VIGRE Algebra QR Review, November 4, 2002

Bilinear maps, quadratic forms, and (skew-)symmetric matrices

Let U, V and W be vector spaces over a field k. A *bilinear* map $B : U \times V \to W$ is a function that is linear in each coordinate separately when the other is held fixed. Thus, for all $u, u' \in U$ and $v, v' \in V$ and all scalars $a, a' \in k$,

$$B(au + a'u', v) = aB(u, v) + a'B(u', v) \quad \text{and} \quad B(u, av + a'v') = aB(u, v) + a'B(u, v').$$

Repeated application shows that

$$B(\sum_{i=1}^{m} a_i u_i, \sum_{j=1}^{n} b_j v_j) = \sum_{i,j} a_i b_j B(u_i, v_j).$$

If one has a basis u_i for U and a basis v_j for V then there is a unique bilinear map to W taking specified values w_{ij} on the pairs (u_i, v_j) as i and j run independently through the index sets for the two bases. This is correct even when the bases are infinite.

A bilinear form B on a k-vector space V is a bilinear map $V \times V \to k$. If one chooses an ordered basis v_i for V, and uses it for both copies of V, then B is uniquely determined by the scalars a_{ij} which are its values on the pairs (v_i, v_j). If the basis is finite with n elements, then we can place these values into an $n \times n$ matrix $A = (a_{ij})$. We use the superscript $^{\mathrm{tr}}$ to denote the transpose of a matrix. If we use coordinates in K^n (written as column vectors) with respect to the chosen ordered basis, then its value on (u, v) is $u^{\mathrm{tr}}Av$ (thought of as a scalar rather than as a 1×1 matrix). Thus, once a basis is fixed, there is a bijection between bilinear forms on the n-dimensional vector space V and and $n \times n$ matrices. The bilinear form is called *non-degenerate* if its matrix is invertible. The rank of bilinear form is the rank of its matrix. Non-degenerate forms have maximum rank.

If one changes basis, and C is the invertible matrix such that multiplying by C converts coordinates in the new basis to coordinates in the original basis, then the matrix of B with respect to the new basis is clearly $C^{\mathrm{tr}}AC$. Two $n \times n$ matrices A, A' are called *cogredient* if there is an $n \times n$ invertible matrix C such that $C^{\mathrm{tr}}AC = A'$. This is an equivalence relation, and the equivalence classes classify bilinear forms.

A bilinear form B is called *symmetric* if $B(u, v) = B(v, u)$ for all choices of u and v, and this is equivalent to having a symmetric matrix in the finite-dimensional case. A bilinear form B is called *alternate* if $B(v, v) = 0$ for all $v \in V$. Then

$$0 = B(u + v, u + v) = B(u, u) + B(u, v) + B(v, u) + B(v, v) = B(u, v) + B(v, u),$$

and so $B(u,v) = -B(v,u)$ is implied. Any bilinear form that satisfies $B(u,v) = -B(v,u)$ for all u, v is alternate provided the base field has characteristic different from 2: $B(u,u) = -B(u,u)$ implies $B(u,u) = 0$ except in characteristic 2. A bilinear form is alternate if and only if its matrix satisfies $A^{\mathrm{tr}} = -A$ and the main diagonal consists entirely of 0's: the latter condition is automatic if the characteristic is not 2. We shall call such matrices *alternate*. Except in characteristic 2, this condition is equivalent to skew-symmetry (i.e., $A^{\mathrm{tr}} = -A$).

In the rest of this discussion the bilinear form will be assumed to be either symmetric or alternate. In these cases we say the $u \perp v$ if $B(u,v) = 0$. Under our hypothesis that the form is symmetric or alternate, this relation is symmetric. Note that a non-zero vector can be perpendicular to itself. The usual inner product on $\mathbb{R}^n$ is the bilinear form corresponding to the identity matrix, and is, of course, symmetric.

