Inverse Problems: homework 3, Spring 2019
Due March 20.
Lindsay Eddy

Instructions. *It is okay, in fact encouraged, to discuss the problems with your peers. However, the work you submit must be your own. For computer problems, please also submit the computer codes used to generate the numerical results.*

1. **FEniCS exercise.** An anisotropic Poisson problem in a two-dimensional domain Ω is given by the strong form

$$-\nabla \cdot (A(\boldsymbol{x})\nabla u) = f \qquad \text{in } \Omega, \tag{1a}$$
$$u = u_0 \qquad \text{on } \partial\Omega, \tag{1b}$$

where $A(\boldsymbol{x}) \in \mathbb{R}^{2\times 2}$ is assumed to be symmetric matrix field that is uniformly bounded and positive definite, f is a given source term, and u_0 is Dirichlet boundary data.

(a) Derive the variational/weak form corresponding to the above problem, and give the energy functional that is minimized by the solution u of (1).

To derive the weak form of the PDE, we multiply by a test function $\tilde{u} \in H_0^1\Omega$:

$$-\int_\Omega \nabla \cdot \big(A(\boldsymbol{x})\nabla u\big)\tilde{u}\ d\boldsymbol{x} - \int_\Omega f\tilde{u}\ d\boldsymbol{x} = 0.$$

Using integration by parts, we obtain

$$-\int_{\partial\Omega} \big[A(\boldsymbol{x})\nabla u\big]\cdot \boldsymbol{n}\ \tilde{u}\ ds + \int_\Omega A(\boldsymbol{x})\nabla u \cdot \nabla \tilde{u}\ d\boldsymbol{x} - \int_\Omega f\tilde{u}\ d\boldsymbol{x} = 0.$$

Since $\tilde{u} = 0$ on $\partial\Omega$, the weak form of the PDE is

$$\int_\Omega A(\boldsymbol{x})\nabla u \cdot \nabla \tilde{u}\ d\boldsymbol{x} - \int_\Omega f\tilde{u}\ d\boldsymbol{x} = 0.$$

The energy functional minimized by the solution u to (1) is

$$\frac{1}{2}\int_\Omega A(\boldsymbol{x})\nabla u \cdot \nabla u\ d\boldsymbol{x} - \int_\Omega fu\ d\boldsymbol{x}.$$

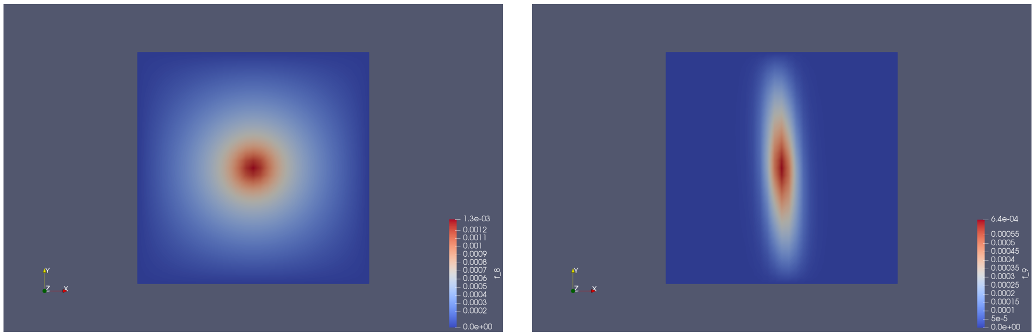
(b) Solve problem (1) in FEniCS using quadratic finite elements. Let $\Omega = [-1,1]\times[-1,1]$, and use

$$f(\boldsymbol{x}) = \exp(-100(x_1^2 + x_2^2)) \quad \text{and} \quad u_0 = 0.$$

Try the following choices for the diffusion matrix:

$$A_1 = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}, \quad \text{and } A_2 = \begin{pmatrix} 1 & -5 \\ -5 & 100 \end{pmatrix}.$$

Compare the results obtained with A_1 and A_2 in (1).



(a) A_1 (b) A_2

Figure 1: Solutions, u, to the PDE decribed by (1), with diffusion matricies A_1 and A_2.

2. The problem of removing noise from an image without blurring sharp edges can be formulated as an infinite-dimensional minimization problem. Given a possibly noisy image $u^{obs}(x, y)$ defined within a square domain Ω, we would like to find the image $u(x, y)$ that is closest in the L_2 sense, i.e. we want to minimize

$$\mathcal{F}_{LS} := \int_\Omega (u - u^{obs})^2 \, d\boldsymbol{x},$$

while also removing noise, which is assumed to comprise very "rough" components of the image. This latter goal can be incorporated as an additional term in the objective, in the form of a penalty, i.e.,

$$\mathcal{R}_{TN} := \frac{1}{2} \int_\Omega k(\boldsymbol{x}) \nabla u \cdot \nabla u \, d\boldsymbol{x},$$

where $k(\boldsymbol{x})$ acts as a "diffusion" coefficient that controls how strongly we impose the penalty, i.e. how much smoothing occurs. Unfortunately, if there are sharp edges in the image, this so-called *Tikhonov (TN) regularization* will blur them. Instead, in these cases we prefer the so-called *total variation (TV) regularization*,

$$\mathcal{R}_{TV} := \int_\Omega k(\boldsymbol{x}) (\nabla u \cdot \nabla u)^{\frac{1}{2}} \, d\boldsymbol{x}$$

where (we will see that) taking the square root is the key to preserving edges. Since $\mathcal{R}_{TV}$ is not differentiable when $\nabla u = \mathbf{0}$, it is usually modified to include a positive parameter ε as follows:

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

We wish to study the some essential properties of the two denoising functionals analytically.

