An Algorithm for Optimal Control of Delay Differential Equations.

Adeshina I. Adekunle, M.Tech.* and Olusegun Olotu, Ph.D.

Department of Mathematical Sciences, Federal University of Technology, PMB 704, Akure, Nigeria.

E-mail: aiadekunle@futa.edu.ng*

ABSTRACT

In this paper we are concerned with optimal control problems whose cost are quadratic and whose state are governed by linear delay differential equations and general boundary conditions. The basic new idea of this paper is to propose an efficient and robust algorithm for the solution of such problems by the conjugate gradient method (CGM) via quadratic programming. Our results are promising as compared with existing algorithms.

(Keywords: algorithm, numerical solution, conjugate gradient method, penalty method, quadratic programming)

INTRODUCTION

Optimal control problems constrained by delay differential equations using the maximum principle[6] involves the solution of a set of $2n$ two-point boundary-value problem in which both delay and advance terms are present. The solution of such problems is impossible computationally or otherwise. Therefore, the main objective of all computational aspects of optimal time-delay systems has been to device a methodology to avoid the solution of the mentioned 2-point boundary-value problem [6].

Jamishidi et al. [6], proposed a near optimum controller for linear systems with input time-delay through the introduction of a small parameter $0 \leq \varepsilon \leq 1$ and McLaurin's series expansion. The control has an exact feedback portion and a truncated series open loop gain, which is only valid for $\varepsilon << 1$. Jaechong [5], introduced a class of linear operators for linear delay differential equations in such a way that the state equation subject to a starting function can be viewed as an inhomogeneous boundary value problem in the linear operator equation. Although the method avoids the usual semi-group theory treatment to the problems but only gives the necessary theory for such problems.

Agrawal et al. [1], gives the necessary theory for constrained variational problem with time delay. Their theoretical consideration can be applied to obtain the analytical solution of certain variational problems. The Control parametrization Enhancing Technique (CPET) is extended to a general class of constrained time-delayed optimal control problems by Wong et al. [14]. A model transformation approach is used to convert the time-delay problem to an optimal control problem involving mixed boundary conditions, but without time- delay. This technique (CPET) increases the number of state and control variables involved over a reduced finite time interval in obtaining the solution of the problem.

Guo-Ping et al. [4], proposed a control method which transforms the differential equation with time delay of the system dynamics into a form without any time delay through a particular transformation. A numerical algorithm for control implementation is presented, since the obtained expression of the optimal controller contains an integral term that is not convenient for online calculation.

Smith [13], proposed an evolutionary algorithm to the optimal control of delay differential equations. Due to complexity and numerical intensity of the problem, a black box solver was developed which gives better solution though with long computational time compared to existing algorithms.

The computational method for solving optimal control problem which is governed by a switched system dynamical system with time delay is developed by Changzhi et al. [2]. They derived the required gradient of the cost function which is obtained via solving a number of delay differential equations forward in time, in which the resulting control problem can be solved as a mathematical programming problem.

An extended discretized scheme is proposed by Olotu et al. [10] to examine the convergence provide of a quadratic control problem constrained by evolution equation with real coefficients. With an unconstrained formulation of the problem via the menalty-multiplier method, the discretization of the time interval, and differential constraint is carried out. An operator, to circumvent the cumberson calculation inherent in some earlier schemes [8, 9], such as the funtion space algorithm, is established and proved.

In this paper, the discretized algorithm via quadratic programming technique [11] is extended to optimal control of delay differential equations. In the proposed algorithm, the optimal control problem is discretized and through the construction of penalty matrix, the optimal control problem becomes large sparse quadratic programming problem. The effectiveness and robustness of the control method is demonstrated by simulation studies of two dynamical models.

MATERIALS AND METHODS

Considering the following class of linear system with output time-delay.

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{x}(t-r) + \mathbf{C}\mathbf{u}(t), \qquad \textbf{(1)}$$
$$\mathbf{x}(t) = \mathbf{h}(t), \quad t \in [-r, t_0],$$

where $\mathbf{x} \in R^n, \mathbf{u} \in R^m$ are the state and control vectors, $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ are constant matrices of appropriate dimensions, $\mathbf{h}(t)$ is the state's initial function, t_0 is the initial process time and r is the time delay, assumed to be constant, but not necessarily small. A control vector $u(t)$ should be obtained which would minimize a quadratic functional,

$$J(\mathbf{u}) = \int_{t_0}^{t_f} (\mathbf{x}^T \mathbf{P}\mathbf{x} + \mathbf{u}^T \mathbf{Q}\mathbf{u})dt, \qquad \textbf{(2)}$$

and satisfy the constraints of Equation (1).

In Equation(2), the matrix $\mathbf{P}$ and $\mathbf{Q}$ are symmetric positive-definite, and t_f is the final process time assumed to be finite.

