A non-classical class of variational problems

Pedro A.F. Cruz and Delfim F.M. Torres*

Department of Mathematics,
University of Aveiro, 3810-193 Aveiro, Portugal
E-mail: pedrocruz@ua.pt   E-mail: delfim@ua.pt
*Corresponding author

Alan S.I. Zinober

Department of Applied Mathematics,
The University of Sheffield, Sheffield, S10 2TN, UK
E-mail: a.zinober@sheffield.ac.uk

Abstract: We study a new non-classical class of variational problems that is motivated by some recent research on the non-linear revenue problem in the field of economics. This class of problem can be set up as a maximising problem in the calculus of variations (CoV) or optimal control. However, the state value at the final fixed time, $y(T)$, is *a priori* unknown and the integrand is a function of the unknown $y(T)$. This is a non-standard CoV problem. In this paper we apply the new costate boundary conditions $p(T)$ in the formulation of the CoV problem. We solve a sample example in this problem class using the numerical shooting method to solve the resulting TPBVP, and incorporate the free $y(T)$ as an additional unknown. Essentially the same results are obtained using symbolic algebra software.

Keywords: non-linear revenue problems; transversality conditions; shooting method; calculus of variations; optimal control.

Biographical notes: Pedro A.F. Cruz is an Assistant Professor in the Department of Mathematics at the University of Aveiro. He obtained his MSc in Computational Science from Instituto Superior Técnico and PhD in Mathematics from the University of Aveiro. He has published journal and conference works in statistics, optimisation, and control theory. His main research interests are in the field of computational optimisation.

Delfim F.M. Torres is an Associate Professor of Mathematics at the University of Aveiro, Scientific Coordinator of the Control Theory Group (cotg), Editor-in-Chief of *IJMS* and *IJAMAS*. He received his Licenciatura from the University of Coimbra, and MSc and PhD from the University of Aveiro. He has written more than 150 scientific and pedagogical publications, and held positions of Invited and Visiting Professor in several countries in Europe, Africa, Caucasus and the USA. His research interests include various topics in the areas of the calculus of variations and optimal control.

Alan Zinober is a Professor of Non-linear Control Theory in the Department of Applied Mathematics at the University of Sheffield. After receiving his BSc and MSc at the University of Cape Town, he received his PhD from the University of Cambridge. He has been the recipient of a number of engineering and physical sciences research council and other research grants. He has published many journal and conference publications and has edited three research monographs. The central theme of his research is in the field of sliding mode control and other areas of non-linear control theory.

1   Introduction

The first result of the calculus of variations (CoV) ever discovered must have been the statement that the shortest path joining two points is a straight line segment. Another classical variational problem consists in finding, amongst all simple closed plane curves of a given fixed length, one that encloses the largest possible area. It is well known since ancient times that the circle is the shape that encloses maximum area for a given length of perimeter. However, it was not until the 18th century that a systematic theory, the CoV, began to emerge. A modern face to the CoV is given by the theory of optimal control. Economics is a source of interesting applications of the theory of CoV and optimal control. Classical examples include the optimal capital spending problem, optimal reservoir control, optimal production subject to royalty payment obligations, optimal maintenance and replacement policy, and optimal drug bust strategy (Léonard and Long, 1992; Sethi and Thompson, 2000).

The following economics problem (explained briefly here) has motivated this paper (Zinober and Kaivanto, 2008). A standard feature of the theory of the firm is that a profit maximising firm facing a downward sloping demand curve reacts to an increase in marginal cost by reducing output and increasing price. In this context, it is well understood that a requirement to pay a flat-rate royalty on sales has just this effect of increasing marginal cost and thereby decreasing output while simultaneously increasing price. However, the effect of permitting the royalty to take on more general forms leads naturally to non-standard CoV problems, and explains why this question has remained unaddressed to date (Zinober and Kaivanto, 2008). Recently the effect of piecewise linear cumulative royalty schedules on the optimal intertemporal production policy, i.e., an optimal economics control problem that does not fit into the classical class of variational problems, has been formulated (Zinober and Kaivanto, 2008). The economics problem lies in the area of repayable launch investment (RLI). For the purposes of this paper we will outline just the mathematical nature of the problem since the precise (non-linear) economic details are of secondary importance here. Consider the system in the time domain modelled by the differential equation

$$y'(t) = u(t), \qquad y(0) \text{ known}$$

with the endpoint state value $y(T)$ at time $t = T$ unknown. We wish to determine the control function $u(t)$ for $t \in [0, T]$ that maximises the return

$$J[u(\cdot)] = \int_0^T f(t, y(t), u(t), y(T)) dt.$$

Note that the integrand depends upon the *a priori* unknown final value $y(T)$. This class of problem is not contained within the classical class of variational problems and the present paper indicates how such problems may be solved.

