



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

LINEAR ALGEBRA
AND ITS
APPLICATIONS

Linear Algebra and its Applications 405 (2005) 239–248

www.elsevier.com/locate/laa

On factor width and symmetric H -matrices

Erik G. Boman^{a,*}, Doron Chen^b, Ojas Parekh^c,
Sivan Toledo^{b,2}

^aDepartment of Discrete Algorithms and Mathematics, Sandia National Laboratories, Albuquerque, NM 87185-1111, USA

^bSchool of Computer Science, Tel-Aviv University, Tel-Aviv 69978, Israel

^cDepartment of Mathematics and Computer Science, Emory University, Atlanta, GA 30322, USA

Received 29 November 2004; accepted 29 March 2005

Available online 17 May 2005

Submitted by M.J. Tsatsomeros

Abstract

We define a matrix concept we call *factor width*. This gives a hierarchy of matrix classes for symmetric positive semidefinite matrices, or a set of nested cones. We prove that the set of symmetric matrices with factor width at most two is exactly the class of (possibly singular) symmetric H -matrices (also known as generalized diagonally dominant matrices) with positive diagonals, H^+ . We prove bounds on the factor width, including one that is tight for factor widths up to two, and pose several open questions.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Combinatorial matrix theory; H -matrix; Generalized diagonally dominant; Factor width

* Corresponding author.

E-mail addresses: egboman@sandia.gov (E.G. Boman), mycroft@tau.ac.il (D. Chen), ojas@mathcs.emory.edu (O. Parekh), stoledo@tau.ac.il (S. Toledo).

¹ Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the US Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

² Supported in part by an IBM Faculty Partnership Award, by grant 572/00 from the Israel Science Foundation (founded by the Israel Academy of Sciences and Humanities), and by grant 2002261 from the United–States–Israel Binational Science Foundation.

1. Introduction

Symmetric positive definite and semidefinite (SPD and SPSD, respectively) matrices arise frequently in applications and have been studied by many authors [1,2]. For instance, it is well known that a Cholesky decomposition $A = LL^T$, where L is lower triangular, exists for any SPD matrix A . In this paper we characterize SPSD matrices in terms of rectangular factorizations of the type $A = VV^T$, where V is typically sparse and may have more columns than rows.

We restrict our attention to real matrices in this paper. In Section 2 we define the factor width of a symmetric matrix and show some basic properties. In Section 3 we show our main result, that factor-width-2 matrices are precisely H^+ matrices. We review a couple of known properties of H -matrices in the process. In Section 4 we prove bounds on the factor width, and show that a lower bound is exact for factor widths one and two. Finally, in Section 6 we pose several open questions.

2. The factor width of a symmetric matrix

Definition 1. The factor width of a real symmetric matrix A is the smallest integer k such that there exists a real (rectangular) matrix V where $A = VV^T$ and each column of V contains at most k non-zeros.

For example, let

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 2 & -2 \\ -1 & -2 & 5 \end{pmatrix}, \quad \text{and let } V = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & -2 & -1 \end{pmatrix}.$$

Then A has factor width at most two because $A = VV^T$. It is easy to see that a matrix has factor width one if and only if it is diagonal and non-negative; hence, the factor width of A is two. The factor width is independent of the ordering of the matrix since $PAP^T = (PV)(PV)^T$ has the same factor width as $A = VV^T$ for any permutation matrix P .

It follows from well-known properties of diagonally dominant matrices [3] that symmetric diagonally dominant matrices with non-negative diagonal have factor width two, which we also prove below. Recall that a real matrix A is diagonally dominant if $|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$ for all i .

Proposition 2. *If A is SPSD and diagonally dominant then A has factor width at most two.*

Proof. Let $P = \{(i, j) | i < j, a_{ij} > 0\}$ and $N = \{(i, j) | i < j, a_{ij} < 0\}$, and let e_i denote the i th unit vector. Then we can write A as a sum of rank-1 matrices,

$$A = \sum_{i=1}^n \left(a_{ii} - \sum_{j \neq i} |a_{ij}| \right) e_i e_i^T + \sum_{(i,j) \in P} a_{ij} (e_i + e_j)(e_i + e_j)^T + \sum_{(i,j) \in N} (-a_{ij})(e_i - e_j)(e_i - e_j)^T.$$

For diagonally dominant matrices all the coefficients are non-negative, and one can readily construct a V such that $A = VV^T$ from the expression above where each column of V is of one of the types $\sqrt{a_{ii} - \sum_{j \neq i} |a_{ij}|} e_i$, $\sqrt{a_{ij}}(e_i + e_j)$, or $\sqrt{-a_{ij}}(e_i - e_j)$. \square

Note that not all factor-width-two matrices are diagonally dominant, as the matrix A in the beginning of this section shows. Any SPSD matrix of order n has factor width at most n . A question arises: Are there matrices of factor width k for all $k \leq n$? The answer is yes.

