TOWARDS BAYESIAN FILTERING ON RESTRICTED SUPPORT

Lenka Pavelková, Miroslav Kárný, Václav Šmídl

Institute of Information Theory and Automation, Prague, Czech Republic

ABSTRACT

Linear state-space model with uniformly distributed innovations is considered. Its state and parameters are estimated under hard physical bounds. Off-line maximum a posteriori probability estimation reduces to linear programming. No approximation is required for sole estimation of either model parameters or states. The noise bounds are estimated in both cases. The algorithm is extended to (i) on-line mode by estimating within a sliding window, and (ii) joint state and parameter estimation. This approach may be used as a starting point for full Bayesian treatment of distributions with restricted support.

1. INTRODUCTION

Statistically inclined research concentrates naturally on stochastic features of the modelled systems and evaluated estimates. However, practically oriented engineers are more concerned with bounds on parameters and states. Respecting physical bounds is a challenging problem especially when the bounds are uncertain. A range of approaches to this problem has been published, including unknown-but-bounded methodology [1] or elliptical approximations [2] to name a few. They often intentionally avoid statistical aspects in order to focus on bounds. This dominating focus restricts their use. Here, we respect both uncertainty and physical bounds by exploiting Bayesian approach. Hard physical bounds are modelled by restricted support of the corresponding joint probability density function (pdf).

We show that off-line evaluation of maximum a posteriori probability (MAP) estimate of linear state-space model with uniform distributions of innovations is equivalent to linear programming (LP). This observation is a starting point for extensions of the approach for more complex models. As a first step, we propose the use of sliding-window for on-line estimation. Moreover, the use of Taylor expansion for approximation of non-linear models yields algorithm for joint parameter and state estimation.

2. LINEAR UNIFORM STATE-SPACE MODEL

We consider the standard linear model

$$x_t = Ax_{t-1} + Bu_t + w_t, \quad y_t = Cx_t + Du_t + v_t, \tag{1}$$

known from Kalman filtering (KF) theory [3]. The notation $x_t \in \Re^n, y_t \in \Re^m, u_t \in \Re^p$, follows traditional conventions for unobserved state, output and input vectors respectively. Also, the identifiers A, B, C, D of the involved parameter matrices of appropriate dimensions are standard in linear filtering as well as common notation Θ for unknown elements in A, B, C, D. Unlike in the KF case, the distributions of vector innovations w_t and v_t are assumed to be uniform

$$f\left(w_t\right) = \mathcal{U}\left(0, r_x\right), \quad f\left(v_t\right) = \mathcal{U}\left(0, r_y\right). \tag{2}$$

$\mathcal{U}\left(\mu, r_x\right)$ denotes uniform pdf on the box with the center μ and half-width of the support interval r_x.

Equations (1) together with the assumptions (2) define the linear uniform state-space model (LU). This model complements its classical Gaussian counterpart (LG) and provides the following advantages: (i) it respects natural bounds on stochastic disturbances, (ii) it allows estimation of the innovation range (unlike KF), and (iii) it allows—without excessive computational demands—to respect hard, physically justified, prior bounds on model parameters and states. Moreover, the presence of finite hard bounds makes the approximate extensions of basic estimation algorithms (such as joint parameter and state estimation) more robust.

3. OFF-LINE ESTIMATION

We assume that the generator of the inputs $u^{1:t} \equiv (u_1, \ldots, u_t)$ meets natural conditions of control [4]. They formalize assumption that information about unknown quantities for generating u_t can only be extracted from the observed data $d^{1:t-1}$, where $d_t = (y_t, u_t)$. Then, for a given initial state x_0, half-widths r_x, r_y and parameters Θ, the joint pdf of data and the state trajectory $x^{1:t}$ of the LU model is

$$f\left(d^{1:t}, x^{1:t} \middle| x_0, r_x, r_y, \Theta\right) \propto \prod_{i=1}^{n} r_{x,i}^{-t} \prod_{j=1}^{m} r_{y,j}^{-t} \chi(\mathcal{S}). \tag{3}$$

$\chi(\mathcal{S})$ is the indicator of the support $\mathcal{S}$. $\propto$ denotes equality up to a constant factor. The convex set $\mathcal{S}$ is defined by inequalities,

$$-r_x \leq x_\tau - Ax_{\tau-1} - Bu_\tau \leq r_x, \; -r_y \leq y_\tau - Cx_\tau - Du_\tau \leq r_y.$$

where $\tau = 1, 2, \ldots, t$. Bayesian estimation of x_0, r_x, r_y requires to complement the conditional pdf (3) by a *prior pdf*

$f(x_0, r_x, r_y|\Theta)$. For known Θ, it can be chosen as uniform pdf on support $\mathcal{S}_0$ defined by inequalities

$$\mathcal{S}_0 = \{\underline{x}_0 \le x_0 \le \overline{x}_0, \quad 0 < r_x \le \overline{r_x}, \quad 0 < r_y \le \overline{r_y}\}. \quad (4)$$

Here, $\overline{x}_0$ denotes a prior upper bound on the initial state. The meaning of other bounds is similar. For unknown Θ, the uniform prior pdf $f(x_0, r_x, r_y, \Theta)$ can be chosen on the set (4) extended by conditions $\underline{\Theta} \le \Theta \le \overline{\Theta}$.

