Journal of the Nigerian Association of Mathematical Physics
***Volume* 20 (March, 2012), pp 49 – 54**

Convergence Estimate of ECGM Algorithm for Reaction Diffusion Control Problem

[1] Omolehin J. O., [2]Rauf K. , [3]Badeggi A. Y. and [3]Aliyu A.I. Ma'ali
[1]Department of Mathematics, University of Ilorin, Ilorin, Nigeria.
[2]Mathematics Department, University of Ilorin, Ilorin, Nigeria.
[3]Mathematics/Computer Science Department, IBB University, Lapai, Nigeria.

Abstract

We consider the convergence rate of ECGM Algorithm for Reaction Diffusion Control Problem of the form

$$\text{Minimize } \int_0^t \{v_1^2(t) + \cdots + v_n^2(t) + u_1^2(t) + \cdots + u_n^2(t)\}\, dt$$

$$\text{Subject to } \dot{u}_i - \dot{v}_i = \bar{C}v_i(t) + \bar{D}u_i(t) \quad i = 1, 2, 3, \ldots, n$$

The work is based on the spectrum analysis of the associated control operator.

Keywords: Reaction diffusion, Convergence, Operator, Control, ECGM

1.0 Introduction

A quadratic functional is defined as:

$$F(x) = F_o + < a, x >_H + \frac{1}{2} < x, Ax >_H ,$$

Where A is an nxn symmetric positive definite matrix operator on the Hilbert space H. a and x are vectors in H and F_0 is a constant term.

The term conjugate descent with F is assumed that a sequence

$$\{p_i\} = p_0, p_1, \ldots, p_k, ..$$

is available with the members of the sequence conjugate with respect to the positive definite linear operator A.

By conjugate with respect to A, we mean that

$$< p_i, Ap_j >_H = \begin{cases} \neq 0, & if\ i \neq j \\ = 0, & if\ i = j \end{cases}$$

In this case, A is assumed positive definite so $< p_i, Ap_i >_H > 0$.

The conventional Conjugate Gradient Method (CGM) was originally designed for the minimization of a quadratic objective functional of the form stated above. Let us briefly define CGM Algorithm:

Stages involved in Conjugate Gradient Method

Stage 1: The first element $x_0 \in H$ of the descent sequence is guessed while the remaining members of the sequence are computed with the aid of the following formulae:

Stage 2: $p_0 = -g_0 = -(a + Ax_0)$

(p_0 is the descent direction and g_0 is the gradient of $F(x)$ when $x = x_0$)

Stage 3:

$x_{i+1} = x_i + \alpha_i p_i,\ \alpha_i = < g_i, g_i >_H / < p_i, Ap_i >_H$

$g_{i+1} = g_i + a_i Ap_i;$

α is the step length

$p_{i+1} = -g_{i+1} + \beta_i p_i;\quad \beta_i = < g_{i+1}, g_{i+1} >_H \ / < g_i, g_i >_H$

Stage 4: if g_i for some i terminate the sequence else, set $i = i + 1$ and go to stage 3.

The CGM has a well worked out theory with an elegant convergence profile [1]. It has been proved that the algorithm converges, at most, in n iterations in a well posed problem and the convergence rate is given as:

[1]Corresponding author: E-mail: omolehin_joseph@yahoo.com, Tel. +2348094335562

$$E(x_n) = \left\{ \frac{1 - \frac{m}{M}}{1 + \frac{m}{M}} \right\}^{2n} E(x_0)$$

Where m and M are smallest and spectrums of matrix A respectively.
That is, for an n dimensional problem, the algorithm will converge in at most n iterations.
The CGM algorithm cannot handle quadratic cost functional of the form:
Minimise

$$\int_0^T \{av^2(t) + bu^2(t)\}dt$$

Subject to

$$\dot{v} = cv(t) + du(t)$$

For the reason that operator A was not known explicitly for continuous cost functional. Researchers came up with different approximation – based techniques such that could estimate α_i that minimizes $F(x_i + \alpha p_i)$ See ref. [2]. In this fashion there came into being various cumbersome techniques to handle quadratic functional. Most popular among such methods is the conventional function space (CFS) algorithm due to Di Pillo et al. [3].
Since CGM algorithm cannot handle quadratic cost functional, Ibiejugba et.al developed an algorithm called Extended Conjugate Method algorithm (ECGM), which is based on the formalism of Conjugate Gradient Method (CGM).

