INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2005; **47**:1129–1135
Published online 29 December 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/fld.851

Approximate iterative methods for variational data assimilation

A. S. Lawless[1,*,†], S. Gratton[2] and N. K. Nichols[1]

[1]*Department of Mathematics, The University of Reading, Whiteknights, P.O. Box 220, Reading, Berkshire, RG6 6AX, U.K.*
[2]*CERFACS, 42 Avenue Gustave Coriolis, 31057, Toulouse, Cedex, France*

SUMMARY

The problem of four-dimensional variational data assimilation can be considered as a non-linear least squares problem. A common method for solving it is equivalent to a Gauss–Newton method. However, approximations must be made in order to solve the problem efficiently. In this paper we examine two approximations to the Gauss–Newton method and their effect on the data assimilation problem. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: Gauss–Newton; data assimilation

1. INTRODUCTION

Data assimilation is the name given to the process of combining observed data with a numerical model to produce the best estimate of the state of a system. It is used for example in numerical weather prediction to obtain the initial conditions from which to run a weather forecasting model. Modern techniques of data assimilation include four-dimensional variational data assimilation (4D-Var), which uses a sequence of observations over a given time window to estimate the best model trajectory within the window. This method considers the data assimilation problem as a non-linear least squares problem, which is solved using an iterative process. However, the problem is very large and so approximations must be made to ensure an efficient solution process. In this paper we consider the effect of some of these approximations on the convergence of the iteration process. In Section 2 we introduce the concept of four-dimensional variational data assimilation. Section 3 introduces the Gauss–Newton method for solving a non-linear least squares problem, which is equivalent to a method commonly used in 4D-Var. In Section 4 we illustrate approximations of the Gauss–Newton method using a

*Correspondence to: A. S. Lawless, Department of Mathematics, The University of Reading, P.O. Box 220, Whiteknights, RG6 6AX, Reading, U.K.
†E-mail: a.s.lawless@reading.ac.uk

Contract/grant sponsor: Natural Environment Research Council

Received 27 April 2004
Revised 4 November 2004
Accepted 18 November 2004

simple 4D-Var system and state some theoretical results which explain the numerical results. Finally we draw some conclusions in Section 5.

2. FOUR-DIMENSIONAL VARIATIONAL DATA ASSIMILATION

The aim of four-dimensional variational data assimilation is to find a model state $\mathbf{x}_0$ at an initial time t_0, such that the distance between the trajectory of the model and a set of observations $\mathbf{y}_i^o$ at times $t_i, i = 0, 1, \ldots, p$, is minimized, subject to $\mathbf{x}_0$ remaining close to a prior estimate $\mathbf{x}^b$, also known as a *background* field. Mathematically this can be expressed as a non-linear least squares problem, for which we wish to find the model state $\mathbf{x}_0$ that minimizes the objective function

$$\mathcal{J}[\mathbf{x}_0] = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^{\mathrm{T}}\mathbf{B}_0^{-1}(\mathbf{x}_0 - \mathbf{x}^b) + \frac{1}{2}\sum_{i=0}^{p}(H_i[\mathbf{x}_i] - \mathbf{y}_i^o)^{\mathrm{T}}\mathbf{R}_i^{-1}(H_i[\mathbf{x}_i] - \mathbf{y}_i^o) \tag{1}$$

subject to $\mathbf{x}_i$ satisfying the discrete non-linear model $\mathbf{x}_i = S(t_i, t_0, \mathbf{x}_0)$, where $S(t_i, t_0, \mathbf{x}_0)$ is the solution operator of the non-linear model. The matrices $\mathbf{B}_0$ and $\mathbf{R}_i$ are the background error and observation error covariance matrices, respectively, and H_i is an operator that maps the model field to observation space.

For the problem of numerical weather prediction the state vector is very large, having the order of 10^7 elements. Thus, efficient methods must be found to minimize function (1). A common way of treating this problem is by a technique known as the *incremental method* [1], which we have previously shown to be equivalent to applying the Gauss–Newton iteration to minimize a non-linear least squares function [2].

