

An Alternative Paradigm for Probabilistic Uncertainty Bounding in Prediction Error Identification

Sippe Douma and Paul Van den Hof

CDC-ECC'05, December 12-15, 2005, Seville, Spain

The message

- In prediction error identification, quantified model uncertainty is usually based on

$$\text{pdf of estimator } z \rightarrow \hat{\theta}_N$$

- “Exact” probabilistic expressions on $\hat{\theta}_N - \theta_0$ are approximated by:
 - Employing asymptotic Gaussian distribution
 - Obtaining **P** through Taylor approximation (OE/BJ)
 - Replacing covariance matrix by estimate
- Probabilistic parameter uncertainty regions can be obtained **without specifying the estimator pdf**, with attractive results even for nonlinear estimators

Contents

- Example for illustration
- Uncertainty bounding in ARX models
- Extension to OE models
- Summary

Example

Data generating system: $y = \theta_0 x_1 + x_2$

$x_2 \in \mathcal{N}(0, 2)$; x_1 correlated with x_2 ; 1 data point (x_1, y)

Estimator: $\theta = y/x_1 = \theta_0 + x_2/x_1$

pdf of θ is very hard to analyze

However: $x_1(\theta - \theta_0) = x_2 \in \mathcal{N}(0, 2)$

After one experiment we have realizations: $x_1, \hat{\theta}$ of x_1, θ

Then $x_1(\hat{\theta} - \theta_0)$ is a realization of $x_2 \in \mathcal{N}(0, 2)$.

Based on test statistic $x_1(\hat{\theta} - \tilde{\theta})$ we select all $\tilde{\theta}$ that are within the α -probability level of x_2 :

$$\theta_0 \in \left\{ \theta \mid (\hat{\theta} - \theta) x_1^2 (\hat{\theta} - \theta) \leq 2c_\chi(\alpha, 1) \right\} \quad \text{w.p. } \alpha$$

Probabilistic parameter bounding without pdf of estimator

Employ statistical properties of random variable

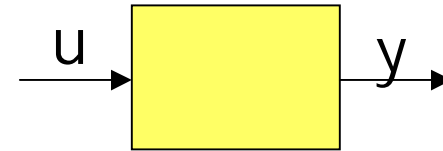
$$x_1(\theta - \theta_0) = x_2 \in \mathcal{N}(0, 2)$$

rather than those of

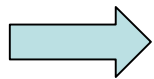
$$\theta - \theta_0$$

Benefit = simplicity of expression / analysis

Issue in i/o dynamical systems:



Question whether to consider measured **input u** as deterministic or stochastic in variance analysis of estimator



Prior or posterior variance

However, the issue goes beyond the role of **u**, and also incorporates the role of **y**.

Uncertainty bounding in ARX models

$$\hat{y}(t|t-1; \theta) = \varphi^T(t)\theta$$

With

$$\Phi = \begin{pmatrix} \varphi^T(1) \\ \vdots \\ \varphi^T(N) \end{pmatrix} \text{ and } \mathbf{y} = [y(1) \cdots y(N)]^T$$

$$\hat{\theta}_N = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{y}$$

If $\mathcal{S} \in \mathcal{M}$: $\mathbf{y} = \Phi \theta_0 + \mathbf{e}$

$$\hat{\theta}_N - \theta_0 = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{e}$$

ARX modelling

$$\text{If } \mathcal{S} \in \mathcal{M}: \hat{\theta}_N - \theta_0 = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{e}$$

Classical approach:

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \rightarrow \mathcal{N}(0, P_{arx})$$

$$P_{arx} = (\mathbb{E}[\frac{1}{N} \Phi^T \Phi])^{-1} \cdot \sigma_e^2$$

$$\theta_0 \in \left\{ \theta \mid (\hat{\theta}_N - \theta) P_{arx}^{-1} (\hat{\theta}_N - \theta) \leq c_\chi(\alpha, n)/N \right\} \quad \text{w.p. } \alpha$$

