SME0305 - 2014

Gustavo Carlos Buscaglia

ICMC - Ramal 738176
gustavo.buscaglia@gmail.com

Prova 7 (5 de maio de 2014)

---

1. Considere o problema de predizer o rendimento de uma aluno universitário no seu segundo ano a partir do rendimento no primeiro ano. Especificamente, seja x o número de notas altas (> 7) que o aluno recebe no primeiro ano, e y o número de notas altas no segundo ano.

   Seja o seguinte "conjunto de treino", correspondente a um conjunto de $m = 4$ alunos que já completaram segundo ano:

x	y
3	4
2	2
1	3
0	2

   (a) (1 pt) Lembrando que nossa definição de função custo é

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

   e que a hipótese usada em regressão linear é

$$h_\theta(x) = \theta_0 + \theta_1\, x$$

   quanto vale $J(0, 1)$?

$$J(0,1) \quad =$$

   (b) (2 pt) Qual é a matriz do sistema linear sobredeterminado equivalente a minimizar a função J? Marque apenas uma das alternativas:

$$(i) \begin{pmatrix} 1 & 3 & 4 \\ 1 & 2 & 2 \\ 1 & 1 & 3 \\ 1 & 0 & 2 \end{pmatrix} \qquad (ii) \begin{pmatrix} 1 & 3 \\ 1 & 2 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$$

$$(iii) \begin{pmatrix} 3 & 4 \\ 2 & 2 \\ 1 & 3 \\ 0 & 2 \end{pmatrix} \qquad (iv) \begin{pmatrix} 1 & 3 \\ 1 & 2 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}^T \begin{pmatrix} 1 & 3 \\ 1 & 2 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$$

   (c) (3 pt) Calcule as derivadas segundas

$$\frac{\partial^2 J}{\partial \theta_0^2} \quad =$$

$$\frac{\partial^2 J}{\partial \theta_0 \, \partial \theta_1} \quad =$$

$$\frac{\partial^2 J}{\partial \theta_1^2} \quad =$$

2. (2 pt) Suponha que, para um problema de regressão linear (como por exemplo o exercício anterior, foram achados valores θ_0^* e θ_1^* tais que $J(\theta_0^*, \theta_1^*) = 0$. Quais das seguintes afirmações são, consequentemente, verdadeiras e quais falsas? (Responda V ou F **à esquerda de cada item**)

   (a) O ponto (θ_0^*, θ_1^*) do plano $\theta_0 - \theta_1$ é um mínimo local de J.

   (b) O fato $J(\theta_0^*, \theta_1^*) = 0$ só pode acontecer se $\theta_0^* = \theta_1^* = 0$.

   (c) O "conjunto de treino" (training set) pode ser ajustado exatamente por uma linha reta.

   (d) O ponto (θ_0^*, θ_1^*) não é necessariamente um mínimo. O que dá para afirmar é que, se não for mínimo, então a função J será negativa (menor que zero) no mínimo verdadeiro.

3. (2 pt) Seja f uma função tal que $f(\theta_0, \theta_1)$ devolve um número real. Suponha que é utilizado o método do gradiente

$$\underline{\theta}^{new} = \underline{\theta}^{old} - \alpha \, \nabla f(\underline{\theta}^{old}), \qquad \text{com} \quad \alpha > 0$$

   para tentar minimizar $f(\theta_0, \theta_1)$. Responda verdadeiro ou falso **à esquerda de cada item**:

   (a) Se θ_0 e θ_1 são inicializados iguais, eles se manterão iguais ao longo das iterações.

   (b) Se a condição inicial é um mínimo *global* de f, então as iterações não mudarão os valores de θ_0 e θ_1.

   (c) Se a condição inicial é um mínimo *local* de f, então as iterações não mudarão os valores de θ_0 e θ_1.

   (d) Se α é suficientemente pequeno, sempre acontecerá que $f(\underline{\theta}^{new}) \leq f(\underline{\theta}^{old})$.

   (e) Se as primeiras iterações do método do gradiente fazem a função f **crescer**, então é muito provável que a "taxa de aprendizado" α seja muito grande.