Quest Journals
Journal of Research in Applied Mathematics
Volume 2~ Issue 5 (2015) pp: 07-13
ISSN(Online) : 2394-0743 ISSN (Print):2394-0735
www.questjournals.org

Research Paper

On The Justification of the Extended Conjugate Gradient Method (ECGM) Algorithm for Optimal Control Problems

J.S. APANAPUDOR and F.M. ADERIBIGBE
Dept of mathematics and computer science delta state university abraka
Dept of mathematical sciences ekiti state university ado – ekiti

Received 31 January, 2015; **A**ccepted 25 June, 2015

ABSTRACT:-The paper examines the Extended Conjugate Gradient Method (ECGM) algorithm for Discrete optimal Control Regulator Problems with special attention on the sequence of gradients and search directions generated. It considered the sequence of gradients generated from the ECGM algorithm for DOCP and observed that by taking the inner product of any two alternate gradients (i.e. $< g_i, g_j >, i \neq j$), we were able to show that the gradients are orthogonal. Furthermore, we showed that all the search directions emanating from this algorithm are as well conjugate with respect to the control operator H.These characteristics enhances the applicability of the algorithm to various problems on Optimal control.

I. INTRODUCTION

The conjugate gradient algorithm requires the computation of $\nabla f(x)$, the gradient of the function at each iteration to determine the direction descent. Each component of $\nabla f(x)$ is about as cumbersome as f(x) since this demands computing n + 1 function values at each iteration. Powell(1964) introduced a technique based on the observation that if x_1 and x_2 are obtained by one – dimensional searches on the samedirection p but from different base points, then in case of a quadratic function the vector x_2 - x_1 is mutually conjugate to p. Thus we have

$$\langle p, A(x_2 - x_1)\rangle = \langle p, g_2 - g_1\rangle = 0, \qquad (1)$$

where $g_i = \nabla f(x_i)$, A is a matrix (second partial derivatives)

Powell approach is an improvement on the one – dimensional search technique. Attempts to generate conjugate search directions has metamorphosed into several techniques such as Fletch – Powell variable metric algorithm, the Conjugate gradient methods in R^n, the Extended Conjugate Gradient Method algorithm. Aderibigbe and Apanapudor(2014) proposed some measures of generating the search directions as demanded in the ECGM algorithm. The ECGM algorithm as proposed by Ibiejugba and Onumanyi(1984) is as follows:

1. Initialize guess $z(0) = (x(0), u(0))^T \in w$

   Set $p_{x,0} = -g_{x,0}, p_{u,0} = -g_{u,0}$ (2)

2. Compute $x_{i+1} = x_i + \alpha_i p_{x,i}$ (3a)

$$u_{i+1} = u_i + \alpha_i p_{u,i} \qquad (3b)$$

$$g_{x,i+1} = g_{x,i} + \alpha_{x,i} A p_{x,i} \qquad (4a)$$

$$g_{u,i+1} = g_{u,i} + \alpha_{u,i} A p_{u,i} \qquad (4b)$$

$$p_{x,i+1} = -g_{x,i+1} + \beta_{x,i} p_{x,i} \qquad (5a)$$

$$p_{u,i+1} = -g_{u,i+1} + \beta_{u,i} p_{u,i} \qquad (5b)$$

*Corresponding Author: J.S. APANAPUDOR
Dept of mathematics and computer science delta state university abraka

where $\alpha_i = \dfrac{\|g_i\|^2}{\langle p_i, Ap_i \rangle} = \dfrac{\langle g_i, g_i \rangle}{\langle p_i, Ap_i \rangle}$ (6)

$$\beta_i = \frac{\|g_{i+1}\|^2}{\|g_i\|^2} = \frac{\langle g_{i+1}, g_{i+1} \rangle}{\langle g_i, g_i \rangle} \quad (7)$$

p_i is the descent search direction at the ith step, $g_i = \begin{pmatrix} g_{x,i} \\ g_{u,i} \end{pmatrix}$, $p_i = \begin{pmatrix} p_{x,i} \\ p_{u,i} \end{pmatrix}$

One of the techniques for generating the search directions as proposed by Aderibigbe and Apanapudor(2014) and starting with $p_0 = -g_0$, is given by

$$p_k = \langle g_k, g_k \rangle M \quad (8)$$

where M is the sum from $j = 0, to, k$,of the ratios of the jth gradient to the inner product of jth gradient with itself.

