

# Gradient Computation

ss2571

May 2020

This document contains the gradient formula for the linear and nonlinear lasso problem.

## 1 Setup

Define two points in spherical coordinates:

$$\mathbf{Q} := (1, \theta, \psi) \in \partial\Omega, \quad \mathbf{P}_i := (r_i, \theta_i, \psi_i) \in \Omega. \quad (1)$$

The normalized forward solution reads:

$$\hat{\Phi}(\mu)(\mathbf{Q}) = \sum_{i=1}^N \frac{I_i}{G_{\mathbf{P}_i}} \phi_{\mathbf{P}_i}(\mathbf{Q}), \quad \mathbf{Q} \in \partial\Omega, \text{ where,} \quad (2)$$

$$G_{\mathbf{P}_i} = \sqrt{\int_{\partial\Omega} |F(\mathbf{Q}, \mathbf{P}_i)|^2 d\mathbf{Q}} = \sqrt{\int_{\partial\Omega} \phi_{\mathbf{P}_i}^2(\mathbf{Q}) d\mathbf{Q}} \quad (3)$$

where,

$$\begin{aligned} \phi_{\mathbf{P}_i}(\mathbf{Q}) &= \frac{2}{\|\mathbf{Q} - \mathbf{P}_i\|_2} - \log(1 - \langle \mathbf{P}_i, \mathbf{Q} \rangle + \|\mathbf{Q} - \mathbf{P}_i\|_2) \\ &= \frac{2}{\sqrt{1 + r_i^2 - 2r_i \cos \gamma_i}} - \log(1 - r_i \cos \gamma_i + \sqrt{1 + r_i^2 - 2r_i \cos \gamma_i}), \text{ where,} \end{aligned} \quad (4)$$

$$\cos \gamma_i = \cos \theta \cos \theta_i + \sin \theta \sin \theta_i \cos(\psi - \psi_i), \quad \langle \mathbf{P}_i, \mathbf{Q} \rangle = r_i \cos \gamma_i, \quad \|\mathbf{Q} - \mathbf{P}_i\|_2 = \sqrt{1 + r_i^2 - 2r_i \cos \gamma_i}.$$

Define the source parameter vector as follows:

$$\mathbf{V} = \{\overbrace{I_1, I_2, \dots}^N, \overbrace{\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_N}^{3N}\}, \quad (5)$$

$$\mathbf{I} = \{I_1, I_2, \dots, I_N\}. \quad (6)$$

And in this report, we use the symbol  $D^i$  to represent the partial derivative with respect to the  $i$ th component of the unknown parameter vector.

## 2 Linear lasso problem

For the linear lasso problem, locations of the source (for example,  $\mathbf{P}_i$ ) are given, we define the estimation function as follows:

$$Est(\mathbf{I}, \mathbf{Q}) = \sum_{i=1}^N \frac{I_i}{G_{\mathbf{P}_i}} \phi_{\mathbf{P}_i}(\mathbf{Q}), \quad \mathbf{Q} \in \partial\Omega, \quad (7)$$

## Linear forward operator and its adjoint operator

Define a vector function  $\mathbf{A}(\mathbf{Q}) \in R^N$  such that,

$$\mathbf{A}(\mathbf{Q}) = \{A_1(\mathbf{Q}), A_2(\mathbf{Q}), \dots, A_N(\mathbf{Q})\}, \text{ and,} \quad (8)$$

$$A_i(\mathbf{Q}) := \frac{\phi_{P_i}(\mathbf{Q})}{G_{P_i}} \in L^2(\partial\Omega). \quad (9)$$

We define the linear forward operator as follows:

$$\begin{aligned} \Phi : R^N &\rightarrow L^2(\partial\Omega), \text{ s.t.,} \\ \Phi(\mathbf{I}) &= \langle \mathbf{A}(\mathbf{Q}), \mathbf{I} \rangle = \sum_{i=1}^N I_i A_i(\mathbf{Q}). \end{aligned} \quad (10)$$

