On interpolations

Patrick Dewilde

TUM Inst. of Advanced Study

IAS TUM Institute for Advanced Study

Program

- Constrained interpolation
- The matrix (Quasi-Separable) setting
- Constrained interpolation in the matrix setting
- Computational issues
- Generalizations

Constrained Interpolation

We concentrate at first on the class of interpolation problems
with a constraint on the norm of the interpolant

- Nevanlinna-Pick
- Hermite-Fejer
- Schur
- Schur-Takagi

and will derive ‘matrix versions’ for them

Motivation

(1) 'low complexity' matrix approximation

(2) solving problems of the type "pre-conditioners for positive definite $C = LL'$ ":

$$\arg\min_{X \text{ of low complexity}} \|LX - I\|_F \quad \text{with } L, X \text{ Cholesky factors}$$

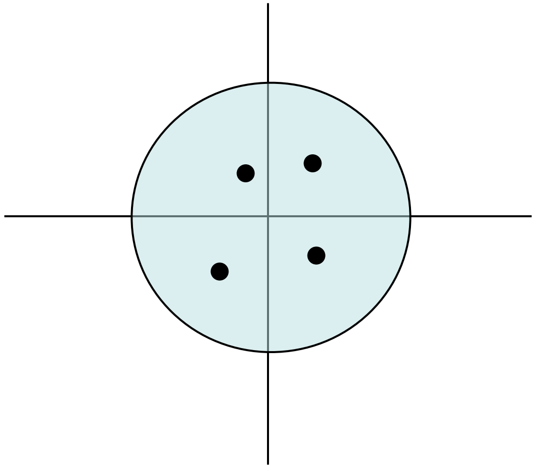
(3) 'Hankel-norm' model reduction of a time-variant (quasi-separable) system

Classical interpolation problems

$H_\infty(z)$ uniformly bounded complex functions that are analytic in the open unit disc **D -** $\{z : |z| < 1\}$ with boundary $\{z : |z| = 1\}$ and norm

$$\|S(z)\| = \sup_{|z|<1} |S(z)| = \sup_\theta |S(e^{i\theta})|$$

Nevanlinna-Pick:



Given (single) points $w_{i=1:n}$
and Values s_i

Find necessary and sufficient condition for S(z) s.t.

(1) $\forall i : S(w_i) = s_i$

(2) $\|S(z)\|_\infty \le 1$ [norm constraint!]

Schur-Takagi

Same types of interpolation as before, but now allow singularities in the unit disc in the solution

(1) let S have a minimal number of poles in the open unit disc **D**
(2) use as norm ‘sup on the boundary circle **T**’

Motivation: turns out ST solves the ‘Hankel norm model reduction problem’

Schur-Takagi Interpolation on the unit circle of the complex plane

Let $\{a_{i=1:n}\}$ be a set of (distinct) points in the open unit disc **D** of the complex plane, and s_i a set of 'interpolation values'. Find a function S(z) that is meromorphic in **D** such that

1. $S(a_i) = s_i$
2. $|S(z)| \leq 1$ for $|z| = 1$ (belongs to L_∞)
3. S is meromorphic in **D** with a minimal number of poles.

An algebraic solution

Let

$$A = \begin{bmatrix} a_1' & & \\ & \ddots & \\ & & a_n' \end{bmatrix}, \; B_1 = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \; B_2 = \begin{bmatrix} s_1' \\ \vdots \\ s_n' \end{bmatrix}$$

and solve the Lyapunov-Stein equation:

$$AMA' + B_1B_1' - B_2B_2' = M$$

(under the given assumptions there will always be a solution)

Then:
(1) If M is singular we have a 'singular case'.... skip it!
(2) If M is non-singular, let (n_1, n_2) be its signature (inertia)
then solutions exist with n_2 poles in **D** (and of degree less or equal to n).

How to construct all solutions?

