Principal component Analysis.

Suppose each data point is stored in a p-dimensional vector. and there are N data pts. The

One may take a data matrix $X = (\vec{x}_1\ \vec{x}_2 \cdots \vec{x}_N) \in \mathbb{R}^{p\times N}$.

<u>Mean:</u> sample mean $M = \frac{1}{N}(\vec{x}_1 + \cdots + \vec{x}_N)$

<u>normalled data:</u> take $\hat{x}_i = \vec{x}_i - M$ and take normalized data matrix $B = (\hat{x}_1 \cdots \hat{x}_N)$.

$\Rightarrow$ The normalized data set has mean zero.

<u>Covariance matrix:</u> $S = \frac{1}{N-1} \underset{p\times N}{B}\underset{N\times p}{B^T} = (S_{ij}). \in \mathbb{R}^{p\times p}$

For $i\neq j$: S_{ij} is called the covariance of x_i and x_j; if $S_{ij}=0$, x_i and x_j is called uncorrelated.

S_{ii} is called the variance of x_i. $\sum_{i=1} S_{ii}$ = total variance = $\mathrm{Tr}(S)$.

<u>Goal:</u> Find uncorrelated directions and get the maximal variance directions.

Linear Algebra: S is symmetric $\Rightarrow$ By Spectral Thm $\exists$ orthogonal matrix P and diagonal matrix $D \in \mathbb{R}^{p\times p}$

such that $S = PDP^T$. In particular we arrange D such that $D = \begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_p\end{pmatrix}$ w/ $\lambda_1 \geq \cdots \geq \lambda_p$.

$(\Rightarrow P^TSP = D \Rightarrow P^TBB^TP = D)$

$(\Leftarrow$ Find SVD of B^T: $B^T = U\Sigma V^T$. $\Rightarrow S = \frac{1}{N-1} V\Sigma U^TU\Sigma V^T = \frac{1}{N-1} V\Sigma^2 V^T.)$

Then $P = V$, $D = \frac{1}{N-1}\Sigma^2$.)

Denote $P = (\vec{u}_1 \cdots \vec{u}_p)$, Then $\vec{u}_1, \vec{u}_2, \cdots \vec{u}_p$ are called the principal components of the data.

orthogonal change of basis.

Take $Y = P^TX$. Take $C = P^T(\hat{x}_1 \cdots \hat{x}_N) = P^TB$

<u>Then:</u> $S_{new} = \frac{1}{N-1} CC^T = \frac{1}{N-1} P^TBB^TP = P^TSP = P^T(PDP^T)P = D$

$\Rightarrow$ y_i and y_j are uncorrelated for any $i\neq j$.

<u>Remarks:</u> · The total variance of $x_1, \cdots, x_p$ = total variance of $y_1, \cdots, y_p$.

$= \lambda_1 + \cdots + \lambda_p$.

- Principal components analysis is potentially valuable for applications in which most of the variation (indicated by the total variance) or dynamic change is due to variations in only the first few of the new variable of $y_1, \cdots y_p$. (Often times the first few of $\lambda_1, \cdots, \lambda_p$ takes a high percentage of $\lambda_1 + \cdots + \lambda_p$).

Jordan Canonical Form.

A matrix $J \in F^{n\times n}$ is called in Jordan canonical form if it is a block diagonal matrix in the following form: $J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}$ where $J_i \in F^{d_i\times d_i}$ is a Jordan block of the form $J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix} \in F^{d_i\times d_i}$.

<u>Main Theorem</u> 1). Let $A \in \mathbb{C}^{n\times n}$. Then there exists an invertible matrix P and and J in Jordan canonical form such that $A = PJP^{-1}$.

2). Let V be a finite dimensional vector space. Let $T: V\to V$ be a linear transformation. Then there exists a basis B of V such that $[T]_B$ is in Jordan canonical form.

Definition: Let $T: V\to V$ be a linear transformation. A non-zero vector $\vec{x}\in V$ is called a generalized eigenvector of T if there exists $\lambda \in F$ and a positive integer $p\in \mathbb{Z}^+$ such that

$$(T-\lambda I)^p \vec{x} = \vec{0}.$$

The set $\{\vec{0}\} \neq K_\lambda = \{\vec{x}\in V: (T-\lambda I)^p\vec{x} = \vec{0} \text{ for some } p\in \mathbb{Z}^+\}$ is called the generalized eigenspace of T associated with eigenvalue λ

Proposition: K_λ is a subspace of V.

