

A multigrid method for large scale inverse problems

Eldad Haber

Dept. of Computer Science,
Dept. of Earth and Ocean Science
University of British Columbia

haber@cs.ubc.ca

July 4, 2003

Collaborator

Uri Ascher



Main Outline

- Preliminaries - all-at-once methods.
- Multigrid for the linearized system
- Discretization
- Smoothers
- The multigrid iteration
- Solving the nonlinear system using an inexact Newton SQP method
- Example
- Conclusions

Preliminaries

- In many applications we need to evaluate coefficients from noisy measured data
- Examples:
 - Resistivity/hydrology measurements:

$$\nabla \cdot \sigma \nabla \phi_k = q_k \quad k = 1 \dots N$$

- Electromagnetics:

$$\nabla \times \mu^{-1} \nabla \times \mathbf{E}_{jk} - i\omega_j \sigma \mathbf{E}_{jk} = i\omega_j q_k$$

- Discretize **PDE** plus **BC** and write *forward problem* as

$$A(m)u = q$$

- In these problems we need to infer the *model* m given measurements of the *field* u

Formulation of the inverse problem

We have:

- The PDE+BC $A(m)u = q$ (physical process)
- Measured data $b^{obs} = Qu$
- A priori information

What are we after?

- Solve the differential problem $A(m)u = q$
- Fit the data, $\|Qu - b^{obs}\| \leq Tol$
- Get a reasonable model: $\|W(m - m_{ref})\|$ not too large

Formulation cont.

There are two (major) ways to approach the problem

Unconstrained:

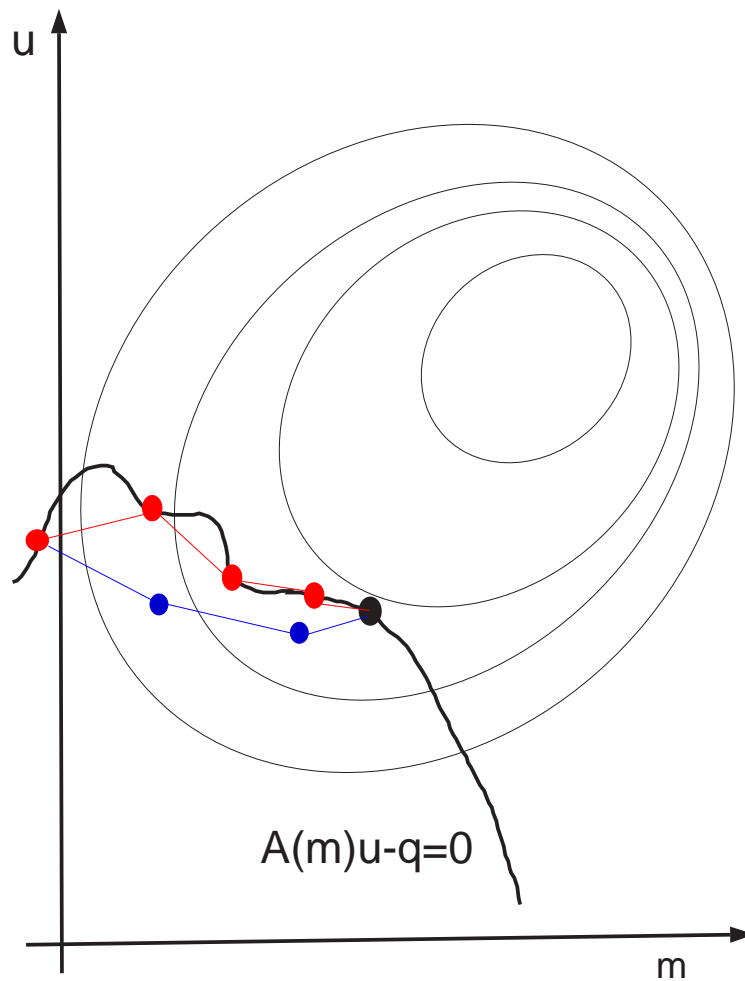
$$\begin{aligned}\phi(m) &= \|Qu - b\|^2 + \beta \|Wm\|^2 \\ &= \|QA(m)^{-1}q - b\|^2 + \beta \|Wm\|^2\end{aligned}$$

Constrained or - simultaneous; all-at-once:

$$\mathcal{L}_{m,u,\lambda} = \|Qu - b\|^2 + \beta \|Wm\|^2 + \lambda^T (A(m)u - q)$$

λ - Vector of Lagrange multipliers.

