ELEC system identification workshop

Exercises

Ivan Markovsky

Outline

Introduction

Behavioral approach

Subspace methods

Optimization methods

SLRA package

Line fitting

problem: fit points $d_1, \ldots, d_N \in \mathbb{R}^2$ by a line

- 1. find condition for existence of a line (any line in \mathbb{R}^2) that passes through the points
- 2. how would you test the condition in MATLAB?
- 3. implement a method for exact line fitting

the points
$$d_i = (a_i, b_i), i = 1, ..., N$$
 lie on line

$$\updownarrow$$
there is $(R_1, R_2, R_3) \neq 0$, such that $R_1 a_i + R_2 b_i + R_3 = 0$, for $i = 1, ..., N$

$$\updownarrow$$
there is $(R_1, R_2, R_3) \neq 0$, such that
$$\begin{bmatrix} R_1 & R_2 & R_3 \end{bmatrix} \begin{bmatrix} a_1 & \cdots & a_N \\ b_1 & \cdots & b_N \\ 1 & \cdots & 1 \end{bmatrix} = 0$$

$$\updownarrow$$

$$\begin{matrix} \\ \\ \\ \\ \\ \\ \end{matrix}$$
rank
$$\begin{pmatrix} \begin{bmatrix} a_1 & \cdots & a_N \\ b_1 & \cdots & b_N \\ 1 & \cdots & 1 \end{bmatrix} \leq 2$$

given matrix d which columns are the data points

exact fitting condition:

```
N = size(d, 2)
dext = [d; ones(1, N)];

if (rank(dext) < 3)
    disp('exact fit exists')
else
    disp('exact fit does not exist')
end</pre>
```

given matrix d which columns are the data points

exact fitting method:

```
N = size(d, 2)
dext = [d; ones(1, N)];
r = null(dext')';
```

Note

$$\mathscr{B} = \{ d \mid Rd = 0 \}$$
 — linear static model

$$\mathscr{B} = \{ d \mid R \begin{bmatrix} d \\ 1 \end{bmatrix} = 0 \}$$
 — affine static model

in exact modeling

affine fitting



data centering + linear modeling

homework: is the same true in approximate modeling?

Conic section fitting

problem: fit points $d_1, \ldots, d_N \in \mathbb{R}^2$ by conic section

$$\mathscr{B}(S, u, v) = \{ d \in \mathbb{R}^2 \mid d^{\top}Sd + u^{\top}d + v = 0 \}$$

- 1. find condition for existence of an exact fit
- 2. propose numerical method for exact fitting
- 3. implement the method and test it on the data

$$d_1 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad d_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad d_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad d_4 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

the points
$$d_i = (a_i, b_i)$$
, $i = 1, ..., N$ lie on conic section \updownarrow

$$\exists \ S = S^{\top}, \ u, \ v, \ \text{at least one of them nonzero, such that}$$

 $d_i^{\top} S d_i + u^{\top} d_i + v = 0, \ \text{for} \ i = 1, \dots, N$

$$\updownarrow$$

there is
$$(s_{11}, s_{12}, s_{22}, u_1, u_2, v) \neq 0$$
, such that

there is
$$(s_{11}, s_{12}, s_{22}, u_1, u_2, v) \neq 0$$
, such that
$$\begin{bmatrix} a_1^2 & \cdots & a_N^2 \\ a_1b_1 & \cdots & a_Nb_N \\ a_1 & \cdots & a_N \\ b_1^2 & \cdots & b_N^2 \\ b_1 & \cdots & b_N \\ 1 & \cdots & 1 \end{bmatrix} = 0$$

Solution for part 1 (continued)

the points $d_i = (a_i, b_i)$, i = 1, ..., N lie on conic section

$$\operatorname{rank} \begin{pmatrix} \begin{bmatrix} a_1^2 & \cdots & a_N^2 \\ a_1b_1 & \cdots & a_Nb_N \\ a_1 & \cdots & a_N \\ b_1^2 & \cdots & b_N^2 \\ b_1 & \cdots & b_N \\ 1 & \cdots & 1 \end{bmatrix} \end{pmatrix} \leq 5$$

