Title
Finding Well-Conditioned Similarities to Block-diagonalize Nonsymmetric Matrices is NP-hard 1

Author
Gu, M.

Publication Date
1994-08-01

LBL-35100
UC-405
Preprint

Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

Physics Division

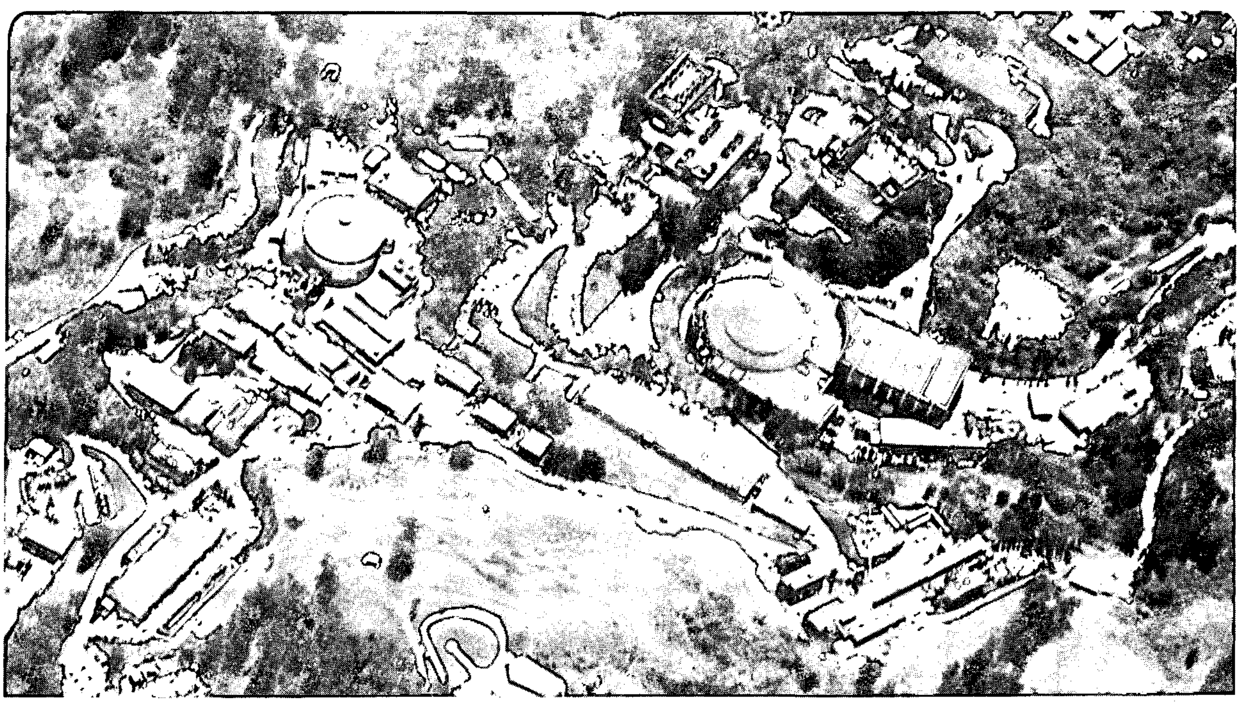
Mathematics Department

To be submitted for publication

Finding Well-conditioned Similarities to Block-diagonalize Nonsymmetric Matrices is NP-hard

M. Gu

August 1994



Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

LBL-35100

Finding Well-conditioned Similarities to Block-diagonalize Nonsymmetric Matrices is NP-hard [1]

Ming Gu
Department of Mathematics and Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

August 1994

[1]Supported by the Applied Mathematical Sciences Subprogram of the Office of Energy Research, U.S. Department of Energy under Contract DE-AC03-76SF00098.

Abstract

Given an upper triangular matrix $A \in \mathbf{R}^{n \times n}$ and a tolerance τ, we show that the problem of finding a similarity transformation G such that $G^{-1} A G$ is block diagonal with the condition number of G being at most τ is NP-hard. Let $f(n)$ be a polynomial in n. We also show that the problem of finding a similarity transformation G such that $G^{-1} A G$ is block diagonal with the condition number of G being at most $f(n)$ times larger than the smallest possible is NP-hard.

1 Introduction

The numerical procedure of computing the Jordan canonical form of a square non-symmetric matrix A can be very unstable, due to the fact that the Jordan canonical form can change drastically under even a very small perturbation [7, page 390]. A more reliable procedure is to partition the eigenvalues into groups and block diagonalize the matrix so that each block has the eigenvalues belonging to one group. Let G be the best conditioned non-singular matrix such that $G^{-1}\,A\,G$ is a block diagonal matrix. It is known [4, 8, 12, 16] that the larger the condition number of G, the more sensitive to perturbations are at least some of these blocks. Algorithms have been proposed to efficiently find partitions whose similarity transformations are sufficiently well-conditioned (see [7, pages 386-389] and the references therein). However, other than the obvious but extremely slow brute force approach, no algorithm has been known to always predict the existence of such partitions, let alone find one. This is true even if we restrict the number of groups (i.e., the number of blocks in $G^{-1}\,A\,G$) to be two.

