

Modelli stocastici ed equazioni di Kolmogorov-Fokker-Planck

A. Barchielli

*Dipartimento di Matematica, Politecnico di Milano,
Piazza Leonardo da Vinci 32, I-20133 Milano, Italy*

Politecnico di Milano — Scuola estiva su:
Campi vettoriali di Hörmander, equazioni differenziali ipoellittiche e
applicazioni
12–16 luglio 2004

Sommario

Le equazioni di Langevin sono equazioni di moto, del tipo delle equazioni di Hamilton, in cui appare una forza proporzionale ad un rumore bianco. La necessità di dare un significato matematico a queste equazioni portò allo sviluppo del calcolo stocastico di Itô e alla nozione di equazione differenziale stocastica. La soluzione di una di queste equazioni è un processo di Markov le cui probabilità di transizione sono legate a oggetti che soddisfano l'equazione di Kolmogorov all'indietro. Quando esistono le densità di probabilità di transizione, sono queste densità che soddisfano l'equazione di Kolmogorov in avanti, o equazione di Fokker-Planck.

La presentazione matematica del calcolo stocastico e delle equazioni differenziali stocastiche richiederebbe troppo tempo per gli scopi di questa Scuola. Queste lezioni intendono solamente dare qualche idea euristica sul calcolo stocastico, per capire cosa sia un'equazione differenziale stocastica e come, a partire da questa, si ottenga l'equazione di Kolmogorov-Fokker-Planck. Questi punti saranno illustrati introducendo esempi di equazioni di interesse per la fisica o, più in generale, per le scienze applicate.

Questi appunti vogliono essere solo una traccia delle lezioni e avranno uno stile estremamente schematico. Per una presentazione più articolata segnalo qualche libro.

Un testo semplice che contiene tutti gli argomenti di cui abbiamo bisogno è:

Z. Schuss, *Theory and Applications of Stochastic Differential Equations* (Wiley, New York, 1980).

Presentazioni matematicamente accurate sono date in:

P. Baldi, *Equazioni Differenziali Stocastiche e Applicazioni*, Quaderni UMI 28 (Pitagora, Bologna, 2000);

Xuerong Mao, *Stochastic Differential Equations and Applications* (Horwood Publishing, Chichester, 1997).

Libri completamente dedicati alle applicazioni sono:

H. Risken, *The Fokker-Planck equation* (Springer, Berlin, 1989);

N.G. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam 1992);

C.W. Gardiner, *Handbook of stochastic methods: for physics, chemistry and the natural sciences* (Springer, Berlin, 1990).

Indice

| | | |
|----------|--|-----------|
| 1 | Introduzione | 3 |
| 1.1 | Le equazioni coinvolte | 3 |
| 1.2 | Un prototipo: le equazioni di Langevin | 4 |
| 2 | L'integrale stocastico | 5 |
| 2.1 | Il processo di Wiener | 5 |
| 2.1.1 | Il processo di Wiener standard unidimensionale | 5 |
| 2.1.2 | Il processo di Wiener d -dimensionale | 7 |
| 2.2 | Necessità di una definizione ad hoc per l'integrale rispetto al processo di Wiener | 7 |
| 2.3 | L'integrale stocastico | 9 |
| 2.3.1 | Filtrazioni e processi adattati | 9 |
| 2.3.2 | L'integrale rispetto al tempo | 10 |
| 2.3.3 | L'integrale di Itô per processi semplici | 10 |
| 2.3.4 | L'isometria di Itô | 12 |
| 2.3.5 | Altre proprietà dell'integrale stocastico | 15 |
| 2.3.6 | Estensione della classe di integrandi | 16 |
| 2.4 | Il differenziale stocastico e la formula di Itô | 16 |
| 2.5 | L'integrale e il differenziale stocastici multidimensionali | 19 |
| 3 | Equazioni differenziali stocastiche | 20 |
| 3.1 | Una classe di EDS | 20 |
| 3.2 | Esistenza e unicità delle soluzioni | 21 |
| 3.3 | Esempi | 22 |
| 3.3.1 | Equazione di Langevin | 22 |
| 3.3.2 | Un mercato finanziario | 23 |
| 3.3.3 | Modello di diffusione di fase per un laser | 23 |

| | | |
|----------|--|-----------|
| 4 | Le equazioni di Kolmogorov e di Fokker-Planck | 23 |
| 4.1 | Il generatore delle diffusioni | 23 |
| 4.2 | Probabilità di transizione e processi di Markov | 24 |
| 4.3 | L'equazione di Kolmogorov all'indietro | 25 |
| 4.4 | L'equazione di Fokker-Planck o di Kolmogorov in avanti | 26 |

1 Introduzione

Scopo di queste lezioni è di spiegare cosa sono le equazioni differenziali stocastiche (EDS), o almeno una sottoclasse di EDS, e mostrare come sono legate alle equazioni di Fokker-Planck e ad altre equazioni alle derivate parziali.

Questo legame aiuta nelle applicazioni; spesso l'esame del problema "fisico" porta a costruire il modello stocastico e si passa poi alla PDE legata. Aiuta anche nelle soluzioni numeriche: la rappresentazione stocastica di una PDE può essere il punto di partenza per una simulazione di tipo MonteCarlo.

1.1 Le equazioni coinvolte

EDS "alla Itô"

$$\begin{cases} dX(t) = b(X(t), t)dt + \sigma(X(t), t)dW(t) \\ X(t_0) = x_0 \end{cases} \quad (1)$$

$t \geq t_0$, $x_0 \in \mathbb{R}^m$, b vettore reale, σ matrice reale

$dW(t)$ "=" $\dot{W}(t)dt$ $\dot{W}(t)$: "rumore bianco" d -dimensionale

Equazione di Fokker-Planck associata:

$$\begin{cases} \frac{\partial}{\partial t} p(x, t|x_0, t_0) \\ = \left[- \sum_i \frac{\partial}{\partial x_i} b_i(x, t) + \frac{1}{2} \sum_{ij} \frac{\partial^2}{\partial x_i \partial x_j} a_{ij}(x, t) \right] p(x, t|x_0, t_0) \\ p(x, t_0|x_0, t_0) = \delta(x - x_0) \end{cases} \quad (2)$$

$$a_{ij} := \sum_{k=1}^d \sigma_{ik} \sigma_{jk} \quad \text{o} \quad a := \sigma \sigma^T \quad (3)$$

Primo problema. Che equazione è (1)? Che senso ha? Si scrive in forma differenziale e non con le derivate per ricordare che non è un'ODE dipendente da un qualche parametro casuale; $\dot{W}(t)$ potrebbe avere un senso solo come distribuzione. Spesso viene scritta in forma integrale

$$X(t) = x_0 + \int_{t_0}^t b(X(s), s)ds + \int_{t_0}^t \sigma(X(s), s)dW(s). \quad (4)$$

Tuttavia l'ultimo "integrale" ha diritto a questo nome perché limite di somme di tanti piccoli termini, ma non c'è nessuna misura rispetto a cui si integri nel senso della teoria della misura e dell'integrazione.

Secondo problema. Come si passa da (1) a (2)? Dovremo dare il legame matematico tra (1) e (2) e anche spiegare le affermazioni: "(1) dà una descrizione 'per traiettorie' del processo stocastico di Markov $X(t)$ ", mentre "(2) fornisce le probabilità di transizione per questo processo"; $p(x, t|x_0, t_0)$ sarà la densità della probabilità di passare da x_0 al tempo t_0 a x al tempo t .

1.2 Un prototipo: le equazioni di Langevin

Consideriamo le equazioni di Hamilton per una particella in un campo di forze $F(q, p, t)$ e immersa in un fluido. Possiamo pensare di rappresentare schematicamente il fluido con una ulteriore forza di natura stocastica $\sigma w(t)$ e scriviamo

$$\begin{cases} \dot{q} = \frac{p}{m} \\ \dot{p} = F(q, p, t) + \sigma w(t) \end{cases} \quad (5)$$

σ è una costante che misura l'intensità del rumore e che introduciamo per convenienza. Il rumore $w(t)$ è casuale e vogliamo che sia il più semplice possibile e il più "caotico" possibile. Chiediamo che sia gaussiano (semplicità) a media nulla (un'eventuale media può essere inglobata in F). Dato che il rumore rappresenta le interazioni della nostra particella con le molecole del fluido, che si comportano in modo caotico e si "dimenticano in fretta" degli urti subiti, possiamo chiedere che i tempi di correlazione del rumore siano piccolissimi, al limite nulli, e chiediamo (nel caso unidimensionale)

$$\text{Cov}[w(t), w(s)] = \delta(t - s). \quad (6)$$

Così facendo però abbiamo che $w(t)$ non è una funzione "liscia" e sorge il problema di dar senso all'equazione di Langevin (5). Finché σ è una costante o, al più, dipende dal tempo, si può aggirare questa difficoltà in vari modi. Se addirittura si suppone F al più lineare in q e p (particella libera o oscillatore armonico con attrito) si può esibire esplicitamente la soluzione $q(t), p(t)$ di (5) come processo gaussiano. Come vedremo, i problemi sorgono quando si assume che l'intensità del rumore possa dipendere dalla posizione e/o dalla velocità: $\sigma = \sigma(q, p, t)$. È in questo caso che si rende necessaria la teoria degli integrali stocastici di Itô.

