Adjoint State Method

Tuesday, October 15, 2024 5:01 PM

we're going to solve inverse problems, just one example of where the adjoint state (or adjoint sensitivity method) is useful

Dates back to at least Pontryagin et al. '62

Widely known in some communities... and always confusing!

We'll follow "Neural Ordinary Diff. Eq" (Chen, Rubanova, Bettencourt, Duvenaud) NeurIPS '18 (Toronto)

Motivation: θ is speed of sound, amount of diffusion or viscosity, initial condition, boundary condition (values, or geometry of boundary)

Warning: derivations are heuristic and sketchy!

Why adjoint state and not $\underbrace{\text{autodiff}}$?

high memory cost when used for ODE/PDE solvers

optimize-then-discretize

discretize-then-optimize

Adjoint State Method (p. 2)

Tuesday, October 15, 2024 5:20 PM

Adjoint State Method

more generally, $\sum \ell(z(t_i), y_i)$ or $\int \ell(z(t), y(t))\,dt$

Suppose we have a loss function $\mathcal{L}(z(T))$,

z solves ODE $z(0) = z_0$, $z'(t) = f(z, t; \underset{\text{weights}}{\theta})$

Abstractly, $z(T) = z_0 + \int_0^T f(z(t), t; \theta)\,dt$

What we want is

$z = z(\theta)$ via implicit fcn thm.

$$\frac{d\mathcal{L}}{d\theta} = \frac{d\mathcal{L}}{dz}\,\frac{\partial z}{\partial \theta}$$ (or a little more complicated...)

(sketchy! at which t ??)

how do we get this? i.e. $\frac{\partial z(t)}{\partial \theta}$ $\forall t$?

Running Example

want $\mathcal{L}(z(T))$

$\mathcal{L}(x) = x^2$, $z' = \theta \cdot z$, $z(0) = 3$ so $z(t) = 3 \cdot e^{\theta t}$

This is simple enough that we can check our answer

Compute $\frac{d}{d\theta}\mathcal{L}(z(T)) = \frac{d}{d\theta}(3e^{\theta T})^2 = 2 \cdot 3e^{\theta T} \cdot \frac{d}{d\theta}(3e^{\theta T})$

i.e. $g(\theta) = \mathcal{L}(z(T))$ Find $\frac{dg}{d\theta}$ $= 18 \cdot T \cdot e^{2\theta T}$

Derivation: first goal, get $\frac{d\mathcal{L}}{dz(t)} =: a(t)$ ← the "adjoint state"

Thm $a(t)$ satisfies its own ODE, $a' = -a \cdot \frac{\partial f(z, t; \theta)}{\partial z}$

We start at $t = T$ with $a(T) = \frac{\partial \mathcal{L}}{\partial z(T)}$ ("initial" condition)

proof

If ODE is "nice" (i.e. f is smooth), which is common, then

$$z(t+h) = z(t) + h\,\underbrace{z'(t)}_{= f(z, t; \theta)} + O(h^2)$$ Taylor Series

$$a(t) := \frac{d\mathcal{L}}{dz(t)} = \frac{d\mathcal{L}}{dz(t+h)} \cdot \frac{dz(t+h)}{dz(t)}$$ Chain Rule

$$= a(t+h) \cdot \frac{d}{dz(t)}\Big(z(t) + h \cdot f(z(t), t; \theta) + O(h^2) \Big)$$

$$= a(t+h) \cdot \Big(I^* + h \cdot \underbrace{\frac{\partial f}{\partial z}}_{\text{computable}} + O(h^2) \Big)$$

* we've been a bit loose but z could be a vector, so this is I not 1.

Think of as row vectors for now. (since Jacobians ($= \nabla^T$) have nice direction of chain rule)

Adjoint State Method (p. 3)

Wednesday, October 16, 2024 9:46 AM

... recall...

$$a(t) = a(t+h)\cdot\left(I + h\cdot\frac{\partial f}{\partial z} + o(h^2)\right)$$

Then

$$a' := \frac{da(t)}{dt} := \lim_{h\to 0}\frac{a(t+h) - a(t)}{h}$$

$$= \lim_{h\to 0}\frac{a(t+h)\cdot\left(I - \left(I + h\frac{\partial f}{\partial z} + o(h^2)\right)\right)}{h}$$

$$= \lim_{h\to 0} -a(t+h)\cdot\frac{\partial f}{\partial z} + o(h)$$ Assuming $a(t)$ is continuous...