Then, the estimation function becomes:

$$Est(\mathbf{I}, \mathbf{Q}) = \Phi(\mathbf{I}) \quad (11)$$

Next we define the adjoint of  $\Phi$ , denote it as  $\Phi^*$ :

$$\Phi^* : L^2(\partial\Omega) \rightarrow R^N, \text{ s.t.,} \quad (12)$$

$$\Phi^*(p(\mathbf{Q})) = \int_{\partial\Omega} \mathbf{A}(\mathbf{Q}) p(\mathbf{Q}) d\mathbf{Q}, \quad p(\mathbf{Q}) \in L^2(\partial\Omega). \quad (13)$$

The composition map  $\Phi^*\Phi$  derives from the following steps, by definition,

$$\Phi^*\Phi : R^N \rightarrow R^N. \quad (14)$$

Take arbitrary vector  $\mathbf{a} \in R^N$ , and apply the map, we get:

$$\Phi^*\Phi(\mathbf{a}) = \int_{\partial\Omega} \mathbf{A}(\mathbf{Q}) \sum_{i=1}^N A_i(\mathbf{Q}) a_i d\mathbf{Q} \in R^N, \quad (15)$$

take the  $k$ th entry of  $\Phi^*\Phi(\mathbf{a})$ , we derive:

$$[\Phi^*\Phi(\mathbf{a})]_k = \int_{\partial\Omega} A_k(\mathbf{Q}) A_i(\mathbf{Q}) a_i d\mathbf{Q} = \sum_{i=1}^N \left( \int_{\partial\Omega} A_k(\mathbf{Q}) A_i(\mathbf{Q}) d\mathbf{Q} \right) a_i = \sum_{i=1}^N M_{k,i} a_i \quad (16)$$

, from above we derive that,

$$\mathcal{M}_{i,j} := (\Phi^*\Phi)_{i,j} = \int_{\partial\Omega} A_i(\mathbf{Q}) A_j(\mathbf{Q}) d\mathbf{Q}. \quad (17)$$

Therefore,

$$\Phi^*\Phi = \mathcal{M} \in R^{N,N}. \quad (18)$$

## Objective function

The objective function reads:

$$Obj(\mathbf{I}) = \frac{1}{2} \|\Phi(\mathbf{I}) - \phi^d(\mathbf{Q})\|_{\partial\Omega}^2 + \lambda \|\mathbf{I}\|_1, \quad (19)$$

$$= \int_{\partial\Omega} [\Phi(\mathbf{I}) - \phi^d(\mathbf{Q})]^2 d\mathbf{Q} + \lambda \sum_i |I_i|. \quad (20)$$

For convenience, we introduce the following two terms:

$$f_{des}(\mathbf{I}) := \frac{1}{2} \|\Phi(\mathbf{I}) - \phi^d(\mathbf{Q})\|_{\partial\Omega}^2 = \frac{1}{2} \int_{\partial\Omega} [\Phi(\mathbf{I}) - \phi^d(\mathbf{Q})]^2 d\mathbf{Q}, \quad (21)$$

$$f_{reg}(\mathbf{I}) := \lambda \|\mathbf{I}\|_1. \quad (22)$$

## 2.1 The gradient and hessian of $f_{des}$

Through a direct calculation, we find that,

$$\nabla f_{des}(\mathbf{I}) = \Phi^* \Phi(\mathbf{I}) - \Phi^*(\phi^d) = \mathcal{M}\mathbf{I} - \mathbf{b}, \text{ where, } \mathbf{b} = \Phi^*(\phi^d). \quad (23)$$

Follows from above, we find the hessian of  $f_{des}$ ,

$$\mathcal{H}(f_{des}(\mathbf{I})) = \mathcal{M}. \quad (24)$$

### 2.1.1 The Lipschitz constant of $\nabla f_{des}$

**Proposition 2.1.** *The gradient  $\nabla f_{des}(\mathbf{I})$  has a Lipschitz constant as follows:*

$$L = \|\Phi^* \Phi\|_2, \text{ where } \|\cdot\|_2 \text{ is the matrix l2 norm.} \quad (25)$$

*Proof.* Let  $\mathbf{I}_a, \mathbf{I}_b \in R^N$  be two arbitrary vectors, then,

$$\|\nabla f_{des}(\mathbf{I}_a) - \nabla f_{des}(\mathbf{I}_b)\|_2 = \|\mathcal{M}(\mathbf{I}_a - \mathbf{I}_b)\|_2 \leq \|\mathcal{M}\|_2 \|\mathbf{I}_a - \mathbf{I}_b\|_2. \quad (26)$$

