Off-line Estimation of System Noise Covariance Matrices by a Special Choice of the Filter Gain

Miroslav Šimandl, Jindřich Duník

Research Centre Data - Algorithms - Decision Making
Department of Cybernetics
Faculty of Applied Sciences
University of West Bohemia
Czech Republic

October 4, 2007

Description of system

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) + \mathbf{w}_k, \ \ k = 0, 1, 2, \ldots$$
$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k, \ \ k = 0, 1, 2, \ldots$$

where

- $\mathbf{f}_k(\cdot)$, $\mathbf{h}_k(\cdot)$ are known vector functions,
- $p(\mathbf{w}_k) = \mathcal{N}\{\mathbf{w}_k : \mathbf{0}, \mathbf{Q}\}$, $p(\mathbf{v}_k) = \mathcal{N}\{\mathbf{v}_k : \mathbf{0}, \mathbf{R}\}$, and $p(\mathbf{x}_0) = \mathcal{N}\{\mathbf{x}_0 : \hat{\mathbf{x}}_0, \mathbf{P}_0\}$ are Gaussian probability density functions (pdf's).

State estimation problem

The aim of the filtering is to find the probability density function of the state $\mathbf{x}_k$ conditioned by the measurements $\mathbf{z}^k = [\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_k]$.

$$p(\mathbf{x}_k | \mathbf{z}^k) = ?$$

Solution of state estimation

General solution of the filtering problem is given by the Bayesian Recursive Relations (BRR's).

Solution of BRR's

- The closed-form solution of the BRR's is possible only for a few special cases, e.g. for linear Gaussian system, which leads to the Kalman Filter (KF).
- In other cases it is necessary to apply some approximative techniques, which leads into local and global filtering methods.

BRR's and description of system noises

The design and application of an arbitrary filter is conditioned by the knowledge of model of the real system, i.e.

noise covariance matrices **Q** and **R** have to be known.

Methods for estimation of noise cov. matrices: on-line

- adaptive filtering methods (Mehra, 1972):
  - often based on the correlation analysis of innovation sequence of the Kalman Filter
  - lot of measurements is required
  - designed for linear systems only
- minimax filtering methods (Verdú and Poor, 1984):
  - based on the game theory
  - uncertainty classes have to be specified
  - designed for linear systems only

Methods for estimation of noise cov. matrices: off-line

- prediction error methods (Ljung, 1999)
- subspace ident. methods (van Overshee and de Moor, 1996)
- designed for linear or special types of nonlinear systems only

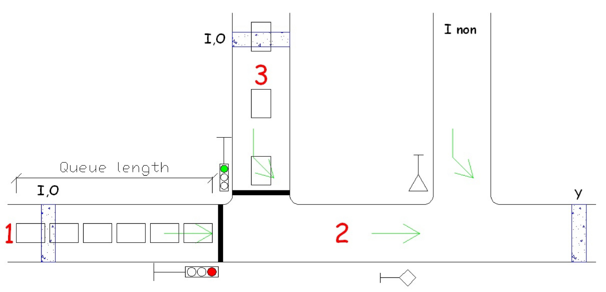
Goal of the paper

The goal of the paper is to propose an alternative technique for off-line estimation of the state and measurement covariance matrices $\mathbf{Q}$ and $\mathbf{R}$. The technique should be applicable for both linear and nonlinear systems.

The proposed technique will be based

- on the possibility to measure the initial period of the monitoring system output repeatedly and
- on the knowledge of system initial condition.

Example of an Intersection



Measurable and immeasurable quantities

- measurable quantity (part of measurement vector $\mathbf{z}_k$):
  - Intensity of traffic flow I - number of cars per time period
  - Intensity is measured repeatedly each day.
- immeasurable quantity (part of state vector $\mathbf{x}_k$):
  - Queue length - number of waiting cars
  - Known system initial condition e.g. at night.

Description of linear system

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{w}_k, \quad k = 0, 1, 2, \ldots$$
$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k, \quad k = 0, 1, 2, \ldots$$

Kalman Filter

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{F}\hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\mathbf{e}_k = (\mathbf{F} - \mathbf{K}_k\mathbf{H})\hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\mathbf{z}_k,$$
$$\mathbf{P}_{k+1|k} = \mathbf{F}\mathbf{P}_{k|k-1}\mathbf{F}^T - \mathbf{K}_k\mathbf{H}\mathbf{P}_{k|k-1}\mathbf{F}^T + \mathbf{Q}^F,$$

where

- Kalman gain is $\mathbf{K}_k = \mathbf{F}\mathbf{P}_{k|k-1}\mathbf{H}^T(\mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^T + \mathbf{R}^F)^{-1}$,
- innovation is $\mathbf{e}_k = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1} = \mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_{k|k-1}$, and
- filter initial condition is $p(\mathbf{x}_0|\mathbf{z}^{-1}) = p(\mathbf{x}_0)$.

