Computational and Variational Methods for Inverse Problems - Homework 3 Solution

Mohammad Afzal Shadab (ms82697)

mashadab@utexas.edu

Due Monday, November 8, 2021

---

1 Image de-noising problem

Given a noisy image $d(x, y)$ defined on a square domain Ω, we want to minimize the data misfit

$$\mathcal{F}_{LS}(u) := \frac{1}{2}\int_\Omega (u-d)^2 \, d\boldsymbol{x} \tag{1.1}$$

while also removing noise, assumed to comprise the highly oscillatory rough components.

We can consider an additional Tikhonov (TN) regularization term of the form

$$\mathcal{R}_{TN}(u) := \frac{\alpha}{2}\int_\Omega \nabla u \cdot \nabla u \, d\boldsymbol{x} \tag{1.2}$$

Unfortunately, the TN regularization tends to blur sharp edges in the image. In these cases, we instead prefer the total-variation (TV) regularization

$$\mathcal{R}_{TV} := \alpha \int_\Omega (\nabla u \cdot \nabla u)^{1/2} \, d\boldsymbol{x} \tag{1.3}$$

However, $\mathcal{R}_{TV}$ is not differentiable at $\nabla u = \mathbf{0}$. Therefore, the functional is modified to include a small parameter $\delta > 0$

$$\mathcal{R}^\delta_{TV} := \alpha \int_\Omega (\nabla u \cdot \nabla u + \delta)^{1/2} \tag{1.4}$$

We wish to study the de-noising functionals $\mathcal{F}_{TN}$ and $\mathcal{F}^\delta_{TV}$ where

$$\min_{u\in\mathcal{U}} \mathcal{F}_{TN}(u) := \mathcal{F}_{LS}(u) + \mathcal{R}_{TN}(u) \tag{1.5}$$

and

$$\min_{u\in\mathcal{U}} \mathcal{F}^\delta_{TV}(u) := \mathcal{F}_{LS}(u) + \mathcal{R}^\delta_{TV}(u) \tag{1.6}$$

where $\mathcal{U}$ is the space of admissible functions for each problem. Boundary conditions are

$$\nabla u \cdot \boldsymbol{n} = 0, \quad \text{on the boundary of the square, } \partial\Omega \tag{1.7}$$

where $\boldsymbol{n}$ is the outward normal with assumption that the image intensity does not change normal to the boundary of the image.

(a) First-order necessary condition for optimality

First, let's derive the first variations of the functionals in weak form. Consider $u \in H^1(\Omega)$ and a corresponding variation $\hat{u} \in H^1(\Omega)$.
TN Regularization: Then the first variations corresponding to $\mathcal{F}_{LS}$ and $\mathcal{R}_{TN}$ are given below

$$\delta_u \mathcal{F}_{LS}(u, \hat{u}) = \int_\Omega (u - d)\hat{u}\, d\boldsymbol{x} \tag{1.8}$$

$$\delta_u \mathcal{R}_{TN}(u, \hat{u}) = \alpha \int_\Omega \nabla u \cdot \nabla \hat{u}\, d\boldsymbol{x} \tag{1.9}$$

The first variation of $\mathcal{R}^\delta_{TV}$ is again computed using calculus of variations where $u \to (u + \delta\hat{u})$

$$\begin{aligned}
\delta_u \mathcal{R}^\delta_{TV}(u, \hat{u}) &= \frac{\mathrm{d}}{\mathrm{d}\delta}\left[\alpha \int_\Omega (\nabla(u + \delta\hat{u}) \cdot \nabla(u + \delta\hat{u}) + \delta)^{1/2}\, d\boldsymbol{x}\right]_{\delta=0} \\
&= \alpha \int_\Omega \frac{1}{2}(\nabla u \cdot \nabla u + \delta)^{-1/2}\, 2(\nabla u \cdot \nabla \hat{u})\, d\boldsymbol{x} \\
&= \alpha \int_\Omega \frac{\nabla u \cdot \nabla \hat{u}}{(\nabla u \cdot \nabla u + \delta)^{1/2}}\, d\boldsymbol{x}
\end{aligned} \tag{1.10}$$

