Approximate Gauss–Newton methods for optimal state estimation using reduced-order models

A. S. Lawless[1], N. K. Nichols[1, *, †], C. Boess[2] and A. Bunse-Gerstner[2]

[1]*Department of Mathematics, University of Reading, P.O. Box 220, Reading RG6 6AX, U.K.*
[2]*Zentrum fuer Technomathematik, Universitaet Bremen, Postfach 330440, D-28334 Bremen, Germany*

SUMMARY

The Gauss–Newton (GN) method is a well-known iterative technique for solving nonlinear least-squares problems subject to dynamical system constraints. Such problems arise commonly in optimal state estimation where the systems may be stochastic. Variational data assimilation techniques for state estimation in weather, ocean and climate systems currently use approximate GN methods. The GN method solves a sequence of linear least-squares problems subject to linearized system constraints. For very large systems, low-resolution linear approximations to the model dynamics are used to improve the efficiency of the algorithm. We propose a new method for deriving low-order system approximations based on model reduction techniques from control theory. We show how this technique can be combined with the GN method to retain the response of the dynamical system more accurately and improve the performance of the approximate GN method. Copyright © 2007 John Wiley & Sons, Ltd.

1. INTRODUCTION

The Gauss–Newton (GN) method is a well-known iterative technique for solving nonlinear least-squares problems subject to dynamical constraints [1]. It may be applied to optimal state estimation

---

*Correspondence to: N. K. Nichols, Department of Mathematics, University of Reading, P.O. Box 220, Reading RG6 6AX, U.K.
†E-mail: n.k.nichols@rdg.ac.uk

problems, where the dynamical system may be stochastic, and is used in variational data assimilation for environmental systems [2–5]. The GN method is essentially an approximation to the Newton method in which only the first-order part of the Hessian is retained. We consider a general nonlinear least-squares problem

$$\min_{\mathbf{x}} \phi(\mathbf{x}) = \mathbf{f}(\mathbf{x})^{\mathrm{T}} \mathbf{f}(\mathbf{x}) \tag{1}$$

where $\mathbf{f}(\mathbf{x})$ is a nonlinear function of $\mathbf{x}$ and we define $\mathbf{J}(\mathbf{x})$ to be the Jacobian of $\mathbf{f}(\mathbf{x})$. Then the GN iteration method for solving this problem consists of the following steps:

$$\text{Solve for } \delta\mathbf{x}^{(k)}\colon \ (\mathbf{J}(\mathbf{x}^{(k)})^{\mathrm{T}}\mathbf{J}(\mathbf{x}^{(k)}))\delta\mathbf{x}^{(k)} = -\mathbf{J}(\mathbf{x}^{(k)})^{\mathrm{T}}\mathbf{f}(\mathbf{x}^{(k)}) \tag{2}$$

$$\text{Update}\colon \ \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \delta\mathbf{x}^{(k)} \tag{3}$$

for $k = 0, 1 \ldots$. For very large systems, such as arise in meteorology and oceanography, it is not generally possible to solve (2) directly. The solution $\delta\mathbf{x}^{(k)}$ is then found by an 'inner' minimization of the linear least-squares function

$$\tilde{\phi}(\delta\mathbf{x}^{(k)}) = \|\mathbf{J}(\mathbf{x}^{(k)})\delta\mathbf{x}^{(k)} + \mathbf{f}(\mathbf{x}^{(k)})\|_2 \tag{4}$$

In the case where dynamical system constraints are imposed on the optimization problem, the objective function $\mathbf{f}(\mathbf{x})$ depends on the dynamical equations. In order to apply the GN method, the Jacobian of the function $\mathbf{f}(\mathbf{x})$ is needed and, hence, the Jacobian of the nonlinear system equations, known as the tangent linear model, is also required. The inner linear minimization problem (4) is then solved subject to the dynamical constraints imposed by the tangent linear model. In order to make the GN method more efficient for large systems, a commonly used approach is to approximate the full tangent linear model by a linearized model at low spatial resolution, while still calculating $\mathbf{f}(\mathbf{x})$ and updating $\mathbf{x}$ at the highest spatial resolution. Thus, in (4) the Jacobian $\mathbf{J}(\mathbf{x}^{(k)})$ is replaced by an approximate low-resolution operator $\tilde{\mathbf{J}}(\mathbf{x}^{(k)})$. Whilst this leads to an algorithm that is practical to compute in real time, the approximate low-resolution model may not correctly reproduce the input–output response of the full system and the accuracy of the state estimate may be compromised.