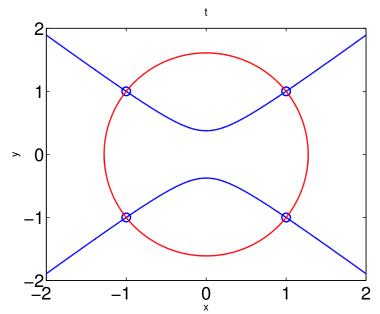
$$f = @(a, b) [a .^2; a .* b; a; b .^2; b; ones(size(a))];$$

Solution for part 2 and 3

```
finding exact models
```

```
R = \text{null}(f(d(1, :), d(2, :))')';
plotting model
function H = plot_model(th, f, ax, c)
H = ezplot(@(a, b) th * f(a, b), ax);
for h = H', set(h, 'color', c, 'linewidth', 2);
show results
plot(d(1, :), d(2, :), 'o', 'markersize', 12)
ax = 2 * axis;
for i = 1:size(R, 1)
 hold on, plot model (R(i, :), f, ax, c(i));
end
```

Solution for part 2 and 3 (continued)



Recursive sequence fitting

problem: fit
$$w = (w(1), ..., w(T))$$
 by model
$$\mathscr{B} = \{ w \mid R_0 w + R_1 \sigma w + \cdots + R_\ell \sigma^\ell w = 0 \}$$

- 1. find condition for existence of an exact fit first, with, and then, without knowledge of ℓ
- 2. propose numerical method for exact fitting find the smallest ℓ , for which exact model exists
- 3. implement the method and test it on the data

$$w = (w(1), ..., w(T)) \in \mathcal{B}$$

$$\mathcal{B} = \{ w \mid R_0 w + R_1 \sigma w + \dots + R_{\ell} \sigma^{\ell} w = 0 \}$$

$$\updownarrow$$

$$R_0 w(t) + R_1 w(t+1) + \dots + R_{\ell} w(t+\ell) = 0$$
for $t = 1, ..., T - \ell$

$$\updownarrow$$

$$rank (\mathcal{H}_{\ell+1}(w)) \leq \ell$$

$$\mathcal{H}_{\ell+1}(w) := \begin{bmatrix} w(1) & w(2) & \dots & w(T-\ell) \\ w(2) & w(3) & \dots & w(T-\ell+1) \\ \vdots & \vdots & & \vdots \\ w(\ell+1) & w(\ell+2) & \dots & w(T) \end{bmatrix}$$

relation at time t=1

$$R_0 w(1) + R_1 w(2) + \cdots + R_\ell w(\ell+1) = 0$$

in matrix form:

$$\begin{bmatrix} R_0 & R_1 & \cdots & R_\ell \end{bmatrix} \begin{bmatrix} w(1) \\ w(2) \\ \vdots \\ w(\ell+1) \end{bmatrix} = 0$$

relation at time t=2

$$R_0 w(2) + R_1 w(3) + \cdots + R_\ell w(\ell + 2) = 0$$

in matrix form:

$$\begin{bmatrix} R_0 & R_1 & \cdots & R_\ell \end{bmatrix} \begin{bmatrix} w(2) \\ w(3) \\ \vdots \\ w(\ell+2) \end{bmatrix} = 0$$

relation at time $t = T - \ell$

$$R_0 w(T-\ell) + R_1 w(T-\ell+1) + \cdots + R_\ell w(T) = 0$$

in matrix form:

$$egin{bmatrix} egin{bmatrix} R_0 & R_1 & \cdots & R_\ell \end{bmatrix} egin{bmatrix} w(T-\ell) \ w(T-\ell+1) \ w(T-\ell+2) \ dots \ w(T) \end{bmatrix} = 0$$

Solution for part 2 and 3

```
with \ell unknown, do the test for \ell = 1, 2, ...
algorithm
for ell = 1:ell_max
    if (rank(H(w, ell + 1)) == ell)
         break
    end
end
in the example, \ell = 3 and R = \begin{bmatrix} 1 & 1 & 1 & -1 \end{bmatrix}
```

