Math 318 Homework 8

(1) (**SVD of Symmetric and PSD matrices**)

(a) Compute the SVD of the symmetric matrix (using Julia or otherwise)

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}.$$

(b) If A is a symmetric matrix of size $n \times n$, argue that $\sigma_i = |\lambda_i|$ for all i. Here σ_i is the ith singular value of A and λ_i is the ith eigenvalue of A.

(c) Based on what you just did, how would you convert a diagonalization of a general symmetric matrix C to the SVD of C? Say in words what steps need to be taken.

(d) If A is a PSD matrix of size $n \times n$ then what is the relationship between its singular values and eigenvalues? What is the SVD of A?

(2) (**Rank one matrices**)

(a) Argue that for any two matrices A and B, $\mathrm{rank}(A + B) \leq \mathrm{rank}(A) + \mathrm{rank}(B)$.
Hint: Think about how the dimension of $\mathrm{Col}(A + B)$ relates to the sum of the dimensions of $\mathrm{Col}(A)$ and $\mathrm{Col}(B)$. Also, if S and T are two sets of vectors in $\mathbb{R}^n$ then $\dim(\mathrm{span}\{S \cup T\}) \leq \dim(\mathrm{span}\{S\}) + \dim(\mathrm{span}\{T\})$.

(b) Use SVD to argue that every rank one matrix in $\mathbb{R}^{m\times n}$ is of the form $\mathbf{u}\mathbf{v}^\top$ for $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n$.

(c) Find two rank one matrices whose sum is still rank 1 and two rank one matrices whose sum has rank 2.

(d) If the columns of $A \in \mathbb{R}^{m\times k}$ are $\mathbf{a}_1, \ldots, \mathbf{a}_k$ and the rows of $B \in \mathbb{R}^{k\times n}$ are $\mathbf{b}_1^\top, \ldots, \mathbf{b}_k^\top$ argue that $AB = \mathbf{a}_1\mathbf{b}_1^\top + \mathbf{a}_2\mathbf{b}_2^\top + \cdots + \mathbf{a}_k\mathbf{b}_k^\top$.
Hint: You could show that the (i, j)-entry on the left side is the same at the (i, j)-entry on the right side. Warm up by checking that if

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} e & f & g \\ h & i & j \end{bmatrix}$$

then

$$AB = \begin{bmatrix} a \\ c \end{bmatrix} \begin{bmatrix} e & f & g \end{bmatrix} + \begin{bmatrix} b \\ d \end{bmatrix} \begin{bmatrix} h & i & j \end{bmatrix}.$$

(3) (**Rank one decomposition of symmetric and PSD matrices**)

(a) Argue that all rank one PSD matrices of size $n \times n$ can be written as $\mathbf{b}\mathbf{b}^\top$ for a vector $\mathbf{b} \in \mathbb{R}^n$.

(b) Describe 3×3 PSD matrices of rank 1 that have 0s and 1s on the diagonal.
Hint: Use part (a) to deduce what any rank one matrix looks like, then find necessary conditions for the diagonals of your psd matrix to be 0 or 1. You do not need to explicitly give all matrices but you should argue how many you could get (double counting is ok) and explain how you get them.

(c) Argue that a PSD matrix of rank r is the sum of r rank one PSD matrices. This means that you can get an expression of the form $M_1 + M_2 + \cdots + M_r$ where each M_i is PSD and rank one.

(d) Can a symmetric matrix of rank r also be written as $M_1 + M_2 + \cdots + M_r$ where each M_i is PSD and rank one? If yes, explain your reason. If no, what kind of rank one decomposition is possible?

(4) † (**Projection with an orthonormal basis**) In class we learned that if $V \subseteq \mathbb{R}^n$ is a subspace with basis $\mathbf{a}_1, \ldots, \mathbf{a}_k$ and $A \in \mathbb{R}^{n\times k}$ is the matrix with columns $\mathbf{a}_1, \ldots, \mathbf{a}_k$, then projection onto V is achieved by the linear transformation with matrix $A(A^\top A)^{-1}A^\top$. In this exercise we are going to see how this formula simplifies if we had started with an orthonormal basis of V.

(a) Suppose $\mathbf{q}_1, \ldots, \mathbf{q}_k$ is an orthonormal basis of V and $Q \in \mathbb{R}^{n\times k}$ is the matrix with columns $\mathbf{q}_1, \ldots, \mathbf{q}_k$.

