On the approximation of positive definite Hankel matrices

Bernhard Beckermann

Laboratoire Painlevé, Université de Lille 1, France

NL2A, CIRM, Oct 27, 2016

Joint work with Alex Townsend (Cornell).

Supported in part by the Labex CEMPI (ANR-11-LABX-0007-01).

Outline

1. Motivation: fast multiplication with structured matrices
2. Our structure: small displacement rank
3. Decay of singular values and the Zolotarev problem
4. Estimates for the Zolotarev problem
5. Consequences for
   - Cauchy matrices
   - Pick/Löwner matrices
   - Vandermonde matrices
   - real positive definite Hankel matrices.

Fast matrix-vector multiplication

Given $X \in \mathbb{C}^{m\times n}$, we define for $k \geq 1$ the singular numbers

$$\sigma_k(X) = \min\{\|X - B\| : \mathrm{rank}(B) < k\},$$

attained for matrix X_k (with euclidean/spectral norm).

In case $m = n$ of a square matrix, we can approach the matrix-vector product Xc by $X_{k+1}c$ in complexity $\mathcal{O}(kn)$, with precision

$$\sup_{c\in\mathbb{C}^n} \frac{\|Xc - X_{k+1}c\|}{\|X\|\,\|c\|} = \frac{\sigma_{k+1}(X)}{\sigma_1(X)}.$$

With ϵ-rank

$$\mathrm{rank}_\epsilon(X) = \min\{k \geq 0 : \frac{\sigma_{k+1}(X)}{\sigma_1(X)} \leq \epsilon\}$$

we get complexity $\mathcal{O}(n\,\mathrm{rank}_\epsilon(X))$ for precision ϵ.

Fast Hadamard matrix-vector multiplication

Hadamard product $T \odot X = \left(T_{j,k} X_{j,k}\right)_{j,k}$.

Observation [Townsend, Webb & Olver'16]: Suppose that matrix-vector product Tc has complexity $\mathcal{O}(n \log(n))$ (e.g., T Toeplitz, Hankel, circulant matrix).

Then we can approach $(T \odot X)c$ by precision ϵ in complexity $\mathcal{O}(n \log(n) \operatorname{rank}_\epsilon(X))$.

Idea of proof: If $X = uv^T$ is of rank 1 then
$(T \odot X)c = \operatorname{diag}(u) T \operatorname{diag}(v) c$ has complexity $\mathcal{O}(n \log(n))$. □

Fast Hadamard matrix-vector multiplication

Hadamard product $T \odot X = \left(T_{j,k} X_{j,k}\right)_{j,k}$.

Observation [Townsend, Webb & Olver'16]: Suppose that matrix-vector product Tc has complexity $\mathcal{O}(n \log(n))$ (e.g., T Toeplitz, Hankel, circulant matrix).

Then we can approach $(T \odot X)c$ by precision ϵ in complexity $\mathcal{O}(n \log(n) \operatorname{rank}_\epsilon(X))$.

Idea of proof: If $X = uv^T$ is of rank 1 then
$(T \odot X)c = \operatorname{diag}(u) T \operatorname{diag}(v) c$ has complexity $\mathcal{O}(n \log(n))$. □

From $\|T \odot Y\|_F \leq \|T\| \, \|Y\|_F$ we get error estimate

$$\frac{\|T \odot X - T \odot X_{k+1}\|_F}{\|T\| \, \|X\|_F} \leq \sup_{j \geq 0} \frac{\sigma_{k+1+j}(X)}{\sigma_{1+j}(X)}.$$

Matrix structure through small displacement rank

Given fixed $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, the quantity

$$\rho = \mathrm{rank}(AX - XB)$$

is called (A, B)-displacement rank of $X \in \mathbb{C}^{m\times n}$ [Heinig & Rost'84].

Matrix structure through small displacement rank

Given fixed $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, the quantity

$$\rho = \mathrm{rank}(AX - XB)$$

is called (A, B)-displacement rank of $X \in \mathbb{C}^{m\times n}$ [Heinig & Rost'84].

