Metoda Monte Carlo

Jerzy Mycielski

grudzien 2012

Przybliżanie całek

- Powiedzmy, że mamy do policzenia następującą całkę:

$$\int_a^b f(x)\, dx = I$$

- Załóżmy, że funkcja $f(x)$ jest tak skomplikowna, że nie jesteśmy w stanie policzyć tej całki analitycznie
- Możemy, zastąpić liczenie tej całki liczeniem wartości oczekiwanej.
- Załóżmy, że mamy pewną zmienną losową $v \in [a, b]$ o znanej funkcji gęstości $p(\cdot)$ i zdefiniujmy $\eta = \frac{f(x)}{p(x)}$
- Wtedy:

$$\mathrm{E}(\eta) = \int_a^b \left[\frac{f(x)}{p(x)}\right] p(x)\, dx = I$$

- Wykorzystując Prawo Wielkich Liczb przy założeniu, że $\mathrm{Var}\left(\frac{f(x)}{p(x)}\right) < \infty$

$$\widehat{I} = \overline{\eta} = N^{-1} \sum_{i=1}^{N} \frac{f(x_i)}{p(x_i)} \xrightarrow{p} \mathrm{E}(\eta) = I$$

Badanie własności estymatorów

- Założmy, że znamy **P**roces **G**enerujący **D**ane (PGD)
- Załóżmy, że PGD zależy od wektora parametrów $\boldsymbol{\theta}$
- Chcemy znaleźć fomułę dla wartości oczekiwanej estymator $\widehat{\boldsymbol{\beta}}$ w małej próbie
- Oznaczmy

$$\mathrm{E}\left(\widehat{\boldsymbol{\beta}}\right) = \mathbf{G}_1\left(\boldsymbol{\theta}, T\right) = \boldsymbol{\psi}_{T1}$$

$$\mathrm{Var}\left(\widehat{\boldsymbol{\beta}}\right) = \mathrm{E}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\psi}_{T1}\right)\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\psi}_{T1}\right) = \mathbf{G}_2\left(\boldsymbol{\theta}, T\right) = \boldsymbol{\psi}_{T2}$$

gdzie T jest liczbą obserwacji

- Na postawie Monte Carlo można $\mathrm{E}\left(\widehat{\boldsymbol{\beta}}\right)$, $\mathrm{Var}\left(\widehat{\boldsymbol{\beta}}\right)$ oszacować dla *znanego* $\boldsymbol{\theta}$, T,

$$\begin{aligned}
\overline{\boldsymbol{\psi}}_{T1} &= \frac{1}{N}\sum_{i=1}^{N}\widehat{\boldsymbol{\beta}}_i \\
\overline{\boldsymbol{\psi}}_{T2} &= \frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left(\widehat{\boldsymbol{\beta}}_i - \overline{\boldsymbol{\psi}}_{T1}\right)\left(\widehat{\boldsymbol{\beta}}_i - \overline{\boldsymbol{\psi}}_{T1}\right)
\end{aligned}$$

Własności liczb pseudolosowych

- Istnieje szereg generatorów liczb pseudolosowych: najczęściej można z nich uzyskać liczby losowe z rozkładu jednostajnego
- Generator liczb losowych rozpoczyna generować liczby losowego wykorzystując ziarno (*seed*). Ciąg liczb pseudolosowych uzyskanych przy tym samym ziarnie jest zawsze ten sam
- Za pomocą generatora liczb losowych z rozkładu jednostajnego można uzyskać także zmienne z innych rozkładów
- Jeśli zmienna η ma rozkład jednostajny, to zmienna $\xi = F^{-1}(\varepsilon)$ ma rozkład dany dystrybunatą F

- Dla rozkładu jednostajnego na przedziale $[0, 1]$

$$\Pr(\xi \leq x) = x$$

uzyskujemy przekształcając

$$\Pr(\xi \leq x) = \Pr(F(\xi) \leq F(x)) = \Pr(\varepsilon \leq F(x)) = F(x)$$

ponieważ $\varepsilon = F(\xi)$, skoro $\xi = F^{-1}(\varepsilon)$

- W przypadku rozkładu normalnego można wykorzystać odwrotność rozkładu bądź przybliżonie z centralnego twierdzenia granicznego

$$\xi = \left(\sum_{j=1}^{12} \eta_j - 6\right)$$

Płaszczyzny odpowiedzi

- Główny problem z eksperymantami Monte Carlo: wnioski dotyczą jedynie PGD określonego przez $\boldsymbol{\theta}$ i T
- Można przeporwadzić wiele eksperymentów, dla różnych $\boldsymbol{\theta}$ i T ale w tym przypadku powstaje pytanie jak w syntetyczny sposób podsumować wnioski
- Rozwiązaniem są płaszczyzny odpowiedzi definiowane przez funkcję

$$\overline{\boldsymbol{\psi}}_{Ti} = \mathbf{H}_i\left(\boldsymbol{\theta}, T\right) + \mathbf{v}_{Ti}$$

- Funkcje te dopasowujemy do uzyskanych z Monte Carlo oszacowań $\overline{\boldsymbol{\psi}}_{Ti}$

