STOR 565 Homework

Show all work. Note: all logarithms are natural logarithms.

1. Let $\mathbf{u}_1 = (-1, 2, 0)^t$ and $\mathbf{u}_2 = (2, 4, 3)^t$. Find the projections of $\mathbf{u}_1$ and $\mathbf{u}_2$ onto $\mathbf{v}$ where:

    1. $\mathbf{v} = (0, 1, 0)^t$
    2. $\mathbf{v} = (1, 1, 1)^t$
    3. $\mathbf{v} = (1, 0, -1)^t$

2. Let $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^d$ be orthonormal vectors with span $V = \{\alpha \mathbf{v}_1 + \beta \mathbf{v}_2 : \alpha, \beta \in \mathbb{R}\}$. For $\mathbf{u} \in \mathbb{R}^d$ define the projection of $\mathbf{u}$ onto V to be the vector $\mathbf{v} \in V$ that is closest to $\mathbf{u}$,

$$\text{proj}_V(\mathbf{u}) = \underset{\mathbf{v} \in V}{\text{argmin}} \, ||\mathbf{u} - \mathbf{v}||.$$

Show that $\text{proj}_V(\mathbf{u}) = \langle \mathbf{u}, \mathbf{v}_1 \rangle \, \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \, \mathbf{v}_2$. Hint: Adapt the argument used in class for the projection onto a one-dimensional subspace.

3. Consider a data set consisting of four points in $\mathbb{R}^2$

$$\mathbf{x}_1 = (1, 2)^t, \ \mathbf{x}_2 = (-1, 2)^t, \ \mathbf{x}_3 = (2, -1)^t, \ \mathbf{x}_4 = (2, 1)^t$$

    1. Replace each observation $\mathbf{x}_i$ by the centered observation $\tilde{\mathbf{x}}_i = \mathbf{x}_i - \frac{1}{4} \sum_{j=1}^4 \mathbf{x}_j$. Draw a plot of the points $\tilde{\mathbf{x}}_i$. Form a data matrix $\mathbf{X}$ from $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_4$.
    2. Calculate the sample covariance matrix $\mathbf{S} = \frac{1}{4} \mathbf{X}^T \mathbf{X}$.
    3. Calculate the eigenvalues of $\mathbf{S}$. Is $\mathbf{S}$ invertible? If so, find $\mathbf{S}^{-1}$.
    4. Find orthonormal eigenvectors of $\mathbf{S}$.
    5. What is the best one-dimensional subspace (line) for approximating the centered observations $\tilde{\mathbf{x}}_i$? Draw this line on your plot.

4. *Measuring the variability of a set of vectors.* Let $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^p$ be a sample of n p-dimensional vectors. We can measure the extent to which a vector $\mathbf{u} \in \mathbb{R}^p$ acts as representative for the sample through the sum of squares

$$S(\mathbf{u}) := \sum_{i=1}^{n} ||\mathbf{x}_i - \mathbf{u}||^2.$$

a. Show that $S(\mathbf{u})$ is minimized when $\mathbf{u}$ is equal to the centroid

$$\overline{\mathbf{x}} \;=\; \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i.$$

If the general case seems difficult, consider first the case when $p = 1$, which you addressed in the previous homework.

Consider the two variance-type quantities

$$V_1 \;=\; \frac{1}{n}\sum_{i=1}^{n} ||\mathbf{x}_i - \overline{\mathbf{x}}||^2 \quad \text{and} \quad V_2 \;=\; \frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} ||\mathbf{x}_i - \mathbf{x}_j||^2.$$

Note that V_1 and V_2 are non-negative.

b. Carefully describe V_1 and V_2 in plain English.

c. Give necessary and sufficient conditions under which $V_1 = 0$.

d. Give necessary and sufficient conditions under which $V_2 = 0$.

e. Show that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{x}_i^t \mathbf{x}_j \;=\; (\sum_{i=1}^{n} \mathbf{x}_i)^t \, (\sum_{j=1}^{n} \mathbf{x}_j) \;=\; n^2 \, ||\overline{\mathbf{x}}||^2$$

f. Using the identity from part e., and some additional calculations, show that

$$V_1 \;=\; V_2 \;=\; \frac{1}{n}\sum_{i=1}^{n} ||\mathbf{x}_i||^2 - ||\overline{\mathbf{x}}||^2$$

5. (Inequalities from Calculus) Use calculus to establish the following inequalities.

a. $(1 + u/3)^3 \geq 1 + u$ for every $u \geq 0$

b. $x + x^{-1} \geq 2$ for $x \geq 1$

c. $\log(1 + x) \geq x - x^2/2$ for $x \geq 0$. Note that this inequality requires taking a second derivative to show that the first derivative is increasing.

6. Show that if $\mathbf{v}_1, \mathbf{v}_2$ are eigenvectors of a symmetric matrix $\mathbf{A}$ with different eigenvalues, then $\mathbf{v}_1, \mathbf{v}_2$ are orthogonal. Hint: Begin by taking transposes to show that $\mathbf{v}_1^t \mathbf{A} \mathbf{v}_2$ and $\mathbf{v}_2^t \mathbf{A} \mathbf{v}_1$ are equal; then use the definition of an eigenvector and simplify.

7. Recall that the trace of an $n \times n$ matrix $\mathbf{A} = \{a_{ij}\}$ is the sum of its diagonal elements, that is $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$.

a. Show that $\mathrm{tr}(\mathbf{A}) = \mathrm{tr}(\mathbf{A}^t)$.

b. Note that $(\mathbf{A}\,\mathbf{B})_{ii} = \sum_{j=1}^{n} a_{ij}\, b_{ji}$ (Why?). Use this to show that $\mathrm{tr}(\mathbf{A}\,\mathbf{B}) = \mathrm{tr}(\mathbf{B}\,\mathbf{A})$.

c. By applying the identity of part b. multiple times, show that

$$\mathrm{tr}(\mathbf{A}\,\mathbf{B}\,\mathbf{C}) \;=\; \mathrm{tr}(\mathbf{B}\,\mathbf{C}\,\mathbf{A}) \;=\; \mathrm{tr}(\mathbf{C}\,\mathbf{A}\,\mathbf{B})$$

d. Suppose that $\mathbf{B} = \{b_{ij}\}$ is an $m \times n$ matrix. By considering $(\mathbf{B}^t\,\mathbf{B})_{ii}$, show that

$$\mathrm{tr}(\mathbf{B}^t\,\mathbf{B}) \;=\; \sum_{i=1}^{m}\sum_{j=1}^{n} b_{ij}^2$$