Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

RATIONAL CRITERIA FOR DIAGONALIZABILITY OF REAL MATRICES THROUGH THE ANALYSIS OF MOMENT AND GRAM MATRICES*

JOÃO FERREIRA ALVES†

Abstract. The purpose of this note is to obtain rational criteria for diagonalizability of real matrices through the analysis of the moment and Gram matrices associated to a given real matrix. These concepts were introduced by Horn and Lopatin in [R.A. Horn and A.K. Lopatin. The moment and Gram matrices, distinct eigenvalues and zeroes, and rational criteria for diagonalizability. *Linear Algebra and its Applications*, 299:153–163, 1999.] for complex matrices. However, when the matrix is real, it is possible to combine their results with the Borchardt-Jacobi Theorem, in order to get new and noteworthy rational criteria.

Key words. Real matrix, Diagonalizability, Moment matrix, Gram matrix, Eigenvalues, Minimal polynomial, Borchardt-Jacobi Theorem.

AMS subject classifications. 15A15, 15A18, 15A21.

1. Introduction. In the literature, the minimal polynomial, $\mu_A(z)$, of an $n \times n$ complex matrix, A, is presented as the central element of the most well-known rational criterion for diagonalizability of complex matrices. In fact, from the formula (see p. 145 of [6])

$$\mu_A(z) = (z-\lambda_1)^{j_1}(z-\lambda_2)^{j_2}\cdots(z-\lambda_d)^{j_d}, \tag{1.1}$$

where $\lambda_1, \ldots, \lambda_d$ denote the distinct eigenvalues of A, and each j_i is the order of the largest Jordan-block of A associated to λ_i $(i = 1, \ldots, d)$, it follows that A is diagonalizable if and only if $\mu_A(z)$ has no multiple zeros, or equivalently, if $d = m$, where m is the degree of $\mu_A(z)$.

So, if we want to decide whether a complex matrix, A, is diagonalizable, we do not need to know its eigenvalues. Instead, we just need to use anyone of the known rational procedures to determine the coefficients of $\mu_A(z)$ and check the formula

$$\gcd(\mu_A(z), \mu_A'(z)) = 1, \tag{1.2}$$

where $\gcd(\mu_A(z), \mu_A'(z))$ denotes the greatest common divisor of the polynomial $\mu_A(z)$ and its derivative, $\mu_A'(z)$.

In order to decide if a given real matrix is diagonalizable over $\mathbb{R}$, further to condition (1.2) we need an additional step. We have to know if all the zeros of $\mu_A(z)$ are real. In other words, we have to check the formula $d_{\mathbb{R}} = m$, where $d_{\mathbb{R}}$ denotes the number of the distinct real eigenvalues of A.

As Abate observed in [1], Sturm Theorem can play an interesting role here. In fact, when $\mu_A(z)$ does not have multiple zeros, we can always use that theorem to determine the number of real zeros of $\mu_A(z)$, obtaining in this way a rational criterion for diagonalizability of real matrices.

*Received by the editors on April 19, 2020. Accepted for publication on August 3, 2020. Handling Editor: Fuad Kittaneh.

†Centro de Análise Matemática Geometria e Sistemas Dinámicos, Department of Mathematics, Technical Institute of Lisbon, University of Lisbon, Av. Rovisco Pais, 1049-001 Lisbon, Portugal (joaolfalves@tecnico.ulisboa.pt). Partially supported by FCT/Portugal through the projects UID/MAT/04459/2019 and UIDB/04459/2020.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

Another classic result in the theory of algebraic equations, which was the culmination of researches of Sturm, Sylvester, Hermite, and others, is the so-called Borchardt-Jacobi Theorem (see [3], [5] or [8]) [1], where the number of distinct roots of a polynomial and its relationship with a specific matrix is studied. Attending its relevance in the results which we will present later and for a matter of completeness, we include it in this section as follows.

Theorem 1.1. (Borchardt-Jacobi) *Let $x_1, \ldots, x_n$ be the roots of a polynomial equation $f(x) = 0$ of degree n with real coefficients. The following statements hold.*

(*i*) *The rank of the matrix*

$$S = \begin{bmatrix} n & s_1 & s_2 & \cdots & s_{n-1} \\ s_1 & s_2 & s_3 & \cdots & s_n \\ s_2 & s_3 & s_4 & \cdots & s_{n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_n & s_{n+1} & \cdots & s_{2n-2} \end{bmatrix},$$

where $s_p = \sum_{i=1}^{n} x_i^p$, is equal to the number of distinct roots of $f(x) = 0$.

(*ii*) *The signature* [2] *of S is equal to the number of distinct real roots of $f(x) = 0$.*

The purpose of this note is to obtain other rational criteria for diagonalizability of a real matrix, A, through the analysis of the moment and Gram matrices associated to A. These concepts were introduced by Horn and Lopatin in [7] for complex matrices. However when the matrix is real, it is possible to combine their results with the Borchardt-Jacobi Theorem, in order to get new rational criteria for diagonalizability of real matrices. In what follows, by diagonalizability of a real matrix (without specifying the field $\mathbb{R}$ or $\mathbb{C}$) we mean diagonalizability over $\mathbb{R}$.

2. Moment and Gram matrices. As we just saw, the diagonalizability over $\mathbb{C}$ or $\mathbb{R}$ of a real matrix is characterized by the relationship between the similarity invariants $d_{\mathbb{R}}$, d and m. In this section, we introduce another similarity invariant, r, in order to obtain other rational criteria for diagonalizability of real matrices based on the relationship between the invariants r and m.

The idea lies in the fact that each of the invariants r, $d_{\mathbb{R}}$, d and m can be determined through moment or Gram matrices. As we will see, this will lead us to two main conclusions. The first tells us that r plays a role similar to the one played by $d_{\mathbb{R}}$ in the sense that a real matrix is diagonalizable over $\mathbb{R}$ if and only if $r = m$. The second conclusion (of computational nature) tells us that among the invariants r, $d_{\mathbb{R}}$, d and m, it is the first whose computation through moment or Gram matrices seems more favorable.

