

Lista de Exercícios

Regressão Linear

1. Regressão Linear por Mínimos Quadrados

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

- (a) Monte X (com coluna de 1's) e \mathbf{y} .
 - (b) Calcule $X^\top X$ e $X^\top \mathbf{y}$.
 - (c) Inverta $X^\top X$ via **Gauss–Jordan**. Documente cada operação de linha.
 - (d) Inverta novamente pelo **método da adjunta**: cofatores, adjunta e determinante.
 - (e) Obtenha os pesos $\hat{\mathbf{w}}$ multiplicando a inversa por $X^\top \mathbf{y}$.
 - (f) Calcule o erro quadrático do modelo com os pesos calculados.
2. Dados $(x_1, x_2, y) \in \{(0, 0, 2), (2, 0, 3), (0, 2, 3), (2, 2, 6)\}$, repita os itens 1–6 acima e verifique se os dois métodos de inversão produzem o mesmo $\hat{\mathbf{w}}$.
 3. Regressão linear com descida do gradiente
 - (a) Derive o gradiente de $J(\mathbf{w}) = \frac{1}{2} \|X\mathbf{w} - \mathbf{y}\|^2$ e execute três iterações manuais de descida do gradiente com $\alpha = 0.1$ no primeiro conjunto da questão anterior.
 - (b) Para cada iteração, calcule o erro quadrado total e compare com o erro quadrado do método analítico. Comente a taxa de convergência.
 4. Formule o problema de regressão linear por mínimos quadrados como um problema de otimização. Explique o significado geométrico da solução obtida.
 5. Derive as equações normais associadas ao problema

$$\min_{\mathbf{w}} \|X\mathbf{w} - \mathbf{y}\|^2$$

e explique sob quais condições elas admitem uma solução única.

6. Mostre que a matriz $X^\top X$ é sempre simétrica e semidefinida positiva. Em que situação ela é definida positiva?
7. Explique a interpretação geométrica do vetor de resíduos

$$\mathbf{r} = \mathbf{y} - X\hat{\mathbf{w}}$$

e sua relação com o espaço coluna de X .

8. Explique por que a função custo

$$J(\mathbf{w}) = \frac{1}{2} \|X\mathbf{w} - \mathbf{y}\|^2$$

é convexa. Qual é a consequência dessa propriedade para a otimização por descida do gradiente?

9. Compare conceitualmente a solução analítica por mínimos quadrados com a solução obtida por descida do gradiente, discutindo vantagens e limitações de cada abordagem.
10. Discuta a influência do passo de aprendizado α na convergência da descida do gradiente. O que pode ocorrer se α for muito grande ou muito pequeno?