Singular value estimates for matrices with small displacement rank

Bernhard Beckermann, Labo Painlevé (ANO-EDP), UST Lille
bbecker@math.univ-lille1.fr
http://math.univ-lille1.fr/∼bbecker

Structured Matrices, Cortona, Sept. 20-24, 2004

Outline

-
-
-
-
-

Outline

-
- Theorem 1: rate of decay of singular values for structured matrices
  Examples: Cauchy, Loewner, Vandermonde, Krylov matrices
  ... and their block generalizations
-
-
-

Outline

- Preliminaries: singular values, Zolotarev problem
- Theorem 1: rate of decay of singular values for structured matrices
  Examples: Cauchy, Loewner, Vandermonde, Krylov matrices
  ... and their block generalizations
-
-
-

Outline

- Preliminaries: singular values, Zolotarev problem
- Theorem 1: rate of decay of singular values for structured matrices
  Examples: Cauchy, Loewner, Vandermonde, Krylov matrices
  ... and their block generalizations
- Theorem 2: the hermitian case (e.g., Hankel or block Hankel)
-
-

Outline

- Preliminaries: singular values, Zolotarev problem
- Theorem 1: rate of decay of singular values for structured matrices
  Examples: Cauchy, Loewner, Vandermonde, Krylov matrices
  ... and their block generalizations
- Theorem 2: the hermitian case (e.g., Hankel or block Hankel)
- Some comments on the "discrete" Zolotarev problem
- Conclusions

Preliminaries: singular values, displacement rank

For a matrix $X \in \mathbb{C}^{m \times n}$ we consider its singular values $\sigma_0(X) \geqslant \sigma_1(X) \geqslant \ldots$ defined by

$$\sigma_k(X) = \min\{||X - Y|| : Y \in \mathbb{C}^{m \times n}, \text{rank}\,(Y) \leqslant k\}.$$

Condition number (if $m = n$): $||X||\,||X^{-1}|| = \sigma_0(X)/\sigma_{n-1}(X)$.

Given $A \in \mathbb{C}^{m \times m}$, $B \in \mathbb{C}^{n \times n}$, the displacement rank is given by

$$\rho_{A,B}(X) = \text{rank}\,(AX - XB).$$

Examples: Cauchy, Vandermonde, Krylov ($\rho = 1$), Loewner, Pick, Toeplitz, Hankel ($\rho = 2$), block counterparts

Algebraic theory: Heinig-Rost, Kailath, Sayed, Morf, Olshevsky,...

In this talk: A, B normal, intersection of spectra $\sigma(A)$, $\sigma(B)$ is empty.

Preliminaries: singular values, displacement rank

For a matrix $X \in \mathbb{C}^{m\times n}$ we consider its singular values $\sigma_0(X) \geqslant \sigma_1(X) \geqslant \dots$ defined by

$$\sigma_k(X) = \min\{||X - Y|| : Y \in \mathbb{C}^{m\times n}, \text{rank}\,(Y) \leqslant k\}.$$

Condition number (if $m = n$): $||X||\,||X^{-1}|| = \sigma_0(X)/\sigma_{n-1}(X)$.

Given $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, the displacement rank is given by

$$\rho_{A,B}(X) = \text{rank}\,(AX - XB).$$

Examples: Cauchy, Vandermonde, Krylov ($\rho = 1$), Loewner, Pick, Toeplitz, Hankel ($\rho = 2$), block counterparts

Algebraic theory: Heinig-Rost, Kailath, Sayed, Morf, Olshevsky,...

In this talk: A, B normal, intersection of spectra $\sigma(A)$, $\sigma(B)$ is empty.

Preliminaries: singular values, displacement rank

For a matrix $X \in \mathbb{C}^{m\times n}$ we consider its singular values $\sigma_0(X) \geqslant \sigma_1(X) \geqslant ...$ defined by

$$\sigma_k(X) = \min\{||X - Y|| : Y \in \mathbb{C}^{m\times n}, \text{rank}\,(Y) \leqslant k\}.$$

Condition number (if $m = n$): $||X||\,||X^{-1}|| = \sigma_0(X)/\sigma_{n-1}(X)$.

Given $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, the displacement rank is given by

$$\rho_{A,B}(X) = \text{rank}\,(AX - XB).$$

Examples: Cauchy, Vandermonde, Krylov ($\rho = 1$), Loewner, Pick, Toeplitz, Hankel ($\rho = 2$), block counterparts

Algebraic theory: Heinig-Rost, Kailath, Sayed, Morf, Olshevsky,...

In this talk: A, B normal, intersection of spectra $\sigma(A)$, $\sigma(B)$ is empty.

Preliminaries: singular values, displacement rank

For a matrix $X \in \mathbb{C}^{m\times n}$ we consider its singular values $\sigma_0(X) \geqslant \sigma_1(X) \geqslant ...$ defined by

$$\sigma_k(X) = \min\{||X - Y|| : Y \in \mathbb{C}^{m\times n}, \text{rank}\,(Y) \leqslant k\}.$$

Condition number (if $m = n$): $||X||\,||X^{-1}|| = \sigma_0(X)/\sigma_{n-1}(X)$.

Given $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, the displacement rank is given by

$$\rho_{A,B}(X) = \text{rank}\,(AX - XB).$$

Examples: Cauchy, Vandermonde, Krylov ($\rho = 1$), Loewner, Pick, Toeplitz, Hankel ($\rho = 2$), block counterparts

Algebraic theory: Heinig-Rost, Kailath, Sayed, Morf, Olshevsky,...

In this talk: A, B normal, intersection of spectra $\sigma(A)$, $\sigma(B)$ is empty.

Preliminaries: the third Zolotarev problem

Given closed $E, F \subset \mathbb{C}$ and $k \geqslant 0$, find a rational function $r = P/Q$, $\deg P \leqslant k$, $\deg Q \leqslant k$, minimizing the expression

$$\frac{\sup_{z \in E} |r(z)|}{\inf_{z \in F} |r(z)|}.$$

Extremal function exists, minimal value: $Z_k(E, F)$.

Clearly, $Z_k(E, F)$ is $\searrow$ in k and $\nearrow$ in E, F. Also, $Z_k(E, F)^{1/k}$ is $\searrow$ in k.
Zolotarev gave explicit solution for E, F real intervals.
If E, F of positive logarithmic capacity:

$$\lim_{k \to \infty} Z_k(E, F)^{1/k} = \exp(-1/\mathrm{cap}\,(E, F)),$$

where $\mathrm{cap}\,(E, F)$ logarithmic capacity of a condenser with plates E and F.

