Enriched Algorithms for Large Scale Unconstrained Optimization

Abbas. Y. Al-Bayati **Omar B. Mohammad**
profabbasalbayati@yahoo.com
College of Computer Sciences and Mathematics
University of Mosul, Iraq

Received on: 02/02/2005 **Accepted on: 09/05/2005**

ABSTRACT

A new method for solving Large-Scale problems in the unconstrained optimization has been proposed in this research depending on the BFGS method.

The limited memory is used in the BFGS method by multiplying the BFGS matrix by a vector to obtain vectors instead of matrices and only two vectors can be stored, by modifying the algorithm given by Nocedal J (1999).

The purpose of this algorithm is to enable us to solving the Large-Scale Problems, as it is obvious to everyone that the computer can store millions of vectors, whereas its ability in storing matrices is limited.

The present method in this research is applied on seven nonlinear functions in order to evaluate the method efficiency in the numbers of iterations (NOI), number of functions (NOF) and function value and comparing it with the standard BFGS method after updating.

This method has been applied on functions with variables till 1000000 and more than that.

From comparing the results, we fined that this algorithm was the best.

Keywords: Unconstrained Optimization, Large Scale, Enriched Algorithms.

الخوارزميات القابلة للتوسع في الأمثلية غير المقيدة ذات القياس العالي

عباس يونس البياتي **عمر بهاء الدين محمد**
كلية علوم الحاسبات والرياضيات، جامعة الموصل

تاريخ استلام البحث: 2005/02/02 **تاريخ قبول البحث: 2005/05/09**

الملخص

يتناول هذا البحث طريقة جديدة في حل المسائل ذات القياس العالي في الأمثلية غير المقيدة بالاعتماد على طريقة BFGS .

في طريقة BFGS تم استخدام الذاكرة المحدودة، حيث قمنا بضرب مصفوفة BFGS بمتجه ليكون حاصل الضرب بشكل متجهات بدلا من أن يكون بشكل مصفوفات، ويتم خزن متجهين فقط، وذلك من خلال تطوير الخوارزمية المعطاة من قبل Nocedal J. (1999) .

إن الغاية من هذا هو أن يصبح بإمكاننا حل المسائل ذات الأبعاد الكبيرة، إذ أن من الواضح للجميع أن الحاسبة بإمكانها خزن الملايين من المتجهات في حين أن إمكانيتها على خزن المصفوفات تكون محدودة.

تم تطبيق الطريقة المقدمة في هذا البحث على سبع دوال لا خطية معروفة في هذا المجال لغرض تقييم كفاءة الطريقة من ناحية عدد التكرارات (NOI) ، عدد مرات حسابات الدالة (NOF) وقيمة الدالة Function Value ومقارنتها بالطريقة التقليدية إلى BFGS بعد التحديث. تم تطبيق الطريقة على دوال ذات متغيرات لغاية 1000000 متغير وأكثر. من مقارنة النتائج تبين أن الخوارزمية الجديدة على العموم كانت الأفضل.

الكلمات المفتاحية: الأمثلية غير المقيدة، القياس العالي، الخوارزميات القابلة للتوسع.

1. طرائق التدرج (Gradient Methods):

طرائق التدرج للامثلية مبنية على توسيع Taylor. يعاد التوسيع هنا بإهمال الحدود التي تتضمن المشتقات من الرتبة الثالثة وما فوق ، هذا يعني

$$f(x_k+\Delta x)\cong f(x)+g^T\Delta x+1/2\Delta x^T G\Delta x \qquad (1)$$

الحدان الأخيران على الجهة اليمنى هما تصحيح قياسي لقيمة الدالة عند النقطة x حيث تنتج تقريبات لقيم الدالة عند النقطة $x+\Delta x$.

(i) طرائق التدرج المترافق (Conjugate Gradient (CG) Methods):

الدالة التربيعية في n من الأبعاد يمكن كتابتها على النحو الآتي:-

$$f(x)=1/2x^T Gx+b^T x+c \qquad (2)$$

مجموعة من n من المتجهات العمودية $d_1,d_2,...,d_n$ يمكن تعريفها على أنها مترافقة تبادليا بدلالة G ، وهذا يعني $d_i^T Gd_j=0$ لكل $i\neq j$

الخوارزمية الأولى المعروفة من هذا النوع تعود إلى Fletcher-Reeves (1964) .

اتجاه البحث الأول لطريقة FR التي تبدأ من نقطة البداية x_1 هي $d_1=-g_1$ (تدعى اتجاه الانحدار السريع).

في مرحلة ألـ $(k+1)th$ نختار d_{k+1} ليكون تركيبا خطيا لـ $-g_{k+1}$ و $d_1,d_2,...,d_k$ مرافقات.

وهكذا فان:

$$d_{k+1} = -g_{k+1} + \sum_{r=1}^{k} \beta_r d_k \qquad k = 1,2,3,... \qquad (3)$$

تبين هذه المعادلة أن كل قيم β_r أصفار ماعدا β_k حيث أن

$$d_{k+1} = -g_{k+1} + \beta_k d_k \; ; \qquad \beta_k = \frac{g_{k+1}^2}{g_k^2}$$

وتسمى بطريقة (FR-CG).

(ii) <u>طرائق المتري المتغير (Variable Metric (VM) Methods):</u>

النموذج التربيعي يمكن الحصول عليه من متسلسلة قطع Taylor الموسعة لـ $f(x)$ حول x_k ، التي يمكن كتابتها:

$$f(x_k + v) = f_k + g_k^T v + \frac{1}{2} v^T G v \qquad (4)$$

حيث $v_k = x_{k+1} - x_k$ وان $x_{k+1} = x_k + v_k$.

واضح أن $(v_k = \lambda_k d_k)$ مصممة لهذا الغرض.

التصحيح v_k يقلل الطرف الأيمن من المعادلة (4) إلى الحد الأدنى والطريقة في الواقع تحتاج إلى مشتقات الدرجة الأولى والثانية من f لتكون متوافرة عند أية نقطة.

نشتق (4) بدلالة v لتعطي

$$\frac{\nabla f}{\nabla v} = G_k v + g_k = 0 \quad \Rightarrow v_k = -G_k^{-1} g_k \qquad (5)$$

حيث G_k هي مصفوفة Hessian .

1-2 <u>الخوارزميات الشبيهة بخوارزمية نيوتن (Quasi-Newton Algorithms):</u>

هذا النوع من الخوارزميات يشبه خوارزمية نيوتن مع خط البحث، ماعدا مصفوفة Hessian G_k التي تقرب بوساطة المصفوفة المتناظرة الموجبة التعريف H_k ، والتي تصحح من تكرار إلى تكرار.

وهكذا للتكرار $k-th$ التركيب الأساسي الآتي:

(a). نضع $d_k = -H_k g_k$.

(b). نجعل خط بحث على طول d_k للحصول على $x_{k+1} = x_k + \lambda_k d_k$

(c). تحديث H_k بوساطة المصفوفة المصححة للحصول على

$$H_{k+1} = H_k + \theta_k$$

حيث θ_k هي المصفوفة المصححة الموجبة التعريف وأنها تحقق شرط QN، وهو:

$$H_{k+1} y_k = \sigma_k v_k \qquad (6)$$

حيث σ_k قياس.

