Problem Set 7

DUE: In class Thursday,Nov. 1 *Late papers will be accepted until 1:00 PM Friday.*

1. [QUADRATIC POLYNOMIALS]

   a) Find a real symmetric (that is, self-adjoint) 3×3 matrix A so that

   $$\langle \vec{x},\, A\vec{x}\rangle = 3x_1^2 + 4x_1x_2 - x_2^2 - x_2x_3.$$

   SUGGESTION: First do the simpler case of finding a 2×2 matrix A so that

   $$\langle \vec{x},\, A\vec{x}\rangle = 3x_1^2 + 4x_1x_2 - x_2^2.$$

   A simple but useful observation is that $4x_1x_2 = 2x_1x_2 + 2x_2x_1$.

   SOLUTION: $A = \begin{pmatrix} 3 & 2 & 0 \\ 2 & -1 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix}$

   b) [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

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

   be a real quadratic polynomial so $\vec{x} = (x_1, \ldots, x_n)$, $\vec{b} = (b_1, \ldots, b_n)$ 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

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

   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.

   Similarly, for any dimension n, if A is invertible 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 has the form

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

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 *again*, this time using the above suggestion, this means using a change of variables $y = x - v$ to change the polynomial $ax^2 + 2bx + c$ to the simpler $ay^2 + \gamma$.]

SOLUTIONS: First the case $n = 1$ again. Then $Q(x) = Ax^2 + 2bx + c$ so

$$\begin{aligned} Q(x) = Q(y+v) =& A(y+v)^2 + 2b(y+v) + c \\ =& Ay^2 + (2Av + 2b)y + Av^2 + 2bv + c. \end{aligned}$$

To kill the linear term, pick v so that $2Av + 2b = 0$, that is, $v = -b/A$. Then $Q(x) = Ay^2 + \gamma$, where

$$\gamma = Ab^2/A^2 - 2b^2/A + c = -b^2/A + c.$$

Next, the case of arbitrary n. It should now feel routine. We are trying the change of variables $\vec{x} == \vec{y} - \vec{v}$ with the thought of picking $\vec{v}$ to simplify the result. The following should be a straightforward computation (the third line uses $A = A^*$):

$$\begin{aligned} Q(\vec{x}) =& Q(\vec{y} + \vec{v}) = \langle \vec{y} + \vec{v},\, A(\vec{y} + \vec{v})\rangle + \langle \vec{b},\, \vec{y} + \vec{v}\rangle + c \\ =& \langle \vec{y},\, A\vec{y}\rangle + \langle \vec{y},\, A\vec{v}\rangle + \langle \vec{v},\, A\vec{y}\rangle + \langle \vec{v},\, A\vec{v}\rangle + 2\langle \vec{b},\, \vec{y}\rangle + 2\langle \vec{b},\, \vec{v}\rangle + c \\ =& \langle \vec{y},\, A\vec{y}\rangle + \langle 2A\vec{v} + 2\vec{b},\, \vec{y}\rangle + \langle \vec{v},\, A\vec{v}\rangle + 2\langle \vec{b},\, \vec{v}\rangle + c. \end{aligned}$$

The term that is linear in $\vec{y}$ will vanish if we pick $\vec{v}$ so that $2A\vec{v} + 2\vec{b} = 0$, that is, $\vec{v} = -A^{-1}\vec{b}$. Then

$$Q(\vec{x}) = \langle \vec{y},\, A\vec{y}\rangle + \gamma$$

where

$$\gamma = \langle A^{-1}\vec{b},\, \vec{b}\rangle - 2\langle \vec{b},\, A^{-1}\vec{b}\rangle + c = -\langle \vec{b},\, A^{-1}\vec{b}\rangle + c.$$

This agrees with what we found in the special case $n = 1$.

c) As an example, apply this to $Q(\vec{x}) = 2x_1^2 + 2x_1x_2 + 3x_2 - 4$.

SOLUTION: Here $Q(\vec{x}) = \langle \vec{x},\, A\vec{x}\rangle + 2\langle \vec{b},\, \vec{x}\rangle + c$, where $A = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} 0 \\ 3/2 \end{pmatrix}$, and $c = -4$. Thus $A^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix}$ so $\vec{v} = -A^{-1}\vec{b} = \begin{pmatrix} 3/2 \\ -3 \end{pmatrix}$.

2. 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.

