The University of Reading

Approximate Gauss-Newton methods for nonlinear least squares problems

S. Gratton[1], A.S. Lawless

Abstract

The Gauss-Newton algorithm is an iterative method regularly used for solving nonlinear least squares problems. It is particularly well-suited to the treatment of very large scale variational data assimilation problems that arise in atmosphere and ocean forecasting. The procedure consists of a sequence of linear least squares approximations to the nonlinear problem, each of which is solved by an `inner' direct or iterative process. In comparison with Newton's method and its variants, the algorithm is attractive because it does not require the evaluation of second-order derivatives in the Hessian of the objective function. In practice the exact Gauss-Newton method is too expensive to apply operationally in meteorological forecasting and various approximations are made in order to reduce computational costs and to solve the problems in real time. Here we investigate the e®ects on the convergence of the Gauss-Newton method of two types of approximation used commonly in data assimilation. Firstly, we examine `truncated' Gauss-Newton methods where the `inner' linear least squares problem is not solved exactly, and secondly, we examine `perturbed' Gauss-Newton methods where the true linearized `inner' problem is approximated by a simpli¯ed, or perturbed, linear least squares problem. We give conditions ensuring that the truncated and perturbed Gauss-Newton methods converge and also derive rates of convergence for the iterations. The results are illustrated by a simple numerical example.

Keywords Nonlinear least squares problems; approximate Gauss-Newton methods; variational data assimilation

1 Introduction

The Gauss-Newton method is a well-known iterative technique used regularly for solving the nonlinear least squares problem (NLSP)

$$\min_x \acute{A}(x) = \frac{1}{2} kf(x)k_2^2; \tag{1}$$

where x is an n-dimensional real vector and f is an m-dimensional real vector function of x [9].

Problems of this form arise commonly from applications in optimal control and ¯ltering and in data ¯tting. As a simple example, if we are given m observed data (t_i (1)

2.1 Statement of the algorithm

We consider the nonlinear least squares problem (NLSP) de¯ned in (1), where we assume that $f : \mathbb{R}^n \to \mathbb{R}^m$ is a nonlinear, twice continuously Fréchet di®erentiable function. We denote the Jacobian of the function f by $J(x) = f'(x)$. The gradient and Hessian of $\phi(x)$ are then given by

$$\nabla\phi(x) = J^T(x)f(x); \tag{2}$$

$$\nabla^2\phi(x) = J^T(x)J(x) + Q(x); \tag{3}$$

where $Q(x)$ denotes the second order terms

$$Q(x) = \sum_{i=1}^{m} f_i(x)\nabla^2 f_i(x). \tag{4}$$

Finding the stationary points of ϕ is equivalent to solving the gradient equation

$$F(x) = \nabla\phi(x) = J^T(x)f(x) = 0. \tag{5}$$

Techniques for treating the NLSP can thus be derived from methods for solving this nonlinear algebraic system.

A common method for solving nonlinear equations of form (5) and hence for solving the NLSP (1) is Newton's method [9]. This method requires the inversion of the full Hessian matrix (3) of function ϕ. For many large scale problems, the second order terms $Q(x)$ of the Hessian are, however, impracticable to calculate and, in order to make the procedure more e±cient, Newton's method is approximated by ignoring these terms. The resulting iterative method is known as the Gauss-Newton algorithm [9] and is de¯ned as follows.

Gauss-Newton Algorithm (GN)

Step 0 : Choose an initial $x_0 \in \mathbb{R}^n$

Step 1 : Repeat until convergence:

Step 1.1 : Solve $J(x_k)^T J(x_k)s_k = -J^T(x_k)f(x_k)$

Step 1.2 : Set $x_{k+1} = x_k + s_k$.

Remarks: We note that at each iteration, Step 1.1 of the method is equivalent to solving the linearized least squares problem

$$\min_s \frac{1}{2}\|J(x_k)s + f(x_k)\|_2^2. \tag{6}$$

We note also that the GN method can be written as a ¯xed-point iteration of the form

$$x_{k+1} = G(x_k) \tag{7}$$

where $G(x) = x - J^+(x)f(x)$ and $J(x)^+ = (J^T(x)J(x))^{-1}J^T(x)$ denotes the Moore-Penrose pseudo-inverse of $J(x)$.

