

APPLICATION OF THE MULTIPLE SHOOTING METHOD TO THE OPTIMIZATION OF A PRODUCTION MODEL

Raul Ikeda Gomes da Silva

Divisão de Engenharia Eletrônica
Instituto Tecnológico da Aeronáutica
São José dos Campos, SP, Brasil, 12228-900
raul@ele.ita.br

Marco Antonio Leonel Caetano

Departamento de Estatística, Matemática Aplicada e Computação
Universidade Estadual Paulista
Rio Claro, SP, Brasil, 13506-900
maleonel@rc.unesp.br

Takashi Yoneyama

Divisão de Engenharia Eletrônica
Instituto Tecnológico da Aeronáutica
São José dos Campos, SP, Brasil, 12228-900
takashi@ele.ita.br

Abstract. The purpose of this paper is to solve a two point boundary value problem (TPBVP) using Multiple Shooting Method. A production model used is a finished good warehouse to store products which are manufactured but not immediately consumed. Once a product is made and put into inventory it incurs the holding costs of two kinds: physically storing and insuring, and the carrying charges of having the firm's money invested in the unsold inventory. The Pontriaguin's Maximum Principle yields the TPBVP which was solve by the Multiple Shooting. The method is based on partitioning the time interval and iteratively reduce the error of guessing the initial conditions at each one interval. The code implemented uses a 7th/8th order Runge-Kutta-Felhberg integrator to guarantee the computation of solutions with low tolerances. The obtained solution is then evaluated in terms of ability to balance the benefits of the regulation of the production process versus the cost of holding inventory.

Keywords. Optimal Control, Multiple Shooting, TPBVP, Optimization.

1. Introduction

Application and optimization methods to production and inventory problems date back at list to the classical EOQ model and the lot size formula. The EOQ is essentially a static model in the sense that the demand is constant and only a stationary solution is sought. The solution methodology used was Dynamic Programming.

An important dynamic production planning model was developed by Holt et al (1958). In their model, referred to as the HMMS model, they considered production costs and inventory costs over time. They used calculus of variations techniques to solve the continuous time version of their model.

In this work, the model was purposed by Sethi et al (1981), which is similar to HMMS model and solved using optimal control theory. It was used the quasilinearization method to solve the problem. Many other methods as steeps descent or boundary iterations are used too. In this work was used the Multiple Shooting Method (Keller, 1976; Hall et al, 1976; Stoer et al, 1980; Deuflhard, 1980 e Deuflhard et al, 1982) which is based in Single Shooting Method. With this method are implemented an 7th/8th order Runge-Kutta-Felhberg integrator to guarantee the computation of solutions with low tolerances. A user-friendly interface was also developed to allow simple interactions in therms of analysis of sensitivities, alternative performance indexes and other what if tests.

2. The Production Inventory Model

Purposing a simple production inventory model, which describes the inventory at time t , it is increased by the production rate and decreased by the sales rate. The Equation (1) shows the stock flow differential equation, and Equation (2) shows the objective function of the model.

$$\dot{I}(t) = P(t) - S(t) \quad (1)$$

$$\text{Maximize}_{P \geq 0} \left\{ J = \int_0^T e^{-\rho t} \left[\frac{h}{2} (I(t) - \hat{I})^2 + \frac{c}{2} (P(t) - \hat{P})^2 \right] dt \right\} \quad (2)$$

Where $I(t)$ is the state variable that describes the inventory level at time t ; $P(t)$ the production rate at time t ; and $S(t)$ the sales rate at time t . It is also considered that $I(0) = I_0$, where I_0 represents the initial inventory level. Some others parameters are: h – inventory holding cost coefficient; c – production cost coefficient; ρ – constant nonnegative

discount rate; \hat{p} - production goal level; and \hat{i} - inventory goal level. It should be assumed the initial time equal zero ($t_0=0$) and T represents the final time, or length of planning period.

The interpretation of the inventory goal level, \hat{i} , is that it is a safety stock that the company wants to keep on hand; e.g., could be two months of sales, one quarter of sales, etc. Similarly, the production goal level, \hat{p} , can be interpreted as the desired or most efficient level at which we want to run the factory.

