

Exercício: regressão múltipla sob a abordagem bayesiana

Disciplina: Modelagem Estatística

Instrutor: Luiz Max Carvalho

Monitor: Isaque Pim

Abril/2023

Notação: Como convenção adotamos $\mathbb{R} = (-\infty, \infty)$, $\mathbb{R}_+ = (0, \infty)$ e $\mathbb{N} = \{1, 2, \dots\}$.

Motivação: Como vimos até aqui, o modelo linear (gaussiano) é extremamente útil para modelar a relação entre possíveis variáveis explanatórias (i.e. covariáveis) e uma variável dependente/resposta contínua. Até agora vimos como fazer inferência para esse modelo sob a ótica clássica/frequentista. Vamos então nos debruçar sobre o tratamento bayesiano do problema.

Tome \mathbf{X} uma matriz real $n \times P$ e $\mathbf{Y} = \{Y_1, \dots, Y_n\}^T \in \mathbb{R}^n$ um vetor contendo os valores da variável dependente.

Nosso modelo é

$$E[Y_i] =: \mu_i(\boldsymbol{\beta}) = \tilde{\mathbf{X}}_i^T \boldsymbol{\beta},$$

onde $\boldsymbol{\beta} \in \mathbb{R}^{P+1}$ é o vetor de coeficientes e parâmetro de interesse e $\tilde{\mathbf{X}}$ é uma matriz obtida adicionando uma coluna de uns, $\mathbf{X}_0 = \{1, \dots, 1\}^T$, a \mathbf{X} . Como antes, vamos assumir que os erros em torno do preditor linear são normalmente distribuídos com variância comum:

$$Y_i \stackrel{\text{i.i.d.}}{\sim} \text{Normal}(\mu_i(\boldsymbol{\beta}), \sigma^2),$$

com $\sigma^2 \in \mathbb{R}_+$ desconhecida. Do ponto de vista bayesiano, precisamos especificar uma distribuição de probabilidade conjunta para todas as quantidades desconhecidas θ do modelo. Neste caso, $\theta = (\boldsymbol{\beta}, \sigma^2) \in \mathbb{R}^{P+1} \times \mathbb{R}_+$.

Questões

1. Mostre que a verossimilhança $f_{\tilde{\mathbf{X}}}(\mathbf{y} \mid \theta)$ pode ser escrita na forma

$$f_{\tilde{\mathbf{X}}}(\mathbf{y} \mid \theta) = g_{\tilde{\mathbf{X}}}(\mathbf{y} \mid \boldsymbol{\beta}, \sigma^2) h_{\tilde{\mathbf{X}}}(\mathbf{y} \mid \sigma^2).$$

2. Utilize o resultado anterior para deduzir que a priori conjugada para este caso é da forma

$$\pi_{B,S}(\boldsymbol{\beta}, \sigma^2) = \pi_{B|S}(\boldsymbol{\beta} \mid \sigma^2) \pi_S(\sigma^2).$$

Em particular, mostre que

$$\pi_{B,S}(\boldsymbol{\beta}, \sigma^2) \propto \left(\frac{1}{\sigma^2} \right)^{a+(P+1)/2+1} \times \exp \left(-\frac{1}{\sigma^2} \left\{ b + \frac{1}{2} (\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)^T \mathbf{V}_\beta^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}_\beta) \right\} \right),$$

onde $\boldsymbol{\mu}_\beta \in \mathbb{R}^{P+1}$, \mathbf{V}_β é uma matriz positiva definida e $a, b \in \mathbb{R}_+$.

Dica: Que escolhas para $\pi_{B|S}$ e π_S eu preciso fazer?

3. A priori anterior chama-se **normal inversa gama** (NIG) e tem quatro parâmetros: \mathbf{m} , \mathbf{V} , a e b . Mostre que a posteriori de θ também é NIG e exiba seus hiperparâmetros.
4. **Distribuições marginais**

Dica: Antes de começar os cálculos para essa seção, vale considerar a seguinte representação do nosso modelo:

$$\begin{aligned} \mathbf{y} &= \tilde{\mathbf{X}}^T \boldsymbol{\beta} + \boldsymbol{\epsilon}_1, \text{ com } \boldsymbol{\epsilon}_1 \sim \text{MVN}_n(\mathbf{0}_n, \boldsymbol{\Sigma}_1), \\ \boldsymbol{\beta} &= \boldsymbol{\mu}_\beta + \boldsymbol{\epsilon}_2, \text{ com } \boldsymbol{\epsilon}_2 \sim \text{MVN}_{P+1}(\mathbf{0}_{P+1}, \boldsymbol{\Sigma}_2), \end{aligned}$$

onde $\boldsymbol{\epsilon}_1$ e $\boldsymbol{\epsilon}_2$ são erros independentes.

