Math 312, Spring 2013 Jerry L. Kazdan

REMARK: We have almost completed Chapter 5, Sections 5.1, 5.2, 5.3, and 5.4 (except for the QR Factorization – which we will skip).

Problem Set 6

DUE: In class Thurs. Mar. 13[*Late papers will be accepted until 1:00 on Friday*].
NOTE: This is now the complete Homework Set 6. Problems 8-10 were added at the end. Despite first appearances, none of these are long.

1. Introduce the following inner product on the space of continuous functions on the interval $-1 \le x \le 1$: $\langle f,\, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx$.

   A real-valued function is called *even* if $f(-x) = f(x)$ for all x, and *odd* if $f(-x) = -f(x)$ for all x. For instance, $2x^4 + x\sin 3x$ is even and $\sin 4x - 7x^5$ is odd. Use the above inner product in the following.

   a) Show that any odd function $f(x)$ is orthogonal to the function 1.
   b) Show that any even function $f(x)$ is orthogonal to $\sin 13x$.
   c) Show that the (pointwise) product $f(x)g(x)$ of an even function $f(x)$ and an odd function $g(x)$ is odd.
   d) Show that any even function $f(x)$ is orthogonal to any odd function $g(x)$.

2. [BRETSCHER, SEC. 5.5 #24]. Using the inner product $\langle f,\, g\rangle := \int_0^1 f(x)g(x)\,dx$, for certain polynomials $\mathbf{f}$, $\mathbf{g}$, and $\mathbf{h}$ say we are given the following table of inner products:

$\langle\ ,\ \rangle$	$\mathbf{f}$	$\mathbf{g}$	$\mathbf{h}$
$\mathbf{f}$	4	0	8
$\mathbf{g}$	0	1	3
$\mathbf{h}$	8	3	50

   For example, $\langle \mathbf{g},\, \mathbf{h}\rangle = \langle \mathbf{h},\, \mathbf{g}\rangle = 3$. Let E be the span of $\mathbf{f}$ and $\mathbf{g}$.

   a) Compute $\langle \mathbf{f},\, \mathbf{g}+\mathbf{h}\rangle$.
   b) Compute $\|\mathbf{g}+\mathbf{h}\|$.
   c) Find $\mathrm{proj}_E\mathbf{h}$. [Express your solution as linear combinations of $\mathbf{f}$ and $\mathbf{g}$.]
   d) Find an orthonormal basis of the span of $\mathbf{f}$, $\mathbf{g}$, and $\mathbf{h}$ [Express your results as linear combinations of $\mathbf{f}$, $\mathbf{g}$, and $\mathbf{h}$.]

3. [LIKE BRETSCHER, SEC. 5.5 #26 & 28]. Use the inner product $\langle f,\, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)\,dx$. Define

$$f(x) = \begin{cases} -1 & \text{if } -\pi < x \le 0, \\ 1 & \text{if } 0 < x \le \pi, \end{cases}$$

   and extend f to all of $\mathbb{R}$ as period is with period 2π: $f(x+2\pi) = f(x)$. This is called a *square wave.*

a) Compute the first N terms in the Fourier Series

$$f(x) = A_0 + \sum_{k=1}^{N} [A_k \cos kx + B_k \sin kx]$$

b) Apply the Pythagorean Theorem 5.5.6 to your answer.

4. [Bretscher, Sec. 5.4 #20] Using pencil and paper, find the least-squares solution to $A\vec{x} = \vec{b}$ where

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \vec{b} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}.$$

5. Use the Method of Least Squares to find the parabola $y = ax^2 + b$ that best fits the following data given by the following four points (x_j, y_j), $j = 1, \ldots, 4$:

$$(-2, 4), \qquad (-1, 3), \qquad (0, 1), \qquad (2, 0).$$

Ideally, you'd like to pick the coefficients a and b so that the four equations $ax_j^2 + b = y_j$, $j = 1, \ldots, 4$ are all satisfied. Since this probably can't be done, one uses least squares to find the best possible a and b.