Outline

Introduction

Behavioral approach

Subspace methods

Optimization methods

SLRA package

Checking whether a sequence is trajectory

- 1. given sequence w and polynomial R, propose method for checking numerically whether $w \in \mathcal{B} = \ker(R(\sigma))$
- 2. implement it in a function w_in_ker(w, r)
- 3. test it on the trajectory

$$w = (u_d, y_d) = ((0,1), (0,1), (0,1), (0,1))$$

and the system

$$\mathscr{B} = \ker(R(\sigma)), \qquad R(z) = \begin{bmatrix} 1 & -1 \end{bmatrix} + \begin{bmatrix} -1 & 1 \end{bmatrix} z$$

$$w \in \ker(R(\sigma))$$

$$\iff R(\sigma)w = 0$$

$$\iff R_0w(t) + R_1w(t+1) + \dots + R_\ell w(t+\ell) = 0$$
for $t = 1, \dots, T - \ell$

$$\iff \underbrace{\begin{bmatrix} R_0 & R_1 & \cdots & R_\ell \\ & R_0 & R_1 & \cdots & R_\ell \\ & & \ddots & \ddots & \ddots \\ & & & R_0 & R_1 & \cdots & R_\ell \end{bmatrix}}_{\mathcal{M}_T(R) \in \mathbb{R}^{p(T-\ell) \times qT}} \underbrace{\begin{bmatrix} w(1) \\ w(2) \\ \vdots \\ w(T) \end{bmatrix}}_{\mathsf{vec}(w)} = 0$$

numerical test: $||\mathcal{M}_T(R) \operatorname{vec}(w)|| < \varepsilon$ (with tolerance ε)

Another solution for part 1

$$w \in \ker(R(\sigma))$$

$$\iff \mathcal{M}_{T}(R)\operatorname{vec}(w) = 0$$

$$\iff R\mathscr{H}_{\ell+1}(w) = 0$$

where

$$\underbrace{\begin{bmatrix} R_0 & R_1 & \cdots & R_\ell \end{bmatrix}}_{R} \underbrace{\begin{bmatrix} w(1) & w(2) & \cdots & w(T-\ell) \\ w(2) & w(3) & \cdots \\ \vdots & \vdots & & \vdots \\ w(\ell+1) & w(\ell+2) & \cdots & w(T) \end{bmatrix}}_{\mathscr{H}_{\ell+1}(w) \in \mathbb{R}^{q(\ell+1) \times (T-\ell)}} = 0$$

numerical test: $||R\mathcal{H}_{\ell+1}(w)|| < \varepsilon$

```
function a = w_in_ker(w, r, ell)
a = norm(r * blkhank(w, ell + 1)) < 1e-8;
block-Hankel matrix \mathcal{H}_{i}(w) constructor
function H = blkhank(w, i, j)
[a, T] = size(w);
if T < q, w = w'; [q, T] = size(w); end
if nargin < 3, j = T - i + 1; end
H = zeros(i * q, j);
for ii = 1:i
  H(((ii - 1) * q + 1) : (ii * q), :) \dots
                = w(:, ii:(ii + i - 1));
end
```

Solution for part 2 (continued)

```
w = [0 0 0 0; 1 1 1 1];
r = [1 -1 -1 1]; ell = 1;
w_in_ker(w, r, 1)
```

homework

use image representation to check

$$w \stackrel{?}{\in} image(P(\sigma))$$
 (w_in_im)

use state space representation to check

$$w \stackrel{?}{\in} \mathscr{B}(A, B, C, D)$$
 (w_in_ss)

Transfer function → kernel representation

1. what model $\mathcal{B}_{tf}(H)$ is specified by transfer function

$$H(z) = \frac{q(z)}{p(z)} = \frac{q_0 + q_1 z^1 + \dots + q_{\ell} z^{\ell}}{p_0 + p_1 z^1 + \dots + p_{\ell} z^{\ell}}$$

2. find R, such that

$$\mathscr{B}_{tf}(H) = \ker(R)$$

3. write function tf2r converting H (tf object) to R and function r2tf converting R to H