The manuscript is organised as follows. In Section 2 we develop the necessary conditions for the extremising solution. The main idea is borrowed from Malinowska and Torres (2008), where appropriate natural boundary conditions are proved for problems of the CoV on time scales (Ferreira and Torres, 2008). We then consider two approaches to obtain the solution of a sample example with a continuous integrand f (Section 3). The first approach considers the numerical shooting solution (Section 4). The results obtained are then validated by symbolic algebra computations (Section 5). We finish with conclusions in Section 6.

2 The non-classical variational problem

We begin by developing the necessary conditions for the extremising solution. Let J be a functional of the form

$$J[y(\cdot)] = \int_a^T f(t, y(t), y'(t), y(T)) dt$$

where

$$(t, y, y', z) \rightarrow f(t, y, y', z)$$

is a smooth function and $T > a$. We consider the problem of determining the functions $y(\cdot) \in C^1$ such that $J[\cdot]$ has an extremum. An initial condition $y(a) = \alpha$ is imposed on $y(\cdot)$, but $y(T)$ is free.

Suppose that $J[\cdot]$ has an extremum at $\tilde{y}(\cdot)$. We can proceed as Lagrange did [cf. Gelfand and Fomin (1963)], by considering the value of J at a nearby function $y = \tilde{y} + \varepsilon h$, where ε is a small parameter, $h(\cdot) \in C^1$ and $h(a) = 0$. Because $y(T)$ is free, we do not require $h(\cdot)$ to vanish at T. Let

$$\phi(\varepsilon) = J[(\tilde{y} + \varepsilon h)(\cdot)] = \int_a^T f(t, \tilde{y}(t) + \varepsilon h(t), \tilde{y}'(t) + \varepsilon h'(t), \tilde{y}(T) + \varepsilon h(T)) dt.$$

A necessary condition for $\tilde{y}(\cdot)$ to be an extremiser is given by

$$\phi'(\varepsilon)\big|_{\varepsilon=0} = 0 \Leftrightarrow \int_a^T [f_y(\cdots)h(t) + f_{y'}(\cdots)h'(t) + f_z(\cdots)h(T)] dt = 0, \tag{1}$$

where $(\cdots) = (t, \tilde{y}(t), \tilde{y}'(t), \tilde{y}(T))$. Integration by parts gives

$$\int_a^T f_{y'}(\cdots)h'(t) dt = f_{y'}(\cdots)h(t)\big]_a^T - \int_a^T \frac{d}{dt}(f_{y'}(\cdots))h(t) dt.$$

Because $h(a) = 0$, the necessary condition (1) can be then written as

$$0 = \int_a^T \left\{ \left[f_y(\cdots) - \frac{d}{dt} f_{y'}(\cdots) \right] h(t) + \left[\frac{f_{y'}(T, \tilde{y}(T), \tilde{y}'(T), \tilde{y}(T))}{T-a} + f_z(\cdots) \right] h(T) \right\} dt \tag{2}$$

for all $h(\cdot) \in C^1$ such that $h(a) = 0$. In particular, equation (2) holds for the subclass of functions $h(\cdot) \in C^1$ that do vanish at $h(T)$. Thus, the classical arguments apply, and therefore

$$f_y(\cdots) - \frac{d}{dt} f_{y'}(\cdots) = 0. \tag{3}$$

Equation (2) must be satisfied for all $h(\cdot) \in C^1$ with $h(a) = 0$, which includes functions $h(\cdot)$ that do not vanish at T. Consequently, equations (2) and (3) imply that

$$\int_a^T \left[\frac{f_{y'}(T, \tilde{y}(T), \tilde{y}'(T), \tilde{y}(T))}{T - a} + f_z(\cdots) \right] h(T) dt = 0$$

$$\Leftrightarrow h(T) \left(f_{y'}(T, \tilde{y}(T), \tilde{y}'(T), \tilde{y}(T)) + \int_a^T f_z(\cdots) dt \right) = 0,$$

that is,

$$f_{y'}(T, \tilde{y}(T), \tilde{y}'(T), \tilde{y}(T)) + \int_a^T f_z(\cdots) dt = 0. \tag{4}$$

We remark that in the classical setting f does not depend on $y(T)$, that is, $f_z = 0$. In that case (4) reduces to the well known natural boundary condition $f_{y'}(T, \tilde{y}(T), \tilde{y}'(T)) = 0$ (or, from an Hamiltonian optimal control perspective, $p(T) = 0$). We have just proved the following result:

Theorem 2.1: Let a and T be given real numbers, $a < T$. If $\tilde{y}(\cdot)$ is a solution of the problem

$$J[y(\cdot)] = \int_a^T f(t, y(t), y'(t), y(T)) dt \to extr$$
$$y(a) = \alpha \qquad (y(T) free) \tag{5}$$
$$y(\cdot) \in C^1,$$

then

$$\frac{d}{dt} f_{y'}(t, \tilde{y}(t), \tilde{y}'(t), \tilde{y}(T)) = f_y(t, \tilde{y}(t), \tilde{y}'(t), \tilde{y}(T)) \tag{6}$$

for all $t \in [a, T]$. Moreover,

$$f_{y'}(T, \tilde{y}(T), \tilde{y}'(T), \tilde{y}(T)) = -\int_a^T f_z(t, \tilde{y}(t), \tilde{y}'(t), \tilde{y}(T)) dt. \tag{7}$$

Remark 2.2: From an optimal control perspective one has

$$p(T) = f_{y'}(T, \tilde{y}(T), \tilde{y}'(T) \tilde{y}(T)),$$

where $p(t)$ is the Hamiltonian multiplier. Theorem 2.1 asserts that the usual necessary optimality conditions [the Euler-Lagrange equation (Gelfand and Fomin, 1963) or the Pontryagin maximum principle (Pontryagin et al., 1986)] hold for problem (5) by substituting the classical transversality condition $p(T) = 0$ with

$$p(T) = -\int_a^T f_z(t, \tilde{y}(t), \tilde{y}'(t), \tilde{y}(T)) dt. \tag{8}$$

3 An illustrative example

We consider an example that illustrates the new class of CoV problem. It has the same form as the complicated non-linear optimal intertemporal production policy problem. Consider the ODE system described by

$$y'(t) = u(t), \qquad y(0) = 0. \tag{9}$$

We wish to maximise

$$J[u(\cdot)] = \int_0^T f(t, y(t), u(t), y(T)) dt \tag{10}$$

where

$$f(t, y, u, z) = a\sqrt{u} - \left(\frac{3}{4} + z\sin(\pi t/10)\right) u \tag{11}$$

is a continuous function. The initial known state is $y(0) = 0$ and final state value $z = y(T)$ is free. In this example we set $T = 10$. The Hamiltonian is $H(t, y, u, p) = -f + p \cdot u$ and

$$\begin{cases} y'(t) = H_p(t, y(t), u(t), p(t)) \\ p'(t) = -H_y(t, y(t), u(t), p(t)). \end{cases}$$

Function f does not depend on y, and for an optimum (maximum in this example), the costate satisfies

$$p' = -H_y \Leftrightarrow p' = 0. \tag{12}$$

The stationarity condition is

$$H_u = 0$$

and this yields

$$u(t) = \frac{1}{4}(z\sin(\pi t/10) - p(t))^2. \tag{13}$$

From (8)

$$p(T) = -\int_0^{10} f_z(t, \tilde{y}(t), \tilde{y}'(t), \tilde{y}(T)) dt$$

holds, i.e.,

$$p(T) = \int_0^{10} \sin(\pi t/10) u(t) dt. \tag{14}$$

4 Numerical shooting algorithms

Let us consider the necessary conditions (NC) that need to be satisfied. For the system of ODEs (9) and (12) with control (13), the known zero initial condition $y(0)$ and a guessed initial value $p(0)$, we need to ensure that the natural boundary condition (14) is satisfied.

We need to solve the two point boundary value problem. Also we need to iterate the value of z used in (13) to ensure that in fact the value z equals the value obtained for $y(t)$ at $t = T$. When one has obtained convergence regarding the values $y(T)$ used in (13) and $p(T)$ (14), then the NC are satisfied and we should have the optimal solution.

Use the Newton shooting method with two guessed values v_1 and v_2 (Betts, 2001). We desire $v_1 = p(0)$ and $v_2 = p(T)$ as specified by equation (14). When the program obtains results with these two equations holding to a very high degree of accuracy, the necessary conditions NC hold and we should have the optimal solution. We have solved the shooting method problem using C++ and the highly accurate Numerical Recipes library routines (Press et al., 2007):

We integrate the system $(y(t), p(t), g_p(t), J(t))$ with $y(t)$ and $p(t)$ satisfying the system of ODEs (9) and (12), and

$$g_p(t) = \sin(\pi t / 10)u(t),$$

$$J'(t) = f(t, y(t), u(t), y(T)).$$

The results are $y(T) = 0.86928249597392515$, $p(T) = -0.46111638323272386$, $g_p(T) = -0.46111638323273074$ and

$$J(T) = 1.85448307363352 \tag{15}$$

Perturbations of the optimal control $u(t)$ by increasing and decreasing the value of $u(t)$ at a single time instant yield smaller $J(T)$ values. See Figure 1 for results on state variable $y(t)$ and control variable $u(t)$.

