(3) OPTIMIZATION FORM 4

$X \in \mathbb{R}^{d \times n}$

minimize $\| X - \phi\phi^T X \|_F^2$

$\phi \in \mathbb{R}^{d \times d}$

$\rightarrow$ frobenius norm.

st. $\phi^T \phi = I$

here X is the matrix from the optimization problem
ϕ is the matrix consisting of colums of eigenvectors

the objective fn. can be simplified as:

$$\| X - \phi\phi^T X \|_F^2 = tr(X^T X - X X^T \phi\phi^T)$$

$\hookrightarrow$ trace of matrix.

its lagrangian is $(f(x) - \lambda g(x))$

$$\mathcal{L} = tr(X^T X) - tr(X X^T \phi\phi^T) - tr(\Lambda^T(\phi^T\phi - I))$$

$\hookrightarrow$ eigen value matrix $\mathbb{R}^{d \times d}$

equating derivative of $\mathcal{L}$ to 0 gives:

$$\frac{\partial \mathcal{L}}{\partial \phi} = 2 X X^T \phi - 2 \phi \Lambda = 0$$

$$\Rightarrow X X^T \phi = \phi \Lambda \quad \Rightarrow \underline{A\phi = \phi\Lambda}$$

$\hookrightarrow$ eigen value problem in vector form.

this can be useful in machine learning, when we want to consider several PCA directions of dimensionality reduction.

(5) 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.

$$\text{Objective: } \max. \ \frac{\overbrace{w^T S_B w}^{\text{max. mean sep.}}}{\underbrace{w^T S_w w}_{\text{min. variance within.}}}$$

w = projection surface

$$S_B = \sum_{j=1}^{c} (\mu_i - \mu_t)(\mu_i - \mu_t)^T$$

$$S_w = \sum_{j=1}^{c} \sum_{i=1}^{n_j} (x_{j,i} - \mu_i)(x_{j,i} - \mu_i)^T$$

this objective can be restated as:

$$\max. \ w^T S_B w$$
$$\text{st. } w^T S_w w = 1.$$

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

$$S_B w = \lambda S_w w \qquad \lambda = \text{lagrange multiplier}$$

- since this is now in generalized eigenvalue problem form, we can now one of the solution methods to obtain $\underline{\underline{w}}$. (projection plane).

(5) LLE PSEUDOCODE

A. select K nearest neighbours for each point
for each point X_i $\quad i = 1 - N$

B. find reconstruction weights for each point
Z = matrix of all neighbours of X_i
Subtract X_i from every column of Z.
local covariance $C = Z^T Z$
solve $Cw = 1$ for w
set $w_{ij} = 0$ if j is not neighbour of i.
remaining elements in row w_i normalized = $w/\text{sum}(w)$

C. map to target dimension (Y using weights W)
$M = (I-W)^T (I-W)$
find bottom $d+1$ eigenvectors of M.
(corresponding to $d+1$ smallest eigenvalues)
Set the i^{th} row of Y to be $i+1^{th}$ ~~sm~~ eigenvector
(discard eigenvector $[1, 1, 1, \ldots]$ with eigenvalue 0)

(6) line: $x - y = 1$

100 points: $(1,0), (2,1), (3,2), \ldots$

the output is computed in python.

covariance matrix:

$$\begin{bmatrix} 842 & 842 \\ 842 & 842 \end{bmatrix}$$

eigen vectors:

1. $[0.707 \quad 0.707]$ 1683

2. $[-0.707 \quad 0.707]$ ~ 0

the main eigen vector is the first one which is parallel with respect to the slope line.

(eigen vectors are computed from the covariance matrix)

7 (A) initialization : 626.1

(B) it. 1 : 16.5

(C) it. 2 : 16.5

⋯3

& conv.

8 (A) if we increase K ⇒ we allow more classes so error can decrease further. agree.

(B) the max. value of K can be N (no. of points). then objective (error) would be 0.

(C) normalized objective (error) would decrease too. (but objective * K might not change).

1. find the graph distance (distance on the nearest neighbour graph) between points x[0] and x[784] for

$K = 1$: -1

$K = 2$: 393.914

$k = 3$: 393.839

$k = 5$: 393.669

true distance: 393.955

euclidean distance: 41.505

x $10\cos(0 - 0.784 \times 50)$
y $10\sin(0 - 0.784 \times 50)$
z: $t(0 - 0.784 \times 50)$

$x = R\cos t$ $\qquad \frac{dx}{dt} = -R\sin t$

$y = R\sin t$

$z = Ht$ $\qquad \frac{dy}{dt} = R\cos t$

$\frac{dH}{dt} = H$

$$S = \int_{x_1}^{x_2} \sqrt{dx^2 + dy^2 + dz^2}$$

$$= \int_{x_1}^{x_2} \sqrt{(R\sin t)^2 + (R\cos t)^2 + H^2}\, dt$$

$$= \int_{x_1}^{x_2} \sqrt{R^2(\sin^2 t + \cos^2 t) + H^2}\, dt$$

$$= \int_{x_1}^{x_2} \sqrt{R^2 + H^2}\, dt.$$

$$= \int \sqrt{1001}\, dt$$

$$\left[\sqrt{1001}\, t\right]_{x_1}^{x_2}$$

$$= \sqrt{1001}\; 50\left(0.784 - 0\right)$$

$$= 393.9455$$