PENALTY FUNCTION METHOD FOR OPTIMIZING CONSTRAINED PROPORTIONAL CONTROL PROBLEMS

Olusegun Olotu
Department of Mathematical Sciences
Federal University of Technology, Akure
Nigeria
oolotu@futa.edu.ng

Emmanuel Jesuyon Dansu*
Department of Mathematical Sciences
Federal University of Technology, Akure
Nigeria
ejdansu@futa.edu.ng

Abstract
The general continuous proportional control problems constrained by ordinary differential equations are considered. We discuss the optimization of proportional control problems with equality constraints. The quadratic penalty function method is used to convert the constrained problems into unconstrained problems. Discretization of the objective functions and the constraints is carried out using the Composite Simpson's Rule and the Fourth-Order Adams-Moulton Technique respectively. The new formulation gives rise to the construction of an operator amenable to the application of the Conjugate Gradient Method (CGM). Analysis of the convergence of the solutions is carried out and the results are found to compare favourably with those obtained from existing algorithms.

Key words
Penalty function method, proportional control, conjugate gradient method, discretization, convergence.

1 Introduction

Given a mathematical programming problem, [Bellmore, Greenberg and Jarvis, 1968] constructed an alternate problem with its feasibility region a superset of the original mathematical problem. The objective function of the problem was constructed so that a penalty is imposed for solutions outside the original feasibility region. An attempt was made to choose an objective function that makes the optimal solution to the new problem the same as the optimal solution to the original mathematical programming problem. Control theory is certainly, at present, one of the most interdisciplinary areas of research and arises in the very first technological discoveries of the industrial revolution as well as the most modern technological applications. On the other hand, Control theory has been a discipline where many mathematical ideas and methods have melted to produce a new body of important Mathematics. Accordingly, it is nowadays a rich crossing point of Engineering and other Sciences with Mathematics [Fernández-Cara and Zuazua, 1979]. The work presented by [Schwartz, 1996] was based on discretizing optimal control problems using explicit, fixed step-size Runge-Kutta integration techniques. The advantage of this scheme over collocation schemes is that the approximating problems that result can be solved very efficiently and accurately. According to [Olotu and Olorunsola, 2006], many earlier schemes, particularly the Function Space Algorithm (FSA) which sidetracks the knowledge of operator for solving quadratic optimal control problems, have been computationally involving and iteratively high. In their research, a new scheme, Discretized Continuous Algorithm (DCA), was proposed with developed associated operator consisting of a series of summation replacing the integrals of the earlier schemes, thus enhancing much more feasible results and lower iterations. [Adekunle and Olotu, 2012] dealt with optimal control problems whose cost is quadratic and whose state is governed by linear delay differential equations and general boundary conditions. The basic new idea of the paper was to propose an efficient and robust algorithm for the solution of such problems by the conjugate gradient method (CGM) via quadratic programming. The results were promising as compared with existing algorithms. [Olotu and Dawodu, 2013] developed a robust algorithm for solving a class of optimal control problems in which the control effort is proportional to the state of the dynamic system. A typical model was studied which generates a constant feedback gain, an estimate of the Riccati equation for large values of the final time. Using the third Simpson's Rule, a discretized unconstrained non-linear problem via the Augmented Lagrangian Method was obtained. This problem was consequently subjected to the Broydon-Fletcher-Goldberg-Shannon (BFGS) Method based on the Quasi-Newton algorithm. The positive-definiteness of the estimated quadratic control operator was analyzed to guarantee its invertibility in the BFGS Method. This paper is necessitated by the fact that there is a dearth of literature on the mathematical theory of proportional control. This research deals with the problem of optimizing an energy cost function of a linear, one-dimensional system, controlled with a proportional controller. The problem has far reaching theoretical and practical applications in mechanical and electrical engineering.

