1 of 1

**SYSTEMS OPTIMIZATION LABORATORY
DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-4022**

**On the Numerical Stability of
Quasi-Definite Systems**

by
Philip E. Gill*, Michael A. Saunders[†] and Joseph R. Shinnerl*

TECHNICAL REPORT SOL 93-4

June 1993

*Department of Mathematics, University of California at San Diego, La Jolla, CA 92093-0112. Partially supported by National Science Foundation Grant DDM-9204547.

[†]Systems Optimization Laboratory, Department of Operations Research, Stanford University, Stanford, CA 94305-4022. Partially supported by Department of Energy Grant DE-FG03-92ER25117, National Science Foundation Grant DDM-9204208, and Office of Naval Research Grant N00014-90-J-1242.

Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the above sponsors.

Also listed as Operations Research Department Technical Report 93-9. Reproduction in whole or in part is permitted for any purposes of the United States Government. This document has been approved for public release and sale; its distribution is unlimited.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

On the Numerical Stability of Quasi-Definite Systems

Philip E. GILL*, Michael A. SAUNDERS† and Joseph R. SHINNERL*

Technical Report SOL 93-4‡

June 1993

Abstract

We discuss the solution of sparse linear equations $Kd = r$, where K is a symmetric and specially structured indefinite matrix that often arises in numerical optimization. For such K , the indefinite factorization $K = LDL^T$ is known to exist in exact arithmetic with 1×1 pivots and no row or column interchanges. It is shown that the stability of the LDL^T factorization of this matrix is naturally connected with the stability of the LDM^T factorization of a closely related unsymmetric positive-definite matrix. Conditions are given that allow the stable numerical solution of this system by Gaussian elimination without row and column interchanges.

Keywords: indefinite systems, symmetric quasi-definite (sqd) systems, unsymmetric positive-definite systems, backward stability, condition number.

1. Introduction

We define a matrix K to be *symmetric quasi-definite*, or “sqd”, if there exists a permutation matrix Π that reorders K to the form

$$\Pi K \Pi^T = \begin{pmatrix} H & A^T \\ A & -G \end{pmatrix}, \quad (1.1)$$

where H and G are $n \times n$ and $m \times m$ symmetric positive definite. It is easily shown that such K are indefinite and nonsingular. Vanderbei [Van92] has shown that these matrices are *strongly factorizable*; i.e., given an sqd matrix K , for *every* permutation matrix P there exist a diagonal matrix D and a unit lower-triangular matrix L such that

$$PKP^T = LDL^T. \quad (1.2)$$

*Department of Mathematics, University of California at San Diego, La Jolla, CA 92093-0112.

†Systems Optimization Laboratory, Department of Operations Research, Stanford University, Stanford, CA 94305-4022.

‡The material contained in this report is based upon research partially supported by Department of Energy Grant DE-FG03-92ER25117, National Science Foundation Grants DDM-9204208 and DDM-9204547, and Office of Naval Research Grant N00014-90-J-1242.

The mere existence of such L and D , however, does not mean they can be computed accurately in finite-precision arithmetic. A well-known example appearing in [Van92] is

$$K = \begin{pmatrix} -\epsilon & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{\epsilon} & 1 \end{pmatrix} \begin{pmatrix} -\epsilon & 0 \\ 0 & 1 + \frac{1}{\epsilon} \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{\epsilon} \\ 0 & 1 \end{pmatrix}.$$

When ϵ is sufficiently small, this factorization cannot be performed accurately in finite precision. If a symmetric interchange is applied to K , however, the numerical LDL^T factorization proceeds without difficulty.

Strong factorizability is particularly attractive when K is large and sparse and a direct factorization method is used to solve the linear system of equations

$$Kd = r. \quad (1.3)$$

In this context, symmetric permutations of K are employed to reduce fill-in during the factorization. That is, one would like to solve system (1.3) by Gaussian elimination, interchanging rows and columns symmetrically in such a way that as many elements as possible are zero at the final step. For these interchanges to be most effective, however, they must not be compromised by other permutations used to maintain numerical stability of the factorization.¹ We are therefore interested in matrix factorizations requiring few, if any, row and column interchanges for numerical stability.