For fixed observations $d^{1:t}$ and uniform prior (4), the expression (3)—on support $\mathcal{S} \cap \mathcal{S}_0$—is proportional to *posterior pdf*. Due to the power $^{-t}$, it is sharply peaked at lower bounds on r_x and r_y. Moreover, the number of vertices of the support is proportional to the number of data. The proportionality factor may be large for realistic systems. Consequently, evaluation of moments of this pdf is computationaly demanding. This motivates our focus on MAP estimation of all unknowns.

Without loss of generality, we assume that elements of r_x and r_y are (significantly) smaller than 1. Under this assumption, the negative logarithm of the posterior pdf can be approximated by sum of elements of r_x and r_y on the convex, linearly restricted set $\mathcal{S} \cap \mathcal{S}_0$. Thus, MAP estimate of states *and* r_x, r_y, for a given Θ, is found by linear programming (LP). Efficient algorithms for solution of high-dimensional LP problems are widely available.

4. ON-LINE ESTIMATION

Standard Bayesian *filtering and smoothing with a fixed lag* $\partial \ge 0$ integrates out from the posterior pdf the superfluous state $x_{t-\partial-1}$ in each time step, t. However, with increasing t, this operation yields increasingly complex support of the posterior pdf and soon becomes intractable. The unknown-but-bounded approaches [1, 2] face this problem by a recursive construction of simple (typically outer) approximation of the support. In order to avoid these approximations, we propose to use a sliding window of length ∂ and apply LP in order to find MAP estimate of the states $x^{t-\partial:t} \equiv (x_{t-\partial}, \ldots, x_t)$ on the intersection of sets $\mathcal{S}$ and $\mathcal{S}_0$ for $\tau = t - \partial, \ldots, t$. This approximation immediately opens a way to feasible *joint parameter and state estimation*. It is sufficient to linearize the product

$$Ax_{t-1} \approx \hat{A}x_{t-1} + A\hat{x}_{t-1} - \hat{A}\hat{x}_{t-1}. \quad (5)$$

By handling Cx_t in the same way, we get the approximate joint estimate as solution of the corresponding LP task. Note that unlike in extended KF, the algorithm updates estimates of the whole window of length ∂ hence, points of expansion $\hat{A}, \hat{x}_{t-1}$ in (5) can be re-evaluated in each time-step as a moving average of MAP estimates on the whole window. Moreover, the expansion (5), which is often used in extended KF is expected to be better conditioned due to the exploited realistic bounds on the estimated quantities. This conjecture is supported by simulation experiments, which will be presented in the final paper. For illustration, estimates of lower bounds on interval of innovations r_x, r_y for a simple LU model with two-dimensional state and one-dimensional observations is presented in Figure 1.

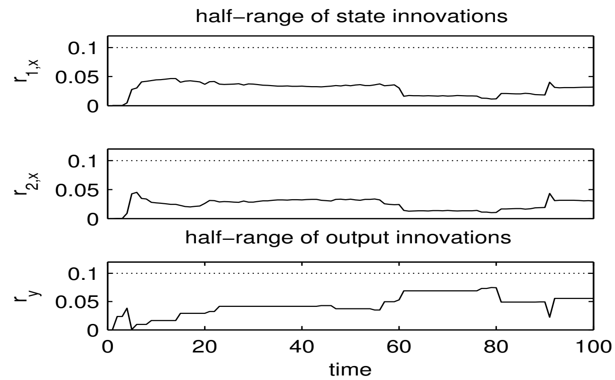

Fig. 1. Illustration of performance of the algorithm for estimation of half-range of innovations. Dashed line denotes simulated and full-lines estimated values.

5. CONCLUDING REMARKS

The proposed approach opens a way for on-line parameter and state estimation for a class of non-uniform distributions with restricted support as well as for Bayesian filtering of non-linear systems. The directly feasible cases are those in which linear programming is replaced by convex programming. Moreover, the outer approximation of the support by ellipsoids or by union of boxes is a good preliminary step for an efficient application of sampling-based estimation algorithms.

Acknowledgement: This work was supported by grants MŠMT 1M0572 (DAR) and AVČR 1ET 100 750 401 (BADDYR).

6. REFERENCES

[1] 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*, vol. 27, no. 2, pp. 408–414, 1982.

[2] B. Polyak, S. Nazin, C. Durieu, and E. Walter, "Ellipsoidal parameter or state estimation under model uncertainty," *Automatica*, vol. 40, no. 7, pp. 1171–1179, 2004.

[3] A. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic Press, 1970.

[4] V. Peterka, "Bayesian system identification," in *Trends and Progress in System Identification* (P. Eykhoff, ed.), pp. 239–304, Oxford: Pergamon Press, 1981.