On a New Proof of the Faber-Manteuffel Theorem

Petr Tichý

joint work with

Jörg Liesen and Vance Faber

Institute of Computer Science,
Academy of Sciences of the Czech Republic

April 28, 2008, Novi Sad, Serbia

Outline

1. Introduction
2. Formulation of the problem
3. The Faber-Manteuffel theorem
4. The ideas of a new proof

Outline

Krylov subspace methods

Given $\mathbf{A} \in \mathbb{R}^{n \times n}$, $v \in \mathbb{R}^n$. Define the jth Krylov subspace

$$\mathcal{K}_j(\mathbf{A}, v) \equiv \operatorname{span}(v, \mathbf{A}v, \ldots, \mathbf{A}^{j-1}v)\,.$$

Krylov subspace methods

Krylov subspace methods

Given $\mathbf{A} \in \mathbb{R}^{n\times n}$, $v \in \mathbb{R}^n$. Define the jth Krylov subspace

$$\mathcal{K}_j(\mathbf{A}, v) \equiv \mathrm{span}(v, \mathbf{A}v, \ldots, \mathbf{A}^{j-1}v)\,.$$

Krylov subspace methods:

- Iterative methods for solving large and sparse linear systems or eigenvalue problems,
- they are based on projection onto the Krylov subspaces,
- examples: Lanczos, CG, Arnoldi, GMRES, BiCG.

Krylov subspace methods
Basis

Each method must generate a basis of $\mathcal{K}_j(\mathbf{A}, v), \quad j = 1, 2, \ldots$

- The trivial choice $v, \mathbf{A}v, \ldots, \mathbf{A}^{j-1}v$ is computationally infeasible (recall the Power Method).

Krylov subspace methods

Basis

Each method must generate a basis of $\mathcal{K}_j(\mathbf{A}, v), \quad j = 1, 2, \dots$

- The trivial choice $v, \mathbf{A}v, \dots, \mathbf{A}^{j-1}v$ is computationally infeasible (recall the Power Method).
- For numerical stability: Well conditioned basis.
- For computational efficiency: Short recurrence.

Each method must generate a basis of $\mathcal{K}_j(\mathbf{A}, v), \quad j = 1, 2, \ldots$

- The trivial choice $v, \mathbf{A}v, \ldots, \mathbf{A}^{j-1}v$ is computationally infeasible (recall the Power Method).
- For numerical stability: Well conditioned basis.
- For computational efficiency: Short recurrence.
- Best of both worlds:
  Orthogonal basis computed by short recurrence.

Optimal Krylov subspace methods
with short recurrences

CG (1952), MINRES, SYMMLQ (1975)

- based on three-term recurrences

$$r_{j+1} = \gamma_j \mathbf{A} r_j - \alpha_j r_j - \beta_j r_{j-1}\,,$$

Optimal Krylov subspace methods

with short recurrences

CG (1952), MINRES, SYMMLQ (1975)

- based on three-term recurrences

$$r_{j+1} = \gamma_j \mathbf{A} r_j - \alpha_j r_j - \beta_j r_{j-1} \,,$$

- generate orthogonal (or $\mathbf{A}$-orthogonal) Krylov subspace basis,

Optimal Krylov subspace methods

with short recurrences

CG (1952), MINRES, SYMMLQ (1975)

- based on three-term recurrences

$$r_{j+1} = \gamma_j \mathbf{A} r_j - \alpha_j r_j - \beta_j r_{j-1}\,,$$

- generate orthogonal (or $\mathbf{A}$-orthogonal) Krylov subspace basis,

- *optimal* in the sense that they minimize some error norm:

  $\|x - x_j\|_{\mathbf{A}}$ in CG,
  $\|x - x_j\|_{\mathbf{A}^T\mathbf{A}} = \|r_j\|$ in MINRES,
  $\|x - x_j\|$ in SYMMLQ -here $x_j \in x_0 + \mathbf{A}\mathcal{K}_j(\mathbf{A}, r_0)$.

Optimal Krylov subspace methods

with short recurrences

CG (1952), MINRES, SYMMLQ (1975)

- based on three-term recurrences

$$r_{j+1} = \gamma_j \mathbf{A} r_j - \alpha_j r_j - \beta_j r_{j-1} \,,$$

- generate orthogonal (or $\mathbf{A}$-orthogonal) Krylov subspace basis,

- *optimal* in the sense that they minimize some error norm:

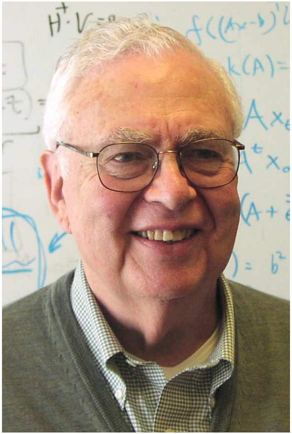
  $\|x - x_j\|_{\mathbf{A}}$ in CG,

  $\|x - x_j\|_{\mathbf{A}^T\mathbf{A}} = \|r_j\|$ in MINRES,

  $\|x - x_j\|$ in SYMMLQ -here $x_j \in x_0 + \mathbf{A}\mathcal{K}_j(\mathbf{A}, r_0)$.

- An important assumption on $\mathbf{A}$:
  $\mathbf{A}$ is symmetric (MINRES, SYMMLQ) & pos. definite (CG).

