Nearest Pencils with Specified Eigenvalues

Emre Mengi

Department of Mathemtics
Koç University
Istanbul, Turkey

Householder Symposium XVIII
June 14, 2011

joint with D. Kressner (EPFL-Lausanne), I. Nakic (Univ of Zagreb) and N. Truhar (Univ of Osijek)

Outline

Kronecker Canonical Form (KCF)

Given a pencil $(A - \lambda B)$ where $A, B \in \mathbb{C}^{m \times n}$ with $m \geq n$.

- There exist invertible matrices $P \in \mathbb{C}^{m \times m}$ and $Q \in \mathbb{C}^{n \times n}$ such that

$$P(A - \lambda B)Q = \operatorname{diag}\left(A_1 - \lambda B_1, A_2 - \lambda B_2, \ldots, A_k - \lambda B_k\right)$$

Each $A_j - \lambda B_j$ is in one of the following forms.

(i) Jordan Block (associated with a finite eigenvalue α)

$$A_j - \lambda B_j = \begin{bmatrix} \lambda - \alpha & 1 & 0 & \ldots & 0 \\ 0 & \lambda - \alpha & 1 & \ldots & 0 \\ & & & & \\ & & & \lambda - \alpha & 1 \\ 0 & & & 0 & \lambda - \alpha \end{bmatrix}$$

Kronecker Canonical Form (KCF)

Given a pencil $(A - \lambda B)$ where $A, B \in \mathbb{C}^{m \times n}$ with $m \geq n$.

- There exist invertible matrices $P \in \mathbb{C}^{m \times m}$ and $Q \in \mathbb{C}^{n \times n}$ such that

$$P(A - \lambda B)Q = \operatorname{diag}\left(A_1 - \lambda B_1, A_2 - \lambda B_2, \ldots, A_k - \lambda B_k\right)$$

Each $A_j - \lambda B_j$ is in one of the following forms.

(i) Jordan Block (associated with a finite eigenvalue α)

$$A_j - \lambda B_j = \begin{bmatrix} \lambda - \alpha & 1 & 0 & \ldots & 0 \\ 0 & \lambda - \alpha & 1 & \ldots & 0 \\ & & & & \\ & & & \lambda - \alpha & 1 \\ 0 & & & 0 & \lambda - \alpha \end{bmatrix}$$

Kronecker Canonical Form (KCF)

Given a pencil $(A - \lambda B)$ where $A, B \in \mathbb{C}^{m \times n}$ with $m \geq n$.

- There exist invertible matrices $P \in \mathbb{C}^{m \times m}$ and $Q \in \mathbb{C}^{n \times n}$ such that

$$P(A - \lambda B)Q = \mathrm{diag}\,(A_1 - \lambda B_1, A_2 - \lambda B_2, \ldots, A_k - \lambda B_k)$$

Each $A_j - \lambda B_j$ is in one of the following forms.

(i) Jordan Block (associated with a finite eigenvalue α)

$$A_j - \lambda B_j = \begin{bmatrix} \lambda - \alpha & 1 & 0 & \ldots & 0 \\ 0 & \lambda - \alpha & 1 & \ldots & 0 \\ & & & & \\ & & & \lambda - \alpha & 1 \\ 0 & & & 0 & \lambda - \alpha \end{bmatrix}$$

Kronecker Canonical Form (KCF)

Given a pencil $(A - \lambda B)$ where $A, B \in \mathbb{C}^{m \times n}$ with $m \geq n$.

- There exist invertible matrices $P \in \mathbb{C}^{m \times m}$ and $Q \in \mathbb{C}^{n \times n}$ such that

$$P(A - \lambda B)Q = \mathrm{diag}\left(A_1 - \lambda B_1, A_2 - \lambda B_2, \ldots, A_k - \lambda B_k\right)$$

Each $A_j - \lambda B_j$ is in one of the following forms.

(i) Jordan Block (associated with a finite eigenvalue α)

$$A_j - \lambda B_j = \begin{bmatrix} \lambda - \alpha & 1 & 0 & \ldots & 0 \\ 0 & \lambda - \alpha & 1 & \ldots & 0 \\ & & & & \\ & & & \lambda - \alpha & 1 \\ 0 & & & 0 & \lambda - \alpha \end{bmatrix}$$

Kronecker Canonical Form (KCF)

(ii) Block associated with an infinite eigenvalue

$$A_j - \lambda B_j = \begin{bmatrix} 1 & \lambda & 0 & \dots & 0 \\ 0 & 1 & \lambda & \dots & 0 \\ & & & & \\ & & & 1 & \lambda \\ 0 & & & 0 & 1 \end{bmatrix}$$

(iii) Singular blocks of the form (of size $(n_j + 1) \times n_j$)

$$A_j - \lambda B_j = \begin{bmatrix} \lambda & 0 & & & 0 \\ 1 & \lambda & & & 0 \\ 0 & 1 & \lambda & & \\ & & & & \\ & & & 1 & \lambda \\ 0 & & & 0 & 1 \end{bmatrix}$$

Kronecker Canonical Form (KCF)

(ii) Block associated with an infinite eigenvalue

$$A_j - \lambda B_j = \begin{bmatrix} 1 & \lambda & 0 & \dots & 0 \\ 0 & 1 & \lambda & \dots & 0 \\ & & & & \\ & & & 1 & \lambda \\ 0 & & & 0 & 1 \end{bmatrix}$$

(iii) Singular blocks of the form (of size $(n_j + 1) \times n_j$)

$$A_j - \lambda B_j = \begin{bmatrix} \lambda & 0 & & & 0 \\ 1 & \lambda & & & 0 \\ 0 & 1 & \lambda & & \\ & & & & \\ & & & 1 & \lambda \\ 0 & & & 0 & 1 \end{bmatrix}$$

Outline

1. Introduction
   - Kronecker Canonical Form
   - Problem Definition
   - Motivation
2. Derivation of a Singular Value Characterization
   - Rank Characterization
   - Construction of an Optimal Perturbation
   - Simultaneous Perturbations
3. Numerical Examples

Problem Definition

- Algebraic multiplicity of an eigenvalue α: Sum of the sizes of the Jordan blocks associated with α in the KCF.
- $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ be given scalars, and $\mathcal{S} = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$
- r be a given positive integer
- $m_j(A, B)$: Algebraic multip of λ_j as an eigenvalue of $A - \lambda B$

Definition (Distance to Pencils with Specified Eigenvalues)

$$\tau_r(A, B, \mathcal{S}) = \inf \left\{ \left\| \begin{bmatrix} \delta A & \delta B \end{bmatrix} \right\|_2 : \sum_{j=1}^{k} m_j(A + \alpha_A \delta A, B + \alpha_B \delta B) \geq r \right\}$$

$$\alpha_A = 1, \alpha_B = 0 \quad \text{or} \quad \alpha_A = \alpha_B = 1$$

Problem Definition

- Algebraic multiplicity of an eigenvalue α: Sum of the sizes of the Jordan blocks associated with α in the KCF.
- $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ be given scalars, and $\mathcal{S} = \{\lambda_1, \lambda_2, \ldots, \lambda_k\}$
- r be a given positive integer
- $m_j(A, B)$: Algebraic multip of λ_j as an eigenvalue of $A - \lambda B$

Definition (Distance to Pencils with Specified Eigenvalues)

$$\tau_r(A, B, \mathcal{S}) = \inf \left\{ \left\| \begin{bmatrix} \delta A & \delta B \end{bmatrix} \right\|_2 : \sum_{j=1}^{k} m_j(A + \alpha_A \delta A, B + \alpha_B \delta B) \geq r \right\}$$

$$\alpha_A = 1, \alpha_B = 0 \quad \text{or} \quad \alpha_A = \alpha_B = 1$$

Problem Definition

- Algebraic multiplicity of an eigenvalue α: Sum of the sizes of the Jordan blocks associated with α in the KCF.
- $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ be given scalars, and $\mathcal{S} = \{\lambda_1, \lambda_2, \ldots, \lambda_k\}$
- r be a given positive integer
- $m_j(A, B)$: Algebraic multip of λ_j as an eigenvalue of $A - \lambda B$

Definition (Distance to Pencils with Specified Eigenvalues)

$$\tau_r(A, B, \mathcal{S}) = \inf \left\{ \left\| \begin{bmatrix} \delta A & \delta B \end{bmatrix} \right\|_2 : \sum_{j=1}^{k} m_j(A + \alpha_A \delta A, B + \alpha_B \delta B) \geq r \right\}$$

$$\alpha_A = 1, \alpha_B = 0 \quad \text{or} \quad \alpha_A = \alpha_B = 1$$

Problem Definition

- Algebraic multiplicity of an eigenvalue α: Sum of the sizes of the Jordan blocks associated with α in the KCF.
- $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ be given scalars, and $\mathcal{S} = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$
- r be a given positive integer
- $m_j(A, B)$: Algebraic multip of λ_j as an eigenvalue of $A - \lambda B$

Definition (Distance to Pencils with Specified Eigenvalues)

$$\tau_r(A, B, \mathcal{S}) = \inf \left\{ \left\| \left[\ \delta A \quad \delta B \ \right] \right\|_2 : \sum_{j=1}^{k} m_j(A + \alpha_A \delta A, B + \alpha_B \delta B) \geq r \right\}$$

$$\alpha_A = 1, \alpha_B = 0 \quad \text{or} \quad \alpha_A = \alpha_B = 1$$

Problem Definition

- Algebraic multiplicity of an eigenvalue α: Sum of the sizes of the Jordan blocks associated with α in the KCF.
- $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ be given scalars, and $\mathcal{S} = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$
- r be a given positive integer
- $m_j(A, B)$: Algebraic multip of λ_j as an eigenvalue of $A - \lambda B$

Definition (Distance to Pencils with Specified Eigenvalues)

$$\tau_r(A, B, \mathcal{S}) = \inf \left\{ \left\| \begin{bmatrix} \delta A & \delta B \end{bmatrix} \right\|_2 : \sum_{j=1}^{k} m_j(A + \alpha_A \delta A, B + \alpha_B \delta B) \geq r \right\}$$

$\alpha_A = 1, \alpha_B = 0$ or $\alpha_A = \alpha_B = 1$

Problem Definition

- Algebraic multiplicity of an eigenvalue α: Sum of the sizes of the Jordan blocks associated with α in the KCF.
- $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ be given scalars, and $\mathcal{S} = \{\lambda_1, \lambda_2, \ldots, \lambda_k\}$
- r be a given positive integer
- $m_j(A, B)$: Algebraic multip of λ_j as an eigenvalue of $A - \lambda B$

Definition (Distance to Pencils with Specified Eigenvalues)

$$\tau_r(A, B, \mathcal{S}) = \inf \left\{ \left\| \begin{bmatrix} \delta A & \delta B \end{bmatrix} \right\|_2 : \sum_{j=1}^{k} m_j(A + \alpha_A \delta A, B + \alpha_B \delta B) \geq r \right\}$$

$$\alpha_A = 1, \alpha_B = 0 \quad \text{or} \quad \alpha_A = \alpha_B = 1$$

Outline

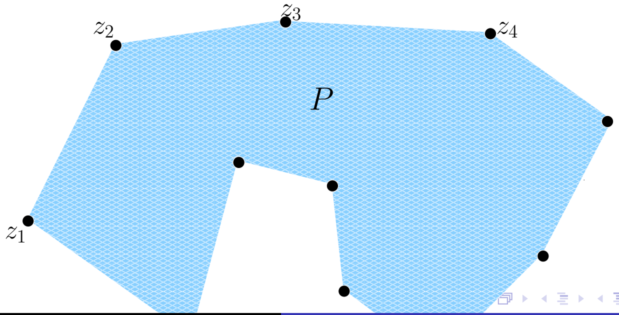
1. Introduction
   - Kronecker Canonical Form
   - Problem Definition
   - Motivation
2. Derivation of a Singular Value Characterization
   - Rank Characterization
   - Construction of an Optimal Perturbation
   - Simultaneous Perturbations
3. Numerical Examples

Shape Estimation from Moments

- The distance to the nearest matrix with a multiple eigenvalue is a special case.

$$\inf_{\lambda \in \mathbb{C}} \tau_2(A, I, \lambda)$$

- Estimating a polygon from moments (Elad, Milanfar, Golub; 2004)



Shape Estimation from Moments

- Given moments
$$\mathcal{M}_k = \int \int_P z^k \, dx \, dy$$
for $k = 1, \ldots, m$.

