A (Very) Brief Introduction to Adjoint Methods

Eric Dow

July 20, 2015

- Optimization problems are often "constrained" by some set of equations (PDEs, ODEs or linear systems)
  - Goal is to minimize or maximize some objective function while satisfying constraint equations (i.e. remaining physical)
- Adjoint methods are useful for computing sensitivities with respect to a large number of parameters (relative to the number of objective functions)
  - Shape optimization: minimize drag around an airfoil by optimizing location of points that define the airfoil surface
  - Mesh adaptation: minimize numerical error in PDE solution by optimizing grid spacing/orientation for every point in the mesh
  - Uncertainty quantification: determine which random variables have the greatest impact on the statistical moments of a stochastic PDE solution

Adjoints for Linear Systems

- Consider a linear system
$$Au = f(p)$$
where f depends on some vector of parameters p
- Define linear objective function
$$J = g^T u$$
- Suppose we want to compute J for a large number parameter vectors $\{p_1, p_2, \dots, p_n\}$
- Naïve approach:
  - Solve linear system $\boldsymbol{n}$ **times** (once for each parameter vector p_i) to get solution vector u_i
  - Compute objective function $J_i = g^T u_i$
- But there's a better way...

Adjoints for Linear Systems

- Introduce Lagrange multiplier vector v (adjoint variable), and form Langrangian equation

$$L(u, v) = g^\mathsf{T} u + v^\mathsf{T}(Au - f)$$

- Since $Au - f = 0$ for any solution u to the linear system, $L(u, v) = J(u)$
- Goal: make evaluation of L (and therefore J) independent of u

$$\begin{aligned} L(u, v) &= g^\mathsf{T} u + v^\mathsf{T}(Au - f) \\ &= u^\mathsf{T}(A^\mathsf{T} v + g) - f^\mathsf{T} v \end{aligned}$$

  If we choose v to satisfy $A^\mathsf{T} v + g = 0$, dependence on u is eliminated!
- Adjoint approach:
  - Solve adjoint equation $A^\mathsf{T} v = -g$ **one time**
  - Compute objective function $J_i = -f(p_i)^\mathsf{T} v$

- Consider a linear convection-diffusion equation defined on a domain Ω with boundary Γ

$$\frac{\partial u}{\partial t} + a \cdot \nabla u = \nabla \cdot (D \nabla u) + s(x; p)$$

$$u|_{\Gamma} = 0 \qquad u(x, 0) = u_0(x)$$

- Linear objective function:

$$J(u) = \int_{\Omega} u(x, T) \; d\Omega$$

- What is the sensitivity of J with respect to the source parameters p?

- Form Lagrangian by introducing adjoint variable $v(x,t)$:

$$
\begin{aligned}
L(u,v) &= \int_\Omega u(x,T)\ d\Omega \\
&+ \int_0^T \int_\Omega v \left[\frac{\partial u}{\partial t} + a \cdot \nabla u - \nabla \cdot (D \nabla u) - s(x;p) \right]\ d\Omega\ dt
\end{aligned}
$$

- Integrate by parts in time and space a couple times:

$$
\begin{aligned}
L(u,v) &= \int_\Omega u(x,T) + u(x,T)v(x,T) - u(x,0)v(x,0)\ d\Omega \\
&+ \int_0^T \int_\Gamma uva \cdot n - u\frac{\partial v}{\partial t} - vD\nabla u \cdot n + uD\nabla v \cdot n\ d\Gamma\ dt \\
&+ \int_0^T \int_\Omega u \left[-\frac{\partial v}{\partial t} - a \cdot \nabla v - \nabla \cdot (D\nabla v) \right]\ d\Omega\ dt \\
&- \int_0^T \int_\Omega s(x;p)v\ d\Omega\ dt
\end{aligned}
$$

- To make the Lagrangian independent of u, we choose

$$
\boxed{
\begin{aligned}
-\frac{\partial v}{\partial t} - a \cdot \nabla v - \nabla \cdot (D\nabla v) = 0 \\
v|_\Gamma = 0 \qquad v(x,T) = -1
\end{aligned}
}
$$

- Adjoint equation is weird: anti-diffusive, terminal boundary condition
- To fix this, introduce new "backwards" time variable $\tau = T - t$:

$$\boxed{\begin{gathered}\frac{\partial v}{\partial \tau} - a \cdot \nabla v - \nabla \cdot (D \nabla v) = 0 \\ v|_\Gamma = 0 \qquad v(x,0) = -1\end{gathered}}$$