THE ECGM ALGORITHM [4]

In other to determine the control operator A that satisfies the requirements for ECGM algorithm the following one-dimensional control problem was considered [4]:

$$\text{Minimize } \int_0^\delta \{(ax^2(t) + bu^2(t)\}\, dt$$

Subject to the dynamic constraint $\dot{x} = cx(t) + du(t)$, $o \leq t \leq \delta$,
Where a, b, c and d are constraints such that a, b > o while c and d are not necessarily positive. The problem can be transformed in to an unconstrained one dimensional cost problem as follows:

$$\langle z, AZ \rangle k = \int_0^\delta \{ (ax^2(t) + bu^2(t) + \mu(\dot{x}(t) - cx(t) - du(t)^2\}\, dt$$

Where μ (μ >o) is the penalty constant and x, u∈R.

It has been established that the control operator A associated with this one-dimensional, equality constrained problem, is such that it satisfies

$$Az(t) = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} = \begin{bmatrix} A_{11}x(t) + A_{12}u(t) \\ A_{21}x(t) + A_{22}u(t) \end{bmatrix},$$

Where A_{11} is such that

$$A_{11}x(t) = \mu(\dot{x}(0) - cx(0))sinht + \mu\int_0^t (\dot{x} - cx)(t)\cosh(t-s)\,ds - \int_0^t [(b + \mu c^T c)x - \mu cx](s)\sinh(t-s)\,ds + \{-\mu sinhT(x(0) + cx(0) + \mu\int_0^T (\dot{x} - cx)(s)\cosh(T-s)\,ds - \int_0^T [(b + \mu c^T c)x - \mu c\dot{x}](s)\sinh(T-s)\,ds\}\exp(T),$$

$0 \leq t \leq T$

$$A_{12}u(t) = au(t) + d^T du(t), \qquad 0 \leq t \leq T$$

$$A_{21} = \mu c^T dx(t) - \mu du(t), \qquad 0 \leq t \leq T$$

$$A_{22}u(t) = \mu u(s)sinht - \mu\int_0^t du(s)\cosh(t-s)\,ds + \int_0^T c^T\, dsinh(t-s)ds + \{\mu du(0)sinhT - \int_0^T \mu du(s)\cosh(t-s)\,ds + \mu\int_0^T du(s)\sinh(T-s)\,ds\}\exp(T), \quad 0 \leq t \leq T$$

We can now obtain the desired AP_i to apply when calculating $g_{i+1} = g_i + \alpha_i AP_i$ in the ECGM algorithm which would hence forth allow us to exploit the simplicity of the Conjugate Gradient Method algorithm (CGM) for continuous cost functional.

Thus, on adopting the "abused" notations.

$J_{x,i} = J_x\,(x_i, u_i, \mu)$,
$J_{u,i} = J_u\,(x_i, u_i, \mu)$
$P_{x,i} = P_x\,(x_i, u_i, \mu)$,
$J_i = J\,(x_i, u_i, \mu)$, (1.26)

$P_x\,(x_i, (t), u_i\,(t), \mu) \quad = \quad \int_o^t \{J_x\,(x_i\,(s), u_i\,(s), \mu)\}ds$,

$P_u\,(x_i, (t), u_i\,(t), \mu) \quad = \quad \int_o^t \{J_u\,(x_i\,(s), u_i\,(s), \mu)\}\,ds$,

$\int_o^t \{c^T du\,(s)\,\mathrm{Sin\,h}\,(\delta\text{-}s)\}ds\,\exp(\delta)$, we can now describe the ECGM Algorithm.