Preliminaries: the third Zolotarev problem

Given closed $E, F \subset \mathbb{C}$ and $k \geqslant 0$, find a rational function $r = P/Q$, $\deg P \leqslant k$, $\deg Q \leqslant k$, minimizing the expression

$$\frac{\sup_{z \in E} |r(z)|}{\inf_{z \in F} |r(z)|}.$$

Extremal function exists, minimal value: $Z_k(E, F)$.

Clearly, $Z_k(E, F)$ is $\searrow$ in k and $\nearrow$ in E, F. Also, $Z_k(E, F)^{1/k}$ is $\searrow$ in k.
Zolotarev gave explicit solution for E, F real intervals.
If E, F of positive logarithmic capacity:

$$\lim_{k \to \infty} Z_k(E, F)^{1/k} = \exp(-1/\mathrm{cap}\,(E, F)),$$

where $\mathrm{cap}\,(E, F)$ logarithmic capacity of a condenser with plates E and F.

Preliminaries: the third Zolotarev problem

Given closed $E, F \subset \mathbb{C}$ and $k \geqslant 0$, find a rational function $r = P/Q$, $\deg P \leqslant k$, $\deg Q \leqslant k$, minimizing the expression

$$\frac{\sup\limits_{z \in E} |r(z)|}{\inf\limits_{z \in F} |r(z)|}.$$

Extremal function exists, minimal value: $Z_k(E, F)$.

Clearly, $Z_k(E, F)$ is $\searrow$ in k and $\nearrow$ in E, F. Also, $Z_k(E, F)^{1/k}$ is $\searrow$ in k.
Zolotarev gave explicit solution for E, F real intervals.
If E, F of positive logarithmic capacity:

$$\lim_{k \to \infty} Z_k(E, F)^{1/k} = \exp(-1/\mathrm{cap}\,(E, F)),$$

where $\mathrm{cap}\,(E, F)$ logarithmic capacity of a condenser with plates E and F.

Theorem 1 (Cauchy, Krylov, Vandermonde, Loewner, Pick)

For any integers $j, k \geqslant 0$ with $\rho = \rho_{A,B}(X)$

$$\frac{\sigma_{j+\rho k}(X)}{\sigma_j(X)} \leqslant Z_k(\sigma(A), \sigma(B)).$$

Result is "sharp" for $j = 0$: for any disjoint closed $E, F \subset \mathbb{C}$ there exist diagonal A, B and a matrix X with

$$\sigma(A) \subset E, \quad \sigma(B) \subset F, \quad \text{and} \quad \rho_{A,B}(X) \leqslant \rho,$$

and

$$\frac{\sigma_{\rho k}(X)}{\sigma_0(X)} \geqslant \epsilon_k \, Z_k(E, F)$$

with $\epsilon_k = 1$ if convex hulls of E and F are real with empty intersection, and $\epsilon_k \geqslant 1/(k+1)^2$ else. BB'03

Theorem 1 (Cauchy, Krylov, Vandermonde, Loewner, Pick)

For any integers $j, k \geqslant 0$ with $\rho = \rho_{A,B}(X)$

$$\frac{\sigma_{j+\rho k}(X)}{\sigma_j(X)} \leqslant Z_k(\sigma(A), \sigma(B)).$$

Result is "sharp" for $j = 0$: for any disjoint closed $E, F \subset \mathbb{C}$ there exist diagonal A, B and a matrix X with

$$\sigma(A) \subset E, \quad \sigma(B) \subset F, \quad \text{and} \quad \rho_{A,B}(X) \leqslant \rho,$$

and

$$\frac{\sigma_{\rho k}(X)}{\sigma_0(X)} \geqslant \epsilon_k \, Z_k(E, F)$$

with $\epsilon_k = 1$ if convex hulls of E and F are real with empty intersection, and $\epsilon_k \geqslant 1/(k+1)^2$ else. BB'03

Example Thm 1: (scaled) Cauchy/Pick matrices

$$X = \Big[\frac{1}{a_j - b_k}\Big]_{j,k}, \quad A = \mathrm{diag}\,(a_1, ..., a_m), \quad B = \mathrm{diag}\,(b_1, ..., b_n), \quad \rho = 1,$$

$$X = \Big[\frac{1}{1 - a_j \overline{a_k}}\Big]_{j,k}, \quad B = A^{-*}, \quad \rho = 1,$$

$$X = \Big[\frac{d_j + d_k}{a_j + a_k}\Big]_{j,k}, \quad B = -A, \quad \rho = 2.$$

Pre- or post-multiplication of X with diagonal matrix: same displacement rank.

Size of $Z_k([-1,-\kappa],[\kappa,1])$ with $\kappa \in (0,1)$? (other intervals by transformation)

$$\gamma^k \leqslant Z_k \leqslant 4\,\gamma^k, \qquad \gamma = \exp(-\frac{1}{\mathrm{cap}\,([-1,-\kappa],[\kappa,1])}) = \exp(-\frac{2\pi K(\kappa)}{K(\sqrt{1-\kappa^2})}).$$

Also,

$$Z_k \leqslant Z_1^k = \Big[\frac{1-\sqrt{\kappa}}{1+\sqrt{\kappa}}\Big]^{2k} \leqslant \Big[\frac{1-\kappa}{1+\kappa}\Big]^{2k}$$

(last bound: Olshevsky & Fasino '02)

Example Thm 1: (scaled) Cauchy/Pick matrices

$$X = \Big[\frac{1}{a_j - b_k}\Big]_{j,k}, \quad A = \text{diag}\,(a_1, ..., a_m), \quad B = \text{diag}\,(b_1, ..., b_n), \quad \rho = 1,$$

$$X = \Big[\frac{1}{1 - a_j \overline{a_k}}\Big]_{j,k}, \quad B = A^{-*}, \quad \rho = 1,$$

$$X = \Big[\frac{d_j + d_k}{a_j + a_k}\Big]_{j,k}, \quad B = -A, \quad \rho = 2.$$

Pre- or post-multiplication of X with diagonal matrix: same displacement rank.
Size of $Z_k([-1,-\kappa],[\kappa,1])$ with $\kappa \in (0,1)$? (other intervals by transformation)

$$\gamma^k \leqslant Z_k \leqslant 4\,\gamma^k, \qquad \gamma = \exp(-\frac{1}{\text{cap}\,([-1,-\kappa],[\kappa,1])}) = \exp(-\frac{2\pi K(\kappa)}{K(\sqrt{1-\kappa^2})}).$$