Gene Golub



G. H. Golub, 1932–2007

- By the end of the 1970s it was unknown if such methods existed also for general unsymmetric $\mathbf{A}$.
- Gatlinburg VIII (now Householder VIII) held in Oxford from July 5 to 11, 1981.
- "A prize of \$500 has been offered by Gene Golub for the construction of a 3-term conjugate gradient like descent method for non-symmetric real matrices or a proof that there can be no such method".

What kind of method Golub had in mind

- We want to solve $\mathbf{A}x = b$ using CG-like descent method: error is minimized in some given inner product norm, $\| \cdot \|_{\mathbf{B}} = \langle \cdot, \cdot \rangle_{\mathbf{B}}^{1/2}$.

What kind of method Golub had in mind

- We want to solve $\mathbf{A}x = b$ using CG-like descent method: error is minimized in some given inner product norm, $\|\cdot\|_{\mathbf{B}} = \langle \cdot, \cdot \rangle_{\mathbf{B}}^{1/2}$.
- Starting from x_0, compute

$$x_{j+1} = x_j + \alpha_j p_j \,, \qquad j = 0, 1, \dots \,,$$

p_j is a direction vector, α_j is a scalar (to be determined),

$$\mathrm{span}\{p_0, \dots, p_j\} = \mathcal{K}_{j+1}(\mathbf{A}, r_0), \qquad r_0 = b - \mathbf{A}x_0 \,.$$

What kind of method Golub had in mind

- We want to solve $\mathbf{A}x = b$ using CG-like descent method: error is minimized in some given inner product norm, $\| \cdot \|_{\mathbf{B}} = \langle \cdot, \cdot \rangle_{\mathbf{B}}^{1/2}$.
- Starting from x_0, compute

$$x_{j+1} = x_j + \alpha_j p_j \,, \qquad j = 0, 1, \dots \,,$$

  p_j is a direction vector, α_j is a scalar (to be determined),

$$\mathrm{span}\{p_0, \dots, p_j\} = \mathcal{K}_{j+1}(\mathbf{A}, r_0), \qquad r_0 = b - \mathbf{A}x_0 \,.$$

- $\|x - x_{j+1}\|_{\mathbf{B}}$ is minimal iff

$$\alpha_j = \frac{\langle x - x_j, p_j \rangle_{\mathbf{B}}}{\langle p_j, p_j \rangle_{\mathbf{B}}} \quad \text{and} \quad \langle p_j, p_i \rangle_{\mathbf{B}} = 0 \,.$$

What kind of method Golub had in mind

- We want to solve $\mathbf{A}x = b$ using CG-like descent method: error is minimized in some given inner product norm, $\| \cdot \|_{\mathbf{B}} = \langle \cdot, \cdot \rangle_{\mathbf{B}}^{1/2}$.
- Starting from x_0, compute

$$x_{j+1} = x_j + \alpha_j p_j \,, \qquad j = 0, 1, \dots \,,$$

p_j is a direction vector, α_j is a scalar (to be determined),

$$\mathrm{span}\{p_0, \dots, p_j\} = \mathcal{K}_{j+1}(\mathbf{A}, r_0), \qquad r_0 = b - \mathbf{A}x_0 \,.$$

- $\|x - x_{j+1}\|_{\mathbf{B}}$ is minimal iff

$$\alpha_j = \frac{\langle x - x_j, p_j \rangle_{\mathbf{B}}}{\langle p_j, p_j \rangle_{\mathbf{B}}} \quad \text{and} \quad \langle p_j, p_i \rangle_{\mathbf{B}} = 0 \,.$$

- $p_0, \dots, p_j$ has to be a $\mathbf{B}$-orthogonal basis of $\mathcal{K}_{j+1}(\mathbf{A}, r_0)$.

Faber and Manteuffel, 1984

SIAM J. NUMER. ANAL.
Vol. 21, No. 2, April 1984

011

NECESSARY AND SUFFICIENT CONDITIONS FOR THE EXISTENCE OF A CONJUGATE GRADIENT METHOD*

VANCE FABER† AND THOMAS MANTEUFFEL†

Abstract. We characterize the class $CG(s)$ of matrices A for which the linear system $A\mathbf{x}=\mathbf{b}$ can be solved by an s-term conjugate gradient method. We show that, except for a few anomalies, the class $CG(s)$ consists of matrices A for which conjugate gradient methods are already known. These matrices are the Hermitian matrices, $A^*=A$, and the matrices of the form $A=e^{i\theta}(dI+B)$, with $B^*=-B$.

- Faber and Manteuffel gave the answer in 1984:
  For a general matrix $\mathbf{A}$ there exists *no* short recurrence for generating orthogonal Krylov subspace bases.
- **What are the details of this statement**?

Outline

Without loss of generality, $\mathbf{B} = \mathbf{I}$. Otherwise change the basis:

$$\langle x, y\rangle_{\mathbf{B}} = \langle \mathbf{B}^{1/2}x, \mathbf{B}^{1/2}y\rangle, \quad \hat{\mathbf{A}} \equiv \mathbf{B}^{1/2}\mathbf{A}\mathbf{B}^{-1/2}, \quad \hat{v} \equiv \mathbf{B}^{1/2}v\,.$$

Formulation of the problem

$\mathbf{B}$-inner product, Input and Notation

Without loss of generality, $\mathbf{B} = \mathbf{I}$. Otherwise change the basis:

$$\langle x, y \rangle_{\mathbf{B}} = \langle \mathbf{B}^{1/2} x, \mathbf{B}^{1/2} y \rangle, \quad \hat{\mathbf{A}} \equiv \mathbf{B}^{1/2} \mathbf{A} \mathbf{B}^{-1/2}, \quad \hat{v} \equiv \mathbf{B}^{1/2} v\,.$$

Input data:

- $\mathbf{A} \in \mathbb{C}^{n \times n}$, a nonsingular matrix.
- $v \in \mathbb{C}^n$, an initial vector.

