STOR 565 Homework: Linear Algebra and Matrices

1. Let $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^t \mathbf{v} = \sum_{i=1}^d u_i \, v_i$ be the usual inner product in $\mathbb{R}^d$. Recall that the norm of a vector $\mathbf{u} \in \mathbb{R}^d$ is defined by $||\mathbf{u}|| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2}$. Use this definition, and the definition of vector sums and scalar multiplication to establish the following.

   a. Show that $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$

   b. Show that $\langle a\mathbf{u}, b\mathbf{v} \rangle = ab \, \langle \mathbf{u}, \mathbf{v} \rangle$

   c. Show that $\langle \mathbf{u} + \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle$

   d. Show that $||\mathbf{u}|| = 0$ if and only if $\mathbf{u} = 0$.

   e. Use the definition of the norm to show that $||\mathbf{u} + \mathbf{v}||^2 = ||\mathbf{u}||^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle + ||\mathbf{v}||^2$.

   f. Use this equation and the Cauchy Schwarz inequality to establish the triangle inequality for the vector norm, namely $||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$.

   g. The standard Euclidean distance between two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^d$ is defined by $d(\mathbf{u}, \mathbf{v}) = ||\mathbf{u} - \mathbf{v}||$. Use part (c) to establish that $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$ for any vectors $\mathbf{u}, \mathbf{v}, z \in \mathbb{R}^d$. Draw a picture illustrating this result.

2. Let $\mathbf{X} \in \mathbb{R}^{n \times p}$ be the data matrix associated with n samples $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^p$ such that $\sum_{i=1}^n \mathbf{x}_i = 0$. Answer the following. You may use arguments from class, but clearly explain your work.

   a. Define the sample covariance matrix $\mathbf{S}$ in terms of $\mathbf{X}$. What are the dimensions of $\mathbf{S}$?

   b. Show that $\mathbf{S} = n^{-1} \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^t$

   c. Show that $\mathbf{S}$ is symmetric and non-negative definite

   d. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ be the eigenvalues of $\mathbf{S}$. Show that $\sum_{k=1}^p \lambda_k = n^{-1} ||\mathbf{X}||^2$

   e. Show that if $p > n$ then $\text{rank}(\mathbf{S}) < p$ and $\mathbf{S}$ is not invertible. Hint: recall that $\text{rank}(\mathbf{S}) = \text{rank}(\mathbf{X}^t \, \mathbf{X}) = \text{rank}(\mathbf{X}) \leq \min(n, p)$.

   f. For any vector $\mathbf{v} \in \mathbb{R}^p$ we have $n^{-1} \sum_{i=1}^n \langle \mathbf{x}_i, \mathbf{v} \rangle^2 = \mathbf{v}^t \mathbf{S} \mathbf{v}$.

3. Let $\mathbf{u} = (u_1, \ldots, u_d)^t$ be a vector in $\mathbb{R}^d$.

a. Show that $||\mathbf{u}|| \leq |u_1| + \cdots + |u_d|$. Hint: use the fact that for $a, b \geq 0$ one has $a \leq b$ if and only if $a^2 \leq b^2$. Give an examples with $d = 2$ where the bound holds with equality, and where one has strict inequality.

b. Use the Cauchy-Schwarz inequality to get the upper bound $|u_1|+\cdots+|u_d| \leq ||\mathbf{u}||\, d^{1/2}$. Find an example where the bound holds with equality.

4. (Norms of outer products) Let $\mathbf{u} \in \mathbb{R}^k$ and $\mathbf{v} \in \mathbb{R}^l$ be vectors. Find an expression relating the Frobenius norm of the outer product $||\mathbf{u}\mathbf{v}^t||$ to the Euclidean norms of the vectors $||\mathbf{u}||$ and $||\mathbf{v}||$.

5. 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:

a. $\mathbf{v} = (0, 1, 0)^t$

b. $\mathbf{v} = (1, 1, 1)^t$

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

6. 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.

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

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

a. 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 $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_4$.

b. Calculate the sample covariance matrix $\mathbf{S} = \frac{1}{4}\mathbf{X}^T\mathbf{X}$.

c. Calculate the eigenvalues of $\mathbf{S}$. Is $\mathbf{S}$ invertible? If so, find $\mathbf{S}^{-1}$.

d. Find orthonormal eigenvectors of $\mathbf{S}$.

e. What is the best one-dimensional subspace (line) for approximating the centered observations $\tilde{\mathbf{x}}_i$? Draw this line on your plot.

8. *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 $p = 1$.

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

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

9. 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.

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

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

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

11. Recall that the Frobenius norm of a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ is given by $||A|| = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$, the square root of the sum of the squares of the entries of the matrix. Establish the following properties of the Frobenius norm for matrices.

(a) $||\mathbf{A}|| = 0$ if and only if $\mathbf{A} = 0$

(b) $||b\mathbf{A}|| = |b|\,||\mathbf{A}||$

(c) $||\mathbf{A}||^2 = \sum_{i=1}^m ||a_{i\cdot}||^2 = \sum_{j=1}^n ||a_{\cdot j}||^2$. Here $a_{i\cdot}$ denotes the ith row of A, and $a_{\cdot j}$ denotes the jth column of A.

