System identification for the errors-in-variables problem

Torsten Söderström

Department of Information Technology
Uppsala University, Sweden
EIV in system identification

Contents

- Background and motivation
Contents

- Background and motivation
  - Problem formulation
Contents

- Background and motivation
  - Problem formulation
  - Line fitting
Contents

- Background and motivation
  - Problem formulation
  - Line fitting
- Identifiability
Contents

- Background and motivation
  - Problem formulation
  - Line fitting
- Identifiability
- Estimators
Contents

- Background and motivation
  - Problem formulation
  - Line fitting
- Identifiability
- Estimators
- Comparisons and conclusions
1.1 Problem formulation

\[ u_0(t) \rightarrow \text{SYSTEM} \rightarrow y_0(t) \rightarrow y(t) \]

\[ \tilde{u}(t), \tilde{y}(t) \] measurement noise.

Determine the system transfer function.
### 1.1 Problem formulation EIV, cont’d

Three cases:
- $v$ and $F$ unknown: True EIV situation
1.1 Problem formulation EIV, cont’d

Three cases:

- \( v \) and \( \mathcal{F} \) unknown: True EIV situation
- \( v \) under control, \( \mathcal{F} \) unknown (repeated exp.)
1.1 Problem formulation EIV, cont’d

Three cases:

- $v$ and $F$ unknown: True EIV situation
- $v$ under control, $F$ unknown (repeated exp.)
- $v$ new control variable, not an EIV problem
1.1 Motivations

- Understand the underlying relations (rather than make a good prediction from noisy data). [The ‘classical’ motivation in e.g. econometrics]

- Approximate a high-dimensional data vector by a small number of factors. [The standard motivation for factor analysis]

- Lack of enough information to classify the available signals into inputs and outputs; use a ‘symmetric’ system model. [Cf. the behavioral approach to modeling]
1.2 Line fitting

Assume that we have a set of points in the $x - y$ plane, that correspond to noisy measurements $(x_1, y_1), \ldots (x_n, y_n)$.

Model

\[
\begin{align*}
y_i &= y_{0i} + \tilde{y}_i, \\
x_i &= x_{0i} + \tilde{x}_i, & i = 1, \ldots, n. \\
y_{0i} &= a_0 x_{0i} + b_0,
\end{align*}
\]

The measurement errors $\{\tilde{y}_i\}$ and $\{\tilde{x}_i\}$: independent random variables of zero mean and variances $\lambda_y$ and $\lambda_x$, respectively.
1.2 Line fitting, cont’d: Data
1.2 Line fitting, cont’d

Least squares estimate

\[
(\hat{a}, \hat{b}) = \arg \min_{a,b} V_1(a, b)
\]

\[
V_1(a, b) = \sum (y_i - ax_i - b)^2
\]
1.2 Line fitting, cont’d

Data least squares estimate

\[ (\hat{a}, \hat{b}) = \arg \min_{a,b} V_2(a, b) \]

\[ V_2(a, b) = \sum (x_i - \frac{y_i}{a} + \frac{b}{a})^2 \]
1.2 Line fitting, cont’d

Total least squares estimate (orthogonal regression)

assume equal uncertainty in $x_i$ and $y_i$

Different options

What to do?
1.2 Line fitting, identifiability analysis

Use first and second order moments. Assume
\[ E(x_{0i}) = m, \quad \text{var}(x_{0i}) = \sigma^2. \]

5 equations 6 unknowns
\[ E(x) = m \quad a, \ b, \]
\[ E(y) = am + b \quad m, \]
\[ \text{var}(x) = \sigma^2 + \lambda_x \quad \sigma^2, \ \lambda_x, \ \lambda_y. \]
\[ \text{var}(y) = a^2 \sigma^2 + \lambda_y \]
\[ \text{cov}(x, y) = a \sigma^2 \]

No unique solution! Unknown uncertainties in both \( x_i \) and \( y_i \) make the problem difficult.
Line fitting, cont’d

Modified problem: Assume known noise variance ratio

\[ \frac{\lambda_y}{\lambda_x} = r \text{ (known)} \]

This problem is feasible:

- **Geometrically**: Scale so that \( \lambda_y = \lambda_u \). Use orthogonal regression.

- **Algebraically**: Use total least squares.