The interpretation of the objective function is that it is wanted to keep the inventory as close as possible to its goal level, \hat{i} , and also keep the production rate, $P(t)$, as close as to its goal level, \hat{p} . The quadratic terms impose penalties for having either $I(t)$ or $P(t)$ not being close to its corresponding goal level.

Associating now an adjoint function $\lambda(t)$ with the Equation (1) and can write the current value Hamiltonian function, which are showed in Equation (3).

$$H(I(t), P(t), \lambda(t), t) = \lambda(t)[P(t) - S(t)] - \frac{h}{2}[I(t) - \hat{I}]^2 - \frac{c}{2}[P(t) - \hat{P}]^2 \quad (3)$$

Applying the Pontryagin maximum principle, differentiating Equation (3) and the resulting expression equal to 0, showed in Equation (4). From this equation this is obtained the decision rule, described in Equation (5).

$$\frac{\partial}{\partial P} H(I(t), P(t), \lambda(t), t) = \lambda(t) - c[P(t) - \hat{P}] = 0 \quad (4)$$

$$P(t) = \hat{P} + \frac{\lambda(t)}{c} \quad (5)$$

This $P(t)$ is valid unless it drives $P(t)$ negative. In most cases $P(t)$ is constrained to be nonnegative, so the form of the optimal control is given by Equation (6).

$$P^*(t) = \text{Maximum}\left[\hat{P} + \frac{\lambda(t)}{c}, 0\right] \quad (6)$$

Assuming that \hat{p} is large enough so that Equation (5) always gives nonnegative production values. With this assumption can substitute Equation (5) on Equation (1) and obtain Equation (7). The equation for the adjoint variable is found as Equation (8).

$$\dot{I}(t) = \hat{P} + \frac{\lambda(t)}{c} - S(t) \quad (7)$$

$$\dot{\lambda}(t) = \rho\lambda(t) - \frac{\partial}{\partial I} H(I(t), P(t), \lambda(t), t) = \rho\lambda(t) + h(I(t) - \hat{I}) \quad (8)$$

It is now a two points boundary value problem to be solved with the initial condition $I(t_0) = I_0$, and the final condition $\lambda(T) = 0$.

3. The Single and Multiple Shooting Methods

In this section we analyze the single and the multiple shooting methods. The former is discussed in the first part of this section. The latter improves the single shooting by considering multiple divisions and is discussed in second part. The last part of this section considers some details for solution of the algebraic nonlinear equation appearing in the multiple shooting method.

3.1. The Single Shooting Method

This is a simple method to solve a two points boundary value showed in Session 2, which consists in guess the the initial value for the adjoint function. The Equation (9) shows a 2nd order differential equation with the boundary conditions $y(0) = 0$ and $y(L) = h$. This problem can be rewritten, in an equivalent form, as a initial value problem with the new initial conditions $y(0) = 0$ and $y'(0) = v$, choosing a value for v which the value $y(L) = h$.

$$\ddot{y} = f(t, y, \dot{y}) \quad (9)$$

It is Gussed a random value for v and with adjusts in this value it is possible to obtain the value of L . Assuming that $y(L;v)$ is the value of y when $t=L$ for a particular choose of v . This way it must just find the roots of the Equation (10).

$$F(v) = y(L;v) - y(L) = y(L;v) - h \quad (10)$$

To obtain the value for $F(v)$, we should use some numerical integrator to solve the Equation (9) with the initial conditions. To choose a new value for v , it can be used a recursive formula showed in Equation (11).

$$v^{(i+1)} = v^{(i)} - \frac{F(v^{(i)})}{\dot{F}(v^{(i)})} \quad (11)$$

Where index i represents the number of iterations. The derivative of $F(v)$ in Equation (11) can be written as Equation (12).

$$\dot{F}(v) \equiv \Delta F(v) \equiv \frac{F(v + \Delta v) - F(v)}{\Delta v} \quad (12)$$

The Δv should be sufficiently small chosen. The function $F(v + \Delta v)$ should be obtained as the same as $F(v)$, but with initial conditions $y(0) = v + \Delta v$. But two problems must be analyzed for this case: If Δv is too big, $\Delta F(v)$ will be a poor approximation for the derivative; If Δv is too small, the quotient will be zero.

