Eigenvalues

$$\underset{\uparrow \text{ eigenvector}}{Ax} = \underset{\nwarrow \text{ eigenvalue}}{\lambda} x, \quad x \neq 0$$

$$A \in \mathbb{C}^{n \times n} \text{ or } \mathbb{R}^{n \times n}$$

$$Ax - \lambda x = 0$$

$$\underbrace{(A - \lambda I)}_{} x = 0$$

$$x \neq 0 \ ?$$

$$\text{null}(A - \lambda I) \neq \{\vec{0}\}$$

If $A - \lambda I$ is invertible

$$x = (A - \lambda I)^{-1} 0 = 0$$

$A - \lambda I$ is not invertible

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

determinant

$$\begin{bmatrix} a_{11}-\lambda & & & \\ & a_{22}-\lambda & & \\ & & \ddots & \\ & & & a_{nn}-\lambda \end{bmatrix}$$

$$\det \begin{bmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{bmatrix} =$$

example

characteristic polynomial

$$|A - \lambda I| = \text{poly}_n(\lambda) = 0$$

At most n

If we allow $\lambda \in \mathbb{C}$

At least one λ exists

At least one eigenvector for each distinct eigenv.

x **Not** unique (can multiply by any constant)

Eigenvectors are <u>directions</u>
(not vectors)

<u>Matrix notation</u>

$$\underline{\overline{X}} = [x_1 \mid x_2 \mid \dots \mid x_m]$$

linearly independent

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \emptyset \\ & \lambda_2 & & \\ & & \ddots & \\ \emptyset & & & \lambda_m \end{bmatrix}$$

capital λ

$$A x_i = \lambda_i x_i , \quad i = 1, \dots, m$$

$$A \underline{\overline{X}} = \underline{\overline{X}} \Lambda$$

$1 \leq m \leq n$

Is $m = n$? $\times$ not in general

Every λ has an algebraic multiplicity α and a geometric multiplicity β (how many linearly independent eigenvectors)

$1 \leq \beta \leq \alpha$

If $m = n$ we call that matrix non-defective or diagonalizable matrix

$m = n$ X is $[n \times n]$

$\Rightarrow X$ is invertible

Eigenvectors span all of $\mathbb{R}^n$

X^{-1} exists

$$X^{-1} \mid \quad AX = X\Lambda \quad \mid X^{-1}$$

$$X^{-1}AX = \Lambda$$

$$\underbrace{\boxed{A = X\Lambda X^{-1}}}_{\text{eigenvalue decomposition}}$$

If A is defective,
Jordan form (decomposition)

$$\Lambda \to \begin{bmatrix} \lambda_1 \; 1 \; 1 \; 1 & 0 & \\ & 1 \; 1 \; 1 \; 1 & 0 \\ & & \lambda_n \; 1 \; 1 \; 1 \; 1 \end{bmatrix}$$

