Sensitivity Analysis

Begin with finite dimensional analysis

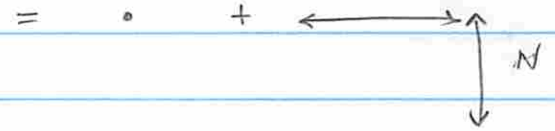
$$\min_{m \in \mathbb{R}^M} f(u(m), m) = \frac{1}{2}(Bu - d)^T (Bu - d) + \frac{1}{2} m^T R m$$

where $r(u,m) = 0$

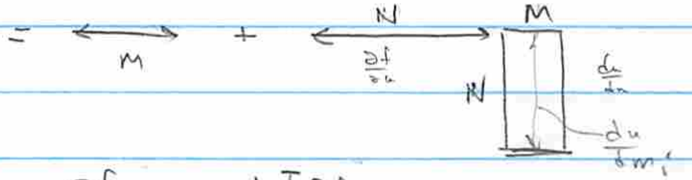
$d \in \mathbb{R}^D$; $u \in \mathbb{R}^N$, $a \in \mathbb{R}^N$; $B \in \mathbb{R}^{D \times N}$; $R \in \mathbb{R}^{M \times M}$

Gradient of f wrt m found from total derivative:

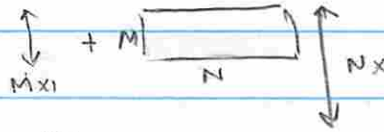
$$g_i = \frac{Df}{Dm_i} = \frac{\partial f}{\partial m_i} + \frac{\partial f}{\partial u}^T \underbrace{\frac{du}{dm_i}}_{\text{sensitivity of state wrt model params}} \quad , i = 1, \ldots, M$$



$$g^T = \frac{\partial f}{\partial m}^T + \frac{\partial f}{\partial u}^T \frac{du}{dm}$$ (derivative are row vectors)



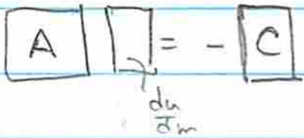
$$g = \frac{\partial f}{\partial m} + \frac{du}{dm}^T \frac{\partial f}{\partial u} =$$



$\frac{\partial f}{\partial m} = Rm$; $\frac{\partial f}{\partial u} = B^T (Bu - d)$

$\frac{du}{dm} = ?$ For: $r(u^{(m)}, m) = 0$

$$\frac{\partial r}{\partial u}\frac{du}{dm} + \frac{\partial r}{\partial m} = 0$$



$$\Rightarrow \frac{du}{dm} = -\left(\frac{\partial r}{\partial u}\right)^{-1} \frac{\partial r}{\partial m}$$

$$= -A(m,u)^{-1} C(m,u)$$

Special case. $A(m)u = b(m)$

$$\frac{\partial (Au)}{\partial m} + A\frac{du}{dm} = \frac{\partial b}{\partial m} \Rightarrow \frac{du}{dm} = A(m)^{-1}\left[\frac{\partial b}{\partial m} - \frac{\partial (Au)}{\partial m}\right]$$

