Variational Principles for Nonlinear Eigenvalue Problems

Heinrich Voss
voss@tuhh.de

Hamburg University of Technology
Institute of Mathematics
For $\lambda \in J \subset \mathbb{R}$, $J$ an interval, let $T(\lambda)$ be a linear self-adjoint and bounded operator on a Hilbert space $\mathcal{H}$. Find $\lambda \in J$ and $x \neq 0$ such that $T(\lambda)x = 0$. Then $\lambda$ is called an eigenvalue of $T(\lambda)$, and $x$ a corresponding eigenvalue.
For $\lambda \in J \subset \mathbb{R}$, $J$ an interval, let $T(\lambda)$ be a linear self-adjoint and bounded operator on a Hilbert space $\mathcal{H}$.

Find $\lambda \in J$ and $x \neq 0$ such that

$$T(\lambda)x = 0.$$  \hfill (1)

Then $\lambda$ is called an eigenvalue of $T(\cdot)$, and $x$ a corresponding eigenelement.
For $\lambda \in J \subset \mathbb{R}$, $J$ an interval, let $T(\lambda)$ be a linear self-adjoint and bounded operator on a Hilbert space $\mathcal{H}$.

Find $\lambda \in J$ and $x \neq 0$ such that

$$T(\lambda)x = 0. \tag{1}$$

Then $\lambda$ is called an eigenvalue of $T(\cdot)$, and $x$ a corresponding eigenelement.

Nonlinear eigenproblems of this type arise in
- dynamic/stability analysis of structures and in fluid mechanics
- vibration of fluid-solid structures
- vibration of sandwich plates
- accelerator design
- vibro-acoustics of piezoelectric/poroelastic structures
- nonlinear integrated optics
- regularization of total least squares problems
- stability of delay differential equations
Introduction

For $\lambda \in J \subset \mathbb{R}$, $J$ an interval, let $T(\lambda)$ be a linear self-adjoint and bounded operator on a Hilbert space $\mathcal{H}$.

Find $\lambda \in J$ and $x \neq 0$ such that

$$T(\lambda)x = 0.$$  \hspace{1cm} (1)

Then $\lambda$ is called an eigenvalue of $T(\cdot)$, and $x$ a corresponding eigenelement.

Nonlinear eigenproblems of this type arise in

- dynamic/stability analysis of structures and in fluid mechanics
- electronic behavior of semiconductor hetero-structures

Nonlinear integrated optics
- regularization of total least squares problems
- stability of delay differential equations
For $\lambda \in J \subset \mathbb{R}$, $J$ an interval, let $T(\lambda)$ be a linear self-adjoint and bounded operator on a Hilbert space $\mathcal{H}$.

Find $\lambda \in J$ and $x \neq 0$ such that

$$T(\lambda)x = 0. \quad (1)$$

Then $\lambda$ is called an eigenvalue of $T(\cdot)$, and $x$ a corresponding eigenelement.

Nonlinear eigenproblems of this type arise in

- dynamic/stability analysis of structures and in fluid mechanics
- electronic behavior of semiconductor hetero-structures
- vibration of fluid-solid structures
For \( \lambda \in J \subset \mathbb{R} \), \( J \) an interval, let \( T(\lambda) \) be a linear self-adjoint and bounded operator on a Hilbert space \( \mathcal{H} \).

Find \( \lambda \in J \) and \( x \neq 0 \) such that

\[
T(\lambda)x = 0. \tag{1}
\]

Then \( \lambda \) is called an eigenvalue of \( T(\cdot) \), and \( x \) a corresponding eigenelement.

Nonlinear eigenproblems of this type arise in
- dynamic/stability analysis of structures and in fluid mechanics
- electronic behavior of semiconductor hetero-structures
- vibration of fluid-solid structures
- vibration of sandwich plates
For $\lambda \in J \subset \mathbb{R}$, $J$ an interval, let $T(\lambda)$ be a linear self-adjoint and bounded operator on a Hilbert space $\mathcal{H}$.

Find $\lambda \in J$ and $x \neq 0$ such that

$$T(\lambda)x = 0.$$ (1)

Then $\lambda$ is called an eigenvalue of $T(\cdot)$, and $x$ a corresponding eigenelement.

Nonlinear eigenproblems of this type arise in
- dynamic/stability analysis of structures and in fluid mechanics
- electronic behavior of semiconductor hetero-structures
- vibration of fluid-solid structures
- vibration of sandwich plates
- accelerator design
For $\lambda \in J \subseteq \mathbb{R}$, $J$ an interval, let $T(\lambda)$ be a linear self-adjoint and bounded operator on a Hilbert space $\mathcal{H}$.

Find $\lambda \in J$ and $x \neq 0$ such that

$$T(\lambda)x = 0. \quad (1)$$

Then $\lambda$ is called an eigenvalue of $T(\cdot)$, and $x$ a corresponding eigenelement.

Nonlinear eigenproblems of this type arise in

- dynamic/stability analysis of structures and in fluid mechanics
- electronic behavior of semiconductor hetero-structures
- vibration of fluid-solid structures
- vibration of sandwich plates
- accelerator design
- vibro-acoustics of piezoelectric/poroelastic structures
For $\lambda \in J \subset \mathbb{R}$, $J$ an interval, let $T(\lambda)$ be a linear self-adjoint and bounded operator on a Hilbert space $\mathcal{H}$.

Find $\lambda \in J$ and $x \neq 0$ such that

$$T(\lambda)x = 0. \quad (1)$$

Then $\lambda$ is called an eigenvalue of $T(\cdot)$, and $x$ a corresponding eigenelement.

Nonlinear eigenproblems of this type arise in
- dynamic/stability analysis of structures and in fluid mechanics
- electronic behavior of semiconductor hetero-structures
- vibration of fluid-solid structures
- vibration of sandwich plates
- accelerator design
- vibro-acoustics of piezoelectric/poroelastic structures
- nonlinear integrated optics
For $\lambda \in J \subset \mathbb{R}$, $J$ an interval, let $T(\lambda)$ be a linear self-adjoint and bounded operator on a Hilbert space $\mathcal{H}$.

Find $\lambda \in J$ and $x \neq 0$ such that

$$T(\lambda)x = 0. \quad (1)$$

Then $\lambda$ is called an eigenvalue of $T(\cdot)$, and $x$ a corresponding eigenelement.

Nonlinear eigenproblems of this type arise in
- dynamic/stability analysis of structures and in fluid mechanics
- electronic behavior of semiconductor hetero-structures
- vibration of fluid-solid structures
- vibration of sandwich plates
- accelerator design
- vibro-acoustics of piezoelectric/poroelastic structures
- nonlinear integrated optics
- regularization of total least squares problems
For $\lambda \in J \subset \mathbb{R}$, $J$ an interval, let $T(\lambda)$ be a linear self-adjoint and bounded operator on a Hilbert space $\mathcal{H}$.

Find $\lambda \in J$ and $x \neq 0$ such that

$$T(\lambda)x = 0. \quad (1)$$

Then $\lambda$ is called an eigenvalue of $T(\cdot)$, and $x$ a corresponding eigenelement.

Nonlinear eigenproblems of this type arise in

- dynamic/stability analysis of structures and in fluid mechanics
- electronic behavior of semiconductor hetero-structures
- vibration of fluid-solid structures
- vibration of sandwich plates
- accelerator design
- vibro-acoustics of piezoelectric/poroelastic structures
- nonlinear integrated optics
- regularization of total least squares problems
- stability of delay differential equations
Outline

1 Two examples

2 Variational characterization of eigenvalue problems
   - Overdamped Problems
   - Nonoverdamped problems

3 Applications
   - Sylvester’s law of inertia
   - Nonlinear low rank modification of a symmetric eigenvalue problem
   - Safeguarded iteration
   - Detecting hyperbolic matrix polynomials
   - Free vibrations of fluid-solid structures
   - Regularization of total least squares problems

4 Concluding remarks
Outline

1. Two examples

2. Variational characterization of eigenvalue problems
   - Overdamped Problems
   - Nonoverdamped problems

3. Applications
   - Sylvester’s law of inertia
   - Nonlinear low rank modification of a symmetric eigenvalue problem
   - Safeguarded iteration
   - Detecting hyperbolic matrix polynomials
   - Free vibrations of fluid-solid structures
   - Regularization of total least squares problems

4. Concluding remarks
Example 1: Vibrations of a rotating tire

A great deal of the overall sound source of road traffic is caused by rolling noise of road vehicles. For passenger cars at speeds above 40 km/h, and for trucks above 60 km/h the major source of traffic noise is due to the sound radiation of rolling tires.
Example 1: Vibrations of a rotating tire

A great deal of the overall sound source of road traffic is caused by rolling noise of road vehicles. For passenger cars at speeds above 40 km/h, and for trucks above 60 km/h the major source of traffic noise is due to the sound radiation of rolling tires.

Small vibrations of a rotating tire are governed by a gyroscopic quadratic eigenvalue problem

\[ Q(\omega)x = Kx + i\omega Gx - Mx = 0, \]

where the stiffness matrix \( K = K^T > 0 \) and the mass matrix \( M = M^T > 0 \) are positive definite and in the gyroscopic term (stemming from the Coriolis force) the matrix \( G = -G^T \) is skew-symmetric.
Two examples

Example 1: Vibrations of a rotating tire

A great deal of the overall sound source of road traffic is caused by rolling noise of road vehicles. For passenger cars at speeds above 40 km/h, and for trucks above 60 km/h the major source of traffic noise is due to the sound radiation of rolling tires.

Small vibrations of a rotating tire are governed by a gyroscopic quadratic eigenvalue problem

\[ Q(\omega)x = Kx + i\omega Gx - Mx = 0, \]

where the stiffness matrix \( K = K^T > 0 \) and the mass matrix \( M = M^T > 0 \) are positive definite and in the gyroscopic term (stemming from the Coriolis force) the matrix \( G = -G^T \) is skew-symmetric.

\[ Q(\omega)x = 0 \]

will be shown to have \( 2n \) real eigenvalues, \( n \) positive and \( n \) negative.
Example 2: Free vibrations of fluid-solid structures

can be modelled in terms of solid displacement and fluid pressure and one obtains the classical form of an eigenproblem

\[
\begin{align*}
\text{div } [\sigma(u)] + \omega^2 \rho_s u &= 0 \text{ in } \Omega_s, \\
\Delta p + \frac{\omega^2}{c^2} p &= 0 \text{ in } \Omega_f, \\
\sigma(u) \cdot n - pn &= 0 \text{ on } \Gamma_I, \\
\nabla p \cdot n + \omega^2 \rho_f u \cdot n &= 0 \text{ on } \Gamma_I, \\
u &= 0 \text{ on } \Gamma_D, \\
\nabla p \cdot n &= 0 \text{ on } \Gamma_N,
\end{align*}
\]

- \(u\): solid displacement
- \(p\): fluid pressure
- \(\lambda = \omega^2\): eigenparameter
- \(\sigma(u)\): linearized stress tensor
- \(\rho_s, \rho_f\): densities of solid and fluid

Interface conditions: equilibrium of accelerations and of force densities.
Find $\lambda := \omega^2 \in \mathbb{C}$ and $(u, p) \in H^1_{\Gamma_D}(\Omega_s)^3 \times H^1(\Omega_f)$ such that

$$a_s(v, u) + c(v, p) = \lambda b_s(v, u) \quad \text{and} \quad a_f(q, p) = \lambda(-c(u, q) + b_f(q, p)).$$

for every $(v, q) \in H^1_{\Gamma_D}(\Omega_s)^3 \times H^1(\Omega_f)$. 

which (using the Lax-Milgram Lemma) can be transformed into a linear (but not self-adjoint) eigenvalue problem

$$K_s u + C p = \lambda M_s u$$

$$K_f p = \lambda(-C' u + M_f p)$$

where $K_s : H^1_{\Gamma_D}(\Omega_s)^3 \to H^1_{\Gamma_D}(\Omega_s)^3$ is self-adjoint, elliptic, bounded, ...
Find \( \lambda := \omega^2 \in \mathbb{C} \) and \((u, p) \in H^1_{\Gamma_D}(\Omega_s)^3 \times H^1(\Omega_f)\) such that

\[
a_s(v, u) + c(v, p) = \lambda b_s(v, u) \quad \text{and} \quad a_f(q, p) = \lambda (-c(u, q) + b_f(q, p)).
\]

for every \((v, q) \in H^1_{\Gamma_D}(\Omega_s)^3 \times H^1(\Omega_f)\).

which (using the Lax-Milgram Lemma) can be transformed into a linear (but not self-adjoint) eigenvalue problem

\[
\begin{align*}
K_su + Cp &= \lambda M_su \\
K_fp &= \lambda (-C'u + M_fp)
\end{align*}
\]

(2a) (2b)

where \(K_s : H^1_{\Gamma_D}(\Omega_s)^3 \rightarrow H^1_{\Gamma_D}(\Omega_s)^3\) is self-adjoint, elliptic, bounded, \ldots
Let \( 0 < \sigma_1 \leq \sigma_2 \leq \ldots \) denote the eigenvalues of the decoupled eigenproblem

\[
K_S u = \sigma M_S u
\]

and denote by \( u_1, u_2, \ldots \) corresponding orthonormal eigenfunctions. Then the spectral theorem yields

\[
(K_S - \lambda M_S)^{-1} u = \sum_{n=1}^{\infty} \frac{1}{\sigma_n - \lambda} \langle u_n, u \rangle u_n.
\]
Let $0 < \sigma_1 \leq \sigma_2 \leq \ldots$ denote the eigenvalues of the decoupled eigenproblem

$$K_S u = \sigma M_S u$$

and denote by $u_1, u_2, \ldots$ corresponding orthonormal eigenfunctions. Then the spectral theorem yields

$$(K_S - \lambda M_S)^{-1} u = \sum_{n=1}^{\infty} \frac{1}{\sigma_n - \lambda} \langle u_n, u \rangle u_n.$$

If $\lambda$ is not contained in the spectrum of the decoupled solid eigenproblem, then $\lambda$ is an eigenvalue of the coupled fluid-solid problem if and only if it is an eigenvalue of the rational eigenvalue problem

$$T(\lambda)p := -K_f p + \lambda M_f p + \sum_{n=1}^{\infty} \frac{\lambda}{\sigma_n - \lambda} C_n p, \quad C_n p := \langle u_n, C p \rangle C' u_n.$$
Let $0 < \sigma_1 \leq \sigma_2 \leq \ldots$ denote the eigenvalues of the decoupled eigenproblem

$$K_S u = \sigma M_S u$$

and denote by $u_1, u_2, \ldots$ corresponding orthonormal eigenfunctions. Then the spectral theorem yields

$$(K_S - \lambda M_S)^{-1} u = \sum_{n=1}^{\infty} \frac{1}{\sigma_n - \lambda} \langle u_n, u \rangle u_n.$$ 

If $\lambda$ is not contained in the spectrum of the decoupled solid eigenproblem, then $\lambda$ is an eigenvalue of the coupled fluid-solid problem if and only if it is an eigenvalue of the rational eigenvalue problem

$$T(\lambda)p := -K_f p + \lambda M_f p + \sum_{n=1}^{\infty} \frac{\lambda}{\sigma_n - \lambda} C_n p, \quad C_n p := \langle u_n, C p \rangle C' u_n.$$ 

$T(\lambda) : H^1(\Omega_f) \to H^1(\Omega_f)$ is self-adjoint and bounded.
Another self-adjoint form of the fluid-solid eigenproblem is obtained if the second equation in (2) is multiplied by $\omega$ and $p$ is substituted by $p := \omega w$. 
Another self-adjoint form of the fluid-solid eigenproblem is obtained if the second equation in (2) is multiplied by $\omega$ and $p$ is substituted by $p =: \omega w$.