التقريب الأكثر عموما هو الذي نشر بوساطة (Huang, 1970)، لقد درس الخوارزمية مع خط البحث المضبوط والتي تستخدم صيغة التحديث العامة من الشكل.

$$H_{k+1}^{Huang} = H_k + w_k v_k v_k^T + \pi_k H_k y_k y_k^T H_k + \rho_k (v_k y_k^T H_k + H_k y_k^T v_k) \qquad (7)$$

حيث π_k, w_k و ρ_k هي قياسات واختيارات مختلفة للمعالم π_k, w_k و ρ_k ستنتج أنواعا مختلفة من الخوارزميات الشبيهة بنيوتن QN-algorithms .

اثبت Huang أن الحصول على اتجاهات بحث مترافق في خطوة التقارب n، للحالة التربيعية، $H_{k+1} y_k$ يجب أن تكون مضروبة بـ v_k. باستعمال هذا الشرط في الصيغة (7) تتكون الصيغة المحدثة الآتية:-

$$H_{k+1}^{Huang} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \frac{\theta_k v_k v_k^T}{v_k^T y_k} + \phi_k w_k w_k^T \qquad (8a)$$

حيث

$$w_k = (y_k^T H_k y_k)^{1/2} \left(\frac{v_k}{v_k^T y_k} - \frac{H_k y_k}{y_k^T H_k y_k} \right) \qquad (8b)$$

θ_k, ϕ_k هما قياسات.

إذا كانت θ_k عددا ثابتا، فان الحالة التربيعية مع مصفوفة Hessian G، تمتلك:

$$H_n = \theta G^{-1}$$

إن صيغة التحديث الأكثر اهتماما التي اكتشفها بصورة مستقلة Broyden, Fletcher, Goldfarb and Shanno في عام (1970) هي صيغة BFGS والتي تم الحصول عليها من (8) بوضع $\theta_k = 1$ و $\phi_k = 1$، نحصل على

$$H_{k+1}^{BFGS} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \frac{v_k v_k^T}{v_k^T y_k} + w_k w_k^T \qquad (9)$$

حيث w_k هو متجه معرف في (8 b).

أو يمكن كتابة (9) كالآتي:

$$H_{k+1}^{BFGS} = H_k - \frac{v_k y_k^T H_k + H_k y_k v_k^T}{v_k^T y_k} + \frac{v_k v_k^T}{v_k^T y_k}\left(1 + \frac{y_k^T H_k y_k}{v_k^T y_k}\right) \qquad (10)$$

وقد اظهر Dixon في العام (1972) أن بوساطة خط البحث المضبوط والدالة التربيعية المعطاة ، كل أعضاء عائلة Huang ينتج نفس عدد التكرارات، وإنجازها متغير عند القيام بخط البحث غير المضبوط.

اكتشف (1973) Biggs انـه يمكن الحصـول علـى تقييم أكثر دقـة مـن هذا الانحنـاء باستخدام أربع قطع مستقلة من المعلومات التي تكون متوافرة على طول v_k ؛ المسماة بقيم الدالـة والمشتقات الاتجاهيـة عند نقطتين متعاقبتين. النموذج التكعيبـي الآتـي من f علـى طول v_k تم تركيبه بوساطة Biggs وهو:

$$\rho_k = \frac{v_k^T y_k}{4v_k^T g_{k+1} + 2v_k^T g_k - 6(f_{k+1} - f_k)} \qquad (11a)$$

ولتحديث الصيغة H_k الشكل الآتي:

$$H_{k+1}^{Biggs} = H_k - \frac{H_k y_k y_k^T H}{y_k^T H_k y_k} + \rho_k \frac{v_k v_k^T}{v_k^T y_k} + w_k w_k^T \qquad (11b)$$

حيث

$$w_k = (y_k^T H_k y_k)^{1/2} \left\{ \frac{v_k}{v_k^T y_k} - \frac{H_k y_k}{y_k^T H_k y_k} \right\}$$

في منتصف السبعينات، أحدث استخدام القياس الذاتي لطرائق VM عندما تم إدخاله، تقدما ملحوظا في الفعالية على طرائق VM السابقة. على وجه الخصوص، اقترح Oren (1974) تحديث VM أكثر أهمية للقياس الذاتي بحيث تم تحديث المصفوفة H_k بـ:

$$H_{k+1}^{Oren} = \left\{ H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + w_k w_k^T \right\} \xi_k + \frac{v_k v_k^T}{v_k^T y_k} \qquad (12a)$$

حيث

$$\xi_k = \frac{v_k^T y_k}{y_k^T H_k y_k} \qquad (12b)$$

وتحديث آخر لـ VM من القياس الذاتي تم إنجازه لاحقا من قبل Al-Bayati (1991) ولأجله كانت لصيغة التحديث H_k الشكل الآتي:

$$H_{k+1}^{Al-Bayati} = H_k + \left\{ I - \frac{v_k y_k^T}{v_k^T y_k} \right\} H_k \left\{ I - \frac{y_k v_k^T}{v_k^T y_k} \right\} + \sigma_k \left\{ \frac{v_k v_k^T}{v_k^T y_k} \right\} + w_k w_k^T \qquad (13a)$$

حيث

$$\sigma_k = \frac{y_k^T H_k y_k}{v_k^T y_k} \qquad (13\ b)$$

<u>2. الطرائق الشبيهة بنيوتن للمسائل الصغيرة وذات الحجم الأوسط:</u>
<u>(Quasi-Newton Methods for Small and Medium-Sized Problems)</u>:

في هذا المبحث سنناقش طرائق QN للمسائل الصغيرة وذات الحجم الأوسط كالتي وضحها Nocedal (1999) وسنأخذ بعين الاعتبار توسيعها إلى القياسات الكبيرة.

<u>(2-1) طريقة BFGS:</u>

خوارزمية QN الأكثر شيوعا هي طريقة BFGS ، سميت باسم مكتشفيها Broyden, Fletcher, Goldfarb, and Shanno. في هذه الفقرة، سنشتق هذه الخوارزمية ثم نصف خصائصها النظرية والإنجازات العملية.

نبدأ الاشتقاق بتشكيل النموذج التربيعي الآتي من دالة الهدف عند التكرار المستمر x_k :

$$m_k(d) = f_k + g_k^T d + 1/2 d^T B_k d \qquad (14)$$

هنا B_k مصفوفة متناظرة وموجبة التعريف $n \times n$ والتي سوف تعدل أو تحدث عند كل تكرار. لاحظ أن القيمة والمشتقة من هذا النموذج عند $d = 0$ تتماثل مع f_k و g_k، بالتتابع. تقليل d_k إلى الحد الأدنى من نموذج التحدب التربيعي، يمكن كتابته بشكل صريح كالأتي:

$$d_k = -B_k^{-1} g_k \qquad (15)$$

ويستعمل كاتجاه بحث، والتكرار الجديد هو

$$x_{k+1} = x_k + \lambda_k d_k \ , \qquad (16)$$

بدلا من حساب B_k من جديد في كل تكرار، اقترح Davidon تحديثها في طريقة ابسط من اجل حساب الانحناء المقاس في معظم الخطوة السابقة. لنفرض أننا أوجدنا تكرارا جديدا x_{k+1} ونتمنى إنشاء نموذج تربيعي جديد، من الشكل:

$$m_{k+1}(d) = f_{k+1} + g_{k+1}^T d + 1/2 d^T B_{k+1} d$$

المتطلبات التي يجب أن نفرضها على B_{k+1}، مبنية على المعرفة التي اكتسبناها من الخطوة السابقة وهي:

$$\nabla m_{k+1}(-\lambda_k d_k) = g_{k+1} - \lambda_k B_{k+1} d_k = g_k$$

بإعادة الترتيب نحصل على،

$$B_{k+1} \lambda_k d_k = g_{k+1} - g_k \qquad (17)$$

لتبسيط الرموز يمكن الاستفادة من المتجهات الآتية:

$$v_k = x_{k+1} - x_k \ , \qquad y_k = g_{k+1} - g_k \qquad (18)$$

لذلك (17) تصبح

$$B_{k+1} v_k = y_k \qquad (19)$$

وهنا نشير إلى هذه الصيغة بمعادلة القاطع.

