Rob 501 - Mathematics for R💀b💀tics
HW #7

Prof. Grizzle

Due Nov. 8, 2018
3PM via Gradescope

Remarks: Problems 1 through 5 involve calculations. Problem 5 is very important. Please spend extra time on it as it is very helpful for understanding the Kalman Filter. Problem 6 explains why we can often obtain recursion relations of the type: $\widehat{x}_{k+1}$ is a linear combination of $\widehat{x}_k$ and the "innovation" or new measured information $(y_{k+1} - \widehat{y}_{k+1|k})$. If you are pressed for time, skip Problem 6 and study the solutions. It is better to spend your time on Problem 5.

1. Classify each matrix as positive definite, positive semi-definite, or neither. In addition, if the matrix is either positive definite or positive semi-definite, find a square root. You may use MATLAB to factor a symmetric matrix as $\Lambda = O^\top P O$ or as $P = O \Lambda O^\top$.

   (a) 🎃 $= \begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix}$.

   (b) 💀 $= \begin{bmatrix} 6 & 10 & 11 \\ 10 & 19 & 19 \\ 11 & 19 & 21 \end{bmatrix}$.

   (c) 👻 $= \begin{bmatrix} 2 & 6 & 10 \\ 6 & 10 & 14 \\ 10 & 14 & 18 \end{bmatrix}$.

2. Use the results on Schur Complements to solve the following problems BY HAND:

   (a) Determine if 🎃 $= \begin{bmatrix} 1 & 3 \\ 3 & 8 \end{bmatrix}$ is positive definite or not.

   (b) Determine if 💀 $= \begin{bmatrix} 1 & 0 & 6 \\ 0 & 4 & 7 \\ 6 & 7 & 10 \end{bmatrix}$ is positive definite or not.

   (c) Find the range of a such that the following matrix is positive definite: 👻 $= \begin{bmatrix} 1 & 2 & 6 \\ 2 & 5 & 7 \\ 6 & 7 & a \end{bmatrix}$

3. Find x of minimum norm that satisfies the equation

$$\begin{bmatrix} 1 & 3 & 2 \\ 3 & 8 & 4 \end{bmatrix} x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

   (a) Use the standard inner product on $\mathbb{R}^3$

(b) Use the inner product $< x, y >= x^\top \begin{bmatrix} 5 & 1 & 9 \\ 1 & 2 & 1 \\ 9 & 1 & 17 \end{bmatrix} y$

4. x has been bewitched to give the following data:

$$y = $$  $$x + \epsilon$$

with

 $= \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 0 \\ 0 & 6 \end{bmatrix} \quad y = \begin{bmatrix} 1.5377 \\ 3.6948 \\ -7.7193 \\ 7.3621 \end{bmatrix} \quad \text{and} \quad E\{\epsilon\epsilon^\top\} = Q = \begin{bmatrix} 1.00 & 0.50 & 0.50 & 0.25 \\ 0.50 & 2.00 & 0.25 & 1.00 \\ 0.50 & 0.25 & 2.00 & 1.00 \\ 0.25 & 1.00 & 1.00 & 4.00 \end{bmatrix}$

As in class, $E\{\epsilon\} = 0$.

(a) Find the Best Linear Unbiased Estimate (BLUE) for x, using only the first two values of y. Also compute the covariance of the estimate.

(b) Find the Best Linear Unbiased Estimate (BLUE) for x, using only the first three values of y. Also compute the covariance of the estimate.

(c) Find the Best Linear Unbiased Estimate (BLUE) for x, using all the values of y. Also compute the covariance of the estimate.

Note: For (a), you use the first 2 rows of y and  and the upper 2×2 part of Q. For (b), you use the first 3 rows of y and C, as well as the upper 3×3 part of Q. You see the pattern, I hope. Do all the calculations in MATLAB. You do not have to turn in your code.