The rest of this paper will consist of section 2 which dwells on some lemma and theorems that will be useful to our work. Section three discusses the main results of demonstrating that the gradients generated by the ECGM algorithm are orthogonal. Furthermore we show that the search directions also generated are conjugate with respect to a given matrix operator. Finally section four concludes the discussion by specifying the contributions of this paper to knowledge.

II. SOME USEFUL TOOLS FOR RELEVANT FOR THIS PAPER

The following tools will be relevant in the ensuing sections and so are in order. These tools are enshrined in the ensuing theorems, lemma and corollary.

TheoremPolak(1971): Let G be a symmetric, positive definite n x n matrix and $g_0 \in R^n$ be arbitrary. Suppose that for $i = 0,1,2,\cdots$

$$g_{i+1} = g_i - \lambda_i G h_i \quad (9a)$$

$$h_{i+1} = g_{i+1} + \gamma_i h_i \quad (9b)$$

with $h_0 = g_0$,where λ_i, γ_i are chosen so that

$$\langle g_{i+1}, g_i \rangle = 0$$

$$\langle h_{i+1}, Gh_i \rangle = 0 \quad (10)$$

for $\lambda_i = \dfrac{\langle g_i, g_i \rangle}{\langle g_i, Ghi \rangle}, \gamma_i = \dfrac{-\langle Gh_i, g_{i-1} \rangle}{\langle Gh_i, h_i \rangle}$ (11)

whenever the denominators are not zero and $\lambda_i = 0, \gamma_i = 0, otherwise$.

Then

$$\langle g_i, g_i \rangle = \delta_{ii} \|g_i\|^2$$

(12a)

$$\langle h_i, Gh_j \rangle = \delta_{ij} \langle h_i, Gh_i \rangle$$

(12b)

where δ_{ij} is the kronecker symbol and $g_i = h_i = 0$, for all $i > m, with, m \le n-1$.

Proof:

Suppose the sequence $g_0, g_{1,} g_2, \cdots, g_m$ and $h_0, h_1, h_2, \cdots, h_m$ are nonzero vectors generated from equation (9) after m iterations. Assume that g_{m+1}, generated from (9a) is zero; then we have from (9a) and (9b) that $h_{m+1} = 0$; $h_{m+1} = 0$ thus $g_{m+j} = h_{m+j} = 0,\ for,\ j \ge 1$.

Now suppose $h_{m+1} = 0$, we show that this implies $g_{m+1} = 0$. To achieve this, we must show that $\langle h_m, g_{m+1} \rangle = 0$. By construction and since $h_0 = g_0$, $\langle h_0, g_1 \rangle = 0$.So let us assume that $\langle h_{j-1}, g_j \rangle = 0$, for $j \in \{1, 2, 3, \cdots, m\}$. Then

$$\langle h_j, g_{j+1} \rangle = \langle h_j, g_j - \lambda_j G h_j \rangle \quad (13)$$
$$= \langle h_j, g_j \rangle - \lambda_j \langle h_j, G h_j \rangle$$
$$= \langle h_j, g_j \rangle - \lambda_j \langle h_j, G h_j \rangle$$
$$= \langle g_j + \gamma_{j-1} h_{j-1}, g_j \rangle - \lambda_j \langle h_j, G h_j \rangle$$
$$= \langle g_j, g_j \rangle - \lambda_j \langle h_j, G h_j \rangle = 0$$

since $\langle g_j, G h_j \rangle = \langle h_j - \gamma_{j-1} h_{j-1}, G h_j \rangle = \langle h_j, G h_j \rangle$

