

1. Questões sobre regressão linear e classificação:

- O que é regressão linear e classificação? Qual seria a saída desejada para cada um destes métodos? Dê um exemplo de regressão e outro de classificação.
- Explique o método do descenso do gradiente e apresente sua fórmula (algoritmo) para o caso de um modelo ou hipótese linear $h_{\theta}(x) = \theta_0 + \theta_1 x$. O algoritmo sempre converge para o mínimo global neste caso? Considere a função de custo quadrática: $J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$ (OBS: você terá que calcular derivadas parciais).
- Qual a função da taxa de aprendizagem α no item anterior? O que acontece quando α é muito pequeno e também quando é muito grande? Explique se é necessário diminuir a taxa de aprendizagem ao longo do tempo.
- Explique para que serve e como se obtém a equação normal $\theta = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$? O que é e como se constrói a matriz \mathbf{X} e o vetor \mathbf{y} ? Cite 2 vantagens de se utilizar a equação normal.
- Quando usar regressão polinomial ao invés de regressão linear?
- Há ótimos locais para a regressão polinomial?
- Em quais casos é recomendado usar redes neurais ao invés de regressão polinomial?
- Considere os três modelos da figura 1. Qual modelo cuja fronteira de decisão generaliza melhor para pontos fora do conjunto de treinamento? Qual modelo tem overfitting e por quê?
- O que é regressão logística?
- Encontre os valores para os parâmetros θ na equação $h_{\theta}(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2)$, onde $g(z) = 1/(1 + \exp(-z))$ é a função sigmóide ou logística, de modo que a fronteira de decisão definida por $h_{\theta}(x)$ separe os pontos X dos pontos O na Figura 2. Suponha que a predição seja 1 (X) se $h_{\theta}(x) \geq 0.5$ e 0 (O) se $h_{\theta}(x) < 0.5$. **Desenhe a fronteira de decisão** em um novo gráfico similar ao da Fig. 2.

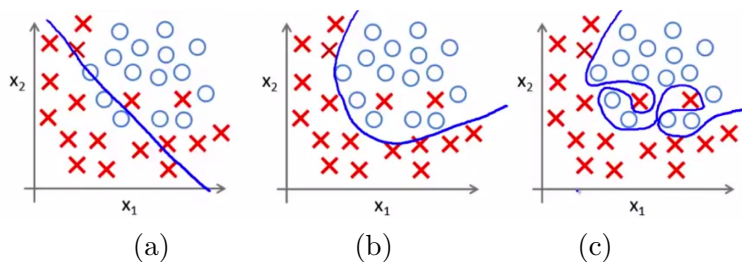


Figura 1: Três modelos para regressão logística.

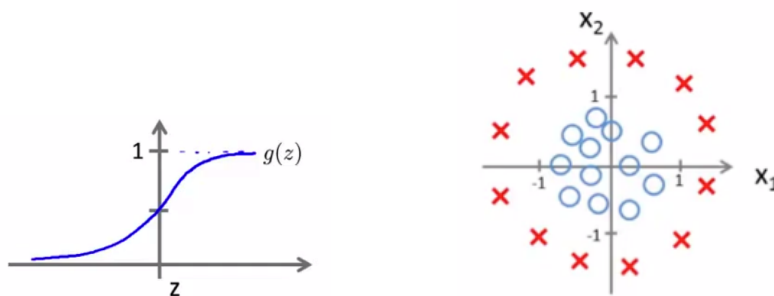


Figura 2: Esquerda: Função sigmóide. Direita: Conjunto de pontos para um problema de classificação hipotético.

2. Questões sobre redes neurais e regularização:

- Explique como funciona o algoritmo de retro-propagação para redes neurais multi-camadas. Ele sempre converge para o mínimo global? Por quê? Ilustre sua justificativa com um gráfico onde o eixo vertical representa a função de custo $J(\theta)$, e o eixo horizontal representa o parâmetro θ .
- O que é overfitting e underfitting? Cite 2 maneiras de amenizar o problema de overfitting.

3. Questões sobre aprendizagem não-supervisionada:

- Quando utilizar métodos de aprendizagem não-supervisionada? Cite uma aplicação.
- O que é K-means? Quais os dois passos principais deste algoritmo?