1 Motivation

2 Nonlinear optimization using gradient based methods

3 Calculation of the gradients for elastic full waveform tomography
Motivation

Resolution: Traveltime Tomography

Traveltime Tomography

True Marmousi-2 Model

Depth [km]
Distance [km]

Vp [m/s]
Motivation of Full Waveform Inversion

First arrival traveltime picks vs. full wavefield

Marmousi-II data shot no. 1
Motivation

Resolution: Traveltime Tomography

![Traveltime Tomography](image1)

**True Marmousi-2 Model**

- **Distance [km]**
- **Depth [km]**
  - 0.5
  - 1
  - 1.5
  - 2
  - 2.5
  - 3

- **Vp [m/s]**
  - 1000
  - 1500
  - 2000
  - 2500
  - 3000
  - 3500
  - 4000
  - 4500
Motivation

Resolution: Waveform Tomography

Waveform Tomography

True Marmousi-2 Model

Distance [km]
Depth [km]

1 2 3 4 5 6 7 8 9 10

0.5
1
1.5
2
2.5
3

Waveform Tomography

V_p [m/s]

1000
1500
2000
2500
3000
3500
4000
4500

Distance [km]
Depth [km]

1 2 3 4 5 6 7 8 9 10

0.5
1
1.5
2
2.5
3

Waveform Tomography

V_p [m/s]

1000
1500
2000
2500
3000
3500
4000
4500
Motivation: Estimate optimum model from data

Seismic Section

Waveform Tomography

Problems

1. What is an "optimum" model?
2. How can this model be found?
3. Is this model unique or are other models existing, which could explain the data equally well?
What is an "optimum" model?

Definition: data residuals and misfit function

Data residuals: \( \delta u = u^{\text{mod}} - u^{\text{obs}} \)

The L2-norm (Residual Energy) of the data residuals: \( E = \frac{1}{2} \delta u^T \delta u \)
How to find an optimum model?

Simple example for non-linear optimization

Let's pick a nice colorful non-linear misfit function with known minima from

https://en.wikipedia.org/wiki/Test_functions_for_optimization

<table>
<thead>
<tr>
<th>Name</th>
<th>Plot</th>
<th>Formula</th>
<th>Global minimum</th>
<th>Search domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rastrigin function</td>
<td><img src="image1" alt="Plot" /></td>
<td>$f(x) = A_n + \sum_{i=1}^{n} [x_i^2 - A \cos(2\pi x_i)]$ where: $A = 10$</td>
<td>$f(0, 0) = 0$</td>
<td>$-5.12 \leq x, y \leq 5.12$</td>
</tr>
<tr>
<td>Ackley's function</td>
<td><img src="image2" alt="Plot" /></td>
<td>$f(x, y) = -20 \exp\left[-0.2 \sqrt{0.5 (x^2 + y^2)}\right] - \exp\left[0.5 (\cos(2\pi x) + \cos(2\pi y))\right] + e + 20$</td>
<td>$f(0, 0) = 0$</td>
<td>$-5 \leq x, y \leq 5$</td>
</tr>
<tr>
<td>Sphere function</td>
<td><img src="image3" alt="Plot" /></td>
<td>$f(x) = \sum_{i=1}^{n} x_i^2$</td>
<td>$f(x_1, \ldots, x_n) = f(0, \ldots, 0) = 0$</td>
<td>$-\infty \leq x_i \leq \infty, 1 \leq i \leq n$</td>
</tr>
<tr>
<td>Rosenbrock function</td>
<td><img src="image4" alt="Plot" /></td>
<td>$f(x) = \sum_{i=1}^{n-1} 100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2$</td>
<td>Min = $\begin{cases} n = 2 &amp; \rightarrow f(1,1) = 0, \ n = 3 &amp; \rightarrow f(1,1,1) = 0, \ n &gt; 3 &amp; \rightarrow f(1,\ldots,1) = 0 \text{ in linear} \end{cases}$</td>
<td>$-\infty \leq x_i \leq \infty, 1 \leq i \leq n$</td>
</tr>
<tr>
<td>Beale's function</td>
<td><img src="image5" alt="Plot" /></td>
<td>$f(x, y) = (1.5 - x + xy)^2 + (2.25 - x + xy)^2 + (2.625 - x + xy)^2$</td>
<td>$f(3,0.5) = 0$</td>
<td>$-4.5 \leq x, y \leq 4.5$</td>
</tr>
<tr>
<td>Goldstein–Price function</td>
<td><img src="image6" alt="Plot" /></td>
<td>$f(x, y) = [1 + (x + y + 1)^2 (19 - 14x + 3x^2 - 14y + 6xy + 3y^2)]$</td>
<td>$f(0, -1) = 3$</td>
<td>$-2 \leq x, y \leq 2$</td>
</tr>
</tbody>
</table>
How to find an optimum model?