Construct a J-unitary causal matrix Θ of dimension 2×2 with

$$\left[\begin{array}{c|cc} A & B_1 & B_2 \end{array}\right]$$

as 'reachability operator' (always exist), then all solutions are given by

$$S = (S_L\Theta_{12} - \Theta_{22})^{-1}(\Theta_{11} - S_L\Theta_{21})$$

in which S_L is in H_∞ and contractive, but otherwise arbitrary.

‘Historical’ note

- the problem and first results go back to independent papers of Schur and Takagi (1910-20)
- a very extensive analysis in the complex case is due to AAK (Adamyan, Arov, Krein)
- many other researchers worked on it, in particular Gohberg, Langer, Dym, Glover, Partington etc...
- Bultheel-D started the ‘system theoretic’ view on it
- the theory being algebraic can ‘easily’ be generalized to ‘just matrices’ and time-varying systems (no complex plane anymore)!

Tangential problems

Now we look for $S(z) \in H_\infty^{n\times m}$, an m×n matrix function, meromorphic and contractive in the unit disc, which also interpolates at certain points $a_{i=1:n} \in \mathbf{D}$ in certain directions ξ_i:

$$S(a_i)\xi_i = \eta_i$$

How are such constrained interpolation problems solved? We shall use a 'modern' method based on 'scattering theory' that can be generalized to matrix problems…

It is solved the same way as before, now with reachability operator

$$\left[\begin{array}{c|c|c} A & B_1 & B_2 \end{array}\right] = \left[\begin{array}{ccc|c|c} a_1' & & & \xi_1' & -\eta_1' \\ & \ddots & & \vdots & \vdots \\ & & a_n' & \xi_n' & -\eta_n' \end{array}\right]$$

Working on (time-variant) matrices?

Motivation for the construction:

$$y(z) = S(z)u(z) \text{ with } S(z) \in H_\infty$$

is equivalent to

$$y_{-\infty:+\infty} = \mathrm{Toe}[\cdots, s_{-1}, s_0, s_1, \cdots]u_{-\infty:+\infty}$$

with

$$S = \mathrm{Toe}[\cdots, s_{-1}, s_0, s_1, \cdots] = \begin{bmatrix} \ddots & \ddots & & & & \\ \ddots & s_1 & s_0 & & & \\ \ddots & s_2 & s_1 & \boxed{s_0} & & \\ \ddots & s_3 & s_2 & s_1 & s_0 & \\ \iddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

In which $S(z)$ translates to a lower, contractive, Toeplitz operator S

More general framework with block matrices works as well

input vectors $u_{-\infty,+\infty}$ with dimensions $m_{-\infty:+\infty}$ in ℓ_2^m
output vectors $y_{-\infty:+\infty}$ with dimensions $n_{-\infty:+\infty}$ in ℓ_2^n