2.2 Convergence of the exact Gauss-Newton method

Su±cient conditions for the convergence of the Gauss-Newton method are known in the case where the normal equations for the linearized least squares problem (6) are solved *exactly* in Step 1:1 at each iteration. We now recall some existing results. The following assumptions are

3 Approximate Gauss-Newton Algorithms

A serious difficulty associated with the use of the Gauss-Newton method in large scale applications, such as data assimilation, is that the linearized least squares problem (6) is computationally too expensive to solve exactly in Step 1.1 of the algorithm at each iteration. The dimensions of the normal matrix equations to be solved in Step 1.1 are often so great that the system coefficients cannot be stored in core memory, even in factored form. Therefore, in order to solve the full nonlinear problem efficiently, in real forecasting time, approximations must be made within the Gauss-Newton procedure.

Two types of approximation are commonly applied. Firstly, the linearized least squares problem (6) is solved only approximately by an `inner' iteration method that is truncated before full accuracy is reached. We refer to this approximate algorithm as the Truncated Gauss-Newton (TGN) method . Secondly, the linearized least squares problem in Step 1.1 is replaced by an approximate, simplified or perturbed, linear problem that can be solved more efficiently in the inner loop. We refer to this algorithm as the Perturbed Gauss-Newton (PGN) method . Here we examine both of these approximate Gauss-Newton methods and also the combined Truncated Perturbed Gauss-Newton (TPGN) method, where both approximations are applied. In the

3.2 Perturbed Gauss-Newton method

For some applications it is desirable to apply a perturbed Gauss-Newton method in which the true Jacobian J

3.3 Truncated Perturbed Gauss-Newton method

In the PGN method, we solve the normal equations in Step 1.1 of the algorithm at each outer iteration k

Inexact Newton Algorithm (IN)

The convergence condition derived in this theorem is less general than that obtained in Theorem 1, which requires a bound only on the spectral radius of the matrix $Q(x)(J^T(x)J(x))^{-1}$ at the ¯xed point $x = x^¤$, rather than on its norm at each iterate x_k. However, the technique used in the proof of Theorem 4 provides a practical test for convergence and is more readily extended to the case of the approximate Gauss-Newton iterations.

4.3 Convergence of the Truncated Gauss-Newton method (I)

We now give a theorem that provides su±cient conditions for the convergence of the truncated Gauss-Newton (TGN) method. It is assumed that the residuals in the TGN method are bounded such that

$$\|r_k\|_2 \leq \bar{\eta}_k \|r\hat{A}(x_k)\|_2; \tag{23}$$

where $\{\bar{\eta}_k\}$ is a nonnegative forcing sequence. The theorem is established by considering the algorithm as an inexact Newton method, as in the proof of Theorem 4.

Theorem 5 *Let assumptions* A1. *and* A2. *hold and let* $r^2\hat{A}(x^¤)$ *be nonsingular. Assume that* $0 \leq \hat{\eta} < 1$ *and select* $\bar{\eta}_k$; $k = 0; 1; \ldots$ *such that*

$$0 \leq \bar{\eta}_k \leq \hat{\eta}$

A1^0. there exists $x^* \in \mathbb{R}^n$ such that $F(x^*) = J^T(x^*)f(x^*) = 0$;

A2^0. the matrix $J(x^*)$ at x^* has full rank n.

We then obtain the following theorem.

Theorem 6 *Let assumptions* A1^0. *and* A2^0. *hold and let* $F^0(x^*) = J(x^*)^T J(x^*) + Q(x^*)$ *be nonsingular. Assume* $0 \le \kappa < 1$. *Then there exists* $\varepsilon > 0$ *such that if* $\|x_0 - x^*\|_2 \le \varepsilon$ *and if*

$$\left\| I - (J(x_k)^T J(x_k) + Q(x_k))(J(x_k)^T J(x_k))^{-1} \right\|_2 \le \gamma_k \le \kappa, \quad k = 0, 1, \ldots, \tag{27}$$

the sequence of perturbed Gauss-Newton iterates $\{x_k\}$ *converges to* x^*.

Proof of Theorem 6 : We can write the PGN method as an IN method by setting

$$\begin{aligned} r_k &= J(x_k)^T f(x_k) - (J(x_k)^T J(x_k) + Q(x_k))(J(x_k)^T J(x_k))^{-1} J(x_k)^T f(x_k) && (28)\\ &= (I - (J(x_k)^T J(x_k) + Q(x_k))(J(x_k)^T J(x_k))^{-1}) J(x_k)^T f(x_k). && (29) \end{aligned}$$

Then, provided the condition (27) holds, we have

$$\|r_k\|_2 \le \kappa \left\| J(x_k)^T f(x_k) \right\|_2; \tag{30}$$

and by Theorem 3 local convergence is guaranteed.

