Sensitivity of a performance measure to a single scalar parameter when a state variable is governed by a differential equation

Let us consider a situation in which a performance measure is a domain integral involving a state variable that is governed by a differential equation. We want to compute the sensitivity of such a performance measure to a scalar parameter. We illustrate this with an example.

Example 1

Compute $\frac{d}{dA}\left(\int_0^L pu(x)\,dx\right)$ where $EAu''(x)+p=0$. Note that only $u(x)$ is a function of x while p, E, and A are scalar parameters that do not depend on x. Let $u(x)$ satisfy the boundary conditions: $u(0)=0$ and $u'(L)=0$.

In this simple problem, the governing differential equation can be solved in closed form. That is,

$$u(x) = -\frac{px^2}{2EA} + \frac{pLx}{EA} \tag{1}$$

satisfies $EAu''(x)+p=0$ for the given boundary conditions. We also note that

$$\frac{du(x)}{dx} = u' = -\frac{px}{EA} + \frac{pL}{EA} \tag{2}$$

$$\frac{d^2u(x)}{dx^2} = u'' = -\frac{p}{EA} \tag{3}$$

We use the known solution of $u(x)$ to compute the sensitivity of $\frac{d}{dA}\left(\int_0^L pu(x)\,dx\right)$ in two ways: (i) integrate w.r.t. x first and then differentiate w.r.t. A. And (ii) differentiate w.r.t. A first and then integrate w.r.t. x.

<u>Method A: Closed-form solution</u>

We integrate w.r.t. x and then differentiate w.r.t. A to get:

$$\begin{aligned}
&\int_0^L pu(x)\,dx = \int_0^L p\left(-\frac{px^2}{2EA} + \frac{pLx}{EA}\right)dx = -\frac{p^2L^3}{6EA} + \frac{p^2L^3}{2EA} = \frac{p^2L^3}{3EA} \\
&\Rightarrow \frac{d}{dA}\left(\int_0^L pu(x)\,dx\right) = \frac{d}{dA}\left(\frac{p^2L^3}{3EA}\right) = -\frac{p^2L^3}{3EA^2}
\end{aligned} \tag{4}$$

<u>Method B: Closed-form solution</u>

We first differentiate w.r.t. A and then integrate w.r.t. x.

$$\frac{d}{dA}\left(\int_0^L pu(x)\,dx\right)=\frac{d}{dA}\left\{\int_0^L p\left(-\frac{px^2}{2EA}+\frac{pLx}{EA}\right)dx\right\}=\int_0^L\frac{d}{dA}\left(-\frac{p^2x^2}{2EA}+\frac{p^2Lx}{EA}\right)dx$$
$$=\int_0^L\left(\frac{p^2x^2}{2EA^2}-\frac{p^2Lx}{EA^2}\right)dx=\frac{p^2L^3}{6EA^2}-\frac{p^2L^3}{2EA^2}=-\frac{p^2L^3}{3EA^2} \tag{5}$$

The results of Methods A and B agree with each other. This tells us that parameter sensitivity and integration commute. That is, the order in which we differentiate and integrate does not matter. This can be stated as a general result to reckon later.

For $\phi=\int_0^L f(x,p)\,dx$, $\frac{d\phi}{dp}=\frac{d}{dp}\left(\int_0^L f(x,p)\,dx\right)=\int_0^L\frac{df(x,p)}{dp}dx$.

More generally, for any domain Ω,

$$\frac{d\phi}{dp}=\frac{d}{dp}\left(\int_\Omega f(x,p)\,d\Omega\right)=\int_\Omega\frac{df(x,p)}{dp}d\Omega$$

But there is a caveat here: the parameter must not change the domain of interest. When the domain changes because of a parameter, that parameter becomes a shape-changing parameter. That would complicate matters and we return to this point later.

The commutative rule applies to differentiation too. If there is a function that depends on a spatial variable x and a parameter p, then too, the order of differentiating w.r.t. x and p does not matter. That is, parameter sensitivity and spatial gradient commute. This leads to another general result to remember.