Not computable numerically

<u>Assume</u> A is non-defective

$$\underbrace{X^{-1} A X}_{\text{similarity transform}} = \Lambda$$

$$A : \mathbb{R}^n \to \mathbb{R}^n$$

linear map

$$y = \sum_{i=1}^{n} a_i x_i$$

$$Ay = A \sum_i a_i x_i =$$

$$= \sum_i a_i (\underbrace{A x_i}_{\lambda_i x_i}) =$$

$$= \sum_i (\underline{\underline{a_i \lambda_i}}) x_i$$

In basis formed by eigenvectors A is diagonal with λ_i's on the diagonal

If x_i's are orthogonal then matrix is called unitarily diagonalizable

$X \rightarrow U$ orthogonal matrix

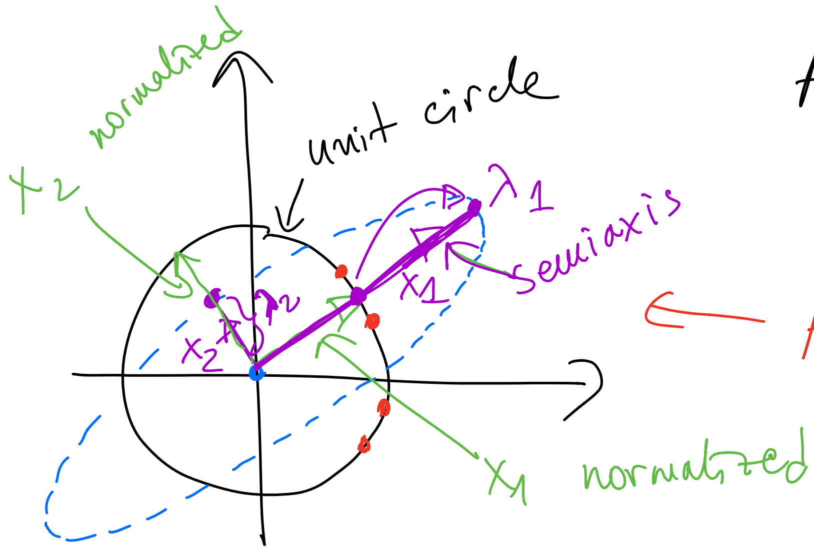
$\|x_i\| = 1$

$U^{-1} = U^*$ ← complex conjugate transpose

If and only if $AA^* = A^*A$

$$\boxed{A = U \Lambda U^*}$$

Assume A can be factorized like this



$A \in \mathbb{R}^{2\times2}$
& unitarily diagonalizable

$\leftarrow$ A is Hermitian/symmetric

$$A = U \Lambda U^*$$

$$A^* = (U \Lambda U^*)^* =$$

$$= (U^*)^* \Lambda^* U^*$$

$$= U \Lambda^* U^*$$

If $A^* = A$ Hermitian matrix or symmetric if in $\mathbb{R}^{n\times n}$

$$\cancel{U} \Lambda^* \cancel{U^*} = \cancel{U} \Lambda \cancel{U^*}$$

$$\Rightarrow \Lambda^* = \Lambda$$

eigenvalues are <u>real</u>

Theorem: If $A^* = A$ then A is unitarily diagonalizable and eigenvalues are real

{ From now on assume A is **H**ermitian }

If A is not unitarily diagonalizable, numerically best to use <u>Schur</u> decomposition

any matrix { $A = U T U^*$

T upper triangular

Eigenvalues on diagonal of T

Eigenvalues

A. Donev, Spring 2021

$$\underbrace{|A - \lambda I|}_{\text{poly}_n(\lambda) = 0} = 0$$

Abel's theorem says no closed-form solution for $n \geq 5$

All eigenvalue methods must be iterative/approximate

[Sidenote: In fact, solving polynomial eqs is done using eigenvalues (Matlab's roots)]

Two cases:

① We only want a few eigenvectors – with smallest or largest $|\lambda|$

Google's PageRank algorithm finds the eigenvector with the largest eigenvalue
(Power Method)

② All eigenvectors (next Wed, pre-recorded)
QR algorithm

Power – Method

$$A = X \Lambda X^{-1}$$

$$A^2 = X \Lambda \underbrace{X^{-1} X}_{= I} \Lambda X^{-1}$$

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

$$A^n = X \Lambda^n X^{-1}, \quad n \geq 1 \text{ integer}$$

Eigenvalalues of A^n are $(\lambda_i)^n$, and eigenvectors are the same.

As $n \to \infty$

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

Largest eigenvalue in modulus will dominate Λ^n

Aside: A^π ? π^π

$$\pi^{3/4} = \left(\pi^3\right)^{1/4} = x$$

$$x^4 = \pi^3$$

$$\ln\left(\pi^{\pi}\right) = \pi \ln(\pi)$$

$$\pi^{\pi} = e^{\pi \ln \pi}$$

$$A \cdot A = O(n^3) \text{ FLOPS}$$

$$\underset{[n \times n]}{A} \cdot \underset{[n \times 1]}{x} = O(n^2) \text{ FLOPS}$$

$$\underbrace{\left(A \ldots \ldots \left(A \cdot \left(A \cdot \left(A \cdot \left(A \cdot x\right)\right)\right)\right)\right)}_{n \text{ times}}$$

this can be computed without forming A

$$= A^n x$$

Choose a random vector q_0

compute $x_n = A^n q_0$

$$x_n = \bar{X} \Lambda^n \underbrace{\bar{X}^{-1} q_0}_{a}$$

$$\bar{X}^{-1} q_0 = a$$

$$\bar{X} \boxed{a} = \boxed{q_0}$$

$\vec{a}$ is $\vec{q_0}$ expressed in the eigenbasis of A

$$x_n = X(\Lambda^n a)$$

$$\lim_{n\to\infty} x_n = ?$$

$$\Lambda^n \xrightarrow[n\to\infty]{} \begin{bmatrix} \lambda_1^n & & \\ & \ddots & \\ & & 0 \end{bmatrix}$$

