Exploiting Tropical Algebra in Numerical Linear Algebra

Françoise Tisseur
School of Mathematics
The University of Manchester

Joint work with

James Hook, Vanni Noferini, Meisam Sharify, Jennifer Pestana

Householder Symposium XIX, June 2014.

Tropical Semirings

By "**tropical**" we refer to a semiring in which the addition operation is **min** or **max**.

In this talk, consider **max-plus semiring** $(\mathbb{R}_{\max}, \oplus, \otimes)$, where $\mathbb{R}_{\max} = \mathbb{R} \cup \{-\infty\}$,

$$a \oplus b = \max(a, b), \quad a \otimes b = a + b, \quad \forall a, b \in \mathbb{R}_{\max},$$

and additive and multiplicative identities $-\infty$ and 0:

$$a \oplus -\infty = a, \quad a \otimes 0 = a.$$

Tropical algebra is the tropical analogue of linear algebra, working with matrices with entries in $\mathbb{R}_{\max}$. If $A, B \in \mathbb{R}_{\max}^{n \times n}$,

$$(A \oplus B)_{ij} = a_{ij} \oplus b_{ij}, \qquad (A \otimes B)_{ij} = \bigoplus_{k=1}^{n} a_{ik} \otimes b_{kj}.$$

When/How Can Tropical Algebra Help NLA?

When? Tropical algebra can help NLA when there are **large variations** in the magnitude of the data.

How? By providing **order of magnitude approximation** to roots, modulus of ei'vals and singular values.

- Offer good starting points for iterative algorithms.
- Can help to reduce condition numbers/backward errors.

error in solution $\lesssim$ condition number $\times$ backward error.

“Tropicalization" of Linear Algebra Problems

- We use **valuations** (provide a measure of size or multiplicity of elements of the field).
  - E.g., $x \in \mathbb{C} \mapsto \mathcal{V}(x) = \log|x| \in \mathbb{R}_{\max}$ ($\log 0 = -\infty$).
  - When $|a| \gg |b|$ or $|a| \ll |b|$ with $a, b \in \mathbb{C}$,

$$\begin{aligned}\mathcal{V}(a+b) &= \log|a+b| \\ &\approx \max(\log|a|, \log|b|) \\ &= \mathcal{V}(a) \oplus \mathcal{V}(b),\end{aligned} \qquad \begin{aligned}\mathcal{V}(ab) &= \log|ab| \\ &= \log|a| + \log|b| \\ &= \mathcal{V}(a) \otimes \mathcal{V}(b).\end{aligned}$$

- Tropicalized linear algebra problems
  - can be easier/cheaper to solve and,
  - does not suffer much from numerical instabilities.

Scalar Polynomials: Classical/Max-Plus

"**Tropicalize**" $p(x) = \sum_{i=0}^{d} a_i x^i$, $a_i \in \mathbb{C}$, i.e., construct

$$\mathrm{t}p(x) = \bigoplus_{i=0}^{d} \log|a_i| \otimes x^{\otimes i} = \max_{0 \le i \le d} (\log|a_i| + ix).$$

Let $\alpha_1 < \cdots < \alpha_p$ be roots of $\mathrm{t}p$ with α_j of multiplicity m_j.

Theorem (Sharify'11)

If $\max(\alpha_j - \alpha_{j-1}, \alpha_{j+1} - \alpha_j) \ge \log 9 \approx 2.2$ *for* $1 \le j \le p$ *then* $p(x)$ *has exactly* m_j *roots in the annulus*

$$\mathcal{A}(x) = \{x \in \mathbb{C} : \tfrac{1}{3}\exp(\alpha_j) \le |x| \le 3\exp(\alpha_j)\}.$$

Max-plus roots of $\mathrm{t}p(x)$ offer order of magnitude approx. to roots of p as long as the α_j are well separated.

