تطوير خوارزمية المتجهات المترافقة المعدلة

عباس حسن تقي [1] ، أمل نشأة شاكر [2]

[1] كلية العلوم ، جامعة كركوك ، كركوك ، العراق

[2] قسم الرياضيات ، كلية التربية للعلوم الصرفة ، جامعة تكريت ، تكريت ، العراق

الملخص

في هذا البحث تم تطوير طريقة V1-CG لطرائق التدرج المترافق المعدلة , لزيادة سرعة تقاربها مع الاحتفاظ بخاصية التقارب الشامل حيث ان اشتقاق هذه الطريقة استند الى دالة تربيعية محدبة صارمة , وقد تم تطوير هذه الطريقة الى الدوال العامة التي فيها المشتقات العليا لاتساوي صفر. تم اثبات خاصية الانحدار الحاد والتقارب الشامل للخوارزمية المقترحة والتجارب العددية على بعض دوال الاختبار واظهرت لنا تحسن واضح على طريقة V1-CG المعدلة في هذا المجال.

1. المقدمة (Introduction)

جرى الاهتمام بطرائق التدرج المترافق لسببين , الاول هو ان هذه الطرائق هي من بين اقدم التقنيات المعروفة وافضلها لحل انظمة المعادلات الخطية ذات الابعاد الكبيرة , والسبب الثاني هو انه بالامكان تكييف هذه الطرائق لحل مسائل الامثلية اللاخطية. [9]

تتمتع هذه الطرائق بميزات تضعها بين طريقة الانحدار الحاد (SD) وطريقة نيوتن , لان هذه الطرائق تتطلب حساب المشتقات الاولى فقط , ولاتحتاج لحساب المشتقات الثانية وخزنها التي تحتاجها طريقة نيوتن , كما انها اسرع من طريقة الانحدار الحاد , أي انها تغلبت على التقارب البطيء لهذه الطريقة, وبما انها لاتحتاج لحساب مصفوفة هيسي او أي من تقريباتها , لذا فهي تستخدم بصورة واسعة لحل مسائل الامثلية ذات القياس الكبير . [13]

ويعرف نوعان لهذه الطرائق , النوع الاول هو طرائق التدرج المترافق الخطية ويعرف ايضاً بطرائق التدرج المترافق التربيعية , كما انه يعرف احياناً باسم طرائق التدرج المترافق النقية (Pure Conjugate Gradient), وتستخدم هذه الطرائق لتصغير الدوال التربيعية المحدبة. اما النوع الثاني فيعرف بطرائق التدرج المترافق اللاخطية , كما يعرف ايضاً بطرائق التدرج المترافق غير التربيعية (Nonqudratic Conjugate Gradient) , وتستخدم لتصغير الدوال العامة المحدبة او الدوال العامة اللاخطية .

اقترح طريقة التدرج المترافق الخطية لاول مرة الباحثان Hestenes و Stiefel سنة (1960) بوصفها طريقة تكرارية استخدمت في الاصل بديلاً لطريقة كاوس للحذف (Gaussian Elmination) لحل انظمة خطية كبيرة بمصفوفة معاملات موجبة التعريف على الحاسوب. [5]

وقام كل من Fletcher و Reeves بتطوير طريقة Hestenes و Stiefel وتقديم اول طريقة تدرج مترافق لاخطية سنة (1960). وهي واحدة من اوائل التقنيات المعروفة لحل مسائل الامثلية اللاخطية ذات الابعاد الكبيرة . وعلى مر السنين اقترحت الكثير من الطرائق المختلفة على غرار الصيغة الاصلية لهذه الطريقة , وبعض من هذه الطرائق ذات استخدام واسع في التطبيق. [9]

2. انواع طرائق التدرج المترافق

2.1 طرائق التدرج المترافق التقليدية (Classical Conjugate Gradient Methods)

طريقة التدرج المترافق الخطية هي طريقة تكرارية لحل مسألة التصغير

$$\text{Min } f(x) = ½\ x^T G x - b^T x + c \qquad (1)$$

اذ b يمثل متجه ثابت , و c تمثل قيمة ثابتة , و G هي مصفوفة متماثلة موجبة التعريف من النمط n × n . ويمكن الحصول على (1) بصورة مكافئة بوصفها نظاماً من المعادلات الخطية وعلى النحو الاتي :

$$G\,x = b \qquad (2)$$

هذا يجعل الحل الوحيد للمسألة (1) هو الحل نفسه للنظام(2) . وبذلك يمكننا ان نعد طريقة التدرج المترافق , اما خوارزمية لحل أنظمة خطية أو لايجاد القيمة الصغرى للدوال المحدبة التربيعية. [10]

نلاحظ ان التدرج (Gradient) للدالة f يساوي المتبقي (Residual) أي التدرج السالب (Negative Gradient) للنظام الخطي , أي ان

$$\nabla f(x) = G\,x - b = g(x) \qquad (3)$$

وعندما $x = x_k$

$$g_k = Gx_k - b \qquad (4)$$

من الصفات الجديرة بالملاحظة ان طريقة التدرج المترافق لها القابلية على توليد مجموعة متجهات لها خاصية الترافق (Conjugacy) , اذ يمكنها ان تحسب الاتجاه الجديد d_{k+1} باستخدام الاتجاه السابق d_k والتدرج الحالي g_k , ويختار بحيث يكون تركيباً خطياً من$(-g_{k+1}$ (اتجاه الانحدار الشديد) و d_k (الاتجاه السابق) ولانحتاج لمعرفة جميع الاتجاهات السابقة $d_0, d_1, \ldots, d_{k-1}$ للمجموعة المترافقة ; اذ d_{k+1} مترافق مع جميع الاتجاهات السابقة .وهذه الصفة تجعل هذه الطريقة تتطلب مساحة خزن وحسابات قليلة . نكتب

$$d_{k+1} = -\,g_{k+1} + \beta_{k+1} d_k \qquad (5)$$

اذ β_{k+1} تمثل كمية عددية , تحدد بحيث يكون d_{k+1} و d_k مترافقين بالنسبة للمصفوفة G .

