Distance problems, spectra and pseudospectra

Francqui Lecture 23-4-2010

Paul Van Dooren
Université catholique de Louvain
CESAME, Louvain-la-Neuve, Belgium

Goal of this lecture

We analyze properties and nearness problems of :

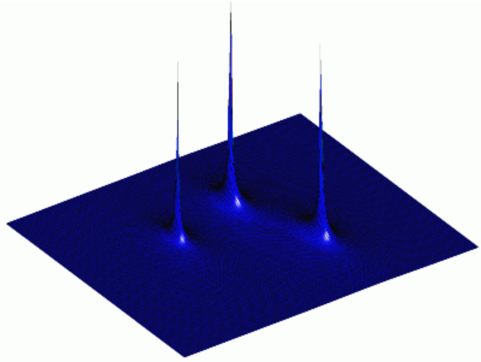
- Constant matrices A
- Pencils $\lambda I - A$
- Polynomial matrices $P(\lambda)$
- Rational matrices $R(\lambda)$

Several problems are related to basic properties of linear systems

We will present basic realization results to make the connections between these different problems

We start by looking at spectra and pseudospectra of matrices

Spectra and pseudospectra



Credit to Bötcher, Pseudospectrum, Scholarpedia

Pseudospectra

Consider the matrix

$$A = \begin{bmatrix} 1+i & 0 & i \\ -i & 0.2 & 0 \\ 0.7i & 0.2 & 0.5 \end{bmatrix}$$

and the level sets of the real function

$$\mathbb{C} \to \mathbb{R}, \lambda \mapsto \|(A - \lambda I)^{-1}\|$$

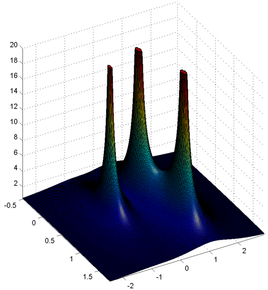
for $\varepsilon^{-1} = 1, \mathbf{2}, 3, \mathbf{4}, \mathbf{6}, 10, 20$

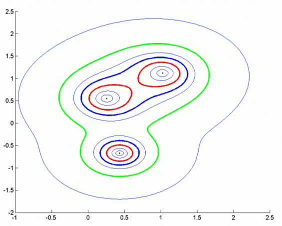
The poles of this function are the spectrum of A

$$\lambda \in \sigma(A) \;\Leftrightarrow\; \|(A - \lambda I)^{-1}\| = \infty$$

The growing level sets are the pseudospectrum of A

$$\sigma_\varepsilon(A) = \{\lambda \in \mathbb{C} : \|(A - \lambda I)^{-1}\| > 1/\varepsilon\}$$





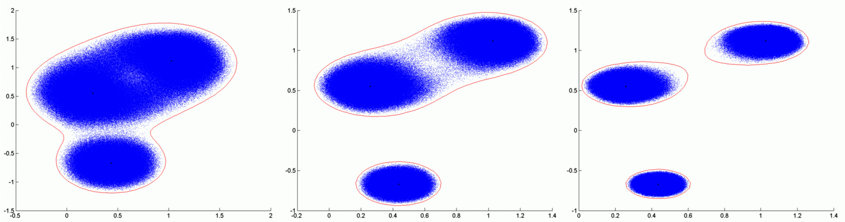
Pseudospectra

Theorem The ε-level set is the union of all eigenvalues of $A+E$ for $\|E\|<\varepsilon$

$$\sigma_\varepsilon(A) = \bigcup_{\|E\|<\varepsilon} \sigma(A+E)$$

Proof $(\lambda I - A - E)v = 0 \;\Leftrightarrow\; Ev = u, (A-\lambda I)^{-1}u = v$ implies that $\|E\| < \varepsilon \;\Rightarrow \|(A-\lambda I)^{-1}\| > 1/\varepsilon$

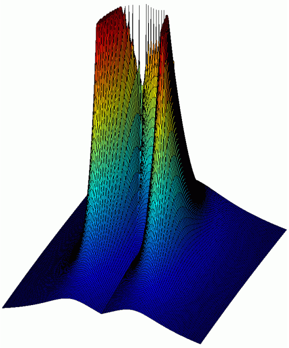
Let us take the above example for $\varepsilon^{-1} = \mathbf{2}, \mathbf{4}, \mathbf{6}$

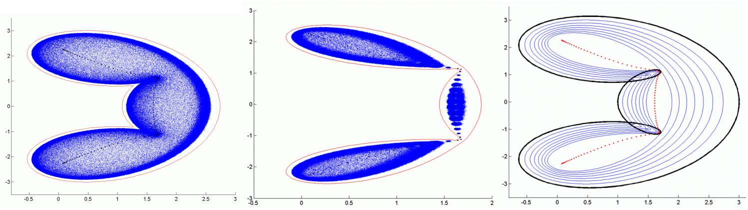


Second example (Grcar matrix)

