Modelos Lineares Generalizados

Disciplina ofertada pelo DECAT/UFS
Código: ESTAT0092
Nível: Graduação
Carga horária: 60h
Período: 2020.2
Professor responsável e ministrante: Luiz Henrique Dore

Aulas 5 e 6: notas complementares

1 Derivada de primeira ordem da log-verossimilhança com relação a $\boldsymbol{\beta}$

Conforme visto no tópico 10, a função de log-verossimilhança é dada por

$$l(\boldsymbol{\beta}, \phi) = \phi^{-1} \left(\sum_{i=1}^{n} Y_i \theta_i - \sum_{i=1}^{n} b(\theta_i) \right) + \sum_{i=1}^{n} c(Y_i, \phi), \tag{1}$$

onde $\boldsymbol{\beta} = (\beta_0, \beta_1, \cdots, \beta_p)$. Vamos calcular $\frac{\partial l}{\partial \beta_j}$, a derivada de $l(\boldsymbol{\beta}, \phi)$ com relação ao j-ésimo componente de $\boldsymbol{\beta}$. Note que $b'(\theta_i) = \mu_i = g^{-1}(\eta_i) = g^{-1}(\boldsymbol{x}_i^T \boldsymbol{\beta})$. Logo, θ_i pode ser visto como a seguinte composição de funções: $\theta_i \equiv \theta_i(\mu_i(\eta_i(\beta_i)))$ [1, p. 71]. Pela regra da cadeia,

$$\frac{\partial \theta_i}{\partial \beta_l} = \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_j}.$$

Utilizando esse resultado e, novamente, a regra da cadeia, tem-se que

$$\frac{\partial l}{\partial \beta_j} = \frac{\partial l}{\partial \theta_i} \frac{\partial \theta_i}{\partial \beta_l} = \frac{\partial l}{\partial \theta_i} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_j}.$$

Portanto,

$$\frac{\partial l}{\partial \beta_j} = \phi^{-1} \sum_{i=1}^{n} Y_i \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_j} - \phi^{-1} \sum_{i=1}^{n} \frac{\partial b(\theta_i)}{\partial \theta_i} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_j}.$$

Note que

$$\frac{\partial \theta_i}{\partial \mu_i} = \frac{1}{\frac{\partial \mu_i}{\partial \theta_i}} = \frac{1}{V(\mu_i)} = V_i^{-1},$$

onde $V_i = V(\mu_i)$ é a função de variância (ver tópico 4). Além disso,

$$\frac{\partial b(\theta_i)}{\partial \theta_i} = b'(\theta_i) = \mu_i \quad \text{e} \quad \frac{\partial \eta_i}{\partial \beta_j} = x_{ij}$$

.

Logo, a derivada de $l(\boldsymbol{\beta}, \phi)$, com relação a β_j, é dada por [2, p. 21]

$$\begin{aligned} \frac{\partial l}{\partial \beta_j} &= \phi^{-1} \sum_{i=1}^{n} Y_i V_i^{-1} \frac{\partial \mu_i}{\partial \eta_i} x_{ij} - \phi^{-1} \sum_{i=1}^{n} \mu_i V_i^{-1} \frac{\partial \mu_i}{\partial \eta_i} x_{ij} = \\ &= \sum_{i=1}^{n} \phi^{-1} \sqrt{\frac{\omega_i}{V_i}} (y_i - \mu_i) x_{ij}, \end{aligned} \tag{2}$$

onde $\omega_i = (\partial \mu_i / \partial \eta_i)^2 / V_i$.