$$S : \ell_2^m \to \ell_2^n : y = Su$$

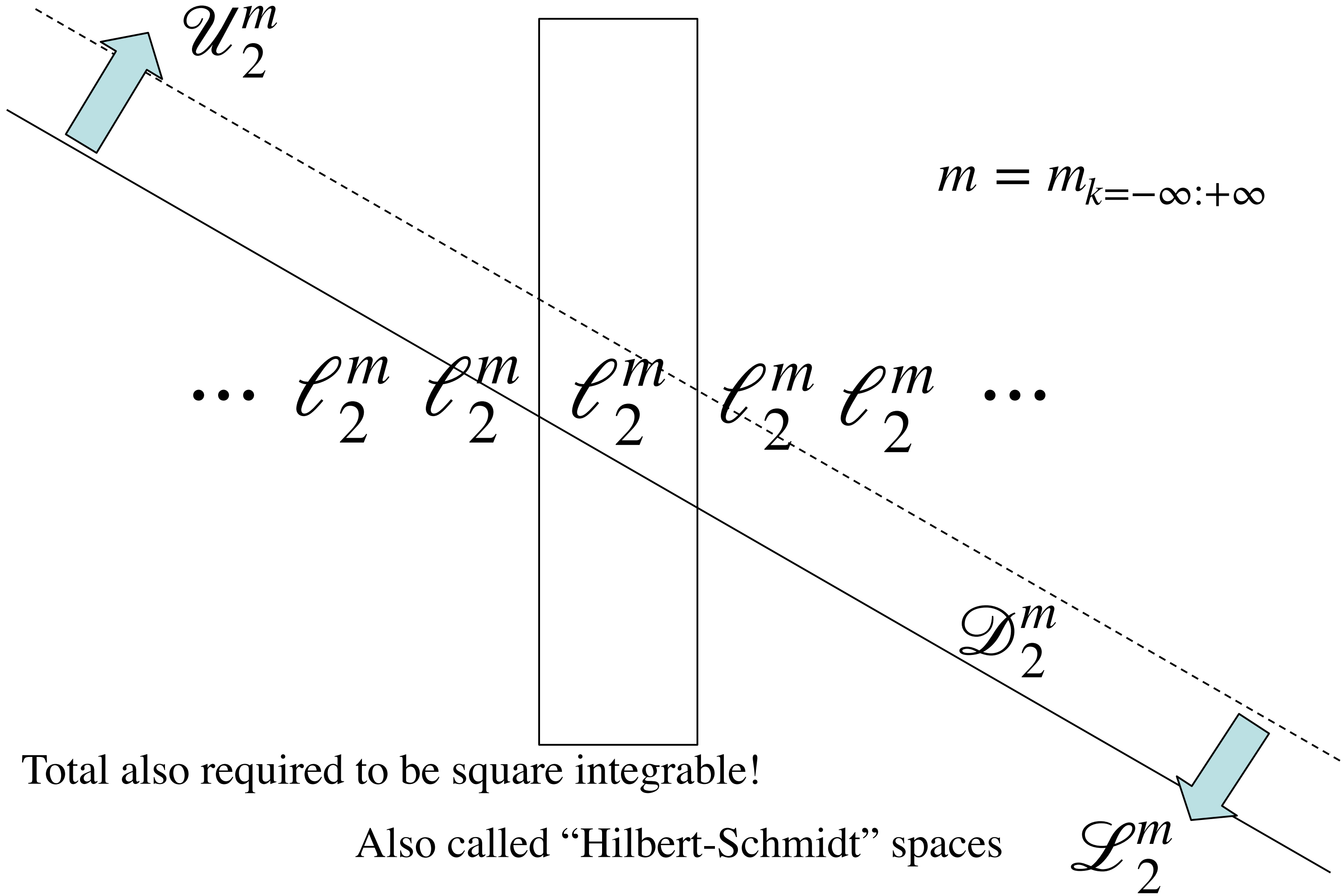
dimensions may disappear: finite matrices!

where, for the causal case:

$$S = \begin{bmatrix} \ddots & & & & \\ \ddots & S_{-1,-1} & & & \\ \ddots & S_{0,-1} & \boxed{S_{0,0}} & & \\ \ddots & S_{1,-1} & S_{1,0} & S_{1.1} & \\ \iddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

upper-lower-diagonal Frobenius spaces

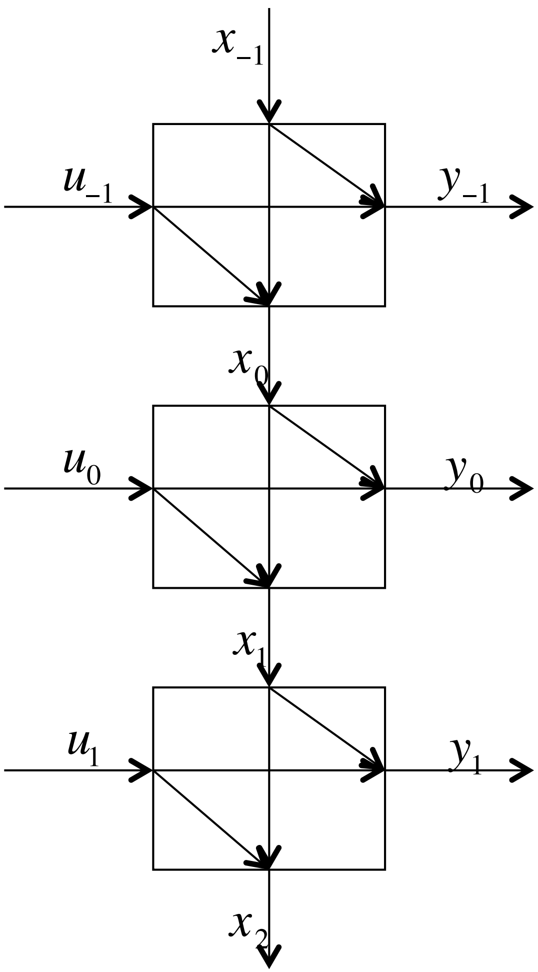
Global input space:



Total also required to be square integrable!

Also called “Hilbert-Schmidt” spaces

Model of computation (linear-lower)



Quasi-separable realization:

$$\begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} = \begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix}$$

State transformation $x_k = R_k \hat{x}_k$ with R_k square invertible:

$$\begin{bmatrix} \hat{A}_k & \hat{B}_k \\ \hat{C}_k & D_k \end{bmatrix} = \begin{bmatrix} R_{k+1}^{-1} A_k R_k & R_{k+1}^{-1} B_k \\ C_k R_k & D_k \end{bmatrix}$$

Global representation:

$$A = \mathrm{diag} A_{-\infty:+\infty}, \; A = \mathrm{diag} A_{-\infty:+\infty}, \text{ etc.}$$

$$\text{causal shift: } [Zx]_{k+1} = x_k$$

Transfer operator:

$$S = D + C(I - ZA)^{-1} ZB$$

Translation rules

Complex plane

scalars w_k, s_k

shift z

shift commutes with scalars

Matrices

block diagonals W_k, S_k

shift Z

shift does'nt commute
but keeps diagonal form:

$$D^{\langle +1 \rangle} = ZDZ'$$

Important spaces related to an upper QS matrix $T = D + C(I - ZA)^{-1}ZB$

Global reachability space:

$$B'Z'(I - A'Z')^{-1}\mathcal{D}_2$$

i.e., all strict past input sequences, at all indices, corresponding to a state ($\mathscr{D}_2$ square integrable (block) diagonals)

Observability space:

$$C(I - ZA)^{-1}\mathcal{D}_2$$

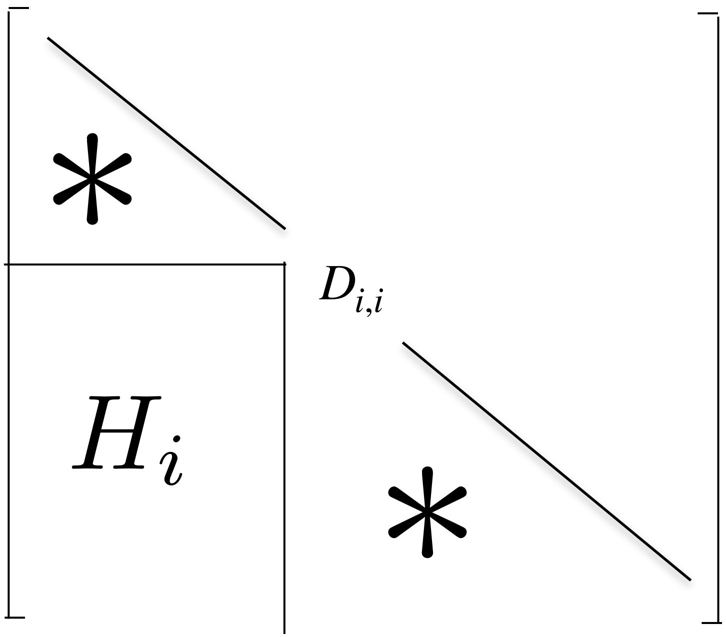
all output sequences, at all index points, that can be generated from strict past inputs

