

ON PROPERTIES OF QUADRATIC MATRICES

Marek Aleksiejczyk

*Institute of Mathematics and Physics, Pedagogical University of
Olsztyn, ul. Żołnierska 14a, 10-561 Olsztyn, Poland*

Alicja Smoktunowicz

*Institute of Mathematics, Warsaw University of Technology, Pl.
Politechniki 1, 00-661 Warsaw, Poland*

Dedicated to Professor Ludwig Reich on his 60th birthday

Received: September 1999

MSC 2000: 15 A 18

Keywords: Projection, involution, eigenvalues, singular values, best normal approximation, numerical range, Schur decomposition.

Abstract: The aim of this paper is to derive new properties of quadratic matrices, i.e. matrices satisfying a quadratic equation $(A - pI)(A - qI) = 0$ where $p, q \in \mathbb{C}$ and I denotes the identity matrix. We focus our attention to numerical range, singular values of quadratic matrices and the closest normal matrix to a quadratic matrix.

1. Introduction and notations

The purpose of this paper is to give a general characterization of quadratic matrices. The set of all n -by- n matrices over \mathbb{C} is denoted by \mathcal{M}_n . We say that $A \in \mathcal{M}_n$ is a quadratic matrix, if there exist $p, q \in \mathbb{C}$

such that $(A - pI)(A - qI) = 0$ where $I = I_n$ denotes the $n \times n$ identity matrix.

Such matrices find many applications in applied linear algebra. The set of quadratic matrices includes the set of projections ($A^2 = A$), involutions ($A^2 = I$), nilpotents ($A^2 = 0$) and elementary matrices ($A = I - uw^*$), eg. Householder reflections and elimination matrices (see [9]).

For $A \in \mathcal{M}_n$, let $\lambda(A)$ and $\sigma(A)$ denote, respectively, the spectrum and the set of singular values of A . The singular values of A are the positive square roots of the eigenvalues of the Hermitian positive semi-definite matrix A^*A (see eg. [4], [11], [18]). Here $A^* \in \mathcal{M}_n$ stands for the matrix formed by conjugating each element and taking the transpose. The singular values of A can be found from the singular value decomposition SVD (see [4], [11]).

Theorem 1.1 (SVD). *Every matrix $A \in \mathcal{M}_n$ of rank r can be written as $A = U\Sigma V^*$, where $U, V \in \mathcal{M}_n$ are unitary and $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0)$ is a diagonal matrix with nonnegative main diagonal entries σ_i , called the singular values of A .*

Let A^\dagger denote the pseudoinverse of $A \in \mathcal{M}_n$ defined as

$$A^\dagger = V \text{diag} \left(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}, 0, \dots, 0 \right) U^*.$$

Then $X = A^\dagger$ is uniquely determined by the Moore–Penrose conditions: $AXA = A$, $XAX = X$, $AX = (AX)^*$ and $XA = (XA)^*$. For more details we refer the reader to [11].

We will consider the 2-norm (spectral norm) and Frobenius norm of A , i.e. $\|A\|_2 = \sigma_1$ and $\|A\|_F = \sqrt{\sigma_1^2 + \dots + \sigma_r^2}$.

We investigate properties of quadratic matrices using known properties of projections. We summarize some basic properties of projections that are relevant to our discussion in the later sections.

A projection (idempotent) $Z \in \mathcal{M}_n$ is a matrix such that $Z^2 = Z$, so $I - Z$ and Z^* are also projections. It is diagonalizable and $\text{rank}(Z) = \text{tr}(Z)$. The minimal equation of a projection Z is $\lambda^2 = \lambda$, hence Z has as only +1 and 0 as eigenvalues.

In Section 2 we prove some inequalities for singular values of quadratic matrices using the Schur theorem (any complex matrix is unitarily similar to a triangular matrix). If $Z \in \mathcal{M}_n$ is a projection then there exists a unitary matrix $U \in \mathcal{M}_n$ such that

$$(1) \quad Z = U \begin{pmatrix} I_r & C \\ 0 & 0 \end{pmatrix} U^*,$$

where $r = \text{rank}(A)$ (see eg. [7]).

