Adjoint state method tutorial
M.J. Yedlin and D. Van Vorst
Pratt's Inversion Method
Adjoint State Method
a) Take Variation of $J(u(p))$
b) Add Zero – Introduce the Adjoint State Variable
c) Comparison with Pratt
Conclusions
Acknowledgements

UBC   CREWES

Tutorial on the continuous and discrete adjoint state method and basic implementation

M.J. Yedlin and D. Van Vorst

Dec. 3, 2010

THANK YOU!

Adjoint state method tutorial

M.J. Yedlin and D. Van Vorst

Pratt's Inversion Method

Adjoint State Method

a) Take Variation of $J(u(p))$

b) Add Zero – Introduce the Adjoint State Variable

c) Comparison with Pratt

Conclusions

Acknowledgements

I would like to thank Dr. Gary Margrave, with the support of the CREWES membership, for inviting me to be an Associate Faculty member of CREWES.

Outline

1 Pratt's Inversion Method

2 Adjoint State Method
- a) Take Variation of $J(u(p))$
- b) Add Zero – Introduce the Adjoint State Variable
- c) Comparison with Pratt

3 Conclusions

4 Acknowledgements

Steps in Pratt's Inversion Method

Adjoint state method tutorial
M.J. Yedlin and D. Van Vorst
Pratt's Inversion Method
Adjoint State Method
a) Take Variation of $J(u(p))$
b) Add Zero – Introduce the Adjoint State Variable
c) Comparison with Pratt
Conclusions
Acknowledgements

Pratt, R.G., 1999. Seismic waveform inversion in the frequency domain, Part 1: Theory, and verification in a physical scale model. Geophysics, 64, 888-901.

1. Forward Modelling
2. Define Cost Functional
3. Compute the Gradient of the Cost Functional
4. Compute the Fréchet Matrix
5. Assemble the Fréchet Matrix $\rightarrow \nabla_{\mathbf{p}} E(\mathbf{p})$
6. Interpretation of $\nabla_{\mathbf{p}} E(\mathbf{p})$

1 Forward Modelling

We start with the forward problem, which is now in discrete form and includes all boundary conditions. It is represented, in the frequency domain, after temporal Fourier transform of the original partial differential equation, as a linear system of equations given by

$$\mathbf{L}(\omega)\mathbf{s}(\omega) = \mathbf{f}(\omega) \tag{1}$$

$\mathbf{L}(\omega) = =$ Linear operator [PDE] ($l \times l$ matrix)
$\mathbf{s}(\omega) =$ State vector – displacements – ($l \times 1$ vector)
$\mathbf{f}(\omega) =$ All sources
Control parameters (model parameters): $\mathbf{p}$ ($m \times 1$ vector)

2 Define Cost Functional – Misfit

Adjoint state method tutorial

M.J. Yedlin and D. Van Vorst

Pratt's Inversion Method

Adjoint State Method

a) Take Variation of $J(u(p))$

b) Add Zero – Introduce the Adjoint State Variable

c) Comparison with Pratt

Conclusions

Acknowledgements

The squared sum of the data residual vector is the norm of the complex data residual $\delta\mathbf{d}(\omega)$, the difference between simulated data, $\mathbf{s}(\omega)$, and recorded data, $\mathbf{d}(\omega)$, denoted as a vector, $\mathbf{s}(\omega) - \mathbf{d}(\omega)$, on a sub-domain of n receiver positions:

$$E(\mathbf{p}) = \frac{1}{2}\delta\mathbf{d}(\omega)^T\delta\mathbf{d}^*(\omega), \qquad (2)$$

with the usual notation of superscript T for transpose and $*$ for complex conjugation, and with the misfit depending on the control vector, $\mathbf{p}$.

Redefintion of Cost Functional

For computational purposes, Pratt redefines the residual at all points of the grid, l padding the residual with $l - n$ zeros. Denote this residual by the ˆ symbol. Thus

$$E(\mathbf{p}) = \frac{1}{2}\delta\hat{\mathbf{d}}(\omega)^T \delta\hat{\mathbf{d}}^*(\omega), \tag{3}$$

Notations have changed since this 1999 paper. Now this is done via a projection operator.

Now compute the gradient of $E(\mathbf{p})$, which will be used in a model update.

Adjoint state method tutorial
M.J. Yedlin and D. Van Vorst
Pratt's Inversion Method
Adjoint State Method
a) Take Variation of $J(u(p))$
b) Add Zero – Introduce the Adjoint State Variable
c) Comparison with Pratt
Conclusions
Acknowledgements

We now compute the gradient of the cost functional:

$$\nabla_{\mathbf{p}} E(\mathbf{p}) = \Re eal\{\hat{\mathbf{J}}^T \delta \hat{\mathbf{d}}^*(\omega)\}. \tag{4}$$

where

$$\hat{J}_{ij} = \frac{\partial s_i}{\partial p_j} \qquad i = 1 \ldots l \text{ and } j = 1 \ldots m \tag{5}$$

$\hat{\mathbf{J}}$ is commonly known as the Fréchet derivative matrix or sensitivity matrix.

