Week 4

Tuesday, February 6, 2024 11:53 AM

Today: - Towards PDE constrained optimization ← hard

- Solving FEM
- Linear Algebra identities
- Karush-Kuhn Tucker conds
- Lagrange multipliers

PDE constrained optimization

to fit unknown model parameters to data, assume parameterized model form for solution

$$\mathcal{L}_\theta[u] = b$$

e.g.

$$\nabla \cdot F = b$$

$$F = -K_\theta(x)\nabla u$$

$$K_\theta(x) \in V_h \leftarrow \text{FEM space}$$

set up and solve

$$\operatorname{argmin} \|u - u_{target}\|^2 + \lambda \|\partial_x u - \partial_x u_{target}\|^2$$

$$\text{s.t. } \mathcal{L}_\theta[u] = b$$

Lecture slides with some examples of tomography, homogenization

---

Need some calculus / linear algebra identities (see "matrix cookbook")

① $\nabla_x \|Ax - b\|^2$

$$= \frac{\partial}{\partial x_\alpha} \sum_i \left[\left(\sum_j A_{ij} x_j\right) - b_i\right]^2$$

$$= \sum_i 2\left[\left(\sum_j A_{ij} x_j\right) - b_i\right] \frac{\partial}{\partial x_\alpha}\left(\sum_k A_{ik} x_k\right)$$

$$= \sum_i 2\left[\left(\sum_j A_{ij} x_j\right) - b_i\right] \sum_k A_{ik}\,\delta_{k\alpha}$$

$$= \sum_i 2\left[\left(\sum_j A_{ij} x_j\right) - b_i\right] A_{i\alpha}$$

$$= 2A^T(Ax - b)$$

Equality Constrained Quadratic Programs

§ 8.5 in Murphy for extra context

To solve:

$$\min_x F(x)$$

$$\text{s.t. } Mx = b$$

Introduce Lagrange Multiplier λ

① $\min_{x,\lambda} F(x) + \lambda^T(Ax - b) := \mathcal{L}(x, \lambda)$

② $Mx = g$

For special case where F is a

(2) $\min Mx = g$

For special case where F is a "quadratic form" $\leftarrow F(x) = \frac{1}{2}x^TAx + b^Tx + c$

Same as calc I: at a minimizer

(1a) $\frac{\partial}{\partial x}\mathcal{L}(x,\lambda) = 0$

(1b) $\frac{\partial}{\partial \lambda}\mathcal{L}(x,\lambda) = 0$

(1a) $\frac{\partial}{\partial x}\mathcal{L}(x,\lambda)$

$$= \frac{\partial}{\partial x}\left[\frac{1}{2}x^TAx + b^Tx + c + \lambda^T(Mx - g)\right]$$

$$= \boxed{Ax + b + M^T\lambda = 0}$$

(1b) $\frac{\partial}{\partial \lambda}\mathcal{L}(x,\lambda)$

$$= \frac{\partial}{\partial \lambda}\left[\frac{1}{2}x^TAx + b^Tx + c + \lambda^T(Mx - g)\right]$$

$$= \boxed{Mx - g = 0}$$

Leaving the system of equations

$$\begin{bmatrix} A & M^T \\ M & 0 \end{bmatrix}\begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -b \\ g \end{bmatrix}$$

Called the KKT equations (§8.5.2 Murphy)

In linear algebra a matrix of the form

$$\begin{bmatrix} A & B \\ B^T & 0 \end{bmatrix}$$

is called a saddle point problem

To solve, we'll do Gauss elimination on a block level (Schur complement)

$$\begin{bmatrix} A & B \\ B^T & 0 \end{bmatrix}\begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Mult. top row by M^TA^{-1} and subtract ~~from~~ bottom row

$$B^Tx - B^TA^{-1}(Ax + B\lambda) = b_2 - B^TA^{-1}b_1$$

$$B^Tx - B^TA^{-1}Ax + B^TA^{-1}B\lambda = \text{''}$$

$$S\lambda = B^TA^{-1}B\lambda = b_2 - B^TA^{-1}b_1$$

Schur [complement] which can be solved ~~for~~ λ

$S\lambda - BA^{-1}B\lambda \quad b_2 \ldots$

Schur complement

which can be solved ~~for~~ λ

then $Ax = b_1 - B\lambda$
can be solved for x

Finally, ready to apply to FEM inverse problem.

Need to turn the optimization problem

$$\min \int (u - u_{target})^2 + \varepsilon^2 \int |\nabla u - \nabla u_{target}|^2$$

$$\text{s.t.} \quad \mathcal{L}_\theta[u] = f$$

into a linear algebra problem

$\nabla \cdot F = f$
$F = -K_\theta(x)\nabla u$ $\Rightarrow$ Model to be fit

$$-\int K_\theta \nabla u \nabla v \, dx = \int v f \, dx$$

$$= \sum_q \partial_x \phi^a(x_q)\, \partial_x \phi^b(x_q)\, \hat{u}_b w_q\, K^\theta(x_q) = \sum_q \phi^a(x_q) f_q w_q$$

$$\boxed{A^\theta \hat{u} = b}$$ Physics Constraint

$$\int (u - u_{target})^2 = (\hat{u} - \hat{u}_{target})^T M (\hat{u} - \hat{u}_{target})$$

$$\int |\nabla u - \nabla u_{target}|^2 = (\hat{u} - \hat{u}_{target})^T S (\hat{u} - \hat{u}_{target})$$

