COUPLING TOPOLOGICAL GRADIENT AND GAUSS-NEWTON METHOD

Jerome Fehrenbach
Mohamed Masmoudi
Laboratoire MIP, CNRS-Universite Paul Sabatier, Toulouse, France

Abstract Topological asymptotic analysis is an emerging method that has been applied with success to shape optimization and shape inverse problems. However, it is not suitable for solving severely ill-posed inverse problems. It is a short term approach and it fails when applied to small inclusions detection in elastographic imaging.

We show in this paper that it is possible to solve this problem by coupling Gauss-Newton and topological gradient methods in a natural way. The data is the displacement under a small compression, only one component of the displacement is given. The inversion problem is solved in two steps. In a first step, we obtain a weight vector for the observations; this is performed by a dual Gauss-Newton method. The second step consists of computing the topological sensitivity relative to the insertion of an inclusion in the medium (a stiff disk), the cost function being weighted with the result of the first step. This method is applied to numerical experiments.

Keywords: Gauss-Newton method, topological gradient, inverse problem, elastography, medical imaging.

1. Introduction

Finding an optimal domain is equivalent to finding its characteristic function. At first sight this $0-1$ optimization problem is not differentiable, but it is possible to obtain the variation of a cost function when the characteristic function is switched from 0 to 1 or from 1 to 0 in a small region. It is called the topological gradient. It has been applied to shape optimization [13, 7, 19, 18] and to shape inverse problems [2, 9, 14, 17] giving very promising results. In comparison with relaxation methods (the characteristic function is replaced by a density function) the topological gradient can be seen as a regularization technique: it reduces the set of admissible solutions from $[0,1]$ to $\{0,1\}$.

Despite this nice property, the topological gradient fails when it is applied to small inclusions detection in elastographic imaging.

We show in this paper that it is possible to solve this difficult problem by coupling topological gradient and Gauss-Newton method.

The second basic idea of this paper is to start with uniform material property. A wrong initial guess does not help to find the solution. For this reason, short term approaches like the steepest descent method are not suitable for this kind of applications. It is also the case of conjugate gradient methods (Fletcher-Reeves and Polak-Ribiere) and BFGS [3], since they behave like the steepest descent method at the first iteration. We show in this paper that the first iteration of the gradient method creates inclusions on the boundary of the domain. This is equivalent to give a wrong initial guess for the remaining iterations. For this reason, our goal is to propose an algorithm having the capability to give a good localization of the inclusions at the first iteration.

We consider a plane stress model for a linear elastic medium. It has a uniform Poisson coefficient ν and its Young's modulus takes two values, E in the background and KE in the inclusion, where $K \neq 1$ is called the contrast.

The data is the displacement under a small compression, only one component of the displacement is given (along the direction of compression), and the objective is to determine the location of the inclusions from these data.

Such an identification is of crucial importance in the medical imaging field. As a matter of fact, prostate or breast cancerous tumors are stiffer than the surroundings and Young's modulus can be 4 to 10 or more times higher. Elastography is an ultrasonic imaging modality that provides an image (called elastogram) of the displacement of tissues after a small external compression [16]. A reliable detection of small inclusion of Young's modulus using the analysis of elastograms would help early detection of cancers.

The paper is organized as follows: in section 2 we present the direct and inverse problems of elasticity, in section 3 we introduce the dual Gauss-Newton method and we show how to build weights for the residual vector. In section 4 we estimate the topological gradient of the weighted cost function with respect to the insertion of a stiff inclusion, and in section 5 we present numerical simulation results.

2. The direct and inverse problems

Let Ω be a 2D domain, its boundary $\partial\Omega$ is divided in two parts of positive measure: $\partial\Omega = \Gamma_N \cup \Gamma_D$. We consider $g \in H^{1/2}(\Gamma_N)$, $u_D \in H^{1/2}(\Gamma_D)$, $E \in L^\infty(\Omega)$ and $\nu \in \mathbf{R}$. The direct problem is (isotropic linear elasticity):

$$\begin{cases} \nabla \cdot \sigma(u) = 0 & \Omega \\ \sigma(u).n = g & \Gamma_N \\ u = u_D & \Gamma_D, \end{cases} \tag{1}$$

where

$$\sigma(u) = \frac{\nu E}{(1+\nu)(1-2\nu)} \mathrm{tr}\, \epsilon(u) I + \frac{E}{2(1+\nu)} \epsilon(u)$$

is the stress tensor and $\epsilon(u) = \frac{1}{2}(\nabla u + \nabla u^T)$ is the linearized strain tensor. The solution u of (1) is the elastic displacement of the body Ω.

