9. Symmetric Matrices, SVD and PCA

Symmetric Matrices:

If $A = A^T$ we say A is symmetric

$$A = \begin{bmatrix} 1 & 4 & 6 \\ 4 & 2 & 5 \\ 6 & 5 & 3 \end{bmatrix}, \quad i \neq j \Rightarrow a_{ij} = a_{ji}$$

Diagonalizing symmetric Matrices:

$$A = \begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}$$

1. Find eig. vals:

$$\begin{vmatrix} 1-\lambda & 5 \\ 5 & 1-\lambda \end{vmatrix} = ((1-\lambda)(1-\lambda)) - 25 = \lambda^2 - 2\lambda - 24$$
$$= (\lambda - 6)(\lambda + 4) \to \lambda = \begin{cases} 6 \\ -4 \end{cases}$$

2. Find eig vecs.

For $\lambda = 6$:

$$\begin{bmatrix} -5 & 5 \\ & \end{bmatrix} \sim \begin{bmatrix} 1 & -1 \\ & \end{bmatrix} \to \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

For $\lambda = -4$

$$\begin{bmatrix} 5 & 5 \\ & \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ & \end{bmatrix} \to \vec{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

3. Create P: $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

4. Find D $\begin{bmatrix} 6 & 0 \\ 0 & -4 \end{bmatrix}$

Theorem:

If A is symmetric, any two eigenvectors from distinct eigenvalues are orthogonal.

Theorem:

If Q is an orthogonal matrix, $Q^{-1} = Q^T$

Definition:

An $n \times n$ matrix A is said to be orthogonally diagonalizable if

$$A = PDP^T = PDP^{-1}$$

Theorem:

An $n \times n$ Matrix A is orthogonally diagonalizable iff. $A = A^T$

Essential difference to "regular" diagonalization:

↳ The Columns of P are orthonormal.

Method: Orthogonal Diagonalization:

1) Find eigenvalues of A.
2) Find basis vectors for the eigenspaces corresponding to each eigenvalue found in (1)
3) Do Gram Schmidt, if necessary, on eigenvectors found in (2)
4) Normalize all vectors in (2+3) and setup P and P^T
5) Write D

Eigenvalues aka spectrum

Theorem (Spectral Theorem)

If $A = A^T$

1) A has n real eig.val. (counting multiplicities)
2) $\dim \operatorname{Nul} A - \lambda_i I$ = alg. mult. of λ_i
3) The eigenspaces are mutually orthogonal
4) A is orthogonally diagonalisable.

Singular value Decomposition:

Problem: I can't go around hoping that that all the matrices that I encounter will be symmetric. I need a method to that will work for <u>any</u> $m \times n$ matrix.

Solution: Singular Value Decomposition allows us to decompose any $m \times n$ Matrix in to three matrices, one of which is a diagonal matrix.

The Decomposition:

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V^T_{n \times n}$$

Normalized, cf. above.

- V consists of $\boxed{\text{eig. vecs}}$ of A^TA (right singular vectors)
- U consists of $\boxed{\text{eig. vecs}}$ of AA^T (left singular vectors)
- Σ is a diagonal matrix with the singular values of A on its diagonal. The singular values are the square roots of the eigenvalues of A^TA / AA^T, and are placed in decreasing order of magnitude. The corresponding vectors in U and V are placed in corresponding columns.

Important: You **must** either find V and then from this derive the vectors that make up U **OR** find U and then from this derive the vectors that make up V.

