ME8873 – Statistical Model Estimation
Homework Set No. 2
Due March 2nd, 2022

Note: Datasets for this assignment are posted under the data locker folder on Canvas. For each problem involving coding, please use either R or MATLAB and attach your commented code at the end of the problem.

Problem 1

The goal of this problem is to investigate uses for PCA and PLSR to interpret multivariate biological data.

a) A cell culture protein expression dataset is posted to the data locker (data2-1 – Sheet1) consisting of 32 cytokine proteins measured across three experimental groups. Use a PCA to determine if there are differences between groups. To do so, compute the PCA (in MATLAB or R) and plot the scores of the data on PC1 and PC2 (2D plot), then draw 95% confidence ellipses using a Hoteling T2 distribution (this is a multivariate t-distribution). Also plot the percent variance captured by each of the first 5 principal components. Finally, create bar graphs of the loadings on PC1 and PC2 to determine the relative contribution of each protein to each PC.

b) Data2-1 also includes a Y response variable (Sheet 2), which is paired with the rows (observations) in the X data. Use a partial least squares discriminant analysis to create a regression model relating X to Y. Compute the percent variance of X and percent variance of Y captured by the first 5 components in the model. Plot the scores on latent variables (LVs) 1 and 2 (2D plot).

c) New X data (data2-1 – Sheet3) becomes available from a fourth experimental group. Use your PLS model from part b) to i) project your new data onto the same scores plot and ii) predict the Y values for the new datapoints. Keep in mind that if you transformed the original dataset using a z-score, this same transform needs to be applied to the new data.

Problem 2

An electrocardiogram dataset is posted to the data locker (data2-2) consisting of three ECG channels measured from a pregnant woman. Each column is one channel measured at 256 Hz. Thus, the ECG data consist of the combination of the mother, the fetus, and sensor noise. Our goal is to use blind source separation to identify "clean" traces for the mother and the fetus.

a) make plots of the temporal trace data from each channel and annotate possible sources (mother, fetus, noise) in each trace. Note that each trace should have components from all three and that the fetus ECG should be faster than that of the mother.

b) use principal component analysis to identify the "sources" mother, fetus, and noise.

- i) Plot the percent variance captured for each channel and overlay the direction in space associated with each PC onto a 3D plot of the data in each channel.
- ii) Plot the temporal data associated with each "source" identified via PCA. Speculate on the identify of each source
- iii) Separately reconstruct the sensor outputs associated with the fetus, the mother, and the noise by zeroing out the other two sources in the PCA computed scores and transform back to sensor measurement space.

c) use independent component analysis to identify the "sources" mother, fetus, and noise. We recommend using the 'PCA and ICA Package" by Brian Moore on the Mathworks file exchange (MATLB). Indicate the cost function you selected as your independence measure (kertosis, negentropy)

- i) Overlay the direction in space associated with each IC onto a 3D plot of the data in each channel. What is different about the directions identified via ICA vs PCA?
- ii) Plot the temporal data associated with each "source" identified via ICA. Speculate on the identity of each source.
- iii) Separately, reconstruct the sensor outputs associated with the fetus, the mother, and the noise by zeroing out the other two sources in the ICA computed scores and transforming back to sensor space.

d) In 2-3 sentences, which approach works better? Why?

Problem 3

Consider a simple scalar random process governed by the following state transition equation:

$$x_{t+1} = x_t + w_t$$

where x_t is the state, and w_t is zero-mean process noise with

$$E[w_t w_s] = \begin{cases} Q > 0 & \forall t = s \\ 0 & \forall t \neq s \end{cases}.$$

The output of the process is observed with a single sensor having the following properties:

$$y_t = x_t + v_t$$
$$E[v_t v_s] = \begin{cases} R > 0 & \forall t = s \\ 0 & \forall t \neq s \end{cases}$$
$$E[v_t w_s] = 0 \qquad \forall t, \forall s$$

We want to build a Kalman filter and examine its characteristics with respect to the state estimation error covariance and the sensor and process noise properties.

(a). Write out all the recursive equations of Kalman filter for this scalar process. Simplify the equations as much as you can.

(b). Given $P_0 = 100$, *Q* = 1, and *R* = 1, plot the values of a posteriori and a priori error covariances, $P_0, P_{1|0}, P_1, P_{2|1}, P_2, P_{3|2}, P_3, \cdots$ against time. Also plot the Kalman gain $K_1, K_2, K_3, \cdots$. Repeat for different values of Q and R, and discuss how the process noise variance and the sensor noise variance are used in the Kalman filter for optimal state estimation.