   SOLUTION: Several examples. Begin with the polynomial, not the matrix.

positive definite: If $\langle \vec{x},\, A\vec{x} \rangle = x_1^2 + x_2^2$ then A is the identity matrix I, and $\langle \vec{x},\, A\vec{x} \rangle = 2x_1^2 + 3x_2^2$ so $A = \left(\begin{smallmatrix} 2 & 0 \\ 0 & 3 \end{smallmatrix}\right)$.
negative definite: For $\langle \vec{x},\, A\vec{x} \rangle = -x_1^2 - x_2^2$, the matrix is $-I$ while for $\langle \vec{x},\, A\vec{x} \rangle = -2x_1^2 - 3x_2^2$, the matrix is $\left(\begin{smallmatrix} -2 & 0 \\ 0 & -3 \end{smallmatrix}\right)$.
indefinite: For $\langle \vec{x},\, A\vec{x} \rangle = x_1^2 - x_2^2$ the matrix is $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ while for $\langle \vec{x},\, A\vec{x} \rangle = -2x_1^2 + 5x_2^2$ the matrix is $\left(\begin{smallmatrix} -2 & 0 \\ 0 & 3 \end{smallmatrix}\right)$.
NOTE: If $\langle \vec{x},\, A\vec{x} \rangle = 3x_2^2$, the matrix is $A := \left(\begin{smallmatrix} 0 & 0 \\ 0 & 3 \end{smallmatrix}\right)$ is *not* positive definite, it is *positive semi-definite*, that is, $\langle \vec{x},\, A\vec{x} \rangle \ge 0$ for all $\vec{x}$ but $\langle \vec{x},\, A\vec{x} \rangle = 0$ for some $\vec{x} \ne 0$.

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.]

SOLUTION: Key step: here

$$\langle \vec{x},\, A\vec{x} \rangle = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2.$$

If we let $\vec{x} = (0,\, 1,\, 0, \ldots, 0)$, clearly $\langle \vec{x},\, A\vec{x} \rangle = \lambda_2$ so if A is positive definite, then $\lambda_2 > 0$. Similarly, if A is positive definite, then all the λ_j are positive.
Conversely, if all the λ_J are positive, it is clear that A is positive definite.
By the same reasoning, A is negative definite if (and only if) all the $\lambda_j < 0$, and indefinite if at least one λ_j is positive and another is negative.
NOTE: the assumption "A is invertible" implies that none of the λ_j are zero.

3. Let $A(t) = (a_{ij}(t))$ and $B(t) = (b_{ij}(t))$ be $n \times n$ matrices whose elements depend smoothly on the real parameter t. As usual, we define the derivative as

$$A'(t) = \lim_{h \to 0} \frac{A(t+h) - A(t)}{h},$$

assuming the limit exists. It is easy to check that this gives $A'(t) = (a'_{ij}(t))$ (it is the same proof that the derivative of a vector $\vec{x}(t)$ is the derivative of its components).

a) Show that $\dfrac{d}{dt}A(t)B(t) = A'(t)B)t) + A(t)B'(t)$. [The proof is identical to the case $n = 1$ in elementary calculus, with due caution since A and B usually don't commute.]

SOLUTION: Here

$$\frac{d}{dt}A(t)B(t) = \lim_{h\to 0} \frac{A(t+h)B(t+h) - A(t)B(t)}{h}$$

But, just as in the case $n = 1$ (and this is the *key* step), begin with

$$A(t+h)B(t+h) - A(t)B(t) = [A(t+h) - A(t)]B(t+h) + A(t)[B(t+h) - B(t)].$$

Thus

$$\begin{aligned}\frac{d}{dt}A(t)B(t) = &\lim_{h\to 0} \frac{[A(t+h) - A(t)]B(t+h)}{h} + \lim_{h\to 0} \frac{A(t)[B(t+h) - B(t)]}{h} \\ =&A'(t)B(t) + A(t)B'(t)\end{aligned}$$

b) If $A(t)$ is invertible, find the formula for the derivative of $A^{-1}(t)$. [Again, The proof is identical to the case $n = 1$ in elementary calculus – with due caution.]

SOLUTION:
Method 1 In the case $n = 1$ we have

$$\begin{aligned}\frac{1}{h}\left[\frac{1}{f(t+h)} - \frac{1}{f(t)}\right] =&\frac{f(t) - f(t+h)}{[f(t+h)f(t)]h} \\ =&\frac{1}{f(t+h)}\left[\frac{f(t) - f(t+h)}{h}\right]\frac{1}{f(t)}\end{aligned}$$

so, taking the limit as $h \to 0$, we find

$$\frac{d}{dt}\frac{1}{f(t)} = -\frac{f'(t)^2}{f(t)^2}.$$

We slavishly imitate this in the general case:

$$\frac{A^{-1}(t+h) - A^{-1}(t)}{h} = A^{-1}(t+h)\left[\frac{A(t) - A(t+h)}{h}\right]A^{-1}(t).$$

Again, taking the limit as $h \to 0$, we find

$$\frac{dA^{-1}(t)}{dt} = -A^{-1}(t)A'(t)A^{-1}(t). \tag{1}$$

Method 2 Use Part a) to differentiate both sides of the identity $A(t)A^{-1}(t) = I$ to find

$$A'(t)A^{-1}(t) + A(t)\left(A^{-1}(t)\right)' = 0.$$

Solving this for $\left(A^{-1}(t)\right)'$ again gives (1).

