Preconditioning Newton's Method

Ali Bouaricha, Jorge J. More', and
Zhijun Wu

CRPC-TR98762
May 1998

Center for Research on Parallel Computation
Rice University
6100 South Main Street
CRPC - MS 41
Houston, TX 77005

Submitted August 1998; Also available as Argonne Preprint ANL/MCS-P715-0598

ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue
Argonne, Illinois 60439

PRECONDITIONING NEWTON'S METHOD

Ali Bouaricha, Jorge J. Moré, and Zhijun Wu

Mathematics and Computer Science Division

Preprint ANL/MCS-P715-0598

May 1998

This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Computational and Technology Research, U.S. Department of Energy, under Contract W-31-109-Eng-38, and by the National Science Foundation, through the Center for Research on Parallel Computation, under Cooperative Agreement No. CCR-9120008.

ABSTRACT

The development of algorithms and software for the solution of large-scale optimization problems ...

PRECONDITIONING NEWTON'S METHOD

Ali Bouaricha, Jorge J. Moré, and Zhijun Wu

1 Introduction

Algorithms for the solution of large-scale unconstrained minimization problems include conjugate gradient methods, limited memory variable metric methods, and Newton methods. In this paper we compare and contrast these algorithms, and show that the efficient and reliable solution of large-scale unconstrained minimization problems requires a Newton method. Moreover, we describe a Newton method that has proved to be efficient and reliable on optimization problems from applications.

Conjugate gradient methods for the minimization of a nonlinear function $f : \mathbb{R}^n \mapsto \mathbb{R}$ generate x_{k+1} from x_k by setting

$$x_{k+1} = x_k + \alpha_k d_k,$$

where the scalar α_k is chosen by a line search, and the direction d_k is generated by a recurrence of the form

$$d_k = -\nabla f(x_k) + \beta_k d_{k-1}$$

for some $\beta_k \geq 0$. These methods only require a small number (5 is typical) of vectors of storage, but tend to require a large number of iterations for convergence. In most cases, better performance is obtained by a limited memory variable metric method.

Limited memory variable metric methods differ from conjugate gradient methods in that the user is allowed to choose the number of vector of storage, and in each iteration,

$$x_{k+1} = x_k - \alpha_k H_k \nabla f(x_k)$$

where the symmetric, positive definite matrix H_k is determined by m vectors obtained during the previous m iterations, and stored in compact form so that the product $H_k \nabla f(x_k)$ only requires order mn flops (floating point operations). The number of iterations usually decreases as m increases, but the cost per iteration increases. In practice a choice of $m = 5$ is reasonable.

Limited memory variable metric methods have several advantages. For example, they only require the user to provide code for the evaluation of the function and the gradient, and the required storage is fixed. On the other hand, they tend to fail on difficult problems.

This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Computational and Technology Research, U.S. Department of Energy, under Contract W-31-109-Eng-38, and by the National Science Foundation, through the Center for Research on Parallel Computation, under Cooperative Agreement No. CCR-9120008.

Even if they converge, their rate of convergence can be too slow to obtain an accurate solution, if compared with a Newton's method. We will come back to this point when we discuss numerical results at the end of this paper.

An important difference between variations of Newton's method is how they obtain the Newton step. Implementations of Newton's method that use a direct method to obtain the step are not suitable for many large problems because of their cost in terms of computing time and storage. We prefer a variation of Newton's method that uses an iterative method to obtain the step because if properly preconditioned, this variation is efficient and reliable.

In this paper we propose a trust region version of Newton's method that uses an incomplete Cholesky decomposition and the conjugate gradient method to obtain the trust region step. In this algorithm the step s_k between iterates is an approximate solution to the trust region subproblem

$$\min \{ q_k(w) : \|D_k w\| \le \Delta_k \} ,$$

where $q_k : \mathbb{R}^n \mapsto \mathbb{R}$ is a quadratic model of the function at the current iterate, D_k is a scaling matrix, Δ_k is the trust region radius, and $\|\cdot\|$ is the l_2 norm. The trust region version of Newton's method is described in Section 2, and the computation of the step in the trust region method is described in Section 3.

Other codes that are suitable for the solution of large-scale unconstrained minimization problems include LBFGS (Liu and Nocedal [15]), TN (Nash [20]), TNPACK (Folgeson and Schlick [26, 27]), SBMIN (Conn, Gould, and Toint [7, 6, 8]), and STENMIN (Bouaricha [4, 5]).

LBFGS is a limited memory variable metric method, TN and TNPACK are truncated Newton methods, and STENMIN is a tensor code that bases each iteration on a tensor model of the objective function. Of these codes, only TN and TNPACK use preconditioners. Both TN and TNPACK algorithms use a preconditioned conjugate gradient method to obtain the Newton step. The preconditioner used in TN is a scaled two-step limited memory BFGS method. In TNPACK however, the user must supply a preconditioner; if the preconditioner is not positive definite, a modified Cholesky decomposition is used to obtain a positive definite preconditioner. The success of TNPACK clearly depends on the preconditioner. If the user supplies the Hessian matrix as a preconditioner, TNPACK reduces to a line search version of a modified Newton's method. With the Hessian as a preconditioner TNPACK would not be suitable for large problems, so the preconditioner must be chosen with care. SBMIN is the implementation of Newton's in the LANCELOT package. Although our algorithm and the SBMIN algorithm rely on the trust region philosophy,the codes are quite different. We detail some of the differences in Section 3.

In this paper we consider a preconditioner based on the incomplete Cholesky factorization of Jones and Plassmann [13] with the modifications proposed by Chin and Moré [14] for the solution of trust region subproblems. This version of the incomplete Cholesky

factorization has predictable storage requirements, a dynamic sparsity pattern, and does not require the specification of a drop tolerance. These advantages make this incomplete Cholesky factorization an attractive ingredient in an optimization code. The SBMIN code, on the other hand, uses the incomplete Cholesky factorization of Munksgaard [19] that has unpredictable storage requirements and requires the specification of a drop tolerance. As shown by Chin and Moré [14], codes that require the specification of a drop tolerance suffer from erratic behavior.

The test problems described in Section 5 come from the MINPACK-2 test problem collection [1] since this collection is representative of large-scale optimization problems arising from applications. We considered using the CUTE collection [3], but found few large scale problems that arise in applications.

The purpose of the numerical experiments in Section 6 is to analyze the performance of our implementation NMTRS of a trust region version Newton's method and to compare this code with the limited memory variable metric code VMLM. The results show that Newton's method is superior to the limited memory variable metric method VMLM in terms of computing time, and usually requires far fewer function and gradient evaluations. We emphasize that the cost of the function and gradient evaluations of the MINPACK-2 test problems is moderate since these are model problems that have been coded with care. On problems with more expensive function and gradient evaluations, Newton's method is likely to be the winner because of it requires a smaller number of function and gradient evaluations.