Choice of filter noise covariance matrices

- If the true noise matrices are known, they are used in the filter design, i.e. $\mathbf{Q}^F = \mathbf{Q}$ and $\mathbf{R}^F = \mathbf{R}$.
- In other cases let the filter noise covariance matrices be chosen as positive definite matrices fulfilling condition $\mathbf{R}^F \gg \mathbf{Q}^F$.

Impact of chosen matrices on innovation sequence

The choice of matrices $\mathbf{Q}^F$ and $\mathbf{R}^F$ causes

- insignificant Kalman gain $\mathbf{K}_k$, i.e. $\mathbf{K}_k \approx \mathbf{0}$, and thus
- innovation in the form

$$\mathbf{e}_k \approx \mathbf{H}\mathbf{F}^k(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|-1}) + \mathbf{H}\sum_{i=0}^{k-1}\mathbf{F}^i\mathbf{w}_{k-i} + \mathbf{v}_k.$$

Mean

With respect to zero mean of noise in the state and measurement equation, the mean of innovation is

$$E[\mathbf{e}_k] = E[\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}] = 0, \forall k.$$

Covariance matrix

The covariance matrix of innovation is

$$\mathrm{cov}[\mathbf{e}_k] \approx \mathbf{H}\mathbf{F}^k\mathbf{P}_{0|-1}(\mathbf{F}^k)^T\mathbf{H}^T + \sum_{i=0}^{k-1}\mathbf{H}\mathbf{F}^i\mathbf{Q}(\mathbf{F}^i)^T\mathbf{H}^T + \mathbf{R},$$

without any dependence on chosen $\mathbf{Q}^F$ and $\mathbf{R}^F$, $\forall k$.

Measured data and sample covariance matrices ($\mathrm{cov}[\mathbf{e}_k]$)

Let m sets of data be measured, i.e. data $\mathbf{z}^{(i)} = [\mathbf{z}_0^{(i)}, \mathbf{z}_1^{(i)}, \ldots, \mathbf{z}_n^{(i)}]$, are at disposal, where $i = 1, 2, \ldots, m$.

On the basis of the data sets, the innovation sequences $\mathbf{e}^{(i)} = [\mathbf{e}_0^{(i)}, \mathbf{e}_1^{(i)}, \ldots, \mathbf{e}_n^{(i)}]$, $\forall i$, can be computed.

Innovation sequences allow to compute sample cov. matrices

$$\mathrm{cov}[\mathbf{e}_k] \approx \hat{\mathbf{P}}_{e_k} = \frac{1}{m-1}\sum_{j=1}^{m} \mathbf{e}_k^{(j)}(\mathbf{e}_k^{(j)})^T, \forall k.$$

Estimation of noise covariance matrices

True matrices $\mathbf{Q}$ and $\mathbf{R}$ (and possibly $\mathbf{P}_{0|-1}$) can be found by standard least square method from relation

$$\hat{\mathbf{P}}_{e_k} \approx \mathbf{H}\mathbf{F}^k\mathbf{P}_{0|-1}(\mathbf{F}^k)^T\mathbf{H}^T + \sum_{i=0}^{k-1}\mathbf{H}\mathbf{F}^i\mathbf{Q}(\mathbf{F}^i)^T\mathbf{H}^T + \mathbf{R}, \forall k.$$

Description of nonlinear system

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) + \mathbf{w}_k, \ \ k = 0, 1, 2, \ldots$$
$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k, \ \ k = 0, 1, 2, \ldots$$

with known system initial state $\mathbf{x}_0$.

Extended Kalman filter

- As an representative of approximative techniques used to the solution of Bayesian recursive relations for nonlinear systems, the Extended Kalman Filter (EKF) was chosen.
- The extended Kalman filter can be understood as the KF applied to a linearised nonlinear system.
- The linearisation is performed by the first-order Taylor expansion.

