Divergence of finite element formulations for inverse problems treated as optimization problems

Carlos Rivas¹, Paul Barbone¹, Assad Oberai²
¹ College of Engineering, Boston University, Boston, MA
² Mechanical, Aerospace, and Nuclear Engineering, Rensselaer Polytechnic Inst., Troy, NY
E-mail: cerivas@bu.edu, barbone@bu.edu, oberaa@rpi.edu

Abstract. Many inverse problems are formulated and solved as optimization problems. In this approach, the data mismatch between a predicted field and a measured field is minimized, subject to a constraint. The constraint represents the “forward” model of the system under consideration. In this paper, the model considered is plane stress incompressible elasticity. This pde is discretized using several standard Galerkin finite element methods. These are known to yield stable and convergent discrete solutions that converge with mesh refinement to the exact solution of the forward problem. It is usually taken for granted that if the constraint equation is discretized by a stable, convergent numerical method, then the inverse problem will also converge to the exact solution with mesh refinement. We show examples in this paper, however, where this is not the case. These are based on inverse problems with interior data, which have provably unique solutions. Even so, the use of classical discretization techniques for the forward constraint within the optimization formulation leads to ill-posed discrete problems. We analyze the discrete systems of equations and show the source of the instability. We discuss variational properties of the continuous inverse optimization problem, and describe a novel B-spline FEM to solve it. We present computational evidence that suggests the B-spline FEM inverse problem solution converges to the exact inverse problem solution with mesh refinement.

1. Introduction

The discrete practical solution of inverse problems often proceeds as follows. A discrete forward model of the experiment is constructed that can be used to compute a predicted data set. The parameters within the forward model are then systematically changed until the predicted data match the measured data. This procedure is most appropriately formulated as a constrained optimization problem.

The constraint represents the “forward” model of the system under consideration. In many problems, it takes the form of an elliptic second order partial differential equation (PDE). When this PDE is discretized using the classical finite element function spaces, standard Galerkin methods yield stable and convergent discrete solutions that converge with mesh refinement to the exact solution of the forward problem. It is usually taken for granted that if the constraint equation is discretized by a stable, convergent numerical method, then the discrete inverse solution will also converge to the exact solution with mesh refinement. In this paper, we show examples, however, where this is not the case. These are based on inverse problems with interior data, which yield provably well-posed inverse problems. Even so, the use of classical discretization techniques within the optimization formulation yields discrete formulations that
fail to converge to the exact solution, even though their forward solutions are provably convergent.

We first recognized this phenomenon in the context of elasticity imaging. In elasticity imaging, we deal with the problem of identifying the mechanical properties of elastic solids based on measurements of interior displacement data, and possibly some boundary conditions. In this paper, we limit our attention to plane stress incompressible elasticity. In this case, an exact analytical solution is available in nearly closed form (up to a quadrature)[1, 5]. Thus it is known that the solutions are unique and that the problems are well-posed, provided the data is sufficiently smooth. As such, the plane stress inverse problem provides a class of ideal test problems to evaluate the performance of numerical methods designed to solve inverse potential problems. Many of the lessons learned here in the context of plane stress elasticity apply to other inverse scalar and vector potential problems where interior data are available.

Problems with interior data tend to be better conditioned than the more classical inverse problem that has only exterior data available. Given that, a discrete formulation that exhibits instability in the relatively well-conditioned problem with interior data, is likely to fail also in the problem with exterior data. Said another way, if a formulation fails with the easy problems, there is little hope for its success in the hard problems.

In the next section we pose a plane stress inverse problem associated with elasticity imaging. We formulate it as an optimization problem, and discuss its numerical solution upon discretization. We then show the results of two typical discretizations; these fail to give the correct solution. In the third section, we present some analysis of the methods used in section 2. We show that the discrete problem fails ellipticity, and that a simple constraint count shows the problem to be underconstrained. We then introduce in section 4 a novel, nonstandard, discretization of the optimization problem. This uses B-splines to represent the $H^2$ displacement field with $C^1$ continuity. Constraint counts indicate that this method is appropriately constrained, and numerical evidence suggests that it is convergent. Finally, in section 5, we discuss the results and try to draw general conclusions for the numerical solution of inverse problems as optimization problems.

