Available online at www.sciencedirect.com
SCIENCE DIRECT®
LINEAR ALGEBRA AND ITS APPLICATIONS
ELSEVIER Linear Algebra and its Applications 396 (2005) 373–384
www.elsevier.com/locate/laa

The geometric mean decomposition[☆]

Yi Jiang [a], William W. Hager [b,*], Jian Li [c]

[a] *Department of Electrical and Computer Engineering, University of Florida, P.O. Box 116130, Gainesville, FL 32611-6130, USA*
[b] *Department of Mathematics, University of Florida, P.O. Box 118105, Gainesville, FL 32611-8105, USA*
[c] *Department of Electrical and Computer Engineering, University of Florida, P.O. Box 116130, Gainesville, FL 32611-6130, USA*

Received 9 December 2003; accepted 27 September 2004

Submitted by L.N. Trefethen

Abstract

Given a complex matrix $\mathbf{H}$, we consider the decomposition $\mathbf{H} = \mathbf{QRP}^*$ where $\mathbf{Q}$ and $\mathbf{P}$ have orthonormal columns, and $\mathbf{R}$ is a real upper triangular matrix with diagonal elements equal to the geometric mean of the positive singular values of $\mathbf{H}$. This decomposition, which we call the *geometric mean decomposition*, has application to signal processing and to the design of telecommunication networks. The unitary matrices correspond to information lossless filters applied to transmitted and received signals that minimize the maximum error rate of the network. Another application is to the construction of test matrices with specified singular values.

AMS classification: 15A23; 65F25; 94A11; 60G35

Keywords: Geometric mean decomposition; Matrix factorization; Unitary factorization; Singular value decomposition; Schur decomposition; **QR** decomposition; MIMO systems

☆ This material is based upon work supported by the National Science Foundation under Grants CCR-0203270 and CCR-0097114.
* Corresponding author. Tel.: +1 352 392 0281; fax: +1 352 392 8357.
E-mail addresses: yjiang@dsp.ufl.edu (Y. Jiang), hager@math.ufl.edu (W.W. Hager), li@dsp.ufl.edu (J. Li).
URL: http://www.math.ufl.edu/∼hager

0024-3795/$ - see front matter
doi:10.1016/j.laa.2004.09.018

1. Introduction

It is well known that any square matrix has a Schur decomposition $\mathbf{QUQ}^*$, where $\mathbf{U}$ is upper triangular with eigenvalues on the diagonal, $\mathbf{Q}$ is unitary, and * denotes conjugate transpose. In addition, any matrix has a singular value decomposition (SVD) $\mathbf{V\Sigma W}^*$ where $\mathbf{V}$ and $\mathbf{W}$ are unitary, and $\mathbf{\Sigma}$ is completely zero except for the singular values on the diagonal of $\mathbf{\Sigma}$. In this paper, we focus on another unitary decomposition which we call the *geometric mean decomposition* or GMD. Given a rank K matrix $\mathbf{H} \in \mathbb{C}^{m\times n}$, we write it as a product $\mathbf{QRP}^*$ where $\mathbf{P}$ and $\mathbf{Q}$ have orthonormal columns, and $\mathbf{R} \in \mathbb{R}^{K\times K}$ is a real upper triangular matrix with diagonal elements all equal to the geometric mean of the positive singular values:

$$r_{ii} = \bar{\sigma} = \left(\prod_{\sigma_j>0} \sigma_j\right)^{1/K}, \quad 1 \leqslant i \leqslant K.$$

Here the σ_j are the singular values of $\mathbf{H}$, and $\bar{\sigma}$ is the geometric mean of the positive singular values. Thus $\mathbf{R}$ is upper triangular and the nonzero diagonal elements are the geometric mean of the positive singular values.

As explained below, we were led to this decomposition when trying to optimize the performance of multiple-input multiple-output (MIMO) systems. However, this decomposition has arisen recently in several other applications. In [7–Prob. 26.3] Higham proposed the following problem:

> Develop an efficient algorithm for computing a unit upper triangular matrix with prescribed singular values σ_i, $1 \leqslant i \leqslant K$, where the product of the σ_i is 1.

A solution to this problem could be used to construct test matrices with user specified singular values.

The solution of Kosowski and Smoktunowicz [10] starts with the diagonal matrix $\mathbf{\Sigma}$, with ith diagonal element σ_i, and applies a series of 2×2 orthogonal transformations to obtain a unit triangular matrix. The complexity of their algorithm is $\mathrm{O}(K^2)$. Thus the solution given in [10] amounts to the statement

$$\mathbf{Q}_0^{\mathsf{T}}\mathbf{\Sigma P}_0 = \mathbf{R}, \tag{1}$$

where $\mathbf{R}$ is unit upper triangular.

