PDE - constraint optimization

Jingyi Liu
University of Southern California

Kanso's Bio-inspired Lab

2024-06-24

Forward and Inverse Problem

Forward problem

$$\mathscr{F}(c;\alpha) = 0$$

$\mathscr{F}(c;\alpha)$ represents a PDE operator acting on c, with coefficient α

Inverse problem

$$\text{find } \alpha, \min_{\alpha} f(c(\alpha))$$
$$\text{s.t. } \mathscr{F}(c;\alpha) = 0$$

f : objective function ($ie.\ |c - c_o|^2$)

α : design variable

c : state variable

https://www.cs.cmu.edu/~ateh/presentations/slides-pde-constrained-optimization.html

https://kailaix.github.io/ADCME.jl/stable/tu_optimization/

Find adjoint variable

$$\text{find } \alpha, \min_{\alpha} f(c(\alpha))$$
$$\text{s.t. } \mathscr{F}(c;\alpha) = 0$$

f : objective function, $(ie. |c - c_o|^2)$
α : design variable
c : state variable

we want to use gradeint based optimization,

we need, $\nabla_\alpha f(c(\alpha))$

$$\nabla_\alpha f = \nabla_c f \, \nabla_\alpha c$$

Obtain by using PDE

from $\mathscr{F}(\mathrm{c}(\alpha);\alpha) = 0$ holds for any α,

we have $\mathrm{d}_\alpha \mathscr{F} = \partial_\mathrm{c} \mathscr{F} \, \partial_\alpha \mathrm{c} + \partial_\alpha \mathscr{F} = 0$

$$\rightarrow \nabla_\alpha c = -(\nabla_c \mathscr{F})^{-1} \nabla_\alpha \mathscr{F}$$

$$\nabla_\alpha f = \lambda^\mathrm{T} \nabla_c f, \quad \text{where } \lambda^T = -(\nabla_c \mathscr{F})^{-1} \nabla_\alpha \mathscr{F}$$

Link to Lagrange multiplier

find α, $\min_{\alpha} f(c(\alpha))$

$\text{s.t.}\ \mathscr{F}(c;\alpha) = 0$

f : objective function, $(ie.\ |c - c_o|^2)$

α : design variable

c : state variable

we want to use gradeint based optimization,

we need, $\partial_\alpha f(c(\alpha))$

$\nabla_\alpha f = \nabla_c f\ \nabla_\alpha c$

Obtain by using PDE

from $\mathscr{F}(\mathrm{x}(\alpha);\alpha) = 0$,

we have $\mathrm{d}_\alpha \mathscr{F} = \partial_\mathrm{c} \mathscr{F}\ \partial_\alpha \mathrm{c} + \partial_\alpha \mathscr{F} = 0$

$\rightarrow \nabla_\alpha c = -(\nabla_c \mathscr{F})^{-1}\ \nabla_\alpha \mathscr{F}$

$\nabla_\alpha f = \lambda^\mathrm{T}\ \nabla_c f$, where $\lambda^T = -(\nabla_c \mathscr{F})^{-1}\ \nabla_\alpha \mathscr{F}$, is adjoint variable

langrange function, $L(c,\alpha,\lambda) = f(c(\alpha)) + \lambda^T\ \mathscr{F}(c;\alpha)$

KKT conditions :

(1) $\partial_\lambda L(c,\alpha,\lambda) = \mathscr{F}(c;\alpha) = 0$

(2) $\partial_c L(c,\alpha,\lambda) = \nabla_c f(c) + \lambda^T\ \nabla_c \mathscr{F}(c;\alpha) = 0$

(3) $\partial_\alpha L(c,\alpha,\lambda) = \nabla_\alpha f(c) + \lambda^T\ \nabla_\alpha \mathscr{F}(c;\alpha) = 0$

for a given α, solve for c, λ, from(1), (2),

$\rightarrow\ \partial_\alpha L(c,\alpha,\lambda) = \nabla_\alpha f - \nabla_c f (\nabla_c \mathscr{F})^{-1}\ \nabla_\alpha \mathscr{F}$

When alpha is optimal, this expression is equal to zero, i.e., all the KKT conditions are satisfied.