Proposition 3. *For any positive $k \leq n$, there exist matrices of order n with factor width k .*

Proof. Let $v_k = (1, 1, \dots, 1, 0, \dots, 0)^T$, where there are k ones. Let $A = v_k v_k^T$. Clearly, the factor width of A is at most k . Since A has rank one then $A = v_k v_k^T$ is the unique symmetric rank-one factorization, and there cannot be any other factorization $A = \bar{V} \bar{V}^T$ with fewer non-zeros. \square

We remark that the lemma above holds even if we restrict our attention to full-rank matrices. A simple example is $A = v_k v_k^T + \epsilon I$ for sufficiently small ϵ .

In conclusion, the concept of factor width defines a family of matrix classes. Let $FW(k)$ denote the set of matrices with factor width k or less. Then $FW(1) \subset FW(2) \subset \dots$. It is easy to verify that $FW(k)$ is a pointed convex cone for any k and $FW(n)$ is precisely the cone of SPSD matrices of order n .

3. Factor-width-2 matrices are H -matrices

The importance of our study of the class of factor-width-2 matrices stems from the fact, which we prove in this section, that this class is exactly the class of H^+ matrices (defined later in this section), which occur frequently in engineering and scientific computation [4,1].

The definition of H -matrices relies on M -matrices. In this paper, we allow both M - and H -matrices to be singular, which is a bit unusual but convenient for us. Following [1], we have:

Definition 4. A real matrix A is an M -matrix if it is of the form $A = sI - B$, where $B \geq 0$ and $s \geq \rho(B)$, where ρ denotes the spectral radius.

For symmetric matrices there is a simpler characterization.

Lemma 5. A real symmetric matrix A is an M -matrix if and only if $a_{ij} \leq 0$ for all $i \neq j$ and A is positive semidefinite.

Definition 6. A matrix A is defined to be an H -matrix if $M(A)$ is an M -matrix, where the comparison matrix $M(A)$ of a matrix A is defined by

$$(M(A))_{ij} = \begin{cases} |a_{ij}|, & i = j, \\ -|a_{ij}|, & i \neq j. \end{cases}$$

We use H^+ to denote H -matrices that have non-negative diagonal. A useful characteristic of non-singular H -matrices (see for instance, [5, Lemma 6.4]) is that they are generalized strictly diagonally dominant, defined as follows:

Definition 7. A square matrix A is *generalized (weakly) diagonally dominant* if there exists a positive vector $y > 0$ such that for every row i ,

$$|a_{ii}|y_i \geq \sum_{j \neq i} |a_{ij}|y_j.$$

If strict inequality holds, we say A is strictly generalized diagonally dominant.

The problem of finding such a vector y is equivalent to the problem of finding a positive diagonal matrix D such that AD (or equivalently, DAD) is diagonally dominant. This problem has been studied in [6,4]; in general y may be found by solving a linear feasibility problem, but potentially faster iterative algorithms were proposed in the papers mentioned.

Theorem 8. A symmetric matrix A is an H -matrix if and only if A is generalized (weakly) diagonally dominant.

This is a well-known equivalence (see, e.g., [1] for a proof for the non-singular case). Now we are ready to prove our main result.

Theorem 9. A matrix has factor width at most two if and only if it is a symmetric H^+ -matrix.

Proof. (\Leftarrow) Suppose that A is a symmetric H^+ -matrix. Then A is generalized diagonally dominant, and hence there is a positive diagonal matrix D such that $\tilde{A} = DAD$

and \tilde{A} is diagonally dominant. We know that diagonally dominant matrices have factor-width at most 2 by Proposition 2. Hence $\tilde{A} = VV^T$ for some V with at most two non-zeros per column. But $A = D^{-1}\tilde{A}D^{-1} = (D^{-1}V)(D^{-1}V)^T$, so A also has factor-width 2. This concludes the first part of the proof.

