

(3) OPTIMIZATION FORM 4

X E Rdx4n minimize $\| X - \phi \phi^T X \|_{\epsilon}^2$ O & IRaxd

St. PT = I Grobenius Norm

here X is the matrix from the optimization problem of is the matrix consisting of columns of eigenvectors

the objective for can be simplified as:

 $|| \times - \phi \phi^{\mathsf{T}} \times ||_{\mathsf{F}}^{2} = \mathsf{tr} \left(\mathsf{x}^{\mathsf{T}} \mathsf{x} - \mathsf{X} \mathsf{x}^{\mathsf{T}} \phi \phi^{\mathsf{T}} \right)$ Frace of matrix

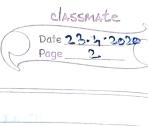
its lagrangian is (for) - 2 gor)

 $L = tr(X^TX) - tr(X^TX^T\varphi\varphi^T)$ $- tr(\Lambda^T(\varphi^T\varphi - I))$ - eigenvalue matrix Rolad

equating derivative of L to 0 gives:

 $\frac{\partial C}{\partial \phi} = 2 \times x^{\mathsf{T}} \phi - 2 \phi \Lambda = 0$

> XXTO = ON => AO = ON > eigenvalue problem this can be useful in machine in vector form learning, when we wante to consider several PCA directions of dimensionality reduction



FISHER DISCRIMINANT ANALYSIS

this is a dimensionality reduction technique

which attempts to classify a set of points in reduced dimensions. it tries to keep the mean of both classes as separate as possible, while trying to minimize variance within a class.

Objective: max. wTSB w min, variance within.

$$w^{T}Sww \leftarrow min$$
, variance with $w = projection surjace$

$$S_{B} = \sum_{i=1}^{E} (\mu_{i} - \mu_{i}) (\mu_{i} - \mu_{i})^{T}$$

 $So = \sum_{i=1}^{c} \sum_{i=1}^{n_i} (x_{i,i} - \mu_i) (x_{i,i} - \mu_i)^T$

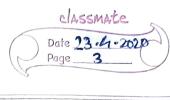
this objective can be restated as:

$$\text{st. } \omega^{\mathsf{T}} \mathbf{s}_{\omega} \omega = 1$$

finding its lagrangian, and equating its derivative to 0 yields:

 $S_B \omega = \pi S_\omega \omega$ $\lambda = lagrange moltiple$

problem form, we can now one of the solution methods to obtain w. (projection plane).



6 LLE PSEUDO CODE

M = (I-w) (I-w)

A select K nearest neighbours for each point for each point X: i= 1-N

B. find reconstruction weights for each point

Z = matrix of all neighbours of X;

Subtract X; from every column of Z.

solve Cw = 1 for wset $w_{ij} = 0$ if j is not neighbour of i.

remaining elements in row W; normalized = W/sum(w)

C. map to target dimension (Y using weights W)

find bottom d+1 eigenvectors of M.

(corresponding to d+1 smallest eigenvalues)

Set the ith row of Y to be i+1th smeigenvector

(diseard eigenvector [:, s, ...] with egenvalue o)

	Date 23.4.2008 Page 4
6	line: 21-y=1
	100 points: (1,0), (2,1), (3,2)
	the output is computed in python.
	Covariance matrix:
	842 842
	842 842
<u>.</u>	The same of the state of the st
	eigenvectors:
	The same of the sa
(•	0.707 0.707]

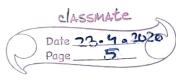
2. [-0.707 0.707] ~ 0

the main eigen vector is the first one

which is parallel with respect to the

Slope line.

(eigen vectors are compoted from the covariance matrix)



(7)	(A) initialization: 626.1		
	(B) it. 1 : 16.5		
	(c) it. 2:16.5	2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	
	3 3		

(8) (A) if	we in	crease	K	=>	we	allow	more	classes
50	error	an c	decre	ase	bur	ther. C	gree.	

(c) normalized objective (error) would decrease too, (but objective * k might not change).

(B) the max value of K can be N (no. of paints).

then objective (error) would be O.