1 CSE512 Fall 2018 - Machine Learning - Homework 6

Name: Rohit Bhal
Solar ID: 112073893
NetID email address: rbhal@cs.stonybrook.edu / rohit.bhal@.stonybrook.edu
Names of people whom you discussed the homework with: Sayontan Ghosh

$\underline{\underline{1}}$

1. To Show: $\tilde{C} = \frac{1}{n}\tilde{X}\tilde{X}^T$

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

$\tilde{X}\tilde{X}^T = (I - v_1v_1^T)X((I-v_1v_1^T)X)^T \quad (\because \tilde{X} = (I-v_1v_1^T)X)$

$= (I-v_1v_1^T)XX^T(I-v_1v_1^T)$

$= (I-v_1v_1^T)(XX^T - XX^Tv_1v_1^T)$

$= (I-v_1v_1^T)(XX^T - n\lambda_1 v_1v_1^T)$

$= XX^T - n\lambda_1 v_1v_1^T - v_1v_1^TXX^T + v_1v_1^Tn\lambda_1v_1v_1^T$

$= XX^T - n\lambda_1v_1v_1^T - v_1v_1^TXX^T + n\lambda_1v_1v_1^Tv_1v_1^T$

$= XX^T - \cancel{n\lambda_1v_1v_1^T} - v_1v_1^TXX^T + \cancel{n\lambda_1v_1v_1^T}$

$= XX^T - v_1v_1^TXX^T$

$= (I-v_1v_1^T)XX^T$

$= (((I-v_1v_1^T)XX^T)^T)^T$

$= ((XX^T)^T(I-v_1v_1^T)^T)^T$

$= (XX^T(I-v_1v_1^T))^T$

$= (XX^T - XX^Tv_1v_1^T)^T$

$= (XX^T - n\lambda_1v_1v_1^T)^T$

$\tilde{X}\tilde{X}^T = XX^T - n\lambda_1v_1v_1^T$

$\Longrightarrow \tilde{C} = \frac{1}{n}\tilde{X}\tilde{X}^T = \frac{1}{n}XX^T - \lambda_1v_1v_1^T.$ Proved.

2. $\tilde{C} = \frac{1}{n} XX^T - \lambda_1 v_1 v_1^T = C - \lambda_1 v_1 v_1^T \quad \left(\because C = \frac{XX^T}{n}\right)$

Now, $\tilde{C} v_j = (C - \lambda_1 v_1 v_1^T) v_j$

$= C v_j - \lambda_1 v_1 v_1^T v_j$

If $j \neq 1$, then $v_1^T v_j = 0$.

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

3. Since, eigen ~~values~~ (vector) are sorted by eigen values.

$\therefore$ If λ_1 & v_1 are no longer the largest, then the v_2 has to be the largest.

$\tilde{C} v_1 = (C - \lambda_1 v_1 v_1^T) v_1 \quad (\because v_1^T v_1 = 1)$

$= C v_1 - \lambda_1 v_1 v_1^T v_1 = C v_1 - \lambda_1 v_1 = \lambda_1 v_1 - \lambda_1 v_1$

$= 0 = \lambda_1' v_1$

Thus, $\lambda_1' = 0$ & ($v_1 \neq 0$ by definition) is now strictly smaller than λ_2.

$\therefore u = v_2$. (principal eigen vector)

4.
```
def compute-kth-eigen (C, k, f):
  lbds = zeros(k,1)
  evs = zeros(k,1)
  for i = 1:k :
     [lbd, ev] = f(C)
      lbds[i] = lbd
       evs[i] = ev
      C = C - (lbd x ev x ev^T)
  end
  return [lbds, evs]
```