$\mathcal{F}_{TN}$ and $\mathcal{F}^{\varepsilon}_{TV}$, where

$$\mathcal{F}_{TN} := \mathcal{F}_{LS} + \mathcal{R}_{TN} \quad \text{and} \quad \mathcal{F}^{\varepsilon}_{TV} := \mathcal{F}_{LS} + \mathcal{R}^{\varepsilon}_{TV}.$$

We will prescribe the homogeneous Neumann condition $\nabla u \cdot \boldsymbol{n} = 0$ on the four sides of a unit square, which amounts to assuming that the image intensity does not change normal to the boundary. Also, for simplicity, we assume a constant diffusion coefficient $k \equiv \kappa > 0$, from now on.

(a) For both $\mathcal{F}_{TN}$ and $\mathcal{F}_{TV}^{\varepsilon}$, derive the first-order necessary condition for optimality using calculus of variations, in both weak form and strong form.

$$\begin{aligned}
\mathcal{F}_{TN}'(u)(\tilde{u}) &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \mathcal{F}_{TN}(u+\varepsilon\tilde{u}) \\
&= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \left[\int_\Omega (u+\varepsilon\tilde{u}-u^{obs})^2 \; d\boldsymbol{x} + \frac{1}{2}\int_\Omega k\nabla(u+\varepsilon\tilde{u})\cdot\nabla(u+\varepsilon\tilde{u}) \; d\boldsymbol{x} \right] \\
&= 2\int_\Omega (u-u^{obs})\tilde{u} \; d\boldsymbol{x} + \int_\Omega k\nabla\tilde{u}\cdot\nabla u \; d\boldsymbol{x} = 0 \;\forall\; \tilde{u}.
\end{aligned}$$

This is the weak form. To get the strong form of the first-order necessary condition for optimality, we integrate by parts. Note that

$$\int_\Omega k\nabla u\cdot\nabla\tilde{u} \; d\boldsymbol{x} = -\int_\Omega \nabla\cdot(k\nabla u)\hat{u} \; d\boldsymbol{x} + \int_{\partial\Omega} (k\nabla u\cdot\boldsymbol{n})\tilde{u} \; ds.$$

Since $\nabla u\cdot\boldsymbol{n} = 0$ on the boundary,

$$\int_\Omega \Big[2(u-u^{obs}) - \nabla\cdot(k\nabla u)\Big]\tilde{u} \; d\boldsymbol{x} = 0.$$

Thus the strong form of the first-order necessary condition for optimality for $\mathcal{F}_{TN}$ is

$$2(u-u^{obs}) = \nabla\cdot(k\nabla u).$$

$$\begin{aligned}
\mathcal{F}_{TV}^{\varepsilon\prime}(u)(\tilde{u}) &= \frac{d}{da}\bigg|_{a=0} \mathcal{F}_{TV}^{\varepsilon}(u+a\tilde{u}) \\
&= \frac{d}{da}\bigg|_{a=0} \left[\int_\Omega (u+a\tilde{u}-u^{obs})^2 \;\; d\boldsymbol{x} + \int_\Omega k\big(\nabla(u+a\tilde{u})\cdot\nabla(u+a\tilde{u})+\varepsilon\big)^{1/2} \;\; d\boldsymbol{x} \right] \\
&= \int_\Omega 2(u-u^{obs})\tilde{u} \;\; d\boldsymbol{x} + \int_\Omega k(\nabla u\cdot\nabla u+\varepsilon)^{-1/2}\nabla u\cdot\nabla\tilde{u} \;\; d\boldsymbol{x} = 0 \;\forall\; \tilde{u}.
\end{aligned}$$

This is the weak form. To get the strong form, we integrate by parts, resulting in

$$\begin{aligned}
\mathcal{F}_{TV}^{\varepsilon\prime}(u)(\tilde{u}) &= \int_\Omega 2(u-u^{obs})\tilde{u} \;\; d\boldsymbol{x} - \int_\Omega \nabla\cdot\Big(k(\nabla u\cdot\nabla u+\varepsilon)^{-1/2}\nabla u\Big)\tilde{u} \;\; d\boldsymbol{x} \\
&\quad + \int_{\partial\Omega} k(\nabla u\cdot\nabla u+\varepsilon)^{-1/2}(\nabla u\cdot\boldsymbol{n})\tilde{u} \;\; d\boldsymbol{x} \\
&= \int_\Omega 2(u-u^{obs})\tilde{u} \;\; d\boldsymbol{x} - \int_\Omega \nabla\cdot\Big(k(\nabla u\cdot\nabla u+\varepsilon)^{-1/2}\nabla u\Big)\tilde{u} \;\; d\boldsymbol{x} = 0
\end{aligned}$$

since $\nabla u\cdot\boldsymbol{n} = 0$. Thus, the strong form of the first-order necessary condition for optimality for $\mathcal{F}_{TV}^{\varepsilon}$ is

$$2(u-u^{obs}) = \nabla\cdot\Big(k(\nabla u\cdot\nabla u+\varepsilon)^{-1/2}\nabla u\Big).$$