(i) Show that the projection matrix $P = Q(Q^\top Q)^{-1}Q^\top$ is $\mathbf{q}_1\mathbf{q}_1^\top + \mathbf{q}_2\mathbf{q}_2^\top + \cdots + \mathbf{q}_k\mathbf{q}_k^\top$.

(ii) Using (i) compute $\mathrm{proj}_V\mathbf{b}$, the projection of $\mathbf{b} \in \mathbb{R}^n$ onto V. (Your answer should be a linear combination of $\mathbf{q}_1, \ldots, \mathbf{q}_k$.)

(iii) From (ii), what are the coordinates of $\mathrm{proj}_V\mathbf{b}$ in the basis $\mathbf{q}_1, \ldots, \mathbf{q}_k$?

(iv) Use your knowledge of orthogonal projectors to write down the matrix that projects onto $V^\perp$.

(v) Using this projector to find $\mathrm{proj}_{V^\perp}\mathbf{b}$.

(b) Suppose we find additional vectors so that $\mathbf{q}_1, \ldots, \mathbf{q}_k, \mathbf{q}_{k+1}, \ldots, \mathbf{q}_n$ is an orthonormal basis of $\mathbb{R}^n$. Check for yourself that $\{\mathbf{q}_{k+1}, \ldots, \mathbf{q}_n\}$ is an orthonormal basis of $V^\perp$.

(i) Apply what you learned in (a) to the basis $\{\mathbf{q}_{k+1}, \ldots, \mathbf{q}_n\}$ of $V^\perp$ to compute $\mathrm{proj}_{V^\perp}\mathbf{b}$, the projection of $\mathbf{b}$ onto $V^\perp$.

(ii) Equating your answer above and the answer in (a) (v), express $\mathbf{b}$ as a linear combination of $\mathbf{q}_1, \ldots, \mathbf{q}_n$.

(iii) What are the coordinates of $\mathbf{b}$ in the basis $\mathbf{q}_1, \ldots, \mathbf{q}_n$?

(c) (4.1, #17) Let L be the line spanned by $(1,1,1)^\top$.

(i) Find a vector $\mathbf{u}$ so that projection onto L is $\mathbf{x} \mapsto \mathbf{u}\mathbf{u}^\top\mathbf{x}$.

(ii) Compute the projection of $\mathbf{b} = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$ onto L and $L^\perp$. Show all work.

(5) **Two Distances***

What is the largest number of points in $\mathbb{R}^2$ such that any two of them have the same distance? Three points are ok, we can put them at the vertices of an equilateral triangle, but there is no set of four points for which all pairwise distances are the same. Do you see why?

What if we allow **two** possible distances? A regular pentagon has two distances among pairs of vertices: all the diagonals have the same length and all the sides have the same length.

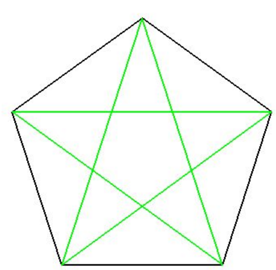

FIGURE 1. A regular pentagon

Question: *What is the maximum number n of points in $\mathbb{R}^d$ such that all pairwise distances among the points are one of two (positive) numbers?*

It seems that n should depend on d, so better to write $n(d)$ instead of n. The above examples are maximal and $n(2) = 5$ which means that in $\mathbb{R}^2$ we can have at most 5 points with two different pairwise distances. In this exercise we will show that $n(d) \le \frac{1}{2}(d^2 + 5d + 4)$.

(a) Let the n points in $\mathbb{R}^d$ be $\mathbf{p}_1, \ldots, \mathbf{p}_n$, and the two allowed distances be a and b. We have that the square of the distance between $\mathbf{p}_i$ and $\mathbf{p}_j$ is

$$\|\mathbf{p}_i - \mathbf{p}_j\|^2 = (p_{i1} - p_{j1})^2 + (p_{i2} - p_{j2})^2 + \cdots + (p_{id} - p_{jd})^2 \in \{a^2, b^2\}$$

Associate to each $\mathbf{p}_i$ the function

$$f_i : \mathbb{R}^d \to \mathbb{R}, \text{ such that } f_i(\mathbf{x}) = (\|\mathbf{x} - \mathbf{p}_i\|^2 - a^2)(\|\mathbf{x} - \mathbf{p}_i\|^2 - b^2)$$

where $\mathbf{x} = (x_1, \ldots, x_d)$. Show that

$$f_i(\mathbf{p}_j) = \begin{cases} 0 & \text{for } i \ne j \\ a^2b^2 & \text{for } i = j \end{cases}$$

Hint: You might choose a few actual points $\mathbf{p}_i$ in $\mathbb{R}^2$ or $\mathbb{R}^3$ and write out the function f_i to get a feel for this question.

