Parameter Estimation

- Now want to also estimate "parameters" of PDE i.e. coefficients appearing in PDE

$$\text{e.g. } \nabla \cdot \underline{\underline{\sigma}} \nabla u = f$$

- Use GLS but new complication, solution, u, is nonlinear in parameters

Now $Ku = b$, we have unknowns in K

i.e. $K(y)\, u = b$

$y \equiv$ parameters to be estimated along w/ b and u

- Want to minimize $\delta = d - Su$ in GLS sense but now w/ unknowns b, u, y

- Proceed as before

$$\min \Omega = b^T W_b b + \delta^T W_\delta \delta + y^T W_y y$$

subject to $K(y)\, u = b$, $\delta = d - Su$

where $W_b = [\mathrm{Cov}(b)]^{-1}$, $W_\delta = [\mathrm{Cov}(\delta)]^{-1}$, $W_y = [\mathrm{Cov}(y)]^{-1}$

Write augmented quadratic form

$$\Omega^{+} = \Omega + \lambda^T (Ku - b)$$

- Get first-order conditions for GLS Extremum

$$\frac{\partial \Lambda^{++}}{\partial b} = 2W_b b - \lambda = 0$$

$$\frac{\partial \Lambda^{++}}{\partial u} = K^T\lambda - 2S^T W_\delta \delta = 0$$

$$\frac{\partial \Lambda^{++}}{\partial \lambda} = Ku - b = 0$$

Same as before except $K(y)$

$$\frac{\partial \Lambda^{++}}{\partial y} = 2W_y y + \lambda^T \frac{\partial K}{\partial y} u = 0$$

new, if K linear in y (typical) $\frac{\partial K}{\partial y}$ is constant

- Gradient in parameter space

let $r \equiv Ku - b$

then $\Lambda^{++} = \Lambda + \lambda^T r = \Lambda + \sum_i \lambda_i r_i$

If each r_i is WR eqn for weighting function ϕ_i then we have

$$\frac{\partial}{\partial y_m}(\lambda^T r) = \sum_i \lambda_i \frac{\partial r_i}{\partial y_m}$$

e.g. assume PDE is $\nabla \cdot y \nabla u = f$

then $r_i = \langle y \nabla u \cdot \nabla \phi_i \rangle - \oint q\, \phi_i\, ds + \langle f \phi_i \rangle$

Type II data part of b

Expand $y = \sum_m y_m \phi_m$ nodal parameters then $\frac{\partial r_i}{\partial y_m} = \langle \phi_m \nabla u \cdot \nabla \phi_i \rangle$

and contribution to gradient becomes

$$\frac{\partial}{\partial y_m}(\lambda^T r) = \langle \phi_m \nabla u \cdot \nabla \lambda \rangle \quad \text{where } \nabla \lambda = \sum \lambda_i \nabla \phi_i$$
$$\text{i.e. } \lambda = \sum \lambda_i \phi_i$$

- Can evaluate this term using standard FE element assembly treating λ and u as data when known

- Since need to iterate (since problem is nonlinear) can re-assemble K, $\frac{\partial K}{\partial y_m}$ in each iteration as λ, u change

and $\frac{\partial \Omega^{++}}{\partial y} = 2 W_y y + \langle \phi_m \nabla u \cdot \nabla \lambda \rangle$

Note: If Type I BCs are involved at those nodes
$r_i \neq f(y)$ i.e. r_i does not depend on y
operationally we can set those $\lambda_i = 0$ in the construction of $\langle \phi_m \nabla u \cdot \nabla \lambda \rangle$ temporarily (i.e. during assembly of this term, <u>only</u>)

- Can prove Adjoint Method for sol'n

1. Prior estimates of b, y are needed (these are the unknowns)

2. Forward Model: Assemble $K, \frac{\partial K}{\partial y}$

   Solve $Ku = b$

3. Model-Data Misfit: Evaluate, $\delta = d - Su$, using new estimate of u

4. Adjoint Model for λ

   Solve $K^T \lambda = 2 S^T W_\delta \delta$

5. Gradient Descent: Evaluate gradients in b and y

$$\frac{\partial \Omega}{\partial b}^{++} = 2 W_b b - \lambda$$

$$\frac{\partial \Omega}{\partial y}^{++} = 2 W_y y + \lambda^T \frac{\partial K}{\partial y} u$$

   If zero, stop, otherwise adjust b, y and repeat steps 2-5

e.g. Gradient Descent

$$b_{k+1} = b_k + \alpha \partial b \qquad \partial b \rightarrow -\frac{\partial \Omega}{\partial b}^{++}$$

$$y_{k+1} = y_k + \alpha \partial y \qquad \partial y \rightarrow -\frac{\partial \Omega}{\partial y}^{++}$$

$$\alpha = -\frac{\partial b^T W_b b_k + \partial \delta^T W_\delta \delta_k + \partial y^T W_y y_k}{\partial b^T W_b \partial b + \partial \delta^T W_\delta \partial \delta + \partial y^T W_y \partial y}$$

} ignores nonlinearity in $\Delta u = K(y+\partial y)^{-1} \partial b$