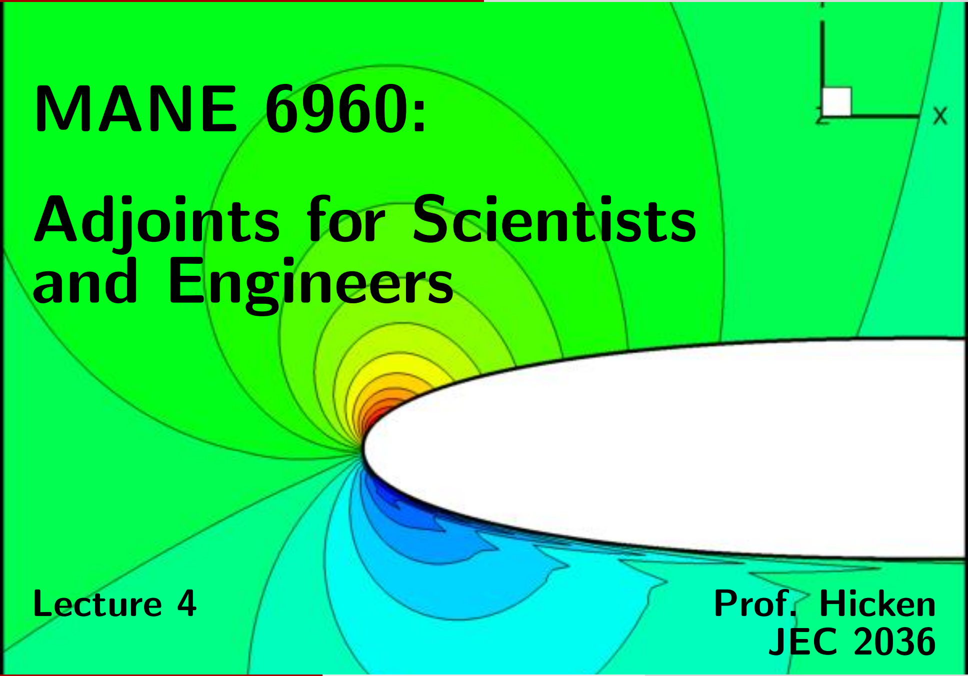


MANE 6960:

Adjoints for Scientists and Engineers

Lecture 4

Prof. Hicken
JEC 2036

The Continuous and Discrete Approaches to the Adjoint

Lecture Objective

Our objective in this lecture is to summarize the two approaches of deriving the (discretized) adjoint equations. We will also start to examine how these approaches differ.

Over the last two classes, we have learned how to derive the following adjoint boundary-value problem given a primal boundary-value problem and a functional.

$$\begin{aligned} L^*\psi &= g, \qquad \forall x \in \Omega, \\ B^*\psi &= c, \qquad \forall x \in \Gamma, \end{aligned} \tag{Adj}$$

Continuous-Adjoint Approach

As with the primal BVP, we cannot solve the adjoint BVP analytically, at least in general. Therefore, we need to discretize the adjoint PDE and BCs.

- you can apply your "favorite" discretization to discretize the adjoint BVP
- however, we shall see next class that some methods are better than others

We will denote the discretized adjoint BVP as

$$(L^*)_h \psi_h = g_h$$

Continuous-Adjoint Approach (cont.)

This sequence — derive the adjoint BVP then discretize it — defines the continuous-adjoint approach.

Definition: Continuous-Adjoint Approach

The continuous-adjoint approach consists of deriving the adjoint BVP first, and then discretizing the adjoint BVP.

- The continuous-adjoint method is also known as the "differentiate-then-discretize" approach or the "optimize-then-discretize" approach.

Discrete Adjoint Approach

The adjoint derivation in the first lecture followed the so-called discrete-adjoint approach. Let's review it again here, following the method we used to derive the continuous adjoint.