Consider the set V of all functions $f : \mathbb{R}^d \to \mathbb{R}$. V is a vector space under the following operations of addition and scalar multiplication. The sum of two functions, $f_1 + f_2$ is defined as $(f_1 + f_2)(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x})$. If f is a function and $\alpha \in \mathbb{R}$, then αf is the function from $\mathbb{R}^d \to \mathbb{R}$ defined as $(\alpha f)(\mathbf{x}) = \alpha(f(\mathbf{x}))$.

(b) Let W be the subspace of V spanned by the functions $f_1, \ldots, f_n$ from (a). Argue that $f_1, \ldots, f_n$ form a basis of W, i.e., they are linearly independent functions in V.
Hint: Suppose they are not, then there is a some linear combination of them $\alpha_1 f_1 + \alpha_2 f_2 + \cdots + \alpha_n f_n = 0$ where 0 is the zero function that sends everything to 0. Use the definition of the functions f_i to show that this forces $\alpha_i = 0$ for all i proving what we want. Think about useful vectors that you could plug into the function $\alpha_1 f_1(\mathbf{x}) + \alpha_2 f_2(\mathbf{x}) + \cdots + \alpha_n f_n(\mathbf{x}) = 0$ to show that $\alpha_i = 0$. A choice of one point will show that one of the $\alpha_i = 0$ and a choice of different point will show that $\alpha_j = 0$ for some $j \ne i$.

(c) Remember we are trying to put an upper bound on $n(d)$. Here is a strategy: suppose we can find another set of functions $g_1, \ldots, g_t$ such that W lies in their span. Argue that $t \ge n(d)$. This means t is an upper bound on $n(d)$.

In the remaining part of this problem we will see how to find such functions $g_1, \ldots, g_t$. Of course we want as small an upper bound as possible so we want a t that is as small as possible.

Note that each f_i is a polynomial of degree 4 in $x_1, \ldots, x_d$. The set of all polynomials in d variables $x_1, \ldots, x_d$, of degree at most 4, is a vector space. Call this vector space P.

(d) Write out all monomials in x_1, x_2 of degree at most 4 and check that all polynomials in x_1, x_2 of degree at most 4 can be written as a linear combination of these monomials.

Let P be the vector space of all monomials of degree at most 4 in $x_1, \ldots, x_d$. The monomials of degree at most 4 form a basis of P since any degree 4 polynomial in d variables can be written as

$$\sum_{i_1 + \cdots + i_d \le 4} a_{i_1, i_2, \ldots, i_d} x_1^{i_1} x_2^{i_2} \cdots x_d^{i_d}$$

There are precisely $\binom{d+4}{4} = \frac{(d+4)(d+3)(d+2)(d+1)}{1 \cdot 2 \cdot 3 \cdot 4}$ monomials of degree at most 4 in d variables. Recall that $\binom{d+4}{4}$ is the number of ways to choose 4 elements from a set of $d+4$ elements.

(e) Using the above, argue that

$$\binom{d+4}{4} \geq n(d).$$

This upper bound is a 4th degree polynomial in d. We need to get to a quadratic in d to get the result we want.

(f) To get a smaller upper bound, we need to look for a smaller set of functions that span W. Maybe we don't need all monomials of degree at most 4 to generate W since the functions f_i have rather special structure. So we should look at it more carefully. Expand f_i and show that it is a linear combination of the following functions:

$$\begin{aligned} (x_1^2 + \cdots + x_d^2)^2 & \\ x_j(x_1^2 + \cdots + x_d^2) & \quad j = 1, 2, \ldots, d \\ x_j^2 & \quad j = 1, 2, \ldots, d \\ x_i x_j & \quad 1 \leq i < j \leq d \\ x_j & \quad j = 1, 2, \ldots, d \\ 1 & \end{aligned}$$

(g) Show that there are $\frac{1}{2}(d^2 + 5d + 4)$ functions in the above list. Why is this number an upper bound on $n(d)$?