All-at-once - Unconstrained: The general idea



Constrained - Unconstrained

Linear systems which evolve from all-at-once methods

The Gradients:

$$\mathcal{L}_u = Q^T(Qu - b) + A(m)^T \lambda = 0$$

$$\mathcal{L}_m = \beta W^T W(m - m_0) + G(m, u)^T \lambda = 0$$

$$\mathcal{L}_\lambda = A(m)u - q = 0$$

where

$$G(m, u) = \frac{\partial(A(m)u)}{\partial m}.$$

Gauss-Newton iteration¹:

$$\begin{pmatrix} A & 0 & G \\ Q^T Q & A^T & 0 \\ 0 & G^T & \beta W^T W \end{pmatrix} \begin{pmatrix} \delta u \\ \delta \lambda \\ \delta m \end{pmatrix} = \begin{pmatrix} \mathcal{L}_u \\ \mathcal{L}_\lambda \\ \mathcal{L}_m \end{pmatrix}$$

¹In the Gauss-Newton case we do not use second order information

The linearized system

It is easy to see that the permuted KKT matrix

$$\begin{pmatrix} A & 0 & G \\ Q^T Q & A^T & 0 \\ 0 & G^T & \beta W^T W \end{pmatrix}$$

corresponds to discretizing the differential operator

$$\begin{pmatrix} \nabla \cdot (e^m \nabla(\cdot)) & 0 & \nabla \cdot ((e^m \nabla u)(\cdot)) \\ \sum \delta(x - x_i) & \nabla \cdot (e^m \nabla(\cdot)) & 0 \\ 0 & -(\nabla u)^T (e^m)(\nabla(\cdot)) & -\beta \nabla^2 \end{pmatrix}$$

plus **BC**

This PDE system can be efficiently solved using multigrid methods

Ellipticity of the PDE system

Assuming we have data everywhere, consider one mode

$$\begin{pmatrix} \delta u \\ \delta \lambda \\ \delta m \end{pmatrix} = \begin{pmatrix} \delta \hat{u} \\ \delta \hat{\lambda} \\ \delta \hat{m} \end{pmatrix} e^{i\theta \cdot x}$$

Denote $\sigma = e^m$ and obtain the symbol

$$C = \begin{pmatrix} -\sigma|\theta|^2 & 0 & i\sigma(\nabla u)^T \theta \\ 1 & -\sigma|\theta|^2 & 0 \\ 0 & -i\sigma(\nabla u)^T \theta & \beta|\theta|^2 \end{pmatrix}$$

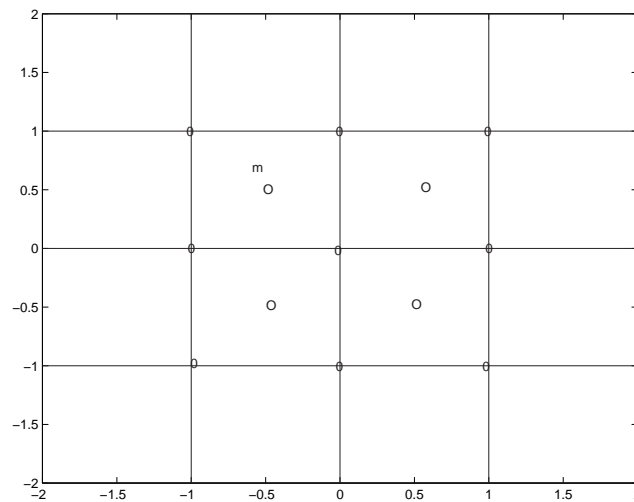
Thus

$$\det(C) = \sigma^2(\beta|\theta|^6 + [(\nabla u)^T \theta]^2)$$

The system is therefore elliptic for any $\beta > 0$

Discretization of the system - I

- Care must be taken to ensure that the discrete version of the Euler-Lagrange equations is h -elliptic.
- Given m at cell centers, there are two possible approaches to the discretization of the forward problem
 - Cell-centered
 - Nodal



Discretization of the system I - cont.

