Tropical aspects of eigenvalue computation problems

Stephane.Gaubert@inria.fr

INRIA and CMAP, École Polytechnique

Séminaire Algo
Lundi 11 Janvier 2010

Synthesis of: Akian, Bapat, SG CRAS 2004, arXiv:0402090; SG, Sharify POSTA 09; and current work...

Tropical / max-plus algebra $\mathbb{R}_{\max} := \mathbb{R} \cup \{-\infty\}$ equipped with

$$\text{“}a + b\text{''} = \max(a, b) \qquad \text{“}ab\text{''} = a + b$$

Tropical algebra is hidden in the three following problems ...

1. Lidskiĭ, Višik, Ljusternik perturbation theory

Theorem (Lidskiĭ 65; also Višik, Ljusternik 60)

Let $a \in \mathbb{C}^{n\times n}$ *be nilpotent, with* m_i *Jordan blocks of size* ℓ_i*. For a generic perturbation* $b \in \mathbb{C}^{n\times n}$*, the matrix*

$$a + \epsilon b$$

has precisely $m_i \ell_i$ *eigenvalues of order* ϵ^{1/ℓ_i} *as* $\epsilon \longrightarrow 0$*.*

$$a = \left[\begin{array}{ccc|ccc|cc|c}
\cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\boxed{\cdot} & \cdot & \cdot & \boxed{\cdot} & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\
\boxed{\cdot} & \cdot & \cdot & \boxed{\cdot} & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array}\right]$$

6 eigenvalues $\sim \omega\epsilon^{1/3}$, $\omega^3 = \lambda$, λ eigenvalue of

$$\begin{bmatrix} b_{31} & b_{34} \\ b_{61} & b_{64} \end{bmatrix}$$

$$a = \left[\begin{array}{ccc|ccc|cc|c}
\cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\boxed{\cdot} & \cdot & \cdot & \boxed{\cdot} & \cdot & \cdot & \boxed{\cdot} & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\
\boxed{\cdot} & \cdot & \cdot & \boxed{\cdot} & \cdot & \cdot & \boxed{\cdot} & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\
\boxed{\cdot} & \cdot & \cdot & \boxed{\cdot} & \cdot & \cdot & \boxed{\cdot} & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array}\right]$$

2 eigenvalues $\sim \omega\epsilon^{1/2}$, $\omega^2 = \lambda$,

$$\lambda = b_{87} - \begin{bmatrix} b_{81} & b_{84} \end{bmatrix} \begin{bmatrix} b_{31} & b_{34} \\ b_{61} & b_{64} \end{bmatrix}^{-1} \begin{bmatrix} b_{37} \\ b_{67} \end{bmatrix}$$

$$
a = \left[\begin{array}{ccc|ccc|cc|c}
\cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\boxdot & \cdot & \cdot & \boxdot & \cdot & \cdot & \boxdot & \cdot & \boxdot \\
\hline
\cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\
\boxdot & \cdot & \cdot & \boxdot & \cdot & \cdot & \boxdot & \cdot & \boxdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\
\boxdot & \cdot & \cdot & \boxdot & \cdot & \cdot & \boxdot & \cdot & \boxdot \\
\hline
\boxdot & \cdot & \cdot & \boxdot & \cdot & \cdot & \boxdot & \cdot & \boxdot
\end{array}\right]
$$

1 eigenvalue $\sim \lambda\epsilon$,

$$
\lambda = b_{99} - \begin{bmatrix} b_{91} & b_{94} & b_{97} \end{bmatrix} \begin{bmatrix} b_{31} & b_{34} & b_{37} \\ b_{61} & b_{64} & b_{67} \\ b_{81} & b_{84} & b_{87} \end{bmatrix}^{-1} \begin{bmatrix} b_{39} \\ b_{69} \\ b_{89} \end{bmatrix}
$$

Lidskiĭ's approach does not give the correct orders in degenerate cases. . .

If the matrix

$$\begin{bmatrix} b_{31} & b_{34} \\ b_{61} & b_{64} \end{bmatrix}$$

has a zero-eigenvalue, then, $a + \epsilon b$ has less than 6 eigenvalues of order $\epsilon^{1/3}$.

Moreover, the Schur complement

$$b_{87} - \begin{bmatrix} b_{81} & b_{84} \end{bmatrix} \begin{bmatrix} b_{31} & b_{34} \\ b_{61} & b_{64} \end{bmatrix}^{-1} \begin{bmatrix} b_{37} \\ b_{67} \end{bmatrix}$$

is not defined, and there may be no eigenvalue of order $\epsilon^{1/2}$

Finding, in general, the correct order of magnitude of all eigenvalues (Puiseux series) $\iff$ characterizing (combinatorially) the Newton polygon of the curve

$$\{(\lambda, \epsilon) \mid \det(a + \epsilon b - \lambda I) = 0\}$$

long standing open problem (see survey Moro, Burke, Overton, SIMAX 97)

This talk: tropical algebra yields the correct order of magnitudes, in degenerate cases (new degenerate cases appear but of a higher order).