6. The water level in the North Sea is mainly determined by the so-called M2 tide, whose period is about 12 hours see [`https://en.wikipedia.org/wiki/Tide`]. The height $H(t)$ thus roughly has the form

$$H(t) = c + a\sin(2\pi t/12) + b\cos(2\pi t/12),$$

where time t is measured in hours (note $\sin(2\pi t/12$ and $\cos(2\pi t/12)$ are periodic with period 12 hours). Say one has the following measurements:

t (hours)	0	2	4	6	8	10
$H(t)$ (meters)	1.0	1.6	1.4	0.6	0.2	0.8

Use the method of least squares to find the constants a, b, and c in $H(t)$ for this data.

7. Let A be a real matrix, not necessarily square.

a) Show that both A^*A and AA^* are self-adjoint.

b) Show that $\ker A = \ker A^*A$. [Hint: Show separately that $\ker A \subset \ker A^*A$ and $\ker A \supset \ker A^*A$. The identity $\langle \vec{x},\, A^*A\vec{x} \rangle = \langle A\vec{x},\, A\vec{x} \rangle$ is useful.]

Quadratic Polynomials Using Inner Products

If A is a real symmetric matrix (so it is self-adjoint), then $Q(\vec{x}) := \langle \vec{x}, A\vec{x} \rangle$ is a quadratic polynomial. Given a quadratic polynomial, it is easy to find the (unique) symmetric symmentic matrix A. Here is an example. Say $Q(\vec{x}) := 3x_1^2 - 8x_1x_2 - 5x_2^2$ To find A, note that $-8x_1x_2 = -4x_1x_2 - 4x_2x_2$ so we can rewite Q as

$$Q(\vec{x}) := 3x_1^2 - 4x_1x_2 - 4x_2x_1 - 5x_2^2.$$

If we let

$$A := \begin{pmatrix} 3 & -4 \\ -4 & -5 \end{pmatrix} \quad \text{[Note } A \text{ is a symmetric matrix],}$$

then it is easy to verify that $Q(\vec{x}) = \langle \vec{x}, A\vec{x} \rangle$. In the remaining problems we will use this to help work with quadratic polynomials.

8. In each of these find a 3×3 *symmetric* matrix A so that $Q(\vec{x}) = \langle \vec{x}, A\vec{x} \rangle$.
   a) $Q(\vec{x}) := 3x_1^2 - 8x_1x_2 - 5x_2^2 + x_3^2$.
   b) $Q(\vec{x}) := 3x_1^2 - 8x_1x_2 - 5x_2^2 - x_2x_3 + x_3^2$.
   c) $Q(\vec{x}) := 3x_1^2 - 8x_1x_2 - 5x_2^2 - x_2x_3$.

9. [Lower order terms and Completing the Square] Which is simpler:

$$z = x_1^2 + 4x_2^2 - 2x_1 + 4x_2 + 2 \qquad \text{or} \qquad z = y_1^2 + 4y_2^2 \ ?$$

If we let $y_1 = x_1 - 1$ and $y_2 = x_2 + 1/2$, they are essentially the same. All we did was translate the origin to $(1, -1/2)$.

The point of this problem is to generalize this to quadratic polynomials in several variables. Let

$$\begin{aligned} Q(\vec{x}) &= \sum a_{ij}x_ix_j + 2\sum b_ix_i + c \\ &= \langle \vec{x}, A\vec{x} \rangle + 2\langle b, \vec{x} \rangle + c \end{aligned}$$

be a real quadratic polynomial so $\vec{x} = (x_1, \dots, x_n)$, $\vec{b} = (b_1, \dots, b_n)$ are real vectors and $A = (a_{ij})$ is a real symmetric $n \times n$ matrix.

In the case $n = 1$, $Q(x) = ax^2 + 2bx + c$ which is clearly simpler in the special case $b = 0$. In this case, if $a \neq 0$, by completing the square we find

$$Q(x) = a\left(x + b/a\right)^2 + c - 2b^2/a = ay^2 + \gamma,$$

where we let $y = x - b/a$ and $\gamma = c - b^2/a$. Thus, by translating the origin: $x = y + b/a$ we can eliminate the linear term in the quadatratic polynomial – so it becomes simpler.

a) Similarly, for any dimension n, if A is invertible, using the above as a model, show there is a change of variables $\vec{y} == \vec{x} - \vec{v}$ (this is a translation by the vector $\vec{v}$) so that in the new $\vec{y}$ variables Q hasthe form

$$\hat{Q}(\vec{y}) := Q(\vec{y} + \vec{v}) = \langle \vec{y},\, A\vec{y} \rangle + \gamma \quad \text{that is,} \quad \hat{Q}(\vec{y}) = \sum a_{ij} y_i y_j + \gamma,$$

where γ involves A, b, and c – but no terms that are linear in $\vec{y}$. [In the case $n = 1$, which you should try *first*, this means using a change of variables $y = x - v$ to change the polynomial $ax^2 + 2bx + c$ to the simpler $ay^2 + \gamma$.]

b) As an example, apply this to $Q(\vec{x}) = 2x_1^2 + 2x_1 x_2 + 3x_2 - 4$.