عند إعطاء البديل v_k وتغيير التدرجات y_k، تتطلب معادلة القاطع المصفوفة B_{k+1} المتناظرة والموجبة التعريف.

هذا سيكون ممكنا فقط إذا v_k و y_k تحققان شرط الانحناء الأتي:

$$v_k^T y_k > 0 \qquad (20)$$

ويمكن رؤيته بسهولة بإعادة ضرب (19) بـ v_k^T. عندما تكون f محدبة بحدة، المتباينة (20) ستتحقق لأي نقطتين x_k و x_{k+1}. ومن ناحية ثانية، لا يتحقق هذا الشرط دائما

للدالة غير المحدبة، وفي هذه الحالة نحتاج إلى اعتماد المعادلة (20) بشكل صريح، بفرض توسيع على إجراء خط البحث الذي يختار λ. الشرط (20) مضمون التحقيق إذا فرضنا شروط Wolfe أو شروط Wolfe القوية على خط البحث. ولتأكيد هذا المطلب، نلاحظ من (18) و (شرط Wolfe) أن $g_{k+1}^T v_k \geq c_2 g_k^T v_k$ ، ولهذا السبب:

$$y_k^T v_k \geq (c_2 - 1)\lambda_k g_k^T d_k \qquad (21)$$

بما أن $c_2 < 1$ و d_k هو اتجاه الانحدار، فالطرف في جهة اليمين يكون موجبا، ويتحقق شرط الانحناء (20).

عندما يتحقق شرط الانحناء، يكون حل معادلة القاطع (19) دائما B_{k+1}. في الواقع، تفسح مجالا لعدد غير منته من الحلول، لان هناك $n(n+1)/2$ من درجات الحرية في مصفوفة متناظرة، ومعادلة القاطع تمثل شروط n فقط.

لتحديد B_{k+1} بصورة وحيدة، نفرض شرط الإضافة (additional condition) وبه تتحقق دالة القاطع B_{k+1} من بين كل مصفوفات التناظر، التي تكون إلى حد ما، اقرب إلى المصفوفة الحالية المستمرة B_k . نحل المسالة،

$$\min_B \|B - B_k\| \qquad (22\ a)$$

وفقا لـ

$$B = B^T,\ Bv_k = y_k \qquad (22\ b)$$

حيث v_k و y_k تحقق (20) و B_k مصفوفة متناظرة وموجبة التعريف. العديد من محددات المصفوفة يمكن استخدامها في (22 a)، وكل محدد يعطي التقدم لطريقة QN المختلفة. المحدد الذي يسمح بحل اسهل من مسالة التقليل إلى الحد الأدنى (22)، والتي تعطي تقدما لقياسات طريقة الامثلية الثابتة، هي شكل محدد Frobenius الموزون

$$\|A\|_W \equiv \left\|W^{1/2} A W^{1/2}\right\|_F \qquad (23)$$

حيث $\|\cdot\|_F$ معرف بـ $\|C\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} cij^2$.

W يمكن اختياره كأي مصفوفة تحقق العلاقة $Wy_k = v_k$. ويستطيع القارئ أن يفرض أن $W = \overline{G}_k^{-1}$ حيث $\overline{G}_k$ هو معدل Hessian المعرف بـ

$$\overline{G}_k = \left[\int_0^1 \nabla^2 f(x_k + \tau\lambda_k d_k) d\tau \right] \qquad (24)$$

وان

$$y_k = \overline{G}_k \lambda_k d_k = \overline{G}_k v_k \qquad (25)$$

نستنتج من مبرهنة Taylor's ، أن الاختيار من مصفوفة الوزن W، المحدد (23) هو بعدي، وهي خاصية مرغوب فيها لأننا لا نرغب في الحل (22) للاعتماد على وحدات المسالة.
بهذه المصفوفة الموزونة وهذا المحدد، يكون الحل الوحيد لـ (22) هو

$$\text{(DFP)} \qquad B_{k+1} = (I - \rho_k y_k v_k^T) B_k (I - \rho_k v_k y_k^T) + \rho_k y_k y_k^T \qquad (26)$$

حيث

$$\rho_k = \frac{1}{y_k^T v_k} \qquad (27)$$

هـذه الصـيغة تسـمى صيغـة DFP المحدثـة، وهـي الصـيغة التـي أقترحهـا أصـلا Davidon, (1959) ، ومن ثم درسها ونفذها وبسطها Fletcher, Powell (1963) .
يمكن تطوير شرط H_k بوساطة النظير الآتي لـ (22):

$$\min_H \| H - H_k \| \qquad (28\ a)$$

وفقا لـ

$$H = H^T,\ Hy_k = v_k \qquad (28\ b)$$

حيث أن المصفوفة W هي الآن أية مصفوفة تحقق $Wv_k = y_k$. (للتأكيد، نفرض ثانيـة أن W معطاة بوساطة معدل Hessian $\overline{G}_k$ معرفة في (24)).
الحل الوحيد H_{k+1} لـ (28) معطاة بـ

$$\text{(BFGS)} \qquad H_{k+1} = (I - \rho_k v_k y_k^T) H_k (I - \rho_k y_k v_k^T) + \rho_k v_k v_k^T \qquad (29)$$

حيث

$$\rho_k = \frac{1}{y_k^T v_k} \qquad (30)$$

2-2 طرائق QN للقياسات الكبيرة والذاكرة المحدودة:
Large-Scale Quasi-Newton Methods and Limited Memory:

في هذه الفقرة سنناقش خوارزمية الاتجاه المترافق الخاصة التي هي مناسبة لمسائل القياس الكبير.

نبـدأ وصـفنا لطريقـة L-BFGS بإعـادة اسـتدعائها مـن أصـلها، وهـي طريقـة BFGS . كل خطوة من طريقة BFGS لها الشكل الآتي:

$$x_{k+1} = x_k - \lambda_k H_k g_k, \qquad k = 1,2,3,... \qquad (31)$$

حيث λ_k هي طول الخطوة، و H_k تحدث في كل تكرار بوساطة الصيغة

$$H_{k+1} = V_k^T H_k V_k + \rho_k v_k v_k^T \qquad (32)$$

(انظر (29)، حيث)

$$\rho_k = \frac{1}{y_k^T v_k}, \qquad V_k = I - \rho_k y_k v_k^T \qquad (33)$$

و

$$v_k = x_{k+1} - x_k, \qquad y_k = g_{k+1} - g_k \qquad (34)$$

نقول أن المصفوفة H_{k+1} يمكن الحصول عليها بتحديث H_k وباستخدام الزوج $\{v_k, y_k\}$.

