Covariance Matching Kalman Filter for observable LTI Systems

Rahul Moghe, Renato Zanetti and Maruthi R. Akella

Aerospace Engineering and Engineering Mechanics
The University of Texas at Austin

rahul.moghe@utexas.edu, renato@utexas.edu, makella@mail.utexas.edu

IEEE Conference on Decision and Control, Miami, Florida, USA

December 19, 2018

Overview

Motivation

System Description

$$x_k = Fx_{k-1} + w_{k-1}, \quad w_{k-1} \sim \mathcal{N}(0, Q),\ x_0 \sim \mathcal{N}(m_x, P_0)$$

$$y_k = Hx_k + v_k, \quad v_k \sim \mathcal{N}(0, R)$$

Common issues with a linear Kalman Filter:

- Q and R matrices are inaccurately known
- F and H matrices are uncertain
- Either w_k or v_k are non Gaussian

Adaptive filters simultaneously estimate the state and the unknown quantities on-line.

Literature Review

Existing adaptive filters:

- Application specific algorithms
- Restrictive assumptions on F and H matrix
- Interdependence between the state and covariance estimate

Literature Review

Existing adaptive filters:

- Application specific algorithms
- Restrictive assumptions on F and H matrix
- Interdependence between the state and covariance estimate

Important properties of our algorithm:

- Either one of the Q or R matrices are completely known
- Does not need the state estimate for covariance estimation
- Assumes observability of the LTI system
- We provide a proof of convergence

Stacked Dynamics

Propagating the state dynamics in time,

$$\begin{aligned} x_{k+1} &= Fx_k + w_k \\ x_{k+2} &= F^2 x_k + F w_k + w_{k+1} \\ x_{k+n-1} &= F^{n-1} x_k + F^{n-2} w_k + \cdots + w_{k+n-1} \end{aligned}$$

Stacked Dynamics

Propagating the state dynamics in time,

$$
\begin{aligned}
x_{k+1} &= Fx_k + w_k \\
x_{k+2} &= F^2 x_k + Fw_k + w_{k+1} \\
x_{k+n-1} &= F^{n-1} x_k + F^{n-2} w_k + \cdots + w_{k+n-1}
\end{aligned}
$$

Predicting future measurements from the current state,

$$
\begin{aligned}
y_k &= Hx_k + v_k \\
y_{k+1} &= HFx_k + Hw_k + v_{k+1} \\
y_{k+2} &= HF^2 x_k + HFw_k + Hw_{k+1} + v_{k+2} \\
y_{k+n-1} &= HF^{n-1} x_k + HF^{n-2} w_k + Hw_{k+n-2} + v_{k+n-1}
\end{aligned}
$$

Coagulating all the measurements

The observability matrix M_o, formed below has full column rank.

$$\underbrace{\begin{bmatrix} y_{k+n-1} \\ y_{k+n-2} \\ \vdots \\ y_k \end{bmatrix}}_{\triangleq \mathcal{Y}_k} = \underbrace{\begin{bmatrix} HF^{n-1} \\ HF^{n-2} \\ \vdots \\ H \end{bmatrix}}_{\triangleq M_o} x_k + \underbrace{\begin{bmatrix} H & HF & HF^2 & \cdots & HF^{n-2} \\ \mathbf{0} & H & HF & \cdots & HF^{n-3} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & H \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}}_{\triangleq M_w} \underbrace{\begin{bmatrix} w_{k+n-2} \\ w_{k+n-3} \\ \vdots \\ w_k \end{bmatrix}}_{\triangleq W_k}$$

$$+ \underbrace{\begin{bmatrix} v_{k+n-1}^T & v_{k+n-2}^T & \cdots & v_k^T \end{bmatrix}^T}_{\triangleq V_k}$$

Time Series

Propagating the stacked measurements in time,

$$\mathcal{Y}_k = M_o x_k + M_w W_k + V_k$$

Time Series

Propagating the stacked measurements in time,

$$\mathcal{Y}_k = M_o x_k + M_w W_k + V_k$$
$$\mathcal{Y}_{k+1} = M_o x_{k+1} + M_w W_{k+1} + V_{k+1}$$