For general $\mathbf{\Sigma}$, where the product of the σ_i is not necessarily 1, one can multiply $\mathbf{\Sigma}$ by the scaling factor $\bar{\sigma}^{-1}$, apply (1), then multiply by $\bar{\sigma}$ to obtain the GMD of $\mathbf{\Sigma}$. And for a general matrix $\mathbf{H}$, the singular value decomposition $\mathbf{H} = \mathbf{V\Sigma W}^*$ and (1) combine to give the $\mathbf{H} = \mathbf{QRP}^*$ where

$$\mathbf{Q} = \mathbf{VQ}_0 \quad \text{and} \quad \mathbf{P} = \mathbf{WP}_0.$$

In a completely different application, [17] considers a signal processing problem: design "precoders for suppressing the intersymbol interference (ISI)." An optimal precoder corresponds to a matrix $\mathbf{F}$ that solves the problem

$$\begin{aligned} &\max_{\mathbf{F}} \min \quad \{u_{ii} : 1 \leqslant i \leqslant K\} \\ &\text{subject to} \quad \mathbf{QU} = \mathbf{HF},\ \operatorname{tr}(\mathbf{F}^*\mathbf{F}) \leqslant p, \end{aligned} \tag{2}$$

where tr denotes the trace of a matrix, p is a given positive scalar called the power constraint, and $\mathbf{QU}$ denotes the factorization of $\mathbf{HF}$ into the product of a matrix with orthonormal columns and an upper triangular matrix with nonnegative diagonal (the $\mathbf{QR}$ factorization). Thus, the goal is to find a matrix $\mathbf{F}$ that complies with the power constraint, and which has the property that the smallest diagonal element of $\mathbf{U}$ is as large as possible. The authors show that if $\mathbf{F}$ is a multiple of a unitary matrix and the diagonal elements of $\mathbf{U}$ are all equal, then $\mathbf{F}$ attains the minimum in (2). Hence, a solution of (2) is

$$\mathbf{F} = \mathbf{P}\sqrt{\frac{p}{K}} \quad \text{and} \quad \mathbf{U} = \mathbf{R}\sqrt{\frac{p}{K}},$$

where $\mathbf{H} = \mathbf{QRP}^*$ is the GMD. In [17] an $\mathrm{O}(K^4)$ algorithm is stated, without proof, for computing an equal diagonal $\mathbf{R}$.

Our own interest in the GMD arose in signal processing in the presence of noise. Possibly, in the next generation of wireless technology, each transmission antenna will be replaced by a cluster of antennas, vastly increasing the amount of data that can be transmitted. The communication network, which can be viewed as a multiple-input multiple-output (MIMO) system, supports significantly higher data rates and offers higher reliability than single-input single-output (SISO) systems [2,12,13].

A MIMO system can be modeled in the following way:

$$\mathbf{y} = \mathbf{Hx} + \mathbf{z}, \tag{3}$$

where $\mathbf{x} \in \mathbb{C}^n$ is the transmitted data, $\mathbf{y} \in \mathbb{C}^m$ is the received data, $\mathbf{z} \in \mathbb{C}^m$ is the noise, and $\mathbf{H}$ is the channel matrix. Suppose $\mathbf{H}$ is decomposed as $\mathbf{H} = \mathbf{QRP}^*$, where $\mathbf{R}$ is a $K \times K$ upper triangular matrix, and $\mathbf{P}$ and $\mathbf{Q}$ are matrices with orthonormal columns that correspond to filters on the transmitted and received signals. With this substitution for $\mathbf{H}$ in (3), we obtain the equivalent system

$$\tilde{\mathbf{y}} = \mathbf{Rs} + \tilde{\mathbf{z}}, \tag{4}$$

with precoding $\mathbf{x} = \mathbf{Ps}$, with decoding $\tilde{\mathbf{y}} = \mathbf{Q}^*\mathbf{y}$, and with noise $\tilde{\mathbf{z}} = \mathbf{Q}^*\mathbf{z}$.

Although the interference with the transmitted signal is known by the transmitter, the off-diagonal elements of $\mathbf{R}$ will inevitably result in interference with the desirable signal at the receiver. However, Costa predicted in [1] the amazing fact that the interference known at the transmitter can be cancelled completely without consuming additional input power. As a realization of Costa's prediction, in [3] the authors applied the so-called Tomlinson–Harashima precodes to convert (4) into K decoupled parallel subchannels:

$$\tilde{y}_i = r_{ii}s_i + \tilde{z}_i, \quad i = 1, \ldots, K.$$

Assuming the variance of the noise on the K subchannels is the same, the subchannel with the smallest r_{ii} has the highest error rate. This leads us to consider the problem of choosing $\mathbf{Q}$ and $\mathbf{P}$ to maximize the minimum of the r_{ii}:

$$\begin{array}{lll} \max\limits_{\mathbf{Q},\mathbf{P}} \min & \{r_{ii} : 1 \leqslant i \leqslant K\} & \\ \text{subject to} & \mathbf{QRP}^* = \mathbf{H},\ \mathbf{Q}^*\mathbf{Q} = \mathbf{I},\ \mathbf{P}^*\mathbf{P} = \mathbf{I}, & (5) \\ & r_{ij} = 0 \text{ for } i > j,\ \mathbf{R} \in \mathbb{R}^{K\times K}, & \end{array}$$

where K is the rank of $\mathbf{H}$. If we take $p = K$ in (2), then any matrix that is feasible in (5) will be feasible in (2). Since the solution of (2) is given by the GMD of $\mathbf{H}$ when $p = K$, and since the GMD of $\mathbf{H}$ is feasible in (5), we conclude that the GMD yields the optimal solution to (5). In [9] we show that a transceiver designed using the GMD can achieve asymptotically optimal capacity and excellent bit error rate performance.

