Linear Algebra & Convex Optimization – Lecture 8

References : Introduction to Linear Algebra, Gilbert Strang; Online References

Analysis of Square Matrices

Eigen Vectors

$A v_i = \lambda_i v_i$

$3 v_i$

$A(3 v_i) = \lambda_i (3 v_i)$

$n \times n$ square matrix **A** is a **function** that maps a point in *n*-dimensional space to another point in the same space

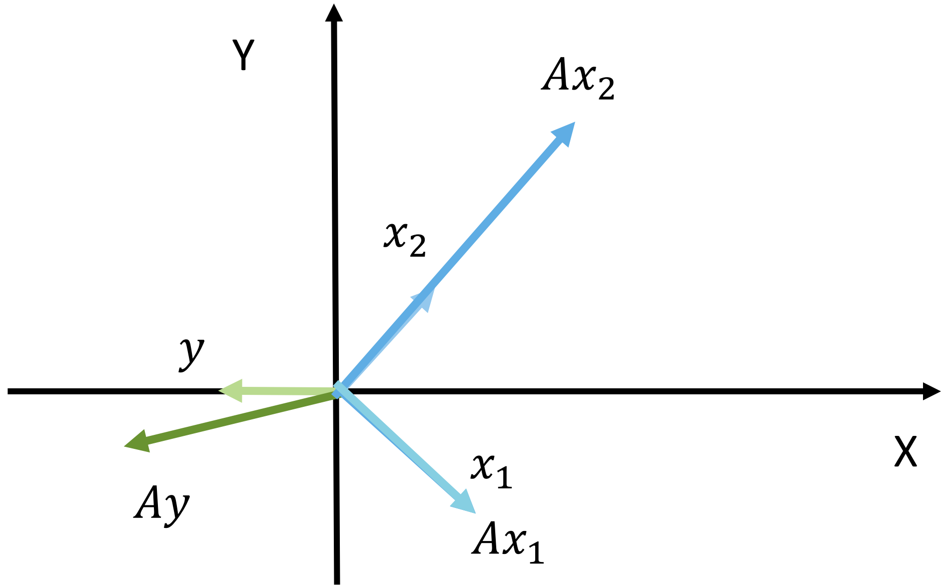
Consider a transformation induced by the Square Matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

E.g. If $y = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$, $Ay = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$

But,

- If $x_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ $Ax_1 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1 . x_1,$
- If $x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, Ax_2 = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3 . x_2,$



For a square Matrix *A*, certain vectors only get linearly scaled and do not change the direction !!
These directions are **eigen vectors**.
The scaling value the **eigen vectors** undergo are **eigen values.**

Definition :

A vector $v \in \mathbb{R}^n$ is an eigen vector of an $n \times n$ square matrix A if

$$Av = \lambda v$$

$Av = \lambda v$ can also be written as $(A - \lambda I)v = 0$

Columns of $(A - \lambda I)$ are dependent $\Longrightarrow$ $|A - \lambda I| = 0$

Poll:

Can we solve the eigen equation system similar to $Ax = b$?

Goal: Find all sets of (λ_i, v_i) where $(A - \lambda_i I)v_i = 0$, where λ_i is associated with the eigen vector v_i

Finding Eigen Values & Vectors :

Goal: Find all sets of (λ_i, v_i) where $(A - \lambda_i I)v_i = 0$, where λ_i is associated with the eigen vector v_i

How many such λ_is we can find?

Example:

Characteristic Polynomial

Roots of Characteristic Equation

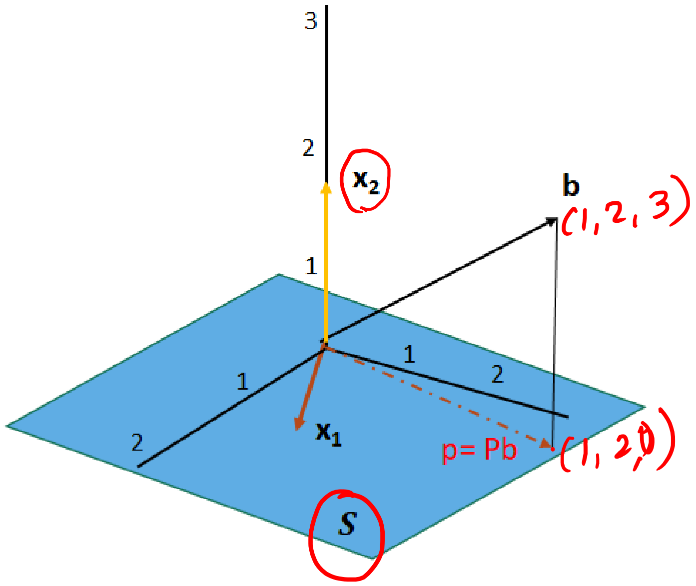
$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad |A - \lambda I| = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 3 - 4\lambda + \lambda^2 = 0. \Longrightarrow (\lambda - 1)(\lambda - 3) = 0$$

