Eigenvalue problem $\boxed{A}x = \lambda x$ ← n pairs (λ, x)

First of all, $Ax = \lambda x$ no explicit solution (neither λ nor x); huge difference from $Ax = b$ for which $x = \boxed{A^{-1}b}$

$A = LU \quad x = U^{-1}L^{-1}b$

- Eigenvalues are roots of characteristic polynomial

  $\lambda \in \mathrm{eig}(A) \Leftrightarrow \boxed{\det(A - \lambda I) = 0}$

- For any polynomial p, $\exists$ (infinitely many) matrices whose eigvals are roots of p

Eigenvalue problem $Ax = \lambda x$

First of all, $Ax = \lambda x$ no explicit solution (neither λ nor x); huge difference from $Ax = b$ for which $x = A^{-1}b$

- Eigenvalues are roots of characteristic polynomial
- For any polynomial p, $\exists$ (infinitely many) matrices whose eigvals are roots of p
- Let $p(x) = \boxed{x^n} + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, $a_i \in \mathbb{C}$. Then

$\frac{1}{a_n}$ $p(\lambda) = 0 \Leftrightarrow \lambda$ eigenvalue of

Companion

$$C = \begin{bmatrix} -a_{n-1} & -a_{n-2} & \dots & -a_1 & -a_0 \\ 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & 0 \end{bmatrix} \in \mathbb{C}^{n\times n}$$

eig(C) = roots(p)

Let $\lambda \in$ roots(p)

$$C \underbrace{\begin{bmatrix} \lambda^{n-1} \\ \vdots \\ \lambda^2 \\ \lambda \\ 1 \end{bmatrix}}_{v} = \begin{bmatrix} -a_{n-1}\lambda^{n-1} - a_{n-2}\lambda^{n-2} - \dots - a_0 \\ \lambda^{n-1} \\ \vdots \\ \lambda^2 \\ \lambda \end{bmatrix} = \lambda \begin{bmatrix} \lambda^{n-1} \\ \vdots \\ \lambda \\ 1 \end{bmatrix} = \lambda v$$

$-a_{n-1}\lambda^{n-1} - a_{n-2}\lambda^{n-2} - \dots - a_0 = \lambda^n$ iff $p(\lambda) = 0$.

$Cv = \lambda v$

Eigenvalue problem $Ax = \lambda x$

First of all, $Ax = \lambda x$ no explicit solution (neither λ nor x); huge difference from $Ax = b$ for which $x = A^{-1}b$ $\frac{2}{3}n^3$

- Eigenvalues are roots of characteristic polynomial
- For any polynomial p, $\exists$ (infinitely many) matrices whose eigvals are roots of p
- So no finite-step algorithm exists for $Ax = \lambda x$

Eigenvalue algorithms are necessarily $\boxed{\text{iterative and approximate}}$

- Same for SVD, as $\sigma_i(A) = \sqrt{\lambda_i(A^T A)}$
- But this doesn't mean they're inaccurate!

S.V. always accurate, λ_i(sym. matrix)

(i) back stability $A + \Delta A$

(ii) well-conditioned σ_i, λ_i(sym) Weyl's thm.

Usual goal: compute the Schur decomposition $A = UTU^*$: U unitary, T upper triangular

eig(A) = eig(T) = diag(T) = T_{ii}

- For normal matrices $A^*A = AA^*$, automatically diagonalised (T diagonal)
- For nonnormal A, if diagonalisation $A = X\Lambda X^{-1}$ really necessary, done via Sylvester equations but nonorthogonal/unstable (nonexaminable)

$AX - XB = C$.

Schur decomposition

Let $A \in \mathbb{C}^{n\times n}$ (square arbitrary matrix). Then $\exists$ unitary $U \in \mathbb{C}^{n\times n}$ s.t.

$$A = UTU^*,$$

with T upper triangular.

- $\text{eig}(A) = \text{eig}(T) = \text{diag}(T)$
- T diagonal iff A normal $A^*A = AA^*$

$\det(A-\lambda I)$, $\lambda \in \mathbb{C}$ $\Leftrightarrow$ $(A-\lambda I)v = 0$, $v \neq 0$.

Proof:

let $Av = \lambda v$

$[v\ V_\perp]^* A [v\ V_\perp] = \begin{bmatrix} \lambda & * \\ 0 & A_2 \end{bmatrix}$, A_2 $(n-1)\times(n-1)$

$\begin{bmatrix} v^* \\ V_\perp^* \end{bmatrix} \lambda v = \begin{bmatrix} \lambda \\ 0 \end{bmatrix}$

square, unitary $[\ ]^*[\ ] = I$

$\begin{bmatrix} 1 & \\ & \ \end{bmatrix}^* \begin{bmatrix} \lambda & \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & \\ & \hat{v}\ \hat{V}_\perp \end{bmatrix}$

$A_2 \hat{v} = \hat{\lambda}\hat{v}$, $\hat{\lambda} \in \text{eig}(A)$.