3. GAUSS–NEWTON ITERATION

To introduce the Gauss–Newton method we consider a general non-linear least squares problem of the form

$$\min_{\mathbf{x}} \mathcal{J}(\mathbf{x}) = \frac{1}{2}\|\mathbf{f}(\mathbf{x})\|_2^2 = \frac{1}{2}\mathbf{f}(\mathbf{x})^{\mathrm{T}}\mathbf{f}(\mathbf{x}) \tag{2}$$

with $\mathbf{x} \in \mathbb{R}^n$ [3]. We can write the 4D-Var objective function (1) in this form by putting $\mathbf{f}(\mathbf{x}) = \mathbf{C}^{-1/2}\hat{\mathbf{d}}$, where $\mathbf{C}^{-1/2}$ represents the symmetric square root of $\mathbf{C}^{-1}$ and

$$\hat{\mathbf{d}}(\mathbf{x}_0) = -\begin{pmatrix} \mathbf{x}_0 - \mathbf{x}^b \\ H_0[\mathbf{x}_0] - \mathbf{y}_0^o \\ \vdots \\ H_p[\mathbf{x}_p] - \mathbf{y}_p^o \end{pmatrix}, \quad \mathbf{C}^{-1} = \begin{pmatrix} \mathbf{B}_0^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}^{-1} \end{pmatrix} \quad \text{with} \quad \mathbf{R} = \mathrm{diag}\{\mathbf{R}_i\} \tag{3}$$

We assume that $\mathcal{J}(\mathbf{x})$ is twice continuously differentiable in an open convex set $D \in \mathbb{R}^n$ and that the minimization problem (2) has a unique solution $\mathbf{x}^* \in D$. Then the first and second

derivatives of $\mathscr{J}(\mathbf{x})$ are given by

$$\nabla \mathscr{J}(\mathbf{x}) = \mathbf{J}^{\mathrm{T}}\mathbf{f}(\mathbf{x}) \tag{4}$$

$$\nabla^2 \mathscr{J}(\mathbf{x}) = \mathbf{J}^{\mathrm{T}}\mathbf{J} + Q(\mathbf{x}) \tag{5}$$

where $\mathbf{J} = \mathbf{J}(\mathbf{x})$ is the Jacobian of $\mathbf{f}(\mathbf{x})$ and $Q(\mathbf{x})$ contains terms involving second derivatives of $\mathbf{f}(\mathbf{x})$. The Gauss–Newton iteration for solving (2) is an approximation to a Newton iteration in which these second-order terms are neglected. Thus we obtain the iterative procedure (GN)

$$\text{Solve} \quad (\mathbf{J}^{\mathrm{T}}\mathbf{J})\delta\mathbf{x}^{(k)} = -\mathbf{J}^{\mathrm{T}}\mathbf{f}(\mathbf{x}^{(k)}) \tag{6}$$

$$\text{Set} \quad \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \delta\mathbf{x}^{(k)} \tag{7}$$

Sufficient conditions for the convergence of the exact Gauss–Newton method (GN) to the solution of the non-linear least squares problem (2) have been derived by different authors. From Reference [4] we know that local convergence to $\mathbf{x}^*$ is guaranteed if $||(\mathbf{J}(\mathbf{x}^*)^{\mathrm{T}}\mathbf{J}(\mathbf{x}^*))^{-1}Q(\mathbf{x}^*)||_2 < 1$. Other sufficient conditions can be found in Reference [3], which also gives a rate of convergence, and in Reference [5], which provides a proof from a geometrical point of view.

For the data assimilation problem it is not normally possible to solve (6) directly and so the solution is found by an iterative minimization of the function $\| \mathbf{J}\delta\mathbf{x} + \mathbf{f} \|_2$. We refer to this procedure as the inner minimization. We note also from (3) that the definition of $\mathbf{f}(\mathbf{x})$ implicitly includes an integration of the non-linear model to calculate the value of $\hat{\mathbf{d}}$ as a function of $\mathbf{x}$ at each time step. Hence, the Jacobian $\mathbf{J}$ includes an integration of the linearization of this non-linear model.

4. APPLICATION TO DATA ASSIMILATION

In practice, the data assimilation problem is very large and approximations must be made to ensure an efficient solution method. Two approximations that are commonly made are to solve the inner minimization problem inexactly and to replace the true linearized model with an approximation to it. We investigate these approximations in simple assimilation experiments.

