

Gradient Computation

ss2571

May 2020

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

1 Gradient computation

Define two points in spherical coordinates as follows:

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

The normalized forward solution reads:

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

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

where,

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

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

Define the source parameter vector as follows:

$$\mathbf{V} = \{\overbrace{I_1, I_2, \dots}^N, \overbrace{r_1, r_2, \dots}^N, \overbrace{\theta_1, \theta_2, \dots}^N, \overbrace{\psi_1, \psi_2, \dots}^N\}, \quad (5)$$

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

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.

1.1 Linear lasso problem

For the linear lasso problem, locations of the source (i.e., \mathbf{P}_i) are given, we define the forward estimation as follows:

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

The objective function reads:

$$Obj = \frac{1}{2} \|Est(\mathbf{I}_{all}, \mathbf{Q}) - \phi^d(\mathbf{Q})\|_2^2 + \lambda \|\mathbf{I}_{all}\|_1, \quad (8)$$

$$= \int_{\partial\Omega} [Est(\mathbf{I}_{all}, \mathbf{Q}) - \phi^d(\mathbf{Q})]^2 d\mathbf{Q} + \lambda \sum_i |I_i|. \quad (9)$$

For convenience, we introduce the following two terms:

$$f_{des}(\mathbf{I}_{all}) := \frac{1}{2} \|Est(\mathbf{I}_{all}, \mathbf{Q}) - \phi^d(\mathbf{Q})\|_2^2 = \frac{1}{2} \int_{\partial\Omega} [Est(\mathbf{I}_{all}, \mathbf{Q}) - \phi^d(\mathbf{Q})]^2 d\mathbf{Q}, \quad (10)$$

$$f_{reg} := \lambda \|\mathbf{I}_{all}\|_1. \quad (11)$$

1.2 Subgradient of the regularization term

$$D^i f_{reg} = \lambda \text{sign}(I_i). \quad (12)$$

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:

$$\begin{aligned} Grad_i &= D^i f_{des}(\mathbf{Q}) = \int_{\partial\Omega} (Est(\mathbf{I}_{all}, \mathbf{Q}) - \phi^d(\mathbf{Q})) D^i Est(\mathbf{I}_{all}, \mathbf{Q}) d\mathbf{Q}, \\ &= \int_{\partial\Omega} (Est(\mathbf{I}_{all}, \mathbf{Q}) - \phi^d(\mathbf{Q})) \frac{1}{G_{P_i}} \phi_{P_i}(\mathbf{Q}) d\mathbf{Q}, \\ &i = 1, 2, \dots, N. \end{aligned} \quad (13)$$

In the minimization iteration, source location P_i is provided by the solver in each search step.

Quasi Newton

For quasi newton solver, gradient of the whole objective function needs to be computed.

$$PseudoGrad_i = \int_{\partial\Omega} (Est(\mathbf{I}_{all}, \mathbf{Q}) - \phi^d(\mathbf{Q})) \frac{1}{G_{P_i}} \phi_{P_i}(\mathbf{Q}) d\mathbf{Q} + \lambda \text{sign}(I_i). \quad (14)$$

The Lipschitz constant of the discrepancy gradient

Proposition 1.1. *Let $\mathbf{I}_{all,1}, \mathbf{I}_{all,2}$ be two intensity solution vectors and $I_{i,1}, I_{i,2}$ be their components, and let:*

$$\|\mathbf{x}\| := \sqrt{x_1^2 + x_2^2 + \dots + x_N^2}, \text{ for any } \mathbf{x} \in R^N.$$

Define a vector as follows:

$$\begin{aligned} \mathbf{A}(\mathbf{Q}) &:= \{D^1 Est(\mathbf{I}_{all}, \mathbf{Q}), D^2 Est(\mathbf{I}_{all}, \mathbf{Q}), \dots, D^N Est(\mathbf{I}_{all}, \mathbf{Q})\} \\ &= \left\{ \frac{1}{G_{P_1}} \phi_{P_1}(\mathbf{Q}), \frac{1}{G_{P_2}} \phi_{P_2}(\mathbf{Q}), \dots, \frac{1}{G_{P_N}} \phi_{P_N}(\mathbf{Q}) \right\}. \end{aligned}$$

