25-5-2023

Bayesian Estimation

$X \sim f(x|\theta) \qquad x \in \mathbb{R}^d$

$\theta \in \Theta \subseteq \mathbb{R}^k$: άγνωστη παράμετρος

Bayesian estimation:

θ : τυχαία μεταβλητή

$p(\theta)$: prior distribution (pdf ή pmf)

Παρατήρηση x από $f(x|\theta)$

$\tilde{p}(\theta|x)$ = posterior distribution

$$\tilde{p}(\theta|x) = \frac{p(\theta)\cdot f(x|\theta)}{f_X(x)} = \frac{p(\theta)\, f(x|\theta)}{\int_{\Theta} p(\theta)\, f(x|\theta)\, d\theta}$$

Εκτίμηση θ : $\mathbb{E}_{\tilde{p}(\theta|x)}[\theta]$

Πρόβλημα: Εκτίμηση μέσω προσομοίωσης

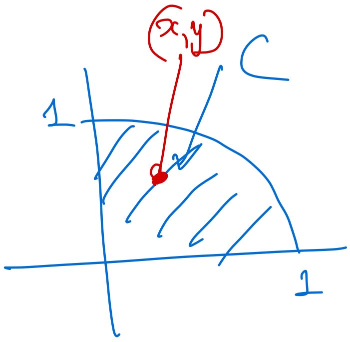
$$\boxed{\mathbb{E}_{\tilde{p}(\theta|x)}[h(\theta)]}$$

Παράδειγμα 1

Έστω X, Y iid $\mathcal{E}xp(\theta)$

$$\boxed{Z = (X,Y) \mid (X,Y) \in C}$$

$$\left(X^2+Y^2 \le 1,\ X,Y \ge 0\right)$$



$$f(x,y \mid \theta) = c\,\theta^2 e^{-\theta x} e^{-\theta y}\, 1(x^2+y^2\le 1)$$

$$c = \frac{1}{P[(X,Y)\in C]}$$

Επίσης $\theta \in (0,1] = \Theta$

prior $\theta \sim \mathcal{U}(0,1)$ $\quad p(\theta) = 1(\theta\in(0,1])$

$$\tilde p(\theta \mid (x,y))$$

$$\tilde p(\theta \mid x,y) = \frac{p(\theta)\cdot f(x,y \mid \theta)}{f(x,y)} =$$

$$= \frac{\theta^2 e^{-\theta(x+y)}\ \cancel{1(x^2+y^2\le1)}^{=1} \cdot \overbrace{1(0<\theta\le1)}^{p(\theta)}}{f(x,y)}$$

$$= \underbrace{C(x,y)\,\theta^2 e^{-\theta(x+y)}}_{\Gamma(3,\,x+y)}\ 1(0<\theta\le 1)$$

$$C \; p_{\Gamma(3,x+y)}(\theta) \cdot 1(0 < \theta \le 1)$$

$$\theta \mid (x,y) \sim \Gamma(3, x+y) \mid \theta \in (0,1]$$

$$E_{\tilde{p}}(\theta) = \int_0^1 \theta \, p(\theta|x,y) \, d\theta$$

$$E_{\tilde{p}}(h(\theta)) = \int_0^1 h(\theta) \, p(\theta|x,y) d\theta .$$

Εκτίμηση $E_{\tilde{p}}(h(\theta))$ μέσω Monte Carlo.

Γεννήτρια από $\theta \sim \Gamma(3, x+y) \mid \theta \in (0,1]$ $\Big\{$ accept/reject