To proceed, it is convenient to introduce some notation.

In what follows, $\mathcal{M}_n$ (resp., $\mathcal{M}_n(\mathbb{R})$) denotes the set of $n \times n$ complex (resp., real) matrices. As before, the number of distinct eigenvalues of $A \in \mathcal{M}_n$ is denoted by d, the degree of the minimal polynomial $\mu_A(z)$ is denoted by m, and the number of distinct real eigenvalues of $A \in \mathcal{M}_n(\mathbb{R})$ is denoted by $d_{\mathbb{R}}$.

[1] In these references, only in [8] we could find this designation for this theorem. The other references to this theorem can be found on p. 202 of [5] (with proof) or more recently on p. 198 of [3] (without proof).

[2] As usual, the signature of a symmetric matrix $S \in \mathcal{M}_n(\mathbb{R})$ is defined by $i_+(S) - i_-(S)$, where $i_+(S)$ (resp., $i_-(S)$) denotes the number of positive (resp., negative) eigenvalues of S.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

By $\mathrm{tr}\,(X)$ and $|X|$ we mean the usual trace and determinant of $X \in \mathcal{M}_n$, and by X^* we mean the conjugate transpose of X.

Let $A \in \mathcal{M}_n$ be a given matrix. The moment matrix of order $p \in \mathbb{N}$ associated to $A \in \mathcal{M}_n$ is defined as the $p \times p$ Hankel matrix

$$K_p\,(A) \equiv \left[\mathrm{tr}\left(A^{i+j-2}\right)\right]_{i,j=1}^{p}.$$

The Frobenius inner product in $\mathcal{M}_n$, $\langle X, Y\rangle = \mathrm{tr}\,(XY^*)$, provides the alternative formula

$$K_p\,(A) = \left[\left\langle A^{i-1}, \left(A^{j-1}\right)^*\right\rangle\right]_{i,j=1}^{p}$$

in analogy with the Gram matrix of $\left\{I, A, \ldots, A^{p-1}\right\}$ with respect to the Frobenius inner product, defined as the $p \times p$ Hermitian matrix

$$L_p\,(A) \equiv \left[\left\langle A^{i-1}, A^{j-1}\right\rangle\right]_{i,j=1}^{p},$$

which we call the Gram matrix of order p associated with A.

So, both matrices, $K_p\,(A)$ and $L_p\,(A)$, are determined by the powers $A^0, \ldots, A^{p-1}$. Despite some analogy between the two definitions, $K_p\,(A)$ has the characteristic of being clearly similarity invariant, which may not happen to $L_p\,(A)$.

At this point, it seems useful to specify the relationship between the matrices $K_p\,(A)$ and $L_p\,(A)$ and the numbers d and m, by recalling preliminarily some of its main properties.

THEOREM 2.1. *For any matrix $A \in \mathcal{M}_n$ the following statements hold.*

(i) $|K_d\,(A)| \neq 0$ *and* $|K_p\,(A)| = 0$ *for all* $p > d$.
(ii) *If all the eigenvalues of A are real, then* $|K_p\,(A)| > 0$ *for all* $p = 1, \ldots, d$.
(iii) $|L_p\,(A)| > 0$ *for all* $p = 1, \ldots, m$, *and* $|L_p\,(A)| = 0$ *for all* $p > m$.

Proof. See Theorems 2, 3 and 7 of [7]. ∎

To point out the implications of this theorem in the context of a real matrix, $A \in \mathcal{M}_n\,(\mathbb{R})$, it should be noted that in such case the matrices $K_p\,(A)$ are also real. Therefore, we can introduce the positive integer

$$r \equiv \min\left\{p = 1, \ldots, n : |K_{p+1}\,(A)| \leq 0\right\},$$

which, by the similarity invariance of the matrices $K_p\,(A)$, is also a similarity invariant. Moreover, as $|K_{d+1}\,(A)| = 0$ (by Theorem 2.1) and $d \leq m$ (by (1.1)), one has

$$r \leq d \leq m, \tag{2.3}$$

for any $A \in \mathcal{M}_n\,(\mathbb{R})$.

In order to establish a first rational criterion for dagonalizability of real matrices based on the invariants r and m, we need a more specific relationship between the moment matrices associated to a real matrix and the invariants d and $d_{\mathbb{R}}$. This is shown in the following theorem through a narrow relationship between the matrix S given in Theorem 1.1 and the moment matrix $K_n\,(A)$.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

THEOREM 2.2. *Let* $A \in \mathcal{M}_n(\mathbb{R})$ *be given. Then* $d = \text{rank}K_n(A)$ *and* $d_{\mathbb{R}} = \text{signature}K_n(A)$*. Moreover, if* $K_p(A)$ *is non-singular for all* $p = r, \ldots, d$*, then* $d_{\mathbb{R}} = d - 2V$*, where* V *is the number of sign variations in the sequence* $|K_r(A)|, \ldots, |K_d(A)|$.