$= \begin{bmatrix} \lambda & * \\ 0 & \hat{\lambda} & * \\ 0 & 0 & * \end{bmatrix} \rightarrow$ repeat.

Schur decomposition

Let $A \in \mathbb{C}^{n\times n}$ (square arbitrary matrix). Then $\exists$ unitary $U \in \mathbb{C}^{n\times n}$ s.t.

$$A = UTU^*,$$

with T upper triangular.

- $\mathrm{eig}(A) = \mathrm{eig}(T) = \mathrm{diag}(T)$
- T diagonal iff A normal $A^*A = AA^*$

Proof: $\boxed{\text{Let } Av = \lambda_1 v}$ and find $U_1 = [v_1, V_\perp]$ unitary. Then

$$AU_1 = U_1 \begin{bmatrix} * & * & * & * & * \\ & * & * & * & * \\ & * & * & * & * \\ & * & * & * & * \\ & * & * & * & * \end{bmatrix} \Leftrightarrow U_1^* A U_1 = \begin{bmatrix} * & * & * & * & * \\ & * & * & * & * \\ & * & * & * & * \\ & * & * & * & * \\ & * & * & * & * \end{bmatrix}.$$

Repeat on the lower-right $(n-1)\times(n-1)$ part to get $U_{n-1}^* U_{n-2}^* \ldots U_1^* A U_1 U_2 \ldots U_{n-1} = T$.

Recap: Matrix decompositions

- SVD $A = U\Sigma V^T$
- Eigenvalue decomposition $A = X\Lambda X^{-1}$
  - Normal: X unitary $X^*X = I$
  - Symmetric: X unitary and Λ real
- Jordan decomposition: $A = XJX^{-1}$, $J = \mathrm{diag}(\begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix})$
- **Schur decomposition** $A = QTQ^*$: Q orthogonal, T upper triangular
- QR: Q orthonormal, U upper triangular
- LU: L lower triangular, U upper triangular

Red: Orthogonal decompositions, stable computation available

Recap: Matrix decompositions

- SVD $A = U\Sigma V^T$
- Eigenvalue decomposition $A = X\Lambda X^{-1}$
  - Normal: X unitary $X^*X = I$
  - Symmetric: X unitary and Λ real
- Jordan decomposition: $A = XJX^{-1}$, $J = \text{diag}\left(\begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}\right)$
- **Schur decomposition** $A = QTQ^*$: Q orthogonal, T upper triangular
- QR: Q orthonormal, U upper triangular
- LU: L lower triangular, U upper triangular
- QZ for $Ax = \lambda Bx$: (genearlised eigenvalue problem) Q, Z orthogonal s.t. QAZ, QBZ are both upper triangular

Red: Orthogonal decompositions, stable computation available

Power method for $Ax = \lambda x$

$x \in \mathbb{R}^n$:=random vector, $x := Ax$, $x = \frac{x}{\|x\|}$, $\lambda = \frac{x^T Ax}{\|x\|^2}$, repeat

$\frac{A^k x}{\|A^k x\|} \to$ eigvec for largest eigval $|\lambda_i|$

Suppose A diagonalisable $A = V \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} V^{-1}$, $V = [v_1 \cdots v_n]$

we can write ($\because V$ spans $\mathbb{R}^n$)

$x = \sum_{i=1}^n c_i v_i$

$(|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|)$

$$\frac{A^k x}{\lambda_1^k} = \sum_{i=1}^n c_i A^k v_i = \sum_{i=1}^n \frac{c_i \lambda_i^k v_i}{\lambda_1^k} = \left(\sum_{i=2}^n c_i \boxed{\left(\frac{\lambda_i}{\lambda_1}\right)^k} v_i\right) + c_1 v_1 \to c_1 v_1$$

$\to 0$ $\quad \boxed{\left(\frac{\lambda_2}{\lambda_1}\right)^k} \to 0$ as $k \to \infty$

conv. rate linear

w. convergence factor $\left|\frac{\lambda_2}{\lambda_1}\right|$

Power method for $Ax = \lambda x$

$x \in \mathbb{R}^n$:=random vector, $x = Ax$, $\boxed{x = \frac{x}{\|x\|},\ \hat{\lambda} = x^TAx,}$ repeat $\to \lambda_1$

- Convergence analysis: suppose A is diagonalisable (generic assumption). We can write $x_0 = \sum_{i=1}^{n} c_i v_i$, $Av_i = \lambda_i v_i$ with $|\lambda_1| > |\lambda_2| \geq \cdots$. Then after k iterations,

$$x = C\sum_{i=1}^{n}\left(\frac{\lambda_i}{\lambda_1}\right)^k c_i v_i \to Cc_1v_1 \quad \text{as } k\to\infty$$

- Converges geometrically $(\lambda, x) \to (\lambda_1, x_1)$ with **linear rate** $\frac{|\lambda_2|}{|\lambda_1|}$
- What does this imply about $A^{\cancel{n}\,k} = QR$ as $\cancel{n}\,k \to \infty$? First vector of $Q \to v_1$

$$\underbrace{A^k \begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}}_{\text{power method with initial vec } = \begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}} = QR\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix} = \begin{bmatrix}q_1\end{bmatrix}\cdot r_{11} \to v_1\cdot D$$

v_1

Power method for $Ax = \lambda x$

$x \in \mathbb{R}^n$:=random vector, $x = Ax$, $x = \frac{x}{\|x\|}$, $\lambda = x^T Ax$, repeat