From induction, we have that $\langle h_m, g_{m+1} \rangle = 0,$ holds. Since h_m and g_{m+1} are orthogonal and that $h_{m+1} = 0$, we conclude from equation (9a) that

$0 = \|g_{m+1}\|^2 + \gamma \langle g_{m+1}, h_m \rangle$, i.e. that for $h_{m+1} = 0$, we must have $g_{m+1} = 0$. Thus we draw our conclusion that the equations (9) will result in two sequences $g_0, g_{1,} g_2, \cdots, g_m$ and $h_0, h_1, h_2, \cdots, h_m$ of vectors such that for some $m \geq 0$ both g_i and h_i are nonzero and for $0 \leq i \leq m$ both g_i and h_i zero for all $i > m$.

Suppose that the integer m is that $g_i \neq 0$ and $h_i \neq 0$ for all $0 \leq i \leq m$ and $g_i = h_i = 0$ for all $i > m$. Obviously the relations in (12) are satisfied trivially whenever $i > m$ or $j > m$. Hence we need only consider the case. The proof of this is presented inductively. We have by Constructionthat , $\langle g_0, g_1 \rangle = 0$ and $\langle h_0, G h_1 \rangle = 0$. Suppose that for some integer $0 \leq k \leq m$,

$$\langle g_i, g_j \rangle = \langle h_i, G h_j \rangle = 0$$

(14a)

for all $i \neq j, 0 \leq i, j \leq k$,

Let $i \in \{1, 2, 3, \cdots, k-1\}$. Then

$$\langle g_{k+1}, g_i \rangle = \langle g_k - \lambda_k G h_k, g_i \rangle$$
$$= -\lambda_k \langle G h_k, g_i \rangle$$
$$= -\lambda_k \langle G h_k, h_i - \gamma_{i-1} h_{i-1} \rangle = 0$$

Also,

$\langle g_{k+1}, g_k \rangle = 0$ by the choice of λ_k

(14b)

and $\langle g_{k+1}, g_0 \rangle = \langle g_k - \lambda_k G h_k, g_0 \rangle = -\lambda_k \langle G h_k, g_0 \rangle = 0$ (15)

Similarly,

$\langle h_{k+1}, G h_k \rangle = 0,$ by the choice of γ_k and for $i \in \{1, 2, 3, \cdots, k-1\}$.

$$\langle h_{k+1}, G h_i \rangle = \langle g_{k+1} + \gamma_k h_k, G h_i \rangle \quad (16)$$
$$= \langle g_{k+1}, G h_i \rangle = \langle h_{k+1}, G h_i \rangle$$
$$= \left\langle g_{k+1}, -\frac{1}{\lambda_k} \langle g_{i+1} - g_i \rangle \right\rangle = 0$$

since $\lambda_i \neq 0$, for $i = 0, 1, 2, \cdots, m.$

Now (14) is true for $k=1$ and thus (12) must be true for all $0 \le i, j \le m$, We conclusively say that the vectors $g_0, g_1, g_2, \cdots$ are orthogonal to each other and since they are nonzero, their total number cannot exceed n.

Corollary

Suppose that $g_0, g_{1,} g_2, \cdots, g_m$ and $h_0, h_1, \cdots, h_m$ are nonzero vectors generated by (9). Then

(i) $\langle h_i, g_k \rangle = 0$, for all $0 \le i < k \le n$. (17)

(ii) $\lambda_i = \dfrac{\langle g_i, g_i \rangle}{\langle g_i, Gh_i \rangle} = \dfrac{\langle h_i, g_i \rangle}{\langle h_i, Gh_i \rangle}, i = 0,1,2,\cdots,k.$ (18)

(iii) $\gamma_i = \dfrac{\langle Gh_i, g_{i+1} \rangle}{\langle Gh_i, h_i \rangle} = \dfrac{\langle g_{i+1}, g_{i+1} \rangle}{\langle g_i, \mathrm{g}_i \rangle} = \dfrac{\langle g_{i+1}, g_{i+1} \rangle \pm \langle g_{i+1}, g_i \rangle}{\langle g_i, g_i \rangle}, i = 0,1,2,\cdots,\mathrm{n}.$ (19)

