Chapter 6
Adjoint-based Local Sensitivity Analysis

6.1 Adjoint equations for linear, steady-state models

Adjoints are useful in sensitivity analysis because they can give information about any perturbed quantity with only one solve of the forward system, and one of the adjoint equations. The cost of these two solutions compare to the $I+1$ "forward" solutions needed to compute the sensitivities as in the previous section. The difference is that an adjoint solve is needed for each QoI, whereas the finite difference approach requires a forward solve for each parameter, regardless of the number of QoIs. On balance, when the number of QoIs is small relative to the number of uncertain parameters, the adjoint approach can be more efficient. The issue with the adjoint approach is that can be difficult to define the adjoint equations. In this section we will deal with linear, time-independent partial differential equations. In the next chapter we relax this assumption with a concomitant increase in complexity.

6.1.1 Definition of Adjoint Operator

To define an adjoint, let's define an inner product:

$$(f,g) = \int_D dV\, fg \tag{6.1}$$

where f and g are functions, D is the phase space domain of the functions, and dV is a differential phase space element. The adjoint for an operator L is typically denoted $L^\dagger$, and is defined as

$$(Lu, u^\dagger) = (u, L^\dagger u^\dagger), \tag{6.2}$$

using the definition of the inner product above. Using this definition, it is easy to show that adjoints make taking inner products of solution variables trivial if the adjoint is known.

For a PDE with differential operator L,

$$Lu = q,$$

with adjoint $L^\dagger$ of L, and an adjoint equation

$$L^\dagger u^\dagger = p,$$

then by the above definition:

$$(Lu, u^\dagger) = (u, L^\dagger u^\dagger) = (u, p) = (q, u^\dagger). \tag{6.3}$$

In other words, the inner product of u and p is the same as the inner product of q and $u^\dagger$.

Now consider a quantity of interest Q given by an integral of the solution u against a weighting function:

$$Q = \int_D dV\, p(\mathbf{r}) u(\mathbf{r}) = (u, p) \tag{6.4}$$

where $p(\mathbf{r})$ can be defined to take the average of u over a particular region of phase space. Equation (6.4) indicates that we can define a QoI as an inner product by picking a weighting function. Using relation Eq. (6.3) above, we see that the Q is just $(u^\dagger, q)$. In other words, the adjoint solution with source $p(\mathbf{r})$ integrated against the source q gives the Q. This is not magical, however, because the adjoint equation is typically as hard to solve as the original PDE as we will see in an example.

We now make the notion of the adjoint concrete for the steady ADR equation with a linear reaction term for $u(x)$ on the domain $(0, X)$ with zero Dirichlet boundary conditions. Under these conditions the ADR equation and boundary conditions are

$$\begin{aligned} v\frac{du}{dx} - \omega\frac{d^2u}{dx^2} + \kappa u &= q \\ u(0) = u(X) &= 0 \end{aligned} \tag{6.5}$$

Using the notation above, we define the operator L as

$$L = v\frac{d}{dx} - \omega\frac{d^2}{dx^2} + \kappa. \tag{6.6}$$

For this domain the inner product given by

$$(u, v) = \int_0^X uv\, dx. \tag{6.7}$$

We will postulate an adjoint form of this system and then show that it satisfies the definition in Eq. (6.2). The form of the adjoint we propose is basically the same equation, with the sign of the advection term flipped:

$$\begin{aligned} L^\dagger &= -v\frac{d}{dx} - \omega\frac{d^2}{dx^2} + \kappa \\ u^\dagger(0) &= u^\dagger(X) = 0. \end{aligned} \tag{6.8}$$

Proof. We need to show $(Lu, u^\dagger) = (u, L^\dagger u^\dagger)$ which is equivalent to:

$$\int_0^X \left(vu^\dagger\frac{du}{dx} - \omega u^\dagger\frac{d^2u}{dx^2} + \kappa u^\dagger u \right) dx = \int_0^X \left(-vu\frac{du^\dagger}{dx} - \omega u\frac{d^2u^\dagger}{dx^2} + \kappa uu^\dagger \right) dx. \tag{6.9}$$