والآن نصف عملية التحديث بتفصيل اكثر. فعند التكرار k، يكون التكرار المستمر هو x_k ومجموعـة أزواج المتجـه تتضـمن $\{v_i, y_i\}$ لكـل $i = k-m,...,k-1$. نختـار أولا بعـض تقريبات Hessian الأولية H_k^0 (لا تشبه تكرار BFGS الاعتيادية، ثم نسمح لهذا التقريب الأولي أن يتنوع من تكرار إلى تكرار). وبعدها نجد بالتطبيقات المتكررة من الصيغة (32) أن تقريب L-BFGS H_k يحقق الصيغة الآتية:

$$\begin{aligned} H_k &= (V_{k-1}^T ... V_{k-m}^T) H_k^0 (V_{k-m} ... V_{k-1}) \\ &+ \rho_{k-m} (V_{k-1}^T ... V_{k-m+1}^T) v_{k-m} v_{k-m}^T (V_{k-m+1} ... V_{k-1}) \\ &+ \rho_{k-m+1} (V_{k-1}^T ... V_{k-m+2}^T) v_{k-m+1} v_{k-m+1}^T (V_{k-m+2} ... V_{k-1}) \qquad (35) \\ &+ ... \\ &+ \rho_{k-1} v_{k-1} v_{k-1}^T. \end{aligned}$$

من هذا التعبير، نستطيع اشتقاق الإجراء المتكرر لحساب حاصل ضرب $H_k g_k$ بصورة فعالة.

الخوارزمية 2-1 (L-BFGS داراتان متكررتان)

Algorithm (L-BFGS two-loop recursion):

الخطوة 1. $q = g_k$

الخطوة 2. لكل $i = k-1, k-2, ..., k-m$

$\lambda_i = \rho_i v_i^T q,$

$q = q - \lambda_i y_i,$

end (for)

الخطوة 3. $r = H_k^0 q,$.

الخطوة 4. لكل $i = k-m, k-m+1, ..., k-1$

$, \beta = \rho_i y_i^T r$

$r^* = r + v_i(\lambda_i - \beta_i)$

end (for)

الخطوة 5. توقف عندما تكون نتيجة $H_k g_k = r$

طريقة اختيار H_k^0 التي أثبتت أنها فعالة في التمرين إذ تضع $H_k^0 = \gamma_k I$ ، حيث

$$\gamma_k = \frac{v_{k-1}^T y_{k-1}}{y_{k-1}^T y_{k-1}} \qquad (36)$$

γ_k هو عامل القياس وهو محاولة لتقييم حجم مصفوفة Hessian الحقيقية على طول معظم اتجاه البحث السابق.

وكنتيجة لهذا يكون طول الخطوة $\lambda_k = 1$ مقبولا في معظم التكرارات. ومن الضروري أن يكون خط البحث مبنيا على شروط Wolfe أو شروط Wolfe القوية، لذلك تحديثBFGS مستقرة.

أما الذاكرة المحدودة لخوارزمية BFGS فيمكن تعيينها بشكل أساسي كما يأتي:

الخوارزمية 2-2 (L-BFGS) :

الخطوة 1. اختر نقطة ابتدائية x_0 ، وعددا صحيحا $m > 0$ ،

الخطوة 2. $k = 0$ ؛ نكرر

الخطوة 3. اختر H_k^0 (على سبيل المثال، باستخدام (36) ؛

الخطوة 4. احسب $d_k = -H_k g_k$ من الخوارزمية (1-2) ؛

الخطوة 5. احسب $x_{k+1} = x_k + \lambda_k d_k$ ، حيث λ_k تختار لتحقق شرط Wolfe ؛

الخطوة 6. إذا كانت $k > m$

أهمل زوج المتجه $\{v_{k-m}, y_{k-m}\}$ من الخزن؛

الخطوة 7. احسب ثم اخزن $v_k = x_{k+1} - x_k$ ، $y_k = g_{k+1} - g_k$

الخطوة 8. $k = k+1$ ؛

الخطوة 9. ارجع إلى الخطوة (3)

لغاية التقارب.

خلال التكرارات $m-1$ الأولى، تكون الخوارزمية (2-2) مساوية لخوارزمية BFGS إذا كانت المصفوفة الأولية H_0 هي نفسها في كلتا الطريقتين، وإذا كانت L-BFGS تختار $H_k^0 = H_0$ في كل تكرار. في الحقيقة، نستطيع إيجاد طريقة B FGS الاعتيادية بوضع m لقيمة كبيرة إلى حد ما في الخوارزمية (2-2) (اكبر من عدد التكرارات المطلوبة لإيجاد الحل). على أية حال، عندما تقترب m من n (بصورة خاصة، $m > n/2$)، فهذا التقريب يكون أكثر كلفة في مصطلحات وقت وخزن الكومبيوتر من تقريب خوارزمية BFGS.

يكون عمل الاستراتيجية للاحتفاظ بمعظم أزواج تصحيح m السابقة $\{v_i, y_i\}$ جيدا في الجانب العملي، ولم تثبت أية استراتيجية أخرى بعد أن تكون مستمرة بشكل أفضل. عندها يمكن اخذ مقاييس أخرى بعين الاعتبار، فعلى سبيل المثال، مقياس واحد يساعد على استمرار مجموعة أزواج التصحيحات بحيث أن المصفوفة المشكلة بوساطة عناصر v_i لها الشرط الأفضل، ومن ثم تودي إلى تجنب مجموعة أزواج المتجهات بحيث أن بعض من v_i تكون قريبة وواقعة على الخط نفسه.

عندما تكون مصفوفة Hessian كثيفة ولكن لدالة الهدف تركيب قابل للاشتقاق جزئيا، الطرائق التي تستخدم هذا التركيب بصورة طبيعية تتغلب على L-BFGS في التطبيق بوضوح، في مجالات حسابات الدالة. في مجال وقت الحساب، من ناحية ثانية، تكونL-BFGS أكثر فعالية بالنسبة إلى الكلفة الواطئة لتكرارها.

الضعف الرئيس لطريقة L-BFGS هو أنها تتقارب ببطء، والتي غالبا ما تقود إلى عدد كبير نسبيا من حسابات الدالة، وتكون غير فعالة على المسائل غير المشروطة بنسبة عالية، حيث تتضمن مصفوفة Hessian وصفا واسعا من القيم الذاتية.

2-3 تبسيط التحديث (Unrolling the update):

في هذا البند نحاول تحديث الذاكرة المحدودة بطرائق ابسط. في الواقع، نظهر هنا، أن الإنجاز الأكثر وضوحا لتحديث الذاكرة المحدودة لـ BFGS يعد أكثر كلفة من التقريب المبني على التمثيلات المدمجة.