There is a deep overlap between projections and quadratic matrices. Notice that $A \in \mathcal{M}_n$ satisfies

$$(2) \quad (A - pI)(A - qI) = 0, \quad p \neq 0 \neq q$$

iff there exists a projection $Z \in \mathcal{M}_n$ such that

$$(3) \quad A = qI + (p - q)Z.$$

It is obvious that

$$(4) \quad Z = \frac{1}{p - q}(A - qI).$$

In case $p = q$ we have $(A - pI)^2 = 0$, so $A - pI$ is a nilpotent. From (1) and (3) we obtain the following theorem.

Theorem 1.2 (Schur form). *For every $p, q \in \mathbb{C}$ a quadratic matrix $A \in \mathcal{M}_n$ such that $(A - pI)(A - qI) = 0$ can be written as*

$$(5) \quad A = URU^*, \quad R = \begin{pmatrix} pI_r & B \\ 0 & qI_{n-r} \end{pmatrix},$$

where $U \in \mathcal{M}_n$ is unitary.

Th. 1.2 find many applications. Notice that if $p \neq q$ then there exists an involution $X \in \mathcal{M}_n$ containing the eigenvectors of R . Then $X^2 = I$ and

$$X = \begin{pmatrix} I_r & \frac{1}{p-q}B \\ 0 & -I_{n-r} \end{pmatrix}.$$

We see that $RX = XD$ where D is a diagonal matrix:

$$D = \begin{pmatrix} pI_r & 0 \\ 0 & qI_{n-r} \end{pmatrix}.$$

Section 2 is devoted to the inverse eigenproblem of quadratic matrices. In Section 3 we show how to find a quadratic matrix with prescribed singular values. In Section 4 the Moore–Penrose inverse of a quadratic matrix is determined. Section 5 deals with the closest normal matrix X to a quadratic matrix A in the 2-norm and the Frobenius norm. Section 6 discusses the case of elementary matrices. In the last section a characterization of the numerical range of quadratic matrices is given.

A part of this work was presented as a poster of Alicja Smoktunowicz and Marek Aleksiejczyk displayed in the SIAM Applied Linear Algebra Conference in Snowbird, Utah, 1997.

2. Eigenvalues of a product of two quadratic matrices

A quadratic matrix $A \in \mathcal{M}_n$ such that $(A - pI)(A - qI) = 0$ has the minimal equation $(\lambda - p)(\lambda - q) = 0$, hence has only p and q as eigenvalues. We are interested in eigenvalues of a product two quadratic matrices satisfying the same quadratic equation.

First we consider projections. In 1956 Afriat (see [1], [18]) proved the following theorem.

Theorem 2.1. *Let $A = A^* = A^2$ and $B = B^* = B^2$. If λ is an eigenvalue of AB , then $\lambda \in [0, 1]$.*

However, the Afriat theorem need not hold for arbitrary projections A and B . Necessary and sufficient conditions for n numbers in the interval $[0, 1]$ to form the spectrum of a product of two orthogonal projections were determined by Nelson and Neumann (see [17]).

Moreover, the following theorem holds.

Theorem 2.2. *For arbitrary nonzero $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ there exist quadratic matrices $A, B \in \mathcal{M}_{2n}$ satisfying the same quadratic equation*

$$(A - pI)(A - qI) = 0, \quad (B - pI)(B - qI) = 0$$

such that $\lambda_1, \dots, \lambda_n \in \lambda(AB)$.

Proof. Let A and B be the direct sum of matrices A_k and B_k :

$$A = A_1 \oplus A_2 \oplus \dots \oplus A_n, \quad B = B_1 \oplus B_2 \oplus \dots \oplus B_n,$$

where

$$A_k = \begin{pmatrix} p & x_k \\ 0 & q \end{pmatrix}, \quad B_k = \begin{pmatrix} p & 0 \\ 1 & q \end{pmatrix}.$$

We would like to determine $x_k \in \mathbb{C}$ such that $\lambda_k \in \lambda(A_k B_k)$.

