22 - 5 - 2023

Αλγόριθμος SIR χρήσιμος στην περίπτωση που η βοηθητική κατανομή $g = C g_0$, g_0: γνωστή, C: άγνωστη σταθερά

Σε αυτή την περίπτωση μπορούμε να προσεγγίσουμε μια γεννήτρια της g μέσω MCMC.

Αλγόριθμος SIR για αυτή την περίπτωση

$$X \sim f(x) = C_1 f_0(x)$$

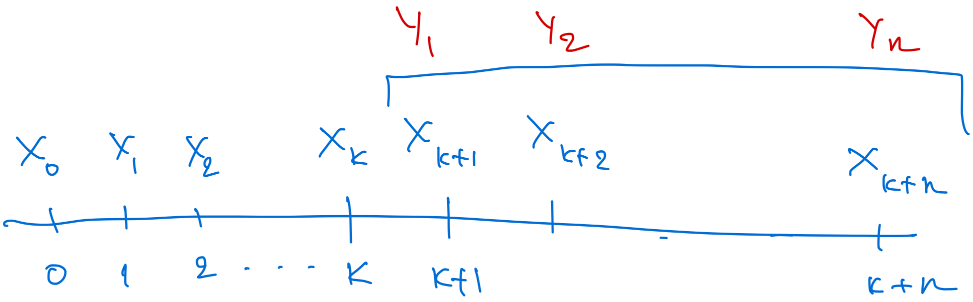
βοηθητική $\quad g(x) = C_2 g_0(x)$

Μέσω MCMC ορίζουμε Μαρκοβιανή διαδικασία $\{X_n, n=0,1,2,\dots\}$ με οριακή κατανομή την g (δεν απαιτείται να γνωρίζουμε την C_2)

Προσομοιώνουμε τις διαδοχικές καταστάσεις από αυτή τη Μαρκοβιανή διαδικασία (θέτοντας και burn-in period k)

$$Y_1, Y_2, \dots, Y_n$$

$$Y_1 = X_{k+1}, \quad Y_2 = X_{k+2} \dots$$



$$\tilde{w}(Y_j) = \frac{f(Y_j)}{g(Y_j)} = \frac{c_1}{c_2} \cdot \frac{f_0(Y_j)}{g_0(Y_j)} \quad , \; j=1,\dots,n$$

Ορίζουμε μια διακριτή τ.μ. $\tilde{X} \in \{Y_1, \dots, Y_n\}$

με $$P(\tilde{X} = Y_j) = \frac{\tilde{w}(Y_j)}{\sum_{j=1}^{n} \tilde{w}(Y_j)} = \frac{f_0(Y_j)/g_0(Y_j)}{\sum_{j=1}^{n} \frac{f_0(Y_j)}{g_0(Y_j)}} = \frac{w(Y_j)}{\sum_j w(Y_j)}$$

$$w(Y_j) = \frac{f_0(Y_j)}{g_0(Y_j)}$$

Σε αυτή την περίπτωση μπορεί να αποδειχθεί ότι

$$\tilde{X} \xrightarrow[n\to\infty]{D} f$$

Πρακτικά Προβλήματα

Το κυριότερο: όταν $n \to \infty$

η κατανομή $\tilde{X} \sim (p_1, \dots, p_n)$ είναι δύσκολο να προσομοιωθεί μέσω αντίστροφου μετ/σμού.

① Δημιουργούμε γεννήτρια accept/reject της p

$\tilde{X} \in \{1, \dots, n\}$

βοηθητική της p: $q(Y_j) = \frac{1}{n}$, $j = 1, \dots, n$

② Αν ο αρχικός μας σκοπός είναι να εκτιμήσουμε $\theta = E(h(X))$, $X \sim f(x)$

αν είχαμε παρατηρήσεις $X_1, \dots, X_N$ από f

$$\hat{\theta} = \frac{1}{N} \sum_{j=1}^{N} h(X_j)$$

Εδώ έχουμε $\tilde{X} \approx X$

<u>Εκτίμηση</u> $\tilde{\theta} = E(\tilde{X})$

Η $\tilde{X} \in \{Y_1, \dots, Y_n\}$ με $P(\tilde{X} = Y_j) = p_j = \frac{w_j}{\sum w_j}$

Επομένως $E(\tilde{X}) = \sum_{j=1}^{n} Y_j \frac{w_j}{\sum_j w_j}$

$$\Rightarrow \boxed{\tilde{\theta} = \frac{\sum_{j=1}^{n} y_j w_j}{\sum_{j=1}^{n} w_j}}$$

δεν χρειάζεται επιπλέον sampling από $\{y_1, \dots, y_n\}$

Εφαρμογή : Bayesian Statistics

X τ.μ. (γενικά διανυσματική)