Proof: To prove (i) we note that for all $0 \le i < k \le n$,

$$\langle h_i, g_k \rangle = \langle h_i, g_{k-1} - \lambda_{k-1} Gh_{k-1} \rangle = \langle h_i, g_{k-1} \rangle = \cdots = \langle h_i, g_{i+1} \rangle = 0. \quad (20)$$

Since the last equality in equation (20) has been established in equation (13) in the theorem above.
Nextto prove (ii) we adopt the following procedure:

$$-\frac{\langle Gh_i, g_{i+1} \rangle}{\langle Gh_i, h_i \rangle} = \frac{-\langle g_{i+1} - g_i, g_{i+1} \rangle}{\langle g_{i+1} - g_i, h_i \rangle}$$

$$= \frac{\langle g_{i+1}, g_{i+1} \rangle}{\langle g_i, h_i \rangle} = \frac{\langle g_{i+1}, g_{i+1} \rangle}{\langle g_i, \mathrm{g}_i + \gamma_{i-1} h_{i-1} \rangle} = \frac{\langle g_{i+1}, g_{i+1} \rangle}{\langle g_i, g_i \rangle}$$

$$= \frac{\langle g_{i+1}, g_{i+1} \rangle \pm \langle g_{i+1}, g_i \rangle}{\langle g_i, g_i \rangle} \quad (21)$$

The nonzero vectors considered in the above theorem and corollary are the gradients and search directions respectively. We intend to use the ideas enshrined in the theorem and corollary to prove that the gradients are orthogonal to each other; while the search directions are conjugate with respect to the matrix operator H of order 2(k+1).

III. MAIN RESULT

Our results in this paper are contained in the following theorem.

3.1 Theorem

The gradients $g_0, g_{1,} g_2, \cdots, g_{n-1}$ and search directions $p_0, p_1, p_2, \cdots, p_{n-1}$ generated by the Extended Conjugate Gradient Method(ECGM) algorithm for discrete optimal control problems satisfy

$$\langle g_i, g_j \rangle = 0, i \ne j \quad (23)$$

then all gradients are orthogonal and

$$\langle p_i, Hp_j \rangle = 0, i \ne j, \quad (24)$$

i.e. all descent search directions are conjugate

Proof: The proof is presented by induction. Consider first the proof that g_0 and g_1 are orthogonal and p_0 and p_1 are conjugate w.r.t H. Then

$\langle g_0, g_1 \rangle = \langle g_0, g_0 + \alpha_0 Hp_0 \rangle$, using (4) or (9)

$= \langle g_0, g_0 \rangle + \alpha_0 \langle g_0, Hp_0 \rangle$

$= \langle g_0, g_0 \rangle - \alpha_0 \langle g_0, Hg_0 \rangle, p_0 = -g_0$

In line with (11) i.e. by the choice of $\alpha_0 = \dfrac{\langle g_0, g_0 \rangle}{\langle p_0, Hp_0 \rangle}$, we have $\langle g_0, g_1 \rangle = 0$ (25)

Thus it is a fact that g_0 and g_1 are orthogonal.

Similarly, consider the inner product,

$$\left\langle p_0, Hp_1 \right\rangle = \left\langle Hp_0, p_1 \right\rangle = \left\langle -Hg_0, -\mathrm{g}_1 - \beta_0 g_0 \right\rangle$$

$$= \left\langle Hg_0, \mathrm{g}_1 \right\rangle + \beta_0 \left\langle Hg_0, g_0 \right\rangle$$

With $\beta_0 = \dfrac{\left\langle g_1, g_1 \right\rangle}{\left\langle g_0, g_0 \right\rangle}$, we have,

$$= \left\langle Hg_0, \mathrm{g}_1 \right\rangle + \frac{< g_1, g_1 >}{< g_0, g_0 >} \left\langle Hg_0, g_0 \right\rangle \tag{25}$$