EX1: Cauchy matrix $X = \left(\frac{1}{a_j - b_k}\right)_{j,k}$, $e = (1, 1, ..., 1)^T$

$$\mathrm{diag}(a_j)X - X\,\mathrm{diag}(b_k) = ee^T \quad \text{of rank 1}$$

EX2: Cauchy matrix pre/post multiplied by diagonal matrix
$X = \left(\frac{f_j g_k}{a_j - b_k}\right)_{j,k}$,

$$\mathrm{diag}(a_j)X - X\,\mathrm{diag}(b_k) = fg^T \quad \text{of rank } \rho = 1$$

EX3: Loewner $\left(\frac{f_j - g_k}{a_j - b_k}\right)_{j,k}$: same A, B, but of rank $\rho = 2$.
EX4: Pick=Loewner with $b_k = -\overline{a_k}$, $f_k = -\overline{g_k}$.

Matrix structure through small displacement rank

Given fixed $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, the quantity

$$\rho = \operatorname{rank}(AX - XB)$$

is called (A, B)-displacement rank of $X \in \mathbb{C}^{m\times n}$ [Heinig & Rost'84].

EX5: Vandermonde matrix $X = \left(a_j^{k-1}\right)_{j,k}$,

$$\operatorname{diag}(a_j)X - XS(\varphi) \quad \text{of rank } \rho = 1, \quad S(\varphi) = \begin{bmatrix} 0 & 0 & \cdots & 0 & \varphi \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

EX6: Krylov matrix $X = \left(A^k b\right)_k$ (including diagonal times Vandermonde)

$$AX - XS(\varphi) = fe_n^T \quad \text{of rank } 1.$$

If $m = n$ and $A = A^*$ then condition number grows [BB'00]

$$\frac{\sigma_1(X)}{\sigma_n(X)} = \|X\|\,\|X^{-1}\| \geq \frac{\exp(\frac{2\text{Catalan}}{\pi}(n-1))}{4\sqrt{n-1}}.$$

Other Examples: Hankel, Toeplitz, block versions,...

Decay of singular values

THM1: [BB, Cortona'08]
Let A, B be normal, with spectra included in $E, F \subset \mathbb{C}$.
If X has (A, B)-displacement rank ρ then for $j, k = 1, 2, \ldots$

$$\frac{\sigma_{j+\rho k}(X)}{\sigma_j(X)} \leq Z_k(E, F)$$

with the Zolotarev number

$$Z_k(E, F) := \inf_{r \in \mathcal{R}_{k,k}} \sup_{z \in E} |r(z)| \sup_{z \in F} \left| \frac{1}{r(z)} \right|.$$

Idea of proof:

Decay of singular values

THM1: [BB, Cortona'08]
Let A, B be normal, with spectra included in $E, F \subset \mathbb{C}$.
If X has (A, B)-displacement rank ρ then for $j, k = 1, 2, ...$

$$\frac{\sigma_{j+\rho k}(X)}{\sigma_j(X)} \leq Z_k(E, F)$$

with the Zolotarev number

$$Z_k(E, F) := \inf_{r \in \mathcal{R}_{k,k}} \sup_{z \in E} |r(z)| \sup_{z \in F} \left| \frac{1}{r(z)} \right|.$$

Idea of proof: Given a rational function $r \in \mathcal{R}_{k,k}$, k iterations of ADI for $AX - XB = FG^T$ with starting matrix X_j gives iterate Y with $\text{rank}(Y) \leq \text{rank}(X_j) + \rho k$ and

$$X - Y = r(A) X_j r(B)^{-1}.$$

Decay of singular values

THM1: [BB, Cortona'08]
Let A, B be normal, with spectra included in $E, F \subset \mathbb{C}$.
If X has (A, B)-displacement rank ρ then for $j, k = 1, 2, ...$

$$\frac{\sigma_{j+\rho k}(X)}{\sigma_j(X)} \leq Z_k(E, F)$$

with the Zolotarev number

$$Z_k(E, F) := \inf_{r \in \mathcal{R}_{k,k}} \sup_{z \in E} |r(z)| \sup_{z \in F} \left| \frac{1}{r(z)} \right|.$$

Idea of proof: Given a rational function $r \in \mathcal{R}_{k,k}$, k iterations of ADI for $AX - XB = FG^T$ with starting matrix X_j gives iterate Y with $\text{rank}(Y) \leq \text{rank}(X_j) + \rho k$ and

$$X - Y = r(A) X_j r(B)^{-1}.$$

COR1: Same bound up to C^2 if E, F are C-spectral for A resp. B.
COR2: Same bound up to $2C_{Crouzeix}$ if $E = W(A)$, $F = W(B)$ [BB'11].