The variational formulation of this problem reads

$$\begin{cases} \text{find } u \in \mathcal{V} \text{ s.t.} \\ \forall v \in \mathcal{V} \qquad a(u, v-u) = \ell(v-u), \end{cases} \tag{2}$$

where $\mathcal{V} = \{u \in H^1(\Omega)^2 \mid u|_{\Gamma_D} = u_D\}$,

$$a(u,v) = \int_\Omega \sigma(u) : \epsilon(v)\, dx \quad \text{and} \quad \ell(v) = \int_{\Gamma_N} g.v\, dl.$$

The functions g and u_D, and the Poisson coefficient ν are assumed to be fixed. The solution u of (1) depends on the spatial distribution of Young's modulus E, and the parameter-to-state map is defined by

$$\begin{array}{rccc} U: & L^\infty(\Omega) & \longrightarrow & \mathcal{V}(\Omega) \\ & E & \longmapsto & u_E, \end{array}$$

where u_E is the solution of (1). It is proved in [6] that the parameter-to-state map U is differentiable, and has a compact differential, at every point where E is bounded below by a positive constant.

We consider also a continuous linear observation map $L : \mathcal{V} \to X$ where X is a Hilbert space. In the case of radial elastography (when the probe gives a radial image), the observation map consists of taking the radial component of a vector field.

The inverse problem is: given an observation, determine Young's modulus distribution E. The data is denoted u_{obs} and it is the measurement estimation of Lu_E.

The inverse problem is

$$\min_E j_0(E) := \frac{1}{2} \left(||LU(E) - u_{obs}||^2 \right).$$

It is highly ill-posed, the classical way to improve its stability is to add a Tikhonov regularization term [11], and look for the minimum of

$$j_\alpha(E) := \frac{1}{2} \left(||LU(E) - u_{obs}||^2 + \alpha ||E||^2 \right).$$

Among known techniques for minimizing the cost-function j_α in the context of elastography, let us mention the steepest descent method [15] and Gauss-Newton method [4–6, 10]. The topological sensitivity [1, 2, 7] of j_α with respect to the insertion of an inclusion has also been considered in the frame of this work, but it doesn't give satisfying results.

These methods don't allow to find very small inclusions, we propose in the following to couple the Gauss-Newton and topological gradient methods.

3. Gauss-Newton methods

In this section, we introduce the dual Gauss-Newton method, that will provide us with a weight vector for the observations.

3.1 Primal Gauss-Newton method

After discretization, we keep the same notations of $j, u_E, E, ...$ Then the cost function $j : \mathbf{R}^p \rightarrow \mathbf{R}$ is of the form

$$j_0(E) = \frac{1}{2}||F(E)||^2,$$

where $F : \mathbf{R}^p \rightarrow \mathbf{R}^n$ is the residual $F := LU(E) - u_{obs}$, it is a differentiable vector-valued function. For the sake of simplicity we do not consider the regularization term $\alpha||E||^2$.

The well-known Gauss-Newton algorithm for the minimization of j_0 is defined as follows:

$$\begin{cases} E_0 \text{ is given} \\ E_{k+1} = E_k + d_k, \text{ where} \\ DF^T(E_k)DF(E_k)d_k = -DF^T(E_k)F(E_k), \qquad k \geq 0 \end{cases} \tag{3}$$