  Estimate the vertices $z_j \in \mathbb{C}$ for $j = 1, \ldots, n$ of P.

- The vertices z_j are the eigenvalues of a pencil $T_0 - \lambda T_1$ where $T_0, T_1 \in \mathbb{C}^{m \times n}$ (with $m \geq n$) are Hankel matrices defined in terms of $\mathcal{M}_k$.

- Because of measurement errors the perturbed pencil $\widetilde{T}_0 - \lambda \widetilde{T}_1$ has generically no eigenvalues.

- Find a nearest pencil with a full set of eigenvalues
$$\inf_{\mathcal{S} \in \mathbb{C}^n} \tau_n(\widetilde{T}_0, \widetilde{T}_1, \mathcal{S}).$$

Shape Estimation from Moments

- Given moments
$$\mathcal{M}_k = \int \int_P z^k \, dx \, dy$$
for $k = 1, \ldots, m$.

  Estimate the vertices $z_j \in \mathbb{C}$ for $j = 1, \ldots, n$ of P.

- The vertices z_j are the eigenvalues of a pencil $T_0 - \lambda T_1$ where $T_0, T_1 \in \mathbb{C}^{m \times n}$ (with $m \geq n$) are Hankel matrices defined in terms of $\mathcal{M}_k$.

- Because of measurement errors the perturbed pencil $\widetilde{T}_0 - \lambda \widetilde{T}_1$ has generically no eigenvalues.

- Find a nearest pencil with a full set of eigenvalues
$$\inf_{\mathcal{S} \in \mathbb{C}^n} \tau_n(\widetilde{T}_0, \widetilde{T}_1, \mathcal{S}).$$

Shape Estimation from Moments

- Given moments
$$\mathcal{M}_k = \int \int_P z^k \, dx \, dy$$
for $k = 1, \ldots, m$.

  Estimate the vertices $z_j \in \mathbb{C}$ for $j = 1, \ldots, n$ of P.
- The vertices z_j are the eigenvalues of a pencil $T_0 - \lambda T_1$ where $T_0, T_1 \in \mathbb{C}^{m \times n}$ (with $m \geq n$) are Hankel matrices defined in terms of $\mathcal{M}_k$.
- Because of measurement errors the perturbed pencil $\widetilde{T}_0 - \lambda \widetilde{T}_1$ has generically no eigenvalues.
- Find a nearest pencil with a full set of eigenvalues
$$\inf_{\mathcal{S} \in \mathbb{C}^n} \tau_n(\widetilde{T}_0, \widetilde{T}_1, \mathcal{S}).$$

Shape Estimation from Moments

- Given moments
$$\mathcal{M}_k = \int\int_P z^k \, dx \, dy$$
for $k = 1, \ldots, m$.

  Estimate the vertices $z_j \in \mathbb{C}$ for $j = 1, \ldots, n$ of P.
- The vertices z_j are the eigenvalues of a pencil $T_0 - \lambda T_1$ where $T_0, T_1 \in \mathbb{C}^{m \times n}$ (with $m \geq n$) are Hankel matrices defined in terms of $\mathcal{M}_k$.
- Because of measurement errors the perturbed pencil $\widetilde{T}_0 - \lambda \widetilde{T}_1$ has generically no eigenvalues.
- Find a nearest pencil with a full set of eigenvalues
$$\inf_{\mathcal{S} \in \mathbb{C}^n} \; \tau_n(\widetilde{T}_0, \widetilde{T}_1, \mathcal{S}).$$

Outline

1. Introduction
   - Kronecker Canonical Form
   - Problem Definition
   - Motivation
2. Derivation of a Singular Value Characterization
   - Rank Characterization
   - Construction of an Optimal Perturbation
   - Simultaneous Perturbations
3. Numerical Examples

Rank Characterization

- Suppose $\alpha_A = 1$ and $\alpha_B = 0$.
- First step in derivation is a rank characterization for $\sum_{j=1}^{k} m_j(A, B) \geq r$.
- We benefit from a Sylvester operator point of view.

Theorem (Kernel of Sylvester Operator)

Let $A, B \in \mathbb{C}^{m \times n}$ be such that $m \geq n$ and $\mathrm{rank}(B) = n$, and $C \in \mathbb{C}^{r \times r}$. The solution space of the generalized Sylvester equation

$$AX - BXC = 0$$

depends on the Kronecker canonical form of $A - \lambda B$ and the Jordan form of C. Specifically suppose that $\mu_1, \ldots, \mu_\ell$ are the common eigenvalues of $A - \lambda B$ and C. Let $c_{j,1}, \ldots, c_{j,\ell_j}$ and $p_{j,1}, \ldots, p_{j,\tilde{\ell}_j}$ denote the sizes of the Jordan blocks of $A - \lambda B$ and C associated with the eigenvalue μ_j, respectively. Then

$$\dim\{X \in \mathbb{C}^{n \times r} : AX - BXC = 0\} = \sum_{j=1}^{\ell} \sum_{i=1}^{\ell_j} \sum_{q=1}^{\tilde{\ell}_j} \min(c_{j,i}, p_{j,q}).$$

Rank Characterization

- Suppose $\alpha_A = 1$ and $\alpha_B = 0$.
- First step in derivation is a rank characterization for $\sum_{j=1}^{k} m_j(A, B) \geq r$.
- We benefit from a Sylvester operator point of view.

Theorem (Kernel of Sylvester Operator)

Let $A, B \in \mathbb{C}^{m \times n}$ be such that $m \geq n$ and $\operatorname{rank}(B) = n$, and $C \in \mathbb{C}^{r \times r}$. The solution space of the generalized Sylvester equation

$$AX - BXC = 0$$

depends on the Kronecker canonical form of $A - \lambda B$ and the Jordan form of C. Specifically suppose that $\mu_1, \dots, \mu_\ell$ are the common eigenvalues of $A - \lambda B$ and C. Let $c_{j,1}, \dots, c_{j,\ell_j}$ and $p_{j,1}, \dots, p_{j,\tilde{\ell}_j}$ denote the sizes of the Jordan blocks of $A - \lambda B$ and C associated with the eigenvalue μ_j, respectively. Then

$$\dim\{X \in \mathbb{C}^{n \times r} : AX - BXC = 0\} = \sum_{j=1}^{\ell} \sum_{i=1}^{\ell_j} \sum_{q=1}^{\tilde{\ell}_j} \min(c_{j,i}, p_{j,q}).$$

Rank Characterization

- Suppose $\alpha_A = 1$ and $\alpha_B = 0$.
- First step in derivation is a rank characterization for $\sum_{j=1}^{k} m_j(A, B) \geq r$.
- We benefit from a Sylvester operator point of view.

Theorem (Kernel of Sylvester Operator)

Let $A, B \in \mathbb{C}^{m \times n}$ be such that $m \geq n$ and $\mathrm{rank}(B) = n$, and $C \in \mathbb{C}^{r \times r}$. The solution space of the generalized Sylvester equation

$$AX - BXC = 0$$

depends on the Kronecker canonical form of $A - \lambda B$ and the Jordan form of C. Specifically suppose that $\mu_1, \dots, \mu_\ell$ are the common eigenvalues of $A - \lambda B$ and C. Let $c_{j,1}, \dots, c_{j,\ell_j}$ and $p_{j,1}, \dots, p_{j,\tilde{\ell}_j}$ denote the sizes of the Jordan blocks of $A - \lambda B$ and C associated with the eigenvalue μ_j, respectively. Then

$$\dim\{X \in \mathbb{C}^{n \times r} : AX - BXC = 0\} = \sum_{j=1}^{\ell} \sum_{i=1}^{\ell_j} \sum_{q=1}^{\tilde{\ell}_j} \min(c_{j,i}, p_{j,q}).$$

Rank Characterization

- Suppose $\alpha_A = 1$ and $\alpha_B = 0$.
- First step in derivation is a rank characterization for $\sum_{j=1}^{k} m_j(A, B) \geq r$.
- We benefit from a Sylvester operator point of view.

Theorem (Kernel of Sylvester Operator)

Let $A, B \in \mathbb{C}^{m \times n}$ be such that $m \geq n$ and $\mathrm{rank}(B) = n$, and $C \in \mathbb{C}^{r \times r}$. The solution space of the generalized Sylvester equation

$$AX - BXC = 0$$

depends on the Kronecker canonical form of $A - \lambda B$ and the Jordan form of C. Specifically suppose that $\mu_1, \ldots, \mu_\ell$ are the common eigenvalues of $A - \lambda B$ and C. Let $c_{j,1}, \ldots, c_{j,\ell_j}$ and $p_{j,1}, \ldots, p_{j,\tilde{\ell}_j}$ denote the sizes of the Jordan blocks of $A - \lambda B$ and C associated with the eigenvalue μ_j, respectively. Then

$$\dim\{X \in \mathbb{C}^{n \times r} : AX - BXC = 0\} = \sum_{j=1}^{\ell} \sum_{i=1}^{\ell_j} \sum_{q=1}^{\tilde{\ell}_j} \min(c_{j,i}, p_{j,q}).$$

Rank Characterization

- Suppose $\alpha_A = 1$ and $\alpha_B = 0$.
- First step in derivation is a rank characterization for $\sum_{j=1}^{k} m_j(A, B) \geq r$.
- We benefit from a Sylvester operator point of view.