Suppose we are given a primal BVP:

$$\begin{aligned} Lu &= f, \qquad \forall x \in \Omega, \\ Bu &= b, \qquad \forall x \in \Gamma. \end{aligned}$$

After discretization, we get the linear system

$$\underbrace{L_h u_h = f_h.}_{\text{This contains both the PDE and the BC discretizations}}$$

Discrete Adjoint Approach (cont.)

We have to discretize the functional as well to derive the discrete adjoint. Recall that the general functional had the form

$$J(u) = (g, u)_\Omega + (c, Cu)_\Gamma$$

The discrete version of this is

$$J_h(u_h) = (g_h, u_h)_h = g_h^T u_h$$

- Like f_h in the primal discretization, g_h combines both volume and boundary terms, i.e. g and c.
- Depending on how $(,)_h$ is defined, g_h will also hold quadrature weights.

Discrete Adjoint Approach (cont.)

As in the continuous BVP, we also have duality in the discrete case:

$$\begin{aligned} J_h(u_h) &= (g_h, u_h)_h \\ &= (g_h, u_h)_h - (\psi_h, \overbrace{L_h u_h - f_h}^{=0})_h \\ &= (g_h, u_h)_h - (\psi_h, L_h u_h)_h + (\psi_h, f_h)_h \\ &= (f_h, \psi_h)_h - (u_h, L_h^T \psi_h - g_h)_h \end{aligned}$$

true for all ψ_h

Setting $L_h^T \psi_h = g_h$, we get that

$$\begin{aligned} J_h(u_h) &= (g_h, u_h)_h \\ &= (f_h, \psi_h)_h \\ &\equiv J_h^*(\psi_h) \end{aligned}$$

Discrete Adjoint Approach (cont.)

This sequence — discretize the primal BVP and functional then derive the adjoint — defines the discrete adjoint approach.

Definition: Discrete-Adjoint Approach

The discrete-adjoint approach refers to discretizing the primal BVP and functional first, and then deriving the adjoint.

- The discrete-adjoint method is also known as the "discretize-then-differentiate" approach or the "discretize-then-optimize" approach.

Summary: Paths to the Adjoint

Primal BVP and functional

$$Lu = f$$
$$Bu = b$$
$$J(u) = (g, u)_\Omega + (c, u)_\Gamma$$

Discretize Primal

$$L_h u_h = f_h$$
$$J_h(u_h) = (g_h, u_h)_h$$

Derive Adjoint BVP

$$L^* \psi = g$$
$$B^* \psi = c$$

Discrete Adjoint

$$L_h^T \psi_h = g_h$$

$\overset{?}{=}$

Continuous Adjoint

$$(L^*)_h \psi_h = g_h$$

these are not equal, in general

Summary: Paths to the Adjoint (cont.)

It is important to recognize that the two paths may produce different solutions. In other words

$$(L^*)_h \neq L_h^T$$

in general.

- The right-hand sides of the adjoint equation, g_h, will also differ in general.

More on this topic next class.

Example

Consider the Poisson's equation in one dimension:

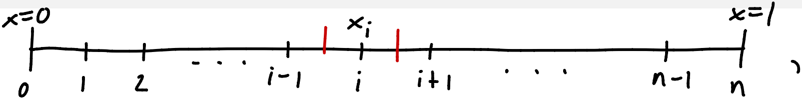
$$Lu = \frac{d}{dx}\left(\nu \frac{du}{dx}\right) = f, \qquad \forall x \in [0,1]$$

$$Bu = \begin{cases} u, & x = 0 \\ \frac{du}{dx}, & x = 1 \end{cases} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

where $\nu(x) > 0$ is a spatially varying diffusion coefficient.

