Journal of the Nigerian Association of Mathematical Physics
Volume **20 (March, 2012), pp 165 – 174**
© J. of NAMP

A Primal-Dual Augmented Lagrangian Method for Optimal Control of Linear-Quadratic Problem with Time Varying Systems

[1]Adekunle A. I.* and *Olotu O.

Department of Mathematical Sciences,
Federal University of Technology,P.M.B 704,Akure, Nigeria

Abstract

In this paper we are concerned with time-varying optimal control problems whose cost is quadratic and whose state is a differential equation and with general boundary conditions. The basic new idea of this paper is to propose a primal-dual augmented Lagrangian method, embedded with a sequential quadratic programming(SQP) for the solution of such problems.The benefit of this approach is that the quality of the dual variables is monitored explicitly during the solution of the subproblem. Moreover, the formulation of a penalized matrix in the primal-dual variables with mesh-refinement strategy guarantees the reliability of the algorithm. Numerical experiments verify the efficiency of the proposed method.

Keywords: Optimal control,primal-dual methods, augmented Lagrangian methods, conjugate gradient method, sequential quadratic programming.

1.0 Introduction

Linear quadratic(LQ) control problems are very important in optimal control. They have many applications, e.g network theory, stability theory, filtering and estimation. There is an extensive body of literature on the study of the stability and the existence of its controller for both linear time-invariant(LTI) and linear time-variant(LTV) systems[1-3]. The contrast is particularly sharp while there is a few papers on numerical solutions of the controller, especially for time varying systems. The point is how to integrate the nonlinear matrix differential Riccati equation with variable coefficient accurately and efficiently. Since Chen and Shih tried to solve the optimal controller by introducing Walsh function[4] to discretize the continuous LTV system, many researchers extended the methodology to block pulse function and segmental linear functions[5].

Birgit[6] considered time-varying linear systems on Hilbert spaces and studied the optimal control problem with indefinite performance criteria over a finite horizon by applying an operator theoretic approach for the unique solvability of the linear quadratic optimization and for the existence of solutions to the integral Riccati equation. Thus, solving the time-dependent, matrix Riccati differential equation computationally remains eminent. Based on the conservative property of the optimal control system of state space, Tan and Zhong[5] developed a symplectic conservative perturbation algorithm, which circumvents solving the continuous time-dependent matrix Riccati differential equation, for solving linear quadratic control with time-varying systems. Since the algorithm is based on perturbation of the LTV systems, there is need to develop algorithm that will solve the time-dependent matrix Riccati equation directly. Variational iteration method(VIM) was applied to the solution of the general Riccati differential equation by Batiha et al[7]. The approximate analytical method only considered the initial value problem(IVP) of the Riccati differential equation, as to the terminal value problem(TVP) obtained in the optimal control system of state space. Dai and Cochran[8] applied the nonlinear programming(NLP) solver to the nonlinear programming problem obtained by using the Haar wavelet technique to transform the state and control variables into nonlinear programming parameters at collocations points. The algorithm is based on the inexact step-size calculation, which affects the convergence of the method[9]. As a result of the development of control operator, which ensures exact computation of the step size in the line search parameter, for LTI systems[3], Olotu and Adekunle[9] developed a discretized continuous algorithm via quadratic programming through quadratic penalty function method for the solution of optimal control of time-invariant system and delay differential equations[10]. The algorithm is robust and efficient for LTI, but suffers from the inherent ill-conditioning and slow-convergence of the quadratic penalty function method as a nonlinear programming algorithm[11].

A new penalty function method for constrained optimization, which is amenable to both equality and inequality

[1]Corresponding author: ***Adekunle A. I.,*** E-mail: aiadekunle@futa.edu.ng, Tel. +2348060406228

constrained, has been developed by Barry and Dimitri[12]. Also, a primal-dual generalization of the Hestenes-Powell augmented Lagrangian function is discussed extensively as subproblem formulation of nonlinear constrained optimization problem by Gill and Robinson [13]. Since the primal-dual augmented-Lagrangian exhibits fast convergence[11] and has been successfully applied to optimal control of large scale dynamical systems[14], it is of interest incorporating the new penalty function method and the primal-dual augmented Lagrangian in the discretized continuous algorithm via quadratic programming and see how it works. Furthermore, mostly real systems are nonlinear and time dependent, numerical solutions of linear quadratic(LQ) control for time varying systems are very important, which deserve further study.

In this paper, the discretized continuous algorithm via quadratic programming technique[9] is extended to optimal control of time-varying systems. In the proposed algorithm, the optimal control problem is discretized and through the construction of penalized matrix in both the primal and dual variables , 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 examples .

2.0 Method of Solution

Consider the linear time-varying(LTV) system:

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t), \qquad \mathbf{x}(t_0) = \mathbf{x}_0. \tag{2.1}$$

The quadratic performance index is given as

$$J(\mathbf{u}) = \frac{1}{2}\int_{t_0}^{t_f} [\mathbf{x}^T\mathbf{P}(t)\mathbf{x} + \mathbf{u}^T\mathbf{Q}(t)\mathbf{u}]dt, \tag{2.2}$$

where $\mathbf{A}(t)$, $\mathbf{B}(t)$, $\mathbf{P}(t)$, and $\mathbf{Q}(t)$ are $n \times n$, $n \times m$, $n \times n$, and $m \times m$ continuous or piecewise continuous matrix time functions respectively. For $\forall t \in [t_0, t_f]$, $\mathbf{P}(t) \geq 0$, $\mathbf{Q}(t) > 0$, symmetric and the end time t_f is fixed. The state $\mathbf{x}(t)$ and control $\mathbf{u}(t)$ are n −vector and m −vector respectively.