An easy calculation gives

$$A_k B_k = \begin{pmatrix} p^2 + x_k & x_k q \\ q & q^2 \end{pmatrix},$$

hence $\lambda_k \in \lambda(A_k B_k)$ iff $(p^2 + x_k - \lambda_k)(q^2 - \lambda_k) - x_k q^2 = 0$. From this we get

$$x_k = \frac{(p^2 - \lambda_k)(q^2 - \lambda_k)}{\lambda_k}.$$

It is obvious that also $\frac{p^2 q^2}{\lambda_k} \in \lambda(A_k B_k)$, so

$$\lambda(AB) = \bigcup_{k=1}^n \lambda(A_k B_k) = \left\{ \lambda_1, \dots, \lambda_n, \frac{p^2 q^2}{\lambda_1}, \dots, \frac{p^2 q^2}{\lambda_n} \right\}.$$

This completes the proof. \diamond

3. Singular values of quadratic matrices

We obtain a generalization of the following estimation of singular values of projections (see eg. [12]): *if $\sigma > 0$ is a singular value of a projection $Z \in \mathcal{M}_n$ then $\sigma \geq 1$.*

Assume that a quadratic matrix A has the Schur form (5).

If $B = U_1 S U_2^*$ is an SVD and $Q = \begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix}$ then Q is unitary and $C = Q^* R Q = \begin{pmatrix} p I_r & S \\ 0 & q I_{n-r} \end{pmatrix}$.

It is evident that there exists a permutation matrix $P \in \mathcal{M}_n$ such that $P^T C P$ is a block diagonal matrix with blocks of size at most two. We may summarize these observations as

Theorem 3.1. *Suppose that $A \in \mathcal{M}_n$ satisfies a quadratic equation $(A - pI)(A - qI) = 0$ where $p, q \in \mathbb{C}$. Then A is unitarily similar to a block diagonal matrix with blocks of size at most two.*

Another proof of this fact was given by A. Zalewska-Mitura and J. Zemánek (see [22]).

Notice that if $X \in \mathcal{M}_2$ has a form $X = \begin{pmatrix} \lambda_1 & s \\ 0 & \lambda_2 \end{pmatrix}$ where $|\lambda_1| \geq |\lambda_2|$, then $\sigma_1 \geq |\lambda_1|$ and $\sigma_2 \leq |\lambda_2|$. Here $\sigma_1 \geq \sigma_2$ are the singular values of X .

From this and Th. 3.1 we get the following theorem.

Theorem 3.2. *Suppose that $A \in \mathcal{M}_n$ satisfies $(A - pI)(A - qI) = 0$ where $p, q \in \mathbb{C}$. If σ is a singular value of A then*

$$(\sigma - |p|)(\sigma - |q|) \geq 0.$$

It is possible to construct a quadratic matrix with prescribed singular values.

Theorem 3.3. *There exists a matrix $A \in \mathcal{M}_{2n}$ satisfying a quadratic equation $(A - pI)(A - qI) = 0$ with given singular values $\sigma_1 \geq \dots \geq \sigma_n > 0$ such that $(\sigma_k - |p|)(\sigma_k - |q|) > 0$ for $k = 1, \dots, n$.*

Proof. Let A be the direct sum of matrices A_k :

$$A = A_1 \oplus A_2 \oplus \dots \oplus A_n,$$

where

$$A_k = \begin{pmatrix} p & s_k \\ 0 & q \end{pmatrix}.$$

We would like to determine $s_k \in \mathbb{R}$ such that $\sigma_k \in \sigma(A_k)$. We have

$$A_k A_k^* = \begin{pmatrix} |p|^2 + s_k^2 & s_k \bar{q} \\ qs_k & |q|^2 \end{pmatrix}.$$

From this we get

$$s_k = \frac{\sqrt{(\sigma_k^2 - |p|^2)(\sigma_k^2 - |q|^2)}}{\sigma_k}.$$

It is clear that also $\frac{|pq|}{\sigma_k} \in \sigma(A_k)$, hence

$$\sigma(A) = \bigcup_{k=1}^n \sigma(A_k) = \left\{ \sigma_1, \dots, \sigma_n, \frac{|pq|}{\sigma_1}, \dots, \frac{|pq|}{\sigma_n} \right\}.$$