Si noti che, se identifichiamo $w(t)dt$ con $dW(t)$ e prendiamo per buona la corrispondenza tra le equazioni (1) e (2), l'equazione di Fokker-Planck associata a (5) risulta essere (indicando con \tilde{p} la densità di probabilità di

transizione per non confonderla con le variabili di momento p)

$$\begin{aligned} \frac{\partial}{\partial t} \tilde{p}(q, p, t | q_0, p_0, t_0) = & \left[- \sum_i \frac{\partial}{\partial q_i} \frac{p_i}{m} - \sum_i \frac{\partial}{\partial p_i} F_i(q, p, t) \right. \\ & \left. + \frac{1}{2} \sum_{ij} \frac{\partial^2}{\partial p_i \partial p_j} (\sigma \sigma^T)_{ij} \right] \tilde{p}(q, p, t | q_0, p_0, t_0). \quad (7) \end{aligned}$$

Vediamo così che il precursore di tutte le EDS porta ad un'equazione parabolica degenera: le derivate seconde non toccano le variabili di tipo q .

2 L'integrale stocastico

2.1 Il processo di Wiener

Il primo passo per dar senso all'EDS (1) è di introdurre il processo $W(t)$ di cui $\dot{W}(T)$ sarà la derivata formale.

Spazio di probabilità: (Ω, \mathcal{F}, P) , uno spazio misurabile con una misura di massa 1.

Variabile aleatoria (reale): $X : \Omega \rightarrow \mathbb{R}$ misurabile. Se \mathbb{R} viene sostituito da uno spazio misurabile qualsiasi abbiamo la più generale variabile aleatoria.

Variabili aleatorie indipendenti: le probabilità congiunte fattorizzano.

Processo stocastico: una collezione di variabili aleatorie, per noi sempre indicizzate dai numeri positivi (il tempo) $\{X(t), t \geq 0\}$. Per noi i processi saranno sempre a valori in un qualche \mathbb{R}^m

q.c. = quasi certamente = quasi ovunque rispetto a P

Modificazioni o versioni: il processo $Y(t)$ è una modificazione o versione di $X(t)$ se sono entrambi definiti sullo stesso spazio di probabilità e per ogni t si ha $Y(t) = X(t)$ q.c., cioè $P[Y(t) = X(t)] = 1, \forall t$.

2.1.1 Il processo di Wiener standard unidimensionale

Un processo stocastico reale $\{W(t), t \geq 0\}$, in uno spazio di probabilità (Ω, \mathcal{F}, P) , si chiama processo di Wiener (standard, unidimensionale) se

- (i) $W(0) = 0$ q.c.,
- (ii) è un processo a incrementi indipendenti, cioè per ogni scelta dell'intero n e dei tempi $0 \leq t_1 < t_2 < \dots < t_n$ le variabili aleatorie $W(t_2) - W(t_1), W(t_3) - W(t_2), \dots, W(t_n) - W(t_{n-1})$ sono indipendenti,
- (iii) per $t > s$, $W(t) - W(s)$ è distribuita come una normale di media 0 e varianza $t - s$.

Il termine ‘standard’ si riferisce alla scelta del coefficiente di $t - s$ nella varianza al punto (iii). Nella letteratura matematica questo processo è spesso chiamato moto Browniano, perché è il più semplice modello del fenomeno fisico del moto Browniano; tuttavia per distinguere il fenomeno da questo primo modello è meglio usare il nome di processo di Wiener.

Le proprietà (i)-(iii) individuano univocamente tutte le probabilità finito-dimensionali del processo; si può verificare che si ha

$$P[W(t_1) \in F_1, W(t_2) \in F_2, \dots, W(t_n) \in F_n] = \int_{F_1 \times F_2 \times \dots \times F_n} p(t_n - t_{n-1}; x_n | x_{n-1}) \cdots p(t_2 - t_1; x_2 | x_1) p(t_1; x_1 | 0) dx_1 dx_2 \cdots dx_n, \quad (8)$$

dove le densità di probabilità di transizione sono date da

$$p(t; x|y) = \frac{1}{\sqrt{2\pi t}} \exp\left[-\frac{(x-y)^2}{2t}\right], \quad t > 0 \quad (9)$$

Le proprietà caratterizzanti il processo di Wiener non fissano lo spazio di probabilità dove è realizzato; esiste tuttavia una realizzazione canonica. Prendiamo come Ω lo spazio di tutte le funzioni continue da \mathbb{R}_+ a \mathbb{R} uscenti da zero e poniamo

$$W(t; \omega) := \omega(t). \quad (10)$$

Sia \mathcal{F} la più piccola σ -algebra che rende misurabile tutte le proiezioni canoniche $\omega \mapsto \omega(t)$: $\mathcal{F} := \sigma(W(t), t \geq 0)$. Si dimostra poi che esiste un’unica misura di probabilità P (la misura di Wiener) rispetto a cui il processo $W(t)$ (10) ha le (8) come distribuzioni finito dimensionali.

Il rumore bianco. Vediamo adesso perché il processo di Wiener ha a che fare con il non ben definito rumore bianco, che è l’oggetto che ci interessa.

Calcoliamo la funzione di covarianza del processo; scegliamo $t > s$ e usiamo l’indipendenza degli incrementi:

$$\begin{aligned} \text{Cov}[W(t), W(s)] &= \text{Cov}[W(t) - W(s) + W(s), W(s)] \\ &= \text{Cov}[W(t) - W(s), W(s)] + \text{Cov}[W(s), W(s)] = \text{Var}[W(s)] = s. \end{aligned}$$

In generale si ha dunque

$$\text{Cov}[W(t), W(s)] = \min\{t, s\} \quad (11)$$

Dato che le trasformazioni lineari mandano gaussiane in gaussiane, l’ipotetico processo $\dot{W}(t)$ sarà gaussiano e, per la relazione precedente, δ -correlato, proprio le proprietà che volevamo dal rumore bianco.

2.1.2 Il processo di Wiener d -dimensionale

Con processo di Wiener d -dimensionale intendiamo semplicemente d processi di Wiener unidimensionali indipendenti. Dunque ciascuna componente di $W(t)$ è un processo di Wiener e vale $\text{Cov}[W_i(t), W_j(s)] = \delta_{ij} \min\{t, s\}$. Assumiamo inoltre di prendere sempre una realizzazione a traiettorie continue: $t \mapsto W(t; \omega) \in \mathbb{R}^d$ continua per ogni $\omega \in \Omega$.

Densità di probabilità di transizione: $x, y \in \mathbb{R}^d, t > 0$,

$$p(t; x|y) = \frac{1}{(2\pi t)^{d/2}} \exp\left[-\frac{|x-y|^2}{2t}\right]. \quad (12)$$

Versione canonica: $\Omega = \{\omega : \omega \in C(\mathbb{R}_+, \mathbb{R}^d), \omega(0) = 0\}$, $W(t; \omega) = \omega(t)$, $\mathcal{F} := \sigma(W(t), t \geq 0)$, la misura di Wiener P è di nuovo univocamente determinata dalle distribuzioni finito-dimensionali.

2.2 Necessità di una definizione ad hoc per l'integrale rispetto al processo di Wiener

Visto che il rumore bianco è un oggetto molto singolare si cerca di dare senso ad un'equazione in forma integrale del tipo (4) e siamo perciò indotti a cercare di dare senso ad integrali rispetto al processo di Wiener. Si può però dimostrare che “le traiettorie $t \mapsto W(t; \omega)$ del processo di Wiener non sono q.c. a variazione finita su alcun intervallo”; dunque non si può certo definire un integrale di Stieltjes $\int_0^t f(t; \omega) dW(t; \omega)$, ω per ω . La definizione di questo integrale non potrà essere fatta ‘per traiettorie’, ma dovrà contenere anche la probabilità come ingrediente essenziale.

In questa sezione $W(t)$ è unidimensionale.

Fissiamo un intervallo $[0, T]$ e una sua partizione $0 = t_0 < t_1 < \dots < t_{n-1} < t_n = T$. Siano $\Delta t_j = t_{j+1} - t_j$ e $\Delta W_j = W(t_{j+1}) - W(t_j)$; quando considereremo un limite sarà per $\max_j \Delta t_j \downarrow 0$.

