Zastosowanie metod MCMC do modeli wyceny

Grzegorz Mika

Wydział Matematyki Stosowanej AGH

23 kwietnia 2017
VI Krakowska Konferencja Matematyki Finansowej

Podstawowy problem

Pewien model

$$\begin{cases} dS_t = S_t(r_t + \mu_t)dt + S_t\sqrt{V_t}dW_t^s + d\left(\sum_{j=1}^{N_t} Z_j\right) \\ dV_t = k_v(\theta_v - V_t)dt + \sigma_v\sqrt{V_t}dW_t^v \end{cases}$$

Cena opcji call

$$C_t = C(S_t, V_t, \Theta) = \mathbb{E}_Q\left[\exp\left(-\int_t^T r_s ds\right)(S_T - K)_+ | V_t, S_t, \Theta\right]$$

Podstawowy problem

Cena opcji call

$$C_t = C(S_t, V_t, \Theta) = \mathbb{E}_Q\left[\exp\left(-\int_t^T r_s ds\right)(S_T - K)_+ | V_t, S_t, \Theta\right]$$

Co umiemy powiedzieć obserwując ceny?

Podstawowy problem

Cena opcji call

$$C_t = C(S_t, V_t, \Theta) = \mathbb{E}_Q \left[\exp\left(-\int_t^T r_s ds \right) (S_T - K)_+ | V_t, S_t, \Theta \right]$$

Co umiemy powiedzieć obserwując ceny?

$p(\Theta|Y)$– niepewność estymacji parametrów
$p(X|Y)$– wydzielenie wpływu poszczególnych czynników
$(X = (V_t, N_t, Z_j)$-"zmienne ukryte")

Podstawowy problem

$p(\Theta|Y)$– niepewność estymacji parametrów
$p(X|Y)$– wydzielenie wpływu poszczególnych czynników

▼

Znamy: $Y|X,\Theta$
Szukamy: $X,\Theta|Y$

Czym jest MCMC?

Rozwiązanie

Czym jest MCMC?

MCMC

MCMC = Markov Chains Monte Carlo

Rozwiązanie

Czym jest MCMC?

MCMC

MCMC = Markov Chains Monte Carlo

Dwa etapy

- generowanie próbki z pewnego ciągu rozkładów (Markov Chain)
- obliczanie parametrów (Monte Carlo)

Rozwiązanie

Łańcuch Markowa

Proces stochastyczny $(X_n)_{n\in\mathbb{N}}$ nazywamy łańcuchem Markowa z rozkładem początkowym λ i macierzą przejścia $P = (p_{ij} : i, j \in I)$, jeżeli:

- X_0 ma rozkład λ,
- dla $n \geqslant 0$, warunkowo przy $X_n = i$, X_{n+1} ma rozkład $(p_{ij} : j \in I)$ i niezależy od $X_0, \ldots, X_{n-1}$.

Ergodyczny łańcuch Markowa

Jednorodny łańcuch Markowa $(X_n)_{n\in\mathbb{N}}$ nazywamy ergodycznym, jeżeli dla dowolnego $j \in I$ istnieją i niezależą od i następujące granice

$$q_j = \lim_{n\to\infty} p_{ij}^{(n)} > 0$$

oraz $\sum_{j\in I} q_j = 1$.

Rozwiązanie

Metody Monte Carlo

- Zadanie: obliczyć całkę $I_d(f) = \int_{[0,1]^d} f(\boldsymbol{x})d\boldsymbol{x}$
- Rozwiązanie: korzystając z Centralnego Twierdzenia Granicznego losujemy punkty $(\boldsymbol{t}_1, \boldsymbol{t}_2, \ldots, \boldsymbol{t}_n)$ z $[0,1]^d$ i przybliżamy $I_d(f)$ przy pomocy $I_d^n(f) = \frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{t}_i)$
- Wynik: $I_d^n(f) \to I_d(f)$, gdy $n \to \infty$

MCMC

MCMC = Markov Chains Monte Carlo

Dwa etapy

- generowanie próbki z pewnego ciągu rozkładów $\pi_n \longrightarrow p(X, \Theta | Y)$
- obliczanie parametrów

Nowy problem

Losowanie z $p(X, \Theta | Y)$ jest trudne.