1.1. Notation

The symbols H , G , and A are reserved for the block components of the arbitrary sqd matrix K , with one important exception. The letter A is also used to denote an arbitrary square nonsingular matrix whose triangular factorization $A = LU = LDM^T$ exists in exact arithmetic without row or column interchanges. Here U is upper triangular, and L and M are lower triangular with every diagonal element equal to 1. The matrix D is diagonal with $DM^T \equiv U$ and therefore $d_{ii} = u_{ii}$ for all i . Although such a factorization does not exist for all nonsingular A , when it does exist, the factors are unique.

2. The Connection to the Unsymmetric Positive-Definite Case

We seek conditions on K that allow, for *every* permutation matrix P , the *stable finite-precision computation* of factors L and D such that $PKP^T = LDL^T$. As we are concerned with arbitrary symmetric permutations of K , there is no loss of generality in assuming $\Pi = I$ in (1.1). With this convention,

$$K = \begin{pmatrix} H & A^T \\ A & -G \end{pmatrix} \equiv \bar{K} \bar{I}, \quad (2.1)$$

¹We emphasize the distinction between *sparsity permutations*, designed to reduce fill-in in the symbolic factorization, and *stability permutations*, designed to maintain accuracy in the numerical computation. We wish to avoid the use of the latter if possible, because they interfere with the former and therefore increase the computation time.

where

$$\bar{K} \equiv \begin{pmatrix} H & -A^T \\ A & G \end{pmatrix} \quad \text{and} \quad \bar{I} \equiv \begin{pmatrix} I_n & \\ & -I_m \end{pmatrix},$$

where I_n and I_m denote the identity matrices of order n and m . The matrix $\bar{K}$ is *unsymmetric positive definite*.² The main idea of this paper is that (2.1) can be used to characterize the stability of algorithms for symmetric quasi-definite matrices in terms of the stability of algorithms for unsymmetric positive-definite matrices.

With $\bar{I}$ as above, let $\tilde{I} \equiv P\bar{I}P^T$ for some permutation P . This is a symmetric permutation of the diagonal matrix $\bar{I}$ and is therefore also diagonal with diagonal entries 1 and -1 . Thus in any product of the form $\tilde{A} = A\tilde{I}$, the matrix $\tilde{A}$ is equal to A with some of its columns scaled by -1 . Now for every permutation P ,

$$PKP^T = P\bar{K}\bar{I}P^T = P\bar{K}P^T(P\bar{I}P^T) = (P\bar{K}P^T)\tilde{I}, \quad (2.2)$$

so that $P\bar{K}P^T = PKP^T\tilde{I}$. Hence,

$$PKP^T = LDL^T \quad \text{if and only if} \quad P\bar{K}P^T = LDL^T\tilde{I} = L(D\tilde{I})\tilde{I}L^T\tilde{I} = L\tilde{D}M^T, \quad (2.3)$$

where $\tilde{D} \equiv D\tilde{I}$ and $M \equiv \tilde{I}L\tilde{I}$. The matrices D and $\tilde{D}$ are diagonal, and L and M are unit lower triangular as required. Relations (2.1) and (2.3) can be construed as an alternative proof of Vanderbei's theorem on the strong factorizability of symmetric quasi-definite matrices. For if K is sqd, then $\bar{K}$ and hence $P\bar{K}P^T$ are positive definite; therefore the LU factorization of $P\bar{K}P^T$ exists; hence, by (2.3), the LDL^T factorization of PKP^T exists as well.

We now show how (2.2) and (2.3) can be used to analyze the stability of the LDL^T factorization for symmetric quasi-definite matrices.

