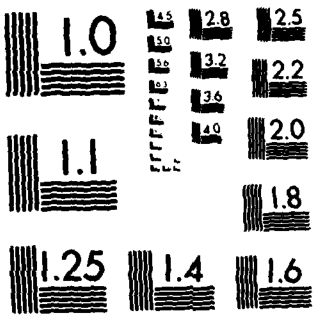
AD-A136 419 ON THE REPRESENTATION OF A BASIS FOR THE NULL SPACE(U) 1/1
STANFORD UNIV CA SYSTEMS OPTIMIZATION LAB
P E GILL ET AL. NOV 83 SOL-83-19 ARO-18424.13-MA
UNCLASSIFIED N00014-75-C-0267 F/G 12/1 NL

END
DATE
FILMED
1-84
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

ARO 18424.13-MA

(12)

Systems
Optimization
Laboratory

AD-A136419

ON THE REPRESENTATION OF A BASIS FOR THE NULL SPACE

by

Philip E. Gill, Walter Murray,
Michael A. Saunders and Margaret H. Wright

TECHNICAL REPORT SOL 83-19
November 1983

Department of Operations Research
Stanford University
Stanford, CA 94305

12

SYSTEMS OPTIMIZATION LABORATORY
DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

ON THE REPRESENTATION OF A BASIS FOR THE NULL SPACE

by

Philip E. Gill, Walter Murray,
Michael A. Saunders and Margaret H. Wright

TECHNICAL REPORT SOL 83-19
November 1983

Research and reproduction of this report were partially supported by the Department of Energy Contract DE-AM03-76SF00326, PA# DE-AT03-76ER72018; Office of Naval Research Contract N00014-75-C-0267; and National Science Foundation Grants, MCS-7926009 and ECS-8012974; and Army Research Office Contract DAAG29-81-K-156.

ON THE REPRESENTATION OF A BASIS FOR THE NULL SPACE

by

Philip E. Gill, Walter Murray,
Michael A. Saunders and Margaret H. Wright

Systems Optimization Laboratory
Department of Operations Research
Stanford University
Stanford, California 94305

ABSTRACT

Given a rectangular matrix $A(x)$ that depends on the independent variables x, many constrained optimization methods involve computations with $Z(x)$, a matrix whose columns form a basis for the null space of $A(x)$. When A is evaluated at a given point, it is well known that a suitable Z (satisfying $AZ = 0$) can be obtained from standard matrix factorizations. However, Coleman and Sorensen have recently shown that standard orthogonal factorization methods may produce orthogonal bases that do not vary continuously with x; they also suggest several techniques for adapting standard factorization schemes so as to ensure continuity of Z in the neighborhood of a given point.

In this note, ~~we~~ *discuss several aspects of the representation* of a basis for the null space. They ~~We~~ describe how an explicit matrix Z can be obtained at any point using a method for updating a factorization with either Householder or stabilized elementary transformations. Under a mild non-singularity assumption, the elements of Z are continuous functions of x. ~~We~~ also show that the chosen form of Z is convenient and efficient when implementing certain methods for nonlinearly constrained optimization.

This research was supported by the U.S. Department of Energy Contract DE-AC03-76SF00326, PA No. DE-AT03-76ER72018; National Science Foundation Grants MCS-7926009 and ECS-8012974; the Office of Naval Research Contract N00014-75-C-0267; and the U.S. Army Research Office Contract DAAG29-79-C-0110.

1. Introduction

Given an $m \times n$ matrix A $(m < n)$, many constrained optimization methods involve computations with a matrix Z whose columns form a basis for the null space of A. It is well known that certain factorizations of A provide stable and efficient means for computing Z. We illustrate this process with the *TQ factorization*, which is defined by

$$AQ = \begin{pmatrix} 0 & T \end{pmatrix}, \tag{1}$$

where Q is an $n \times n$ nonsingular matrix, and T is an $m \times m$ "reverse" triangular matrix such that $T_{ij} = 0$ for $i + j \leq m$. (The reverse-triangular form of T has advantages in implementing certain constrained optimization methods; see Gill *et al.*, 1982). If A has full rank, T is nonsingular; in this case it follows from (1) that the *first* $n - m$ columns of Q can be taken as the columns of the matrix Z. If Q is orthogonal, the TQ factorization is simply a permutation of the QR factorization of A^T (or the LQ factorization of A).