- Forma funkcyjna $\mathbf{H}_i\left(\boldsymbol{\theta}, T\right)$ powinna być dobrana tak, by dopasowanie było jak najlepsze
- Całkowity błąd $\mathbf{v}_{Ti}$ płaszczyzny odpowiedzi można zdekomponować na
  - błędem pomiaru:
  $$\mathbf{v}_{1\,Ti} = \boldsymbol{\psi}_{Ti} - \overline{\boldsymbol{\psi}}_{Ti}$$
  przy czym $\mathrm{E}\left(\mathbf{v}_{1\,Ti}\right) = 0$
  - błąd aproksymacji
  $$\mathbf{v}_{2\,Ti} = \mathbf{G}_i\left(\boldsymbol{\theta}, T\right) - \mathbf{H}_i\left(\boldsymbol{\theta}, T\right)$$

Forma funkcyjna płaszczyzn odpowiedzi

- Forma $\mathbf{H}_i\left(\boldsymbol{\theta}, T\right)$ powinna być tak dobrana, by wartości asympotyczne zgadzały się wielkościami uzyskanymi z płaszczyzny odpowiedzi dla $T \rightarrow \infty$
- Typowa forma funkcyjna płaszczyzny odpowiedzi dla wartości oczekiwanej

$$H_1\left(\boldsymbol{\theta}, T\right) - \beta = \gamma_{01} I + \gamma_{11}\frac{1}{T} + \gamma_{21} I \frac{1}{T} + \sum_{j=1}^{k_1} \gamma_{j+2,1}\phi_{j,1}\frac{1}{T}$$

gdzie $\phi_{j,1}$ są funkcjami θ, $\frac{\theta}{T}$, θ^2, $\frac{1}{T}$, $\frac{1}{T^2}$ a $I = \operatorname{plim}\widehat{\beta} - \beta$

- Jeśli $\gamma_{01} = 0$, to zarówno $H_1\left(\boldsymbol{\theta}, T\right)$ jak i $G_1\left(\boldsymbol{\theta}, T\right)$ dążą do prawdziwego β

Sposoby zwiększania dokładności Monte Carlo

- Znaleź parametry, które nie wpływają na wielkość statystyki (niezmienniczość)
- Można to sprawdzić licząc wielkość statystyki dla kilku zbiorów parametrów ale *tego samego* zbioru liczb pseudolosowych
- Antytetyczne liczby pseudolosowe:
  - załóżmy, że mamy dwa estymatory $\widehat{\psi}$, $\widetilde{\psi}$ i estymator wymieszany $\overline{\psi} = \frac{1}{2}\left(\widehat{\psi} + \widetilde{\psi}\right)$ taki, że $\mathrm{E}\left(\overline{\psi}\right) = \psi$
  - wariancja estymatora $\overline{\psi}$ jest równa $\frac{1}{4}\left[\mathrm{Var}\left(\widehat{\psi}\right) + \mathrm{Var}\left(\widetilde{\psi}\right) + 2\,\mathrm{Cov}\left(\widehat{\psi}, \widetilde{\psi}\right)\right]$
  - jeśli $\mathrm{Cov}\left(\widehat{\psi}, \widetilde{\psi}\right) < 0$ to wariancja estymatora wymieszanego mniejsza od estymatorów pierwotnych
  - czasami jesteśmy w stanie tak przekształcić zbiór liczb pseudolosowych $\{\varepsilon\}$, że estymator policzony dla przekształconego zbioru jest ujemnie skorelowany z estymatorem ze zbioru oryginalnego
  - w takim przypadku możliwe będzie uzyskanie dokładniejszych oszacowań
  - najczęściej stosujemy zbiory $\{\varepsilon_i\}$ i $\{-\varepsilon_i\}$

- zmienne kontrolne:
  - załóżmy, że znamy takie $\boldsymbol{\psi}^*$, dla którego znamy $\mathrm{E}\left(\boldsymbol{\psi}^*\right)$
  - załóżmy ponadto, że $\widehat{\boldsymbol{\psi}}$ i $\boldsymbol{\psi}^*$ są dodatnio skorelowane ($\boldsymbol{\psi}^*$ zachowuje się podobnie do $\widehat{\boldsymbol{\psi}}$)
  - zdefiniujmy $\overline{\boldsymbol{\psi}} = \widetilde{\boldsymbol{\psi}} - \boldsymbol{\psi}^* - \mathrm{E}\left(\boldsymbol{\psi}^*\right)$
  - wtedy
$$\mathrm{Var}\left(\overline{\boldsymbol{\psi}}\right) = \mathrm{Var}\left(\widetilde{\boldsymbol{\psi}}\right) + \mathrm{Var}\left(\boldsymbol{\psi}^*\right) - 2\,\mathrm{Cov}\left(\widetilde{\boldsymbol{\psi}}, \boldsymbol{\psi}^*\right)$$
  - i $\mathrm{Var}\left(\overline{\boldsymbol{\psi}}\right) < \mathrm{Var}\left(\widetilde{\boldsymbol{\psi}}\right)$ jeśli $\mathrm{Cov}\left(\widetilde{\boldsymbol{\psi}}, \boldsymbol{\psi}^*\right) > \frac{1}{2}\,\mathrm{Var}\left(\boldsymbol{\psi}^*\right)$