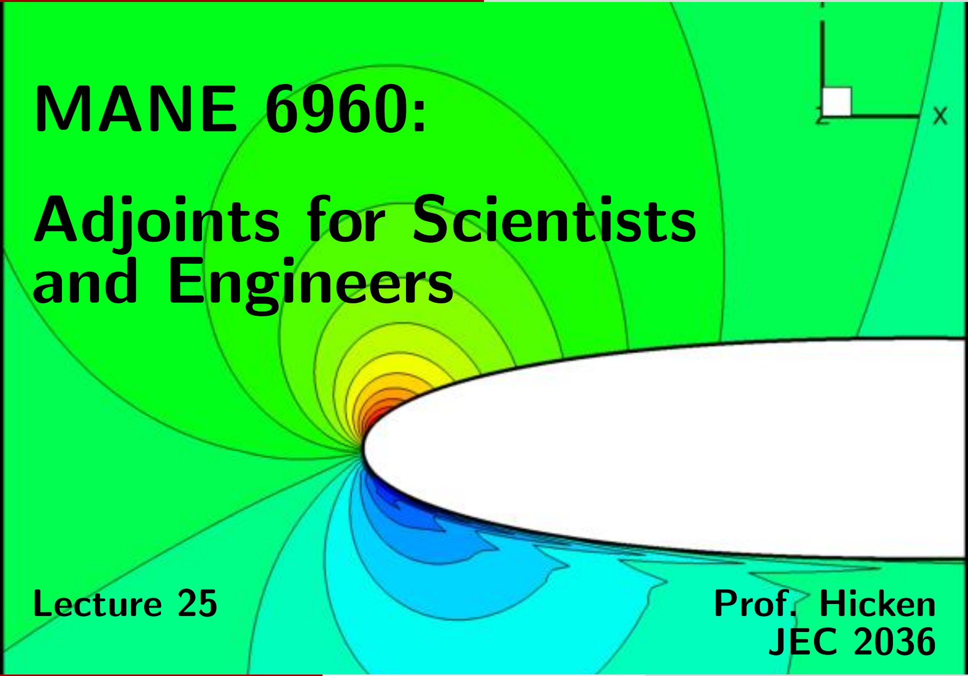


MANE 6960: Adjoints for Scientists and Engineers

Lecture 25

Prof. Hicken
JEC 2036

Trying to “fix” the adjoint

In today’s final lecture, I will highlight two methods that attempt to address the “failure” of the adjoint to provide useful sensitivity information when the dynamics are chaotic.

- The first is relatively inexpensive and inaccurate and the other is relatively expensive and accurate.
- Consequently, the best “fix” for chaotic adjoints remains an open problem.

Ensemble Adjoint

Average of Averages

The ensemble adjoint [LAH00] was the first idea proposed to deal with unstable adjoints that arise from chaotic dynamical systems. The idea is simple:

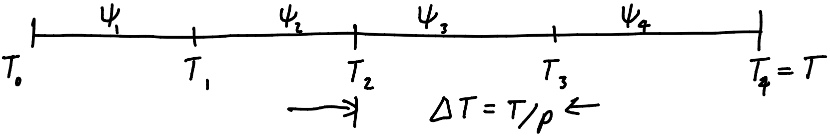
- adjoint is well-behaved over sufficiently short periods of time, even for chaotic problems.
- break the full simulation period into shorter periods, and compute the adjoint on each.
- use each adjoint to compute a sensitivity, and average this ensemble of sensitivities.

Average of Averages (cont.)

The ensemble adjoint avoids the long-time instability problem because each time period is short and treated separately.

- the adjoints are independent; no continuity is enforced between periods.

Example: break $[0, T]$ into P partitions of period $\Delta T = T/P$.



ψ_p is the solution of

$$\frac{\partial J}{\partial u}^T + \left(\frac{\partial R}{\partial u}\right)^T \psi_p = 0, \quad \forall t \, [T_{p-1}, T_p]$$

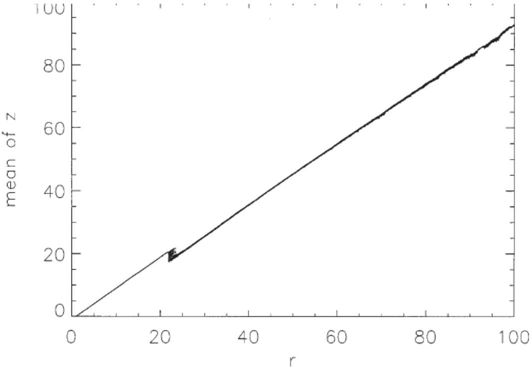
Average of Averages (cont.)