- **Statistically**: ML loss is well behaved and has a maximum for true parameter values of \( a \) and \( b \).
1.2 Line fitting, generalizations

- Multivariable model

\[ y = b + \sum a_j x_j \]
1.2 Line fitting, generalizations

- Multivariable model

$$y = b + \sum a_j x_j$$

- Time series model

$$y(t) = -a_1 y(t-1) - \cdots - a_{n_a} y(t-n_a) + b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b)$$
1.2 Line fitting, generalizations

- **Multivariable model**
  \[ y = b + \sum a_j x_j \]

- **Time series model**
  \[ y(t) = -a_1 y(t-1) - \cdots - a_{n_a} y(t-n_a) + b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b) \]

  - Regressor variables will be correlated
1.2 Line fitting, generalizations

- **Multivariable model**
  \[ y = b + \sum a_j x_j \]

- **Time series model**
  \[ y(t) = -a_1 y(t-1) - \cdots - a_{n_a} y(t-n_a) + b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b) \]
  - Regressor variables will be correlated
  - Uncertainties in different regressor variables are related!!
1.2 Consequences of input noise

The presence of input noise $\tilde{u}(t)$ will cause most identification methods to give biased estimates. This may be tolerable in closed-loop design:

- $G_c$ is usually insensitive to variations in $G_0$ (Examples: OP amplifiers, approximate linearization by feedback, etc)

- The robustness condition

$$\| \Delta G T \|_\infty < 1$$

gives tolerance for deviations of open loop transfer function.
Contents

- Background and motivation
- Identifiability
  - Problem formulation, basic assumptions
  - Nonparametric models
  - How to handle lack of identifiability
  - Parametric models
- Estimators
- Comparisons and conclusions
2.1 Problem formulation EIV cont’d

\[ G(q^{-1}) = \frac{B(q^{-1})}{A(q^{-1})}, \quad q^{-1}x(t) = x(t - 1). \]
2.1 Identifiability: Assumptions

The available signals are time-discrete

\[ u(t) = u_0(t) + \tilde{u}(t), \]
\[ y(t) = y_0(t) + \tilde{y}(t). \]

- The system is linear [causal] and asymptotically stable.
- \( \tilde{u}(t), \tilde{y}(t) \) are uncorrelated stationary processes, with zero means and spectra \( \phi_{\tilde{u}}(\omega) \) and \( \phi_{\tilde{y}}(\omega) \), respectively.
- \( u_0(t) \) is persistently exciting and uncorrelated with \( \tilde{u}(t) \) and \( \tilde{y}(t) \).
2.2 Identifiability nonparametric models

Use second order statistics of $\zeta(t) = (y(t), u(t))^T$:

$$
\Phi_{\zeta} = \begin{pmatrix} GG^* & G \\ G^* & 1 \end{pmatrix} \phi_{u0} + \begin{pmatrix} \phi_{\tilde{y}} & 0 \\ 0 & \phi_{\tilde{u}} \end{pmatrix}
= \begin{pmatrix} \hat{G}\hat{G}^* & \hat{G} \\ \hat{G}^* & 1 \end{pmatrix} \hat{\phi}_{u0} + \begin{pmatrix} \hat{\phi}_{\tilde{y}} & 0 \\ 0 & \hat{\phi}_{\tilde{u}} \end{pmatrix}.
$$

Note that for each frequency there are 3 equations with 4 unknowns. There is hence one degree of freedom (for each frequency) in the solution. Cf static cases!
2.3 How to handle the lack of identifiability?

At least four options

1. ‘Accept’ the status. Do not make further assumptions. Instead of looking for a unique estimate, deal with the whole set of possible estimates. [Set membership estimation]
2.3 How to handle the lack of identifiability?

At least four options

1. ‘Accept’ the status. Do not make further assumptions. Instead of looking for a unique estimate, deal with the whole set of possible estimates. [Set membership estimation]

2. Impose more detailed, parametric models of $u_0(t), \tilde{u}(t), \tilde{y}(t)$, say ARMA processes of specified orders.
2.3 Identifiability, cont’d

3. **Modify** at least one of the assumptions on Gaussian distributed data \((u_0, \tilde{u}, \tilde{y})\). Use higher order statistics to gain additional information. Deistler(1986), Tugnait(1992).
2.3 Identifiability, cont’d

4. **Use more than one experiment.** [Assume the user can control the signal $v(t)$]

- $\phi_{u_0}(\omega)$ differs between the different experiments,

or

- $u_0(t)$ is (well) correlated between experiments, but $\tilde{y}(t)$, $\tilde{u}(t)$ are uncorrelated between experiments.
2.4 Identifiability parametric models

Model $\tilde{u}(t)$, $\tilde{y}(t)$, $u_0(t)$ as ARMA processes (or white noise as a special case), and analyze identifiability.

