LV 11.4500 – UBUNG 7

U7.1 In this exercise we will study the matrix inversion used to derive the Kalman gain on page 23, Lecture 14:

$$C_n = (\hat{C}_n^{-1} + H^T\Gamma^{-1}H)^{-1} = \hat{C}_n - \underbrace{\hat{C}_n H^T (H\hat{C}_n H^T + \Gamma)^{-1}}_{=:K_n} H\hat{C}_n \tag{1}$$

where $\Gamma \in \mathbb{R}^{k\times k}$, $H \in \mathbb{R}^{k\times d}$, $\hat{C}_n \in \mathbb{R}^{d\times d}$ and $\Gamma, \hat{C}_n > 0$.

a) Verify that $\hat{C}_n^{-1} + H^T\Gamma^{-1}H$ is positive definite (and thus invertible).

b) Verify that the right-inverse property of (1) holds

$$(\hat{C}_n^{-1} + H^T\Gamma^{-1}H)(\hat{C}_n - K_n H\hat{C}_n) = I$$

Hint: make use of the obvious identity

$$(H\hat{C}_n H^T + \Gamma)(H\hat{C}_n H^T + \Gamma)^{-1} = I$$

c)Consider the following matrix equality

$$\begin{bmatrix} -\Gamma & H \\ H^T & \hat{C}_n^{-1} \end{bmatrix} = \begin{bmatrix} -\Gamma & 0 \\ H^T & I \end{bmatrix} \begin{bmatrix} I & -\Gamma^{-1}H \\ 0 & \hat{C}_n^{-1} + H^T\Gamma^{-1}H \end{bmatrix},$$

Using properties of determinants of block upper and lower matrices, verify that the left-hand side matrix is invertible.

Next, considering the inverse

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} := \begin{bmatrix} -\Gamma & H \\ H^T & \hat{C}_n^{-1} \end{bmatrix}^{-1}$$

verify that $B_{22} = (\hat{C}_n^{-1} + H^T\Gamma^{-1}H)^{-1}$ for for instance by using

$$\begin{aligned} B_{21}H + B_{22}\hat{C}_n^{-1} &= I_d \\ -B_{21}\Gamma + B_{22}H^T &= 0_k. \end{aligned} \tag{2}$$

Last part of exercise: verify (1)**.**

Hint: First determine B_{11} and B_{21} from a system of equations, and use the alternative form of B_{21} in (2)

d) Show that

$$C_n^{-1} m_n = \hat{C}_n^{-1}\hat{m}_n + H^T\Gamma^{-1}y_n.$$

implies that

$$m_n = (I - K_n H)\hat{m}_n + K_n y_n.$$

Hint: $H\hat{C}_n H^T = H\hat{C}_n H^T + \Gamma - \Gamma$

U7.2 Consider the linear-Gaussian setting

$$V_{j+1} = AV_j + \xi_j, \qquad j = 0, 1, \dots$$
$$V_0 \sim N(m_0, C_0)$$

with $\xi_j \overset{iid}{\sim} N(0, \Sigma)$, and observations

$$Y_j = HV_j + \eta_j, \quad j = 1, 2, \dots$$

with $\eta_j \overset{iid}{\sim} N(0, \Gamma)$ and the usual independence assumptions $V_0 \perp \{\xi_j\} \perp \{\eta_j\}$.

In order to derive the Kalman filtering predict-analysis steps, one does not require generally that both Σ and C_0 are positive definite, as we assumed in Lecture 14. When A is a non-singular matrix, for instance, it suffices that only C_0 is positive definite.

With this in mind, derive the Kalman predict-analysis formulas for the dynamics on $\mathbb{R}$ with $A = -1$, $C_0 = 10$, $\Sigma = 0$, $H = 1$ and $\Gamma = 2$.

How many iterations are needed to ensure that $\mathbb{E}\left[\,(m_j - V_j)^2 | Y_{1:j} = y_{1:j}\right] < 10^{-3}$? Same question when $C_0 = 10^8$? Motivate the answer.