Simple example for non-linear optimization

... for example the 2D Rosenbrock function:

\[ E = (1 - x_1)^2 + 100(x_2 - x_1^2)^2 \]

To remind you that the aim is to optimize material parameters of the underground I substitute:

\[ x_1 \rightarrow \text{P-wave velocity } V_p \]
\[ x_2 \rightarrow \text{density } \rho \]

\[ E = (1 - V_p)^2 + 100(\rho - V_p^2)^2 \]

Keep in mind though that this is geophysical non-sense.
How to find an optimum model?

Parameter space

\[ E = (1 - Vp)^2 + 100(\rho - Vp^2)^2 \]  [Rosenbrock, 1960]
How to find an optimum model?

Minimize misfit function $E$ to find the "optimum" model.
How to find an optimum model?

Define starting point in the parameter space

Start search with an initial guess $m_1$. 
How to find an optimum model?

Iterative model update

Update the model iteratively: $m_2 = m_1 + \mu_1 \delta m_1$. $\delta m_1$ denotes the direction and $\mu_1$ the steplength.
How to find an optimum model?

Iterative model update

Which direction $\delta m_1$ should be choosen?
How to find an optimum model?

Find optimum search direction ...

Taylor series expansion:

\[
E(m_1 + \delta m_1) \approx E(m_1) + \delta m_1 \left( \frac{\partial E}{\partial m} \right)_1 + \frac{1}{2} \delta m_1^T H_1 \delta m_1
\]

with the Hessian \( H_{ij} = \frac{\partial^2 E}{\partial m_i \partial m_j} \). Set derivative to zero:

\[
\frac{\partial E(m_1 + \delta m_1)}{\partial \delta m_1} = \left( \frac{\partial E}{\partial m} \right)_1 + \delta m_1 H_1 = 0
\]

... which leads to:

\[
\delta m_1 = -H_1^{-1} \left( \frac{\partial E}{\partial m} \right)_1
\]

where \((\partial E/\partial m)_1\) denotes the gradient direction of the objective function and \(H_1^{-1}\) the inverse Hessian matrix.
How to find an optimum model?

Newton Method

Newton method: \[ m_{n+1} = m_n - \mu_n H_n^{-1} \left( \frac{\partial E}{\partial m} \right)_n \]
How to find an optimum model?

**Gradient Method**

Gradient method: \( m_{n+1} = m_n - \mu_n P_n \left( \frac{\partial E}{\partial m} \right)_n \)
Gradient calculation by the adjoint state method

- Adjoint state gradient estimation using Lagrange multipliers [Plessix, 2006]
- Adjoint state gradient estimation in the frequency domain [Pratt et al., 1998]
- Adjoint state gradient estimation using perturbation theory [Tarantola, 2005, Mora, 1987]
Rewrite misfit function

\[ E = \frac{1}{2} \delta \mathbf{u}^T \delta \mathbf{u} = \frac{1}{2} \sum_{\text{sources}} \int \delta \mathbf{u}^2(x_r, x_s, t) \sum_{\text{receiver}} \delta \mathbf{u} \]
Gradient