Identifiability is often, but not always achieved. Example:

$$\hat{G}\hat{\phi}_u \equiv G\phi_u$$

Assume that $G = B/A$ does not contain any pair of zeros, reflected in the unit circle, in order to avoid ambiguities.
2.4 Identifiability parametric models, cont’d

Assume $\phi_{\tilde{u}}, \phi_{\tilde{y}}$ general, $G$ no reflected pair of zeros.

\[ \hat{G} \hat{\phi}_{u_0} \equiv G \phi_{u_0} \]

\[ \Rightarrow \hat{G} = \alpha G, \hat{\phi}_{u_0} = \phi_{u_0} / \alpha \]

$\alpha > 0$ (constant)

(cf static case)

Conditions

\[ \hat{\phi}_{\tilde{u}}(\omega) \geq 0, \hat{\phi}_{\tilde{y}}(\omega) \geq 0, \forall \omega \]

\[ \Rightarrow \alpha \in \text{a small interval around } \alpha = 1. \]
Contents

- Background and motivation
- Identifiability
- Example of estimators
  - Least squares (LS), Instrumental variables (IV), Bias-compensated LS (BCLS), The Frisch scheme, Total least squares (TLS)
  - Prediction error method (PEM) and maximum likelihood (ML) method
  - Accuracy aspects
- Comparisons and conclusions
3.1 Notations for parametric estimators

System description

\[ A(q^{-1})y_0(t) = B(q^{-1})u_0(t), \]

\[ A(q^{-1}) = 1 + a_1q^{-1} + \cdots + a_nq^{-n_a}, \]
\[ B(q^{-1}) = b_1q^{-1} + \cdots + b_nq^{-n_b}. \]

Parameter vector \( \theta \) and regressor vector \( \varphi(t) \):

\[ \theta = (a_1 \ldots a_{n_a} b_1 \ldots b_{n_b})^T, \]
\[ \varphi(t) = (-y(t - 1) \ldots - y(t - n_a)) \]
\[ u(t - 1) \ldots u(t - n_b))^T. \]
3.1 Notations for parametric estimators, cont’d

System description

\[ A(q^{-1})y(t) - B(q^{-1})u(t) = A(q^{-1})y_0(t) - B(q^{-1})u_0(t) \]
\[ + A(q^{-1})\tilde{y}(t) - B(q^{-1})\tilde{u}(t). \]
\[ \Delta = \varepsilon(t) \]

Hence, the system can be written as a linear regression

\[ y(t) = \varphi^T(t)\theta + \varepsilon(t). \]

Note that \( \varphi(t) \) and \( \varepsilon(t) \) are correlated.
3.1 Notations for parametric estimators, cont’d

Denote covariance matrices and their estimates as

\[
R_\varphi = E \left\{ \varphi(t)\varphi^T(t) \right\}, \quad \hat{R}_\varphi = \frac{1}{N} \sum_{t=1}^{N} \varphi(t)\varphi^T(t).
\]

Conventions:

- \( \theta_0 \) denotes the true parameter vector
- \( \hat{\theta} \) denotes its estimate.
- \( \varphi_0(t) \) denotes the noise-free part of the regressor vector.
- \( \tilde{\varphi}(t) \) denotes the noise-contribution to the regressor vector. (\( \varphi(t) = \varphi_0(t) + \tilde{\varphi}(t) \))
3.2 The least squares estimate is biased

Model

\[ y(t) = \varphi^T(t) \theta + \varepsilon(t). \]

Assume \( \tilde{u}(t) \) and \( \tilde{y}(t) \) are white.

The least squares (LS) estimate

\[
\hat{\theta}_{LS} = \hat{R}_\varphi^{-1} \hat{r}_\varphi y \to R_\varphi^{-1} r_\varphi y, \quad N \to \infty \\
= (R_{\varphi_0} + R_{\tilde{\varphi}})^{-1} r_{\varphi_0 y_0} = (R_{\varphi_0} + R_{\tilde{\varphi}})^{-1} R_{\varphi_0} \theta_0
\]

Bias due to \( R_{\tilde{\varphi}}. \)
3.3 Relation to factor models

Extend regressor and parameter vectors

\[ \overline{\varphi}(t) = \begin{pmatrix} -y(t) \\ \varphi(t) \end{pmatrix}, \quad \overline{\theta} = \begin{pmatrix} 1 \\ \theta \end{pmatrix} \]

