(Adjoint) Sensitivities of a Linear System

Linear System

reference solution x_r

$$\underline{\underline{A}}\ \underline{x} = \underline{b}(\underline{\theta}) \xrightarrow{\text{solve}} \underline{x} \longrightarrow \text{loss function } J = \frac{1}{2}(\underline{x}-\underline{x}_r)^T(\underline{x}-\underline{x}_r)$$

sensitivities

$$\frac{dJ}{d\underline{\theta}} = ?$$

Here:

$\underline{\underline{A}}$ fixed
$$\underline{\underline{A}} = \begin{bmatrix} 10 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

$\underline{b}$ variable (parameter-dependent)
$$\underline{b} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix}$$

but we have guess $\hat{\underline{b}} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

[A] Solve "classical" problem

[1] $\boxed{\text{Solve } \underline{\underline{A}}\,\underline{x} = \hat{\underline{b}} \text{ for } \underline{x}}$

[2] $\boxed{\text{Evaluate } J = \frac{1}{2}(\underline{x}-\underline{x}_r)^T(\underline{x}-\underline{x}_r)}$

[B] Obtain gradients: $\frac{dJ}{d\underline{\theta}}$

(Note: gradient is a row vector)

1st option: Finite Differences

$$\frac{dJ}{d\underline{\theta}} = \begin{bmatrix} \frac{dJ}{d\theta_0} & \frac{dJ}{d\theta_1} & \frac{dJ}{d\theta_2} \end{bmatrix}$$

for $\frac{dJ}{d\theta_0}$: $\tilde{\underline{b}} = \hat{\underline{b}} + \varepsilon\, \underline{e}_0$ (usually $1e^{-6}$)

Solve $\underline{\underline{A}}\,\tilde{\underline{x}} = \tilde{\underline{b}}$ for $\tilde{\underline{x}}$

Evaluate: $\tilde{J} = \frac{1}{2}(\tilde{\underline{x}}-\underline{x}_r)^T(\tilde{\underline{x}}-\underline{x}_r)$

$$\boxed{\frac{dJ}{d\theta_0} \approx \frac{\tilde{J}-J}{\varepsilon}}$$

2nd option: Forward Sensitivities:

not correct? wrong shape

Solve $\underline{\underline{A}}\frac{d\underline{x}}{d\underline{\theta}} = \frac{d\underline{b}}{d\underline{\theta}} - \frac{d\underline{\underline{A}}}{d\underline{\theta}}\underline{x}$ for $\frac{d\underline{x}}{d\underline{\theta}}$

here: $\frac{d\underline{b}}{d\underline{\theta}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$\frac{d\underline{\underline{A}}}{d\underline{\theta}} = \underline{\underline{\underline{0}}}$$

$$\boxed{\text{solve } \underline{\underline{A}}\frac{d\underline{x}}{d\underline{\theta}} = \text{eye}(3) \text{ for } \frac{d\underline{x}}{d\underline{\theta}}}$$

then: $\frac{dJ}{d\underline{\theta}} = \frac{\partial J}{\partial \underline{\theta}} + \frac{\partial J}{\partial \underline{x}}\frac{d\underline{x}}{d\underline{\theta}}$

$$\frac{\partial J}{\partial \underline{\theta}} = \underline{0}^T \qquad \frac{\partial \underline{x}}{\partial \underline{x}} = \underline{\underline{I}}_{3\times3} \qquad \frac{\partial \underline{x}^T}{\partial \underline{x}} = \underline{\underline{I}}_{1\times3\times3}$$

$$\frac{\partial J}{\partial \underline{x}} = \frac{\partial}{\partial \underline{x}}\left(\frac{1}{2}(\underline{x}-\underline{x}_r)^T(\underline{x}-\underline{x}_r)\right)$$

$$= \frac{1}{2}\left(\underbrace{\underbrace{\frac{\partial(\underline{x}-\underline{x}_r)^T}{\partial \underline{x}}}_{\underline{\underline{I}}_{1\times3\times3}}\underbrace{(\underline{x}-\underline{x}_r)}_{3}}_{1\times3} + \underbrace{\underbrace{(\underline{x}-\underline{x}_r)^T}_{1\times3}\underbrace{\frac{\partial(\underline{x}-\underline{x}_r)}{\partial \underline{x}}}_{\underline{\underline{I}}_{3\times3}}}_{1\times3}\right)$$

$$= \frac{1}{2}\left((\underline{x}-\underline{x}_r)^T + (\underline{x}-\underline{x}_r)^T\right)$$

$$= \frac{1}{\cancel{2}}\cdot\cancel{2}\cdot(\underline{x}-\underline{x}_r)^T = (\underline{x}-\underline{x}_r)^T$$

$$\boxed{\text{then } \frac{dJ}{d\underline{\theta}} = \underline{0}^T + (\underline{x}-\underline{x}_r)^T\frac{d\underline{x}}{d\underline{\theta}}}$$

3rd option: Adjoint / Backward Sensitivities

Solve $\underline{\underline{A}}^T\underline{\lambda} = \left(\frac{\partial J}{\partial \underline{x}}\right)^T$ for $\underline{\lambda}$

$$\boxed{\text{Solve } \underline{\underline{A}}^T\underline{\lambda} = \underline{x}-\underline{x}_r \text{ for } \underline{\lambda}}$$

then: $\frac{dJ}{d\underline{\theta}} = \underbrace{\frac{\partial J}{\partial \underline{\theta}}}_{\underline{0}^T} + \underline{\lambda}^T\left(\underbrace{\frac{d\underline{b}}{d\underline{\theta}}}_{\text{eye}(3)} - \underbrace{\frac{d\underline{\underline{A}}}{d\underline{\theta}}}_{\underline{\underline{0}}}\underline{x}\right)$, with $\frac{d\underline{\underline{A}}}{d\underline{\theta}}\underline{x} = \underline{\underline{0}}$

$$\boxed{\text{then } \frac{dJ}{d\underline{\theta}} = \underline{0}^T + \underline{\lambda}^T\text{eye}(3)}$$