Max-min and min-max approximation problems for normal matrices revisited

Petr Tichý
Czech Academy of Sciences,
University of West Bohemia

joint work with

Jörg Liesen
TU Berlin

January 30, SNA 2014,
Nymburk, Czech Republic

Bounding GMRES residual norm

$\mathbf{A}x = b$, $\mathbf{A} \in \mathbb{C}^{n\times n}$ is nonsingular, $b \in \mathbb{C}^n$,

$x_0 = \mathbf{0}$ and $\|b\| = 1$ for simplicity.

GMRES computes $x_k \in \mathcal{K}_k(\mathbf{A}, b)$ such that $r_k \equiv b - \mathbf{A}x_k$ satisfies

$$\begin{aligned} \|r_k\| &= \min_{p\in\pi_k} \|p(\mathbf{A})b\| && \text{(GMRES)} \\ &\le \max_{\|b\|=1} \min_{p\in\pi_k} \|p(\mathbf{A})b\| && \text{(worst-case GMRES)} \\ &\le \min_{p\in\pi_k} \|p(\mathbf{A})\| && \text{(ideal GMRES)} \end{aligned}$$

where π_k = degree $\le k$ polynomials with $p(0) = 1$.

Two bounds on the GMRES residual norm

$$\max_{\|b\|=1} \min_{p \in \pi_k} \|p(\mathbf{A})b\| \;\leq\; \min_{p \in \pi_k} \|p(\mathbf{A})\|$$

- They are **equal** if $\mathbf{A}$ is **normal**.

  [Greenbaum, Gurvits '94; Joubert '94].

- The inequality can be **strict** if $\mathbf{A}$ is **non-normal**.

  [Toh '97; Faber, Joubert, Knill, Manteuffel '96].

How to prove the equality for normal matrices?

If $\mathbf{A}$ is normal, then

$$\max_{\|b\|=1} \min_{p \in \pi_k} \|p(\mathbf{A})b\| \;=\; \min_{p \in \pi_k} \|p(\mathbf{A})\| .$$

- [Joubert '94] Proof using analytic methods of **optimization theory**, for real or complex data, only in the GMRES context.
- [Greenbaum, Gurvits '94]: Proof based mostly on **matrix theory**, only for real data but in a more general form.
- These proofs are quite **complicated**.
- Is there a straightforward proof that uses, e.g., known classical results of **approximation theory**?

Outline

1. Normal matrices and classical approximation problems
2. Best polynomial approximation for f on Γ
3. Proof
4. Connection to results by Greenbaum and Gurvits

Link to classical approximation problems

- $\mathbf{A}$ is normal iff $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^*$, $\mathbf{Q}^*\mathbf{Q} = \mathbf{I}$.
- $\Gamma \equiv \{\lambda_1, \ldots, \lambda_n\}$ is the set of eigenvalues of $\mathbf{A}$.
- For any function g defined on Γ denote

$$\|g\|_\Gamma \equiv \max_{z\in\Gamma} |g(z)|.$$

- $p \in \pi_k$ means

$$p(z) = 1 - \sum_{i=1}^{k} \alpha_i\, z^i\,.$$

- Then

$$\begin{aligned}\min_{p\in\pi_k} \|p(\mathbf{A})\| &= \min_{p\in\pi_k} \|\mathbf{Q}p(\mathbf{\Lambda})\mathbf{Q}^*\| \;=\; \min_{p\in\pi_k} \max_{\lambda_i} |p(\lambda_i)| \\ &= \min_{\alpha_1,\ldots,\alpha_k} \left\| 1 - \sum_{i=1}^{k} \alpha_i\, z^i \right\|_\Gamma .\end{aligned}$$

Generalization

- Instead of 1 we consider a general **function** f defined on Γ. Instead of $\{z^i\}_{i=1}^k$ we consider general **basis functions** φ_i.

  We ask whether

  $$\max_{\|b\|=1} \min_{p \in \mathcal{P}_k} \|f(\mathbf{A})b - p(\mathbf{A})b\| \; = \; \min_{p \in \mathcal{P}_k} \|f(\mathbf{A}) - p(\mathbf{A})\|$$

  where $\mathbf{A}$ is normal and p is of the form

  $$p(z) = \sum_{i=1}^{k} \alpha_i \, \varphi_i(z) \quad \in \quad \mathcal{P}_k \, .$$

- A comment on $\mathbb{R}$ versus $\mathbb{C}$ $\rightarrow$ coefficients α_i.

