

# SENSITIVITY ANALYSIS FOR DIFFERENTIAL ALGEBRAIC EQUATIONS

G. BORONI<sup>†</sup>, P. LOTITO<sup>†</sup> and A. CLAUSSE<sup>†‡</sup>

<sup>†</sup>CONICET and Universidad Nacional del Centro, Tandil, Argentina.

<sup>‡</sup>Comisión Nacional de Energía Atómica, Argentina.  
{gboroni, plotito, clausse}@exa.unicen.edu.ar

**Abstract**— In this paper we propose a new method based on adjoint systems for parametric sensitivity analysis of DAE's. This method is employed in a series of experiments, and the results are compared with the estimation of sensitivity by a finite differences method, widely used because of the simplicity of its implementation. In addition, we propose a method extension based on the use of estimation functions that allows transforming the original adjoint system into an algebraic system, making it suitable for many efficient numerical methods.

**Keywords**— Simulation, Parametric sensitivity, DAE system.

## I. INTRODUCTION

The mathematical models that are currently being used to investigate physical phenomena are becoming more and more realistic. The new characteristic of these models is that they often use parameters whose values may not be accurately known calling for parametric sensitivity analysis. Areas of application include optimization, parameter estimation, optimal control, model simplification, process sensitivity, uncertainty analysis, and experimental design for a wide range of scientific and engineering problems.

Recent work on methods and software for sensitivity analysis of differential algebraic system (DAE) (Cao *et al.*, 2003; Feehery *et al.*, 1997; Li and Petzold, 2000; Li and Petzold, 1999; Li *et al.*, 2000; and Maly and Petzold, 1997), has demonstrated that subsequent sensitivity values can be computed reliably and efficiently via automatic differentiation in combination with DAE solution techniques.

However, there are some difficult issues when trying to make sensitivity analysis of DAE systems with a large number of parameters compared with the number of variables. In these cases, the extension of the original system with the equations required for the sensitive analysis makes the resulting system grow significantly, and therefore, the calculation efficiency decreases.

The first contribution of this paper is the employment of a new method based on adjoint systems for parametric sensitivity analysis, whose theory was recently proposed for differential algebraic equations (Cao *et al.*, 2003). Following this method, it has been proved that it is possible to calculate the sensitivity of

parameters using linear adjoint systems, with no need to extend the original system, what allows improving the calculation efficiency considerably.

This method is used in a series of experiments, where the results are compared with the estimation of sensitivity by the finite differences method, widely used for the simplicity of its implementation.

The second contribution of this paper is to propose an extension of the method based on the use of estimation functions that allows transforming the original adjoint system into an algebraic system, making it suitable for the implementation of many efficient numerical methods.

## II. SENSITIVITY CALCULATION FOR DIFFERENTIAL ALGEBRAIC SYSTEMS BY ADJOINT METHOD

The numerical solution of initial value problems modeled by DAE has attracted the interest of the numerical community for the last 30 years (Brenan *et al.*, 1996 and Kees and Millen, 1999). Many engineering and scientific problems can be naturally modeled with DAE's, mostly for balance equations of the general form:

$$\begin{aligned} F(x, \dot{x}, p, t) &= 0, \\ x(t_0) &= x_0(p), \end{aligned} \quad (1)$$

composed by a combination of differential and algebraic equations, where  $x \in R^{N_x}$  is the state variable vector, and  $p \in R^{N_p}$  is a parameter vector.

The problem of calculating the sensitivity of parameters of a DAE system takes the following form: for a parameter dependent DAE system of type (1), find

$$\frac{dx}{dp_j} \text{ at time } T, \text{ for } j: 1, \dots, N_p.$$

Their solution requires the simultaneous solution of the original DAE system with the  $N_p$  sensitivity systems obtained by differentiating the original DAE with respect to each parameter in turn (Cao *et al.*, 2003).

For large systems this may look like a lot of work, but it can be done efficiently, if  $N_p$  is relatively small, by exploiting the fact that the sensitivity systems are linear and all share the same Jacobian matrices with the original system (Li *et al.*, 2000).

However, some problems require calculating the sensitivities with respect to a large number of parameters. For these problems, particularly if the

number of state variables  $N_x$  is also large, the previous sensitivity approach is out of reach.