By rearranging (25), we obtain

$$= \left\langle Hg_0, \mathrm{g}_1 \right\rangle + < g_1, g_1 > \frac{\left\langle Hg_0, g_0 \right\rangle}{< g_0, g_0 >}$$

$$= \left\langle Hg_0, \mathrm{g}_1 \right\rangle + \frac{1}{\alpha_0} < g_1, g_1 > \tag{26}$$

From equations(7) and (11), we have

$g_1 = g_0 + \alpha_0 Hp_0 = g_0 - \alpha_0 Hg_0$ and by some rearrangement and substitution into (26) yields

$$\left\langle p_0, Hp_1 \right\rangle = \left\langle \frac{g_1 - g_0}{\alpha_0}, \mathrm{g}_1 \right\rangle + \frac{1}{\alpha_0} \left\langle g_1, g_1 \right\rangle$$

$$= 0, \text{ since } \left\langle \frac{g_0, g_1}{\alpha_0} \right\rangle = 0 \tag{26}$$

which indeed shows that p_0 , and p_1 are conjugate w.r.t. H.

Next, we assume that $g_0, g_1, g_2, \cdots, g_k$ are orthogonal and that $p_0, p_1, p_2, \cdots, p_{k-1}$ are conjugate w.r.t. H. To complete our proof we need only show that g_{k-1} and p_k can be added to these orthogonal and conjugate sets respectively. We start by showing p_k conjugate to $p_0, p_1, p_2, \cdots, p_{k-1}$. The following reminders are useful and thus are in order.

$$\left\langle g_i, g_j \right\rangle = 0, i \neq j, 0 \le i \le k, 0 \le j \le k \tag{27}$$

$$\left\langle p_i, Hp_j \right\rangle = 0, i \neq j, 0 \le i \le k-1, 0 \le j \le k-1 \tag{28}$$

Equation (7) or (11) empowered us to write

$$g_{i+1} = g_i + \alpha_i Hp_i$$

and so

$$\left\langle g_{i+1}, p_k \right\rangle = \left\langle g_i + \alpha_i Hp_i, p_k \right\rangle = \left\langle g_i, p_k \right\rangle + \alpha_i \left\langle p_k, Hp_i \right\rangle \tag{29}$$

Substituting (8) into (29)

$$\left\langle g_{i+1}, -\left\langle g_k, g_k \right\rangle \sum_{j=0}^{k} \frac{g_j}{\left\langle g_j, g_j \right\rangle} \right\rangle = \left\langle g_i + \alpha_i Hp_i, -\left\langle g_k, g_k \right\rangle \sum_{j=0}^{k} \frac{g_j}{\left\langle g_j, g_j \right\rangle} \right\rangle + \alpha_i \left\langle p_k, Hp_i \right\rangle \tag{30}$$

Using (28) the LHS of (30) lead us to,

$$\left\langle g_{i+1}, -\left\langle g_k, g_k \right\rangle + \left\langle g_{k-1}, g_{k-1} \right\rangle \sum_{j=0}^{k-1} \frac{g_j}{\left\langle g_j, g_j \right\rangle} \right\rangle = -\left\langle g_k, g_k \right\rangle$$ and the RHS of (30) gives

$$= -\left\langle g_k, g_k \right\rangle + \alpha_i \left\langle p_k, Hp_i \right\rangle, \; 0 \le i \le k-1 \tag{31}$$

Resolving the LHS and RHS, yields

$$+\alpha_i \left\langle p_k, Hp_i \right\rangle = 0, \; 0 \le i \le k-1 \tag{32}$$

But $\alpha_i = \dfrac{\langle g_i, g_i \rangle}{\langle p_i, Hp_i \rangle} \neq 0, 0 \leq i \leq k-1$

Since in (32) $\langle p_k, Hp_i \rangle = \langle p_i, Hp_k \rangle = 0, 0 \leq i \leq k-1$

