Rational bases for system identification

Adhemar Bultheel
Department Computer Science
Numerical Approximation and Linear Algebra Group (NALAG)
K.U.Leuven, Belgium
adhemar.bultheel@cs.kuleuven.ac.be
http://www.cs.kuleuven.ac.be/∼ade/



September 2002

Summary

1. Definitions from signals, systems

2. Vector orthogonal polynomials (Jacobi/unitary Hessenberg matrices)

3. Orthogonal rational functions in $L^2(\mathbb{R}, \mu)$, $L^2(\mathbb{T}, \mu)$

4. ORF and inverse (Hessenberg) eigenvalue problem (semiseparable matrices)

Definitions

- A signal is a complex function of a time variable.
  -discrete (f_n), $n \in \mathbb{Z}$ (discrete signal) or
  -continuous $f(t)$, $t \in \mathbb{R}$ (analog signal)

- Define a delay operator $(\mathcal{D}f)_n = f_{n-1}$ $\qquad (\mathcal{D}f(t) = f(t-1))$

- A pulse δ is the Kronecker delta $\qquad$ (Dirac delta function)

- A pulse decomposition:
  $f = (\sum_m f_m \mathcal{D}^m)\delta$ $\qquad (f = \int f(s)\mathcal{D}^s\delta(\cdot)ds)$

- An inner product: $\langle f, g\rangle = \sum \overline{f_n} g_n$ $\qquad (\langle f, g\rangle = \int \overline{f(t)} g(t)dt)$

- The convolution $h = f * g$:
  $h_n = \sum_m f_m g_{n-m}$ $\qquad (h(t) = \int f(s)g(s-t)ds)$

- The (auto)correlation function:
  $(r_{fg})_n = \sum_m \overline{f_m} g_{n+m}$ $\qquad (r_{fg}(t) = \int \overline{f(s)} g(s+t) ds)$
  $(r_{ff})_n = r_n = \sum_m \overline{f_m} f_{n+m}$ $\qquad (r_{ff}(t) = r(t) = \int \overline{f(s)} f(s+t) ds)$

- The Fourier transform: $F = \mathcal{F} f$
  $F(e^{i\omega}) = \sum_m f_m e^{-im\omega}$ $\qquad (F(\omega) = \int f(t) e^{-it\omega} dt)$

- The $\mathcal{Z}$-transform:
  $F(z) = \sum_m f_m z^{-m}$ $\qquad (F(z) = \int f(t) e^{-itz} dt)$
  (compare with (2-sided) Laplace transform $\qquad F(p) = \int f(t) e^{-tp} dt)$
  $F(z)|_{z=e^{i\omega} \in \mathbb{T}} = (\mathcal{F} f)(e^{i\omega})$ $\qquad F(z)|_{z=\omega \in \mathbb{R}} = (\mathcal{F} f)(\omega)$

- power spectrum: $R = \mathcal{F} r = |F|^2$
  white noise: $R =$ constant; normalized constant $= 1$
  colored noise: $g = \mathcal{H} f$, f white noise.

- A (linear) system $\mathcal{H}$ is a (linear) operator that maps a signal (input) to another signal (output) $\mathcal{H}u = y$

- It is stationary or time invariant if $\mathcal{H}\mathcal{D} = \mathcal{D}\mathcal{H}$

- The impulse response: $h = \mathcal{H}\delta$
  $y = \mathcal{H}u = \mathcal{H}\left(\sum u_m \mathcal{D}^m\right)\delta = \sum u_m \mathcal{D}^m(\mathcal{H}\delta) = \left(\sum u_m h_{n-m}\right) = h * u$
  $y = \mathcal{H}u = \mathcal{H}\left(\int u(s)\mathcal{D}^s\right)\delta = \int u(s)\mathcal{D}^s(\mathcal{H}\delta) = \int u(s)h(\cdot - s) = h * u$
  $y = h * u = u * h = \sum h_m \mathcal{D}^m u = H(\mathcal{D}^{-1})u$ hence $\mathcal{H} = H(\mathcal{D}^{-1})$.

- The transfer function: $H(z) = (\mathcal{Z}h)(z) = |H(z)|e^{i\Phi(z)}$
  amplitude: $|H(z)|$ and phase: $\Phi(z)$.

- A causal system: if $u = 0$ before time 0, then $y = 0$ before time 0.

- A (BIBO) stable system: if $\|u\| < \infty \Rightarrow \|\mathcal{H}u\| < \infty$

- For a causal stable system: $H \in H^2(\mathbb{E})$ $\qquad$ $(H \in H^2(\mathbb{L}))$

- The inverse system: if $H(z) \neq 0$, $z \in \mathbb{T}$ ($z \in \mathbb{R}$), inverse $= 1/H(z)$.
  Stable & causal then inverse if poles of H in $\mathbb{D}$ (in $\mathbb{U}$)
  Inverse is table & causal if zeros of H in $\mathbb{D}$ (in $\mathbb{U}$)
  the latter is called minimal phase

- A system is all pass if $|H(z)| = 1$ (hence $\|Y\| = \|HU\| = \|U\|$)

- FIR/MA: H is polynomial in z^{-1} (otherwise IIR)
  AR: $H =$ cnst/polynomial in z^{-1}
  ARMA: $H =$ rational in z^{-1}
  proper: deg(numerator) $\leq$ deg(denominator) ($\Leftarrow$ stable & causal)

State Space

If H is proper rational then partial fraction expansion:

$$H(z) = D + \frac{c_1}{z - \alpha_1} + \cdots + \frac{c_n}{z - \alpha_n} = C(zI - A)^{-1}B + D$$

where

$$A = \mathrm{diag}(\alpha_1, \ldots, \alpha_n), \quad B = [1, \ldots, 1]^T, \quad C = [c_1, \ldots, c_n], \quad D = H(\infty).$$

This is called a state space description of the system. It is characterized by the 4 matrices (A, B, C, D). Such a state space description is not unique.

If T is an invertible matrix, then $(A,B,C,D) \equiv (A',B',C',D')$ if

$$\begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix} = \begin{bmatrix} T & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} T & 0 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} TAT^{-1} & TB \\ CT^{-1} & D \end{bmatrix}$$

since $C(zI-A)^{-1}B + D = C'(zI-A')^{-1}B' + D'$.
A minimal realization: size A is minimal.
Then $\sigma(A) =$ poles of the transfer function.

$$H(z) = \sum \frac{h_k}{z^k} = D + C(zI-A)^{-1}B = D + \frac{CB}{z} + \cdots + \frac{CA^{k-1}B}{z^k} + \cdots.$$

$Y(z) = [C(zI-A)^{-1}B + D]U(z) = CX(z) + DU(z)$ with
$X(z) = (zI-A)^{-1}BU(z)$ or $zX(z) = AX(z) + BU(z)$. This

$$\begin{aligned} zX(z) &= AX(z) + BU(z) \\ Y(z) &= CX(z) + DU(z) \end{aligned}$$

is the z-domain formulation of

$$x_{n+1} = Ax_n + Bu_n \qquad \dot{x}(t) = Ax(t) + Bu(t)$$
$$y_n = Cx_n + Du_n \qquad y(t) = Cx(t) + Du(t)$$

$$\dot{x}(t) = \frac{1}{i}\frac{dx(t)}{dt}$$

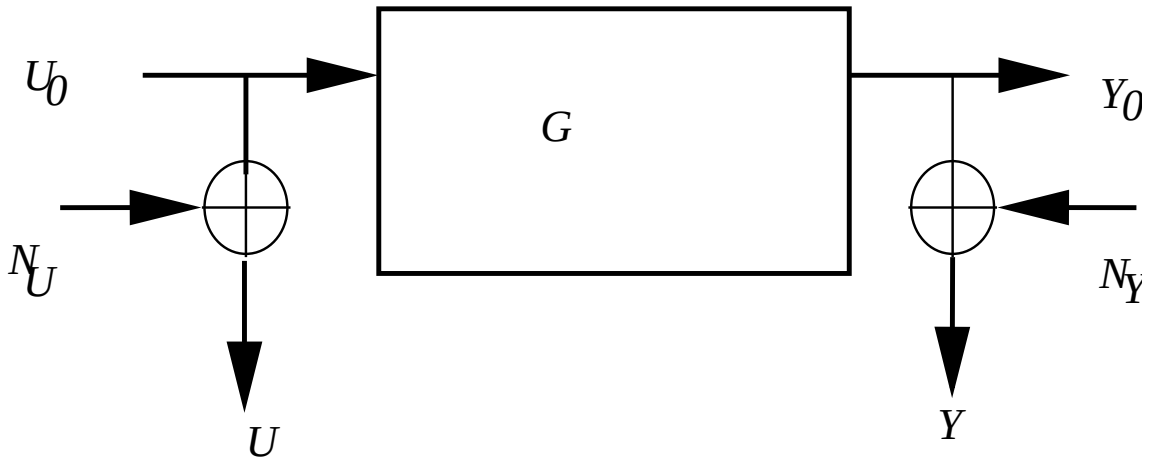
if system is causal and stable and state is zero at time zero.
x is the state.

SS description very powerful because of
- linear algebra techniques applicable
- generalization of SISO to MIMO is relatively simple

	n	n_u
n	A	B
n_y	C	D

- many other advantages

Identification problem



$Y = G(U - N_U) + N_Y = GU + (N_Y - GN_U) = GU + V, \quad V = H_0 E$, E white noise, H_0 some coloring filter.

It is assumed we know Y and U (or G) in a number of frequency points: $z_k \in \mathbb{T}$ or $\mathbb{R}$ and some information about the noise (covariances)

Frequency Domain Identification

Estimate G as $\hat{G}$ such that we minimize $\|Y - \hat{Y}\|_{w_y}$ or $\|G - \hat{G}\|_{w_g}$ where $\|F\|_w^2 = \sum_{k=1}^N w_k^2 |F(z_k)|^2$.
Note with $w_g = w_y U$:

$$\|Y - \hat{Y}\|_{w_y} = \|GU - \hat{G}U\|_{w_y} = \|(G - \hat{G})U\|_{w_y} = \|G - \hat{G}\|_{w_g},$$

E.g. measurements $\{G(z_j)\}_{j=1}^N$, variance $\{\sigma_j\}_{j=1}^N$, then set $w_g = 1/\sigma$.

$$\|G - \hat{G}\|_{w_g} = \left\| \frac{Y}{U} - \frac{B}{A} \right\|_{w_g} = \left\| \frac{YA - BU}{AU} \right\|_{w_g} = \|YA - BU\|_w, \quad w = \frac{w_g}{AU}$$

In ML estimation:
$w^{-2} = \mathrm{Cov}(AY - BU) = \sigma_Y^2 |A|^2 + \sigma_U^2 |B|^2 - 2\mathrm{Re}(\sigma_{YU}^2 A\overline{B})$.

Linear/Nonlinear problem

Suppose $\hat{G}$ parametrised by parameter set θ.
Cost function: $C(\theta) = C(L(\theta), N(\theta))$, L linear, N nonlinear.
If N does not depend on θ, the problem becomes a linear LSQ problem.

Solving for $\min C(\theta)$ directly is costly and not convex so convergence to global min is not guaranteed, unless we use a good starting point.

The method of solution: ITERATE $\theta^{(k+1)} = \operatorname{argmin} C(L(\theta), N(\theta^{(k)}))$.
May work in some cases to find global solution.
Requires a stable and efficient method to solve the linear LSQ problem.