Problem 4

A vector discrete-time random sequence x_t (of dimension 2x1) is given by

$$x_{t+1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} x_t + w_t$$
$$w_t \sim N(0, 1) \text{ and uncorrelated (White)}$$

The observation equation is given by

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} x_t + v_t$$
$$v_t \sim N(0,\ 2 + (-1)^t) \text{ and uncorrelated (White)}$$
$$E[v_t w_s] = 0 \qquad \forall t,\ \forall s$$

Calculate the values of $P_t,\ P_{t+1|t},\ K_t$ for $t = 1, \cdots$ with

$$P_0 = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}.$$

Problem 5

A discrete-time Kalman filter is applied estimate the location of an autonomous vehicle. The state equation and measurement equation of the system are given by

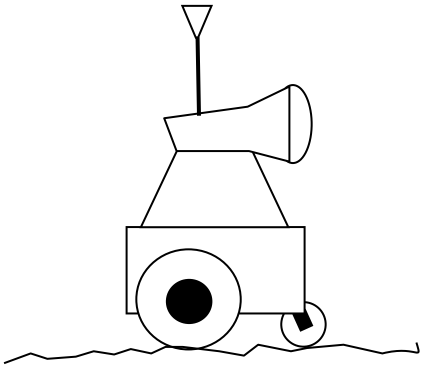
$$x_{t+1} = A \cdot x_t + w_t$$
$$y_t = H \cdot x_t + v_t$$

where $x_t \in R^{nx1}$, $y_t \in R^{\ell x1}$, $w_t \in R^{nx1}$, $v_t \in R^{\ell x1}$, $A \in R^{n \times n}$, and $H \in R^{\ell xn}$. Experiments were conducted to identify the characteristics of both process noise w_t and measurement noise v_t. It turns out that:

- The process noise can be treated as zero-mean value, uncorrelated (white) noise: $E[w_t] = 0$, and $E[w_t \cdot w_s^T] = \begin{cases} 0 & \forall t \neq s \\ Q & \forall t = s \end{cases}$
- The measurement noise, too, can be treated as zero-mean, uncorrelated (white) noise: $E[v_t] = 0$, and $E[v_t \cdot v_s^T] = \begin{cases} 0 & \forall t \neq s \\ R & \forall t = s \end{cases}$, but
- The process noise and the measurement noise are correlated to each other, since both process and measurement system are disturbed by the same source, i.e. the uneven road surface, so that:

  $$E[w_t \cdot v_s^T] = \begin{cases} 0 & \forall t \neq s-1 \\ C & \forall t = s-1 \end{cases}.$$

  Note that v_t is correlated with w_{t-1} alone: $E[w_{t-1} v_t^T] = C \in R^{nx\ell}$.



The standard Kalman filter, where the process noise and the measurement noise are assumed to be uncorrelated, must be modified for application to this system. Answer the following questions.

a) Let ε_t be *a priori* state estimation error, $\varepsilon_t \equiv \hat{x}_{t|t-1} - x_t$. Compute the correlation between the a priori error and the measurement noise: $E[\varepsilon_t v_t^T]$.

b) Obtain the modified Kalman gain as a function of matrices R, Q, C, A, H, and the covariance matrix $P_{t|t-1} = E[\varepsilon_t \varepsilon_t^T]$. Use the following equation that is obtained by differentiating the squared a posteriori error of state estimation, $e_t = \hat{x}_t - x_t$:

$$\frac{1}{2}\frac{d}{dK} e_t^T e_t = K_t H \varepsilon_t \varepsilon_t^T H^T - [K_t H \varepsilon_t v_t^T + K_t v_t \varepsilon_t^T H^T] + K_t v_t v_t^T + [\varepsilon_t v_t^T - \varepsilon_t \varepsilon_t^T H^T].$$

Problem 6

A heat pump is an important component involved in various energy systems ranging from air conditioners and refrigerators to fuel cells. Figure 1 depicts its principle, the refrigeration cycle. Refrigerant is first compressed at the compressor. The resultant high-pressure, high-temperature refrigerant is led to a heat exchanger, called a condenser, where the refrigerant becomes liquid while losing heat. The liquid refrigerant is then delivered to a location that we wish to cool. At the inlet, the high-pressure, liquid refrigerant is discharged through an expansion valve, which brings the refrigerant temperature to low. Then the low-temperature refrigerant is led to another heat exchanger, called an evaporator, where the refrigerant absorbs heat and changes its phase back to vapor. As this cycle is continued, heat is removed from the evaporator side to the condenser side.