Then substituting (1.8 & 1.9) in (1.5), we have the following optimality condition in the *weak form*

$$\boxed{\delta_u \mathcal{F}_{TN}(u, \hat{u}) = \int_\Omega (u - d)\hat{u}\, d\boldsymbol{x} + \alpha \int_\Omega \nabla u \cdot \nabla \hat{u}\, d\boldsymbol{x} = 0 \quad \forall \hat{u} \in H^1(\Omega)} \tag{1.11}$$

Now, to derive the corresponding strong form for TN regularization, let's integrate (1.11) by parts and applying Gauss divergence theorem to get

$$\delta_u \mathcal{F}_{TN}(u, \hat{u}) = \int_\Omega (u - d)\hat{u}\, d\boldsymbol{x} - \alpha \int_\Omega \Delta u \hat{u}\, d\boldsymbol{x} + \alpha \int_{\partial\Omega} \hat{u} \nabla u \cdot \boldsymbol{n}\, ds = 0 \quad \forall \hat{u} \in H^1(\Omega)$$

where $\Delta \equiv \nabla^2$, which gives the *strong form*

$$\alpha \Delta u \quad = (u - d), \quad \text{in } \Omega \tag{1.12}$$

$$\nabla u \cdot \boldsymbol{n} \quad = 0, \quad \text{on } \partial\Omega \tag{1.13}$$

TV Regularization: Plugging (1.8 & 1.10) in (1.6), we get the optimality condition for

$$\boxed{\delta_u \mathcal{F}^\delta_{TV}(u, \hat{u}) = \int_\Omega (u - d)\hat{u}\, d\boldsymbol{x} + \alpha \int_\Omega \frac{\nabla u \cdot \nabla \hat{u}}{(\nabla u \cdot \nabla u + \delta)^{1/2}}\, d\boldsymbol{x} = 0 \quad \forall \hat{u} \in H^1(\Omega)} \tag{1.14}$$

which is the required *weak form* of the first-order necessary optimal condition. Now, again performing the integration by parts on the second term with $\nabla\hat{u}$, we get

$$\begin{aligned}
\delta_u \mathcal{F}^\delta_{TV}(u, \hat{u}) = \int_\Omega (u - d)\hat{u}\, d\boldsymbol{x} &- \alpha \int_\Omega \nabla \cdot \left(\frac{\nabla u}{(\nabla u \cdot \nabla u + \delta)^{1/2}}\right) \hat{u}\, d\boldsymbol{x} \\
&+ \alpha \int_{\partial\Omega} \hat{u} \frac{\nabla u \cdot \boldsymbol{n}}{(\nabla u \cdot \nabla u + \delta)^{1/2}}\, ds = 0 \quad \forall \hat{u} \in H^1(\Omega)
\end{aligned} \tag{1.15}$$

Finally, it gives the *strong form*

$$\alpha \nabla \cdot \left(\frac{\nabla u}{(\nabla u \cdot \nabla u + \delta)^{1/2}}\right) = u - d \quad \text{in } \Omega \tag{1.16}$$

$$\nabla u \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega \tag{1.17}$$

(b) When $\nabla u = 0$, $\mathcal{R}_{TV}$ is not differentiable but $\mathcal{R}_{TV}^{\delta}$ is.

