On Symmetric Ultrametric Matrices

Richard S. Varga and Reinhard Nabben†*

Abstract. In a recent paper by S. Martínez, G. Michon, and J. San Martín [1], it was shown that if $A := [a_{i,j}]$ in $\mathbb{R}^{n\times n}$ is a symmetric strictly ultrametric matrix, then its inverse $A^{-1} := [\alpha_{i,j}]$ in $\mathbb{R}^{n\times n}$ is a strictly diagonally dominant Stieltjes matrix, with the additional property that

$$a_{i,j} = 0 \text{ if and only if } \alpha_{i,j} = 0.$$

Here, a generalization of this result to symmetric ultrametric matrices is given.

Key words. Stieltjes matrix, strictly ultrametric matrices.

AMS(MOS) subject classification. 15A57, 15A48.

1 Introduction

In a recent paper, S. Martínez, G. Michon, and J.San Martín [1] gave the following definition. For notation, let $N := \{1, 2, \cdots, n\}$ for any positive integer n.

Definition 1.1. *A matrix $A = [a_{i,j}]$ in $\mathbb{R}^{n\times n}$ is a symmetric strictly ultrametric matrix if*

$$\begin{array}{rl} i) & A \text{ is symmetric with nonnegative entries;} \\ ii) & a_{i,j} \geq \min\{a_{i,k}; a_{k,j}\} \text{ for all } i,k,j \text{ in } N; \\ iii) & a_{i,i} > \max\{a_{i,k} : k \in N\backslash\{i\}\} \text{ for all } i \in N, \end{array} \tag{1}$$

where, if $n = 1$, then $(1iii)$ is interpreted as $a_{1,1} > 0$.

The result of [1, Theorem 1] is

*Institute for Computational Mathematics, Kent State University, Kent, OH 44242, USA. The research of this author was partially supported by the National Science Foundation.

†Fakultät für Mathematik, Universität Bielefeld, D-4800 Bielefeld 1, FRG. The research of this author was partially supported by the Deutsche Forsungsgemeinschaft.

Numerical Linear Algebra

Theorem 1.1. ([1]). *If $A = [a_{i,}]$ in $\mathbb{R}^{n\times n}$ is a symmetric strictly ultrametric matrix, then A is nonsingular and its inverse, $A^{-1} := [\alpha_{i,j}]$ in $\mathbb{R}^{n\times n}$, is a strictly diagonally dominant Stieltjes matrix (i.e., $\alpha_{i,j} \leq 0$ for all $i \neq j$ and $\alpha_{i,i} > \sum_{\substack{k=1\\k\neq i}}^{n} |\alpha_{i,k}|$ for all $1 \leq i, j \leq n$), with the additional property that*

$$a_{i,j} = 0 \quad if\ and\ only\ if \quad \alpha_{i,j} = 0. \tag{2}$$

For a shorter linear algebra proof of Theorem 1.1, see Nabben and Varga [2].

The result from Theorem 1, that the inverse of a symmetric strictly ultrametric matrix is a *strictly* diagonally dominant Stieltjes matrix, suggested that a possible weakening of the hypotheses of Definition 1.1 might be possible. Our modest goal here is to specifically weaken (1.*iii*) of Definition 1.1 to obtain a generalization, Theorem 2.2 below, of Theorem 1.1.

2 Main Result

To begin, we first state the following results of Nabben and Varga [2], which will be used in our constructions below. For additional notation, set

$$\xi_n := (1, 1, \cdots, 1)^T \text{in } \mathbb{R}^n.$$

Proposition 2.1. ([2]). *Let $A = [a_{i,j}]$ in $\mathbb{R}^{n\times n}$ be symmetric with all its entries nonnegative, and set*

$$\tau(A) := \min\{a_{i,j} : i, j \in N\}. \tag{3}$$

If $n > 1$, then A is a symmetric strictly ultrametric matrix if and only if $A - \tau(A)\xi_n\xi_n^T$ is completely reducible, i.e., there exist a positive integer r with $1 \leq r < n$ and a permutation matrix P in $\mathbb{R}^{n\times n}$ such that

$$P\left[A - \tau(A)\xi_n\xi_n^T\right]P^T = \begin{bmatrix} C & O \\ O & D \end{bmatrix}, \tag{4}$$

where $C \in \mathbb{R}^{r\times r}$ and $D \in \mathbb{R}^{(n-r)\times(n-r)}$ are each a symmetric strictly ultrametric matrix.

