DYSCO course on low-rank approximation and its applications

Exercises

Ivan Markovsky

Vrije Universiteit Brussel

Vrije Universiteit Brussel

Outline

Line fitting

problem: give a condition on the data

$$\mathscr{D} = \{ d_1, \ldots, d_N \} \subset \mathbb{R}^2$$

that is equivalent to the condition that

the points $d_1, \ldots, d_N$ are on a line in $\mathbb{R}^2$

Conic section fitting

- conic section (with parameters $S = S^\top$, u, v)

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

1. give a condition on the data

$$\mathscr{D} = \{\, d_1, \ldots, d_N \,\} \subset \mathbb{R}^2$$

   that is equivalent to the condition that

   $d_1, \ldots, d_N$ are lying on a conic section

2. find a conic section fitting the points

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

Subspace clustering

- union of two lines (with parameters $R^1, R^2 \in \mathbb{R}^{1\times 2}$)

$$\mathscr{B}(R^1, R^2) = \{\, d \in \mathbb{R}^2 \mid (R^1 d)(R^2 d) = 0 \,\}$$

1. give a condition on the data

$$\mathscr{D} = \{\, d_1, \ldots, d_N \,\} \subset \mathbb{R}^2$$

   that is equivalent to the condition that

   $d_1, \ldots, d_N$ are lying on a union of two lines

2. find a union of two lines model fitting the points

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

Recursive sequences

- $w = \big(w(1), \dots, w(T)\big)$ is recursive of order ℓ if

$$R_0 w(t) + R_1 w(t+1) + \cdots + R_\ell w(t+\ell) = 0,$$
$$\text{for } t = 1, \dots, T-\ell \text{ and some } R_0, R_1, \dots, R_\ell \in \mathbb{R}$$

1. give a condition on w that is equivalent to

   w is a recursive of order ℓ

2. find the minimal recursive order of

$$(1, 2, 4, 7, 13, 24, 44, 81)$$

Polynomial common divisor

- the polynomials

$$p(z) = p_0 + p_1 z + \cdots + p_{\ell_p} z^{\ell_p}$$
$$q(z) = q_0 + q_1 z + \cdots + q_{\ell_q} z^{\ell_q}$$

  have a common divisor

$$c(z) = c_0 + c_1 z + \cdots + c_{\ell_c} z^{\ell_c}$$

  iff $p = ca$ and $q = cb$ for some polynomials a and b

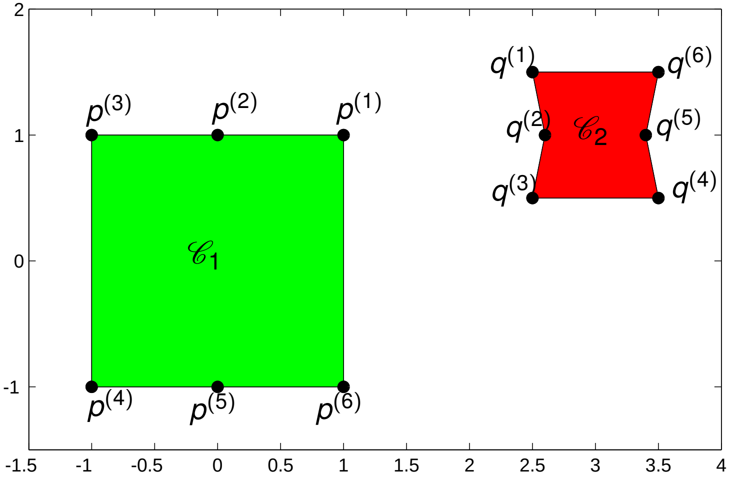
- give a condition on p, q that is equivalent to

  p, q have a common divisor of degree ℓ_c

Outline

Least squares contour alignment

given contours $\mathscr{C}_1, \mathscr{C}_2$, specified by matching points $p^{(i)} \leftrightarrow q^{(i)}$