Need constraints. E.g. system stability needs system poles in $\mathbb{D}$. So have to solve a constrained and weightes NLSQ problem in $\mathbb{C}$.

ITERATE converges, then use as starting point for $\min C(\theta)$ with standard routine.

Polynomial method

$\min C = \min \|YA - UB\|_w^2$ w may depend upon (estimated) solution.

Set $W = w[Y \ -U], P = [A \ B]^T \Rightarrow w(YA - UB) = WP$

$$C = \|WP\|^2 = \|P\|_W^2 = \sum_{j=1}^{N} P(z_j)^H W_j^H W_j P(z_j), \quad P \in \mathbb{P}_n^{2\times 1}.$$

$\langle f, g\rangle_W = \sum f(z_j)^H W_j^H W_j g(z_j)$ a p.d. inner product in $\mathbb{P}_n^{2\times 1}$ if $n \leq N$.

$\min \|P\|_W$ needs constraint: e.g. P "monic".

How to parametrize P?

$\dim \mathbb{P}^{2\times 1} = 2(n+1)$:

$$\left[\begin{array}{cccc|cccc} 1 & z^{-1} & \cdots & z^{-n} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & z^{-1} & \cdots & z^{-n} \end{array}\right]$$

or

$$\left[\begin{array}{cc|cc|c|cc} 1 & 0 & z^{-1} & 0 & \cdots & z^{-n} & 0 \\ 0 & 1 & 0 & z^{-1} & \cdots & 0 & z^{-n} \end{array}\right]$$

Simpler to orthogonalize $\{I, z^{-1}I, \ldots, z^{-n}I\}$ w.r.t. matrix-valued SP $\langle f, g\rangle_W = \sum f(z_j)^H W_j^H W_j g(z_j)$ in $\mathbb{P}^{2\times 2}$

Gives orthogonal block polynomials ϕ_k: $\langle \phi_k, \phi_l\rangle_W = \delta_{k-l} I$
to make unique choose e.g. l.c. upper triangular.

Write $P = \sum_{k=0}^{n} \phi_k \lambda_k = \Phi(z)\boldsymbol{\lambda}$, $\phi_k \in \mathbb{P}_k^{2\times 2}$, $\lambda_k \in \mathbb{C}^{2\times 1}$, $\lambda_n \neq 0$.

$\|P\|_W^2 = \|\Phi\boldsymbol{\lambda}\|_W = \boldsymbol{\lambda}^H \boldsymbol{\Phi}^H \mathbf{W}^H \mathbf{W} \boldsymbol{\Phi} \boldsymbol{\lambda}$, $\boldsymbol{\Phi} = [\phi_j(z_i)]$, $\mathbf{W} = \operatorname{diag}(W_k)$.

ϕ_k orthonormal $\Rightarrow \boldsymbol{\Phi}^H \mathbf{W}^H \mathbf{W} \boldsymbol{\Phi} = \mathbf{Q}^H \mathbf{Q} = \mathbf{I}$, $\mathbf{Q} = \mathbf{W}\boldsymbol{\Phi}$ is unitary

$\min \|P\|_W^2 = \min \|\boldsymbol{\lambda}\|^2 = \lambda_n^H \lambda_n \Rightarrow P = \phi_n \lambda_n$.

Recurrence relation

$$z^{-1}\phi_{k-1}(z) = \phi_k(z)\eta_{kk} + \cdots + \phi_0(z)\eta_{0k}, \quad k = 1, \ldots, n+1$$

$$z^{-1}[\phi_0 \cdots \phi_{n-1}\ \phi_n] = [\phi_0 \cdots \phi_n \phi_{n+1}] \begin{bmatrix} \eta_{01} & \cdots & \eta_{0n} & \eta_{0,n+1} \\ \eta_{11} & \cdots & \eta_{1n} & \eta_{1,n+1} \\ & \ddots & \vdots & \vdots \\ & & \eta_{nn} & \eta_{n,n+1} \\ & & & \eta_{n+1,n+1} \end{bmatrix}$$

$$z^{-1}\Phi(z) = \Phi(z)\mathbf{H} + [0 \cdots 0 \phi_{n+1}\eta_{n+1,n+1}]$$

Now write this down for $\mathbf{z} = [z_1, \ldots, z_N]^T$.

$$\tilde{\mathbf{Z}}^{-1}\mathbf{\Phi} = \mathbf{\Phi}\mathbf{H} + [0 \ \cdots \ 0 \ \phi_{n+1}(\mathbf{z})\eta_{n+1,n+1}]$$

$$\tilde{\mathbf{Z}} = \mathrm{diag}(\mathbf{z}I_2), \quad \mathbf{\Phi} = [\phi_j(z_i)] = \Phi(\mathbf{z})$$

$$\tilde{\mathbf{Z}}^{-1}\mathbf{\Phi} = \mathbf{\Phi}\mathbf{H} + [0 \ \cdots \ 0 \ \phi_{n+1}(\mathbf{z})\eta_{n+1,n+1}]$$

multiply with $\mathbf{W}$ and set $\mathbf{Q} = \mathbf{W}\mathbf{\Phi}$ $(\mathbf{W}\tilde{\mathbf{Z}}^{-1} = \mathbf{Z}^{-1}\mathbf{W})$

$$\mathbf{Z}^{-1}\mathbf{Q} = \mathbf{Q}\mathbf{H} + [0 \ \cdots \ 0 \ \mathbf{W}\phi_{n+1}(\mathbf{z})\eta_{n+1,n+1}]$$

multiply from the left with $\mathbf{Q}^H$ and because $\phi_{n+1} \perp_W \mathbb{P}_n$

$$\mathbf{Q}^H\mathbf{Z}^{-1}\mathbf{Q} = \mathbf{H}.$$

For $N = n + 1$, $\mathbf{Q}$ is square unitary matrix and thus

$$\boxed{\mathbf{Q}^H\mathbf{Z}^{-1}\mathbf{Q} = \mathbf{H} \Leftrightarrow \mathbf{Z}^{-1} = \mathbf{Q}\mathbf{H}\mathbf{Q}^H}$$

In the analog case: $\mathbf{Z}$ is real $\Rightarrow \mathbf{H} = \mathbf{Q}^H \mathbf{Z}^{-1} \mathbf{Q} = \mathbf{H}^H$.
$\Rightarrow \mathbf{H}$ block tridiagonal $\Rightarrow$ scalar pentadiagonal
(block 3-term recurrence relation)

In the digital case: $\mathbf{Z}^{-H}\mathbf{Z}^{-1} = \mathbf{I} \Rightarrow \mathbf{H}^H\mathbf{H} = \mathbf{Q}^H\mathbf{Z}^{-1}\mathbf{Q}\mathbf{Q}^H\mathbf{Z}^{-H}\mathbf{Q} = \mathbf{I}$.
$\Rightarrow \mathbf{H}$ is unitary (block) upper Hessenberg.
(Szegő-type recurrence relation)

$\mathbf{Z}^{-1} = \mathbf{Q}\mathbf{H}\mathbf{Q}^H$ is Inverse eigenvalue problem.
Not completely defined. Need some condition on the $\mathbf{W}$

Set $\mathbf{w} = \begin{bmatrix} W_1 \\ \vdots \\ W_N \end{bmatrix} \Rightarrow \langle \phi_0, \phi_0 \rangle_W = \phi_0^H \sum W_k^H W_k \phi_0 = \phi_0^H \mathbf{w}^H \mathbf{w} \phi_0 = I_2.$

With $\eta_{00} = \phi_0^{-1} = \mathrm{sqrt}(\mathbf{w}^H\mathbf{w})$: $\langle \phi_0, I \rangle = \eta_{00}$. So $\mathbf{Q}^H\mathbf{w} = \mathbf{\Phi}^H\mathbf{W}^H\mathbf{w} = \mathbf{\Phi}^H\mathbf{W}^H\mathbf{W}[I, \ldots, I]^H = [\eta_{00}^H, 0, \ldots, 0]^H \Rightarrow \boxed{\mathbf{Q}^H\mathbf{w} = \mathbf{e}_1\eta_{00}}$

Inverse Eigenvalue Problem

$$\mathbf{Q}^H \mathbf{Z}^{-1} \mathbf{Q} = \mathbf{H} \quad \& \quad \mathbf{Q}^H \mathbf{w} = \mathbf{e}_1 \eta_{00} \quad \& \quad \eta_{00}^H \eta_{00} = \mathbf{w}^H \mathbf{w}$$

or

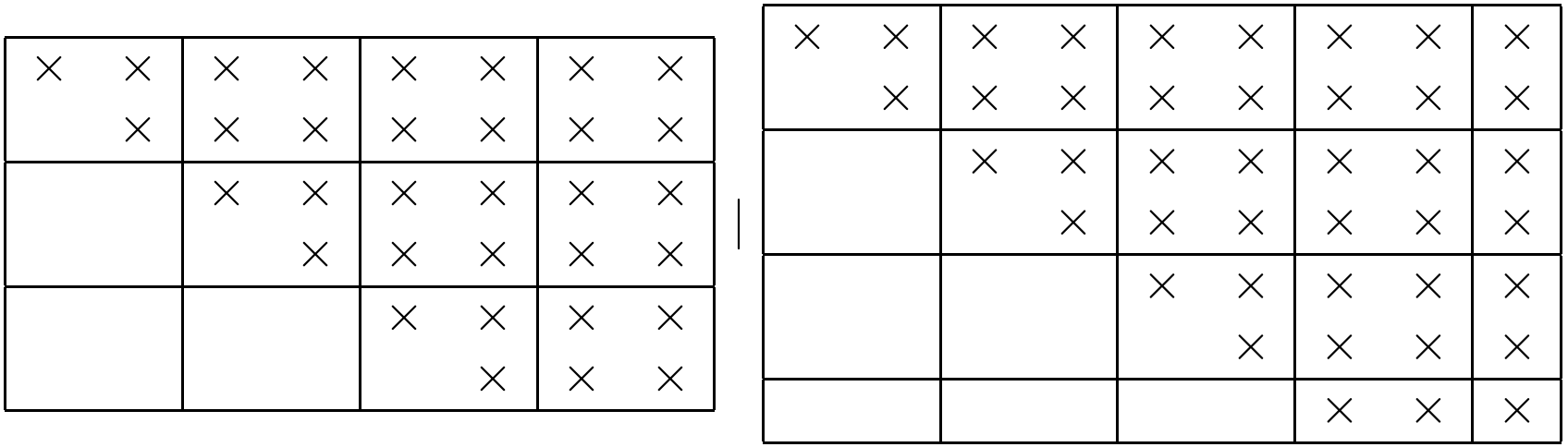
$$\mathbf{Q}^H \left[\begin{array}{c|c} \begin{matrix} W_1 \\ \vdots \\ W_N \end{matrix} & \mathbf{Z}^{-1} \end{array}\right] \left[\begin{array}{c|c} I_2 & \\ \hline & \mathbf{Q} \end{array}\right] = \left[\begin{array}{c|c} \begin{matrix} \eta_{00} \\ \\ 0 \end{matrix} & \mathbf{H} \end{array}\right]$$