- (a) Determine $\boldsymbol{\Sigma}_1$ e $\boldsymbol{\Sigma}_2$;
- (b) Compute a verossimilhança marginal com respeito a σ^2 :

$$\tilde{f}_{\tilde{\mathbf{X}}}(\mathbf{y} \mid \sigma^2) := \int_{\mathbb{R}^{P+1}} f_{\tilde{\mathbf{X}}}(\mathbf{y} \mid \mathbf{b}, \sigma^2) \pi_{B|S}(\mathbf{b} \mid \sigma^2) d\mathbf{b}.$$

- (c) Usando o item anterior, compute a verossimilhança marginal ou *pre-ditiva a priori*:

$$m_{\tilde{\mathbf{X}}}(\mathbf{y}) := \int_0^\infty \tilde{f}_{\tilde{\mathbf{X}}}(\mathbf{y} \mid s) \pi_S(s) ds.$$

- (d) Mostre $\bar{f}_{\tilde{\mathbf{X}}}(\beta \mid \mathbf{y})$ e comente sobre como calcular, por exemplo, $\Pr(\beta_1 > a \mid \mathbf{y})$, para $a \in \mathbb{R}$.
5. Suponha que eu coletei uma nova matriz de desenho $m \times P$, \mathbf{X}' e quero prever o valor de \mathbf{y}' a partir do que eu aprendi usando \mathbf{X} e \mathbf{y} . Compute

$$\bar{p}_{\tilde{\mathbf{X}}, \mathbf{X}'}(\mathbf{y}' \mid \mathbf{y}) := \int_{\mathbb{R}^{P+1} \times \mathbb{R}_+} p_{\tilde{\mathbf{X}}}(\mathbf{b}, s \mid \mathbf{y}) f_{\mathbf{X}'}(\mathbf{y}' \mid \mathbf{b}, s) db ds,$$

e esboce o seu gráfico para uma observação (linha de \mathbf{X}') de um conjunto de dados da sua escolha.

Dica: use um conjunto de dados que você conheça bem. Bons exemplos são os bancos de 'peso ao nascer' e 'kid score', que já analisamos em sala.

Resultados úteis

Aqui estão enunciados alguns resultados úteis para o desenvolvimento das questões acima. Estes são dados sem demonstração, que você está convidada a fazer.

- **Completando o “quadrado” em múltiplas dimensões:** tome \mathbf{A} matriz simétrica positiva definida $d \times d$ e $\boldsymbol{\alpha}, \mathbf{u} \in \mathbb{R}^d$. Vale que:

$$\mathbf{u}^T \mathbf{A} \mathbf{u} - 2\boldsymbol{\alpha}^T \mathbf{u} = (\mathbf{u} - \mathbf{A}^{-1} \boldsymbol{\alpha})^T \mathbf{A} (\mathbf{u} - \mathbf{A}^{-1} \boldsymbol{\alpha}) - \boldsymbol{\alpha}^T \mathbf{A} \boldsymbol{\alpha}. \quad (1)$$

Dica: Expanda o produto e procure por cancelamentos de termos da forma $\mathbf{a}^T \mathbf{M}^{-1} \mathbf{a}$.

- **Sherman-Woodbury-Morrisson:** tome \mathbf{A} matriz quadrada $d \times d$ inversível, \mathbf{B} matriz $k \times d$, \mathbf{C} matriz $d \times k$ e \mathbf{D} matriz quadrada $k \times k$ inversível. Então

$$(\mathbf{A} + \mathbf{B} \mathbf{D} \mathbf{C})^{-1} = \mathbf{A}^{-1} - (\mathbf{D}^{-1} + \mathbf{C} \mathbf{A}^{-1} \mathbf{B})^{-1} \mathbf{C} \mathbf{A}^{-1}.$$

- **Determinantes:** tome \mathbf{A} , \mathbf{B} , \mathbf{C} e \mathbf{D} como antes. Então,

$$\det(\mathbf{A} + \mathbf{B} \mathbf{D} \mathbf{C}) = \det(\mathbf{A}) \det(\mathbf{D}) \det(\mathbf{D}^{-1} + \mathbf{C} \mathbf{A}^{-1} \mathbf{B}).$$