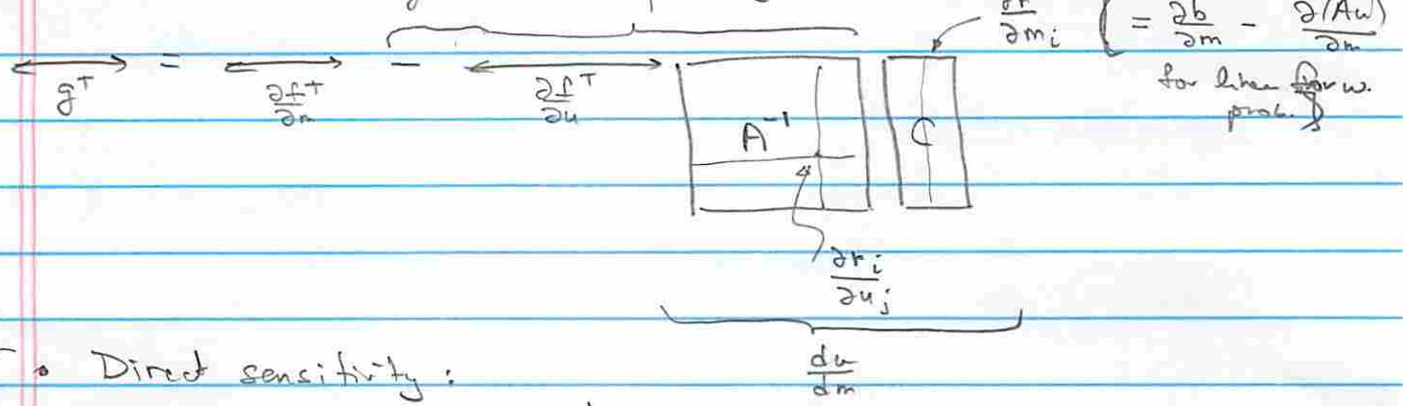
$$g^T = \frac{\partial f}{\partial m}^T + \frac{\partial f}{\partial u}^T \frac{du}{dm}$$

$$= \underset{\sim}{\frac{\partial f}{\partial m}^T} - \underset{\sim}{\frac{\partial f}{\partial u}^T} \boxed{A}^{-1} \underset{\sim}{\frac{\partial r}{\partial m}}$$

Jacobian (of res w.r.t. state); common component of Newton forw solver

straightforward

p^T: adjoint variable



In both cases, first need to solve forward problem $c(u,m)=0$ to obtain u

- Direct sensitivity:

  1) first compute $\frac{du}{dm_i}$, $i = 1, \ldots, M$

     $\Rightarrow$ Solve $A^{-1}C$

     disadvantage: M solves with A needed (not bad if A is factored first) $\Rightarrow$ if iterative solver used, restrict to small M

     advantage: same solver as Newton for forw prob.

  2) then compute $\frac{\partial f}{\partial u}^T \frac{du}{dm}$ and subtract from $\frac{\partial f}{\partial u}^T$

     $\Rightarrow$ cost: linear algebra (M inner products)

- Adjoint sensitivity:

  1) first. $p = -A^{-T} \frac{\partial f}{\partial u} = - \boxed{A^{-T}} \updownarrow \frac{\partial f}{\partial u}$

     i.e. $A^T p = -\frac{\partial f}{\partial u}$ $\Rightarrow$ one solve! (but with A^T)

  2) then $C^T p$ (M inner products, just linear algebra) and add to $\frac{\partial f}{\partial u}$

     advantage: just one solve — independent of M!

     disadvantage: not exactly same operator as Newton forward solver

Adjoint method can be derived via Lagrangian:

define Lagrangian $\mathcal{L}(u,m,p) := \overbrace{f(u,m)}^{\text{objective}} + \overbrace{p}^{\text{Lag. mult.}}{}^{T} \overbrace{r(u,m)}^{\text{forw residual}}$

$p \in \mathbb{R}^N$ adj

1) $\frac{\partial \mathcal{L}}{\partial p} = 0 \Rightarrow$ Solve $r(u,m) = 0$ forward problem

2) $\frac{\partial \mathcal{L}}{\partial u} = 0 \Rightarrow \frac{\partial f}{\partial u}^T + p^T \underbrace{\frac{\partial r}{\partial u}}_{\hookrightarrow A} = 0$

$\Rightarrow$ Solve $A^T p = -\frac{\partial f}{\partial u}$ adjoint problem

3) $\frac{\partial \mathcal{L}}{\partial m} = \text{gradient} \Rightarrow \boxed{\frac{\partial f}{\partial m} + p^T \underbrace{\frac{\partial r}{\partial m}}_{C}}$

Given some m, solve forw; $\nearrow u$ then $\overset{\text{with } u \,\&\, m,}{\text{solve}}$ adjoint; then $\underset{\text{with } m, u, p}{\text{evaluate}}$ gradient

Lagrangian is a device for facilitating gradient computation
$\Rightarrow$ only partial derivatives needed; no dependence of
state variable of parameters!

