1o LISTA DE EXERCÍCIOS - CAP 5

Para obter arquivos de dados, visite:

https://github.com/jermwatt/machine_learning_refined

CAP 5

1) Considere a regressão linear com N=1, ou seja, determinar a melhor reta, no sentido de mínimos quadrados, que aproxima um conjunto de pontos no plano. Mais especificamente, dados um conjuntos de pontos no plano $D = \{(x_i, y_i),\ i = 1, 2, ..., n\}$ onde x[i] e y[i] representam os vetores de pontos, buscamos determinar *a* e *b* de forma a minimizar o valor de

$$\sum_{i=1}^{n} (ax[i] + b - y[i])^2$$

a) Reescreva o problema acima, na formulação

$$\min_{x} \| Ax - c \|_2$$

Onde *A* é a matriz n x 2 e *c* o vetor dimensão n.

b) Aplique a equação normal ao problema de mínimos quadrados para verificar que este problema corresponde a resolver o sistema 2x2 abaixo:

$$\begin{bmatrix} \sum_{i=1}^{n} x[i]^2 & \sum_{i=1}^{n} x[i] \\ \sum_{i=1}^{n} x[i] & n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} x[i].y[i] \\ \sum_{i=1}^{n} y[i] \end{bmatrix}$$

2) Considere

$$A = \begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 2 & 1 \\ -2 & 1 \end{pmatrix} \quad \text{e} \quad b = \begin{pmatrix} 3 \\ -3 \\ -3 \\ -3 \end{pmatrix}$$

Calcule $\min \| Ax - b \|_2$ usando equações normais.

3) Usando a definição de convexidade [Bortolossi, 2002]:

Definição 11.6 (Funções convexas e côncavas)

(a) Dizemos que uma função $f\colon D \subset \mathbb{R} \to \mathbb{R}$ é *convexa* (ou *côncava para cima*) em um intervalo $I \subset D$ se, e somente se,

$$f((1-t)\cdot p + t\cdot q) \leq (1-t)\cdot f(p) + t\cdot f(q), \tag{11.3}$$

para todo $p, q \in I$ e todo $t \in [0,1]$.

Verifique que *f:R->R* definida por $f(x) = x^2$ é uma função convexa.

4) Usando o teorema abaixo [Bortolossi 2002]:

Teorema 11.14 Seja $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$ uma função de classe C^2 definida em um subconjunto convexo e aberto U de $\mathbb{R}^n$.

(a) f é uma função convexa em U se, e somente se, a matriz hessiana $D^2 f(\mathbf{p})$ é positiva semidefinida *para todo* $\mathbf{p} \in U$.

(b) f é uma função côncava em U se, e somente se, a matriz hessiana $D^2 f(\mathbf{p})$ é negativa semidefinida *para todo* $\mathbf{p} \in U$.

Verifique se as funções abaixo são convexas:

a) $f\colon \mathbb{R}^2 \to \mathbb{R}$, definida por $f(x,y) = x^2+y^2$

b) $f\colon \mathbb{R}^2 \to \mathbb{R}$, definida por $f(x,y) = x^4+y^4$

c) $f\colon \mathbb{R}^2 \to \mathbb{R}$, definida por $f(x,y) = x^2+xy$

5) Visualize os gráficos das funções do exercício 4, usando o código:

```
import numpy as np
import matplotlib.pyplot as plt

def f(x,y):
    return x*x-x*y

x = np.linspace(-8,8,30)
y = np.linspace(-8,8,30)

X,Y = np.meshgrid(x,y)
Z = f(X,Y)
fig = plt.figure()
ax = plt.axes(projection='3d')
#ax.plot_wireframe(X, Y, Z, color='black')
ax.plot_surface(X, Y, Z, rstride=1, cstride=1,
                cmap='viridis', edgecolor='none')
```

6) Partindo dos dados no arquivo 3d_linregress_data.csv conforme abaixo:

```
# load in data
csvname = datapath + '3d_linregress_data.csv'
data = np.loadtxt(csvname,delimiter=',')
x = data[:-1,:]
y = data[-1:,:]

print(np.shape(x))
print(np.shape(y))
```

a) Faça a regressão $y = w_0 + x_1 w_1 + x_2 w_2$ determinando w_0 , w_1 e w_2.
b) Exiba o valor de da função custo obtido após a otimização.
c) Calcule o MSE e MAD.

7) Resolver os exercícios 5.2, 5.6, 5.8, 5.11

Exercícios a serem entregues: 1) ,3) ,6) e 7) apenas o 5.8

Entrega da Lista: 11/Setembro