1. Derive the discrete-adjoint equation; use a second-order accurate finite-difference method for the primal BVP
2. Compare the resulting adjoint discretization with the continuous BVP

Example: Primal BVP Discretization



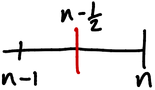
$h = 1/n$

$$\left[\frac{d}{dx}\left(\nu \frac{du}{dx}\right) - f\right]_i \approx \frac{1}{h}\left[\nu_{i+\frac{1}{2}}\left(\frac{u_{i+1}-u_i}{h}\right) - \nu_{i-\frac{1}{2}}\left(\frac{u_i - u_{i-1}}{h}\right)\right] - \underbrace{f_i}_{f(x_i)}$$

$$= \underbrace{\left(\frac{\nu_{i-\frac{1}{2}}}{h^2}\right)}_{(L_h)_{i,i-1}} u_{i-1} - \underbrace{\left(\frac{\nu_{i-\frac{1}{2}} + \nu_{i+\frac{1}{2}}}{h^2}\right)}_{(L_h)_{i,i}} u_i + \underbrace{\left(\frac{\nu_{i+\frac{1}{2}}}{h^2}\right)}_{(L_h)_{i,i+1}} u_{i+1} - f_i$$

$$(Bu)_{i=0} = u_0 = 0 \quad \Rightarrow \quad (L_h)_{1,1} = 1$$

$$(Bu)_{i=n} = \left(\frac{du}{dx}\right) = 0$$



$$\frac{1}{h/2}\left[\nu_n\left(\frac{du}{dx}\right) - \nu_{n-\frac{1}{2}}\left(\frac{u_n - u_{n-1}}{h}\right)\right] - f_n = 0$$

Example: Primal BVP Discretization (cont.)

$$
\begin{bmatrix}
1 & & & & & \\
\frac{\nu_{\frac{1}{2}}}{h^2} & -\frac{\left(\nu_{\frac{1}{2}} + \nu_{\frac{3}{2}}\right)}{h^2} & \frac{\nu_{\frac{3}{2}}}{h^2} & & & \\
& \ddots & \ddots & \ddots & & \\
& & \frac{\nu_{n-\frac{3}{2}}}{h^2} & -\frac{\left(\nu_{n-\frac{3}{2}} + \nu_{n-\frac{1}{2}}\right)}{h^2} & \frac{\nu_{n-\frac{1}{2}}}{h^2} \\
& & & \frac{2\nu_{n-\frac{1}{2}}}{h^2} & -\frac{2\nu_{n-\frac{1}{2}}}{h^2}
\end{bmatrix}
\begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n \end{bmatrix}
=
\begin{bmatrix} 0 \\ f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{bmatrix}
$$

Example: Functional Discretization

Recall that the general (compatible) functional in this case is

$$J(u) = \int_\Omega g u \, d\Omega - c_0 \left(\nu \frac{du}{dx} \right)_{x=0} - c_1 \left(\nu u \right)_{x=1}$$

This can be discretized as

$$J_h(u_h) = \sum_{i=0}^{n} g(x_i)\, u_i\, w_i \; - c_0 \left(\nu_0 \frac{u_1 - u_0}{h} \right) - c_1 \nu_n u_n$$

where
$$w_i = \begin{cases} h, & i = 1, 2, \ldots, n-1 \\ h/2, & i = 0, n \end{cases}$$

ie. we use trapezoid quadrature

Example: Functional Discretization (cont.)

We need to get the discrete functional in the form

$$J_h(u_h) = (g_h, u_h)_h = g_h^T u_h$$

$$(g_h)_i = g(x_i)\,h \quad , \quad i = 2, 3, \ldots, n-1$$

$$(g_h)_0 = g(x_0)\frac{h}{2} + \frac{c_0 \nu_0}{h}$$

$$(g_h)_1 = g(x_1)\,h - \frac{c_0 \nu_0}{h}$$

$$(g_h)_n = g(x_n)\frac{h}{2} - c_1 \nu_n$$

Example: Discrete Adjoint

The bilinear identity gives us

$$
\begin{aligned}
(\psi_h, L_h u_h)_h &= \sum_{i=0}^{n} \psi_i (L_h u_h)_i \\
&= \sum_{i=0}^{n} \psi_i \sum_{j=0}^{n} (L_h)_{ij} u_j \\
&= \sum_{j=0}^{n} u_j \sum_{i=0}^{n} (L_h)_{ij} \psi_i \\
&= \sum_{j=0}^{n} u_j (L_h^T \psi_h)_j = (u_h, L_h^T \psi_h)_h
\end{aligned}
$$

i.e. $\quad L_h^T \psi_h = g_h$

Example: Discrete Adjoint (cont.)

