Estatística Computacional - Lista 3

Gabriel Lima Novais

November 11, 2019

Exercise 1 (Gibbs Sampler)

(Q1) Sabemos que o Kernel pode ser descrito como função de $\pi_{X|Y}$ e $\pi_{Y|X}$ tal como se segue abaixo:

$$K^S_{X,Y}((x,y),(x',y')) = \pi_{X|Y}(x'|y')\pi_{Y|X}(y'|x)$$

e também temos que:

$$K^S_{X,Y}((x',y'),(x,y)) = \pi_{X|Y}(x|y)\pi_{Y|X}(y|x')$$

Para que o Kernel seja reversível é preciso que:

$$\pi_{X,Y}(x,y)K^S_{X,Y}((x,y),(x',y')) = \pi_{X,Y}(x',y')K^S_{X,Y}((x',y'),(x,y))$$

Mas realizando as contas veja que:

$$\pi_{X,Y}(x,y)K^S_{X,Y}((x,y),(x',y')) = \pi_{X,Y}(x,y)\pi_{X|Y}(x'|y')\pi_{Y|X}(y'|x) =$$
$$= \pi_{X,Y}(x,y)\pi_{X,Y}(x',y')\pi_Y(y')\pi_{Y|X}(y'|x)$$

e que:

$$\pi_{X,Y}(x',y')K^S_{X,Y}((x',y'),(x,y)) = \pi_{X,Y}(x',y')\pi_{X|Y}(x|y)\pi_{Y|X}(y|x') =$$
$$= \pi_{X,Y}(x',y')\pi_{X,Y}(x,y)\pi_Y(y)\pi_{Y|X}(y|x')$$

Portanto, fica perceptível que ambas são diferentes, pois $\pi_Y(y')\pi_{Y|X}(y'|x) \neq \pi_Y(y)\pi_{Y|X}(y|x')$. Logo é não reversível.

(Q2) Pelo item anterior temos que:

$$K^S_{X,Y}((x,y),(x',y')) = \pi_{X|Y}(x'|y')\pi_{Y|X}(y'|x)$$

Logo:

$$\int K^S_{X,Y}((x,y),(x',y'))dy' = \int \pi_{X|Y}(x'|y')\pi_{Y|X}(y'|x)dy'$$

Considerando que o Kernel $K^S_X(x,x')$ seja apenas em função de x então podemos tomar que a cadeia de Markov $(X^{(t)})$ será independente de Y e assim:

$$K_X^S(x,x') = \int K_{X,Y}^S((x,y),(x',y'))dy' = \int \pi_{X|Y}(x'|y')\pi_{Y|X}(y'|x)dy'$$

Mas utilizando Bayes e rearranjando a equação acima teremos que:

$$K_X^S(x,x') = \int \pi_{X|Y}(x'|y')\pi_{Y|X}(y'|x)dy' = \int \frac{\pi_{X,Y}(x',y')}{\pi_Y(y')}\frac{\pi_{X,Y}(x,y')}{\pi_X(x)}dy'$$

Ou seja:

$$\pi_X(x)K_X^S(x,x') = \int \frac{\pi_{X,Y}(x',y')}{\pi_Y(y')}\pi_{X,Y}(x,y')dy'$$

Por outro lado temos também que de maneira análoga:

$$\pi_X(x')K_X^S(x',x) = \int \frac{\pi_{X,Y}(x,y)}{\pi_Y(y)}\pi_{X,Y}(x',y)dy$$

Fazendo a troca de variável y para y' temos que as duas equações elaboradas são iguais, ou seja:

$$\pi_X(x)K_X^S(x,x') = \pi_X(x')K_X^S(x',x)$$

O que demonstra a reversibilidade do Kernel em questão.