The continuous linearized Euler-Lagrange equations are

$$\nabla \cdot (e^m \nabla(\delta u)) - \nabla \cdot ((e^m \nabla u) \delta m) = g_u$$

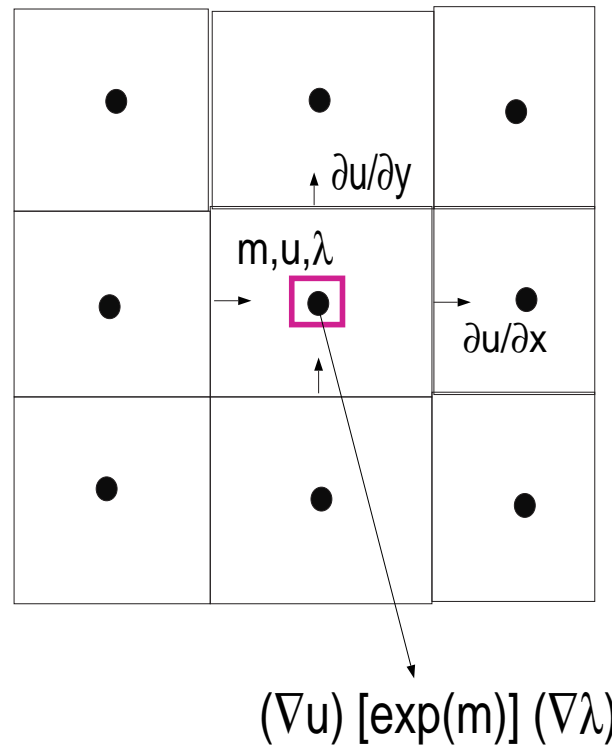
$$\sum \delta(x - x_i) \delta u + \nabla \cdot (e^m \nabla(\delta \lambda)) = g_\lambda$$

$$(\nabla u)^T (e^m) (\nabla(\delta \lambda)) - \beta \nabla^2 \delta m = g_m$$

If we use a cell-centered discretization for the forward problem, then the term $(\nabla u)^T e^m (\nabla(\delta \lambda))$ must be calculated using long differences in both u and λ

In this case, for very small regularization parameter β , the h -ellipticity of the system can be lost on coarser grids.

Discretization of the system I - cont.



Note - Compact discretization for the forward problem may still lead to non-compact discretization of the inverse problem!

Discretization of the system I - cont.

To avoid the difficulty associated with long differencing when the regularization parameter is small, we may introduce an $\mathcal{O}(h^2)$ “artificial regularization” (similar to Levenberg-Marquardt for the reduced Hessian and to pressure correction techniques in CFD)

$$\nabla \cdot (e^m \nabla(\delta u)) - \nabla \cdot ((e^m \nabla u) \delta m) = g_u$$

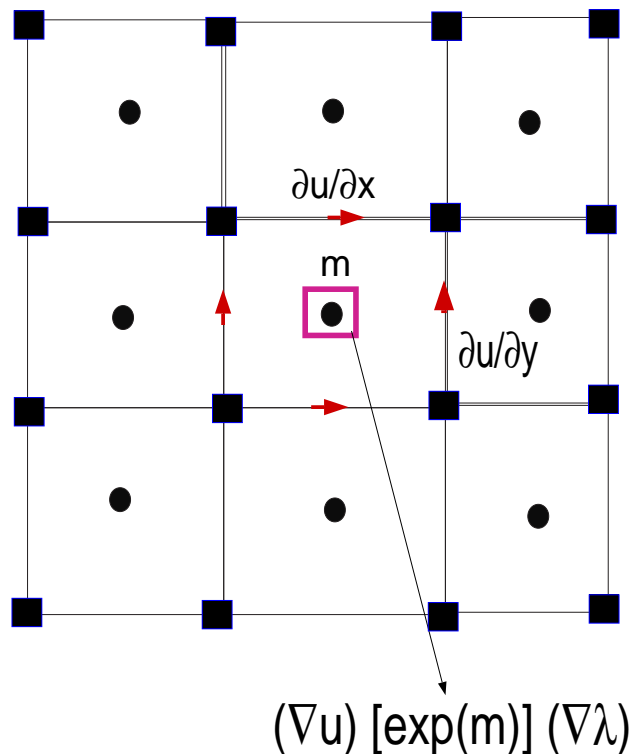
$$\sum \delta(x - x_i) \delta u + \nabla \cdot (e^m \nabla(\delta \lambda)) = g_\lambda$$

$$(\nabla u)^T (e^m) (\nabla(\delta \lambda)) - (\beta + \omega h^2) \nabla^2 \delta m = g_m$$

