Metodi Stocastici per la Finanza

Tiziano Vargiolu
vargiolu@math.unipd.it[1]

[1]Università degli Studi di Padova

Anno Accademico 2014-2015
Lezione 5

Indice

Due errori

Quando dobbiamo simulare con metodi Monte Carlo una EDS che non ha soluzione esatta, e quindi usiamo una discretizzazione (es. schema di Eulero), ci esponiamo a due errori.
Difatti, volendo approssimare X soluzione di una EDS:

- lo approssimiamo con Y^n, con n fissato, ottenuto tramite uno schema di Eulero,
- che poi simuliamo attraverso M simulazioni Monte Carlo, ottenendo i processi stocastici Y^n_m, $m = 1, \dots, M$, tutti con legge uguale a Y^n e indipendenti fra loro.

Avremo quindi due errori:

$$\left| \mathbb{E}[f(X(\cdot))] - \frac{1}{M} \sum_{m=1}^{M} f(Y^n_M(\cdot)) \right| \leq$$

$$|\mathbb{E}[f(X)] - \mathbb{E}[f(Y^n)]| \quad + \quad \left| \mathbb{E}[f(Y^n)] - \frac{1}{M} \sum_{m=1}^{M} f(Y^n_M) \right|$$

(errore di discretizzazione) (errore Monte Carlo)

Errore di discretizzazione

Per quanto riguarda il primo tipo di errore, un risultato teorico afferma che, per f sufficientemente regolare, l'errore di discretizzazione é del tipo

$$\mathbb{E}[f(X_T)] - \mathbb{E}[f(Y_T^n)] = \frac{C_1}{n} + O(n^{-2})$$

dove C_1 é una costante che dipende da b, σ, T, f ma non da n. Questo permette di ottenere stime di ordine superiore in modo molto semplice: infatti basta considerare

$$\begin{aligned}
&\mathbb{E}[f(X_T)] - 2\mathbb{E}[f(Y_T^{n/2})] + \mathbb{E}[f(Y_T^n)] = \\
&= \quad 2\mathbb{E}[f(X_T)] - 2\mathbb{E}[f(Y_T^{n/2})] - \mathbb{E}[f(X_T)] + \mathbb{E}[f(Y_T^n)] = \\
&= \quad \frac{C_1}{n} - \frac{C_1}{n} + O(n^{-2}) = O(n^{-2})
\end{aligned}$$

e quindi la quantità $2\mathbb{E}[f(Y_T^{n/2})] - \mathbb{E}[f(Y_T^n)]$ approssima al secondo ordine la quantità cercata $\mathbb{E}[f(X_T)]$.

Errore Monte Carlo

L'errore Monte Carlo, invece, é come di consueto proporzionale a $\frac{1}{\sqrt{M}}$, dove M é il numero di simulazioni, quindi converge a 0 in modo molto piú lento.

$$\mathrm{Var}\left[\mathbb{E}[f(Y^n)] - \frac{1}{M}\sum_{m=1}^{M} f(Y^n_M)\right] =$$

$$= \frac{1}{M^2}\sum_{m=1}^{M}\mathrm{Var}\left[\mathbb{E}[f(Y^n)] - f(Y^n_M)\right] = \frac{1}{M}\mathrm{Var}[f(Y^n)]$$

e quindi la deviazione standard media dell'errore MC decresce come $\frac{1}{\sqrt{M}}$, ed è proporzionale alla deviazione standard di $f(Y^n)$.
Soluzione: ridurre la varianza di $f(Y^n)$!

Metodi di riduzione della varianza

Ci sono molti metodi di ridurre la varianza nel metodo Monte Carlo. Ne esamineremo, con suggerimenti per l'implementazione, principalmente tre:

1. metodo delle variabili antitetiche;
2. metodo della variabile di controllo;
3. campionamento di importanza (importance sampling).