$$\lambda_1 = 1, \lambda_2 = 3$$

$$(A - 3I)v_{\lambda=3} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad -1v_1 + 1v_2 = 0; 1v_1 - 1v_2 = 0 \qquad v_{\lambda=3} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

- Determinant of $n \times n$ square matrix $A - \lambda I$ will have characteristic polynomial of degree n
- A characteristic polynomial on λ of degree n will have n roots (real/complex , repeated/distinct)

Eigen Values & Vectors of Projection Matrix



Projection Matrix P is defined with respect a sub-space S

Pb will modify all $b \notin S$

For all $b \in S$, $Pb = 1.b$. This implies all $b \in S$ are eigenvectors with $\lambda = 1$

For all $b \perp S$, $Pb = 0.b$. This implies all $b \perp S$ are eigenvectors with $\lambda = 0$

Example: Projection Matrix for *X-Y* Plane:

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

$$P \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

$$|P - \lambda I| = 0$$

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

$$\lambda(1-\lambda)^2 = 0$$

$$\lambda_1 = 0, \lambda_2 = 1$$

λ_2 has an algebraic multiplicity of 2 as it occurs 2 times in the root of characteristic polynomial.

λ_2 has a geometric multiplicity of 2 as we can find 2 independent eigen vectors corresponding to λ_2 .

Geometric Multiplicity < = Algebraic Multiplicity

Eigen Values & Vectors of Rotation Matrix

Rotation Matrix:

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Example:

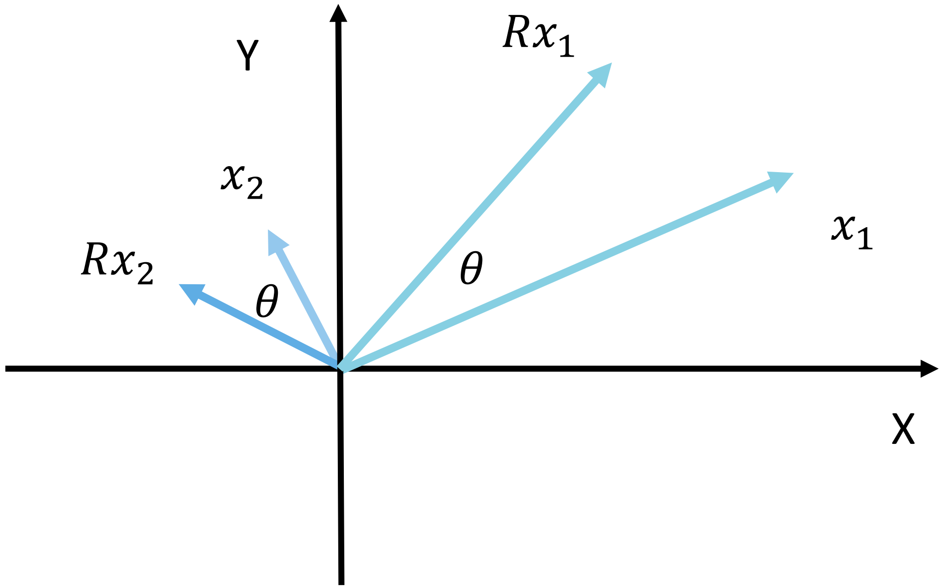
$$\theta = 90^0$$

$$R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad |R - \lambda I| = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix}$$

$$= \lambda^2 + 1 = 0$$

$$\lambda_1 = i; \lambda_2 = -i$$

$$v_1 = \begin{bmatrix} 1 \\ i \end{bmatrix} \qquad v_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$$



