

# Sensitivity of differential-algebraic equations

Yingbo Ma  
mayingbo5@gmail.com

March 12, 2020

## 1 Notation

- $M \in \mathbb{C}^{m \times m}$  denotes a mass matrix.
- $t_0 < t_1 \in \mathbb{C}$  denote constants.
- $t \in \mathbb{C}$  denotes independent variable.
- $p \in \mathbb{C}^n$  denotes parameters.
- $u(p, t) : \mathbb{C}^n \times \mathbb{C} \mapsto \mathbb{C}^m$  denotes states.
- $f(u(p, t), p, t) : \mathbb{C}^m \times \mathbb{C}^n \times \mathbb{C} \mapsto \mathbb{C}^m$  denotes the right-hand-side of a differential equation.
- $S(p, t) : \mathbb{C}^n \times \mathbb{C} \mapsto \mathbb{C}^{m \times n}$  denotes  $\frac{\partial u}{\partial p}$  (sensitivity with respect to the states).
- $J(u(p, t), p, t) : \mathbb{C}^m \times \mathbb{C}^n \times \mathbb{C} \mapsto \mathbb{C}^{m \times m}$  denotes  $\frac{\partial f}{\partial u}$  (Jacobian of the differential equation with respect to the states).
- $g(u(p, t), p, t) : \mathbb{C}^m \times \mathbb{C}^n \times \mathbb{C} \mapsto \mathbb{C}$  is a cost function that is sufficiently smooth.
- $\lambda(u(p, t), p, t, t_0, t_1) : \mathbb{C}^m \times \mathbb{C}^n \times \mathbb{C} \times \mathbb{C} \times \mathbb{C} \mapsto \mathbb{C}^n$  denotes Lagrange multiplier.
- $\lambda_\tau$  denotes  $\frac{\partial \lambda}{\partial \tau}$ .

## 2 Introduction and Forward Sensitivity Analysis

In many applications, we may wish to compute the gradient of the continuous cost function

$$G[u] = \int_{t_0}^{t_1} g(u(p, t), p, t) dt \quad (1)$$

with respect to the parameters  $\frac{dG}{dp}$ , where  $u$  is a function that satisfies the differential-algebraic equation

$$M\dot{u} = f(u(p, t), p, t), \quad (2)$$

$$u(t_0) = u_0(p). \quad (3)$$

Naïvely, we could apply Leibniz rule for integration and obtain:

$$\frac{dG}{dp} = \frac{d}{dp} \int_{t_0}^{t_1} g(u(p, t), p) dt \quad (4)$$

$$= \int_{t_0}^{t_1} \frac{d}{dp} g(u(p, t), p) dt \quad (5)$$

$$= \int_{t_0}^{t_1} \frac{\partial g}{\partial u} \frac{\partial u}{\partial p} + \frac{\partial g}{\partial p} dt \quad (6)$$

$$= \int_{t_0}^{t_1} \frac{\partial g}{\partial u} S + \frac{\partial g}{\partial p} dt \quad (7)$$

We can get  $S$  by differentiating eq. (2) with respect to  $p$  both sides,

$$\frac{d}{dp} \frac{d}{dt} M u = \frac{d}{dp} f(u(p, t), p, t) \quad (8)$$

$$\implies M \frac{d}{dt} \frac{du}{dp} = \frac{\partial f}{\partial u} \frac{du}{dp} + \frac{\partial f}{\partial p} \quad (9)$$

$$\implies M \dot{S} = JS + \frac{\partial f}{\partial p}. \quad (10)$$

eq. (10) is often referred as the forward sensitivity equation. It is apparent that computing  $S$  can be intractable when there are large number of parameters, because we would have to solve a  $m \times n$  differential equation.

### 3 Adjoint Sensitivity Analysis with Continuous Cost Function

To alleviate the computational cost of the forward sensitivity equation, we could add a “0” to  $\frac{dG}{dp}$ , and get:

$$\frac{dG}{dp} = \frac{dI}{dp} \quad (11)$$

$$= \int_{t_0}^{t_1} \frac{\partial g}{\partial u} S + \frac{\partial g}{\partial p} - \lambda^* \left( \underbrace{M \dot{S} - JS - \frac{\partial f}{\partial p}}_{0 \quad eq. (10)} \right) dt. \quad (12)$$

The motivation of this step is to introduce  $\dot{S}$  and an *arbitrary*  $\lambda$  function, and the hope is to use the classic technique of integration by parts to group the  $S$  terms and choose the  $\lambda$  function such that the gradient is *independent* of  $S$ .