Extension to Matrix Polynomials

Let $P(\lambda) = \sum_{i=0}^{d} A_i \lambda^i \in \mathbb{C}[\lambda]^{n\times n}$ and $\mathrm{t}p(x) = \bigoplus_{i=0}^{d} \log \|A_i\| \otimes x^{\otimes i}$

with max-plus roots $\alpha_1 < \cdots < \alpha_q$, α_j of multiplicity m_j.

$k_0 < \cdots < k_q$: corresponding indices in Newton polygon.

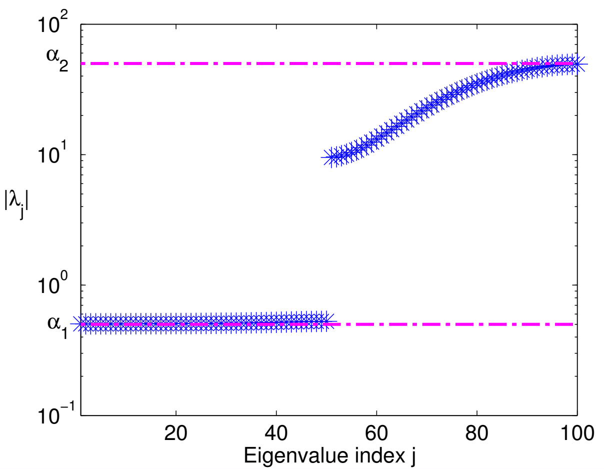
Theorem (Noferini, Sharify, T.' 14)

If $\alpha_\ell - \alpha_{\ell-1} \geq 2\log(1+2\kappa(A_{k_\ell}))$, $\ell = j-1, j$ then $P(\lambda)$ has exactly nm_j ei'vals inside the annulus $\mathcal{A}((1+2\kappa(A_{k_{j-1}}))^{-1}\exp(\alpha_j), (1+2\kappa(A_{k_j}))\exp(\alpha_j))$.

For $A_{k_{j-1}}$, A_{k_j} well conditioned and $\alpha_{j-1}, \alpha_j, \alpha_{j+1}$ sufficiently well separated, P has nm_j ei'vals of modulus close to $\exp(\alpha_j)$.

Illustration with Spring Problem

`spring`: quadratic matrix polynomial from NLEVP.
$\kappa_2(A_j) \le 5$, $j = 0\colon 2$.



Use of Max-Plus Roots in NLA

- Max-plus roots used to select **starting points** in the Ehrlich-Aberth method for polynomial eigenproblems. [Bini, Noferini, Sharify'13].

- Useful for Betcke's **diagonal scaling** aimed at improving the conditioning of P's ei'vals near ω.

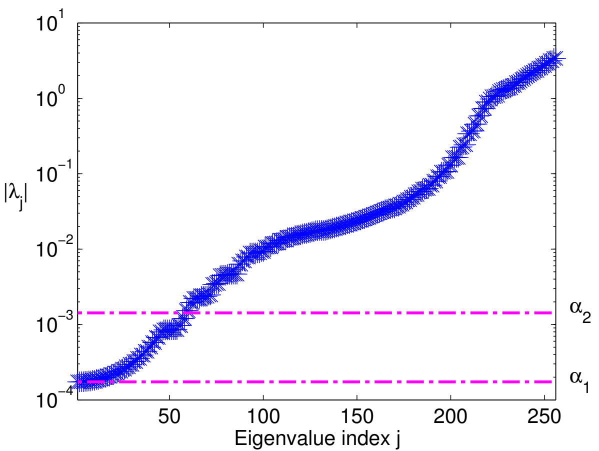
- Define **eigenvalue parameter scalings** ($\lambda = \alpha_j \mu$) for polynomial eigenvalue solvers based on linearizations.

$$\widetilde{P}(\mu) := \delta^{-1} P(\alpha_j \mu), \quad \delta = \|A_{k_{j-1}}\| \exp(k_{j-1}\alpha_j).$$

  - Allow computation of ei'pairs with small b'err for $|\lambda|$ near α_j.
  - Linearization process does not affect ei'val condition number of ei'vals near α_j.