The objective of this paper is to show that this problem is inherently hard by proving that the existence problem alone is NP-complete. To understand what this means, we first briefly review some related concepts. A problem is called a decision problem if the solution is either "yes" or "no". There are two well-known classes of decision problems, P and NP. Roughly speaking, every problem in P can be *solved* in time proportional to a polynomial in the input size (polynomial time); while a "yes" solution to every problem in NP can be *verified* in polynomial time. It is known that $\mathrm{P} \subseteq \mathrm{NP}$, and there is strong evidence that $\mathrm{P} \neq \mathrm{NP}$. An NP-complete problem is a problem in NP whose solution is as "hard" to find as any other problem in NP. An NP-hard problem is a problem whose solution is as "hard" to find as any NP-complete problem. There does not exist a polynomial time algorithm for *any* NP-complete or NP-hard problem unless $\mathrm{P} = \mathrm{NP}$. For an extensive treatment of this subject, the reader is referred to [6].

To fix the notation, we take $\kappa(G) = \|G\|_2\,\|G^{-1}\|_2$ to be the condition number of G. We assume that the input numbers are all integers. Since entries in A can be rational numbers, we require that there is an integer β in the input such that the (i,j) entry of A is of the form $a_{i,j}/\beta$, where $a_{i,j}$ is an integer in the input. To simplify the problem, in the rest of this paper we assume that A is upper triangular[1]; and that the diagonal elements of A are distinct. We adopt the convention that a block diagonal matrix has at least two diagonal blocks. We will use *poly*(n) to denote a positive valued polynomial in n.

[1]It is relevent to mention that in finite precision computation, an arbitrary square matrix can in general be brought to upper triangular form (the Schur form) by using orthogonal or unitary transformations in cost cubic in the matrix size. See [7] for details.

Consider the following two decision problems:

DICHOTOMY

instance: An upper triangular matrix $A \in \mathbf{R}^{n \times n}$ with distinct diagonal elements, and a tolerance $\tau \geq 1$.

problem: Does there exist a non-singular matrix G such that $G^{-1}\, A\, G$ is 2×2 block-diagonal with $\kappa(G) \leq \tau$?

and

INVARIANT SUBSPACE

instance: An upper triangular matrix $A \in \mathbf{R}^{n \times n}$ with distinct diagonal elements, and a tolerance $\tau \geq 1$.

problem: Does there exist a non-singular matrix G such that $G^{-1}\, A\, G$ is block-diagonal and $\kappa(G) \leq \tau$?

DICHOTOMY is a slightly restricted version of INVARIANT SUBSPACE. It is introduced because it is easier to handle. We will show that DICHOTOMY, and hence INVARIANT SUBSPACE, is NP-complete. It follows that the problem of finding a matrix G such that $G^{-1}\, A\, G$ is block-diagonal and $\kappa(G) \leq \tau$ is NP-complete.

Let G_{opt} be a matrix such that $G_{opt}^{-1}\, A\, G_{opt}$ is block diagonal and $\kappa(G_{opt})$ is as small as possible. We will further show that the approximation problem of finding a matrix G such that $G^{-1}\, A\, G$ is block-diagonal with $\kappa(G) \leq f(n) \cdot \kappa(G_{opt})$ is NP-hard, where $f(n)$ is a fixed polynomial in n.

In this paper, we will exclusively consider real matrices, even though our results hold for complex matrices as well. Our results also generalize to regular matrix pencils. We only use 2-norm in this paper, but the results hold for any norm $\|\cdot\|$ such that

$$\frac{\|x\|_2}{p_1(n)} \leq \|x\| \leq p_2(n) \cdot \|x\|_2 ,$$

where $x \in \mathbf{R}^n$ and $p_1(n)$ and $p_2(n)$ are positive-valued polynomials in n.

It has been conjectured by Demmel [5] that INVARIANT SUBSPACE is NP-complete. Our results confirm this conjecture.

We emphasize that the fact that DICHOTOMY is NP-complete only implies that it is very unlikely that there is any general program that runs efficiently on all instances of DICHOTOMY. It is still possible that there are algorithms that can solve most pratical cases of DICHOTOMY efficiently.

The rest of this paper is organized as follows. Section 2 introduces some results in the literature and provides some necessary tools. Section 3 shows the NP-completeness of DICHOTOMY and the NP-hardness of the corresponding approximation problem.

2 Theoretical background

In this section we consider some properties of similarity transformations that 2×2 block-diagonalize A. For a matrix X, $\sigma_i(X)$ denotes its i-th largest singular value, and $\mathrm{diag}(X)$ denotes its main diagonal. The following lemma is classical.

Lemma 1 (Golub and van Loan [7, pages 335-338]) *Let $A \in \mathbf{R}^{n\times n}$ be an upper triangular matrix with distinct diagonal elements. Then there exists a non-singular upper triangular matrix G such that*

$$G^{-1} \cdot A \cdot G = \mathrm{diag}(A)\,. \tag{2.1}$$

Remark 1: The diagonal elements of A are the eigenvalues of A and the columns of G are the corresponding eigenvectors. Since A has distinct eigenvalues, its eigenvector matrix G is unique up to scalings of its columns. Hence any non-singular matrix G satisfying (2.1) is upper triangular. We will always choose G such that each of its columns has norm roughly of the order between 1 and n.