$\frac{\partial f'}{\partial p}=\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial p}\right)$ where $f'=\frac{\partial f}{\partial x}$

If there are multiple spatial variables, e.g., x,y,z, we can write

$$\nabla f'_p=\left(\nabla f\right)'_p$$

Let us now return to solving the example without using the explicit analytical solution of the governing equation. Here too, there are two ways to do it: (i) adjoint method and (ii) direct method. In the adjoint method, we do not compute the sensitivity of the state variable (here, $u(x)$) whereas we do that in the direct method.

Direct method

Recall that in this example, we need to compute $\frac{d}{dA}\left(\int_0^L pu(x)\,dx\right)$ where $EAu''(x)+p=0$ with boundary conditions: $u(0)=0$ and $u'(L)=0$. Note that $\frac{d}{dA}\left(\int_0^L pu(x)\,dx\right)=\int_0^L p\frac{du}{dA}dx$. In the direct method, we try to compute $\frac{du}{dA}$.

First, we differentiate the governing differential equation w.r.t. A to get

$$Eu''+EA\frac{du''}{dA}(x)=0 \Rightarrow Eu''+EA\frac{d^2}{dx^2}\left(\frac{du}{dA}\right)=0 \tag{6}$$

Now, we have a differential equation involving $\frac{du}{dA}$. The loading term in the equation is (Eu''). If this is known, we can solve for $\frac{du}{dA}$. We already solved for $u(x)$ using $EAu''(x)+p=0$, here in closed form, to get $u(x)=-\frac{px^2}{2EA}+\frac{pLx}{EA}$ (see Eq. (1)), yielding $u''=-\frac{p}{EA}$ (Eq. (3)). Thus, we have

$$\begin{aligned}
&Eu''+EA\frac{d^2}{dx^2}\left(\frac{du}{dA}\right)=0\\
&\Rightarrow E\left(-\frac{p}{EA}\right)+EA\frac{d^2}{dx^2}\left(\frac{du}{dA}\right)\\
&\Rightarrow \frac{d^2}{dx^2}\left(\frac{du}{dA}\right)=\frac{p}{EA^2}\\
&\Rightarrow \frac{d}{dx}\left(\frac{du}{dA}\right)=\frac{px}{EA^2}+C_1\\
&\Rightarrow \frac{du}{dA}=\frac{px^2}{2EA^2}+C_1x+C_2
\end{aligned} \tag{7}$$

We know the boundary conditions on $\frac{du}{dA}$ as

$$\left.\frac{du}{dA}\right|_{x=0}=0 \text{ and } \left.\frac{du'}{dA}\right|_{x=L}=0 \tag{8}$$

since u and u' are zero at $x=0$ and $x=L$, respectively. As they are specified, their derivatives are zero w.r.t. A. With the boundary conditions in Eq. (8), we can obtain C_1 and C_2:

$$\left.\frac{du}{dA}\right|_{x=0} = 0 \Rightarrow C_2 = 0 \text{ and } \left.\frac{du'}{dA}\right|_{x=L} = 0 \Rightarrow C_1 = -\frac{pL}{EA^2}$$

leading to

$$\frac{du}{dA} = \frac{px^2}{2EA^2} - \frac{pLx}{EA^2} \tag{9}$$

Now, we have

$$\frac{d}{dA}\left(\int_0^L pu(x)\,dx\right) = \int_0^L p\frac{du}{dA}dx = \int_0^L p\left(\frac{px^2}{2EA^2} - \frac{pLx}{EA^2}\right)dx = \frac{p^2L^3}{6EA^2} - \frac{p^2L^3}{2EA^2} = -\frac{p^2L^3}{3EA^2}$$

which is the correct sensitivity we got earlier using the closed-form solution.

Adjoint method

Here, we add the left-side expression of the governing equation (here, $\left(EAu'' + p\right)$) to the functional whose sensitivity is to be found. There is no problem in doing so because that expression is zero and hence does not change the value of functional. In fact, we multiply the expression by $\lambda(x)$, an arbitrary function and then add to the functional. The reason for multiplying with $\lambda(x)$ becomes clear in the later steps. For now, we define $\mathrm{L} = \int_0^L \left\{pu + \lambda\left(EAu'' + p\right)\right\}dx$ and note that the sensitivity of L will be the same as that of the functional.