(d) $||\mathbf{AB}|| \leq ||\mathbf{A}||\,||\mathbf{B}||$. Hint: Use Cauchy-Schwarz.

12. Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are orthogonal vectors in $\mathbb{R}^n$. Show that $||\sum_{i=1}^k \mathbf{v}_i||^2 = \sum_{i=1}^k ||\mathbf{v}_i||^2$. Interpret this in terms of the Pythagorean formula relating the length of the hypotenuse of a right triangle to the lengths of the other edges.

13. Show that if $A \in \mathbb{R}^{n\times n}$ is non-negative definite then all its eigenvalues are non-negative. Hint: Apply the definition of non-negative definite to the eigenvectors of A.

14. Let $\mathbf{A}$ and $\mathbf{B}$ be invertible $n \times n$ matrices. Argue that $(\mathbf{A}\,\mathbf{B})^{-1} = \mathbf{B}^{-1}\,\mathbf{A}^{-1}$.

15. Let $\mathbf{A}$ be an $n \times n$ matrix. Show that if $\mathbf{A}$ has rank n then $\mathbf{Ax} = 0$ if and only if $\mathbf{x} = 0$. Hint: If $\mathbf{A}$ has rank n then its columns are linearly independent.

16. Let $A \in \mathbb{R}^{d\times d}$ be symmetric. The spectral theorem tells us that there is an orthonormal basis $v_1, \ldots, v_d$ for $\mathbb{R}^d$ such that each v_i is an eigenvector of A.

a. Show that the $d \times d$ matrix $\Gamma = [v_1, \ldots, v_d]$ is orthogonal, that is $\Gamma^t \Gamma = I$. Note that this implies $\Gamma \Gamma^t = I$, though you do not need to show this.

b. Let $D = \text{diag}(\lambda_1, \ldots, \lambda_d)$ be the $d \times d$ diagonal matrix with D_{ii} equal to the ith eigenvalue of A and all other entries equal to zero. Show that $A\Gamma = \Gamma D$.

c. Conclude from the expression above that A can be written in the form $A = \Gamma D \Gamma^t$

17. In the previous homework problem you showed that any symmetric matrix $A \in \mathbb{R}^{d \times d}$ can be written in the form $A = \Gamma D \Gamma^t$, where $\Gamma \in \mathbb{R}^{d \times d}$ is an orthogonal matrix and $D = \text{diag}(\lambda_1, \ldots, \lambda_d)$ is a diagonal matrix with D_{ii} equal to the ith eigenvalue of A and all other entries equal to zero. Suppose that A is non-negative definite, so that each $\lambda_i \geq 0$. Define $A^{1/2} = \Gamma D^{1/2} \Gamma^t$ where $D^{1/2} = \text{diag}(\lambda_1^{1/2}, \ldots, \lambda_d^{1/2})$. Show that $A^{1/2}$ is symmetric and satisfies $A^{1/2} A^{1/2} = A$.

18. Let $a_1, \ldots, a_n$ be positive numbers. Use the Cauchy-Schwartz inequality for inner products to show that $n^2 \leq (\sum_{k=1}^n a_k)(\sum_{k=1}^n a_k^{-1})$. Hint: Begin with the identity $1 = a_k^{1/2} a_k^{-1/2}$ which holds for $k = 1, \ldots, n$.

19. Let $A, B \in \mathbb{R}^{m \times n}$ be a matrices.

a. Show that $A = B$ iff $Ax = Bx$ for all $x \in \mathbb{R}^n$.

b. Let $v_1, \ldots, v_n$ be a basis for $\mathbb{R}^n$. Show that if $Av_i = Bv_i$ for $1 \leq i \leq n$ then $Ax = Bx$ for all $x \in \mathbb{R}^n$.

20. (Non-negative definite matrices) Recall that a symmetric matrix $\mathbf{A} \in \mathbb{R}^{d \times d}$ is non-negative definite (written $\mathbf{A} \geq 0$) if $\mathbf{u}^t \mathbf{A} \mathbf{u} \geq 0$ for every vector $\mathbf{u} \in \mathbb{R}^d$, and is positive definite (written $\mathbf{A} > 0$) if $\mathbf{u}^t \mathbf{A} \mathbf{u} > 0$ for every non-zero vector $\mathbf{u} \in \mathbb{R}^d$.

a. Show that if a matrix $\mathbf{A} \geq 0$ then its diagonal entries are non-negative. Hint: consider (basis) vectors $\mathbf{u}$ having one component equal to 1 and all other components equal to 0.

b. Show that if $\mathbf{A} \geq 0$ then all its eigenvalues are non-negative.

c. It is tempting to think that if $\mathbf{A} \geq 0$ then all its entries are non-negative, but this is not the case. Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

Show that $\mathbf{A}$ is non-negative definite, but not positive definite. What is the rank of $\mathbf{A}$?

d. Modify the $(1,1)$ entry of $\mathbf{A}$ to produce a positive definite matrix $\mathbf{B}$. What is the rank of $\mathbf{B}$?