In many applications, the elements of A are functions of an independent variable x $(x \in \Re^n)$ – for example, A often represents the Jacobian of a set of constraints. In such a context, it is desirable that the basis for the null space can be represented by a *continuous* matrix function $Z(x)$. For example, a continuous representation of Z is advantageous in proving local convergence of methods for nonlinear constraints that maintain a quasi-Newton approximation to the projected Hessian of the Lagrangian function, since the operation of projection is defined by Z.

Coleman and Sorensen (1982) recently posed the following question: *Let $\mathcal{B}$ be a ball around a point $\hat{x}$ $(\hat{x} \in \Re^n)$. Suppose that $A(x)$ is an $m \times n$ matrix of rank m whose elements vary continuously with x in $\mathcal{B}$. Is it possible to construct, in a stable and efficient manner, a matrix $Z(x)$ with elements that vary continuously with x in $\mathcal{B}$?* They observe that a "standard" method of computing the orthogonal factorization through Householder matrices may not provide the required continuous matrix $Z(x)$, and then propose several alternative strategies for ensuring a continuous Z in a neighborhood of $\hat{x}$.

This note summarizes the main features of a scheme for computing Z that has been used in software for quadratic programming and nonlinear programming (Gill *et al.*, 1983a and 1983b). In Section 2, we describe the algorithm for computing Z. Section 3 shows that the method produces a continuous basis for the null space under the conditions given by Coleman and Sorensen. Finally, in Section 4 we discuss some advantages of the chosen representation for Z in the context of algorithms for constrained optimization.

We emphasize that the procedures given in this note do *not* depend on use of the TQ rather than other factorizations. We have chosen to specify the computations in terms of the TQ factorization in order to describe precisely the implementations mentioned above. Use of the TQ factorization leads to no increase in complexity of the implementation, although it does slightly complicate the exposition.

2. Representation and computation of Z

The algorithm for computing Z is essentially a method for *updating* the TQ factorization (1) when the matrix Q is stored *explicitly* (see Gill *et al.*, 1982).

We assume that the m rows of $A(x)$ can be partitioned into two groups: the *first* m_L rows (denoted by A_L) are constant, and the last m_N rows (denoted by $A_N(x)$) vary with x. (This partition corresponds to the case when $A(x)$ is the Jacobian of a mixture of linear and nonlinear *constraints.*) Thus, A and T in (1) have the forms

$$A = \begin{pmatrix} A_L \\ A_N \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} 0 & T_L \\ T_N & W \end{pmatrix},$$

where T_L and T_N are reverse-triangular.

Consider now a different matrix $\bar{A}$ that is *not* necessarily close to A. Observe that the old Q reduces the first m_L rows of $\bar{A}$ to reverse-triangular form, i.e.

$$\bar{A}Q = \begin{pmatrix} A_L \\ \bar{A}_N \end{pmatrix} Q = \begin{pmatrix} 0 & T_L \\ S & \bar{W} \end{pmatrix}, \tag{2}$$

where S is $m_N \times (n - m_L)$. In order to obtain the new TQ factorization, S is triangularized by a sequence of m_N transformations. The procedure can also be viewed as applying m_N *updates* to the TQ factorization of A_L while the rows of $\bar{A}_N$ are added one at a time.

The triangularization of S is accomplished in a standard fashion by applying a sequence of elementary transformations to its rows. As usual, each step may be viewed as transforming the "first" row of a successively smaller matrix. To illustrate the procedure, consider the definition of the i-th transformation ($i = 1, \ldots, m_N$). At this stage, rows 1 through $i - 1$ of S have already been triangularized, and the corresponding $i - 1$ transformations have been applied to the rows of S. Let j denote the index $m_L + i - 1$ (j is the number of rows of $\bar{A}$ already in reverse-triangular form). Let $\bar{s}_i^T$ denote the "reduced" vector to be triangularized (i.e., the first $n - j$ components of row i), and assume that $\bar{s}_i^T$ is given by

$$\bar{s}_i^T = (\, s_i^T \quad \sigma_i \,), \tag{3}$$

where s_i^T is a vector with $n - j - 1$ components and σ_i is the i-th reverse-diagonal element. The triangularization of $\bar{s}_i^T$ is achieved by an elementary matrix P_i of the form

$$P_i = I - \frac{1}{\beta_i} u_i v_i^T. \tag{4}$$

We give two alternative definitions of P_i. First, P_i can be taken as a *Householder matrix*, in which case

$$u_i = v_i = \begin{pmatrix} s_i \\ \sigma_i + \operatorname{sign}(\sigma_i)\|\bar{s}_i\|_2 \end{pmatrix} \quad \text{and} \quad \beta_i = \frac{1}{2}\|u_i\|_2^2. \tag{5}$$

(For a detailed discussion of Householder matrices, see Stewart, 1973.) If regarded as a function of the elements of $\bar{s}_i$, the matrix P_i is *discontinuous* at any point where σ_i vanishes.