The LQ control problems are to find the control input $\mathbf{u}(t)$ to minimize the performance index (2.2) subjected to the dynamic equation(2.1). First we partition the interval $[t_0, t_f]$ into s sub-intervals with knots $t_0 < t_1 < t_2 \cdots < t_s$ and $t_k = t_0 + k\Delta t_k$, where Δt_k is the mesh size of k^{th} sub-interval . If these sub-intervals are small enough, we can assume that at any collocation point, the values $\mathbf{x}(t)$ and $\mathbf{u}(t)$ can be approximated by zero order spline $\mathbf{x}_k$ and $\mathbf{u}_k$ respectively.

Applying trapezoidal discretization for s grid points to the optimal control problem(2.1) and (2.2), we have

$$\min J(\mathbf{u}) = \sum_{k=0}^{s} (\mathbf{x}_{k+1}^T\mathbf{M}(t_{k+1})\mathbf{x}_{k+1} + \mathbf{x}_k^T\mathbf{M}(t_k)\mathbf{x}_k + \mathbf{u}_{k+1}^T\mathbf{N}(t_{k+1})\mathbf{u}_{k+1} + \mathbf{u}_k^T\mathbf{N}(t_k)\mathbf{u}_k) \tag{2.3}$$

subject to

$$(\mathbf{I_{n\times n}} - \frac{\Delta t_k}{2}\mathbf{A}(t_{k+1}))\mathbf{x}_{k+1} - (\mathbf{I_{n\times n}} + \frac{\Delta t_k}{2}\mathbf{A}(t_k))\mathbf{x}_k - \frac{\Delta t_k}{2}(\mathbf{B}(t_{k+1})\mathbf{u}_{k+1} + \mathbf{B}(t_k)\mathbf{u}_k) = \mathbf{0} \tag{2.4}$$

where $\mathbf{M}(t_k) = \mathbf{P}(t_k)\frac{\Delta t_k}{4}$, $\mathbf{N}(t_k) = \mathbf{Q}(t_k)\frac{\Delta t_k}{4}$, and $\mathbf{I_{n\times n}}$ is $n \times n$ identity matrix .

By parameter optimization[9, 15], the discretized problem becomes a large sparse quadratic programming problem. We give a matrix representation

$$\min J(z) = \mathbf{z}^T\mathbf{D}\mathbf{z} + \mathbf{c} \tag{2.5}$$

subject to

$$\mathbf{E}\mathbf{z} = \mathbf{k} \tag{2.6}$$

and

$$\mathbf{z}^T = (\mathbf{x}_1^T, \mathbf{x}_2^T, \cdots, \mathbf{x}_s^T, \mathbf{u}_0^T, \mathbf{u}_1^T, \mathbf{u}_2^T \cdots, \mathbf{u}_s^T) \tag{2.7}$$

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

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

where i^{th} element corresponds to i^{th} block, and $\mathbf{c} = \mathbf{x}_k^T(0)\mathbf{M}(t_0)\mathbf{x}_k(0)$. The matrix $\mathbf{E}$ is block matrix of order $ns \times (n+m)s + m$ with the representation

$$\mathbf{E} = (\mathbf{G} \;\; \vdots \;\; \mathbf{H}), \tag{2.8}$$

where $\mathbf{G}$ is an $ns \times ns$ block bidiagonal matrix with principal block diagonal elements $[\mathbf{G}_{ii}] = \mathbf{I_{n\times n}} - \frac{\Delta t_k}{2}\mathbf{A}(t_i)$ and lower block principal diagonal elements $[\mathbf{G}_{ij}] = -\left(\mathbf{I_{n\times n}} + \frac{\Delta t_k}{2}\mathbf{A}(t_{i-1})\right)$, $\forall i,j$ block such that $i = j+1$.The matrix $\mathbf{H}$ is an $ns \times (s+1)m$ block bidiagonal matrix with principal block diagonal elements $[\mathbf{H}]_{ii} = -\frac{\Delta t_k}{2}\mathbf{B}(t_{i-1})$ and upper block principal diagonal elements $[\mathbf{H}_{ij}] = -\frac{\Delta t_k}{2}\mathbf{B}(t_j)$, $\forall i,j$ block such that $j = i+1$. The column vector $\mathbf{k}$ is of order $ns \times 1$ with entries given by; $[\mathbf{k}]_{1:n,1} = (\mathbf{I_{n\times n}} + \frac{\Delta t_k}{2}\mathbf{A}(t_0))\mathbf{x}_0$ and $[\mathbf{k}_{i,1}] = 0, i = n+1, n+2, \cdots, ns$.

Given $\mathbf{y} \in \Re^{ns}$, $\mathbf{y}_\lambda \in \Re^{ns}$, and $\rho > 0$, then by the primal-dual augmented Lagrangian method, the unconstrained minimization problem of the discretized optimal control problem(2.5) and (2.6) is

$$\min_{\mathbf{z},\mathbf{y}} P(\mathbf{z},\mathbf{y};\mathbf{y}_\lambda,\rho,\varepsilon) = \mathbf{z}^T\mathbf{D}\mathbf{z} + \mathbf{c} + \phi_\varepsilon[(\mathbf{E}\mathbf{z}-\mathbf{k})_i]^T\mathbf{y}_\lambda + \frac{1}{2\rho}\|\mathbf{E}\mathbf{z}-\mathbf{k}\|^2 + \frac{1}{2\rho}\|\mathbf{E}\mathbf{z}-\mathbf{k}+\rho(\mathbf{y}-\mathbf{y}_\lambda)\|^2, \tag{2.9}$$

where ϕ_ε is a real valued function of a single variable, depending on a positive parameter ε satisfying the following properties[12]

$(P1)$: $\phi(0) = 0$,
$(P2)$: $\phi'(0) = 0$,
$(P3)$: $\lim_{t\to\infty}\phi'(t) = +\infty$,
$(P4)$: $\lim_{t\to-\infty}\phi'(t) = -\infty$,
$(P5)$: $\phi''(t) > 0 \quad \forall t \in \Re\backslash\{0\}$,
$(P6)$: $\phi''(0) = 1 \quad \forall t \in \Re$.