2 General Problem Formulation

The problem is modelled in order to find the state and control paths that minimize the objective function of the following problem.

$$Minimize\ J(x,w) = \frac{1}{2}\int_0^T f(t,x(t),w(t))dt \quad (1)$$

$$Subject\ to\ \dot{x}(t) = g(t,x(t),w(t)),\ w(t) = mx(t) \quad (2)$$

$$t \in [0,T], \quad x(0) = x_0 \quad (3)$$

$$where\ p,q,a,b,m \in \mathbf{R};\ p,q > 0$$

$$also\ x,w,f,g \in \mathbf{R}\ and$$

$$f\ and\ g\ are\ twice\ differentiable$$

$$m\ is\ the\ proportional\ control\ constant.$$

where x is the state variable which describes the system, w is the control variable which directs how the system evolves. The numerical solution is obtained by applying the Conjugate Gradient Method (CGM) to the discretized form of the problem. The objective function is discretized using the Composite Simpson's Rule. The Fourth-Order Adams-Moulton Technique is used in discretizing the constraint.

3 Methodology

Given the optimal control model with feedback law having the form

$$Minimize\ J(x,w) = \frac{1}{2}\int_0^T (px^2(t) + qw^2(t))dt \quad (4)$$

$$Subject\ to\ \dot{x}(t) = ax(t) + bw(t),\ w(t) = mx(t) \quad (5)$$

$$t \in [0,T],\ x(0) = x_0 \quad (6)$$

$$where\ p,q,a,b,m \in \mathbf{R};\ p,q > 0\ and$$

$$m\ is\ the\ proportional\ control\ constant.$$

In discretizing the objective function

$$J(x,w) = \frac{1}{2}\int_0^T (px^2(t) + qw^2(t))dt \quad (7)$$

we use the Composite Simpson's Rule which is given as

$$\int_0^n f(t)dt = \frac{h}{3}\left\{ f(0) + 2\sum_{j=1}^{\frac{n}{2}-1} f(t_{2j}) + 4\sum_{j=1}^{\frac{n}{2}} f(t_{2j-1}) + f(n)\right\} - \left(\frac{n}{180}\right)h^4 f^4(\zeta) \quad (8)$$

Since $w(t) = mx(t)$,

$$J(x,w) = \frac{1}{2}\int_0^T (px^2(t) + q(mx(t))^2)dt \quad (9)$$

$$= \frac{(p+qm^2)}{2}\int_0^T x^2(t)dt \quad (10)$$

Given that $n = \frac{T}{h}$, where n is the number of partitions and h is the step length, we apply the Composite Simpson's Rule as follows

$$J(x,w) = \frac{(p+qm^2)}{2}\left(\frac{h}{3}\right)\left\{ f(0) + 2\sum_{j=1}^{\frac{n}{2}-1} x_{2j}^2 + 4\sum_{j=1}^{\frac{n}{2}} x_{2j-1}^2 + f(n)\right\} \quad (11)$$

$$J(x,w) = \frac{h(p+qm^2)}{6}\left\{ x_0^2 + 2\sum_{j=1}^{\frac{n}{2}-1} x_{2j}^2 + 4\sum_{j=1}^{\frac{n}{2}} x_{2j-1}^2 + x_n^2 \right\} \quad (12)$$

Setting $\frac{M}{2} = \frac{h(p+qm^2)}{6}$ and putting the expression in matrix form, we have

$$\mathbf{J}(\mathbf{X}) = \frac{1}{2}X^T AX + C \quad (13)$$

where $X^T = (x_1, x_2, x_3, x_4, ..., x_{n-2}, x_{n-1}, x_n)$ is an n-dimensional row vector,

$$A = \begin{pmatrix} 4M & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 2M & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 4M & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & 2M & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 2M & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 4M & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & M \end{pmatrix}$$

is an $n \times n$ dimensional coefficient matrix defined below as:

$$A = A_{ij} = \begin{cases} 4M, & \text{i=j (odd);} \\ 2M, & \text{i=j (even);} \\ M, & \text{i=j=n;} \\ 0, & i \neq j, \end{cases} \quad (14)$$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{pmatrix} \text{ and } C = \frac{M}{2} x_0^2.$$

To discretize the constraint

$$\dot{x}(t) = ax(t) + bw(t) \quad (15)$$

we use the Fourth-Order Adams-Moulton Technique which is defined as

$$x_{i+1} = x_i + \frac{h}{24}\Big\{9f(t_{i+1}, x_{i+1}) + 19f(t_i, x_i) - 5f(t_{i-1}, x_{i-1}) + f(t_{i-2}, x_{i-2})\Big\} \quad (16)$$

Again, since $w(t) = mx(t)$, we have

$$\dot{x}(t) = ax(t) + bmx(t) \quad (17)$$
$$\dot{x}(t) = Bx(t) \quad (18)$$

where $B = a + bm$.