Theorem 1

Given the optimal control $\boldsymbol{u}^*$ and solution $\boldsymbol{x}(t)^*$ of the state system (1) that minimizes $J(\boldsymbol{u})$ over U; where U is set of admissible controls, then there exists adjoint variable $\mu(t) \in R^n$ satisfying:

$$\begin{cases} \dot{\mu}(t) = 2\mathbf{P}\mathbf{x}(t) - \mathbf{A}^T\mu(t) - \mathbf{B}^T\mu(t+r), & t_0 \le t \le t_f - r \\ \dot{\mu}(t) = 2\mathbf{P}\mathbf{x}(t) - \mathbf{A}^T\mu(t), & t_f - r \le t \le t_f \end{cases} \qquad \textbf{(3)}$$

and with transversality condition

$$\mu(T) = 0 \qquad \textbf{(4)}$$

$$\mathbf{u}^*(t) = \frac{\mathbf{Q}^{-1}\mathbf{C}^T}{2}\mu(t) \qquad \textbf{(5)}$$

Proof 1

By introducing adjoint variable $\mu(t) \in R^n$, the required augmented functional from Equation (1) and Equation (2) can be formed. The Hamiltonian function is given as:

$$\mathbf{H}(\mathbf{x}, \mathbf{u}, \mu, \dot{\mathbf{x}}) = \mathbf{x}^T(t)P\mathbf{x}(t) + \mathbf{u}^T(t)Q\mathbf{u}(t) + \mu^T(t)(\dot{\mathbf{x}} - A\mathbf{x}(t) - B\mathbf{x}(t-r) - C\mathbf{u}(t)). \qquad \textbf{(6)}$$

Obtaining the necessary conditions for optimal control problem using the Euler-Lagrangian(E-L) equations for $\mathbf{H}$, regarded as function of the four vector variables($\mathbf{x}, \mathbf{u}, \mu, \dot{\mathbf{x}}$), we have:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{x}(t-r) + \mathbf{C}\mathbf{u}(t), \qquad t \in [0, T] \qquad \textbf{(7)}$$

$$\dot{\mu}(t) = 2\mathbf{P}\mathbf{x}(t) - \mathbf{A}^T \mu(t) - \mathbf{B}^T \mu(t+r), 0 \le t \le T - r \quad \textbf{(8)}$$

$$\dot{\mu}(t) = 2\mathbf{P}\mathbf{x}(t) - \mathbf{A}^T \mu(t), \qquad T - r < t \le T \quad \textbf{(9)}$$

$$u(t) = \frac{\mathbf{Q}^{-1}\mathbf{C}^T}{2} \mu(t), \qquad t \in [0, T], \quad \textbf{(10)}$$

which are the required optimality conditions.

The optimality system of Equations (7) - (10) represents a system of linear 2-point boundary values problems involving both delay and advance terms. Clearly, the coupling that exists between $\mathbf{x}(t)$ and $\mu(t)$ and the fact that variables with t, $t-r$, $t+r$ arguments are involved, make the solution of such problems impossible, both analytically and numerically by indirect method.

DEVELOPMENT OF THE ALGORITHM

In order to obtain numerical solutions to Equations (1) and (2) by direct method, we shall replace the constrained problem by appropriate discrete optimal control problem. Partitioning the interval $[t_0, t_f]$ into s sub-intervals with knots $t_0 < t_1 < t_2 \cdots < t_f$ and $t_k = k\Delta t_k$, where Δt_k is is the mesh size of k_{th} sub-interval . If these sub-intervals are small enough, we can assume that in any sub-interval $[k-1, k]$, the values $x(t)$ and $u(t)$ can be approximated by zero order spline $\mathbf{x}_k$ and $\mathbf{u}_k$, respectively.

Our optimal control problem(1) and (2) is then approximated by:

$$\min J(\mathbf{u}) = \sum_{k=0}^{s} (\mathbf{x}_k^T \ t_k)\mathbf{M}\mathbf{x}_k(t_k) + \mathbf{u}_k^T(t_k)\mathbf{N}\mathbf{u}_k(t_k) \quad \textbf{(11)}$$

subject to,

$$\dot{\mathbf{x}}_k(t_k) = \mathbf{A}\mathbf{x}_k(t_k) + \mathbf{B}\mathbf{x}_k(t_k - r) + \mathbf{C}\mathbf{u}_k(t_k) \quad \textbf{(12)}$$

where $\mathbf{M} = \mathbf{P}\Delta t_k$ and $\mathbf{N} = \mathbf{Q}\Delta t_k$. Furthermore, we shall use finite difference approximation to write:

$$\mathbf{x}_{k+1}(t_k) = \mathbf{x}_k(t_k) + \dot{\mathbf{x}}_k(t_k)\Delta t_k, \quad \textbf{(13)}$$

thus, the constraint(12) becomes:

$$\mathbf{x}_{k+1}(t_k) = \mathbf{F}\mathbf{x}_k(t_k) + \mathbf{G}\mathbf{x}_k(t_k - r) + \mathbf{H}\mathbf{u}_k(t_k) \quad \textbf{(14)}$$

where $\mathbf{F} = I_{n\times n} + \mathbf{A}\Delta t_k$, $\mathbf{G} = \mathbf{B}\Delta t_k$, and $\mathbf{H} = \mathbf{C}\Delta t_k$.