2 Derivada de primeira ordem da log-verossimilhança com relação a ϕ

Derivando a log-verossimilhança (1), com relação a ϕ, obtém-se [1, p. 114]

$$\frac{\partial l}{\partial \phi} = -\phi^{-2} \left(\sum_{i=1}^{n} Y_i \theta_i - \sum_{i=1}^{n} b(\theta_i) \right) + \sum_{i=1}^{n} \frac{\partial c(Y_i, \phi)}{\partial \phi}. \tag{3}$$

3 Derivadas de segunda ordem da log-verossimilhança com relação a $\boldsymbol{\beta}$

Conforme visto na seção 1, a derivada de $l(\boldsymbol{\beta}, \phi)$, com relação a β_j, pode ser escrita como

$$\frac{\partial l}{\partial \beta_j} = \phi^{-1} \sum_{i=1}^{n} Y_i \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} x_{ij} - \phi^{-1} \sum_{i=1}^{n} \mu_i \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} x_{ij}$$

Logo, a derivada parcial de $l(\boldsymbol{\beta}, \phi)$, com relação a β_j e β_l, é

$$\begin{aligned}\frac{\partial^2 l}{\partial \beta_j \beta_l} = \frac{\partial}{\partial \beta_l}\left(\frac{\partial l}{\partial \beta_j}\right) &= \phi^{-1} \sum_{i=1}^{n} Y_i \frac{\partial}{\partial \beta_l}\left(\frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i}\right) x_{ij} - \phi^{-1} \sum_{i=1}^{n} \frac{\partial}{\partial \beta_l}\left(\mu_i \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i}\right) x_{ij} = \\ &= \phi^{-1} \sum_{i=1}^{n} Y_i \left[\frac{\partial^2 \theta_i}{\partial \mu_i \partial \beta_l} \frac{\partial \mu_i}{\partial \eta_i} + \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial^2 \mu_i}{\partial \eta_i \beta_l}\right] x_{ij} - \\ &- \phi^{-1} \sum_{i=1}^{n} \left[\frac{\partial \mu_i}{\partial \beta_l} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} + \mu_i \frac{\partial^2 \theta_i}{\partial \mu_i \partial \beta_l} \frac{\partial \mu_i}{\partial \eta_i} + \mu_i \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial^2 \mu_i}{\partial \eta_i \beta_l}\right] x_{ij}. \end{aligned} \tag{4}$$

Aplicando a regra da cadeia, pode-se escrever

$$\begin{aligned}\frac{\partial^2 \theta_i}{\partial \mu_i \partial \beta_l} &= \frac{\partial}{\partial \beta_l}\left(\frac{\partial \theta_i}{\partial \mu_i}\right) = \frac{\partial}{\partial \mu_i}\left(\frac{\partial \theta_i}{\partial \mu_i}\right) \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_l} = \frac{\partial^2 \theta_i}{\partial \mu_i^2} \frac{\partial \mu_i}{\partial \eta_i} x_{il}, \\ \frac{\partial^2 \mu_i}{\partial \eta_i \partial \beta_l} &= \frac{\partial}{\partial \beta_l}\left(\frac{\partial \mu_i}{\partial \eta_i}\right) = \frac{\partial}{\partial \eta_i}\left(\frac{\partial \mu_i}{\partial \eta_i}\right) \frac{\partial \eta_i}{\partial \beta_l} = \frac{\partial^2 \mu_i}{\partial \eta_i^2} x_{il} \quad \text{e} \\ \frac{\partial \mu_i}{\partial \beta_l} &= \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_l} = \frac{\partial \mu_i}{\partial \eta_i} x_{il}. \end{aligned}$$

Substituindo essas expressões em (4), obtém-se [2, p. 21]

