Assignment 5/MATH 247/Winter 2010

Due: Friday, February 19 *in class (!)*

(*answers will be posted right after class*)

As usual, there are pieces of text, before the questions [1], [2], … *themselves.*

Recall: For the quadratic form

$$q(X) \;=\; q(x_1, x_2, ..., x_n) \;=\; \sum_{i,j=1}^{n} a_{ij} \cdot x_i \cdot x_j \qquad (X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}),$$

we have the matrix-form

$$q(X) \;=\; X^{tr} \cdot A \cdot X \qquad\qquad (A = (a_{ij})^{n\times n}).$$

Moreover, if P is an invertible $n \times n$ matrix, and $X = P \cdot Y$ $(Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix})$, then

$$q(X) \;=\; \hat{q}(Y) \;=\; \hat{q}(y_1, y_2, ..., y_n) \;=\; Y^{tr} \cdot B \cdot Y$$

for the matrix $B \;=\; P^{tr} \cdot A \cdot P$. *We assume that* $A = (a_{ij})^{n\times n}$ *is symmetric*: $a_{ij} = a_{ji}$ for all $i, j = 1,...,n$; as a consequence, then B is also symmetric.

The quadratic form $q(X)$ is 1) *positive definite*, 2) *negative definite*, 3) *indefinite* if, respectively: 1) for all $X \neq 0$, we have $q(X) > 0$; 2) for all $X \neq 0$, we have $q(X) < 0$; 3) there are $X_1 \neq 0,\; X_2 \neq 0$ such that $q(X_1) > 0$ and $q(X_2) < 0$. These three qualities are mutually exclusive; but a quadratic form may be none of the above three kinds.

$q(X)$ is 4) *positive semidefinite*, 5) *negative semidefinite* if, respectively: 4) for all $X \neq 0$, we have $q(X) \geq 0$; 5) for all $X \neq 0$, we have $q(X) \leq 0$. Our text says *non-negative semidefinite* for positive semidefinite.

The terminology is such that "positive semidefinite" is more general than "positive definite"; and similarly for "negative semidefinite". However, the important possibility is that $q(X)$ may be positive semidefinite, without being positive definite. For instance,

$q(x,y) = x^2 + 2xy + y^2$ is positive semidefinite since $q(x,y) = (x+y)^2 \geq 0$ always; but $\begin{pmatrix} 1 \\ -1 \end{pmatrix} \neq 0$ and $q(\begin{pmatrix} 1 \\ -1 \end{pmatrix}) = q(1,-1) = (1-1)^2 = 0$, thus q is not positive definite.

Of course, if q is positive semidefinite, then it cannot be either negative definite, or indefinite.

(In the main application in analysis (calculus), when the quadratic form formed by the second partials at a stationary point is positive definite, one has a local minimum; when the quadratic form is negative definite, one has a local maximum; when it is indefinite, one *definitely* has no local minimum, nor a maximum; and when it is semidefinite but not definite, *one does not know* about local extremum without further investigation. This will be looked at later in class.)

With the quadratic form $q(X) = X^{tr} \cdot A \cdot X$, we associate the symmetric bilinear form $\beta(X_1, X_2)$ defined by $\beta(X_1, X_2) = X_1^{tr} \cdot A \cdot X_2$. β is symmetric in the sense that $\beta(X_1, X_2) = \beta(X_2, X_1)$ We have the identity $q(X) = \beta(X, X)$.

[1] Let $q(X) = q(x,y,z,w) = x^2 - y^2 - 3z^2 + 5w^2 + 5xz + 2xw + wz \quad (X = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix})$.

Throughout this question, q denotes this particular quadratic form.

1) **Write** $q(X)$ in the symmetric matrix form.

