RECURSIVE BAYESIAN ESTIMATION OF MODELS WITH UNIFORM INNOVATIONS

Lenka Pavelková, Miroslav Kárný

Institute of Information Theory and Automation, Prague, Czech Republic

ABSTRACT

Here, a new view on models with bounded errors is presented. These model are standardly estimated by min-max algorithms without statistical interpretation. Here the Bayesian approach is used to estimate their parameters. The autoregressive model with uniform innovations is defined for this purpose. If also unobservable quantities (states) are considered, the state model with uniform innovations is introduced. An approximation of the posterior probability density for both models is proposed so the estimation can run recursively as required in many application.

1. ARX MODEL WITH UNIFORM INNOVATIONS

1.1. Description

The parameterized model of the system with a single output y_t is described by the probability density function (**pdf**):

$$f(y_t|\psi_t,\Theta) \equiv \mathcal{U}_{y_t}(\theta'\psi, r) \equiv \frac{\chi_{y_t}(-r \leq y_t - \theta'\psi_t \leq r)}{2r} \tag{1}$$

where

ψ_t is regression vector made of past observed data $d(t-1) = d_1, \ldots, d_{t-1}$, $d_i \equiv (y_i, u_i)$ and the current system input u_t; $\psi_t' \equiv [u_t', d_{t-1}, \ldots, d_{t-\partial}, 1]$ with the model order $\partial \geq 0$,

θ is vector of regression coefficients,

$r > 0$ is a positive scalar half-width of the range of the innovations $e_t \equiv y_t - \theta'\psi_t$,

$\Theta \equiv (\theta, r)$ are unknown parameters of the model,

$\mathcal{U}_y(\mu, r)$ is a uniform pdf of y given by expectation μ and half-width $r > 0$,

$\chi_x(x^*)$ is an indicator function of the set x^* evaluated at value x; it equals 1 if $x \in x^*$ and it is zero otherwise.

1.2. Parameter estimation

Parameters are described by the posterior pdf

$$f(\Theta|d(t)) \propto \frac{1}{r^{\nu_t}}\chi_r(\overline{r} \geq r \geq 0)\chi_\Theta(-\mathbf{1}_{\nu_t} r \leq W_t[-1,\theta']' \leq \mathbf{1}_{\nu_t} r)$$

its statistics evolve

$\nu_t = \nu_{t-1} + 1,\ \nu_0 \geq \mathring{\Psi} + 1$ is chosen a priori (data counter)

$W_t' = \left[W_{t-1}', \Psi_t\right],\ W_0$ is chosen a priori (data matrix)

where

$\propto$ denotes proportionality,

This work was supported by Center DAR, MŠMT 1M0572, project Bayes MŠMT 2C06001 and project BADDYR AVČR 1ET 100 750 401

Ψ_t is data vector; $\Psi_t' \equiv [y_t, \psi_t']$,

$\mathbf{1}_{\nu_t}$ is column vector consisting of ν_t units,

$\underline{r}$ is a sure upper bound on r, $\infty \geq r \geq 0$.

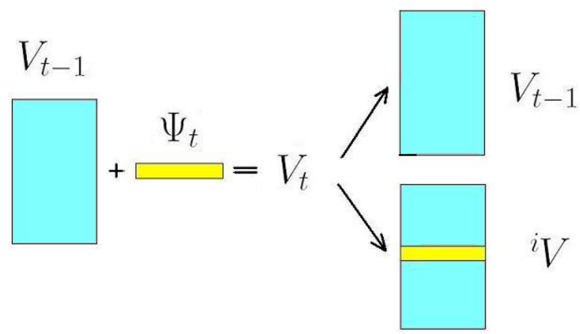
Methods used:
Point estimation - maximum likelihood (ML) estimate $\approx$ linear programming (LP)

1.3. Approximation

The dimension of W_t is time increasing therefore the recursive estimation needs an approximation. The original statistic W_t is replaced by the approximate one V_t.

Problems to be solved:

- Choice of the dimension of the matrix $V_t \Leftrightarrow$ memory length k
- Update and subsequent approximation
  $V_{t-1} +$ new data vector $\Psi_t \rightarrow V_t$ (Kullback-Leibler divergence used as an approximation measure)



The best option minimizes the upper bound $r \in (0, \overline{r})$.

Statistics in time t:

- $V_t = \{V_{t-1}, {}^1V, \ldots, {}^iV, \ldots, {}^kV\}$
- $\nu_t = \{\nu_{t-1}, \nu_{t-1} + 1\}$

1.4. Algorithm

Initialization

- Select the model structure (size of Ψ_t) + the dimension of the statistic V.
- Select lower and upper bounds on the estimated parameters and noise boundary (prior information).
- Construct V_0, choose ν_0 and set $t = 0$.

Recursive mode

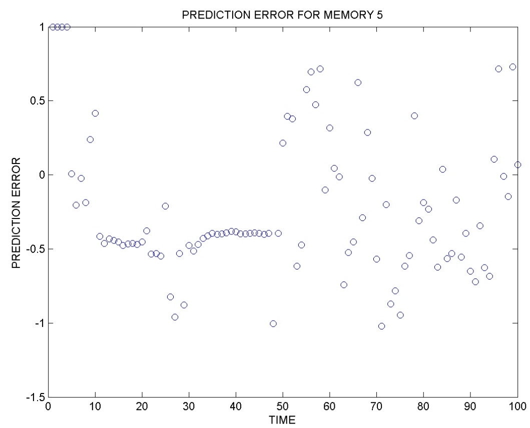
1. Set $t = t + 1$, acquire data d_t and create the data vector Ψ_t.
2. Update the matrix V_{t-1} to the matrix V_t by Ψ_t.