$B'Z'(I - A'Z')^{-1}\mathcal{D}_2$ forms a 'sliced basis' if the realization is strictly 'reachable' (a basis at each index point)

$C(I - ZA)^{-1}\mathcal{D}_2$ forms a 'sliced basis' if the realization is strictly 'observable'

(Nerode equivalence)

Matrix interpretation



a minimal factorization of each H_i corresponds to finding a realization

column space = reachability sp.
row space = observability space

low rank i^{th} 'Hankel matrix'

$$H_i = \begin{bmatrix} C_i \\ C_{i+1}A_i \\ C_{i+2}A_{i+1}A_i \\ \vdots \end{bmatrix} \begin{bmatrix} B_{i-1} & A_{i-1}B_{i-2} & A_{i-1}A_{i-2}B_{i-3} & \cdots \end{bmatrix}$$

Point evaluations

Complex plane

$$S(w_i) = s_i$$

equivalent to

$S(z) = s_i + (z - w_i)S_i(z),\ S_i(z) \in H_\infty$

or

$(z - w_i)^{-1}(S(z) - s_i) \in H_\infty$

Matrices

$S^\wedge(W_i) = S_i$ for W_i and S_i diagonals

equivalent to (defined by)

$S = S_i + S_1(Z - W_i)$, S_1 lower and bounded

or, equivalently

$(S - S_i)(Z - W_i)^{-1} \in \mathscr{L}$

Generalized interpolation for matrices

(because of the general formalism, subsumes Nevanlinna-Pick, Hermite-Fejer and Schur. For Schur-Takagi, see later)

Data (block diagonals): $W,\ \xi,\ \eta$ all bounded, with V u.e.s.

Find S such that .

(1) S is lower, contractive

(2) $(S\xi - \eta)(Z - W)^{-1} \in \mathcal{L}$

matching dimensions needed!

Examples

Tangential Nevanlinna-Pick:

$$W = \begin{bmatrix} W_1 & & & \\ & W_2 & & \\ & & \ddots & \\ & & & W_n \end{bmatrix}, \quad \begin{matrix} \xi & = & \begin{bmatrix} \xi_1 & \xi_2 & \cdots & \xi_n \end{bmatrix} \\ \eta & = & \begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_n \end{bmatrix} \end{matrix}$$

Schur:

$$W := \begin{bmatrix} 0 & I & & 0 \\ & \ddots & \ddots & \\ & & 0 & I \\ 0 & & & 0 \end{bmatrix}, \quad \begin{matrix} \xi & = & \begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix} \\ \eta & = & \begin{bmatrix} S_{[0]} & S_{[1]} & \cdots & S_{[n]} \end{bmatrix}. \end{matrix}$$

Hermite-Fejer (mix of the preceding!)

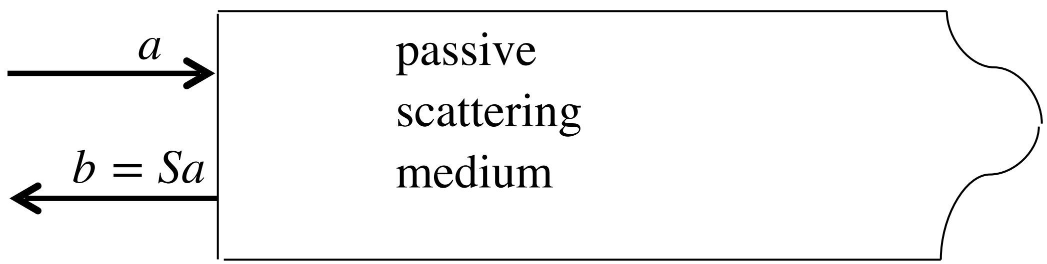
$$W_k := \begin{bmatrix} V_k & I & & 0 \\ & \ddots & \ddots & \\ & & V_k^{<m_k-1>} & I \\ 0 & & & V_k^{<m_k>} \end{bmatrix}, \quad \begin{matrix} \xi_k := \begin{bmatrix} [\xi_k]_{[0]} & 0 & \cdots & 0 \end{bmatrix} \\ \eta_k := \begin{bmatrix} [\eta_k]_{[0]} & [\eta_k]_{[1]} & \cdots & [\eta_k]_{[m_k]} \end{bmatrix}. \end{matrix} \tag{1}$$