Discretization of the system II

The problem of long differences can be solved by nodal discretization for u and λ while m remains at cell centers.

In this case we use only short differences and therefore obtain an h -elliptic discretization even for very small regularization parameters.



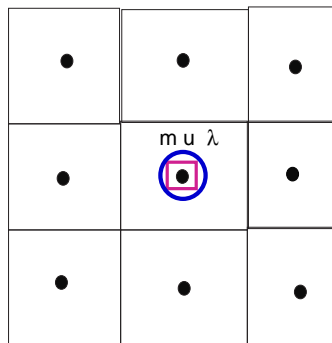
Discretization of the system and optimization

- We discretize the optimization problem rather than the Euler-Lagrange equations directly, thereby retaining KKT symmetry in the discrete necessary conditions.
- We obtain a discrete optimization problem, and use optimization techniques (e.g. variants of SQP [Nocedal & Wright 1999])

Smoothing I

Cell-centered discretization

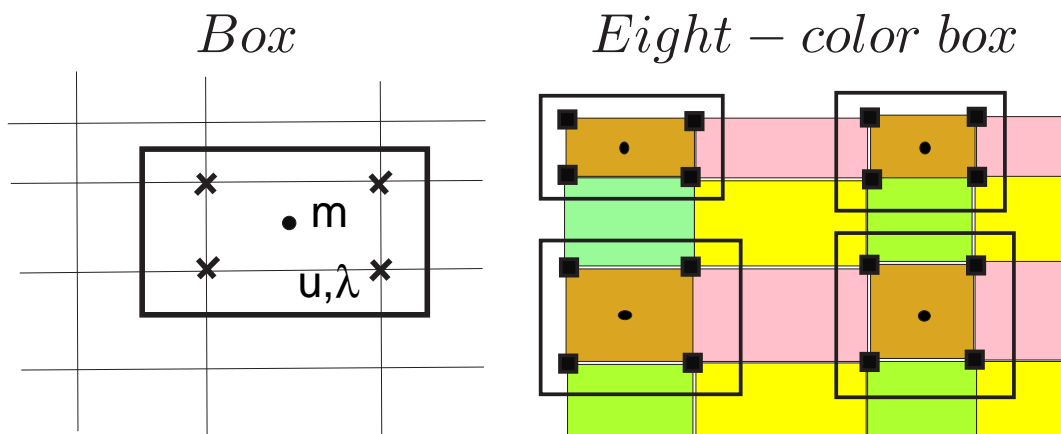
- In this case all unknowns are at the same point.
- For small regularization parameter values, the system is tightly coupled.
- We therefore use Collective Symmetric Gauss-Seidel relaxation.
- During a relaxation sweep invert a 3×3 block at each grid point.



Smoothing II

Cell-node discretization

- The grid is staggered and we therefore use box relaxation.
- Box relaxation I [Vanka 1986]
- Box relaxation II - Eight-color box relaxation (in the spirit of red-black boxes)



The Multigrid iteration

- Prolongation and restriction
 - Nodal unknowns - Trilinear interpolation and full-weight restriction.
 - Cell unknowns - Trilinear interpolation and full-weight restriction.
- Coarse grid operator by Galerkin coarsening.

We perform a classical $W(2,2)$ cycle.

The Multigrid iteration in an Inexact Newton solver

Reminder

We are solving a nonlinear problem where at each iteration the linear system

$$H\delta y = -g$$

must be solved.

In inexact Multigrid-Newton methods, we solve this system to a very rough tolerance (say **0.1** [Kelley 1998]).

This can be accomplished by one W(2,2) cycle per Inexact Newton iteration.

Note

The cost of calculating the Jacobian is negligible compared with one W-cycle.