Innovation sequence

Let inequality $\mathbf{R}^F \gg \mathbf{Q}^F$ holds, then innovation of the EKF is

$$\mathbf{e}_k \approx \sum_{i=1}^{k-1} \mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1}) \prod_{j=i}^{k-1} \mathbf{F}_l(\hat{\mathbf{x}}_{l|l})\mathbf{w}_l + \mathbf{v}_k,$$

where $l = k - j + i$ and linearised system matrices are

$\mathbf{F}_k(\hat{\mathbf{x}}_{k|k}) = \frac{\partial \mathbf{f}_k(\mathbf{x}_k)}{\partial \mathbf{x}_k}|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k}}$, $\mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1}) = \frac{\partial \mathbf{h}_k(\mathbf{x}_k)}{\partial \mathbf{x}_k}|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k-1}}$.

Covariance matrix of innovation sequence

Covariance matrix of innovation sequence of the EKF is

$$\mathrm{cov}[\mathbf{e}_k] \approx \sum_{i=1}^{k-1} \mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1}) \prod_{j=i}^{k-1} \mathbf{F}_l(\hat{\mathbf{x}}_{l|l})\mathbf{Q}\left(\mathbf{H}_k(\cdot) \prod_{j=i}^{k-1} \mathbf{F}_l(\cdot)\right)^T +$$

$$+ \mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1})\mathbf{Q}\mathbf{H}_k^T(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{R}.$$

Measured data and sample covariance matrices

m sets of measured data $\mathbf{z}^{(i)} = [\mathbf{z}_0^{(i)}, \mathbf{z}_1^{(i)}, \ldots, \mathbf{z}_n^{(i)}]$ allow to compute sets of the innovation sequences $\mathbf{e}^{(i)} = [\mathbf{e}_0^{(i)}, \mathbf{e}_1^{(i)}, \ldots, \mathbf{e}_n^{(i)}]$, $i = 1, 2, \ldots, m$ by means of the EKF with properly chosen matrices $\mathbf{Q}^F$ and $\mathbf{R}^F$ and to find sample covariance matrices to

$$\mathrm{cov}[\mathbf{e}_k] \approx \hat{\mathbf{P}}_{e_k} = \frac{1}{m-1}\sum_{j=1}^{m} \mathbf{e}_k^{(j)}(\mathbf{e}_k^{(j)})^T, \forall k.$$

Estimation of noise covariance matrices

True matrices $\mathbf{Q}$ and $\mathbf{R}$ can be found by the least square method from relation

$$\hat{\mathbf{P}}_{e_k} \approx \sum_{i=1}^{k-1} \mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1}) \prod_{j=i}^{k-1} \mathbf{F}_l(\hat{\mathbf{x}}_{l|l})\mathbf{Q}\Big(\mathbf{H}_k(\cdot)\prod_{j=i}^{k-1}\mathbf{F}_l(\cdot)\Big)^T +$$
$$+ \mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1})\mathbf{Q}\mathbf{H}_k^T(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{R}.$$

Description of linear system

The system is defined as

$$\mathbf{x}_{k+1} = \begin{bmatrix} 0.8 & 0.9 \\ -0.4 & 0.5 \end{bmatrix} \mathbf{x}_k + \mathbf{w}_k,$$

$$z_k = [1 \quad 0]\mathbf{x}_k + v_k,$$

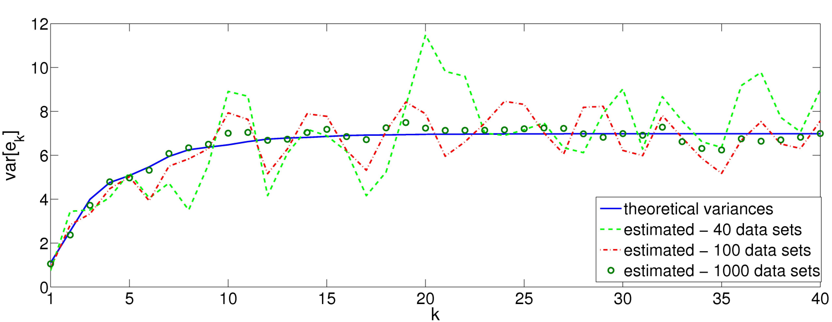
where

- $p(\mathbf{w}_k) = \mathcal{N}\{\mathbf{w}_k : [0,0]^T, \mathbf{I}\}$, $p(v_k) = \mathcal{N}\{v_k : 0, 0.1\}$, $\forall k$,
- $p(\mathbf{x}_0|z^{-1}) = p(\mathbf{x}_0) = \mathcal{N}\{\mathbf{x}_0 : [20,20]^T, 0.1 \times \mathbf{I}\}$,
- $\mathbf{Q}^F = 10^{-18} \times \mathbf{I}$, and $R^F = 10^{18}$.