Hessenberg matrix

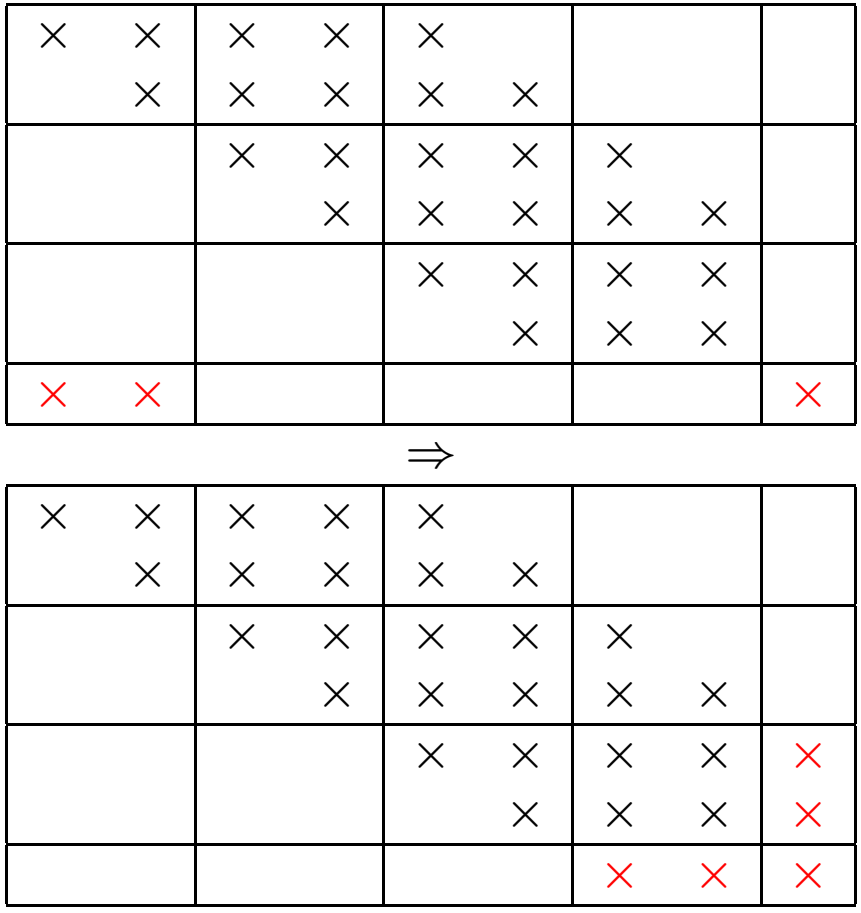
$$\left[\begin{array}{c|ccc} w_1 & z_1^{-1} & & \\ w_2 & & z_2^{-1} & \\ & & & \ddots \\ w_N & & & z_N^{-1} \end{array}\right]$$

unitary similarity $\Rightarrow$ transf.

Block Upper Hessenberg



Recursively: continuous time



Recursively: discrete time

Szegő-type recurrence:
$$\begin{array}{rcl} \phi_n \sigma_n & = & z^{-1}\phi_{n-1} + \phi_{n-1}^* \gamma_n \\ \phi_n^* \sigma_n^H & = & z^{-1}\phi_{n-1}\gamma_n^H + \phi_{n-1}^* \end{array}$$

$\sigma_n, \gamma_n \in \mathbb{C}^{2\times 2}$ (block Schur parameters)

Store the block upper Hessenberg matrix in factored form $H = G_1 G_2 \cdots G_m$

$$G_k = \begin{bmatrix} I_{2(k-1)} & & & \\ & -\gamma_k & \hat{\sigma}_k & \\ & \sigma_k & \hat{\gamma}_k & \\ & & & I_{\ldots} \end{bmatrix}$$

Chasing the elements to block upper Hessenberg form by similarity transformations percolating through the product requires operations on 3×3 or 5×5 blocks $\Rightarrow$ fast algorithm.

Remarks

- Fast algorithm ($O(N^2)$ operations)
- Economize if $n < N$ (operate on $N \times n$ matrix)
- can do in real arithmetic for complex data
- Works for MIMO systems (OVP are $s \times 1$, $s = (n_u + n_y)n_y$)
- Several degree structures are possible (column permutations)
- Representation of $\hat{G}$ is optimal: representation = orthogonal basis
  If the weights change then basis changes and is optimal.
  The condition number of systems to solve = 1.
- nonlinear (ML) solution: fix last basis, condition number grows moderately

real computations for $\mathbb{T}$

We can preprocess complex data to do real arithmetic.

2 data points $e^{i\omega_k}, e^{-i\omega_k} \in \mathbb{T}$ with corresponding weights W_k and $\overline{W}_k$.

Define $Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -i \end{bmatrix}$

Then $Q^H \begin{bmatrix} W_k \\ \overline{W}_k \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} W_k \\ \overline{W}_k \end{bmatrix} = \sqrt{2} \begin{bmatrix} \mathrm{Re}(W_k) \\ \mathrm{Im}(W_k) \end{bmatrix}$

and $Q^H \begin{bmatrix} e^{i\omega_k} & 0 \\ 0 & e^{-i\omega_k} \end{bmatrix} Q = \begin{bmatrix} \cos\omega_k & \sin\omega_k \\ -\sin\omega_k & \cos\omega_k \end{bmatrix}$

MIMO system

$\vec{M}$= stack columns in long vector

model $A^{-1}B$: A is $n_y \times n_y$, B is $n_y \times n_u$

$$\vec{P} = \begin{bmatrix} \vec{A} \\ \vec{B} \end{bmatrix}$$ this is $n_y(n_y + n_u)$ long

$$E = AY - BU \ (n_y \times 1) \Rightarrow \min \sum_k \vec{E}^H(z_k) C_{\vec{E}}^{-1}(z_k) \vec{E}(z_k)$$

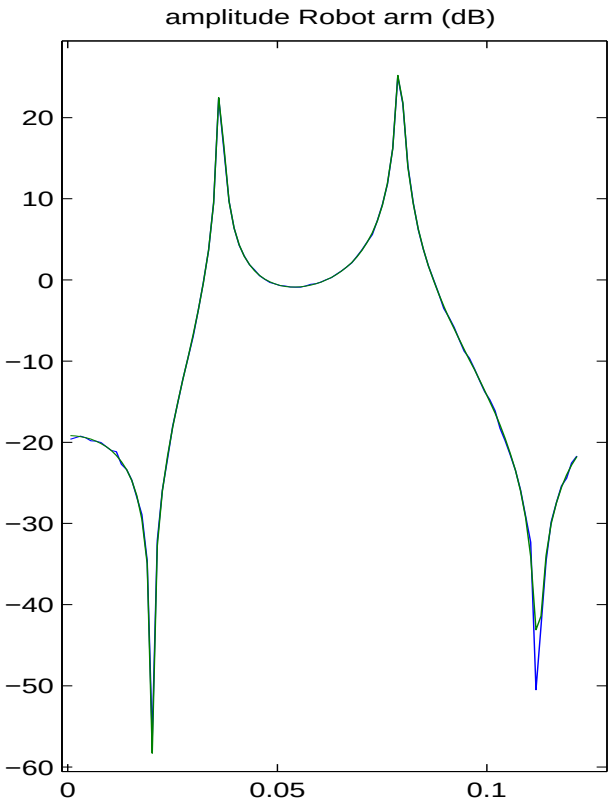
$$\vec{E} = ([Y^T \quad -U^T] \otimes I_{n_y})\vec{P}$$

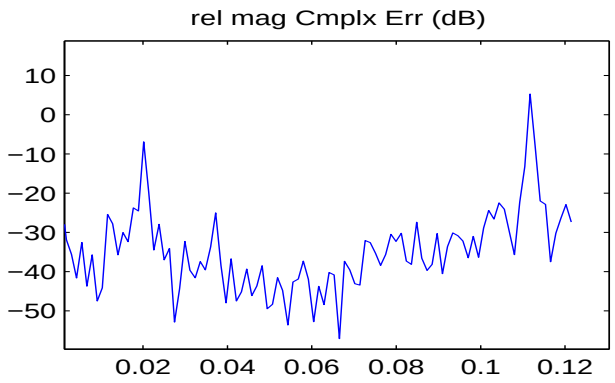
$$C_{\vec{E}} = \mathrm{Cov}(\vec{E}) = C_1 + C_2 - 2C_3 \quad \begin{cases} C_1 = (I_{n_u} \otimes A) C_{\vec{Y}} (I_{n_u} \otimes A^H) \\ C_2 = (I_{n_u} \otimes B) C_{\vec{U}} (I_{n_u} \otimes B^H) \\ C_3 = \mathrm{Herm}[(I_{n_u} \otimes A) C_{\vec{Y}\vec{U}} (I_{n_u} \otimes B^H)] \end{cases}$$

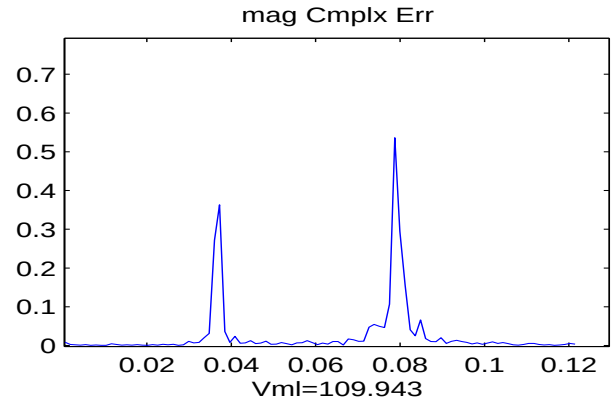
$\min \|\vec{P}\|_W$, $W = ([Y^T \quad -U^T] \otimes I_{n_y}) C_{\vec{E}}^{-1/2}$ has size $1 \times (n_y + n_u) n_y$