(33)

Thus p_k is conjugate w.r.t. H to $p_0, p_1, p_2, \cdots, p_{k-1}$. We are therefore left to show that g_{k+1} is orthogonal to $g_0, g_1, g_2, \cdots, g_k$. With the assumption that $g_0, g_1, g_2, \cdots, g_k$ are orthogonal, we show that

$$\langle g_i, g_{k+1} \rangle = 0, i < k \tag{34}$$

Applying equation (4) or (9) produces,

$$\langle g_i, g_{k+1} \rangle = \langle g_i, g_k + \alpha_k Hp_k \rangle$$
$$= \langle g_i, g_k \rangle + \alpha_k \langle g_i, Hp_k \rangle \tag{35}$$

We shall examine equation (35) in two different cases:
Case I: i < k, equation (35) becomes

$$\langle g_i, g_{k+1} \rangle = \alpha_k \langle g_i, Hp_k \rangle \tag{36}$$

Hence is orthogonal to $g_0, g_1, g_2, \cdots, g_{k-1}$ by the assumption. Again, we have from equation (33),

$$\langle p_{i-1}, Hp_k \rangle = 0, i < k$$

Thus we can say that by the use of (36) and (34),

$$\langle g_i, g_{k+1} \rangle = \alpha_k \langle g_i - \beta_{i-1} p_{i-1}, Hp_i \rangle$$
$$= -\alpha_k \langle p_i, Hp_k \rangle$$
$$= 0, i < k. \tag{37}$$

which indicates the conjugacy of p_i and p_k for $i < k$..

Case II, $i = k$.
In this case, equation (35) yields,

$$\langle g_k, g_{k+1} \rangle = \langle g_k, g_k \rangle + \alpha_k \langle g_k, Hp_k \rangle \tag{38}$$

$$= -\langle g_k, g_k \rangle + \frac{\langle g_k, g_k \rangle}{\langle p_k, Hp_k \rangle} \langle g_k, Hp_k \rangle \tag{39}$$

Applying (4), we consider

$$\langle g_k, Hp_k \rangle = \langle g_k - \beta_{k-1} p_{k-1}, Hp_k \rangle, \tag{40}$$

g_k is being considered as p_k. i.e. $p_k = g_k - \beta_{k-1} p_{k-1}$; since p_{k+1} and p_k are conjugate w.r.t. H by (33). Thus applying (9) in (40) yields

$$\langle g_k, Hp_k \rangle = -\langle p_k, Hp_k \rangle \tag{41}$$

and so, we have

$$\langle g_k, g_{k+1} \rangle = 0.$$

This completes the proof of the theorem.

IV. CONCLUSION

This paper has been able to show that the sequence of gradients generated by the ECGM algorithm is orthogonal. Similarly the set of search directions produced by this same algorithm is conjugate with respect to H. These features coupled with the convergence of the sequence of iterates generated by ECGM algorithm for DOCP have justified the robustness and outstanding nature of the ECGM algorithm for DOCP amongst algorithms for optimal control problems.

REFERENCES

[1]. Aderibigbe, F.M. and Apanapudor, J.S.(2014) On the Extended Conjugate Gradient Method (ECGM) algorithm for discrete optimal control problems and some of its features,, IIOSR Journal of Mathematics, Vol. 10, Issue 3 version IV, pp. 16 – 22.
[2]. Chow, J.H., Frederick, D.K. and Chats, N.W.(2003) Discrete Time Control Problems, Brooks Cole
[3]. Thomson Learning Publishing Co. USA.
[4]. Macki, J. and Straus, A, (1980) Introduction to Optimal Control Theory, Springer – Verlag, N.Y.Inc.
[5]. Otunta, F.O.(1991) Optimization Techniques for a class of Regulator Problems, Ph.D. Thesis,University of Ilorin, Nigeria
[6]. Polak, E.(1971) Computational Methods in Optimization, Academic Press London.