Where the Delta Functions in the Adjoint Source Come From

Bill Menke, June 16, 2023

The L_2 total wavefield error is defined as

$$E' \equiv \sum_i \left(e(\mathbf{x}^{(i)}, t), e(\mathbf{x}^{(i)}, t) \right)_t \quad \text{with} \quad e(\mathbf{x}^{(i)}, t) \equiv u^{obs}(\mathbf{x}^{(i)}, t) - u(\mathbf{x}^{(i)}, t, m) \tag{1}$$

Here the sum is over observation points, $\mathbf{x}^{(i)}$, the subscript on the inner product indicates the variable(s) of integration, and m is a model parameter. The goal is to determine a formula for $\partial E / \partial m$ that can be used in the gradient method to determine the model parameter, m, that minimizes the error, E'.

The adjoint formulas require that the error be expressed as an inner product over $(\mathbf{x}, t)$, whereas (1) contains an inner product over t, only. This problem is handled by introducing a Dirac comb (sum of delta functions), with the spikes at the observer locations

$$E \equiv \left(e(\mathbf{x}, t), e(\mathbf{x}, t) \sum_i \delta(\mathbf{x} - \mathbf{x}^{(i)}) \right)_{\mathbf{x},t} \tag{2}$$

Then, $E = E'$, as

$$\left(e(\mathbf{x}, t), e(\mathbf{x}, t) \sum_i \delta(\mathbf{x} - \mathbf{x}^{(i)}) \right)_{\mathbf{x},t} = \sum_i \left(e(\mathbf{x}^{(i)}, t), e(\mathbf{x}^{(i)}, t) \right)_t \tag{3}$$

Differentiating (2) with respect to the model parameter, m, yields

$$\frac{\partial E}{\partial m} = \frac{\partial}{\partial m} \left(e(\mathbf{x}, t), e(\mathbf{x}, t) \sum_i \delta(\mathbf{x} - \mathbf{x}^{(i)}) \right)_{\mathbf{x},t}$$

$$= -2 \left(\frac{\partial u}{\partial m}, e(\mathbf{x}, t) \sum_i \delta(\mathbf{x} - \mathbf{x}^{(i)}) \right)_{\mathbf{x},t}$$

$$= -2 \left(\frac{\partial u}{\partial m}, \sum_i e(\mathbf{x}^{(i)}, t) \delta(\mathbf{x} - \mathbf{x}^{(i)}) \right)_{\mathbf{x},t}$$

$$= -2 \left(\frac{\partial u}{\partial m}, s(\mathbf{x}, t) \right)_{\mathbf{x},t} \quad \text{with} \quad s(\mathbf{x}, t) \equiv \sum_i e(\mathbf{x}^{(i)}, t) \delta(\mathbf{x} - \mathbf{x}^{(i)})$$

(4)

Presuming that the wavefield satisfies the linear partial differential equation, $\mathcal{L}(m)u(m) = f$, with the force, f, independent of m, the chain rule gives the wavefield derivative

$$\frac{\partial \mathcal{L}}{\partial m} u(m) + \mathcal{L}(m) \frac{\partial u}{\partial m} = 0 \quad \text{or} \quad \frac{\partial u}{\partial m} = -\mathcal{L}^{-1} \frac{\partial \mathcal{L}}{\partial m} u(m)$$

(5)

The error derivative evaluated at reference model, m_0, is

$$\left.\frac{\partial E}{\partial m}\right|_{m_0} = 2 \left. \left(\mathcal{L}^{-1} \frac{\partial \mathcal{L}}{\partial m} u, s(\mathbf{x}, t) \right)_{\mathbf{x},t} \right|_{m_0} = 2 \left(\frac{\partial \mathcal{L}_0}{\partial m} u_0, \mathcal{L}_0^{\dagger -1} s(\mathbf{x}, t) \right)_{\mathbf{x},t}$$

(6)

Here we have used the shorthand, $u_0 = u(m_0)$, $\mathcal{L}_0 = (m_0)$, and $\partial \mathcal{L}_0 / \partial m \equiv \partial \mathcal{L} / \partial m|_{m_0}$ and the fact that $\mathcal{L}^{-1\dagger} = \mathcal{L}^{\dagger -1}$. This result can be simplified by introducing an adjoint field, λ, and a function, ξ, defined as

$$\xi \equiv 2 \frac{\partial \mathcal{L}_0}{\partial m} u_0 \quad \text{and} \quad \lambda = \mathcal{L}_0^{\dagger -1} s(\mathbf{x}, t)$$

(7)

From whence it follows that

$$\left.\frac{\partial E}{\partial m}\right|_{m_0} = (\xi, \lambda)_{\mathbf{x},t} \quad \text{with} \quad \mathcal{L}_0^{\dagger} \lambda = s(\mathbf{x}, t)$$

(8)

Note that the adjoint field satisfies a differential equation with a source, $s(\mathbf{x}, t)$, that contains a Dirac comb (sum of delta functions).

So, the upshot is that the delta functions in the adjoint source arise due to a mathematical manipulation that adapts the case of sparce observations (observations only at points, $\mathbf{x}^{(i)}$) to a method based on inner products over spatial coordinates.