American Journal of Applied Sciences 5 (2): 89-92, 2008
ISSN 1546-9239

Convergence Ratio Profile for Optimal Control Problems Governed by Ordinary Differential Equations with Matrix Coefficients

[1]O. Olotu and [2]S.A. Olorunsola
[1]The Federal University of Technology, Mathematical Science Department
P.M.B. 704, Akure, Ondo State, Nigeria
[2]The University Of Ado-Ekiti, Mathematical Science, Department
P.M.B.5383, Ado-Ekiti, Ekiti, State, Nigeria

Abstract: The geometric convergence ratio, the main focus of a discretized scheme for constrained quadratic control problem was examined. In order to allow for the numerical applications of the developed scheme, discretizing the time interval and using Euler's scheme for its differential constraint obtained a finite dimensional approximation. Applying the penalty function method, an unconstrained problem was obtained on function minimization with bilinear form expression. This finally led to the construction of an operator. The Scheme was applied to a sampled problem and it exhibited geometric convergence ratio, α, in the open interval (0, 1) as depicted in column 6 of Table 1.

Key words: Quadratic, discretization, bilinear form expression, associated operator, convergence profile and geometric convergence

INTRODUCTION

In[1,2], the scheme establishing the solution of optimal control problems constrained by evolution equation of the delay type with matrix coefficients was presented, without addressing the geometric convergence ratio profile. Here, a class of optimal control problems constrained by ordinary differential equation with matrix coefficients is considered. Discretization of the generalized problem is obtained by discretizing its objective function and using[3] for its differential constraint. Using[4], a penalty function method is applied to convert the constrained problem into an unconstrained formulation problem. With this formulation, an associated control operator was constructed as in[2]. Here, again, we state the constructed operator as a consequence of a theorem in this paper to allow for brevity and compactness of the paper. Consequently, consideration of the convergence profile and the geometric convergence ratio profile as they relate to such class of problems is highlighted by a sample problem for the confirmation of the success of the scheme.

Generalized problem 1

$$Min\int_0^Z \left(x(t)^T Px(t) + u(t)^T Qu(t)\right)dt \qquad (1.1)$$

Subject to

$$\dot{x}(t) = Ax(t) + Cu(t) \quad x(0) = x_0 \quad 0 \le t \le Z, \qquad (1.2)$$

where, $x(t)$, $u(t)\varepsilon R^n$, $x(t)^T, u(t)^T$ *denote the transposes of* $x(t)$ *and* $u(t)$ *respectively*
P, Q *are n by n symmetric matrices*
A *and* C *are* n *by* n *and* n *by* m *matrices not necessarily symmetric respectively.*

MATERIALS AND METHODS

2.1 Discretization: By discretizing (1.1) and (1.2), we have $\dot{X}(t) = Ax(t) + Cu(t)$

$$\left(x(t_{k+1}) - x(t_k)\right)/\Delta_k = Ax_k(t_k) + Cu_k(t_k) \qquad (1.3)$$

X(0)=0

We then have the discretized generalized problem in the form;

$$\begin{aligned} &\min J = \sum_{k=0}^{n} \Delta_k \left(x_k(t)^T Px_k(t_k) + u_k(t_k)^T Qu_k(t_k)\right) \\ &subject\ to\ \left(x(t_{k+1}) - x(t_k)\right)/\Delta_k = Ax_k(t_k) + Cu_k(t_k) \\ &x(0) = 0 \end{aligned} \qquad (1.4)$$

Penalty method's application: Using the standard penalty function[4], we obtain the unconstrained problem 1

$$MinJ(x,u,\mu) = \sum_{k=0}^{n} \left\{ \begin{aligned} &\Delta_k \left(x(t_k)^T Px_k(t_k) + u(t_k)^T Qu_k(t_k)\right) \\ &+ \mu[x_{K+1}(t_{k+1}) - x_k(t_k) - \Delta_k Ax_k(t_k) - C\Delta_k u_k(t_k)]^T \\ &x_{K+1}(t_{k+1}) - x_k(t_k) - \Delta_k Ax_k(t_k) - C\Delta_k u_k(t_k)] \end{aligned} \right\}.$$

