Final exam IIA2217 2023

Solution proposal 26.05.2023

Task 1

a)

Using (2) $x_k = y_k - e_k$ in (1) and putting $k := k-1$

$$\overbrace{y_k - e_k}^{x_k} = \theta_1 \big(\overbrace{y_{k-1} - e_{k-1}}^{x_{k-1}}\big) + \theta_2 u_{k-1} + \theta_1 e_{k-1}$$

$$\Downarrow$$

$$y_k = \theta_1 y_{k-1} + \theta_2 u_{k-1} + e_k$$

$$\Downarrow$$

$$y_k = \overbrace{\begin{bmatrix} y_{k-1} & u_{k-1} \end{bmatrix}}^{\varphi_k^T} \overbrace{\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}}^{\theta_0} + e_k$$

We have

$$\underline{\theta_0 = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}}, \quad \underline{\varphi_k = \begin{bmatrix} y_{k-1} \\ u_{k-1} \end{bmatrix}}$$

b)

- Prediction error $\varepsilon_k = y_k - \bar{y}_k$
  $\bar{y}_k = \varphi_k^T \theta$

- The OLS solution

$$\underline{\hat{\theta}_N = \left(\sum_{k=1}^{N} \varphi_k \Lambda \varphi_k^T\right)^{-1} \sum_{k=1}^{N} \varphi_k \Lambda y_k}$$

Proof: Solving

$$\frac{\partial V_N(\theta)}{\partial \theta} = -2 \frac{1}{N} \sum_{k=1}^{N} \varphi_k \Lambda \left(y_k - \varphi_k^T \theta\right) = 0$$

$$\Rightarrow \sum_{k=1}^{N} \varphi_k \Lambda y_k = \sum_{k=1}^{N} \varphi_k \Lambda \varphi_k^T \cdot \theta$$

$$\Rightarrow \hat{\theta}_N = \left(\sum_{k=1}^{N} \varphi_k \Lambda \varphi_k^T\right)^{-1} \sum_{k=1}^{N} \varphi_k \Lambda y_k$$

c) Consider

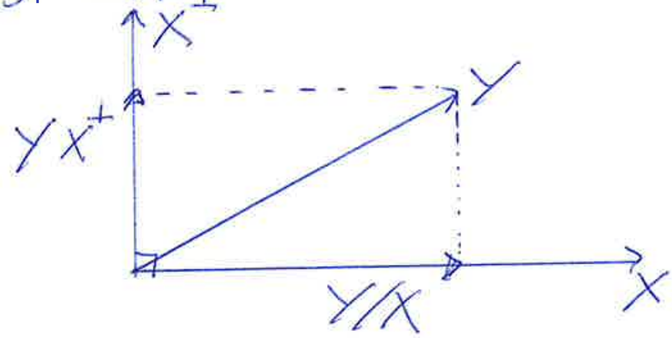
$$Y = \Theta X + E$$

- $Y/X \overset{\text{def}}{=} YX^T(XX^T)^{-1}X$

  $X^\perp = I_{K\times K} - X^T(XX^T)^{-1}X$

  $YX^\perp = Y - YX^T(XX^T)^{-1}X$

- Illustration



- Yes, $Y = Y/X + YX^\perp$ is correct

Proof

$$Y/X + YX^\perp = \cancel{YX^T(XX^T)^{-1}X} + Y - \cancel{YX^T(XX^T)^{-1}X}$$

$$= Y$$

d) $YX^T = OXX^T + EX^T$

$\Downarrow$

- $O_{OLS} = YX^T(XX^T)^{-1}$

We have used that $\frac{1}{N}EX^T \approx 0$ when N large

- Prediction

$$\bar{Y} = O_{OLS}X = YX^T(XX^T)^{+}X$$

- The prediction $\bar{Y} = O_{OLS}X = \hat{O}X = Y/X$

Hence, prediction $\bar{Y}$ equal Y/X, i.e., the prediction of Y onto X!

