EP4 - Monte Carlo via Cadeias de Markov (MCMC) - algoritmo de Metropolis-Hastings

Guilherme Navarro - 8943160

Maio de 2018

1 Monte Carlo via Cadeias de Markov

O método de Monte Carlo tem se destacado em diversas áreas de estudo devido a sua ampla aplicação. No EP3, foi visto o exemplo de sua utilização em que, a partir do método de simulação via *Importance Sampling*, é obtida amostra de distribuições a posteriori num único passo. Essa forma de integração é denominado método não iterativo, uma vez que os valores da amostra são gerados de forma independente e não há uma preocupação em relação à sua convergência.

Entretanto, existem vários casos em que as funções de densidade das distribuições de interesse são complicadas e de difícil amostragem. Uma alternativa para isso são os métodos iterativos de Monte Carlo via Cadeias de Markov (MCMC) nos quais, a princípio, obtém-se amostra através de técnicas de simulação iterativa, baseadas em cadeias de Markov ergódicas. Vale ressaltar que, nesses casos, os valores gerados não serão mais independentes, uma vez que a geração do próximo valor da amostra depende de seu valor anterior.

O presente relatório tem como objetivo simular o método de MCMC segundo o algoritmo de Metropolis-Hastings.

2 Constantes utilizadas na simulação

Para dar continuidade a simulação do método de integração *importance sampling*, serão considerados, neste relatório, os seguintes valores:

1. $RG = 50057666 \Rightarrow a = 0,50057666$
2. $CPF = 43396344847 \Rightarrow b = 0,43396344847$
3. $NUSP = 8943160 \Rightarrow c = 1,8535791$

3 Algoritmo de simulação

Para a simulação do método de MCMC, considere a seguinte função:

$$f(x) = y^{c-1}e^{-y}$$

sendo $y = x + |sin(ax + b)|$, $a = 0.RG$, $b = 0.CPF$ e $c = 1.NUSP$ utilizarei o MCMC para gerar amostras com o intuito de obter uma estimativa do valor de I pelo método iterativo.

$$I = \int_0^\infty f(x)dx$$

Utilizando uma função $g(x)$ é uma função de densidade de probabilidade da variável aleatória X, então, para gerar uma amostra dessa variável, sendo a seguinte função:

$$g(x) = \frac{x^{c-1}e^{-x}}{\Gamma(c)}$$

Sabe-se que a variável aleatória X, neste caso, possui uma complicada função, portanto, a geração de sua amostra x_1, x_2, ..., x_n deve ser realizada a partir de uma amostra ϵ_1, ϵ_2, ..., ϵ_n, da variável aleatória ε com distribuição do núcleo e com o mesmo domínio D de X.

3.1 Metropolis-Hastings

Para a realização do algoritmo de Metropolis-Hastings, suponha que a cadeia esteja no estado x_i e que $x_{i+1} = x_i + \epsilon_i$ é gerado através da distribuição do núcleo. A escolha do próximo valor da amostra é realizada a partir dos seguintes passos:

1. Especifique o valor inicial x_1;

2. Gere um valor ϵ_i da distribuição do núcleo;

3. Calcule a probabilidade de aceitação $\alpha(x_i, x_{i+1})$ e gere um valor u_i da distribuição Uniforme[0, 1], com

$$R(x_i, x_{i+1}) = \frac{g(x_{i+1})Q(x_{i+1} - x_i)}{g(x_i)Q(x_i - x_{i+1})}$$

$$\alpha(x_i, x_{i+1}) = min(1, R(x_i, x_{i+1}))$$

4. Se $u_i \leq \alpha(x_i, x_{i+1})$, então aceite o novo valor $x_{i+1} = x_i + \epsilon_i$, caso contrário, $x_{i+1} = x_i$, ou seja

$$x_{i+1} = \begin{cases} x_i + \epsilon_i & \text{se } u_i \leq \alpha \\ x_i & \text{se } u_i > \alpha \end{cases}$$