By parameter optimization [7], the discretized problem becomes a large sparse quadratic programming problem. We give a matrix representation:

$$\min J(z) = \mathbf{z}_k^T(t_k)\mathbf{D}\mathbf{z}_k(t_k) + \mathbf{c} \quad \textbf{(15)}$$

subject to,

$$\mathbf{E}\mathbf{z}_k(t_k) = \mathbf{k} \quad \textbf{(16)}$$

and,

$$\mathbf{z}_k^T(t_k) = (\mathbf{x}_1^T(t_k), \mathbf{x}_2^T(t_k), \cdots, \mathbf{x}_s^T(t_k), \mathbf{u}_0^T(t_k), \mathbf{u}_1^T(t_k), \cdots, \mathbf{u}_s^T(t_k)) \quad \textbf{(17)}$$

where $\mathbf{D}$ is a block diagonal matrix of order $(n+m)s+m$, with entries given by:

$$[\mathbf{D}]_{ii} = \begin{cases} \mathbf{M}, & i = 1, 2, \cdots, s \\ \mathbf{N}, & i = s+1, s+2, \cdots, 2s+1 \end{cases}$$

where i^{th} element corresponds to i^{th} block, and $\mathbf{c} = \mathbf{x}_k^T(0)\mathbf{M}\mathbf{x}_k(0)$. The matrix $\mathbf{E}$ is block matrix of order $ns \times (n+m)s+m$ with the representation:

$$\mathbf{E} = \begin{pmatrix} \mathbf{K} & \vdots & \mathbf{L} & \vdots & 0 \end{pmatrix}. \quad \textbf{(18)}$$

Let $n_1 = \frac{r}{\Delta t_k}$ and $\mathbf{x}_{k-n_1}(t_k) = x_k(t_k - r)$, then $\mathbf{K}$ is an $ns \times ns$ sparse block matrix with principal block diagonal elements $[\mathbf{K}]_{ii} = I_{n\times n}$ and lower block principal diagonal elements $[\mathbf{K}]_{ij} = -\mathbf{F}, \forall \quad i, j$ block, such that $i = j+1$

and

$[\mathbf{K}]_{ij} = -\mathbf{G}, \forall \quad i = n_1+1, \cdots, n, \quad and \quad j = 1, 2, \cdots, n_1$.

$[\mathbf{L}]$ is an $ns \times ms$ block diagonal matrix with block diagonal elements $[\mathbf{L}]_{ii} = -\mathbf{H}$, and 0 is an $ns \times m$ zero matrix. The column vector $\mathbf{k}$ is of order $ns \times 1$ with entries given by:

$[\mathbf{k}]_{1:n,1} = \mathbf{F}\mathbf{x}_k(0) + G\mathbf{x}_{-n_1}(-r)$,

$[\mathbf{k}]_{nj:n(j+1)} = \mathbf{G}\mathbf{x}_{i-n_1}(t_i - r), i = 2, 3, \cdots, n_1, j = i-1$,

and $[\mathbf{K}]_{i1} = 0, i = n_1+1, \cdots, ns.$

Using proposition 2.8 of [3], the quadratic programming (QP) problem (15) and (16) is equivalent to the solution of the saddle point system of linear equations:

$$\begin{pmatrix} \mathbf{D} & \mathbf{E}^T \\ \mathbf{E} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{z}_k(t_k) \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{k} \end{pmatrix} \qquad \textbf{(19)}$$

where $\lambda \in R^{ns}$ is the Lagrange multipliers. If $\mathbf{E}$ is a full row rank matrix, we can solve Equation (19) effectively by the Gaussian elimination with suitable pivoting strategy, or by a symmetric factorization which takes into account that Equation (19) is indefinite. Alternatively, we can use MINRES, a Krylov space method which generates the iterates minimizing the Euclidean of the residual in the Krylov space. The performance of the MINRES depend on the spectrum of the KKT system (19), similarly as the performance of the conjugate gradient method. Hence, unconstrained minimization problem by penalty function method is:

$$\min L_\rho(\mathbf{z}_k(t_k)) = \mathbf{z}^T(t_k)\mathbf{D}\mathbf{z}(t_k) + \mathbf{c} + \rho\mathbf{E}\mathbf{z}(t_k) - \mathbf{k}, \mathbf{E}\mathbf{z}(t_k) - \mathbf{k} \qquad \textbf{(20)}$$

on expansion, we have:

$$\min L_\rho(\mathbf{z}(t_k)) = \mathbf{z}^T(t_k)\mathbf{A}_\rho\mathbf{z}(t_k) + \mathbf{B}_\rho\mathbf{z}(t_k) + \mathbf{C}_\rho \qquad \textbf{(21)}$$

Equation (21) is the quadratic form representation for the unconstrained minimization problem, where $L_\rho(\mathbf{z}(t_k))$ is penalized Lagrangian, ρ is penalty parameter, the penalized matrix $\mathbf{A}_\rho = \left[\mathbf{D} + \rho\mathbf{E}^T\mathbf{E}\right], \mathbf{B}_\rho = -2\rho\mathbf{k}^T\mathbf{E}$ and $\mathbf{C}_\rho = \rho\mathbf{k}^T\mathbf{k} + \mathbf{c}$.

Lemma 1 (Olotu et al.)

Consider the continuous optimal control problem(1) and (2) and the associated discretized optimal control problem (15) and (16), the matrix $\boldsymbol{D}$ defined in (21) is positive symmetric definite and well-conditioned.

The property of a problem being well-conditioned indicates the problem is independent of the numerical method that is being used to obtain its solution. Since we have established the positive symmetric definiteness of $\mathbf{D}$, we state the following lemmas.

Lemma 2 (Dostial)

Let $\boldsymbol{D} \in R^{((n+m)s+m)\times((n+m)s+m)}$ be a symmetric positive definite matrix, let $\boldsymbol{E} \in R^{(ns)\times((n+m)s+m)}, \rho > 0$, and let $Ker\boldsymbol{D} \bigcap Ker\boldsymbol{E} = 0$. Then the penalized matrix $\boldsymbol{A}_\rho$ is positive definite.