Then problem (2) is equivalent to the quadratic eigenvalue problem

$$
\left( \begin{pmatrix} K_s & 0 \\ 0 & K_f \end{pmatrix} + \omega \begin{pmatrix} O & C \\ C' & O \end{pmatrix} - \omega^2 \begin{pmatrix} M_s & O \\ O & M_f \end{pmatrix} \right) \begin{pmatrix} u \\ w \end{pmatrix} = 0.
$$
Outline

1 Two examples

2 Variational characterization of eigenvalue problems
   - Overdamped Problems
   - Nonoverdamped problems

3 Applications
   - Sylvester’s law of inertia
   - Nonlinear low rank modification of a symmetric eigenvalue problem
   - Safeguarded iteration
   - Detecting hyperbolic matrix polynomials
   - Free vibrations of fluid-solid structures
   - Regularization of total least squares problems

4 Concluding remarks
Outline

1. Two examples

2. Variational characterization of eigenvalue problems
   - Overdamped Problems
   - Nonoverdamped problems

3. Applications
   - Sylvester’s law of inertia
   - Nonlinear low rank modification of a symmetric eigenvalue problem
   - Safeguarded iteration
   - Detecting hyperbolic matrix polynomials
   - Free vibrations of fluid-solid structures
   - Regularization of total least squares problems

4. Concluding remarks
Variational characterization for linear eigenproblems

Let $A : \mathcal{H} \to \mathcal{H}$ a bounded linear and self-adjoint operator in a Hilbert space $\mathcal{H}$. Then those eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots$ above the essential spectrum of $A$ (if there are any) can be characterized by three fundamental variational principles,

1. Rayleigh's principle (1873)
   \[ \lambda_n = \max_{\{x : \langle x, x \rangle = 0, i = 1, \ldots, n-1\}} \{ R(x) : \langle Ax, x \rangle \langle x, x \rangle \} \]

2. the maxmin characterization due to Poincaré (1890)
   \[ \lambda_n = \max_{\dim V = n} \min_{x \in V, x \neq 0} R(x) \]

3. the minmax characterization due to Courant (1920), Fischer (1905), and Weyl (1912).
   \[ \lambda_n = \min_{\dim V = n-1} \max_{x \in V^\perp, x \neq 0} R(x) \]

Variational characterizations are very powerful tools when studying self-adjoint linear operators on a Hilbert space $\mathcal{H}$. Bounds for eigenvalues, comparison theorems, interlacing results and monotonicity of eigenvalues can be proved easily with these characterizations, to name just a few.
Let $A : \mathcal{H} \to \mathcal{H}$ a bounded linear and self-adjoint operator in a Hilbert space $\mathcal{H}$. Then those eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots$ above the essential spectrum of $A$ (if there are any) can be characterized by three fundamental variational principles, by

- Rayleigh’s principle (1873)

$$\lambda_n = \max \{ R(x) : \langle x, x_i \rangle = 0, \ i = 1, \ldots, n - 1 \}, \ R(x) := \frac{\langle Ax, x \rangle}{\langle x, x \rangle}. $$
Let $A : \mathcal{H} \to \mathcal{H}$ be a bounded linear and self-adjoint operator in a Hilbert space $\mathcal{H}$. Then those eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots$ above the essential spectrum of $A$ (if there are any) can be characterized by three fundamental variational principles:

- **Rayleigh’s principle (1873)**

  $$\lambda_n = \max \{ R(x) : \langle x, x_i \rangle = 0, \ i = 1, \ldots, n-1 \}, \ R(x) := \frac{\langle Ax, x \rangle}{\langle x, x \rangle}.$$

- **The maxmin characterization due to Poincaré (1890)**

  $$\lambda_n = \max_{\dim V = n} \min_{x \in V, x \neq 0} R(x).$$
Variational characterization for linear eigenproblems

Let $A : \mathcal{H} \to \mathcal{H}$ a bounded linear and self-adjoint operator in a Hilbert space $\mathcal{H}$. Then those eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots$ above the essential spectrum of $A$ (if there are any) can be characterized by three fundamental variational principles, by

- Rayleigh’s principle (1873)
  $$\lambda_n = \max \{ R(x) : \langle x, x_i \rangle = 0, \ i = 1, \ldots, n-1 \}, \ R(x) := \frac{\langle Ax, x \rangle}{\langle x, x \rangle}.$$ 
- the maxmin characterization due to Poincaré (1890)
  $$\lambda_n = \max_{\dim V = n} \min_{x \in V, x \neq 0} R(x).$$ 
- the minmax characterization due to Courant (1920), Fischer (1905), and Weyl (1912).
  $$\lambda_n = \min_{\dim V = n-1} \max_{x \in V \perp, x \neq 0} R(x).$$
Variational characterization for linear eigenproblems

Let \( A : \mathcal{H} \to \mathcal{H} \) a bounded linear and self-adjoint operator in a Hilbert space \( \mathcal{H} \). Then those eigenvalues \( \lambda_1 \geq \lambda_2 \geq \ldots \) above the essential spectrum of \( A \) (if there are any) can be characterized by three fundamental variational principles, by

- Rayleigh’s principle (1873)

\[
\lambda_n = \max \left\{ R(x) : \langle x, x_i \rangle = 0, \ i = 1, \ldots, n - 1 \right\}, \ R(x) := \frac{\langle Ax, x \rangle}{\langle x, x \rangle}.
\]

- the maxmin characterization due to Poincaré (1890)

\[
\lambda_n = \max_{\dim V = n} \min_{x \in V, \ x \neq 0} R(x).
\]

- the minmax characterization due to Courant (1920), Fischer (1905), and Weyl (1912).

\[
\lambda_n = \min_{\dim V = n - 1} \max_{x \in V^\perp, \ x \neq 0} R(x).
\]

Variational characterizations are very powerful tools when studying self-adjoint linear operators on a Hilbert space \( \mathcal{H} \). Bounds for eigenvalues, comparison theorems, interlacing results and monotonicity of eigenvalues can be proved easily with these characterizations, to name just a few.
Rayleigh functional

Let

$$f : \begin{cases} \mathbb{J} \times \mathcal{H} & \to \mathbb{R} \\ (\lambda, x) & \mapsto \langle T(\lambda)x, x \rangle \end{cases}$$

be continuous, and assume that for every fixed $x \in \mathcal{H}$, $x \neq 0$, the real equation

$$f(\lambda, x) = 0$$

has at most one solution in $\mathbb{J}$. 

(3)
Rayleigh functional

Let

\[ f : \left\{ \begin{array}{c} J \times \mathcal{H} \\ (\lambda, x) \end{array} \rightarrow \mathbb{R} \\ \langle T(\lambda)x, x \rangle \right\} \]

be continuous, and assume that for every fixed \( x \in \mathcal{H}, x \neq 0 \), the real equation

\[ f(\lambda, x) = 0 \quad (3) \]

has at most one solution in \( J \).

Then equation (3) implicitly defines a functional \( p \) on some subset \( D \) of \( \mathcal{H} \setminus \{0\} \) which we call the Rayleigh functional, and which is exactly the Rayleigh quotient in case of a linear eigenproblem \( T(\lambda) = \lambda I - A \).
Let
\[ f : \begin{cases} J \times \mathcal{H} &\rightarrow \mathbb{R} \\ (\lambda, x) &\mapsto \langle T(\lambda)x, x \rangle \end{cases} \]
be continuous, and assume that for every fixed \( x \in \mathcal{H}, x \neq 0 \), the real equation
\[ f(\lambda, x) = 0 \]
has at most one solution in \( J \).

Then equation (3) implicitly defines a functional \( p \) on some subset \( \mathcal{D} \) of \( \mathcal{H} \setminus \{0\} \) which we call the Rayleigh functional, and which is exactly the Rayleigh quotient in case of a linear eigenproblem \( T(\lambda) = \lambda I - A \).

We further assume that for every \( x \in \mathcal{D}, x \neq 0 \) and \( \lambda \in J, \lambda \neq p(x) \) it holds that
\[ f(\lambda, x)(\lambda - p(x)) > 0 \]
which generalizes the definiteness of the operator \( B \) for the generalized linear eigenproblem \( T(\lambda) := \lambda B - A \).
Overdamped problems

If the Rayleigh functional \( p \) is defined on the entire space \( \mathcal{H} \setminus \{0\} \) then the eigenproblem \( T(\lambda)x = 0 \) is called **overdamped**.
Overdamped problems

If the Rayleigh functional $p$ is defined on the entire space $\mathcal{H}\setminus\{0\}$ then the eigenproblem $T(\lambda)x = 0$ is called **overdamped**.

This notation is motivated by the finite dimensional quadratic eigenvalue problem

$$T(\lambda)x = \lambda^2 Mx + \lambda \alpha Cx + Kx = 0$$

where $M$, $C$ and $K$ are symmetric and positive definite matrices.
Overdamped problems

If the Rayleigh functional $p$ is defined on the entire space $\mathcal{H} \setminus \{0\}$ then the eigenproblem $T(\lambda)x = 0$ is called overdamped.

This notation is motivated by the finite dimensional quadratic eigenvalue problem

$$T(\lambda)x = \lambda^2 Mx + \lambda \alpha Cx + Kx = 0$$

where $M$, $C$ and $K$ are symmetric and positive definite matrices.

$\alpha = 0$ all eigenvalues on imaginary axis

$\alpha > 0$ eigenvalues go into left half plane as conjugate complex pairs all eigenvalues going to the left are smaller than all eigenvalues going to the right system is overdamped
Overdamped problems

If the Rayleigh functional $p$ is defined on the entire space $\mathcal{H} \setminus \{0\}$ then the eigenproblem $T(\lambda)x = 0$ is called *overdamped*.

This notation is motivated by the finite dimensional quadratic eigenvalue problem

$$T(\lambda)x = \lambda^2 Mx + \lambda \alpha Cx + Kx = 0$$

where $M$, $C$ and $K$ are symmetric and positive definite matrices.

- $\alpha = 0$: all eigenvalues on imaginary axis
- increase $\alpha$: eigenvalues go into left half plane as conjugate complex pairs
- increase $\alpha$: all eigenvalues going to the left are smaller than all eigenvalues going to the right
- system is overdamped
Overdamped problems

If the Rayleigh functional $p$ is defined on the entire space $\mathcal{H} \setminus \{0\}$ then the eigenproblem $T(\lambda)x = 0$ is called overdamped.

This notation is motivated by the finite dimensional quadratic eigenvalue problem

$$T(\lambda)x = \lambda^2 Mx + \lambda \alpha CX + Kx = 0$$

where $M$, $C$ and $K$ are symmetric and positive definite matrices.

- $\alpha = 0$  all eigenvalues on imaginary axis
- increase $\alpha$ eigenvalues go into left half plane as conjugate complex pairs
- increase $\alpha$ complex pairs reach real axis, run in opposite directions
Overdamped problems

If the Rayleigh functional $p$ is defined on the entire space $\mathcal{H} \setminus \{0\}$ then the eigenproblem $T(\lambda)x = 0$ is called **overdamped**.

This notation is motivated by the finite dimensional quadratic eigenvalue problem

$$T(\lambda)x = \lambda^2 Mx + \lambda \alpha Cx + Kx = 0$$

where $M$, $C$ and $K$ are symmetric and positive definite matrices.

- $\alpha = 0$ all eigenvalues on imaginary axis
- increase $\alpha$ eigenvalues go into left half plane as conjugate complex pairs
- increase $\alpha$ complex pairs reach real axis, run in opposite directions
- increase $\alpha$ all eigenvalues on the negative real axis
Overdamped problems

If the Rayleigh functional $p$ is defined on the entire space $\mathcal{H} \setminus \{0\}$ then the eigenproblem $T(\lambda)x = 0$ is called over-damped.

This notation is motivated by the finite dimensional quadratic eigenvalue problem

$$T(\lambda)x = \lambda^2 Mx + \lambda \alpha Cx + Kx = 0$$

where $M$, $C$ and $K$ are symmetric and positive definite matrices.

- $\alpha = 0$: all eigenvalues on imaginary axis
- increase $\alpha$: eigenvalues go into left half plane as conjugate complex pairs
- increase $\alpha$: complex pairs reach real axis, run in opposite directions
- increase $\alpha$: all eigenvalues on the negative real axis
- increase $\alpha$: all eigenvalues going to the left are smaller than all eigenvalues going to the right
- system is over-damped
Quadratic overdamped problems

For quadratic overdamped systems the two solutions

\[ p_{\pm}(x) = \frac{1}{2} \left( -\alpha \langle Cx, x \rangle \pm \sqrt{\alpha^2 \langle Cx, x \rangle^2 - 4 \langle Mx, x \rangle \langle Kx, x \rangle} \right) / \langle Mx, x \rangle. \]

of the quadratic equation

\[ \langle T(\lambda)x, x \rangle = \lambda^2 \langle Mx, x \rangle + \lambda \alpha \langle Cx, x \rangle + \langle Kx, x \rangle = 0 \]  

are real, and they satisfy \( \sup_{x \neq 0} p_-(x) < \inf_{x \neq 0} p_+(x) \).
Quadratic overdamped problems

For quadratic overdamped systems the two solutions
\[ p_\pm(x) = \frac{1}{2} \left( -\alpha \langle Cx, x \rangle \pm \sqrt{\alpha^2 \langle Cx, x \rangle^2 - 4 \langle Mx, x \rangle \langle Kx, x \rangle} \right) / \langle Mx, x \rangle. \]
of the quadratic equation
\[ \langle T(\lambda)x, x \rangle = \lambda^2 \langle Mx, x \rangle + \lambda \alpha \langle Cx, x \rangle + \langle Kx, x \rangle = 0 \quad (4) \]
are real, and they satisfy \( \sup_{x \neq 0} p_-(x) < \inf_{x \neq 0} p_+(x) \).

Hence, equation (4) defines two Rayleigh functionals \( p_- \) and \( p_+ \) corresponding to the intervals
\[ J_- := (-\infty, \inf_{x \neq 0} p_+(x)) \quad \text{and} \quad J_+ := (\sup_{x \neq 0} p_-(x), \infty). \]
For quadratic overdamped systems the two solutions
\[ p_{\pm}(x) = \frac{1}{2} \left( -\alpha \langle Cx, x \rangle \pm \sqrt{\alpha^2 \langle Cx, x \rangle^2 - 4 \langle Mx, x \rangle \langle Kx, x \rangle} \right) / \langle Mx, x \rangle. \]
of the quadratic equation
\[ \langle T(\lambda) x, x \rangle = \lambda^2 \langle Mx, x \rangle + \lambda \alpha \langle Cx, x \rangle + \langle Kx, x \rangle = 0 \] (4)
are real, and they satisfy \( \sup_{x \neq 0} p_-(x) < \inf_{x \neq 0} p_+(x) \).

Hence, equation (4) defines two Rayleigh functionals \( p_- \) and \( p_+ \) corresponding to the intervals
\[ J_- := (-\infty, \inf_{x \neq 0} p_+(x)) \quad \text{and} \quad J_+ := (\sup_{x \neq 0} p_-(x), \infty). \]

Notice, however, that for \( J_- \) the operator \( T(\cdot) \) has to replaced by \( -T(\cdot) \) to satisfy the second condition.
Rayleigh’s principle

For general (not necessarily quadratic) overdamped problems Hadeler (1967 for the finite dimensional case, and 1968 for \( \dim \mathcal{H} = \infty \)) generalized Rayleigh’s principle proving that the eigenvectors are orthogonal with respect to the generalized scalar product

\[
[x, y] := \begin{cases} 
\langle \frac{T(p(x)) - T(p(y))}{p(x) - p(y)} x, y \rangle, & \text{if } p(x) \neq p(y) \\
\langle T'(p(x))x, y \rangle, & \text{if } p(x) = p(y)
\end{cases}
\]  

(5)

which is symmetric, definite and homogeneous, but in general is not bilinear.
Rayleigh’s principle (Hadeler 1967, 1968)

Let $T(\lambda) : \mathcal{H} \to \mathcal{H}$, $\lambda \in J$ be a family of linear self-adjoint and bounded operators such that (1) is over-damped, and assume that for $\lambda \in J$ there exists $\nu(\lambda) > 0$ such that $T(\lambda) + \nu(\lambda)I$ is compact.
Rayleigh’s principle (Haderel 1967,1968)

Let \( T(\lambda) : \mathcal{H} \to \mathcal{H} \), \( \lambda \in J \) be a family of linear self-adjoint and bounded operators such that (1) is over-damped, and assume that for \( \lambda \in J \) there exists \( \nu(\lambda) > 0 \) such that \( T(\lambda) + \nu(\lambda)I \) is compact.

Let \( T(\cdot) \) be continuously differentiable and suppose that

\[
\frac{\partial}{\partial \lambda} \langle T(\lambda)x, x \rangle \bigg|_{\lambda=p(x)} > 0 \quad \text{for ever } x \neq 0.
\]
Rayleigh’s principle (Haderler 1967, 1968)

Let $T(\lambda) : \mathcal{H} \to \mathcal{H}$, $\lambda \in J$ be a family of linear self-adjoint and bounded operators such that (1) is over-damped, and assume that for $\lambda \in J$ there exists $\nu(\lambda) > 0$ such that $T(\lambda) + \nu(\lambda)I$ is compact.

Let $T(\cdot)$ be continuously differentiable and suppose that

$$\frac{\partial}{\partial \lambda} \langle T(\lambda)x, x \rangle \bigg|_{\lambda = \rho(x)} > 0 \quad \text{for every } x \neq 0.$$

Then problem $T(\lambda)x = 0$ has at most a countable set of eigenvalues in $J$ which we assume to be ordered by magnitude $\lambda_1 \leq \lambda_2 \leq \ldots$ and where each eigenvalue is counted according to its multiplicity.
Rayleigh’s principle (Haderler 1967,1968)

Let $T(\lambda) : \mathcal{H} \to \mathcal{H}$, $\lambda \in J$ be a family of linear self-adjoint and bounded operators such that (1) is over-damped, and assume that for $\lambda \in J$ there exists $\nu(\lambda) > 0$ such that $T(\lambda) + \nu(\lambda)I$ is compact.

Let $T(\cdot)$ be continuously differentiable and suppose that

$$\frac{\partial}{\partial \lambda} \langle T(\lambda)x, x \rangle \bigg|_{\lambda = \rho(x)} > 0 \quad \text{for every } x \neq 0.$$

Then problem $T(\lambda)x = 0$ has at most a countable set of eigenvalues in $J$ which we assume to be ordered by magnitude $\lambda_1 \leq \lambda_2 \leq \ldots$ and where each eigenvalue is counted according to its multiplicity.

The corresponding eigenvectors $x_1, x_2, \ldots$ can be chosen orthonormally with respect to the generalized scalar product (5), and the eigenvalues can be determined recurrently by

$$\lambda_n = \min \{ \rho(x) : [x, x_i] = 0, \ i = 1, \ldots, n-1, \ x \neq 0 \}.$$
Poincaré’s maxmin characterization was first generalized by Duffin (1955) to overdamped quadratic eigenproblems of finite dimension, and for more general overdamped problems of finite dimension it was proved by Rogers (1964).
Poincaré’s maxmin characterization was first generalized by Duffin (1955) to overdamped quadratic eigenproblems of finite dimension, and for more general overdamped problems of finite dimension it was proved by Rogers (1964).

Infinite dimensional eigenvalue problems were studied by Turner (1967), Langer (1968), and Weinberger (1969) who proved generalizations of both, the maxmin characterization of Poincaré and of the minmax characterization of Courant, Fischer and Weyl for quadratic (and by Turner (1968) for polynomial) overdamped problems.
Poincaré’s maxmin characterization was first generalized by Duffin (1955) to overdamped quadratic eigenproblems of finite dimension, and for more general overdamped problems of finite dimension it was proved by Rogers (1964).