Some facts on Zolotarev numbers

$$Z_k(E,F) := \inf_{r\in\mathcal{R}_{k,k}} \sup_{z\in E} |r(z)| \sup_{z\in F} \left|\frac{1}{r(z)}\right|.$$

1. for all $k \geq 1$
$$\exp(\frac{-k}{\mathrm{cap}(E,F)}) \leq Z_k(E,F) \leq Z_1(E,F)^k,$$
2. Asymptotically,
$$\lim_{k\to\infty} Z_k(E,F)^{1/k} = \exp(\frac{-1}{\mathrm{cap}(E,F)}).$$
3. For any Moebius transform T we have $Z_k(E,F) = Z_k(T(E),T(F))$, $\mathrm{cap}(E,F) = \mathrm{cap}(T(E),T(F))$.
4. If $E = [-1,-\lambda]$, $F = [\lambda, 1]$ for some $\lambda \in (0,1)$ then [Zolotarev'1877]
$$Z_k(E,F) \leq 4\exp(\frac{-k}{\mathrm{cap}(E,F)}) \leq 4\exp(\frac{-k\pi^2}{2\log(4/\lambda)}).$$

Some more facts on Zolotarev numbers

With the decreasing Groetsch modulus

$$\mu(\lambda) = \frac{\pi}{2}\frac{K(\sqrt{1-\lambda^2})}{K(\lambda)}, \quad K(\lambda) = \int_0^1 \frac{1}{\sqrt{(1-t^2)(1-\lambda^2 t^2)}}dt,$$

Zolotarev found out that

$$\mu(Z_k([-1,-\lambda],[\lambda,1])) = \frac{k}{\mathrm{cap}([-1,-\lambda],[\lambda,1])},$$

it remains to apply a formula for the inverse of μ

$$\kappa = 4\sqrt{q}\prod_{j=1}^{\infty}\frac{(1+q^{2j})^4}{(1+q^{2j-1})^4}, \quad q = \exp(-2\mu(\kappa)).$$

Some more facts on Zolotarev numbers

With the decreasing Groetsch modulus

$$\mu(\lambda) = \frac{\pi}{2}\frac{K(\sqrt{1-\lambda^2})}{K(\lambda)}, \quad K(\lambda) = \int_0^1 \frac{1}{\sqrt{(1-t^2)(1-\lambda^2 t^2)}} dt,$$

Zolotarev found out that

$$\mu(Z_k([-1,-\lambda],[\lambda,1])) = \frac{k}{\mathrm{cap}([-1,-\lambda],[\lambda,1])},$$

it remains to apply a formula for the inverse of μ

$$\kappa = 4\sqrt{q}\prod_{j=1}^{\infty}\frac{(1+q^{2j})^4}{(1+q^{2j-1})^4}, \quad q = \exp(-2\mu(\kappa)).$$

Mistake in [Lebedev'76], reproduced in [Medovikov & Lebedev'05], [Osedelets'07], [Druskin, Knizhnerman, Zaslavsky'09], [Güttel et.al.'14], [Nakatsukasa & Freund'15], [Bini, Massei & Robol,16],....

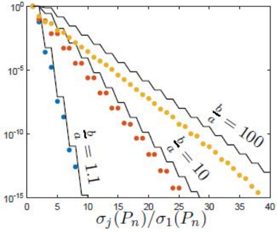
Numerical rank for Cauchy matrices ($\rho = 1$), Pick and Loewner matrices ($\rho = 2$)

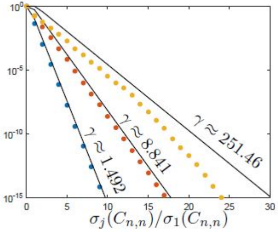
Here E, F do not depend on dimensions m, n of X.

