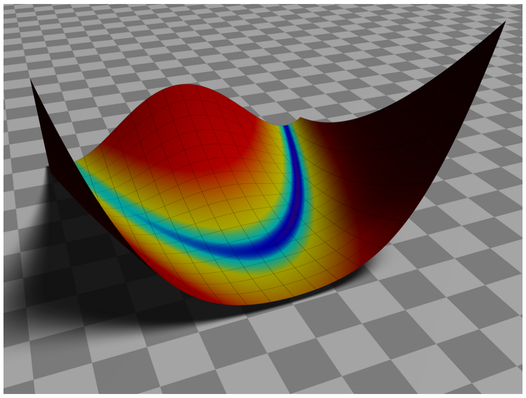
Direct/Adjoint Methods

Lecture 11



ME EN 575
Andrew Ning
aning@byu.edu

Outline

Analytic Sensitivity Equations

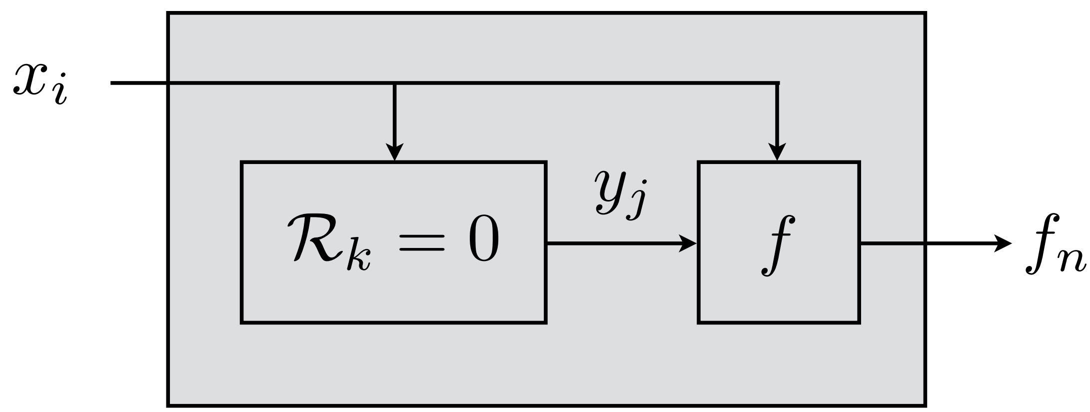
Direct/Adjoint

x_i : design variables

y_j : state variables

$\mathcal{R}_k$: residuals

f_n : outputs (objectives and constraints)



Our end goal is to get

$$\frac{df_n}{dx_i}$$

for all i and n.

$$f_n = f(x_i, y_j(x_i))$$

$$\frac{df_n}{dx_i} = \frac{\partial f_n}{\partial x_i} + \frac{\partial f_n}{\partial y_j}\frac{dy_j}{dx_i}$$

$$\mathcal{R}(x_i, y_j(x_i)) = 0$$

$$\frac{d\mathcal{R}_k}{dx_i} = \frac{\partial \mathcal{R}_k}{\partial x_i} + \frac{\partial \mathcal{R}_k}{\partial y_j}\frac{dy_j}{dx_i} = 0$$

Two equations:

$$\frac{df_n}{dx_i} = \frac{\partial f_n}{\partial x_i} + \frac{\partial f_n}{\partial y_j}\boxed{\frac{dy_j}{dx_i}}$$

$$\frac{d\mathcal{R}_k}{dx_i} = \frac{\partial \mathcal{R}_k}{\partial x_i} + \frac{\partial \mathcal{R}_k}{\partial y_j}\boxed{\frac{dy_j}{dx_i}} = 0$$

Rearrange second equation:

$$\frac{\partial \mathcal{R}_k}{\partial y_j}\frac{dy_j}{dx_i} = -\frac{\partial \mathcal{R}_k}{\partial x_i}$$

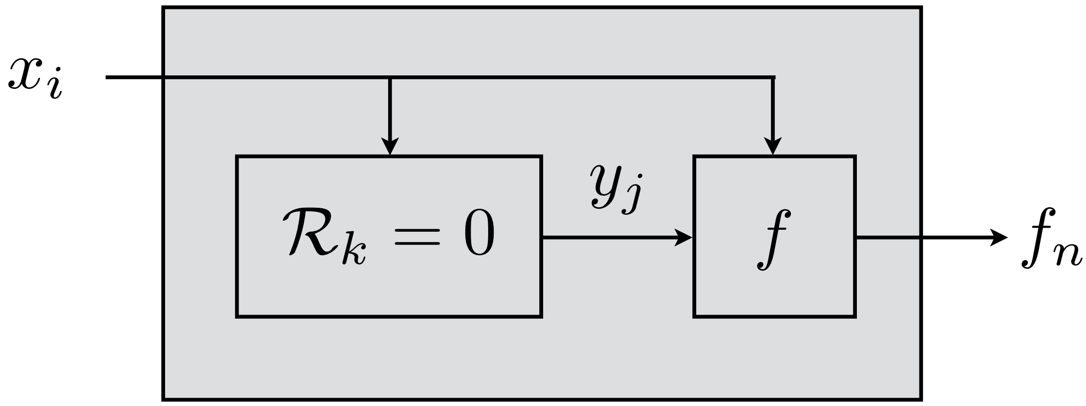
Sub into first:

$$\frac{df_n}{dx_i} = \frac{\partial f_n}{\partial x_i} - \frac{\partial f_n}{\partial y_j}\left[\frac{\partial \mathcal{R}_k}{\partial y_j}\right]^{-1}\frac{\partial \mathcal{R}_k}{\partial x_i}$$

