Mathematics Department Stanford University
Math 61DM Homework 5

Due at Section, Friday Oct. 26

1. Let V be a subspace of $\mathbb{R}^n$ and let $P_V : \mathbb{R}^n \to \mathbb{R}^n$ be the orthogonal projection of $\mathbb{R}^n$ onto V as defined in lecture. [**Hint: remember that for any subspace** U **and** $u \in \mathbb{R}^n$, $u = \mathcal{P}_U(u) + \mathcal{P}_{U^\perp}(u)$**.**]

   (a) If I is the identity transformation of $\mathbb{R}^n$ (i.e. $I(\underline{x}) = \underline{x} \,\forall\, \underline{x} \in \mathbb{R}^n$), prove that $I - P_V = P_{V^\perp}$.

   (b) If V is the 1-dimensional subspace span$\{\underline{v}\}$, where $\underline{v}$ is a given vector in $\mathbb{R}^n$ with $\|\underline{v}\| = 1$, find (i) the matrix of P_V and (ii) the matrix of $P_{V^\perp}$.

2. Let $(x_1, y_1), \ldots, (x_n, y_n)$ be n points in $\mathbb{R}^2$. This question is about *linear regression*, which means finding real numbers r, t such that the line $y = rx + t$ is the best linear approximation to $(x_1, y_1), \ldots, (x_n, y_n)$, in the sense that the total square error

$$E = \sum_{i=1}^{n} \big(y_i - (rx_i + t)\big)^2$$

   is as small as possible.

   (i) Let

$$A = \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ 1 & 1 & \ldots & 1 \end{pmatrix}^{\mathrm{T}}, \qquad y = \begin{pmatrix} y_1 & y_2 & \ldots & y_n \end{pmatrix}^{\mathrm{T}}.$$

   Show that (r, t) achieves a minimum value of E if and only if $A \begin{pmatrix} r \\ t \end{pmatrix} = \mathcal{P}_{C(A)}(y)$.

   (ii) When is $N(A) = \{0\}$ (i.e. when is the solution (r, t) unique)?

3. Suppose A is an $m \times n$ real matrix. Prove:

   (a) $A^{\mathrm{T}}A$ is positive semi-definite, and

   (b) $A^{\mathrm{T}}A$ is positive definite if $N(A) = \{\underline{0}\}$.

   Here we use the following terminology: an $n \times n$ matrix $B = (b_{ij})$ is positive semi-definite if $\underline{x}^{\mathrm{T}} B \underline{x} \geq 0$ for all $\underline{x} \in \mathbb{R}^n$ and B is positive definite if $\underline{x}^{\mathrm{T}} B \underline{x} > 0$ for all $\underline{x} \in \mathbb{R}^n \setminus \{\underline{0}\}$. Notice that $\underline{x}^{\mathrm{T}} B \underline{x} = \sum_{i,j=1}^{n} b_{ij} x_i x_j$; such an expression is called a *quadratic form*. Similarly, for a linear map $T \in \mathcal{L}(V, V)$ where V is a finite dimensional inner product space, one says that T is positive semi-definite if $Tx \cdot x \geq 0$ for all $x \in V$, and one says that T is positive definite if $Tx \cdot x > 0$ for all $x \in V \setminus \{0\}$.

4. (a) Let $\theta \in [0, 2\pi)$ and let T be the linear transformation of $\mathbb{R}^2$ defined by $T(\underline{x}) = Q(\theta)\underline{x}$, where $Q(\theta)$ is the 2×2 matrix $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Prove that if $\underline{x} = \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix}$ (with $r \geq 0$ and $\alpha \in [0, 2\pi)$) then $T(\underline{x}) = \begin{pmatrix} r\cos(\alpha+\theta) \\ r\sin(\alpha+\theta) \end{pmatrix}$. With the aid of a sketch, give a geometric interpretation of this.

   (b) What is the matrix of the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ which takes the point (x, y) to its "reflection in the line $y = x$" i.e. the transformation $T(x, y) = (y, x)$.

   Caution: In part (b) we are writing points in $\mathbb{R}^2$ as row vectors, but in order to represent T in terms of matrix multiplication you should first rewrite everything in terms of column vectors.

5. Let S be an alphabet with q letters, and consider the set S^n of words of length n over S.

   (a) Prove that if $x \in S^n$ and r is a nonnegative integer, the Hamming ball $B_r(x)$ of radius r (this is the set of words in S^n of distance at most r from x) has $\sum_{k=0}^{r} \binom{n}{k}(q-1)^k$ elements.

   (b) Prove that any code $C \subset S^n$ that corrects for t errors has at most $q^n / \sum_{k=0}^{r} \binom{n}{k}(q-1)^k$ codewords. This is known as the *sphere packing bound.*

6. Let $\ell \geq 2$ be an integer, $n = 2^\ell - 1$, and $k = n - \ell$. Recall that generalized Hamming code $\mathbb{F}_2^n$ has 2^k elements and can correct for a single error. Prove that any code in $\mathbb{F}_2^n$ with more than 2^k elements cannot correct for a single error.

7. Let $C \subset \mathbb{F}_2^9$ consist of all strings $abcabc\bar{a}\bar{b}\bar{c}$ with $a, b, c \in \mathbb{F}_2$ and $\bar{a} = 1 - a$ for $a \in \mathbb{F}_2$.
   (a) How many codewords are in C?
   (b) Is C a linear code? Why?
   (c) How many errors can C detect? [**Hint: A code can detect t errors if $d(C) \geq 2t$.**]

8. Suppose I wanted to send you a message which is just a nonnegative integer that is at most 15. I first convert it to binary as $abcd$ with $a, b, c, d \in \{0, 1\}$, which I then view as an element of $\mathbb{F}_2^4$. Next, using the Hamming code, I encode it as a seven bit codeword in $\mathbb{F}_2^7$. I then transmit it to you along a channel. In the message you receive, at most one error might occur in which one of the bits is misinterpreted. The received word is 1100010.

   (i) How many errors occured in the received word?

   (ii) What was the sent code word?

   (iii) What was the integer I wanted to send you?

   Show your work.