Theorem (Kernel of Sylvester Operator)

Let $A, B \in \mathbb{C}^{m\times n}$ be such that $m \geq n$ and $\mathrm{rank}(B) = n$, and $C \in \mathbb{C}^{r\times r}$. The solution space of the generalized Sylvester equation

$$AX - BXC = 0$$

depends on the Kronecker canonical form of $A - \lambda B$ and the Jordan form of C. Specifically suppose that $\mu_1, \dots, \mu_\ell$ are the common eigenvalues of $A - \lambda B$ and C. Let $c_{j,1}, \dots, c_{j,\ell_j}$ and $p_{j,1}, \dots, p_{j,\tilde{\ell}_j}$ denote the sizes of the Jordan blocks of $A - \lambda B$ and C associated with the eigenvalue μ_j, respectively. Then

$$\dim\{X \in \mathbb{C}^{n\times r} : AX - BXC = 0\} = \sum_{j=1}^{\ell} \sum_{i=1}^{\ell_j} \sum_{q=1}^{\tilde{\ell}_j} \min(c_{j,i}, p_{j,q}).$$

Rank Characterization

- Suppose $\alpha_A = 1$ and $\alpha_B = 0$.
- First step in derivation is a rank characterization for $\sum_{j=1}^{k} m_j(A, B) \geq r$.
- We benefit from a Sylvester operator point of view.

Theorem (Kernel of Sylvester Operator)

Let $A, B \in \mathbb{C}^{m \times n}$ be such that $m \geq n$ and $\mathrm{rank}(B) = n$, and $C \in \mathbb{C}^{r \times r}$. The solution space of the generalized Sylvester equation

$$AX - BXC = 0$$

depends on the Kronecker canonical form of $A - \lambda B$ and the Jordan form of C. Specifically suppose that $\mu_1, \ldots, \mu_\ell$ are the common eigenvalues of $A - \lambda B$ and C. Let $c_{j,1}, \ldots, c_{j,\ell_j}$ and $p_{j,1}, \ldots, p_{j,\tilde{\ell}_j}$ denote the sizes of the Jordan blocks of $A - \lambda B$ and C associated with the eigenvalue μ_j, respectively. Then

$$\dim\{X \in \mathbb{C}^{n \times r} : AX - BXC = 0\} = \sum_{j=1}^{\ell} \sum_{i=1}^{\ell_j} \sum_{q=1}^{\tilde{\ell}_j} \min(c_{j,i}, p_{j,q}).$$

Rank Characterization

- Suppose $C = C(\mu, \Gamma) = \begin{bmatrix} \mu_1 & -\gamma_{21} & \dots & -\gamma_{r1} \\ 0 & \mu_2 & \dots & -\gamma_{r2} \\ & & \ddots & \\ 0 & & & \mu_r \end{bmatrix}$, and

$$\mathcal{G} = \{\Gamma : C(\mu, \Gamma) \text{ has Jordan blocks of maximal size.}\}$$

Theorem (Sylvester Characterization)

Given a pencil $A - \lambda B$ with $A, B \in \mathbb{C}^{m \times n}$ such that $m \geq n$ and $\operatorname{rank}(B) = n$, a set $\mathcal{S} = \{\lambda_1, \dots, \lambda_k\}$ consisting of distinct complex scalars and a positive integer r. Then the following two statements are equivalent.

1. $\sum_{j=1}^{k} m_j(A, B) \geq r$
2. *There exists a $\mu \in \mathcal{S}^r$ such that*

$$\dim\{X \in \mathbb{C}^{n \times r} : AX - BXC(\mu, \Gamma) = 0\} \geq r$$

for all $\Gamma \in \mathcal{G}$.

Rank Characterization

- Suppose $C = C(\mu, \Gamma) = \begin{bmatrix} \mu_1 & -\gamma_{21} & \dots & -\gamma_{r1} \\ 0 & \mu_2 & \dots & -\gamma_{r2} \\ & & \ddots & \\ 0 & & & \mu_r \end{bmatrix}$, and

$$\mathcal{G} = \{\Gamma : C(\mu, \Gamma) \text{ has Jordan blocks of maximal size.}\}$$

Theorem (Sylvester Characterization)

Given a pencil $A - \lambda B$ with $A, B \in \mathbb{C}^{m \times n}$ such that $m \geq n$ and $\operatorname{rank}(B) = n$, a set $\mathcal{S} = \{\lambda_1, \dots, \lambda_k\}$ consisting of distinct complex scalars and a positive integer r. Then the following two statements are equivalent.

1. $\sum_{j=1}^{k} m_j(A, B) \geq r$
2. *There exists a $\mu \in \mathcal{S}^r$ such that*

$$\dim\{X \in \mathbb{C}^{n \times r} : AX - BXC(\mu, \Gamma) = 0\} \geq r$$

for all $\Gamma \in \mathcal{G}$.

Rank Characterization

- Suppose $C = C(\mu, \Gamma) = \begin{bmatrix} \mu_1 & -\gamma_{21} & \dots & -\gamma_{r1} \\ 0 & \mu_2 & \dots & -\gamma_{r2} \\ & & \ddots & \\ 0 & & & \mu_r \end{bmatrix}$, and

$$\mathcal{G} = \{\Gamma : C(\mu, \Gamma) \text{ has Jordan blocks of maximal size.}\}$$

Theorem (Sylvester Characterization)

Given a pencil $A - \lambda B$ with $A, B \in \mathbb{C}^{m \times n}$ such that $m \geq n$ and $\text{rank}(B) = n$, a set $\mathcal{S} = \{\lambda_1, \dots, \lambda_k\}$ consisting of distinct complex scalars and a positive integer r. Then the following two statements are equivalent.

1. $\sum_{j=1}^{k} m_j(A, B) \geq r$
2. *There exists a $\mu \in \mathcal{S}^r$ such that*

$$\dim\{X \in \mathbb{C}^{n \times r} : AX - BXC(\mu, \Gamma) = 0\} \geq r$$

for all $\Gamma \in \mathcal{G}$.

Rank Characterization

- Suppose $C = C(\mu, \Gamma) = \begin{bmatrix} \mu_1 & -\gamma_{21} & \cdots & -\gamma_{r1} \\ 0 & \mu_2 & \cdots & -\gamma_{r2} \\ & & \ddots & \\ 0 & & & \mu_r \end{bmatrix}$, and
$\mathcal{G} = \{\Gamma : C(\mu, \Gamma) \text{ has Jordan blocks of maximal size.}\}$

Theorem (Sylvester Characterization)

Given a pencil $A - \lambda B$ with $A, B \in \mathbb{C}^{m \times n}$ such that $m \geq n$ and $\mathrm{rank}(B) = n$, a set $\mathcal{S} = \{\lambda_1, \ldots, \lambda_k\}$ consisting of distinct complex scalars and a positive integer r. Then the following two statements are equivalent.

1. $\sum_{j=1}^{k} m_j(A, B) \geq r$
2. *There exists a $\mu \in \mathcal{S}^r$ such that*

$$\dim\{X \in \mathbb{C}^{n \times r} : AX - BXC(\mu, \Gamma) = 0\} \geq r$$

for all $\Gamma \in \mathcal{G}$.

Rank Characterization

Kroneckerization of the Sylvester Equation

- Recall the basic Identity for $X = \begin{bmatrix} X_1 & \dots & X_r \end{bmatrix} \in \mathbb{C}^{n \times r}$

$$\mathrm{vec}(FXG) = (G^T \otimes F)\mathrm{vec}(X), \quad \text{where} \quad \mathrm{vec}(X) = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_r \end{bmatrix} \in \mathbb{C}^{nr}$$

- In particular
$AX - BXC(\mu, \Gamma) = 0 \Leftrightarrow ((I \otimes A) - (C^T(\mu, \Gamma) \otimes B))\mathrm{vec}(X) = 0.$

- Consequently

$$\dim\{X \in \mathbb{C}^{n \times r} : AX - BXC(\mu, \Gamma) = 0\} \geq r$$
$$\Longleftrightarrow$$
$$\mathrm{rank}\left(((I \otimes A) - (C^T(\mu, \Gamma) \otimes B))\right) \leq nr - r$$

Rank Characterization

Kroneckerization of the Sylvester Equation

- Recall the basic Identity for $X = \begin{bmatrix} X_1 & \dots & X_r \end{bmatrix} \in \mathbb{C}^{n\times r}$

$$\mathrm{vec}(FXG) = (G^T \otimes F)\mathrm{vec}(X), \quad \text{where} \quad \mathrm{vec}(X) = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_r \end{bmatrix} \in \mathbb{C}^{nr}$$

- In particular
$AX - BXC(\mu,\Gamma) = 0 \Leftrightarrow ((I \otimes A) - (C^T(\mu,\Gamma) \otimes B))\mathrm{vec}(X) = 0.$

- Consequently

$$\dim\{X \in \mathbb{C}^{n\times r} : AX - BXC(\mu,\Gamma) = 0\} \geq r$$
$$\Longleftrightarrow$$
$$\mathrm{rank}\left(((I \otimes A) - (C^T(\mu,\Gamma) \otimes B))\right) \leq nr - r$$

Rank Characterization

Kroneckerization of the Sylvester Equation

- Recall the basic Identity for $X = \begin{bmatrix} X_1 & \dots & X_r \end{bmatrix} \in \mathbb{C}^{n \times r}$

$$\mathrm{vec}(FXG) = (G^T \otimes F)\mathrm{vec}(X), \quad \text{where} \;\; \mathrm{vec}(X) = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_r \end{bmatrix} \in \mathbb{C}^{nr}$$

- In particular
$AX - BXC(\mu, \Gamma) = 0 \Leftrightarrow ((I \otimes A) - (C^T(\mu, \Gamma) \otimes B))\mathrm{vec}(X) = 0.$

- Consequently

$$\dim\{X \in \mathbb{C}^{n \times r} : AX - BXC(\mu, \Gamma) = 0\} \geq r$$
$$\Longleftrightarrow$$
$$\mathrm{rank}\left(((I \otimes A) - (C^T(\mu, \Gamma) \otimes B))\right) \leq nr - r$$

Rank Characterization

Kroneckerization of the Sylvester Equation

$$\begin{aligned}\mathcal{L}(\mu, \Gamma, A, B) &:= \Big(((I \otimes A) - (C^T(\mu,\Gamma) \otimes B)\Big) \\ &= \begin{bmatrix} A - \mu_1 B & 0 & & & 0 \\ \gamma_{21} B & A - \mu_2 B & & & 0 \\ & & \ddots & & \\ & & & A - \mu_{r-1} B & 0 \\ \gamma_{r1} B & \gamma_{r2} B & & \gamma_{r(r-1)} B & A - \mu_r B \end{bmatrix}.\end{aligned}$$

Theorem (Rank Characterization)

Given a pencil $A - \lambda B$ with $A, B \in \mathbb{C}^{n \times m}$ such that $n \geq m$, a set $\mathcal{S} = \{\lambda_1, \dots, \lambda_k\}$ consisting of distinct complex scalars and a positive integer r. Then the following two statements are equivalent.

1. $\sum_{j=1}^{k} m_j(A, B) \geq r$
2. *There exists a $\mu \in \mathcal{S}^r$ such that*

$$\operatorname{rank}\left(\mathcal{L}(\mu, \Gamma, A, B)\right) \leq nr - r$$

for all $\Gamma \in \mathcal{G}$.