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & & & \\ -1 & 1 & 1 & 1 & 1 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & -1 & 1 & 1 & 1 & 1 \\ & & & -1 & 1 & 1 & 1 \\ & & & & -1 & 1 & 1 \\ & & & & & -1 & 1 \end{bmatrix}$$

$n = 100, \varepsilon^{-1} = 10^{-i}, i = 2, 3, \ldots, 8$





Growth and decay

Example the linear diff. equations

$\dot{x}(t) = Ax(t) \quad x_{k+1} = Ax_k$

have an asymptotic decay rate

$\|e^{tA}\| \leq c_c e^{t\alpha(A)} \quad \|A^k\| \leq c_d \rho(A)^k$

(spectral abscissa) (spectral radius)

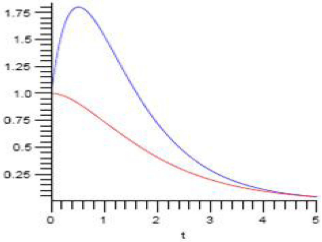
$\alpha(A) = \max_i \Re\lambda_i(A) \quad \rho(A) = \max_i |\lambda_i(A)|$

Exc Which is which ?

$A_1 = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}, \; A_2 = \begin{bmatrix} -1 & 5 \\ 0 & -2 \end{bmatrix}$

If A is a normal matrix then $c_c = c_d = 1$ but in general we can only guarantee the bound $c_c = c_d = \kappa(V)$ where $AV = V\Lambda$ is the eigen-decomposition

What if $\kappa(V)$ is ∞ (Jordan) ?



Non-normality

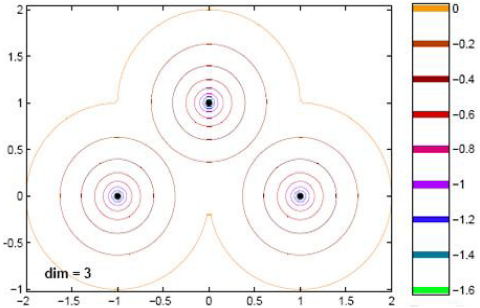
Let A be normal then the pseudo-spectrum for a unitarily invariant norm consists of the union of concentric circles of radius ε around the spectrum of A

Proof $A = U\Lambda U^*$ implies $\|(A - \lambda I)^{-1}\| = \|(\Lambda - \lambda I)^{-1}\|$ and $\|(\Lambda - \lambda I)^{-1}\| = \max_i |\lambda - \lambda_i|^{-1}$

Example

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$\sigma(A) = \{1, -i, i\}$$



About nearness problems

We measure distance only in unitarily invariant norms

The Frobenius (or Euclidean norm) has additional advantages

One wants to find the distance to certain sets of matrices :

- Hankel matrices is easy
- Toeplitz matrices is easy
- Hermitian matrices is easy
- Semidefinite matrices is easy
- Orthogonal matrices is easy
- Low rank matrices is easy
- Distance matrices is relatively easy
- Normal matrices is not so easy
- Stable matrices is difficult
- Unstable matrices is easy

Intersections of these constraints are usually difficult as well

Frobenius norm

We look at unitarily invariant norms only

The Frobenius (or Euclidean norm) has additional advantages

- related to the inner product $\langle X, Y \rangle_{\mathbb{C}} := \langle X_r, Y_r \rangle_{\mathbb{R}} + \langle X_i, Y_i \rangle_{\mathbb{R}}$
- where $\langle X_r, Y_r \rangle = trace Y_r^T X_r$
- and $\langle X, Y \rangle_{\mathbb{C}} = \Re\, trace Y^* X = trace(Y_r^T X_r + Y_i^T X_i)$
- $\|M\|_F^2 = \langle M, M \rangle_{\mathbb{C}} = trace M^* M$
- it is strictly convex
  ($\|aM + bN\| < a\|M\| + b\|N\|$ unless $M = cN, a, b, c > 0$)

Consequence :
The nearest matrix in this norm to a linear subspace is just the orthogonal projection in this inner product

Nearest Hankel and Toeplitz matrix

Hankel and Toeplitz matrices form a linear subspace since

$H = \sum_{i=1}^{2n-1} c_i H_i \quad$ (H_i is a Hankel basis matrix and $c_i \in \mathbb{R}$ or $\mathbb{C}$)
$T = \sum_{i=1}^{2n-1} c_i T_i \quad$ (T_i is a Toeplitz basis matrix and $c_i \in \mathbb{R}$ or $\mathbb{C}$)
$T_h = a_0 I_n + \sum_{i=1}^{n-1}(a_i R_i + b_i C_i) \quad$ (R_i and C_i are real and imaginary Hermitian basis Toeplitz matrices, and $a_i, b_i \in \mathbb{R}$)