Steps involved in ECGM algorithm:
Step 1

It involves guessing the first sequence x_0. The remaining members of the sequence are then calculated as follows:
Step 2

$$p_0 = -g_0$$

(p_0 is the descent direction and g_0 is the gradient of the cost functional when $x = x_0$)

Step3

$$x_{i+1} = x_i + \alpha_i P_{x,i}$$
$$u_{i+1} = u_i + \alpha_i P_{u,i}$$
$$g_{x,i+1} = g_{x,i} + \alpha_i AP_{x,i}$$
$$g_{u,i+1} = g_{u,i} + \alpha_i AP_{u,i}$$
$$P_{x,i+1} = -g_{x,i+1} + \beta_i P_{x,i}$$
$$P_{u,i+1} = -g_{u,i+1} + \beta_i P_{u,i},$$

Where

$$\alpha_i = \frac{\langle g_i, g_i\rangle}{\langle P_i, AP_i\rangle} \text{ and } \beta_i = \frac{\langle g_{i+1}, g_{i+1}\rangle}{\langle g_i, g_i\rangle}$$

$$AP_i = \Bigg(-\mu \sinh t\big(J_{x,0} - cp_{x,0}\big) + \mu \int_0^t (cj_i - cp_{x,i}) \cosh(t-s)\,ds - \int_0^t [q + \mu c^T c]P_{x,i} - \mu c J_{x,i} \sinh(t-s)\,ds$$
$$+ \Big\{\mu \sinh\sigma\big[-j_{x,0} - cp_{x,0}\big]$$
$$+ \mu \int_0^T \big(J - cp_{x,i}\big)\cosh(\sigma - s)\,ds - \int_0^t [q + \mu c^T c]p_{x,i} - \mu c j_{x,i}]\sinh(\delta - s)ds\Big\} \exp(\sigma) + cu_i(t)$$
$$+ \mu d^T du(t);\; \mu\big[c^T dp_{x,i} - dJ_{x,i}\big]$$
$$+ \mu\Big\{Du(0)\mathrm{Sinh}t - \int_o^t Du_i\,(s)\cosh(\delta - s)\,ds + \mu \int_O^\delta c^T\,Du_i(s)\mathrm{Sinh}(\sigma - s)ds\Big\}\exp(\delta)\Bigg),$$

and where we have used the following notations:

$$J_i = J(x_i, u_i, \mu)$$
$$J_{x,i} = J_x(x_i, u_i, \mu)$$
$$J_{u,i} = J_u(x_i, u_i, \mu)$$
$$p_{x,i} = p_x(x_i, u_i, \mu)$$
$$p_{u,i} = p_u(x_i, u_i, \mu)$$
$$p_{u,i} = p_u(x_i, u_i, \mu)$$
$$p_x(x_i, u_i, \mu) = \int_o^t J_x\,(x_i(s), u_i(s), \mu)ds$$

and

$$p_u(x_i, u_i, \mu) = \int_o^t J_u\,(x_i(s), u_i(s), \mu)ds$$

Step 4
If g_i satisfies the tolerance, for some i terminate the sequence else, set $i = i + 1$ and go to step 3.

Main Result
Consider

Minimize $\int_0^t \{v_1^2(t) + \cdots + v_n^2(t) + u_1^2(t) + \cdots + u_n^2(t)\}\,dt$

Subject to $\dot{u}_i - \dot{v}_i = \bar{C}v_i(t) + \bar{D}u_i(t) \quad i = 1,2,3,\ldots,n$ (1)

Problem (1) is transformed into a penalized cost functional with a penalty constant μ:
Minimize

$$\int_0^t \{(v_1^2(t) + \cdots + v_n^2(t) + u_1^2(t) + \cdots + u_n^2(t) + \mu\|\dot{u}_i(t) - v_i - cv_i(t) - \bar{d}u_i(t)\|)^2\}\,dt \qquad (2)$$