An outline of the paper is as follows: In Section 2, we strengthen the maximin property of the GMD by showing that it achieves the optimal solution of a less constrained problem where the orthogonality constraints are replaced by less stringent trace constraints. Section 3 gives an algorithm for computing the GMD of a matrix, starting from the SVD. Our algorithm, like that in [10], involves permutations and multiplications by 2×2 orthogonal matrices. Unlike the algorithm in [10], we use Givens rotations rather than orthogonal factors gotten from the SVD of 2×2 matrices. And the permutations in our algorithm are done as the algorithm progresses rather than in a preprocessing phase. One advantage of our algorithm is that we are able to use it in [8] to construct a factorization $\mathbf{H} = \mathbf{QRP}^*$, where the diagonal of $\mathbf{R}$ is any vector satisfying Weyl's multiplicative majorization conditions [15]. Section 4 shows how the GMD can be computed directly, without first evaluating the SVD, by a process that combines Lanczos' method with Householder matrices. This version of the algorithm would be useful if the matrix is encoded in a subroutine that returns the product of $\mathbf{H}$ with a vector. In Section 5 we examine the set of 3×3 matrices for which a bidiagonal $\mathbf{R}$ is possible.

2. Generalized maximin properties

Given a rank K matrix $\mathbf{H} \in \mathbb{C}^{m\times n}$, we consider the following problem:

$$\begin{array}{lll} \max\limits_{\mathbf{F},\mathbf{G}} \min & \{|u_{ii}| : 1 \leqslant i \leqslant K\} & \\ \text{subject to} & \mathbf{U} = \mathbf{G}^*\mathbf{HF},\ u_{ij} = 0 \text{ for } i > j,\ \mathbf{U} \in \mathbb{C}^{K\times K}, & (6) \\ & \operatorname{tr}(\mathbf{G}^*\mathbf{G}) \leqslant p_1,\ \operatorname{tr}(\mathbf{F}^*\mathbf{F}) \leqslant p_2. & \end{array}$$

Again, tr denotes the trace of a matrix. If $p_1 = p_2 = K$, then each $\mathbf{Q}$ and $\mathbf{P}$ feasible in (5) is feasible in (6). Hence, problem (6) is less constrained than problem (5) since the set of feasible matrices has been enlarged by the removal of the orthogonality constraints. Nonetheless, we now show that the solution to this relaxed problem is the same as the solution of the more constrained problem (5).

Theorem 1. *If $H \in \mathbb{C}^{m\times n}$ has rank K, then a solution of* (6) *is given by*

$$\mathbf{G} = \mathbf{Q}\sqrt{\frac{p_1}{K}}, \quad \mathbf{U} = \left(\frac{\sqrt{p_1 p_2}}{K}\right)\mathbf{R}, \quad \text{and} \quad \mathbf{F} = \mathbf{P}\sqrt{\frac{p_2}{K}},$$

where $\mathbf{QRP}^*$ *is the GMD of* $\mathbf{H}$.

Proof. Let $\mathbf{F}$ and $\mathbf{G}$ satisfy the constraints of (6). Let $\bar{\mathbf{G}}$ be the matrix obtained by replacing each column of $\mathbf{G}$ by its projection into the column space of $\mathbf{H}$. Let $\bar{\mathbf{F}}$ be the matrix obtained by replacing each column of $\mathbf{F}$ by its projection into the column space of $\mathbf{H}^*$. Since the columns of $\mathbf{G} - \bar{\mathbf{G}}$ are orthogonal to the columns of $\bar{\mathbf{G}}$, we have

$$\mathbf{G}^*\mathbf{G} = (\mathbf{G} - \bar{\mathbf{G}})^*(\mathbf{G} - \bar{\mathbf{G}}) + \bar{\mathbf{G}}^*\bar{\mathbf{G}}.$$

Since the diagonal of a matrix of the form $\mathbf{M}^*\mathbf{M}$ is nonnegative and since $\operatorname{tr}(\mathbf{G}^*\mathbf{G}) \leqslant p_1$, we conclude that

$$\operatorname{tr}(\bar{\mathbf{G}}^*\bar{\mathbf{G}}) \leqslant p_1.$$

By similar reasoning,

$$\operatorname{tr}(\bar{\mathbf{F}}^*\bar{\mathbf{F}}) \leqslant p_2.$$

Since the columns of $\mathbf{G} - \bar{\mathbf{G}}$ are orthogonal to the columns of $\mathbf{H}$ and since the columns of $\mathbf{F} - \bar{\mathbf{F}}$ are orthogonal to the columns of $\mathbf{H}^*$, we have $(\mathbf{G} - \bar{\mathbf{G}})^*\mathbf{H} = \mathbf{0} = \mathbf{H}(\mathbf{F} - \bar{\mathbf{F}})$; it follows that

$$\mathbf{U} = \mathbf{G}^*\mathbf{H}\mathbf{F} = \bar{\mathbf{G}}^*\mathbf{H}\bar{\mathbf{F}}.$$

In summary, given any $\mathbf{G}$, $\mathbf{F}$, and $\mathbf{U}$ that are feasible in (6), there exist $\bar{\mathbf{G}}$ and $\bar{\mathbf{F}}$ such that $\bar{\mathbf{G}}$, $\bar{\mathbf{F}}$, and $\mathbf{U}$ are feasible in (6), the columns of $\bar{\mathbf{G}}$ are contained in the column space of $\mathbf{H}$, and the columns of $\bar{\mathbf{F}}$ are contained in the column space of $\mathbf{H}^*$.