The second option is to define P_i as an *elementary* but non-orthogonal matrix (as in Gaussian elimination). In this case,

$$u_i = e_{n-j}, \quad v_i = \begin{pmatrix} s_i \\ 0 \end{pmatrix}, \quad \text{and} \quad \beta_i = \sigma_i, \tag{6}$$

where e_{n-j} is the $(n - j)$-th coordinate vector. In constructing a stabilized elementary transformation, requirements of numerical stability forbid the use of a pivot element σ_i that is "small" relative to the elements of s_i. Hence, if σ_i is too small, a column interchange must be performed, which introduces a discontinuity in P_i as a function of $\bar{s}_i$.

We now place the triangularization of S within the framework of obtaining the TQ factorization of $\bar{A}$. The successively smaller matrices $\{P_i\}$ defined above are embedded in a sequence of $n \times n$ matrices $\{\tilde{P}_i\}$, $i = 1, \ldots, m_N$, defined by

$$\tilde{P}_i = \begin{pmatrix} P_i & 0 \\ 0 & I_j \end{pmatrix}. \tag{7}$$

A matrix $\tilde{Q}_i$ that triangularizes the first $j+1$ rows of $\bar{A}$ is given by

$$\tilde{Q}_i = Q\tilde{P}_1 \cdots \tilde{P}_i = \tilde{Q}_{i-1}\tilde{P}_i, \tag{8}$$

where $\tilde{Q}_0 = Q$. The final result (after m_N updates) is

$$\bar{A}\tilde{Q}_{m_N} = \begin{pmatrix} 0 & 0 & T_L \\ 0 & \bar{T}_N & \bar{W} \end{pmatrix},$$

so that $\bar{Q}$ is defined by

$$\bar{Q} = \tilde{Q}_{m_N} = Q\tilde{P}_1 \cdots \tilde{P}_{m_N}. \tag{9}$$

The total number of operations required to obtain $\bar{T}$ and $\bar{Q}$ includes the following: $n^2 m_N$ to apply the transformations to S; $\frac{2}{3}m_N^3 + m_N^2(n-m)$ to perform the reduction with Householder transformations ($\frac{1}{3}m_N^3 + \frac{1}{2}m_N^2(n-m)$ with elementary matrices); $m_N^2 + 2m_N(n-m)$ to transform Q with Householder matrices ($\frac{1}{2}m_N^2 + m_N(n-m)$ with elementary matrices).

3. Continuity of Z

In this section we sketch a proof that Z as computed by the method of Section 2 is continuous in a neighborhood of a point $\hat{x}$ where $A(\hat{x})$ has full rank.

Let A_N denote $A_N(\hat{x})$, and $\bar{A}_N$ denote $A_N(\hat{x}+\delta x)$, where $\hat{x}+\delta x$ is in the ball $\mathcal{B}$. By assumption, A is continuous, and hence

$$\bar{A}_N = A_N + \delta A, \tag{10}$$

where δA can be made as small as desired by restricting the size of $\mathcal{B}$.

We assume that the TQ factorization (1) of A is given, and that Q is bounded (if Q is orthogonal, the latter assumption is satisfied automatically). It follows from (10) and the boundedness of Q that the result of applying Q to $\bar{A}_N$ is a matrix that is "almost" in reverse-triangular form. Thus, the first $n - m_L$ columns of $\bar{A}_N Q$ may be written as

$$\tilde{T} = (\, 0 \quad T_N \,) + \delta T, \tag{11}$$

where the elements of δT can be made as small as desired by appropriate definition of $\mathcal{B}$. Let $\tilde{t}_i^T$ denote the i-th row of $\tilde{T}$. We now consider the continuity properties of the matrices $\{P_i\}$ that will be applied to restore $\tilde{T}$ to reverse-triangular form.