3.2 Dual Gauss-Newton method

In our context, we have $p > n$, hence the matrix $DF^T DF$ is singular, and the resolution of (3) requires the use of techniques [3, 11] like Levenberg-Marquardt, discrepancy principle, ... There exists an alternative that we call the dual Gauss-Newton method [12]. When $p > n$ and DF is surjective, the matrix $DFDF^T$ is invertible, and the following problem:

$$\begin{cases} \min ||d|| \qquad \text{s.t.} \\ DFd = -F \end{cases} \tag{4}$$

is equivalent to its Euler-Lagrange optimality condition

$$\begin{cases} DF\,DF^T\lambda = F \\ d = -DF^T\lambda. \end{cases} \tag{5}$$

A heuristic justification for the introduction of (4) is the following: the asymptotic expansion of F around E_k is

$$F(E_k + d) = F(E_k) + DF(E_k)d + R(E_k, d),$$

where the rest $R(E_k, d)$ is small: $||R(E_k, d)|| = o(||d||)$. In order to cancel the two first terms, the descent direction has to satisfy $DF(E_k)d = -F(E_k)$. Among the affine vector space of solutions, taking d with minimal norm should make the rest $R(E_k, d)$ small.

In this paper we do not iterate the dual Gauss-Newton method, but we solve (5) once and use the vector $\lambda \in \mathbf{R}^n$ as a weight for the residual $LU(E) - u_{obs}$. More precisely, we are interested in the topological sensitivity of the following "weighted" cost function

$$j_W(E) = F(E)^T \lambda,$$

where

$$F(E) = LU(E) - u_{obs}. \tag{6}$$

This makes sense for the following reason: the descent direction given by the dual Gauss-Newton algorithm is $d = -DF^T \lambda = -D(F^T \lambda)$. Then $-d$ is the gradient of the cost function $F^T \lambda$, and it is natural to consider the topological variation of $F^T \lambda$. Notice that, thanks to the positivity of $DF\, DF^T$, the cost function $F^T \lambda$ is positive.

3.3 A zero memory Gauss-Newton implementation

The linear system in (5) is solved using the conjugate gradient (CG) method. It gives a good approximation of λ after a few iterations because $DF\, DF^T$ is the discrete version of a compact operator: its eigenvalues tend rapidly towards zero. Each iteration of CG requires performing the product of $DF\, DF^T$ by a vector, this can be performed without computing and storing the whole Jacobian [8, 6]. More precisely, when a vector x is given, $y = DF\, DF^T x$ is computed in two steps: $t = DF^T x$ is evaluated by solving an adjoint equation, and $y = DFt$ is evaluated by solving a direct problem.

4. The topological gradient

In this section, we study the topological sensitivity of the cost function $J_W(u) := (Lu - u_{obs})^T \lambda$ with respect to the insertion at point x_0 of a very stiff disk. The Young's modulus takes two values, 1 in the background and $K \gg 1$ in the inclusions.

The cost function J_W depends linearly on u, hence the adjoint is the solution $p \in \mathcal{V}_0$ of

$$\forall w \in \mathcal{V}_0, \qquad a(w, p) = -(L\, w)^T \lambda, \tag{7}$$

where $\mathcal{V}_0 = \{u \in H^1(\Omega) \mid u|_{\Gamma_D} = 0\}$.

The inclusion is εB where B is the unit disk. The first term in the asymptotic expansion of the cost function j_W is

$$\varepsilon^2 |B| \epsilon(u) M \epsilon(p),$$

where M is the elastic moment tensor of the disk inclusion.It is proved in [1] that the coefficients of the tensor M satisfy $m_{pq}^{ij} = m_{qp}^{ij}$, $m_{pq}^{ij} = m_{ij}^{pq}$ and have the following expression depending on the Lame coefficients of the material $\lambda_0 = \frac{E\nu}{(1+\nu)(1-2\nu)}$ and $\mu_0 = \frac{E}{2(1+\nu)}$, and on $\kappa = \frac{\lambda_0 + 3\mu_0}{\lambda_0 + \mu_0}$:

$$m_{11}^{11} = m_{22}^{22} = (\lambda_0 + 2\mu_0)(K-1)\frac{(\lambda_0 + \mu_0)(1 + 2\kappa K - K) + \mu_0(\kappa - 1)}{1 + K(\lambda_0 + \mu_0)(1 + \kappa K)},$$

$$m_{11}^{22} = (\lambda_0 + 2\mu_0)\frac{\lambda_0(K-1)(1+\kappa K) + \mu_0(K-1)^2}{(\mu_0 + K(\lambda_0 + \mu_0))(1 + \kappa K)},$$

$$m_{12}^{12} = \frac{\mu_0(K-1)(1+\kappa)}{1 + \kappa K}.$$

The topological gradient of j at the point x_0 is:

$$T(x_0) = \epsilon(u) M \epsilon(p) + (K-1)\sigma(u) : \epsilon(p). \tag{8}$$

The procedure for calculating the topological sensitivity of J_W goes as follows:

1- The initial guess is $E \equiv 1$. Compute the direct state u,

2- The weight λ is determined by solving

$$(DF\, DF^T + \alpha I)\lambda = F, \tag{9}$$

where α is a Tikhonov regularization parameter and $F = Lu - u_{obs}$,

3- compute the adjoint state p solution of (7),

4- compute $\epsilon(u)$ and $\epsilon(p)$,

5- at each point x_0, determine the value of $T(x_0)$, using (8).

The higher isovalues of T give the location of the inclusion.

5. Application to a numerical experiment

A numerical simulation was conducted in order to simulate an in vitro experiment [20]. The method described in section 4 is applied as follows. The domain Ω is an annulus, on the inner boundary Γ_D a uniform radial displacement is applied, and the outer boundary Γ_N is free, see figure 1.

The elasticity problem is discretized using P1 finite elements. It is implemented using `Matlab` and `GetFem++`, the mesh has 3077 elements and 1606 nodes. The total computation time for the procedure described above is about

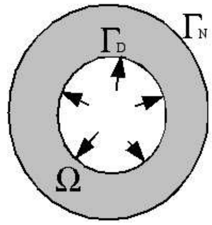

Figure 1. The domain Ω and the Dirichlet boundary

1 minute on a desktop computer (Intel Pentium M processor, 1.6 GHz, 512Mo RAM).

The Poisson coefficient is $\nu = 0.45$. We insert 1 or 2 inclusions in the domain, where Young's modulus is 4 times higher than in the background. The data are perturbed with 0.1% noise, and with 1% noise (white Gaussian noise). The weight λ is computed by solving (9). Once the weight λ is known, the adjoint state p is computed. The topological gradient is computed with $K = 50$ (value of the contrast).

The results on different geometries are displayed on figures 2-7.

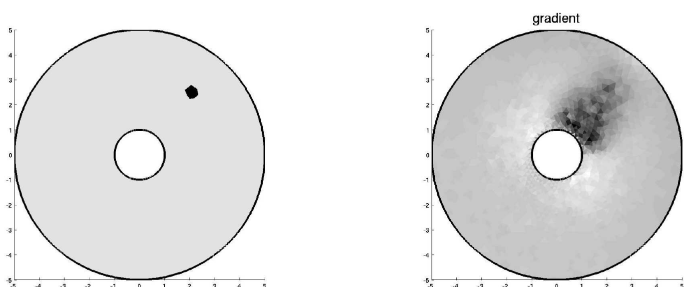

Figure 2. The inclusion (left) ; the gradient of j_0 (right)

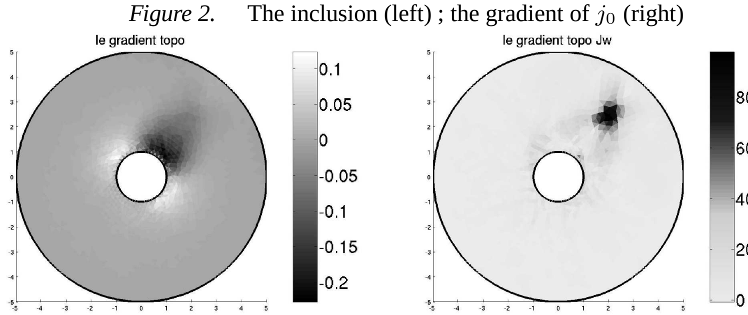

Figure 3. The topological gradient of j_0 with 0.1% noise (left) ; the topological gradient of j_W (right)