Moreover, in all of these cases the basis matrices are orthogonal !
$\langle H_i, H_j \rangle = \langle T_i, T_j \rangle = \langle R_i, R_j \rangle = \langle C_i, C_j \rangle = \langle R_i, C_j \rangle = 0$

Consequence :
The distance problem in the Frobenius norm is solved by the inner products with the basis matrices : just take the averages along anti-diagonals (Hankel) or diagonals (Toeplitz)

The solution is unique because of the strict convexity of this norm

Nearest Hankel and Toeplitz matrix

In the 2-norm this is not the case

Let $A_h = \begin{bmatrix} 0 & -1 & -1 \\ 1 & 1 & -1 \\ 0 & 1 & 0 \end{bmatrix}$ and $A_t = \begin{bmatrix} -1 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

In the Frobenius norm, the (unique) optimal Hankel approximation to A_h and optimal Toeplitz approximation to A_t are the zero matrix

Both errors A_h and A_t have 2-norm 1.88 since the eigenvalues of A_t are $\{-1.53, -0.35, 1.88\}$

But $H = \begin{bmatrix} 0 & 0 & \delta \\ 0 & \delta & 0 \\ \delta & 0 & 0 \end{bmatrix}$ and $T = \begin{bmatrix} \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ with $\delta = 0.17$

give error norms equal to 1.7 (and hence smaller). It is not clear how to solve this problem in the 2-norm because it is not strictly convex

Nearest substochastic matrix

What is the minimum distance of A to the (row) substochastic matrices $\mathcal{S}$?

This is a closed convex set since $S \geq 0$ and $S\mathbf{1} \leq \mathbf{1}$ is the closed convex set of matrices with each row in the standard $n-$dimensional simplex

If $A \notin \mathcal{S}$ we need to find for each row in A the closest vector in this simplex.

Since this is a closed convex set, it is the projection on this convex set and will be on the boundary of $\mathcal{S}$. The nearest point is unique in the Frobenius norm.

Exc : If $A > (A\mathbf{1} - \mathbf{1})\mathbf{1}^T/n$ and $A\mathbf{1} > \mathbf{1}$, then the nearest matrix in the 2-norm and the Frobenius norm is given by $S = A - (A\mathbf{1} - \mathbf{1})\mathbf{1}^T/n$

Nearest Hermitian matrix

What is the minimum distance to the set of Hermitian matrices $\mathcal{H}$?

Write $A = A_h + A_a$, where $A_h = (A + A^*)/2$ and $A_a = (A - A^*)/2$ are the Hermitian and anti-Hermitian parts of A

Then $\langle A_h, A_a \rangle = 0$, hence *arg* $\min_{H\in\mathcal{H}} \|A - H\|_F = A_h$ for Frobenius

This is the orthogonal projection on the linear subspace of Hermitian matrices

This result holds also for all unitarily invariant norms since

$$\|A - A_h\| = \|A_a\| = (\|(A - H) + (H^* - A^*)\|)/2 \leq (\|A - H\| + \|H^* - A^*\|)/2 = \|A - H\|$$

for $H = H^*$ and since $\|M^*\| = \|M\|$ (for unitarily invariant norms)

Nearest semi-definite Hermitian matrix

What is the minimum distance to the set of semi-definite matrices $\mathcal{H}_0$?

This is a closed convex ($H_0 \succeq 0$) set unlike the positive definite ones ($H_0 \succ 0$) : their boundary contains the semi-definite ones

Let $A_h = HQ, H = H^* \succeq 0, QQ^* = I_n$ be the polar decomposition of A_h

Then $(A_h + H)/2 = arg\min_{H_0 \in \mathcal{H}_0} \|A - H_0\|_F$ is the unique minimizer in the Frobenius norm

The error satisfies $\delta_F(A)^2 = \sum_{\lambda_i(A_h)<0} \lambda_i(A_h)^2 + \|A_a\|_F^2$

In the 2-norm the solution is known (Halmos) but more complicated to describe

Nearest rank k matrix

This is given by the singular value decomposition (Eckart-Young)

Let $A \in \mathbb{C}^{m \times n}$ (with $m \geq n$) then $A = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^*$ with $\Sigma = diag\{\sigma_i\}$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$

The set of rank k matrices is a smooth manifold (one can define a tangent plane and an orthogonal projection). The nearest point is unique in the Frobenius norm and is given by

$A_k = \sum_{i=1}^{k} u_i \sigma_i v_i^*$ and $\delta_F(A)^2 = \sum_{i=k+1}^{n} \sigma_i^2$