2. Computing the roots of matrix pencils

$$P(\lambda) = \lambda^2 10^{-18} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \lambda \begin{pmatrix} -3 & 10 \\ 16 & 45 \end{pmatrix} + 10^{-18} \begin{pmatrix} 12 & 15 \\ 34 & 28 \end{pmatrix}$$

- Apply the QZ algorithmb to the companion form of $P(\lambda)$ Matlab (7.3.0) [similar in Scilab]
- We get: $-Inf, -7.731e-19, Inf, 3.588e-19$

2. Computing the roots of matrix pencils

$$P(\lambda) = \lambda^2 10^{-18} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \lambda \begin{pmatrix} -3 & 10 \\ 16 & 45 \end{pmatrix} + 10^{-18} \begin{pmatrix} 12 & 15 \\ 34 & 28 \end{pmatrix}$$

- Apply the QZ algorithmb to the companion form of $P(\lambda)$ Matlab (7.3.0) [similar in Scilab]
- We get: $-Inf, -7.731e-19, Inf, 3.588e-19$
- Scaling of Fan, Lin and Van Dooren (2004): $-Inf, Inf, -3.250e-19, 3.588e-19$

2. Computing the roots of matrix pencils

$$P(\lambda) = \lambda^2 10^{-18} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \lambda \begin{pmatrix} -3 & 10 \\ 16 & 45 \end{pmatrix} + 10^{-18} \begin{pmatrix} 12 & 15 \\ 34 & 28 \end{pmatrix}$$

- Apply the QZ algorithmb to the companion form of $P(\lambda)$ Matlab (7.3.0) [similar in Scilab]
- We get: $-Inf, -7.731e-19, Inf, 3.588e-19$
- Scaling of Fan, Lin and Van Dooren (2004): $-Inf, Inf, -3.250e-19, 3.588e-19$
- tropical scaling (this talk): $-7.250E-18 \pm 9.744E-18i$, $-2.102E+17 \pm 7.387E+17i$ the correct answer (agrees with `Pari`).

3. Location of roots of polynomials

Given

$$f(z) = a_0 + a_1 z + \cdots + a_k z^k + \cdots + a_n z^n, \qquad a_i \in \mathbb{C}$$

Let $\zeta_1, \ldots, \zeta_n$ be the solutions of $f(z) = 0$, ordered by $|\zeta_1| \geq \cdots \geq |\zeta_n|$. Bound $|\zeta_i|$? E.g.,

Cauchy (1829)

$$|\zeta_1| \leq 1 + \max_{0 \leq k \leq n-1} \frac{|a_k|}{|a_n|} \ .$$

Fujiwara (1916)

$$|\zeta_1| \leq 2 \max_{0 \leq k \leq n-1} \sqrt[n-k]{\frac{|a_k|}{|a_n|}} \ .$$

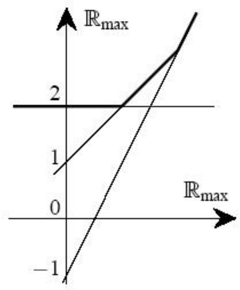
This talk:

Fujiwara's inequality is of a tropical nature

the tropical point of view yields other inequalities

Tropical polynomial functions. . .

are convex piecewise-linear with nonnegative integer slopes



$$p(x) = \text{“}(-1)x^2 + 1x + 2\text{”} = \max(-1 + 2x, 1 + x, 2)$$

“Fondamental theorem of algebra”

A tropical polynomial function

$$p(x) = \text{“}\sum_{0\leq k\leq n} b_k x^k\text{”} = \max_{0\leq k\leq n} b_k + kx \enspace .$$

can be factored uniquely (Cuninghame-Green & Meijer, 80) as

$$p(x) = \text{“}b_n \prod_{1\leq k\leq n} (x+\alpha_k)\text{”}$$
$$= b_n + \sum_{1\leq k\leq n} \max(x, \alpha_k) \enspace .$$

The points $\alpha_1, \ldots, \alpha_n$ are the tropical roots: the maximum is attained twice.

The Newton polygon Δ is the concave hull of the points (k, b_k), $k = 0, \dots, n$.

Proposition

Two formal (tropical) polynomials yield the same polynomial function iff their Newton polygons coincide

The Newton polygon Δ is the concave hull of the points (k, b_k), $k = 0, \ldots, n$.

