MATRIX RECONSTRUCTION WITH PRESCRIBED DIAGONAL ELEMENTS, EIGENVALUES, AND SINGULAR VALUES

DRAFT AS OF April 30, 2013

SHENG-JHIH WU* AND MOODY T. CHU†

Abstract. Diagonal entries and eigenvalues of a Hermitian matrix, diagonal entries and singular values of a general matrix, and eigenvalues and singular values of a general matrix satisfy necessarily some majorization relationships which turn out also to be sufficient conditions. The inverse problem of numerically reconstructing a matrix to satisfy any given set of data meeting one of these relationships has been solved. In this paper, we take one step further to construct a matrix satisfying prescribed diagonal elements, eigenvalues, and singular values simultaneously. Theory of the existence of such a matrix is established and a numerical method is proposed.

Key words. inverse eigenvalue problem, Sing-Thompson theorem, Weyl-Horn theorem, Mirsky theorem, majorization, matrix reconstruction, projected gradient flow

AMS subject classifications. 65F18, 15A29, 90C52, 68W25,

1. Introduction. Let $\mathbf{d} \in \mathbb{R}^n$ be a real vector with entries arranged in the order $|d_1| \geq \ldots \geq |d_n|$, $\boldsymbol{\sigma} \in \mathbf{R}^n$ be a nonnegative vector with entries in the order $\sigma_1 \geq \ldots \geq \sigma_n \geq 0$, and $\boldsymbol{\lambda} \in \mathbb{C}^n$ be a complex vector whose entries are closed under complex conjugation and are ordered as $|\lambda_1| \geq \ldots \geq |\lambda_n|$. Represented in Figure 1.1 are classical results in the field of matrix theory concerning the relationships among diagonal entries, eigenvalues, and singular values. For the completeness of this paper, we summarize these theorems as follows. An interesting fact is that all these relationships involve a sequence of inequalities known as the notion of majorization [20].

THEOREM 1.1. *(Sing-Thompson Theorem [23, 24]) There exists a real matrix $A \in \mathbb{R}^{n\times n}$ with singular values $\boldsymbol{\sigma}$ and main diagonal entries* $\mathbf{d}$*, possibly in different order, if and only if*

$$\sum_{i=1}^{k} |d_i| \leq \sum_{i=1}^{k} \sigma_i \tag{1.1}$$

for all $k = 1, 2, \ldots, n$ and

$$\sum_{i=1}^{n-1} |d_i| - |d_n| \leq \sum_{i=1}^{n-1} \sigma_i - \sigma_n. \tag{1.2}$$

THEOREM 1.2. *(Weyl-Horn Theorem [13, 25]) There exists a real matrix $A \in \mathbb{R}^{n\times n}$ with singular values $\boldsymbol{\sigma}$ and eigenvalues $\boldsymbol{\lambda}$ if and only if*

$$\prod_{i=1}^{k} |\lambda_i| \leq \prod_{i=1}^{k} \sigma_i, \tag{1.3}$$

for all $k = 1, 2, \ldots, n-1$ and

$$\prod_{i=1}^{n} |\lambda_i| = \prod_{i=1}^{n} \sigma_i. \tag{1.4}$$

*Institute of Mathematics, Academia Sinica, 6F, Astronomy-Mathematics Building, No.1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan (sjw@math.sinica.edu.tw)

†Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205. (chu@math.ncsu.edu) This research was supported in part by the National Science Foundation under grant DMS-1014666.

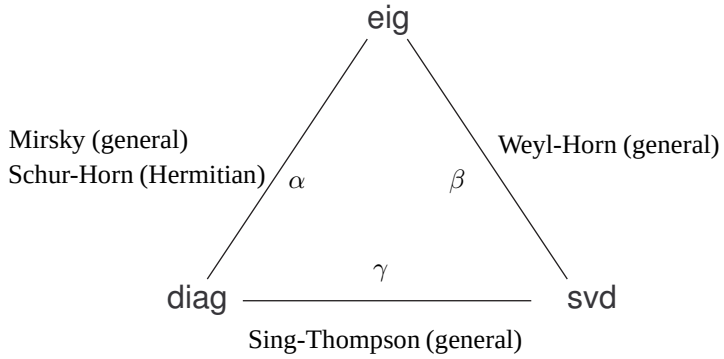


FIGURE 1.1. *Existing results concerning majorization.*

In regard to the relationship between diagonal entries and eigenvalues, we have two separate results, depending upon whether the underlying matrix is Hermitian or not.

THEOREM 1.3. *(Mirsky [21]) There exists a real matrix $A \in \mathbb{R}^{n\times n}$ with eigenvalues $\boldsymbol{\lambda}$ and main diagonal entries $\mathbf{d}$, possibly in different order, if and only if*

$$\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} d_i. \tag{1.5}$$

THEOREM 1.4. *(Schur-Horn Theorem [2]) Suppose that entries of $\boldsymbol{\lambda} \in \mathbb{R}^n$ are arranged in the order $\lambda_1 \geq \ldots \geq \lambda_n$ and entries of $\mathbf{d} \in \mathbb{R}^n$ in the order $d_1 \geq \ldots \geq d_n$. There exists a Hermitian matrix H with eigenvalues $\boldsymbol{\lambda}$ and diagonal entries $\mathbf{d}$, possibly in different order, if and only if*

$$\sum_{i=1}^{k} \lambda_{n-i+1} \leq \sum_{i=1}^{k} d_{n-i+1} \tag{1.6}$$

for all $k = 1, 2, \ldots, n$ and

$$\sum_{i=1}^{n} \lambda_{n-i+1} = \sum_{i=1}^{n} d_{n-i+1}. \tag{1.7}$$

What makes these results significant is that the conditions specified in each of the four theorems are both necessary and sufficient. Given a set of data satisfying a necessary condition, then a challenging but interesting task is to numerically construct a matrix with the prescribed characteristics. Such an inverse problem has been studied in the literature [3, 5, 6, 7, 11, 17, 26]. One fundamental trait associated with these inverse problems is that the solution is not unique. The algorithms developed thus far can serve as constructive proofs of the above theorems, namely, the existence of a matrix, but can hardly pinpoint a specific matrix. For instance, starting with a given matrix $A \in \mathbb{R}^{n\times n}$, we can calculate its eigenvalues $\boldsymbol{\lambda}$ and singular values $\boldsymbol{\sigma}$ which necessarily satisfy the inequalities (1.3) and (1.4). Applying the divide-and-conquer algorithm proposed in [7] to the set of data $\boldsymbol{\lambda}$ and $\boldsymbol{\sigma}$, we can construct a matrix B which has the very same eigenvalues $\boldsymbol{\lambda}$ and singular values $\boldsymbol{\sigma}$. More often than not, however, B is entirely different from the original A. Such a discretion can easily be explained — The matrix to be constructed has more degrees of freedom than the prescribed data can characterize. Generally speaking, the inverse problem is ill-posed. In practice, additional constraints, such as that of being a correlation matrix [10, 19] or having a few extra fixed entries, may be needed to further narrow down the reconstruction. How to construct a matrix satisfying both the sufficient conditions and other structural constraints remains to be an interesting open question.

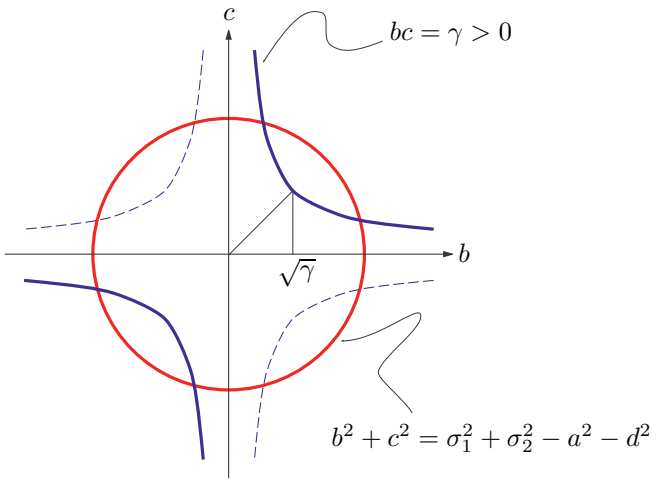


FIGURE 2.1. *Existence of $A \in \mathbb{R}^{2\times 2}$ satisfying the MWHST condition.*

Considered in this paper is the problem of matrix reconstruction subject to prescribed diagonal entries, eigenvalues, and singular values concurrently. We shall refer to this structure as the Mirsky-Weyl-Horn-Sing-Thompson (MWHST) condition. Referring to Figure 1.1, it is natural to ask whether a matrix exists to meet all three prescribed data at the same time. Obviously, any two of these three prescribed data sets must satisfy the necessary conditions required in the corresponding theorem outlined above. While a matrix satisfying two sets of prescribed data does exist and can be constructed numerically, it is not clear whether the third set of prescribed data can be met simultaneously. So far as we know, no theory has been developed in this regard. Our contribution in this paper is that we set forth a theoretical proof of the existence and a numerical algorithm for the reconstruction.