This is also optimal in the 2-norm ($\delta_2(A) = \sigma_{k+1}$) but it is not unique

Exc : Describe all optimal rank-1 approximations to diag{2,1,0}

Multiplicative versus additive perturbations

There are two versions of this perturbation result

Let $A \in \mathbb{C}^{m \times n}$ with $m \geq n$ and $\Delta \in \mathbb{C}^{m \times n}$

The set of closest rank $n-1$ matrices satisfies

$\min_{\Delta} \|\Delta\|_2$ such that $\sigma_n(A - \Delta) = 0$

It equals $\sigma_n(A)$ and $\Delta = u_n \sigma_n v_n^*$

There is also a multiplicative version where $\Delta \in \mathbb{C}^{n \times m}$

$\min_{\Delta} \|\Delta\|_2$ such that $\sigma_n(I_n - \Delta A) = 0$

It equals $\sigma_1^{-1}(A)$ and $\Delta = v_1 \sigma_1^{-1} u_1^*$

Link when A is invertible : $det(A - \Delta) = det(A) det(I - \Delta A^{-1})$

Nearest unitary matrix

What is the minimum distance of A to the set of unitary matrices $\mathcal{U}$?

This requires the minimization of $\|A - Q\|_F, \; Q \in \mathcal{U}$, which is related to

$$\min_{Q \in \mathcal{U}} \|A - BQ\|_F^2 = \min_{Q \in \mathcal{U}} \langle A - BQ, A - BQ \rangle_{\mathbb{C}}$$

This Procrustes problem is solved by the polar factor of $B^* A = HQ$ since $\langle A - BQ, A - BQ \rangle_{\mathbb{C}} = \Re \, trace(A^* A + B^* B - 2Q^* B^* A)$

For $A = U \Sigma V^*$, the polar factor is given by $Q = UV^*$ and we have

$$\|A - UV^*\|_F^2 = \|U(\Sigma - I_n)V^*\|_F^2 = \sum_i (\sigma_i - 1)^2$$
$$\|A - UV^*\|_2^2 = \|U(\Sigma - I_n)V^*\|_2^2 = \max_i (\sigma_i - 1)^2$$

Exc: Are these solutions unique for each norm ?

Structured perturbations and rank defficiency

Let us consider a structured perturbation of the form

$$\Delta = diag\{\Delta_1, \cdots, \Delta_r\}, \Delta_i \in \mathbb{C}^{m_i \times m_i}$$

and try to find the nearest singular matrix $M + \Delta$ to a given matrix M

$$\mu_{\mathbb{C}}(M) = \min_{\Delta} \|\Delta\| : \det(M - \Delta) = 0$$

For $r = 1$ (the one block case) this is given by $\Delta = u_{min}\sigma_{min}v^*_{min}$ for both the Frobenius norm and the 2-norm and $\|\Delta\|_2 = \|\Delta\|_F = \sigma_n$

For $r = 2$ and $r = 3$ this can still be solved and is related to the computation of so-called "structured singular values"

For $r \geq 4$ this becomes an NP hard problem

Equal modulus property

Clearly $\|\Delta\|_F \geq \|\Delta\|_2 \geq \sigma_{min}(M)$ because of Eckart-Young

But also $\|\Delta\|_F \geq \|\Delta\|_2 \geq \sigma_{min}(M_D)$ with $M_D := D^{-1}MD$ and $D^{-1}\Delta D = \Delta$ (why ?) and this implies $D = diag\{d_1 I_{m_1}, \cdots, d_r I_{m_r}\}$

Let us try to find a scaling that maximizes $\sigma_{min}(M_D)$ for all D

If that maximum is smooth, it follows that $M_D v_{min} = \sigma_{min} u_{min}$ with

$\|u_i\|_2^2 = \|v_i\|_2^2$ for all m_i-subvectors of u_{min} and v_{min} (EM property)

This follows from differentiating $\sigma_{min}(M_D)$ versus d_i (equal to 0)

Then $\Delta := \sigma_{min} diag\{u_1 v_1^* / \|v_1\|_2^2, \cdots, u_r v_r^* / \|v_r\|_2^2\}$
satisfies $\Delta v_{min} = \sigma_{min} u_{min}$ and has norm σ_{min}

Therefore $M_D - \Delta$ is singular and so is $M - \Delta$ (why ?)

Exc : What is the corresponding multiplicative version ?

