The wonder of matrices in scientific computing

V. Simoncini

Dipartimento di Matematica, Università di Bologna (Italy)

valeria.simoncini@unibo.it

Some preliminary definitions

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

- A Hermitian: $\quad A = A^* \quad (A = A^T$ if A real)

Some preliminary definitions

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

- A Hermitian: $\quad A = A^*$ $\quad (A = A^T$ if A real)
- A normal: $\quad AA^* = A^*A$

  (also includes A Hermitian, skew-Hermitian, unitary, ...)

Some preliminary definitions

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

- A Hermitian: $\quad A = A^* \quad (A = A^T$ if A real)
- A normal: $\quad AA^* = A^*A$
  (also includes A Hermitian, skew-Hermitian, unitary, ...)
- A non-normal: ....none of the above

Some preliminary definitions

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

- A Hermitian: $\quad A = A^*$ $\quad (A = A^T$ if A real)
- A normal: $\quad AA^* = A^*A$

  (also includes A Hermitian, skew-Hermitian, unitary, ...)
- A non-normal: ....none of the above
- A highly non-normal: ....the nasty ones...

What "characterizes" a matrix?

If A is Hermitian, then it is its spectrum. $\lambda \in \mathbb{R}$:

$$Ax = x\lambda$$

with variational inequalities for the Rayleigh quotient:

$$\lambda_{\min} \leq \frac{x^* A x}{x^* x} \leq \lambda_{\max}, \qquad \forall 0 \neq x \in \mathbb{C}^n$$

What "characterizes" a matrix?

If A is non-normal, the spectrum is not enough ... But we can still use the Rayleigh quotient.

Field of values

$$W(A) = \left\{ \frac{x^* A x}{x^* x}, \quad 0 \neq x \in \mathbb{C}^n \right\} \subset \mathbb{C}$$

(or, Numerical range)

$\Rightarrow \sigma(A) \subset W(A)$

Possibly much broader region than for A Hermitian...

What “characterizes” a matrix?

If A is non-normal, the spectrum is not enough ... But we can still use the Rayleigh quotient.

Field of values

$$W(A) = \left\{ \frac{x^* A x}{x^* x}, \quad 0 \neq x \in \mathbb{C}^n \right\} \subset \mathbb{C}$$

(or, Numerical range)

$\Rightarrow \sigma(A) \subset W(A)$

Possibly much broader region than for A Hermitian...

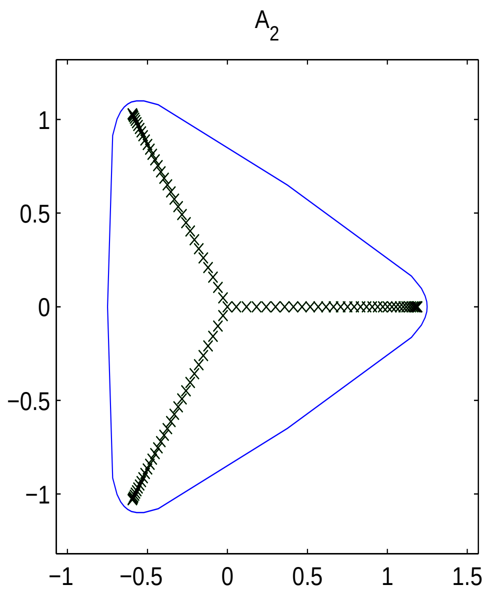
- $W(A)$ convex (Toeplitz-Haudorff theorem)
- $W(A + \alpha I) = W(A) + \alpha$ and $W(\alpha A) = \alpha W(A), \quad \forall \alpha \in \mathbb{C}$
- $W(A) \subset \mathbb{C}^+$ if and only if $A + A^* > 0$ (positive definite)