The first variation of $\mathcal{R}_{TV}$ from (1.3) is

$$\delta_u \mathcal{R}_{TV}(u, \hat{u}) = \alpha \int_{\Omega} \frac{1}{2} (\nabla u \cdot \nabla u)^{-1/2} 2(\nabla u \cdot \hat{u}) \, d\boldsymbol{x} = \alpha \int_{\Omega} \frac{\nabla u \cdot \hat{u}}{(\nabla u \cdot \nabla u)^{1/2}} \, d\boldsymbol{x} \tag{1.18}$$

At $\nabla u = \mathbf{0}$, the denominator of (1.18) goes to ∞. Therefore, $\mathcal{R}_{TV}$ is *not differentiable.*

Using (1.10),

$$\delta_u \mathcal{R}_{TV}^{\delta}(u, \hat{u}) = \alpha \int_{\Omega} \frac{\nabla u \cdot \nabla \hat{u}}{(\nabla u \cdot \nabla u + \delta)^{1/2}} \, d\boldsymbol{x} \tag{1.19}$$

Here, for $\nabla u = \mathbf{0}$, the denominator of (1.19) does not go to ∞ but it is actually zero. Therefore, $\mathcal{R}_{TV}^{\delta}$ is *differentiable.*

(c) Infinite-dimensional Newton step

Let's begin with the second variations of the different components as evaluated earlier by using a perturbation $\tilde{u} \in H^1(\Omega)$. Using calculus of variation on (1.8) and (1.9), we get

$$\delta_u^2 \mathcal{F}_{LS}(u, \hat{u}, \tilde{u}) = \int_{\Omega} \tilde{u}\hat{u} \, d\boldsymbol{x} \tag{1.20}$$

$$\delta_u^2 \mathcal{R}_{TN}(u, \hat{u}, \tilde{u}) = \alpha \int_{\Omega} \nabla \tilde{u} \cdot \nabla \hat{u} \, d\boldsymbol{x} \tag{1.21}$$

Let's find the second variation of $\mathcal{R}_{TV}^{\delta}$ again using the calculus of variation

$$\begin{aligned} \delta_u^2 \mathcal{R}_{TV}^{\delta}(u, \hat{u}, \tilde{u}) &= \int_{\Omega} \frac{\alpha \nabla \tilde{u} \cdot \nabla \hat{u}}{(\nabla u \cdot \nabla u + \delta)^{1/2}} - \frac{\alpha \nabla u \cdot \nabla \hat{u}}{2(\nabla u \cdot \nabla u + \delta)^{3/2}} (2 \nabla u \cdot \nabla \tilde{u}) \, d\boldsymbol{x} \\ &= \int_{\Omega} \frac{\alpha}{(\nabla u \cdot \nabla u + \delta)^{1/2}} \left(\mathcal{I} - \frac{\nabla u \otimes \nabla u}{\nabla u \cdot \nabla u + \delta} \right) \nabla \tilde{u} \cdot \nabla \hat{u} \, d\boldsymbol{x} \end{aligned} \tag{1.22}$$

The Newton step $\tilde{u}$ is given by

$$\delta_u^2 \mathcal{F}(u, \hat{u}, \tilde{u}) = -\delta_u \mathcal{F}(u, \hat{u}, \tilde{u}) \quad \forall \tilde{u} \in H^1(\Omega) \tag{1.23}$$

Now, let's apply it for the different regularizations:

TN Regularization: Plugging (1.20), (1.21), (1.11) and (1.5) in (1.23) gives the *weak form*

$$\int_{\Omega} \tilde{u}\hat{u} \, d\boldsymbol{x} + \alpha \int_{\Omega} \nabla \tilde{u} \cdot \nabla \hat{u} \, d\boldsymbol{x} = -\int_{\Omega} (u - d)\hat{u} \, d\boldsymbol{x} - \alpha \int_{\Omega} \nabla u \cdot \nabla \hat{u} \, d\boldsymbol{x} = 0 \quad \forall \hat{u} \in H^1(\Omega) \quad \forall \hat{u}, \tilde{u} \in H^1(\Omega) \tag{1.24}$$

Now, again using the integration by parts,

$$\int_{\Omega} \tilde{u}\hat{u} \, d\boldsymbol{x} - \alpha \int_{\Omega} \Delta \tilde{u}\hat{u} \, d\boldsymbol{x} + \alpha \int_{\partial\Omega} \hat{u} \nabla \tilde{u} \cdot \boldsymbol{n} \, ds = -\int_{\Omega} (u - d)\hat{u} \, d\boldsymbol{x} + \alpha \int_{\Omega} \Delta u \hat{u} \, d\boldsymbol{x} - \alpha \int_{\partial\Omega} \hat{u} \nabla u \cdot \boldsymbol{n} \, ds \tag{1.25}$$