The theorem gives explicit conditions on the perturbed Jacobian J that are sufficient to guarantee the convergence of the perturbed Gauss-Newton method. The requirement is that $J(x_k)^T J(x_k)$ should be a good approximation to the derivative $F^0(x) = J(x)^T J(x) + Q(x)$ of the perturbed gradient equation (14).

4.5 Fixed point of the Perturbed Gauss-Newton method

We now consider how close the solution x^* of the perturbed gradient equation (14) is to the solution x^* of the original NLSP. To answer this question we treat the GN method as a stationary fixed-point iteration of the form (7).

We assume that the GN iteration converges locally to x^* for all x_0 in an open convex set D containing x^* (defined as in Theorem 1) and that $G(x)$ satisfies

$$\|G(x) - G(x^*)\|_2 \le \omega \|x - x^*\|_2; \quad \forall\, x \in D; \quad \text{with } \omega < 1. \tag{31}$$

Then we have the following theorem, which bounds the distance between the solutions of the exact and perturbed iterations.

Theorem 7 *Let assumptions* A1., A2., A1^0. *and* A2^0. *hold and assume* $\theta < 1$. *Let (31) be satisfied and let* $x^* \in D$: *Then*

$$\|x^* - x^*\|_2 \le \frac{1}{1-\omega} \left\| (J^+(x^*) - J^+(x^*)) f(x^*) \right\|_2. \tag{32}$$

Proof of Theorem 7 : We de¯ne $G($

Proof of Theorem 8 : We can write TPGN in the same form as IN by setting

$$\begin{aligned} r_k &= (I - (J(x_k)^T J(x_k) + Q(x_k))(J(x_k)^T J(x_k))^{-1})J(x_k)^T f(x_k) \\ &+ (J(x_k)^T J(x_k) + Q(x_k))(J(x_k)^T J(x_k))^{-1} r_k. \end{aligned} \tag{35}$$

Q

Let c be such that $1 < c < \,¸=¾$. *Then there exists* $" > 0$ *such that, if* $kx_0 \,¡\, x¤k_2 < "$, *the iterates* fx_kg *generated by the Gauss-Newton algorithm converge to* $x¤$. *Additionally, the following inequality holds*

$$kx_k$$

5.2 Convergence of the Truncated Gauss-Newton method (II)

By an extension of Theorem 9, we now establish alternative conditions for the truncated Gauss-Newton (TGN) method to converge. We assume, as previously, that the residuals in the TGN method are bounded such that

$$\|r_k\|_2 \leq \bar{\eta}_k \|J(x_k)^T f(x_k)\|_2; \tag{44}$$

where $\{\bar{\eta}_k\}$ is a nonnegative forcing sequence.

Theorem 11 *Let the conditions of Theorem 9 hold and let c be such that $1 < c < \sigma/\gamma$. Select $\bar{\eta}_k$, $k = 0, 1, \ldots$ to satisfy*

$$0 \leq \bar{\eta}_k \leq \bar{\eta} < \frac{\sigma - c\gamma}{c(\gamma + \beta^2)}, \quad k = 0, 1, \ldots \tag{45}$$

Then there exists $\varepsilon > 0$ such that if $\|x_0 - x^*\|_2 < \varepsilon$, the sequence of truncated Gauss-Newton iterates $\{x_k\}$ satisfying (44) converges to x^*. Additionally, the following inequality holds :

$$\|x_{k+1} - x^*\|_2 \leq \frac{c}{\sigma}(\gamma + \bar{\eta}_k(\gamma + \beta^2))\,\|x_k - x^*\|_2 + C\,\|x_k - x^*\|_2^2; \tag{46}$$

where $C = \dfrac{c\beta\omega}{2\sigma}(1 + \bar{\eta})$.

Proof of Theorem 11 : The proof is by induction. Let us denote by J_0, f_0, J^* and f^* the quantities $J(x_0)$, $f(x_0)$, $J(x^*)$ and $f(x^*)$. From the proof of Theorem 9 (see [5, Theorem 10.2.1]), there exists a positive quantity ε_1 such that, if $\|x_0 - x^*\|_2 < \varepsilon_1$, then $x_0 \in D$, $J_0^T J_0$ is nonsingular, $\|(J_0^T J_0)^{-1}\|_2 < c/\sigma$, and

$$\|x_0 - (J_0^T J_0)^{-1} J_0^T f_0 - x^*\|_2 \leq \frac{c\gamma}{\sigma}\|x_0 - x^*\|_2 + \frac{c\beta\omega}{2\sigma}\|x_0 - x^*\|_2^2. \tag{47}$$

Let

$$\varepsilon = \min\left(\varepsilon_1; \frac{\sigma - c(\gamma + \bar{\eta}(\gamma + \beta^2))}{c\beta\omega(1 + \bar{\eta})}\right); \tag{48}$$

where $\sigma - c(\gamma + \bar{\eta}(\gamma + \beta^2)) > 0$ by (45).