To estimate the gradient direction $\partial E / \partial \mathbf{m}$ the residual energy is rewritten as:

$$E = \frac{1}{2} \delta \mathbf{u}^T \delta \mathbf{u} = \frac{1}{2} \sum_{\text{sources}} \int \, dt \sum_{\text{receiver}} \delta \mathbf{u}^2(x_r, x_s, t)$$

After derivation with respect to a model parameter $\mathbf{m}$ we get

$$\frac{\partial E}{\partial \mathbf{m}} = \sum_{\text{sources}} \int \, dt \sum_{\text{receiver}} \frac{\partial \delta \mathbf{u}}{\partial \mathbf{m}} \delta \mathbf{u}$$

$$= \sum_{\text{sources}} \int \, dt \sum_{\text{receiver}} \frac{\partial (\mathbf{u}^{\text{mod}}(\mathbf{m}) - \mathbf{u}^{\text{obs}})}{\partial \mathbf{m}} \delta \mathbf{u}$$

$$= \sum_{\text{sources}} \int \, dt \sum_{\text{receiver}} \frac{\partial \mathbf{u}^{\text{mod}}(\mathbf{m})}{\partial \mathbf{m}} \delta \mathbf{u}$$
Calculation of the gradient direction $\frac{\partial E}{\partial m}$

Mapping model space $\rightarrow$ data space (forward problem)
Calculation of the gradient direction $\frac{\partial E}{\partial m}$

**Mapping model space $\rightarrow$ data space (forward problem)**

Taylor series:

$$\tilde{u}(m + \delta m) \approx \tilde{u}(m) + \frac{\partial u}{\partial m} \delta m + O(\delta m^2)$$
Calculation of the gradient direction $\frac{\partial E}{\partial m}$

Mapping model space $\rightarrow$ data space (forward problem)

$$
\delta \tilde{u}(m + \delta m) - \tilde{u}(m) = \frac{\partial u}{\partial m} \delta m
$$
Calculation of the gradient direction \( \frac{\partial E}{\partial m} \)

Mapping model space \( \rightarrow \) data space (forward problem)

\[ \delta \tilde{u}(x_s, x_r, t) = \frac{\partial u}{\partial m} \delta m \]

sensitivity kernel
Calculation of the gradient direction \( \frac{\partial E}{\partial m} \)

**Mapping model space \( \rightarrow \) data space (forward problem)**

\[
\delta \tilde{u}(x_s, x_r, t) = \int_V dV \frac{\partial u}{\partial m} \delta m
\]

- **Data Space**
  - \( x \)-coordinate \( \rightarrow \)
  - \( \delta \tilde{u} \)
  - \( \leftarrow \) time

- **Model Space**
  - \( x \)-coordinate \( \rightarrow \)
  - \( \delta m \)
  - \( \leftarrow \) depth
Calculation of the gradient direction $\frac{\partial E}{\partial m}$

Mapping data space $\rightarrow$ model space (inverse problem)

Data Space $\rightarrow$ Model Space

$\delta m' = \sum \int dt \sum_{\text{receiver}} \frac{\partial m}{\partial u} \delta \tilde{u}'$

$\delta \tilde{u}'$
Example: Fourier transform

For a given time-dependent function $f(t)$, the forward operator is defined as:

$$\mathcal{F}\{f(t)\}(\omega) = \int_{-\infty}^{\infty} f(t) \exp(-i\omega t) dt$$

For the frequency-dependent function $f(\omega)$, the inverse operator is defined as:

$$\mathcal{F}^{-1}\{f(\omega)\}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\omega) \exp(i\omega t) d\omega$$

with the circular frequency $\omega$ and $i^2 = -1$. 
Exercise 1

Prove that the Fourier transform

\[ \mathcal{F}\{f(t)\}(\omega) = \int_{-\infty}^{\infty} f(t) \exp(-i\omega t) dt \]

is a linear operator:

\[ \mathcal{F}\{af(t)\} = a\mathcal{F}\{f(t)\} \]
\[ \mathcal{F}\{f(t) + g(t)\} = \mathcal{F}\{f(t)\} + \mathcal{F}\{g(t)\} \]

for time-dependent functions \( f(t) \), \( g(t) \) and \( a \in \mathbb{R} \).
Kernels of the Fourier transform and its inverse