Robot arm

order [6/6], 100 data points, 200Hz, discrete time

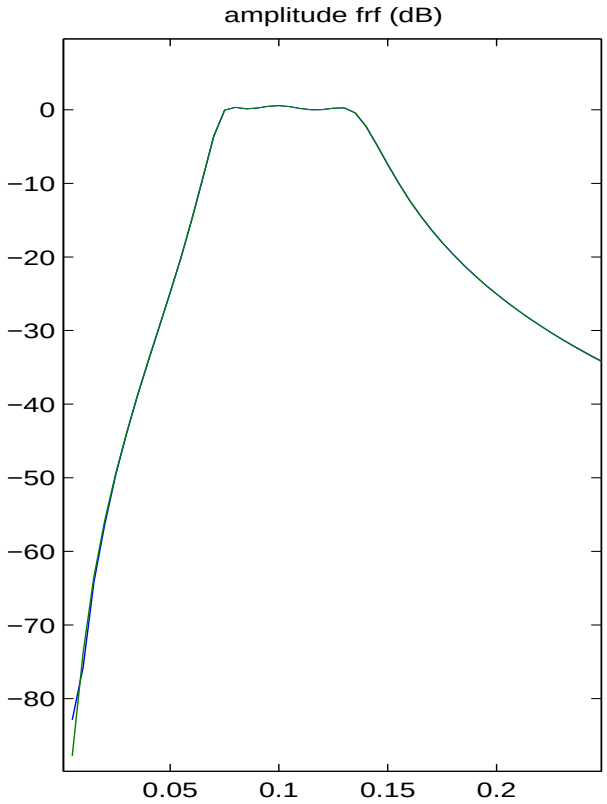


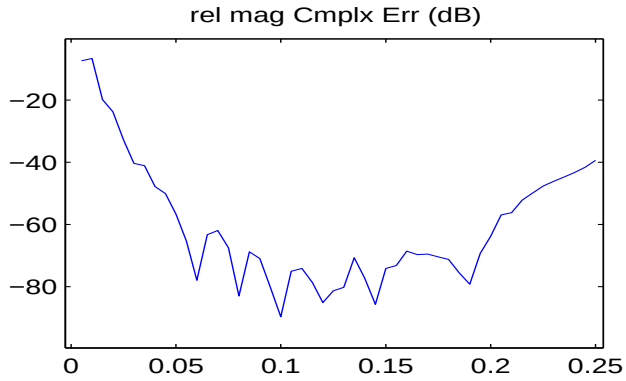


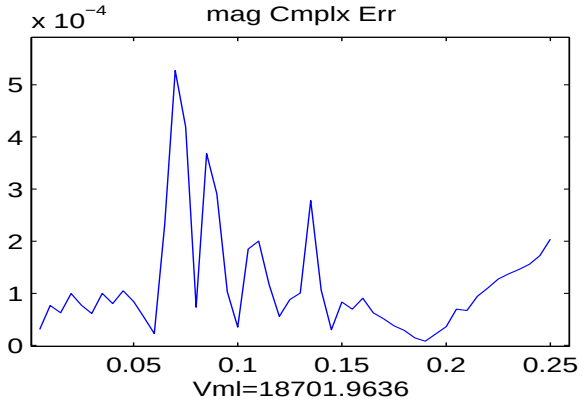


Band pass filter

order [6/6], 50 data points, 20KHz, discrete time

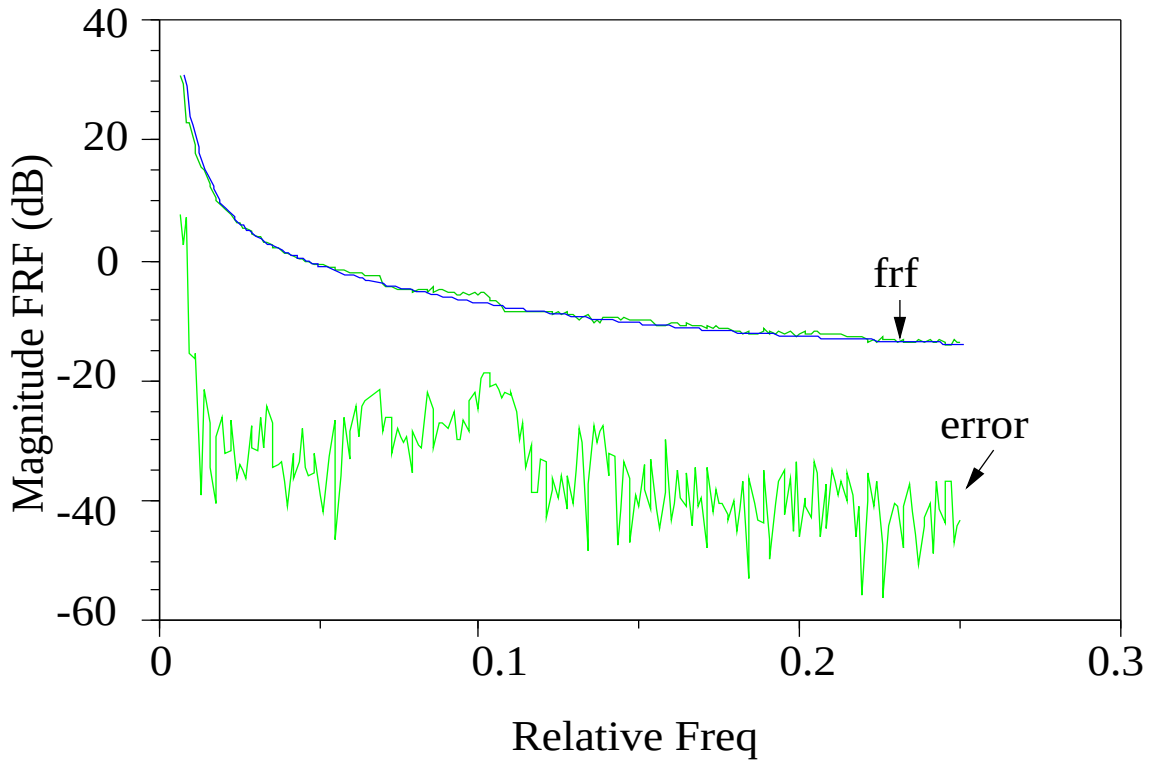






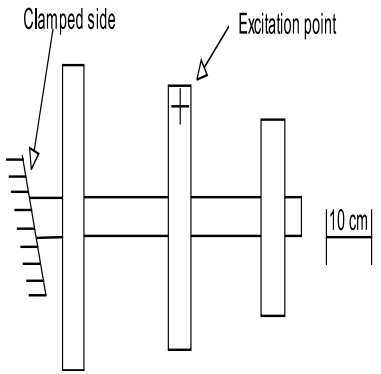
CD's radial servo system

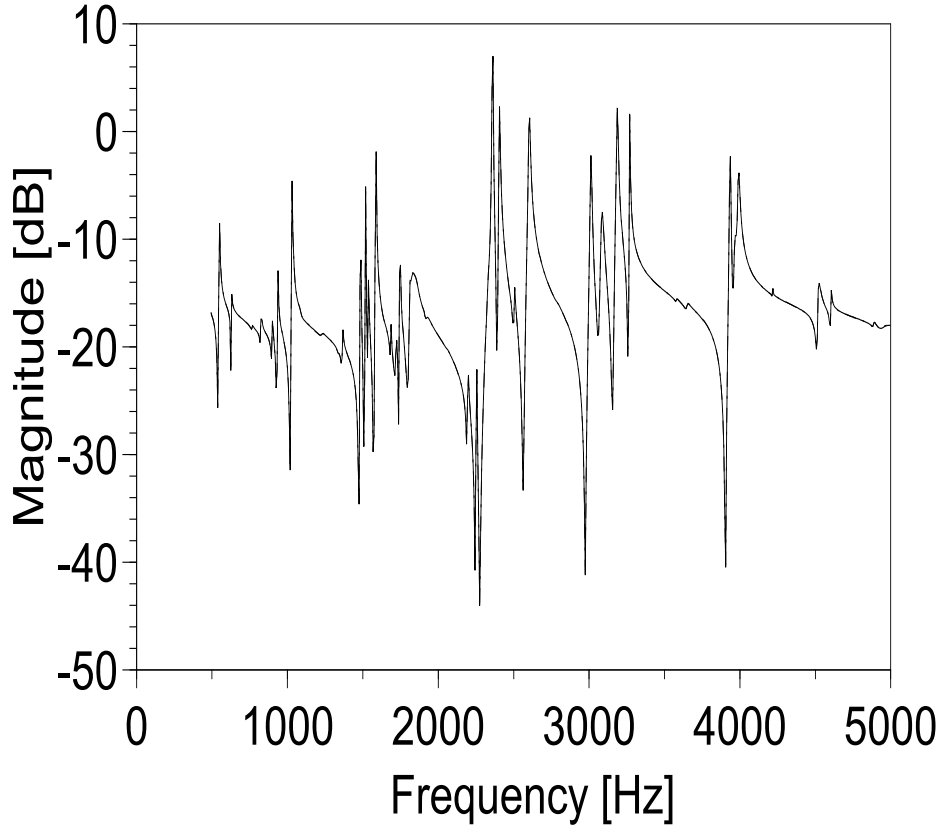
sampling frequency 9.7 kHz, order [5/5]

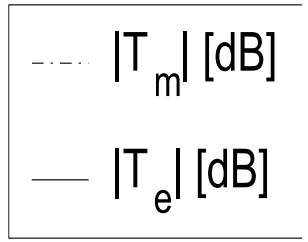


Mechanical structure

order [120/120]







References OPV

[1] A. Bultheel, M. Van Barel, Y. Rolain, J. Schoukens. Robust rational approximation for identification, *Proc. CDC conference*, 2001

[2] R. Pintelon, Y. Rolain, A. Bultheel, M. Van Barel. *Numerically robust frequency domain identification of multivariable systems*, (2002).

[3] M. Van Barel, A. Bultheel, Discrete linearized least squares approximation on the unit circle, *J. Comput. Appl. Math.* 50 (1994) 965-972.

[4] A. Bultheel, M. Van Barel. Vector orthogonal polynomials and least squares approximation, *SIAM J. Matrix Anal. Appl.* 16 (1995) 863-885.

[5] M. Van Barel, A. Bultheel, Orthogonal polynomial vectors and least squares approximation for a discrete inner product, *Electron. Trans. Numer. Anal.* 3 (1995) 1-23.

Identification and ORF

Identification problem: find rational function $\hat{G}$ such that

$$\min \sum_{k=1}^{N} |w_k(G(z_k) - \hat{G}(z_k))|^2$$

The w_k can possibly depend on the solution, but in the present situation we use an estimate of the solution so that we can consider w_k as known.

Looking for a rational $\hat{G}$:

$$\hat{G} \in \mathcal{L}_n = \left\{ \frac{p_n(z)}{\prod_{j=1}^{n}(1 - \frac{\alpha_j}{z})} : p_k \in \mathbb{P}_n \right\}$$

Note that if all $\alpha_k = 0 \Rightarrow \mathcal{L}_n = \mathbb{P}_n$.

So why not represent it w.r.t. a basis of rational functions:

$$\mathcal{L}_n = \text{span}\{B_0, \ldots, B_n\} \quad \text{with} \quad B_0 = 1, \quad B_k = \prod_{j=1}^{k} \frac{1}{z - \alpha_j}, \quad k \geq 1$$

For numerical stability reasons: use orthogonalize the basis $\{B_k\} \rightarrow \{\phi_k\}$

$$\hat{G}(z) = \sum_{k=0}^{n} \lambda_k \phi_k(z),$$

ϕ_k orthogonal basis functions with respect to an appropriate inner product. The inner product can be discrete or continuous, but is in general with respect to a weight (measure) on $\mathbb{T}$ or $\mathbb{R}$.

$$\langle f, g \rangle_\mu = \int f(z)\overline{g(z)} d\mu(z) \quad \text{or} \quad \langle f, g \rangle_w = \sum_{j=1}^{N} f(z_j)\overline{g(z_j)} w_j.$$

Takenaka-Malmquist basis

Work by Ninnes, Van den Hof, Heuberger, Bokor, de Hoog, and others:

Use ORF with respect to Lebesgue measure on $\mathbb{T}$. These have explicit expressions

$$\phi_n(z) = \frac{\sqrt{1-|\alpha_n|^2}}{z-\alpha_n} \prod_{k=1}^{n-1} \frac{1-\overline{\alpha}_k z}{z-\alpha_k}$$

but that does not help for the condition number of $\mathbf{\Phi}^H \mathbf{W}^H \mathbf{W} \mathbf{\Phi}$ in the linear least squares problem.

And/or use a finite number of α_k that are cyclically repeated.