Examples of such functions are $\phi(t) = \frac{t^2}{2}, \frac{e^t+e^{-t}-2}{2}$, and $\frac{1}{2}(\frac{t^2}{2}+e^t-t)$. For $\varepsilon > 0$ define $\phi_\varepsilon(t) = \varepsilon\phi(\frac{t}{\varepsilon})$.

On expansion, we have

$$\min_{\mathbf{z},\mathbf{y}} P_\rho(\mathbf{z},\mathbf{y}) = \mathbf{z}^T\mathbf{A}_\rho\mathbf{z} + \rho\mathbf{y}^T\mathbf{I}_{ns\times ns}\mathbf{y} + \mathbf{B}_\rho\mathbf{z} + \mathbf{B}_\rho^*\mathbf{y} + c \tag{2.10}$$

Equation (2.10) is the quadratic form representation in both $\mathbf{z}$ and $\mathbf{y}$ for the unconstrained minimization problem(2.9), where $P_\rho(\mathbf{z},\mathbf{y})$ is primal-dual augmented Lagrangian, ρ is penalty parameter, the penalized matrix $\mathbf{A}_\rho = \left[\mathbf{D} + \frac{1}{\rho}\mathbf{E}^T\mathbf{E}\right]$, $\mathbf{B}_\rho = \mathbf{y}^T\mathbf{E} - \frac{2}{\rho}\mathbf{k}^T\mathbf{E} - \mathbf{y}_\lambda^T\mathbf{E}$, and $\mathbf{B}_\rho^* = -\rho\mathbf{y}_\lambda^T - \mathbf{k}^T$ and $c = \mathbf{c} + \phi_\varepsilon^T\mathbf{y}_\lambda + \frac{1}{\rho}\mathbf{k}^T\mathbf{k} + (\mathbf{k}^T + 2\rho\mathbf{y}_\lambda^T)\mathbf{y}_\lambda$. Let the ns-vector $\pi(\mathbf{z}) = \mathbf{y}_\lambda - \frac{1}{\rho}(\mathbf{E}\mathbf{z}-\mathbf{k})$, then the gradient and Hessian for $P(\mathbf{z},\mathbf{y})$ may be written as

$$\nabla P(\mathbf{z},\mathbf{y}) = \begin{pmatrix} 2\mathbf{D}\mathbf{z} + \mathbf{y}_\lambda^T\phi_\varepsilon'[(\mathbf{E}\mathbf{z}-\mathbf{k})_i]\nabla(\mathbf{E}\mathbf{z}-\mathbf{k})_i + \mathbf{J}^T(\mathbf{y}-\mathbf{y}_\lambda) \\ \rho(\mathbf{y}-\pi(\mathbf{z})) \end{pmatrix}, \tag{2.11}$$

$$\nabla^2 P(\mathbf{z},\mathbf{y}) = \begin{pmatrix} 2\mathbf{D} + \mathbf{y}_\lambda^T\phi_\varepsilon''\nabla(\mathbf{E}\mathbf{z}-\mathbf{k})_i + \mathbf{y}_\lambda^T\phi_\varepsilon'\nabla^2(\mathbf{E}\mathbf{z}-\mathbf{k})_i + \mathbf{H}(\mathbf{y}-\mathbf{y}_\lambda) & \mathbf{J}^T \\ \mathbf{J} & \rho\mathbf{I}_{ns\times ns} \end{pmatrix}, \tag{2.12}$$

where $\mathbf{J}(\mathbf{z})$ and $\mathbf{H}(\mathbf{z})$ is the Jacobian and Hessian matrices of $c(\mathbf{z}) = \mathbf{E}\mathbf{z} - \mathbf{k}$ with appropriate dimensions. Observe that the first-order multipliers $\pi(\mathbf{z}) = \mathbf{y}_\lambda - \frac{1}{\rho}(\mathbf{E}\mathbf{z}-\mathbf{k})$ minimize $P(\mathbf{z},\mathbf{y})$ with respect to $\mathbf{y}$ for fixed values of $\mathbf{z}$.

Theorem 2.1 *Assume that $(\mathbf{z}^*,\mathbf{y}^*)$ satisfies the following conditions associated with problem(2.10):*

(i) $c(\mathbf{z}^*) = 0$,
(ii) $\mathbf{D}\mathbf{z}^* = 0$,
(iii) there exists a positive scalar ω such that $p^T H(\mathbf{z}^*,\mathbf{y}^*)p \geq \omega\|p\|^2$ for all p satisfying $J(\mathbf{z}^*)p = 0$.

Then $(\mathbf{z}^*,\mathbf{y}^*)$ is a stationary point of the primal-dual function

$$P(\mathbf{z},\mathbf{y};\mathbf{y}^*,\rho,\varepsilon) = \mathbf{z}^T\mathbf{D}\mathbf{z} + \mathbf{c} + \phi_\varepsilon[(\mathbf{E}\mathbf{z}-\mathbf{k})_i]^T\mathbf{y}^* + \frac{1}{2\rho}\|\mathbf{E}\mathbf{z}-\mathbf{k}\|^2 + \frac{1}{2\rho}\|\mathbf{E}\mathbf{z}-\mathbf{k}+\rho(\mathbf{y}-\mathbf{y}^*)\|^2.$$

Proof. We must show that ∇P is zero and $\nabla^2 P$ is positive definite at the primal-dual point $(\mathbf{z},\mathbf{y}) = (\mathbf{z}^*,\mathbf{y}^*)$. Assumption (i) and the definition $\pi(\mathbf{z}) = \mathbf{y}^* - \frac{1}{\rho}(\mathbf{E}\mathbf{z}-\mathbf{k})$ implies that $\pi(\mathbf{z}^*) = \mathbf{y}^*$. Substituting for $\pi, \mathbf{z}$ and $\mathbf{y}$ in the gradient(2.11) and using the assumptions $P2$ and (ii), gives $\nabla P(\mathbf{z}^*,\mathbf{y}^*;\mathbf{y}^*,\rho)$=0 directly.

