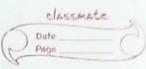
## SET-6



## 6.1) Deep Learning

- 2) Exercise 5.18 (Bishop (2006))
  - the infuls to the outputs, our model can be represented when it is a significant to the outputs, our model can be represented on:

- \* The equations in Section 5.3.2 mostly rumain the same, except Eq. (5.64), which is now  $y = \frac{\pi}{2} \frac{\omega_{ki}^{(3)}}{\omega_{ki}^{(3)}} z_{i} + \frac{\pi}{2} \frac{\omega_{ki}^{(3)}}{\omega_{ki}^{(3)}} z_{i}$
- \* The eq. for the derivative of the error function w.r.t.  $\frac{\partial F_n}{\partial x_n} = (y_n t_n) \partial_{y_n} y_k = (y_n t_n) z_n$   $\frac{\partial F_n}{\partial y_n} = (y_n t_n) \partial_{y_n} y_k = (y_n t_n) z_n$

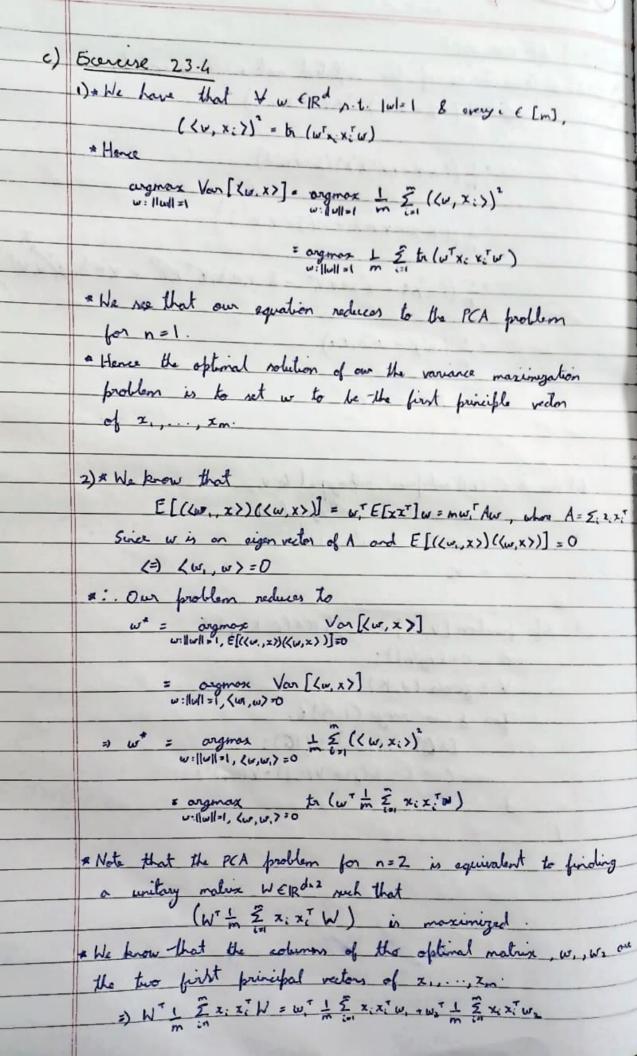
## 6.2) Representation Learning

1) of Exercise 23.1

1) . We know that if V, W are finite dimensional vector spaces, & T is a bien transformation from V to W, Then the image of T is a finite-dimensional subspace of W and dim (V) = dim (mell (T)) + dim (emage (T))

- :. for dim(null(A)) 7,1, 7 v +0 s.t. Av = Ao = 0.
  - · . 7 u + v ER nt. A=Av.
- 2) We have that A = Ax for some u + v ER.
  - =) For any necovery function of, ((Au) = 6 (Av)
  - . Exact recovery of a linear compression rehere is impossible.

b) Exercise 23.3 \* Let x be a matrix with its j-th column as y(xi). We find the spectral decomposition of X'X, using the results from section 23.1.1 (A more efficient solution for the case of >> m). · Note that (x'x) ij = K(xi, xi), thus the eigendecomposition of XTX can be found in polynomial time. \* Let V be the matrix whose solumns are the aleading eigenvectors of X \* X, and let D be a diagonal nx o matrix whose digonal consits of the corresponding eigenvalues. \* Denote by U be the matrix whose columns are the n Looding eigenvectors of XXT. We now how how to \* We now show how to project the data without maintaining the matrix U. For every x EX, the projection UT d(x) is calculated as follows:  $V^{T}\phi(x) = D^{\frac{1}{2}}V^{T}\chi^{T}\phi(x) = D^{\frac{1}{2}}V^{T}\begin{pmatrix} K(x_{i}, x) \\ K(x_{i}, x) \end{pmatrix}$ 



with \$\frac{\pi}{m} \text{ and } w, are orthonormal, we obtain that

with \$\frac{\pi}{m} \text{ x; x; w, w, w, \frac{\pi}{m}} \frac{\pi}{m} \text{ x; x; w, w, \frac{\pi}{m}} \frac{\pi}{m} \text{ x; x; w, \w, \w, \frac{\pi}{m}} \frac{\pi}{m} \text{ x; x; w, \w, \w, \frac{\pi}{m}} \frac{\pi}{m} \text{ x; x; w.} : we conclude that w\* = w2.

```
Hybre Exercise 20.5
(205) a) Covarience of the deflated matrix is given by
                                                          = 1 ((xx-x,v, xx)(I-x,v, T))
                                                          = 1 ((xTx - v, n x, v, T)(I - v, v, T))
                                                           = 1 ((xxx - xxxv,v, - x v,nh,v, - x v,nh,v
                                                           = 1 (xx - n 1, v, v)
                                                              = 1 X X - 1, V, V, T
        b) As X & (d-1) subspace orthogonal to v,
                                                   =) u must be orthogonal to v.
                            :. => u v = 0 & u u = 1
                                  . . u = V2
       a) def function [V, X] = simple PCA (C, K, 6) {
                                                d = c. length();
                                                 V = zeroes (d, K);
                                                  for is in range (1, K) (
                                                                            [x(j), v(:,i)] = ((6);
                                                                                C = C - X(j) * V(:, j) * V(:, j) ,
```