

Métodos dos Mínimos Quadrados

Gustavo Aparecido de Souza Viana

Abstract—Métodos de regressão e classificação são utilizados a fim de prever algo. Assim, neste trabalho será abordado a implementação do Método dos Mínimos quadrados caracterizado como um tipo de Regressão, onde a saída do mesmo é quantitativa. Para os experimentos, foram utilizados três bases numéricas e além disso três variações do método. Os resultados mostraram que as três técnicas tem capacidade de prever alguma informação baseada em uma base de dados pequena com uma boa acurácia.

I. INTRODUÇÃO

A necessidade de predições está sendo bem comum e usada na atualidade, podemos dar o exemplo de predições relacionadas ao mercado de ações ou até mesmo predições relacionadas a área médica, prever se o paciente tem tendências de desenvolver algum tipo de câncer.

Conseguimos prever algum informação baseada em uma base conhecida com machine learning ou com métodos estatísticos. É de extrema importância conhecer o ramo de atividade da aplicação pois métodos de machine learning geralmente estão relacionados a rede neural, consequentemente necessitam de mais dados para que as predições tenham uma maior acurácia. De contra partida, métodos estatísticos não necessitam.

Neste trabalho foi abordado a implementação do Métodos dos mínimos quadrados com objetivo de criar métodos para aplicar uma regressão e assim prever alguma informação, as técnicas utilizadas foram: Linear, Quadrática e Robusta. Para testar essas técnicas foram utilizados três bases *Alps Water*, *Books x Grades* e *US Census Dataset*.

II. CONCEITOS FUNDAMENTAIS

Nesta seção, será apresentado os conceitos básicos de Método dos Mínimo Quadrados juntamente com as técnicas de Mínimo Quadrados linear, quadrática e robusta.

A. Método dos Mínimo Quadrados

O método consiste em prever uma saída por meio de uma entrada contendo dados em um vetor, e assim por meio da equação fechada, prever. Ou seja, temos como entrada as *features* e a saída espera para o mesmo, assim conseguimos descobrir uma matriz contendo os valores de que serão utilizados para prever algo. Representação do Método dos Mínimo Quadrados na Equação 1.

$$\hat{Y} = \hat{\beta}_0 + \sum_{j=1}^p X_j \hat{\beta}_j \quad (1)$$

B. Mínimo Quadrados Linear

O método utiliza a base do Método dos Mínimo Quadrados mas com adicionando uma coluna com valor = 1 na matrix X e assim ajustando somente a equação a fim de minimizar o erro. Assim temos a seguinte equação de ajuste no modelo linear, Equação 2, onde X é a matrix de entrada somente com as *features* e y é a matrix de tamanho $N \times 1$ com a saída esperada. E a Equação 3 para efetuar a predição.

$$X = \begin{bmatrix} 1 & a \\ 1 & b \\ 1 & c \\ \vdots & \vdots \\ \vdots & \vdots \end{bmatrix}$$

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (2)$$

$$predição = \begin{bmatrix} 1 & a \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \quad (3)$$

C. Mínimo Quadrados Quadrático

O método utiliza a base do Método dos Mínimo Quadrados Linear mas com um pequeno ajuste na matrix X necessitando adicionar uma coluna com o valor de X^2 . Assim temos a seguinte equação de ajuste no modelo quadrático, Equação 4, onde X é a matrix de entrada somente com as *features* e y é a matrix de tamanho $N \times 1$ com a saída esperada. E a Equação 5 para efetuar a predição.

$$X = \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \end{bmatrix}$$

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (4)$$

$$predição = \begin{bmatrix} 1 & a & a^2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \quad (5)$$

D. Mínimo Quadrados Robusto

O método utiliza a base do Método dos Mínimo Quadrados Linear mas o modelo passa por um pré processamento com intuito de ajustar o valor de y e a matrix X previamente. Assim temos a seguinte equação de ajuste no modelo robusto, Equação 7, onde X é a matrix de entrada somente com as *features*, y é a matrix de tamanho $N \times 1$ com a saída esperada e W é o coeficiente para ajustar melhor a matrix X e o valor de y . E a Equação 3 para efetuar a predição.

$$w_i = 1/|(y_{real} - y_{estimado})| \quad (6)$$

$$\hat{\beta} = (X^T W X)^{-1} X^T W y \quad (7)$$

$$predição = \begin{vmatrix} 1 & a \\ \beta_1 & \beta_2 \end{vmatrix} \quad (8)$$

III. METODOLOGIA

Nesta seção será apresentado a metodologia utilizada para implementar o Método dos Mínimo Quadrados juntamente com as técnicas de Mínimo Quadrados linear, quadrática e robusta, implementados em Python. O código fonte pode ser encontrado em <https://github.com/apparecido/master-special-learning-topic>.

Vale ressaltar que as técnicas para a regressão necessitam de manipulações com matrizes, para isso foi criado algumas funções, como inversão, determinante, transposta e multiplicação de matrizes, que seus respectivos pseudocódigos não serão citados neste trabalho.