Time Series

Propagating the stacked measurements in time,

$$\mathcal{Y}_k = M_o x_k + M_w W_k + V_k$$
$$\mathcal{Y}_{k+1} = M_o x_{k+1} + M_w W_{k+1} + V_{k+1}$$

Eliminating the state,

$$FM_o^\dagger \mathcal{Y}_k = Fx_k + FM_o^\dagger M_w W_k + FM_o^\dagger V_k$$
$$M_o^\dagger \mathcal{Y}_{k+1} = Fx_k + w_k + M_o^\dagger M_w W_{k+1} + M_o^\dagger V_{k+1}$$

Time Series

Propagating the stacked measurements in time,

$$\mathcal{Y}_k = M_o x_k + M_w W_k + V_k$$
$$\mathcal{Y}_{k+1} = M_o x_{k+1} + M_w W_{k+1} + V_{k+1}$$

Eliminating the state,

$$FM_o^{\dagger}\mathcal{Y}_k = Fx_k + FM_o^{\dagger}M_w W_k + FM_o^{\dagger}V_k$$
$$M_o^{\dagger}\mathcal{Y}_{k+1} = Fx_k + w_k + M_o^{\dagger}M_w W_{k+1} + M_o^{\dagger}V_{k+1}$$

State independent linear strictly stationary time series,

$$\underbrace{M_o^{\dagger}\mathcal{Y}_{k+1} - FM_o^{\dagger}\mathcal{Y}_k}_{\triangleq \mathcal{Z}_k} = \underbrace{w_k + M_o^{\dagger}M_w W_{k+1} - FM_o^{\dagger}M_w W_k}_{\triangleq \mathcal{W}_k} +$$
$$+ \underbrace{M_o^{\dagger}V_{k+1} - FM_o^{\dagger}V_k}_{\triangleq \mathcal{V}_k}$$

Covariance Estimation

$\mathcal{Z}_k$ is zero mean and has a constant covariance

$$\mathcal{Z}_k = N_0 y_k + N_1 y_{k+1} + \cdots + N_n y_{k+n} =$$
$$A_1 w_k + A_2 w_{k+1} + \cdots + A_{n-1} w_{k+n-1} + B_0 v_k + B_1 v_{k+1} + \cdots + B_n v_{k+n}$$

Covariance Estimation

$\mathcal{Z}_k$ is zero mean and has a constant covariance

$$\mathcal{Z}_k = N_0 y_k + N_1 y_{k+1} + \cdots + N_n y_{k+n} =$$
$$A_1 w_k + A_2 w_{k+1} + \cdots + A_{n-1} w_{k+n-1} + B_0 v_k + B_1 v_{k+1} + \cdots + B_n v_{k+n}$$

Covariance calculation

$$Cov(\mathcal{Z}_k) = A_1 Q A_1^T + \cdots + A_{n-1} Q A_{n-1}^T + B_0 R B_0^T + \cdots + B_n R B_n^T$$
$$Cov(\mathcal{Z})_i = \sum_{k=1}^{k=i} \frac{1}{k} \bar{\mathcal{Z}}_k \bar{\mathcal{Z}}_k^T$$
$$Cov(\mathcal{Z})_i - A_1 Q A_1^T - \cdots - A_{n-1} Q A_{n-1}^T = B_0 \hat{R}_i B_0^T + \cdots + B_n \hat{R}_i B_n^T$$

Q or R Estimation

Vectorizing the equation,

$$vec(C_k) = \underbrace{(B_0 \otimes B_0 + \cdots B_n \otimes B_n)}_{\triangleq S} vec(\hat{R}_k)$$

$$vec(D_k) = \underbrace{(A_1 \otimes A_1 + \cdots A_n \otimes A_n)}_{\triangleq T} vec(\hat{Q}_k)$$

Repetitive elements of the symmetric matrix and corresponding columns of S and T matrix are averaged out.

