An optimal control technique for solving differential equations

L. Bayón, J.M. Grau, M.M. Ruiz and P.M. Suárez

University of Oviedo, Department of Mathematics, E.U.I.T.I. Campus of Viesques, Gijón, 33203, Spain.

Abstract. In this paper we present a new method for solving systems of ordinary nonlinear differential equations with initial conditions. The method is based on transforming the problem to an optimal control problem. We then solve it with a technique based on the use of an integral form of the Euler equation combined with the shooting method and the cyclic coordinate descent method.

Keywords: Ordinary Differential Equations, Numerical methods, Optimal Control, Shooting method, Cyclic Coordinate Descent
PACS: 02.30.Hq

INTRODUCTION

For hundreds of years, ordinary differential equations (ODEs) have been used to model continuous systems in all scientific and engineering disciplines. Many mathematicians have studied the nature of these equations and many well-developed solution techniques exist. The mathematical formulation most frequently used is that of an initial value problem (IVP) for a first-order system of ordinary nonlinear differential equations:

$$\left\{ \begin{array}{l} x_i'(t) = f_i(t, x_1(t), \ldots, x_n(t)) \\ x_i(a) = y_i \end{array} \right. \tag{1}$$

with $i = 1, \ldots n$. The numerical solution of ODEs is a well-studied problem in numerical analysis and there are many books on the subject [1]. The numerical methods most frequently employed fall into the following categories: Taylor methods, Runge-Kutta methods, Multistep methods, Extrapolation methods and Adaptive techniques. In this paper we present a new method for solving (1) based on transforming the IVP to an optimal control problem (OCP).

This basic idea has also been developed in [2], though these authors use iterative dynamic programming (IDP) to solve the OCP and obtain a piecewise-constant optimal control function. We, however, propose a new methodology for solving the OCP, which has already been proven successful within the framework of Hydrothermal Optimization [3]. Our method uses a variety of mathematical techniques, well-known for the case of functions, though now adapted to the case of functionals, which are efficiently combined to afford a novel contribution. The technique is based on the use of an integral form of the Euler equation, combined with an adapted version of the shooting method and the cyclic coordinate descent method. We shall compare the error between our approach, classical methods and [2].

STATEMENT

Let us define the function $F : [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$:

$$F(t, x_1(t), \ldots, x_n(t), x_1'(t), \ldots, x_n'(t)) = \sum_{i=1}^{n} \left[x_i'(t) - f_i(t, x_1(t), \ldots, x_n(t)) \right]^2 \tag{2}$$

Next, we define the following minimization problem:

$$\begin{array}{ll} \text{Minimize:} & E(x_1, \ldots, x_n, x_1', \ldots, x_n') = \displaystyle\int_a^b F(t, x_1(t), \ldots, x_n(t), x_1'(t), \ldots, x_n'(t)) dt \\ \text{subject to:} & x_1(a) = y_1; x_2(a) = y_2; \ldots; x_n(a) = y_n \end{array} \tag{3}$$

where $E(x_1, \ldots, x_n, x_1', \ldots, x_n')$ is called the *error functional*. If the optimal solution of (3) is zero, since the function F is continuous and non-negative, then $F \equiv 0$. Thus, the first-order system (2) will hold for all t, and the solution of (1) is obtained. We now formulate problem (3) as an OCP. We denote $\mathbf{x}(t) = (x_1(t), ..., x_n(t))$, $\mathbf{y} = (y_1, ..., y_n)$, and $\mathbf{u}(t) = (u_1(t), ..., u_n(t))$. We consider the state variables to be $\mathbf{x}(t)$ and the control variables to be $\mathbf{x}'(t)$. The OCP is thus:

$$\begin{array}{ll} \min\limits_{\mathbf{u}(t)} & E(\mathbf{x},\mathbf{u}) = \int_a^b F(t,\mathbf{x}(t),\mathbf{u}(t))dt \\ \text{s. t.} & \mathbf{x}'(t) = \mathbf{u}(t) \\ & \mathbf{x}(a) = \mathbf{y} \end{array} \tag{4}$$

OPTIMAL CONTROL ALGORITHM

In this section the standard Lagrange type problem (4) is formulated within the framework of Optimal Control [4]. The classical approach involves using Pontryagin's Minimum Principle (PMP), which results in a two-point boundary value problem (TPBVP).