Notice that the kernels of the Fourier transform
\( K_{\mathcal{F}} = \exp(-i\omega t) \)
and the inverse transform
\( K_{\mathcal{F}^{-1}} = \exp(i\omega t) \)

\[
\mathcal{F}\{f(t)\}(\omega) = \int_{-\infty}^{\infty} f(t)K_{\mathcal{F}}\,dt
\]

\[
\mathcal{F}^{-1}\{f(\omega)\}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\omega)K_{\mathcal{F}^{-1}}\,d\omega
\]

are related by
\( K_{\mathcal{F}^{-1}} = K^*_{\mathcal{F}} \),
where * denotes the complex conjugate.
This property can also be used for the FWI problem.
Calculation of the gradient direction $\frac{\partial E}{\partial m}$

Mapping data space $\rightarrow$ model space (inverse problem)

$$\delta m' = \sum_{\text{sources}} \int \text{dt} \sum_{\text{receiver}} \frac{\partial m}{\partial u} \delta \tilde{u}' = \sum_{\text{sources}} \int \text{dt} \sum_{\text{receiver}} \left[ \frac{\partial u}{\partial m} \right]^* \delta \tilde{u}'$$
Properties of linear operators

By introducing the linear operator $\hat{L}$ the integrals can be written as:

$$\delta \tilde{u} = \hat{L} \delta m := \int_V dV \frac{\partial u}{\partial m} \delta m.$$ 

and

$$\delta m' = \hat{L}^* \delta \tilde{u}' := \sum_{\text{sources}} \int dt \sum_{\text{receiver}} \left[ \frac{\partial u}{\partial m} \right]^* \delta \tilde{u}'.$$

Because the operator $\hat{L}$ is linear, the kernel of $\hat{L}$ and its adjoint counterpart $\hat{L}^*$ are identical (see chapter 5.4.2 in [Tarantola, 2005])

$$\left[ \frac{\partial u}{\partial m} \right]^* = \left[ \frac{\partial u}{\partial m} \right].$$
\[ \delta m' = \frac{\partial E}{\partial m} \]

Therefore the mapping from the data to the model space is equal to the gradient of the residual energy:

\[ \delta m' = \sum_{\text{sources}} \int dt \sum_{\text{receiver}} \left[ \frac{\partial u_i}{\partial m} \right]^* \delta \tilde{u}' \]

\[ = \sum_{\text{sources}} \int dt \sum_{\text{receiver}} \left[ \frac{\partial u_i}{\partial m} \right] \delta u \]

\[ = \frac{\partial E}{\partial m} \]

if the perturbation of the data space \( \delta \tilde{u}' \) is interpreted as data residuals \( \delta u \).
Calculation of the gradient direction $\frac{\partial E}{\partial m}$

General approach to estimate the gradient via adjoint state method

So the approach to estimate the gradient direction $\frac{\partial E}{\partial m}$ can be split into 3 parts

1. Find a solution to the forward problem
   \[ \delta u = \hat{L} \delta m. \]

2. Identify the Frechét kernels $\frac{\partial u}{\partial m}$

3. Use the property, that a linear operator $\hat{L}$ and its adjoint $\hat{L}^*$ have the same kernels and calculate the gradient direction by using:
   \[ \frac{\partial E}{\partial m} = \delta m' = \hat{L}^* \delta u'. \]
Possible applications

This approach is very general and can be applied to a wide range of physical problems ...

- Quantum Mechanics: Schrödinger's Equation
- Hydrodynamics: Oceanography, Meteorology
- Electrodynamics: GPR
- Medical Diagnostics: Acoustic Wave Equation
- Magnetohydrodynamics: Astrophysics, Plasmaphysics
- General Relativity: Source Parameters of Gravity Waves

... if the problem is linear.