So small rounding errors of the computer numerical integrator can make an impact over the result of $\Delta F(v)$. Although the $F(v)$ must be obtained with the best precision possible.

3.2. The Multiple Shooting Method

Some generic differential equation is represented by Equation (13), and its boundary conditions are represented by Equation (14).

$$\dot{y}(t) \equiv \frac{d}{dt} y(t) = f(t, y) \quad (13)$$

$$r[y(a), y(b)] = 0 \quad (14)$$

Where a is different of b and they are fixed. There are two observations. First, it is more common use separated boundary conditions, but the Equation (14) described a more generic form to represent it, including nonlinear or connected conditions. Second, it is not considered contact intervals, so $f(t, y)$ is unique for all interval $[a,b]$.

In session 3.1, the Simple Shooting Method was analyzed, which had purposed a solution for just one extreme points of the interval. However Osborne (1969) observed that there are two great difficulties in that technique: the problem of initial value can be intrinsically instable or bad conditioned mathematically. For example, equations with great Lipschitz constants. The second difficulty is that there is no good estimative for the initial conditions, which can cause divergence of the iterative method of nonlinear algebraic equations solution.

The Multiple Shooting Method try to shorten these difficulties. So the time interval $[a,b]$ is divided in $m-1$ subintervals ($m \geq 3$). The new interval: $a = t_1 < t_2 < t_3 < \dots < t_{m-1} < t_m = b$. The t_j , $j = 1, 2, 3, \dots, m$, are called nodes. Considering Y_j are an solution estimative for $y(t)$ in nodes t_j , now there are $m-1$ initial values problems for Equation (13).

Considering $y(t; t_j, Y_j)$ the solution of Equation (13) in point t with initial conditions $y(t_j) = Y_j$, this way the problem consist just to determine the Y_j values, such that the Equation (13) satisfies the conditions defined by Equation (15).

$$\begin{aligned} y(t_{j+1}; t_j, Y_j) &= Y_{j+1} \\ r[Y_1, y(t_m; t_{m-1}, Y_{m-1})] &= 0 \\ j &= 1, 2, \dots, m-1 \end{aligned} \quad (15)$$

These $(m-1)n$, where n are the n th differential equation order, conditions represents a algebraic nonlinear equation system showed in Equation (16), with $(m-1)n$ incognitos, which are $Y = [Y_1 \dots Y_{m-1}]^T$.

$$F(Y) = \begin{bmatrix} F_1(Y_1, Y_2) \\ F_1(Y_1, Y_2) \\ \vdots \\ \vdots \\ F_{m-1}(Y_1, Y_{m-1}) \end{bmatrix} = \begin{bmatrix} y(t_2; t_1, Y_1) - Y_2 \\ y(t_3; t_2, Y_2) - Y_3 \\ \vdots \\ \vdots \\ r[Y_1, y(t_m; t_{m-1}, Y_{m-1})] \end{bmatrix} = 0 \quad (16)$$

3.3. Algebraic Nonlinear Equations Solution

The Equation (16) described in Session 3.2., can be solved using the Newton Method, showed in Equation (17).

$$Y^{(i+1)} = Y^{(i)} - [J(Y^{(i)})]^{-1} F(Y^{(i)}) \quad (17)$$

Where $[J(Y)]$ is the jacobian matrix and in this case, by the structure of $F(Y)$ vector, is given by Equation (18).

$$[J(Y)] = \begin{bmatrix} G_1 & -I & 0 & \dots & 0 \\ 0 & G_2 & -I & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & G_{m-2} & -I \\ A & 0 & \dots & 0 & BG_{m-1} \end{bmatrix} \quad (18)$$

Where $[A]$, $[B]$, $[G_j]$, $j = 1, 2, \dots, m-1$ are jacobian matrix too, $n \times n$ order. Your expressions are described in Equation (19).