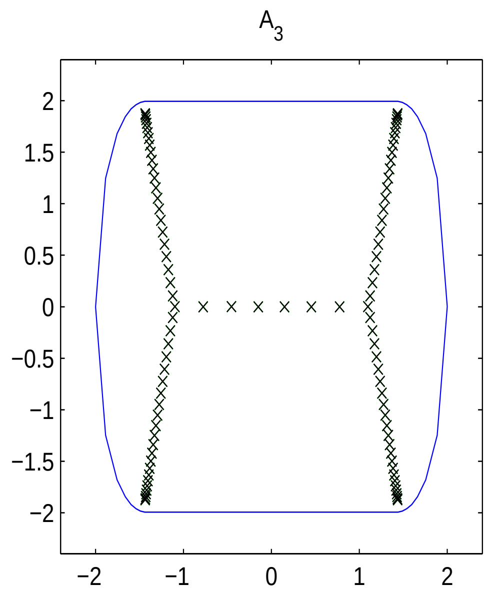
More properties of the field of values $W(A) = \{\frac{x^*Ax}{x^*x}, \quad 0 \neq x \in \mathbb{C}^n\}$

- $W(A+B) \subseteq W(A) + W(B)$
- If U unitary, then $W(A) = W(U^*AU)$.
  If $U = [U_1, *]$, then $W(U_1^*AU_1) \subseteq W(A)$
- If A normal, then $W(A) = Co(\sigma(A))$

More properties of the field of values $W(A) = \{\frac{x^*Ax}{x^*x}, \quad 0 \neq x \in \mathbb{C}^n\}$

- $W(A+B) \subseteq W(A) + W(B)$
- If U unitary, then $W(A) = W(U^*AU)$.
  If $U = [U_1, *]$, then $W(U_1^*AU_1) \subseteq W(A)$
- If A normal, then $W(A) = Co(\sigma(A))$





Some bad news

$$W(AB) \not\subseteq W(A)W(B)$$

An example:

$$A = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}$$

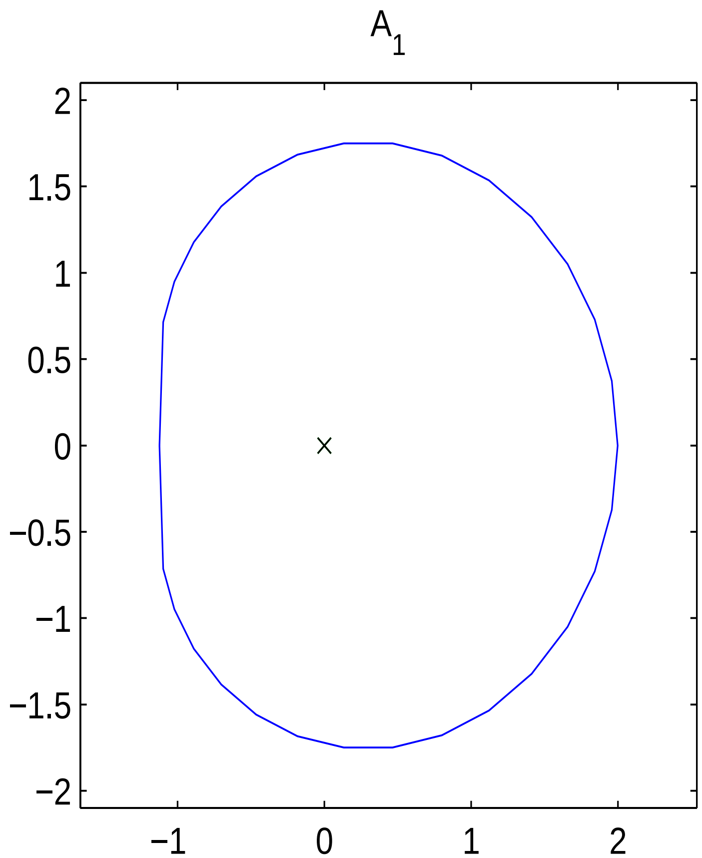
We obtain

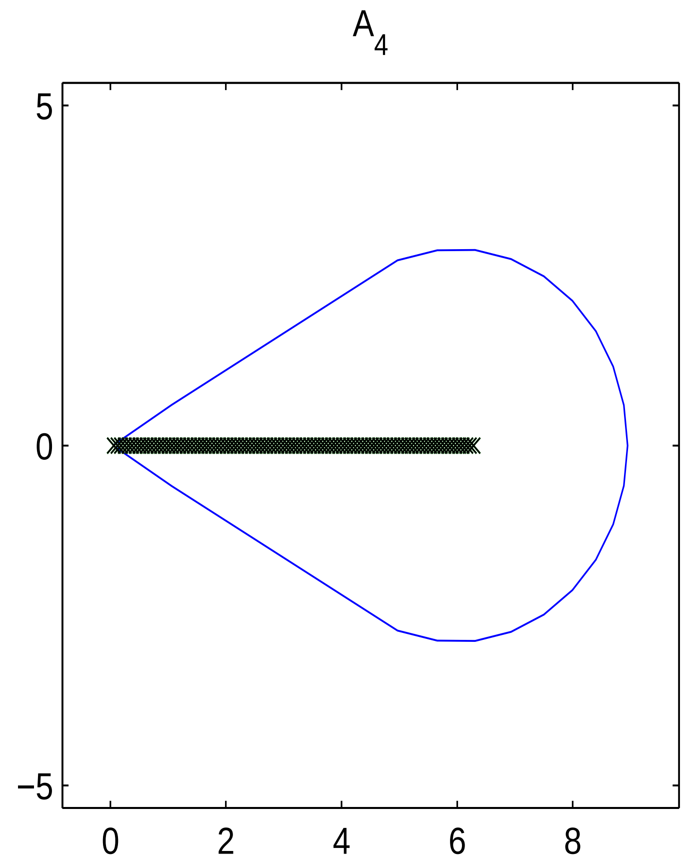
$$W(A) = W(B) = W(A)W(B) = D_1(0) \qquad \text{unit disk}$$

whereas:

$$W(AB) = W\left(\begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}\right) = [0, 4]$$

Some more intriguing examples for $W(A)$





Explicit expressions for $W(A)$ are known for general 2×2 matrices, for Jordan blocks, (oblique) projectors, etc...

Some applications: The `Tolosa` matrix.

From the "Matrix Market" repository:

"The Tolosa matrix arises in the stability analysis of a model of an airplane in flight. The interesting modes of this system are described by complex eigenvalues whose imaginary parts lie in a prescribed frequency range. The task is to compute the eigenvalues with largest imaginary parts. The problem has been analyzed at CERFACS (Centre European de Recherche et de Formation Avancee en Calcul Scientifique) in cooperation with the Aerospatiale Aircraft division.

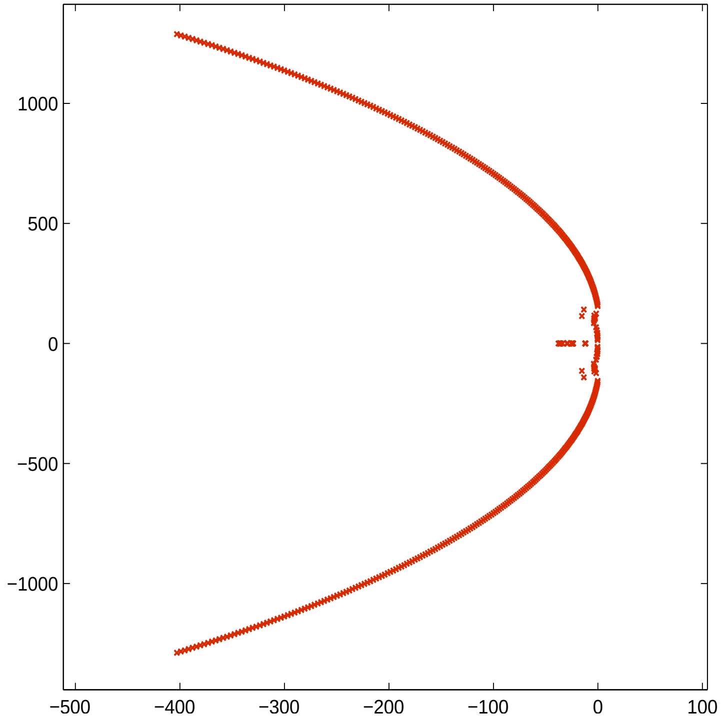
Approximation of the eigenvalues of interest, $\{\lambda_{k_1}, \ldots, \lambda_{k_m}\}$, via a projection technique:

