

Square roots of matrices

Definition:

matrix A , $B = \sqrt{A}$ if $B \cdot B = \sqrt{A} \sqrt{A} = A$

A matrix can have several (many) square roots

Example¹⁾ $A = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (identity matrix)

has infinitely many square roots:

$$\begin{pmatrix} 1 & 0 \\ 0 & \pm 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & \pm 1 \end{pmatrix}, \frac{1}{c} \begin{pmatrix} b & a \\ a & -b \end{pmatrix}, \frac{1}{c} \begin{pmatrix} b & -a \\ -a & -b \end{pmatrix},$$

$$\frac{1}{c} \begin{pmatrix} -b & a \\ a & b \end{pmatrix}, \frac{1}{c} \begin{pmatrix} -b & -a \\ -a & b \end{pmatrix}$$

$$\text{with } a^2 + b^2 = c^2$$

2) a positive semi-definite matrix (eigenvalues $\lambda \geq 0$)
has only one positive semi-definite square root:
"principal square root"

3) a 2×2 matrix with eigenvalues $\lambda_1 \neq 0, \lambda_2 \neq 0$
and $\lambda_1 \neq \lambda_2$
has four square roots

4) a general $n \times n$ matrix with $\lambda_1 \neq \lambda_2 \neq \dots \neq \lambda_n \neq 0$
has 2^n square roots

5) $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has no square roots ---- ETCETERA!

Matrix - Newton

remember: $x^2 - a = 0 \rightarrow \begin{cases} x_0 = a \\ x_{k+1} = \frac{1}{2} \left(x_k + \frac{a}{x_k} \right), k=0,1,2,\dots \\ = \frac{1}{2} \left(x_k + x_k^{-1} a \right) \end{cases}$

check $\rightarrow x_k - \frac{x_k^2 - a}{2x_k}$

Newton-Raphson to find $\pm\sqrt{a}$

matrix variant of this iteration method:

$X^2 - A = O \rightarrow \begin{cases} X_0 = A \\ X_{k+1} = \frac{1}{2} \left(X_k + X_k^{-1} A \right) \\ k=0,1,2,\dots \end{cases}$

$\begin{pmatrix} 0 & \dots & 0 \\ \vdots & \Theta & \vdots \\ 0 & \dots & 0 \end{pmatrix}$

"matrix-Newton" for square roots of matrices.

Theorem: if $\frac{1}{2} \left| 1 - \left(\frac{\lambda_j}{\lambda_i} \right)^{1/2} \right| \leq 1 \quad \forall (i,j)$ ($n \times n$ matrix)

(Higham, 1986)

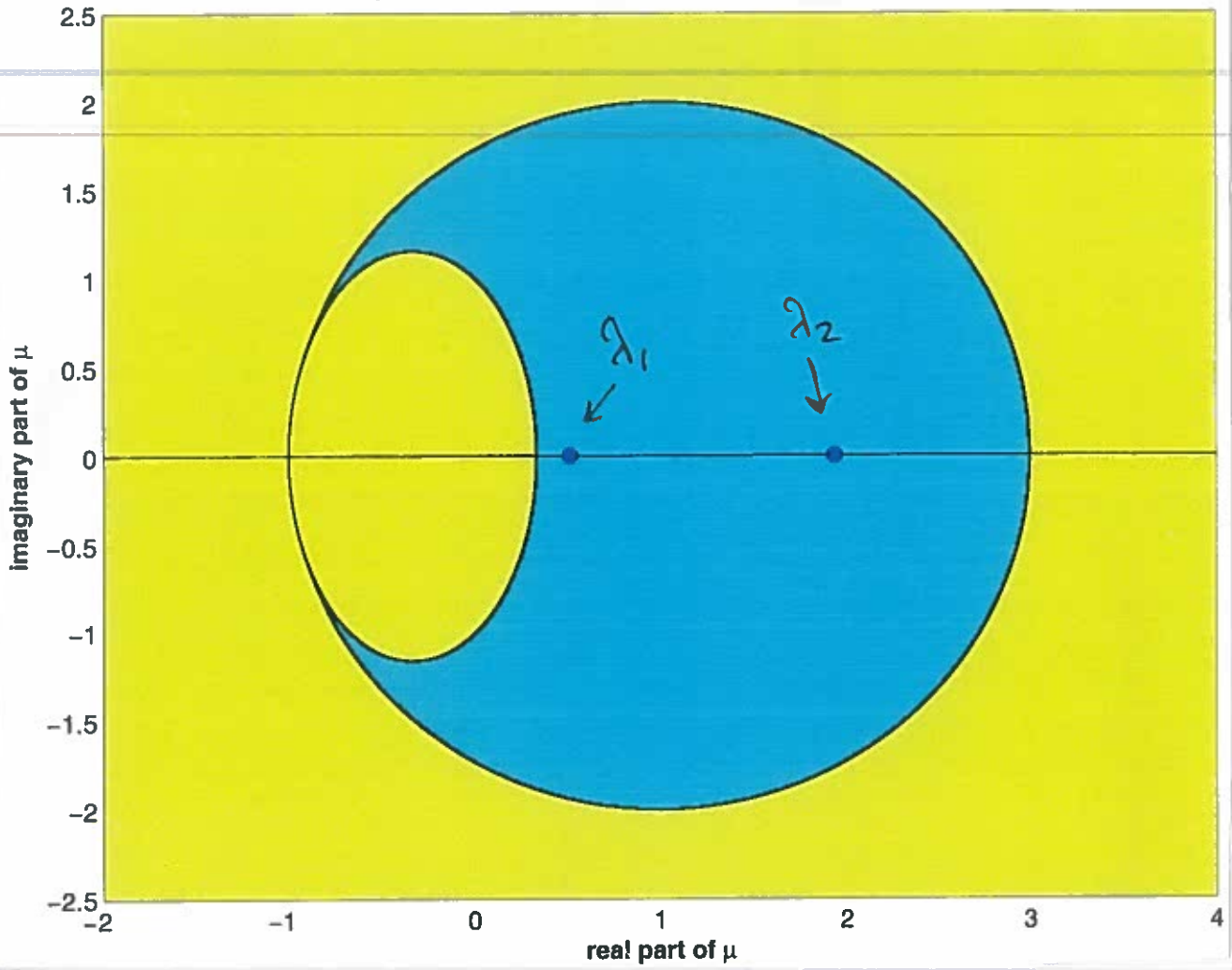
then matrix-Newton is stable
(small roundoff errors won't be amplified)