Nowe rozwiązanie

Losowe pole Markowa

Niech G będzie grafem reprezentującym pewien łańcuch Markowa $(X_n)_{n \in \mathbb{N}}$. Wtedy G nazywamy losowym polem Markowa, jeżeli dla dowolnego wierzchołka X_i zachodzi

$$P(X_i | X_{G \setminus i}) = P(X_i | X_{N_i}),$$

gdzie $X_{G \setminus i}$ oznacza wszystkie wierzchołki oprócz X_i, a X_{N_i} sąsiedztwo wierzchołka X_i.

Nowe rozwiązanie

Rozkład Gibbsa

Rozkład prawdopodobieństwa $P(X)$ na grafie G nazywamy rozkładem Gibbsa, jeżeli może zostać sfaktoryzowany na iloczyn dodatnich funkcji zdefiniowanych na klikach pokrywających wszytkie wierzchołki i krawądzie grafu G, tj.

$$P(X) = \frac{1}{Z} \prod_{c \in C_G} \phi_c(X_c),$$

gdzie C_G jest zbiorem klik grafu G, a $Z = \sum_x \prod_{c \in C_G} \phi_c(X_c)$ jest stałą normalizującą.

Nowe rozwiązanie

Twierdzenie Clifforda– Hammerslay'a

Graf G reprezentujący proces $(X_n)_{n\in\mathbb{N}}$ jest losowym polem Markowa wtedy i tylko wtedy, gdy istnieje dla niego rozkład Gibbsa.

Nowe rozwiązanie

Twierdzenie Clifforda– Hammerslay'a

Graf G reprezentujący proces $(X_n)_{n\in\mathbb{N}}$ jest losowym polem Markowa wtedy i tylko wtedy, gdy istnieje dla niego rozkład Gibbsa.

Wniosek

Rozkład $p(X, \Theta|Y)$ jest jednoznacznie definiowany przez iloczyn rozkładów $p(X|\Theta, Y)$ oraz $p(\Theta|X, Y)$.

Sukces

Gibbs sampling

- ustalmy (X_0, Θ_0),
- wylosujmy

$$\Theta_n \sim p(\Theta|X_{n-1}, Y)$$
$$X_n \sim p(X|\Theta_n, Y)$$

Otrzymujemy łańcuch Markowa $(K_n)_{n=1}^{N}$, gdzie $K_n = (\Theta_n, X_n)$.

Sukces

Gibbs sampling

- ustalmy (X_0, Θ_0),
- wylosujmy

$$\Theta_n \sim p(\Theta | X_{n-1}, Y)$$
$$X_n \sim p(X | \Theta_n, Y)$$

Otrzymujemy łańcuch Markowa $(K_n)_{n=1}^N$, gdzie $K_n = (\Theta_n, X_n)$.

Zbieżność MCMC

Łańcuch $(K_n)_{n=1}^\infty$ jest ergodyczny z rozkładem stacjonarnym $p(\Theta, X | Y)$.

Ostatnia prosta

Niech $(K_n)_{n\in\mathbb{N}}$ będzie ergodycznym łańcuchem Markowa z rozkładem stacjonarnym π oraz $f \in \mathcal{L}(\mathbb{R})$. Wtedy

Ergodyczne uśrednianie

Przy dowolnym warunku początkowym K_0

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(K_k) = \int f(K) d\pi, \;\; \pi - pw$$

Centralne twierdzenie graniczne

$$\exists_{\sigma(f)} \; \sqrt{n} \left(\frac{1}{n} \sum_{k=1}^{n} f(K_k) - \int f(K) d\pi \right) \xrightarrow{d} Z \sim \mathcal{N}(0, \sigma)$$

Bibliografia

- Norris J.R., *Markov Chain*, Cambridge University Press, 1997
- Iwanik A., Misiewicz J.K., *Wykłady z procesów stochastycznych z zadaniami. Cześć pierwsza: łańcuchy Markowa*, Oficyna Wydawnicza Uniwersytetu Zielonogórskiego, 2009
- Johannes M., Polson N., *MCMC Methods for Continuous– Time Financial Econometrics*
- Cheung S., *Proof of Hammerslay– Clifford Theorem*