Note Title 2020-04-03

Newton for the matrix sign

Scalar version: $x_{k+1} = \frac{1}{2}\left(x_k + x_k^{-1}\right) \qquad x_0 = \lambda$

<u>Theorem</u>: $\lim_{k\to\infty} x_k = \operatorname{sign}(\lambda)$

(and the convergence is quadratic).

$$y = \frac{1+x}{1-x} \qquad \mathbb{C}\cup\{\infty\} \longrightarrow \mathbb{C}\cup\{\infty\}$$

$$y_k := \frac{1+x_k}{1-x_k} \qquad y_{k+1} = \frac{1+x_{k+1}}{1-x_{k+1}} = \frac{1+\frac{1}{2}(x_k+x_k^{-1})}{1-\frac{1}{2}(x_k+x_k^{-1})} = \frac{x_k^2+2x_k+1}{-x_k^2+2x_k-1}$$

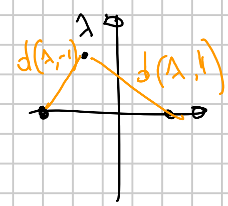
$$= -\left(\frac{1+x_k}{1-x_k}\right)^2 = -y_k^2$$

$$\begin{array}{ccccc} x & \longmapsto & \frac{1}{2}\left(x+\frac{1}{x}\right) & \longmapsto & \cdots \\ \downarrow & & \downarrow & & \\ y & \longmapsto & -y^2 & \longmapsto & \cdots \end{array}$$

$x_0 = \lambda \qquad y_0 = \frac{1+\lambda}{1-\lambda} \qquad |y_0| > 1 \text{ if } \operatorname{Re}(\lambda) > 0$

$\qquad\qquad\qquad\qquad\qquad |y_0| < 1 \text{ if } \operatorname{Re}(\lambda) < 0$

$$|y_0| = \frac{|1+\lambda|}{|1-\lambda|} = \frac{d(\lambda,-1)}{d(\lambda,1)}$$



$$y_1 = -y_0^2, \quad y_2 = -y_0^4, \quad y_3 = -y_0^8, \ \ldots$$

$$|y_k| = \left|y_0^{2^k}\right| \longrightarrow \begin{cases} 0 & \text{if } |y_0| < 1 \\ \infty & \text{if } |y_0| > 1 \end{cases}$$

so $y_k \to \begin{cases} 0 \\ \infty \end{cases}$ quadratically $\quad$ if $0 = y = \frac{1+x}{1-x} \iff x = -1$

if $\infty = y = \frac{1+x}{1-x} \iff x = \infty$

$x_k \to \begin{cases} -1 \\ 1 \end{cases}$ quadratically

Behavior of the corresponding matrix iteration

$$X_{k+1} = \frac{1}{2}\left(X_k + X_k^{-1}\right) \qquad X_0 = A$$

$$Y_k = (X_k - S)(X_k + S)^{-1} \qquad S = \mathrm{sign}(A)$$

if $A = VJV^{-1}$

$$Y_0 = (A-S)(A+S)^{-1} = V\left(J - \mathrm{sign}(J)\right)V^{-1}\,V\underbrace{\left(J + \mathrm{sign}(J)\right)^{-1}}_{\text{on diagonal: } \lambda_j + \mathrm{sign}(\lambda_j) \neq 0}V^{-1}$$

$$= V \begin{bmatrix} \ddots & & \times \\ & \overbrace{\left(\lambda_j - \mathrm{sign}(\lambda_j)\right)\left(\lambda_j + \mathrm{sign}(\lambda_j)\right)^{-1}}^{\lambda_j^{(1)}} & \\ 0 & & \ddots \quad \times \end{bmatrix} V^{-1}$$

$\lambda_j = a + bi$ $\qquad \lambda_j + \mathrm{sign}(\lambda_j) = (a + \mathrm{sign}(a)) + bi$

$\lambda_j - \mathrm{sign}(\lambda_j) = (a - \mathrm{sign}(a)) + bi$

$\Rightarrow$ all eigvls of Y_0 are $|\lambda_j| < 1$

$$Y_{k+1} = (X_{k+1} - S)(X_{k+1} + S)^{-1} = \left(\tfrac{1}{2}(X_k + X_k^{-1}) - S\right)\left(\tfrac{1}{2}(X_k + X_k^{-1}) + S\right)^{-1} =$$

$$= \underline{\left(\tfrac{1}{2}(X_k + X_k^{-1}) - S\right)(2X_k)}(2X_k)^{-1}\left(\tfrac{1}{2}(X_k + X_k^{-1}) + S\right)^{-1}$$

$$= \left(X_k^2 + I - 2SX_k\right)\left[\left(\tfrac{1}{2}(X_k + X_k^{-1}) + S\right)(2X_k)\right]^{-1}$$

$$= \left(X_k^2 + \underbrace{I}_{= S^2} - 2\underbrace{SX_k}_{SX_k = X_kS}\right)\left(X_k^2 + \underbrace{I}_{= S^2} + 2SX_k\right)^{-1}$$

$$= (X_k - S)^2 (X_k + S)^{-2} = \left[(X_k - S)(X_k + S)^{-1}\right]^2 = Y_k^2$$