$$\begin{aligned}\frac{\partial^2 l}{\partial \beta_j \beta_l} &= \phi^{-1} \sum_{i=1}^{n} Y_i \left[\frac{\partial^2 \theta_i}{\partial \mu_i^2} \left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 x_{il} + \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial^2 \mu_i}{\partial \eta_i^2} x_{il}\right] x_{ij} - \\ &- \phi^{-1} \sum_{i=1}^{n} \left[\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 \frac{\partial \theta_i}{\partial \mu_i} x_{il} + \mu_i \frac{\partial^2 \theta_i}{\partial \mu_i^2} \left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 x_{il} + \mu_i \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial^2 \mu_i}{\partial \eta_i^2} x_{il}\right] x_{ij} = \\ &= \phi^{-1} \sum_{i=1}^{n} (Y_i - \mu_i) \frac{\partial^2 \theta_i}{\partial \mu_i^2} \left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 x_{ij} x_{il} + \\ &+ \phi^{-1} \sum_{i=1}^{n} (Y_i - \mu_i) \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial^2 \mu_i}{\partial \eta_i^2} x_{ij} x_{il} - \phi^{-1} \sum_{i=1}^{n} \left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 \frac{\partial \theta_i}{\partial \mu_i} x_{ij} x_{il}. \end{aligned} \tag{5}$$

4 Derivadas de segunda ordem da log-verossimilhança com relação a $\boldsymbol{\beta}$ e a ϕ

Derivando (2) com relação a ϕ, obtém-se [1, p. 114]

$$\frac{\partial^2 l}{\partial \beta_j \partial \phi} = -\phi^{-2} \sum_{i=1}^{n} \sqrt{\frac{\omega_i}{V_i}} (Y_i - \mu_i) x_{ij} \tag{6}$$

5 Derivadas de segunda ordem da log-verossimilhança com relação a ϕ

Derivando (3) com relação a ϕ, obtém-se

$$\frac{\partial^2 l}{\partial \phi^2} = 2\phi^{-3}\left(\sum_{i=1}^{n} Y_i\theta_i - \sum_{i=1}^{n} b(\theta_i)\right) + \sum_{i=1}^{n} \frac{\partial^2 c(Y_i,\phi)}{\partial \phi^2}. \tag{7}$$

6 Valor esperado das derivadas de segunda ordem da log-verossimilhança

Utilizando a fórmula (5), o valor esperado de $\frac{\partial^2 l}{\partial \beta_j \beta_l}$ fica dado por

$$\begin{aligned} E\left[\frac{\partial^2 l}{\partial \beta_j \beta_l}\right] &= \phi^{-1}\sum_{i=1}^{n} E(Y_i-\mu_i)\frac{\partial^2\theta_i}{\partial\mu_i^2}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2 x_{ij}x_{il} + \\ &+ \phi^{-1}\sum_{i=1}^{n} E(Y_i-\mu_i)\frac{\partial\theta_i}{\partial\mu_i}\frac{\partial^2\mu_i}{\partial\eta_i^2}x_{ij}x_{il} - \phi^{-1}\sum_{i=1}^{n}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2\frac{\partial\theta_i}{\partial\mu_i}x_{ij}x_{il} \\ &= -\phi^{-1}\sum_{i=1}^{n}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2\frac{\partial\theta_i}{\partial\mu_i}x_{ij}x_{il} = -\phi^{-1}\sum_{i=1}^{n}\omega_i x_{ij}x_{il}. \end{aligned} \tag{8}$$

Utilizando a fórmula (6), o valor esperado de $\frac{\partial^2 l}{\partial \beta_j \phi}$ fica dado por

$$E\left[\frac{\partial^2 l}{\partial \beta_j \partial \phi}\right] = -\phi^{-2}\sum_{i=1}^{n}\left(\sqrt{\frac{\omega_i}{V_i}}E(Y_i-\mu_i)x_{ij}\right) = 0. \tag{9}$$

Utilizando a fórmula (7), o valor esperado de $\frac{\partial^2 l}{\partial \phi^2}$ fica dado por

$$E\left[\frac{\partial^2 l}{\partial \phi^2}\right] = 2\phi^{-3}\left(\sum_{i=1}^{n} E(Y_i)\theta_i - \sum_{i=1}^{n} b(\theta_i)\right) + \sum_{i=1}^{n} E\left(\frac{\partial^2 c(Y_i,\phi)}{\partial \phi^2}\right) = \tag{10}$$

$$= 2\phi^{-3}\left(\sum_{i=1}^{n} \mu_i\theta_i - \sum_{i=1}^{n} b(\theta_i)\right) + \sum_{i=1}^{n} E\left(\frac{\partial^2 c(Y_i,\phi)}{\partial \phi^2}\right) \tag{11}$$

