

Machine Learning

Sommersemester2020

Exercise 0

Ciheng Zhang (3473188) zch3183505@gmail.com
Gang Yu stxxxxx@stud.uni-stuttgart.de
Huipanjun Tian (3471607) Thpjpyl5111217@gmail.com

April 26, 2020

1 Matrix equations

1.1

$$\begin{aligned}XA + A^T &= I \\XAA^{-1} + A^TA^{-1} &= A^{-1} \\X + A^TA^{-1} &= A^{-1} \\X &= A^{-1} - A^TA^{-1}\end{aligned}$$

1.2

$$\begin{aligned}X^TC &= [2A(X + B)]^T \\X^TC &= 2(x + B)^TA^T \\X^TC &= 2X^TA^T + 2B^TA^T \\x^T(C - 2A^T) &= 2B^TA^T \\x &= (2B^TA^T(C - 2A)^{-1})^T \\X &= 2((C - 2A)^{-1})^TAB\end{aligned}$$

1.3

$$\begin{aligned}(Ax)^{-1}A - y^TA &= 0_n^T \\x^TA^TA &= y^TA \\x^T &= y^TAA^{-1}(A^T)^{-1} \\x^T &= y^T(A^T)^{-1} \\x &= A^{-1}y\end{aligned}$$

1.4

$$\begin{aligned}(Ax - y)^TA + x^TB &= 0_n^T \\(Ax)^TA - y^TA + A^TB &= 0_n^T \\x^TA^TA - y^TA &= -A^TB \\x^TA^TA &= y^TA - A^TB \\x^T &= Ay^T(A^T)^{-1} - A^TBA^{-1}(A^T)^{-1} \\x &= A^{-1}y - A^T(A^{-1})^TB^TA\end{aligned}$$

2 Vector derivatives

2.1

a is a nxn Matrix

2.2

$$\frac{\partial}{\partial x}[x^T x] = x$$

2.3

let

$$f(x) = (Ax - y)(Ax - y)^T + x^T Bx$$

when we want to calculate the minimum of $f(x)$, we should at first the derivatives calculation

$$\frac{df(x)}{dx} = A(Ax - y) + Bx$$

$$\frac{df(x)}{dx} = A^2 x + Bx - Ay$$

So the minimum values is the time $f(x) = 0$. And $A^2 > 0$ and B is positive definite. Therefore the minimum of the equations is

$$(A^2 + B)x = Ay$$

$$x = Ay(A^2 + B)^{-1}$$

3 Error Measures

3.1

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2$$

$$MSE = \frac{1}{n} (y - \hat{y})(y - \hat{y})^T$$

$$MAE = \frac{1}{N} \sum_{i=1}^n |y_i - \hat{y}| = \frac{1}{n} \|y - \hat{y}\|_1$$

3.2

MAE is more robust to outliers. MSE calculation is simple. Also if the outliers are important for the Problem, people should choose MSE. If the outliers are just wrong data, people should use MAE.

3.3

both y and \hat{y} have 2 situations: $=0$ or $=1$. So if the true values and predicted values are equal, both MSE and MAE are equal to 0. If the true values equal to 0 and predicted values equal to 0, both MSE and MAE are equal to 0 and vice versa. Therefore in this situation MSE and MAE are equal.