

# SCC0276 - Aprendizado de Máquina

## Aula Regressão Multivariada

Profa. Dra. Roseli Aparecida Francelin Romero  
SCC - ICMC - USP

2019

# Sumário

## 1 Introduction

## 2 Modelos de Regressão

## 3 Regressão Linear Ponderada Localmente

# O que é Regressão Linear

- Regressão é um tipo de algoritmo supervisionado no qual temos um alvo ou algo que desejamos prever.
- A diferença entre regressão e classificação é que na regressão, o alvo é um valor numérico e contínuo.

# Regressão Linear

- LR é usado para encontrar uma relação linear entre o alvo e um ou mais preditores.
- A variável que estamos prevendo é chamada de (variável de critério, variável de resultado, variável endógena ou independente)
- a variável (s) na qual baseamos nossas previsões é chamada de (variável preditora, variável exógena ou dependente).

# Regressão Linear

- Existem dois tipos de regressão linear: Regressão Linear Univariada ou Regressão Multivariada.
- Quando há apenas uma variável preditora, o método de previsão é chamado Regressão Univariada. O gráfico de uma Regressão Univariada sempre forma uma linha reta.

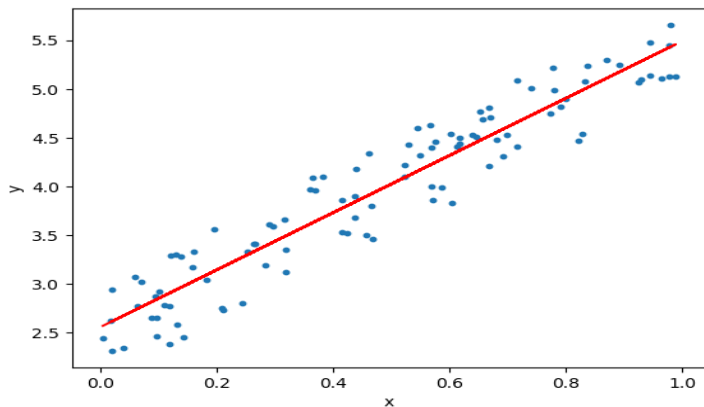


Figura 1: Regressão Linear

# Sumário

- 1 Introduction
- 2 Modelos de Regressão
- 3 Regressão Linear Ponderada Localmente

# Modelos de Regressão

- Regressão Linear:  $Energy = 0.0015 * food - 0.99 * water$
- Regressao Não Linear:  $Energy = 0.0015 * food/water$



# Modelos de Regressão

- Compute a reta de melhor ajuste Suponha que nossos dados de entrada estejam na matriz  $X$  e nossos pesos de regressão no vetor  $w$ .
- Para um dado dado  $X_1$ , nosso valor previsto é dado por  $y_1 = X_1^T w$

# Modelos de Regressão

- Nós temos Xs e Ys, mas como podemos encontrar os ws?
- Uma maneira é encontrar os ws que minimizam o erro.
- Nós definimos erro como a diferença entre y previsto e o y atual. Usar apenas o erro permitirá que valores positivos e negativos sejam cancelados, então usamos o erro ao quadrado como:

- $\sum_{i=0}^m (y_i - x_i^T w)^2$

- Escrevendo isso em notação matricial, temos  $(y - Xw)^T (y - Xw)$

# Modelos de Regressão

- Se tomarmos a derivada disso em relação a  $w$ , obteremos  $X^T(y - Xw)$
- Podemos definir isso como zero e resolver para obter a seguinte equação final:



$$w = (X^T X)^{-1} X^T y$$

- A equação final tem  $(X^T X)^{-1}$ , que é uma matriz inversa.

# Modelos de Regressão

- Portanto, temos que primeiro verificar a existência inversa da matriz antes de usá-la ou podemos ter um erro

# Regressão Linear Ponderada Localmente (LWLR)

- Uma maneira de reduzir o erro médio quadrático é uma técnica conhecida como LWLR.
- Com o LWLR, damos um peso aos pontos de dados próximos ao ponto de interesse de dados; então calculamos uma regressão de mínimos quadrados.
- A formula agora se torna:

$$w = (X^T W X)^{-1} X^T W y$$

## Caso Geral

- Queremos aproximar  $y$  por uma combinação linear de funções quaisquer,  $n$  é no. de pontos e  $m$  é o no. de funções ( $m \leq n$ )
- $y = w_1 g_1(x) + w_2 g_2(x) + \dots + w_m g_m(x)$
- A solução é dada por:

$$\begin{bmatrix} (g_1, g_1) & (g_1, g_2) & \dots & (g_1, g_m) \\ (g_2, g_1) & (g_2, g_2) & \dots & (g_2, g_m) \\ \dots & \dots & \dots & \dots \\ (g_n, g_1) & (g_n, g_2) & \dots & (g_n, g_m) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_m \end{bmatrix} = \begin{bmatrix} (y, g_1) \\ (y, g_2) \\ \dots \\ (y, g_m) \end{bmatrix}$$