$X \sim f(x|\theta)$

θ : άγνωστη παράμετρος (γενικά διανυσματική)

π.χ. $X \sim \mathcal{N}(\mu, \sigma^2)$, $\theta = (\mu, \sigma^2)$

Bayesian approach: Η θ θεωρείται η ίδια επίσης τ.μ. $\theta \in \Theta$

$\theta \sim p(\theta)$: αρχική (εκ των προτέρων) κατανομή (prior)

Παρατηρούμε μια τιμή $X = x$

Εκ των υστέρων (posterior) κατανομή του θ

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

$$p(\theta|x) = C(x)\cdot f(x|\theta)\, p(\theta) = C(x)\, f_0(x,\theta)$$

Από Bayesian statistics

Bayes estimator του $\theta \mid x$

$$\tilde{\theta} = E(\theta|x) = \int_{\theta} \theta\, p(\theta|x)\, dx$$

Γενικά αν μπορούμε να δημιουργήσουμε γεννήτρια από $p(\theta|x)$

$$\hat{\theta} = \frac{1}{N}\sum_{j=1}^{n} \theta_j \quad , \quad \underline{\theta_1, \dots, \theta_n \sim p(\theta|x)}$$

Αν δε γνωρίζουμε την $C(x)$

μπορούμε να κατασκευάσουμε γεννήτριες που δεν απαιτούν να γνωρίζουμε το $C(x)$.

π.χ. accept reject

$$p(\theta|x) = C f(x|\theta)\, p(\theta)$$

βοηθητική $g(\theta) = p(\theta)$

Δημιουργούμε παρατήρηση $\Lambda \sim p(\theta)$

Αν $\Lambda = \lambda$: $\to$ accept με πιθ.

$$\frac{p(\lambda|x)}{C\,p(\lambda)} \sim \frac{C f(x|\lambda)\, p(\lambda)}{c\, p(\lambda)}$$

$$\sim f(x|\lambda)$$

Πρόβλημα: Αν η $p(\theta)$ επίσης περίπλοκη

prior : $\underline{p(\theta) = C_0 \cdot p_0(\theta)}$, C_0: άγνωστη

π.χ. η prior $p(\theta)$ είναι η posterior δεδομένων παρατηρήσεων σε προηγούμενα βήματα.

Εφαρμόζεται η SIR

$$p(\theta|x) = C(x)\cdot f(x|\theta)\, p(\theta)$$
$$= \underbrace{C(x)\cdot C_0}_{\text{σταθερά}} \cdot \underbrace{f(x|\theta)\cdot p_0(\theta)}_{\text{γνωστή}}$$

<u>βοηθητική</u> $p(\theta) = C_0 \cdot p_0(\theta)$

SIR: $\{Y_n, n=0,1,\dots\} \sim MCMC$ με οριακή $p(\theta)$

Παίρνουμε m διαδοχικές παρατηρήσεις
$Y_1, \dots, Y_m$

$$w_j = \frac{p(Y_j|x)}{p(Y_j)} = \frac{C(x)\cdot f(x|Y_j)\, p_0(\cancel{Y_j})\cdot \cancel{C_0}}{\cancel{C_0}\, p_0(\cancel{Y_j})}$$

$$p_j = \frac{w_j}{\sum_{j=1}^{m} w_j} = \frac{f(x|Y_j)}{\sum_{j=1}^{m} f(x|Y_j)}, \quad \tilde{Y} \sim P$$

$$\tilde{Y} \xrightarrow[m\to\infty]{\mathcal{D}} p(\theta|x)$$

$$\tilde{\theta} = E(\tilde{Y}) = \frac{\sum_{j=1}^{m} Y_j\, f(x|y_j)}{\sum_{j=1}^{m} f(x|y_j)}$$

Επαναληπτική εφαρμογή

Έστω prior $\theta \sim p(\theta)$

και παίρνουμε δύο παρατηρήσεις $X_1, X_2 \overset{iid}{\sim} f(x|\theta)$

$$p(\theta | x_1, x_2) = C(x_1, x_2) \cdot \underbrace{f(x_1|\theta) f(x_2|\theta)}_{L(\theta; x_1, x_2)} p(\theta)$$

Εναλλακτικά :

$p_1(\theta, x_1) = p(\theta | x_1)$ posterior του $\theta \,|\, X_1 = x_1$

Η $p_1(\theta)$ μπορεί να θεωρηθεί prior για τη δεύτερη παρατήρηση

$$P_2(\theta | x_1, x_2) = \frac{f(x_2|\theta) \cdot P_1(\theta, x_1)}{C_2(x_2, x_1)}$$

$$P_2(\theta | x_1, x_2, P) = p(\theta | x_2, p(\theta | x_1))$$