5. Read the Handout `GaussianRandomVariablesAndVectors.pdf`, which you can find on CANVAS. We consider three jointly normal random variables (X, Y, Z), with

$$\text{mean} \quad \mu = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad \text{and covariance} \quad \Sigma = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

(a) Compute the conditional distribution of $\begin{bmatrix} X \\ Y \end{bmatrix} |\{Z = z\}$, the conditional distribution of the vector $[X, Y]^\top$ given $Z = z$, which is the same as the joint distribution of the normal random variables $X|\{Z = z\}$ and $Y|\{Z = z\}$. To be extra clear, give the mean vector and covariance matrix for $[X, Y]^\top$ given $Z = z$.

(b) Compute the distribution of $X|_{\{Z=z\}}$ conditioned on $Y|_{\{Z=z\}} = y$.

(c) Compute the conditional distribution of $X|_{\begin{bmatrix} Y = y \\ Z = z \end{bmatrix}}$, or more compactly, $X_{|Y=y,Z=z}$, the conditional distribution of X given the vector $[Y = y, Z = z]^\top$.

(d) Compare your answers for (b) and (c).

6. Let $(\mathcal{X}, \mathbb{R}, < \cdot, \cdot >)$ be a finite-dimensional inner product space. Our objective is to understand recursion relations similar to what we see in RLS and the Kalman filter. With that in mind, let $\{y_1, \cdots, y_N\}$ be a linearly independent set in $\mathcal{X}$. For $1 \leq k \leq N$ define

$$M_k := \text{span}\{y_1, \cdots, y_k\}$$

and for $x \in \mathcal{X}$ define $\widehat{x}_k := \arg\min_{m \in M_k} ||x - m||$, which is the orthogonal projection of x onto M_k.

(a) Suppose that for some $1 \leq k < N$, we have $y_{k+1} \perp M_k$ (interpretation: the new measurement y_{k+1} is orthogonal to the previous measurements). Show that there exists $\beta \in \mathbb{R}$ such that

$$\widehat{x}_{k+1} = \widehat{x}_k + \beta y_{k+1},$$

and give a formula for β.

(b) We no longer make any hypothesis about y_{k+1} being orthogonal to M_k. What we do now is, for each $1 \leq k < N$, define

$$\widehat{y}_{k+1|k} = \arg\min_{m \in M_k} ||y_{k+1} - m||,$$

which we interpret as the orthogonal projection of the new measurement y_{k+1} onto the subspace generated by the previous measurements. Show that there exists $\beta \in \mathbb{R}$ such that

$$\widehat{x}_{k+1} = \widehat{x}_k + \beta(y_{k+1} - \widehat{y}_{k+1|k}),$$

and give a formula for β.

Remark: The error term $(y_{k+1} - \widehat{y}_{k+1|k})$ is the "innovations" in the case of the Kalman filter and RLS. What is particularly nice in the case of RLS and the Kalman filter, where we have a model such as $y_k = C_k x_k + v_k$, we can compute $\widehat{y}_{k+1|k}$ directly from $\widehat{x}_k$, often with a formula such as $\widehat{y}_{k+1|k} = C_{k+1}\widehat{x}_k$. In other words, we do not have to solve an extra optimization problem in order to compute $\widehat{y}_{k+1|k}$; instead, we can bootstrap from the previous solution to the optimization problem.

Hints

Hints: Prob. 1 Use the `help eig` command in MATLAB.

Hints: Prob. 2 Recall that for a symmetric matrix $M = \begin{bmatrix} A & B \\ B^\top & C \end{bmatrix}$ the following are equivalent:

(a) $M \succ 0$

(b) $A \succ 0$ and $C - B^\top A^{-1} B \succ 0$

(c) $C \succ 0$ and $A - BC^{-1}B^\top \succ 0$

Hints: Prob. 3 This is an under determined system of equations and not an over determined system of equations.

Hints: Prob. 4 Recall our formulas (using $C :=$  for convenience):

$$\widehat{K} = (C^\top Q^{-1} C)^{-1} C^\top Q^{-1} \quad \text{and} \quad E\{(\widehat{x} - x)(\widehat{x} - x)^\top\} = (C^\top Q^{-1} C)^{-1}$$

For (a), you use the first 2 rows of y and C and the upper 2×2 part of Q. For (b), you use the first 3 rows of y and C, as well as the upper 3×3 part of Q. You see the pattern, I hope. Do all the calculations in MATLAB. You do not have to turn in your code.