Proof. Let $x_1, \ldots, x_n$ be the eigenvalues of $A \in \mathcal{M}_n(\mathbb{R})$ repeated according to its algebraic multiplicity. As $x_1, \ldots, x_n$ are the roots of the polynomial equation with real coefficients

$$|xI - A| = 0$$

and

$$\text{tr}(A^p) = \sum_{i=1}^{n} x_i^p, \quad \text{for } p \in \mathbb{N},$$

we have that

$$K_n(A) = \begin{bmatrix} n & s_1 & s_2 & \cdots & s_{n-1} \\ s_1 & s_2 & s_3 & \cdots & s_n \\ s_2 & s_3 & s_4 & \cdots & s_{n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_n & s_{n+1} & \cdots & s_{2n-2} \end{bmatrix},$$

with $s_p = \sum_{i=1}^{n} x_i^p$, and by Theorem 1.1, one obtains

$$d = \text{rank}K_n(A) \quad \text{and} \quad d_{\mathbb{R}} = \text{signature}K_n(A).$$

On the other hand, under the assumption that the first d principal submatrices of $K_n(A)$ are non-singular, it follows by Jacobi Theorem (see p 303 of [4]) that

$$\begin{aligned} d_{\mathbb{R}} &= \text{signature}K_n(A) \\ &= \text{rank}K_n(A) - 2V \\ &= d - 2V, \end{aligned}$$

where V denotes the number of sign variations in the sequence

$$1,\ |K_1(A)|,\ \ldots,\ |K_d(A)|.$$

Finally, taking into account that $r \leq d$, $|K_1(A)| = n > 0$, and also $|K_i(A)| > 0$ for $i = 2, \ldots, r$, it follows that V coincides with the number of sign variations in the sequence $|K_r(A)|, \ldots, |K_d(A)|$, as stated. ∎

We are now in a position to prove the main result of this section.

THEOREM 2.3. *A matrix* $A \in \mathcal{M}_n(\mathbb{R})$ *is diagonalizable over* $\mathbb{R}$ *if and only if* $r = m$.

Proof. As $A \in \mathcal{M}_n(\mathbb{R})$ is diagonalizable over $\mathbb{R}$ if and only if $d_{\mathbb{R}} = m$, we just need to prove that $d_{\mathbb{R}} = m$ is equivalent to $r = m$.

Assume $d_{\mathbb{R}} = m$. As $d_{\mathbb{R}} \leq d \leq m$, one obtains $d_{\mathbb{R}} = d = m$. By Theorem 2.1, this implies: $|K_{m+1}(A)| = 0$ and $|K_p(A)| > 0$ for all $p = 1, \ldots, m$. Hence, $r = m$.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

Next assume $r = m$. By (2.3) this implies $r = d = m$. Therefore, as $K_r(A)$ is non-singular (by the definition of r), it follows by Theorem 2.2 that $d_{\mathbb{R}} = d = m$. □

Before concluding this section, observe that Theorem 2.1 enables us to deduce that in analogy with the definition of r, one has

$$\begin{aligned} m &= \min\{p = 1, \ldots, n : |L_{p+1}(A)| \leq 0\} \\ &= \min\{p = 1, \ldots, n : |L_{p+1}(A)| = 0\}. \end{aligned}$$

From this point of view, and because $m \geq r$, it becomes clear that the determination of m (involving the powers $A^0, \ldots, A^m$, required to obtain $L_{m+1}(A)$) is less advantageous than the one of r (which involves the powers $A^0, \ldots, A^r$, required to obtain $K_{r+1}(A)$).

We notice that this last observation points out to an important distinction between the number of distinct eigenvalues of $A \in \mathcal{M}_n(\mathbb{R})$ and the numbers m and r. In fact, by Theorems 2.1 and 2.2, we may write

$$d = \max\{p = 1, \ldots, n : |K_p(A)| \neq 0\} = \operatorname{rank} K_n(A),$$

but, as Horn and Lopatin observe in [7], the equality

$$d = \min\{p = 1, \ldots, n : |K_{p+1}(A)| = 0\}$$

can fail. Thus, unlike the invariants r and m, the determination of d requires the matrix $K_n(A)$, which in turn involves the powers $A^0, \ldots, A^{n-1}$. Of course the same observation applies to the number $d_{\mathbb{R}}$ whose determination through Theorem 2.2 also requires the matrix $K_n(A)$.

Thus, among the invariants r, $d_{\mathbb{R}}$, d and m associated with a given $A \in \mathcal{M}_n(\mathbb{R})$, it is the first whose determination through moment or Gram matrices seems more advantageous.

3. Main criteria and examples. Theorem 2.3 and the observations before suggest that it may be interesting to establish rational criteria for diagonalizability of real matrices underlied by the relationship between the numbers r and m.

Of the three criteria we present in this section, the first two are based, not on the minimal polynomial of the matrix, $A \in \mathcal{M}_n(\mathbb{R})$, but upon the powers A^k, $k = 0, \ldots, r$, whose computation is indissociable from the number r. Two reasons motivate the use of these criteria. The first concerns the difference between the numbers m and r which can be very significant when A has complex eigenvalues or exhibits large Jordan-blocks. The second reason, less obvious, concerns the classification of a diagonalizable matrix: as we will see, these are diagonalizability criteria that simultaneously allow us to decide whether two real matrices, both diagonalizable over $\mathbb{R}$, are similar.

This last aspect is relevant in the classification of large diagonalizable matrices with few eigenvalues and makes a clear distinction between the first two criteria and the last one, which only assumes that the coefficients of $\mu_A(z)$ are known. The idea is to apply Theorem 2.3 to the companion matrix of $\mu_A(z)$, and, like any other criterion of this type, its use is justified whenever the determination of the coefficients of $\mu_A(z)$, through any of the known algorithms, is considerably more advantageous than the computation of the powers of A involved in the previous criteria.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

3.1. The matrices $K_{r+1}(A)$ and $L_{r+1}(A)$. The first of these criteria states that to decide whether a given matrix $A \in \mathcal{M}_n(\mathbb{R})$ is diagonalizable over $\mathbb{R}$, we do not need to determine the coefficients of $\mu_A(z)$, or any other associated polynomial. Instead one can determine the powers A^k, $k = 0, \ldots, r$, in order to get the matrices $K_{r+1}(A)$ and $L_{r+1}(A)$ and check whether their determinants coincide.

