

A Generalized Autocovariance Least-Squares Method for Covariance Estimation

Bernt M. Åkesson
CAPEC

Department of Chemical Engineering
Technical University of Denmark
DK-2800 Kgs. Lyngby
Email: baa@kt.dtu.dk

John Bagterp Jørgensen
Informatics and Mathematical Modelling

Technical University of Denmark
DK-2800 Kgs. Lyngby
Email: jbj@imm.dtu.dk

Sten Bay Jørgensen
CAPEC

Department of Chemical Engineering
Technical University of Denmark
DK-2800 Kgs. Lyngby
Email: sbj@kt.dtu.dk

Abstract—A generalization of the autocovariance least-squares method for estimating noise covariances is presented. The method can estimate mutually correlated system and sensor noise and can be used with both the predicting and the filtering form of the Kalman filter.

Index Terms—Covariance estimation, optimal estimation, state estimation.

I. INTRODUCTION

The Kalman filter requires knowledge about the noise statistics. In practical applications, however, the noise covariances are generally not known. The autocovariance least-squares (ALS) method was presented by Odelson et al. [1] as a technique for estimating the system and sensor noise covariances from plant data. The technique was shown to give unbiased estimates with smaller variance than previously proposed methods, such as the correlation method by Mehra [2]. The objective of this paper is to demonstrate how the method can be extended to systems where the system noise and the sensor noise are mutually correlated. Moreover, the generalized method works with both the predicting and the filtering form of the Kalman filter.

II. GENERALIZED AUTOCOVARANCE LEAST-SQUARES ESTIMATION

Consider a linear time-invariant system in discrete-time,

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + Gw_k \\ y_k &= Cx_k + v_k \end{aligned} \quad (1)$$

where $A \in \mathbb{R}^{n_x \times n_x}$, $B \in \mathbb{R}^{n_x \times n_u}$, $G \in \mathbb{R}^{n_x \times n_w}$ and $C \in \mathbb{R}^{n_y \times n_x}$. The process noise w_k and the measurement noise v_k are zero-mean white noise processes according to

$$\begin{bmatrix} w_k \\ v_k \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q_w & S_{wv} \\ S_{wv}^T & R_v \end{bmatrix} \right). \quad (2)$$

Assume that a suboptimal stationary Kalman filter is used to estimate the state. The filter is based on initial guesses of covariances Q_w , R_v and S_{wv} . The filter can be either in the one-step predicting form,

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_k + K_p(y_k - C\hat{x}_{k|k-1}), \quad (3)$$

or in the filtering form,

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_f(y_k - C\hat{x}_{k|k-1}), \quad (4)$$

TABLE I

DEFINITIONS OF SYMBOLS IN (8) FOR EACH KALMAN FILTER FORM

Symbol	Predicting Form	Filtering Form
e_k	$y_k - C\hat{x}_{k k-1}$	$y_k - C\hat{x}_{k k}$
\bar{A}	$A - K_p C$	$A - K_p C$
\bar{G}	$[G \quad -K_p]$	$[G \quad -K_p]$
\bar{C}	C	$C - CK_f C$
\bar{H}	I	$I - CK_f$

and the Kalman filter gains are defined as

$$K_p = (AP_p C^T + GS_{wv})(CP_p C^T + R_v)^{-1} \quad (5)$$

$$K_f = P_p C^T (CP_p C^T + R_v)^{-1} \quad (6)$$

and P_p is the covariance of the state prediction error, $\tilde{x}_{k|k-1} = x_k - \hat{x}_{k|k-1}$. The covariance $P_p = E[\tilde{x}_{k|k-1}\tilde{x}_{k|k-1}^T]$ is obtained as the solution to the Riccati equation

$$\begin{aligned} P_p &= AP_p A^T + GQ_w G^T \\ &\quad - (AP_p C^T + GS_{wv})(CP_p C^T + R_v)^{-1}(CP_p A^T + S_{wv}^T G^T). \end{aligned} \quad (7)$$

