INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 13, Number 1, Pages 114–121

QR VERSUS CHOLESKY: A PROBABILISTIC ANALYSIS

MARK LIRA, RAM IYER, A. ALEXANDRE TRINDADE, AND VICTORIA HOWLE

Abstract. Least squares solutions of linear equations $A\,x = b$ are very important for parameter estimation in engineering, applied mathematics, and statistics. There are several methods for their solution including QR decomposition, Cholesky decomposition, singular value decomposition (SVD), and Krylov subspace methods. The latter methods were developed for sparse A matrices that appear in the solution of partial differential equations. The QR (and its variant the RRQR) and the SVD methods are commonly used for dense A matrices that appear in engineering and statistics. Although the Cholesky decomposition is backward stable and known to have the least operational count, several authors recommend the use of QR in applications. In this article, we take a fresh look at least squares problems for dense A matrices with full column rank using numerical experiments guided by recent results from the theory of random matrices. Contrary to currently accepted belief, comparisons of the sensitivity of the Cholesky and QR solutions to random parameter perturbations for various low to moderate condition numbers show no significant difference to within machine precision. Experiments for matrices with artificially high condition numbers reveal that the relative difference in the two solutions is on average only of the order of 10^{-6}. Finally, Cholesky is found to be markedly computationally faster than QR – the mean computational time for QR is between two and four times greater than Cholesky, and the standard deviation in computation times using Cholesky is about a third of that of QR. Our conclusion in this article is that for systems with $Ax = b$ where A has full column rank, if the condition numbers are low or moderate, then the normal equation method with Cholesky decomposition is preferable to QR.

Key words. Least squares problems, QR decomposition, Choleksy decomposition, random matrix, statistics.

1. Introduction

The solution of linear equations of the type $A\,x = b$, where $A \in \mathbb{R}^{m\times n}$ is fundamental to problems in science, engineering, applied mathematics and statistics. However, depending on the area, the problems have different features. For instance, linear PDEs in applied mathematics are characterized by a sparse matrix A with a large value of n (typically at least in the thousands), whereas classical parameter estimation problems in engineering and statistics are characterized by a dense matrix A with moderately large value for n (in the tens or hundreds). Furthermore, problems in engineering tend to be minimum norm and least squares if there is periodicity in the data, or have $m \geq n$ and $\text{rank}(A) = n$. In recent years, the area of smart materials and structures have yielded linear compact operator equations, which upon discretization result in least squares problems of moderately sized A matrix [1, 2, 3, 4].

Methods for the solution of linear equations include QR decomposition, Cholesky decomposition, singular value decomposition (SVD), Krylov subspace and Multigrid methods. Krylov subspace methods such as the generalized minimal residual method (GMRES) [5, 6] and the Lanczos method were developed for sparse A matrices that appear in the solution of partial differential equations [7]. Multigrid methods are useful in solving discretized differential equations [7]. The QR method of Francis [8, 9] and the singular value decomposition (SVD) methods are commonly

Received by the editors July 31, 2015.
2000 *Mathematics Subject Classification.* 65F99.

used for dense A matrices that appear in engineering and statistics [3]. Although the Cholesky decomposition is backward stable [10], several authors recommend the use of QR in applications [14, 15].

Assuming $\text{rank}(A) = n$, the Cholesky method for the solution of $A\,x = b$ involves the formulation of the normal equations $A^T\,A\,x = A^T\,b$, decomposing $A^T\,A = L\,L^T$ where L is a lower triangular matrix, and then solving for x using forward and backward substitutions. The basic QR method involves the solution of $R\,x = Q^T\,b$.

A second class of applications where the Cholesky method might find favor are the minimum norm – least squares problems. Consider a linear system $A\,x = b$, where $A \in \mathbb{R}^{m\times n}$ and $\text{rank}(A) = p < \min\{m, n\}$. We assume that p is unknown. A variant of the QR — the rank revealing QR (RRQR) [11, 12, 13] — may be used to find p and obtain a thin QR decomposition of A. Suppose $A = Q\,R$ where Q is a $m \times p$ matrix with orthonormal columns and R is an upper-triangular $p \times n$ matrix. The normal equation then reduces to $(R\,R^T)\,v = Q^T\,b$ and $x = R^T\,v$.

(i) One method for solving for x, which we refer to as QRC, computes a Cholesky factorization of the reduced normal equations. The matrix $R\,R^T$ is a non-singular $p \times p$ matrix. Therefore, we may compute a Cholesky factorization $L\,L^T = R\,R^T$ and proceed to solve for v using forward and backward substitutions. Once v is found, x is computed.

(ii) Another procedure to solve for x is the complete orthogonal factorization method (COF) [14], in which a QR factorization of R^T is computed. Suppose $R^T = U\,S$, where U is a $n \times p$ matrix with orthonormal columns and S is a nonsingular and upper-triangular $p \times p$ matrix. Then $x = U\,S\,v$ and the normal equation yields $S^T\,S\,v = Q^T\,b$. We may solve for $z = S\,v$ from $S^T\,z = Q^T\,b$ and then find $x = U\,z$.

In [14, 15] one finds a sensitivity analysis of the normal equation method, computing the sensitivity of the system to perturbations. The analysis looks at the upper bounds, which are not indicative of the behavior of the normal equation method for low to moderate conditioned systems. A perturbation analysis for the QR decomposition can be found in [16]. A related analysis is found in [17]. An error analysis of the Cholesky method is done in [18] and for positive semidefinite matrices in [10]. In Trefethen and Bau [15], an artificial example is constructed to show that the QR method should be considered to be superior to the normal equation method. The argument presented is that the normal equation method is susceptible to larger errors in the solution if the condition number κ_2 is at least as large as $1/\sqrt{\epsilon}$, where ϵ is the machine precision. Golub and Van Loan [14] state that the normal equation method is less accurate than a stable QR approach, though when the systems are ill-conditioned with large residuals, both methods are apt to produce comparable inaccurate results, which is a somewhat different statement than that of Trefethen. On the other hand, Higham[10] states that the Cholesky decomposition is one of the most numerically stable of all matrix algorithms, but the normal equation method is guaranteed to be backward stable only for well-conditioned matrices [19]. Trefethen [15] asserts that the SVD method is the only fully stable algorithm for solving rank-deficient problems. For high condition number systems (that is, $1/\sqrt{\epsilon} \le \kappa(A) \le 1/\epsilon$), it is possible for the solution of the normal equations to be highly erroneous for some vectors $\bar{b}$. An example verifying this is presented in Trefethen [15].

From the above discussion, it may be gathered that the authors were very concerned about backward stability *for all* matrices, and one can categorically say that for condition numbers greater than $\frac{1}{\sqrt{\epsilon}}$, the QR method is preferable to the