THM2: If $a_j \in T([-1, -\lambda])$ and $b_k \in T([\lambda, 1])$ for some $\lambda \in (0, 1)$ and some Moebius transform T and X like in EX1–EX4 then

$$\mathrm{rank}_\epsilon(X) \le r := \rho \left\lceil \frac{1}{\pi^2} \log(\frac{4}{\epsilon}) \log(\frac{4}{\lambda}) \right\rceil$$

and more precisely $\sigma_{j+r}(X) \le \epsilon\, \sigma_j(X)$ for all $j \ge 1$.





Numerical rank for Vandermonde/Krylov ($\rho = 1$)

Problem with Vandermonde with real abscissa or Krylov with hermitian A: here $\sigma(A) \subset E = \mathbb{R}$ but

$$\sigma(B) = \sigma(S(-1)) = \{\exp(\frac{\pi}{n}(2j-1)) : j = 1, ..., n\} =: \Lambda_n$$

depends on n. Here n even !
First Approach : Use asymptotic results from [Gryson,BB'10] giving $\lim_{n,k\to\infty,\frac{k}{n}\to t>0} Z_k(\mathbb{R}, \Lambda_n)^{1/n}$. Problem: k gets too large.
Our approach

$$\sigma(B) \subset F_{\pi/n} \cup F_{\pi+\pi/n}, \quad F_\varphi = \{e^{it} : \varphi \le t \le \pi - \varphi\}$$

Lemma: [BB & Townsend'16] For even n,

$$\begin{aligned} Z_{2k}(\mathbb{R}, F_{\pi/n} \cup F_{\pi+\pi/n}) &\le 2\sqrt{Z_k(F_{\pi/n}, F_{\pi+\pi/n})} \\ &= 2\sqrt{Z_k([-1, -\tan^2(\frac{\pi}{2n})], [\tan^2(\frac{\pi}{2n}), 1])}. \end{aligned}$$

Numerical rank for Vandermonde/Krylov (bis)

THM3: With $X \in \mathbb{C}^{m \times n}$ Vandermonde/Krylov like in EX5–EX6

$$\text{rank}_\epsilon(X) \leq r \quad := \quad 2 + 2 \left\lceil \frac{4}{\pi^2} \log(\frac{4}{\epsilon}) \log(\frac{8\lfloor \frac{n}{2} \rfloor}{\pi}) \right\rceil$$

and more precisely $\sigma_{j+r}(X) \leq \epsilon\, \sigma_j(X)$ for all $j \geq 1$.

Numerical rank for Vandermonde/Krylov (bis)

THM3: With $X \in \mathbb{C}^{m \times n}$ Vandermonde/Krylov like in EX5–EX6

$$\operatorname{rank}_\epsilon(X) \le r \quad := \quad 2 + 2 \left\lceil \frac{4}{\pi^2} \log(\frac{4}{\epsilon}) \log(\frac{8\lfloor \frac{n}{2} \rfloor}{\pi}) \right\rceil$$

and more precisely $\sigma_{j+r}(X) \le \epsilon \, \sigma_j(X)$ for all $j \ge 1$.

Idea of proof: If n is odd then denote by $\widetilde{X} \in \mathbb{C}^{m \times (n-1)}$ (also Vandermonde/Krylov) the first $(n-1)$ columns of X. Then interlacing of singular values gives $\operatorname{rank}_\epsilon(X) \le \operatorname{rank}_\epsilon(\widetilde{X}) + 1$ and it remains to discuss the case of even n. THM1 and Lemma give bound with $\lambda = \tan^2(\frac{\pi}{4\lfloor n/2 \rfloor})$. □

Numerical rank for real semi pos. def. Hankel

THM3: With $X \in \mathbb{C}^{m \times n}$ Vandermonde/Krylov like in EX5–EX6

$$\mathrm{rank}_\epsilon(X) \leq r \quad := \quad 2 + 2\left\lceil \frac{4}{\pi^2} \log(\frac{4}{\epsilon}) \log(\frac{8\lfloor \frac{n}{2} \rfloor}{\pi}) \right\rceil$$

and more precisely $\sigma_{j+r}(X) \leq \epsilon\, \sigma_j(X)$ for all $j \geq 1$.

