.31)$$

oubm

(questa identità va intesa come identità tra processi: i due processi, pur essendo definiti su diversi spazi di probabilità, hanno le stesse distribuzioni finito dimensionali e quindi la stessa legge).

Dalla precedente proposizione segue in particolare che le proprietà locali (come la non differenziabilità \mathbb{P} -q.s.) delle traiettorie browniane e del processo di Ornstein–Uhlenbeck sono le stesse: nell’equazione stocastica (2.18) per tempi “piccoli” conta solo il termine di diffusione.

Dimostrazione. Per la gaussianità dei due processi è sufficiente verificare l’identità della media e della funzione di covarianza. Poichè $\mathbb{E}(W_{S(t)}W_{S(t')}) = S(t) \wedge S(t') = S(t \wedge t')$ si conclude subito che la funzione di covarianza del secondo membro in (2.31) coincide con quella ottenuta in (2.30). \square

Armati di proposizione 2.12, siamo ora in grado di calcolare la probabilità di transizione per il processo di Ornstein–Uhlenbeck, ovvero vogliamo calcolare la funzione $q_t(x, y)$, $(t, x, y) \in \mathbb{R}_+ \times \mathbb{R}^2$ tale che per un qualunque $B \in \mathcal{B}(\mathbb{R})$, $x \in \mathbb{R}$ e $t > 0$ si ha

$$\mathbb{P}(X_t^x \in B) = \int_B dy q_t(x, y)$$

Quando “scopriremo” il legame tra soluzioni di equazioni stocastiche e operatori differenziali del second’ordine identificheremo q con la soluzione fondamentale dell’equazione

$$\partial_t u(t, x) = \frac{1}{2} \partial_{xx}^2 u(t, x) - \lambda x \partial_x u(t, x) \quad (2.32)$$

oupar

Vedremo inoltre come, attraverso un cambio di variabili (*trasformazione dello stato fondamentale*), il precedente operatore differenziale sia unitariamente equivalente all’operatore $1/2 \partial_x^2 + \lambda x^2/2$. In questo senso il processo di Ornstein–Uhlenbeck è uno dei modi di “risolvere” l’oscillatore armonico in meccanica quantistica.

Ricordando che

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

è la probabilità di transizione del moto browniano, vedi (1.1), abbiamo, grazie alla proposizione 2.12 (qui sotto P è la legge del browniano W ovvero la misura di Wiener)

$$\begin{aligned}\mathbb{P}(X_t^x \in B) &= P\left(e^{-\lambda t}[x + W_{S(t)}]\right) = \int_{e^{\lambda t}B} dy p_{S(t)}(x, y) \\ &= \int_B dy \frac{\sqrt{\lambda}}{\sqrt{\pi(1 - e^{-2\lambda t})}} \exp\left\{-\frac{\lambda(y - e^{-\lambda t}x)^2}{1 - e^{-2\lambda t}}\right\}\end{aligned}$$

che conclude il calcolo.

Verificare “a mano” che la soluzione dell’equazione (2.32) con dato iniziale $u_0(x)$ ammette la rappresentazione

$$u(t, x) = \int dy q_t(x, y)u_0(y)$$

Il seguente lemma “ovvio” ci sarà utile nella prossima sezione.

t:ovvio

Lemma 2.13. *Sia X_t il processo di Ornstein–Uhlenbeck con dato iniziale $x = 0$. Allora per ogni $T > 0$*

$$\lim_{\lambda \rightarrow +\infty} \sup_{t \in [0, T]} |X_t| = 0 \quad \mathbb{P} \text{ q.s.}$$

Dimostrazione. La dimostrazione non ha nulla di probabilistico. Si tratta di analizzare l’equazione

$$X_t = -\lambda \int_0^t ds X_s + w_t$$

dove w_t è una funzione continua assegnata. Dimostriamo che, detto $M_\lambda := \sup_{t \in [0, T]} X_t$ si ha $\lim_{\lambda \rightarrow \infty} M_\lambda = 0$. Sia τ_λ tale che $M_\lambda = X_{\tau_\lambda}$ (possiamo supporre $M_\lambda > 0$); preso ora un $\delta \in (0, M_\lambda)$ introduciamo $\sigma_\lambda < \tau_\lambda$ tale che per $t \in [\sigma_\lambda, \tau_\lambda]$ si abbia $\delta \leq X_t \leq M_\lambda$. (Ex. σ_λ e τ_λ sono tempi di arresto?) Abbiamo

$$M_\lambda = \delta - \lambda \int_{\sigma_\lambda}^{\tau_\lambda} dt X_t + w_{\tau_\lambda} - w_{\sigma_\lambda} \leq \delta - \delta\lambda(\tau_\lambda - \sigma_\lambda) + w_{\tau_\lambda} - w_{\sigma_\lambda}$$

Supponiamo ora che $\overline{\lim}_\lambda M_\lambda =: M > 0$ e ricaviamo un assurdo. Dichiariamo $\ell := \underline{\lim}_\lambda \lambda(\tau_\lambda - \sigma_\lambda)$; passando ad opportune sottosuccessioni possiamo supporre $M_\lambda \rightarrow M$ e $\lambda(\tau_\lambda - \sigma_\lambda) \rightarrow \ell$. Se $\ell = +\infty$ troviamo, per la limitatezza di w_t , il palese assurdo $M \leq -\infty$. Se invece $\ell \in [0, \infty)$ necessariamente $\tau_\lambda - \sigma_\lambda \rightarrow 0$ e quindi, per la continuità di w_t , troviamo $M \leq \delta(1 - \ell)$, ovvero $M \leq 0$ per l’arbitrarietà di δ . Ripendo lo stesso argomento per il minimo concludiamo la dimostrazione. \square

s:smo

2.6. Approssimazione di Smoluchowski

Ritorniamo al fenomeno fisico che ci ha motivato l’introduzione del moto browniano, vedi sezione 1; cerchiamo di fare un modello un po’ più raffinato (realistico?). Consideriamo una particella in un fluido viscoso (e.g. olio) soggetta ad un campo di forze assegnato

b ed agli urti delle molecole del fluido in agitazione termica. Scriviamo l'equazione di Newton indicando con X_t la posizione della particella:

$$m\ddot{X}_t = b(t, X_t) - \dot{X}_t + f_t$$

dove m è la massa della particella, $b(t, X_t)$ il campo di forza esterno, $-\dot{X}_t$ l'attrito e f_t la forza dovuta agli urti delle molecole del fluido. Scriviamo l'equazione di Newton come un sistema del prim'ordine e schematizziamo f_t come un termine impulsivo (cambia istantaneamente la velocità della particella) che modelliamo come \dot{w}_t , ove w_t è il moto browniano. Otteniamo allora

$$\begin{cases} dX_t &= Y_t dt \\ mdY_t &= b(t, X_t)dt - Y_t dt + dw_t \\ X_0 &= x \\ Y_0 &= y \end{cases} \quad (2.33) \quad \boxed{\text{aou}}$$

Ove, nell'applicazione considerata, X_t, Y_t sono processi a valori in \mathbb{R}^3 . Per complicare ulteriormente la terminologia il processo (X_t, Y_t) viene alle volte detto processo di Ornstein-Uhlenbeck!

In questa sezione vediamo come nel limite $m \rightarrow 0$ la soluzione di (2.33) converga all'equazione *aristotelica*

$$\begin{cases} dX_t &= b(t, X_t)dt + dw_t \\ X_0 &= x \end{cases} \quad (2.34) \quad \boxed{\text{arist}}$$

t:ma **Proposizione 2.14.** *Supponiamo che il campo vettoriale b sia globalmente Lipschitz (soddisfi cioè la condizione (2.22)). Poniamo $m = 1/\lambda$ e denotiamo con (X^λ, Y^λ) l'unica soluzione di (2.33) e con X la soluzione di (2.34). Allora per ogni $x, y \in \mathbb{R}^3$ e $T > 0$ si ha*

$$\lim_{\lambda \rightarrow \infty} \sup_{t \in [0, T]} |X_t^\lambda - X_t| = 0 \quad \mathbb{P} \text{ q.s.}$$

Dimostrazione. Dimostriamo la proposizione con l'ipotesi aggiuntiva che campo vettoriale b sia limitato, rimandando a [11] per il caso generale.

Per la variazione delle costanti possiamo scrivere la seconda equazione in (2.33) come

$$Y_t^\lambda = e^{-\lambda t} y + \int_0^t ds \lambda e^{-\lambda(t-s)} b(s, X_s^\lambda) + \int_0^t \lambda e^{-\lambda(t-s)} dw_s$$

che inserita nella prima equazione in (2.33) produce

$$\begin{aligned} X_t^\lambda &= x + \int_0^t ds e^{-\lambda s} y + \int_0^t ds \int_0^s du \lambda e^{-\lambda(s-u)} b(u, X_u^\lambda) + \int_0^t ds \int_0^s \lambda e^{-\lambda(s-u)} dw_u \\ &= x + \frac{1}{\lambda} (1 - e^{-\lambda t}) y + \int_0^t du (1 - e^{-\lambda(t-u)}) b(u, X_u^\lambda) + \int_0^t (1 - e^{-\lambda(t-u)}) dw_u. \end{aligned}$$

Scrivendo l'equazione (2.34) in forma integrale e sottraendo membro a membro troviamo

$$X_t^\lambda - X_t = \frac{1}{\lambda} (1 - e^{-\lambda t}) y - \int_0^t e^{-\lambda(t-s)} dw_s - \int_0^t ds e^{-\lambda(t-s)} b(s, X_s^\lambda) + \int_0^t ds [b(s, X_s^\lambda) - b(s, X_s)]. \quad (2.35) \quad \boxed{\text{stim4}}$$

In virtù della limitatezza di b e del lemma 2.13 abbiamo

$$\lim_{\lambda \rightarrow \infty} \sup_{t \in [0, T]} \left[\left| \int_0^t e^{-\lambda(t-s)} dw_s \right| + \left| \int_0^t ds e^{-\lambda(t-s)} b(s, X_s^\lambda) \right| \right] = 0 \quad \mathbb{P} \text{ q.s.}$$

s:lmv

Pertanto la dimostrazione si conclude utilizzando il lemma di Gronwall in (2.35). \square

2.7. Limite di McKean-Vlasov

Consideriamo N diffusioni in \mathbb{R}^d con interazione di campo medio descritta da un potenziale a due corpi $V: \mathbb{R}^d \rightarrow \mathbb{R}$, ovvero il sistema di equazioni stocastiche

$$dX_t^i = \frac{1}{N} \sum_{j=1}^N \nabla V(X_t^i - X_t^j) dt + dw_t^i \quad (2.36) \quad \text{mvN}$$

dove $w^i, i = 1, \dots, N$ sono browniani indipendenti in \mathbb{R}^d . Vogliamo descrivere tale sistema nel limite di infiniti gradi di libertà $N \rightarrow \infty$. Nel caso in cui non ci sia perturbazione stocastica, tale limite è descritto dall'equazione di Vlasov.

Per prima cosa dobbiamo decidere quale sia l'osservabile rilevante: non vogliamo conoscere la posizione della singola particella, ma solo le proprietà 'medie' del sistema. Definiamo quindi la *misura empirica* $\pi^N: \mathbb{R}^{dN} \rightarrow \mathcal{P}(\mathbb{R}^d)$ come la mappa che a $x = (x^1, \dots, x^N) \in \mathbb{R}^{dN}$ associa la probabilità π^N su \mathbb{R}^d definita da

$$\pi^N := \frac{1}{N} \sum_{i=1}^N \delta_{x^i}.$$

Dato $T > 0$, con lieve abuso di notazioni chiamiamo anche misura empirica la mappa $\pi^N: C([0, T]; \mathbb{R}^{dN}) \rightarrow C([0, T]; \mathcal{P}(\mathbb{R}^d))$ definita da $(\pi_t^N)(x) := \pi^N(x_t), t \in [0, T]$. Nel seguito consideriamo $\mathcal{P}(\mathbb{R}^d)$ con la topologia indotta dalla convergenza debole (convergono integrali di funzioni continue e limitate). Con tale topologia $\mathcal{P}(\mathbb{R}^d)$ è uno spazio polacco, i.e. metrizzabile completo e separabile. Una possibile metrica che descrive la convergenza debole è quella definita dalla distanza di Levy.

Sia $X^N = (X^{N,1}, \dots, X^{N,N})$ la soluzione di (2.36) con una successione di dati iniziali (deterministici o aleatori) $x_0^N = (x_0^{N,1}, \dots, x_0^{N,N})$ per (2.36) organizzati in modo che la successione di probabilità su \mathbb{R}^d data da $\pi^N(x_0^N)$ converga (in probabilità nel caso di dati aleatori) a $\rho_0(r)dr$ (dr è la misura di Lebesgue in \mathbb{R}^d) per un'opportuna densità di probabilità $\rho_0: \mathbb{R}^d \rightarrow \mathbb{R}_+$. Ad esempio ciò accade se x^N, i_0 sono variabili i.i.d. con legge $\rho_0(r)dr$.

Vogliamo dimostrare che tale proprietà si mantiene nel tempo, ovvero esistere una famiglia di densità di probabilità $\rho_t, t \in [0, T]$ per cui, dato $t \in [0, T]$, la successione di variabili aleatoria a valori in $\mathcal{P}(\mathbb{R}^d)$ data da $\pi^N(X_t^N)$ converge in probabilità a $\rho_t(r)dr$. Vogliamo anche scoprire come calcolare ρ_t in funzione di ρ_0 : risolvendo l'equazione di McKean-Vlasov

$$\begin{cases} \partial_t \rho_t - \nabla \cdot (\rho_t \rho_t * \nabla V) = \frac{1}{2} \Delta \rho_t \\ \rho_{t=0} = \rho_0 \end{cases}$$

dove $\rho * \nabla V(r) := \int dr' \rho(r') \nabla V(r - r')$ è la convoluzione. L'interpretazione di (2.7) è la seguente: nel limite $N \rightarrow \infty$ la misura empirica evolve come un'equazione di trasporto nel campo di forza "efficace" $-\rho * \nabla V$ e con un termine di viscosità.

L'enunciato che dimostreremo è in effetti lievemente diverso.

t:mV **Teorema 2.15.** *Sia $V \in C^2(\mathbb{R}^d)$ con ∇V Lipschitz. Sia inoltre X^N la soluzione di (2.36) con dati iniziali come sopra e tali che*

$$\sup_N \mathbb{E} \left(\frac{1}{N} \sum_{i=1}^N |x_0^{N,i}|^2 \right) < +\infty.$$

Dato $T > 0$ sia \mathcal{P}^N , probabilità su $C([0, T]; \mathbb{R}^{dN})$, la legge di X^N . La successione di probabilità su $C([0, T]; \mathcal{P}(\mathbb{R}^d))$ definita da $\mathcal{P}^N \circ (\pi^N)^{-1}$ (ovvero la legge della misura empirica) converge debolmente alla probabilità concentrata su $\bar{\pi}$ ove $\bar{\pi}_t(r) = \rho_t(r) dr$ con $\rho = \rho_t(r)$ unica soluzione debole di (2.7).

La formulazione debole di (2.7) è semplicemente ottenuta moltiplicando per una funzione di prova liscia e integrando per parti. Rimandando ai corsi di equazioni alle derivate parziali, assumeremo qui che tale formulazione ammette un'unica soluzione. L'esistenza si può invece ottenere come corollario del precedente teorema.

La strategia di dimostrazione di Teorema 2.15 è canonica: prima dimostriamo la compattezza di $\{\mathcal{P}^N \circ (\pi^N)^{-1}\}$, poi facciamo vedere che i suoi punti di accumulazione sono supportati sulle soluzioni deboli di (2.7) ed infine ci appelliamo all'unicità.

Per dimostrare la compattezza, applicando Prohorov allo spazio $C([0, T]; \mathcal{P}(\mathbb{R}^d))$, è sufficiente il seguente lemma in cui $\pi = (\pi_t)_{t \in [0, T]}$ è un elemento di $C([0, T]; \mathcal{P}(\mathbb{R}^d))$ e si indica con $\langle \pi, \phi \rangle$ la dualità tra probabilità e funzioni.

t:pcp **Lemma 2.16.** *Sia $\{P_n\}$ una successione di probabilità su $C([0, T]; \mathcal{P}(\mathbb{R}^d))$. Supponiamo esistere una successione di compatti $K_\ell \subset \subset \mathcal{P}(\mathbb{R}^d)$ per cui*

$$\lim_{\ell \rightarrow \infty} \sup_n P_n(\exists t \in [0, T]: \pi_t \notin K_\ell) = 0. \quad (2.37) \quad \boxed{\text{c1}}$$

Supponiamo inoltre che per ogni $\phi: \mathbb{R}^d \rightarrow \mathbb{R}$ liscia (C^∞) a supporto compatto e per ogni $\eta > 0$ valga

$$\lim_{\delta \downarrow 0} \sup_n P_n \left(\sup_{|t-s| < \delta} |\langle \pi_t, \phi \rangle - \langle \pi_s, \phi \rangle| > \eta \right) = 0. \quad (2.38) \quad \boxed{\text{c2}}$$

Allora la successione $\{P_n\}$ è tight.

Il punto essenziale è che in (2.37) abbiamo lo stesso compatto K_ℓ per tutti i tempi, questo ci permette di guardare in (2.38) il modulo di continuità solo integrando una funzione alla volta. Tralasciando la dimostrazione del lemma vediamo come si usa nel nostro caso.

Per il teorema di Prohorov per le probabilità su \mathbb{R}^d , la stima (2.37) segue dalla seguente affermazione.

t:dc1 **Lemma 2.17.** Sia $\varphi: \mathbb{R}^d \rightarrow \mathbb{R}_+$ data da $\varphi(x) = |x|^2/2$. Allora

$$\lim_{\ell \rightarrow \infty} \overline{\lim}_{N \rightarrow \infty} \mathcal{P}^N \left(\sup_{t \in [0, T]} \langle \pi_t^N, \varphi \rangle > \ell \right) = 0.$$

Dimostrazione. Per la formula di Ito,

$$\varphi(X_t^i) = \varphi(X_0^i) + \int_0^t ds \left[-\frac{1}{N} \sum_j \nabla \varphi(X_s^i) \cdot \nabla V(X_s^i - X_s^j) + \frac{1}{2} \Delta \varphi(X_s^i) \right] + M_t^{\varphi, i}$$

dove $M^{\varphi, i}$, $i = 1, \dots, N$ sono martingale con “angoletto”

$$[M^{\varphi, i}, M^{\varphi, j}]_t = \delta_{i, j} \int_0^t ds |\nabla \varphi(X_s^i)|^2.$$

Osserviamo che, per la definizione della misura empirica, $\pi^N * \nabla V(r) = \frac{1}{N} \sum_j \nabla V(r - X^j)$, sommando su i e dividendo per N ricaviamo

$$\pi_t^N(\varphi) = \pi_0^N(\varphi) + \int_0^t ds \left[-\pi_s^N(\nabla \varphi \cdot \pi^N * \nabla V) + \frac{1}{2} d \right] + M_t^\varphi \quad (2.39) \quad \boxed{\text{dain}}$$

dove M^φ è una martingala con variazione quadratica

$$[M^\varphi]_t = \frac{1}{N^2} \sum_i \int_0^t ds |\nabla \varphi(X_s^i)|^2 = \frac{1}{N} \int_0^t ds \langle \pi_s^N, |\nabla \varphi|^2 \rangle.$$

Per la disuguaglianza di Doob,

$$\mathbb{E} \left(\sup_{t \in [0, T]} |M_t^\varphi|^2 \right) \leq 4\mathbb{E}[M^\varphi]_T.$$

D'altra parte, poiché ∇V è Lipschitz,

$$\begin{aligned} |\nabla \varphi(r) \cdot \pi^N * \nabla V(r)| &\leq C|r| \frac{1}{N} \sum_j [|\nabla V(0)| + |r - X^j|] \\ &\leq C \left(1 + |r|^2 + \frac{1}{N} \sum_j |X^j|^2 \right) \leq C(1 + \varphi(r) + \langle \pi^N, \varphi \rangle). \end{aligned}$$

Per ipotesi sui dati iniziali, $\mathbb{E}(\langle \pi_0^N, \varphi \rangle)$ è limitato uniformemente in N . Prendendo quindi in (2.39) l'estremo superiore in t ed utilizzando il lemma di Gronwall è immediato ricavare

$$\mathbb{E} \left(\sup_{t \in [0, T]} \langle \pi_t^N, \varphi \rangle \right) < +\infty$$

che implica la tesi. □

Come per la convergenza della passeggiata aleatoria al moto browniano, la stima (2.38) si ricava dall'affermazione seguente.

t:dc2 **Lemma 2.18.** Sia $\phi \in C_K^2(\mathbb{R}^d)$. Allora per ogni $\eta > 0$

$$\lim_{\delta \downarrow 0} \overline{\lim}_{N \rightarrow \infty} \sup_{s \in [0, T-\delta]} \frac{1}{\delta} \mathcal{P}^N \left(\sup_{t \in [s, s+\delta]} |\langle \pi_t^N, \phi \rangle - \langle \pi_s^N, \phi \rangle| > \eta \right) = 0.$$

Dimostrazione. Fissato $s \in [0, T]$, applichiamo la formula di Ito nell'intervallo $[s, t]$,

$$\phi(X_t^i) = \phi(X_s^i) + \int_s^t du \left[-\frac{1}{N} \sum_j \nabla \phi(X_u^i) \cdot \nabla V(X_u^i - X_u^j) + \frac{1}{2} \Delta \phi(X_u^i) \right] + M_t^{s, \phi, i}$$

dove $M_t^{s, \phi, i}$, $i = 1, \dots, N$, $t \in [s, T]$ sono martingale con “angoletto”

$$[M^{\varphi, i}, M^{\phi, j}]_t = \delta_{i, j} \int_s^t ds |\nabla \phi(X_u^i)|^2.$$

Sommando su i e dividendo per N ricaviamo

$$\pi_t^N(\phi) - \pi_s^N(\phi) = \int_s^t du \left[-\langle \pi_u^N, \nabla \phi \cdot \pi_u^N * \nabla V \rangle + \frac{1}{2} \langle \pi_u^N, \Delta \phi \rangle \right] + M_t^{s, \phi}$$

dove $M^{s, \phi}$ è una martingala con variazione quadratica

$$[M^\varphi]_t = \frac{1}{N} \int_s^t du \langle \pi_u^N, |\nabla \phi|^2 \rangle$$

che si annulla per $N \rightarrow \infty$. Utilizzando la disuguaglianza di Doob ed il lemma precedente per controllare ∇V è ora facile concludere la dimostrazione. \square

t:idlim **Lemma 2.19.** Sia P , probabilità su $C([0, T]; \mathcal{P}(\mathbb{R}^d))$, un punto di accumulazione di $\{\mathcal{P}^N \circ (\pi^N)^{-1}\}$. Allora, con P probabilità uno, per ogni $t \in [0, T]$ e $\phi \in C_K^2(\mathbb{R}^d)$,

$$\langle \pi_t, \phi \rangle = \langle \bar{\pi}_0, \phi \rangle + \int_0^t ds \left\{ \langle \pi_s, \nabla \phi \cdot \pi_s * \nabla V \rangle + \frac{1}{2} \langle \pi_s, \Delta \phi \rangle \right\}.$$

ove $\bar{\pi}_0(dr) = \rho_0(r)dr$.