Next, we differentiate L with respect to A and do integration by parts twice to get

$$\begin{aligned}
\frac{d\mathrm{L}}{dA} &= \frac{d}{dA}\left[\int_0^L \left\{pu + \lambda\left(EAu'' + p\right)\right\}dx\right] = \int_0^L \left[\frac{d}{dA}\left\{pu + \lambda\left(EAu'' + p\right)\right\}\right]dx \\
&= \int_0^L \left(p\frac{du}{dA} + \lambda Eu'' + \lambda EA\frac{du''}{dA}\right)dx \\
&= \int_0^L \left(p\frac{du}{dA} + \lambda Eu'' - \lambda' EA\frac{du'}{dA}\right)dx + \left.\left(\lambda EA\frac{du'}{dA}\right)\right|_0^L \\
&= \int_0^L \left(p\frac{du}{dA} + \lambda Eu'' + \lambda'' EA\frac{du}{dA}\right)dx - \left.\left(\lambda' EA\frac{du}{dA}\right)\right|_0^L + \left.\left(\lambda EA\frac{du'}{dA}\right)\right|_0^L
\end{aligned} \tag{10}$$

where the first line entails switching the order of integration and differentiation and the last two lines involve integration by parts. The integral in Eq. (10) can be reduced to avoid the computation of $\frac{du}{dA}$ by requiring that its coefficient term be made equal to zero:

$$EA\lambda'' + p = 0 \tag{11}$$

This means that we need to find $\lambda(x)$ such that Eq. (11) is satisfied along with the boundary conditions that appear in the last line of Eq. (10). This is the reason why we multiplied by $\lambda(x)$ while defining L. Note that L is called the Lagrangian and $\lambda(x)$ is called the adjoint variable.

By examining Eq. (11) and the governing equation, we observe that $\lambda = u$. In order to satisfy the boundary conditions in Eq. (10), i.e., $\left(\lambda' EA \frac{du}{dA}\right)\Big|_0^L = 0$ and $\left(\lambda EA \frac{du'}{dA}\right)\Big|_0^L = 0$, we take $\lambda'(L) = 0$ and $\lambda(0) = 0$. This follows the boundary conditions of $u(x)$. We also note that since $u(0)$ and $u'(L)$ are specified, their sensitivities w.r.t. A are zero. Hence, both boundary conditions are satisfied at $x = 0, L$. Thus, Eq. (10) reduces to

$$\frac{dL}{dA} = \int_0^L \left(\lambda E u''\right) dx = \int_0^L \left(u E u''\right) dx \tag{12}$$

With the help of Eqs. (1) and (3), Eq. (12) can be evaluated as follows.

$$\begin{aligned} \frac{dL}{dA} &= \int_0^L \left(u E u''\right) dx = \int_0^L E\left(-\frac{px^2}{2EA} + \frac{pLx}{EA}\right)\left(-\frac{p}{EA}\right) dx \\ &= \int_0^L \left(-\frac{px^2}{2EA} + \frac{pLx}{EA}\right)\left(-\frac{p}{A}\right) dx = \frac{p^2}{EA^2}\left(\frac{L^3}{6} - \frac{L^3}{2}\right) = -\frac{p^2 L^3}{3EA^2} \end{aligned} \tag{13}$$

Thus, we see that we got the same result for $\frac{d}{dA}\left(\int_0^L p u(x)\, dx\right)$ without having to compute $\frac{du}{dA}$.

This example showed us that the derivative of a performance measure that includes a variable governed by a differential equation can be easily calculated by introducing an adjoint variable (here, λ) and the corresponding adjoint equation that governs λ. To see this, let us consider another example.

Example 2

Compute $\frac{d}{dp}\left\{\int_0^T \left(p^2 + g^2\right) dt\right\}$ where $\dot{g} + pg = 0$ and $g(0) = 1$.