Solution for part 1 and 2

the transfer function H represents model

$$\mathscr{B}_{tf}(H) = \{ w = \begin{bmatrix} u \\ y \end{bmatrix} \mid p(\sigma)y = q(\sigma)u \}$$

the corresponding kernel representation is

$$\underbrace{\left[q(\sigma) \quad -p(\sigma)\right]}_{R(\sigma)} \begin{bmatrix} u \\ y \end{bmatrix} = 0$$

note: $y = \mathcal{Z}^{-1}(H\mathcal{Z}(u))$ assumes zero initial conditions

homework: include initial conditions in $y = \mathcal{Z}^{-1}(H\mathcal{Z}(u))$

```
function r = tf2r(H)
[Q P] = tfdata(tf(H), 'v');
R = vec(fliplr([Q; -P]))';

function H = r2tf(R)
Q = fliplr(R(1:2:end));
P = - fliplr(R(2:2:end));
H = tf(Q, P, -1);
```

note: MATLAB uses descending order of coefficients

Initial conditions specification by trajectory

```
LSIM(SYS,U,T,X0) specifies the initial state vector X0 at time T(1)
(for state-space models only).
```

problem: given minimal $\mathscr{B} = \mathscr{B}(A, B, C, D) \in \mathscr{L}_{m,\ell}$

- 1. show that $\underbrace{\left(w(-\ell+1),\ldots,w(0)\right)}_{w_p}\in\mathscr{B}$ determines x(0)
- 2. explain how to use w_0 to "set" given x(0)
- 3. implement and test $w_p \leftrightarrow x(0)$ (wp2x0/x02wp)

$$y_{p} = \underbrace{\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\ell-1} \end{bmatrix}}_{\mathscr{O}} x(-\ell+1) + \underbrace{\begin{bmatrix} H(0) \\ H(1) & H(0) \\ \vdots & \ddots & \ddots \\ H(\ell-1) & \cdots & H(1) & H(0) \end{bmatrix}}_{\mathscr{T}} u_{p}$$

$$w_p = \begin{bmatrix} u_p \\ y_p \end{bmatrix} \in \mathscr{B} \quad \Longrightarrow \quad \text{solution } x(-\ell+1) \text{ exists}$$

minimal repr. \implies \mathscr{O} full rank \implies $x(-\ell+1)$ unique

$$x(0) = A^{\ell-1}x(-\ell+1) + \underbrace{\begin{bmatrix} A^{\ell-2}B & \cdots & BA^0 & 0 \end{bmatrix}}_{C} u_p$$

Solution for part 2 and 3

in order to set x(0), we include a prefix $w_p \wedge w_f$

Solution for part 2 and 3

```
construct \mathscr{C}
C = [sys.b zeros(ell, 1)];
for i = 1: (ell - 2)
  C = [sys.a * C(:, 1) C];
end
construct \mathscr{T}
h = impulse(sys, ell - 1);
T = \text{toeplitz}(h, [h(1) zeros(1, ell - 1)]);
```

Solution for part 2 and 3 (continued)

$$x(0) = \begin{bmatrix} \mathscr{C} - A^{\ell-1} \mathscr{O}^{+} \mathscr{T} & A^{\ell-1} \mathscr{O}^{+} \end{bmatrix} \begin{bmatrix} u_{\mathsf{p}} \\ y_{\mathsf{p}} \end{bmatrix}$$

```
function wp = x02wp(x0, sys)
ell = size(sys, 'order');
<<construct-C>>
<<construct-T>>
O = obsv(sys);
AO = sys.a ^ (ell -1) * pinv(O);
wp = pinv([C - AO * T , AO]) * x0;
wp = reshape(wp, ell, 2);
```

Solution (continued)

simulate data

```
n = 2; sys = drss(n);
T = 20; u = rand(T, 1); xini = rand(n, 1);
[y, t, x] = lsim(sys, u, [], xini); w = [u y];
test wp2x0 and x02wp
<<simulate-data>>
wp = w (end - n + 1:end, :); x0 = x (end, :)';
wp2x0(wp, sys) - x0
wp2x0(x02wp(x0, sys), sys) - x0
```

Outline

Introduction

Behavioral approach

Subspace methods

Optimization methods

SLRA package

Exact identification of a kernel representation

let $w \in \mathcal{B} \in \mathcal{L}^2_{1,\ell}$ (SISO system)

implement the method $w \mapsto R$ (slide 19)

test it on examples (use drss)

