ISSN 0032-9460, Problems of Information Transmission, 2009, Vol. 45, No. 3, pp. 258–263. © Pleiades Publishing, Inc., 2009.
Original Russian Text © A.V. Seliverstov, V.A. Lyubetsky, 2009, published in Problemy Peredachi Informatsii, 2009, Vol. 45, No. 3, pp. 73–78.

LARGE SYSTEMS

On Symmetric Matrices with Indeterminate Leading Diagonals[1]

A. V. Seliverstov and V. A. Lyubetsky

Kharkevich Institute for Information Transmission Problems, RAS, Moscow
slvstv@iitp.ru lyubetsk@iitp.ru

Received December 4, 2008; in final form, April 10, 2009

Abstract—We consider properties of the matrix of a real quadratic form that takes a constant value on a sufficiently large set of vertices of a multidimensional cube centered at the origin given that the corresponding quadric does not separate vertices of the cube. In particular, we show that the number of connected components of the graph of the matrix of such a quadratic form does not change when one edge of the graph is deleted.

DOI: 10.1134/S0032946009030065

We consider a quadratic form $A(\boldsymbol{x}) = \boldsymbol{x}^t A \boldsymbol{x}$, where $\boldsymbol{x}$ runs over the real space of dimension n, $\boldsymbol{x}^t$ denotes transposition, and A is a symmetric real $n \times n$ matrix with entries a_{ij}. We assume throughout what follows that $n \geq 3$. A quadric given by a pair $Q = \langle A, c \rangle$ is the affine variety defined by the equation $A(\boldsymbol{x}) = c$, c being a real number. In fact, we consider the intersection of a quadric Q with the set $\mathbb{U}$ of the *n-cube* whose vertices have coordinates ± 1. We denote the intersection of a quadric Q and the set $\mathbb{U}$ by M_Q. The variable D denotes a diagonal matrix with arbitrary real numbers on the main diagonal, unless otherwise specified. We denote by $G(A)$ the undirected loop-free graph with no multiple edges and with n vertices where the ith vertex is connected by an edge with the jth vertex whenever the entry a_{ij} of A is nonzero. Clearly, the graph $G(A)$ is independent of diagonal entries of A. By rk we denote the rank of a matrix; $|M|$ is the cardinality of a set M.

Properties of $G(A)$ were studied earlier, e.g., what is the smallest rank of a matrix that defines a given graph and how this matrix can be found [1]. The present paper uses an important theorem by Fiedler [2]: *if* $(\forall D)\,(\mathrm{rk}(A + D) \geq n - 1)$, *then there exists a permutation matrix* P *such that* $P^t A P$ *is a tridiagonal irreducible matrix.* Moreover, this holds over any field except for the field of three elements, for which there are some exceptions [3]. Reducibility of a matrix A means that it is block-diagonal, i.e., can be represented as a direct sum of two matrices $A = B \oplus C$.

We are interested in the following problems, partial answers to which are given below. Note that all our proofs, as well as the proof of Fiedler's theorem in [3], are elementary.

1. Find a maximum (or minimum) point of a quadratic form $A(\boldsymbol{x})$ under the constraint $\boldsymbol{x} \in \mathbb{U}$ by means of replacing A with a "simpler" matrix B with the same maximum point. Namely, we consider the transition from A to a matrix $A + D$, the maximum point being unchanged: $(\forall \boldsymbol{x} \in \mathbb{U})\,(\forall D)\,[(A + D)(\boldsymbol{x}) = A(\boldsymbol{x}) + \mathrm{tr}\,D]$, where tr is the trace of a matrix. Here "simple" means that in the general case B is of a lower rank than A. Note that heuristic or approximate (though efficient) optimization algorithms for matrices of low ranks are known [4].