So, by using the arbitrariness of $\hat{u}$ again, we will get the *strong form*

$$\tilde{u} - \alpha \Delta \tilde{u} = -(u - d) + \alpha \Delta u \quad \text{in } \Omega \tag{1.26}$$

$$\alpha \nabla \tilde{u} \cdot \boldsymbol{n} = -\alpha \nabla u \cdot \boldsymbol{n} \quad \text{on } \partial\Omega \tag{1.27}$$

TV Regularization: Similarly, plugging (1.20), (1.22), (1.48) and (1.6) in (1.23) gives the *weak form*

$$\begin{aligned}&\int_\Omega \tilde{u}\hat{u}\, d\boldsymbol{x} + \int_\Omega \frac{\alpha}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\left(\mathcal{I}-\frac{\nabla u\otimes\nabla u}{\nabla u\cdot\nabla u+\delta}\right)\nabla\tilde{u}\cdot\nabla\hat{u}\, d\boldsymbol{x}\\ &= -\int_\Omega (u-d)\hat{u}\, d\boldsymbol{x} - \alpha\int_\Omega \frac{\nabla u\cdot\nabla\hat{u}}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\, d\boldsymbol{x} = 0 \quad \forall\hat{u}\in H^1(\Omega) \quad \forall\hat{u},\tilde{u}\in H^1(\Omega)\end{aligned} \tag{1.28}$$

Integration by parts leads to gives

$$\begin{aligned}&\int_\Omega \tilde{u}\hat{u}\, d\boldsymbol{x} - \int_\Omega \nabla\cdot\left(\frac{\alpha}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\left(\mathcal{I}-\frac{\nabla u\otimes\nabla u}{\nabla u\cdot\nabla u+\delta}\right)\nabla\tilde{u}\right)\hat{u}\, d\boldsymbol{x}\\ &+\int_{\partial\Omega}\hat{u}\frac{\alpha}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\left(\mathcal{I}-\frac{\nabla u\otimes\nabla u}{\nabla u\cdot\nabla u+\delta}\right)\nabla\tilde{u}\cdot\boldsymbol{n}\, ds = -\int_\Omega (u-d)\hat{u}\\ &+\int_\Omega \frac{\alpha\Delta u}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\hat{u} - \int_{\partial\Omega}\hat{u}\frac{\alpha\nabla u\cdot\boldsymbol{n}}{(\nabla u\cdot\nabla u+\delta)^{-1/2}}\, ds \quad \forall\tilde{u},\hat{u}\in H^1(\Omega)\end{aligned} \tag{1.29}$$

Again the arbitrariness of $\hat{u}$ leads to the *strong form*

$$\begin{aligned}&\tilde{u} - \nabla\cdot\left(\frac{\alpha}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\left(\mathcal{I}-\frac{\nabla u\otimes\nabla u}{\nabla u\cdot\nabla u+\delta}\right)\nabla\tilde{u}\right)\\ &\qquad = -(u-d) + \nabla\cdot\left(\frac{\alpha\nabla u}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\right) \quad \text{in } \Omega\end{aligned} \tag{1.30}$$

$$\frac{\alpha}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\left(\mathcal{I}-\frac{\nabla u\otimes\nabla u}{\nabla u\cdot\nabla u+\delta}\right)\nabla\tilde{u}\cdot\boldsymbol{n} = -\frac{\alpha\nabla u\cdot\boldsymbol{n}}{(\nabla u\cdot\nabla u+\delta)^{1/2}} \quad \text{on } \partial\Omega \tag{1.31}$$