$$[G_j] \equiv \frac{\partial}{\partial Y_j} y(t_{j+1}; t_j, Y_j)$$

$$j = 1, 2, \dots, m-1$$

$$[A] \equiv \frac{\partial}{\partial Y_m} r(Y_1, Y_m) \quad (19)$$

$$[B] \equiv \frac{\partial}{\partial Y_1} r(Y_1, Y_m)$$

These derivatives must also be calculated by numerical method, as shown in Session 3.1. To obtain the solution of the problem it just have to calculate the correct value of vector Y , which can be done by the condensation algorithm. The Equation (20) denotes the ΔY vector. And rewriting the Equation (17) can obtain Equation (21). Equation (22) shows the multiplication between Equations (20) and (21).

$$\begin{bmatrix} \Delta Y_1 \\ \vdots \\ \vdots \\ \Delta Y_{m-1} \end{bmatrix} \equiv Y^{(i+1)} - Y^{(i)} \quad (20)$$

$$F_j \equiv F_j(Y_j^{(i)}, Y_{j+1}^{(i)}) = y(t_{j+1}; t_j, Y_j^{(i)}) - Y_{j+1}^{(i)}$$

$$[J(Y^{(i)})] \begin{bmatrix} \Delta Y_1 \\ \vdots \\ \vdots \\ \Delta Y_{m-1} \end{bmatrix} = -F(Y^{(i)}) \quad (21)$$

$$\begin{aligned}
G_1 \Delta Y_1 - \Delta Y_2 &= -F_1 \\
G_2 \Delta Y_2 - \Delta Y_3 &= -F_2 \\
&\vdots \\
G_{m-2} \Delta Y_{m-2} - \Delta Y_{m-1} &= -F_{m-2} \\
A \Delta Y_1 + B G_{m-1} \Delta Y_{m-1} &= -F_{m-1}
\end{aligned} \tag{22}$$

All the terms of the ΔY vector in Equation (22) can be expressed by the terms of ΔY_1 , as shown in Equation (23). And from the Equation (22) is obtained the Equations (24) (25) and (26) which represents a algebraic linear system for the vector ΔY_1 .

$$\begin{aligned}
\Delta Y_2 &= G_1 \Delta Y_1 + F_1 \\
\Delta Y_3 &= G_2 G_1 \Delta Y_1 + G_2 F_1 + F_2 \\
&\vdots
\end{aligned} \tag{23}$$

$$\Delta Y_{m-1} = G_{m-2} \dots G_2 G_1 \Delta Y_1 + \sum_{j=1}^{m-2} \left[\left(\prod_{\ell=j+1}^{m-2} G_\ell \right) F_j \right]$$

$$E \Delta Y_1 = W \tag{24}$$

$$W = -(F_{m-1} + B G_{m-1} F_{m-2} + \dots + B G_{m-1} G_{m-2} \dots G_2 F_1) \tag{25}$$

$$E = A + B G_{m-1} G_{m-2} \dots G_1 \tag{26}$$

Once obtained ΔY_1 , it is easily possible obtain $\Delta Y_2, \Delta Y_3, \dots, \Delta Y_{m-1}$ and with Equations (17) and (20) can have the new guess for Y vector, which is represented by Equation (27), and I represents a new iteration of the algorithm.

$$Y^{(i+1)} = Y^{(i)} + \begin{bmatrix} \Delta Y_1 \\ \Delta Y_1 \\ \vdots \\ \Delta Y_{m-1} \end{bmatrix} \tag{27}$$

4. Results

The Multiple Shooting Method was applied to solve the TBPVP described in Session 2, and was used a time varying demand showed in Equation (28). Assuming the inventory goal level, $\hat{i} = 15$, initial inventory condition $I_0 = 10$, final time $T = 8$, inventory holding cost coefficient $h = 1$ and production cost coefficient $c = 1$.

$$S(t) = t(t - 4)(t - 8) + 30 \tag{28}$$

Considering the discount rate $\rho = 0$ and a production goal level sufficiently big to not having negative production rate, $\hat{p} = 30$. Figure (1) illustrates the solution of this problem after the convergence in 16 iterations. The inventory was lead to the goal in the final time purposed, the production didn't have any negative point, and the initial condition for inventory was the as defined. The tolerance and derivative step was 0.1. The integrator tolerance was 0.15 and its step 0.001.