$$
\underbrace{\begin{bmatrix}
1 & \frac{\nu_{\frac{1}{2}}}{h^2} & & \\
0 & -\frac{\left(\nu_{\frac{1}{2}} + \nu_{\frac{3}{2}}\right)}{h^2} & \frac{\nu_{\frac{3}{2}}}{h^2} & \\
 & \ddots & \ddots & \ddots \\
 & \frac{\nu_{n-\frac{3}{2}}}{h^2} & -\frac{\left(\nu_{n-\frac{3}{2}} + \nu_{n-\frac{1}{2}}\right)}{h^2} & \frac{2\nu_{n-\frac{1}{2}}}{h^2} \\
 & & \frac{\nu_{n-\frac{1}{2}}}{h^2} & -\frac{2\nu_{n-\frac{1}{2}}}{h^2}
\end{bmatrix}}_{L_h^T}
\underbrace{\begin{bmatrix}
\psi_0 \\ \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{n-1} \\ \psi_n
\end{bmatrix}}_{\psi_h}
=
\underbrace{\begin{bmatrix}
g_0 \frac{h}{2} + \frac{c_0 \nu_0}{h} \\
g_1 h - \frac{c_0 \nu_0}{h} \\
g_2 h \\
\vdots \\
g_{n-1} h \\
g_n \frac{h}{2} - c_1 \nu_n
\end{bmatrix}}_{g_h}
$$

Example: Continuous Adjoint

Recall that the adjoint BVP was self-adjoint; therefore, the natural discretization for the adjoint BVP is the same one used for the primal BVP:

$$(L^*)_h = L_h$$

And the right-hand side would be given by

$$g_h^T = \begin{bmatrix} c_0 & g_1 & g_2 & \dots & g_{n-1} & g_n - \frac{2\nu_n c_1}{h} \end{bmatrix}$$

Comparing $(L^*)_h$ with L_h^T, it is clear that they differ. The same is true for the g_h.

- Let's analyze the discrete adjoint to see what it approximates.

Example: Adjoint, Interior Scheme

$$i = 2, 3, \ldots, n-1$$

$$(L_h^T \psi_h)_i = \left(\frac{\nu_{i-\frac{1}{2}}}{h^2}\right)\psi_{i-1} - \left(\frac{\nu_{i-\frac{1}{2}} + \nu_{i+\frac{1}{2}}}{h^2}\right)\psi_i + \left(\frac{\nu_{i+\frac{1}{2}}}{h^2}\right)\psi_{i+1} = g(x_i)\,h$$

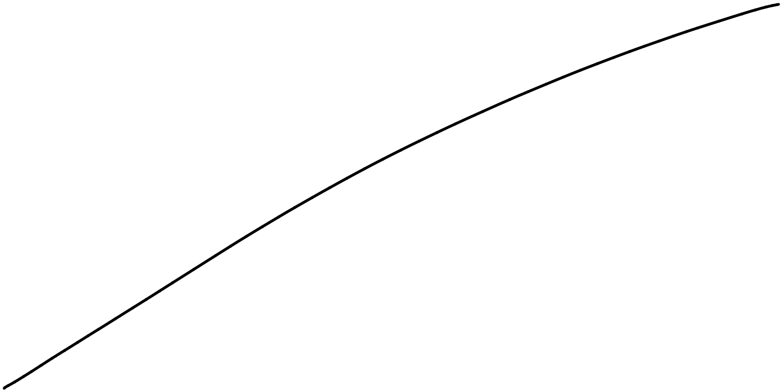
as $h \to 0$, $RHS \to 0$

Correct by defining $\overline{\psi_i} = \begin{cases} \psi_i / h & , i = 1, 2, \ldots n-1 \\ \psi_i/(h/2) & , i = 0, n \end{cases}$

$$\Rightarrow \left(\frac{\nu_{i-\frac{1}{2}}}{h}\right)\overline{\psi}_{i-1} - \left(\frac{\nu_{i-\frac{1}{2}} + \nu_{i+\frac{1}{2}}}{h}\right)\overline{\psi}_i + \left(\frac{\nu_{i+\frac{1}{2}}}{h}\right)\overline{\psi}_{i+1} = g(x_i)\,h$$

$$\approx \left[h \frac{d}{dx}\left(\nu \frac{d\psi}{dx}\right) = g\,h\right]_{x=x_i} , \quad i = 2, 3, \ldots, n-1$$