We start from

$$\|J_0^T f_0$$

Gathering the partial results (47) and (51), we obtain

$$\begin{aligned} \|x_1 - x^*\|_2 &= \left\| x_0 - (J_0^T J_0)^{-1} J_0^T f_0 + (J_0^T J_0)^{-1} r_0 - x^* \right\|_2 \\ &\cdot \left\| x_0 - (J_0^T J_0)^{-1} J_0^T f_0 - x^* \right\|_2 + \|r_0\|_2 \left\| (J_0^T J_0)^{-1} \right\| \end{aligned}$$

We remark that Theorem 12 establishes the convergence of the PGN method to the ¯xed point $x^¤$ of the *exact* Gauss-Newton method. At the ¯xed point, the perturbed Jacobian J must, therefore, be such that $J(x^¤)^T f(x^¤) = 0$ in order to be able to satisfy the conditions of the theorem; that is, at the ¯xed point $x^¤$, the null space of $J(x^¤)^T$ must contain $f(x^¤)$. In contrast the convergence results of Theorem 6 only require that a point $x^¤$ exists such that $J(x^¤)^T f(x^¤) = 0$ and $J(x^¤)$ is full rank.

5.4 Convergence of the Truncated Perturbed Gauss-Newton method (II)

In the following theorem we consider the truncated perturbed Gauss-Newton iteration where an approximate Jacobian J is used and the inner linear least squares problem (15) is not solved exactly on each outer step. The residuals in the inner normal equations at each outer iteration are assumed to be bounded such that

$$\|r_k\|_2 \le \bar{\eta}_k \left\| J(x_k)^T f(x_k) \right\|_2; \tag{57}$$

where $\{\bar{\eta}_k\}$ is a nonnegative forcing sequence. Su±cient conditions for the TPGN method to converge are then given as follows.

Theorem 13 *Let the conditions of Theorem 9 hold and let $J(x)$ be an approximation to $J(x)$. Let c be such that $1 < c < \sigma/\gamma$. Assume that $\eta_k \le \hat{\eta} < (\sigma - c\gamma)/(c(\gamma + \beta^2))$ and select $\bar{\eta}_k$, $k = 0, 1, \ldots$ such that*

$$0 \le \bar{\eta}_k \le \left(\eta_k \left\| J(x_k)^T f(x_k) \right\|_2 - \left\| J(x_k)^T J(x_k)(J^+(x_k) - \hat{J}^+(x_k)) f(x_k) \right\|_2\right) \Big/ \left(3 \left\| J(x_k)^T J(x_k)(\hat{J}(x_k)^T \hat{J}(x_k))^{-1} \right\|_2 \left\| \hat{J}(x_k)^T f(x_k) \right\|_2\right), \tag{58}$$

for $k = 0, 1, \ldots$. Then there exists $\varepsilon > 0$ such that if $\|x_0 - x^¤\|_2 < \varepsilon$, the sequence of perturbed Gauss-Newton iterates $\{x_k\}$ satisfying (57) converges to $x^¤$. Additionally, the following inequality holds :

$$\|x_{k+1} - x^¤\|_2 \le \frac{c}{\sigma}(\gamma + \eta_k(\gamma + \beta^2)) \|x_k - x^¤\|_2 + C \|x_k - x^¤\|_2^2; \tag{59}$$

where $C = c\beta^\circ(1 + \hat{\kappa})/(2\sigma)$.

