Estimation of State and Measurement Noise Covariance Matrices by Multi-Step Prediction

Jindřich Duník and Miroslav Šimandl

Department of Cybernetics, Faculty of Applied Sciences, University of West Bohemia, Czech Republic

System description

The system is specified by the state and measurement equations

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

where

- $\mathbf{x}_k \in \mathbb{R}^{n_x}$ and $\mathbf{z}_k \in \mathbb{R}^{n_z}$ represent the state of the system and the measurement, respectively,
- $\mathbf{f}_k(\cdot)$ and $\mathbf{h}_k(\cdot)$ are the known vector functions,
- the system initial condition is given by the known mean $\bar{\mathbf{x}}_0$ and covariance matrix $\mathbf{P}_0$, and
- $\mathbf{w}_k$ and $\mathbf{v}_k$ are the state and measurement noises with the zero means and with unknown noise covariance matrices $\mathbf{Q}$ and $\mathbf{R}$, respectively.

Short overview of noise covariance matrices estimation techniques

The noise covariance matrices estimation techniques can roughly be divided into two groups, namely

- off-line methods, e.g.
  prediction error method, subspace identification method,
- on-line methods, e.g.
  adaptive filtering methods, minimax filtering methods.

However, both off-line and on-line methods are suitable mainly for linear systems or for special types of nonlinear systems.

Goal of the paper

The goal of the paper is to propose an off-line technique for estimation of the noise covariance matrices for both linear and nonlinear systems.

- The proposed technique is based on the multi-step prediction error and on knowledge of the system initial condition. The technique takes an advantage of the well-known standard relations from the area of state estimation methods and least square method.
- At first, the technique will be introduced for linear systems and then it will be extended into the area of nonlinear systems.

Noise covariance matrix estimation for linear systems

Multi-step prediction

Let the linear t-invariant system (1), (2) with $\mathbf{f}_k(\mathbf{x}_k) = \mathbf{F}\mathbf{x}_k$ and $\mathbf{h}_k(\mathbf{x}_k) = \mathbf{H}\mathbf{x}_k$ be considered.

Then, the multi-step prediction of the state is specified by

$$\hat{\mathbf{x}}_{k|-1} = \mathbf{F}\hat{\mathbf{x}}_{k-1|-1}$$

and the multi-step prediction of the measurement by

$$\hat{\mathbf{z}}_{k|-1} = \mathbf{H}\hat{\mathbf{x}}_{k|-1} = \mathbf{H}\mathbf{F}^k\hat{\mathbf{x}}_{0|-1}, \quad (3)$$

where $\hat{\mathbf{x}}_{0|-1} = \bar{\mathbf{x}}_0$.

Statistical properties of multi-step measurement prediction error

With respect to (1), (2), and (3), the prediction error of the measurement is of the form

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

The measurement prediction error $\mathbf{e}_k$ at time instant k is the zero mean variable with the covariance matrix

$$\mathrm{cov}[\mathbf{e}_k] = \mathrm{cov}[\mathbf{z}_k - \hat{\mathbf{z}}_{k|-1}] = \mathbf{H}\mathbf{F}^k\mathbf{P}_0(\mathbf{F}^k)^T\mathbf{H}^T + \sum_{i=0}^{k-1}\mathbf{H}\mathbf{F}^i\mathbf{Q}(\mathbf{F}^i)^T\mathbf{H}^T + \mathbf{R}. \quad (4)$$

The cross-covariance matrix of the innovations $\mathbf{e}_l$ and $\mathbf{e}_k$ is given by

$$\mathrm{cov}[\mathbf{e}_l; \mathbf{e}_k] = \mathrm{cov}[\mathbf{z}_l - \hat{\mathbf{z}}_{l|-1}; \mathbf{z}_k - \hat{\mathbf{z}}_{k|-1}] = \mathbf{H}\mathbf{F}^l\mathbf{P}_0(\mathbf{F}^k)^T\mathbf{H}^T + \sum_{i=1}^{l}\mathbf{H}\mathbf{F}^{l-i}\mathbf{Q}(\mathbf{F}^{\mathbf{k}-\mathbf{i}})^\mathbf{T}\mathbf{H}^\mathbf{T}, \quad (5)$$

$\forall k, l, l < k$.