Similarly, the Hessian(2.12) becomes

$$\nabla^2 P(\mathbf{z}^*,\mathbf{y}^*) = \begin{pmatrix} \mathbf{D} & 0 \\ 0 & \rho\mathbf{I}_{ns\times ns} \end{pmatrix}. \tag{2.13}$$

It is sufficient to show that $\mathbf{D}$ is positive definite. By definition it is easily seen that $\mathbf{D}$ is positive definite and hence $\nabla^2 P$ is positive definite.

Theorem(2.1) indicates that if an estimate of $\mathbf{y}^*$ is known for problem(2.9), then an approximate minimization of P with respect to both $\mathbf{z}$ and $\mathbf{y}$ is likely to provide an even better estimate.

Lemma 2.1 ([9]) *Let $\boldsymbol{D}(t_k) \in R^{((n+m)s+m)\times((n+m)s+m)}$ be a symmetric positive definite matrix, let $\boldsymbol{E}(t_k) \in R^{(ns)\times((n+m)s+m)}$, $\rho > 0$, and let $Ker\boldsymbol{D} \cap Ker\boldsymbol{E} = \mathbf{0}$. Then the penalized matrix $\boldsymbol{A}_\rho(t_k)$ is positive definite.*

Lemma 2.2 ([9]) *Let $\boldsymbol{D}(t_k) \in R^{((n+m)s+m)\times((n+m)s+m)}$ be a symmetric positive definite matrix, let $\boldsymbol{E}(t_k) \in R^{(ns)\times((n+m)s+m)}$, $\mu > 0$ 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 sufficient condition for $\mathbf{z}^*(t_k) \in R^{((n+m)s+m)}$ to a be local minimum point. Hence, the sequential minimization problem(2.10) for a fixed set values of $\mathbf{y_\lambda}$ and a sequence of values $\varepsilon_k \to 0$ behaves like the classical Hestenes-Powell augmented Lagrangian method. Also for $\rho_k \to 0$, the discretized problem(2.10) yields the primal-dual quadratic penalty function method . Thus, when $\rho_k \to 0$ and $\varepsilon_k \to 0$, we have the primal-dual augmented Lagrangian method. Then our formulation is a combination of three robust and efficient techniques.

Our aim is to solve the unconstrained minimization equation(2.10) by conjugate gradient algorithm in the inner loop and enforce the feasibility condition in the outer loop in both the primal and the dual variables and reduce the discretization error.

3.O Algorithm of The Scheme

In this section, we are proposing an algorithm on the basis of the above discussions. This algorithm is based on the sequential minimization in both the primal and the dual variables using the conjugate gradient algorithm[3, 9], as stated below:

Algorithm 4.1: A Primal-Dual Algorithm for Constrained Optimal Control Problem

Step 1. Set $k = 0$ and choose $\varepsilon_0 > 0$, $\rho_0 \in (0,1)$, and $c > 1$. Initialize $\mathbf{y}_{(\lambda,0)}$, $\mathbf{y}_{0,0} \in \Re^{ns}$ and $\mathbf{z}_{0,0} \in \Re^{(n+m)s+m}$. With $\mathbf{y}$ fixed.

Step 2. Set $i = 0$ and set $p_0 = -g_0 = -\nabla P_{\rho_k,\varepsilon_k}(\mathbf{z}_{0,0})$.

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

Step 4. Set $\mathbf{z}_{k,i+1} = \mathbf{z}_{k,i} + \alpha_i p_i$

Step 5. Compute $\nabla P_{\rho_k,\varepsilon_k}(\mathbf{z}_{k,i+1})$

Step 6. If $\nabla P_{\rho_k,\varepsilon_k}(\mathbf{z}_{k,i+1}) = 0$ and $\mathbf{E}\mathbf{z}_{k,i+1} = \mathbf{k}$, go to step 8b ;else go to step 7.

Step 7a. If $\nabla P_{\rho_k,\varepsilon_k}(\mathbf{z}_{k,i+1}) \neq 0$, set

$g_{i+1} = \nabla P_{\rho_k,\varepsilon_k}(\mathbf{z}_{k,i+1})$,

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

Step 7b. Set $i = i + 1$, and go to step 3.

Step 8a. Else if $\mathbf{E}\mathbf{z}_{k,i+1} \neq \mathbf{k}$,

Step 8b. With $\mathbf{z}_{k,i}$ fixed, set $j = 0$ and set $p_0^* = -g_0^* = -\nabla P_{\rho_k,\varepsilon_k}(\mathbf{y}_{0,0})$.

Step 8c. Compute $\alpha_j^* = \frac{{p_j^*}^T p_j^*}{{p_j^*}^T \mathbf{I}_\rho p_j^*}$

Step 8d. Set $\mathbf{y}_{k,j+1} = \mathbf{y}_{k,j} + \alpha_j^* p_j$

Step 8e. Compute $\nabla P_{\rho_k,\varepsilon_k}(\mathbf{y}_{k,j+1})$

Step 8f. If $\nabla P_{\rho_k,\varepsilon_k}(\mathbf{y}_{k,j+1}) \neq 0$, set

$g_{j+1}^* = \nabla P_{\rho_k,\varepsilon_k}(\mathbf{y}_{k,j+1})$,

$p_{j+1}^* = -g_{j+1}^* + \gamma_j^* p_j$, with $\gamma_j^* = \frac{{g_{j+1}^*}^T g_{j+1}^*}{{g_j^*}^T g_j^*}$

Step 8g. Set $j = j + 1$, and go to step 8c.

Step 8h. Else if $\nabla P_{\rho_k,\varepsilon_k}(\mathbf{y}_{k,j+1}) = 0$ and $\mathbf{E}\mathbf{z}_{k,i+1} = \mathbf{k}$, End.

Step 9. $\rho_{k+1} = c\rho_k$, $\varepsilon_{k+1} = \frac{\varepsilon_k}{c}$, $\mathbf{y}_{\lambda,k+1} = \mathbf{y}_{\lambda,k}(\phi_\varepsilon'[(\mathbf{Ez} - \mathbf{k})_i] - 1) + \frac{1}{\rho_k}(\mathbf{E}\mathbf{z}_{k,i} - \mathbf{k}) + \mathbf{y}_{k,j}$ set $k = k + 1$ and go to step 2.