Applying the Adams-Moulton Technique, we have

$$x_{i+1} = x_i + B\Big\{\frac{h}{24}(9x_{i+1} + 19x_i - 5x_{i-1} + x_{i-2})\Big\} \quad (19)$$

$$x_{i+1} = \frac{24 + 19Bh}{24 - 9Bh}x_i - \frac{5Bh}{24 - 9Bh}x_{i-1} + \frac{Bh}{24 - 9Bh}x_{i-2} \quad (20)$$

If $D = \frac{24+19Bh}{24-9Bh}$, $E = -\frac{5Bh}{24-9Bh}$, $F = \frac{Bh}{24-9Bh}$, we have

$$x_{i+1} = Dx_i + Ex_{i-1} + Fx_{i-2} \quad (21)$$

Taking values from $i = 2$ to $i = n - 1$ and putting the expression in matrix form, we have

$$\begin{pmatrix} -E & -D & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ -F & -E & -D & 1 & \dots & 0 & 0 & 0 & 0 \\ 0 & -F & -E & -D & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & -F & -E & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & -D & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & -E & -D & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & -F & -E & -D & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_{n-3} \\ x_{n-2} \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} Fx_0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (22)$$

So,

$$GX = K \quad (23)$$

where G is of dimension $(n-2) \times n$, X is of dimension $n \times 1$ and K is of dimension $(n - 2) \times 1$.

By parametric representation, the discretized proportional control problem becomes

$$Minimize\ \mathbf{J}(\mathbf{X}) = \frac{1}{2}X^T AX + C \quad (24)$$
$$Subject\ to\ GX = K \quad (25)$$

The unconstrained form of the original problem is obtained by using the Quadratic Penalty Function Method as follows:

$$L(X, \mu) = \frac{1}{2}X^T AX + C + \mu||GX - K||^2 \quad (26)$$

Expanding and collecting like terms, we have

$$\begin{aligned} L(X, \mu) &= \frac{1}{2}X^T AX + C + \mu(GX - K)^T (GX - K) \quad (27 \\ &= \frac{1}{2}X^T AX + C + \mu((GX)^T - K^T)(GX - K) \\ &= \frac{1}{2}X^T AX + C + \mu(X^T G^T - K^T)(GX - K) \\ &= (\frac{1}{2}X^T AX + \mu X^T XG^T G) - 2\mu K^T GX + (C + \mu K^T K \\ &= X^T(\frac{1}{2}A + \mu G^T G)X - 2\mu K^T GX + (C + \mu K^T K) \\ &= X^T A_p X + U^T X + V \quad (28 \end{aligned}$$

where $A_p = \frac{1}{2}A + \mu G^T G$ is of dimension $n \times n$, $U^T = -2\mu K^T G$ is of dimension $1 \times n$ and $V = C + \mu K^T K$ is of dimension 1×1.

Equation (28) is a quadratic programming problem which can be solved by using the Conjugate Gradient Method (CGM).

4 The Conjugate Gradient Algorithm

The Conjugate Gradient Algorithm is given as follows [Yang, Cao, Chung and Morris, 2005]:

Step 0. With the iteration number $k = 0$, find the objective function value $f_0 = f(x_0)$ for the initial point x_0.
Step 1. Initialize the inside loop index, the temporary solution and the search direction vector to $n = 0$, $x(n) = x_k$ and $s(n) = -g_k = -g(x_k)$, respectively, where $g(x) =$ is the gradient of the objective function $f(x)$.
Step 2. For $n = 0$ to $N-1$, repeat the following things:
Find the (optimal) step size

$$\alpha_n = ArgMin_{\alpha} f((x_n) + \alpha s(n)) \tag{29}$$

and update the temporary solution to

$$x(n+1) = x(n) + \alpha_n s(n) \tag{30}$$

and the search direction vector to

$$s(n+1) = -g_{n+1} + \beta_n s(n) \tag{31}$$

with

$$\beta_n = \frac{[g_{n+1} - g_n]^T g_{n+1}}{g_n^T g_n} \quad (Polak - Ribi\grave{e}re) \tag{32}$$

or

$$\beta_n = \frac{g_{n+1}^T g_{n+1}}{g_n^T g_n} \quad (Fletcher - Reeves) \tag{33}$$

Step 3. Update the approximate solution point to $x_{k+1} = x(N)$, which is the last temporary one.
Step 4. If $x_k \approx x_{k-1}$ and $f(x_k) \approx f(x_{k-1})$, then declare x_k to be the minimum and terminate the procedure. Otherwise, increment k by one and go back to Step 1.

5 The Analytical Solution

Lemma 5.1. *Consider*

$$J(w) = \int_{t_0}^{t_1} f(t, x(t), w(t))dt \tag{34}$$

$$Subject\ to\ \dot{x}(t) = g(t, x(t), w(t)),\ x(t_0) = x_0 \tag{35}$$

Suppose that $f(t,x,w)$ and $g(t,x,w)$ are both continuously differentiable functions in their three arguments and concave in x and w. Suppose w^* is a control, with associated state x^*, and λ a piecewise differentiable function, such that w^*, x^*, and λ together satisfy on $t_0 \leq t \leq t_1$:

$$f_w + \lambda g_w = 0, \tag{36}$$

$$\lambda' = -(f_x + \lambda g_x), \tag{37}$$

$$\lambda(t_1) = 0, \tag{38}$$

$$\lambda(t) \geq 0. \tag{39}$$

Then for all controls w, we have

$$J(w^*) \geq J(w). \tag{40}$$

Proof. See [Lenhart and Workman, 2007]

Lemma 5.2. *Given the optimal control $w^*(t)$ proportional to the solution $x^*(t)$ of the state system at a constant rate $m \in \mathbf{R}$ that minimizes the performance index $J(x,w)$ over $[0,T]$, then there exists a unique solution that satisfies the condition $a + bm < 0$ with the proportional control constant and optimal objective values defined as $m = -\frac{1}{b}\left\{a + \sqrt{\frac{pb^2+qa^2}{q}}\right\}$ and $J^*(m) = \frac{x_0^2(p+qm^2)}{4(a+bm)}\left\{e^{2(a+bm)t} - 1\right\}$ respectively.*

Proof. See [Olotu and Dawodu, 2013]

The proof gives the general analytical solution to the proportional control problem under consideration as follows:

$$m = -\frac{1}{b}\left\{a + \sqrt{\frac{pb^2 + qa^2}{q}}\right\} \tag{41}$$

$$x(t) = x_0 e^{(a+bm)t},\ t \in [0,T] \tag{42}$$

$$w(t) = mx(t) \tag{43}$$

$$J^*(m) = \frac{x_0^2(p + qm^2)}{4(a + bm)}\left\{e^{2(a+bm)t} - 1\right\} \tag{44}$$

where m is the *proportional control constant*, $x(t)$ is the *state variable*, $w(t)$ is the *control variable* and $J^*(m)$ is the *optimal objective value*. In order to control the exponential growth of $J^*(m)$ for infinitely large values of t, the restriction $a + bm < 0$ was imposed in the process of obtaining the solution. This helps to guarantee the existence, convergence and asymptotic stability of the solution.

6 Numerical Examples and Results

6.1 Example 1

Consider a one-dimensional optimal control problem

$$Minimize\ J(x,w) = \frac{1}{2}\int_0^5 (2x^2(t) + w^2(t))dt$$
$$Subject\ to\ \dot{x}(t) = 2x(t) + 3w(t),\ x(0) = 1,\ 0 \le t \le 5$$

Here, $p = 2$, $q = 1$, $a = 2$, $b = 3$ and $x_0 = 1$.