Lemma 2 (Demmel [4]) *Let G satisfy (2.1). For a given $1 \le k \le n-1$, partition*

$$\mathrm{diag}(A) = \begin{pmatrix} \Lambda_1 & \\ & \Lambda_2 \end{pmatrix} \quad \textit{and} \quad G = \begin{pmatrix} E & F \\ & K \end{pmatrix},$$

where $\Lambda_1 \in \mathbf{R}^{k\times k}$; $\Lambda_2 \in \mathbf{R}^{(n-k)\times(n-k)}$; $E \in \mathbf{R}^{k\times k}$; $F \in \mathbf{R}^{k\times(n-k)}$; and $K \in \mathbf{R}^{(n-k)\times(n-k)}$. Define $P = F \cdot K^{-1} \in \mathbf{R}^{k\times(n-k)}$ and

$$G_{\mathrm{opt}} = \begin{pmatrix} I & P\cdot(I+P^T\cdot P)^{-\frac{1}{2}} \\ & (I+P^T\cdot P)^{-\frac{1}{2}} \end{pmatrix}.$$

Then G_{opt} is one of the similarity transformations that 2×2 block diagonalize A with the upper left block containing the eigenvalues of Λ_1, and its condition number, $\kappa(G_{\mathrm{opt}}) = \|P\|_2 + \sqrt{1+\|P\|_2^2}$, is the smallest among such similarity transformations.

Remark 2: Let Z be a projection that projects $\mathbf{R}^n$ onto the space spanned by eigenvectors pertaining to eigenvalues in Λ_1, then we also have (see [4])

$$\kappa(G_{opt}) = \|Z\|_2 + \sqrt{\|Z\|_2^2 - 1}\,. \tag{2.2}$$

Now we consider similarity transformations that 2×2 block diagonalize A with the upper left block having eigenvalues different from those of Λ_1. Let Π be a permutation matrix and partition

$$\Pi^T \cdot \mathrm{diag}(A) \cdot \Pi = \begin{pmatrix} \hat{\Lambda}_1 & \\ & \hat{\Lambda}_2 \end{pmatrix}.$$

To block-diagonalize A such that the upper left block contains the eigenvalues in $\hat{\Lambda}_1$, we first re-arrange the diagonal elements of A.

Lemma 3 (Golub and van Loan [7, page 335]) *Let $A \in \mathbf{R}^{n\times n}$ be an upper triangular matrix with distinct diagonal elements, and let Π be a permutation matrix. Then there exists an orthogonal matrix Q such that*

$$\mathrm{diag}(Q^T \cdot A \cdot Q) = \Pi^T \cdot \mathrm{diag}(A) \cdot \Pi \,. \tag{2.3}$$

The orthogonal matrix Q can be efficiently computed [1, 2, 14, 15]. Let Q be the orthogonal matrix in Lemma 3. Then relation (2.1) is equivalent to

$$(Q^T \cdot G \cdot \Pi)^{-1} \cdot (Q^T \cdot A \cdot Q) \cdot (Q^T \cdot G \cdot \Pi) = \mathrm{diag}(Q^T \cdot A \cdot Q) \,.$$

According to Lemma 1 and Remark 1, this relation implies that $\hat{G} \equiv Q^T \cdot G \cdot \Pi$ is upper triangular. We write this simple relation in a more informative way:

$$G \cdot \Pi = Q \cdot \hat{G} \,. \tag{2.4}$$

In other words, the orthogonal matrix Q in (2.3) and the upper triangular matrix $\hat{G}$ such that

$$\hat{G}^{-1} \cdot (Q^T \cdot A \cdot Q) \cdot \hat{G} = \mathrm{diag}(Q^T \cdot A \cdot Q)$$

are simply the Q and R factors in the QR factorization of $G \cdot \Pi$.

Relation (2.4) is QR factorization with column permutation. It is known that one can use QR factorization to reveal the ill-conditioning of G by choosing a special permutation matrix Π [3, 9, 10, 11]

Theorem 1 (Gu and Eisenstat [11]) *Let G be a non-singular upper triangular matrix. Then for any given integer $1 \le k \le n-1$, there exists a permutation Π such that in the QR factorization*[2]

$$G \cdot \Pi = Q \cdot \hat{G} = Q \cdot \begin{pmatrix} \hat{E} & \hat{F} \\ & \hat{K} \end{pmatrix} ,$$

with $\hat{E} \in \mathbf{R}^{k\times k}$, $\hat{F} \in \mathbf{R}^{k\times(n-k)}$, and $\hat{K} \in \mathbf{R}^{(n-k)\times(n-k)}$, we have

$$\hat{F} = \hat{E} \cdot \hat{L} \quad \textit{with} \quad \|\hat{L}\|_2 \le n^2$$

and

$$\sigma_i(\hat{E}) \ge \frac{\sigma_i(G)}{n^2} \quad \textit{and} \quad \sigma_j(\hat{K}) \le \sigma_{j+k}(G) \cdot n^2 \,,$$

where $1 \le i \le k$ and $1 \le j \le n-k$.

In particular, this theorem states that if $\sigma_k(G)$ is much larger than $\sigma_{k+1}(G)$, then the first k columns of $G \cdot \Pi$ are linearly independent and the last $n-k$ columns are close to being in the subspace spanned by the first k columns.

[2]This factorization can be computed in cost proportional to n^3. See [11] for details.

