 ISSN: 1813 – 1662 (Print)
E-ISSN: 2415 – 1726 (On Line)

Partial Pearson-two (PP2) of quasi newton method for unconstrained optimization

Basheer M. Salih[1] , Khalil K. Abbo[2] , Zeyad M. Abdullah[3]
[1] *College of Education , University of Mosul , Mosul , Iraq*
[2] *College of Computers Sciences and Math. , University of Mosul , Mosul , Iraq*
[3] *College of Computers Sciences and Math. , University of Tikrit , Tikrit , Iraq*

Abstract:
In this paper, we developing new quasi-Newton method for solving unconstrained optimization problems .The nonlinear Quasi-newton methods is widely used in unconstrained optimization[1]. However,. We consider once quasi-Newton which is (Pearson-two) update formula [2], namely, Partial P2. Most of quasi-Newton methods don't always generate a descent search directions, so the descent or sufficient descent condition is usually assumed in the analysis and implementations [3] . Descent property for the suggested method is proved. Finally, the numerical results show that the new method is also very efficient for general unconstrained optimizations [4].

Key words: Unconstrained optimization; Pearson-two QN method ; global convergence.

1.Introduction:

we consider the following unconstrained optimization problem $\min_{x\in R^n} f(x)$ (1)

Where $f : R^n \rightarrow R$ is continuously differentiable function.

Quasi-Newton method is a well-known and useful method for solving unconstrained convex programming and the BFGS method is the most effective quasi-Newton type methods for solving unconstrained optimization problems from the computation point of view. For the current iterate $x_k \in R^n$ and symmetric positive definite matrix $B_k \in R^{n\times n}$, the next iterate is obtained by

$x_{k+1} = x_k + \alpha_k d_k$ (2)

where $\alpha_k > 0$ is a step-size obtained by a one-dimensional line search, and

$d_k = -B_k^{-1}\nabla f(x_k)$ (3)

Is a descent direction B_k^{-1} being available and approximating the inverse of the Hessian matrix of f at x_k. throughout this paper, we use $\|\ \|$ to denote Euclidean vector or matric norm and denote $\nabla f(x_k)$ by g_k .

2. Rank-One Quasi-Newton Methods.

As we have seen the key points of the QN methods is to generate H_{k+1} by means of QN equation. In this section we introduce Pearson-two update that satisfies the quasi -Newton equation.

Let H_k be the inverse Hessian approximation of the $k-th$ iterations. We try updating H_k into H_{k+1} i.e.

$$H_{k+1} = H_k + E_k \quad (4)$$

Where usually E_k is a matrix with lower rank. In the case of rank- one, we have

$$H_{k+1} = H_k + uv^T \quad (5)$$

Where $u, v \in R^n$ by QN equation we obtain

$$H_{k+1}y_k = (H_k + uv^T)y_k = s_k$$

That is

$$(v^T y_k)u = s_k - H_{k+1}y_k \quad (6)$$

This indicate that u must be in the direction of $S_k - H_k y_k$. Assume that $S_k - H_k y_k \neq 0$ and that the vector v satisfies $v^T y_k \neq 0$, then it follows from (5) and (6) that, and put we put $v = s_k$ in (6) we obtain the following updating formula

$$H_{k+1} = H_k + \frac{(s_k - H_k y_k)s_k^T}{s_k^T y_k} \quad (7)$$

Which called Pearson-two (P2) formula [5] .It is easy to see P2 is not symmetric . The main drawback of the Pearson -two QN update (P2) in general does not retain the positive definiteness of H_k hence , the search directions generated by them in general not descent directions to over to this drawback , in the following section we will introduce new type of algorithms based on Pearson-two QN optimization techniques called partial Pearson -two (PP2) methods . We end this section with PP2 QN algorithms.

Algorithm (Pearson-two QN method) [6],[7].

step 1: Given initial point $x \in R^n$ and a positive definite matrix $H_1 \in R^{n\times n}$. Let $\varepsilon > 0$ and set $k = 0$.

Step 2: calculate $g_1 = g(x_1)$ test a criterion for stopping the iterations for example $\|g_k\| < \varepsilon$, then stop otherwise let $d_1 = -H_1 g_1$ and continue with step3.

Step 3: calculate step length α_k such that wolf condition

$$f(x_k + \alpha d_k) \leq f(x_k) + \rho\alpha g_k^T d_k$$

and satisfied .

Step 4: set $x_{k+1} = x_k + \alpha_k d_k$

Step 5: calculate g_{k+1}

Step 6: Test a criterion for stopping the iterations , for example $\|g_k\| < \varepsilon$ then stop.

Step 7: update H_{k+1} ,P2 let $d_{k+1} = -H_k g_{k+1}$

Set $k = k+1$ go to step 3.