Requires:

- (asymptotic) normality of $(\Phi^T \Phi)^{-1} \Phi^T \mathbf{e}$
- Replacement of P_{arx} by an estimate \hat{P}_{arx}

ARX modelling

$$\text{If } \mathcal{S} \in \mathcal{M}: \hat{\theta}_N - \theta_0 = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{e}$$

Alternative:

$$\text{Consider } \beta := \frac{1}{\sqrt{N}} \Phi^T \Phi (\hat{\theta}_N - \theta_0) = \frac{1}{\sqrt{N}} \Phi^T \mathbf{e}.$$

$$\rightarrow \mathcal{N}(0, Q) \quad Q = \mathbb{E}\left[\frac{1}{N} \Phi^T \Phi\right] \cdot \sigma_e^2$$

ARX modelling

Consider $\beta := \frac{1}{\sqrt{N}} \Phi^T \Phi (\hat{\theta}_N - \theta_0) = \frac{1}{\sqrt{N}} \Phi^T \mathbf{e}$.

$$\rightarrow \mathcal{N}(0, Q) \quad Q = \mathbb{E}\left[\frac{1}{N} \Phi^T \Phi\right] \cdot \sigma_e^2$$

Result

$$\theta_0 \in \left\{ \theta \mid (\hat{\theta}_N - \theta)^T P_{arx,n}^{-1} (\hat{\theta}_N - \theta) \leq \frac{c_\chi(\alpha, n)}{N} \right\} \text{ w.p. } \alpha$$

$$\text{with } P_{arx,n} = \left(\frac{1}{N} \Phi^T \Phi\right)^{-1} Q \left(\frac{1}{N} \Phi^T \Phi\right)^{-1}$$

Requires:

- (asymptotic) normality of $\Phi^T \mathbf{e}$
- Replacement of Q by an estimate

ARX modelling

Implementable scheme:

Replace $Q = \mathbb{E}[\frac{1}{N}\Phi^T\Phi] \cdot \sigma_e^2$ by $\frac{1}{N}\Phi^T\Phi\hat{\sigma}_e^2$

Then $\hat{P}_{arx,n} = (\frac{1}{N}\Phi^T\Phi)^{-1}\hat{\sigma}_e^2$

Same expression as used in the classical situation

Result is related to likelihood method, determined by

$$\{\theta \mid V_N(\theta) - V_N(\hat{\theta}_N) \leq c_\chi(\alpha, n)/N\}$$

ARX modelling

Implementable scheme:

Replace $Q = \mathbb{E}[\frac{1}{N}\Phi^T\Phi] \cdot \sigma_e^2$ by $\frac{1}{N}\Phi^T\Phi\hat{\sigma}_e^2$

Then $\hat{P}_{arx,n} = (\frac{1}{N}\Phi^T\Phi)^{-1}\hat{\sigma}_e^2$

Same expression as used in the classical situation

Conclusion

Classical results with P_{arx} approximated by sample estimates, has stronger theoretical support than often considered.

Benefit: relaxation of conditions for normality

Simulation example:

First order ARX system:

$$y(t) = \frac{0.5}{1 + 0.9q^{-1}}u(t) + \frac{1}{1 + 0.9q^{-1}}e(t)$$

identified with 1st order ARX model.