نختار اتجاه البحث الاول عند النقطة الابتدائية $d_0 = -\,g_0$ (اتجاه الانحدار الشديد), [9] ان معدل التقارب لطرائق التدرج المترافق يكون خطياً مالم يسترجع التكرار . [11]

يمكن ان تتوسع طريقة التدرج المترافق لتشمل دوال لاخطية عامة باستخدام متسلسلة تايلر في تقريب الدالة الى الرتبة لدالة الهدف . وبالقرب من الحل تسلك هذه الدوال سلوكاً مشابهاً للدوال التربيعية [4].

خوارزمية (1) (طريقة التدرج المترافق التقليدية (Classical Conjugate Gradient Method)

الخطوة 1 : اختر قيمة ابتدائية x_0 , ضع $d_0 = - g_0$, $\varepsilon > 0$ و k = 0 .

الخطوة 2 : احسب طول الخطوة $\alpha_k > 0$ يحقق شرطي Wolfe

$f(x_k + \alpha d_k) \le f(x_k) + c_1 \alpha g_k^T d_k$ و

$g_{k+1}^T d_k \ge c_2 g_k^T d_k$

حيث ان $0 < c_1 < c_2 < 1$.

الخطوة 3 : احسب $x_{k+1} = x_k + \alpha_k d_k$. اذا كان $\|g_{k+1}\| < 0$, عندئذ توقف .

الخطوة 4 : احسب β_{k+1} وولد الاتجاه $d_{k+1} = - g_{k+1} + \beta_{k+1} d_k$.

الخطوة 5 : ضع k = k + 1 واذهب الى الخطوة 2 .

تختلف خوارزميات التدرج المترافق على وفق الاختيارات المختلفة للمعلمة β_{k+1} في الخطوة 4 . واكثر الصيغ المشهورة لخصت بالجدول الاتي :

الجدول ذو الرقم (1) الخيارات المختلفة للمعلمة β_{k+1} في خوارزميات التدرج المترافق التقليدية [1]

Hestenes – Stiefel (HS)1952	$\beta_{K+1}^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k}$
Fletcher- Reeves (RE) 1964	$\beta_{K+1}^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}$
Polak- Ribiere (PR) 1969	$\beta_{K+1}^{PR} = \frac{g_{k+1}^T y_k}{g_k^T g_k}$
Dai- Yuan (DY) 1999	$\beta_{k+1}^{DY} = \frac{g_{k+1}^T g_{k+1}}{d_k^T y_k}$

2.2 خصائص خوارزميات التدرج المترافق التقليدية Properties (of Classical Conjugate Gradient Algorithms)

تمتلك طرائق التدرج المترافق الخواص الاتية عندما دالة الهدف دالة تربيعية والبحث الخطي مضبوط

1. الترافق (Conjugacy) $d_i^T G d_j = 0$, j = 0,1, …. , i-1 .
2. التعامد (Orthogonality) $g_i^T g_j = 0$, j = 0,1, ….. , i-1 .
3. الانحدار (Descent) $d_i^T g_i = - g_i^T g_i = - \|g_i\|^2 < 0$.
4. خط البحث المضبوط (Exact Line Search, ELS) $g_{k+1}^T d_k = 0$, k ≥ 0 .
5. خاصية التوقف التربيعي (Quadratic Termination Property). [13]
6. تتطلب مضاعفات O(n) من العمليات الحسابية لكل تكرار للتقارب الى الحد الادنى (هذا يعني ان لديها خوارزميات التقارب الشامل) . [4]
7. $span\{d_0, d_1, \ldots, d_{k+1}\} = span\{g_0, g_1, \ldots, g_{k+1}\}$
$span\{g_0, G g_1, \ldots, G^k g_{k+1}\}$

ولهذا السبب تدعى طرائق التدرج المترافق بطريقة فضاء Krylov الجزئي (Krylov Subspace Method). [3]

نظرية التقارب الشامل لطرائق المتجهات المترافقة (Global Convergence of Descent Method)

قبل النظرية نحتاج الى التعريف التالي:

ليكن $f(x)$ دالة مستمرة ومشتقتها مستمرة ومنتظمة في المجموعة المحددة

$L = \{x \in R^n : f(x) \le f(x_1)\}$

وليكن θ الزاوية المحصورة بين اتجاه البحث d_k و $-g_k$ عندئذ

$\cos\theta = \frac{-g_k^T d_k}{\|d_k\|\|g_k\|}$

حيث ان d_k تحقق خاصية الانحدار وان المتتابعة المتولدة بـ $x_{k+1} = x_k + \alpha_k d_k$ تتقارب الى النقطة الحرجة أي ان

$g_k = 0 \quad or \quad f(x_k) \to -\infty \quad or \quad g_k \to 0$

اذا كانت $\theta \le \frac{\pi}{2} - \mu$ والبحث الخطي مضبوط . INT[1]

2.3 طرائق التدرج المترافق المعدلة Modified Conjugate Gradient Methods)