2. Background and Motivation

2.1. Inverse Problem Formulation

2.1.1. Strong Formulation Here we consider an incompressible elastic sheet occupying the two dimensional domain $\Omega$, with piecewise smooth boundary $\Gamma$. The elastic properties of the material are characterized by the shear modulus $\mu(x)$ and Poisson’s ratio, $\nu$. The condition of incompressibility implies that the Poisson’s ratio takes the value $\nu = 1/2$. Thus the only unknown material parameter is the shear modulus, $\mu$.

We suppose that we are given the measured displacement field $u^m(x)$, everywhere inside $\Omega$. We seek to find a modulus distribution $\mu(x)$ that is "most consistent" with $u^m$ in the following sense. Given $\mu(x)$, we define the “predicted” displacement field $u(x)$ to be the solution of the plane stress elasticity equations:

$$2 \nabla(\mu \nabla \cdot u) + \nabla \cdot (\mu \nabla u + \mu \nabla u^T) = 0 \quad x \in \Omega$$

$$u = u^m \quad x \in \Gamma.$$  

Our solution is that modulus distribution $\mu(x)$ that leads to the minimum value of $\pi_0[\mu] = \frac{1}{2} \| u - u^m \|^2$.

When the given displacement field $u^m(x)$ satisfies certain solvability conditions, an exact (strong) solution for $\mu(x)$ exists, which is given by [1]:

$$\mu(x) = \mu(x_o) \exp \left\{ - \frac{x}{\int_{x_o}^{x_o} [2(\nabla \cdot u)1 + 2\epsilon]^{-1}[2\nabla(\nabla \cdot u) + 2\nabla \cdot \epsilon] \cdot dy} \right\}.$$  

2
In (3), $\mu(x_o)$ is the single specified constant required to make the solution unique, and $\epsilon$ is the (2D) measured strain tensor. It is thus clear from the form of equation (3) that the solution of the inverse problem is unique and well-defined for any reasonably smooth measured displacement field that satisfies the solvability conditions described in [1].

2.1.2. Weak Formulation

The typical approach, which we adopt here, is to formulate the optimization problem variationally, and to enforce the constraint through a Lagrange multiplier. To that end we introduce the Lagrangian $L[u, \lambda, \mu]$, and weak form of the elasticity constraint:

$$L[u, \lambda, \mu] = \frac{1}{2} \| \mu - u^m \|^2_N + a_1(\lambda, u; \mu)$$

$$a_1(\lambda, u; \mu) = \left( \nabla \lambda, 2 \mu \nabla \cdot u + 2 \mu \epsilon \right)$$

The choice of the data mismatch norm, $\| \cdot \|_N$, is user and application dependent, but $N = 0$ (i.e. the $L^2$ norm) is typical. Here, $(\cdot, \cdot)$ is the $L^2$ inner product over $\Omega$.

The typical approach is to make this Lagrangian stationary in the function spaces consistent with classical finite elements, but we will have cause later to consider generalizations of these. For now, we introduce the following function spaces for the solutions and their variations:

$$S_1 = \{ u \in H^s(\Omega) \mid u = u^m \text{ on } \Gamma_g \}$$

$$V_1 = \{ v \in H^s(\Omega) \mid v = 0 \text{ on } \Gamma_g \}$$

$$P_1 = \{ \lambda \in H^l(\Omega) \mid \lambda = 0 \text{ on } \Gamma_g \}$$

$$A_1 = \{ \mu \in H^m(\Omega) \mid \int_\Omega \mu \, d\Omega = \mathcal{P} \}$$

$$B_1 = \{ \gamma \in H^m(\Omega) \mid \int_\Omega \gamma \, d\Omega = 0 \}$$

The mnemonic employed here is that $s$ is the order of the solution space, $l$ is the order of the Lagrange multiplier space, and $m$ is the order of the modulus space. In all practical applications, the case $s = l$ is used. For second order elliptic equations like elasticity, the values $s = l = 1$ are the obvious choices motivated by the forward problem. We next consider two discretizations of this problem obtained from this “obvious” choice of the function spaces.

2.2. Test Problem

As a test problem, we consider the displacement field corresponding to simple shear, $u_x = \varphi y$, and $u_y = 0$; ($\varphi$ = constant). Substituting this into the exact solution (3) shows that the (unique) solution to this inverse problem is $\mu = \text{constant}$.