λ 's are eigenvalues of A

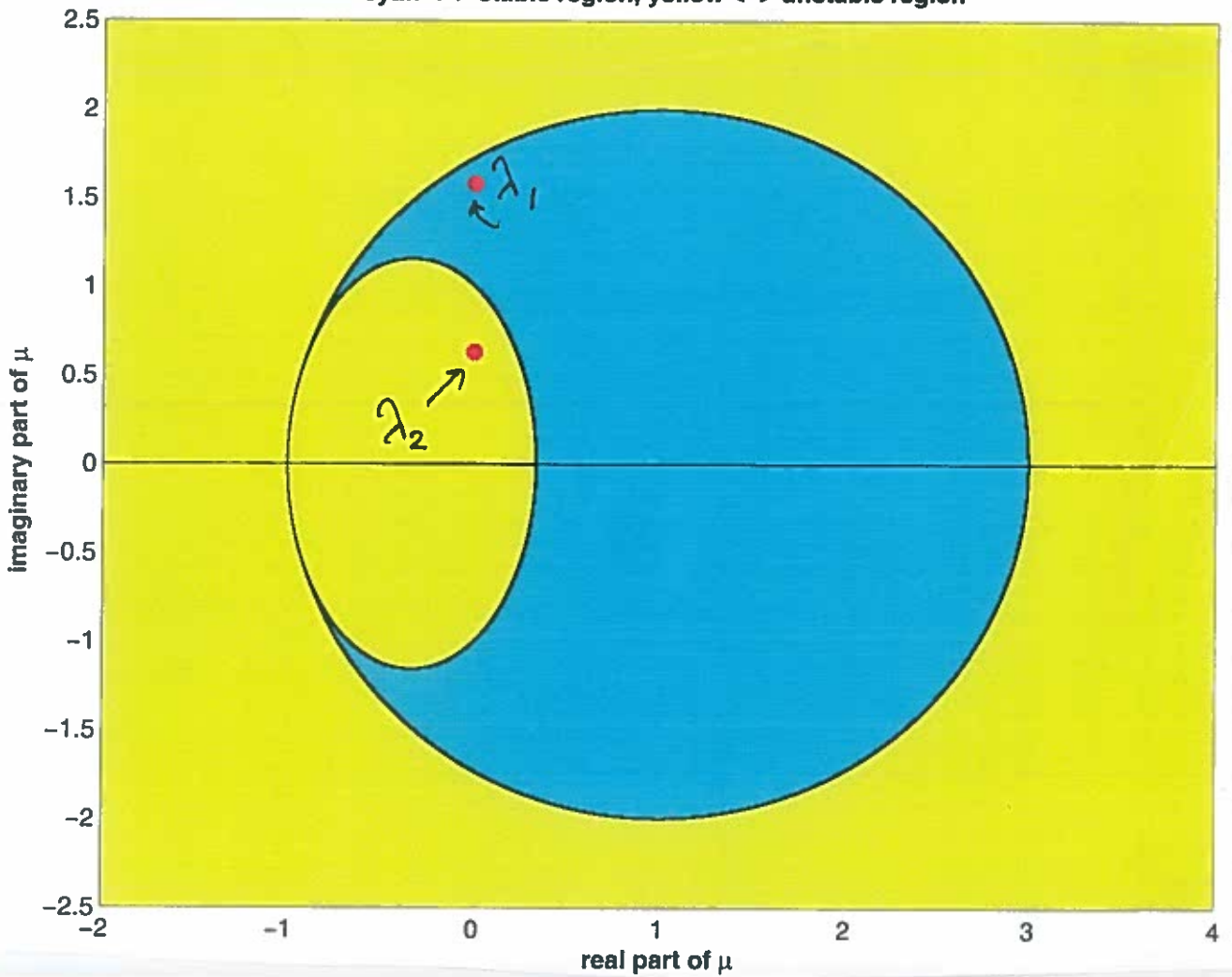
(otherwise, it may be unstable)