- Adjoint equation runs backwards in time relative to the primal equation
- After solving the adjoint equation **one time**, objective function can be evaluated for any vector p as

$$J = -\int_0^T \int_\Omega s(x;p) v(x,t)\; d\Omega\; dt$$

- Since the adjoint solution $v(x,t)$ is independent of p,

$$\nabla_p J = -\int_0^T \int_\Omega \nabla_p s\; v(x,t)\; d\Omega\; dt$$

- 1D Burgers' equation on $x \in [0, 1]$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0$$

$$u(0, t) = 0 \qquad u(x, 0) = u_0(x)$$

- Want to determine the sensitivity of J to the initial condition $u_0(x)$ where

$$J(u) = \int_0^1 \frac{1}{2}u^2(x, T)\ dx$$

- To form the Lagrangian, first linearize the PDE and objective around the solution $u(x, t)$:

$$\frac{\partial \hat{u}}{\partial t} + \hat{u}\frac{\partial u}{\partial x} + u\frac{\partial \hat{u}}{\partial x} = 0$$

$$\hat{u}(0, t) = 0$$

$$\hat{J}(\hat{u}) = \int_0^1 u(x, T)\hat{u}(x, T)\ dx$$

Adjoints for Non-linear PDEs

- Form Lagrangian by introducing adjoint variable $v(x,t)$:

$$
\begin{aligned}
L(\hat{u},v) &= \int_0^1 u(x,T)\hat{u}(x,T)\;dx \\
&- \int_0^T \int_0^1 v\left[\frac{\partial \hat{u}}{\partial t} + \hat{u}\frac{\partial u}{\partial x} + u\frac{\partial \hat{u}}{\partial x}\right]\;dx\;dt
\end{aligned}
$$

- Integrate by parts in time and space:

$$
\begin{aligned}
L(\hat{u},v) &= \int_0^1 u(x,T)\hat{u}(x,T) - \hat{u}(x,T)v(x,T) + \hat{u}(x,0)v(x,0)\;dx \\
&+ \int_0^T u(0,t)\hat{u}(0,t)v(0,t) - u(1,t)\hat{u}(1,t)v(1,t)\;dt \\
&+ \int_0^T \int_0^1 \hat{u}\left[\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x}\right]\;dx\;dt
\end{aligned}
$$

- To make the Lagrangian independent of $\hat{u}$, we choose

$$
\boxed{
\begin{aligned}
&\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} = 0 \\
v(1,t) = 0 \qquad &v(x,T) = u(x,T)
\end{aligned}
}
$$

- Again introduce "backwards" time variable $\tau = T - t$:

$$\boxed{\begin{array}{r} \dfrac{\partial v}{d\tau} - u\dfrac{\partial v}{\partial x} = 0 \\ v(1,\tau) = 0 \qquad v(x,0) = u(x,t=T) \end{array}}$$

- Perturbation in objective function can be computed from the adjoint solution:

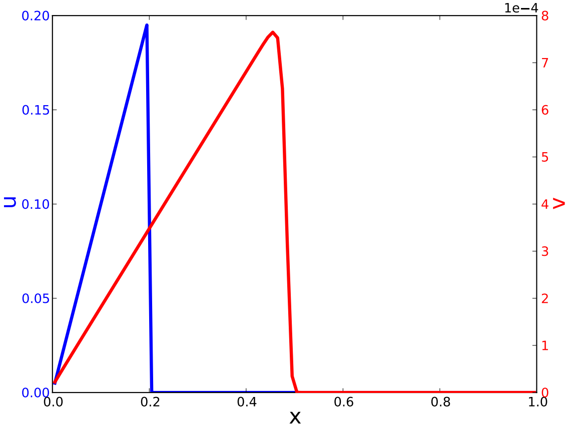
$$\hat{J}(\hat{u}) = \int_0^1 \hat{u}(x,t=0)v(x,\tau=T)\ dx$$

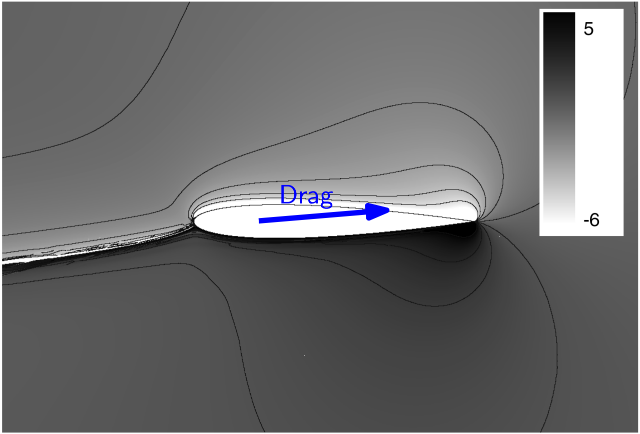
- Frechét derivative with respect to perturbations in the initial condition is therefore

$$\frac{\delta J}{\delta u_0} = v(x,\tau=T)$$

$$J = \frac{1}{2}\int_0^1 u(x,T)^2 \; dx$$

What is the sensitivity of J with respect to to the initial condition $u_0(x)$?





x-momentum adjoint for a NACA 0012 at $M = 0.4$, $\alpha = 5^\circ$ [Fidkowski, 2011]

- **Continuous Adjoint**: form adjoint PDE from primal PDE, discretize the adjoint PDE how you like
  - Get to choose discretization for both problems, so easier to ensure stability
  - Computed gradients are not, in general, exact with respect to the numerical approximation of the objective function

- **Discrete Adjoint**: discretize the primal PDE, solve adjoint of discretized primal
  - Computed gradients are exact with respect to the numerical approximation of the objective function
  - Discrete adjoint is not guaranteed to be stable

- **Dual consistent** schemes link the two: choose a discretization of the primal equations such that the discrete adjoint is a consistent discretization of the continuous adjoint PDE
  - This is most easily done in the context of finite element methods

Questions?