Dimostrazione. Per continuità e densità basta verificare l'affermazione per una ϕ fissata. Sia $N \rightarrow \infty$ una sottosuccessione (non rinominata) per cui $\mathcal{P}^N \circ (\pi^N)^{-1} \rightarrow P$. In virtù dei precedenti calcoli, con \mathcal{P}^N probabilità uno vale

$$\langle \pi_t^N, \phi \rangle = \langle \pi_0^N, \phi \rangle + \int_0^t ds \left\{ \langle \pi_s^N, \nabla \phi \cdot \pi_s^N * \nabla V \rangle + \frac{1}{2} \langle \pi_s^N, \Delta \phi \rangle \right\} + M_t^\phi$$

dove M_t^ϕ è un martingala con variazione quadratica

$$[M^\phi]_t = \frac{1}{N} \int_0^t ds \langle \pi_s^N, |\nabla \phi|^2 \rangle.$$

Sia

$$E_t(\pi, \phi) := \langle \pi_t, \phi \rangle - \langle \bar{\pi}_0, \phi \rangle - \int_0^t ds \left\{ \langle \pi_s, \nabla \phi \cdot \pi_s * \nabla V \rangle - \frac{1}{2} \langle \pi_s, \Delta \phi \nabla \rangle \right\}.$$

Poiché $\mathcal{P}^N \circ (\pi^N)^{-1} \rightarrow P$, per ogni $\eta > 0$

$$\begin{aligned} P \left(\sup_{t \in [0, T]} |E_t(\pi, \phi)| > \eta \right) &\leq \varliminf_{N \rightarrow \infty} \mathcal{P}^N \left(\sup_{t \in [0, T]} |E_t(\pi^N, \phi)| > \eta \right) \\ &= \varliminf_{N \rightarrow \infty} \mathcal{P}^N \left(\sup_{t \in [0, T]} |M_t^\phi| > \eta \right) = 0 \end{aligned}$$

grazie (ancora!) alla diseguaglianza di Doob. Concludiamo per l'arbitrarietà di η .

In realtà c'è una minima truffa, se ∇V non è limitato la funzione $\pi \mapsto \langle \pi, \nabla V \rangle$ non è continua in $\mathcal{P}(\mathbb{R}^d)$ e quindi l'insieme $\{\pi : |E_t(\pi, \phi)| > \eta\}$ non è aperto (o meglio non sappiamo esserlo). Poiché ∇V ha crescita lineare la stima in Lemma 2.17, che controlla funzione con crescita quadratica, è più che sufficiente ad emendare la truffa. Si omettono i dettagli. \square

Il limite di McKean-Vlasov si può formulare il modo diverso, forse più evocativo, che si discute brevemente senza dettagli. Consideriamo il caso con dati iniziali i.i.d. Allora le diffusioni X^i , $i = 1, \dots, N$ sono scambiabili, ovvero la legge congiunta è invariante per permutazioni. Fissamo l'attenzione su una di queste, diciamo X^1 e cerchiamo di capire come sia la sua legge. Per N finito dipende dalle altre e non possiamo dire nulla in più, nel limite $N \rightarrow \infty$ accade invece che converga alla legge della seguente diffusione “non lineare”

$$\begin{cases} dX_t = -(\pi_t * \nabla V)(X_t)dt + dw_t \\ \pi_t = \text{Legge di } X_t \end{cases}$$

che si interpreta naturalmente come una condizione di punto fisso: a π data la prima equazione determina X , la seconda impone che π al tempo t sia proprio la legge di X al tempo t . Evidentemente π_t , $t \in [0, T]$ evolve con l'equazione di McKean-Vlasov (2.7).

s:lt

2.8. Tempi locali di un browniano e rappresentazione del browniano riflesso

Introduciamo l'insieme degli zeri di un moto browniano:

$$Z(\omega) := \{t \in \mathbb{R}_+ : w_t(\omega) = 0\}$$

che risulta un sottoinsieme aleatorio di \mathbb{R}_+ . La seguente proposizione afferma che Z ha, con probabilità uno, una struttura tipo insieme di Cantor.

t:zbm

Proposizione 2.20. *Per \mathbb{P} quasi tutti gli $\omega \in \Omega$, l'insieme $Z(\omega)$ è chiuso, ha misura di Lebesgue zero e non ha punti isolati.*

Dimostrazione (incompleta). Essendo l'insieme degli zeri di una funzione continua \mathbb{P} -q.s., $Z(\omega)$ è chiuso. Per vedere che, \mathbb{P} -q.s., ha misura di Lebesgue zero basta scrivere

$$\mathbb{E} \int_0^\infty dt \mathbb{1}_Z(t) = \int_0^\infty dt \mathbb{P}(w_t = 0) = 0$$

Poiché $w_0 = 0$ si ha $0 \in Z$, dalla legge dei logaritmi iterati (vedi [9, Thm. 2.9.23] o [12, Thm. II.1.9]) segue che $\{0\}$ è un punto di accumulazione di Z . Dato che l'incremento di un browniano, $w_{t_0+t} - w_{t_0}$, con t_0 fissato e $t \geq 0$, è ancora un browniano, il risultato segue. \square

Vogliamo ora costruire una misura (aleatoria) su \mathbb{R}_+ il cui supporto sia proprio Z ; tale misura descriverà quindi il tempo che il browniano ha trascorso in zero. Formalmente tale misura dovrebbe essere $\delta(w_t)dt$ con $\delta(x)$ la “funzione” di Dirac. Non possiamo sperare di dare senso direttamente a $\delta(w_t)$, ma riconoscendo in $2\delta(x)$ la derivata seconda di $|x|$ la formula di Ito ci direbbe

$$|w_t| = \int_0^t \operatorname{sgn}(w_s)dw_s + \int_0^t ds \delta(w_s)$$

Il problema è diventato quindi di dare senso alla formula di Ito per funzioni non C^2 . La convessità di $|x|$ ci aiuterà.

Sia $f \in C(\mathbb{R})$ una funzione convessa e indichiamo con f'_- la derivata sinistra (che esiste per tutti gli $x \in \mathbb{R}$). Il seguente risultato, noto come *formula di Tanaka*, estende la formula di Ito a tali funzioni.

t:tf **Teorema 2.21.** *Data f convessa, esiste un processo continuo e crescente A_t^f tale che*

$$f(w_t) = f(w_0) + \int_0^t f'_-(w_s)dw_s + A_t^f$$

Osservazione. Il risultato è vero (con la stessa dimostrazione) per una qualunque semimartingala continua (ovvero un processo X_t tale che $dX_t = \psi_t dt + \varphi_t dw_t$). Diversamente da Ito, la formula di Tanaka è però speciale per processi unidimensionali (solo in una dimensione le funzioni convesse hanno derivata monotona).

Dimostrazione. Se $f \in C^2(\mathbb{R})$ Tanaka si riduce ad Ito, approssimiamo quindi f con una funzione $C^2(\mathbb{R})$ come segue. Sia $j \in C^\infty(\mathbb{R})$ positiva con supporto in $(-\infty, 0]$ tale che $\int dx j(x) = 1$. Per $\varepsilon > 0$ definiamo $j_\varepsilon(x) := \varepsilon^{-1}j(\varepsilon^{-1}x)$ e $f_\varepsilon(x) := j_\varepsilon * f(x) = \int dy j_\varepsilon(x - y)f(y)$. Evidentemente $f_\varepsilon(x) \rightarrow f(x)$ per ogni $x \in \mathbb{R}$ e $f'_\varepsilon(x) \uparrow f'_-(x)$. Ne segue che $f_\varepsilon(w_t) \rightarrow f(w_t)$ e $\int_0^t f'_\varepsilon(w_s)dw_s \rightarrow \int_0^t f'_-(w_s)dw_s$ in probabilità rispetto a \mathbb{P} . Applicando la formula di Ito a f_ε possiamo allora definire

$$A_t^\varepsilon := \frac{1}{2} \int_0^t ds f''_\varepsilon(w_s) = f_\varepsilon(w_t) - f_\varepsilon(0) - \int_0^t f'_\varepsilon(w_s)dw_s$$

che risulta continuo e crescente (per la convessità di f_ε). Per quanto detto prima possiamo passare al limite per $\varepsilon \downarrow 0$; definendo $A_t^f = \lim_\varepsilon A_t^\varepsilon$ completiamo allora la dimostrazione della formula di Tanaka. \square

Considerando la funzione convessa $f_a(x) = |x - a|$, con $a \in \mathbb{R}$ abbiamo

$$f'_a(x) = \operatorname{sgn}(x - a) := \begin{cases} 1 & \text{se } x > a \\ -1 & \text{se } x \leq a \end{cases}$$

Chiamiamo allora *tempo locale* del browniano in a il processo continuo crescente L_t^a tale che

$$|w_t - a| = |a| + \int_0^t \operatorname{sgn}(w_s - a) dw_s + L_t^a$$

Vediamo ora che la misura dL^a descrive effettivamente il tempo passato dal browniano in a .

t:slt **Proposizione 2.22.** *La misura dL_t^a ha supporto contenuto in $\{t \in \mathbb{R}_+ : w_t = a\}$.*

Dimostrazione. Applicando la formula di Ito al processo w_t ed alla funzione $(x - a)^2$ troviamo

$$(w_t - a)^2 = a^2 + 2 \int_0^t (w_s - a) dw_s + t$$

Applicando invece la formula di Ito al processo $|w_t - a|$ (qui stiamo leggermente barando, non l'abbiamo veramente dimostrata in questa generalità) ed alla funzione x^2 troviamo

$$(w_t - a)^2 = a^2 + 2 \int_0^t |w_s - a| \operatorname{sgn}(w_s - a) dw_s + 2 \int_0^t |w_s - a| dL_s^a + \int_0^t ds [\operatorname{sgn}(w_s - a)]^2$$

che confrontata con la precedente fornisce $\int_0^t |w_s - a| dL_s^a = 0$. □

Con qualche sforzo in più si può dimostare (vedi [12, VI, Cor. 1.9]) che, \mathbb{P} -q.s.

$$L_t^a = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_0^t ds \mathbb{1}_{[a, a+\varepsilon]}(w_s) \tag{2.40} \quad \boxed{\text{lccl}}$$

il che conferma l'affermazione che il tempo locale L_t^a misura il tempo passato in a dal browniano.

Utilizzando il corollario 1.29, non è difficile verificare, vedi [12, Thm. VI.1.7] che si può scegliere $a \mapsto L_t^a$ continuo con probabilità uno, in particolare misurabile. Come corollario della formula di Tanaka si arriva facilmente alla seguente *formula del tempo di occupazione*; per ogni $f : \mathbb{R} \rightarrow [0, \infty)$ boreliana

$$\int_0^t ds f(w_s) = \int_{\mathbb{R}} da f(a) L_t^a$$

che conferma l'idea che L_t^a misuri il tempo passato dal browniano in a .

Non abbiamo nessuna difficoltà a definire un moto browniano riflesso in zero: basta dichiararlo $|w_t|$. Vorremo però definire la riflessione per una qualunque equazione stocastica (unidimensionale) e trovare una diversa caratterizzazione del browniano riflesso. Il risultato essenziale è il seguente *Lemma di Skorohod* in cui la probabilità non appare.

t:ls **Lemma 2.23.** *Sia $y \in C(\mathbb{R}_+)$ tale che $y(0) = 0$. Esiste un'unica coppia di funzioni continue $z, a \in C(\mathbb{R}_+)$ tali che:*

(i) per ogni $t \geq 0$ si ha $z(t) = y(t) + a(t)$;

(ii) si ha $z(t) \geq 0$;

(iii) la funzione a è crescente, $a(0) = 0$ e la misura $da(t)$ ha supporto contenuto in $\{t \in \mathbb{R}_+ : z(t) = 0\}$.

L'idea è che, dato y , la funzione a fornisce la spinta necessaria per mantenere z positivo, ma nulla in più: appena $z > 0$ a non agisce più.

Dimostrazione. L'esistenza della coppia (z, a) viene dimostrata per esibizione: detto $y_- := \max\{0, -y\}$ poniamo

$$a(t) = \sup_{s \in [0, t]} (y(s))_- , \quad z(t) = y(t) + a(t)$$

Si verifica immediatamente (Sugg. fare il disegno) che soddisfa le richieste del lemma.

Per dimostare l'unicità ci diamo un'altra coppia (\tilde{z}, \tilde{a}) che soddisfa (i)–(iii). Abbiamo che $z - \tilde{z} = a - \tilde{a}$ è a variazione limitata (come differenza di due funzioni crescenti), poiché $z(0) = \tilde{z}(0)$ abbiamo, per il teorema fondamentale del calcolo,

$$[z(t) - \tilde{z}(t)]^2 = 2 \int_0^t [z(s) - \tilde{z}(s)] d[a(s) - \tilde{a}(s)] = -2 \int_0^t z(s) d\tilde{a}(t) - 2 \int_0^t \tilde{z}(s) da(t) \leq 0$$

in cui abbiamo usato le proprietà di supporto di a, \tilde{a} e $z, \tilde{z} \geq 0$. □

Possiamo ora definire cosa sia la soluzione di un'equazione stocastica $dX_t = b(X_t)dt + \sigma(X_t)dw_t$ con dato iniziale $x \geq 0$ e riflessione in zero: è la coppia (X_t, ℓ_t) di processi adattati continui \mathbb{P} -q.s. tale che, \mathbb{P} -q.s.,

(i) per ogni $t \geq 0$ si ha

$$X_t = x + \int_0^t ds b(X_s) + \int_0^t \sigma(X_s)dw_s + \ell_t$$

(ii) si ha $X_t \geq 0$;

(iii) il processo ℓ_t è crescente, $\ell_0 = 0$ e la misura ℓ_t ha come supporto un sottoinsieme di $\{t \in \mathbb{R}_+ : X_t = 0\}$.

In tal caso si dice che (X_t, ℓ_t) è soluzione del *problema di Skorohod*. Con appropriate condizioni su b e σ si può dimostrare l'esistenza e unicità della soluzione, vedi [4, §23] per i dettagli.

t:t12

Teorema 2.24. *Sia w un moto browniano e poniamo $S_t = \sup_{s \in [0, t]} w_s$. Allora i processi $(S_t, S_t - w_t)$ e $(L_t, |w_t|)$ (in cui L_t è il tempo locale in zero di w) hanno la stessa legge.*

Dimostrazione. Poniamo $B_t := -\int_0^t \text{sgn}(w_s)dw_s$ e notiamo che, grazie al teorema di Levy (teorema 1.21), B_t è - in legge - ancora un moto browniano. La formula di Tanaka dice allora

$$|w_t| = \int_0^t \text{sgn}(w_s)dw_s + L_t = -B_t + L_t$$

grazie al lemma 2.23 abbiamo allora

$$L_t = \sup_{s \in [0, t]} (B_s) \quad |w_t| = -B_t + \sup_{s \in [0, t]} (B_s)$$

s:arbm

che conclude la dimostrazione. □

2.9. Approssimazione del browniano riflesso

In questa sezione vediamo come, se mettiamo il campo vettoriale $b(x)$ sufficientemente forte per $x < 0$, la soluzione dell'equazione stocastica (unidimensionale) $dX = b(X_t)dt + dw_t$ converga al moto browniano riflesso in zero. Ricordando che $x_- := \max\{0, -x\}$, dato il solito spazio di probabilità filtrato $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ con un moto browniano w e $\lambda > 0$ chiamiamo X^λ l'unica soluzione dell'equazione stocastica

$$\begin{cases} dX_t^\lambda &= \lambda(X_t^\lambda)_- dt + dw_t \\ X_0^\lambda &= 0 \end{cases} \quad (2.41) \quad \boxed{1p1}$$

t:cbmr

Teorema 2.25. *Sia X^λ la soluzione di (2.41) e*

$$X_t := w_t + \sup_{s \in [0, t]} (-w_s)$$

(notiamo che, grazie al teorema 2.24, X_t ha la legge di un browniano riflesso in zero). Allora per ogni $T > 0$

$$\lim_{\lambda \rightarrow \infty} \sup_{t \in [0, T]} |X_t^\lambda - X_t| = 0 \quad \mathbb{P}\text{-q.s.}$$

Cominciamo con un lemma generale che, come caso particolare, garantisce la monotonia di X^λ rispetto a λ .

t:mes

Lemma 2.26. *Siano b^1, b^2 funzioni globalmente Lipschitz da \mathbb{R} a \mathbb{R} . Sia inoltre $X^i, i = 1, 2$ la soluzione di*

$$\begin{cases} dX_t^i &= b^i(X_t)dt + dw_t \\ X_0^i &= x^i \end{cases}$$

Se $b^1(x) \leq b^2(x)$ per ogni $x \in \mathbb{R}$ e $x^1 \leq x^2$ allora

$$\mathbb{P}(X_t^1 \leq X_t^2, \text{ per ogni } t \geq 0) = 1$$

Questo lemma fornisce un esempio di *accoppiamento*, ovvero costruiamo due processi stocastici nello stesso spazio di probabilità. In questo caso, poiché X^1 e X^2 sono costruiti in funzione dello stesso browniano w_t l'accoppiamento è banale (usiamo lo stesso rumore), vedremo poi esempi di accoppiamenti un po' più astuti.

Il risultato del lemma rimane vero se supponiamo che i processi X^i soddisfino l'equazione $dX_t^i = b^i(X_t)dt + \sigma(X_t^i)dw_t$ (i.e. con lo stessa diffusione), vedi [9, Prop. 5.2.18].

Dimostrazione. Per la continuità di X_t^i è sufficiente mostrare che per ogni $t \in \mathbb{R}_+$ vale $\mathbb{P}(X_t^2 \geq X_t^1) = 1$. Ricordando che per $x \in \mathbb{R}$ la parte positiva è definita da $x_+ := \max\{0, x\}$, abbiamo

$$\begin{aligned} \mathbb{E}(X_t^1 - X_t^2)_+ &= (x^1 - x^2)_+ + \mathbb{E} \int_0^t ds \mathbb{1}_{[0, \infty]}(X_s^1 - X_s^2) [b^1(X_s^1) - b^2(X_s^2)] \\ &= (x^1 - x^2)_+ + \mathbb{E} \int_0^t ds \mathbb{1}_{[0, \infty]}(X_s^1 - X_s^2) [b^1(X_s^1) - b^2(X_s^1) + b^2(X_s^1) - b^2(X_s^2)] \\ &\leq L \mathbb{E} \int_0^t ds \mathbb{1}_{[0, \infty]}(X_s^1 - X_s^2) (X_s^1 - X_s^2) = L \int_0^t ds \mathbb{E}(X_s^1 - X_s^2)_+ \end{aligned}$$

in cui abbiamo usato che, per ipotesi, $(x^1 - x^2)_+ = 0$. Applicando il lemma di Gronwall nella precedente disuguaglianza troviamo $\mathbb{E}(X_t^1 - X_t^2)_+ = 0$. \square

Dimostrazione teorema 2.25. Per la monotonia di $\lambda \mapsto \lambda x_-$ ed il lemma 2.26 abbiamo che $\lambda_1 \leq \lambda_2$ implica $X^{\lambda_1} \leq X^{\lambda_2}$ \mathbb{P} -q.s. Affermiamo che vale inoltre $X_t^\lambda \leq X_t$ per ogni $t \geq 0$. Questo si verifica come segue. Osserviamo che $X \geq 0$; preso $t \geq 0$, se $X_t^\lambda \leq 0$ non dobbiamo fare nulla, altrimenti introducendo $\tau = \sup\{s \in [0, t] : X_s^\lambda = 0\}$ troviamo

$$X_t^\lambda = X_t^\lambda - X_\tau^\lambda = w_t - w_\tau + \int_\tau^t ds \lambda (X_s^\lambda)_- = w_t - w_\tau \leq X_t$$

dove abbiamo usato $(X_s^\lambda)_- = 0$ per $s \in [\tau, t]$.

