Generalized and Sparse Least Squares Problems

ÅKE BJÖRCK
Department of Mathematics
Linköping University
S-581 83 Linköping
Sweden

ABSTRACT. Least squares problems arise frequently in optimization, e.g., in interior point methods. This paper surveys methods for solving least squares problems of non-standard form such as generalized and sparse problems. Algorithms for standard and banded problems are first studied. Methods for solving generalized least squares problems are then surveyed. The special case of weighted problems is treated in detail. Iterative refinement is discussed as a general technique for improving the accuracy of computed solutions. Least squares problems where the solution is constrained by linear equality constraints or quadratic constraints are also treated.

Graph theoretic methods for reordering rows and columns to reduce fill in when solving sparse least squares problems are surveyed. The numerical phase of sparse Cholesky and sparse QR factorization is then discussed. In particular the multifrontal method, which currently is the most efficient implementation, is described.

1. Least Squares Problems

1.1. INTRODUCTION

Let \( A \in \mathbb{R}^{m \times n} \) be a rectangular matrix, and \( b \in \mathbb{R}^m \) a vector. A fundamental computational problem is the linear least squares problem

\[
\min_{x \in \mathbb{R}^n} \|Ax - b\|_2, \quad \mathcal{S} = \{x \in \mathbb{R}^n \mid \|Ax - b\|_2 = \min\}. \tag{1.1}
\]

Least squares problems arise frequently in applications. An important example in optimization is in interior point methods, where the Karush-Kuhn-Tucker optimality condition give rise to a generalized least squares problem, see Wright [54].

Many surveys over computational methods for problem (1.1) exist, see, e.g., Lawson and Hanson [37], Björck [9], [10]. Here we focus on methods for solving least squares problems of non-standard form. In particular we consider in detail methods for weighted and sparse problems.

In Section 1 we first briefly review the method of normal equations and methods based on the QR factorization of \( A \) for the standard least squares problem. We then

consider problems where $A$ is a (rectangular) banded matrix, and show that these can be handled by simple modifications of methods for dense problems.

In Section 2.1 we discuss the method of Paige for generalized least squares problems involving two matrices $A$ and $W$, where $W = B^T B$ is symmetric positive definite matrix. In Section 2.2 we consider methods which rely on factorizing the symmetric indefinite system matrix directly, and discuss its numerical stability. The important special case when $W$ is diagonal, the weighted least squares problem, is treated in Section 2.3. Iterative refinement in fixed precision is an important technique to improve solutions computed by methods which are not backward stable. This is the topic of Section 2.4.

In the simplest constrained least squares problem the solution is required to satisfy a subsystem of equations exactly. In Section 3 we compare three different methods for solving such problems. An important technique for regularizing ill-conditioned least squares problems is to include a quadratic inequality constraint. In Section 4 we give an overview of methods for solving such constrained problems.

Sparse least squares problems are treated in Sections 5 and 6. The symbolic phase of the computation is covered in Section 5 and the numerical phase in Section 6. Graph theoretic tools are introduced in 5.2 and transformation to block triangular form in 5.3. Techniques for reordering column for sparsity are discussed in Section 5.4. The Cholesky and QR factorization both aim at computing a sparse upper triangular factor $R$, and share the same symbolic phase. In the numerical phase the difference is that in the Cholesky factorization one first forms numerically the matrix $A^T A$ whereas in the QR factorization the matrix $A$ is transformed directly after the rows have been preordered. The row sequential method for QR factorization is described in Section 6.1. The currently most efficient implementations of both sparse Cholesky and QR factorizations use a multifrontal approach. This is described for the QR factorization in Section 6.2. Finally, in Section 6.3 methods are studied for updating the solution to a sparse problem when a few dense equations are added.

We have treated generalized problems separately from sparse problems. It should be stressed that in practice the interest is often in generalized and sparse problems. Unfortunately, in many cases efficient algorithms for such problems remain to be developed.

1.2. THE METHOD OF NORMAL EQUATIONS

The set $S$ of all least squares solutions is characterized by $x \in S \iff A^T (b - Ax) = 0$, i.e., the residual $r = b - Ax$ is orthogonal to $\mathcal{R}(A)$. Hence any least squares solution satisfies the normal equations

$$A^T Ax = A^T b. \quad (1.2)$$

The solution $x$ is unique if and only if rank $(A) = n$. In this case $A^T A$ is positive definite, and the Cholesky factorization $A^T A = R^T R$, $R$ upper triangular, diag $(R) >$