The continuity argument is essentially an inductive demonstration that the "almost" reverse-triangular form of $\tilde{T}$ is not destroyed by application of the transformations $\{P_i\}$. To why, consider the special character of the Householder transformation that reduces the first row of $\tilde{T}$. It follows from (11) that, when $\mathcal{B}$ is sufficiently small, the first row of $\tilde{T}$ is of the form

$$\tilde{t}_1^T = (\, \delta^T \quad \tau + \epsilon \,), \tag{12}$$

where δ is a vector with $n - m_L - 1$ components, τ is the first reverse-diagonal element of T_N, and $\|\delta\|$ and $|\epsilon|$ are small. From (5), the Householder vector v_1 is given by

$$v_1 = \begin{pmatrix} \delta \\ \tau + \epsilon + \operatorname{sign}(\tau + \epsilon)\|\tilde{t}_1\| \end{pmatrix}. \tag{13}$$

Since A has full rank, τ is non-zero, and hence the choice of sign in (13) will be constant in a neighborhood of $\hat{x}$. Therefore, v_1 (and P_1) will converge to a limit as $\|\delta x\|$ approaches zero.

Now consider the effect of P_1 on the remaining rows of $\tilde{T}$. It follows from (4) that the i-th row transformed by P_1 is given by

$$\tilde{t}_i^T - \gamma_i v_1^T, \quad \text{where} \quad \gamma_i = \tilde{t}_i^T v_1/\beta_i. \tag{14}$$

Since (13) shows that the first $n - m_L - 1$ components of v_1 are small, (14) indicates that the effect of transforming $\tilde{T}$ by P_1 is merely to add *small* perturbations to the first $n - m_L - 1$ components of rows 2 through m_N. Thus, $\tilde{T}P_1$ remains in "almost" reverse-triangular form and a relationship like (12) holds for the next row to be triangularized. Using a similar argument, we can show that each vector v_i is continuous in a neighborhood of $\hat{x}$. (In the limit, (13) shows that v_i becomes a multiple of e_{n-j}.)

An analogous argument can be made for continuity of each P_i defined by a stabilized elementary transformation. In this case, it can be shown that for a sufficiently small neighborhood, the pivot elements do not become too small because they are small perturbations of the reverse-diagonal elements of T. Using (6) and (12), we also see that the vector v_i converges to zero as $\|\delta x\|$ approaches zero.

Having shown that the matrices $\{P_i\}$ are continuous in a neighborhood of $\hat{x}$, we must now consider the continuity of Z. Recall from (9) that

$$\bar{Q} = Q\tilde{P}_1 \cdots \tilde{P}_{m_N}.$$

In the stabilized elementary case, it is easy to see that each $\tilde{P}_i$ converges to the identity matrix as $\|\delta x\|$ approaches zero. Hence, Q (and Z) are continuous.

Unfortunately, the same simple argument does not apply in the Householder case, since it is well known that the identity matrix is not a Householder matrix. However, because the first $n - j - 1$ components of v_i are converging to zero, the application of $\tilde{P}_i$ causes *changes that converge to zero* in columns 1 through $n - j - 1$, where j assumes the values $m_L, \ldots, m - 1$. Since Z comprises the first $n - m$ columns of Q, this implies that the changes in the columns of Z converge to zero as $\|\delta x\|$ approaches zero. Thus, Z *is* continuous in the Householder case, although Q is not.

Despite the continuity of Z, it may be considered undesirable for the remaining columns of Q (and the reverse-diagonals of T) to undergo the changes in sign that result when a Householder reduction is applied to a nearly triangular matrix. If so, the difficulty can be avoided by systematically changing the relevant sign after each transformation is applied — for example, by applying the composite transformation $P_i D_i$, where the $(n - j)$-dimensional diagonal matrix D_i is defined by

$$D_i = \operatorname{diag}(1, \ldots, 1, -1).$$

With this technique, T and Q would become continuous in a neighborhood of $\hat{x}$. Coleman and Sorensen (1982) suggest an alternative procedure that avoids alternating signs in the diagonals of T and the last m columns of Q.

4. Further properties

The approach of Section 2 is not the only possible way to obtain a representation of Z. The most obvious alternative is to apply a "standard" Householder triangularization. In this case, the Householder vectors are stored in *compact form* along with T in an $n \times m$ matrix (plus an m-vector) (see, e.g., Stewart, 1973), so that no explicit matrix Q is stored. (Products of the form Qv or $Q^T v$ are obtained by applying the stored transformations to v.)