Infinite dimensional eigenvalue problems were studied by Turner (1967), Langer (1968), and Weinberger (1969) who proved generalizations of both, the maxmin characterization of Poincaré and of the minmax characterization of Courant, Fischer and Weyl for quadratic (and by Turner (1968) for polynomial) overdamped problems.

The corresponding generalizations for general overdamped problems of infinite dimension were derived by Hadeler (1968). Similar results (weakening the compactness or smoothness requirements) are contained in Rogers (1968), Werner (1971), Abramov (1973), Hadeler (1975), Markus (1985), Maksudov & Gasanov (1992), and Hasanov (2002).
Let $T(\lambda) : \mathcal{H} \to \mathcal{H}$, $\lambda \in J$ be a family of linear self-adjoint and bounded operators such that (1) is over-damped, and assume that for $\lambda \in J$ there exists $\nu(\lambda) > 0$ such that $T(\lambda) + \nu(\lambda)I$ is compact.
Let $T(\lambda) : \mathcal{H} \rightarrow \mathcal{H}, \lambda \in J$ be a family of linear self-adjoint and bounded operators such that (1) is over-damped, and assume that for $\lambda \in J$ there exists $\nu(\lambda) > 0$ such that $T(\lambda) + \nu(\lambda)I$ is compact.

Let $T(\cdot)$ be continuously differentiable and suppose that

$$\frac{\partial}{\partial \lambda} \langle T(\lambda)x, x \rangle \bigg|_{\lambda = \rho(x)} > 0 \quad \text{for ever } x \neq 0.$$
Minmax and Maxmin principle (Haderler 1968)

Let $T(\lambda) : \mathcal{H} \to \mathcal{H}$, $\lambda \in J$ be a family of linear self-adjoint and bounded operators such that (1) is over-damped, and assume that for $\lambda \in J$ there exists $\nu(\lambda) > 0$ such that $T(\lambda) + \nu(\lambda)I$ is compact.

Let $T(\cdot)$ be continuously differentiable and suppose that

$$\frac{\partial}{\partial \lambda} \langle T(\lambda)x, x \rangle \bigg|_{\lambda = \rho(x)} > 0 \quad \text{for ever } x \neq 0.$$

Let the eigenvalues $\lambda_n$ of $T(\lambda)x = 0$ be numbered in non-decreasing order regarding their multiplicities. Then they can be characterized by the following two variational principles

$$\lambda_n = \min_{\dim V = n} \max_{x \in V, \ x \neq 0} \rho(x)$$
$$= \max_{\dim V = n-1} \min_{x \in V^\perp, \ x \neq 0} \rho(x).$$
Example 1: Gyroscopic eigenvalue problem

For the gyroscopic eigenvalue problem the family of operators $Q(\lambda)$ satisfies the general conditions of the variational characterizations.
Example 1: Gyroscopic eigenvalue problem

For the gyroscopic eigenvalue problem the family of operators $Q(\lambda)$ satisfies the general conditions of the variational characterizations.

For $x \neq 0$ it holds for $x := x_1 + ix_2$ that

$$f(\omega, x) = x^H Q(\omega) x = x^H K x - 2\omega x_1^T G x_1 - \omega^2 x^H M x = 0,$$

since

$$x^H G x = x_1^T H x_1 - ix_2^T G x_1 + ix_1^T G x_2 + x_2^T G x_2 = 2ix_1^T G x_2.$$
Example 1: Gyroscopic eigenvalue problem

For the gyroscopic eigenvalue problem the family of operators $Q(\lambda)$ satisfies the general conditions of the variational characterizations.

For $x \neq 0$ it holds for $x := x_1 + ix_2$ that

$$f(\omega, x) = x^H Q(\omega) x = x^H K x - 2\omega x_1^T G x_1 - \omega^2 x^H M x = 0,$$

since

$$x^H G x = x_1^T H x_1 - ix_2^T G x_1 + ix_1^T G x_2 + x_2^T G x_2 = 2ix_1^T G x_2.$$

Hence, $f(\omega, x) = 0$ has exactly one solution $p_+(x) \in J_+ = (0, \infty)$ for $x \neq 0$, such that the sign condition

$$\frac{\partial}{\partial \omega} \langle Q(\omega)x, x \rangle \bigg|_{\omega=p_+(x)} > 0 \quad \text{for ever} \ x \neq 0$$

is satisfied. Therefore $Q(\omega)x = 0$ has $n$ positive eigenvalues $0 < \omega_1 \leq \cdots \leq \omega_n$ which can be characterized by all three variational principles.
Example 1: Gyroscopic eigenvalue problem

For the gyroscopic eigenvalue problem the family of operators $Q(\lambda)$ satisfies the general conditions of the variational characterizations.

For $x \neq 0$ it holds for $x := x_1 + ix_2$ that

$$ f(\omega, x) = x^H Q(\omega) x = x^H Kx - 2\omega x_1^T Gx_1 - \omega^2 x^H Mx = 0, $$

since

$$ x^H Gx = x_1^T Hx_1 - ix_2^T Gx_1 + i x_1^T Gx_2 + x_2^T Gx_2 = 2ix_1^T Gx_2. $$

Hence, $f(\omega, x) = 0$ has exactly one solution $p_+(x) \in J_+ = (0, \infty)$ for $x \neq 0$, such that the sign condition

$$ \left. \frac{\partial}{\partial \omega} \langle Q(\omega) x, x \rangle \right|_{\omega = p_+(x)} > 0 \quad \text{for every} \ x \neq 0 $$

is satisfied. Therefore $Q(\omega)x = 0$ has $n$ positive eigenvalues $0 < \omega_1 \leq \cdots \leq \omega_n$ which can be characterized by all three variational principles.

Likewise, it has $n$ negative eigenvalues, which allow also for all three variational characterizations.
Example 2: Fluid-solid vibration

For the rational eigenproblem governing free vibrations the family of operators $T(\lambda)$ satisfies the general conditions of the variational characterizations in every interval $J_n := (\sigma_{n-1}, \sigma_n)$. 
Example 2: Fluid-solid vibration

For the rational eigenproblem governing free vibrations the family of operators $T(\lambda)$ satisfies the general conditions of the variational characterizations in every interval $J_n := (\sigma_{n-1}, \sigma_n)$.

$$f(\lambda, p) := \langle T(\lambda)p, p \rangle = -\langle K_f p, p \rangle + \lambda \langle M_f p, p \rangle + \sum_{n=1}^{\infty} \frac{\lambda}{\sigma_n - \lambda} |\langle u_n, Cp \rangle|^2$$

is monotonically increasing, such that $f(\lambda, p) = 0$ has at most one solution in $J_n$, but the Rayleigh functional is not defined on the entire space.
Outline

1. Two examples

2. Variational characterization of eigenvalue problems
   - Overdamped Problems
   - Nonoverdamped problems

3. Applications
   - Sylvester’s law of inertia
   - Nonlinear low rank modification of a symmetric eigenvalue problem
   - Safeguarded iteration
   - Detecting hyperbolic matrix polynomials
   - Free vibrations of fluid-solid structures
   - Regularization of total least squares problems

4. Concluding remarks
For nonoverdamped eigenproblems (i.e. $D(p) \neq \mathcal{H} \setminus \{0\}$) the natural ordering to call the smallest eigenvalue the first one, the second smallest the second one, etc., is not appropriate.
For nonoverdamped eigenproblems (i.e. $\mathcal{D}(p) \neq \mathcal{H} \setminus \{0\}$) the natural ordering to call the smallest eigenvalue the first one, the second smallest the second one, etc., is not appropriate.

This is obvious if we make a linear eigenvalue

$$T(\lambda)x := (\lambda I - A)x = 0$$

nonlinear by restricting it to an interval $J$ which does not contain the smallest eigenvalue of $A$. 
Non-overdamped problems

For nonoverdamped eigenproblems (i.e. $\mathcal{D}(p) \neq \mathcal{H} \setminus \{0\}$) the natural ordering to call the smallest eigenvalue the first one, the second smallest the second one, etc., is not appropriate.

This is obvious if we make a linear eigenvalue

$$T(\lambda)x := (\lambda I - A)x = 0$$

nonlinear by restricting it to an interval $J$ which does not contain the smallest eigenvalue of $A$.

Then all conditions are satisfied, $p$ is the restriction of the Rayleigh quotient $R_A$ to $\mathcal{D}(p) := \{x \neq 0 : R_A(x) \in J\}$, and $\inf_{x \in \mathcal{D}(p)} p(x)$ will not be an eigenvalue.
Enumeration of eigenvalues

\( \lambda \in J \) is an eigenvalue of \( T(\cdot) \) if and only if \( \mu = 0 \) is an eigenvalue of the linear problem \( T(\lambda)y = \mu y \). The key idea is to orientate the number of \( \lambda \) on the location on the eigenvalue \( \mu = 0 \) in the spectrum of the linear operator \( T(\lambda) \).
\( \lambda \in J \) is an eigenvalue of \( T(\cdot) \) if and only if \( \mu = 0 \) is an eigenvalue of the linear problem \( T(\lambda)y = \mu y \). The key idea is to orientate the number of \( \lambda \) on the location on the eigenvalue \( \mu = 0 \) in the spectrum of the linear operator \( T(\lambda) \).

We assume that for every \( \lambda \in J \) it holds that the supremum of the essential spectrum of \( T(\lambda) \) is negative (for instance: there exists \( \nu(\lambda) > 0 \) such that \( T(\lambda) + \nu(\lambda)I \) is compact).
\( \lambda \in J \) is an eigenvalue of \( T(\cdot) \) if and only if \( \mu = 0 \) is an eigenvalue of the linear problem \( T(\lambda)y = \mu y \). The key idea is to orientate the number of \( \lambda \) on the location on the eigenvalue \( \mu = 0 \) in the spectrum of the linear operator \( T(\lambda) \).

We assume that for every \( \lambda \in J \) it holds that the supremum of the essential spectrum of \( T(\lambda) \) is negative (for instance: there exists \( \nu(\lambda) > 0 \) such that \( T(\lambda) + \nu(\lambda)I \) is compact).

If \( \lambda \in J \) is an eigenvalue of \( T(\cdot) \) then there exists \( n \in \mathbb{N} \) such that

\[
0 = \max_{\dim V = n} \min_{x \in V, x \neq 0} \frac{\langle T(\lambda)x, x \rangle}{\langle x, x \rangle}.
\]
\[ \lambda \in J \text{ is an eigenvalue of } T(\cdot) \text{ if and only if } \mu = 0 \text{ is an eigenvalue of the linear problem } T(\lambda)y = \mu y. \] The key idea is to orientate the number of \( \lambda \) on the location on the eigenvalue \( \mu = 0 \) in the spectrum of the linear operator \( T(\lambda) \).

We assume that for every \( \lambda \in J \) it holds that the supremum of the essential spectrum of \( T(\lambda) \) is negative (for instance: there exists \( \nu(\lambda) > 0 \) such that \( T(\lambda) + \nu(\lambda)I \) is compact).

If \( \lambda \in J \) is an eigenvalue of \( T(\cdot) \) then there exists \( n \in \mathbb{N} \) such that

\[ 0 = \max_{\dim V = n} \min_{x \in V, x \neq 0} \frac{\langle T(\lambda)x, x \rangle}{\langle x, x \rangle}. \]

In this case we assign \( n \) to the eigenvalue \( \lambda \) of problem \( T(\lambda)x = 0 \) as its number and call \( \lambda \) an \( n \)-th eigenvalue of \( T(\cdot) \).
Minmax characterization (V.&B.Werner 1982, V. 2010)

Let $T(\lambda), \lambda \in J$ be a family of linear self-adjoint and bounded operators on a Hilbert space $\mathcal{H}$ depending continuously on a parameter $\lambda \in J$ where $J$ is an open real (not necessarily bounded) interval.
Minmax characterization (V.&B.Werner 1982, V. 2010)

Let $T(\lambda), \lambda \in J$ be a family of linear self-adjoint and bounded operators on a Hilbert space $\mathcal{H}$ depending continuously on a parameter $\lambda \in J$ where $J$ is an open real (not necessarily bounded) interval.

Assume that

- for $x \neq 0$ $f(\lambda, x) := \langle T(\lambda)x, x \rangle = 0$ has at most one solution $p(x) \in J$, and let $\mathcal{D}$ be the domain of the Rayleigh functional $p$,
- $(\lambda - p(x))f(\lambda, x) > 0$ for every $x \in \mathcal{D}$ and $\lambda \neq p(x)$,
- and that the supremum of the essential spectrum of $T(\lambda)$ is negative for every $\lambda \in J$. 

Let $T(\lambda), \lambda \in J$ be a family of linear self-adjoint and bounded operators on a Hilbert space $\mathcal{H}$ depending continuously on a parameter $\lambda \in J$ where $J$ is an open real (not necessarily bounded) interval.

Assume that

- for $x \neq 0$ $f(\lambda, x) := \langle T(\lambda)x, x \rangle = 0$ has at most one solution $p(x) \in J$, and let $D$ be the domain of the Rayleigh functional $p$,
- $(\lambda - p(x))f(\lambda, x) > 0$ for every $x \in D$ and $\lambda \neq p(x)$,
- and that the supremum of the essential spectrum of $T(\lambda)$ is negative for every $\lambda \in J$.

Then the nonlinear eigenvalue problem $T(\lambda)x = 0$ has at most a countable set of eigenvalues in $J$, and it holds that:

If $\lambda_n \in J$ is an $n$-th eigenvalue then

$$\lambda_n = \min_{\dim V = n, \, V \cap D \neq \emptyset} \sup_{x \in D \cap V} p(x).$$

(6)
Let $T(\lambda), \lambda \in J$ be a family of linear self-adjoint and bounded operators on a Hilbert space $\mathcal{H}$ depending continuously on a parameter $\lambda \in J$ where $J$ is an open real (not necessarily bounded) interval.

Assume that

- for $x \neq 0$ $f(\lambda, x) := \langle T(\lambda)x, x \rangle = 0$ has at most one solution $p(x) \in J$, and let $D$ be the domain of the Rayleigh functional $p$,
- $(\lambda - p(x))f(\lambda, x) > 0$ for every $x \in D$ and $\lambda \neq p(x)$,
- and that the supremum of the essential spectrum of $T(\lambda)$ is negative for every $\lambda \in J$.

Then the nonlinear eigenvalue problem $T(\lambda)x = 0$ has at most a countable set of eigenvalues in $J$, and it holds that:

If $\lambda_n \in J$ is an $n$-th eigenvalue then

$$\lambda_n = \min_{\dim V = n, \ V \cap D \neq \emptyset} \sup_{x \in D \cap V} p(x). \quad (6)$$

If conversely

$$\lambda_n = \inf_{\dim V = n, \ V \cap D \neq \emptyset} \sup_{x \in D \cap V} p(x) \in J$$

then $\lambda_n$ is an $n$-th eigenvalue of $T(\lambda)x = 0$ and (6) holds.
Sketch of proof

Step 1 (technical): Let $\lambda \in J$, and assume that $V$ is a finite dimensional subspace of $\mathcal{H}$ such that $V \cap \mathcal{D} \neq \emptyset$. Then it holds that

$$
\lambda \left\{ \begin{array}{c}
< \\
= \\
> \\
\end{array} \right\} \sup_{x \in V \cap \mathcal{D}(p)} p(x) \iff \min_{x \in V} \langle T(\lambda)x, x \rangle \left\{ \begin{array}{c}
< \\
= \\
> \\
\end{array} \right\} 0 \quad (7)
$$
Sketch of proof

Step 1 (technical): Let $\lambda \in J$, and assume that $V$ is a finite dimensional subspace of $\mathcal{H}$ such that $V \cap D \neq \emptyset$. Then it holds that

$$
\lambda \left\{ \begin{array}{c}
< \\
= \\
> 
\end{array} \right\} \sup_{x \in V \cap D(p)} p(x) \iff \min_{x \in V} \langle T(\lambda)x, x \rangle \left\{ \begin{array}{c}
< \\
= \\
> 
\end{array} \right\} 0 \quad (7)
$$

Step 2: If $\lambda_n$ is an $n$-th eigenvalue, then $\mu_n(\lambda_n) = 0$, and

$$
\mu_n(\lambda_n) = \max_{\dim V = n} \min_{x \in V, \|x\|=1} \langle T(\lambda_n)x, x \rangle = \min_{x \in \overline{V}, \|x\|=1} \langle T(\lambda_n)x, x \rangle.
$$
Sketch of proof

Step 1 (technical): Let $\lambda \in J$, and assume that $V$ is a finite dimensional subspace of $\mathcal{H}$ such that $V \cap \mathcal{D} \neq \emptyset$. Then it holds that

\[
\lambda \begin{cases}
\quad < \\
\quad = \\
\quad > 
\end{cases} \sup_{x \in V \cap \mathcal{D}(p)} p(x) \iff \min_{x \in V} \langle T(\lambda)x, x \rangle \begin{cases}
\quad < \\
\quad = \\
\quad > 
\end{cases} 0 \quad (7)
\]

Step 2: If $\lambda_n$ is an $n$-th eigenvalue, then $\mu_n(\lambda_n) = 0$, and

\[
\mu_n(\lambda_n) = \max_{\dim V = n} \min_{x \in V, \|x\|=1} \langle T(\lambda_n)x, x \rangle = \min_{x \in \bar{V}, \|x\|=1} \langle T(\lambda_n)x, x \rangle.
\]

Hence, $\min_{x \in V, \|x\|=1} \langle T(\lambda_n)x, x \rangle \leq 0$ for every $V$ with $\dim V = n$, and (7) implies

\[
\sup_{x \in V \cap \mathcal{D}} p(x) \geq \lambda_n = \sup_{x \in \bar{V} \cap \mathcal{D}} p(x).
\]

Hence, $\lambda_n$ is a minmax value of $p$. 
Theorem (minmax for extreme eigenvalues)

Assume that the conditions of the minmax characterization hold and that

$$\inf_{x \in D} p(x) \in J.$$
Assume that the conditions of the minmax characterization hold and that

\[
\inf_{x \in D} p(x) \in J.
\]

If \( \lambda_n \in J \) for some \( n \in \mathbb{N} \) then every \( V \in H_j \) with \( V \cap D(p) \neq \emptyset \) and \( \lambda_j = \sup_{x \in V \cap D(p)} p(x) \) is contained in \( D \cup \{0\} \), and the characterization (6) can be replaced by

\[
\lambda_j = \min_{\text{dim } V = j} \max_{V \subset D \cup \{0\}, \, v \in V, \, x \neq 0} p(v), \quad j = 1, \ldots, n.
\]
Maxmin characterization (V. 2003)

Assume that the conditions of the minmax characterization are satisfied.
Maxmin characterization (V. 2003)

Assume that the conditions of the minmax characterization are satisfied.