> **Don't DO** $U = AAt.\text{eigenvects}()$
> $V = AtA.\text{eigenvects}()$
> and use both.

$$A = USV^T = USV^{-1} \iff AV = US$$

If we take one column at a time, we get:

$$A\bar{v}_1 = \sigma_1 \cdot \bar{u}_1$$

$$A\bar{v}_2 = \sigma_1 \cdot \bar{u}_2$$

$$A\bar{v}_3 = \sigma_1 \cdot \bar{u}_3$$

$$\vdots$$

$$A\bar{v}_n = \sigma_n \cdot \bar{u}_n$$

So we can find $\bar{u}_i$:

$$\bar{u}_i = \frac{1}{\sigma_i} \cdot A\bar{v}_i = \sigma_i^{-1} \cdot A\bar{v}_i$$

Note: We run into problems when

$\sigma_i = 0$

and

$m > n$

Ex:

$$A_{5\times2} = U_{5\times5}\, S_{5\times2}\, V^T_{2\times2}$$

If we depart from A^TA, we will only get 2 eigenvectors, and will need to use the "trick" to obtain the three other vectors + GramSchmidt.

If we depart from AA^T we will get 5 eigenvectors.

Protip:

If $m>n$ (more rows than columns)

shape(U)[0] > shape(V)[0]

If we depart from A^TA we get too few eig. vecs. Depart from AA^T.

If $n>m$ (more columns than rows)

shape(U)[0] < shape(V)[0]

If we depart from AA^T we get too few eig. vecs. Depart from A^TA.

Method:

1) Find eig. vals of

   a) AA^T if $m > n$

   b) A^TA if $m < n$

2) Depending on choice in (1), find eig. vecs of and normalize them.

   a) $AA^T \rightarrow \bar{u}$'s Use GSP if $\lambda_i = \lambda_j$ for $i \neq j$

   b) $A^TA \rightarrow \bar{v}$'s Use GSP if $\lambda_i = \lambda_j$ for $i \neq j$

3) Find $s_i = \sqrt{\lambda_i}$ found in step (1). Define S

   → use padding of zeroes to get right size.

4) Depending on (2) setup U or V:

   a) Deriving $\bar{v}$'s from $\bar{u}$'s : $\bar{v}_i^T = s_i^{-1} \cdot u_i^T \cdot A$

      ↳ you may have to use GSP again if $\lambda_i = 0$

   b) Deriving $\bar{u}$'s from $\bar{v}$'s : $u_i = s_i^{-1} \cdot A \cdot \bar{v}_1$

      ↳ you may have to use GSP again if $\lambda_i = 0$

5) Test $A == U \cdot S \cdot V^T$

Constructing orthogonal vectors:

Use Case:

i) Departed from "wrong" AA^T/A^TA and are missing one or more eigenvectors.

ii) One of the eig. vals is 0 meaning it's not possible to derive vectors from each other.

The "trick":

Assume you have two orthogonal vectors $\bar{u}_1$ and $\bar{u}_2$ and need a third one orthogonal to $\bar{u}_1$ and $\bar{u}_2$:

$$\begin{cases} \bar{u}_1 \cdot \bar{u}_3 = 0 \\ \bar{u}_2 \cdot \bar{u}_3 = 0 \end{cases} \quad , \quad \bar{u}_3 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\Downarrow$$

$$\left.\begin{aligned} \bar{u}_1^T \cdot \bar{u}_3 &= 0 \\ \bar{u}_2^T \cdot \bar{u}_3 &= 0 \end{aligned}\right\} \begin{aligned} u_{11}x_1 + u_{12}x_2 + u_{13}x_3 &= 0 \\ u_{21}x_1 + u_{22}x_2 + u_{23}x_3 &= 0 \end{aligned}$$

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \end{bmatrix} \xrightarrow{\text{Find}} \text{Nullspace}$$

so find Nullspace of

$$\begin{bmatrix} \bar{u}_1^T \\ \bar{u}_2^T \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \bar{u}_1 & \bar{u}_2 \end{bmatrix}^T$$