Formulation of the problem

$\mathbf{B}$-inner product, Input and Notation

Without loss of generality, $\mathbf{B} = \mathbf{I}$. Otherwise change the basis:

$$\langle x, y \rangle_{\mathbf{B}} = \langle \mathbf{B}^{1/2} x, \mathbf{B}^{1/2} y \rangle, \quad \hat{\mathbf{A}} \equiv \mathbf{B}^{1/2} \mathbf{A} \mathbf{B}^{-1/2}, \quad \hat{v} \equiv \mathbf{B}^{1/2} v\,.$$

Input data:

- $\mathbf{A} \in \mathbb{C}^{n \times n}$, a nonsingular matrix.
- $v \in \mathbb{C}^{n}$, an initial vector.

Notation:

- $d_{\min}(\mathbf{A})$ ... the degree of the minimal polynomial of $\mathbf{A}$.
- $d = d(\mathbf{A}, v)$ ... the grade of v with respect to $\mathbf{A}$,
  the smallest d s.t. $\mathcal{K}_d(\mathbf{A}, v)$ is invariant under mult. with $\mathbf{A}$.

- Generate a basis $v_1, \ldots, v_d$ of $\mathcal{K}_d(\mathbf{A}, v)$ s.t.
  1. $\operatorname{span}\{v_1, \ldots, v_j\} = \mathcal{K}_j(A, v),$ for $j = 1, \ldots, d$,
  2. $\langle v_i, v_j \rangle = 0,$ for $i \neq j$, $i, j = 1, \ldots, d$.

- Generate a basis $v_1, \dots, v_d$ of $\mathcal{K}_d(\mathbf{A}, v)$ s.t.
  1. $\mathrm{span}\{v_1, \dots, v_j\} = \mathcal{K}_j(A, v),$ for $j = 1, \dots, d$,
  2. $\langle v_i, v_j \rangle = 0,$ for $i \neq j, \quad i, j = 1, \dots, d$.

Arnoldi's method:
Standard way for generating the orthogonal basis
(no normalization for convenience): $v_1 \equiv v$,

$$v_{j+1} = \mathbf{A} v_j - \sum_{i=1}^{j} h_{i,j}\, v_i\,, \qquad h_{i,j} = \frac{\langle \mathbf{A} v_j, v_i \rangle}{\langle v_i, v_i \rangle}\,,$$

$j = 0, \dots, d-1$.

In matrix notation:

$$
\begin{aligned}
v_1 &= v\,, \\
\mathbf{A}\,\underbrace{[v_1,\ldots,v_{d-1}]}_{\equiv\,\mathbf{V}_{d-1}} &= \underbrace{[v_1,\ldots,v_d]}_{\equiv\,\mathbf{V}_d}\underbrace{\begin{bmatrix} h_{1,1} & \cdots & h_{1,d-1} \\ 1 & \ddots & \vdots \\ & \ddots & h_{d-1,d-1} \\ & & \\ & & 1 \end{bmatrix}}_{\equiv\,\mathbf{H}_{d,d-1}}\,,
\end{aligned}
$$

$$\mathbf{V}_d^*\mathbf{V}_d \ \text{ is diagonal}\,, \quad d = \dim \mathcal{K}_n(\mathbf{A}, v)\,.$$

Formulation of the problem

Optimal short recurrences (Definition - Liesen and Strakoš, 2008)

$\mathbf{A}$ admits an optimal $(s+2)$-term recurrence, if

- for any v, $\mathbf{H}_{d,d-1}$ is at most $(s+2)$-band Hessenberg, and
- for at least one v, $\mathbf{H}_{d,d-1}$ is $(s+2)$-band Hessenberg.

$$\mathbf{A}\,\mathbf{V}_{d-1} \;=\; \mathbf{V}_d \;\underbrace{\left[\begin{array}{ccccc} \overbrace{\bullet \quad \cdots \quad \bullet}^{s+1} & & & & \\ \bullet \quad \ddots & & & \ddots & \\ \ddots & \ddots & & & \bullet \\ & \ddots & & \ddots & \vdots \\ & & & \ddots & \bullet \\ & & & & \bullet \end{array}\right]}_{d-1}$$

What are sufficient and necessary conditions for $\mathbf{A}$ to admit an optimal $(s+2)$-term recurrence?

What are sufficient and necessary conditions for $\mathbf{A}$ to admit an optimal $(s+2)$-term recurrence?

In other words, how can we characterize matrices $\mathbf{A}$ such that for any v, Arnoldi's method applied to $\mathbf{A}$ and v generates an orthogonal basis via a short recurrence of length $s+2$.

Formulation of the problem

Basic question

What are sufficient and necessary conditions for $\mathbf{A}$ to admit an optimal $(s+2)$-term recurrence?

