Journal of Mathematics and Statistics 6 (1): 42-46, 2010
ISSN 1549-3644

Quadratic Convergence Ratio of Constrained Optimal Control Problems

O. Olotu,
Department of Mathematical Sciences, The Federal University of Technology,
PMB 704, Akure, Ondo State, Nigeria

Abstract: Problem statement: Earlier Researchers developed algorithms for solving quadratic control problems, but did not address the quadratic convergence ratio profile. **Approach:** Hence, the objective of this work is using our developed scheme, Discretized Continuous Algorithm (DCA), to examine both convergence profile and quadratic ratio profile demonstrating the effectiveness and efficiency of the scheme. Methodologically, we obtained generalized unconstrained formulation of the constrained problem. Using Conjugate Gradient Method (CGM) and linear ratio of the deviation of the generated iterates from the exact solution, we obtained respectively the objective values and the quadratic ratio values. **Results:** If these quadratic ratio values fall within the interval [1, ∞], then the convergence is quadratically convergent. Two examples were examined and satisfied the condition for quadratic convergence. **Conclusion:** Thus, we can conclude that the algorithm is optimally effective and efficient.

Key words: Convergence profile, discretized, quadratic ratio, quadratic convergent and effective

INTRODUCTION

In (Olorunsola and Olotu, 2004; Olotu and Olorunsola, 2005), we presented the development of the discretized algorithm for the solution of a class of constrained optimal control problems with real coefficients. As in (Balakrishman and Neustadt, 1964; Bertsekas, 1996), we resort to a finite approximation of it by discretizing the time interval and using finite difference method for its differential constraint. By (Bertsekas, 1996), a penalty function method was used to obtain the unconstrained formulation of the given constrained problem. Using (Fletcher and Reeves, 1964; Gutknecht and Rozloznik, 2002), reviewed by (Ibiejugba, 1984), with the bilinear form expression, an associated operator as in (Ibiejugba, 1984) circumventing the rigorous calculations inherent in the Function Space Algorithm (FSA) (Balakrishman and Neustadt, 1964; Olorunsola, 1996) was stated as a theorem which aided the numerical method as a tool for solving the problems and for investigating quadratic convergence ratio profile.

MATERIALS AND METHODS

Generalized problem (P):

$$\text{Minimize} \int_0^T \left(ax^2(t) + bu^2(t)\right)dt$$

subject to (1)

$$\dot{x}(t) = cx(t) + du(t),\ 0 \le t \le T$$

$$x(0) = 0 \text{ and } a, b, c, d \text{ are real numbers}$$

We now provide the ingredients for the development of the disretized scheme which gives the framework for the construction of the associated operator A.

Discretization: Using discretization, finite difference method as in (Olorunsola and Olotu, 2004; Olotu and Olorunsola, 2005) and applying the penalty function method, we obtain the equivalent unconstrained and discretized formulation of the penalized problem:

$$\operatorname{Min} J(x,u,\mu) = \sum_{k=0}^{n} \left\{ \begin{aligned} &\Delta_k\left(ax_k^{\,2}(t_k) + bu_k^{\,2}(t_k)\right) \\ &+\mu[x_{K+1}(k+1) \\ &-x_k(t_k) - \Delta_k cx_k(t_k) \\ &-d\Delta_k u_k(t_k)]^2 \end{aligned} \right\}$$

$$= \sum_{k=0}^{n} \left\{ \begin{aligned} &\Delta_k\left(ax_k^{\,2}(t_k) + bu_k^{\,2}(t_k)\right) \\ &+\mu[x_{k+1}^{\,2}(t_k) + x_k^{\,2}(t_k) \\ &+\Delta_k^{\,2}c^2x_k^{\,2}(t_k) \\ &+d^2\Delta_k^{\,2}u_k^{\,2}(t_k) + 2c\Delta_k x_k^{\,2}(t_k) \\ &+2d\Delta_k x_k(t_k)u_k(t_k) \\ &+2cd\Delta_k^{\,2}x_k(t_k)u_k(t\ ) \\ &-2x_{k+1}(t_k)x_k(t_k) \\ &-2c\Delta_k x_{k+1}x_k(t_k) \\ &-2d\Delta_k x_{k+1}(t_k)u_k(t_k)] \end{aligned} \right\} \tag{2}$$