State space formulation

- parahermitian conjugate $f_*(z) = [f(1/\overline{z})]^H$ or $f_*(z) = [f(\overline{z})]^H$
- All-pass/inner function: $G(z)G_*(z) = I/$ analytic
- balanced state space: $G(z) = C(zI - A)^{-1}B + D$
  $X_c = \sum_{k=0}^{\infty}(A^kB)(A^kB)^H = \Sigma = X_o = \sum_{k=0}^{\infty}(CA^k)^H(CA^k)$
- all-pass+balanced $\Rightarrow \Sigma = I$ and

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}^H = I = \begin{bmatrix} A & B \\ C & D \end{bmatrix}^H\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

- matrix valued inner product: $\langle F, G\rangle = \frac{1}{2\pi}\int F(e^{i\omega})[G(e^{i\omega})]^H d\omega$

$X_c = \left\langle (zI - A)^{-1}B, (zI - A)^{-1}B \right\rangle,$
$X_o = \left\langle [C(zI - A)^{-1}]^H, [C(zI - A)^{-1}]^H \right\rangle$

• Thus: take balanced SS realization (A_n, B_n, C_n, D_n) of Blaschke product of degree n, and set

$$(zI - A_n)^{-1} B_n = [\phi_1(z), \phi_2(z), \ldots, \phi_n(z)]^T$$

and the ϕ_k are ORF.

• Product of Blaschke factors $\Rightarrow$
recursive SS $(A_k, B_k, C_k, D_k) \Rightarrow$ Takenaka-Malmquist ORF.
SS for $(1 - \overline{\alpha} z)/(z - \alpha)$ is

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \alpha & \sqrt{1 - |\alpha|^2} \\ \sqrt{1 - |\alpha|^2} & -\overline{\alpha} \end{bmatrix}$$

• Hambo basis: inner fct = (Blaschke prod)^k $\Rightarrow$ cyclic repetition of finite number of poles.

Special case (1)

Laguerre basis: all $\alpha_k = \alpha$:

$$\phi_n(x) = \frac{\sqrt{1-|\alpha|^2}}{z-\alpha}\left(\frac{1-\overline{\alpha}z}{z-\alpha}\right)^{n-1}.$$

Laguerre polynomials $L_n(x)$: $\int_0^\infty e^{-x} L_n(x) L_m(x) dx = \delta_{n-m}$.

Define $\hat{\phi}_n(x) = e^{-x/2} L_n(x)$

Take Fourier transform: $c\phi_n(x) = \frac{\sqrt{2\alpha}}{jx+\alpha}\left(\frac{jx-\alpha}{jx+\alpha}\right)^n$.

This is the continuous time Laguerre basis (orthogonal on $\mathbb{R}$).

A bilinear (Jukowky) transform to $\mathbb{T}$ gives the discrete time Laguerre basis.

Special case (2)

Kautz basis: pairs of complex conjugate poles.

$$\begin{array}{lcl} \phi_{2k-1}(z) & = & C_k(1-\beta_k z)\psi_k(z), \\ \phi_{2k}(z) & = & C_k'(1-\beta_k' z)\psi_k(z), \end{array} \quad \psi_k(z) = \frac{\prod_{j=1}^{k-1}(1-\overline{\alpha}_j z)(1-\alpha_j z)}{\prod_{j=1}^{k}(z-\alpha_j)(z-\overline{\alpha}_j)}$$

Obtained from Legendre polynomials: $\int_{-1}^{1} P_n(x)P_m(x)dx = \delta_{n-m}$.

Define $\hat{\phi}(t) = e^{-\gamma t}P_n(2e^{-\gamma t}-1) \Rightarrow \perp \text{ on } [0,\infty)$

Taking Fourier transform: $c_n\phi_n(x) = \frac{c_n}{jx+(n+1/2)\gamma}\prod_{k=0}^{n-1}\frac{jx-(k+1/2)\gamma}{jx+(k+1/2)\gamma}$.

This is the continuous time Kautz basis (orthogonal on $\mathbb{R}$).

Set $\alpha_k = \frac{2-\gamma(2k+1)}{2+\gamma(2k+1)}$ and bilinear transform to $\mathbb{T}$ gives discrete time Kautz system for a specific choice of the α_k.

Recursive state space

$\Phi_n(z)^T = (zI - A_n)^{-1}B_n$ with (A_n, B_n, C_n, D_n) a balanced state space realization of a general Blaschke product $\prod_{k=1}^{n}(1-\bar{\alpha}_k z)/(z-\alpha_k)$

If $G_i \leftrightarrow (A_i, B_i, C_i, D_i)$, $i = 1, 2$ (inner $\leftrightarrow$ balanced) then
$G = G_1 G_2 \leftrightarrow (A, B, C, D)$ with

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ B_2C_1 & A_2 & B_2D_1 \\ \hline D_2C_1 & C_2 & D_2D_1 \end{array}\right]$$

Thus with
$G_1 = \prod_{k=1}^{n-1} \frac{1-\bar{\alpha}_k z}{z-\alpha_k} \leftrightarrow (A_{n-1}, B_{n-1}, C_{n-1}, D_{n-1})$ and
$G_2 = \frac{1-\bar{\alpha}_n z}{z-\alpha_n} \leftrightarrow (\alpha_n, \sqrt{1-|\alpha_n|^2}, \sqrt{1-|\alpha_n|^2}, -\bar{\alpha}_n)$

$$A_n = \begin{bmatrix} A_{n-1} & 0 \\ X_n & \alpha_n \end{bmatrix}, \quad B_n = \begin{bmatrix} B_{n-1} \\ b_n \end{bmatrix}$$

$$A_n = \begin{bmatrix} A_{n-1} & 0 \\ X_n & \alpha_n \end{bmatrix}, \quad B_n = \begin{bmatrix} B_{n-1} \\ b_n \end{bmatrix}$$

$$X_n = \sqrt{1-|\alpha_n|^2}C_{n-1}, \qquad b_n^T = \sqrt{1-|\alpha_n|^2}D_{n-1},$$
$$C_n = [(-\bar{\alpha}_n)C_{n-1} \mid \sqrt{1-|\alpha_n|^2}], \qquad D_n = (-\bar{\alpha}_n)D_{n-1}.$$

Hambo basis

Take $G(z) = \prod_{j=1}^{d} \frac{1-\bar{\alpha}_j z}{z-\alpha_j}$ with balanced state space realization (A, B, C, D).

Define $V_k(z) = (zI - A)^{-1} B G^{k-1}(z)$, set $V_k(z) = [\phi_{k1}, \ldots, \phi_{kd}]^T$. Then the $\{\phi_{k,i} : i = 1, \ldots, d; k = 0, 1, \ldots\}$ form an orthonormal set ORF for $\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\omega}) \overline{g(e^{i\omega})} d\omega$.

Note that $v_k(j) = V_k(\mathcal{D})\delta_j$ is a vector in the state space.

Orthogonal state space basis in the sense that $[V_k, V_k] = I \Leftrightarrow \{v_k, v_l\} = I$ with
$[f, g] = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\omega})[g(e^{i\omega})]^H d\omega$ and $\{f, g\} = \sum_j f(j)[g(j)]^H$.

Identification

Make row vector of ORF: $\Phi(z) = [(zI - A)^{-1}B]^T$.

Model: $\hat{G} = \Phi\boldsymbol{\lambda}$, set $\boldsymbol{\Phi} = \Phi(\mathbf{z})$, $\mathbf{G} = G(\mathbf{z})$, $\mathbf{W} = \mathrm{diag}(\mathbf{w})$, $\mathbf{E} = \mathbf{G} - \boldsymbol{\Phi}\boldsymbol{\lambda}$.

minimize $\|G - \hat{G}\|_w = \sum_k |w_k[G(z_k) - \Phi(z_k)\boldsymbol{\lambda}]|^2 = \mathbf{E}^H\mathbf{W}^H\mathbf{W}\mathbf{E} = \|\mathbf{E}\|_w^2$.

Hence solve $\boldsymbol{\Phi}\boldsymbol{\lambda} = \mathbf{G}$ in weighted least squares sense: $\boldsymbol{\lambda} = \boldsymbol{\Phi}^\dagger\mathbf{G}$,

$\boldsymbol{\Phi}^\dagger = [\boldsymbol{\Phi}^H\mathbf{W}^H\mathbf{W}\boldsymbol{\Phi}]^{-1}[\boldsymbol{\Phi}^H\mathbf{W}^H\mathbf{W}]$

However $\delta_{k-l} = \langle\phi_k, \phi_l\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} \overline{\phi_k(e^{i\omega})}\phi_l(e^{i\omega})d\omega \neq \sum_j |w_j|^2\overline{\phi_k(z_j)}\phi_l(z_j)$.

Hence $\boldsymbol{\Phi}^H\mathbf{W}^H\mathbf{W}\boldsymbol{\Phi} \neq \mathbf{I}$.

Condition number of $\mathbf{W}\boldsymbol{\Phi}$ can be very large.

Loss of correct digits in $\boldsymbol{\lambda}$ (! iterate on the nonlinear parameters $\boldsymbol{\alpha}$)

ORF recurrence

OPS: $z^{-1}\phi_k(z) \in \text{span}\{\phi_l : l = k+1, k, \ldots, 0\}$

ORF:(generic case = regular system: $\phi_k(\alpha_{k-1}) \neq 0$)

$$z^{-1}\phi_k(z) \in \text{span}\{[1 - \alpha_l/z]\phi_l(z) : l = k+1, \ldots, 0\} \quad (\alpha_0 = 0)$$

$$\phi_k(z) \in \text{span}\{[z - \alpha_l]\phi_l(z) : l = k+1, \ldots, 0\} \quad (\alpha_0 = 0)$$

$$\Phi(z) = \Phi(zI - \mathbf{A})\mathbf{H} + [0 \ldots 0 \ (z - \alpha_{n+1})\phi_{n+1}b_n]$$

$$\Phi = [\phi_0, \ldots, \phi_n], \quad \mathbf{A} = \text{diag}(\alpha_0, \alpha_1, \ldots, \alpha_n), \quad \mathbf{H} = \begin{bmatrix} \eta_{01} & \cdots & \eta_{0n} & \eta_{0,n+1} \\ \eta_{11} & \cdots & \eta_{1n} & \eta_{1,n+1} \\ & \ddots & \vdots & \vdots \\ & & \eta_{nn} & \eta_{n,n+1} \end{bmatrix}$$

$n = N$: ϕ_{N+1} does not exist ($\|\phi_{N+1}\|_w = 0$): choose $\Phi_{N+1}(z_k) = 0$, $k = 1, \ldots, N$.

ORF recurrence (continuous time)

In case of real line: 3-term recurrence for OPS. Similar for ORF (if $\alpha_k \in \mathbb{R}$).

$$\phi_k(z) \in \text{span}\{[z-\alpha_l]\phi_l(z) : l = k-1, k, k+1\}, \quad (\phi_{-1} = 0)$$

$$\phi_n(z) = b_{n-1}(z-\alpha_{n-1})\phi_{n-1} + a_n(z-\alpha_n)\phi_n + b_n(z-\alpha_{n+1})\phi_{n+1}.$$

$$\Phi = \Phi(zI - \mathbf{A})\mathbf{T} + [0 \ldots 0 \ (z-\alpha_{n+1})\phi_{n+1}b_n]$$

$$\Phi = [\phi_0, \ldots, \phi_n], \quad \mathbf{A} = \text{diag}(\alpha_0, \alpha_1, \ldots, \alpha_n), \quad \mathbf{T} = \begin{bmatrix} a_0 & b_0 & & & \\ b_0 & a_1 & b_1 & & \\ & b_1 & \ddots & \ddots & \\ & & \ddots & \ddots & b_{n-1} \\ & & & b_{n-1} & a_n \end{bmatrix}$$

If α_k are complex, $\mathbf{T}$ not tridiagonal but . . . (see later).

