Adjoint Models

Christian Melsheimer
IUP, University of Bremen

4th International Radiative Transfer Modeling Workshop,
Bredbeck, 9 July, 2002

Prologue

Let $< u, v >$ be some scalar product, L some linear operator[1].
Then the *adjoint operator* L^* defined by

$$< u, L^* v > = < Lu, v >$$

If L is a real matrix, then $L^* = L^T$ (transpose); if L is a complex matrix, then $L^* = \bar{L}^T$ (complex conjugate, transpose).

Basic Idea

Forward Model:

$$\vec{F}(\vec{x}) = \vec{y}$$

$$\left(\quad \Leftrightarrow F_i(x_1, \ldots x_n) = y_i \quad \right)$$

where
$\vec{F}$ is a complicated, non-linear, nasty function ("the Model");
$\vec{x}$ is the model input (often the retrieval variable);
$\vec{y}$ is the model output (simulated measurement).

Sensitivity Analysis:

$$\delta x \Rightarrow \delta y = ?$$

where δx is a small disturbance;
or, the other way round:

$$\delta y \Rightarrow \delta x = ?$$

where δy is a small disturbance;
or even:

$$\nabla_y J \Rightarrow \nabla_x J = ?$$

where $J(\vec{y}(\vec{x}))$ is, e.g., a cost function, a measure for some error of $\vec{y}(\vec{x})$ with respect to the measurement $\vec{y}_m$. ∇_y and ∇_y are the gradients with respect to $\vec{x}$ and $\vec{y}$, respectively.

[1] The usual conditions apply, such as $u \in$ vector space U and $v \in$ vector space V etc.

Tangent-Linear Model:

$$\delta y_i = \sum_k \frac{\partial F_i}{\partial x_k} \delta x_k$$

$$\overset{\frac{\partial F_i}{\partial x_k} \equiv K_{ik}}{\Longleftrightarrow} \quad \delta \vec{y} = K \cdot \delta \vec{x}$$

where K is the **Jacobian** of $\vec{F}$!
Here we have a linear mapping from δy to δx that involves the first derivative (something like a tangent) of the forward model, hence the name "tangent-linear".

Adjoint (Tangent-Linear) Model:

$$\frac{\partial J(\vec{y}(\vec{x}))}{\partial x_i} \overset{\vec{y}=\vec{F}(\vec{x})}{=} \sum_k \frac{\partial F_i}{\partial x_k} \frac{\partial J}{\partial y_i} \qquad \text{(chain rule)}$$

$$\equiv \nabla_x J(\vec{y}) = K^T \nabla_y J(\vec{y})$$

where K^T is the adjoint of K, i.e. the adjoint of the tangent linear model, hence the name. It is usually just called "adjoint model", although the correct name is "adjoint tangent-linear model".
If:

$$\vec{F}(\vec{x}) = \vec{F}^{(N)}(\ldots(\vec{F}^{(1)}(\vec{F}^{(0)}(\vec{x}))..) \quad \text{(a sequence of operations)}$$

then:

$$K = K^{(N)} \cdots K^{(1)} K^{(0)}$$

and:

$$K^T = {K^{(0)}}^T {K^{(1)}}^T \cdots {K^{(N)}}^T$$

So What?

- Adjoint models are widely used in meteorology, particularly for assimilation schemes
- ECMWF, MétéoFrance and others have adopted official coding conventions for adjoint models
- There are adjoint model compilers (at least in FORTRAN77) to automatically generate the code for the adjoint model from the code for the forward model, e.g. TAMC (Tangent Linear and Adjoint Model Compiler) by Ralf Giering (see below)

BUT:

- How to apply adjoint models to the actual retrieval in atmospheric sounding (OEM, Levenberg-Marquardt etc.)?
- Would it really make things easier (we still need the Jacobian ...)?

[...]

Sources of Information, References

- Errico, R.M.: What is an adjoint model? *Bull. Am. Met. Soc.*, vol. 78, p. 2577, 1997.
- Rodgers, C.D.: *Inverse Methods for Atmospheric Sounding: Theory and Practice*, 238 pp., World Scientific, Singapore, 2000. ISBN 981-02-2740-X
- Giering, R.: Automatic Adjoint and Tangent Linear Code Generation, http://www.npaci.edu/Research/ESS/seminars/giering.htm
- Kornblueh, L., Working group on Model evaluation and data assimilation, Technical Explanations, http://www.mpimet.mpg.de/working_groups/wg3/node3.html