The analtyical objective value is $J_A = 0.37168976$ and the numerical objective value from the Conjugate Gradient based Penalty Function Method using MATLAB® is $J_N = 0.37167224$. We take $\mu = 100$, $Tol = 10^{-5}$, and $h = 0.25$.

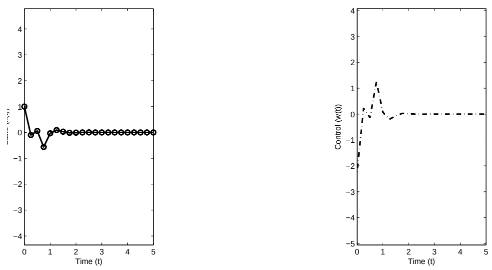

Figure 1. Separate Graphs of State and Control against Time

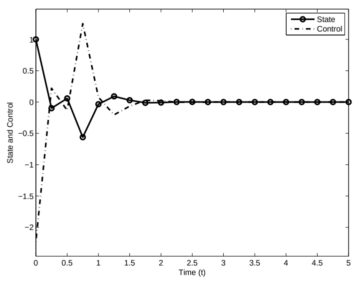

Figure 2. Joint Graphs of State and Control against Time

6.2 Example 2

Minimize a population, with exponential growth, modelled by

$$Minimize\ J(x,w) = \frac{1}{2}\int_0^1 (x^2(t) + w^2(t))dt$$
$$Subject\ to\ \dot{x}(t) = x(t) + w(t),\ x(0) = 1.$$

In this case, $p = 1$, $q = 1$, $a = 1$, $b = 1$ and $x_0 = 1$.

The analtyical objective value is $J_A = 1.13575983$ and the numerical objective value from the Conjugate Gradient based Penalty Function Method using MATLAB® is $J_N = 1.13577134$. We take $\mu = 100$, $Tol = 10^{-5}$, and $h = 0.25$.

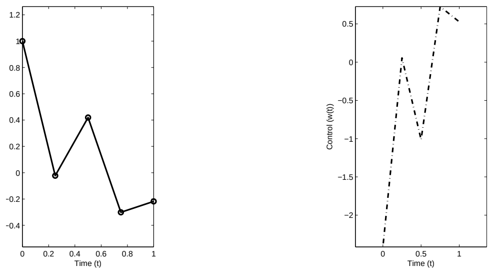

Figure 3. Separate Graphs of State and Control against Time

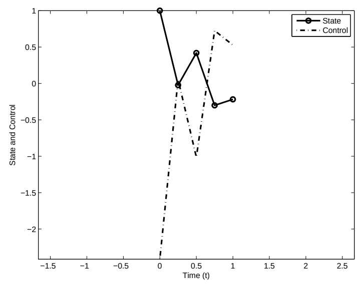

Figure 4. Joint Graphs of State and Control against Time

6.3 Example 3

Let $x(t)$ represent the number of tumor cells at time t (with exponential growth factor α) and $w(t)$ the drug concentration. Minimize, simultaneously, the number of tumor cells at the end of the treatment period and the accumulated harmful effects of the drug on the body if the general form of the problem is

$$Minimize\ J(x,w) = \frac{1}{2}\int_0^T (2x^2(t) + 2w^2(t))dt$$
$$Subject\ to\ \dot{x}(t) = \alpha x(t) - w(t),\ x(0) = x_0 > 0.$$

Take $T = 4$, $\alpha = 0.35$ and $x_0 = 2$.

Now, $p = 2$, $q = 2$, $a = 0.35$ and $b = -1$.

The analtyical objective value is $J_A = 5.63674883$ and the numerical objective value from the Conjugate Gradient based Penalty Function Method using MATLAB® is $J_N = 5.63669538$. We take $\mu = 100$, $Tol = 10^{-5}$, and $h = 0.25$.