We now have a big problem! We have to solve a forward problem for each of p_j, in order to do a model update at each iteration: $\mathbf{p}^{(k+1)} = \mathbf{p}^{(k)} + \alpha \nabla_{\mathbf{p}} E(\mathbf{p})$, for some step size, α.

4 Compute the Fréchet Matrix

Adjoint state method tutorial

M.J. Yedlin and D. Van Vorst

Pratt's Inversion Method

Adjoint State Method

a) Take Variation of $J(u(p))$

b) Add Zero – Introduce the Adjoint State Variable

c) Comparison with Pratt

Conclusions

Acknowledgements

We differentiate both sides of (1), assuming that the sources are not dependent on **p** and obtain

$$\frac{\partial}{\partial p_i}\left(\mathbf{L}(\omega)\mathbf{s}(\omega)\right) = 0. \tag{6}$$

which becomes

$$\mathbf{L}(\omega)\frac{\partial \mathbf{s}(\omega)}{\partial p_i} = -\frac{\partial \mathbf{L}(\omega)}{\partial p_i}\mathbf{s}(\omega) \tag{7}$$

and solve (7) for the elements of $\hat{\mathbf{J}}$ in terms of virtual sources $\mathbf{f}^{(i)}_{virt}(\omega)$ to obtain

$$\frac{\partial \mathbf{s}(\omega)}{\partial p_i} = -\mathbf{L}^{-1}(\omega)\frac{\partial \mathbf{L}(\omega)}{\partial p_i}\mathbf{s}(\omega) = \mathbf{L}^{-1}(\omega)\mathbf{f}^{(i)}_{virt}(\omega). \tag{8}$$

5 Assemble the Fréchet Matrix $\rightarrow \nabla_{\mathbf{p}} E(\mathbf{p})$

We now assemble the matrix $\hat{\mathbf{J}}$ to obtain

$$\begin{aligned}\hat{\mathbf{J}} &= \left[\frac{\partial \mathbf{s}(\omega)}{\partial p_1} \frac{\partial \mathbf{s}(\omega)}{\partial p_2} \cdots \frac{\partial \mathbf{s}(\omega)}{\partial p_m}\right] \\ &= \mathbf{L}^{-1}(\omega)\left[\mathbf{f}_{virt}^{(1)}(\omega)\mathbf{f}_{virt}^{(2)}(\omega)\cdots\mathbf{f}_{virt}^{(m)}(\omega)\right], \qquad (9)\end{aligned}$$

which can be written in a compact form given by

$$\hat{\mathbf{J}} = \mathbf{L}^{-1}(\omega)\mathbf{F}. \qquad (10)$$

$$\begin{aligned}\textbf{Finally} : \nabla_{\mathbf{p}} E(\mathbf{p}) &= \Re eal\{\hat{\mathbf{J}}^t \delta\hat{\mathbf{d}}^*(\omega)\} \\ &= \Re eal\{\mathbf{F}^T \left[\mathbf{L}^{-1}(\omega)\right]^T \delta\hat{\mathbf{d}}^*(\omega)\}. \qquad (11)\end{aligned}$$

6 Interpretation of $\nabla_{\mathbf{p}} E(\mathbf{p})$

Adjoint state method tutorial
M.J. Yedlin and D. Van Vorst
Pratt's Inversion Method
Adjoint State Method
a) Take Variation of $J(u(p))$
b) Add Zero – Introduce the Adjoint State Variable
c) Comparison with Pratt
Conclusions
Acknowledgements

We note the association of the new field shown below:

$$\nabla_{\mathbf{p}} E(\mathbf{p}) = \Re eal\{\mathbf{F}^T \underbrace{\left[\mathbf{L}^{-1}(\omega)\right]^T \delta\hat{\mathbf{d}}^*(\omega)}_{\text{Adjoint Field } \mathbf{v}}\}. \quad (12)$$

Substitute for the source term $\mathbf{F}^T$, for the m'th gradient component in (12) to get

$$\frac{\partial E(\mathbf{p})}{\partial p_m} = \Re eal\{-\mathbf{s}(\omega)^T \frac{\partial \mathbf{L}(\omega)}{\partial p_m}^T \mathbf{v}\}. \quad (13)$$

We recognize the above as Claerbout's principle of reflector mapping.