2. Existence. We first address the question of the existence for a matrix satisfying the MWHST condition. The original proofs of both Theorems 1.1 and 1.2 employ the induction principle that requires a partition of the problem into two subproblems of smaller sizes. The criteria used for the partition are different, so the resulting subproblems are of different nature. There does not seem to have any common ground allowing us to mend the proofs together. Instead, we offer a proof of existence as follows.

We begin with the 2×2 case which seems simple but is already causing an issue. We then employ stability argument to show that the existence is ensured when $n > 2$.

2.1. The 2×2 case. Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

denote the 2×2 real matrix to be constructed. The MWHST condition requires that

$$\lambda_1 + \lambda_2 = d_1 + d_2; \qquad \text{(Mirsky)} \tag{2.1}$$

$$\begin{cases} |\lambda_1| \geq |\lambda_2|, \\ \sigma_1 \geq \sigma_2, \\ |\lambda_1| \leq \sigma_1, \\ |\lambda_1||\lambda_2| = \sigma_1\sigma_2; \end{cases} \qquad \text{(Weyl} - \text{Horn)} \tag{2.2}$$

$$\begin{cases} |d_1| \geq |d_2|, \\ |d_1| + |d_2| \leq \sigma_1 + \sigma_2, \\ |d_1| - |d_2| \leq \sigma_1 - \sigma_2. \end{cases} \qquad \text{(Sing} - \text{Thompson)} \tag{2.3}$$

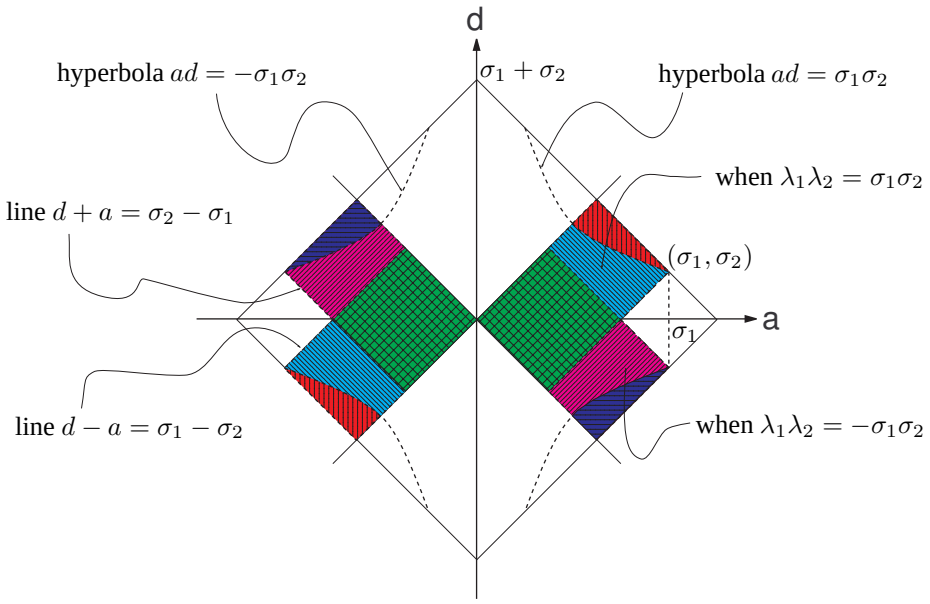


FIGURE 2.2. *Domain of feasible diagonal entries* (a, d)*, given* $\{\sigma_1, \sigma_2\}$ *and* $\{\lambda_1, \lambda_2\}$.

The Frobenius norm of A necessarily implies the equality

$$a^2 + b^2 + c^2 + d^2 = \sigma_1^2 + \sigma_2^2.$$

Assuming that the main diagonal entries are already fixed, our goal is to determine the off-diagonal entries b and c satisfying the system

$$\begin{cases} bc &= ad - \lambda_1\lambda_2, \\ b^2 + c^2 &= \sigma_1^2 + \sigma_2^2 - a^2 - d^2. \end{cases} \tag{2.4}$$

The existence of a matrix A satisfying the MWHST condition therefore boils down to finding the intersection of a hyperbola and a circle as is indicated in Figure 2.1.

Obviously, the system (2.4) is solvable for the off-diagonal entries b and c only if the vertex of the hyperbola lies within the disk, that is, when

$$2|ad - \lambda_1\lambda_2| \le \sigma_1^2 + \sigma_2^2 - a^2 - d^2. \tag{2.5}$$

Given singular values $\boldsymbol{\sigma}$ and eigenvalues $\boldsymbol{\lambda}$, we plot in Figure 2.2 the region of $(a, d) \in \mathbb{R}^2$ for which the inequality (2.5) holds. Since a significant amount of information is contained in the drawing, we briefly explain its interpretation.

1. First, the shape of "kissing fish" in Figure 2.2 represents the feasible region of the diagonal (a, d), given $\boldsymbol{\sigma}$, in order to satisfy the Sing-Thompson condition (2.3) alone. See also [6] for details.
2. Taking into account the Weyl-Horn condition (2.2), we have two mutually exclusive cases — either $\lambda_1\lambda_2 = \sigma_1\sigma_2$ or $\lambda_1\lambda_2 = -\sigma_1\sigma_2$.
3. When eigenvalues λ_1 and λ_2 are either complex conjugate $\lambda_1 = \overline{\lambda}_2$ or real valued with one sign,
   (a) The only feasible diagonal entries must be further restricted to the union of the red, cyan, and green regions in Figure 2.2.
   (b) When $ad \ge \sigma_1\sigma_2$, the (a, d) must come from the red region.

4. When eigenvalues λ_1 and λ_2 are of opposite sign,
   (a) The only feasible diagonal entries must be further restricted to the union of the blue, purple, and green regions in Figure 2.2.
   (b) When $ad \leq -\sigma_1\sigma_2$, the (a, d) must come from the blue region.
5. The green region is when both $|a+d| \leq \sigma_1 - \sigma_2$ and $|a-d| \leq \sigma_1 - \sigma_2$. When (a, d) is in this region, any $\boldsymbol{\lambda}$ satisfying the Weyl-Horn condition (2.2) will guarantee the existence of a matrix A.

The above elaboration on the 2×2 case is significant in that it evidences that merely satisfying the MWHST condition is not sufficient to guarantee the existence of a 2×2 matrix with prescribed diagonal entries, eigenvalues, and singular values. We argue that additional constraints, such as the location of (a, d), depending on $\lambda_1\lambda_2 = \pm\sigma_1\sigma_2$, must be imposed in order to construct a 2×2 matrix satisfying the MWHST condition simultaneously. This result is of interest in itself.

2.2. General case. What is even more astonishing is that the case of 2×2 discussed above is only an exception. In the following, we shall argue that the reconstruction for the case $n > 2$ amounts to an underdetermined system which almost surely guarantees the existence of multiple solutions.

A rough count of the dimensionality might insinuate the reason, although in reality the MWHST condition gives rise to a bunch of inequalities, making the dimensionality analysis not so straightforward. For the case $n = 2$, we need to determine the two off-diagonal entries b and c which are required to meet the conditions of having two eigenvalues and two singular values. Using the equalities in the Mirsky condition (2.1) and the Weyl-Horn condition (2.2), we can reduce the number of eigenvalue condition and singular value condition to one and one, respectively. Thus, the reconstruction problem amounts to solving two equations in two unknowns, such as that of (2.4). To ensure that this square problem has a solution, we find out in the preceding section that some additional constraints are required. For the case $n > 2$, it is impossible to use the geometric argument directly. Instead, we shall resort to parametric representation of the underlying matrix and do the counting. Details along this line are given in Section 3.