After integration by parts, the  $\lambda^* M \dot{S}$  term is

$$\int_{t_0}^{t_1} \lambda^* M \dot{S} dt = \lambda^* M S \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \dot{\lambda}^* M S dt. \quad (13)$$

The gradient expression after grouping is then:

$$\begin{aligned} \frac{dG}{dp} &= \int_{t_0}^{t_1} \frac{\partial g}{\partial u} S + \frac{\partial g}{\partial p} + \dot{\lambda}^* M S + \lambda^* J S + \lambda^* \frac{\partial f}{\partial p} dt - \lambda^* M S \Big|_{t_0}^{t_1} \\ &= \int_{t_0}^{t_1} \left( \frac{\partial g}{\partial u} + \dot{\lambda}^* M + \lambda^* J \right) S + \lambda^* \frac{\partial f}{\partial p} + \frac{\partial g}{\partial p} dt - \lambda^* M S \Big|_{t_0}^{t_1} \end{aligned} \quad (14)$$

It becomes obvious that we can impose the condition

$$\frac{\partial g}{\partial u} + \dot{\lambda}^* M + \lambda^* J = \mathbf{0} \quad \text{and} \quad \lambda(t_1)^* M = \mathbf{0}^*, \quad (15)$$

to make eq. (14) independent of  $S$ .

After rearranging eq. (14) and eq. (15), we have

$$M^* \dot{\lambda} = -J^* \lambda - \frac{\partial g}{\partial u}^*, \quad (16)$$

$$M^* \lambda(t_1) = \mathbf{0} \quad (17)$$

$$\frac{dG}{dp} = \int_{t_0}^{t_1} \lambda^* \frac{\partial f}{\partial p} + \frac{\partial g}{\partial p} dt + \lambda^* M S \Big|_{t=t_0}. \quad (18)$$

Here, we want to remark that the  $\lambda^* M S \Big|_{t=t_0}$  term is zero if the initial condition of eq. (2) is independent of the parameters.

## 4 Adjoint Sensitivity Analysis with Discrete Cost Function

In many cases, we only want to compute sensitivity at specific a time point  $\tau$ , i.e.  $\frac{\partial g(u(p,t), p, t)}{\partial p} \Big|_{t=\tau}$ . In these cases, we call the cost function discrete.

To reuse our previous result, we need to find a way to relate  $\frac{dg}{dp}$  to  $\frac{dG}{dp}$ . There are two options, using Dirac delta distribution or using Leibniz integration rule. We pick the latter because it is more straight forward.

Note we have

$$\frac{d}{dp} \frac{d}{d\tau} G[u] = \frac{d}{dp} \frac{d}{d\tau} \int_{t_0}^{\tau} g(u(p, t), p, t) dt \quad (19)$$

$$= \frac{d}{dp} \left[ g(u, p, \tau) + \int_{t_0}^{\tau} \frac{\partial g}{\partial \tau} dt \right] \quad (20)$$

$$= \frac{dg}{dp} + \int_{t_0}^{\tau} \frac{\partial^2 g}{\partial \tau \partial p} dt, \quad (21)$$

and

$$\frac{d}{d\tau} \frac{d}{dp} G[u] = \frac{d}{d\tau} \left[ \int_{t_0}^{\tau} \lambda^* \frac{\partial f}{\partial p} + \frac{\partial g}{\partial p} dt + \lambda^* MS \Big|_{t=t_0} \right] \quad \text{eq. (18)} \quad (22)$$

$$= \left( \lambda^* \frac{\partial f}{\partial p} + \frac{\partial g}{\partial p} \right) \Big|_{t=\tau} + \int_{t_0}^{\tau} \lambda_{\tau}^* \frac{\partial f}{\partial p} + \underbrace{\lambda^* \frac{\partial^2 f}{\partial p \partial \tau}}_{=0} dt + \int_{t_0}^{\tau} \frac{\partial^2 g}{\partial \tau \partial p} dt + \lambda_{\tau}^* MS \Big|_{t=t_0}. \quad (23)$$

By comparing eq. (21) and eq. (23), we can conclude that

$$\frac{dg}{dp} = \left( \lambda^* \frac{\partial f}{\partial p} + \frac{\partial g}{\partial p} \right) \Big|_{t=\tau} + \int_{t_0}^{\tau} \lambda_{\tau}^* \frac{\partial f}{\partial p} dt + \lambda_{\tau}^* MS \Big|_{t=t_0}. \quad (24)$$

Now we need  $\lambda_{\tau}$  to compute the sensitivity integral eq. (24). It can be obtained by differentiating eq. (16),

$$M^* \dot{\lambda}_{\tau} = -J^* \lambda_{\tau} - \frac{\partial g}{\partial \tau} \quad (25)$$

$$= -J^* \lambda_{\tau} \quad g \text{ does not depend on } \tau \quad (26)$$

We want to remark that from eq. (26), we can see that the initialization of  $\lambda_{\tau}$  cannot be trivial, since  $\lambda_{\tau}(\tau) = \mathbf{0}$  can only result in uninteresting dynamics.