Infinite dimensions : direct sensitivity

Take model problem:

$$\min_m \mathcal{J}(u(m), m) := \frac{1}{2}\int_\Omega (b(x)u - d)^2 \, dx + \frac{\alpha}{2}\int_\Omega \nabla m \cdot \nabla m \, dx$$

where $-\nabla \cdot (m \nabla u) + v \cdot \nabla u = 0$ in Ω

$u = 0$ on Γ

$u \in H^1_0(\Omega)$; $m \in H^1(\Omega)$ (from regularization term)

First step: BVP $\Rightarrow$ weak form

$$\int_\Omega p\left[-\nabla\cdot(m\nabla u)\overset{v\cdot\nabla u}{\swarrow}-f\right]dx = 0 \quad \forall p\in H_0^1$$

Green's identity:

$$\Rightarrow \int_\Omega\left[m\nabla u\cdot\nabla p - pf\overset{pv\cdot\nabla u}{\downarrow}\right]dx = \int_\Gamma \overset{0}{\cancel{p}}\, m\nabla u\cdot n\, ds = 0 \quad \forall p\in H_0^1$$

$$\Rightarrow \int_\Omega\left[m\nabla u\cdot\nabla p - pf\overset{pv\cdot\nabla u}{\uparrow}\right]dx = 0 \quad \forall p\in H_0^1$$

Replace m with family of variations:

$$m \rightarrow m+\varepsilon\hat{m}\ ; \quad \hat{m}\in H^1$$

this leads to $u\rightarrow u+\varepsilon\hat{u}$ $\quad \hat{u}\in H_0^1$

$$\Rightarrow \int_\Omega\left[(m+\varepsilon\hat{m})(\nabla u+\varepsilon\nabla\hat{u})\cdot\nabla p + pv\cdot\nabla(u+\varepsilon\hat{u}) - pf\right]dx = 0$$

$$\forall p\in H_0^1$$

Now take derivative wrt ε:

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} = \underbrace{\int_\Omega\Big[\overbrace{\hat{m}\nabla u\cdot\nabla p}^{\eta^T\frac{\partial r}{\partial m}} + \overbrace{m\nabla\hat{u}\cdot\nabla p + pv\cdot\nabla\hat{u}}^{\eta^T A\frac{du}{dm}}\Big]dx = 0}_{\text{weak form of direct sensitivity equation}} \quad \forall p\in H_0^1$$

Solve for $\hat{u}$ given $\hat{m}$ and u

Strong form:
$$\int_\Omega p\left[-\nabla\cdot(\hat{m}\nabla u) - \nabla\cdot(m\nabla\hat{u}) + v\cdot\nabla\hat{u}\right]dx + \underbrace{\int_\Gamma\left(p\hat{m}\nabla u\cdot n + pm\nabla\hat{u}\cdot n\right)ds}_{0} = 0 \quad \forall p\in H_0^1$$

$$\Rightarrow \boxed{\begin{array}{c} -\nabla\cdot(m\nabla\hat{u}) + v\cdot\nabla\hat{u} = +\nabla\cdot(\hat{m}\nabla u) \quad \text{in } \Omega \\ \hat{u} = 0 \quad \text{on } \Gamma \end{array}}$$

strong form of sensitivity equation

5)

Do same for objective function:

$$\mathcal{J}(u+\varepsilon\hat{u},\ m+\varepsilon\hat{m}) = \frac{1}{2}\int_\Omega \left(b(x)(u+\varepsilon\hat{u}) - d\right)^2 dx$$

$$+ \frac{\alpha}{2}\int_\Omega (\nabla m + \varepsilon\nabla\hat{m})\cdot(\nabla m + \varepsilon\nabla\hat{m})\, dx$$

$$\underbrace{\left.\frac{d\mathcal{J}}{d\varepsilon}\right|_{\varepsilon=0}} = \int_\Omega \left(b(x)(u+\cancel{\varepsilon\hat{u}}^{\,0}) - d\right) b(x)\hat{u}\, dx + \alpha\int_\Omega \nabla\hat{m}\cdot\nabla m\, dx$$