Simplifying (2), we have:

$$\sum_{k=0}^{m}\left\{\begin{array}{l} x_k{}^2(t_k)[a\Delta_k+\mu+\mu\Delta_k{}^2c^2 \\ +\mu 2c\Delta_k]+u_k{}^2(t_k)[b\Delta_k+\mu d^2\Delta_k{}^2] \\ +\mu x_{k+1}{}^2(t_k)+x_k(t_k)u_k(t_k) \\ [2d\Delta_k\mu+2cd\Delta_k{}^2\mu] \\ +x_{k+1}(t_k)x_k(t_k)[-2\mu-2\mu c\Delta_k] \\ +x_{k+1}(t_k)u_k(t_k)[-2\mu d\Delta_k]: \end{array}\right\} \quad (3)$$

Let $Z_k=\begin{pmatrix} x_k(t_k) \\ u_k(t_k)\end{pmatrix}$ and $y_k(t_k)=x_{k+1}(t_k)$

Let
$$\begin{aligned} \alpha_k &= a\Delta_k+\mu+\mu\Delta_k{}^2c^2+\mu 2c\Delta_k \\ \beta_k &= b\Delta_k+\mu d^2\Delta_k{}^2 \\ \lambda_k &= 2\mu d\Delta_k+2\mu cd\Delta_k{}^2 \\ \delta_k &= -2\mu(1+c\Delta_k),\cdots \rho_k=-2\mu d\Delta_k \end{aligned} \quad (4)$$

Now using (3 and 4) becomes:

$$\sum_{k=0}^{n}\left\{\begin{array}{l} \alpha_k x_k{}^2(t_k)+\beta_k u_k{}^2(t_k)+y_k{}^2(t_k)\mu \\ +x_k(t_k)u_k(t_k)\lambda_k+y_k(t_k)x_k(t_k)\delta_k \\ +y_k(t_k)u_k(t_k)\rho_k \end{array}\right\} \quad (5)$$

For a full detail of the construction of operator A, we refer to (Olorunsola and Olotu, 2004; Olotu and Olorunsola, 2005), while we state below the contents of theorem 1, establishing the operator A.

Theorem 1 (Operator A): Let the initial guess of the conjugate algorithm be $Z_0(t_0)$ so that:

$$z_0(t_0)=(x_0(t_k),u_0(t_k))$$

Then the control operator A associated with problem 1 is given by:

$$\begin{aligned} &\langle Z_{K1}(t_k),AZ_{K2}(t_k)\rangle_H = \\ &\sum_{k=0}^{n}\{\alpha_k x_{K1}(t_k)x_{K2}(t_k)+\beta_k u_{K1}(t_k) \\ &u_{K2}(t_k)+y_{k1}(t_k)y_{k2}(t_k)\mu+\lambda_K x_{K1}(t_k)u_{K2}(t_k) \\ &+\lambda_K u_{K1}(t_k)x_{K2}(t_k)+\delta_K y_{k1}(t_k)x_{k2}(t_k) \\ &+\partial_k y_{k2}(t_k)x_{k1}(t_k)+\rho_k y_{k1}(t_k)u_{k2}(t_k) \\ &+\rho_k u_{k1}(t_k)y_{k2}(t_k)\} \end{aligned} \quad (6)$$

$$AZ_{K2}(t_k)=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{pmatrix}\begin{pmatrix} x_{k2} \\ u_{k2}\end{pmatrix}=\begin{pmatrix} A_{11}x_{k2}+A_{12}u_{k2} \\ A_{21}x_{k2}+A_{22}u_{K2}\end{pmatrix} \quad (7)$$