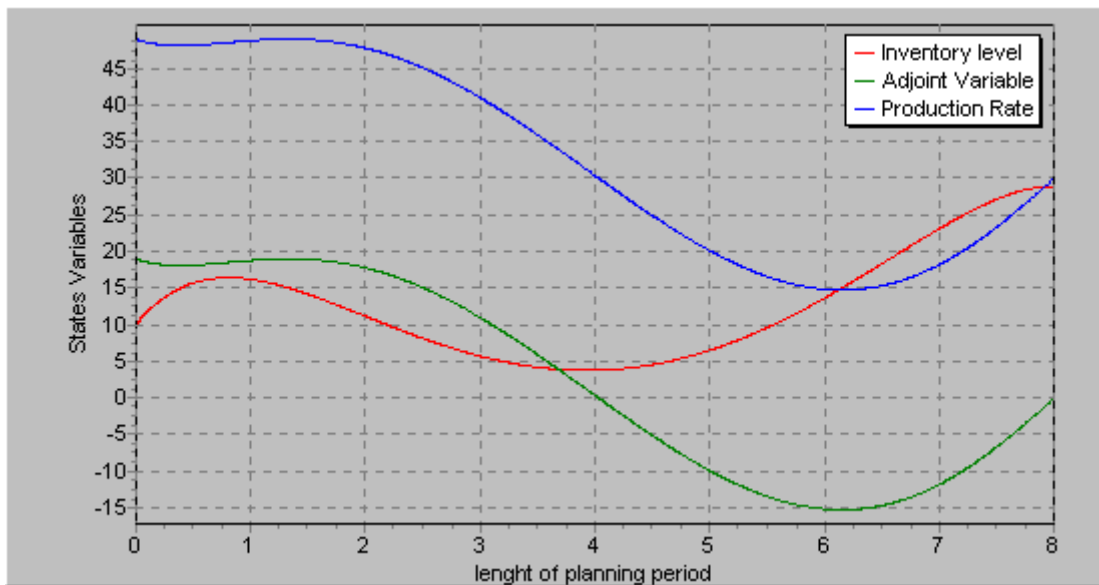


Figure 1. The optimal solution after the algorithm's convergence.

The number of time divisions, m , was 8, and the initial conditions was set equal to 0 for each Y_j , $j = 1, 2, \dots, 7$, even the I_0 condition. Figure (2) illustrates the algorithm after the first iteration. After a few iterations it is possible to observe that the algorithm tries to join the intervals first and then try to reach the boundary conditions. Figure (3) shows the algorithm after 2 iterations.

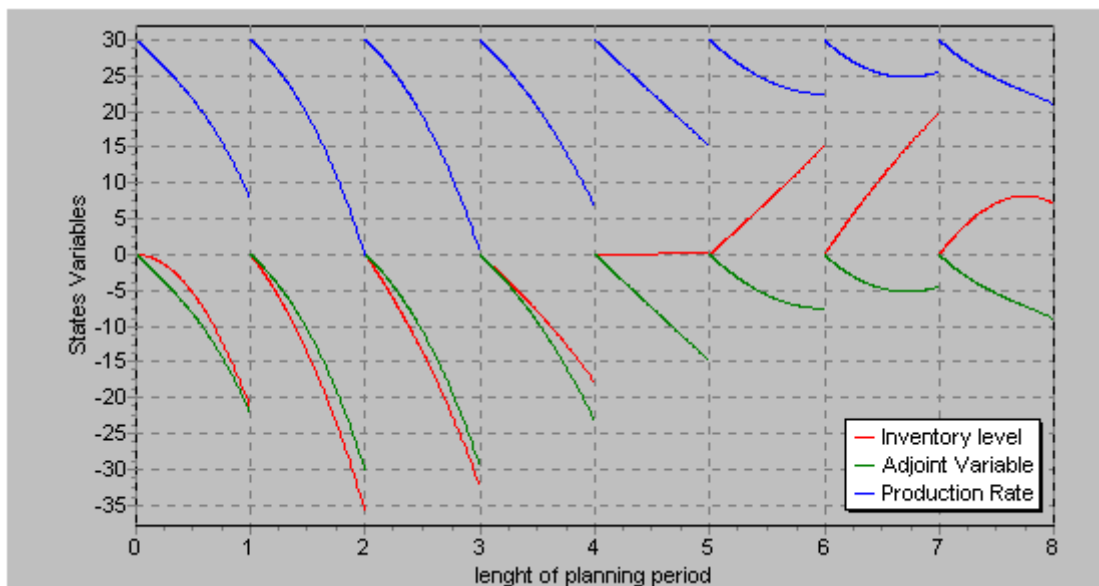


Figure 2. The states variables divided in m parts after 1st algorithm iteration.