# Regressão Multivariada

- Ocorre quando  $x$  é uma variável multi-dimensional, isto é,  
 $X = X_1, X_2, \dots, X_n$
- A solução é a mesma, dada por:  $W = (X^T X)^{-1} \cdot (X^T Y)$  onde  $X$  é um vetor  $n$ -dimensional
- Caso Bidimensional: Se  $X = (X_1, X_2)$ , teremos:

$$Y = W_1 \cdot 1 + W_2 \cdot X_1 + W_3 \cdot X_2 + W_4 \cdot X_1^2 + W_5 \cdot X_2^2 + W_6 \cdot X_1 \cdot X_2$$

$$g_1 = (1, 1, 1, \dots, 1)^T$$

$$g_2 = (X_1)^T$$

$$g_3 = (X_2)^T$$

$$g_4 = (X_1^2)^T$$

$$g_5 = (X_2^2)^T$$

$$g_6 = (X_1 \cdot X_2)^T$$

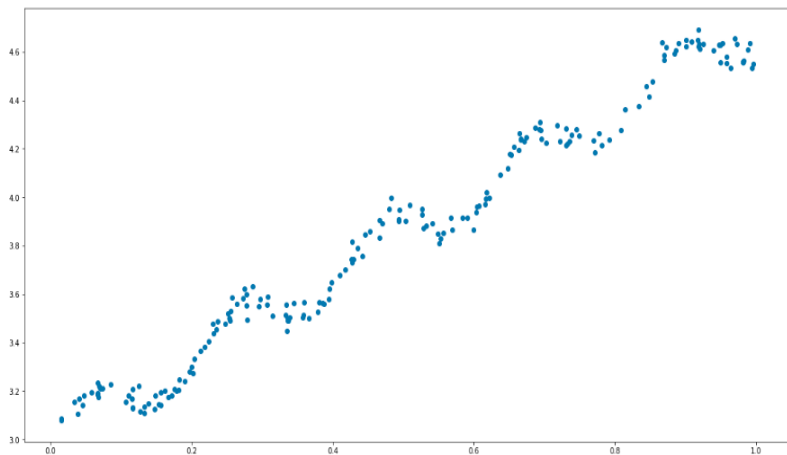


Figura 2: Regressão Linear



# Sumário

- 1 Introduction
- 2 Modelos de Regressão
- 3 Regressão Linear Ponderada Localmente

# Regressão Linear Ponderada Localmente (LWLR)

- A regressão linear tem um problema, é que ela tende a "underfit" os dados.
- Isso nos dá o menor erro médio-quadrado para estimadores não-biased. Portanto, com o underfitting, não estamos obtendo as melhores previsões.

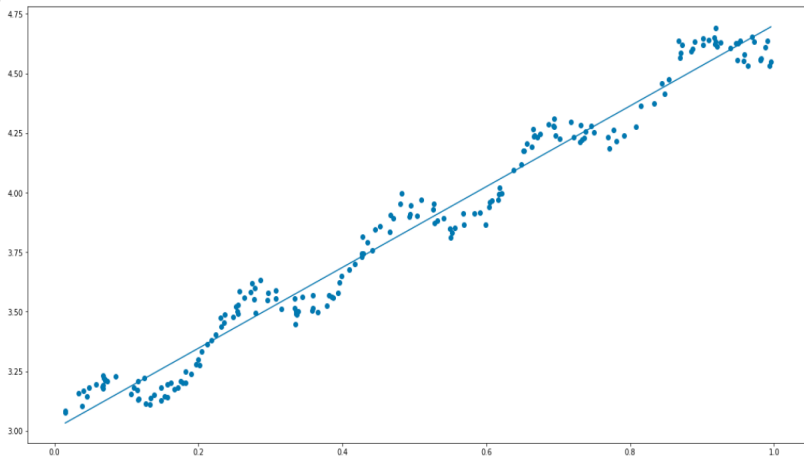


Figura 3: Regressão Linear com Underfitting

## Regressão Linear Ponderada Localmente (LWLR)

- W aqui é uma matriz que é usada para ponderar os pontos de dados.
- O LWLR usa um kernel similar ao do SVM para pesar pontos próximos mais fortemente que outros pontos.
- O kernel mais comum para usar é o kernel Gaussiano. Isso atribui um peso dado por:
- 

$$w(i, i) = \exp\left(\frac{|x^{(i)} - \bar{x}|}{-2k^2}\right)$$

onde  $x^{(i)}$  são os pontos dados e k uma constante.

## Regressão Linear Ponderada Localmente (LWLR)

- A partir da fórmula acima, quanto mais próximo o ponto de dados  $x$  estiver dos outros pontos, maior  $w(i, i)$  será.
- Também vemos uma constante  $k$ , que é uma constante definida pelo usuário, que determinará quanto pesa pontos próximos.
- Este é o único parâmetro que temos que nos preocupar com o LWLR.