Incominciamo col tentativo di definire $\int_0^T f(t) dW(t)$ con f funzione deterministica (non è aleatoria, non dipende da ω). Possiamo pensare di definire l'integrale come un qualche limite delle somme integrali

$$S_n = \sum_j f(t_j^*) \Delta W_j, \quad (13)$$

dove t_j^* è un qualche punto intermedio fra t_j e t_{j+1} .

S_n è lineare negli incrementi del Wiener, f è deterministica $\Rightarrow S_n$ è gaussiana e questa proprietà si mantiene anche nei passaggi al limite.

Basta dunque trovare media e varianza per avere la legge di S_n . Banalmente $\mathbb{E}[S_n] = 0$ perché gli incrementi del Wiener hanno media nulla; resta da trovare la varianza. Per l'indipendenza degli incrementi

$$\mathbb{E}[\Delta W_j \Delta W_i] = \delta_{ij} \text{Var}[\Delta W_j] = \delta_{ij} \Delta t_j \quad (14)$$

e poi

$$\begin{aligned}\text{Var}[S_n] &= \mathbb{E}[S_n^2] = \sum_{ij} f(t_j^*)f(t_i^*) \mathbb{E}[\Delta W_j \Delta W_i] \\ &= \sum_j f(t_j^*)^2 \Delta t_j \rightarrow \int_0^T f(t)^2 dt. \quad (15)\end{aligned}$$

Dunque, se f è quadrato integrabile, è ragionevole e matematicamente consistente dire che $\int_0^T f(t)dW(t)$ segue una legge normale di media 0 e varianza $\int_0^T f(t)^2 dt$. Si può perfezionare questa definizione e si ottiene così l'“integrale di Wiener”; però i possibili integrandi sono solo le funzioni deterministiche. Si noti che nella definizione entra il fatto che W sia un processo gaussiano, dunque viene coinvolta la probabilità e non solo le traiettorie.

Per dare senso all'equazione (4) abbiamo però bisogno di integrare anche processi stocastici e qui i problemi aumentano come mostra il seguente esempio. Consideriamo $\int_0^t W(t)dW(t)$; comunque lo si possa definire, consideriamo due diverse discretizzazioni:

- integrando nel tempo iniziale

$$I_1 = \sum_j W(t_j) \Delta W_j; \quad (16)$$

- integrando nel tempo a metà intervallo

$$I_2 = \sum_j W(t_j^*) \Delta W_j, \quad t_j^* := \frac{t_j + t_{j+1}}{2}. \quad (17)$$

Ancora prima di fare ogni tipo di limite abbiamo

$$\begin{aligned}\mathbb{E}[I_1] &= \sum_j \mathbb{E}[W(t_j)(W(t_{j+1}) - W(t_j))] \\ &= \sum_j \mathbb{E}[W(t_j)] \mathbb{E}[W(t_{j+1}) - W(t_j)] = 0, \quad (18)\end{aligned}$$

$$\begin{aligned}\mathbb{E}[I_2] &= \sum_j \mathbb{E}[W(t_j^*)(W(t_{j+1}) - W(t_j))] = \sum_j \mathbb{E}[W(t_j^*)(W(t_j^*) - W(t_j))] \\ &= \sum_j \mathbb{E}[(W(t_j^*) - W(t_j))^2] = \sum_j \left(\frac{t_j + t_{j+1}}{2} - t_j \right) = \frac{1}{2} \sum_j \Delta t_j = \frac{T}{2}.\end{aligned}$$

Questo semplice calcolo mostra già che ci saranno più possibili definizioni, dipendenti dal tipo di discretizzazione; questo è un riflesso del fatto che le traiettorie non sono a variazione finita. La discretizzazione in I_1 porterà all'integrale di Itô e la discretizzazione in I_2 porterà all'integrale di Stratonovich (che non vedremo).

2.3 L'integrale stocastico

L'idea di integrale stocastico è di partire da un processo costante a tratti $X(t) = \sum_{j=0}^{n-1} e_j 1_{[t_j, t_{j+1})}(t)$ e di definire l'integrale come $\int_0^T X(t) dW(t) = \sum_j e_j \Delta W_j$. Poi si studia a quale classe di integrandi si può estendere questa definizione usando opportuni limiti.

Una prima osservazione è che si parte da partizioni di $[0, T]$ in sottointervalli, come per l'integrale alla Riemann; non si usano partizioni più generali, come nell'integrale alla Lebesgue.

Una seconda osservazione è che la proprietà che rendeva semplice l'integrale di tipo I_1 è il fatto che in ciascun termine l'integrando e l'incremento del Wiener fossero indipendenti. Questo può essere mantenuto anche nel caso generale se chiediamo che $X(t)$ dipenda dal Wiener fino al tempo t e non dagli incrementi $W(t+s) - W(t)$, $s > 0$. Si dice che l'integrando deve essere adattato o non anticipativo. Vediamo di formalizzare queste idee.

2.3.1 Filtrazioni e processi adattati

Sia $W(t)$ un processo di Wiener d -dimensionale nello spazio di probabilità completo (Ω, \mathcal{F}, P) . Sia \mathcal{F}_t la σ -algebra generata da $\{W(s), 0 \leq s \leq t\}$ e da tutti gli insiemi di probabilità nulla di \mathcal{F} ; \mathcal{F}_t è l'insieme degli eventi di cui sappiamo se si sono avverati o no se abbiamo osservato il processo di Wiener fino a t .

- Per l'indipendenza degli incrementi del Wiener, $W(t+s) - W(t)$, $s > 0$, è indipendente da \mathcal{F}_t , cioè $W(t+s) - W(t)$ è indipendente da ogni variabile aleatoria \mathcal{F}_t -misurabile.
- La famiglia $\{\mathcal{F}_t, t \geq 0\}$ è una *filtrazione*, cioè una famiglia crescente di sotto- σ -algebre: $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ per $s \leq t$.
- Si dimostra che tale filtrazione è *continua* nel senso che $\mathcal{F}_t = \bigvee_{0 \leq s < t} \mathcal{F}_s = \bigcap_{u > t} \mathcal{F}_u$.
- ◇ Un processo $\{X(t), t \geq 0\}$ si dice *adattato* o non anticipativo se $X(t)$ è \mathcal{F}_t -misurabile per ogni t .

Oltre agli integrali stocastici, noi avremo da fare anche integrali sul tempo, valori attesi (cioè integrali su ω), scambiare questi integrali utilizzando Fubini... Avremo perciò bisogno della misurabilità congiunta in t e ω .

- ◇ Un processo $\{X(t), t \geq 0\}$ si dice *misurabile* se $(t, \omega) \mapsto X(t; \omega)$ è $\mathcal{B}_{\mathbb{R}_+} \otimes \mathcal{F}$ -misurabile su $\mathbb{R}_+ \times \Omega$.
- ◇ Un processo $\{X(t), t \geq 0\}$ si dice *progressivamente misurabile* o progressivo se, $\forall T > 0$, $(t, \omega) \mapsto X(t; \omega)$ è $\mathcal{B}_{(0, T]} \otimes \mathcal{F}_T$ -misurabile su $[0, T] \times \Omega$. Un processo progressivo è misurabile e adattato.

- Per le proprietà della nostra filtrazione, si dimostra che un processo continuo a destra e adattato è progressivamente misurabile.

2.3.2 L'integrale rispetto al tempo

Se X è un processo misurabile, posso considerare gli integrali fatti traiettoria per traiettoria (ω fissato)

$$\left(\int_a^b X(t) dt \right) (\omega) := \int_a^b X(t; \omega) dt; \quad (19)$$

l'integrale può esistere o non esistere, si chiede la misurabilità congiunta di $X(t; \omega)$ in t e ω per garantire che l'integrale sia una variabile aleatoria. Se

$$\int_a^b \mathbb{E}[|X(t)|] dt < +\infty, \quad (20)$$

si ha allora per il teorema di Fubini

$$\int_a^b \mathbb{E}[X(t)] dt = \mathbb{E} \left[\int_a^b X(t) dt \right]. \quad (21)$$

Sia $X(t)$, $0 \leq t \leq T$ un processo reale progressivamente misurabile con

$$\int_0^T |X(t; \omega)| dt < +\infty \quad \text{q.c.}$$

- Si modifica (su insiemi di probabilità nulla) il processo in modo che l'integrale esista certamente; la versione di X può essere presa ancora progressiva.
- Per costruzione $\int_0^t X(s) ds$, $t \in [0, T]$, è un processo continuo e adattato, dunque è anche progressivo.