Figure 1 Optimal pair $(y(t), u(t))$ to problem (9)–(10) obtained from both symbolic computation and the shooting method
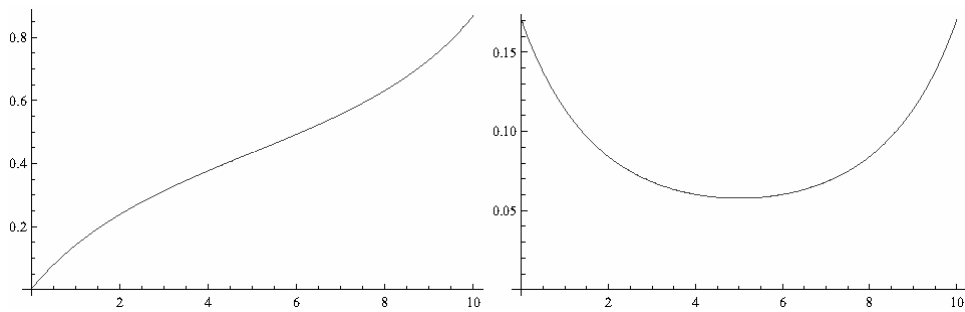
A completely different approach, using a non-linear programming technique, was also used. This technique may be useful for the actual piecewise constant economics problem. We solved this problem using Euler and Runge-Kutta discretisation, and an optimisation algorithm to solve for the unknown control variables u_k at each time instant t_k. We computed the non-linear programming problem using AMPL (Fourer et al., 2002) with

the MINOS solver and NEOS (2009). Using 40 time steps yields a good approximation very similar to the optimal results obtained using the precise approach described here.

5 Symbolic algebra solution

Consider the ODE system (9) and the associated optimal control problem described by (10) and (11). We set this as a minimisation problem. From (10) define

$$J_m[u(\cdot)] = \int_0^T g(t, y(t), u(t), y(T))dt \tag{16}$$

where

$$g(t, y, u, z) = -f(t, y, u, z) = \left(\frac{3}{4} + z\sin(\pi t / 10)\right)u - a\sqrt{u} \tag{17}$$

with the final state value $z = y(T)$ free and $T = 10$. We now use the Euler-Lagrange equation (6) to find candidate solutions:

$$\frac{d}{dt} g_{y'}(t, \tilde{y}(t), \tilde{y}'(t), \tilde{y}(T)) = g_y(t, \tilde{y}(t), \tilde{y}'(t), \tilde{y}(T))$$

for all $t \in [0, T]$. Set $u = y'$ and $z = y(T)$ so

$$\begin{aligned}
\frac{d}{dt} g_{y'}(t, y, y', y(T)) &= \frac{d}{dt} g_u(t, y, u, z) \\
&= \frac{d}{dt}\left(\frac{3}{4} - \frac{1}{2\sqrt{u}} + z\sin\left(\frac{\pi t}{10}\right)\right) \\
&= \frac{1}{10}\pi z\cos\left(\frac{\pi t}{10}\right) + \frac{u'}{4u^{3/2}}.
\end{aligned}$$

Since $g_y = 0$, using (6) we can find u by solving

$$\frac{1}{10}\pi z\cos\left(\frac{\pi t}{10}\right) + \frac{u'(t)}{4u(t)^{3/2}} = 0.$$

The next result was obtained using Maple:

$$u(t) = \frac{1}{\left(c + 2z\sin\left(\dfrac{\pi t}{10}\right)\right)^2}.$$

We find easily y from $y' = u$ using integration, e.g., in Maple or Mathematica:

$$y(t) = \frac{20\left(-\tan^{-1}(2z/D)g(t)c^2 + h(t)g(t)c^2 - zD\left(-\cos\left(\dfrac{\pi t}{10}\right)c + g(t)\right)\right)}{c\pi D^3 g(t)} \tag{18}$$

where

$$D = \sqrt{c^2 - 4z^2},$$

$$g(t) = \left(c + 2z \sin\left(\frac{\pi t}{10} \right) \right),$$

$$h(t) = \tan^{-1} \left(\frac{2z + c \tan\left(\frac{\pi t}{20} \right)}{\sqrt{c^2 - 4z^2}} \right).$$

Some comments:

- The function h is not defined for $t = 10$. However one can define h for $t = 10$ as

$$\lim_{t \to T} \tan^{-1} \left(\frac{2z + c \tan\left(\frac{\pi x}{20} \right)}{\sqrt{c^2 - 4z^2}} \right) = \frac{\pi}{2}.$$

- Theorem 2.1 assumes $u \in C^0$.