We now prove that

$$\min_{1\leqslant i\leqslant K} |u_{ii}| \leqslant \frac{\bar{\sigma}\sqrt{p_1 p_2}}{K},$$

whenever $\mathbf{U}$ is feasible in (6). By the analysis given above, there exist $\mathbf{G}$ and $\mathbf{F}$ associated with the feasible $\mathbf{U}$ in (6) with the columns of $\mathbf{G}$ and $\mathbf{F}$ contained in the column spaces of $\mathbf{H}$ and $\mathbf{H}^*$ respectively. Let $\mathbf{V\Sigma W}^*$ be the singular value decomposition of $\mathbf{H}$, where $\mathbf{\Sigma} \in \mathbb{R}^{K\times K}$ contains the K positive singular values of $\mathbf{H}$ on the diagonal. Since the column space of $\mathbf{V}$ coincides with the column space of $\mathbf{H}$, there exists a square matrix $\mathbf{A}$ such that $\mathbf{G} = \mathbf{VA}$. Since the column space of $\mathbf{W}$ coincides with the column space of $\mathbf{H}^*$, there exists a square matrix $\mathbf{B}$ such that $\mathbf{F} = \mathbf{WB}$. Hence, we have

$$\begin{aligned}\mathbf{U} &= \mathbf{G}^*\mathbf{H}\mathbf{F} = (\mathbf{VA})^*\mathbf{H}(\mathbf{WB}) \\ &= (\mathbf{VA})^*\mathbf{V\Sigma W}^*(\mathbf{WB}) = \mathbf{A}^*\mathbf{\Sigma}\mathbf{B}.\end{aligned}$$

It follows that

$$\min_{1\leqslant i\leqslant K} |u_{ii}|^{2K} \leqslant \prod_{i=1}^{K} |u_{ii}|^2 = \det(\mathbf{U}^*\mathbf{U})$$
$$= \det(\mathbf{\Sigma}^*\mathbf{\Sigma})\det(\mathbf{A}^*\mathbf{A})\det(\mathbf{B}^*\mathbf{B})$$
$$= \det(\mathbf{\Sigma}^*\mathbf{\Sigma})\det(\mathbf{G}^*\mathbf{G})\det(\mathbf{F}^*\mathbf{F})$$
$$= \bar{\sigma}^{2K}\det(\mathbf{G}^*\mathbf{G})\det(\mathbf{F}^*\mathbf{F}), \tag{7}$$

where det denotes the determinant of a matrix. By the geometric mean inequality and the fact that the determinant (trace) of a matrix is the product (sum) of the eigenvalues,

$$\det(\mathbf{G}^*\mathbf{G}) \leqslant \left(\frac{\mathrm{tr}(\mathbf{G}^*\mathbf{G})}{K}\right)^K \leqslant \left(\frac{p_1}{K}\right)^K \quad \text{and}$$
$$\det(\mathbf{F}^*\mathbf{F}) \leqslant \left(\frac{\mathrm{tr}(\mathbf{F}^*\mathbf{F})}{K}\right)^K \leqslant \left(\frac{p_2}{K}\right)^K.$$

Combining this with (7) gives

$$\min_{1\leqslant i\leqslant K} |u_{ii}| \leqslant \frac{\bar{\sigma}\sqrt{p_1 p_2}}{K}. \tag{8}$$

Finally, it can be verified that the choices for $\mathbf{G}$, $\mathbf{U}$, and $\mathbf{F}$ given in the statement of the theorem satisfy the constraints of (6) and the inequality (8) is an equality. □

3. Implementation based on SVD

In this section, we prove the following theorem by providing a constructive algorithm for computing the GMD:

Theorem 2. *If* $\mathbf{H} \in \mathbb{C}^{m\times n}$ *has rank* K, *then there exist matrices* $\mathbf{P} \in \mathbb{C}^{n\times K}$ *and* $\mathbf{Q} \in \mathbb{C}^{m\times K}$ *with orthonormal columns, and an upper triangular matrix* $\mathbf{R} \in \mathbb{R}^{K\times K}$ *such that* $\mathbf{H} = \mathbf{QRP}^*$, *where the diagonal elements of* $\mathbf{R}$ *are all equal to the geometric mean of the positive singular values.*

Our algorithm for evaluating the GMD starts with the singular value decomposition $\mathbf{H} = \mathbf{V\Sigma W}^*$, and generates a sequence of upper triangular matrices $\mathbf{R}^{(L)}$, $1 \leqslant L < K$, with $\mathbf{R}^{(1)} = \mathbf{\Sigma}$. Each matrix $\mathbf{R}^{(L)}$ has the following properties:

(a) $r_{ij}^{(L)} = 0$ when $i > j$ or $j > \max\{L, i\}$,
(b) $r_{ii}^{(L)} = \bar{\sigma}$ for all $i < L$, and the geometric mean of $r_{ii}^{(L)}$, $L \leqslant i \leqslant K$, is $\bar{\sigma}$.

We express $\mathbf{R}^{(k+1)} = \mathbf{Q}_k^{\top}\mathbf{R}^{(k)}\mathbf{P}_k$ where $\mathbf{Q}_k$ and $\mathbf{P}_k$ are orthogonal for each k.

These orthogonal matrices are constructed using a symmetric permutation and a pair of Givens rotations. Suppose that $\mathbf{R}^{(k)}$ satisfies (a) and (b). If $r_{kk}^{(k)} \geqslant \bar{\sigma}$, then let $\mathbf{\Pi}$ be a permutation matrix with the property that $\mathbf{\Pi R}^{(k)}\mathbf{\Pi}$ exchanges the $(k+1)$st diagonal element of $\mathbf{R}^{(k)}$ with any element r_{pp}, $p > k$, for which $r_{pp} \leqslant \bar{\sigma}$. If $r_{kk}^{(k)} < \bar{\sigma}$, then let $\mathbf{\Pi}$ be chosen to exchange the $(k+1)$st diagonal element with any element r_{pp}, $p > k$, for which $r_{pp} \geqslant \bar{\sigma}$. Let $\delta_1 = r_{kk}^{(k)}$ and $\delta_2 = r_{pp}^{(k)}$ denote the new diagonal elements at locations k and $k+1$ associated with the permuted matrix $\mathbf{\Pi R}^{(k)}\mathbf{\Pi}$.