<u>Proof</u>: ① $\vec{0}\in K_\lambda$

② $\forall \vec{x}, \vec{y} \in K_\lambda$: $(T-\lambda I)^p(\vec{x}) = \vec{0}$ and $(T-\lambda I)^q(\vec{y}) = \vec{0}$ for some $p, q\in \mathbb{Z}^+$

$\forall c\in F$: $(T-\lambda I)^{p+q}(c\vec{x}+\vec{y}) = c(T-\lambda I)^{p+q}(\vec{x}) + (T-\lambda I)^{p+q}(\vec{y}) = \vec{0}$.

$\Rightarrow c\vec{x}+\vec{y} \in K_\lambda$.

Remarks: 1). $E_\lambda \subseteq K_\lambda$.

2). For $\vec{0}\neq \vec{x}\in K_\lambda$, let $p_x \in \mathbb{Z}^+$ be the smallest positive integer such that $(T-\lambda I)^{p_x}(\vec{x}) = \vec{0}$.

$\Rightarrow$ define $\vec{y} = (T-\lambda I)^{p_x-1}(\vec{x}) \neq \vec{0}$

Then $(T-\lambda I)\vec{y} = (T-\lambda I)^{p_x}(\vec{x}) = \vec{0}$

$\Rightarrow \vec{y} \in E_\lambda$.

Example Let $A \in \mathbb{C}^{6\times 6}$. and $P \in \mathbb{C}^{n\times n}$ such that $A = PJP^{-1}$ where
$$J = \left(\begin{array}{ccc|c|cc} 2&1&0&0&0&0\\ 0&2&1&0&0&0\\ 0&0&0&0&0&0\\ \hline 0&0&0&2&0&0\\ \hline 0&0&0&0&3&1\\ 0&0&0&0&0&3 \end{array}\right)_{6\times 6}$$

Then $f(x) = \det(A-\lambda I) = \det(J-\lambda I) = (2-\lambda)^4(3-\lambda)^2$.

Eigenvalues: $\lambda = 2$ (multiplicity $=4$), $\lambda = 3$ (multiplicity $=2$).

Denote $P = (\vec{p}_1\ \vec{p}_2\ \vec{p}_3\ \vec{p}_4\ \vec{p}_5\ \vec{p}_6)$. Then $A = PJP^{-1} \Leftrightarrow AP = PJ$

$$\Leftrightarrow A\vec{p}_1 = 2\vec{p}_1 \Leftrightarrow (A-2I)\vec{p}_1 = \vec{0} \Rightarrow \vec{p}_1 \in Null(A-2I)$$
$$A\vec{p}_2 = \vec{p}_1 + 2\vec{p}_2 \Leftrightarrow (A-2I)\vec{p}_2 = \vec{p}_1 \Rightarrow \vec{p}_2 \in Null(A-2I)^2$$
$$A\vec{p}_3 = \vec{p}_2 + 2\vec{p}_3 \Leftrightarrow (A-2I)\vec{p}_3 = \vec{p}_2 \Rightarrow \vec{p}_3 \in Null(A-2I)^3$$
$$A\vec{p}_4 = 2\vec{p}_4 \Leftrightarrow (A-2I)\vec{p}_4 = \vec{p}_4 \Rightarrow \vec{p}_4 \in Null(A-2I)$$
$$A\vec{p}_5 = 3\vec{p}_5 \Leftrightarrow (A-3I)\vec{p}_5 = \vec{0} \Rightarrow p_5 \in Null(A-3I)$$
$$A\vec{p}_6 = \vec{p}_5 + 3\vec{p}_6 \Leftrightarrow (A-3I)\vec{p}_6 = \vec{p}_5 \Rightarrow p_6 \in Null(A-3I)^2$$

Remarks: 1) $\{\vec{p}_1, \vec{p}_2, \cdots, \vec{p}_6\}$ is a set of generalized eigenvectors

2). $\{\vec{p}_1, \vec{p}_2, \vec{p}_3, \vec{p}_4\} \subseteq K_2$. $\dim K_2 = 4 =$ multiplicity of $\lambda = 2$?

. $\{p_5, p_6\} \subseteq K_3$, $\dim K_6 = 2 =$ multiplicity of $\lambda = 3$?

3). $\{\vec{p}_1, \vec{p}_2, \vec{p}_3, \vec{p}_4\} \subseteq Null(A-2I)^3$ $\quad K_2 = Null((A-2I)^4)$?

$\{\vec{p}_5, \vec{p}_6\} \subseteq Null(A-3I)^2$ $\quad K_4 = Null(A-3I)^2$?

4). # of Jordan blocks with diagonal $\lambda = 2$ $= 2 = \dim E_2$

# of Jordan blocks with diagonal $\lambda = 3$ $= 1 = \dim E_3$. // end of March 5

Conjectures: Let $T: V \to V$ be a linear transformation.

1). There exists a basis of V which consists of generalized eigenvectors of T.