Determine $\mathcal{V} = \mathrm{range}(V)$ such that

$$\sigma(V^* A V) \approx \{\lambda_{k_1}, \ldots, \lambda_{k_m}\}$$

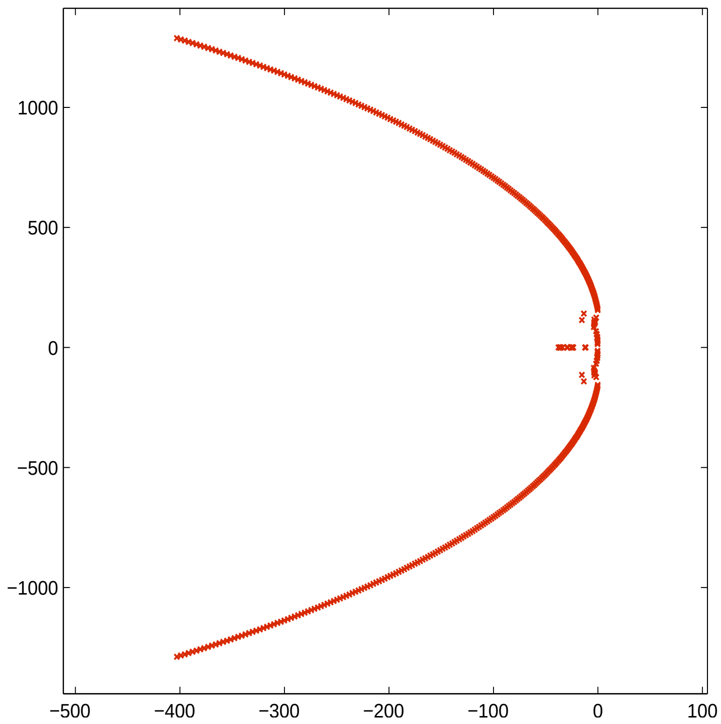
Spectral properties of the Tolosa matrix

$\sigma(A)$

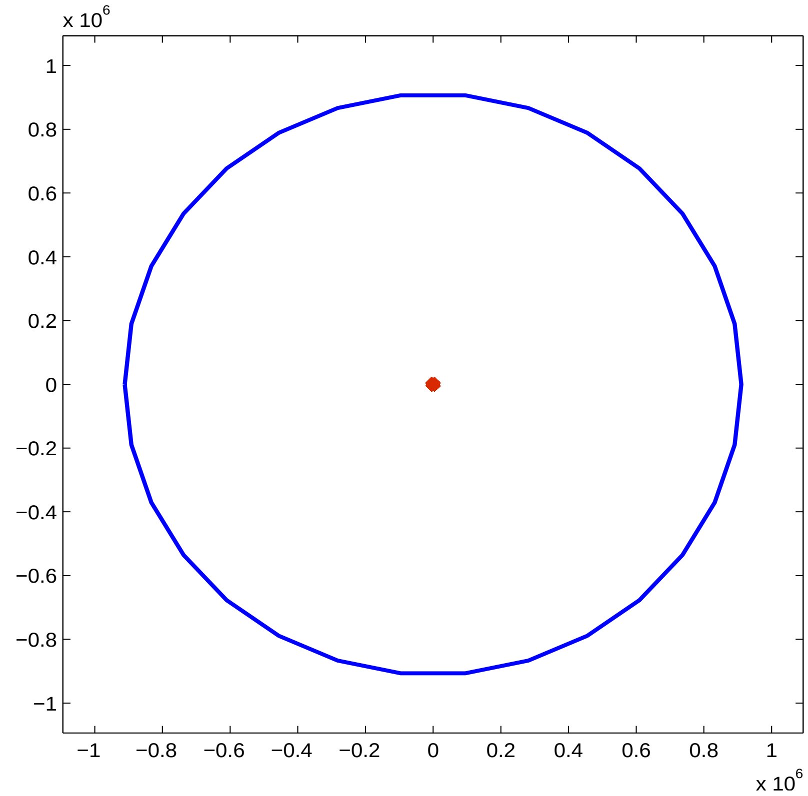


Spectral properties of the Tolosa matrix

$\sigma(A)$



$W(A)$



The role of the field of values in applications

Convergence rate for iterative linear system solvers:

