A brief motivation to study Subspace System Identification

Henri P. Gavin

Department of Civil & Environmental Engineering
Duke Univ.

2017-08-30
What is System Identification?
What is System Identification?

- the practice of **extracting models from measurements**
- a sub-set of parameter estimation specific to **dynamical systems**
- **modern** methods (1990’s to today) (e.g., ERA and SSID)
  - linear systems theory
  - **geometric** perspective of projections
    ...as opposed to least squares
  - leverages bases of linear vector spaces computed from **matrix decompositions** e.g., QR (LQ) and SVD
  - a **projection** of matrices of data onto space of models
  - characterized by **direct** matrix operations
    ...*no iterative guess-and-check*
  - multi-input, multi-output (**MIMO**) systems
  - the **order** of the system is contained in the data
    ...**high order** models (a dozen or more states)
Find a dynamic model from within the measurements

**Measurements**

- sequence of inputs $u(k)$

**Models**

- **MIMO FIR filters**
  
  $$y(k) = \sum_i H(i) u(k - i)$$

- **state-space models**
  
  $$x(k + 1) = A \ x(k) + B \ u(k) + w(k)$$
  $$y(k) = C \ x(k) + D \ u(k) + v(k)$$

- **transfer functions**
  
  $$\bar{y}(z) = \bar{H}(z) \ \bar{u}(z)$$

- **nonlinearities . . . restricted to . . .**
  
  $$u(k) = f(\hat{u}(k))$$
  $$\hat{y}(k) = g(y(k))$$
  
  . . . $x(k + 1)$ is linear in $x(k)$
dynamical systems

Subspace System ID

What is System Identification?

CEE 629

H.P. Gavin

2017-08-30
Geometric interpretation of ordinary least squares — fit a two-term polynomial to three data points

- model equation \( y_{model} = a_1 x + a_2 x^2 \),
- using values of the independent variable (error free, by assumption)
  \( x^T = [0.6, -0.7, 0.9] \) ... (and \( (x^2)^T = [0.36, 0.49, 0.81] \))
- and using values of the dependent variable (measured, with some error)
  \( y_{data}^T = [0.2, 0.3, 1.2] \)
- model basis ... \( y_{model} = [x, x^2]a = Xa \),
  where the columns of \( X \) form a basis of the model.
- what are the OLS estimates for model parameters, \( a_1, a_2 \)?
- ordinary least squares parameter estimates ... \( \hat{a} = [X^T X]^{-1} (X^T y_{data}) \)
- ordinary least squares model ...
  \( y_{model} = X \hat{a} = (X[X^T X]^{-1} X^T) y_{data} = P_X^\perp y_{data} \)
  where \( P_X^\perp \) is the orthogonal projection matrix onto the basis spanned by the columns of \( X \). (It's idempotent.)

- The orthogonal projection of the data onto the space of models minimizes the Euclidean length of \( y_{data} - y_{model} \).
- The shortest distance from a point to a plane.
Geometric interpretation of ordinary least squares — fit a two-term polynomial to three data points

-0.5 0 0.5 1 1.5
-0.5 0 0.5 1
-0.4 -0.2 0 0.2 0.4 0.6 0.8 1
-0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1

$a_1 = 0.223, a_2 = 1.030$
The outputs $y$ are linear in the states $x$ and the inputs $u$. $y = Cx + Du$

But the states are not measured!

This is the fundamental problem in LTI system ID.

How do we compute a state sequence corresponding to measured $u$ and $y$?
oblique projection of data onto space of models

\[
\begin{bmatrix}
U_f \\
Y_f
\end{bmatrix}
\]

\[
\begin{bmatrix}
U_p \\
Y_p
\end{bmatrix}
\]

\[
D_k U_f
\]

\[
\mathcal{P}_k X_f
\]

span \(U_f\)

span \(U_p\)

msmnt error

span \(U_p\)

Subspace System ID

What is System Identification?