Further Olorunsola (1996) simplifying (6) and using (Gutknecht and Rozloznik, 2002), we have:

$$\begin{aligned} &\langle Z_{K1}(t_k),AZ_{K2}(t_k)\rangle_H = \\ &\sum_{k=0}^{n}\{\alpha_k x_{K1}(t_k)x_{K2}(t_k)+\beta_k u_{K1}(t_k)u_{K2} \\ &(t_k)+\mu[(\Delta_K\dot{x}_{K1}+x_{K1})(\Delta_K\dot{x}_{K21}+x_{K2})] \\ &+\lambda_K x_{K1}u_{K2}+\lambda_K u_{K1}x_{K2}+\delta_K(\Delta_K\dot{x}_{K1}+x_{K1})x_{K2} \\ &+\delta_K x_{K1}(\Delta_K\dot{x}_{K2}+x_{K2})+\rho_K(\Delta_K\dot{x}_{K1}+x_{K1})u_{K2} \\ &+\rho_K u_{K1}(\Delta_K\dot{x}_{K2}+x_{K2})\} \end{aligned} \quad (8)$$

$$\begin{aligned} &\sum_{k=0}^{n}\{\alpha_k x_{K1}(t_k)x_{K2}(t_k)+\beta_k u_{K1}(t_k)u_{K2}(t_k) \\ &+\mu\Delta_K{}^2\dot{x}_{K1}(t_k)\dot{x}_{K2}(t_k)+\mu\Delta_K\dot{x}_{K1}(t_k)x_{K2}(t_k) \\ &+\mu\Delta_K x_{K1}(t_k)\dot{x}_{K2}(t_k)\dot{x}_{k2}(t_k) \\ &+\mu x_{K1}(t_k)x_{K2}(t_k)+\lambda_K x_{K1}(t_k)u_{K2}(t_k) \\ &+\lambda_K u_{K1}(t_k)x_{K2}(t_k)+\delta_K\Delta_K\dot{x}_{K1}(t_k)x_{K2}(t_k) \\ &+\delta_K x_{K1}(t_k)x_{K2}(t_k) \\ &+\delta_K\Delta_K x_{K1}(t_k)\dot{x}_{K2}(t_k)+\delta_K x_{K1}(t_k)x_{K2}(t_k) \\ &+\rho_K\Delta_K u_{K2}(t_k)\dot{x}_{K1}(t_k)+\rho_K u_{K2}(t_k)x_{K1}(t_k) \end{aligned} \quad (9)$$

Proof: Setting u_{k2} (t) = 0, in (9) and by (Olorunsola, 1996) we have:

$$\begin{pmatrix} A_{11}x_{k2} \\ A_{21}x_{k2}\end{pmatrix}=\begin{pmatrix} V_{11} \\ V_{21}\end{pmatrix} \quad (10)$$

and

$$\langle Z_{k1}(t_k),W_{k2}(t_k)\rangle_H = \sum_{k=0}^{n}\left\{\begin{array}{l} a_k x_{k1}(t_k)x_{k2}(t_k)+\mu\Delta_k{}^2\dot{x}_{k1}(t_k)\dot{x}_{k2}(t_k) \\ +\mu\Delta_k\dot{x}_{k1}(t_k)x_{k2}(t_k)+\mu\Delta_k x_{k1}(t_k)\dot{x}_{k2}(t_k) \\ +\mu x_{k1}(t_k)x_{k2}(t_k)+\lambda_k u_{k1}(t_k)x_{k2}(t_k) \\ +\delta_k\Delta_k\dot{x}_{k1}(t_k)x_{k2}(t_k)+\delta_k\Delta_k x_{k1}(t_k)\dot{x}_{k2}(t_k) \\ +\delta_k x_{k1}(t_k)x_{k2}(t_k) \end{array}\right\} \quad (11)$$