Proposition

Two formal (tropical) polynomials yield the same polynomial function iff their Newton polygons coincide

Indeed, the function $x \mapsto \max_{0 \leq k \leq n} b_k + kx$ is the Legendre-Fenchel transform of $k \mapsto -b_k$.

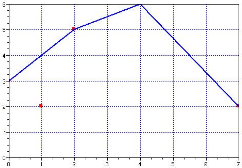
The Newton polygon Δ is the concave hull of the points (k, b_k), $k = 0, \ldots, n$.

Proposition

Two formal (tropical) polynomials yield the same polynomial function iff their Newton polygons coincide

Indeed, the function $x \mapsto \max_{0 \leq k \leq n} b_k + kx$ is the Legendre-Fenchel transform of $k \mapsto -b_k$.

The tropical roots $\alpha_1, \ldots, \alpha_k$ are the opposite of the slopes of Δ. They can be computed in $O(n)$ time.



$$p(x) = \max(2+7x, 6+4x, 5+2x, 2+x, 3)$$
$$= 2 + 2\max(-1, x) + 2\max(-1/2, x) + \max(4/3, x)$$

Associate to $f = a_0 + \cdots + a_n z^n$, $a_i \in \mathbb{C}$, the tropical polynomial

$$p(x) = \max_{0\leq k\leq n} \log|a_k| + kx \enspace .$$

The maximal tropical root is

$$\alpha_1 = \max_{1\leq k\leq n-1} \frac{\log|a_k| - \log|a_n|}{n-k}$$

Fujiwara's bound readsa

$$|\zeta_1| \leq 2 \max_{0\leq k\leq n-1} \sqrt[n-k]{\frac{|a_k|}{|a_n|}} \enspace .$$

Associate to $f = a_0 + \cdots + a_n z^n$, $a_i \in \mathbb{C}$, the tropical polynomial

$$p(x) = \max_{0\leq k\leq n} \log|a_k| + kx \enspace .$$

The maximal tropical root is

$$\alpha_1 = \max_{1\leq k\leq n-1} \frac{\log|a_k| - \log|a_n|}{n-k}$$

Fujiwara's bound readsa

$$|\zeta_1| \leq 2\exp(\alpha_1) \enspace .$$

Explanation . . . amoebas

K a field, $v : K \to \mathbb{R} \cup \{-\infty\}$ "valuation"

Ex. $K = \mathbb{C}$, $v(z) = \log|z|$,
$v(z_1 + z_2) \leq \log 2 + \max(v(z_1), v(z_2))$,
$v(z_1 z_2) = v(z_1) + v(z_2)$.

Ex. $K = \mathbb{C}\{\{\epsilon\}\}$, $v(s) = -\operatorname{val}(s)$, eg.
$v(\epsilon^{-1/2} + 3 - 8\epsilon^2 + \cdots) = 1/2$
$v(s_1 + s_2) \leq \max(v(s_1), v(s_2))$, $v(s_1 s_2) = v(s_1) + v(s_2)$,

The amoeba of $V \subset (K^*)^n$ is the set
$\{(v(z_1), \ldots, v(z_n)) \mid (z_1, \ldots, z_n) \in V\}$ (Gelfand, Kapranov, Zelevinsky)

Theorem (Kapranov)

If $f(z) = \sum_k f_k z^k \in \mathbb{C}\{\{\epsilon\}\}[z_1, \dots, z_n]$, *the amoeba of* $f = 0$ *is the set of points* $x \in \mathbb{R}^n$ *at which the maximum*

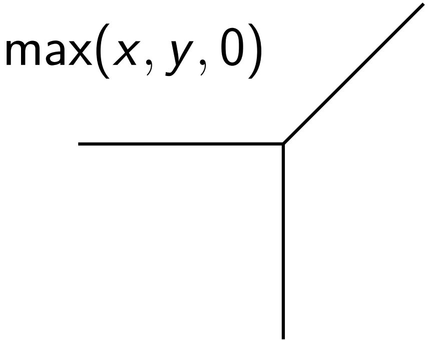
$$\max_k v(f_k) + \langle k, x \rangle$$

is attained at least twice.

Follows from Puiseux theorem when $n = 1$. Inclusion $\subset$ obvious. Converse: reduction to Puiseux.