In other words, how can we characterize matrices $\mathbf{A}$ such that for any v, Arnoldi's method applied to $\mathbf{A}$ and v generates an orthogonal basis via a short recurrence of length $s+2$.

Example of sufficiency: If $\mathbf{A}^* = \mathbf{A}$, then $s = 1$ and $\mathbf{A}$ admits an optimal 3-term recurrence.

Formulation of the problem

Basic question

What are sufficient and necessary conditions for $\mathbf{A}$ to admit an optimal $(s+2)$-term recurrence?

In other words, how can we characterize matrices $\mathbf{A}$ such that for any v, Arnoldi's method applied to $\mathbf{A}$ and v generates an orthogonal basis via a short recurrence of length $s+2$.

Example of sufficiency: If $\mathbf{A}^* = \mathbf{A}$, then $s=1$ and $\mathbf{A}$ admits an optimal 3-term recurrence.

Definition. If

$$\mathbf{A}^* = p_s(\mathbf{A}),$$

where p_s is a polynomial of the smallest possible degree s, $\mathbf{A}$ is called normal(s).

Outline

The Faber-Manteuffel theorem

Theorem. [Faber and Manteuffel, 1984], [Liesen and Strakoš, 2008]

Let $\mathbf{A}$ be a nonsingular matrix with minimal polynomial degree $d_{\min}(\mathbf{A})$. Let s be a nonnegative integer, $s + 2 < d_{\min}(\mathbf{A})$:

$\mathbf{A}$ admits an optimal $(s+2)$-term recurrence

if and only if

$\mathbf{A}$ is normal(s).

The Faber-Manteuffel theorem

Theorem. [Faber and Manteuffel, 1984], [Liesen and Strakoš, 2008]

Let $\mathbf{A}$ be a nonsingular matrix with minimal polynomial degree $d_{\min}(\mathbf{A})$. Let s be a nonnegative integer, $s + 2 < d_{\min}(\mathbf{A})$:

$\mathbf{A}$ admits an optimal $(s+2)$-term recurrence

if and only if

$\mathbf{A}$ is normal(s).

- Sufficiency is rather straightforward, necessity *is not*. Key words from the proof of necessity in (Faber and Manteuffel, 1984) include: "continuous function" (analysis), "closed set of smaller dimension" (topology), "wedge product" (multilinear algebra).

The Faber-Manteuffel theorem

Why is necessity so hard?

Optimal $(s+2)$-term recurrence:

$$\mathbf{A}\,\mathbf{V}_{d-1} \quad = \quad \mathbf{V}_d \underbrace{\begin{array}{c} \overbrace{\hphantom{\bullet \quad \cdots \quad \bullet}}^{s+1} \hphantom{\quad\quad\quad\quad\quad\quad} \\ \begin{bmatrix} \bullet & \cdots & \bullet & & \\ \bullet & \ddots & & \ddots & \\ & \ddots & \ddots & & \bullet \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \bullet \\ & & & & \bullet \end{bmatrix} \end{array}}_{d-1}$$

Prove something about the linear operator $\mathbf{A}$, without complete knowledge of the structure of its matrix representation.

The Faber-Manteuffel theorem

Why is necessity so hard?

Since $\mathcal{K}_d(\mathbf{A}, v)$ is invariant, $\mathbf{A}v_d \in \mathcal{K}_d(\mathbf{A}, v)$ and

$$\mathbf{A}{\color{green}v_d} = \sum_{i=1}^{d} {\color{blue}h_{id}}\,{\color{red}v_i}.$$

$$\mathbf{A}\,\mathbf{V}_d \quad = \quad \mathbf{V}_d \begin{array}{l} \quad\overbrace{\hphantom{{\color{blue}\bullet} \quad \cdots \quad {\color{blue}\bullet}}}^{s+1} \\ \left[\begin{array}{cccccc} {\color{blue}\bullet} & \cdots & {\color{blue}\bullet} & & & {\color{red}\bullet} \\ {\color{blue}\bullet} & \ddots & & \ddots & & \vdots \\ & \ddots & \ddots & & {\color{blue}\bullet} & {\color{red}\bullet} \\ & & \ddots & \ddots & \vdots & \vdots \\ & & & \ddots & {\color{blue}\bullet} & {\color{red}\bullet} \\ & & & & {\color{blue}\bullet} & {\color{red}\bullet} \end{array}\right] \\ \quad\underbrace{\hphantom{{\color{blue}\bullet} \quad \ddots \quad \ddots \quad \ddots \quad {\color{blue}\bullet}}}_{d-1} \end{array}$$

Outline

- Motivated by the paper [J. Liesen and Z. Strakoš, 2008] which contains a completely reworked theory of short recurrences for generating orthogonal Krylov subspace bases.

  "It is unknown if a simpler proof of the necessity part can be found. In view of the fundamental nature of the Faber-Manteuffel Theorem, such proof would be a welcome addition to the existing literature. It would lead to a better understanding of the theorem by enlightening some (possibly unexpected) relationships, and it would also be more suitable for classroom teaching."

V. Faber, J. Liesen and P. Tichý., 2008

The Faber-Manteuffel Theorem for Linear Operators

- Motivated by the paper [J. Liesen and Z. Strakoš, 2008] which contains a completely reworked theory of short recurrences for generating orthogonal Krylov subspace bases.