Frechet derivative of $\mathcal{J}$ at m in direction of $\hat{m}$

[infinite dimensional version of $g^T\hat{m}$]

$$\overbrace{\underbrace{\int_\Omega (bu-d)b\hat{u}\, dx}_{\sim \frac{\partial f}{\partial u}\frac{du}{dm}} + \underbrace{\alpha\int_\Omega \nabla\hat{m}\cdot\nabla m\, dx}_{\sim \frac{\partial f}{\partial m}}}^{\text{weak form of "gradient"}} \equiv \text{gradient } \forall \hat{m}\in H^1.$$

Adjoint method in infinite dimensions

weak form of adjoint eqn

$$\underbrace{\int_\Omega [m\nabla\hat{u}\cdot\nabla p + p v\cdot\nabla\hat{u}]\, dx}_{\sim p^T\left(A\frac{du}{dm}\right)} = \underbrace{-\int_\Omega (bu-d)b\hat{u}\, dx}_{\sim \frac{\partial f}{\partial u}\frac{du}{dm}} \quad \forall \hat{u}\in H^1_0$$

Solve for p

Then gradient (given $u, p, \hat{m}$):

$$\int_\Omega \hat{m}\nabla u\cdot\nabla p\, dx + \alpha\int_\Omega \nabla\hat{m}\cdot\nabla m\, dx \equiv \underbrace{\text{grad.}}_{\text{Frechet derivative}} \forall \hat{m}\in H^1$$

Strong form of adjoint eqn:

$$\Rightarrow -\nabla\cdot(m\nabla p) - \nabla\cdot(p\vec{v}) = -b(bu-d) \text{ in } \Omega$$

$$p = 0 \text{ on } \Gamma$$

Strong form of gradient

$$-\alpha\Delta m \cancel{+} - \nabla u\cdot\nabla p \text{ in } \Omega\ ; \quad \alpha\frac{\partial m}{\partial n} = 0 \text{ on } \Gamma$$

Lagrangian approach to adjoint method:

$$\mathcal{L}(u,p,m) = \frac{1}{2}\int_\Omega (bu-d)^2\,dx + \frac{\alpha}{2}\int \nabla m\cdot\nabla m\,dx$$

$u \in H^1_0$; $p \in H^1_0$; $m \in H^1$

$$+ \underbrace{\int_\Omega \left[m\nabla u\cdot\nabla p + p v\cdot\nabla u - pf \right] dx}_{\text{weak form of forw prob}}$$

Forward eqn: $\delta_p \mathcal{L} \Rightarrow$

Weak form of forw problem $\}$: $\int_\Omega \left[m\nabla u\cdot\nabla\hat{p} + \hat{p} v\cdot\nabla u - \hat{p} f \right] dx = 0 \quad \forall \hat{p} \in H^1_0$

$$\Rightarrow \int_\Omega \hat{p}\left[-\nabla\cdot(m\nabla u) + v\cdot\nabla u - f \right] dx = 0 \quad \forall \hat{p} \in H^1_0$$

$$\Rightarrow \quad -\nabla\cdot(m\nabla u) + v\cdot\nabla u = f \text{ in } \Omega, \quad u = 0 \text{ on } \Gamma$$

Adjoint equation $\delta_u \mathcal{L} \Rightarrow \int_\Omega (bu-d)\,b\hat{u}\,dx + \int_\Omega \left[m\nabla\hat{u}\cdot\nabla p + p v\cdot\nabla\hat{u} \right] dx = 0$