2.3.3 L'integrale di Itô per processi semplici

Incominciamo a costruire l'integrale stocastico su $(0, T)$ rispetto ad una componente del processo di Wiener. Fissiamo dunque un tempo finale $T > 0$ e una componente del nostro processo di Wiener, diciamo la prima; poniamo $B(t) := W_1(t)$.

- Un processo reale $X(t)$, $t \in [0, T]$ è detto *semplice* se può essere rappresentato come

$$X(t) = \sum_{j=0}^{n-1} e_j 1_{[t_j, t_{j+1})}(t) \quad (22)$$

dove $0 = t_0 < t_1 < t_2 < \dots < t_n = T$ e le e_j sono variabili aleatorie \mathcal{F}_{t_j} -misurabili e quadrato integrabili, $\mathbb{E}[|e_j|^2] < +\infty$. Indichiamo con $\mathcal{S}(T)$ l'insieme dei processi semplici con $t \in [0, T)$.

- È immediato vedere che un processo semplice è adattato, misurabile, continuo da destra e dunque anche progressivo.
- Definiamo l'integrale stocastico di un processo semplice con rappresentazione (22) come

$$\int_0^T X(t)dB(t) := \sum_{j=0}^{n-1} e_j \Delta B_j, \quad \Delta B_j := B(t_{j+1}) - B(t_j). \quad (23)$$

Nel seguente Lemma raccogliamo le proprietà cruciali dell'integrale stocastico, proprietà che in particolare permetteranno di dare una prima estensione della definizione.

Lemma 1. *Siano X e Y due processi semplici ed a e b due numeri reali; si ha allora*

(i) $\int_0^T X(t)dB(t) \in L^2(\Omega, \mathcal{F}_T, P),$

(ii) $\mathbb{E} \left[\int_0^T X(t)dB(t) \right] = 0,$

(iii) $\mathbb{E} \left[\left| \int_0^T X(t)dB(t) \right|^2 \right] = \mathbb{E} \left[\int_0^T |X(t)|^2 dt \right],$

(iv) $\int_0^T (aX(t) + bY(t)) dB(t) = a \int_0^T X(t)dB(t) + b \int_0^T Y(t)dB(t).$

Dimostrazione. Dato che e_j è \mathcal{F}_{t_j} -misurabile e ΔB_j è indipendente da \mathcal{F}_{t_j} , e_j e ΔB_j sono indipendenti; dato che entrambi sono integrabili, per l'indipendenza è integrabile anche il prodotto e la media del prodotto è il prodotto delle medie. Per la linearità del valore atteso si ha dunque

$$\mathbb{E} \left[\int_0^T X(t)dB(t) \right] = \sum_{j=0}^{n-1} \mathbb{E} [e_j \Delta B_j] = \sum_{j=0}^{n-1} \mathbb{E}[e_j] \mathbb{E}[\Delta B_j] = 0$$

e questo dimostra (ii).

$\int_0^T X(t)dB(t)$ è banalmente \mathcal{F}_T -misurabile.

$$\left| \int_0^T X(t)dB(t) \right|^2 = \sum_i e_i^2 (\Delta B_i)^2 + 2 \sum_{i,j:i < j} (e_i \Delta B_i)(e_j \Delta B_j)$$

e_i è quadrato integrabile e così ΔB_i ; e_i^2 e $(\Delta B_i)^2$ sono integrabili e indipendenti: dunque il prodotto è integrabile ed $e_i \Delta B_i$ è quadrato integrabile. Infine $e_i \Delta B_i e_j \Delta B_j$ è integrabile essendo un prodotto di due variabili aleatorie quadrato integrabili. Questo dimostra (i).

Per $i < j$, $e_i e_j \Delta B_i$ e ΔB_j sono indipendenti; così si ha

$$\begin{aligned} \mathbb{E} \left[\left| \int_0^T X(t) dB(t) \right|^2 \right] &= \sum_i \mathbb{E} [e_i^2 (\Delta B_i)^2] + 2 \sum_{i,j:i < j} \mathbb{E} [(e_i \Delta B_i)(e_j \Delta B_j)] \\ &= \sum_i \mathbb{E} [e_i^2] \mathbb{E} [(\Delta B_i)^2] + 2 \sum_{i,j:i < j} \mathbb{E} [e_i e_j \Delta B_i] \mathbb{E} [\Delta B_j] \\ &= \sum_i \mathbb{E} [e_i^2] (t_{i+1} - t_i) = \mathbb{E} \left[\int_0^T |X(t)|^2 dt \right] \end{aligned}$$

e questo dimostra (iii). Infine (iv) è banale. \square

Se pensiamo all'integrale stocastico come a una qualche operazione su uno spazio di processi, la (iv) ci dice che quest'operazione è lineare, la (i) ci dice che il risultato finale vive in uno spazio di Hilbert e il lato sinistro di (iii) ne rappresenta la norma al quadrato. Se strutturiamo lo spazio degli integrandi come spazio di Hilbert con norma quadra il lato destro di (iii), abbiamo subito che l'integrale stocastico si estende come isometria.

2.3.4 L'isometria di Itô

L'intervallo $[0, T]$ è fissato.

Consideriamo equivalenti due processi se

$$P \left[\int_0^T |X(t) - X'(t)| dt = 0 \right] = 1. \quad (24)$$

- Indichiamo con $M^2(T)$ lo spazio delle classi di equivalenza di processi reali, progressivamente misurabili e tali che

$$\mathbb{E} \left[\int_0^T |X(t)|^2 dt \right] < +\infty. \quad (25)$$

- Si poteva sostituire “progressivo” con “misurabile e adattato” e far vedere che in ogni classe di equivalenza c'era almeno un rappresentativo progressivo. È usuale passare dalla classe di equivalenza al rappresentativo e viceversa senza dichiararlo ogni volta.
- Si ha $\mathcal{S}(T) \subset M^2(T)$.
- Evidentemente $M^2(T)$ è un sottospazio lineare di $L^2((0, T) \times \Omega, \mathcal{B}_{(0,T)} \otimes \mathcal{F}_T, Leb \otimes P)$; si dimostra che è chiuso. Dunque $M^2(T)$ è uno spazio di Hilbert reale con norma

$$\|X\|_{M^2(T)}^2 = \mathbb{E} \left[\int_0^T |X(t)|^2 dt \right] \equiv \int_0^T \mathbb{E} [|X(t)|^2] dt. \quad (26)$$

Lemma 2. Per ogni $X \in M^2(T)$ esiste una successione $\{X_n\}$ di processi semplici e limitati (in ω oltre che in t) tale che

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\int_0^T |X(t) - X_n(t)|^2 dt \right] = 0. \quad (27)$$

Diamo solo una vaga idea dei tre passi della dimostrazione.

1. Un processo in $M^2(T)$ può essere approssimato da una successione di processi limitati in $M^2(T)$.
2. Un processo limitato in $M^2(T)$ può essere approssimato da una successione di processi limitati continui in $M^2(T)$.
3. Un processo limitato continuo in $M^2(T)$ può essere approssimato da una successione di processi semplici limitati.

Questo Lemma ci dice in particolare che $\mathcal{S}(T)$ è denso in $M^2(T)$.

Poniamo $I_T(X) := \int_0^T X(t) dB(t)$ per $X \in \mathcal{S}(T)$; per la proprietà (iv) I_T è lineare. La proprietà (iii) diviene

$$\|I_T(X)\|_{L^2(\Omega, \mathcal{F}_T, P)} = \|X\|_{M^2(T)}. \quad (28)$$

Dunque

$$I_T : \mathcal{S}(T) \rightarrow L^2(\Omega, \mathcal{F}_T, P) \quad (29)$$

è un operatore lineare tra spazi di Hilbert, densamente definito, limitato: per continuità può essere univocamente esteso ad un operatore lineare limitato definito su tutto $M^2(T)$. L'*isometria di Itô* (28) continua a valere per l'estensione; dunque $I_T(\cdot) = \int_0^T \cdot dB(t)$ è un operatore lineare isometrico da $M^2(T)$ a $L^2(\Omega, \mathcal{F}_T, P)$; questo estende le proprietà (i), (iii) e (iv) del Lemma 1. Per continuità vale anche la proprietà (ii): il valore atteso di un integrale stocastico con integrando di tipo M^2 è nullo.

Implicitamente abbiamo definito l'integrale stocastico come limite in media quadratica di successioni di integrali di processi semplici, ma abbiamo ributtato sulla teoria degli operatori lineari limitati tra spazi di Hilbert l'onere di garantire l'indipendenza dalla successione approssimante e del fatto che le proprietà essenziali sopravvivono al passaggio al limite.