The Algorithm(4.1) exhibits at least Q-linear convergence if $\mathbf{y}_{\lambda,k}$ is bounded and super-linear convergence if $\mathbf{y}_{\lambda,k}$ is unbounded ([11],pg.118), as it inherits the convergence properties of the augmented Lagrangian for fixed $\mathbf{y}$.

Since we aim at reducing the discretization error, adding large number of points, increases the size of the quadratic programming problem to be solved,and thereby causing a significant computational penalty. Hence, we introduce grid-refinement strategy [15] to assess the accuracy of the proposed method.

Suppose $\mathbf{z}^*(t) = (\mathbf{x}(t), \mathbf{u}(t))^T$ be the optimal solution of Algorithm(4.1), Then $\mathbf{z}^*(t)$ can be approximated as follows;

$$\begin{aligned} \mathbf{x}(t) &\approx \tilde{\mathbf{x}}(t) = \textstyle\sum_{i=0}^{2s} \gamma_i D_i(t) \\ \mathbf{u}(t) &\approx \tilde{\mathbf{u}}(t) = \textstyle\sum_{i=0}^{s} \beta_i C_i(t) \end{aligned} \qquad (3.1)$$

where the functions $D_i(t)$ form the basis for C^1 cubic B-splines. The coefficients γ_i in the state variable representation are uniquely defined by Hermite interpolation of the discrete solution. The functions $C_i(t)$ form the basis for C^0 piece-wise linear B-splines, and the coefficients β_i in the control variable representation are uniquely defined by interpolation of the discrete solution. Let

$$\varepsilon_k \approx \|\mathbf{c}_k h^3\|, \tag{3.2}$$

be the local error for the trapezoidal discretization at step k, where the coefficients $\mathbf{c}_k$ depends on the partial derivatives of the right-hand side of equation(2.1). Consider a single interval $t_k \leq t \leq t_k + h_k$, the absolute error is defined as,

$$\boldsymbol{\eta}_{i,k} = \int_{t_k}^{t_k+1} |\psi(v)|dv, \tag{3.3}$$

where

$$\boldsymbol{\psi}(t) = \dot{\tilde{\mathbf{x}}}(t) - \mathbf{A}(t)\tilde{\mathbf{x}}(t) - \mathbf{B}(t)\tilde{\mathbf{u}}(t) \tag{3.4}$$

defines the error in the differential equation as a function of t. Thus, the relative local error is defined by

$$\varepsilon_k \approx \max_i \frac{\boldsymbol{\eta}_{i,k}}{w_i+1}, \tag{3.5}$$

where the scale weight

$$w_i = \max_k^s[\tilde{x}_{i,k}, \dot{\tilde{x}}_{i,k}] \tag{3.6}$$

defines the maximum value for the ith state variable or its derivative over the s grid points in the phase. By equating (3.2) with (3.5), we obtain

$$\|\mathbf{c}_k\| = \max_i \frac{\eta_{i,k}}{w_i+1} h^3. \tag{3.7}$$

Let the integer I_k be the number of points to add to interval k, so that,

$$\varepsilon_k \approx \|c_k\| \left(\frac{h}{1+I_k}\right)^3 = \max_i \frac{\eta_{i,k}}{w_i+1} \left(\frac{1}{1+I_k}\right)^3, \tag{3.8}$$

then the new mesh can be constructed by choosing the set of integers I_k to minimize

$$\phi(I_k) = \max_k \varepsilon_k, \tag{3.9}$$

and satisfy the constraints

$$\sum_{k=1}^{n_s} I_k \leq s - 1, \tag{3.10}$$

and

$$I_k \leq s_1, \tag{3.11}$$

where equation(3.9)-(3.11) define a nonlinear integer programming problem.

Hence, if we set the desired discretization error tolerance as δ, then the mesh refinement algorithm is proposed as follows;

Algorithm (4.2): Mesh-Refinement
Step 1: Set δ, $\kappa \in (0,1)$ and $s_1 = 4$. Compute the cubic spline representation (3.1) from the discrete solution $\mathbf{z}^*$.
Step 2: Compute an estimate for the discretization error ε_k in each segment of the current mesh using (3.3) and the average error $\bar{\varepsilon}$.
Step 3: If the error is equi-distributed, subdivide each interval of the current mesh and terminate.
Step 4: Else construct a new mesh as follows

(a) Compute the interval α with maximum error i.e;

$$\varepsilon_k = \max_k \varepsilon_k. \tag{3.12}$$

(b) Terminate if

- $s_1{}'$ points have been added $s_1{}' \geq \min[s_1, \kappa s]$
- the error is within tolerance: $\varepsilon_\alpha \leq \delta$ and $I_\alpha = 0$ or
- the predicted error is safely within tolerance: $\varepsilon_\alpha \leq \kappa\delta$ and $0 < I_\alpha < s_1$ or
- $s - 1$ points have been added or
- s_1 points have been added to a single interval.

(c) Add a point to interval α, i.e $I_\alpha = I_\alpha + 1$.

(d) Update the predicted error for interval α using (3.8) and go to step 5a.

4.O Numerical Examples

In order to illustrate the effectiveness of the proposed primal-dual augmented Lagrangian method, two examples for LQ control of time-varying system are used for simulation and comparisons. All simulation 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(R)Pentium(R) Dual processor running at 1.87GHz.