Grazie alla monotonia e limitatezza, esiste un processo Y_t , $t \in \mathbb{R}_+$ tale che $X^\lambda \uparrow Y$ per $\lambda \rightarrow \infty$. Per completare la dimostrazione faremo vedere che \mathbb{P} -q.s.: Y è continuo, $Y \geq 0$ ed esiste un processo continuo crescente ℓ per cui $Y_t = w_t + \ell_t$ con $\int_0^\infty d\ell(t) Y_t = 0$. Dalle precedenti proprietà, grazie al lemma 2.23, concludiamo che $Y = X$

Continuità di Y_t . Ricordiamo che per $f \in C(\mathbb{R}_+)$, $\delta > 0$ e $T > 0$, il modulo di continuità di f nell'intervallo $[0, T]$ è definito come

$$\omega_{\delta, T}(f) := \sup_{\substack{s, t \in [0, T] \\ |t-s| < \delta}} |f(t) - f(s)|.$$

Cominciamo con il mostrare la *stima a priori*

$$\inf_{t \in [0, T]} X_t^\lambda \geq -2\omega_{\delta, T}(w) - 4e^{-\lambda\delta} \sup_{t \in [0, T]} |w(t)|. \quad (2.42) \quad \boxed{\text{ygs1}}$$

che ci mostra come X_t^λ non possa diventare “troppo” negativo.

Sia $\tau \in [0, T]$ tale che $\inf_{t \in [0, T]} X_t^\lambda = X_\tau^\lambda$. Se $X_\tau^\lambda = 0$ non c'è nulla da dimostrare, altrimenti sia $\sigma = \sup\{t \in [0, \tau] : X_t^\lambda = 0\}$. Per $t \in [\sigma, \tau]$ abbiamo

$$X_t^\lambda = X_\sigma^\lambda - \lambda \int_\sigma^t ds X_s^\lambda + w_t - w_\sigma$$

che possiamo integrare facilmente ottenendo

$$\begin{aligned}
X_\tau^\lambda &= w_\tau - w_\sigma - \int_\sigma^\tau ds \lambda e^{-\lambda(\tau-s)} [w_s - w_\sigma] \\
&= e^{-\lambda(\tau-\sigma)} [w_\tau - w_\sigma] + \int_\sigma^\tau ds \lambda e^{-\lambda(\tau-s)} [w_\tau - w_s] \\
&= e^{-\lambda(\tau-\sigma)} [w_\tau - w_\sigma] + \int_\sigma^{\sigma \vee (\tau-\delta)} ds \lambda e^{\lambda(\tau-s)} [w_\tau - w_s] \\
&\quad + \int_{\sigma \vee (\tau-\delta)}^\tau ds \lambda e^{-\lambda(\tau-s)} [w_\tau - w_s] \\
&\geq -4 e^{-\lambda\delta} \sup_{t \in [0, T]} |w_t| - 2 \omega_{\delta, T}(w).
\end{aligned}$$

Possiamo ora stimare il modulo di continuità di X^λ . Affermiamo che

$$\omega_{\delta, T}(X^\lambda) \leq 8 \left[\omega_{\delta, T}(w) + e^{-\lambda\delta} \sup_{t \in [0, T]} |w_t| \right]. \quad (2.43) \quad \boxed{\text{ygs2}}$$

Siano $t, s \in [0, T]$ con $|t - s| < \delta$. Consideriamo prima il caso in cui $X_u^\lambda \leq 0$ per ogni $u \in [s, t]$. Come prima possiamo integrare l'equazione e troviamo

$$X_t^\lambda - X_s^\lambda = (e^{-\lambda(t-s)} - 1) X_s^\lambda + w_t - w_s - \int_s^t du \lambda e^{-\lambda(t-u)} [w_u - w_s],$$

cosicché, grazie a (2.42),

$$|X_t^\lambda - X_s^\lambda| \leq -X_s^\lambda + 2 \omega_{\delta, T}(w) \leq 4 \left[\omega_{\delta, T}(w) + e^{-\lambda\delta} \sup_{t \in [0, T]} |w_t| \right]. \quad (2.44) \quad \boxed{\text{ygs3}}$$

Se invece $X_u^\lambda \geq 0$ per ogni $u \in [s, t]$ troviamo subito $|X_t^\lambda - X_s^\lambda| \leq \omega_{\delta, T}(w)$. Gli altri casi si riconducono facilmente ai precedenti; vediamo come funziona se $X_s^\lambda < 0$, $X_t^\lambda < 0$. Sia $\sigma = \inf\{u > s : X_u^\lambda = 0\}$ e $\tau = \sup\{u < t : X_u^\lambda = 0\}$. Basta allora scrivere

$$|X_t^\lambda - X_s^\lambda| = |X_t^\lambda - X_\tau^\lambda| + |X_\sigma^\lambda - X_s^\lambda|$$

ed utilizzare (2.44) negli intervalli $[s, \sigma]$ e $[\tau, t]$ per ottenere (2.43).

Prendendo il limite $\lambda \rightarrow \infty$ in (2.43) troviamo che il processo limite Y è \mathbb{P} -q.s. continuo. Per la convergenza monotona di X^λ a Y abbiamo allora che X^λ converge a Y uniformemente per t in compatti.

Conclusione della dimostrazione. Per verificare che $Y \geq 0$ osserviamo che

$$\int_0^t ds (X_s^\lambda)_- = \frac{1}{\lambda} [X_t^\lambda - w_t],$$

che nel limite $\lambda \rightarrow \infty$ implica $\int_0^\infty ds (Y_s)_- = 0$, ovvero $Y \geq 0$ \mathbb{P} -q.s. per la continuità di Y .

Introduciamo il processo crescente

$$\ell_t^\lambda := \int_0^t ds \lambda(X_s^\lambda)_- = X_t^\lambda - w_t$$

Grazie alla convergenza di X^λ al processo continuo Y

$$\ell_t := \lim_{\lambda \rightarrow \infty} \ell_t^\lambda = Y_t - w_t$$

è un processo continuo crescente. In particolare la misura di Lebesgue-Stieltjes $d\ell^\lambda$ converge debolmente a $d\ell$ per $\lambda \rightarrow \infty$. Rimane da mostrare che $\int_0^\infty d\ell_t Y_t = 0$; osserviamo che il supporto della misura $d\ell^\lambda$ è un sottoinsieme di $\{t \geq 0 : X_t^\lambda \leq 0\}$. Grazie alla convergenza uniforme di X^λ a Y e la convergenza debole di $d\ell^\lambda$ a $d\ell$, troviamo, per ogni $T > 0$

$$\int_0^T d\ell_t Y_t = \lim_{\lambda \rightarrow \infty} \int_0^T d\ell_t^\lambda X_t^\lambda \leq 0,$$

che conclude la dimostrazione poiché $Y \geq 0$. □

Note bibliografiche

Una teoria abbastanza completa delle equazioni stocastiche si può trovare in [9, Ch. 5], per ulteriori dettagli vedi [8] in cui vengono anche considerate equazioni stocastiche su varietà. Per un punto di vista meno astratto è invece consigliato [4]; in particolare vi si possono trovare i teoremi di continuità rispetto ai dati iniziali e condizioni che sostituiscono l'ipotesi (di gran lunga troppo restrittiva nelle applicazioni) di crescita lineare per $x \rightarrow \infty$ di $b(x)$ e $\sigma(x)$. La formulazione delle equazioni stocastiche come problema alle martingale è dovuta a Stroock e Varadhan [15], vedi anche [9, §5.5.4] in cui non sono però dimostrati i risultati di equazioni alle derivate parziali utilizzati. Vedi [9, §5.5.6] per il processo di Ornstein–Uhlenbeck n -dimensionale. La sezione 2.6 è una soluzione di [9, Ex. 5.2.26], vedi [11] per una discussione più ampia (compresa la motivazione fisica). Nel caso in cui la diffusione è non costante l'approssimazione di Smoluchowski non è così facile, vedi [7]. I tempi locali per il moto browniano sono stati introdotti da Levy. La teoria sviluppata in [9, § 3.6] definisce il tempo locale mediante (2.40), qui si è invece seguito l'approccio in [12, Ch. VI]. La dimostrazione del teorema 2.25 (suggerita da L. Zambotti) è tratta da [1].

S:m

3. Processi di Markov

I processi di Markov sono un modello matematico per un'evoluzione aleatoria in cui lo stato del sistema al tempo t (futuro) dipende dalla storia dal tempo 0 al tempo s (passato) solo tramite lo stato del sistema al tempo s (presente). Sono quindi esclusi tutti i fenomeni con "isteresi". Per specificare un processo di Markov basta assegnare la probabilità delle possibili transizioni da uno stato x al tempo s allo stato y al tempo t . Considereremo solo il caso di *processi di Markov omogenei* in cui tale probabilità dipende solo da $t - s$. Come vedremo, il moto browniano e, più in generale la soluzione di equazioni stocastiche sono esempi di processi di Markov. Un esempio più elementare è dato dalla passeggiata aleatoria. Cominciamo dal caso di *catene di Markov* (tempo discreto) con spazio degli stati finito o numerabile.

S:mc

3.1. Catene di Markov

Sia S un insieme finito o numerabile. Un'evoluzione (a tempo discreto) deterministica su S è semplicemente un'applicazione $F : S \rightarrow S$ con cui costruiamo, data una condizione iniziale $x \in S$, la successione $X_{n+1} = F(X_n)$, $n \geq 0$, con $X_0 = x$. Un esempio di processo di Markov è ottenuto aggiungendo un po' di stocasticità al modo seguente. Sia U_n , $n \geq 0$ una successione di variabili i.i.d. con distribuzione uniforme in $[0, 1]$ ed $F : S \times [0, 1] \rightarrow S$ un'applicazione misurabile. Dato $x \in S$ consideriamo la successione di variabili aleatorie (a valori in S) definita da

$$X_{n+1} = F(X_n, U_n), \quad X_0 = x \quad (3.1) \quad \text{mcm}$$

appare allora evidente che lo stato al tempo $n + 1$ (ovvero X_{n+1}) dipende dalla storia passata (le U_n rappresentano rumore "fresco") solo tramite X_n . Vedremo come questo esempio rappresenti in realtà la generica catena di Markov.

Vediamo ora come si costruisce una catena di Markov con probabilità di transizione data. Sia $\nu \in \mathcal{P}(S)$ una probabilità su S , ovvero una successione di numeri $\nu(x)$, $x \in S$ con $\nu(x) \geq 0$ e $\sum_x \nu(x) = 1$ e $p_{\cdot, \cdot} : S \times S \rightarrow [0, 1]$ una probabilità di transizione, ovvero

$$p_{x,y} \geq 0 \quad \forall x, y \in S \quad \sum_y p_{x,y} = 1 \quad \forall x \in S \quad (3.2) \quad \text{ms}$$

che rappresenta la probabilità di andare (in tempo 1) da x a y .

Sia $\Omega := S^{\mathbb{N}}$ la collezione delle traiettorie (a tempo discreto) a valori in S . Chiamiamo X_n , $n \geq 0$ la coordinata canonica in Ω , ovvero $X_n(\omega) = \omega_n$, $\omega = \{\omega_n\}_{n \geq 0} \in \Omega$. Consideriamo Ω equipaggiato con la σ -algebra dei cilindri $\sigma\{X_0, \dots, X_n, \dots\}$ e costruiamo la probabilità \mathbb{P}_ν su Ω tale che sul cilindro $\{X_0 = x_0, \dots, X_n = x_n\}$ valga

$$\mathbb{P}_\nu(X_0 = x_0, \dots, X_n = x_n) = \nu(x_0) p_{x_0, x_1} \cdots p_{x_{n-1}, x_n} \quad (3.3) \quad \text{rcm}$$

chiamiamo \mathbb{P}_ν il processo di Markov con probabilità di transizione p e stato iniziale ν . Se $\nu = \delta_x$, con $x \in S$ scriviamo \mathbb{P}_x invece di \mathbb{P}_{δ_x} .

Con una piccola ambiguità, diciamo anche che una successione di variabili aleatorie $X = \{X_0, X_1, \dots\}$ (definite su un qualunque spazio di probabilità con valori in S) la cui legge sia \mathbb{P}_ν è un processo di Markov con probabilità di transizione p e stato iniziale ν . Più propriamente la definizione (3.3) si dovrebbe chiamare *realizzazione canonica* del processo di Markov; per il momento non facciamo però troppa differenza tra un processo di Markov e la sua realizzazione canonica.

Un semplice calcolo mostra come dalla definizione (3.3) segua la proprietà di Markov, ovvero per ogni $y \in S$

$$\mathbb{P}_\nu(X_{n+1} = y | X_0, \dots, X_n) = \mathbb{P}_\nu(X_{n+1} = y | X_n) = p_{X_n, y} \quad (3.4) \quad \boxed{\text{pm}}$$

che formalizza il fatto che lo stato al tempo $n+1$ (futuro) dipenda dalla storia precedente (passato) solo attraverso X_n (presente).

Nelle definizioni assiomatiche di catena di Markov (qui si è privilegiato l'aspetto costruttivo) è data una successione $X = \{X_0, X_1, \dots\}$ di variabili aleatorie a valori in S per cui vale (con probabilità uno) la prima uguaglianza in (3.4). La seconda uguaglianza diviene allora la definizione della probabilità di transizione $p_{x,y}$ e, detta ν la legge di X_0 , la legge di X soddisfa (3.3).

La proprietà di Markov nella forma (3.4), così come la costruzione (3.3), privilegia una direzione del tempo; in realtà proprietà di Markov si può formulare in modo simmetrico tra passato e futuro. Siano $n \geq 1$ (il presente), $A \in \sigma\{X_0, \dots, X_{n-1}\}$ (un evento nel passato) e $B \in \sigma\{X_{n+1}, \dots\}$ (un evento nel futuro). La proprietà di Markov si può allora enunciare in modo equivalente (verificare) dicendo che, dato il presente, passato e futuro sono indipendenti, ovvero

$$\mathbb{P}_\nu(A \cap B | X_n) = \mathbb{P}_\nu(A | X_n) \cdot \mathbb{P}_\nu(B | X_n)$$

Vediamo come ogni catena di Markov, nel senso di (3.3), si possa rappresentare nella forma (3.1) per un'opportuna F .

t:rcm

Teorema 3.1. *Sia X una catena di Markov con probabilità di transizione p e stato iniziale ν ; allora esiste $F : S \times [0, 1] \rightarrow S$ per cui X si può rappresentare come in (3.1) ove x è una variabile aleatoria, indipendente dalle U_n con legge ν .*

Dimostrazione. Lo stato iniziale ν è già stato incorporato in (3.1), dobbiamo ora costruire la funzione F a partire dalla probabilità di transizione p . Chiamiamo $(\Omega, \mathcal{F}, \mathbb{P})$ lo spazio di probabilità ove sono definite le v.a. x , con legge ν , e U_n , con legge uniforme in $[0, 1]$; tutte indipendenti. Dato $x \in S$, possiamo costruire una partizione $I_{x,y}$, $y \in S$ di $[0, 1]$ in modo che $\text{Leb}(I_{x,y}) = p_{x,y}$. Poiché S è numerabile lo possiamo identificare con \mathbb{N} , si può allora prendere $I_{x,1} = [0, p_{x,1})$, $I_{x,2} = [p_{x,1}, p_{x,1} + p_{x,2})$ e così via. A questo punto dichiariamo

$$F(x, u) := \sum_y y \mathbb{1}_{I_{x,y}}(u)$$

Si tratta ora di verificare che la catena X definita da (3.1) è Markov (ovvio) con probabilità di transizione p . Si ha

$$\mathbb{P}(X_{n+1} = y | X_n = x) = \mathbb{P}(F(x, U_n) = y) = \mathbb{P}(U_n \in I_{x,y}) = \text{Leb}(I_{x,y}) = p_{x,y}$$

che conclude la dimostrazione. \square

Data una catena di Markov X gli associamo naturalmente un semigrupp (markoviano) sulle funzioni su S (osservabili) ed il semigrupp duale sulle probabilità su S .

Per $f : S \rightarrow \mathbb{R}$ introduciamo l'operatore lineare

$$Pf(x) := \sum_y p_{x,y} f(y)$$

Consideriamo ora $f_n(x) := \mathbb{E}_x(f(X_n))$ ovvero il valore medio di f per la catena al tempo n . Qui \mathbb{E}_x è il valore di attesa rispetto a \mathbb{P}_x , grazie alla proprietà di Markov troviamo subito

$$f_n(x) = \sum_y \mathbb{P}_x(X_1 = y) \mathbb{E}_x(f(X_n) | X_1 = y) = \sum_y p_{x,y} \mathbb{E}_y(f(X_{n-1})) = Pf_{n-1}(x) \quad (3.5) \quad \boxed{\text{kin}}$$

e quindi $f_n = P^n f$.

Indichiamo con P anche l'operatore aggiunto che agisce su $\nu \in \mathcal{P}(S)$ come

$$\nu P(y) := \sum_x \nu(x) p_{x,y}$$

Prendiamo ora la catena di Markov con stato iniziale $\nu \in \mathcal{P}(S)$ ed indichiamo con ν_n la legge di X_n , ovvero della stato al tempo n . Allora, sempre per la proprietà di Markov,

$$\begin{aligned} \nu_n(y) &= \mathbb{P}_\nu(X_n = y) = \sum_x \mathbb{P}_\nu(X_{n-1} = x) \mathbb{P}_\nu(X_n = y | X_{n-1} = x) \\ &= \sum_x \nu_{n-1}(x) p_{x,y} = \nu_{n-1} P(y) \end{aligned} \quad (3.6)$$

Notiamo come per ottenere l'evoluzione di valori medi di funzioni abbiamo condizionato rispetto X_1 mentre per l'evoluzione delle misure rispetto a X_{n-1} . Per questo motivo ci si riferisce a (3.5), rispettivamente a (3.6), come equazioni di Kolmogorov in indietro, rispettivamente in avanti. Se avessimo considerato probabilità di transizione dipendenti dal tempo, tale terminologia sarebbe stata più chiara.

Osserviamo infine come (3.5) e (3.6) siano ottenibili l'una dall'altra per dualità:

$$\nu_n(f) := \sum_y \nu_n(y) f(y) = \mathbb{E}_\nu(f(X_n)) = \sum_x \nu(x) \mathbb{E}_x(f(X_n)) = \sum_x \nu(x) f_n(x) =: \nu(f_n)$$

Nel contesto delle dinamiche deterministiche sono rilevanti le soluzioni stazionarie (posizioni di equilibrio); nel constesto delle catene di Markov, se la probabilità di transizione non degenera, non vi è modo di rimanere sempre sullo stesso stato; definiamo allora una misura invariante μ richiedendo che, qualora lo stato iniziale sia μ allora la legge di X_1 rimane μ e quindi, per la proprietà di Markov X_n avrà legge μ per ogni $n \geq 0$. Formalmente diciamo che $\mu \in \mathcal{P}(S)$ è una *misura invariante* (o misura stazionaria) sse $\mu = \mu P$.

Il primo problema naturale riguarda l'esistenza di misure invarianti, facciamo vedere che a questo fine è sufficiente una proprietà di compattezza.

t:emi **Lemma 3.2.** *Sia S un insieme finito e X una catena di Markov, allora esiste una misura invariante per X .*

Dimostrazione. Prendiamo $\nu \in \mathcal{P}(S)$ e, dato $N \geq 1$, consideriamo la somma ergodica

$$\nu^{(N)} := \frac{1}{N} \sum_{n=0}^{N-1} \nu_n = \frac{1}{N} \sum_{n=0}^{N-1} \nu P^n$$

poichè S è compatto, possiamo estrarre una sottosuccessione convergente da $\nu^{(N)} \in \mathcal{P}(S)$ (visto che S è un insieme finito tutte le nozioni di convergenza coincidono). Sia $\mu \in \mathcal{P}(S)$ un punto limite, passando al limite lungo la sottosuccessione convergente nella relazione

$$\nu^{(N)} P := \frac{1}{N} \sum_{n=0}^{N-1} \nu P^{n+1} = \nu^{(N)} + \frac{1}{N} [\nu_0 - \nu_N]$$

troviamo $\mu P = \mu$, che dimostra l'invarianza di μ . □

Stabilita l'esistenza di misure invarianti, discutiamo l'unicità e le proprietà di convergenza di X_n alla misura invariante nel limite $n \rightarrow \infty$. Nel caso di spazio degli stati finiti, $|S| < \infty$ vediamo come si ottiene, suppondo abbastanza aleatorietà nella catena, una risposta completa. Diciamo che una catena di Markov è *irriducibile* se esistono $m \geq 1$ e $\varepsilon > 0$ tali che

$$\inf_{x,y \in S} (P^m)_{x,y} \geq \varepsilon \tag{3.7} \quad \text{irr}$$

ovvero possiamo andare da uno stato ad ogni altro in m passi con probabilità strettamente positiva. Vediamo allora il *teorema ergodico* per catene di Markov su spazio degli stati finiti. Per $\nu, \mu \in \mathcal{P}(S)$ introduciamo la norma in variazione totale come

$$\|\mu - \nu\|_{\text{TV}} := \sum_{x \in S} |\mu(x) - \nu(x)| = 2 \sup_{A \subset S} |\mu(A) - \nu(A)| \tag{3.8} \quad \text{vt}$$

Esercizio: dimostrare la seconda identità in (3.8).

tefmc **Teorema 3.3.** *Sia $X_n, n \geq 0$ una catena di Markov irriducibile. Allora esiste un'unica misura invariante $\mu \in \mathcal{P}(S)$. Inoltre $\mu(x) > 0$ per ogni $x \in S$ ed esiste $\gamma = \gamma(m, \varepsilon) \in (0, 1)$ tale che per ogni $\nu \in \mathcal{P}(S)$ vale*

$$\|\nu P^n - \mu\|_{\text{TV}} \leq \gamma^n \tag{3.9} \quad \text{civt}$$

Osserviamo come la stima (3.9) dica che, uniformemente nello stato iniziale ν , la legge della variabile aleatoria X_n converga, in modo esponenziale, in variazione totale alla misura invariante μ .