Solution

implementation

```
function r = w2r(w, ell)
r = null(blkhank(w, ell + 1)')';

test

<<simulate-data>>
sysh = r2tf(w2r(w, n));
norm(sys - sysh)
```

homework: generalize to the MIMO case

Impulse response estimation

let $w \in \mathcal{B} \in \mathcal{L}^2_{1,\ell}$ (SISO system) implement the method $w \mapsto H$ (slide 20–21) test it on examples (use drss)

implementation

```
function h = uy2h(u, y, ell, t)
L = ell + t;
H = [blkhank(u, L); blkhank(y, L)];
wini_uf = zeros(2 * ell + t, 1);
wini uf(ell + 1) = 1;
h = H(2 * ell + t + 1:end, :) * ...
      pinv(H(1:(2 * ell + t), :)) * wini_uf;
test
<<simulate-data>>
t = 5;
h = impulse(sys, t - 1);
hh = uy2h(u, y, n, t);
norm(h - hh)
```

38/52

Outline

Introduction

Behavioral approach

Subspace methods

Optimization methods

SLRA package

Misfit computation using image repr.

given

- data w = (w(1), ..., w(T)) and
- ▶ LTI system \mathscr{B} = image $(P(\sigma))$

derive method for computing

$$\mathsf{misfit}(w,\mathscr{B}) := \min_{\widehat{w} \in \mathscr{B}} \|w - \widehat{w}\|_2$$

i.e., find the orthogonal projection of w on \mathscr{B}

$$w \stackrel{?}{\in} image(P(\sigma))$$

$$\iff$$
 there is v , such that $w = P(\sigma)v$

$$\iff$$
 there is v , such that for $t=1,\ldots,T$
$$w(t)=P_0v(t)+P_1v(t+1)+\cdots+P_\ell v(t+\ell)$$

 \iff there is solution v of the system

$$\begin{bmatrix} w(1) \\ w(2) \\ \vdots \\ w(T) \end{bmatrix} = \underbrace{\begin{bmatrix} P_0 & P_1 & \cdots & P_\ell \\ & P_0 & P_1 & \cdots & P_\ell \\ & & \ddots & \ddots & \ddots \\ & & & P_0 & P_1 & \cdots & P_\ell \end{bmatrix}}_{\mathcal{M}_{T+\ell}(P)} \begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(T+\ell) \end{bmatrix}$$

we showed that

$$\widehat{w} \in \ker(R(\sigma)) \iff \widehat{w} = \mathscr{M}_T(P)v$$
, for some v

then the misfit computation problem

$$\mathsf{misfit}(\textit{w}, \mathscr{B}) := \min_{\widehat{\textit{w}} \in \mathscr{B}} \| \textit{w} - \widehat{\textit{w}} \|$$

becomes

minimize over
$$v \parallel w - \mathcal{M}_T(P)v \parallel$$

this is standard least-norm problem

projector on $\mathcal{B} = \text{image}(P)$

$$\Pi_{\mathsf{image}(P)} := \mathscr{M}_{\mathcal{T}}(P) \big(\mathscr{M}_{\mathcal{T}}^{\top}(P) \mathscr{M}_{\mathcal{T}}(P) \big)^{-1} \mathscr{M}_{\mathcal{T}}^{\top}(P)$$

misfit

$$\mathsf{misfit}(w,\mathscr{B}) := \sqrt{w^\top \big(I - \mathsf{\Pi}_{\mathsf{image}(P)}\big) w}$$

and optimal approximation

$$\widehat{w} = \Pi_{\mathrm{image}(P)} w$$

homework: misfit computation with $\mathscr{B} = \ker(R(\sigma))$

Misfit computation using I/S/O representation

given

- data w = (w(1), ..., w(T)) and
- ▶ LTI system $\mathscr{B} = \mathscr{B}(A, B, C, D)$

derive method for computing

$$\mathsf{misfit}(w,\mathscr{B}) := \min_{\widehat{w} \in \mathscr{B}} \|w - \widehat{w}\|_2$$

i.e., find the orthogonal projection of w on \mathscr{B}

$$w \stackrel{?}{\in} \mathscr{B}(A,B,C,D)$$

$$\mathscr{B}(A,B,C,D) = \{(u,y) \mid \sigma x = Ax + Bu, \ y = Cx + Du\}$$

$$(u_d,y_d) \in \mathscr{B}(A,B,C,D) \iff \exists x_{\text{ini}} \in \mathbb{R}^n, \text{ such that}$$