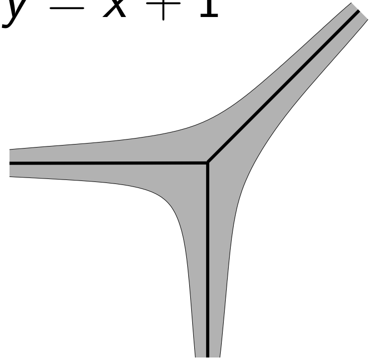
When $n = 1$: the set of tropical roots is a zero-dimensional amoeba

Example. $y = x + 1$, $K = \mathbb{C}\{\{\epsilon\}\}$



$K = \mathbb{C}$.

$$y = x + 1$$



Cf. Passare, Rüllgaard; Purbhoo

Geom. interp. of Fujiwara's bound : the open polyhedron

$$\log|z| > \log 2 + \max_{0\leq k\leq n-1}(\log|a_k| - \log|a_n|)/(n-k)$$

in the variables $\log|z|$, $\log|a_k|$, $0 \leq k \leq n$, is included in the complement of the amoeba of $a_0 + \cdots + a_n z^n = 0$ (tought of as an hypersurface of $(\mathbb{C}^*)^{n+2}$ in the variables $a_0, \ldots, a_n, z$).

The components of the complement of an amoeba are convex.

There is also a lower bound (G. Birkhoff (1914))
$\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n$ tropical roots of

$$\max_{0 \leq m \leq n} \log|a_m| - \log C_n^m + mx \quad .$$

$$\frac{1}{n}\exp(\alpha_1) \leq \exp(\beta_1) \leq |\zeta_1| \leq 2\exp(\alpha_1) \quad .$$

Theorem

$$\frac{1}{C_n^k}\exp(\alpha_1+\cdots+\alpha_k)\leq\exp(\beta_1+\cdots+\beta_k)$$
$$\leq|\zeta_1\cdots\zeta_k|\leq cst_k\exp(\alpha_1+\cdots+\alpha_k)$$

Corollary

$$cst''_{n,k}\exp(\alpha_k)\leq|\zeta_k|\leq cst'_{n,k}\exp(\alpha_k)$$

Ostrowski: $\text{cst}_k \leq 2k+1$ (hidden in his memoir on the Graeffe's method (1940), the "numerical newton polygon" is Δ). Early (simpler) instance of Viro's patchworking.

Hadamard: $\text{cst}_k \leq k+1$ (1891)

Polya: $\text{cst}_k < e\sqrt{k+1}$ (reproduced in Ostrowski).

Specht: $|\zeta_1 \cdots \zeta_k| \leq (k+1)\exp(k\alpha_1)$ (1938, weaker!), followup by Mignotte and Moussa.

Akian, Brandjesky, SG: β part of the Theorem; Akian, SG: other inequalities, eg. $\text{cst}_k \leq \sqrt{\sharp \text{ of monomials}}$

Matrix proof

Start from Kingman's inequality (61): Let $A, B \geq 0$, and $C = A^{(s)} \circ B^{(t)}$, with $s + t = 1, s, t \geq 0$ [entrywise product and exponent] then

$$\rho(C) \leq \rho(A)^s \rho(B)^t \quad .$$

I.e.

$$\log \circ \rho \circ \text{entrywise exp} \qquad \text{is convex}$$

Indeed, $\log \rho(C) = \lim_m \log \|C^m\|/m$ is a pointwise limit of convex functions of $(\log C_{ij})$, for any monotone norm. □

So

$$\rho(A \circ B) \leq \rho(A^{(p)})^{1/p} \rho(B^{(q)})^{1/q} \qquad 1/p + 1/q = 1$$

Friedland (88) observed that

$$\rho(B^{(q)})^{1/q} \to \max_{i_1,\dots,i_m} (B_{i_1 i_2} \cdots B_{i_{m-1} i_m})^{1/m} =: \rho_\infty(B)$$

and so for all $A \in \mathbb{C}^{n\times n}$,

$$\rho(A) \leq \rho(\mathrm{pattern}(A)) \rho_\infty(|A|) \leq n \rho_\infty(|A|)$$

denoting pattern(A) the corresponding $0/1$ matrix.

Apply $\rho(A) \leq \rho(\text{pattern}(A))\rho_\infty(|A|)$ to the kth exterior power of the companion matrix of f (the eigenvalues of which are $\zeta_{i_1} \dots \zeta_{i_k}$).