Proof of Theorem 13 : The TPGN iteration takes the form $x_{k+1} = x_k + s_k$, where $s_k = -\hat{J}^+(x_k) f(x_k) + (\hat{J}(x_k)^T \hat{J}(x_k))^{-1} r_k$. Therefore, using the notation of Theorem 11, we may consider the TPGN method as a truncated Gauss-Newton method with the residual de¯ned as

$$\hat{r}_k = J(x_k)^T J(x_k)(J^+(x_k) - \hat{J}^+(x_k)) f(x_k) + J(x_k)^T J(x_k)(\hat{J}(x_k)^T \hat{J}(x_k))^{-1} r_k. \tag{60}$$

Then, provided the condition (57) holds, we have

$$\begin{aligned} \|\hat{r}_k\|_2 &\le \left\| J(x_k)^T J(x_k)(J^+(x_k) - \hat{J}^+(x_k)) f(x_k) \right\|_2 + \\ &\quad \left\| J(x_k)^T J(x_k)(\hat{J}(x_k)^T \hat{J}(x_k))^{-1} \right\|_2 \left\| \bar{\eta}_k J(x_k)^T f(x_k) \right\|_2 \\ &\le \eta_k \left\| J(x_k)^T f(x_k) \right\|_2. \end{aligned} \tag{61}$$

The conclusion then follows from Theorem 11.

We remark that to ensure $\bar{}_k \; , \; 0$, we require that the relation given by equation (55) holds. This is simply the condition of Theorem 12 that guarantees the convergence of the PGN method in the case where the inner loop is solved exactly without truncation.

Theorem 13 gives conditions for the truncated perturbed Gauss-Newton method to converge to the $\bar{}$xed point $x^{¤}$ of the *exact* Gauss-Newton method, and is therefore more restrictive than the theorem developed in Section 4. Here the allowable form of the perturbed Jacobian is constrained to satisfy $J(x^{¤})^T f(x^{¤}) = J(x^{¤})^T f(x^{¤}) = 0$ in order that the conditions of the theorem may be met. The theorem does, however, establish that the method converges with rates of convergence higher than linear in certain cases. These cases are discussed in the next subsection.

5.5 Rates of convergence of the approximate Gauss-Newton methods

>From Theorems 11, 12 and 13, the expected convergence rates of the approximate Gauss-Newton methods may be established for various cases. The convergence rates are shown in (46), (56) and (59) for the truncated Gauss-Newton, the perturbed Gauss-Newton and the truncated perturbed Gauss-Newton methods, respectively. These rates are dependent on the parameters $¼; \; ,$ and ®, de$\bar{}$ned as in Theorem 9, and can be contrasted directly with the convergence rates of the exact Gauss-Newton method, given by (37). We observe the following.

1. *Linear convergence.* The theorems show that in general the GN, TGN, PGN and TPGN methods converge linearly. In comparison with the exact GN algorithm, we see that the price paid for the inaccurate solution of the linear least squares problem in the inner step of the approximate methods is a degradation of the local linear rate of convergence.

2. *Super-linear convergence.* As previously noted, if $¼ = 0$, which holds, for example, in the zero-residual case where $f(x^{¤}) = 0$, the convergence of the exact GN method is quadratic. In this same case, if $¼ = 0$ and if the forcing sequence $f\bar{}_k g$ satis$\bar{}$es $\lim_{k \, ! \, +1} \bar{}_k = 0$, then the convergence rates of the TGN and TPGN methods are super-linear. For the PGN method to converge super-linearly in this case, the sequence $f´_k g$ must satisfy $\lim_{k \, ! \, +1} ´_k = 0$.

3. *Quadratic convergence.* From the proof of Theorem 11, we see that the convergence of the TGN method is quadratic if $¼ = 0$ and if the normal equation residual is such that

$$kr_k k_2 \; ´ \; \left\| J(x_k)^T J(x_k) s_k + J(x_k)^T f(x_k) \right\|_2 \; \cdot \; C_1 \left\| J(x_k)^T f(x_k) \right\|_2^2;$$

   for some positive constant C_1. Similarly, in the case $¼ = 0$, the PGN method converges quadratically if

$$\left\| J(x_k)^T J(x_k) \right)^3 J^+(x_k) \; ¡ \; J^+(x_k) \right)´ f(x_k) \Big\|_2 \; \cdot \; C_2 \left\| J(x_k)^T f(x_k) \right\|_2^2;$$

   as does the TPGN method in this case if

$$\left\| (J(x_k)^T J(x_k))((J(x_k)^+ \; ¡ \; J(x_k)^+) f(x_k) + (J(x_k)^T J(x_k))^{¡1} r_k) \right\|_2 \; \cdot \; C_3 \left\| J(x_k)^T f(x_k) \right\|_2^2;$$

   for positive constants $C_2; C_3$.

4. *E®ect of nonlinearity.*

of the nonlinearity and the residual size in [redacted] e therefore that, in order to guarantee convergence of the approximat [redacted] arized equation must be solved more accurately when the proble [redacted] when there is a large residual at the optimal.