Numerical results for limited memory variable metric method and truncated Newton methods have been obtained by Phua [23], Nash and Nocedal [21], Gilbert and Lemarechal [10], Zou *et al.* [30], Papadrakakis and Pantazopoulos [22], and Lucidi and Roma [16]. Interestingly enough, the implementations of Newton's method in these papers do not use preconditioning. **Is this true?**

2 Newton's Method

At each iteration of a trust region Newton method for the minimization of $f : \mathbb{R}^n \mapsto \mathbb{R}$ we have an iterate x_k, a bound Δ_k, a scaling matrix D_k, and a quadratic model

$$q_k(w) = \nabla f(x_k)^T w + \tfrac{1}{2} w^T B_k w$$

of the possible reduction $f(x_k + w) - f(x_k)$ for $\|D_k w\| \le \Delta_k$. Given a step s_k, the test for acceptance of the trial point $x_k + s_k$ depends on a parameter $\eta_0 > 0$. The following algorithm summarizes the main computational steps:

For $k = 0, 1, \ldots,$ maxiter

Compute the quadratic model q_k.

Compute a scaling matrix D_k.
Compute an approximate solution s_k to the trust region subproblem.
Compute the ratio ρ_k of actual to predicted reduction.
Set $x_{k+1} = x_k + s_k$ if $\rho_k \geq \eta_0$; otherwise set $x_{k+1} = x_k$. Update Δ_k.

In this section we elaborate on this outline and on our implementation of the trust region Newton method for large-scale problems.

The computation of the model q_k requires the gradient and the approximate Hessian matrix B_k. In this work, B_k is usually either the Hessian matrix $\nabla^2 f(x_k)$, or a symmetric approximation to the Hessian matrix obtained by differences of the gradient.

The iterate x_k and the bound Δ_k are updated according to rules that are standard in trust region methods. Given a step s_k such that $\|D_k s_k\| \leq \Delta_k$ and $q_k(s_k) < 0$, these rules depend on the ratio

$$\rho_k = \frac{f(x_k + s_k) - f(x_k)}{q_k(s_k)} \tag{2.1}$$

of the actual reduction in the function to the predicted reduction in the model. Since the step s_k is chosen so that $q_k(s_k) < 0$, a step with $\rho_k > 0$ yields a reduction in the function. Given $\eta_0 > 0$, the iterate x_k is updated as in the basic algorithm, that is, $x_{k+1} = x_k + s_k$ if $\rho_k \geq \eta_0$, otherwise $x_{k+1} = x_k$. The updating rules for Δ_k depend on constants η_1 and η_2 such that

$$0 < \eta_0 < \eta_1 < \eta_2 < 1,$$

while the rate at which Δ_k is either increased or decreased depend on constants σ_1, σ_2, and σ_3 such that

$$0 < \sigma_1 < \sigma_2 < 1 < \sigma_3.$$

The trust region bound Δ_k is updated by setting

$$\begin{array}{lll} \Delta_{k+1} \in [\sigma_1 \Delta_k, \sigma_2 \Delta_k] & \text{if} & \rho_k \leq \eta_1 \\ \Delta_{k+1} \in [\sigma_1 \Delta_k, \sigma_3 \Delta_k] & \text{if} & \rho_k \in (\eta_1, \eta_2) \\ \Delta_{k+1} \in [\Delta_k, \sigma_3 \Delta_k] & \text{if} & \rho_k \geq \eta_2. \end{array}$$

We choose a step s_k that provides an approximate solution s_k to the trust region subproblem

$$\min \left\{ q_k(w) : \|D_k w\| \leq \Delta_k \right\}.$$

There is no need to compute an accurate minimizer; the main requirement is that the step s_k give as much reduction in the model q_k as the Cauchy step s_k^C, that is, a solution to the problem

$$\min \left\{ q_k(w) : \|D_k w\| \leq \Delta_k, \ w = -\nu \nabla f(x_k), \ \nu \in \mathbb{R} \right\}.$$

This requirement guarantees global convergence of the trust region method under suitable conditions.

The above outline of a trust region Newton method is standard. The main differences appear in the method used to compute the step s_k and the use of the scaling matrix D_k. We discuss the algorithm used to compute the step s_k in Section 3, and the scaling matrix in Section 4.

3 Computation of the step

The computation of the step in an unconstrained trust region method requires an approximate solution of the subproblem

$$\min \{ q(s) : \|Ds\| \le \Delta \}, \tag{3.1}$$

where D is the scaling matrix and $q : \mathbb{R}^n \mapsto \mathbb{R}$ is the quadratic

$$q(s) = g^T s + \tfrac{1}{2} s^T B s.$$

In this formulation g is the gradient at the current iterate and B is the approximate Hessian matrix. In this section we describe the method used to compute the step, and contrast our approach with others that have appeared in the literature.

In our approach we transform the ellipsoidal trust region into a spherical trust region and then apply the conjugate gradient method to the transformed problem. Thus, given the problem (3.1), we transform this problem into

$$\min \{ \widehat{q}(w) : \|w\| \le \Delta \}, \tag{3.2}$$

where

$$\widehat{q}(w) = \widehat{g}^T w + \tfrac{1}{2} w^T \widehat{B} w, \qquad \widehat{g} = D^{-1} g, \qquad \widehat{B} = D^{-1} B D^{-T}.$$

Given an approximate solution of (3.2), the corresponding solution of (3.1) is $s = D^{-T} w$.

We are interested in the solution of large-scale problems, and thus the conjugate gradient method, with suitable modifications that take into account the trust region constraint and the possible indefiniteness of $\widehat{B}$, is a natural choice for computing a step w. The global convergence theory of the trust region method only requires that

$$\widehat{q}(w) \le \beta \widehat{q}(s^C), \tag{3.3}$$

where s^C is the Cauchy step and β is a positive constant, but for fast local convergence we need

$$\|\nabla \widehat{q}(w)\| \le \xi \|\nabla \widehat{q}(0)\|, \tag{3.4}$$

where $\xi \in (0,1)$ is a constant that may depend on the iteration.

At the moment we leave the matrix D unspecified. The main requirement on D is that the resulting $\widehat{B}$ must have clustered eigenvalues. As we shall see in Section 4, this requirement is satisfied by choosing D from an incomplete Cholesky factorization of $\widehat{B}$.

The classical conjugate gradient method for minimizing the quadratic $\widehat{q}$ generates a sequence of iterates $\{w_k\}$ as follows:

Let $w_0 \in \mathbb{R}^n$ be given. Set $r_0 = -(\widehat{g} + \widehat{B}w_0)$ and $d_0 = r_0$.
For $k = 0, 1, \ldots,$
 Compute $\alpha_k = \|r_k\|^2/(d_k^T \widehat{B} d_k)$.
 Update the iterate: $w_{k+1} = w_k + \alpha_k d_k$.
 Update the residual: $r_{k+1} = r_k - \alpha_k \widehat{B} d_k$.
 Compute $\beta_k = \|r_{k+1}\|^2/\|r_k\|^2$.
 Update the direction: $d_{k+1} = r_{k+1} + \beta_k d_k$.