Lemma 3.3 (Dostial)

Let $\boldsymbol{D} \in R^{((n+m)s+m)\times((n+m)s+m)}$ be a symmetric positive definite matrix, let $\boldsymbol{E} \in R^{(ns)\times((n+m)s+m)}, \mu > 0$ such that $\mathbf{z}^T(t_k)\mathbf{D}\mathbf{z}(t_k) \geq \mu\mathbf{z}(t_k), \mathbf{z}(t_k) \in Ker\mathbf{D}$.Then $\mathbf{A}_\rho$ is positive definite for sufficiently large ρ.

The lemmas ensure the sficient condition for $\mathbf{z}^*(t_k) \in R^{((n+m)s+m)}$ to a be local minimum point. We solve the unconstrained minimization Equation (21) by conjugate gradient algorithm in the inner loop and enforce the feasibility condition in the outer loop as stated in the following algorithm.

Algorithm 1

Step 1. Select a $\mathbf{z}_{0,0}(t_k) \in IR^{(n+m)s+m}$, $c > 1$ and $\rho_0 > 0$. Set $k = 0$.

Step 2. Set $i = 0$ and set $p_0 = -g_0 = -\Delta L_\rho(\mathbf{z}_{0,0}(t_k))$

Step 3. Compute $\alpha_i = \dfrac{p_i^T p_i}{p_i^T \mathbf{A}_\rho p_i}$

Step 4. Set $\mathbf{z}_{0,i+1}(t_k) = \mathbf{z}_{0,i} + \alpha_i p_i$

Step 5. Compute $\Delta L_\rho(\mathbf{z}_{i+1}(t_k))$

Step 6.If $\Delta L_\rho(\mathbf{z}_{0,i+1}(t_k)) = 0$ and $\mathbf{E}\mathbf{z}_{0,i+1}(t_k) = \mathbf{k}$ stop;else go to step 7.

Step7.If $\Delta L_\rho(\mathbf{z}_{0,i+1}(t_k)) \neq 0$,set

$g_{i+1} = \Delta L_\rho(\mathbf{z}_{0,i+1}(t_k))$,

$p_{i+1} = -g_{i+1} + \gamma_i p_i$, with $\gamma_i = \dfrac{g_{i+1}^T g_{i+1}}{g_i^T g_i}$

Step 8. Set $i = i + 1$, and go to step 3.

Step 9. Else if $\mathbf{E}\mathbf{z}_{0,i+1}(t_k) \neq \mathbf{k}$, set $\rho_{k+1} = c\rho_k$; set $k = k + 1$ and go to step 2.

ERROR AND CONVERGENCE ANALYSIS

In this section, the error and convergence analysis of the Algorithm 1 is given as a computational algorithm via quadratic programming for obtaining optimal solutions to optimal control of delay differential equations. The feasibility error estimates for the penalty approximation of quadratic programming problem (15) and (16) is stated in the theorem below. The error bound is proportional to $\dfrac{1}{\rho}$, but dependent on $\mathbf{E}$ and $\mathbf{k}$.

Theorem 2

Let $\boldsymbol{D}$, $\boldsymbol{E}$, $\boldsymbol{c}$ and $\boldsymbol{k}$ be those of the definition of problem(15) to (16) with $\boldsymbol{E} \neq 0$ not necessarily a full rank matrix, let $\beta_{min} > 0$ denote the smallest nonzero eigenvalues of $\boldsymbol{E}\boldsymbol{D}^{-1}\boldsymbol{E}^T$, let ε denote a given positive number, and let $\rho > 0$ .If $\mathbf{z}_k(t_k)$ is such that:

$$\nabla L_\rho \leq \varepsilon, \qquad \textbf{(22)}$$

then the feasibility error satisfies

$$\mathbf{E}\mathbf{z}_k(t_k) - \mathbf{k} \leq (1 + \beta_{\min}\,\rho)^{-1}\left(\frac{\varepsilon}{2}\mathbf{E}\mathbf{D}^{-1} + \mathbf{k}\right), \qquad \textbf{(23)}$$

Proof 2

Let us recall that for any vector $\mathbf{z}_k(t_k)$

$$\nabla L_\rho(\mathbf{z}_k(t_k)) = 2\left(\mathbf{D} + \rho\mathbf{E}^T\mathbf{E}\right)\mathbf{z}_k(t_k) - 2\rho\mathbf{E}^T\mathbf{k},$$

so that, after denoting $\mathbf{g} = \nabla L_\rho(\mathbf{z}_k(t_k))$ and $\mathbf{A}_\rho = \left(\mathbf{D} + \rho\mathbf{E}^T\mathbf{E}\right)$,

$$\mathbf{z}_k(t_k) = \frac{1}{2}\mathbf{A}_\rho^{-1}(\mathbf{g} + 2\rho\mathbf{E}^T\mathbf{k}).$$

It follows that,

$$\mathbf{E}\mathbf{z}_k(t_k) = \frac{1}{2}\mathbf{E}\mathbf{A}_\rho^{-1}(\mathbf{g}) + \rho\mathbf{E}\mathbf{A}_\rho^{-1}\mathbf{E}^T\mathbf{k}$$