For every alternate bilinear form B there is a basis such that the matrix of B is the direct sum of 2×2 blocks $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and 1×1 blocks that are (0). Thus, every alternate matrix is cogredient to a matrix of this form. Two alternate matrices are cogredient if and only if they have the same rank, which must be even. The determinant of an alternate matrix must be a perfect square in the field, because cogredience preserves determinants up to multiplication by a nonzero square in the field.

Exercises

(1) Prove that if A is a skew-symmetric or a symmetric matrix over $\mathbb{R}^n$ that maps a subspace V into itself, then A also maps the orthogonal complement of V, with respect to the usual inner product, into itself. (Note that u is orthogonal, in the usual inner product, to the value of A on v iff $u^{\mathrm{tr}}Av = 0$ iff $(u^{\mathrm{tr}}Av)^{\mathrm{tr}} = 0$ iff $v^{\mathrm{tr}}A^{\mathrm{tr}}u = 0$ iff $\pm v^{\mathrm{tr}}Au = 0$ iff v is orthogonal, in the usual inner product, to the value of A on u.)

(2) Recall that, by a previous problem, every real square matrix maps either a one-dimensional or a two-dimensional subspace of $\mathbb{R}^n$ into itself (this was Problem 7. from the session of October 21: one uses the existence of complex eigenvalue and eigenvector). Show that A as in part (1) is orthogonally conjugate (since, when C is orthogonal, $C^{-1} = C^{\mathrm{tr}}$, this also gives cogredience) to a direct sum of matrices, each of which has the same kind of symmetry as A, and each of which is either 2×2 or 1×1. Thus, in the skew-symmetric case, each block is either $\begin{pmatrix} 0 & c \\ -c & 0 \end{pmatrix}$ for some real number c or (0), while in the symmetric case each block is either of the form $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ or is 1×1.

(3) Show that a two by two symmetric matrix over $\mathbb{R}$ always has a real eigenvalue, and, hence, can be diagonalized orthogonally. Conclude that every real symmetric matrix is orthogonally conjugate to a diagonal matrix. In particular, a real symmetric matrix of arbitrary size has real eigenvalues.

(4) Use the result of (2) to show that every real skew-symmetric matrix has pure imaginary eigenvalues ($0 = 0i$ is included among the pure imaginary numbers.)

(5) An orthogonal matrix over the real numbers also has the property that if it maps a subspace V of $\mathbb{R}^n$ into itself, then it maps the orthogonal complement of V into itself as well. Using this and the fact stated at the beginning of (2), show that every orthogonal matrix over $\mathbb{R}$ is a direct sum of orthogonal matrices each of which is either a 2×2 block or a 1×1 block, (± 1).

(6) One can define an inner product on $\mathbb{C}^n$ by letting the product (u, v) of $u = (u_j)$ and $v = (v_j)$ be $\sum_{j=1}^n u_j \overline{v_j}$, where the bar indicates complex conjugation. This inner product is not bilinear: it is linear in the first coordinate, but conjugate linear in the second, i.e., $(u, cv) = \overline{c}(u, v)$. This gives a notion of orthogonality and a norm. Every subspace of $\mathbb{C}^n$ has a unitary basis, i.e., a basis of mutually orthogonal vectors of length 1. Denote by A^* the conjugate of the transpose of the $n \times n$ matrix A. Then the columns of A form a unitary basis for $\mathbb{C}^n$ iff $A^*A = I$ iff $AA^* = I$. Such a matrix is called *unitary*. A matrix A is called *Hermitian* if $A^* = A$. A real unitary matrix is orthogonal. A real Hermitian matrix is symmetric. Show that every Hermitian matrix is conjugate, by a unitary matrix, to a diagonal matrix with real eigenvalues. Show that every unitary matrix is conjugate, by a unitary matrix, to a diagonal matrix whose eigenvalues have absolute value 1. (Imitate the argument in (2) for the Hermitian case and the argument in (5) for the unitary case. In a way, matters are simpler over $\mathbb{C}$, because every square matrix has an eigenvalue.)