The final form of the total derivative with respect to α is

$$\frac{DJ}{D\alpha} = \sum_{p=1}^{P} \left(\frac{\partial J}{\partial \alpha} + \int_{T_{p-1}}^{T_p} (\psi_p, N'[\alpha])_\Omega \, dt \right)$$

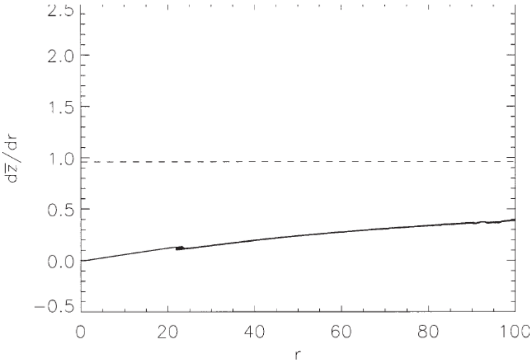
- In principle, the ensemble approach is no more expensive than the convectional adjoint; in fact, it could be implemented in parallel with respect to the time partition.
- Unfortunately, the resulting sensitivities converge very slowly: the convergence rate with the number of samples P is slower than $1/P$, i.e. worse than Monte-Carlo methods.

Results of the Ensemble Adjoint



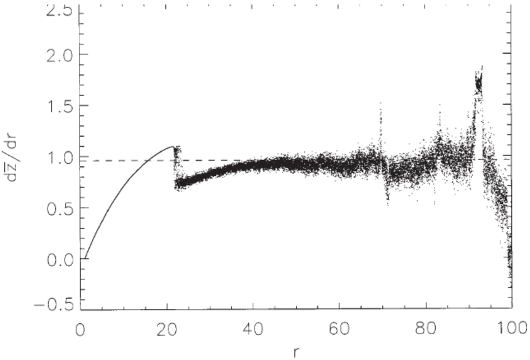
Behavior of $\bar{z} = 1/T \int_0^T z\,dt$, where z is from the Lorenz system, using $T = 131.36$ [LAH00].

Results of the Ensemble Adjoint (cont.)



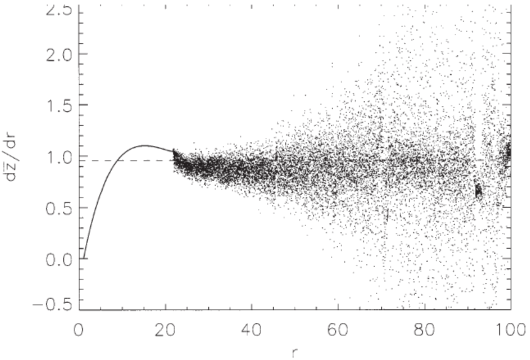
Ensemble-adjoint-based sensitivity $d\bar{z}/d\rho$ using $P = 1314$ and $\Delta T = 0.1$. [LAH00].

Results of the Ensemble Adjoint (cont.)



Ensemble-adjoint-based sensitivity $d\bar{z}/d\rho$ using $P = 299$ and $\Delta T = 0.44$. [LAH00].

Results of the Ensemble Adjoint (cont.)



Ensemble-adjoint-based sensitivity $d\bar{z}/d\rho$ using $P = 199$ and $\Delta T = 0.66$. [LAH00].

Least-Squares Shadowing

Shadowing Trajectory

The least-squares shadowing method for sensitivity analysis is quite different in its approach. The rough idea is as follows:

- Problem with chaotic system is that small perturbation of a parameter α leads to a large change in the state; recall

$$\|\Delta u(t)\| \approx e^{\lambda t} \|\Delta u_0\|$$

- Can we perturb the dynamics in such a way that the perturbed solution remains close to the reference solution?

This approach (perturbing the dynamics) is justified by the shadowing lemma:

Shadowing Trajectory (cont.)

Theorem: Shadowing Lemma

For any $\delta > 0$ and reference solution $u(t;\alpha)$, there exists an $\epsilon > 0$, a smooth time transformation $\tau(t)$, and a solution $\tilde{u}(\tau(t);\alpha+\epsilon)$ such that

$$\begin{aligned} \|\tilde{u}(\tau(t);\alpha+\epsilon) - u(t;\alpha)\| &< \delta, \qquad \forall t, \\ \left|1 - \frac{d\tau}{dt}\right| &< \delta, \qquad \forall t. \end{aligned}$$

- Requires uniform hyperbolicity and ergodicity.