- As in the previous

  $$\min_{p \in \mathcal{P}_k} \|f(\mathbf{A}) - p(\mathbf{A})\| \quad = \quad \min_{p \in \mathcal{P}_k} \|f(z) - p(z)\|_{\Gamma} \, .$$

A polynomial of best approximation for f on Γ

Definition and notation

$p_* \in \mathcal{P}_k$ is a **polynomial of best approximation** for f on Γ when

$$\|f - p_*\|_\Gamma \;=\; \min_{p \in \mathcal{P}_k} \|f - p\|_\Gamma.$$

For $p \in \mathcal{P}_k$, define

$$\Gamma(p) \;\equiv\; \{z \in \Gamma :\; |f(z) - p(z)| = \|f - p\|_\Gamma\}.$$

Characterization of best approximation for f on Γ

[Chebyshev, Berstein, de la Vallée Poussing, Haar, Remez, Zuhovickiĭ, Kolmogorov]

[Rivlin, Shapiro '61], [Lorentz '86]

Characterization theorem (complex case)

$p_* \in \mathcal{P}_k$ is **a polynomial of best approximation** for f on Γ

if and only if

there exist ℓ points $\mu_i \in \Gamma(p_*)$ where $1 \leq \ell \leq 2k+1$, and
ℓ real numbers $\omega_1, \ldots, \omega_\ell > 0$ with $\omega_1 + \cdots + \omega_\ell = 1$, such that

$$\sum_{j=1}^{\ell} \omega_j \, \overline{p(\mu_j)} \, [f(\mu_j) - p_*(\mu_j)] = 0, \quad \forall \, p \in \mathcal{P}_k.$$

Denote

$$\delta \equiv \|f - p_*\|_\Gamma = |f(\mu_j) - p_*(\mu_j)|, \qquad j = 1, \ldots, \ell\,.$$

Proof I

It suffices to prove that

$$\max_{\|b\|=1} \min_{p\in\mathcal{P}_k} \|f(\mathbf{A})b - p(\mathbf{A})b\| \quad \geq \quad \min_{p\in\mathcal{P}_k} \|f(\mathbf{A}) - p(\mathbf{A})\|$$
$$= \quad \min_{p\in\mathcal{P}_k} \|f(z) - p(z)\|_\Gamma \,.$$

Suppose that the eigenvalues of $\mathbf{A}$ are sorted such that

$$\lambda_j = \mu_j, \qquad j = 1, \ldots, \ell.$$

Define the vector w

$$w = \mathbf{Q}\,\xi, \quad \xi \equiv \left[\sqrt{\omega_1}, \ldots, \sqrt{\omega_\ell}, 0, \ldots, 0\right]^T \,.$$

Then

$$\begin{aligned} 0 &= \sum_{j=1}^{\ell} \omega_j \,\overline{p(\mu_j)}\,[f(\mu_j) - p_*(\mu_j)] \\ &= \xi^H p(\mathbf{\Lambda})^H [f(\mathbf{\Lambda}) - p_*(\mathbf{\Lambda})]\,\xi \\ &= w^H p(\mathbf{A})^H [f(\mathbf{A}) - p_*(\mathbf{A})]\,w \,. \end{aligned}$$

Proof II

In other words,

$$f(\mathbf{A})b - p_*(\mathbf{A})w \perp p(\mathbf{A})w\,, \quad \forall\, p \in \mathcal{P}_k\,,$$

or, equivalently,

$$\|f(\mathbf{A})w - p_*(\mathbf{A})w\| \;=\; \min_{p\in\mathcal{P}_k} \|f(\mathbf{A})w - p(\mathbf{A})w\|\,.$$

Moreover

$$\begin{aligned}
\|f(\mathbf{A})w - p_*(\mathbf{A})w\|^2 &= \|\,[f(\mathbf{\Lambda}) - p_*(\mathbf{\Lambda})]\,\xi\|^2 \\
&= \sum_{j=1}^{\ell} \xi_j^2\,|f(\mu_j) - p_*(\mu_j)|^2 \\
&= \sum_{j=1}^{\ell} \omega_j \delta^2 \;=\; \delta^2 \\
&= \|f(\mathbf{A}) - p_*(\mathbf{A})\|^2\,.
\end{aligned}$$

Proof III

In summary, **for** $p_* \in \mathcal{P}_k$ we have constructed $w \in \mathbb{C}^n$ such that

$$
\begin{aligned}
\min_{p \in \mathcal{P}_k} \|f(\mathbf{A}) - p(\mathbf{A})\| &= \|f(\mathbf{A}) - p_*(\mathbf{A})\| \\
&= \|f(\mathbf{A})w - p_*(\mathbf{A})w\|^2 \\
&= \min_{p \in \mathcal{P}_k} \|f(\mathbf{A})w - p(\mathbf{A})w\| \\
&\leq \max_{\|b\|=1} \min_{p \in \mathcal{P}_k} \|f(\mathbf{A})b - p(\mathbf{A})b\| .
\end{aligned}
$$

The proof for **complex** $\mathbf{A}$ is finished.