  "It is unknown if a simpler proof of the necessity part can be found. In view of the fundamental nature of the Faber-Manteuffel Theorem, such proof would be a welcome addition to the existing literature. It would lead to a better understanding of the theorem by enlightening some (possibly unexpected) relationships, and it would also be more suitable for classroom teaching."

- We give two new proofs of the Faber-Manteuffel theorem that use more elementary tools,
- first proof - improved version of the Faber-Manteuffel proof,
- second proof - completely new proof based on orthogonal transformations of upper Hessenberg matrices.

Idea of the second proof (1)

V. Faber, J. Liesen and P. Tichý, 2008

(for simplicity, we omit indices by $\mathbf{V}_d$ and $\mathbf{H}_{d,d}$)

Let $\mathbf{A}$ admit an optimal $(s+2)$-term recurrence

$$\mathbf{A}\,\mathbf{V} \;=\; \mathbf{V}\,\mathbf{H}, \quad \mathbf{V}^*\mathbf{V} = \mathbf{I}\,.$$

Up to the last column, $\mathbf{H}$ is $(s+2)$-band Hessenberg.

Idea of the second proof (1)

V. Faber, J. Liesen and P. Tichý, 2008

(for simplicity, we omit indices by $\mathbf{V}_d$ and $\mathbf{H}_{d,d}$)

Let $\mathbf{A}$ admit an optimal $(s+2)$-term recurrence

$$\mathbf{A}\,\mathbf{V} \;=\; \mathbf{V}\,\mathbf{H}, \quad \mathbf{V}^*\mathbf{V} = \mathbf{I}\,.$$

Up to the last column, $\mathbf{H}$ is $(s+2)$-band Hessenberg.
Let $\mathbf{G}$ be a $d \times d$ unitary matrix, $\mathbf{G}^*\mathbf{G} = \mathbf{I}$. Then

$$\mathbf{A}\,\underbrace{(\mathbf{VG})}_{\mathbf{W}} \;=\; \underbrace{(\mathbf{VG})}_{\mathbf{W}}\,\underbrace{(\mathbf{G}^*\mathbf{HG})}_{\widetilde{\mathbf{H}}}\,.$$

$\mathbf{W}$ is unitary.

Idea of the second proof (1)

V. Faber, J. Liesen and P. Tichý, 2008

(for simplicity, we omit indices by $\mathbf{V}_d$ and $\mathbf{H}_{d,d}$)

Let $\mathbf{A}$ admit an optimal $(s+2)$-term recurrence

$$\mathbf{A}\,\mathbf{V} \;=\; \mathbf{V}\,\mathbf{H}, \quad \mathbf{V}^*\mathbf{V} = \mathbf{I}\,.$$

Up to the last column, $\mathbf{H}$ is $(s+2)$-band Hessenberg.
Let $\mathbf{G}$ be a $d \times d$ unitary matrix, $\mathbf{G}^*\mathbf{G} = \mathbf{I}$. Then

$$\mathbf{A}\,\underbrace{(\mathbf{VG})}_{\mathbf{W}} \;=\; \underbrace{(\mathbf{VG})}_{\mathbf{W}}\,\underbrace{(\mathbf{G}^*\mathbf{HG})}_{\widetilde{\mathbf{H}}}\,.$$

$\mathbf{W}$ is unitary. If $\mathbf{G}$ is chosen such that $\widetilde{\mathbf{H}}$ is again unreduced upper Hessenberg matrix, then

$$\mathbf{A}\,\mathbf{W} \;=\; \mathbf{W}\,\tilde{\mathbf{H}}\,.$$

represents the result of Arnoldi's method applied to $\mathbf{A}$ and w_1.
Up to the last column, $\widetilde{\mathbf{H}}$ has to be $(s+2)$-band Hessenberg.

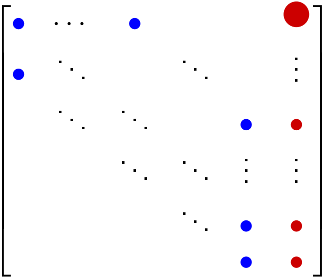
Idea of the second proof (2)

V. Faber, J. Liesen and P. Tichý, 2008

Proof by contradiction. Let $\mathbf{A}$ admit an optimal $(s+2)$-term recurrence and $\mathbf{A}$ not be normal(s).

Then there exists a starting vector v such that $h_{1,d} \neq 0$.

$$\mathbf{A}\ \mathbf{V} \quad = \quad \mathbf{V}$$



Idea of the second proof (2)

V. Faber, J. Liesen and P. Tichý, 2008

Proof by contradiction. Let $\mathbf{A}$ admit an optimal $(s+2)$-term recurrence and $\mathbf{A}$ not be normal(s).

Then there exists a starting vector v such that $h_{1,d} \neq 0$.

$$\mathbf{A}\,(\mathbf{VG}) \;=\; (\mathbf{VG})\,\mathbf{G}^{*} \begin{bmatrix} \bullet & \cdots & \bullet & & & \bullet \\ \bullet & \ddots & & \ddots & & \vdots \\ & \ddots & \ddots & & \bullet & \bullet \\ & & \ddots & \ddots & \vdots & \vdots \\ & & & \ddots & \bullet & \bullet \\ & & & & \bullet & \bullet \end{bmatrix} \mathbf{G}$$

Idea of the second proof (2)

V. Faber, J. Liesen and P. Tichý, 2008

Proof by contradiction. Let $\mathbf{A}$ admit an optimal $(s+2)$-term recurrence and $\mathbf{A}$ not be normal(s).

