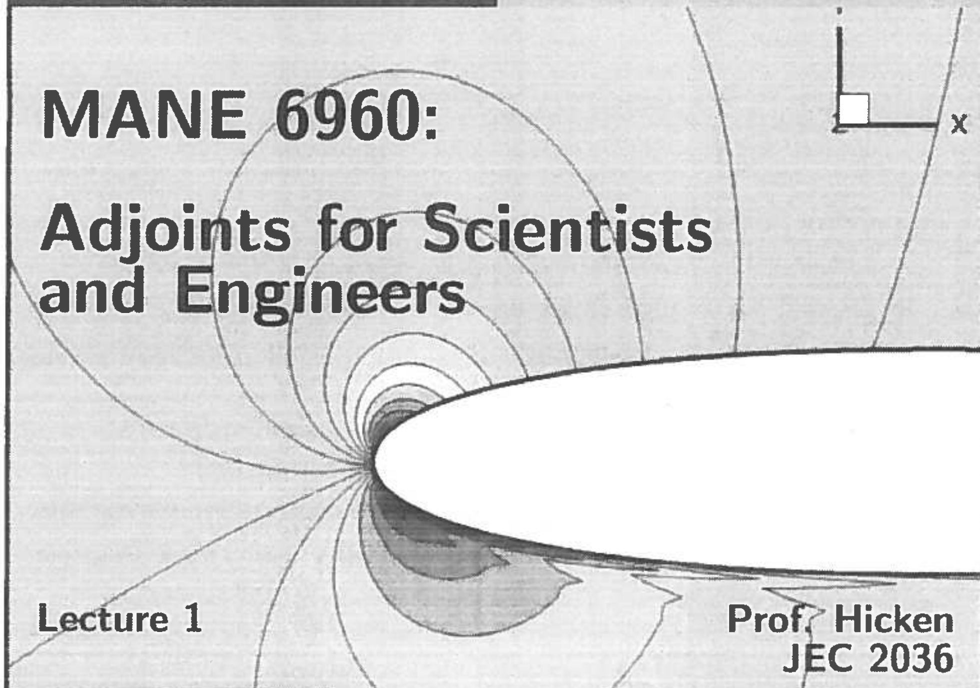


MANE 6960:

Adjoints for Scientists and Engineers

Lecture 1

Prof. Hicken
JEC 2036

Course Objectives

MANE 6960, "Adjoints for Scientists and Engineers," aims to help you:

- be able to derive the adjoint equation for any given primal problem and functional;
- use the adjoint for sensitivity analysis and output error estimation; and,
- implement and solve adjoint problems in software.

Instructor

Prof. Jason Hicken

- Office Location: JEC 2036
- email: hickej2@rpi.edu
- Office Hours: TBD

Prerequisites

To take this course, your previous course work should have included

- multivariate and vector calculus,
- ordinary and partial differential equations,
- numerical methods, and
- programming.

If you are missing one of these, you might be able to get by...

Course Texts

No required text(s)

Supplemental References:

- C. Lanczos, *"Linear Differential Operators,"* SIAM, 1996
- J. L. Lions, *"Optimal Control of Systems Governed by Partial Differential Equations,"* Springer-Verlag, 1971
- A. Borzi and V. Schulz, *"Computational Optimization of Systems Governed by Partial Differential Equations,"* SIAM, 2012

Grading Breakdown

There are four major assignments/projects that will make up the bulk of your grade

- $100\% = 4\times\ 25\%$
- Each will require extensive programming
- I will expect a LaTeX'ed report for each

I will introduce the first assignment next class.

Class Policies

See the syllabus for further details.

Late Assignments: 10% penalty if submitted within 24hrs; 25% penalty if submitted within a week; 100% penalty otherwise.

Please read the Academic Integrity statement in the syllabus:

- first violation = grade of zero on assignment
- second violation = grade of F in the course

mention schedule here

Motivation

Applications

In science and engineering, we frequently encounter problems for which we need to determine parameters in a system that is governed by a partial differential equation (PDE).

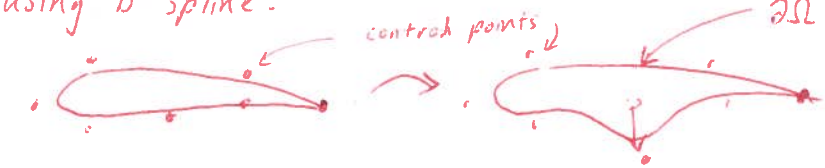
- simulation-based design optimization
- PDE-constrained inverse problems

Let's consider some concrete examples...

Example 1: drag minimization

Find the airfoil shape that minimizes drag subject to the incompressible Navier-Stokes equations

Suppose boundary, $\partial\Omega$, is parameterized using B-spline:



Let $\alpha \in \mathbb{R}^n$ denote the control point positions

Let $\vec{u}, p$ denote the velocity and pressure, respectively

Example 1: drag minimization (cont.)

Problem statement:

$$\min_{\alpha,\vec{u},p} D(\alpha,\vec{u},p) = -\int_{\partial\Omega(\alpha)} (\underline{\Sigma}\,\hat{n})\cdot\hat{z}_\infty \, d\Gamma$$

subject to

$$\vec{u}\cdot\vec{\nabla}\vec{u} = -\nabla p + \mu\nabla^2\vec{u}, \quad \forall\, x \in \Omega$$

$$\vec{\nabla}\cdot\vec{u} = 0, \quad \forall\, x \in \Omega$$

$$\vec{u} = 0, \quad \forall\, x \in \partial\Omega$$

Example 2: inverse problem in elastography

Find the shear modulus such that computed displacements are close, in some sense, to a set of measured displacements.