صيغة BFGS المباشرة هي:

$$B_{k+1} = B_k - \frac{B_k v_k v_k^T B_k}{v_k^T B_k v_k} + \frac{y_k y_k^T}{y_k^T v_k} \qquad \text{(BFGS)} \qquad (37)$$

ويمكن كتابتها كآلاتي:

$$B_{k+1} = B_k - a_k a_k^T + b_k b_k^T, \qquad (38)$$

حيث أن المتجهات a_k و b_k تعرف بـ

$$a_k = \frac{B_k v_k}{(v_k^T B_k v_k^T)^{1/2}}, \qquad b_k = \frac{y_k}{(y_k^T v_k)^{1/2}} \qquad (39)$$

نستطيع الاستمرار بحفظ أزواج المتجه $\{v_i, y_i\}$، ولكن استخدام (38) لحساب حاصل ضرب مصفوفة–متجه. وطريقة الذاكرة المحدودة للـ BFGS التي تستخدم هذا التقريب ستنبثق بوساطة تعريف المصفوفة الرئيسة B_k^0 عند كل تكرار ومن ثم التحديث طبقا للصيغة

$$B_k = B_k^0 + \sum_{i=k-m}^{k-1} \left[b_i b_i^T - a_i a_i^T \right]. \qquad (40)$$

أزواج المتجه $\{a_i, b_i\}$، $i = k-m, \ldots, k-1$ تستعاد لاحقا من أزواج المتجه المخزونة $\{v_i, y_i\}$، $i = k-m, \ldots, k-1$ بوساطة الإجراء الأتي:

2-4 تبسيط صيغة BFGS (Unrolling the BFGS formula):

for $i = k-m, ..., k-1$

$b_i \leftarrow y_i / (y_i^T v_i)^{1/2}$;

$a_i \leftarrow B_k^0 v_i + \sum_{j=k-m}^{i-1} \left[(b_j^T v_i) b_j - (a_j^T v_i) a_j \right]$;

$a_i \leftarrow a_i / (v_i^T a_i)^{1/2}$;

end (for)

نلاحظ أن المتجهات a_i يجب إعادة حسابها عند كل تكرار، لأنها تعتمد على زوج المتجه $\{v_{k-m}, y_{k-m}\}$، الذي يزال (removed) في نهاية التكرار k. من ناحية أخرى، يمكن حفظ المتجهات b_i وضربها الداخلي $b_j^T v_i$ من التكرار السابق، لذلك فقط القيم الجديدة b_{k-1} و $b_j^T v_{k-1}$، تحتاج إلى أن تحسب عند التكرار المستمر.

3. خوارزمية قياس ذاتي جديدة مع الصيغة المكررة

New Self-Scaling Algorithm with a recursive formulae

تحديث المصفوفات الشبيهة بنيوتن بخزن محدود:

Updating Quasi-Newton Matrices with limited storage:

درس العالم Nocedal (1980) كيفية استعمال مصفوفات BFGS الشبيهة بنيوتن لتقليل الخزن إلى الحد الأدنى للمسائل التي يكون فيها الخزن حرجا . فقد أعطى صيغة التحديث التي تولد مصفوفات باستخدام معلومات من تكرارات m الأخيرة، حيث m هي أي رقم مجهز بوساطة المستخدم. المصفوفة الشبيهة بنيوتن تحدث في كل تكرار بوساطة رمي المعلومة القديمة واستبدالها بالمعلومة الجديدة. واظهر أن المصفوفات المتولدة لها بعض الخصائص المرغوب فيها. الخوارزميات الناتجة يتم اختبارها عدديا ومقارنتها ببعض الطرائق المعروفة.

لتكن f الدالة التي يتم تقليلها إلى الحد الأدنى، g تدرجه. نعرف

$$v_k = x_{k+1} - x_k \quad \text{و} \quad y_k = g_{k+1} - g_k$$

صيغة تحديث BFGS معطاة بـ

$$H_{k+1} = H_k + \frac{v_k v_k^T}{y_k^T v_k}\left[\frac{y_k^T H_k y_k}{y_k^T v_k} + 1\right] - \frac{1}{y_k^T v_k}\left[v_k y_k^T H_k + H_k y_k v_k^T\right] \qquad (41)$$

ويمكن كتابتها بالشكل الآتي

$$H_{k+1} = H_k + U(v_k, y_k, H_k) \qquad (42)$$

من السهولة إثبات أنه إذا كانت H_k موجبة التعريف و $y_k^T v_k > 0$ ، فان H_{k+1} موجبة التعريف. من الآن فصاعدا سنفرض أن $y_k^T v_k > 0$ لكل k. في طرائق التدرج المترابط يمكن إنجاز ذلك دائما، بتجنب خط البحث المضبوط بشكل كاف. في إنجازاتنا الاعتيادية للطرائق الشبيهة بنيوتن التي تتطلب بصورة عامة $n^2/2 + n/2$ مواقع خزن. وفي إنجازاتنا سنبقي التصحيحات U (المعرفة في (42)) على حدة. من (41) سنرى أن كل تصحيح جديد من H يتطلب خزن المتجهين، المسميين v و y.

وبإعطاء مصفوفة موجبة التعريف و قطرية H_0 يتم إنشاء

$$H_1 = H_0 + U(v_0, y_0, H_0)$$
$$H_2 = H_0 + U(v_0, y_0, H_0) + U(v_1, y_1, H_1)$$

.

.

.

etc.,

حيث $\{v_k\}$ تتولد بوساطة بعض تكرارات التقليل إلى الحد الأدنى. لتكن m الرقم الأقصى من مصفوفات التصحيح U التي يمكن خزنها. بما أن H_0 قطرية، فهذا يعني أن الحد الأقصى لـ متجهات$-n$ التي يمكن استعمالها لتعريف المصفوفة الشبيهة بنيوتن هي $2m+1$. في الوقت الذي تتولد فيه H_m نكون قد وصلنا إلى الحد الأدنى من الخزن المحدود،

$$H_m = H_0 + U(v_0, y_0, H_0) + \ldots + U(v_{m-1}, y_{m-1}, H_{m-1}).$$

يمكننا رمي المصطلح $U(v_0, y_0, H_0)$ واستبداله بواحد يستخدم v_m و y_m. (نفرض أن H_m يتم استخدامه لإيجاد v_m و y_m). لاحظ أن التصحيحات

$U(v_1, y_1, H_1),..., U(v_{m-1}, y_{m-1}, H_{m-1})$ تعتمد على $U(v_0, y_0, H_0)$. ولهذا، إذا أهملنا هذا المصطلح، يجب علينا تغيير الجميع؛ وهذا غير مرغوب فيه . نستطيع تجنب هذه المشكلة بخزن v ، $y^T H$ و $v^T y$ عند كل تكرار. ثم رمي $U(v_0, y_0, H_0)$ والتصحيحات الأخرى لا تتأثر . على أية حال، تقودنا هذه التقريبات إلى صعوبات أخرى: خسارة التعريف الموجب للمصفوفات و/ أو الحصول على الطرائق المتكررة بدون نهاية تربيعية.

والآن سنرى أن استخدام وصف مختلف لتحديث الاستراتيجية لنبذ التحديث يمكن إيجادها بسهولة.

3-1 صيغة التحديث الخاصة لـ BFGS (A Special BFGS Update Formula):

صيغة BFGS (41) يمكن كتابتها أيضا بصيغة الضرب

$$H_{k+1} = (I - \rho_k v_k y_k^T) H_k (I - \rho_k y_k v_k^T) + \rho_k v_k v_k^T \equiv V_k^T H_k V_k + \rho_k v_k v_k^T \qquad (43)$$

حيث $\rho = \dfrac{1}{y_k^T v_k}$

بصورة عامة، لدينا لـ $k+1 \le m$ التحديث العام لـ BFGS

$$\begin{aligned} H_{k+1} &= V_k^T V_{k-1}^T ... V_0^T H_0 V_0 ... V_{k-1} V_k \\ &+ V_k^T ... V_1^T \rho_0 v_0 v_0^T V_1 ... V_k \\ &. \\ &. \\ &+ V_k^T V_{k-1}^T \rho_{k-2} v_{k-2} v_{k-2}^T V_{k-1} V_k \\ &+ V_k^T \rho_{k-1} v_{k-1} v_{k-1}^T V_k \\ &+ \rho_k v_k v_k^T . \end{aligned} \qquad (44)$$

للحالة $k+1 > m$ لدينا التحديث الخاص

$$\begin{aligned} H_{k+1} &= V_k^T V_{k-1}^T ... V_{k-m+1}^T H_0 V_{k-m+1} ... V_{k-1} V_k \\ &+ V_k^T ... V_{k-m+2}^T \rho_{k-m+1} v_{k-m+1} v_{k-m+1}^T V_{k-m+2} ... V_k \\ &. \\ &. \\ &+ V_k^T \rho_{k-1} v_{k-1} v_{k-1}^T V_k \\ &+ \rho_k v_k v_k^T . \end{aligned} \qquad (45)$$

المصفوفات المعرفة بـ (44) و (45) ستسمى مصفوفات BFGS الخاصة.