(\Rightarrow) Suppose A has factor width two or less. A symmetric matrix A is an H -matrix if and only if its comparison matrix $M(A)$ is an M -matrix. Given a factor-width-two factorization $A = VV^T$, we can obtain a width-two factorization of $M(A)$ by simply flipping the sign of one non-zero in each column of V that contains two non-zeros with the same sign. By this factorization, $M(A)$ is positive semi-definite. Because $M(A)$ is a comparison matrix, it has non-negative diagonals and non-positive off-diagonals. Therefore, $M(A)$ satisfies the conditions of Lemma 5, so it is an M -matrix and hence A is an H -matrix. Since A is also SPSD, A must be in H^+ . \square

One consequence of this theorem is that for any $FW(2)$ matrix $A = VV^T$, there exists a positive diagonal D such that $A = (DU)(DU)^T$, where U has ≤ 2 non-zeros per column and entries of unit magnitude. However, this does not imply that $V = DU$.

4. Bounding the factor width

We do not know if the factor width of a given matrix can be efficiently computed, except in special cases. From our characterizations, it follows that $FW(k)$ matrices can be recognized in linear time for $k = 1$ and in polynomial time for $k = 2$, but already recognition for $k = 3$ is of unknown complexity and may be NP-hard.

In this section we derive several bounds that can be used to efficiently estimate the factor width of a matrix.

One upper bound on the factor width is easy to obtain: the largest number of non-zeros in a column of a Cholesky factor of PAP^T for some permutation matrix P . Many sparse matrices have a sparse Cholesky factor, and effective algorithms exist to find a permutation P that leads to a sparse factor. We note, however, that this bound may be very loose. For example, all the Cholesky factors of symmetric permutations of the Laplacian of the complete graph K_n have n non-zeros in their first column, giving a trivial upper bound of n , even though the Laplacian matrix actually has factor width 2.

The lower bounds that we present relate the factor width of a matrix A to the 2-norm of a matrix derived from A . The derivations are computationally trivial. One of the bounds is tight for matrices with factor widths one or two.

We use two tools to derive from A a matrix whose 2-norm lower bounds the factor width of A . The first tool is *diagonal normalization*, or symmetric diagonal scaling. The factor width of A is invariant under symmetric diagonal scalings of the form DAD , where D is diagonal, but the norm is not. If, however, we always

symmetrically scale A so that the diagonal elements of DAD are all 1's (except for diagonal elements in zero rows of A , which remain zero in DAD), then $\|DAD\|_2$ bounds from below the factor width of A . The second tool is perhaps more surprising. We show that if we also replace the elements of A by their absolute values, we get a tighter lower bound.

Definition 10. Let A be an SPSD matrix. Let D_A be the diagonal matrix whose diagonal elements are those of A , and let D_A^+ be the Moore–Penrose pseudo-inverse of D_A , that is, $(D_A^+)_{ii} = 1/(D_A)_{ii}$ for all i where $(D_A)_{ii} \neq 0$. The *diagonal normalization* $\text{dn}(A)$ is the matrix

$$\text{dn}(A) = (D_A^+)^{1/2} A (D_A^+)^{1/2}.$$

Our lower bounds depend on the following lemma, which provides a sufficient condition for a real function s of a matrix to be a lower bound on factor width.

Lemma 11. Let s be a function which assigns a real value to an SPSD matrix. Let s satisfy:

- (1) $s(uu^T) \leq k$ for any column vector u with k non-zeros,
- (2) $s(A + B) \leq \max(s(A), s(B))$.

Then for any SPSD matrix A , the factor width of A is at least $\lceil s(A) \rceil$.

Proof. We prove the theorem by first showing that if the factor-width of A is bounded by k , then $s(A) \leq k$.

Let A be a matrix in $FW(k)$ (A has factor width at most k). Let $A = UU^T$ be a factor-width- k representation of A . The number of non-zeros in a column u of U , which we denote by $\text{nnz}(u)$, is at most k . For notational convenience, let $u \in U$ mean “ u is a column of U ”.

$$\begin{aligned} s(A) &= s(UU^T) \\ &= s\left(\sum_{u \in U} uu^T\right) \\ &\leq \max_{u \in U} (s(uu^T)) \\ &\leq \max_{u \in U} (\text{nnz}(u)) \\ &\leq k. \end{aligned}$$

We have shown that if the factor-width of A is at most k , then $s(A) \leq k$. Therefore, if $s(A) > k$ then the factor-width of A is larger than k . Thus, the factor width of A is greater or equal to $\lceil s(A) \rceil$. \square

4.1. The diagonal normalization bound

We now show that the factor-width of an SPSD matrix A is bounded from below by $\lceil \|\text{dn}(A)\|_2 \rceil$.

Theorem 12. For any SPSD matrix A , the factor width of A is bounded from below by $\lceil \|\text{dn}(A)\| \rceil$.

Proof. In our proof we will use the two results below, which we state without proofs since they both can be easily verified. \square

Lemma 13. Suppose a, b, c, d are non-negative and $c > 0, d > 0$.