We discretize by choosing $s = l = 1$, and choosing finite dimensional subsets $S_1^h \subset S_1$, $V_1^h \subset V_1$, etc, in accordance with standard FEM interpolation. We choose the Newton method to solve the system of nonlinear equations that comes from this minimization problem and the Galerkin approximation to discretize the linear sub-problem that is formed at each iteration. At each step in the Newton iteration, the updates for $u$ and $\lambda$ are eliminated from the discrete equations to obtain an equation for the update to the modulus, $\mu$. The coefficient matrix in this equation, $M$, is the “reduced Hessian”, and represents the (discrete) second derivative of $\pi_\sigma$ with respect to the modulus. Based on the fact that the exact (strong) solution of the inverse problem is unique, we can expect the Hessian of the minimization problem to be positive definite, particularly near the exact solution.
Here we consider the most straightforward discretization of the functional (4). This is formed by taking $s = l = 1$, $m = 0$, and choosing subsets $S^h \subset S_1$, $V^h \subset V_1$, etc. We chose continuous bilinear shape functions for the displacements and Lagrange multipliers, and discontinuous piecewise constant shape functions for the shear modulus. A mesh of $10 \times 10$ elements was employed for this reconstruction. For this solution we prescribed $\mu = 1$ in one node to fix the one constant. Figure (1a) shows the reconstruction for the shear modulus in a square domain. Clearly this reconstruction is inconsistent with the exact solution, $\mu = 1$, and even violates positivity of the modulus. The eigenvalues of the reduced Hessian at the last iteration are shown in figure (1b). From this figure, we can see that this matrix is not positive definite as might be expected from the analytical solution. We can see clearly that there is one zero-eigenvalue that contaminates our solution.

2.4. Formulation 2
The nature of the "spurious" solution exhibited in the previous example seems to depend upon the fact that the discretization for $\mu^h$ is discontinuous. Indeed, the derivation of the exact strong solution depends upon the assumption that $\mu$ is differentiable. Therefore, one might expect that using a continuous basis for $\mu^h$ might eliminate the spurious solution. Figure (2a) shows the reconstruction obtained for the same case as above, but using a bilinear $C^0$ basis for $\mu^h(x)$. This solution is also clearly wrong. In fact, the Hessian is more singular than before, with seven zero eigenvalues; see figure (2b). Thus using $C^0$ interpolation for $\mu$ made the problem worse rather than better.

3. Analysis
Some understanding of the issues here may be reached by analyzing the equations for the Newton update. At the $k^{th}$ iteration we have:

$$x_{k+1} = x_k + \delta x$$ (11)
Here \( \delta x = \{ \delta u, \delta \mu, \delta \lambda \} \). We note that updates are in the same spaces as their variations; i.e. \( \delta u \in V_1, \delta \mu \in B_1 \), and \( \delta \lambda \in P_1 \). These variables satisfy the following saddle point problem:

\[
\begin{align*}
    a(\delta u, \delta \mu; v, \gamma) + b(\delta \lambda; v, \gamma) &= (f_1, v) + (g_1, \gamma) \quad \forall (v, \gamma) \in V_1 \times B_1 \\
    b(w; \delta u, \delta \mu) &= (s_1, w) \quad \forall w \in P_1
\end{align*}
\]

Here the operators \( a(\cdot, \cdot) \) and \( b(\cdot, \cdot) \) are defined as:

\[
\begin{align*}
    a(\delta u, \delta \mu; v, \gamma) &= (v, \delta u)_N + a_1(\lambda_k, v; \delta \mu) + a_1(\lambda_k, \delta u; \gamma) \\
    b(w; \delta u, \delta \mu) &= a_1(w, \delta u; \mu_k) + a_1(w, u_k; \delta \mu)
\end{align*}
\]

and, \( u_k, \mu_k \) and \( \lambda_k \) are the current guesses for \( u, \mu \) and \( \lambda \), respectively. The source terms, \( f_1 \), \( g_1 \) and \( s_1 \) are defined as:

\[
\begin{align*}
    (f_1, v) &= -(u_k - u^m, v)_N - a_1(\lambda_k, v; \mu_k) \\
    (g_1, \gamma) &= -a_1(\lambda_k, u_k; \gamma) \\
    (s_1, w) &= -a_1(w, u_k; \mu_k)
\end{align*}
\]

The well-posedness of the saddle-point problem (12-13) may be addressed along the lines of [3]. It turns out that the critical step in this analysis is whether for any choice of \( \delta \mu \in B_1 \), then \( b(w; \delta u, \delta \mu) = 0 \) \( \forall w \in P_1 \) implies

\[
\|\delta u\| \geq C\|\delta \mu\|_m
\]

for some \( C > 0 \) and appropriate choice of norms. When (19) is violated, the functional loses ellipticity.