Lemma 2.1. *Suppose $A = LU$, where L is unit lower triangular and U is upper triangular. If $\hat{L}$ and $\hat{U}$ are the computed factors for A obtained by Gaussian elimination without interchanges, then $\hat{L}$ and $\hat{U}\tilde{I}$ are the computed factors of $A\tilde{I}$.*

Proof. Consider the application of Gaussian elimination to the matrix A in stages, the k th stage being the elimination of all nonzero elements in the first column of the $(n - k + 1) \times (n - k + 1)$ remaining matrix. It is easy to show inductively that if the signs of the entries in the k th column of A are changed, then the signs of the entries in the k th column of U will also be changed, but all other columns of U will be unaffected. ■

From this lemma and (2.2), if $\hat{L}$ and $\hat{D}$ are the computed factors of PKP^T , then $\hat{L}$, $\hat{D}\tilde{I}$, and $\hat{M}^T = \tilde{I}\hat{L}^T\tilde{I}$ are the computed LU factors of $P\bar{K}P^T$. Hence any conditions that ensure stability for the factorization $P\bar{K}P^T = LDM^T$ will also ensure stability for $PKP^T = LDL^T$. In particular, it is safe to factor the quasi-definite matrix PKP^T *without stability interchanges* if and only if it is safe to factor the unsymmetric positive-definite matrix $P\bar{K}P^T$ *without stability interchanges*.

²Positive definiteness of $\bar{K}$ follows from the positive definiteness of its symmetric part, $\frac{1}{2}(\bar{K} + \bar{K}^T) = \text{diag}(H, G)$.

3. When Stability Interchanges are Unnecessary

Throughout this section, we assume that C is an unsymmetric positive-definite matrix. Let $T \equiv (C + C^T)/2$ and $S \equiv (C - C^T)/2$. Then it is safe to factor C without stability interchanges if

- (i) S is not too large compared to T , and
- (ii) T is not too ill-conditioned.

That this is so follows from the results of Golub and Van Loan, [GV79] and [GV89], summarized in the next section.

3.1. Theorems of Golub and Van Loan

Let C be factored $C = LDM^T$, where L and M are unit lower-triangular and D is diagonal. In the backward error analysis of Gaussian elimination, it is shown that the computed solution $\hat{x}$ to the system $Cx = r$ is the exact solution of the perturbed system $(C + \Delta C)\hat{x} = r$, where the size of ΔC is bounded by an expression involving the sizes of the computed factors of C ; say, $\hat{L}$, $\hat{D}$, and $\hat{M}^T$ (cf. (3.2) below). Algorithms that produce $\hat{L}$, $\hat{D}$, and $\hat{M}^T$ of sufficiently bounded size are therefore considered stable. For general C , stability interchanges are necessary to prevent these factors from having large elements.³ For positive-definite C , however, the following theorems can be used with the assumption

$$\|\hat{L}\|\|\hat{D}\|\|\hat{L}^T\|_F \leq \gamma \|L\|\|D\|\|L^T\|_F, \quad \gamma \text{ scalar}, \quad (3.1)$$

to obtain a satisfactory bound on the sizes of the computed factors without stability interchanges.⁴

Theorem 3.1. ([GV79], p. 88) *Let $C \in \mathbb{R}^{n \times n}$ be positive definite and set $T = (C + C^T)/2$ and $S = (C - C^T)/2$. If $C = LDM^T$, then*

$$\|L\|\|D\|\|M^T\|_F \leq n \left(\|T\|_2 + \|ST^{-1}S\|_2 \right).$$

■

In the following theorem, the symbol $\mathbf{u}$ denotes the unit round-off, and A denotes any $n \times n$ matrix whose LDM^T factorization exists.