Also,

$$Z_k \leqslant Z_1^k = \Big[\frac{1-\sqrt{\kappa}}{1+\sqrt{\kappa}}\Big]^{2k} \leqslant \Big[\frac{1-\kappa}{1+\kappa}\Big]^{2k}$$

(last bound: Olshevsky & Fasino '02)

Example Thm 1: (scaled) Vandermonde/Krylov matrices

A diagonal or not, $b \in \mathbb{C}^m$, $B = S_n(\theta)$ defined by

$$S_n(\theta) = \begin{bmatrix} 0 & \cdots & 0 & \theta \\ 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & \cdots 0 & 1 & 0 \end{bmatrix} \in \mathbb{C}^{n \times n}, \quad |\theta| = 1 \text{ (spectrum rotated } n\text{th roots of unity).}$$

Example Krylov: $X = K_n(A, b) = (b, Ab, A^2 b, ..., A^{n-1} b)$, we have $\rho_{A,B}(X) \leqslant 1$.

Example Vandermonde: here A diagonal, $b = (1, ..., 1)^*$ (different b: row scaling).

Different results by Gautschi '75-90, Inglese '88, Tyrtyshnikov '94.

$$\sigma(A) \subset \mathbb{R} : \quad \frac{\sigma_{n-1}(K_n(A, b))}{\sigma_0(K_n(A, b))} \leqslant 4\sqrt{n}\, 1.792^{1-n} \qquad \text{BB'00}$$

sharp up to $\mathcal{O}(\sqrt{n})$ (slightly larger bound is obtained via $Z_{n-1}(\mathbb{R}, \sigma(S_n(\theta)))$).

Example Thm 1: (scaled) Vandermonde/Krylov matrices

A diagonal or not, $b \in \mathbb{C}^m$, $B = S_n(\theta)$ defined by

$$S_n(\theta) = \begin{bmatrix} 0 & \cdots & 0 & \theta \\ 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & \cdots 0 & 1 & 0 \end{bmatrix} \in \mathbb{C}^{n \times n}, \quad |\theta| = 1 \ \text{(spectrum rotated } n\text{th roots of unity).}$$

Example Krylov: $X = K_n(A, b) = (b, Ab, A^2 b, ..., A^{n-1} b)$, we have $\rho_{A,B}(X) \leqslant 1$.

Example Vandermonde: here A diagonal, $b = (1, ..., 1)^*$ (different b: row scaling).
Different results by Gautschi '75-90, Inglese '88, Tyrtyshnikov '94.

$$\sigma(A) \subset \mathbb{R}: \quad \frac{\sigma_{n-1}(K_n(A,b))}{\sigma_0(K_n(A,b))} \leqslant 4\sqrt{n}\, 1.792^{1-n} \qquad \text{BB'00}$$

sharp up to $\mathcal{O}(\sqrt{n})$ (slightly larger bound is obtained via $Z_{n-1}(\mathbb{R}, \sigma(S_n(\theta)))$).

Application Thm 1: linear system theory

Transfer function of MIMO linear time-invariant discrete-time system

$$\text{approach } D + \underbrace{C}_{\rho' \text{ rows}} (zI_m - A)^{-1} \underbrace{B}_{\rho \text{ columns}} \text{ by } \widetilde{D} + \underbrace{\widetilde{C}}_{\rho' \text{ rows}} (\underbrace{zI_k - \widetilde{A}}_{k \times k,\, k \ll m})^{-1} \underbrace{\widetilde{B}}_{\rho \text{ columns}}$$

on unit circle, here $A = A^*$ with $\sigma(A) \subset (-1, 1)$ (stable system).

Symmetric case $B = C^*$: balanced truncation max-norm error bounded by $2 \sum_{j \geqslant k} \sigma_j(Y)$, with $Y = (B^* A^{j+\ell} B)$ pos. semi-def. Hankel operator. Also,

$$\sigma_j(Y) = \sigma_j(X), \qquad AXA^* - X = -BB^*.$$

Hence in the symmetric case we have that

$$\text{error}_{\rho k} \leqslant Z_k(\sigma(A), \sigma(A^{-*}))\, \text{trace}(Y), \quad \text{trace}(Y) \leqslant ||B||_F^2/(1 - ||A||^2).$$

General MIMO case: extend to symmetric system $(\hat{B} = \hat{C}^* = (B, C^*))$. Hence

$$\text{error}_{(\rho+\rho')k} \leqslant Z_k(\sigma(A), \sigma(A^{-*}))\, (||B||_F^2 + ||C||_F^2)/(1 - ||A||^2).$$

Application Thm 1: linear system theory

Transfer function of MIMO linear time-invariant discrete-time system

$$\text{approach } D + \underbrace{C}_{\rho' \text{ rows}} (zI_m - A)^{-1} \underbrace{B}_{\rho \text{ columns}} \text{ by } \widetilde{D} + \underbrace{\widetilde{C}}_{\rho' \text{ rows}} (\underbrace{zI_k - \widetilde{A}}_{k \times k,\, k \ll m})^{-1} \underbrace{\widetilde{B}}_{\rho \text{ columns}}$$

on unit circle, here $A = A^*$ with $\sigma(A) \subset (-1, 1)$ (stable system).
Symmetric case $B = C^*$: balanced truncation max-norm error bounded by $2 \sum_{j \geqslant k} \sigma_j(Y)$, with $Y = (B^* A^{j+\ell} B)$ pos. semi-def. Hankel operator. Also,

$$\sigma_j(Y) = \sigma_j(X), \qquad AXA^* - X = -BB^*.$$

Hence in the symmetric case we have that

$$\text{error}_{\rho k} \leqslant Z_k(\sigma(A), \sigma(A^{-*}))\, \text{trace}(Y), \quad \text{trace}(Y) \leqslant ||B||_F^2/(1 - ||A||^2).$$

General MIMO case: extend to symmetric system ($\hat{B} = \hat{C}^* = (B, C^*)$). Hence

$$\text{error}_{(\rho+\rho')k} \leqslant Z_k(\sigma(A), \sigma(A^{-*}))\, (||B||_F^2 + ||C||_F^2)/(1 - ||A||^2).$$

