Ολοκλήρωση Monte Carlo - (MC)

Χρίστος Μέρκατας

Τμήμα Μαθηματικών
Πανεπιστήμιο Αιγαίου

19 Μαΐου 2017

Εισαγωγή I

Σε αρκετές εφαρμογές πρέπει να υπολογίσουμε ολοκληρώματα σε χώρους μεγάλης διάστασης ή πολύπλοκα ολοκληρώματα που ίσως δεν έχουν έχουν αναλυτική έκφραση σε μία διάσταση. Κάποια παραδείγματα, μεταξύ άλλων περιλαμβάνουν

- **Στατιστική κατά Bayes:** Για παράμετρο (διάνυσμα) $\theta \in \Theta$ και δεδομένα $y \in \mathbb{Y}$ η posterior υπολογίζεται από

$$p(\theta \mid y) = \frac{p(y \mid \theta)p(\theta)}{\int_\Theta p(y|\theta')p(\theta')\mathrm{d}\theta'}$$

- **Εύρεση περιθώριας κατανομής-Missing data representation:** Όλα τα μοντέλα δύναται να θεωρηθούν ως μοντέλα ελλιπών δεδομένων, δηλαδή η πυκνότητα των παρατηρήσεων μπορεί να δοθεί από ολοκλήρωμα της μορφής:

$$p(y|\theta) = \int_{\mathcal{Z}} p(y, z|\theta)\nu(\mathrm{d}z).$$

Εισαγωγή II

- **Υπολογισμός αναμενόμενης τιμής:** $\mathbb{E}[f(y)] = \int_{\mathcal{Y}} f(y)\nu(\mathrm{d}y)$.
- ΄Αλλα προβλήματα από τη Στατιστική Μηχανική, model selection, computer vision.

Είναι φανερό ότι αν η τυχαία μεταβλητή είναι διακριτή τα παραπάνω ολοκληρώματα παριστούν αθροίσματα. Τότε η Monte Carlo μέθοδος απαλλάσει από το υπολογιστικό κόστος του αθροίσματος τιμών `δύσκολων΄ συναρτήσεων.

Υπολογισμός ολοκληρωμάτων με τη μέθοδο MC I

Έστω ότι θέλουμε να υπολογίσουμε ολοκλήρωμα της μορφής

$$I = \int_0^1 g(x)\mathrm{d}x.$$

Αν $U \sim \mathcal{U}(0, 1)$ με $f_U(u) = 1,\ u \in (0, 1)$ τότε

$$I = \int_0^1 g(u) f_U(u)\mathrm{d}u = \int_0^1 g(u) \cdot 1\mathrm{d}u = \mathbb{E}[g(U)].$$

Αν $U_1, \ldots, U_n \overset{\text{iid}}{\sim} \mathcal{U}(0, 1)$ τότε και $g(U_1), \ldots, g(U_n)$, iid με κοινή μέση τιμή I και διακύμανση έστω $\mathrm{Var}[g(U_i)] = \sigma^2$. Τότε από τον ΙΝΜΑ έπεται ότι

$$\frac{1}{n}\sum_{i=1}^{n} g(U_i) \overset{\text{a.s}}{\to} \mathbb{E}[g(U)] = I.$$

Υπολογισμός ολοκληρωμάτων με τη μέθοδο MC II

Άρα λοιπόν για τον υπολογισμό του ολοκληρώματος με τη μέθοδο Monte Carlo παράγουμε ένα ‘μεγάλο’ πλήθος n τυχαίων αριθμών $U_i \sim \mathcal{U}(0,1)$ και έχουμε

$$I = \int_0^1 g(u)\mathrm{d}u = \frac{1}{n}\sum_{i=1}^{n} g(U_i).$$

Σφάλμα της μεθόδου MC I

Οι τυχαίες μεταβλητές $g(U_i), i = 1, \ldots, n$ αποτελούν τυχαίο δείγμα με μέση τιμή I και διακύμανση $\mathrm{Var}[g(U_i)] = \sigma^2$. Τότε από το Κεντρικό Οριακό Θεώρημα για `μεγάλο' n η τ.μ $S_n = \sum_{i=1}^{n} g(U_i)$ ακολουθεί κανονική κατανομή

$$S_n \sim \mathcal{N}(nI, n\sigma^2).$$

Από τις ιδιότητες έπεται ότι

$$\mathrm{P}(nI - 3\sigma\sqrt{n} < S_n < nI + 3\sigma\sqrt{n}) \approx 0.997,$$

ή ισοδύναμα

$$\mathrm{P}\left(\left|\frac{g(U_1) + \ldots + g(U_n)}{n} - I\right| < \frac{3\sigma}{\sqrt{n}}\right) \approx 0.997.$$

- Η μέθοδος Monte Carlo εκτός από τον τρόπο υπολογισμού ολοκληρωμάτων μας παρέχει και μία εκτίμηση του σφάλματος.

