COMMUNICATIONS IN INFORMATION AND SYSTEMS
Vol. 5, No. 1, pp. 123-130, 2005 005

EFFICIENT SOLUTION OF LINEAR MATRIX EQUATIONS WITH APPLICATION TO MULTISTATIC ANTENNA ARRAY PROCESSING

HANOCH LEV-ARI*

Abstract. We present a computationally-efficient matrix-vector expression for the solution of a matrix linear least squares problem that arises in multistatic antenna array processing. Our derivation relies on an explicit new relation between Kronecker, Khatri-Rao and Schur-Hadamard matrix products, which involves a *selection matrix* (i.e., a subset of the columns of a permutation matrix). Moreover, we show that the same selection matrix also relates the vectorization-by-columns operator to the *diagonal extraction* operator, which plays a central role in our computationally-efficient solution.

1. Introduction. Linear matrix equations show up in a variety of engineering, mathematics and physics problems, including linear system analysis, modeling of non-stationary covariances, and multistatic antenna array processing. For instance, the Lyapunov equations $A^H X + XA + Q = 0$ and $X - A^H XA = Q$ (where the superscript ^{H} denotes conjugate transpose) are used to analyze the stability of continuous-time and discrete-time systems, respectively [1]. The generalized Lyapunov equation

$$AXB^T + CXD^T = Q$$

has been used to characterize structured covariance matrices, and to construct efficient matrix factorization and inversion algorithms [2, 3, 4]. Such equations can be readily converted into the standard linear equation format by using the well-known identity [5]

$$\text{vec}\{AXB^T\} = (B \otimes A)\,\text{vec}\{X\} \qquad (1)$$

where $\text{vec}\{\cdot\}$ denotes *vectorization by columns* of a matrix. This results in the linear equation

$$(\,B \otimes A + D \otimes C\,)\,\text{vec}\{X\} = \text{vec}\,Q$$

which can be solved for the unknown $\text{vec}\{X\}$.

A linear matrix equation of a somewhat different flavor arises in *multistatic antenna array processing* applications. An unknown medium is probed by transmitting energy into it from a multi-element antenna array, and recording the scattered signal received by (another) multi-element antenna array. The resulting measurements are

*Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115. E-mail: levari@ece.neu.edu

arranged into a matrix $H = \{h_{ij}\}$, where h_{ij} is the response (at a single frequency) from the j-th transmitting element to the i-th receiving element [6]. When the medium consists of reasonably spaced point scatterers in a uniform background, the distorted wave Born approximation [7] provides a simple characterization of the multistatic data matrix H in terms of the scatterer locations $\{\chi_i\}$ and scattering coefficients $\{\tau_i\}$, viz.,

$$H = \sum_{i=1}^{L} g_{rec}(\chi_i)\, \tau_i\, g_{tr}^T(\chi_i) \tag{2}$$

where L denotes the number of point scatterers, and where $g_{tr}(\chi_i)$ (resp. $g_{rec}(\chi_i)$) is the so-called *steering vector* associated with wave propagation between the transmitting (resp. receiving) array and the i-th scatterer. The acoustics community usually refers to multistatic array processing as (mathematical) "time-reversal" [6].

The multistatic antenna array processing problem amounts to recovering the scatterer locations and scattering coefficients from the acquired data matrix H. A subspace analysis technique can be used to determine the scatterer locations via a MUSIC-like pseudo-distribution [8]. Once the locations are known, the linear equation (2) can be solved for the unknown $\{\tau_i\}$. This equation can be written in matrix notation as

$$H = G_{rec}\, X\, G_{tr}^T, \qquad X \triangleq \operatorname{diag}\{\tau_i\,;\ 1 \le i \le L\} \tag{3a}$$

where

(3b)
$$G_{tr} = \Big[g_{tr}(\chi_1) \;\; g_{tr}(\chi_2) \;\; \ldots \;\; g_{tr}(\chi_L) \Big],\ G_{rec} = \Big[g_{rec}(\chi_1) \;\; g_{rec}(\chi_2) \;\; \ldots \;\; g_{rec}(\chi_L) \Big].$$