In the following, we turn our attention to a G matrix of the form

$$G = \begin{pmatrix} E & E \cdot L \\ & K \end{pmatrix},$$

where $E \in \mathbf{R}^{k\times k}$ is relatively well-conditioned; $L \in \mathbf{R}^{k\times(n-k)}$ with $\|L\|_2$ bounded by a low degree polynomial in n; and $K \in \mathbf{R}^{(n-k)\times(n-k)}$ with $\|K\|_2$ very small. It is easy to check that under these conditions $\sigma_k(G)$ is indeed much larger than $\sigma_{k+1}(G)$.

We first examine how E, L and K change when permuting the first k columns of G. Let $\Gamma \in \mathbf{R}^{k\times k}$ be a permutation matrix, let $\Pi = \begin{pmatrix} \Gamma & \\ & I_{n-k} \end{pmatrix}$, and let $E\cdot\Gamma = W\cdot\hat{E}$ be the QR factorization of $E \cdot \Gamma$. Then

$$\begin{aligned} G\cdot\Pi &= \begin{pmatrix} E\cdot\Gamma & E\cdot\Gamma\cdot\Gamma^T\cdot L \\ & K \end{pmatrix} \\ &= \begin{pmatrix} W & \\ & I_{n-k} \end{pmatrix} \begin{pmatrix} \hat{E} & \hat{E}\cdot\hat{L} \\ & K \end{pmatrix}, \end{aligned}$$

where $\hat{L} = \Gamma^T \cdot L$. Thus the elements of L are invariant under permutations of the first k columns of G.

We want to examine the smallest condition number for any similarity transformation that brings A to 2×2 block diagonal form with the upper left block having the first s eigenvalues of $\Pi^T \cdot \mathrm{diag}(A) \cdot \Pi$, for some $1 \le s \le k-1$. To this end, partition

$$\hat{E} = \begin{pmatrix} \hat{E}_{11} & \hat{E}_{12} \\ & \hat{E}_{22} \end{pmatrix} \quad \text{and} \quad \hat{L} = \begin{pmatrix} \hat{L}_1 \\ \hat{L}_2 \end{pmatrix},$$

where $\hat{E}_{11} \in \mathbf{R}^{s\times s}$; $\hat{E}_{12} \in \mathbf{R}^{s\times(k-s)}$; $\hat{E}_{22} \in \mathbf{R}^{(k-s)\times(k-s)}$; $\hat{L}_1 \in \mathbf{R}^{s\times(n-k)}$; and $\hat{L}_2 \in \mathbf{R}^{(k-s)\times(n-k)}$. Then

$$G\cdot\Pi = \begin{pmatrix} W & \\ & I_{n-k} \end{pmatrix} \begin{pmatrix} \hat{E}_{11} & \hat{E}_{12} & \hat{E}_{11}\cdot\hat{L}_1 + \hat{E}_{12}\cdot\hat{L}_2 \\ & \hat{E}_{22} & \hat{E}_{22}\cdot\hat{L}_2 \\ & & K \end{pmatrix}.$$

By Lemma 2, such smallest condition number has the form $\|P\|_2 + \sqrt{1+\|P\|_2^2}$, where

$$\begin{aligned} P &= \begin{pmatrix} \hat{E}_{12} & \hat{E}_{11}\cdot\hat{L}_1 + \hat{E}_{12}\cdot\hat{L}_2 \end{pmatrix} \begin{pmatrix} \hat{E}_{22} & \hat{E}_{22}\cdot\hat{L}_2 \\ & K \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \hat{E}_{12} & \hat{E}_{11}\cdot\hat{L}_1 + \hat{E}_{12}\cdot\hat{L}_2 \end{pmatrix} \begin{pmatrix} \hat{E}_{22}^{-1} & -\hat{L}_2\cdot K^{-1} \\ & K^{-1} \end{pmatrix} \\ &= \begin{pmatrix} \hat{E}_{12}\cdot\hat{E}_{22}^{-1} & \hat{E}_{11}\cdot\hat{L}_1\cdot K^{-1} \end{pmatrix}. \end{aligned}$$

It follows that

$$\|\hat{E}_{11} \cdot \hat{L}_1 \cdot K^{-1}\|_2 \le \|P\|_2 \le \|\hat{E}_{12} \cdot \hat{E}_{22}^{-1}\|_2 + \|\hat{E}_{11} \cdot \hat{L}_1 \cdot K^{-1}\|_2 \,. \tag{2.5}$$

To simplify this relation, we first observe that

$$\|\hat{E}_{12} \cdot \hat{E}_{22}^{-1}\|_2 \le \|\hat{E}_{12}\|_2 \cdot \|\hat{E}_{22}^{-1}\|_2 \le \kappa(\hat{E}) = \kappa(E) \,.$$

On the other hand, let E_S be the columns of G that are permuted to become the first s columns of $G \cdot \Pi$, and let L_S be the corresponding rows in L. Then

$$E_S = \begin{pmatrix} W & \\ & I_{n-k} \end{pmatrix} \begin{pmatrix} \hat{E}_{11} \\ 0 \end{pmatrix} \quad \text{and} \quad L_S = \hat{L}_1 \,,$$

and hence

$$\|\hat{E}_{11} \cdot \hat{L}_1 \cdot K^{-1}\|_2 = \left\| \begin{pmatrix} W & \\ & I_{n-k} \end{pmatrix} \begin{pmatrix} \hat{E}_{11} \\ 0 \end{pmatrix} \hat{L}_1 \cdot K^{-1} \right\|_2 = \|E_S \cdot L_S \cdot K^{-1}\|_2 \,.$$