CEE 629

H.P. Gavin

2017-08-30 7 / 44
Requirements

- persistent measured input . . .  
or an (assumed) persistent (unmeasured) noise disturbance
- the system responds persistently to persistent inputs
- the outputs do not feed back to inputs  
  (there’s a way around this)
What could possibly go wrong?

... these problems are cousins ...

- excessive measurement noise and/or process noise
- nearly linearly-dependent bases (ill-conditioning)
- over-parameterization
- non-linearity ... Hammerstein systems
What to do about it.

...we will start the term examining some of these methods ...

- pre-conditioning data (filtering, detrending) ... but be careful!
- $\ell_2$ regularization ... bias and covariance
- $\ell_1$ regularization (as a QP)
- rank reduction and total least squares
- *structured* rank reduction ... a lot of opportunity here, I think.
- ... nonlinearity ... ??
What good is an identified LTI system . . . in an arbitrary basis?

- What is $x$? . . . It’s a good question!
- If you don’t know what the identified state sequence $x$ physically corresponds to, can you do anything with the model? . . . e.g., control?
- Maybe prediction is enough.
- Maybe you know the mechanisms at play within the system . . . physics-based modeling.
- estimate (a few) physical parameters, e.g., $L, C, R, K, M, C$, from the identified LTI system dynamics, e.g., $p_j = -\zeta_j \omega_{nj} + i\omega_{dj}$
- usually a non-linear least squares problem
  Levenberg-Marquardt, etc.
  but with vastly reduced spaces of parameters and data!
System ID seems so essential! Why did it take until now to figure out how to do it right?
Lineage

- Jacopo Riccati (1676 - 1754)
- Alexandre-Theophile Vandermonde (1735 - 1796)
- Carl Friedrich Gauss (1777 - 1855)
- Camille Jordan (1838 - 1922)
- Hermann Hankel (1839 - 1873)
- Ferdinand Georg Frobenius (1849 - 1917)
- Andrey Markov (1856 - 1922)
- Aleksandr Lyapunov (1857 - 1918)
- Issai Schur (1875 - 1941)
- Otto Toeplitz (1881 - 1940)
- Norbert Wiener (1894 - 1964)
- Alston Scott Householder (1904 - 1993)
- Rudolf E. Kálmán (1930 - 2016)
- Gene H. Golub (1932 - 2007)
Subspace System Identification
Deterministic Subspace System Identification

discrete-time LTI in state-space

\[ x(k + 1) = A x(k) + B u(k) + w(k), \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^r \]
\[ y(k) = C x(k) + D u(k) + v(k), \quad y \in \mathbb{R}^m \]

disturbance - noise covariance

\[
E \left[ \begin{pmatrix} w(p) \\ v(p) \end{pmatrix}^T \begin{pmatrix} w(q)^T \\ v(q)^T \end{pmatrix} \right] = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \delta_{pq}
\]

apply to sequences

\[
\begin{bmatrix}
 x(i + 1) & \cdots & x(i + j) \\
 y(i) & \cdots & y(i + j - 1)
\end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix}
 x(i) & \cdots & x(i + j - 1) \\
 u(i) & \cdots & u(i + j - 1)
\end{bmatrix} + \begin{bmatrix} w(i) & \cdots & w(i + j - 1) \\
 v(i) & \cdots & v(i + j - 1)
\end{bmatrix}
\]
\[
\begin{bmatrix}
 X_{i+1,j} \\
 Y_{i,j}
\end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix}
 X_{i,j} \\
 U_{i,j}
\end{bmatrix} + \begin{bmatrix} W_{i,j} \\
 V_{i,j}
\end{bmatrix}
\]

So, a realization \( A, B, C, D \) may be obtained from input / output sequences with a direct matrix decomposition, 
*if* a state vector sequence \([x(i) \cdots x(i + j)]\) is known.
Subspace ID : notation and data matrices