Equation (2 can be put into the following equivalent form:
Minimize $\int_0^t \{(v_1^2(t) + \cdots + v_n^2(t) + u_1^2(t) + \cdots + u_n^2(t) + \mu\|\dot{u}_i(t) - v_i - cv_i(t) - du_i(t)\|)^2\}\,dt$
= Minimize $\int_0^t \left\{\left(v(t), \dot{v}(t), u(t), \dot{u}(t)\right)B\left(v(t), \dot{v}(t), u(t), \dot{u}(t)\right)^T\right\}dt$ (3)

Where

$$B = \begin{pmatrix} a + \mu c^2 & \mu & \mu cd & -\mu \\ \mu & \mu & \mu d & -\mu \\ \mu cd & \mu d & b + \mu d^2 & -\mu \\ -\mu & \mu & -\mu d & \mu \end{pmatrix}$$

Define

$$H(\mu) = \begin{pmatrix} \frac{a}{\mu} + c^2 & 1 & cd & -1 \\ 1 & 1 & d & -1 \\ cd & d & \frac{b}{\mu} + d^2 & -1 \\ -1 & 1 & -d & 1 \end{pmatrix}$$

Then

$$B = \mu H(\mu).$$

Denote the eigenvalues of constant matrix operator $H(\mu)$ by $\lambda_1, \lambda_2, \lambda_3, \lambda_4,$ then the eigenvalues of B are $\mu\lambda_1, \mu\lambda_2, \mu\lambda_3, \mu\lambda_4,$ respectively because $B = \mu H$. The limit of $H(\mu)$ is taken as $\mu \to \infty$.
i.e

$$\lim_{\mu\to\infty} H(\mu) = \lim_{\mu\to\infty} \begin{pmatrix} \frac{a}{\mu} + c^2 & 1 & cd & -1 \\ 1 & 1 & d & -1 \\ cd & d & \frac{b}{\mu} + d^2 & -1 \\ -1 & 1 & -d & 1 \end{pmatrix}$$

$$= \begin{pmatrix} c^2 & 1 & cd & -1 \\ 1 & 1 & d & -1 \\ cd & d & d^2 & -1 \\ -1 & 1 & -d & 1 \end{pmatrix}$$

= H.

Given that $H(\varepsilon) \to H \ as \ \varepsilon \to 0, where \ \mu^{-1} = \varepsilon.$
It implies that $\frac{M_{H(\varepsilon)}}{m_{H(\varepsilon)}} \to \frac{M_{\dot{H}}}{m_{\dot{H}}}$ as $\varepsilon \to 0$
due to the fact that zeros of a polynomial are continuous functions of its coefficients. M_H and m_H denote respectively, the largest and smallest eigenvalues of the matrix operator H. we now proceed to find the eigenvalues as follows:

$$\begin{vmatrix} c^2 - \lambda & 1 & cd & -1 \\ 1 & 1 - \lambda & d & -1 \\ cd & d & d^2 - \lambda & -1 \\ -1 & 1 & -d & 1 - \lambda \end{vmatrix} = 0 \qquad (4)$$