4. Combine the rank-nullity Theorem 3.3.7 with Theorem 5.4.1, which says $(\text{im}\, A)^\perp = \ker(A^*)$, to show that $\text{rank}\, A = \text{rank}\, A^*$, that is, $\dim\ \text{im}\, A = \dim\ \text{im}\, A^*$.

   SOLUTION: Say $A : R^n \to \mathbb{R}^k$, so $A^* : \mathbb{R}^k \to \mathbb{R}^n$. Then the Rank-Nullity theorem applied to A and A^* gives

$$n = \dim \text{im} A + \dim \ker A, \qquad \text{and} \qquad k = \dim \text{im} A^* + \dim \ker A^* \tag{2}$$

   Theorem 4.5.1 states that $(\text{im} A)^\perp = \ker A^*$. This, and the same identity interchanging the roles of A and A^*, imply that

$$k - \dim \text{im} A = \dim \ker A^* \qquad \text{and} \qquad n - \dim \text{im} A^* = \dim \ker A. \tag{3}$$

   The first of (2)and the second of (3) show that $\dim \text{im} A = \dim \text{im} A^*$. Note that one can also get this by using the second of (2) and the first of (3).

Bonus Problem

[Please give this directly to Professor Kazdan]

1-B This problem concerns BLOCK MATRICES as discussed in the text on pages 75–77 and pages 87–88. They are often useful to break a problem involving larger matrices into ones with smaller matrices. This technique is essential in the computations Google uses to search the web.

   NOTATION: Let $M = \left(\begin{array}{c|c} \text{A} & \text{B} \\ \hline \text{C} & \text{D} \end{array}\right)$ be an $(n+k) \times (n+k)$ block matrix partitioned into the $n \times n$ matrix A, the $n \times k$ matrix B, the $k \times n$ matrix C and the $k \times k$ matrix D.

   Let $N = \left(\begin{array}{c|c} \text{W} & \text{X} \\ \hline \text{Y} & \text{Z} \end{array}\right)$ is another matrix with the same "shape" as M. The text (p. 75–77) shows that the naive matrix multiplication

$$MN = \left(\begin{array}{c|c} \text{AW+BY} & \text{AX+BZ} \\ \hline \text{CW+DY} & \text{CX+DZ} \end{array}\right)$$

   is correct. In the special case when $C = 0$, the text (p. 87–88) shows that if A is invertible, then M is invertible if (and only if) D is invertible and gives a formula for M^{-1} (note that this is applicable in the special case of upper triangular matrices). Taking the transpose we also get formulas in the special case of M where $B = 0$.

   a) More generally, if A is invertible, show that M is invertible if and only if the matrix $H := D - CA^{-1}B$ is invertible – in which case

$$M^{-1} = \begin{pmatrix} A^{-1} + A^{-1}BH^{-1}CA^{-1} & -A^{-1}BH^{-1} \\ -H^{-1}CA^{-1} & H^{-1} \end{pmatrix}.$$

b) Similarly, if D is invertible, show that M is invertible if and only if the matrix $K := A - BD^{-1}C$ is invertible – in which case

$$M^{-1} = \begin{pmatrix} K^{-1} & -K^{-1}BD^{-1} \\ -D^{-1}CK^{-1} & D^{-1} + D^{-1}CK^{-1}BD^{-1} \end{pmatrix}.$$

c) For which values of a, b, and c is the following matrix invertible? What is the inverse?

$$S := \begin{pmatrix} a & b & b & \cdots & b & b \\ c & a & 0 & & 0 & 0 \\ c & 0 & a & & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ c & 0 & 0 & \cdots & a & 0 \\ c & 0 & 0 & \cdots & 0 & a \end{pmatrix}$$

d) Let the square matrix M have the block form $M := \left(\begin{smallmatrix} A & B \\ C & 0 \end{smallmatrix}\right)$, so $D = 0$. If B and C are square, show that M is invertible if and only if both B and C are invertible, and find an explicit formula for M^{-1}. [Answer: $M^{-1} := \begin{pmatrix} 0 & C^{-1} \\ B^{-1} & -B^{-1}AC^{-1} \end{pmatrix}$].

[Last revised: March 21, 2014]