Then there exists a starting vector v such that $h_{1,d} \neq 0$.

$$\mathbf{A}\,(\mathbf{V}\mathbf{G}) \;=\; (\mathbf{V}\mathbf{G})\,\mathbf{G}^* \begin{bmatrix} \bullet & \cdots & \bullet & & & \bullet \\ \bullet & \ddots & & \ddots & & \vdots \\ & \ddots & \ddots & & \bullet & \bullet \\ & & \ddots & \ddots & \vdots & \vdots \\ & & & \ddots & \bullet & \bullet \\ & & & & \bullet & \bullet \end{bmatrix} \mathbf{G}$$

Find unitary $\mathbf{G}$ (a product of Givens rotations) such that $\widetilde{\mathbf{H}}$ is unreduced upper Hessenberg, but $\widetilde{\mathbf{H}}$ is not $(s+2)$-band (up to the last column) - **contradiction**.

Idea of the second proof (3)

V. Faber, J. Liesen and P. Tichý, 2008

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\mathbf{G}_{7,8}$.

$$\begin{bmatrix} \bullet & \bullet & \bullet & 0 & 0 & 0 & 0 & \bullet \\ \bullet & \bullet & \bullet & \bullet & 0 & 0 & 0 & \bullet \\ & \bullet & \bullet & \bullet & \bullet & 0 & 0 & \bullet \\ & & \bullet & \bullet & \bullet & \bullet & 0 & \bullet \\ & & & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & & \bullet & \bullet & \bullet & \bullet \\ & & & & & \bullet & \bullet & \bullet \\ & & & & & & \bullet & \bullet \end{bmatrix} \mathbf{G}_{7,8}$$

Idea of the second proof (3)

V. Faber, J. Liesen and P. Tichý, 2008

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\mathbf{G}_{7,8}$.

$$\mathbf{G}_{7,8}^{*} \begin{bmatrix} \bullet & \bullet & \bullet & 0 & 0 & 0 & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & 0 & 0 & \bullet & \bullet \\ & \bullet & \bullet & \bullet & \bullet & 0 & \bullet & \bullet \\ & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & & \bullet & \bullet & \bullet & \bullet \\ & & & & & \bullet & \bullet & \bullet \\ & & & & & & \bullet & \bullet \end{bmatrix}$$

Idea of the second proof (3)

V. Faber, J. Liesen and P. Tichý, 2008

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\color{blue}{\mathbf{G}_{7,8}}$.

$$
\begin{bmatrix}
{\color{blue}\bullet} & {\color{blue}\bullet} & {\color{blue}\bullet} & 0 & 0 & 0 & {\color{green}\bullet} & {\color{green}\bullet} \\
{\color{blue}\bullet} & {\color{blue}\bullet} & {\color{blue}\bullet} & {\color{blue}\bullet} & 0 & 0 & {\color{green}\bullet} & {\color{green}\bullet} \\
 & {\color{blue}\bullet} & {\color{blue}\bullet} & {\color{blue}\bullet} & {\color{blue}\bullet} & 0 & {\color{green}\bullet} & {\color{green}\bullet} \\
 & & {\color{blue}\bullet} & {\color{blue}\bullet} & {\color{blue}\bullet} & {\color{blue}\bullet} & {\color{green}\bullet} & {\color{green}\bullet} \\
 & & & {\color{blue}\bullet} & {\color{blue}\bullet} & {\color{blue}\bullet} & {\color{green}\bullet} & {\color{green}\bullet} \\
 & & & & {\color{blue}\bullet} & {\color{blue}\bullet} & {\color{green}\bullet} & {\color{green}\bullet} \\
 & & & & & {\color{green}\bullet} & {\color{green}\bullet} & {\color{green}\bullet} \\
 & & & & & {\color{red}\bullet} & {\color{green}\bullet} & {\color{green}\bullet}
\end{bmatrix}
\mathbf{G}_{6,7}
$$

Idea of the second proof (3)

V. Faber, J. Liesen and P. Tichý, 2008

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\mathbf{G}_{7,8}$.

$$\mathbf{G}_{6,7}^{*} \begin{bmatrix} \bullet & \bullet & \bullet & 0 & 0 & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & 0 & \bullet & \bullet & \bullet \\ & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & & \bullet & \bullet & \bullet & \bullet \\ & & & & & \bullet & \bullet & \bullet \\ & & & & & & \bullet & \bullet \end{bmatrix}$$

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\mathbf{G}_{7,8}$.

$$\begin{bmatrix} \bullet & \bullet & \bullet & 0 & 0 & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & 0 & \bullet & \bullet & \bullet \\ & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & & \bullet & \bullet & \bullet & \bullet \\ & & & & \bullet & \bullet & \bullet & \bullet \\ & & & & & & \bullet & \bullet \end{bmatrix} \mathbf{G}_{5,6}$$

Idea of the second proof (3)

V. Faber, J. Liesen and P. Tichý, 2008

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\color{blue}{\mathbf{G}_{7,8}}$.