Using,

$$\mathbf{E}\mathbf{A}_\rho^{-1} = (I + \rho\mathbf{E}\mathbf{D}^{-1}\mathbf{E}^T)^{-1}\mathbf{E}\mathbf{D}^{-1}, \qquad \textbf{(24)}$$

which can be shown from Sherman Morrison-Woodbury formula and the fact that $\mathbf{D}$ is symmetric positive definite, and by simple manipulation, we get:

$$\begin{aligned}\mathbf{E}\mathbf{z}_k(t_k) - \mathbf{k} &= \frac{1}{2}\mathbf{E}\mathbf{A}_\rho^{-1}(\mathbf{g}) + \rho(I + \rho\mathbf{E}\mathbf{D}^{-1}\mathbf{E}^T)^{-1}\mathbf{E}\mathbf{D}^{-1}\mathbf{E}^T\mathbf{k} - \mathbf{k} \\ &= \frac{1}{2}\mathbf{E}\mathbf{A}_\rho^{-1}(\mathbf{g}) + (I + \rho\mathbf{E}\mathbf{D}^{-1}\mathbf{E}^T)^{-1}\left((I + \rho\mathbf{E}\mathbf{D}^{-1}\mathbf{E}^T) - I\right)\mathbf{k} - \mathbf{k} \\ &= \frac{1}{2}\mathbf{E}\mathbf{A}_\rho^{-1}(\mathbf{g}) - (I + \rho\mathbf{E}\mathbf{D}^{-1}\mathbf{E}^T)^{-1}\mathbf{k}.\end{aligned}$$

By using the assumptions that $\mathbf{k} \in Im(\mathbf{E})$ and $g \le \varepsilon$, and Lemma 1.6 of [3], and properties of norm, then Equation(23) is immediate.

The bounds on the approximation error of the discretized continuous algorithm via quadratic programming using the feasibility error estimates (23) can be improved. The improvement on the approximation error of the optimal point is stated as follows.

Theorem 3

Let $\boldsymbol{D}$, $\boldsymbol{E}$, $\boldsymbol{c}$ and $\boldsymbol{k}$ be those of the definition of problem (15) and (16) with $\boldsymbol{E}$ not necessarily a full rank matrix, let λ_{min} denote the least eigenvalue of $\boldsymbol{D}$, let σ_{min} denote the least nonzero singular value of $\boldsymbol{E}$, let $(\hat{\mathbf{z}}_k(t_k), \boldsymbol{\lambda}_{LS})$ denote the least square KKT pair for problem(15) and (16), let $\beta_{min} > 0$ denote nonzero eigenvalue of the matrix $\boldsymbol{E}\boldsymbol{D}^{-1}\boldsymbol{E}^T$, let $\varepsilon > 0$, $\rho > 0$, and

$$\boldsymbol{\lambda} = \rho(\mathbf{E}\mathbf{z}_k(t_k) - \mathbf{k}). \quad \textbf{(25)}$$

If $\mathbf{z}_k(t_k)$ is such that

$$\nabla L_\rho \le \varepsilon,$$

then

$$\boldsymbol{\lambda} - \boldsymbol{\lambda}_{LS} \le \varepsilon \frac{\kappa(\mathbf{D})}{\sigma_{\min}} + \frac{D(\frac{\varepsilon}{2}\mathbf{E}\mathbf{D}^{-1} + \mathbf{k})}{\sigma_{\min}^2 (1+\rho)\beta_{min}} \quad \textbf{(26)}$$

and

$$\mathbf{z}(t_k) - \hat{\mathbf{z}}_k(t_k) \le \varepsilon \frac{\kappa(\mathbf{D}) + 1}{\lambda_{\min}} + \frac{\kappa(\mathbf{D})(\frac{\varepsilon}{2}\mathbf{E}\mathbf{D}^{-1} + \mathbf{k})}{\sigma_{\min}(1+\rho\beta_{min})} \quad \textbf{(27)}$$

Proof 3

Let us denote $\mathbf{g} = \nabla L_\rho(\mathbf{z}_k(t_k))$ and $\mathbf{e} = \mathbf{E}\mathbf{z}_k(t_k) - \mathbf{k}$, so that,

$$2(\mathbf{D}\mathbf{z}_k(t_k) + \mathbf{E}^T \lambda) = \mathbf{g} and \mathbf{E}\mathbf{z}_k(t_k) = \mathbf{k} + \mathbf{e}$$

If

$$\mathbf{g} = L_\rho(\mathbf{z}_k(t_k)) \le \varepsilon,$$

then by Theorem(2)

$$\mathbf{E}\mathbf{z}_k(t_k) - \mathbf{k} \le (1 + \beta_{\min}\rho)^{-1}(\frac{\varepsilon}{2}\mathbf{E}\mathbf{D}^{-1} + \mathbf{k}).$$

Substituting into the estimates(2.47) and (2.48) of Proposition 2.12 of [3], we get:

$$\mathbf{E}^T(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{LS}) \le \varepsilon\kappa(\mathbf{D}) + \frac{D(\frac{\varepsilon}{2}\mathbf{E}\mathbf{D}^{-1} + \mathbf{k})}{\sigma_{\min}(1+\rho\beta_{min})} \quad \textbf{(28)}$$

and Equation (27). To complete the proof, notice that $\lambda - \lambda_{LS} \in Im\mathbf{E}$. , so that,

$$\sigma_{\min} \boldsymbol{\lambda} - \boldsymbol{\lambda}_{LS} \le \mathbf{E}^T(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{LS})$$

In investigating Algorithm 1, we are often interested in the rate at which it converges to a limit. Given a sequence $\{\mathbf{z}_k(t_k)\} \subset R^{(n+m)s+m}$ with $\mathbf{z}_k(t_k) \to \mathbf{z}^*$, the typical approach is to measure the rate of convergence in terms of error function,

$$e : R^{(n+m)s+m} \to R$$

such that $e(\mathbf{z}_k(t_k)) \ge 0$. , $\forall \mathbf{z}_k(t_k) \in R^{(n+m)s+m}$. and $e(\mathbf{z}^*(t_k)) = 0$.