Theorem 2.1. ([2]). *Given any symmetric strictly ultrametric matrix A in $\mathbb{R}^{n\times n}$ $(n \geq 1)$, there is an associated rooted tree for $N = \{1, 2, \cdots, n\}$, consisting of $2n - 1$, vertices, such that*

$$A = \sum_{\ell=1}^{2n-1} \tau_\ell\, \mathbf{u}_\ell\, \mathbf{u}_\ell^T, \tag{5}$$

where the vectors $\mathbf{u}_\ell$ *in* (5), *determined from the vertices of the tree, are nonzero vectors in* $\mathbb{R}^n$ *having only* 0 *and* 1 *components, and where the* τ_ℓ*'s in* (5) *are nonnegative and, with the notation that*

$$\chi(\mathbf{u}_\ell) := \text{ sum of the components of } \mathbf{u}_\ell, \tag{6}$$

satisfy the property that $\tau_\ell > 0$ *when* $\chi(\mathbf{u}_\ell) = 1$. *Conversely, given any rooted tree for* $N = \{1, 2, \cdots, n\}$ *with* $\tau_\ell > 0$ *when* $\chi(\mathbf{u}_\ell) = 1$, *then* $\sum_{\ell=1}^{2n-1} \tau_\ell \, \mathbf{u}_\ell \, \mathbf{u}_\ell^T$ *is a symmetric strictly ultrametric matrix in* $\mathbb{R}^{n \times n}$.

To generalize Theorem 1.1, (1*iii*) of Theorem 1.1 is weakened to allow for the case of equality.

Definition 2.1. *A matrix* $A = [a_{i,j}]$ *in* $\mathbb{R}^{n \times n}$ *is a symmetric pre-ultrametric matrix if*

$$\begin{array}{rl} i) & A \text{ is symmetric with nonnegative entries;} \\ ii) & a_{i,j} \geq \min\{a_{i,k}; a_{k,j}\} \text{ for all } i,k,j \in N; \\ iii) & a_{i,i} \geq \max\{a_{i,k} : k \in N \backslash \{i\}\} \text{ for all } i \in N. \end{array} \tag{7}$$

It is evident that a symmetric pre-ultrametric matrix *can* be singular, as choosing $A = O$ shows. Now, it easily follows from Theorem 2.1 that A is a symmetric pre-ultrametric matrix if and only if the representation of (5) is valid where the τ_ℓ's in (5) are just nonnegative numbers (i.e., no further restrictions on the τ_ℓ's are necessary). But this shows that if A is a symmetric pre-ultrametric matrix, then for each $\epsilon > 0$,

$$A(\epsilon) := A + \epsilon I_n \text{ (where } I_n \text{ is the identity matrix in } \mathbb{R}^{n \times n}) \tag{8}$$

is a symmetric strictly ultrametric matrix which, from Theorem 1.1, is necessarily nonsingular. Then on applying (5) of Theorem 2.1, $A(\epsilon)$ can be represented as

$$A(\epsilon) = \sum_{\ell=1}^{2n-1} \tau_\ell(\epsilon) \mathbf{u}_\ell \mathbf{u}_\ell^T \quad (\epsilon > 0), \tag{9}$$

where it is important to note that the vectors $\{\mathbf{u}_\ell\}_{\ell=1}^{2n-1}$ in (9) are *independent* of ϵ. (This is a consequence of the fact that the complete reduction steps of (4) of Proposition 2.1, which are applied to principal submatrices of $A(\epsilon)$ to obtain the representation of (5), yield vectors $\mathbf{u}_\ell$ which depend only on the vertices of the associated tree and are *independent* of ϵ. This will also be illustrated in the example in Section 2.)