THEOREM 3.1. *A matrix $A \in \mathcal{M}_n(\mathbb{R})$ is diagonalizable over $\mathbb{R}$ if and only if $|K_{r+1}(A)| = |L_{r+1}(A)|$.*

Proof. Suppose that $A \in \mathcal{M}_n(\mathbb{R})$ is diagonalizable over $\mathbb{R}$. By Theorem 2.3, this implies $r = m$. Therefore, $|K_{r+1}(A)| = |K_{m+1}(A)|$ and $|L_{r+1}(A)| = |L_{m+1}(A)|$. But by Theorem 2.1, we also have $|K_{m+1}(A)| = 0$ and $|L_{m+1}(A)| = 0$. Hence, $|K_{r+1}(A)| = 0 = |L_{r+1}(A)|$.

Now assume that $|K_{r+1}(A)| = |L_{r+1}(A)|$. As $|K_{r+1}(A)| \leq 0$ (by the definition of r) and $|L_{r+1}(A)| \geq 0$ (by Theorem 2.1), one obtains $|L_{r+1}(A)| = 0$, which (again by Theorem 2.1) implies $r \geq m$. But by (2.3) this implies $r = m$, and by Theorem 2.3, it follows that A is diagonalizable over $\mathbb{R}$. $\square$

The approach based on the determinants $|K_{r+1}(A)|$ and $|L_{r+1}(A)|$, whose calculation only involves sums and products, is, for its conceptual simplicity, worthy of note and gives a new insight into the relationship between diagonalizability and symmetry of real matrices. Indeed, as $K_{r+1}(A) = L_{r+1}(A)$ holds trivially when $A \in \mathcal{M}_n(\mathbb{R})$ is symmetric, the well-known fact that any real symmetric matrix is diagonalizable over $\mathbb{R}$ is an immediate consequence of Theorem 3.1.

By the proof of Theorem 3.1, we may conclude that $|K_{r+1}(A)| = |L_{r+1}(A)|$ if and only if both determinants are null. So, to decide whether a given real matrix A is diagonalizable over $\mathbb{R}$, we do not need to calculate the determinants of $K_{r+1}(A)$ and $L_{r+1}(A)$, but only their signs. In particular we have the following.

COROLLARY 3.2. *Let $A \in \mathcal{M}_n(\mathbb{R})$ be given. If $|K_{r+1}(A)| < 0$, then A is not diagonalizable over $\mathbb{R}$.*

Two examples illustrate the use of Theorem 3.1 and Corollary 3.2 in the setting of a large matrix with complex eigenvalues or large Jordan-blocks, where the difference between m and r may be relevant.

EXAMPLE 3.3. Consider the $n \times n$ matrix

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 1 & a \\ 1 & 1 & \cdots & 1 & 1 & b & 1 \\ 1 & 1 & \cdots & 1 & a & 1 & 1 \\ 1 & 1 & \cdots & b & 1 & 1 & 1 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ 1 & a & \cdots & 1 & 1 & 1 & 1 \\ b & 1 & \cdots & 1 & 1 & 1 & 1 \end{bmatrix},$$

with $n = 4, 6, 8, \ldots$, $a = \frac{n}{2}$ and $b = -2$. As $\operatorname{tr}(A) = n$ and $\operatorname{tr}(A^2) = -n$, one obtains

$$|K_2(A)| = \begin{vmatrix} n & n \\ n & -n \end{vmatrix} < 0.$$

Hence, $r = 1$ and $|K_{r+1}(A)| < 0$, and by Corollary 3.2, it follows that A cannot be diagonalizable over $\mathbb{R}$.

In the previous example, we have $m = 4$ and the coefficients of $\mu_A(z)$ are polynomials in n. Therefore, the determination of $\mu_A(z)$, even for large values of n, would not be difficult and would lead us, in a different

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

way, to the same conclusion. A second example illustrates a different situation, where any attempt to obtain the coefficients of $\mu_A(z)$ would necessarily result in a much more complicated process.

EXAMPLE 3.4. Consider the $n \times n$ matrix

$$A = \begin{bmatrix} 4 & 1 & 1 & \cdots & 1 \\ 2 & 3 & 1 & \ddots & \vdots \\ 1 & 2 & 3 & \ddots & 1 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 2 & 3 \end{bmatrix},$$

with $n > 2$. In this case, the matrix is non-derogatory ($m = n$) and some of the coefficients of $\mu_A(z)$, unlike the previous example, grow exponentially with n. Thus, the computation of $\mu_A(z)$, even for small values of n, can be complicated. Nevertheless the entries of the matrices

$$K_3(A) = \begin{bmatrix} n & 3n+1 & n^2+10n+5 \\ * & * & n^3+9n^2+35n+19 \\ * & * & n^4+12n^3+54n^2+124n+65 \end{bmatrix}$$

and

$$L_3(A) = \begin{bmatrix} n & 3n+1 & n^2+10n+5 \\ * & n^2+11n+4 & n^3+9n^2+39n+18 \\ * & * & n^4+12n^3+54n^2+143n+78 \end{bmatrix}$$

are polynomial, and by Theorem 3.1, this is all we need to conclude that A cannot be diagonalizable over $\mathbb{R}$. Indeed, as $|K_2(A)| > 0$, $|K_3(A)| = 0$ and $|L_3(A)| > 0$, one gets $r = 2$ and $|K_{r+1}(A)| \neq |L_{r+1}(A)|$.

The computational aspects observed in the previous examples stem from the difference between m and r, and therefore (by Theorem 2.3), do not apply when the matrix is diagonalizable over $\mathbb{R}$. Even so, the use of Theorem 3.1 in this setting can be interesting, not only for its conceptual simplicity, but also because, in contrast to any criterion based on the minimal polynomial of the matrix, it plays a relevant role in the classification of a diagonalizable real matrix.

This follows from the next result which (informally) states that when we use Theorem 3.1 to conclude that two real matrices are diagonalizable over $\mathbb{R}$, we have everything we need to see if they are similar or not.