ORF recurrence (discrete time)

Forward recursion

$$\begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = e_n \frac{z-\alpha_{n-1}}{z-\alpha_n} \begin{bmatrix} 1 & \overline{L}_n \\ L_n & 1 \end{bmatrix} \begin{bmatrix} \zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix}$$

$$L_n = -\frac{\left\langle \phi_k, \frac{1-\overline{\alpha}_{n-1}z}{z-\alpha_n}\phi_{n-1} \right\rangle}{\left\langle \phi_k, \frac{z-\overline{\alpha}_{n-1}}{z-\alpha_n}\phi_{n-1}^* \right\rangle}, \quad e_n = \left(\frac{1-|\alpha_n|^2}{1-|\alpha_{n-1}|^2} \frac{1}{1-|L_n|^2} \right)^{1/2}$$

$$\zeta_k(z) = \frac{1-\overline{\alpha}_k z}{z-\alpha_k}, \quad B_n = \zeta_1 \cdots \zeta_n, \quad \phi_n^*(z) = B_n(z)\overline{\phi_n(1/\overline{z})}.$$

Backward recursion = Nevanlinna-Pick algorithm

Recurrence (discrete time)

Eliminate the $\phi_k^* \Rightarrow \phi_k(z) = \sum_{j=0}^{k+1} \eta_{jk}(z-\alpha_j)\phi_j(z)$

$$\Phi(z) = \Phi(z)(zI - \mathbf{A})\mathbf{H} + [0 \ldots 0 \ (z-\alpha_{n+1})\phi_{n+1}(z)b_n]$$

$\mathbf{H}$ is not unitary (see later)

Conclusion

continuous time: $\mathbf{H}$ almost tridiagonal.

discrete time: $\mathbf{H}$ almost unitary Hessenberg

Recursive state space

$\Phi_n(z)^T = (zI - A_n)^{-1}B_n$ with (A_n, B_n, C_n, D_n) a state space realization for the ORF

$$A_n = \begin{bmatrix} A_{n-1} & 0 \\ X_n & \alpha_n \end{bmatrix}, \quad B_n = \begin{bmatrix} B_{n-1} \\ b_n \end{bmatrix}$$

If $X_n = [x_{n1}, \ldots, x_{nn}]$, then

$$\rho_{nn} = 1, \quad \rho_{n,n-k-1} = e_{n-k}(-\bar{\alpha}_{n-k-1})(\rho_{n,n-k} + L_{n-k}\sigma_{n,n-k})$$
$$\sigma_{nn} = 0, \quad \sigma_{n,n-k-1} = e_{n-k}(\bar{L}_{n-k}\rho_{n,n-k} + \sigma_{n,n-k}),$$
$$x_{n,n-k} = e_{n-k}(1 - |\alpha_{n-k-1}|^2)(\rho_{n,n-k} + L_{n-k}\sigma_{n,n-k})$$

and

$$b_n = e_1(L_1\rho_{n,1} + \sigma_{n,1}) \langle 1, 1 \rangle_w^{-1/2}$$

Inner product evaluation

If data are available in $\mathbf{z} = \{z_j\}_{j=1}^N \subset \mathbb{T}$ with corresponding weights $\mathbf{w} = \{w_j\}_{j=1}^N$ then choose the discrete inner product.

For a continuous inner product, choose points on the circle, e.g. equidistant $z_j = \exp\{2ij\pi/N\}$, $j = 1, \ldots, N$ and evaluate weight $w_j = w(z_j)$ and use discrete inner product.

For example $w(z) = |U(z)|^2$, and z_j equidistant on $\mathbb{T}$, then given time domain data u_k, this can be computed very efficiently by FFT.

We can evaluate the moments c_l in $|U(z)|^2 = \sum_{l\in\mathbb{Z}} c_l z^l$ by convolution and use a Nevanlinna-Pick type algorithm on the PR function $\Omega(z) = c_0/2 + \sum_{l=1}^{\infty} c_l z^l$. (approximately)

The linear least squares problem

So we can compute

$$\begin{aligned} \boldsymbol{\Phi}(\boldsymbol{\alpha}) = \boldsymbol{\Phi}(\mathbf{z};\boldsymbol{\alpha}) &= [\phi_0(\mathbf{z}) \mid \phi_1(\mathbf{z}) \mid \cdots \mid \phi_n(\mathbf{z})] \in \mathbb{C}^{N\times n+1} \\ \mathbf{W} &= \operatorname{diag}(\mathbf{w}), \quad \boldsymbol{\lambda} = \{\lambda_k\}_{k=0}^n, \quad \mathbf{G} = G(\mathbf{z}) \end{aligned}$$

Setting $\hat{G}_n = \sum_{k=0}^n \lambda_k \phi_k$, solve in least squares sense

$$\mathbf{W}\boldsymbol{\Phi}(\boldsymbol{\alpha})\boldsymbol{\lambda}(\boldsymbol{\alpha}) = \mathbf{W}\mathbf{G} \equiv \mathbf{Q}(\boldsymbol{\alpha})\boldsymbol{\lambda}(\boldsymbol{\alpha}) = \mathbf{W}\mathbf{G}, \quad \mathbf{Q}(\boldsymbol{\alpha}) = \mathbf{W}\boldsymbol{\Phi}(\boldsymbol{\alpha})$$

by orthogonality however,

$$\boldsymbol{\lambda}(\boldsymbol{\alpha}) = [\mathbf{Q}(\boldsymbol{\alpha})^H \mathbf{Q}(\boldsymbol{\alpha})]^{-1} \mathbf{Q}(\boldsymbol{\alpha})^H \mathbf{W}\mathbf{G} = \mathbf{Q}(\boldsymbol{\alpha})^H \mathbf{W}\mathbf{G}$$

Nonlinear problem

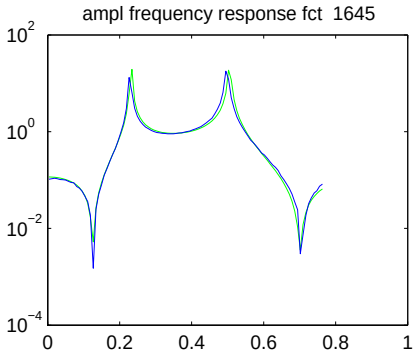
$$
\begin{aligned}
K(\boldsymbol{\alpha}) &= \sum_j |G(z_j) - \sum_{k=0}^{n} \lambda_k(\boldsymbol{\alpha})\phi_k(z_j;\boldsymbol{\alpha})|^2 w_j \Rightarrow \min_{\boldsymbol{\alpha}\subset\mathbb{D}} K(\boldsymbol{\alpha}) \\
&= \|(\mathbf{I} - \mathbf{\Phi}(\boldsymbol{\alpha})\mathbf{Q}(\boldsymbol{\alpha})^H\mathbf{W})\mathbf{G}\|_{\mathbf{W}}^2 = \|(\mathbf{I} - \mathbf{Q}(\boldsymbol{\alpha})\mathbf{Q}(\boldsymbol{\alpha})^H)\mathbf{W}\mathbf{G}\|^2
\end{aligned}
$$

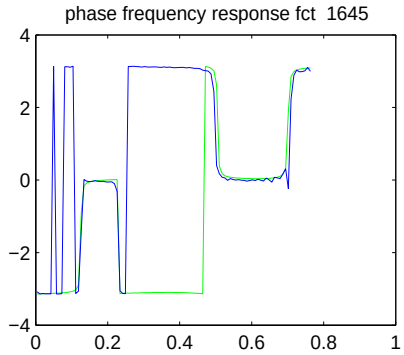
constraint: set $\alpha_j = r_j e^{i\omega_j}$ with $-1 \leq r_j \leq 1$, $j = 1, \ldots, n$.

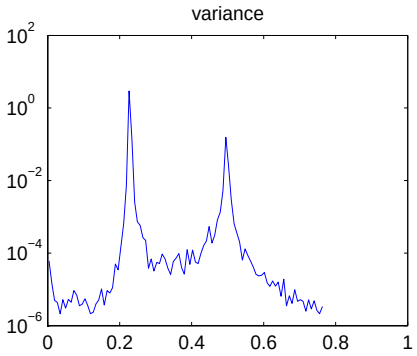
α_j can be forced to be real or complex conjugate. Therefore use parameters b_i, c_i where $(z - \alpha_i)(z - \overline{\alpha}_i) = z^2 + b_i z + c_i$ with appropriate constraints.

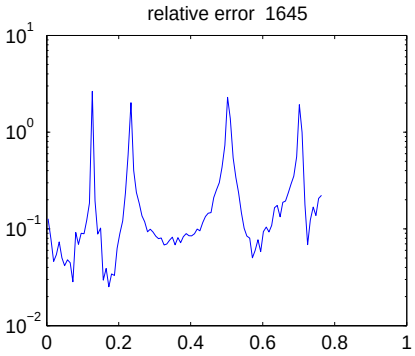
Robot arm

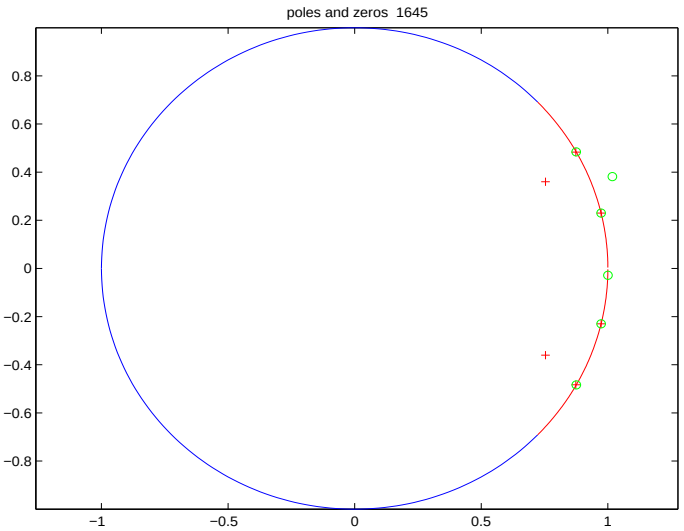
100 data points, 200 Hz, condensed in the beginning, order [6/6] discrete time