This brings us to our next

Definition 2.2. *A matrix* $A = [a_{i,j}]$ *in* $\mathbb{R}^{n \times n}$ *is a symmetric ultrametric matrix if* A *is a symmetric pre-ultrametric matrix and if, from the representation* (9), *the vectors* $\{\mathbf{u}_\ell\}_{\ell=1}^{2n-1}$ *satisfy*

$$(iv) \text{ span } \{\mathbf{u}_\ell : \tau_\ell(0) > 0\} = \mathbb{C}^n, \tag{10}$$

where $\tau_\ell(0) := \lim_{\epsilon \to 0} \tau_\ell(\epsilon)$.

To couple Definition 2.1 with Definition 1.1, let $\{\mathbf{e}_j\}_{n=1}^{n}$ denote the set of unit basis vectors in $\mathbb{R}^n$ (i.e., $\mathbf{e}_1 = (1, 0, \cdots, 0)^T$, $\mathbf{e}_2 = (0, 1, 0, \cdots, 0)^T$, etc.). If A is a symmetric strictly ultrametric matrix in $\mathbb{R}^{n\times n}$, then from Theorem 2.1, each $\mathbf{e}_j$ $(1 \le j \le n)$ is some $\mathbf{u}_\ell$ in $\{u_\ell\}_{\ell=1}^{2n-1}$ and its associated multiplier, τ_ℓ in (5), is necessarily positive. Consequently, as (10) is then obviously satisfied, each symmetric strictly ultrametric matrix in $\mathbb{R}^{n\times n}$ is necessarily a symmetric ultrametric matrix.

We next establish

Lemma 2.1. *Assume that $A = [a_{i,j}]$ in $\mathbb{R}^{n\times n}$ is a symmetric pre-ultrametric matrix. Then, A is positive definite if and only if (10) is valid, i.e, if and only if A is a symmetric ultrametric matrix.*

Proof. Since A is by hypothesis a symmetric pre-ultrametric matrix, then on letting $\epsilon \to 0$, it follows from (9) and (10) that

$$A = \sum_{\ell=1}^{m} \tau_\ell(0)\ \mathbf{u}_\ell\ \mathbf{u}_\ell^T \text{ where } \tau_\ell(0) > 0 \text{ for all } 1 \le \ell \le m, \tag{11}$$

where $m \le 2n-1$. (This amounts to throwing out those $\tau_\ell(0)$ in (9) which are zero, and then renumbering the remaining terms in (9)). Then for any $\mathbf{x} \in \mathbb{R}^n$,

$$\mathbf{x}^H A\mathbf{x} = \sum_{\ell=1}^{m} \tau_\ell(0) \left(\mathbf{x}^H \mathbf{u}_\ell\right) \left(\mathbf{u}_\ell^T \mathbf{x}\right) = \sum_{\ell=1}^{m} \tau_\ell(0) |\mathbf{u}_\ell^T \mathbf{x}|^2 \ge 0, \tag{12}$$

so that A is at least positive semi-definite. Equality holds in (12) (since $\tau_\ell(0) > 0$ for all $1 \le \ell \le m$) only if $\mathbf{u}_\ell^T \mathbf{x} = 0$ for all $1 \le \ell \le m$. Thus, $\mathbf{x}$ is orthogonal to every linear combination of the $\mathbf{u}_\ell$'s, i.e.,

$$\mathbf{x} \perp \text{span } \{\mathbf{u}_\ell : \tau_\ell(0) > 0\}.$$

If (10) is valid, then equality in (12) implies $\mathbf{x} = \mathbf{0}$ and $\mathbf{y}^H A\mathbf{y} > 0$ for all $\mathbf{y} \ne \mathbf{0}$ in $\mathbb{C}^n$, i.e., A is positive definite. Conversely, if (10) is not true, there is an $\mathbf{x} \ne \mathbf{0}$ with $\mathbf{x}^H A\mathbf{x} = \mathbf{0}$, so that A is singular. □

This brings us to our main result.