Il teorema precedente ha un'interpretazione puramente analitica, è allora noto come *teorema di Perron–Frobenius*. Data una matrice stocastica $p_{x,y}$, soddisfacente cioè (3.2), irriducibile allora uno è autovalore sinistro e l'autovettore corrispondente può essere scelto con tutte le componenti strettamente positive.

Dimostrazione. Dimostreremo la seguente stima: esiste $\gamma \in (0, 1)$ tale che per ogni $y, y' \in S$ si ha

$$\|\delta_y P^n - \delta_{y'} P^n\|_{\text{TV}} \leq \gamma^n \quad (3.10)$$

bcivt

La precedente disuguaglianza mostra che partendo da due stati diversi, y e y' la distribuzione del corrispondente stato al tempo n differisce di poco; è questa *perdita di memoria* il meccanismo per l'ergodicità in catene irriducibili.

Vediamo prima come le affermazioni del teorema seguano da (3.10). Introduciamo $\underline{\mu}_n(x) := \inf_y (P^n)_{y,x}$ e $\bar{\mu}_n(x) := \sup_y (P^n)_{y,x}$. È immediato verificare che $\underline{\mu}_n(x)$ è una successione crescente mentre $\bar{\mu}_n(x)$ è decrescente. Esiste quindi il limite di entrambe, che indichiamo con $\underline{\mu}$ e $\bar{\mu}$. In virtù di (3.10) abbiamo ora $\underline{\mu} = \bar{\mu} =: \mu$ e vale $\lim_{n \rightarrow \infty} (P^n)_{y,x} = \mu(x)$ uniformemente in y ; è ora facile verificare che μ è l'unica misura invariante della catena, soddisfa $\mu(x) > 0$ per ogni $x \in S$. Infine la stima (3.9) segue da (3.10).

Per dimostrare la perdita di memoria codificata dalla stima (3.10), l'idea è di considerare insieme le due catene di Markov con stati iniziali y e y' (ovvero costruirle sul medesimo spazio di probabilità) in modo che si mettano insieme con buona probabilità. In modo più formale, date le due catene \mathbb{P}_y e $\mathbb{P}_{y'}$ (due probabilità su $S^{\mathbb{N}}$), consideriamo lo spazio prodotto $S \times S$; diciamo che un *accoppiamento* delle misure \mathbb{P}_y e $\mathbb{P}_{y'}$ è una probabilità $\mathbb{Q}_{y,y'}$ su $(S \times S)^{\mathbb{N}}$ i cui marginali sono \mathbb{P}_y e $\mathbb{P}_{y'}$. Vediamo come costruendo un'accoppiamento si ha naturalmente una stima sulla distanza in variazione totale. Chiamiamo (X_n, Y_n) la coordinata canonica per $(S \times S)^{\mathbb{N}}$

$$\begin{aligned} |\delta_y P^n(x) - \delta_{y'} P^n(x)| &= |(P^n)_{y,x} - (P^n)_{y',x}| = |\mathbb{P}_y(X_n = x) - \mathbb{P}_{y'}(X_n = x)| \\ &= |\mathbb{Q}_{y,y'}(X_n = x) - \mathbb{Q}_{y,y'}(Y_n = x)| \\ &= |\mathbb{Q}_{y,y'}(X_n = x, Y_n \neq x) - \mathbb{Q}_{y,y'}(X_n \neq x, Y_n = x)| \end{aligned}$$

da cui

$$\begin{aligned} \|\delta_y P^n - \delta_{y'} P^n\|_{\text{TV}} &= \sum_x |\mathbb{Q}_{y,y'}(X_n = x, Y_n \neq x) - \mathbb{Q}_{y,y'}(X_n \neq x, Y_n = x)| \\ &\leq 2 \mathbb{Q}_{y,y'}(X_n \neq Y_n) \end{aligned} \quad (3.11)$$

Si tratta ora di costruire la misura $\mathbb{Q}_{y,y'}$ in modo che la probabilità che $X_n = Y_n$ tenda a uno per $n \rightarrow \infty$. L'idea è abbastanza semplice. Costruiamo (X_n, Y_n) , $n \geq 0$ come una catena di Markov su $S \times S$ con la seguente regola: X e Y si muovono indipendentemente (e quindi ognuno con probabilità di transizione p) finché sono diversi, quando sono uguali si muovono invece insieme. Formalizziamo quest'idea nella definizione

$$q_{(y,y'),(x,x')} := \begin{cases} p_{y,x} p_{y',x'} & \text{se } y \neq y' \\ p_{y,x} \delta_{x,x'} & \text{se } y = y' \end{cases}$$

è immediato verificare che q è una probabilità di transizione su $S \times S$. I marginali di q sono inoltre $\sum_{x'} q_{(y,y'),(x,x')} = p_{y,x}$ e $\sum_x q_{(y,y'),(x,x')} = p_{y',x'}$. Se dichiariamo $\mathbb{Q}_{y,y'}$ come la catena con dato iniziale (y, y') , segue che è un accoppiamento tra \mathbb{P}_y e $\mathbb{P}_{y'}$. Infatti, ad esempio,

$$\mathbb{Q}_{y,y'}(X_1 = x) = \sum_{x'} q_{(y,y'),(x,x')} = p_{y,x}$$

L'ipotesi di irriducibilità (3.7) ci garantirà che la catena $\mathbb{Q}_{y,y'}$ ha, uniformemente nei dati iniziali, probabilità maggiore di ε di far attaccare X e Y in m passi; useremo poi la proprietà di Markov e (3.11) per dedurre da ciò la stima (3.10). La morale è che se un'evento ha probabilità strettamente positiva di capitare in tempo finito allora prima o poi capita di certo. Qui non sono però variabili indipendenti (in cui è quasi ovvio), ma catene di Markov: sarà essenziale l'uniformità rispetto al dato iniziale in (3.7). Vediamo i dettagli.

Rispetto alla catena congiunta introduciamo il tempo di arresto τ come il primo tempo in cui le due copie finiscono nello stesso stato, i.e. $\tau := \inf\{n \geq 0 : X_n = Y_n\}$. È facile verificare che (3.7) implica

$$\mathbb{Q}_{y,y'}(\tau \leq m) \geq \sum_x \mathbb{Q}_{y,y'}(X_m = Y_m = x) \geq \sum_x (P^m)_{y,x} (P^m)_{y',x} \geq \varepsilon$$

Utilizzando la proprietà di Markov troviamo allora

$$\begin{aligned} \mathbb{Q}_{y,y'}(\tau > (k+1)m) &= \mathbb{Q}_{y,y'}\left(\mathbb{Q}_{y,y'}(\tau > (k+1)m \mid (X_{km}, Y_{km}))\right) \\ &= \mathbb{Q}_{y,y'}\left(\mathbb{Q}_{X_{km}, Y_{km}}(\tau > m)\right) \\ &= \mathbb{Q}_{y,y'}\left(\mathbb{1}_{X_{km} \neq Y_{km}} \mathbb{Q}_{X_{km}, Y_{km}}(\tau > m)\right) \\ &\leq \sup_{x,x'} \mathbb{Q}_{x,x'}((\tau > m) \mid X_{km} \neq Y_{km}) \end{aligned}$$

da cui troviamo

$$\mathbb{Q}_{y,y'}(\tau > (k+1)m) \leq (1 - \varepsilon) \mathbb{Q}_{y,y'}(\tau > km)$$

che implica la stima esponenziale (3.10). \square

Vediamo un altro esempio in cui può essere utile costruire un accoppiamento tra due catene di Markov con dati iniziali diversi. Consideriamo il caso in cui S sia numerabile. Diciamo che una funzione $f : S \rightarrow \mathbb{R}$ è armonica per la probabilità di transizione p sse si ha $Pf = f$, ovvero

$$\sum_y p_{x,y} [f(y) - f(x)] = 0 \quad \forall x \in S \quad (3.12) \quad \boxed{\text{far}}$$

Se pensiamo alla passeggiata aleatoria semplice su \mathbb{Z}^d , definita da $p_{x,y} = 1/(2d)$ se $|x-y| = 1$ e $p_{x,y} = 0$ altrimenti, l'equazione (3.12) si riduce all'equazione di Laplace discreta e giustifica il nome dato.

Vogliamo trovare delle condizioni che implicino un analogo del teorema di Liouville che dice che le sole funzioni armoniche limitate sono le costanti. Vedremo come l'esistenza di un accoppiamento "riuscito" garantisca tale proprietà. Chiamiamo accoppiamento (markoviano) della catena una probabilità di transizione $q_{(y,y'),(x,x')}$ su $S \times S$ i cui marginali siano rispettivamente $p_{y,x}$ e $p_{y',x'}$. Diciamo inoltre che è un accoppiamento con *successo* sse per ogni $y, y' \in S$ si ha, con le stesse notazioni introdotte in precedenza,

$$\mathbb{Q}_{y,y'}(X_n = Y_n \text{ definitivamente}) = \mathbb{Q}_{y,y'}\left(\bigcup_n \bigcap_{k \geq n} \{X_k = Y_k\}\right) = 1$$

t:liouv

Teorema 3.4. *Sia p una probabilità di transizione per cui esiste un accoppiamento con successo, allora ogni funzione armonica limitata è costante.*

Dimostrazione. Se f è una funzione armonica limitata, per ogni $n \geq 0$ abbiamo

$$f(x) = P^n f(x) = \mathbb{E}_x(f(X_n))$$

Indicando con $\tilde{\mathbb{E}}_{y,y'}$ il valore di attesa rispetto all'accoppiamento $\mathbb{Q}_{y,y'}$ e con $\tau := \inf\{n \geq 0 : X_k = Y_k \forall k \geq n\}$ il tempo in cui ha successo, otteniamo subito

$$\begin{aligned} |f(x) - f(x')| &= \left| \mathbb{E}_x(f(X_n)) - \mathbb{E}_{x'}(f(X_n)) \right| \\ &= \left| \tilde{\mathbb{E}}_{x,x'}(f(X_n) - f(Y_n)) \right| \leq 2 \sup_x |f(x)| \mathbb{Q}_{x,x'}(\tau > n) \end{aligned}$$

e prendendo il limite $n \rightarrow \infty$ concludiamo la dimostrazione. \square

Utilizzando la ricorrenza della passeggiata aleatoria semplice in una o due dimensioni è immediato verificare che l'accoppiamento costruito nella dimostrazione di teorema 3.2 è con successo in questo caso. Per la passeggiata aleatoria semplice in \mathbb{Z}^d , con $d \geq 3$, una piccola variante di tale accoppiamento funziona: basta scegliere prima una coordinata a caso, se il valore di tale coordinata è diverso la si cambia indipendentemente per X e Y altrimenti tale coordinata cambia allo stesso modo per X e Y . Esercizio: scrivere per bene questo accoppiamento e verificare che ha successo.

Consideriamo una probabilità di transizione p su S numerabile e consideriamo il corrispondente problema di Dirichlet. Dato $D \subset S$ ed un dato al bordo $\varphi : D^c \rightarrow \mathbb{R}$ vogliamo trovare $f : S \rightarrow \mathbb{R}$ tale che

$$\begin{cases} Pf(x) = f(x) & x \in D \\ f(x) = \varphi(x) & x \notin D \end{cases} \quad (3.13) \quad \text{pdir}$$

Notiamo come nel caso della passeggiata aleatoria semplice in \mathbb{Z}^d questa sia proprio l'equazione di Laplace discreta. Se consideriamo il caso di φ limitata, possiamo immediatamente scrivere una soluzione di (3.13) in termini della catena di Markov X .

t:rpdd

Teorema 3.5. *Sia $\tau := \inf\{n \geq 0 : X_n \notin D\}$ il tempo di uscita da D e $\varphi : D^c \rightarrow \mathbb{R}$ limitata; allora la funzione*

$$f(x) = \mathbb{E}_x\left(\varphi(X_\tau) \mathbb{1}_{\{\tau < \infty\}}\right) \quad (3.14) \quad \text{spdd}$$

è una soluzione di (3.13).

La rappresentazione (3.14) mostra come possiamo ottenere $f(x)$ facendo partire la catena da x , guardare dove esce da D e mediare (sulle tutte possibili traiettorie) il corrispondente valore di φ . Vediamo inoltre come il dato al bordo φ vada assegnato solo laddove la catena ha probabilità strettamente positiva di arrivare. È altresì chiaro che in

questa generalità non abbiamo alcuna informazione su τ la catena potrebbe uscire da D (prima o poi) con probabilità uno oppure rimanervi intrappolata indefinitivamente.

Dimostrazione. La dimostrazione consiste in un calcolo diretto. Cominciamo con lo stabilire precisamente cosa si intende per X_τ , è definito solo sull'evento $\{\tau < \infty\}$ come

$$(X_\tau)(\omega) := \sum_{n=0}^{\infty} \mathbb{1}_{\{\tau=n\}}(\omega) X_n(\omega) = \sum_{n=0}^{\infty} X_n(\omega) \prod_{k=0}^{n-1} \mathbb{1}_D(X_k(\omega)) \mathbb{1}_{D^c}(X_n(\omega))$$

Il fatto che f definita da (3.14) soddisfi il dato al bordo in (3.13) è ovvio, vediamo che soddisfa anche l'equazione per $x \in D$. Condizionando rispetto a X_1 ed utilizzando la proprietà di Markov troviamo (se $x \in D$ allora $\tau \geq 1$)

$$\begin{aligned} f(x) &= \sum_{n=1}^{\infty} \mathbb{E}_x \left(\varphi(X_n) \mathbb{1}_{\{\tau=n\}} \right) \\ &= \sum_{n=1}^{\infty} \sum_y p_{x,y} \mathbb{E}_x \left(\varphi(X_n) \mathbb{1}_{\{\tau=n\}} \middle| X_1 = y \right) \\ &= \sum_y p_{x,y} \sum_{n=1}^{\infty} \mathbb{E}_y \left(\varphi(X_{n-1}) \mathbb{1}_{\{\tau=n-1\}} \right) \\ &= \sum_y p_{x,y} f(y) = Pf(x) \end{aligned}$$

che conclude la verifica. □

Un calcolo analogo a quello fatto in precedenza mostra come la funzione

$$u(x) := \mathbb{P}_x(\tau = \infty) \tag{3.15} \quad \boxed{\text{usd}}$$

risolve il problema (3.13) con $\varphi = 0$. In particolare se accade che $\mathbb{P}_x(\tau = \infty) > 0$ per qualche $x \in D$ non ci può essere unicità nel problema di Dirichlet.

Se, al contrario, per ogni $x \in D$ accade che la catena esce da D con probabilità uno, l'ottimismo ci fa pensare che esista un'unica soluzione di (3.13) che si rappresenti come in (3.14). Vediamo che questo succede effettivamente, almeno per soluzioni limitate. *Esercizio:* costruire un esempio in cui vi è un'unica soluzione limitata, ma non c'è unicità in quelle non limitate.

Teorema 3.6. *Sia $\mathbb{P}_x(\tau = \infty) = 0$ per ogni $x \in D$; allora l'unica soluzione limitata di (3.13) è data da (3.14).*

Premettiamo il seguente lemma generale, che mostra come sia possibile esibire delle martingale a partire da una catena di Markov.

Lemma 3.7. *Sia X una catena di Markov con probabilità di transizione p e g una funzione limitata su S ; allora la successione*

$$M_n^g := g(X_n) - g(X_0) - \sum_{k=0}^{n-1} [Pg(X_k) - g(X_k)]$$

è una martingala (rispetto alla filtrazione naturale $\mathcal{F}_n := \sigma\{X_0, \dots, X_n\}$).

Dimostrazione. Si deve verificare che per ogni $n \geq m$ accade (l'integrabilità è garantita dalla limitatezza di g)

$$0 = \mathbb{E}(M_n^g - M_m^g | \mathcal{F}_m) = \mathbb{E}\left(g(X_n) - g(X_m) - \sum_{k=m}^{n-1} [Pg(X_k) - g(X_k)] \middle| \mathcal{F}_m\right)$$

Dalla proprietà di Markov segue che per $\ell \geq m$ si ha $\mathbb{E}(g(X_\ell) | \mathcal{F}_m) = P^{\ell-m}g(X_m)$. Pertanto l'ultimo membro nell'equazione precedente è dato da

$$P^{n-m}g(X_m) - g(X_m) - \sum_{k=m}^{n-m-1} [P^{k+1-m}g(X_m) - P^{k-m}g(X_m)] = 0$$

in quanto somma telescopica. □

Dimostrazione del teorema 3.7. Consideriamo la martingala costruita come nel lemma precedente a partire da f , soluzione limitata di (3.13), e la arrestiamo al tempo τ . Poniamo ovvero

$$M_n^{(\tau)} := f(X_{\tau \wedge n}) - f(X_0) - \sum_{k=0}^{(\tau-1) \wedge (n-1)} [Pf(X_k) - f(X_k)]$$

e notiamo che la somma a secondo membro è identicamente nulla poichè per $k \leq \tau - 1$ si ha $X_k \in D$ e quindi $Pf(X_k) = f(X_k)$. Prendendo il valore di attesa rispetto a \mathbb{P}_x con $x \in D$ (verificare che $\mathbb{E}_x(M_n^{(\tau)}) = 0$ anche per la martingala arrestata) troviamo allora

$$f(x) = \mathbb{E}_x(f(X_{\tau \wedge n}))$$

Passando al limite per $n \rightarrow \infty$, notando che l'ipotesi $\tau < \infty$ con \mathbb{P}_x probabilità uno implica $X_{\tau \wedge n} \rightarrow X_\tau \notin D$ ed usando $f(x) = \varphi(x)$ per $x \notin D$, la formula (3.14) segue. □

Consideriamo il caso in cui D non è limitato, X esce con probabilità uno da ogni insieme limitato, ma può rimanere in D con probabilità strettamente positiva. Se consideriamo $\varphi \geq 0$ la formula (3.14) definisce allora la soluzione minimale di (3.13), nel senso che ogni altra soluzione positiva è più grande di quella definita da (3.14). Si capisce: le altre si ottengono sommando la u in (3.15). Esercizio: dimostrare l'affermazione precedente (sugg. il lemma di Fatou può essere utile).

Una formula di rappresentazione analoga a quella ottenuta per il problema di Dirichlet si può ottenere per l'equazione di Poisson. La dimostrazione del teorema seguente è da considerarsi un esercizio.

t:rppd

Teorema 3.8. Sia $D \subset S$ per $g : D \rightarrow \mathbb{R}$ limitata si consideri il problema

$$\begin{cases} Pf(x) - f(x) = -g(x) & x \in D \\ f(x) = 0 & x \notin D \end{cases}$$

Sia τ il primo tempo in cui la catena X esce da D . Se $\mathbb{E}_x(\tau) < \infty$ per ogni $x \in D$, l'unica soluzione limitata è data da

$$f(x) = \mathbb{E}_x\left(\sum_{n=0}^{\tau-1} g(X_n)\right)$$

Nel caso particolare in cui $g = 1$ troviamo l'equazione soddisfatta dal valore di attesa del tempo di uscita. Può essere utile per calcolare esplicitamente tale valore risolvendo l'equazione.

s:pm

3.2. Processi di Markov

Cominciamo nel dare le definizioni rilevanti nel caso di tempo continuo, \mathbb{R}_+ invece di \mathbb{N} , e spazio degli stati S dato da uno spazio polacco (metrico, completo e separabile). Per le applicazioni a processi di diffusione (soluzioni di equazioni differenziali stocastiche) lo spazio degli stati sarà \mathbb{R}^n , ma per un po' consideriamo un caso più generale.

t:pdt

Definizione 3.9. Una probabilità di transizione $p_t(x, B)$ con $t \in \mathbb{R}_+$, $x \in S$ e $B \in \mathcal{B}(S)$ è una funzione (a valori in $[0, 1]$) tale che:

(i) $p_0(x, \cdot) = \delta_x$;

(ii) per ogni $t \in \mathbb{R}_+$, $B \in \mathcal{B}(S)$ la funzione $x \mapsto p_t(x, B)$ è $\mathcal{B}(S)$ misurabile;

(iii) per ogni $t \in \mathbb{R}_+$, $x \in S$ $p_t(x, \cdot)$ è una probabilità su $\mathcal{B}(S)$;

(iv) vale l'equazione di Chapman–Kolmogorov

$$p_{t+s}(x, B) = \int p_t(x, dy) p_s(y, B)$$

L'interpretazione dell'equazione di Chapman–Kolmogorov è ovvia: la probabilità di andare da $(0, x)$ a $(t+s, B)$ la posso ottenere integrando su tutti i possibili modi di andare prima da $(0, x)$ a (t, dy) e poi da (t, y) a $(t+s, B)$.

Diciamo che la probabilità di transizione è *submarkoviana* se $p_t(x, S) < 1$ per qualche t e x . Contempliamo cioè la possibilità che il processo “sparisca” da S .

