A Generalized Autocovariance Least-Squares Method for Covariance Estimation

Åkesson, Bernt Magnus; Jørgensen, John Bagterp; Poulsen, Niels Kjølstad; Jørgensen, Sten Bay

Published in:
American Control Conference 2007

Link to article, DOI:
10.1109/ACC.2007.4282878

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
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 AUTOCOVARIANCE LEAST-SQUARES ESTIMATION

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

\[ x_{k+1} = Ax_k + Bu_k + Gw_k \]
\[ y_k = Cx_k + v_k \]

where \( A \in \mathbb{R}^{n_x \times n_x} \), \( B \in \mathbb{R}^{n_x \times n_u} \), \( G \in \mathbb{R}^{n_x \times n_v} \), 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

\[ \begin{bmatrix} w_k \\ v_k \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q_w & S_{wv} \\ S_{wv}^T & R_v \end{bmatrix} \right) \] (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,

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

or in the filtering form,

\[ \hat{x}_{k|k} = \hat{x}_{k|k-1} + K_f(y_k - C\hat{x}_{k|k-1}), \]

and the Kalman filter gains are defined as

\[ K_p = (AP_pC^T + GS_{wv})(CP_pC^T + R_v)^{-1} \]
\[ K_f = P_pC^T(CP_pC^T + R_v)^{-1} \] (5) (6)

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

\[ P_p = AP_pA^T + GQ_wG^T \]
\[ - (AP_pC^T + GS_{wv})(CP_pC^T + R_v)^{-1}(CP_pA^T + S_{wv}^T G^T) \] (7)

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

\[ \tilde{x}_{k+1|k} = A\tilde{x}_{k|k-1} + G\tilde{w}_k \]
\[ e_k = C\tilde{x}_{k|k-1} + \tilde{H}v_k \] (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 \( \tilde{w}_k \) in (8) is defined as

\[ \tilde{w}_k = \begin{bmatrix} w_k \\ v_k \end{bmatrix} \]

with properties

\[ \begin{bmatrix} \tilde{w}_k \\ v_k \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tilde{Q}_w & \tilde{S}_{wv} \\ \tilde{S}_{wv}^T & \tilde{R}_v \end{bmatrix} \right) \] (9)

where

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

1-4244-0989-6/07/$25.00 ©2007 IEEE.

Table I

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Predicting Form</th>
<th>Filtering Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e_k )</td>
<td>( y_k - C\hat{x}_{k</td>
<td>k-1} )</td>
</tr>
<tr>
<td>( \hat{A} )</td>
<td>( A - K_pC )</td>
<td>( A - K_fC )</td>
</tr>
<tr>
<td>( \hat{C} )</td>
<td>( [G - K_p] )</td>
<td>( [G - K_f] )</td>
</tr>
<tr>
<td>( \hat{R} )</td>
<td>( I )</td>
<td>( I - CK_f )</td>
</tr>
</tbody>
</table>
The autocovariance of the measurement prediction or estimate error is given by
\[ R_{e,0} = E[e_k e^T_k] = \bar{C}\bar{P}_p\bar{C}^T + \bar{C}\bar{A}\bar{L}^{-1}G\bar{S}_{wv}\bar{H}^T - \bar{C}\bar{A}^{-1}K_p\bar{R}_p\bar{H}^T, \quad j \geq 1 \] (11)

The autocovariance matrix is defined as
\[ \mathcal{R}(L) = \begin{bmatrix} R_{e,0} & R_{e,1} & \cdots & R_{e,L-1} \\ R_{e,1} & R_{e,0} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ R_{e,L-1} & R_{e,L-2} & \cdots & R_{e,0} \end{bmatrix} \] (12)

and can be written as
\[
\mathcal{R}(L) = OP_pO^T + Z \left[ \bigoplus_{i=1}^{L} Q_w \right] Z^T \\
+ Z \left[ \bigoplus_{i=1}^{L} S_{wv} \right] \Psi^T + \Psi \left[ \bigoplus_{i=1}^{L} S_{wv}^T \right] Z^T \\
+ \Psi \left[ \bigoplus_{i=1}^{L} R_{v} \right] \Psi^T,
\] (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
\[ O = \begin{bmatrix} \bar{C} \\ \bar{C}\bar{A} \\ \vdots \\ \bar{C}\bar{A}^{L-1} \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \bar{C} & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \bar{C}\bar{A}^{L-2} & \cdots & \bar{C}\bar{A} & 0 \end{bmatrix}. \]

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.
\[ \text{vec}(A) = A_s. \]
\[ \text{vec}(P_p) = \text{vec}(\bar{C}P_p\bar{A}^T) + \text{vec}(\bar{C}\bar{Q}_wG^T) = (I_{n_d} - \bar{A} \otimes \bar{A})^{-1}(\bar{G}\bar{Q}_wG^T) \] (14)

We introduce three permutation matrices. For an \( mnL^2 \times mn \) matrix \( A \) we define a permutation matrix \( U_{m,n,L} \) which is an \( mnL^2 \times mn \) matrix of zeros and ones satisfying
\[ \text{vec} \left( \bigoplus_{i=1}^{L} A \right) = U_{m,n,L} \text{vec} (A). \] (15)

For a square matrix of size \( p \times p \) we have the permutation matrix \( U_{p,p} = U_{p,p,L} \). Finally, there is the vec-permutation matrix (or commutation matrix) \( T_{m,n} \), such that for an \( m \times n \) matrix \( A \) [3], [4],
\[ \text{vec} (A^T) = T_{m,n} \text{vec} (A). \] (16)

Applying the vec operator to (13) yields
\[
\begin{align*}
\mathcal{R}(L) &= ([Z \otimes Z] U_{n_d,L} + D(G \otimes G)(Q_w)_s \\
&+ [([\Psi \otimes \Psi] U_{m,n,L} + D(K_p \otimes K_p))(R_e)_s \\
&+ [([\Psi \otimes Z] U_{n_d,n_d,L} - D(I_{n_d}^2 + T_{m,n})(K_p \otimes G) \\
&+ (Z \otimes \Psi) U_{m,n_d,L} T_{m,n})(S_{wv})_s \quad (17)
\end{align*}
\]
in which
\[ D = (O \otimes O)(I_{n_d^2} - \bar{A} \otimes \bar{A})^{-1}. \] (18)

Given a sequence of data \( \{ e_{l,j} \}_{l=1}^{N_d} \), the estimate of the autocovariance can be computed by
\[ \hat{R}_{e,j} = \frac{1}{N_d - j} \sum_{i=1}^{N_d-j} e_{i+j} e_i^T, \] (19)
where \( N_d \) is the length of the data sequence. The estimated autocovariance matrix \( \hat{R}(L) \) can be formed analogously to (12) using the estimates (19). Now (17) can be written in the form of a linear least-squares problem
\[ \Phi = \min_{Q_w,S_{wv},R_e} \left\| \begin{bmatrix} (Q_w)_s \\
(S_{wv})_s \\
(R_e)_s \end{bmatrix} - (\hat{R}(L))_s \right\|^2 \] (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