, then the gradient $\nabla f_{des}(\mathbf{I}_{all})$ has a Lipschitz constant as follows:

$$L = \left| \sum_i \int_{\partial\Omega} \|\mathbf{A}(\mathbf{Q})\| \|\mathbf{A}_i(\mathbf{Q})\| d\mathbf{Q} \right|. \quad (15)$$

Proof. Notice that:

$$\begin{aligned} \|\nabla f_{des}(\mathbf{I}_{all,1}) - \nabla f_{des}(\mathbf{I}_{all,2})\| &= \left\| \int_{\partial\Omega} [Est(\mathbf{I}_{all,1}, \mathbf{Q}) - Est(\mathbf{I}_{all,2}, \mathbf{Q})] \cdot \nabla Est(\mathbf{I}_{all}, \mathbf{Q}) d\mathbf{Q} \right\| \\ &= \left\| \int_{\partial\Omega} \langle \mathbf{A}(\mathbf{Q}), \mathbf{I}_{all,1} - \mathbf{I}_{all,2} \rangle \mathbf{A}(\mathbf{Q}) d\mathbf{Q} \right\|, \end{aligned} \quad (16)$$

Using the Cauchy Schwarz inequality to the vector l2 norm, we get:

$$\int_{\partial\Omega} \langle \mathbf{A}(\mathbf{Q}), \mathbf{I}_{all,1} - \mathbf{I}_{all,2} \rangle \mathbf{A}(\mathbf{Q}) d\mathbf{Q} \leq \int_{\partial\Omega} \|\mathbf{A}(\mathbf{Q})\| \|\mathbf{I}_{all,1} - \mathbf{I}_{all,2}\| \|\mathbf{A}(\mathbf{Q})\| d\mathbf{Q}. \quad (17)$$

Return to Equation (16), we derive the following:

$$\|\nabla f_{des}(\mathbf{I}_{all,1}) - \nabla f_{des}(\mathbf{I}_{all,2})\| = \left\| \int_{\partial\Omega} \langle \mathbf{A}(\mathbf{Q}), \mathbf{I}_{all,1} - \mathbf{I}_{all,2} \rangle \mathbf{A}(\mathbf{Q}) d\mathbf{Q} \right\| \quad (18)$$

$$\leq \sqrt{\sum_i [\|\mathbf{I}_{all,1} - \mathbf{I}_{all,2}\| \int_{\partial\Omega} \|\mathbf{A}(\mathbf{Q})\| \|\mathbf{A}_i(\mathbf{Q})\| d\mathbf{Q}]^2} \quad (19)$$

$$= \left| \sum_i \int_{\partial\Omega} \|\mathbf{A}(\mathbf{Q})\| \|\mathbf{A}_i(\mathbf{Q})\| d\mathbf{Q} \right| \|\mathbf{I}_{all,1} - \mathbf{I}_{all,2}\| \quad (20)$$

therefore, we obtain an estimation of the Lipschitz constant as follows,

$$L = \left| \sum_i \int_{\partial\Omega} \|\mathbf{A}(\mathbf{Q})\| \|\mathbf{A}_i(\mathbf{Q})\| d\mathbf{Q} \right|. \quad (21)$$

□

1.3 Nonlinear lasso problem

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

$$Est(\mathbf{V}, \mathbf{Q}) = \sum_{i=1}^N \frac{I_i}{G_{\mathbf{P}_i}} \phi_{\mathbf{P}_i}(\mathbf{Q}), \quad \mathbf{Q} \in \partial\Omega, \quad (22)$$

The objective function reads:

$$Obj = \frac{1}{2} \|\text{Est}(\mathbf{V}, \mathbf{Q}) - \phi^d(\mathbf{Q})\|_2^2 + \lambda \|\mathbf{I}_{all}\|_1, \quad (23)$$

$$= \int_{\partial\Omega} [\text{Est}(\mathbf{V}, \mathbf{Q}) - \phi^d(\mathbf{Q})]^2 d\mathbf{Q} + \lambda \sum_i |I_i|. \quad (24)$$

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

The discrepancy gradient

In this section we compute the gradient term ∇f_{des} .