Metodo delle variabili antitetiche

Supponiamo di dover calcolare $\mathbb{E}[f(X)]$, con X simmetrica (tale cioè che X e $-X$ abbiano la stessa legge). Allora possiamo scrivere

$$\mathbb{E}[f(X)] = \mathbb{E}\left[\frac{f(X)+f(-X)}{2}\right]$$

Il vantaggio di questa rappresentazione è che

$$\begin{aligned}\mathrm{Var}\left[\frac{f(X)+f(-X)}{2}\right] &= \frac{1}{2}\mathrm{Var}[f(X)] + \frac{1}{2}\mathrm{Cov}(f(X), f(-X)) < \\ &< \mathrm{Var}[f(X)]\end{aligned}$$

per Cauchy-Schwarz: si avrebbe l'uguaglianza solo se $\mathrm{Cov}(f(X), f(-X)) = \sqrt{\mathrm{Var}[f(X)]\mathrm{Var}[f(-X)]} = \mathrm{Var}[f(X)]$, cosa molto rara: generalmente difatti si ha addirittura che $\mathrm{Cov}(f(X), f(-X)) < 0$, riuscendo quindi perlomeno a dimezzare la varianza.

Stima della covarianza

Per averne un'idea, supponiamo che $f \in C^1(\mathbb{R})$ e approssimiamo $f(x) \simeq f(0) + xf'(0)$ (lo sviluppo intorno a 0 è il più adatto poichè X è simmetrica, e quindi tutti i suoi momenti di ordine dispari sono nulli): allora

$$
\begin{aligned}
&\mathrm{Cov}(f(X), f(-X)) \simeq \\
&\quad \simeq \quad \mathbb{E}[(f(0) + Xf'(0) - \mathbb{E}[f(0) + Xf'(0)]) \times \\
&\qquad\qquad \times (f(0) - Xf'(0) - \mathbb{E}[f(0) - Xf'(0)])] = \\
&\quad = \quad \mathbb{E}[(f'(0)(X - \mathbb{E}[X])(-f'(0)(X - \mathbb{E}[X]))] = -f'(0)^2 \mathrm{Var}[X]
\end{aligned}
$$

Questo termine è quindi di ordine paragonabile a $\mathrm{Var}[X]$.

Guadagno nella varianza

Per vedere quindi quanto si guadagna con questo metodo, supponiamo che $f \in C^2(\mathbb{R})$ ed approssimiamo $f(x) \simeq f(0) + xf'(0) + \frac{1}{2}x^2 f''(0)$: allora

$$\begin{aligned}
\mathrm{Var}\left[\frac{f(X)+f(-X)}{2}\right] &\simeq \mathrm{Var}\left[\frac{Xf'(0)+\frac{1}{2}X^2f''(0)-Xf'(0)+\frac{1}{2}X^2f''(}{2}\right. \\
&= \frac{1}{4}\mathrm{Var}\left[X^2f''(0)\right] = \frac{1}{4}f''(0)^2\mathrm{Var}[X^2]
\end{aligned}$$

Facendo la stessa approssimazione per $\mathrm{Var}[f(X)]$, si ottiene:

$$\begin{aligned}
\mathrm{Var}\left[f(X)\right] &\simeq \mathrm{Var}\left[f(0)+Xf'(0)+\frac{1}{2}X^2f''(0)\right] = \\
&= \mathrm{Var}[Xf'(0)] + \frac{1}{4}\mathrm{Var}\left[X^2f''(0)\right] + \mathrm{Cov}(Xf'(0), X^2f''(0)) \\
&= f'(0)^2\mathrm{Var}[X] + \frac{1}{4}f''(0)^2\mathrm{Var}[X^2]
\end{aligned}$$

poichè $\mathrm{Cov}(X, X^2) = \mathbb{E}[X^3] - \mathbb{E}[X]\mathbb{E}[X^2] = 0$.

Esempi di variabili aleatorie simmetriche

Esempio. In una simulazione semplice di una variabile lognormale, dovremo simulare due variabili:

$$\begin{aligned} S^1 &= s_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma X}, \\ S^2 &= s_0 e^{(r-\frac{1}{2}\sigma^2)T-\sigma X}, \end{aligned}$$

con $X \sim N(0, T)$.

Esempio. In una implementazione del metodo di Eulero per un moto browniano geometrico, il processo potrebbe essere generato, usando lo stesso processo guida $(w_n)_n$, da due variabili:

$$\begin{aligned} S^1_{n+1} &= S^1_n(1 + r\delta t + \sigma w_{n+1}), \\ S^2_{n+1} &= S^2_n(1 + r\delta t - \sigma w_{n+1}). \end{aligned}$$

In questo caso, le variabili aleatorie simmetriche in questione sono le $(w_n)_n$.