$\forall \hat{u} \in H^1_0$

weak form of adjoint equation

$$\Rightarrow \int \hat{u}\left[b(bu-d) - \nabla\cdot(m\nabla p) - \nabla\cdot(pv) \right] dx = 0 \quad \forall \hat{u} \in H^1_0$$

$$\Rightarrow \quad -\nabla\cdot(m\nabla p) - \nabla\cdot(pv) = -b(bu-d) \text{ in } \Omega$$

$$p = 0 \text{ on } \Gamma$$

Gradient expression: $\delta_m \mathcal{L} \rightarrow \underbrace{\alpha\int_\Omega \nabla m\cdot\nabla\hat{m}\,dx + \int_\Omega \hat{m}\nabla u\cdot\nabla p\,dx}_{\text{weak form of grad}} \equiv \text{grad} \quad \forall \hat{m} \in H^1$

$$\Rightarrow \left.\begin{array}{l} -\alpha\Delta m + \nabla u\cdot\nabla p \text{ in } \Omega \\ \dfrac{\partial m}{\partial \vec{n}} = 0 \text{ on } \Gamma \end{array}\right\} \text{Strong form of gradient}$$

Summary of infinite dimensions:

Direct:

State: find $u \in H^1_0$ s.t. $\int_\Omega [m \nabla u \cdot \nabla p + p v \cdot \nabla u - p f] \, dx = 0 \quad \forall p \in H^1_0$

Sensitivity eqn: Find $\hat{u} \in H^1_0$ s.t. $\int_\Omega [\hat{m} \nabla u \cdot \nabla p + m \nabla \hat{u} \cdot \nabla p + p v \cdot \nabla \hat{u}] \, dx = 0 \quad \forall p \in H^1_0$

Gradient for $\hat{m} \in H^1$: $\int_\Omega (bu - d) b \hat{u} \, dx + \alpha \int_\Omega \nabla \hat{m} \cdot \nabla m \, dx$ (Frechet deriv of $\mathcal{J}$ in direction of $\hat{m}$)

Adjoint

State find $u \in H^1_0$ $\int_\Omega [m \nabla u \cdot \nabla p + p v \cdot \nabla u - p f] \, dx = 0 \quad \forall p \in H^1_0$

Adjoint: Find $p \in H^1_0$ s.t. $\int_\Omega [m \nabla \hat{u} \cdot \nabla p + p v \cdot \nabla \hat{u} + (bu - d) b \hat{u}] \, dx = 0 \quad \forall \hat{u} \in H^1_0$

Gradient $\forall \hat{m} \in H^1$: $\int_\Omega \hat{m} \nabla u \cdot \nabla p \, dx + \alpha \int_\Omega \nabla \hat{m} \cdot \nabla m \, dx$

Hessian:

$$\xleftarrow{\quad} (Bu-d)^T (Bu-d) + m^T R m$$

$$\min_{m \in \mathbb{R}^M} f(u(m), m) \quad \text{s.t.} \quad r(u(m), m) = 0$$

$$\mathcal{L}^G(u, m, p) := f(u, m) + p^T r(u, m)$$

$$\frac{\partial \mathcal{L}^G}{\partial p} = 0 \Rightarrow r(u, m) = 0 \quad \text{forw}$$

$$\frac{\partial \mathcal{L}^G}{\partial u} = 0 \Rightarrow \left(\frac{\partial f}{\partial u}\right)^T + p^T \underset{\hookrightarrow A}{\boxed{\frac{\partial r}{\partial u}}} = 0 \Rightarrow A^T p = -\frac{\partial f}{\partial u} \quad \text{adjoint}$$

$$\frac{\partial \mathcal{L}^G}{\partial m} = 0 \Rightarrow \frac{\partial f}{\partial m}^T + p^T \underset{\hookrightarrow C}{\boxed{\frac{\partial r}{\partial m}}} = 0 \Rightarrow C^T p + \frac{\partial f}{\partial m} = \text{gradient}$$

Hessian? Use Lagrangian idea again

directional derivative: $g^T \tilde{m}$
(in direction $\tilde{m}$)

$$\mathcal{L}^H(u, m, p; \tilde{u}, \tilde{m}, \tilde{p}) := \tilde{m}^T \overbrace{\left[C^T p + \frac{\partial f}{\partial m}\right]}^{g} + \tilde{u}^T\left(\frac{\partial f}{\partial u} + A^T p\right) + \tilde{p}^T r$$

$$\frac{\partial \mathcal{L}^H}{\partial p} = A\tilde{u} + C\tilde{m} = 0 \quad \text{(incremental forward eqn)}$$