In this work we propose a new method for approximating the inner step of the GN method, based on the ideas of model reduction. Model reduction has been used in the field of state-space control theory to approximate very large dynamical systems with low-order models [6]. In the context of Kalman filtering, such methods have been applied to optimal state estimation under certain simplifying assumptions [7]. In the context of variational data assimilation low-order models for optimal state estimation have been obtained using proper orthogonal decomposition [8]. Model reduction techniques have not, however, been applied in the context of the GN method for state estimation. In this paper we show how model reduction may be used in the inner step of the GN algorithm to give an approximate iteration procedure that retains the most important properties of the dynamical system response. Employing the example of the balanced-truncation technique, we demonstrate the application of this technique to state estimation and present experiments with a shallow-water model that show the benefits that may be obtained by using a reduced-order model instead of a standard low-resolution model within the GN algorithm. We show that the use of model reduction allows the system to be approximated accurately by models of much smaller dimension than can be used with low-resolution approximations. In the next section, we describe the application of approximate GN methods to the variational data assimilation state estimation

problem and in Section 3 we present the model reduction techniques that we use here. In Section 4 the test model is defined and experimental results are shown. Conclusions are given in Section 5.

2. FOUR-DIMENSIONAL VARIATIONAL DATA ASSIMILATION

The aim of state estimation is to match the output response of a dynamical system model to observed measurements of the outputs over a specified time window. For a discrete dynamical system, we let $\mathbf{x}_j \in \mathbb{R}^n$ be the model state vector, $\mathbf{y}_j \in \mathbb{R}^{p_j}$ be a vector of p_j observations and $\mathbf{h}_j : \mathbb{R}^n \to \mathbb{R}^{p_j}$ be a nonlinear function that relates the system states to the observations at time t_j. The state estimation data assimilation problem is defined as follows.

Problem 1
Minimize, with respect to $\mathbf{x}_0$, the objective function

$$\mathcal{J}[\mathbf{x}_0] = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}})^{\mathrm{T}} \mathbf{B}_0^{-1} (\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}}) + \frac{1}{2} \sum_{j=0}^{N} (\mathbf{h}_j(\mathbf{x}_j) - \mathbf{y}_j)^{\mathrm{T}} \mathbf{R}_j^{-1} (\mathbf{h}_j(\mathbf{x}_j) - \mathbf{y}_j) \tag{5}$$

subject to $\mathbf{x}_j$, $j = 0, \ldots, N$, satisfying the discrete dynamical system equations.

The background or *prior* estimate, $\mathbf{x}_0^{\mathrm{b}}$, of the initial state, $\mathbf{x}_0$, is known. If the initial errors $(\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}})$ and the observational errors $\mathbf{d}_j = \mathbf{h}_j(\mathbf{x}_j) - \mathbf{y}_j$ are assumed to be unbiased, Gaussian random vectors with covariance matrices $\mathbf{B}_0$ and $\mathbf{R}_j$, respectively, then the optimization delivers the maximum posterior Bayesian estimate of the initial state [9].

In practice, the constrained minimization problem is solved iteratively by the GN method. At each step of this method the inner linear least-squares problem (4) must be solved for the increment $\delta\mathbf{x}^{(k)}$. We write the linearized discrete system equations as

$$\delta\mathbf{x}_{j+1} = \mathbf{M}_j \delta\mathbf{x}_j, \quad \mathbf{d}_j = \mathbf{H}_j \delta\mathbf{x}_j \tag{6}$$

where $\delta\mathbf{x}_0 = \mathbf{B}_0^{1/2}\boldsymbol{\omega}$ with random Gaussian noise input $\boldsymbol{\omega} \sim \mathcal{N}(0, \mathbf{I})$, and $\mathbf{M}_j$ and $\mathbf{H}_j$ are, respectively, the tangent linear system model and the Jacobian of the observation operator obtained by linearizing around the state $\mathbf{x}_j$. The inner minimization problem is then given by the following problem.