ان طرائق التدرج المترافقة المعدلة تعتبر من الطرائق المهمة في هذا المجال لانها ترتبط بعلاقة ضمنية مع طرائق الشبيهة بنيوتن والذي تمتلك سرعة تقارب تربيعية اذا كانت دالة الهدف دالة تربيعية وبحث خطي مضبوط وتكون سرعتها فوق الخطية في الدوال العامة الا انها تحتاج الى مصفوفة وعمليات حسابية كبيرة وللتغلب على المشاكل الموجودة في طرائق الشبيهة بنيوتن وزيادة سرعة تقارب طرائق المتجهات المترافقة تم استحداث خوارزميات التدرج المترافق المعدلة فعلى سبيل المثال Shanno في [12] استخدم صيغة BFGS شبيهة بنيوتن والمعرفة في المعادلة (6) حيث اعتبر $H_k = I_{n*n}$ وتوصل الى الطريقة الاتية :

$$d_{k+1} = - g_{k+1} + \frac{s_k^T g_{k+1}}{s_k^T y_k} y_k + \left[\frac{y_k^T g_{k+1}}{s_k^T y_k} - \left(1 + \frac{y_k^T y_k}{s_k^T y_k}\right) \frac{s_k^T g_{k+1}}{s_k^T y_k} \right] s_k \qquad (6)$$

وسمي هذه الطريقة BFGS معدوم الذاكرة (Memory Less).
وفي عام 2010 اقترح Khalil في اطروحته [7] طريقة اخرى من طرائق التدرج المترافق وسماها V1-CG وعرفها بالشكل الاتي

$$\beta^{v1} = \left(1 - \frac{s_k^T y_k}{y_k^T y_k}\right) \frac{y_k^T g_{k+1}}{s_k^T y_k} \qquad (7):$$

$$d_{k+1} = -g_{k+1} + \beta^{v1} s_k \qquad (8)$$

2.4 تطوير خوارزمية V1 – CG

من المعلوم ان طرائق الامثلية العددية هي طرائق تكرارية ولايوجد طريقة معينة ملائمة لجميع انواع المسائل فلكل طريقة محاسنها ومميزاتها الجيدة وكذلك بعض الصفات غير الجيدة وتكون كفوءة لبعض انواع المسائل وغير كفوءة لأنواع اخرى من المسائل.

ان اغلب طرائق الامثلية العددية وبالتحديد طرائق الرتبة الاولى وطرائق الرتبة الثانية يتم اشتقاقها استناداً الدالة تربيعية بطريقة ما وعليه تواجه بعض المشاكل عند تطبيقها على الدوال العامة .

يتركز اهتمامنا في هذه الفقرة على طريقة V1 – CG حيث نحاول تطوير هذه الطريقة وبالتالي زيادة سرعة تقاربها مع الاحتفاظ بخاصية التقارب الشامل لهذه الطريقة ان اشتقاق هذه الطريقة استند الى دالة تربيعية محدبة صارمة بالشكل الاتي: [7]

$$f(x) = \frac{1}{2}\, x^T G x$$

حيث $x \in R^n$ و G مصفوفة هيسي بسعة n × n متناظرة وموجبة التعريف باستخدام متسلسلة تايلر بالنسبة لمشتقة الدالة f نحصل على

$$g_{k+1} = g_k + G\,(x_{k+1} - x_k) \qquad (9)$$

أو بصيغة مختصرة

$$y_k = G s_k \qquad (10)$$

حيث $y_k = \ _{k+1} - g_k$ و $s_k = x_{k+1} - x_k$. نلاحظ ان المعادلة (10) غير صحيحة بالنسبة للدالة العامة (غير التربيعية) لان المشتقات العليا في هذه الحالة لاتساوي صفروقد جرت محاولات عديدة لتصحيح المعادلة (10) بالنسبة للدالة غير التربيعية فعلى سبيل المثال اقترح العالم Li في [8] الصيغة الاتية :

Z = GS (11)

حيث

$$Z = y_k + t\,|g_k|\, s_k \qquad (12)$$

$$t = 1 + Max\left[-\frac{s_k^T y_k}{|s_k|}, 0\right] \qquad (13)$$

وكذلك اقترح العالم Wei في [14] عدة صيغ منها :

1. $Z = y_k + \frac{\theta}{s_k^T u_k}\, u_k \qquad (14)$

حيث u أي متجه في R^n بحيث $\ ^T_k s_k \neq 0$ وان

$$\theta = 2\,[f_k - f_{k+1}] + [g_{k+1} + g_k]^T s_k \qquad (15)$$

نلاحظ ان الصيغة (14) يتضمن قيمة دالة الهدف وكذلك المشتقة عند نقطتين مختلفتين هما x_k و x_{k+1} لذا يمثل افضل تقريب للمشتقة الثالثة

2. $Z = y_k + A\, s_k \qquad (16)$

حيث A أي مصفوفة متناظرة وموجبة التعريف . ولتطوير خوارزمية V1-CG نستخدم المعادلات β^{v1} و (14) أي بتعويض عن كل y بـ Z المعرفة في المعادلة (14) ولنسمي الطريقة الناتجة GV1-CG

$$\beta^{GV1} = \left(1 - \frac{s_k^T z}{z^T z}\right) \frac{z^T g_{k+1}}{s_k^T z} \qquad (17)$$

