AA 549: Estimation and Kalman Filtering
Homework #1
Due: Sunday, Apr 11, 2021 @ 11:59 pm

1. **Block Matrix Computations**

   Multiply the following block matrices together. In each case give the required dimensions of the sub-blocks of B. If the dimensions are not determined by the shapes of A, then pick a dimension that works.

   (a) **(PTS: 0-2)**

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & & \vdots \\ A_{M1} & \cdots & A_{MN} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & \cdots & B_{1K} \\ \vdots & & \vdots \\ B_{N1} & \cdots & B_{NK} \end{bmatrix}, \qquad AB = ? \tag{1}$$

   where $A_{11} \in \mathbb{R}^{m_1 \times n_1}$, $A_{1N} \in \mathbb{R}^{m_1 \times n_N}$, $A_{M1} \in \mathbb{R}^{m_M \times n_1}$, and $A_{MN} \in \mathbb{R}^{m_M \times n_N}$.

   (b) **(PTS: 0-2)**

$$A = \begin{bmatrix} - & A_1 & - \\ \vdots & & \vdots \\ - & A_m & - \end{bmatrix}, \quad B = \begin{bmatrix} | & \cdots & | \\ B_1 & & B_k \\ | & \cdots & | \end{bmatrix}, \qquad AB = ? \tag{2}$$

   where $A_1 \in \mathbb{R}^{1 \times n}$ and $A_m \in \mathbb{R}^{1 \times n}$.

   (c) **(PTS: 0-2)**

$$\begin{bmatrix} | & \cdots & | \\ A_1 & & A_n \\ | & \cdots & | \end{bmatrix}, \quad B = \begin{bmatrix} - & B_1 & - \\ \vdots & & \vdots \\ - & B_n & - \end{bmatrix}, \qquad AB = ? \tag{3}$$

   where $A_1 \in \mathbb{R}^{m \times 1}$ and $A_n \in \mathbb{R}^{m \times 1}$.

   (d) **(PTS: 0-2)**

$$A = \begin{bmatrix} - & A_1 & - \\ \vdots & & \vdots \\ - & A_m & - \end{bmatrix}, \quad D \in \mathbb{R}^{n \times n}, \quad B = \begin{bmatrix} | & \cdots & | \\ B_1 & & B_k \\ | & \cdots & | \end{bmatrix}, \qquad ADB = ? \tag{4}$$

   where $A_1 \in \mathbb{R}^{1 \times n}$, $A_m \in \mathbb{R}^{1 \times n}$.

   (e) **(PTS: 0-2)**

$$\begin{bmatrix} | & \cdots & | \\ A_1 & & A_n \\ | & \cdots & | \end{bmatrix}, \quad D = \begin{bmatrix} d_{11} & \cdots & d_{1n} \\ \vdots & & \vdots \\ d_{n1} & \cdots & d_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} - & B_1 & - \\ \vdots & & \vdots \\ - & B_n & - \end{bmatrix}, \qquad ADB = ? \tag{5}$$

   where $A_1 \in \mathbb{R}^{m \times 1}$, $A_n \in \mathbb{R}^{m \times 1}$, $d_{ij} \in \mathbb{R}$.

(f) **(PTS: 0-2)**

$$A \in \mathbb{R}^{m\times n}, \quad \begin{bmatrix} B_1 & \cdots & B_k \end{bmatrix}, \qquad AB = ? \tag{6}$$

(g) **(PTS: 0-2)**

$$A = \begin{bmatrix} -A_1- \\ \vdots \\ -A_m- \end{bmatrix}, \quad B, \qquad AB = ? \tag{7}$$

where $A_1, A_m \in \mathbb{R}^{1\times n}$.

2. **Projections**

- **(PTS: 0-2)** Consider two vectors

$$A_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

Compute two projection matrices $P_1 \in \mathbb{R}^{3\times 3}$ and $P_2 \in \mathbb{R}^{3\times 3}$ such that $P_1 x$ is the projection of x onto the range of A_1 and $P_2 x$ is the projection of x onto the range of A_2. (You can use the formula $P_1 = A_1(A_1^T A_1)^{-1} A_1^T$, etc.)

- **(PTS: 0-2)** Now consider a matrix $A \in \mathbb{R}^{3\times 2}$

$$A = \begin{bmatrix} | & | \\ A_1 & A_2 \\ | & | \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Compute the projection matrix onto the range of A given by $P = A(A^T A)^{-1} A^T$. How does P relate to P_1 and P_2? What property of the columns of A makes this relationship fairly simple?

- **(PTS: 0-2)** Now modify A_2 to be

$$A_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

What is the projection matrix $A(A^T A)^{-1} A^T$ now? In particular how did the quantity $A^T A$ change? How does this relate to the orthogonality of A_1 and A_2?