$$= -a(t)\cdot\frac{\partial f}{\partial z}(z(t), t; \theta) \qquad \square$$

Running Example

$\mathcal{L}(x) = x^2$, $z' = \underbrace{\theta\cdot z}_{=f(z,t,\theta)}$, $z(0) = 3$ so $z(t) = 3\cdot e^{\theta t}$

Then $f(z,t,\theta)$ so $\frac{\partial f}{\partial z} = \theta$

$a(t) := \frac{d\mathcal{L}}{dz(t)}$

satisfies $a' = -a\cdot\frac{\partial f}{\partial z}$ i.e. $a' = -\theta\cdot a$

Full loss is $\mathcal{L}(z(T))$, $\frac{d\mathcal{L}}{dz(T)} = \underbrace{2\cdot z(T)}_{=6\cdot e^{\theta T}}$ (explicit) since $\mathcal{L}'(x) = 2x$

(we know since we did forward solve)

So $a(T) = \; = 6\cdot e^{\theta T}$

our initial/final condition.

Hence solve $a' = -\theta\cdot a$ on $t\in[0,T]$ and "I.C." $a(T) = 6e^{\theta T}$

... i.e., solve backwards! (if your ODE solver is unhappy, do a change of variables $t \mapsto \tilde{t} = T - t$ and adjust...)

So... in general, solve $a' = -a\cdot\frac{df}{dz}$

on $[0,T]$ "starting" at $a(T) = \frac{d\mathcal{L}}{dz(T)}$ (known)

But... we want $\frac{d\mathcal{L}}{d\theta}$. What to do?

Adjoint State Method (p. 4)

Wednesday, October 16, 2024 10:04 AM

⚠ SUBTLE

So, we defined $\frac{d\mathcal{L}}{dz(t)} = a(t)$, showed $\boxed{a' = -a \cdot \frac{\partial f}{\partial z} \quad a(T) = \frac{\partial \mathcal{L}}{\partial z(T)}}$

i.e. $\frac{d\mathcal{L}}{dz} := a_z$, $a_z' = -a_z \frac{\partial f}{\partial z}$

where $z' = f(z, t, \theta)$

Let's augment: $\frac{d}{dt}\begin{bmatrix} z \\ \theta \end{bmatrix} = \underbrace{\begin{bmatrix} f(z,t,\theta) \\ 0 \end{bmatrix}}_{F}$ } temporarily in column vectors

(i.e. $\theta' = 0$ since it doesn't depend on time)

$\frac{d\mathcal{L}}{d[z,\theta]} := A(t) = [a_z, a_\theta]$

so... $\boxed{A' = -A \cdot \frac{\partial F}{\partial [z,\theta]}}$ i.e.

$$\left[\frac{d}{dt} a_z, \frac{d}{dt} a_\theta\right] = -[a_z, a_\theta] \cdot \begin{bmatrix} \frac{\partial f}{\partial z} & \frac{\partial f}{\partial \theta} \\ \cancel{\frac{\partial 0}{\partial z}} & \cancel{\frac{\partial 0}{\partial \theta}} \end{bmatrix}$$

So solve augmented system in your ODE solver

$$= \Big[\underbrace{-a_z \cdot \frac{\partial f}{\partial z}}_{\text{original adjoint ODE}}, -a_z \frac{\partial f}{\partial \theta}\Big]$$

these might depend on state $z(t)$, so in practice, augment further and solve for z (backwards) So you have it. $z(T)$ known from forward pass

i.e. $\frac{d}{dt} a_\theta = \underbrace{-a_z \frac{\partial f}{\partial \theta}}$ doesn't depend on a_θ so can integrate

$$a_\theta(0) = \underbrace{a_\theta(T)} - \int_0^T a_\theta' \, dt$$

set to 0 as "I.C." since $\mathcal{L}$ doesn't depend explicitly on θ } sketchy!

$$= \int_T^0 a_\theta' \, dt = \int_T^0 -a_z \cdot \frac{\partial f}{\partial \theta} \, dt$$

calculate in adjoint ODE ("backward pass")

and $a_\theta(0) = \frac{d\mathcal{L}}{d\theta(0)}$

$= \frac{d\mathcal{L}}{d\theta}$ since θ is constant

which is what we wanted!