Orr-Sommerfeld Problem

Quartic matrix polynomial from NLEVP collection.



Here e'vals of large magnitude are not captured by the max-plus roots.

Max-Plus Eigenvalues

- The **max-plus ei'vals** of $M \in \mathbb{R}_{\max}^{n\times n}$ are the roots of

$$\chi_M(\lambda) = \mathrm{perm}(M \oplus \lambda \otimes I).$$

  Here I is the identity matrix in $\mathbb{R}_{\max}^{n\times n}$.

- The n max-plus ei'vals of M can be computed in $O(zn)$ ops, where $z = nnz(M)$ using a network flow algorithm [Gassner & Klinz'10]

- The max-plus ei'vals of a max-plus matrix polynomial

$$P(\lambda) = P_0 \oplus P_1 \otimes \lambda \oplus \cdots \oplus P_d \otimes \lambda^{\otimes d}, \quad P_j \in \mathbb{R}_{\max}^{n\times n}$$

  are the max-plus roots of $\chi_P(\lambda) = \mathrm{perm}\big(P(\lambda)\big)$.

Matrices Depending on a Parameter

Let $A(t)_{ij} = b_{ij}\exp(m_{ij}t)$, $B \in \mathbb{C}^{n\times n}$, $M \in \mathbb{R}_{\max}^{n\times n}$ $(\exp(-\infty) = 0)$.

Use valuation $\mathcal{V}[f(t)] = \lim_{t\to\infty} \dfrac{\log|f(t)|}{t}$ $\implies \mathcal{V}[A] = M$.

Theorem (Akian, Bapat, Gaubert' 04)

*For all $M \in \mathbb{R}_{\max}^{n\times n}$ and all **generic** $B \in \mathbb{C}^{n\times n}$, the ei'vals $\lambda_1(t), \dots, \lambda_n(t)$ of nonsing. $A(t)$ satisfy*

$$\mathcal{V}[\lambda_i(t)] = \lim_{t\to\infty} \frac{\log|\lambda_i(t)|}{t} = \mu_i,$$

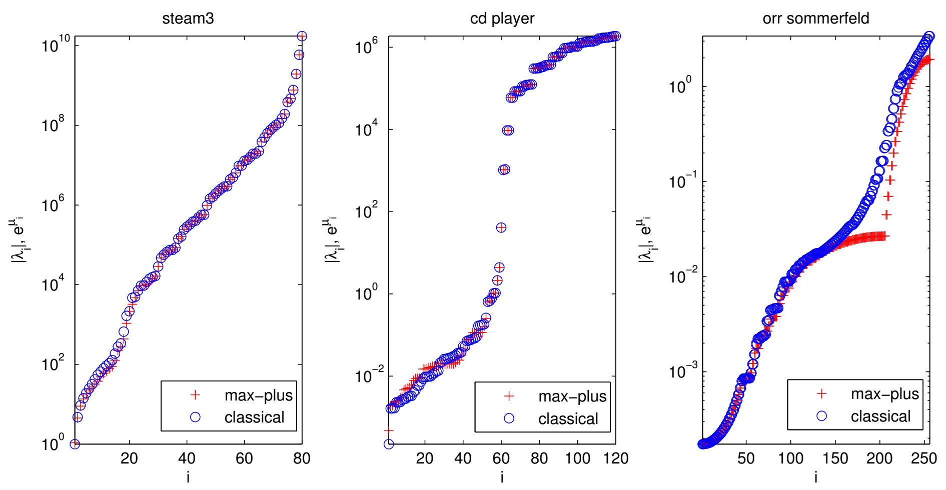
where $\mu_1 \dots, \mu_n$ are max-plus ei'vals of M's.

For $A \in \mathbb{C}^{n\times n}$ use $\mathcal{V}[x] = \log|x| \implies \big(\mathcal{V}[A]\big)_{ij} = \log|a_{ij}|$.