Next, we construct orthogonal matrices $\mathbf{G}_1$ and $\mathbf{G}_2$ by modifying the elements in the identity matrix that lie at the intersection of rows k and $k+1$ and columns k and $k+1$. We multiply the permuted matrix $\mathbf{\Pi R}^{(k)}\mathbf{\Pi}$ on the left by $\mathbf{G}_2^{\mathsf{T}}$ and on the right by $\mathbf{G}_1$. These multiplications will change the elements in the 2×2 submatrix at the intersection of rows k and $k+1$ with columns k and $k+1$. Our choice for the elements of $\mathbf{G}_1$ and $\mathbf{G}_2$ is shown below, where we focus on the relevant 2×2 submatrices of $\mathbf{G}_2^{\mathsf{T}}$, $\mathbf{\Pi R}^{(k)}\mathbf{\Pi}$, and $\mathbf{G}_1$:

$$\underset{(\mathbf{G}_2^{\mathsf{T}})}{\bar{\sigma}^{-1}\begin{bmatrix} c\delta_1 & s\delta_2 \\ -s\delta_2 & c\delta_1 \end{bmatrix}} \underset{(\mathbf{\Pi R}^{(k)}\mathbf{\Pi})}{\begin{bmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{bmatrix}} \underset{(\mathbf{G}_1)}{\begin{bmatrix} c & -s \\ s & c \end{bmatrix}} = \underset{(\mathbf{R}^{(k+1)})}{\begin{bmatrix} \bar{\sigma} & x \\ 0 & y \end{bmatrix}}. \tag{9}$$

If $\delta_1 = \delta_2 = \bar{\sigma}$, we take $c = 1$ and $s = 0$; if $\delta_1 \neq \delta_2$, we take

$$c = \sqrt{\frac{\bar{\sigma}^2 - \delta_2^2}{\delta_1^2 - \delta_2^2}} \quad \text{and} \quad s = \sqrt{1 - c^2}. \tag{10}$$

In either case,

$$x = \frac{sc(\delta_2^2 - \delta_1^2)}{\bar{\sigma}} \quad \text{and} \quad y = \frac{\delta_1\delta_2}{\bar{\sigma}}. \tag{11}$$

Since $\bar{\sigma}$ lies between δ_1 and δ_2, s and c are nonnegative real scalars.

Fig. 1 depicts the transformation from $\mathbf{R}^{(k)}$ to $\mathbf{G}_2^{\mathsf{T}}\mathbf{\Pi R}^{(k)}\mathbf{\Pi G}_1$. The dashed box is the 2×2 submatrix displayed in (9). Notice that c and s, defined in (10), are real scalars chosen so that

$$c^2 + s^2 = 1 \quad \text{and} \quad (c\delta_1)^2 + (s\delta_2)^2 = \bar{\sigma}^2.$$

With these identities, the validity of (9) follows by direct computation. Defining $\mathbf{Q}_k = \mathbf{\Pi G}_2$ and $\mathbf{P}_k = \mathbf{\Pi G}_1$, we set

$$\mathbf{R}^{(k+1)} = \mathbf{Q}_k^{\mathsf{T}}\mathbf{R}^{(k)}\mathbf{P}_k. \tag{12}$$

It follows from Fig. 1, (9), and the identity $\det(\mathbf{R}^{(k+1)}) = \det(\mathbf{R}^{(k)})$, that (a) and (b) hold for $L = k+1$. Thus there exists a real upper triangular matrix $\mathbf{R}^{(K)}$, with $\bar{\sigma}$ on the diagonal, and unitary matrices $\mathbf{Q}_i$ and $\mathbf{P}_i$, $i = 1, 2, \ldots, K-1$, such that

$$\mathbf{R}^{(K)} = (\mathbf{Q}_{k-1}^{\mathsf{T}} \cdots \mathbf{Q}_2^{\mathsf{T}}\mathbf{Q}_1^{\mathsf{T}})\mathbf{\Sigma}(\mathbf{P}_1\mathbf{P}_2 \cdots \mathbf{P}_{k-1}).$$

$$\begin{array}{cc} & \begin{array}{cccccc} & & \text{Column } k & & & \\ & & \downarrow & & & \end{array} \\ \begin{array}{c} \\ \\ \text{Row } k \rightarrow \\ \\ \\ \\ \end{array} & \begin{array}{cccccc} X & X & X & 0 & 0 & 0 \\ & X & X & 0 & 0 & 0 \\ & & X & 0 & 0 & 0 \\ & & & X & 0 & 0 \\ & & & & X & 0 \\ & & & & & X \end{array} \end{array} \qquad \begin{array}{cccccc} X & X & X & X & 0 & 0 \\ & X & X & X & 0 & 0 \\ & & X & X & 0 & 0 \\ & & & X & 0 & 0 \\ & & & & X & 0 \\ & & & & & X \end{array}$$

$$\mathrm{R}^{(k)} \longrightarrow \mathrm{G}_2^* \Pi \mathrm{R}^{(k)} \Pi \mathrm{G}_1$$

Fig. 1. The operation displayed in (9).