Substitute state equation into output equation, recursively, to obtain

\[
\begin{bmatrix}
  y(i) & \cdots & y(i+j-1) \\
  y(i+1) & \cdots & y(i+j) \\
  \vdots & & \vdots \\
  y(i+k-1) & \cdots & y(i+j+k-2) \\
\end{bmatrix}
\begin{bmatrix}
  C \\
  CA \\
  \vdots \\
  CA^{k-1}
\end{bmatrix}
= 
\begin{bmatrix}
  x(i) & \cdots & x(i+j-1)
\end{bmatrix}
\begin{bmatrix}
  \vdots \\
  \vdots \\
\end{bmatrix}
\begin{bmatrix}
  \vdots \\
  \vdots \\
\end{bmatrix}
\begin{bmatrix}
  u(i) & \cdots & u(i+j-1) \\
  \vdots & & \vdots \\
  u(i+k-1) & \cdots & u(i+j+k-2)
\end{bmatrix}
\]

\[
Y_{i,j,k} = P_k X_{i,j} + D_k U_{i,j,k} + V_{i,j,k}
\]

\[
Y_p \equiv Y_{1,j,k} \quad U_p \equiv U_{1,j,k} \quad X_p \equiv X_{1,j} \quad \text{past I/O data and states}
\]

\[
Y_f \equiv Y_{k+1,j,k} \quad U_f \equiv U_{k+1,j,k} \quad X_f \equiv X_{k+1,j} \quad \text{future I/O data and states}
\]

\[
Y_p = P_k X_p + D_k U_p + V_p
\]

\[
Y_f = P_k X_f + D_k U_f + V_f
\]
Subspace ID: three necessary assumptions

1. The input is persistent
2. The states respond persistently
3. There is no feedback $u \xleftrightarrow{y}$
   (there are ways around this)

So,

$$\text{rank} \begin{bmatrix} U_p \\ Y_p \end{bmatrix} = kr + n$$

and any input / output sequence

$$[\tilde{u}(i) \cdots \tilde{u}(i + j - 1)] , [\tilde{y}(i) \cdots \tilde{y}(i + j - 1)]$$

is in the basis of

$$\begin{bmatrix} U_p \\ Y_p \end{bmatrix}$$
Subspace ID : main lemma

\[ \text{span} [X_f] = \text{span} \begin{bmatrix} U_p \\ Y_p \end{bmatrix} \cap \text{span} \begin{bmatrix} U_f \\ Y_f \end{bmatrix} \]

\(X_f\) lies in the span of \( \begin{bmatrix} U_p \\ Y_p \end{bmatrix} \)

\(\hat{Y}_f = \mathcal{P}_k X_f + \mathcal{D}_k U_f\)

\(\mathcal{P}_k X_f\) lies in \( \text{span} \begin{bmatrix} U_p \\ Y_p \end{bmatrix} \), and \( \text{span} \begin{bmatrix} U_f \\ Y_f \end{bmatrix} \),

\(\mathcal{P}_k X_f\) is the oblique projection of \( Y_f \) into \( \begin{bmatrix} U_p \\ Y_p \end{bmatrix} \) parallel to \( U_f \).
Subspace ID : computation

Compute the oblique projection via $LQ$ decomposition

\[
\begin{bmatrix}
U_f \\
U_p \\
Y_p \\
Y_f
\end{bmatrix} =
\begin{bmatrix}
L_{11} & 0 & 0 & Q_1 \\
L_{21} & L_{22} & 0 & Q_2 \\
L_{31} & L_{32} & L_{33} & Q_3
\end{bmatrix}
\]

\[P_k X_f = L_{32} L_{22}^{-1} \begin{bmatrix} U_p \\ Y_p \end{bmatrix}\]

Compute the state sequence via SVD

\[
P_k X_f = U \Sigma V^T = U \Sigma^{1/2} T \cdot T^{-1} \Sigma^{1/2} V^T
\]

\[P_k = U_n \Sigma_n^{1/2} T\]

\[X_f = T^{-1} \Sigma_n^{1/2} V_n^T = \begin{bmatrix} x(k+1) & \cdots & x(k+j) \end{bmatrix}\]

The state sequence is thus obtained, but in an arbitrary basis. . . . n4sid
primary reference texts for this course