Since the unknown matrix X is *diagonal*, eq. (3a) is over-determined (provided that the number of elements in H exceeds L), which suggests using a least squares approach, viz.,

$$X_{opt} \triangleq \arg\min_X \Big\| H - G_{rec}\, X\, G_{tr}^T \Big\|_F^2 \tag{4}$$

subject to the constraint that X is a diagonal matrix [9].

Applying the direct vectorization transformation (1) to $H - G_{rec}\, X\, G_{tr}^T$ results in a highly inefficient least squares problem, because $\operatorname{vec}\{X\}$ is very sparse. In this paper we describe an alternative approach based on:

- a known vectorization identity, viz.,

  $$\operatorname{vec}\{AXB^T\} = (B \odot A)\operatorname{vecd}\{X\}, \quad X \text{ is diagonal} \tag{5}$$

  which involves the so-called *Khatri-Rao matrix product* $\odot$ [5], as well as the *diagonal extraction* operator $\operatorname{vecd}\{X\}$, which forms a column vector consisting of the diagonal elements of the square matrix X, viz.,

$$\mathrm{vecd}\,\{X\} \triangleq \begin{bmatrix} x_{11} & x_{22} & \dots & x_{LL} \end{bmatrix}^T \tag{6}$$

instead of the much longer column vector $\mathrm{vec}\,\{X\}$;

- several new results about the relation between Kronecker, Khatri-Rao and Schur-Hadamard matrix products, which lead to a very efficient computational procedure for solving the matrix least squares problem (4).

We formulate the problem and present our main results in Sec. 2. New results about the "Kronecker to Khatri-Rao to Schur-Hadamard" conversion are derived in Sec. 3, and some concluding remarks are provided in Sec. 4.

2. Problem Formulation and Main Results. We consider the matrix linear least squares (LLS) problem

$$\min_X \left\| Q - A\,X\,B^T \right\|_F^2 \tag{7}$$

where A, B, Q are given (complex valued) matrices of sizes $N_A \times L$, $N_B \times L$, and $N_A \times N_B$, respectively, and where the unknown $L \times L$ matrix X is *diagonal*. We also assume that $L < N_A\,N_B$, so that the linear matrix equation $AXB^T = Q$ is over-determined.

Using the identity (1) we can transform (7) into the vector LLS form

$$\min_X \left\| \mathrm{vec}\,\{Q\} - (B \otimes A)\,\mathrm{vec}\,\{X\} \right\|_2^2$$

which has the well-known solution

$$\mathrm{vec}\,\{X\} = \Big[(B \otimes A)^H (B \otimes A) \Big]^{-1} (B \otimes A)^H\,\mathrm{vec}\,\{Q\}.$$

As we have observed earlier, when the unknown matrix X is diagonal, solving for $\mathrm{vec}\,\{X\}$ is highly inefficient, since most of the elements of X vanish.

Instead we can use the more compact vectorization identity (5) to rewrite the matrix LLS problem (7) in the *reduced-order vector form*

$$\min_X \left\| \mathrm{vec}\,\{Q\} - (B \odot A)\,\mathrm{vecd}\,\{X\} \right\|_2^2 \tag{8}$$

where $\odot$ denotes the *Khatri-Rao matrix product* [5]: the k-th column of $B \odot A$ is the Kronecker product of the k-th column of B by the k-th column of A, for $k = 1, 2, \dots L$. Notice that $\mathrm{vecd}\,\{X\}$ consists of only the nontrivial (i.e., diagonal) elements of the matrix X. The explicit solution of (8) is

$$\mathrm{vecd}\,\{X\} = \Big[(B \odot A)^H (B \odot A) \Big]^{-1} (B \odot A)^H\,\mathrm{vec}\,\{Q\}. \tag{9}$$

It turns out that this expression can also be implemented using Schur-Hadamard products (i.e., element-wise array multiplication), resulting in a significant reduction in computational cost, as implied by the following result.

