Application of LSQR to Calibration of a MODFLOW Model: A Synthetic Study

Chris Muffels[1,2], Matthew Tonkin[2,3], Haijiang Zhang[1], Mary Anderson[1], Tom Clemo[4]
[1]*University of Wisconsin-Madison, muffels@geology.wisc.edu, hjzhang@geology.wisc.edu, andy@geology.wisc.edu, Madison, WI, USA*
[2]*S.S. Papadopulos and Associates, Inc., matt@sspa.com, Bethesda, MD, USA*
[3]*The University of Queensland, Australia*
[4]*Boise State University, tomc@cgiss.boisestate.edu, Boise, ID, USA*

ABSTRACT

The inverse problem in groundwater modeling is often made numerically tractable and computationally practical by estimating only a small fraction of the many unknown system parameters. However, this parsimonious approach restricts the solution of the inverse problem to a pre-determined subspace of the true parameter space. To reflect detailed local variations in hydraulic conductivity or recharge, it may be desirable to estimate a very large number of parameters during calibration, which requires an inversion technique that can accommodate highly parameterized models. The least-squares QR (LSQR) decomposition is an iterative solution method that can solve for many hundreds or thousands of parameters. LSQR has been used successfully in seismic tomography inversion problems. As an iterative method, LSQR can solve sparse and dense inverse problems of the form **Ax=b** using significantly less computer storage than direct solution methods. We test the applicability of the LSQR method for solving the inverse problem for groundwater flow using a synthetic model and compare results with those obtained using the more commonly employed method, the singular value decomposition (SVD). Parameter sensitivities are calculated using forward differences and the adjoint-state method.

INTRODUCTION

The inverse problem in groundwater modeling is often posed as a parsimonious one by estimating a small fraction of the many unknown system parameters. Difficulties and limitations of this approach are well known (e.g. Moore and Doherty 2005, deMarsily et al. 2005). The development of computationally efficient methods for calculating parameter sensitivities (Townley and Wilson 1985) combined with the application of regularization methodologies (Engl et al. 1996; Tonkin and Doherty 2005) suggests that the rapid and stable solution of large groundwater inverse problems is within reach and parsimony may not be the 'norm' in model parameterization for much longer. However, the solution of highly-parameterized inverse problems in the least-squares sense using the approach common to most programs - direct matrix inversion – suffers enormously in the presence of round-off error and other sources of noise. The focus of this paper is the application of LSQR to the calibration of a MODFLOW2000 (Harbaugh et al. 2000) model and a comparison with SVD. It is important to note that large inverse problems (100,000 parameters) are routinely solved in many applications including seismic tomography. This paper introduces readers to the theory behind the SVD and LSQR methods and then illustrates their application using a simple synthetic groundwater flow model. Discussions focus on the contrasting computational burden, convergence behavior, and parameter estimates of the two methods for the simple case. The analysis described herein is one step in ongoing research examining the applicability and versatility of the LSQR decomposition technique.

METHODS EMPLOYED

The linear least squares inverse problem can be illustrated by:

$$\text{minimize } |\mathbf{Ax} - \mathbf{b}| \tag{1}$$

where **A** is a matrix of *m* rows and *n* columns, **x** is an *n*-row vector of parameters, and **b** is an *m*-row vector of observations. In non-linear problems the solution is obtained through the iterative formation and solution of a (local) linear approximation to the non-linear problem. In non-linear models, the parameter

upgrade vector, $\Delta\mathbf{x}$, is often calculated through the iterative minimization of a residual sum-of-squares objective function (phi) using the direct solution of the normal equations:

$$\Delta\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{r} \quad (2)$$

where **r** lists the residuals for the current parameter set, T indicates the matrix transpose, and (–1) indicates the matrix inverse. Each linearization requires the construction of **A**, which typically forms the most computationally intensive aspect of the process (Carrera et al. 1990). While matrix inversion is effective for well-conditioned, typically over-determined systems comprising a small number of parameters, if $\mathbf{A}^T\mathbf{A}$ is near-singular it cannot be inverted and the likelihood of a near-singular $\mathbf{A}^T\mathbf{A}$ increases as *n* increases. Alternative methods for solving (1) are the SVD and the LSQR methods.