(Q3) Tomando o Scan Gibbs Sampler descrito no enunciado da questão, ou seja, com $P(J=1) = P(J=2) = 1/2$ teremos que:

$$K_{X,Y}^R((x,y),(x',y')) = \frac{1}{2}\pi_{X|Y}(x'|y)\delta(y') + \frac{1}{2}\pi_{Y|X}(y'|x)\delta(x')$$

Por outro lado temos que a forma reversa do Kernel seria tal que:

$$K_{X,Y}^R((x',y'),(x,y)) = \frac{1}{2}\pi_{X|Y}(x|y')\delta(y) + \frac{1}{2}\pi_{Y|X}(y|x')\delta(x)$$

Utilizando Bayes nessa utlima equação obtemos:

$$K_{X,Y}^R((x',y'),(x,y)) = \frac{1}{2}\frac{\pi_{X,Y}(x,y')}{\pi_Y(y')}\delta(y) + \frac{1}{2}\frac{\pi_{X,Y}(x',y)}{\pi_X(x')}\delta(x)$$

Mas:

$$\pi_{X,Y}(x',y')K_{X,Y}^R((x',y'),(x,y)) = \frac{1}{2}\pi_{X|Y}(x'|y')\pi_{X,Y}(x,y')\delta(y) + \frac{1}{2}\pi_{Y|X}(y'|x')\pi_{X,Y}(x',y)\delta(x)$$

Dividindo por $\pi_{X,Y}(x,y)$ teremos que:

$$\frac{\pi_{X,Y}(x',y')}{\pi_{X,Y}(x,y)}K_{X,Y}^R((x',y'),(x,y)) = \frac{1}{2}\frac{1}{\pi_{X|Y}(x|y)\pi_Y(y)}\pi_{X|Y}(x'|y')\pi_{X,Y}(x,y')\delta(y) +$$
$$+\frac{1}{2}\frac{1}{\pi_{Y|X}(y|x)\pi_X(x)}\pi_{Y|X}(y'|x')\pi_{X,Y}(x',y)\delta(x)$$

Por Kronecker, o valor só será relevante quando so mesmos forem o que se encontram entre parênteses, e assim:

$$\frac{\pi_{X,Y}(x',y')}{\pi_{X,Y}(x,y)}K^R_{X,Y}((x',y'),(x,y)) = \frac{1}{2}\pi_{X|Y}(x'|y)\delta(y') + \frac{1}{2}\pi_{Y|X}(y'|x)\delta(x') = K^R_{X,Y}((x,y),(x',y'))$$

Ou seja, o Kernel em questão é de fato reversível.

Exercise 2 (Metropolis-within-Gibbs)

(Q1) Queremos analisar o que ocorre quando o caso de amostrar da outra desnidade condicional é possível e desta forma o MWG se torna sinplesmente o Random Scan Gibbs Sampler. Para isso vamos analisar a probabiliade α dada no algoritmo do Metropolis-Hastings:

$$\alpha(X_1|X_1^{t-1}, X_2^{t-1}) = min\left(1, \frac{\pi(X_1, X_2^{t-1})q(X_1^{t-1}|X_1, X_2^{t-1})}{\pi(X_1^{t-1}, X_2^{t-1})q(X_1|X_1^{t-1}, X_2^{t-1})}\right)$$

Mas fazendo $q(x_1'|x_1, x_2) = \pi_{X_1|X_2}(x_1'|x_2)$ teremos que:

$$\alpha(X_1|X_1^{t-1}, X_2^{t-1}) = min\left(1, \frac{\pi(X_1, X_2^{t-1})\pi_{X_1|X_2}(X_1^{t-1}|X_2^{t-1})}{\pi(X_1^{t-1}, X_2^{t-1})\pi_{X_1|X_2}(X_1|X_2^{t-1})}\right)$$

Abrido o primeiro termo do numerador e do denominador com a regra de Bayes chegamos em:

$$\alpha(X_1|X_1^{t-1}, X_2^{t-1}) = min\left(1, \frac{\pi_{X_2}(X_2^{t-1})\pi_{X_1|X_2}(X_1|X_2^{t-1})\pi_{X_1|X_2}(X_1^{t-1}|X_2^{t-1})}{\pi_{X_2}(X_2^{t-1})\pi_{X_1|X_2}(X_1^{t-1}|X_2^{t-1})\pi_{X_1|X_2}(X_1|X_2^{t-1})}\right)$$

Cancelando os termos iguais na razão dentro do mínimo podemos encontrar que:

$$\alpha(X_1|X_1^{t-1}, X_2^{t-1}) = min\,(1,1) = 1$$

E assim concluimos o que desejávamos.