If there is an $n$-th eigenvalue $\lambda_n \in J$ of $T(\lambda)x = 0$, then

$$\lambda_n = \max_{V \in H_{n-1}} \inf_{V \perp \cap D \neq \emptyset} p(v),$$

and the maximum is attained by $W := \text{span}\{u_1, \ldots, u_{n-1}\}$ where $u_j$ denotes an eigenvector corresponding to the $j$-largest eigenvalue $\mu_j(\lambda_n)$ of $T(\lambda_n)$. 
Assume that the conditions of the minmax characterization are satisfied.

If there is an $n$-th eigenvalue $\lambda_n \in J$ of $T(\lambda)x = 0$, then

$$\lambda_n = \max_{V \in H_{n-1}} \inf_{V \in V^\perp \cap D} p(v),$$

and the maximum is attained by $W := \text{span}\{u_1, \ldots, u_{n-1}\}$ where $u_j$ denotes an eigenvector corresponding to the $j$-largest eigenvalue $\mu_j(\lambda_n)$ of $T(\lambda_n)$.

Proof takes advantage of the following Lemma:
Let $\lambda \in J$, and let $V$ be a finite dimensional subspace of $\mathcal{H}$ such that $V^\perp \cap D \neq \emptyset$. Then it holds that

$$\lambda \left\{ \begin{array}{ll} < \\ = \\ > \end{array} \right\} \inf_{x \in V^\perp \cap D(p)} p(x) \iff \max_{x \in V^\perp, \|x\|=1} \langle T(\lambda)x, x \rangle \left\{ \begin{array}{ll} < \\ = \\ > \end{array} \right\} 0$$
Under the conditions of the minmax characterization assume that \( J \) contains \( n \geq 1 \) eigenvalues \( \lambda_1 \leq \cdots \leq \lambda_n \) (where \( \lambda_i \) is an \( i \)th eigenvalue) with corresponding \( [\cdot, \cdot] \) orthogonal eigenvectors \( x_1, \ldots, x_n \).

If there exists \( x \in D \) with \( [x_i, x] = 0 \) for \( i = 1, \ldots, n \) then \( J \) contains an \((n + 1)\)th eigenvalue, and

\[
\lambda_{n+1} = \inf\{ p(x) : [x_j, x] = 0, \ i = 1, \ldots, n \}. \tag{9}
\]
Under the conditions of the minmax characterization assume that $J$ contains $n \geq 1$ eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$ (where $\lambda_i$ is an $i$th eigenvalue) with corresponding $[\cdot, \cdot]$ orthogonal eigenvectors $x_1, \ldots, x_n$. If there exists $x \in D$ with $[x_i, x] = 0$ for $i = 1, \ldots, n$ then $J$ contains an $(n+1)$th eigenvalue, and

$$\lambda_{n+1} = \inf \{ p(x) : [x_j, x] = 0, \ i = 1, \ldots, n \}. \quad (9)$$

The generalized scalar product (5) has to be modified according to

$$[x, y] := \begin{cases} \langle T(p(x)) - T(p(y)) \rangle x, y, & \text{if } p(x) \neq p(y) \\ \langle x, y \rangle, & \text{if } p(x) = p(y) \end{cases}$$

if $T$ is not differentiable at $p(x) = p(y)$. 


Sketch of proof

Let \( \{y_k\} \) be a minimizing sequence of (9) such that

\[
\|y_k\| = 1, \quad [x_j, y_k] = 0, \quad j = 1, \ldots, n, \quad p(y_k) \to \lambda_{n+1}.
\]
Sketch of proof

Let \( \{y_k\} \) be a minimizing sequence of (9) such that

\[
\|y_k\| = 1, \quad [x_j, y_k] = 0, \quad j = 1, \ldots, n, \quad p(y_k) \to \lambda_{n+1}.
\]

There exist unique

\[
\tilde{y}_k = y_k + \sum_{j=1}^{n} c_{kj} x_j
\]

such that

\[
\langle T(\lambda_{n+1})x_j, \tilde{y}_k \rangle = 0, \quad j = 1, \ldots, n,
\]

and it holds that

\[
\|y_k - \tilde{y}_k\| \to 0 \quad \text{and} \quad p(\tilde{y}_k) \to \lambda_{n+1}.
\]
Sketch of proof

Let \( \{y_k\} \) be a minimizing sequence of (9) such that

\[
\|y_k\| = 1, \quad [x_j, y_k] = 0, \quad j = 1, \ldots, n, \quad p(y_k) \to \lambda_{n+1}.
\]

There exist unique

\[
\tilde{y}_k = y_k + \sum_{j=1}^{n} c_{kj} x_j
\]

such that \( \langle T(\lambda_{n+1}) x_j, \tilde{y}_k \rangle = 0, \quad j = 1, \ldots, n \), and it holds that

\[
\|y_k - \tilde{y}_k\| \to 0 \quad \text{and} \quad p(\tilde{y}_k) \to \lambda_{n+1}.
\]

For \( V_k := \text{span}\{x_1, \ldots, x_n, \tilde{y}_k\} \) it can be shown that

\[
\lim_{k \to \infty} \sup \{p(z) : z \in V_k\} = \lambda_{n+1},
\]

and it follows from the minmax characterization that \( \lambda_{n+1} \) is an eigenvalue of \( T(\cdot) \).
Minmax characterizations for non-overdamped nonlinear eigenvalue problems were proved (independently from our work) by

- **Barston** (1974) for some extreme eigenvalues of finite dimensional quadratic eigenproblems
Minmax characterizations for non-overdamped nonlinear eigenvalue problems were proved (independently from our work) by

- **Barston** (1974) for some extreme eigenvalues of finite dimensional quadratic eigenproblems

and for the infinite dimensional case by

- **Griniv & Mel’nik** (1996): $T(\lambda) = A(\lambda) - I$, $A(\lambda)$ compact
Minmax characterizations for non-overdamped nonlinear eigenvalue problems were proved (independently from our work) by

- **Barston** (1974) for some extreme eigenvalues of finite dimensional quadratic eigenproblems

and for the infinite dimensional case by

- **Griniv & Meln’ik** (1996): $T(\lambda) = A(\lambda) - I$, $A(\lambda)$ compact
- **Binding, Eschwé & H. Langer** (2000)
Further literature

Minmax characterizations for non-overdamped nonlinear eigenvalue problems were proved (independently from our work) by

- **Barston** (1974) for some extreme eigenvalues of finite dimensional quadratic eigenproblems

and for the infinite dimensional case by

- **Griniv & Meln’ik** (1996): $T(\lambda) = A(\lambda) - I$, $A(\lambda)$ compact
- **Binding, Eschwé & H. Langer** (2000)
- **Eschwé & M. Langer** (2004) unbounded operators
Outline

1. Two examples

2. Variational characterization of eigenvalue problems
   - Overdamped Problems
   - Nonoverdamped problems

3. Applications
   - Sylvester’s law of inertia
   - Nonlinear low rank modification of a symmetric eigenvalue problem
   - Safeguarded iteration
   - Detecting hyperbolic matrix polynomials
   - Free vibrations of fluid-solid structures
   - Regularization of total least squares problems

4. Concluding remarks
Outline

1. Two examples

2. Variational characterization of eigenvalue problems
   - Overdamped Problems
   - Nonoverdamped problems

3. Applications
   - Sylvester’s law of inertia
   - Nonlinear low rank modification of a symmetric eigenvalue problem
   - Safeguarded iteration
   - Detecting hyperbolic matrix polynomials
   - Free vibrations of fluid-solid structures
   - Regularization of total least squares problems

4. Concluding remarks
Number of eigenvalues

The inertia of a Hermitian matrix $A$ is the triplet of nonnegative integers

$$\text{In}(A) := (n_p, n_n, n_z)$$

where $n_p$, $n_n$, and $n_z$ are the number of positive, negative, and zero eigenvalues of $A$ (counting multiplicities).
Number of eigenvalues

The inertia of a Hermitian matrix $A$ is the triplet of nonnegative integers

$$\text{In}(A) := (n_p, n_n, n_z)$$

where $n_p$, $n_n$, and $n_z$ are the number of positive, negative, and zero eigenvalues of $A$ (counting multiplicities).

Sylvester’s classical law of inertia states that two Hermitian matrices $A, B \in \mathbb{C}^{n \times n}$ are congruent (i.e. $A = S^H BS$ for some nonsingular matrix $S$) if and only if they have the same inertia $\text{In}(A) = \text{In}(B)$. 
Number of eigenvalues

The inertia of a Hermitian matrix $A$ is the triplet of nonnegative integers

$$\ln(A) := (n_p, n_n, n_z)$$

where $n_p$, $n_n$, and $n_z$ are the number of positive, negative, and zero eigenvalues of $A$ (counting multiplicities).

Sylvester’s classical law of inertia states that two Hermitian matrices $A, B \in \mathbb{C}^{n \times n}$ are congruent (i.e. $A = S^H B S$ for some nonsingular matrix $S$) if and only if they have the same inertia $\ln(A) = \ln(B)$.

An obvious consequence of the law of inertia is the following corollary: If $A$ has an $LDL^H$ factorization $A = LDL^H$, then $n_p$ and $n_n$ equals the number of positive and negative entries of $D$, respectively, and if only a block $LDL^H$ factorization exists where $D$ is a block diagonal matrix with $1 \times 1$ and indefinite $2 \times 2$ blocks on its diagonal, then one has to increase the number of positive and negative $1 \times 1$ blocks of $D$ by the number of $2 \times 2$ blocks to get $n_p$ and $n_n$, respectively. Hence, the inertia of $A$ can be computed easily. This is particularly advantageous if the matrix is banded.
Sylvester’s law

Theorem (overdamped case)
Assume that $T : J \to \mathbb{C}^{n \times n}$ satisfies the conditions of the minmax characterization, and assume that the nonlinear eigenvalue problem $T(\lambda)x = 0$ is overdamped.

For $\sigma \in J$ let $(n_p, n_n, n_z)$ be the inertia of $T(\sigma)$. Then the nonlinear eigenproblem $T(\lambda)x = 0$ has $n$ eigenvalues in $J$, $n_p$ of which are smaller than $\sigma$, $n_n$ exceed $\sigma$, and for $n_z > 0$, $\sigma$ is an eigenvalue of geometric multiplicity $n_z$. 
Sylvester’s law

Theorem (overdamped case)
Assume that $T : J \to \mathbb{C}^{n \times n}$ satisfies the conditions of the minmax characterization, and assume that the nonlinear eigenvalue problem $T(\lambda)x = 0$ is overdamped.

For $\sigma \in J$ let $(n_p, n_n, n_z)$ be the inertia of $T(\sigma)$. Then the nonlinear eigenproblem $T(\lambda)x = 0$ has $n$ eigenvalues in $J$, $n_p$ of which are smaller than $\sigma$, $n_n$ exceed $\sigma$, and for $n_z > 0$, $\sigma$ is an eigenvalue of geometric multiplicity $n_z$.

Theorem (extreme eigenvalues)
Assume that $T : J \to \mathbb{C}^{n \times n}$ satisfies the conditions of the minmax characterization, and let $(n_p, n_n, n_z)$ be the inertia of $T(\sigma)$ for some $\sigma \in J$.

(i) If $\lambda_1 := \inf_{x \in D} p(x) \in J$, then the nonlinear eigenproblem $T(\lambda)x = 0$ has exactly $n_p$ eigenvalues $\lambda_1 \leq \cdots \leq \lambda_{n_p}$ in $J$ which are less than $\sigma$.

(ii) If $\sup_{x \in D} p(x) \in J$, then the nonlinear eigenproblem $T(\lambda)x = 0$ has exactly $n_n$ eigenvalues $\lambda_{n-n_n+1} \leq \cdots \leq \lambda_n$ in $J$ exceeding $\sigma$. 
**Sylvester’s law**

**Theorem (general case):** Let $T : J \rightarrow \mathbb{C}^{n \times n}$ satisfy the conditions of the minmax characterization, and let $\sigma, \tau \in J$, $\sigma < \tau$.

Let $(n_{p\sigma}, n_{n\sigma}, n_{z\sigma})$ and $(n_{p\tau}, n_{n\tau}, n_{z\tau})$ be the inertia of $T(\sigma)$ and $T(\tau)$, respectively. Then $n_{p\sigma} \leq n_{p\tau}$, and the eigenvalue problem $T(\lambda)x = 0$ has exactly $n_{p\tau} - n_{p\sigma}$ eigenvalues $\lambda_{n_{p\sigma}+1} \leq \cdots \leq \lambda_{n_{p\tau}}$ in $(\sigma, \tau)$. 
Sylvester’s law

Theorem (general case): Let $T : J \rightarrow \mathbb{C}^{n \times n}$ satisfy the conditions of the minmax characterization, and let $\sigma, \tau \in J$, $\sigma < \tau$. Let $(n_{p_\sigma}, n_{n_\sigma}, n_{z_\sigma})$ and $(n_{p_\tau}, n_{n_\tau}, n_{z_\tau})$ be the inertia of $T(\sigma)$ and $T(\tau)$, respectively. Then $n_{p_\sigma} \leq n_{p_\tau}$, and the eigenvalue problem $T(\lambda) x = 0$ has exactly $n_{p_\tau} - n_{p_\sigma}$ eigenvalues $\lambda_{n_{p_\sigma} + 1} \leq \cdots \leq \lambda_{n_{p_\tau}}$ in $(\sigma, \tau)$.

Example: Consider the rational eigenvalue problem

$$T(\lambda) := -K + \lambda M + \sum_{j=1}^{p} \frac{\lambda}{\sigma_j - \lambda} C_j C_j^T,$$

where $K, M \in \mathbb{R}^{n \times n}$ are symmetric and positive definite, $C_j \in \mathbb{R}^{n \times k_j}$ has rank $k_j$, and $0 < \sigma_1 < \cdots < \sigma_p$, which models the free vibrations of certain fluid–solid structures.
Sylvester’s law

**Theorem (general case):** Let $T : J \rightarrow \mathbb{C}^{n \times n}$ satisfy the conditions of the minmax characterization, and let $\sigma, \tau \in J$, $\sigma < \tau$.

Let $(n_{p\sigma}, n_{n\sigma}, n_{z\sigma})$ and $(n_{p\tau}, n_{n\tau}, n_{z\tau})$ be the inertia of $T(\sigma)$ and $T(\tau)$, respectively. Then $n_{p\sigma} \leq n_{p\tau}$, and the eigenvalue problem $T(\lambda)x = 0$ has exactly $n_{p\tau} - n_{p\sigma}$ eigenvalues $\lambda_{n_{p\sigma}+1} \leq \cdots \leq \lambda_{n_{p\tau}}$ in $(\sigma, \tau)$.

**Example:** Consider the rational eigenvalue problem

$$T(\lambda) := -K + \lambda M + \sum_{j=1}^{p} \frac{\lambda}{\sigma_j - \lambda} C_j C_j^T,$$

where $K, M \in \mathbb{R}^{n \times n}$ are symmetric and positive definite, $C_j \in \mathbb{R}^{n \times k_j}$ has rank $k_j$, and $0 < \sigma_1 < \cdots < \sigma_p$, which models the free vibrations of certain fluid–solid structures.

For the first interval $J_0 := (0, \sigma_1)$ we get: if $\tau \in J_0$ and $(n_p, n_n, n_z)$ is the inertia of $T(\tau)$, then there are exactly $n_p$ eigenvalues in $J_0$ which are less than $\tau$. 
Sylvester’s law

**Theorem (general case):** Let $T : J \rightarrow \mathbb{C}^{n \times n}$ satisfy the conditions of the minmax characterization, and let $\sigma, \tau \in J$, $\sigma < \tau$.

Let $(n_{p\sigma}, n_{n\sigma}, n_{z\sigma})$ and $(n_{p\tau}, n_{n\tau}, n_{z\tau})$ be the inertia of $T(\sigma)$ and $T(\tau)$, respectively. Then $n_{p\sigma} \leq n_{p\tau}$, and the eigenvalue problem $T(\lambda)x = 0$ has exactly $n_{p\tau} - n_{p\sigma}$ eigenvalues $\lambda_{n_{p\sigma}+1} \leq \cdots \leq \lambda_{n_{p\tau}}$ in $(\sigma, \tau)$.

**Example:** Consider the rational eigenvalue problem

$$T(\lambda) := -K + \lambda M + \sum_{j=1}^{p} \frac{\lambda}{\sigma_j - \lambda} C_j C_j^T,$$

where $K, M \in \mathbb{R}^{n \times n}$ are symmetric and positive definite, $C_j \in \mathbb{R}^{n \times k_j}$ has rank $k_j$, and $0 < \sigma_1 < \cdots < \sigma_p$, which models the free vibrations of certain fluid–solid structures.

For the first interval $J_0 := (0, \sigma_1)$ we get: if $\tau \in J_0$ and $(n_p, n_n, n_z)$ is the inertia of $T(\tau)$, then there are exactly $n_p$ eigenvalues in $J_0$ which are less than $\tau$.