Se $X \in M^2(T)$ ed è continuo si può prendere come processo semplice approssimante una sua discretizzazione. Dunque prendiamo una partizione $0 = t_0 < t_1 < \dots < t_n = T$ e indichiamo con "lim" il limite, in L^2 o M^2 a seconda dei casi, quando il passo della discretizzazione va a zero. Poniamo

$$X_d(t) = \sum_j X(t_j) 1_{[t_j, t_{j+1})}(t); \quad (30)$$

si ha in M^2

$$\lim X_d(t) = X(t). \quad (31)$$

Dunque per processi continui vale la definizione “elementare” di integrale stocastico

$$\int_0^T X(t)dB(t) = \lim \sum_j X(t_j) [B(t_{j+1}) - B(t_j)]. \quad (32)$$

Due esempi fondamentali.

$$\int_0^T B(t)dB(t) = \frac{1}{2} B(T)^2 - \frac{T}{2}, \quad (33)$$

$$\int_0^T tdB(t) = TB(T) - \int_0^T B(t)dt. \quad (34)$$

Dobbiamo dimostrare queste relazioni a partire dalla definizione; usiamo l'abbreviazione $\Delta Y_i = Y(t_{i+1}) - Y(t_i)$. Iniziamo dalla seconda relazione. Abbiamo

$$\begin{aligned} \int_0^T tdB(t) &= \lim \sum_i t_i \Delta B_i, \\ \int_0^T B(t)dt &= \lim \sum_i B(t_{i+1}) \Delta t_i; \end{aligned}$$

nell'ultimo integrale abbiamo inserito $B(t_{i+1})$ per convenienza: in un integrale di Riemann di una funzione continua non conta il punto intermedio scelto. Sommando le due relazioni abbiamo

$$\begin{aligned} \int_0^T tdB(t) + \int_0^T B(t)dt &= \lim \sum_i [B(t_{i+1})t_{i+1} - B(t_{i+1})t_i \\ &\quad + t_i B(t_{i+1}) - t_i B(t_{i+1})] = \lim \sum_i [B(t_{i+1})t_{i+1} - t_i B(t_{i+1})] = TB(T). \end{aligned}$$

Per quanto riguarda il primo integrale

$$\begin{aligned} \int_0^T B(t)dB(t) &= \lim \sum_i B(t_i) \Delta B_i \\ &= \frac{1}{2} \lim \sum_i \left\{ [B(t_{i+1})^2 - B(t_i)^2] - [B(t_{i+1}) - B(t_i)]^2 \right\} \\ &= \frac{1}{2} B(T)^2 - \frac{1}{2} \lim \sum_i [B(t_{i+1}) - B(t_i)]^2 = \frac{1}{2} B(T)^2 - \frac{T}{2} - \frac{1}{2} \lim \sum_i Y_i, \end{aligned}$$

dove abbiamo posto $Y_i := [B(t_{i+1}) - B(t_i)]^2 - (t_{i+1} - t_i)$. Le Y_i sono indipendenti a media nulla e varianza

$$\text{Var}[Y_i] = (t_{i+1} - t_i)^2 \mathbb{E}[(Z_i^2 - 1)^2], \quad Z_i := \frac{\Delta B_i}{\sqrt{\Delta t_i}}.$$

Ma Z_i è normale standard; dunque $\mathbb{E}[(Z_i^2 - 1)^2] = c$, una costante indipendente da i e

$$\lim \text{Var} \left[\sum_i Y_i \right] = \lim \sum_i (t_{i+1} - t_i)^2 c = 0,$$

che dimostra la convergenza in media quadratica.

2.3.5 Altre proprietà dell'integrale stocastico

Se abbiamo $X \in M^2(T)$ e $0 \leq a < b \leq T$, si può facilmente definire $\int_a^b X(t)dB(t)$; non ripetiamo tutta la procedura. Valgono le seguenti proprietà.

- Formula del punto intermedio: se $a < u < b$

$$\int_a^b X(t)dB(t) = \int_a^u X(t)dB(t) + \int_u^b X(t)dB(t). \quad (35)$$

- Se Z è una variabile aleatoria limitata \mathcal{F}_a -misurabile, allora

$$Z \int_a^b X(t)dB(t) = \int_a^b ZX(t)dB(t). \quad (36)$$

La dimostrazione è banale: le due proprietà valgono per processi semplici e restano vere al limite. Si noti il significato della seconda: se Z è limitata e \mathcal{F}_T -misurabile, ma non \mathcal{F}_a -misurabile, $Z \int_a^b X(t)dB(t)$ ha perfettamente senso e sta in $L^2(\Omega, \mathcal{F}_T, P)$, mentre $\int_a^b ZX(t)dB(t)$ non è definito perché $ZX(t)$ non è adattato.

È utile considerare l'integrale stocastico come un processo:

$$\int_0^t X(s)dB(s), \quad 0 \leq t \leq T.$$

- Si dimostra che esiste una versione continua di questo processo e si assume di aver sempre scelto questa versione continua.
- Per costruzione questo processo è anche adattato; essendo continuo e adattato è anche progressivo.

La dimostrazione della continuità non è banale; coinvolge i concetti di attesa condizionale e di martingala e la disuguaglianza di martingala di Doob. I concetti di attesa condizionale e di martingala sono strumenti di grande importanza nel calcolo stocastico e nello studio dei processi; il doverli saltare per mancanza di tempo è certo molto limitativo.

2.3.6 Estensione della classe di integrandi

Si può estendere l'integrale di Itô alla classe di processi $\Lambda^2(T)$: i processi progressivi su $[0, T]$ tali che

$$P \left[\int_0^T |X(t)|^2 dt < +\infty \right] = 1. \quad (37)$$

Ovviamente $M^2 \subset \Lambda^2$. Per ottenere questa estensione si può lavorare con i 'tempi di arresto', come su [Mao], o con limiti in probabilità, come su [Baldi]. Dò un cenno a questo secondo approccio.

- Si dice che una successione di variabili aleatorie Y_n tende in probabilità a una variabile aleatoria Y , $Y_n \xrightarrow{P} Y$, se $\forall \eta > 0$, $\lim_{n \rightarrow \infty} P[|Y_n - Y| > \eta] = 0$.

Questa definizione va bene per variabili aleatorie a valori in spazi metrici mettendo la distanza al posto del modulo della differenza.

1. I processi semplici in $\Lambda^2(T)$ sono del tipo $\sum_j e_j 1_{[t_j, t_{j+1})}(t)$, con e_j variabile aleatoria \mathcal{F}_{t_j} -misurabile.
2. Per ogni $X \in \Lambda^2(T)$ esiste una successione X_n di processi semplici in $\Lambda^2(T)$ tale che $P\text{-}\lim_{n \rightarrow \infty} \int_0^T |X_n(t) - X(t)|^2 dt = 0$.
3. Si definisce $\int_0^T X_n(t) dB(t) = \sum_j e_j^{(n)} \Delta B_j$ e si dimostra che la successione di variabili aleatorie ottenute tende in probabilità ad una variabile aleatoria che dipende da X e non dalla successione approssimante. Tale limite è la definizione di integrale stocastico:

$$\int_0^T X(t) dB(t) = P\text{-}\lim_{n \rightarrow \infty} \int_0^T X_n(t) dB(t). \quad (38)$$

4. Su $M^2(T)$ le due definizioni di integrale stocastico coincidono.
5. Valgono tutte le proprietà dell'integrale stocastico già viste, tranne quelle dove sono coinvolti i valori attesi.
6. Esiste una versione continua e con $\int_0^t X(s) dB(s)$ si indica la versione continua.

2.4 Il differenziale stocastico e la formula di Itô

Introduciamo un'ultima classe di processi, quelli che integriamo rispetto al tempo.

- Indichiamo con $\Lambda^1(T)$ l'insieme dei processi $F(t)$ su $[0, T]$ progressivi, tali che

$$P \left[\int_0^T |F(t)| dt < +\infty \right] = 1. \quad (39)$$

- Si noti che $\Lambda^1(T) \supset \Lambda^2(T) \supset M^2(T)$.
- Se $F \in \Lambda^1(T)$, allora $\int_0^t F(s)ds$ è un processo continuo e adattato, dunque anche progressivo. Le traiettorie, essendo continue, sono anche integrabili; dunque $\int_0^\bullet F(s)ds \in \Lambda^1(T)$.

Definizione 1. Sia X un processo reale tale che $\forall t \in [0, T]$ si abbia

$$X(t) = X_0 + \int_0^t F(u)du + \int_0^t G(u)dB(u) \quad (40)$$

con X_0 \mathcal{F}_0 -misurabile, $F \in \Lambda^1(T)$, $G \in \Lambda^2(T)$; allora si dice che X è un processo di Itô o che *ammette differenziale stocastico* (unidimensionale) e si scrive

$$dX(t) = F(t)dt + G(t)dB(t). \quad (41)$$

Si noti che per noi \mathcal{F}_0 è banale e dunque X_0 è quasi certamente una costante; tuttavia si può generalizzare la costruzione anche a situazioni con \mathcal{F}_0 non banale, ma indipendente dal processo di Wiener.