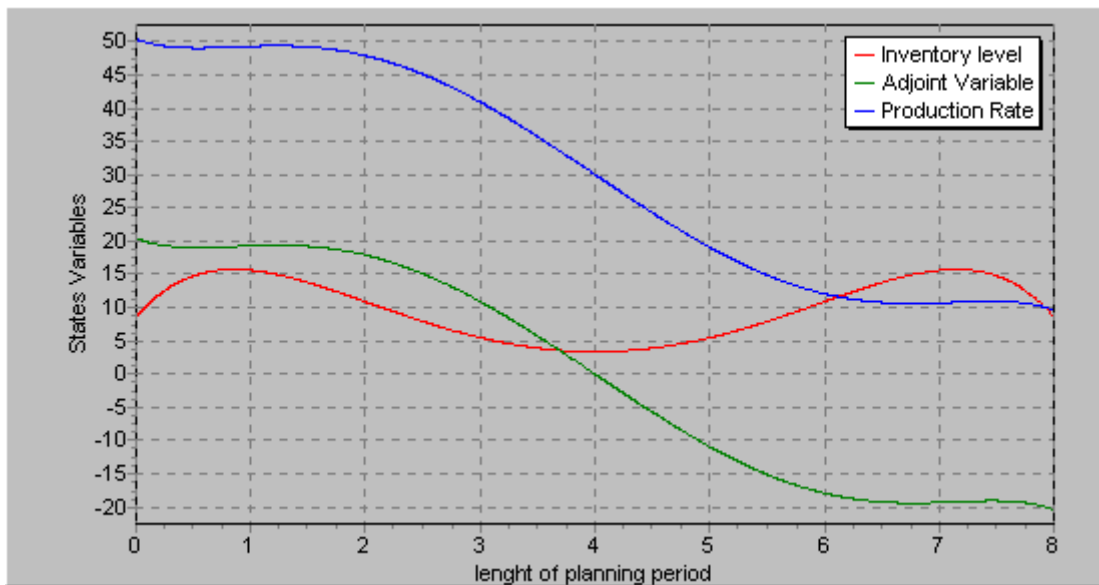


Figure 3. The states variables divided in m parts after 2nd algorithm iteration.

Figure (4) illustrates the final interface constructed for interact the user to the program. It is possible to set the tolerance and derivative pass for the solution of derivative, the tolerance and maximum pass for the integrator, the final time and the initial conditions for each division. At right side there are the results of inventory rate, adjoint variable state and the optimal control along the time. The legend shows the initial condition after each iteration.