وباستخدام التعويض عن z نحصل على

$$\beta^{GV1} = \frac{\left(y_k + \frac{\theta_k}{s_k^T u_k} u_k\right)^T g_{k+1}}{s_k^T\left(y_k + \frac{\theta_k}{s_k^T u_k} u_k\right)} - \frac{\left(y_k + \frac{\theta_k}{s_k^T u_k} u_k\right)^T g_{k+1}}{\left(y_k + \frac{\theta_k}{s_k^T u_k} u_k\right)^T \left(y_k + \frac{\theta_k}{s_k^T u_k} u_k\right)}$$

$$= \frac{y_k^T g_{k+1} + \frac{\theta_k}{s_k^T u_k} u_k^T g_{k+1}}{S_k^T y_k + \frac{\theta_k}{s_k^T u_k} s_k^T u_k} - \frac{y_k^T g_{k+1} + \frac{\theta_k}{s_k^T u_k} u_k^T g_{k+1}}{y_k^T y_k + \frac{\theta_k}{s_k^T u_k} y_k^T u_k + \frac{\theta_k}{s_k^T u_k} u_k^T y_k + \left(\frac{\theta_k}{s_k^T u_k}\right)^2 u_k^T u_k}$$

$$= \frac{y_k^T g_{k+1} + \frac{\theta_k}{s_k^T u_k} u_k^T g_{k+1}}{S_k^T y_k + \frac{\theta_k}{s_k^T u_k} s_k^T u_k} - \frac{y_k^T g_{k+1} + \frac{\theta_k}{s_k^T u_k} u_k^T g_{k+1}}{y_k^T y_k + 2\frac{\theta_k}{s_k^T u_k} u_k^T y_k + \left(\frac{\theta_k}{s_k^T u_k}\right)^2 u_k^T u_k} \qquad (18)$$

نلاحظ من المعادلة (18) المتجه u متجه اختياري وهناك عدة طرق لاختيارها منها على سبيل المثال اذا اخذنا $u_k = y_k$ ونعوض عنها في المعادلة (18) نحصل على :

$$\beta^{GV1} = \frac{y_k^T g_{k+1} + \frac{\theta_k}{s_k^T y_k} y_k^T g_{k+1}}{S_k^T y_k + \frac{\theta_k}{s_k^T y_k} s_k^T y_k} - \frac{y_k^T g_{k+1} + \frac{\theta_k}{s_k^T y_k} y_k^T g_{k+1}}{y_k^T y_k + 2\frac{\theta_k}{s_k^T y_k} y_k^T y_k + \left(\frac{\theta_k}{s_k^T y_k}\right)^2 y_k^T y_k} =$$

$$= \frac{\left(1 + \frac{\theta_k}{s_k^T y_k}\right) y_k^T g_{k+1}}{\left(1 + \frac{\theta_k}{s_k^T y_k}\right) s_k^T y_k} - \frac{\left(1 + \frac{\theta_k}{s_k^T y_k}\right) y_k^T g_{k+1}}{\left[\left(1 + \frac{\theta_k}{s_k^T y_k}\right)\left(1 + \frac{\theta_k}{s_k^T y_k}\right)\right] y_k^T y_k}$$

$$\therefore\ \beta^{GV1} = \frac{y_k^T g_{k+1}}{s_k^T y_k} - \frac{y_k^T g_{k+1}}{\left(1 + \frac{\theta_k}{s_k^T y_k}\right) y_k^T y_k} \qquad (19)$$

اذن ممكن تعريف اتجاه البحث للخوارزمية المعدلة الجديدة بالشكل التالي :

$$d_{k+1} = -g_{k+1} + \beta^{GV1} s_k \qquad ...(20)$$

2.4.1 خاصية الانحدار للخوارزمية GV1-CG

مبرهنة (2.1) (برهان خاصية الانحدار لـ GV1-CG)

نفرض α تحقق شروط Wolfe $f(x_k + \alpha d_k) \le f(x_k) +$ $c_1 \alpha g_k^T d_k$ و $g_{k+1}^T d_k \ge c_2 g_k^T d_k$ حيث ان $0 < c_1 < c_2 <$ 1 وان g تحقق شرط ليبشز لكل x وان $0 < L < 1$ عندئذ خوارزمية GV1-CG تحقق خاصية الانحدار .

البرهان :

سوف نبرهن هذه النظرية عن طريق الاستقراء الرياضي عندما K = 1 فان

$$d_1 = -g_1$$

$$d_k^T g_k = -g_1^T g_1 = -\|g_1\|_2^2 < 0$$

نفرض $d_k^T g_k < 0$ لكل K

نبرهن $d_{k+1}^T g_{k+1} < 0$

وعليه

$$d_{k+1}^T g_{k+1} = -g_{k+1}^T g_{k+1} + \beta^{GV1} s_k^T g_{k+1}$$

$$= -g_{k+1}^T g_{k+1} + \left(\frac{y_k^T g_{k+1}}{s_k^T y_k} - \frac{y_k^T g_{k+1}}{\left(1+\frac{\theta_k}{s_k^T y_k}\right) y_k^T y_k}\right) s_k^T g_{k+1}$$

$$=- g_{k+1}^T g_{k+1} + \frac{y_k^T g_{k+1}}{s_k^T y_k} s_k^T g_{k+1} - \frac{y_k^T g_{k+1}}{\left(1+\frac{\theta_k}{s_k^T y_k}\right) y_k^T y_k} s_k^T g_{k+1}$$

$$=- g_{k+1}^T g_{k+1} + (y_k^T g_{k+1})(s_k^T g_{k+1})\left(\frac{1}{s_k^T y_k} - \frac{1}{\left(1+\frac{\theta_k}{s_k^T y_k}\right) y_k^T y_k}\right)$$