A general state-space model of the measurement prediction/estimate error can be defined,

$$\begin{aligned} \tilde{x}_{k+1|k} &= \bar{A}\tilde{x}_{k|k-1} + \bar{G}\bar{w}_k \\ e_k &= \bar{C}\tilde{x}_{k|k-1} + \bar{H}v_k \end{aligned} \quad (8)$$

where e_k and the system matrices have different definitions, as shown in Table I, depending on which form of the filter is used. The noise \bar{w}_k in (8) is defined as

$$\bar{w}_k = \begin{bmatrix} w_k \\ v_k \end{bmatrix}$$

with properties

$$\begin{bmatrix} \bar{w}_k \\ v_k \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \bar{Q}_w & \bar{S}_{wv} \\ \bar{S}_{wv}^T & R_v \end{bmatrix} \right) \quad (9)$$

where

$$\bar{Q}_w = E[\bar{w}_k \bar{w}_k^T] = \begin{bmatrix} Q_w & S_{wv} \\ S_{wv}^T & R_v \end{bmatrix}, \quad \bar{S}_{wv} = E[\bar{w}_k v_k^T] = \begin{bmatrix} S_{wv} \\ R_v \end{bmatrix}.$$

The autocovariance of the measurement prediction or estimate error is given by

$$R_{e,0} = E[e_k e_k^T] = \bar{C} P_p \bar{C}^T + \bar{H} R_v \bar{H}^T \quad (10)$$

$$R_{e,j} = E[e_{k+j} e_k^T] = \bar{C} \bar{A}^j P_p \bar{C}^T + \bar{C} \bar{A}^{j-1} G S_{wv} \bar{H}^T - \bar{C} \bar{A}^{j-1} K_p R_v \bar{H}^T, \quad j \geq 1 \quad (11)$$

The autocovariance matrix is defined as

$$\mathcal{R}_e(L) = \begin{bmatrix} R_{e,0} & R_{e,1}^T & \cdots & R_{e,L-1}^T \\ R_{e,1} & R_{e,0} & & R_{e,L-2}^T \\ \vdots & & \ddots & \vdots \\ R_{e,L-1} & R_{e,L-2} & \cdots & R_{e,0} \end{bmatrix} \quad (12)$$

and can be written as

$$\begin{aligned} \mathcal{R}_e(L) = & O P_p O^T + Z \left[\bigoplus_{i=1}^L Q_w \right] Z^T \\ & + Z \left[\bigoplus_{i=1}^L S_{wv} \right] \Psi^T + \Psi \left[\bigoplus_{i=1}^L S_{wv}^T \right] Z^T \\ & + \Psi \left[\bigoplus_{i=1}^L R_v \right] \Psi^T, \end{aligned} \quad (13)$$

where

$$Z = \Gamma \left[\bigoplus_{i=1}^L G \right], \quad \Psi = \Gamma \left[\bigoplus_{i=1}^L (-K_p) \right] + \bigoplus_{i=1}^L \bar{H}$$

and

$$O = \begin{bmatrix} \bar{C} \\ \bar{C} \bar{A} \\ \vdots \\ \bar{C} \bar{A}^{L-1} \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 \\ \bar{C} & & 0 & 0 & 0 \\ \bar{C} \bar{A} & & 0 & 0 & 0 \\ \vdots & \ddots & & & \vdots \\ \bar{C} \bar{A}^{L-2} & \cdots & \bar{C} \bar{A} & \bar{C} & 0 \end{bmatrix}.$$

We apply the vec operator to (13) in order to state the problem as a linear least-squares problem. The vec operator performs stacking of the matrix columns to form a column matrix [3]. By applying the rules for the vec operator, we write the Lyapunov equation for P_p in stacked form, with the subscript s used as shorthand for the vec operator, i.e. $\text{vec}(A) = A_s$.

$$\begin{aligned} \text{vec}(P_p) &= \text{vec}(\bar{A} P_p \bar{A}^T) + \text{vec}(\bar{G} \bar{Q}_w \bar{G}^T) \\ &= (I_{n_x^2} - \bar{A} \otimes \bar{A})^{-1} (\bar{G} \bar{Q}_w \bar{G}^T)_s \end{aligned} \quad (14)$$