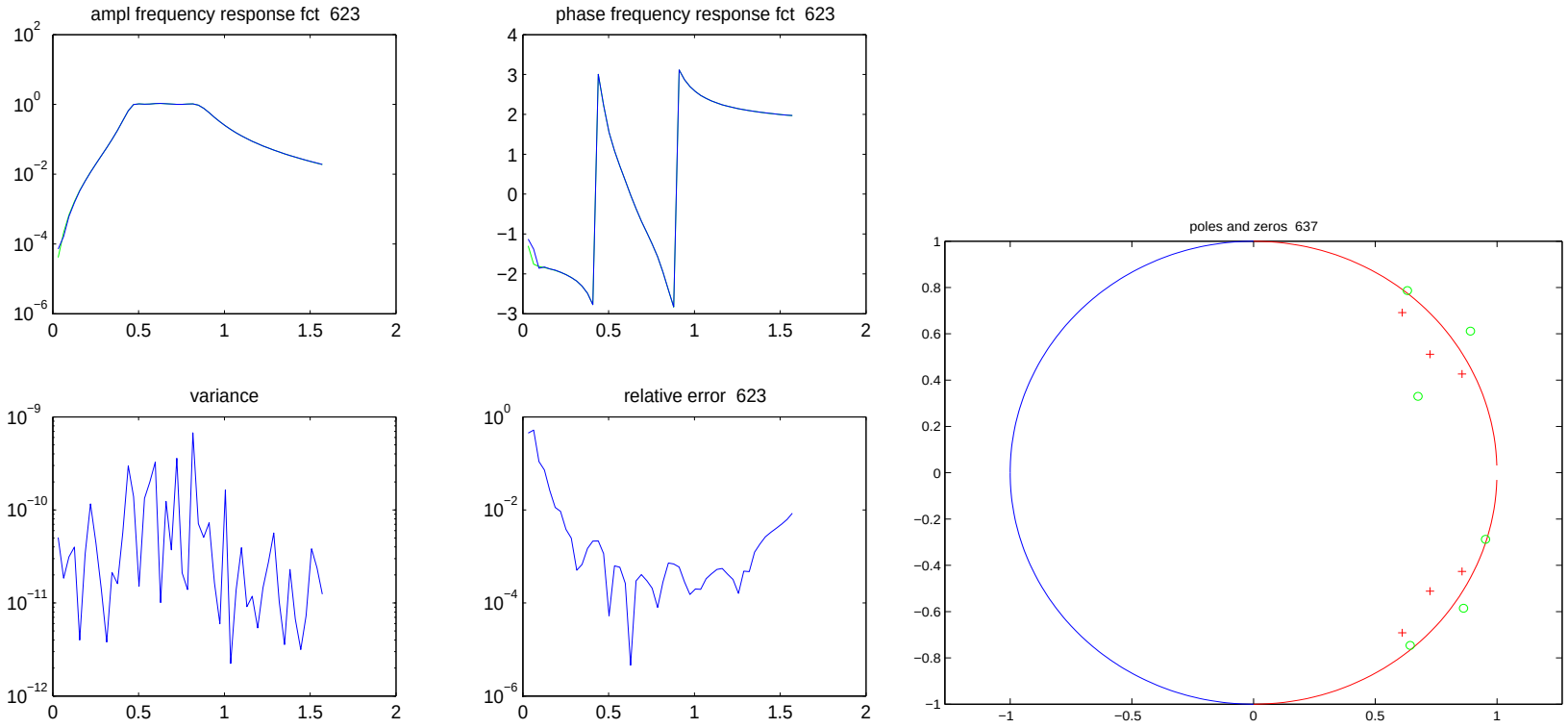






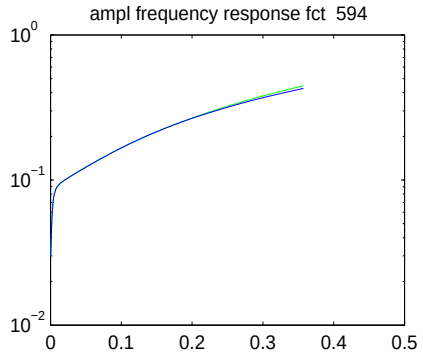
Bandpass filter

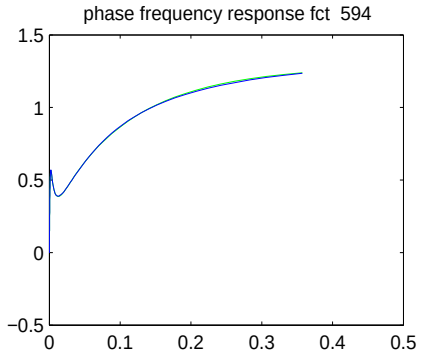
order [6/6], 50 data points, 20KHz, discrete time

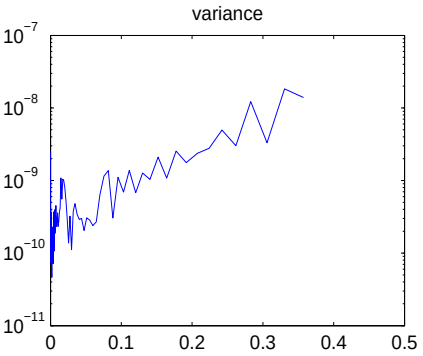


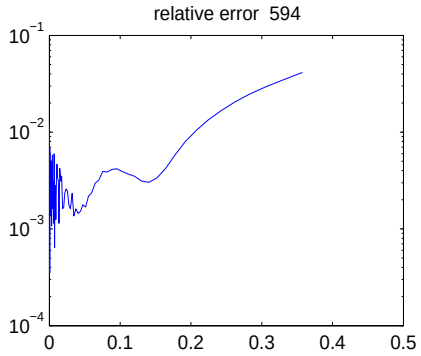
Electrical Machine

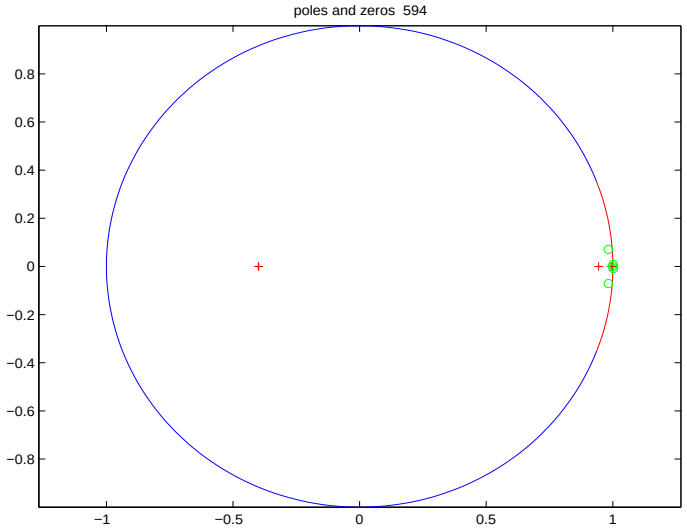
110 data points, 4000 Hz, condensed in the beginning, order [5/5], discrete time



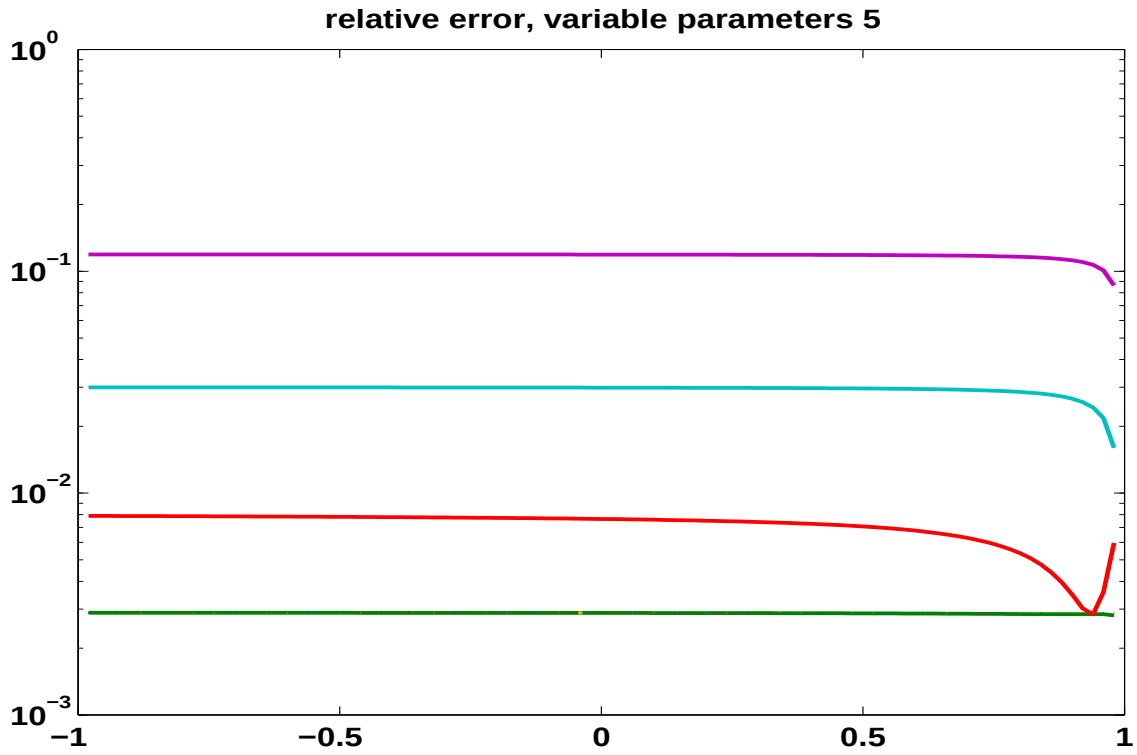








Sensitivity, Electrical Machine



poles: [-0.3999 -0.3989 0.9437 0.9935 0.9971]

Conclusion

- Rational approximation in weighted discrete least squares sense.
- Linear and nonlinear parameters. Solve for linear in terms of nonlinear. Use orthogonal basis to minimize the condition number.
- Either direct orthogonal rational basis or linearized vector polynomial as a combination of orthogonal block polynomials.

References ORF

[1] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. *Orthogonal rational functions*. Cambridge University Press, 1999.

[2] Th. de Hoog. Rational orthonormal bases and related transforms in linear system modeling PhD, T.U. Delft, 2001

[3] P. Van gucht, A. Bultheel. Using orthogonal rational functions for system identification, Report TW314, Dept. Computer Science, K.U.Leuven, September 2000.

[4] P. Van gucht and A. Bultheel. State-space realization and orthogonal rational functions. Report TW292, Dept. Computer Science, K.U.Leuven, September 1999.

[5] P. Van gucht, A. Bultheel. Matlab routines for system identification using orthogonal rational functions.
http://www.cs.kuleuven.ac.be/~nalag/research/software/ORF/ORFidentification.html.

ORF as inverse eigenvalue problem

$$\Phi(z) = \Phi(z)(zI - \mathbf{A})\mathbf{H} + [0 \ldots 0 \ (z - \alpha_{n+1})\phi_{n+1} b_n]$$

write for all z_j: (recall $\phi_{N+1}(\mathbf{z}) = 0$): $\mathbf{\Phi} = (\mathbf{Z\Phi} - \mathbf{\Phi A})\mathbf{H}$

multiply with $\mathbf{W}$ with $\mathbf{Q} = \mathbf{W\Phi}$: $\mathbf{Q} = (\mathbf{ZQ} - \mathbf{QA})\mathbf{H}$

assuming $\mathbf{S} = \mathbf{H}^{-1}$ exists: $\mathbf{Q}(\mathbf{S} + \mathbf{A}) = \mathbf{ZQ}$

multiply with $\mathbf{Q}^H$: $\mathbf{S} + \mathbf{A} = \mathbf{Q}^H\mathbf{ZQ}$

as before: $\mathbf{Q}^H\mathbf{w} = \mathbf{e}_1\eta_{00}$, $\mathbf{w} = [W_1, \ldots, W_N]^T$, $\eta_{00} = \|\mathbf{w}\|$

$$\mathbf{Q}^H \left[\begin{array}{c|c} \mathbf{w} & \mathbf{Z} \end{array}\right] \left[\begin{array}{c|c} 1 & \\ \hline & \mathbf{Q} \end{array}\right] = \left[\begin{array}{c|c} \|\mathbf{w}\| & \\ 0 & \mathbf{S} + \mathbf{A} \end{array}\right]$$

Semiseparable matrices

A semiseparable matrix is of the form

$\mathbf{S} = \mathrm{tril}(\mathbf{u}\mathbf{v}^H, 0) + \mathbf{R}$, e.g.

$$\mathbf{S} = \begin{bmatrix} u_1\bar{v}_1 & r_{12} & r_{13} & r_{14} \\ u_2\bar{v}_1 & u_2\bar{v}_2 & r_{23} & r_{24} \\ u_3\bar{v}_1 & u_3\bar{v}_2 & u_3\bar{v}_3 & r_{34} \\ u_4\bar{v}_1 & u_4\bar{v}_2 & u_4\bar{v}_3 & u_4\bar{v}_4 \end{bmatrix}$$

its inverse is upper Hessenberg $\mathbf{H}$.