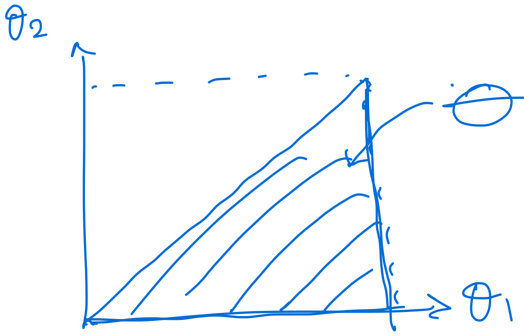
Παράδειγμα 2

(X,Y) independent. $X \sim Exp(\theta_1)$

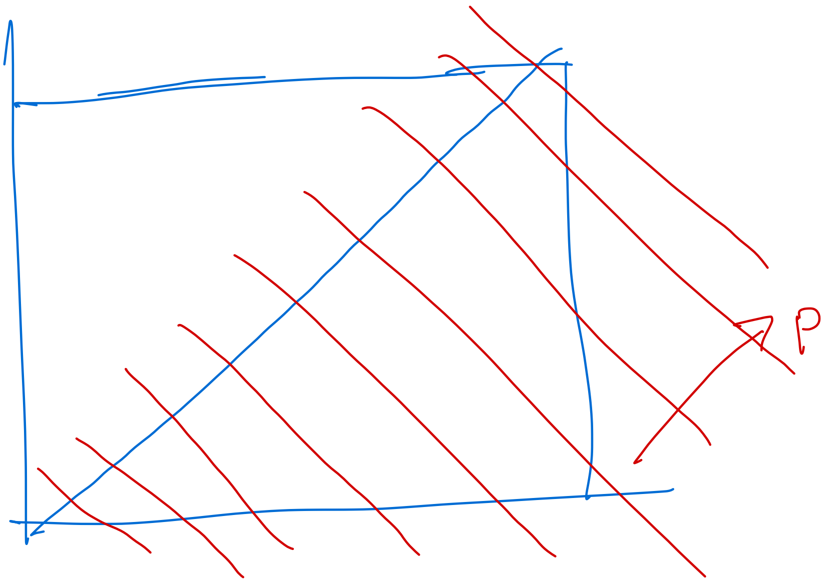
$Y \sim Exp(\theta_2)$

$\theta = (\theta_1, \theta_2)$ άγνωστη παράμετρος

$\theta \in \Theta : \{ 0 < \theta_2 < \theta_1 < 1 \}$ $\quad (EY > EX)$



prior $\quad p(\theta_1, \theta_2) = C\,(\theta_1 + \theta_2)\, 1(\theta \in \Theta)$



Έστω ζεύγος (x,y) $\quad x>0,\ y>0$

$\tilde{p}(\theta_1, \theta_2 \mid x, y) \longrightarrow E_{\tilde{p}}[h(\theta_1,\theta_2)]$

$$f((x,y)\mid\theta) = \theta_1 e^{-\theta_1 x}\,\theta_2 e^{-\theta_2 y}$$

$$\tilde{p}(\theta\mid x,y) = \frac{p(\theta)\cdot f(x,y\mid\theta)}{f(x,y)} =$$

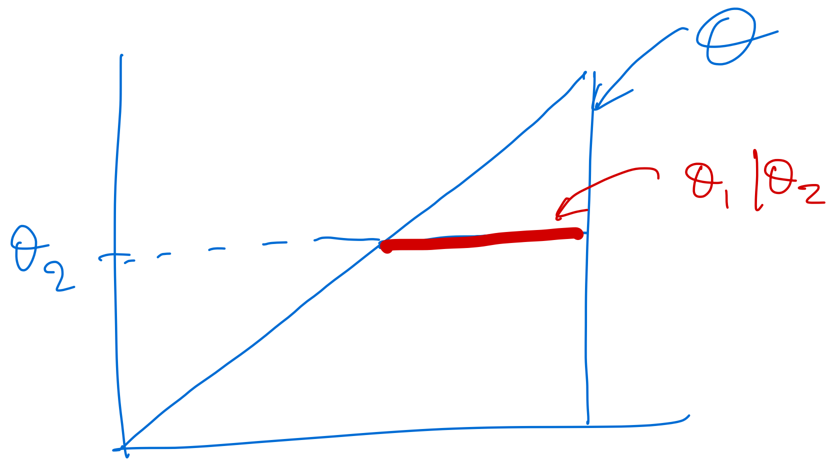
$$= \frac{\theta_1\theta_2 e^{-(\theta_1 x+\theta_2 y)}\cdot C\,(\theta_1+\theta_2)\,\mathbb{1}(0<\theta_2<\theta_1<1)}{f(x,y)} =$$

$$= C_1(x,y)\,\boxed{\theta_1\theta_2 e^{-\theta_1 x} e^{-\theta_2 y}(\theta_1+\theta_2)\,\mathbb{1}(0<\theta_2<\theta_1<1)}$$

$$= C_1\cdot b(\theta_1,\theta_2)$$

Γεννήτρια από $\tilde{p}(\theta\mid x,y)$

(1) Gibbs sampler $\longrightarrow$??