THEOREM 2.1. *Given two matrices, A (of size $N_A \times L$) and B (of size $N_B \times L$), we have*

$$(A \odot B)^H (A \odot B) = (A^H A) \circ (B^H B) \tag{10}$$

where $\circ$ denotes a Schur-Hadamard matrix product. In addition, if Q is any matrix of size $N_A \times N_B$, then

$$\text{vecd}\,\{A^T Q B\} = (B \odot A)^T \,\text{vec}\,\{Q\}. \tag{11}$$

□

COROLLARY. *When $L < \min\{N_A, N_B\}$ it follows from (10) that*

$$\text{rank}\,\{A \odot B\} = L \quad \Longleftrightarrow \quad (A^H A) \circ (B^H B) > 0 \tag{12a}$$

and thus also

$$\text{rank}\,\{A\} = L = \text{rank}\,\{B\} \quad \Longrightarrow \quad \text{rank}\,\{A \odot B\} = L. \tag{12b}$$

□

The proof of this theorem relies on certain properties of the Khatri-Rao product and the diagonal extraction operator $\text{vecd}\{\cdot\}$, which we establish in the following section. We observe that the left-hand-side expression in (10) requires $N_A N_B L + N_A N_B L(L+1)/2$ multiplications, while forming the equivalent right-hand-side expression requires only $(N_A + N_B + 1)L(L+1)/2$ multiplications. Thus the latter offers significant computational savings, especially when $N_A N_B \gg N_A + N_B + 1$.

Now, using (10) and (11) we can rewrite (9) in the more compact form

$$\text{vecd}\,\{X\} = \Big[(B^H B) \circ (A^H A) \Big]^{-1} \text{vecd}\,\{A^H Q \,\text{conj}(B)\}. \tag{13}$$

The expression (13), which requires $\mathcal{O}(L^3) + \mathcal{O}([N_A + N_B]L^2)$ (multiply and add) operations is much more efficient than (9), which requires $\mathcal{O}(L^3) + \mathcal{O}([N_A N_B]\, L^2)$ operations. The computational advantage of using (13) is particularly evident when the LLS problem (7) is "strongly over-determined," i.e., when

$$L \ll \min(N_A, N_B) \tag{14}$$

which implies that $N_A\, N_B \gg N_A + N_B \gg L$.

In order to be able to use (13) we must ascertain that the matrix $(B^H B) \circ (A^H A)$ is invertible. This will hold, for instance, when both A and B have full column rank. Such is indeed the case in multistatic antenna array processing: both G_{tr} and G_{rec} have full column rank (except in very rare pathological cases [10]). In the full rank case $A^H A > 0$ and $B^H B > 0$, so that their Schur-Hadamard product is positive definite as well [11]. In general, for any two Hermitian positive semidefinite matrices $R = [r_{ij}]$ and $Q = [q_{ij}]$ we have [11]

$$(\min_i q_{ii})\,\lambda_{min}(R) \le \lambda_{min}(R \circ Q) \le \lambda_{max}(R \circ Q) \le (\max_i q_{ii})\,\lambda_{max}(R).$$

In particular, when both matrices are *positive definite*, then $\lambda_{min}(R) > 0$, as well as $q_{ii} > 0$ for all i, so that $\lambda_{min}(R \circ Q) > 0$ and , therefore, $R \circ Q > 0$, as stated.