Since the trust region (3.2) is centered at the origin, we choose $w_0 = 0$. With this choice, the conjugate gradient method generates iterates $w_1, \ldots, w_m$ where m is the largest integer such that $d_k^T \widehat{B} d_k > 0$ for $0 \le k < m$, such that

$$\widehat{q}(w_k) = \min\left\{\widehat{q}(w) : w \in \operatorname{span}\{r_0, \widehat{B}r_0, \ldots, \widehat{B}^{k-1}r_0\}\right\}, \qquad 0 \le k \le m.$$

The conjugate gradient method terminates with $r_m = 0$ or with $d_m^T \widehat{B} d_m \le 0$. These properties of the conjugate gradient method are well-known, see, for example, Hestenes and Steifel [12], Hackbusch [11], Axelsson [2], or Saad [25]. We also need to know that since we have chosen $w_0 = 0$, the iterates satisfy

$$\|w_k\| < \|w_{k+1}\|, \qquad 0 \le k < m.$$

This result was first obtained by Steihaug [28] from the inequality

$$(w_{k+1} - w_k)^T (w_{j+1} - w_j) > 0, \qquad j \le k.$$

As we shall see, this result is of importance to our development.

We claim that as the conjugate gradient method generates iterates for the trust region problem (3.2), there is an index k with $k \le m$ such that $\|w_k\| \le \Delta$ and one of the following three conditions holds:

$$\|r_k\| \le \xi \|\widehat{g}\| \qquad d_k^T \widehat{B} d_k \le 0 \qquad \|w_{k+1}\| > \Delta. \tag{3.5}$$

Clearly, if we continue to generate iterates with $\|w_k\| \le \Delta$ then one of the first two conditions in (3.5) will hold for some $k \le m$; otherwise, we must exit the trust region at some iterate, and then the third condition holds.

Given an iterate w_k that satisfies the conditions in (3.5), we can compute a suitable step w. If for some $\xi \in (0, 1)$ and $0 \le k \le m$ we have $\|w_k\| \le \Delta$ and $\|r_k\| \le \xi \|g\|$, then we accept $w = w_k$ as the step. In this case w is a truncated Newton step. If $\|w_k\| \le \Delta$ and

$d_k^T \widehat{B} d_k \le 0$, or if $\|w_k\| \le \Delta$ and $\|w_{k+1}\| > \Delta$, let $\tau_k > 0$ be the unique positive solution to the quadratic equation

$$\|w_k + \tau d_k\| = \Delta,$$

and set $w = w_k + \tau_k d_k$. In all cases we have

$$\widehat{q}(w) \le \min\{\widehat{q}(w_k), \widehat{q}(s^C)\},$$

for the final step w, and thus this choice of step satisfies the sufficient decrease condition (3.3) of trust region methods.

Steihaug [28] suggested the trust region step described above, but his description was in terms of the preconditioned conjugate gradient method applied to the problem

$$\min\left\{ q(s) : (s^T C s)^{\frac{1}{2}} \le \Delta \right\},$$

where C is a positive definite matrix. The preconditioned conjugate gradient method for minimizing q takes the following form:

> Let $s_0 \in \mathbb{R}^n$ be given. Set $r_0 = -(g + Bs_0)$ and $d_0 = q_0$ where $Cq_0 = r_0$.
> For $k = 0, 1, \ldots,$
>> Compute $\alpha_k = (r_k^T q_k)/(d_k^T B d_k)$.
>> Update the iterate: $s_{k+1} = s_k + \alpha_k d_k$.
>> Update the residual: $r_{k+1} = r_k - \alpha_k B d_k$.
>> Solve $Cq_{k+1} = r_{k+1}$.
>> Compute $\beta_k = (r_{k+1}^T q_{k+1})/(r_k^T q_k)$.
>> Update the direction: $d_{k+1} = q_{k+1} + \beta_k d_k$.

The relationship between this algorithm for generating $\{s_k\}$, and our algorithm for generating $\{w_k\}$ is that if $C = D^T D$ then $Ds_k = w_k$. This can be verified by an induction argument. Thus, the two methods are equivalent. However, our approach based on the subproblem (3.2) is more direct. Moreover, we deal with the scaling matrix D directly, and not through the matrix $C = D^T D$.

Conn, Gould, and Toint [7, 6] and [8, Sections 3.2 and 3.3] also use the preconditioned conjugate gradient method to compute steps in a trust region method. In their approach, the trust region subproblem

$$\min\{q(s) : \|s\|_\infty \le \Delta\}$$

has no scaling matrix and uses the l_∞ norm. A preconditioned conjugate gradient method is used to generate the steps $\{s_k\}$; the preconditioner is a diagonal matrix in [7, 6], while more general preconditioners are used in [8].

A disadvantage of the Conn, Gould, and Toint [8] approach is that they cannot rely on the property that $\{\|s_k\|\}$ is monotonically increasing; instead, they have that $\{s_k^T C s_k\}$ is

monotonically increasing, where C is the preconditioner. As a consequence iterates may enter and leave the trust region several times. This does not destroy the global convergence properties of the algorithm, but can lead to poor behavior because we would expect later iterates to produce better steps. In particular, if the first iterate exits the trust region, then their approach generates a steepest descent iterate.

4 Incomplete Cholesky factorizations as preconditioners

The efficiency and reliability of the algorithm for computing the trust region step depends on the preconditioner. There are many widely used preconditioners, ranging from diagonal to full-matrix preconditioners. However, choosing a good preconditioner for a given problem has always been a difficult task. Diagonal or band preconditioners may not be successful if the essential part of the Hessian matrix is not contained within the diagonal or the band, respectively. Full-matrix preconditioners, on the other hand, may be prohibitive if it is too expensive to factor or store the Hessian matrix. In this section we describe a preconditioner based on the incomplete Cholesky factorization.

Given a symmetric matrix B and a symmetric sparsity pattern $\mathcal{S}$, an incomplete Cholesky factorization is a lower triangular matrix L such that

$$B = LL^T + R, \qquad l_{i,j} = 0 \text{ if } (i,j) \notin \mathcal{S}, \quad r_{i,j} = 0 \text{ if } (i,j) \in \mathcal{S}. \tag{4.1}$$

The incomplete Cholesky factorization may not exist for a general symmetric matrix B. Existence of the incomplete Cholesky factorization is only guaranteed if B is an H-matrix with positive diagonal elements. For additional information on incomplete Cholesky factorizations, see Axelsson [2], or Saad [25].