In Cao *et al.* (2003) it was demonstrated that these problems can often be handled more efficiently by the adjoint method (Errico 1997). In this approach, we are interested in calculating the sensitivity  $\frac{dG}{dp}$  of an objective function of the perturbed solution and the parameter:

$$G(x, p) = \int_0^T g(x, t, p) dt, \quad (2)$$

or alternatively the sensitivity  $\frac{dg}{dp}$  of a function  $g(x, t, p)$  defined only at time  $T$ . For example, in the mathematical models of nuclear reactors, a common analysis is calculating the sensitivity of speed  $u_i$  of the cooling liquid, product of the perturbation of the demanded thermal power  $Q$ . Thus, the following objective function can be formulated:

$$G(u_i, Q) = \int_0^T (u_i(t, Q) - \bar{u}_i(t, Q))^2 dt, \quad (3)$$

where  $\bar{u}$  is the average of the speed of the liquid. Next, we describe how to employ the adjoint method to compute the sensitivity.

**A. Sensitivity of  $G(x, p)$**

To find an expression for  $\frac{dG}{dp}$ , we define the augmented objective function  $I(x, p)$  as:

$$I(x, p) = G(x, p) - \int_0^T \lambda(t) F(x, \dot{x}, p, t) dt, \quad (4)$$

where  $\lambda$  is a Lagrange multiplier, and  $F(x, \dot{x}, p, t) = 0$  by (1). Formally differentiating both members with respect to  $p$ , the sensitivity of  $G(x, p)$  with respect to  $p$  is:

$$\frac{dI}{dp} = \frac{dG}{dp} = \int_0^T (g_p + g_x x_p) dt - \int_0^T \lambda(t) (F_p + F_x x_p + F_{\dot{x}} \dot{x}_p) dt. \quad (5)$$

Then, using integration by parts we obtain an equivalent expression for the term  $\lambda(t) F_x \dot{x}_p$  depending on  $x_p$  instead:

$$\int_0^T \lambda(t) F_x \dot{x}_p dt = \lambda(t) F_x x_p \Big|_0^T - \int_0^T (\lambda(t) F_x)' x_p dt. \quad (6)$$

If the Eq. (6) is replaced in (5) and the terms regrouped, we obtain that:

$$\begin{aligned} \frac{dG}{dp} = & \int_0^T (g_p - \lambda(t) F_p) dt \\ & - \int_0^T \left( -g_x + \lambda(t) F_x - (\lambda(t) F_x)' \right) x_p dt \\ & - \lambda(t) F_x x_p \Big|_0^T, \end{aligned} \quad (7)$$

and if we require that the following (adjoint) equation be satisfied

$$\begin{aligned} (\lambda(t) F_x)' - \lambda(t) F_x &= -g_x, \\ \lambda(t) F_x \Big|_{t=T} &= 0, \end{aligned} \quad (8)$$

Eq. (7) becomes

$$\frac{dG}{dp} = \int_0^T (g_p - \lambda(t) F_p) dt + \lambda(t) F_x x_p \Big|_{t=0}, \quad (9)$$

Note that (8) is the adjoint equation, and  $\lambda$  is the adjoint variable of the system. In this work we used the notation of augmented adjoint system corresponding to (8):

$$\begin{cases} \dot{\alpha}(t) - \lambda(t) F_x = -g_x, \\ \alpha(t) - \lambda(t) F_x = 0, \\ \alpha(T) = 0, \end{cases} \quad (10)$$

where we have introduced the function  $\alpha(t)$ .

From (10) onwards, it is observed that the adjoint DAE must be solved backward in time. To calculate the integral term  $\int_0^T (g_p - \lambda(t) F_p) dt$  of the Eq. (9), we use a quadrature variable  $\beta$ , expanding again the adjoint system:

$$\begin{cases} \dot{\alpha}(t) - \lambda(t) F_x = -g_x, \\ \alpha(t) - \lambda(t) F_x = 0, \\ \dot{\beta} + (g_p - \lambda(t) F_p) = 0, \\ \alpha(T) = 0, \\ \beta(T) = 0. \end{cases} \quad (11)$$

In this way, for  $t=0$ ,

$$\beta(0) = - \int_0^T (g_p - \lambda(t) F_p) dt, \quad (12)$$

yielding the sensitivity equations for  $\frac{dG}{dp}$ :

$$\frac{dG}{dp} = \beta(0) + \lambda(t) F_x x_p \Big|_{t=0}. \quad (13)$$

**B. Sensitivity of  $g(x, t, p)$**

Now let us consider the computation of  $\frac{dg}{dp}$ .