e)

- $$H_{21L} = \overbrace{\begin{bmatrix} H_2 & H_3 & H_4 \\ H_3 & H_4 & H_5 \end{bmatrix}}^{J=3 \text{ cols.}}, \quad H_{11L} = \overbrace{\begin{bmatrix} H_1 & H_2 & H_3 \\ H_2 & H_3 & H_4 \end{bmatrix}}^{J=3 \text{ columns}}$$

- $H_{11L} = O_L C_J$, $H_{21L} = O_L A C_J$

$J=3$ in this case

- SVD of H_{11L}

$$H_{11L} = USV^T = [U_1\ U_2]\begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix}[V_1\ V_2]^T \approx U_1 S_1 V_1^T$$

≈ 0

where $S_1 = \operatorname{diag}(s_1, s_2 \cdots s_n)$, S_2 from small/zero singular values

May chose $U_1 = O_L$, $S_1 V_1^T = C_J$

- Solving

$$H_{21L} = O_L A C_J = U_1 A S_1 V_1^T$$
$$\Downarrow$$
$$U_1^T H_{21L} V_1 = A S_1 \implies \underline{\hat{A} = U_1^T H_{21L} V_1 S_1^{-1}}$$

or

$$A = (O_L^T O_L)^{-1} O_L^T H_{21L} C_J^T (C_J C_J^T)^{-1}$$

Task 2

a)

An observer estimating the state x_k, eg. calculating an estimate $\hat{x}_k$ in an optimal sense, e.g. The Kalman filter minimizing the mean square error

$$\min \; E\left((x_k - \hat{x}_k)(x_k - \hat{x}_k)^T\right) = \hat{X},$$

i.e. the covariance matrix $\hat{X}$ of the error $x_k - \hat{x}_k$

b)

$\bar{y}_k = D\bar{x}_k$ (prediction of y_k based on apriori ~~state~~ state estimate $\bar{x}_k$)

$\hat{x}_k = \bar{x}_k + K(y_k - \bar{y}_k)$ (Aposterior state estimate, K = kalman gain)

$\bar{x}_{k+1} = A\hat{x}_k + Bu_k$

c) Eleminating $\hat{x}_k$ in above algorithm

$$\bar{x}_{k+1} = A\bar{x}_k + Bu_k + AK(y_k - \bar{y}_k)$$

$$\bar{x}_{k+1} = A\bar{x}_k + Bu_k + \hat{K}\varepsilon_k$$

$$y_k = \bar{y}_k + \varepsilon_k$$

when $\hat{K} = AK$ is kalman gain in innovations formulation.

d) The relationship is

$$\underline{\tilde{K} = AK}$$

when K is Kalman gain in algorithm in 2b) and $\tilde{K}$ Kalman gain in innovations formulation in 2c).

e) $\bar{y}_k = g(\bar{x}_k)$ (Initial predicted, apriori state $\bar{x}_{k=0}$ for starting alg.)

$\hat{x}_k = \bar{x}_k + K(y_k - \bar{y}_k)$ (May use const. K)

$\hat{x}_{k+1} = f(\hat{x}_k, u_k)$

Above alg for constant K

f) • Kalman filter on prediction form

$\bar{y}_k = D\bar{x}_k$

$\bar{x}_{k+1} = Ax_k + Bu_k + \tilde{K}\overbrace{(y_k - \bar{y}_k)}^{\varepsilon_k}$

• Prediction error $\varepsilon_k = y_k - \bar{y}_k$

– Prediction error criterion

$$V_N(\theta) = \frac{1}{N}\sum_{k=1}^{N} \varepsilon_k^T \varepsilon_k$$

Find $\hat{\theta}_N$ minimizing $V_N(\theta)$

$n = 2$

$A = \begin{bmatrix} 0 & 1 \\ \theta_1 & \theta_2 \end{bmatrix}$, $B = \begin{bmatrix} \theta_3 \\ \theta_4 \end{bmatrix}$, $K = \begin{bmatrix} \theta_5 \\ \theta_6 \end{bmatrix}$, $D = [1 \ 0]$

