Spring 2024: Computational and Variational Methods for Inverse Problems
CSE 393P/GEO 391/ME 397/ORI 397
Assignment 4: Inverse Problems Governed by PDE Forward Models
Due April 29, 2024

The problems below require a mix of paper-and-pencil work and FEniCS implementation. Jupyter notebooks that provide example implementations in FEniCS of the steepest descent and inexact Newton CG methods are available on our JupyterHub cloud in the `05_Poisson_SD` and `06_Poisson_INCG` folders. Feel free to build on these, as well as the example code in `Assignments/Assignment4` implementing the Helmholtz equation, in solving Problems II and III below.

Please turn in a single PDF document with your answers (including figures and plots when requested) to the problems below. For Problems II and III, please do **not** include printouts of the notebooks in your solution document. Instead, please only include plots and figures (like you would report them in a journal article) and selected code snippets showing your Fenics implementation of the cost functional, gradient computation, and Hessian action. Solution documents that include raw printouts of your notebook will **not** be graded.

I. Frequency-domain inverse wave propagation problem

Here we formulate and solve an inverse acoustic wave propagation problem in the frequency domain, which is governed by the Helmholtz equation. Let $\Omega \subset \mathbb{R}^d$ be a bounded domain ($d \in \{2, 3\}$) with boundary Γ. The idea is to insonify the object with harmonic waves $u_{ij}^{\text{inc}}(\boldsymbol{x})$ from multiple sources ($j = 1, \ldots, N_s$) at multiple frequencies ω_i ($i = 1, \ldots, N_f$) and measure the amplitude $u_{ij}(\boldsymbol{x})$ of the scattered wavefields along the boundary of Ω for each source and frequency, with the goal of inferring the soundspeed of the medium, $c(\boldsymbol{x})$. In particular, the harmonic waves $u_{ij}^{\text{inc}}(\boldsymbol{x})$ are the solution of the Helmholtz equation in an acoustic homogeneous medium with speed of sound c_0.

We can formulate this inverse problem with Tikhonov regularization as follows:

$$\min_m \Phi(m) := \frac{1}{2} \sum_i^{N_f} \sum_j^{N_s} \int_\Gamma (u_{ij}(m) - d_{ij})^2 \, d\boldsymbol{s} + \frac{\beta}{2} \int_\Omega \nabla m \cdot \nabla m \, d\boldsymbol{x}$$

where $u_{ij}(\boldsymbol{x})$ depends on the medium parameter field $m(\boldsymbol{x})$, which is equal to $1 - c_0^2/c(\boldsymbol{x})^2$, through the solution of Helmholtz problem

$$-\Delta u_{ij} - k_{0,i}^2 \, (1 - m) \, u_{ij} = k_{0,i}^2 \, m \, u_{ij}^{\text{inc}} \quad \text{in } \Omega, \quad i = 1, \ldots, N_f, \quad j = 1, \ldots, N_s,$$
$$\frac{\partial u_{ij}}{\partial n} = 0 \quad \text{on } \Gamma.$$

In the problem above, $k_{0,i} = \frac{\omega_i}{c_0}$ denotes the wave number, $d_{ij}(\boldsymbol{x})$ denotes given measurements (for frequency i and source j), u_{ij} is the amplitude of the scattered acoustic wavefield, and $\beta > 0$ is the regularization parameter.

1. Derive an expression for the (infinite dimensional) gradient of Φ with respect to the medium m using the Lagrangian method for a single source and frequency, i.e., for $N_f = N_s = 1$. Give both weak and strong forms of the gradient.

2. Derive an expression for the (infinite dimensional) action of the Hessian of Φ in a direction $\tilde{m}$ in the single source and frequency case. Give both weak and strong forms of the Hessian action.

3. Derive an expression for the (infinite dimensional) gradient for an arbitrary number of sources and frequencies.[1] How many state and adjoint equations have to be solved for a single gradient computation?