Data una probabilità di transizione p_t gli associamo il semigruppoo P_t che agisce sulle funzioni limitate $f : S \rightarrow \mathbb{R}$ misurabili (o positive misurabili) come

$$P_t f(x) = \int p_t(x, dy) f(y) \tag{3.16} \quad \text{dsg}$$

Dato il solito spazio di probabilità filtrato $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ diciamo che il processo adattato X_t , $t \geq 0$ (ovvero $X_t \in \mathcal{F}_t$) è Markov (o submarkov) con probabilità di transizione p sse per ogni $f \in \mathcal{B}(S)$ positiva e ogni $s, t \in \mathbb{R}_+$ accade che

$$\mathbb{E}(f(X_{t+s}) | \mathcal{F}_s) = P_t f(X_s) \quad \mathbb{P}\text{-q.s} \tag{3.17} \quad \text{dpm}$$

Osserviamo che se introduciamo la *filtrazione naturale* $\mathcal{F}_t^X := \sigma\{X_s, s \leq t\} \subset \mathcal{F}_t$ (poiché X è adattato) e X è Markov (rispetto a \mathcal{F}_t) lo è anche rispetto alla sua filtrazione naturale \mathcal{F}_t^X . La verifica è un esercizio (basta usare le proprietà delle attese condizionali).

Dire che un processo è Markov rispetto alla sua filtrazione naturale equivale ad assegnare le sue distribuzioni finito dimensionali. Vedi [12, Thm. III.1.4] per la dimostrazione (facile) del seguente risultato.

t:cpmn

Proposizione 3.10. *Un processo X è Markov rispetto a \mathcal{F}_t^X con probabilità di transizione p e stato iniziale $\nu \in \mathcal{P}(S)$ sse per ogni $0 = t_0 < t_1 < \dots < t_n$ e $f_i \in \mathcal{B}(S)$ positive*

$$\mathbb{E}\left(\prod_{i=0}^n f(X_{t_i})\right) = \int \nu(dx_0) f_0(x_0) \int p_{t_1-t_0}(x_0, dx_1) f_1(x_1) \cdots \int p_{t_n-t_{n-1}}(x_{n-1}, dx_n) f_n(x_n)$$

In altri termini, con notazioni poverine,

$$\mathbb{P}(X_{t_0} \in dx_0, X_{t_1} \in dx_1, \dots, X_{t_n} \in dx_n) = \nu(dx_0) p_{t_1-t_0}(x_0, dx_1) \cdots p_{t_n-t_{n-1}}(x_{n-1}, dx_n) \quad (3.18)$$

dfdpm

Mediante il teorema di estensione di Kolmogorov menzionato in sezione 1.2, a partire dalla probabilità di transizione p e la legge iniziale ν possiamo costruire la realizzazione canonica \mathbb{P}_ν del processo X sullo spazio $\Omega = S^{\mathbb{R}^+}$. Il vantaggio di tale realizzazione è che possiamo introdurre l'operatore di traslazione. Dato $t \geq 0$ definiamo $\vartheta_t : \Omega \rightarrow \Omega$ come $(\vartheta_t \omega)(s) := \omega(t+s)$. Palesemente $X_s(\vartheta_t \omega) = X_{t+s}(\omega)$. Utilizzando la proposizione precedente è facile verificare (vedi [12, Prop. III.1.7] in caso di dubbio) che possiamo formulare, equivalentemente, la proprietà di Markov dicendo che per ogni variabile aleatoria $Z \in \mathcal{F}_\infty^X := \bigvee_{t \geq 0} \mathcal{F}_t^X$ limitata (o positiva) si ha

$$\mathbb{E}_\nu(Z \circ \vartheta_t | \mathcal{F}_t^X) = \mathbb{E}_{X_t}(Z) \quad (3.19)$$

pmpe

che è la forma più conveniente della proprietà di Markov. Nella sezione precedente l'abbiamo quasi sempre usato Markov in questo modo.

Soluzione di equazioni stocastiche

Siano b, σ funzioni Lipschitz su \mathbb{R}^n a valori campi vettoriali, rispettivamente matrici, e $X_t^{(x)}$, $t \geq 0$ (in cui evidenziamo la dipendenza dal dato iniziale nella notazione) la soluzione di

$$X_t^{(x)} = x + \int_0^t ds b(X_s^{(x)}) + \int_0^t \sigma(X_s^{(x)}) dw_s \quad (3.20)$$

ei

Notiamo, che rispetto alla teoria svolta in sezione 2.3, abbiamo considerato il caso in cui non vi è dipendenza esplicita dal tempo nei coefficienti, questo per avere un processo di Markov omogeneo nel tempo, ovvero la probabilità di transire da x al tempo s a B al tempo t (con $t > s$) dipende solo $t - s$.

Vediamo che $X_t^{(x)}$, soluzione di (3.20), è un processo di Markov con probabilità di transizione

$$p_t(x, B) := \mathbb{P}\left(X_t^{(x)} \in B\right) \quad (3.21)$$

tpsde

Cominciamo con il verificare la proprietà di Markov. Evidentemente

$$\begin{aligned} X_{t+s}^{(x)} &= x + \int_0^{t+s} du b(X_u^{(x)}) + \int_0^{t+s} \sigma(X_u^{(x)}) dw_u \\ &= X_s^{(x)} + \int_s^{t+s} du b(X_u^{(x)}) + \int_s^{t+s} \sigma(X_u^{(x)}) dw_u \end{aligned}$$

Per $y \in \mathbb{R}^n$ introduciamo il processo $Y_t^{(y)}$ come l'unica soluzione di

$$Y_t^{(y)} = y + \int_0^t du b(Y_u^{(y)}) + \int_0^t \sigma(Y_u^{(y)}) d\tilde{w}_u$$

dove $\tilde{w}_u := w_{u+s} - w_s$, $u \geq 0$ è un moto browniano rispetto alla filtrazione $\tilde{\mathcal{F}}_u := \mathcal{F}_{s+u}$.

In virtù dell'unicità della soluzione di (3.20), confrontando le due precedenti equazioni troviamo

$$X_{t+s}^{(x)} = Y_t^{(X_s^{(x)})}$$

e quindi

$$\mathbb{E}\left(f(X_{t+s}^{(x)}) \middle| \mathcal{F}_s\right) = \mathbb{E}\left(f(Y_t^{(X_s^{(x)})}) \middle| \tilde{\mathcal{F}}_0\right) = P_t f(X_s^{(x)})$$

in cui abbiamo usato che il moto browniano \tilde{w} è indipendente da $\tilde{\mathcal{F}}_0$.

Rimangono da verificare le altre proprietà richieste in definizione 3.9. Vediamo che per ogni $B \in \mathcal{B}(\mathbb{R}^n)$ e $t \geq 0$ l'applicazione $x \mapsto p_t(x, B)$ è misurabile. Poiché possiamo approssimare boreliani con aperti è sufficiente verificare che per ogni $f : \mathbb{R}^n \rightarrow \mathbb{R}$ limitata e continua e $t \geq 0$, l'applicazione $x \mapsto P_t f(x)$ è continua. Scrivendo l'equazione (3.20) con y al posto di x , utilizzando la lipshizianità di b, σ , la disuguaglianza di Doob ed il lemma di Gronwall troviamo

$$\lim_{\delta \downarrow 0} \sup_{|x-y| < \delta} \mathbb{E}\left(\sup_{t \in [0, T]} |X_t^{(x)} - X_t^{(y)}|^2\right) = 0$$

Grazie a tale stima è facile verificare la continuità enunciata in precedenza.

Il fatto che per ogni $t \geq 0$ e $x \in \mathbb{R}^n$ l'applicazione $\mathcal{B}(\mathbb{R}^n) \ni B \mapsto p_t(x, B)$ sia una probabilità è ovvio, rimane quindi da verificare l'equazione di Chapman–Kolmogorov. In virtù della proprietà di Markov, abbiamo

$$\begin{aligned} p_{t+s}(x, B) &= \mathbb{P}(X_{t+s}^{(x)} \in B) = \mathbb{E}\mathbb{E}(\mathbb{1}_B(X_{t+s}^{(x)}) | \mathcal{F}_s) = \mathbb{E}\mathbb{E}\left(\mathbb{1}_B(Y_t^{(X_s^{(x)})}) \middle| \tilde{\mathcal{F}}_0\right) \\ &= \mathbb{E}\left(p_t(X_s^{(x)}, B)\right) = \int p_s(x, dy) p_t(y, B) \end{aligned}$$

che conclude la dimostrazione della markovianità di $X_t^{(x)}$.

Nelle applicazioni discusse in sezione 3.1 abbiamo spesso usato la proprietà di Markov (3.17) non ma tempi deterministici, ma per tempi stocastici (di arresto). Nel caso di catene di Markov (tempo discreto) ciò è sicuramente lecito, nel caso di tempo continuo è invece una richiesta più forte (proprietà di Markov forte). Vedremo come tale proprietà sia implicata da una proprietà analitica sul semigrupp P_t .

s:sf

3.3. Semigrupp di Feller

Nel contesto di spazio degli stati S polacco, introduciamo lo spazio delle funzioni continue che si annullano all'infinito come $C_0(S) := \overline{C_K(S)}$ ove $C_K(S)$ denota l'insieme

delle funzioni continue su S a supporto compatto e la chiusura è nella norma uniforme, $|f| := \sup_{x \in S} |f(x)|$. Ovviamente nel caso in cui S sia compatto abbiamo $C_0(S) = C(S)$. Indichiamo con $\|\cdot\|$ la norma di un operatore lineare su $C_0(S)$; ovvero $\|A\| := \sup_{f: |f|=1} |Af|$.

t:sf

Definizione 3.11. *Un semigruppato di Feller su $C_0(S)$ è una famiglia ad un parametro $\{P_t\}_{t \geq 0}$ di operatori lineari su $C_0(S)$ tali che*

- (i) *per ogni $t, s \geq 0$ si ha $P_{t+s} = P_t P_s$ (P_t è un semigruppato);*
- (ii) *$\|P_t\| \leq 1$ per ogni $t \geq 0$ e $P_0 = \mathbb{1}$ (P_t è una contrazione);*
- (iii) *se $f \geq 0$ allora $P_t f \geq 0$ per ogni $t \geq 0$ (P_t preserva la positività);*
- (iv) *per ogni $f \in C_0(S)$ si ha $\lim_{t \downarrow 0} |P_t f - f| = 0$ (P_t è fortemente continuo).*

In questo contesto non siamo in grado di distinguere (ancora!) tra Markov e submarkov; a livello del semigruppato la conservazione della probabilità è espressa da $P_t 1 = 1$ (ove 1 è la funzione costante uguale ad uno), ma in generale $1 \notin C_0(S)$. In ogni caso, qualora avessimo una probabilità di transizione submarkoviana possiamo sempre ricondurci al caso markoviano aggiungendo il *cimitero*. Introduciamo cioè un punto in più in S ponendo $S_\Delta := S \cup \{\Delta\}$; poniamo inoltre $\mathcal{B}(S_\Delta) := \sigma\{\mathcal{B}(S), \{\Delta\}\}$ ed estendiamo ogni funzione su S ad una funzione su S_Δ dichiarando $f(\Delta) = 0$.

Se abbiamo una probabilità di transizione p_t submarkoviana su S la possiamo allora estendere ad una \tilde{p}_t markoviana su S_Δ dichiarando $\tilde{p}_t(x, B) = p_t(x, B)$ con $x \in S$ e $B \in \mathcal{B}(S)$, $\tilde{p}_t(x, \{\Delta\}) = 1 - p_t(x, S)$ con $x \in S$ e $\tilde{p}_t(\{\Delta\}, \{\Delta\}) = 1$.

Utilizzando il teorema di rappresentazione di Riesz possiamo associare ad un semigruppato di Feller una probabilità di transizione.

t:dfapt

Lemma 3.12. *Sia P_t un semigruppato di Feller, allora esiste un'unica probabilità di transizione (submarkoviana in generale) tale che*

$$P_t f(x) = \int p_t(x, dy) f(y) \tag{3.22}$$

Pp

Dimostrazione. Per $x \in S$ e $t \geq 0$ fissati, l'applicazione $f \mapsto P_t f(x)$ è un funzionale lineare continuo e positivo su $C_0(S)$. Per il teorema di rappresentazione di Riesz esiste allora una misura positiva $p_t(x, \cdot)$ tale che (3.22) vale. Inoltre, poiché $\|P_t\| \leq 1$, troviamo anche che $p_t(x, \cdot)$ è una subprobabilità ovvero $p_t(x, S) \leq 1$.

Rimangono da verificare le altre richieste in definizione 3.9. L'unica non immediatamente ovvia è la misurabilità di $x \mapsto p_t(x, B)$. Per $f \in C_0(S)$ abbiamo che $x \mapsto P_t f(x)$ è un'applicazione continua e quindi anche misurabile; d'altra parte con $f \in C_0(S)$ si può approssimare $\mathbb{1}_B$ con $B \in \mathcal{B}(S)$ e verificare la misurabilità richiesta. \square

Vediamo ora quando il semigruppato costruito a partire da una probabilità di transizione risulta di Feller. Ovviamente servirà che $P_t C_0(S) \subset C_0(S)$. Il seguente risultato mostra come la continuità forte del semigruppato sia in realtà implicata dalla semplice continuità puntuale.

t:ptsf

Teorema 3.13. *Sia p_t una probabilità di transizione e P_t il semigruppato associato. Il semigruppato P_t è Feller sse $P_t C_0(S) \subset C_0(S)$ e per ogni $f \in C_0(S)$, $x \in S$ si ha*

$$\lim_{t \downarrow 0} P_t f(x) = f(x) \tag{3.23}$$

ccpt

Dimostrazione. L'unica affermazione da dimostrare è la continuità forte a partire da (3.23). Poiché $\|P_t\| \leq 1$ è inoltre sufficiente verificarla per un insieme D denso in $C_0(S)$; per costruire tale insieme abbiamo bisogno di introdurre il risolvete. Fissata $f \in C_0(S)$ abbiamo che per $t \geq 0$ l'applicazione $x \mapsto P_t f(x)$ è continua; d'altra parte, in virtù di (3.23) e la proprietà di semigruppato, si ha la continuità da destra di $t \mapsto P_t f(x)$

$$\lim_{s \downarrow 0} P_{t+s} f(x) = P_t f(x)$$

Concludiamo che $(t, x) \mapsto P_t f(x)$ è misurabile; possiamo quindi definire il risolvete R_λ , $\lambda > 0$ come

$$R_\lambda f(x) := \int_0^\infty dt e^{-\lambda t} P_t f(x) \tag{3.24}$$

risolvente

Poiché $\|P_t\| \leq 1$, per convergenza dominata segue che $R_\lambda f \in C_0(S)$, inoltre

$$|\lambda R_\lambda f| \leq \int_0^\infty dt \lambda e^{-\lambda t} |P_t f| \leq |f|$$

ovvero $\|\lambda R_\lambda\| \leq 1$.

Vediamo che per ogni $x \in S$ si ha $\lim_{\lambda \rightarrow \infty} \lambda R_\lambda f(x) = f(x)$. Per $\delta > 0$ scriviamo

$$\lambda R_\lambda f(x) - f(x) = \int_0^\delta dt \lambda e^{-\lambda t} [P_t f(x) - f(x)] + \int_\delta^\infty dt \lambda e^{-\lambda t} [P_t f(x) - f(x)]$$

e troviamo

$$\lim_{\lambda \rightarrow \infty} |\lambda R_\lambda f(x) - f(x)| \leq \sup_{t \in [0, \delta]} |P_t f(x) - f(x)|$$

e concludiamo grazie a (3.23).

Un semplice calcolo (esercizio) mostra infine che vale l'identità del risolvete, ovvero per ogni $\lambda, \mu > 0$ si ha

$$R_\lambda - R_\mu = (\mu - \lambda) R_\lambda R_\mu = (\lambda - \mu) R_\mu R_\lambda$$

Introduciamo ora $D := R_\lambda C_0(S)$, per quanto mostrato prima $D \subset C_0(S)$ e, per l'identità del risolvete, D non dipende da λ . Per concludere la dimostrazione facciamo vedere che D è denso e che P_t è fortemente continuo su D . Grazie al teorema di Riesz, per verificare che D è denso basta mostrare che ogni misura limitata nulla su D è identicamente nulla. Sia m una misura limitata nulla su D , grazie a quanto dimostrato in precedenza per $f \in C_0(S)$ abbiamo $R_\lambda f \in D$ e ricaviamo

$$\int dm f = \lim_{\lambda \rightarrow \infty} \int dm \lambda R_\lambda f = 0$$

ovvero m è identicamente nulla.

Per mostrare la continuità forte di P_t su D cominciamo con l'osservare come

$$P_t R_\lambda f = \int_0^\infty ds e^{-\lambda s} P_t P_s f(x) = e^{\lambda t} \int_t^\infty ds e^{-\lambda s} P_s f$$

Per $g \in D$ esiste $f \in C_0(S)$ tale che $g = R_\lambda f$; abbiamo quindi

$$\begin{aligned} |P_t g - g| &= |P_t R_\lambda f - R_\lambda f| = \left| e^{\lambda t} \int_t^\infty ds e^{-\lambda s} P_s f - R_\lambda f \right| \\ &\leq (e^{\lambda t} - 1) |R_\lambda f| + e^{\lambda t} \int_0^t ds e^{-\lambda s} |P_s f| \leq (e^{\lambda t} - 1) |R_\lambda f| + t e^{\lambda t} |f| \end{aligned}$$

s:smp

e concludiamo prendendo il limite $t \downarrow 0$. □

3.4. Processi di Feller e proprietà di Markov forte

Diciamo che un processo di Markov è di Feller se il semigruppato associato è di Feller. Più precisamente dato uno spazio filtrato $(\Omega, \mathcal{F}, \mathcal{F}_t)$, una collezione di probabilità \mathbb{P}_ν , $\nu \in \mathcal{P}(S)$ ed un processo $X_t : \Omega \rightarrow S$, $t \geq 0$ diciamo che è un *processo di Feller* se per ogni $\nu \in \mathcal{P}(S)$ e $f \in C_0(S)$ si ha che

$$\mathbb{E}_\nu(f(X_{t+s}) | \mathcal{F}_s) = P_t f(X_s) \quad \mathbb{P}_\nu\text{-q.s.}$$

per un semigruppato di Feller P_t .

Mentre per processi di Markov non abbiamo nessuna regolarità delle traiettorie $t \mapsto X_t$, i processi di Feller hanno traiettorie continue da destra. Più precisamente vale il seguente risultato.

t:rtpf

Teorema 3.14. *Sia X_t un processo di Feller. Allora esiste un processo \tilde{X}_t (una modifica di X_t) tale che: per ogni $\nu \in \mathcal{P}(S)$ si ha*

$$\mathbb{P}_\nu(X_t = \tilde{X}_t) = 1 \quad \forall t \geq 0$$

con traiettorie continue da destra con probabilità uno, ovvero

$$\mathbb{P}_\nu\left(\lim_{s \downarrow 0} \tilde{X}_{t+s} = \tilde{X}_t \quad \forall t \geq 0\right) = 1$$

Idea della dimostrazione. Rimandando a [12, Thm. III.2.7] per la dimostrazione completa, accenniamo l'idea base. Dato X vogliamo costruire \tilde{X} con traiettorie continue da destra. Il programma è questo: ci prendiamo un \mathbb{Q} denso in \mathbb{R}_+ e vorremmo porre

$$\tilde{X}_t := \lim_{s \downarrow t, s \in \mathbb{Q}} X_s \tag{3.25}$$

versione

se riusciamo a verificare che il limite sopra esiste con probabilità uno è poi abbastanza ragionevole credere che \tilde{X} abbia le proprietà volute. Per dimostrare l'esistenza del limite in (3.25) esibiremo delle supermartingale ed usiamo una versione del teorema di convergenza,

vedi teorema 1.17. Invece di guardare X_s possiamo guardare $h(X_s)$ per opportune $h : S \rightarrow \mathbb{R}_+$; basta dimostrare l'esistenza del limite $s \downarrow t$ per "abbastanza" (sufficienti a separare i punti) h . Scegliamo h della forma $h = R_\lambda f$ ove $f \geq 0$ e R_λ è il risolvente di P_t . Utilizzando il fatto che P_t è Feller si può verificare che queste h bastano.

Utilizzando la proprietà di Markov, vediamo ora che, dati $\lambda > 0$ e $f \geq 0$, il processo $Y_t := e^{-\lambda t} R_\lambda f(X_t)$ è una supermartingala, i.e. per $0 \leq s < t$ abbiamo $\mathbb{E}_\nu(Y_t | \mathcal{F}_s^X) \leq Y_s$. Qui \mathcal{F}^X è la filtrazione generata da X_t . Ricordando che il risolvente è stato definito in (3.24), grazie alla proprietà di Markov otteniamo

$$\begin{aligned} \mathbb{E}_\nu(Y_t | \mathcal{F}_s^X) &= e^{-\lambda t} \mathbb{E}_{X_s}(R_\lambda f(X_{t-s})) = e^{-\lambda t} P_{t-s} \int_0^\infty du e^{-\lambda u} P_u f(X_s) \\ &= e^{-\lambda s} \int_{t-s}^\infty du e^{-\lambda u} P_u f(X_s) \leq e^{-\lambda s} R_\lambda f(X_s) = Y_s \end{aligned}$$

È anche possibile trovare una condizione sul semigruppato P_t sufficiente a garantire la continuità delle traiettorie e non la sola continuità da destra. Vale infatti il seguente risultato, vedi [12, Ex. III.2.27]. Ricordiamo che $B_r(x) := \{y \in S : d(y, x) < r\}$ indica la palla di raggio r centrata in x .

t:ctpf

Proposizione 3.15. *Sia P_t un semigruppato di Feller con probabilità di transizione p_t che soddisfa la seguente condizione. Per ogni $\varepsilon > 0$ e $K \subset S$ compatto risulta*

$$\lim_{t \downarrow 0} \sup_{x \in K} \frac{1}{t} p_t(x, B_\varepsilon(x)^c) = 0$$

allora esiste un processo di Feller con probabilità di transizione p_t e traiettorie continue con probabilità uno.