We will show that these are equivalent term by term. While the $\kappa uu^\dagger$ term is obvious, the advection term needs integration by parts:

$$\int_0^X vu^\dagger\frac{du}{dx}\, dx = \cancel{vu^\dagger u\Big|_0^X} - v\int_0^X u\frac{du^\dagger}{dx}\, dx = \int_0^X -vu\frac{du^\dagger}{dx}\, dx \tag{6.10}$$

which is the term on the RHS of Eq. (6.9). The diffusion term just needs integration by parts twice:

$$\int_0^X u^\dagger \frac{d^2u}{dx^2}\,dx = \cancel{u^\dagger \frac{du}{dx}\Big|_0^X} - \int_0^X \frac{du}{dx}\frac{du^\dagger}{dx}\,dx = \cancel{\frac{du^\dagger}{dx}u\Big|_0^X} + \int_0^X u\frac{d^2u^\dagger}{dx^2}\,dx \tag{6.11}$$

which matches the diffusion term on the RHS of Eq. (6.9). □

With the known value of the adjoint ADR equation, we can use it to compute a QoI. As an example, if our QoI is the average of u over the middle third of the domain, this would make $p(x)$:

$$p(x) = \begin{cases} \frac{3}{X} & x \in [\frac{X}{3}, \frac{2}{3}X] \\ 0 & \text{otherwise} \end{cases}, \tag{6.12}$$

which leads to a Q:

$$Q = \int_{\frac{X}{3}}^{\frac{2}{3}X} \frac{3}{X}u(x)\,dx. \tag{6.13}$$

To get our QoI we could solve

$$Lu = q \quad \text{or} \quad L^\dagger u^\dagger = p, \tag{6.14}$$

and compute

$$Q = (u,p) = (q,u^\dagger). \tag{6.15}$$

The choice of which equation to solve is seemingly immaterial: each involves solving an ADR-like equation and then computing the inner product. There are instances where having an estimate of the adjoint solution can make the forward problem easier to solve, a salient example being source-detector problems in Monte Carlo particle transport simulations [Wagner and Haghighat(1998)]. The reason that this works is that the adjoint solution is, in a sense, a measure of how important a region of space is to the QoI.

6.1.2 Adjoints for Computing Derivatives

Our interest in the adjoint solution arises from the manner in which they allow first-order sensitivities to be computed. In some situations, this is called perturbation analysis, but as we will see it is the same as the sensitivity analyses discussed above. Consider the perturbed problem:

$$(L+\delta L)(u+\delta u) = q + \delta q \tag{6.16}$$

where δL and δq are perturbations to the original problem and δu is the change in the solution due to changing the problem. In the ADR example, the δL would involve changing the advection speed, diffusion coefficient, or reaction operator and δq would be a change to the source.

Expanding the product on the LHS of Eq. (6.16) we get

$$Lu + L\delta u + \delta L u = q + \delta q + \mathrm{O}(\delta^2). \tag{6.17}$$

Henceforth, we will ignore second order perturbations, i.e., the δ^2 terms. Now, $Lu = q$ so those terms can be cancelled to give:

$$L\delta u + \delta L u = \delta q. \tag{6.18}$$

Upon multiplying by $u^\dagger$ and taking the inner product, this becomes

$$(L\delta u, u^\dagger) + (\delta L u, u^\dagger) = (\delta q, u^\dagger). \tag{6.19}$$

This equation is useful, except we do not know what δu is. It is simple to compute the perturbation to L and apply it to a known forward solution u (this just involves taking derivatives). Similarly, we can compute δq easily because q is a parameter.

To remove the δu from Eq. (6.19) we will use the property of the adjoint that we can "switch" L and $L^\dagger$ in the inner product to make the relation

$$(L\delta u, u^\dagger) = (\delta u, L^\dagger u^\dagger) = (\delta u, p), \tag{6.20}$$

where $L^\dagger u^\dagger = p$ was used in the second equality. This makes Eq. (6.19)

$$(\delta u, p) + (\delta L u, u^\dagger) = (\delta q, u^\dagger). \tag{6.21}$$

Therefore, if we can get another relation for $(\delta u, p)$, then we can eliminate δu from our equations.

The definition of the perturbed QoI is

$$Q + \delta(Q) = \int_D dV\, pu + \int_D dV\, p\delta u + \int_D dV\, (\delta p)u. \tag{6.22}$$

We can rearrange this equation to get

$$(\delta u, p) = \delta(Q) - (u, \delta p).$$

Using this result in Eq. (6.21) gives an equation for the perturbation to the QoI in terms of perturbations to parameters and the forward and adjoint solution:

$$\delta(Q) = (\delta q, u^\dagger) + (u, \delta p) - (\delta L u, u^\dagger) \tag{6.23}$$

That is, if we know u and $u^\dagger$, we can compute $\delta(Q)$. In general, for a quantity θ, we interpret the quotient $\delta Q/\delta\theta$ as the partial derivative of the QoI with respect to θ. This interpretation is reasonable because the perturbation can be as small as we like because we placed no restrictions on its size. Therefore we write

$$\frac{\partial Q}{\partial \theta} = \left(\frac{\partial q}{\partial \theta}, u^\dagger\right) + \left(u, \frac{\partial p}{\partial \theta}\right) - \left(\frac{\partial L}{\partial \theta} u, u^\dagger\right). \tag{6.24}$$

This derivative formula gives us a way to compute sensitivity coefficients without taking finite difference derivatives. Also, for each parameter θ we can use the same $u^\dagger$ and u to compute the sensitivity by changing what goes into the inner product.