3.1. Formulation 1 fails \( K^h \) ellipticity

It is relatively easy to see that Formulation 1 violates (19) in the test problem considered. To see this, suppose \( u^h_k = u^m \), and \( \mu^h_k = 1 \); that is, we are at the exact solution. Then (15) can be zero for \( \delta u^h = 0 \), provided:

\[
a_1(w^h, u^m; \delta \mu^h) = 0. \tag{20}
\]
Now we choose \( w^h = e_i N_A(x) \), where \( e_i \) is the Cartesian basis vector in the \( x_i \) direction, and \( N_A(x) \) is the bilinear shape function for interior node \( A \). Integrating (20) by parts thus requires:

\[
\sum_{\text{elements}} \int_{\Gamma^e} 2e_i N_A \epsilon_{ij} n_j \delta \mu \, d\Gamma = 0 \quad (21)
\]

Here we are integrating over the boundaries of those elements that share node \( A \). For the given strain field, \( \epsilon_{xy} = \epsilon_{yx} = \text{constant} \), and equation (21) is satisfied identically on a regular rectangular mesh when \( \delta \mu \) takes on a checkerboard pattern. Interestingly, the checkerboard solution depends not on the cancellation of contributions from adjacent elements, but rather on cancellation of contributions from diagonally opposed elements. That is, the “black squares” of a checkerboard cancel the contributions of the other black squares.

The analysis shown here shows that we can have a nontrivial \( \delta u^h \) with a trivial \( \delta \mu^h \), which is clearly a violation of the condition (19) at the discrete level. The fact that the elasticity equation leads to a unique solution of the inverse problem in the strong form, but not the discrete form, implies that the discrete elasticity constraint is being underenforced. That the saddle point problem is underconstrained for the two formulations considered above may be demonstrated by a simple constraint count.

### 3.2. Constraint count test

Constraint counting provides a simple heuristic to test the viability of mixed finite elements [2]. We define the constraint ratio, \( r \), by

\[
r = \frac{n_{\text{unk}}}{n_{\text{con}}} \quad (22)
\]

Here, \( n_{\text{unk}} \) represents the total number of unknowns in the problem (without considering Lagrange multiplier), while \( n_{\text{con}} \) represents the number of constraint equations. To explain this test we introduce the standard mesh illustrated in figure 3. The conjecture is that as the number of elements per side, \( n_{\text{es}} \), tends to infinity, \( r \) should tend to the corresponding ratio of the number of equilibrium PDEs. So in two dimensions the ideal ratio would be 3/2. If \( r < 3/2 \), then our discrete problem is overconstrained. It means that we have excessive Lagrange multiplier variables or a deficiency of shear modulus variables. If \( r > 3/2 \), our system is underconstrained and we can anticipate a tendency to have spurious solutions.

---

**Figure 3.** Standard mesh for constraint count test
For Formulation 1, that uses bilinear shape functions for the displacement and Lagrange multiplier, and constant shear modulus in each element, \( r \) is:

\[
r_1 = \frac{n_{\text{unk}}}{n_{\text{con}}} = \frac{2(n_{es} - 1)^2 + n_{es}^2}{2(n_{es} - 1)^2} = \frac{3}{2} + \frac{2n_{es} - 1}{2(n_{es} - 1)^2} > \frac{3}{2}.
\]

In particular when \( n_{es} = 10 \), as in figure (1a), then \( r_1 \approx 1.6 > 3/2 \).

