EECS 127 Lecture 3.

Today:
- PCA
- Symmetric Matrices.
- Spectral thm.

Admin
- Midterm, March 12. Keep 5-9pm free.
- Enrollment.
- Discussion Sections.

Principal Component Analysis.

- High-dimensional data
- "Dimesionality reduction"

p-dimensional data

"Maximize" the variance of the data.

$$\max_{\|\vec{w}\|=1} \vec{w}^T C \cdot \vec{w}$$

$\vec{x}_1, \vec{x}_2 \cdots \vec{x}_n \in \mathbb{R}^p$ (zero-mean)

Uncover underlying lower-dimensional structure.

$(x, y) \rightarrow (x, y, z)$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \; \end{bmatrix}$$

Goal: Project our data into lower-dimensin and recorer low-dim. structure.

Project: $\vec{w} \in \mathbb{R}^p$ s.t. the projected vectors are as close to the original vectors as possible. $\|\vec{w}\|_2^2 = 1$.

$\vec{x}_1, \vec{x}_2 \ldots \vec{x}_n$: **Projections:** $\langle \vec{x}_i, \vec{w} \rangle \cdot \vec{w}$

Error: $\| \vec{x}_i - \langle \vec{w}, \vec{x}_i \rangle \vec{w} \|^2 = e_i^2$ on $\vec{x}_i$

$$\boxed{\text{Average proj. error: } \quad \frac{1}{n} \sum_{i=1}^{n} e_i^2 = \text{MSE}(\vec{w})}$$

Zero-mean

What is the mean of the projection?

$$\frac{1}{n} \sum_{i=1}^{n} \langle \vec{x}_i, \vec{w} \rangle \vec{w} = \left(\frac{1}{n} \sum_{i=1}^{n} \vec{x}_i^T \vec{w} \right) \vec{w}$$

$$= \left(\underbrace{\left(\frac{1}{n} \sum \vec{x}_i^T \right)}_{0} \vec{w} \right) \vec{w}$$

$$= 0$$

$$\| \underbrace{\vec{x}_i - \langle \vec{w}, \vec{x}_i \rangle \vec{w}}_{\vec{v}_i} \|^2 = (\vec{v}_i^{\,T} \cdot \vec{v}_i)$$

$$= (\vec{x}_i - \langle \vec{w}, \vec{x}_i \rangle \vec{w})^T (\vec{x}_i - \langle \vec{w}, \vec{x}_i \rangle \cdot \vec{w}).$$

$$= \|\vec{x}_i\|^2 - 2\langle \vec{w}, \vec{x}_i \rangle \cdot \langle \vec{w}, \vec{x}_i \rangle + \langle \vec{w}, \vec{x}_i \rangle^2 \cdot \|\vec{w}\|_2^2$$

$$= \|\vec{x}_i\|^2 - \langle \vec{w}, \vec{x}_i \rangle^2.$$

$$MSE(\vec{w}) = \frac{1}{n} \sum \|\vec{x}_i\|^2 - \frac{1}{n} \sum \langle w, \vec{x}_i \rangle^2.$$

$$E[X^2] = E[X]^2 + Var(X)$$

Goal: **Maximize:** $\frac{1}{n} \sum \langle \vec{w}, \vec{x}_i \rangle^2.$

$$= \Big(\underbrace{\frac{1}{n} \sum \langle \vec{x}_i, \vec{w} \rangle}_{0}\Big)^2 + Var(\langle \vec{w}, \vec{x}_i \rangle).$$

Data is zero mean.

$$\frac{1}{n}\sum \langle \vec{w}, \vec{x}_i \rangle^2 = \frac{1}{n} \| X \cdot \vec{w} \|^2$$

$$= \frac{1}{n} (X\vec{w})^T (X\vec{w})$$

$$X = \underbrace{\begin{bmatrix} - \vec{x}_1^T - \\ - \vec{x}_2^T - \\ \vdots \\ - \vec{x}_n^T - \end{bmatrix}}_{n \times p} \qquad \vec{w} = \begin{bmatrix} \\ \\ \\ \end{bmatrix}_{p \times 1}$$

$$= \frac{1}{n} \vec{w}^T X^T X \vec{w}.$$

$$= \vec{w}^T \cdot \underbrace{\left(\frac{X^T X}{n}\right)}_{C} \cdot \vec{w}$$

$$\boxed{\begin{array}{l} \text{maximize} \quad \vec{w}^T \cdot C \cdot \vec{w} \\ \|\vec{w}\|_2^2 = 1 \end{array}}$$

where C is the covariance matrix of your data.

$$C = \begin{bmatrix} \sum x_{i1}^2 & \sum x_{i1} x_{i2} & & \\ \sum x_{i1} x_{i2} & \sum x_{i2}^2 & & \\ & & \ddots & \\ & & & \sum x_{ip}^2 \end{bmatrix}$$

Symmetric matrices

- Matrix $A \in S^n$ if $A^T = A$.

$A \in \mathbb{R}^{n\times n}$ $\quad A_{ij} = A_{ji} \quad \forall\, i,j\,,\ 1 \le i,j \le n.$

e.g. Graph Laplacian

Diagonalization of Matrices

Condition for diagonalization: "Algebraic" multiplicity" = "Geometric" multiplicity.

$\det(A-\lambda I)$ → # of roots $= \lambda$. $\qquad \dim(N(A-\lambda I))$

e.g. $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ 0 \end{bmatrix}$ $\qquad$ $N(A)$ dim 1.

e.g. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

- Properties of Symmetric Matrices. (C.EG: Thm 4.1).

  - $A \in \mathbb{R}^{n\times n}$, S^n is a symmetric matrix.
  - $\lambda_1, \lambda_2, \cdots \lambda_k$ are all eigenvalues.
  - $\mu_1, \mu_2 \cdots \mu_k$ are algebraic multiplicities.
  - $\phi_i = N(A - \lambda_i I)$.

Then:

1. $\lambda_i \in \mathbb{R}$
2. $\phi_i \perp \phi_j$
3. $\dim(\phi_i) = \mu_i$

"Spectral" Thm.

$A = U \Lambda U^T$

U: orthonormal matrix.

Λ: diagonal.

Proof of 3:

Lemma: $(\lambda, \vec{u})$ eigenpair for A. Then, there exists an orthornmal U such that.

$$U^T A U = \begin{bmatrix} \lambda & 0 \\ 0 & \boxed{B} \end{bmatrix}$$

$B \in S^{n-1}$

Proof:

$$U = \left[\begin{array}{c|c} \vec{u} & \underbrace{U_1}_{\text{Gram-Schmidt}} \end{array}\right]$$

Gram-Schmidt.

$$U^T A U = \left[\begin{array}{c} \vec{u}^T \\ \hline \\ \\ \end{array}\right] \left[\begin{array}{c} \\ A \\ \\ \end{array}\right] \left[\begin{array}{c|c} \vec{u} & U_1 \end{array}\right]$$

$$= \left[\begin{array}{c} \vec{u}^T \\ \hline U_1^T \end{array}\right] \left[\begin{array}{c|c} \lambda\vec{u} & AU_1 \end{array}\right]$$

$$= \left[\begin{array}{c|ccc} \lambda & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & U_1^T A U_1 & \\ 0 & & & \end{array}\right]$$

$B = U_1^T A U_1$

$B^T = B.$

$\therefore$ B is ~~g~~ symmetric. $\blacksquare$

Proceed by Induction.