ADR Example: Perturbed κ

Let's see how this works on an example. In the ADR equation, set $\kappa \longrightarrow \kappa + \delta\kappa$ and assume there is no change to the source or the weighting function p. This makes $\delta L = \delta\kappa$. Therefore,

$$\delta Q = (\cancel{\delta q}, u^\dagger) + (u, \cancel{\delta p}) - (\delta\kappa u, u^\dagger), \tag{6.25}$$

where u solves $Lu = q$ and $u^\dagger$ solves $L^\dagger u^\dagger = p$. This result directly leads to

$$\frac{\partial Q}{\partial \omega} = -\left(\frac{\partial L}{\partial \omega} u, u^\dagger\right) = -(u, u^\dagger). \tag{6.26}$$

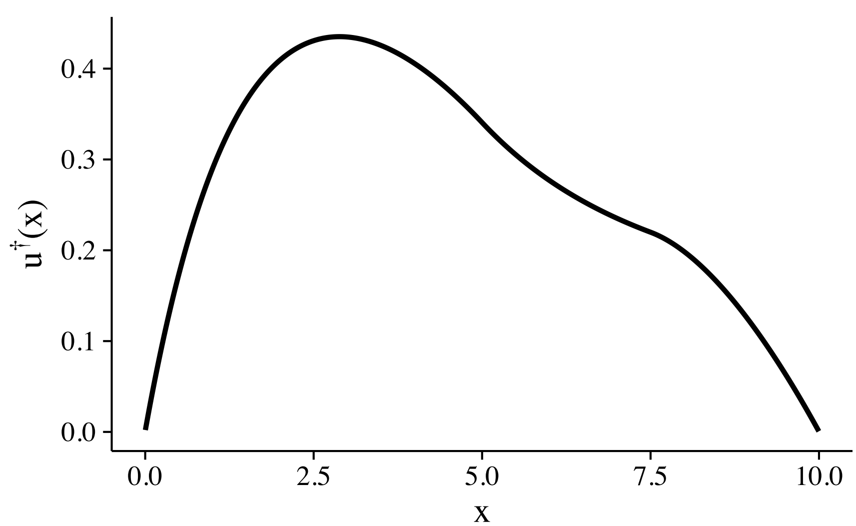

Fig. 6.1 The solution $u^\dagger(x)$ evaluated at $\bar{\theta}$.

ADR Example: Computing Derivatives from Each Parameter

Using the same data as the example in the previous chapter (Section 4.2) where the source and κ varied over space and the QoI was the total reaction rate, we can compute the sensitivities using the adjoint. Also, we can use the code from that example by simply running the code with $v \to -v$. The adjoint solution $u^\dagger$ at the mean of the all the parameters is shown in Figure 6.1. In this case if we compute the QoI using the forward or adjoint solution we get a match to 12 digits for the

$$(u,\kappa) = 52.3903954692 \qquad (S,u^\dagger) = 52.3903954692.$$

To compute the inner products we use simple quadrature based on the midpoint rule.

To compute the derivative of Q with respect to v we use Eq. (6.24) to get

$$\frac{\partial Q}{\partial v} = -\left(\frac{\partial u}{\partial x}, u^\dagger\right) = -1.74049052049.$$

The derivative of u must be estimated from the forward solution using finite differences. Similarly, the derivative with respect to ω involves the integral of the second-derivative of the forward solution times the adjoint solution:

$$\frac{\partial Q}{\partial \omega} = \left(\frac{\partial^2 u}{\partial x^2}, u^\dagger\right) = -0.970207772262.$$

For κ_l the derivative is based on an integral only over the range $x \in (5,7.5)$:

$$\frac{\partial Q}{\partial \kappa_l} = \int_5^{7.5} u(x)\,dx - \int_5^{7.5} u(x)u^\dagger(x)\,dx = 12.862742303.$$

The sensitivity to κ_h is an integral over the other parts of the problem:

$$\frac{\partial Q}{\partial \kappa_h} = \int_0^5 u(x)\,dx + \int_{7.5}^{10} u(x)\,dx - \int_0^5 u(x)u^\dagger(x)\,dx - \int_{7.5}^{10} u(x)u^\dagger(x)\,dx = 17.7613932101.$$

The final sensitivity to compute is involves the source strength, q. From Eq. (6.24) we get

$$\frac{\partial Q}{\partial q} = (x(10-x), u^\dagger) = 52.3903954692.$$

These results all agree with the first-order derivative results in Table 4.2 to several digits.