Corresponding Author: O. Olotu, The Federal University of Technology, Mathematical Science Department, P.M.B. 704, Akure, Ondo State, Nigeria

$$= \sum_{k=0}^{nn} \left\{ \begin{array}{l} x_k^T(t_k)[P\Delta_k + \mu + \mu A^T \Delta_k^T A + 2\mu A \Delta_k] x(t_k) + u_k(t_k)^T \\ [Q\Delta_k + \mu C^T \Delta_k^T \Delta_k c \ ] u(t_k) + \mu x(t_{k+1})^T (t_k) x(t_{k+1}) \\ + x_k^T(t_k)[2C^T \Delta_k^T \mu + 2C^T \Delta_k^T \Delta_k A\mu] u_k(t_k) \\ + x_{k+1}^T(t_{k+1})[-2\mu - 2\mu A \Delta_k] x_k(t_k) + x_{k+1}^T(t_{k+1})[-2\mu C \Delta_k] u_k(t_k) \end{array} \right\} \quad (1.5)$$

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

In (1.5), let

$$Mk = P\Delta_k + \mu + \mu A^t \Delta_k^t A + 2\mu A \Delta_k$$

$$Nk = Q\Delta_k + \mu C^T \Delta_k^T \Delta_k c$$

$$Bk = 2C^T \Delta_k^T \mu + 2C^T \Delta_k^T \Delta_k A\mu$$

$$Pk = -2\mu - 2\mu A \Delta_k$$

$$Ak = -2\mu C \Delta_k$$

Now, (1.5) becomes

$$\sum_{k=0}^{n} \left\{ \begin{array}{l} Mkx_k^2(t_k) + N_k u_k^2(t_k) + vy_k^2(t_k) + Akx_k(t_k)u_k(t_k) \\ + Bky_k(t_k)x_k(t_k) + Pky_k(t_k)u_k(t_k) \end{array} \right\} \quad (1.6)$$

2.3. Theorem 1: The exact control operator G satisfying generalized problem 1 is given by

$$GZ_k(t_k) = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \begin{pmatrix} x_{k2}(t_k) \\ u_{k2}(t_k) \end{pmatrix} \quad (1.7)$$

where

$$\begin{aligned} G_{12}U_{k2}(t_k) &= e_o \sinh(t_k) + u_{k2}(0)(P_k + A_k)\cosh(t_k) - \sinh(T)(u_{k2}(0)\Delta_k^T P_k) \\ &+ \int_0^T (u_{k2}(s_k)\Delta_k^T P_k)\cosh(T - s_k)ds_k \\ &- \int_0^T (u_{k2}(s_k)(P_k + A_k)\sinh(T - s_k)ds_k \end{aligned} \quad (1.8)$$

$$G_{22}u_{k2}(t_k) = u_{k2}(t_k)N_k \quad (1.9)$$

$$\begin{aligned} G_{11}x_{k2}(t_k) &= \tau_0 \sinh(t_k)[(M_k + \mu + 2B_k)x_{K2}(0) + (\mu\Delta_K + \Delta_K^T B_k)\dot{x}_{K2}(0)]\cosh(t_k) + \\ &- \sinh T\{x_{k2}(0)(\Delta_k^T \mu + \Delta_k^T \beta_k) + \dot{x}_{k2}(0)(\mu\Delta_k^T \Delta_k)\} \\ &+ \int_0^T \{x_{k2}(sk)(\mu\Delta_K + \Delta_K^T B_k) + \dot{x}_{k2}(sk)(\mu\Delta_k^T \Delta_k)\}\cosh(T - sk)ds_k \\ &- \int \{(M_k + \mu + 2B_k)x_{K2}(sk) + (\mu\Delta_K + \Delta_K^T B_k)\dot{x}_{K2}(sk)\}\sinh(T - sk)ds_k \end{aligned} \quad (1.10)$$

$$G_{21}x_{k2}(t_k) = x_{k2}(t_k)(p_k + A_k) + \Delta^T{}_k \dot{x}_{k2}(t_k)p_k \quad (1.11)$$

2.4. Remark 1: The scheme converges at the 4th iteration for each penalty parameter constant μ as depicted in Table 1 for the numerical calculation.