In this section, we are to prove the existence by an entirely different strategy. We begin with the assumption that a real-valued matrix $n \times n$ matrix A satisfying Weyl-Horn condition is already in hand. The existence of such a matrix with prescribed singular values $\boldsymbol{\sigma}$ and eigenvalues $\boldsymbol{\lambda}$ is guaranteed in theory and is obtainable numerically. Referring to the diagram in Figure 1.1, we assume that the β side of the triangular relationship is already satisfied and it remains to establish the α side in the same diagram. If both the α side and the β side of the diagram are kept, then the γ side follows automatically. In other words, given the desirable main diagonal elements $\mathbf{d}$ satisfying both the Mirsky condition and the Sing-Thompson condition, our goal now is to somehow transform the matrix A so that the resulting diagonal elements are those of $\mathbf{d}$.

It is important to understand what kinds of transformations are allowed. Since the given A is not necessarily symmetric, to maintain the eigenvalues we have to employ similarity transformations. Likewise, to preserve the singular values we must perform orthogonal equivalence transformations. To keep both eigenvalues and singular values, the only option is to apply orthogonal similarity transformations to the matrix A.

Let $\mathcal{O}(n) \subset \mathbb{R}^{n\times n}$ denote the group of $n \times n$ orthogonal matrices. Our idea of driving the diagonal of $Q^\top AQ$ to that of the specified diagonal $\mathbf{d}$ can be formulated as the minimization problem

$$\min_{Q\in\mathcal{O}(n)} F(Q) := \frac{1}{2}\|\mathsf{diag}(Q^\top AQ) - \mathsf{diag}(\mathbf{d})\|_F^2, \tag{2.6}$$

where $\|\cdot\|_F$ stands for the Frobenius matrix norm. Since the matrix A is real, $\mathsf{diag}(Q^\top AQ) = \mathsf{diag}(Q^\top A^\top Q)$. Therefore,

$$\mathsf{diag}(Q^\top AQ)) = \mathsf{diag}(Q^\top \frac{A + A^\top}{2} Q). \tag{2.7}$$

Define the matrix

$$S := \frac{A + A^\top}{2}. \tag{2.8}$$

It is more convenient to work on the (symmetrized) optimization problem

$$\min_{Q\in\mathcal{O}(n)} F(Q) := \frac{1}{2}\|\mathsf{diag}(Q^\top SQ) - \mathsf{diag}(\mathbf{d})\|_F^2. \tag{2.9}$$

The optimizer of problem (2.6) is the same as that of problem (2.9), and vice versa. Denote

$$\eta(Q) := \mathsf{diag}(Q^\top SQ) - \mathsf{diag}(\mathbf{d}). \tag{2.10}$$

Ideally, we are interested in finding an orthogonal matrix $Q \in \mathcal{O}(n)$ such that $\eta(Q) = 0$. When this happens, then the diagonal entries, the eigenvalues, and the singular values of the matrix $Q^\top AQ$ satisfy the MWHST condition simultaneously.

To solve problem (2.9), we apply the conventional optimization techniques by directly working on the matrices. A useful resource for the matrix calculus maneuvers used below is our earlier paper [8] where theoretical justifications on how the projected gradient and the projected Hessian can be computed without resorting to the Lagrange multiplier theory are carefully documented. Rewriting the objective function $F(Q)$ in terms of the Frobenius inner product for matrices,

$$\langle A, B\rangle := \sum_{i,j} a_{ij}b_{ij},$$

we find that the Fréchet derivative of F at Q acting on an arbitrary matrix $H \in \mathbb{R}^{n\times n}$ is given by

$$\begin{aligned} F'(Q).H &= \langle \eta(Q), \eta'(Q).H\rangle = 2\left\langle \eta(Q), \mathsf{diag}(Q^\top SH)\right\rangle \\ &= 2\left\langle \eta(Q), Q^\top SH\right\rangle = 2\left\langle SQ\eta(Q), H\right\rangle. \end{aligned} \tag{2.11}$$

In (2.11), the second equality follows from the linearity of the operator diag; the second equality results from the fact $\eta(Q)$ is a diagonal matrix; and the third equality is due to the adjoint property. By the Riesz representation theorem, the gradient ∇F at Q can be represented as

$$\nabla F(Q) = 2SQ\eta(Q). \tag{2.12}$$

Because of the constraint $Q \in \mathcal{O}(n)$, the projected gradient $g(Q)$ of $\nabla F(Q)$ onto the tangent space $\mathcal{T}_Q\mathcal{O}(n)$ of $\mathcal{O}(n)$ is important. Taking advantage of the Lie structure of $\mathcal{O}(n)$, it turns out that its tangent space at Q can be identified as either the left translation or the right translation of the subspace of skew-symmetric matrices, that is,

$$\mathcal{T}_Q\mathcal{O}(n) = \{QK \mid K \text{ is skew-symmetric}\} = \{KQ \mid K \text{ is skew-symmetric}\}. \tag{2.13}$$

The following lemma, proved in [8], serves as a recipe to compute the projected gradient onto the tangent space of $\mathcal{O}(n)$.

LEMMA 2.1. *Let $Q \in \mathcal{O}(n)$ be a fixed orthogonal matrix. The projection of any given matrix $X \in \mathbb{R}^{n\times n}$ onto the tangent space $\mathcal{T}_Q\mathcal{O}(n)$ is given by*

$$\mathbb{P}_{\mathcal{T}_Q\mathcal{O}(n)}(X) = \frac{1}{2}Q\left\{Q^\top X - X^\top Q\right\} = \frac{1}{2}\left\{XQ^\top - QX^\top\right\}Q. \tag{2.14}$$

Taking the left translation, the projected gradient can be calculated explicitly as

$$\begin{aligned} g(Q) &= \frac{1}{2}Q\left(Q^\top\nabla F(Q) - \nabla F(Q)^\top Q\right) \\ &= Q\left[Q^\top SQ, \eta(Q)\right], \end{aligned} \tag{2.15}$$

where

$$[A, B] = AB - BA$$

denotes the Lie bracket. If we define the dynamical system

$$\dot{Q} = -g(Q), \quad Q((0) = I, \tag{2.16}$$

then the solution flow $Q(t)$ to (2.16) moves in the steepest flow on $\mathcal{O}(n)$ for the objective function $F(Q)$. We want to show that the asymptotically stable equilibrium of this flow provides the orthogonal matrix Q at which $\eta(Q) = 0$ and establishes the existence of a matrix satisfying the MWHST condition.

The problem (2.9) may have many stationary points. To fully analyze the local behavior of its stationary points, we need to know the action of the projected Hessian of $F(Q)$ over the tangent space at the stationary point. For general optimization, computing the projected Hessian is a desirable but difficult task. But for problem (2.9), we can compute the projected Hessian by the technique developed in [8]. First, we extend the function $g(Q)$ for $Q \in \mathcal{O}(n)$ to the function $G : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ formally via

$$G(Z) := Z \left[Z^\top SZ, \eta(Z)\right]. \tag{2.17}$$

Trivially, the Fréchet derivative of G at $Z \in \mathcal{O}(n)$ acting on $H \in \mathbb{R}^{n\times n}$ is given by

$$G'(Z).H = H \left[Z^\top SZ, \eta(Z)\right] + Z \left[H^\top SZ + Z^\top SH, \eta(Z)\right] + Z \left[Z^\top SZ, \mathsf{diag}(H^\top SZ + Z^\top SH)\right].$$

We are interested in the case when $Q \in \mathcal{O}(n)$ is a stationary point and $H \in \mathcal{T}_Q\mathcal{O}(n)$. Because Q is a stationary point of (2.9) if and only if $\left[Q^\top SQ, \eta(Q)\right] = 0$ by (2.15), also because $H \in \mathcal{T}_Q\mathcal{O}(n)$ if and only if H is of the form $H = QK$ for some skew-symmetric matrix $K \in \mathbb{R}^{n\times n}$, the projected Hessian $G'(Q)$ acting on QK is given by

$$\begin{aligned}
\langle QK, G'(Q).QK\rangle &= \left\langle QK, Q\left[\left[Q^\top SQ, K\right], \eta(Q)\right] + Q\left[Q^\top SQ, \operatorname{diag}\left[Q^\top SQ, K\right]\right]\right\rangle \\
&= \left\langle K, \left[\left[Q^\top SQ, K\right], \eta(Q)\right] + \left[Q^\top SQ, \operatorname{diag}\left[Q^\top SQ, K\right]\right]\right\rangle \\
&= \left\langle \left[Q^\top SQ, K\right], [K, \eta(Q)] + \operatorname{diag}\left[Q^\top SQ, K\right]\right\rangle.
\end{aligned} \tag{2.18}$$

The second-order necessary optimality condition at a stationary point $Q \in \mathcal{O}(n)$ is now characterized by

$$\langle QK, G'(Q).QK\rangle \geq 0 \quad \text{for all skew-symmetric matrices } K. \tag{2.19}$$

It is known that the condition becomes sufficient if the strict inequality holds.