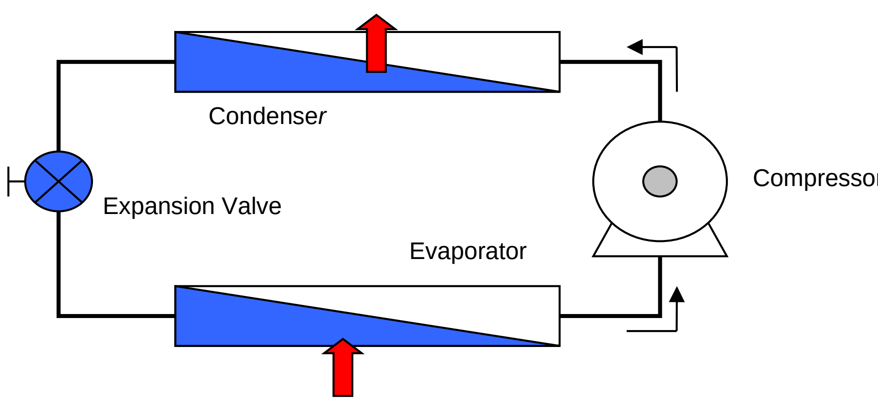


Figure 1: Refrigeration cycle.

In both condenser and evaporator, refrigerant is changing phase between liquid and gas. One side is all liquid and the other side is all gas. In between the refrigerant is in two-phase where liquid and gas coexist and the temperature is uniform throughout the phase transition. One of the critical state variables involved in this refrigeration cycle is the length of the two-phase section, L_2, as shown in the figure.

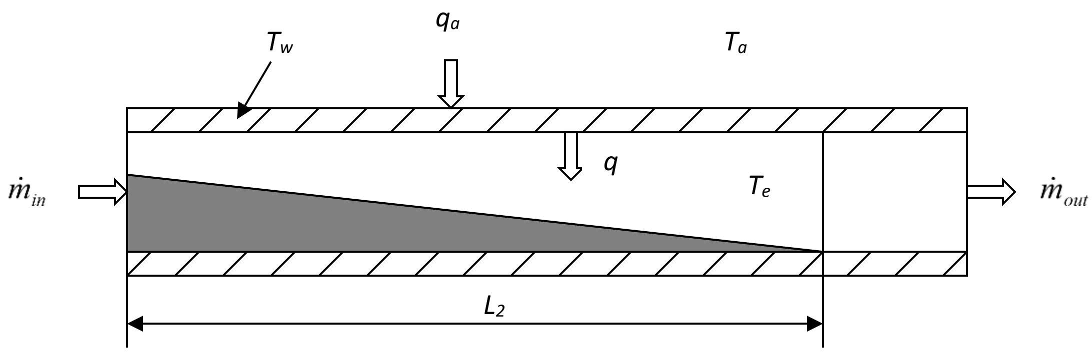


Figure 2 Evaporator: T_e is the evaporating temperature, L_2 is the length of the two-phase section. T_w is the wall temperature of the tube. T_a is the room air temperature. $\dot{m}_{in}$ and $\dot{m}_{out}$ are the inlet and outlet refrigerant mass flow rates, respectively. q is the heat transfer rate from the tube wall to the two-phase refrigerant. q_a is the heat transfer rate from the room to the tube wall.

If we know the two-phase length, we can easily obtain the total amount of heat transferred through the heat exchanger, and the control of the heat pump can be significantly improved. Unfortunately, however, no sensor is commercially available for directly measuring the two-phase length. The problem can be solved by using the Kalman filter to estimate the two-phase length based on the process model and available low-cost sensors, such as thermocouples.

The real heat exchanger dynamics is nonlinear. A low-order, linearized evaporator model is given by

$$\begin{pmatrix} \partial \dot{T}_e \\ \partial \dot{T}_w \\ \partial \dot{L}_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \partial T_e \\ \partial T_w \\ \partial L_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ 0 \\ b_2 \end{pmatrix} u + \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} w$$

$$y = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \partial T_e \\ \partial T_w \\ \partial L_2 \end{pmatrix} + v$$

The discrete version of the above model is

$$X_{t+1} = A_t X_t + B_t U_t + G_t w_t$$
$$y_t = H_t \cdot x_t + v_t$$

Using the data posted in the course locker (data2-6), implement the Kalman filter of this system and reconstruct the two-phase length L_2.