Λογισμος ΙΙ, Ολοκλήρωση συναρτήσεων στον $\mathbb{R}^d$ Εφαρμογές

Α. Ν. Γιαννακόπουλος
Τμήμα Στατιστικής

Ο.Π.Α

Εαρινό Εξάμηνο 2018

Εφαρμογές στις πιθανοτητες

Αν έχουμε μια τυχαία μεταβλητή $X : \Omega \to \mathbb{R}$, και $P(X \leq x)$ είναι η πιθανότητα η τυχαία μεταβλητή X να παίρνει τιμές μικροτερες ή ίσες του $x \in \mathbb{R}$,

$$P(X \leq x) = P(\{\omega \in \Omega \ : \ X(\omega) \leq x\},$$

τότε μπορούμε να ορίσουμε μια συνάρτηση $F : \mathbb{R} \to [0,1] \subset \mathbb{R}$ ως

$$F(x) := P(X \leq x).$$

Η συνάρτηση αυτή ονομαζεται συνάρτηση κατανομής της τυχαίας μεταβλητής X και την χαρακτηρίζει πλήρως.

Αν υπάρχει συνάρτηση $f : \mathbb{R} \to \mathbb{R}$ τέτοια ώστε

$$F(x) = \int_{-\infty}^{x} f(y)dy, \quad \forall\ x \in \mathbb{R},$$

τότε η f ονομαζεται συνάρτηση πυκνότητας πιθανοτητας της X.

Ας υποθέσουμε τώρα οτι εχουμε μια τυχαια μεταβλητή X με συνάρτηση κατανομής F_X και για μια συναρτηση $\phi : \mathbb{R} \to \mathbb{R}$ ορίζουμε την καινούργια τυχαία μεταβλητή $Y = \phi(X)$.

Ποιά θα είναι η κατανομή F_Y της Y;

Ας υποθέσουμε τώρα οτι εχουμε μια τυχαια μεταβλητή X με συνάρτηση κατανομής F_X και για μια συναρτηση $\phi : \mathbb{R} \to \mathbb{R}$ ορίζουμε την καινούργια τυχαία μεταβλητή $Y = \phi(X)$.

Ποιά θα είναι η κατανομή F_Y της Y;

Απο τον ορισμό, για κάθε $y \in \mathbb{R}$,

$$F_Y(y) = P(Y \leq y) = P(\phi(X) \leq y) = P(X \leq \phi^{-1}(y)) = F_X(\phi^{-1}(y)).$$

Άρα, αν η ϕ ειναι αντιστρέψιμη

$$F_Y(y) = F_X(\phi^{-1}(y)), \ \forall \ y \in \mathbb{R}.$$

Μία πολύ θεμελιώδης μέθοδος προσομοίωσης, η μέθοδος της αντιστροφής βασίζεται στην πολύ απλή αυτή παρατήρηση.

Τι θα μπορούσαμε να πούμε για τις πυκνοτητες πιθανότητας;

Τι θα μπορούσαμε να πούμε για τις πυκνοτητες πιθανότητας;

$$f_Y(y) = \frac{d}{dy} F_X(\phi^{-1}(y)) = F_X'(\phi^{-1}(y)) \frac{d}{dy} \phi^{-1}(y)$$

Όμως,

$$F_X'(\phi^{-1}(y) = f_X(\phi^{-1}(y)),$$

και

$$\frac{d}{dy} \phi^{-1}(y) = \frac{1}{\phi'(y)}.$$

Άρα αν $\phi \in C^1$ έχουμε επιπλέον ότι

$$f_Y(y) = \frac{1}{\phi'(y)} f_X(\phi^{-1}(y)).$$

Ας υποθέσουμε ότι έχουμε μια τυχαία μεταβλητή $X : \Omega \to \mathbb{R}^d$.

Την τυχαία αυτη μεταβλητή μπορούμε να την κατανοήσουμε και σαν ένα τυχαίο διάνυσμα

$$X = (X_1, \cdots, X_d), \;\; X_i : \Omega \to \mathbb{R}, \;\; i = 1, \cdots, d.$$

Τότε μας ενδιαφέρει η πιθανότητα

$$P(X \in U), \;\; \forall\, U \subset \mathbb{R}^d.$$

Ορισμός

Αν υπάρχει μια συνάρτηση $f_X : \mathbb{R}^d \to \mathbb{R}_+$ *τέτοια ώστε*

$$P(X \in U) = \int_U f_X(x_1, \cdots, x_d) dx_1 \cdots dx_d, \ \forall\, U \subset \mathbb{R}^d,$$

τότε η συνάρτηση f_X *ονομαζεται* ***απο κοινού*** *συνάρτηση κατανομής της* X.

Ορισμός

Οι συναρτήσεις $f_i : \mathbb{R} \to \mathbb{R}_+$ *που ορίζονται ως*

$$f_i(x_i) = \int_{-\infty}^{\infty} [\int_{-\infty}^{\infty} [\int_{-\infty}^{\infty} [\int_{-\infty}^{\infty}$$

$$\cdots [\int_{-\infty}^{\infty} f(x_1, \cdots, x_{i-1}, {\color{red}x_i}, x_{i+1}, \cdots, x_d) dx_1] \cdots] dx_{i-1}] dx_{i+1} \cdots] dx_d$$

δηλαδή ως τα επαναλαμβανόμενα ολοκληρώματα ως προς ολες τις μεταβλητές εκτός της i ονομαζονται περιθώριες κατανομές (marginal distributions).