We now apply the above theory to establish the existence of a matrix satisfying the MWHST condition. Recall that our goal is accomplished if we can find an orthogonal matrix Q such that $\eta(Q) = 0$. We first argue that any stationary point Q at which $\eta(Q) = 0$ satisfies the necessary condition of being a minimal of the optimization problem (2.9).

THEOREM 2.2. *Let $Q \in \mathcal{O}(n)$ be a stationary point of (2.9). If $\eta(Q) = 0$, then for all skew-symmetric matrices K, the projected Hessian $G'(Q)$ is positive semi-definite.*

Proof. By (2.18), we see that if $\eta(Q) = 0$, then

$$\begin{aligned}
\langle QK, G'(Q).QK\rangle &= \left\langle \left[Q^\top SQ, K\right], \operatorname{diag}\left[Q^\top SQ, K\right]\right\rangle = \left\langle \operatorname{diag}\left[Q^\top SQ, K\right], \operatorname{diag}\left[Q^\top SQ, K\right]\right\rangle \\
&= \|\operatorname{diag}\left[Q^\top SQ, K\right]\|_F^2 \geq 0
\end{aligned}$$

for all skew-symmetric matrices K. □

Though the sufficient condition is much harder to verify, keep in mind that we have a descent flow $Q(t)$ to follow. We now argue that if $\eta(Q) \neq 0$ for a stationary point Q, then there is a direction along which the value

of the objective function F will be further downgraded. As such, our flow $Q(t)$ will continue to descend and bypass[1] this stationary point until a local minimizer is found.

For simplicity, we assume the generic situation that all eigenvalues of S are distinct. The analysis for the case of equal eigenvalues is slightly involved, but our numerical experiences seem to suggest that the asymptotic behavior should be similar.

THEOREM 2.3. *Assume that the matrix S has distinct eigenvalues. Let $Q \in \mathcal{O}(n)$ be a stationary point. If $\eta(Q) \neq 0$, then there exists a skew-symmetric matrix K such that $\langle QK, G'(Q)QK\rangle < 0$.*

Proof. Since S is symmetric, let $S = U^\top \Pi U$ denote its spectral decomposition where $U^\top$ is the orthogonal matrix of eigenvectors and $\Pi = \mathrm{diag}\{\pi_1, \ldots, \pi_n\}$ is the diagonal matrix of eigenvalues. If $\eta(Q) \neq 0$, then we may assume, after performing a similarity transformation by permutation matrices, if necessary, that $\eta(Q)$ is of the form

$$\eta(Q) = \mathrm{diag}\{\eta_1 I_{n_1}, \cdots, \eta_k I_{n_k}\},$$

where I_{n_i} is the $n_i \times n_i$ identical matrix for $i = 1, \cdots, k$, and $\eta_1 > \cdots > \eta_k$. Because Q is a stationary point, by (2.15) we have $\left[Q^\top SQ, \eta(Q)\right] = 0$. It follows that the matrix $Y := Q^\top SQ$ is of block diagonal form

$$Y = \mathrm{diag}\{Y_{11}, \cdots, Y_{kk}\}, \tag{2.20}$$

where Y_{ii} is a $n_i \times n_i$ symmetric matrix for $i = 1, \cdots, k$. Define the matrix $V := (UQ)\eta(Q)(UQ)^\top$. Then $[\Pi, V] = 0$. Since the diagonal matrix Π has distinct diagonal entries, V must also be a diagonal matrix. Write $V = \mathrm{diag}\{v_1, \cdots, v_n\}$. Observe that V is permutation-similar to $\eta(Q)$. Since $(UQ)^\top$ is the orthogonal matrix of eigenvectors of Y, the matrix UQ must have the same block structure as that of Y.

Let $K \in \mathbb{R}^{n\times n}$ be a skew-symmetric matrix partitioned in block form $K = [K_{ij}]$ whose block sizes are conformal to those of Y as in (2.20) and $K_{ii} = 0$ for $i = 1, \cdots, k$. Clearly, $\mathrm{diag}\left[Q^\top SQ, K\right] = 0$. With respect to this matrix K, the projected Hessian (2.18) becomes

$$\begin{aligned}\langle QK, G'(Q).QK\rangle &= \left\langle \left[Q^\top SQ, K\right], [K, \eta(Q)]\right\rangle = -\left\langle V\tilde{K} - \tilde{K}V, \Pi\tilde{K} - \tilde{K}\Pi\right\rangle \\ &= -2\sum_{i<j}(\pi_i - \pi_j)(v_i - v_j)\tilde{k}_{ij}^2, \end{aligned} \tag{2.21}$$

where $\tilde{K} = \left[\tilde{k}_{ij}\right] := (UQ)K(UQ)^\top$ remains to be skew-symmetric since UQ is orthogonal. With both sets $\{\pi_1, \ldots, \pi_n\}$ and $\{v_1, \ldots, v_n\}$ fixed, it is obvious from (2.21) that we may choose appropriate values of $\tilde{k}_{ij}$ such that $\langle QK, G'(Q).QK\rangle < 0$. ∎

The implication of Theorem 2.3 is that if Q is a stationary point with $\eta(Q) \neq 0$, then at this point there exists at least one direction along which $F(Q)$ can be further decreased. This point is an unstable equilibrium for the gradient dynamics. The flow must continue until an isolated limit point at which $\eta(Q) = 0$ is found. Based on this understanding from Theorems 2.2 and 2.3, we conclude that the convergence of $Q(t)$ to an asymptotically stable equilibrium point Q at which $\eta(Q) = 0$ is guaranteed. When this limit point is achieved, the corresponding matrix $Q^\top AQ$ will maintain the prescribed diagonal entries, eigenvalues, and singular values. The existence of a matrix satisfying the MWHST condition is hereby established.

3. Variational formulation. In this section, we outline a computational framework that can be implemented to carry out the task of matrix reconstruction. Let $\Sigma := \mathsf{diag}(\boldsymbol{\sigma})$ denote the diagonal matrix of singular values. Since the given eigenvalues are closed under complex conjugation, we may convert $\mathsf{diag}(\boldsymbol{\lambda})$ to a block diagonal matrix $\Lambda \in \mathbb{R}^{n\times n}$ whose block sizes are at most of 2×2 when complex conjugate eigenvalues are present. For simplicity, we shall consider only the case when the prescribed eigenvalues are semi-simple.

[1] Unless $Q(t)$ happens to stay on a heteroclinic orbit, which is numerically unlikely due to ubiquitous floating-point arithmetic errors.

Referring to Figure 1.1, define a triplet of functions

$$\begin{cases} \alpha(P) &:= \mathsf{diag}(P\Lambda P^{-1}) - \mathsf{diag}(\mathbf{d}), \\ \beta(P,U,V) &:= P\Lambda P^{-1} - U\Sigma V^\top, \\ \gamma(U,V) &:= \mathsf{diag}(U\Sigma V^\top) - \mathsf{diag}(\mathbf{d}), \end{cases} \tag{3.1}$$

where, without causing ambiguity, we use the same notation $\mathsf{diag}(A)$ and $\mathsf{diag}(\mathbf{v})$ to indicate a diagonal matrix with elements either from the main diagonal of the matrix A or from entries of the vector $\mathbf{v}$. We propose to cast the reconstruction problem as the minimization of the objective function defined by

$$f(P,U,V) := \frac{\omega_1}{2}\|\alpha(P)\|_F^2 + \frac{\omega_2}{2}\|\beta(P,U,V)\|_F^2 + \frac{\omega_3}{2}\|\gamma(U,V)\|_F^2, \tag{3.2}$$

subject to the constraints that $U, V \in \mathbb{R}^{n\times n}$ are orthogonal matrices and $P \in \mathbb{R}^{n\times n}$ is an invertible matrix with $\|P\|_F = 1$. The idea is to parameterize the target matrix in two different ways,

$$\begin{aligned} X &:= U\Sigma V^\top, \\ Y &:= P\Lambda P^{-1}, \end{aligned}$$

preserving singular values and eigenvalues, respectively; and ideally we want the objective value to become zero. The nonnegative weights $\boldsymbol{\omega} = [\omega_1, \omega_2, \omega_3]$ in (3.2) can be used to control the minimization for various purposes. Table 3.1 summarizes a few selections. More importantly, the weights can also be modified adaptively during the integration process as control parameters, recasting the formulation into an interesting control problem. We shall describe some of the control strategies in Section 4.2.