Problem 2
Minimize, with respect to $\delta\mathbf{x}_0^{(k)}$, the objective function

$$\tilde{\mathcal{J}}^{(k)}[\delta\mathbf{x}_0^{(k)}] = \frac{1}{2}(\delta\mathbf{x}_0^{(k)} - [\mathbf{x}^{\mathrm{b}} - \mathbf{x}_0^{(k)}])^{\mathrm{T}} \mathbf{B}_0^{-1} (\delta\mathbf{x}_0^{(k)} - [\mathbf{x}^{\mathrm{b}} - \mathbf{x}_0^{(k)}])$$

$$+ \frac{1}{2} \sum_{j=0}^{N} (\mathbf{H}_j \delta\mathbf{x}_j^{(k)} - \mathbf{d}_j^{(k)})^{\mathrm{T}} \mathbf{R}_j^{-1} (\mathbf{H}_j \delta\mathbf{x}_j^{(k)} - \mathbf{d}_j^{(k)}) \tag{7}$$

subject to the states $\delta\mathbf{x}_j^{(k)}$ satisfying the discrete linear system equations (6).

Problem 2 is solved by an inner iteration process, such as the conjugate gradient method. Each inner iteration requires one forward solution of the tangent linear model equations (6), starting

from the current best estimate of the initial states, and one backward solution of the corresponding linear adjoint equations to determine the gradient of the objective function. In order for this process to be operationally feasible for very large systems such as weather and ocean systems, which involve as many as 10^8 state variables, the computational cost is reduced by using low-dimensional approximations to the linear models.

Low-dimensional system models can be obtained by using low-resolution approximations to the full dynamical system. Significant features of the system behaviour are often lost, however, in such approximations. In particular optimal error growth modes may not be captured by these models. In the next section we propose an alternative method for generating low-order system approximations using techniques of model reduction.

3. MODEL REDUCTION BY BALANCED TRUNCATION

To find low-order approximations to the linearized system model (6), we project the system into a low-dimensional subspace. We introduce linear restriction operators $\mathbf{U}_j^{\mathrm{T}} \in \mathbb{R}^{r \times n}$ that project the state variables into the subspace $\mathbb{R}^r$, where $r \ll n$. We define variables $\delta \hat{\mathbf{x}}_j \in \mathbb{R}^r$, such that $\delta \hat{\mathbf{x}}_j = \mathbf{U}_j^{\mathrm{T}} \delta \mathbf{x}_j$, and define prolongation operators $\mathbf{V}_j \in \mathbb{R}^{n \times r}$ that project the variables back into the original space $\mathbb{R}^n$. The restriction and prolongation operators $\mathbf{U}_j^{\mathrm{T}}$ and $\mathbf{V}_j$ satisfy $\mathbf{U}_j^{\mathrm{T}} \mathbf{V}_j = \mathbf{I}_r$ and $\mathbf{V}_j \mathbf{U}_j^{\mathrm{T}}$ is a projection operator. We write the projected linear system as

$$\delta \hat{\mathbf{x}}_{j+1} = \mathbf{U}_j^{\mathrm{T}} \mathbf{M}_j \mathbf{V}_j \delta \hat{\mathbf{x}}_j, \quad \hat{\mathbf{d}}_j = \mathbf{H}_j \mathbf{V}_j \delta \hat{\mathbf{x}}_j \tag{8}$$