Hints: Prob. 5 Print out the handout on Jointly Gaussian Random Vectors, and read it carefully. Note **Fact 1:** Conditional Distributions of Gaussian Random Vectors

(a) Identify $X_1 = \begin{bmatrix} X \\ Y \end{bmatrix}$ and $X_2 = Z$. Based on this, identify and write down Σ_{11}, Σ_{12}, Σ_{21}, and Σ_{22}, and then μ_1 and μ_2, and finally, note that $x_2 = z$. Now apply the formulas for $\mu_{1|2}$ and $\Sigma_{1|2}$. These are the mean and covariance of the jointly normally distributed random variables $X_{|Z=z}$ and $Y_{|Z=z}$.

(b) From (a), we know $X_{|Z=z}$ and $Y_{|Z=z}$ are jointly distributed normal random variables, and we know their mean and covariance. Rename the mean μ and the covariance Σ (i.e., $\mu_{1|2} \to \mu$ and $\Sigma_{1|2} \to \Sigma$). Using Fact 1, identify $X_1 = X_{|Z=z}$ and $X_2 = Y_{|Z=z}$, and then identify and write down Σ_{11}, Σ_{12}, Σ_{21}, and Σ_{22}, and then μ_1 and μ_2, and finally, note that $x_2 = y$. Now apply the formulas for $\mu_{1|2}$ and $\Sigma_{1|2}$. These are the mean and covariance of a normally distributed random variable. Which one? If you do not know, work part (c) and then return here. If you do know, still work part (c).

(c) Go back to the very beginning with our three jointly normal random variables, and this time identify $X_1 = X$ and $X_2 = \begin{bmatrix} Y \\ Z \end{bmatrix}$. Based on this, identify and write down Σ_{11}, Σ_{12}, Σ_{21}, and Σ_{22}, and then μ_1 and μ_2, and finally, note that $x_2 = \begin{bmatrix} y \\ z \end{bmatrix}$. Now apply the formulas for $\mu_{1|2}$ and $\Sigma_{1|2}$. These are the mean and covariance of the normally distributed random variable $X_{|Y=y,Z=z}$.

(d) Read again the handout on Jointly Gaussian Random Vectors. Go back to the beginning and repeat as necessary: **FACT 4:** *If we have jointly distributed normal random vectors, when we condition one* *block of vectors on another, we always obtain either a jointly distributed normal random vector or, if only a scalar quantity is left, a normally distributed random variable.* This is an amazingly useful property of Gaussian (i.e., normal) random variables.

Hints: Prob. 6 There are several ways to approach part (a):

- Method 1: Let G_k be the Gram matrix for M_k and G_{k+1} be the Gram matrix for M_{k+1}. Then, using $y_{k+1} \perp M_k$, you deduce that
$$G_{k+1} = \begin{bmatrix} G_k & 0_{1\times k} \\ 0_{k\times 1} & < y_{k+1}, y_{k+1} > \end{bmatrix}$$
Use this block diagonal structure to relate the solution of the normal equations for M_{k+1} to the solution of the normal equations for M_k.

- Method 2: The solution of the optimization problem depends on the subspaces M_k and not on the bases you use for them. Apply Gram Schmidt to produce an ortho**normal** basis for M_k. Relate it to an orthornormal basis for M_{k+1}. Then use what you know about the orthogonal projection of x onto sets with orthonormal bases.

For part (b), we know from the Projection Theorem that $y_{k+1} - \widehat{y}_{k+1|k}$ is orthogonal to M_k. If you need a second hint, note that
$$M_{k+1} = M_k \oplus \text{span}\{y_{k+1}\} = M_k \oplus \text{span}\{y_{k+1} - v\},$$
for any $v \in M_k$. Hence, because $\widehat{y}_{k+1|k} \in M_k$ by definition, we have
$$M_{k+1} = M_k \oplus \text{span}\{y_{k+1} - \widehat{y}_{k+1|k}\} \;\text{ and }\; M_k \perp (y_{k+1} - \widehat{y}_{k+1|k}).$$
You can now apply your result from (a).