Real perturbations of complex matrices

Now consider a real perturbation $\Delta \in \mathbb{R}^{n\times m}$ of a complex matrix $M \in \mathbb{C}^{m\times n}$ and try to find the smallest (multiplicative) perturbation Δ such that $I - \Delta M$ is singular

$$\mu_{\mathbb{R}} = \min_{\Delta} \|\Delta\| : \det(I - \Delta M) = 0$$

The solution is given by

$$\mu_{\mathbb{R}}(M) = \inf_{\gamma\in(0,1]} \sigma_2 \begin{bmatrix} \Re M & -\gamma \Im M \\ \gamma^{-1}\Im M & \Re M \end{bmatrix}$$

The corresponding additive version maximizes $\sigma_{n-2}(\gamma)$ over γ.

Notice that $M - \Delta$ is then singular iff $\begin{bmatrix} \Re M - \Delta & -\gamma \Im M \\ \gamma^{-1}\Im M & \Re M - \Delta \end{bmatrix}$ has nullity 2 (why ?)

Off diagonal perturbations of Hermitian matrices

Now consider a perturbation Δ_h of a Hermitian matrix H of the form

$$H := \begin{bmatrix} S & R \\ R^* & T \end{bmatrix}, \qquad \Delta_h := \begin{bmatrix} 0 & \Delta \\ \Delta^* & 0 \end{bmatrix}$$

and look for the nearest singular perturbation $H - \Delta_h$ of H

$$\mu_{\mathbb{R}} = \min_{\Delta} \|\Delta\|_2 : \det(H - \Delta_h) = 0$$

Clearly, $\Gamma = diag\{\gamma, \gamma^{-1}\}$ yields the invariance $\Gamma\Delta_h\Gamma = \Delta_h$, so that

$$\|\Delta_h\|_F \geq \|\Delta_h\|_2 \geq \max_{\gamma} \sigma_{min}(\Gamma H \Gamma)$$

for all γ but we will get a better bound

Exc : Prove that there exists a Δ_H such that $\det(H - \Delta_h) = 0$ if and only if $diag\{S, -T\}$ is not (positive or negative) definite

Inertia preservation

Consider

$$H_\gamma := \begin{bmatrix} \gamma^2 S & R \\ R^* & T/\gamma^2 \end{bmatrix} = \Gamma \begin{bmatrix} S & R \\ R^* & T \end{bmatrix} \Gamma, \quad \Gamma := \begin{bmatrix} \gamma I_n & 0 \\ 0 & I_n/\gamma \end{bmatrix}$$

with γ positive real.

Then $\Gamma \Delta_h \Gamma = \Delta_h$ and

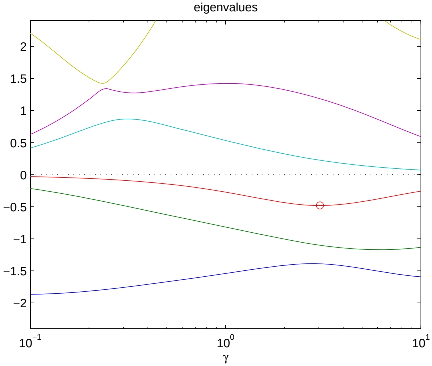
$$\det(H - \Delta_h) = \det \Gamma (H - \Delta_h) \Gamma = \det(H_\gamma - \Delta_h)$$

and

$$In\{H_\gamma - \Delta_h\} = In\{H - \Delta_h\} \quad \forall \gamma \in (0, \infty)$$

Hence bounds for $\det(H_\gamma - \Delta_h) = 0$ must hold for all values of γ

$$\inf_{\substack{\text{struct indef } \Delta_H \\ \det(H+\Delta H)=0}} \|\Delta_h\|_2 = \min\{\sup_\gamma \underline{\mu}(H_\gamma), \sup_\gamma \underline{\nu}(H_\gamma)\}$$

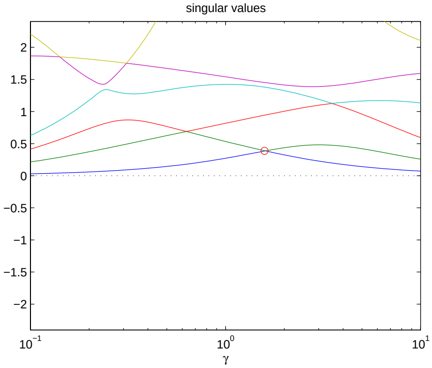


where

$$\underline{\mu}(H) := \min_i\{\lambda_i(H) : \lambda_i(H) > 0\}, \quad \underline{\nu}(H) := \min_i\{-\lambda_i(H) : \lambda_i(H) < 0\}$$

Notice that

$$\inf_{\substack{\text{struct indef } \Delta_H \\ \det(H+\Delta H)=0}} \|\Delta_h\|_2 = \min\{\sup_\gamma \underline{\mu}(H_\gamma), \sup_\gamma \underline{\nu}(H_\gamma)\} \geq \sup_\gamma \sigma_{min}(H_\gamma)$$