Theorem 2.2. *Let $A = [a_{i,j}]$ in $\mathbb{R}^{n\times n}$ be a symmetric ultrametric matrix, in the sense of Definition 2.1. Then, A is positive definite and its inverse, $A^{-1} := [\alpha_{i,j}]$ in $\mathbb{R}^{n\times n}$, is a diagonally dominant Stieltjes matrix, i.e., $\alpha_{i,j} \le 0$ for all $i \ne j$ and*

$$\alpha_{i,i} \ge \sum_{\substack{j=1 \\ j\ne i}}^{n} |\alpha_{i,j}| \quad \textit{for all} \quad i \in N, \tag{13}$$

with strict inequality holding in (13) *for at least one i in N. Moreover,*

$$a_{i,j} = 0 \quad \textit{implies} \quad \alpha_{i,j} = 0 \qquad \textit{(but not necessarily conversely).} \tag{14}$$

Proof. For each $\epsilon > 0, A(\epsilon) := A + \epsilon I_n$ is a symmetric strictly ultrametric matrix. Hence from Theorem 1.1, $(A + \epsilon I_n)^{-1} := [\alpha_{i,j}(\epsilon)]$ in $\mathbb{R}^{n \times n}$ is a strictly diagonally dominant Stieltjes matrix, so that

$$\alpha_{i,j}(\epsilon) \leq 0 \text{ for all } i \neq j \text{ in } N \text{ and for each } \epsilon > 0, \tag{15}$$

$$\alpha_{i,j}(\epsilon) > \sum_{\substack{j=1 \\ j \neq i}}^{n} |\alpha_{i,j}(\epsilon)| \text{ for all } i \in N \text{ and for each } \epsilon > 0, \text{ and} \tag{16}$$

$$a_{i,j}(\epsilon) = 0 \text{ if and only if } \alpha_{i,j}(\epsilon) = 0 \text{ for each } \epsilon > 0. \tag{17}$$

On letting $\epsilon \downarrow 0, A = A(0)$ is positive definite from Lemma 2.1, and its inverse, $A^{-1}(0) = [\alpha_{i,j}(0)]$ in $\mathbb{R}^{n \times n}$, is then well-defined. Again, letting $\epsilon \downarrow 0$ in (15) and (16) gives

$$\alpha_{i,j}(0) \leq 0 \text{ for all } i \neq j \text{ in } N, \text{ and} \tag{15$'$}$$

$$\alpha_{i,i}(0) \geq \sum_{\substack{j=1 \\ j \neq i}}^{n} |\alpha_{i,j}(0)| \text{ for all } i \in N. \tag{16$'$}$$

That strict inequality must hold, for some i, in (16$'$) is clear, for otherwise, $A^{-1}(0)\xi_n = \mathbf{0}$, which contradicts the fact, from Lemma 2.1, that $A = A(0)$ and its inverse are both positive definite.

Finally, to establish (14), first note from (7*iii*) that no diagonal element $a_{i,i}$ of A can vanish, as this would force A to have a zero row and to be singular. Thus, the condition in (17) necessarily pertains only to off-diagonal entries of the matrices $A(\epsilon)$ and $(A(\epsilon))^{-1}$. But as $a_{i,j}(\epsilon) = a_{i,j}$ for all $i \neq j$ from (8), (17) can be expressed as

$$a_{i,j} = 0 \text{ if and only if } \alpha_{i,j}(\epsilon) = 0 \text{ for each } \epsilon > 0, \tag{17$'$}$$

and on letting $\epsilon \downarrow 0$, we can only deduce (cf. (14)) that

$$a_{i,j} = 0 \text{ implies } \alpha_{i,j}(0) = 0,$$

for it could be the case that $a_{i,j} > 0$ and $\alpha_{i,j}(\epsilon) > 0$ for each $\epsilon > 0$, while $\alpha_{i,j}(0) = 0$. □