Outline of Proof

Time Series

$$\mathcal{Z}_k = N_0 y_k + N_1 y_{k+1} + \cdots + N_n y_{k+n} =$$
$$A_1 w_k + A_2 w_{k+1} + \cdots + A_{n-1} w_{k+n-1} + B_0 v_k + B_1 v_{k+1} + \cdots + B_n v_{k+n}$$

- We prove that $Cov(\mathcal{Z})_k \xrightarrow{\mathcal{P}} Cov(\mathcal{Z})$ using the CLT for time series with decaying autocovariance
- Then, $\hat{R}_k \xrightarrow{\mathcal{P}} R$ follows from the full rank assumption

State Error Covariance Convergence

- Formulate the following 3 Covariance matrices:

$$\hat{P}_{k+1} = \hat{\bar{F}}_k \hat{P}_k \hat{\bar{F}}_k^T + \hat{K}_k \hat{R}_k \hat{K}_k^T + Q$$

$$P_{k+1} = \hat{\bar{F}}_k P_k \hat{\bar{F}}_k^T + \hat{K}_k R \hat{K}_k^T + Q$$

$$P_{k+1,opt} = \bar{F}_k P_{k,opt} \bar{F}_k^T + K_k R K_k^T + Q$$

- We prove that $\hat{P}_k \xrightarrow{\mathcal{P}} P_k$
- We prove that $P_k \xrightarrow{\mathcal{P}} P_{k,opt}$

Simulation Problem

A marginally stable system with $T = 0.05$,

$$x_k = \begin{bmatrix} 1 & T & 0.5T^2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix} x_{k-1} + w_{k-1}$$

$$y_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} x_k + v_k$$

We use the same example to estimate R and then Q

$$R = \begin{bmatrix} 0.25 & 0 \\ 0 & 0.25 \end{bmatrix} \text{ and } Q = \begin{bmatrix} 0.25 & 0.04 & 0.04 \\ 0.04 & 0.25 & 0.04 \\ 0.04 & 0.04 & 0.25 \end{bmatrix}$$

Simulation for covariance matrices

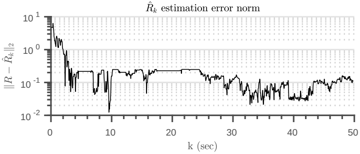

Figure: R estimate vs. time

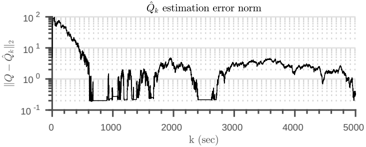

Figure: Q estimate vs. time

The Q matrix was assumed to be a constant

The R matrix was assumed to be a constant

In both the cases, the initial estimate of the covariance was 5I with suitable dimensions.

Simulation for state estimation (unknown R case)

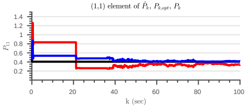

Figure: P_{11} vs. t

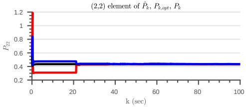

Figure: P_{22} vs. t

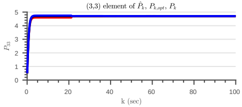

Figure: P_{33} vs. t

Simulation for state estimation (unknown Q case)

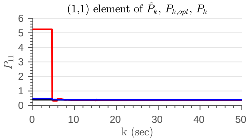

Figure: P_{11} vs. t

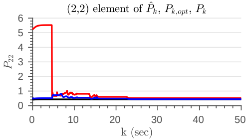

Figure: P_{22} vs. t

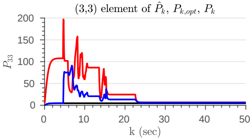

Figure: P_{33} vs. t

Discussion

- Linearly dependent measurements (non-invertible S matrix) introduce ambiguity in R matrix estimation
- Methods like linear matrix inequalities and weighting the initial estimate can be used for covariance estimation
- The rate of convergence depends on difference in order of magnitude of the known and the unknown covariance

$$Cov(\mathcal{Z}_k) = (\text{terms with R}) + (\text{terms with Q})$$