Application Thm 1: linear system theory

Transfer function of MIMO linear time-invariant discrete-time system

$$\text{approach } D + \underbrace{C}_{\rho' \text{ rows}} (zI_m - A)^{-1} \underbrace{B}_{\rho \text{ columns}} \text{ by } \widetilde{D} + \underbrace{\widetilde{C}}_{\rho' \text{ rows}} (\underbrace{zI_k - \widetilde{A}}_{k \times k,\, k \ll m})^{-1} \underbrace{\widetilde{B}}_{\rho \text{ columns}}$$

on unit circle, here $A = A^*$ with $\sigma(A) \subset (-1, 1)$ (stable system).
Symmetric case $B = C^*$: balanced truncation max-norm error bounded by $2 \sum_{j \geqslant k} \sigma_j(Y)$, with $Y = (B^* A^{j+\ell} B)$ pos. semi-def. Hankel operator. Also,

$$\sigma_j(Y) = \sigma_j(X), \qquad AXA^* - X = -BB^*.$$

Hence in the symmetric case we have that

$$\text{error}_{\rho k} \leqslant Z_k(\sigma(A), \sigma(A^{-*}))\, \text{trace}(Y), \quad \text{trace}(Y) \leqslant ||B||_F^2 / (1 - ||A||^2).$$

General MIMO case: extend to symmetric system ($\hat{B} = \hat{C}^* = (B, C^*)$). Hence

$$\text{error}_{(\rho+\rho')k} \leqslant Z_k(\sigma(A), \sigma(A^{-*}))\, (||B||_F^2 + ||C||_F^2)/(1 - ||A||^2).$$

Ideas for proof of Thm 1

Following idea of Penzl'00: if $\deg P \leqslant k, \deg Q \leqslant k$,

$$\Delta := Q(A)XP(B) - P(A)XQ(B) \quad \text{is of rank} \leqslant \rho\, k.$$

Thus with $r(z) := P(z)/Q(z)$

$$X - Y = r(A)Xr(B)^{-1}, \qquad Y := Q(A)^{-1}\Delta P(B)^{-1} \quad \text{of rank} \leqslant \rho\, k.$$

Sharpness for $\rho = 1$: $X = D_1CD_2$, C Cauchy, $X^{-t} = X$.

Real intervals: use alternants.

General case: use rational Fekete points.

Ideas for proof of Thm 1

Following idea of Penzl'00: if $\deg P \leqslant k, \deg Q \leqslant k$,

$$\Delta := Q(A)XP(B) - P(A)XQ(B) \quad \text{is of rank} \leqslant \rho\, k.$$

Thus with $r(z) := P(z)/Q(z)$

$$X - Y = r(A)Xr(B)^{-1}, \qquad Y := Q(A)^{-1}\Delta P(B)^{-1} \quad \text{of rank} \leqslant \rho\, k.$$

Sharpness for $\rho = 1$: $X = D_1 C D_2$, C Cauchy, $X^{-t} = X$.

Real intervals: use alternants.

General case: use rational Fekete points.

Theorem 2 (Hankel, Loewner)

Let $A, X \in \mathbb{C}^{m\times m}$, A being normal, and X being hermitian, with signature sign(X) = number str. pos. eigenvals - number str. neg. eigenvals, and

$$AX - XA^* = MN^* - NM^*, \qquad M, N \in \mathbb{C}^{m\times\rho}.$$

We furthermore suppose that (A, M) is reachable, that is, $M, AM, A^2M, \ldots$ span the whole $\mathbb{C}^m$.

Then for any integers $j, k \geqslant 0$ with $j + \rho k < |\mathrm{sign}(X)|$ there holds

$$\frac{\sigma_{j+\mathrm{rank}(X)-|\mathrm{sign}(X)|+\rho k}(X)}{\sigma_j(X)} \leqslant Z_k(\sigma(A), \mathbb{R})^2.$$

(only interesting if $\sigma(A) \cap \mathbb{R} = \emptyset$ and $|\mathrm{sign}(X)| \approx \mathrm{rank}\,(X)$). BB'04

Example Thm 2: Hankel matrices

$$X = \begin{bmatrix} h_0 & h_1 & \cdots & h_{m-1} \\ h_1 & h_2 & \cdots & h_m \\ \vdots & \vdots & & \vdots \\ h_{m-1} & h_m & \cdots & h_{2m-2} \end{bmatrix} \in \mathbb{R}^{m \times m},$$

$$S_m(\Theta)^* X - X S_m(\Theta) = e_\downarrow N^* - N e_\downarrow^*, \qquad e_\downarrow = (0, 0, ..., 0, 1)^*.$$

Thm 2 for Hankel matrices: BB'01 (Marrakesh)

X counteridentity: $\sigma_0(X) = \sigma_{m-1}(X) = 1$, sign$(X) \in \{0, 1\}$.

Quantity $Z_k(\mathbb{R}, S_m(\Theta))$ occurred before for Krylov with hermitian argument....

Sharpness? If positive definite (sign(X) = rank $(X) = m$)

$$\frac{\sigma_{m-1}(X)}{\sigma_0(X)} \leqslant 16m\, 3.210^{1-m} \quad \text{sharp up to } \mathcal{O}(m) \qquad \text{BB'00.}$$

Example Thm 2: Hankel matrices

$$X = \begin{bmatrix} h_0 & h_1 & \cdots & h_{m-1} \\ h_1 & h_2 & \cdots & h_m \\ \vdots & \vdots & & \vdots \\ h_{m-1} & h_m & \cdots & h_{2m-2} \end{bmatrix} \in \mathbb{R}^{m\times m},$$

$$S_m(\Theta)^* X - X S_m(\Theta) = e_\downarrow N^* - N e_\downarrow^*, \qquad e_\downarrow = (0,0,...,0,1)^*.$$

Thm 2 for Hankel matrices: BB'01 (Marrakesh)

X counteridentity: $\sigma_0(X) = \sigma_{m-1}(X) = 1$, sign$(X) \in \{0,1\}$.

Quantity $Z_k(\mathbb{R}, S_m(\Theta))$ occurred before for Krylov with hermitian argument....

Sharpness? If positive definite (sign(X) = rank $(X) = m$)

$$\frac{\sigma_{m-1}(X)}{\sigma_0(X)} \leqslant 16m\,3.210^{1-m} \quad \text{sharp up to } \mathcal{O}(m) \qquad \text{BB'00.}$$

Example Thm 2: Hankel matrices

$$X = \begin{bmatrix} h_0 & h_1 & \cdots & h_{m-1} \\ h_1 & h_2 & \cdots & h_m \\ \vdots & \vdots & & \vdots \\ h_{m-1} & h_m & \cdots & h_{2m-2} \end{bmatrix} \in \mathbb{R}^{m\times m},$$

$$S_m(\Theta)^* X - X S_m(\Theta) = e_\downarrow N^* - N e_\downarrow^*, \qquad e_\downarrow = (0,0,...,0,1)^*.$$

Thm 2 for Hankel matrices: BB'01 (Marrakesh)

X counteridentity: $\sigma_0(X) = \sigma_{m-1}(X) = 1$, $\mathrm{sign}(X) \in \{0,1\}$.