Definition: Let $\{z_n\}$ be a sequence of vectors in a Hilbert Space H with limit z* in H such that

$$\frac{\left\| z_{n+1} - z^* \right\|}{\left\| z_n - z^* \right\|} = y < 1 \quad as \quad n \to \infty$$

Then $\{z_n\}$ is said to converge geometrically to z* with a convergence ratio y as reported by[5]. Here, we recall the various steps of the conjugate gradient algorithm that generates the convergent sequence $\{z_n\}$ of solutions of problem 1 according to[3]. The algorithm employs the explicit knowledge of the control operator G developed in theorem 1

Step 1: Choose initial values for the conjugate descent algorithm ;

$$z_0^T(t_k) = (x_0(t_k), u_0(t_k)), \quad z_0(t_k) \in H$$

Compute $p_0 = -g_0$

While the remaining members are computed as follows;

Step 2: Update x_0 and u_0 such that

$$x_{n+1}(t_k) = x_n(t_k) + \alpha_n p_{x,n}$$

$$u_{n+1}(t_k) = u_n(t_k) + \alpha_n p_{u,n}$$

where $\alpha_n, p_{\bullet,n}$ *are the step length and the descent direction respectively*

$$and \quad \alpha_n = \frac{\langle g_{x,n}, g_{x,n} \rangle}{\langle p_{x,n}, Gp_{x,n} \rangle}$$

Step 3: Update gradient and descent directions with the updating rule

$$g_{x,n+1} = g_{x,n} + \alpha_n Gp_{x,n}$$

$$g_{u,n+1} = g_{u,n} + \alpha_n Gp_{u,n}$$

$$p_{x,n+1} = -g_{x,n} + \beta_n p_{x,n}$$

$$p_{u,n+1} = -g_{u,n} + \beta_n p_{u,n}$$

Where $g_{\bullet,n+1}$ *and* $p_{\bullet,n+1}$ are the gradient and descent direction at the (n+1) th iteration respectively and

G is the control operator in theorem 1,

$$\beta_n = \frac{\langle g_{n+1}, g_{n+1} \rangle}{\langle g_n, g_n \rangle} \text{ and}$$

$$g^T{}_{\bullet,n} = (g_{x,n}, g_{u,n}) \quad and \quad p^T{}_{\bullet,n} = (p_{x,n}, p_{u,n})$$

are the transposes of $g_{\bullet,n}$ *and* $p_{\bullet,n}$ respectively.

Now,

$$Gp_n(t_k) = \begin{pmatrix} G_{11}p_{x,n}(t_k) + G_{12}p_{u,n}(t_k) \\ G_{21}p_{x,n}(t_k) + G_{22}p_{u,n}(t_k) \end{pmatrix} = \begin{pmatrix} Gp_1(t_k) \\ Gp_2(t_k) \end{pmatrix}$$

Setting

$J_{x,k} = \nabla_{x,k} J(x_n(t_k), u_n(t_k), \mu)$ and $J_{u,k} = \nabla_{u,k} J(x_n(t_k), u_n(t_k), \mu)$

where

$$J(x_k(t_k), u_k(t_k), \mu) = \int_0^T (x^T{}_k(t_k)Px_k(t_k) + u^T{}_k(t_k)Qu_k(t_k))dt_k$$

$$+\mu\int_0^T \|\dot{x}_k(t_k) - Ax_k(t_k) - Cu_k(t_k)\| \;\|\dot{x}_k(t_k) - Ax_k(t_k) - Cu_k(t_k)\| dt_k$$

$\nabla_{\bullet,k} J(x_k(t_k), u_k(t_k), \mu)$

is the gradient of $J(x_k(t_k), u_k(t_k), \mu)$ *at the kth step.*

We obtain the following relations,

$$p_{x,k} = -\int_0^T J_{x,k}(x_k(t_k), u_k(t_k), \mu)dt_k$$

or more generally

$$p_{\bullet,k} = -\int_0^T J_{\bullet,k}(x_k(t_k), u_k(t_k), \mu)dt_k$$

Assuming the following remarks, we state and prove the following theorem 2

z* is the optimizer for problem 1

The expression $\left| z_{n+1} - z^* \right\| z_n - z^* \Big|^{-1}$ is the convergence ratio of the sequence $\{z_n(t)\}$ in the Hilbert space **H.**

In[6], a general quadratic functional in the Hilbert Space H to be minimized as

$$F(z) = F_0 + \langle a, z\rangle + \frac{1}{2}\langle z, Az\rangle_H$$

was given, where A is a symmetric positive definite n-square matrix operator of theorem 1

If $F_0 = \langle a, z\rangle_H = 0$, then $\langle z, Az\rangle_H$

(iv) $\{P_{z,n}\}$ are conjugate with respect to the linear operator A i.e. $\langle P_{z,n}, AP_k\} = 0,\ n < k$