4.1. Assimilation system

The model we use to test approximations to the Gauss–Newton method is a one-dimensional non-linear shallow water system for the flow of a fluid over an obstacle in the absence of rotation. The model equations are given by

$$\frac{Du}{Dt} + \frac{\partial\phi}{\partial x} = -g\frac{\partial\bar{h}}{\partial x}, \qquad \frac{1}{\phi}\frac{D\phi}{Dt} + \frac{\partial u}{\partial x} = 0 \tag{8}$$

with $D/Dt = \partial/\partial t + u\partial/\partial x$. In these equations $\bar{h} = \bar{h}(x)$ is the height of the bottom orography, u is the velocity of the fluid and $\phi = gh$ is the geopotential, where g is the constant surface

gravitational acceleration and $h>0$ the depth of the fluid above the orography. The problem is defined on the domain $x \in [0, L]$, with periodic boundary conditions, and we let $t \in [0, T]$.

The model is discretized using a two-time-level, semi-implicit, semi-Lagrangian scheme. Further details of the numerics can be found in Reference [6]. We use a periodic domain of 1000 grid points, with a spacing $\Delta x = 0.01m$, so that $x \in [0m, 10m]$. The model time step is taken to be 9.2×10^{-3} s and we consider an assimilation over a window of 100 time steps.

In order to test the assimilation we run 'identical twin' experiments, in which observations are generated from a run of the model defined to be the truth. These observations are then assimilated using 4D-Var, starting from an incorrect prior estimate. The inner minimization is performed using a conjugate gradient method and is considered to have converged when the relative change in the objective function between iterations is reduced by a factor γ. Using such a system we test the effect of approximations to the Gauss–Newton method when applied to the assimilation problem.

4.2. Truncation of inner mininimization

The first approximation that is often made is to perform an inexact solution of the inner problem (6). Intuitively we may expect that this is a reasonable approximation, since each successive inner problem is only an approximation to the true problem and so we may want to solve these only to within the accuracy of that approximation. This is equivalent to applying a *truncated Gauss–Newton iteration* (TGN), in which the first step (6) of the GN method is solved approximately, such that the solution $\delta\mathbf{x}^{(k)}$ satisfies

$$(\mathbf{J}^{\mathrm{T}}\mathbf{J})\delta\mathbf{x}^{(k)} = -\mathbf{J}^{\mathrm{T}}\mathbf{f}(\mathbf{x}^{(k)}) + \mathbf{r}^{(k)} \tag{9}$$

where $\mathbf{r}^{(k)}$ is the residual due to the inexact solution procedure.

To investigate the effect of this approximation we run an assimilation experiment for four iterations of the Gauss–Newton method, in which we have perfect observations of all variables at each time step. We compare two experiments, one in which the convergence parameter γ is held constant at a value of 10^{-4} and one in which it is varied between Gauss–Newton iterations, from 10^{-2} on the first iteration, decreasing by a factor of 10 on each successive iteration. The convergence of the objective function and its gradient is shown in Figure 1. We see that by using a variable convergence criterion with greater truncation on earlier iterations we are able to obtain faster overall convergence. An examination of the two solutions from these experiments shows that they both approximate the truth to within the same accuracy.

By adapting the theory from Reference [7] we may prove a new result which shows that the TGN algorithm converges to the true solution $\mathbf{x}^*$ of the non-linear problem, provided that the truncation residual $\mathbf{r}^{(k)}$ is small enough. We obtain the following theorem, which is proved in Reference [8].

Theorem 1

Assume that $\hat{\eta} < 1$ and that on each iteration the Gauss–Newton method is truncated with $\| \mathbf{r}^{(k)} \|_2 \leqslant \eta_k \| \mathbf{J}^{\mathrm{T}}\mathbf{f}(\mathbf{x}^{(k)}) \|_2$, where

$$\eta_k \leqslant \frac{\hat{\eta} - \| (\mathbf{J}^{\mathrm{T}}(\mathbf{x}^{(k)})\mathbf{J}(\mathbf{x}^{(k)}))^{-1} Q(\mathbf{x}^{(k)}) \|_2}{1 + \| (\mathbf{J}^{\mathrm{T}}(\mathbf{x}^{(k)})\mathbf{J}(\mathbf{x}^{(k)}))^{-1} Q(\mathbf{x}^{(k)}) \|_2} \tag{10}$$