Problem

- find a transformation (rotation + scaling + translation)

$$\mathscr{A}_{a,\theta,s}(p) = s \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} p + a$$

  that minimizes the LS distance between

$$\min_{a\in\mathbb{R}^2, \theta\in[0,2\pi), s\in\mathbb{R}_+} \sum_{i=1}^{N} \|p^{(i)} - \mathscr{A}_{a,\theta,s}(q^{(i)})\|_2^2$$

  $\mathscr{C}_1$ and $\mathscr{A}_{a,\theta,s}(\mathscr{C}_2)$

- apply the solution on the data in the example

Data in the example

```
p = [1      0     -1     -1      0      1
     1      1      1     -1     -1     -1];

q = [2.5     2.6     2.5     3.5     3.4     3.5
     1.5     1.0     0.5     0.5     1.0     1.5];
```

Hint

- use the change of variables

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = s \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \theta \\ s \end{bmatrix} = \begin{bmatrix} \sin^{-1}(b_2/\sqrt{b_1^2 + b_2^2}) \\ \sqrt{b_1^2 + b_2^2} \end{bmatrix}$$

- to obtain an equivalent problem

$$\min_{(a_1, a_2, b_1, b_2) \in \mathbb{R}^4} \left\| \begin{bmatrix} p_1^{(1)} \\ p_2^{(1)} \\ \vdots \\ p_1^{(N)} \\ p_2^{(N)} \end{bmatrix} - \begin{bmatrix} 1 & 0 & q_1^{(1)} & -q_2^{(1)} \\ 0 & 1 & q_2^{(1)} & q_1^{(1)} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & q_1^{(N)} & -q_2^{(N)} \\ 0 & 1 & q_2^{(N)} & q_1^{(N)} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{bmatrix} \right\|_2$$

Orthogonal Procrustes problem

- HW: alignment by reflection + scaling + translation
- rigid transformation =
  rotation + reflection + scaling + translation
- contour alignment by rigid transformation is related to the orthogonal Procrustes problem:
  - given $m \times n$ real matrices C_1 and C_2

    minimize over Q $\quad \|C_1 - QC_2\|_F \quad$ subject to $\quad Q^\top Q = I$
  - solution: $Q = UV^\top$, where $U\Sigma V^\top$ is the SVD of $C_1^\top C_2$

Outline

Check whether $w \overset{?}{\in} \mathscr{B}$

- $w = (u_d, y_d) = \big((0,1),(0,1),(0,1),(0,1)\big)$

```
w = [0 0 0 0; 1 1 1 1];
```

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

```
R = [1 -1 -1 1]; ell = 1;
```

Affine time-invariant system

- an LTI system $\mathscr{B} \in \mathscr{L}_{\mathrm{m},\ell}$ admits a kernel repr.

$$\mathscr{B} = \ker\big(R(\sigma)\big) := \{w \mid R(\sigma)w = 0\}$$

  for some $R(z) = R_0 z^0 + R_1 z^1 + \cdots + R_\ell z^\ell$

- show that

$$\mathscr{B}_c := \{\, w \mid R(\sigma)w = c \,\}$$

  is an affine time-invariant system, *i.e.*, $\mathscr{B}_c = \mathscr{B} + w_{\mathrm{p}}$ for LTI model $\mathscr{B} \in \mathscr{L}_{\mathrm{m},\ell}$ and trajectory w_{p}

- find $\mathscr{B}$ and w_{p}, s.t. $\mathscr{B} + w_{\mathrm{p}} = \{\, w \mid (0.5 + \sigma)w = 1 \,\}$

Transfer function $\mapsto$ kernel representation

- what model $\mathscr{B}_{\text{tf}}(H)$ is specified by a transfer function

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

- find R, such that

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

- write a function `tf2ker` converting H (`tf` object) to R

Specification of initial conditions

- initial conditions are explicitly specified in I/S/O repr.
- in MATLAB

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

- in transfer function representation initial conditions are often set to 0
- explain how to specify initial conditions in a representation free manner
- what is the link to $x_{\text{ini}} = x(1)$ in I/S/O repr?