Shadowing Trajectory (cont.)

The shadow trajectory $\tilde{u}$ can be approximated by solving the following least-squares optimization problem:

$$\min_{\tilde{u},\tau} J_\epsilon(\tilde{u},\tau) = \frac{1}{2}\int_0^T \|\tilde{u}(\tau(t)) - u(t)\|^2\, dt + \frac{\mu}{2}\int_0^T \left(1 - \frac{d\tau}{dt}\right)^2 dt,$$

$$\text{s.t.} \quad \frac{d\tilde{u}}{d\tau} + N(\tilde{u};\alpha+\epsilon) = 0.$$

- The first term in the objective forces the shadow trajectory to lie close to the reference
- The second term encourages the time transformation to be close to the identity; $\mu > 0$ is a parameter
- $\tilde{u}$ converges to a shadow trajectory as $T \to \infty$

Tangent form of the LSS

Assuming we can obtain a shadow trajectory based on the parameter $\alpha + \epsilon$, the next step is to consider taking the limit $\epsilon \to 0$ in order to define the sensitivity $Du/D\alpha$.

Recall that the conventional direct (a.k.a. tangent) sensitivity $Du/D\alpha \equiv v$ satisfies

$$\frac{D}{D\alpha}\underbrace{\left[\frac{du}{dt} + N(u,\alpha)\right]}_{0} = \frac{dv}{dt} + N'[u]v + N'[\alpha] = 0$$

Tangent form of the LSS (cont.)

Before we start taking limits, it will be helpful to express $d\tau/dt$ as a perturbation to the identity mapping:

time dilation

$$\frac{d\tau}{dt} = 1 + \epsilon\eta(t).$$

will look for η rather that τ

It will also be useful to keep in mind that

$$\lim_{\epsilon\to 0} \frac{\tilde{u} - u}{\epsilon} = \frac{Du}{D\alpha} \equiv v$$

here, this refers to the LSS tangent sens. (not conventional sens.)

Tangent form of the LSS (cont.)

First, we divide the objective J_ϵ by ϵ^2 and take the limit:

$$\lim_{\epsilon\to 0} \frac{J_\epsilon(\tilde{u},\tau)}{\epsilon^2} = \lim_{\epsilon\to 0} \frac{1}{\epsilon^2}\frac{1}{2}\int_0^T (\tilde{u}(\tau)-u,\ \tilde{u}(\tau)-u)\,dt$$

$$+ \lim_{\epsilon\to 0} \frac{1}{\epsilon^2}\frac{\mu}{2}\int_0^T \Big(1 - \underbrace{\frac{d\tau}{dt}}_{=1+\epsilon\eta(t)}\Big)^2 dt$$

$$= \lim_{\epsilon\to 0} \frac{1}{2}\int_0^T \left(\frac{\tilde{u}-u}{\epsilon}, \frac{\tilde{u}-u}{\epsilon}\right) dt$$

$$+ \lim_{\epsilon\to 0} \frac{\mu}{2}\int_0^T \frac{\cancel{\epsilon^2}\,\eta(t)^2}{\cancel{\epsilon^2}}\,dt$$

Tangent form of the LSS (cont.)

$$\lim_{\varepsilon \to 0} \frac{J_\varepsilon(\tilde{u}, \tau)}{\varepsilon} = \frac{1}{2}\int_0^T (v, v)\, dt + \frac{\mu}{2}\int_0^T \eta(t)^2\, dt$$

$$= \frac{1}{2}\int_0^T \|v\|^2 dt + \frac{\mu}{2}\int_0^T \eta(t)^2 dt$$

Tangent form of the LSS (cont.)