6.2 Adjoints for nonlinear, time-dependent equations

In the previous section we had to make some strong assumptions about the underlying mathematical model to use adjoints. In this section we relax that assumption and show how an adjoint equation can be formed. To begin we have a time-dependent PDE of the form

$$F(u,\dot{u}) = 0, \tag{6.27}$$

where $\dot{u}$ is the time derivative of u. We also have boundary conditions such that the solution goes to zero on the boundary of the domain of interest. As before, we define an inner product (u,v) as the integral of uv over phase-space. It will be useful to write the QoI as an integral over phase-space and time separately. In particular

$$Q = \int_{t_0}^{t_\mathrm{f}} (u,p)\,dt. \tag{6.28}$$

We can modify this equation for the QoI by adding a Lagrange multiplier, $u^\dagger$, times F without changing the QoI because $F(u,\dot{u}) = 0$. We call this new quantity the adjoined metric and write it as

$$\mathscr{L} = \int_{t_0}^{t_\mathrm{f}} \left[(u,p) - (F,u^\dagger)\right]\,dt. \tag{6.29}$$

The first-order sensitivity (i.e., first variation) to a parameter for this equation for a functional $(g(u),u^\dagger)$ is defined as is

$$\frac{d}{d\theta}(g(u),u^\dagger) = \frac{\partial}{\partial u}(g(u),u^\dagger)\frac{\partial u}{\partial \theta}.$$

Using this definition we get the first-order sensitivity as

$$\frac{d\mathscr{L}}{d\theta} = \int_{t_0}^{t_\mathrm{f}} \left[(1,p)u_\theta + (u,p_\theta) - \frac{\partial}{\partial \dot{u}}(F,u^\dagger)\dot{u}_\theta - \frac{\partial}{\partial u}(F,u^\dagger)u_\theta - \frac{\partial}{\partial \theta}(F,u^\dagger)\right]\,dt. \tag{6.30}$$

where subscripts indicate partial derivatives. As before we would like to eliminate the u_θ and $\dot{u}_\theta$ terms because we do not know the derivative of the solution or its time derivative with respect to the parameter. To make this elimination we first integrate the $\dot{u}_\theta$ term by parts to get

$$\begin{aligned}\frac{d\mathscr{L}}{d\theta} = &- \frac{\partial}{\partial \dot{u}}(F,u^\dagger)u_\theta\Big|_{t_0}^{t_\mathrm{f}} + \\ &+ \int_{t_0}^{t_\mathrm{f}} \left[(1,p)u_\theta + (u,p_\theta) + u_\theta\frac{d}{dt}\frac{\partial}{\partial \dot{u}}(F,u^\dagger) - \frac{\partial}{\partial u}(F,u^\dagger)u_\theta - \frac{\partial}{\partial \theta}(F,u^\dagger)\right]\,dt.\end{aligned} \tag{6.31}$$

In order to eliminate u_θ from this equation, we will define the Lagrange multiplier so that

$$(1,p) + \frac{d}{dt}\frac{\partial}{\partial \dot{u}}(F,u^\dagger) - \frac{\partial}{\partial u}(F,u^\dagger) = 0. \tag{6.32}$$

We have ignored the boundary term at t_0 and t_f for now. We will state that Eq. (6.32) holds at all points in space so that we get an equation for the adjoint $u^\dagger$.

The boundary term in Eq. (6.31) evaluates u_θ at $t = t_0$ and t_f. At the final time, we can state that $u^\dagger(t_\mathrm{f}) = 0$ to represent the fact that anything that happens beyond the final time does not contribute to the quantity of interest. The issue of the final condition on $u^\dagger$ indicates a subtlety of the adjoint equation: it runs backward in time. One has to solve the adjoint equation starting at the final time and solve backward to get to t_0. At $t = t_0$ this term indicates how the initial conditions for u are perturbed by the parameter. Therefore, we only need to consider this quantity if the initial conditions are dependent on θ.

