Universidade do Estado de Santa Catarina

Cálculo Numérico

Exercícios para a Prova 2

Fernando Deeke Sasse

28 de setembro de 2025

Resolva os exercícios com o auxílio de calculadora. Use ao menos 4 casas decimais nas respostas finais.

1. Dados os pontos na Tabela 2, determine, passo a passo, o spline cúbico do tipo *clamped*, com derivadas em $x = 1$ e $x = 4$ dadas por 1 e 0, respectivamente.

x	1	2	3	4
f	1.8	2.6	4.5	5.7

Tabela 1: Dados para spline cúbico.

Resposta. Denotando os polinômios em cada intervalo por

$$p_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i\,, \quad i = 1, \ldots 12\,,$$

as equações que determinam os coeficientes são

$$\begin{cases} a_1 + b_1 + c_1 + d_1 = 1.8 \\ 8a_1 + 4b_1 + 2c_1 + d_1 = 2.6 \\ 8a_2 + 4b_2 + 2c_2 + d_2 = 2.6 \\ 27a_2 + 9b_2 + 3c_2 + d_2 = 4.5 \\ \ldots \end{cases}$$

Os polinômios cúbicos interpolantes resultantes são (não necessários numa prova manual):

$$\begin{cases} 0.8 + 1.0x - 0.8733\,(x-1)^2 + 0.6733\,(x-1)^3 & x < 2 \\ 0.05333 + 1.2733x + 1.146\,(x-2)^2 - 0.52\,(x-2)^3 & x < 3 \\ -1.52 + 2.0067x - 0.4133\,(x-3)^2 - 0.3933\,(x-3)^3 & x > 3 \end{cases}$$

2. Considerando os dados na Tabela 2, use mínimos quadrados, passo a passo, para ajustar a função $f(x) = a\mathrm{e}^x + bx$. Calcule também o coeficiente de determinação.

x	1	2	3	4	5
y	6.7	12.6	21.2	38.9	62.1

Tabela 2: Dados para ajuste por mínimos quadrados.

Resposta: $a = 0.217256$, $b = 6.10986$, $r^2 = 0.9946$.

Notas: A *soma do quadrado dos resíduos* é dada por

$$\text{SSE} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \,,$$

sendo y_i os valores tabelados e $\hat{y}_i$ os valores do modelo (previstos). A *soma total total dos quadrados* é definida por

$$\text{SST} = \sum_{i=1}^{n} (y_i - \bar{y}_i)^2 \,,$$

sendo $\bar{y}$ a média de y_i. O *coeficiente de determinação* é dado por

$$r^2 = 1 - \frac{\text{SSE}}{\text{SST}} \,.$$

Notemos que:

- $r^2 = 1$: o modelo explica toda a variabilidade nos dados de resposta.
- $r^2 = 0$: o modelo não explica nenhuma variabilidade; não é melhor do que prever a média.
- $0 < r^2 < 1$: a proporção da variabilidade explicada pelo modelo.

Por exemplo, se $r^2 = 0.81$, os resultados indicam que 81% da incerteza original foi explicada por esse modelo. Na regressão linear simples (um preditor), r^2 também é o quadrado do coeficiente de correlação de Pearson entre x e y. A quantidade r é denominada *coefficiente de correlação.*

3. Sejam os dados mostrados na Tabela 4

x	1	2	3	4
y	2.3	5.4	8.1	17.2

Tabela 3: Dados gradientes descendentes em regressão linear univariada.

Use o método dos gradientes descendentes e construa método iterativo que determina os coeficientes de uma reta $y = \theta_0 + \theta_1 x$ de ajuste. Comece o processo iterativo com $(\theta_0^{(0)}, \theta_1^{(0)}) = (0, 0)$. Faça três iterações. Use ao menos 4 dígitos decimais. Calcule $\epsilon_0 = |\theta_0^{(j+1)} - \theta_0^{(j)}|$ e $\epsilon_1 = |\theta_1^{(j+1)} - \theta_1^{(j)}|$ em cada iteração. Escolha o valor parâmetro de taxa de aprendizado como $\alpha = 0.1$.

Resposta: $\theta^{(3)} = [\theta_0^{(3)}\ \theta_1^{(3)}]^T = [0.90375, 3.1125]^T$.

4. Sejam os dados mostrados na Tabela 4

x_1	1.1	2.6	3.2	4.6
x_2	4.1	6.5	8.2	8.8
y	2.3	6.4	8.1	17.2

Tabela 4: Dados gradientes descendentes em regressão linear multivariada.

Use o método dos gradientes descendentes e construa método iterativo que determina os coeficientes de um plano $y = \theta_0 + \theta_1 x_1 + \theta_2 x_2$ de ajuste. Comece o processo iterativo com $(\theta_0^{(0)}, \theta_1^{(0)}, \theta_2^{(0)}) = (0, 0, 0)$. Faça três iterações. Use ao menos 4 dígitos decimais. Escolha o valor parâmetro de taxa de aprendizado como $\alpha = 0.05$.

Resposta: $\theta^{(3)} = [\theta_0^{(3)}\, \theta_1^{(3)}\, \theta_2^{(3)}]^T = [1.02323453, 1.4351872, 1.58913611]^T$.