Esercizio. Sia p_t il nucleo del calore su \mathbb{R}^n , ovvero

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

Verificare che è il corrispondente semigruppato di Feller su $C_0(\mathbb{R}^n)$ e che le traiettorie del processo associato (il moto browniano n -dimensionale) sono continue con probabilità uno.

Osserviamo come nell'esercizio precedente (o per soluzioni di equazioni stocastiche) non abbiamo in realtà bisogno di evocare proposizione 3.15 (non dimostrata) per la regolarità delle traiettorie in quanto l'abbiamo già dimostrata per altra via.

Il prossimo passo consiste nel dimostrare la proprietà di Markov forte per processi di Feller. Questa si può enunciare in diversi modi, la più conveniente è utilizzando la versione canonica del processo ed introducendo l'operazione di traslazione (nel tempo) direttamente sullo spazio di probabilità Ω . Prima di enunciare la proprietà di Markov forte dobbiamo decidere la filtrazione di riferimento. In molti dei risultati discussi è

cruciale avere una filtrazione \mathcal{F}_t che soddisfi la condizione usuale, sia cioè completa e continua da destra; non è affatto ovvio come soddisfare queste condizioni conservando la proprietà di Markov, tanto più che vorremmo \mathcal{F}_t completa non solo rispetto ad una probabilità \mathbb{P} fissata, ma rispetto a tutta la famiglia Markoviana \mathbb{P}_ν , $\nu \in \mathcal{P}(S)$.

Vediamo come organizzare una filtrazione che soddisfi la condizione usuale. Abbiamo un semigruppato di Feller P_t , vogliamo costruire la versione *canonica* del corrispondente processo sullo spazio (di Skorohod) $\Omega := D(\mathbb{R}_+, S)$ delle funzioni su \mathbb{R}_+ a valori S continue da destra. Come al solito chiamiamo X_t la coordinata canonica su tale spazio, cioè $X_t(\omega) := \omega(t)$, ed introduciamo la filtrazione canonica $\mathcal{F}_t^0 := \sigma\{X_s, s \in [0, t]\}$ insieme a $\mathcal{F}_\infty^0 := \bigvee_{t \geq 0} \mathcal{F}_t^0$. Fissata $\nu \in \mathcal{P}(S)$ chiamiamo \mathbb{P}_ν la probabilità su Ω la cui distribuzione sui cilindri è data da (3.18). Introduciamo ora \mathcal{F}_∞^ν come il completamento di \mathcal{F}_∞^0 rispetto a \mathbb{P}_ν

$$\mathcal{F}_\infty^\nu := \mathcal{F}_\infty^0 \bigvee \{N \subset \Omega : \exists M \in \mathcal{F}_\infty^0, M \supset N, \mathbb{P}_\nu(M) = 0\}$$

Definiamo inoltre la filtrazione \mathcal{F}_t^ν aggiungendo a \mathcal{F}_t^0 gli insiemi di \mathbb{P}_ν probabilità zero in \mathcal{F}_∞^ν , ovvero

$$\mathcal{F}_t^\nu := \mathcal{F}_t^0 \bigvee \{N \in \mathcal{F}_\infty^\nu : \mathbb{P}_\nu(N) = 0\}$$

Introduciamo infine

$$\mathcal{F}_\infty := \bigcap_{\nu \in \mathcal{P}(S)} \mathcal{F}_\infty^\nu, \quad \mathcal{F}_t := \bigcap_{\nu \in \mathcal{P}(S)} \mathcal{F}_t^\nu$$

e notiamo che \mathcal{F}_t è la filtrazione “massimale” per avere la proprietà di Markov per ogni \mathbb{P}_ν .

Fino a questo punto non abbiamo fatto nulla di straordinario: ci siamo presi la filtrazione canonica \mathcal{F}_t^0 ed abbiamo fatto il completamento massimale, l’aspetto quasi miracoloso è che dopo questa operazione troviamo anche una filtrazione continua da destra; in realtà non è poi così miracoloso se teniamo conto che le traiettorie di un processo di Feller sono continue da destra.

t:rccf

Proposizione 3.16. *Per ogni $\nu \in \mathcal{P}(S)$ la filtrazione \mathcal{F}_t^ν e la filtrazione \mathcal{F}_t sono continue da destra: $\mathcal{F}_{t+} := \bigcap_{s>0} \mathcal{F}_{t+s} = \mathcal{F}_t$ e $\mathcal{F}_{t+}^\nu := \bigcap_{s>0} \mathcal{F}_{t+s}^\nu = \mathcal{F}_t^\nu$.*

Idea della dimostrazione. Rimandando a [12, Prop.III.2.10] per la dimostrazione completa, acceniamo l’idea base. Fissiamo $\nu \in \mathcal{P}(S)$. Poiché sia \mathcal{F}_t^ν sia \mathcal{F}_{t+}^ν sono complete rispetto a \mathbb{P}_ν , non è difficile verificare che se per ogni variabile aleatoria $Z \in \mathcal{F}_\infty^\nu$ limitata abbiamo

$$\mathbb{E}_\nu(Z|\mathcal{F}_t^\nu) = \mathbb{E}_\nu(Z|\mathcal{F}_{t+}^\nu) \quad \mathbb{P}_\nu\text{-q.s.}$$

allora $\mathcal{F}_t^\nu = \mathcal{F}_{t+}^\nu$. Possiamo prendere Z della forma $Z = f_1(X_{t_1}) \cdots f_n(X_{t_n})$; utilizzando la proprietà di Markov, il fatto che semigruppato è di Feller e la continuità da destra delle traiettorie si riesce a verificare la precedente identità. Nella dimostrazione della proprietà di Markov forte (teorema 3.18) vedremo un argomento analogo.

La proprietà di Markov (normale) l’abbiamo enunciata nella forma (3.17), ma come già osservato in sezione 3.2 possiamo introdurre l’operatore di traslazione $\theta_t : \Omega \rightarrow \Omega$ e formularla come in (3.19). Nel contesto di processi di Feller non c’è alcuna difficoltà a passare a completamento della filtrazione, in caso di dubbio vedi [12, Prop. III.2.14].

t:pmcs

Proposizione 3.17. *Sia X_t la realizzazione canonica di un processo di Feller e Z una variabile aleatoria (limitata o positiva) in \mathcal{F}_∞ , allora*

$$\mathbb{E}_\nu(Z \circ \theta_t | \mathcal{F}_t) = \mathbb{E}_{X_t}(Z), \quad \mathbb{P}_\nu\text{-q.s.} \quad (3.26)$$

pmcs

Enunciamo ora la proprietà di Markov forte e vediamo come sia garantita per processi di Feller. Ricordiamo che un *tempo di arresto* è una variabile aleatoria $\tau : \Omega \rightarrow [0, +\infty]$ tale che $\{\tau \leq t\} \in \mathcal{F}_t$ (è la filtrazione completata che soddisfa la condizione usuale!). Definiamo la filtrazione arrestata come

$$\mathcal{F}_\tau := \{B \in \mathcal{F}_\infty : B \cap \{\tau \leq t\} \in \mathcal{F}_t\}$$

il cui significato intuitivo è il seguente: \mathcal{F}_τ è fatta dagli eventi che possiamo decidere guardando la traiettoria X_t per $t \leq \tau$.

Definiamo naturalmente il processo arrestato al tempo τ come

$$X_\tau(\omega) := \begin{cases} X_{\tau(\omega)}(\omega) & \omega : \tau(\omega) < +\infty \\ \Delta & \omega : \tau(\omega) = +\infty \end{cases}$$

in cui ricordiamo che Δ è il cimitero. Anche l'operatore di traslazione θ_τ per un tempo di arresto τ è definito naturalmente come

$$\theta_\tau(\omega) := \begin{cases} \omega(s+t)(\omega) & \omega : \tau(\omega) = t \\ \Delta & \omega : \tau(\omega) = +\infty \end{cases}$$

Esercizio. Verificare che \mathcal{F}_τ è una σ -algebra e che $X_\tau \in \mathcal{F}_\tau$.

La proprietà di Markov forte si enuncia come in (3.26) sostituendo al tempo deterministico t un tempo di arresto τ . Mostriamo ora che vale per processi di Feller.

t:mfpf

Teorema 3.18. *Sia X_t la realizzazione canonica di un processo di Feller, τ un tempo di arresto e Z una variabile aleatoria (limitata o positiva) in \mathcal{F}_∞ , allora*

$$\mathbb{E}_\nu(Z \circ \theta_\tau | \mathcal{F}_\tau) = \mathbb{E}_{X_\tau}(Z), \quad \mathbb{P}_\nu\text{-q.s. sull'evento } \{X_\tau \neq \Delta\} \quad (3.27)$$

mfpf

Dimostrazione. La dimostrazione consiste in tre passi. Nel primo facciamo vedere che (3.27) vale se τ ha valori in un insieme numerabile (in particolare Markov equivale a Markov forte per processi a tempo discreto). Nel secondo facciamo vedere che (3.27) vale se $Z \in \mathcal{F}_\infty^0$: è questo il passo cruciale in cui utilizzeremo la proprietà di Feller del semigruppato associato. Infine facciamo vedere che vale anche per il completamento di \mathcal{F}_∞^0 .

Passo 1. Supponiamo $\tau : \Omega \rightarrow Q$ ove Q è un sottoinsieme numerabile di $[0, +\infty]$.

Dobbiamo far vedere che se $B \in \mathcal{F}_\tau$ allora

$$\mathbb{E}_\nu\left(\mathbb{1}_{\{X_\tau \neq \Delta\}} \mathbb{1}_B Z \circ \theta_\tau\right) = \mathbb{E}_\nu\left(\mathbb{1}_{\{X_\tau \neq \Delta\}} \mathbb{1}_B \mathbb{E}_{X_\tau}(Z)\right)$$

Il primo membro è uguale a

$$\begin{aligned} \sum_{q \in Q} \mathbb{E}_\nu \left(\mathbb{1}_{\{\tau=q\}} \mathbb{1}_{\{X_q \neq \Delta\}} \mathbb{1}_B Z \circ \theta_q \right) &= \sum_{q \in Q} \mathbb{E}_\nu \left(\mathbb{1}_{\{\tau=q\}} \mathbb{1}_{\{X_q \neq \Delta\}} \mathbb{1}_B \mathbb{E}_{X_q}(Z) \right) \\ &= \mathbb{E}_\nu \left(\mathbb{1}_{\{X_\tau \neq \Delta\}} \mathbb{1}_B \mathbb{E}_{X_\tau}(Z) \right) \end{aligned}$$

dove abbiamo usato $\{X_q \neq \Delta\} \cap \{\tau = q\} \cap B \in \mathcal{F}_q$ e la proprietà di Markov tout court (3.26).

Passo 2. Supponiamo $Z \in \mathcal{F}_\infty^0$ e $Z \geq 0$ (il caso Z limitato segue immediatamente).

Sia $\tau_n := 2^{-n}([2^n \tau] + 1)$; assume valori in un insieme numerabile e $\tau_n \downarrow \tau$ per $n \rightarrow \infty$. Poiché $Z \in \mathcal{F}_\infty^0$ è sufficiente considerare il caso $Z = \prod_{i=1}^k f_i(X_{t_i})$ con $0 \leq t_1 < \dots < t_k$, $f_i \in C_0(S)$ con $f_i \geq 0$. Per il passo 1 abbiamo, sull'insieme $X_\tau \neq \Delta$,

$$\mathbb{E}_\nu \left(\prod_{i=1}^k f_i(X_{t_i}) \circ \theta_{\tau_n} \middle| \mathcal{F}_{\tau_n} \right) = \mathbb{E}_{X_{\tau_n}} \left(\prod_{i=1}^k f_i(X_{t_i}) \right) =: g(X_{\tau_n})$$

evidentemente

$$g(x) = \int p_{t_1}(x, dx_1) f_1(x_1) \cdots \int p_{t_k - t_{k-1}}(x_{k-1}, dx_k) f_k(x_k)$$

e, grazie alla proprietà di Feller di P_t , abbiamo $g \in C_0(S)$. Utilizzando la continuità da destra delle traiettorie X_t e della filtrazione \mathcal{F}_t , possiamo ora prendere il limite per $n \rightarrow \infty$ nell'espressione precedente dimostrando la proprietà di Markov forte per $Z \in \mathcal{F}_\infty^0$.

Passo 3. Caso generale $Z \in \mathcal{F}_\infty$ e $Z \geq 0$ (il caso Z limitato segue immediatamente).

Introduciamo

$$\tilde{\mathbb{P}}_\nu(B) := \frac{\mathbb{P}_\nu(B \cap \{X_\tau \neq \Delta\})}{\mathbb{P}_\nu(X_\tau \neq \Delta)}$$

ed osserviamo che le probabilità \mathbb{P}_ν e $\tilde{\mathbb{P}}_\nu$ condizionate a \mathcal{F}_τ coincidono sull'evento $\{X_\tau \neq \Delta\}$ (verificare). Basta pertanto dimostrare (3.27) per $\tilde{\mathbb{P}}_\nu$ omettendo “sull'evento $\{X_\tau \neq \Delta\}$ ”.

Sia $\mu(B) := \tilde{\mathbb{P}}_\nu(X_\tau \in B)$ la legge di X_τ . Per definizione del completamento \mathcal{F}_∞ , possiamo trovare due variabili aleatorie $Z', Z'' \in \mathcal{F}_\infty^0$ positive tali che $Z' \leq Z \leq Z''$ e $\mathbb{P}_\mu(Z'' - Z' > 0) = 0$. Grazie al passo 2 abbiamo allora

$$\tilde{\mathbb{P}}_\nu(Z'' \circ \theta_\tau - Z' \circ \theta_\tau > 0) = \tilde{\mathbb{E}}_\nu \left(\tilde{\mathbb{E}}_{X_\tau} \mathbb{1}_{\{Z'' - Z' > 0\}} \right) = 0$$

e quindi, per la completezza di \mathcal{F}_∞ e l'arbitrarietà di ν , troviamo $Z \circ \theta_\tau \in \mathcal{F}_\infty$. A questo punto

$$\tilde{\mathbb{E}}_\nu(Z' \circ \theta_\tau) \leq \tilde{\mathbb{E}}_\nu(Z \circ \theta_\tau) \leq \tilde{\mathbb{E}}_\nu(Z'' \circ \theta_\tau)$$

e concludiamo utilizzando nuovamente il passo 2. □

Se consideriamo un processo di Feller X_t con dato iniziale $x \in S$, evidentemente sappiamo tutto degli eventi che dipendono solo da X_0 . A priori gli eventi che dipendono solo da X_t per $t \in [0, \varepsilon)$ ed $\varepsilon > 0$ arbitrariamente piccolo potrebbero essere meno banali, ma la continuità da destra della filtrazione \mathcal{F}_t implica che questo non succede. È sbagliato pensare che la continuità da destra della filtrazione sia una condizione puramente tecnica: dice che non guadagniamo alcuna informazione in più guardando il processo ad un tempo successivo purché infinitesimo.

t:01 **Teorema 3.19.** Legge 0–1 di Blumenthal.

Sia $x \in S$ e \mathbb{P}_x la realizzazione canonica di un processo di Feller con dato iniziale x . Se $B \in \mathcal{F}_{0+} = \bigcap_{\varepsilon > 0} \mathcal{F}_\varepsilon$ allora $\mathbb{P}_x(B) = 0$ oppure $\mathbb{P}_x(B) = 1$.

Dimostrazione. Per la continuità da destra della filtrazione \mathcal{F}_t abbiamo $\mathcal{F}_{0+} = \mathcal{F}_0$, il teorema è allora ovvio. Una possibile dimostrazione formale è la seguente

$$\mathbb{P}_x(B) = \mathbb{E}_x(\mathbb{1}_B \mathbb{1}_B) = \mathbb{E}_x(\mathbb{1}_B \mathbb{1}_B \circ \theta_0) = \mathbb{E}_x(\mathbb{1}_B \mathbb{P}_{X_0}(B)) = [\mathbb{P}_x(B)]^2$$

in cui abbiamo usato la proprietà di Markov. La precedente uguaglianza implica che $\mathbb{P}_x(B) = 0$ oppure $\mathbb{P}_x(B) = 1$. □

s:gsf

3.5. Generatori di semigrupp di Feller

Nella costruzione di un processo di Feller siamo partiti dal semigrupp corrispondente P_t . Tale procedura non è affatto pratica, solo in pochi esempi (come il moto browniano) conosciamo esattamente la probabilità di transizione p_t . Vorremmo riuscire a costruire un processo partendo da una descrizione della dinamica per un intervallo di tempo infinitesimo, come accade per i sistemi dinamici in cui costruiamo il flusso di fase associato ad un dato campo vettoriale. Tale desiderio si realizza nel modo seguente. Sia P_t un semigrupp di Feller su $C_0(S)$. Formalmente il *generatore* di P_t è l'operatore lineare L per cui $P_t = e^{tL}$. Se S è un insieme finito L è una matrice e non è difficile verificare che ogni semigrupp continuo si può scrivere in tal modo per un opportuno L . In generale L sarà però un operatore non limitato e non è completamente chiaro cosa intendiamo per e^{tL} , dobbiamo pure dichiarare quale sia il dominio di L .

Dato un semigrupp di Feller P_t dichiariamo $\mathcal{D}(L)$ l'insieme delle funzioni $f \in C_0(S)$ per cui esiste il limite per $t \downarrow 0$ di $t^{-1}[P_t f - f]$. A questo punto definiamo il generatore L con dominio $\mathcal{D}(L)$ come

$$Lf := \lim_{t \downarrow 0} \frac{P_t f - f}{t} \tag{3.28} \quad \text{gsg}$$

Vediamo ora che in che senso il generatore L descriva la dinamica in un tempo infinitesimo. Sia \mathbb{P}_ν un processo di Feller e $f \in \mathcal{D}(L)$, dalla proprietà di Markov ricaviamo immediatamente

$$\mathbb{E}_\nu(f(X_{t+h}) - f(X_t) | \mathcal{F}_t) = P_h f(X_t) - f(X_t) = h Lf(X_t) + o(h)$$

ove abbiamo usato (3.28). Identifichiamo quindi $Lf(X_t)$ come l'incremento aspettato di $f(X_t)$.

Le seguenti due proposizioni identificano alcune proprietà analitiche di semigrupp di contrazione fortemente continui su spazi di Banach (la proprietà di Markov non ha alcun ruolo), per la dimostrazione si veda, ad esempio, [12, Prop. VII.1.2–1.4]. La prima stabilisce in che senso abbiamo $P_t = e^{tL}$: la funzione $f(t) = P_t f$ risolve l'equazione differenziale $\frac{d}{dt}f(t) = Lf(t)$ con dato iniziale $f(0) = f$. La seconda indentifica il risolvente di P_t (che abbiamo introdotto nella dimostrazione di teorema 3.13) come l'inverso di $\lambda - L$.

t:sga1 **Proposizione 3.20.** *Sia $f \in \mathcal{D}(L)$ allora:*

(i) *per ogni $t \in \mathbb{R}_+$, $P_t f \in \mathcal{D}(L)$;*

(ii) *l'applicazione $t \mapsto P_t f$ è fortemente differenziabile e $\frac{d}{dt}P_t f = LP_t f = P_t Lf$;*

(iii) *vale*

$$P_t f = f + \int_0^t ds P_s Lf = f + \int_0^t ds LP_s f$$

t:sga2 **Proposizione 3.21.** *L'insieme $\mathcal{D}(L)$ è denso in $C_0(S)$ e l'operatore $(L, \mathcal{D}(L))$ è chiuso (nella norma del grafico). Inoltre l'operatore $R_\lambda := \int_0^\infty dt e^{-\lambda t} P_t$ è il risolvente di L , ovvero per ogni $\lambda > 0$ l'applicazione $\mathcal{D}(L) \ni f \mapsto \lambda f - Lf \in C_0(S)$ è biunivoca e l'inversa è data da $R_\lambda f$.*

L'enunciato seguente dipende invece dalla struttura probabilistica (P_t preserva le funzioni positive) ed individua la condizione corrispondente sul generatore L come un principio del massimo.

t:sga3 **Proposizione 3.22.** *Il generatore L soddisfa il principio del massimo positivo, ovvero per ogni $f \in \mathcal{D}(L)$, se $\bar{x} \in S$ è tale che $0 \leq f(\bar{x}) = \sup_{x \in S} f(x)$ allora $Lf(\bar{x}) \leq 0$.*

Dimostrazione. Poiché un semigrupp di Feller si rappresenta in termini di una probabilità di transizione p_t abbiamo

$$P_t f(\bar{x}) - f(\bar{x}) = \int p_t(\bar{x}, dy) f(y) - f(\bar{x}) \leq f(\bar{x}) [p_t(\bar{x}, S) - 1] \leq 0$$

ricordando (3.28) l'enunciato segue. □

Analogamente a lemma 3.7 è possibile costruire una famiglia di martingale partendo da un semigrupp di Feller.

t:mdsf **Teorema 3.23.** *Sia X_t un processo di Feller, $\nu \in \mathcal{P}(S)$ e $f \in \mathcal{D}(L)$ allora*

$$M_t^f := f(X_t) - f(X_0) - \int_0^t ds Lf(X_s)$$

è una martingala rispetto a \mathbb{P}_ν .