- Convergence analysis: suppose A is diagonalisable (generic assumption). We can write $x_0 = \sum_{i=1}^n c_i v_i$, $Av_i = \lambda_i v_i$ with $|\lambda_1| > |\lambda_2| > \cdots$. Then after k iterations,

$$x = C \sum_{i=1}^{n} \left(\frac{\lambda_i}{\lambda_1} \right)^k c_i v_i \to C c_1 v_1 \quad \text{as } k \to \infty$$

- Converges geometrically $(\lambda, x) \to (\lambda_1, x_1)$ with **linear rate** $\frac{|\lambda_2|}{|\lambda_1|}$
- What does this imply about $A^n = QR$ as $n \to \infty$? First vector of $Q \to v_1$

Notes:

- Google pagerank & Markov chain linked to power method
- As we'll see, power method is basis for refined algs (QR algorithm, Krylov methods (Lanczos, Arnoldi,...))

Why compute eigenvalues? Google PageRank

'Importance' of websites via dominant eigenvector of column-stochastic matrix

$$A = \alpha P + (1-\alpha)\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$$

P: adjacency matrix, $\alpha \in (0,1)$

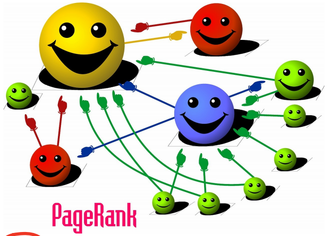


image from wikipedia

Google does (did) a few steps of Power method: with initial guess x_0, $k = 0, 1, \ldots$

1. $x_{k+1} = Ax_k$
2. $x_{k+1} = x_{k+1}/\|x_{k+1}\|_2$, $\quad k \leftarrow k+1$, repeat.

- $x_k \to$ PageRank vector $v_1 : Av_1 = \lambda_1 v_1$

Inverse power method (Shift-invert power)

$x \leftarrow Ax.$

$x \leftarrow (A-\mu I)^{-1}x$

Inverse power method: $x := (A-\mu I)^{-1}x$, $x = x/\|x\|$

- Converges with improved **linear rate** $\boxed{\frac{|\lambda_{\sigma(2)}-\mu|}{|\lambda_{\sigma(1)}-\mu|}}$ to eigval closest to μ (σ: permutation)

$(A-\mu I)y = x$

$\left|\frac{\lambda_2}{\lambda_1}\right|$

$\sigma: \{1,2,\dots,n\} \to \{1,\dots,n\}.$

$\lambda \in \mathrm{eig}(A)$

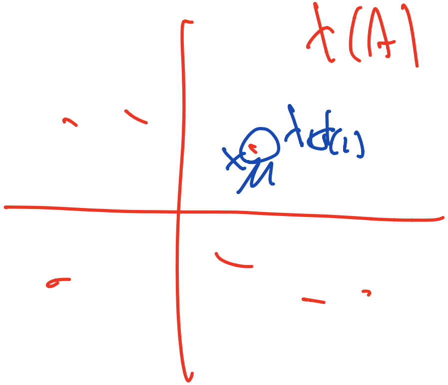
$\iff \frac{1}{\lambda-\mu} \in \mathrm{eig}(A-\mu I)$

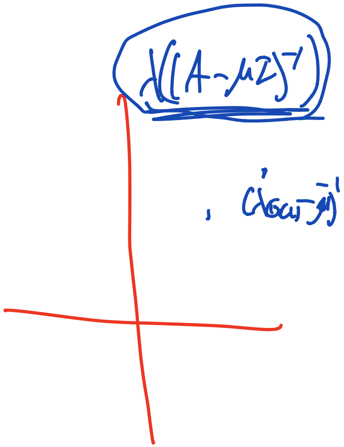
$Av = \lambda v$

$(A-\mu I)v = (\lambda-\mu)v$

$v = (\lambda-\mu)\cdot(A-\mu I)^{-1}v$

$(A-\mu I)^{-1}v = \boxed{\frac{1}{\lambda-\mu}}\,v$





Inverse power method

Inverse power method: $x := (A - \mu I)x$, $x = x/\|x\|$

- Converges with improved **linear rate** $\frac{|\lambda_{\sigma(2)} - \mu|}{|\lambda_{\sigma(1)} - \mu|}$ to eigval closest to μ (σ: permutation)
- μ can change adaptively with the iterations. The choice $\mu := x^T A x$ gives Rayleigh quotient iteration, with **quadratic** convergence $\|Ax^{(k+1)} - \lambda^{(k+1)} x^{(k+1)}\| = O(\|Ax^{(k)} - \lambda^{(k)} x^{(k)}\|^2)$ (cubic if A symmetric)