2) Apply the coordinate transformation $X = P \cdot Y$, $Y = \begin{pmatrix} r \\ s \\ t \\ u \end{pmatrix}$, with the

transforming matrix

$$P = \begin{pmatrix} 1 & 0 & -5/2 & -17/37 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -8/37 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and **write** $q(X)$ in the form $q(x,y,z,w) = \hat{q}(r,s,t,u) = \sum_{i,j=1}^{4} b_{ij} y_i y_j$ (where we wrote $y_1 = r,\ y_2 = s,\ y_3 = t,\ y_4 = u$) with suitable coefficients b_{ij}. Use precise fraction calculation, rather than decimals; the result should come out precisely *diagonal*.

3) Using the matrix P appropriately, **write** each of the variables r,s,t,u as a linear expression of x,y,z,w; and **verify** *directly* the resulting identity $q(x,y,z,w) = \hat{q}(r,s,t,u)$.

4) Using the work done so far, **decide** what kind q is in the definiteness classification (see above!). **Prove** your assertion: show that q satisfies the *definition* given above of the quality (positive definite, …, indefinite, …). Appealing to some theorem about that quality is not enough.

5) Consider the symmetric bilinear form β associated with q. Let

$$X = \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix}, \quad Y = \begin{pmatrix} -2 \\ 0 \\ 1 \\ 3 \end{pmatrix}.$$

Calculate $\beta(X,X)(=q(X))$, $\beta(Y,Y)$ and $\beta(X,Y)$, and **give** a formula for $q(aX+bY)$ in the form of a homogeneous quadratic polynomial (quadratic form!) of the variables a and b.

The Gram-Schmidt orthogonalization process

Read section 7.7, p. 246, in the text. When you see "inner product space V", take V to be any subspace of a standard space $\mathbb{R}^n$, and for the "inner product " $\langle u,v \rangle$, read dot product: $\langle X,Y \rangle = X^{tr} \cdot Y$. Note Remark 2 on page 246 especially.

Read Example 7.10 on the same page, but omit the next example 7.11 (which is in a vector space of a kind that we have not looked at yet). May also read solved problem 7.21, p.261.

You will need to use Gram-Schmidt in some problems that follow.

Recall: An $n \times n$ matrix P is called *orthogonal* if $P^{tr} \cdot P = I_n$; equivalently, if $P^{-1} = P^{tr}$; and equivalently, if the columns $P_1, ..., P_n$ of P form an *orthonormal* system: $P_i^{tr} \cdot P_j = 0$ when $i \neq j$; and $P_i^{tr} \cdot P_i = \|P_i\|^2 = 1$. If $\mathcal{U} = (u_1, ..., u_n)$ and $\mathcal{V} = (v_1, ..., v_n)$ *both* are *orthonormal* bases of a subspace V of some $\mathbb{R}^m$ (in particular when $V = \mathbb{R}^n$, $\mathcal{U} = \mathcal{E}$, the standard basis, and $\mathcal{V}$ is another *orthonormal* basis of $\mathbb{R}^n$), then $P = [\mathcal{U} \to \mathcal{V}]$, the change-of-basis matrix, is an orthogonal matrix.

[2] **1) Prove** that if P and Q are orthogonal $n \times n$ matrices, then so are P^{-1}, P^{tr}, $P \cdot Q$.

2) Use Gram-Schmidt to **find** a 4×4 orthogonal matrix P the first column of which is $P_1 = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}$.

Instructions: Start with any basis $\mathcal{B}$ of $\mathbb{R}^4$ the first of whose vectors is $2P_1$. You may use some of the standard vectors $e_1, ..., e_4$ as the last three vectors of $\mathcal{B}$. Apply Gram-Schmidt to $\mathcal{B}$. Clear denominators in *intermediate* results too. Use the symbol $\sim$ to denote a scaled vector: e.g., $\begin{pmatrix} 1/2 \\ 3/4 \\ -3/2 \\ 1/4 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 2 \\ 3 \\ -6 \\ 1 \end{pmatrix} \sim \begin{pmatrix} 2 \\ 3 \\ -6 \\ 1 \end{pmatrix}$. Scaling should be used also to get smaller numbers: e.g., $\begin{pmatrix} 3 \\ 6 \\ -9 \\ 12 \end{pmatrix} \sim \begin{pmatrix} 1 \\ 2 \\ -3 \\ 4 \end{pmatrix}$. Normalize the system at the end only. Write precise fractions and square-roots when necessary; do not use decimal approximation.