The Inexact Newton solver, Optimization issues.

We incorporate the method into an inexact Sequential Quadratic Programming (SQP) version [Heinkenshlus 1996] .

Add a globalization strategy (using a line search and the l_1 merit function) [Fletcher 1981] .

Further stabilization is obtained by inexact projection to the constraint [Nocedal & Wright 1999]

Numerical experiments I - The linearized problem

The Setting

- The PDE is

$$\nabla \cdot e^m \nabla u = q$$

in the interval $[-1, 1]^3$. Apply natural BC.

- We use cell-centered and nodal finite volume discretizations for u , where m is at cell centers.
- We measure u on a 6^3 -grid uniformly spaced in the interval $[-0.6, 0.6]$ (contaminated with noise)
- The fine grid has 32^3 cells which leads to roughly 10^5 unknowns

Numerical experiments I - The linearized problem

To test the multigrid scheme, we solve the system which arises in the first iteration to accuracy of 10^{-6}

Rate of convergence for different β 's and different discretizations

β	Levels	CC	SCC	CN
10^{-2}	4	0.11	0.06	0.04
	3	0.05	0.05	0.04
10^{-3}	4	NC	0.09	0.04
	3	0.35	0.09	0.04
10^{-4}	4	NC	0.12	0.04
	3	NC	0.12	0.04

CC - Cell Centered

SCC - Stabilized Cell Centered

CN - Cell-Node

Numerical experiments I - The linearized problem, cont.

- The stabilization of the cell-centered discretization is needed for small values of the regularization parameter
- Cell-Node performs better than Cell Centered (possible due to h -ellipticity effects)
- Using cell-node, the multigrid method is only slightly sensitive to the size of the regularization parameter

Solving the nonlinear problem

In the second experiment we solve the nonlinear problem using inexact Newton.

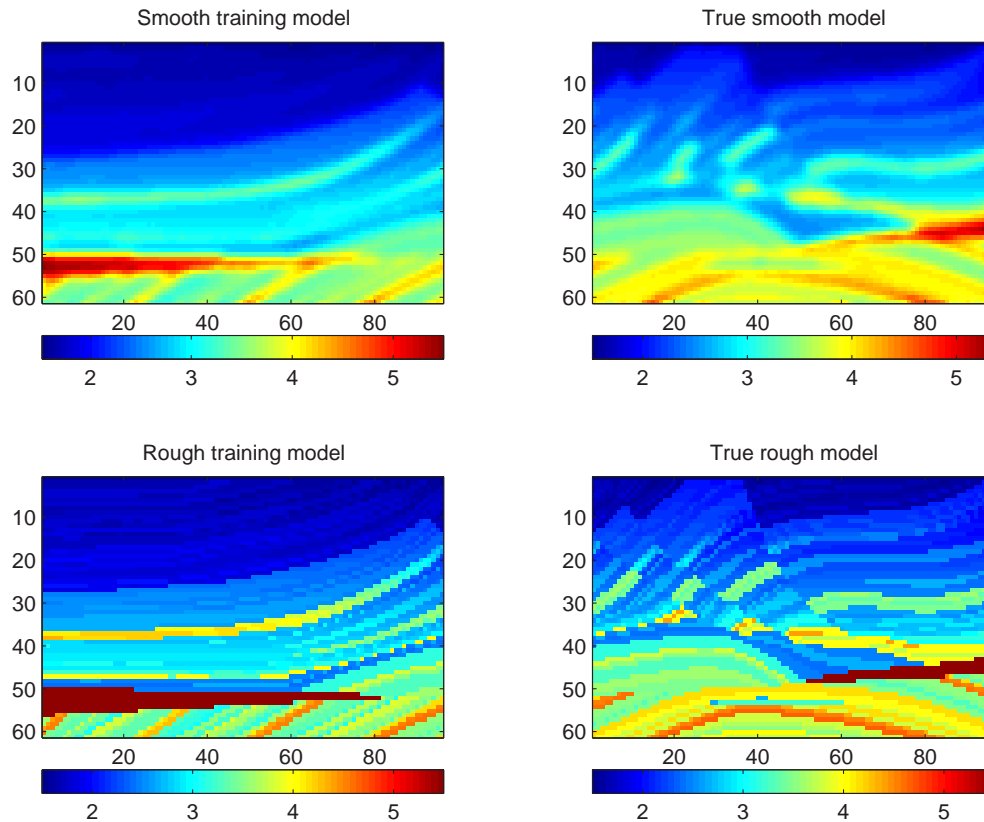
Number of nonlinear iterations and of W-cycles needed to solve the nonlinear problem when using an inexact Newton method.