Σφάλμα της μεθόδου MC II

- Το σφάλμα της μεθόδου είναι της τάξης $\mathcal{O}(n^{-1/2})$.
- Για παράδειγμα αν $n = 100$ έχουμε σφάλμα της τάξης $\mathcal{O}(100^{-1/2}) = 0.1$.
- Πρακτικά, αν θέλουμε να αυξήσουμε την ακρίβεια κατά 10 φορές (ισοδύναμα να ελαπώσουμε την τυπική απόκλιση), θα πρέπει να χρησιμοποιήσουμε επιπλέον 100 τυχαίους αριθμούς $g(U_i)$.
- Εν γένει, η σύγκλιση της μεθόδου είναι πιο γρήγορη σε σχέση με άλλες μεθόδους αριθμητικής ανάλυσης.

Παραδείγματα I

Για ολοκληρώματα της μορφής

$$I = \int_a^b g(x)\mathrm{d}x,$$

θέτουμε $y = \frac{x-a}{b-a}$ με $\mathrm{d}y = \frac{\mathrm{d}x}{b-a}$ και έχουμε ότι

$$I = \int_a^b g(x)\mathrm{d}x = \int_0^1 g(a + (b-a)y)(b-a)\mathrm{d}y$$

Παραδείγματα II

Για ολοκληρώματα της μορφής

$$I = \int_0^\infty g(x)\mathrm{d}x,$$

θέτουμε $y = \frac{1}{x+1}$ με $\mathrm{d}y = -\frac{\mathrm{d}x}{(x+1)^2}$ και έχουμε ότι

$$I = \int_0^\infty g(x)\mathrm{d}x = \int_0^1 g\left(\frac{1-y}{y}\right) y^{-2}\mathrm{d}y$$

Τον ίδιο μετασχηματισμό εφαρμόζουμε και για ολοκληρώματα της μορφής

$$I = \int_{-\infty}^0 g(x)\mathrm{d}x$$

Εφαρμογή I

Έστω $X \sim \mathcal{E}(1)$. Να υπολογιστούν

1. Η πιθανότητα η τ.μ X να πάρει τιμές μικρότερες ή ίσες του 2.
2. Η αναμενόμενη τιμή της τυχαίας μεταβλητής X^2.

Λύση.

1. Για τον αναλυτικό υπολογισμό της πιθανότητας έχουμε

$$\mathrm{P}(X \leq 2) = \int_0^2 e^{-x}\mathrm{d}x = -e^{-x}|_0^2 = 1 - e^{-2} \approx 0.8647.$$

   Για τον υπολογισμό της πιθανότητας με χρήση Monte Carlo εφαρμόζουμε το μετασχηματισμό $y = (x-0)/2$ με $x = 2y$ και $\mathrm{d}y = \mathrm{d}x/2$. Άρα έχουμε ότι

$$\int_0^2 e^{-x}\mathrm{d}x = \int_0^1 2e^{-2y}\mathrm{d}y.$$

Εφαρμογή II

- Παραγωγή $U_1, \ldots, U_n \sim \mathcal{U}(0,1)$
- Υπολογισμός $n^{-1}\sum_{i=1}^{n} 2e^{-2u_i}$.

2. Για την αναμενόμενη τιμή έχουμε

$$\mathbb{E}[X^2] = \int_0^\infty x^2 f_X(x)\mathrm{d}x = \int_0^\infty x^2 e^{-x}\mathrm{d}x = \Gamma(3) = (3-1)! = 2.$$

Για τον υπολογισμό της αναμενόμενης τιμής με χρήση Monte Carlo εφαρμόζουμε το μετασχηματισμό $y = 1/(x+1)$ με $x = (1-y)/y$ και $\mathrm{d}y = -1/(x+1)^2\mathrm{d}x$. Έπεται ότι

$$I = \int_0^\infty x^2 e^{-x}\mathrm{d}x = \int_0^1 \left(\frac{1-y}{y}\right)^2 e^{\frac{y-1}{y}} y^{-2}\mathrm{d}y.$$

- Παραγωγή $U_1, \ldots, U_n \sim \mathcal{U}(0,1)$
- Υπολογισμός $n^{-1}\sum_{i=1}^{n} \left(\frac{1-u_i}{u_i}\right)^2 e^{\frac{u_i-1}{u_i}} u_i^{-2}$.

Βιβλιογραφία I

Ευφροσύνη Ε. Μακρή (2009)
Μέθοδοι Προσομοίωσης
Πανεπιστημιακές Παραδόσεις, Πανεπιστήμιο Πατρών

Cristophe Andrieu, Nando De Freitas, Arnaud Doucet, Michael I. Jordan (2003)
An Introduction to MCMC for Machine Learning
Machine Learning, 50,5–43.