Unfortunately, the interesting problems are nonlinear 😞
Calculation of the gradient direction $\frac{\partial E}{\partial m}$

A short side note on FWI of gravity waves ...

from B.P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration) 2016,
Elastic equations of motion for anisotropic media

\[ \rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial}{\partial x_j} \sigma_{ij} = f_i, \]

\[ \sigma_{ij} - c_{ijkl} \epsilon_{kl} = T_{ij}, \]

\[ \epsilon_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \]

+ initial and boundary conditions,

where \( \rho \) denotes the density, \( u_i \) the displacement, \( \sigma_{ij} \) the stress tensor, \( \epsilon_{ij} \) the strain tensor, \( c_{ijkl} \) the stiffness tensor, \( f_i, T_{ij} \) source terms for volume and surface forces, respectively.
First order perturbations

In the next step every parameter and variable in the elastic wave equation is perturbated by a first order perturbation:

\[ u_i \rightarrow u_i + \delta u_i, \]
\[ \sigma_{ij} \rightarrow \sigma_{ij} + \delta \sigma_{ij} \]
\[ \epsilon_{ij} \rightarrow \epsilon_{ij} + \delta \epsilon_{ij} \]
\[ \rho \rightarrow \rho + \delta \rho \]
\[ c_{ijkl} \rightarrow c_{ijkl} + \delta c_{ijkl} \]
Calculation of the gradient direction $\frac{\partial E}{\partial m}$

Perturbed elastic equations of motion

$$
\rho \frac{\partial^2 \delta u_i}{\partial t^2} - \frac{\partial}{\partial x_j} \delta \sigma_{ij} = \Delta f_i
$$

$$
\delta \sigma_{ij} - c_{ijkl} \delta \epsilon_{kl} = \Delta T_{ij}
$$

$$
\delta \epsilon_{ij} = \frac{1}{2} \left( \frac{\partial \delta u_i}{\partial x_j} + \frac{\partial \delta u_j}{\partial x_i} \right)
$$

+ perturbated initial and boundary conditions

The new source terms are

$$
\Delta f_i = -\delta \rho \frac{\partial^2 u_i}{\partial t^2}, \quad \Delta T_{ij} = \delta c_{ijkl} \epsilon_{kl}.
$$
Calculation of the gradient direction $\frac{\partial E}{\partial m}$

Pertubated elastic equations of motion

Data Space

Model Space

$x$-coordinate $\rightarrow$

$\delta \tilde{u}$

$\delta m$

$\leftarrow$ time

$\leftarrow$ depth

$x$-coordinate $\rightarrow$
Calculation of the gradient direction $\frac{\partial E}{\partial m}$

**Definition: Green’s function**

If a unit impulse is applied as a source term at $x = x'$ at time $t = t'$ in the n-direction, then we denote the $i$th component of the displacement field at any point $(x, t)$ as Green’s function $G_{in}(x, t; x', t')$ ([Aki and Richards, 1980]).

$$\rho \frac{\partial^2 G_{in}}{\partial t^2} - \frac{\partial \sigma_{ij}}{\partial x_j} = \delta_{in} \delta(x - x') \delta(t - t')$$

$$\sigma_{ij} = c_{ijkl} \epsilon_{kl}.$$
Solution of the perturbated wave equation in terms of Green’s function

The solution of the perturbated elastic equations of motion in terms of the elastic Green’s function $G_{ij}(\mathbf{x}, t; \mathbf{x}', t')$ can be written as:

$$
\delta u_i(\mathbf{x}, t) = \int_V dV \int_0^T dt' G_{ij}(\mathbf{x}, t; \mathbf{x}', t') \Delta f_j(\mathbf{x}', t') - \int_V dV \int_0^T dt' \frac{\partial G_{ij}}{\partial x'_k}(\mathbf{x}, t; \mathbf{x}', t') \Delta T_{jk}(\mathbf{x}', t').
$$
Simple example: Green’s function

Greens function $G_{ij}(x,t;x',t')$

Source wavelet $f_j$

Seismogram $u_i$
Solution of the perturbed wave equation in terms of Green’s function