The reduced-dimension inner minimization problem then becomes

Problem 3
Minimize, with respect to $\delta \hat{\mathbf{x}}_0^{(k)}$, the objective function

$$\hat{\mathscr{J}}^{(k)}[\delta \hat{\mathbf{x}}_0^{(k)}] = \frac{1}{2} (\delta \hat{\mathbf{x}}_0^{(k)} - \mathbf{U}_0^{\mathrm{T}} [\mathbf{x}^{\mathrm{b}} - \mathbf{x}_0^{(k)}])^{\mathrm{T}} \hat{\mathbf{B}}_0^{-1} (\delta \hat{\mathbf{x}}_0^{(k)} - \mathbf{U}_0^{\mathrm{T}} [\mathbf{x}^{\mathrm{b}} - \mathbf{x}_0^{(k)}])$$

$$+ \frac{1}{2} \sum_{j=0}^{N} (\mathbf{H}_j \mathbf{V}_j \delta \hat{\mathbf{x}}_j^{(k)} - \mathbf{d}_j^{(k)})^{\mathrm{T}} \mathbf{R}_j^{-1} (\mathbf{H}_j \mathbf{V}_j \delta \hat{\mathbf{x}}_j^{(k)} - \mathbf{d}_j^{(k)}) \tag{9}$$

subject to the states $\delta \hat{\mathbf{x}}_j^{(k)}$ satisfying the discrete linear system equations (8).

The prolongation operator $\mathbf{V}_0$ is then applied to lift the solution $\delta \hat{\mathbf{x}}_0^{(k)}$ back into the space $\mathbb{R}^n$. Here, the matrix $\hat{\mathbf{B}}_0 = \mathbf{U}_0^{\mathrm{T}} \mathbf{B}_0 \mathbf{U}_0$ models the background error statistics in the reduced space.

The projection matrices are chosen to ensure that the low-order linear model (8) accurately approximates the output response of the linear system (6) to the input data over a full frequency range. Specifically, the aim of the model reduction is to design a model of order $r \ll n$ of form (8) such that the expected value of the distance between the system outputs $\mathbf{d}_j$ and the reduced-order model outputs $\hat{\mathbf{d}}_j$, expressed as

$$\lim_{j \to \infty} \mathscr{E}\{\|\hat{\mathbf{d}}_j - \mathbf{d}_j\|_2^2\} = \lim_{j \to \infty} \mathscr{E}\{[\hat{\mathbf{d}}_j - \mathbf{d}_j]^{\mathrm{T}} [\hat{\mathbf{d}}_j - \mathbf{d}_j]\} \tag{10}$$

is minimized over all random Gaussian white-noise inputs $\boldsymbol{\omega}$ of normalized unit length such that $\mathscr{E}\{(1/n)\|\boldsymbol{\omega}\|_2^2\}=1$, where $\mathscr{E}\{\cdot\}$ denotes the expected value.

The response of a discrete linear system is represented by its Hankel matrix and a good approximation to the optimal reduced-order system can be found by matching the Hankel singular values of the reduced-order model to those of the full linear system model. The projection matrices $\mathbf{U}_j$ and $\mathbf{V}_j$ are therefore selected to ensure that the maximum Hankel singular values of the full system are retained by the reduced-order model. A global error bound on the expected error between the frequency responses of the full and reduced systems, based on the neglected Hankel singular values, then exists [6]. A reduced-order system derived by this method is therefore expected to capture well the modes of optimal error growth. Efficient and accurate numerical techniques are readily available for finding the restriction and prolongation operators for systems of moderately large size based on balanced-truncation methods [10, 11], and for very large systems, Krylov subspace methods and rational interpolation methods can be implemented [12, 13]. (See [14] for further details and preliminary results.)

4. NUMERICAL EXPERIMENTS

To illustrate the benefit of using reduced-order models we apply the method of balanced truncation to a discrete model of the nonlinear 1D shallow-water equations with rotation and compare the results to those obtained using a low-resolution model of the system. The continuous system is given by

$$
\begin{aligned}
&\frac{\partial u}{\partial t}+(U_c+u)\frac{\partial u}{\partial x}-f v = -g\frac{\partial(h+\bar{h})}{\partial x} \\
&\frac{\partial v}{\partial t}+(U_c+u)\frac{\partial v}{\partial x}+f u = 0, \qquad \frac{\partial h}{\partial t}+\frac{\partial h(U_c+u)}{\partial x}=0
\end{aligned}
\tag{11}
$$

where u denotes the departure of the velocity in the x-direction from a known constant forcing the mean flow, U_c, $\bar{h}=\bar{h}(x)$ is the height of the orography, f is the Coriolis parameter and g is the gravitational force. The model domain is periodic in the x-direction. The system is discretized as described in [15] using a semi-implicit semi-Lagrangian integration scheme.