الخوارزمية التي وصفت الآن تستخدم صيغة الضرب. سنعبر عنها باستخدام صيغة الجمع. مرة أخرى، لتكن m عدد التصحيحات المخزونة. تحديث m الأول سيكون تحديث ألـ BFGS الاعتيادية.

$$H_{i+1} = H_i + U(v_i, y_i, H_i)\,, \qquad i = 0,1,..., m-1.$$

من (45) لدينا H_{m+1} ويمكن حسابها كالآتي:

لتكن $\hat{H}_1 = H_0$ لكل $j = 1,2,..., m$

$$\hat{H}_{j+1} = \hat{H}_j + U(v_j, y_j, \hat{H}_j)\,.$$

$$H_{m+1} = \hat{H}_{m+1}$$

بصورة عامة، لأي $k > m$ لتكن $v = k+1-m$ ، فان H_{k+1} معطاة بـ

$$\begin{cases} \hat{H}_v = H_0. \\ \text{For } j = v, v+1,..., k, \qquad \hat{H}_{j+1} = \hat{H}_j + U(v_j, y_j, \hat{H}_j). \\ H_{k+1} = \hat{H}_{k+1}. \end{cases} \tag{46}$$

باستخدام وصف صيغة الجمع من الضروري إعادة حساب كل التصحيحات عند كل خطوة.

3-2 الخوارزمية الجديدة التي تحقق شرط QN:

The new algorithm which satisfies Quasi-Newton condition:

$$H_{k+1} = \left(H_k - \frac{H_k y_k v_k^T + v_k y_k^T H_k}{y_k^T v_k} \right) + \left(\sigma + \frac{y_k^T H_k y_k}{y_k^T v_k} \right) \frac{v_k v_k^T}{y_k^T v_k}$$

$$H_{k+1} = H_k - \frac{H_k y_k v_k^T}{y_k^T v_k} - \frac{v_k y_k^T H_k}{y_k^T v_k} + \sigma \frac{v_k v_k^T}{y_k^T v_k} + \frac{y_k^T H_k y_k}{y_k^T v_k} \cdot \frac{v_k v_k^T}{y_k^T v_k}$$

$$H_{k+1}y_k = H_k y_k - \frac{H_k y_k v_k^T y_k}{y_k^T v_k} - \frac{v_k y_k^T H_k y_k}{y_k^T v_k} + \sigma \frac{v_k v_k^T y_k}{y_k^T v_k} + \frac{y_k^T H_k y_k}{y_k^T v_k} \cdot \frac{v_k v_k^T y_k}{y_k^T v_k}$$
$$= H_k y_k - H_k y_k - H_k y_k + \sigma v_k + H_k y_k$$

$\therefore H_{k+1}y_k = \sigma_k v_k$ (QN شرط).

لقد تبين عمليا أن σ هو قياس $\sigma > 0$. شرط QN يتحقق لكل قيم σ . لقد اقترحنا العديد من القيم لـ σ ، على سبيل المثال، إذا كانت

$\sigma = 2$ (الخوارزمية الجديدة) (47)

لاحظ أن

إذا كانت $\sigma = 1$ ، فإنها تسمى BFGS القياسية.

إذا كانت $\sigma = \dfrac{y_k^T H_k y_k}{y_k^T v_k}$ ، فإنها تسمى Al-Bayati (1991) .

إذا كانت $\sigma = \dfrac{y_k^T v_k}{4v_k^T g_{k+1} + 2v_k^T g_k - 6(f_{k+1} - f_k)}$ ، فإنها تسمى Biggs (1973) .

أما إذا كانت $\sigma = 2$ فهي تمثل الدراسة الجديدة (present study).

3-3 اشتقاق الخوارزمية الجديدة (The derivation of new algorithm):

عندما $\sigma = 2$ نحصل على

$$H^* = \left(H - \frac{Hyv^T + vy^T H}{y^T v} \right) + \left(2 + \frac{y^T Hy}{y^T v} \right) \frac{vv^T}{y^T v}$$

$$H^* g = \left(Hg - \frac{Hyv^T g + vy^T Hg}{y^T v} \right) + \left(2 + \frac{y^T Hy}{y^T v} \right) \frac{vv^T g}{y^T v}$$

$$H^* g = Hg - \frac{Hyv^T g}{y^T v} - \frac{vy^T Hg}{y^T v} + 2\frac{vv^T g}{y^T v} + \frac{y^T Hy}{y^T v}\frac{vv^T g}{y^T v} \tag{48}$$

$$r^* = r - \beta v + 2\lambda v$$

$$\therefore r^* = r + 2\lambda v - \beta v$$

$$\therefore r^* = r + v(2\lambda - \beta)$$

حيث

$$r = Hg$$
$$\lambda = \rho v^T g$$
$$\beta = \rho y^T r$$
$$\rho = 1 / y^T v$$

<u>3-4 الخطوط العريضة للخوارزمية الجديدة Outlines of the new algorithm</u>:-

الخطوة (1): نبدأ بـ x_0 ، عدد صحيح $m > 0$ ،

الخطوة (2): نضع $k = 0$ ، نكرر؛

الخطوة (3): نحسب H_k^0 باستخدام المعادلة (36) .

الخطوة (4): نحسب $d_k = -g_k$ من الخوارزمية (2-2).

الخطوة (5): نضع $x_{k+1} = x_k + \lambda_k d_k$ ، حيث λ_k تختار لتحقق شروط Wolfe.

الخطوة (6): نضع $r = -H_k g_k$ ،

الخطوة (7): $r^* = r + v(2\lambda - \beta)$ ،

الخطوة (8): إذا كانت $k > m$ ،

أهمل زوج المتجه $\{v_{k-m}, y_{k-m}\}$ من الخزن،

الخطوة (9): نحسب ونخزن $v_k = x_{k+1} - x_k$ ، $y_k = g_{k+1} - g_k$ ،

الخطوة (10): نضع $k = k + 1$ ،

الخطوة (11): اختبر التقارب.

وارجع إلى الخطوة (2).

4- النتائج العملية والاستنتاجات

في هذا البحث تم استخدام (7) دوال غير خطية وكل دالة لها (6) حالات مختلفة.

ففي الجدول (4-1) تم استخدام دالة Cantrel عندما: $m = 3$ و

. $n = 4, n =100, n=1000, n= 10000, n= 100000, n= 1000000$

ومن النتائج الإجمالية لهذه الدالة تبين أن هناك نسبة تحسين قدرها % (86.4) بالنسبة إلى NOI مع نسبة تحسين قدرها % (90.2) بالنسبة إلى NOF .