This completes the proof. \diamond

4. The Moore–Penrose inverse

Assume that $A \in \mathcal{M}_n$ is a quadratic matrix satisfying a quadratic equation $(A - pI)(A - qI) = 0$, $p, q \in \mathbb{R}$. Then A is nonsingular iff $p \neq 0 \neq q$. From $A^{-1}(A^2 - (p+q)A + pqI) = 0$ we find that

$$(6) \quad A^\dagger = A^{-1} = \frac{1}{pq}((p+q)I - A), \quad p \neq 0 \neq q.$$

From Th. 1.2 we have $A^\dagger = (URU^*)^\dagger = UR^\dagger U^*$, where R has a form

$$(7) \quad R = \begin{pmatrix} pI_r & B \\ 0 & qI_{n-r} \end{pmatrix}.$$

It is easy to check that

$$(8) \quad R^\dagger = R^{-1} = \begin{pmatrix} \frac{1}{p}I_r & \frac{-1}{pq}B \\ 0 & \frac{1}{q}I_{n-r} \end{pmatrix}, \quad p \neq 0 \neq q.$$

If $p = q = 0$ then R is a nilpotent and

$$R^\dagger = \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}^\dagger = \begin{pmatrix} 0 & 0 \\ B^\dagger & 0 \end{pmatrix}.$$

Now we consider the case $q = 0$ and $p \neq 0$. Then using the formula for the Moore–Penrose inverse for partitioned matrices from [15] we get

$$(9) \quad R^\dagger = \begin{pmatrix} pI_r & B \\ 0 & 0 \end{pmatrix}^\dagger = \begin{pmatrix} \frac{1}{p}(I_r + \frac{1}{|p|^2}BB^*)^{-1} & 0 \\ \frac{1}{|p|^2}B^*(I_r + \frac{1}{|p|^2}BB^*)^{-1} & 0 \end{pmatrix}, \quad p \neq 0.$$

5. Nearest normal matrix

We seek normal approximants in the 2-norm and Frobenius norm to a quadratic matrix $A \in \mathcal{M}_n$ such that $(A - pI)(A - qI) = 0$ where $p, q \in \mathbb{C}$ and $p \neq q$. We find a formula for nearest normal matrix in terms of A, p and q .

We remind that a normal matrix $A \in \mathcal{M}_n$ is any matrix satisfying $AA^* = A^*A$. Normal matrices include Hermitian ($A^* = A$), skew-Hermitian ($A^* = -A$), unitary ($A^*A = I$) and real symmetric, skew-symmetric and orthogonal matrices.

Let $A \in \mathcal{M}_n$, and denote by $\nu_2(A)$ ($\nu_F(A)$) its distance from the set of normal matrices in the 2-norm (and Frobenius norm), see [10], [13], [20].

$$(10) \quad \nu_2(A) = \inf \{ \| A - N \|_2 : N \text{ is normal} \},$$

$$(11) \quad \nu_F(A) = \inf \{ \| A - N \|_F : N \text{ is normal} \}.$$

First, we present some results on projections. Assume that $Z \in \mathcal{M}_n$ is a projection (idempotent). It was proved by Phillips (see [19]) that the infimum $\nu_2(Z)$ is attained at $\frac{1}{2}(Z + Z^*)$, a Hermitian part of Z . The same holds for the Frobenius norm (see [3]). Bhatia, Horn and Kittaneh (see [3]) generalized a result of Phillips to the binormal operators with respect to every unitarily invariant norm. We use these results to exhibit nearest normal to a quadratic matrix. It is obvious that if $N \in \mathcal{M}_n$ is a nearest normal to $A \in \mathcal{M}_n$ in the 2-norm or Frobenius norm, then for any $\alpha, \beta \in \mathbb{C}$ the matrix $\alpha N + \beta I$ is a nearest normal to $\alpha A + \beta I$. Now we write a quadratic matrix $A \in \mathcal{M}_n$ such that $(A - pI)(A - qI) = 0$ in a form $A = qI + (p - q)Z$, where Z is a projection, and $Z = \frac{1}{p-q}(A - qI)$.