3) With P determined in 2), and denoting the columns of P by P_1, P_2, P_3, P_4, let $U = \text{span}(P_1, P_2)$. Use 2) (that is: don't start from scratch!) and give a basis of $U^{\perp}$, the orthogonal complement of U.

4) For U as in 3), and $Y = \begin{pmatrix} 7 \\ 2 \\ 0 \\ -1 \end{pmatrix}$, **determine** $\operatorname{proj}_U(Y)$ and $\operatorname{proj}_{U^\perp}(Y)$ (**Hint**: read Theorem 7.8 , page 245 (just before the section on Gram-Schmidt), and the Remark following it. Take note of the term "Fourier coefficient" just before the remark. It will come back later.)

[3] We consider the following quadratic forms:

$$q_1(X) = -x^2 + 3y^2 - z^2 - w^2 + 2xy + 2xz + 2xw - 2yz - 2yw + 6zw \ ,$$

$$q_2(X) = -x^2 + y^2 + 2xy + 2xz + 2xw + 2zw$$

$$q_3(X) = -4x^2 - 3z^2 - 3w^2 - 4yz - 4yw + 2zw$$

$$q_4(X) = 3x^2 + 5y^2 + 5z^2 + 5w^2 - 2yz - 2yw - 2zw$$

Treat each of the quadratic forms $q = q_1,\ q_2,\ \dots$ as follows.

1) Find an orthogonal transforming matrix P that turns $q(X)$ into a pure-square (diagonal) form $\hat{q}(Y)$ in terms of the new variables Y; write out what $\hat{q}(Y)$ is, and what the relationship is between the old and new coordinate vectors X and Y.

2) Determine which kind q is in the definiteness classification. For each case when the q is indefinite, or semidefinite-non-definite, **supply** examples of vectors X that show the property in question.

3) Determine $\max \dfrac{q(X)}{\|X\|^2}$ and $\min \dfrac{q(X)}{\|X\|^2}$.

4) Suppose that A is a symmetric matrix, P is an orthogonal matrix, and $P^{tr} \cdot A \cdot P$ is a diagonal. Let B be a polynomial expression of A:
$B = b_0 I + b_1 A + \dots + b_k A^k$. **Prove** that B is symmetric, and $P^{tr} \cdot B \cdot P$ is diagonal.
(**Hints:** if you find it too hard, verify the assertion first for the special cases
$B = A^2,\ 2A^2,\ I + A\ , I + 3A + 2A^2, \dots$ until the pattern becomes clear.)

5) Consider the quadratic form

$$q_5(X) \quad = \quad 4x^2 + 4y^2 + 3z^2 + 3w^2 + 2xy + 2xz + 2xw + 2yz + 2yw + 4zw \ .$$

Verify that $q_5(X) \ = X^{tr} \cdot B \cdot X$ for the matrix $B = A^2 + A + I$, where A is the matrix for which $q_2(X) \ = X^{tr} \cdot A \cdot X$, with q_2 from above. Using knowledge on $q_2(X)$, **do** the work asked for in 1), 2) and 3) for $q_5(X)$, with a minimal amount of calculation only.

6) Let A be an $n \times n$ symmetric matrix, $q(X) = X^{tr} \cdot A \cdot X$ the corresponding quadratic form. **Prove** that the function $Q(X)$ defined by $Q(X) \ = \ q(A \cdot X - X)$ is a quadratic form, and any orthogonal matrix P that transforms q into pure-squares form, does the same for Q. **Determine** the values μ_i for which $q(X) = \sum_{i=1}^{n} \mu_i y_i^2$ ($X = P \cdot Y$) from the values λ_i for which $q(X) = \sum_{i=1}^{n} \lambda_i y_i^2$.