

Trabalho 1 - Base Lógica

Modelos Compartimentais em Epidemiologia e Inferência Bayesiana

25 de janeiro de 2022

Considere o problema proposto no Trabalho 1. Seguindo a mesma notação, temos um conjunto de dados composto por $n = 40$ termos da sequência $\{y_j\}$, com $j = 1, \dots, n$. Sabe-se que a função de verossimilhança segue distribuição normal. A função densidade de probabilidade de um termo genérico da sequência é, portanto,

$$f_Y(y_j; \mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{1}{2} \frac{(y_j - \mu)^2}{\sigma^2}\right),$$

com média μ e variância σ^2 . Assumindo que Y é uma variável aleatória cujas observações são i.i.d., a função de verossimilhança pode ser escrita como

$$\begin{aligned} L(\mu, \sigma^2; y_1, \dots, y_n) &= \prod_{j=1}^n f_Y(y_j; \mu, \sigma^2) \\ &= \prod_{j=1}^n (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{1}{2} \frac{(y_j - \mu)^2}{\sigma^2}\right) \\ &= (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{j=1}^n (y_j - \mu)^2\right) \end{aligned}$$

Tomando o logaritmo natural da função de verossimilhança, obtemos

$$\begin{aligned} \mathcal{L}(\mu, \sigma^2; y_1, \dots, y_n) &= \ln(L(\mu, \sigma^2; y_1, \dots, y_n)) \\ &= \ln\left((2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{j=1}^n (y_j - \mu)^2\right)\right) \\ &= \ln\left((2\pi\sigma^2)^{-n/2}\right) + \ln\left(\exp\left(-\frac{1}{2\sigma^2} \sum_{j=1}^n (y_j - \mu)^2\right)\right) \\ &= -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{j=1}^n (y_j - \mu)^2 \\ &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{j=1}^n (y_j - \mu)^2 \end{aligned}$$

No caso particular dos dados analisados, vamos assumir que o modelo capaz de descrever o comportamento de $\{y_j\}_{j=1}^{40}$ é definido pela função exponencial crescente, dada por

$$\mathcal{M}(t, \theta) = y_1 \exp(\theta_1 t),$$

para $t_j \in \{1, \dots, 40\}$, onde $y_1 = I(t_1)$, o primeiro valor da série temporal do número de indivíduos infectados diariamente, é adotado como condição inicial e θ_1 é um parâmetro do modelo (taxa de crescimento dos indivíduos infectados).

Considere inicialmente que o vetor θ a ser estimado tenha apenas como componente a taxa de crescimento θ_1 . Para $\theta = (a)$ escolhido arbitrariamente, se avaliarmos $\mathcal{M}(t, \theta)$ para todo $t = t_j$, teremos um vetor $\mu = (\mathcal{M}(t_1, \theta), \dots, \mathcal{M}(t_{40}, \theta))$, isto é, $\mu = (y_1 \exp(at_1), \dots, y_1 \exp(at_{40}))$. No problema proposto, a função de verossimilhança está centrada em $\mathcal{M}(t, \theta)$. Explicitamente, isso significa que

$$\begin{aligned}\mathcal{L}(\mathcal{M}(t, \theta), \sigma^2; y_1, \dots, y_n) &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{j=1}^n (y_j - \mu_j)^2 \\ &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{j=1}^n (y_j - y_1 \exp(at_j))^2\end{aligned}$$

Para avaliarmos $\mathcal{L}(\mathcal{M}(t, \theta), \sigma^2; y_1, \dots, y_n)$ na expressão anterior, a variância σ^2 da distribuição normal é desconhecida—não conhecemos explicitamente o valor do ruído presente nos dados. Portanto, σ^2 tem que ser tratado como um *hiperparâmetro* do modelo e deve ser inferido. Com isso, o vetor de parâmetros assume $\theta = (\theta_1, \sigma^2)$.