Compare empirical distributions of

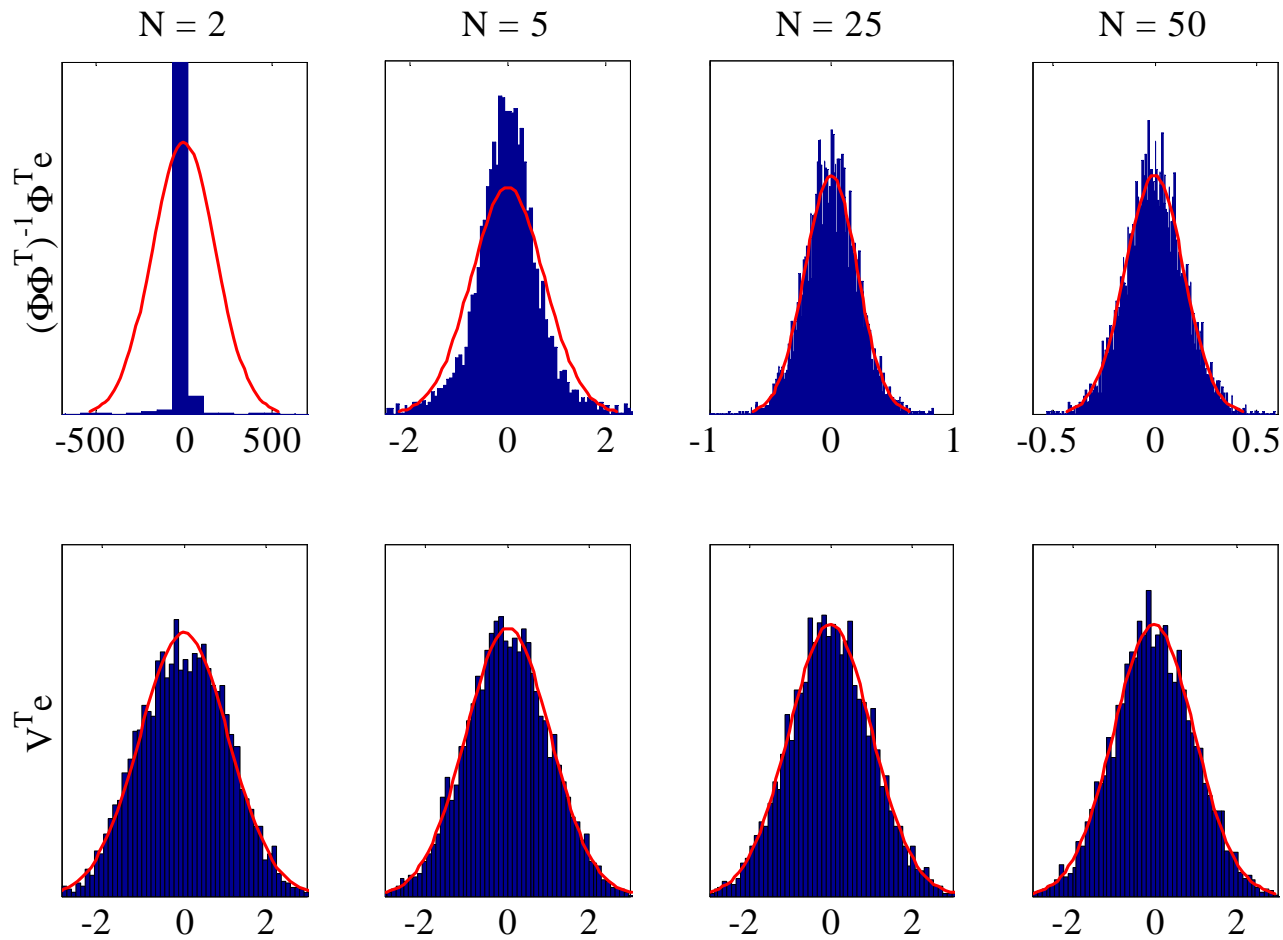
$$\hat{\theta}_N - \theta_0 = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{e}$$