<u>Using the closed-form solution</u>

The solution of $\dot{g} + pg = 0$ for the given boundary condition can be obtained as

$$g = e^{-pt} \tag{14}$$

By substituting the preceding solution into the given domain integral and then differentiation w.r.t. p, we get

$$\frac{d}{dp}\left\{\int_0^T \left(p^2+g^2\right)dt\right\} = \int_0^T \frac{d}{dp}\left(p^2+g^2\right)dt = \int_0^T \frac{d}{dp}\left(p^2+e^{-2pt}\right)dt$$
$$= \int_0^T \left(2p-2te^{-2pt}\right)dt = 2pT + T\frac{e^{-2pT}}{p} + \frac{e^{-2pT}}{2p^2} - \frac{1}{2p^2} \tag{15}$$

We can also do this by differentiating w.r.t. p first and then integrating.

$$\frac{d}{dp}\left\{\int_0^T \left(p^2+g^2\right)dt\right\} = \int_0^T \frac{d}{dp}\left(p^2+g^2\right)dt = \int_0^T \left(2p+2g\frac{dg}{dp}\right)dt \tag{16}$$

Using Eq. (10), we get

$$g = e^{-pt}$$
$$\Rightarrow \frac{dg}{dp} = -te^{-pt} \tag{17}$$

Eqs. (16) and (17) give

$$\frac{d}{dp}\left\{\int_0^T \left(p^2+g^2\right)dt\right\} = \int_0^T \left(2p-2te^{-2pt}\right)dt$$
$$= 2pT + T\frac{e^{-2pT}}{p} + \frac{e^{-2pT}}{2p^2} - \frac{1}{2p^2} \tag{18}$$

Let us now use the adjoint method where we do not compute $\frac{dg}{dp}$.

<u>Example 2 using the adjoint method</u>

Consider, as before,

$$\mathrm{L} = \int_0^T \left\{(p^2+g^2) + \lambda(\dot{g}+gp)\right\}dt \tag{19}$$

Now, we have

$$\frac{d\mathrm{L}}{dp} = \int_0^T \left\{\left(2p+2g\frac{dg}{dp}\right) + \lambda\left(\frac{d\dot{g}}{dp} + g + p\frac{dg}{dp}\right) + \frac{d\lambda}{dp}\left(\dot{g}+pg\right)\right\}dt \tag{20a}$$

where the last term in the integrand is zero because $\dot{g} + pg = 0$. Therefore, $\frac{d\lambda}{dp}$ does not enter the picture. By using integration by parts for $\left(\lambda \frac{d\dot{g}}{dp}\right)$ term in the integrand of Eq. (20a), we get

$$\frac{d\mathrm{L}}{dp} = \int_0^T \left\{ \left(2p + \lambda g\right) + \left(2g + p\lambda - \dot{\lambda}\right) \frac{dg}{dp} \right\} dt + \lambda \frac{dg}{dp}\bigg|_0^T \qquad (20b)$$

wherein we also collected terms that involve $\frac{dg}{dp}$ together. We now equate the terms that multiply $\frac{dg}{dp}$ to zero to get a differential equation that enables us to solve for λ. That is, we have

$$2g + p\lambda - \dot{\lambda} = 0 \text{ and } \lambda(L) = 0 \qquad (21)$$

Note that $\frac{dg}{dp}\bigg|_{x=0} = 0$ because $g(0) = 1$, a value that does not depend on p.

In the adjoint method, we use the solution of the governing equation. That is, we use $g = e^{-pt}$ from Eq. (14). By substituting this result in Eq. (21), we solve it to get

$$\lambda = -\frac{e^{-pt}}{p} + \left(\frac{e^{-2pT}}{p}\right) e^{pt} \qquad (22)$$

It may be verified that

$$\dot{\lambda} = e^{-pt} + \left(e^{-2pT}\right) e^{pt}$$

$$2g + p\lambda - \dot{\lambda} = 0$$

$$\Rightarrow 2e^{-pt} + p\left\{-\frac{e^{-pt}}{p} + \left(\frac{e^{-2pT}}{p}\right) e^{pt}\right\} - e^{-pt} - \left(e^{-2pT}\right) e^{pt} = 0$$