THEOREM 3.5. *Two matrices $A, B \in \mathcal{M}_n(\mathbb{R})$, both diagonalizable over $\mathbb{R}$, are similar if and only if $K_{r+1}(A) = K_{r+1}(B)$.*

Of course for small matrices this may not be very interesting since in this case, we can always use the characteristic polynomial to decide whether two diagonalizable matrices are similar, but for large matrices we have a different situation and, in this case, Theorem 3.5 can be useful when dealing with diagonalizable matrices with few eigenvalues.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

Example 3.6. For the matrix of Example 3.3, with $n = 6, 8, 10, \ldots$, $a = -1$ and $b = 0$, one obtains

$$K_5(A) = \begin{bmatrix} n & * & * & * & n^4 - 6n^3 + \frac{25}{2}n^2 - 8n \\ * & * & * & * & n^5 - \frac{15}{2}n^4 + \frac{85}{4}n^3 - 30n^2 + 20n \\ * & * & * & * & n^6 - 9n^5 + \frac{129}{4}n^4 - \frac{243}{4}n^3 + 63n^2 - 28n \\ * & * & * & * & n^7 - \frac{21}{2}n^6 + \frac{91}{2}n^5 - \frac{861}{8}n^4 + \frac{301}{2}n^3 - 126n^2 + 56n \\ * & * & * & * & n^8 - 12n^7 + 61n^6 - 174n^5 + \frac{2449}{8}n^4 - 342n^3 + 236n^2 - 80n \end{bmatrix}$$

and

$$L_5(A) = \begin{bmatrix} n & * & * & * & n^4 - 6n^3 + \frac{25}{2}n^2 - 8n \\ * & * & * & * & n^5 - \frac{15}{2}n^4 + \frac{89}{4}n^3 - 33n^2 + 24n \\ * & * & * & * & n^6 - 9n^5 + \frac{133}{4}n^4 - \frac{255}{4}n^3 + 67n^2 - 28n \\ * & * & * & * & n^7 - \frac{21}{2}n^6 + \frac{93}{2}n^5 - \frac{897}{8}n^4 + 161n^3 - 138n^2 + 64n \\ * & * & * & * & n^8 - 12n^7 + 62n^6 - 180n^5 + \frac{2585}{8}n^4 - 366n^3 + 252n^2 - 80n \end{bmatrix},$$

where, for ease of writing, some entries have been omitted [3]. As $|K_p(A)| > 0$ for $p = 1, \ldots, 4$, and $|K_5(A)| = |L_5(A)| = 0$, one gets $r = 4$ and $|K_{r+1}(A)| = |L_{r+1}(A)|$. Therefore, A is diagonalizable over $\mathbb{R}$, and so, any other matrix $B \in \mathcal{M}_n(\mathbb{R})$, diagonalizable over $\mathbb{R}$, is similar to A if and only if $K_5(B) = K_5(A)$. Of course, for small values of n we could always solve the problem by comparing the characteristic polynomials of A and B. However for large values of n the situation would necessarily be complicated because, in contrast with the polynomial entries of $K_5(A)$, some of the coefficients of the characteristic polynomial of A grow exponentially with n.

The proof of Theorem 3.5 will be presented in the next subsection and is based upon the fact that the minimal polynomial of a diagonalizable matrix over $\mathbb{R}$ is determined by the matrix $K_{r+1}(A)$. As we will see, this will give rise to another rational criterion for diagonalizability of real matrices.

We end this subsection with a final observation regarding the application of Theorem 3.5 in the context of real symmetric matrices.

Corollary 3.7. *Two symmetric matrices $A, B \in \mathcal{M}_n(\mathbb{R})$ are similar if and only if $K_{r+1}(A) = K_{r+1}(B)$.*

Example 3.8. For the real symmetric matrix of Example 3.3, with $n = 4, 6, 8, \ldots$, and $a = b = 1 - \frac{n}{2}$,

[3] The omitted entries in the first row of $K_5(A)$ (and $L_5(A)$) are: $\operatorname{tr}(A) = n$, $\operatorname{tr}(A^2) = n^2 - n$ and $\operatorname{tr}(A^3) = n^3 - \frac{9}{2}n^2 + 6n$. The other omitted entries of $L_5(A)$ are: $\langle A, A \rangle = n^2 - \frac{1}{2}n$, $\langle A, A^2 \rangle = n^3 - \frac{9}{2}n^2 + 7n$, $\langle A, A^3 \rangle = n^4 - 6n^3 + \frac{27}{2}n^2 - \frac{17}{2}n$, $\langle A^2, A^2 \rangle = n^4 - 6n^3 + \frac{27}{2}n^2 - 8n$, $\langle A^2, A^3 \rangle = n^5 - \frac{15}{2}n^4 + \frac{89}{4}n^3 - \frac{63}{2}n^2 + 22n$ and $\langle A^3, A^3 \rangle = n^6 - 9n^5 + \frac{133}{4}n^4 - \frac{255}{4}n^3 + \frac{277}{4}n^2 - 32n$.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

one obtains $r = 2$ and

$$K_3(A) = \begin{bmatrix} n & n & \frac{1}{4}n^3 \\ n & \frac{1}{4}n^3 & \frac{1}{4}n^3 \\ \frac{1}{4}n^3 & \frac{1}{4}n^3 & \frac{1}{16}n^5 \end{bmatrix}.$$

Thus, to decide whether a symmetric matrix $B \in \mathcal{M}_n(\mathbb{R})$ is similar to A, we do not need to know the characteristic polynomial of A (whose coefficients grow exponentially with n), instead, we just need check the formula

$$K_3(B) = \begin{bmatrix} n & n & \frac{1}{4}n^3 \\ n & \frac{1}{4}n^3 & \frac{1}{4}n^3 \\ \frac{1}{4}n^3 & \frac{1}{4}n^3 & \frac{1}{16}n^5 \end{bmatrix}.$$

3.2. The polynomial $\psi_A(z)$. In this subsection, we present another rational criterion for diagonalizability of real matrices which implicitly involves the determination of the minimal polynomial of the matrix, but only when the matrix is diagonalizable over $\mathbb{R}$. In other words, it is a rational criterion for diagonalizability of real matrices that can also be seen as an algorithm for determining the coefficients of $\mu_A(z)$ when $A \in \mathcal{M}_n(\mathbb{R})$ is diagonalizable over $\mathbb{R}$.