Relations (4) and (5) depend on the known matrices $\mathbf{F}$ and $\mathbf{H}$ and on the unknown noise covariance matrices $\mathbf{Q}$ and $\mathbf{R}$.

Computation of sample prediction error covariance matrices

The prediction error (cross-)covariance matrices can be computed by the following two ways.

- Computation by means of repeatedly measured data sets

  Let N repeatedly measured data sets be supposed, i.e. $[\mathbf{z}_0^{(i)}, \mathbf{z}_1^{(i)}, \ldots, \mathbf{z}_n^{(i)}]$, $i = 1, 2, \ldots, N$. Then, N sets of the prediction errors $[\mathbf{e}_0^{(i)}, \mathbf{e}_1^{(i)}, \ldots, \mathbf{e}_n^{(i)}]$, $i = 1, 2, \ldots, N$, can be found. On the basis of these sets, the sample prediction error (cross-)covariance matrices can be computed according to

$$\hat{\mathbf{P}}_{e_k} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{e}_k^{(i)}(\mathbf{e}_k^{(i)})^T \approx \mathrm{cov}[\mathbf{e}_k], \quad (6)$$

$$\hat{\mathbf{P}}_{e_l, e_k} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{e}_l^{(i)}(\mathbf{e}_k^{(i)})^T \approx \mathrm{cov}[\mathbf{e}_l; \mathbf{e}_k]. \quad (7)$$

- Computation by means of single measured data set

  Let a measured data set be supposed $[\mathbf{z}_0, \ldots, \mathbf{z}_n]$ and the set of the prediction errors $[\mathbf{e}_0, \ldots, \mathbf{e}_n]$ be computed. Due to the fact, that for stable linear systems the equality $\mathrm{cov}[\mathbf{e}_k] = \mathrm{cov}[\mathbf{e}_{k+1}]$, $k = L, L+1, \ldots, n$, holds, the sample prediction error covariance matrix can be computed as

$$\hat{\mathbf{P}}_{e_L} = \frac{1}{n-L+1}\sum_{k=L}^{n}\mathbf{e}_k(\mathbf{e}_k)^T \approx \mathrm{cov}[\mathbf{e}_L]. \quad (8)$$

  Similarly, the sample cross-covariance matrix can be computed.

Estimation of state and measurement noise covariance matrices for linear systems

Substitution of the sample (cross-)covariance matrices (6)–(8) into (4), (5) leads to the relations

$$\hat{\mathbf{P}}_{e_k} = \mathbf{H}\mathbf{F}^k\mathbf{P}_0(\mathbf{F}^k)^T\mathbf{H}^T + \sum_{i=0}^{k-1}\mathbf{H}\mathbf{F}^i\hat{\mathbf{Q}}(\mathbf{F}^i)^T\mathbf{H}^T + \hat{\mathbf{R}},$$

$$\hat{\mathbf{P}}_{\mathbf{e}_l, \mathbf{e}_k} = \mathbf{H}\mathbf{F}^l\mathbf{P}_0(\mathbf{F}^k)^T\mathbf{H}^T + \sum_{i=1}^{l}\mathbf{H}\mathbf{F}^{l-i}\hat{\mathbf{Q}}(\mathbf{F}^{\mathbf{k}-\mathbf{i}})^\mathbf{T}\mathbf{H}^\mathbf{T},$$

representing the system of linear equations, where the unknown variables are the elements of the noise covariance matrices. Solution of the system of linear equations leads to the estimates of the state noise covariance matrix $\hat{\mathbf{Q}}$ and the measurement noise covariance matrix $\hat{\mathbf{R}}$.