We get (after some combinatorics):

$$|\zeta_1 \dots \zeta_k| \leq (k+1) \exp(\alpha_1 + \cdots + \alpha_k)$$

Polya's proof uses an idea (variation on Jensen) due to Lindelöf (1902) and Landau (1905)

$$\frac{|a_0|R^k}{|\zeta_n \cdots \zeta_{n-k+1}|} \leq \exp(\frac{1}{2\pi}\int_0^{2\pi} \log|f(Re^{i\theta})|d\theta), \quad \forall R > 0$$

and setting $R = \exp(t)$,

$$\log|\zeta_n \cdots \zeta_{n-k+1}| \geq \sup_{t\in\mathbb{R}} tk - M(t)$$

where

$$M(t) = \frac{1}{2\pi}\int_0^{2\pi} \log\Big|\sum_{0\leq m\leq n} \frac{a_m}{a_0}\exp(m(i\theta + t))\Big|d\theta$$

Then, bound the geometric mean by the L_2 mean, apply Parseval, and take a well chosen t. In this way one gets $\mathrm{cst}_k \leq e\sqrt{k+1}$.

The β-bound in the theorem uses different (log-convexity) arguments.
WLOG, $a_n = 1$.

Then, denoting by cav the concave hull,

$$(\operatorname{cav} \log a)_k = \alpha_1 + \cdots + \alpha_k$$

Uses in particular $k \mapsto \log C_m^k$ convex

Application to scaling of matrix pencils

$$P(\lambda) = A_0 + A_1\lambda + \cdots + A_d\lambda^d, \qquad A_k \in \mathbb{C}^{n\times n}$$

Considering the tropical polynomial

$$p(x) = \max_{0\leq m\leq d}(\log\|A_m\| + mx)$$

with tropical roots α_i,

If each of the matrices A_k is well conditioned, we expect precisely n roots of order $\exp(\alpha_i)$, for all $1 \leq i \leq d$.

Substitute $\lambda = \exp(\alpha_i)\mu$, rescale

$$\begin{aligned}\tilde{P}(\mu) &= \exp(-p(\alpha_i))P(\lambda)\\ &= \tilde{A}_0 + \tilde{A}_1\mu + \cdots + \tilde{A}_d\mu^d\\ \tilde{A}_k &= \exp(k\alpha_i - p(\alpha_i))A_k\end{aligned}$$

For at least two indices r, s (belonging to the edge of the Netwon polygon corresponding to α_i)

$$\|\tilde{A}_r\| = \|\tilde{A}_s\| = 1$$

and

$$\|\tilde{A}_k\| \leq 1, \ \ \textit{for } k \neq r, s$$

Idea: perform such a scaling for each α_i, QZ is expected to compute accurately the group of eigenvalues of order $\exp(\alpha_i)$.

Ex, for

$$P(\lambda) = \lambda^2 10^{-18} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \lambda \begin{pmatrix} -3 & 10 \\ 16 & 45 \end{pmatrix} + 10^{-18} \begin{pmatrix} 12 & 15 \\ 34 & 28 \end{pmatrix}$$

two tropical eigenvalues, approx $-18 \log 10$ and $18 \log 10$. We called QZ once for each tropical eigenvalue, that's how we got the four complex eigenvalues:

$-7.250E - 18 \pm 9.744E - 18i$,
$-2.102E + 17 \pm 7.387E + 17i$

In the quadratic case, Fan, Lin and Van Dooren (2004) proposed a scaling with a unique call to QZ which coincides with our only when the two tropical roots coincide. When these are far away from each other, a single scaling cannot work!

Tropical splitting of eigenvalues

Definition (eigenvalue variation)

Let $\lambda_1, \ldots \lambda_n$ and $\mu_1 \ldots \mu_n$ denote two sequences of complex numbers. The variation between λ and μ is defined by

$$v(\lambda, \mu) := \min_{\pi \in S_n} \{\max_i |\mu_{\pi(i)} - \lambda_i|\} \enspace ,$$

where S_n is the set of permutations of $\{1, 2, \ldots, n\}$. If $A, B \in \mathbb{C}^{n \times n}$, the eigenvalue variation of A and B is defined by $v(A, B) := v(\operatorname{spec} A, \operatorname{spec} B)$.

Theorem (quadratic case, SG, Sharify POSTA 09)

Let $P(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0, \quad A_i \in \mathbb{C}^{n\times n}$, $\alpha^+ > \alpha^-$ *tropical roots*, $\delta := \exp(\alpha^+ - \alpha^-)$. *Let* $\zeta_1, \dots, \zeta_n$ *denote the eigenvalues of the pencil* $\lambda A_2 + A_1$, *and* $\zeta_{n+1} = \cdots = \zeta_{2n} = 0$. *Then,*

$$v(\operatorname{spec} P, \zeta) \leq \frac{C\alpha^+}{\delta^{1/2n}}$$

where

$$C := 4 \times 2^{-1/2n}\Big(2 + 2\operatorname{cond} A_2 + \frac{\operatorname{cond} A_2}{\delta}\Big)^{1-1/2n}\big(\operatorname{cond} A_2\big)^{1/2n}$$

$$\alpha^+(\operatorname{cond} A_1)^{-1} \leq |\zeta_i| \leq \alpha^+ \operatorname{cond} A_2, \qquad 1 \leq i \leq n$$

So, there are precisely n eigenvalues of the order of the maximal tropical root if

- δ (measuring the separation between the two tropical roots) is sufficiently large, and
- the matrices A_2, A_1 are well conditioned,

Under the dual assumption (A_0, A_1 well conditioned), there are precisely n eigenvalues of the order of the minimal tropical root.