Interchanging differentiations we have that  $\frac{dg}{dp} = \frac{d}{dp} \frac{dG}{dT} = \frac{d}{dT} \frac{dG}{dp}$  and from (9), we obtain:

$$\begin{aligned} \frac{dg}{dp} = & (g_p - \lambda(t, T) F_p) \Big|_{t=T} \\ & - \int_0^T (\lambda_t(t, T) F_p) dt + \lambda_t(t, T) F_x x_p \Big|_{t=0}, \end{aligned} \quad (14)$$

where the adjoint variable  $\lambda(t, T)$  depends on both  $t$  and  $T$ , and  $\lambda_t(t, T)$  denotes  $\frac{d\lambda(t, T)}{dT}$ . The corresponding adjoint equation is:

$$\begin{cases} \dot{\alpha}(t) - \lambda_T(t, T)F_x = 0, \\ \alpha(t) - \lambda_T(t, T)F_x = 0, \\ \alpha(T) = (g_x - \lambda F_x)|_{t=T}, \end{cases} \quad (15)$$

and it is solved in a similar way to system (10).

Again, to calculate the integral term of Eq. (14), we used a quadrature variable  $\beta$ , and we defined the following adjoint system:

$$\begin{cases} \dot{\alpha}(t) - \lambda_T(t, T)F_x = 0, \\ \alpha(t) - \lambda_T(t, T)F_x = 0, \\ \alpha(T) = (g_x - \lambda F_x)|_{t=T}, \\ \dot{\beta} - \lambda_T(t, T)F_p = 0. \end{cases} \quad (16)$$

Finally, the expression of the sensitivity equation  $\frac{dg}{dp}$  is given by

$$\frac{dg}{dp} = (g_p - \lambda(t, T)F_p)|_{t=T} - \beta(0) + \lambda_T(t, T)F_{x,p}|_{t=0}. \quad (17)$$

### III. NUMERICAL EXPERIMENTS

Nowadays, there is a variety of numerical methods for solving IVP's given by DAE's, which can be classified into three classes: single-step, multistep, and extrapolation methods. The efficiency of each type of approximation depends on the specific problem characteristics.

In this work the systems (11) and (16) are solved transforming the DAE's into purely algebraic systems by means of numerical approximations computing through estimating functions as in Boroni and Clausse (2005). In order to do that, let us define the functions  $\Lambda(\lambda_n, \lambda_{n-1}, \lambda_{n-2}, \dots)$  and  $D\Lambda(\lambda_n, \lambda_{n-1}, \lambda_{n-2}, \dots)$ , that represent multistep estimators of the state vector and its derivatives, *i.e.*,  $\lambda$  and  $\dot{\lambda}$ . Thus:

$$\begin{aligned} \lambda(t) &\rightarrow \sum_{i=0}^k a_i \lambda(t - i\Delta t) \\ \dot{\lambda}(t) &\rightarrow \sum_{i=0}^k b_i \lambda(t - i\Delta t) \end{aligned} \quad (18)$$

Specifically, in this work we use the adjoint variable and implicit linear multistep estimator (Jackson and Sacks-Davis, 1980) for the derivatives:

$$\lambda = \lambda(t), \quad \dot{\lambda} = \sum_{i=0}^k \frac{\alpha_i \lambda(t - i\Delta t)}{\Delta t}. \quad (19)$$

To analyze the application of the adjoint method in the calculation of sensitivity, a series of experiments was made, where the numerical results and errors were examined. In this study two systems were defined: a linear differential system with explicit solution, and a nonlinear conservative oscillating system. So as to simplify the presentation of the results, only the calculation of the sensitivity of the objective function  $G(x, p)$  is considered.