Given the large linear system

$$Ax = b, \qquad A \in \mathbb{R}^{n\times n}$$

It is possible to determine an approximation x_k s.t. the residual $r_k := b - Ax_k$ satisfies

$$\|r_k\|_2 = \min_{p_k \in \mathbb{P}_k, p_k(0)=1} \|p_k(A) r_0\|_2$$

($r_0 := b - Ax_0,\ x_0$ starting approximation)

The role of the field of values in applications

Convergence rate for iterative linear system solvers:

Given the large linear system

$$Ax = b, \qquad A \in \mathbb{R}^{n \times n}$$

It is possible to determine an approximation x_k s.t. the residual $r_k := b - Ax_k$ satisfies

$$\|r_k\|_2 = \min_{p_k \in \mathbb{P}_k, p_k(0)=1} \|p_k(A) r_0\|_2$$

($r_0 := b - Ax_0,\ x_0$ starting approximation)

Then, if $0 \notin W(A)$,

$$\frac{\|r_k\|_2}{\|r_0\|_2} \leq 2 \min_{p_k \in \mathbb{P}_k, p_k(0)=1} \max_{z \in W(A)} \|p_k(z)\|$$

...often accurate bound for k small and highly non-normal A.

The role of the field of values in applications

Estimates for the norm of the resolvent

$$\frac{1}{\mathrm{dist}(z, \sigma(A))} \le \|(zI - A)^{-1}\| \le \frac{1}{\mathrm{dist}(z, W(A))}, \quad \forall z \notin W(A)$$

(for A normal, $\|(zI - A)^{-1}\| = \frac{1}{\mathrm{dist}(z,\sigma(A))}$)

that is: If z is close to $W(A)$ we may expect large $\|(zI - A)^{-1}\|$

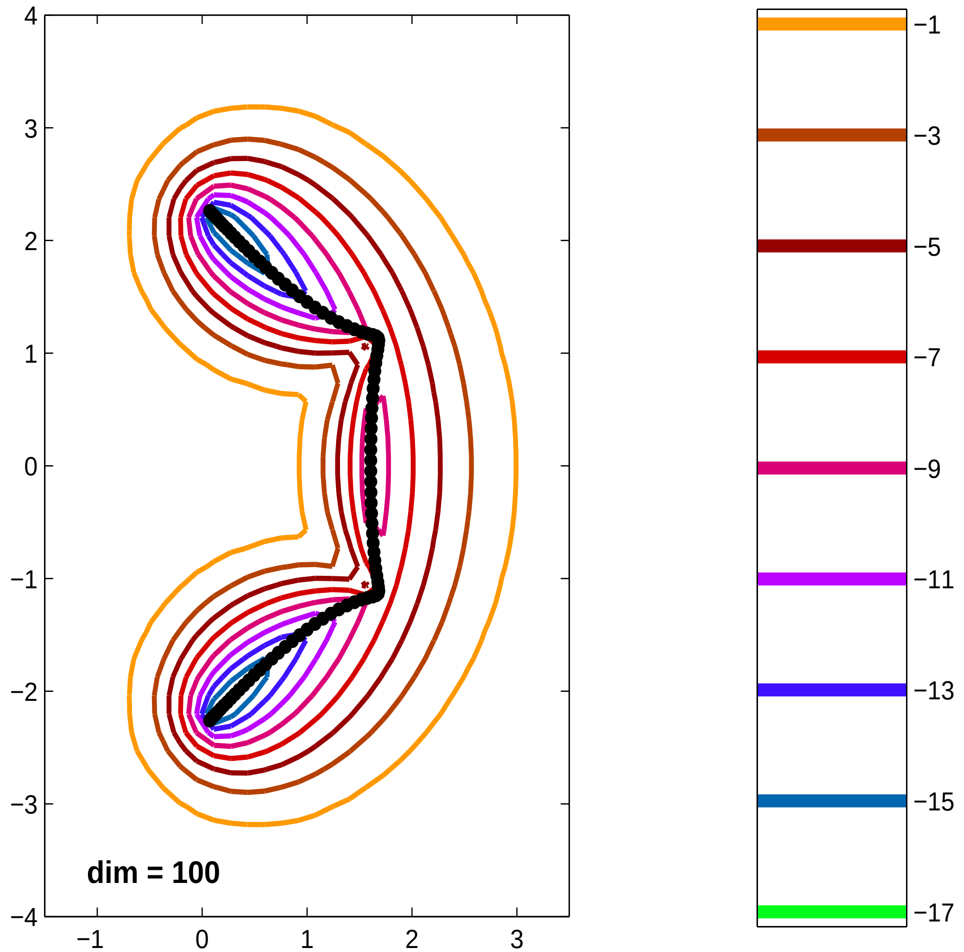
$\Rightarrow$ A very sensitive to perturbations