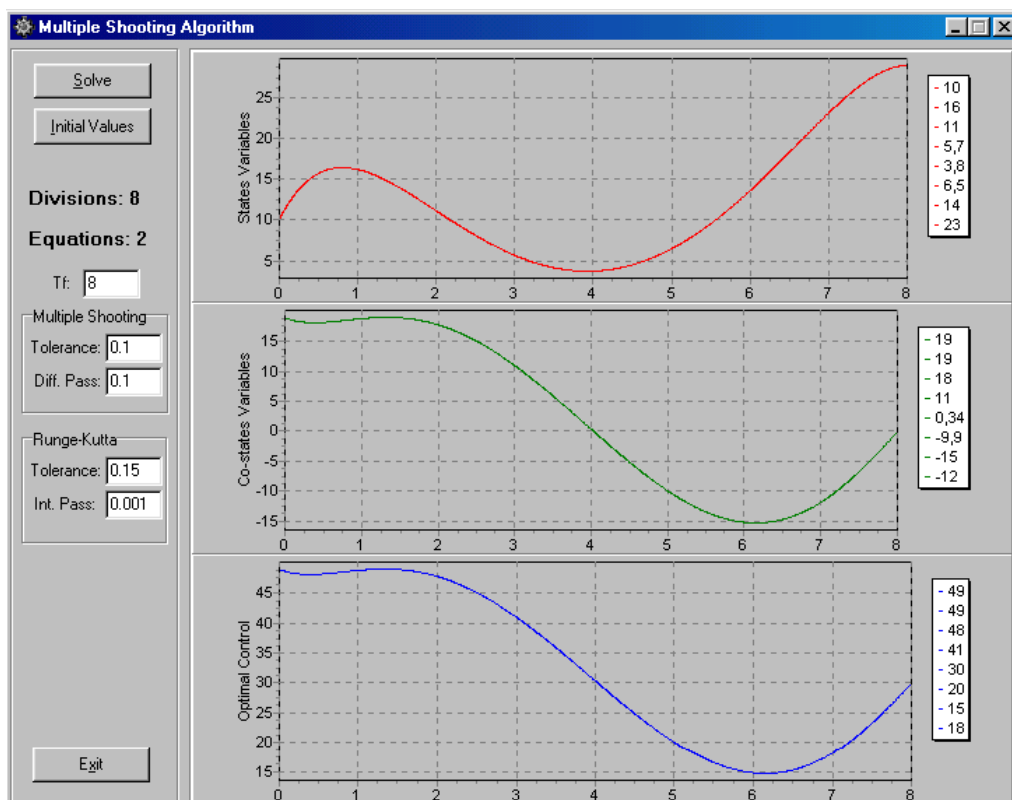


Figure 4. The user-friendly program interface.

5. Conclusion

For this simple time variant model the Multiple Shooting method could solve the production inventory problem. The time for convergence depends of the derivative step and tolerance. For more or less than eight to ten divisions there was a more difficulty to converge, but it converged too. Varying the final time would be interesting, and some cases the solution can diverge. The time to find the solution depends on the computational capability of the workstation too.

This method likes very simple to use when implemented than others, how qualiseliarization which can be very uncomfortable if the model has lots of state variables. Otherwise if the jacobian matrix is singular, or there is no optimal solution for the problem, can turn the method to diverge. The computer language implemented seemed very easy to work with high precision necessary. Future works can be tests of other models more complexes analyzing the robustness of this method.

6. Acknowledgement

The first author would like to thank CAPES for the financial support.

7. References

- Deuflhard, P. and Bader, G., 1982, "Multiple Shooting Techniques Revisited", SFB 123, No 163, University of Heidelberg.
- Deuflhard, P., 1980, "Recent Advances in Multiple Shooting Techniques", Computational Techniques for Ordinary Differential Equations, Academic Press, London.
- Fehlberg, E., 1969, "Low order classical Runge-Kutta formulas with step-size control and their application to some heat transfer problems", Nasa, TR-R 315, Huntsville.
- Keller, H.B., 1976, "Numerical Solution of Two-Point Boundary Value Problems", Society for Industrial and Applied Mathematics, Philadelphia.
- Kirk, E.D., 1970, "Optimal Control Theory an Introduction", Prentice-Hall, Englewood Cliffs, New Jersey.
- Hall, G. and Watt, J.M., 1976, "Modern Numerical Methods for Ordinary Differential Equations", Clarendon Press, Oxford.
- Holt, C., Modigliani, F., Muth, J. and Simon, H., 1960, "Planning Production, Inventories and Work Force", Prentice-Hall, Englewood Cliffs, New Jersey.
- Osborne, M.R., 1969, "On Shooting Methods for Boundary Value Problem", Journal of Mathematical Analysis and Applications, No. 27, pp. 417-433.
- Sethi, S.P. and Thompson, G.L., 1981, "Optimal Control Theory Applications to Management Science", Martinus Nijhoff Publishing, Boston.
- Stoer, J. and Bulirsch, R., 1980, "Introduction to Numerical Analysis", Springer-Verag, New York.

8. Copyright Notice

The author is the only responsible for the printed material included in his paper.