Classical/Max-Plus Eigenvalues

steam 3 from Florida sparse,

cd_player, orr_sommerfeld from the NLEVP.



Max-Plus Singular Values

Let $M \in \mathbb{R}_{\max}^{n\times n}$, $B \in \mathbb{C}^{n\times n}$, $A(t) = (a_{ij}(t))$, $a_{ij}(t) = b_{ij}\exp(m_{ij}t)$.

Theorem (De Schutter, De Moor' 02)

Let $A(t) = U(t)\Sigma(t)V(t)$ be the analytic SVD of $A(t)$ with $\Sigma = diag(\sigma_1(t), \dots, \sigma_n(t))$, $\sigma_j(t) = 0$, $j = n-k+1\colon n$. Then for all G and all ***generic*** *B*

$$\lim_{t\to\infty} \frac{\log \sigma_i(t)}{t} =: s_i, \quad i = 1\colon n-k$$

exists and is independent of the choice of B.

Definition

The **max-plus singular values** of M are $s_1, \dots, s_n$, with $s_1, \dots, s_{n-k}$ defined as above and $s_{n-k+1}, \dots, s_n = -\infty$.

Max-Plus Singular Values (cont.)

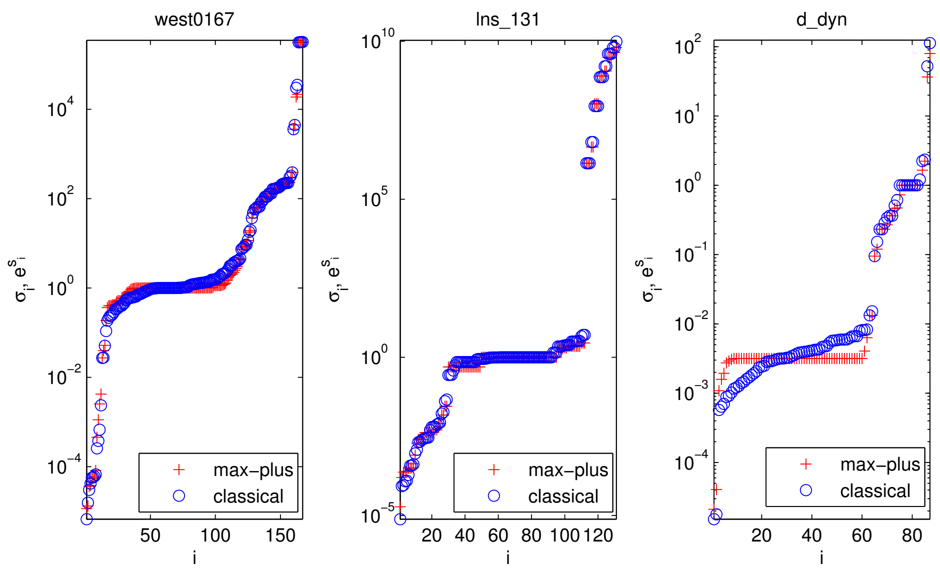
Theorem (Hook' 14)

The ***max-plus singular values*** *of $M \in \mathbb{R}_{\max}^{n\times n}$ are the max-plus ei'vals of the max-plus pencil,*

$$M \oplus \sigma \otimes 0.$$

- Charaterization extends to rectangular case.
- Max-plus singular values and e'vals of symmetric max-plus matrices are equal.
- If $A \in \mathbb{C}^{n\times n}$ with ith singular value σ_i we expect $\exp(s_i) \approx \sigma_i$, where s_i is the ith max-plus singular value of $M = (\log|a_{ij}|) \in \mathbb{R}_{\max}^{n\times n}$.