Esempio: calcolo di una call

Supponiamo di dover calcolare il prezzo di una call *in the money*, cioè tale che $s_0 > K$. Se simuliamo la variabile aleatoria $X \sim N(0, \sigma^2 T)$ e quindi poniamo $f(x) := (s_0 e^{(r-\frac{1}{2}\sigma^2)T+x} - K)^+$, allora ovviamente $f \notin C^1(\mathbb{R})$, e quindi il calcolo sopra non è applicabile rigorosamente. Un calcolo formale dà comunque, per $x \simeq 0$,

$$f^{(n)}(x) = s_0 e^{(r-\frac{1}{2}\sigma^2)T+x} \simeq f(x) + K$$

per derivate di ogni ordine $n \geq 1$. Abbiamo quindi

$$\begin{aligned} f'(0)\mathrm{Var}[X] &\simeq s_0^2 e^{2(r-\frac{1}{2}\sigma^2)T} \sigma^2 T, \\ \frac{1}{4} f''(0)^2 \mathrm{Var}[X^2] &\simeq \frac{1}{4} s_0^2 e^{2(r-\frac{1}{2}\sigma^2)T} 2\sigma^4 T^2 = \frac{1}{2} s_0^2 e^{2(r-\frac{1}{2}\sigma^2)T} \sigma^4 T^2, \end{aligned}$$

Esempio: guadagno sulla varianza di una call

Confrontiamo le varianze con e senza l'applicazione del metodo delle variabili antitetiche:

$$
\begin{aligned}
\mathrm{Var}\left[\frac{f(X)+f(-X)}{2}\right] &\simeq \frac{1}{2}e^{2(r-\frac{1}{2}\sigma^2)T}\sigma^4 T^2, \\
\mathrm{Var}[f(X)] &\simeq s_0^2 e^{2(r-\frac{1}{2}\sigma^2)T}\sigma^2 T + \frac{1}{2}s_0^2 e^{2(r-\frac{1}{2}\sigma^2)T}\sigma^4 T^2 = \\
&= s_0^2 e^{2(r-\frac{1}{2}\sigma^2)T}\sigma^2 T\left(1+\frac{1}{2}\sigma^2 T\right)
\end{aligned}
$$

Abbiamo quindi fattorizzato la varianza in un termine proporzionale a $\frac{1}{2}\sigma^2 T$, presente in ogni caso, che viene aumentato di 1 se non si usano le variabili antitetiche. Questo significa che, se σ è bassa e/o T è basso, si ha un notevole guadagno nella varianza.

Implementazione

Quando implementiamo questo metodo, per avere una stima della varianza dobbiamo calcolare

$$\mathrm{Var}\left[\frac{1}{2N}\sum_{i=1}^{N}(f(X_i)+f(-X_i))\right]=\frac{1}{4N^2}\sum_{i=1}^{N}\mathrm{Var}[f(X_i)+f(-X_i)]=$$

$$=\frac{1}{4N^2}\sum_{i=1}^{N}(\mathbb{E}[(f(X_i)+f(-X_i))^2]-\mathbb{E}[f(X_i)+f(-X_i)]^2)$$

Supponendo di avere già calcolato

$$\mathbb{E}[f(X_i)]=\mathbb{E}[f(-X_i)]\simeq E:=\sum_{i=1}^{N}\frac{f(X_i)+f(-X_i)}{2}$$

abbiamo

$$\mathrm{Var}\left[\frac{1}{2N}\sum_{i=1}^{N}(f(X_i)+f(-X_i))\right]\simeq\frac{1}{4N^2}\sum_{i=1}^{N}(f(X_i)+f(-X_i))^2-\frac{1}{N}E^2$$

Metodo della variabile di controllo: idea

Supponiamo di dover calcolare $\mathbb{E}[f(X)]$, e che sia possibile decomporre $f = g + h$, con $\mathbb{E}[g(X)]$ calcolabile facilmente e

$$\mathrm{Var}[h(X)] < \mathrm{Var}[f(X)]$$

Se si riesce a trovare una tale g, allora conviene simulare $h(X) = f(X) - g(X)$ e calcolare

$$\mathbb{E}[f(X)] = \mathbb{E}[g(X)] + \mathbb{E}[h(X)]$$

Metodo della variabile di controllo: più flessibilità

Per avere la certezza di ottenere una variabile di controllo che abbassi sempre la varianza, un modo migliore per decomporre f è il seguente: fissata una candidata variabile di controllo $g(X)$ (con l'ovvia proprietà che $\mathbb{E}[g(X)]$ sia facilmente calcolabile), definiamo