Example 4.1 *Consider a one-dimensional linear time-varying system referring to [5]*

$$Minimize \quad J = \frac{1}{2}\int_0^1 (9tx^2(t) + u^2(t))\,dt \tag{4.1}$$

subject to

$$\dot{x}(t) = 6\sqrt{t}u(t), \quad x(0) = 1.0, \quad t \in [0, t_f]. \tag{4.2}$$

The Riccati equation and the state feedback equation of the system are;

$$C(t) = \frac{1}{2}\left(\frac{1-e^{-18(t_f^2-t^2)}}{1+e^{-18(t_f^2-t^2)}}\right), \quad x(t) = \frac{1-e^{-18(t_f^2-t^2)}}{1+e^{-18(t_f^2)}}e^{-9t^2}[5]. \tag{4.3}$$

The solution $C(t)$ contains the fast decaying component $(e^{-18(t_f^2-t^2)})$ and slow decaying component (constant value 1), so that it is stiff and can be used to test stability and precision of different numerical algorithms.

The numerical solution $x(t)$ is compared in the interval $0.8 < t < 1$, in which $x(t)$ varies rapidly. The function $\phi(t)$ chosen was $\frac{t^2}{2}$. The Table below gives the comparison for the state feedback equation.

Table 1: Comparison of solution of state $x(t)(\Delta t = 0.05)$

t	0	0.80	0.85	0.90	0.95	1.0
Analytical solution	1.0	0.0031559	0.0015098	0.0007046	0.0003481	0.00024681
Proposed Algorithm	1.0	0.0031558	0.0015097	0.0007096	0.0003481	0.00024681
Sympl. Consv. Approx.	1.0	0.0031557	0.0015097	0.0007046	0.0003480	0.00024680
4th Runge-Kutta	1.0	0.0031788	0.0015248	0.0007090	0.0003314	0.00017988

Remark 4.1 *Table 1 demonstrates the effect of the proposed algorithm on the low-order trapezoidal approximation of the state equation. The precision is six valid numbers compared to the 4th order Runge-Kutta method which has 2 valid numbers and compared favorably with the symplectic conservative method, which requires the solution of $(s + 1)$ partial differential equations to obtain the discrete interval energy matrices, with embedded 4th Runge-Kutta and discrete forward-backward algorithm to obtain the state and adjoint solution respectively. Unlike the symplectic conservative method, the proposed primal-dual algorithm takes computation of the objective function into consideration and sensitivity analysis can be implemented on the constrain equation(discretized state equation) to design efficient control system.*

The Figures 1 and 2 show the variation of the primal variable($x(t)$), dual variable($y(t)$) and the control variable($u(t)$) with time.

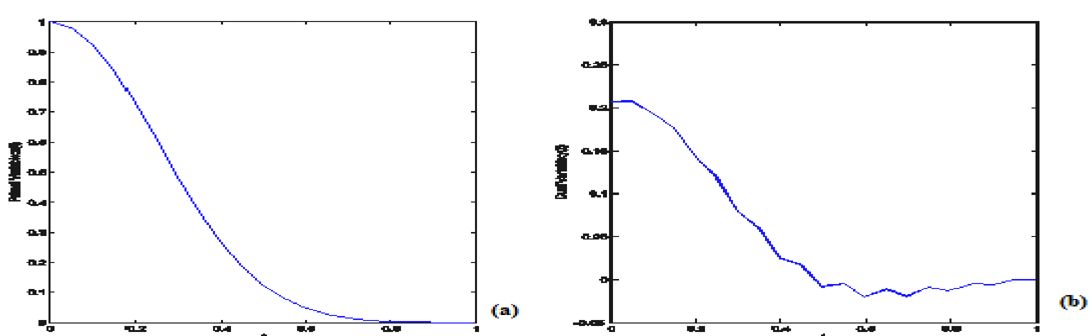

Figure 1: (a) Variation of primal trajectory and (b) variation of dual trajectory, with time for example 4.1

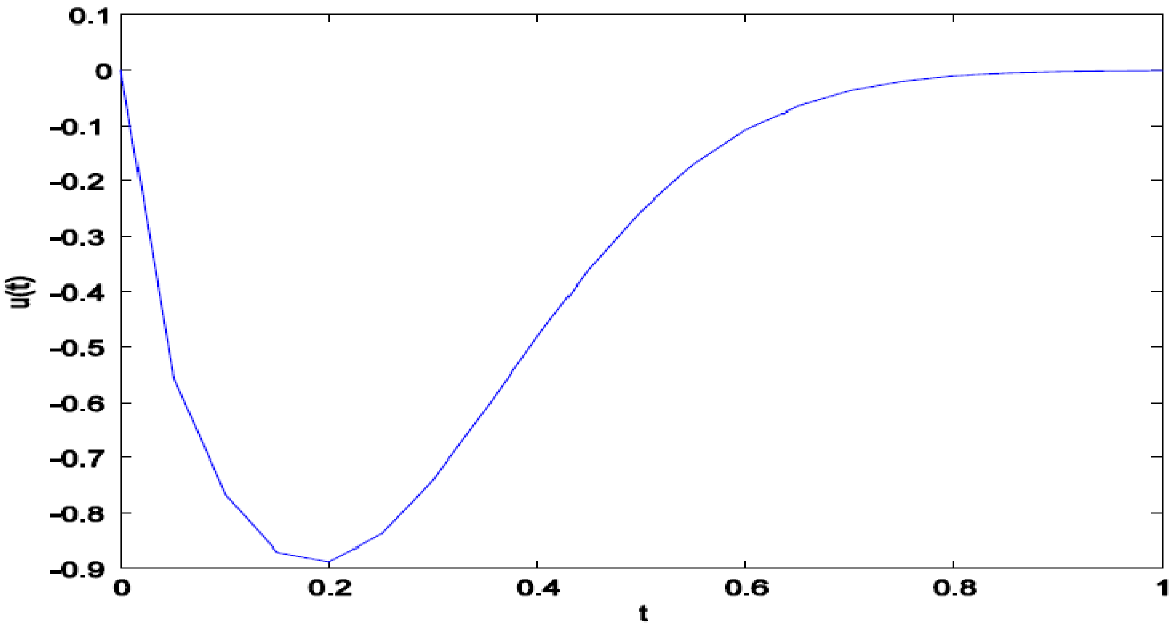

Figure 2: Variation of control trajectory for example(4.1)