We introduce three permutation matrices. For an $m \times n$ matrix A we define a permutation matrix $U_{m,n,L}$, which is an $mnL^2 \times mn$ matrix of zeros and ones satisfying

$$\text{vec} \left(\bigoplus_{i=1}^L A \right) = U_{m,n,L} \text{vec}(A). \quad (15)$$

For a square matrix of size $p \times p$ we have the permutation matrix $U_{p,L} = U_{p,p,L}$. Finally, there is the vec-permutation matrix (or commutation matrix) $T_{m,n}$, such that for an $m \times n$ matrix A [3], [4],

$$\text{vec}(A^T) = T_{n,m} \text{vec}(A). \quad (16)$$

Applying the vec operator to (13) yields

$$\begin{aligned} (\mathcal{R}_e(L))_s = & [(Z \otimes Z) U_{n_w, L} + D(G \otimes G)] (Q_w)_s \\ & + [(\Psi \otimes \Psi) U_{n_y, L} + D(K_p \otimes K_p)] (R_v)_s \\ & + [(\Psi \otimes Z) U_{n_w, n_y, L} - D(I_{n_x^2} + T_{n_x, n_x}) (K_p \otimes G) \\ & + (Z \otimes \Psi) U_{n_y, n_w, L} T_{n_y, n_w}] (S_{wv})_s \end{aligned} \quad (17)$$

in which

$$D = (O \otimes O) (I_{n_x^2} - \bar{A} \otimes \bar{A})^{-1}. \quad (18)$$

Given a sequence of data $\{e_i\}_{i=1}^{N_d}$, the estimate of the autocovariance can be computed by

$$\hat{R}_{e,j} = \frac{1}{N_d - j} \sum_{i=1}^{N_d-j} e_{i+j} e_i^T, \quad (19)$$

where N_d is the length of the data sequence. The estimated autocovariance matrix $\hat{\mathcal{R}}_e(L)$ can be formed analogously to (12) using the estimates (19). Now (17) can be written in the form of a linear least-squares problem

$$\Phi = \min_{Q_w, S_{wv}, R_v} \left\| \begin{bmatrix} (Q_w)_s \\ (S_{wv})_s \\ (R_v)_s \end{bmatrix} - (\hat{\mathcal{R}}_e(L))_s \right\|_2^2 \quad (20)$$

where additional constraints may be necessary in order to ensure positive semidefiniteness of the covariance matrices. As noted in [1], a short data sequence or significant model error may result in covariance estimates that are not positive definite. This problem can be remedied by stating (20) as a convex semidefinite programming problem [1].

The optimal Kalman filter gain can then be computed from the estimated covariances by (5) or (6) after solving the Riccati equation (7).

Note that for the case with a predicting Kalman filter and with $S_{wv} = 0$, (17) takes the form presented in [1].

III. CONCLUSIONS AND FUTURE WORKS

A. Conclusions

A generalization of the autocovariance least-squares method by Odelson et al. has been presented. The generalization is applicable to systems with mutually correlated disturbances and also works with data generated by the filtering form of the Kalman filter.

B. Future Works

Solution methods for the constrained least-squares problem will be investigated. The estimation method will be applied to realistic examples.

REFERENCES

- [1] B. J. Odelson, M. R. Rajamani, and J. B. Rawlings, "A new autocovariance least-squares method for estimating noise covariances," *Automatica*, vol. 42, no. 2, pp. 303–308, 2006.
- [2] R. K. Mehra, "On the identification of variances and adaptive kalman filtering," *IEEE Transactions on Automatic Control*, vol. AC-15, no. 12, pp. 175–184, 1970.
- [3] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Transactions on Circuits and Systems*, vol. 25, no. 9, pp. 772–781, 1978.
- [4] J. R. Magnus and H. Neudecker, "The commutation matrix: some properties and applications," *The Annals of Statistics*, vol. 7, no. 2, pp. 381–394, 1979.