في الجدول (4-2) تم استخدام دالة Miele عندما: $m = 3$ و

.$n = 4, n =100, n=1000, n= 10000, n= 100000, n= 1000000$

ومن النتائج الإجمالية لهذه الدالة تبين أن هناك نسبة تحسين قدرها % (75.4) بالنسبة إلى NOI مع نسبة تحسين قدرها % (88.2) بالنسبة إلى NOF .

في الجدول (4-3) تم استخدام دالة Wood عندما: $m = 3$ و

.$n = 4, n =100, n=1000, n= 10000, n= 100000, n= 1000000$

ومن النتائج الإجمالية لهذه الدالة تبين أن هناك نسبة تحسين قدرها % (75.5) بالنسبة إلى NOI مع نسبة تحسين قدرها % (85.8) بالنسبة إلى NOF .

في الجدول (4-4) تم استخدام دالة Powell عندما: $m = 3$ و

.$n = 4, n =100, n=1000, n= 10000, n= 100000, n= 1000000$

ومن النتائج الإجمالية لهذه الدالة تبين أن هناك نسبة تحسين قدرها % (92.3) بالنسبة إلى NOI مع نسبة تحسين قدرها % (92.7) بالنسبة إلى NOF .

في الجدول (4-5) تم استخدام دالة Sum عندما: $m = 3$ و

.$n = 4, n =100, n=1000, n= 10000, n= 100000, n= 1000000$

ومن النتائج الإجمالية لهذه الدالة تبين أن هناك نسبة تحسين قدرها % (61.7) بالنسبة إلى NOI مع نسبة تحسين قدرها % (67.4) بالنسبة إلى NOF .

في الجدول (4-6) تم استخدام دالة Osp عندما: $m = 3$ و

.$n = 4, n =100, n=1000, n= 10000, n= 100000, n= 1000000$

ومن النتائج الإجمالية لهذه الدالة تبين أن هناك نسبة تحسين قدرها % (95.0) بالنسبة إلى NOI مع نسبة تحسين قدرها % (93.9) بالنسبة إلى NOF .

في الجدول (4-7) تم استخدام دالة Cubic عندما: m = 3 و

.n = 4, n =100, n=1000, n= 10000, n= 100000, n= 1000000

ومن النتائج الإجمالية لهذه الدالة تبين أن هناك نسبة تحسين قدرها % (83.1) بالنسبة إلى NOI مع عدم وجود نسبة تحسين بالنسبة إلى NOF .

ومن النتائج الإجمالية لهذه الدوال، تبين أن هناك تحسين إلى (NOI) لجميع الدوال التي تم تنفيذها.

أما بالنسبة إلى (NOF) فقد تبين أن (6) دوال من مجموع (7) دوال تكون ذات نتائج أفضل ودالة بنتائج مكافئة تقريبا.

كما إننا نلاحظ تحسينا واضحا بالنسبة إلى الوصول إلى الحد الأدنى للدالة.

الجدول (4-1) مقارنة بين طريقة BFGS والطريقة الجديدة لدالة Cantrel

BFGS σ =1				(New) (47)			
Function	**No. I**	**No. F**	Fun. Value		**No. I**	**No. F**	Fun. Value
n=4 m=3	**38**	**46**	$3.926*10^{-13}$		**37**	**45**	$2.867*10^{-14}$
n=100 m=3	**40**	**46**	$5.684*10^{-12}$		**35**	**38**	$9.315*10^{-13}$
n=1000 m=3	**38**	**42**	$2.621*10^{-10}$		**34**	**41**	$9.869*10^{-12}$
n=10000 m=3	**38**	**45**	$1.688*10^{-9}$		**44**	**56**	$6.324*10^{-11}$
n=100000 m=3	**45**	**49**	$8.932*10^{-10}$		**30**	**39**	$1.871*10^{-9}$
n=1000000 m=3	**44**	**58**	$7.956*10^{-8}$		**30**	**39**	$9.440*10^{-9}$
Total	**243**	**286**		Total	**210**	**258**	

Tools	BFGS $\sigma = 1$	(New) **(47)**
No. I	100%	86.4 %
No. F	100%	90.2 %

الجدول (4-2) مقارنة بين طريقة BFGS والطريقة الجديدة لدالة Miele

	BFGS σ =1			**(New) (47)**			
Function	**No. I**	**No. F**	Fun. Value		**No. I**	**No. F**	Fun. Value
n=4 m=3	**378**	**436**	$2.170*10^{-16}$		**181**	**246**	$1.155*10^{-18}$
n=100 m=3	**58**	**70**	$6.359*10^{-17}$		**124**	**176**	$4.729*10^{-16}$
n=1000 m=3	**129**	**160**	$1.325*10^{-12}$		**99**	**131**	$3.587*10^{-15}$
n=10000 m=3	**184**	**203**	$1.263*10^{-11}$		**90**	**119**	$2.178*10^{-11}$
n=100000 **m=3**	**106**	**124**	$6.979*10^{-12}$		**164**	**224**	$2.054*10^{-15}$
n=1000000 m=3	**107**	**128**	$1.136*10^{-10}$		**67**	**93**	$3.522*10^{-9}$
Total	**962**	**1121**		Total	**725**	**989**	

Tools	BFGS $\sigma = 1$	(New) **(47)**
No. I	100%	75.4 %
No. F	100%	88.2 %

الجدول (4-3) مقارنة بين طريقة BFGS والطريقة الجديدة لدالة Wood

BFGS σ =1				(New) (47)			
Function	**No. I**	**No. F**	Fun. Value		**No. I**	**No. F**	Fun. Value
n=4 m=3	**110**	**142**	$7.646*10^{-24}$		**109**	**156**	$9.873*10^{-24}$
n=100 m=3	**130**	**165**	$6.693*10^{-20}$		**63**	**93**	$8.811*10^{-21}$
n=1000 m=3	**139**	**174**	$8.050*10^{-21}$		**107**	**159**	$4.133*10^{-21}$
n=10000 m=3	**124**	**154**	$7.322*10^{-20}$		**127**	**172**	$5.693*10^{-21}$
n=100000 **m=3**	**155**	**196**	$5.085*10^{-18}$		**85**	**128**	$5.657*10^{-25}$
n=1000000 m=3	**127**	**160**	$6.279*10^{-18}$		**102**	**142**	$4.793*10^{-17}$
Total	**785**	**991**		Total	**593**	**850**	

Tools	BFGS σ = 1	(New) **(47)**
No. I	100%	75.5 %
No. F	100%	85.8 %

الجدول (4-4) مقارنة بين طريقة BFGS والطريقة الجديدة لدالة Powell

BFGS σ =1				**(New) (47)**			
Function	**No. I**	**No. F**	Fun. Value		**No. I**	**No. F**	Fun. Value
n=4 m=3	**118**	**140**	$1.204*10^{-19}$		**118**	**148**	$2.739*10^{-20}$
n=100 m=3	**154**	**180**	$1.499*10^{-15}$		**155**	**211**	$1.596*10^{-20}$
n=1000 m=3	**167**	**199**	$2.599*10^{-14}$		**328**	**398**	$3.043*10^{-18}$
n=10000 m=3	**398**	**483**	$2.912*10^{-17}$		**223**	**300**	$1.534*10^{-19}$
n=100000 **m=3**	**292**	**461**	$1.011*10^{-21}$		**251**	**321**	$1.081*10^{-17}$
n=1000000 m=3	**285**	**342**	$3.558*10^{-19}$		**230**	**295**	$1.561*10^{-19}$
Total	**1414**	**1805**		Total	**1305**	**1673**	