3. Diagonal Extraction and the Khatri-Rao Product. Given two matrices, A (of size $N_A \times L$) and B (of size $N_B \times L$), let $\{a_i\,;\ 1 \le i \le L\}$ denote the columns of A, and $\{b_i\,;\ 1 \le i \le L\}$ denote the columns of B, namely,

$$A = [\,a_1 \quad a_2 \quad \dots a_L\,], \qquad B = [\,b_1 \quad b_2 \quad \dots b_L\,].$$

The columns of the Kronecker product $A \otimes B$ are $\{a_i \otimes b_j\}$ for all i, j combinations in lexicographic order, namely,

$$A \otimes B = \Big[\, a_1 \otimes b_1 \quad a_1 \otimes b_2 \quad \dots \quad a_1 \otimes b_L \quad a_2 \otimes b_1 \quad a_2 \otimes b_2 \quad \dots \quad a_L \otimes b_L \,\Big].$$

Thus, the Khatri-Rao product

$$A \odot B \triangleq \Big[\, a_1 \otimes b_1 \quad a_2 \otimes b_2 \quad \dots a_L \otimes b_L \,\Big] \tag{15}$$

consists of a subset of the columns of $A \otimes B$. This observation can be expressed in the form $\big(A \otimes B\big)\, S_L = A \odot B$, where the *selection matrix* S_L is

$$S_L \triangleq [\, e_1 \quad e_{L+2} \quad e_{2L+3} \quad \dots \quad e_{L^2} \,] \tag{16a}$$

and e_k is an $L^2 \times 1$ column vector with a unity element in the k-th position and zeros elsewhere, viz.,

$$e_k \triangleq [\, \underbrace{0 \quad \dots \quad 0 \quad 1}_{k} \quad 0 \quad \dots \quad 0 \,]^T, \qquad 1 \le k \le L^2. \tag{16b}$$

Applying the $(L^2 \times L)$ matrix S_L from the right selects only $a_i \otimes b_j$ combinations with $i = j$ so that indeed $\big(A \otimes B\big)\, S_L = A \odot B$.

Next, we observe that for any two sets of columns of the same length N, say $\{a_j\,;\ 1 \le j \le L\}$ and $\{b_j\,;\ 1 \le j \le L\}$, we have

$$a_j \odot b_j \equiv a_j \otimes b_j = \begin{pmatrix} a_{1j}\, b_j \\ a_{2j}\, b_j \\ \vdots \\ a_{Nj}\, b_j \end{pmatrix}.$$

Now, the elements of the $N \times 1$ column vector

$$a_j \circ b_j = \begin{pmatrix} a_{1j}\, b_{1j} \\ a_{2j}\, b_{2j} \\ \vdots \\ a_{Nj}\, b_{Nj} \end{pmatrix}$$

are clearly a subset of the elements of $a_j \otimes b_j$ and, in fact,

$$a_j \circ b_j = S_N^T \left(a_j \otimes b_j \right)$$

so that $S_N^T \left(A \odot B \right) = A \circ B$ for any two matrices A, B *of the same size.*

In summary, we have the following fundamental result, which relates Kronecker, Khatri-Rao and Schur-Hadamard products.

THEOREM 3.1. *Given two matrices,* A *(of size* $N_A \times L$*) and* B *(of size* $N_B \times L$*), we have*

$$\left(A \otimes B \right) S_L = A \odot B \tag{17a}$$

where the selection matrix S_L *is as defined in (16). In addition, if both matrices have the same size (i.e.,* $N_A = N_B = N$*) then we also have*

$$S_N^T \left(A \odot B \right) = A \circ B \tag{17b}$$

and thus also

$$S_N^T \left(A \otimes B \right) S_L = A \circ B. \tag{17c}$$

□

As for the diagonal extraction operator $\mathrm{vecd}\{\cdot\}$, we observe that

$$\mathrm{vecd}\{A\} = S_N^T \, \mathrm{vec}\{A\}$$

for any square $(N \times N)$ matrix $A = \{a_{ij}\,;\ 1 \le i \le N\,,\, 1 \le j \le N\}$. This is so because $\mathrm{vec}\{\cdot\}$ vectorizes a matrix by columns, so that