A difficulty with an incomplete Cholesky factorization is that it is not clear how to choose the sparsity pattern $\mathcal{S}$ for the numerical factorization. We could choose $\mathcal{S}$ to be the sparsity pattern of B, but it is usually advantageous to allow some fill. Several approaches for defining $\mathcal{S}$ have been proposed. Fixed fill strategies (Meijerink and Van Der Vorst [18]) fix the nonzero structure of the incomplete factor prior to the factorization. Drop tolerance strategies (Munksgaard [19], for example) include nonzeros in the incomplete factor if they are larger than some threshold parameter. Thus, the number of nonzeros in the incomplete factor is unknown prior to the factorization.

Another difficulty with an incomplete Cholesky factorization is that the factorization may fail or be unstable for a general positive definite matrix. The standard solution for this situation is to increase the size of the elements in the diagonal until a satisfactory factorization is obtained. See, for example, Manteuffel [17] and Munksgaard [19].

The strategies described above require that the user impose a fixed sparsity pattern on the Cholesky factor or specify a drop tolerance. Jones and Plassmann [13] and Saad [24, 25] have proposed factorizations in which the sparsity pattern is determined by the relative size

of the elements in the factorization, and by the available storage. Our approach is closely related to the factorization proposed by Jones and Plassmann, but the ILUT factorization of Saad is similar.

Jones and Plassmann retain the m_k largest nonzeroes in the off-diagonal part of the k-th column of L, where m_k is the number of nonzero off-diagonal elements in the k-th column of the lower triangular part of B. Thus, the number of nonzeros in the incomplete factor is the same as in the original matrix. The numerical results of Jones and Plassmann show that significant improvements in performance are obtained on problems generated from finite element models, and on problems from the Harwell-Boeing sparse matrix collection. Their algorithm can be summarized as follows:

Algorithm 4.1 Let B be a symmetric matrix, and let m_k be the number of nonzero off-diagonal elements in the k-th column of the lower triangular part of B.

For $k = 0, 1, \ldots,$

Compute the elements in the k-th column of L.

Select the m_k nonzeros in the k-th column of L with the largest magnitude.

Update the last $n - k$ diagonal elements of L.

Algorithm 4.1 is an outline of the incomplete Cholesky factorization that we use. For additional information, see Lin and Moré [14].

The performance of Algorithm 4.1 depends on the size of the elements of B. For a general positive definite matrix B, Jones and Plassmann [13] recommended computing a scaled matrix

$$\widehat{B} = D^{-1/2} B D^{-1/2}, \qquad D = \operatorname{diag}(b_{i,i}),$$

and using Algorithm 4.1 on $\widehat{B}_k = \widehat{B} + \alpha_k I$ for some $\alpha_k \geq 0$. If the incomplete Cholesky factor of $\widehat{B}_k$ is $\widehat{L}_k$, then $L_k = D^{1/2}\widehat{L}_k$ is the factor of $B + \alpha_k D$. The algorithm used to generate α_k was to start with $\alpha_0 = 0$, and increment α_k by a constant factor (0.01) until Algorithm 4.1 succeeded.

This approach does not extend to general indefinite matrices. In the general case B may have zero or small diagonal elements, and then the scaling above is almost certain to produce a badly scaled matrix. We also note that the strategy of adding a constant factor to the diagonal is not likely to be efficient in general.

In our approach we scale the initial matrix by the l_2 norm of the columns of B. Scaling by the l_2 norm of the columns of B works better than setting

$$D = \operatorname{diag}\{|b_{i,i}|, \epsilon\}, \quad \text{or} \quad D = \operatorname{diag}\{\max(b_{i,i}, \epsilon)\},$$

for some $\epsilon > 0$, which are scalings that are commonly used in optimization (see, for example, Gay [9], Conn, Gould, and Toint [8, page 125]. The following algorithm specifies our strategy in detail.

Algorithm 4.2 Let B be a symmetric matrix.

> Choose $\alpha_S > 0$.
> Compute $\widehat{B} = D^{-1/2} B D^{-1/2}$ where $D = \text{diag}(\|Be_i\|_2)$.
> Set $\alpha_0 = 0$ if $\min(\widehat{b}_{ii}) > 0$; otherwise $\alpha_0 = -\min(\widehat{b}_{ii}) + \alpha_S$.
> For $k = 0, 1, \ldots,$
>> Use Algorithm 4.1 on $\widehat{B}_k = \widehat{B} + \alpha_k I$; exit if successful.
>> Set $\alpha_{k+1} = \max(2\alpha_k, \alpha_S)$

For any α_0, this algorithm is guaranteed to generate an H-matrix $\widehat{B}_k$ for some $k \geq 0$. Hence, the incomplete Cholesky factorization exists. The choice of $\alpha_0 = 0$ is certainly reasonable if B is positive definite, or more generally, if B has positive diagonal elements. A reasonable initial choice for α_0 is not clear if B is an indefinite matrix, but our choice of α_0 with $\alpha_S = 10^{-3}$ seems to be adequate.

Algorithm 4.2 produces the incomplete Cholesky factor $\widehat{L}_k$ of the final $\widehat{B}_k$, and thus $L_k = D^{1/2}\widehat{L}_k$ is the factor of $B + \alpha_k D$. Since we are interested in using L_k as a preconditioner for B in the conjugate gradient method, we want the preconditioned matrix

$$\widehat{L}_k^{-1} \widehat{B} \widehat{L}_k^{-T} = L_k^{-1} B L_k^{-T}$$

to have clustered eigenvalues. We give an idea of the behavior of L_k as a preconditioner by plotting the eigenvalues of B and of the preconditioned matrix $L_k^{-1} B L_k^{-T}$ for two problems.

Figure 4.1 is a plot of the eigenvalues of B and of the preconditioned matrix for the msa problem with $n = 100$ variables. In this problem the Hessian matrix is positive definite and Algorithm 4.2 produces the incomplete Cholesky factorization with $\alpha_0 = 0$. Note that the eigenvalues of the preconditioned matrix are clustered and that the condition number of the preconditioned matrix is reduced. The important effect is the clustering of the eigenvalues.

The plot in Figure 4.1 is fairly typical of the positive definite problems that we have tried, and thus we expect good behavior from the conjugate gradient method. For more information on the effect of clustering on the convergence behavior of the conjugate gradient method, see Van der Vorst [29]. We now consider indefinite problems.

Figure 4.2 has the plots of the eigenvalues of B and of the preconditioned matrix for a randomly generated sparse symmetric matrix of order 100. We used the `sprandsym` function in MATLAB with a density of 0.05 and a distribution of eigenvalues defined by

$$\lambda_k = \rho_k 10^{s_k}, \qquad s_k = \text{ncond}\left(\frac{k-1}{n-1}\right)$$

where ρ_k is uniformly distributed in $[-1, 1]$ and ncond $= 8$. For this problem Algorithm 4.2 produced the incomplete Cholesky factorization with $\alpha_2 = \beta$.