Expanding equation (2.4), we obtain

$$(c^2-\lambda)\left\{(1-\lambda)\begin{vmatrix} d^2-\lambda & -1 \\ -d & 1-\lambda \end{vmatrix} - d\begin{vmatrix} d & -1 \\ -1 & 1-\lambda \end{vmatrix} - \begin{vmatrix} d & d^2-\lambda \\ -d & -1 \end{vmatrix}\right\}$$
$$-\left\{\begin{vmatrix} d^2-\lambda & \lambda \\ -d & 1-\lambda \end{vmatrix} - d\begin{vmatrix} cd & -1 \\ -1 & 1-\lambda \end{vmatrix} - \begin{vmatrix} cd & d^2-\lambda \\ -1 & -d \end{vmatrix}\right\}$$
$$+\left\{cd\begin{vmatrix} d & -1 \\ -1 & 1-\lambda \end{vmatrix} - (1-\lambda)\begin{vmatrix} cd & -1 \\ -1 & 1-\lambda \end{vmatrix} - \begin{vmatrix} cd & d \\ -1 & -1 \end{vmatrix}\right\}$$
$$\left\{\begin{vmatrix} d & d^2-\lambda \\ -1 & -d \end{vmatrix} - (1-\lambda)\begin{vmatrix} cd & d^2-\lambda \\ -1 & d \end{vmatrix} - d\begin{vmatrix} cd & d \\ -1 & -1 \end{vmatrix}\right\} = 0.$$
$$\Rightarrow (c^2-\lambda)\{(1-\lambda)[(d^2-\lambda)(1-\lambda)-d] - d[(1-\lambda)-1] - (d^2+d^2-\lambda)\}$$
$$-\ \{(d^2-\lambda)(1-\lambda) - d - d[cd(1-\lambda)-1] - [-d^2c+d^2-\lambda]\} +\ cd\{d(1-\lambda) - 1 - (1-\lambda)[cd(1-\lambda) - 1] - [-cd+d]\} + \{-d^2+d^2-\lambda-(1-\lambda)[-cd^2+d^2-\lambda] - d[-cd+d]\} = 0$$
$$\Rightarrow\ c^2d - \lambda c^2 d + \lambda^2c^2d + 2\lambda^2c^2 - \lambda^3c^2 - c^2d^2 + \lambda c^2d^2 - \lambda d + \lambda d^2 - \lambda^3 - 2\lambda^3 + \lambda^4 + \lambda d^2 - \lambda^2 d^2 + \lambda d^2 - \lambda^2 - \lambda cd^2 - \lambda cd^2 + 2\lambda c^2d^2 - \lambda^2c^2d^2 - \lambda cd - \lambda cd^2 + \lambda d^2 - \lambda^2 = 0$$
(5)

Expanding and rearranging (5), we obtain

$$\lambda^4 + \lambda^3(-c^2-d-2) + \lambda^2(c^2d + 2c^2 + d - d^2 - 2 - c^2d^2)$$
$$+\lambda(-c^2d + c^2d^2 - d - 2d^2 - 3cd^2 + 2c^2d^2 - cd + d^2)$$
$$+c^2d + c^2d^2 = 0. \qquad (6)$$

It is the roots of this last quadratic equation that will give us the required eigenvalues. Using the results due to Abramowitz and Stegun for the solution of quadratic equation [5]

let

$$a_3 = -(c^2+d+2)$$
$$a_2 = (c^2d + 2c^2 + d - d^2 - 2 - c^2d^2)$$
$$a_1 = (-c^2d + c^2d^2 - d + 2d^2 - 3cd^2 + 2c^2d^2 - cd + d^2).$$
$$a_0 = c^2d - c^2d^2. \qquad (7)$$

Then equation (1.6) becomes

$$\lambda^4 - a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = 0. \qquad (8)$$

Find the real root u_1 of the cubic equation $u^3 - a_2u^2 + (a_1a_3 - 4a_0)u - (a_1^2 + a_0a_3^2 - 4a_0a_2) = 0$

i.e

$$u_1 = (n_1+n_2) - \tfrac{a_2}{3}$$

where

$$n_1 = \left[r + (q^3+r^2)^{1/2}\right]^{1/3}, \quad n_2 = \left[r - (q^3+r^2)^{1/2}\right]^{1/3}.$$

and

$$q = \tfrac{1}{3}(a_1a_3 - 4a_0) - \tfrac{1}{9}$$
$$r = \tfrac{1}{6}\left(-a_2(a_1a_3 - 4a_0) - 3(a_1^2 + a_0a_3^2 - 4a_0a_2) + \tfrac{1}{27}a_2^3\right)$$