The solution of the perturbated elastic equations of motion in terms of the elastic Green’s function $G_{ij}(x, t; x', t')$ can be written as:

\[
\delta u_i(x, t) = \int_V dV \int_0^T dt' G_{ij}(x, t; x', t') \Delta f_j(x', t') - \int_V dV \int_0^T dt' \frac{\partial G_{ij}}{\partial x_k'}(x, t; x', t') \Delta T_{jk}(x', t').
\]

The new source terms are

\[
\Delta f_j = -\delta \rho \frac{\partial^2 u_j}{\partial t^2}, \quad \Delta T_{jk} = \delta c_{jklm} \epsilon_{lm}.
\]
Substitute source terms of the perturbated equations of motion

Substituting the force and traction source terms yields after some rearranging

\[ \delta u_i(x, t) = - \int_V dV \int_0^T dt' G_{ij}(x, t; x', t') \frac{\partial^2 u_j}{\partial t^2}(x', t') \delta \rho - \int_V dV \int_0^T dt' \frac{\partial G_{ij}}{\partial x'_k}(x, t; x', t') \epsilon_{lm}(x', t') \delta c_{jklm} \]
Calculation of the gradient direction $\frac{\partial E}{\partial m}$

**Born approximation**

Introducing isotropy via

$$\delta c_{jklm} = \delta_{jk} \delta_{lm} \delta \lambda + (\delta_{jl} \delta_{km} + \delta_{jm} \delta_{kl}) \delta \mu$$

leads to:

$$\delta u_i(x, t) = - \int_V dV \left[ \int_0^T dt' G_{ij}(x, t; x', t') \frac{\partial^2 u_j}{\partial t^2}(x', t') \right] \delta \rho$$

$$- \int_V dV \left[ \int_0^T dt' \frac{\partial G_{ij}}{\partial x'_k} (x, t; x', t') \epsilon_{lm}(x', t') \delta_{jk} \delta_{lm} \right] \delta \lambda$$

$$- \int_V dV \left[ \int_0^T dt' \frac{\partial G_{ij}}{\partial x'_k} (x, t; x', t') \epsilon_{lm}(x', t') (\delta_{jl} \delta_{km} + \delta_{jm} \delta_{kl}) \right] \delta \mu.$$  

This equation has the same form as the desired expression for the forward problem:

$$\delta u = \int_V dV \frac{\partial u}{\partial m} \delta m.$$
Calculate the gradient direction $\frac{\partial E}{\partial m}$

Identify Frechét kernels

Therefore the Frechét kernels $\frac{\partial u_i}{\partial m(x)}$ for the individual material parameters can be identified as:

$$
\frac{\partial u_i}{\partial \rho} = - \int_0^T dt' G_{ij}(x, t; x', t') \frac{\partial^2 u_j}{\partial t^2}(x', t')
$$

$$
\frac{\partial u_i}{\partial \lambda} = - \int_0^T dt' \frac{\partial G_{ij}}{\partial x'_k}(x, t; x', t') \epsilon_{lm}(x', t') \delta_{jk} \delta_{lm}
$$

$$
\frac{\partial u_i}{\partial \mu} = - \int_0^T dt' \frac{\partial G_{ij}}{\partial x'_k}(x, t; x', t') \epsilon_{lm}(x', t')(\delta_{jl} \delta_{km} + \delta_{jm} \delta_{kl})
$$
Definition of the adjoint operator and gradient

By definition the adjoint of the operator can be written as

\[
\frac{\partial E}{\partial m} = \delta m'(x) = \sum_{\text{sources}} \int_0^T \sum_{\alpha=1}^{N_{\text{rec}}} \left[ \frac{\partial u_i}{\partial m} \right]^* \delta u_i'(x_\alpha, t'),
\]
Calculate adjoint operators