7 Valor esperado de U_ϕ

Conforme visto no tópico 11, U_ϕ é a derivada da função de log-verossimilhança com relação ao parâmetro ϕ. No tópico 10, viu-se que a função de log-verossimilhança é dada por

$$l(\boldsymbol{\beta}, \phi) = \ln\left[L(\boldsymbol{\beta}, \phi)\right] = \ln\left[\prod_{i=1}^{n} f(Y_i; \theta_i, \phi)\right] = \sum_{i=1}^{n} \ln\left[f(Y_i; \theta_i, \phi)\right].$$

Dessa forma, U_ϕ pode ser escrita como

$$U_\phi = \frac{\partial l}{\partial \phi} = \sum_{i=1}^{n} \frac{\partial \ln\left[f(Y_i; \theta_i, \phi)\right]}{\partial \phi} = \sum_{i=1}^{n} \frac{\partial f(Y_i; \theta_i, \phi)/\partial \phi}{f(Y_i; \theta_i, \phi)} = \sum_{i=1}^{n} U_{i,\phi},$$

onde $U_{i,\phi} = \frac{\partial f(Y_i;\theta_i,\phi)/\partial\phi}{f(Y_i;\theta_i,\phi)}$. Segue-se que

$$E\left[U_\phi\right] = \sum_{i=1}^{n} E[U_{i,\phi}]. \tag{12}$$

Agora,

$$E[U_{i,\phi}] = \int_{-\infty}^{\infty} U_{i,\phi} f(y_i; \theta_i, \phi) dy_i = \int_{-\infty}^{\infty} \frac{\partial f(y_i; \theta_i, \phi)/\partial \phi}{f(y_i; \theta_i, \phi)} \cdot f(y_i; \theta_i, \phi) dy_i = \int_{-\infty}^{\infty} \frac{\partial f(y_i; \theta_i, \phi)}{\partial \phi} dy_i.$$

Conforme mencionado no tópico 2, supõe-se que a família exponencial satisfaz algumas condições de regularidade. Uma dessas condições diz que o suporte da família exponencial não depende dos parâmetros. Essa condição garante que, quando aplicadas a $f(y_i; \theta, \phi)$, as operações de diferenciação com relação aos parâmetros e integração com relação a y_i são permutáveis. Isso significa que a última integral acima pode ser expressa como

$$\int_{-\infty}^{\infty} \frac{\partial f(y_i; \theta_i, \phi)}{\partial \phi} dy_i = \frac{\partial}{\partial \phi}\left(\int_{-\infty}^{\infty} f(y_i; \theta_i, \phi) dy_i\right).$$

Mas $\int_{-\infty}^{\infty} f(y_i; \theta_i, \phi) dy_i = 1$ é constante, pois $f(y_i; \theta_i, \phi)$ é uma função de densidade de probabilidade. Como a derivada de uma função constante é zero, segue-se que

$$E[U_{i,\phi}] = \int_{-\infty}^{\infty} \frac{\partial f(y_i; \theta_i, \phi)}{\partial \phi} dy_i = \frac{\partial}{\partial \phi}\left(\int_{-\infty}^{\infty} f(y_i; \theta_i, \phi) dy_i\right) = 0.$$

Substituindo em (12), obtém-se

$$E\left[U_\phi\right] = \sum_{i=1}^{n} E[U_{i,\phi}] = 0.$$

Referências

[1] G. M. Cordeiro e C. G. B. Demétrio, *Modelos lineares generalizados e extensões*, 2013, disponível em https://docs.ufpr.br/~taconeli/CE22518/LivClarice.pdf.

[2] G. A. Paula, *Modelos lineares generalizados com apoio computacional*, 2013, disponível em https://www.ime.usp.br/~giapaula/texto_2013.pdf.