SVD

Singular value decomposition is a direct solution method that decomposes an arbitrary matrix **B** into:

$$\mathbf{B} = \mathbf{USV}^T \quad (3)$$

where **U** and **V** contain the left and right eigenvectors of **B**, respectively (Lawson and Hanson 1995; Anderson et al. 1999). The singular value decomposition is general in the sense that it can be applied to any $m \times n$ matrix. The first phase of the SVD is to compute **U** and **V** such that **USV** is bi-diagonal. In the special case of the square symmetric positive-definite matrix, $\mathbf{B}=\mathbf{A}^T\mathbf{A}$, **S** is diagonal and lists the *n* singular values or eigenvalues of $\mathbf{A}^T\mathbf{A}$. In addition the *n* column vectors of matrix **V** are the eigenvectors of $\mathbf{A}^T\mathbf{A}$. Because $\mathbf{A}^T\mathbf{A}$ is symmetric positive semi-definite, its eigenvectors are orthogonal. Truncated singular value decomposition (TSVD) is a mechanism for determining $\Delta\mathbf{x}$ from the *k* most dominant eigenvectors, where $k < n$ is a subset of the full complement of eigenvectors. The TSVD is then not an exact decomposition of **B** but represents the closest approximation that can be achieved by a matrix of rank *k*. The eigenvectors represent unique weighted combinations of parameters that influence the calculated observation values. The eigenvalues indicate the magnitude of the influence. Therefore, the eigenvectors associated with small eigenvalues have little influence on the calculated observations and can be safely neglected. As a corollary, because they have a small influence on the observation values, the eigenvectors are very susceptible to noise in the data. Errors in the estimates of these vectors can significantly degrade the parameter distribution. Given matrix **A** of size *m*-by-*n* the SVD requires $O(m \times n)^2$ or $O(n)^3$ operations depending on the initial factorization approach taken (e.g., the Givens or Householder approaches). SVD also requires full storage of the matrices **U**, **V** and the vector **S**.

LSQR

LSQR (Paige and Saunders 1982a,b) is an iterative solution method related to conjugate-gradient (CG) methods. LSQR can be shown to be algebraically equivalent to applying the symmetric CG method to the normal equations, but possesses superior numerical properties when **A** is ill-conditioned that can lead to more rapid and accurate convergence. The LSQR was developed for solving large sparse problems, but has been applied to the solution of dense problems such as described in this paper. The LSQR is applicable to the solution of the non-symmetric problem **Ax=b** or for linear least squares of the form (1). When used to solve (1) the sensitivity matrix **A** is only used to compute matrix-vector products. The LSQR is based on the Lanczos process and the bi-diagonalization procedure of Golub and Kahan (1965). Variants of the Golub-Kahan-Lanczos bi-diagonalization procedure are employed in some implementations of the SVD, illustrating the relation between these approaches. However, the LSQR uses an iterative method to approximate the solution, with the number of iterations required to reach convergence depending on the condition number of the problem and the desired accuracy. For well conditioned **A** ($m \times n$), the LSQR can theoretically solve the system in *n* operations. However, where **A** is ill-conditioned preconditioning of $O(n)$ is required and the number of required operations increases linearly with *n*. A satisfactory solution may be obtained sooner, since the LSQR captures spectral components of the solution in order of increasing frequency. Therefore, initial iterates are 'smoother' than later iterates and can be described as regularized in a manner comparable with TSVD. Since the least-squares solution is obtained through matrix-vector multiplications using matrix **A** and matrix $\mathbf{A}^T\mathbf{A}$ is never

formed and(or) decomposed, the condition number of the inverse problem is lower than when operating on the normal equations as is the case with the implementation of the SVD used here (Subroutine `DGESVD`, Anderson et al. 1999). CG-like methods are appealing due to their low storage requirements and the LSQR requires storage only of **A** and of two vectors of size *n* and *m* respectively, unless the Lanczos vectors are stored for later use. Under certain circumstances the ortho-normal vectors formed in the LSQR solution process may approximate the eigenvectors calculated using the SVD (Yao et al. 1999).

