Eigenvalues

$$Ax = \lambda x \quad , \quad x \neq 0$$

x: eigenvector

λ: eigenvalue

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

$$Ax - \lambda x = 0$$

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

$x \neq 0$?

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

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.

× <u>Not</u> unique (can multiply by any constant)

Eigenvectors are directions
(not vectors)

Matrix notation

$$\underline{\overline{X}} = \left[x_1 \,|\, x_2 \,|\, \dots \,|\, x_m \right]$$

linearly independent

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

capital λ

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

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

$1 \le m \le n$

Is $m = n$? ✗ not in general

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

$1 \le \beta \le \alpha$

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

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

$\Rightarrow X$ is invertible

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

X^{-1} exists

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

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

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

If A is defective,

Jordan form (decomposition)

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

Not computable numerically

Assume A is non-defective

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

similarity transform

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

linear map

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

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

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

$$= \sum_i (a_i \lambda_i) x_i$$

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

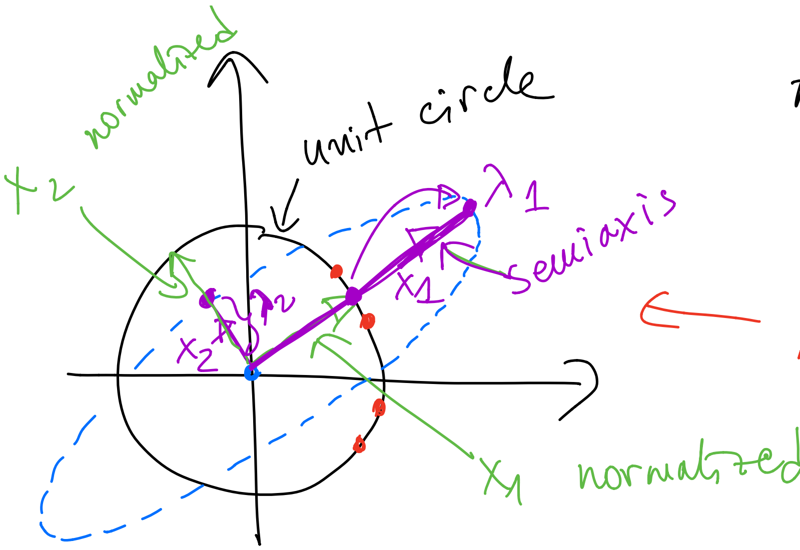
$\|x_i\| = 1$

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

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

$$\boxed{A = U \Lambda U^*}$$ ←

Assume A can be factorized like this



$A \in \mathbb{R}^{2 \times 2}$
& 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>

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

From now on assume A is Hermitian

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

any matrix $\quad A = U T U^*$

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 $n \geq 5$

$\boxed{\text{All eigenvalue methods must be iterative/approximate}}$

[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}

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

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

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

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

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

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

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

this can be computed <u>without</u> forming A

$$= A^n x$$

Choose a random vector q_0

compute $x_n = A^n q_0$

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

$$\overline{\underline{X}}^{-1} q_0 = a$$

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

$\vec{a}$ 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

$$\begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix} \begin{bmatrix} b_1 & & \\ & \ddots & \\ & & b_n \end{bmatrix} = \begin{bmatrix} ab_1 & & \\ & \ddots & \\ & & a_n b_n \end{bmatrix}$$

$$x_n \rightarrow X_1$$

How about λ_1

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

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

<u>Rayleigh quotient</u>

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

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

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

$$= \lambda_1$$

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

If x_n is normalized

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

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

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

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

<u>Algorithm</u> (Power method)

1) Choose random q_0

$$\left\{\begin{array}{l} \tilde{q}_k = A\, q_{k-1} \quad \text{(One matrix vector multiply)} \\[2ex] q_k = \dfrac{\tilde{q}_k}{\|\tilde{q}_k\|} \\[2ex] \lambda_k = q_k^T A\, q_k \\[2ex] r_k = \|A\, q_k - \lambda_k\, q_k\|_2 \\ \quad\quad \text{residual} \end{array}\right.$$

$\rightarrow$ 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)

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

all eigs & vectors

$O(n^3)$ but expensive

(much more than LU)

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

sparse

a few of eigs / vectors

(Power method)

QR algorithm

for eigenvalues

$A = QR$

unitary

upper triangular

Gram-Schmidt orthogonalization

$O(n^3)$ FLOPS

Similarity transformations

$A_k \longrightarrow A_{k+1}$

iteration

$A_1 = A$

Assume A is symmetric

$$A_{k+1} = P_k^{-1} A_k P_k$$

similarity transformation

or $P_k A_k P_k^{-1}$

invertible

$\left\{ A_{k+1} \text{ has the same eigenvalues as } A_k \right.$

$$A_{k+1} = P_k^{-1} A_k P_k$$

$$A_k = \underset{\text{unitary}}{X_k} \underset{\substack{\text{diagonal} \\ \text{(eigenvalues)}}}{\Lambda_k} X_k^*$$

$$A_{k+1} = \left(P_k^{-1} X_k\right) \Lambda_k \left(X_k^{-1} P_k\right)$$

$$X_{k+1} = P_k^{-1} X_k$$

$$X_{k+1}^{-1} = X_k^{-1} P_k$$

$$\Rightarrow \quad A_{k+1} = \underbrace{X_{k+1} \Lambda_k X_{k+1}^{-1}}_{\substack{\text{eigenvalue} \\ \text{decomposition}}}$$

$$A_{k+1} = X_{k+1} \Lambda X_{k+1}^{-1}$$

has the same eigenvalues as A

We want

(?) $A_k \longrightarrow$ diagonal matrix

<u>QR method</u> $A_1 = A$

Algorithm:

$$\begin{cases} A_k = Q_k R_k \\ A_{k+1} = R_k Q_k \end{cases}$$

swap order

$$A_{k+1} = Q_k^{-1} A_k Q_k \quad \text{(i.e. } P_k = Q_k^{-1}\text{)}$$

$$= \underbrace{(Q_k^* Q_k)}_{\text{identity}} R_k Q_k = R_k Q_k$$

$$A_k \longrightarrow \Lambda$$ diagonal matrix (of eigenvalues)

$$Q_1 Q_2 \cdots Q_k = \prod_{i=1}^{k} Q_i \longrightarrow X$$

s.t. $$A = X^{-1} \Lambda X$$

Why it works? $A = X^{-1} \Lambda X$

computes the same sequence as QR

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

$$= Q_k R_k$$

eigenvectors as $k \to \infty$

diagonal matrix

$$A_k = Q_k^* A Q_k$$ ← simil. transf.

similar matrices: same eigenvalues

$$A_k \longrightarrow \text{diagonal matrix}$$

$A = QR$

orthonormal basis for range / column space of A