Estatística Computacional

Norah Jones

2024-10-02

Índice

1 Introdução: Problemas de interesse

Encontrar soluções de equações não lineares onde não é possível obter uma solução analítica;

Obter integrais que apresentam uma forma complicada que inviabiliza encontrar uma solução analítica;

Gerar artificialmente amostras a partir de modelos estatísticos;

Aplicar a metodologia estudada na resolução de problemas de inferência.

2 Introduction

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

Parte I

Solução Numérica de Equações

Motivação - O Estimador de Máxima Verossimilhança

No que segue o termo densidade, significa ou uma densidade de probabilidade (caso absolutamente contínuo) ou uma função de probabilidade (caso discreto).

Sejam $X_1, \ldots, X_n \overset{iid}{\sim} f(.|\theta),\ \theta \in\ \Theta$, onde $f(.|\theta)$ é uma densidade, θ é um parâmetro que desejamos estimar e Θ é o espaço paramétrico;

Suponha que observamos os valores $x_1, \ldots, x_n$. A função de verossimilhança é definida por:

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta),\ \theta \in \Theta$$

A função de log-verossimilhança é dada por:

$$l(\theta) = logL(\theta) = \sum_{i=1}^{n} logf(x_i|\theta),\ \theta\ \in \Theta$$

Seja $\hat{\theta}\ \in \Theta$ um valor do parâmetro que maximiza a função de verossimilhança, ou seja, tal que

$$L(\hat{\theta})\ \geq L(\hat{\theta}),\ \text{para todo}\ \theta\ \in\ \Theta$$

Então dizemos que $\hat{\theta}$ é uma estimativa de máxima verossimilhança de θ.

A interpretação no caso discreto: é mais provável que $\hat{\theta}$ tenha gerado os dados $x_1, \ldots, x_n$

Como $\hat{\theta}$ depende da amostra, escrevemos $\hat{\theta}(x_1, \ldots, x_n)$. Neste caso, $\hat{\theta}(X_1, \ldots, X_n)$ é o estimador de máxima verossimilhança (EMV).

> **Nota**
>
> Para cada amostra observada $\mathbf{x} =\ (x_1, \ldots, x_n)$
> A definição nos diz que $\hat{\theta}(\mathbf{x})$ é um ponto de máximo global. Podemos ter nenhum ou mais de um máximo global
> Suponha que Θ é um intervalo e que o ponto $\hat{\theta}$ é um ponto interior de Θ que é ponto de máximo de L, podendo ser um máximo local. Se L tem derivada em $\hat{\theta}$, então $L'(\hat{\theta}) = 0$.
> Ou seja, $\hat{\theta}$ é um ponto estacionário de L (também dizemos $\hat{\theta}$ é um zero da função L'). Este resultado é conhecido no Cálculo como Teorema de Fermat para Pontos Estacionários.
> Ou seja, sob as condições acima, se $\hat{\theta}$ for EMV, então a derivada de L se anula neste ponto. A recíproca pode não ser verdadeira.

Assim, em muitos casos encontrados na prática, encontrar o EMV é um problema relacionado a encontrar soluções em θ para a equação $L'(\hat{\theta}) = 0$ ou $l'(\theta) = 0$.

Por exemplo, considere uma amostra aleatória $X_1, \ldots, X_n$ proveniente de uma distribuição $exp(\theta)$. Assim, cada X_i tem densidade de probabilidade.

$$f(x|\theta) = \begin{cases} \theta \exp(-\theta x), & \text{se} \quad x > 0 \\ 0, & \text{se} \quad x \leq 0. \end{cases}$$