Similarly, we can determine the constraint on v by taking the difference between the dynamical systems, dividing by ϵ, and taking the limit:

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\overbrace{\frac{d\tilde{u}}{d\tau}+N(\tilde{u};\alpha+\epsilon)}^{0}\overbrace{-\frac{du}{dt}-N(u;\alpha)}^{0}\right] \doteq 0$$

First,
$$\frac{d\tilde{u}}{d\tau} = \frac{d\tilde{u}}{dt}\frac{dt}{d\tau} = \frac{1}{(1+\varepsilon\eta(t))}\frac{d\tilde{u}}{dt}$$

so,
$$\frac{1}{\varepsilon}\left(\frac{d\tilde{u}}{d\tau}-\frac{du}{dt}\right) = \frac{1}{\varepsilon(1+\varepsilon\eta)}\left[\frac{d\tilde{u}}{dt} - \frac{du}{dt} - \varepsilon\eta\frac{du}{dt}\right]$$

$$\lim_{\varepsilon\to 0}\frac{1}{(1+\varepsilon\eta)}\left[\frac{d}{dt}\left(\frac{\tilde{u}-u}{\varepsilon}\right) - \eta\frac{du}{dt}\right] = \frac{dv}{dt} - \eta\frac{du}{dt}$$

Tangent form of the LSS (cont.)

Furthermore, $\tilde{u} = u + \epsilon v + O(\epsilon^2)$

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon}\left[N(\tilde{u}; \alpha + \epsilon) - N(u;\alpha)\right]$$

$$= \lim_{\epsilon \to 0} \frac{1}{\cancel{\epsilon}}\left[\cancel{N(u,\alpha)} + N'[u]\cancel{\epsilon} v + N'[\alpha]\cancel{\epsilon} \underbrace{}_{O(\epsilon^2)} - \cancel{N(u,\alpha)}\right]$$

$$= N'[u]v + N'[\alpha]$$

Together we get

$$\underbrace{\frac{dv}{dt} + N'[u]v + N'[\alpha]}_{\text{usual direct/tangent equation}} \underbrace{\overbrace{- \eta \frac{dy}{dt}}^{= \eta N(u,\alpha)}}_{\text{LSS term}} = 0$$

Tangent form of the LSS (cont.)

Definition: LSS, tangent form

The LSS tangent (i.e. direct) sensitivity is given by the solution to the least-squares problem

$$\min_{v,\eta} \quad \frac{1}{2}\int_0^T \|v(t)\|^2 \, dt + \frac{\mu}{2}\int_0^T \eta(t)^2 \, dt,$$

$$\text{s.t.} \quad \frac{dv}{dt} + N'[u]v + N'[\alpha] + \eta N(u,\alpha) = 0$$

* note the absence of I.C. on v

Tangent form of the LSS (cont.)

- An adjoint version of the LSS can also be derived [WHB14]
- The LSS adjoint optimization statement is similar, but the constraint is replaced with the adjoint equation (with an LSS-specific term).

Solving for the LSS tangent

The LSS tangent (and adjoint) optimization problem is "nice," because the constraint is linear and the objective is a convex quadratic.

- this guarantees a unique solution, which can be found by solving the (linear) first-order optimality conditions

Unfortunately, the linear system that must be solved is huge.

Solving for the LSS tangent (cont.)

To obtain additional insight into the LSS problem, we now derive the first-order optimality conditions.

$$L(v, w, \eta) = \frac{1}{2}\int_0^T (v, v)\, dt + \frac{\mu}{2}\int_0^T \eta(t)^2\, dt$$

$$+ \int_0^T \left(w, \frac{dv}{dt} + N'[u]v + N'[\alpha] + \eta N(u,\alpha)\right) dt$$

$$D_w L\, \delta w = \int_0^T \left(\delta w, \underbrace{\frac{dv}{dt} + N'[u]v + N'[\alpha] + \eta N(u,\alpha)}_{\therefore\ =0}\right) dt = 0$$

$$\Rightarrow \quad \frac{dv}{dt} + N'[u]v + N'[\alpha] + \eta N(u,\alpha) = 0$$

Solving for the LSS tangent (cont.)

Recall $(w, v) = w^T v$

$$D_\eta L \delta\eta = \int_0^T \left(\mu \eta \delta\eta + w^T N(u,\alpha) \delta\eta \right) dt = 0$$

$$\Rightarrow \quad \mu \eta(t) + w^T N(u,\alpha) = 0$$

$$D_v L \delta v = \int_0^T \left(\delta v^T v + w^T \frac{d}{dt} \delta v + w^T N'[u] \delta v \right) dt = 0$$

$$= \int_0^T \left(\delta v, -\frac{dw}{dt} + N'[u]^* w + v \right) dt$$

$$+ \left[w^T \delta v \right]_{t=0}^{T} = 0$$

Solving for the LSS tangent (cont.)