Combining this identity with the singular value decomposition, we obtain $\mathbf{H} = \mathbf{QRP}^*$ where

$$\mathbf{Q} = \mathbf{V}\left(\prod_{i=1}^{K-1} \mathbf{Q}_i\right), \quad \mathbf{R} = \mathbf{R}^{(K)}, \quad \text{and} \quad \mathbf{P} = \mathbf{W}\left(\prod_{i=1}^{K-1} \mathbf{P}_i\right).$$

In summary, our algorithm for computing the GMD, based on an initial SVD, is the following:

(1) Let $\mathbf{H} = \mathbf{V\Sigma W}^*$ be the singular value decomposition of $\mathbf{H}$, and initialize $\mathbf{Q} = \mathbf{V}$, $\mathbf{P} = \mathbf{W}$, $\mathbf{R} = \mathbf{\Sigma}$, and $k = 1$,
(2) If $r_{kk} \geqslant \bar{\sigma}$, choose $p > k$ such that $r_{pp} \leqslant \bar{\sigma}$. If $r_{kk} < \bar{\sigma}$, choose $p > k$ such that $r_{pp} \geqslant \bar{\sigma}$. In $\mathbf{R}$, $\mathbf{P}$, and $\mathbf{Q}$, perform the following exchanges:
$r_{k+1,k+1} \leftrightarrow r_{pp}$,
$\mathbf{P}_{:,k+1} \leftrightarrow \mathbf{P}_{:,p}$,
$\mathbf{Q}_{:,k} \leftrightarrow \mathbf{Q}_{:,p}$,
(3) Construct the matrices $\mathbf{G}_1$ and $\mathbf{G}_2$ shown in (9). Replace $\mathbf{R}$ by $\mathbf{G}_2^{\mathsf{T}}\mathbf{R}\mathbf{G}_1$, replace $\mathbf{Q}$ by $\mathbf{QG}_2$, and replace $\mathbf{P}$ by $\mathbf{PG}_1$,
(4) If $k = K - 1$, then stop, $\mathbf{QRP}^*$ is the GMD of $\mathbf{H}$. Otherwise, replace $k \times k + 1$ and go to step 2.

A Matlab implementation of this algorithm for the GMD is posted at the following web site:

http://www.math.ufl.edu/~hager/papers/gmd.m

Given the SVD, this algorithm for the GMD requires $O((m+n)K)$ flops. For comparison, reduction of $\mathbf{H}$ to bidiagonal form by the Golub–Kahan bidiagonalization scheme [4] (also see [5,6,14,16]), often the first step in the computation of the SVD, requires $O(mnK)$ flops.

4. The unitary update

In Section 3, we construct the successive columns of the upper triangular matrix $\mathbf{R}$ in the GMD by applying a unitary transformation. We now give a different view of this unitary update. This new view leads to a direct algorithm for computing the GMD, without having to compute the SVD.

Since any matrix can be unitarily reduced to a real matrix (for example, by using the Golub–Kahan bidiagonalization [4]), we assume, without loss of generality, that $\mathbf{H}$ is real. The first step in the unitary update is to generate a unit vector $\mathbf{p}$ such that $\|\mathbf{Hp}\| = \bar{\sigma}$, where the norm is the Euclidean norm. Such a vector must exist for the following reason: If $\mathbf{v}_1$ and $\mathbf{v}_2$ are right singular vectors of unit length associated with the largest and smallest singular values of $\mathbf{H}$ respectively, then

$$\|\mathbf{H}\mathbf{v}_1\| \geqslant \bar{\sigma} \geqslant \|\mathbf{H}\mathbf{v}_2\|, \tag{13}$$

that is, the geometric mean of the singular values lies between the largest and the smallest singular value. Let $\mathbf{v}(\theta)$ be the vector obtained by rotating $\mathbf{v}_1$ through an angle θ toward $\mathbf{v}_2$. Since $\mathbf{v}_1$ is perpendicular to $\mathbf{v}_2$, we have $\mathbf{v}(0) = \mathbf{v}_1$ and $\mathbf{v}(\pi/2) = \mathbf{v}_2$. Since $\mathbf{v}(\theta)$ is a continuous function of θ and (13) holds, there exists $\bar{\theta} \in [0, \pi/2]$ such that $\|\mathbf{H}\mathbf{v}(\bar{\theta})\| = \bar{\sigma}$. We take $\mathbf{p} = \mathbf{v}(\bar{\theta})$, which is a unit vector since a rotation does not change length.

Note that we do not need to compute extreme singular vectors, we only need to find approximations satisfying (13). Such approximations can be generated by an iterative process such as Lanczos' method (see [11–Chapter 13]). Let $\mathbf{P}_1$ and $\mathbf{Q}_1$ be orthogonal matrices with first columns $\mathbf{p}$ and $\mathbf{Hp}/\bar{\sigma}$ respectively. These orthogonal matrices, which must exist since $\mathbf{p}$ and $\mathbf{Hp}/\bar{\sigma}$ are unit vectors, can be expressed in terms of Householder reflections [6–p. 210]. By the design of $\mathbf{P}_1$ and $\mathbf{Q}_1$, we have

$$\mathbf{Q}_1^{\mathsf{T}}\mathbf{H}_1\mathbf{P}_1 = \begin{bmatrix} \bar{\sigma} & \mathbf{z}_2 \\ \mathbf{0} & \mathbf{H}_2 \end{bmatrix},$$

where $\mathbf{H}_1 = \mathbf{H}$, $\mathbf{z}_2 \in \mathbb{R}^{n-1}$, and $\mathbf{H}_2 \in \mathbb{R}^{(m-1)\times(n-1)}$.

