"Beyond Forward Simulation": intro to Inverse Problems

Monday, October 21, 2024 9:05 AM

Entire field of research, so hard to give just a few references. A lot of today's material is from the Zulkeefal Dar et al. chapter "Reduced Order Modeling" (ch. 8 in "Machine Learning in Modeling and Simulation", Rabczuk and Bathe, eds, Springer 2023)

Classic works by many authors. A small sampling:

- UQ (Ralph Smith or T. Sullivan's books),
- approximation theory in context of parametric PDEs (Ron DeVore, Albert Cohen, etc)
- Surrogate modeling/Model reduction (Willcox, Ghatas, Owhadi, etc.)

mathematical "tasks":
- How to sample?
- How to reconstruct?
- Existence, uniqueness
- For Bayesian: how to efficiently sample?

(Applied)
Tasks / Goals of Inverse Problems

① System identification

ex: $u(t)$ = % population infected w/ COVID,
$u' = \theta \cdot u(1-u)$ ODE, we have (noisy) observations.
Q: what is θ?

ex: biology or chemical reactions,
Q: what is the ODE or PDE?

ex: physics: we know full PDE but too slow to solve
Q: what's an approximate PDE (that's cheap to solve)?

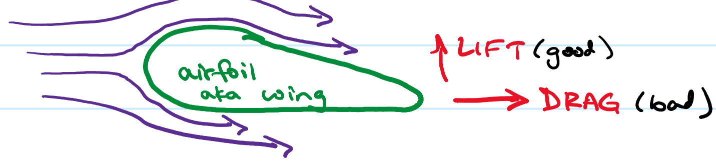
② "plain" inverse problem (kinda the same as #1)

ex: Full Waveform Inversion (FWI) in seismics
(see 1st lecture of class)
Q: what's the speed of sound? (as a function of space)
...given measurements...

③ Optimization and control
feedback into system

ex:

Solve Navier-Stokes



optimize shape of
airfoil to minimize DRAG $- \lambda \cdot$ LIFT

Inverse Problems (p. 2)

Monday, October 21, 2024 9:27 AM

④ Uncertainty Quantification

"uncertainty" can be aleatoric or epistemic

A key (useful) idea in science, and UQ in particular, is to treat epistemic uncertainty as stochastic

i.e. "what's the chance they serve hot dogs at the cafeteria today?" We may say "70%" but it's on a fixed lunch schedule and is deterministic: 0% or 100%

ex what's expected energy output from a wind farm next year?

· we have unknown (or uncertain) inputs, like wind speed.

Commonalities:

We need to solve the forward problem many times, each for a different parameter θ

Ex: θ is growth rate for COVID, $\theta \in \mathbb{R}^1$

Ex: θ is speed of sound, θ a function (or discretized, so high-dimensional)

Ex: θ is the geometry of a boundary, θ a curve (inside $\mathbb{R}^2$) or a surface (inside $\mathbb{R}^3$) or discretized

Ex: θ is wind speed, a function of time (or discretized)

Broad categories of solution methods:

ⓐ Forward-solve based

For given θ, solve forward problem, compare to true data

Compute loss, then take gradient $\frac{\partial}{\partial \theta}$ (eg. adjoint-state)

ⓑ Purely data-driven

eg. SINDy and co. for system id., model discovery ...

eg. neural net based ...

Inverse Problems (p. 3)

Monday, October 21, 2024 9:38 AM

Surrogate Models

If solving the forward problem for each value of θ is expensive, a natural idea is to make a cheap approximation.

Obvious idea, so a huge literature, and inconsistent naming.

R. Smith's organization:

"Surrogate model" is the general idea, with specific subcategory:

a) "Regression"/"Interpolation" models
- "Data-fit"
- "Response Surface"
- "Emulators"

Ex: polynomial chaos expansion PCE

Gaussian Process Regression

Karhunen-Loeve decomposition

- often model only the "Quantity of Interest" QOI

ex: airfoil, classically we solve Navier-Stokes to get $\vec{u}(\vec{x},t)$ (3D/2D velocity), then a simple post-processing gives $\underbrace{\text{LIFT and DRAG}}_{\text{QOI}}$

$\theta \nearrow \hat{u} \longrightarrow \widehat{QOI}$

high-dim.

vs.

$\theta \longrightarrow \widehat{QOI}$

Rather than $u_{approx} \to LIFT_{approx.}, DRAG_{approx}$

we skip ↗ and directly model ↖ ↗

since much lower dimensional

- often allow stochastic models, good for UQ, Bayesian Inverse Problems
- "non-intrusive"

b) "Reduced Order Models" (ROM)

usually meaning projection-based ROM

- model u directly ... but in lower-dimensional space

- "offline" step to find basis, via training data
- "online" step to deploy (for new θ)

intrusive (physics-based, access to equations, source code)

Ex: Proper Orthogonal Decomposition (POD) based Galerkin projection

POD-G

Reduced Basis (a greedy method)

non-intrusive (data-driven)

Ex: Dynamic Mode Decomposition (DMD)

Inverse Problems (p. 4)

Monday, October 21, 2024 9:56 AM

c) "hierarchical models", sometimes "low-fidelity"