Formulation 2 uses bilinear shape functions for all three variables. Thus \( r \) is calculated:

\[
r_2 = \frac{n_{\text{unk}}}{n_{\text{con}}} = \frac{2(n_{es} - 1)^2 + (n_{es} + 1)^2}{2(n_{es} - 1)^2} = \frac{3}{2} + \frac{2n_{es}}{(n_{es} - 1)^2} > \frac{3}{2}.
\]

In this case for \( n_{es} = 10 \), we find \( r_2 \approx 1.7 > 3/2 \). We see that for both formulations, \( r > 3/2 \) and the methods are underconstrained. Furthermore, we see that Formulation 2 is worse than Formulation 1 in this measure, consistent with the observations in the examples from section 2.

**4. Numerical formulation with high order continuity**

We have recently proven (unpublished) that the critical inequality (19) can be satisfied with the choice \( s = 2 \), e.g. when the displacement field \( u \in H^2 \). This motivates a discretization that respects this higher degree of continuity in our dependent variables. To that end, we employ quadratic-spline shape functions for our displacement. The B-splines guarantee \( C^1 \) continuity of the displacement, hence the \( H^2 \) condition is satisfied.

**4.1. Formulation 3: Displacement in \( H^2 \) and constant \( \mu \)**

For this case the constraint ratio, \( r \) is:

\[
r = \frac{n_{\text{unk}}}{n_{\text{con}}} = \frac{2((n_{es} + 2)^2 - (4n_{es} + 4)) + n_{es}^2}{2((n_{es} + 2)^2 - (4n_{es} + 4))} = \frac{3}{2}.
\]

For this case, the constraint ratio has its ideal value and is independent of the mesh size. So we expect good behaviour of this element.

**4.1.1. Numerical Test**

This formulation yielded the exact solution on the test problem of the previous section. We therefore chose to evaluate its convergence on a problem that lies outside the discrete space. This case was tested by solving the inverse problem in which our measured data is:

\[
\begin{align*}
  u_x &= \frac{1}{\alpha} \log(1 + \alpha y) \quad (26) \\
  u_y &= 0 \quad (27)
\end{align*}
\]

Here, \( \alpha \) is a constant. This example is a generalization of the example shown in section 2. This inverse problem has also an analytical solution for the shear modulus:

\[
\mu = \mu_0 (1 + \alpha y) \quad (28)
\]

Here, \( \mu_0 \) is a constant. Figure (4a) shows the shear modulus reconstruction for a small mesh and Figure (4b) shows the performance of this element with mesh refinement. The discrete solution is evidently converging toward the exact solution.
5. Discussion and Conclusions

Inverse problem practitioners typically employ the “guess and check” method of solving an inverse problem. That is, a forward model of the experiment is constructed that can be used to compute a predicted data set. The parameters within the forward model are then systematically changed until the predicted data matches the measured data.

Here we demonstrated that using a stable, convergent numerical method in the forward model does not guarantee a stable convergent inverse solution. That is, even though the forward solution may converge with mesh refinement, the inverse solution may not. This is true even in the context of a “nice” inverse problem such as that studied here, where we have interior data and a provably unique solution. The continuous problem has a unique solution; the discrete problem exhibits spurious solutions indicative of instability of the discretization. That is, the discrete problem arising from standard discretization techniques fails to represent the mathematical structure of the continuous problem.

Of course, whether a variational problem is well-posed or not depends upon the choice of function spaces in which the problem is cast. One observation made here is that the inverse problem seems to require the displacement field to be in $H^2$, while the forward problem is well-posed with a displacement field in $H^1$. Formulation 4 respects this distinction and converges; the other formulations fail to converge.

Though our understanding of this problem is limited at present, the results here may indicate that in order to create a stable and convergent numerical method for inverse problems by the “guess and check” method, standard forward solvers may be inadequate. Instead, fundamentally different forward solvers may be required in order to accurately capture in the discrete problem the mathematical structure of the continuous inverse problem.

References

[1] Barbone P and Oberai A, Elastic modulus imaging: Some exact solutions of the compressible elastography inverse problem, Physics in Medicine and Biology 52 (2007), 1577–1593.
[2] Hughes T J 1987 The Finite Element Method, linear static and Dynamic finite element analysis pp 209
[3] Brezzi F, On the existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers, R.A.I.R.O 8 (1974), 129–151.
[4] Sabin M 1997 Spline Finite elements, PhD Dissertation, Cambridge University, U.K.
[5] Suzuki A Sumi C and Nakayama K, Estimation of shear modulus distribution in soft tissue from strain distribution, IEEE Trans.Biomed.Eng 42 (1995), 193–202.