5. Volte ao passo 2.

Nota-se que o valor de $Q(x_{i+1} - x_i)$, neste caso, é equivalente a $Q(x_i - x_{i+1})$, já que as distribuições do núcleo (normal e uniforme) são simétricas, portanto, o calculo de $R(x_i, x_{i+1})$ pode ser simplificado para

$$R(x_i, x_{i+1}) = \frac{g(x_{i+1})}{g(x_i)}$$

Assim, com esse procedimento, a amostra de X pode ser obtido. No entanto, uma questão importante de ordem prática é como os valores iniciais influenciam o comportamento da cadeia. É importante que, conforme o número de iterações aumenta, a cadeia esquece gradualmente os valores iniciais e eventualmente converge para uma distribuição de equilíbrio. Com isso, em termos práticos, é comum gerar várias amostras de X e selecionar sempre o último valor delas para compor a amostra final da variável X.

Compreende-se que a meta da simulação é encontrar uma estimativa para o valor de I definido anteriormente. Para isso, sabe-se que pelo método do $ImportanceSampling$:

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

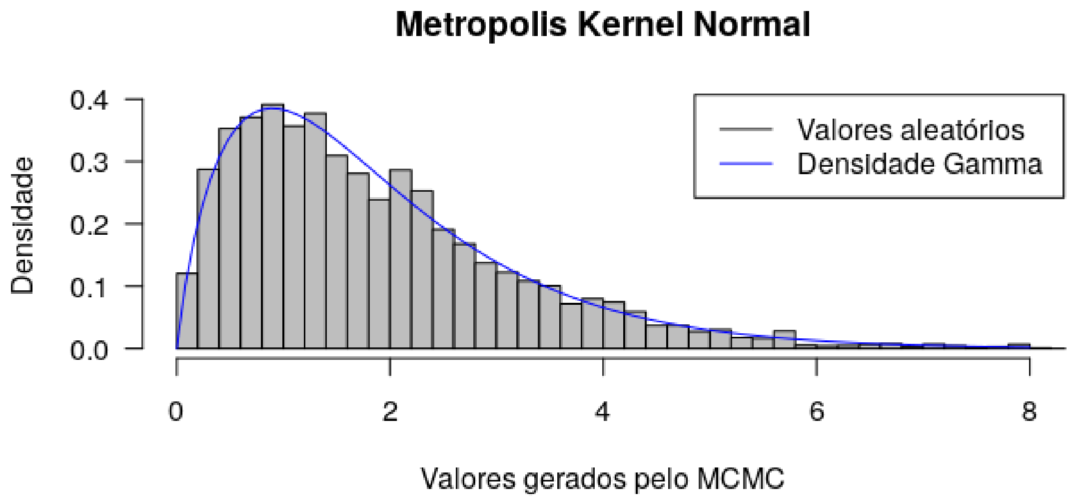
$$\Rightarrow \hat{I} \approx \frac{1}{n}\sum_{i=1}^{n} \frac{f(x_i)g(x_i)}{g(x_i)}$$

Portanto, estimar o valor de I é equivalente a estimar $\hat{I}$. Para tal, será utilizado o método de integração *Importance Sampling* visto no EP3 com o erro definido como sendo: Erro $= \frac{\sigma}{\sqrt{n}}$. Escolhendo uma amostra de B = 50000, ou seja, teremos 50000 iterações mas, como sabemos que temos que descartar os valores iniciais, no caso, descartarei os 20000 primeiros valores, restando assim um vetor de 30000 elementos para a convergência da integral. Utilizando 50 simulações e tomando a média delas.