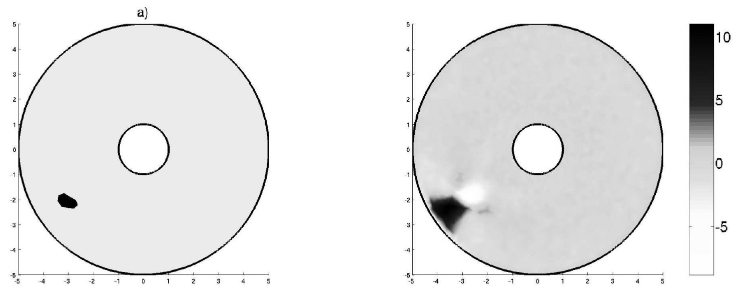

Figure 4. The inclusion (left) ; the weight λ for 0.1% noise (right)

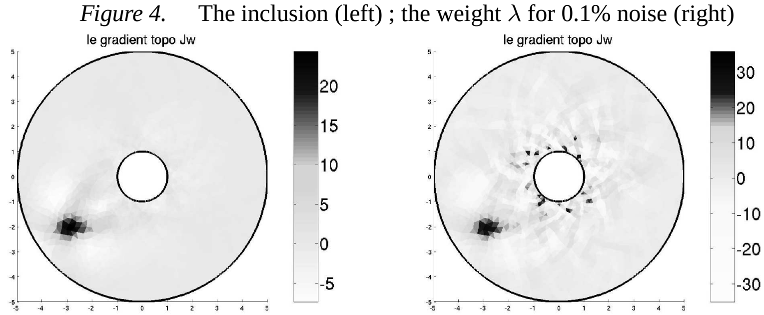

Figure 5. The topological gradient of j_W with 0.1% noise(left) and 1% noise (right)

The following observations can be made:

1) The gradient of the classical cost function $j_0 = \frac{1}{2}||F||^2$ gives no interesting information: its maximum is on the boundary, see figure 2 right,

2) The topological gradient of j_0 gives no interesting results, see figure 3 left,

3) The topological gradient of the weighted cost function $j_W = F^T\lambda$ detects the location of the inclusion with a reasonable accuracy, even for noisy measures (figure 3 right, figures 5 and 7),

4) Our method to determine the weight is justified *a-posteriori*: this weight gives more importance to the measures that are located “in front” and “behind” the inclusion (relatively to the compression direction), see figures 4 right and 6 right. These are precisely the points where the displacements are the more affected by the inclusion, and the data are the more relevant to detect it. In the points “in front” of the inclusion, the weight takes negative values and this makes sense: at these points F also takes negative values, hence $F^T\lambda$ decreases when F gets closer to zero,

5) When more noise is added, the detection of inclusions becomes much less accurate.

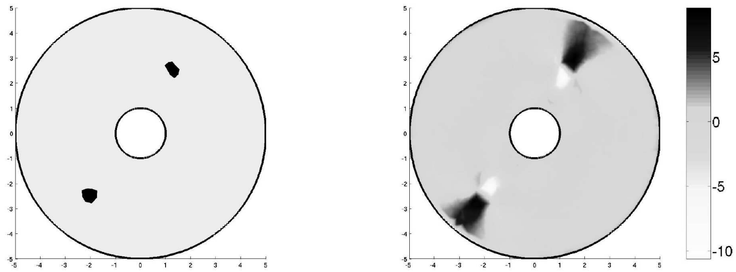

Figure 6. The inclusion (left) ; the weight λ for 0.1% noise (right)

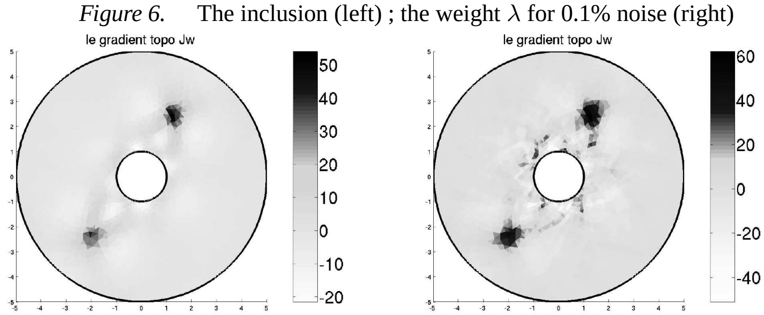

Figure 7. The topological gradient of j_W with 0.1% noise (left) and 1% noise (right)

Conclusions

We presented in this paper a method for the detection of small inclusions in an elastic medium, when the data is the radial displacement at each point. The first step consists of finding a weight for the measurements that gives importance to the data in the relevant zones, this weight is determined using a dual Gauss-Newton method and an efficient implementation. The second step consists of computing the topological sensitivity of the so weighted cost function.