#### A. Linear System

An example of a linear differential equation of the first order is set out, given by:

$$\begin{cases} F(\dot{x}, x, p, t) = \dot{x} + px = 0 \\ x(t_0) = x_0(p) = p \end{cases} \quad (20)$$

This system has only one parameter  $p$ , and the explicit solution is:

$$x(t, p) = pe^{-tp}. \quad (21)$$

The purpose of working with this system was to study the error between the explicit solution and the numerical solution of the sensitivity computed as in (13), for the following objective function

$$G(x, p) = \int_0^T x(t, p) dt. \quad (22)$$

Using (13) and (22), we obtain that:

$$\frac{dG}{dp} = \frac{d(1 - e^{-Tp})}{dp} = Te^{-Tp}. \quad (23)$$

The temporary evolution of the sensitivity  $\frac{dG}{dp}$  can be seen in Fig. 1, where it has the form of a left-biased distribution.

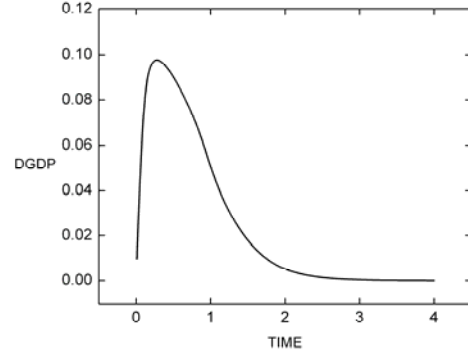


Fig. 1. Temporal evolution of the sensitivity  $\frac{dG}{dp}$ .

In Fig. 2 we can see the absolute error  $\varepsilon$  between the exact value and the computed value of the sensitivity  $dG/dp$  with respect to the explicit solution. In this figure it can be seen that the calculation of the  $dG/dp$  generates a error with a maximum close to  $1.0 \times 10^{-4}$ , which stabilizes for a time  $t \geq 3$ .

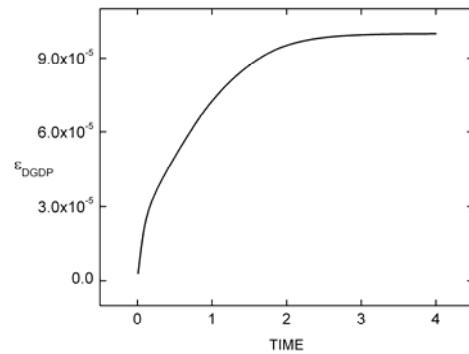


Fig. 2. Calculation of the absolute error for  $\frac{dG}{dp}$ .

**B. Elastic pendulum**

The second example is the elastic pendulum (Fig. 3), consisting of a no-linear differential equation system of the fourth order whose natural variables are the string length, the inclination angle with respect to the vertical, and their respective temporal derivatives, that is:

$$\begin{aligned} \dot{r} &= z \\ \dot{\theta} &= w \\ \dot{z} - rw^2 &= -\frac{k}{m}(r-L) + g \cos(\theta) \\ r\dot{w} + 2zw &= -g \sin(\theta) \end{aligned} \quad (24)$$

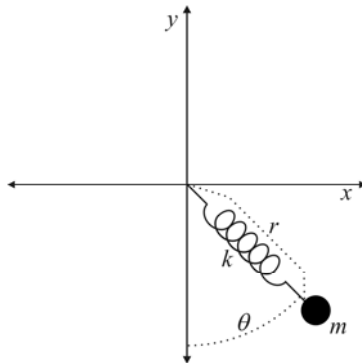


Fig. 3. Elastic Pendulum.

Since in this case, no explicit solution is available, relative comparisons were made between the explicit and implicit numerical solutions for the following parameters and initial conditions:

$$k = 7, L = 1, m = 0.1, g = 9.8, r(0) = 1, \theta(0) = \pi/2, \dot{r}(0) = 0 \text{ y } \dot{\theta}(0) = 0.$$

Figure 4 shows the trajectory of the mass in the (x, y) plane.