Quantity $Z_k(\mathbb{R}, S_m(\Theta))$ occurred before for Krylov with hermitian argument....

Sharpness? If positive definite ($\mathrm{sign}(X) = \mathrm{rank}\,(X) = m$)

$$\frac{\sigma_{m-1}(X)}{\sigma_0(X)} \leqslant 16m\, 3.210^{1-m} \quad \text{sharp up to } \mathcal{O}(m) \qquad \text{BB'00.}$$

Example Thm 2: Hankel matrices

$$X = \begin{bmatrix} h_0 & h_1 & \cdots & h_{m-1} \\ h_1 & h_2 & \cdots & h_m \\ \vdots & \vdots & & \vdots \\ h_{m-1} & h_m & \cdots & h_{2m-2} \end{bmatrix} \in \mathbb{R}^{m\times m},$$

$$S_m(\Theta)^* X - X S_m(\Theta) = e_\downarrow N^* - N e_\downarrow^*, \qquad e_\downarrow = (0,0,...,0,1)^*.$$

Thm 2 for Hankel matrices: BB'01 (Marrakesh)

X counteridentity: $\sigma_0(X) = \sigma_{m-1}(X) = 1$, sign$(X) \in \{0,1\}$.

Quantity $Z_k(\mathbb{R}, S_m(\Theta))$ occurred before for Krylov with hermitian argument....

Sharpness? If positive definite (sign$(X) =$ rank $(X) = m$)

$$\frac{\sigma_{m-1}(X)}{\sigma_0(X)} \leqslant 16m\, 3.210^{1-m} \quad \text{sharp up to } \mathcal{O}(m) \qquad \text{BB'00.}$$

Example Thm 2: Hankel matrices

$$X = \begin{bmatrix} h_0 & h_1 & \cdots & h_{m-1} \\ h_1 & h_2 & \cdots & h_m \\ \vdots & \vdots & & \vdots \\ h_{m-1} & h_m & \cdots & h_{2m-2} \end{bmatrix} \in \mathbb{R}^{m \times m},$$

$$S_m(\Theta)^* X - X S_m(\Theta) = e_\downarrow N^* - N e_\downarrow^*, \qquad e_\downarrow = (0, 0, ..., 0, 1)^*.$$

Thm 2 for Hankel matrices: BB'01 (Marrakesh)

X counteridentity: $\sigma_0(X) = \sigma_{m-1}(X) = 1$, sign$(X) \in \{0, 1\}$.

Quantity $Z_k(\mathbb{R}, S_m(\Theta))$ occurred before for Krylov with hermitian argument....

Sharpness? If positive definite (sign$(X) =$ rank $(X) = m$)

$$\frac{\sigma_{m-1}(X)}{\sigma_0(X)} \leqslant 16m\, 3.210^{1-m} \quad \text{sharp up to } \mathcal{O}(m) \qquad \text{BB'00}.$$

Example Thm 2: back to Pick matrices

A positive definite Pick matrix X verifies

$$BX + XB^* = N(1, ..., 1) + (1, ..., 1)^* N^*, \qquad N \in \mathbb{C}^m,$$

with s.p.d. B (diagonal with distinct elements). Apply Theorem 2 with $A = iB$, $-A^* = iB$, and thus

$$\frac{\sigma_{m-1}(X)}{\sigma_0(X)} \leqslant Z_{m-1}(\sigma(B), i\mathbb{R})^2 \leqslant Z_{m-1}(\sigma(B), -\sigma(B)) \leqslant \left(\frac{\sqrt{\mathrm{cond}(B)} - 1}{\sqrt{\mathrm{cond}(B)} + 1} \right)^{2m-2}.$$

This is (up to the square root) the Olshevsky & Fasino bound.

Penzl'00 proposes (implicitly) the following upper bound

$$Z_k(\sigma(B), -\sigma(B)) \leqslant \left(\prod_{j=0}^{k-1} \frac{\mathrm{cond}(B)^{(2j+1)/(2k)} - 1}{\mathrm{cond}(B)^{(2j+1)/(2k)} + 1} \right)^2.$$

Example Thm 2: back to Pick matrices

A positive definite Pick matrix X verifies

$$BX + XB^* = N(1, ..., 1) + (1, ..., 1)^* N^*, \qquad N \in \mathbb{C}^m,$$

with s.p.d. B (diagonal with distinct elements). Apply Theorem 2 with $A = iB$, $-A^* = iB$, and thus

$$\frac{\sigma_{m-1}(X)}{\sigma_0(X)} \leqslant Z_{m-1}(\sigma(B), i\mathbb{R})^2 \leqslant Z_{m-1}(\sigma(B), -\sigma(B)) \leqslant \left(\frac{\sqrt{\text{cond}(B)} - 1}{\sqrt{\text{cond}(B)} + 1} \right)^{2m-2}.$$

This is (up to the square root) the Olshevsky & Fasino bound.
Penzl'00 proposes (implicitly) the following upper bound

$$Z_k(\sigma(B), -\sigma(B)) \leqslant \left(\prod_{j=0}^{k-1} \frac{\text{cond}(B)^{(2j+1)/(2k)} - 1}{\text{cond}(B)^{(2j+1)/(2k)} + 1} \right)^2.$$

Ideas for proof of Thm 2 for $\rho = 1$

1.
2.
3.
4.
5.
6.

Ideas for proof of Thm 2 for $\rho = 1$

1. (A, M) reachable $\Longrightarrow K_m(A, M)$ nonsingular, let $L^* = E_\downarrow^* K_m(A, M)^{-1}$.
2.
3.
4.
5.
6.

Ideas for proof of Thm 2 for $\rho = 1$

1. (A, M) reachable $\Longrightarrow K_m(A, M)$ nonsingular, let $L^* = E_\downarrow^* K_m(A, M)^{-1}$.
2. The matrix $K := K(A^*, L)$ is nonsingular, $K^* M = E_\downarrow$.
3.
4.
5.
6.