② SIR (Sampling - Importance Resampling)

γεννήτρια → $f(x) = c_1 f_0(x)$

$g(x) = c_2 g_0(x)$

Δημιουργούμε παρατηρήσεις $x_1, \dots, x_M \sim g(x)$ $\left(\begin{matrix}\text{ακριβώς} \\ \text{ή} \\ \text{MCMC}\end{matrix}\right)$

$$w_j = \frac{f_0(x_j)}{g_0(x_j)}, \quad j = 1, \dots, M$$

βάρη $q_j = \dfrac{w_j}{\sum_{j=1}^{M} w_j}$ ⇒ $\left.\begin{matrix}\text{γεννήτρια από } \{x_1, \dots, x_M\} \\ \text{με πιθανότητα } \{q_1, \dots, q_M\}\end{matrix}\right\} Z$

$$Z \xrightarrow[M \to \infty]{\emptyset} f$$

$$\boxed{\widehat{E_f(h(X))} = \frac{1}{M} \sum_{j=1}^{M} q_j h(x_j)}$$

← δεν χρειάζεται προσομοίωση της Z

Στο δικό μας πρόβλημα

γεννήτρια $\longrightarrow$

$$\tilde{p}(\theta) = C \underbrace{\theta_1\theta_2(\theta_1+\theta_2)\, e^{-\theta_1 x}\, e^{-\theta_2 y}\, 1(0<\theta_2<\theta_1<1)}_{p_0(\theta)}$$

$$\text{"}g(\theta)\text{"} = p(\theta) = (\theta_1+\theta_2)\, 1(0<\theta_2<\theta_1<1)$$

① Δημιουργούμε $\theta^1, \theta^2, \dots, \theta^M \sim p(\theta)$. $\left(\theta^j = (\theta_1^j, \theta_2^j)\right)$

② Υπολ. $w_j = \dfrac{\tilde{p}(\theta^j)}{p(\theta^j)} = \theta_1^j\theta_2^j\, e^{-\theta_1^j x}\, e^{-\theta_2^j y}$

$$q_j = \frac{w_j}{\sum_{j=1}^M w_j}$$

③ $\hat{E}_{\tilde{p}}(h(\theta)) = \dfrac{1}{M}\sum_{j=1}^M q_j\, h(\theta^j)$

Αν θέλουμε παρατηρήσεις από $\tilde{p}$:

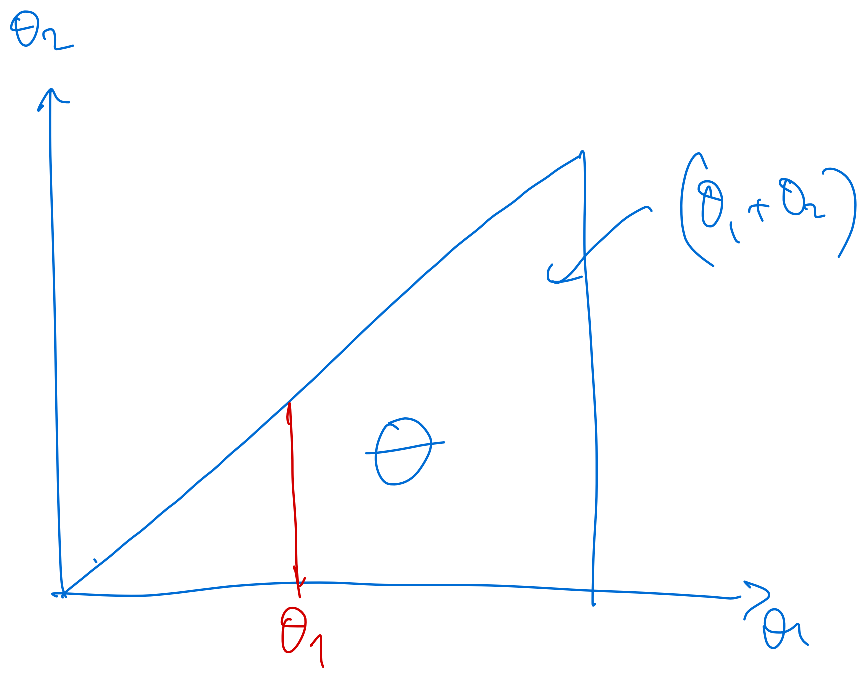
Γεννήτρια από $Z \sim \left\{\begin{array}{c}\theta^1, \dots, \theta^M \\ q_1, \dots, q_M\end{array}\right\}$ [Στο R παραγγελία την sample με βάρη]

Γεννήτρια από:

$$p(\theta_1,\theta_2) = c_1(\theta_1+\theta_2), \quad 0<\theta_2<\theta_1<1$$

(A) Gibbs sampler: πρέπει να υπολογιστούν οι $p_{\theta_1|\theta_2}(\theta_1|\theta_2)$, $p_{\theta_2|\theta_1}(\theta_2|\theta_1)$

(B)



accept/reject

(B1) accept/reject με $g(\theta)$ $\theta \in \Theta$

$\theta_1 \sim \mathcal{U}(0,1)$ $\qquad g_1(\theta_1) = 1$

$\theta_2|\theta_1 \sim \mathcal{U}(0,\theta_1)$ $\qquad g(\theta_2|\theta_1) = \frac{1}{\theta_1}, \ \theta_2 \in (0,\theta_1)$

$$\boxed{g(\theta_1, \theta_2) = \frac{1}{\theta_1}, \quad 0 < \theta_2 < \theta_1 < 1}$$