Rank Characterization

Kroneckerization of the Sylvester Equation

$$\mathcal{L}(\mu, \Gamma, A, B) := \left(((I \otimes A) - (C^T(\mu,\Gamma) \otimes B)\right)$$

$$= \begin{bmatrix} A - \mu_1 B & 0 & & & 0 \\ \gamma_{21} B & A - \mu_2 B & & & 0 \\ & & \ddots & & \\ & & & A - \mu_{r-1} B & 0 \\ \gamma_{r1} B & \gamma_{r2} B & & \gamma_{r(r-1)} B & A - \mu_r B \end{bmatrix}.$$

Theorem (Rank Characterization)

Given a pencil $A - \lambda B$ with $A, B \in \mathbb{C}^{n \times m}$ such that $n \geq m$, a set $\mathcal{S} = \{\lambda_1, \dots, \lambda_k\}$ consisting of distinct complex scalars and a positive integer r. Then the following two statements are equivalent.

1. $\sum_{j=1}^{k} m_j(A, B) \geq r$
2. *There exists a $\mu \in \mathcal{S}^r$ such that*

$$\operatorname{rank}(\mathcal{L}(\mu, \Gamma, A, B)) \leq nr - r$$

for all $\Gamma \in \mathcal{G}$.

Rank Characterization

Kroneckerization of the Sylvester Equation

$$\mathcal{L}(\mu, \Gamma, A, B) := \left(((I \otimes A) - (C^T(\mu, \Gamma) \otimes B) \right)$$
$$= \begin{bmatrix} A - \mu_1 B & 0 & & & 0 \\ \gamma_{21} B & A - \mu_2 B & & & 0 \\ & & \ddots & & \\ & & & A - \mu_{r-1} B & 0 \\ \gamma_{r1} B & \gamma_{r2} B & & \gamma_{r(r-1)} B & A - \mu_r B \end{bmatrix}.$$

Theorem (Rank Characterization)

Given a pencil $A - \lambda B$ with $A, B \in \mathbb{C}^{n \times m}$ such that $n \geq m$, a set $\mathcal{S} = \{\lambda_1, \ldots, \lambda_k\}$ consisting of distinct complex scalars and a positive integer r. Then the following two statements are equivalent.

1. $\sum_{j=1}^{k} m_j(A, B) \geq r$
2. *There exists a $\mu \in \mathcal{S}^r$ such that*
$$\operatorname{rank}\left(\mathcal{L}(\mu, \Gamma, A, B)\right) \leq nr - r$$
for all $\Gamma \in \mathcal{G}$.

Rank Characterization

Kroneckerization of the Sylvester Equation

$$\mathcal{L}(\mu, \Gamma, A, B) := \Big(((I \otimes A) - (C^T(\mu,\Gamma) \otimes B) \Big)$$
$$= \begin{bmatrix} A - \mu_1 B & 0 & & & 0 \\ \gamma_{21} B & A - \mu_2 B & & & 0 \\ & & \ddots & & \\ & & & A - \mu_{r-1} B & 0 \\ \gamma_{r1} B & \gamma_{r2} B & & \gamma_{r(r-1)} B & A - \mu_r B \end{bmatrix}.$$

Theorem (Rank Characterization)

Given a pencil $A - \lambda B$ with $A, B \in \mathbb{C}^{n \times m}$ such that $n \geq m$, a set $\mathcal{S} = \{\lambda_1, \dots, \lambda_k\}$ consisting of distinct complex scalars and a positive integer r. Then the following two statements are equivalent.

1. $\sum_{j=1}^{k} m_j(A, B) \geq r$
2. *There exists a $\mu \in \mathcal{S}^r$ such that*
$$\operatorname{rank}(\mathcal{L}(\mu, \Gamma, A, B)) \leq nr - r$$
for all $\Gamma \in \mathcal{G}$.

Rank Characterization

The rank characterization yields a lower bound on the distance.

Theorem (Minimal Rank)

Given $C \in \mathbb{C}^{\ell \times q}$ and a positive integer $p < \min(\ell, q)$. Then

$$\inf\{\|\delta C\|_2 : \mathrm{rank}(C + \delta C) \le p\} = \sigma_{p+1}(C).$$

- From an application of the theorem above

$$\begin{aligned}\tau_r(A, B, \mathcal{S}) = \quad & \inf_{\mu \in \mathcal{S}^r} \ \inf\{\|\delta A\| : \forall \Gamma \in \mathcal{G} \ \ \mathrm{rank}\left(\mathcal{L}(\mu, \Gamma, A + \delta A, B)\right) \le nr - r\} \\ \ge \quad & \inf_{\mu \in \mathcal{S}^r} \ \sup_{\Gamma} \sigma_{nr-r+1}\left(\mathcal{L}(\mu, \Gamma, A, B)\right)\end{aligned}$$

Rank Characterization

The rank characterization yields a lower bound on the distance.

Theorem (Minimal Rank)

Given $C \in \mathbb{C}^{\ell \times q}$ *and a positive integer* $p < \min(\ell, q)$. *Then*

$$\inf\{\|\delta C\|_2 : \operatorname{rank}(C + \delta C) \leq p\} = \sigma_{p+1}(C).$$

- From an application of the theorem above

$$\begin{aligned}\tau_r(A, B, \mathcal{S}) = \quad & \inf_{\mu \in \mathcal{S}^r} \ \inf\{\|\delta A\| : \forall \Gamma \in \mathcal{G} \ \operatorname{rank}\left(\mathcal{L}(\mu, \Gamma, A + \delta A, B)\right) \leq nr - r\} \\ \geq \quad & \inf_{\mu \in \mathcal{S}^r} \ \sup_{\Gamma} \sigma_{nr-r+1}\left(\mathcal{L}(\mu, \Gamma, A, B)\right)\end{aligned}$$

Rank Characterization

The rank characterization yields a lower bound on the distance.

Theorem (Minimal Rank)

Given $C \in \mathbb{C}^{\ell \times q}$ and a positive integer $p < \min(\ell, q)$. Then

$$\inf\{\|\delta C\|_2 : \mathrm{rank}(C + \delta C) \leq p\} = \sigma_{p+1}(C).$$

- From an application of the theorem above

$$\begin{aligned} \tau_r(A, B, \mathcal{S}) = & \quad \inf_{\mu \in \mathcal{S}^r} \inf\{\|\delta A\| : \forall \Gamma \in \mathcal{G} \;\; \mathrm{rank}\,(\mathcal{L}(\mu, \Gamma, A + \delta A, B)) \leq nr - r\} \\ \geq & \quad \inf_{\mu \in \mathcal{S}^r} \sup_{\Gamma} \sigma_{nr-r+1}\,(\mathcal{L}(\mu, \Gamma, A, B)) \end{aligned}$$

Rank Characterization

The rank characterization yields a lower bound on the distance.

Theorem (Minimal Rank)

Given $C \in \mathbb{C}^{\ell \times q}$ *and a positive integer* $p < \min(\ell, q)$*. Then*

$$\inf\{\|\delta C\|_2 : \mathrm{rank}(C + \delta C) \leq p\} = \sigma_{p+1}(C).$$

- From an application of the theorem above

$$\begin{aligned}\tau_r(A, B, \mathcal{S}) = & \quad \inf_{\mu \in \mathcal{S}^r} \ \inf\{\|\delta A\| : \forall \Gamma \in \mathcal{G} \ \ \mathrm{rank}\,(\mathcal{L}(\mu, \Gamma, A + \delta A, B)) \leq nr - r\} \\ \geq & \quad \inf_{\mu \in \mathcal{S}^r} \ \sup_{\Gamma} \sigma_{nr-r+1}\,(\mathcal{L}(\mu, \Gamma, A, B))\end{aligned}$$

Outline

Construction of an Optimal Perturbation

We deduce the other direction (for all μ)

$$\inf\{\|\delta A\| : \forall \Gamma \in \mathcal{G} \;\; \text{rank}\left(\mathcal{L}(\mu, \Gamma, A + \delta A, B)\right) \leq nr - r\}$$
$$\leq$$
$$\sup_{\Gamma} \sigma_{nr-r+1}\left(\mathcal{L}(\mu, \Gamma, A, B)\right)$$

under mild multiplicity, linear independence, and full Jordan block assumptions by constructing an optimal perturbation.

Construction of an Optimal Perturbation

Let $\kappa_r(\mu) := \sup_\Gamma \sigma_{nr-r+1}\left(\mathcal{L}(\mu, \Gamma, A, B)\right)$.

Construct a perturbation δA_* such that

1. $\|\delta A_*\| = \kappa_r(\mu)$
2. $\text{rank}\left(\mathcal{L}(\mu, \Gamma, A + \delta A_*, B)\right) \leq nr - r \quad \exists \Gamma \in \mathcal{G}$

Construction of an Optimal Perturbation

Let $\kappa_r(\mu) := \sup_\Gamma \sigma_{nr-r+1}\left(\mathcal{L}(\mu, \Gamma, A, B)\right)$.

Construct a perturbation δA_* such that

1. $\|\delta A_*\| = \kappa_r(\mu)$
2. $\operatorname{rank}\left(\mathcal{L}(\mu, \Gamma, A + \delta A_*, B)\right) \leq nr - r \quad \exists \Gamma \in \mathcal{G}$

Construction of an Optimal Perturbation

Let $\kappa_r(\mu) := \sup_\Gamma \sigma_{nr-r+1}\left(\mathcal{L}(\mu, \Gamma, A, B)\right)$.

Construct a perturbation δA_* such that

1. $\|\delta A_*\| = \kappa_r(\mu)$
2. $\text{rank}\left(\mathcal{L}(\mu, \Gamma, A + \delta A_*, B)\right) \leq nr - r \quad \exists \Gamma \in \mathcal{G}$

Construction of an Optimal Perturbation

Specifically $\delta A_* := -\kappa_r(\mu)\mathcal{U}\mathcal{V}^+$.

Above $\mathcal{U} \in \mathbb{C}^{m\times r}$ and $\mathcal{V} \in \mathbb{C}^{n\times r}$ are defined as follows.

- Let Γ_* be the optimal Γ satisfying

$$\kappa_r(\mu) = \sigma_{nr-r+1}\left(\mathcal{L}(\mu, \Gamma_*, A, B)\right)$$

- Consider the left and right singular vectors assoc with $\kappa_r(\mu)$

$$\mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \quad \text{and} \quad U^*\ \mathcal{L}(\mu, \Gamma_*, A, B) = V^*\ \kappa_r(\mu).$$

- Then $\mathcal{U}$ and $\mathcal{V}$ are such that

$$\text{vec}\,(\mathcal{U}) = U \quad \text{and} \quad \text{vec}\,(\mathcal{V}) = V.$$

Construction of an Optimal Perturbation

Specifically $\delta A_* := -\kappa_r(\mu)\mathcal{U}\mathcal{V}^+$.

Above $\mathcal{U} \in \mathbb{C}^{m\times r}$ and $\mathcal{V} \in \mathbb{C}^{n\times r}$ are defined as follows.

- Let Γ_* be the optimal Γ satisfying

$$\kappa_r(\mu) = \sigma_{nr-r+1}\left(\mathcal{L}(\mu, \Gamma_*, A, B)\right)$$

- Consider the left and right singular vectors assoc with $\kappa_r(\mu)$

$$\mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \quad \text{and} \quad U^*\ \mathcal{L}(\mu, \Gamma_*, A, B) = V^*\ \kappa_r(\mu).$$

- Then $\mathcal{U}$ and $\mathcal{V}$ are such that

$$\text{vec}\,(\mathcal{U}) = U \quad \text{and} \quad \text{vec}\,(\mathcal{V}) = V.$$

Construction of an Optimal Perturbation

Specifically $\delta A_* := -\kappa_r(\mu)\mathcal{U}\mathcal{V}^+$.

