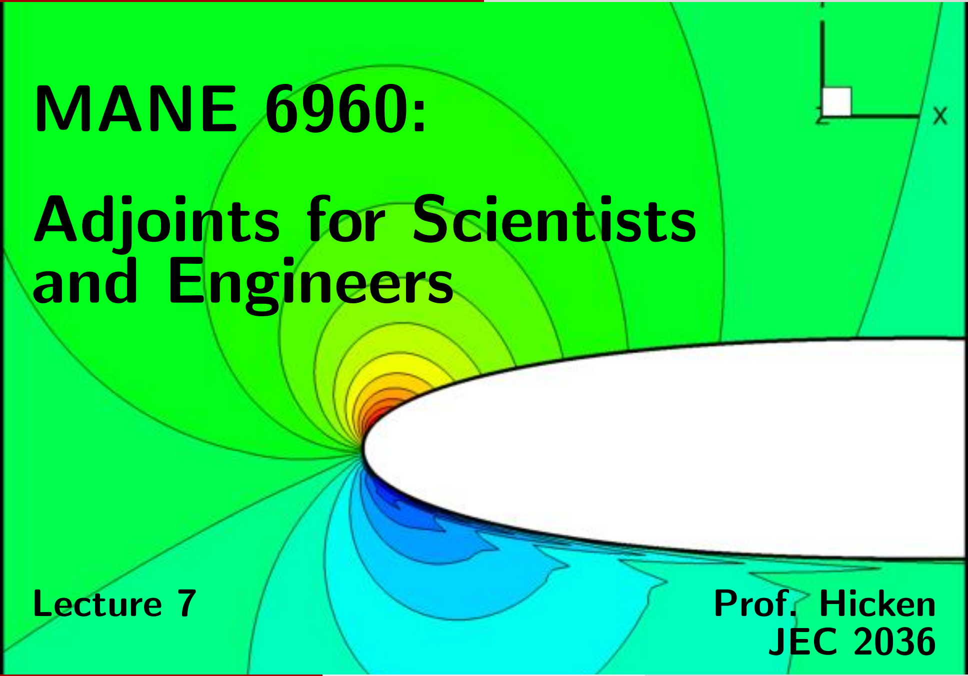


MANE 6960:

Adjoints for Scientists and Engineers

Lecture 7

Prof. Hicken
JEC 2036

Lecture Objective

Many physical systems are modelled using nonlinear boundary-value problems. Furthermore, the quantity of interest in such systems is often a nonlinear functional that depends on the solution of the (nonlinear) BVP.

This motivates the generalization of the concept of an adjoint to nonlinear problems, which is the topic we will begin to discuss in this lecture.

e.g
- Fluid Flows
- Nonlinear Structural Problems

Nonlinear Boundary-Value Problems

In this lecture, and those that follow, we will be concerned with the generic nonlinear BVP defined below:

$$\begin{aligned} N(u) &= 0, \qquad \forall x \in \Omega, \\ B(u) &= 0, \qquad \forall x \in \Gamma. \end{aligned} \tag{$\star$}$$

- N is a nonlinear differential operator
- B is a nonlinear boundary operator

Examples: $\vec{\nabla} \cdot \vec{F}_i = 0$, where $\vec{F}_i$ is the flux of the ith conserved quantity in the Euler equations

Nonlinear Functionals

In addition, we will consider generic nonlinear functionals of the form

$$\begin{aligned} J(u) &= \int_\Omega g(u)\, d\Omega + \int_\Gamma c(C(u))\, d\Gamma \\ &= (g(u), 1)_\Omega + (c(C(u)), 1)_\Gamma. \end{aligned}$$

- $g(u)$ is a nonlinear function of u
- $C(u)$ is a nonlinear boundary operator
- $c(C(u))$ is a nonlinear boundary function

Fréchet and Gâteaux Derivatives

Motivation

Recall the definition of the discrete adjoint from the first lecture:

$$\left(\frac{\partial R_h}{\partial u_h}\right)^T \psi_h = \frac{\partial J_h}{\partial u_h}$$

This suggests that we will need to differentiate N, B, and J with respect to u, in order to generalize the adjoint BVP to nonlinear problems.

