Statistical Learning: Distribution-free Prediction and Confidence Intervals for ARX Systems with Instrumental Variables

Babolcsay, Barbara
Supervisor: Csáji, Balázs Csanád

Eötvös Loránd University
Faculty of Science

2024/05/30

1. Previously

2. This Semester

Problem Setting from Last Semester

- Consider a linear regression system:

$$\left.\begin{array}{c} \phi_1^T\theta^* + N_1 = Y_1 \\ \phi_2^T\theta^* + N_2 = Y_2 \\ \dots \\ \phi_n^T\theta^* + N_n = Y_n \end{array}\right\} \; \Phi^T\theta^* + N = Y$$

- The Least-Squares Estimate (LSE) of θ^*:

$$\hat{\theta} = (\Phi^T\Phi)^{-1}\Phi^T Y$$

- We aim to build non-asymptotic distribution-free confidence regions around the point estimate

Main Idea of SPS[1]

- Find $\hat{\theta}$, the root of the normal equation:

$$0 = \sum_{t=1}^{n} \phi_t (Y_t - \phi_t^T \theta) = \sum_{t=1}^{n} \phi_t \phi_t^T (\theta^* - \theta) + \sum_{t=1}^{n} \phi_t N_t = H_0(\theta)$$

- Perturb the signs of the prediction errors:

$$H_i(\theta) = \sum_{t=1}^{n} \alpha_{i,t} \phi_t \phi_t^T (\theta^* - \theta) + \sum_{t=1}^{n} \alpha_{i,t} \phi_t N_t$$

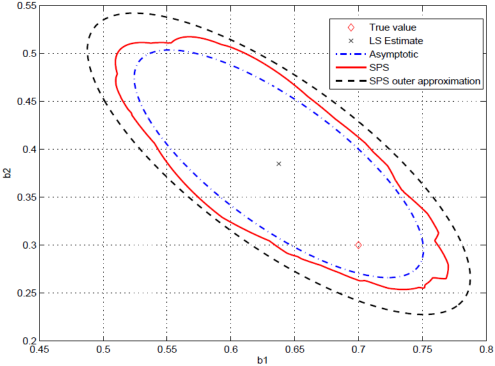
- Order them in some measure and define a subset over which θ is accepted

Remark

If θ is "close" to θ^*, the prediction error is "close" to the noise term so with symmetrically distributed noise the perturbed ones' distribution should remain the same.

Outer Approximation

- To get regions that are easier to calculate: ellipsoidal outer approximation
- It leads to $m-1$ convex minimization problem and the qth largest optimum gives the proper radius of the ellipsoid in the parameter space $(p = 1 - \frac{m}{q})$



[1]

Generalization of SPS

Assumptions we would like to relax:

exogenous regressors (independent from the noise terms)	$\Longleftrightarrow$	endogenous regressors
symmetrically distributed noise	$\Longleftrightarrow$	exchangeable noise
scalar valued data	$\Longleftrightarrow$	vector valued data

Scalar ARX Systems and Instrumental Variables

- Scalar ARX problem:

$$Y_t = \sum_{i=1}^{d_1} a_i^* Y_{t-i} + \sum_{i=1}^{d_2} b_i^* U_{t-i} + N_t \iff Y_t = \Phi_t^T \theta^* + N_t$$

- To handle the endogenous regressors - instrumental variables (ψ_t):
  1. must be correlated with the regressors
  2. cannot be correlated with the noise terms
- How to generate IVs:
  1. use only previous inputs ($\psi_t = [U_{t-1}, U_t]$)
  2. with least-squares estimation ($\psi_t = [\hat{Y}_t, U_t]$)
- The IV-estimate:

$$\hat{\theta}_{IV} = (\Psi^T \Phi)^{-1} \Psi^T Y$$

IV-SPS for ARX Systems [4]

endogenous regressors, symmetrically distributed noise, scalar valued data

- Find $\hat{\theta}_{IV}$, the root of the normal equation:

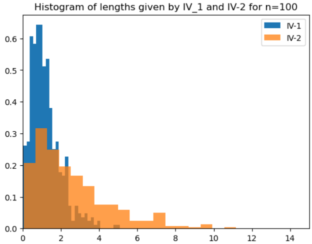
$$0 = \sum_{t=1}^{n} \psi_t (Y_t - \phi_t^T \theta) = \sum_{t=1}^{n} \psi_t \phi_t^T (\theta^* - \theta) + \sum_{t=1}^{n} \psi_t N_t = H_0(\theta)$$

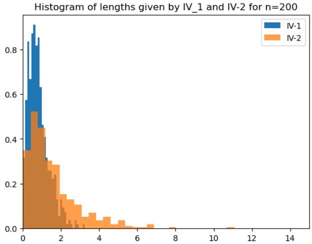
- Perturb the signs of the prediction errors:

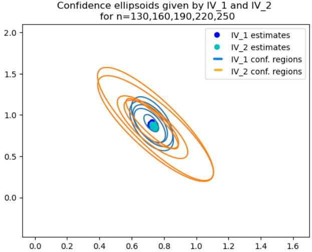
$$H_i(\theta) = \sum_{t=1}^{n} \alpha_{i,t} \psi_t \phi_t^T (\theta^* - \theta) + \sum_{t=1}^{n} \alpha_{i,t} \psi_t N_t$$

- Order them in some measure and define a subset over which θ is accepted

Performance of Different IV-Generating Methods







$$Y_t = a^* Y_{t-1} + b^* U_{t-1} + N_t$$
$$U_t = c^* U_{t-1} + V_t,$$

$$a^* = 0.75$$
$$b^* = 1$$
$$c^* = 0.75$$
$$p = 0.9$$

IV-PEM for ARX Systems[2]

endogenous regressors, exchangeable noise, scalar valued data

- Find $\hat{\theta}_{IV}$, the root of the normal equation:

$$0 = \sum_{t=1}^{n} \psi_t (Y_t - \phi_t^T \theta) = \sum_{t=1}^{n} \psi_t \phi_t^T (\theta^* - \theta) + \sum_{t=1}^{n} \psi_t N_t = H_0(\theta)$$

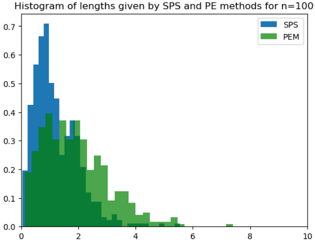
- Permute the signs of the prediction errors:

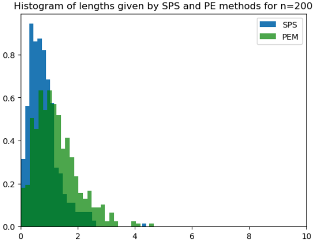
$$H_i(\theta) = P_i \sum_{t=1}^{n} \psi_t \phi_t^T (\theta^* - \theta) + P_i \sum_{t=1}^{n} \psi_t N_t,$$

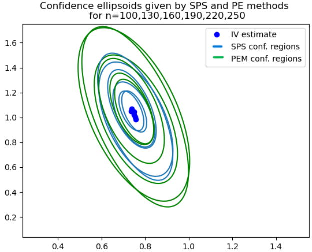
  where P_i is a random permutation matrix

- Order them in some measure and define a subset over which θ is accepted

Performance of SPS and PEM with i.i.d. Noise







$$Y_t = a^* Y_{t-1} + b^* U_{t-1} + N_t$$
$$U_t = c^* U_{t-1} + V_t,$$

$$a^* = 0.75$$
$$b^* = 1$$
$$c^* = 0.75$$
$$p = 0.9$$

Vector Valued ARX Systems and Instrumental Variables

- Vector valued ARX problem
  (Y_t, U_t, N_t are vectors, A^*, B^* are matrices):

$$Y_t = \sum_{i=1}^{d_1} A_i^* Y_{t-i} + \sum_{i=1}^{d_2} B_i^* U_{t-i} + N_t \iff Y = \Phi\Theta^* + N$$

- The IV-estimate:

$$\hat{\Theta}_{IV} = (\Psi^T \Phi)^{-1} \Psi^T Y,$$

  matrix valued!

MIV-SPS [3]

endogenous regressors, symmetrically distributed noise, vector valued data

- Find $\hat{\Theta}_{IV}$, the root of the normal equation:

$$0 = \Psi^T(Y - \Phi\Theta) = \Psi^T\Phi(\Theta^* - \Theta) + \Psi^T N = H_0(\Theta)$$

- Perturb the signs of the prediction errors:

$$H_i(\Theta) = \Psi^T W_i \Phi(\Theta^* - \Theta) + \Psi^T W_i N,$$

where $W_i = \mathrm{diag}(\alpha_i)$

- Order them in some measure (Frobenius-norm) and define a subset over which Θ is accepted

References

B. Cs. Csáji, M. C. Campi, and E. Weyer.
Sign-perturbed sums: a new system identification approach for constructing exact non-asymptotic confidence regions in linear regression models.
IEEE Trans. Signal Process., 63(1):169–181, 2015.

S. Kolumbán, I. Vajk, and J. Schoukens.
Perturbed datasets methods for hypothesis testing and structure of corresponding confidence sets.
Automatica J. IFAC, 51:326–331, 2015.

S. Szentpéteri and B. Cs. Csáji.
Non-asymptotic state-space identification of closed-loop stochastic linear systems using instrumental variables.
Systems Control Lett., 178:Paper No. 105565, 11, 2023.

V. Volpe, B. Cs. Csáji, A. Carè, E. Weyer, and M. C. Campi.
Sign-perturbed sums (SPS) with instrumental variables for the identification of ARX systems.
In *54th IEEE Conference on Decision and Control (CDC 2015)*, pages 2115–2120. IEEE, Osaka, 2015.

Thank You for Your Attention!