(b) Recall that for an objective functional $\mathcal{F}$, the Newton direction $\hat{u}$ is found by solving an equation of the following form

$$\mathcal{F}''(u)(\tilde{u},\hat{u}) = -\mathcal{F}'(u)(\tilde{u}), \quad \forall\tilde{u}.$$

(Here we have stated the equation for the Newton step in a generic weak form, and use $\tilde{u}$ to denote test functions.) For both $\mathcal{F}_{TN}$ and $\mathcal{F}_{TV}^{\varepsilon}$, derive the equation for the infinite-dimensional Newton step in both weak and strong forms. The strong form of the second variation of $\mathcal{F}_{TV}^{\varepsilon}$ will give an anisotropic diffusion operator of the form $-\text{div}(\boldsymbol{A}(u)\nabla\tilde{u})$, where $\boldsymbol{A}(u)$ is an anisotropic tensor that plays the role of the diffusivity coefficient[1]. (In contrast, you can think of the second variation of $\mathcal{F}_{TN}$ giving an *isotropic* diffusion operator, i.e. with $\boldsymbol{A} = \alpha\boldsymbol{I}$ for some α.)

The second variation of $\mathcal{F}_{TN}$ is

$$\begin{aligned}
\mathcal{F}_{TN}''(u)(\tilde{u},\hat{u}) &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \mathcal{F}_{TN}'(u+\varepsilon\hat{u})(\tilde{u}) \\
&= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \left[2\int_\Omega (u+\varepsilon\hat{u}-u^{obs})\tilde{u} \;\; d\boldsymbol{x} + \int_\Omega k\nabla\tilde{u}\cdot\nabla u \;\; d\boldsymbol{x} + \varepsilon\int_\Omega k\nabla\tilde{u}\cdot\nabla\hat{u} \;\; d\boldsymbol{x}\right] \\
&= 2\int_\Omega \tilde{u}\hat{u} \;\; d\boldsymbol{x} + \int_\Omega k\nabla\tilde{u}\cdot\nabla\hat{u} \;\; d\boldsymbol{x}.
\end{aligned}$$

Thus the weak form of the infinite dimensional Newton step for $\mathcal{F}_{TN}$ is

$$2\int_\Omega \tilde{u}\hat{u} \;\; d\boldsymbol{x} + \int_\Omega k\nabla\tilde{u}\cdot\nabla\hat{u} \;\; d\boldsymbol{x} = -2\int_\Omega (u-u^{obs})\tilde{u}\; d\boldsymbol{x} - \int_\Omega k\nabla\tilde{u}\cdot\nabla u\; d\boldsymbol{x} \;\forall\; \tilde{u}.$$

To get the strong form of the second variation of $\mathcal{F}_{TN}$, note that

$$-\int_\Omega k\nabla\cdot(\nabla\hat{u})\tilde{u}\; d\boldsymbol{x} = -\int_{\partial\Omega} k\tilde{u}(\nabla\hat{u}\cdot\boldsymbol{n})\; ds + \int_\Omega k\nabla\tilde{u}\cdot\nabla\hat{u}\; d\boldsymbol{x}.$$

Since $\nabla u\cdot\boldsymbol{n} = 0$, $\nabla(u+\varepsilon\hat{u})\cdot\boldsymbol{n} = 0$, and $\varepsilon > 0$, $\nabla\hat{u}\cdot\boldsymbol{n} = 0$. We have already obtained the strong form of the right hand side in part (a). Thus

$$\begin{aligned}
&\int_\Omega \big[2\hat{u} - \nabla\cdot(k\nabla\hat{u})\big]\tilde{u}\; d\boldsymbol{x} = \int_\Omega \big[-2(u-u^{obs}) + \nabla\cdot(k\nabla u)\big]\tilde{u}\; d\boldsymbol{x} \\
&\Rightarrow 2\hat{u} - \nabla\cdot(k\nabla\hat{u}) = -2(u-u^{obs}) + \nabla\cdot(k\nabla u) \\
&\Rightarrow \hat{u} + u - u^{obs} = \nabla\cdot(k\nabla\hat{u}).
\end{aligned}$$

That is, the strong form of the infinite dimensional Newton step for $\mathcal{F}_{TN}$ is

$$\hat{u} + u - u^{obs} = \nabla\cdot(k\nabla\hat{u}).$$

[1]Hint: For vectors $a, b, c \in \mathbb{R}^n$ holds $(a\cdot b)c = (ca^T)b$, where $a\cdot b \in \mathbb{R}$ is the inner product and $ca^T \in \mathbb{R}^{n\times n}$ is a matrix of rank one.