$$\mathrm{vec}\{A\} = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{N1} & a_{12} & \dots & a_{N2} & \dots & a_{NN} \end{bmatrix}^T$$

and we notice that the diagonal elements $\{a_{11}, a_{22}, \dots, a_{NN}\}$ are evenly spaced within $\mathrm{vec}\{A\}$, occupying the 1-st, $(N+2)$-nd, $(2N+3)$-rd, …, N^2-th positions. Pre-multiplying $\mathrm{vec}\{A\}$ by S_N^T selects the 1-st, $(N+2)$-nd, $(2N+3)$-rd, etc. elements of this vector, which results in the (much shorter) column vector $\begin{bmatrix} a_{11} & a_{22} & \dots & a_{NN} \end{bmatrix}^T \equiv \mathrm{vecd}\{A\}$. Conversely, for a diagonal matrix D, the $N^2 \times 1$ column vector $\mathrm{vec}\{D\}$ is sparse, and can be generated by inserting zeros into $\mathrm{vecd}\{D\}$, viz.,

$$\mathrm{vec}\{D\} = S_N \, \mathrm{vecd}\{D\}.$$

Notice that combining the two last results produces $\mathrm{vecd}\{D\} = S_N^T S_N \, \mathrm{vecd}\{D\}$, which holds true for every $(N \times N)$ diagonal matrix D, so that we must have $S_N^T S_N = I_N$.

In summary, we have established the following result, which relates the vectorization-by-columns operator $\mathrm{vec}\{\cdot\}$ to the diagonal extraction operator $\mathrm{vecd}\{\cdot\}$.

THEOREM 3.2. *Given a square $(N \times N)$ matrix A, we have*

$$\text{vecd}\,\{A\} = S_N^T \,\text{vec}\,\{A\}. \tag{18a}$$

If A is diagonal, then also

$$\text{vec}\,\{A\} = S_N \,\text{vecd}\,\{A\}, \;\; A \textit{ is diagonal.} \tag{18b}$$

Moreover, the columns of the $(N^2 \times N)$ selection matrix S_N are mutually orthonormal, viz.,

$$S_N^T S_N = I_N. \tag{18c}$$

□

Proof of Theorem 2.1. From $\big(A \otimes B\big)\, S_L = A \odot B$ it follows that

$$(A \odot B)^H (A \odot B) = S_L^T\,(A \otimes B)^H (A \otimes B)\, S_L = S_L^T \Big[\,(A^H A) \otimes (B^H B)\Big]\, S_L.$$

Applying (17c) results in $S_L^T \Big[\,(A^H A) \otimes (B^H B)\Big]\, S_L = (A^H A) \circ (B^H B)$, so that

$$(A \odot B)^H (A \odot B) = (A^H A) \circ (B^H B)$$

which establishes (10). Next, observe that for any given matrices A, B, and Q of sizes $N_A \times L$, $N_B \times L$, and $N_A \times N_B$, respectively, we have

$$\begin{aligned} \text{vecd}\,\{A^T Q B\} &= S_L^T \,\text{vec}\,\{A^T Q B\} \\ &= S_L^T\,(B^T \otimes A^T)\,\text{vec}\,\{Q\} = \Big[(B \otimes A) S_L\Big]^T \text{vec}\,\{Q\} \end{aligned}$$

where we used the identities (18a) and (1). In view of (17a) we conclude that

$$\text{vecd}\,\{A^T Q B\} = (B \odot A)^T \;\text{vec}\,\{Q\}$$

which establishes (11). Finally, (17c) is obtained by combining (17a) and (17b), which concludes our proof of the theorem. □