2D CHECKERBOARD TEST CASE

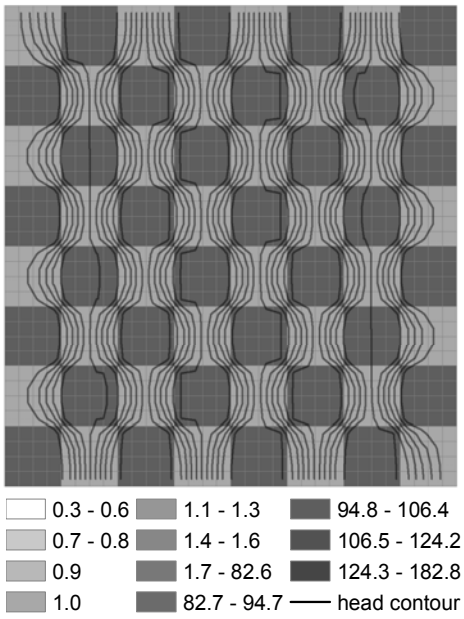

Figure 1. Actual hydraulic conductivity distribution

A simple two-dimensional model was constructed with 32 rows and 32 columns. Constant-head boundaries were used on the left and right sides of the model with no-flow along the top and bottom. Two zones of hydraulic conductivity were evenly distributed throughout the model in a checkerboard pattern (Figure 1). The lighter zones were assigned a hydraulic conductivity of 1 ft/day and the darker zones 100 ft/day. Each zone is 4 rows by 4 columns. The resulting head distribution is illustrated in Figure 1. To test LSQR, the simple model was calibrated to heads using hydraulic conductivity as the only variable parameter. Each model cell was used as a parameter and each node-center as a head-observation, resulting in 1024 parameters and 1024 observations. The starting value for each parameter was calculated as a random fraction (up to ±50 %) of the true value (Figure 2). The sensitivity matrix was calculated using forward-differences with an increment of 1% of the parameter value. Parameters were restricted to values between 0.0001 and 200.0 and could not change by more than a factor of 10.0 within any optimization iteration. Formal regularization was not used and LSQR worked with the sensitivity matrix **A** directly, not the normal matrix ($\mathbf{A}^T\mathbf{A}$).