Background – History of the Method and Related Developments

Adjoint state method tutorial
M.J. Yedlin and D. Van Vorst
Pratt's Inversion Method
Adjoint State Method
a) Take Variation of $J(u(p))$
b) Add Zero – Introduce the Adjoint State Variable
c) Comparison with Pratt
Conclusions
Acknowledgements

1. 1755 – Lagrange formulated the adjoint operator – integration by parts;
2. 1910 – Hadamard – Calculus of Variations;
3. 1913 – Gâteaux – Variation of functionals – functional derivatives;
4. 1937 – Fréchet – Smooth functional derivatives;
5. 1968 – Lions – Optimal control;
6. 1974 – Marchuk – Neutron diffusion (originally done in the late 1940's);
7. 1971 – Claerbout – Theory of reflector mapping;
8. 1984 – Tarantola – Linearized Inversion of Seismic Reflection Data;
9. 1989 – Jameson – Airfoil design
10. 2006 – Liu and Tromp – Finite-Frequency Kernels Based on Adjoint Methods

Inner Product

Adjoint state method tutorial
M.J. Yedlin and D. Van Vorst
Pratt's Inversion Method
Adjoint State Method
a) Take Variation of $J(u(p))$
b) Add Zero – Introduce the Adjoint State Variable
c) Comparison with Pratt
Conclusions
Acknowledgements

Our misfit, which we called $E(\mathbf{p})$ in the Pratt paper, we will now call $J(p)$, where $J(p)$ is a cost function, usually measured in L^2 space. It will be denoted by

$$J(p) = \int_{domain} \left| u(p)^{pred} - u^{meas} \right|^2 d\Omega \tag{14}$$

This time we will keep things simple by initially only considering one control parameter p. Set $u(p)^{pred} \equiv u(p)$ and $u^{meas} \equiv u$. In (14) we have the Euclidean distance between prediction and measurements. It can be written as a scalar (dot product)

$$J(p) = \langle u(p) - u, u(p) - u \rangle \tag{15}$$

Compute the Variation of $J(u(p))$ With Respect to $u(p)$

Adjoint state method tutorial
M.J. Yedlin and D. Van Vorst
Pratt's Inversion Method
Adjoint State Method
a) Take Variation of $J(u(p))$
b) Add Zero – Introduce the Adjoint State Variable
c) Comparison with Pratt
Conclusions
Acknowledgements

Symbolically,

$$\begin{aligned} \delta J(u(p)) &= \left\langle \frac{\partial J}{\partial u}, \delta u \right\rangle \\ &= \langle g, \delta u \rangle \end{aligned} \tag{16}$$

where

$$\delta u = \frac{\partial u}{\partial p}\, dp \tag{17}$$

From a forward model,

$$A(p)u(p) = f, \tag{18}$$

we can compute δu.

Finish the Variation

To compute δu we have

$$\delta(A(p)u(p) - f) = 0 = \delta A(p)u(p) + A(p)\delta u(p) \tag{19}$$

So

$$A(p)\delta u(p) = -\delta A(p)u(p) = \tilde{f} \tag{20}$$

Our problem is "simple" now. Just compute $\delta J(u(p)) = \langle g, \delta u \rangle$ subject to $A(p)\delta u(p) = \tilde{f}$.

Notice that if we just have one parameter – no problem. **If we have many, we have to repeat the forward solution for** $\delta u(p)$. Note that if we are at a minimum of $J(u(p))$, then $\delta J(u(p)) = 0$ and we are done. Otherwise, we descend the gradient of the cost functional. What's next?

Enter Lions and Lagrange and the Adjoint State, V

Adjoint state method tutorial

M.J. Yedlin and D. Van Vorst

Pratt's Inversion Method

Adjoint State Method

a) Take Variation of $J(u(p))$

b) Add Zero – Introduce the Adjoint State Variable

c) Comparison with Pratt

Conclusions

Acknowledgements

The idea that Lions had to solve the above problem was to use optimal control theory and solve a dual problem by adding the forward modelling equation as a constraint, via the introduction of the adjoint state variable, v. Thus

$$\delta J(u(p)) = \langle g, \delta u \rangle - \left\langle v, A(p)\, \delta u(p) - \tilde{f} \right\rangle \tag{21}$$

Expand (21) and use the adjoint operator property, (integration by parts), $\langle v, A(p)\delta u \rangle = \left\langle A(p)^\dagger v, \delta u \right\rangle$ to get

$$\delta J(u(p)) = \left\langle g - A(p)^\dagger v, \delta u \right\rangle + \left\langle v, \tilde{f} \right\rangle \tag{22}$$

Coups de grâce: choose $g - A(p)^\dagger v = 0$ to get

$$\delta J(u(p)) = \left\langle v, \tilde{f} \right\rangle \tag{23}$$

Lions vs. Pratt – Why do it this way?