$\theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_6 \end{bmatrix}$ if $\bar{x}_{k=0} = 0$

If we want an estimate of $\bar{x}_0 = \begin{bmatrix} \theta_7 \\ \theta_8 \end{bmatrix}$

$\theta = [\theta_1 \cdots \theta_8]^T$

Task 3

a) $N=10$, $L=2$ and $g=0$, $y_k \; \forall k=0,1,\dots,9$

$$Y_{013} = O_3 X_0 + H_3^d U_{012} \qquad (1)$$

We have

$$Y_{013} = \begin{bmatrix} y_0 & y_1 & y_2 & y_3 & y_4 & y_5 & y_6 & y_7 \\ y_1 & y_2 & y_3 & y_4 & y_5 & y_6 & y_7 & y_8 \\ y_2 & y_3 & y_4 & y_5 & y_6 & y_7 & y_8 & y_9 \end{bmatrix} \in \mathbb{R}^{3m \times 8}$$

$K=8$ columns in Y_{013}, X_0 and U_{012}

$$X_0 = \begin{bmatrix} x_0 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \end{bmatrix}, \quad O_3 = \begin{bmatrix} D \\ DA \\ DA^2 \end{bmatrix}$$

$$U_{012} = \begin{bmatrix} u_0 & u_1 & u_2 & u_3 & u_4 & u_5 & u_6 & u_7 \\ u_1 & u_2 & u_3 & u_4 & u_5 & u_6 & u_7 & u_8 \end{bmatrix}$$

$g=0$ means that $E=0$

$$H_3^d = \begin{bmatrix} 0 & 0 \\ DB & 0 \\ DAB & DB \end{bmatrix}$$

b) Multiply from right in (1) with projection matrix

N.B. Many students miss this definition

$$P = U_{0|2}^{\perp} = I_K - U_{0|2}^T \left(U_{0|2} U_{0|2}^T \right)^{+} U_{0|2}$$

such that $U_{0|2} P = U_{0|2} U_{0|2}^{\perp}$

$$= U(I - U^T(UU^T)^{+}U) = U - U = 0$$

Then

$$Z_{0|L+1} = Y_{0|L+1} P = Y_{0|L+1} U_{0|L+g}^{\perp} = Y_{0|3} U_{0|2}^{\perp}$$

c) • SVD of $Z_{0|L+1}$

$$Z_{0|L+1} = USV^T = [U_1 \; U_2] \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix} [V_1 \; V_2]^T \approx U_1 S_1 V_1^T$$

• System order n = # of non-zero singular values in $Z_{0|L+1}$

• $O_{L+1} = U_1$

• D as the upper $m \times n$ submatrix in O_{L+1}

- A may be found using the shift invariance principle, when we know O_{L+1} and noticing that

$$O_{L+1} = \begin{bmatrix} O_L \\ \hline X \end{bmatrix} = \begin{bmatrix} \overset{D}{X} \\ \hline O_L A \end{bmatrix} \qquad (2)$$

- Define $\underline{O}_{L+1}$ as O_{L+1} omitting the last $m \times n$ submatrix (indicated as x a

$\Rightarrow \underline{O}_{L+1} = O_L$

- Define $\overline{O}_{L+1}$ as O_{L+1} omitting the upper ($m \times m$) submatrix in O_{L+1} (se (2) above)

$\Rightarrow \overline{O}_{L+1} = O_L A = \underline{O}_{L+1} A$

- This gives

$$A = (\underline{O}_{L+1}^T \underline{O}_{L+1})^{-1} \underline{O}_{L+1}^T \overline{O}_{L+1} = \underline{O}_{L+1}^{+} \overline{O}_{L+1}$$

d)