But what is the meaning of $\partial N/\partial u$ when N is a nonlinear operator and u is a function, i.e. an infinite-dimensional vector?

Motivation (cont.)

The problem with using the usual derivative in the context of functions can be appreciated by trying to apply the definition of a partial derivative.

For a function $J_h : \mathbb{R}^n \rightarrow \mathbb{R}$, the ith partial derivative w.r.t. u_h is given by

$$\frac{\partial J_h}{\partial u_{h,i}} = \lim_{\epsilon \rightarrow 0} \frac{J_h(u_h + e_i \epsilon) - J_h(u_h)}{\epsilon}$$

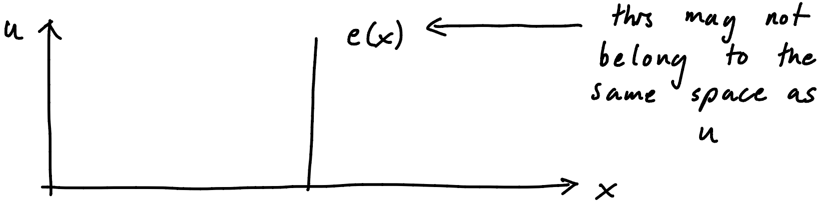
where

$$(e_i)_j = \begin{cases} 1, & \text{if } j = i \\ 0, & \text{if } j \neq i \end{cases}$$

$$e_i = \begin{array}{cl} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow \text{ith row} \\ \\ \\ \end{array} \end{array}$$

Motivation (cont.)

Defining e_i becomes problematic when $u \in \mathcal{V}$, i.e. when $\mathcal{V}$ is a function space and u is infinite-dimensional.



We need a more general definition for a derivative with respect to a function.

Fréchet Derivative

$N'[u]$ denotes that derivative is evaluated at u

Definition: Fréchet Derivative

Let $\mathcal{V}$ and $\mathcal{W}$ be two Banach spaces, and let $\mathcal{U} \subset \mathcal{V}$ be an open subset. The function $N : \mathcal{U} \rightarrow \mathcal{W}$ is Fréchet differentiable at $u \in \mathcal{U}$ if

$$\lim_{v \to 0} \frac{\|N(u+v) - N(u) - N'[u]v\|_{\mathcal{W}}}{\|v\|_{\mathcal{V}}} = 0,$$

for some bounded linear operator $N'[u] : \mathcal{V} \rightarrow \mathcal{W}$, which we call the Fréchet derivative of N at u.

- Banach space: a complete, normed vector space.

note, v is a <u>vector</u>; can take any path to 0

Fréchet Derivative (cont.)

I used the nonlinear differential operator, N, in the Fréchet derivative definition, but I just as easily could have used B or J. The only difference in the present context is the range space.

- For N, $\mathcal{W}$ is a function space, for example, the space of functions with bounded 2-norms on Ω, i.e. $L^2(\Omega)$.
- For B, $\mathcal{W}$ is the space of bounded functions on Γ.
- For J, $\mathcal{W} = \mathbb{R}$.

Gâteaux Derivative

The Fréchet derivative definition is not very helpful in practice. Instead, I usual rely on the directional derivative to compute the Fréchet derivative.

Definition: Directional (Gâteaux) Derivative

Let $\mathcal{V}$ and $\mathcal{W}$ be two Banach spaces. The directional, or Gâteaux, derivative of the function $N : \mathcal{U} \subset \mathcal{V} \to \mathcal{W}$ at $u \in \mathcal{U}$ in the direction $v \in \mathcal{U}$ is

$$D_v N(u) = \lim_{\epsilon \to 0} \frac{N(u + \epsilon v) - N(u)}{\epsilon}$$

$\varepsilon \in \mathbb{R}$

Gâteaux Derivative (cont.)

Why is the directional derivative useful?

If N is Fréchet differentiable at u (which we will assume), then the directional derivative of N exists at u in all directions v,[1] and

$$N'[u]v = D_v N(u).$$

This is analogous to the finite-dimensional formula $(\nabla f) \cdot v = D_v f$

Furthermore, we have the following practical formula:

$$D_v N(u) = \left[\frac{d}{d\epsilon} N(u + \epsilon v) \right]_{\epsilon = 0}.$$

[1] The converse is not true; the directional derivative can exist in all directions, but N may not be Fréchet differentiable.