These may be stacked again for different k's and we obtain the full Hermite-Fejer interpolation problem with

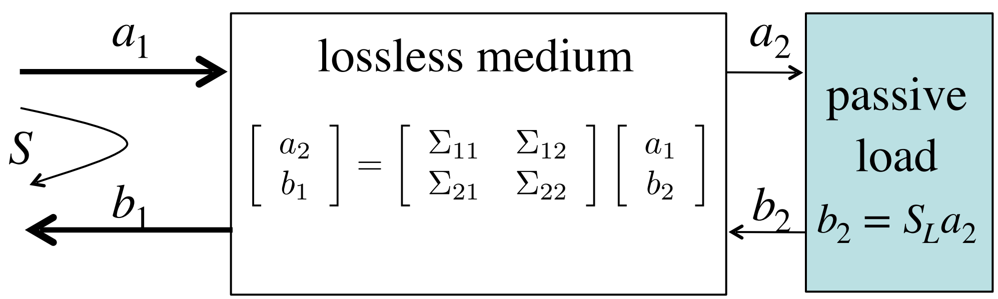
$$W := \begin{bmatrix} W_1 & & & \\ & W_2 & & \\ & & \ddots & \\ & & & W_n \end{bmatrix}, \quad \begin{matrix} \xi := \begin{bmatrix} \xi_1 & \xi_2 & \cdots & \xi_n \end{bmatrix} \\ \eta := \begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_n \end{bmatrix}. \end{matrix} \tag{2}$$

Solution methodology: scattering theory

S contractive lower



lossless transfer:

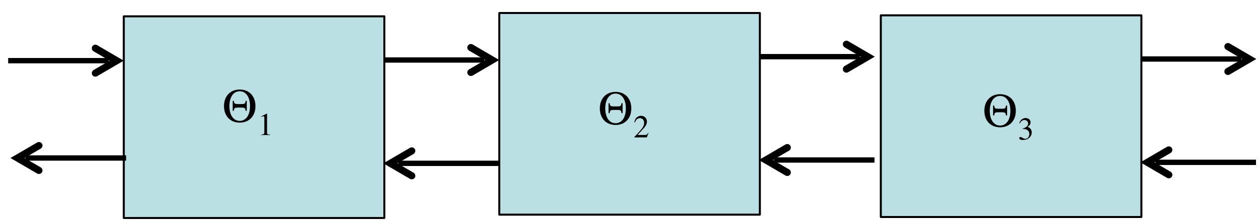


Σ unitary: $\Sigma\Sigma' = \Sigma'\Sigma = I$

$$S = \Sigma_{21} + \Sigma_{22} S_L (I - \Sigma_{12} S_L)^{-1} \Sigma_{11}$$

Chain scattering matrices

$$\begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \end{bmatrix}$$



$$\Theta = \Theta_3 \Theta_2 \Theta_1$$

Central properties of chain scattering matrices

- Θ is J-unitary for $J = \begin{bmatrix} I & \\ & -I \end{bmatrix}$ (conservation of energy)
- (classical case:) transmission zeros = (modes)* i.e. the frequencies where the input scattering is independent of the load, or $\Sigma_{21}(w_i) = S(w_i)$ (interpolation) provide the poles $1/w_i^*$ of Θ. For the matrix case this translates to:

$$\Theta J B'_{\Theta} Z'(I - A'_{\Theta} Z')^{-1} \mathcal{D}_2 = C_{\Theta}(I - Z A_{\Theta})^{-1} \mathcal{D}_2$$

reachability space maps to *observability space* by multiplication with ΘJ

- loading formula:

$$\begin{bmatrix} S & -I \end{bmatrix} = (\Theta_{22} - S_L \Theta_{12})^{-1} \begin{bmatrix} S_L & -I \end{bmatrix} \Theta$$

