Introduction to ML (CS771), Autumn 2018
Indian Institute of Technology Kanpur
Homework Assignment Number 4

Student Name: Avijit Roy
Roll Number: 18111404
Date: November 11, 2018

QUESTION

1

1 Eigenchangers!

X is an $N \times D$ matrix over $\mathbb{R}$. An Singular Value Decomposition(SVD) of X is $X = U\Sigma V^T$,where-

1.The columns of U and V are orthonormal.

2. if Σ =(σ_{ij}) then $\sigma_{ij} = 0$ for $i \neq 0$ and $\sigma_{11} \geq \sigma_{22} \geq ... \geq 0$. Let $\sigma_{ii} = \sigma_i$ are called singular values of X.

Theorem: Let X = $U\Sigma V^T$ be an SVD of A. Then,

1. σ_i's are the square roots of the eigenvalues of X^TX (As well as XX^T).
2. The columns of V are the eigenvectors of X^TX.
3. The columns of U are the eigenvectors of XX^T

Proof:

Columns of V are $(v_1, ..., v_n)$

Columns of U are $(u_1, ..., u_n)$

$$X^TXv_i = (U\Sigma V^T)^T U\Sigma V^T v_i$$
$$= (V^T)^T\Sigma^T U^T U\Sigma V^T v_i$$
$$= V\Sigma^T\Sigma V^T v_i = V\Sigma^T\Sigma e_i$$
$$= V\Sigma_i^2 e_i = \sigma_i^2 v_i$$

Here, We get the value of $\Sigma^T\Sigma$ as-

$$\Sigma^T\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & 0 & \dots & 0 \\ & & \dots\dots\dots & & \\ 0 & 0 & 0 & 0 & \sigma_r \\ & & \dots\dots\dots & & \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & 0 & \dots & 0 \\ & & \dots\dots\dots & & \\ 0 & 0 & 0 & 0 & 0 \\ & & \dots\dots\dots & & \\ 0 & 0 & 0 & \sigma_r & 0 \end{bmatrix} e_i$$

$$= \begin{bmatrix} \sigma_1^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & 0 & \dots & 0 \\ & & \dots\dots\dots & & \\ 0 & 0 & 0 & \sigma_r^2 & 0 \\ & & \dots\dots\dots & & \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}_i = \sigma_i^2 \times e_i$$

$\therefore$ v_i is an eigenvector of X^TX with eigenvalue σ_i^2

The diagonal entries of Σ are the roots of the eigenvalues of X^TX.

Same argument shows that u_i's are the eigenvectors of X^TX with eigenvalues σ_i^2.

All the columns of U and V are linearly independent(as they are orthogonal). So, U and V invertible.

$$XV = U\Sigma \qquad [V^T = V^{-1}]$$

$$= (u_1, ..., u_m)\Sigma$$

$$= (\sigma u_1, \sigma u_2, ..., \sigma_m u_m)$$

$$\therefore Xv_i = \sigma_i u_i$$

$$\boxed{u_i = \frac{Xv_i}{\sigma_i}}$$

Advantage: Determining the eigenvalue of $XX^T(N \times N)$ will take $O(N^3)$ and for $X^T X(D \times D)$ it will take $O(D^3)$. As $N \leq D$, the time complexity will be less for XX^T case.

Student Name: Avijit Roy
Roll Number: 18111404
Date: November 11, 2018

QUESTION

2

2 A General Activation Function

1. $\beta = kx$ where $k \to \infty$ will approximate the Identity function.
2. $\beta \to \infty$ then h(x) will approximate ReLU.

Student Name: Avijit Roy
Roll Number: 18111404
Date: November 11, 2018

QUESTION 3

My solution to problem 3

Introduction to ML (CS771), Autumn 2018
Indian Institute of Technology Kanpur
Homework Assignment Number 4

Student Name: Avijit Roy
Roll Number: 18111404
Date: November 11, 2018

QUESTION 4

My solution to problem 4

Introduction to ML (CS771), Autumn 2018
Indian Institute of Technology Kanpur
Homework Assignment Number 4

QUESTION 5

Student Name: Avijit Roy
Roll Number: 18111404
Date: November 11, 2018

My solution to problem 5

Introduction to ML (CS771), Autumn 2018
Indian Institute of Technology Kanpur
Homework Assignment Number 4

Student Name: Avijit Roy
Roll Number: 18111404
Date: November 11, 2018

QUESTION 6

My solution to problem 6