Noise covariance matrix estimation for nonlinear systems

Multi-step prediction by extended Kalman filter

- Let the nonlinear system (1), (2) with known initial state $\mathbf{x}_0$ be considered. Then, the extended Kalman filter (EKF) is defined as follows

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

  where
  - $\hat{\mathbf{x}}_{k|k}$, $\mathbf{P}_{k|k}$ and $\hat{\mathbf{x}}_{k+1|k}$, $\mathbf{P}_{k+1|k}$ are the filtering and predictive state means and covariance matrices,
  - $\hat{\mathbf{z}}_{k|k-1}$ is the predictive mean of the measurement,
  - $\mathbf{F}_k(\hat{\mathbf{x}}_{k|k}) = \frac{\partial \mathbf{f}_k(\mathbf{x}_k)}{\partial \mathbf{x}_k}|_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k}}$ and $\mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1}) = \frac{\partial \mathbf{h}_k(\mathbf{x}_k)}{\partial \mathbf{x}_k}|_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1}}$,
  - $\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\hat{\mathbf{x}}_{k|k-1})\left(\mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1})\mathbf{P}_{k|k-1}\mathbf{H}_k^T(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{R}^F\right)^{-1}$ is the filter gain,
  - $\mathbf{e}_k = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}$ is the innovation, and
  - the filter initial condition is given by $\hat{\mathbf{x}}_{0|-1} = \mathbf{x}_0$, $\mathbf{P}_{0|-1} \to \mathbf{0}$.
- The innovation sequence of the EKF remains bounded if the system is observable, the system and filter initial states are the same, and the filter noise covariance matrices are greater than true ones, i.e. $\mathbf{Q}^F > \mathbf{Q}$, $\mathbf{R}^F > \mathbf{R}$.
- Additionally, if the filter noise covariance matrices will be chosen so that the inequality $\mathbf{Q}^F \ll \mathbf{R}^F$ holds, then the gain will be close to the zero, i.e. $\mathbf{K}_k \to \mathbf{0}$, for a few first time instants and the EKF works as a multi-step predictor.

Statistical properties of extended Kalman filter innovation sequence

Then, the innovation of the EKF is the zero mean variable with the covariance matrix

$$\mathrm{cov}[\mathbf{e}_0] \approx \mathbf{R}, \quad (9)$$
$$\mathrm{cov}[\mathbf{e}_1] \approx \mathbf{H}_1(\hat{\mathbf{x}}_{1|0})\mathbf{Q}\mathbf{H}_1^T(\hat{\mathbf{x}}_{1|0}) + \mathbf{R}, \quad (10)$$
$$\mathrm{cov}[\mathbf{e}_k] \approx \sum_{i=1}^{k-1}\mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1})\prod_{j=i}^{k-1}\mathbf{F}_m(\hat{\mathbf{x}}_{m|m})\mathbf{Q}\left(\mathbf{H}_k(\cdot)\prod_{j=i}^{k-1}\mathbf{F}_m(\cdot)\right)^T + \mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1})\mathbf{Q}\mathbf{H}_k^T(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{R}, \quad (11)$$

where $k = 2, 3\ldots$, $m = k - j + i - 1$. The cross-covariance matrices can be computed as well.

Estimation of state and measurement noise covariance matrices for nonlinear systems

The noise covariance matrices can be estimated analogously to the linear case. In short, the following steps are necessary.

- To compute the sample covariance matrices of the innovation $\mathbf{e}_k$ on the basis of the repeatedly measured data sets.
- To substitute the sample covariance matrices $\hat{\mathbf{P}}_{e_k}$ into relations (9)–(11) and to solve the system of linear equations with respect to the elements of the noise covariance matrices $\mathbf{Q}$ and $\mathbf{R}$.

Numerical example

- Let the following nonlinear system be supposed

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

  where
  - the system and filter initial state is $\mathbf{x}_0 = \hat{\mathbf{x}}_{0|-1} = [40,\ 0.95]^T$,
  - the state noise is described by $p(\mathbf{w}_k) = \mathcal{N}\{[w_{1,k},\ w_{2,k}]^T : [0,\ 0]^T, \mathrm{diag}([0.25,\ 0.004])\}$, $\forall k$, and
  - the measurement noise is described by $p(v_k) = \mathcal{N}\{v_k : 0, 0.01\}$, $\forall k$.
- The system output was repeatedly measured for first three time instant, i.e. data sets $[z_0^{(i)}, z_1^{(i)}, z_2^{(i)}]$ for $i = 1, 2, \ldots, 1000$ were available.
- The EKF with $\mathbf{Q}^F = \mathrm{diag}([40,\ 40])$ and $R^F = 10^8$ was applied to all sets of measured data.