Because a linear operator and its transpose have the same kernels \( \partial u_i / \partial m \), the only difference arise in the variables of sum/integration, which are complementary. Inserting the integral kernels yields

\[
\frac{\partial E}{\partial \rho} = - \sum_{\text{sources}} \int_0^T \, dt \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T \, dt' G_{ij}(x_\alpha, t'; x, t) \frac{\partial^2 u_j}{\partial t^2}(x, t) \delta u'_i(x_\alpha, t')
\]

\[
\frac{\partial E}{\partial \lambda} = - \sum_{\text{sources}} \int_0^T \, dt \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T \, dt' \frac{\partial G_{ij}}{\partial x_k}(x_\alpha, t'; x, t) \epsilon_{lm}(x, t) \delta_{jk} \delta_{lm} \delta u'_i(x_\alpha, t')
\]

\[
\frac{\partial E}{\partial \mu} = - \sum_{\text{sources}} \int_0^T \, dt \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T \, dt' \frac{\partial G_{ij}}{\partial x_k}(x_\alpha, t'; x, t) \epsilon_{lm}(x, t)(\delta_{jl} \delta_{km} + \delta_{jm} \delta_{kl}) \delta u'_i(x_\alpha, t')
\]
Some rearrangements ...

The terms only depending on time $t$ and the positions $x$ can be moved in front of the sum over the receivers

\[
\frac{\partial E}{\partial \rho} = - \sum_{\text{sources}} \int_0^T \text{d}t \frac{\partial^2 u_j}{\partial t^2} (x, t) \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T \text{d}t' G_{ij}(x_\alpha, t'; x, t) \delta u'_i(x_\alpha, t'),
\]

\[
\frac{\partial E}{\partial \lambda} = - \sum_{\text{sources}} \int_0^T \text{d}t \epsilon_{lm}(x, t) \delta_{jk} \delta_{lm} \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T \text{d}t' \frac{\partial G_{ij}}{\partial x_k}(x_\alpha, t'; x, t) \delta u'_i(x_\alpha, t'),
\]

\[
\frac{\partial E}{\partial \mu} = - \sum_{\text{sources}} \int_0^T \text{d}t \epsilon_{lm}(x, t) (\delta_{jl} \delta_{km} + \delta_{jm} \delta_{kl}) \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T \text{d}t' \frac{\partial G_{ij}}{\partial x_k}(x_\alpha, t'; x, t) \delta u'_i(x_\alpha, t').
\]
Calculation of the gradient direction $\frac{\partial E}{\partial m}$

Introducing the wavefield $\Psi_j$

Defining the wavefield

$$\Psi_j(x, t) = \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T \text{d}t' G_{ij}(x_\alpha, t'; x, t) \delta u'_i(x_\alpha, t'),$$

yields

$$\frac{\partial E}{\partial \rho} = - \sum_{\text{sources}} \int_0^T \text{d}t \frac{\partial^2 u_j}{\partial t^2}(x, t) \Psi_j,$$

$$\frac{\partial E}{\partial \lambda} = - \sum_{\text{sources}} \int_0^T \text{d}t \epsilon_{lm}(x, t) \delta_{jk} \delta_{lm} \frac{\partial \Psi_j}{\partial x_k},$$

$$\frac{\partial E}{\partial \mu} = - \sum_{\text{sources}} \int_0^T \text{d}t \epsilon_{lm}(x, t) (\delta_{jl} \delta_{km} + \delta_{jm} \delta_{kl}) \frac{\partial \Psi_j}{\partial x_k}.$$
Calculate implicit sums ...

Writing out the implicit sums in the gradients of the Lamé parameters $\delta \lambda'$ and $\delta \mu'$

\[
\frac{\partial E}{\partial \lambda} = - \sum_{\text{sources}} \int_0^T dt \sum_{l} \sum_{k} \sum_{j} \sum_{m} \epsilon_{lm}(x, t) \delta_{jk} \delta_{lm} \frac{\partial \Psi_j}{\partial x_k},
\]

\[
\frac{\partial E}{\partial \mu} = - \sum_{\text{sources}} \int_0^T dt \sum_{l} \sum_{k} \sum_{j} \sum_{m} \epsilon_{lm}(x, t) (\delta_{jl} \delta_{km} + \delta_{jm} \delta_{kl}) \frac{\partial \Psi_j}{\partial x_k}.
\]
Neglecting all wavefield components and derivatives in z-direction leads to