$$\begin{aligned} &=\sum_{k=0}^{n}\{x_{K1}(t_k)[\alpha_k x_{k2}(t_k)+\mu\Delta_K\dot{x}_{K2}(t_k) \\ &\quad+\mu x_{K2}(t_K)+\delta_K x_{K2}(t_k) \\ &\quad+\delta_K\Delta_K\dot{x}_{K2}(t_k)+\delta_K x_{K2}(t_k)] \\ &\quad+\dot{x}_{K1}(t_k)[\mu\Delta_k{}^2\dot{x}_{K2}(t_k)+\mu\Delta_K x_{K2}(t_k) \\ &\quad+\delta_K\Delta_K x_{K2}(t_k)]+u_{K1}(t_k)[\lambda_K x_{k2}(t_k)]]\} \end{aligned} \quad (12)$$

$$=\sum_{k=0}^{n}\{x_{K1}(t_k)V_{11}(t_k)+\dot{x}_{K1}(t_k)\dot{V}_{11}(t_k)+u_{K1}(t_k)V_{21}\} \quad (13)$$

Define:

$$\Omega(t_k) = (\alpha_K + \mu + 2\delta_K)x_{K2}(t_k) + (\mu\Delta_K + \delta_K\Delta_K)\dot{x}_{K2}(t_k)$$

And:

$$f(t_k) = \mu\Delta_K{}^2\dot{x}_{K2}(t_k) + (\mu\Delta_K + \delta_K)x_{K2}(t_k)$$

Comparing and we have:

$$A_{21}u_{k1}(t_k) = V_{21}(t_k) = \lambda_k x_{k2}(t_k) \quad (14)$$

To obtain the component $A_{11}x_{K1}(t_k) = V_{11}(tk)$, where:

$$\Omega(t_k) - V_{11}(t_k) \quad \text{and} \quad f(t_k) - \dot{V}_{11}(t_k)$$

Are both continuous functions on [0, T] and choosing:

$$x_{K1}(\bullet) \in D[0,T] \ni x_{K1}(0) = x_{K1}(T) = 0$$

$$\begin{aligned}&\int_0^T \{x_{K1}(t_k)[\Omega(t_k) - V_{11}(t_k)] + \dot{x}_{K1}(t_k)[f(t_k)\\ &-\dot{V}_{11}(t_k)]\}dt_k = 0\end{aligned} \quad (15)$$

From (15), obtain a second order bounded value problem (Olorunsola, 1996):

$$\ddot{V}_{11}(t_k) - V_{11}(t_k) = q(t_k) = \dot{f}(t_k) - \Omega(t_k) \quad (16)$$

With:

$$V_{11}(0) = p_0 \quad \text{and} \quad \dot{V}_{11}(0) = r_0$$

Obtain the Laplace transform of the ordinary differential equation, with:

$$L\{V_{11}(t_k)\} = \hat{V}_{11}(s) \quad \text{and} \quad L\{q(t_k)\} = Q(s) \quad (17)$$

Now, using inverse Laplace transform with the convolution theorem, we have:

$$\begin{aligned}V_{11}(t_k) &= A_{11}(t_k) = \tau_0 \sinh(t_k)\\ &+[(\alpha_k + \mu + 2\delta_{kk})x_{K2}(0)\\ &+(\mu + \delta_k)\Delta_K \dot{x}_{K2}(0)]\cosh t_k\\ &-\sinh T\{\mu\Delta_k^2 \dot{k}_{k2}(0)\\ &+\Delta_K(\mu + \delta_K)x_{K2}(0)\}\\ &+\int_0^T\{\left(\mu\Delta_K{}^2\dot{x}_{K2}(s_k)\right)\\ &+\Delta_K(\mu + \delta_K)x_{K2}(s_k)\}\cosh(t_k - s_k)ds_k -\\ &-\int_0^T\{(\alpha_K + \mu + 2\delta_K)x_{K2}(s_k)\\ &+\Delta_K(\mu + \delta_K)\dot{x}_{K2}(s_k)\}\sinh(t_{kk} - s_k)ds_k\end{aligned} \quad (18)$$