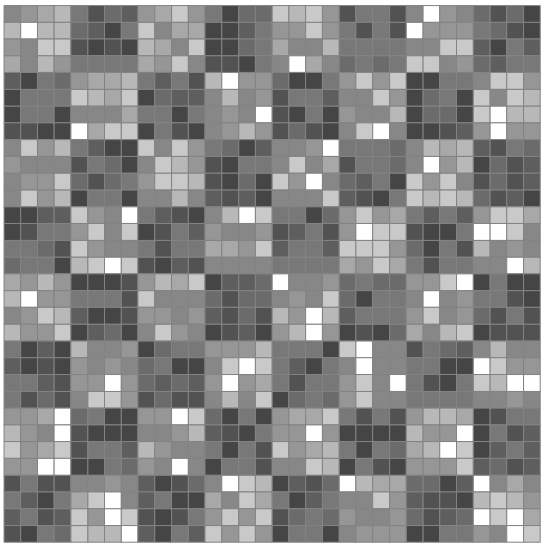

Figure 2. Starting hydraulic conductivity distribution

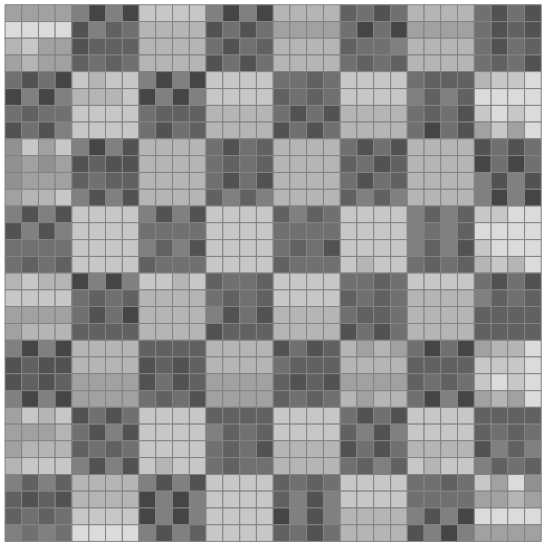

Figure 3. LSQR calibrated hydraulic conductivity distribution

LSQR required 63 optimization iterations - a total of 64,575 forward model executions - to reduce phi from 7.6266 to 0.0002. The resulting hydraulic conductivity distribution is shown in Figure 3. This checkerboard model was calibrated with the same parameters, observations, and parameter change limitations, using SVD as employed in PEST (Doherty 2005) with one exception: SVD operates on the normal matrix $\mathbf{A}^T\mathbf{A}$ and it is therefore expected that the SVD encountered greater noise in obtaining a solution. The SVD was implemented both with and without truncation. The SVD truncation level, k, was set by specifying the ratio of highest-to-lowest acceptable singular value (variable EIGTHRESH). Initially EIGTHRESH was set to 1.0E-4, approximately 200 singular values (SVD 200) were used in the solution, and phi was reduced to 0.0015 in 3 optimization iterations (3076 model runs) (Figure 5). By reducing EIGTHRESH to 1.0E-10, approximately 600 singular values (SVD 600) were used in the solution and in 4 optimization iterations (4101 model runs) phi was reduced to 0.0004; the resulting parameter distribution is shown in Figure 6. The limited improvement demonstrates the minimal influence of eigenvectors associated with small eigenvalues. It is worth noting that the ratio of the highest to lowest singular values in the full problem was 1.0E-25 and that when all 1024 singular values were used the problem was too ill-conditioned to obtain a solution. The starting parameter vector had a variance of 0.09 while LSQR reduced this to 0.02 and SVD 200 to 0.06. The variance resulting from SVD 600 actually increased to 0.13 owing to two parameters with a relative residual value of less than -6.

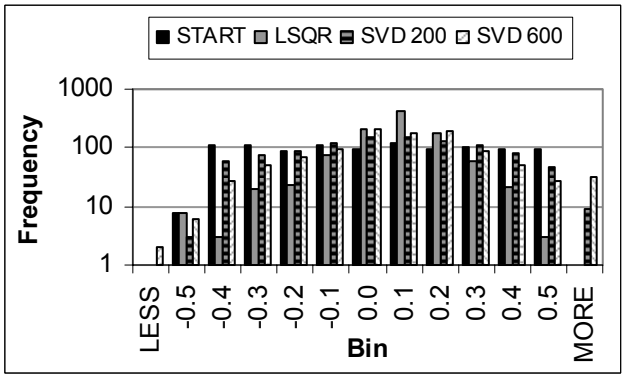

Figure 4. Histogram of relative parameter residuals

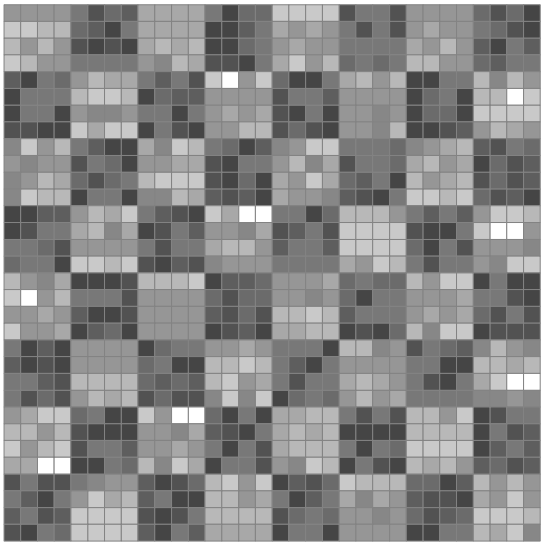

Figure 5. SVD 200 calibrated conductivity distribution

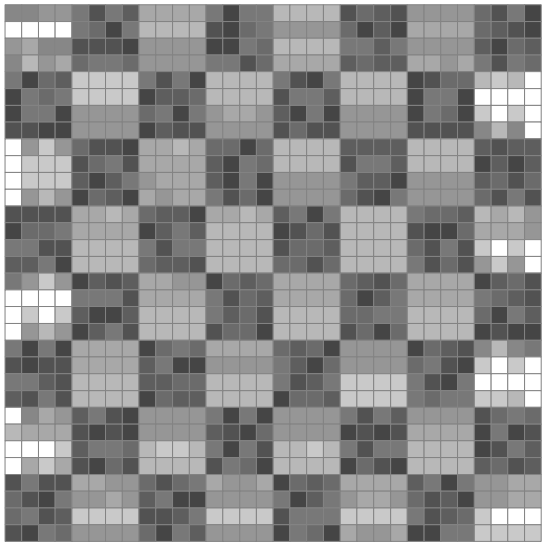

Figure 6. SVD 600 calibrated conductivity distribution

After the sensitivity matrix was calculated, it took LSQR less than 1.5 minutes to solve the linear system while SVD 200 took about 2.5 minutes and SVD 600 about 4.5 minutes. While this result is expected, it can be misleading given the number of optimization iterations required to reduce phi for LSQR compared with SVD, as discussed later. All simulations were completed on a laptop computer with a P4 3.0 GHz CPU and 1.5 GB of RAM.