Theorem 2: The sequence $\{z_n(t)\}$ of solutions to problem 1 using the explicit knowledge of the control operator A in theorem 1 converges geometrically to {z*} with ratio

λ=1-α**, where**

$$\alpha = \frac{1}{\|z_0\|}.Max\|Az_n\|^3(\langle AP_{z,n}, P_{z,n}\rangle)^{-1} \quad (1.12)$$

Proof:

Let $F(z) = \langle z - z^*, Az\rangle_H$, $z^T(t) = (x(t), u(t), t)$

At the optimality condition, we have $Az^*(t) = 0$

Let z, z_n *be in* H. $F(z_n) = \langle z_n - z^*, A(z - z^*)\rangle$

$$= \langle z_n, Az_n\rangle - \langle z^*, Az^*\rangle \quad (1.13)$$

and

$F(z_{n+1}) = \langle (z_n + \alpha_n P_{z,n}) - z^*, A(z_n + \alpha_n P_{z,n} - z^*\rangle$

$= \langle z_n, Az_n\rangle + \alpha_n\langle z_n, AP_{z,n}\rangle + \alpha_n\langle Pz_n, Az_n\rangle + \alpha_n{}^2\langle P_{z,n}, AP_{z,n}\rangle$

$-\langle z^*, Az_n\rangle - \alpha_n\langle z^*, AP_{z,n}\rangle$ (1.14)

From (1.13) and (1.14),

$$F(z_n) - F(z_{n+1}) = \frac{\|Az_n\|^2 . F(z_n)}{\langle P_{z,n}, AP_{z,n}\rangle . \langle Az_n, z_n\rangle} \quad (1.15)$$

Since A is a self-adjoint operator,[2,7,8], Az*=0

Again, $\langle Az_n, z_n\rangle = \langle Az_n, z_0\rangle + \sum_{k=0}^{n-1}\alpha_k\langle Az_n, p_{z,k}\rangle$

But, $\sum_{k=0}^{n-1}\alpha_n\langle Az_n, P_{z,n}\rangle = 0$ *for* $n \neq k$

Hence, $\langle Az_n, z_n\rangle = \langle Az_n, z_0\rangle$

A is bounded, meaning there exists m, M>0 m, M in R such that for every z in H

$$m\|z - z^*\|^2 \leq \|A(z - z^*)\|^2 \leq M\|z - z^*\|^2 \quad (1.16a)$$

So,

$$\langle Az_n, z_n\rangle \leq \|Az_n\|\|z_0\| \quad (1.16b)$$

Substituting, (1.16b) in (1.15), we have,

$$F(z_n) - F(z_{n+1}) = \frac{\|Az_n\|^2 . F(z_n)}{\langle P_{z,n}, AP_{z,n}\rangle\|Az_n\|\|z_n\|}$$

So,

$$F(z_{n+1}) \leq [1 - \frac{\|Az_n\|^3}{\langle P_{z,n}, AP_{z,n}\rangle\|z_0\|}]F(z_n) \quad (1.17)$$

By (1.16a), $F(z_n) \geq m\|z_n - z^*\|^2$,

So that (1.17) becomes

$$\|z_{n+1} - z\|^2(\|z_n - z^*\|^2)^{-1} \leq \frac{F(z_{n+1})}{F(z_n)} \leq 1 - \frac{\|Az_n\|^3\|z_0\|^{-1}}{\langle P_{z,n}, AP_{z,n}\rangle}$$

Thus end the analytic proof of the geometric convergence ratio.

DATA AND ANALYSIS

Hypothetical example: Now, we shall consider an example to test the efficiency of the developed scheme.

Example 1:

$$Min\int_0^1 \left(x(t)^T Px(t) + u(t)^T Qu(t)\right)dt$$

Table 1: Convergence profile and geometric ratio profile for example

Penalty constant	Number of iterations	Objective function	Constraint satisfaction	Penalty function	Geometric ratio
1	2	3	4	5	6
μ=.0001	1	5	10	5.001	.9483081
	2	3.6554451	7.115152	3.655163	.9363912
	3	3.568452	7.05019	3.569159	.9266359
	4	4.07852	8.54454	4.079097	.82418
μ=.0002	1	5	10	5.001999	.952805
	2	3.248526	5.669127	3.244966	.9430421
	3	2.225528	3.799845	2.2262288	.9420196
	4	1.931005	3.392155	1.931683	.8998834
	5	2.364958	4.446056	2.3658446	
μ=.0003	1	5	10	5.003	.9514626
	2	3.144416	5.39738	3.145815	.9357017
	3	2.079541	3.393059	2.080559	.9305661
	4	1.806033	2.9870036	2.08693	.8250323
	5	2.323675	4.19311	2.324929	
μ=.0004	1	5	10	5.004001	.9524865
	2	3.211735	5.59364	3.213973	.9386605
	3	2.161946	3.662971	2.1634411	.9263953
	4	1.8850631	3.207022	1.851913	.8250323
μ=.0005	1	5	10	5.005001	.9516159
	2	3.244694	5.681146	3.24535	.9415714
	3	2.24664	3.78466	2.206568	.9281042
	4	1.879938	3.31896	1.881597	.8084613

$$\dot{x}(t) = Ax(t) + Cu(t) \quad x(0) = x_0 \quad 0 \le t \le 1, \qquad (1.18)$$

where, $x(t)$, $u(t) \varepsilon R^2$, $x(t)^T, u(t)^T$ *denote the transposes of* $x(t)$ *and* $u(t)$ *respectively*
P, Q *are 2 by 2 symmetric matrices*
A *and* C *are 2 by 2 and 2by 2 matrices not necessarily symmetric respectively.*

and

$$P = \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix}, \; Q = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \; A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \; C = \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix}, x_0 \begin{pmatrix} 1 \\ .5 \end{pmatrix}, u_0 = \begin{pmatrix} 1 \\ .5 \end{pmatrix}$$

RESULTS

As seen in Table 1, column 3, the scheme shows a good convergence profile, particularly at the 3rd iteration for μ=.0001 and at the 4th iteration for .0002≤μ≤.0005, where the iterates are 3.568452, 1.931005, 1.808033, 1.8850631, 1.879938 respectively. However, the first cycle at the 3rd iteration, is not comparable enough, since its iterate value 3.568452 is comparatively higher than any of the other cycles.

Column 6 shows the geometric convergence ratio profile for each μ per cycle. It is seen that $\frac{|z_{n+1} - z^*|}{|z_n - z^*|} < 1, \; for \; n \; l\arg e,$ as depicted in the convergence ratio column 6 of Table 1.

For μ=.0001, the geometric ratio convergence starting at .9483081 and ending at .82418, shows values lying between 0 and 1. This characterizes a geometric ratio convergence. Similarly, this same trend holds for $.0002 \le \mu \le .0005$ in column 6 of Table 1.

DISCUSSION

In this study, the scheme has demonstrated its objective having its geometric ratio convergence established between 0 and 1 as seen in Table 1. Hence, its success and reliability as it relates to problems of this class in terms of convergence profile and geometric convergence ratio via this sampled problem is being demonstrated.

REFERENCES

1. Olorunsola, S.A. and O. Olotu, 2004. A discretized algorithm for the solution of a constrained, continuous quadratic control problem. J. Nigeria Assoc. Math. Phys., Uni. Benin, Benin, 8: 295-300.
2. Olorunsola, S.A., 2004. Analysis of convergence for control problems governed by evolution equations inequality and equality constraints with multipliers Imbedded. Ife J. Sci., Obafemi Awolowo Uni., Ile-Ife, Nigeria, 6: 63.
3. Di pillo, G., L. Grippo and F. Lampariello, 1974. The Multiplier method for optimal control problems. Conf. Optimiz. Engg. Econom., Naples, ITALY
4. Glad, S.T., 1979. A combination of the penalty function and the multiplier methods for solving optimal control problems. J. Optimiz. Theor. Applic., 28: 303-329.
5. Ibiejugba, M.A., F. Otunta and S.A. Olorunsola, 1992. The role of the multiplier in then multiplier method. J. Math. Soc., 11, n.2. part 3.
6. Hasdorff, L., 1976. Gradient optimization and Nonlinear control. Wiley, New York.
7. Omolehin, J.O., 2006. Hessian spectrum to pertubation factor for gradient method algorithm. J. Nigerian Assoc. Math. Phys., 10: 355-362.
8. Shannon, D.F., 1978. On the convergence of new conjugate gradient algorithm, SIAM. J. Numer. Anal., 15: 1247-1257.