2. Derive properties of the graph $G(A)$ from properties of the set M_Q. Among properties of a graph, one can mention the following: to have at most one connected component (assumed

[1] Supported in part by the International Science & Technology Center, project no. 3807.

throughout to be a *non-one-element* set of graph vertices); to be an open (connected, non-self-intersecting) path passing through all graph vertices; to be a cycle, i.e., a closed (connected, non-self-intersecting) path passing through all graph vertices, etc. Among properties of M_Q we note the following one: a quadric Q *does not separate* vertices of the cube if either $(\forall \boldsymbol{x} \in \mathbb{U})\,(A(\boldsymbol{x}) \geq c)$ or $(\forall \boldsymbol{x} \in \mathbb{U})\,(A(\boldsymbol{x}) \leq c)$. Usually, quadrics are considered that contain all the maximum (or minimum) points of the corresponding quadratic form $A(\boldsymbol{x})$ (with the constraint $\boldsymbol{x} \in \mathbb{U}$); i.e., the form attains the value c. Another condition on M_Q is to contain sufficiently many vertices of the cube. This condition can be expressed by a lower bound on the cardinality of M_Q, or by the condition that M_Q is not contained in any hyperplane, or by the condition of maximality of this set in the partially ordered set described in the next paragraph.

3. Consider a partially ordered set (lower semilattice) consisting of all sets of the form M_Q except for $\mathbb{U}$ with the set-theoretic inclusion relation $\subseteq$. It is desired to describe this semilattice; the first step towards this is describing its maximum elements. We call them *rigid* sets M_Q and call the corresponding quadrics Q *rigid* quadrics. Instead of quadrics centered at the origin, it is of interest to consider other families of hypersurfaces. A trivial example is the family of hypersurfaces given by $A(\boldsymbol{x}) + L(\boldsymbol{x}) = c$, where $L(\boldsymbol{x})$ is a linear form of $\boldsymbol{x}$. This case can be reduced to the case of quadrics centered at the origin by increasing the dimension of the cube by 1: consider quadrics $A(\boldsymbol{x}) + yL(\boldsymbol{x}) = c$. In dimension two (though we do not treat it below), rigid quadrics are diagonal pairs of vertices of the square; such a pair lies on a straight line.

4. Describe the quotient set consisting of equivalence classes of all quadrics with the following equivalence relation: $Q \sim Q'$ if $M_Q = M_{Q'}$.

Below we use the following obvious remarks: $G(A)$ is a connected graph if and only if the matrix A is permutation similar to an irreducible matrix; $G(A)$ is an open non-self-intersecting path if and only if A is permutation similar to a tridiagonal irreducible matrix.

Now we turn to statements and proofs.

Theorem 1. *If a quadric $Q = \langle A, c\rangle$ does not separate vertices of the ± 1-cube and $|M_Q| \geq 3$, then the graph $G(A)$ is not an open path passing through all vertices and $(\exists D)\,(\mathrm{rk}(A+D) \leq n-2)$.*

Proof. Let $G(A)$ be an open path. Without loss of generality we may assume that c is the maximum value of the quadratic form $A(\boldsymbol{x})$ on $\mathbb{U}$ and that the matrix A is tridiagonal and irreducible. At a vertex $\boldsymbol{x} \in \mathbb{U}$, the value of the quadratic form is

$$A(\boldsymbol{x}) = 2\sum_{i=1}^{n-1} a_{i,i+1}x_i x_{i+1} + \operatorname{tr} A;$$

it attains the maximum value at precisely two opposite vertices where each term $a_{i,i+1}x_i x_{i+1}$ is positive. Indeed, the sign of the first coordinate of a vertex can be chosen arbitrarily, and signs of the other coordinates are then determined uniquely since the coefficients $a_{i,i+1}$ are nonzero by irreducibility of A. The contradiction with the condition of the theorem proves that the graph $G(A)$ is not an open path. By Fiedler's theorem, there exists a diagonal matrix D such that the matrix $A + D$ is of rank at most $n - 2$. $\triangle$

Example 1. In Theorem 1, we cannot reduce the bound on the rank to $n - 3$ for any matrix A: the quadric $(x_1+x_2+2x_4)(x_3+x_4) = 0$ does not separate vertices of the cube, and for any matrix D the matrix $A + D$ for this quadric has a 2×2 submatrix with determinant $-1/4$.