A note on the real case

- Assume that $\mathbf{A}$, $f(\mathbf{A})$ and $\varphi_i(\mathbf{A})$ are **real**. We look for a polynomial of a best approximation with **real coefficients**.

- **Technical problem**: $\mathbf{A}$ can have **complex eigenvalues** but we look for a **real vector** b that maximizes

$$\min_{p \in \mathcal{P}_k} \|f(\mathbf{A})b - p(\mathbf{A})b\| .$$

- Γ is a set of points that appear in **complex conjugate pairs**.

- This **symmetry** with respect to the real axes has been used to find a real b and to prove the equality [Liesen, T. 2013].

Results by Greenbaum and Gurvits, Horn and Johnson

Theorem [Greenbaum, Gurvits '94]

Let $\mathbf{A}_0, \mathbf{A}_1, \ldots, \mathbf{A}_k$ be **normal matrices** that **commute**. Then

$$\max_{\|v\|=1} \min_{\alpha_1,\ldots,\alpha_k} \|\mathbf{A}_0 v - \sum_{i=1}^{k} \alpha_i \mathbf{A}_i v\| \;=\; \min_{\alpha_1,\ldots,\alpha_k} \|\mathbf{A}_0 - \sum_{i=1}^{k} \alpha_i \mathbf{A}_i\|.$$

Theorem [Theorem 2.5.5, Horn, Johnson '90]

Commuting normal matrices **can be simultaneously unitarily diagonalized**, i.e., there exists a unitary $\mathbf{U}$ so that

$$\mathbf{U}^H \mathbf{A}_i \mathbf{U} \;=\; \mathbf{\Lambda}_i\,, \quad i = 0, 1, \ldots, k.$$

Connection to results by Greenbaum and Gurvits

Using the theorem by Horn and Johnson we can equivalently rewrite the problem

$$\min_{\alpha_1,\ldots,\alpha_k} \|\mathbf{A}_0 - \sum_{i=1}^{k} \alpha_i \, \mathbf{A}_i\|$$

in our notation

$$\min_{\alpha_1,\ldots,\alpha_k} \|f(\mathbf{A}) - \sum_{i=1}^{k} \alpha_i \, \varphi_i(\mathbf{A})\|$$

where $\mathbf{A}$ is any diagonal matrix with distinct eigenvalues and f and φ_i are any functions satisfying

$$f(\mathbf{A}) = \mathbf{\Lambda}_0, \qquad \varphi_i(\mathbf{A}) = \mathbf{\Lambda}_i, \qquad i = 1, \ldots, k.$$

Summary

- Inspired by the convergence analysis of GMRES **we formulated** two general approximation problems involving normal matrices.

- We used a **direct link** between
  - approximation problems involving normal matrices,
  - classical approximation problems

  and **proved** that

$$\max_{\|b\|=1} \min_{p \in \mathcal{P}_k} \|f(\mathbf{A})b - p(\mathbf{A})b\| = \min_{p \in \mathcal{P}_k} \|f(\mathbf{A}) - p(\mathbf{A})\|\,.$$

- Our results
  - represent a **generalization** of results by [Joubert '94],
  - offer **another point of view** to [Greenbaum, Gurvits '94].

Related papers

- J. Liesen and P. Tichý, [Max-min and min-max approximation problems for normal matrices revisited, submitted to ETNA (2013).]
- A. Greenbaum and L. Gurvits, [Max-min properties of matrix factor norms, SISC, 15 (1994), pp. 348–358.]
- W. Joubert, [A robust GMRES-based adaptive polynomial preconditioning algorithm for nonsymmetric linear systems, SISC, 15 (1994), pp. 427–439.]
- M. Bellalij, Y. Saad, and H. Sadok, [Analysis of some Krylov subspace methods for normal matrices via approximation theory and convex optimization, ETNA, 33 (2008/09), pp. 17–30.]

Thank you for your attention!