Above $\mathcal{U} \in \mathbb{C}^{m\times r}$ and $\mathcal{V} \in \mathbb{C}^{n\times r}$ are defined as follows.

- Let Γ_* be the optimal Γ satisfying

$$\kappa_r(\mu) = \sigma_{nr-r+1}\left(\mathcal{L}(\mu, \Gamma_*, A, B)\right)$$

- Consider the left and right singular vectors assoc with $\kappa_r(\mu)$

$$\mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \ \text{ and } \ U^*\ \mathcal{L}(\mu, \Gamma_*, A, B) = V^*\ \kappa_r(\mu).$$

- Then $\mathcal{U}$ and $\mathcal{V}$ are such that

$$\mathrm{vec}\left(\mathcal{U}\right) = U \ \text{ and } \ \mathrm{vec}\left(\mathcal{V}\right) = V.$$

Construction of an Optimal Perturbation

Specifically $\delta A_* := -\kappa_r(\mu)\mathcal{U}\mathcal{V}^+$.

Above $\mathcal{U} \in \mathbb{C}^{m\times r}$ and $\mathcal{V} \in \mathbb{C}^{n\times r}$ are defined as follows.

- Let Γ_* be the optimal Γ satisfying

$$\kappa_r(\mu) = \sigma_{nr-r+1}\left(\mathcal{L}(\mu, \Gamma_*, A, B)\right)$$

- Consider the left and right singular vectors assoc with $\kappa_r(\mu)$

$$\mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \ \text{ and } \ U^*\ \mathcal{L}(\mu, \Gamma_*, A, B) = V^*\ \kappa_r(\mu).$$

- Then $\mathcal{U}$ and $\mathcal{V}$ are such that

$$\text{vec}\,(\mathcal{U}) = U \ \text{ and } \ \text{vec}\,(\mathcal{V}) = V.$$

Construction of an Optimal Perturbation

Theorem (Rellich, 1937-1942)

Let $f : \mathbb{R} \to \mathbb{R}$ be defined as $f(\omega) = \sigma_j(\mathcal{A}(\omega))$ (σ_j denoting the jth largest singular value) where $\mathcal{A}(\omega)$ is an analytic matrix-valued function of ω. If the multiplicity of $\sigma_j(\mathcal{A}(\tilde{\omega}))$ is one and $\sigma_j(\mathcal{A}(\tilde{\omega})) \neq 0$, then $f(\omega)$ is real analytic at $\tilde{\omega}$ with the derivative

$$f'(\tilde{\omega}) = \text{Real}\left(u^* \frac{d\mathcal{A}(\tilde{\omega})}{d\omega} v \right)$$

where u and v consist of a consistent pair of a unit left and a right singular vector associated with $\sigma_j(\mathcal{A}(\tilde{\omega}))$.

Construction of an Optimal Perturbation

Theorem (Rellich, 1937-1942)

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be defined as $f(\omega) = \sigma_j(\mathcal{A}(\omega))$ (σ_j denoting the jth largest singular value) where $\mathcal{A}(\omega)$ is an analytic matrix-valued function of ω. If the multiplicity of $\sigma_j(\mathcal{A}(\tilde{\omega}))$ is one and $\sigma_j(\mathcal{A}(\tilde{\omega})) \neq 0$, then $f(\omega)$ is real analytic at $\tilde{\omega}$ with the derivative

$$f'(\tilde{\omega}) = \mathrm{Real}\left(u^* \frac{d\mathcal{A}(\tilde{\omega})}{d\omega} v \right)$$

where u and v consist of a consistent pair of a unit left and a right singular vector associated with $\sigma_j(\mathcal{A}(\tilde{\omega}))$.

Construction of an Optimal Perturbation

(1) $\|\delta A_*\| = \| -\kappa_r(\mu)\mathcal{U}\mathcal{V}^+\| = \kappa_r(\mu)$

- Applying the theorem (Rellich) to $\sigma_{nr-r+1}(\mathcal{L}(\mu, \Gamma, A, B))$, the optimality of Γ_* implies

$$\mathcal{U}^*\mathcal{U} = \mathcal{V}^*\mathcal{V}$$

- Consider

$$\mathcal{U}\mathcal{V}^+ = \mathcal{U}\mathcal{V}^+\mathcal{V}\mathcal{V}^+$$

  - $\mathcal{V}\mathcal{V}^+$ first orthogonally projects a vector onto $\mathrm{Range}(\mathcal{V})$
  - $\mathcal{U}\mathcal{V}^+$ then changes the coordinates from $\mathcal{V}$ to $\mathcal{U}$
  - Consequently $\|\mathcal{U}\mathcal{V}^+\| = 1$.

Construction of an Optimal Perturbation

(1) $\|\delta A_*\| = \| - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+\| = \kappa_r(\mu)$

- Applying the theorem (Rellich) to $\sigma_{nr-r+1}(\mathcal{L}(\mu, \Gamma, A, B))$, the optimality of Γ_* implies

$$\mathcal{U}^*\mathcal{U} = \mathcal{V}^*\mathcal{V}$$

- Consider

$$\mathcal{U}\mathcal{V}^+ = \mathcal{U}\mathcal{V}^+\mathcal{V}\mathcal{V}^+$$

  - $\mathcal{V}\mathcal{V}^+$ first orthogonally projects a vector onto $\mathrm{Range}(\mathcal{V})$
  - $\mathcal{U}\mathcal{V}^+$ then changes the coordinates from $\mathcal{V}$ to $\mathcal{U}$
  - Consequently $\|\mathcal{U}\mathcal{V}^+\| = 1$.

Construction of an Optimal Perturbation

(1) $\|\delta A_*\| = \| - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+\| = \kappa_r(\mu)$

- Applying the theorem (Rellich) to $\sigma_{nr-r+1}\left(\mathcal{L}(\mu, \Gamma, A, B)\right)$, the optimality of Γ_* implies

$$\mathcal{U}^*\mathcal{U} = \mathcal{V}^*\mathcal{V}$$

- Consider

$$\mathcal{U}\mathcal{V}^+ = \mathcal{U}\mathcal{V}^+\mathcal{V}\mathcal{V}^+$$

- $\mathcal{V}\mathcal{V}^+$ first orthogonally projects a vector onto $\mathrm{Range}(\mathcal{V})$
- $\mathcal{U}\mathcal{V}^+$ then changes the coordinates from $\mathcal{V}$ to $\mathcal{U}$
- Consequently $\|\mathcal{U}\mathcal{V}^+\| = 1$.

Construction of an Optimal Perturbation

(1) $\|\delta A_*\| = \| - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+\| = \kappa_r(\mu)$

- Applying the theorem (Rellich) to $\sigma_{nr-r+1}(\mathcal{L}(\mu, \Gamma, A, B))$, the optimality of Γ_* implies

$$\mathcal{U}^*\mathcal{U} = \mathcal{V}^*\mathcal{V}$$

- Consider

$$\mathcal{U}\mathcal{V}^+ = \mathcal{U}\mathcal{V}^+\mathcal{V}\mathcal{V}^+$$

  - $\mathcal{V}\mathcal{V}^+$ first orthogonally projects a vector onto $\mathrm{Range}(\mathcal{V})$
  - $\mathcal{U}\mathcal{V}^+$ then changes the coordinates from $\mathcal{V}$ to $\mathcal{U}$
  - Consequently $\|\mathcal{U}\mathcal{V}^+\| = 1$.

Construction of an Optimal Perturbation

(1) $\|\delta A_*\| = \| -\kappa_r(\mu)\mathcal{U}\mathcal{V}^+\| = \kappa_r(\mu)$

- Applying the theorem (Rellich) to $\sigma_{nr-r+1}\,(\mathcal{L}(\mu, \Gamma, A, B))$, the optimality of Γ_* implies

$$\mathcal{U}^*\mathcal{U} = \mathcal{V}^*\mathcal{V}$$

- Consider

$$\mathcal{U}\mathcal{V}^+ = \mathcal{U}\mathcal{V}^+\mathcal{V}\mathcal{V}^+$$

  - $\mathcal{V}\mathcal{V}^+$ first orthogonally projects a vector onto $\mathrm{Range}(\mathcal{V})$
  - $\mathcal{U}\mathcal{V}^+$ then changes the coordinates from $\mathcal{V}$ to $\mathcal{U}$
  - Consequently $\|\mathcal{U}\mathcal{V}^+\| = 1$.

Construction of an Optimal Perturbation

(1) $\|\delta A_*\| = \| - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+\| = \kappa_r(\mu)$

- Applying the theorem (Rellich) to $\sigma_{nr-r+1}\left(\mathcal{L}(\mu, \Gamma, A, B)\right)$, the optimality of Γ_* implies

$$\mathcal{U}^*\mathcal{U} = \mathcal{V}^*\mathcal{V}$$

- Consider

$$\mathcal{U}\mathcal{V}^+ = \mathcal{U}\mathcal{V}^+\mathcal{V}\mathcal{V}^+$$

  - $\mathcal{V}\mathcal{V}^+$ first orthogonally projects a vector onto $\mathrm{Range}(\mathcal{V})$
  - $\mathcal{U}\mathcal{V}^+$ then changes the coordinates from $\mathcal{V}$ to $\mathcal{U}$
  - Consequently $\|\mathcal{U}\mathcal{V}^+\| = 1$.