The data for the problem are also given in [15]. An initial perturbation $\delta\mathbf{x}_0$ is defined and the outputs $\mathbf{d}_j$ satisfying the discrete linear equations (6) are determined. A time-invariant linear model that approximates the tangent linear model of the system is used in the experiments. The solution in $\mathbb{R}^r$ of the reduced inner minimization problem, Problem 3, is found using (i) a low-resolution model and (ii) a reduced-order model obtained by balanced truncation. In case (i) the corresponding solution in $\mathbb{R}^n$ is found by interpolation, whilst in case (ii) the solution in $\mathbb{R}^n$ is found by applying the prolongation operator.

The results are compared with the exact solution to the full linear least-squares problem, Problem 2. In the example where $n=1500$ and $r=750$, the errors between the exact solution and the solutions obtained using the two different low-dimensional models are shown in Figure 1(a). The least-square norms of the errors in the two cases are, respectively, (i) 0.0396 and (ii) 0.0057. It is clear that for the same size, the optimal reduced-order models are significantly more accurate than the low-resolution models. This benefit can be partially explained by examining the eigenstructure of the reduced-dimensional systems shown in Figure 2. More of the significant eigenvalues of
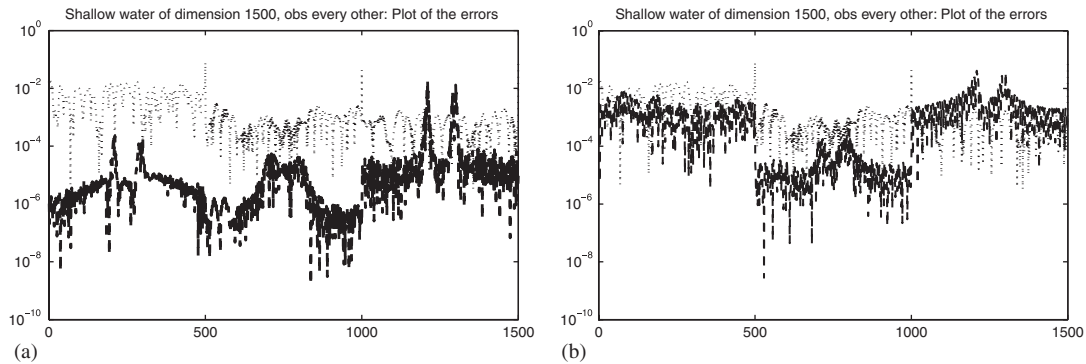
Figure 1. Errors in solutions to reduced linear least-squares problem for low-resolution model (dotted) of order $r = 750$ and optimal reduced-order model (dashed) of order: (a) $r = 750$ and (b) $r = 250$.
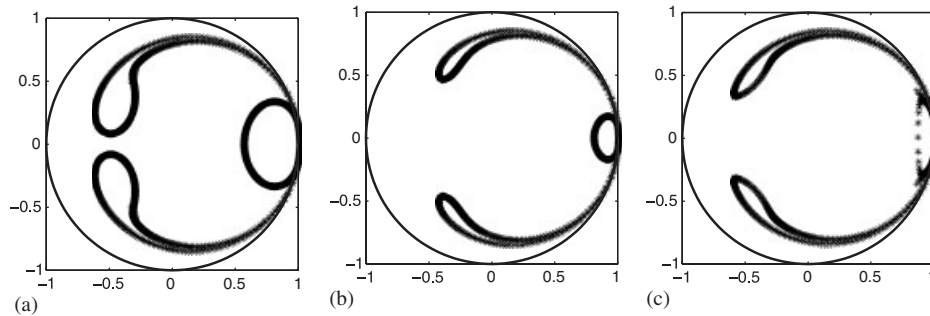

Figure 2. Eigenvalues of: (a) full matrix; (b) low-resolution matrix; and (c) low-order matrix.

the optimal reduced-order model match those of the full system than is the case for the low-resolution model, showing that the modes of the system are more accurately captured by the balanced-truncation method than by using low-resolution models.