So, for A non-normal, closeness to singularity is substituted by

size of $\|(zI - A)^{-1}\|$

An example. The ε-pseudospectrum.

Level curves $\|(zI - A)^{-1}\| = \varepsilon$, for various values of ε



Non-normality in matrix functions

Cauchy integral formula

$$f(A) = \frac{1}{2\pi i} \int_\Gamma f(z)(zI - A)^{-1} \mathrm{dz}, \qquad \Gamma = \partial\Omega, \quad \sigma(A) \subseteq \Omega$$

Then

$$\|f(A)\| \leq \frac{|\Gamma|}{2\pi} \max_{z\in\Gamma} \|f(z)\| \, \|(zI - A)^{-1}\|$$

Non-normality in matrix functions

Cauchy integral formula

$$f(A) = \frac{1}{2\pi i} \int_{\Gamma} f(z)(zI - A)^{-1} \mathrm{dz}, \qquad \Gamma = \partial\Omega, \quad \sigma(A) \subseteq \Omega$$

Then

$$\|f(A)\| \le \frac{|\Gamma|}{2\pi} \max_{z\in\Gamma} \|f(z)\| \, \|(zI - A)^{-1}\|$$

Let:

$$\lambda_0 = \max_{\lambda\in\sigma(A)} \Re(\lambda), \qquad \alpha_0 = \max_{\lambda\in\sigma(\frac{1}{2}(A+A^*))} \lambda$$

(for real $A < 0$, α_0 is closest point of $W(A)$ to zero)

For $f(x) = e^{tx}$, it holds that

$$e^{t\lambda_0} \le \|e^{tA}\| \le e^{t\alpha_0}$$

A practical example

Evolution equation (A discretization of a spatial PDE):

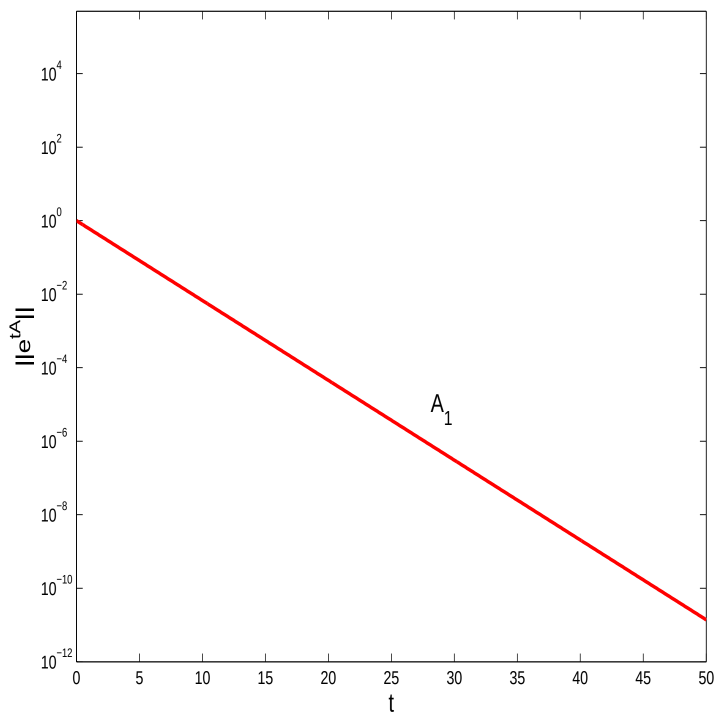
$$y' = Ay, \qquad y(t_0) = y_0$$

Solution: $y(t) = e^{tA} y_0$. Consider

$$A_1 = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 \\ -1 & -1 & -1 & 0 & 0 \\ -1 & -1 & -1 & -1 & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}, \quad A_2 = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 \\ 1 & 1 & 1 & -1 & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