Different from singular value bound

Construction

At the optimum value, one shows again that EM holds

$$H_\gamma \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix}$$

and can then construct

$$\Delta_h := \begin{bmatrix} 0 & \Delta \\ \Delta^* & 0 \end{bmatrix}, \qquad \Delta = \lambda u v^* / \|v\|_2^2$$

Clearly $(H - \Delta_h) \begin{bmatrix} u \\ v \end{bmatrix} = 0$, and since $\|u\|_2^2 = \|v\|_2^2$ it follows that

$\|\Delta_h\|_{2,F} = \|\Delta\|_{2,F} = |\lambda|$, which is optimal in the 2-norm and the Frobenius norm

There is also a multiplicative version, related to the additive case since H is square and normally invertible

Other nearness problems

Nearest normal matrix ($NN^* = N^*N$) : no formula but a property that can be turned into an algorithm (Ruhe)

Nearest correlation matrix (positive semi-definite with unit diagonal) : no formula but a property that can be turned into an algorithm (Higham)

Nearest unstable matrix (based on pseudo-spectra, see later)

Nearest unstable polynomial

Nearest stable matrix or polynomial (distance to a non-convex set, hard)

Nearest singular matrix for elementwise norm (hard, Rohn)
Can be reduced to a structured singular value problem

Linear systems

Consider the explicit state space equations

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases}$$

and

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) + Du(k), \end{cases}$$

The input to output responses are (linear) convolutions :

$$y(t) = \int_{-\infty}^{\infty} Ce^{A(t-\tau)} Bu(\tau) d\tau.$$

and

$$y(k) = \sum_{-\infty}^{\infty} CA^{(k-j)} Bu(j).$$

State space

Taking Laplace and z-transforms yields state-space equations

$$\begin{cases} \lambda x(\cdot) = Ax(\cdot) + Bu(\cdot) \\ y(\cdot) = Cx(\cdot) + Du(\cdot), \end{cases}$$

where λ stands for the Laplace variable s in continuous-time and for the shift operator z in discrete-time.

Eliminating the state $x(\cdot)$ from this yields the transfer function

$$H(\lambda) := C(\lambda I_n - A)^{-1}B + D, \qquad y(\lambda) = H(\lambda)u(\lambda)$$

relating the input and output in the transformed domain.
This is a rational matrix function in the variable λ.

Conversely, does every rational matrix function correspond to a state-space sytem ?

Minimal realizations

Every $m \times p$ proper (i.e. bounded at $\lambda = \infty$) rational matrix function $H(\lambda)$ of McMillan degree n can be realized as $\{A, B, C, D\}$ such that $H(\lambda)$ is the Schur complement of the so-called system matrix $S(\lambda)$

$$H(\lambda) = S_c \begin{bmatrix} A - \lambda I_n & B \\ C & D \end{bmatrix} = D + C(\lambda I_n - A)^{-1} B$$

This realization is minimal (in the sense that no smaller realization can be found) if and only if the following properties hold

$rank \begin{bmatrix} A - \lambda I_n & B \end{bmatrix} = n, \ \forall \lambda \in \mathbb{C}$ (controllability)

$rank \begin{bmatrix} A - \lambda I_n \\ C \end{bmatrix} = n, \ \forall \lambda \in \mathbb{C}$ (observability)

All minimal realizations of a same transfer function are related by an invertible state-space transformation T

$$S_c \begin{bmatrix} A - \lambda I_n & B \\ C & D \end{bmatrix} = S_c \begin{bmatrix} T^{-1}(A - \lambda I_n) T & T^{-1} B \\ CT & D \end{bmatrix}$$

Minimal realizations

Every $m \times p$ arbitrary rational matrix function $H(\lambda)$ of McMillan degree n can be realized as $\{A, B, C, D, E, F\}$ with $(\lambda E - A)$ $n \times n$ and non-singular, such that $H(\lambda)$ is the Schur complement of the so-called system matrix $S(\lambda)$

$$H(\lambda) = S_c \left[\begin{array}{cc} A - \lambda E & B \\ C - \lambda F & D \end{array} \right] = D + (C - \lambda F)(\lambda E - A)^{-1} B$$

This realization is minimal (in the sense that no smaller realization of this type can be found) if and only if the following properties hold

$rank \left[\begin{array}{cc} A - \lambda E & B \end{array} \right] = n, \ \forall \lambda \in \mathbb{C}, \ \lambda \neq \infty$ (finite controllability)

$rank \left[\begin{array}{cc} E & B \end{array} \right] = n$, (controllability at infinity)

$rank \left[\begin{array}{c} A - \lambda E \\ C - \lambda F \end{array} \right] = n, \ \forall \lambda \in \mathbb{C}, \ \lambda \neq \infty$ (finite observability)