System dynamics

\[ \overline{\varphi}_0^T \overline{\theta}_0 = 0 \]

\[ R_{\overline{\varphi}} = R_{\overline{\varphi}_0} + R_{\overline{\varphi}} \]

Note \( R_{\overline{\varphi}_0} \) is singular.
3.3 Relation to factor models

\[ R_{\tilde{\varphi}} = R_{\tilde{\varphi}0} + R_{\tilde{\varphi}} \]

\( \tilde{y}(t), \tilde{u}(t) \) white \( \Rightarrow \) \( R_{\tilde{\varphi}} \) diagonal
\( R_{\varphi_0} \) singular (co-rank = 1)

Factor model

- \( R_{\tilde{\varphi}_0} \) rank-deficient (‘low rank’) - due to latent variables
- \( R_{\tilde{\varphi}} \) diagonal

Generalizations: include dynamics!
3.4 Bias-compensating least squares, BCLS

Idea: Find additional equations for determining $\lambda_u$ and $\lambda_y$ and modify the normal equations to

\[
\begin{pmatrix}
\hat{R}_{\varphi} - \begin{pmatrix}
\hat{\lambda}_y I_{n_a} & 0 \\
0 & \hat{\lambda}_u I_{n_b}
\end{pmatrix}
\end{pmatrix}
\begin{pmatrix}
\hat{\theta}
\end{pmatrix} = \hat{r}_{\varphi y}
\]

Many possibilities exist.

Nonlinear equations with structure (often bilinear equations). Hence iterative schemes are necessary.
3.4 BCLS, cont’d

There are many variants

- $\tilde{u}(t), \tilde{y}(t)$ may be white or ARMA
- Different additional equations
- Different algorithms for solving the equations


Interrelations: Hong et al(2009)
3.4 BCLS, cont’d

Some possibilities for additional equations:
- Minimal LS loss
- LS estimates for an extended model
- Residual covariance function

Some possibilities for algorithms (note equations are often bilinear!)
- Relaxation algorithms (solve repeatedly linear equations)
- Variable projection algorithms → low dimensional optimization problem
3.5 Generalized IV estimate (GIVE)

Introduce a noise parameter vector $\rho$:

- $\tilde{y}$ white:
  $$\rho = \begin{pmatrix} \lambda_y & \lambda_u \end{pmatrix}^T$$

- $\tilde{y}$ correlated:
  $$\rho = \begin{pmatrix} r_{\tilde{y}}(0) & \ldots & r_{\tilde{y}}(n_a) & \lambda_u \end{pmatrix}^T$$
3.5 GIVE, cont’d

‘Biased-compensated IV equations’

\[
\left( \hat{R}_{z\varphi} - \tilde{R}_{z\tilde{\varphi}}(\rho) \right) \theta = \hat{r}_{zy} - r_{\tilde{z}\tilde{y}}(\rho)
\]

where

\[
z(t) = \begin{pmatrix}
y(t) \\
\vdots \\
y(t - n_a - p_y) \\
u(t - 1) \\
\vdots \\
u(t - n_b - p_u)
\end{pmatrix}
\]
3.5 GIVE, cont’d

Special cases

- Bias-eliminating least squares (Zheng, and others)
- Different Frisch schemes (Bologna group, and others)
- Extended compensated least squares (Ekman)
3.5 GIVE: Algorithmic aspects

Total parameter vector

\[ \vartheta = \left( \begin{array}{c} \theta^T \\ \rho^T \end{array} \right)^T \]

Biased-compensated IV equations

\[ f(\vartheta) \triangleq \frac{1}{N} \sum_{t=1}^{N} z(t, \theta) \varepsilon(t, \theta) - r_{z\varepsilon}(\theta, \rho) \approx 0 \]

Criterion

\[ V(\theta, \rho) = \| f(\vartheta) \|_{W(\theta)}^2 \]

\[ \hat{\theta}, \hat{\rho} = \arg \min_{\theta, \rho} V(\theta, \rho) \]
3.5 GIVE: Algorithmic aspects, cont’d