Output matching

- given y_{f} and $\mathscr{B}$
- find u_{f}, such that $(u_{\mathrm{f}}, y_{\mathrm{f}}) \in \mathscr{B}$

Setup

- random SISO unstable system $\mathscr{B}$

```
clear all, n = 3;
Br = drss(n); [Qr, Pr] = tfdata(Br, 'v');
B = ss(tf(fliplr(Qr), fliplr(Pr), -1));
```

- reference output

```
T = 100; yf = ones(T, 1);
```

Outline

Misfit computation using image repr.

- given
  - data $w = \big(w(1), \dots, w(T)\big)$ and
  - LTI system $\mathscr{B} = \text{image}\,\big(P(\sigma)\big)$
- derive a method for computing

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

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

Misfit computation using I/S/O repr.

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

- derive a method for computing

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

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

Latency computation

- given
  - data w and
  - LTI system $\mathscr{B}_{\text{ext}} = \ker\big(R(\sigma)\big)$ $\qquad (w_{\text{ext}} := \left[\begin{smallmatrix} \widehat{e} \\ w \end{smallmatrix}\right])$
- find an algorithm for computing

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

- HW: latency computation with $\mathscr{B}_{\text{ext}} = \mathscr{B}(A, B, C, D)$
  (this is the ordinary Kalman filter)

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 * ranc
[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 * ranc
opt = oeOptions('InitialCondition', 'estimate')
B1 = oe(iddata(y, u), [n + 1 n 0], opt);
B2 = ident([u y], 1, n, struct('exct', 1));
norm(B1 - B2) / norm(B1)

ans =

   1.4760e-07
```

Outline

Identification without PE input

- given exact data $w = (u_d, y_d) \in \mathscr{B} \in \mathscr{L}^2_{1,\ell}$
- assuming controllability and PE of u_d of order $2\ell + 1$
- $\text{left}\ker\left(\mathscr{H}_{\ell+1}(w)\right)$ completely specifies $\mathscr{B}$
  (the model is identifiable from the data)
- what "goes wrong" when u_d is not PE of order $\ell + 1$?
- verify it numerically

Spurious poles

- given exact data $w \in \mathscr{B} \in \mathscr{L}^{\mathrm{p}}_{0,\ell}$ (autonomous system)
- assume that w is PE of maximal order, *i.e.*, $\operatorname{rank}\big(\mathscr{H}_{\ell+1}(w)\big) = \ell$
- the roots of $P \neq 0$, $P\mathscr{H}_{\ell+1}(w) = 0$ are the poles of $\mathscr{B}$
- consider now $\operatorname{left\,ker}\big(\mathscr{H}_{\ell+2}(w)\big)$ (ℓ is over-specified)
- what are the roots of $P \neq 0$, $P\mathscr{H}_{\ell+2}(w) = 0$?
- how to recover $\mathscr{B}$ from $\operatorname{left\,ker}\big(\mathscr{H}_{\ell+2}(w)\big)$?

Identification from short record(s)

- what is the minimum number $T_{\min}$ of sequential samples needed for identification of a model in $\mathscr{L}^{q}_{\mathtt{m},\ell}$ from $N = 1$ trajectory?
- what is the minimum number $T'_{\min}$ of sequential samples for identification of a model in $\mathscr{L}^{q}_{\mathtt{m},\ell}$ if $N' > 1$ trajectories with $T'_{\min}$ samples can be used?

MPUM for noisy data

- consider "noisy data"

$$w = \overline{w} + \widetilde{w}, \qquad \text{where} \quad \overline{w} \in \overline{\mathscr{B}} \in \mathscr{L}_{\mathrm{m},\ell}$$

  and $\widetilde{w}$ is white noise

- under the usual assumptions $\mathscr{B}_{\mathrm{mpum}}(\overline{w}) = \overline{\mathscr{B}}$
- what is $\mathscr{B}_{\mathrm{mpum}}(w)$ for the noisy data?
- suggest modifications of exact identification methods that make them suitable for approximation

Impulse response estimation

- implement the method $w \to H$
- compare it with `impulseest`