Claim. *For any complex matrix A there exists a complex diagonal matrix D such that $A + D$ has the zero eigenvalue of multiplicity $n - 1$.*

Proof. This is obvious for $n = 1$. Let $n \geq 2$. Consider the family of matrices of the form $\{uA + \mathrm{diag}(d_1, \ldots, d_n) \mid u, d_1, \ldots, d_n \in \mathbb{C}\}$, where $\mathrm{diag}(d_1, \ldots, d_n)$ stands for a diagonal matrix.

To any such matrix, we associate its characteristic polynomial

$$x^n + F_{n-1}x^{n-1} + \ldots + F_1 x + F_0 = \det(-uA + \operatorname{diag}(x - d_1, \ldots, x - d_n)),$$

whose coefficients F_k are forms of $u, d_1, \ldots, d_n$. In particular, $F_{n-1} = -d_1 - \ldots - d_n - u \operatorname{tr} A$. By the Hilbert theorem, any $n \geq 2$ forms over $\mathbb{C}$, each of $n+1$ variables, have a common zero different from the origin. In particular, the system of equations

$$\begin{cases} F_{n-1} - u = 0, \\ F_{n-2} = 0, \\ \ldots\ldots\ldots \\ F_1 = 0, \\ F_0 = 0 \end{cases}$$

has a nonzero solution $\check{u}, \check{d}_1, \ldots, \check{d}_n$. Assume that $\check{u} = 0$. Then

$$F_0(\check{u}, \check{d}_1, \ldots, \check{d}_n) = \ldots = F_{n-1}(\check{u}, \check{d}_1, \ldots, \check{d}_n) = 0.$$

Therefore, the characteristic polynomial of $\operatorname{diag}(\check{d}_1, \ldots, \check{d}_n)$ is x^n. This means that $\check{d}_1 = \ldots = \check{d}_n = 0$. But the solution $\check{u}, \check{d}_1, \ldots, \check{d}_n$ is nonzero. The contradiction proves that the system has a solution with $u \neq 0$. By homogeneity, it also has a solution with $u = 1$. Then the corresponding values of $d_1, \ldots, d_n$ satisfy the claim. △

Let us say that a property holds for *almost any* matrix if the set of exceptions is of zero Lebesgue measure. By $[\,\cdot\,]$ we denote the integral part of a number.

Conjecture. *For almost any complex matrix A of order n there exists a complex diagonal matrix D such that* $\operatorname{rk}(A + D) \leq n - [\sqrt{n}]$.

The conjecture is based on the following arguments. If A is a diagonal matrix, we set $D = -A$. Let A be a nondiagonal matrix. Consider the complex projective space $W \cong \mathbb{CP}^{n^2-1}$ of nonzero $n \times n$ matrices up to a scalar factor. The subvariety $W_r \subseteq W$ of matrices of rank at most r is determined by vanishing of all minors of order $r+1$. Its codimension is $(n-r)^2$ (see [5]). With a nondiagonal matrix A of order n, we associate the projective space $L_A \subset W$ given by the equations $a_{ij}x_{k\ell} = a_{k\ell}x_{ij}$ for all $i \neq j$ and $k \neq \ell$. A point in L_A corresponds to nonzero matrices of the form $\lambda(A + D)$ or to diagonal matrices. In particular, for any nondiagonal matrices A and B, the corresponding subspaces L_A and L_B intersect by the $(n-1)$-plane H consisting of diagonal matrices; the embedding $H \subset W$ does not depend on the choice of A. The dimension of L_A is $\dim L_A = n$. If $r = n - [\sqrt{n}]$, the sum of dimensions of W_r and L_A is not less than the dimension of W. By the projective dimension theorem [6], the varieties $W_{n-[\sqrt{n}]}$ and L_A do intersect. An intersection point corresponds to matrices of the form $\lambda A + D$ of rank at most $n - [\sqrt{n}]$. It remains to show that for *almost any* matrix A, in the intersection of $W_{n-[\sqrt{n}]}$ and L_A there exists a point that does not belong to the $(n-1)$-plane H.