The second variation of $\mathcal{F}^{\varepsilon}_{TV}$ is

$$
\begin{aligned}
\mathcal{F}^{\varepsilon''}_{TV}(u)(\tilde{u},\hat{u}) &= \frac{d}{da}\bigg|_{a=0} \mathcal{F}^{\varepsilon'}_{TV}(u+a\hat{u})(\tilde{u}) \\
&= \frac{d}{da}\bigg|_{a=0} \Bigg[\int_\Omega 2(u+a\hat{u}-u^{obs})\tilde{u}\; d\boldsymbol{x} \\
&\quad + \int_\Omega k\big(\nabla(u+a\hat{u})\cdot\nabla(u+a\hat{u})+\varepsilon\big)^{-1/2}\nabla(u+a\hat{u})\cdot\nabla\tilde{u}\; d\boldsymbol{x}\Bigg] \\
&= \Bigg[2\int_\Omega \hat{u}\tilde{u}\; d\boldsymbol{x} + \int_\Omega k\big(\nabla(u+a\hat{u})\cdot\nabla(u+a\hat{u})+\varepsilon\big)^{-1/2}\nabla\hat{u}\cdot\nabla\tilde{u}\; d\boldsymbol{x} \\
&+ \int_\Omega k\big(2\nabla u\cdot\nabla\hat{u}+2a\nabla\hat{u}\cdot\nabla\hat{u}\big)(-1/2)\big(\nabla(u+a\hat{u})\cdot\nabla(u+a\hat{u})\big)^{-3/2}(\nabla u\cdot\nabla\tilde{u}+a\nabla\hat{u}\cdot\nabla\tilde{u})\; d\boldsymbol{x}\Bigg]_{a=0} \\
&= 2\int_\Omega \hat{u}\tilde{u}\; d\boldsymbol{x} + k\int_\Omega \big[(\nabla u\cdot\nabla u+\varepsilon)^{-1/2}\nabla\hat{u} - (\nabla u\cdot\nabla\hat{u})(\nabla u\cdot\nabla u)^{-3/2}\nabla u\big]\cdot\nabla\tilde{u}\; d\boldsymbol{x}.
\end{aligned}
$$

Thus the weak form of the infinite dimensional Newton step for $\mathcal{F}^{\varepsilon}_{TV}$ is

$$
\begin{aligned}
2\int_\Omega \hat{u}\tilde{u}\; d\boldsymbol{x} &+ k\int_\Omega \big[(\nabla u\cdot\nabla u+\varepsilon)^{-1/2}\nabla\hat{u} - (\nabla u\cdot\nabla\hat{u})(\nabla u\cdot\nabla u)^{-3/2}\nabla u\big]\cdot\nabla\tilde{u}\; d\boldsymbol{x} \\
&= -2\int_\Omega (u-u^{obs})\tilde{u}\;\; d\boldsymbol{x} - k\int_\Omega (\nabla u\cdot\nabla u+\varepsilon)^{-1/2}\nabla u\cdot\nabla\tilde{u}\;\; d\boldsymbol{x}.
\end{aligned}
$$

Integrating by parts, and noting that $\nabla u\cdot\boldsymbol{n} = \nabla\hat{u}\cdot\boldsymbol{n} = 0$, we obtain the strong form of the infinite dimensional Newton step for $\mathcal{F}^{\varepsilon}_{TV}$:

$$
\begin{aligned}
2\hat{u} \;-\; k\nabla\cdot\big[(\nabla u\cdot\nabla u+\varepsilon)^{-1/2}\nabla\hat{u} - (\nabla u\cdot\nabla\hat{u})(\nabla u\cdot\nabla u)^{-3/2}\nabla u\big] \\
= -2(u-u^{obs}) + k\nabla\cdot\big[(\nabla u\cdot\nabla u+\varepsilon)^{-1/2}\nabla u.\big]
\end{aligned}
$$

(c) Derive expressions for the two eigenvalues and corresponding eigenvectors of $\boldsymbol{A}$. Based on these expressions, give an explanation of why $\mathcal{F}^{\varepsilon}_{TV}$ is effective at preserving sharp edges in the image, while $\mathcal{F}_{TN}$ is not. Consider a single Newton step for this argument.

$$
\mathbf{A}(u) = k\big[(\nabla u\cdot\nabla u+\varepsilon)^{-1/2}I - (\nabla u\cdot\nabla u)^{-3/2}(\nabla u\nabla u^T)\big]
$$

I found the eigenvalues and eigenvectors of $\mathbf{A}$ using Mathematica. The eigenvalues of $\mathbf{A}$ are

$$
\Big\{\Big(2\,k\,ux^2\sqrt{ux^2+uy^2} + 2\,k\,uy^2\sqrt{ux^2+uy^2} - ux^2\sqrt{eps+ux^2+uy^2} - uy^2\sqrt{eps+ux^2+uy^2} - \sqrt{eps\,ux^4+ux^6+2\,eps\,ux^2\,uy^2+3\,ux^4\,uy^2+eps\,uy^4+3\,ux^2\,uy^4+uy^6}\Big)\Big/ \Big(2\,(ux^2+uy^2)^{3/2}\sqrt{eps+ux^2+uy^2}\Big),
$$

$$
\Big(2\,k\,ux^2\sqrt{ux^2+uy^2} + 2\,k\,uy^2\sqrt{ux^2+uy^2} - ux^2\sqrt{eps+ux^2+uy^2} - uy^2\sqrt{eps+ux^2+uy^2} + \sqrt{eps\,ux^4+ux^6+2\,eps\,ux^2\,uy^2+3\,ux^4\,uy^2+eps\,uy^4+3\,ux^2\,uy^4+uy^6}\Big)\Big/ \Big(2\,(ux^2+uy^2)^{3/2}\sqrt{eps+ux^2+uy^2}\Big)\Big\}
$$