for a 2x2 matrix this condition can be visualized;

cyan \leftrightarrow stable region, yellow \leftrightarrow unstable region



cyan \leftrightarrow stable region, yellow \leftrightarrow unstable region



Stabilizing matrix - Newton

→ "Denman-Beavers"

$$X_{k+1} = \frac{1}{2} (X_k + \underline{X_k^{-1} A}), \quad X_0 = A$$

"symmetrize" the iteration: $\frac{1}{2} (X_k + \dots)$ → ? or $A X_k^{-1}$? (may be different!)

$$1) \quad X_{k+1} = \frac{1}{2} (X_k + \underline{A^{\frac{1}{2}} X_k^{-1} A^{\frac{1}{2}}}), \quad X_0 = A$$

Define: $2) \quad Y_k = \underline{A^{-\frac{1}{2}} X_k A^{-\frac{1}{2}}} \Rightarrow Y_k^{-1} = \underline{(A^{-\frac{1}{2}} X_k A^{-\frac{1}{2}})^{-1}} \quad 3)$

$$= (A^{-\frac{1}{2}})^{-1} X_k^{-1} (A^{-\frac{1}{2}})^{-1} \\ = A^{\frac{1}{2}} X_k^{-1} A^{\frac{1}{2}}$$

then: $5) \quad X_{k+1} = \frac{1}{2} (X_k + Y_k^{-1})$

and $6) \quad Y_{k+1} = \underline{A^{-\frac{1}{2}} X_{k+1} A^{-\frac{1}{2}}}$

$$= A^{-\frac{1}{2}} \cdot \frac{1}{2} (X_k + Y_k^{-1}) A^{-\frac{1}{2}} \\ = \frac{1}{2} \underbrace{A^{-\frac{1}{2}} X_k A^{-\frac{1}{2}}}_{= Y_k} + \frac{1}{2} \underbrace{A^{-\frac{1}{2}} Y_k^{-1} A^{-\frac{1}{2}}}_{= X_k^{-1}}$$

$$7) \quad = \frac{1}{2} (Y_k + X_k^{-1})$$

Note: $8) \quad Y_0 = I$ satisfies: $Y_0 = \underline{A^{-\frac{1}{2}} X_0 A^{-\frac{1}{2}}}$

$$= \underline{A^{-\frac{1}{2}} A A^{-\frac{1}{2}}} \\ 9) \quad = I$$

check: $10) \quad \left. \begin{aligned} A^{\frac{1}{2}} I A^{\frac{1}{2}} &= \\ \parallel A^{\frac{1}{2}} A^{-\frac{1}{2}} A A^{-\frac{1}{2}} A^{\frac{1}{2}} &= \\ A &= I A I \\ &= A \end{aligned} \right\}$

$$4) \quad A^{\frac{1}{2}} Y_k = X_k A^{-\frac{1}{2}} \\ A^{\frac{1}{2}} Y_k A^{\frac{1}{2}} = X_k \\ \Rightarrow X_k^{-1} = (A^{\frac{1}{2}} Y_k A^{\frac{1}{2}})^{-1} \\ = A^{\frac{1}{2}} Y_k^{-1} A^{\frac{1}{2}}$$

⇒ Denman-Beavers, 1976
(stable iterations, but twice as expensive)

$$\begin{cases} X_{k+1} = \frac{1}{2}(X_k + Y_k^{-1}), & X_0 = A \\ Y_{k+1} = \frac{1}{2}(Y_k + X_k^{-1}), & Y_0 = I \end{cases}$$

$k=0, 1, 2, \dots$

For checking the convergence, one could use the Frobenius-norm of a matrix: $\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^m |a_{ij}|^2}$

(in Matlab: norm(A, 'fro'))

and check $\|X_{k+1} - X_k\|_F$, for example.

[Matlab \sqrt{A} : sqrtn.m \rightsquigarrow sqrt(A)]