$\boldsymbol{\omega}$	Purpose	Theorem
$[0,0,1]$	singular values + diagonal	Sing-Thompson
$[0,1,0]$	singular values + eigenvalues	Weyl-Horn
$[1,0,0]$	eigenvalues + diagonal	Mirsky
$[1,0,0]$	eigenvalues + diagonal	Schur-Horn, if P is orthogonal
$[0,1,1]$	singular values + eigenvalues + diagonal	
$[1,1,0]$	singular values + eigenvalues + diagonal	
$[1,0,1]$	not interesting	
$[1,1,1]$	singular values + eigenvalues + diagonal	

TABLE 3.1
A few weighting schemes

The calculation being tedious but straightforward by using matrix calculus, we first claim that the gradient of f given by the following expressions.

LEMMA 3.1. *The gradient ∇f can be characterized segmentally as*

$$\frac{\partial f}{\partial P} = -\left[Y^\top, \omega_1\alpha(P) + \omega_2\beta(P,U,V)\right] P^{-\top}, \tag{3.3}$$

$$\frac{\partial f}{\partial U} = -\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right) V\Sigma^\top, \tag{3.4}$$

$$\frac{\partial f}{\partial V} = -\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right)^\top U\Sigma. \tag{3.5}$$

We then take into account that the variable P is restricted to the sphere $\mathcal{S}(n) \subset \mathbb{R}^{n\times n}$ of unit Frobenius norm, while the other two variables U and V are restricted to the group $\mathcal{O}(n) \subset \mathbb{R}^{n\times n}$ of orthogonal matrices.

Only the projected gradient of f onto the tangent space of the feasible set can determine the role of the variables P, U and V in the optimization.

Trivially, the projection of any given matrix $A \in \mathbb{R}^{n\times n}$ onto the tangent space $\mathcal{T}_Z\mathcal{S}(n)$ of $\mathcal{S}(n)$ at the point $Z \in \mathcal{S}(n)$ is given by

$$\mathbb{P}_{\mathcal{T}_Z\mathcal{S}(n)}(A) = A - \langle A, Z\rangle Z. \tag{3.6}$$

When applied to our $\frac{\partial f}{\partial P}$, we find that it is already in the tangent space $\mathcal{T}_P\mathcal{S}(n)$ because

$$\left\langle \frac{\partial f}{\partial P}, P \right\rangle = -\mathrm{trace}\left(\left[Y^\top, \omega_1\alpha(P) + \omega_2\beta(P,U,V)\right]\right)$$

and the trace of any communicator vanishes. Using the right translation formula in Lemma 2.1, we also obtain the projected gradients for (3.4) and (3.5) as

$$\begin{aligned}
\mathbb{P}_{\mathcal{T}_U\mathcal{O}(n)}\left(\frac{\partial f}{\partial U}\right) &= \frac{-1}{2}\left\{\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right)X^\top - X\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right)^\top\right\}U,\\
\mathbb{P}_{\mathcal{T}_V\mathcal{O}(n)}\left(\frac{\partial f}{\partial V}\right) &= \frac{-1}{2}\left\{\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right)^\top X - X^\top\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right)\right\}V,
\end{aligned}$$

respectively.

All together, we can formulate a steepest descent flow $(P(t), U(t), V(t))$ on the manifold $\mathcal{S}(n)\times\mathcal{O}(n)\times\mathcal{O}(n)$ via the dynamical system

$$\left\{\begin{array}{rcl}
\dot{P} &=& \left[Y^\top, \omega_1\alpha(P) + \omega_2\beta(P,U,V)\right]P^{-\top},\\
\dot{U} &=& \frac{1}{2}\left\{\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right)X^\top - X\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right)^\top\right\}U,\\
\dot{V} &=& \frac{1}{2}\left\{\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right)^\top X - X^\top\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right)\right\}V.
\end{array}\right. \tag{3.7}$$

The solution flow moves to reduce the objective value $f(P,U,V)$ which might have multiple stationary points. In the next section, we propose a mechanism of "steeling control" with the hope of driving this flow to an equilibrium where $f(P,U,V) = 0$.

4. Convergence and control. As a descent flow, the dynamics will stop only in two scenarios — either $P(t)^{-1}$ fails to exist in finite time or the objective value $f(P(t),U(t),V(t))$ converges to a constant. The first scenario can easily be detected and, when it happens, a restart might fix the problem. Our main concern is about the second scenario. In this section, we characterize a few properties of the limit points and propose a control strategy to effectuate the convergence behavior.

4.1. Properties of equilibria. First of all, we rule out the possibility of having limit cycles.

LEMMA 4.1. *Suppose that $P(t)^{-1}$ does not present a problem. Then the flow $(P(t),U(t),V(t))$ starting from any initial point converges to an isolated limit point.*

Proof. Clearly the vector field defined in (3.7) is analytic in the variables P, U, and V wherever the field is defined. Being an analytic gradient flow, the isolation of limit points is guaranteed by using the Łojasiewicz inequalities [4, 18]. ∎

Secondly, any stationary point of (3.2) must satisfy the system of homogeneous equations

$$\left\{\begin{array}{rcl}
Y^\top\left(\omega_1\alpha(P) + \omega_2\beta(P,U,V)\right) - \left(\omega_1\alpha(P) + \omega_2\beta(P,U,V)\right)Y^\top &=& 0,\\
\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right)X^\top - X\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right)^\top &=& 0,\\
\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right)^\top X - X^\top\left(\omega_2\beta(P,U,V) - \omega_3\gamma(U,V)\right) &=& 0.
\end{array}\right. \tag{4.1}$$

The system (4.1) is nonlinear in (P, U, V) as $X = U\Sigma V^\top$ and $Y = P\Lambda P^{-1}$ themselves are functions of (P, U, V). Still, we could gain some insight into its solution by recasting these equations in some parametric form. To characterize a parametric expression for the solution to (4.1), we shall make use of the following result [1, Corollary 2].

LEMMA 4.2. *Suppose that $A \in \mathbb{R}^{n\times n}$ is nonsingular. Then the general solution of the system*

$$A^\top Z - Z^\top A = B \tag{4.2}$$

is given by

$$Z = \Phi^\top \left(W + \frac{1}{2}\Psi^\top B \Psi \right) \Psi^{-1} \tag{4.3}$$

where W is an arbitrary symmetric matrix and $\Phi A \Psi = I$.

We now illustrate the idea of parametrization by assuming that the prescribed eigenvalues $\boldsymbol{\lambda}$ are simple and the prescribed singular values $\boldsymbol{\sigma}$ are all positive. Analysis for the cases such as multiple eigenvalues or zero singular values are more involved and will not be discussed in this paper.

LEMMA 4.3. *There exist a block diagonal matrix D having the same structure as Λ and symmetric matrices S, T such that*

$$\left\{ \begin{array}{rcl} \omega_1\alpha(P) + \omega_2\beta(P,U,V) & = & P^{-\top} D P^\top, \\ \omega_2\beta(P,U,V) - \omega_3\gamma(U,V) & = & SX, \\ \omega_2\beta(P,U,V) - \omega_3\gamma(U,V) & = & XT. \end{array} \right. \tag{4.4}$$

Proof. The first equation in (4.1) is similar to the classical homogeneous Sylvester equation, except that $Y^\top$ is also part of the unknown. By the fact that Λ has distinct eigenvalues, the matrix $\omega_1\alpha(P)+\omega_2\beta(P,U,V)$ commutes with $Y^\top = (P\Lambda P^{-1})^\top$ if and only if it is of the first asserted form in (4.3). Indeed, it is also necessary that the matrix $(\omega_1\alpha(P) + \omega_2\beta(P,U,V))^\top$, if not identically zero, should have the same eigenvectors as those of Y.