من خلال شرط ليبشز

$$y_k^T g_{k+1} \le L s_k^T g_{k+1}$$

$$d_{k+1}^T g_{k+1} \le -g_{k+1}^T g_{k+1} + L (s_k^T g_{k+1})^2 \left(\frac{1}{s_k^T y_k} - \frac{1}{\left(1+\frac{\theta_k}{s_k^T y_k}\right) y_k^T y_k}\right)$$

$$=- g_{k+1}^T g_{k+1} + L (s_k^T g_{k+1})^2 \left(\frac{\left(1+\frac{\theta_k}{s_k^T y_k}\right) y_k^T y_k - s_k^T y_k}{s_k^T y_k \left(1+\frac{\theta_k}{s_k^T y_k}\right) y_k^T y_k}\right)$$

=

نلاحظ ان $s_k^T y_k > 0$ من خلال شرط Wolfe ، $y_k^T y_k >$ 0 و $y_k^T y_k \le L s_k^T y_k$ وبتعويضها في المعادلة اعلاه نحصل على :

$$d_{k+1}^T g_{k+1} \le - g_{k+1}^T g_{k+1} + L (s_k^T g_{k+1})^2 \left(\frac{\left(1+\frac{\theta_k}{s_k^T y_k}\right)^{L-1}}{\left(1+\frac{\theta_k}{s_k^T y_k}\right) y_k^T y_k}\right) < 0 \quad \forall L \in (0,1)$$

حيث ان $\left(1+\frac{\theta_k}{s_k^T y_k}\right)$ كمية موجبة , وهذا يؤدي الى ان $d_{k+1}^T g_{k+1} < 0$ لذلك فان اتجاه البحث المولدة من خوارزمية GV1-CG هي متجه منحدر لكل

2.4.2.K التقارب الشامل للخوارزمية GV1-CG

مبرهنة (2.2) (برهان التقارب للخوارزمية GV1-CG)

ليكن α_k تحقق شروط Wolfe وان الدالة f مقيدة من الاسفل وان d_k هي اتجاه منحدر لكل k أي ان $g_k^T d_k < 0$ وان g_k تحقق شرط ليبشز في فترة مفتوحة N تحتوي L بحيث $L \equiv \{x : f(x) \le f(x_0)\}$ عندها الخوارزمية اما تتوقف عند النقطة $\|g_k\|_2 = 0$ او $\lim_{k\to\infty} inf \ \|g_k\|_2 = 0$.

البرهان : سنبرهن عن طريق التناقض , أي بمعنى نفرض المبرهنة ليست صحيحة فان $\|g_k\| \neq 0$ وعليه يوجد ثابت موجب وليكن $\lambda > 0$ بحيث

$$\|g_k\|_2 \ge \lambda \qquad (21)$$

فان

$$d_{k+1}^T g_{k+1} = -g_{k+1}^T g_{k+1} + \frac{y_k^T g_{k+1}}{s_k^T y_k} s_k^T g_{k+1} - \frac{y_k^T g_{k+1}}{\left(1+\frac{\theta_k}{s_k^T y_k}\right) y_k^T y_k} s_k^T g_{k+1}$$

باستخدام شرط Wolfe الثاني (الاساسي) $g_{k+1}^T d_k \ge c_2 g_k^T d_k$ لـ $s_k^T g_{k+1}$ وشرط ليبشز ل

$$y_k^T y_k \le L s_k^T y_k \Rightarrow \frac{L}{y_k^T y_k} \ge \frac{1}{s_k^T y_k}.$$

لذلك

$$d_{k+1}^T g_{k+1} \le -g_{k+1}^T g_{k+1} + \sigma L \frac{y_k^T g_{k+1}}{s_k^T y_k} s_k^T g_k - \sigma \frac{y_k^T g_{k+1}}{\left(1+\frac{\theta_k}{s_k^T y_k}\right) y_k^T y_k} s_k^T g_k$$

$$= -g_{k+1}^T g_{k+1} + \sigma \left(L - \frac{1}{\left(1+\frac{\theta_k}{s_k^T y_k}\right)}\right) \frac{y_k^T g_{k+1}}{y_k^T y_k} s_k^T g_k$$

اذا كانت

$$y_k^T y_k = g_{k+1}^T g_{k+1} - 2 g_{k+1}^T g_k + g_k^T g_k \ge g_{k+1}^T g_{k+1} - g_{k+1}^T g_k = y_k^T g_{k+1}$$

$$1 \ge \frac{y_k^T g_{k+1}}{y_k^T y_k}$$

$$\therefore \ d_{k+1}^T g_{k+1} \ge -\|g_{k+1}\|^2 + \omega s_k^T g_k \quad (22)$$

عندما $\omega = \sigma\left(L - \frac{1}{\left(1+\frac{\theta_k}{s_k^T y_k}\right)}\right)$, وبقسمة طرفي المتباينة (22) على $\omega \|g_{k+1}\|_2^2$

وبتربيع الطرفين وبعد تبسيط المعادلة نحصل على :

بـ $\frac{1}{\omega^2}\left[\frac{d_{k+1}^T g_{k+1}}{\|g_{k+1}\|_2^2} + 1\right]^2 \ge \frac{(g_k^T s_k)^2}{\|g_{k+1}\|_2^4}$

بضرب الطرفين في $\|g_{k+1}\|_2^4$ واستخدام الحقيقة

$$(g_k^T s_k)^2 = \|s_k\|_2^2 \|g_k\|_2^2 \cos^2\theta_k$$

فان

$$\frac{1}{\omega^2} \frac{\|g_{k+1}\|_2^4}{\|s_k\|_2^2}\left[\frac{d_{k+1}^T g_{k+1}}{\|g_{k+1}\|_2^2} + 1\right]^2 \ge \frac{(g_k^T s_k)^2}{\|s_k\|_2^2} = \|g_k\|_2^2 \cos^2\theta_k \ge \lambda^2 \cos^2\theta_k$$