$rank \left[\begin{array}{c} E \\ F \end{array} \right] = n$, (observability at infinity)

All minimal realizations of a same transfer function are also related

Example

The following (non-proper) $m \times p$ polynomial matrix

$$P(\lambda) := P_0 + \lambda P_1 + \lambda^2 P_2 + \cdots + \lambda^d P_d$$

can be realized as follows

$$P(\lambda) = S_c \left[\begin{array}{ccccc|c} -I_m & & & & & P_d \\ \lambda I_m & -I_m & & & & \vdots \\ & \ddots & \ddots & & & \vdots \\ & & \lambda I_m & -I_m & & P_1 \\ \hline & & & \lambda I_m & & P_0 \end{array}\right]$$

This realization is minimal iff $rank P_d = m$

Variant: Every rational matrices of degree n can be realized by a quintuple $\{A, B, C, D, E\}$ with non-zero $\det(A - \lambda E)$ $rankE = n$ such that

$$H(\lambda) = S_c \begin{bmatrix} A - \lambda E & B \\ C & D \end{bmatrix} = D + C(\lambda E - A)^{-1} B$$

These realizations are now used to verify algebraically three important system properties :

A system is minimal if it is controllable and observable

A minimal system is stable if the solutions of $\lambda x(\cdot) = Ax(\cdot)$ converge to zero for all initial conditions

A minimal system is passive if it is stable and its “supply" dissipates energy

We will present numerical algorithms to test this

Symmetry

A matrix function $\Phi(\lambda) : \mathbb{C} \to \mathbb{C}^{n\times n}$ is parahermitian with respect to Γ iff it is its own paraconjugate transpose, i.e. if $\Phi_*(\lambda) = \Phi(\lambda)$, where

$\Phi_*(s) \triangleq \Phi^*(-s)$ for $\Gamma = j\,\mathbb{R}$ (in continuous-time)

$\Phi_*(z) \triangleq \Phi^*(1/z)$ for $\Gamma = e^{j\,\mathbb{R}}$ (in discrete-time)

and $\Phi^*(.)$ is just the conjugate transpose $\Phi(.)$

Such transfer functions represent typically spectral density functions

We expect that such functions also have symmetric realizations

$\Phi(s) = \Phi_2^*(-s)^2 + \Phi_1^*(-s) + \Phi_0 + \Phi_1 s + \Phi_2 s^2$
and
$\Phi(z) = \Phi_2^* z^{-2} + \Phi_1^* z^{-1} + \Phi_0 + \Phi_1 z + \Phi_2 z^2$

are for instance both parahermitian if $\Phi_0 = \Phi_0^*$.

Symmetric realizations

A parahermitian transfer function can always be realized by a parahermitian system matrix $S(\lambda)$ of the form

$$S(s) = \begin{bmatrix} 0 & A^* + s\,E^* & C^* + s\,F^* \\ A - s\,E & H_{11} & H_{12} \\ C - s\,F & H_{21} & H_{22} \end{bmatrix} \text{ (continuous-time)}$$

$$S(z) = \begin{bmatrix} 0 & z\,A^* - E^* & z\,C^* - F^* \\ A - z\,E & H_{11} & H_{12} \\ C - z\,F & H_{21} & H_{22} \end{bmatrix} \text{ (discrete-time)}$$

where the matrix $H := \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$ is Hermitian

Defining $T(\lambda) = (C - \lambda\,F)(\lambda\,E - A)^{-1}$ then the transfer function is

$$\Phi(\lambda) = S_c S(\lambda) = \begin{bmatrix} T(\lambda) & I \end{bmatrix} H \begin{bmatrix} T_*(\lambda) \\ I \end{bmatrix}$$

Examples

The parahermitian transfer functions (with $\Phi_0 = \Phi_0^*$)

$$\Phi(s) = \Phi_2^*(-s)^2 + \Phi_1^*(-s) + \Phi_0 + \Phi_1 s + \Phi_2 s^2$$

and

$$\Phi(z) = \Phi_2^* z^{-2} + \Phi_1^* z^{-1} + \Phi_0 + \Phi_1 z + \Phi_2 z^2$$

are respectively realized by the system matrices

$$\left[\begin{array}{cccc|c} 0 & 0 & -I & -sI & 0 \\ 0 & 0 & 0 & -I & -sI \\ -I & 0 & 0 & 0 & \Phi_2 \\ sI & -I & 0 & 0 & \Phi_1 \\ \hline 0 & sI & \Phi_2^* & \Phi_1^* & \Phi_0 \end{array}\right], \quad \left[\begin{array}{cccc|c} 0 & 0 & -zI & I & 0 \\ 0 & 0 & 0 & -zI & I \\ -I & 0 & 0 & 0 & \Phi_2 \\ zI & -I & 0 & 0 & \Phi_1 \\ \hline 0 & zI & \Phi_2^* & \Phi_1^* & \Phi_0 \end{array}\right]$$

which are both minimal if Φ_2 has full rank.