(Q2) Sabemos que o Kernel de transição (conforme é apresentado nos slides) é dado por:

$$K((x_1,x_2),(x_1',x_2')) = (\alpha(x_1'|x_1,x_2)q(x_1'|x_1,x_2) + (1-a(x_1,x_2))\delta_{x_1}(x_1'))\pi_{X_2|X_1}(x_2'|x_1')$$

Onde sabemos que:

$$a(x_1,x_2) = \int \alpha(u|x_1,x_2)q(u|x_1,x_2)du$$

Para verificar a invariância de π precisamos integrar o Kernel, e para isso será realizada a integração duas vezes, pois ele é definido com o par (x_1, x_2). Assim temos que:

$$\int\int \pi(x_1,x_2)K((x_1,x_2),(x_1',x_2'))dx_1dx_2$$

Substituindo nessa integral dupla os valores anteriores temos que:

$$\int\int \pi(x_1,x_2)K((x_1,x_2),(x_1',x_2'))dx_1dx_2 =$$

$$= \int\int \pi(x_1,x_2)(\alpha(x_1'|x_1,x_2)q(x_1'|x_1,x_2) + (1-a(x_1,x_2))\delta_{x_1}(x_1'))\pi_{X_2|X_1}(x_2'|x_1')dx_1dx_2 =$$

$$= \int\int \pi_{X_2}(x_2)\pi_{X_1|X_2}(x_1|x_2)(\alpha(x_1'|x_1,x_2)q(x_1'|x_1,x_2) + (1-a(x_1,x_2))\delta_{x_1}(x_1'))\pi_{X_2|X_1}(x_2'|x_1')dx_1dx_2 =$$

Separando em cada componente e deslocando a integral:

$$= \int \pi_{X_2}(x_2)\int \pi_{X_1|X_2}(x_1|x_2)(\alpha(x_1'|x_1,x_2)q(x_1'|x_1,x_2) + (1-a(x_1,x_2))\delta_{x_1}(x_1'))dx_1\pi_{X_2|X_1}(x_2'|x_1')dx_2 =$$

Porém veja que:

$$\int \pi_{X_1|X_2}(x_1|x_2)(\alpha(x_1'|x_1,x_2)q(x_1'|x_1,x_2) + (1-a(x_1,x_2))\delta_{x_1}(x_1'))dx_1 = \pi_{X_1|X_2}(x_1'|x_2)$$

Então utilizando este fato e Bayes encontramos que:

$$\int\int \pi(x_1,x_2)K((x_1,x_2),(x_1',x_2'))dx_1dx_2 = \int \pi_{X_2}(x_2)\pi_{X_1|X_2}(x_1'|x_2)\pi_{X_2|X_1}(x_2'|x_1')dx_2 \implies$$

$$\implies \int\int \pi(x_1,x_2)K((x_1,x_2),(x_1',x_2'))dx_1dx_2 = \int \pi(x_1',x_2)\pi_{X_2|X_1}(x_2'|x_1')dx_2 = \pi_{X_1}(x_1')\pi_{X_2|X_1}(x_2'|x_1') \implies$$

$$\implies \int\int \pi(x_1,x_2)K((x_1,x_2),(x_1',x_2'))dx_1dx_2 = \pi(x_1',x_2')$$

Portanto chegamos no fato de que o Kernel deixa a distribuição π invariante.

Exercise 3 (Metropolis-Hastings and Gibbs Sampler)

(Q1) Queremos mostrar que:

$$\pi(x)T(x,y) = \pi(y)T(y,x)$$

Veja que:

$$T(x,y) = \alpha(x,y)q(x,y) + (1-\sum\alpha(x,z)q(x,z))\delta(y) = \frac{\gamma(x,y)}{\pi(x)q(x,y)}q(x,y) + (1-\sum\frac{\gamma(x,z)}{\pi(x)q(x,z)}q(x,z))\delta(y) \implies$$

$$\implies T(x,y) = \frac{\gamma(x,y)}{\pi(x)} + (1-\sum\frac{\gamma(x,z)}{\pi(x)})\delta(y)$$

Agora veja que:

$$T(y,x) = \frac{\gamma(y,x)}{\pi(y)} + (1-\sum\frac{\gamma(y,z)}{\pi(y)})\delta(x)$$

Retirando o denominador de ambos os T's obtemos as duas equações que seguem:

(1) $\pi(x)T(x,y) = \gamma(x,y) + (\pi(x) - \sum\gamma(x,z))\delta(y)$

(2) $\pi(y)T(y,x) = \gamma(y,x) + (\pi(y) - \sum \gamma(y,z))\delta(x)$

Como sabemos que $\gamma(x,y) = \gamma(y,x)$ e que δ é um síbolo de Kronecker o resultado pedido foi encontrado, ou seja, (1)=(2).