Solutions to the reduced least-squares problem obtained for smaller values of r demonstrate that balanced truncation can be applied to find much smaller systems with accuracy equal to that of the low-resolution model. A comparison of the errors obtained with the low-resolution model of dimension $r = 750$ and an optimal reduced-order model of dimension $r = 250$ is shown in Figure 1(b). The least-squares error norm obtained using the optimal reduced-order model is now 0.0288. This demonstrates that an optimal reduced-order model with one-third the dimension can achieve the same accuracy as a low-resolution model. This represents a considerable increase in computational efficiency in practice.

5. CONCLUSIONS

We have described a new approach to finding low-dimensional linear models that can be used to improve the efficiency of approximate GN methods for solving nonlinear least-squares problem.

The new approach applies techniques of optimal reduced-order modelling from control theory and is shown to give better accuracy with significant improvement in performance in solving the inner linear least-squares problem. The challenge is now to develop these techniques to apply to the very large problems arising in environmental systems.

REFERENCES

1. Gill PE, Murray W, Wright MR. *Practical Optimization*. Academic Press: New York, 1986.
2. Lawless AS, Gratton S, Nichols NK. Approximate Gauss–Newton methods for nonlinear least squares problems. *SIAM Journal on Optimization* 2007; **18**:106–132.
3. Lawless AS, Gratton S, Nichols NK. An investigation of incremental 4D-Var using non-tangent linear models. *Quarterly Journal of the Royal Meteorological Society* 2005; **131**:459–476.
4. Lawless AS, Gratton S, Nichols NK. Approximate iterative methods for variational data assimilation. *International Journal for Numerical Methods in Fluids* 2005; **47**:1129–1135.
5. Courtier P, Thépaut 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.
6. Antoulas AC. *Approximation of Large-scale Dynamical Systems*. SIAM: Philadelphia, PA, 2005.
7. Farrell BF, Ioannou PJ. State estimation using a reduced-order Kalman filter. *Journal of Atmospheric Science* 2001; **58**:3666–3680.
8. Cao Y, Zhu J, Navon IM, Luo Z. A reduced order approach to four-dimensional variational data assimilation using proper orthogonal decomposition. *International Journal for Numerical Methods in Fluids* 2007; **53**:1571–1583.
9. Lewis JM, Lakshmivarahan S, Dhall S. *Dynamic Data Assimilation*: *A Least Squares Approach*. Cambridge University Press: Cambridge, 2006.
10. Laub AJ, Heath MT, Paige CC, Ward RC. Computation of system balancing transformations and other applications of simultaneous diagnolization algorithms. *IEEE Transactions on Automatic Control* 1987; **AC-32**:115–122.
11. Chahlaoui Y, Van Dooren P. Model reduction of time-varying systems. In *Dimension Reduction of Large-scale Systems*, Benner P, Mehrmann V, Sorensen D (eds). Springer: Berlin, 2005.
12. Freund RW. Model reduction methods based on Krylov subspaces. *Acta Numerica* 2003; **12**:267–319.
13. Gugercin S, Sorensen DC, Antoulas AC. A modified low-rank Smith method for large-scale Lyapunov equations. *Numerical Algorithms* 2003; **32**:27–55.
14. Lawless AS, Nichols NK, Boess C, Bunse-Gerstner A. Using model reduction methods within incremental 4D-Var. *Monthly Weather Review*, in press.
15. Katz D, Lawless AS, Nichols NK, Cullen MJP, Bannister RN. A comparison of potential vorticity-based and vorticity-based control variables. *Numerical Analysis Report*, *8/2006*, Department of Mathematics, The University of Reading, 2006.