Let $A \in \mathcal{M}_n(\mathbb{R})$ be given. Notice that if $|K_{r+1}(A)| = 0$, then the null-space of $K_{r+1}(A)$ is one-dimensional and it contains an unique vector with unitary last component. Therefore, we can define a monic polynomial

$$\psi_A(z) \equiv \begin{cases} c_0 + c_1 z + \cdots + c_{r-1}z^{r-1} + z^r & \text{if } |K_{r+1}(A)| = 0 \\ 1 & \text{if } |K_{r+1}(A)| < 0 \end{cases},$$

where $(c_0, \ldots, c_{r-1}, 1)$ denotes the unique vector with unitary last component lying in the null-space of $K_{r+1}(A)$.

Since the coefficients of $\psi_A(z)$ are determined by the matrix $K_{r+1}(A)$, it follows that $\psi_A(z)$ is invariant by similarity. Moreover, by the definition of $\psi_A(z)$ and (2.3) we have

$$\deg \psi_A(z) \le r \le d \le \deg \mu_A(z), \tag{3.4}$$

and by Theorem 2.3, we may conclude that

$$\deg \psi_A(z) < \deg \mu_A(z)$$

whenever A is not diagonalizable over $\mathbb{R}$.

At this moment it seems useful to notice also that the observations made initially on the minimal polynomial of a matrix $A \in \mathcal{M}_n$ can be reformulated in terms of the minimal polynomial for the eigenvalues of A, defined by

$$\varphi_A(z) \equiv (z - \lambda_1) \cdots (z - \lambda_d).$$

In fact, denoting by $\mathbb{C}[z]$ the ring of polynomials in z with complex coefficients, as $\mu_A(z)$ is the (unique)

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

monic generator of the proper ideal

$$\{P(z) \in \mathbb{C}[z] : P(A) = 0\},$$

and by consequence, the following theorem holds.

THEOREM 3.9. *For any matrix $A \in \mathcal{M}_n$ the following statements are equivalent: (i) $d = m$; (ii) $\varphi_A(z) = \mu_A(z)$; (iii) $\varphi_A(A) = 0$.*

At the same time we recall another result of [7] establishing an important relationship between the polynomials $\varphi_A(z)$ and $\mu_A(z)$ and the null-spaces of $K_{d+1}(A)$ and $L_{m+1}(A)$, which by Theorem 2.1 are one-dimensional and both contain an unique vector with unitary last component.

THEOREM 3.10. *For any matrix $A \in \mathcal{M}_n$ the following statements hold.*

(i) $\varphi_A(z) = a_0 + \cdots + a_{d-1}z^{d-1} + z^d$, *where* $(a_0, \ldots, a_{d-1}, 1)$ *is the unique vector with unitary last component lying in the null-space of* $K_{d+1}(A)$.

(ii) $\mu_A(z) = \bar{b}_0 + \cdots + \bar{b}_{m-1}z^{m-1} + z^m$, *where* $(b_0, \ldots, b_{m-1}, 1)$ *is the unique vector with unitary last component lying in the null-space of* $L_{m+1}(A)$.

Proof. See Theorems 4 and 7 of [7]. □

We are now in a position to prove the main result of this subsection which provides another rational criterion for diagonalizability of real matrices that involves the computation of the powers A^p, with $p = 0, \ldots, r$, needed to determine the polynomial $\psi_A(z)$ and the associated matrix $\psi_A(A)$.

THEOREM 3.11. *For any matrix $A \in \mathcal{M}_n(\mathbb{R})$ the following statements are equivalent: (i) A is diagonalizable over $\mathbb{R}$; (ii) $\psi_A(A) = 0$; (iii) $\psi_A(z) = \mu_A(z)$.*

Proof. Assume that $A \in \mathcal{M}_n(\mathbb{R})$ is diagonalizable over $\mathbb{R}$. By Theorems 2.3 and 3.9, we have $r = d = m$ and $\varphi_A(A) = 0$ But by the definition of $\psi_A(z)$ and Theorem 3.10, this implies $\psi_A(z) = \varphi_A(z)$. Hence, $\psi_A(A) = \varphi_A(A) = 0$.

Next assume that $\psi_A(A) = 0$. This means that the polynomial $\psi_A(z)$ is a multiple of $\mu_A(z)$, but this together with (3.4) implies $\psi_A(z) = \mu_A(z)$.

Finally, if $\psi_A(z) = \mu_A(z)$, then $r = \deg \psi_A(z) = \deg \mu_A(z) = m$, and by Theorem 2.3, it follows that A is diagonalizable over $\mathbb{R}$. □

Thus, as $\psi_A(z) = \mu_A(z)$ if and only if $\psi_A(A) = 0$, we are in the presence of a rational criterion for diagonalizability of real matrices that simultaneously provides an explicit procedure for determining the minimal polynomial of the matrix, but only when $A \in \mathcal{M}_n(\mathbb{R})$ is diagonalizable over $\mathbb{R}$.

EXAMPLE 3.12. Consider the matrix of Example 3.4. As $r = 2$ and the vector $(2n+6, -n-5, 1)$ lies in the null-space of $K_3(A)$, one obtains

$$\psi_A(z) = 2n + 6 - (n+5)z + z^2$$

and

$$\psi_A(A) = (2n+6)I - (n+5)A + A^2.$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

Hence, as the entry $(1,1)$ of $\psi_A(A)$ is $2-n$, it follows that A cannot be diagonalizable over $\mathbb{R}$, as we already knew.