Adjoint State Method (p. 5)

Wednesday, October 16, 2024 10:29 AM

Running Example

$\mathcal{L}(x) = x^2$, $z' = \underbrace{\theta \cdot z}_{= f(z,t,\theta)}$, $z(0) = 3$ so $z(t) = 3 \cdot e^{\theta t}$

Then so $\frac{\partial f}{\partial z} = \theta$

$a(t) := \frac{d\mathcal{L}}{dz(t)}$

satisfies $a' = -a \cdot \frac{\partial f}{\partial z}$ i.e. $a' = -\theta \cdot a$

Full loss is $\mathcal{L}(z(T))$, $\frac{d\mathcal{L}}{dz(T)} = \underbrace{2 \cdot z(T)}_{= 6 \cdot e^{\theta T}}$ (explicit) from forward numerical solve

(we know since we did forward solve)

so $a(T) = 6 \cdot e^{\theta T}$

Hence solve $a' = -\theta \cdot a$ on $t \in [0, T]$ and "I.C." $a(T) = 6e^{\theta T}$

... i.e., solve backwards!

Solve explicitly

So...

$a = c\,e^{-\theta t}$

$a(T) = c \cdot e^{-\theta \cdot T}$

$= 6e^{\theta T}$

so $c = 6 \cdot e^{2\theta T}$

so $a(t) = 6\,e^{2\theta T - \theta t}$

Find $\frac{\partial f}{\partial \theta} = z = 3e^{\theta t}$

(for init/final condition)

Add in augmented dynamics...

$$a_\theta(0) = -\int_T^0 a \cdot \frac{\partial f}{\partial \theta}\, dt$$

$$= \int_0^T \underbrace{6e^{2\theta T - \theta t}}_{a} \cdot \underbrace{3e^{\theta t}}_{\frac{\partial f}{\partial \theta} (= z)}\, dt$$

$$= 18 \cdot e^{2\theta T} \cdot \int_0^T dt = 18 \cdot T \cdot e^{2\theta T}$$

matches!

Adjoint State Method (p. 6)

Wednesday, October 16, 2024 10:42 AM

In practice, 1) $z(t)$ often a vector

2) loss might not just be $\mathcal{L} = (z(T) - y_T)^2$

it might be $\mathcal{L} = \sum_{i=1}^{n} (z(t_i) - y_i)^2$

then solve (backward) adjoint equation

from t_n to t_{n-1} using (explicit) $\frac{\partial \mathcal{L}}{\partial z(t_n)}$

as "initial condition",

then solve from t_{n-1} to t_n using $\frac{\partial \mathcal{L}}{\partial z(t_{n-1})}$

as "initial condition",

etc.

See Dolfin-Adjoint

part of FEniCS, article at tinyurl.com/m9yfp7w

or SUNDIALS/CVODES from LLNL

typically adjoint eq'n is no harder than original ODE/PDE

doesn't need PDE to be linear

doesn't always "work", e.g. tough w/ unstructured grids

also, need not be consistent w/ discretization

(optimize-then-discretize

$\neq$ discretize-then-optimize)

Adjoint State Method (p. 7)

Friday, April 25, 2025 10:33 AM

His p is my θ

Ex. from Steven Johnson's notes for 18.335 ('06, '12, '21) tinyurl.com/mkv43kb

$G = g(u(\theta), \theta)$

$$\min_{\theta \in \mathbb{R}^n} g(u,\theta) \quad \text{s.t.} \quad \underbrace{f(u,\theta) = 0}_{\text{PDE}, \ldots}$$

$u \in \mathbb{R}^m$ state variable — or in a function space, but we'll do $m < \infty$ in this ex.

$\theta \in \mathbb{R}^n$ control

Find $\nabla_\theta G$, aka $G\nabla_\theta^T = \dfrac{dG}{d\theta} = G_\theta$ Jacobian

our example: $f(u,\theta) = 0$ is $A \cdot u = b$, A is a function of θ, e.g. $A = V \text{diag}(\theta) V^T$

maybe $b = b(\theta)$ too

$$\frac{dG}{d\theta} = \overset{1\times n}{g_\theta} + \overset{1\times m}{g_u} \cdot \overset{m\times n}{u_\theta}$$

total deriv. / chain rule

not obvious

explicit (easy)

u_θ is Jacobian matrix, $m \times n$

$u = A^{-1}b$ but A, b depend on θ, do implicit diff. of $Au = b$

$$\underbrace{\frac{d}{d\theta_i}(Au)}_{A_{\theta_i} \cdot u + A \cdot u_{\theta_i}} = \frac{d}{d\theta_i}(b) = \text{"} b_{\theta_i} \text{"}$$