Therefore

$$n_{1,2} = \left[\frac{1}{6}(c^2d^2 + c^2d + 2d^2 - 3cd^2 + 2d)(2c^2 + d^2 - c^2d^2 - d - 3) - 3(c^2d^2 + 2cd^2 - cd - 2d) - \frac{1}{27}(2c^2 + d^2 - c^2d^2 - 3)\right.$$
$$\left.\pm\left(\left(\tfrac{1}{3}(c^2d^2 + c^2d + 2d^2 - 3cd^2 - 3cd + 2d) - \tfrac{1}{9}\right)^3 + \tfrac{1}{6}\left((c^2d^2 + c^2d + 2d^2 - 3cd^2 - 3cd + 2d) - 3(c^2d^2 + 2cd^2 - cd - 2d)\right) - \tfrac{1}{27}(2c^2 + d^2 - c^2d^2 - d - 3)^2\right)^{1/2}\right]^{1/3} \qquad (9)$$

The roots of equation (8) are as follows:

$$\lambda_{1,2} = \frac{\left(\frac{a_3}{2} + \left(\frac{a_3^2}{4} + u_1 - a_2\right)^{1/2}\right)}{4\left(\frac{u_1}{2} + \left(\left(\frac{u_1}{2}\right)^2 - a_0\right)^{1/2}\right)} \pm \sqrt{\frac{a_3}{2} + \frac{1}{2}\left(\frac{a_3^2}{4} + u_1 - a_2\right)^{1/2}} \qquad (10)$$

$$\lambda_{3,4} = \frac{\left(\frac{a_3}{2} - \left(\frac{a_3^2}{4} + u_1 - a_2\right)^{1/2}\right)}{4\left(\frac{u_1}{2} + \left(\left(\frac{u_1}{2}\right)^2 - a_0\right)^{1/2}\right)} \pm \sqrt{\frac{a_3}{2} + \frac{1}{2}\left(\frac{a_3^2}{4} + u_1 - a_2\right)^{1/2}} \qquad (11)$$

For these eigenvalues $\lambda_i\ (i = 1,2,3,,4)$ to have real values,

$$u_1 > a_2 \qquad a_0 < a_2. \qquad (12)$$

Hence, the estimated value of the convergence rate of ECGM algorithm for reaction diffusion problem is given as:

$$J(v,u,\mu) = \left\{\frac{1-\frac{\min(\lambda_i)}{\max(\lambda_i)}}{1+\frac{\min(\lambda_i)}{\max(\lambda_i)}}\right\}^{2n} J(v_0,u_0,\mu) \qquad (13)$$

Conclusion

The spectrum analysis of our control problem has allowed us to estimate the convergence rate of the ECGM algorithm for our control problem as given in equation (13).

Therefore, research will continue on the numerical solution to our control problem..

References

[1] Omolehin J. O. (1991); *On the control of reaction diffusion equation.* University of Ilorin, Ilorin, Nigeria, PhD Thesis.

[2] Ibiejugba M. A. (1980); *Computing methods in optimal control theory*. University of Leeds, Leeds, England, PhD thesis.

[3] Ibiejugba, M. A. and Onumanyi, P. (1994). *On control operator and some of its application,* J. Math. Analysis & Applic. 103, 31 - 47

[4] J.O**.**Omolehin (2006),"Sine transformation for reaction diffusion problem", Analele stiintifice ale Universitatii, A1.I. Cuza lasi, Tomul LII, S.I, Matimatica, f.pp.105-112

[5] Abramowitz, M. and Stegun, I. A (1970); *Handbook of Mathematical Functions,* New York: Dover publications, Inc.

[6] Gelfand, I.M and Fomin, S. Y. (1963); *calculus of Variation,* Eaglewood, N. J.: Prentice-Hall.

[7] Hill, J. M (1980); *An integral arising in double diffusion theory*. J. Math. Analysis & Applc. 77, 278 – 287.

[8] Hill, J. M. (1981); *On the solution of reaction diffusion equation,* IMA Journal of Applied Mathematics, 27, 177 – 194.