Assume that $e_k \ne 0$. , $\forall k$ and

$$\lim_{k\to\infty} \frac{e_{k+1}}{e_k^p} = \lim_{k\to\infty} \frac{\mathbf{z}_{k+1}(t_k) - \mathbf{z}^*(t_k)}{\mathbf{z}_k(t_k) - \mathbf{z}^*(t_k)^p} \le \beta < 1 \quad \textbf{(29)}$$

If $p = 1$ and $0 \le \beta \le 1$ then $\{\mathbf{z}_k(t_k)\}$ converges linearly with convergence ratio β. If $\beta = 0$ then $\{\mathbf{z}_k(t_k)\}$ converges super linearly with convergence ratio β. Also, if $p = 2$ and $0 \le \beta \le 1$ then z$\{\mathbf{z}_k(t_k)\}$ converges quadratically with convergence ratio β.

Most optimization algorithms of interest produce sequence converging either quadratically or super linearly. It is known in literature [12] that algorithms using conjugate gradient method are said to converge quadratically. Thus,in this paper we also examine the quadratic convergence of the developed algorithms for numerical examples below.

RESULTS AND DISCUSSION

In this section, we demonstrate the reliability of Algorithm (3) to optimal control of delay differential equation to other methods such as control parametrization technique (CPT) [14] and evolutionary algorithms [13]. All computations in the following examples were performed in the MATLAB® environment, Version 7.6.0324 Release (2008a) running on a Microsoft Windows $Vista^{TM}$. Home Premium operating system with an Intel® Pentium® Dual processor running at 1.87GHz.

In order to investigate the performance of the new algorithm, we consider the following continuous optimal control problems constrained by delay differential equations due to Smith[13];

Example 1

$$\text{Minimize} \qquad J = \int_0^2 (x^2(t) + u^2(t))\, dt \qquad \textbf{(30)}$$

subject to

$$\dot{x}(t) = tx(t) + x(t-1) + u(t), \qquad \textbf{(31)}$$
$$x(t) = 1, \quad t \in [-1, 0].$$

The optimality system is:

$$\dot{x}(t) = tx(t) + x(t-r) + u(t), \qquad t \in [0,2], \qquad \textbf{(32)}$$

$$\dot{\mu}(t) = 2x(t) - t\mu(t) + \mu(t+1), \qquad t \in [0,1], \qquad \textbf{(33)}$$

$$\dot{\mu}(t) = 2x(t) - t\mu(t), \qquad t \in [1,2], \qquad \textbf{(34)}$$

$$u(t) = \frac{\mu(t)}{2}, \qquad t \in [0,2], \qquad \textbf{(35)}$$

where $\mu(t)$ is the adjoint variable. The optimality system (32)-(35) is not amenable to both analytical method of solutions and indirect numerical methods. Applying classical control parametrization technique proposed by Wong[14], we have, by letting,

$$y_1(t) = x(t), \quad t \in [0,1]$$
$$y_2(t) = x(t+1), \quad t \in [0,1]$$
$$v_1(t) = u(t), \quad t \in [0,1]$$
$$v_2(t) = u(t+1), \quad t \in [0,1]$$

the resulting non-delayed re-formulation of Example 1 as:

$$\text{Minimize} \qquad J = \int_0^1 (y_1^2(t) + y_2^2(t) + v_1^2(t) + v_2^2(t))\, dt \qquad \textbf{(36)}$$

subject to the non-delayed differential equations on the interval $[0,1]$

$$\dot{y}_1(t) = ty_1(t) + 1 + v_1(t), \qquad \textbf{(37)}$$
$$\dot{y}_2(t) = ty_2(t) + y_1(t) + v_2(t) \qquad \textbf{(38)}$$

together with boundary constraints

$$y_1(0) = 1, \qquad \textbf{(39)}$$
$$y_1(1) = y_2(0). \qquad \textbf{(40)}$$

The solution of optimal control problem (36)-(40) is $J = 6.1391$, while the new algorithm gives $J = 4.7971$ in 12 seconds, which compared favorably with evolutionary algorithm which has best value of $J = 4.796817$ and worst value of $J = 4.84888$ in hours. We also investigated the quadratic convergence of the new algorithm on Example 1 by setting $p = 2$ in Equation (29).

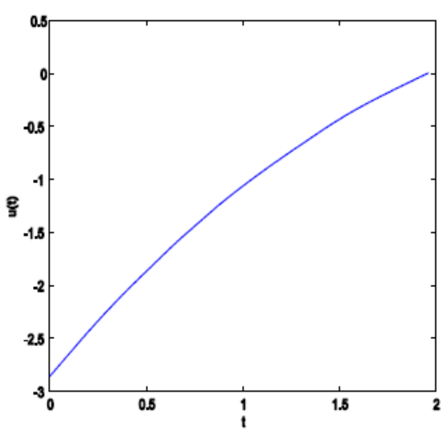

Figure 1: Optimal Control Curve for Example 1.