Unlike classical methods, this method gives satisfying results for small inclusions.

Acknowledgments

This work was supported in part by ACINIM-IMEG CNRS project (France).

References

[1] H.Ammari, H.Kang (2004) *Reconstruction of small inclusion from boundary measurements*, Lecture Notes in Mathematics, vol.1846, Springer

[2] S.Amstutz, I.Horchani, M.Masmoudi (2005) *Crack detection by the topological gradient method*, Control and Cybernetics **34**(1), 81–101.

[3] J.Dennis, R.Schnabel (1996) *Numerical methods for Unconstrained Optimization and Nonlinear Equations*, Classics in Appl.Math.Sc, SIAM.

[4] M.Doyley, P.Meaney, J.Bamber (2000) *Evaluation of an iterative reconstruction method for quantitative elastography*, Phys. Med. Biol **45**, 1521–1540.

[5] M.Doyley, S.Srinivasan, S.Pendergrass, Z.Wu, J.Ophir (2005) *Comparative evaluation of strain-based and model-based modulus elastography*, Ultrasound Med. Biol **31**(6), 787–802.

[6] J.Fehrenbach, M.Masmoudi, R.Souchon, P.Trompette *Detection of inclusion of Young's modulus*, preprint.

[7] S.Garreau, Ph.Guillaume, M.Masmoudi (2001) *The topological asymptotic for PDE systems: the elasticity case*, SIAM J. Control. Optim **39**(6), 1756-1778.

[8] A.Griewank (2000) *Evaluating derivatives, Principles and Techniques of Algorithmic Differentiation*, Frontiers in Applied Mathematics, SIAM.

[9] M.Hassine, M.Masmoudi (2004) *The topological asymptotic expansion for the Quasi-Stokes problem*, ESAIM Control Optim. Calc. Var.**10**, (4), 478–504.

[10] F.Kallel, M.Bertrand (1996)*Tissue elasticity reconstruction using linear perturbation method*, IEEE Transactions on medical imaging **15**(3), 299–313.

[11] A.Kirsch (1996) *An introduction to the mathematical theory of inverse problems*, Springer.

[12] M.Masmoudi (1990) *Conception de circuits passifs de tres haute fidelite*, Matapli **31**.

[13] M. Masmoudi (2002) *The Toplogical Asymptotic*, in Computational Methods for Control Applications. International Series GAKUTO.

[14] M.Masmoudi, J.Pommier, B.Samet (2005) *The topological asymptotic expansion for the Maxwell equation and some applications* Inverse Problems **21** 547–564.

[15] A.Oberai, N.Gokhale, G.Feijoo (2003) *Solution of inverse problems in elasticity imaging using the adjoint method*, Inverse Problems **19**, 297–313.

[16] J.Ophir, I.Cespedes, H.Ponnekanti, Y.Yazdi, X.Li (1991) *Elastography: a quantitative method for imaging the elasticity of biological tissues*, Ultrasonic imaging **13**, 111–134.

[17] B.Samet, S.Amstutz, M.Masmoudi (2003) *The topological asymptotic for the Helmholtz equation.* SIAM J. Control. Optim.**42**(5), 1523–1544.

[18] A.Schumacher (1995) *Topologieoptimisierung von Bauteilstrukturen unter Verwendung von Lopchpositionierungkrieterien.* Thesis, Universitat-Gesamthochschule-Siegen.

[19] J.Sokolowski, A.Zochowski (1999) *On the topological derivative in shape optimization.* SIAM J. Control Optim.**37**, 1241–1272.

[20] R.Souchon, O.Soualmi, M.Bertrand, J.Y.Chapelon, F.Kallel, J.Ophir (2002) *Ultrasonic elastography using sector scan imaging and a radial compression* Ultrasonics **40**(1-8) 867–871.