and the eigenvectors of $\mathbf{A}$ are

$$\left\{\left\{-\frac{-\mathrm{ux}^2\sqrt{\mathrm{eps}+\mathrm{ux}^2+\mathrm{uy}^2}+\mathrm{uy}^2\sqrt{\mathrm{eps}+\mathrm{ux}^2+\mathrm{uy}^2}-\sqrt{\left(\mathrm{ux}^2+\mathrm{uy}^2\right)^2\left(\mathrm{eps}+\mathrm{ux}^2+\mathrm{uy}^2\right)}}{2\,\mathrm{ux}\,\mathrm{uy}\sqrt{\mathrm{eps}+\mathrm{ux}^2+\mathrm{uy}^2}},1\right\},\right.$$

$$\left.\left\{-\frac{-\mathrm{ux}^2\sqrt{\mathrm{eps}+\mathrm{ux}^2+\mathrm{uy}^2}+\mathrm{uy}^2\sqrt{\mathrm{eps}+\mathrm{ux}^2+\mathrm{uy}^2}+\sqrt{\left(\mathrm{ux}^2+\mathrm{uy}^2\right)^2\left(\mathrm{eps}+\mathrm{ux}^2+\mathrm{uy}^2\right)}}{2\,\mathrm{ux}\,\mathrm{uy}\sqrt{\mathrm{eps}+\mathrm{ux}^2+\mathrm{uy}^2}},1\right\}\right\}$$

Note that the eigenvalues of $\mathbf{A}$ are not equal. In $\mathcal{F}_{TN}$, in the matrix αI, the eigenvalues are of course equal. The anisotropy afforded by the unequal eigenvalues of $\mathbf{A}$ preserves sharp edges.

3. **Boundary control of elliptic PDE.** Consider the objective functional

$$\mathcal{J}(m) = f(u(m), m) := \frac{1}{2}\int_\Omega (u - u_d)^2\, d\boldsymbol{x} + \frac{\gamma}{2}\int_{\partial\Omega} m^2\, ds,$$

where

$$-\Delta u = 0, \quad \text{in } \Omega,$$
$$\nabla u \cdot \boldsymbol{n} + \alpha u = \alpha m, \quad \text{in } \partial\Omega.$$

Here Ω is bounded domain in $\mathbb{R}^2$, u_d is a given function, and α is a positive constant. The goal is to derive the gradient of $\mathcal{J}$ with respect to the boundary control m using a Lagrangian approach.

(a) Write down the Lagrangian functional for this problem. Be precise with function spaces.

First, we derive the weak form of the state equation. Let $v \in H^1(\Omega)$ be a test function. Then

$$-\int_\Omega \nabla \cdot (\nabla u) v\; d\boldsymbol{x} = 0.$$

Using multivariate integration by parts,

$$0 = -\int_\Omega \nabla \cdot (\nabla u) v\; d\boldsymbol{x} = -\int_{\partial\Omega} (\nabla u \cdot \boldsymbol{n}) v\; ds + \int_\Omega \nabla v \cdot \nabla u\; d\boldsymbol{x}.$$

Thus, by the boundary condition, the weak form of the state equation is: Find $u \in H^1(\Omega)$ such that

$$0 = \alpha \int_{\partial\Omega} (u - m) v\; ds + \int_\Omega \nabla u \cdot \nabla v\; d\boldsymbol{x}$$

for all $v \in H^1(\Omega)$.

Thus the Lagrangian functional is

$$\mathcal{L}(u, m, p) = \frac{1}{2}\int_\Omega (u - u_d)^2\; d\boldsymbol{x} + \frac{\gamma}{2}\int_{\partial\Omega} m^2\; ds + \alpha\int_{\partial\Omega} (u - m) p\; ds + \int_\Omega \nabla u \cdot \nabla p\; d\boldsymbol{x}.$$

Note that $u, p \in H^1(\Omega)$.

(b) Derive the adjoint equation. State both the weak and strong form the adjoint equation.

Given m, as well as u, the solution to the state equation, find $p \in H^1(\Omega)$ such that

$$\begin{aligned}
\mathcal{L}_u(u, m, p)(\tilde{u}) &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \mathcal{L}(u + \varepsilon\tilde{u}, m, p) \\
&= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \left[\frac{1}{2}\int_\Omega (u + \varepsilon\tilde{u} - u_d)^2\; d\boldsymbol{x} + \alpha\int_{\partial\Omega} (u + \varepsilon\tilde{u} - m) p\; ds + \int_\Omega \nabla(u + \varepsilon\tilde{u}) \cdot \nabla p\; d\boldsymbol{x}\right] \\
&= \int_\Omega (u - u_d)\tilde{u}\; d\boldsymbol{x} + \alpha\int_{\partial\Omega} \tilde{u} p\; ds + \int_\Omega \nabla\tilde{u} \cdot \nabla p\; d\boldsymbol{x} = 0
\end{aligned}$$

for all $\tilde{u} \in H^1(\Omega)$.