- Proof relies on Bathia, Elsner, and Krause (1990):

$$\nu(\operatorname{spec} A, \operatorname{spec} B) \leq 4 \times 2^{-1/n}(\|A\| + \|B\|)^{1-1/n}\|A - B\|^{1/n}$$

Experimental results

To estimate the accuracy of computing an eigenpair, we consider the normwise backward error (Tisseur 1999)

$$\eta(\tilde{x}, \tilde{\lambda}) = \min\{\epsilon : (P(\tilde{\lambda}) + \Delta P(\tilde{\lambda}))\tilde{x} = 0, \|\Delta A_l\|_2 \leq \epsilon \|E_l\|_2\}$$

$$\eta(\tilde{x}, \tilde{\lambda}) = \frac{\|r\|_2}{\tilde{\alpha}\|\tilde{x}\|_2}$$

where $r = P(\tilde{\lambda})\tilde{x}$, $\tilde{\alpha} = \sum |\tilde{\lambda}|^l \|E_l\|_2$ and the matrices E_l represent tolerances

Backward error for quadratic pencils

- η: no scaling, η_s Fan, Lin, and Van Dooren (2004), η_t tropical
- Backward error for the 5 smallest eigenvalues of 100 randomly generated quadratic pencils,
  - $P(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0 \quad A_i \in \mathbb{C}^{10\times 10}$
  - $\|A_2\|_2 \approx 5.54 \times 10^{-5}, \|A_1\|_2 \approx 4.73 \times 10^{3}, \|A_0\|_2 \approx 6.01 \times 10^{-3}$

$\vert\lambda\vert$	$\eta(\zeta,\lambda)$	$\eta_s(\zeta,\lambda)$	$\eta_t(\zeta,\lambda)$
2.98E-07	1.01E-06	5.66E-09	6.99E-16
5.18E-07	1.37E-07	8.48E-10	2.72E-16
7.38E-07	5.81E-08	4.59E-10	2.31E-16
9.53E-07	3.79E-08	3.47E-10	2.08E-16
1.24E-06	3.26E-08	3.00E-10	1.98E-16

Backward error for a matrix pencil with an arbitrary degree

Backward error for 20 randomly generated matrix pencils

- $P(\lambda) = \lambda^5 A_5 + \lambda^4 A_4 + \lambda^3 A_3 + \lambda^2 A_2 + \lambda A_1 + A_0 \quad A_i \in \mathbb{C}^{20\times 20}$
- $\|A_5\|_2 \approx 10^5, \|A_4\|_2 \approx 10^{-4}, \|A_3\|_2 \approx 10^{-1}, \|A_2\|_2 \approx 10^2, \|A_1\|_2 \approx 10^2, \|A_0\|_2 \approx 10^{-3}$

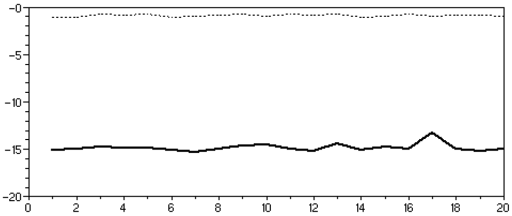


Figure: Backward error before and after scaling for the smallest eigenvalue

Backward error for a matrix pencil with an arbitrary degree

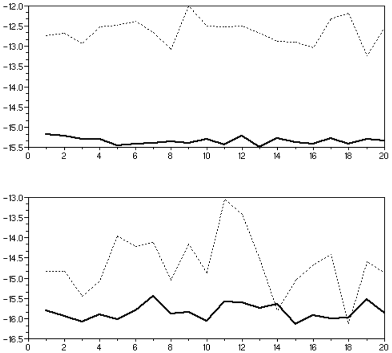


Figure: Backward error before and after scaling for the "central" 50^{th} eigenvalue and the maximum one from top to down

- The tropical roots could be used as a warning, if they are too separated, the "naive" computations are likely to be inaccurate.
- What precedes is suboptimal: calls $O(d)$ times QZ (number of times equal to the different orders of tropical eigenvalues), so the execution time can be slowed down by a factor d.
- If the matrices A_i are badly conditioned, we cannot estimate the eigenvalues based only on the norms $\|A_i\|$ ... but then we can use a finer estimation, the tropical eigenvalues...