Ideas for proof of Thm 2 for $\rho = 1$

1. (A, M) reachable $\Longrightarrow K_m(A, M)$ nonsingular, let $L^* = E_\downarrow^* K_m(A, M)^{-1}$.
2. The matrix $K := K(A^*, L)$ is nonsingular, $K^* M = E_\downarrow$.
3. $AX - XA = MN^* - NM^*$, $A^* K - K S_m = N' E_\downarrow^* \Longrightarrow$
4.
5.
6.

Ideas for proof of Thm 2 for $\rho = 1$

1. (A, M) reachable $\Longrightarrow K_m(A, M)$ nonsingular, let $L^* = E_\downarrow^* K_m(A, M)^{-1}$.
2. The matrix $K := K(A^*, L)$ is nonsingular, $K^* M = E_\downarrow$.
3. $AX - XA = MN^* - NM^*$, $A^*K - KS_m = N' E_\downarrow^* \Longrightarrow$
   the matrix $Y := K^* X K$ satisfies $S_m^* Y - Y S_m = E_\downarrow \hat{N}^* - \hat{N} E_\downarrow^*$
4.
5.
6.

Ideas for proof of Thm 2 for $\rho = 1$

1. (A, M) reachable $\Longrightarrow K_m(A, M)$ nonsingular, let $L^* = E_\downarrow^* K_m(A, M)^{-1}$.
2. The matrix $K := K(A^*, L)$ is nonsingular, $K^* M = E_\downarrow$.
3. $AX - XA = MN^* - NM^*$, $A^*K - KS_m = N' E_\downarrow^* \Longrightarrow$
   the matrix $Y := K^* X K$ satisfies $S_m^* Y - Y S_m = E_\downarrow \hat{N}^* - \hat{N} E_\downarrow^*$
   therefore Y is (real) Hankel of signature $\mathrm{sign}(X) \geqslant 0$
4.
5.
6.

Ideas for proof of Thm 2 for $\rho = 1$

1. (A, M) reachable $\Longrightarrow K_m(A, M)$ nonsingular, let $L^* = E_\downarrow^* K_m(A, M)^{-1}$.
2. The matrix $K := K(A^*, L)$ is nonsingular, $K^* M = E_\downarrow$.
3. $AX - XA = MN^* - NM^*$, $A^*K - KS_m = N' E_\downarrow^* \Longrightarrow$
   the matrix $Y := K^* X K$ satisfies $S_m^* Y - Y S_m = E_\downarrow \hat{N}^* - \hat{N} E_\downarrow^*$
   therefore Y is (real) Hankel of signature $\mathrm{sign}(X) \geqslant 0$
4. Factorization result of Fiedler: $Y = V^t D V$,
5.
6.

Ideas for proof of Thm 2 for $\rho = 1$

1. (A, M) reachable $\Longrightarrow K_m(A, M)$ nonsingular, let $L^* = E_\downarrow^* K_m(A, M)^{-1}$.
2. The matrix $K := K(A^*, L)$ is nonsingular, $K^* M = E_\downarrow$.
3. $AX - XA = MN^* - NM^*$, $A^*K - KS_m = N' E_\downarrow^* \Longrightarrow$
   the matrix $Y := K^* X K$ satisfies $S_m^* Y - Y S_m = E_\downarrow \hat{N}^* - \hat{N} E_\downarrow^*$
   therefore Y is (real) Hankel of signature sign$(X) \geqslant 0$
4. Factorization result of Fiedler: $Y = V^t D V$,
   or with $\vec{p} = (p_0, ..., p_{m-1})^t$, $p(z) = \sum p_j z^j$: $\vec{p}^* Y \vec{p} = \sum_{j=1}^m d_j \, p(z_j) \, \overline{p(\overline{z_j})}$.
5.
6.

Ideas for proof of Thm 2 for $\rho = 1$

1. (A, M) reachable $\Longrightarrow K_m(A, M)$ nonsingular, let $L^* = E_\downarrow^* K_m(A, M)^{-1}$.
2. The matrix $K := K(A^*, L)$ is nonsingular, $K^* M = E_\downarrow$.
3. $AX - XA = MN^* - NM^*$, $A^*K - KS_m = N' E_\downarrow^* \Longrightarrow$
   the matrix $Y := K^* X K$ satisfies $S_m^* Y - Y S_m = E_\downarrow \hat{N}^* - \hat{N} E_\downarrow^*$
   therefore Y is (real) Hankel of signature sign$(X) \geqslant 0$
4. Factorization result of Fiedler: $Y = V^t D V$,
   or with $\vec{p} = (p_0, ..., p_{m-1})^t$, $p(z) = \sum p_j z^j$: $\vec{p}^* Y \vec{p} = \sum_{j=1}^m d_j \, p(z_j) \, \overline{p(\overline{z_j})}$.
   non-real z_j occur in conjugate pairs, $d_j \neq 0$,
5.
6.

Ideas for proof of Thm 2 for $\rho = 1$

1. (A, M) reachable $\Longrightarrow K_m(A, M)$ nonsingular, let $L^* = E_\downarrow^* K_m(A, M)^{-1}$.
2. The matrix $K := K(A^*, L)$ is nonsingular, $K^* M = E_\downarrow$.
3. $AX - XA = MN^* - NM^*$, $A^*K - KS_m = N'E_\downarrow^* \Longrightarrow$
   the matrix $Y := K^*XK$ satisfies $S_m^* Y - YS_m = E_\downarrow \hat{N}^* - \hat{N} E_\downarrow^*$
   therefore Y is (real) Hankel of signature $\mathrm{sign}(X) \geqslant 0$
4. Factorization result of Fiedler: $Y = V^t D V$,
   or with $\vec{p} = (p_0, ..., p_{m-1})^t$, $p(z) = \sum p_j z^j$: $\vec{p}^* Y \vec{p} = \sum_{j=1}^m d_j \, p(z_j) \, \overline{p(\overline{z_j})}$.
   non-real z_j occur in conjugate pairs, $d_j \neq 0$,
   at most $(m - \mathrm{sign}(X))/2$ many z_j have imaginary part > 0 or are real and $d_j < 0$.
5.
6.