II. The steepest descent method for solving the Helmholtz inverse problem

We would like to solve the frequency-domain inverse wave propagation problem we discussed above, with some slight modifications. In particular, we consider a single frequency and a single (superposed) source. Additionally, the inversion parameter m representing the medium properties is such that $\tanh(m(\boldsymbol{x})) = 1 - \frac{c_0^2}{c(\boldsymbol{x})^2}$. Here, the hyperbolic tangent reparameterization ensures that the squared slowness, $c(\boldsymbol{x})^{-2}$ is nonnegative and no more than twice the reference value $c_0(\boldsymbol{x})^{-2}$. With the above changes, the inverse problem is:

$$\min_m \Phi(m) := \frac{1}{2}\int_\Gamma (u(m) - d)^2 \, d\boldsymbol{s} + \frac{\beta}{2}\int_\Omega \nabla m \cdot \nabla m \, d\boldsymbol{x}, \tag{1}$$

where the scattered wave field $u(\boldsymbol{x})$ depends on the medium properties $m(\boldsymbol{x})$ through the solution of the Helmholtz equation

$$-\Delta u - k_0^2\,(1 - \tanh(m))\,u = k_0^2\,\tanh(m)\,u^{\text{inc}} \quad \text{in } \Omega \tag{2}$$

$$\frac{\partial u}{\partial n} = 0 \quad \text{on } \Gamma. \tag{3}$$

Above, the computational domain $\Omega = \{\boldsymbol{x} \in \mathbb{R}^2 | \, \|x\|^2 \leq 1\}$ is the unit circle, $k_0 = 5$ is the wave number, and $\beta > 0$ is the regularization parameter. The incident wavefield $u^{\text{inc}}(\boldsymbol{x})$ is the superposition of three spherical waves with wave number k_0 generated by point-like sources. That is,

$$u^{\text{inc}}(\boldsymbol{x}) = \sum_{j=1}^{3} -e_i \frac{\cos\left(k_0|\boldsymbol{x} - \boldsymbol{x}_i|\right)}{4\pi|\boldsymbol{x} - \boldsymbol{x}_i|}, \tag{4}$$

where $\boldsymbol{x}_i$ denote the locations of the three point-like sources and $e_i \sim N(0,1)$ are randomly chosen coefficients.

We synthesize the measurement data $d(\boldsymbol{x})$ by solving the Helmholtz equation with the "true" medium properties

$$m(x,y) = \begin{cases} 0.2 & \text{if } \sqrt{(x-0.1)^2 + 2(y+0.2)^2} < 0.5 \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Independent and identically distributed Gaussian noise is added to this "data" to simulate actual instrument noise. The mesh of the computational domain Ω, as well as the FEniCS implementation of the forward Helmholtz equation in (2)–(3) with the true medium properties given in (5) and the incident wave given in (4), are provided on the JupyterHub cloud in the folder `Assignments/Assignment4`.

On the JupyterHub, we also provide the FEniCS code `05_Poisson_SD/Poisson_SD.ipynb`, which is an implementation of the steepest descent method[2] for an inverse problem governed by the Poisson

[1] Hint: Every state equation with solution u_{ij} has a corresponding adjoint variable p_{ij}. The Lagrangian functional now involves the sum over all state equations.

[2] Note that we do not advocate the use of the steepest descent method for variational inverse problems, since its convergence will generally be very slow. You will see how slow it is! For all cases, please use a maximum number of iterations of at least 1000—you will need it. In Problem III, we will solve this problem using an inexact Newton-CG method, which is far more efficient.

equation. Note that the steepest descent method for infinite dimensional functions is stated in weak form (using our lecture notes notation). The k-th iteration reads

$$\int_\Omega \hat{m}\tilde{m}_k \, d\boldsymbol{x} = -(\mathcal{G}(m_k), \hat{m}) \quad \forall \hat{m} \in \mathcal{M}, \tag{6}$$

where $\tilde{m}_k$ is the steepest descent search direction, $\hat{m}$ is a variation of the parameter field m, $\mathcal{M}$ is the space of admissible functions (for example, H^1), and $(\mathcal{G}(m_k), \hat{m})$ is the Fréchet derivative of $\Phi(m)$ with respect to m, in the direction $\hat{m}$, evaluated at m_k.[3] When the $L^2(\Omega)$-inner product on left side of this equation is discretized with finite elements, a "mass matrix" $\boldsymbol{M}_{ij} = \int_\Omega \phi_i \phi_j$ arises, which means a linear system with coefficient matrix $\boldsymbol{M}$ must be solved to compute the steepest descent direction.