How do you interpret complex eigen values ? Check this: https://haoye.us/post/2019-12-05-interpreting-complex-eigenvalues/

Linear Independence of Eigen Vectors:

If all eigenvalues $\lambda_1, \ldots, \lambda_n$ are different then all eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly independent

Proof:

① consider $c_1\mathbf{x}_1 + \ldots + c_n\mathbf{x}_n = \mathbf{0}$ $\Rightarrow$ We should prove $c_1 = c_2 = \cdots c_n = 0$

$c_1x_1 + c_2x_2 + \ldots c_nx_n = 0$

$c_1\underbrace{Ax_1}_{\lambda_1x_1} + c_2\underbrace{Ax_2}_{\lambda_2x_2} + \ldots c_nAx_n = 0$

multiply by A to get $c_1\lambda_1\mathbf{x}_1 + \ldots + c_n\lambda_n\mathbf{x}_n = \mathbf{0}$ ① × A → ②

multiply by λ_n to get $c_1\lambda_n\mathbf{x}_1 + \ldots + c_n\lambda_n\mathbf{x}_n = \mathbf{0}$ ① × λ_n → ③

subtract them, get $\mathbf{x}_n$ removed and have the following: ② − ③

- $c_1(\lambda_1 - \lambda_n)\mathbf{x}_1 + \ldots + c_{n-1}(\lambda_{n-1} - \lambda_n)\mathbf{x}_{n-1} = \mathbf{0}$ → ④

do the same again: multiply it with A and with λ_{n-1} to get

- $c_1(\lambda_1 - \lambda_n)\lambda_1\mathbf{x}_1 + \ldots + c_{n-1}(\lambda_{n-1} - \lambda_n)\lambda_{n-1}\mathbf{x}_{n-1} = \mathbf{0}$ ⑤ ④ × A
- $c_1(\lambda_1 - \lambda_n)\lambda_{n-1}\mathbf{x}_1 + \ldots + c_{n-1}(\lambda_{n-1} - \lambda_n)\lambda_{n-1}\mathbf{x}_{n-1} = \mathbf{0}$ ④ × λ_{n-1} → ⑥
- subtract to get rid of $\mathbf{x}_{n-1}$ ⑤ − ⑥

eventually, have this:

$(\lambda_1 - \lambda_2) \cdot (\lambda_1 - \lambda_3) \cdot \ldots \cdot (\lambda_1 - \lambda_n) \cdot c_1 \cancel{\mathbf{x}_n}^{x_1} = \mathbf{0}$

since all λ_i are distinct and $\boldsymbol{x}_n \neq \mathbf{0}$, conclude that $c_1 = 0$

can show the same for the rest $c_2, \ldots, c_n$

$c_1 = c_2 = \cdots c_n = 0$

implies

$x_1, x_2, \ldots x_n$ are independent

Eigen Decomposition (Spectral Decomposition):

$$AB = A\begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} = \begin{bmatrix} Ab_1 & \cdots & Ab_n \end{bmatrix} = \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix}$$

For $n \times n$ square matrix A , assume there are n distinct eigen values

$$Av = \lambda v \quad \Longrightarrow \quad \begin{matrix} Av_1 = \lambda_1 v_1 \\ Av_2 = \lambda_2 v_2 \\ \vdots \\ Av_n = \lambda_n v_n \end{matrix} \quad \Longrightarrow \quad A\underbrace{\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ v_1 & v_2 & \cdots & v_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}}_{S} = \underbrace{\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_n v_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}} = \underbrace{\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ v_1 & v_2 & \cdots & v_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}}_{S} \underbrace{\begin{bmatrix} \lambda_1 & & & \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}}_{\Lambda}$$

$$S = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \qquad AS = A\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} = S\Lambda \qquad \text{where, } \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