We now continue with a summary of the theory of symmetric bilinear forms (without proofs). A symmetric bilinear form B always has an orthogonal basis, i.e., a basis consisting of vectors v_j such that $B(v_i, v_j) = 0$ if $v_i \neq v_j$. This means that it has a matrix that is diagonal, and for a suitable ordering of the basis the nonzero diagonal entries will precede the diagonal entries. Over an algebraically closed field, the nonzero diagonal entries can all be chosen to be 1. Over the real numbers, they call all be chosen to be 1 or -1.

A symmetric bilinear form over $\mathbb{R}$ is called *positive definite* if $B(u, u) > 0$ unless $u = 0$ and *negative definite* if $B(u, u) < 0$ unless $u = 0$. The form is called positive (resp., negative) *semidefinite* if $B(u, u) \geq 0$ (resp., ≤ 0) for all u. What do these conditions say about the eigenvalues?

Over the real numbers, every symmetric bilinear form B has a unique matrix that is diagonal, and such that the diagonal entries consist of a string of 1's, a string of -1's, and then a string of 0's. Some of these strings may be empty. This is called Sylvester's theorem. Thus, every real symmetric matrix is cogredient to a matrix of the type described. Let p, q, and r be the numbers of 1's, -1's, and 0's on the diagonal in such a matrix. The rank of the matrix is $p + q$. Some authors call $p - q$ the *signature* of the matrix. The rank and signature determine the cogredience class (assuming that one knows the size). Other authors refer to (p, q, r) as the signature, which certainly determines the cogredience class. Here p represents the dimension of the maximum subspace on which the restriction of the form is positive definite, q represents the dimension of the maximum subspace on which the restriction of the form is negative definite, and r represents the dimension of the subspace of all vectors that are perpendicular under B to every other vector (there may be a larger subspace on which the form is identically 0).

A quadratic form on a vector space V over k is a function $Q : V \to k$ which, when represented in coordinates, is given by a homogeneous polynomial of degree 2 in the coordinates. Thus, if $x_1, \ldots, x_n$ represent the variable coordinates, there is a formula for the function of the form $Q(x_1, \ldots, x_n) = \sum_{1\le i\le j\le n} c_{ij}x_ix_j$, where the c_{ij} are scalars in k. If one uses a different basis, there is a linear change of coordinates, but the function is still represented by a homogeneous polynomial of degree 2. If B with matrix (a_{ij}) is a bilinear form on k, then $Q(x) = B(x,x)$ is a quadratic form with $Q(x_1, \ldots, x_n) = \sum_{1\le i,j\le n} a_{ij}x_ix_j$. Note that the coefficient of x_ix_j for $i < j$ (after one combines terms) is $a_{ij} + a_{ji}$, while the coefficient of x_i^2 is a_{ii}.

For the rest of this discussion we assume that the field k does not have characteristic 2. We want to show that every quadratic form arises from a unique bilinear form. Given a quadratic form with coefficients c_{ij} as above, let $a_{ii} = c_{ii}$ for every i, but let $a_{ij} = a_{ji} = \frac{1}{2}c_{ij}$ for all $i < j$. The matrix $A = (a_{ij})$ gives a symmetric bilinear form B, and one calculates that $B(x, x) = x^{\mathrm{tr}}Ax = Q(x)$, noting that $\frac{1}{2}a_{ij}x_ix_j + \frac{1}{2}a_{ij}x_jx_i = a_{ij}x_ix_j$. But one can recover B from Q as follows:

$$Q(u+v) = B(u+v, u+v) = B(u,u)+B(u,v)+B(v,u)+B(v,v) = Q(u)+2B(u,v)+Q(v),$$

and so $B(u,v) = \frac{1}{2}\big(Q(u+v) - Q(u,u) - Q(v,v)\big)$. Thus, over a field of characteristic different from 2, there is a bijection between bilinear forms and quadratic forms.[1] The effect of change of basis is captured by the notion of cogredience of matrices, and the classification of quadratic forms is equivalent to the classification of symmetric matrices. In particular, for every quadratic form there is a basis for which it is a linear combination of squares of variables. Over an algebraically closed field the coefficients may all be taken to be 1 or 0. Over the reals, the coefficients may all be taken to be -1, 1, or 0. Over the reals one may even choose an orthonormal basis such that the form is a linear combination of squares of variables (but the coefficients might be any real numbers). All the terminology of bilinear forms (rank, signature, positive and negative definiteness, etc.) immediately transfers to quadratic forms.