In Eq. 9, setting $x_{k2}(t_k) = 0 \rightarrow \dot{x}_{k2}(t_k) = 0$

And following the same steps as from Eq. 10-18, we have:

$$V_{22}(t_k) = A_{22}u_{K2}(t_k) = \beta_K u_{K2}(t_k) \quad (19)$$

$$\begin{aligned}V_{12}(t_k) &= (\rho_k\Delta_k)u_{K2}(0)\sinh(t_k)\\ &-\int_0^{t_k}(\rho_k\Delta_k)u_{K2}\cosh(t_k - s_k)ds_k\\ &-\int_0^{t_k}(\lambda_k + \rho_k)u_{K2}(s_K)\sinh(t_k - s_k)ds_k\\ &+(\lambda_k + \rho_k)u_{K2}(0)\cosh t_k\\ &+\frac{\sinh t}{\sinh T}\{(\lambda_k + \rho_k)u_{K2}(T)\\ &-(\lambda_k + \rho_k)u_{K2}(0)\cosh T\\ &-(\rho_k\Delta_k)u_{K2}(0)\sinh(T)\\ &+\int_0^{T_k}(\rho_K\Delta_K)u_{K2}(s_K)\cosh(T - s_k)ds_k\\ &+\int_0^T(\lambda_K + \rho_K)u_{K2}\sinh(T - s_k)ds_k\}\end{aligned} \quad (20)$$

Having constructed operator A, written as:

$$A = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}$$

where, v_{11} (18), v_{12} is (20), v_{21} is (14), v_{22} is (19).

The discretized algorithm was now applied to the following hypothetical problems P1 and P2 whose quadratic convergence ratios were examined and found to satisfy the definition of quadratic convergence stated as follows:

Definition 1: Let the sequence $\{x_k\}$ in the normed linear space:

$$\{R^n, \|.\|\} \text{ converge to } \dot{x}$$

Then the order of convergence is p, if:

$$\lim_{k\to\infty} \frac{\left\|x_{k+1} - \dot{x}\right\|}{\left\|x_k - \dot{x}\right\|^p} \le R$$

where, R is a positive constant not necessarily less than one. If p = 2, then the convergence of $\{x_k\}$ is called quadratic.

Examples:
Example problem P1:

$$\text{Minimize} \int_0^1 \left(x^2(t) + u^2(t)\right) dt$$

Such that:

$$\dot{x}(t) = 2.095x(t) + 1.904u(t),\ 0 \le t \le 1$$

The analytic solution is 1.0647 given by (Olorunsola and Olotu, 2004). The numerical solution, in Table 1, to this problem is obtained by assuming the following initial values and parameters:

$X_0 = 1$, $u_0 = 0.5$, a = b =1, c = 2.095 and d = 1.904

Example problem P2:

$$\text{Minimize} \int_0^1 \left(x^2(t) + u^2(t)\right) dt$$

Such that:

$$\dot{x}(t) = u(t),\ x(0) = 0,\ 0 \le t \le 1$$

The numerical solution to this problem is obtained by assuming the following initial values for the variables:

$X_0 = 1$, u_0, = 1

The analytic solution is 0.7641. Applying the same algorithm to problem P2 and solving by extended conjugate gradient method, we have the following Table 2.

Table 1: Numerical solution and quadratic ratio for problem P1

Pamt	Stpsize	Itr	Obj.V	Csat	Funct	QR
0.5	0.2	9	1.212418	5.280804	3.852670	6.7643290
1.00	0.2	6	1.256018	4.345545	5.601494	0.6829638
1.50	0.2	6	0.939635	3.605450	6.347810	2.7583200
2.00	0.2	5	1.728642	4.195213	10.11907	1.5061560
2.50	0.2	3	1.428916	3.642317	10.53471	2.7456270

Where: Pamt: μ, penalty parameter, Stp: Step size for the discretization, Itr: Number of iterations, Obj.V: Objective Value, Csat: Constraint satisfaction, Funct: Obj.V+μ*(Csat), QR: Quadratic ratio