باخذ المجموع عندما $k \ge 1$ ينتج :

$$\frac{1}{\omega^2} \sum_{k=1}^{\infty} \frac{\|g_{k+1}\|_2^2}{\|s_k\|_2^2}\left[\frac{d_{k+1}^T g_{k+1}}{\|g_{k+1}\|_2^2} + 1\right]^2 \ge \sum_{k=1}^{\infty} \frac{(g_k^T s_k)^2}{\|s_k\|_2^2} =$$

$$\sum_{k=1}^{\infty} \|g_k\|_2^2 \cos^2\theta_k \ge \sum_{k=1}^{\infty} \lambda^2 \cos^2\theta_k = \infty$$

وهذا يناقض شرط Zoutendijk [15] لذلك اما $\|g_k\|_2 = 0$

أو $\lim_{k\to\infty} inf \ \|g_k\|_2 = 0$

3. التجارب العددية (Numerical Experiments)

فـي هـذه الفقـرة سـنقدم أداء لبرنـامج الفـورتران (COMPAQ VICUAL FORTRAN)لخوارزمية التدرج المترافق GV1-CG في مجموعة من الدوال القياسية في الامثلية غير المقيدة مأخوذة من [2] . علمـاً ان البرنـامج اساسـاً للعـالم Andrei وقمنا بتطـويره وقمنا باختيار 22 دالة من الدوال القياسية ذات النطاق الواسع(انظر الملحق) ولكل دالة اعتبرنا عدد المتغيرات للتجارب العددية عندما $n = 100$ و $n = 1000$.

قمنا بمقارنـة أداء هـذه الخوارزميـة مـع β^{GV1} المـذكورة فـي المعادلـة (19) (التـي هـي افضـل مـن β^{V1} المـذكورة فـي المعادلـة (7)) .[6] وهـذه الخوارزميـة تحقـق شـرطي Wolfe القياسـية بحيـث ان $c_1 =$ 0.001 و $c_2 = 0.09$ وطـول الخطـوة الابتدائيـة $\alpha_1 = \frac{1}{\|g_1\|}$ والتخمـين للتكـرارات الاخـرى بمعنـى $k > 1$ هـو $\alpha_k = \alpha_{k-1} * \frac{\|d_{k-1}\|}{\|d_k\|}$. وفي جميع الحالات فان معيار التوقف هو $\|g_{k+1}\| < 10^{-6} * max[1, |f_{k+1}|]$ وان اكبر عدد ممكن من التكـرارات هو 2000 . والمقارنات تتضمن ماياتي :

1- NOI : عدد التكرارات

2- Line : عـدد مـرات اسـتدعاء البرنـامج الفرعـي لحسـاب حجـم الخطوة α_k .

3- Time : الوقت الذي يستغرقه تنفيذ الدالة .

ان الجـدول رقـم (1) يوضـح تفاصـيل النتـائج عنـدما $n = 100$ والجـدول رقـم (2) يوضـح تفاصـيل النتـائج عنـدما $n = 1000$. والجدول رقم (3) يوضح اداء الخوارزمية بالنسب المئوية اذ ان طريقة V1-CG عدّت 100% ومن الجدول نرى ان الخوارزمية حسنت من طريقـة V1-CG تقريبـاً (16% - 40%) مـن ناحيـة التكـرارات NOI.

جدول رقم (1) تفاصيل نتائج NOI و Lins و time عندما N = 100

	V1-CG	GV1-CG
	NOI / Lins / Time	NOI / Lins / Time
1	34 / 18 / 0.02	31 / 0 / 0.02
2	84 / 25 / 0.12	93 / 0 / 0.11
3	89 / 36 / 0.34	64 / 0 / 0.23
4	25 / 11 / 5.48	28 / 0 / 1
5	21 / 6 / 0.05	23 / 0 / 0.06
6	14 / 8 / 0.03	7 / 0 / 0.03
7	34 / 11 / 0.03	58 / 0 / 0.03
8	1299 / 470 / 4	263 / 0 / 1
9	65 / 30 / 0.03	53 / 1 / 0.02
10	40 / 17 / 0.03	31 / 0 / 0.02
11	150 / 82 / 5.33	190 / 2 / 1
12	338 / 101 / 2	321 / 0 / 0.61
13	9 / 5 / 0.02	7 / 0 / 0.01
14	10 / 10 / 0.03	11 / 0 / 0.01
15	74 / 19 / 3.99	75 / 0 / 4.49
16	166 / 59 / 1.93	86 / 0 / 0.18
17	26 / 9 / 0.45	23 / 0 / 0.28
18	34 / 18 / 0.01	30 / 1 / 0.02
19	203 / 143 / 6	168 / 0 / 6
20	24 / 19 / 0.02	26 / 0 / 0.01
21	55 / 55 / 0.05	23 / 0 / 0.04
22	213 / 104 / 0.08	219 / 0 / 0.25
Total	3007 / 1246 / 30.04	1830 / 4 / 15.42