10. For $\vec{x} \in \mathbb{R}^n$ let $Q(\vec{x}) := \langle \vec{x},\, A\vec{x} \rangle$, where A is a real symmetric matrix. We say that A is *positive definite* if $Q(\vec{x}) > 0$ for all $\vec{x} \neq 0$, *negative definite* if $Q(\vec{x}) < 0$ for all $\vec{x} \neq 0$, and *indefinite* if $Q(\vec{x}) > 0$ for some $\vec{x}$ but $Q(\vec{x}) < 0$ for some other $\vec{x}$.

a) In the special case $n = 2$ give (simple!) examples of matrices A that are positive definite, negative definite, and indefinite.

b) In the special case where A is an invertible *diagonal* matrix,

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

under what conditions is $Q(\vec{x})$ positive definite, negative definite, and indefinite? [REMARK: We will see that the general case can *always* be reduced to this special case where A is diagonal.]

Bonus Problem

[Please give this directly to Professor Kazdan]

If you are measuring something like the temperature $T(t)$ at time t, then you know the time t fairly accurately the the measurements of the temperature may have experimental errors. For instance, if we anticipate that $T(t)$ should roughly be a straight line, $T(t) = a + bt$, then in the method of least squares with k data points we measure the square of the discrepancy by

$$Q(a, b) := [T_1 - (a + bt_1)]^2 + [T_2 - (a + bt_2)]^2 + \cdots + [T_k - (a + bt_k)]^2.$$

However, if the data is something like height vs weight, then both the height and weight may have errors, so the above procedure does not seem to be so wise. The next problem presents an alternate method.

1-B Let P_1, P_2,..., P_k be k points (think of them as *data*) in $\mathbb{R}^3$ and let $\mathcal{S}$ be the plane

$$\mathcal{S} := \left\{ X \in \mathbb{R}^3 \, : \, \langle X, \, N \rangle = c \right\},$$

where $N \neq 0$ is a unit vector normal to the plane and c is a real constant.

This problem outlines how to find the plane that *best approximates the data points* in the sense that it minimizes the function

$$Q(N, c) := \sum_{j=1}^{k} \text{distance}\,(P_j, \mathcal{S})^2.$$

Determining this plane means finding N and c.

a) Show that for a given point P, then

$$\text{distance}\,(P, \mathcal{S}) = |\langle P - X, \, N \rangle| = |\langle P, \, N \rangle - c\,|,$$

where X is any point in $\mathcal{S}$

b) First do the special case where the center of mass $\overline{P} := \frac{1}{k}\sum_{j=1}^{k} P_j$ is at the origin, so $\overline{P} = 0$. Show that for any P, then $\langle P, \, N \rangle^2 = \langle N, \, PP^* N \rangle$. Here view P as a column vector so PP^* is a 3×3 matrix.

Use this to observe that the desired plane $\mathcal{S}$ is determined by letting N be an eigenvector of the matrix

$$A := \sum_{j=1}^{k} P_j P_j^T$$

corresponding to it's lowest eigenvalue. What is c in this case?

c) Reduce the general case to the previous case by letting $V_j = P_j - \overline{P}$.

d) Find the equation of the line $ax + by = c$ that, in the above sense, best fits the data points $(-1, 3)$, $(0, 1)$, $(1, -1)$, $(2, -3)$.

e) Let $P_j := (p_{j1}, \ldots, p_{j3})$, $j = 1, \ldots, k$ be the coordinates of the j^{th} data point and $Z_\ell := (p_{1\ell}, \ldots, p_{k\ell})$, $\ell = 1, \ldots, 3$ be the vector of ℓ^{th} coordinates. If a_{ij} is the ij element of A, show that $a_{ij} = \langle Z_i, \, Z_j \rangle$.

[Last revised: May 5, 2013]