\[
\frac{\partial E}{\partial \lambda} = - \sum_{\text{sources}} \int_{0}^{T} dt \left( \epsilon_{xx} + \epsilon_{yy} \right) \left( \frac{\partial \Psi_x}{\partial x} + \frac{\partial \Psi_y}{\partial y} \right),
\]

\[
\frac{\partial E}{\partial \mu} = - \sum_{\text{sources}} \int_{0}^{T} dt \left[ \left( \epsilon_{xy} + \epsilon_{yx} \right) \left( \frac{\partial \Psi_x}{\partial y} + \frac{\partial \Psi_y}{\partial x} \right) \right]
+ 2 \left( \epsilon_{xx} \frac{\partial \Psi_x}{\partial x} + \epsilon_{yy} \frac{\partial \Psi_y}{\partial y} \right).
\]
Introducing the strain tensor

Using the definition of the strain tensor $\epsilon_{ij}$ we get

$$\frac{\partial E}{\partial \lambda} = - \sum_{\text{sources}} \int_0^T dt \left( \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} \right) \left( \frac{\partial \Psi_x}{\partial x} + \frac{\partial \Psi_y}{\partial y} \right),$$

$$\frac{\partial E}{\partial \mu} = - \sum_{\text{sources}} \int_0^T dt \left[ \left( \frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \right) \left( \frac{\partial \Psi_x}{\partial y} + \frac{\partial \Psi_y}{\partial x} \right) \right] + 2 \left( \frac{\partial u_x}{\partial x} \frac{\partial \Psi_x}{\partial x} + \frac{\partial u_y}{\partial y} \frac{\partial \Psi_y}{\partial y} \right).$$
Calculation of the gradient direction $\frac{\partial E}{\partial m}$

**Final gradients**

Finally the gradients for the Lamé parameters $\lambda$, $\mu$ and the density $\rho$ can be written as

\[
\frac{\partial E}{\partial \lambda} = - \sum_{\text{sources}} \int dt \left( \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} \right) \left( \frac{\partial \Psi_x}{\partial x} + \frac{\partial \Psi_y}{\partial y} \right)
\]

\[
\frac{\partial E}{\partial \mu} = - \sum_{\text{sources}} \int dt \left( \frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \right) \left( \frac{\partial \Psi_x}{\partial y} + \frac{\partial \Psi_y}{\partial x} \right) + 2 \left( \frac{\partial u_x}{\partial x} \frac{\partial \Psi_x}{\partial x} + \frac{\partial u_y}{\partial y} \frac{\partial \Psi_y}{\partial y} \right)
\]

\[
\frac{\partial E}{\partial \rho} = \sum_{\text{sources}} \int dt \left( \frac{\partial^2 u_x}{\partial t^2} \Psi_x + \frac{\partial^2 u_y}{\partial t^2} \Psi_y \right)
\]
For each shot solve the forward problem for the actual model $m_n$ to generate a synthetic dataset $u_{\text{mod}}$ and the wavefield $u_{\text{mod}}(x, t)$.

2. Calculate the residual seismograms $\delta u = u_{\text{mod}} - u_{\text{obs}}$.

3. Generate the wavefield $\Psi(x, t)$ by backpropagating the residuals from the receiver positions.

4. Calculate the gradients $\frac{\partial E}{\partial m}$ for each material parameter.

5. Estimate the step length $\mu_n$ by a line search.

6. Update the material parameters using the gradient method

$$m_{n+1} = m_n - \mu_n P_n \left( \frac{\partial E}{\partial m} \right)_n$$
Teaser for the next lecture …

Modified elastic Marmousi-II model: Full Waveform Inversion

Click here for cool FWI movie
*Quantitative seismology.*  
W.H. Freeman and Company.

Nonlinear two-dimensional elastic inversion of multioffset seismic data.  
*Geophysics, 52*:1211 – 1228.

A review of the adjoint-state method for computing the gradient of a functional with geophysical applications.  

Gauss-Newton and full Newton methods in frequency domain seismic waveform inversion.  

An automatic method for finding the greatest or least value of a function.  

*Inverse Problem Theory.*  
SIAM.