In our case $\mathbf{S} = \mathbf{H}^{-1}$ and from the solution of the inverse eigenvalue problem it follows that for z_i all on $\mathbb{R}$ or all on $\mathbb{T}$, $\mathbf{R}$ is also "rank 1": $\mathbf{R} = \mathrm{triu}(\mathbf{s}\mathbf{t}^H, 1)$. Hence fast algorithms exist of $O(n^2)$.

Special cases

continuous time case: $z_k \in \mathbb{R}$

$\mathbf{S} + \mathbf{A} = \mathbf{Q}^H \mathbf{Z} \mathbf{Q}, \quad \Rightarrow \mathbf{S} + \mathbf{A} = (\mathbf{S} + \mathbf{A})^H$

$\mathbf{S}^H - \mathbf{S} = \mathbf{A} - \mathbf{A}^H = 2i\mathrm{Im}(\mathbf{A}),$

α's are real or come in pairs: $(a + ib, -a + ib)$.

Hence the off-diagonal part of $\mathbf{S}$ is hermitian $\Rightarrow \mathbf{R} = \mathrm{triu}(\mathbf{v}\mathbf{u}^H, 1)$

$\mathbf{S} = \mathrm{tril}(\mathbf{u}\mathbf{v}^H, 0) + \mathrm{triu}(\mathbf{v}\mathbf{u}^H, 1)$ and $\mathrm{Im}(\mathrm{diag}(\mathbf{S})) = -\mathrm{Im}(\mathbf{A})$.

Such an $\mathbf{H} = \mathbf{S}^{-1} \neq$ tridiagonal (only tridiagonal when α_k's real)

discrete time case: $z_k \in \mathbb{T}$

$\mathbf{Q}^H \mathbf{Z} \mathbf{Q}$ unitary $\Rightarrow \mathbf{S} + \mathbf{A}$ unitary.

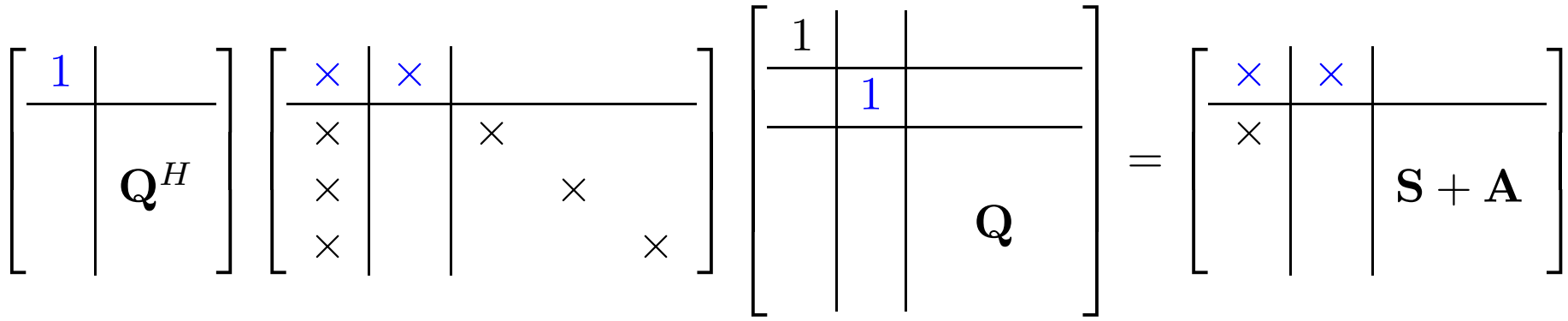
α's are real or come in complex conjugate pairs: $(a + bi, a - bi)$

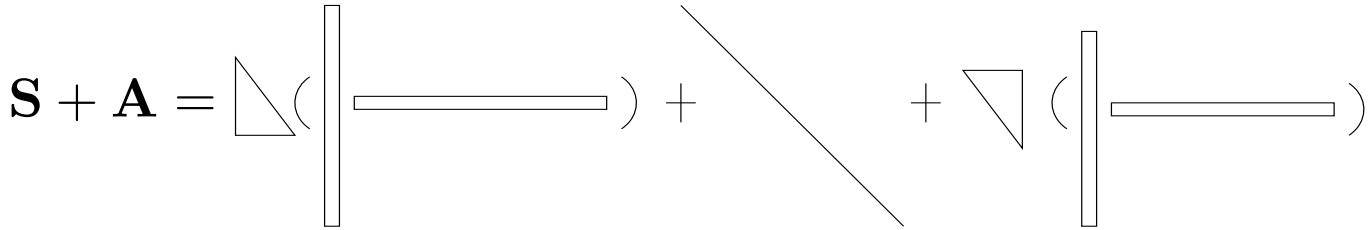
This implies some structure for $\mathbf{S}$, hence for $\mathbf{H}$

This structure is a consequence of the solution of the inverse eigenvalue problem. It is sufficient to know that the upper $\mathrm{triu}(\mathbf{S}^{-1}, 1)$ has rank 1.

Algorithm

Update can be done recursively in $O(n)$ operations $\Rightarrow$ total of $O(n^2)$ operations $=$ fast algorithm.
Possibly $O(nN)$ if degree n with N data points.





$$\left[\begin{array}{c|c||c} w_4 & z_4 & \\ \hline * & & \\ & & M \end{array}\right] = \left[\begin{array}{c|c||c} w_4 & z_4 & \\ \hline * & & \begin{array}{ccc} u_1\bar{v}_1+\alpha_0 & v_1\bar{u}_2 & v_1\bar{u}_3 \\ u_2\bar{v}_1 & u_2\bar{v}_2+\alpha_1 & v_2\bar{u}_3 \\ u_3\bar{v}_1 & u_3\bar{v}_2 & u_3\bar{v}_3+\alpha_2 \end{array} \end{array}\right]$$

$$Q = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}, \quad Q^H \begin{bmatrix} w_4 \\ * \end{bmatrix} = \begin{bmatrix} * \\ 0 \end{bmatrix}$$

updates: $\begin{bmatrix} \cdot \\ u_1 \\ u_2 \\ u_3 \end{bmatrix} \to \begin{bmatrix} u_1' \\ u_2'' \\ u_2 \\ u_3 \end{bmatrix} \to \begin{bmatrix} u_1' \\ u_2' \\ u_3'' \\ u_3 \end{bmatrix} \to \begin{bmatrix} u_1' \\ u_2' \\ u_3' \\ u_4'' \end{bmatrix} \to \begin{bmatrix} u_1' \\ u_2' \\ u_3' \\ u_4' \end{bmatrix}$

similar for the v's and α's.

Applying $\left[\begin{array}{c|c} Q^H & \\ \hline & I \end{array}\right] \left[\begin{array}{c|c} z_4 & \\ \hline & M \end{array}\right] \left[\begin{array}{c|c} Q & \\ \hline & I \end{array}\right]$ gives

$$\begin{bmatrix} a & \bar{e} & (-\bar{s}v_1)\bar{u}_2 & (-\bar{s}v_1)\bar{u}_3 \\ e & d & (\bar{c}v_1)\bar{u}_2 & (\bar{c}v_1)\bar{u}_3 \\ u_2(-s\bar{v}_1) & u_2(c\bar{v}_1) & u_2\bar{v}_2+\alpha_1 & v_2\bar{u}_3 \\ u_3(-s\bar{v}_1) & u_3(c\bar{v}_1) & u_3\bar{v}_2 & u_3\bar{v}_3+\alpha_2 \end{bmatrix} =$$

$$\left[\begin{array}{c|ccc} u_1'\bar{v}_1'+\alpha_0 & v_1'\bar{u}_2'' & v_1'\bar{u}_2 & v_1'\bar{u}_3 \\ \hline u_2''\bar{v}_1' & u_2''\bar{v}_2''+\alpha_1' & v_2''\bar{u}_2 & v_2''\bar{u}_3 \\ u_2\bar{v}_1' & u_2\bar{v}_2'' & u_2\bar{v}_2+\alpha_1 & v_2\bar{u}_3 \\ u_3\bar{v}_1' & u_3\bar{v}_2'' & u_3\bar{v}_2 & u_3\bar{v}_3+\alpha_2 \end{array}\right]$$

$v_1' = -\bar{s}v_1, \quad v_1'' = \bar{c}v_1$
$a = u_1'\bar{v}_1' + \alpha_0 \Rightarrow u_1' = (a-\alpha_0)/\bar{v}_1'$
$e = u_2''\bar{v}_1' \Rightarrow u_2'' = e/\bar{v}_1'$
$d = u_2''\bar{v}_2'' + \alpha_1' \Rightarrow \alpha_1' = d - u_2''\bar{v}_2''$

Is SS+D, but (2,2) diagonal element α_1' should be α_1.

Needs another unitary similarity transform on rows/columns 2 and 3.
We can use a matrix of the form

$$Q = \frac{1}{\sqrt{1+|t|^2}} \begin{bmatrix} 1 & \bar{t} \\ -t & 1 \end{bmatrix}$$

where $t = \frac{\alpha_1 - \alpha_1'}{\bar{v}_2 u_2'' - \bar{u}_2 v_2''}$

gives updates: $Q^H \begin{bmatrix} u_2'' \\ u_2 \end{bmatrix} = \begin{bmatrix} u_2' \\ u_3'' \end{bmatrix}, \quad [\bar{v}_2'' \;\; \bar{v}_2] Q = [\bar{v}_2' \;\; \bar{v}_3''], \;\; \alpha_2' = \alpha_1'.$

To be repeated until the last diagonal element is $u_4'' \bar{v}_4'' + \alpha_3'$.

To transform this in $u_4' \bar{v}_4' + \alpha_3$, use $u_4' = u_4''$ and set $\bar{v}_4' = (u_4'' \bar{v}_4'' + \alpha_3' - \alpha_3)/u_4'$.

First versions implemented in matlab (M. Van Barel). Fast and numerically stable.

Algorithms with complex poles yet to be designed.

We may expect that

• complex "conjugate" poles can be combined into 2×2 real blocks which means that we can construct real basis functions, which will guarantee that the poles and zeros will have symmetry.

• generalization to MIMO systems become possible.

References SS matrices

[1] D. Fasino, L. Gemignani. *Direct and inverse eigenvalue problems for diagonal-plus-semiseparable matrices* (2001) Submitted

[2] D. Fasino, L. Gemignani. *A Lanczos-type algorithm for the QR factorization of Cauchy-like matrices* (2002) Submitted

[3] M. Van Barel, D. Fasino, L. Gemignani, N. Mastronardi. *Orthogonal rational functions and diagonal-plus-semiseparable matrices*, (2002) Submitted