So equiv to $\quad \min \ (\hat{u} - \hat{u}_{target})^T (M + \varepsilon^2 S)(\hat{u} - \hat{u}_{target})$

Combining every thing:

$$\begin{pmatrix} M + \varepsilon^2 S & A^{\theta T} \\ A^\theta & 0 \end{pmatrix} \begin{pmatrix} \hat{u} \\ \lambda \end{pmatrix} = \begin{pmatrix} (M + \varepsilon^2 S)\hat{u}_{target} \\ b \end{pmatrix}$$

After solving KKT for $\hat{u}, \lambda$

Solve a step of $\quad \min_\theta \ L(\theta \,|\, \hat{u}, \lambda)$

$$L = (\hat{u} - \hat{u}_t)^T (M + \varepsilon^2 S)(u - \hat{u}_t) + \lambda^T (A^\theta \hat{u} - b)$$

Thurs 2/8

Thurs 2/8

PINNs vs FEM
Cage Match

- HW questions?

Recall Task $\min_{\theta} \|u - u_t\|^2 + \varepsilon^2 \|u' - u_t'\|^2$
s.t $\mathcal{L}_\theta[u] = f$

We say that our solution data on u is indirectly supervising our model form θ

- Overview of PDE constrained FEM code
- Live coding for turning PINN code from forward solver into inverse solver

Final aspects of FEM theory

Given $u'' = f$

Turn into variational problem

$$a(u,v) = (u', v') = \int u' v' \, dx$$
$$(f, v) = \int f v \, dx$$

Solve, for any $v \in V_h$
$u \in V_h$ s.t $a(u,v) = (f,v)$ ☆

Thm For $f \in L^2$, ☆ is unique

Pf - If two soln exist
$$a(u_1, v) = (f, v)$$
$$a(u_2, v) = (f, v)$$

$\forall v$ $a(u_2 - u_1, v) = (0, v)$

Since $u_1, u_2 \in V_h$, $u_2 - u_1 \in V_h$
Pick $v = u_2 - u_1$

$$a(u_2 - u_1, u_2 - u_1) = 0$$
$$\int_0^1 (u_2 - u_1)'^2 \, dx = 0$$
$$\Rightarrow (u_2 - u_1)' = 0$$

$(u_2 - u_1) = C$

But $v \in V_h \Rightarrow v(0) = v(1) = 0$

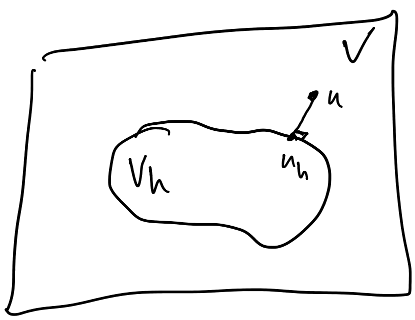
$(u_2 - u_1)(0) = 0 \Rightarrow C = 0$

$u_2 = u_1$

Galerkin Orthogonality

Consider true solution $u \in V$
and FEM solution $u_h \in V_h$

$$a(u - u_h, w) = 0 \quad \forall w \in V_h$$



<u>def</u> $\|v\|_E = \sqrt{a(v,v)}$ is <u>energy norm</u>

<u>thm</u> Cauchy-Schwartz inequality

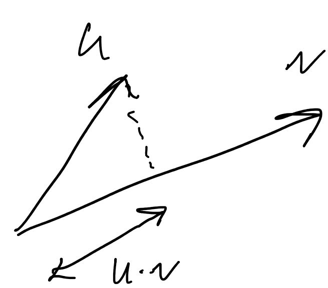
$$(u,v) \leq \|u\|_0 \, \|v\|$$

and consequently

$$|a(u,v)| \leq \|v\|_E \, \|u\|_E$$

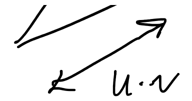
<u>Some intution</u>

Take 2 vector $u, v \in \mathbb{R}^N$



$$\|u+v\|^2 = (u+v)\cdot(u+v)$$
$$= u\cdot u + v \cdot v + 2u\cdot v$$

$$= u \cdot u + v \cdot v + 2 u \cdot v$$
$$= \|u\|^2 + \|v\|^2 + 2 u \cdot v$$



If C.S. $\leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\|$
$$\leq \|u\|^2 + \|v\|^2$$

<u>thm</u> $$\|u-u_h\|_E = \inf_{v \in V_h} \|u-v\|_E$$

<u>pf</u> $$\|u-u_h\|_E^2 = a(u-u_h, u-u_h) \quad \forall v \in V_h$$
$$= a(u-u_h, u-v+v-u_h)$$
$$= a(u-u_h, u-v) + \underbrace{a(u-u_h, v-u_h)}_{=0 \text{ by galerkin orthogonality}}$$
$$= a(u-u_h, u-v)$$

$$\|u-u_h\|_E^2 \leq \|u-u_h\|_E \|u-v\|_E$$

$$\|u-u_h\|_E \leq \min_v \|u-v\|_E$$

Take v to be the interpolant

$$\|u-u_h\|_E \leq \|u-u_I\|$$
$$\leq C h^2 \max_{x \in [0,1]} \|u''\|$$