- At this stage $\tilde{A}_L$ is known

$$\tilde{A}_L = O_L A (O_L^T O_L)^{-1} O_L^T$$

- Compute $\hat{B}_L$ from (17)

$$\tilde{B}_L = (Y_{1|L} - \hat{A}_L Y_{0|L}) U_{0|L+g}^T (U_{0|L+g} U_{0|L+g}^T)^+$$

Now using the structure of $\hat{B}_L$ to find E and $O_L B$

$$\hat{B}_L = [O_L B \quad H_L^d] - \tilde{A}_L [H_L^d \quad 0]$$

Take the example to illustrate

- We find

$$\tilde{B}_2 = \left[\begin{bmatrix} D \\ DA \end{bmatrix} B - \tilde{A}_2 \begin{bmatrix} 0 \\ DB \end{bmatrix} \;\middle|\; \begin{bmatrix} 0 \\ DB \end{bmatrix} \right]$$

and $\begin{bmatrix} 0 \\ DB \end{bmatrix}$ directly as the last column in $\hat{B}_L$. Then we solve the 1st column in $\tilde{B}_2$ for $O_2 B = \begin{bmatrix} D \\ DA \end{bmatrix} B$ and then we find B

Task 4

a) Parameters : $N=10$, $L=J=2$, $g=0$ ($E=$

Start with matrices in (22).

$$Y_{J+1|L} = Y_{3|2} = \begin{bmatrix} y_3 & y_4 & y_5 & y_6 & y_7 & y_8 \\ y_4 & y_5 & y_6 & y_7 & y_8 & y_9 \end{bmatrix}$$

$K=6$ columns in all data matrices!

$$Y_{J|L} = \begin{bmatrix} y_2 & y_3 & y_4 & y_5 & y_6 & y_7 \\ y_3 & y_4 & y_5 & y_6 & y_7 & y_8 \end{bmatrix}$$

$$U_{J|L+g} = U_{2|2} = \begin{bmatrix} u_2 & u_3 & u_4 & u_5 & u_6 & u_7 \\ u_3 & u_4 & u_5 & u_6 & u_7 & u_8 \end{bmatrix}$$

$$E_{J|L+1} = E_{2|3} = \begin{bmatrix} e_2 & e_3 & e_4 & e_5 & e_6 & e_7 \\ e_3 & e_4 & e_5 & e_6 & e_7 & e_8 \\ e_4 & e_5 & e_6 & e_7 & e_8 & e_9 \end{bmatrix}$$

$$\tilde{A}_L = O_L A (O_L^T O_L)^{-1} O_L^T \,, \quad \tilde{B}_L = [O_L B \;\; H_L^d] - \tilde{A}_L [H_L^d \;\; 0]$$

$$\tilde{C}_J = [O_L C \;\; H_L^s] - \tilde{A}_L [H_L^s \;\; 0]$$

$$H_L^d = H_2^d = \begin{bmatrix} 0 \\ DB \end{bmatrix} , \quad H_L^s = \begin{bmatrix} F & 0 \\ DC & F \end{bmatrix}$$

Matrices in (21) 12/18

$$Y_{g|L} = Y_{2|2} = \begin{bmatrix} y_2 & y_3 & y_4 & y_5 & y_6 & y_7 \\ y_3 & y_4 & y_5 & y_6 & y_7 & y_8 \end{bmatrix}$$

$$X_g = X_2 = \begin{bmatrix} x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \end{bmatrix}$$

$$O_L = \begin{bmatrix} D \\ DA \end{bmatrix}$$

$$U_{g|L+g-1} = U_{2|1} = \begin{bmatrix} u_2 & u_3 & u_4 & u_5 & u_6 & u_7 \end{bmatrix}$$

$$E_{g|L} = E_{2|2} = \begin{bmatrix} e_2 & e_3 & e_4 & e_5 & e_6 & e_7 \\ e_3 & e_4 & e_5 & e_6 & e_7 & e_8 \end{bmatrix}$$

b) We may put $g=0$ in a deterministic system and when no noise!

$$Y_{0|L} = O_L X_0 + H_L^d U_{0|L+g-1} \qquad (3)$$

$$Y_{1|L} = \hat{A}_L Y_{0|L} + \hat{B}_L U_{0|L+g} \qquad (4)$$

Multiply (3) and (4) with projection matrix, such that

$$U^{\perp}_{0|L+g} = P = I - U_x^{T} (U_x U_x^{T})^{+} U_x$$

when x represents indexes $0|L+g$!