Upon solving Eq. (**??**) we can compute the sensitivity of the QoI to parameter θ through the equation:

$$\frac{d\mathscr{L}}{d\theta} = -\left.\frac{\partial}{\partial \dot{u}}(F,u^\dagger)u_\theta\right|_{t_0} + \int_{t_0}^{t_\mathrm{f}} \left[(u,p_\theta) - \frac{\partial}{\partial \theta}(F,u^\dagger)\right] dt. \tag{6.33}$$

Notice that to evaluate the equation we need the full forward solution and adjoint solution for all time to compute the integrals. This could represent a storage problem for large scale systems.

Linear ADR Equation

The above derivation of the adjoint for a time-dependent problem was fairly abstract. We shall show it applies to the ADR problem we have seen before. For the linear ADR equation we saw before, the system can be written in as $F(u,\dot{u}) = 0$ where

$$F(u,\dot{u}) = \dot{u} + v\frac{\partial u}{\partial x} - \omega\frac{\partial^2 u}{\partial x^2} + \kappa u - S.$$

We also consider a problem domain given by $x \in [0,X]$, with $u(0,t) = u(X,t) = 0$. For a generic QoI weighting function p, terms in Eq. (6.32) are

$$\frac{d}{dt}\frac{\partial}{\partial \dot{u}}(F,u^\dagger) = \frac{d}{dt}\frac{\partial}{\partial \dot{u}}\int_0^X u^\dagger(x,t)\left(\dot{u} + v\frac{\partial u}{\partial x} - \omega\frac{\partial^2 u}{\partial x^2} + \kappa u - S\right)dx = \int_0^X \frac{\partial u^\dagger}{\partial t}\,dx.$$

In this equation the derivative is simple to compute because $\dot{u}$ only appears in a single term. The other term in the definition of the adjoint will require integration by parts. This term is

$$\begin{aligned}\frac{\partial}{\partial u}(F,u^\dagger) &= \frac{\partial}{\partial u}\int\limits_0^X u^\dagger(x,t)\left(\dot{u}+v\frac{\partial u}{\partial x}-\omega\frac{\partial^2 u}{\partial x^2}+\kappa u - S\right)dx\\ &= \frac{\partial}{\partial u}\int\limits_0^X u(x,t)\left(-v\frac{\partial u^\dagger}{\partial x}-\omega\frac{\partial^2 u^\dagger}{\partial x^2}+\kappa u^\dagger\right)dx\\ &= \int\limits_0^X \left(-v\frac{\partial u^\dagger}{\partial x}-\omega\frac{\partial^2 u^\dagger}{\partial x^2}+\kappa u^\dagger\right)dx.\end{aligned}$$

Where we used integration by parts to move the derivatives onto the adjoint variables. In doing so we relied on the fact that $u(0,t)=u(X,t)=0$ and that we are free to define the boundary conditions for $u^\dagger$ to be $u^\dagger(0,t)=u^\dagger(X,t)=0$.

If we assert that Eq. (6.32) hold at every point in the medium, then we have the adjoint equation

$$-\frac{\partial u^\dagger}{\partial t}-v\frac{\partial u^\dagger}{\partial x}-\omega\frac{\partial^2 u^\dagger}{\partial x^2}+\kappa u^\dagger = p,$$

with boundary and final conditions,

$$u^\dagger(0,t)=u^\dagger(X,t)=0, \qquad u(x,t_\mathrm{f})=0.$$

This equation is the time-dependent version of Eq. (6.8). As a result our new approach to deriving an adjoint is equivalent to the previous one, except now we can, in principle handle more complicated equations.

Nonlinear ADR Equation

As an example of a more complicated PDE that we can derive an adjoint for, we will look at the ADR equation with a nonlinear reaction term. In this case we redefine $F(u,\dot{u})$ as

$$F(u,\dot{u})=\dot{u}+v\frac{\partial u}{\partial x}-\omega\frac{\partial^2 u}{\partial x^2}+\kappa u^4 - S.$$

Notice that the reaction is now proportional to the fourth-power of u. For this new form of F we will have to compute

$$\frac{\partial}{\partial u}\int\limits_0^X u^\dagger(x,t)\kappa u^4(x,t)\,dx = \int\limits_0^X 4u^3(x,t)\kappa u^\dagger(x,t).$$

As a result, the adjoint equation is

$$-\frac{\partial u^\dagger}{\partial t}-v\frac{\partial u^\dagger}{\partial x}-\omega\frac{\partial^2 u^\dagger}{\partial x^2}+4u^3\kappa u^\dagger = p.$$

This is a linear equation, but it requires knowledge of $u(x,t)$ at every time and point in space to evaluate.