$\{\tilde{u}_i\}_{i=1}^m$ = set of measured displacements (from ultra sound, e.g.)

Goal is to find μ (shear modulus)
$\quad\;\; \| $
$\quad\;\; \alpha$ here

Example 2: inverse problem in elastography (cont.)

Problem statement:

$$\min_{\alpha=\mu,u} \mathcal{J}(\alpha,u) = \sum_{i=1}^{m} \|\vec{u}-\tilde{u}\|^2 + \underbrace{\frac{\sigma}{2}\int \mu^2 \, d\Omega}_{\text{regularization}}$$

subject to

$$\nabla\cdot(\mu\, C\, \nabla\vec{u}) = 0, \quad \forall x \in \Omega$$

$$\vec{u} = \vec{y}, \quad \forall x \in \partial\Omega$$

Problem Characteristics

Both the above examples share the same basic characteristics.

1. There are a (potentially) large number of parameters that must be determined; in some applications the parameters may be infinite dimensional.
2. The problems are governed by a PDE constraint.

> Gradient descent is the most efficient means of solving these types of problems, due to the large number of parameters; however, how do we find the gradient?

Generic Problem

To answer the above question, let's consider a more general (abstract) problem. (algebraic)

$$\begin{aligned} \min_{\alpha,u} \quad & \mathcal{J}(\alpha,u) \\ \text{s.t.} \quad & \mathcal{R}(\alpha,u) = 0 \end{aligned}$$

where

- $\alpha \in \mathbb{R}^n$ parameter vector to be determined,
- $u \in \mathbb{R}^s$ is the state,
- $\mathcal{J}$ is the objective, or cost function; and
- $\mathcal{R}$ is the state equation.

Generic Problem (cont.)

In order to use a gradient-based method to solve the problem, we need the gradient:

(total) gradient $= \frac{D\mathcal{J}}{D\alpha}$, with respect to α

But $\mathcal{J}(\alpha,u)$ is also a function of the state, u. Because of this, we need to account for how changes in α impact u: Thus

$$(1) \quad \frac{D\mathcal{J}}{D\alpha} = \frac{\partial \mathcal{J}}{\partial \alpha} + \frac{\partial \mathcal{J}}{\partial u}\underbrace{\frac{Du}{D\alpha}}_{}$$

$:=$ direct sensitivities

Generic Problem (cont.)

Q: How do we find $Du/D\alpha$?

Appeal to implicit function theorem

Assuming $R(\alpha,u)$ is continuously differentiable and that $\partial R/\partial u$ is invertible ~~in Nbr of~~ at (α,u), then $u=u(\alpha)$ and

$$(2)\qquad \frac{DR}{D\alpha} = \frac{\partial R}{\partial \alpha} + \frac{\partial R}{\partial u}\frac{Du}{D\alpha} = 0 \leftarrow R.H.S = 0 \quad \because R=0\ \forall \alpha$$

Solve for $Du/D\alpha$ in (2), then substitute into (1)

Generic Problem (cont.)

$$\frac{DJ}{D\alpha} = \frac{\partial J}{\partial \alpha} - \frac{\partial J}{\partial u}\overbrace{\left(\frac{\partial R}{\partial u}\right)^{-1}\frac{\partial R}{\partial \alpha}}^{-Du/D\alpha}$$

Q: What is the (practical) problem with this?

$$\frac{Du}{D\alpha} = -\left(\frac{\partial R}{\partial u}\right)^{-1}\frac{\partial R}{\partial \alpha}$$

$\left(\frac{\partial R}{\partial \alpha}\right)$ has n columns,

$\therefore$ we need to solve n systems whose size is the same as $R(\alpha,u)$.

i.e. s

Generic Problem (cont.)

Solution? Introduce the adjoint!

$$\frac{DJ}{D\alpha} = \frac{\partial J}{\partial \alpha} \cancel{+} \underbrace{-\frac{\partial J}{\partial u}\left(\frac{\partial R}{\partial u}\right)^{-1}}_{\text{define this vector to be } \psi^T}\frac{\partial R}{\partial \alpha}$$

$$\therefore \quad \psi^T := -\frac{\partial J}{\partial u}\left(\frac{\partial R}{\partial u}\right)^{-1}$$

$$\Longrightarrow \boxed{\left(\frac{\partial R}{\partial u}\right)^T \psi = -\left(\frac{\partial J}{\partial u}\right)^T} \leftarrow \text{RHS has just 1 column!}$$

(algebraic/discrete) adjoint equation

Take-away message

We only need one adjoint for each $\mathcal{J}$ to get the gradient with respect to any number of parameters, including infinite-dimensional parameters.

The reason for this is that

$$\frac{D\mathcal{J}}{D\alpha} = \frac{\partial \mathcal{J}}{\partial \alpha} + \psi^T\frac{\partial R}{\partial \alpha}$$

involves only (relatively cheap) products.

What's next?

There is not much more to say regarding the algebraic case, but there are a whole host of questions that arise if we dig deeper:

- What does $\partial\mathcal{R}/\partial u$ mean when $\mathcal{R}$ is a PDE?
- What is $(\partial\mathcal{R}/\partial u)^T$ mean when $\mathcal{R}$ is a PDE?
- What role do boundary conditions ~~in~~ play the adjoint?
- How does one compute ψ in practice when there are thousands or millions of state equations?
- Does this work for time dependent problems?

This course aims to answer these questions and more.