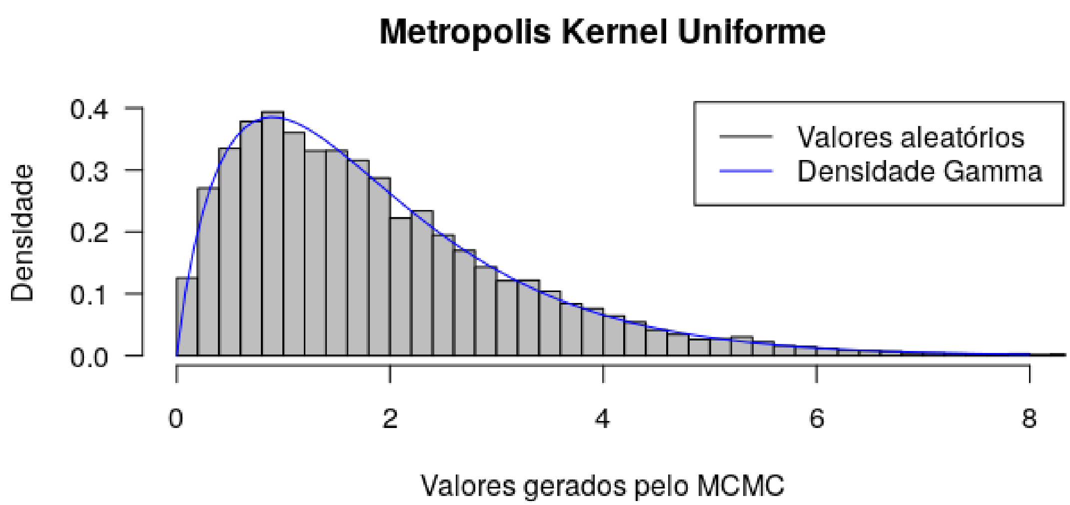
A tabela a seguir apresenta os resultados obtidos nas simulações.

Núcleo	Estimativa	σ	Erro
Normal	0.74330	0.61396	0.00274
Uniforme	0.73957	0.59613	0.00266

A figura abaixo mostram os valores da amostra de X e a curva da função $g(x)$, resultados das simulações com distribuições do núcleo normal com parâmetros $\mu = 2, \sigma = 10$. Percebe-se que essas amostras estão condizendo com a curva, uma vez que os pontos estão distribuídos de acordo com a linha.



Analogamente para figura abaixo que mostram os valores da amostra de X e a curva da função $g(x)$, resultados das simulações com distribuições do núcleo uniforme com parâmetros de $inicio = 0$ e $fim = 5$. Percebe-se que essas amostras estão condizendo com a curva, uma vez que os pontos estão distribuídos de acordo com a linha.



4 Conclusão

A partir dos valores obtidos no cálculo da integral, podemos concluir que utilizando MCMC em problemas com funções compliciadas, é possível chegar à uma solução de uma maneira prática com uma computação relativamente simples. Porém, observei que pelo método iterativo, a sua convergência pode demorar um certo tempo, pois este EP chegou no mesmo resultado do anterior, mas não de forma instantânea.

Referências

[1] YILDIRIM, I., Bayesian Inference: Metropolis-Hastings Sampling (2012). *Disponível em:* `http://www.mit.edu/~ilkery/papers/MetropolisHastingsSampling.pdf`. **acesso em 06/05/2018.**

[2] CHRISTIAN,C.; CASELA, G. ,Intorducing Monte Carlo Methods with R (2009). *Disponível em:* `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.703.5878&rep=rep1&type=pdf`. **acesso em 09/05/2018.**

[3] GENTLE, J.E. Statistics and Computuing. Random Number Generation and Monte Carlo Methods.1^{st} edition. Fairfax. Springer. (1998). 247p.

[4] STERN, J.M.,Cognitive Constructivism and the Epistemic Significance of Sharp Statistical Hypotheses in Natural Sciences(2010). *Disponível em:* `https://www.ime.usp.br/~jstern/books/evli.pdf`. **acesso em 05/05/2018.**

[5] Notas de Aula do professor J. Stern - MAP2212 - Maio 2018