Example 3.13. Consider the matrix of Example 3.6 with $n=100$, $a=-1$ and $b=0$. As $r=4$ and the vector $(-296, 200, 146, -100, 1)$ lies in the null-space of

$$K_5(A) = \begin{bmatrix} 100 & 100 & 9900 & 955600 & 94124200 \\ * & * & * & * & 9270952000 \\ * & * & * & * & 913164877200 \\ * & * & * & * & 89944386745600 \\ * & * & * & * & 8859290272852000 \end{bmatrix},$$

one obtains

$$\psi_A(z) = -296 + 200z + 146z^2 - 100z^3 + z^4$$

and

$$\psi_A(A) = -296I + 200A + 146A^2 - 100A^3 + A^4 = 0.$$

By Theorem 3.11, this not only shows that A is diagonalizable over $\mathbb{R}$, as we have already seen, but also ensures that equation

$$-296 + 200z + 146z^2 - 100z^3 + z^4 = 0 \tag{3.5}$$

has four distinct real roots which coincide with the eigenvalues of the matrix.

Now we can prove Theorem 3.5.

Assume that $A \in \mathcal{M}_n(\mathbb{R})$ is diagonalizable over $\mathbb{R}$. By Theorem 3.11, we have $\psi_A(z) = \mu_A(z)$. Therefore, the equation $\psi_A(z) = 0$ has r distinct real roots, $\lambda_1, \ldots, \lambda_r$, which coincide with the eigenvalues of A. By the definition of $\psi_A(z)$ this proves that the spectrum of A is determined by $K_{r+1}(A)$. On the other hand, as the Vandermonde matrix

$$V = \begin{bmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_r \\ \vdots & & \vdots \\ \lambda_1^{r-1} & \cdots & \lambda_r^{r-1} \end{bmatrix}$$

is invertible and

$$\operatorname{tr}(A^p) = \sum_{i=1}^{r} \alpha_i \lambda_i^p, \quad \text{for } p \geq 0,$$

where α_i denotes the geometric multiplicity of λ_i, one obtains

$$V \begin{bmatrix} \alpha_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_r \end{bmatrix} V^T = K_r(A). \tag{3.6}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

Thus, the eigenvalues of A and corresponding geometric multiplicities are determined by $K_{r+1}(A)$. This ends the proof of Theorem 3.5.

Theorem 3.11 combined with (3.6) may be useful in computing the canonical Jordan form of a real diagonalizable matrix because it does not require the characteristic polynomial of the matrix and avoids the direct computation of the ranks of the matrices $\lambda_1 I - A, \ldots, \lambda_r I - A$, which as one knows (see p. 132 of [6]) is an inherently unstable process.

Example 3.14. Consider again the matrix of Example 3.13. As the eigenvalues of A coincide with the four real roots of the equation (3.5), one obtains:

$$\lambda_1 = -1.4142\ldots, \quad \lambda_2 = 1.4142\ldots, \quad \lambda_3 = 1.5025\ldots, \quad \lambda_4 = 98.4974\ldots$$

Although this is useless to determine the rank of matrices $\lambda_i I - A$, $i = 1, \ldots, 4$, it is sufficient to obtain the geometric multiplicities through (3.6). Indeed, as

$$K_4(A) = \begin{bmatrix} 100 & 100 & 9900 & 955\,600 \\ * & * & * & 94124200 \\ * & * & * & 9270952\,000 \\ * & * & * & 913164877200 \end{bmatrix},$$

one gets

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \lambda_4^2 \\ \lambda_1^3 & \lambda_2^3 & \lambda_3^3 & \lambda_4^3 \end{bmatrix}^{-1} \begin{bmatrix} 100 \\ 100 \\ 9900 \\ 955600 \end{bmatrix} \approx \begin{bmatrix} 49.00 \\ 48.99 \\ 1.00 \\ 1.00 \end{bmatrix},$$

hence $\alpha_1 = \alpha_2 = 49$ and $\alpha_3 = \alpha_4 = 1$.

Example 3.15. Consider the symmetric matrix of Example 3.8. As $r = 2$ and the vector $\left(-\frac{1}{4}n^2, 0, 1\right)$ lies in the null-space of

$$K_3(A) = \begin{bmatrix} n & n & \frac{1}{4}n^3 \\ n & \frac{1}{4}n^3 & \frac{1}{4}n^3 \\ \frac{1}{4}n^3 & \frac{1}{4}n^3 & \frac{1}{16}n^5 \end{bmatrix},$$

one obtains $\psi_A(z) = z^2 - \frac{1}{4}n^2$. Therefore, the eigenvalues are $\lambda_1 = -\frac{1}{2}n$ and $\lambda_2 = \frac{1}{2}n$, with multiplicities given by

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -\frac{1}{2}n & \frac{1}{2}n \end{bmatrix}^{-1} \begin{bmatrix} n \\ n \end{bmatrix} = \begin{bmatrix} \frac{1}{2}n - 1 \\ \frac{1}{2}n + 1 \end{bmatrix}.$$

We conclude this subsection with a final remark about the polynomials $\psi_A(z)$ and $\mu_A(z)$. Apart the case $\deg \psi_A(z) = 0$, Theorem 3.10 establishes a formal parallelism between the construction of these polynomials that goes beyond the formulas

$$\deg \psi_A(z) = r = \min\{p \in \mathbb{N} : |K_{p+1}(A)| \leq 0\}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

and

$$\deg \mu_A(z) = m = \min\{p \in \mathbb{N} : |L_{p+1}(A)| \leq 0\}.$$

In fact, just as the coefficients of $\psi_A(z)$ are obtained from the null-space of $K_{r+1}(A)$, also the coefficients of $\mu_A(z)$ can be obtained in a similar manner, from the null-space of $L_{m+1}(A)$. From this point of view it is therefore clear that the determination of the polynomial $\psi_A(z)$ is always more favorable than the one of $\mu_A(z)$.