$Y_k = Y_0^{2^k} \to 0$ because $\rho(Y_0) < 1$

$$Y_k = (X_k - S)(X_k + S)^{-1} \Leftrightarrow Y_k(X_k + S) = X_k - S$$

$$= (Y_k - I)X_k = (-I - Y_k)S \Rightarrow X_k = (Y_k - I)^{-1}(-I - Y_k)S$$

$\to S$ if $Y_k \to 0$

Algorithm:

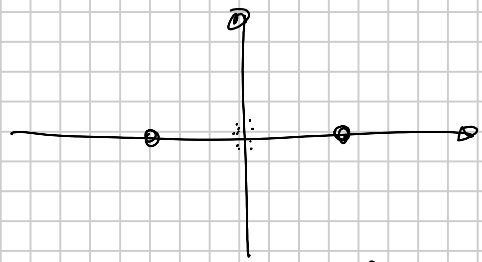
$X_0 = A$

Repeat $X_{k+1} = \frac{1}{2}(X_k + X_k^{-1})$.

you do need an inverse!

inv(X

Convergence slow when far from ± 1



How to speed up the iteration when this happens? Scaling!

$\mathrm{sign}(X) = \mathrm{sign}(\alpha X)$ for each $\alpha > 0$

Can we choose α to make sure that all eigenvalues of A have absolute value ≈ 1

Might be impossible if the eigvls of A have different scales, e.g. $\mathrm{eig}(A) = 10^5, 2, 10^{-6}(i+1)$

But still, it is better to have $10^{-5}, 2, 10^{-6}(i+1)$

than $i+1, 2\cdot 10^6, 10^{11}$ $\quad (\alpha = 10^6)$

(Q: what happens if you start the iteration $X_{k+1} = \frac{1}{2}(X_k + X_k^{-1})$ from a point on the imaginary axis?) e.g. $X_0 = i$

Idea: we want to balance the orders of magnitudes of eigenvalues so that they are "centered in 1"
(10^5 is as slow-converging as 10^{-5})

Idea 1: make geometric mean of eigenvalues $= 1$:

Cheaper! $X^{-1} = U^{-1}L^{-1}$ gives $\det(..) = \mathrm{prod}(\mathrm{diag}(U))$

$$(\lambda_1 \lambda_2 \cdots \lambda_n)^{1/n} = (\det A)^{1/n} = 1 \iff \det A = 1$$

"determinantal scaling".

Idea 2: make $[\lambda_{min}, \lambda_{max}]$ "centered around 1"

so you try to make $\alpha \lambda_{min} = (\alpha \lambda_{max})^{-1}$

need some steps of power iteration on A, A^{-1}

$$\alpha \lambda_{min} = \alpha^{-1} \lambda_{max}^{-1}$$

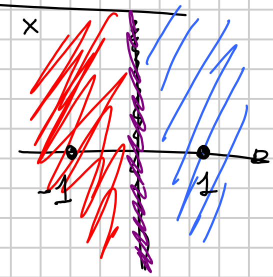
$$\alpha^2 = \frac{1}{\lambda_{min}\lambda_{max}}$$

"Spectral scaling"

Idea 3: $\sigma_{min}(\alpha A)\,\sigma_{max}(\alpha A) = 1$ "norm scaling"

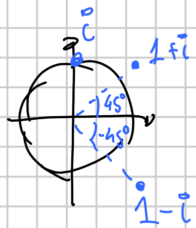
$$\overset{\shortparallel}{\|\alpha A\|}$$

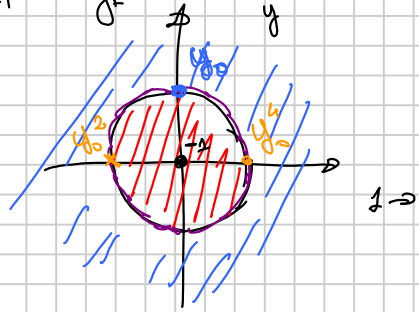
need some steps of power it. on A^TA and $A^{-T}A^{-1}$



$x_0 = i$ $\qquad x_{k+1} = \frac{1}{2}\left(x_k + x_k^{-1}\right)$?