Example: Adjoint, Interior Scheme (cont.)



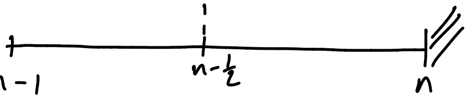
Example: Adjoint, Right Boundary

$i = n$:

$$\left(\frac{\nu_{n-\frac{1}{2}}}{h^2}\right)\psi_{n-1} - \left(\frac{2\nu_{n-\frac{1}{2}}}{h^2}\right)\psi_n = g_n\frac{h}{2} - c_1\nu_n$$

First, introduce scaled adjoint $\bar{\Psi}_i$:

$$\left(\frac{\nu_{n-\frac{1}{2}}}{h}\right)\bar{\Psi}_{n-1} - \left(\frac{\nu_{n-\frac{1}{2}}}{h}\right)\bar{\Psi}_n + (\nu_n)c_1 = g_n\frac{h}{2}$$



$$\left[\nu\frac{d\bar{\Psi}}{dx}\right]_{n-\frac{1}{2}} \approx \nu_{n-\frac{1}{2}}\left(\frac{\Psi_n - \Psi_{n-1}}{h}\right), \quad \left[\nu\frac{d\bar{\Psi}}{dx}\right]_n = \nu_n c_1$$

This is a first order discretization of $\frac{d}{dx}\left(\nu\frac{d\bar{\Psi}}{dx}\right) = g$ at $x = 1$

Example: Adjoint, Left Boundary

$i = 1$:

$$-\left(\frac{\nu_{\frac{1}{2}} + \nu_{\frac{3}{2}}}{h^2}\right)\psi_1 + \left(\frac{\nu_{\frac{3}{2}}}{h^2}\right)\psi_2 = g_1 h - \frac{c_0 \nu_0}{h}$$

First, insert scaled adjoint and rearrange:

$$\underbrace{\left(\nu_{\frac{3}{2}}\right)\left(\frac{\bar{\psi}_2 - \bar{\psi}_1}{h}\right) - \left(\frac{\nu_{\frac{1}{2}}\bar{\psi}_1 - \nu_0 c_0}{h}\right)} = g_1 h$$

$\nu_0 \rightarrow$ $O(h)$ approx to $\nu_{\frac{1}{2}}$

$$\therefore \quad h\frac{d}{dx}\left(\nu\frac{d\psi}{dx}\right)_{i=1} + O(h) = g_1 h$$

Example: Adjoint, Left Boundary (cont.)

$i=0$:

$$\psi_0 + \left(\frac{\nu_{\frac{1}{2}}}{h^2}\right)\psi_1 = g_0\frac{h}{2} + \frac{c_0\nu_0}{h}$$

First, insert scaled variables

$$\frac{h}{2}(\bar{\psi}_0) + \left(\frac{\nu_{\frac{1}{2}}}{h}\right)\bar{\psi}_1 = g_0\frac{h}{2} + \frac{c_0\nu_0}{h}$$

$$\Rightarrow \quad \nu_{\frac{1}{2}}\bar{\psi}_1 - \nu_0 c_0 = \frac{h^2}{2}(g_0 - \bar{\psi}_0) = O(h^2)$$

$$\Rightarrow \quad \nu_0(\bar{\psi}_0 - c_0) + O(h) = O(h^2)$$

$$\Rightarrow \quad \bar{\psi}_0 = c_0 \quad , \text{ as } \quad h \to 0$$

References