Theorem 3.2. ([GV89], p. 136, eqn. (4.1.3)) *Let $\hat{L}$, $\hat{D}$, and $\hat{M}$ denote the computed factors of A . Let $\hat{x}$ denote the computed solution to the system $Ax = b$, obtained by solving the triangular systems $\hat{L}y = b$ and $\hat{D}\hat{M}^T x = \hat{y}$ in sequence by the usual methods of forward and backward substitution (cf. [GV89], p. 97, Algorithm 3.2.3). Then $(A + \Delta A)\hat{x} = b$, with*

$$\|\Delta A\| \leq n\mathbf{u} \left(3\|A\| + 5\|\hat{L}\|\|\hat{D}\|\|\hat{M}^T\| \right) + \mathcal{O}(\mathbf{u}^2). \quad (3.2)$$

■

³When C is not known to be positive definite, interchanges are also necessary to ensure the existence of the factors.

⁴When applied to vectors or matrices, the symbols $|\cdot|$ and $\leq$ are to be interpreted componentwise.

By (3.1) and these theorems, it follows that the computed solution $\hat{x}$ to the positive-definite system $Cx = r$ satisfies $(C + \Delta C)\hat{x} = r$, with

$$\|\Delta C\|_2 \leq \|\Delta C\|_F \leq u \left(3n\|C\|_F + 5\gamma n^2 (\|T\|_2 + \|ST^{-1}S\|_2) \right) + \mathcal{O}(u^2). \quad (3.3)$$

Note that $\|T\|_2 \leq \|C\|_2$, and thus

$$\|T\|_2 + \|ST^{-1}S\|_2 \leq \left(1 + \frac{\|ST^{-1}S\|_2}{\|C\|_2} \right) \|C\|_2. \quad (3.4)$$

Hence, for positive-definite C , the computation of L , D , and M is stable if the factor

$$\omega(C) \equiv \frac{\|ST^{-1}S\|_2}{\|C\|_2} \quad (3.5)$$

is not too large.

Because it may not always be clear how the structure of the matrix $ST^{-1}S$ depends on the structure of the original matrix C , we observe that

$$\omega(C) \leq \theta(C) \equiv \left(\frac{\|S\|_2}{\|T\|_2} \right)^2 \kappa_2(T), \quad (3.6)$$

where $\kappa_2(T) \equiv \|T\|_2\|T^{-1}\|_2$. While the bound $\theta(C)$ will often be somewhat larger than $\omega(C)$, its straightforward dependence on the symmetric and skew-symmetric parts of C makes it easier to estimate.

3.2. Application to Quasi-Definite Matrices

For our purposes, the role of C is played by $P\bar{K}P^T$. As it is easily shown that ω and θ are invariant under symmetric permutations of their arguments, we assume $C = \bar{K}$. In this case,

$$S = \frac{1}{2}(\bar{K} - \bar{K}^T) = \begin{pmatrix} 0 & -A^T \\ A & 0 \end{pmatrix} \quad \text{and} \quad T = \frac{1}{2}(\bar{K} + \bar{K}^T) = \begin{pmatrix} H & \\ & G \end{pmatrix}.$$

Therefore,

$$ST^{-1}S = \begin{pmatrix} -A^T G^{-1} A & \\ & -A H^{-1} A^T \end{pmatrix},$$

and stability of the factorization can be guaranteed if $\omega(\bar{K}) \equiv \|ST^{-1}S\|_2/\|\bar{K}\|_2$ is not too large. Note that a single relatively dense column of A will destroy the sparsity of the matrices $A^T G^{-1} A$ and $A H^{-1} A^T$ and will make the computation of $\|A^T G^{-1} A\|_2$ and $\|A H^{-1} A^T\|_2$ relatively expensive. However, the alternative bound $\theta(\bar{K})$ is readily computed in terms of the block components A , H , H^{-1} , G , and G^{-1} . For example, suppose $\|H\|_2 \geq \|G\|_2$ and $\|G^{-1}\|_2 \geq \|H^{-1}\|_2$. Then $\|T\|_2 = \|H\|_2$, and