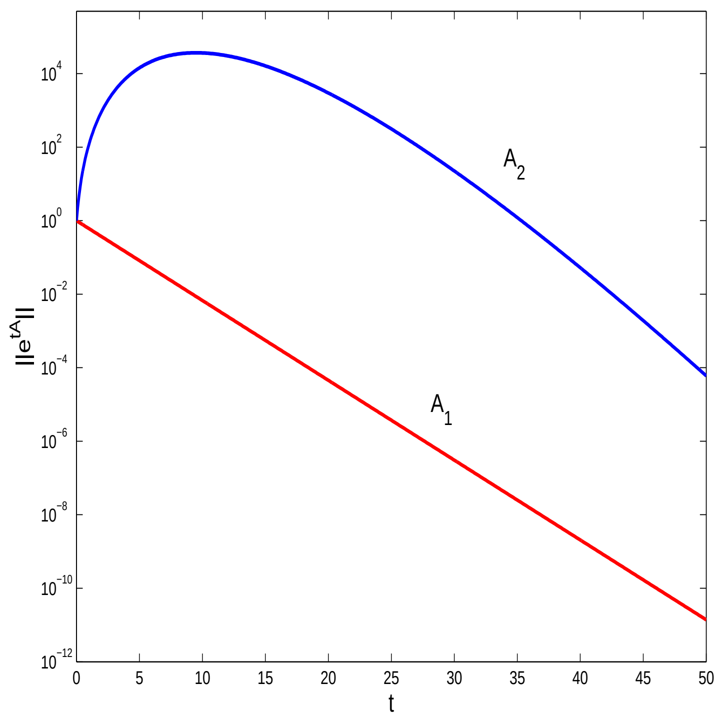
A practical example. The solution to $y' = Ay$