(Q2) Para o algoritmo de Metropolis-Hastings, temos que a probabilidade $\alpha(x,y)$ é dada por:

$$\alpha(x,y) = min(1, \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)})$$

Multiplicando dos dois lados por $\pi(x)q(x,y)$ encontramos que:

$$\pi(x)q(x,y)\alpha(x,y) = \pi(x)q(x,y)min(1, \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}) \implies$$

$$\implies \pi(x)q(x,y)\alpha(x,y) = min(\pi(x)q(x,y), \pi(y)q(y,x)) \implies$$

$$\gamma(x,y) = min(\pi(x)q(x,y), \pi(y)q(y,x))$$

Agora fazendo o processo inverso para o algoritmo de Baker, isto é, dividindo ambos os lados por $\pi(x)q(x,y)$, encontramos o seguinte:

$$\gamma(x,y) = \frac{\pi(x)q(x,y)\pi(y)q(y,x)}{\pi(x)q(x,y) + \pi(y)q(y,x)} \implies$$

$$\implies \frac{\gamma(x,y)}{\pi(x)q(x,y)} = \frac{1}{\pi(x)q(x,y)}\frac{\pi(x)q(x,y)\pi(y)q(y,x)}{\pi(x)q(x,y) + \pi(y)q(y,x)} \implies$$

$$\implies \alpha(x,y) = \frac{\pi(y)q(y,x)}{\pi(x)q(x,y) + \pi(y)q(y,x)}$$

E esta é a probabilidade de aceitação pedida.

(Q3) Nessa questão queremos verificar qual dos dois Kernels de Transição é o menor. Para isso analisemos o caso em que $x \neq y$ (e lembrando das respostas obtidas no item anterior) e assim temos então:

(MH) $TMH = \alpha(x,y)q(x,y) = \frac{\gamma(x,y)}{\pi(x)} = \frac{min(\pi(x)q(x,y), \pi(y)q(y,x))}{\pi(x)}$

(B) $TB = \alpha(x,y)q(x,y) = \frac{\gamma(x,y)}{\pi(x)} = \frac{1}{\pi(x)}\frac{\pi(x)q(x,y)\pi(y)q(y,x)}{\pi(x)q(x,y)+\pi(y)q(y,x)}$

Agora precisamos analisar quem é o maior deles. Suponha sem perda de generalidade (para o outro caso o problema é análogo) que $\pi(x)q(x,y) \leq \pi(y)q(y,x)$.Então teremos que :

$$TMH = q(x,y)$$

$$TB = q(x,y)\frac{\pi(y)q(y,x)}{\pi(x)q(x,y) + \pi(y)q(y,x)}$$

Logo teremos que:

$$TB = TMH \times r$$

onde $r = \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)+\pi(y)q(y,x)}$.

Veja que $r \leq 1$ e desta forma temos que $TB = r \times TMH \leq 1 \times TMH$. Portanto temos que $TB \leq TMH$. Por outro lado sabemos do Teorema citado no enunciado da questão que o T que for maior possui uma variância assintótica menor. Logo o TMH possui essa variância assintótica menor

(Q4) Vamos utilizar a definição dada no enunciado para o cálculo do Kernel de Transição.Verificar que os Kernels de Transição são π reversíveis segue a mesma lógica apresentada em itens anteriores.