and

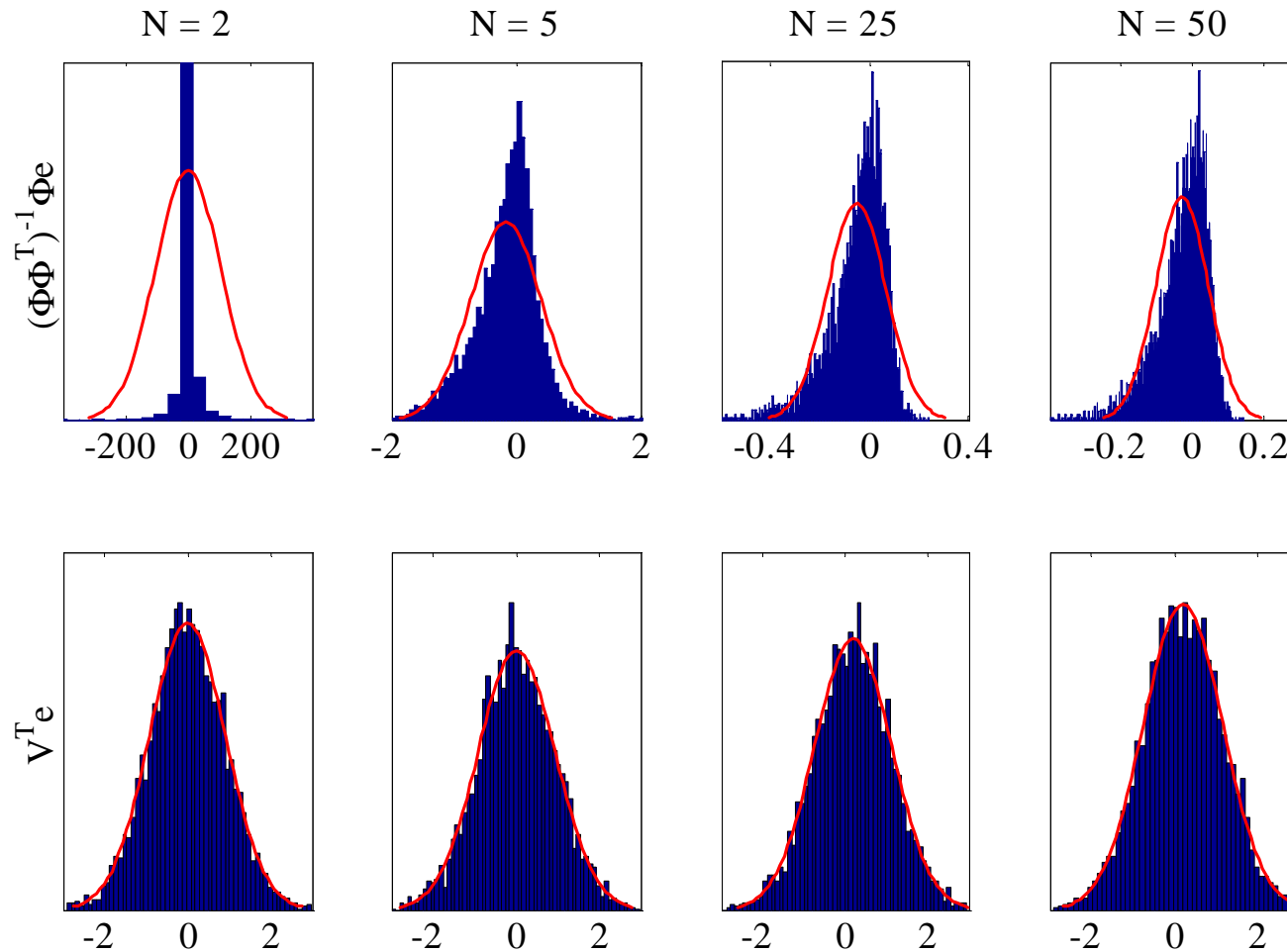
$$\frac{1}{\sqrt{N}} \Phi^T \mathbf{e}$$

for different values of N,
on the basis of 5000 Monte Carlo simulations

Component related to **numerator** parameter:



Component related to **denominator** parameter:



OE modelling

$$\hat{y}(t|t-1; \theta) = \frac{B(q, \theta)}{F(q, \theta)} u(t)$$

Then $V'_N(\hat{\theta}_N) = 0$ can be written as

$$\frac{1}{N} \sum_{t=1}^N \left[y(t) - \frac{B(q, \hat{\theta}_N)}{F(q, \hat{\theta}_N)} u(t) \right] \cdot \psi(t, \hat{\theta}_N) = 0$$

$$\psi(t, \theta) = \frac{\partial}{\partial \theta} \hat{y}(t|t-1; \theta)$$

and $\frac{1}{N} \sum_{t=1}^N [F(q, \hat{\theta}_N) y_F(t) - B(q, \hat{\theta}_N) u_F(t)] \cdot \psi(t, \hat{\theta}_N) = 0$

with $y_F(t) = F(q, \hat{\theta}_N)^{-1} y(t)$; $u_F(t) = F(q, \hat{\theta}_N)^{-1} u(t)$

OE modelling

Linear regression type of equation; solution satisfies

$$\hat{\theta}_N = (\Psi^T \Phi)^{-1} \Psi^T y_F$$

with

$$\Phi^T = [\varphi_F^T(1, \hat{\theta}_N), \dots, \varphi_F^T(N, \hat{\theta}_N)]; \quad \Psi^T = [\psi^T(1, \hat{\theta}_N) \dots \psi^T(N, \hat{\theta}_N)]$$

$$\varphi_F^T(t, \hat{\theta}_N) = [-y_F(t-1) \dots -y_F(t-n_f) \quad u_F(t) \dots u_F(t-n_b+1)]$$

$$\frac{1}{N} \sum_{t=1}^N [F(q, \hat{\theta}_N) y_F(t) - B(q, \hat{\theta}_N) u_F(t)] \cdot \psi(t, \hat{\theta}_N) = 0$$

$$\text{with } y_F(t) = F(q, \hat{\theta}_N)^{-1} y(t); \quad u_F(t) = F(q, \hat{\theta}_N)^{-1} u(t)$$

OE modelling

Linear regression type of equation; solution satisfies

$$\hat{\theta}_N = (\Psi^T \Phi)^{-1} \Psi^T y_F$$

Not fit for parameter estimation, since r.h.s. is parameter-dependent.

However since r.h.s. is known once $\hat{\theta}_N$ is determined, similar uncertainty analysis can be made as for ARX

With $y_F = \Phi \theta_0 + e_F$

$$\frac{1}{\sqrt{N}} (\Psi^T \Phi) (\hat{\theta}_N - \theta_0) = \frac{1}{\sqrt{N}} \Psi^T e_F$$

OE modelling

$$\theta_0 \in \left\{ \theta \mid (\hat{\theta}_N - \theta)^T P_{oe,n}^{-1} (\hat{\theta}_N - \theta) \leq \frac{c_\chi(\alpha, n)}{N} \right\} \quad \text{w.p. } \alpha$$

$$P_{oe,n} = \left(\frac{1}{N} \Psi^T \Phi \right)^{-1} Q \left(\frac{1}{N} \Phi^T \Psi \right)^{-1}, \quad Q = \sigma_{e_F}^2 \mathbb{E} \left[\frac{1}{N} \Psi^T \Psi \right]$$

Requires:

- (asymptotic) normality of $\Psi^T e_F / \sqrt{N}$
- Replacement of Q by an estimate
- No 1st order Taylor approximation involved

Classical:
$$P_{oe} = \sigma_e^2 \left[\mathbb{E} \frac{1}{N} \Psi(\theta_0)^T \Psi(\theta_0) \right]^{-1}$$

Summary

- There is an alternative paradigm for parameter uncertainty bounding, without constructing pdf of estimator
- Applicable to ARX, OE and also BJ models
- Leading to simpler and less approximative expressions
- Can be extended to OE models, even when $S \notin \mathcal{M}$
- Relation with Bayesian and likelihood based uncertainty intervals needs to be explored