Ένα ισοδύναμος τροπος να ορίσουμε τις περιθώριες κατανομες είναι ως εξής.

Σταθεροποιούμε το x_i και οριζουμε τις συναρτήσεις $\phi_i : \mathbb{R}^{d-1} \to \mathbb{R}_+$ ως

$$\phi_i(x_1, \ldots, x_{i-1}, x_{i+1}, \cdots, x_d; x_i) := f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \cdots, x_d)$$

Τότε,

$$\begin{aligned} f_i(x_i) &= \int_{\mathbb{R}^{d-1}} \phi_i(x_1, \ldots, x_{i-1}, x_{i+1}, \cdots, x_d; x_i) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_d \\ &= \int_{\mathbb{R}^{d-1}} f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \cdots, x_d) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_d . \end{aligned}$$

Παράδειγμα

1. *Υπολογίστε την σταθερά* A *έτσι ώστε η συναρτηση*

$$f(x_1, \cdots, x_d) = A \exp\left(-\frac{x_1^2 + \cdots + x_d^2}{2}\right)$$

να είναι συνάρτηση κατανομής πιθανοτητας στον $\mathbb{R}^d$.

2. *Υπολογίστε τις περιθώριες κατανομες για την* f.

Μετασχηματισμοί τυχαίων μεταβλητών

Ας υποθέσουμε ότι $X^* = (X_1^*, \cdots, X_d^*)$ ειναι μια τυχαια μεταβλητή με απο κοινού συναρτηση κατανομής f_{X^*} και $T : \mathbb{R}^d \to \mathbb{R}^d$ ένας C^1 μετασχηματισμός.

Φτιάχνουμε την τυχαία μεταβλητή $X = T(X^*)$.

Ποιά θα είναι η απο κοινού συνάρτηση κατανομής f_X της X;

Ας πάρουμε ένα οποιοδηποτε $U^* \subset \mathbb{R}^d$ και έστω $U = T(U^*)$.

Απο τον τύπο αλλαγής μεταβλητών σε πολλαπλά ολοκληρώματα γνωρίζουμε ότι

$$\int_U f_X(x)dx = \int_{U^*} (f_X \circ T)(x^*)|Det(DT(x^*)|dx^*$$

Όμως

$$\int_U f_X(x)dx = P(X \in U) = P(T(X^*) \in U) = P(X^* \in T^{-1}(U)) \\ = \int_{U^*} f_{X^*}(x^*)dx^*.$$

Συνεπώς

$$f_{X^*}(x^*) = (f_X \circ T)(x^*)|Det(DT(x^*)|$$

Για να συνδέσετε την f_X με την f_{X^*} αρκεί να θυμηθείτε την σχέση μεταξύ της $|Det(DT(x^*)|$ και της $|Det(DT^{-1})(x)|$, για $x = T(x^*)$.

$$\begin{aligned} f_X(x) &= \frac{1}{|Det(DT(x^*)|} f_{X^*}(x^*) \\ &= |Det(DT^{-1}(x))| f_{X^*}(T^{-1}(x)) \end{aligned}$$

Παράδειγμα (Η μέθοδος Box-Muller για την παραγωγή κανονικά κατανεμημένων τυχαίων μεταβλητών)

Οι τυχαίες μεταβλητές X_1^* και X_2^* είναι ανεξάρτητες και ομοιόμορφα κατανεμημένες στο $[0, 1]$.

Με βάση αυτές, ορίζουμε τις τυχαίες μεταβλητές X_1, X_2 ως

$$X_1 = \sqrt{-2\ln(X_1^*)}\cos(2\pi X_2^*),$$
$$X_2 = \sqrt{-2\ln(X_1^*)}\sin(2\pi X_2^*).$$

Δείξτε ότι η $X = (X_1, X_2)$ ακολουθεί την τυπική κανονική κατανομή.

Αν $X^* = (X_1^*, X_2^*)$ και $X = (X_1, X_2)$ τότε $X = T(X^*)$ με

$$\begin{aligned} T &: [0,1] \times [0,1] \subset \mathbb{R}^2 \to \mathbb{R}^2, \\ T(x_1^*, x_2^*) &= (T_1(x_1^*, x_2^*), T_2(x_1^*, x_2^*)), \\ T_1(x_1^*, x_2^*) &= \sqrt{-2\ln(x_1^*)}\cos(2\pi x_2^*), \\ T_2(x_1^*, x_2^*) &= \sqrt{-2\ln(x_1^*)}\sin(2\pi x_2^*) \end{aligned}$$

Επίσης,

$$f_{X^*}(x_1^*, x_2^*) = 1, \ \ \forall (x_1^*, x_2^*) \in [0,1] \times [0,1].$$

Θα πρέπει να είστε σε θέση να αποδείξετε οτι

$$Det(DT(x_1^*, x_2^*)) = -\frac{2\pi}{x_1^*}.$$

Συνεπώς

$$f_X(x_1, x_2) = \frac{1}{2\pi} x_1^*.$$

Απομένει να εκφράσουμε το x_1^* σαν συνάρτηση των x_1, x_2 (δηλαδη να βρούμε τον T^{-1}!).

Μπορούμε ευκολα να δούμε οτι

$$x_1^* = \exp\left(-\frac{x_1^2 + x_2^2}{2}\right),$$

συνεπώς

$$f_X(x_1, x_2) = \frac{1}{2\pi} \exp\left(-\frac{x_1^2 + x_2^2}{2}\right).$$