Here, the anisotropic tensor $\boldsymbol{A}(u)$ that plays the role of the diffusion coefficient which is defined as

$$\boldsymbol{A}(u) = \frac{\alpha}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\left(\mathcal{I}-\frac{\nabla u\otimes\nabla u}{\nabla u\cdot\nabla u+\delta}\right) \tag{1.32}$$

(d) Eigenvalues and corresponding eigenvectors of $\boldsymbol{A}(u)$

The tensor $\boldsymbol{A}(u)\in\mathbb{R}^{2\times 2}$ hase two eigenvalues λ_1 and λ_2 and corresponding eigenvectors (or eigenfunctions) $\mathbf{v}_1$ and $\mathbf{v}_2$. Let's start with $\mathbf{v}_1 = \mu\nabla u = \left(\mu\frac{\partial u}{\partial x}, \mu\frac{\partial u}{\partial y}\right)$.

Applying it to (1.32),

$$\boldsymbol{A}\mathbf{v}_1 = \frac{\alpha}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\left(\mathcal{I}-\frac{\nabla u\otimes\nabla u}{\nabla u\cdot\nabla u+\delta}\right)\mu\nabla u \tag{1.33}$$

$$= \frac{\alpha}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\left(1-\frac{\nabla u\cdot\nabla u}{\nabla u\cdot\nabla u+\delta}\right)\mu\nabla u \tag{1.34}$$

Therefore, the first eigenvalue is

$$\lambda_1 = \frac{\alpha}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\left(1-\frac{\nabla u\cdot\nabla u}{\nabla u\cdot\nabla u+\delta}\right) \tag{1.35}$$

Next, find the next eigenvalue λ_2. Assume, the second eigenvectors to be orthogonal to the previous eigenvectors as $\boldsymbol{A}$ is symmetric. So, let $\mathbf{v}_2 = \gamma(\nabla u)^\perp$ and again applying it to (1.32),

$$\boldsymbol{A}\mathbf{v}_2 = \frac{\alpha}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\left(\mathcal{I}-\frac{\nabla u\otimes\nabla u}{\nabla u\cdot\nabla u+\delta}\right)\gamma(\nabla u)^\perp \tag{1.36}$$

$$= \frac{\alpha}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\left(\gamma(\nabla u)^\perp - \frac{\cancelto{0}{\nabla u\cdot(\nabla u)^\perp}}{\nabla u\cdot\nabla u+\delta}\gamma\nabla u\right) \tag{1.37}$$

$$= \frac{\alpha}{(\nabla u\cdot\nabla u+\delta)^{1/2}}\gamma(\nabla u)^\perp \tag{1.38}$$

Therefore, the second eigenvalue is

$$\lambda_2 = \frac{\alpha}{(\nabla u \cdot \nabla u + \delta)^{1/2}} \tag{1.39}$$

So, the eigen values and the corresponding eigenvectors are

$$\lambda_1 = \frac{\alpha}{(\nabla u \cdot \nabla u + \delta)^{1/2}} \left(1 - \frac{\nabla u \cdot \nabla u}{\nabla u \cdot \nabla u + \delta}\right), \quad \mathbf{v}_1 = \nabla u \tag{1.40}$$

$$\lambda_2 = \frac{\alpha}{(\nabla u \cdot \nabla u + \delta)^{1/2}}, \quad \mathbf{v}_2 = (\nabla u)^{\perp} \tag{1.41}$$

For a very small δ, $\lambda_1 \sim 0$ whereas $\lambda_2 \sim \frac{\alpha}{\nabla u \cdot \nabla u}$. Therefore, the diffusion is much smaller in the direction of the gradient when compared to its perpendicular direction. It can be observed from the strong form that the regularization is a diffusion equation where the diffusion is related to the laplacian Δ. This operator gives rise to a de-noising effect, which reduce the oscillations in the solution.