β	Newton iterations	SCC	CN
10^{-1}	8	8	8
10^{-2}	8	9	8
10^{-3}	14	17	14
10^{-4}	17	21	17

Note that we manage to solve the entire nonlinear inverse problem in a total cost which is not much larger than that of solving a few forward problems.

Solving the nonlinear problem cont.

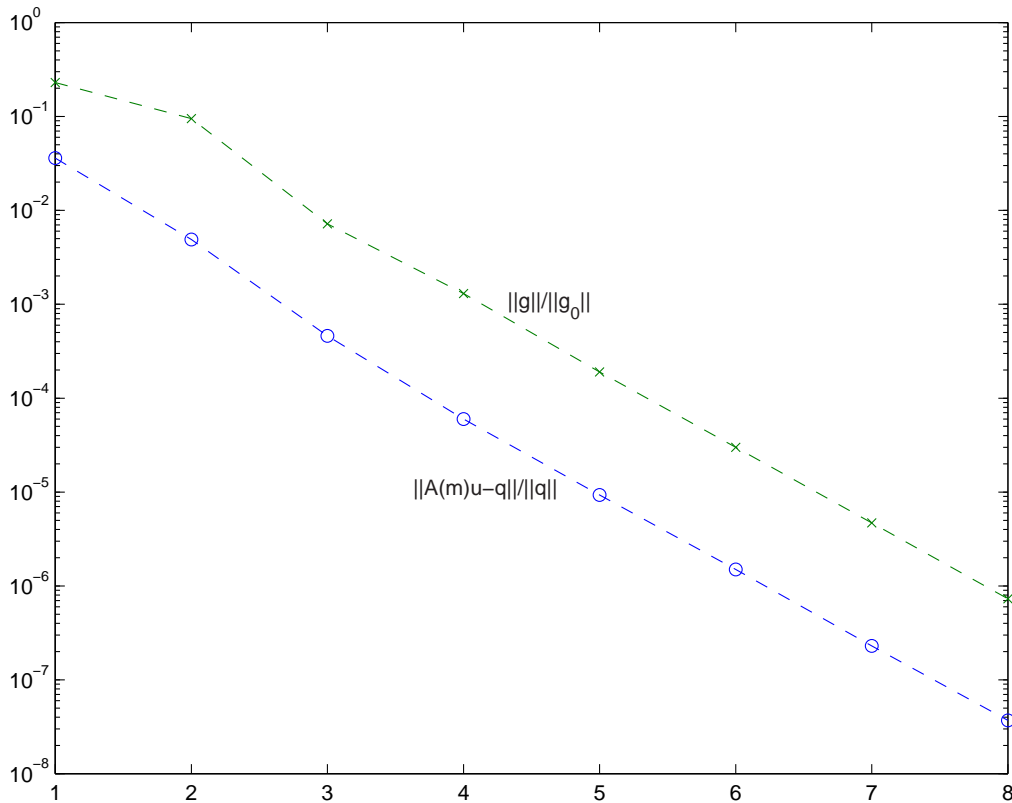
The solutions



Slices in the true model (left) and the reconstructed model for $z = 0.3, 0.55, 0.75$

Solving the nonlinear problem cont.

Convergence of the iteration



Relative gradient and residual of the constraint (PDE) as a function of iteration for $\beta = 10^{-2}$.

Conclusions & some related questions

Using an h -elliptic discretization, it is possible to solve the inverse problem in a cost which is not much larger than of a forward problem

Compact discretizations for the forward problem do not necessarily lead to compact discretizations of the inverse problem.

Related questions

- Incorporating a-priori information
- Robust statistics - choosing β using GCV, L-curve, etc.
- Incorporation of better SQP-type schemes
- Parallelism in the case of multi-experiments