To see the second parametric representation, we apply Lemma 4.2 to the last equation in (4.1) by identifying $Z = \omega_2\beta(P,U,V) - \omega_3\gamma(U,V)$, $A = X$ and $B = 0$. Note that the general solution (4.3) may be rewritten as $Z = \Phi^\top W \Phi A$. So $S = \Phi^\top W \Phi$ is symmetric. The third parametric representation can be obtained by a similar argument. □

The term "almost surely" has been used in the literature to characterize an event that happens with probability one. When this event depends on some parameters, we say that these parameters are generic when those for which the event fails to hold form a set of measure zero. Since the representation (4.3) involves some parameters with structures, we are able to count the dimensionality of solution set precisely.

LEMMA 4.4. *For generically prescribed data $\boldsymbol{\lambda}$, $\boldsymbol{\sigma}$, and $\mathbf{d}$, the solutions to the nonlinear algebraic system (4.1) form a manifold of dimension n.*

Proof. It is known that orthogonal matrices can be characterized by $\frac{n(n-1)}{2}$ free parameters. The variables (P, U, V) therefore constitute a total of $n^2 + \frac{n(n-1)}{2} + \frac{n(n-1)}{2} = 2n^2 - n$ unknowns. On the other hand, the freedom of D means that first equation in the system (4.3) gives rise to only $n^2 - n$ independent equations, whereas the freedom of the symmetric matrices S and T in the second and the third (4.3) implies that each gives rise to $\frac{n(n-1)}{2}$ independent equations. Thus, the system (4.1) contains precisely $2n^2 - 2n$ independent nonlinear equations. The system is under-determined and the assertion follows from the Thom transversality theorem. □

The conclusion in Lemma 4.4 is significant. It suggests that, if $P(t)^{-1}$ is never a problem, then we might be able to prescribe at least n additional conditions so as to reduce the solution set, even to the extent of dimension zero, that is, a discrete set. We have mentioned in Section 1 that one issue associated with the inverse problems is that we can hardly reconstruct a specific matrix, even if the prescribed data are feasible. Now we know that

there are n degrees of freedom for the variables (P, U, V) in the reconstruction. By imposing extra conditions we might have a chance to pinpoint some more specifically structured matrices.

To demonstrate the possibility of introducing extra constraints, consider the scenario of imposing prescribed entries to the reconstructed matrix at locations in addition to the main diagonal. The setup is straightforward. Let $\mathcal{L} := \{(i_t, j_t\}_{t=1}^{\ell}$ denote a subset of double indices including all diagonal positions. Let $\mathbf{d} \in \mathbb{R}^{\ell}$ be a given real vector. Suppose that we wish to reconstruct a matrix A which not only satisfies the MWHST conditions but also has the prescribed entry values $a_{i_t j_t} = d_t$ for all $t = 1, \ldots, \ell$. Similar to the operator diag that has been used in (3.1) to identify the diagonal entries, let $\mathscr{P}$ denote the operator that picks out all entries with indices in $\mathcal{L}$. Then all we need to do is to replace every reference to diag by $\mathscr{P}$, particularly in the definitions of $\alpha(P)$ and $\beta(U, V)$, and our gradient flow is ready to go.

A subtle point need be made. At first glance, merely replacing diag by $\mathscr{P}$ does not alter the appearance of the nonlinear system (4.1) and, thus, seems not to have any effect on Lemma 4.4. The reason is that the added condition $\mathscr{P}$ as we have just described above are not applied to the variables (P, U, V) directly. It does affect, however, the objective function and ultimately alter the dynamics of the flow and associated limit points. Some numerical examples will be given in Section 5 . Given $\boldsymbol{\lambda}$, $\boldsymbol{\sigma}$, and $\mathbf{d}$, questions such as how extra conditions should be imposed, how many, where, and, more importantly, whether there are some inherent relationships that must be satisfied, such as those required in the Sing-Thompson theorem between $\boldsymbol{\sigma}$ and $\mathbf{d}$, remain to be wide open problems. Some related discussions dealing with only the inverse problems of eigenvalues and prescribed entries can be found in [14] and our book [9, Section 4.7].

4.2. Control. The objective function being non-convex, it is possible that the flow $(P(t), U(t), V(t))$ converges to a stationary point at which the objective value is not zero. This is not serving our purpose of matrix reconstruction. We prefer to see that the gradient flow converges to a point at which $\alpha(P) = 0$, $\beta(P, U, V) = 0$, and $\gamma(U, V) = 0$. The approach of analyzing the projected Hessian information such as that adopted in Section 2.2 for the problem (2.9) does not work because we have numerical evidence that the problem (3.2) does have undesirable local minimizers. Instead, we propose two possible ways to change the course of integration of the gradient flow.

The passive approach is to employ the global optimization technique to roughly estimate a good starting point with the hope that the associated flow will converge to a desirable stationary point. The objective of global optimization is to find the globally best solution in the presence of multiple local optima. Over the years, many strategies for global optimization have been proposed. See, for example, the book [22] for a general discussion. One prevailing scheme that allows general constraints and makes possible straightforward generalizations of existent local algorithms is the notion of adaptive partition, such as the MultiStart method available in MATLAB. To avoid distraction from our main focus of this paper, we shall not elaborate further details of global optimization here. We simply mention that applying the global optimization technique to the problem (3.2) at its full power is not practical because it does not respect the matrix structure well. Instead, we propose using global optimization with low precision only for the purpose of estimating a starting point for our gradient flow.

We are more interested in the active approach that adaptively adjusts the weights $\boldsymbol{\omega}$ to drive the flow to a desirable stationary point. One heuristic strategy we have been employing is to use the norms of $\alpha(P)$, $\beta(P, U, V)$, and $\gamma(U, V)$ periodically or intermittently as the feedback information. The idea is to penalize the term, say, $\beta(P, U, V)$, by increasing its weight ω_2 when its norm is relatively too large when comparing to the other two terms. One weighting scheme of this nature is

$$\boldsymbol{\omega} := \frac{[\|\alpha(P)\|, \|\beta(P, U, V)\|, \|\gamma(U, V)\|]}{\|\alpha(P)\| + \|\beta(P, U, V)\| + \|\gamma(U, V)\|}. \tag{4.5}$$

Such an adaptive strategy is somewhat similar to the spirit of simulated annealing. It is particularly effective at the initial stage of integration as it quickly brings down the objective values and increases the chance of converging to a global minimum. Summarized in Figure 4.1 is the flowchart of our control scheme with the following highlights:

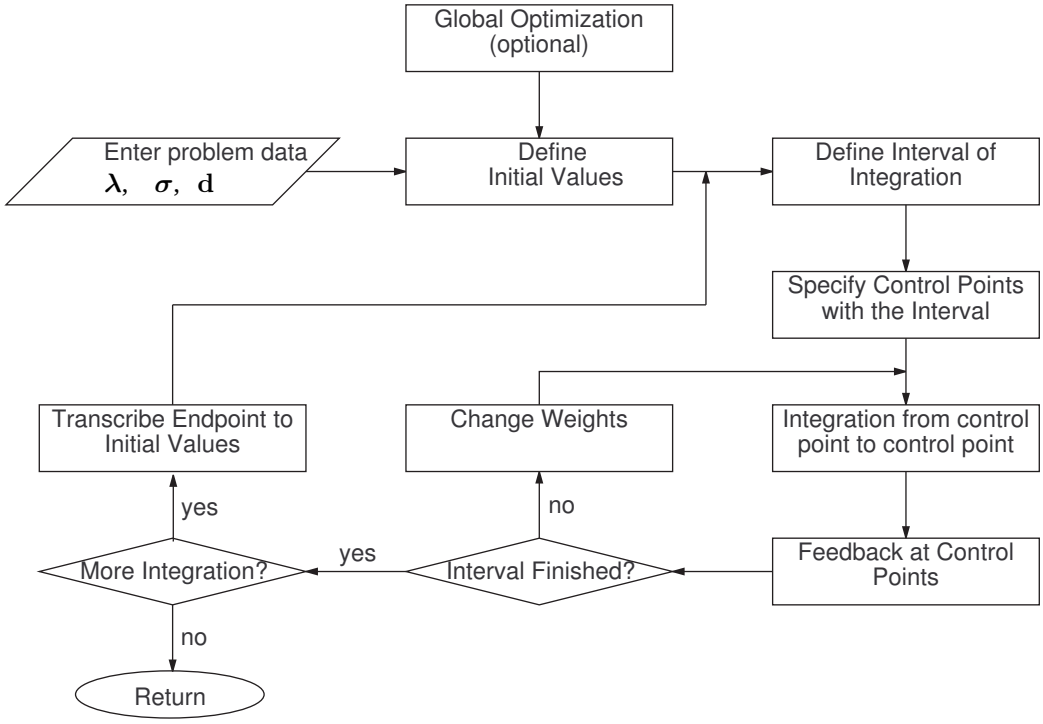


FIGURE 4.1. *Flowchart of objective function control.*

- We integrate the gradient flow from interval to interval;
- Within each interval we assign a few control points at which the norms $\|\alpha(P)\|$, $\|\beta(P,U,V)\|$, and $\|\gamma(U,V)\|$ are gathered as feedback information;
- At the passing of each control point, we update the weights by the scheme (4.5) and continue the integration.