It is important to note that  $\lambda$  depends on not only  $t$  but also  $\tau$  with a discrete cost function, since  $\lambda_{\tau}$  is non-zero.

To obtain the initialization of  $\lambda_{\tau}|_{t=\tau}$ , we can differentiate  $\lambda(t, \tau)|_{t=\tau}$  in the constraint eq. (17) with respect to  $\tau$ ,

$$\frac{d}{d\tau} \left( \lambda(t, \tau)^* M \Big|_{t=\tau} \right) = \left( \underbrace{\frac{\partial \lambda}{\partial t} \frac{\partial t}{\partial \tau} \Big|_{t=\tau}}_{\dot{\lambda}} \right)^* M + \lambda_{\tau}^* M \Big|_{t=\tau} = \mathbf{0}^*. \quad (27)$$

Hence, we have

$$\lambda_{\tau}^* M \Big|_{t=\tau} = - \dot{\lambda}^* M \Big|_{t=\tau} \quad (28)$$

$$= \left( J^* \lambda + \frac{\partial g}{\partial u}^* \right)^* \Big|_{t=\tau} \quad \text{eq. (16)} \quad (29)$$

Together, the system of equations is,

$$M^* \dot{\lambda}_{\tau} = -J^* \lambda_{\tau} \quad (30)$$

$$M^* \lambda_{\tau} \Big|_{t=\tau} = \left( J^* \lambda + \frac{\partial g}{\partial u}^* \right) \Big|_{t=\tau} \quad (31)$$

$$\frac{dg}{dp} = \left( \lambda^* \frac{\partial f}{\partial p} + \frac{\partial g}{\partial p} \right) \Big|_{t=\tau} + \int_{t_0}^{\tau} \lambda_{\tau}^* \frac{\partial f}{\partial p} dt + \lambda_{\tau}^* MS \Big|_{t=t_0}. \quad (32)$$

#### 4.1 Example: Index-I differential-algebraic equation

To handle singular mass matrix, we need to split the problem system into differential variables  $\{\cdot\}^d$  and algebraic variables  $\{\cdot\}^a$ , so that  $\widetilde{M}$  is fully ranked,

$$\begin{pmatrix} \widetilde{M} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \dot{u}^d \\ \dot{u}^a \end{pmatrix} = \begin{pmatrix} f(u^d, u^a, p, t) \\ h(u^d, u^a, p, t) \end{pmatrix} \quad (33)$$

$$u^d(t_0) = u_{d0}(p). \quad (34)$$

We have

$$J^* = \begin{pmatrix} \frac{\partial f}{\partial u^d}^* & \frac{\partial h}{\partial u^d}^* \\ \frac{\partial f}{\partial u^a}^* & \frac{\partial h}{\partial u^a}^* \end{pmatrix}. \quad (35)$$

The initialization step is

$$\begin{pmatrix} \widetilde{M}^* & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_\tau^d \\ \lambda_\tau^a \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial u^d}^* & \frac{\partial h}{\partial u^d}^* \\ \frac{\partial f}{\partial u^a}^* & \frac{\partial h}{\partial u^a}^* \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \lambda^a \end{pmatrix} + \begin{pmatrix} \frac{\partial g^d}{\partial u}^* \\ \frac{\partial g^a}{\partial u}^* \end{pmatrix}. \quad (36)$$

Expanding the equation we have

$$\widetilde{M}^* \lambda_\tau^d = \frac{\partial h}{\partial u^d}^* \lambda^a + \frac{\partial g^d}{\partial u}^* \quad (37)$$

$$\mathbf{0} = \frac{\partial h}{\partial u^a}^* \lambda^a + \frac{\partial g^a}{\partial u}^* \implies \lambda^a = - \left( \frac{\partial h}{\partial u^a}^* \right)^{-1} \frac{\partial g^a}{\partial u}^* \quad (38)$$

Plugging eq. (38) into eq. (37), we obtain the initialization of  $\lambda_\tau$

$$\widetilde{M}^* \lambda_\tau^d(\tau) = \left( - \frac{\partial h}{\partial u^d}^* \left( \frac{\partial h}{\partial u^a}^* \right)^{-1} \frac{\partial g^a}{\partial u}^* + \frac{\partial g^d}{\partial u}^* \right) \Big|_{t=\tau}. \quad (39)$$

## References

- [1] Y. CAO, S. LI, L. PETZOLD, AND R. SERBAN, *Adjoint sensitivity analysis for differential-algebraic equations: The adjoint dae system and its numerical solution*, SIAM journal on scientific computing, 24 (2003), pp. 1076–1089.