Ideas for proof of Thm 2 for $\rho = 1$

1. (A, M) reachable $\Longrightarrow K_m(A, M)$ nonsingular, let $L^* = E_\downarrow^* K_m(A, M)^{-1}$.
2. The matrix $K := K(A^*, L)$ is nonsingular, $K^* M = E_\downarrow$.
3. $AX - XA = MN^* - NM^*$, $A^*K - KS_m = N' E_\downarrow^* \Longrightarrow$
   the matrix $Y := K^* X K$ satisfies $S_m^* Y - Y S_m = E_\downarrow \hat{N}^* - \hat{N} E_\downarrow^*$
   therefore Y is (real) Hankel of signature $\mathrm{sign}(X) \geqslant 0$
4. Factorization result of Fiedler: $Y = V^t D V$,
   or with $\vec{p} = (p_0, ..., p_{m-1})^t$, $p(z) = \sum p_j z^j$: $\vec{p}^* Y \vec{p} = \sum_{j=1}^m d_j \, p(z_j) \, \overline{p(\overline{z_j})}$.
   non-real z_j occur in conjugate pairs, $d_j \neq 0$,
   at most $(m - \mathrm{sign}(X))/2$ many z_j have imaginary part > 0 or are real and $d_j < 0$.
5. Also $\vec{p}^* K^* K \vec{p} = \sum_{a \in \sigma(A^*)} |L_a p(a)|^2$.
6.

Ideas for proof of Thm 2 for $\rho = 1$

1. (A, M) reachable $\Longrightarrow K_m(A, M)$ nonsingular, let $L^* = E_\downarrow^* K_m(A, M)^{-1}$.
2. The matrix $K := K(A^*, L)$ is nonsingular, $K^* M = E_\downarrow$.
3. $AX - XA = MN^* - NM^*$, $A^* K - K S_m = N' E_\downarrow^* \Longrightarrow$
   the matrix $Y := K^* X K$ satisfies $S_m^* Y - Y S_m = E_\downarrow \hat{N}^* - \hat{N} E_\downarrow^*$
   therefore Y is (real) Hankel of signature $\text{sign}(X) \geqslant 0$
4. Factorization result of Fiedler: $Y = V^t D V$,
   or with $\vec{p} = (p_0, ..., p_{m-1})^t$, $p(z) = \sum p_j z^j$: $\vec{p}^* Y \vec{p} = \sum_{j=1}^m d_j \, p(z_j) \, \overline{p(\overline{z_j})}$.
   non-real z_j occur in conjugate pairs, $d_j \neq 0$,
   at most $(m - \text{sign}(X))/2$ many z_j have imaginary part > 0 or are real and $d_j < 0$.
5. Also $\vec{p}^* K^* K \vec{p} = \sum_{a \in \sigma(A^*)} |L_a p(a)|^2$.
6. Now play with the Courant minmax principle, and sets of polynomials...

Ideas for proof of Thm 2 for $\rho = 1$

1. (A, M) reachable $\Longrightarrow K_m(A, M)$ nonsingular, let $L^* = E_\downarrow^* K_m(A, M)^{-1}$.
2. The matrix $K := K(A^*, L)$ is nonsingular, $K^* M = E_\downarrow$.
3. $AX - XA = MN^* - NM^*$, $A^*K - KS_m = N' E_\downarrow^* \Longrightarrow$
   the matrix $Y := K^* X K$ satisfies $S_m^* Y - Y S_m = E_\downarrow \hat{N}^* - \hat{N} E_\downarrow^*$
   therefore Y is (real) Hankel of signature $\text{sign}(X) \geqslant 0$
4. Factorization result of Fiedler: $Y = V^t D V$,
   or with $\vec{p} = (p_0, ..., p_{m-1})^t$, $p(z) = \sum p_j z^j$: $\vec{p}^* Y \vec{p} = \sum_{j=1}^m d_j \, p(z_j) \, \overline{p(\overline{z_j})}$.
   non-real z_j occur in conjugate pairs, $d_j \neq 0$,
   at most $(m - \text{sign}(X))/2$ many z_j have imaginary part > 0 or are real and $d_j < 0$.
5. Also $\vec{p}^* K^* K \vec{p} = \sum_{a \in \sigma(A^*)} |L_a p(a)|^2$.
6. Now play with the Courant minmax principle, and sets of polynomials...
   with prescribed zeros, and as factor the numerator/denominator of an Zolotarev-extr

Difficulties in proof of Thm 2 for $\rho > 1$

How to construct an invertible $(M_1, AM_1, ..., M_2, AM_2, ...)$?

Classical: MIMO linear system in controller form with denominator being in Popov normal form

We find $\vec{n} = (\vec{n}_1, ..., \vec{n}_\rho)$ such that $K_{\vec{n}}(A, M) = (K_{\vec{n}_1}(A, M_1), ..., K_{\vec{n}_\rho}(A, M_\rho))$ nonsingular.

$K = K_{\vec{n}}(A^*, L)$, construction of L more complicated.

Now $Y = K^* X K$ is a hermitian $\rho \times \rho$ block matrix, each block of Hankel structure. We need new factorization result! In terms of $K_{\vec{n}}(D, G)$, D diagonal with "many" real entries.

The rest can be generalized.... with vector-valued polynomials.

Extension: discrete Zolotarev problems

Decay of the $m/3$th singular value or the condition number for $\rho > 1$ or ...

$$\text{Find} \quad \Gamma_\tau := \lim_{n,k\to\infty, n/k\to\tau} Z_k(E_n, F_n)^{1/n},$$

given some asymptotic behavior of the sets E_n, F_n (e.g., $E_n = \mathbb{R}$ and F_n rotated nth roots of unity). Tool: logarithmic potential theory.

With Borel measures σ_E, σ_F, logarithmic potential $U^\mu(z) := \int \log(\frac{1}{|x-z|})\, d\mu(x)$

$$\log(\Gamma_\tau) \leqslant \min_{(\mu,\nu)\in\mathcal{M}} \max_{\mathrm{supp}(\sigma_E-\mu)} U^{\nu-\mu} + \max_{\mathrm{supp}(\sigma_F-\nu)} U^{\mu-\nu}$$

$$\text{where} \quad \mathcal{M} = \Big\{(\mu,\nu) : 0 \leqslant \mu \leqslant \sigma_E,\, 0 \leqslant \nu \leqslant \sigma_F,\, ||\mu||, ||\nu|| \leqslant \tau\Big\}.$$

Case Hankel: $\sigma_E = +\infty$ on $\mathbb{R}$, σ_F Lebesgue measure on unit circle, and

$$\log(\Gamma_\tau) = -\int_0^\pi \log\left|\frac{i-e^{is}}{i+e^{is}}\right| w_\psi(s)\, ds, \quad \int_0^{2\pi} w_\psi(s)\, ds = \tau,$$

where $w_\psi(s) = 1/(2\pi)$ if $|\sin(s)| \leqslant \sin(\psi)$, and else $w_\psi(s) = \arcsin(\frac{\sin(\psi)}{|\sin(s)|})/\pi^2$.