Some of the other topics in this course
Total Least Squares

- solve the over-determined noisy system
  \[ y \approx Xa \]

- errors in \( y \) only . . . o.l.s.
  \[ y + \tilde{y} = Xa \]

- errors in \( y \) and \( X \) . . . t.l.s.
  \[ y + \tilde{y} = [X + \tilde{X}]a \]

- the t.l.s. problem
  \[ \min \left\| [\tilde{X} \tilde{y}] \right\|_F^2 \quad \text{such that} \quad y + \tilde{y} = [X + \tilde{X}]a \]

- Eckart-Young thm.
OLS, $\ell_1$, and $\ell_2$ regularization

\[ \hat{y}(x; a_1, a_2) = a_1 x + a_2 x^2 \quad y = \hat{y} + \eta \quad \sigma_\eta = 0.5 \]

OLS criterion

\[ \min_a \| y - Xa \|_2^2 \]

$\ell_1$ regularization criterion

\[ \min_a \| y - Xa \|_2^2 + \alpha \| a \|_1 \]

$\ell_2$ regularization criterion

\[ \min_a \| y - Xa \|_2^2 + \beta \| a \|_2 \]
Nonlinear Least Squares - Levenberg Marquardt

Example 5(a)

Example 5(b)
Singular Spectrum Analysis

- rank reduction of a Hankel matrix of a time series

\[ Y = \begin{bmatrix}
  y_1 & y_2 & \cdots & \cdots & y_{j-1} & y_j \\
  y_2 & y_3 & \cdots & \cdots & y_j & y_{j+1} \\
  \vdots & \vdots & & & \vdots & \vdots \\
  y_{k-1} & y_k & \cdots & \cdots & y_{k+j-3} & y_{k+j-2} \\
  y_k & y_{k+1} & \cdots & \cdots & y_{k+j-2} & y_{k+j-1}
\end{bmatrix} \]

- truncated SVD expansion

\[ Y_n = U_n \Sigma_n V_n^T \]

- reconstruct a Hankel form . . .
  average cross-diagonals (?) . . . or structured low-rank approximation

- extract a reduced basis from a very noisy measurement
Singular Spectrum Analysis

Subspace System ID  other topics  CEE 629  H.P. Gavin  2017-08-30 23 / 44
MIMO Wiener Filtering

- Estimate the (matrix-valued) Wiener coefficients (Markov parameters) $H(i)$
  \[ y(k) = \sum_{i} H(i) u(k - i) \]
- Filter a noisy measurement from broad-band input-output data
- Recover a reduced basis from a very noisy measurement
- Identify the impulse response of a MIMO system from noisy I/O data
Wiener filtering

model prediction and true signal

\( y + \text{noise} \)
\( \hat{y} \)

power spectral density

frequency, Hz

\( N = 8000 \quad \text{SNR} = 2.00 \quad K = 20 \)
MIMO Wiener Filter ID

- Model prediction (-) and true output (o)
- Unit impulse responses
- FFT
- True
- Identified

Subspace System ID
other topics
CEE 629
H.P. Gavin
2017-08-30 26 / 44
preliminaries needed for subspace identification

- linear vector spaces; decompositions; truncation; projections; ordinary least squares
- total least squares; SSA and Wiener filters
- nonlinear least squares; SQP; and $\ell_1$ regularization
- linear time invariant systems, Kalman filter; controllability and observability; Lyapunov equations; balanced realizations and model reduction;
Total Least Squares

- solve the over-determined noisy system
  \[ y \approx Xa \]

- errors in \( y \) only \( \ldots \) o.l.s.
  \[ y + \tilde{y} = Xa \]

- errors in \( y \) and \( X \) \( \ldots \) t.l.s.
  \[ y + \tilde{y} = [X + \tilde{X}] a \]

- the t.l.s. problem
  \[ \min \left\| [\tilde{X} \tilde{y}] \right\|_F^2 \quad \text{such that} \quad y + \tilde{y} = [X + \tilde{X}] a \]