We use Fiedler factorization $Y = X^H X$ with X as in THM3 and obtain:

COR3: With $Y \in \mathbb{R}^{n \times n}$ semi pos. def. Hankel

$$\mathrm{rank}_\epsilon(Y) \leq r \quad := \quad 2 + 2\left\lceil \frac{2}{\pi^2} \log(\frac{16}{\epsilon}) \log(\frac{8\lfloor \frac{n}{2} \rfloor}{\pi}) \right\rceil$$

and more precisely $\sigma_{j+r}(Y) \leq \epsilon\, \sigma_j(Y)$ for all $j \geq 1$.

Further reading: arXiv:1609.09494

MERCI !

Further details about Zolotarev (1)

If $E = [-1, -\lambda]$, $F = [\lambda, 1]$ for some $\lambda \in (0, 1)$ then [Zolotarev'1877]

$$\mu(Z_k(E, F)) = k \frac{\pi^2}{2\mu(\lambda)} = \frac{k}{\text{cap}(E, F)}$$

with the decreasing Groetsch modulus

$$\mu(\lambda) = \frac{\pi}{2} \frac{K(\sqrt{1 - \lambda^2})}{K(\lambda)}, \quad K(\lambda) = \int_0^1 \frac{1}{\sqrt{(1 - t^2)(1 - \lambda^2 t^2)}} dt,$$

in particular

$$Z_k(E, F) \leq 4 \exp\left(\frac{-k}{\text{cap}(E, F)}\right) \leq 4 \exp\left(\frac{-k\pi^2}{2 \log(4/\lambda)}\right).$$

since $\mu(\lambda) \leq \log(4/\lambda)$.

Proof of Lemma

Set $E = [-1, -\lambda]$, $F = [\lambda, 1]$, $\lambda = \tan^2(\frac{\pi}{2n})$, n even.
There exists $R \in \mathcal{R}_{k,k}$ real-valued on $\mathbb{R}$ with $R(-z) = 1/R(z)$ and thus $|R(z)| \leq 1$ for $z \in i\mathbb{R}$ extremal for $Z_k(E, F)$.

Thus there exists $r \in \mathcal{R}_{k,k}$ real-valued on $\partial\mathbb{D}$ with $r(1/z) = 1/r(z)$ and $|r(z)| \leq 1$ for $z \in \mathbb{R}$ extremal for $Z_k(F_{\pi/n}, F_{\pi+\pi/n})$. Hence

$$
\begin{aligned}
&Z_{2k}(\mathbb{R}, F_{\pi/n} \cup F_{\pi/n+\pi}) \\
&\leq \max_{z\in\mathbb{R}} \left|\frac{r(w) + 1/r(w)}{2}\right| \max_{w\in F_{\pi/n}\cup F_{\pi/n+\pi}} \left|\frac{2}{r(w) + 1/r(w)}\right| \\
&= \frac{2\sqrt{Z_k(E, F)}}{1 + Z_k(E, F)} \leq 2\sqrt{Z_k(E, F)}
\end{aligned}
$$

Further details about Zolotarev (2)

Denote $\widetilde{F}_{\phi} = \{e^{it} : \varphi \leq t \leq 2\pi - \varphi\} = F^2_{\varphi/2} = F^2_{\varphi/2+\pi}$ then by considering only rational functions in z^2

$$Z_{2k}(\mathbb{R}, F_{\pi/n} \cup F_{\pi/n+\pi}) \leq Z_k([0, +\infty), \widetilde{F}_{2\pi/n})$$

Sending the circle to the imaginary axis and -1 to ∞

$$Z_{2k}([0, +\infty), \widetilde{F}_{\pi/n}) = Z_{2k}([-1, 1], i\mathbb{R} \setminus [-\tan^{-1}(\frac{\pi}{2n}), \tan^{-1}(\frac{\pi}{2n})])$$
$$\leq Z_k([0, 1], (-\infty, -\tan^{-2}(\frac{\pi}{2n})])$$

...