

EP 04: MCMC

Danilo Brito; NUSP 10693250; William Veloso NUSP 10801513

Introdução

MCMC é uma abreviação para Markov Chain Monte Carlo. Esta técnica utiliza conceitos de Cadeias de Markov e Métodos de Monte Carlo para gerar amostras de variáveis aleatórias. Podemos utilizar estas amostras para calcular aproximações de integrais.

O problema

Seja $g(x) \propto \text{gamma}(C, x)|\cos(Rx)|$ onde gamma é a densidade exponencial com parâmetro C . Seja $f(x) = e^{-Cx}|\cos(Rx)|$. Temos então que $g(x) \propto f(x)$. Desta forma, X é uma variável aleatória com f.d.p.

$$g(x) = \frac{1}{K}f(x) = \frac{1}{K}e^{-Cx}|\cos(Rx)|$$

onde K é a constante de integração dada por:

$$K = \int_0^\infty f(x)dx = \int_0^\infty e^{-Cx}|\cos(Rx)|dx$$

Temos interesse em calcular a integral:

$$z = \int_0^\infty h(x)g(x)dx = \int_1^2 \frac{1}{K}e^{-Cx}|\cos(Rx)|dx = \frac{1}{K} \int_1^2 e^{-Cx}|\cos(Rx)|dx = \frac{1}{K} \int_1^2 f(x)dx$$

Onde $h(x) = I(1 < x < 2)$ é a função indicadora do intervalo $[1, 2]$. A solução analítica da integral necessária para encontrar K pode ser muito difícil. Podemos, então, utilizar métodos numéricos para encontrar o valor de K :

MCMC

Discutiremos brevemente nesta seção o método MCMC, um algoritmo para gerar amostras de uma variável aleatória X , com distribuição $g(x)$. Uma vantagem deste algoritmo é que ele não exige saber a distribuição exata da variável aleatória, bastando conhecer uma função $f(x) \propto g(x)$, isto é, uma função $f(x) = Kg(x)$. Nesta implementação, K é a constante de integração de $g(x)$.

A ideia básica do algoritmo iniciando no estado i é, segundo **STERN (2010)**¹, dada por:

1. Um candidato ao próximo estado X_{i+1} é proposto com probabilidade $P(X_{i+1} = x|X_n = x_i)$.
2. Dar um passo da seguinte forma:
 - i) A cadeia se move para o candidato X_{i+1} com *probabilidade de aceitação* $\alpha(i, i+1)$.
 - ii) Caso contrário, o candidato X_{i+1} é rejeitado e a cadeia permanece no estado X_i .
3. Voltar ao passo 1.

Vamos usar a distribuição $Normal(x_{atual}, s)$ como núcleo. Esta distribuição tem a propriedade de simetria, isto é $P(X|Y) = P(Y|X)$. Desta forma, podemos simplificar os cálculos das probabilidades de aceitação e o passo é implementado em R como segue:

Este algoritmo baseia-se na Propriedade de Markov das transições:

$$P(X_{n+1} = x|X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = P(X_{n+1} = x|X_n = x_n)$$

¹“Cognitive Constructivism and the Epistemic Significance of Sharp Statistical Hypotheses in Natural Sciences”. Julio Michael Stern, (2010)

Isto é, as transições entre estados dependem exclusivamente do estado anterior. No nosso caso, isso decorre do fato de que cada novo passo é gerado por uma Normal independente centrada no passo anterior. Prova-se que, se $P(X)$ se aproxima assintoticamente de uma *distribuição limite* $\pi(X)$ e esta é única, então os valores de $\mathbf{x} = (x_1, x_1, \dots, x_N)$ aproximam-se de uma amostra aleatória de X .

Prova-se que o uso de uma função $f(x) = Kg(x)$ é equivalente ao uso da própria $g(x)$ durante a etapa do cálculo da probabilidade de aceitação, onde cancela-se a constante de integração. No caso da probabilidade de Metropolis:

$$\frac{f(x_{prox})}{f(x)} = \frac{Kg(x_{prox})}{Kg(x)} = \frac{g(x_{prox})}{g(x)}$$

E no caso da probabilidade de Barker:

$$\frac{f(x_{prox})}{f(x) + f(x_{prox})} = \frac{Kg(x_{prox})}{Kg(x) + Kg(x_{prox})} = \frac{g(x_{prox})}{g(x) + g(x_{prox})}$$

Um pouco de teoria

Seja X uma variável aleatória que assume valores em $A \subseteq \mathbb{R}$, com função densidade de probabilidade $g(x)$.

Recordemos que como g é uma f.d.p., valem as seguintes propriedades:

$$g(x) \geq 0, \forall x \in A$$

$$\int_A g(x) dx = 1$$

Suponha uma amostra aleatória $\mathbf{x} = (x_1, x_1, \dots, x_N)$ de X de tamanho N . Recordemos a seguinte propriedade, estudada nos exercícios-programa anteriores:

$$E \left[\frac{1}{N} \sum_{i=1}^N f(x_i) \right] = \int_A f(x) g(x) dx$$

Imediatamente, notamos que esta amostra não é independente, visto que cada passo depende do passo anterior por meio da distribuição Normal e, no caso de rejeição, o passo usado na amostra é igual ao passo anterior. Apesar disso, vamos continuar utilizando essa amostra:

No script anexo, repetimos a análise para a probabilidade de aceitação de Barker.

Agradecemos a colaboração de todos os estudantes da disciplina e da Monitora que forneceram essencial colaboração no desenvolvimento deste trabalho.