Central interpolation theorem

Let W, ξ, η be the interpolation data and assume that the pair (V, ξ) is strictly reachable, then the interpolation problem has (strict) solutions, iff

$$\left[\begin{array}{c|cc} W' & \xi' & -\eta' \end{array} \right]$$

is the reachability pair of a lossless chain scattering matrix Θ, i.e. iff the Lyapunov-Stein equation

$$P^{\langle -1 \rangle} = W'PW + \xi'\xi - \eta'\eta$$

has a strictly positive definite (diagonal) solution P (which is called the *Pick matrix*)

Furthermore:
All solutions are given by loading Θ with an arbitrary upper contractive S_L.

Sketch of proof

Consider the following facts:

1. J-unitary, upper operators have J-unitary realizations and vice versa. Given a reachability pair $\left[\begin{array}{c|cc} A_\Theta & B_{\Theta,1} & B_{\Theta,2} \end{array}\right]$: there should exist a state transformation *R* which makes

$$(R^{\langle -1 \rangle})^{-1} \left[\begin{array}{c|cc} A_\Theta R & B_{\Theta,1} & B_{\Theta,2} \end{array}\right]$$

J-isometric with block signature +,+,- .

This will happen iff the Lyapunov-Stein equation

$$P^{\langle -1 \rangle} = A_\Theta P A'_\Theta + B_{\Theta,1} B'_{\Theta,1} - B_{\Theta,2} B'_{\Theta,2}$$

has a strictly positive definite solution *P*, which defines *R* as $P = RR'$

(modulo an irrelevant right unitary factor!)

Sketch of proof (2)

2. The requested interpolation formulates as

$$(S\xi - \eta)(Z - W)^{-1} \in \text{causal}$$

I.e.

$$\begin{bmatrix} S & I \end{bmatrix} \begin{bmatrix} \xi \\ -\eta \end{bmatrix} (Z - W)^{-1} \in \mathcal{L}$$

one recognizes a reachability space:

$$\begin{bmatrix} \xi \\ -\eta \end{bmatrix} Z'(I - WZ')^{-1}\mathcal{D}_2$$

Sketch of proof (3)

($\Rightarrow$) S contractive in the above formula requires the basis

$$\begin{bmatrix} \xi \\ -\eta \end{bmatrix} Z'(I - WZ')^{-1}$$

to be J-positive. This is: requiring a positive definite solution to the Lyapunov-Stein equation: $P^{\langle -1 \rangle} = W'PW + \xi'\xi - \eta'\eta$

($\Leftarrow$) if

$$\begin{bmatrix} S & -I \end{bmatrix} = (\Theta_{22} - S_L\Theta_{12})^{-1} \begin{bmatrix} S_L & -I \end{bmatrix} \Theta$$

with Θ having reachability data

$$\left[\begin{array}{c|cc} W' & \xi' & -\eta' \end{array} \right]$$

then interpolation holds!

Norm approximation (summary)

There is a "Caratheodory-Fejer" version of the previous theory: interpolations of a related strictly positive definite matrix C defined by

$$C = \frac{1}{2}(G + G') = LL' = M'M \quad \text{(Cholesky)}$$

(with G a causal part of , and L, M Cholesky factors).

Any interpolation on data from G via the Cayley transform $S = (G + I)^{-1}(G - I)$ produces a low complexity Θ_a and low complexity approximations

$$C_a = \frac{1}{2}(G_a + G_a') = L_a L_a' = M_a' M_a$$

such that $\arg\,\min_X \|LX - I\|_F$ is given by $X = L_a^{-1} d$ where d is a diagonal tending to I when the interpolation proceeds.