5. Numerical Experiments. We have pointed out earlier that the goal of this study is to construct a matrix with prescribed diagonal entries, eigenvalues, and singular values simultaneously, whereas the given data necessarily satisfying the MWHST condition. We have argued in Section 2.2 that such a matrix does exist. This existence problem is of theoretic interest in its own right. We are also interested in the numerical problem of actually constructing such a matrix by means of the controlled gradient flow approach. In this section, we demonstrate the working of such an algorithm.

First, we comment on the choice of the integrator. To preserve the orthogonality of $(U(t), V(t))$ in the long run, structure preserving techniques from the field of geometric integration should have been used [12, 15]. Also, for gradient flows, the so called pseudo-transient continuation might also be an effective integrator [16]. Since we only want to demonstrate the dynamical behavior of the differential system (3.7) and explore its limit points, we choose for the convenience of quick implementation to use some standard routines such as ode113 or ode15s from MATLAB as our integrator. The local error tolerance is set at $\mathsf{AbsTol} = \mathsf{RelTol} = 10^{-10}$.

Our code, available upon request, is interactive where users can specify the interval of integration per call and decide whether the integration should be continued. For each specified interval of integration, regardless of its length, we choose 1 to 10 uniformly spaced control points at which we update the weights. We stress that too many control points, causing the objective function (3.2) to change too frequently, will degrade the steepest descent property of the flow, whereas too few control points might allow the flow to drift toward and be trapped in the basin of some undesirable local minimizer, making the subsequent control less effective.

To create feasible test data, we randomly generate a matrix A_0 and use its eigenvalues, singular values, and diagonal elements as the target $\boldsymbol{\lambda}$, $\boldsymbol{\sigma}$ and $\mathbf{d}$. The MWHST condition therefore is automatically satisfied. To further confine the reconstruction, we also require that some entries of A_0 be fixed as additional prescribed constraints. In the examples below, the initial values for the gradient flow are also generated randomly. If the initial values are more selectively chosen, say, by estimates from the global optimization, the needed length of integration might be shorter and the control might be easier. For the ease of reporting data, we display only five digits of the resulting matrices.

Example 1. Consider the randomly generated 6×6 matrix

$$A_0 = \begin{bmatrix} 0.7754 & 0.1212 & -0.3174 & 1.1949 & -0.9096 & -0.9090 \\ -0.3998 & -0.7560 & 0.1206 & -1.0767 & 1.2196 & -1.3019 \\ -1.2898 & -0.5780 & -0.6221 & -1.9055 & -0.3125 & 2.2231 \\ -1.4958 & -0.5997 & 0.2040 & -0.3973 & -0.0775 & -0.2939 \\ 0.1929 & -0.1977 & 0.3752 & 1.7697 & 0.1069 & 0.9239 \\ -1.6462 & 1.3532 & 0.9897 & 0.5379 & 0.0005 & -0.0382 \end{bmatrix}.$$

If both the first row and the first column of A_0 are fixed as additional prescribed entries, our algorithm returns the matrix

$$X_1 = Y_1 = \begin{bmatrix} 0.7754 & 0.1212 & -0.3174 & 1.1949 & -0.9096 & -0.9090 \\ -0.3998 & -0.7560 & -2.1078 & 0.0203 & 1.1121 & -0.5526 \\ -1.2898 & -0.0284 & -0.6221 & -0.1919 & 0.9255 & 2.9329 \\ -1.4958 & -1.0197 & 0.4272 & -0.3973 & -0.3038 & -0.7649 \\ 0.1929 & -0.9298 & 1.1917 & 0.7851 & 0.1069 & -0.3944 \\ -1.6462 & -0.8000 & -0.0791 & 0.1115 & -0.1535 & -0.0382 \end{bmatrix}$$

as its limit point. If both the first row and the last row of A_0 are fixed as prescribed entries, then we obtain

$$X_2 = Y_2 = \begin{bmatrix} 0.7754 & 0.1212 & -0.3174 & 1.1949 & -0.9096 & -0.9090 \\ -2.3396 & -0.7560 & -1.1553 & -1.6924 & 0.9209 & 0.0898 \\ -1.1965 & -0.8044 & -0.6221 & 1.7406 & -0.6763 & -0.0808 \\ -0.1879 & -0.1144 & -1.8697 & -0.3973 & 0.7414 & 0.3612 \\ -0.2473 & 0.1778 & 0.4978 & 0.4616 & 0.1069 & 1.1131 \\ -1.6462 & 1.3532 & 0.9897 & 0.5379 & 0.0005 & -0.0382 \end{bmatrix}$$

as the reconstructed matrix. Even if we fix the entire boarder entries of A_0, we are able to construct the matrix

$$X_3 = Y_3 = \begin{bmatrix} 0.7754 & 0.1212 & -0.3174 & 1.1949 & -0.9096 & -0.9090 \\ -0.3998 & -0.7560 & 0.6758 & -0.5000 & 0.4905 & -1.3019 \\ -1.2898 & -0.5982 & -0.6221 & -1.1451 & -0.0252 & 2.2231 \\ -1.4958 & -0.4443 & 0.8822 & -0.3973 & -0.8479 & -0.2939 \\ 0.1929 & -1.1083 & 1.3881 & -1.7252 & 0.1069 & 0.9239 \\ -1.6462 & 1.3532 & 0.9897 & 0.5379 & 0.0005 & -0.0382 \end{bmatrix}$$

with the same $\boldsymbol{\lambda}$, $\boldsymbol{\sigma}$ and $\mathbf{d}$ as A_0. All three experiments show that the extra constraints still cannot uniquely determine a solution. Starting from $t = 0$, we specify our intervals of integration at the intermittence of powers of 2 with one control at the endpoint. As the length of each interval gets protracted, so are the distances between control points at which the weights are changed. The above limit points are found to the precision of approximately 10^{-9} without any difficulty.

Example 2. Consider another randomly generated 6×6 matrix

$$A_0 = \begin{bmatrix} 1.6903 & 0.4434 & 0.5309 & -0.4648 & 2.2449 & -2.2784 \\ -2.1574 & -1.0637 & -0.4337 & -0.1651 & -1.2057 & -0.4199 \\ -0.2099 & -0.3928 & -0.5548 & -2.2547 & -0.4961 & 0.6704 \\ -0.5790 & 0.7162 & 1.8638 & -0.2952 & -0.1529 & 0.3959 \\ -1.4898 & -0.7578 & 1.1751 & -2.1534 & -0.2094 & -0.2145 \\ 0.8561 & 1.6079 & -0.2854 & -0.6930 & -0.4523 & -0.0405 \end{bmatrix}.$$

The matrix

$$X_1 = Y_1 = \begin{bmatrix} 1.6903 & 0.4434 & 0.5309 & -0.4648 & 2.2449 & -2.2784 \\ -0.5220 & -1.0637 & 0.7395 & -1.1429 & 1.1652 & -0.3124 \\ -1.6229 & -0.6284 & -0.5548 & 2.1024 & 0.3180 & 1.0329 \\ -1.1242 & 2.0171 & -0.3040 & -0.2952 & -0.5820 & -2.4021 \\ -0.8558 & -1.5404 & -1.0452 & -0.2008 & -0.2094 & -0.2747 \\ 0.1635 & -0.2898 & -0.0073 & 0.8985 & 0.9796 & -0.0405 \end{bmatrix}$$