$A_{\theta_i} \cdot u + A \cdot u_{\theta_i}$ product rule so... $u_{\theta_i} = A^{-1}(b_{\theta_i} - A_{\theta_i} \cdot u)$

hence

$$G_\theta = \underset{1\times n}{g_\theta} + \underset{1\times m}{g_u} \cdot \underbrace{\Big(\underset{m\times m}{A^{-1}} \big(\underbrace{b_\theta - A_\theta \cdot u}_{m\times n} \big) \Big)}$$

in this order, we solve n equations

Expensive if A isn't explicit (ex: ODE/PDE) since you can't cache inverse

better:

$$= g_\theta + \underbrace{\Big(\overset{1\times m}{g_u} \overset{m\times m}{A^{-1}} \Big)}_{\lambda^T} \cdot \overset{m\times n}{(b_\theta - A_\theta u)}$$

λ^T adjoint state, i.e. $A^{-T} g_u = \lambda$

transpose or... adjoint

aka solve $A^T \lambda = g_u$

A single ODE/PDE solve

Same trick as reverse-mode AD: cleverly group parentheses

Adjoint State Method (p. 8)

Friday, April 25, 2025 10:47 AM

What's the "adjoint" of a differential operator?

Ex From ch 8 of Gil Strang's "Comp. Sci. + Engineering" '07

$A(u) = u' - \varepsilon \cdot u''$ advection-diffusion eq'n in 1D

$u(0) = u(1) = 0$ homogeneous (if not, "subtract them off")

$u = u(x)$, solve on $x \in [0,1]$

The adjoint depends not just on the operator, but on the initial/boundary conditions (and domain)

cf. §10.2 "Applied Analysis", Hunter + Nachtergaele

$A: D \to \mathcal{H}$, $D \subsetneq \mathcal{H}$

$A^*: D^* \to \mathcal{H}$ st. $\langle Ax, y\rangle = \langle x, A^* y\rangle$ $\forall x \in D, \forall y \in D^*$

(A^*, D^*) pair is adjoint

For a matrix, A, adjoint A^* is st. $\forall\, v, w$ $\langle Av, w\rangle = \langle v, A^* w\rangle$

$\langle f, g\rangle = \int_0^1 f(x) g(x)\, dx$

$\int u\, dv = u \cdot v \big| - \int v\, du$

Same trick used to find "weak form"

$$\langle Au, \lambda\rangle = \int_0^1 \lambda \cdot (u' - \varepsilon u'')\, dx$$

$$= \int_0^1 \lambda \cdot u'(x)\, dx - \varepsilon \cdot \int_0^1 \lambda \cdot u''(x)\, dx$$

integrate both by parts

$$= \underbrace{\lambda \cdot u \big|_0^1}_{=0 \text{ B.C.}} - \int_0^1 \lambda' u - \varepsilon \Big[\lambda \cdot u' \big|_0^1 \underbrace{- \int_0^1 \lambda' u'}_{\text{IBP again}} \Big]$$

$$- \underbrace{\lambda' u \big|_0^1}_{=0 \text{ B.C}} + \int_0^1 \lambda'' u$$

$$= -\int_0^1 \lambda' u - \varepsilon \cdot \int_0^1 \lambda'' u - \varepsilon \cdot \lambda u' \big|_0^1$$

$$= \int_0^1 u \cdot \underbrace{(-\lambda' - \varepsilon \lambda'')}_{= A^*(\lambda)} - \varepsilon \cdot \underbrace{u' \lambda \big|_0^1}$$

impose $\lambda(0) = \lambda(1) = 0$ so that this vanishes

$$= \langle u, A^*(\lambda)\rangle$$

Ex $Au = 3u' + 4u$, $u(0) = 0$, on domain $[0, T]$

$$\langle Au, v\rangle = \int_0^T 3u'v\, dt + \int_0^T 4uv\, dt$$

$$= 3uv \big|_0^T - \int_0^T 3uv'\, dt + \int_0^T 4uv\, dt$$

so

$$A^*(v) = -3v' + 4v, \quad v(T) = 0, \quad t \in [0, T]$$

(and you can change $t \to T - t$ if you wanted...)