<u>EX</u>

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$$

1. Find eig.vals of $A \cdot A^T$:

$$A \cdot A^T = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{bmatrix} \to \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{vmatrix} = 0$$

$$\lambda_1 = 3 \; , \; \lambda_2 = 1 \; , \; \lambda_3 = 0$$

2. Find eigenspaces:

$\lambda = 3$: $\begin{bmatrix} 1-3 & 0 & 1 \\ 0 & 1-3 & 1 \\ 1 & 1 & 2-3 \end{bmatrix} \overset{RREF}{\sim} \begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix}$ $\sqrt{(1/2)^2 + (1/2)^2 + 1^2}$

$\sqrt{\frac{3}{2}}$

$\lambda = 1$: $\begin{bmatrix} 1-1 & 0 & 1 \\ 0 & 1-1 & 1 \\ 1 & 1 & 2-1 \end{bmatrix} \to \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

$\frac{1}{\frac{2}{\frac{\sqrt{3}}{\sqrt{2}}}} = \frac{\sqrt{2}}{2 \cdot \sqrt{3}}$

$\lambda = 0$: $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$

$\frac{1}{\sqrt{6}}$

$\frac{1}{\sqrt{6}}$

$\frac{\sqrt{2}}{\sqrt{3}}$

3. Find S:

$s_1 = \sqrt{3}$, $s_2 = 1$, $s_3 = 0$

$S = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

$\frac{3}{2}$

4. Derive $\bar{v}$'s:

$\bar{v}_1^T = (\sqrt{3})^{-1} \cdot \bar{u}_1^T \cdot A = (\sqrt{3})^{-1} \begin{bmatrix} 1/2 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$

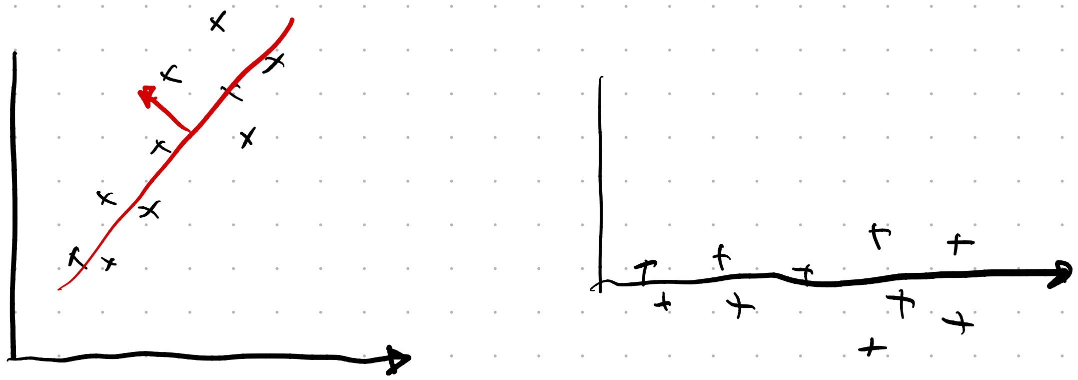
$= \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$

$\bar{v}_2^T = 1 \cdot \begin{bmatrix} -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \end{bmatrix}$

$$\begin{bmatrix} 1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ \sqrt{2}/\sqrt{3} & 0 & 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

Principal Component Analysis

The most used dimensionality reduction in statistics and ML.



It is the same as compressing the data!

1. Center data: subtract mean from all features (columns). Call this Data matrix X.

2. Find Covariance matrix C:

$$C = X^T X$$

3. Find the eigenvects and eigenvals of C
   ↳ eigendecomposition of A.

4. Sort the eig. vals from largest to smallest.

5. $C \cdot V = V \cdot D$

   $C = VDV^{-1}$

   eigenvect

   eigenvals

$y = \beta_1 x + \beta_2 x^2$

6. Principal component

$$T = X \cdot V$$

Principal Components

loadings

Using SVD for PCA:

Assume: $X = U \Sigma V^T$

centered

$$T = U \Sigma$$

Fraction of variance explained.

$$\frac{\sum_{i=1}^{r} = \lambda_i}{\sum_{i=1}^{n} = \lambda_i} \approx [0.7; 0.8]$$