If $\tau_1 < \tau_2$ are contained in one interval $J_j$ then the number of eigenvalues in the interval $(\tau_1, \tau_2)$ can be obtained from the inertia of $T(\tau_1)$ and $T(\tau_2)$ according to the last Theorem.
Outline

1. Two examples

2. Variational characterization of eigenvalue problems
   - Overdamped Problems
   - Nonoverdamped problems

3. Applications
   - Sylvester’s law of inertia
   - Nonlinear low rank modification of a symmetric eigenvalue problem
   - Safeguarded iteration
   - Detecting hyperbolic matrix polynomials
   - Free vibrations of fluid-solid structures
   - Regularization of total least squares problems

4. Concluding remarks
For $A = A^T \in \mathbb{R}^{n \times n}$ let $\alpha_1 \leq \cdots \leq \alpha_n$ be its eigenvalues, and let $\alpha_0 := -\infty$ and $\alpha_{n+1} := \infty$. 
For $A = A^T \in \mathbb{R}^{n \times n}$ let $\alpha_1 \leq \cdots \leq \alpha_n$ be its eigenvalues, and let $\alpha_0 := -\infty$ and $\alpha_{n+1} := \infty$.

The following interlacing theorem for constant rank-one modifications of Hermitian matrices is well known.
Low rank modification

For $A = A^T \in \mathbb{R}^{n \times n}$ let $\alpha_1 \leq \cdots \leq \alpha_n$ be its eigenvalues, and let $\alpha_0 := -\infty$ and $\alpha_{n+1} := \infty$.

The following interlacing theorem for constant rank-one modifications of Hermitian matrices is well known.

**Theorem**

Let $B := A + \tau cc^T$, $c \in \mathbb{R}^n$ and $\tau \in \mathbb{R}$, with eigenvalues $\beta_1 \leq \cdots \leq \beta_n$. Then it holds that

\[ \alpha_i \leq \beta_i \leq \alpha_{i+1} \quad \text{for} \quad \tau > 0, \quad i = 1, \ldots, n \quad (10) \]

\[ \alpha_{i-1} \leq \beta_i \leq \alpha_i \quad \text{for} \quad \tau < 0, \quad i = 1, \ldots, n. \quad (11) \]
Low rank modification

For $A = A^T \in \mathbb{R}^{n \times n}$ let $\alpha_1 \leq \cdots \leq \alpha_n$ be its eigenvalues, and let $\alpha_0 := -\infty$ and $\alpha_{n+1} := \infty$.

The following interlacing theorem for constant rank-one modifications of Hermitian matrices is well known.

**Theorem**

Let $B := A + \tau cc^T$, $c \in \mathbb{R}^n$ and $\tau \in \mathbb{R}$, with eigenvalues $\beta_1 \leq \cdots \leq \beta_n$. Then it holds that

\[ \alpha_i \leq \beta_i \leq \alpha_{i+1} \quad \text{for} \quad \tau > 0, \quad i = 1, \ldots, n \tag{10} \]

\[ \alpha_{i-1} \leq \beta_i \leq \alpha_i \quad \text{for} \quad \tau < 0, \quad i = 1, \ldots, n. \tag{11} \]

If $A$ is diagonal with distinct diagonal entries $\alpha_1 < \cdots < \alpha_n$, and if all components of $c$ are different from zero then (10) and (11) even hold with strict inequalities.
Low rank modification

From the last Theorem we immediately obtain the following existence result for the nonlinear rank-one modification \( B := A + \phi(\lambda)cc^T \) of \( A \).
Applications

Nonlinear low rank modification of a symmetric eigenvalue problem

Low rank modification

From the last Theorem we immediately obtain the following existence result for the nonlinear rank-one modification $B := A + \phi(\lambda)cc^T$ of $A$.

**Theorem**

(i) For $k \in \{1, \ldots, n-1\}$ let $\phi \in C[\alpha_k, \alpha_{k+1}]$ be nonnegative. Then the nonlinear eigenvalue problem

$$ (A + \phi(\lambda)cc^T)x = \lambda x \quad (12) $$

has an eigenvalue $\hat{\lambda} \in [\alpha_k, \alpha_{k+1}]$.

(ii) For $k \in \{2, \ldots, n\}$ let $\phi \in C[\alpha_{k-1}, \alpha_k]$ be nonpositive. Then the nonlinear eigenvalue problem (12) has an eigenvalue $\hat{\lambda} \in [\alpha_{k-1}, \alpha_k]$. 

Low rank modification

From the last Theorem we immediately obtain the following existence result for the nonlinear rank-one modification \( B := A + \phi(\lambda)cc^T \) of \( A \).

**Theorem**

(i) For \( k \in \{1, \ldots, n-1\} \) let \( \phi \in C[\alpha_k, \alpha_{k+1}] \) be nonnegative. Then the nonlinear eigenvalue problem

\[
(A + \phi(\lambda)cc^T)x = \lambda x
\]  

has an eigenvalue \( \hat{\lambda} \in [\alpha_k, \alpha_{k+1}] \).

(ii) For \( k \in \{2, \ldots, n\} \) let \( \phi \in C[\alpha_{k-1}, \alpha_k] \) be nonpositive. Then the nonlinear eigenvalue problem (12) has an eigenvalue \( \hat{\lambda} \in [\alpha_{k-1}, \alpha_k] \).

Notice that differently from a constant rank-one modification of \( A \) there does not necessarily exist an eigenvalue of (12) in \([\alpha_n, \infty)\) for \( \phi \geq 0 \).
Applications
Nonlinear low rank modification of a symmetric eigenvalue problem

Low rank modification

From the last Theorem we immediately obtain the following existence result for the nonlinear rank-one modification $B := A + \phi(\lambda)cc^T$ of $A$.

Theorem

(i) For $k \in \{1, \ldots, n - 1\}$ let $\phi \in C[\alpha_k, \alpha_{k+1}]$ be nonnegative. Then the nonlinear eigenvalue problem

$$ (A + \phi(\lambda)cc^T)x = \lambda x \quad (12) $$

has an eigenvalue $\hat{\lambda} \in [\alpha_k, \alpha_{k+1}]$.

(ii) For $k \in \{2, \ldots, n\}$ let $\phi \in C[\alpha_{k-1}, \alpha_k]$ be nonpositive. Then the nonlinear eigenvalue problem (12) has an eigenvalue $\hat{\lambda} \in [\alpha_{k-1}, \alpha_k]$.

Notice that differently from a constant rank-one modification of $A$ there does not necessarily exist an eigenvalue of (12) in $[\alpha_n, \infty)$ for $\phi \geq 0$.

If $c$ is an eigenvector of $A$ corresponding to $\alpha_n$ then

$$ (A + \phi(\lambda)cc^T)c = (\alpha_n + \phi(\lambda)\|c\|^2)c, \text{ and } \alpha_n + \phi(\lambda)\|c\|^2 \geq \alpha_n \text{ is the maximal eigenvalue of } (A + \phi(\lambda)cc^T)x = \mu x. \text{ Hence, if } \phi(\lambda)\|c\|^2 > \lambda - \alpha_n \text{ for every } \lambda \geq \alpha_n \text{ then there does not exist an eigenvalue of (12) in } [\alpha_n, \infty). \)
Low rank modification

The following theorem ensures the existence of eigenvalues in the extreme intervals \((-\infty, \alpha_1]\) and \([\alpha_n, \infty)\), respectively.

\[ \phi(\lambda) - \phi(\eta) \lambda - \eta \| c \|_2 \leq 1 - \delta \quad \text{for every} \quad \lambda > \eta. \quad (13) \]
Low rank modification

The following theorem ensures the existence of eigenvalues in the extreme intervals \((-\infty, \alpha_1]\) and \([\alpha_n, \infty)\), respectively.

**Theorem**

(i) Let \(\phi \in C[\alpha_n, \infty)\) be nonnegative and assume that there exists \(\eta \geq \alpha_n\) such that for some \(\delta > 0\) it holds that

\[
\frac{\phi(\lambda) - \phi(\eta)}{\lambda - \eta} \|c\|^2 \leq 1 - \delta \quad \text{for every } \lambda > \eta. \tag{13}
\]

Then the nonlinear eigenvalue problem \((12)\) has an eigenvalue \(\hat{\lambda} \in [\alpha_n, \infty)\) which is the largest eigenvalue of \((A + \phi(\lambda)cc^T)x = \mu x\) with \(\lambda = \hat{\lambda}\).
Low rank modification

The following theorem ensures the existence of eigenvalues in the extreme intervals $(-\infty, \alpha_1]$ and $[\alpha_n, \infty)$, respectively.

Theorem

(i) Let $\phi \in C[\alpha_n, \infty)$ be nonnegative and assume that there exists $\eta \geq \alpha_n$ such that for some $\delta > 0$ it holds that

$$\frac{\phi(\lambda) - \phi(\eta)}{\lambda - \eta} \|c\|^2 \leq 1 - \delta \quad \text{for every } \lambda > \eta. \quad (13)$$

Then the nonlinear eigenvalue problem (12) has an eigenvalue $\hat{\lambda} \in [\alpha_n, \infty)$ which is the largest eigenvalue of $(A + \phi(\lambda)c c^T)x = \mu x$ with $\lambda = \hat{\lambda}$.

(ii) A corresponding result holds for the interval $(-\infty, \alpha_1]$. 
The following theorem ensures the existence of eigenvalues in the extreme intervals \((-\infty, \alpha_1]\) and \([\alpha_n, \infty)\), respectively.

**Theorem**

(i) Let \(\phi \in C[\alpha_n, \infty)\) be nonnegative and assume that there exists \(\eta \geq \alpha_n\) such that for some \(\delta > 0\) it holds that

\[
\frac{\phi(\lambda) - \phi(\eta)}{\lambda - \eta} \|c\|^2 \leq 1 - \delta \quad \text{for every} \quad \lambda > \eta.
\]  

(13)

Then the nonlinear eigenvalue problem (12) has an eigenvalue \(\hat{\lambda} \in [\alpha_n, \infty)\) which is the largest eigenvalue of \((A + \phi(\lambda)cc^T)x = \mu x\) with \(\lambda = \hat{\lambda}\).

(ii) A corresponding result holds for the interval \((-\infty, \alpha_1]\).

It is shown that the maximal eigenvalue \(\mu_n(\lambda)\) of problem \((A + \phi(\lambda)cc^T)x = \mu x\) satisfies \(\mu_n(\lambda) \geq \alpha_n\), and in particular \(\mu_n(\alpha_n) \geq \alpha_n\), and that for sufficiently large \(\lambda\) it holds that \(\mu_n(\lambda) \leq \lambda\) such that \(\mu_n(\cdot)\) has a fixed point in \([\alpha_n, \lambda]\).
Low rank modification

Notice that condition (13) can not be relaxed to

$$\frac{\phi(\lambda) - \phi(\eta)}{\lambda - \eta} \|c\|^2 < 1 \quad \text{for every } \lambda > \eta.$$  

Choosing in the last example $\phi(\lambda) = (\lambda - \alpha_n + e^{-\lambda})/\|c\|^2$, then the maximum eigenvalue $\mu_n(\lambda) = \lambda + e^{-\lambda}$ of $(A + \phi(\lambda)cc^T)x = \mu x$ has no fixed point in $[\alpha_n, \infty)$. 

Low rank modification

Notice that condition (13) can not be relaxed to

\[ \frac{\phi(\lambda) - \phi(\eta)}{\lambda - \eta} \|c\|^2 < 1 \quad \text{for every } \lambda > \eta. \]

Choosing in the last example \( \phi(\lambda) = (\lambda - \alpha_n + e^{-\lambda})/\|c\|^2 \), then the maximum eigenvalue \( \mu_n(\lambda) = \lambda + e^{-\lambda} \) of \((A + \phi(\lambda)cc^T)x = \mu x\) has no fixed point in \([\alpha_n, \infty)\).

The following uniqueness theorem also holds for the unbounded interval \((\alpha_n, \infty)\).
Low rank modification

Notice that condition (13) can not be relaxed to

\[
\frac{\phi(\lambda) - \phi(\eta)}{\lambda - \eta} \|c\|^2 < 1 \quad \text{for every } \lambda > \eta.
\]

Choosing in the last example \( \phi(\lambda) = (\lambda - \alpha_n + e^{-\lambda})/\|c\|^2 \), then the maximum eigenvalue \( \mu_n(\lambda) = \lambda + e^{-\lambda} \) of \((A + \phi(\lambda)cc^T)x = \mu x\) has no fixed point in \([\alpha_n, \infty)\).

The following uniqueness theorem also holds for the unbounded interval \((\alpha_n, \infty)\).

**Theorem**

For \( k \in \{1, \ldots, n\} \) let \( \phi \in C[\alpha_k, \alpha_{k+1}] \) be nonnegative and assume the condition

\[
\frac{\phi(\lambda) - \phi(\tilde{\lambda})}{\lambda - \tilde{\lambda}} \|c\|^2 < 1
\]

is satisfied for \( \lambda, \tilde{\lambda} \in I := (\alpha_k, \alpha_{k+1}), \lambda \neq \tilde{\lambda} \).
Low rank modification

Notice that condition (13) can not be relaxed to

\[
\frac{\phi(\lambda) - \phi(\eta)}{\lambda - \eta} \|c\|^2 < 1 \quad \text{for every } \lambda > \eta.
\]

Choosing in the last example \( \phi(\lambda) = (\lambda - \alpha_n + e^{-\lambda})/\|c\|^2 \), then the maximum eigenvalue \( \mu_n(\lambda) = \lambda + e^{-\lambda} \) of \((A + \phi(\lambda)cc^T)x = \mu x\) has no fixed point in \([\alpha_n, \infty)\).

The following uniqueness theorem also holds for the unbounded interval \((\alpha_n, \infty)\).

**Theorem**

For \( k \in \{1, \ldots, n\} \) let \( \phi \in C[\alpha_k, \alpha_{k+1}] \) be nonnegative and assume the condition

\[
\frac{\phi(\lambda) - \phi(\tilde{\lambda})}{\lambda - \tilde{\lambda}} \|c\|^2 < 1
\]

(14)

is satisfied for \( \lambda, \tilde{\lambda} \in I := (\alpha_k, \alpha_{k+1}), \lambda \neq \tilde{\lambda} \).

Then problem (12) has at most one eigenvalue \( \hat{\lambda} \in I \).
Low rank modification

We now consider a perturbation of a symmetric eigenvalue problem of the form

\[(A + \tau H)x = \lambda x\]

where \(A, H \in \mathbb{R}^{n \times n}\) are symmetric, \(H\) has small rank \(r \ll n\), and \(\tau \in \mathbb{R}\).
Low rank modification

We now consider a perturbation of a symmetric eigenvalue problem of the form

$$(A + \tau H)x = \lambda x$$

where $A, H \in \mathbb{R}^{n \times n}$ are symmetric, $H$ has small rank $r \ll n$, and $\tau \in \mathbb{R}$. The inertia of $H$ is denoted by $(\pi, \nu, \zeta)$, and its eigenvalues by $\sigma_1 \leq \cdots \leq \sigma_\nu < 0 < \sigma_{n-\pi+1} \leq \cdots \leq \sigma_n$. The following Theorem of Weyl (1912) can also be generalized to nonlinear small rank perturbations.

Theorem
Let $\beta_1 \leq \cdots \leq \beta_n$ denote the eigenvalues of $B := A + \tau H$. Then it holds that $\alpha_i - \nu \leq \beta_i \leq \alpha_i + \pi$ for $\tau > 0$, $i = 1, \ldots, n$.

These results can be generalized to small rank perturbations of nonlinear eigenvalue problems allowing for a minmax characterization of its eigenvalues.
Low rank modification

We now consider a perturbation of a symmetric eigenvalue problem of the form

\[(A + \tau H)x = \lambda x\]

where \(A, H \in \mathbb{R}^{n \times n}\) are symmetric, \(H\) has small rank \(r \ll n\), and \(\tau \in \mathbb{R}\). The inertia of \(H\) is denoted by \((\pi, \nu, \zeta)\), and its eigenvalues by \(\sigma_1 \leq \cdots \leq \sigma_\nu < 0 < \sigma_{n-\pi+1} \leq \cdots \leq \sigma_n\).

The following Theorem of Weyl (1912) can also be generalized to nonlinear small rank perturbations.
Low rank modification

We now consider a perturbation of a symmetric eigenvalue problem of the form

$$(A + \tau H)x = \lambda x$$

where $A, H \in \mathbb{R}^{n \times n}$ are symmetric, $H$ has small rank $r \ll n$, and $\tau \in \mathbb{R}$. The inertia of $H$ is denoted by $(\pi, \nu, \zeta)$, and its eigenvalues by

$$\sigma_1 \leq \cdots \leq \sigma_\nu < 0 < \sigma_{n-\pi+1} \leq \cdots \leq \sigma_n.$$ 

The following Theorem of Weyl (1912) can also be generalized to nonlinear small rank perturbations.

**Theorem**

Let $\beta_1 \leq \cdots \leq \beta_n$ denote the eigenvalues of $B := A + \tau H$. Then it holds that

$$\alpha_{i-\nu} \leq \beta_i \leq \alpha_{i+\pi} \quad \text{for} \quad \tau > 0, \quad i = 1, \ldots, n$$
Low rank modification

We now consider a perturbation of a symmetric eigenvalue problem of the form

$$(A + \tau H)x = \lambda x$$

where $A, H \in \mathbb{R}^{n \times n}$ are symmetric, $H$ has small rank $r \ll n$, and $\tau \in \mathbb{R}$. The inertia of $H$ is denoted by $(\pi, \nu, \zeta)$, and its eigenvalues by $\sigma_1 \leq \cdots < \sigma_{\nu} < 0 < \sigma_{n-\pi+1} \leq \cdots \leq \sigma_n$.