We have the following theorem (see [10] for a discussion on 2×2 case).

Theorem 5.1. *Assume $A \in \mathcal{M}_n$ satisfies a quadratic equation $(A - pI)(A - qI) = 0$ where $p, q \in \mathbb{C}$ and $p \neq q$. Then*

$$(12) \quad N = qI + \frac{p - q}{2}(Z + Z^*), \quad Z = \frac{1}{p - q}(A - qI)$$

is a nearest normal to A in the 2-norm and the Frobenius norm.

6. Elementary matrices

The set of quadratic matrices includes the set of elementary matrices of a form $A = I - uw^*$, $u, w \in \mathbb{C}^n$.

They are frequently used in numerical linear algebra (see eg. [4], [9], [21]). A matrix $A = I - uw^*$ satisfies a quadratic equation $(A - pI)(A - qI) = 0$ where $p = 1 - w^*u$ and $q = 1$ (see [9]).

We would like to find the singular values of A . We write A^*A in a form

$$A^*A = I - B, \quad B = pw^* + wp^*, \quad p = u - \frac{u^*u}{2}w.$$

We see that we need to determine the eigenvalues of a Hermitian matrix B . By simple computations we get the following result.

Theorem 6.1. *Assume that $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$ are the singular values of $A = I - uw^*$ where $u, w \in \mathbb{C}^n$ and $w^*u \in \mathbb{R}$. Then*

$$(13) \quad \sigma_2 = \dots = \sigma_{n-1} = 1$$

and

$$(14) \quad \sigma_1 = \frac{1}{2}(\sqrt{(u^*u)(w^*w)} + \sqrt{(u^*u)(w^*w) + 4 - 4(w^*u)}),$$

$$(15) \quad \sigma_n = \frac{|1 - w^*u|}{\sigma_1}.$$

7. Numerical range

If $A \in \mathcal{M}_n$ is a given matrix, then the set

$$W(A) = \{x^*Ax : x \in \mathbb{C}^n, x^*x = 1\}$$

is called the *field of values* or *numerical range* of A .

In the following lemma we collect some basic properties of $W(A)$ (see [5], [6], [11], [16]).

Lemma 7.1. *Let $A \in \mathcal{M}_n$. Then*

- (a) $W(A)$ is compact and convex subset of the complex plane.
- (b) $W(UAU^*) = W(A)$ for every unitary $U \in \mathcal{M}_n$.

- (c) $W(\alpha A + \beta I) = \beta + \alpha W(A)$ for every $\alpha, \beta \in \mathbb{C}$.
 (d) If $A \in \mathcal{M}_2$ has eigenvalues $p, q \in \mathbb{C}$ then the numerical range of A is an elliptical disc with eigenvalues of A as foci and a minor axis of the length $\sqrt{\operatorname{tr}(A^*A) - |p|^2 - |q|^2}$.

In 1993 J.I. Fuji and Y. Seo (see [8]) proved the following theorem.

Theorem 7.1. *If $Y \in \mathcal{M}_n$ is a nilpotent of a form $Y = \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}$ then the numerical range of Y is a circular disc centered at 0 and the radius equal to $\frac{1}{2} \|B\|_2$.*

If $Z \in \mathcal{M}_n$ is a projection of a form $Z = \begin{pmatrix} I_r & B \\ 0 & 0 \end{pmatrix}$ then the numerical range of Z is an elliptical disc:

$$W(Z) = \left\{ x + iy : \frac{(x - \frac{1}{2})^2}{a^2} + \frac{y^2}{b^2} \leq 1 \right\}$$

where

$$a = \frac{1}{2} \sqrt{1 + \|B\|_2^2}, \quad b = \frac{1}{2} \|B\|_2.$$

Using (2)–(4), Lemma 7.1, Th. 7.1 and applying the Schur form (5) we get the following theorem.