Transformations of realizations

One can define (generalized) state space transformations that leave the transfer function (and $\{E, A, C, F\}$) invariant :

$$\begin{bmatrix} I & 0 & 0 \\ EX & I & 0 \\ FX & 0 & I \end{bmatrix} S(s) \begin{bmatrix} I & XE^* & XF^* \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} = \begin{bmatrix} 0 & A^* + sE^* & C^* + sF^* \\ A - sE & H_{11}(X) & H_{12}(X) \\ C - sF & H_{21}(X) & H_{22}(X) \end{bmatrix}$$

$$\begin{bmatrix} I & 0 & 0 \\ EX & I & 0 \\ FX & 0 & I \end{bmatrix} S(z) \begin{bmatrix} I & XA^* & XC^* \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} = \begin{bmatrix} 0 & zA^* - E^* & zC^* - F^* \\ A - zE & H_{11}(X) & H_{12}(X) \\ C - zF & H_{21}(X) & H_{22}(X) \end{bmatrix}$$

where the matrices $H(X)$ are given by

$$H(X) \doteq \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} + \begin{bmatrix} E \\ F \end{bmatrix} X \begin{bmatrix} A^* & C^* \end{bmatrix} + \begin{bmatrix} A \\ C \end{bmatrix} X \begin{bmatrix} E^* & F^* \end{bmatrix}$$

$$H(X) \doteq \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} + \begin{bmatrix} A \\ C \end{bmatrix} X \begin{bmatrix} A^* & C^* \end{bmatrix} - \begin{bmatrix} E \\ F \end{bmatrix} X \begin{bmatrix} E^* & F^* \end{bmatrix}$$

KYP lemma

Lemma (KYP) A parahermitian function $\Phi(\lambda)$ is non-negative on Γ ($\Phi(\lambda) \succeq 0$) iff there exists a Hermitian matrix X such that $H(X) \succeq 0$

Proof (If) Let $H(X) \succeq 0$ then the transfer function equals

$$\Phi(\lambda) = S_c S(\lambda) = \left[\begin{array}{cc} T(\lambda) & I \end{array}\right] H(X) \left[\begin{array}{c} T_*(\lambda) \\ I \end{array}\right] \succeq 0$$

Non-negative parahermitian transfer functions appear in several problems

- **Boundedness** $\|G(\lambda)\|_2 \leq \gamma$ iff $\gamma^2 I - G(\lambda) G_*(\lambda) \succeq 0$
- **Passivity** $G(\lambda)$ is passive iff $G(\lambda) + G_*(\lambda) \succeq 0$
- **Positive** polynomial matrices imposes a condition $\Phi(\lambda) \succeq 0$

KYP and zeros

Theorem A parahermitian transfer function is positive on Γ iff $\Phi(\lambda_0) \succ 0$ for $\lambda_0 \in \Gamma$ and it has no zeros on Γ

The zeroes of $\Phi(\lambda_0)$ (with $H_{22} \succ 0$ and $F = 0$) are the generalized eigenvalues of the Schur complements of the system matrices :

$$\begin{bmatrix} 0 & A^* + s\,E^* \\ A - s\,E & H_{11} \end{bmatrix} - \begin{bmatrix} C^* \\ H_{12} \end{bmatrix} H_{22}^{-1} \begin{bmatrix} C & H_{21} \end{bmatrix}$$

$$\begin{bmatrix} 0 & z\,A^* - E^* \\ A - z\,E & H_{11} \end{bmatrix} - \begin{bmatrix} z\,C^* \\ H_{12} \end{bmatrix} H_{22}^{-1} \begin{bmatrix} C & H_{21} \end{bmatrix}$$

and are symmetric with respect to Γ (i.e. s_i, $-\bar{s}_i$ or z_i, $1/z_i$)

These are known as the Hamiltonian and the symplectic zero pencils

Condition $F = 0$ can be obtained via an orthogonal transformation

Some references

N. J. Higham. Matrix nearness problems and applications. In M. J. C.. Gover and S. Barnett, editors, Applications of Matrix Theory, (Institute of Mathematics and its Applications Conference Series, 1989

Y. Genin, Y. Hachez, Y. Nesterov, R. Stefan, P. Van Dooren, and S. Xu. Positivity and linear matrix inequalities. European Journal of Control, 8(3):275–298, 2002

Lloyd N. Trefethen, Mark Embree. Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators, Princeton University Press, Princeton, USA, 2005