Hence,

$$Z_{0|L} = Y_{0|L} U^{\perp}_{0|L+g} = O_L X_g^{a}$$

when $X_g^{a} = X_g U^{\perp}_{0|L+g}$

and

$$\underbrace{Y_{1|L} U^{\perp}_{0|L+g}}_{Z_{1|L}} = \hat{A}_L \underbrace{Y_{0|L} U^{\perp}_{0|L+g}}_{= Z_{0|L}}$$

15,1

For a general system we first remove the noise term by projecting down on

$$W = \begin{bmatrix} U_{g|L+g} \\ U_{0|g} \\ Y_{0|g} \end{bmatrix}$$

and remove deterministic term with $U_{0|L+g}^{\perp}$

$$\underline{\underline{(Y_{g|L}/W) U_{0|L+g}^{\perp} = O_L X_g^a = Z_{g|L}}}$$

$$\underbrace{(Y_{g+1|L}/W) U_{0|L+g}^{\perp}}_{= Z_{g+1|L}} = \tilde{A}_L \underbrace{(Y_{g|L}/W) U_{0|L}^{\perp}}_{Z_{g|L}}$$

and were the projection A/B is defined as

$$\underline{\underline{A/B = AB^T(BB^T)^+B}}$$ Should be defined

when $(\ )^+$ denotes the pseudo inverse

c) • n and O_L from svd of

$Z_{y|L} = USV^T \approx U_1 S_1 V_1^T$

• n the number of non-zero singular values in $Z_{y|L}$, dimension of S_1 etc

• $O_L = U_1$

• A from $Z_{y+1|L} = O_L A O_L^+ Z_{y|L}$

gives $Z_{y+1|L} = U_1 A U_1^+ U_1 S_1 V_1^T$

$Z_{y+1|L} = U_1 A S_1 V_1^T \Rightarrow Z_{y+1|L} V_1 = U_1 A S_1$

$$\Rightarrow \underline{\underline{A = U_1^T Z_{y+1|L} V_1 S_1^{-1}}}$$

d) Comparing the two variants of the Kalman filter we obtain

$$C e_k = K \varepsilon_k$$

$$F e_k = \varepsilon_k$$

and $E(e_k e_k^T) = I$.

Gives $e_k = F^{-1} \varepsilon_k$

and $C e_k = C F^{-1} \varepsilon_k = K \varepsilon_k$

and hence

$$\underline{\underline{K = C F^{-1}}}$$

when the Kalman filter exist!

e). • We may estimate the innovations process $\varepsilon_k = F e_k$ for $k \geq J$ with the projection

$$Z^s_{J|1} = Y_{J|1} - Y_{J|1} / \begin{bmatrix} U_{0|J} \\ Y_{0|J} \end{bmatrix}$$

$$= [\varepsilon_J \; \varepsilon_{J+1} \; \cdots \; \varepsilon_{N-1}]$$

• And when ε_k is known we simply may solve a deterministic subspace problem as in Task 3

$$x_{k+1} = A x_k + \tilde{B} \tilde{u}_k$$

$$\tilde{y}_k = D x_k$$

where $\tilde{y}_k = y_k - \varepsilon_k$, $\tilde{u}_k = \begin{bmatrix} u_k \\ \varepsilon_k \end{bmatrix}$

$\tilde{B} = [B \; K]$, and notice that $E=0$ when feedback in the data!