Pseudo-Inverse Matrix for Data Scientists

Manish Rai

2017-08-03

1 Intuition behind Pseudo-inverse

The inverse of a matrix depends on it's adjunct (adj(A)) and it's determinant (det(A)), and is given by $A^{-1} = \frac{adj(A)}{|A|}$. This formula shows that for the matrix A had to be a non-singular square matrix for it's inverse to exist. But we don't always encounter square matrix in our lives and $Moore - Penrose - Pseudoinverse$ of a matrix gives a fair approximation to the inverse of a non-square matrix.

2 Applications for a Data Scientists

As a data scientists we come across problems which require computing the inverse of matrices. We have to deal with problem $Y = XW$, where Y is a column matrix of the response/target variables and X is the N x D matrix with N-datasets and D-dimensions.

Each regression problem tries to solve for the weights, W, that minimizes the expected error. And the elegant way to solve for W is through linear algebra like

$$W_{ML} = X^{-1}Y \tag{1}$$

where X inverse is the input matrix and depends on the size of the input dataset. This matrix is not necessarily a square matrix and hence to compute the weights we do need a Pseudo-inverse matrix for X which is given as

$$X^{+} = (X^{T}X)^{-1}X^{T} \tag{2}$$

This was the simplest form of the regression equation. But, a majority of times we preprocess the input data (X) or do some feature extraction from it, before feeding it into the model. This extracted feature can be expressed in terms of the basis function $\phi(x)$.

In this case also, we need to calculate the inverse of basis $\phi(x)$ matrix. This is done easily through the pseudo-inverse calculation.

$$\phi^{+} = (\phi^{T}\phi)^{-1}\phi^{T} \tag{3}$$

$$W_{ML} = (\phi^{T}\phi)^{-1}\phi^{T}Y \tag{4}$$