

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 εύκολη 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

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

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

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

Εισαγωγή II

- ▶ **Υπολογισμός αναμενόμενης τιμής:** $\mathbb{E}[f(y)] = \int_{\mathcal{Y}} f(y) \nu(dy)$.
- ▶ Άλλα 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

$$I = \int_0^1 g(x) dx.$$

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

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

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

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

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

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

Σφάλμα της μεθόδου MC I

Οι τυχαίες μεταβλητές $g(U_i)$, $i = 1, \dots, n$ αποτελούν τυχαίο δείγμα με μέση τιμή I και διακύμανση $\text{Var}[g(U_i)] = \sigma^2$. 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

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

ή ισοδύναμα

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

- ▶ 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 ελαττώσουμε 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

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

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

$$I = \int_a^b g(x) dx = \int_0^1 g(a + (b-a)y)(b-a) dy$$

Παραδείγματα II

Για ολοκληρώματα της μορφής

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

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

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

Τον ίδιο μετασχηματισμό εφαρμόζουμε και για ολοκληρώματα της μορφής

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