$$\frac{\partial \mathcal{L}^H}{\partial u} = \frac{\partial}{\partial u}\left(\tilde{m}^T C^T p\right) + \frac{\partial}{\partial u}\left(\tilde{m}^T \frac{\partial f}{\partial m}\right) + \frac{\partial}{\partial u}\left(\tilde{u}^T \frac{\partial f}{\partial u}\right) + \frac{\partial}{\partial u}\left(\tilde{u}^T A^T p\right) + A^T \tilde{p}$$

$$\Rightarrow A^T \tilde{p} = -\Big(\underbrace{L_{mu}^T}_{L_{um}} \tilde{m} + L_{uu} \tilde{u}\Big) \quad \text{incremental adjoint}$$

$$\frac{\partial \mathcal{L}^H}{\partial m} \Rightarrow \underbrace{H\tilde{m}}_{= \left(\frac{\partial g}{\partial m}\right)\tilde{m}} = \frac{\partial}{\partial m}\left[\tilde{m}^T C^T p\right] + \frac{\partial^2 f}{\partial m \partial m}\tilde{m} + \frac{\partial}{\partial m}\left(\tilde{u}^T \frac{\partial f}{\partial u}\right) + \frac{\partial}{\partial m}\left(\tilde{u}^T A^T p\right) + C^T \tilde{p}$$

$$= L_{mm}\tilde{m} + \underbrace{L_{um}^T}_{L_{mu}}\tilde{u} + C^T \tilde{p}$$

Given m, solve forw ⇒ u; solve adjoint ⇒ p, evaluate gradient

Given m, u, p, and $\tilde{m}$, solve incr. forw ⇒ $\tilde{u}$; solve incr. adj ⇒ $\tilde{p}$

⇒ evaluate $H\tilde{m}$

Note: $\tilde{u} = -A^{-1} C \tilde{m}$

$$\tilde{p} = -A^{-T} L_{um} \tilde{m} - A^{-T} L_{uu} \tilde{u} = -A^{-T} L_{um} \tilde{m} + A^{-T} L_{uu} A^{-1} C \tilde{m}$$

$$H \tilde{m} = L_{mm} \tilde{m} - L_{mu} (A^{-1} C) \tilde{m} - C^T A^{-T} L_{um} \tilde{m} + \underline{C^T A^{-T} L_{uu} A^{-1} C \tilde{m}}$$

$$= \underbrace{\left[C^T A^{-T} L_{uu} A^{-1} C + L_{mm} - L_{mu} A^{-1} C - C^T A^{-T} L_{um} \right]}_{H} \tilde{m}$$

$$L_{uu} = \underbrace{\frac{\partial^2 f}{\partial u \partial u}}_{B^T B} + \frac{\partial}{\partial u} (A^T p)$$

$$L_{mm} = R + \frac{\partial}{\partial m} (C^T p)$$

$$L_{mu} = L_{um}^T = \underbrace{\frac{\partial}{\partial m} \left(\frac{\partial f}{\partial u} \right)}_{= 0} + \frac{\partial}{\partial m} (A^T p)$$

Gauss-Newton approximation of Hessian:

$$p \approx 0 \quad [\text{motivation: adj: } A^T p = -B^T \underbrace{(Bu - d)}_{\approx 0 \text{ for low noise/good model}}$$

$$\Rightarrow L_{uu} = B^T B \; ; \; L_{mm} = R \; ; \; L_{mu} = L_{um}^T = 0$$

$$\Rightarrow H^{GN} = C^T A^{-T} B^T B A^{-1} C + R$$

Infinite dimensions

$$\mathcal{L}^G(u,p,m) := \frac{1}{2}\int_\Omega (bu-d)^2\,dx + \frac{\alpha}{2}\int_\Omega \nabla m\cdot\nabla m\,dx$$