□

## 2.2 The subgradient of $f_{reg}$

Denote the subgradient of  $f_{reg}$  as  $\partial f_{reg}$ , we have that,

$$\partial f_{reg}(\mathbf{I}) = \lambda \text{ sign}(\mathbf{I}). \quad (27)$$

## 2.3 Gradient calculation for linear solvers

### FBS

For the FBS solver, only the gradient of the discrepancy term is required, and locations of the source are given. Therefore we only compute:

$$\nabla f_{des}(\mathbf{I}) = \mathcal{M}\mathbf{I} - \Phi^*(\phi^d(\mathbf{Q})), \text{ where,} \quad (28)$$

$$\mathcal{M}_{i,j} = \int_{\partial\Omega} A_i(\mathbf{Q}) A_j(\mathbf{Q}) d\mathbf{Q}, \text{ and, } \Phi^*(\phi^d(\mathbf{Q})) = \int_{\partial\Omega} \mathbf{A}(\mathbf{Q}) \phi^d(\mathbf{Q}) d\mathbf{Q}. \quad (29)$$

### Quasi Newton

For quasi newton solver, we compute the gradient of the entire objective function, i.e.,

$$\nabla_{pseudo} Obj(\mathbf{I}) = \nabla f_{des} + \partial f_{reg}. \quad (30)$$

*Proof.*

□

## 3 Nonlinear lasso problem

For the nonlinear lasso problem, both locations and intensities are unknown, the forward estimation reads:

$$Est(\mathbf{V}, \mathbf{Q}) = \sum_{i=1}^N \frac{I_i}{G_{P_i}} \phi_{P_i}(\mathbf{Q}), \quad \mathbf{Q} \in \partial\Omega, \quad (31)$$

The objective function reads:

$$Obj = \frac{1}{2} \|Est(\mathbf{V}, \mathbf{Q}) - \phi^d(\mathbf{Q})\|_{\partial\Omega}^2 + \lambda \|\mathbf{I}\|_1, \quad (32)$$

$$= \frac{1}{2} \int_{\partial\Omega} [Est(\mathbf{V}, \mathbf{Q}) - \phi^d(\mathbf{Q})]^2 d\mathbf{Q} + \lambda \sum_i |I_i|. \quad (33)$$

The pseudo gradient of the nonlinear lasso problem consists of two component: the gradient of the discrepancy term and the subgradient of the l1 regularizatoion term.

## The discrepancy gradient

In this section we compute the gradient term  $\nabla f_{des}$ .

For  $i = 1 \rightarrow N$ ,

$$(\nabla f_{des})_i = D^i f_{des} = \int_{\partial\Omega} (Est(\mathbf{V}, \mathbf{Q}) - \phi^d(\mathbf{Q})) \frac{1}{G_{\mathbf{P}_i}} \phi_{\mathbf{P}_i}(\mathbf{Q}) d\mathbf{Q}, \quad (34)$$

for  $i = N + 1 \rightarrow 4N$ ,

$$(\nabla f_{des})_i = D^i f_{des} = \int_{\partial\Omega} (Est(\mathbf{V}, \mathbf{Q}) - \phi^d(\mathbf{Q})) D^i Est(\mathbf{V}, \mathbf{Q}) d\mathbf{Q}, \quad (35)$$

$$D^i Est(\mathbf{V}, \mathbf{Q}) = I_i \frac{G_{\mathbf{P}_i} D^i \phi_{\mathbf{P}_i} - \phi_{\mathbf{P}_i} D^i G_{\mathbf{P}_i}}{G_{\mathbf{P}_i}^2}, \quad (36)$$

Recall that:

$$\phi_{\mathbf{P}_i}(\mathbf{Q}) = \frac{2}{\|\mathbf{Q} - \mathbf{P}_i\|_2} - \log(1 - \langle \mathbf{P}_i, \mathbf{Q} \rangle + \|\mathbf{Q} - \mathbf{P}_i\|_2),$$

$$G_{\mathbf{P}_i} = \sqrt{\int_{\partial\Omega} |F(\mathbf{Q}, \mathbf{P}_i)|^2 d\mathbf{Q}} = \sqrt{\int_{\partial\Omega} \phi_{\mathbf{P}_i}^2(\mathbf{Q}) d\mathbf{Q}}.$$