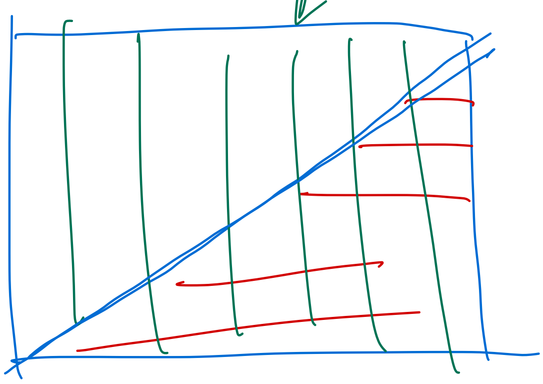
$$C = \max_{(\theta_1, \theta_2) \in \Theta} \frac{p(\theta_1, \theta_2)}{g(\theta_1, \theta_2)} = \frac{c_1(\theta_1 + \theta_2)}{\frac{1}{\theta_1}}$$

$$= c_1 \cdot \left\{ \max_{0 < \theta_2 < \theta_1 < 1} \theta_1 (\theta_1 + \theta_2) \right\} = \text{- - - -}$$

B2

g : θ_1, θ_2 iid $\mathcal{U}(0,1)$

$\theta_1, \theta_2 \sim \mathcal{U}(0,1)$



$$g = 1, \quad (\theta_1, \theta_2) \in [0,1]^2$$

$$c = \sup \frac{p(\theta)}{g(\theta)} = c_1 \sup \frac{(\theta_1 + \theta_2)\, 1(0 < \theta_2 < \theta_1 < 1)}{1} = 2 \cdot c_1$$

Αλγόριθμος για $p(\Theta)$

Δημ. $Y_1, Y_2 \sim \mathcal{U}(0,1)$ ανεξ.

(Y_1, Y_2) δεκτή με πιθ.

$$a = \frac{c_1 (Y_1 + Y_2) \cdot \mathbb{1}((Y_1, Y_2) \in \Theta)}{c\, 2} =$$

$$= \frac{Y_1 + Y_2}{2} \cdot \mathbb{1}((Y_1, Y_2) \in \Theta)$$

Άσκηση 2 Simulated Annealing for TSP

Όχημα πρέπει να κινηθεί στις πόλεις $1, 2, \dots, n$ με οποιαδήποτε σειρά (1 φορά σε κάθε πόλη)

Κέρδος u_{ij} για διαδρομή $i \to j$

$$\max V(x) = \sum_{i=1}^{n-1} u_{x_i x_{i+1}}, \quad x = (x_1, \dots, x_n)$$

μετάθεση του $\{1, \dots, n\}$

$$U = \begin{pmatrix} u_{11} & \cdots & u_{1n} \\ \vdots & \ddots & \\ u_{n1} & \cdots & u_{nn} \end{pmatrix} \quad = 0$$

πίνακας κέρδους

$n!$ μεταθέσεις (λύνεται με απαρίθμηση για μικρά n)

Simulated annealing

Markov Chain $\quad S = \{\text{μεταθέσεις}\}$

$\forall x \in S$: $N(x)$: γειτονιά της x.

εναλλαγή 2 στοιχείων του x.

πχ. $\quad x = (1, 2, 3, 4)$

$$N(x) : \left\{ \begin{array}{l} (2,1,3,4) \\ (3,2,1,4) \\ (4,2,3,1) \\ (1,3,2,4) \\ (1,4,3,2) \\ (1,2,4,3) \end{array} \right\} \qquad \binom{4}{2}$$

Αν $X_n = x \Rightarrow X_{n+1} = y \in N(x)$ ισοπίθανα.

$X_{n+1} = y$:

→ αν $V(y) \geq V(x)$ δεκτή $X_{n+1} = y$

→ αν $V(y) < V(x)$ δεκτή με πιθ. $e^{\lambda_n (V(y) - V(x))}$ (< 1)

όπου $\lambda_1, \lambda_2, \ldots$ ακολουθία $\lambda_n \to \infty$ αργά

Συνήθως $\lambda_n \sim \log(1+n)$

Έστω $\lambda_n = \log(1+n)$

Πιθ. αποδοχής $(1+n)^{V(y)-V(x)}$

Προγραμματισμός function (n)

1. Προσομοιώνουμε $U_{n\times n}$ $\quad \overset{n.x}{u_{ij}} \sim U(0,A)$

2. " το simulated annealing για δοσμένο U

Δημιουργούμε υλοποιήσεις

$X_1, X_2, \dots\dots$