Analytic Sensitivity Equations

$$\frac{df_n}{dx_i} = \frac{\partial f_n}{\partial x_i} - \frac{\partial f_n}{\partial y_j}\left[\frac{\partial \mathcal{R}_k}{\partial y_j}\right]^{-1}\frac{\partial \mathcal{R}_k}{\partial x_i}$$

- Can get total derivatives from partial derivatives that are much more easily obtained.
- We don't actually invert a matrix. We often don't even store the factorization.
- Order of operations is extremely important (more on this next).



Two ways to solve. Partial derivatives are always the same, but order of operations is not.

$$\frac{df_n}{dx_i} = \frac{\partial f_n}{\partial x_i} - \underbrace{\frac{\partial f_n}{\partial y_j} \overbrace{\left[\frac{\partial \mathcal{R}_k}{\partial y_j}\right]}^{\Phi}}_{\Psi}{}^{-1} \frac{\partial \mathcal{R}_k}{\partial x_i}$$

Direct method:

$$\frac{\partial \mathcal{R}_k}{\partial y_j}\Phi_j = -\frac{\partial \mathcal{R}_k}{\partial x_i}$$

then:

$$\frac{df_n}{dx_i} = \frac{\partial f_n}{\partial x_i} + \frac{\partial f_n}{\partial y_j}\Phi_j$$

One linear solve for each input x_i,
but can reuse for all outputs f_n

Adjoint method:

$$\left[\frac{\partial \mathcal{R}_k}{\partial y_j}\right]^T \Psi_k = -\frac{\partial f_n}{\partial y_j}.$$

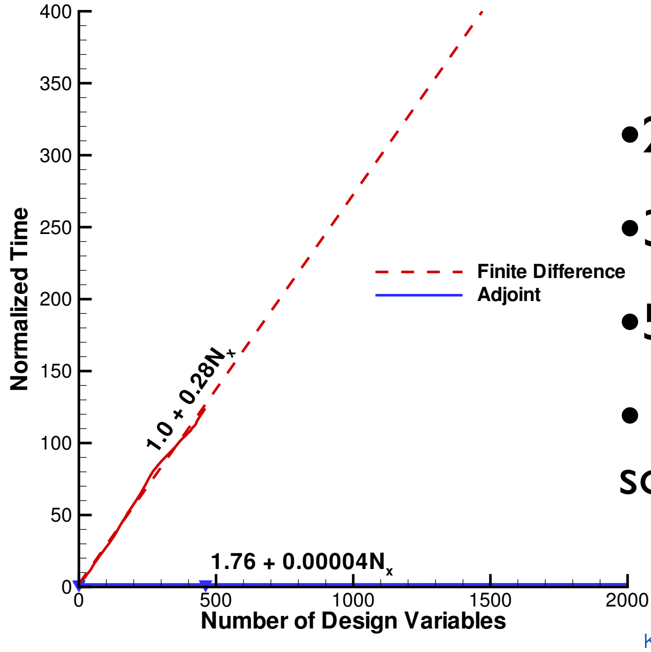
then

$$\frac{df_n}{dx_i} = \frac{\partial f_n}{\partial x_i} + \Psi_k^T \frac{d\mathcal{R}_k}{dx_i}$$

One linear solve for each output f_n,
but can reuse for all inputs x_i

Step	Direct	Adjoint
Matrix Factorization	same	same
Back-solve	N_x times	N_f times
Multiplication	same	same

Important implication: For the adjoint method, computing gradients is practically independent of the number of design variables.



- 2M CFD cells
- 300k CSM DOFs
- 56 processors
- 1 aerostructural solution = 5.5 min