In Section 6 we give numerical results demon [redacted] behaviour of the approximate Gauss-Newton methods. The rates of [redacted] oximate methods are also illustrated for various cases.

5.6 Summary

In this section we have established theory ensuring local convergence of the Gauss-Newton, the truncated Gauss-Newton, the perturbed Gauss-Newton and the truncated perturbed Gauss-Newton methods based on the theory of [5] for exact Gauss-Newton methods. The conditions for convergence derived in this section are more restrictive than those of Section 4, but enable the rates of convergence to be established. Numerical examples illustrating the results for the three approximate Gauss-Newton methods are shown in the next section.

6 Numerical example

We examine the theoretical results of Sections 4 and 5 using a simple initial value problem discretized by a second-order Runge-Kutta scheme. The example is based on that in [7, Chapter 4] and is used because it provides a clear way of producing a perturbed Jacobian. We consider the ordinary di®erential equation

$$\frac{dz}{dt} = z^2; \tag{62}$$

where $z = z(t)$ and $z(0) = z_0$ is given. Application of a second order Runge-Kutta scheme gives a discrete nonlinear model

$$x^{n+1} = x^n + (x^n)^2 ¢ t + (x^n)^3 ¢ t^2 + \frac{1}{2}(x^n)^4 ¢ t^3; \tag{63}$$

where $¢t$ denotes the model time step and $x^n ¼ z(t_n)$ at time $t_n = n¢t$. We de¯ne a least squares problem

$$\min_{x^0} Á(x) = \frac{1}{2}(x^0 \; i \; y [redacted] \tag{64}$$

subject to (63), where $y^0; \; y^1$ are values of obse [redacted] This is of the same form as (1), with

$$f = \begin{bmatrix} x^0 \\ x^1 \end{bmatrix} \tag{65}$$

Then the Jacobian of f is given b [redacted]

$$J \begin{bmatrix} 1 & 0 \\ y^1 & 1 \end{bmatrix}$$

/F9 10.91 Tf 9.090 TD [(t)] TJ/F [redacted]

Table 1: Perfect observations, exact Jacobian

²	Iterations	Error	Gradient
0.00	5	0.000000e+00	0.000000e+00
0.25	20	9.015011e-14	1.364325e-13
0.50	37	7.207568e-13	1.092931e-12
0.75	84	2.246647e-12	3.407219e-12
0.90	210	8.292034e-12	1.257587e-11
0.95	401	1.857048e-11	2.816403e-11
1.00	1000	3.143301e-04	4.765072e-04
1.05	431	2.652062e-02	3.880614e-02
1.10	231	5.357142e-02	7.568952e-02
1.15	163	8.101821e-02	1.106474e-01
1.20	130	1.093852e-01	1.444877e-01
1.25	112	1.394250e-01	1.781241e-01

We now use this example to test some of the theorems we have derived in Section 4. For the experiments the true value of x^0 is set to be $¡2.5$ and we begin with an initial estimate of $¡2.3$. Observations are generated using the truth at the initial time t_0 and using the discrete numerical model (63) to calculate the `truth' at time t_1. The time step is set to be $¢t = 0.5$. We begin by testing the convergence of the TGN algorithm.

6.1 Truncated Gauss-Newton method - numerical results

Table 2: Imperfect observations, exact Jacobian

^{2}	Iterations	Error	Gradient
0.00	10	4.440892e-15	7.778500e-15
0.25	17	9.503509e-14	1.806853e-13
0.50	32	6.181722e-13	1.176347e-12
0.75	66	1.671552e-12	3.180605e-12
0.90	128	4.250822e-12	8.088735e-12
0.95	181	6.231016e-12	1.185694e-11
1.00	359	1.052936e-11	2.003732e-11
1.05	157	6.324736e-02	1.093406e-01
1.10	116	8.697037e-02	1.452842e-01
1.15	93	1.103473e-01	1.783861e-01
1.20	79	1.336149e-01	2.092708e-01
1.25	69	1.570351e-01	2.384890e-01

If we apply the second order Runge-Kutta scheme to this equation, we obtain

$$\pm x^{n+1} = (1 + 2x^n \¢ t + 3(x^n)^2 \¢ t^2 + 3(x^n)^3 \¢ t^3 + \frac{5}{2}(x^n)^4 \¢ t^4 + (x^n)^5 \¢ t^5)\pm x^n: \qquad (71)$$