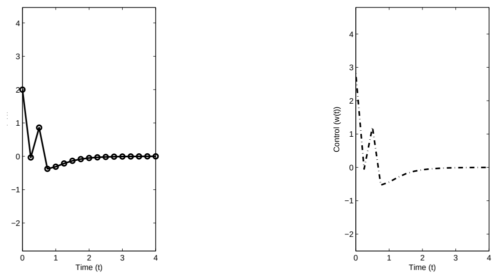

Figure 5. Separate Graphs of State and Control against Time

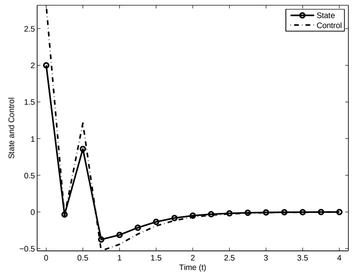

Figure 6. Joint Graphs of State and Control against Time

6.4 Convergence Analysis of Results

Definition 6.1. *Given that $x_k \subset \mathcal{R}^n$ is a sequence of solutions x_k that approaches a limit x^* (i.e. $x_k \to x^*$), then the error $e(x_k) = e_k$ such that $e(x_k) = e_k = ||x_k - x^*|| \geq 0$, $\forall x_k \subset \mathcal{R}^n$ and $e(x^*) \neq 0$.*
The convergence ratio β can be expressed as:

$$\beta = \lim_{k\to\infty} \frac{e_{k+1}}{e_k} = \lim_{k\to\infty} \frac{||x_{k+1} - x^*||}{||x_k - x^*||}, \quad \forall k \tag{45}$$

If $\beta = 0$, convergence is said to be superlinear.
If $0 < \beta < 1$, convergence is said to be linear.
If $\beta = 1$, convergence is said to be sublinear.

Considering the results for Problem 1, we obtained the convergence ratio profile of our scheme as shown in the table below. The profile is computed with respect to the Penalty Parameter μ. J is the Objective Value and β is the Convergence Ratio.

Table 1. The Convergence Ratio Profile

μ	**J**	β
1.0×10^2	0.37147256	0.09932549
1.0×10^3	0.37165241	0.09092986
1.0×10^4	0.37167044	0.00000000
1.0×10^5	0.37167224	0.00000000

The convergence ratio profile shows that for increasing values of the penalty parameter μ, the convergence ratio $\beta \to 0$ very fast. The analysis shows that the convergence is superlinear which is an improvement on the work presented by [Olotu and Dawodu, 2013]. This is an agreeable convergence for optimization algorithms.

7 Conclusion

This research work has shown the efficiency of the conjugate gradient method in solving constrained proportional control problems which are transformed to unconstrained problems by using the much celebrated penalty function method. A wide array of problems in the fields of engineering and the life sciences can be solved reliably by using the new algorithm developed which converges faster than existing algorithms.

References

Adekunle, A.I. and Olotu, O. (2012). An algorithm for optimal control of delay differential equations. *Pacific Journal of Science and Technology*, **13**(1), pp. 228–237.

Bellmore, M., Greenberg, H.J., and Jarvis, J.J. (1968). *Generalized Penalty-Function Concepts in Mathematical Optimization*. Maryland.

Fernández-Cara, E., and Zuazua, E. (1979). *Control Theory: History, Mathematical Achievements and Perspectives*. Washington D.C..

Lenhart, S., and Workman, J.T. (2007). *Optimal Control Applied to Biological Models*, Taylor and Francis Group, Boca Raton, Florida.

Olotu, O., and Dawodu, K.A. (2013). Quasi-Newton embedded augmented lagrangian algorithm for discretized optimal proportional control problems. *Journal of Mathematical Theory and Modeling*, **3**(3).

Olotu, O., and Olorunsola, S.A. (2006). An algorithm for a discretized constrained, continuous quadratic control problem. *Journal of Applied Sciences*, **9**(1), pp. 6249–6260.

Schwartz, A.L. (1996). *Theory and Implementation of Numerical Methods Based on Runge-Kutta Integration for Solving Optimal Control Problems*. Berkeley.

Yang, W.Y., Cao, W., Chung, T., and Morris, J. (2005). *Applied Numerical Methods Using MATLAB®*. John Wiley and Sons, Inc., Hoboken, New Jersey.