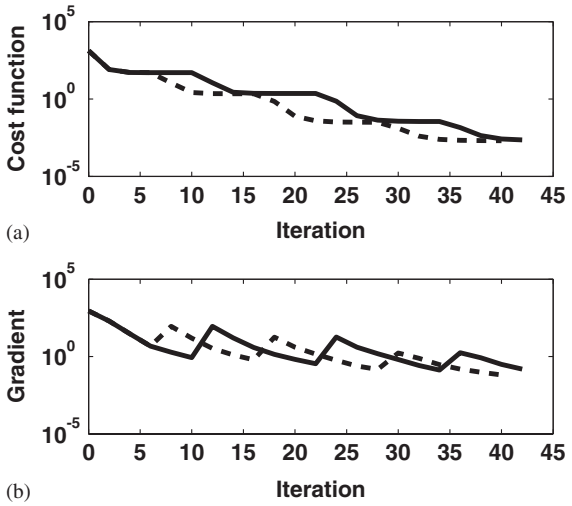

Figure 1. Comparison of convergence of: (a) objective function and (b) its gradient, for a constant convergence criterion (solid line) and a variable criterion (dashed line).

Then there exists $\varepsilon > 0$ such that, if $\| \mathbf{x}_0 - \mathbf{x}^* \|_2 < \varepsilon$, the TGN iteration converges to the solution $\mathbf{x}^*$ of the non-linear least squares problem (2).

We note that (10) provides a bound on the maximum truncation possible on each iteration of the TGN algorithm in order to ensure convergence.

4.3. *Approximation of linear model*

A second approximation we may wish to make in 4D-Var is to replace the true linearization of the non-linear model with an approximation to it. In the context of the Gauss–Newton iteration, this is equivalent to replacing the Jacobian $\mathbf{J}(\mathbf{x})$ with an approximation $\tilde{\mathbf{J}}(\mathbf{x})$. Thus, we obtain the *perturbed Gauss–Newton iteration* (PGN), in which the first step (6) of the method is replaced by

$$\text{Solve} \quad (\tilde{\mathbf{J}}^{\mathrm{T}}\tilde{\mathbf{J}})\delta\mathbf{x}^{(k)} = -\tilde{\mathbf{J}}^{\mathrm{T}}\mathbf{f}(\mathbf{x}^{(k)}) \tag{11}$$

where $\tilde{\mathbf{J}}$ is the perturbed Jacobian.

To test the effect of such an approximation we perform an assimilation using the same case as in Section 4.2, but where we add noise to the observations with a maximum amplitude of 5%. Assimilations are performed with an exact and approximate linear model, as derived previously in Reference [6] and used for assimilation experiments in Reference [2]. In Figure 2 we show the errors in the u field at the centre of the time window in comparison with the true solution and the difference between the two solutions. We find that both solutions agree with the truth to within the accuracy of the observations and the difference between them is two orders of magnitude less than this.

Again it is possible to provide a theoretical understanding of these results by an analysis of the PGN method. For a particular application of this method to the 4D-Var objective function

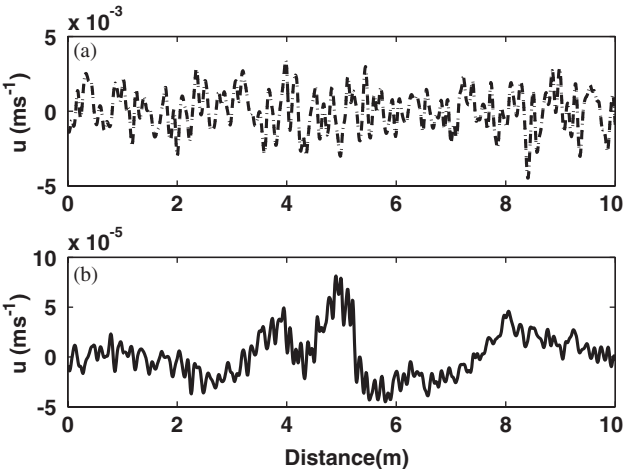

Figure 2. Plot (a) shows the error in the u field in the centre of time window for the assimilation with true linearization (dotted line) and approximate linearization (dashed line). Plot (b) shows the difference between the two solutions.

the authors of Reference [9] derive a sufficient condition for the iteration to converge, under the assumption that derivatives of the approximate Jacobian can be neglected. We propose a new theorem by considering the PGN method as a way of solving the approximate normal equations

$$\tilde{\mathbf{J}}(\tilde{\mathbf{x}}^*)^{\mathrm{T}}\mathbf{f}(\tilde{\mathbf{x}}^*) = 0 \tag{12}$$

Then based on the work of References [7, 10] we can prove that the PGN algorithm converges to a solution of (12).