Kenway, Kennedy and Martins, AIAA Journal, 2012

Example: Finite Element Analysis

Force-displacment relationship:

$$F = Ku$$

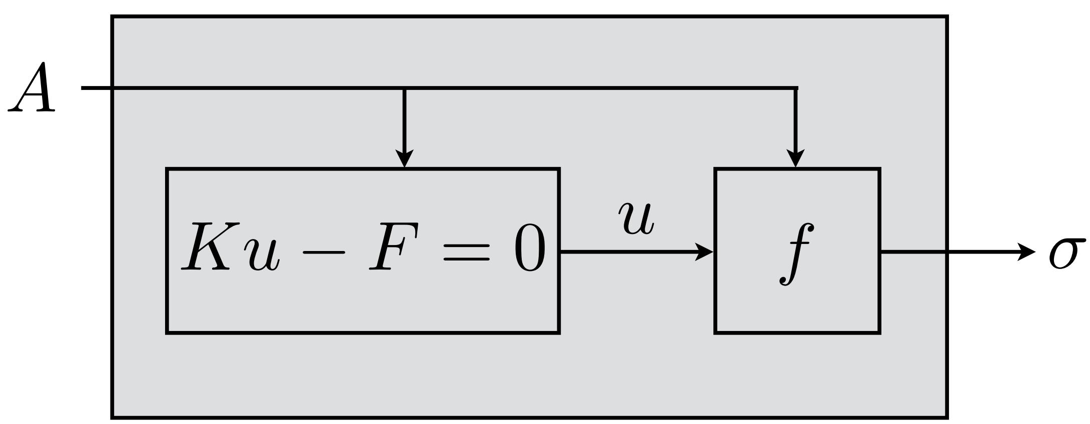
Residual form:

$$\mathcal{R}(u) = Ku - F = 0$$

Stress:

$$\sigma = Su$$

Can change cross-sectional areas of truss.



x_i : design variables

y_j : state variables

$\mathcal{R}_k$: residuals

f_n : outputs (objectives and constraints)

A : design variables (cross-sectional areas)

y_j : state variables

$\mathcal{R}_k$: residuals

f_n : outputs (objectives and constraints)

$$
\begin{aligned}
A &: \text{design variables (cross-sectional areas)} \\
u &: \text{state variables (deflections)} \\
\mathcal{R}_k &: \text{residuals} \\
f_n &: \text{outputs (objectives and constraints)}
\end{aligned}
$$

$$
\begin{aligned}
A &: \text{design variables (cross-sectional areas)} \\
u &: \text{state variables (deflections)} \\
Ku - F = 0 &: \text{residuals (stiffness relationship)} \\
f_n &: \text{outputs (objectives and constraints)}
\end{aligned}
$$

$$
\begin{aligned}
A &: \text{design variables (cross-sectional areas)} \\
u &: \text{state variables (deflections)} \\
Ku - F = 0 &: \text{residuals (stiffness relationship)} \\
\sigma &: \text{outputs (stress)}
\end{aligned}
$$

Partial derivatives:

$$
\begin{aligned}
\frac{\partial f}{\partial x} &= \\
\frac{\partial f}{\partial y} &= \\
\frac{\partial \mathcal{R}}{\partial x} &= \\
\frac{\partial \mathcal{R}}{\partial y} &=
\end{aligned}
$$

Partial derivatives:

$$\frac{\partial f}{\partial x} = \frac{\partial \sigma}{\partial A} = 0$$

$$\frac{\partial f}{\partial y}$$

$$\frac{\partial \mathcal{R}}{\partial x}$$

$$\frac{\partial \mathcal{R}}{\partial y}$$

Partial derivatives:

$$\frac{\partial f}{\partial x} = \frac{\partial \sigma}{\partial A} = 0$$

$$\frac{\partial f}{\partial y} = \frac{\partial \sigma}{\partial u} = S$$

$$\frac{\partial \mathcal{R}}{\partial x}$$

$$\frac{\partial \mathcal{R}}{\partial y}$$

Partial derivatives:

$$\frac{\partial f}{\partial x} = \frac{\partial \sigma}{\partial A} = 0$$

$$\frac{\partial f}{\partial y} = \frac{\partial \sigma}{\partial u} = S$$

$$\frac{\partial \mathcal{R}}{\partial x} = \frac{\partial \mathcal{R}}{\partial A} = \left[\frac{\partial K}{\partial A}\right] u$$

$$\frac{\partial \mathcal{R}}{\partial y}$$

Partial derivatives:

$$\frac{\partial f}{\partial x} = \frac{\partial \sigma}{\partial A} = 0$$

$$\frac{\partial f}{\partial y} = \frac{\partial \sigma}{\partial u} = S$$

$$\frac{\partial \mathcal{R}}{\partial x} = \frac{\partial \mathcal{R}}{\partial A} = \left[\frac{\partial K}{\partial A}\right] u$$

$$\frac{\partial \mathcal{R}}{\partial y} = \frac{\partial \mathcal{R}}{\partial u} = K$$

$$\frac{df_n}{dx_i} = \frac{\partial f_n}{\partial x_i} - \frac{\partial f_n}{\partial y_j}\left[\frac{\partial \mathcal{R}_k}{\partial y_j}\right]^{-1}\frac{\partial \mathcal{R}_k}{\partial x_i}$$

$$\frac{d\sigma}{dA} = \frac{\partial \sigma}{\partial A} - \frac{\partial \sigma}{\partial u}\left[\frac{\partial \mathcal{R}_k}{\partial u}\right]^{-1}\frac{\partial \mathcal{R}_k}{\partial A}$$

$$\frac{d\sigma}{dA_i} = -SK^{-1}\left[\frac{\partial K}{\partial A_i}\right]u$$