Table 2: Computational results with the proposed algorithm as a particular algorithm

	$\varepsilon_k = 0.1^k, \rho_k = 1 \times 10^{-5}$		$\varepsilon_k = 1 \times 10^{-5}, \rho_k = 0.1^k$		$\varepsilon_k = 0.1^k, \rho_k = 0.1^k$	
	Hestenes-Powell Type		**Primal-dual quad.**		**Proposed method**	
K	**Constrain satisfaction**	$J(u)$	**Constrain satisfaction**	$J(u)$	**Constrain satisfaction**	$J(u)$
1	2.1968×10^{-2}	1.9985	1.0828×10^{-5}	0.1040	2.3209×10^{-5}	0.1036
2	3.4009×10^{-3}	1.9280	6.8143×10^{-7}	0.1034	3.84994×10^{-6}	0.1036
3	4.0201×10^{-4}	1.7669	4.7005×10^{-8}	0.4765	4.0712×10^{-7}	0.2084
4	4.3317×10^{-5}	1.0512	4.4285×10^{-9}	0.3916	4.3575×10^{-8}	0.2606
5	4.4877×10^{-6}	0.8071	4.4478×10^{-10}	0.3146	4.5001×10^{-9}	0.2858
6	4.5657×10^{-7}	0.6077	4.5530×10^{-11}	0.3105	4.5724×10^{-10}	0.2983
7	4.6048×10^{-8}	0.5290	4.6056×10^{-12}	0.3102	4.3361×10^{-11}	0.3045
8	4.6243×10^{-9}	0.3073	4.6289×10^{-13}	0.3123	4.4461×10^{-12}	0.3076
9	4.6341×10^{-10}	0.3090	4.6000×10^{-14}	0.3086	4.7000×10^{-14}	0.3086
10	4.6378×10^{-11}	0.3098				
11	4.7393×10^{-12}	0.3103				
12	4.0878×10^{-13}	0.3119				
13	4.4000×10^{-14}	0.3203				

Remark 4.2 *The Table 2 below shows the computational result of the proposed method as a combination of three optimization techniques. The proposed primal-dual augmented Lagrangian method gives better convergence and objective function value compared to the Hestenes-Powell augmented Lagrangian like method and the primal-dual quadratic penalty method. It is easily seen that the number of minimization required for all the methods decrease with increase in the penalty parameter* ρ_k. *However, the effects of ill-conditioning are felt more under these circumstances when the unconstrained minimization is carried out with very small* $\rho_k \geq 10^{-15}$.

Example 4.2 *Consider a 2-dimensional LTV system which is a model of aerospace trajectory control[5]. The system data is given as follows:*

$$\mathbf{A}(t) = \begin{bmatrix} 2t & 1 \\ 0 & t+1 \end{bmatrix}, \mathbf{B}(t) = \begin{bmatrix} 1 \\ \frac{(2t+2)}{2t+3} \end{bmatrix}, \mathbf{P}(t) = \mathbf{I}, \quad Q(t) = 1.0; \quad t \in [0,4.0].$$

And the initial conditions are $\mathbf{x}(0) = (1,1)^T$.

It is difficult to obtain analytical solution.Figures 3 and 4 show the variation of state variables and control variable for Example 4.2 as obtained by the primal-dual augmented Lagrangian method at computing interval ($\Delta t = 0.2$). The conclusion coincides with that of Example 4.1, which illustrates the effectiveness of the proposed method.

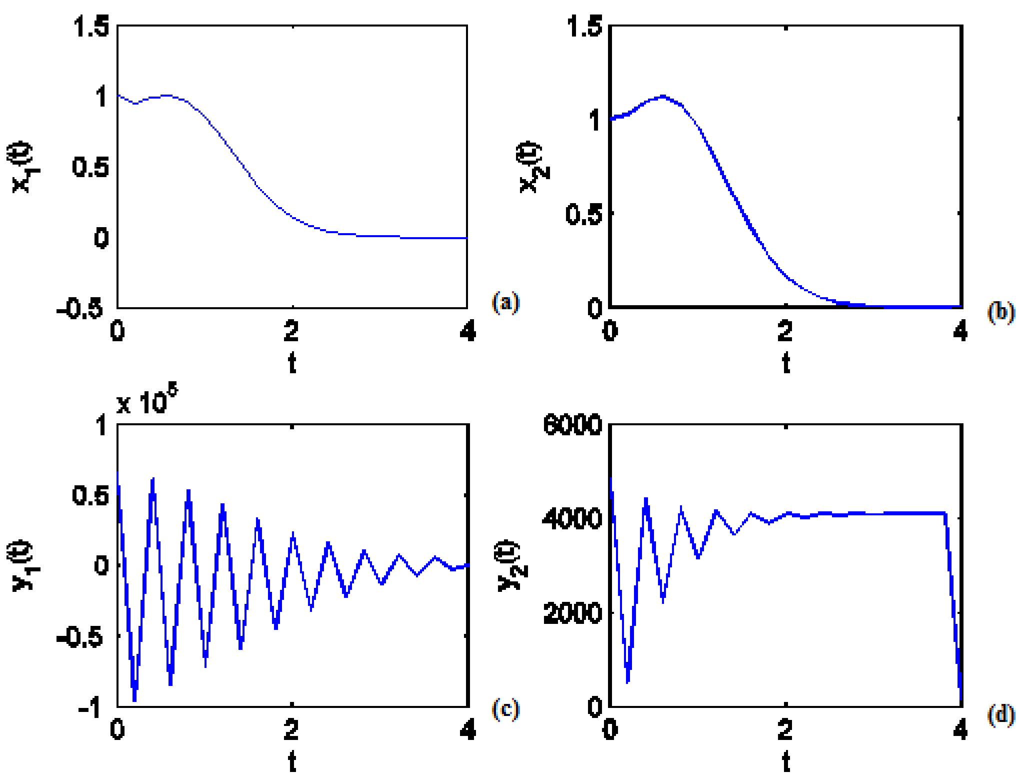

Figure 3: (a)Variation of optimal state trajectory $x_1(t)$,(b) variation of optimal state trajectory $x_2(t)$, (c) variation of dual variable $y_1(t)$ and (d) variation of dual variable $y_2(t)$, with time for example 4.2