Assume $W$ does not depend on $\theta$: A variable projection algorithm is possible:

\[
\theta(\rho) \triangleq \arg \min_{\theta} V(\theta, \rho)
\]

\[
\overline{V}(\rho) = V(\theta(\rho)), \rho)
\]

\[
\hat{\rho} = \arg \min_{\rho} \overline{V}(\rho)
\]

\[
\hat{\theta} = \theta(\hat{\rho})
\]
3.6 Covariance matching

Define

\[ z_0(t) = \frac{1}{A(q^{-1})} u_0(t) \]

\[ a_0 = 1, \quad r_0(\tau) = E\{z_0(t + \tau)z_0(t)\} \]

Covariance functions of the measured data

\[ r_u(\tau) = \sum \sum a_i a_j r_0(\tau - i + j), \quad (\tau > 0) \]

\[ r_y(\tau) = \sum \sum b_i b_j r_0(\tau - i + j), \quad (\tau > 0) \]

\[ r_{yu}(\tau) = \sum \sum b_i a_j r_0(\tau - i + j) \]
3.6 Covariance matching, cont’d

**Step 1** Estimate the covariance vectors

$$
\mathbf{r}_y = \begin{pmatrix}
  r_y(1) \\
  \vdots \\
  r_y(p_y)
\end{pmatrix},
\quad
\mathbf{r}_u = \begin{pmatrix}
  r_u(1) \\
  \vdots \\
  r_u(p_u)
\end{pmatrix},
\quad
\mathbf{r}_{yu} = \begin{pmatrix}
  r_{yu}(p_1) \\
  \vdots \\
  r_{yu}(p_2)
\end{pmatrix}
$$

from data.

**Step 2** Determine $\theta$ and

$$
\mathbf{r}_z = \begin{pmatrix}
  r_0(0) \\
  \vdots \\
  r_0(k)
\end{pmatrix}
$$
3.6 Covariance matching, cont’d

Model

$$\mathbf{r} \triangleq \begin{pmatrix} r_y \\ r_u \\ r_{yu} \end{pmatrix} = \begin{pmatrix} F_y(\theta) \\ F_u(\theta) \\ F_{yu}(\theta) \end{pmatrix} \mathbf{r}_z \triangleq F(\theta)\mathbf{r}_z$$
3.6 Covariance matching, cont’d

Variable projection algorithm

\[
\{\hat{\theta}, \hat{r}_z\} = \arg \min_{\theta, r_z} ||\hat{r} - F(\theta)r_z||_W^2
\]

\[
\Rightarrow \hat{r}_z = (F^T(\theta)WF(\theta))^{-1}F^T(\theta)W\hat{r}
\]

\[
\hat{\theta} = \arg \min_{\theta} \left[ \hat{r}^T W\hat{r} - \hat{r}^T WF(\theta) \times (F^T(\theta)WF(\theta))^{-1}F^T(\theta)W\hat{r} \right]
\]
3.7 Maximum likelihood

Model noise and noise-free input as well as the system.

Example with $\tilde{y}(t), \tilde{u}(t)$ white, $u_0(t)$ ARMA process:

$$\zeta(t) = \begin{pmatrix} y(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} \frac{B(q^{-1})C(q^{-1})}{A(q^{-1})D(q^{-1})} & 1 & 0 \\ \frac{C(q^{-1})}{D(q^{-1})} & 0 & 1 \end{pmatrix} \begin{pmatrix} v(t) \\ \tilde{y}(t) \\ \tilde{u}(t) \end{pmatrix}.$$

$$\theta = \text{coeff} \ (A, B, C, D, \lambda_u, \lambda_y, \lambda_v)$$
3.7 ML, innovations form

\[ \Phi_\zeta(\omega) \equiv H(e^{i\omega}; \theta)Q(\theta)H^*(e^{i\omega}; \theta) \]

\( H \) monic; \( H(q^{-1}), H^{-1}(q^{-1}) \) asymptotically stable

Consequence: Optimal one-step prediction errors

\[ \varepsilon(t, \theta) = \zeta(t) - \hat{\zeta}(t|t-1; \theta) \]
\[ = H^{-1}(q^{-1}; \theta)\zeta(t) \]

(can be computed using a Kalman filter; spectral factorization using an algebraic Riccati equation)
3.7 PEM and ML, cont’d

Prediction errors

\[ \varepsilon(t, \theta) = \zeta(t) - \hat{\zeta}(t|t-1; \theta) \]
\[ = \mathbf{H}^{-1}(q^{-1}; \theta)\zeta(t). \]