$$
\mathbf{G}_{5,6}^{*}
\begin{bmatrix}
\color{blue}{\bullet} & \color{blue}{\bullet} & \color{blue}{\bullet} & 0 & \color{green}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} \\
\color{blue}{\bullet} & \color{blue}{\bullet} & \color{blue}{\bullet} & \color{blue}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} \\
 & \color{blue}{\bullet} & \color{blue}{\bullet} & \color{blue}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} \\
 & & \color{blue}{\bullet} & \color{blue}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} \\
 & & & \color{blue}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} \\
 & & & & \color{green}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} \\
 & & & & & \color{green}{\bullet} & \color{green}{\bullet} & \color{green}{\bullet} \\
 & & & & & & \color{green}{\bullet} & \color{green}{\bullet}
\end{bmatrix}
$$

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\color{blue}{\mathbf{G}_{7,8}}$.

$$
\begin{bmatrix}
\color{blue}\bullet & \color{blue}\bullet & \color{blue}\bullet & 0 & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\
\color{blue}\bullet & \color{blue}\bullet & \color{blue}\bullet & \color{blue}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\
 & \color{blue}\bullet & \color{blue}\bullet & \color{blue}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\
 & & \color{blue}\bullet & \color{blue}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\
 & & & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\
 & & & \color{red}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\
 & & & & & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\
 & & & & & & \color{green}\bullet & \color{green}\bullet
\end{bmatrix}
\mathbf{G}_{4,5}
$$

Idea of the second proof (3)

V. Faber, J. Liesen and P. Tichý, 2008

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\mathbf{G}_{7,8}$.

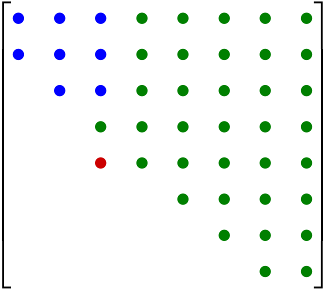
$$\mathbf{G}_{4,5}^{*} \begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & & \bullet & \bullet & \bullet & \bullet \\ & & & & & \bullet & \bullet & \bullet \\ & & & & & & \bullet & \bullet \end{bmatrix}$$

Idea of the second proof (3)

V. Faber, J. Liesen and P. Tichý, 2008

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\mathbf{G}_{7,8}$.



$\mathbf{G}_{3,4}$

Idea of the second proof (3)

V. Faber, J. Liesen and P. Tichý, 2008

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\color{blue}{\mathbf{G}_{7,8}}$.

$$\mathbf{G}_{3,4}^* \begin{bmatrix} \color{blue}\bullet & \color{blue}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ \color{blue}\bullet & \color{blue}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ & \color{blue}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ & & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ & & & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ & & & & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ & & & & & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ & & & & & & \color{green}\bullet & \color{green}\bullet \end{bmatrix}$$

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\color{blue}{\mathbf{G}_{7,8}}$.

$$\begin{bmatrix} \color{blue}\bullet & \color{blue}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ \color{blue}\bullet & \color{blue}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ & \color{red}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ & & & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ & & & & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ & & & & & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ & & & & & & \color{green}\bullet & \color{green}\bullet \end{bmatrix} \mathbf{G}_{2,3}$$

Idea of the second proof (3)

V. Faber, J. Liesen and P. Tichý, 2008

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\mathbf{G}_{7,8}$.

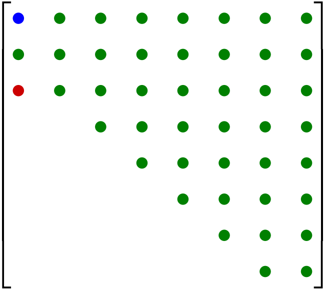
$$\mathbf{G}_{2,3}^{*} \begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & & \bullet & \bullet & \bullet & \bullet \\ & & & & & \bullet & \bullet & \bullet \\ & & & & & & \bullet & \bullet \end{bmatrix}$$

Idea of the second proof (3)

V. Faber, J. Liesen and P. Tichý, 2008

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\mathbf{G}_{7,8}$.



$\mathbf{G}_{1,2}$

Idea of the second proof (3)

V. Faber, J. Liesen and P. Tichý, 2008

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\mathbf{G}_{7,8}$.

$$\mathbf{G}_{1,2}^{*} \begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & & \bullet & \bullet & \bullet & \bullet \\ & & & & & \bullet & \bullet & \bullet \\ & & & & & & \bullet & \bullet \end{bmatrix}$$

Idea of the second proof (3)

V. Faber, J. Liesen and P. Tichý, 2008

Let v be a starting vector such that $h_{1,8} \neq 0$.

Choose Givens rotation $\mathbf{G}_{7,8}$.

$$\begin{bmatrix}
\bullet & \bullet & \bullet & \circ & \circ & \circ & \circ & \bullet \\
\bullet & \bullet & \bullet & \bullet & \circ & \circ & \circ & \bullet \\
 & \bullet & \bullet & \bullet & \bullet & \circ & \circ & \bullet \\
 & & \bullet & \bullet & \bullet & \bullet & \circ & \bullet \\
 & & & \bullet & \bullet & \bullet & \bullet & \bullet \\
 & & & & \bullet & \bullet & \bullet & \bullet \\
 & & & & & \bullet & \bullet & \bullet \\
 & & & & & & \bullet & \bullet
\end{bmatrix}$$

$$\mathbf{G} \equiv \mathbf{G}_{7,8}\mathbf{G}_{6,7} \dots \mathbf{G}_{1,2}, \qquad \widetilde{\mathbf{H}} \equiv \mathbf{G}^* \mathbf{H} \, \mathbf{G}.$$