$u \in H_0^1$; $p \in H_0^1$; $m \in H^1$

$$+ \int_\Omega \left[m\nabla u\cdot\nabla p + p v\cdot\nabla u - pf\right] dx$$

$$\delta_p \mathcal{L}^G \rightarrow \text{forw}: \int_\Omega \left[m\nabla u\cdot\nabla\hat{p} + \hat{p}\, v\cdot\nabla u - \hat{p} f\right] dx = 0 \quad \forall\, \hat{p}\in H_0^1$$

$$\delta_u \mathcal{L}^G \rightarrow \text{adj}: \int_\Omega (bu-d)\, b\hat{u}\, dx + \int_\Omega \left[m\nabla\hat{u}\cdot\nabla p + p v\cdot\nabla\hat{u}\right] dx = 0 \quad \forall \hat{u}\in H_0^1$$

$$\delta_m \mathcal{L}^G \rightarrow \text{"grad"}: \quad \alpha\int_\Omega \nabla m\cdot\nabla\hat{m}\,dx + \int_\Omega \hat{m}\nabla u\cdot\nabla p\,dx \qquad \text{grad } \forall\, \hat{m}\in H^1$$

(Frechet derivative in dir $\hat{m}$)

$$\mathcal{L}^H(u,p,m;\tilde{u},\tilde{p},\tilde{m}) := \overbrace{\alpha\int_\Omega \nabla m\cdot\nabla\tilde{m}\,dx + \int_\Omega \tilde{m}\nabla u\cdot\nabla p\,dx}^{\text{Frechet deriv in dir } \tilde{m}}$$

$$+ \underbrace{\int_\Omega \left[m\nabla u\cdot\nabla\tilde{p} + \tilde{p}\, v\cdot\nabla u - \tilde{p} f\right] dx}_{\text{forw}} + \underbrace{\int_\Omega (bu-d)\, b\tilde{u}\,dx + \int_\Omega \left[m\nabla\tilde{u}\cdot\nabla p + p v\cdot\nabla\tilde{u}\right]}_{\text{adj}}$$

incr. adj

$$\delta_u \mathcal{L}^H = \int_\Omega \left[\tilde{m}\nabla\hat{u}\cdot\nabla p + m\nabla\hat{u}\cdot\nabla\tilde{p} + \tilde{p}\, v\cdot\nabla\hat{u} + (b\hat{u})(b\tilde{u})\right] dx = 0 \quad \forall\, \hat{u}\in H_0^1$$

$$\Rightarrow -\nabla\cdot(\tilde{m}\nabla p) - \nabla\cdot(m\nabla\underline{\tilde{p}}) - \nabla\cdot(\underline{\tilde{p}}\vec{v}) + bb\tilde{u} = 0 \text{ in } \Omega$$

$$\tilde{p} = 0 \text{ on } \Gamma$$

incr. forw

$$\delta_p \mathcal{L}^H = \int_\Omega \left[\tilde{m}\nabla u\cdot\nabla\hat{p} + m\nabla\tilde{u}\cdot\nabla\hat{p} + \hat{p}\, v\cdot\nabla\tilde{u}\right] dx = 0 \text{ in } \Omega \quad \forall\, \hat{p}\in H_0^1$$

$$\Rightarrow -\nabla\cdot(m\nabla\tilde{u}) + v\cdot\nabla\tilde{u} = \nabla\cdot(\tilde{m}\nabla u) \text{ in } \Omega$$

$$\tilde{u} = 0 \quad \text{on } \Gamma$$

Hessian action: in dir of $\tilde{m}$

$$\delta_m \mathcal{L}^H = \int_\Omega \left[\alpha\nabla\hat{m}\cdot\nabla\tilde{m} + \hat{m}\nabla u\cdot\nabla\tilde{p} + \hat{m}\nabla\tilde{u}\cdot\nabla p\right] dx \equiv \mathcal{H}\tilde{m} \quad \forall\, \hat{m}\in H^1$$

$$\Rightarrow \left.\begin{array}{l} -\nabla\cdot(\alpha\nabla\tilde{m}) + \nabla u\cdot\nabla\tilde{p} + \nabla\tilde{u}\cdot\nabla p \quad \text{in } \Omega \\ \partial\tilde{m}/\partial n = 0 \quad \text{on } \Gamma \end{array}\right\} \mathcal{H}\tilde{m}$$