Plugging these relations into (2.5) we obtain

$$\|E_S \cdot L_S \cdot K^{-1}\|_2 \le \|P\|_2 \le \kappa(E) + \|E_S \cdot L_S \cdot K^{-1}\|_2 \,.$$

To appreciate this relation, we consider the special case where $K = (\gamma) \in \mathbf{R}^{1\times 1}$. In this case L is a column vector. We have

$$\frac{\|E_S \cdot L_S\|_2}{|\gamma|} \le \|P\|_2 \le \kappa(E) + \frac{\|E_S \cdot L_S\|_2}{|\gamma|} \,. \tag{2.6}$$

Assume that $|\gamma| \ll 1/\tau$. To 2×2 block-diagonalize A, relation (2.6) suggests that it is necessary to find a subset S such that

$$\|E_S \cdot L_S\|_2 \le \tau \cdot |\gamma| \ll 1 \,.$$

Note that minimizing $\|E_S \cdot L_S\|_2$ is 0-1 integer programming, which is in general NP-hard [6].

3 DICHOTOMY is NP-complete

A standard process of devising an NP-completeness proof for a decision problem $\mathbf{D}$ consists the following four steps (see [6, page 45]):

1. showing that $\mathbf{D}$ is in NP;
2. selecting a known NP-complete $\bar{\mathbf{D}}$;
3. constructing a reduction from $\bar{\mathbf{D}}$ to $\mathbf{D}$;

4. proving that the reduction is a *polynomial transformation.*

Now we explain the term polynomial transformation. The reduction from $\bar{\mathbf{D}}$ to $\mathbf{D}$ is an algorithm that solves $\bar{\mathbf{D}}$ by calling the algorithm for solving $\mathbf{D}$ as a subroutine. Informally, the reduction is a polynomial transformation if the difference between its overall cost and the total cost on all the subroutine calls for solving $\mathbf{D}$ is at most a polynomial in the input size.

We first need to show that DICHOTOMY is in NP. Given a partition, we consider the projection Z that accordingly dichotomizes the spectrum of A. Z can be written as a sum of projections onto certain individual eigenspaces defined by the partition (see [12] for details) and hence is a rational function of entries of A. Note that according to equation (2.2), the solution to DICHOTOMY is 'yes' if and only if

$$\|Z\|_2 + \sqrt{\|Z\|_2^2 - 1} \leq \tau,$$

or, equivalently,

$$\|Z\|_2 \leq \frac{\tau^2 + 1}{2\tau},$$

which holds if and only if the matrix

$$\hat{Z} \equiv \left(\frac{\tau^2 + 1}{2\tau}\right)^2 \cdot I - Z^T Z$$

is semi-positive definite, where I is the identity matrix. In other words, the solution to DICHOTOMY is 'yes' if and only if all the principle submatrices of $\hat{Z}$ has non-negative determinants. It is straightforward to show that checking whether all the determinants of the principle submatrices of $\hat{Z}$ are non-negative can be done in time polynomial in n and $\log_2 b$. Therefore DICHOTOMY is in NP.

Now we focus on the remaining three steps. We pick the known NP-complete problem to be SUBSET SUM.

SUBSET SUM

instance: Positive integers $a_1, \ldots, a_m$ and b, with $b > \max_{1 \leq i \leq m} a_i$.

problem: Does there exist a subset $\mathcal{S} \subseteq \{1, \ldots, m\}$ such that

$$\sum_{i \in \mathcal{S}} a_i = b\ ?$$

It is known [13] that SUBSET SUM is NP-complete. It is also known [6, page 223] that SUBSET SUM can be solved in time polynomial in bn, which can be exponentially larger than the input size $O(n \log_2 b)$. Note that SUBSET SUM is solvable in time polynomial in $O(n \log_2 b)$ if b is bounded by a polynomial in n [6, page 223].

Our goal is to reduce SUBSET SUM to a special case of DICHOTOMY. Since this section involves the detailed reduction construction, it is not as insightful as the previous one.

Now we construct a simple class of matrices to which SUBSET SUM can be reduced. We consider the eigenvector matrix $G \in \mathbf{R}^{(m+3)\times(m+3)}$ of the form

$$G = \begin{pmatrix} E & E \cdot l \\ & \delta^2 \end{pmatrix} ,$$

where $l \in \mathbf{R}^{m+2}$ is a column vector with $\|l\|_2 = poly(m)$; and $0 < 1/\delta < \tau < 1/\delta^2$. We further choose $E \in \mathbf{R}^{(m+2)\times(m+2)}$ to be

$$E = \begin{pmatrix} 1 & e^T \\ & 2\delta \cdot I \end{pmatrix} ,$$

where $e = (1, \ldots, 1)^T \in \mathbf{R}^{m+1}$. It can be easily verified that $\kappa(E) \leq 2(m+3)/\delta$. Partition $l = (g_1, g^T)^T$. Now G can be written as

$$G = \begin{pmatrix} 1 & e^T & g_1 + e^T g \\ & 2\delta \cdot I & 2\delta \cdot g \\ & & \delta^2 \end{pmatrix} . \tag{3.7}$$