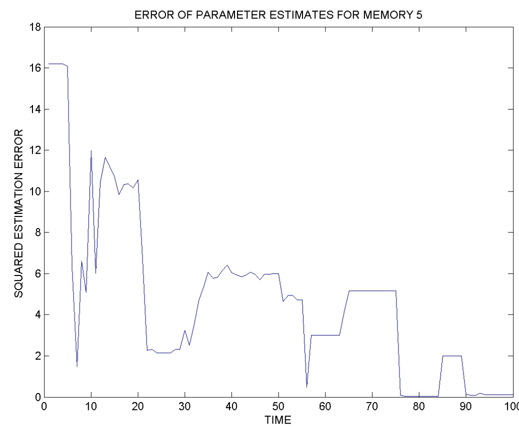
3. If $V_t = V_{t-1}$, then set $\nu_t = \nu_{t-1}$ and preserve the point estimate $\hat{\Theta}_t \equiv \hat{\Theta}_{t-1}$ of parameters Θ otherwise set $\nu_t = \nu_{t-1} + 1$ and update point estimates $\hat{\Theta}_t$. Increase $\overline{r}$ if the above LP fails.

4. Go to the step 1. while data are available.

1.5. Illustrative example

The system described by the model (1) was simulated
with $\psi_t = [y_{t-1}, y_{t-2}, y_{t-3}, 1]$, $r = 1$, $\theta' = [2.85, -2.7075, 0.8574, 0]$.
Parameters of the uniform ARX model were estimated using 100 data samples for memory length $k = 5$.
Prediction errors and the trajectories of $\rho_t \equiv (\hat{\theta}_t - \theta)'(\hat{\theta}_t - \theta)$ are on the following figures:





2. STATE MODEL WITH UNIFORM INNOVATIONS

2.1. Description

The state model with single output y_t and vector state x_t is described by the following pdf's

$$\begin{aligned} f(x_t|A, B, x_{t-1}, u_t) &\equiv \mathcal{U}_{x_t}(Ax_{t-1} + Bu_t, R) \equiv \frac{\chi_{x_t}(-R \le x_t - Ax_{t-1} - Bu_t \le R)}{\prod_i^{dim(R)} 2R_i} \\ f(y_t|C, D, x_t, u_t) &\equiv \mathcal{U}_{y_t}(Cx_t + Du_t, r) \equiv \frac{\chi_{y_t}(-r \le y_t - Cx_t - Du_t \le r)}{2r} \end{aligned} \quad (2)$$

where
$A,\ B,\ C,\ D$ are model parameters, matrices of appropriate dimensions,
$R > 0$ is a positive vector half-widths of the range of innovations $w_t \equiv x_t - Ax_{t-1} - Bu_t$,
$r > 0$ is a positive scalar half-width of the range of innovations $e_t \equiv y_t - Cx_t - Du_t$.

Joint pdf $f(d(t), x(t)|A, B, C, D, R, r)$ can be constructed from conditional pdfs (2)

2.2. Parameters a state estimation

Tasks to be solved:

- Point estimation of the state x_t and the noise boundaries R, r (known A, B, C, D)
- Point estimation of the model parameters A, B, C, D, and the noise boundaries R, r (known states)
- Complete point estimation of the state and parameters (Taylor expansion used)

Methods used:
Point estimation - maximum likelihood (ML) estimate $\approx$ linear programming (LP)

2.3. Approximation

The method of the "sliding window" is used here - LP uses only a finite number past data items.

3. CONCLUSIONS

- An alternative to the deterministic model with bounded errors is proposed.
- Exploitation in the cases when the models with unbounded support do not suite.
- Possibility of the noise-boundary estimation.
- The approximation of the posterior pdf allows the recursive run.
- Restriction on the states decreases parameterization ambiguity.

4. FUTURE PLANS

- Choice of the optimal memory length.
- Search for more expediential expansion point.
- Extension to the multi-dimensional data.
- Application to the traffic data.

The ARX model with uniform innovations is described in [2].

Acknowledgement. This work is supported by grants MŠMT 1M0572, GAČR 1ET100750401, MDČR 1F43A/003/120

5. REFERENCES

[1] A. Bemporad, C. Filippi, and F. Torrisi. Inner and outer approximations of polytopes using boxes. *Computational Geometry*, 27:151–178, 2004.

[2] M. Kárný and L. Pavelková. Projection-based Bayesian Recursive Estimation of ARX Model with Uniform Innovations. *Systems & Control Letters*, 2005 - submitted.

[3] M. Milanese and G. Belforte. Estimation theory and uncertainty intervals evaluation in presence of unknown but bounded errors linear families of models and estimators. *IEEE Transactions on Automatic Control*, 27(2):408—-414, 1982.

[4] V. Peterka. Bayesian system identification. In P. Eykhoff, editor, *Trends and Progress in System Identification*, pages 239–304. Pergamon Press, Oxford, 1981.