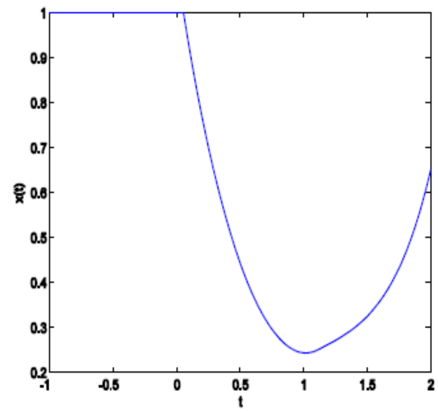

Figure 2: Optimal State Trajectory for Example 1.

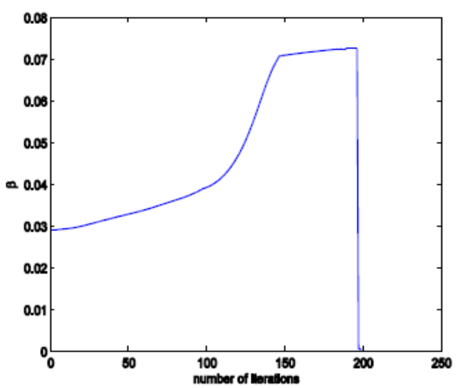

Figure 3: Quadratic Convergence for Example 1.

The graph above shows that $0 \le \beta \le 1$, which shows that the new algorithm exhibits quadratic convergence.

Example 2

$$\textit{Minimize} \qquad J = \int_0^5 (\frac{1}{2}\ 10x_1^2 + x_2^2 + u^2)dt \qquad \textbf{(41)}$$

subject to,

$$\dot{x}_1(t) = x_2(t) \qquad \textbf{(42)}$$

$$\dot{x}_2(t) = -10x_1(t) - 5x_2(t) - 2x_1(t-\tau) - x_2(t-\tau) + u(t) \qquad \textbf{(43)}$$

$$x_1(t) = 1.0 \qquad -\tau \le t \le 0 \qquad \textbf{(44)}$$

$$x_2(t) = 1.0 \qquad -\tau \le t \le 0 \qquad \textbf{(45)}$$

For this system; $\mathbf{P} = \begin{pmatrix} 5 & 0 \\ 0 & 0.5 \end{pmatrix}$, $\mathbf{Q} = (0.5)$, $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -10 & -5 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 0 & 0 \\ -2 & -1 \end{pmatrix}$ and $\mathbf{P} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

The optimality system for optimal control problem(41)-(45) is:

$$\dot{x}_1(t) = x_2(t), \qquad t \in [0,5] \qquad \textbf{(46)}$$

$$\dot{x}_2(t) = -10x_1(t) - 5x_2(t) - 2x_1(t-\tau) - x_2(t-\tau) + u(t) \quad t \in [0,5] \qquad \textbf{(47)}$$

$$\dot{\mu}_1(t) = \begin{cases} 10x_1(t) + 10\mu_2(t) + 2\mu_2(t+\tau) & t \in [0, 5-\tau] \\ 10x_1(t) + 10\mu_2(t) & t \in [5-\tau, 5] \end{cases} \qquad \textbf{(48)}$$

$$\dot{\mu}_2(t) = \begin{cases} x_2(t) - \mu_1(t) + 5\mu_2(t) + \mu_2(t+\tau) & t \in [0, 5-\tau] \\ x_2(t) - \mu_1(t) + 5\mu_2(t) & t \in [0, 5-\tau] \end{cases}. \qquad \textbf{(49)}$$

$$u(t) = 2\mu_2(t) \qquad t \in [0,5] \qquad \textbf{(50)}$$

where $\mu_1(t)$ and $\mu_2(t)$. The optimality system cannot be solved by both analytical method and indirect numerical method due to coupling that exist between the state variables and adjoint variables.

The optimal control problem(41)-(45) is solved for two different values of τ . , namely 0.1 and 1.0 using the new algorithm. We obtained:

$J = 2.5586$ for $\tau = 0.1$ in 65 seconds and
$J = 2.9189$ for $\tau = 1.0$ in 93 seconds,

compared with the evolutionary algorithm which gives $J = 2.5628$ for $\tau = 0.1$ and $J = 2.9277$ for $\tau = 1.0$ both in hours.

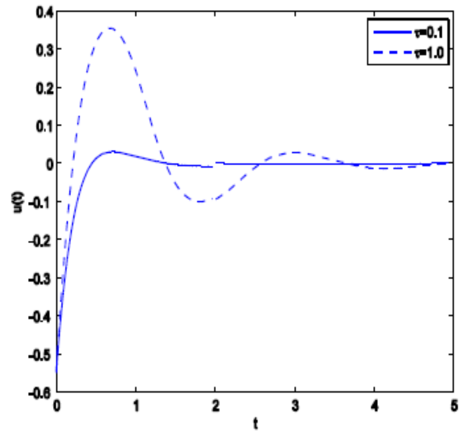

Figure 4: Optimal State Trajectory for Example 2.

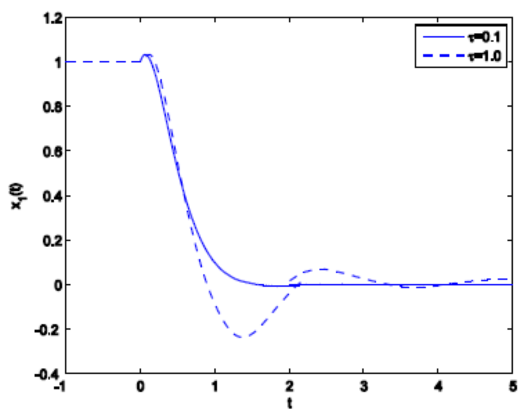

Figure 5: Optimal State Trajectory $(x_1(t))$ for Example 2.