Tools	BFGS σ = 1	(New) **(47)**
No. I	100%	92.3 %
No. F	100%	92.7 %

الجدول (4-5) مقارنة بين طريقة BFGS والطريقة الجديدة لدالة Sum

BFGS σ =1				**(New) (47)**			
Function	**No. I**	**No. F**	Fun. Value		**No. I**	**No. F**	Fun. Value
n=4 M=3	**29**	**30**	$6.818*10^{-14}$		**19**	**20**	$1.837*10^{-14}$
n=100 m=3	**40**	**43**	$1.124*10^{-10}$		**25**	**28**	$1.147*10^{-10}$
n=1000 m=3	**46**	**52**	$1.255*10^{-8}$		**29**	**35**	$9.694*10^{-9}$
n=10000 m=3	**51**	**59**	$4.791*10^{-6}$		**32**	**40**	$3.462*10^{-6}$
n=100000 **M=3**	**57**	**68**	$5.089*10^{-4}$		**35**	**46**	$1.224*10^{-3}$
n=1000000 m=3	**54**	**61**	$4.331*10^{-4}$		**31**	**42**	$3.527*10^{-4}$
Total	**277**	**313**		Total	**171**	**211**	

Tools	BFGS σ = 1	(New) **(47)**
No. I	100%	61.7 %
No. F	100%	67.4 %

الجدول (4-6) مقارنة بين طريقة BFGS والطريقة الجديدة لدالة Osp

BFGS σ =1				**(New) (47)**			
Function	**No. I**	**No. F**	Fun. Value		**No. I**	**No. F**	Fun. Value
n=4 m=3	**33**	**34**	$3.173*10^{-15}$		**23**	**24**	$2.403*10^{-15}$
n=100 m=3	**74**	**77**	$2.340*10^{-15}$		**66**	**68**	$5.697*10^{-16}$
n=1000 m=3	**209**	**219**	$1.037*10^{-15}$		**220**	**229**	$7.418*10^{-16}$

n=10000 m=3	**740**	**788**	$9.729*10^{-16}$		**740**	**762**	$6.020*10^{-16}$
n=100000 m=3	**3421**	**3603**	$9.038*10^{-16}$		**3196**	**3325**	$2.032*10^{-16}$
n=1000000 m=3	**3247**	**3423**	$8.256*10^{-16}$		**3091**	**3237**	$3.673*10^{-16}$
Total	**7724**	**8144**		Total	**7336**	**7645**	

Tools	BFGS σ = 1	(New) **(47)**
No. I	100%	95.0 %
No. F	100%	93.9 %

الجدول (4-7) مقارنة بين طريقة BFGS والطريقة الجديدة لدالة Cubic

BFGS σ =1				**(New) (47)**			
Function	**No. I**	**No. F**	Fun. Value		**No. I**	**No. F**	Fun. Value
n=4 m=3	**36**	**45**	$8.657*10^{-25}$		**28**	**47**	$2.465*10^{-32}$
n=100 m=3	**39**	**53**	$8.678*10^{-25}$		**34**	**57**	$1.011*10^{-28}$
n=1000 m=3	**43**	**56**	$6.513*10^{-23}$		**34**	**56**	$1.163*10^{-23}$
n=10000 m=3	**38**	**53**	$6.296*10^{-22}$		**33**	**59**	$1.686*10^{-24}$
n=100000 m=3	**39**	**50**	$6.145*10^{-22}$		**34**	**58**	$8.007*10^{-23}$
n=1000000 m=3	**42**	**62**	$8.066*10^{-20}$		**34**	**58**	$2.221*10^{-23}$
Total	**237**	**319**		Total	**197**	**335**	

Tools	**BFGS σ = 1**	**(New) (47)**
No. I	**100%**	**83.1 %**
No. F	**100%**	**105.0 %**

4-1 الاستنتاجات:

1. إن الطريقة المقدمة (المقترحة) في هذا البحث تمتاز بإمكانيتها على حل المسائل ذات الأبعاد الكبيرة باستخدام الحاسبات الشخصية ذات السعة الاعتيادية. حيث تم حل مسائل لمختلف الدوال وبعدد متغيرات يصل إلى حد **1000000** متغير .
2. تمتاز هذه الطريقة بإمكانية التوصل إلى حل المسالة بأقل عدد من التكرارات NOI حيث تراوحت نسبة التحسين ما بين % 95.0 و % 61.7 لجميع الدوال التي تم تنفيذها.
وكذلك تم الحصول على النتائج بعدد حسابات الدالة NOF أقل حيث تراوحت نسبة التحسين مابين % 93.9 و % 67.4 بالنسبة إلى ست دوال من اصل سبع دوال لا خطية تم تنفيذها.
3. أعطت هذه الطريقة نتائج أفضل لقيمة الدالة Function value. حيث كانت هناك فروقات واضحة بين نتائج هذه الطريقة وطريقة BFGS الاعتيادية بعد التحديث.

المصادر

[1] Al-Bayati, A.Y. (1991) "A New Family of Self-Scaling Variable Metric Algorithms for unconstrained Optimization", Mosul University, **Journal of Education and Science**, Vol. 12, PP. 25-54.

[2] Biggs, M.C. (1973) "A note on minimization algorithms which make use of the non-quadratic properties of the objective function", **Journal of Institute of Mathematics and Its Application**, Vol. 12, PP. 337-338.

[3] Broyden, C.G. (1970) "The Convergence of a class of double-rank minimization algorithms part I and part II ", **Journal of Institute of Mathematics and Its Applications** ,Vol. 6, PP. 76-90 and 222-231.

[4] Dixon, L.C.W. (1972) **Non-Linear Optimization**, London , English Universities Press.

[5] Fletcher, R. and C.M. Reeves (1964) "Function Minimization by Conjugate Gradients". **Computer Journal**, Vol. 7, PP.149-154.

[6] Fletcher, R. (1970) "A New Approach to Variable Metric Algorithms", **The Computer Journal** ,Vol. 13, PP. 317-322.

[7] Fletcher, R. (1987) **Practical Methods of Optimization** , New York, Brisbane, Toronto and Singapore, John Wiley and Sons, Chi Chester, PP. 75.

[8] Goldfarb, D. (1970) "A Family of Variable Metric Methods Derived by Variational Means", **Mathematics of Computation** ,Vol. 24, PP. 23-26.

[9] Huang, H.Y. (1970) "Unified Approach to Quadratically Convergent Algorithms for Function Minimization". **JOTA**. Vol.5, PP. 405-423.

[10] Nocedal. J. (1980) "Updating Quasi-Newton Matrices With Limited Storage", **Mathematical of Computation**,Vol.35,No.151,PP.773-782.

[11] Nocedal, J. (1999) **Numerical Optimization**, PP.173-192 and 199-209.

[12] Oren, S.S. (1974) "Self-Scaling Variable Metric Algorithm, Part II". **Management Science**, 20, PP. 863-874.

[13] Shanno, D.F. (1970) "Conditioning of Quasi-Newton Methods for Function Minimization", **Mathematics of Computation**, Vol.24, PP. 647-657.