Then

$$\frac{a + b}{c + d} \leq \max\left(\frac{a}{c}, \frac{b}{d}\right).$$

Lemma 14. Let A be SPSD. Then

$$\|\text{dn}(A)\| = \lambda_{\max}(A, D_A) = \max_x \frac{x^T A x}{x^T D_A x},$$

where $\lambda(A, B)$ denotes a generalized eigenvalue.

We define the function s_1 to be $s_1(A) = \|\text{dn}(A)\|_2$ and show that s_1 satisfies the conditions of Lemma 11. We begin with condition 1, and show that for any vector u with k non-zeros, $s_1(uu^T)$ is exactly k .

Let u be a column vector with k non-zero entries. If $u = 0$, then $s_1(uu^T) = 0 = k$. Otherwise, uu^T is a rank-1 matrix. The matrix

$$\text{dn}(uu^T) = \left(D_{uu^T}^+\right)^{1/2} uu^T \left(D_{uu^T}^+\right)^{1/2} = \left[\left(D_{uu^T}^+\right)^{1/2} u\right] \left[\left(D_{uu^T}^+\right)^{1/2} u\right]^T$$

also has rank-1, because $(D_{uu^T})_{ii} = u_i^2$. The norm of $\text{dn}(uu^T)$ is the only non-zero eigenvalue of $\left(D_{uu^T}^+\right)^{1/2} uu^T \left(D_{uu^T}^+\right)^{1/2}$. Let v be the sign vector of u ,

$$v_i = \begin{cases} 1 & u_i > 0, \\ -1 & u_i < 0, \\ 0 & u_i = 0. \end{cases}$$

We now show that v is an eigenvector corresponding to the eigenvalue k . We have

$$\left(D_{uu^T}^+\right)_{ii} = \begin{cases} u_i^{-2} & u_i \neq 0, \\ 0 & u_i = 0, \end{cases}$$

so

$$\left(D_{uu^T}^+\right)^{1/2} u = v.$$

Therefore, $\text{dn}(uu^T) = vv^T$, so $\text{dn}(uu^T)v = (vv^T)v = v(v^T v) = vk = kv$.

All that remains is to prove that $s_1(A + B) \leq \max(s_1(A), s_1(B))$.

$$\begin{aligned}
 s_1(A + B) &= \|\text{dn}(A + B)\|_2 \\
 &= \max_x \frac{x^T(A + B)x}{x^T(D_A + D_B)x} \\
 &= \max_x \frac{x^T A x + x^T B x}{x^T D_A x + x^T D_B x} \\
 &\leq \max_x \max \left(\frac{x^T A x}{x^T D_A x}, \frac{x^T B x}{x^T D_B x} \right) \\
 &\leq \max \left(\max_y \frac{y^T A y}{y^T D_A y}, \max_z \frac{z^T B z}{z^T D_B z} \right) \\
 &= \max(\|\text{dn}(A)\|_2, \|\text{dn}(B)\|_2) \\
 &= \max(s_1(A), s_1(B)),
 \end{aligned}$$

where we used Lemmas 13 and 14.

4.2. A tighter lower bound

The lower bound can be made tighter. Let $|A|$ denote the matrix whose i, j entry is $|a_{ij}|$.

Theorem 15. For any SPSD matrix A , the factor width of A is bounded from below by $\lceil \|\text{dn}(|A|)\|_2 \rceil$.

Proof. Let $s_2(A) = \lceil \|\text{dn}(|A|)\|_2 \rceil$. One can show that s_2 satisfies the first condition in Lemma 11 in the same way as for s_1 (cf. proof of Theorem 12). For the second condition, we only need to prove that

$$\max_x \frac{x^T(|A + B|)x}{x^T(D_A + D_B)x} \leq \max_x \frac{x^T(|A| + |B|)x}{x^T(D_A + D_B)x},$$

because the rest follows from the previous proof. Without loss of generality we can assume A and B already have been diagonally normalized (scaled) so $D_A = D_B = I$. From the Perron–Frobenius Theorem we know that the largest eigenvector of a non-negative matrix is non-negative, so $x \geq 0$. From the triangle inequality $|a_{ij} + b_{ij}| \leq |a_{ij}| + |b_{ij}|$, it follows that $x^T(|A + B|)x \leq x^T(|A| + |B|)x$ for any $x \geq 0$. \square

This second bound is tighter (or at least as tight) as our first bound (Theorem 12). This follows from the fact that $\|A\|_2 \leq \| |A| \|_2$ for any SPSD A .