Matrices from Florida Sparse Collection



Hungarian Pair

Optimal assignment problem for $M \in \mathbb{R}_{\max}^{n \times n}$: compute

$$\mathrm{perm}(M) = \max_{\pi \in P(n)} \sum_{i=1}^{n} m_{\pi(i),i}.$$

Can be expressed as a **linear programming problem** (LPP)

$$\max\Big\{ \sum_{i,j=1}^{n} m_{ij} d_{ij} : d_{ij} > 0, \sum_{j=1}^{n} d_{ij} = \sum_{j=1}^{n} d_{ji} = 1 \ \forall i \Big\},$$

with **dual problem**

$$\min\{\sum_{i=1}^{n} u_i + v_i : u, v \in \mathbb{R}^n : m_{ij} - u_i - v_j \le 0\}.$$

An **Hungarian pair** is a solution (u, v) to the dual LPP.

Hungarian Scaling

For $A \in \mathbb{C}^{n\times n}$ let $M = \mathcal{V}[A] \in \mathbb{R}_{\max}^{n\times n}$ with $\mathcal{V}[x] = \log|x|$.

Let (u, v) be a Hungarian pair for M.

Define diagonal matrices $D_u, D_v \in \mathbb{R}^{n\times n}$ by

$$(D_u)_{ii} = \exp(-u_i), \quad (D_v)_{ii} = \exp(-v_i).$$

The **Hungarian scaling** of A is $D_u A D_v =: H$ ($|h_{ij}| \le 1$).

- Implemented in HSL-MC64 (together with some reordering).
- Commonly used before solving highly indefinite and nonsymmetric linear systems.
- Can show $\kappa_2(H) \le n \min_{D_1,D_2\in\mathcal{D}_n} \kappa_2(D_1 A D_2)$ for normal A (tends to also be the case for nonnormal A).

Numerical Experiments

- $H = \mathrm{diag}(e^{-u}) A \,\mathrm{diag}(e^{-v})$: Hungarian scaling.
- Row and column scalings with DGEEQU, $B = D_R A D_C$.
- $\min_{D_1,D_2 \in \mathcal{D}_n} \kappa_2(D_1 A D_2) \le \rho(|A||A^{-1}|) \le n \min_{D_1,D_2 \in \mathcal{D}_n} \kappa_2(D_1 A D_2)$, [Rump' 03].

Problem	$\kappa_2(A)$	$\kappa(B)$	$\kappa_2(H)$	$\rho(\lvert A\rvert\lvert A^{-1}\rvert)$
d_dyn	7.4e+6	8.4e+2	1.6e+1	6.2e+0
rotor1	2.4e+12	1.6e+8	3.0e+1	1.3e+0
lns_131	1.3e+15	1.3e+6	6.0e+1	1.8e+1
west0132	4.2e+11	2.8e+6	2.1e+2	6.5e+1
nnc261	2.9e+14	1.4e+12	4.5e+4	1.4e+4

Hungarian Scaling and Max-Plus Sing. Val.

Theorem (Hook'14)

Let $u, v \in \mathbb{R}^n$ and $M \in \mathbb{R}_{\max}^{n\times n}$. The matrix

$$\mathrm{diag}(-u) \otimes M \otimes \mathrm{diag}(-v)$$

has all of its max-plus singular values equal to zero iff (u, v) is a Hungarian pair for M.

Now if $A \in \mathbb{C}^{n\times n}$ has entries which varies a lot in magnitude and (u, v) is a Hungarian pair for $M = \mathcal{V}[A] \in \mathbb{R}_{\max}^{n\times n}$ then (heuristically)

$$\mathrm{diag}(e^{-u})A\,\mathrm{diag}(e^{-v})$$

should have its singular values map to $\exp(0) = 1$.

Summary

- Max-plus roots, eigenvalues and singular values can offer order of magnitude approximations to their "classical" analogs for problems with large variations in the magnitude of the data.
- Plan to use tropical algebra to produce a polynomial eigensolver with better numerical stability property than eigensolvers such as `polyeig`.
- Investigate the effect of Hungarian scaling on eigenvalue condition numbers and backward errors.