Since the Hessian matrix is indefinite, we have plotted the magnitude of both the positive and the negative eigenvalues. The plots of the eigenvalues of the original matrix are on the

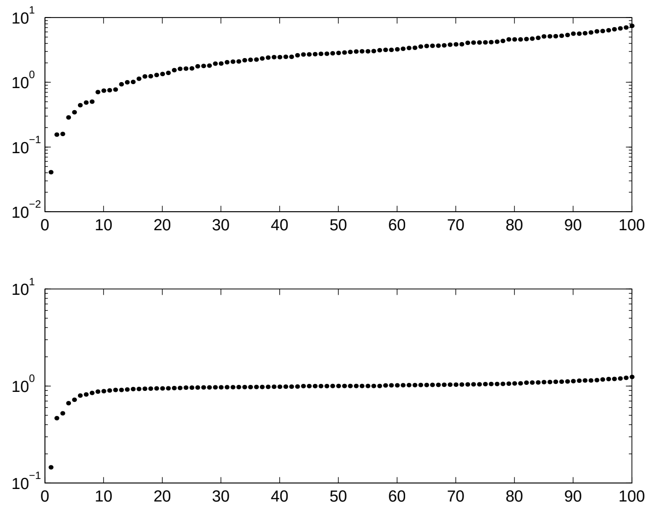

Figure 4.1: Eigenvalues of B (top) and $L_k^{-1} B L_k^{-T}$ (bottom) for the msa problem

top of Figure 4.2, while the plots of the eigenvalues of the preconditioned matrix are on the bottom. These plots show that Algorithm 4.2 tends to cluster the eigenvalues of largest magnitude, and tends to reduce the condition number of the preconditioned problem. In general we do not get the almost spectacular clustering shown in Figure 4.2, but clustering of the largest positive eigenvalues does seem to be a general trend. In a trust region method we would want to have clustering of the positive eigenvalues because the algorithm for computing the step in Section 3 terminates successfully once the conjugate gradient method determines a direction of negative curvature.

5 Large-Scale Optimization Problems

Most of our test problems come from the MINPACK-2 test problem collection [1] since this collection is representative of large-scale optimization problems arising from applications. Below we give brief descriptions of the infinite dimensional version of these problems, and of the finite dimensional formulations. The MINPACK-2 report contains additional information on these problems; in particular, parameter values are chosen as in this report.

The optimization problems that we consider arise from the need to minimize a function f of the form

$$f(v) = \int_{\mathcal{D}} \Phi(x, v, \nabla v)\, dx, \tag{5.1}$$

where $\mathcal{D}$ is some domain in either $\mathbb{R}$ or $\mathbb{R}^2$, and Φ is defined by the application. In all cases f is well defined if $v : \mathcal{D} \mapsto \mathbb{R}^p$ belongs to $H^1(\mathcal{D})$, the Hilbert space of functions such

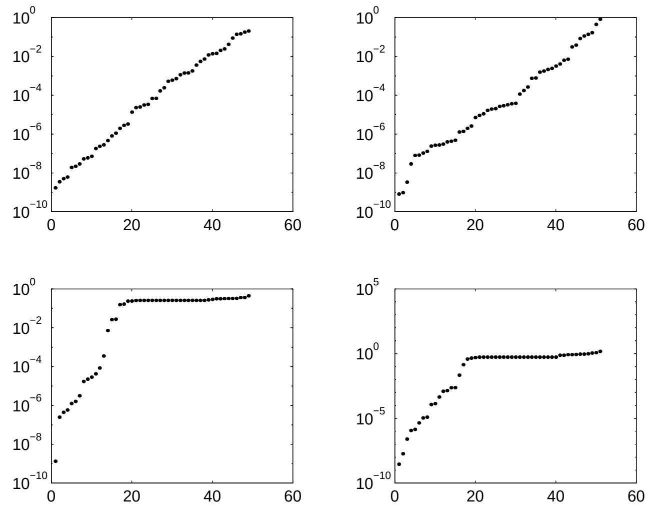

Figure 4.2: Magnitude of the eigenvalues of B and $L_k^{-1}BL_k^{-T}$ for a randomly generated sparse indefinite matrix of order 100. The positive and negative eigenvalues of B are on the top left and top right, respectively. Similarly, the positive and negative eigenvalues of $L_k^{-1}BL_k^{-T}$ are on the bottom left and bottom right, respectively.

that v and $\|\nabla v\|$ belong to $L^2(\mathcal{D})$. This is the proper setting for examining existence and uniqueness questions for the infinite-dimensional problem.

Finite element approximations to these problems are obtained by minimizing f over the space of piecewise linear functions v with values $v_{i,j}$ at $z_{i,j}$, $0 \le i \le n_y+1$, $0 \le j \le n_x+1$, where $z_{i,j} \in \mathbb{R}^2$ are the vertices of a triangulation of $\mathcal{D}$ with grid spacings h_x and h_y. The vertices $z_{i,j}$ are chosen to be a regular lattice so that there are n_x and n_y interior grid points in the coordinate directions, respectively.

Lower triangular elements T_L are defined by vertices $z_{i,j}, z_{i+1,j}, z_{i,j+1}$, while upper triangular elements T_U are defined by vertices $z_{i,j}, z_{i-1,j}, z_{i,j-1}$. The values $v_{i,j}$ are obtained by solving the minimization problem

$$\min\left\{\sum\left(f_{i,j}^L(v) + f_{i,j}^U(v)\right) : v \in \mathbb{R}^n\right\},$$

where $f_{i,j}^L$ and $f_{i,j}^U$ are the finite element approximation to the integrals in the elements T_L and T_U, respectively.

The elastic plastic torsion (**ept**) and the journal bearing problem (**jbp**) are quadratic problems of the form

$$\min\{f(v) : v \in K\}, \tag{5.2}$$

where $f : K \mapsto \mathbb{R}$ is the quadratic

$$f(v) = \int_{\mathcal{D}} \left\{\tfrac{1}{2} w_q(x)\|\nabla v(x)\|^2 - w_l(x)v(x)\right\} dx,$$

where $w_q : \mathcal{D} \mapsto \mathbb{R}$ and $w_l : \mathcal{D} \mapsto \mathbb{R}$ are functions defined on the rectangle $\mathcal{D}$. In the **ept** problem $w_q \equiv 1$ and $w_l \equiv c$. For our numerical results we use $c = 5$. In the **jbp** problem

$$w_q(\xi_1, \xi_2) = (1 + \epsilon \cos \xi_1)^3, \quad w_l(\xi_1, \xi_2) = \epsilon \sin \xi_1$$

for some constant ϵ in $(0, 1)$, and $\mathcal{D} = (0, 2\pi) \times (0, 2b)$ for some constant $b > 0$. For our numerical results we use $\epsilon = 0.1$ and $b = 10$.