Vogliamo arrivare ora a introdurre un attrezzo fondamentale nel calcolo stocastico: la formula di Itô; per i differenziali stocastici, questa formula è l'analogo della derivata di funzione di funzione.

Il differenziale del prodotto. Sia

$$dX_i(t) = F_i(t)dt + G_i(t)dB(t), \quad i = 1, 2; \quad (42)$$

vogliamo trovare il differenziale di $X_1(t)X_2(t)$. Consideriamo degli incrementi finiti

$$\Delta X_i(t) := X_i(t + \Delta t) - X_i(t); \quad (43)$$

con questa definizione si ha banalmente

$$\Delta(X_1(t)X_2(t)) = X_1(t)\Delta X_2(t) + X_2(t)\Delta X_1(t) + (\Delta X_1(t))(\Delta X_2(t)). \quad (44)$$

Vogliamo mostrare che per “incrementi infinitesimi” si ha

$$d(X_1(t)X_2(t)) = X_1(t)dX_2(t) + X_2(t)dX_1(t) + (dX_1(t))(dX_2(t)). \quad (45)$$

dove i $dX_i(t)$ vengono presi dalla (42) e la “correzione di Itô” va calcolata con le regole

$$(dB(t))^2 = dt, \quad dt dB(t) = 0, \quad (dt)^2 = 0. \quad (46)$$

Si ricordi che la deviazione standard di $dB(t)$ è \sqrt{dt} . Più esplicitamente

$$d(X_1(t)X_2(t)) = [X_1(t)F_2(t) + F_1(t)X_2(t) + G_1(t)G_2(t)] dt + [X_1(t)G_2(t) + G_1(t)X_2(t)] dB(t). \quad (47)$$

Si noti che il lato destro ha senso; infatti $G_1, G_2 \in \Lambda^2(T) \Rightarrow G_1G_2 \in \Lambda^1(T)$, $F_1 \in \Lambda^1(T)$ e X_2 continuo adattato $\Rightarrow F_1X_2 \in \Lambda^1(T)$, eccetera.

Teorema 3. *Se X_1 e X_2 hanno differenziale stocastico (42), $F_i \in \Lambda^1(T)$, $G_i \in \Lambda^2(T)$, allora X_1X_2 ammette differenziale stocastico, che risulta dato dalla (47).*

Dimostrazione. Se F_1, F_2, G_1, G_2 sono costanti nel tempo la tesi segue dalla proprietà (36), dai due esempi (33), (34) e da $d(t^2) = 2tdt$. Se F_1, F_2, G_1, G_2 sono processi semplici ci si restringe agli intervalli su cui sono costanti e poi si somma. Al caso generale si arriva per approssimazione con integrandi semplici. \square

Vediamo ora la formula di Itô per una funzione di un solo processo; il caso generale lo lasciamo alla prossima sezione dove anche il processo di Wiener sarà multidimensionale. Si noti che la formula (49) corrisponde a uno sviluppo di Taylor arrestato al secondo ordine e all'uso delle regole di moltiplicazione (46).

Teorema 4 (Formula di Itô). *Sia X un processo che ammette differenziale stocastico*

$$dX(t) = F(t)dt + G(t)dB(t) \quad (48)$$

e sia $f(x, t)$, $f : \mathbb{R} \otimes \mathbb{R}_+ \rightarrow \mathbb{R}$, una funzione continua e derivabile con continuità una volta in t e due volte in x . Allora il processo $f(X(t), t)$ ammette il differenziale stocastico

$$df(X(t), t) = \frac{\partial f(X(t), t)}{\partial t} dt + \frac{\partial f(X(t), t)}{\partial x} dX(t) + \frac{1}{2} \frac{\partial^2 f(X(t), t)}{\partial x^2} G(t)^2 dt. \quad (49)$$

Dimostrazione.

Passo 1. Supponiamo $f(x, t) = x^n$; per ricorrenza, usando il Teorema 3 si dimostra

$$d(X(t)^n) = nX(t)^{n-1}dX(t) + \frac{1}{2} n(n-1)X(t)^{n-2}G(t)^2 dt. \quad (50)$$

Per linearità la formula di Itô è vera anche se $f(x, t) = P(x)$ (un polinomio).

Passo 2. Supponiamo ora $f(x, t) = P(x)g(t)$. Si ha

$$dP(X(t)) = P'(X(t))dX(t) + \frac{1}{2}P''(X(t))G(t)^2dt, \quad (51)$$

$$dg(t) = g'dt. \quad (52)$$

Anche in questo caso la formula di Itô segue dal Teorema 3. Per linearità la formula di Itô è vera anche per $f(x, t) = \sum_{r=1}^l P_r(x)g_r(t)$.

Passo 3. Per approssimazioni si dimostra il caso generale. □

2.5 L'integrale e il differenziale stocastici multidimensionali

Gli integrali stocastici multidimensionali sono solo somme di integrali unidimensionali fatti rispetto alle varie componenti del processo di Wiener. Solo la formula di Itô necessita di una piccola generalizzazione. Ricordiamo che abbiamo un processo di Wiener $W(t)$ d -dimensionale come descritto nella Sezione 2.1.2 con la sua filtrazione \mathcal{F}_t descritta nella Sezione 2.3.1.

Lasciando perdere la specificazione del tempo finale T , diremo che un processo $F(t)$ a valori in \mathbb{R}^m appartiene a Λ_m^1 se ogni sua componente F_i appartiene a Λ^1 . Diremo che un processo $G(t)$ a valori in $\mathbb{R}^{m \times d}$ appartiene a $\Lambda_{m,d}^2$ se ogni sua componente G_{ij} appartiene a Λ^2 ; lo stesso per $M_{m,d}^2$. Sono poi definiti gli integrali $\int_0^t F(s)ds$, vettore di componenti $\int_0^t F_i(s)ds$, e $\int_0^t G(s)dW(s)$, vettore di componenti $\sum_{j=1}^d \int_0^t G_{ij}(s)dW_j(s)$.

Diremo che un processo $X(t)$ a valori in \mathbb{R}^m ammette differenziale stocastico

$$dX(t) = F(t)dt + G(t)dW(t), \quad F \in \Lambda_m^1, \quad G \in \Lambda_{m,d}^2, \quad (53)$$

se esiste X_0 , \mathcal{F}_0 -misurabile, per cui si possa scrivere $\forall t$

$$X(t) = X_0 + \int_0^t F(s)ds + \int_0^t G(s)dW(s). \quad (54)$$

Proposizione 5. *Con le notazioni precedenti si ha*

$$d(X_i(t)X_j(t)) = X_i(t)dX_j(t) + X_j(t)dX_i(t) + \sum_{h=1}^d G_{ih}(t)G_{jh}(t)dt. \quad (55)$$

La dimostrazione è analoga a quella del caso unidimensionale, una volta che si sia mostrato che per $h \neq k$ si ha $d(W_h(t)W_k(t)) = W_h(t)dW_k(t) + W_k(t)dW_h(t)$, che a sua volta si dimostra con le solite discretizzazioni e l'indipendenza delle componenti del processo di Wiener. Le regole di calcolo sono ora riassunte da

$$dW_h(t)dW_k(t) = \delta_{hk}dt, \quad dt dW_h(t) = 0, \quad (dt)^2 = 0. \quad (56)$$

Passando a polinomi e poi usando approssimazioni si può dimostrare la formula di Itô multidimensionale.