Theorem 2. *If a quadric $Q = \langle A, c\rangle$ does not separate vertices of the ± 1-cube and $|M_Q| \geq 2n+1$, then the graph $G(A)$ is not a cycle of length n.*

Proof. In the three-dimensional space, the quadric satisfying the conditions of the theorem passes through all the eight vertices of the cube, and its graph consists of three isolated vertices.

Let $n \geq 4$. We may assume that c is the maximum value of the quadratic form $A(\boldsymbol{x})$ on $\mathbb{U}$. Assume that $G(A)$ is a cycle of length n and vertices of this cycle correspond to coordinate indices. Up to an additive constant, values of the quadratic form at points of $\mathbb{U}$ are equal to the sum of the terms corresponding to edges of the cycle $G(A)$. Then the maximum is attained in at most $2n$

points of $\mathbb{U}$. Indeed, if a value of some coordinate is fixed, then, moving along the cycle $G(A)$, we uniquely determine values of $n-2$ more coordinates by taking the partial sum along the path to be the maximum possible for the previously chosen signs of coordinates. If the initial choice is unfortunate, the sum over the cycle may happen to be less that c. However, every point where the maximum is attained is obtained for some choice of the index and sign of the "starting" coordinate. We may start the cycle traversal from any of the n coordinates. Therefore, arbitrariness of choice remains only for at most $2n$ pairs of an index and sign of a coordinate. △

Example 2. In the condition of Theorem 2, it is not sufficient to assume $|M_Q| \geq 2n$: the quadric $x_1x_2 + \ldots + x_{n-1}x_n - x_nx_1 = n-2$ does not separate vertices of the cube, contains $2n$ vertices, and corresponds to a cycle.

Theorem 3. *If a quadric $Q = \langle A, c\rangle$ does not separate vertices of the ±1-cube and the set M_Q is not contained in any hyperplane, then the graph $G(A)$ and the graph $G(A)$ with one edge deleted have the same number of connected components.*

Proof. Assume that deleting an edge e of $G(A)$ breaks a component of $G(A)$ into two components.

We may assume that c is the maximum value of $A(\boldsymbol{x})$ on $\mathbb{U}$ (since the quadric does not separate vertices) and that the matrix A (up to reindexing the coordinates) is of the form

$$A = \begin{pmatrix} B & S \\ S^t & C \end{pmatrix},$$

where B and C are square symmetric matrices of sizes $k \times k$ and $(n-k)\times(n-k)$ and the rectangular matrix S has a single nonzero entry, which corresponds to the edge e.

Let the only nonzero entry a_{ij} in S be positive, where $i \leq k$ and $j \geq k+1$. The quadratic form $A(\boldsymbol{x})$ attains its maximum value at vertices $\boldsymbol{v}$ where values of the coordinates v_i and v_j coincide. Indeed, assume the contrary: for some vertex $\boldsymbol{v}$, the value of the form is $A(\boldsymbol{v}) = c$ and $v_i = -v_j$. Consider the vertex $\boldsymbol{w}$ with $w_\ell = -v_\ell$ if $1 \leq \ell \leq k$ and $w_\ell = v_\ell$ if $k+1 \leq \ell \leq n$. The value of the form is $A(\boldsymbol{w}) = A(\boldsymbol{v}) + 2a_{ij} > c$. This contradicts the maximality of c. Thus, the form attains the value c in vertices belonging to the hyperplane $x_i = x_j$. This contradicts the condition.

