9. Symmetric Matrices, SVD and PCA

Symmetric Matrices:

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

$$A = \begin{bmatrix} 1 & 4 & 6 \\ 4 & 2 & 5 \\ 6 & 5 & 3 \end{bmatrix}$$

Diagonalizing symmetric Matrices:

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

1. Find eig. vals:

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

For $\lambda = 6$:

$$\begin{bmatrix} -5 & 5 \\ & \end{bmatrix} \sim \begin{bmatrix} 1 & -1 \\ & \end{bmatrix} \rightarrow \overset{\bar{v}_1}{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}$$

For $\lambda = -4$:

$$\begin{bmatrix} 5 & 5 \\ & \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ & \end{bmatrix} \rightarrow \overset{\bar{v}_2}{\begin{bmatrix} -1 \\ 1 \end{bmatrix}}$$

$P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 6 & \\ & -4 \end{bmatrix}$ Notice $\bar{v}_1 \perp \bar{v}_2$

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) Use 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 that will work for *any* $m \times n$ matrix.

Solution: Singular Value Decomposition allows us to decompose any $m \times n$ Matrix into 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 corresponing 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 $U = AAt.eigenvects()$
$V = AtA.eigenvects()$
and use both.

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

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

$$A\bar{v}_1 = \sigma_1 \cdot \bar{u}_1$$
$$A\bar{v}_2 = \sigma_2 \cdot \bar{u}_2$$
$$A\bar{v}_3 = \sigma_3 \cdot \bar{u}_3$$
$$\vdots$$
$$A\bar{v}_n = \sigma_n \cdot \bar{u}_n$$

so we can find $\bar{u}_i$ as:

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

Note: We run into problems when

$n < m$

or

$\sigma_i = 0$

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]

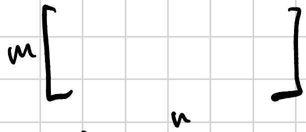
If we depart from AA^T we get too few eig. vecs. Depart from A^TA.

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) $AA^T \to \bar{u}$'s

b) $A^TA \to \bar{v}$'s

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

$\to$ 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 : $\bar{v}_i^T = s_i^{-1} \cdot \bar{u}_i^T \cdot A$

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

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 $\bar{u}_2$:

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

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

$$\downarrow$$

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

So find nullspace of:

$$\begin{bmatrix} u_1^T \\ u_2^T \end{bmatrix}$$

Ex

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

Step 1:

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

$$(1-\lambda) \begin{vmatrix} 1-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} + \begin{vmatrix} 0 & 1-\lambda \\ 1 & 1 \end{vmatrix} = (1-\lambda)\Big((1-\lambda)(2-\lambda) - 1\Big) - (1-\lambda)$$

$$= (1-\lambda)\left(1 - 3\lambda + \lambda^2\right) - (1-\lambda)$$

$$= (1-\lambda)\left(1-3\lambda+\lambda^2-1\right) = (1-\lambda)(\lambda^2-3\lambda) = (1-\lambda)(\lambda-3)\lambda$$

$$\lambda = \begin{cases} 3 \\ 1 \\ 0 \end{cases}$$

Step 2:

$\lambda = 3$:

$$\begin{bmatrix} -2 & 0 & 1 \\ 0 & -2 & 1 \\ 1 & 1 & -1 \end{bmatrix} \overset{r_1 \to r_1 + 2r_3}{\sim} \begin{bmatrix} 0 & 2 & -1 \\ 0 & -2 & 1 \\ 1 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & -1/2 \\ 0 & 0 & 0 \end{bmatrix}$$

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

$\lambda = 1$:

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

$\lambda = 0$:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

Step 3:

$$\sigma_1 = \sqrt{3} \quad , \sigma_2 = \sqrt{1} = 1 \quad , \sigma_3 = 0$$

step 4: $\bar{v}_i^T = s_i^{-1} \cdot u_i^T \cdot A$

$$\bar{v}_1^T = (\sqrt{3})^{-1} \cdot \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{18}} \begin{bmatrix} 3 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

$$\bar{v}_2^T = 1 \cdot \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$$V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

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

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

Pseudo-Inverse (Moore-Penrose):

Problem: Not all matrices have invers
↳ Rectangular: $m \neq n$
↳ Singular $n \times n$ matrices

solution: The pseudoinverse, A^+, gives a way to "invert" such matrices in a least-squares sense

Definition:

For a matrix $A \in \mathbb{R}^{m \times n}$, its Moore-Penrose pseudoinverse is the unique matrix that satisfies:

1. $A A^+ A = A$
2. $A^+ A A^+ = A^+$
3. $(AA^+)^T = AA^+$
4. $(A^+ A)^T = A^+ A$

How to compute:

Given $A = U \Sigma V^T$, $A^+ = V \Sigma^+ U^T$

what built in functions use.

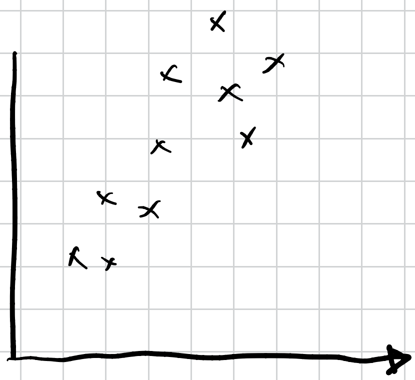
where Σ^+:

- $\sigma_i \mapsto 1/\sigma_i$, $\sigma_i \neq 0$
- $\Sigma = \Sigma^T$

(see notebook)

Principal Component Analysis If time permits.

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 S:

3. Find the eigenvects and eigenvals of S

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

5. $C \cdot V = VD$

6. Principal component

Using SVD for PCA:

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