- **(PTS: 0-2)** Now for A_1 and the new A_2, let

$$B = \begin{bmatrix} | & | \\ B_1 & B_2 \\ | & | \end{bmatrix} = A(A^T A)^{-1/2}$$

What is the angle between the columns of B? How does the range of B relate to the range of A?

- **(PTS: 0-2)** Write the projection matrix $A(A^T A)^{-1} A^T$ in terms of B and then again specifically in terms of B_1 and B_2.

3. **Joint Gaussian Distributions**

Consider the 2D joint Gaussian for $x \in \mathbb{R}^2$, $x \sim \mathcal{N}(\mu, \Sigma)$ with $\mu \in \mathbb{R}^2$ and $\Sigma \in \mathbb{R}^{2\times 2}$

$$\mu = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

with joint probability density function

$$f(x) = \frac{1}{\sqrt{(2\pi)^2 \det(\Sigma)}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}$$

- **(PTS: 0-2)** Plot the joint distribution pdf $f(x)$ (a surface in 2D).
(Suggestion: use `https://www.geogebra.org/calculator`)
- **(PTS: 0-2)** Plot the level set $f(x) = 0.01$.
- **(PTS: 0-2)** What are the eigenvectors of Σ? How do they relate to the axes of the ellipse plotted above?
- **(PTS: 0-2)** What is the ratio between the lengths of the major and minor axes of the ellipse? How does this quantity relate to the eigenvalues of Σ?

Now consider the 1D normal distibution for $z \in \mathbb{R}$, $z \sim \mathcal{N}(\nu, \sigma^2)$ for $\nu \in \mathbb{R}$ and $\sigma^2 \in \mathbb{R}_+$. Let

$$x = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} z$$

- **(PTS: 0-2)** What is the distribution of $x \in \mathbb{R}^3$? ie. find $\mu \in \mathbb{R}^3$ and $\Sigma \in \mathbb{R}^{3\times 3}$ (in terms of ν and σ^2 such that $x \sim \mathcal{N}(\mu, \Sigma)$.

Finally, consider the joint distribution for $x \in \mathbb{R}^3$, $x \sim \mathcal{N}(\mu, \Sigma)$ where

$$\mu = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{bmatrix}$$

- **(PTS: 0-2)** What are the marginal densities of x_1? x_2? x_3?
- **(PTS: 0-2)** What is the probability density function conditioned on the fact that $\sum_i x_i = 1$?

4. **Vector & Matrix Derivatives**

- **(PTS: 0-2)** Let

$$f(x) = \frac{1}{(2\pi)^2 \det(\Sigma)} e^{(x-\mu)^T \Sigma^{-1} (x-\mu)}$$

What is an expression for $\frac{\partial f}{\partial x}$? What is an expression for $\frac{\partial f}{\partial \Sigma}$?

- **(PTS: 0-2)**
  Consider the least-squares optimization problem for $A \in \mathbb{R}^{m \times n}$ $m > n$ full col rank.

$$\min_x \quad f(x) = ||y - Ax||_2^2$$

  Use the optimality condition $\frac{\partial f}{\partial x} = 0$ to solve for the optimal x.
- **(PTS: 0-2)** Consider the minimum-norm optimization problem for $A \in \mathbb{R}^{m \times n}$ $m < n$ full row rank.

$$\begin{aligned} \min_x \quad & \tfrac{1}{2} x^T x \\ \text{s.t.} \quad & y = Ax \end{aligned}$$

  with Lagrangian

$$\mathcal{L}(x, v) = \tfrac{1}{2} x^T x + v^T (Ax - y)$$

  Write out the expressions $\frac{\partial \mathcal{L}}{\partial x} = 0$ and $\frac{\partial \mathcal{L}}{\partial v} = 0$ and use them to solve for the optimal x and v.
- **(PTS: 0-2)** Now consider the optimization problem for $M \in \mathbb{R}^{m \times n}$

$$\begin{aligned} \min_{M \in \mathbb{R}^{m \times n}} \quad & \tfrac{1}{2} \text{Tr}(MRM^T) \\ \text{s.t.} \quad & I = MH \end{aligned}$$

  for $R \succ 0$, $H \in \mathbb{R}^{n \times m}$ with $m < n$ full column rank with Lagrangian

$$\mathcal{L}(M, V) = \tfrac{1}{2} \text{Tr}(MRM^T) + \text{Tr}\big(V^T (I - MH)\big)$$

  Write out expressions for $\frac{\partial \mathcal{L}}{\partial M} = 0$ and $\frac{\partial \mathcal{L}}{\partial V} = 0$ and use them to solve for the optimal M and V.