The steady state combustion (**ssc**) problem and the optimal design with composites (**odc**) are formulated in terms of a family of minimization problems of the form

$$\min\{f_\lambda(v) : v \in H_0^1(\mathcal{D})\}.$$

For the **ssc** problem $f_\lambda : H_0^1(\mathcal{D}) \mapsto \mathbb{R}$ is the functional

$$f_\lambda(v) = \int_{\mathcal{D}} \left\{ \tfrac{1}{2} \|\nabla v(x)\|^2 - \lambda \exp[v(x)] \right\} dx,$$

and $\lambda \geq 0$ is a parameter. For the **odc** problem

$$f_\lambda(v) = \int_{\mathcal{D}} \left\{ \psi_\lambda \left(\|\nabla v(x)\| \right) + v(x) \right\} dx,$$

and $\psi_\lambda : \mathbb{R} \mapsto \mathbb{R}$ is the piecewise quadratic

$$\psi_\lambda(t) = \begin{cases} t^2, & 0 \leq t \leq t_1, \\ 2t_1(t - \frac{1}{2}t_1), & t_1 \leq t \leq t_2, \\ \frac{1}{2}(t^2 - t_2^2) + 2t_1(t_2 - \frac{1}{2}t_1), & t_2 \leq t, \end{cases}$$

with the breakpoints t_1 and t_2 defined by $t_1^2 = \lambda$, $t_2^2 = 2\lambda$. In our numerical results we consider the problem of minimizing f_λ for a fixed value of λ; for the **ssc** problem $\lambda = 2$, while for the **odc** problem $\lambda = 0.008$.

The minimal surface area (**msa**) problem is an optimization problem of the form (5.2) where $f : K \mapsto \mathbb{R}$ is the functional

$$f(v) = \int_{\mathcal{D}} \left(1 + \|\nabla v(x)\|^2 \right)^{1/2} dx,$$

and the set K is defined by

$$K = \left\{ v \in H^1(\mathcal{D}) : v(x) = v_D(x) \text{ for } x \in \partial\mathcal{D} \right\}$$

for the boundary data function $v_D : \partial\mathcal{D} \mapsto \mathbb{R}$ that specifies the Enneper minimal surface.

In all the problems so far v has been defined in some domain in $\mathbb{R}^2$. For the one-dimensional Ginzburg-Landau problem (**gl1**) v is defined in a subset of $\mathbb{R}$. This problem is defined by

$$\min\{f(v) : v(-d) = v(d), \ v \in C^1[-d, d]\},$$

where $2d$ is a constant (the width of the superconducting material), and

$$f(v) = \frac{1}{2d} \int_{-d}^{d} \left\{ \alpha(\xi)|v(\xi)|^2 + \tfrac{1}{2}\beta(\xi)|v(\xi)|^4 + \gamma|v'(\xi)|^2 \right\} \, d\xi.$$

The functions α and β are piecewise constant, and γ is a (universal) constant.

The Ginzburg-Landau problem (gl2) can be phrased as an optimization problem of the general form (5.1), where v is defined in $\mathbb{R}^4$. The first two components represent a complex-valued function ψ (the order parameter), and the other two components a vector-valued function (the vector potential). The dimensionless form of this problem is

$$\min\{f_1(\psi) + f_2(\psi, A) : \psi, A \in H_0^1(\mathcal{D})\},$$

where $\mathcal{D}$ is a two-dimensional region,

$$f_1(\psi) = \int_{\mathcal{D}} \left\{ -|\psi(x)|^2 + \tfrac{1}{2}|\psi(x)|^4 \right\} \, dx,$$

$$f_2(\psi, A) = \int_{\mathcal{D}} \left\{ \left\| [\nabla - iA(x)] \, \psi(x) \right\|^2 + \kappa^2 \left\| (\nabla \times A)(x) \right\|^2 \right\} \, dx,$$

and κ is the Ginzburg-landau constant.

6 Numerical Experiments

Our aim in this section is to analyze the performance of the NMTRS code that implements a trust region version of Newton's method. In particular, we compare the performance of the trust region code with the limited memory variable metric code VMLM since several numerical studies have shown that algorithms of this type compare well with other codes for the solution of large-scale problems.

We consider the behavior of the algorithm as the number of variables n varies between $2,500$ and $40,000$ since one of our purposes is to study the behavior of these algorithms as the number of variables increases and to show that the trust region code can easily produce solutions for problems where the number of variables is in this range.

For most of our results we imposed a limit of one hour of computing time per problem, and a limit of $5,000$ gradient evaluations. Our computations were performed on a Sun SPARC 10 workstation using double precision arithmetic. The termination test used in these results was

$$\|\nabla f(x)\| \le \tau \|\nabla f(x_0)\|, \qquad \tau = 10^{-5}.$$

This test is scale invariant, and scales as the number of variables increases, but we do not recommend this test as a general termination test for optimization algorithms. The choice of $\tau = 10^{-5}$ tests the ability to solve optimization problems to moderate accuracy.

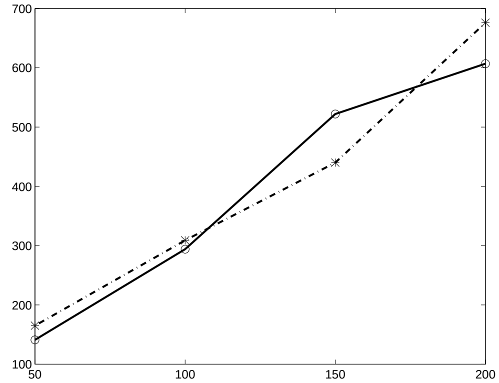

Figure 6.1: Number of iterations of VMLM as a function of $n^{1/2}$

In the trust region algorithm we accept x_{k+1} as a next iterate if the ratio ρ_k in (2.1) is bigger than 10^{-4}; otherwise we update the trust region radius as follows:

$$\begin{aligned} \Delta_{k+1} &= 0.5\ \Delta_k && \text{if} \quad \rho_k < 0.25 \\ \Delta_{k+1} &= \Delta_k && \text{if} \quad \rho_k \in (0.25, 0.5] \\ \Delta_{k+1} &= 2.0\ \Delta_k && \text{if} \quad \rho_k \in (0.5, 0.9] \\ \Delta_{k+1} &= 4.0\ \Delta_k && \text{if} \quad \rho_k \geq 0.9 \end{aligned}$$

We chose $\Delta_0 = \min(1000.0 \parallel g(x_0) \parallel, 1000.0)$, where $g(x_0)$ is the gradient value at the initial guess x_0. We selected this value of Δ_0 based on the experimental test results. In the conjugate gradient iterative algorithm, the value of ξ in (3.4) is set to 10^{-2} for all our tests.

We have to change the way Δ is chosen; the above choice stinks!