Theorem 7.2. *Assume that a quadratic matrix $A \in \mathcal{M}_n$ has the Schur form*

$$A = U \begin{pmatrix} pI_r & B \\ 0 & qI_{n-r} \end{pmatrix} U^*$$

where U is unitary. Then the numerical range of A is an elliptical disc with p, q as foci and $\|B\|_2$ as the length of a minor axis.

Acknowledgements. The authors would like to express their gratitude to Professor Jaroslav Zemánek for suggesting the problem and for many stimulating conversations. We are indebted to Professor Lajos László for valuable comments and M-script in MATLAB which allowed us to find the closest normal matrix to a quadratic matrix.

References

- [1] AFRIAT, S.N.: On the Latent Vectors and Characteristic Values of Products of Pairs of Symmetric Idempotents, *Quart. J. Math.* **7** (1956), 76–78.
- [2] ALEKSIEJCZYK, M.: On the diameter of the generalized numerical range, *Demonstratio Math.* **XXX/1** (1997), 129–136.

- [3] BHATIA, R., HORN, R.A. and KITTANEH, F.: Normal approximants to binormal operators, *Linear Algebra Appl.* **147** (1991), 169–179.
- [4] BJÖRCK, .A.: Numerical Methods for Least Squares Problems, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1996.
- [5] BONSALL, F.F., DUNCAN J.: Numerical ranges in studies in functional analysis (editor R. G. Bartle), *MAA Studies in Math.* **21**, 1980.
- [6] DAVIS, C.: The Toeplitz–Hausdorff theorem explained, *Canad. Math. Bull.* **14** (1971), 245–246.
- [7] FUJI, J.I. and FURUTA, T.: Holub’s factorization and normal approximants of idempotent operators, *Math. Japonica* **25/1** (1980), 143–145.
- [8] FUJI, J.I. and SEO, Y.: Graphs and idempotent operators, *Math. Japonica* **37/6** (1992), 1079–1085.
- [9] GASTINEL, N.: Utilisation de matrices vérifiant une équation de degré 2 pour la transmutation de matrices, *C.R. Acad. Sci. Paris* **250** (1960), 1960–1961.
- [10] HIGHAM, N.J.: Matrix nearest problems and applications, In *Applications of Matrix Theory*, M.J.C. Gover and S. Barnett, 1989, 1–27.
- [11] HORN, R.A. and JOHNSON, C.R.: Topics in Matrix Analysis, Cambridge University Press, 1991.
- [12] HOUSEHOLDER, A.S. and CARPENTER, J.A.: The Singular Values of Involutory and of Idempotent Matrices, *Numer. Math.* **5** (1963), 234–237.
- [13] LÁSZLÓ, L.: An attainable lower bound for the best normal approximation, *SIAM J. Matrix Anal. Appl.* **15/3** (1994), 1035–1043.
- [14] LI, C.K.: C–numerical ranges and C–numerical radii, *Linear and Multilinear Algebra* **37** (1994), 51–82.
- [15] MIAO, J.M.: General expressions for the Moore–Penrose inverse of a 2×2 block matrix, *Linear Algebra Appl.* **151** (1991), 1–15.
- [16] MURNAGHAN, E.D.: On the field of values of the square matrix, *Proc. Nat. Acad. Sci. USA* **18** (1932), 246–248.
- [17] NELSON, S. and NEUMANN, M.: Generalizations of the projection method with applications to SOR theory for Hermitian Positive Semidefinite Linear Systems, *Numer. Math.* **51** (1987), 123–141.
- [18] PRASOLOV, V.V.: Problems and Theorems in Linear Algebra, American Math.Society, 1994.
- [19] PHILLIPS, J.: Nearest normal approximation for certain operators, *Proc. Amer. Math. Soc.* **61** (1976), 44–48.
- [20] RUHE, A.: Closest normal matrix finally found!, *BIT* **27** (1987), 585–598.
- [21] SMOKTUNOWICZ, A. and GRABARSKI, A.: Numerically stable deflation of complex symmetric matrices, Manuscript, 1999.
- [22] ZALEWSKA–MITURA, A. and ZEMÁNEK, J.: The Gerschgorin discs under unitary similarity, *Linear Operators, Banach Center Publications* **38**, Warszawa 1997, 427–441.