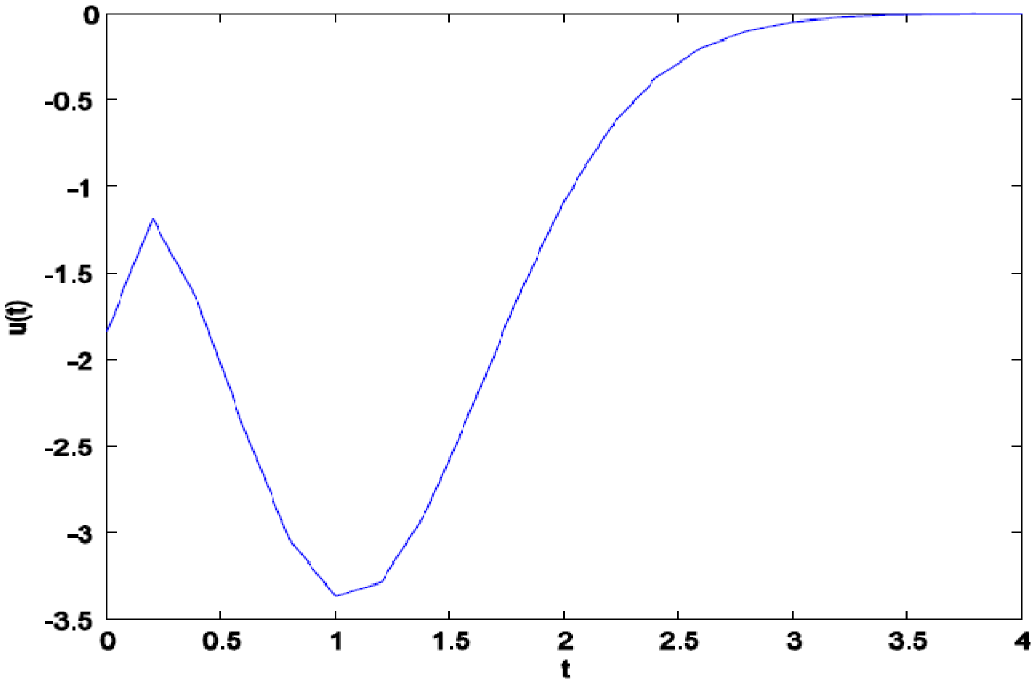

Figure 4: variation of optimal control trajectory with time for example 4.2

5.O Conclusions

This paper presents the methodology of primal-dual augmented Lagrangian method. It gives a uniform way to solve the computational problems of Linear-Quadratic control for time varying systems, which is based on sequential minimization with respect to both primal and dual variables. The result obtained by the new algorithm for optimizing time-varying systems compares favorably with the symplectic conservative perturbation method. It has the advantage of reduced computational efforts and sensitivity analysis can be carried on the system design to obtain preferable objective function value. Thus, we have shown that conjugate gradient method for solving constrained quadratic programming problem is well suited for solving a certain class of discretized optimal control problems with time-varying system. Hence, the algorithm is attractive computationally and can be easily extended to nonlinear optimal control and deserves a further study.

References

[1] Anderson B. D. O. and Moore J. B,(1990). Optimal Control: Quadratic Methods. Englewood Cliffs, N J: Prentice Hall.
[2] Chen C. T,(1984). Linear System Theory and Design. New York: CBS College, 1984.
[3] Ibeijugba M A and Onumanyi P, (1984). A Control Operator and Some of its Applications. Journal of Mathematical Analysis and Application, **103**:31-47.
[4] Chen. W. L, and Shih Y P, (1978). Analysis and Optimal Control of Time-Varying Linear Systems via Walsh Functions. Int J Control, **27**(6):917-932.
[5] Tan S and Zhong W,(2007). Numerical Solution of Linear Quadratic Control for Time-Varying Systems via Symplectic Conservative Perturbation. Applied Mathematics and Mechanics(English Edition), **28**: 227-287
[6] Birgit J, (1995). Linear Quadratic Optimal Control of Time Varying Systems with Indefinite Costs on Hilbert Spaces: The Finite Horizon Problem.Journal of Mathematical Systems,Estimation and Control, **6**(1):1-28.
[7] Batiha B, Noorani M. S. M and Hashinm I, (2007). Application of Variational Iteration Method to a General Riccati Differential Equation. International Mathematics Forum, **2**(56):2759-2770
[8] Dai R and Cochran J E, Jr,(2009). Wavelet Collocation Method for Optimal Control Problems. J. Optim. Theory, **143**:265-278

[9] 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,**Vol. 12**:119-131.

[10] Adekunle A. I. (2011).Algorithms for a Class of Discretized Constrained Optimal Control Problems. M.Tech Thesis, Federal University of Technology, Akure.

[11] Dimitri P B, (1996). Constrained Optimization and Lagrange multiplier Methods.Dimitri P. Athena Scientific,Isbn 1-886529-04-3,Belmont

[12] Barry W K and Dimitri P B,(2010). A New Penalty Function Method for Constrained Minimization. Decision and Control,Issues 13-15, Current Version:162-166.

[13] Gil P E and Robinson D P, (2008). A Primal-Dual Augmented Lagrangian. UCSD Department of Mathematics,Technical Report, NA-08-02.

[14] Gil P E, Jay L O,Leonard M W, Petzold L R, and Sharma V, (2000). An SQP Method for the Optimal Control of Large-Scale Dynamical Systems. Journal of Computational and Applied Mathematics, **Vol. 120**. Issues 1-2:197-213.

[15] John B T,(2001). Practical Methods for Optimal Control Problem Using Nonlinear Programming. SIAM, Philadelphia.