Assuming that the m_L transformations corresponding to constant rows of A are retained, the matrices $\bar{T}$ and $\bar{Q}$ of Section 2 can be computed using the standard Householder procedure in $m_N(2nm_L - m_L^2)$ operations to apply the m_L fixed transformations to A_N, and $\frac{2}{3}m_N^3 + m_N^2(n-m)$ operations to produce the desired reverse-triangular form. If the explicit matrix Q is needed, it can be formed (in $2nm(n-m) + \frac{2}{3}m^3$ operations) by multiplying the transformations together in reverse order. Thus, the standard Householder approach requires less storage and work than the method of Section 2.

Nonetheless, an explicit representation of Q has certain advantages because of its convenience and flexibility under more general circumstances that occur in constrained optimization methods. *Our particular concern is to perform updates efficiently and uniformly when simple bounds and linear constraints are treated separately from nonlinear constraints* — as in the methods described in Gill *et al.* (1983b). Special treatment of these constraints involves not only exploiting the presence of constant rows in A (which can be done with the standard Householder procedure, as noted above), but also performing operations on A other than simply adding rows. In particular, a given matrix A will often be subject to the *deletion of rows* and the *addition or deletion of columns* — say, in computing Lagrange multiplier estimates for different versions of the working set at the same point. With Q stored explicitly, all four types of updates can be performed in an efficient uniform manner (see Gill *et al.*, 1982, for details). However, it is not possible to perform the latter three updates within the framework of a compact representation of Q.

Representing each *change* to Q as in (8) rather than recomputing Q from a compact representation is useful in other situations – for example, modifying a projected Hessian (or approximate Hessian) to reflect the new subspace of projection. Finally, an explicit representation of Q is convenient if finite-differences are to be taken along the columns of Z.

References

Coleman, T. F. and Sorensen, D. C. (1982). A note on the computation of an orthogonal basis for the null space of a matrix, Report, Computer Science Department, Cornell University, Ithaca, New York. To appear in *Math. Prog.*

Gill, P. E., Murray, W., Saunders, M. A., and Wright, M. H. (1982). Procedures for optimization problems with a mixture of bounds and general linear constraints, Report SOL 82-6, Department of Operations Research, Stanford University, California. To appear in *ACM Trans. Math. Software.*

Gill, P. E., Murray, W., Saunders, M. A., and Wright, M. H. (1983a). User's guide to SOL/QPSOL, Report SOL 83-7, Department of Operations Research, Stanford University, California.

Gill, P. E., Murray, W., Saunders, M. A., and Wright, M. H. (1983b). User's guide to SOL/NPSOL, Report SOL 83-11, Department of Operations Research, Stanford University, California.

Stewart, G. W. (1973). *Introduction to Matrix Computations*, Academic Press, London and New York.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER SOL 83-19	2. GOVT ACCESSION NO. AD-A136 419	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ON THE REPRESENTATION OF A BASIS FOR THE NULL SPACE		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Philip E. Gill, Walter Murray, Michael A. Saunders and Margaret H. Wright		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0267 DAAG29-81-K-0156
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Operations Research - SOL Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-047-143
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research - Dept. of the Navy 800 N. Quincy Street Arlington, VA 22217		12. REPORT DATE November 1983
		13. NUMBER OF PAGES 5 pp.
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

THE VIEW, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF THE AUTHOR(S) AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER DOCUMENTATION.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonlinear optimization
Matrix factorization
Null-space continuity

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

ABSTRACT ON REVERSE SIDE

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

SOL 83-19, "ON THE REPRESENTATION OF A BASIS FOR THE NULL SPACE," by Philip E. Gill, Walter Murray, Michael A. Saunders and Margaret H. Wright.

Given rectangular matrix $A(x)$ that depends on the independent variables x, many constrained optimization methods involve computations with $Z(x)$, a matrix whose columns form a basis for the null space of $A(x)$. When A is evaluated at a given point, it is well known that a suitable Z (satisfying $AZ = 0$) can be obtained from standard matrix factorizations. However, Coleman and Sorensen have recently shown that standard orthogonal factorization methods may produce orthogonal bases that do not vary continuously with x; they also suggest several techniques for adapting standard factorization schemes so as to ensure continuity of Z in the neighborhood of a given point.

In this note, we discuss several aspects of the representation of a basis for the null space. We describe how an explicit matrix Z can be obtained at any point using a method for updating a factorization with either Householder or stabilized elementary transformations. Under a mild non-singularity assumption, the elements of Z are continuous functions of x. We also show that the chosen form of Z is convenient and efficient when implementing certain methods for nonlinearly constrained optimization.