The following Theorem of Weyl (1912) can also be generalized to nonlinear small rank perturbations.

**Theorem**

Let $\beta_1 \leq \cdots \leq \beta_n$ denote the eigenvalues of $B := A + \tau H$. Then it holds that

$$\alpha_{i-\nu} \leq \beta_i \leq \alpha_{i+\pi} \quad \text{for } \tau > 0, \ i = 1, \ldots, n$$

These results can be generalized to small rank perturbations of nonlinear eigenvalue problems allowing for a minmax characterization of its eigenvalues.
1. Two examples

2. Variational characterization of eigenvalue problems
   - Overdamped Problems
   - Nonoverdamped problems

3. Applications
   - Sylvester’s law of inertia
   - Nonlinear low rank modification of a symmetric eigenvalue problem
   - **Safeguarded iteration**
   - Detecting hyperbolic matrix polynomials
   - Free vibrations of fluid-solid structures
   - Regularization of total least squares problems

4. Concluding remarks
**Safeguarded iteration**

**Require:** initial vector $x_0 \in D(p)$

1. compute $\sigma_0 = p(x_0)$
2. for $k = 1, 2, \ldots$ until convergence do
3. determine an eigenvector $x_k$ corresponding to the $j$ largest eigenvalues of $T(\sigma_{k-1})$
4. evaluate Rayleigh function $\sigma_k := p(x_k)$, i.e. solve $x_k^H T(\sigma_k) x_k = 0$ for $\sigma_k$
5. end for
Safeguarded iteration

**Require:** initial vector $x_0 \in D(p)$

1. compute $\sigma_0 = p(x_0)$
2. for $k = 1, 2, \ldots$ until convergence do
3. determine an eigenvec. $x_k$ corresp. to $j$ largest eigenval. of $T(\sigma_{k-1})$
4. evaluate Rayleigh funct. $\sigma_k := p(x_k)$, i.e. solve $x_k^H T(\sigma_k) x_k = 0$ for $\sigma_k$
5. end for

**Theorem:** Let $J \subset \mathbb{R}$ be an open interval, let $T(\lambda) \in \mathbb{C}^{n \times n}$, $\lambda \in J$, be a family of Hermitian matrices allowing for the minmax characterization.

(i) If $\lambda_1 := \inf_{x \in D(p)} p(x) \in J$ and $x_0 \in D(p)$ then the safeguarded iteration for $j = 1$ converges globally and monotonically decreasing to $\lambda_1$. 
Safeguarded iteration

**Require:** initial vector $x_0 \in \mathcal{D}(p)$
1. compute $\sigma_0 = p(x_0)$
2. **for** $k = 1, 2, \ldots$ **until** convergence **do**
3. determine an eigenvec. $x_k$ corresp. to $j$ largest eigenval. of $T(\sigma_{k-1})$
4. evaluate Rayleigh funct. $\sigma_k := p(x_k)$, i.e. solve $x_k^H T(\sigma_k) x_k = 0$ for $\sigma_k$
5. **end for**

**Theorem:** Let $J \subset \mathbb{R}$ be an open interval, let $T(\lambda) \in \mathbb{C}^{n \times n}$, $\lambda \in J$, be a family of Hermitian matrices allowing for the minmax characterization.

(i) If $\lambda_1 := \inf_{x \in \mathcal{D}(p)} p(x) \in J$ and $x_0 \in \mathcal{D}(p)$ then the safeguarded iteration for $j = 1$ converges globally and monotonically decreasing to $\lambda_1$.

(ii) If $T(\lambda)$ is holomorphic on a neighborhood $U \subset \mathbb{C}$ of a $j$th eigenvalue of $T(\cdot)$ and $\lambda_j$ is a simple eigenvalue, then the safeguarded iteration converges locally and quadratically to $\lambda_j$. 
Safeguarded iteration

**Require:** initial vector $x_0 \in \mathcal{D}(p)$

1. compute $\sigma_0 = p(x_0)$
2. for $k = 1, 2, \ldots$ until convergence do
3. determine an eigenvector $x_k$ corresponding to $j$ largest eigenvalues of $T(\sigma_{k-1})$
4. evaluate Rayleigh function $\sigma_k := p(x_k)$, i.e. solve $x_k^H T(\sigma_k) x_k = 0$ for $\sigma_k$
5. end for

**Theorem:** Let $J \subset \mathbb{R}$ be an open interval, let $T(\lambda) \in \mathbb{C}^{n \times n}$, $\lambda \in J$, be a family of Hermitian matrices allowing for the minmax characterization.

(i) If $\lambda_1 := \inf_{x \in \mathcal{D}(p)} p(x) \in J$ and $x_0 \in \mathcal{D}(p)$ then the safeguarded iteration for $j = 1$ converges globally and monotonically decreasing to $\lambda_1$.

(ii) If $T(\lambda)$ is holomorphic on a neighborhood $U \subset \mathbb{C}$ of a $j$th eigenvalue of $T(\cdot)$ and $\lambda_j$ is a simple eigenvalue, then the safeguarded iteration converges locally and quadratically to $\lambda_j$.

(iii) Under the conditions of (ii) the convergence is even cubic if $T'(\lambda)$ is positive definite for $\lambda \in U \cap J$, and $x_k$ in step 3 of the Algorithm is chosen to be an eigenvector corresponding to the $j$ largest eigenvalue of the generalized eigenproblem $T(\sigma_{k-1}) x = \mu T'(\sigma_{k-1}) x$. 
Outline

1. Two examples

2. Variational characterization of eigenvalue problems
   - Overdamped Problems
   - Nonoverdamped problems

3. Applications
   - Sylvester’s law of inertia
   - Nonlinear low rank modification of a symmetric eigenvalue problem
   - Safeguarded iteration
   - Detecting hyperbolic matrix polynomials
   - Free vibrations of fluid-solid structures
   - Regularization of total least squares problems

4. Concluding remarks
Hyperbolic (or more generally definite) matrix polynomials are important classes of Hermitian matrix polynomials. They allow for a definite linearization and can therefore be solved by the QR algorithm for Hermitian matrices.
Detecting hyperbolic matrix polynomials

Hyperbolic (or more generally definite) matrix polynomials are important classes of Hermitian matrix polynomials. They allow for a definite linearization and can therefore be solved by the QR algorithm for Hermitian matrices. They have only real eigenvalues which can be characterized as minmax and maxmin values of Rayleigh functionals, but there is no easy way to test if a given polynomial is hyperbolic or definite or not.

\[ Q(\lambda) \mathbf{x} = (\lambda^2 \mathbf{A} + \lambda \mathbf{B} + \mathbf{C}) \mathbf{x} = 0 \]

with \( \mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{C}^{n \times n} \) Hermitian, \( \mathbf{A} > 0 \), is hyperbolic, if

\[
\lambda^2 \mathbf{x}^H \mathbf{A} \mathbf{x} + \lambda \mathbf{x}^H \mathbf{B} \mathbf{x} + \mathbf{x}^H \mathbf{C} \mathbf{x} = 0
\]

for every \( \mathbf{x} \in \mathbb{C}^n \), \( \mathbf{x} \neq \mathbf{0} \) has two distinct real roots

\[
p_{\pm}(\mathbf{x}) = \frac{-\mathbf{x}^H \mathbf{B} \mathbf{x} \pm \sqrt{\left(\mathbf{x}^H \mathbf{B} \mathbf{x}\right)^2 - 4 \left(\mathbf{x}^H \mathbf{A} \mathbf{x}\right) \left(\mathbf{x}^H \mathbf{C} \mathbf{x}\right)}}{2 \mathbf{x}^H \mathbf{A} \mathbf{x}}.
\]
Detecting hyperbolic matrix polynomials

Hyperbolic (or more generally definite) matrix polynomials are important classes of Hermitian matrix polynomials. They allow for a definite linearization and can therefore be solved by the QR algorithm for Hermitian matrices. They have only real eigenvalues which can be characterized as minmax and maxmin values of Rayleigh functionals, but there is no easy way to test if a given polynomial is hyperbolic or definite or not.

Taking advantage of the safeguarded iteration which converges globally and monotonically to extreme eigenvalues we obtain an efficient algorithm that identifies hyperbolic or definite polynomials and enables the transformation to an equivalent definite linear pencil.
Detecting hyperbolic matrix polynomials

Hyperbolic (or more generally definite) matrix polynomials are important classes of Hermitian matrix polynomials. They allow for a definite linearization and can therefore be solved by the QR algorithm for Hermitian matrices. They have only real eigenvalues which can be characterized as minmax and maxmin values of Rayleigh functionals, but there is no easy way to test if a given polynomial is hyperbolic or definite or not.

Taking advantage of the safeguarded iteration which converges globally and monotonically to extreme eigenvalues we obtain an efficient algorithm that identifies hyperbolic or definite polynomials and enables the transformation to an equivalent definite linear pencil.

\[ Q(\lambda)x := (\lambda^2 A + \lambda B + C)x = 0 \]

with \( A, B, C \in \mathbb{C}^{n \times n} \) Hermitian, \( A > 0 \), is hyperbolic, if

\[ \lambda^2 x^H Ax + \lambda x^H Bx + x^H Cx = 0 \]

for every \( x \in \mathbb{C}^n, x \neq 0 \) has two distinct real roots

\[ p_{\pm}(x) = \left(-x^H Bx \pm \sqrt{(x^H Bx)^2 - 4(x^H Ax)(x^H Cx)})/2x^H Ax. \right] \]
The ranges \( J_{\pm} := p_{\pm}(\mathbb{C}^n \setminus \{0\}) \) are disjoint real intervals with \( \max J_- < \min J_+ \). \( Q(\lambda) \) is positive definite for \( \lambda < \min J_- \) and \( \lambda > \max J_+ \), and it is negative definite for \( \lambda \in (\max J_-, \min J_+) \).
The ranges $J_{\pm} := p_{\pm}(\mathbb{C}^n \setminus \{0\})$ are disjoint real intervals with $\max J_- < \min J_+$, $Q(\lambda)$ is positive definite for $\lambda < \min J_-$ and $\lambda > \max J_+$, and it is negative definite for $\lambda \in (\max J_-, \min J_+)$. Each of the intervals $J_-$ and $J_+$ contains $n$ eigenvalues

$$\lambda_n^- \leq \lambda_{n-1}^- \leq \cdots \leq \lambda_1^- < \lambda_1^+ \leq \cdots \leq \lambda_n^+$$

(15)

(notice that in $J_-$ the sign condition is satisfied for $-Q(\lambda)$, and therefore the smallest eigenvalue is an $n$th eigenvalue) which can be characterized as (Duffin 1955)

$$\lambda_j^- = \max_{\dim V=j} \min_{x \in V, x \neq 0} p_-(x), \quad \lambda_j^+ = \min_{\dim V=j} \max_{x \in V, x \neq 0} p_+(x).$$
Detecting hyperbolic matrix polynomials

The ranges $J_{\pm} := p_{\pm}(\mathbb{C}^n \setminus \{0\})$ are disjoint real intervals with $\max J_- < \min J_+$, $Q(\lambda)$ is positive definite for $\lambda < \min J_-$ and $\lambda > \max J_+$, and it is negative definite for $\lambda \in (\max J_-, \min J_+)$. Each of the intervals $J_-$ and $J_+$ contains $n$ eigenvalues

$$\lambda_n^- \leq \lambda_{n-1}^- \leq \cdots \leq \lambda_1^- < \lambda_1^+ \leq \cdots \leq \lambda_n^+$$

(15)

(notice that in $J_-$ the sign condition is satisfied for $-Q(\lambda)$, and therefore the smallest eigenvalue is an $n$th eigenvalue) which can be characterized as (Duffin 1955)

$$\lambda_j^- = \max_{\dim V=j} \min_{x \in V, x \neq 0} p_-(x), \quad \lambda_j^+ = \min_{\dim V=j} \max_{x \in V, x \neq 0} p_+(x).$$

The safeguarded iteration for $\lambda_1^+$ converges globally and monotonically decreasing for every initial vector $x_0 \in \mathbb{C}^n \setminus \{0\}$. This suggests the following Algorithm for detecting the hyperbolicity of $Q(\lambda)$. 
Detecting hyperbolic matrix polynomials

**Require:** initial vector $x_0 \neq 0$

1. if $d(x_0) = (x_0^H B x_0)^2 - 4(x_0^H A x_0)(x_0^H C x_0) < 0$ then
2. STOP: $Q(\lambda)$ is not hyperbolic
3. end if
4. determine $\sigma_0 = p_+(x_0)$
5. for $k = 1, 2, \ldots$ until convergence do
6. determine eigenvector $x_k$ of $Q(\sigma_{k-1})$ corresponding to its largest eigenvalue
7. if $d(x_k) = (x_k^H B x_k)^2 - 4(x_k^H A x_k)(x_k^H C x_k) < 0$ then
8. STOP: $Q(\lambda)$ is not hyperbolic
9. end if
10. determine $\sigma_k = p_+(x_k)$
11. if $\sigma_k \geq \sigma_{k-1}$ then
12. STOP: $Q(\lambda)$ is not hyperbolic
13. end if
14. if $Q(2\sigma_k - \sigma_{k-1})$ is negative definite then
15. STOP: $Q(\lambda)$ is hyperbolic
16. end if
17. end for
Detecting hyperbolic matrix polynomials

Once a parameter $\mu$ is found such that $Q(\mu)$ is negative definite the following transformation yields a definite linearization of $Q(\lambda)$. 

$$\tilde{Q}(\lambda) := Q(\lambda + \mu) = \lambda^2 A + \lambda (B + 2\mu A) + (C + \mu^2 A + \mu B) =: \lambda^2 \tilde{A} + \lambda \tilde{B} + \tilde{C}$$

where $\tilde{C} = \tilde{Q}(0) = Q(\mu)$ is negative definite.

Well known linearizations (cf. Lancaster 1966) $L_1(\lambda) := \lambda (\tilde{A} - \tilde{C}) + (\tilde{B} \tilde{C} \tilde{C})$ and $L_2(\lambda) := \lambda (\tilde{A} \tilde{A} - \tilde{C}) - (\tilde{A} \tilde{C} \tilde{C})$ of $\tilde{Q}(\lambda)$ are obviously definite.

Employing the Cholesky factorization of $\text{diag}\{\tilde{A}, -\tilde{C}\}$ it can be transformed to a standard eigenvalue problem and solved by the QR algorithm preserving the reality of its eigenvalues.
Detecting hyperbolic matrix polynomials

Once a parameter $\mu$ is found such that $Q(\mu)$ is negative definite the following transformation yields a definite linearization of $Q(\lambda)$.

Shifting $Q(\lambda)$ by $\mu$ yields a quadratic matrix polynomial

$$\tilde{Q}(\lambda) := Q(\lambda + \mu) = \lambda^2 A + \lambda(B + 2\mu A) + (C + \mu^2 A + \mu B) =: \lambda^2 \tilde{A} + \lambda \tilde{B} + \tilde{C}$$

where $\tilde{C} = \tilde{Q}(0) = Q(\mu)$ is negative definite.
Detecting hyperbolic matrix polynomials

Once a parameter $\mu$ is found such that $Q(\mu)$ is negative definite the following transformation yields a definite linearization of $Q(\lambda)$.

Shifting $Q(\lambda)$ by $\mu$ yields a quadratic matrix polynomial

$\tilde{Q}(\lambda) := Q(\lambda + \mu) = \lambda^2 A + \lambda(B + 2\mu A) + (C + \mu^2 A + \mu B) =: \lambda^2 \tilde{A} + \lambda \tilde{B} + \tilde{C}$

where $\tilde{C} = \tilde{Q}(0) = Q(\mu)$ is negative definite.

Well known linearizations (cf. Lancaster 1966)

$L_1(\lambda) := \lambda \begin{pmatrix} \tilde{A} & 0 \\ 0 & -\tilde{C} \end{pmatrix} + \begin{pmatrix} \tilde{B} & \tilde{C} \\ \tilde{C} & 0 \end{pmatrix}$

and

$L_2(\lambda) := \lambda \begin{pmatrix} 0 & \tilde{A} \\ \tilde{A} & \tilde{B} \end{pmatrix} + \begin{pmatrix} -\tilde{A} & 0 \\ 0 & \tilde{C} \end{pmatrix}$

of $\tilde{Q}(\lambda)$ are obviously definite.
Detecting hyperbolic matrix polynomials

Once a parameter $\mu$ is found such that $Q(\mu)$ is negative definite the following transformation yields a definite linearization of $Q(\lambda)$.

Shifting $Q(\lambda)$ by $\mu$ yields a quadratic matrix polynomial

$$\tilde{Q}(\lambda) := Q(\lambda + \mu) = \lambda^2 A + \lambda(B + 2\mu A) + (C + \mu^2 A + \mu B) =: \lambda^2 \tilde{A} + \lambda \tilde{B} + \tilde{C}$$

where $\tilde{C} = \tilde{Q}(0) = Q(\mu)$ is negative definite.

Well known linearizations (cf. Lancaster 1966)

$$L_1(\lambda) := \lambda \begin{pmatrix} \tilde{A} & 0 \\ 0 & -\tilde{C} \end{pmatrix} + \begin{pmatrix} \tilde{B} & \tilde{C} \\ \tilde{C} & 0 \end{pmatrix} \quad \text{and} \quad L_2(\lambda) := \lambda \begin{pmatrix} 0 & \tilde{A} \\ \tilde{A} & \tilde{B} \end{pmatrix} + \begin{pmatrix} -\tilde{A} & 0 \\ 0 & \tilde{C} \end{pmatrix}$$

of $\tilde{Q}(\lambda)$ are obviously definite.