$\rightarrow S^{-1} AS = S\Lambda$

$\boldsymbol{S^{-1}}$ exists for A

Matrix Factorization

$$\boldsymbol{A = S\Lambda S^{-1}}$$

Matrix Diagonalization

$$\underline{\boldsymbol{S^{-1}AS}} = \underline{\boldsymbol{\Lambda}}$$

Eigen Decomposition : Application

Efficiently Computing Matrix Powers:

if $A\mathbf{x} = \lambda\mathbf{x}$ then

$$A^2\mathbf{x} = \lambda A\mathbf{x} = \lambda^2\mathbf{x}$$

λs are squared when A is squared

Eigenvectors stay the same and don't mix up, only eigenvalues grow

for diagonalizable matrices:

$$A^2 = AA$$

$$= (S\Lambda S^{-1})(S\Lambda S^{-1})$$

$$= S\Lambda^2 S^{-1}$$

$$\boxed{A^k = S\Lambda^k S^{-1}}$$

$A = \begin{bmatrix} & & \\ & & \end{bmatrix}_{10000 \times 10000}$ $\quad A^{100}$

$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$

$\Lambda \cdot \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$

Similar Matrices

Two square matrices A and B are *similar* if $B = M^{-1}AM$ for some matrix M.

E.g. If A has distinct eigenvalues, $S^{-1}AS = \Lambda \implies$ A and Λ is similar, in this case $M = S$

$$\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$\Lambda x = \lambda_1 x$

$$\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{e_1} = \lambda_1 \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{e_1}$$

Why are they called "similar" ?

suppose $A\mathbf{x} = \lambda\mathbf{x}$ $= A I x = \lambda x$

$A M M^{-1} x = \lambda x$

$$AMM^{-1}\mathbf{x} = \lambda\mathbf{x} \quad \text{— (1)}$$

$$M^{-1}AMM^{-1}\mathbf{x} = \lambda M^{-1}\mathbf{x} \quad M^{-1} \times (1)$$

$$\underbrace{BM^{-1}\mathbf{x}}_{} = \lambda \underbrace{M^{-1}\mathbf{x}}_{y}$$

$B\,y = \lambda y$

Similar Matrices have same eigenvalues

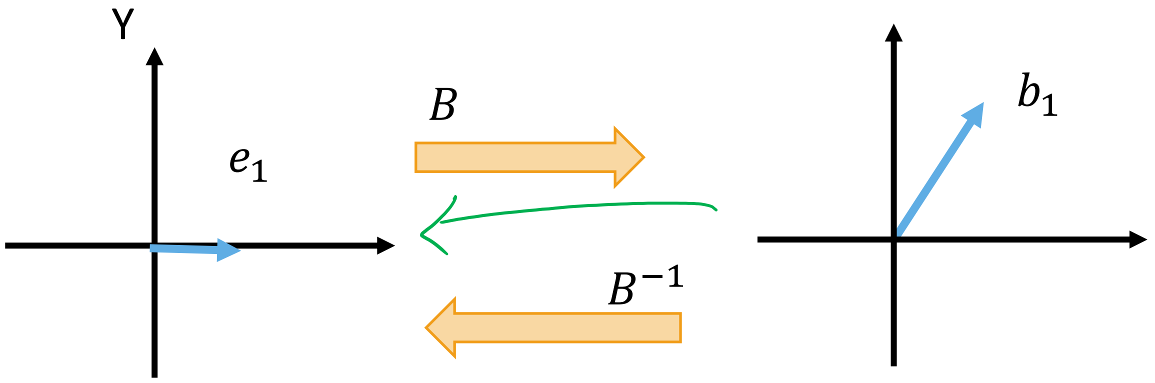
Why Geometric Multiplicity <= Algebraic Multiplicity ?

$Be_1 = b_1$ $\quad e_1 = B^{-1} b_1$

B transforms unit vector e_1 to b_1.
Also $e_1 = B^{-1}Be_1 = B^{-1}b$

$B: \mathbb{R}^n \rightarrow \mathbb{R}^n$

$$B = \begin{bmatrix} b_1 & b_2 \dots b_n \end{bmatrix} = \begin{bmatrix} Be_1 & Be_2 \dots Be_n \end{bmatrix}$$



$Av_i = \lambda_0 v_i$ and geometric multiplicity of λ_0 is k

$\{v_1, v_2, \dots v_k\}$ forms the set of independent eigen vectors

Create a matrix P such that $P = [v_1 \; v_2 \; . \; . \; . \; v_k \;\; u_{k+1} \; . \; . \; . \; u_n]$

$u_{k+1}, . \; . \; . \; , u_n$ are any set of independent vectors that span rest of $\mathbb{R}^n$