Veja que para valores de $y_k \neq x_k$, a equação de $T(x,y)$ vira $T(x,y) = \alpha(x,y)q(x,y)$. Logo para o Random Scan Gibbs Sampler o valor do Kernel de Transição pode ser descrito como:

$$TRSG(x,y) = \frac{\pi_{X_K|X_{-K}}(y_K|x_{-K})\delta_{x_{-K}}(y_{-K})}{d}$$

Vale ressaltar que o valor d que está dividindo a equação acima provém de uma distribuição uniforme, que equivale em termos de notação ao valor de $\pi(x)$ dos itens anteriores. Agora seguindo o mesmo procedimento para o Modified Random Scan Gibbs Sampler, podemos escrever:

$$TMRSG(x,y) = \frac{q(y_K|x_{-K})min\left(1, \frac{1-\pi_{X_K|X_{-K}}(x_K|x_{-K})}{1-\pi_{X_K|X_{-K}}(y_K|x_{-K})}\right)\delta_{x_{-K}}(y_{-K})}{d} \implies$$

$$\implies TMRSG(x,y) = \frac{\frac{\pi_{X_K|X_{-K}}(y_K|x_{-K})}{1-\pi_{X_K|X_{-K}}(x_K|x_{-K})}min\left(1, \frac{1-\pi_{X_K|X_{-K}}(x_K|x_{-K})}{1-\pi_{X_K|X_{-K}}(y_K|x_{-K})}\right)\delta_{x_{-K}}(y_{-K})}{d} \implies$$

$$\implies TMRSG(x,y) = \pi_{X_K|X_{-K}}(y_K|x_{-K})\frac{min\left(\frac{1}{1-\pi_{X_K|X_{-K}}(y_K|x_{-K})}, \frac{1}{1-\pi_{X_K|X_{-K}}(x_K|x_{-K})}\right)\delta_{x_{-K}}(y_{-K})}{d}$$

Agora tal como foi feito no item anterior, temos como objetivo calacular quem será o maior deles, pois de acordo com o teorema do enunciado do item 3, este será o que possui menor variância assintótica e portanto para o nosso caso isto será o melhor. Vamos então a comparação:

$$TMRSG(x,y) \geq TRSG(x,y)$$

Logo, o MRSG é o que possui menor variância assintótica.

Exercise 4 (Metropolis-Hastings)

(Q1) A "proposal" é simplesmente uma integral em θ tal que:

$$q(x'|x) = \int g(x'|\theta)f(\theta|x)d\theta$$

Isso se deve ao fato de que $V \sim f(.)$ e que $X \sim g(.)$. O maior problema da função $f(.)$ é que não sabemos avaliar a mesma de forma pontual e isso gera uma dificuldade em se calcular a probabilidade α de aceitação do algoritmo de Metropolis-Hastings.

(Q2) Considerando a forma randomizada do algoritmo de Metropolis-Hastings descrita no enunciado, podemos verificar que a razão presente na probabilidade de aceitação seria justamente:

$$r = \frac{\pi(x')g(x|\theta')}{\pi(x)g(x'|\theta)}$$

Agora façamos algumas manipulações algébricas para encontrar a razão r no formato que seria no caso de um "target" de $\bar{\pi}(x, \theta)$:

$$r = \frac{\pi(x')g(x|\theta')}{\pi(x)g(x'|\theta)} \implies r = \frac{\pi(x')g(x|\theta')}{\pi(x)g(x'|\theta)} \frac{f(\theta|x)f(\theta'|x')}{f(\theta|x)f(\theta'|x')} \implies$$

Agora juntando os termos convenientemente e depois lembrando que $\bar{\pi}(x) = \pi(x)$ podemos encontrar:

$$r = \frac{\pi(x')g(x|\theta')}{\pi(x)g(x'|\theta)} \frac{f(\theta|x)f(\theta'|x')}{f(\theta|x)f(\theta'|x')} \implies r = \frac{\pi(x')f(\theta'|x')g(x|\theta')f(\theta|x)}{\pi(x)f(\theta|x)g(x'|\theta)f(\theta'|x')} \implies$$

$$\implies r = \frac{\pi(x',\theta')g(x|\theta')f(\theta|x)}{\pi(x,\theta)g(x'|\theta)f(\theta'|x')} = \frac{\bar{\pi}(x',\theta')g(x|\theta')f(\theta|x)}{\bar{\pi}(x,\theta)g(x'|\theta)f(\theta'|x')}$$

Ou seja, caso tomemos uma proposal tal que $q((x',\theta')|(x, theta)) = g(x'|\theta)f(\theta'|x')$, então encontramos que o $\bar{r} = r$. Logo podemos concluir que o Kernel de transição possui uma distribuição invariante $\bar{\pi}$.

