Cover page for answers.pdf
CSE512 Fall 2018 - Machine Learning - Homework 6

Your Name: SAIF SULEMAN VAZIR

Solar ID: 112072061

NetID email address: saifsuleman.vazir@stonybrook.edu

Names of people whom you discussed the homework with: -

HOMEWORK 6

1) To prove:

$$\bar{C} = \frac{1}{n} X X^T - \lambda_1 v_1 v_1^T$$

$$\bar{X} = (I - v_1 v_1^T) X$$

$$\therefore \bar{X}^T = X^T (I - v_1 v_1^T)$$

$$\bar{C} = \frac{1}{n} \bar{X} \bar{X}^T \Rightarrow \frac{1}{n} \left[(I - v_1 v_1^T) X \right] \left[X^T (I - v_1 v_1^T) \right]$$

$$= \frac{1}{n} \left[(X - v_1 v_1^T X)(X^T - X^T v_1 v_1^T) \right]$$

$$= \frac{1}{n} \left[X X^T - (X X^T v_1) v_1^T - v_1 v_1^T X X^T + v_1 v_1^T (X X^T v_1) v_1^T \right]$$

using $X X^T v_1 = n \lambda_1 v_1$, we get

$$\therefore \bar{C} = \frac{1}{n} \left[X X^T - n \lambda_1 v_1 v_1^T - v_1 v_1^T X X^T + v_1 v_1^T n \lambda_1 v_1 v_1^T \right]$$

$$= \frac{1}{n} X X^T + \frac{1}{n} \left[-n \lambda_1 v_1 v_1^T - v_1 v_1^T X X^T + n \lambda_1 v_1 (v_1^T v_1) v_1^T \right]$$

using $v_1^T v_1 = I$

$$\bar{C} = \frac{1}{n} X X^T + \frac{1}{n} \left[\cancel{-n \lambda_1 v_1 v_1^T} - v_1 v_1^T X X^T + \cancel{n \lambda_1 v_1 v_1^T} \right]$$

$$= \frac{1}{n} X X^T + \frac{1}{n} \left[- \underbrace{(v_1 v_1^T)}_{B^T} \underbrace{(X X^T)}_{A^T} \right]$$

$$\therefore \bar{C} = \frac{1}{n} X X^T - \frac{1}{n} \left[(A \cdot B)^T \right]$$

$$= \frac{1}{n} X X^T - \frac{1}{n} \left[(\underbrace{X X^T v_1}_{n \lambda_1 v_1} v_1^T)^T \right]$$

$$= \frac{1}{n} X X^T - \frac{1}{\cancel{n}} \left[\cancel{n} \lambda_1 v_1 v_1^T \right]^T = \frac{1}{n} X X^T - \lambda_1 v_1 v_1^T$$

Hence proved.

Scanned by CamScanner

2) To show that for $\forall j \neq 1$, if v_j is a principal eigenvector of C with eigenvalue λ_j, v_j is also ~~the~~ a principal eigenvector of $\bar{C}$ with value λ_j.

Sol^n: we know that:

$$\bar{C} = \frac{1}{n} XX^T - \lambda_1 v_1 v_1^T$$

Multiplying throughout by v_j

$$\bar{C} v_j = \frac{1}{n} XX^T v_j - \lambda_1 v_1 (v_1^T v_j)$$

using $XX^T v_j = n \lambda_j v_j$ and $v_1^T v_j = 0$ $[j \neq 1]$

we get:

$$\bar{C} v_j = \frac{1}{\cancel{n}} [\cancel{n} \lambda_j v_j] - \lambda_1 v_1 (0)$$

$$\bar{C} v_j = \lambda_j v_j$$

Hence proved that v_j is a principal eigenvector of $\bar{C}$ with value λ_j.

3) To show that first principal eigenvector of $\bar{C}$ is $u = v_2$.

Sol^n:

we know that:

$$\bar{C} = \frac{1}{n} XX^T - \lambda_1 v_1 v_1^T$$

Multiplying by v_1 ~~to~~ we get:

$$\bar{C} v_1 = \frac{1}{n} \underbrace{XX^T v_1}_{n\lambda_1 v_1} - \lambda_1 v_1 \underbrace{(v_1^T v_1)}_{I}$$

$$\therefore \bar{C} v_1 = \frac{1}{\cancel{n}} \cancel{n} \lambda_1 v_1 - \lambda_1 v_1 (I)$$

$$= \lambda_1 v_1 - \lambda_1 v_1 = \underline{0}$$

$$\therefore \bar{C} v_1 = \underline{\underline{0 \cdot v_1}}$$

This can be explained by saying that as $\bar{C}$ is the deflated matrix, the eigenvalue corresponding to first principal eigenvector in C has changed to 0 in $\bar{C}$.

Now, the first (leading) principal vector of $\bar{C}$ will be v_2. [as $v_1 > v_2 > v_3$ ... in original C matrix]

4)
```
import numpy as np
def get_eigenvects (C, k, f):
    l = [ ]
    eigen_vects = [ ]
    for i in range(k):
        l_temp, eigen_temp = f(C)
        eigen_vects.append(eigen_temp)
        l.append (l_temp)
        C = C - (l_temp * eigen_temp *
                    np.transpose (eigen_temp))
    return l, eigen_vects
```