$M = P^{-1}AP$ is a similar matrix of A with same eigen values

$$M = P^{-1}AP \quad = P^{-1}A\,[\underbrace{v_1 \; v_2 \; . \; . \; . \; v_k \;\; u_{k+1} \; . \; . \; . \; u_n}_{P}]$$

$$= P^{-1}\,[Av_1 \; Av_2 \; . \; . \; . \; Av_k \; Au_{k+1} \; . \; . \; . \; Au_n]$$

$$= P^{-1}\,[\lambda_0 v_1 \; \lambda_0 v_2 \; . \; . \; . \; \lambda_0 v_k \; Au_{k+1} \; . \; . \; . \; Au_n]$$

$$= [\lambda_0 P^{-1} v_1 \;\; \lambda_0 P^{-1} v_2 \; . \; . \; . \; \lambda_0 P^{-1} v_k \; P^{-1}Au_{k+1} \; . \; . \; . \; P^{-1}Au_n]$$

$$= [\lambda_0 e_1 \;\; \lambda_0 e_2 . \; . \; . \;\; \lambda_0 e_k \;\; P^{-1}Au_{k+1} \; . \; . \; . \; P^{-1}Au_n] = (M)$$

$$M = \begin{bmatrix} \lambda_0 & \cdots & 0 & \star & \cdots & \star \\ \vdots & \ddots & \vdots & \vdots & Q_1 & \vdots \\ 0 & \cdots & \lambda_0 & \star & \cdots & \star \\ 0 & \cdots & 0 & \star & \cdots & \star \\ \vdots & & \vdots & \vdots & Q_2 & \vdots \\ 0 & \cdots & 0 & \star & \cdots & \star \end{bmatrix}$$

$\lambda_M = \text{Soln. of } \det(M - \lambda I)$

$\lambda_M = Soln. of \; (\lambda_0 - \lambda)^k . \det(Q_2 - \lambda I)$

*λ_0 *will be a root of* $(\lambda_0 - \lambda)^k$. λ_0 *may or maynot be root of* $det(Q_2 - \lambda I)$. *Hence Algebraic Multiplicity of* λ_0 *is at least* k

Eigen Analysis: Symmetric Matrices (A = A^T)

$n \times n$ real Symmetric Matrix A will have
1. n real eigen values (not necessarily distinct)
2. All eigen vectors are orthogonal to each other

Proof: (Real Eigen Values)

$$z = a + bi \qquad \bar{z} = a - bi \qquad z\bar{z} = (a + bi)(a - bi) = a^2 + b^2.$$

if w, z are complex numbers, then $\overline{wz} = \bar{w}\,\bar{z}$.

$$\underbrace{\begin{bmatrix} a_1 - b_1 i \\ a_2 - b_2 i \\ \vdots \\ a_n - b_n i \end{bmatrix}^T}_{\bar{\mathbf{v}}^T} \cdot \underbrace{\begin{bmatrix} a_1 + b_1 i \\ a_2 + b_2 i \\ \vdots \\ a_n + b_n i \end{bmatrix}}_{\mathbf{v}} = (a_1^2 + b_1^2) + (a_2^2 + b_2^2) + \cdots + (a_n^2 + b_n^2)$$

$$\mathbf{v} \neq \mathbf{0} \;\Rightarrow\; \bar{\mathbf{v}} \cdot \mathbf{v} \neq 0$$

$$A\mathbf{v} = \lambda\mathbf{v}.$$

$$\overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} \Rightarrow A\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}.$$

$$\bar{\mathbf{v}}^T A\mathbf{v} = \bar{\mathbf{v}}^T (A\mathbf{v}) = \bar{\mathbf{v}}^T (\lambda\mathbf{v}) = \lambda(\bar{\mathbf{v}}^T \cdot \mathbf{v})$$

$$\bar{\mathbf{v}}^T A\mathbf{v} = (A\bar{\mathbf{v}})^T \mathbf{v} = (\bar{\lambda}\bar{\mathbf{v}})^T \mathbf{v} = \bar{\lambda}(\bar{\mathbf{v}}^T \cdot \mathbf{v})$$

$$\Rightarrow \qquad \lambda = \bar{\lambda}$$

Eigen Analysis: Symmetric Matrices (A = A^T)