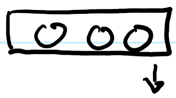
Ex · solve but to a looser tolerance

· coarser discretization
· simplified physics

$\Big\}$ "obvious" ideas

· model u_{approx} like type (6)

e.g. LES for Navier-Stokes, continuum models for particles



beam via Euler-Bernoulli beam theory (2D, uniform)

· analysis is often more application dependent

and as of ~2019, we'll add...

d) Neural-net based operator learning, hybrid of above approaches

Ex Neural Operators (several groups)

Physics-Informed NO (PINO)
Derivative-Informed NO (DINO)
Random Features
Fourier NO (FNO)
DeepONets

Misc. Math Background

θ often a function, so we'll talk about "function spaces"

type of vector space but ∞-dimensional

"∞-dimensional" vector space

· $\mathbb{R}^n$, $x = (x_1, x_2, \dots, x_n)$

· "$\mathbb{R}^\infty$", $x = (x_1, x_2, \dots)$ ← by itself, doesn't have nice structure

"ℓ^2" = "vectors" (or "sequences") such that if $x = (x_1, x_2, \dots)$ then $\sum_{i=1}^{\infty} |x_i|^2 < \infty$

Inverse Problems (p. 5)

Monday, October 21, 2024 10:07 AM

a **Banach** space is a vector space that also has a norm
(to measure "size", $\|x\|$, and distance $\|x-y\|$ ... so also a metric space)
(... and is "complete", i.e. no holes like $\mathbb{Q}$)

a **Hilbert** space is a vector space with an inner product (... and complete...)
$\langle x, y \rangle$ Also is Banach

Ex: $x, y \in \mathcal{H} = \mathbb{R}^n$, $\langle x,y \rangle = x^T y = \sum_{i=1}^{n} x_i y_i$

Ex: $f, g \in$ function space, $\langle f,g \rangle = \int_\Omega f(s) g(s)\, ds$

A **compact set** means closed and totally bounded in our context
↑ valley-girl speak

⚠ Analysis 101 pop-quiz: what does "closed" mean ?
If a set isn't open, does that mean it's closed ?

Function space examples

$f: \Omega \to \mathbb{R}$

$C(\Omega)$ = continuous functions,
a Banach space w/ norm $\|f\|_\infty := \sup_{x \in \Omega} |f(x)|$

$C^1(\Omega)$ = differentiable functions,
derivative is continuous

$H^1(\Omega)$ = 1st Sobolev Space, a fancy version of C^1
that relaxes "derivative" to "weak derivative" Hilbert Space

$L^2(\Omega)$ = f s.t. $\|f\|_{L^2} \left(:= \sqrt{\langle f,f \rangle} \right)$ is finite
$\langle f,g \rangle = \int f(s) g(s)\, ds$ like "RMS"

Inverse Problems (p. 6)

Monday, October 21, 2024 10:19 AM

Weak derivative (super brief)

A vector $x \in \mathbb{R}^n$ is zero iff $\forall y \in \mathbb{R}^n$, $y^T x = 0$.

So, $x = \tilde{x}$ iff $\forall y \in \mathbb{R}^n$, $\langle y, x \rangle = \langle y, \tilde{x} \rangle$

Take an ODE like $\underbrace{u' = 3u}$ on $t \in [0, 1]$

We say u is a weak solution of ↑ if

$\forall \varphi \in$ "test function space" ← depends on context, usually something dense, smooth w/ appropriate I.C./B.C.

$$\langle \varphi, u' \rangle = \langle \varphi, 3u \rangle$$

magic: $\int u\,dv = uv| - \int v\,du$

$$\langle \varphi, u' \rangle = \int \varphi(s) u'(s)\, ds$$

$$= \underbrace{\varphi(s) u(s) \Big|_{s=0}^{1} - \int \varphi'(s) u(s)\, ds}_{\text{no mention of } u' \text{ here!}}$$

punchline...

u can satisfy $u' = 3u$

even if u' doesn't exist (in the classical sense)?

It's not a free-for-all... not all functions have weak derivatives.

Intro to P.O.D

Basic idea to represent $u(t,x) \cong \sum_{k=1}^{\text{rank}} \overbrace{\Psi_k(x) \cdot \underbrace{u_k(t)}_{\text{coefficients}}}^{\text{"separable"}}$

learn $\{\Psi_k\}$ from data snapshots at different times $\{u_1, \dots, u_T\}$

and do SVD/PCA

For a new time point, simulate using $\{\Psi\}$ basis

If rank small, this is cheaper

In practice, our equations (from the weak form) are now overdetermined

so reduce them (Galerkin or Petrov-Galerkin style)

Inverse Problems (p. 7)

Monday, October 21, 2024 10:37 AM

Intro to DMD (non-intrusive) Schmid '10, relatively recent

Snapshot matrix $S = \underbrace{\left[u_1, \ldots, u_{T-1} \right]}_{\text{time}} \Big\}$ span

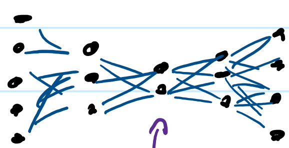
$S_+ = \left[u_2, \ldots, u_T \right]$

and look for operator T (linear, low-rank)

such that $S_+ = T \cdot S$ Find T via SVD

ROM in ML ... other than neural operators.

① Finding reduced-bases usually via Autoencoders



"bottleneck" layer / latent space

② Closure Modeling

to stabilize existing ROM methods