We will specify the parameters g_1 and $g = (g_2, \ldots, g_{m+2})^T \in \mathbf{R}^{m+1}$ later. The matrix G is non-zero only on its diagonal and its first row and last column. The matrix G^{-1} has a similar form:

$$G^{-1} = \begin{pmatrix} 1 & -e^T/(2\delta) & -g_1/\delta^2 \\ & I/(2\delta) & -g/\delta^2 \\ & & 1/\delta^2 \end{pmatrix} .$$

Now we construct the matrix A. Since $G^{-1} \cdot A \cdot G$ is diagonal, A will be completely determined once we have chosen these diagonal elements (the eigenvalues of A). To simplify matters, we want them to be distinct and evenly distributed. To make the reduction more interesting, we also want $\|A\|_2$ to be of the order m. Due to these considerations, we choose the eigenvalues to be $\lambda_1 = 0$, $\lambda_2 = -2\delta$, $\lambda_{i+2} = 2\delta \cdot i$ for $1 \leq i \leq m$, and $\lambda_{m+3} = \delta$. We claim that A has the following form:

$$A = \begin{pmatrix} 0 & e^T\Omega & 0 \\ & 2\delta \cdot \Omega & -2(2 \cdot \Omega - I)g \\ & & \delta \end{pmatrix} , \tag{3.8}$$

where $\Omega = \mathrm{diag}(\omega_2, \ldots, \omega_{m+2}) \in \mathbf{R}^{(m+1)\times(m+1)}$ is diagonal, with $\omega_2 = -1$ and $\omega_j = j - 2$ for $3 \leq j \leq m+2$; and the parameter g_1 in equation (3.7) is taken to be

$$g_1 = e^T(2 \cdot \Omega - I)g = \sum_{j=2}^{m+2} (2\omega_j - 1)g_j .$$

Now we verify that G is indeed an eigenvector matrix for A. In fact

$$\begin{aligned}
G^{-1}AG &= G^{-1}\begin{pmatrix} 0 & e^T\Omega & 0 \\ & 2\delta\cdot\Omega & -2(2\cdot\Omega - I)g \\ & & \delta \end{pmatrix}\begin{pmatrix} 1 & e^T & g_1 + e^T g \\ & 2\delta\cdot I & 2\delta\cdot g \\ & & \delta^2 \end{pmatrix} \\
&= \begin{pmatrix} 1 & -e^T/(2\delta) & -g_1/\delta^2 \\ & I/(2\delta) & -g/\delta^2 \\ & & 1/\delta^2 \end{pmatrix}\begin{pmatrix} 0 & 2\delta e^T\Omega & 2\delta e^T\Omega g \\ & 4\delta^2\cdot\Omega & 2\delta^2\cdot g \\ & & \delta^3 \end{pmatrix} \\
&= \begin{pmatrix} 0 & & \\ & 2\delta\cdot\Omega & \\ & & \delta \end{pmatrix}.
\end{aligned}$$

We note that the eigenvalues of A are $O(\delta)$ apart but have condition numbers[3] $O(1/\delta^2)$.

Theorem 2 DICHOTOMY *is* NP-*complete.*

Proof: We have shown that DICHOTOMY is in NP. Now we reduce SUBSET SUM to DICHOTOMY. Given positive integers $a_1, \ldots, a_m$ and b, with $b > \max_{1\le i\le m} a_i$, we set $\delta = 1/b^2$, $\tau = \lfloor b^{2.5} \rfloor$, $g_2 = -1$ and $g_{i+2} = a_i/b$ for $1 \le i \le m$. We assume that $b > 64(m+3)^2$. If $b \le 64(m+3)^2$, then SUBSET SUM can be solved in time polynomial in the input size. Our goal is to show that under this reduction, the solution to SUBSET SUM is "yes" if and only if the solution to DICHOTOMY is "yes", and hence solving SUBSET SUM is equivalent to solving DICHOTOMY.

Let $\mathcal{S}$ be a non-empty subset of $\{1, \ldots, m+2\}$. Let $G_{\mathcal{S}}$ be a similarity transformation with the smallest condition number such that $G_{\mathcal{S}}^{-1}\cdot A\cdot G_{\mathcal{S}}$ is 2×2 block diagonal; and let λ_i be an eigenvalue in the upper left block if and only if $i \in \mathcal{S}$. Let $E_{\mathcal{S}}$ be the matrix consisting of the corresponding columns of E, and let $l_{\mathcal{S}}$ be the vector consisting of the corresponding components of l. By Lemma 2 and relation (2.6), we have $\kappa(G_{\mathcal{S}}) = \|P\|_2 + \sqrt{1 + \|P\|_2^2}$, where

$$\frac{\|E_{\mathcal{S}}\cdot l_{\mathcal{S}}\|_2}{\delta^2} \le \|P\|_2 \le \frac{2(m+3)}{\delta} + \frac{\|E_{\mathcal{S}}\cdot l_{\mathcal{S}}\|_2}{\delta^2}.$$