4. Concluding Remarks. We have established an explicit characterization of the mappings

$$A \otimes B \quad \Longrightarrow \quad A \odot B \quad \Longrightarrow \quad A \circ B$$

in terms of the selection matrix S_L (Theorem 3.1). We have also observed that the same matrix relates the two operators $\text{vec}\,\{\cdot\}$ and $\text{vecd}\,\{\cdot\}$ (Theorem 3.2). We used these relations to derive our main result (Theorem 2.1) and, subsequently, to construct a computationally-efficient solution of the matrix least-squares problem (8), requiring $\mathcal{O}(L^3) + \mathcal{O}([N_A + N_B]L^2)$ (multiply and add) operations. In contrast, the most efficient known alternative (i.e., eq. (9)) requires $\mathcal{O}(L^3) + \mathcal{O}([N_A N_B]\, L^2)$

operations, which is significantly higher when $L \ll \min(N_A, N_B)$. Furthermore, preliminary inquiries indicate that our (Schur-Hadamard type) solution (13) is less sensitive to roundoff errors than the known (Khatri-Rao type) solution (9).

The fundamental relations presented in Theorems 3.1 and 3.2 can be exploited to derive a variety of useful results. For instance, (11) implies that, for a *diagonal* matrix D,

$$\begin{aligned}\mathrm{vecd}\,\{A^T DB\} &= (B \odot A)^T\ \mathrm{vec}\,\{D\} = (B \odot A)^T\ S_L\,\mathrm{vecd}\,\{D\} \\ &= \Big[\, S_L^T (B \odot A) \Big]^T\ \mathrm{vecd}\,\{D\} = (B \circ A)^T\ \mathrm{vecd}\,\{D\}\end{aligned}$$

where we used (18b) and (17b). Thus, we get the new identity

$$\mathrm{vecd}\,\{A^T DB\} = (B \circ A)^T\ \mathrm{vecd}\,\{D\} \tag{19}$$

which should be contrasted with the known result (5).

REFERENCES

[1] T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[2] H. Lev-Ari and T. Kailath, *State-Space Approach to Factorization of Lossless Transfer Functions and Structured Matrices*, Linear Algebra and Its Applications, 162-164(1992), pp. 273–295.

[3] H. Lev-Ari, *Displacement Structure: Two Related Perspectives*, in: Communications, Computing, Control and Signal Processing: A Tribute to Thomas Kailath, A. Paulraj, V. Roychowdhury and C. Schaper, eds., pp. 233–241, Kluwer Academic Publishers, Norwell, MA, 1997.

[4] T. Kailath and A.H. Sayed, *Displacement Structure: Theory and Applications*, SIAM Review, 37:3(1995), pp. 297–3865.

[5] J. W. Brewer, *Kronecker Products and Matrix Calculus in System Theory*, IEEE Trans. Circ. & Syst., CAS-25(1978), pp. 772–781.

[6] C. Prada, S. Manneville, D. Spoliansky, and M. Fink, *Decomposition of the Time Reversal Operator: Detection and Selective Focusing on Two Scatterers*, Journal of the Acoustical Society of America, 99(1996), pp. 2067–2076.

[7] J. R. Taylor, *Scattering Theory*, Wiley, New York, 1972.

[8] H. Lev-Ari and A.J. Devaney, *The Time Reversal Technique Re-interpreted: Subspace-Based Signal Processing for Multistatic Target Location*, Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop, Cambridge, MA, March 2000, pp. 509–513.

[9] R. Zandifar, *Super-Resolution Multistatic Target Location*, M.S. Thesis, Department of Electrical and Computer Engineering, Northeastern University, June 2001.

[10] A. J. Devaney, *Super-resolution Processing of Multi-static Data Using Time Reversal and MUSIC*, Journal of the Acoustical Society of America, to appear. Preprint available on the author's web site.

[11] R. A. Horn, *The Hadamard Product*, in: Matrix Theory and Applications, Proceedings of Symposia in Applied Mathematics, Vol. 40, C.R. Johnson (ed.), American Mathematical Society, Providence, RI, 1990.