Let H be the Hamiltonian function associated with the problem

$$H(t,\mathbf{x},\mathbf{u},\lambda) = F(t,\mathbf{x},\mathbf{u}) + \lambda \cdot \mathbf{u} \tag{5}$$

where $\lambda = (\lambda_1(t), ..., \lambda_n(t))$ is the costate vector. In order for $\mathbf{u}$ to be optimal, a nontrivial function λ must necessarily exist, such that for almost every $t \in [a,b]$, $i = 1, ...n$

$$x_i' = H_{\lambda_i}(t,\mathbf{x},\mathbf{u},\lambda_1, ..., \lambda_{i-1}, \cdot, \lambda_{i+1}, ..., \lambda_n) \tag{6}$$

$$-\lambda_i' = H_{x_i}(t, x_1, ..., x_{i-1}, \cdot, x_{i+1}, ..., x_n, \mathbf{u}, \lambda) \tag{7}$$

$$H(t,\mathbf{x},\mathbf{u},\lambda) = \min_{\mathbf{v}(t)} H(t,\mathbf{x},\mathbf{v},\lambda) \tag{8}$$

$$\mathbf{x}(a) = \mathbf{y};\ \lambda(b) = \mathbf{0} \tag{9}$$

From (7), there exists a piecewise C^1 function λ_i that satisfies:

$$\lambda_i' = -H_{x_i} = -F_{x_i} \tag{10}$$

From (10), it follows that

$$\lambda_i(t) = K_i - \int_a^t F_{x_i}(s,\mathbf{x},\mathbf{u})ds \tag{11}$$

with $K_i = \lambda_i(a)$. From (8), it follows that $u_i(t)$ minimizes $H(t,\mathbf{x},u_1, ..., u_{i-1}, \cdot, u_{i+1}, ..., u_n, \lambda)$, for each t. Hence we have

$$F_{u_i} + \lambda_i(t) = 0 \tag{12}$$

From (11) and (12), we have

$$K_i = -F_{x_i'} + \int_a^t F_{x_i}(s,\mathbf{x},\mathbf{u})ds \tag{13}$$

Thus, for each i (assuming the rest of the variables are fixed), the problem consists in finding the K_i and the function $x_i(t)$ that satisfy (13) and (9). From the computational point of view, the construction can be performed using the same procedure as the simple shooting method [5], employing a discretized version of Equation (13). Therefore, the method which we have developed to obtain the solution is based on the use of an integral form of the Euler equation, combined with the simple shooting method.

To solve the variational problem (3) (with $i = 1, ...n$), we propose an algorithm of its numerical resolution using a particular strategy related to the cyclic coordinate descent (CCD) method [6]. The classic CCD method minimizes a function of n variables cyclically with respect to the coordinate variables. With our method, the problem could be solved like a sequence of problems whose error functional converges to zero.

NUMERICAL EXAMPLE

The method presented in this paper allows us to solve a wide range of nth order ordinary nonlinear differential equations with initial conditions. Let us consider the following IVP:

$$y'' + t^2 y = 0; \;\; y(0) = 0; y'(0) = 0.1 \tag{14}$$

an example likewise presented in [2]. Let

$$x_1(t) = y(t); \;\; x_2(t) = y'(t) \tag{15}$$

Then (14) transforms into a first-order system:

$$\begin{cases} x_1'(t) = x_2(t) \\ x_2'(t) = -t^2 x_1(t) \\ x_1(0) = 0; x_2(0) = 0.01 \end{cases} \tag{16}$$