(Q3) Utilizando a decomposição conhecida do mínimo dada por $min(U,V) = \frac{U+V}{2} - \frac{|V-U|}{2}$ temos que:

$$\begin{aligned} E(min(U,V)) &= E(\frac{U+V}{2} - \frac{|V-U|}{2}) = E(\frac{V-U+2U}{2} - \frac{|V-U|}{2}) \\ &= E(U + \frac{V-U}{2} - \frac{|V-U|}{2}) = E(U) + E(\frac{V-U}{2} - \frac{|V-U|}{2}) = \\ &= E(U) + \frac{E(V-U)}{2} - \frac{E(|V-U|)}{2} \end{aligned}$$

ou de forma análoga poderíamos ter que:

$$E(min(U,V)) = E(V) + \frac{E(U-V)}{2} - \frac{E(|U-V|)}{2}$$

Se assumirmos que $E(V) \geq E(U)$ então $E(|V-U|) \geq E(V-U)$ e então sabemos que $E(min(U,V) \leq E(U)$ e desta maneira conseguimos achar que $min(E(U), E(V)) = E(U)$ e assim provamos o que foi pedido. Por outro lado caso se tomarmos o caso contrário, $E(V) \leq E(U)$, podemos verificar que $E(|V-U|) \leq E(V-U)$ e assim $min(E(U), E(V)) = E(V)$, o que também prova o que foi pedido.

(Q4) Dúvida!

Exercise 5 (Thinning of a Markov chain)

(Q1) Apesar de podermos considerar o valor esperado um operador Linear e então demonstrar com os conhecimentos de álgebra linear, vamos utilizar a dica. Veja que:

$$(Y - \alpha Z)^2 \geq 0 \implies E((Y - \alpha Z)^2) \geq 0 \implies$$

$$\implies E(Y^2 - 2\alpha YZ + \alpha^2 Z^2) \geq 0 \implies$$
$$\implies E(Y^2) - 2\alpha E(YZ) + \alpha^2 E(Z^2) \geq 0$$

Veja que a equação acima se trata de um polinômio do segundo grau em termsos de α. Portanto para que o valor seja positivo tal como está descrito acima precisamos que o valor do $\triangle$ seja menor do que zero para todo número do domínio. Assim, calculemos este valor:

$$\triangle \leq 0 \implies (2E(YZ))^2 - 4(E(Z^2))(E(Y^2)) \leq 0 \implies$$
$$\implies (E(YZ))^2 \leq E(Z^2)E(Y^2)$$

Fornecendo exatamente a desigualdade que queriamos encontrar.

(Q2) Utilizando a fórmula da covariância temos que:

$$Cov(YZ) = E(YZ) - E(Y)E(Z)$$

Mas considerando os valores de $E(Y) = E(Z)$ então podemos fazer (associado ao Cauchy-Schwartz) a seguinte manipulação:

$$Cov(YZ) = E(YZ) - E(Y)E(Z) \leq \sqrt{E(Y^2)E(Z^2)} - E(Y)E(Z) =$$

$$= \sqrt{E(Y^2)E(Y^2)} - E(Y)E(Y) = E(Y^2) - E^2(Y) = Var(Y)$$

E assim obtemos o resultado pedido.

(Q3) Dúvida!

Simulation question (Probit model — Gibbs and M-H)

(Q1) A minha escolha dos parâmetros foram:

- $\beta^T = [3, -2]$
- $n = 50$
- $qtd(Y_i = 0) = 20$ e $qtd(Y_i = 1) = 30$

Para conseguir estruturar o raciocínio e o próprio código utilizou-se a "seed" $= 50$.

(Q2) A função de verossimilhança pode ser dada por:

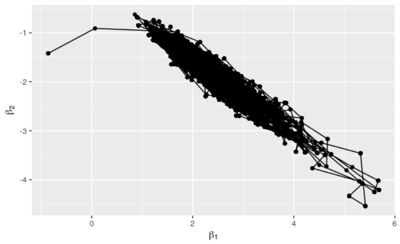
$$P(Y_1 = k_1, \ldots, Y_n = k_n \mid \beta) = \prod_{i=1}^{n} \Phi(X_i^T \beta)^{k_i} (1 - \Phi(X_i^T \beta))^{1-k_i}$$