Employing the Cholesky factorization of $\text{diag}\{\tilde{A}, -\tilde{C}\}$ it can be transformed to a standard eigenvalue problem and solved by the QR algorithm preserving the reality of its eigenvalues.
The matrices $Q(\sigma_k)$ converge as the sequence $\{\sigma_k\}$ converges. Hence, one should reuse as much information as possible from previous steps when solving the eigenvalue problem $Q(\sigma_{k-1})x_k = \mu x_k$. 
The matrices $Q(\sigma_k)$ converge as the sequence $\{\sigma_k\}$ converges. Hence, one should reuse as much information as possible from previous steps when solving the eigenvalue problem $Q(\sigma_{k-1})x_k = \mu x_k$.

Rayleigh quotient iteration or (implicitly restarted) Krylov subspace methods are able to use the eigenvector $x_{k-1}$ of the last step as initial vector.
The matrices $Q(\sigma_k)$ converge as the sequence $\{\sigma_k\}$ converges. Hence, one should reuse as much information as possible from previous steps when solving the eigenvalue problem $Q(\sigma_{k-1})x_k = \mu x_k$.

Rayleigh quotient iteration or (implicitly restarted) Krylov subspace methods are able to use the eigenvector $x_{k-1}$ of the last step as initial vector.

The nonlinear Arnoldi method (Voss 2004) is an iterative projection method which can use the entire search space of the previous step as initial information.
The matrices $Q(\sigma_k)$ converge as the sequence $\{\sigma_k\}$ converges. Hence, one should reuse as much information as possible from previous steps when solving the eigenvalue problem $Q(\sigma_{k-1})x_k = \mu x_k$.

Rayleigh quotient iteration or (implicitly restarted) Krylov subspace methods are able to use the eigenvector $x_{k-1}$ of the last step as initial vector.

The nonlinear Arnoldi method (Voss 2004) is an iterative projection method which can use the entire search space of the previous step as initial information.

The technique can be generalized to definite matrix polynomials (hyperbolic, but without $A > 0$) and matrix polynomials of both types of higher degree.
Outline

1. Two examples

2. Variational characterization of eigenvalue problems
   - Overdamped Problems
   - Nonoverdamped problems

3. Applications
   - Sylvester’s law of inertia
   - Nonlinear low rank modification of a symmetric eigenvalue problem
   - Safeguarded iteration
   - Detecting hyperbolic matrix polynomials
   - Free vibrations of fluid-solid structures
   - Regularization of total least squares problems

4. Concluding remarks
Free vibrations of fluid-solid structures

Recall

\[ T(\lambda)u := -K_f u + \lambda M_f u + \sum_{n=1}^{\infty} \frac{\lambda}{\sigma_n - \lambda} C_n u, \quad C_n u := \langle u_n, Cu \rangle C' u_n. \quad (17) \]
Recall

\[ T(\lambda)u := -K_f u + \lambda M_f u + \sum_{n=1}^{\infty} \frac{\lambda}{\sigma_n - \lambda} C_n u, \quad C_n u := \langle u_n, Cu \rangle C'u_n. \]  

(17)

\( f(\cdot, u) = \langle T(\cdot)u, u \rangle \) is strictly monotonically increasing and \( f(0, u) < 0 \) for every \( u \neq 0 \) and \( \lambda \neq \sigma_\ell, \ell \in \mathbb{N}. \)

Hence, in each interval \( J_\ell := (\sigma_{\ell-1}, \sigma_\ell) \) (with \( \sigma_0 := 0 \)) the eigenvalues allow for a minmax and maxmin characterization.
Recall

\[ T(\lambda)u := -K_f u + \lambda M_f u + \sum_{n=1}^{\infty} \frac{\lambda}{\sigma_n - \lambda} C_n u, \quad C_n u := \langle u_n, Cu \rangle C'u_n. \quad (17) \]

\( f(\cdot, u) = \langle T(\cdot)u, u \rangle \) is strictly monotonically increasing and \( f(0, u) < 0 \) for every \( u \neq 0 \) and \( \lambda \neq \sigma_{\ell}, \ell \in \mathbb{N} \).

Hence, in each interval \( J_\ell := (\sigma_{\ell-1}, \sigma_\ell) \) (with \( \sigma_0 := 0 \)) the eigenvalues allow for a minmax and maxmin characterization.

For \( \kappa \in J_\ell \) consider the linear comparison problem

\[ (K_f + \sum_{j=1}^{\ell-1} \frac{\kappa}{\kappa - \sigma_j} C_j)u = \mu(M_f + \sum_{j=\ell}^{\infty} \frac{1}{\sigma_j - \kappa} C_j)u \quad (18) \]

and denote by \( R_\kappa(u) \) the corresponding Rayleigh quotient.
Let $\mu_n(\kappa)$ be the $n$ smallest eigenvalue of the comparison problem (18). Then $\kappa \mapsto \mu_n(\kappa)$ is monotonically increasing and $\lambda$ is an $n$-th eigenvalue of the rational eigenproblem (17) if and only if $\lambda$ is a fixed point of $\mu_n(\cdot)$. 
Let $\mu_n(\kappa)$ be the $n$ smallest eigenvalue of the comparison problem (18). Then $\kappa \mapsto \mu_n(\kappa)$ is monotonically increasing and $\lambda$ is an $n$-th eigenvalue of the rational eigenproblem (17) if and only if $\lambda$ is a fixed point of $\mu_n(\cdot)$.

**Theorem**

Let $\sigma_{\ell-1} < \kappa_1 < \kappa_2 < \sigma_{\ell}$ and $N(\kappa) := \max\{n \in \mathbb{N} : \mu_n(\kappa) \leq \kappa\}$. Then the nonlinear eigenvalue problem (17) has exactly $N(\kappa_2) - N(\kappa_1)$ eigenvalues in $(\kappa_1, \kappa_2]$. 
Bounds to Rayleigh functional

Lemma
Assume that $\kappa \in J_\ell$ and $R_\kappa(u) \in J_\ell$ for some $u \in \mathcal{H}$, $u \neq 0$. Then $u \in D_\ell$, and

$$
\min\{\kappa, R_\kappa(u)\} \leq p_\ell(u) \leq \max\{\kappa, R_\kappa(u)\}.
$$
Lemma
Assume that $\kappa \in J_\ell$ and $R_\kappa(u) \in J_\ell$ for some $u \in \mathcal{H}$, $u \neq 0$. Then $u \in D_\ell$, and

$$\min\{\kappa, R_\kappa(u)\} \leq p_\ell(u) \leq \max\{\kappa, R_\kappa(u)\}.$$

Proof:

$$f(R_\kappa(u), u) = \cdots =$$

$$= \sum_{j=1}^{\ell-1} \frac{\sigma_\ell(R_\kappa(u) - \kappa)}{(\sigma_\ell - R_\kappa(u))(\sigma_\ell - \kappa)} \langle u, C_j u \rangle + \sum_{j=\ell}^{\infty} \frac{R_\kappa(u)(R_\kappa(u) - \kappa)}{(\sigma_\ell - R_\kappa(u))(\sigma_\ell - \kappa)} \langle u, C_j u \rangle$$

i.e. $f(R_\kappa(u), u) \geq 0$ for $R_\kappa(u) \geq \kappa$, and $f(R_\kappa(u), u) \leq 0$ for $R_\kappa(u) \leq \kappa$. 
Bounds to Rayleigh functional

Lemma
Assume that $\kappa \in J_\ell$ and $R_\kappa (u) \in J_\ell$ for some $u \in \mathcal{H}$, $u \neq 0$. Then $u \in D_\ell$, and

$$\min \{ \kappa, R_\kappa (u) \} \leq p_\ell (u) \leq \max \{ \kappa, R_\kappa (u) \}.$$ 

Proof:

$$f(R_\kappa (u), u) = \cdots =$$

$$= \sum_{j=1}^{\ell-1} \frac{\sigma_\ell (R_\kappa (u) - \kappa)}{(\sigma_\ell - R_\kappa (u))(\sigma_\ell - \kappa)} \langle u, C_j u \rangle + \sum_{j=\ell}^{\infty} \frac{R_\kappa (u)(R_\kappa (u) - \kappa)}{(\sigma_\ell - R_\kappa (u))(\sigma_\ell - \kappa)} \langle u, C_j u \rangle$$

i.e. $f(R_\kappa (u), u) \geq 0$ for $R_\kappa (u) \geq \kappa$, and $f(R_\kappa (u), u) \leq 0$ for $R_\kappa (u) \leq \kappa$. And

$$f(\kappa, u) = \cdots = (\kappa - R_\kappa (u)) \left( \sum_{j=\ell}^{\infty} \frac{1}{\sigma_j - \kappa} \langle u, C_j u \rangle + \langle u, M_f u \rangle \right),$$

i.e. $f(R_\kappa (u), u) \leq 0$ for $R_\kappa (u) \geq \kappa$, and $f(R_\kappa (u), u) \geq 0$ for $R_\kappa (u) \leq \kappa$. 
Lower end of spectrum

Theorem
Assume that for $\kappa \in J_1 = (0, \sigma_1)$ the comparison problem (18) has $r$ eigenvalues $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_r$ in $J_1$. Then the rational eigenvalue problem (17) has $r$ eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_r$ (where $\lambda_j$ is a $j$-th eigenvalue), and it holds that

$$\min\{\mu_j, \kappa\} \leq \lambda_j \leq \max\{\mu_j, \kappa\} \quad \text{for} \ j = 1, \ldots, r.$$  \hspace{1cm} (19)
Theorem
Assume that for $\kappa \in J_1 = (0, \sigma_1)$ the comparison problem (18) has $r$ eigenvalues $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_r$ in $J_1$.
Then the rational eigenvalue problem (17) has $r$ eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_r$ (where $\lambda_j$ is a $j$-th eigenvalue), and it holds that

$$\min\{\mu_j, \kappa\} \leq \lambda_j \leq \max\{\mu_j, \kappa\} \quad \text{for } j = 1, \ldots, r. \quad (19)$$

Proof: $\inf_{u \in D_1} p_1(u) > 0$ follows by contradiction from a comparison with (18) for $\kappa = \mu_1$
Lower end of spectrum

Theorem
Assume that for $\kappa \in J_1 = (0, \sigma_1)$ the comparison problem (18) has $r$ eigenvalues $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_r$ in $J_1$.
Then the rational eigenvalue problem (17) has $r$ eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_r$ (where $\lambda_j$ is a $j$-th eigenvalue), and it holds that

$$\min\{\mu_j, \kappa\} \leq \lambda_j \leq \max\{\mu_j, \kappa\} \quad \text{for } j = 1, \ldots, r.$$  \hfill (19)

Proof: $\inf_{u \in D_1} p_1(u) > 0$ follows by contradiction from a comparison with (18) for $\kappa = \mu_1$

The existence of $\lambda_j, j = 1, \ldots, r$ follows from minmax Theorem for extreme eigenvalues since the invariant subspace $W$ of (18) corresponding to $\mu_r$ is contained in $D_1$ and $\max_{v \in W, x \neq 0} R_{\kappa}(v) \in J_1$. 

TUHH
Heinrich Voss
Variational Principles for Nonlinear EVPs
Isfahan, July, 2016
55 / 74
Applications

Free vibrations of fluid-solid structures

Lower end of spectrum

Theorem
Assume that for $\kappa \in J_1 = (0, \sigma_1)$ the comparison problem (18) has $r$ eigenvalues $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_r$ in $J_1$. Then the rational eigenvalue problem (17) has $r$ eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_r$ (where $\lambda_j$ is a $j$-th eigenvalue), and it holds that

$$\min\{\mu_j, \kappa\} \leq \lambda_j \leq \max\{\mu_j, \kappa\} \quad \text{for} \quad j = 1, \ldots, r. \quad (19)$$

Proof: $\inf_{u \in D_1} p_1(u) > 0$ follows by contradiction from a comparison with (18) for $\kappa = \mu_1$

The existence of $\lambda_j, j = 1, \ldots, r$ follows from minmax Theorem for extreme eigenvalues since the invariant subspace $W$ of (18) corresponding to $\mu_r$ is contained in $D_1$ and $\max_{v \in W, x \neq 0} R_\kappa(v) \in J_1$.

(19) follows from the last Lemma and the minmax Theorem for extreme eigenvalues by comparison with an invariant subspace corresponding to $\mu_1, \ldots, \mu_j$. 
Interior intervals $J_\ell$

**Theorem**

Assume that for $\kappa \in J_\ell$ the $n$-th eigenvalue of the comparison problem (18) satisfies $\mu_n(\kappa) \in (\sigma_{\ell-1}, \sigma_\ell)$. Then the rational eigenvalue problem (17) has an $n$-th eigenvalue $\lambda_n \in (\sigma_{\ell-1}, \sigma_\ell)$, and it holds that

$$
\min\{\kappa, \mu_n(\kappa)\} \leq \lambda_n \leq \max\{\kappa, \mu_n(\kappa)\}.
$$

(20)
Theorem
Assume that for \( \kappa \in J_\ell \) the \( n \)-th eigenvalue of the comparison problem (18) satisfies \( \mu_n(\kappa) \in (\sigma_{\ell-1}, \sigma_\ell) \). Then the rational eigenvalue problem (17) has an \( n \)-th eigenvalue \( \lambda_n \in (\sigma_{\ell-1}, \sigma_\ell) \), and it holds that

\[
\min\{\kappa, \mu_n(\kappa)\} \leq \lambda_n \leq \max\{\kappa, \mu_n(\kappa)\}.
\] (20)

Proof: Show that

- There exist \( V \) with \( \dim V = n \) such that \( V \cap D_\ell \neq \emptyset \) and \( \sup_{u \in V \cap D_\ell} p_\ell(u) \leq \max\{\kappa, \mu_n(\kappa)\} \)
**Theorem**
Assume that for $\kappa \in J_\ell$ the $n$-th eigenvalue of the comparison problem (18) satisfies $\mu_n(\kappa) \in (\sigma_{\ell-1}, \sigma_\ell)$. Then the rational eigenvalue problem (17) has an $n$-th eigenvalue $\lambda_n \in (\sigma_{\ell-1}, \sigma_\ell)$, and it holds that

$$\min\{\kappa, \mu_n(\kappa)\} \leq \lambda_n \leq \max\{\kappa, \mu_n(\kappa)\}. \quad (20)$$

**Proof:** Show that
- There exist $V$ with dim $V = n$ such that $V \cap D_\ell \neq \emptyset$ and $\sup_{u \in V \cap D_\ell} p_\ell(u) \leq \max\{\kappa, \mu_n(\kappa)\}$
- $\sup_{u \in V \cap D_\ell} p_\ell(u) \geq \min\{\kappa, \mu_n(\kappa)\}$ for every $V$ with dim $V = n$ such that $V \cap D_\ell \neq \emptyset$.  


Theorem
Assume that for $\kappa \in J_\ell$ the $n$-th eigenvalue of the comparison problem (18) satisfies $\mu_n(\kappa) \in (\sigma_{\ell-1}, \sigma_\ell)$. Then the rational eigenvalue problem (17) has an $n$-th eigenvalue $\lambda_n \in (\sigma_{\ell-1}, \sigma_\ell)$, and it holds that

$$\min\{\kappa, \mu_n(\kappa)\} \leq \lambda_n \leq \max\{\kappa, \mu_n(\kappa)\}. \quad (20)$$

Proof: Show that

- There exist $V$ with $\dim V = n$ such that $V \cap D_\ell \neq \emptyset$ and $\sup_{u \in V \cap D_\ell} p_\ell(u) \leq \max\{\kappa, \mu_n(\kappa)\}$
- $\sup_{u \in V \cap D_\ell} p_\ell(u) \geq \min\{\kappa, \mu_n(\kappa)\}$ for every $V$ with $\dim V = n$ such that $V \cap D_\ell \neq \emptyset$.

Then

$$\lambda_n = \inf_{\dim V=n, \ V\cap D_\ell\neq\emptyset} \sup_{u \in V \cap D_\ell} p_\ell(u) \in J_\ell,$$

i.e. $\lambda_n$ is an $n$-th eigenvalue of (17) and inequalities (20) holds.
Outline

1. Two examples

2. Variational characterization of eigenvalue problems
   - Overdamped Problems
   - Nonoverdamped problems

3. Applications
   - Sylvester’s law of inertia
   - Nonlinear low rank modification of a symmetric eigenvalue problem
   - Safeguarded iteration
   - Detecting hyperbolic matrix polynomials
   - Free vibrations of fluid-solid structures
   - Regularization of total least squares problems

4. Concluding remarks
Total Least Squares Problem

The ordinary Least Squares (LS) method assumes that the system matrix $A$ of a linear model is error free, and all errors are confined to the right hand side $b$. In practical applications it may happen that all data are contaminated by noise.
Total Least Squares Problem

The ordinary Least Squares (LS) method assumes that the system matrix $A$ of a linear model is error free, and all errors are confined to the right hand side $b$. In practical applications it may happen that all data are contaminated by noise.

If the true values of the observed variables satisfy linear relations, and if the errors in the observations are independent random variables with zero mean and equal variance, then the total least squares (TLS) approach often gives better estimates than LS.
The ordinary Least Squares (LS) method assumes that the system matrix $A$ of a linear model is error free, and all errors are confined to the right hand side $b$. In practical applications it may happen that all data are contaminated by noise.

If the true values of the observed variables satisfy linear relations, and if the errors in the observations are independent random variables with zero mean and equal variance, then the total least squares (TLS) approach often gives better estimates than LS.

Given $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $m \geq n$

Find $\tilde{A} \in \mathbb{R}^{m \times n}$, $\tilde{b} \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$ such that

$$\|(A, b) - (\tilde{A}, \tilde{b})\|_F^2 = \min$$

subject to $\tilde{A}x = \tilde{b}$,

where $\| \cdot \|_F$ denotes the Frobenius norm.
If $A$ and $(A, b)$ are ill-conditioned, regularization is necessary.
If $A$ and $(A, b)$ are ill-conditioned, regularization is necessary.

