Use and implementation of adjoint methods in ice sheet models

Jesse Johnson, Jean Utke

30 June, 2010
Land Ice Working Group Session
Objective function
Using $g(x, p)$

Given model output $x$, it is desirable to compute a scalar valued function

$$g(x, p)$$

that also depends on model parameters $p$. Examples of $g(x, p)$ include:

- flow rate,
- stresses,
- energy, or
- model agreement with data.

Refer to $g(x, p)$ as an objective function.
Utility of derivatives

Why $\frac{dg}{dt}$ is needed, why its hard to get.

Differentiating the objective function, $\frac{dg}{dp}$ provides:

- the *sensitivity* of the objective function with respect to the parameters, and
- the *search direction* to be used in conjunction with conjugate gradients to determine the minimum of $g(x, p)$.

The chain rule gives:

$$\frac{dg}{dp} = g_x x_p + g_p$$

Underscoring the problem, $x_p$ is tough to evaluate!

Assuming $x$ can be written $Ax = b$, and its derivative is $A_{p_i} x + A x_{p_i} = b_{p_i}$, each $x_{p_i}$ is solved with $x_{p_i} = A^{-1}(b_{p_i} - A_{p_i}x)$

...one linear system for each parameter!
Avoiding the M linear solves
Add zero in a cunning way

- Rewrite the objective function

\[ \tilde{g} = g - \lambda^T f \]

where \( f = Ax - b \), which is zero, making \( \lambda \) arbitrary.

- Strategy is to choose \( \lambda \) such that \( x_p \) is eliminated

\[ \left. \frac{dg}{dp} \right|_{f=0} = \left. \frac{d\tilde{g}}{dp} \right|_{f=0} = g_p - \lambda^T f_p + (g_x - \lambda^T f_x) x_p \]

- \( x_p \) is eliminated if

\[ f_x^T \lambda = g_x^T \]

- \( A = f_x \), so what we really require is that \( \lambda \) satisfies the adjoint equation

\[ A^T \lambda = g_x^T. \]

Hence \( \frac{dg}{dp} \) comes from the evaluation of a single linear system!
The conclusion

Having solved the adjoint system for $\lambda$, the gradient is written

$$\frac{dg}{dp} = g_p - \lambda^T (A_p x - b_p)$$

noting that;

- $x$ is the result of solving the forward model,
- Computing $A_p$ and $b_p$ are assumed to be analytic expressions, and can be treated “automatically”.
- Automatic differentiation (AD) is done with openAD (Utke).
Greenland ice sheet
Velocities from Joughin 2010
Greenland ice sheet
Velocities from Joughin 2010
Greenland ice sheet
Profile

Isunnguata
Sermia
Russel
Glacier

Kilometers
0 10 20 40 60 80

University of Montana
Adjoint LIWG Session 7 / 11
Conservation of energy

\[
\frac{1}{\rho c_p} \nabla \cdot k_i \nabla \theta - \mathbf{u} \cdot \nabla \theta + 2\eta \dot{\varepsilon}_n^2 = 0
\]

Conservation of momentum

\[
\nabla \cdot 2\eta \dot{\varepsilon} - \nabla p = \rho \mathbf{g}
\]

Boundary conditions

\[
[-p \mathbf{l} + 2\eta \dot{\varepsilon}] \hat{\mathbf{n}} = 0 \quad \text{(Free surface)},
\]

\[
\tau_b = \beta^2 \cdot \mathbf{u} \quad \text{(Basal traction)},
\]

\[
-\hat{\mathbf{n}} k_i \nabla \theta = Q \quad \text{(Basal heat flow)}.
\]
Objective function minimization
Using quasi-Newton method
Resulting fields
Temperature and velocity

![Graph showing temperature and velocity fields over distance and altitude.](chart.png)
Sensitivity
Sensitivity of temperature to heat flow

Contour: Sensitivity of $T$ to $\alpha_{000}$

Max: 246.13
Min: -1.225

(University of Montana) Adjoint LIWG Session 11 / 11
Sensitivity
Sensitivity of velocity to heat flow