Further, the special form of the matrix E implies that

$$|\sum_{i\in\mathcal{S}} g_i| \le \|E_{\mathcal{S}}\cdot l_{\mathcal{S}}\|_2 \le 2\delta\sqrt{m+3} + |\sum_{i\in\mathcal{S}} g_i|$$

Combining these relations, we have

$$\frac{2\cdot|\sum_{i\in\mathcal{S}} g_i|}{\delta^2} \le \kappa(G_{\mathcal{S}}) \le \frac{8(m+3)}{\delta} + \frac{2\cdot|\sum_{i\in\mathcal{S}} g_i|}{\delta^2}. \tag{3.9}$$

[3]Let λ be an eigenvalue of A. Then there exist unit vectors x and y such that $Ax = \lambda x$ and $y^T A = \lambda y^T$. The condition number of λ is defined to be $1/|y^T x|$ (see [7, page 344]).

By construction, we see that $b \cdot \sum_{i\in\mathcal{S}} g_i$ is always an integer. Hence we have either $|\sum_{i\in\mathcal{S}} g_i| \geq 1/b$ or $\sum_{i\in\mathcal{S}} g_i = 0$.

If $|\sum_{i\in\mathcal{S}} g_i| \geq 1/b$, then relation (3.9) implies that

$$\kappa(G_{\mathcal{S}}) > \frac{2}{b\delta^2} > \tau .$$

On the other hand, if $|\sum_{i\in\mathcal{S}} g_i| = 0$, then relation (3.9) implies that

$$\kappa(G_{\mathcal{S}}) < \frac{8(m+3)}{\delta} < \tau .$$

Summarizing, we have shown that $\kappa(G_{\mathcal{S}}) < \tau$ if and only if $|\sum_{i\in\mathcal{S}} g_i| = 0$.

Since $|\sum_{i\in\mathcal{S}} g_i| > 1$ as long as $1 \in \mathcal{S}$, it follows that $|\sum_{i\in\mathcal{S}} g_i| = 0$ if and only if $1 \notin \mathcal{S}$, $2 \in \mathcal{S}$, and

$$\sum_{i\in\hat{\mathcal{S}}} a_i = b \,, \quad \text{where} \quad \hat{\mathcal{S}} = \{i \mid \quad i > 0 \quad \text{and} \quad i+2 \in \mathcal{S}\} .$$

The solution to DICHOTOMY is "yes" if and only if there exists an index set $\mathcal{S}$ such that $\kappa(G_{\mathcal{S}}) \leq \tau$; and the solution to SUBSET SUM is "yes" if and only if there exists an index set $\hat{\mathcal{S}}$ such that $\sum_{i\in\hat{\mathcal{S}}} a_i = b$. This shows that the solution to DICHOTOMY is the same as the solution to SUBSET SUM. Hence we have reduced SUBSET SUM to DICHOTOMY and thus DICHOTOMY is NP-complete. ∎

It now follows that finding the least conditioned similarity transformation G such that $G^{-1} \cdot A \cdot G$ is block-diagonal is NP-hard. So we consider the following less ambitious problem. Let $f(n)$ be a positive-valued polynomial in n. Given an upper triangular matrix $A \in \mathbf{R}^{n\times n}$ with distinct diagonal elements, G_{opt} denotes the matrix with the smallest condition number that block-diagonalizes A.

APPROXIMATION

instance: An upper triangular matrix $A \in \mathbf{R}^{n\times n}$ with distinct diagonal elements.

problem: Find a non-singular matrix G_{approx} such that $G_{approx}^{-1} \cdot A \cdot G_{approx}$ is block-diagonal with $\kappa(G_{approx}) \leq f(n) \cdot \kappa(G_{opt})$.

The following theorem says that APPROXIMATION is still very hard.

Theorem 3 APPROXIMATION *is* NP-*hard.*

Proof: We reduce SUBSET SUM to APPROXIMATION using the reduction in the proof of Theorem 2. Given positive integers $a_1, \ldots, a_m$ and b, with $b > \max_{1\leq i\leq m} a_i$, we set $\delta = 1/b^2$, $\tau = \lfloor b^{2.5} \rfloor$, $g_2 = -1$ and $g_{i+2} = a_i/b$ for $1 \leq i \leq m$. We assume that $b > 64 \cdot f^2(m) \cdot (m+3)^2$. If $b \leq 64 \cdot f^2(m) \cdot (m+3)^2$, then SUBSET SUM can be solved in time polynomial in the input size. We will show that under this reduction, the solution to SUBSET SUM is "yes" if and only if $\kappa(G_{approx}) < b^3$.

Let $\mathcal{S}$ be a non-empty subset of $\{1, \ldots, m+2\}$. Let $G_{\mathcal{S}}$ be a similarity transformation with the smallest condition number such that $G_{\mathcal{S}}^{-1} \cdot A \cdot G_{\mathcal{S}}$ is block diagonal; and let λ_i be an eigenvalue in the upper left block if and only if $i \in \mathcal{S}$. Let $E_{\mathcal{S}}$ be the matrix consisting of the corresponding columns of E, and let $l_{\mathcal{S}}$ be the vector consisting of the corresponding components of l. Similar to the proof of Theorem 2, we have

$$\frac{2 \cdot |\sum_{i \in \mathcal{S}} g_i|}{\delta^2} \leq \kappa(G_{\mathcal{S}}) \leq \frac{8(m+3)}{\delta} + \frac{2 \cdot |\sum_{i \in \mathcal{S}} g_i|}{\delta^2} . \tag{3.10}$$

As before, $b \cdot \sum_{i \in \mathcal{S}} g_i$ is always an integer. Hence we either have $|\sum_{i \in \mathcal{S}} g_i| \geq 1/b$ or $\sum_{i \in \mathcal{S}} g_i = 0$.