You should modify the implementation in the notebook to solve the Helmholtz inverse problem with boundary information. Please turn in your modified code along with numerical results and discussion of the following:

1. Report the solution of the inverse problem and the number of required iterations (for a 10^3 relative reduction in the norm of the gradient) for the following cases:

   (a) Noise level of 0.01 (roughly 1% noise), regularization $\beta = 10^{-5}$, and initial guess of $m(\boldsymbol{x}) = 0$. Use an initial step length $\alpha = 1$. Does it converge within 1,000 iterations?

   (b) Same as (a), but using the H^1-inner product to precondition the steepest descent direction. That is the descent direction $\tilde{m}_k$ is computed by solving

   $$\int_\Omega \hat{m}\tilde{m}_k \, d\boldsymbol{x} + \int_\Omega \nabla\hat{m} \cdot \nabla\tilde{m}_k \, d\boldsymbol{x} = -(\mathcal{G}(m_k), \hat{m}) \quad \forall \hat{m} \in \mathcal{M} \tag{7}$$

   in place of (6)

   (c) Same as (b), but increase the wave number to $k_0 = 10$. What do you observe?

2. (**Optional, for extra credit.**) Since the coefficient m is discontinuous, a better choice of regularization is total variation rather than Tikhonov regularization, to prevent an overly smooth reconstruction. Modify the implementation and plot the result for a reasonably chosen regularization parameter. Specifically (using the same notation as in Assignment 3), consider the smoothed total variation regularization functional:

   $$\mathcal{R}_{TV}^{\varepsilon} := \beta \int_\Omega (\nabla m \cdot \nabla m + \delta)^{\frac{1}{2}} \, d\boldsymbol{x},$$

   where the parameter $\delta = 0.01$ is introduced to make the non-differentiable TV regularization term differentiable. When computing the steepest descent direction use the $L^2(\Omega)$-inner product as in equation (6).

   Just to be sure, note that the first variation of $\mathcal{R}_{TV}^{\delta}$ with respect to $m(\boldsymbol{x})$, i.e., the weak form of the gradient, is given by

   $$\delta_m \mathcal{R}_{TV}^{\delta} := \beta \int_\Omega (\nabla m \cdot \nabla m + \delta)^{-\frac{1}{2}} \, \nabla m \cdot \nabla\tilde{m} \, d\boldsymbol{x} \quad \forall \tilde{m} \in H^1.$$

3. (**Optional, for extra credit.**) Solve the Helmholtz inverse problem using the three sources imposed independently (i.e., the multiple source inverse problem) rather than as a single superposed source as you did above. What do you observe about the quality of the inverse solution?

[3] We have referred to the right side of (6) as (the negative of) the "weak form of the gradient."

III. The inexact Newton-CG method for solving the Helmholtz inverse problem

Here, we continue solution of the Helmholtz inverse problem by employing the inexact Newton-CG method in place of steepest descent.

The forward model, observation operator, and noise model are the same as those in Problem II. The reference wave number k_0 should be set to $k_0 = 10$, unless otherwise specified.

1. Modify `05_Poisson_INCG/Poisson_INCG.ipynb` (the FEniCS implementation of the inexact Newton and Gauss-Newton methods[4]) to solve the Helmholtz inverse problem defined in (1)–(2). Initialize the random generator seed by including the following line in the beginning of the notebook: `np.random.seed(seed=42)`. Solve the inverse problem for discretizations of the domain using the original mesh and up to two levels of uniform mesh refinement by changing the value of the variable `nrefinements=0` to 1 and then to 2. Report the numbers of Gauss-Newton and total CG iterations for each case. Note that the state and adjoint variables and their incremental variants are discretized with quadratic finite elements, while the inversion parameter field is discretized with linear elements. Discuss how the number of iterations changes as the mesh is refined, i.e., as the parameter dimension increases. In these experiments, the noise level should be fixed to the default value (1%) while the mesh is refined. The "optimal" regularization parameter can be found manually (i.e., by experimenting with a few different values and finding the one that results in a reconstruction that best matches the "true" medium properties), or else by the discrepancy principle (if you are so inclined).