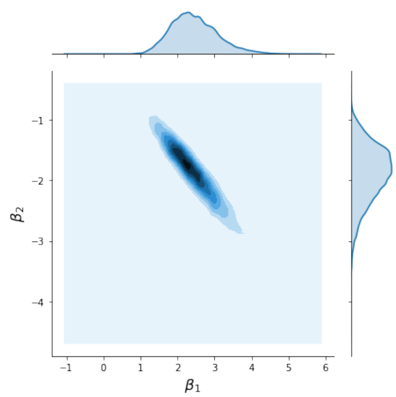
Para a realização do cálculo da distribuição da prior utilizou-se a seguinte matriz de covariância:

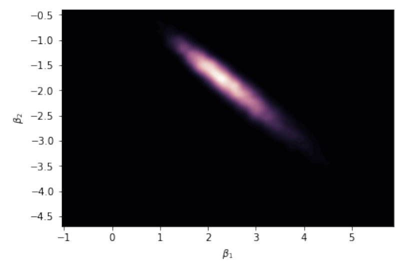
$$\sum = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Então tendo em posse a função que calculava a prior de β e uma função que calculava a função de verossimilhança de Y condicionada a β foi possível construir a função que retornava o log da posterior.

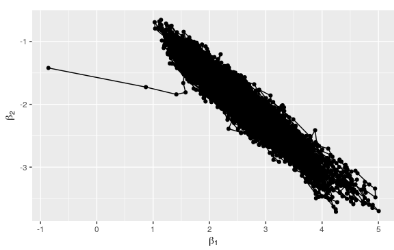
(Q3) Após construir o algoritmo de Metropolis-Hastings foram realizadas 100.000 interações, e o resultado que deveria chegar próximo do valor do $\beta = [3, -2]$ conseguiu atingir a região mais importante, como pode ser observado nos gráficos abaixo:

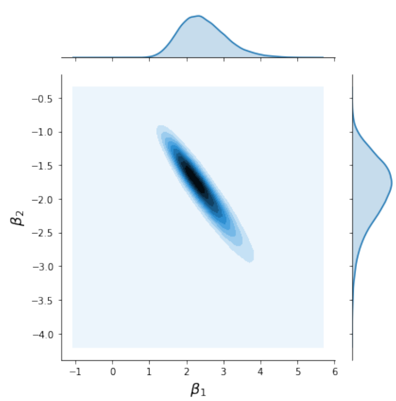


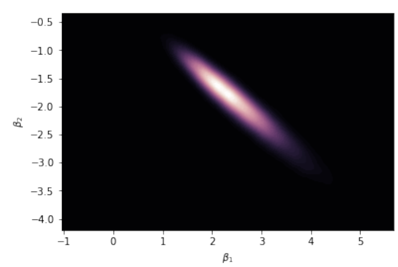




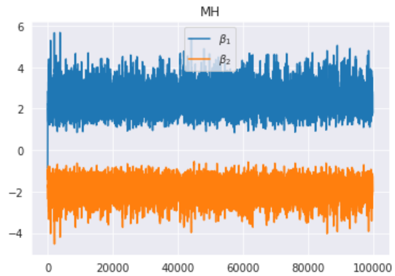
(Demais) Utilizando uma variável no auxiliar no Gibbs sampler os resultados obtidos são os que se seguem nos gráficos abaixo:

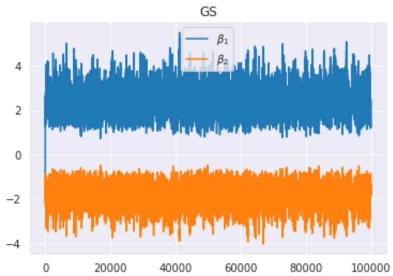


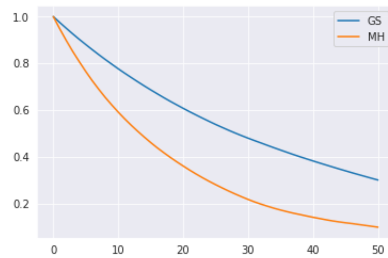




Para a comparação entre os dois métodos podemos observar a evolução da cadeia de Markov da simulação e depois verificar as autocorrelações, cujos resultados estão disponíveis abaixo:







E como podemos verificar o Metropolis-Hastings de fato ficou menos autocorrelacionado. Logo o Metropolis-Hasting performou melhor.