Extension: discrete Zolotarev problems

Decay of the $m/3$th singular value or the condition number for $\rho > 1$ or ...

$$\text{Find} \quad \Gamma_\tau := \lim_{n,k\to\infty, n/k\to\tau} Z_k(E_n, F_n)^{1/n},$$

given some asymptotic behavior of the sets E_n, F_n (e.g., $E_n = \mathbb{R}$ and F_n rotated nth roots of unity). Tool: logarithmic potential theory.

With Borel measures σ_E, σ_F, logarithmic potential $U^\mu(z) := \int \log(\frac{1}{|x-z|})\, d\mu(x)$

$$\log(\Gamma_\tau) \leqslant \min_{(\mu,\nu)\in\mathcal{M}} \max_{\mathrm{supp}(\sigma_E-\mu)} U^{\nu-\mu} + \max_{\mathrm{supp}(\sigma_F-\nu)} U^{\mu-\nu}$$

$$\text{where} \quad \mathcal{M} = \Big\{(\mu,\nu) : 0 \leqslant \mu \leqslant \sigma_E,\, 0 \leqslant \nu \leqslant \sigma_F,\, ||\mu||, ||\nu|| \leqslant \tau\Big\}.$$

Case Hankel: $\sigma_E = +\infty$ on $\mathbb{R}$, σ_F Lebesgue measure on unit circle, and

$$\log(\Gamma_\tau) = -\int_0^\pi \log|\frac{i-e^{is}}{i+e^{is}}|\, w_\psi(s)\, ds, \qquad \int_0^{2\pi} w_\psi(s)\, ds = \tau,$$

where $w_\psi(s) = 1/(2\pi)$ if $|\sin(s)| \leqslant \sin(\psi)$, and else $w_\psi(s) = \arcsin(\frac{\sin(\psi)}{|\sin(s)|})/\pi^2$.

Extension: discrete Zolotarev problems

Decay of the $m/3$th singular value or the condition number for $\rho > 1$ or ...

$$\text{Find} \quad \Gamma_\tau := \lim_{n,k\to\infty, n/k\to\tau} Z_k(E_n, F_n)^{1/n},$$

given some asymptotic behavior of the sets E_n, F_n (e.g., $E_n = \mathbb{R}$ and F_n rotated nth roots of unity). Tool: logarithmic potential theory.

With Borel measures σ_E, σ_F, logarithmic potential $U^\mu(z) := \int \log(\frac{1}{|x-z|})\, d\mu(x)$

$$\log(\Gamma_\tau) \leqslant \min_{(\mu,\nu)\in\mathcal{M}} \max_{\mathrm{supp}(\sigma_E-\mu)} U^{\nu-\mu} + \max_{\mathrm{supp}(\sigma_F-\nu)} U^{\mu-\nu}$$

$$\text{where} \quad \mathcal{M} = \Big\{(\mu,\nu) : 0 \leqslant \mu \leqslant \sigma_E,\, 0 \leqslant \nu \leqslant \sigma_F,\, ||\mu||, ||\nu|| \leqslant \tau\Big\}.$$

Case Hankel: $\sigma_E = +\infty$ on $\mathbb{R}$, σ_F Lebesgue measure on unit circle, and

$$\log(\Gamma_\tau) = -\int_0^\pi \log|\frac{i - e^{is}}{i + e^{is}}|\, w_\psi(s)\, ds, \quad \int_0^{2\pi} w_\psi(s)\, ds = \tau,$$

where $w_\psi(s) = 1/(2\pi)$ if $|\sin(s)| \leqslant \sin(\psi)$, and else $w_\psi(s) = \arcsin(\frac{\sin(\psi)}{|\sin(s)|})/\pi^2$.

Conclusion

Conclusion

How bad are matrices X with "small" rank $(AX - XB)$?

Conclusion

How bad are matrices X with "small" rank $(AX - XB)$?

Answer 1: quite ill-conditioned if A, B normal with "separated" spectrum.

Conclusion

How bad are matrices X with "small" rank $(AX - XB)$?

Answer 1: quite ill-conditioned if A, B normal with "separated" spectrum.

Answer 2 (if $B = A^*$ normal and $X = X^*$):
quite ill-conditioned if "high" signature and $\sigma(A)$ "separated" from $\mathbb{R}$

Conclusion

How bad are matrices X with "small" rank $(AX - XB)$?

Answer 1: quite ill-conditioned if A, B normal with "separated" spectrum.

Answer 2 (if $B = A^*$ normal and $X = X^*$):
quite ill-conditioned if "high" signature and $\sigma(A)$ "separated" from $\mathbb{R}$

More precise answers? Solve (approximately) the corresponding Zolotarev problem.

Conclusion

How bad are matrices X with "small" rank $(AX - XB)$?

Answer 1: quite ill-conditioned if A, B normal with "separated" spectrum.

Answer 2 (if $B = A^*$ normal and $X = X^*$):
quite ill-conditioned if "high" signature and $\sigma(A)$ "separated" from $\mathbb{R}$

More precise answers? Solve (approximately) the corresponding Zolotarev problem.

And if A, B are no longer normal but diagonalizable?

Conclusion

How bad are matrices X with "small" $\text{rank}\,(AX - XB)$?

Answer 1: quite ill-conditioned if A, B normal with "separated" spectrum.

Answer 2 (if $B = A^*$ normal and $X = X^*$):
quite ill-conditioned if "high" signature and $\sigma(A)$ "separated" from $\mathbb{R}$

More precise answers? Solve (approximately) the corresponding Zolotarev problem.

And if A, B are no longer normal but diagonalizable?

Multiply right-hand side with the condition numbers of both eigenvector matrices!

Conclusion

How bad are matrices X with "small" rank $(AX - XB)$?

Answer 1: quite ill-conditioned if A, B normal with "separated" spectrum.

Answer 2 (if $B = A^*$ normal and $X = X^*$):
quite ill-conditioned if "high" signature and $\sigma(A)$ "separated" from $\mathbb{R}$

More precise answers? Solve (approximately) the corresponding Zolotarev problem.

And if A, B are no longer normal but diagonalizable?

Multiply right-hand side with the condition numbers of both eigenvector matrices!

And structured singular values?
In our approch, the row rank approximation has a displacement rank twice as large...