$\|e^{tA}\|$



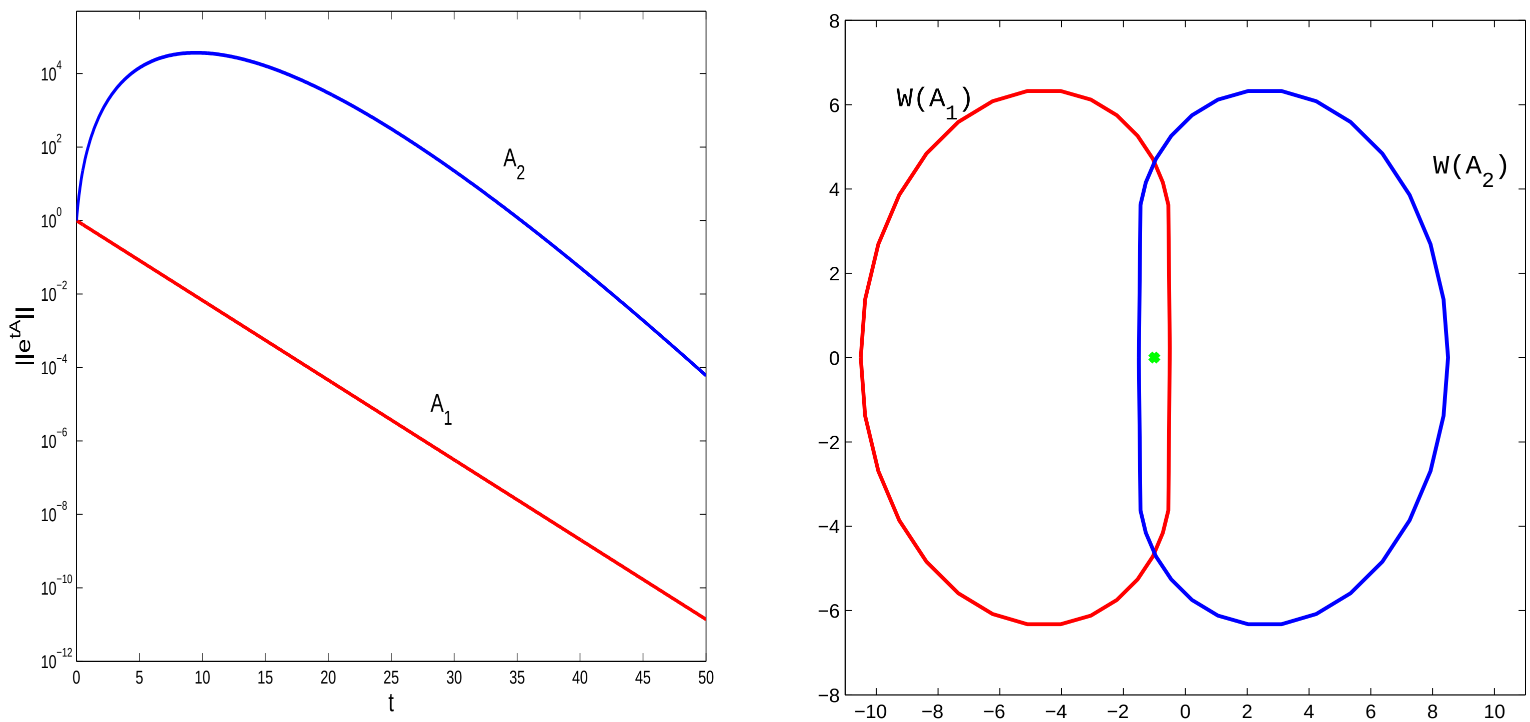
A practical example. The solution to $y' = Ay$

$$\|e^{tA}\|$$



A practical example. The solution to $y' = Ay$

$\|e^{tA}\|$ $W(A)$



Continuous time-independent dynamical systems

$$\begin{cases} x' = Ax + Bu \\ y = Cx \end{cases}$$

u: inputs, y: outputs, x : state vector

The solutions X, Y to the Lyapunov matrix equations:

$$A^*X + XA + BB^* = 0, \qquad AY + YA^* + C^*C = 0$$

provide information on the stability of the system

$\star$ There exists a unique solution if and only if $\sigma(A)$ is on one side of $\mathbb{C}$ (usually $\sigma(A) \in \mathbb{C}^-$).

Continuous time-independent dynamical systems

Crucial facts

- In the frequency/time domain:

$$X = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega I - A)^{-1}BB^*(i\omega I - A^*)^{-1}\mathrm{d}\omega = \int_0^{\infty} e^{tA}BB^*e^{tA^*}\,\mathrm{dt}$$

Continuous time-independent dynamical systems

Crucial facts

- In the frequency/time domain:

$$X = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega I - A)^{-1}BB^*(i\omega I - A^*)^{-1}\mathrm{d}\omega = \int_0^{\infty} e^{tA}BB^*e^{tA^*}\mathrm{dt}$$

- Norm of solution matrix:

$$\|X\| \leq \frac{\|BB^*\|_2}{\mathrm{sep}(A, -A^*)}$$

where

$$\mathrm{sep}(L_1, L_2) := \min_{\|P\|=1} \|L_1 P + P L_2\|$$

  - ⋆ $\mathrm{sep}(L_1, L_2)$ may be much smaller than $\min|\lambda(L_1) - \lambda(L_2)|$
  - ⋆ $\mathrm{sep}(A, -A^*)$ related to distance of $W(A)$ from the origin!

Conclusions and References

- Non-normality is far more than non-symmetry
- Intriguing phenomena occur in many different contexts

(Some) References:

R.A. Horn & C.R. Johnson, *Topics in Matrix Analysis*, Cambridge Univ. Press, 1991

L. N. Trefethen & M. Embree, *Spectra and pseudospectra*, Princeton Univ. Press, 2005