3.3. The companion matrix of $\mu_A(z)$. In this subsection, we present a rational criterion for diagonalizability of real matrices of a different type from the previous ones in the sense that involves the minimal polynomial of the matrix without assuming the calculation of any of its powers. Naturally, the use of any criteria of this type is justified whenever the determination of the coefficients of $\mu_A(z)$, through any of the known algorithms, is considerably more advantageous than the computation of the powers of A involved in the previous criteria.

The well-known fact that the companion matrix of a monic polynomial is non-derogatory plays an important role in what follows. For that purpose we recall that a matrix $A \in \mathcal{M}_n$ is said to be non-derogatory if each of its eigenvalues has geometric multiplicity one, or equivalently, if its characteristic polynomial, $|zI - A|$, and its minimal polynomial coincide (see [2], [9] or [6]).

For a given $A \in \mathcal{M}_n$, with $\mu_A(z) = c_0 + c_1 z + \cdots + c_{m-1}z^{m-1} + z^m$, let

$$C \equiv \begin{bmatrix} 0 & \cdots & 0 & -c_0 \\ 1 & \cdots & 0 & -c_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & -c_{m-1} \end{bmatrix} \in \mathcal{M}_m$$

be the companion matrix of $\mu_A(z)$. As $|zI - C| = \mu_A(z)$, the classical Newton identities,

$$\text{(3.7)} \qquad \operatorname{tr}\left(C^k\right) = \begin{cases} -c_{m-1}, & \text{if } k = 1 \\ -kc_{m-k} - \sum_{i=1}^{k-1} c_{m-k+i} \operatorname{tr}\left(C^i\right), & \text{if } m \geq k > 1 \\ -\sum_{i=1}^{m} c_{m-i} \operatorname{tr}\left(C^{k-i}\right), & \text{if } k > m \end{cases},$$

provide a recursive formula to compute the matrices $K_p(C)$ in terms of the coefficients of $\mu_A(z)$.

This means that when the coefficients of $\mu_A(z)$ are given we do not need to calculate any of the powers of C to obtain the matrix $K_m(C)$, and as Horn and Lopatin prove in [7] this is almost all we need to decide whether A is diagonalizable.

THEOREM 3.16. *A matrix $A \in \mathcal{M}_n$ is diagonalizable if and only if $K_m(C)$ is non-singular.*

Proof. See Corollary 1 of [7]. ∎

We will end this note with the following version of this theorem for real matrices.

THEOREM 3.17. *A matrix $A \in \mathcal{M}_n(\mathbb{R})$ is diagonalizable over $\mathbb{R}$ if and only if $K_m(C)$ is positive definite.*

Proof. As $|zI - C| = \mu_A(z)$ and C is non-derogatory, we have $\mu_A(z) = \mu_C(z)$. Hence, A is diagonalizable over $\mathbb{R}$ if and only if C is diagonalizable over $\mathbb{R}$. But on the other hand, as $\deg \mu_C(z) = m$ we have by

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 36, pp. 664-677, September 2020.

Theorem 2.3 that C is diagonalizable over $\mathbb{R}$ if and only if

$$\min\{p \in \mathbb{N} : |K_{p+1}(C)| \leq 0\} = m,$$

which in turn is equivalent to $|K_p(C)| > 0$ for $p = 1, \ldots, m$. ∎

Example 3.18. Consider the matrix of Example 3.3 with $n = 20$. Using any of the known methods for determining the minimal polynomial of a matrix, one obtains

$$\mu_A(z) = z^4 - 20z^3 - 6z^2 - 540z - 891.$$

Therefore, $C \in \mathcal{M}_4(\mathbb{R})$ and by (3.7) we do not need to calculate any of the powers of C to obtain

$$K_4(C) = \begin{bmatrix} 4 & 20 & 412 & 9980 \\ 20 & 412 & 9980 & 216436 \\ 412 & 9980 & 216436 & 4628900 \\ 9980 & 216436 & 4628900 & 99632908 \end{bmatrix}.$$

As $|K_3(C)| < 0$, it follows by Theorem 3.17 that A cannot be diagonalizable over $\mathbb{R}$. Nevertheless, as $K_4(C)$ is non-singular, it follows by Theorem 3.16 that A is diagonalizable over $\mathbb{C}$.

Acknowledgments. The author wishes to express his gratitude to José Manuel Ferreira for his invaluable insight and helpful discussions on the subject which contributed to improve the final version of this work. Special thanks to the referees for their comments and suggestions.

REFERENCES

[1] M. Abate. When is a linear operator diagonalizable? *American Mathematical Monthly*, 104:824–830, 1997.
[2] Ch.G. Cullen. *Matrices and Linear Transformations*. Dover, New York, 1990.
[3] M. Fiedler. *Special Matrices and Their Applications in Numerical Mathematics*. Dover, New York, 2008.
[4] F.R. Gantmacher. *The Theory of Matrices*, Vol I. Chelsea Publishing Company, New York, 1964.
[5] F.R. Gantmacher. *The Theory of Matrices*, Vol II. Chelsea Publishing Company, New York, 1964.
[6] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.
[7] R.A. Horn and A.K. Lopatin. The moment and Gram matrices, distinct eigenvalues and zeroes, and rational criteria for diagonalizability. *Linear Algebra and its Applications*, 299:153–163, 1999.
[8] R.F. Rinehart. An interpretation of the index of inertia of the discriminant matrices of a linear associative algebra. *Transactions of the American Mathematical Society*, 46:307–327, 1939.
[9] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. Springer, New York, 1993.