Table 6.1 contains a summary of the results for the NMTRS and VMLM codes on the MINPACK-2 large-scale problems [1]. In this table `iters` is the number of iterations, `nfev` is the number of function and gradient evaluations, `nhev` is the number of Hessian evaluations, `ncg` is the number of conjugate gradient iterations, and `time` is the computing time (in seconds).

An important difference between a limited memory variable metric method and a Newton method is that for the Newton method the number of iterations required to solve a variational problem (such as those described in Section 5) can be independent of n. This can be seen in Table 6.1 for most of the problems. In contrast, the number of iterations required by a limited memory variable metric method is likely to grow with n. Table 6.1 suggests that for these problems the number of iterations grows like $n^{1/2}$. This is verified in Figure 6.1 by plotting the number of gradient evaluations required to solve the ssc and msa problems.

Since the computing time for evaluating the function and gradient of these problems grows linearly with n, and the number of iterations grows like $n^{1/2}$, the computing time

Table 6.1: Performance of NMTRS versus VMLM on the MINPACK-2 test problems

		NMTRS					VMLM		
Problem	n	`iters`	`nfev`	`nhev`	`ncg`	`time`	`iters`	`nfev`	`time`
ept	2500	3	4	4	27	0.8	124	133	2.8
ept	10000	3	4	4	46	3.9	280	293	25.1
ept	40000	3	4	4	88	22.4	610	629	217.0
pjb	2500	3	4	4	32	0.8	226	236	5.1
pjb	10000	3	4	4	61	4.6	423	445	38.2
pjb	40000	3	4	4	120	28.0	823	855	310.2
gl2	2500	8	15	9	439	25.6	2710	2786	50.1
gl2	10000	11	17	12	1109	171.5	†4831	†5000	†368.0
gl2	40000	7	13	8	1052	506.9	†4839	†5000	†1535.0
msa	2500	6	7	7	68	2.8	138	144	4.7
msa	10000	6	7	7	98	12.4	269	277	34.8
msa	40000	10	14	11	229	91.2	609	623	307.0
ssc	2500	3	4	4	33	1.4	167	176	7.5
ssc	10000	3	4	4	59	6.8	345	358	60.8
ssc	40000	3	4	4	113	35.6	588	615	414.8
gl1	2500	14	23	14	62	1.5	†4844	†5000	†61.7
gl1	10000	15	17	15	44	5.5	†4847	†5000	†246.3
gl1	40000	21	30	21	61	30.9	†4854	†5000	†933.0
odc	2500	40	61	41	253	20.3	260	268	11.3
odc	10000	187	274	188	858	362.4	410	415	69.3
odc	40000	882	1217	883	3946	7114.0	1209	1228	815.9

†Limit of 5000 function evaluations reached.

Table 6.2: Percentages of the computing time for NMTRS and $n = 10,000$

Problem	$f(x)$, $\nabla f(x)$	$\nabla^2 f(x)$	ICF	CG
ept	6	28	14	49
pjb	5	26	12	55
gl2	1	19	47	33
msa	7	50	9	33
ssc	11	44	8	36
gl1	4	28	30	33
odc	9	68	10	11

to solve these problems with VMLM grows like $n^{3/2}$. In particular, this implies that if we quadruple the number of variables in VMLM, then the computing time grows by a factor of 8. This is verified by the results in Table 6.1.

For the Newton method the computing time depends on the cost of the four main components of the Newton iteration: the function and gradient evaluation, the evaluation of the Hessian matrix, the incomplete Cholesky factor of the Hessian matrix, and the conjugate gradient iteration. In general we cannot expect a linear growth in the computing cost for Newton's method because the number of conjugate gradient iterations is likely to grow with n; this holds even for simple model problems like ept.

The results in Table 6.1 show that the computing time of NMTRS grows almost like $n^{3/2}$. The reason for this is that for these problems the cost of the conjugate gradient iterations tends to dominate, and the number of conjugate gradient iterations tends to grow like $n^{1/2}$.

An analysis of the computing time reveals where the computing time is spent. In Table 6.2 we present the percentages of the overall computing time spent by the various components of the Newton method: the function and gradient evaluation, the Hessian evaluation, the incomplete Cholesky factorization, and the conjugate gradient iteration. In this table $n = 10,000$.

Table 6.2 shows that, with the exception of the odc and msa problems, the cost of the conjugate gradient iterations dominates the computing time. Elaborate on this.

We tried to improve the performance of the VMLM code by increasing the number of vectors saved. The results for $m = 5, 10, 15, 20$ appear in Table 6.3. These results show, in particular, that the number of function and gradient evaluations usually decreases as we increase m from $m = 5$ to $m = 10$. The 45% reduction for the gl2 problem is significant, but the reductions are less than 20% reduction for the other problems.

The results in Table 6.3 also show that the computing time usually increases as we increase m. This is certainly the case if we compare the results for $m = 5$ with $m = 10, 15, 20$.

Table 6.3: Performance of VMLM as a function of m for $n = 10,000$

	$m=5$			$m=10$			$m=15$			$m=20$		
Problem	iters	nfev	time	iters	nfev	time	iters	nfev	time	iters	nfev	time
ept	281	293	25	245	249	27	240	248	30	192	198	27
pjb	424	445	38	395	404	43	393	400	48	340	350	48
gl2	6067	6277	471	3345	3441	328	3538	3622	400	2701	2780	357
msa	270	277	34	222	226	32	218	226	35	220	225	39
ssc	346	358	60	279	289	54	237	242	49	241	250	55
odc	411	415	67	476	489	89	396	403	80	473	480	104

We would have expected a decrease in computing time if the cost of evaluating the function and gradient dominated, but this is not the case for these problems. The VMLM algorithm requires $(8m + 12)n$ operations per iteration; the number of operations for the function and gradient is harder to estimate since most of the problems require the evaluation of intrinsic functions. The cost of the function and gradient evaluations relative to the internal arithmetic of the algorithm can be measured by computing the ratio t_f/τ_m, where t_f is the time required to evaluate the function and gradient, and τ_m is the time required for the $(8m + 12)n$ operations. These ratios are shown in Table 6.4 for $m = 5$.

Given these ratios, we can predict the required reduction in the number of iterations before we see an improvement in the execution time for VMLM. The prediction is based on expressing the time required for convergence as

$$t_m = (t_f + \tau_m)n_m,$$

where n_m is the number of iterations required for convergence. We have ignored the time required by the algorithm when the line search requires more than one function and gradient evaluation since most iterations only require one function and gradient evaluation. Since we expect τ_m to be proportional to the number of floating point operations, $\tau_{10} \geq 1.7\,\tau_5$. Hence, $t_{10} \leq t_5$ implies that we must have

$$\frac{n_{10}}{n_5} \leq \frac{t_f + \tau_5}{t_f + \tau_{10}} \equiv \frac{1 + \nu_5}{(\tau_{10}/\tau_5) + \nu_5} \leq \frac{1 + \nu_5}{1.7 + \nu_5},$$

where ν_m is the ratio t_f/τ_m. This inequality gives a bound on the required reduction in iterations needed for a decrease in computing time. For example, this inequality shows that for any function with ν in the interval $[1.4, 2.3]$ we can only have $t_{10} \leq t_5$ if we have a reduction of approximately 20% in the number of iterations; for $\nu_5 = 1.4$ the reduction should be 23%, while for $\nu_5 = 2.3$ the reduction should be 17%. The results in Table 6.3 confirm this analysis.