Teorema 6 (Formula di Itô). *Sia $f : \mathbb{R}^m \times \mathbb{R}_+ \rightarrow \mathbb{R}$ una funzione continua in (x, t) con le sue derivate $f_t, f_{x_i}, f_{x_i x_j}$. Sia X un processo di differenziale stocastico (53). Allora il processo $f(X(t), t)$ ammette differenziale stocastico*

$$\begin{aligned} df(X(t), t) &= f_t(X(t), t)dt + \sum_{i=1}^m f_{x_i}(X(t), t)dX_i(t) \\ &\quad + \frac{1}{2} \sum_{i,j=1}^m f_{x_i x_j}(X(t), t) \sum_{h=1}^d G_{ih}(t)G_{jh}(t)dt. \end{aligned} \quad (57)$$

Per un vettore o una matrice A indichiamo con $|A|^2$ la somma dei moduli quadri di tutte le componenti e applichiamo la formula di Itô a $\left| \int_0^t G(s)dW(s) \right|^2$; otteniamo

$$\begin{aligned} d \left| \int_0^t G(s)dW(s) \right|^2 &= 2 \sum_{i=1}^m \sum_{h,k=1}^d \left(\int_0^t G_{ih}(s)dW_h(s) \right) G_{ik}(t)dW_k(t) \\ &\quad + \sum_{i=1}^m \sum_{h=1}^d |G_{ih}(t)|^2 dt. \end{aligned} \quad (58)$$

Se $G \in M_{m,d}^2$, il primo differenziale (integrale) sul lato destro ha media nulla e così abbiamo la generalizzazione dell'isometria di Itô

$$\mathbb{E} \left[\left| \int_0^t G(s)dW(s) \right|^2 \right] = \int_0^t \mathbb{E} [|G(s)|^2] ds. \quad (59)$$

3 Equazioni differenziali stocastiche

3.1 Una classe di EDS

Finalmente abbiamo dato senso all'EDS da cui siamo partiti

$$\begin{cases} dX(t) = b(X(t), t)dt + \sigma(X(t), t)dW(t) \\ X(0) = x_0 \end{cases} \quad (60)$$

dove $x_0 \in \mathbb{R}^m$ e b e σ sono funzioni deterministiche e misurabili

$$\begin{aligned} b &: \mathbb{R}^m \times [0, T] \rightarrow \mathbb{R}^m, \\ \sigma &: \mathbb{R}^m \times [0, T] \rightarrow \mathbb{R}^{m \times d}. \end{aligned}$$

Tale equazione va intesa come una condizione sul differenziale stocastico di X e dunque è un'equazione integrale

$$X(t) = x_0 + \int_0^t b(X(s), s) ds + \int_0^t \sigma(X(s), s) dW(s) \quad (61)$$

coinvolgente integrali ordinari e stocastici. Una soluzione di questa equazione è un processo continuo, definito nello spazio di probabilità dove vive il processo di Wiener d -dimensionale; tale processo deve essere adattato alla filtrazione descritta nella Sezione 2.3.1, deve essere tale per cui gli integrali coinvolti abbiano senso e deve rendere vera la (61). In realtà nella teoria più avanzata delle EDS si possono dare concetti di soluzione un po' più articolati (soluzioni deboli e forti), ma noi tralasciamo questa parte.

Il termine in dt si chiama termine di drift e b coefficiente di drift, il termine in $dW(t)$ si chiama termine di diffusione e σ coefficiente di diffusione.

Questa sezione è intitolata “una classe” di EDS non solo perché si possono considerare rumori più generali del processo di Wiener, ma soprattutto perché si potrebbero considerare dei coefficienti b e σ casuali, o dipendenti da X per tutti i tempi antecedenti $t...$

3.2 Esistenza e unicità delle soluzioni

Ipotesi (I). Diremo che b e σ soddisfano le ipotesi (I) se sono misurabili in (x, t) e se

- (i) sono a crescita sublineare, cioè esiste una costante $M > 0$ tale che per ogni $x \in \mathbb{R}^m$, $t \in [0, T]$

$$|b(x, t)| + |\sigma(x, t)| \leq M(1 + |x|) \quad (62)$$

- (ii) sono localmente lipschitziani in x , cioè per ogni $N > 0$ esiste $L_N > 0$ tale che se $x, y \in \mathbb{R}^m$, $|x| \leq N$, $|y| \leq N$, $t \in [0, T]$, si abbia

$$|b(x, t) - b(y, t)| + |\sigma(x, t) - \sigma(y, t)| \leq L_N |x - y|. \quad (63)$$

Teorema 7. *Nelle ipotesi (I) esiste un processo X che soddisfa (61) in $[0, T]$. Inoltre $X_i \in M^2(T)$, $i = 1, \dots, m$, ed esiste una costante $C(M, T)$ tale che*

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |X(t)|^2 \right] \leq C(M, T) (1 + |x_0|^2). \quad (64)$$

Infine vale l'unicità della soluzione nel senso che se X' è un'altra soluzione allora

$$P [X(t) = X'(t), \forall t \in [0, T]] = 1. \quad (65)$$

La dimostrazione usa tecniche del tipo di quelle impiegate per le ODE come il lemma di Gronwall e il metodo iterativo di Picard; intervengono però anche ingredienti di tipo più probabilistico come le disuguaglianze di Doob per martingale, il lemma di Borel-Cantelli,...

Abbiamo già osservato che le EDS non possono essere intese traiettoria per traiettoria per via dell'integrale stocastico che coinvolge la probabilità e perciò tutti i valori del processo. Questo però non impedisce che tali equazioni abbiano un carattere locale.

Tempo d'arresto. Un tempo d'arresto è una variabile aleatoria τ non negativa estesa (può prendere il valore $+\infty$) tale che $\{\omega \in \omega : \tau(\omega) \leq t\} \in \mathcal{F}_t, \forall t \geq 0$.

Tempo d'uscita. Sia $X(t)$ un processo continuo m -dimensionale e D un aperto in \mathbb{R}^m . Poniamo

$$\tau(\omega) = \inf\{t \geq 0 : X(t; \omega) \notin D\}. \quad (66)$$

Si dimostra che τ è un tempo d'arresto che chiamiamo tempo d'uscita di X da D .

Teorema 8 (di localizzazione). *Supponiamo che $b^{(i)}, \sigma^{(i)}, i = 1, 2$, siano misurabili su $\mathbb{R}^m \times [0, T]$ e siano $X^{(i)}, i = 1, 2$, soluzioni delle EDS*

$$\begin{cases} dX^{(i)}(t) = b^{(i)}(X(t), t)dt + \sigma^{(i)}(X(t), t)dW(t) \\ X^{(i)}(0) = x_0 \end{cases} \quad (67)$$

Sia D un aperto in \mathbb{R}^m tale che su $D \times [0, T]$ sia $b^{(1)} = b^{(2)}, \sigma^{(1)} = \sigma^{(2)}$ e si abbia, per ogni $x, y \in D, 0 \leq t \leq T$,

$$\left| b^{(i)}(x, t) - b^{(i)}(y, t) \right| + \left| \sigma^{(i)}(x, t) - \sigma^{(i)}(y, t) \right| \leq L_N |x - y|. \quad (68)$$

Allora, se τ_i indica il tempo d'uscita di $X^{(i)}$ da D ,

$$\tau_1 = \tau_2 \quad \text{q.c.} \quad (69)$$

$$P \left[X^{(1)}(t) = X^{(2)}(t), \forall t \in [0, \tau_1] \right] = 1. \quad (70)$$

3.3 Esempi

3.3.1 Equazione di Langevin

Una particella in un mezzo viscoso, omogeneo, isotropo:

$$\begin{cases} d\vec{x} = \vec{v}dt \\ d\vec{v} = -\gamma\vec{v}dt - \vec{\nabla}\Phi(\vec{x})dt + \sqrt{\frac{2\gamma kT}{m}} d\vec{W}(t) \end{cases} \quad (71)$$

A seconda della scelta del potenziale Φ , questa equazione può rappresentare il moto browniano, un modello di diffusione per reazioni chimiche, migrazione di atomi in cristalli...

3.3.2 Un mercato finanziario

$$\frac{dS_i(t)}{S_i(t)} = b_i(S(t), t)dt + \sum_{i,j}^m \sigma_{ij}(S(t), t)dW_j(t), \quad i = 1, \dots, m; \quad (72)$$

$S(t) = (S_1(t), \dots, S_m(t))$ e gli $S_i(t)$ rappresentano prezzi di titoli. Se i valori iniziali sono positivi, le soluzioni hanno componenti positive.

3.3.3 Modello di diffusione di fase per un laser

$$\begin{cases} dX(t) = X(t) \left[(i\nu - \frac{1}{2} b^2) dt + ibdW(t) \right] \\ X(0) = \lambda \in \mathbb{C} \end{cases} \quad (73)$$

$\nu > 0, b > 0$. Soluzione (provare a verificare usando la formula di Itô):

$$X(t) = \lambda \exp \{ i\nu t + ibW(t) \}. \quad (74)$$

$$\overline{X(s)} X(t) = |\lambda|^2 \exp \{ i\nu(t-s) + ib[W(t) - W(s)] \} \quad (75)$$

$W(t) - W(s)$ è normale di media zero e varianza $|t - s|$; dalla funzione caratteristica delle leggi normali:

$$\mathbb{E} \left[\overline{X(s)} X(t) \right] = |\lambda|^2 \exp \left\{ i\nu(t-s) - \frac{1}{2} b^2 |t-s| \right\} \quad (76)$$

dipende solo dalla differenza dei tempi.