The difference between TV and TN regularization lies in the isotropic diffusion tensor. From the eigen-decomposition, we found that the eigenvalue parallel to the gradient (∇u) is much smaller than the in the direction perpendicular to it ($\nabla u^{\perp}$). This leads to the Newton steps that encourage diffusion and subsequently, smoothness. Moreover, it discourages diffusion and preserves sharp edges in the direction of the gradients. Thus, $\mathcal{F}^{\delta}_{TV}$ is effective at preserving the sharpness of the images. The TN regularization yields only an isotropic diffusion tensor, which results diffusion in all the directions. Therefore, TN regularization also blurs the image while de-noising.

(e) Edge preservation for large δ

Using (1.15) and (1.22), the strong forms of Hessian $\mathcal{H}^{\delta}_{TV}(u)$ and gradient $\mathcal{G}^{\delta}_{TV}(u)$ can be calculated as follows

$$\mathcal{H}^{\delta}_{TV}(u)\tilde{u} = -\nabla \cdot \left(\frac{\alpha}{(\nabla u \cdot \nabla u + \delta)^{1/2}} \left(\mathcal{I} - \frac{\nabla u \otimes \nabla u}{\nabla u \cdot \nabla u + \delta}\right) \nabla \tilde{u}\right) \tag{1.42}$$

$$\mathcal{G}^{\delta}_{TV}(u) = -\nabla \cdot \left(\frac{\alpha \nabla u}{(\nabla u \cdot \nabla u + \delta)^{1/2}}\right) \tag{1.43}$$

For large δ, one can observe that $\frac{\alpha}{(\nabla u \cdot \nabla u + \delta)^{1/2}} \approx \alpha/\sqrt{\delta}$ and $\left(\mathcal{I} - \frac{\nabla u \otimes \nabla u}{\nabla u \cdot \nabla u + \delta}\right) \approx \mathcal{I}$. Therefore, $\mathcal{H}^{\delta}_{TV}(u)\tilde{u} \approx -\nabla \cdot \left(\frac{\alpha}{\sqrt{\delta}} \nabla \tilde{u}\right)$ and $\mathcal{G}^{\delta}_{TV}(u) \approx -\nabla \cdot \left(\frac{\alpha}{\sqrt{\delta}} \nabla u\right)$. For Tikhanov, $\mathcal{H}_{TN}(u)\tilde{u} \approx -\alpha \Delta \tilde{u}$ and $\mathcal{G}_{TN}(u) = -\alpha \Delta u$. So, both Hessians and Gradients are comparable. As a result, $\mathcal{R}_{TN}$ behaves like $\mathcal{R}^{\delta}_{TV}$.

For $\delta = 0$, $\lambda_1 = 0$ from (1.40) and therefore the isotropic diffusion tensor $\boldsymbol{A}$ is singular. This renders the diffusion problem from the Hessian of $\mathcal{R}^{\delta}_{TV}$ ill-posed.

(f) Equivalence between *Optimize-then-discretize* (OTD) and *Discretize-then-optimize* (DTO) approaches

Optimize-then-discretize (Galerkin method):

Taking p to be the Newton step in infinite dimensions, we solve the system

$$\delta^2_u \mathcal{F}^{\delta}_{TV}(u, p, \hat{u}) = -\delta_u \mathcal{F}^{\delta}_{TV}(u, \hat{u}) \quad \forall \hat{u} \in H^1(\Omega)$$

Choosing the same discretization for both the trial and test spaces, we take $u_h = \sum_i u_i \phi_i$, $p_h = \sum_k p_k \phi_k$, and consider the basis function individually $\hat{u} = \phi_j$. Substituting this into the weak form of the Newton step (1.28), we arrive at

$$\int_{\Omega} \phi_j \left(\sum_k p_k \phi_k\right) + \frac{\alpha}{(|\sum_i u_i \nabla \phi_i|^2 + \delta)^{1/2}} \left(\mathcal{I} - \frac{(\sum_i u_i \nabla \phi_i) \otimes (\sum_i u_i \nabla \phi_i)}{|\sum_i u_i \nabla \phi_i|^2 + \delta}\right) \left(\sum_k p_k \nabla \phi_k\right) \cdot \nabla \phi_j \, d\boldsymbol{x}$$
$$= -\int_{\Omega} (\sum_i u_i \phi_i - d)\phi_j + \frac{\alpha(\sum_i u_i \nabla \phi_i \cdot \nabla \phi_j)}{(|\sum_i u_i \nabla \phi_i|^2 + \delta)^{1/2}} \, d\boldsymbol{x}$$