Construction of an Optimal Perturbation

(2) $\operatorname{rank}\left(\mathcal{L}(\mu, \Gamma_*, A + \delta A_*, B)\right) \leq nr - r$

- Suppose that the columns of $\mathcal{V}$ are linearly independent.

$$\begin{aligned}
\mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \quad & \iff \quad ((I \otimes A) - (C^T(\mu, \Gamma_*) \otimes B))\ \ V = \kappa_r(\mu)\ U \\
& \iff \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U} \\
& \iff \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U}\ (\mathcal{V}^+\mathcal{V}) \\
& \implies \quad (A - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0 \\
& \implies \quad (A + \delta A_*)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0.
\end{aligned}$$

- Now any matrix in the subspace

$$\{\mathcal{V}D : C(\mu, \Gamma)D - DC(\mu, \Gamma) = 0\}$$

of dimension at least r is a solution for the Sylvester equation *i.e.*,

$$(A + \delta A_*)\mathcal{V}D - B\mathcal{V}C(\mu, \Gamma_*)D = 0 \iff (A + \delta A_*)\,(\mathcal{V}D) - B\,(\mathcal{V}D)\,C(\mu, \Gamma_*) = 0$$

- Consequently

$$\operatorname{rank}\Big((I \otimes (A + \delta A_*)) - (C^T(\mu, \Gamma_*) \otimes B)\Big) \leq nr - r$$

Construction of an Optimal Perturbation

(2) $\text{rank}\left(\mathcal{L}(\mu, \Gamma_*, A + \delta A_*, B)\right) \leq nr - r$

- Suppose that the columns of $\mathcal{V}$ are linearly independent.

$$\begin{aligned}
\mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \quad &\Longleftrightarrow \quad \left((I \otimes A) - (C^T(\mu, \Gamma_*) \otimes B)\right)\ V = \kappa_r(\mu)\ U \\
&\Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U} \\
&\Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U}\ (\mathcal{V}^+\mathcal{V}) \\
&\Longrightarrow \quad (A - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0 \\
&\Longrightarrow \quad (A + \delta A_*)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0.
\end{aligned}$$

- Now any matrix in the subspace

$$\{\mathcal{V}D : C(\mu, \Gamma)D - DC(\mu, \Gamma) = 0\}$$

of dimension at least r is a solution for the Sylvester equation *i.e.*,

$$(A + \delta A_*)\mathcal{V}D - B\mathcal{V}C(\mu, \Gamma_*)D = 0 \Longleftrightarrow (A + \delta A_*)\,(\mathcal{V}D) - B\,(\mathcal{V}D)\,C(\mu, \Gamma_*) = 0$$

- Consequently

$$\text{rank}\left(\left(I \otimes (A + \delta A_*)\right) - \left(C^T(\mu, \Gamma_*) \otimes B\right)\right) \leq nr - r$$

Construction of an Optimal Perturbation

(2) $\text{rank}\left(\mathcal{L}(\mu, \Gamma_*, A + \delta A_*, B)\right) \leq nr - r$

- Suppose that the columns of $\mathcal{V}$ are linearly independent.

$$\begin{aligned}
\mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \quad &\Longleftrightarrow \quad ((I \otimes A) - (C^T(\mu, \Gamma_*) \otimes B))\ \ V = \kappa_r(\mu)\ U \\
&\Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U} \\
&\Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U}\ (\mathcal{V}^+\mathcal{V}) \\
&\Longrightarrow \quad (A - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0 \\
&\Longrightarrow \quad (A + \delta A_*)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0.
\end{aligned}$$

- Now any matrix in the subspace

$$\{\mathcal{V}D : C(\mu, \Gamma)D - DC(\mu, \Gamma) = 0\}$$

of dimension at least r is a solution for the Sylvester equation *i.e.*,

$$(A + \delta A_*)\mathcal{V}D - B\mathcal{V}C(\mu, \Gamma_*)D = 0 \Longleftrightarrow (A + \delta A_*)\,(\mathcal{V}D) - B\,(\mathcal{V}D)\,C(\mu, \Gamma_*) = 0$$

- Consequently

$$\text{rank}\left(((I \otimes (A + \delta A_*)) - (C^T(\mu, \Gamma_*) \otimes B)\right) \leq nr - r$$

Construction of an Optimal Perturbation

(2) $\operatorname{rank}\left(\mathcal{L}(\mu, \Gamma_*, A + \delta A_*, B)\right) \leq nr - r$

- Suppose that the columns of $\mathcal{V}$ are linearly independent.

$$\begin{aligned}
\mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \quad &\Longleftrightarrow \quad \left((I \otimes A) - (C^T(\mu, \Gamma_*) \otimes B)\right)\ V = \kappa_r(\mu)\ U \\
&\Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U} \\
&\Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U}\ (\mathcal{V}^+\mathcal{V}) \\
&\Longrightarrow \quad (A - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0 \\
&\Longrightarrow \quad (A + \delta A_*)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0.
\end{aligned}$$

- Now any matrix in the subspace

$$\{\mathcal{V}D : C(\mu, \Gamma)D - DC(\mu, \Gamma) = 0\}$$

of dimension at least r is a solution for the Sylvester equation *i.e.*,

$$(A + \delta A_*)\mathcal{V}D - B\mathcal{V}C(\mu, \Gamma_*)D = 0 \Longleftrightarrow (A + \delta A_*)\,(\mathcal{V}D) - B\,(\mathcal{V}D)\,C(\mu, \Gamma_*) = 0$$

- Consequently

$$\operatorname{rank}\left(\left(I \otimes (A + \delta A_*)\right) - \left(C^T(\mu, \Gamma_*) \otimes B\right)\right) \leq nr - r$$

Construction of an Optimal Perturbation

(2) $\text{rank}\left(\mathcal{L}(\mu, \Gamma_*, A + \delta A_*, B)\right) \leq nr - r$

- Suppose that the columns of $\mathcal{V}$ are linearly independent.

$$\begin{aligned}
\mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \quad &\Longleftrightarrow \quad ((I \otimes A) - (C^T(\mu, \Gamma_*) \otimes B))\ \ V = \kappa_r(\mu)\ U \\
&\Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U} \\
&\Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U}\left(\mathcal{V}^+\mathcal{V}\right) \\
&\Longrightarrow \quad (A - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0 \\
&\Longrightarrow \quad (A + \delta A_*)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0.
\end{aligned}$$

- Now any matrix in the subspace

$$\{\mathcal{V}D : C(\mu, \Gamma)D - DC(\mu, \Gamma) = 0\}$$

of dimension at least r is a solution for the Sylvester equation *i.e.*,

$$(A + \delta A_*)\mathcal{V}D - B\mathcal{V}C(\mu, \Gamma_*)D = 0 \Longleftrightarrow (A + \delta A_*)\left(\mathcal{V}D\right) - B\left(\mathcal{V}D\right)C(\mu, \Gamma_*) = 0$$

- Consequently

$$\text{rank}\left(\left(I \otimes (A + \delta A_*)\right) - \left(C^T(\mu, \Gamma_*) \otimes B\right)\right) \leq nr - r$$

Construction of an Optimal Perturbation

(2) $\operatorname{rank}\left(\mathcal{L}(\mu, \Gamma_*, A + \delta A_*, B)\right) \leq nr - r$

- Suppose that the columns of $\mathcal{V}$ are linearly independent.

$$
\begin{aligned}
\mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \quad &\Longleftrightarrow \quad ((I \otimes A) - (C^T(\mu, \Gamma_*) \otimes B))\ \ V = \kappa_r(\mu)\ U \\
&\Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U} \\
&\Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U}\left(\mathcal{V}^+\mathcal{V}\right) \\
&\Longrightarrow \quad (A - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0 \\
&\Longrightarrow \quad (A + \delta A_*)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0.
\end{aligned}
$$

- Now any matrix in the subspace

$$\{\mathcal{V}D : C(\mu, \Gamma)D - DC(\mu, \Gamma) = 0\}$$

of dimension at least r is a solution for the Sylvester equation *i.e.*,

$$(A + \delta A_*)\mathcal{V}D - B\mathcal{V}C(\mu, \Gamma_*)D = 0 \Longleftrightarrow (A + \delta A_*)\,(\mathcal{V}D) - B\,(\mathcal{V}D)\,C(\mu, \Gamma_*) = 0$$

- Consequently

$$\operatorname{rank}\left(\left(I \otimes (A + \delta A_*)\right) - \left(C^T(\mu, \Gamma_*) \otimes B\right)\right) \leq nr - r$$

Construction of an Optimal Perturbation

(2) $\operatorname{rank}\left(\mathcal{L}(\mu, \Gamma_*, A + \delta A_*, B)\right) \leq nr - r$

- Suppose that the columns of $\mathcal{V}$ are linearly independent.

$$
\begin{aligned}
\mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \quad &\Longleftrightarrow \quad \left((I \otimes A) - (C^T(\mu, \Gamma_*) \otimes B)\right)\ V = \kappa_r(\mu)\ U \\
&\Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U} \\
&\Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U}\left(\mathcal{V}^+\mathcal{V}\right) \\
&\Longrightarrow \quad (A - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0 \\
&\Longrightarrow \quad (A + \delta A_*)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0.
\end{aligned}
$$

- Now any matrix in the subspace

$$\{\mathcal{V}D : C(\mu, \Gamma)D - DC(\mu, \Gamma) = 0\}$$

of dimension at least r is a solution for the Sylvester equation *i.e.*,

$$(A + \delta A_*)\mathcal{V}D - B\mathcal{V}C(\mu, \Gamma_*)D = 0 \Longleftrightarrow (A + \delta A_*)(\mathcal{V}D) - B(\mathcal{V}D)\, C(\mu, \Gamma_*) = 0$$

- Consequently

$$\operatorname{rank}\left(\left(I \otimes (A + \delta A_*)\right) - \left(C^T(\mu, \Gamma_*) \otimes B\right)\right) \leq nr - r$$

Construction of an Optimal Perturbation

(2) $\operatorname{rank}\left(\mathcal{L}(\mu, \Gamma_*, A + \delta A_*, B)\right) \leq nr - r$

- Suppose that the columns of $\mathcal{V}$ are linearly independent.

$$\begin{aligned} \mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \quad & \Longleftrightarrow \quad \left((I \otimes A) - (C^T(\mu, \Gamma_*) \otimes B)\right)\ V = \kappa_r(\mu)\ U \\ & \Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U} \\ & \Longleftrightarrow \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U}\left(\mathcal{V}^+\mathcal{V}\right) \\ & \Longrightarrow \quad (A - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0 \\ & \Longrightarrow \quad (A + \delta A_*)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0. \end{aligned}$$

- Now any matrix in the subspace

$$\{\mathcal{V}D : C(\mu, \Gamma)D - DC(\mu, \Gamma) = 0\}$$

of dimension at least r is a solution for the Sylvester equation *i.e.*,

$$(A + \delta A_*)\mathcal{V}D - B\mathcal{V}C(\mu, \Gamma_*)D = 0 \Longleftrightarrow (A + \delta A_*)\,(\mathcal{V}D) - B\,(\mathcal{V}D)\,C(\mu, \Gamma_*) = 0$$

- Consequently

$$\operatorname{rank}\left(\left(I \otimes (A + \delta A_*)\right) - \left(C^T(\mu, \Gamma_*) \otimes B\right)\right) \leq nr - r$$

Construction of an Optimal Perturbation

(2) $\text{rank}\left(\mathcal{L}(\mu, \Gamma_*, A + \delta A_*, B)\right) \leq nr - r$

- Suppose that the columns of $\mathcal{V}$ are linearly independent.

$$\begin{aligned} \mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \quad & \iff \quad ((I \otimes A) - (C^T(\mu, \Gamma_*) \otimes B))\ \ V = \kappa_r(\mu)\ U \\ & \iff \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U} \\ & \iff \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U}\left(\mathcal{V}^+\mathcal{V}\right) \\ & \implies \quad (A - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0 \\ & \implies \quad (A + \delta A_*)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0. \end{aligned}$$

- Now any matrix in the subspace

$$\{\mathcal{V}D : C(\mu, \Gamma)D - DC(\mu, \Gamma) = 0\}$$

of dimension at least r is a solution for the Sylvester equation *i.e.*,

$$(A + \delta A_*)\mathcal{V}D - B\mathcal{V}C(\mu, \Gamma_*)D = 0 \iff (A + \delta A_*)\,(\mathcal{V}D) - B\,(\mathcal{V}D)\,C(\mu, \Gamma_*) = 0$$

- Consequently

$$\text{rank}\left(\left(I \otimes (A + \delta A_*)\right) - \left(C^T(\mu, \Gamma_*) \otimes B\right)\right) \leq nr - r$$