# Regressão Linear Ponderada Localmente (LWLR)

- **Prós:** Com um valor  $k$  adequado, podemos ter um melhor ajuste para nossos dados, sem "overfitting" e "underfitting"
- **Contra:**  
Envolve muita computação. Você deve usar os dados inteiros para refinar uma única estimativa

# Regressão Linear Ponderada Localmente (LWLR)

- K (1.0, 0.01, 0.003)

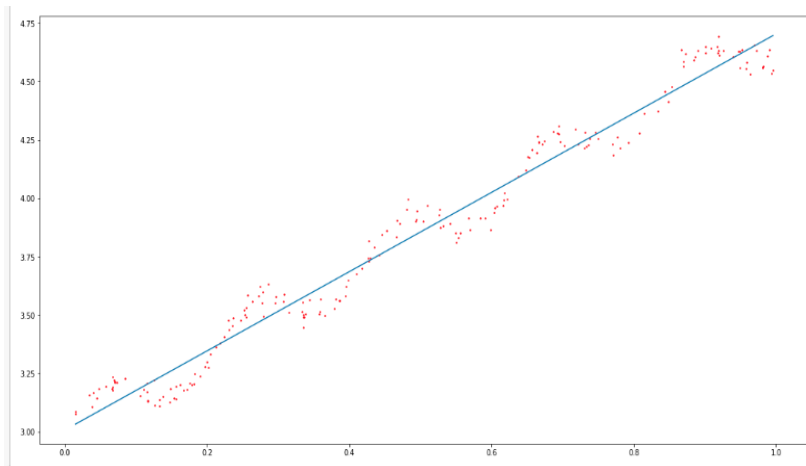


Figura 4:  $K=1.0$  - com Underfitting



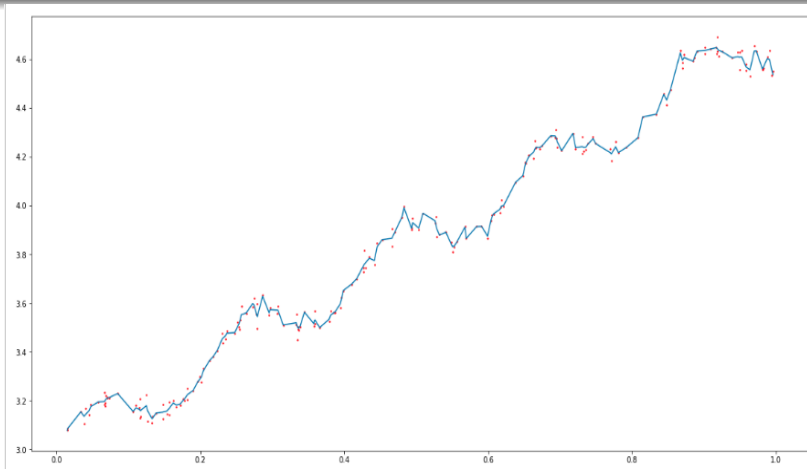


Figura 5:  $K=0.003$  - Overfitting

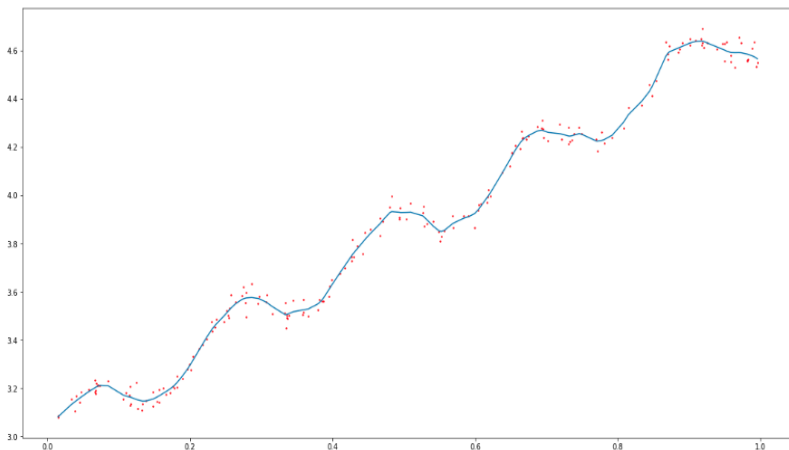


Figura 6:  $K=0.01$  - sem Underfitting and Overfitting

# Métricas de Regressão

- MAE: Mean Absolute Error:

$$\text{MAE} = \frac{\sum_{i=1}^n |Y_i - \hat{Y}_i|}{n} = \frac{\sum_{i=1}^n |e_i|}{n}.$$

# Métricas de Regressão

- MSE: Mean Squared Error: 
$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2.$$

# Métricas de Regressão

- $R^2$ : média:  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ ;
- soma total dos quadrados:  $SS_{\text{tot}} = \sum_i (y_i - \bar{y})^2$ ,
- soma dos quadrados dos resíduos:  
$$SS_{\text{res}} = \sum_i (y_i - f_i)^2 = \sum_i e_i^2$$
- Coeficiente de Determinação:  $R^2 \equiv 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$

# Métricas de Classificação

- Accuracy.
- Logarithmic Loss.
- Area Under ROC Curve.

# Métricas para Resultados de Predição na Classificação

- Confusion Matrix.
- Classification Report.