Assume $\lambda_1 > \lambda_2 \geq \lambda_3 \geq \lambda_4 \cdots$

strict inequality

$$\Lambda^n a \longrightarrow \begin{bmatrix} \lambda_1^n a_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$X \begin{bmatrix} \lambda_1^n a_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{x}_1 \cdot (\lambda_1^n a_1)$$

$$x_n \longrightarrow (\lambda_1^n a_1)\, x_1$$

$$\frac{x_n}{\|x_n\|} \longrightarrow x_1$$

aside

$$\begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix} \begin{bmatrix} b_1 & & \\ & \ddots & \\ & & b_n \end{bmatrix} = \begin{bmatrix} a_1 b_1 & & \\ & \ddots & \\ & & a_n b_n \end{bmatrix}$$

$$x_n \to X_1$$

How about λ_1

$$\underline{A X_n} \approx \underline{\lambda_1} \, \underline{X_n}$$

$$\lambda_1 = \frac{(A X_n)_{1,2,3,\dots,n}}{(X_n)_{1,2,3\dots,n}}$$

Rayleigh quotient

$$x_n \cdot A x_n = x_n^T A x_n \Rightarrow$$

$$x_1^T (\lambda_1 x_1)$$

$$= \lambda_1 \underbrace{(x_1^T x_1)}$$

$$= \lambda_1$$

$$\lambda_1 \approx x_n \cdot (A x_n)$$

If x_n is normalized

$$\lambda_1 \approx \frac{x_n \cdot (A x_n)}{x_n \cdot x_n}$$

$$\min_{x \neq 0} \frac{x^T A x}{x^T x} = \lambda_{min}$$

$$\max_{x \neq 0} \frac{x^T A x}{x^T x} = \lambda_{max}$$

$$\Rightarrow |\lambda_1| \le |\lambda_{max}|$$

Algorithm (Power method)

1) Choose random q_0

$$\tilde{q}_k = A q_{k-1}$$

(One matrix vector multiply)

$$q_k = \frac{\tilde{q}_k}{\|\tilde{q}_k\|}$$

$$\lambda_k = q_k^T A q_k$$

$$r_k = \|A q_k - \lambda_k q_k\|_2$$

residual

Stop if $r_k < \overset{10^{-9}}{\varepsilon} \cdot \lambda_k$

This guarantees 9 digits in λ
if $\varepsilon = 10^{-9}$

$$\vec{q}_0 = \underline{\overline{X}}\, \vec{a}$$

$$q_0 = \sum_{i=1}^{n} a_i x_i$$

$a_i \neq 0 \quad \forall i$

(why we choose q_0 random)

$$q_n = A^n q_0 = A^n \sum a_i x_i$$

$$= \sum a_i (A^n x_i) =$$

$$= a_i \quad (\lambda_i^n x_i)$$

$$= \sum (a_i \lambda_i^n) x_i$$

$$\tilde{q}_n = \sum a_i \lambda_1^n \cdot \left(\frac{\lambda_i}{\lambda_1}\right)^n x_i$$

$$\delta = \left|\frac{\lambda_2}{\lambda_1}\right| \qquad \delta < 1$$

$$\tilde{q}_n \longrightarrow a_1 \lambda_1^n x_1 + O(\delta^n)$$

normalize it

$$\| q_n - (\pm x_1) \| = O(\delta^n)$$

<u>linearly</u> convergent

$$q_n \longrightarrow x_1 + O(\delta) x_2 + O(\delta) x_3 + \dots$$

$$\lambda_n = q_n \cdot (A q_n)$$

$$= \underbrace{x_1 \cdot (A x_1)}_{\lambda_1} + O(\delta^2) \sum_{i,j \neq 1} x_i \cdot (A x_j)$$

$$\lambda_n = \lambda_1 + O(\delta^2)$$

$$\|\lambda_n - \lambda_1\| = O(\delta^{2n})$$

Eigenvalue is more accurate

If $\delta << 1$, iteration (power method) converges very fast, especially for eigenvalue

Eigenvalue algorithms converge well (rapidly) iff eigenvalues are (well) separated

In Matlab

(QR method)

$$[X, \Lambda] = \text{eig}(A)$$

<u>all</u> eigs & vectors

$O(n^3)$ but expensive

(much more than LU)

$$[X, \Lambda] = \underbrace{\text{eigs}}_{=}(A, n_{eig})$$

sparse

a <u>few</u> of eigs / vectors

(Power method)