If $|\sum_{i \in \mathcal{S}} g_i| \geq 1/b$, then relation (3.10) implies that

$$\kappa(G_{\mathcal{S}}) > \frac{2}{b\delta^2}.$$

Hence, in the case $|\sum_{i \in \mathcal{S}} g_i| \geq 1/b$ for *all* non-empty index sets $\mathcal{S}$, we have

$$\kappa(G_{approx}) \geq \kappa(G_{opt}) > \frac{1}{b\delta^2} = b^3 .$$

On the other hand, if $|\sum_{i \in \mathcal{S}} g_i| = 0$ for *some* non-empty index set $\mathcal{S}$, then relation (3.10) implies that

$$\kappa(G_{\mathcal{S}}) < \frac{8(m+3)}{\delta} ,$$

and hence

$$\kappa(G_{approx}) \leq \kappa(G_{opt}) \cdot f(m) < \frac{8(m+3) \cdot f(m)}{\delta} \leq b^3 .$$

Summarizing, we have shown that $\kappa(G_{approx}) < b^3$ if and only if there exists a non-empty index set $\mathcal{S}$ such that $|\sum_{i \in \mathcal{S}} g_i| = 0$. Similar to the arguments in the proof of Theorem 2, this implies that $\kappa(G_{approx}) < b^3$ if and only if the solution to SUBSET SUM is "yes". Hence APPROXIMATION is NP-hard. ∎

By using the standard techniques for reducing the optimization problem to its corresponding decision problem (see [6, pages 110-117]), we can show that APPROXIMATION and the problem of finding the best conditioned similarity transformation G can be reduced to INVARIANT SUBSPACE by polynomial transformations. Hence these two problems are as "hard" as INVARIANT SUBSPACE and thus *NP-equivalent* (see [6, pages 110-117]).

Acknowledgements The author thanks Professor J. Demmel for sharing with him many thoughtful insights into the subject of this paper, Professors L. Blum, A. Edelman, W. Kahan, and B. Parlett for helpful discussions, and Professor S. Eisenstat for suggestions that improved the presentation.

References

[1] Z. Bai and J. Demmel, *On swapping diagonal blocks in real Schur form*, Lin. Alg. Appl., 186 (1993), pp. 73–96.

[2] A. Bojanczyk and P. Van Dooren, *Reordering diagonal blocks in real schur form*, in Linear Algebra for Large Scale and Real-Time Applications, M. Moonen, G. Golub, and B. de Moor, eds., Kluwer Academic Publishers, 1993, pp. 351–352. NATO-ASI Series E: Applied Sciences, Vol. 232.

[3] T. Chan, *Rank revealing QR factorizations*, Lin. Alg. Appl., 88/89 (1987), pp. 67–82.

[4] J. Demmel, *The condition number of equivalence transformations that block diagonalize matrix pencils*, SIAM J. Num. Anal., 20 (1983), pp. 599–610.

[5] ——. personal communication, 1993.

[6] M. Garey and D. Johnson, *Computers and Intractability*, W. H. Freeman, San Francisco, 1979.

[7] G. Golub and C. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, 2nd ed., 1989.

[8] G. Golub and J. H. Wilkinson, *Ill-conditioned eigensystems and the computation of the Jordan canonical form*, SIAM Review, 18 (1976), pp. 578–619.

[9] G. H. Golub, *Numerical methods for solving linear least squares problems*, Numer. Math., 7 (1965), pp. 206–216.

[10] G. H. Golub, V. Klema, and G. W. Stewart, *Rank degeneracy and least squares problems*, Tech. Rep. TR–456, Department of Computer Science, University of Maryland, 1976.

[11] M. Gu and S. Eisenstat, *An efficient algorithm for computing a rank-revealing QR decomposition*, Computer Science Dept. Report YALEU/DCS/RR-967, Yale University, June 1993.

[12] W. Kahan, *Conserving confluence curbs ill-condition*, Computer Science Dept. Technical Report 6, University of California, Berkeley, CA, August 1972.

[13] R. M. Karp, *Reducibility among combinatorial problems*, in Complexity of Computer Computations, R. E. Miller and J. W. Thatcher, eds., Plenum Press, 1972, pp. 85–103.

[14] A. Ruhe, *An algorithm for numerical determination of the structure of a general matrix*, BIT, 10 (1970), pp. 196–216.

[15] G. W. STEWART, *Algorithm 506 HQR3 and EXCHANG: Fortran subroutine for calculating and ordering the eigenvalues of a real upper Hessenberg matrix [F2]*, ACM Trans. Math. Soft., 2 (1976), pp. 275–280.

[16] J. H. WILKINSON, *Sensitivity of eigenvalues*, Utilitas Math., 25 (1984), pp. 5–76.

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
TECHNICAL INFORMATION DEPARTMENT
BERKELEY, CALIFORNIA 94720