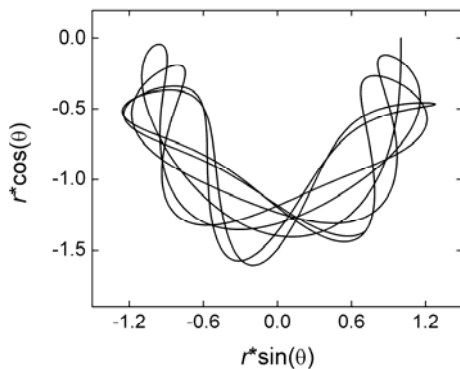


Fig. 4. Trajectory of the mass in the (x, y) plane.

For the calculation of the sensitivity we defined an objective function related with the energy of the pendulum:

$$E = \frac{1}{2}mv^2 + \frac{1}{2}k(r-L)^2 + mgr(1-\cos(\theta)), \quad (25)$$

where the term  $\frac{1}{2}k(r-L)^2$  is associated to the harmonic potential energy,  $mgr(1-\cos(\theta))$  represents the

gravitational potential energy and  $\frac{1}{2}mv^2$  is associated to the kinetic energy. Initially, when the pendulum is standing, the total energy is:

$$E_0 = \frac{1}{2}k(r-L)^2. \quad (26)$$

The purpose of this example is to determine an error's measure, considering that the objective function has an explicit solution given by the principle of conservation of the energy. Therefore, the following objective function is defined:

$$\begin{aligned} G(r, \theta, \dot{r}, \dot{\theta}, k, m) &= \int_0^T E(r, \theta, \dot{r}, \dot{\theta}, k, m, t) dt \\ &= \int_0^T E_0(r, \theta, \dot{r}, \dot{\theta}, k, m, t) dt \end{aligned} \quad (27)$$

Developing (27) for the pendulum in movement, we obtain that:

$$G(r, \theta, \dot{r}, \dot{\theta}, k, m) = \int_0^T \left[ \frac{1}{2}mv^2 + \frac{1}{2}k(r-L)^2 + mgr(1-\cos(\theta)) \right] dt, \quad (28)$$

where  $v^2 = \dot{r}^2 + r^2\dot{\theta}^2$ . When the pendulum is standing,

the sensitivity  $\frac{dG}{dk}$  is:

$$\frac{dG}{dk} = \frac{d(k(r_0-L)^2 T / 2)}{dk} = 0, \quad (29)$$

obtaining as a result a constant value equal to 0 ( $r_0 = L$ ). With respect to parameter  $m$ , the sensitivity

$\frac{dG}{dm}$  is expressed by:

$$\frac{dG}{dm} = \frac{d(mgr_0(1-\cos(\theta_0))T)}{dm} = gr_0T. \quad (30)$$

For (30) the results indicate that sensitivity grows linearly with time. Using the sensitivity expressions (29) and (30), we made a comparative analysis of the numerical results obtained employing the adjoint method and the finite differences method. In the finite difference method, the system DAE is solved numerically with different values of parameter: a value of perturbed parameter and the rest of the parameters with original values. Thus, it is possible to obtain for

example an estimation of the sensitivity  $\frac{dG}{dk}$  equal to:

$$\frac{dG}{dk} \approx \frac{\int_0^T E(r, \theta, \dot{r}, \dot{\theta}, k + \delta_k, m, t) - \int_0^T E(r, \theta, \dot{r}, \dot{\theta}, k, m, t)}{\delta_k}. \quad (32)$$

The advantage of this method is the simplicity of its implementation, whereas the main disadvantage is that the estimation error is significant.

Figure 5 shows the error defined by the absolute difference  $\epsilon_{DGDK}$  between the explicit solution and numerical solution. In this figure it can be seen that when adopting the adjoint method the error is close to 1,

whereas when adopting the finite differences method, the error is tripled with respect to the adjoint method.

Figure 6 shows the error defined by the absolute difference  $\varepsilon_{DGDM}$ , where the absolute error calculated by the adjoint method is very small if we compare it with the error obtained by finite differences. It can also be observed that for values of time  $T$  between 7.7 and 7.9, the sensitivity by finite differences reaches values between 125 and 130, with an  $\varepsilon_{DGDM}$  error between 55 and 60, what practically represents a 50%. This is a clear indicator that the adoption of the finite difference method can produce wrong results.