$$h(X;\lambda) := f(X) + \lambda(g(X) - \mathbb{E}[g(X)])$$

Se invece di simulare $f(X)$ simuliamo $h(X)$, otteniamo

$$\mathbb{E}[h(X;\lambda)] = \mathbb{E}[f(X)] + \lambda\mathbb{E}[g(X) - \mathbb{E}[g(X)]] = \mathbb{E}[f(X)]$$

e quindi otteniamo quello che vogliamo indipendentemente da λ. La varianza invece dipende da λ:

$$\mathrm{Var}[h(X;\lambda)] = \mathrm{Var}[f(X)] + 2\lambda\mathrm{Cov}(f(X), g(X)) + \lambda^2\mathrm{Var}[g(X)]$$

Metodo della variabile di controllo: varianza minima

Poichè lo scopo è ottenere la minima varianza nello stimatore Monte Carlo, minimizzando

$$\mathrm{Var}[h(X;\lambda)] = \mathrm{Var}[f(X)] + 2\lambda \mathrm{Cov}(f(X), g(X)) + \lambda^2 \mathrm{Var}[g(X)]$$

si trova che il punto di minimo si raggiunge per

$$\lambda^* := \frac{\mathrm{Cov}(f(X), g(X))}{\mathrm{Var}[g(X)]}$$

e quindi

$$\mathrm{Var}[h(X;\lambda^*)] = \mathrm{Var}[f(X)] - \frac{\mathrm{Cov}(f(X), g(X))^2}{\mathrm{Var}[g(X)]} \leq \mathrm{Var}[f(X)]$$

Nota: se $\mathrm{Cov}(f(X), g(X)) = 0$, non si ha nessun guadagno nella varianza!

Esempio: prezzo di una call

Supponiamo di dover calcolare il prezzo di una call; allora, siccome $(S_T - K)^+ = S_T - K + (K - S_T)^+$, si ha

$$\mathbb{E}[e^{-rT}(S_T - K)^+] = s_0 - e^{-rT}K + \mathbb{E}[e^{-rT}(K - S_T)^+]$$

Quindi, se $\mathrm{Var}[(K - S_T)^+] < \mathrm{Var}[(S_T - K)^+]$, conviene simulare la put. Chiaramente se questo non è vero conviene simulare la call. Per avere un'idea di quale delle due varianze è più grande, calcoliamo

$$\begin{aligned}
\mathrm{Var}[e^{-rT}(K - S_T)^+] &\leq e^{-2rT}K^2, \\
\mathrm{Var}[e^{-rT}(S_T - K)^+] &\leq \mathbb{E}[(e^{-rT}(S_T - K)^+)^2] \leq e^{-2rT}\mathbb{E}[S_T^2] = \\
&= e^{-2rT}s_0^2\mathbb{E}[e^{2(r-\frac{1}{2}\sigma^2)T + 2\sigma W_T}] = e^{-2rT}s_0^2 e^{2rT} e \\
&= s_0^2 e^{-\sigma^2 T + \frac{1}{2}\cdot 4\sigma^2 T} = s_0^2 e^{\sigma^2 T}
\end{aligned}$$

Quindi, se $e^{-2rT}K^2 << s_0^2 e^{\sigma^2 T}$ conviene simulare la put, se $e^{-2rT}K^2 >> s_0^2 e^{\sigma^2 T}$ conviene simulare la call, e se nessuna delle due relazioni è evidente, è indifferente simulare la call o la put.

Esempio: prezzo di una call II

Volendo usare come variabile di controllo $g(S_T) := S_T$, ci chiediamo se $\lambda = -1$ sia quello ottimale. Allora si ha

$$\mathrm{Var}[g(S_T)] = \mathbb{E}[S_T^2] - \mathbb{E}[S_T]^2 = s_0^2 e^{2rT}(e^{\sigma^2 T} - 1)$$

e la covarianza tra $f(S_T)$ e $g(S_T)$, se $s_0 >> K$, può essere stimata come

$$\begin{aligned} \mathrm{Cov}(f(S_T), g(S_T)) &= \mathbb{E}[(S_T^2 - KS_T)^+] - \mathbb{E}[(S_T - K)^+]\mathbb{E}[S_T] \simeq \\ &\simeq \mathbb{E}[S_T^2 - KS_T] - \mathbb{E}[S_T - K]\mathbb{E}[S_T] = \mathbb{E}[S_T^2] - \mathbb{E}[S_T]\mathbb{E}[S_T] = \\ &= \mathrm{Var}[S_T] = \mathrm{Var}[g(S_T)] \end{aligned}$$

e quindi $\lambda^* = -1$.