Construction of an Optimal Perturbation

(2) $\text{rank}\left(\mathcal{L}(\mu, \Gamma_*, A + \delta A_*, B)\right) \leq nr - r$

- Suppose that the columns of $\mathcal{V}$ are linearly independent.

$$\begin{aligned}
\mathcal{L}(\mu, \Gamma_*, A, B)\ V = \kappa_r(\mu)\ U \quad & \iff \quad ((I \otimes A) - (C^T(\mu, \Gamma_*) \otimes B))\ \ V = \kappa_r(\mu)\ U \\
& \iff \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U} \\
& \iff \quad A\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = \kappa_r(\mu)\mathcal{U}\left(\mathcal{V}^+\mathcal{V}\right) \\
& \implies \quad (A - \kappa_r(\mu)\mathcal{U}\mathcal{V}^+)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0 \\
& \implies \quad (A + \delta A_*)\mathcal{V} - B\mathcal{V}C(\mu, \Gamma_*) = 0.
\end{aligned}$$

- Now any matrix in the subspace

$$\{\mathcal{V}D : C(\mu, \Gamma)D - DC(\mu, \Gamma) = 0\}$$

of dimension at least r is a solution for the Sylvester equation *i.e.*,

$$(A + \delta A_*)\mathcal{V}D - B\mathcal{V}C(\mu, \Gamma_*)D = 0 \iff (A + \delta A_*)\,(\mathcal{V}D) - B\,(\mathcal{V}D)\,C(\mu, \Gamma_*) = 0$$

- Consequently

$$\text{rank}\left(\left(I \otimes (A + \delta A_*)\right) - \left(C^T(\mu, \Gamma_*) \otimes B\right)\right) \leq nr - r$$

Construction of an Optimal Perturbation

Theorem (Nearest Pencils with Specified Eigenvalues)

Let $A - \lambda B$ be an $m \times n$ pencil with $m \geq n$, r be a positive integer and $\mathcal{S} = \{\lambda_1, \ldots, \lambda_k\}$ be a set of distinct complex scalars.

1. *Then the equality*

$$\tau_r(A, B, \mathcal{S}) = \inf_{\mu \in \mathcal{S}^r} \sup_{\Gamma} \sigma_{nr-r+1} \left(\mathcal{L} \left(\mu, \Gamma, A, B \right) \right)$$

holds provided that the optimization problem on the right is attained at a (μ_*, Γ_*) where the multiplicity, linear independence and full Jordan block qualifications hold.

2. *Furthermore a minimal perturbation δA_* such that $\sum_{j=1}^{k} m(A + \delta A_*, B) \geq r$ is given by*

$$\delta A_* = -\tau_r(A, B, \mathcal{S}) \mathcal{U} \mathcal{V}^+$$

where the matrices $\mathcal{U}$ and $\mathcal{V}$ are formed by the left and right singular vectors at the optimal (μ_*, Γ_*).

Construction of an Optimal Perturbation

Theorem (Nearest Pencils with Specified Eigenvalues)

Let $A - \lambda B$ be an $m \times n$ pencil with $m \geq n$, r be a positive integer and $\mathcal{S} = \{\lambda_1, \ldots, \lambda_k\}$ be a set of distinct complex scalars.

1. *Then the equality*

$$\tau_r(A, B, \mathcal{S}) = \inf_{\mu \in \mathcal{S}^r} \sup_{\Gamma} \sigma_{nr-r+1} \left(\mathcal{L}\left(\mu, \Gamma, A, B\right)\right)$$

holds provided that the optimization problem on the right is attained at a (μ_*, Γ_*) where the multiplicity, linear independence and full Jordan block qualifications hold.

2. *Furthermore a minimal perturbation δA_* such that $\sum_{j=1}^{k} m(A + \delta A_*, B) \geq r$ is given by*

$$\delta A_* = -\tau_r(A, B, \mathcal{S}) \mathcal{U} \mathcal{V}^+$$

where the matrices $\mathcal{U}$ and $\mathcal{V}$ are formed by the left and right singular vectors at the optimal (μ_*, Γ_*).

Construction of an Optimal Perturbation

Theorem (Nearest Pencils with Specified Eigenvalues)

Let $A - \lambda B$ be an $m \times n$ pencil with $m \geq n$, r be a positive integer and $\mathcal{S} = \{\lambda_1, \ldots, \lambda_k\}$ be a set of distinct complex scalars.

1. *Then the equality*

$$\tau_r(A, B, \mathcal{S}) = \inf_{\mu \in \mathcal{S}^r} \sup_{\Gamma} \sigma_{nr-r+1}\left(\mathcal{L}\left(\mu, \Gamma, A, B\right)\right)$$

holds provided that the optimization problem on the right is attained at a (μ_*, Γ_*) where the multiplicity, linear independence and full Jordan block qualifications hold.

2. *Furthermore a minimal perturbation δA_* such that $\sum_{j=1}^{k} m(A + \delta A_*, B) \geq r$ is given by*

$$\delta A_* = -\tau_r(A, B, \mathcal{S})\mathcal{U}\mathcal{V}^+$$

where the matrices $\mathcal{U}$ and $\mathcal{V}$ are formed by the left and right singular vectors at the optimal (μ_*, Γ_*).

Outline

Simultaneous Perturbations

When $\alpha_A = \alpha_B = 1$ and $\tau_r(A, B, \mathcal{S})$ defined w.r.t. the Frobenius norm

$$\tau_r(A, B, \mathcal{S}) \geq \inf_{\mu \in \mathcal{S}^r} \sup_{\Gamma} \frac{\sigma_{nr-r+1}\left(\mathcal{L}\left(\mu, \Gamma, A, B\right)\right)}{\sqrt{1 + \|C(\mu, \Gamma)\|_2^2}}$$

Numerical Examples

(1) Distance from $A - \lambda B$ to pencils with multiple eigenvalues

$$\inf_{\lambda \in \mathbb{C}} \sup_{\gamma \in \mathbb{C}} \sigma_{2n-1}\left(\begin{bmatrix} A - \lambda B & 0 \\ \gamma B & A - \lambda B \end{bmatrix}\right)$$

e.g.

$$A - \lambda B = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} -1 & 2 & 3 \\ 2 & -1 & 2 \\ 4 & 2 & -1 \end{bmatrix}$$

The nearest pencil

$$\begin{bmatrix} 1.91465 & -0.57896 & -1.21173 \\ -1.32160 & 1.93256 & -0.57897 \\ -0.72082 & -1.32160 & 1.91466 \end{bmatrix} - \lambda \begin{bmatrix} -1 & 2 & 3 \\ 2 & -1 & 2 \\ 4 & 2 & -1 \end{bmatrix}$$

at a distance of 0.59299 with multiple eigenvalue $\lambda_* = -0.85488$.

Numerical Examples

(1) Distance from $A - \lambda B$ to pencils with multiple eigenvalues

$$\inf_{\lambda \in \mathbb{C}} \sup_{\gamma \in \mathbb{C}} \sigma_{2n-1}\left(\begin{bmatrix} A - \lambda B & 0 \\ \gamma B & A - \lambda B \end{bmatrix} \right)$$

e.g.

$$A - \lambda B = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} -1 & 2 & 3 \\ 2 & -1 & 2 \\ 4 & 2 & -1 \end{bmatrix}$$

The nearest pencil

$$\begin{bmatrix} 1.91465 & -0.57896 & -1.21173 \\ -1.32160 & 1.93256 & -0.57897 \\ -0.72082 & -1.32160 & 1.91466 \end{bmatrix} - \lambda \begin{bmatrix} -1 & 2 & 3 \\ 2 & -1 & 2 \\ 4 & 2 & -1 \end{bmatrix}$$

at a distance of 0.59299 with multiple eigenvalue $\lambda_* = -0.85488$.

Numerical Examples

(1) Distance from $A - \lambda B$ to pencils with multiple eigenvalues

$$\inf_{\lambda \in \mathbb{C}} \sup_{\gamma \in \mathbb{C}} \sigma_{2n-1}\left(\begin{bmatrix} A - \lambda B & 0 \\ \gamma B & A - \lambda B \end{bmatrix} \right)$$

e.g.

$$A - \lambda B = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} -1 & 2 & 3 \\ 2 & -1 & 2 \\ 4 & 2 & -1 \end{bmatrix}$$

The nearest pencil

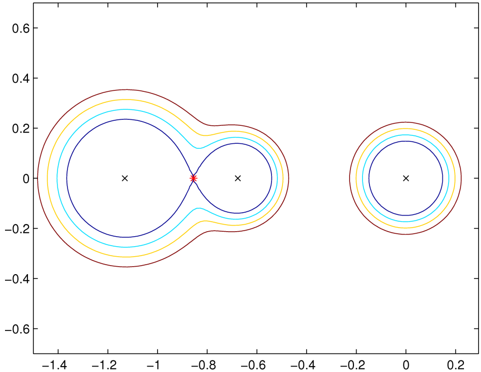
$$\begin{bmatrix} 1.91465 & -0.57896 & -1.21173 \\ -1.32160 & 1.93256 & -0.57897 \\ -0.72082 & -1.32160 & 1.91466 \end{bmatrix} - \lambda \begin{bmatrix} -1 & 2 & 3 \\ 2 & -1 & 2 \\ 4 & 2 & -1 \end{bmatrix}$$

at a distance of 0.59299 with multiple eigenvalue $\lambda_* = -0.85488$.

Numerical Examples

The ϵ-pseudospectrum of $A - \lambda B$

$$\Lambda_\epsilon(A, B) = \{\lambda \in \mathbb{C} : \exists\, (\delta A) \quad \det(A + \delta A - \lambda B) = 0 \ \ \textit{and} \ \ \|\delta A\|_2 \leq \epsilon\}$$

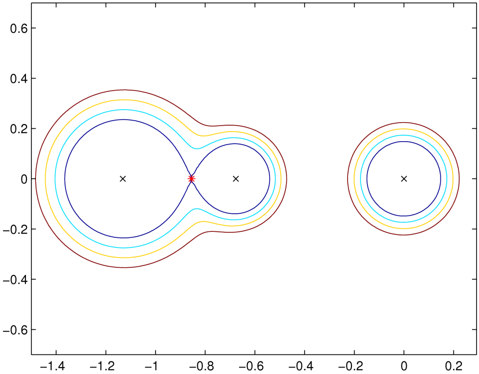


Blue Curve: $\Lambda_\epsilon(A, B)$ for $\epsilon = 0.59299$; Red asterisks marks $\lambda_* = -0.85488$.

Numerical Examples

The ϵ-pseudospectrum of $A - \lambda B$

$$\Lambda_\epsilon(A, B) = \{\lambda \in \mathbb{C} : \exists\, (\delta A) \quad \det(A + \delta A - \lambda B) = 0 \;\; and \;\; \|\delta A\|_2 \leq \epsilon\}$$



Blue Curve: $\Lambda_\epsilon(A, B)$ for $\epsilon = 0.59299$; Red asterisks marks $\lambda_* = -0.85488$.

