CSE512 Fall 2018 - Machine Learning - Homework 6

Your Name: Amit Dharmadhikari

Solar ID: 112044244

NetID email address: amit.dharmadhikari@stonybrook.com
adharmadhika@cs.stonybrook.edu

Names of people whom you discussed the homework with:
Mihir Parulekar

Q1 PCA via Successive Deflation

1. Show that the covariance of the deflated matrix $\tilde{C} = \frac{1}{n}\tilde{X}\tilde{X}^T$ is given by

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

Ans. Given :- $\tilde{X} = (I - v_1 v_1^T)X$

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

$$= \frac{1}{n}(I - v_1 v_1^T)XX^T(I - v_1 v_1^T)$$

$$= \frac{1}{n}(I - v_1 v_1^T)XX^T - \frac{1}{n}(I - v_1 v_1^T)XX^T v_1 v_1^T$$

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

$$+ \frac{1}{n}v_1 v_1^T XX^T v_1 v_1^T$$

Given that :- $XX^T v_1 = n\lambda_1 v_1$

$$\therefore \tilde{C} = \frac{1}{n}XX^T - \frac{1}{n}v_1 v_1^T XX^T - \lambda v_1 v_1^T + v_1 v_1^T \lambda_1 v_1 v_1^T$$

$$= \frac{1}{n}XX^T - \lambda_1 v_1 v_1^T - \frac{1}{n}v_1 v_1^T XX^T \cancel{+ v_1 v_1^T \lambda_1 v_1^T}$$

$$+ v_1 v_1^T \lambda_1 v_1 v_1^T$$

Now, we know $XX^T v_1 = n\lambda_1 v_1$

i.e. $v_1^T XX^T = n\lambda_1 v_1^T$

$$\therefore \tilde{C} = \frac{1}{n}XX^T - \lambda_1 v_1 v_1^T - \frac{1}{\cancel{n}}v_1 \cancel{n}\lambda_1 v_1^T$$

$$+ v_1 v_1^T \lambda_1 v_1 v_1^T$$

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

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

2. Show that for $j \neq 1$, if v_j is a principal eigenvector of C with corresponding eigenvalue d_j (that is, $Cv_j = d_j v_j$), then v_j is also a principal eigenvector of $\tilde{C}$ with the same eigenvalue d_j.

Ans. $Cv_j = d_j v_j$

$\tilde{C} = \frac{1}{n} XX^T - d_1 v_1 v_1^T v_j$

It is given that $j \neq 1$ and also that

$v_i^T v_j \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$

Let's consider $v_i^T v_j$ $\quad i = 1, i \neq j, \therefore v_i^T v_j = 0$

$v_i^T v_j = 0$

$\therefore \tilde{C} v_j = \frac{1}{n} XX^T v_j - 0 = \frac{1}{n} XX^T v_j$

But, $\frac{1}{n} XX^T v_j = C$

$\therefore \tilde{C} v_j = Cv_j$

However, it is given that,

$Cv_j = d_j v_j$

$\therefore \tilde{C} v_j = d_j v_j$

~~i.e. v~~ i.e. v_j is also the principal eigen vector of $\tilde{C}$ with same eigen value d_j

3. Let u be the first principal eigenvector of $\tilde{C}$. Explain why $u = v_2$. (You may assume u is unit norm.)

Ans. Let us see the proof for the second subquestion.

$$\tilde{C} v_j = XX^T v_j - \lambda_1 v_1 v_1^T v_j$$

if $j = 1$ ... first eigenvector

$$\tilde{C} v_1 = C v_1 - \lambda_1 v_1 v_1^T v_1$$

$$= \lambda_1 v_1 - \lambda_1 v_1 \quad \text{as } v_1^T v_1 = 1$$

$$\tilde{C} v_1 = \bar{0}$$

This is not an eigenvector (v_1)

$\therefore$ The first eigen vector starts from $j=2$ (i.e. v_2)

The first principal eigen vector has the largest eigen value.

$\therefore$ v_2 is the first principal eigen vector of $\tilde{C}$

$\therefore$ $u = v_2$

Hence, proved.

4. Suppose we have a simple method f for finding the leading eigen vector and eigen value of a positive-definite matrix, denoted by $[\lambda, u] = f(C)$. Write some pseudocode for finding the first k principal basis vectors of X that only uses the special f function and simple vector arithmetic.

→

```
Algorithm (c, k):
    vectors_list = []
    for i = 1 to k do:
        (λi, vi) = f(c)
        vectors_list.append(vi)
        c = c - λi vi
    end for
    return vectors_list
end algorithm
```