The (algebraic) tropical eigenvalues of a matrix $A \in \mathbb{R}_{\max}^{n\times n}$ are the roots of

$$\text{“}\operatorname{per}(A + xI)\text{”}$$

where

$$\text{“}\operatorname{per}(M)\text{”} := \text{“}\sum_{\sigma\in S_n}\prod_{i\in[n]} M_{i\sigma(i)}\text{”}$$

All geom. eigenvalues λ (“$Au = \lambda u$”) are algebraic eigenvalues, but the converse does not hold.

The (algebraic) tropical eigenvalues of a matrix $A \in \mathbb{R}_{\max}^{n\times n}$ are the roots of

$$\text{“}\mathrm{per}(A + xI)\text{”}$$

where

$$\text{“}\mathrm{per}(M)\text{”} := \max_{\sigma\in S_n} \sum_{i\in[n]} M_{i\sigma(i)}$$

- All geom. eigenvalues λ (“$Au = \lambda u$”) are algebraic eigenvalues, but the converse does not hold.
- Trop. eigs. can be computed in $O(n)$ calls to an optimal assignment solver (Butkovič and Burkard) (not known whether the formal characteristic polynomial can be computed in polynomial time).

Coming back to Lidskiĭ's theory

We associate to $a + \epsilon b$, $a, b \in \mathbb{C}^{n\times n}$ the matrix $A = v(a + \epsilon b) \in \mathbb{R}_{\max}^{n\times n}$, i.e.,

$$A_{ij} = \begin{cases} 0 & \text{if } a_{ij} \neq 0 \\ -1 & \text{if } a_{ij} = 0, b_{ij} \neq 0 \\ -\infty & \text{otherwise} \end{cases}$$

Let $\gamma_1 \geq \cdots \geq \gamma_n$ trop. eigs., and let $\mathcal{L}_1(\epsilon), \ldots, \mathcal{L}_n(\epsilon)$ denote the eigenvalues of $a + \epsilon b$, $v(\mathcal{L}_1) \geq \cdots \geq v(\mathcal{L}_n)$.

Theorem (Majorization, Akian, Bapat, SG, arXiv:0402090)

$$v(\mathcal{L}_1) + \cdots + v(\mathcal{L}_k) \leq \gamma_1 + \cdots + \gamma_k$$

and $=$ for generic values of b.

Retrospectively, the bounds on the modulus of polynomial roots appear as “log-majorization” inequalities.

Not only the valuations of the eigenvalues, but their leading coefficients can be obtained: Akian, Bapat, SG CRAS 2004. . .

The idea

Replace

$$\det(a + \epsilon b - \lambda I)$$

by

$$\mathrm{diag}(\epsilon^{-U})(a + \epsilon b - \epsilon^{\gamma}\mu)\,\mathrm{diag}(\epsilon^{-V}) \to \text{nonsingular limit}(\mu)$$

How to find the scaling U, V, μ?

Dual variables

$A(\gamma) =$ "$A + \gamma I$". The dual of the linear programming formulation of the optimal assignment problem reads:

$$\text{"per } A(\gamma)\text{"} = \min \sum_i U_i + \sum_j V_j; \quad A(\gamma)_{ij} \leq U_i + V_j \ .$$

Let U, V be "Hungarian" (optimal dual) variables. By complementary slackness, a permutation σ is optimal iff it is supported by $G^s := \{(i,j) \mid A_{ij}(\gamma) = U_i + V_j\}$.

$$(a+\epsilon b)_{ij} \sim c_{ij}\epsilon^{-v(A_{ij})}$$

$c_{ij} = a_{ij}$ if $v(A_{ij}) = 0$, $c_{ij} = b_{ij}$ if $v(A_{ij}) = -1$, $c_{ij} = 0$ otherwise.

$G^0 = \{(i,j) \in G^s \mid \text{“}(A+\gamma I)_{ij}\text{”} = A_{ij}\}$,
$G^1 = \{(i,i) \in G^s \mid \text{“}(A+\gamma I)_{ii}\text{”} = \gamma\}$

$(c^G)_{ij} := c_{ij}$ if $(i,j) \in G$, 0 otherwise.