U7.3 In Lecture 14, extended Kalman filtering was presented for nonlinear mappings Ψ and linear observation mappings $h(v) = Hv$. In this exercise we will extend ExKF to settings where also h may be nonlinear. Initial condition $V_0 \sim N(m_0, C_0)$ and for $j = 0, 1, \dots$

$$\begin{aligned} V_{j+1} &= \Psi(V_j) + \xi_j, \\ Y_{j+1} &= h(V_{j+1}) + \eta_{j+1}, \end{aligned} \tag{3}$$

and Gaussian noise with the usual distribution and independence assumptions.

a) Given ExKF moments (m_j, C_j) we recall that ExKF proceeds by linearizing

$$\Psi_L(v; m_j) = \Psi(m_j) + D\Psi(m_j)(v - m_j)$$

and applying the "Kalman-filter-like" prediction step:

$$\hat{m}_{j+1} = \Psi_L(m_j) = \Psi(m_j)$$
$$\hat{C}_{j+1} = D\Psi(m_j) C_j D\Psi(m_j)^T + \Sigma.$$

We extend the analysis step of ExKF to settings with nonlinear observations by also linearing h:

$$h_L(v; \hat{m}_{j+1}) = h(\hat{m}_{j+1}) + Dh(\hat{m}_{j+1})(v - \hat{m}_{j+1}),$$

and applying a "Kalman filter-like" analysis step to determine the analysis moments. The updated mean can for instance be determined by the variational principle

$$m_{j+1} = \arg\min_{u \in \mathbb{R}^d} \frac{1}{2}|y_{j+1} - h_L(u; \hat{m}_{j+1})|^2_{\Gamma} + \frac{1}{2}|u - \hat{m}_{j+1}|^2_{\hat{C}_{j+1}}.$$

Task: Determine m_{j+1}, K_{j+1} and C_{j+1} for ExKF.

The calculations need not be done in detail, but motivate the results convincingly.

Hint: For K_{j+1} and C_{j+1} it may be helpful to observe that according to the KF approach, we are equating the first and second order terms of

$$\frac{1}{2}|(y_{j+1} - h(\hat{m}_{j+1}) + Dh(\hat{m}_{j+1})\hat{m}_{j+1}) - Dh(\hat{m}_{j+1})u|_\Gamma^2 + \frac{1}{2}|u - \hat{m}_{j+1}|_{\hat{C}_{j+1}}^2$$

to an ansatz Gaussian density exponent

$$\frac{1}{2}|u - m_{j+1}|_{C_{j+1}}^2.)$$

b) Summarize the predict-analysis steps for ExKF and extend the filtering approach with nonlinear observations to 3DVAR as well.

Remark: This extension may work quite well in settings with weak nonlinearities in h, but it is generally not robust. It turns out to be far easier to extend EnKF and particle filters to settings with nonlinear observations. See Jazwinski, "Stochastic processes and filtering theory" for more on ExKF with nonlinear observations.

U7.4 Consider the filtering problem

$$\begin{aligned} V_{j+1} &= \cos(V_j) + \xi_j \\ V_0 &\sim N(0,1) \end{aligned} \tag{4}$$

where $\xi_j \sim N(0, 0.25)$ and with nonlinear observations

$$Y_j = V_j + V_j^2/20 + \eta_j, \quad j = 1, 2, \ldots,$$

with $\eta_j \sim N(0,1)$.

a) Generate a sequence of $J = 10000$ observations by synthetic data $v_{1:J}^\dagger$, and implement a 3DVAR algorithm for the filtering problem following the extension to nonlinear observations in U7.3. You may assume that $m_0 = 0$ for 3DVAR, and study the effect of different values of $\hat{C}$ in terms of the time-averaged mean-squared tracking error

$$\frac{1}{J+1}\sum_{k=0}^{J} |v_k^\dagger - m_k|^2$$

b) Repeat the exercise for the ExKF method. You may then assume $(m_0, C_0) = (0, 1)$.