

Lista de Exercícios referente ao Cap. 1 do livro do Bishop

Embora a lista de exercícios seja extensa, envolvendo ao todo 22 exercícios (19 + 3 extras), eles são bem simples na maioria das vezes. Contudo, são importantes para revisar e sedimentar conhecimentos de Álgebra Linear, Variáveis Aleatórias e valores esperados, estimador ML (MLE), teoria da decisão e teoria da informação.

Exercícios do livro texto:

- 01) 1.1
- 02) 1.2
- 03) 1.5
- 04) 1.6
- 05) 1.7
- 06) 1.8
- 07) 1.9
- 08) 1.10
- 09) 1.11
- 10) 1.13
- 11) 1.21
- 12) 1.22
- 13) 1.23
- 14) 1.25 (veja Apêndice D, em especial a equação (D-8))
- 15) 1.31
- 16) 1.33
- 17) 1.37
- 18) 1.39
- 19) 1.41

Exercícios Extra:

E1) Considere a função custo/objetivo dada na equação (1.2) em que \mathbf{y} está definido na equação (1.1). Mostre que (1.2) pode ser escrito como

$$E(\mathbf{w}) = \frac{1}{2} \|\mathbf{y} - \mathbf{t}\|_2^2$$

onde

$$\mathbf{y} = [y(x_1, \mathbf{w}) \dots y(x_N, \mathbf{w})]^T$$

$$\mathbf{t} = [t_1 \dots t_N]^T$$

$$\mathbf{y} = \mathbf{A} \mathbf{w}$$

Determine as dimensões e os elementos (também chamados de entradas) da matriz \mathbf{A} . Mostre que o vetor de coeficientes que minimiza esta função objetivo pode ser escrito como

$$\mathbf{w}^* = \mathbf{A}^\dagger \mathbf{t} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{t}$$

Compare com o exercício 1.1 e note a vantagem de se usar Álgebra Linear para trabalhar com uma notação mais compacta.

E2) Mesma ideia de E1, porém agora considerando a função objetivo dada na equação (1.4) do livro. Escrevê-la de forma matricial. Encontre o vetor de coeficientes ótimos (em fórmula fechada).

E3) (Exercício Computacional)

Replique o experimento computacional denominado “Polynomial Curve Fitting” usado diversas vezes no livro texto (veja páginas 4 e 5 do livro, bem como Apêndice A).

Faça:

(a) Replique os resultados da Figura 1.4 e da Figura 1.6 para validar seu código (i.e., ter certeza de que ele está funcionando adequadamente);

(b) Simule uma base de dados que não tenha relevância estatística, isto é, que seja uma amostra que NAO representa bem o todo (a população). Verifique alguns resultados experimentais para compreender a importância de ter uma amostra relevante. Explique qual a relação entre o caso simulado e casos práticos envolvendo vetores de dimensão elevada.

Dica: Para a simulação, ao invés de pegar dados igualmente espaçados no intervalo $[0,1]$, você pode forçar com que seus dados sejam amostrados apenas do semiciclo positivo (ou apenas do negativo) do modelo gerador.

(c) Simule uma base de dados em que 1 dos dados seja *outlier*. O que ocorre com a curva vermelha, estimativa da curva verde (modelo gerador), neste caso?

Dica: Para a simulação, você pode gerar seus dados de treinamento normalmente, igual feito no item (a), e ao final do processo escolher 1 desses dados pra atribuir um novo valor de *target* que seja completamente “maluco” (por exemplo, $target = 10$).