جدول رقم (2) تفاصيل نتائج NOI و Lins و time عندما N = 1000

	V1-CG	GV1-CG
	NOI / Lins / Time	NOI / Lins / Time
1	34 / 18 / 0.02	34 / 0 / 0.02
2	346 / 99 / 3	300 / 0 / 2
3	384 / 185 / 8	255 / 0 / 4
4	166 / 136 / 59	181 / 0 / 66
5	23 / 3 / 0.05	27 / 0 / 1
6	11 / 6 / 0.03	9 / 0 / 1
7	59 / 19 / 0.03	52 / 0 / 1
8	1097 / 413 / 45	470 / 0 / 19
9	63 / 29 / 0.03	53 / 0 / 0.02
10	37 / 19 / 1	29 / 0 / 0.02
11	1138 / 1045 / 122	574 / 1 / 31
12	929 / 269 / 8	1535 / 0 / 13
13	5 / 4 / 0.02	7 / 0 / 0.01
14	10 / 10 / 1	12 / 0 / 0.01
15	239 / 55 / 109	305 / 0 / 139
16	1014 / 408 / 36	243 / 0 / 4
17	140 / 125 / 9	84 / 0 / 5
18	35 / 21 / 0.01	36 / 1 / 0.02
19	2001 / 1846 / 1197	2001 / 0 / 1197
20	26 / 10 / 0.02	20 / 0 / 1
21	52 / 51 / 0.05	21 / 1 / 1
22	199 / 37 / 2	519 / 0 / 5
Total	8008 / 4808 / 1600.26	6767 / 3 / 1490.1

جدول رقم(3) مقارنة الخوارزمية بالنسبة المئوية بالنسبال الى NOI

N	Measure	V1-CG	GV1-CG
100	NOI	100%	60.85%
1000	NOI	100%	84.50%

الملحق (Appendix)

1- Extended Rosenbrok Function

$$f(x) = \sum_{i=1}^{n/2} c(x_{2i} - x_{2i-1}^3)^2 + (1 - x_{2i-1})^2$$

$c = 100\,,\quad x_1 = [-1.2,1,\dots,-1.2,1]$

2- Perturbed Quadratic Function

$$f(x) = \sum_{i=1}^{n} i x_i^2 + \frac{1}{10}\left(\sum_{i=1}^{n} x_i\right)^2$$

$x_1 = [0.5,\dots,0.5]$

3- Raydan (1) Function

$$f(x) = \sum_{i=1}^{n} \frac{i}{10}(exp(x_i) - x_i)$$

$x_1 = [1,1,\dots,1]$

4- Hager Function

$$f(x) = \sum_{i=1}^{n} \left(exp(x_i) - \sqrt{i x_i}\right)$$

$x_1 = [1,1,\dots,1]$ 5-

Generalized Tridiagonal-1 Function

$$f(x) = \sum_{i=1}^{n-1} (x_i + x_{i+1} - 3\,)^2 + (x_i - x_{i+1} + 1)^4$$

$x_1 = [1,1,\dots,1]$

6- Extended Three Exponential Terms Function

$$f(x) = \sum_{i=1}^{n} \left(exp(x_{2i-1} + 3x_{2i} - 0.1) + exp(x_{2i-1} - 3x_{2i} - 0.1) + exp(-x_{2i-1} - 0.1)\right)$$

$x_1 = [0.1,\dots,0.1]$

7- Generalized Tridiagonal 2 Function

$$f(x) = \left((5 - 3x_1 - x_1^2)x_1 - 3x_2 + 1\right)^2 + \sum_{i=1}^{n-1} \left((5 - 3x_i - x_i^2)x_i - 3x_{i+1} + 1\right)^2 + \left((5 - 3x_n - x_n^2)x_n - 3x_{n-1} + 1\right)^2 ,$$

$x_1 = [-1,\dots,-1]$

8- Extended Powell Function

$$f(x) = \sum_{i=1}^{n-1} (x_{4i-3} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 + (x_{4i-2} - 2x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4$$

$x_1 = [3,-1,0,1,\dots,3,-1,0,1]$

9- Extended Maratos Function

$$f(x) = \sum_{i=1}^{n-1} x_{2i-1} + c(x_{2i-2}^2 + x_{2i}^2 - 1)^2$$

$x_1 = [1.1,0.1,\dots,1,0.1]$

10- ARWHEAD Function

$$f(x) = \sum_{i=1}^{n-1} (-4x_i + 3) + \left(\sum_{i=1}^{n-1} x_i^2 + x_n^2\right)^2$$

$$x_1 = [1, \ldots, 1]$$

11- NONDIA Function

$$f(x) = (x_1 - 1)^2 + \sum_{i=2}^{n} 100(x_1 - x_{i-1}^2)^2$$

$$x_1 = [-1, \ldots, -1]$$

12- Extended Wood Function

$$f(x) = \sum_{i=1}^{n/4} 100(\,x_{4i-3}^2 - x_{4i-2})^2 + (x_{4i-3} - 1)^2 + 90(\,x_{4i-1}^2 - x_{4i})^2 + (1 - x_{4i-1})^2$$
$$+ 10.1\{(x_{4i-2} - 1)^2 + (x_{4i} - 1)^2\} + 19.8(x_{4i-2} - 1)(x_{4i} - 1)$$

$$x_0 = [-3, -1, -3, -1, \ldots, -3, -1, -3, -1]$$

13- Extended Hiebert Function

$$f(x) = \sum_{i=1}^{n/2} (x_{2i-1} - 10)^2 + (x_{2i-1}x_{2i} - 50000)^2$$

$$x_1 = [0, \ldots, 0]$$

14- Almost Perturbed Quadratic

$$f(x) = \sum_{i=1}^{n} i\, x_i^2 + \frac{1}{100}(x_i + x_n)^2$$

$$x_1 = [0.5, \ldots, 0.5]$$

15- ENGVALI Function (CUTE)

$$f(x) = \sum_{i=1}^{n-1} (\,x_i^2 + \,x_{i+1}^2)^2 \;+\; \sum_{i=1}^{n-1} (-4x_i + 3)$$