Idea: $A(\gamma)_{ij} \leq U_i + V_j$ implies, as $\epsilon \to 0$,

$$\text{diag}(\epsilon^{-U})(a+\epsilon b - \epsilon^{\gamma}\mu)\,\text{diag}(\epsilon^{-V}) \to c^{G^0} - \mu I^{G^1} \quad .$$

Theorem (Akian, Bapat, SG CRAS 2004)

If the pencil $c^{G^0} - \mu I^{G^1}$ determined from the optimal dual variables for the tropical eigenvalue γ has m non-zero eigenvalues $\lambda_1, \dots, \lambda_m$, then $a + b\epsilon$ has m eigenvalues $\sim \lambda_i \epsilon^\gamma$, and all the other ones are either $o(\epsilon^\gamma)$ or $\omega(\epsilon^\gamma)$.

Generically, $m =$ tropical multiplicity of γ, so we get the eigenvalues.

$\det(c^{G^0} - \mu I^{G^1})$ indep. of the choice of optimal dual variables (only optimal permutations matter).

Murota 90, alternative algorithmic approach: "combinatorial relaxation".

This extends Lidskiĭ's theorem: when a is nilpotent in Jordan form, $-1/\ell_i$ are precisely the tropical eigenvalues, and the boxes in

$$
\left[\begin{array}{ccc|ccc|cc|c}
\cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\boxdot & \cdot & \cdot & \boxdot & \cdot & \cdot & \boxdot & \cdot & \boxdot \\
\hline
\cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\
\boxdot & \cdot & \cdot & \boxdot & \cdot & \cdot & \boxdot & \cdot & \boxdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\
\boxdot & \cdot & \cdot & \boxdot & \cdot & \cdot & \boxdot & \cdot & \boxdot \\
\hline
\boxdot & \cdot & \cdot & \boxdot & \cdot & \cdot & \boxdot & \cdot & \boxdot
\end{array}\right]
$$

correspond to the union of the saturation digraphs for the different tropical eigenvalues.

This theorem also extends: Ma, Edelman 98; Najman 99

This explains why some attempts to extend Lidskiĭ failed:

the perturbed eigenvalues are controlled by pencils. . .

even if the original problem is a standard eigenvalue problem (not all eigenvalues of pencils can be expressed as eigenvalues of Schur complements).

Example

$$\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22}\epsilon & b_{23}\epsilon \\ b_{31} & b_{32}\epsilon & b_{33}\epsilon \end{bmatrix}$$

$$A(x) = \begin{bmatrix} 0 \oplus x & 0 & 0 \\ 0 & -1 \oplus X & -1 \\ 0 & -1 & -1 \oplus X \end{bmatrix}$$

$$P_A(x) = (x \oplus 0)^2 (x \oplus -1) \ .$$

Tropical roots: $\gamma_1 = \gamma_2 = 0$, $\gamma_3 = -1$.

If $\gamma = 0$, then $U = V = (0,0,0)$

$$A(0) = \begin{bmatrix} 0_{01} & 0_0 & 0_0 \\ 0_0 & 0_1 & 1 \\ 0_0 & 1 & 0_1 \end{bmatrix} ,$$

(0 and 1 subscripts correspond to G^0 and G^1).

$$\det \begin{bmatrix} b_{11} - \lambda & b_{12} & b_{13} \\ b_{21} & -\lambda & 0 \\ b_{31} & 0 & -\lambda \end{bmatrix} = \lambda(-\lambda^2 + \lambda b_{11} + b_{12} b_{21} + b_{31} b_{31})$$

The theorem predicts that this equation has, for generic values of the parameters b_{ij}, two non-zero roots, λ_1, λ_2, which yields two eigenvalues of $a + \epsilon b$, $\sim \lambda_m \epsilon^0 = \lambda_m$, for $m = 1, 2$.

Tropical eigenvalue $\gamma = -1$. $U = (0, -1, -1)$, $V = (1, 0, 0)$,

$$A(1) = \begin{bmatrix} 0 & 0_0 & 0_0 \\ 0_0 & -1_{01} & -1_0 \\ 0_0 & -1_0 & -1_{01} \end{bmatrix}$$

$$\det \begin{bmatrix} 0 & b_{12} & b_{13} \\ b_{21} & b_{22} - \lambda & b_{23} \\ b_{31} & b_{31} & b_{33} - \lambda \end{bmatrix} = 0 \ .$$

This yields $\lambda(b_{12}b_{21} + b_{13}b_{31}) + b_{12}b_{23}b_{31} + b_{13}b_{32}b_{21} - b_{21}b_{12}b_{33} - b_{31}b_{13}b_{22} = 0$. The theorem predicts that this equation has generically a unique nonzero root, λ_1, and that $a + \epsilon b$ has a third eigenvalue $\sim \lambda_1 \epsilon$.

Conclusion.

Simpler results in the non-archimedian case (Puiseux series).

Much remains to do in the case of log-glasses:

- finer location of the spectrum
- numerical applications.

Thank you!