For $i = 1 \rightarrow N$,

$$Grad_i = D^i f_{des}(\mathbf{Q}) = D^i f_{des}(\mathbf{Q}) = \int_{\partial\Omega} (\text{Est}(\mathbf{V}, \mathbf{Q}) - \phi^d(\mathbf{Q})) \frac{1}{G_{\mathbf{P}_i}} \phi_{\mathbf{P}_i}(\mathbf{Q}) d\mathbf{Q}, \quad (25)$$

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

$$Grad_i = D^i f_{des}(\mathbf{Q}) = D^i f_{des}(\mathbf{Q}) = \int_{\partial\Omega} (\text{Est}(\mathbf{V}, \mathbf{Q}) - \phi^d(\mathbf{Q})) D^i \text{Est}(\mathbf{V}, \mathbf{Q}) d\mathbf{Q}, \quad (26)$$

$$D^i \text{Est}(\mathbf{V}, \mathbf{Q}) = I_i \frac{G_{\mathbf{P}_i} D^i \phi_{\mathbf{P}_i} - \phi_{\mathbf{P}_i} D^i G_{\mathbf{P}_i}}{G_{\mathbf{P}_i}^2}, \quad (27)$$

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,

$$D^i \phi_{\mathbf{P}_i} = -\frac{1}{\|\mathbf{Q} - \mathbf{P}_i\|_2^3} D^i \|\mathbf{Q} - \mathbf{P}_i\|_2 - \frac{1}{1 - \langle \mathbf{P}_i, \mathbf{Q} \rangle + \|\mathbf{Q} - \mathbf{P}_i\|_2} D^i (-\langle \mathbf{P}_i, \mathbf{Q} \rangle + \|\mathbf{Q} - \mathbf{P}_i\|_2). \quad (28)$$

$$D^i G_{\mathbf{P}_i} = \frac{\int_{\partial\Omega} \phi_{\mathbf{P}_i}(\mathbf{Q}) D^i \phi_{\mathbf{P}_i}(\mathbf{Q}) d\mathbf{Q}}{G_{\mathbf{P}_i}}. \quad (29)$$

Next we compute the following terms:

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

For $i = N + 1 \rightarrow 2N$:

$$D^i \langle \mathbf{Q}, \mathbf{P}_i \rangle = \cos \gamma_i, \quad (30)$$

$$D^i \|\mathbf{Q} - \mathbf{P}_i\|_2 = \frac{\partial \|\mathbf{Q} - \mathbf{P}_i\|_2}{\partial r_i} = \frac{r_i - \cos \gamma_i}{\|\mathbf{Q} - \mathbf{P}_i\|_2}, \quad (31)$$

for $i = 2N + 1 \rightarrow 3N$:

$$D^i \langle \mathbf{Q}, \mathbf{P}_i \rangle = r_i (-\cos \theta \sin \theta_i + \sin \theta \cos \theta_i \cos(\psi - \psi_i)), \quad (32)$$

$$D^i \|\mathbf{Q} - \mathbf{P}_i\|_2 = \frac{\partial \|\mathbf{Q} - \mathbf{P}_i\|_2}{\partial \theta_i} = \frac{-r_i (-\cos \theta \sin \theta_i + \sin \theta \cos \theta_i \cos(\psi - \psi_i))}{\|\mathbf{Q} - \mathbf{P}_i\|_2}, \quad (33)$$

for $i = 3N + 1 \rightarrow 4N$:

$$D^i \langle \mathbf{Q}, \mathbf{P}_i \rangle = r_i (\cos \theta \cos \theta_i - \sin \theta \sin \theta_i \sin(\psi - \psi_i)), \quad (34)$$

$$D^i \|\mathbf{Q} - \mathbf{P}_i\|_2 = \frac{\partial \|\mathbf{Q} - \mathbf{P}_i\|_2}{\partial \psi_i} = \frac{-r_i (-\sin \theta \sin \theta_i \sin(\psi_i - \psi))}{\|\mathbf{Q} - \mathbf{P}_i\|_2}. \quad (35)$$

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.,

$$PseudoGrad_i = \begin{cases} Grad_i + \lambda * sign(I_i), & i = 1 \rightarrow N, \\ Grad_i, & i > N. \end{cases} \quad (36)$$