$$\theta(\bar{K}) = \frac{\|A\|_2^2}{\|H\|_2} \|G^{-1}\|_2. \quad (3.7)$$

The factorization $P\bar{K}P^T = LDM^T$ will be numerically stable for every permutation P when $\omega(\bar{K})$ or $\theta(\bar{K})$ is not too large. This means that (i) the norm of the skew-symmetric part of $\bar{K}$, namely $\|A\|_2$, must not be too large compared to $\|H\|_2$ and $\|G\|_2$; and (ii) $T \equiv \text{diag}(H, G)$ must not be too ill-conditioned. From Section 2, the stability of the LDL^T factorization of PKP^T for all permutations P is also ensured by these conditions.

4. The Condition Number of a Quasi-Definite System

To assess the accuracy of computed solutions to the quasi-definite system $Kd = r$, we must consider both the backward stability of factorization (1.2) and the forward sensitivity of d to perturbations in K . That is, given that our computed solution $\hat{d}$ satisfies the perturbed system

$$(K + \Delta K)\hat{d} = r, \quad (4.1)$$

how close is $\hat{d}$ to d , the true solution? The answer is that the bound on $\|d - \hat{d}\|/\|d\|$ is proportional to the product of $\|\Delta K\|/\|K\|$ and $\kappa(K)$. For general matrices K , the relative perturbation $\|\Delta K\|/\|K\|$ cannot be suitably bounded without the use of stability interchanges. When K is sqd, however, (3.3) and (3.4) give a bound on this perturbation that is essentially proportional to $1 + \omega(\bar{K})$.

Suppose therefore that $\hat{d}$ is the computed solution of $Kd = r$, and, as before, let $\kappa_2(K) = \|K\|_2\|K^{-1}\|_2$. Note that $\kappa_2(K) = \kappa_2(\bar{K}) = \kappa_2(PK P^T)$. The following inequality is adapted from [GVL79], p. 92, where it appeared for unsymmetric positive-definite systems. Written for the quasi-definite context under assumption (3.1) for the computed factors of $\bar{K}$, it reads

$$\frac{\|d - \hat{d}\|_2}{\|d\|_2} \leq \mathbf{u} n \gamma c_n (1 + \omega(\bar{K})) \kappa_2(K), \quad (4.2)$$

where $\mathbf{u}$ is the unit round-off and c_n depends linearly on n .

Under assumption (3.1), then, arbitrary symmetric permutations (i.e., those that reduce fill-in) of the system $Kd = r$ can be solved stably without further permutations as long as the product

$$\phi(K) \equiv (1 + \omega(\bar{K})) \kappa_2(K) \quad (4.3)$$

is not too large. We therefore interpret ϕ to be the condition number of an sqd system when stability interchanges are not allowed. In algorithms where sequences of sqd systems are solved, techniques that either reduce ϕ or prevent or delay its increase will, by postponing the need for stability permutations and hence allowing the unhampered use of sparsity permutations, decrease the total computation time for solving the sqd system $Kd = r$.

It should be noted that the reduction of ϕ is sufficient, but not necessary, for ensuring the accurate solution of $Kd = r$ without interchanging rows and columns for stability. Indeed, Golub and Van Loan [GV79] exhibit a family of unsymmetric positive-definite systems for which ω increases without bound but whose computed

solutions remain accurate without the use of stability interchanges. Their example suggests that in special cases it may be possible to refine the above results further to obtain a sharper bound.