Using the symmetry of the tensor, we finally get the following form Newton step for OTD

$$\sum_k p_k \int_\Omega \phi_j\phi_k + \frac{\alpha}{(|\sum_i u_i\nabla\phi_i|^2+\delta)^{1/2}}\left(\mathcal{I} - \frac{(\sum_i u_i\nabla\phi_i)\otimes(\sum_i u_i\nabla\phi_i)}{|\sum_i u_i\nabla\phi_i|^2+\delta}\right)\nabla\phi_j\cdot\nabla\phi_k \tag{1.44}$$

$$= -\int_\Omega \left(\sum_i u_i\phi_i - d\right)\phi_j + \frac{\alpha(\sum_i u_i\nabla\phi_i\cdot\nabla\phi_j)}{(|\sum_i u_i\nabla\phi_i|^2+\delta)^{1/2}}\,d\boldsymbol{x} \tag{1.45}$$

Discretize-then-optimize (Ritz method): From (1.1) and (1.4)

$$\mathcal{F}^\delta_{TV}(u) = \mathcal{F}_{LS} + \mathcal{R}^\delta_{TV} = \frac{1}{2}\int_\Omega (u-d)^2\,d\boldsymbol{x} + \int_\Omega \alpha(\nabla u\cdot\nabla u+\delta)\,d\boldsymbol{x} \tag{1.46}$$

Consider a function approximation for $u \in H^1(\Omega)$ as $u \approx u_h = \sum_i u_i\phi_i$. Substituting this into the energy functional gives

$$\mathcal{G}^\delta_{TV}(\mathbf{u}) = \int_\Omega \frac{1}{2}(\sum_i u_i\phi_i - d)^2 + \alpha\left[(\sum_i u_i\nabla\phi_i)\cdot(\sum_i u_i\nabla\phi_i)+\delta\right]^{1/2}\,d\boldsymbol{x} \tag{1.47}$$

Taking the gradient of the j-th component g_j

$$g_j = \frac{\partial}{\partial u_j}\mathcal{G}^\delta_{TV}(\mathbf{u}) = \int_\Omega \left(\sum_i u_i\phi_i - d\right)\phi_j + \alpha\frac{\sum_i(u_i\nabla\phi_i\cdot\nabla\phi_j)}{(|\sum_i u_i\nabla\phi_i|^2+\delta)^{1/2}}\,d\boldsymbol{x}$$

The (j,k) component of the Hessian $\mathbf{H}$ is calculated as

$$\begin{aligned}
H_{jk} &= \frac{\partial g_j}{\partial u_k}\\
&= \int_\Omega \phi_j\phi_k + \frac{\alpha}{(|\sum_i u_i\nabla\phi_i|^2+\delta)^{1/2}}\left(\nabla\phi_j\cdot\nabla\phi_k - \frac{\nabla\phi_k\cdot(\sum_i u_i\nabla\phi_i)\nabla\phi_j\cdot(\sum_i u_i\nabla\phi_i)}{|\sum_i u_i\nabla\phi_i+\delta|}\right)\,d\boldsymbol{x}\\
&= \int_\Omega \phi_j\phi_k + \frac{\alpha}{(|\sum_i u_i\nabla\phi_i|^2+\delta)^{1/2}}\left(\mathcal{I} - \frac{(\sum_i u_i\nabla\phi_i)\otimes(\sum_i u_i\nabla\phi_i)}{|\sum_i u_i\nabla\phi_i+\delta|}\right)\nabla\phi_j\cdot\nabla\phi_k\,d\boldsymbol{x}
\end{aligned}$$