Problems new this week

1. For each real number λ, find the rank and signature of the real quadratic form

$$x^2 + y^2 + z^2 + 4xy + 2xz + 2\lambda yz.$$

(2)(a) Up to the equivalence relation that comes from linear changes of basis, how many equivalence classes of quadratic forms are there on a real vector space of dimension n. Your answer should be a function of n, and should be explained.

(b) If x, y, u, v are variables and $a > 0$, b, c, $d > 0$, e, f are real constants such that $b^2 > 4ac$ and $e^2 < 4df$, what are the rank and signature of the quadratic form $ax^2 + bxy + cy^2 + du^2 + euv + fv^2$?

[1] Because it makes sense even in characteristic 2, many authors define the bilinear form associated with a quadratic form Q to be $B(u,v) = Q(u+v, u+v) - Q(u,u) - Q(v,v)$. If one does this, then $Q(u) = \frac{1}{2}B(u,u)$.

3. Consider finite dimensional vector spaces over the real numbers with symmetric bilinear forms. Sylvester's theorem assigns to each form a signature (p, q, r) where p, q, r give the numbers of vectors of positive, negative, and zero norm, resp., in an orthogonal basis. Suppose that the real vector space V has a form with signature (p, q, r), and the its subspace W has signature (p', q', r') with respect to the restriction of this form. Prove or give a counterexample to each of the statements: (1) $p' \leq p$, (2) $q' \leq q$, and (3) $r' \leq r$.

4. Let A be a skew-symmetric real $n \times n$ matrix. Prove that (a) $\det(A + I_n) \geq 1$ and that (b) for any positive definite symmetric $n \times n$ matrix B, $\det(A + B) \geq \det B$. [The examiners gave the following hint: Use that a hermitian matrix (a complex matrix whose complex conjugate is its transpose) has real eigenvalues. The problem can be done in other ways, but some theorem is needed.]

5. Let V be the real vector space of all 2×2 real matrices.

(a) Prove that the operation $(A, B) = \text{trace}\,(AB)$ is a bilinear form on V. Is it symmetric?

(b) Find a basis for V containing the identity matrix and consisting of matrices mutually orthogonal with respect to this form. Evaluate (A, A) for each element A of your basis.

(c) Find the rank and the signature of this form on V and also on the subspace W of V consisting of all matrices of trace zero.

6. Let $\mathbb{R}$ be the field of real numbers and let n be a positive integer. Let V_n be the vector space of bilinear maps from $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^3$. Let W_n be the subspace of V consisting of bilinear maps T that satisfy $T(u, v) = -T(v, u)$ for all choices of u and v in $\mathbb{R}^n$. Determine, with justification, formulas for the dimensions of V and W as $\mathbb{R}$-vector spaces.

7. Let a be a variable real number. Consider the matrix $M_a = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & a \end{pmatrix}$. Determine, as a varies, the rank and signature of this matrix. For which values of a is it diagonalizable? For which values of a is it positive definite?

8. Let A be an $n \times n$ skew-symmetric real matrix.

(a) Prove or disprove: If A is a matrix of a non-degenerate alternating bilinear form, then every non-zero eigenvalue of A is purely imaginary.

(b) Prove or disprove: If n is odd, the determinant of A is zero.

9. Determine, with proof, for which positive integers n the finite field $\text{GF}(5^n)$ with 5^n elements has the property that every element has a unique cube root.