5. Identifying factor-width-2 matrices

Since $FW(2)$, the set of all matrices with factor width at most two, is a subset of H -matrices, any algorithm to identify H -matrices (generalized diagonally dominant matrices) can easily be adapted to recognize matrices in $FW(2)$. There are many such algorithms, see for instance, [6,4]. Since $FW(2)$ matrices are also SPSD, it may in fact be easier to identify such matrices than general H -matrices.

We show that we can use Theorem 15 to easily identify matrices with factor-width at most 2. The following theorem shows that $FW(2)$ is exactly the set of symmetric matrices with non-negative diagonals satisfying $\|\text{dn}(|A|)\| \leq 2$.

Theorem 16. *Matrix A has factor-width at most 2 if and only if it is symmetric with non-negative diagonals, and satisfies $\|\text{dn}(|A|)\| \leq 2$.*

Proof. (\Rightarrow) Let A have factor-width at most 2. Then A is symmetric with non-negative diagonals. By Theorem 15, $\|\text{dn}(|A|)\| \leq 2$.

(\Leftarrow) Let A be symmetric with non-negative diagonals satisfying $\|\text{dn}(|A|)\| \leq 2$. Since $\|\text{dn}(|A|)\| = \max_x \frac{x^T |A| x}{x^T D_A x} \leq 2$, it follows that $x^T (2D_A - |A|)x \geq 0$ so $2D_A - |A|$ is positive semidefinite. $2D_A - |A|$ is exactly A 's comparison matrix. Since A 's comparison matrix is symmetric and positive semidefinite, then it is an M -matrix. Therefore A is an H -matrix. Furthermore, A is symmetric with non-negative diagonals, and therefore A is an H^+ -matrix. Since A is an H^+ -matrix, it has factor-width at most 2. \square

This result is in fact just a special case of one of many known characterizations of H -matrices:

Theorem 17. *The following are equivalent:*

- (i) A is a non-singular H -matrix.
- (ii) Let $D = \text{diag}(A)$. Then $\rho(|I - D^{-1}A|) < 1$.

This theorem was stated in a slightly different (more general) form for M -matrices in [7, Theorem 1]. Note that this result holds for all H -matrices (even non-symmetric matrices). Since we allow singular H -matrices, the inequality in case (ii) should be modified to $\rho(|I - D^{-1}A|) \leq 1$. This condition is then equivalent to Theorem 16 in the SPSD case.

We conclude that our lower bound (Theorem 16) is always tight for factor width two. We do not know if the bound is tight for factor width three. For large factor widths it is easy to construct examples where the bound is not tight, that is, the factor width is strictly greater than the lower bound.

6. Conclusions and open problems

We have defined factor width for symmetric matrices and characterized the matrix classes $FW(1)$ and $FW(2)$. An obvious question is, does $FW(k)$ correspond to any known matrix class for other values of k ? In particular, what is $FW(3)$? We note that the finite element method naturally produces matrices of low factor width since each element has a small number of degrees of freedom. This indicates that the study of (low) factor width may have practical applications.

Other open problems are the complexity of computing the factor width (Section 4), and proving better upper bounds. It could be interesting to study how many columns in V are needed to realize a factor width k decomposition $A = VV^T$. This number can be denoted the “factor width k rank”.

Finally, we ask if there is any useful generalization of factor-width for non-symmetric matrices. A simple but naive choice is to consider factorizations $A = UV^T$ and count the non-zeros in columns of U and V . However, with such a definition any matrix would have “factor width” one since any non-zero a_{ij} in A can be represented by the scaled outer product $a_{ij}e_i e_j^T$.

Acknowledgements

We thank Bruce Hendrickson and Bryan Shader for helpful discussions on parts of this work.

References

- [1] A. Berman, R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, SIAM, Philadelphia, 1994.
- [2] R. Hill, S. Waters, On the cone of positive semidefinite matrices, *Linear Algebra Appl.* 90 (1987) 81–88.
- [3] G. Barker, D. Carlson, Cones of diagonally dominant matrices, *Pacific J. Math.* 57 (1) (1975) 15–32.
- [4] L. Li, On the iterative criterion for generalized diagonally dominant matrices, *SIAM J. Matrix Anal. Appl.* 24 (1) (2002) 17–24.
- [5] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, 1994.
- [6] B. Li, L. Li, M. Harada, M. Tsatsomeros, An iterative criterion for H -matrices, *Linear Algebra Appl.* 278 (1998) 179–190.
- [7] R.S. Varga, On recurring theorems on diagonal dominance, *Linear Algebra Appl.* 13 (1976) 1–9.