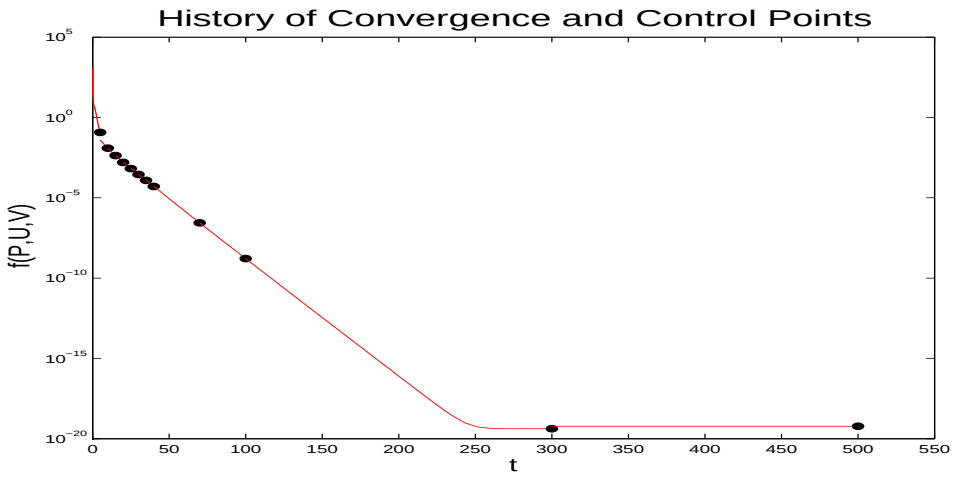


FIGURE 5.1. *History of convergence and control points.*

is found to have the same $\boldsymbol{\lambda}$, $\boldsymbol{\sigma}$, $\mathbf{d}$, and the first row as A_0. For this example, we integrate from $t = 0$ to $t = 10$, and then successively to $t = 20, 30, 40, 100, 500$ with 2 controls per interval. The bullets plotted in Figure 5.1 indicate the control points. A close examination reveals that the graph of the objective function $f(P(t), U(t), V(t))$ is only piecewise continuous. The more frequent controls, i.e., 2 control points over shorter interval, at the initial stage is crucial in preventing the flow $(P(t), U(t), V(t))$ from getting trapped too soon in the domain of attraction by an undesirable equilibrium. Since our code works interactively, at around $t = 40$ we observe that the objective value probably indicates convergence and thus decide to relax the control by using longer intervals. At present, we do not have a definitive rule on how the control should be made. We restart the calculation when one control strategy fails. With a few trials, we can reconstruct the matrix within the expected precision.

6. Conclusion. We study the inverse problem of constructing a matrix with prescribed diagonal entries, eigenvalues, and singular values when these data satisfy the majorization conditions entailed by the Mirsky, Weyl-Horn, and Sing-Thompson theorems simultaneously. Our main contributions are twofold. First, we establish the existence of such a matrix. Second, we propose a controlled gradient flow approach to solve this problem. Both results are new in the field. Also, we show that generically there are n degrees of freedom to impose additional constraints. We demonstrate that our numerical scheme is general enough to handle the situation when some extra off-diagonal entries are fixed.

REFERENCES

[1] H. W. BRADEN, *The equations $A^T X \pm X^T A = B$*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 295–302 (electronic).

[2] E. A. CARLEN AND E. H. LIEB, *Short proofs of theorems of Mirsky and Horn on diagonals and eigenvalues of matrices*, Electron. J. Linear Algebra, 18 (2009), pp. 438–441.

[3] N. N. CHAN AND K. H. LI, *Diagonal elements and eigenvalues of a real symmetric matrix*, J. Math. Anal. Appl., 91 (1983), pp. 562–566. See also *Algorithm 115: A FORTRAN subroutine for finding a real symmetric matrix with prescribed diagonal elements and eigenvalues* in Algorithms Supplement, The Computer Journal (1983) 26(2): 184-186.

[4] R. CHILL, *On the Łojasiewicz-Simon gradient inequality*, J. Funct. Anal., 201 (2003), pp. 572–601.

[5] M. T. CHU, *Constructing a Hermitian matrix from its diagonal entries and eigenvalues*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 207–217.

[6] ———, *On constructing matrices with prescribed singular values and diagonal elements*, Linear Algebra Appl., 288 (1999), pp. 11–22.

[7] ———, *A fast recursive algorithm for constructing matrices with prescribed eigenvalues and singular values*, SIAM J. Numer. Anal., 37 (2000), pp. 1004–1020 (electronic).

[8] M. T. CHU AND K. R. DRIESSEL, *The projected gradient method for least squares matrix approximations with spectral constraints*, SIAM J. Numer. Anal., 27 (1990), pp. 1050–1060.
[9] M. T. CHU AND G. H. GOLUB, *Inverse eigenvalue problems: theory, algorithms, and applications*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2005.
[10] P. I. DAVIES AND N. J. HIGHAM, *Numerically stable generation of correlation matrices and their factors*, BIT, 40 (2000), pp. 640–651.
[11] I. S. DHILLON, R. W. HEATH, JR., M. A. SUSTIK, AND J. A. TROPP, *Generalized finite algorithms for constructing Hermitian matrices with prescribed diagonal and spectrum*, SIAM J. Matrix Anal. Appl., 27 (2005), pp. 61–71 (electronic).
[12] E. HAIRER, C. LUBICH, AND G. WANNER, *Geometric numerical integration*, vol. 31 of Springer Series in Computational Mathematics, Springer, Heidelberg, 2010. Structure-preserving algorithms for ordinary differential equations, Reprint of the second (2006) edition.
[13] A. HORN, *On the eigenvalues of a matrix with prescribed singular values*, Proc. Amer. Math. Soc., 5 (1954), pp. 4–7.
[14] K. D. IKRAMOV AND V. N. CHUGUNOV, *Inverse matrix eigenvalue problems*, J. Math. Sci. (New York), 98 (2000), pp. 51–136. Algebra, 9.
[15] A. ISERLES, H. Z. MUNTHE-KAAS, S. P. NØRSETT, AND A. ZANNA, *Lie-group methods*, in Acta numerica, 2000, vol. 9 of Acta Numer., Cambridge Univ. Press, Cambridge, 2000, pp. 215–365.
[16] C. T. KELLEY, L.-Z. LIAO, L. QI, M. T. CHU, J. P. REESE, AND C. WINTON, *Projected pseudotransient continuation*, SIAM J. Numer. Anal., 46 (2008), pp. 3071–3083.
[17] P. KOSOWSKI AND A. SMOKTUNOWICZ, *On constructing unit triangular matrices with prescribed singular values*, Computing, 64 (2000), pp. 279–285.
[18] S. ŁOJASIEWICZ AND M.-A. ZURRO, *On the gradient inequality*, Bull. Polish Acad. Sci. Math., 47 (1999), pp. 143–145.
[19] G. MARSAGLIA AND I. OLKIN, *Generating correlation matrices*, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 470–475.
[20] A. W. MARSHALL AND I. OLKIN, *Inequalities: theory of majorization and its applications*, vol. 143 of Mathematics in Science and Engineering, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1979.
[21] L. MIRSKY, *Matrices with prescribed characteristic roots and diagonal elements*, J. London Math. Soc., 33 (1958), pp. 14–21.
[22] J. PINTÉR, *Global optimization in action*, vol. 6 of Nonconvex Optimization and its Applications, Kluwer Academic Publishers, Dordrecht, 1996. Continuous and Lipschitz optimization: algorithms, implementations and applications.
[23] F. Y. SING, *Some results on matrices with prescribed diagonal elements and singular values*, Canad. Math. Bull., 19 (1976), pp. 89–92.
[24] R. C. THOMPSON, *Singular values, diagonal elements, and convexity*, SIAM J. Appl. Math., 32 (1977), pp. 39–63.
[25] H. WEYL, *Inequalities between the two kinds of eigenvalues of a linear transformation*, Proc. Nat. Acad. Sci. U. S. A., 35 (1949), pp. 408–411.
[26] H. ZHA AND Z. ZHANG, *A note on constructing a symmetric matrix with specified diagonal entries and eigenvalues*, BIT, 35 (1995), pp. 448–451.