If the only nonzero entry a_{ij} in S is negative, the form attains the value c in vertices that belong to the hyperplane $x_i = -x_j$. This also contradicts the condition. △

Theorem 4. 1. *If a quadric $Q = \langle A \oplus B, c\rangle$ does not separate vertices of the ±1-cube and is rigid, then one of the matrices A or B is diagonal.*

2. *If a quadric $Q = \langle A, c\rangle$ does not separate vertices of the ±1-cube and is rigid, then the graph $G(A)$ has at most one connected component* (*different from an isolated point*).

Proof. Clearly, items 1 and 2 of the theorem are equivalent. Let us prove item 1. We may assume that c is the maximum value of the quadratic form $(A \oplus B)(\boldsymbol{x})$ on $\mathbb{U}$. Since the forms $(A \oplus 0)(\boldsymbol{x})$ and $(0 \oplus B)(\boldsymbol{x})$ depend on different variables, they independently attain their maximum values α and β on $\mathbb{U}$. Hence, $\alpha + \beta = c$. If both matrices A and B are nondiagonal, then the quadrics $(A \oplus 0)(\boldsymbol{x}) = \alpha$ and $(0 \oplus B)(\boldsymbol{x}) = \beta$ pass through proper sets $M_{\langle A\oplus 0, \alpha\rangle}$ and $M_{\langle 0 \oplus B, \beta\rangle}$ of vertices of the cube, which are proper extensions of M_Q. This contradicts the rigidity of M_Q. △

Theorem 5. 1. *If a quadric $Q = \langle A, c\rangle$ is rigid, then it contains at least $n(n-1)$ vertices of the ±1-cube that do not belong to any hyperplane.*

2. *If a quadric $Q = \langle A, c\rangle$ is rigid, then*

$$(\forall Q')\,[M_Q \subseteq M_{Q'} \iff (\exists \mu \neq 0)\,(\exists \lambda, D)\,(Q' = \langle \mu(\lambda A + D), \mu c\rangle)].$$

Proof. 1. Values of a quadratic form at the vertices $\boldsymbol{v}$ and $-\boldsymbol{v}$ coincide. Hence, values of a quadratic form at all vertices are determined by its values at vertices belonging to an arbitrarily chosen face of the cube. By the rigidity, coefficients of the form must satisfy $n(n-1)/2$ linearly independent relations, each of them corresponding to a vertex of the chosen face of the cube. Indeed, extend the set M_Q to M' by adding a new vertex. Denote the corresponding systems of equations in $(n+1)n/2$ matrix elements a_{ij} by $\mathrm{Sys}(M_Q)$ and $\mathrm{Sys}(M')$. They consist of equations of the form

$$\sum_{i=1}^{n} a_{ii} + 2\sum_{i>j} a_{ij}x_ix_j = c.$$

Clearly, the systems $\mathrm{Sys}(M_Q)$ and $\mathrm{Sys}(M')$ are consistent, and their ranks (i.e., ranks of the corresponding matrices with $(n+1)n/2$ columns and, respectively, $|M_Q|$ and $|M'|$ rows) either coincide ($\operatorname{rk}\mathrm{Sys}(M_Q) = \operatorname{rk}\mathrm{Sys}(M')$) or differ by one ($\operatorname{rk}\mathrm{Sys}(M_Q) = \operatorname{rk}\mathrm{Sys}(M') - 1$). If the ranks coincide, solution spaces of the systems $\mathrm{Sys}(M_Q)$ and $\mathrm{Sys}(M')$ coincide. However, the quadric satisfying the system $\mathrm{Sys}(M')$ contains the vertex of M' that do not belong to M. The contradiction shows that $\operatorname{rk}\mathrm{Sys}(M_Q) = \operatorname{rk}\mathrm{Sys}(M') - 1$. By the rigidity of M_Q, the quadric passing through all vertices of M' passes through every vertex of the cube. Since the matrix corresponding to a quadric passing through all vertices of the cube is diagonal, the affine space of solutions of $\mathrm{Sys}(M')$ consists of diagonal matrices with trace c. Hence, its rank is $\operatorname{rk}\mathrm{Sys}(M') = (n(n+1)/2) - (n-1)$. Finally, we get the value $\operatorname{rk}\mathrm{Sys}(M_Q) = n(n-1)/2$. Hence, the cardinality of M_Q is at least $n(n-1)/2$, and the total number of vertices on the quadric, including vertices on the opposite face, is at least $n(n-1)$.