Adjoint state method tutorial

M.J. Yedlin and D. Van Vorst

Pratt's Inversion Method

Adjoint State Method

a) Take Variation of $J(u(p))$

b) Add Zero – Introduce the Adjoint State Variable

c) Comparison with Pratt

Conclusions

Acknowledgements

Let's compare our results:

Pratt:

$$\frac{\partial E(\mathbf{p})}{\partial p_m} = \Re eal\{-\mathbf{s}(\omega)^T \frac{\partial \mathbf{L}(\omega)}{\partial p_m}^T \mathbf{v}\} \tag{24}$$

Lions:

$$\begin{aligned} \delta J(u(p)) &= \left\langle v, \tilde{f} \right\rangle \\ &= \left\langle v, -\frac{\partial A(p)}{\partial p} \delta p u(p) \right\rangle \\ &= \left\langle -\frac{\partial A(p)}{\partial p} u(p), v \right\rangle \delta p \\ &= \nabla_p J(u(p)) \delta p \end{aligned} \tag{25}$$

By inspection the gradients in (24) and (25) are equal. Done!

Adjoint state method tutorial

M.J. Yedlin and D. Van Vorst

Pratt's Inversion Method

Adjoint State Method

a) Take Variation of $J(u(p))$

b) Add Zero – Introduce the Adjoint State Variable

c) Comparison with Pratt

Conclusions

Acknowledgements

Adjoint Method – More General

The adjoint method is more extensible, including issues of:

1. Norms
2. Variation operators as introduced by Gâteaux
3. Regularization
4. Non-linear forward modelling operators
5. Generality (from Liu and Tromp –Elastic Adjoint)

Elastic Adjoint Gradient

Adjoint state method tutorial

M.J. Yedlin and D. Van Vorst

Pratt's Inversion Method

Adjoint State Method

a) Take Variation of $J(u(p))$

b) Add Zero – Introduce the Adjoint State Variable

c) Comparison with Pratt

Conclusions

Acknowledgements

$$
\begin{aligned}
\delta\chi \quad = \quad & \int_\Omega \delta\rho(\mathbf{x}) \underbrace{\left[- \int_0^T \mathbf{s}^{adjoint}(\mathbf{x}, T-t) \cdot \partial_t^2\, \mathbf{s}(\mathbf{x}, t)\, \mathrm{d}t \right]}_{\text{Fréchet kernel for density}} \mathrm{d}^3\mathbf{x} \\
+ \quad & \int_\Omega \delta\mathbf{c}(\mathbf{x}) :: \underbrace{\left[- \int_0^T \nabla \mathbf{s}^{adjoint}(\mathbf{x}, T-t)\, \nabla \mathbf{s}(\mathbf{x}, t)\, \mathrm{d}t \right]}_{\text{Fourth order Fréchet kernel tensor}} \mathrm{d}^3\mathbf{x} \\
+ \quad & \int_0^T \int_\Omega \left[\mathbf{s}^{adjoint}(\mathbf{x}, T-t) \cdot \delta\mathbf{f} \right] \mathrm{d}^3\mathbf{x}\mathrm{d}t. \qquad (26)
\end{aligned}
$$

René Gâteaux (1889-1914) – Father of Functional Derivatives

Pierre Poyet dans sa « Turne » de Normale.
Gateaux, Langlamet, Gonthiez, Guadet, Poyet (à la table).

René Gâteaux (1889-1914), seated at far left in 1908, is the grandfather of the adjoint state method. His work created a formal setting for taking the variation of a functional.

Conclusions

Adjoint state method tutorial
M.J. Yedlin and D. Van Vorst
Pratt's Inversion Method
Adjoint State Method
a) Take Variation of $J(u(p))$
b) Add Zero – Introduce the Adjoint State Variable
c) Comparison with Pratt
Conclusions
Acknowledgements

1. Pratt's method is another way of implementing the adjoint state method;
2. The adjoint state method is a very powerful technique for handling a variety of inverse problems, with the forward modelling operator being possibly nonlinear;
3. The setting of the adjoint state method offers mathematical flexiblity not realizable in conventional approaches.

Acknowledgements

Adjoint state method tutorial

M.J. Yedlin and D. Van Vorst

Pratt's Inversion Method

Adjoint State Method

a) Take Variation of $J(u(p))$

b) Add Zero – Introduce the Adjoint State Variable

c) Comparison with Pratt

Conclusions

Acknowledgements

- CREWES sponsors and NSERC.
- G. Margrave, R.J. Ferguson and Kevin W. Hall.