The Newton step for $\mathbf{p}$ is given by $\sum_k H_{jk}p_k = -g_j$. Substituting this into the expression (1.48) gives

$$\sum_k p_k \int_\Omega \phi_j\phi_k + \frac{\alpha}{(|\sum_i u_i\nabla\phi_i|^2+\delta)^{1/2}}\left(\mathcal{I} - \frac{(\sum_i u_i\nabla\phi_i)\otimes(\sum_i u_i\nabla\phi_i)}{|\sum_i u_i\nabla\phi_i+\delta|}\right)\nabla\phi_j\cdot\nabla\phi_k \tag{1.48}$$

$$= -\int_\Omega \left(\sum_i u_i\phi_i - d\right)\phi_j + \frac{\alpha\sum_i(u_i\nabla\phi_i\cdot\nabla\phi_j)}{(|\sum_i u_i\nabla\phi_i|^2+\delta)^{1/2}}\,d\boldsymbol{x} \tag{1.49}$$

Therefore, OTD and DTO approaches are equivalent.

2 Installing FEniCS

FEniCS has been installed using `conda` commands. It can be accessed directly through terminal or jupyter notebook.

3 FEniCS to solve 2D Poisson problem (BVP)

$$\nabla\cdot(\boldsymbol{A}\nabla u) = f \quad \text{in } \Omega \tag{3.1}$$

$$u = u_0 \quad \text{on } \Gamma \tag{3.2}$$

where the conductivity tensor $\boldsymbol{A}(\boldsymbol{x}) \in \mathbb{R}^{2\times 2}$ is assumed to be symmetric and positive definite for all $\boldsymbol{x}, f(\boldsymbol{x})$ is a given distributed source, and $u_0(\boldsymbol{x})$ is the source on the boundary Γ.

(a) Weak form

We consider a trial space, $u \in H^1(\Omega)$ such that $u|_\Gamma = u_0$. For the weak form, multiply (3.2) by a test function $v(\boldsymbol{x}) \in H_0^1(\Omega)$ and integrate it over the domain Ω.

$$-\int_\Omega \nabla \cdot (\boldsymbol{A}\nabla u) v d\boldsymbol{x} = \int_\Omega f v \; d\boldsymbol{x} \tag{3.3}$$

Using Green's identity,

$$\int_\Omega -\nabla \cdot (\boldsymbol{A}\nabla u v) + \boldsymbol{A}\nabla u \cdot \nabla v d\boldsymbol{x} = \int_\Omega f v \; d\boldsymbol{x} \tag{3.4}$$

Using Gauss divergence theorem in the first term,

$$-\int_\Gamma \boldsymbol{A} v \nabla u \cdot \boldsymbol{n} d\boldsymbol{s} + \int_\Omega \boldsymbol{A}\nabla u \cdot \nabla v d\boldsymbol{x} = \int_\Omega f v \; d\boldsymbol{x} \tag{3.5}$$

On the boundary $v = 0$ as it is a Dirichlet boundary condition, therefore we get the *weak form*

$$\int_\Omega \boldsymbol{A}\nabla u \cdot \nabla v d\boldsymbol{x} = \int_\Omega f v \; d\boldsymbol{x} \tag{3.6}$$

Energy functional: The energy functional can be found by substituting u in the place of v. It is defined as follows

$$\min_{u \in H^1(\Omega), u|_\Gamma = u_0} \mathcal{F}(u) := \int_\Omega \boldsymbol{A}\nabla u \cdot \nabla u d\boldsymbol{x} - \int_\Omega f u \; d\boldsymbol{x} \tag{3.7}$$

It can be verified by the optimal condition on the first variation $\delta_u \mathcal{F}(u, v) = 0 \; \forall v \in H_0^1(\Omega)$.

(b) Solving using FEniCS

In the Jupyter notebook titled '`HW3_Q3b.ipynb`'.

4 Implementation of image de-noising using Tikhonov and Total Variation Regularizations in FEniCS

In the Jupyter notebook titled '`HW3_Q4.ipynb`'.