DISCUSSION AND CONCLUSIONS

The LSQR method shows promise in solving large groundwater inverse problems as demonstrated by the reduction in the residual sum of squares and variance in the parameter residuals compared with SVD 200 and SVD 600. At the time this paper was written an attempt to include all 1024 singular values in the SVD solution was not successful but is the subject of ongoing research. In addition, attempts will be made to use LSQR in the solution of the normal equations $\mathbf{A}^T\mathbf{A}$ to determine how it behaves with the additional noise. It is believed that the ill-conditioning of the system explains why LSQR required more optimization iterations in this test. There is some inherent damping in LSQR due to its internal stopping criteria and investigations of this are part of ongoing research. Also, even though the maximum a parameter could change was 10, neither of the SVD runs had a parameter change by more than 2.5 where as LSQR changed a parameter by 10 each optimization iteration. Given the similarities in the core processes underlying both SVD and LSQR and the success of the regularized inversion approach embodied in TSVD, it is believed that LSQR will demonstrate similar success with the additional advantage of reduced noise and reduced computational burden. This in turn may support the estimation of many more parameters, allowing more complex and larger models to be parameterized with greater confidence, and allowing the inverse process itself to identify where heterogeneity must exist, further reducing the necessity of relying on *a-priori* parsimony.

REFERENCES

Anderson, E., Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammerling, A. McKenney, and D. Sorenson, 1999. *LAPACK Users' Guide Third Edition*, 407 pp, SIAM Publications, Philadelphia, PA, (http://www.netlib.org/lapack/index.html)

Carrera, J., F. Navarrina, L. Vives, J. Heredia, and A. Medina, 1990. Computational aspects of the inverse problem, *Computational methods in subsurface hydrology*, *Proceedings of the Eighth International Conference on Computational Methods in Water Resources*, *Venice, Italy,* pp. 513-522, Springer-Verlag, Boston, Massachusetts.

Doherty, J., 2005. PEST, Model Idenpendent Parameter Estimation. User Manual: 5th Edition. Watermark Numerical Computing, Brisbane, Queensl., Australia.

Engl, H. W., M. Hanke, and A. Neubauer, 1996. Regularization of Inverse Problems, 321 pp., Springer, New York.

de Marsily, G., F. Delay, J. Goncalves, Ph. Renard, V. Teles, S. Violette, 2005. Dealing with spatial heterogeneity, Hydrogeology Journal 13:161–183, DOI 10.1007/s10040-004-0432-3

Golub, G. H., and W. Kahan, 1965. *Calculating the Singular Values and Pseudoinverse of a Matrix*. Journal of the Society for Industrial and Applied Mathematics: Series B, Numerical Analysis, 2(2):205-224.

Harbaugh, A.W., E.R. Banta, M.C. Hill, and M.G. McDonald, 2000. MODFLOW-2000, the U.S. Geological Survey modular ground-water model -- User guide to modularization concepts and the Ground-Water Flow Process: U.S. Geological Survey Open-File Report 00-92, 121 p.

Lawson, C. L. and J, R. Hanson, 1995. *Solving Least Squares Problems, SIAM Classics in Mathematics*, pp 337, Philadelphia, Pennsylvania.

Moore, C., and J. Doherty, 2005. Role of the calibration process in reducing model predictive error, Water Resour. Res., 41, W05020, doi:10.1029/2004WR003501.

Paige, C.C. and M.A. Saunders, 1982a. LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares. ACM Trans. Math. Softw. vol. 8: 43-71.

Paige, C.C. and M.A. Saunders, 1982b. LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares. ACM Trans. Math. Softw. vol. 8: 192-209.

Tonkin, M.J., and J. Doherty. 2005. A Hybrid Regularized Inversion Methodology for Highly Parameterized Environmental Models: Water Resources Research. 41: W10412, doi:10.1029/2005WR003995.

Townley, L. R., and J. L. Wilson, 1985. Computationally efficient algorithms for parameter estimation and uncertainty propagation in numerical models of groundwater flow, Water Resources Research., 21(12), 1851– 1860.

Yao, Z.S., R.G. Roberts and A. Tryggvason, 1999. Calculating resolution and covariance matrices for seismic tomography with the LSQR method, Geophys. J. Int. 138, 886-894.