We proved: It is possible to choose $\mathbf{G}_{7,8}$ such that

$$h_{1,8} \neq 0 \quad \Longrightarrow \quad \tilde{h}_{1,7} \neq 0 \ \text{ or } \ \tilde{h}_{2,7} \neq 0\,.$$

Summary

Generating of orthogonal basis of $\mathcal{K}_d(\mathbf{A}, v)$ via short recurrences

$$\boxed{\begin{array}{l}\text{Arnoldi-type recurrence}\\ (s+2)\text{-term}\end{array}}$$

$$\Updownarrow$$

$$\boxed{\begin{array}{l}\mathbf{A}\text{ is normal(s)}\\ \mathbf{A}^* = p(\mathbf{A})\end{array}}$$

- When is $\mathbf{A}$ normal(s)?

Summary

Generating of orthogonal basis of $\mathcal{K}_d(\mathbf{A}, v)$ via short recurrences

$\boxed{\begin{array}{l}\text{Arnoldi-type recurrence} \\ (s+2)\text{-term}\end{array}}$

$\Updownarrow$

$\boxed{\begin{array}{l}\mathbf{A} \text{ is normal(s)} \\ \mathbf{A}^* = p(\mathbf{A})\end{array}}$

- When is $\mathbf{A}$ normal(s)?
- $\mathbf{A}$ is normal and
  [Faber and Manteuffel, 1984],
  [Khavinson and Świątek, 2003]
  [Liesen and Strakoš, 2008]
  1. $s = 1$ if and only if the eigenvalues of $\mathbf{A}$ lie on a line in $\mathbb{C}$.
  2. If the eigenvalues of $\mathbf{A}$ are *not* on a line, then $d_{\min}(\mathbf{A}) \leq 3s - 2$.

Summary

Generating of orthogonal basis of $\mathcal{K}_d(\mathbf{A}, v)$ via short recurrences

$\boxed{\text{Arnoldi-type recurrence } (s+2)\text{-term}}$

$\Updownarrow$

$\boxed{\mathbf{A} \text{ is normal(s) } \mathbf{A}^* = p(\mathbf{A})}$

$\updownarrow$

$\boxed{\text{the only interesting case is } s = 1, \text{ collinear eigenvalues}}$

- When is $\mathbf{A}$ normal(s)?
- $\mathbf{A}$ is normal and [Faber and Manteuffel, 1984], [Khavinson and Świątek, 2003] [Liesen and Strakoš, 2008]
  1. $s = 1$ if and only if the eigenvalues of $\mathbf{A}$ lie on a line in $\mathbb{C}$.
  2. If the eigenvalues of $\mathbf{A}$ are *not* on a line, then $d_{\min}(\mathbf{A}) \leq 3s - 2$.

Summary

Generating of orthogonal basis of $\mathcal{K}_d(\mathbf{A}, v)$ via short recurrences

$\boxed{\begin{array}{l}\text{Arnoldi-type recurrence} \\ (s+2)\text{-term}\end{array}}$

$$\Updownarrow$$

$\boxed{\begin{array}{l}\mathbf{A} \text{ is normal(s)} \\ \mathbf{A}^* = p(\mathbf{A})\end{array}}$

$$\updownarrow$$

$\boxed{\begin{array}{l}\text{the only interesting case} \\ \text{is } s = 1, \\ \text{collinear eigenvalues}\end{array}}$

- When is $\mathbf{A}$ normal(s)?
- $\mathbf{A}$ is normal and
  [Faber and Manteuffel, 1984],
  [Khavinson and Świątek, 2003]
  [Liesen and Strakoš, 2008]
  1. $s = 1$ if and only if the eigenvalues of $\mathbf{A}$ lie on a line in $\mathbb{C}$.
  2. If the eigenvalues of $\mathbf{A}$ are *not* on a line, then $d_{\min}(\mathbf{A}) \leq 3s - 2$.
- All classes of "interesting" matrices are known.

Related papers

- J. Liesen and Z. Strakoš, [On optimal short recurrences for generating orthogonal Krylov subspace bases, to appear in SIAM Review, 2008].
  Completely reworked theory of short recurrences for generating orthogonal Krylov subspace bases

- V. Faber, J. Liesen and P. Tichý, [The Faber-Manteuffel Theorem for Linear Operators, SIAM J. Numer. Anal., 2008, 46, 1323-1337].
  New proofs of the fundamental theorem of Faber and Manteuffel

More details can be found at

http://www.cs.cas.cz/tichy
http://www.math.tu-berlin.de/~liesen
http://www.cs.cas.cz/strakos

Related papers

- J. Liesen and Z. Strakoš, [On optimal short recurrences for generating orthogonal Krylov subspace bases, to appear in SIAM Review, 2008].
  Completely reworked theory of short recurrences for generating orthogonal Krylov subspace bases

- V. Faber, J. Liesen and P. Tichý, [The Faber-Manteuffel Theorem for Linear Operators, SIAM J. Numer. Anal., 2008, 46, 1323-1337].
  New proofs of the fundamental theorem of Faber and Manteuffel

More details can be found at

`http://www.cs.cas.cz/tichy`
`http://www.math.tu-berlin.de/~liesen`
`http://www.cs.cas.cz/strakos`

Thank you for your attention!