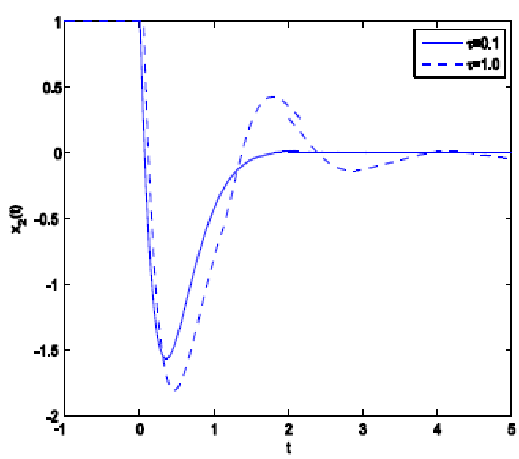

Figure 6: Optimal State Trajectory $(x_2(t))$ for Example 2

CONCLUSION

An algorithm for obtaining optimal control for delay differential equations is obtained in this paper. The result obtained by the new algorithm for optimizing Delay differential equations compares favorably with the Evolutionary algorithm and existing algorithms.The new algorithm exhibits quadratic convergence, which is an advantage over existing methods that inculcate conjugate gradient for determination of near optimal control to the optimization of delay differential equations. Thus, we have shown that conjugate gradient method for solving constrained quadratic programming probm is well suited for solving a certain class of discretized optimal control problems with delay term. Thus, the algorithm is attractive computationally and can be easily extended to nonlinear and time varying systems.

REFERENCES

1. Agrawal, O.P., Gregory, J., and Spector, K.P. 1997. "A Bliss-type Multiplier Rule for Constrained Variational Problems with Time Delay". *J. Math. Anal. and Appl*. 210:702–711

2. Changzhi, W., Kok, L.T., Rui, L., and Zhao, Y. 2006. "Optimal Control of Swithced Systems with Time Delay". *Applied Mathematics Letters*. 19:1062-1067.

3. Dostial, Z. 2009. *Optimal Quadratic Programming Algorithms with Applications to Variational Inequalities*. Springer Science and Business Media: New York, NY.

4. Guo-Ping, C., Jin-Zhi, H., and Simon, X.Y. 2003. "An Optimal Control Method for Linear Systems with Time Delay". *Computers and Structures*. 81:1539-1546.

5. Jaedong, S. 1994. "The Characterization of Optimal Control Using Delay Differential Operator". *Journal of the Chungcheong Mathematical Society*. 7:123-139.

6. Jamshidi, M. and Razzaghi, M. 1975. "Optimization of Linear Systems with Input Time-Delay". Kybernetika. 11(5): 375-384

7. John, B.T. 2001. *Practical Methods for Optimal Control Problem Using Nonlinear Programming*. SIAM: Philadelphia, PA.

8. Olotu, O. and Olorusola, S.A. 2006. "An Algorithm for a Discretized Constrained, Continuous Quadratic Control Problem". *Journal of Applied Sciences*. 9(1):6249-6260.

9. Olotu, O. and Olorusola, S.A. 2008. "Convergence Ratio Profile for Optimal Control Problems Governed by Ordinary Differential Equations with Matrix Coefficients". *American Journal of Applied Sciences*. 5(2):89-92.

10. Olotu, O. and Olorusola, S.A. 2009. "Convergence Profile of a Discretized Scheme for Constrained Problem via Penalty Multiplier Method". *Journal of the Nigerian Association of Mathematical Physics*.14:341-348.

11. Olotu, O. and Adekunle, A.I. 2010. "Analytic and Numeric Solution of Discretized Constrained Optimal Control Problem with Vector and Matrix coefficients". *Advanced Modeling and Optimization*.12:119-131.

12. Rao, S.S.1996. *Engineering Optimization, Theory and Practise.* John Wiley& Sons: New York, NY.

13. Smith, S.F. 2005. "Optimal Control of Delay Differential Equations using Evolutionary Algorithms". *Complexity International*.12:1-10.

14. Wong, K.H., Jennings, L.S. and Benyah, F. 2002. "The Control Parametrization Enhancing Transform for Constrained Time-Delayed Optimal Control Problems". *ANZIAM J*. 43:154-182.

ABOUT THE AUTHORS

Adeshina I. Adekunle, is an Assistant Lecturer at the Department of Mathematical Sciences, Federal University of Technology, Akure. He is a member of the Nigeria Association of Mathematical Physics and International Association of Applied Mathematics and Mechanics (GAMM)[Activity Group: Optimization with PDE Constraints. He holds a Master of Technology (M.Tech.) in Industrial Mathematics from the Federal University of Technology, Akure. His research interests are in optimization techniques, optimal control of differential equations and mathematical biology.

Dr. Olusegun Olotu, is an Associate Professor at the Department of Mathematical Sciences, Federal University of Technology, Akure. He holds a Ph.D. degree in Mathematical Sciences from the University of Ado-Ekiti. His research interests are in the areas of optimization techniques and optimal control of differential equations.

SUGGESTED CITATION

Adekunle, A.I. and O. Olotu. 2012. "An Algorithm for Optimal Control of Delay Differential Equations". *Pacific Journal of Science and Technology*. 13(1):228-237.

Pacific Journal of Science and Technology