O espaço paramétrico é $\Theta = (0, \infty)$. Temos então a seguinte função de log-verassimilhança:

$$\ell(\theta) = n \log \theta - \theta \sum_{i=1}^{n} x_i, \ \ \theta > 0,$$

implicando em

$$\ell'(\theta) = n/\theta - \sum_{i=1}^{n} x_i, \ \ \theta > 0.$$

A solução da equação $\ell'(\theta) = 0$ é dada por $\hat{\theta} = n / \sum_{i=1}^{n} x_i$

```
theta = 5
amostra = rexp(100, rate = theta);


loglik <- function(theta = NULL) {
    return(length(amostra)*log(theta) - theta*sum(amostra))
}

score_exp = function(theta = NULL) {
    n = length(amostra)

    return(n/theta - sum(amostra))
}

theta_x <- seq(from = 1, to = 10, by = 0.1)  # Sequência de valores de -10 a 10 com incremen

y <- loglik(theta = theta_x);

maximo_estimado = optimise(loglik, interval = c(1, 10), maximum = T)
y_score = score_exp(theta = theta_x)
```

```r
#Plotando a log-verossimilhança
plot(theta_x, y, type = "l", col = "blue", lwd = 2,
     main = paste("Gráfico da log-verossimilhança de: f(x) =", theta, "exp(-", theta, "x)"),
     xlab = "x", ylab = "f(x)")
abline(v=theta, col="blue")
abline(v=maximo_estimado, col="red")

#Plotando a função score
plot(theta_x, y_score, type = "l", col = "blue", lwd = 2,
     main = paste("Gráfico da funcao score de: f(x) =", theta, "exp(-", theta, "x)"),
     xlab = "x", ylab = "f(x)")
abline(h=0, col="red")
abline(v=maximo_estimado, col="red")
```

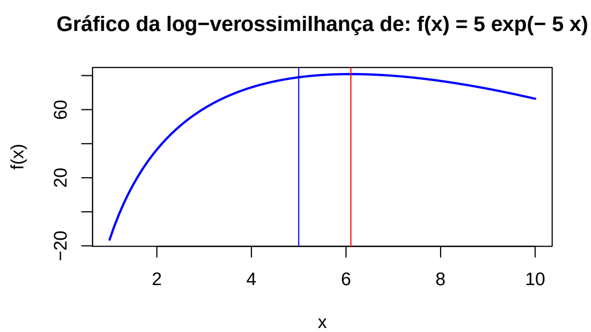

Figura 2.1: log-verossimilhança1

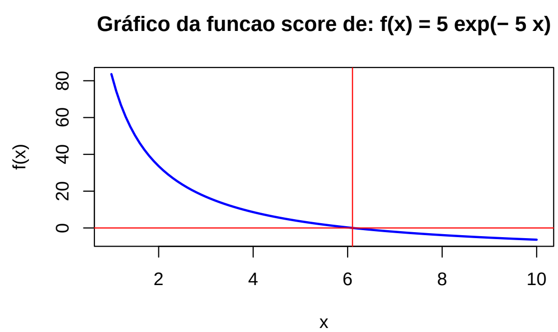

Figura 2.2: Função score

A função $S(\theta) = \ell'(\theta)$, $\theta \in \Theta$, é denominada função escore. Assim, em geral, encontrar o EMV é equivalente a resolver em θ a equação $S(\theta) = 0$.

No exemplo acima obtivemos uma solução analítica para esta equação. Mas em algumas situações isto não é possível.

Considere o seguinte exemplo (Bolfarine e Sandoval, 2010): sejam $X_1, \ldots, X_n \overset{iid}{\sim} f(.|\theta)$, onde

$$f(x|\theta) = \begin{cases} \frac{1}{2}(1+\theta x), & \text{se} \quad x > 0 \\ 0, & \text{se} \quad x \leq 0. \end{cases} \tag{2.1}$$

Onde $-1 < \theta < 1$ verifique que $f(\cdot|\theta)$ é uma densidade.

Então é possível mostrar que

$$S(\theta) = \sum_{i=1}^{n} \frac{x_i}{1 + \theta x_i}.$$

No entanto, não há solução analítica para $S(\theta) = 0$. Os gráficos a seguir mostram o comportamento das funções de log-verossimilhança (à esquerda) e escore (à direita). A amostra utilizada foi gerada artificialmente a partir do modelo (Equação 2.1), com $n = 40$ e $\theta = -0,35$ (como gerar esta amostra? voltaremos em breve a este assunto).