Dimostrazione. Poiché $f \in \mathcal{D}(L)$ osserviamo che $f, Lf \in C_0(S)$, in particolare sono limitate e pertanto M_t^f è integrabile. Grazie alla proprietà di Markov, per $0 \leq s \leq t$ abbiamo

$$\begin{aligned}\mathbb{E}_\nu(M_t^f - M_s^f | \mathcal{F}_s) &= \mathbb{E}_\nu\left(f(X_t) - f(X_s) - \int_s^t du Lf(X_u) \middle| \mathcal{F}_s\right) \\ &= \mathbb{E}_{X_s}\left(f(X_{t-s}) - f(X_0) - \int_0^{t-s} du Lf(X_u)\right)\end{aligned}$$

D'altra parte, per proposizione 3.20, per ogni $y \in S$

$$0 = P_{t-s}f(y) - f(y) - \int_0^{t-s} du Lf(y) = \mathbb{E}_y\left(f(X_{t-s}) - f(X_0) - \int_0^{t-s} du Lf(X_u)\right)$$

che conclude la dimostrazione. \square

Soluzioni di equazioni stocastiche come processi di Feller.

In sezione 3.2 abbiamo visto come la soluzione di un'equazione stocastica sia un processo di Markov, mostriamo adesso che, previe ipotesi aggiuntive sui coefficienti b e σ , risulta anche un processo di Feller e caratterizziamo il corrispondente generatore.

t:sdefp

Proposizione 3.24. *Siano b e σ funzioni da \mathbb{R}^n a valori campi vettoriali, rispettivamente matrici, globalmente lipshiziane e limitate. Allora $X_t^{(x)}$, soluzione di (3.20), è un processo di Feller.*

Dimostrazione. Abbiamo già mostrato che $X_t^{(x)}$ soddisfa la proprietà di Markov, il semigruppato associato P_t ha nucleo integrale come in (3.21). Dobbiamo ora mostrare che è un semigruppato di Feller su $C_0(\mathbb{R}^n)$. Abbiamo già mostrato che $P_t C_0(\mathbb{R}^n) \subset C(\mathbb{R}^n)$, in virtù di teorema 3.13 basta verificare che $\lim_{|x| \rightarrow \infty} P_t f(x) = 0$ e che (3.23) è soddisfatta.

Poiché b e σ sono funzioni limitate da (3.20) segue (verificare) che esiste $C < \infty$ tale che per ogni $x \in \mathbb{R}^n$ e $t \in \mathbb{R}_+$ abbiamo

$$\mathbb{E} \sup_{s \in [0, t]} |X_s^{(x)} - x|^2 \leq C(t + t^2) \quad (3.29) \quad \boxed{\text{sxx}}$$

Per $r > 0$ scriviamo allora

$$|P_t f(x)| \leq \mathbb{E}\left(|f(X_t^{(x)})| \left[\mathbb{1}_{|X_t^{(x)} - x| \leq r} + \mathbb{1}_{|X_t^{(x)} - x| > r} \right]\right) \leq \sup_{y: |y-x| \leq r} |f(y)| + |f| \mathbb{P}(|X_t^{(x)} - x| > r)$$

Grazie alla stima (3.29) e la disuguaglianza di Chebyshev, possiamo maggiorare, uniformemente in x , il secondo termine con $C(t + t^2)r^{-2}$. D'altra parte, poiché $f \in C_0(\mathbb{R}^n)$, per ogni $r > 0$ il primo termine si annulla per $|x| \rightarrow \infty$. Ciò conclude la dimostrazione che $P_t C_0(S) \subset C_0(S)$.

Per verificare (3.23) è sufficiente osservare che da (3.29) segue $X_t^{(x)} \rightarrow x$ in probabilità per $t \downarrow 0$. Per convergenza dominata (f è limitata) troviamo allora

$$\lim_{t \downarrow 0} \mathbb{E}\left(f(X_t^{(x)}) - f(x)\right) = 0$$

che, per (3.21), equivale a (3.23). □

Come abbiamo osservato in sezione 2.4, la formula di Ito ci dice se introduciamo \tilde{L} come l'operatore differenziale

$$\tilde{L}u(x) := \frac{1}{2} \sum_{1 \leq i, j \leq n} a(x)_{i,j} \partial_i \partial_j u(x) + \sum_{i=1}^n b_i(x) \partial_i u(x)$$

ove $a_{i,j}(x) = \sum_{k=1}^m \sigma_{i,k}(x) \sigma_{j,k}(x)$, allora la formula di Ito implica che per ogni $f \in C_K^2(\mathbb{R}^n)$ il processo

$$\tilde{M}_t^f := f(X_t^{(x)}) - f(X_0^{(x)}) - \int_0^t ds \tilde{L}f(X_s^{(x)})$$

è una martingala.

Se confrontiamo tale risultato con teorema 3.23 possiamo trarre le seguenti conclusioni. Da una parte la formula di Ito fornisce una realizzazione della martingala \tilde{M}_t^f in termini di integrali stocastici. Dall'altra il generatore L di $X_t^{(x)}$ (ogni processo di Feller ha il suo generatore) deve essere un'estensione di \tilde{L} . Infatti \tilde{L} definito su $C_K^2(\mathbb{R}^n)$ non è un operatore chiuso (nella norma del grafico), mentre, per proposizione 3.21, L lo è sempre.

Note bibliografiche.

Il teorema 3.1 è tratto da [6], a cui si fa riferimento per una brillante esposizione delle catene di Markov. La dimostrazione, via accoppiamento, del teorema ergodico per catene di Markov su spazio degli stati finito è tratta da [10, Thm. II.1.2] a cui si fa riferimento (capitolo II) per ulteriori esempi di accoppiamento e per la dimostrazione che una funzione limitata e armonica rispetto alla passeggiata aleatoria in \mathbb{Z}^d è costante. Vedi ad esempio [14, Ch. VIII] per il teorema ergodico per catene di Markov su spazio degli stati numerabile e per la corrispondente decomposizione ergodica. Vedi [16, §13.1] per il problema di Dirichelet con spazio degli stati numerabile. Per i cenni della teoria generale dei processi di Markov e di Feller si è seguito [12, Chap. III, VII] a cui si fa riferimento per ulteriori sviluppi e per le dimostrazioni omesse. A tale proposito è anche utile l'esposizione in [16] e in [9, § 2.5–2.7].

s:pde

s:arm 4. Processi di diffusione e PDE

4.1. Problema di Dirichlet e moto browniano

In questa sezione fissiamo la realizzazione canonica del moto browniano n -dimensionale, $n \geq 2$. Ovvero su $\Omega = C(\mathbb{R}_+; \mathbb{R}^n)$ con la filtrazione canonica $\mathcal{F}_t^0 := \sigma\{\omega(s), s \in [0, t]\}$ introduciamo la famiglia di misure \mathbb{P}_x , $x \in \mathbb{R}^n$ come $\mathbb{P}_x(B) := P(x + \omega(\cdot) \in B)$ ove P è la misura di Wiener n -dimensionale; evidentemente $\mathbb{P}_x(\omega(0) = x) = 1$. Come in sezione 3.4 chiamiamo \mathcal{F}_t il completamento “massimale” di \mathcal{F}_t^0 . In virtù di quanto fatto prima, possiamo dire che vale la proprietà di Markov forte come enunciata in teorema 3.18. Per ricordarci che si tratta di un browniano chiamiamo però $w_t(\omega) := \omega(t)$ la coordinata canonica in Ω . Sotto \mathbb{P}_x il processo w ha allora la legge di un browniano con dato iniziale x . Se A è un aperto di \mathbb{R}^n diciamo infine $\tau_A := \inf\{t \geq 0 : w_t \in A^c\}$ il tempo di uscita da A che risulta un tempo di arresto. Come al solito chiamiamo \bar{A} la chiusura di A e ∂A la sua frontiera. Poiché sappiamo che il browniano esce da ogni palla con probabilità uno, basta guardare un componente ed usare e.g. proposizione 1.7, se A è limitato allora $\tau_A < \infty$ con probabilità uno.

Sia D un aperto di \mathbb{R}^n ricordiamo che una funzione $u : D \rightarrow \mathbb{R}$ si dice *armonica* in D se $u \in C^2(D)$ e vale $\Delta u(x) = 0$ per ogni $x \in D$, ovviamente $\Delta = \sum_{i=1}^n \partial_i^2$ è il Laplaciano. Un'altra possibile definizione di funzioni armoniche (senza ipotesi di derivabilità) è ottenuta richiedendo che u soddisfi la proprietà del valor medio.

Sia $B_r(x) := \{y \in \mathbb{R}^n : |x - y| < r\}$ la palla di \mathbb{R}^n (qui è importante che $|\cdot|$ sia la norma euclidea). Introduciamo la misura μ_x^r su $\partial B_r(x)$ come

$$\mu_x^r(A) := \mathbb{P}_x(w_{\tau_{B_r(x)}} \in A)$$

Utilizzando l'invarianza per rotazioni del moto browniano n -dimensionale è facile verificare che μ_x^r è la misura di superficie normalizzata su $\partial B_r(x)$.

t:pvm

Definizione 4.1. Diciamo che una funzione Boreliana $u : D \rightarrow \mathbb{R}$ soddisfa la proprietà del valor medio in D sse per ogni $x \in D$ ed ogni $r > 0$ tale che $\bar{B}_r(x) \subset D$ si ha

$$u(x) = \int_{\partial B_r(x)} \mu_x^r(dy) u(y) \tag{4.1} \quad \text{pvm}$$

La seguente proposizione è un risultato classico di analisi (via teorema della divergenza), ma la dimostrazione proposta (che usa invece il moto browniano) è di un certo interesse.

t:arvm

Proposizione 4.2. Sia u armonica in D allora u soddisfa la proprietà del valor medio in D .

Notiamo che il principio del massimo (le funzioni armoniche assumo massimo sulla frontiera) è un corollario immediato.

Dimostrazione. Consideriamo il processo arrestato $w_{t \wedge \tau_{B_r(x)}}$ e scriviamo la formula di Ito (ci è concesso poiché $u \in C^2(D)$)

$$u(w_{t \wedge \tau_{B_r(x)}}) = u(w_0) + \frac{1}{2} \int_0^{t \wedge \tau_{B_r(x)}} ds \Delta u(w_s) + \int_0^{t \wedge \tau_{B_r(x)}} \nabla u(w_s) \cdot dw_s$$

Ora $\nabla u(x)$ è limitata in $B_r(x)$ e quindi l'integrale stocastico è una martingala. Prendendo il valore di attesa rispetto a \mathbb{P}_x e usando che $\Delta u = 0$ in D troviamo quindi

$$u(x) = \mathbb{E}_x(u(w_0)) = \mathbb{E}_x(u(w_{t \wedge \tau_{B_r(x)}}))$$

Passando ora al limite per $t \uparrow \infty$ ed usando che $\tau_{B_r(x)} < \infty$ \mathbb{P}_x -q.s. troviamo

$$u(x) = \mathbb{E}_x(u(w_{\tau_{B_r(x)}})) = \int_{\partial B_r(x)} \mu_x^r(dy) u(y)$$

cioè la proprietà del valor medio. □

Con un po' di pazienza (convoluzioni) si mostra che se u soddisfa la proprietà del valor medio allora $u \in C^\infty(D)$ ed è ivi armonica, vedi e.g. [9, Prop. 4.2.5].

Come ben noto in elettrostatica, per *problema di Dirichlet* si intende il seguente problema. Dati un dominio D ed una funzione continua $f : \partial D \rightarrow \mathbb{R}$ trovare una funzione $u : \bar{D} \rightarrow \mathbb{R}$ in $C(\bar{D}) \cap C^2(D)$ tale che

$$\begin{cases} \Delta u(x) = 0 & x \in D \\ u(x) = f(x) & x \in \partial D \end{cases} \quad (4.2) \quad \boxed{\text{dirp}}$$

Utilizzando il moto browniano abbiamo un candidato naturale per una soluzione.

t: canar **Proposizione 4.3.** *Sia f tale che per ogni $x \in D$*

$$\mathbb{E}_x(|f(w_{\tau_D})|) < \infty$$

allora la funzione

$$u(x) := \mathbb{E}_x(f(w_{\tau_D})) \quad (4.3) \quad \boxed{\text{carm}}$$

è armonica in D .

N.B. Non stiamo dicendo che la u definita in (4.3) è continua su \bar{D} (non abbiamo nemmeno assunto f continua), solo che soddisfa l'equazione differenziale in (4.2).

Dimostrazione. Per quanto detto sopra, basta verificare che la u definita in (4.3) soddisfa la proprietà del valor medio. Siano $x \in D$ e $r > 0$ tale che $\bar{B}_r(x) \subset D$; evidentemente $\tau_D \geq \tau_{\bar{B}_r(x)}$. Per la proprietà di Markov forte (in versione sofisticatissima, con due tempi di arresto! Convicersi che è legale) troviamo

$$\begin{aligned} u(x) &= \mathbb{E}_x(f(w_{\tau_D})) = \mathbb{E}_x\left(\mathbb{E}_x(f(w_{\tau_D})) \Big| \mathcal{F}_{\tau_{\bar{B}_r(x)}}\right) \\ &= \mathbb{E}_x\left(\mathbb{E}_{w_{\tau_{\bar{B}_r(x)}}}(f(w_{\tau_D}))\right) \\ &= \mathbb{E}_x(u(w_{\tau_{\bar{B}_r(x)}})) = \int_{\partial B_r(x)} \mu_x^r(dy) u(y) \end{aligned}$$

che conclude la dimostrazione. \square

D'altra parte vediamo come ogni soluzione limitata del problema di Dirichlet (con qualche ipotesi aggiuntiva su D e f) possa scriversi nella forma (4.3).

t:arcs

Proposizione 4.4. *Siano f limitata e D tale che per ogni $x \in D$ valga $\mathbb{P}_x(\tau_D < \infty) = 1$ (in particolare basta D limitato). Se u è una soluzione limitata del problema di Dirichlet allora u si rappresenta come in (4.3).*

Dimostrazione. Si tratta semplicemente di scrivere la formula di Ito e prendere il valore di attesa rispetto a \mathbb{P}_x . Siccome dobbiamo garantirci che il pezzo di integrale stocastico ha media nulla, servono un paio di troncature. Introduciamo $D_n := \{x \in D : d(x, \partial D) > 1/n\}$ e $B_n := B_n(0)$, consideriamo il processo arrestato $w_{t \wedge \tau_{D_n} \wedge \tau_{B_n}}$ e scriviamo Ito per la funzione u . Troviamo

$$u(w_{t \wedge \tau_{D_n} \wedge \tau_{B_n}}) = u(w_0) + \int_0^{t \wedge \tau_{D_n} \wedge \tau_{B_n}} \nabla f(w_s) \cdot dw_s$$

prendendo ora il valore di attesa rispetto a \mathbb{P}_x , notando che $u \in C^2(\bar{D}_n \cap \bar{B}_n)$ cosicché ∇u è limitato, troviamo allora

$$u(x) = \mathbb{E}_x(u(w_{t \wedge \tau_{D_n} \wedge \tau_{B_n}}))$$

Passiamo infine al limite per $n \rightarrow \infty$ e poi $t \rightarrow \infty$. Usando la continuità e limitatezza di u e l'ipotesi $\mathbb{P}_x(\tau_D < \infty) = 1$, per convergenza dominata troviamo (4.3). \square

Le proposizioni 4.3 e 4.4 risolvono quasi complemente il problema di Dirichlet, o meglio fanno capire che il cuore di tutta la faccenda è la continuità di u fin sul bordo. Nella precedente formulazione dipende dalla funzione al bordo f che ci è stata data; vogliamo ora riguardarlo come un problema di regolarità del solo D . Dato un dominio D vogliamo cioè capire per quali punti $a \in \partial D$ capita che

$$\lim_{x \rightarrow a, x \in D} \mathbb{E}_x(f(w_{\tau_D})) = f(a) \tag{4.4}$$

rpb1

per ogni $f : \partial D \rightarrow \mathbb{R}$ limitata e continua. Se tutti i punti di ∂D sono siffatti vuol dire che (4.3) è l'unica soluzione limitata del problema di Dirichlet (supponendo che $\mathbb{P}_x(\tau_D < \infty) = 1$). Se vi sono degli $a \in \partial D$ per cui (4.4) fallisce vuol dire che abbiamo formulato male il problema: non dobbiamo dare il dato al bordo in tali punti, ma lasciar decidere a u quello che deve valere. Si può considerare il seguente esempio facile supponiamo D sia la palla "bucata" di \mathbb{R}^2 , i.e. $D = \{x \in \mathbb{R}^2 : x \neq 0, |x| < 1\}$. Poiché vi è un'unica soluzione del problema di Dirichlet sulla palla, visto il dato al bordo su $\partial B_1(0)$, la soluzione u decide di suo quanto deve valere nell'origine, è assai buffo pretendere di assegnare pure il valore di u nell'origine. In $n \geq 3$ ci sono esempio dello stesso fenomeno non così banali, ovvero per un D connesso.

Discutiamo qui l'interpretazione dei punti regolari (ovvero vale (4.4)) in termini del browniano. Introduciamo il tempo di arresto $\sigma_D := \inf\{t > 0 : w_t \in D^c\}$. A priori è

diverso da τ_D : se $w_0 \in \partial D$ abbiamo $\tau_D = 0$, ma potrebbe succedere che w_t entra subito in D cosicché $\sigma_D > 0$. È facile esibire degli $\omega \in \Omega$ per cui questo succede il problema sarà capire se questo può capitare con probabilità positiva.

Dato il dominio D , diciamo che $a \in \partial D$ è un *punto regolare* sse $\mathbb{P}_a(\sigma_a = 0) = 1$, cioè il browniano entra subito in $D^c \setminus \partial D$. La rilevanza di tale definizione nel contesto del problema di Dirichlet è chiarita dal seguente risultato, si veda [9, Thm. 4.2.12] per la dimostrazione.

t:prpd **Teorema 4.5.** *Le seguenti affermazioni sono equivalenti:*

- l:r1** 1. (4.4) vale per ogni $f : \partial D \rightarrow \mathbb{R}$ misurabile, limitata e continua in a ;
- l:r2** 2. $a \in \partial D$ è un punto regolare;
- l:r3** 3. per ogni $\varepsilon > 0$ si ha $\lim_{x \rightarrow a, x \in D} \mathbb{P}_x(\tau_D > \varepsilon) = 0$.

Usando la precedente caratterizzazione ed il fatto che il moto browniano entra immediatamente in ogni cono si ricava il criterio di “cono esterno”. Dato $a \in \partial D$, se esiste un cono con vertice in a contenuto in D^c allora a è un punto regolare.

s:fk

4.2. Formula di Fenyman-Kac

Dato $V : \mathbb{R}^n \rightarrow \mathbb{R}_+$, consideriamo il problema, con $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^n$

$$\begin{cases} \partial_t u(t, x) = \frac{1}{2} \Delta u(t, x) - V(x)u(t, x) \\ u(0, x) = f(x) \end{cases} \quad (4.5) \quad \text{cp}$$

che modella la diffusione del calore (u rappresenta la temperatura) con pozzi descritti da V .

Con opportune ipotesi, la soluzione ammette la rappresentazione

$$u(t, x) = \mathbb{E}_x \left(e^{-\int_0^t ds V(w_s)} f(w_t) \right) \quad (4.6) \quad \text{fk}$$

osservando che proprietà di monotonia ed principio del massimo, si ricavano direttamente.

t:fk

Proposizione 4.6. *Siano $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $V : \mathbb{R}^n \rightarrow [0, \infty)$ continue. Sia inoltre $u \in C^{1,2}((0, +\infty) \times \mathbb{R}^n) \cap C([0, +\infty) \times \mathbb{R}^n)$ una soluzione di (4.5) tale che per ogni $T > 0$ esistono $K > 0$ $\alpha < 1/(2nT)$ per cui*

$$\sup_{s \in [0, t]} |u(s, x)| \leq K e^{\alpha |x|^2}. \quad (4.7) \quad \text{cru}$$

Allora u ammette la rappresentazione (4.6), in particolare è unica.