$$x_1 = [2, 2, \ldots, 2]$$

16- Extended Quadratic Penalty QP2 Function

$$f(x) = \sum_{i=1}^{n-1} (\,x_i^2 + sinx_i)^2 + \left(\sum_{i=1}^{n} x_i^2 - 100\right)^2$$

$$x_1 = [1, \ldots, 1]$$

17- Extended Tridiagonal - 2 Function

$$f(x) = \sum_{i=1}^{n} (x_i x_{i+1} - 1)^2 + c(x_i + 1)(x_{i+1} + 1)$$

$$x_1 = [1, \ldots, 1]\;, \qquad c = 0.1$$

18- Extended White & Holst Function

$$f(x) = \sum_{i=1}^{n/2} c(x_{2i} - x_{2i-1}^3)^2 + \;(1 - x_{2i-1})^2$$

$$x_0 = [-1.2, 1, \ldots, -1.2, 1]\;, \qquad c = 100$$

19- Diagonal 2 Function

$$f(x) = \sum_{i=1}^{n/2} \left(exp(x_i) - \frac{x_i}{i}\right)$$

$$x_0 = [1/1, 1/2, \ldots, 1/n]$$

20- Extended Tridiagonal - 1 Function

$$f(x) = \sum_{i=1}^{n/2} (x_{2i-1} + x_{2i} - 3\,)^2 + \;(x_{2i-1} - \,x_{2i} + 1)^4$$

$$x_0 = [2, 2, \ldots, 2]$$

21- Extended Block- Diagonal BD1 Function

$$f(x) = \sum_{i=1}^{n/2} (x_{2i-1}^2 + x_{2i}^2 - 2\,)^2 + \;(exp(x_{2i-1} - 1) - \,x_{2i})^2$$

$$x_0 = [0.1, 0.1, \ldots, 0.1]$$

22- Extended Cliff Function

$$f(x) = \sum_{i=1}^{n/2} \left(\frac{x_{2i-1} - 3}{100}\right)^2 - (x_{2i-1} - x_{2i}) + exp(20(x_{2i-1} - x_{2i}))$$
$$x_0 = [0, -1, \dots, 0, -1]$$

المصادر

[1] Andrei, N. (2007c), Conjugate Gradient Algorithms for Molecular Formation under Pair Wise Potential Minimization, Center for Advanced Modeling and Optimization.

[2] Andrei, N. (2008), An Unconstrained Optimization Test Function collection, Advance Model. Optimization, Vol.(10),PP.147-161.

[3] Chong, E. K. P. and Zak, S. H., (2001), An Introduction to Optimization, Jone Wiley & Sons, Inc., Canada.

[4] Fletcher, R., (1987), Practical Methods of Opimization, A Wiley Inter Science Publication, John-Wiley & Sons, Inc., New York.

[5] Hestenes, M. R. and Stiefel, E., (1952), Methods of Conjugate Gradients for Solving Linear Systems, Journal of Research of the National Bureau of Standards, Vol. (5), No.(49), pp. 409-436.

[6] Hiroshi Y. and Naoki S. (2005), Anew nonlinear conjugate gradient method for unconstrained optimization, Journal of the Operations Research, Vol.(48), No.(4),PP. 284-296.

[7] Khalil, A. K., (2010), Developing of Gradient Algorithms for Solving Unconstrained Non-Linear Problem with Artificial Neural Networks, Doctoral thesis Faculty of Computing and Mathematics University of Mosal Sciences.

[8] Li, D. and M. Fukushima, (2001), Amodified BFGC Method and its global Convergence in nonconvex minimization J. comput. Appl. Math,129.

[9] Nocedal, J. and Wright, S. J., (2006), Numerical Optimization Springer Series in Operation Research, 2^{nd} edition, Springer-Verlag, New York.

[10] Pedregal, P., (2004), Introduction to Optimization, Springer-Verlag, Inc., New York, USA.

[11] Powell, M. J. D., (1977), Restart Procedure for Conjugate Gradient Method, Mathematical Programming, 12, 241-254.

[12] Shanno, D. (1978), On the Convergence of a Conjugate Gradient Algorithm, SIAM.J. Number. Anal, 15.

[13] Sun, W. and Yuan, Y., (2006), Optimization Theory and Methods, Nonlinear Programming, Springer Science, Business Media, LIC., New York.

[14] Wei Z., Li. G. and Qi L. (2006), New Quasi Newton Methods for Unconstrained Opt. Problem. $\bar{O}$ Applied Math. Comput 175.

[15] Zoutendijk, G.(1970), Nonliear programming, computationl methods, In Integer and Nonlinear Programming, J. Abadie (ED), North – Holand.

مصادر الانترنت

INT[1]: Kinsella. J. (2009) Course Notes for MS4327 Optimization http:// jkcray. Maths.ul.ie/ms4327/slides.pdf2011.

Development Modified Conjugate Gradiente Algorithm

Abbas H. Taqi [1] , Amal N. Shaker [2]

[1] *College of Science , Kirkuk University , Kirkuk , Iraq*

[2] *Department of Mathematic , College of Education for Pure Sciences , Tikrit University , Tikrit , Iraq*

Abstract

In this paper was to develop method V1-CG methods associated gradient amended to increase the speed of the convergence while retaining the charactenstic mass convergence as the derivation of this method was based on strict convex quadratic function has been develop this way to public function in whichthe supreme derivatives not equal zero.

The Desent property and global convergence for the proposed algorithm are established, oure numerical experiment on some test function and it showed us a clear improvement on the way V1-CG modified.