$|X(t)|^2 = |\lambda|^2$: intensità costante

$$\int_{-\infty}^{+\infty} \mathbb{E} \left[\overline{X(0)} X(\tau) \right] e^{-i\kappa\tau} d\tau = \frac{|\lambda|^2 b^2}{(\kappa - \nu)^2 + b^2/4} : \quad (77)$$

lo spettro è lorentziano.

4 Le equazioni di Kolmogorov e di Fokker-Planck

4.1 Il generatore delle diffusioni

Passiamo ora al problema delle PDE associate alle EDS; si possono ottenere rappresentazioni stocastiche per soluzioni di varie PDE, noi vediamo solo le equazioni di Kolmogorov.

- C_b^2 : funzioni su \mathbb{R}^m continue e limitate insieme alle loro derivate prime e seconde.
- C_K^2 : funzioni su \mathbb{R}^m continue a supporto compatto insieme alle loro derivate prime e seconde.

- Assumiamo la validità delle ipotesi **(I)** e poniamo per $f \in C_b^2$

$$(L_t f)(x) := \sum_{i=1}^m b_i(x, t) \frac{\partial f(x)}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^m a_{ij}(x, t) \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \quad (78)$$

$$a_{ij}(x, t) := \sum_{h=1}^d \sigma_{ih}(x, t) \sigma_{jh}(x, t). \quad (79)$$

- Per $x \in \mathbb{R}^m$, $0 \leq s \leq t \leq T$ indichiamo con $X^{x,s}(t)$ la soluzione della nostra EDS con condizione iniziale x al tempo s :

$$\begin{cases} dX^{x,s}(t) = b(X^{x,s}(t), t)dt + \sigma(X^{x,s}(t), t)dW(t) \\ X^{x,s}(s) = x \end{cases} \quad (80)$$

Con $f \in C_b^2$ possiamo applicare la formula di Itô a $f(X^{x,s}(t))$; raggruppando i termini si ottiene immediatamente

$$\begin{aligned} f(X^{x,s}(t)) &= f(x) + \int_s^t (L_u f)(X^{x,s}(u)) du \\ &\quad + \sum_{i=1}^m \sum_{j=1}^d \int_s^t \frac{\partial f(X^{x,s}(u))}{\partial x_i} \sigma_{ij}(X^{x,s}(u), u) dW_j(u). \end{aligned} \quad (81)$$

L'integrando nell'integrale stocastico è di classe $M^2(T)$; infatti $\frac{\partial f}{\partial x_i}$ è limitata, σ_{ij} è a crescita sublineare e $X^{x,s} \in M^2(T)$. Dunque l'integrale stocastico ha valore atteso nullo. In modo analogo si vede che l'integrale sul tempo ammette media e per Fubini si possono scambiare integrale temporale e valore atteso. Posto

$$(T_{s,t} f)(x) = \mathbb{E} [f(X^{x,s}(t))], \quad (82)$$

si ha allora

$$(T_{s,t} f)(x) = f(x) + \int_s^t (T_{s,u} \circ L_u f)(x) du. \quad (83)$$

L'operatore L_t è chiamato il generatore infinitesimale della diffusione associata alla nostra EDS. Si noti che questa equazione vale per ogni $f \in C_b^2$ nelle ipotesi **(I)**; non è necessaria nessuna ipotesi aggiuntiva su a e b . Dunque, sappiamo che per costruzione (83) ammette almeno una soluzione T .

4.2 Probabilità di transizione e processi di Markov

Per proseguire abbiamo bisogno di qualche nozione sui processi di Markov, che diamo in modo molto schematico.

Sia D un aperto di \mathbb{R}^m con i suoi borelliani \mathcal{B}_D ; una probabilità di transizione su D è una funzione $p(A, t|x, s)$, dove $0 \leq s \leq t$, $x \in D$, $A \in \mathcal{B}_D$, tale che

- (i) per s, t, A fissati, $x \mapsto p(A, t|x, s)$ è \mathcal{B}_D -misurabile;
- (ii) per s, t, x fissati, $p(\cdot, t|x, s)$ è una misura di probabilità;
- (iii) p soddisfa l'equazione di Chapman-Kolmogorov: $\forall s < u < t$

$$p(A, t|x, s) = \int_D p(A, t|y, u)p(dy, u|x, s). \quad (84)$$

Data una probabilità di transizione e $x_0 \in D$ esiste un processo X con leggi finito-dimensionali

$$P[X(t_1) \in F_1, \dots, X(t_n) \in F_n] = \int_{F_1 \times \dots \times F_{n-1}} p(F_n, t_n|x_{n-1}, t_{n-1}) \\ \times p(dx_{n-1}, t_{n-1}|x_{n-2}, t_{n-2}) \cdots p(dx_1, t_1|x_0, 0). \quad (85)$$

Chiamiamo markoviano un processo con questa struttura.

L'analisi della nostra EDS porta a stabilire che la sua soluzione è un processo di Markov su \mathbb{R}^m ; la caratteristica essenziale è che i coefficienti di drift e di diffusione dipendono dalla soluzione all'ultimo istante, il presente, e non dalla soluzione calcolata nei tempi precedenti, il passato. Con qualche precisazione si riesce ad ottenere che per $t_0 < s < t$ si ha

$$X^{x_0, t_0}(t; \omega) = X^{X^{x_0, t_0}(s; \omega), s}(t; \omega); \quad (86)$$

per ottenere questo risultato si devono dimostrare delle proprietà di regolarità della soluzione nei dati iniziali e la "proprietà di Markov forte". Le probabilità di transizione sono poi date da

$$p(A, t|x, s) = P[X^{x, s}(t) \in A] \equiv \mathbb{E} [1_A(X^{x, s}(t))]. \quad (87)$$

Inoltre si ha

$$(T_{s, t}f)(x) = \int_{\mathbb{R}^m} f(y)p(dy, t|x, s) \quad (88)$$

e l'equazione di Chapman-Kolmogorov, tradotta in termini degli operatori $T_{s, t}$, dà

$$T_{r, s} \circ T_{s, t} = T_{r, t}, \quad r \leq s \leq t. \quad (89)$$

4.3 L'equazione di Kolmogorov all'indietro

Poniamo

$$u(x, t) := \mathbb{E} [f(X^{x, t}(T))] \equiv (T_{t, T}f)(x). \quad (90)$$

Si ha allora

$$u(x, t-h) = (T_{t-h, T}f)(x) = (T_{t-h, t}T_{t, T}f)(x) \\ = u(x, t) + \int_{t-h}^t dr (T_{t-h, r}L_r u(\cdot, t))(x),$$

$$\begin{aligned}\lim_{h \downarrow 0} \frac{1}{h} [u(x, t) - u(x, t - h)] &= -\lim_{h \downarrow 0} \frac{1}{h} \int_{t-h}^t dr (T_{t-h, r} L_r u(\cdot, t))(x) \\ &= -(L_t u(\cdot, t))(x) =: -L_t u(x, t).\end{aligned}$$

Per rendere rigorosi questi passaggi bisogna mostrare che $u(t)$ sta nel dominio di L_t . Per questo bisogna utilizzare risultati di regolarità delle soluzioni delle EDS rispetto ai dati iniziali e la “proprietà di Markov forte”; è necessaria una qualche ulteriore ipotesi sui coefficienti dell’EDS.

Mettendo insieme quanto abbiamo ottenuto, abbiamo l’equazione di Kolmogorov all’indietro:

$$\begin{cases} \frac{\partial}{\partial t} u(x, t) + L_t u(x, t) = 0 \\ u(x, T) = f(x) \end{cases} \quad (91)$$

4.4 L’equazione di Fokker-Planck o di Kolmogorov in avanti

Dalle equazioni (83) e (88) ottengo

$$\frac{\partial}{\partial t} \int_{\mathbb{R}^m} f(y) p(dy, t|x, s) = \int_{\mathbb{R}^m} (L_t f)(y) p(dy, t|x, s), \quad (92)$$

che è la forma debole dell’equazione di Fokker-Planck per la densità di probabilità $p(dy, t|x, s) = p(y, t|x, s)dy$:

$$\begin{cases} \frac{\partial}{\partial t} p(x, t|y, s) = L_t^* p(x, t|y, s), \\ p(x, s|y, s) = \delta(x - y) \end{cases} \quad (93)$$

$$L_t^* g(x, t) := -\sum_{i=1}^m \frac{\partial}{\partial x_i} (b_i(x, t)g(x, t)) + \frac{1}{2} \sum_{i,j=1}^m \frac{\partial^2}{\partial x_i \partial x_j} (a_{ij}(x, t)g(x, t)). \quad (94)$$

Per dimostrare l’esistenza delle densità basta dimostrare che l’equazione di Fokker-Planck (93) ha una soluzione fondamentale regolare.