Thus for the example we obtain the perturbed Jacobian

$$\mathcal{J}(x^0) = \begin{pmatrix} 1 \\ 1 + 2x^0 \¢ t + 3(x^0)^2 \¢ t^2 + 3(x^0)^3 \¢ t^3 + \frac{5}{2}(x^0)^4 \¢ t^4 + (x^0)^5 \¢ t^5 \end{pmatrix}: \qquad (72)$$

Using this perturbed Jacobian we apply the PGN algorithm on our example, where on each iteration we con¯rm that the su±cient condition (27) is satis¯ed. For this example we ¯nd that that the second order terms $\mathcal{Q}$ are given by

$$\begin{aligned} \mathcal{Q}(x^0) \quad = \quad & (x^0 + (x^0)^2 \¢ t + (x^0)^3 \¢ t^2 + \frac{1}{2}(x^0)^4 \¢ t^3 \ ¡ \ y^1) \¢ \\ & (2 \¢ t + 6x^0 \¢ t^2 + 9(x^0)^2 \¢ t^3 + 10(x^0)^3 \¢ t^4 + 5(x^0)^4 \¢ t^5): \end{aligned} \qquad (73)$$

For the case in which we have perfect observations we ¯nd that (27) is satis¯ed on each iteration and the PGN method converges to the true solution in 18 iterations. When error is added on to the observations, as in the previous section, the PGN method converges in 9 iterations and again we ¯nd that the condition for convergence is always satis¯ed. This time the converged solution is not the same as that of the exact Gauss-Newton method. The solution di®ers from the true solution $x_0 = \ ¡ 2{:}5$ by approximately 0:01.

In order to examine a case in which the su±cient condition (27) is not satis¯ed on each iteration, we change the time step to $\¢ t = 0{:}6$, keeping all other parameters of the problem the same as before. For the case of perfect observations the PGN converges to the correct solution in 23 iterations, compared to 5 iterations for the exact GN and 6 iterations for the Newton method. We ¯nd that the condition for convergence is satis¯ed on each iteration, with the maximum value of the left hand side of (27) reaching 0:994. However, when error is present on the observed values, the convergence condition fails by the second iteration and we ¯nd that the PGN fails to converge in 1000 iterations. For this case the exact GN, using the true Jacobian, converges to the correct solution in 8 iterations.

6.3 Truncated Perturbed Gauss-Newton method - numerical results

Finally in this section we consider the case in which the perturbed Gauss-Newton method is also truncated. Following the same method as in the previous two sections, we solve on each iteration the approximate equation

$$\mathcal{J}(x_k^0)^T \mathcal{J}(x_k^0) s_k = \ ¡ \mathcal{J}(x_k^0)^T f(x_k^0) + r_k; \qquad (74)$$

where we choose the residual r_k. We choose

$$r_k = \ ² \left(\frac{\tilde{A} \ ¡ \ j1 \ ¡ \ (\mathcal{J}(x_k)^T J(x_k) + \mathcal{Q}(x_k))(\mathcal{J}(x_k)^T \mathcal{J}(x_k))^{\ ¡ 1} j}{j(\mathcal{J}(x_k)^T J(x_k) + \mathcal{Q}(x_k))(\mathcal{J}(x_k)^T \mathcal{J}(x_k))^{\ ¡ 1} j} \right) \ j r \acute{A}(x_k^0) j; \qquad (75)$$

where ²

Table 3: Imperfect observations, inexact Jacobian

^{2}	Iterations	Error	Residual
0.00	21	8.215650e-14	6.693951e-14
0.25	33	4.911627e-13	4.007662e-13
0.50	56	1.217249e-12	9.930633e-13
0.75	121	3.732126e-12	3.044658e-12
0.90	306	1.105871e-11	9.021988e-12
0.95	596	2.444178e-11	1.993989e-11
1.00	1000	1.260007e-01	9.382085e-02
1.05	90	1.714365e+00	1.765471e+00
1.10	53	1.842029e+00	1.934063e+00
1.15	36	1.940084e+00	2.069636e+00
1.20	25	2.019233e+00	2.184031e+00
1.25	23	2.085381e+00	2.283791e+00

show the convergence results for the TPGN method using various levels of truncation. The third column now shows the di®erence between the TPGN solution and the exact Newton method applied to the perturbed problem, and the fourth column gives the residual in the perturbed