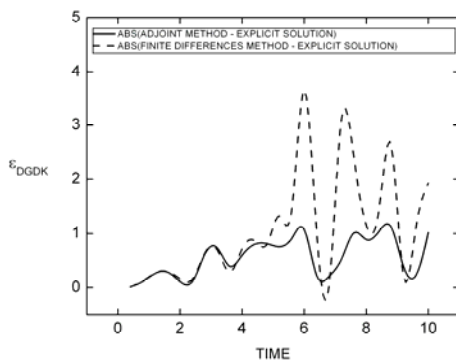


Fig. 5. Absolute error of  $\frac{dG}{dk}$ .

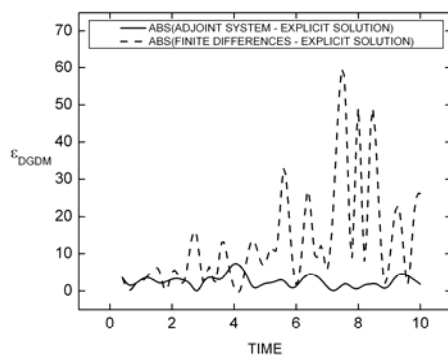


Fig. 6. Absolute error of  $\frac{dG}{dm}$ .

#### IV. CONCLUSIONS

It can be seen in the previous examples that the calculation of sensitivity is very often obtained with quite small precision errors. For the linear system (III.A), the proposed method allows obtaining a reasonably accurate solution. When the DAE system is nonlinear (III.B), the calculation of parametric sensitivity by the adjoint method produces a quite acceptable solution.

Finally, the most important advantage of the proposed method is the possibility of calculating the sensitivity of parameters of DAE's using linear adjoint systems, with no need to extend the original system, thus improving the efficiency in the calculation.

#### REFERENCES

- Boroni, G. and A. Clause, "Una formulaci3n matem3tica orientada a objetos para simulaci3n continua," *Mec3nica Computacional*, **24**, 261 (2005).
- Brenan, K., S. Campbell and L. Petzold, *Numerical Solution of Initial Value Problems in Differential-Algebraic Equations*, SIAM, Philadelphia, PA (1996).
- Cao, Y., S. Li, L. Petzold and R. Serban, "Adjoint sensitivity analysis for differential-algebraic equations: The adjoint DAE system and its numerical solution," *SIAM J. Sci. Comput.*, **24**, 1076–1089 (2003).
- Errico, R.M., "What is an adjoint model?," *Bulletin of the American Meteorological Society*, **78**, 2577–2591 (1997).
- Feehery, W.F., J.E. Tolsma and P.I. Barton, "Efficient sensitivity analysis of large-scale differential-algebraic systems," *Appl. Numer. Math.*, **25**, 41–54 (1997).
- Jackson, R. and R. Sacks-Davis, "An Alternative Implementation of Variable Step-Size Multistep Formulas for Stiff ODEs," *ACM Trans. on Mathematical Software*, **6**, 295–318 (1980).
- Kees, C. and C. Miller, "C11 Implementations of Numerical Methods for Solving Differential-Algebraic Equations: Design and Optimization Considerations," *ACM Trans. on Mathematical Software*, **25**, 377–403 (1999).
- Li, S. and L.R. Petzold, "Design of New DASPK for Sensitivity Analysis," *Technical report*, Dept. of Computer Science, UCSB, Santa Barbara (1999).
- Li, S. and L.R. Petzold, "Software and algorithms for sensitivity analysis of large-scale differential-algebraic systems," *J. Comput. Appl. Math.*, **125**, 131–145 (2000).
- Li, S., L.R. Petzold and W. Zhu, "Sensitivity analysis of differential-algebraic equations: A comparison of methods on a special problem," *Appl. Numer. Math.*, **32**, 161–174 (2000).
- Maly, T. and L.R. Petzold, "Numerical methods and software for sensitivity analysis of differential-algebraic systems," *Appl. Numer. Math.*, **20**, 57–79 (1997).

Received: April 12, 2007.

Accepted: August 14, 2007.

Recommended by Subject Editor: Jorge Solsona