This is the weak form of the adjoint equation. To obtain the strong form, we note that by intergration by parts, the adjoint equation becomes

$$\begin{aligned}
0 &= \int_\Omega (u - u_d)\tilde{u}\; d\boldsymbol{x} + \alpha\int_{\partial\Omega} \tilde{u} p\; ds + \int_{\partial\Omega} \tilde{u}(\nabla p \cdot \boldsymbol{n})\; ds - \int_\Omega \tilde{u} \nabla \cdot (\nabla p)\; d\boldsymbol{x} \\
&= \int_\Omega \big[u - u_d - \nabla \cdot \nabla p\big]\tilde{u}\; d\boldsymbol{x} + \int_{\partial\Omega} \big[\alpha p + \nabla p \cdot \boldsymbol{n}\big]\tilde{u}\; ds \;\forall\; \tilde{u}.
\end{aligned}$$

Thus, the strong form of the adjoint equation is

$$\begin{aligned} u - u_d &= \Delta p, \quad \text{in } \Omega, \\ \alpha p + \nabla p \cdot \boldsymbol{n} &= 0, \quad \text{on } \partial\Omega. \end{aligned}$$

(c) Derive the expression for the gradient.
Given m, u, p,

$$\begin{aligned} g(m)(\tilde{m}) = \mathcal{L}_m(u, m, p)(\tilde{m}) &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \mathcal{L}(u, m + \varepsilon\tilde{m}, p) \\ &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \left[\frac{\gamma}{2} \int_{\partial\Omega} (m + \varepsilon\tilde{m})^2 \, ds + \alpha \int_{\partial\Omega} (u - m - \varepsilon\tilde{m}) p \, ds \right] \\ &= \gamma \int_{\partial\Omega} m\tilde{m} \, ds - \alpha \int_{\partial\Omega} \tilde{m} p \, ds. \end{aligned}$$

4. **Frequency-domain inverse wave problem.** Let $\Omega \subset \mathbb{R}^n$ ($n \in \{2,3\}$) be a bounded domain. A Fourier transformation of the time-dependent acoustic wave equation with locally varying wave speed results in a stationary equation for each frequency k_i (we assume $N_f \geq 1$ frequencies). Additionally, one often has data available from several experiments, each of which uses a different source f_j (we assume $N_s \geq 1$ sources). The inverse problem of estimating the locally varying slowness (i.e., the inverse of the wave speed) in frequency domain is:

$$\min_m \mathcal{J}(m) := \frac{1}{2}\sum_{i=1}^{N_f}\sum_{j=1}^{N_s}\int_\Omega (u_{ij} - u_{ij}^{obs})^2 \, d\boldsymbol{x} + \frac{\alpha}{2}\int_\Omega \nabla m \cdot \nabla m \, d\boldsymbol{x}$$

where u_{ij} is the solution of

$$-\Delta u_{ij}(\boldsymbol{x}) - k_i^2 m(\boldsymbol{x}) u_{ij}(\boldsymbol{x}) = f_j(\boldsymbol{x}) \text{ in } \Omega, \quad i = 1, \dots, N_f, \quad j = 1, \dots, N_s,$$
$$u_{ij} = 0 \text{ on } \partial\Omega.$$

Above, u_{ij}^{obs} denotes given measurements, u_{ij} is the wave field for frequency k_i and source f_j, and $\alpha > 0$ is the regularization parameter.

(a) Compute the gradient of $\mathcal{J}$ using the Lagrangian method for a single source and frequency, i.e., for $N_f = N_s = 1$.

To compute the weak form of the PDE to which u is a solution, we first multiply by a test function $v \in V = \{H^1(\Omega) \mid v = 0 \text{ on } \partial\Omega\}$ and integrate over Ω.

$$\int_\Omega \nabla \cdot (\nabla u) v \; d\boldsymbol{x} + \int_\Omega k^2 muv \; d\boldsymbol{x} + \int_\Omega fv \; d\boldsymbol{x} = 0$$

Then we use (multivariate) integration by parts, and the fact that $v = 0$ on $\partial\Omega$ to obtain the weak form

$$\int_\Omega \nabla u \cdot \nabla v \; d\boldsymbol{x} - \int_\Omega k^2 muv \; d\boldsymbol{x} - \int_\Omega fv \; d\boldsymbol{x}.$$

Now, the Lagrangian is

$$\mathcal{L}(u,m,p) = \frac{1}{2}\int_\Omega (u - u^{\text{obs}})^2 \, d\boldsymbol{x} + \frac{\alpha}{2}\int_\Omega \nabla m \cdot \nabla m \, d\boldsymbol{x} + \int_\Omega \nabla u \cdot \nabla p \, d\boldsymbol{x} - \int_\Omega k^2 mup \, d\boldsymbol{x} - \int_\Omega fp \, d\boldsymbol{x}.$$

The state equation is

$$\begin{aligned}
\mathcal{L}_p(u,m,p)(\tilde{p}) &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathcal{L}(u,m,p+\varepsilon\tilde{p}) \\
&= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \left(\frac{1}{2}\int_\Omega (u - u^{\text{obs}})^2 \; d\boldsymbol{x} + \frac{\alpha}{2}\int_\Omega \nabla m \cdot \nabla m \; d\boldsymbol{x} + \int_\Omega \nabla u \cdot \nabla p \; d\boldsymbol{x} + \right. \\
&\left. \varepsilon \int_\Omega \nabla u \cdot \nabla \tilde{p} \; d\boldsymbol{x} - \int_\Omega k^2 mu(p + \varepsilon\tilde{p}) \; d\boldsymbol{x} - \int_\Omega f(p+\varepsilon\tilde{p}) \; d\boldsymbol{x}\right) \\
&= \int_\Omega \nabla u \cdot \nabla \tilde{p} \; d\boldsymbol{x} - \int_\Omega k^2 mu\tilde{p} \; d\boldsymbol{x} - \int_\Omega f\tilde{p} \; d\boldsymbol{x} = 0.
\end{aligned}$$

To compute the gradient $g(m)(\tilde{m})$ for a given m, first find the $u \in V$ which solves the state equation for all $\tilde{p} \in V$.