Table 6.4: Ratios $\nu_m = t_f/\tau_m$ of VMLM for $n = 10,000$ and $m = 5$

Problem	ept	pjb	gl2	msa	ssc	odc
t_f/τ_m	1.4	1.4	1.0	2.3	3.5	3.3

The above analysis shows that if the cost of the function and gradient evaluation is moderate relative to the cost of the arithmetic in the algorithm, then it probably does not pay to increase the number of vectors saved. If the cost of the function and gradient evaluation increases, then it does pay to increase m. However, then the VMLM code requires storage comparable with Newton's method. Moreover, for problems with expensive function and gradient evaluations the cost of the overall algorithm will be determined by the number of function and gradient evaluations, and as we have already seen, Newton's method is likely to require a smaller number of function and gradient evaluations.

Of course, there is no guarantee that increasing m leads to a reduction in computing time. This can be seen in Table 6.3 by comparing the results of $m = 5$ with those for $m = 10, 15, 20$. Note, in particular, that gl2 is the only problem where we obtained more than a 20% reduction in computing time by increasing m.

References

[1] B. M. Averick, R. G. Carter, J. J. Moré, and G.-L. Xue, *The MINPACK-2 test problem collection*, Preprint MCS-P153-0694, Mathematics and Computer Science Division, Argonne National Laboratory, 1992.

[2] O. Axelsson, *Iterative Solution Methods*, Cambridge Univeristy Press, 1994.

[3] I. Bongartz, A. R. Conn, N. I. M. Gould, and P. L. Toint, *CUTE: Constrained and unconstrained testing environment*, ACM Trans. Math. Software, 21 (1995), pp. 123–160.

[4] A. Bouaricha, *Algorithm 765; STENMIN: A software package for large, sparse unconstrained optimization using tensor methods*, ACM Trans. Math. Software, 23 (1997), pp. 81–90.

[5] ——, *Tensor methods for large, sparse unconstrained optimization*, SIAM J. Optim., 7 (1997), pp. 732–756.

[6] A. R. Conn, N. I. M. Gould, M. Lescrenier, and P. L. Toint, *Performance of a multifrontal scheme for partially separable optimization*, Report 88-4, Namur University, Namur, Belgium, 1988.

[7] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Testing a class of methods for solving minimization problems with simple bounds on the variables*, Math. Comp., 50 (1988), pp. 399–430.

[8] ——, LANCELOT, Springer Series in Computational Mathematics, Springer-Verlag, 1992.

[9] D. M. Gay, *Subroutines for unconstrained minimization using a model/trust region approach*, ACM Trans. Math. Software, 9 (1983), pp. 503–524.

[10] J. C. Gilbert and J. Nocedal, *Global convergence properties of conjugate gradient methods for optimization*, SIAM J. Optimization, 2 (1992), pp. 21–42.

[11] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, Applied Mathematical Sciences 95, Springer-Verlag, 1994.

[12] M. R. Hestenes and E. Steifel, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 409–436.

[13] M. T. Jones and P. E. Plassmann, *An improved incomplete Cholesky factorization*, ACM Trans. Math. Software, 21 (1995), pp. 5–17.

[14] C.-J. Lin and J. J. Moré, *Incomplete cholesky factorizations with limited memory*, Preprint MCS-P682-0897, Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439, 1997.

[15] D. C. Liu and J. Nocedal, *On the limited memory BFGS method for large scale optimization*, Math. Programming, 45 (1989), pp. 503–528.

[16] S. Lucidi and M. Roma, *Numerical experience with new truncated newton methods in large scale unconstrained optimization*, Comp. Optim. Appl., 7 (1997), pp. 71–87.

[17] T. A. Manteuffel, *An incomplete factorization technique for positive definite linear systems*, Math. Comp., 34 (1980), pp. 307–327.

[18] J. A. Meijerink and H. A. Van der Vorst, *An iterative solution method for linear equations systems of which the coefficient matrix is a symmetric M-matrix*, Math. Comp., 31 (1977), pp. 148–162.

[19] N. Munksgaard, *Solving sparse symmetric sets of linear equations by preconditioned conjugate gradients*, ACM Trans. Math. Software, 6 (1980), pp. 206–219.

[20] S. G. Nash, *Newton-like methods minimization via the Lanczos method*, SIAM J. Numer. Anal., 21 (1984), pp. 770–788.

[21] S. G. Nash and J. Nocedal, *A numerical study of the limited memory BFGS method and the truncated Newton method for large scale optimization*, SIAM J. Optimization, 1 (1991), pp. 358–372.

[22] M. Papadrakakis and G. Pantazopoulos, *A survey of quasi-Newton methods with reduced storage*, Internat. J. Numer. Methods Engrg., 36 (1993), pp. 1573–1596.

[23] K. H. Phua, *An evaluation of nonlinear optimization codes on supercomputers*, International Journal of High Speed Computing, 3 (1991), pp. 199–213.

[24] Y. Saad, *ILUT: A dual threshold incomplete LU factorization*, Numer. Linear Algebra Appl., 4 (1994), pp. 387–402.

[25] ——, *Iterative Methods for Sparse Linear Systems*, PWS Publishing Company, Boston, 1996.

[26] T. Schlick and A. Fogelson, *TNPACK – A truncated Newton minimization package for large-scale problems: I. Algorithms and usage*, ACM Trans. Math. Software, 18 (1992), pp. 46–70.

[27] ——, *TNPACK – A truncated Newton minimization package for large-scale problems: II. Implementations examples*, ACM Trans. Math. Software, 18 (1992), pp. 71–111.

[28] T. Steihaug, *The conjugate gradient method and trust regions in large scale optimization*, SIAM J. Numer. Anal., 20 (1983), pp. 626–637.

[29] H. A. Van der Vorst, *The convergence behavior of the preconditioned CG and CG-S in the presence of rounding errors*, in Preconditioning Conjugate Gradient Methods, L. Y. K. O. Axelsson, ed., Springer-Verlag, 1990, pp. 126–136.

[30] X. Zou, I. M. Navon, M. Berger, K. H. Phua, T. Schlick, and F. X. Le Dimet, *Numerical experience with limited-memory quasi-Newton and truncated Newton methods*, SIAM J. Optimization, 3 (1993), pp. 582–608.