The Kalman filter and the square root Kalman filter using only QR decompositions

Koichi (Koiti) Yano

Introduction

In this note, I outline the algorithms for the Kalman filter and the square root Kalman filter using only QR decompositions, which is proposed by Tracy (2022). The Kalman filter (Kalman (1960)) is a recursive algorithm that estimates the state vector of a linear Gaussian state space model. The square root Kalman filter using only QR decompositions is an alternative to the plain-vanilla Kalman filter. The square root Kalman filter is numerically more stable than the Kalman filter (See Anderson and Moore (1979)).

The Kalman filter and the square root Kalman filter using only QR decompositions

A linear Gaussian state space model

A (invariant) linear Gaussian state space model is defined by the following two equations.

$$x_t = Fx_{t-1} + Eu_t + w_t$$

and

$$y_t = Hx_t + v_t,$$

where x_t is the $(k \times 1)$ state vector , y_t is the $(l \times 1)$ observation vector, u_t is the $(n \times 1)$ exogenous vector, v_t is the $(k \times 1)$ state noise, $w(t)$ is the $(l \times 1)$ observation noise, t is a time index. The matrices F, E, and H are $(k \times k)$, $(k \times n)$, and $(l \times k)$, respectively. The system noise w_t and the observation noise v_t are sampled from $N(0, W)$ and $N(0, V)$, respectively, where W is a $(k \times k)$ covariance matrix and V is an $(l \times l)$ covariance matrix. The Kalman Filter is used to estimate the state vector x_t given the observations y_t.

The aims of the problem obtain the estimate $x_{t|t}$ of the state vector x_t and the estimate P_t of the covariance matrix of the x_t at time t. The two estimates are defined as follows:

$$x_{t|t} = E[x_t | y_1, y_2, \dots, y_t]$$

$$P_{t|t} = \text{cov}\big[(x_t - x_{t|t})\big] = \text{E}\big[(x_t - x_{t|t})(x_t - x_{t|t})^t\big]$$

The Kalman filter

The Kalman filter is implemented using the following recursion:

1. Prediction step:
    1. $x_{t|t-1} = F x_{t-1|t-1} + E u_t$
    2. $P_{t|t-1} = F P_{t-1|t-1} F^t + V$
2. Innovation step:
    1. $e_t = y_t - H x_{t|t-1}$
    2. $s_t = H P_{t-1|t-1} H^t + W$
3. Update step:
    1. $K_t = P_{t|t-1} H^t s_t^{-1}$
    2. $x_{t|t} = x_{t|t-1} + K_t e_t$
    3. $P_{t|t} = (I - K_t H) P_{t|t-1}$

where $x_{t|t}$ is the estimated state vector at, $P_{t|t}$ is the estimated state covariance matrix, and K_t is the Kalman gain at time t.

The outline of derivation of $P_{t|t}$ and K_t

$$
\begin{aligned}
P_{t|t} &= \mathrm{cov}[x_t - x_{t|t}] \\
P_{t|t} &= \mathrm{cov}[x_t - (x_{t|t-1} + K_t e_t)] \\
P_{t|t} &= \mathrm{cov}[x_t - (x_{t|t-1} + K_t(y_t - Hx_{t|t-1}))] \\
P_{t|t} &= \mathrm{cov}[x_t - (x_{t|t-1} + K_t(Hx_t + v_t - Hx_{t|t-1}))] \\
P_{t|t} &= \mathrm{cov}[(I - K_t H)(x_t - x_{t|t-1}) - K_t v_t)] \\
P_{t|t} &= (I - K_t H)\mathrm{cov}[(x_t - x_{t|t-1})](I - K_t H)^t + K_t \mathrm{cov}[v_t] K_t^t \\
P_{t|t} &= (I - K_t H) P_{t|t-1} (I - K_t H)^t + K_t R_t K_t^t,
\end{aligned}
$$

where $R_t = \mathrm{cov}[v_t]$.

$$
\begin{aligned}
&P_{t|t} = P_{t|t-1} - K_t H P_{t|t-1} - P_{t|t-1} H^t {K_t}^t + K_t R K_t^t \\
&\frac{\partial tr(P_{t|t})}{\partial K_t} = -2(HP_{t|t-1})^t + 2K_t S_t = 0 \\
&K_t = P_{t|t-1} H^t S_t^{-1},
\end{aligned}
$$

See Anderson and Moore (1979) and Kitagawa (2010) for more details.

The square root Kalman filter using only QR decompositions

The square root Kalman filter is an alternative to the Kalman filter that is numerically more stable. Tracy (2022) proposes the following recursion for The square root Kalman filter using only QR decompositions (hereafter the QR Kalman filter).

The QR Kalman filter is implemented using the following recursion

1. Prediction step
    1. $x_{t|t-1} = Fx_{t-1|t-1} + Eu_t$
    2. $\Sigma_{t|t-1} = gr_r(\Sigma_{t-1|t-1} F^t, \Gamma_v)$
2. Innovation step:
    1. $e_t = y_t - Hx_{t|t-1}$
    2. $G_t = gr_r(\Sigma_{t|t-1} H^t, \Gamma_w)$

3. Update step:

    1. $K_t = [G_t^{-1}(G_t^{-t}H)(\Sigma_{t|t-1})^t(\Sigma_{t|t-1})]^t$
    2. $x_{t|t} = x_{t|t-1} + K_t e_t$
    3. $\Sigma_{t|t} = gr_r(\Sigma_{t|t-1}(I - K_t H)^t, \Gamma_w K_t^{\ t})$,

where $\Sigma_{t|t} = \sqrt{P_{t|t}}$, the Cholesky decompositions of the covariance matrices $P_{t|t}$. The function gr_r returns the matrix R of the QR decomposition of the matrix. The matrices Γ_v and Γ_w are the Cholesky decompositions of V and W, respectively. See Tracy (2022) for more details.

Running Code

See "run_kf_qr_1.R" and "run_kf_qr_2.R" in "tools/".

To be added:

```
1 + 1
```

```
[1] 2
```