https://en.wikipedia.org/wiki/Karush–Kuhn–Tucker_conditions#:~:text=In

Apply adjoint method

$$\text{find } \boldsymbol{\beta}, \max_{\boldsymbol{\beta}} \text{Sh}(c(\boldsymbol{\beta}))$$

$$\text{s.t. } \mathcal{F}(c;\boldsymbol{\beta}) = \text{Pe}\,\mathbf{u}(\boldsymbol{\beta}) \cdot \nabla c - \triangle c = 0,$$

$$||\boldsymbol{\beta}^2|| = 1$$

where $Sh = \int_S \nabla c \cdot \hat{\mathbf{n}} dS$, represents flux,

$c(r, \theta, \phi)$ represents concentrationfield in spherical coordinates

$\mathcal{F}$ represents advection – diffusion equation

$$\mathscr{L} = \text{Sh}(c(\boldsymbol{\beta})) + \int g\big(\text{Pe}\,\mathbf{u}(\boldsymbol{\beta}) \cdot \nabla c - \triangle c\big) dV = 0,$$

See details in: https://arxiv.org/abs/2404.13467

Automatic differentiation in computation graph

$$\text{find } \alpha, \min_{\alpha,\mathbf{c}_1} J = f(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4)$$

$$\text{s.t. } \mathbf{c}_2 = \mathscr{F}_1(\mathbf{c}_1; \alpha)$$

$$\mathbf{c}_3 = \mathscr{F}_2(\mathbf{c}_2; \alpha)$$

$$\mathbf{c}_4 = \mathscr{F}_3(\mathbf{c}_3; \alpha)$$

f is loss function

adjoint variables : $\lambda_4^T = \dfrac{\partial f}{\partial \mathbf{c}_4}$

$$\lambda_3^T = \frac{\partial f}{\partial \mathbf{c}_3} + \lambda_4^T \frac{\partial \mathscr{F}_3}{\partial \mathbf{c}_3}$$

$$\lambda_2^T = \frac{\partial f}{\partial \mathbf{c}_2} + \lambda_3^T \frac{\partial \mathscr{F}_2}{\partial \mathbf{c}_2}$$

langrange function, $L = f + \lambda_2^T (\mathscr{F}_1(\mathbf{c}_1; \alpha) - \mathbf{c}_2) + \lambda_3^T (\mathscr{F}_2(\mathbf{c}_2; \alpha) - \mathbf{c}_3) + \lambda_4^T (\mathscr{F}_3(\mathbf{c}_3; \alpha) - \mathbf{c}_4)$

gradient of objective function, $\dfrac{\partial L}{\partial \alpha} = \lambda_2^T \dfrac{\partial \mathscr{F}_1}{\partial \alpha} + \lambda_3^T \dfrac{\partial \mathscr{F}_2}{\partial \alpha} + \lambda_4^T \dfrac{\partial \mathscr{F}_3}{\partial \alpha}$

---

relations : $\dfrac{\partial f}{\partial \mathbf{c}_2} = \dfrac{\partial J_{\text{total}}}{\partial \mathbf{c}_3} \dfrac{\partial \mathscr{F}_2}{\partial \mathbf{c}_2}, \qquad \dfrac{\partial f}{\partial \alpha} = \dfrac{\partial J_{\text{total}}}{\partial \mathbf{c}_3} \dfrac{\partial \mathscr{F}_2}{\partial \alpha}$

total gradient of $\mathbf{c}_2$: $\dfrac{\partial J_{\text{total}}}{\partial \mathbf{c}_2} = \dfrac{\partial f}{\partial \mathbf{c}_2} + \dfrac{\partial J_{\text{total}}}{\partial \mathbf{c}_3} \dfrac{\partial \mathscr{F}_2}{\partial \mathbf{c}_2}$ $\quad \rightarrow \lambda_i^T = \dfrac{\partial J_{\text{total}}}{\partial \mathbf{c}_i}$

Thank you!

https://kailaix.github.io/ADCME.jl/stable/tu_optimization/

https://www.cs.cmu.edu/~ateh/presentations/slides-pde-constrained-optimization.html#/4