Agora, se tomarmos duas escolhas arbitrárias $\theta_1 = a$ e $\sigma^2 = b$, de forma que $\theta = (a, b)$, temos que

$$\mathcal{L}(\mathcal{M}(t, \theta), \sigma^2; y_1, \dots, y_n) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(b) - \frac{1}{2b} \sum_{j=1}^n (y_j - y_1 \exp(at_j))^2.$$

A avaliação de $\mathcal{L}(\mathcal{M}(t, a), b; y_1, \dots, y_n)$ vai fornecer uma medida do quão bem o modelo $\mathcal{M}(t, \theta)$ descreve o comportamento dos dados $\{y_j\}$, considerando tais parâmetros a e b . Neste contexto, devemos avaliar a combinação de diversas amostras para a e b , de maneira a obter o modelo que possibilite o *melhor ajuste dos dados*, o objetivo principal do trabalho. Neste ponto, entra o papel fundamental do método de amostragem, para auxiliar na formação do conjunto de amostras que aproximam as distribuições de θ_1 e σ^2 .

Para o nosso problema, a intuição inicial é de que os parâmetros a serem inferidos têm distribuição uniforme, com

$$\begin{aligned}p(\theta_1) &= \mathcal{U}(0, 1, 0, 2) \text{ dia}^{-1} \\ p(\sigma) &= \mathcal{U}(10, 15) \text{ indivíduos}^2.\end{aligned}$$

Explicitamente, portanto, sem considerar o conhecimento adquirido com os dados $\{y_j\}$, temos $p(\theta_1) = 10$ e $p(\sigma) = 0,2^2$. Neste caso, podemos usar o Teorema de Bayes para inferir a distribuição *a posteriori* de θ_1 e σ^2 . Desconsiderando o fator multiplicativo da evidência, $1/p(\mathbf{y})$, o Teorema de Bayes pode ser escrito como

$$p(\theta | \mathbf{y}) \propto p(\mathbf{y} | \theta) p(\theta).$$

Agora vamos dar enfoque ao papel do método de amostragem na aproximação de $p(\theta | \mathbf{y})$. Abaixo é descrita uma rotina computacional genérica para gerar um conjunto de S amostras, das quais esperamos que $\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(S)}\} \sim \theta | \mathbf{y}$. Vamos considerar o algoritmo de Metropolis na descrição abaixo. Entretanto, note que qualquer amostrador pode ser adotado, com as devidas adequações referentes a cada metodologia.

1. Defina uma estimativa inicial $\mathbf{x}^{(1)} = (x_1, x_2)$ para θ_1 e σ^2 , respectivamente. Atente-se à distribuição *a priori*.
2. Para $s = 2$ até S , faça:
 - (a) Gere uma nova amostra candidata \mathbf{x}^* , respeitando uma distribuição simétrica de proposta. Por exemplo, suponha que tenhamos escolhido uma distribuição de proposta uniforme, $J(\mathbf{x}^* | \mathbf{x}^{(s)}) = \text{Uniforme}(\mathbf{x}^{(s)} - \delta, \mathbf{x}^{(s)} + \delta)$. Isso significa que a amostra $\mathbf{x}^* \in [\mathbf{x}^{(s)} - \delta, \mathbf{x}^{(s)} + \delta]$, ou ainda, $x_1^* \in [x_1^{(s)} - \delta, x_1^{(s)} + \delta]$ e $x_2^* \in [x_2^{(s)} - \delta, x_2^{(s)} + \delta]$ com igual probabilidade.

(b) Calcule a taxa de aceitação r :

$$r = \frac{p(\mathbf{x}^* \mid y_1, \dots, y_{40})}{p(\mathbf{x}^{(s)} \mid y_1, \dots, y_{40})}$$

(c) Sorteie $u \sim \text{Uniforme}(0, 1)$.

(d) Se $u \leq \min(r, 1)$:

Aceite \mathbf{x}^* , isto é, faça $\mathbf{x}^{(s)} = \mathbf{x}^*$.

Caso contrário:

Rejeite \mathbf{x}^* , isto é, faça $\mathbf{x}^{(s)} = \mathbf{x}^{(s-1)}$.

(e) Faça $s = s + 1$.

3. Retorne $\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(s)}\}$.