$n \times n$ real Symmetric Matrix A will have

1. n real eigen values (not necessarily distinct)
2. All eigen vectors are orthogonal to each other

Proof: (Orthogonal Eigen Vectors)

x is an eigenvector of A corresponding to the eigenvalue λ_1

y is an eigenvector of A corresponding to the eigenvalue λ_2

$$\lambda_1 \neq \lambda_2.$$

$$Ax = \lambda_1 x$$
$$Ay = \lambda_2 y$$

After taking into account the fact that A is symmetric

$$y^\intercal A x = \lambda_1 y^\intercal x$$
$$x^\intercal A^\intercal y = \lambda_2 x^\intercal y$$

$$y^\intercal A x - x^\intercal A^\intercal y = \lambda_1 y^\intercal x - \lambda_2 x^\intercal y$$
$$0 = (\lambda_1 - \lambda_2) y^\intercal x$$

Hence x and y are orthogonal.

Eigendecomposition of A is $A = Q\Lambda Q^T$
instead of $A = S\Lambda S^{-1}$

$$A = \overset{Q}{\begin{bmatrix} | & | & & | \\ q_1 & \cdots & & q_n \\ | & | & & | \end{bmatrix}} \begin{bmatrix} \ddots \end{bmatrix} \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix}$$

$$Ax = \underbrace{\overset{Q}{\begin{bmatrix} \quad \end{bmatrix}}} \underbrace{\overset{\Lambda}{\begin{bmatrix} \ddots \end{bmatrix}}} \underbrace{\begin{bmatrix} q_1^T \\ \vdots \\ q_n^T \end{bmatrix} \begin{bmatrix} x \end{bmatrix}}$$

Eigen Analysis: Symmetric Matrices

Every Symmetric Matrix can be factorized as $A = Q\Lambda Q^T$
with real eigenvalues Λ and orthonormal eigenvectors in the columns of Q

$$A = Q\Lambda Q^T = \begin{bmatrix} | & & | \\ \mathbf{q}_1 & \cdots & \mathbf{q}_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} - & \mathbf{q}_1^T & - \\ & \vdots & \\ - & \mathbf{q}_n^T & - \end{bmatrix}$$

$$\begin{bmatrix} | & & | \\ q_1 & \cdots & q_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 q_1^T \\ \lambda_2 q_2^T \\ \vdots \\ \lambda_n q_n^T \end{bmatrix}$$

$$A = Q\Lambda Q^T = \sum \lambda_i \mathbf{q}_i \mathbf{q}_i^T \Rightarrow$$

Sum of Rank-1 Matrices

Spectral Decomposition of Matrix A

$$= \lambda_1 q_1 q_1^T + \cdots \lambda_n q_n q_n^T$$

a projection matrix is $P_i = \frac{\mathbf{q}_i \mathbf{q}_i^T}{\|\mathbf{q}_i\|^2} = \mathbf{q}_i \mathbf{q}_i^T$

$$P_i = (q_i q_i^T)$$

$$Ax = (\lambda_1 q_1 q_1^T + \cdots \lambda_n q_n q_n^T)\,x$$

each of these outer products can be seen as a Projection Matrix

so symmetric matrix can be represented as a combination of mutually orthogonal projection matrices

$$= \lambda_1 q_1 \underbrace{q_1^T x}_{c_1} + \cdots \lambda_n q_n \underbrace{q_n^T x}_{c_n} = \lambda_1 c_1 q_1 + \cdots + \lambda_n c_n q_n$$

$$A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \qquad B = \begin{bmatrix} \tilde{b}_1^T \\ \vdots \\ \tilde{b}_n^T \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}\tilde{b}_1^T + a_{12}\tilde{b}_2^T + \cdots \end{bmatrix}$$

$$\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} \tilde{b}_1^T \\ \vdots \\ \tilde{b}_n^T \end{bmatrix} = a_1 \tilde{b}_1^T + \cdots a_n \tilde{b}_n^T$$

Positive Semi-Definite (PSD) Matrices

What is the intuition on Matrix transformation for PSD Matrices ?

A matrix is **positive semi-definite** if it is symmetric and all eigen values are non-negative

- A matrix M is positive semi definite if $x^T M x \geq 0$ for all $x \neq 0$

$$x^T \underbrace{Mx}_{y}$$

$$x^T y \geq 0$$