Now assume that vertices from M_Q belong to a hyperplane. By the symmetry around the center of the cube, coordinates of these vertices satisfy a homogeneous linear relation. Without loss of generality, we may assume that

$$x_n = \sum_{i=1}^{n-1} \lambda_i x_i.$$

If we substitute this into the system of linear equations $\mathrm{Sys}(M_Q)$, we obtain a system of equations of the form

$$\sum_{i=1}^{n} a_{ii} + 2\sum_{j=1}^{n-1} a_{nj}\left(\sum_{i=1}^{n-1} \lambda_i x_i x_j\right) + 2\sum_{n>i>j} a_{ij}x_ix_j = c.$$

Here the first n columns, which correspond to the elements a_{ii}, coincide; $n-1$ more columns, which correspond to elements a_{ni}, are linearly expressed through the other columns. Thus, the column rank of the obtained system is at most $((n-2)(n-1)/2)+1$, which is less by at least two than the rank $\operatorname{rk}\mathrm{Sys}(M_Q) = n(n-1)/2$. The contradiction shows that the vertices from M_Q do not belong to a hyperplane.

2. Consider two affine subspaces in the space of symmetric matrices of order n:

$$N = \{B \mid (\forall \boldsymbol{v} \in M_Q)\, B(\boldsymbol{v}) = c\} \qquad \text{and} \qquad L = \{D + \lambda A \in N\}.$$

Dimensions of these subspaces are as follows: $\dim N = n$, since this subspace is determined by the system $\mathrm{Sys}(M_Q)$ of linear equations of rank $n(n-1)/2$ in the space of dimension $n(n+1)/2$, and $\dim L = n$, since the condition $(D+\lambda A)(\boldsymbol{v}) = c$ is equivalent on vertices of the cube to one linear equation $\operatorname{tr} D = c(1-\lambda)$. Since these affine subspaces are embedded ($L \subseteq N$) and their dimensions coincide, the subspaces themselves coincide: $L = N$. △

Example 3. The number of vertices $n(n-1)$ given in item 1 of Theorem 5 is attained for a rigid quadric $(x_1 + x_2 + x_3)^2 = 1$.

The authors are grateful to K.Yu. Gorbunov and S.A. Pirogov for a discussion of results of the paper.

REFERENCES

1. Fallat, S.M. and Hogben, L., The Minimum Rank of Symmetric Matrices Described by a Graph: A Survey, *Linear Algebra Appl.*, 2007, vol. 426, no. 2–3, pp. 558–582.

2. Fiedler, M., A Characterization of Tridiagonal Matrices, *Linear Algebra Appl.*, 1969, vol. 2, no. 2, pp. 191–197.

3. Bento, A. and Leal Duarte, A., On Fiedler's Characterization of Tridiagonal Matrices over Arbitrary Fields, *Linear Algebra Appl.*, 2005, vol. 401, pp. 467–481.

4. Schrijver, A., *Theory of Linear and Integer Programming*, Chichester: Wiley, 1986. Translated under the title *Teoriya lineinogo i tselochislennogo programmirovaniya*, Moscow: Mir, 1991.

5. Harris, J., *Algebraic Geometry: A First Course*, New York: Springer, 1995, 3rd ed. Translated under the title *Algebraicheskaya geometriya. Nachal'nyi kurs*, Moscow: MCCME, 2005.

6. Hartshorne, R., *Algebraic Geometry*, New York: Springer, 1977. Translated under the title *Algebraicheskaya geometriya*, Moscow: Mir, 1981.