5. An Application in Numerical Optimization

Consider the problem of minimizing a smooth function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ subject to the smooth equality constraints $c : \mathbb{R}^n \rightarrow \mathbb{R}^m$:

$$\begin{aligned} & \underset{x \in \mathbb{R}^n}{\text{minimize}} && f(x) \\ & \text{subject to} && c(x) = 0. \end{aligned} \quad (5.1)$$

The optimality conditions for this problem are

$$\begin{aligned} g(x) - A(x)^T \pi &= 0 \\ c(x) &= 0, \end{aligned} \quad (5.2)$$

where $g(x)$ denotes the gradient of f ; $A(x)$ is the Jacobian matrix of $c(x)$, $A_{ij}(x) \equiv \partial c_i / \partial x_j$; and π is the vector of Lagrange multipliers. Applying Newton's method for nonlinear equations to system (5.2) produces the system

$$\begin{pmatrix} W(x) & A(x)^T \\ A(x) & 0 \end{pmatrix} \begin{pmatrix} -\Delta x \\ \Delta \pi \end{pmatrix} = \begin{pmatrix} g(x) - A(x)^T \pi \\ 0 \end{pmatrix}, \quad (5.3)$$

where $W(x) \equiv \nabla^2 f(x) - \sum \pi_i \nabla^2 c_i(x)$ approximates the Hessian matrix of the Lagrangian function $f(x) - c(x)^T \pi$.

Evaluating the components of (5.3) at a particular iterate x_k yields the familiar *KKT system* of linear equations. In algorithms where sequences of KKT systems are solved, the *diagonally perturbed KKT system*

$$K_\delta \begin{pmatrix} -\Delta x \\ \Delta \pi \end{pmatrix} = \begin{pmatrix} w \\ r \end{pmatrix}, \quad \text{where} \quad K_\delta \equiv \begin{pmatrix} W & A^T \\ A & -\delta^2 I \end{pmatrix}, \quad (5.4)$$

is frequently used as an alternative for reasons of computational stability and efficiency. Here δ^2 is a small regularization parameter typically held fixed at some value near 10^{-8} , while W and A change as the sequence $\{x_k\}$ converges to a stationary point x^* . (For further details, see [GMPS91].)

When W is taken as a positive-definite approximation to $W(x_k)$, K_δ is sqd, and the analysis of the previous sections can be used to accelerate the numerical solution of (5.4), provided specialized techniques are developed for the treatment of asymptotic ill-conditioning in K_δ . When problem (5.1) has special structure, δ^2 may be dispensed with by transforming (5.4) to an sqd system of smaller dimension. For example, Vanderbei [Van92] has discussed a quasi-definite formulation of the solution of (5.1) when f and c are derived from an interior-point method for linear programming. Further such techniques are the subject of a forthcoming paper [GSS93].

6. Conclusions

By means of a column scaling (2.1) and careful record-keeping of symmetric permutations (2.3), we have identified a numerical equivalence between symmetric quasi-definite linear systems of equations and a special class of unsymmetric positive-definite systems. Using the existing results of Golub and Van Loan for unsymmetric positive-definite systems, we have established an effective condition number (4.3) for the finite-precision numerical solution of the sqd system $Kd = r$ *without* the use of row and column interchanges. When these conditions are satisfied, they can be used in conjunction with direct methods for the factorization of sparse matrices to accelerate the numerical solution of sqd systems.

References

- [GMPS91] P. E. Gill, W. Murray, D. B. Ponceleón, and M. A. Saunders. Solving reduced KKT systems in barrier methods for linear and quadratic programming. Report SOL 91-7, Department of Operations Research, Stanford University, 1991.
- [GSS93] P. E. Gill, M. A. Saunders, and J. R. Skinnerl. On the numerical stability of perturbed KKT systems. To appear, 1993.
- [GV79] G. H. Golub and C. F. Van Loan. Unsymmetric positive definite linear systems. *Linear Algebra and its Applications*, 28, 85–98, 1979.
- [GV89] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, Maryland, second edition, 1989.
- [Van92] R. J. Vanderbei. Symmetric quasi-definite matrices. Report SOR 91-10, Department of Civil Engineering and Operations Research, Princeton University, 1992.

**DATE
FILMED**

10 / 20 / 93

END