$y_0 = \frac{1+x_0}{1-x_0} = \frac{1+i}{1-i} = i \qquad y_{k+1} = -y_k^2$

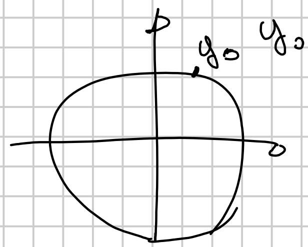




$y_0 = i$

$y_1 = -y_0^2 = -(-1) = 1$ $\qquad y_3 = -1, \quad y_4 = -1, \dots$

$y_2 = -1^2 = -1$



$x_0 = 1.1i$

Stability: complicated. Under small perturbations,

Assume (up change of basis) $A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$ $\quad \Lambda(A_{11}) \subseteq LHP$, $\Lambda(A_{22}) \subseteq RHP$

$\text{sign}(A) = \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} \qquad V_1 = \begin{bmatrix} I \\ 0 \end{bmatrix}$

$$\|\text{sign}(A+E) - \text{sign}(A)\| \lesssim \frac{\|E\|}{\left(\text{sep}(A_{11}, A_{22})\right)^3}$$

(Scales as separation^3)

However, often one is only interested in V_1, V_2 stable/antistable invariant subspaces $\quad \left(\text{sign}(A) = [V_1 | V_2] \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} [V_1 | V_2]^{-1}\right)$

If one uses a computed $\text{sign}(A+E)$ to extract these two subspaces, $\operatorname{Ker}\left(\operatorname{sign}(A+E)+I\right)=\begin{bmatrix}I\\X\end{bmatrix}$

$$\|X\| \leq \frac{\|E\|}{\operatorname{sep}(A_{11},A_{22})}$$ without the cube!

[Bai, Demmel '98, Byers, Mehrmann, He '97]

Suppose you are given M, N s.t. $A=M^{-1}N$

Can you compute $\operatorname{sign}(A)$ just from M,N without forming A?

(Natural question if you think about generalized eigenvalues / QZ algorithm)
QZ is, essentially, a method to compute eigenvalues of $A=M^{-1}N$ just from M,N without forming A)

We can implement the Newton iteration in a similar "inverse-free" fashion

input : M,N s.t. $A=M^{-1}N=X_0$

$$X_1=\tfrac12\left(A+A^{-1}\right) \overset{?}{=} M_1^{-1}N_1$$

$$\tfrac12\left(M^{-1}N+N^{-1}M^{-1}\right)=\tfrac12 M^{-1}\left(N+MN^{-1}M\right)$$

Suppose we find $\hat M,\hat N$
$MN^{-1}=\hat M^{-1}\hat N$

$$=\tfrac12 M^{-1}\left(N+\hat M^{-1}\hat N M\right)=\tfrac12 M^{-1}\hat M^{-1}\left(\hat M N+\hat N M\right)$$

$$=\underbrace{\left(\hat M M\right)^{-1}}_{M_1^{-1}}\underbrace{\left(\tfrac12\left(\hat M N+\hat N M\right)\right)}_{N_1}$$

$(M_0,N_0)\to(M_1,N_1)\to(M_2,N_2)\to(M_3,N_3)\to\dots\to(M_\infty,N_\infty)$

How to find
$MN^{-1}=\hat M^{-1}\hat N$?

$$\Leftrightarrow \hat M M=\hat N N \Leftrightarrow \begin{bmatrix}\hat M & \hat N\end{bmatrix}\begin{bmatrix}M\\-N\end{bmatrix}=0$$

$$\begin{bmatrix}\hat M & \hat N\end{bmatrix}^T=\operatorname{ker}\left(\begin{bmatrix}M\\-N\end{bmatrix}^T\right)$$ (computed e.g. with a QR decomposition)

There are cases in which computing this kernel is a lot more stable than computing M^{-1} or N^{-1}:

for instance,

$$\begin{bmatrix} M \\ -N \end{bmatrix} = \begin{bmatrix} Q & \\ & Q \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \\ \varepsilon & 0 \\ 0 & 1 \end{bmatrix} Q^T \quad \text{for a small } \varepsilon > 0$$

$$\begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \\ \varepsilon & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \cdot \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$$

$$Q_1, Q_2 \in \mathbb{R}^{4\times 2} \quad R_1 \in \mathbb{R}^{2\times 2}$$

$$Q_2^T Q_1 = 0 \Rightarrow Q_2^T = \begin{bmatrix} \hat{M} & \hat{N} \end{bmatrix}$$

("linear algebra on pencils") matrix pencil $= (-M\lambda + N)$

$$x_{k+1} = \frac{1}{2}\left(x_k + x_k^{-1}\right) = \frac{x_k^2 + 1}{2x_k}$$

← "even terms" of $(1+x_k)^2$

← "odd terms" of $(1+x_k)^2$

$$= \frac{(1+x_k)^2 + (1-x_k)^2}{(1+x_k)^2 - (1-x_k)^2}$$

The same works for higher powers:

$$x_{k+1} = \frac{(1+x_k)^3 + (1-x_k)^3}{(1+x_k)^3 - (1-x_k)^3} = \frac{1+3x_k+3x_k^2+x_k^3 + (1-3x_k+3x_k^2-x_k^3)}{1+3x_k+3x_k^2+x_k^3 - (1-3x_k+3x_k^2-x_k^3)}$$

$$= \frac{2(1+3x_k^2)}{2(3x_k+x_k^3)}$$

has fixed points in ± 1 of order 3

Problem: it is no longer true that $x_k \to \begin{cases} -1 & \text{if } \operatorname{Re} x_0 < 0, \\ 1 & \text{if } \operatorname{Re} x_0 > 0 \end{cases}$