Numerical Examples

(ii) Distance from rectangular $A - \lambda B$ to pencils with r eigenvalues

$$\inf_{\lambda_j \in \mathbb{C}} \ \sup_{\gamma_{ik} \in \mathbb{C}} \ \sigma_{nr-r+1} \left(\begin{bmatrix} A - \lambda_1 B & 0 & & 0 \\ \gamma_{21} B & A - \lambda_2 B & & 0 \\ & & \ddots & \\ \gamma_{r1} B & \gamma_{r2} B & & A - \lambda_r B \end{bmatrix} \right)$$

e.g.

$$A - \lambda B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.1 & 2 & 1 \\ 0 & 0 & 0.3 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The nearest pencil

$$\begin{bmatrix} 0.99847 & 0 & 0 & 0.00007 \\ -0.03697 & 0.08698 & 2.00172 & 1.00095 \\ -0.01283 & 0.03689 & 0.30078 & 2.00376 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

at a distance of 0.03927 with two eigvals $\lambda_1^* = 2.55144$ and $\lambda_2^* = 1.45405$.

Numerical Examples

(ii) Distance from rectangular $A - \lambda B$ to pencils with r eigenvalues

$$\inf_{\lambda_j \in \mathbb{C}} \ \sup_{\gamma_{ik} \in \mathbb{C}} \ \sigma_{nr-r+1} \left(\begin{bmatrix} A - \lambda_1 B & 0 & & 0 \\ \gamma_{21} B & A - \lambda_2 B & & 0 \\ & & \ddots & \\ \gamma_{r1} B & \gamma_{r2} B & & A - \lambda_r B \end{bmatrix} \right)$$

e.g.

$$A - \lambda B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.1 & 2 & 1 \\ 0 & 0 & 0.3 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The nearest pencil

$$\begin{bmatrix} 0.99847 & 0 & 0 & 0.00007 \\ -0.03697 & 0.08698 & 2.00172 & 1.00095 \\ -0.01283 & 0.03689 & 0.30078 & 2.00376 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

at a distance of 0.03927 with two eigvals $\lambda_1^* = 2.55144$ and $\lambda_2^* = 1.45405$.

Numerical Examples

(ii) Distance from rectangular $A - \lambda B$ to pencils with r eigenvalues

$$\inf_{\lambda_j \in \mathbb{C}} \ \sup_{\gamma_{ik} \in \mathbb{C}} \ \sigma_{nr-r+1} \left(\begin{bmatrix} A - \lambda_1 B & 0 & & 0 \\ \gamma_{21} B & A - \lambda_2 B & & 0 \\ & & \ddots & \\ \gamma_{r1} B & \gamma_{r2} B & & A - \lambda_r B \end{bmatrix} \right)$$

e.g.

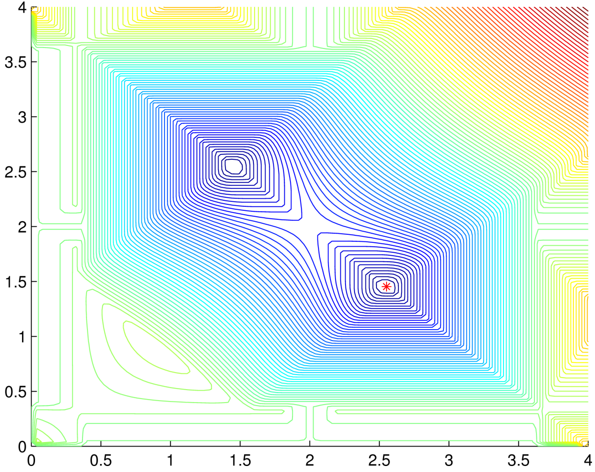
$$A - \lambda B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.1 & 2 & 1 \\ 0 & 0 & 0.3 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The nearest pencil

$$\begin{bmatrix} 0.99847 & 0 & 0 & 0.00007 \\ -0.03697 & 0.08698 & 2.00172 & 1.00095 \\ -0.01283 & 0.03689 & 0.30078 & 2.00376 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

at a distance of 0.03927 with two eigvals $\lambda_1^* = 2.55144$ and $\lambda_2^* = 1.45405$.

Numerical Examples



Level sets of $g(\lambda_1, \lambda_2)$; Red asterisks marks $(\lambda_1^*, \lambda_2^*) = (2.5514, 1.4541)$.

Numerical Examples

(iii) Distance from $A - \lambda B$ to stable pencils

$$\inf_{\lambda_j \in \mathbb{C}^-} \sup_{\gamma_{p\ell} \in \mathbb{C}} \sigma_{nr-r+1}\left(\begin{bmatrix} A - \lambda_1 B & 0 & & 0 \\ \gamma_{21} B & A - \lambda_2 B & & 0 \\ & & \ddots & \\ \gamma_{n1} B & \gamma_{n2} B & & A - \lambda_n B \end{bmatrix} \right)$$

e.g.

$$A = \begin{bmatrix} 0.6 - \frac{i}{3} & -0.2 + \frac{4i}{3} \\ -0.1 + \frac{2i}{3} & 0.5 + \frac{i}{3} \end{bmatrix}$$

is unstable with eigenvalues $0.7 - i$ and $0.4 + i$ at a distance of 0.6610 to stability. The matrix

$$A + \delta A_* = \begin{bmatrix} 0.0681 - 0.3064i & -0.4629 + 1.2524i \\ 0.2047 + 0.5858i & -0.1573 + 0.3064i \end{bmatrix}$$

at a distance of 0.6610 has the eigvals $-0.9547i$ and $-0.0885 + 0.9547i$.

Numerical Examples

(iii) Distance from $A - \lambda B$ to stable pencils

$$\inf_{\lambda_j \in \mathbb{C}^-} \sup_{\gamma_{p\ell} \in \mathbb{C}} \sigma_{nr-r+1}\left(\begin{bmatrix} A - \lambda_1 B & 0 & & 0 \\ \gamma_{21} B & A - \lambda_2 B & & 0 \\ & & \ddots & \\ \gamma_{n1} B & \gamma_{n2} B & & A - \lambda_n B \end{bmatrix} \right)$$

e.g.

$$A = \begin{bmatrix} 0.6 - \frac{i}{3} & -0.2 + \frac{4i}{3} \\ -0.1 + \frac{2i}{3} & 0.5 + \frac{i}{3} \end{bmatrix}$$

is unstable with eigenvalues $0.7 - i$ and $0.4 + i$ at a distance of 0.6610 to stability. The matrix

$$A + \delta A_* = \begin{bmatrix} 0.0681 - 0.3064i & -0.4629 + 1.2524i \\ 0.2047 + 0.5858i & -0.1573 + 0.3064i \end{bmatrix}$$

at a distance of 0.6610 has the eigvals $-0.9547i$ and $-0.0885 + 0.9547i$.

Numerical Examples

(iii) Distance from $A - \lambda B$ to stable pencils

$$\inf_{\lambda_j \in \mathbb{C}^-} \sup_{\gamma_{p\ell} \in \mathbb{C}} \sigma_{nr-r+1} \left(\begin{bmatrix} A - \lambda_1 B & 0 & & 0 \\ \gamma_{21} B & A - \lambda_2 B & & 0 \\ & & \ddots & \\ \gamma_{n1} B & \gamma_{n2} B & & A - \lambda_n B \end{bmatrix} \right)$$

e.g.

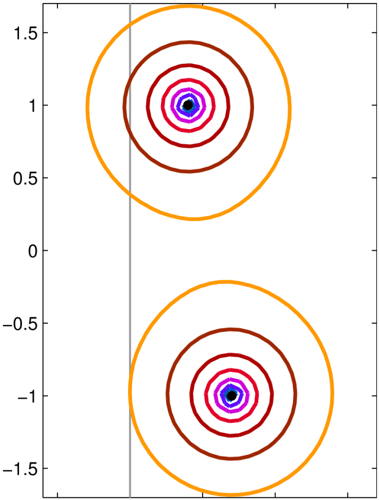
$$A = \begin{bmatrix} 0.6 - \frac{i}{3} & -0.2 + \frac{4i}{3} \\ -0.1 + \frac{2i}{3} & 0.5 + \frac{i}{3} \end{bmatrix}$$

is unstable with eigenvalues $0.7 - i$ and $0.4 + i$ at a distance of 0.6610 to stability. The matrix

$$A + \delta A_* = \begin{bmatrix} 0.0681 - 0.3064i & -0.4629 + 1.2524i \\ 0.2047 + 0.5858i & -0.1573 + 0.3064i \end{bmatrix}$$

at a distance of 0.6610 has the eigvals $-0.9547i$ and $-0.0885 + 0.9547i$.

Numerical Examples



Orange Curve:
$\Lambda_\epsilon(A)$ for $\epsilon = 0.6610$

Summary

- A singular value characterization is derived for $\tau_r(A, B, \mathcal{S})$.
- Charac makes the numerical computation possible.

- Open Problems
  - A singular value charac for $\tau_r(A, B, \mathcal{S})$ when $\alpha_A = \alpha_B = 1$
  - Lipschitz-based computation of $\tau_r(A, B, \mathcal{S})$ efficiently

Summary

- A singular value characterization is derived for $\tau_r(A, B, \mathcal{S})$.
- Charac makes the numerical computation possible.

- Open Problems
  - A singular value charac for $\tau_r(A, B, \mathcal{S})$ when $\alpha_A = \alpha_B = 1$
  - Lipschitz-based computation of $\tau_r(A, B, \mathcal{S})$ efficiently

Summary

- A singular value characterization is derived for $\tau_r(A, B, \mathcal{S})$.
- Charac makes the numerical computation possible.

- Open Problems
  - A singular value charac for $\tau_r(A, B, \mathcal{S})$ when $\alpha_A = \alpha_B = 1$
  - Lipschitz-based computation of $\tau_r(A, B, \mathcal{S})$ efficiently

Summary

- A singular value characterization is derived for $\tau_r(A, B, \mathcal{S})$.
- Charac makes the numerical computation possible.

- Open Problems
  - A singular value charac for $\tau_r(A, B, \mathcal{S})$ when $\alpha_A = \alpha_B = 1$
  - Lipschitz-based computation of $\tau_r(A, B, \mathcal{S})$ efficiently

Summary

- A singular value characterization is derived for $\tau_r(A, B, \mathcal{S})$.
- Charac makes the numerical computation possible.

- Open Problems
  - A singular value charac for $\tau_r(A, B, \mathcal{S})$ when $\alpha_A = \alpha_B = 1$
  - Lipschitz-based computation of $\tau_r(A, B, \mathcal{S})$ efficiently

References

- A.N. Malyshev, A formula for the 2-norm distance from a matrix to the set of matrices with multiple eigenvalues, *Numer Math*, 83:443-454, 1999
- G. Boutry, M. Elad, G.H. Golub and P. Milanfar, Generalized eigenvalue problem for nonsquare pencils using a minimal perturbation approach, *SIAM J Matrix Anal App*, 27(2):582-601, 2005
- E. Mengi, Locating a nearest matrix with an eigenvalue of prespecified algebraic multiplicity, *Numer Math*, 118:109-135, 2011
- D. Kressner, E. Mengi, I. Nakic and N. Truhar, Generalized eigenvalue problems with specified eigenvalues, *math arXiv*