$$y = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{T-1} \end{bmatrix} \underbrace{X_{\text{ini}} + \begin{bmatrix} D \\ CB & D \\ CAB & CB & D \\ \vdots & \ddots & \ddots & \ddots \\ CA^{T-1}B & \cdots & CAB & CB & D \end{bmatrix}}_{U}$$

we showed that

$$\widehat{w} \in \mathscr{B}(A, B, C, D) \iff \widehat{y} = \mathscr{O}_{\mathcal{T}}(A, C)\widehat{x}_{\mathsf{ini}} + \mathscr{T}_{\mathcal{T}}(H)\widehat{u}$$

then the misfit computation problem

$$\min_{\widehat{x}_{\mathsf{ini}},\widehat{u}} \quad \left\| \begin{bmatrix} u_d \\ y_d \end{bmatrix} - \begin{bmatrix} 0 & I \\ \mathscr{O}_T(A,C) & \mathscr{T}_T(H) \end{bmatrix} \begin{bmatrix} \widehat{x}_{\mathsf{ini}} \\ \widehat{u} \end{bmatrix} \right\|$$

exploiting the structure in the problem

→ EIV Kalman filter

Latency computation using kernel repr.

given

- data w and
- LTI system $\mathscr{B}_{\mathsf{ext}} = \ker \big(R(\sigma) \big)$ $(w_{\mathsf{ext}} := \left[\begin{smallmatrix} \widehat{e} \\ \mathsf{w} \end{smallmatrix} \right])$

find an algorithm for computing

minimize over e $\|\widehat{e}\|$ subject to $(\widehat{e}, w) \in \mathscr{B}_{ext}$

partition
$$R = \begin{bmatrix} R_e & R_w \end{bmatrix}$$
 conformably with $w_{\text{ext}} = \begin{bmatrix} e \\ w \end{bmatrix}$

by analogy with the derivation on page 41, we have

$$\begin{bmatrix} e \\ w \end{bmatrix} \in \ker (R(\sigma)) \iff \begin{bmatrix} \mathscr{M}_T(R_e) & \mathscr{M}_T(R_w) \end{bmatrix} \begin{bmatrix} e \\ w \end{bmatrix} = 0$$

the latency computation problem is

$$\min_e \quad \|e\|_2 \quad \text{subject to} \quad \mathscr{M}_T(R_e)e = -\mathscr{M}_T(R_w)w$$

the solution is given by

$$\widehat{e} = -\underbrace{\left(\mathscr{M}_T(R_e)^\top \mathscr{M}_T(R_e)\right)^{-1} \mathscr{M}_T(R_e)^\top}_{\mathscr{M}_T(R_e)^+} \mathscr{M}_T(R_w) w$$

Outline

Introduction

Behavioral approach

Subspace methods

Optimization methods

SLRA package

Software

mosaic-Hankel low-rank approximation

```
http://slra.github.io/software.html
```

```
[sysh,info,wh] = ident(w, m, ell, opt)
```

- sysh I/S/O representation of the identified model
- opt.sys0 I/S/O repr. of initial approximation
- ▶ opt.wini initial conditions
- opt.exct exact variables
- ▶ info.Rh parameter R of kernel repr.
- ▶ info.M misfit

```
[M, wh, xini] = misfit(w, sysh, opt)
```

demo file

Variable permutation

verify that permutation of the variables doesn't change the optimal misfit

```
T = 100; n = 2; B0 = drss(n);
u = randn(T, 1); y = lsim(B0, u) + 0.001 * rand
[B1, info1] = ident([u y], 1, n); disp(info1.M)
    2.9736e-05
[B2, info2] = ident([y u], 1, n); disp(info2.M)
    2.9736e-05
disp(norm(B1 - inv(B2)))
    5.8438e-12
```

Output error identification

verify that the results of oe and ident coincide

```
T = 100; n = 2; B0 = drss(n);
u = randn(T, 1); y = lsim(B0, u) + 0.001 * random (B0, u) + 0.001 * r
opt = oeOptions('InitialCondition', 'estimate')
B1 = oe(iddata(y, u), [n + 1 n 0], opt);
B2 = ident([u \ v], 1, n, struct('exct', 1));
norm(B1 - B2) / norm(B1)
ans =
                           1.4760e-07
```