- Eckart-Young thm.
$\hat{y}(x; a_1, a_2) = a_1 x + a_2 x^2 \quad y = \hat{y} + \eta \quad \sigma_\eta = 0.5$

**OLS criterion**
$$\min_a \|y - Xa\|^2_2$$

**$\ell_1$ regularization criterion**
$$\min_a \|y - Xa\|^2_2 + \alpha \|a\|_1$$

**$\ell_2$ regularization criterion**
$$\min_a \|y - Xa\|^2_2 + \beta \|a\|_2$$

![Graph showing OLS, $\ell_1$, and $\ell_2$ regularization](image-url)
Nonlinear Least Squares - Levenberg Marquardt

Example 5(a)

Example 5(b)
Singular Spectrum Analysis

- rank reduction of a Hankel matrix of a time series

\[ Y = \begin{bmatrix}
  y_1 & y_2 & \cdots & \cdots & y_{j-1} & y_j \\
  y_2 & y_3 & \cdots & \cdots & y_j & y_{j+1} \\
  \vdots & \vdots & & & \vdots & \vdots \\
  y_{k-1} & y_k & \cdots & \cdots & y_{k+j-3} & y_{k+j-2} \\
  y_k & y_{k+1} & \cdots & \cdots & y_{k+j-2} & y_{k+j-1}
\end{bmatrix} \]

- truncated SVD expansion

\[ Y_n = U_n \Sigma_n V_n^T \]

- reconstruct a Hankel form ... average cross-diagonals (??) ... or structured low-rank approximation

- extract a reduced basis from a very noisy measurement
Singular Spectrum Analysis

Subspace System ID

other topics

CEE 629

H.P. Gavin

2017-08-30 32 / 44
MIMO Wiener Filtering

- Estimate the (matrix-valued) Wiener coefficients (Markov parameters) $H(i)$

$$y(k) = \sum_i H(i) \ u(k - i)$$

- filter a noisy measurement from broad-band input-output data
- recover a reduced basis from a very noisy measurement
- identify the impulse response of a MIMO system from noisy I/O data
Wiener filtering

model prediction and true signal

\[ y + \text{noise} \]

\[ y \]

\[ \hat{y} \]

Power spectral density

Frequency, Hz

\[ N = 8000 \quad \text{SNR} = 2.00 \quad K = 20 \]

Subspace System ID

other topics

CEE 629

H.P. Gavin

2017-08-30 34 / 44
preliminaries needed for subspace identification

- linear vector spaces; decompositions; truncation; projections; ordinary least squares
- total least squares; SSA and Wiener filters
- nonlinear least squares; SQP; and $\ell_1$ regularization
- linear time invariant systems, Kalman filter; controllability and observability; Lyapunov equations; balanced realizations and model reduction;
Modeling Field Data
System Identification - Natural Periods

identified natural periods, s

peak foundation velocity, mm/s
System Identification - Damping Ratios

Identified damping ratios, %

Peak foundation velocity, mm/s

Subspace System ID
Modeling Field Data
CEE 629
H.P. Gavin
2017-08-30
Subspace System ID
Modeling Field Data
CEE 629
H.P. Gavin
2017-08-30 39 / 44
System Identification - modeling low-amplitude behavior

2011-11-18 23:11:14 UTC nMax=160, order=16, err=0.312

Subspace System ID
Modeling Field Data
CEE 629
H.P. Gavin
2017-08-30 40 / 44
System Identification - modeling high-amplitude behavior

2011-12-23 02:17:47 - M 6.00 nMax=160, order=16, err=0.424

Subspace System ID
Modeling Field Data
CEE 629
H.P. Gavin
2017-08-30
System Identification - Response Prediction

Record #132 Parameters Modeling Record #114 Output

Subspace System ID

Modeling Field Data

CEE 629

H.P. Gavin

2017-08-30

43 / 44
System Identification and Response Prediction

Subspace System ID
Modeling Field Data
CEE 629
H.P. Gavin
2017-08-30