Para cada técnica, foram implementados dois métodos com a responsabilidade de calcular a matrix de β_i e o método para fazer uma predição baseado na matrix de β_i encontrada. Ambos métodos seguem as fórmulas citadas na sessão de Conceitos Fundamentais.

Os método de regressão para cada técnica, chamados de "linear_least_squares", "quadratic_least_squares" e "robust_least_squares", recebem a matrix X de *features* e como segundo parâmetro o y que é o valor esperado, e retornam a matrix de β_i . Como podemos ver abaixo:

Algorithm 1: linear_least_squares(matrix, y)

```
add_begin_column_value_one(matrix)
// (X^T * X)^-1 * X^T * y
return inverse((transpose(matrix) * matrix)) *
transpose(matrix) * y
```

Algorithm 2: quadratic_least_squares(matrix, y)

```
add_begin_column_value_one(matrix)
matrix_add_end_column_square(matrix)
// (X^T * X)^-1 * X^T * y
return inverse((transpose(matrix) * matrix)) *
transpose(matrix) * y
```

Algorithm 3: robust_least_squares(matrix, y)

```
betas = linear_least_squares(matrix, y)
newy = linear_predict(matrix, betas)
w = calculate_w(y)
y = matrix_escalar_matrix_multiplication(y, w)
matrix = matrix_escalar_matrix_multiplication(matrix,
w)
// (X^T * X)^-1 * X^T * y
return inverse((transpose(matrix) * matrix)) *
transpose(matrix) * y
```

Os métodos "linear_predict" e "quadratic_predict" são responsáveis pelas predições, sendo assim recebem como

parâmetro uma matrix X com um caso a ser previsto e como segundo parâmetro a matrix de β_i .

Algorithm 4: linear_predict(matrix_test, betas)

```
add_begin_column_value_one(matrix_test)
return matrix_multiplication(matrix_test, betas)
```

Algorithm 5: quadratic_predict(matrix_test, betas)

```
add_begin_column_value_one(matrix_test)
matrix_add_end_column_square(matrix_test)
return matrix_multiplication(matrix_test, betas)
```

IV. EXPERIMENTOS E RESULTADOS

Nesta seção será apresentado os experimentos e seus respectivos resultados.

O experimento consiste na aplicação das três técnicas citadas baseadas no Método dos Mínimos Quadrados, sendo elas linear, quadrática e robusta para cada base de dados, sendo elas a *Alps Water*, *Books x Grades* e *US Census Dataset*.

A Figura 3 representa os resultados contendo os valores de β e o resultado de uma predição para cada técnica e base de dados utilizada.

```
>>>>> ../dataset/alpswater.csv <<<<<

Beta for linear_least_squares
[-81.0637]
[0.5229]
>> Prediction for:
[190]
Is: 18.2858290204276

Beta for quadratic_least_squares
[38.8293]
[-0.6548]
[0.0029]
>> Prediction for:
[190]
Is: 18.741470484161326

Beta for robust_least_squares
[-81.393]
[0.5245]
>> Prediction for:
[190]
Is: 18.261107478859074
```

Fig. 1: Representação dos resultados da matrix de β e uma predição para o dataset *Alps Water*

```
>>>>> ../dataset/books_attend_grade.csv <<<<<<
```

```
Beta for linear_least_squares  
[37.3792]  
[4.0369]  
[1.2835]  
>> Prediction for:  
[0, 9]  
Is: 48.930480676960784
```

```
Beta for quadratic_least_squares  
[64.0916]  
[4.1689]  
[-2.9102]  
[0.1484]  
>> Prediction for:  
[0, 9]  
Is: 49.92158646276325
```

```
Beta for robust_least_squares  
[33.8229]  
[3.9131]  
[1.5655]  
>> Prediction for:  
[0, 9]  
Is: 47.91216941632925
```

Fig. 2: Representação dos resultados da matrix de β e uma predição para o dataset *Books x Grades*

V. CONCLUSÃO

Após a análise do conceito e resultados obtidos após a aplicação dos métodos e comparando os mesmos, podemos concluir que todas as técnicas utilizadas conseguiram atingir o objetivo de prever uma informação na base de dados com uma boa acurácia e o método robusto teve uma melhor eficiência que muito provável se dá ao *fitting* que a matrix X e y recebem.

REFERENCES

```
>>>>> ../dataset/us_census.csv <<<<<<
```

```
Beta for linear_least_squares  
[-3783.9456]  
[2.0253]  
>> Prediction for:  
[2010.0]  
Is: 286.9128909090773
```

```
Beta for quadratic_least_squares  
[32294.0174]  
[-34.9875]  
[0.0095]  
>> Prediction for:  
[2010.0]  
Is: 311.5880717733089
```

```
Beta for robust_least_squares  
[-3782.3332]  
[2.0246]  
>> Prediction for:  
[2010.0]  
Is: 287.21107201161976
```

Fig. 3: Representação dos resultados da matrix de β e uma predição para o dataset *US Census Dataset*