The reduction to triangular form continues in this way; after $k-1$ steps, we have

$$\left(\prod_{j=1}^{k-1} \mathbf{Q}_j\right)^{\mathsf{T}} \mathbf{\Sigma} \left(\prod_{j=1}^{k-1} \mathbf{P}_j\right) = \begin{bmatrix} \mathbf{R}_k & \mathbf{Z}_k \\ \mathbf{0} & \mathbf{H}_k \end{bmatrix}, \tag{14}$$

where $\mathbf{R}_k$ is a $k \times k$ upper triangular matrix with $\bar{\sigma}$ on the diagonal, the $\mathbf{Q}_j$ and $\mathbf{P}_j$ are orthogonal, $\mathbf{0}$ denotes a matrix whose entries are all 0, and the geometric mean of the singular values of $\mathbf{H}_k$ is $\bar{\sigma}$. In the next step, we take

$$\mathbf{P}_k = \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{P}} \end{bmatrix} \quad \text{and} \quad \mathbf{Q}_k = \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{Q}} \end{bmatrix}, \tag{15}$$

where $\mathbf{I}_k$ is a $k \times k$ identity matrix. The first column $\bar{\mathbf{p}}$ of $\bar{\mathbf{P}}$ is chosen so that $\|\mathbf{H}_k\bar{\mathbf{p}}\| = \bar{\sigma}$, while the first column of $\bar{\mathbf{Q}}$ is $\mathbf{H}_k\bar{\mathbf{p}}/\bar{\sigma}$.

This algorithm generates the GMD directly from **H**, without first computing the SVD. In the case that **H** is **Σ**, the algorithm of Section 3 corresponds to a vector **p** in this section that is entirely zero except for two elements containing c and s. Note that the value of $\bar{\sigma}$ could be computed from an **LU** factorization (for example, see [14–Theorem 5.6]).

5. Bidiagonal GMD?

From dimensional analysis, one could argue that a bidiagonal **R** is possible. The K singular values of a rank K matrix represent K degrees of freedom. Likewise, a $K \times K$ bidiagonal matrix with constant diagonal has K degrees of freedom, the diagonal element and the $K-1$ superdiagonal elements. Nonetheless, we now observe that a bidiagonal **R** is not possible for all choices **H**.

Let us consider the collection of 3×3 diagonal matrices **Σ** with diagonal elements σ_1, σ_2, and σ_3 normalized so that

$$\sigma_1\sigma_2\sigma_3 = 1. \tag{16}$$

In this case, $\bar{\sigma} = 1$. Suppose that

$$\mathbf{\Sigma} = \mathbf{QBP}^*,$$

where **B** is an upper bidiagonal matrix with $\bar{\sigma} = 1$ on the diagonal. Since the singular values of **Σ** and **B** coincide, there exist values a and b for the superdiagonal elements b_{12} and b_{23} for which the singular values of **B** are σ_1, σ_2, and σ_3.

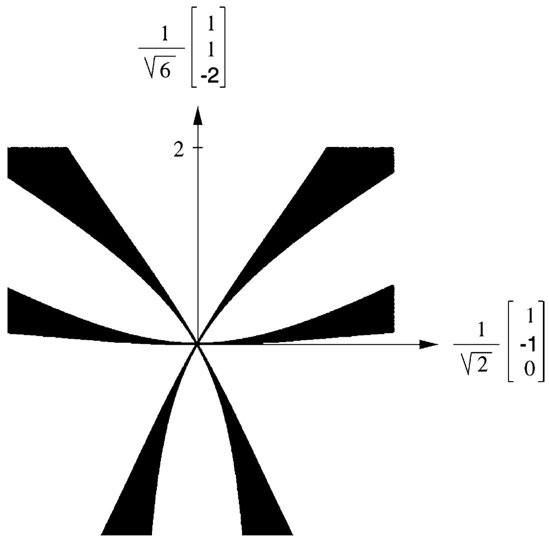

Fig. 2. Projection of feasible singular values onto a 4×4 square oriented perpendicular to (1,1,1).

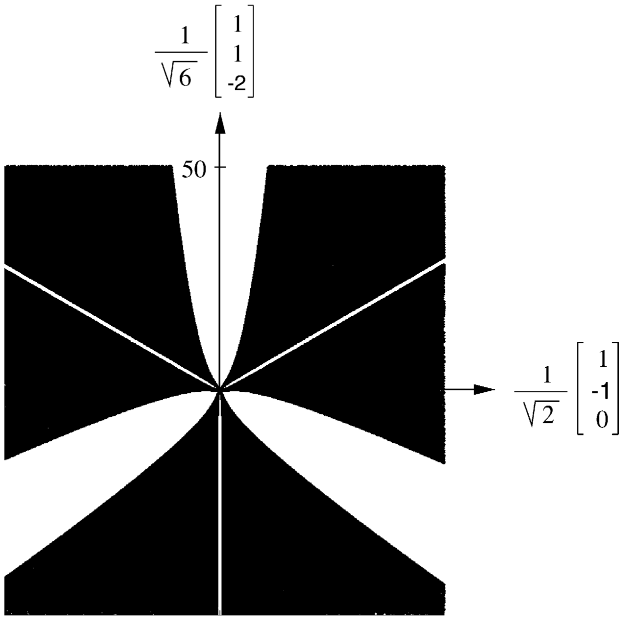

Fig. 3. Projection of feasible singular values onto a 100×100 square oriented perpendicular to (1,1,1).