Table 2: Numerical solution and quadratic ratio for problem P2

Pamt	Itr	Stps	Obj.V	Csat	Funct	QR
0.50	1	0.2	1.0347120	5.691094E0-2	1.063167	0
1.00	3	0.2	0.8676740	6.353983E-02	0.8938423	1
1.50	7	0.2	0.8769342	6.53919E0-2	0.9657592	1
2.00	2	0.2	0.8769342	0.0489593	0.9748530	1
2.50	3	0.2	0.8747541	3.892236E-02	0.9720601	1

RESULTS AND DISCUSSION

For example problem P1 in Table 1, the analytic solution is 1.0647 at the optimum. The initial objective functional value is 1.212418 with corresponding constraint 5.280804, 9 iterations, step size 2 and quadratic ratio 6.7643290, for the first cycle. As the parameters increase to 1.00 and 1.50, particularly at 1.50, the numerical solutions get closer to the analytic solution. We notice that as the parameters get beyond 1.50, the objective values deviate much more from the analytic solution. However, for all cycles, the quadratic ratios, 6.7643290, 0.6829638, 2.7583200, 1.501560 and 2.7456270, except 0.6829638 at the second cycle, obtained as the infimum per cycle of iterations, fall within the required interval [1, ∞].

For example problem P2 in Table 2, the analytic solution is 0.7641 at the optimum. The initial objective functional value is 1.034712 with corresponding constraint 6.691094402, 2 iterations, step size 2 and zero quadratic ratio, for the first cycle. As the parameters increase to 1.00 and 2.50, the objective functional values 0.8676740, 0.8769342, 0.8769342, 0.8769342 and 0.8774541 compare much more favorably to the analytic solution. We notice that in all cycles, except the first, the quadratic ratio profile is 1. Thus, this shows that the convergence is quadratic.

CONCLUSION

In each cycle and for each example, the results obtained with Discredited Continuous Algorithm (DCA) show that the scheme confirms quadratic ratio profile. We notice, particularly in the second example, that the quadratic ratio value, 1, obtained per cycle, except at the first cycle, shows that the convergence is also linearly convergent. This is as a result of the closeness of the numerical solutions to the analytic solution.

Consequently, we can say that the scheme has demonstrated its effectiveness and efficiency, since its quadratic ratios fall within the specified interval [1, ∞]. It is recommended that the best guessed choices of initial values of the state and control variables and the choice of smaller step size can enhance faster quadratic convergence profile and quadratic ratio profile respectively.

REFERENCES

Balakrishman, A.V. and I.V. Neustadt, 1964. Computing Methods in Optimization Problems, Academic press, Inc., New York.

Bertsekas, D. P., 1996. Constrained Optimization and Lagrange multiplier methods, Academic Press Inc., Massachusetts.

Fletcher, R. and C.M. Reeves, 1964. Function minimization by conjugate gradients. Comput. J., 7: 149-154. http://comjnl.oxfordjournals.org/cgi/content/abstract/7/2/149Gutknecht, M.H. and M., Rozloznik, 2002. A framework for generalized conjugate methods with special emphasis on contributions by Rudiger Weiss. Applied Num. Math., 41: 7-22. DOI: 10.1016/S0168-9274(01)00107-6

Ibiejugba, M.A., 1984. A control operator and some of its applications. J. Optimiz., 103: 31-47. http://cat.inist.fr/?aModele=afficheN&cpsidt=8958544

Olorunsola, S.A., 1996. On two updating methods of the multipliers in the multipliers imbedding algorithm. Syst. Anal. Model. Simulat., 21: 51-73. http://portal.acm.org/citation.cfm?id=243058

Olorunsola, S.A. and O. Olotu, 2004. A discretizexd algorithm for the solution of a constrained, continuous quadratic control problem. J. Nigeria Assoc. Math. Phys. 8: 295-300. http://ajol.info/index.php/jonamp/article/view/40017

Olotu, O and S.A. Olorunsola, 2005. A discretized scheme for continuous optimal control problems constrained by evolution equation of the retarded restoring type with real coefficients. Ife J. Sci., 7: 193-201. http://ajol.info/index.php/ijs/article/view/32177