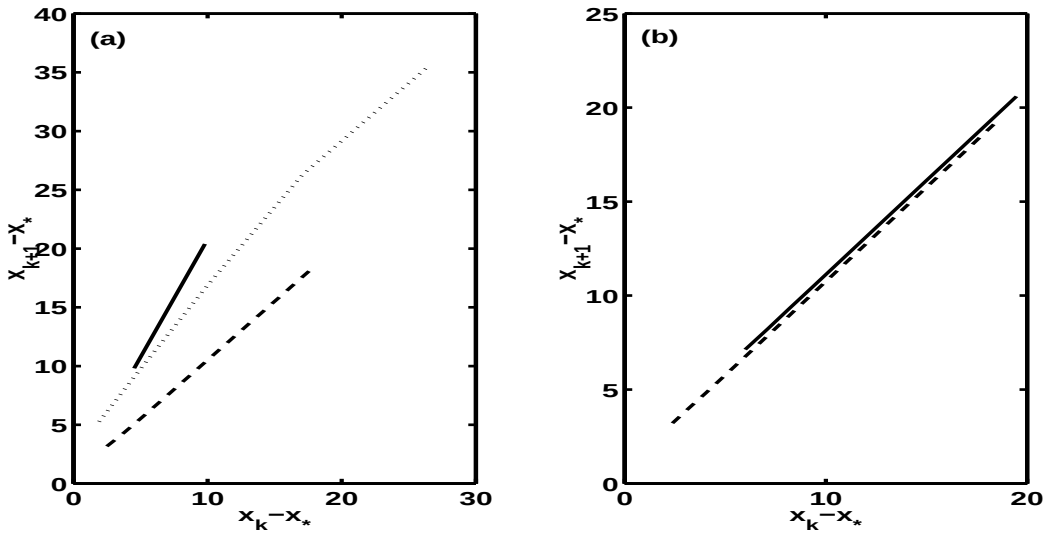

Figure 1: Convergence rates for the cases of (a) exact Jacobian and (b) perturbed Jacobian for the zero residual case. The solid line is for no truncation, the dashed line for constant truncation and the dotted line in plot (a) is for variable truncation.

7 Conclusions

We have described here three approximate Gauss-Newton methods, the truncated, the perturbed and the truncated perturbed Gauss-Newton methods, for solving the nonlinear least squares problem (NLSP). We have derived conditions for the convergence of these approximate methods by treating them as inexact Newton methods, following the theory of [4]. More restricted convergence results, including rates of convergence, have also been derived for the approximate methods by extending the theory of [5] for the exact Gauss-Newton method. In practice, the approximate Gauss-Newton methods are used to treat very large data assimilation problems arising in atmosphere and ocean modelling and prediction. The convergence properties of these algorithms have not previously been investigated. We show by a simple numerical example that the bounds established by the theory are precise, in a certain sense, and that the approximate methods are convergent if the conditions of the theory hold.

References

[1] Åke Björck. *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, 1996.

[2] E. Catinas. Inexact perturbed Newton methods and applications to a class of Krylov solvers. *Journal of Optimization Theory and Applications*, 3:543{571, 2001.

[3] P. Courtier, J.N. Thepaut, A. Hollingsworth. A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120:1367{1387, 1994.

[4] R.S. Dembo, S.C. Eisenstat and T. Steihaug. Inexact Newton methods. *SIAM Journal on Numerical Analysis*, 19:400{408, 1982.

[5] J.E. Dennis and R.B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cli®s, NJ, 1983.

[6] S. Gratton. *Outils th¶eoriques d'analyse du calcul a pr¶ecision ¯nie*, Institut National Polytechnique de Toulouse, PhD Dissertation, Note: TH/PA/98/30, 1998.

[7] A.S. Lawless. *Development of linear models for data assimilation in numerical weather prediction*, The University of Reading, Department of Mathematics, PhD Thesis, 2001.

[8] A.S. Lawless, S. Gratton, N.K. Nichols. An investigation of incremental 4D-Var using non-tangent linear models. *Quarterly Journal of the Royal Meteorological Society* (to appear).

[9] J.M. Ortega and W.C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[10] A. Ostrowsky. *Solution of Equations and Systems of Equations*, 2nd Edition, Academic Press, New York, 1966.

[11] V. Pereyra. Iterative methods for solving nonlinear least squares problems. *SIAM Journal on Numerical Analysis*, 4:27{36, 1967.

[12] J.S. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*, Springer, New York/Berlin, 1980.

[13] P-Å. Wedin. On the Gauss-Newton method for the nonlinear least-squares problems. Institute for Applied Mathematics, Stockolm, Sweden, Working Paper 24, 1974.