PEM (prediction error method) estimate

\[ \hat{\theta}_N = \arg \min_{\theta} V_N(\theta). \]
\[ V_N(\theta) = \det \left( \frac{1}{N} \sum_{t=1}^{N} \varepsilon(t, \theta)\varepsilon^T(t, \theta) \right). \]
3.7 PEM and ML, cont’d

ML estimate (Gaussian data)

\[ V_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} \ell(\epsilon(t, \theta), \theta, t), \]

with

\[ \ell(\epsilon, \theta, t) = \frac{1}{2} \log \det Q(\theta) + \frac{1}{2} \epsilon^T(t, \theta) Q^{-1}(\theta) \epsilon(t, \theta), \]

\[ Q(\theta) = E \left\{ \epsilon(t, \theta) \epsilon^T(t, \theta) \right\}. \]
3.7 ML, cont’d

The ML estimate can alternatively be computed in the frequency domain, Pintelon-Schoukens (2005), Goodwin et al (2010), [some differences in how transient effects are handled]

The inherent spectral factorization is easier to carry out in the frequency domain.
3.7 ML, cont’d

General properties:

- (Very) high accuracy.
- The numerical optimization procedure is, in general, quite complex.
- Computations in the frequency domain is possible.
- The procedure may fail to give good results if only poor initial parameter estimates are available.
- A prediction error method leads to another criterion, and somewhat degraded accuracy.
3.8 How good can the estimates be?

The asymptotic distribution of $\hat{\theta}$ is known in many cases:

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \xrightarrow{\text{dist}} N(0, P),$$

The covariance matrix $P$ depends on

- the method (and the user parameters),
3.8 How good can the estimates be?

The asymptotic distribution of $\hat{\theta}$ is known in many cases:

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \overset{\text{dist}}{\rightarrow} \mathcal{N}(0, \mathcal{P}),$$

The covariance matrix $\mathcal{P}$ depends on

- the method (and the user parameters),
- the system,
3.8 How good can the estimates be?

The asymptotic distribution of $\hat{\theta}$ is known in many cases:

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \xrightarrow{\text{dist}} N(0, P),$$

The covariance matrix $P$ depends on

- the method (and the user parameters),
- the system,
- the dynamics for $u_0(t), \tilde{u}(t), \tilde{y}(t)$. 
3.8 How good can the estimates be?, cont’d

The Cramér-Rao lower bound $P_{\text{CRLB}}$

$$\text{cov} (\hat{\theta} - \theta_0) \geq P_{\text{CRLB}} = J^{-1} = P_{\text{ML}},$$

$$J = E \left\{ \left( \frac{\partial \log L(\theta)}{\partial \theta} \right)^T \left( \frac{\partial \log L(\theta)}{\partial \theta} \right) \right\},$$

where $L(\theta)$ is the likelihood function. The matrix $J$ is the Fisher information matrix.

Algorithms exist for computing $P_{\text{CRLB}}$, Söderström(2006).
Contents

- Background and motivation
- Identifiability
- Estimators
- Comparisons and conclusions
  - Example: Computational load
  - Example: Statistical accuracy
  - Some comparisons
  - Open issues
4.1 Some comparisons - computational load

A second order system;

![Graph showing comparisons between # parameters and # flops for PEM, Frisch, and IV.](image-url)
4.2 Some comparisons - performance

A second order system; $N \var(\hat{b}_1)$ vs. $\lambda_e^2$; other parameters behave similarly.
4.3 Some comparisons - performance

![Graph showing performance and computational complexity of different methods: IV, GIVE, GIVE\text{opt}, CM, CM\text{opt}, ML, ML (no input noise), SML, PEM.]}
4.4 Some open issues and future work

- Undermodeling
4.4 Some open issues and future work

- Undermodeling
- More of unification and relation between methods
4.4 Some open issues and future work

- Undermodeling
- More of unification and relation between methods
- Extensions to the multivariate case
4.4 Some open issues and future work

- Undermodeling
- More of unification and relation between methods
- Extensions to the multivariate case
- Modeling in continuous-time
4.4 Some open issues and future work

- Undermodeling
- More of unification and relation between methods
- Extensions to the multivariate case
- Modeling in continuous-time
- Model order determination
4.4 Some open issues and future work

- Undermodeling
- More of unification and relation between methods
- Extensions to the multivariate case
- Modeling in continuous-time
- Model order determination
- Recursive algorithms
4.4 Some open issues and future work

- Undermodeling
- More of unification and relation between methods
- Extensions to the multivariate case
- Modeling in continuous-time
- Model order determination
- Recursive algorithms
Thanks for listening!