Aim

The aim is to estimate diagonal elements of state noise covariance matrix $\mathbf{Q}$ and measurement noise variance R.

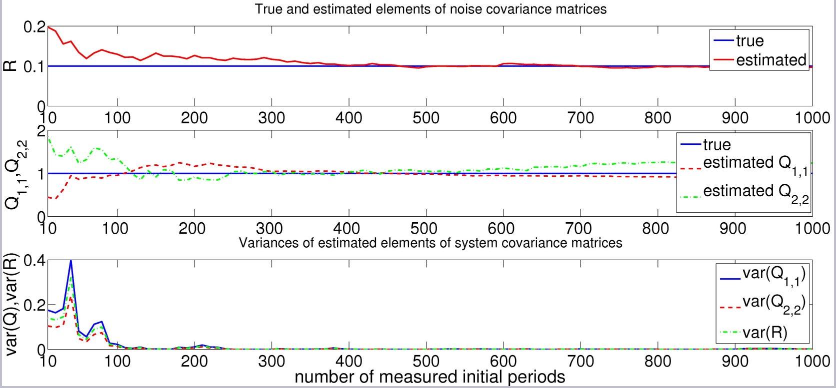
Theoretical and estimated variances of innovation seq.



Estimation of noise covariance matrices

Estimation of the noise covariance matrices was based on three equations: cov$[e_0]$, cov$[e_1]$, and cov$[e_2]$.

True end estimated elements of noise covariance matrices



Description of nonlinear system

The system is defined as

$$x_{1,k+1} = x_{1,k}x_{2,k} + w_{1,k},$$
$$x_{2,k+1} = x_{2,k} + w_{2,k},$$
$$z_k = x_{1,k}^2 + v_k,$$

where

- $p(v_k) = \mathcal{N}\{v_k : 0, 0.01\}$,
  $p(\mathbf{w}_k) = \mathcal{N}\{\mathbf{w}_k : [0,0]^T, diag([0.25, 10^{-3}])\}$,
- $p(\mathbf{x}_0) = p(\mathbf{x}_0|z^{-1}) = \mathcal{N}\{\mathbf{x}_0 : [40, 0.9]^T, 10^{-8} \times \mathbf{I}\}$,
- $\mathbf{Q}^F = \mathbf{I}$, and $R^F = 10^{18}$.

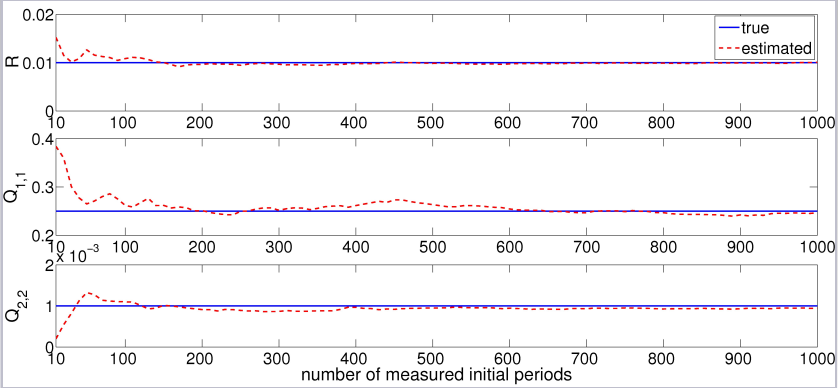
Aim

The aim is to estimate diagonal elements of state noise covariance matrix $\mathbf{Q}$ and measurement noise variance R.

Estimation of noise covariance matrices

Estimation of the noise covariance matrices was based on three equations: $\mathrm{cov}[e_0]$, $\mathrm{cov}[e_1]$, and $\mathrm{cov}[e_0; e_1]$.

True end estimated elements of noise covariance matrices



Conclusion Remarks

- Techniques for estimation of the system noise covariance matrices were discussed.
- The novel off-line estimation technique for linear and nonlinear systems was proposed.
- The technique is based on the known system initial condition and on the possibility to measure initial period repeatedly.
- The technique allows to determine a sufficient number of independent equations to estimate all elements of the system noise covariance matrices.