- From (17) $u(t) \geq 0$. If $u = 0$ we have $J_m = 0$ and we see that this is not the best solution. So $u > 0$ and $z = y(T) > 0$.

- We must verify $c^2 - 4z^2 > 0$ [see D in (18)] so two cases are to be investigated: $c > 2z$ and $c < -2z$.

Recall (7):

$$g_{y'}(T, \tilde{y}(T), \tilde{y}'(T), \tilde{y}(T)) = -\int_a^T g_z(t, \tilde{y}(t), \tilde{y}'(t), \tilde{y}(T)) dt. \tag{19}$$

We have

$$g_z(t, y, u, z) = u \sin\left(\frac{\pi t}{10} \right)$$

and using $u(\cdot)$ we obtain

$$g_z(t, y, u, z) = \frac{\sin\left(\frac{\pi t}{10} \right)}{\left(c + 2z \sin\left(\frac{\pi t}{10} \right) \right)^2}.$$

Integrating for the branch $c > 2z$ we obtain

$$\int_0^{10} g_z(t, y(t), y'(t), y(T)) dt = \frac{20 \left(2 \tan^{-1}\left(\frac{2z}{D} \right) z - \pi z + D \right)}{\pi D^3}.$$

The left-hand side of (19) is

$$g_{y'}(T, y(T), y'(T), y(T)) = \frac{1}{4}(3 - 2c).$$

Solving equation (19) numerically we get $c = 2.42223$ and $z = 0.869282$. The objective value is $J_m = -1.85448$ obtained using numerical integration over (16). This compares favourably with the result (15). Note that similar calculations for the branch $c < 2z < 0$ provide a worse solution with $J_m = +6.62857$ ($c = -7.21816$ and $z = 3.14287$).

The results obtained here by Symbolic Algebra Computations (SAC) agree with the shooting method results in Section 4.

6   Conclusions

In this note we have shown how the standard necessary optimality conditions and numerical procedures for problems of the CoV and optimal control should be adapted in order to cover Lagrangians depending on the free end-point. The numerical techniques were validated with a simple sample example that allows symbolic calculations using a modern computer algebra system. In the actual optimal intertemporal production policy economics problem the Lagrangian may be piecewise continuous and this requires amended numerical techniques, such as non-linear programming, for its solution. This numerical solution approach will be important for solution of the actual non-linear economics problem.

Acknowledgements

References

Betts, J.T. (2001) 'Practical methods for optimal control using nonlinear programming', *Advances in Design and Control*, Vol. 3, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Ferreira, R.A.C. and Torres, D.F.M. (2008) 'Higher-order calculus of variations on time scales', *Mathematical Control Theory and Finance*, pp.149–159, Springer, Berlin.

Fourer, R., Gay, D.M. and Kernighan, B.W. (2002) *AMPL: A Modeling Language for Mathematical Programming*, Duxbury Press/Brooks/Cole Publishing Company.

Gelfand, I.M. and Fomin, S.V. (1963) *Calculus of Variations*, revised English edition translated and edited by Richard A. Silverman, Prentice Hall, Englewood Cliffs, N.J.

Léonard, D. and Long, N.V. (1992) *Optimal Control Theory and Static Optimization in Economics*, Cambridge University Press, Cambridge.

Malinowska, A.B. and Torres, D.F.M. (2008) 'Natural boundary conditions in the calculus of variations', submitted, available at arXiv:0812.0705v1 [math.OC].

NEOS server for optimization (2009) available at http://www-neos.mcs.anl.gov/neos/, (accessed on July).

Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., Mishchenko, E.F. (1986) *The Mathematical Theory of Optimal Processes*, Selected works, Vol. 4, translated from the Russian by K.N. Trirogoff, translation edited by L.W. Neustadt, reprint of the 1962 English translation, Gordon & Breach, New York.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (2007) *Numerical Recipes – The Art of Scientific Computing*, 3rd ed., Cambridge Univ. Press, Cambridge.

Sethi, S.P. and Thompson, G.L. (2000) *Optimal Control Theory – Applications to Management Science and Economics*, 2nd ed., Kluwer Academic Publishers, Boston, MA.

Zinober, A. and Kaivanto, K. (2008) *Optimal Production Subject to Piecewise Continuous Royalty Payment Obligations*.