Theorem 2
Let the first derivative of $\tilde{\mathbf{J}}(\mathbf{x})^{\mathrm{T}}\mathbf{f}(\tilde{\mathbf{x}})$ be written as

$$\tilde{\mathbf{J}}(\mathbf{x})^{\mathrm{T}}\mathbf{J}(\mathbf{x}) + \tilde{Q}(\mathbf{x}) \tag{13}$$

where $\tilde{Q}(\mathbf{x})$ represents second-order terms arising from the derivative of $\tilde{\mathbf{J}}(\mathbf{x})$. Suppose that on each iteration k of the PGN method

$$||(\tilde{\mathbf{J}}(\mathbf{x}^{(k)})^{\mathrm{T}}\tilde{\mathbf{J}}(\mathbf{x}^{(k)}) - \tilde{\mathbf{J}}(\mathbf{x}^{(k)})^{\mathrm{T}}\mathbf{J}(\mathbf{x}^{(k)}) - \tilde{Q}(\mathbf{x}^{(k)}))(\tilde{\mathbf{J}}(\mathbf{x}^{(k)})^{\mathrm{T}}\tilde{\mathbf{J}}(\mathbf{x}^{(k)}))^{-1}||_2 \leqslant \hat{\eta} < 1 \tag{14}$$

Then there exists $\varepsilon > 0$ such that, if $\| \mathbf{x}_0 - \tilde{\mathbf{x}}^* \|_2 \leqslant \varepsilon$, the PGN iteration converges to the solution $\tilde{\mathbf{x}}^*$ of the perturbed problem (12).

A full proof of this theorem may be found in Reference [8], where the distance between the fixed points $\tilde{\mathbf{x}}^*$ and $\mathbf{x}^*$ is also discussed.

5. CONCLUSIONS

In order to use 4D-Var for data assimilation in any realistic system it is necessary to make some approximations. In this work we have examined two such approximations, truncation of the inner minimization and approximation of the linear model. Numerical results with a simple model showed that the assimilation system may still be valid when such approximations are included. By examining the underlying algorithm in the context of approximations to the Gauss–Newton iteration we have been able to provide theoretical results which support the numerical results. Preliminary results on the convergence of the truncated and perturbed Gauss–Newton iterations have been presented. Proofs of the theorems are not given here, but full details are provided in Reference [8].

ACKNOWLEDGEMENTS

Support for this research was provided by the Natural Environment Research Council through the NERC Centre of Excellence in Data Assimilation (DARC).

REFERENCES

1. Courtier P, Thepaut J-N, Hollingsworth A. A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society* 1994; **120**:1367–1387.
2. Lawless AS, Gratton S, Nichols NK. An investigation of incremental 4D-Var using non-tangent linear models. *Quarterly Journal of the Royal Meteorological Society*, to appear.
3. Dennis JE, Schnabel R. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall: Englewood Cliffs, NJ, 1983.
4. Ortega JM, Rheinboldt WC. *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press: New York, 1970.
5. Wedin P-Å. On the Gauss–Newton method for the nonlinear least-squares problem. *Working Paper 24*, Institute for Applied Mathematics, Stockholm, Sweden, 1974.
6. Lawless AS, Nichols NK, Ballard SP. A comparison of two methods for developing the linearization of a shallow water model. *Quarterly Journal of the Royal Meteorological Society* 2003; **129**:1237–1254.
7. Dembo RS, Eisenstat SC, Steihaug T. Inexact Newton methods. *SIAM Journal on Numerical Analysis* 1982; **19**(2):400–408.
8. Gratton S, Lawless AS, Nichols NK. Approximate Gauss–Newton methods for nonlinear least squares problems. *Numerical Analysis Report 9/04*, Department of Mathematics, The University of Reading, 2004.
9. Veersé F, Thépaut J-N. Multiple-truncation incremental approach for four-dimensional variational data assimilation. *Quarterly Journal of the Royal Meteorological Society* 1998; **124**:1889–1908.
10. Cătinaş E. Inexact perturbed Newton methods and applications to a class of Krylov solvers. *Journal of Optimization Theory and Applications* 2001; **108**(3):543–571.