Campionamento d'importanza (*importance sampling*)

Questa tecnica si basa su un cambio della probabilità rispetto a cui si simula. Consideriamo $g(X) > 0$ $\mathbb{P}$-q.c., con g funzione deterministica. Allora possiamo scrivere

$$\mathbb{E}[f(X)] = \mathbb{E}\left[f(X)\frac{g(X)}{g(X)}\right] = \mathbb{E}_g\left[\frac{f(X)}{g(X)}\right] \cdot \mathbb{E}[g(X)]$$

dove $\mathbb{E}_g$ è la speranza rispetto alla probabilità $\mathbb{P}_g := \frac{g(X)}{\mathbb{E}[g(X)]} \cdot \mathbb{P}$. Chiaramente, affinchè tutto questo sia efficiente, $\mathbb{E}[g(X)]$ deve essere calcolabile in modo facile. Il problema si riduce quindi a simulare X sotto la nuova probabilità $\mathbb{P}_g$.

Minimizzare la varianza rispetto a g

Cerchiamo ora una funzione g che minimizzi la varianza della nuova funzione da simulare sotto la probabilità $\mathbb{P}_g$: dobbiamo quindi minimizzare

$$\begin{aligned}\mathrm{Var}_g\left[\frac{f(X)}{g(X)}\right] &= \mathbb{E}_g\left[\frac{f^2(X)}{g^2(X)}\right] - \mathbb{E}_g\left[\frac{f(X)}{g(X)}\right]^2 = \\ &= \mathbb{E}\left[\frac{f^2(X)}{g(X)}\right]/\mathbb{E}[g(X)] - \frac{\mathbb{E}[f(X)]^2}{\mathbb{E}[g(X)]^2}\end{aligned}$$

È tuttavia evidente che un tale problema, esteso a tutte le possibili funzioni g, è mal posto: difatti, se si prende $g \equiv f$, si ha che $\mathrm{Var}_g\left[\frac{f(X)}{g(X)}\right] = \mathrm{Var}_g[1] = 0$, e quindi il minimo viene raggiunto proprio per $g = f$, cosa che però non ci dà alcun vantaggio dal punto di vista computazionale poichè si ha

$$\mathbb{E}[f(X)] = \mathbb{E}_g\left[\frac{f(X)}{g(X)}\right] \cdot \mathbb{E}[g(X)] = 1 \cdot \mathbb{E}[g(X)] = \mathbb{E}[f(X)]$$

e quindi siamo ricondotti al problema iniziale.

Minimizzare la varianza rispetto a g: un approccio

L'uso pratico del campionamento d'importanza è quindi questo: si cerca una g di una certa classe $g(x) = g(x;\theta)$, e si minimizza $\mathrm{Var}_g\left[\frac{f(X)}{g(X;\theta)}\right]$ rispetto a θ.

Esempio. Se S evolve secondo un moto browniano geometrico, di solito si sceglie $g(S;\alpha) := S_T^\alpha$, con $\alpha > 0$. Allora

$$
\begin{aligned}
\mathbb{E}[g(S;\alpha)] &= \mathbb{E}[S_T^\alpha] = \mathbb{E}[s_0^\alpha e^{\alpha((r-\frac{1}{2}\sigma^2)T+\sigma W_T)}] = \\
&= s_0^\alpha e^{\alpha(r-\frac{1}{2}\sigma^2)T}\mathbb{E}[e^{\alpha\sigma W_T)}] = \\
&= s_0^\alpha e^{\alpha(r-\frac{1}{2}\sigma^2)T+\frac{\alpha^2}{2}\sigma^2 T} = s_0^\alpha e^{\alpha r T-\frac{1}{2}\alpha(\alpha-1)\sigma^2 T}
\end{aligned}
$$

e, a seconda del contingent claim che si vuole prezzare, si minimizza la varianza rispetto ad α.