Aula 06 - Parte 02 - Regressão

Álgebra Linear e Teoria da Informação

Prof. Tiago Tavares

Revisando a Regressão Linear

Até o momento, trabalhamos com um modelo linear simples:

$$f(x) = \hat{y} = ax + b$$

Para cada medição (x_i, y_i), o erro individual é:

$$e_i = y_i - \hat{y}_i$$

E a métrica de desempenho do modelo, o Erro Quadrático Médio (EQM), é calculada como:

$$\text{EQM} = \frac{1}{n} \sum_{i=1}^{n} ||e_i||^2$$

Onde n é o número total de medições.

Formulação Matricial do Modelo Linear

Podemos reescrever o modelo linear de forma matricial. Para um conjunto de n medições (x_i, y_i), as estimativas $\hat{y}_i$ podem ser representadas por um vetor:

$$\begin{aligned}\hat{\boldsymbol{y}} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} &= \begin{bmatrix} ax_1 + b \\ ax_2 + b \\ \vdots \\ ax_n + b \end{bmatrix} \\ &= \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}\end{aligned}$$

Formulação Matricial do Erro (EQM)

A diferença entre as predições e os valores medidos também pode ser expressa em um vetor:

$$\boldsymbol{e} = \boldsymbol{y} - \boldsymbol{\hat{y}} = \begin{bmatrix} y_1 - \hat{y}_1 \\ y_2 - \hat{y}_2 \\ \vdots \\ y_n - \hat{y}_n \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

O Erro Quadrático Médio (EQM) pode ser calculado utilizando a transposta do vetor de erros:

$$\mathrm{EQM} = \frac{1}{n}\boldsymbol{e}^T\boldsymbol{e}$$

Modelos com Múltiplas Entradas

A formulação matricial permite representar modelos mais complexos! Por exemplo, para um modelo polinomial de segundo grau: $f(x) = ax^2 + bx + c$:

$$\hat{\boldsymbol{y}} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} = \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Para calcular uma única predição, por exemplo, $f(3)$:

$$f(3) = \begin{bmatrix} 3^2 & 3 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 9 & 3 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Estrutura das Matrizes

Para o modelo $\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{w}$:

- **Vetor de Estimativas ($\hat{\boldsymbol{y}}$):**
- **Matriz de Entradas ($\boldsymbol{X}$):**
- **Vetor de Pesos ($\boldsymbol{w}$):**

Nosso problema central é estimar os elementos de $\boldsymbol{w}$ de forma a minimizar o erro ($\boldsymbol{e}^T\boldsymbol{e}$) entre as predições ($\hat{\boldsymbol{y}}$) e os valores medidos ($\boldsymbol{y}$).

Calculando Parâmetros com a Pseudoinversa

Se o modelo fosse perfeito, as predições seriam exatamente os valores medidos:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{w}$$

Geralmente, $\boldsymbol{X}$ não é uma matriz quadrada e, portanto, não é inversível diretamente. Para resolver $\boldsymbol{w}$, multiplicamos ambos os lados pela transposta de $\boldsymbol{X}$, $\boldsymbol{X}^T$:

$$\boldsymbol{X}^T\boldsymbol{y} = \boldsymbol{X}^T\boldsymbol{X}\boldsymbol{w}$$

A matriz $(\boldsymbol{X}^T\boldsymbol{X})$ é quadrada e pode ser inversível. Então, podemos chegar a:

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

A Pseudoinversa de uma Matriz

A expressão $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$ é conhecida como a **pseudoinversa** de $\boldsymbol{X}$, denotada por $\boldsymbol{X}^+$.

Assim, o vetor de pesos $\boldsymbol{w}$ que minimiza o Erro Quadrático Médio pode ser encontrado diretamente:

$$\boldsymbol{w} = \boldsymbol{X}^+\boldsymbol{y}$$

Esta solução de forma fechada é extremamente útil para encontrar os parâmetros ideais de um modelo linear nos parâmetros.

Hora da chamada!

Hoje estamos na Aula 06, Parte 02!