Let $L \in \mathbb{R}^{p \times n}$, $p \leq n$ and $\delta > 0$. Then the quadratically constrained formulation of the Regularized Total Least Squares (RTLS) problems reads:

Find $\tilde{A} \in \mathbb{R}^{m \times n}$, $\tilde{b} \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$ such that

$$\| (A, b) - (\tilde{A}, \tilde{b}) \|^2_F = \min! \quad \text{subject to} \quad \tilde{A}x = \tilde{b}, \quad \|Lx\|^2_2 \leq \delta^2.$$
Regularized Total Least Squares Problem

If $A$ and $(A, b)$ are ill-conditioned, regularization is necessary.

Let $L \in \mathbb{R}^{p \times n}$, $p \leq n$ and $\delta > 0$. Then the quadratically constrained formulation of the Regularized Total Least Squares (RTLS) problems reads:

Find $\tilde{A} \in \mathbb{R}^{m \times n}$, $\tilde{b} \in \mathbb{R}^{m}$ and $x \in \mathbb{R}^{n}$ such that

$$
\|(A, b) - (\tilde{A}, \tilde{b})\|^2_F = \min \quad \text{subject to} \quad \tilde{A}x = \tilde{b}, \quad \|Lx\|^2_2 \leq \delta^2.
$$

Using the orthogonal distance this problems can be rewritten as (cf. Van Huffel, Vandevalle)

Find $x \in \mathbb{R}^{n}$ such that

$$
\frac{\|Ax - b\|^2_2}{1 + \|x\|^2_2} = \min \quad \text{subject to} \quad \|Lx\|^2_2 \leq \delta^2.
$$
If $\delta > 0$ is chosen small enough, the constraint $\|Lx\|_2^2 \leq \delta^2$ is active, and the RTLS problem reads:

Find $x \in \mathbb{R}^n$ such that

$$\frac{\|Ax - b\|_2^2}{1 + \|x\|_2^2} = \min \quad \text{subject to} \quad \|Lx\|_2^2 = \delta^2.$$
If $\delta > 0$ is chosen small enough, the constraint $\|Lx\|_2^2 \leq \delta^2$ is active, and the RTLS problem reads

Find $x \in \mathbb{R}^n$ such that

$$\frac{\|Ax - b\|_2^2}{1 + \|x\|_2^2} = \min \quad \text{subject to} \quad \|Lx\|_2^2 = \delta^2.$$ 

The first order optimality conditions are

$$B(x)x + \lambda L^T Lx = d(x), \quad \|Lx\|_2^2 = \delta^2$$

where

$$B(x) = \frac{1}{1 + \|x\|_2^2} \left(A^T A - f(x) I_n\right), \quad f(x) = \frac{\|Ax - b\|_2^2}{1 + \|x\|_2^2}, \quad d(x) = \frac{A^T b}{1 + \|x\|_2^2}.$$
A quadratic eigenproblem


The first order conditions can be solved via the maximal positive eigenvalue and corresponding eigenvector of a quadratic eigenproblem

\[ ((W + \lambda I)^2 - \delta^{-2}hh^T)u = 0 \quad \text{(QEP)} \]

where \( W \in \mathbb{R}^{p \times p} \) is symmetric, and \( h \in \mathbb{R}^p \).
A quadratic eigenproblem


The first order conditions can be solved via the maximal positive eigenvalue and corresponding eigenvector of a quadratic eigenproblem

\[(W + \lambda I)^2 - \delta^{-2} hh^T) u = 0 \quad (QEP)\]

where \( W \in \mathbb{R}^{p \times p} \) is symmetric, and \( h \in \mathbb{R}^p \).

\[f(\lambda, x) = x^H T(\lambda) x = \lambda^2 \|x\|^2_2 + 2\lambda x^H W x + \|Wx\|^2_2 - |x^H h|^2 / \delta^2, \quad x \neq 0\]

is a parabola which attains its minimum at

\[\lambda = -\frac{x^H W x}{x^H x}.\]
A quadratic eigenproblem


The first order conditions can be solved via the maximal positive eigenvalue and corresponding eigenvector of a quadratic eigenproblem

\[
((W + \lambda I)^2 - \delta^{-2} hh^T)u = 0 \quad (QEP)
\]

where \( W \in \mathbb{R}^{p \times p} \) is symmetric, and \( h \in \mathbb{R}^p \).

\[
f(\lambda, x) = x^H T(\lambda) x = \lambda^2 \|x\|^2 + 2 \lambda x^H W x + \|W x\|^2 - |x^H h|^2 / \delta^2, \ x \neq 0
\]

is a parabola which attains its minimum at

\[
\lambda = -\frac{x^H W x}{x^H x}.
\]

Let \( J = (-\lambda_{\text{min}}, \infty) \) where \( \lambda_{\text{min}} \) is the minimum eigenvalue of \( W \). Then \( f(\lambda, x) = 0 \) has at most one solution \( p(x) \in J \) for every \( x \neq 0 \), the Rayleigh functional \( p \) of (QEP) is defined on \( J \), and the general conditions are satisfied.
Characterization of maximal real eigenvalue

Let \( x_{\text{min}} \) be an eigenvector of \( W \) corresponding to \( \lambda_{\text{min}} \). Then

\[
f(-\lambda_{\text{min}}, x_{\text{min}}) = x_{\text{min}}^H (W - \lambda_{\text{min}})^2 x_{\text{min}} - |x_{\text{min}}^H h|^2 / \delta^2 = -|x_{\text{min}}^H h|^2 / \delta^2 \leq 0
\]
Let $x_{\text{min}}$ be an eigenvector of $W$ corresponding to $\lambda_{\text{min}}$. Then

$$f(-\lambda_{\text{min}}, x_{\text{min}}) = x_{\text{min}}^H(W - \lambda_{\text{min}})^2 x_{\text{min}} - |x_{\text{min}}^H h|^2 / \delta^2 = -|x_{\text{min}}^H h|^2 / \delta^2 \leq 0$$

Hence, if $x_{\text{min}}^H h \neq 0$ then $x_{\text{min}} \in D$. 
Characterization of maximal real eigenvalue

Let $x_{\text{min}}$ be an eigenvector of $W$ corresponding to $\lambda_{\text{min}}$. Then

$$f(-\lambda_{\text{min}}, x_{\text{min}}) = x_{\text{min}}^H (W - \lambda_{\text{min}})^2 x_{\text{min}} - |x_{\text{min}}^H h|^2 / \delta^2 = -|x_{\text{min}}^H h|^2 / \delta^2 \leq 0$$

Hence, if $x_{\text{min}}^H h \neq 0$ then $x_{\text{min}} \in D$.

If $x_{\text{min}}^H h = 0$, and the minimum eigenvalue $\mu_{\text{min}}$ of $T(-\lambda_{\text{min}})$ is negative, then for the corresponding eigenvector $y_{\text{min}}$ it holds

$$f(-\lambda_{\text{min}}, y_{\text{min}}) = y_{\text{min}}^H T(-\lambda_{\text{min}}) y_{\text{min}} = \mu_{\text{min}} \|y_{\text{min}}\|_2^2 < 0,$$

and $y_{\text{min}} \in D$. 
Characterization of maximal real eigenvalue

Let $x_{\text{min}}$ be an eigenvector of $W$ corresponding to $\lambda_{\text{min}}$. Then

$$f(-\lambda_{\text{min}}, x_{\text{min}}) = x_{\text{min}}^H(W - \lambda_{\text{min}})^2 x_{\text{min}} - |x_{\text{min}}^H h|^2 / \delta^2 = -|x_{\text{min}}^H h|^2 / \delta^2 \leq 0$$

Hence, if $x_{\text{min}}^H h \neq 0$ then $x_{\text{min}} \in D$.

If $x_{\text{min}}^H h = 0$, and the minimum eigenvalue $\mu_{\text{min}}$ of $T(-\lambda_{\text{min}})$ is negative, then for the corresponding eigenvector $y_{\text{min}}$ it holds

$$f(-\lambda_{\text{min}}, y_{\text{min}}) = y_{\text{min}}^H T(-\lambda_{\text{min}}) y_{\text{min}} = \mu_{\text{min}} \|y_{\text{min}}\|^2 \leq 0,$$

and $y_{\text{min}} \in D$.

If $x_{\text{min}}^H h = 0$, and $T(-\lambda_{\text{min}})$ is positive semi-definite, then

$$f(-\lambda_{\text{min}}, x) = x^H T(-\lambda_{\text{min}}) x \geq 0 \quad \text{for every } x \neq 0,$$

and $D = \emptyset$. 
Assume that $D \neq \emptyset$. For $x^H h = 0$ it holds that

$$f(\lambda, x) = \|(W + \lambda I)x\|_2^2 > 0 \quad \text{for every } \lambda \in J,$$

i.e. $x \notin D$. 
Assume that $D \neq \emptyset$. For $x^H h = 0$ it holds that

$$f(\lambda, x) = \|(W + \lambda I)x\|_2^2 > 0 \quad \text{for every } \lambda \in J,$$

i.e. $x \not\in D$.

Hence, $D$ does not contain a two-dimensional subspace of $\mathbb{R}^n$, and therefore $J$ contains at most one eigenvalue of (QEP).
Assume that $D \neq \emptyset$. For $x^H h = 0$ it holds that
\[ f(\lambda, x) = \| (W + \lambda I) x \|^2 > 0 \quad \text{for every } \lambda \in J, \]
i.e. $x \not\in D$.

Hence, $D$ does not contain a two-dimensional subspace of $\mathbb{R}^n$, and therefore $J$ contains at most one eigenvalue of (QEP).

If $\lambda \in \mathbb{C}$ is a non-real eigenvalue of (QEP) and $x$ a corresponding eigenvector, then
\[ x^H T(\lambda) x = \lambda^2 \| x \|^2 + 2\lambda x^H W x + \| W x \|^2 - |x^H h|^2 / \delta^2 = 0. \]
Hence, the real part of $\lambda$ satisfies
\[ \text{real}(\lambda) = -\frac{x^H W x}{x^H x} \leq -\lambda_{\min}. \]
Theorem

Let $\lambda_{\min}$ be the minimal eigenvalue of $W$, and $x_{\min}$ be a corresponding eigenvector.
Theorem

Let $\lambda_{\min}$ be the minimal eigenvalue of $W$, and $x_{\min}$ be a corresponding eigenvector. If $x_{\min}^H h = 0$ and $T(-\lambda_{\min})$ is positive semi-definite, then $\hat{\lambda} := -\lambda_{\min}$ is the maximal real eigenvalue of (QEP).
Let $\lambda_{\text{min}}$ be the minimal eigenvalue of $W$, and $x_{\text{min}}$ be a corresponding eigenvector.

If $x_{\text{min}}^H h = 0$ and $T(-\lambda_{\text{min}})$ is positive semi-definite, then $\hat{\lambda} := -\lambda_{\text{min}}$ is the maximal real eigenvalue of $(QEP)$.

Otherwise, the maximal real eigenvalue is the unique eigenvalue $\hat{\lambda}$ of $(QEP)$ in $J = (-\lambda_{\text{min}}, \infty)$, and it holds

$$\hat{\lambda} = \max_{x \in D} p(x).$$
Let $\lambda_{\min}$ be the minimal eigenvalue of $W$, and $x_{\min}$ be a corresponding eigenvector.

If $x_{\min}^H h = 0$ and $T(-\lambda_{\min})$ is positive semi-definite, then $\hat{\lambda} := -\lambda_{\min}$ is the maximal real eigenvalue of (QEP).

Otherwise, the maximal real eigenvalue is the unique eigenvalue $\hat{\lambda}$ of (QEP) in $J = (-\lambda_{\min}, \infty)$, and it holds

$$\hat{\lambda} = \max_{x \in D} p(x).$$

$\hat{\lambda}$ is the right most eigenvalue of (QEP), i.e.

$$\text{real}(\lambda) \leq -\lambda_{\min} \leq \hat{\lambda} \quad \text{for every eigenvalue } \lambda \text{ of (QEP)}.$$
Example

shaw(2000): eigenvalues of \((W + \lambda I)^2 - hh^T \delta^2)y = 0\)
Example: close up

\text{shaw(2000)}: eigenvalues of \((W+\lambda I)^2hh^T/\delta^2)y=0\n
-1 -0.5 0 0.5 1

imaginary part

real part

-0.05 -0.04 -0.03 -0.02 -0.01 0 0.01 0.02 0.03 0.04 0.05
In the general situation the maximum in

\[ \lambda_j = \max_{\dim V = j, \ V \cap D \neq \emptyset} \inf_{v \in V, \ v \neq 0} p(v) \]

is attained by the invariant subspace of \( T(\lambda_j) \) corresponding to the \( j \) smallest eigenvalues, and the minimum by every eigenvector corresponding to the eigenvalue 0. This suggests the following algorithm for computing an \( \ell \)-th eigenvalue

1: Start with an approximation \( \mu_1 \) to the \( \ell \)-th eigenvalue of \( T(\lambda) x = 0 \)
2: for \( k = 1, 2, ... \) until convergence
3: determine eigenvector \( u \) corresponding to the \( \ell \)-smallest eigenvalue of \( T(\mu_k) \)
4: evaluate \( \mu_{k+1} = p(u) \)
5: end for

Convergence is quadratic to simple eigenvalues; global to the first eigenvalue, local to further eigenvalues.
Safeguarded iteration

In the general situation the maximum in

$$\lambda_j = \max_{\dim V = j, \ V \cap D \neq \emptyset} \inf_{v \in V, \ v \neq 0} p(v)$$

is attained by the invariant subspace of $T(\lambda_j)$ corresponding to the $j$ smallest eigenvalues, and the minimum by every eigenvector corresponding to the eigenvalue 0. This suggests the following algorithm for computing an $\ell$-th eigenvalue

**Safeguarded iteration**

1. Start with an approximation $\mu_1$ to the $\ell$-th eigenvalue of $T(\lambda)x = 0$
2. for $k = 1, 2, \ldots$ until convergence do
3. determine eigenvector $u$ corresponding to the $\ell$-smallest eigenvalue of $T(\mu_k)$
4. evaluate $\mu_{k+1} = p(u)$
5. end for
Safeguarded iteration

In the general situation the maximum in

$$\lambda_j = \max_{\dim V=j, \ V \cap D \neq \emptyset} \inf_{v \in V, v \neq 0} p(v)$$

is attained by the invariant subspace of $T(\lambda_j)$ corresponding to the $j$ smallest eigenvalues, and the minimum by every eigenvector corresponding to the eigenvalue 0. This suggests the following algorithm for computing an $\ell$-th eigenvalue

**Safeguarded iteration**

1: Start with an approximation $\mu_1$ to the $\ell$-th eigenvalue of $T(\lambda)x = 0$
2: for $k = 1, 2, \ldots$ until convergence do
3: determine eigenvector $u$ corresponding to the $\ell$-smallest eigenvalue of $T(\mu_k)$
4: evaluate $\mu_{k+1} = p(u)$
5: end for

Convergence is quadratic to simple eigenvalues; global to the first eigenvalue, local to further eigenvalues.
Solving first order condition

1: Start with approximation $x^0$
2: compute $B_0 = B(x^0)$, $d_0 = d(x^0)$
3: for $k = 0, 1, 2, \ldots$ until convergence do
4: compute $W_k$, $h_k$
5: $[v, \mu] = \text{eigs}(W, 1, 'SA')$
6: compute $\lambda = p(v)$
7: while $|\mu| > \text{tol}$ do
8: $T = (W_k + \lambda I)^2 - \delta^{-2} h_k h_k^T$
9: $[v, \mu] = \text{eigs}(T, 1, 'SA')$
10: $\lambda = p(v)$
11: end while
12: compute $B_{k+1} = B(x_{k+1})$, $d_{k+1} = d(x_{k+1})$
13: transform back $x^{k+1} \leftarrow (\lambda, v)$
14: compute $B_{k+1} = B(x^{k+1})$, $d_{k+1} = d(x^{k+1})$
15: check termination property
16: end for
Example: shaw(2000)

shaw(2000): Safeguarded iteration
Example: shaw(2000)
Concluding remarks

Outline

1. Two examples

2. Variational characterization of eigenvalue problems
   - Overdamped Problems
   - Nonoverdamped problems

3. Applications
   - Sylvester’s law of inertia
   - Nonlinear low rank modification of a symmetric eigenvalue problem
   - Safeguarded iteration
   - Detecting hyperbolic matrix polynomials
   - Free vibrations of fluid-solid structures
   - Regularization of total least squares problems

4. Concluding remarks
Applications of variational principles

Infinite dimensional problems

- locating eigenvalues of fluid–solid type structures (V. 2003, 2005)
- quadratic problems of restricted rank (V. 2004)
Applications of variational principles

Infinite dimensional problems
- locating eigenvalues of fluid–solid type structures (V. 2003, 2005)
- quadratic problems of restricted rank (V. 2004)

finite dimensional problems
- safeguarded iteration (V. & Werner 1982)
- a priori bounds for condensation methods (V. 1984)
- dynamic element methods (V. 1987)
- nonlinear Arnoldi method (V. 2003)
- Jacobi–Davidson type method for nonlinear evps (T. Betcke & V. 2004)
- a priori bounds for AMLS (Elssel & V. 2006, Stammberger & V. 2012)
- regularized total least squares problems (Lampe & V. 2007, 2008)
- detecting hyperbolic and definite matrix polynomials (Niendorf & V. 2010)
- Low rank modifications of symmetric evps (V., Yildiztekin & Huang 2011)
- Computable error bounds for symmetric problems (V., Yildiztekin 2012)
- generalization of Sylvester’s law (Kostic & V. 2013)
- shape optimization (S.A. Mohammadi, V. 2015)


References