Exercise

Determine the Fréchet derivative of the functional

$$J(u) = \int_\Omega (\nabla u) \cdot (\nabla u) \, d\Omega.$$

$$D_v J(u) = \left[\frac{\partial}{\partial \varepsilon} \int_\Omega (\nabla(u+\varepsilon v)) \cdot (\nabla(u+\varepsilon v)) \, d\Omega \right]_{\varepsilon=0}$$

$$= \left[\frac{\partial}{\partial \varepsilon} \int_\Omega \sum_{i=1}^{d} \left(\frac{\partial}{\partial x_i} (u+\varepsilon v) \right)^2 d\Omega \right]_{\varepsilon=0}$$

Exercise (cont.)

$$D_v J(u) = \left[\int_\Omega \sum_{i=1}^{d} 2 \left(\frac{\partial}{\partial x_i}(u+\varepsilon v) \right) \frac{\partial}{\partial \varepsilon} \frac{\partial}{\partial x_i}(u+\varepsilon v) \, d\Omega \right]_{\varepsilon=0}$$

$$= \left[\int_\Omega \sum_{i=1}^{d} 2 \left(\frac{\partial}{\partial x_i}(u+\varepsilon v) \right) \frac{\partial v}{\partial x_i} \, d\Omega \right]_{\varepsilon=0}$$

$$= \int_\Omega 2 (\nabla u) \cdot (\nabla v) \, d\Omega$$

$$= J'[u] v$$

Note, this is linear in v, as expected

Exercise

Determine the Fréchet derivative of the nonlinear operator

$$N(u) = u\frac{\partial u}{\partial x}.$$

$$D_v N(u) = \left[\frac{\partial}{\partial \varepsilon} \left\{ (u + \varepsilon v) \frac{\partial}{\partial x}(u + \varepsilon v) \right\} \right]_{\varepsilon = 0}$$

$$= \left[v \frac{\partial}{\partial x}(u + \varepsilon v) + (u + \varepsilon v) \frac{\partial v}{\partial x} \right]_{\varepsilon = 0}$$

Exercise (cont.)

$$D_v N(u) = v \frac{\partial u}{\partial x} + u \frac{\partial v}{\partial x}$$

$$= N'[u]\, v \qquad \text{(again, linear)}$$

We get the same answer using

$$N(u) = u \frac{\partial u}{\partial x} = \frac{1}{2} \frac{\partial}{\partial x}(u^2)$$

$$N'[u]\, v = \frac{\partial}{\partial x}(uv) = v \frac{\partial u}{\partial x} + u \frac{\partial v}{\partial x}$$

Deriving the Adjoint BVP for Nonlinear Problems

Green's Identity, revisited

With Fréchet differentiation in our toolbox, we can now derive the adjoint for nonlinear BVPs and functionals.

First, we will need to identify the relevant terms in the extended Green's identity:

$$(\psi, Lu)_\Omega - (u, L^*\psi)_\Omega = (Cu, B^*\psi)_\Gamma - (Bu, C^*\psi)_\Gamma.$$

For the nonlinear BVP, we have

- $L \to N'[u]$;
- $B \to B'[u]$; and
- $C \to C'[u]$.

Green's Identity, revisited (cont.)

Definition: Adjoint Operators of Nonlinear BVPs

Consider the nonlinear boundary value problem $(\star)$, and assume that N and B are Fréchet differentiable for all $u \in \mathcal{U} \subset \mathcal{V}$. Then we have the following identity:

ψ is not necessarly the adjoint

$$(\psi, N'[u]v)_\Omega - (v, N'[u]^*\psi)_\Omega =$$
$$(C[u]'v, B'[u]^*\psi)_\Gamma - (B'[u]v, C'[u]^*\psi)_\Gamma, \qquad \forall v, \psi \in \mathcal{U},$$

where $N'[u]^*$ denotes the adjoint differential operator, and $B'[u]^*$ denotes the adjoint boundary-condition operator.