Dimostrazione. Fissato $t > 0$ consideriamo il processo

$$u(t - s, w_s) \exp \left\{ - \int_0^s du V(w_u) \right\}$$

al variare di $s \in [0, r]$ per $r < t$. Introduciamo inoltre il solito tempo di arresto $\tau_\ell := \inf\{t \geq 0: |w_t| \geq \sqrt{n\ell}\}$. Applicando la formula di Itô (è legale poiché abbiamo supposto $u \in C^{1,2}$) ricaviamo

$$\begin{aligned} u(t-r, w_{r \wedge \tau_\ell}) \exp \left\{ - \int_0^{r \wedge \tau_\ell} du V(w_u) \right\} &= u(t, w_0) \\ + \int_0^{r \wedge \tau_\ell} ds \left[-\partial_s u(t-s, w_s) + \frac{1}{2} \Delta u(t-s, w_s) - u(t-s, w_s) V(w_s) \right] &e^{-\int_0^s du V(w_u)} \\ + M_{r \wedge \tau_\ell}^t & \end{aligned}$$

dove M^t è una martingala. Facendo il valore di attesa ed utilizzando l'equazione (4.5) ricaviamo (grazie all'arresto il termine di martingala ha attesa nulla)

$$u(t, x) = \mathbb{E}_x \left(u(t-r, w_{r \wedge \tau_\ell}) \exp \left\{ - \int_0^{r \wedge \tau_\ell} du V(w_u) \right\} \right).$$

Vogliamo ora passare al limite per $\ell \rightarrow \infty$ e $r \uparrow t$ utilizzando l'ipotesi (4.7). Poiché $V \geq 0$ basta dimostrare che

$$\lim_{\substack{\ell \rightarrow \infty \\ r \uparrow t}} \mathbb{E}_x \left(|u(t-r, w_{r \wedge \tau_\ell})| \mathbb{1}_{\tau_\ell > r} \right) = 0.$$

Per il principio di riflessione, ponendo $w = (w^1, \dots, w^n)$,

$$\begin{aligned} \mathbb{P}_x(\tau_\ell > r) &= \mathbb{P}_x \left(\sup_{t \in [0, r]} |w_t| > \sqrt{n\ell} \right) \leq \sum_{i=1}^n \mathbb{P}_x \left(\sup_{t \in [0, r]} |w_t^i| > \ell \right) \\ &\leq 2 \sum_{i=1}^n [\mathbb{P}_x(w_r^i > \ell) + \mathbb{P}_x(-w_r^i > \ell)] \end{aligned}$$

e concludiamo utilizzando la condizione di crescita (4.7) e stime gaussiane. \square

Si discute ora la formula di Feynman-Kac nel contesto generale di semigruppdi di Feller. Sia S uno spazio polacco e $P_t, t \geq 0$ un semigruppdi di Feller su $C_0(S)$ con generatore $(L, \mathcal{D}(L))$. Sia inoltre $V: S \rightarrow \mathbb{R}$ continua e limitata. Indicando con $(X_t)_{t \geq 0}$ il processo associato, per $t \geq 0$ e $f \in C_0(S)$ definiamo

$$(P_t^V f)(x) := \mathbb{E}_x \left(e^{\int_0^t ds V(X_s)} f(X_t) \right). \quad (4.8) \quad \boxed{\text{sgV}}$$

$\boxed{\text{t: fks}}$

Proposizione 4.7. *La famiglia $P_t^V, t \geq 0$ è un semigruppdi fortemente continuo su $C_0(S)$ con generatore $L + V$ (qui V indica l'operatore di moltiplicazione: $(Vf)(x) := V(x)f(x), f \in C_0(S)$) con dominio $\mathcal{D}(L)$.*

Per la teoria generale di semigruppdi fortemente continui, $P_t^V f$ è quindi la soluzione "semigruppale" del problema di Cauchy

$$\begin{cases} \partial_t u = Lu + Vu \\ u(0) = f. \end{cases}$$

Osserviamo inoltre che, avendo supposto V continua e limitata, l'operatore di moltiplicazione per V è un operatore limitato su $C_0(S)$; coerentemente il dominio di $L + V$ è lo stesso del dominio di L .

Dimostrazione. Cominciamo verificando che P_t^V è un semigruppato. Dati $t, s \geq 0$, per la proprietà di Markov,

$$\begin{aligned} P_{t+s}^V f(x) &= \mathbb{E}_x \left(e^{\int_0^{t+s} du V(X_u)} f(X_{t+s}) \right) = \mathbb{E}_x \left[\mathbb{E}_x \left(e^{\int_0^{t+s} du V(X_u)} f(X_{t+s}) \middle| \mathcal{F}_t \right) \right] \\ &= \mathbb{E}_x \left[e^{\int_0^t du V(X_u)} \mathbb{E}_x \left(e^{\int_t^{t+s} du V(X_u)} f(X_{t+s}) \middle| \mathcal{F}_t \right) \right] \\ &= \mathbb{E}_x \left[e^{\int_0^t du V(X_u)} \mathbb{E}_{X_t} \left(e^{\int_0^s du V(X_u)} f(X_s) \right) \right] \\ &= \mathbb{E}_x \left[e^{\int_0^t du V(X_u)} (P_s^V f)(X_t) \right] = P_t^V P_s^V f(x). \end{aligned}$$

Mostriamo ora che P_t^V è fortemente continuo. Evidentemente,

$$P_t^V f - f = P_t^V f - P_t f + P_t f - f.$$

Per la continuità forte di P_t il secondo termine si annulla per $t \downarrow 0$. D'altra parte, per la definizione di P_t^V ,

$$\left| P_t^V f(x) - P_t f(x) \right| = \left| \mathbb{E}_x \left(\left[e^{\int_0^t ds V(X_s)} - 1 \right] f(X_t) \right) \right| \leq e^{t|V|} t |V| |f|$$

che per $t \downarrow 0$ converge a zero uniformemente per $x \in S$.

Rimane infine da verificare che $L + V$ è il generatore di P_t^V . Dobbiamo dimostrare che se $f \in \mathcal{D}(L)$ allora

$$\lim_{t \downarrow 0} \frac{P_t^V f - f}{t} = Lf + Vf.$$

Poichè L è il generatore di P_t , è sufficiente

$$\lim_{t \downarrow 0} \frac{P_t^V f - P_t f}{t} = Vf.$$

A tal fine,

$$\begin{aligned} \frac{P_t^V f(x) - P_t f(x)}{t} - V(x)f(x) &= \mathbb{E}_x \left(\frac{e^{\int_0^t ds V(X_s)} - 1}{t} f(X_t) \right) - V(x)f(x) \\ &= \mathbb{E}_x \left(\frac{e^{\int_0^t ds V(X_s)} - 1}{t} - V(x) \right) f(x) + \mathbb{E}_x \left(\frac{e^{\int_0^t ds V(X_s)} - 1}{t} [f(X_t) - f(x)] \right). \end{aligned}$$

Poichè $t^{-1}(e^{\int_0^t ds V(X_s)} - 1)$ è limitato, il secondo termine si annulla per $t \downarrow 0$, uniformemente per $x \in S$, grazie alla forte continuità di P_t . Analizziamo ora il primo termine. Poichè $f \in C_0(S)$ possiamo considerare il caso in cui x è in un compatto. Di nuovo grazie alla continuità forte di P_t , per ogni compatto $K \subset\subset S$ e $\varepsilon > 0$,

$$\limsup_{t \downarrow 0} \sup_{x \in K} \mathbb{P}_x(\text{dist}(X_t, x) > \varepsilon) = 0.$$

s:1a Usando la continuità e la limitatezza di V concludiamo quindi la dimostrazione. \square

4.3. Legge dell'arcoseno

Sia w il moto browniano unidimensionale con $w_0 = 0$, fissato $t > 0$ vogliamo scoprire quale sia la frazione del tempo in $[0, t]$ in cui il moto browniano è positivo. Ovvero introduciamo la variabile aleatoria reale A_t (in realtà a valori in $[0, 1]$) definita da

$$A_t := \frac{1}{t} \int_0^t ds \mathbb{1}_{(0, +\infty)}(w_s) \quad (4.9) \quad \text{At}$$

di cui vogliamo ottenere la distribuzione.

Si osserva che l'invarianza di scala del moto browniano implica (verificare!) che la legge di A_t non dipende da t . Poiché $-w$ è un browniano tanto quanto w , la legge di A_t sarà simmetrica rispetto a $1/2$. Si potrebbe pensare che la situazione più probabile sia $A_t \approx 1/2$, ma in effetti i valori più probabili di A_t sono quelli vicini a 0 o a 1. Il seguente risultato è noto come legge dell'arcoseno.

t:1as **Teorema 4.8.** Sia \mathbb{P}_0 la misura di Wiener e A_t la variabile aleatoria definita in (4.9). Per ogni $t > 0$ e $u \in [0, 1]$,

$$\mathbb{P}_0(A_t \leq u) = \int_0^u dy \frac{1}{\pi \sqrt{y(1-y)}} = \frac{2}{\pi} \arcsin \sqrt{u}. \quad (4.10) \quad \text{1as}$$

Dimostrazione. Consideriamo il moto browniano con dato iniziale $w_0 = x$, per $\beta > 0$ introduciamo la funzione

$$u(t, x) := \mathbb{E}_x \left(e^{-\beta t A_t} \right) = \mathbb{E}_x \left(e^{-\beta \int_0^t ds \mathbb{1}_{(0, +\infty)}(w_s)} \right).$$

Grazie alla formula di Feynman-Kac (con truffa: $\mathbb{1}_{(0, +\infty)}$ non è esattamente una funzione continua), u risolve l'equazione

$$\begin{cases} \partial_t u = \frac{1}{2} \partial_x^2 u - \beta \mathbb{1}_{(0, +\infty)} u \\ u(0, \cdot) = 1. \end{cases}$$

Osservando che, al variare di β , $u(t, 0)$ identifica la legge di A_t , dimostreremo il risultato risolvendo esplicitamente l'equazione precedente.

In effetti è conveniente passare alla trasformata di Laplace, ovvero per $\lambda > 0$ introdurre

$$z_\lambda(x) := \int_0^\infty dt e^{-\lambda t} u(t, x) = \int_0^\infty dt e^{-\lambda t} \mathbb{E}_x \left(e^{-\beta t A_t} \right).$$

Con calcolo diretto si verifica che $z \in C^1(\mathbb{R})$ e risolve

$$\begin{cases} \lambda z_\lambda = \frac{1}{2} z_\lambda'' - \beta z_\lambda + 1 & x > 0 \\ \lambda z_\lambda = \frac{1}{2} z_\lambda'' + 1 & x < 0. \end{cases}$$

Poiché z è limitata, troviamo

$$z_\lambda(x) = \begin{cases} C_0 e^{-x\sqrt{2(\lambda+\beta)}} + \frac{1}{\lambda+\beta} & x > 0 \\ C_1 e^{x\sqrt{2\lambda}} + \frac{1}{\lambda} & x < 0 \end{cases}$$

e determiniamo le costanti C_0 e C_1 imponendo $z \in C^1(\mathbb{R})$. Ricaviamo

$$z_\lambda(0) = \frac{1}{\sqrt{\lambda(\lambda+\beta)}}.$$

Poiché la trasformata di Laplace identifica univocamente la distribuzione, per concludere la dimostrazione è sufficiente assumere che (4.10) sia giusta e verificare che

$$\int_0^\infty dt e^{-\lambda t} \mathbb{E}_0(e^{-\beta t A_t}) = \frac{1}{\sqrt{\lambda(\lambda+\beta)}}.$$

Tale calcolo diretto è lasciato al lettore. □

s:minv

s:emi

5. Misure invarianti per processi di diffusione

5.1. Esistenza di misure invarianti

Sia S uno spazio polacco e P_t un semigruppato di Feller su $C_0(S)$. Come in sezione 3.1, diciamo che $\mu \in \mathcal{P}(S)$ è *invariante* (o stazionaria) per P_t (o per il processo associato) sse $\mu = \mu P_t$, $t \geq 0$. Diamo una caratterizzazione infinitesima dell'invarianza in termini del generatore.

t:mii

Lemma 5.1. *Sia P_t un semigruppato di Feller con generatore L con dominio $\mathcal{D}(L)$. La probabilità $\mu \in \mathcal{P}(S)$ è invariante se e solo se per ogni $f \in \mathcal{D}(L)$*

$$\int d\mu Lf = 0. \quad (5.1) \quad \text{mii}$$

Diciamo che $D \subset \mathcal{D}(L)$ è un *core* per L se la chiusura (nella norma del grafico) della restrizione di L a D restituisce L con dominio $\mathcal{D}(L)$. Evidentemente per dimostrare che μ è invariante basta verificare (5.1) per ogni f in un core per L .

Dimostrazione. Se μ è invariante, per ogni $t > 0$ e $f \in C_0(S)$,

$$\int d\mu(x) \frac{1}{t} [P_t f(x) - f(x)] = 0.$$

Se ora $f \in \mathcal{D}(L)$ abbiamo che $[P_t f(x) - f(x)]/t$ converge a $Lf(x)$ uniformemente per $x \in S$. Passando al limite $t \downarrow 0$ ricaviamo quindi $\mu(Lf) = 0$.

Per dimostrare l'implicazione inversa, osserviamo che per $\lambda > 0$ e $f \in \mathcal{D}(L)$, da (5.1) ricaviamo

$$\mu((\lambda - L)f) = \lambda\mu(f).$$

Fissata $g \in C_0(S)$ e detto $R_\lambda = (\lambda - L)^{-1}$ il risolvente di L scegliamo ora $f = R_\lambda g$. Troviamo

$$\mu(g) = \lambda\mu(R_\lambda g) = \mu((1 - \lambda^{-1}L)^{-1}g) = \dots = \mu((1 - \lambda^{-1}L)^{-n}g).$$

Per semigruppato fortemente continui possiamo recuperare il semigruppato dal risolvente, ovvero per ogni $t \geq 0$ e $f \in C_0(S)$

$$\lim_{n \rightarrow \infty} \left(1 - \frac{t}{n}L\right)^{-n} = P_t f.$$

Scegliendo $\lambda = n/t$ e passando al limite per $n \rightarrow \infty$ concludiamo la dimostrazione. \square

Consideriamo il processo di diffusione che risolve l'equazione stocastica

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

dove w è un moto browniano in \mathbb{R}^m , b un campo vettoriale in \mathbb{R}^n e σ un campo a valori matrici $n \times m$. Come al solito chiamiamo $a = \sigma\sigma^\top$ la matrice $n \times n$ di diffusione.

t:emid

Proposizione 5.2. *Sia a limitata e b tale che esistono $R, \alpha > 0$ tali che se $|x| \geq R$ (qui $|\cdot|$ è la norma euclidea in \mathbb{R}^n)*

$$b(x) \cdot \frac{x}{|x|} \leq -\alpha.$$

Allora esiste una misura invariante per il processo X .

Dimostrazione. Fissata $\nu \in \mathcal{P}(\mathbb{R}^n)$ sia ν^T la probabilità definita da

$$\nu^T := \frac{1}{T} \int_0^T dt \nu P_t.$$

Come in Lemma 3.2, è sufficiente mostrare che la famiglia $\{\nu^T\}_{T>0}$ è tight. A tal fine scegliamo $\nu = \delta_0$ ed esibiremo una funzione $\phi: \mathbb{R}^n \rightarrow [0, +\infty)$ divergente all'infinito per cui

$$\sup_{T>0} \nu^T(\phi) < +\infty. \quad (5.2) \quad \text{sufphi}$$

Sia L il generatore di X . Per verificare (5.2) basta mostrare che esistono $\gamma > 0$ e $K < +\infty$ per cui

$$L\phi \leq -\gamma\phi + K. \quad (5.3) \quad \text{cond1}$$

Dalla formula di Ito ricaviamo infatti

$$\phi(X_t) = \phi(X_0) + \int_0^t ds L\phi(X_s) + M_t^\phi$$

dove M^ϕ è una martingala locale. Prendendo il valore di attesa (introducendo prima un tempo di arresto) ed utilizzando (5.3) ricaviamo

$$\mathbb{E}_0(\phi(X_t)) \leq \phi(0) + \int_0^t ds [-\gamma\mathbb{E}_0(\phi(X_s)) + K]$$

che implica

$$\sup_{t>0} \mathbb{E}_0(\phi(X_t)) < +\infty.$$

Si tratta ora di costruire la funzione ϕ con le proprietà desiderate. Scegliamo $\psi \in C^2(\mathbb{R}_+; \mathbb{R}_+)$ tale che per $r > R$ si abbia $\psi(r) = r$ e poniamo

$$\phi(x) := \exp \{ \lambda \psi(|x|) \}$$

con $\lambda > 0$ scelto strada facendo. Via calcolo diretto, dalle ipotesi ricaviamo che esiste una costante $C > 0$ per cui, se $|x| > R$

$$L\phi(x) \leq -\lambda\alpha\phi(x) + C \left[\lambda^2 + \frac{\lambda}{|x|} \right] \phi(x)$$

è ora facile concludere la dimostrazione. □

t:memi

Osservazione 5.3. *La precedente dimostrazione implica invero (esercizio) che la misura invariante costruita μ ha momenti esponenziali: esiste $\gamma > 0$ per cui $\mu(\exp\{\gamma|x|\}) < +\infty$.*

Supponiamo di scegliere un dato iniziale con distribuzione assolutamente continua rispetto a Lebesgue, $\nu = \rho_0 dx$, e vediamo come evolve, assumendo che rimanga assolutamente continua a tutti i tempi: $\nu P_t = \rho_t dx$. Presa f nel dominio di L e derivando rispetto al tempo l'identità $\int dx \rho_0(x) P_t f(x) = \int dx \rho_t(x) f(x)$ troviamo

$$\int dx \partial_t \rho_t(x) f(x) = \int dx \rho_0(x) P_t L f(x) = \int dx \rho_t(x) L f(x)$$

Indicando con L^\top l'aggiunto di L rispetto alla misura di Lebesgue, ovvero

$$L^\top g(x) = \frac{1}{2} \sum_{i,j} \partial_i \partial_j (a_{i,j}(x) g(x)) - \sum_i \partial_i (b_i(x) g(x))$$

“ricaviamo” (è tutto formale) che ρ_t evolve secondo l'equazione di Fokker-Planck

$$\partial_t \rho_t = L^\top \rho_t. \tag{5.4}$$

fp

Supponiamo ora che L sia uniformemente ellittico: esiste $c > 0$ tale che per ogni $x \in \mathbb{R}^n$ e ogni $v \in \mathbb{R}^n$ si ha $v \cdot a(x)v \geq c|v|^2$. In tal caso (5.4) è un'evoluzione uniformemente parabolica che alliscia il dato iniziale. Anche se prendiamo ν singolare (per esempio la delta in un punto) appena $t > 0$ la probabilità νP_t diventa assolutamente continua, con densità strettamente positiva e liscia. In particolare scegliendo ν la misura invariante troviamo che la misura invariante è assolutamente continua rispetto a Lebesgue, ha densità strettamente positiva e liscia. Tale densità è infine soluzione stazionaria di (5.4). In questa situazione (L uniformemente ellittico) si può dimostrare che in effetti la misura invariante è unica, vedi ????

s:rev

5.2. Misure reversibili

Consideriamo un processo di Feller con misura invariante μ . In tal caso il semigruppso associato P_t estende ad un semigruppso di contrazione fortemente continuo su $L^2(S, \mu)$. Infatti, per $f \in C_0(S)$ abbiamo, indicando con $\|\cdot\|$ la norma in $L^2(S, \mu)$,

$$\|P_t f\|_2^2 = \int d\mu (P_t f)^2 = \int d\mu(x) \left[\mathbb{E}_x(f(X_t)) \right]^2 \leq \int d\mu(x) \mathbb{E}_x(f(X_t)^2) = \|f\|_2^2$$

dove abbiamo usando l'invarianza di μ nell'ultimo passaggio. Poiché $C_0(S)$ è denso in $L^2(S, \mu)$, possiamo estendere P_t ad un operatore limitato (di norma minore o uguale a 1) su $L^2(S, \mu)$. Dalla continuità forte di P_t ricaviamo immediatamente la continuità forte di P_t come semigruppso in $L^2(S, \mu)$. Con lieve abuso di notazione, indichiamo con L anche il generatore di P_t come semigruppso in $L^2(S, \mu)$.

Bibliografia

- [BBB] [1] L. Bertini, P. Buttà, S. Brassesco: *Soft and hard wall in a stochastic reaction diffusion equation*. Preprint 2006.
- [B] [2] P. Billingsley: *Convergence of Probability Measures*. Wiley, New York 1968. [N.B. sono state pubblicate edizioni aggiornate]
- [D] [3] J.L. Doob: *Stochastic processes*. Wiley, New York 1953.
- [GS] [4] I.I. Gihman, A.V. Skorohod: *Stochastic differential equations*. Springer, New York, 1972.
- [E] [5] A. Einstein: *Investigations on the theory of the Brownian movement*. Edited with notes by R. Fürth. Translated by A. D. Cowper. Dover Publications, Inc., New York, 1956.
- [FG] [6] P.A. Ferrari, A. Galves: *Costruction of stochastic processes, coupling and regeneration*. Note della “XIII Escuela Venezolana de Matemáticas”, 2000.
- [Fr] [7] M. Freidlin; *Some remarks on the Smoluchowski-Kramers approximation*. J. Statist. Phys. **117**, 617–634 (2004).
- [IW] [8] N. Ikeda, S. Watanabe: *Stochastic differential equations and diffusion processes*. North-Holland Publishing Co., Amsterdam-New York, Tokyo, 1981.
- [KS] [9] I. Karatzas, S.E. Shreve: *Brownian motion and stochastic calculus*. Second edition. Graduate Texts in Mathematics, 113. Springer, New York, 1991.
- [L] [10] T.M. Liggett: *Interacting particle systems*. Springer, New York, 1985.
- [N] [11] E. Nelson: *Dynamical theories of Brownian motion*. Princeton University Press, Princeton, 1967.
- [RY] [12] D. Revuz, M. Yor: *Continuous Martingales and Brownian Motion*. Springer, New York, 1991. [N.B. sono state pubblicate edizioni aggiornate]
- [RW] [13] L.C.G. Rogers, D. Williams: *Diffusions, Markov processes, and martingales. Vol. 1,2*. Cambridge University Press, Cambridge, 2000.
- [S] [14] A.N. Shiryaev: *Probability*. Springer, New York, 1984.
- [SV] [15] D.W. Stroock, S.R.S. Varadhan: *Multidimensional diffusion processes*. Springer, New York, 1979.
- [V] [16] A.D. Ventsel: *Teoria dei processi stocastici*. Edizioni Mir, Mosca 1983.
- [W] [17] D. Williams: *Probability with martingales*. Cambridge University Press, 1991.