Schur-Takagi interpolation

In this case, the interpolation problem does not lead to a lossless Θ, but only to a fully non-singular solution of the Lyapunov-Stein equation and a causal J-unitary Θ:

$$P^{\langle -1 \rangle} = A_\Theta P A'_\Theta + B_{\Theta,1} B'_{\Theta,1} - B_{\Theta,2} B'_{\Theta,2}$$

Let the inertia of *P*:

$$P_k = R_k \begin{bmatrix} I_{p_k} & \\ & -I_{q_k} \end{bmatrix} R'_k$$

This will produce an interpolating and contractive S which is not ‘upper’, but whose lower part is quasi-separable of dimension q (i.e. has a realization of dimension q, that is hence q_k at each index point *k*.)

Application: model order reduction

Given:

- T - a strictly upper matrix to be reduced
- $T = C(I - ZA)^{-1}ZB$ - a "high order" model for T (e.g., a power series)
- Γ - an invertible diagonal, measure of accuracy

Asked:

$T_a - C_a(I - ZA_a)^{-1}ZB_a$ of lowest possible complexity

such that

$$\|(T - T_a)\Gamma^{-1}\|_{Hankel} \leq I$$

where the Hankel norm is 'sup' of the norms of the matrices H_k;
it is a 'strong' norm.

Solution

Assume the original given in 'output normal form', i.e. $\begin{bmatrix} A \\ C \end{bmatrix}$ isometric.

Do the following steps:

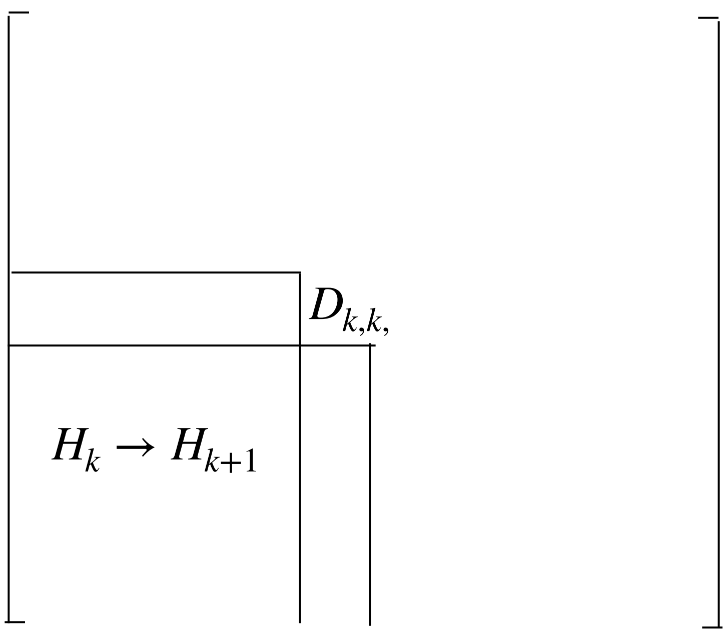
(1) find an orthogonal completion of [A C]: $\begin{bmatrix} A & B_U \\ C & D_U \end{bmatrix}$

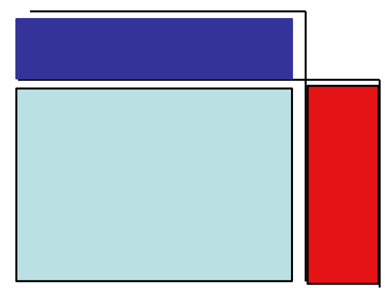
(2) solve the Pick equation with the data $\left[\begin{array}{c|cc} A_U & B_U & B\Gamma^{-1} \end{array} \right]$

(3) find T_a by projecting the causal part out of the Schur-Takagi interpolating operator

Practical solution: essential Hankels

Although the solution presented is exact, it has its drawbacks:
it is somewhat difficult to compute and requires treatment of all
the data. A practical approach (advocated by Chandrasekaran-Gu
as well as Van Dooren e.a.) consists in computing nested SVD's:





subtract blue add red