Note Title 8/10/2022

ISSUES / CONCERNS with PCA

- TIME COMPLEXITY – Finding the Eigen vectors and Eigen values.

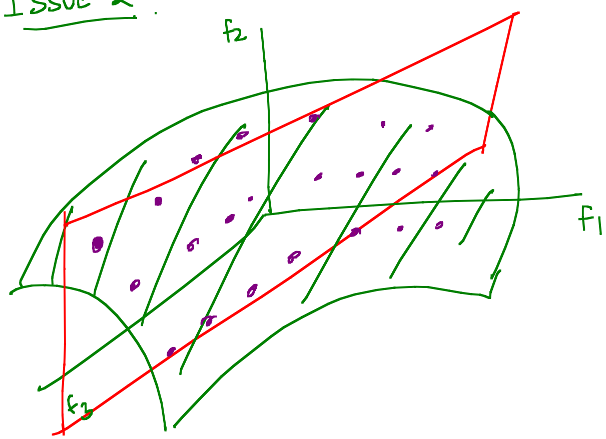
  $C \in \mathbb{R}^{d \times d}$ Typically $O(d^3)$

- Issue when d is large

  Example: Face recognition (Eigen Faces)

Note Title 8/10/2022

- ISSUE 2 :



Data may not necessarily live in a low-dimensional LINEAR subspace.

SURPRISING RESULT : Same Solution to both issues!

Note Title 8/10/2022

Issue 1: Large d

$[d >> n]$

$d \to$ # features

$n \to$ # datapoints.

$$X = \begin{bmatrix} | & | & | & & | \\ x_1 & x_2 & x_3 & \cdots & x_n \\ | & | & | & & | \end{bmatrix} \qquad \underline{X} \in \mathbb{R}^{d \times n}$$

$$\underline{C = \frac{1}{n}\left(\sum_{i=1}^{n} x_i x_i^T\right)}$$

$$XX^T = \begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix} \begin{bmatrix} — & x_1 & — \\ & \vdots & \\ — & x_n & — \end{bmatrix} = \sum_{i=1}^{n} x_i x_i^T$$

exercise: [Show this]

$$\Rightarrow \boxed{C = \frac{1}{n} XX^T}$$

Note Title 8/10/2022

Let w_k be the eigenvector corresponding to the k^{th} largest eigenvalue of C (λ_k) ← e.value

$$C w_k = \lambda_k w_k \quad \text{[by definition]}$$

$$\left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T\right) w_k = \lambda_k w_k$$

$$\boxed{w_k = \sum_{i=1}^{n} \left(\frac{x_i^T w_k}{n\lambda_k}\right) \cdot x_i}$$

w_k is a LINEAR COMBINATION of data points!

Note Title 8/10/2022

$$\begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix} \begin{bmatrix} \alpha_{k1} \\ \vdots \\ \alpha_{kn} \end{bmatrix} = \sum_{i=1}^{n} \alpha_{ki} x_i$$

$w_k = X \alpha_k$ for some $\alpha_k \in \mathbb{R}^n$

$\boxed{\text{How to get } \alpha_k \text{ ?}}$

Some Algebra:

$w_k = X \alpha_k \quad \forall k$

$C w_k = \lambda_k w_k$

$\left(\frac{1}{n} X X^T\right)(X \alpha_k) = \lambda_k X \alpha_k$

Note Title 8/10/2022

$$(XX^T)X\alpha_k = n\lambda_k X\alpha_k$$

Pre multiply by X^T

$$X^T((XX^T)X\alpha_k) = X^T(n\lambda_k X\alpha_k)$$

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

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

$$(X^TX)(X^TX)\alpha_k = n\lambda_k (X^TX)\alpha_k$$

Call $X^TX := K$

$$K^2\alpha_k = n\lambda_k K\alpha_k$$

If we can find α_k that satisfies

$$\boxed{K \alpha_k = (n \lambda_k) \alpha_k} \leftarrow$$

$\hookrightarrow$ Eigen equation

We know $\quad w_k = X \alpha_k$

$$w_k^T w_k = (X \alpha_k)^T (X \alpha_k) = \alpha_k^T (X^T X) \alpha_k$$

$$\boxed{1 = \alpha_k^T K \alpha_k}$$

Note Title 8/10/2022

LINEAR ALGEBRA FACT : The non-zero eigenvalues of $\underline{XX^T}$ and $\underline{X^TX}$ are exactly the same !

$XX^T \to \mathbb{R}^{d \times d}$, $X^TX \to \mathbb{R}^{n \times n}$.

$C = \frac{1}{n} XX^T$ Eigenvectors = $\{w_1, \ldots, w_\ell\}$

$\forall k, \|w_k\|^2 = 1$

Eigen values = $\{\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell\}$

$\underline{XX^T} = nC$ Eigenvectors = $\{w_1, \ldots, w_\ell\}$

Note Title 8/10/2022

Eigenvalues = $\{ n\lambda_1 \geq n\lambda_2 \geq \cdots \geq n\lambda_\ell \}$

$$\frac{X^T X}{K}$$

Eigenvectors = $\{ \beta_1, \ldots, \beta_\ell \}$

$\|\beta_k\|^2 = 1 \quad \forall k$

Eigenvalues = $\{ n\lambda_1 \geq n\lambda_2 \cdots \geq n\lambda_\ell \}$

$$\boxed{K \beta_k = (n\lambda_k)\, \beta_k}$$

IS $\beta_k = \alpha_k$?

Note Title 8/10/2022

$$\underline{\beta_k^T K \beta_k} = \beta_k^T (n\lambda_k \beta_k) = n\lambda_k \underline{\beta_k^T \beta_k}$$

$$= \underline{n\lambda_k}$$

Set $\forall k$ $\alpha_k := \dfrac{\beta_k}{\sqrt{n\lambda_k}}$

Now

① $K\alpha_k = (n\lambda_k)\alpha_k$

② $\alpha_k^T K \alpha_k = 1$

$$= \frac{\beta_k^T K \beta_k}{n\lambda_k} = \frac{n\lambda_k}{n\lambda_k} = 1$$

Note Title 8/10/2022

Input : $D = \{x_1, \dots, x_n\}$ $x_i \in \mathbb{R}^d$ $[\; d >> n \;]$

$\forall i, j$

$$\boxed{K_{ij} = x_i^T x_j}$$

- Step 1 : Compute $\boxed{K = X^T X}$ $K \in \mathbb{R}^{n \times n}$
- Step 2 : Compute eigendecomposition of K
  e.vectors $\{\beta_1, \dots, \beta_\ell\}$ corresponding to e.values $\{n\lambda_1, \dots, n\lambda_\ell\}$ $\boxed{O(n^3)}$
- Step 3 set $\alpha_k = \dfrac{\beta_k}{\sqrt{n\lambda_k}}$ $\forall k = 1, \dots, \ell$
- Step 4, $\boxed{w_k = X\alpha_k}$ $\forall k.$