Next, we compute the adjoint equation, which is

$$
\begin{aligned}
\mathcal{L}_u(u,m,p)(\tilde{u}) &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \mathcal{L}(u+\varepsilon\tilde{u},m,p) \\
&= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \bigg(\frac{1}{2}\int_\Omega (u+\varepsilon\tilde{u}-u^{\mathsf{obs}})^2 \; d\boldsymbol{x} + \frac{\alpha}{2}\int_\Omega \nabla m\cdot\nabla m \; d\boldsymbol{x} + \int_\Omega \nabla u\cdot\nabla p \; d\boldsymbol{x} \\
&+ \varepsilon\int_\Omega \nabla\tilde{u}\cdot\nabla p \; d\boldsymbol{x} - \int_\Omega k^2 m(u+\varepsilon\tilde{u})p \; d\boldsymbol{x} - \int_\Omega fp \; d\boldsymbol{x}\bigg) \\
&= \int_\Omega (u-u^{\mathsf{obs}})\tilde{u} \; d\boldsymbol{x} + \int_\Omega \nabla\tilde{u}\cdot\nabla p \; d\boldsymbol{x} - \int_\Omega k^2 m\tilde{u}p \; d\boldsymbol{x}.
\end{aligned}
$$

The next step in computing the gradient, given m and u, is to find $p\in V$ for which the adjoint equation holds for all $\tilde{u}\in V$.
Finally, given m, u, and p, we compute the gradient

$$
\begin{aligned}
g(m)(\tilde{m}) = \mathcal{L}_m(u,m,p)(\tilde{m}) &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \mathcal{L}(u,m+\varepsilon\tilde{m},p) \\
&= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \bigg(\frac{1}{2}\int_\Omega (u-u^{\mathsf{obs}})^2 \; d\boldsymbol{x} + \frac{\alpha}{2}\int_\Omega (\nabla m+\varepsilon\nabla\tilde{m})\cdot(\nabla m+\varepsilon\nabla\tilde{m}) \; d\boldsymbol{x} \\
&+ \int_\Omega \nabla u\cdot\nabla p \; d\boldsymbol{x} - \int_\Omega k^2(m+\varepsilon\tilde{m})up \; d\boldsymbol{x} - \int_\Omega fp \; d\boldsymbol{x}\bigg) \\
&= \alpha\int_\Omega \nabla m\cdot\nabla\tilde{m} \; d\boldsymbol{x} - \int_\Omega k^2\tilde{m}up \; d\boldsymbol{x}.
\end{aligned}
$$

(b) Compute the gradient for an arbitrary number of sources and frequencies[2]. How many state and adjoint equations have to be solved for a single gradient computation? In your solution be precise about the function spaces you use for state and adjoint variables, and the parameter m.
Let $V=\{H^1(\Omega)\mid v=0 \text{ on } \partial\Omega\}$.
First, we consider the Lagrangian with an arbitrary number of sources and frequencies:

$$
\begin{aligned}
\mathcal{L}(u,m,p) = &\frac{1}{2}\sum_{i=1}^{N_f}\sum_{j=1}^{N_s}\int_\Omega (u_{ij}-u_{ij}^{obs})^2 \, d\boldsymbol{x} + \frac{\alpha}{2}\int_\Omega \nabla m\cdot\nabla m \, d\boldsymbol{x} \\
&\sum_{i=1}^{N_f}\sum_{j=1}^{N_s}\bigg(\int_\Omega \nabla u_{ij}\cdot\nabla p_{ij} \; d\boldsymbol{x} - \int_\Omega k_i^2 m u_{ij}p_{ij} \; d\boldsymbol{x} - \int_\Omega f_j p_{ij} \; d\boldsymbol{x}\bigg).
\end{aligned}
$$

[2] For every state equation with solution u_{ij}, use the adjoint variable p_{ij}.

Now, given m, for each $i = 1, \dots, N_f$, $j = 1, \dots, N_s$, we solve a state equation:

$$\begin{aligned}
\mathcal{L}_{p_{ij}}(u, m, p)(\tilde{p}) &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \Bigg(\frac{1}{2} \sum_{I=1}^{N_f} \sum_{J=1}^{N_s} \int_\Omega (u_{IJ} - u_{IJ}^{obs})^2 \, d\boldsymbol{x} + \frac{\alpha}{2} \int_\Omega \nabla m \cdot \nabla m \, d\boldsymbol{x} \\
&+ \sum_{\substack{I=1 \\ I \neq i}}^{N_f} \sum_{\substack{J=1 \\ J \neq j}}^{N_s} \left(\int_\Omega \nabla u_{IJ} \cdot \nabla p_{IJ} \; d\boldsymbol{x} - \int_\Omega k^2 m u_{IJ} p_{IJ} \; d\boldsymbol{x} - \int_\Omega f p_{IJ} \; d\boldsymbol{x} \right) \\
&+ \int_\Omega \nabla u_{ij} \cdot \nabla (p_{ij} + \varepsilon \tilde{p}) \; d\boldsymbol{x} - \int_\Omega k_i^2 m u_{ij} (p_{ij} + \varepsilon \tilde{p}) \; d\boldsymbol{x} - \int_\Omega f_j (p_{ij} + \varepsilon \tilde{p}) \; d\boldsymbol{x} \Bigg) \\
&= \int_\Omega \nabla u_{ij} \cdot \nabla \tilde{p} \; d\boldsymbol{x} - \int_\Omega k_i^2 m u_{ij} \tilde{p} \; d\boldsymbol{x} - \int_\Omega f_j \tilde{p} \; d\boldsymbol{x} = 0.
\end{aligned}$$

That is, for each $i = 1, \dots, N_f$, $j = 1, \dots, N_s$, find $u_{ij} \in V$ for which the $\{i, j\}$th state equation holds for all $\tilde{p} \in V$.
Next, for each $i = 1, \dots, N_f$, $j = 1, \dots, N_s$, solve an adjoint equation, which is

$$\begin{aligned}
\mathcal{L}_{u_{ij}}(u, m, p) = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} & \Bigg(\frac{1}{2} \sum_{\substack{I=1 \\ I \neq i}}^{N_f} \sum_{\substack{J=1 \\ J \neq j}}^{N_s} \int_\Omega (u_{IJ} - u_{IJ}^{obs})^2 \, d\boldsymbol{x} + \frac{1}{2} \int_\Omega (u_{ij} + \varepsilon \tilde{u} - u_{ij}^{\mathsf{obs}})^2 \\
&+ \frac{\alpha}{2} \int_\Omega \nabla m \cdot \nabla m \, d\boldsymbol{x} + \sum_{\substack{I=1 \\ I \neq i}}^{N_f} \sum_{\substack{J=1 \\ J \neq j}}^{N_s} \int_\Omega \nabla u_{IJ} \cdot \nabla p_{IJ} \; d\boldsymbol{x} + \int_\Omega \nabla (u_{ij} + \varepsilon \tilde{u}) \cdot \nabla p_{ij} \; d\boldsymbol{x} \\
&- \sum_{\substack{I=1 \\ I \neq i}}^{N_f} \sum_{\substack{J=1 \\ J \neq j}}^{N_s} \int_\Omega k_I^2 m u_{IJ} p_{IJ} \; d\boldsymbol{x} - \int_\Omega k_i^2 m (u_{ij} + \varepsilon \tilde{u}) p_{ij} \; d\boldsymbol{x} - \sum_{I=1}^{N_f} \sum_{J=1}^{N_s} \int_\Omega f_J p_{IJ} dx \Bigg) \\
&= \int_\Omega (u_{ij} - u_{ij}^{\mathsf{obs}}) \tilde{u} \; d\boldsymbol{x} + \int_\Omega \nabla \tilde{u} \cdot \nabla p_{ij} \; d\boldsymbol{x} - \int_\Omega k_i^2 m \tilde{u} p_{ij} \; d\boldsymbol{x} = 0.
\end{aligned}$$

That is, for each $i = 1, \dots, N_f$, $j = 1, \dots, N_s$, given m and u_{ij}, find $p_{ij} \in V$ for which the $\{i, j\}$th state equation holds for all $\tilde{u} \in V$.
Finally, given m, as well as p_{ij} and u_{ij} for all $i = 1, \dots, N_f$, $j = 1, \dots, N_s$, we compute the gradient,

$$\begin{aligned}
g(m)(\tilde{m}) = \mathcal{L}_m(u, m, p)(\tilde{m}) &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \mathcal{L}(u, m + \varepsilon \tilde{m}, p) \\
&= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \Bigg(\frac{1}{2} \sum_{i=1}^{N_f} \sum_{j=1}^{N_s} \int_\Omega (u_{ij} - u_{ij}^{\mathsf{obs}})^2 \; d\boldsymbol{x} + \frac{\alpha}{2} \int_\Omega \nabla (m + \varepsilon \tilde{m}) \cdot \nabla (m + \varepsilon \tilde{m}) \; d\boldsymbol{x} \\
&+ \sum_{i=1}^{N_f} \sum_{j=1}^{N_s} \left(\int_\Omega \nabla u_{ij} \cdot \nabla p_{ij} \; d\boldsymbol{x} - \int_\Omega k^2 (m + \varepsilon \tilde{m}) u_{ij} p_{ij} \; d\boldsymbol{x} - \int_\Omega f p_{ij} \; d\boldsymbol{x} \right) \Bigg) \\
&= \alpha \int_\Omega \nabla m \cdot \nabla \tilde{m} \; d\boldsymbol{x} - \sum_{i=1}^{N_f} \sum_{j=1}^{N_s} \int_\Omega k^2 \tilde{m} u_{ij} p_{ij} \; d\boldsymbol{x}.
\end{aligned}$$

$N_f N_s$ state equations and $N_f N_s$ adjoint equations have to be solved for a single gradient computation.