The LSS tangent solution v, time-dilation η, and multipliers w satisfy

$$\begin{aligned} \frac{dv}{dt} + N'[u]v + N'[\alpha] + \eta N(u,\alpha) &= 0, \qquad \forall\, t \in [0,T] \\ \mu\eta + w^T N(u,\alpha) &= 0, \qquad \forall\, t \in [0,T] \\ -\frac{dw}{dt} + N'[u]^* w + v &= 0, \qquad \forall\, t \in [0,T], \\ w(0) = 0, &\quad w(T) = 0. \end{aligned}$$

Solving for the LSS tangent (cont.)

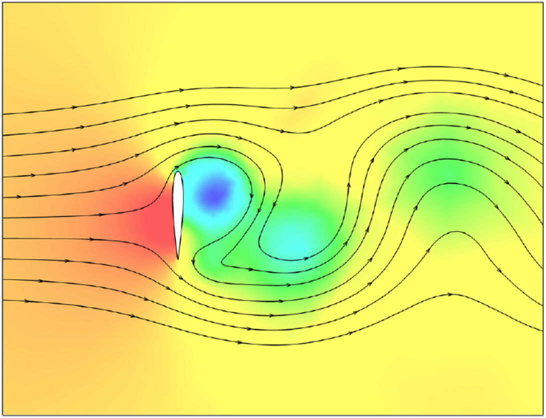
The most notable aspect of the problem defining v is that it is a boundary-value problem in time.

- To see this, one can eliminate v and η and obtain a second-order ODE in time for w with both initial and terminal conditions.

There have been limited uses of the LSS method, given the size of the linear system involved

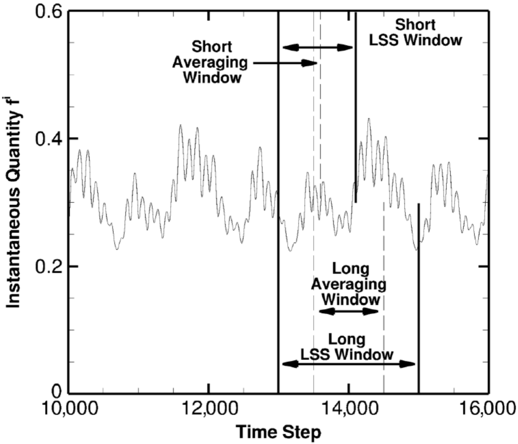
- For example, in [BWND16] the authors solve for the chaotic flow around a 2D airfoil.
- Although the number of spatial DOF and time steps are relatively small, the memory requirements for the linear system is in the terabytes.

Results using LSS



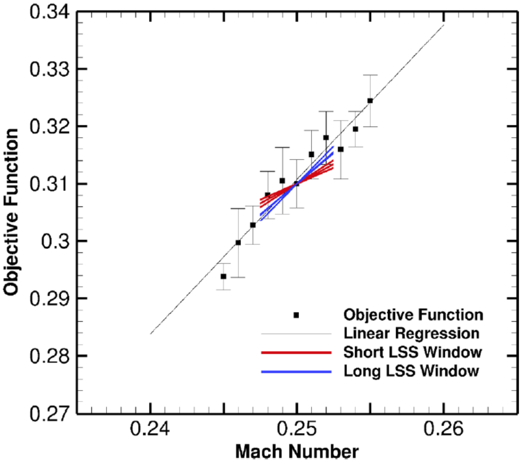
Chaotic airfoil flow studied in [BWND16].

Results using LSS (cont.)



Definition of the LSS and averaging windows [BWND16].

Results using LSS (cont.)



LSS-based sensitivities compared to trend [BWND16].

References

[BWND16] Patrick J. Blonigan, Qiqi Wang, Eric J. Nielsen, and Boris Diskin, *Least Squares Shadowing Sensitivity Analysis of Chaotic Flow around a Two-Dimensional Airfoil*, 54th AIAA Aerospace Sciences Meeting, American Institute of Aeronautics and Astronautics, January 2016.

[LAH00] Daniel J. Lea, Myles R. Allen, and Thomas W. N. Haine, *Sensitivity analysis of the climate of a chaotic system*, Tellus A **52** (2000), no. 5, 523–532.

[WHB14] Qiqi Wang, Rui Hu, and Patrick Blonigan, *Least Squares Shadowing sensitivity analysis of chaotic limit cycle oscillations*, Journal of Computational Physics **267** (2014), 210–224.