From Eq. (22) and the last line of Eq. (20b) give

$$
\begin{aligned}
\frac{d\mathrm{L}}{dp} &= \int_0^T (2p + \lambda g) dt \\
&= \int_0^T \left\{ 2p + \left(-\frac{e^{-pt}}{p} + \frac{e^{-2pT}}{p} e^{pt} \right) e^{-pt} \right\} dt \\
&= \int_0^T \left\{ 2p - \frac{e^{-2pt}}{p} + \frac{e^{-2pT}}{p} \right\} dt \\
&= \left(2pt + \frac{e^{-2pt}}{2p^2} + \frac{e^{-2pT}}{p} t \right) \Bigg|_0^T \\
&= 2pT + \frac{e^{-2pT}}{2p^2} - \frac{1}{2p^2} + \frac{e^{-2pT}}{p} T
\end{aligned}
\tag{23}
$$

which is the same as the result in Eq. (18).

It may be noted that the adjoint method works even if a closed-form solution of the governing differential equation is not possible. In such a case both the governing and adjoint differential equations can be solved numerically. The important thing to understand here is that in the adjoint method, we do not solve the sensitivity of the state variable.

To contrast the adjoint method with the direct method of sensitivity analysis, let us solve the same example using the direct method.

<u>Example 2 using the direct method</u>

Recall that in this example, we need to compute $\frac{d}{dp}\left(\int_0^T \left(p^2 + g^2 \right) dt \right)$ where $\dot{g} + pg = 0$ with boundary conditions: $g(0) = 1$.

First, we differentiate the governing differential equation w.r.t. p to get

$$
\frac{d\dot{g}}{dp} + g + p\frac{dg}{dp} = 0 \tag{24}
$$

In this simple example, we know $g = e^{-pt}$ and $\frac{dg}{dp} = -te^{-pt}$ in closed form. Thus, Eq. (24) becomes

$$
\begin{aligned}
&\frac{d\dot{g}}{dp} + e^{-pt} - pte^{-pt} = 0 \\
&\Rightarrow \frac{d\dot{g}}{dp} = pte^{-pt} - e^{-pt} \\
&\Rightarrow \frac{d}{dt}\left(\frac{dg}{dp}\right) = pte^{-pt} - e^{-pt} \\
&\Rightarrow \frac{dg}{dp} = \int \left(pte^{-pt} - e^{-pt} \right) dt + C
\end{aligned}
\tag{25}
$$

We can now substitute this into $\frac{d}{dp}\left(\int_0^T \left(p^2+g^2\right)dt\right) = \int_0^T \frac{d}{dp}\left(p^2+g^2\right)dt = \int_0^T \left(2p + 2g\frac{dg}{dp}\right)dt$.

Direct and adjoint methods: which is better?

In the direct method, we differentiate the governing differential equation w.r.t. the parameter, and obtain a new differential equation where the variable is the derivative of the state variable. We need to solve this new differential equation and substitute it directly into the sensitivity expression. It is important to notice that in the adjoint method, we avoid computing the sensitivity of the state variable. Instead, we solve an adjoint differential equation. Both the direct method and the adjoint method are amenable for numerical solution when closed-form expressions are not available. Is there an advantage of one method over the other? The adjoint method is advantageous if we need to take the sensitivity w.r.t. many parameters. This is because, in the direct method, we need to solve the differential equation governing the sensitivity of the state variable w.r.t. each parameter. On the other hand, in the adjoint method, we need to solve the differential equation governing the adjoint variable only once. Consider the following problem to see this.

Example 3 as an exercise

Compute $\frac{d}{dA_0}\left(\int_0^L pu(x)\,dx\right)$ and $\frac{d}{dA_1}\left(\int_0^L pu(x)\,dx\right)$ where $\frac{d}{dx}\left\{EAu'(x)\right\} + p = 0$ with $A(x) = A_1\left(L-x\right) + A_0$. Note that only $u(x)$ is a function of x while p, E, A_0 and A_1 are scalar parameters that do not depend on x. Let $u(x)$ satisfy the boundary conditions: $u(0)=0$ and $u'(L)=0$.