Figs. 2 and 3 are generated in the following way: For each pair (a, b), we compute the singular values σ_1, σ_2, and σ_3 of $\mathbf{B}$, and we project the computed 3-tuple $(\sigma_1, \sigma_2, \sigma_3)$ onto the plane through the origin that is perpendicular to the vector (1, 1, 1). A dot is placed on the plane at the location of the projected point; as an increasing number of pairs (a, b) is considered, the dots form a shaded region associated with the singular values for which a bidiagonal GMD can be achieved. The white region corresponds to singular values for which a bidiagonal $\mathbf{R}$ is impossible.

Figs. 2 and 3 show the shaded regions associated with 4×4 and 100×100 squares centered at the origin in the plane perpendicular to (1,1,1). The origin of either square corresponds to the singular values $\sigma_1 = \sigma_2 = \sigma_3 = 1$. Fig. 2 is a speck at the center of Fig. 3. These figures indicate that for well-conditioned matrices, with singular values near $\sigma_1 = \sigma_2 = \sigma_3 = 1$, a bidiagonal GMD is not possible, in most cases. As the matrix becomes more ill-conditioned (see Fig. 3), the relative size of the shaded region, associated with a bidiagonal GMD, increases relative to the white region, where a bidiagonal GMD is impossible.

6. Conclusions

The geometric mean decomposition $\mathbf{H} = \mathbf{QRP}^*$, where the columns of $\mathbf{Q}$ and $\mathbf{P}$ are orthonormal and $\mathbf{R}$ is a real upper triangular matrix with diagonal elements all equal to the geometric mean of the positive singular values of $\mathbf{H}$, yields a solution to the maximin problem (6); the smallest diagonal element of $\mathbf{R}$ is as large as possible. In MIMO systems, this minimizes the worst possible error in the transmission process. Other applications of the GMD are to precoders for suppressing intersymbol interference, and to the generation of test matrices with prescribed singular values. Starting with the SVD, we show in Section 3 that the GMD can be computed using a series of Givens rotations, and row and column exchanges. Alternatively, the GMD could be computed directly, without performing an initial SVD, using a Lanczos process and Householder matrices. In a further extension of our algorithm for the GMD, we show in [8] how to compute a factorization $\mathbf{H} = \mathbf{QRP}^*$ where the diagonal of $\mathbf{R}$ is any vector satisfying Weyl's multiplicative majorization conditions [15].

References

[1] M. Costa, Writing on dirty paper, IEEE Trans. Inf. Theory 29 (1983) 439–441.
[2] G.J. Foschini, M.J. Gans, On limits of wireless communications in a fading environment when using multiple antennas, Wireless Person. Commun. 6 (1998) 311–335.
[3] G. Ginis, J.M. Cioffi, Vectored transmission for digital subscriber line systems, IEEE J. Select. Areas Commun. 20 (2002) 1085–1104.
[4] G.H. Golub, W. Kahan, Calculating the singular values and pseudo-inverse of a matrix, SIAM J. Numer. Anal. 2 (1965) 205–224.
[5] G.H. Golub, C.F. Van Loan, Matrix Computations, Johns Hopkins University Press, Baltimore, MD, 1996.
[6] W.W. Hager, Applied Numerical Linear Algebra, Prentice-Hall, Englewood Cliffs, NJ, 1988.
[7] N.J. Higham, Accuracy and Stability of Numerical Algorithms, SIAM, Philadelphia, 1996.
[8] Y. Jiang, W.W. Hager, J. Li, The generalized triangular decomposition, SIAM J. Matrix Anal. Appl., submitted for publication.
[9] Y. Jiang, J. Li, W.W. Hager, Joint transceiver design for MIMO systems using geometric mean decomposition, IEEE Trans. Signal Process., accepted for publication.
[10] P. Kosowski, A. Smoktunowicz, On constructing unit triangular matrices with prescribed singular values, Computing 64 (2000) 279–285.
[11] B.N. Parlett, The Symmetric Eigenvalue Problem, Prentice-Hall, Englewood Cliffs, NJ, 1980.
[12] G.G. Raleigh, J.M. Cioffi, Spatial–temporal coding for wireless communication, IEEE Trans. Commun. 46 (1998) 357–366.
[13] I.E. Telatar, Capacity of multi-antenna Gaussian channels, European Trans. Telecommun. 10 (1999) 585–595.
[14] L.N. Trefethen, D. Bau III, Numerical Linear Algebra, SIAM, Philadelphia, 1997.
[15] H. Weyl, Inequalities between two kinds of eigenvalues of a linear transformation, Proc. Nat. Acad. Sci. USA 35 (1949) 408–411.
[16] J.H. Wilkinson, C. Reinsch, Linear algebra, in: F.L. Bauer (Ed.), Handbook for Automatic Computation, vol. 2, Springer-Verlag, Berlin, 1971.
[17] J.-K. Zhang, A. Kavčić, X. Ma, K.M. Wong, Design of unitary precoders for ISI channels, in: Proceedings IEEE International Conference on Acoustics Speech and Signal Processing, vol. 3, IEEE, Orlando, FL, 2002, pp. 2265–2268 .