Thus,

$$D^i \phi_{\mathbf{P}_i} = -\frac{2}{\|\mathbf{Q} - \mathbf{P}_i\|_2^2} D^i \|\mathbf{Q} - \mathbf{P}_i\|_2 - \frac{1}{1 - \langle \mathbf{P}_i, \mathbf{Q} \rangle + \|\mathbf{Q} - \mathbf{P}_i\|_2} D^i (-\langle \mathbf{P}_i, \mathbf{Q} \rangle + \|\mathbf{Q} - \mathbf{P}_i\|_2). \quad (37)$$

$$D^i G_{\mathbf{P}_i} = \frac{\int_{\partial\Omega} \phi_{\mathbf{P}_i}(\mathbf{Q}) D^i \phi_{\mathbf{P}_i}(\mathbf{Q}) d\mathbf{Q}}{G_{\mathbf{P}_i}}. \quad (38)$$

Next we compute the following terms:

$D^i \|\mathbf{Q} - \mathbf{P}_i\|_2$ ,  $D^i \langle \mathbf{Q}, \mathbf{P}_i \rangle$ , where,

$$\langle \mathbf{P}_i, \mathbf{Q} \rangle = r_i \cos \gamma_i, \quad \|\mathbf{Q} - \mathbf{P}_i\|_2 = \sqrt{1 + r_i^2 - 2r_i \cos \gamma_i},$$

$$\cos \gamma_i = \cos \theta \cos \theta_i + \sin \theta \sin \theta_i \cos(\psi - \psi_i).$$

For  $i = N + 1 : 3 : 4N - 2$ ,  $V_i = r_i$ ,  $D^i = \frac{\partial}{\partial r_i}$  therefore :

$$D^i \langle \mathbf{Q}, \mathbf{P}_i \rangle = \cos \gamma_i, \quad (39)$$

$$D^i \|\mathbf{Q} - \mathbf{P}_i\|_2 = \frac{\partial \|\mathbf{Q} - \mathbf{P}_i\|_2}{\partial r_i} = \frac{r_i - \cos \gamma_i}{\|\mathbf{Q} - \mathbf{P}_i\|_2}, \quad (40)$$

for  $i = N + 2 : 3 : 4N - 1$ ,  $V_i = \theta_i$ ,  $D^i = \frac{\partial}{\partial \theta_i}$  therefore :

$$D^i \langle \mathbf{Q}, \mathbf{P}_i \rangle = r_i (-\cos \theta \sin \theta_i + \sin \theta \cos \theta_i \cos(\psi - \psi_i)), \quad (41)$$

$$D^i \|\mathbf{Q} - \mathbf{P}_i\|_2 = \frac{\partial \|\mathbf{Q} - \mathbf{P}_i\|_2}{\partial \theta_i} = \frac{-r_i (-\cos \theta \sin \theta_i + \sin \theta \cos \theta_i \cos(\psi - \psi_i))}{\|\mathbf{Q} - \mathbf{P}_i\|_2}, \quad (42)$$

for  $i = N + 3 : 3 : 4N$ ,  $V_i = \psi_i$ ,  $D^i = \frac{\partial}{\partial \psi_i}$  therefore :

$$D^i \langle \mathbf{Q}, \mathbf{P}_i \rangle = r_i (-\sin \theta \sin \theta_i \sin(\psi_i - \psi)), \quad (43)$$

$$D^i \|\mathbf{Q} - \mathbf{P}_i\|_2 = \frac{\partial \|\mathbf{Q} - \mathbf{P}_i\|_2}{\partial \psi_i} = \frac{-r_i (-\sin \theta \sin \theta_i \sin(\psi_i - \psi))}{\|\mathbf{Q} - \mathbf{P}_i\|_2}. \quad (44)$$

## Pseudo Gradient

Whenever the gradient of the whole objective function is required, we compute the discrepancy gradient and add the subgradient of the regularization term to it, i.e.,

$$PsudoGrad_i = \begin{cases} Grad_i + \lambda * sign(I_i), & i = 1 \rightarrow N, \\ Grad_i, & i > N. \end{cases} \quad (45)$$