Nota. Regressão Linear e Gradientes Descendentes. Nosso objetivo é estabelecer a relação entre uma variável dependente escalar y e uma variável independente x da forma:

$$y \approx \hat{y} = \theta_0 + \theta_1 x\,.$$

Devemos encontrar os parâmetros θ_0 e θ_1 que minimizem a discrepância entre os valores previstos $\hat{y}$ e os valores observados y. Essa discrepância é quantificada pela função de custo (Erro Quadrático Médio):

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(\hat{y}^{(i)} - y^{(i)}\right)^2 = \frac{1}{2m} \sum_{i=1}^{m} \left(\theta_0 + \theta_1 x^{(i)} - y^{(i)}\right)^2\,.$$

Na expressão acima $y^n(i)$ e $x^n(i)$, $i = 1, \ldots m$ denotam os valores dados na Tabela 4. O método dos gradientes descendentes é um algoritmo de otimização iterativa para minimizar a função de custo $J(\theta_0, \theta_1)$. Os parâmetros são atualizados da seguinte forma:

$$\theta_j^{(i+1)} := \theta_j^{(i)} - \alpha \left.\frac{\partial J}{\partial \theta_j}\right|_{\theta_j^{(i)}}, \quad j = 0, 1\,,$$

sendo α a taxa de aprendizado (tomando como positivo, pois buscamos um mínimo).

As derivadas parciais são:

$$\frac{\partial J}{\partial \theta_0} = \frac{1}{m} \sum_{i=1}^{m} \left(\hat{y}^{(i)} - y^{(i)}\right)$$

$$\frac{\partial J}{\partial \theta_1} = \frac{1}{m} \sum_{i=1}^{m} \left(\hat{y}^{(i)} - y^{(i)}\right) x^{(i)}$$

De modo mais compacto, para fins computacionais, podemos reescrever:

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2$$

onde $y^{(i)}$ são os valores observados e $h_\theta(x)$ é a hipótese, dada pelo modelo linear.

$$h_\theta(x) = \theta^T \mathbf{x} = \theta_0 + \theta_1 x$$

sendo

$$\theta^T = [\theta_0 \ \theta_1], \qquad \mathbf{x}^T = [1, x].$$

O método descrito acima pode ser estendido para casos de regressão linear multivariada. Nessas situações, desejamos modelar a relação entre uma variável dependente escalar y e um vetor de variáveis independentes $\mathbf{x} \in \mathbb{R}^n$ por meio de um modelo linear:

$$y \approx \hat{y} = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n.$$

Nosso objetivo é encontrar os parâmetros $\theta_0, \theta_1, \ldots, \theta_n$ que minimizem a discrepância entre os valores previstos $\hat{y}$ e os valores observados y. Assim como no caso univariado, essa discrepância é quantificada pela função de custo (Erro Quadrático Médio):

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(\mathbf{x}^{(i)}) - y^{(i)} \right)^2,$$

onde m é o número de exemplos, $y^{(i)}$ são os valores observados e $h_\theta(\mathbf{x})$ é a hipótese do modelo, dada por

$$h_\theta(\mathbf{x}) = \theta^T \mathbf{x} = \theta_0 + \theta_1 x_1 + \cdots + \theta_n x_n.$$

Aqui utilizamos a convenção de incluir um termo adicional $x_0 = 1$ em cada vetor de entrada, de forma que:

$$\theta^T = [\theta_0 \ \theta_1 \ \cdots \ \theta_n], \qquad \mathbf{x}^T = [1 \ x_1 \ \cdots \ x_n].$$

O método dos gradientes descendentes consiste em um algoritmo iterativo para minimizar a função de custo. A atualização dos parâmetros é dada por:

$$\theta_j := \theta_j - \alpha \frac{\partial J}{\partial \theta_j}, \qquad j = 0, 1, \ldots, n,$$

onde $\alpha > 0$ é a taxa de aprendizado. As derivadas parciais assumem a forma geral:

$$\frac{\partial J}{\partial \theta_j} = \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(\mathbf{x}^{(i)}) - y^{(i)} \right) x_j^{(i)}.$$

Assim, cada parâmetro θ_j é ajustado na direção negativa do gradiente, levando em consideração sua contribuição relativa por meio da variável x_j.

Para facilitar implementações computacionais, podemos escrever as atualizações de forma vetorial. Definindo:

- $X \in \mathbb{R}^{m \times (n+1)}$: matriz de dados, em que cada linha é $(x_0^{(i)}, x_1^{(i)}, \ldots, x_n^{(i)})$ com $x_0^{(i)} = 1$;
- $\mathbf{y} \in \mathbb{R}^m$: vetor dos valores observados;
- $\theta \in \mathbb{R}^{n+1}$: vetor de parâmetros.

A hipótese pode ser escrita como:

$$\hat{\mathbf{y}} = X\theta,$$

e o gradiente da função de custo é:

$$\nabla_\theta J(\theta) = \frac{1}{m} X^T (X\theta - \mathbf{y}).$$

O processo iterativo de gradientes descendentes é então expresso como:

$$\theta := \theta - \alpha \nabla_\theta J(\theta) = \theta - \frac{\alpha}{m} X^T (X\theta - \mathbf{y}).$$