2. Ill-posedness of the inverse problem is closely related to the spectrum (i.e., the eigenvalues) of the Hessian operator.[5]

   (a) Compute the eigenvalues of the regularization-preconditioned Hessian at the solution of the inverse problem for the three levels of mesh refinement considered above. What do you observe?

   (b) Consider the original coarse mesh without refinement and compute the eigenvalues of the regularization-preconditioned Hessian for different values of the reference wave number $k_0 = 5, 10, 12.5$. Use the solution of the inverse problem for the smaller value of k_0 as initial guess for solving the inverse problem with a larger k_0.[6] What do you observe?

   For both case (a) and (b) use the regularization parameter $\beta = 10^{-4}$. To find the solution of the inverse problem use the inexact Newton conjugate gradient method with the Gauss-Newton Hessian (at least for the first 10–20 iterations).

3. (**Optional, for extra credit.**) Replace Tikhonov regularization with total variation regularization and repeat sub-problem 1 above (i.e., report number of Gauss-Newton and CG iterations as the mesh is refined).

IV. (Optional, for extra credit.) An inverse problem for Burgers' equation

Consider the inverse problem for the viscosity field m in the one-dimensional viscous Burgers' equation (this equation is often taken as a one-dimensional proxy for the Navier-Stokes equations). Taking $[0, L]$

[4] Recall that the Gauss-Newton approximation of the Hessian drops all terms (in the incremental adjoint equation and the Hessian action expression) that involve the adjoint variable p.

[5] The eigenvalues usually decay rapidly such that inversion of the Hessian can be done in a stable manner only with the use of regularization.

[6] This technique, known as frequency hopping, helps to mitigate the nonlinearity of the high-frequency Helmholtz inverse problem.

as the spatial domain and $[0, T]$ as the temporal interval, the solution $u = u(x,t) : [0, L] \times [0, T]$ satisfies

$$u_t + uu_x - (mu_x)_x = f \quad \text{in } (0, L) \times (0, T), \tag{8a}$$
$$u(0,t) = u(L,t) = 0 \quad \text{for all } t \in [0, T], \tag{8b}$$
$$u(x,0) = 0 \quad \text{for all } x \in [0, L]. \tag{8c}$$

Here, $m = m(x) : [0, L] \to \mathbb{R}$ is the spatially-dependent viscosity field we wish to invert for, $f = f(x,t)$ is a given source term, and subscripts t and x indicate partial derivatives with respect to time and space coordinates. The conditions (8b) and (8c) are the boundary and initial conditions, respectively.[7] We are given observations $d = d(x,t)$ for *a portion of the time interval*, i.e., for $t \in [T_1, T]$, where $T_1 > 0$. To invert for the viscosity field m, we thus minimize the functional

$$\mathcal{F}(m) := \frac{1}{2} \int_{T_1}^{T} \int_0^L (u - d)^2 \, dx \, dt + \frac{\beta}{2} \int_0^L \frac{dm}{dx} \frac{dm}{dx} \, dx \tag{9}$$

with regularization parameter $\beta > 0$. An efficient optimization method for (9) requires the gradient of $\mathcal{F}$ with respect to m.

1. Derive a weak form of (8) by multiplying (8a) with a test function $p(x,t) : [0, L] \times [0, T]$ that satisfies Dirichlet boundary conditions analogous to (8b), and integrating over space and time. There is no need to impose the initial condition explicitly via a Lagrange multiplier (as we did in class for the initial condition inversion problem), since the inversion parameter (m) does not appear in the initial condition. Instead, you should build the satisfaction of the initial condition into the definition of the solution space (just as you would do with a Dirichlet boundary condition). Use integration by parts on just the viscous term to derive the weak form of the Burgers' equation.

2. Using the Lagrangian approach, derive expressions for the adjoint equation and for the gradient of $\mathcal{F}$ with respect to m. Give weak and strong forms of these expressions. Note that m as well as its variation $\hat{m}$ are functions of space only, while u and the adjoint p are functions of space and time.

[7] Note that the boundary Γ of the one-dimensional interval $\Omega = (0, L)$ is simply the points $x = 0$ and $x = L$, i.e., $\Gamma = \{0, L\}$.