Lista de Exercícios
Regressão Logística

1. Considere o conjunto de dados abaixo, com duas variáveis explicativas e uma variável resposta binária:

i	x_{1i}	x_{2i}	y_i
1	0	0	0
2	1	0	0
3	0	1	0
4	1	1	1

Considere o modelo com intercepto:

$$p_i = \mathbb{P}(y_i = 1 \mid \boldsymbol{x}_i) = \sigma(\boldsymbol{w}^\top \tilde{\boldsymbol{x}}_i), \qquad \sigma(z) = \frac{1}{1+e^{-z}},$$

onde $\tilde{\boldsymbol{x}}_i = (1, x_{1i}, x_{2i})^\top$.

(a) Monte a matriz X (com coluna de 1's) e o vetor $\boldsymbol{y}$.

(b) Escreva a função custo (log-verossimilhança negativa):

$$J(\boldsymbol{w}) = -\sum_{i=1}^{n} \Big(y_i \log p_i + (1-y_i)\log(1-p_i) \Big).$$

(c) Mostre que o gradiente de $J(\boldsymbol{w})$ pode ser escrito como

$$\nabla J(\boldsymbol{w}) = X^\top (\boldsymbol{p} - \boldsymbol{y}),$$

onde $\boldsymbol{p} = (p_1, \ldots, p_n)^\top$.

(d) Considere $\boldsymbol{w}^{(0)} = \boldsymbol{0}$. Calcule explicitamente os valores p_i, o gradiente $\nabla J(\boldsymbol{w}^{(0)})$ e o custo $J(\boldsymbol{w}^{(0)})$.

(e) Execute **duas iterações manuais** de descida do gradiente:

$$\boldsymbol{w}^{(t+1)} = \boldsymbol{w}^{(t)} - \alpha \nabla J(\boldsymbol{w}^{(t)}),$$

usando passo de aprendizado $\alpha = 0.5$.

(f) Para cada iteração, calcule $J(\boldsymbol{w}^{(t)})$ e verifique se o custo diminui.

2. Considere agora o conjunto de dados

$$(x_1, x_2, y) \in \{(0,0,0), (2,0,0), (0,2,0), (2,2,1)\}.$$

(a) Repita os itens (a)–(f) da questão anterior, usando o mesmo passo α.

(b) Compare o comportamento da descida do gradiente nos dois conjuntos de dados.

(c) Explique a influência da escala das variáveis explicativas na convergência do método.

3. Formule o ajuste da regressão logística como um problema de otimização irrestrita. Explique por que a solução não admite forma fechada, ao contrário do caso de mínimos quadrados.

4. Mostre que a Hessiana da função custo pode ser escrita como

$$\nabla^2 J(\boldsymbol{w}) = X^\top W X, \qquad W = \mathrm{diag}\big(p_i(1-p_i)\big).$$

(a) Mostre que W é semidefinida positiva.

(b) Conclua que $J(\boldsymbol{w})$ é uma função convexa.

(c) Explique a consequência dessa propriedade para a convergência da descida do gradiente.

5. Sob quais condições a solução do problema de regressão logística é única? Discuta o papel do posto da matriz X e do comportamento dos termos $p_i(1-p_i)$.

6. Explique o conceito de separação perfeita em regressão logística. Mostre, conceitualmente, por que nesse caso a função custo não possui minimizador finito. Relacione esse fenômeno com o comportamento da descida do gradiente.

7. Considere o problema regularizado:

$$\min_{\boldsymbol{w}} \; J(\boldsymbol{w}) + \frac{\lambda}{2}\|\boldsymbol{w}\|^2.$$

(a) Derive o gradiente da função custo regularizada.

(b) Explique como a regularização afeta a convergência da descida do gradiente.

(c) Discuta o papel da regularização na presença de separação perfeita.

8. Mostre que

$$\log \frac{p(\boldsymbol{x})}{1-p(\boldsymbol{x})} = \boldsymbol{w}^\top \tilde{\boldsymbol{x}}.$$

Explique a interpretação dos coeficientes do modelo em termos de:

(a) log-odds;
(b) odds ratio.

9. Defina o vetor $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{p}$.
   (a) Mostre como $\boldsymbol{r}$ aparece na expressão do gradiente.
   (b) Compare a interpretação desse resíduo com o resíduo da regressão linear por mínimos quadrados.
   (c) Explique por que não existe, no caso logístico, uma interpretação direta em termos de projeção ortogonal.