Applying the above development, problem (14) changes to the following form:

$$\begin{array}{ll} \min\limits_{u_1(t),u_2(t)} & E(x_1,x_2,u_1,u_2) = \int_0^b \left[(u_1(t) - x_2(t))^2 + (u_2(t) + t^2 x_1(t))^2 \right] dt \\ \text{s. t.} & x_1'(t) = u_1; \;\; x_2'(t) = u_2 \\ & x_1(0) = 0; \;\; x_2(0) = 0.01 \end{array} \tag{17}$$

We consider $b = 1$, and a discretization of 1000 subintervals. In 12 iterations, our method gives the approximate optimal solution presented in Figure 1.

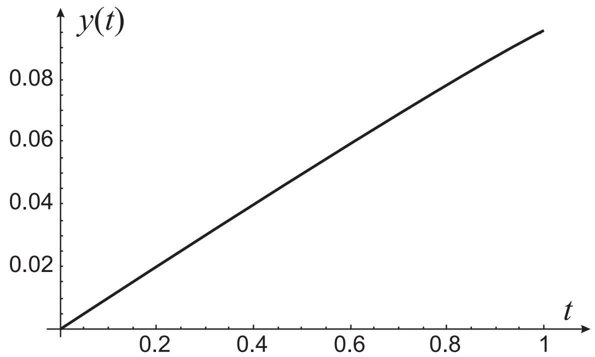

Fig. 1. Optimal solution.

We now compare our solution $y(t)$ with that obtained using IDP in [2] and with that obtained using the NDSolve instruction in the commercial software package, Mathematica. Said software incorporates a variety of classical methods, which we have employed using their default parameters. Instead of the classical global error, we consider the error functional E.

Table 1. Value of the error functional.

	E
Our solution	$3.1\ 10^{-17}$
IDP	$1.2\ 10^{-3}$
Euler	$4.9\ 10^{-3}$
Midpoint	$6.4\ 10^{-5}$
Runge-Kutta	$8.1\ 10^{-14}$
Predictor-corrector Adams	$1.5\ 10^{-11}$

It can be seen from Table 1 that the error functional obtained by our method is lower than that obtained by classical methods. This result follows from the fact that the aim of the classical methods is to minimize the global error. Nonetheless, the global error obtained by means of our method is reasonably satisfactory (in the order of $2.4\ 10^{-5}$), having obtained $y(1) = 0.0950929$ versus $y(1) = 0.095069$ using Runge-Kutta, not to mention, $y(1) = 0.0881$ obtained using IDP.

CONCLUSIONS

In this paper we have presented a new optimal control technique for solving ordinary differential equations. Our method substantially improves a previous approach that uses iterative dynamic programming to solve the associated optimal control problem. We consider the error functional instead of the classical global error, and the error functional obtained by our method is lower than that obtained by classical methods. The global error, which is the one normally considered, is not always really important, especially in problems of variational origin. Finally, it is worth noting that our method may be applicable to initial value problems of a very general nature, as well as to boundary value problems.

REFERENCES

1. R. Burden and J. Faires, *Numerical Analysis*, Thomson Brooks/Cole, Belmont CA, 2005.
2. S. Effati, H. Roohparvar, Iterative dynamic programming for solving linear and nonlinear differential equations. *Appl. Math. Comput.* 175, 247–257, 2006.
3. L. Bayon, J.M. Grau, M.M Ruiz and P.M. Suarez, A Bolza Problem in Hydrothermal Optimization, *Appl. Math. Comput.* 184(1), 12-22, 2007.
4. R. Vinter, *Optimal control*, Birkhäuser, Boston, 2000.
5. U.M. Ascher, R. Mattheij, and R.D. Russell, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, Classics in Applied Mathematics, 13, SIAM, Philadelphia, 1995.
6. Z.Q. Luo and P. Tseng, On the convergence of the coordinate descent method for convex differentiable minimization, *J. Optim. Theory Appl.* 72(1), 7-35, 1992.