3 An Example

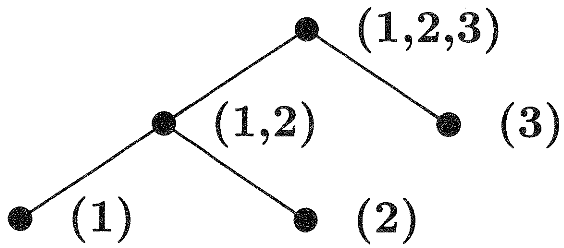
Consider the matrix

$$A_1 := \begin{bmatrix} 3 & 3 & 2 \\ 3 & 4 & 2 \\ 2 & 2 & 3 \end{bmatrix}, \tag{18}$$

which can be verified to be a symmetric pre-ultrametric matrix in $\mathbb{R}^{3\times 3}$. Then, for each $\epsilon > 0$,

$$A_1(\epsilon) := A_1 + \epsilon I_3 = \begin{bmatrix} 3+\epsilon & 3 & 2 \\ 3 & 4+\epsilon & 2 \\ 2 & 2 & 3+\epsilon \end{bmatrix}, \tag{19}$$

whose associated rooted tree (cf. Theorem 2) can be verified to be



Then, the representation for $A_1(\epsilon)$ in (9) holds with the following definitions:

$$\begin{cases} \tau_1(\epsilon) := 2; \mathbf{u}_1 := (1,1,1)^T & \tau_2(\epsilon) := 1; \mathbf{u}_2 := (1,1,0)^T \quad \tau_3(\epsilon) := \epsilon; \mathbf{u}_3 := (1,0,0)^T \\ \tau_4(\epsilon) := 1+\epsilon; \mathbf{u}_4 := (0,1,0)^T & \tau_5(\epsilon) := 1+\epsilon; \mathbf{u}_5 := (0,0,1)^T. \end{cases}$$

In this case (cf. (10)),

$$\text{span } \{u_\ell : \tau_\ell(0) > 0\} = \text{ span } \{(1,1,1)^T; (1,1,0)^T; (0,1,0)^T; (0,0,1)^T\} = \mathbb{C}^3,$$

so that A_1 is then a symmetric ultrametric matrix. Now, the inverse of A_1, given by

$$A_1^{-1} = \begin{bmatrix} +1.6 & -1 & -0.4 \\ -1 & +1 & 0 \\ -0.4 & 0 & 0.6 \end{bmatrix}, \tag{20}$$

is such that A_1^{-1} has *some* zero off-diagonal entries, while A_1 has *no* zero off-diagonal entries. This shows, in particular, that the implication of (14), which holds vacuously for A_1, is true, but that the inverse implication in (14) does *not* hold. We also see from (20) that strict diagonal dominance holds in the first and third rows of A_1^{-1}, while only diagonal dominance holds in the second row of A_1^{-1}.

As a final comment, it was our original thought that the case of strict inequality in $(1iii)$ could be weakened to the case of inequality in $(1iii)$, while still preserving the main results of Theorem 1.1, if some irreducibility-like additional hypothesis were added. As it turned out, our additional hypothesis, that of (10), is what resulted. It is perhaps interesting to note that the assumption of (10) and the assumption of irreducibility of a matrix (cf. [3, p. 19]) both can be viewed as *global properties* of a matrix.

References

[1] S.Martínez, G. Michon, and J. San Martín, *Inverses of ultrametric matrices are of Stieltjes type.* SIAM J. Matrix Analysis and Appl. (to appear).

[2] R. Nabben and R.S. Varga, *A linear algebra proof that the inverse of a strictly ultrametric matrix is a strictly diagonally dominant Stieltjes matrix.* SIAM J. Matrix Analysis and Appl. (to appear).

[3] R.S. Varga, *Matrix Iterative Analysis.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962.