Adjoint BVP

We now have a means of finding the adjoint differential and boundary operators; however, in order to define the adjoint BVP we also need the boundary values and source term.

To this end, we need the Fréchet derivative of the functional $J(u)$:

$$J'[u]v = \int_\Omega g'[u]v \, d\Omega + \int_\Gamma c'[Cu]C'[u]v \, d\Gamma.$$

Then we can make the following associations with the linear adjoint problem:

- $g \to g'[u]$;
- $c \to c'[Cu]$; and
- $C \to C'[u]$.

Adjoint BVP (cont.)

Definition: Adjoint Problem (nonlinear BVP)

Let u be the solution to the nonlinear BVP $(\star)$, and consider the nonlinear functional $J(u)$ defined earlier. Then the associated adjoint boundary-value problem is

$$\begin{aligned} N'[u]^*\psi &= g'[u], & \forall x \in \Omega, \\ B'[u]^*\psi &= c'[Cu], & \forall x \in \Gamma. \end{aligned} \tag{Adj}$$

* even though N, B, and J are potentially nonlinear, the adjoint problem is always linear

Connection to Sensitivity Analysis

We conclude by demonstrating that the adjoint problem (Adj) is precisely the one we want for sensitivity analysis (we will delve into this topic more in a few lectures).

$N(u,\alpha) = 0$
$B(u,\alpha) = 0$

Consider a nonlinear BVP and functional that depends on a parameter $\alpha \in \mathbb{R}^n$:

$$J(u,\alpha) = \int_\Omega g(u,\alpha)\, d\Omega + \int_\Gamma c(C(u),\alpha)\, d\Gamma$$

Then

$$\frac{\partial J_h}{\partial u_h} \frac{D u_h}{D\alpha}$$

$$\frac{DJ}{D\alpha} = \int_\Omega \nabla_\alpha g \, d\Omega + \int_\Gamma \nabla_\alpha c \, d\Gamma$$

$$+ \int_\Omega g'[u,\alpha] v \, d\Omega + \int_\Gamma c'[C(u),\alpha]\, C'[u] v \, d\Gamma$$

Connection to Sensitivity Analysis (cont.)

where $v = \frac{Du}{D\alpha}$ is the direct sensitivity (we would need to solve n linearized PDEs to get all n of the v).

Using the solution to the adjoint BVP

$$\frac{DJ}{D\alpha} = \int_\Omega \nabla_\alpha g \, d\Omega + \int_\Gamma \nabla_\alpha c \, d\Gamma$$

$$+ \underbrace{\int_\Omega (N'[u]^* \psi) v \, d\Omega}_{(N'[u]^*\psi,\, v)_\Omega} + \underbrace{\int_\Gamma (B'[u]^* \psi) C'[u] v \, d\Gamma}_{(B'[u]^*\psi,\, C'[u]v)_\Gamma}$$

Connection to Sensitivity Analysis (cont.)

Next, we use Green's Identity to get

$$\frac{DJ}{D\alpha} = \int_\Omega \nabla_\alpha g \, d\Omega + \int_\Gamma \nabla_\alpha c \, d\Gamma$$
$$+ \int_\Omega \psi (N'[u]v) \, d\Omega + \int_\Gamma (B'[u]v)(C'[u]^* \psi) \, d\Gamma$$

But

$$\frac{DN}{D\alpha} = \nabla_\alpha N + N'[u] \frac{Du}{D\alpha} = \nabla_\alpha N + N'[u] v = 0$$

$$\frac{DB}{D\alpha} = \nabla_\alpha B + B'[u] \frac{Du}{D\alpha} = \nabla_\alpha B + B'[u] v = 0$$

Connection to Sensitivity Analysis (cont.)

$$\frac{DJ}{D\alpha} = \int_{\Omega} \nabla_{\alpha} g \, d\Omega + \int_{\Gamma} \nabla_{\alpha} c \, d\Gamma$$

$$- \int_{\Omega} \psi \, \nabla_{\alpha} N \, d\Omega - \int_{\Gamma} (\nabla_{\alpha} B) \, C'[u]^{*} \psi \, d\Gamma$$

Total derivative has been reduced to a partial derivative.

References