3 Método da Bisseção

Vamos expor a seguir métodos numéricos para encontrarmos pontos estacionários de uma função g, ou seja, pontos $x \in \mathbb{R}$ tais que $g'(x) = 0$. Primeiramente vamos considerar um exemplo simples (Givens & Hoeting, 2013). Seja

$$g(x) = \frac{\log x}{1+x}, \quad x > 0. \tag{3.1}$$

Então, para todo $x > 0$,

$$g'(x) = \frac{\frac{1+x}{x} - \log x}{(1+x)^2}. \tag{3.2}$$

A equação $g'(x) = 0$ implica em

$$\log x - \frac{1}{x} - 1 = 0,$$

Que não possui solução analítica.

3.1 Métodos Iterativos

Este tipo de problema pode ser resolvido usando o que chamamos de *Método iterativo*. Observemos o gráfico da função g:

```
library(ggplot2)
funcao_g <- function(x) {
    return(log(x)/(1+x))
}



x <- seq(0.01, 6, 0.01)
y <- funcao_g(x)
```

```
resultados <- data.frame(x = x, imagem = y);

ggplot(data = resultados, aes(x = x, y = imagem)) +
geom_line()
```

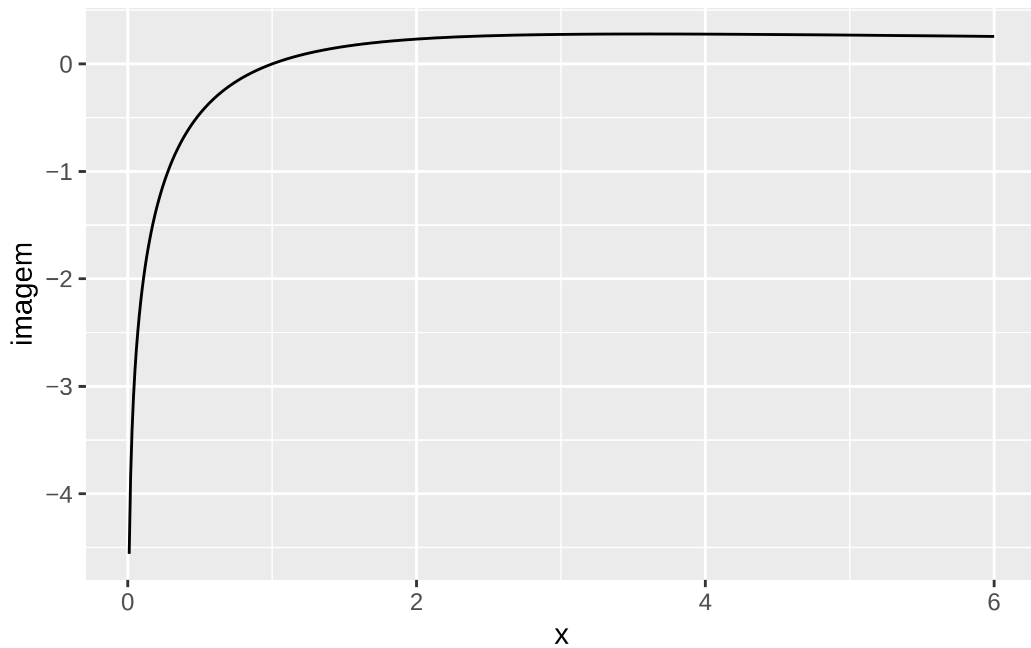

Figura 3.1: Gráfico da função g(x)

4 Summary

In summary, this book has no content whatsoever.

References

Knuth, Donald E. 1984. «Literate Programming». *Comput. J.* 27 (2): 97–111. https://doi.org/10.1093/comjnl/27.2.97.