3. Partial Pearson-two (PP2) Quasi-Newton Methods

This section is concerned with developing partial Pearson-two QN methods for solving unconstrained optimization problem defined in equation (1) , where the objective function $f(x), x \in R^n$ is continuously differentiable and bounded from below , starting from an initial point x , and a position definite matrix H_1. The classical Pearson-two QN method with line search is as follows

$$x_{k+1} = x_k + \alpha_k d_k \quad (8)$$

Where

$$d_1 = -H_1 g_1$$

And

$$d_{k+1} = -H_{k+1} g_{k+1} \quad (9)$$

Where H_{k+1} defined in equation

$$H_{k+1} = H_k + \frac{(s_k - H_k y_k) v_k^T}{v_k^T y_k} \quad (10)$$

therefore ,

$$d_{k+1} = -[H_k + \frac{(s_k - H_k y_k) v_k^T}{v_k^T y_k}] g_{k+1}$$

Or

$$d_{k+1} = -[H_k g_{k+1} + \frac{v_k^T g_{k+1}}{v_k^T y_k}(s_k - H_k y_k)] \quad (11)$$

Since

$$H_k g_{k+1} = H_k g_{k+1} - H_k g_k + H_k g_k$$

$$H_k g_{k+1} = H_k y_k + H_k g_k \quad (12)$$

from equations (11) and (12) we get

$$d_{k+1} = -[H_k y_k + H_k g_k + \frac{v_k^T g_{k+1}}{v_k^T y_k} s_k - \frac{v_k^T g_{k+1}}{v_k^T y_k} H_k y_k] \quad (13)$$

Or

$$d_{k+1} = -[H_k g_k + (1 - \frac{v_k^T g_{k+1}}{v_k^T y_k}) H_k y_k + \frac{v_k^T g_{k+1}}{v_k^T y_k} s_k]$$

We call the algorithms defined by equation (9) and (13) general partial Pearson-two (PP2) algorithms where $v_k = s_k$. At this summarize the proposed general partial Pearson-two algorithm as follows:

algorithm (Partial Pearson-two QN method)

step 1: Given initial point $x \in R^n$ and a positive definite matrix $H_1 \in R^{n \times n}$. Let $\varepsilon > 0$ and set $k \le 1$.

Step 2: calculate g_1 test a criterion for stopping the iterations, if satisfied $\|g_k\| < \varepsilon$, then stop otherwise let $d_1 = -H_1 g_1$ and continue with step3.

Step 3: calculate the step size α_k such that Wolfe conditions

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

And

$$d_k^T g(x_k + \alpha_k d_k) \ge \sigma d_k^T g_k$$

satisfied .

Step 4: set $x_{k+1} = x_k + \alpha_k d_k$

Calculate g_{k+1}, f_{k+1}

Step 5: Test a criterion for stopping the iteration, if satisfied stop otherwise go to step 6.

Step 6: Calculate search direction if $v_k^T y_k \ne 0$ compute search direction from equation (12) with $v = s_k$ go to step7

Otherwise $d_{k+1} = -H_k g_{k+1}$, go to step 3.

Step 7: update H_{k+1} via equation (10) with v as in step 6 . set $k = k+1$ go to step 3.

4. Analysis of the Partial Pearson-two (PP2).

In this subsection we will analysis the partial Pearson –two (PP2) algorithm. Throughout this section we will assume that the objective function $f(x)$ is twice continuously differentiable and denote its matrix of second derivatives by $G(x)$. The starting point of the PP2 algorithm is x_1 and we define the level set.

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

Where $f(x)$ is uniformly convex on D, which implies that f has a unique minimizer x in D.

Assumption (A):

The level set D is convex and there exists positive constants m and M such that

$$m\|z\|^2 \le z^T G(x) z \le M\|z\|^2$$

For all $x \in D$ and $z \in R^n$

The gradient of the $f(x)$ is Lipschitz continuous i.e $\exists L > 0$ such that

$$\|g(x) - g(y)\| \le \|x - y\| \qquad \forall x, y \in D$$

An immediate consequence of assumption (A.1) is that if we define

$$\bar{G} = \int_0^1 G(x_k + \tau s_k) d\tau \quad (14)$$

Then we have

$$y_k = \bar{G} s_k \text{ and } \bar{G}^{-1} y_k = s_k \quad (15)$$

Which implies

$$m_1\|s\|^2 \le y^T s_k \le M_1\|s\|^2 \quad (16)$$

And

$$m_1\|y\|^2 \le y^T s_k \le M_1\|y\|^2 \quad (17)$$

We will denote θ by the angle between the steepest descent direction $-g_k$ and displacement s_k hence

$$-g_k^T s_k = \|g_k\|\|s_k\| \cos\theta \quad (18)$$

As a sequence of the Wolfe conditions,

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

And

$$d_k^T g(x_k + \alpha_k d_k) \ge \sigma d_k^T g_k$$

The angle θ_k will determine important properties about the length of the displacement and decrease in the function per step. Many of these conditions have been proved see [8].

In the following theorems we will show that partial Pearson-two (PP2) generates conjugate direction and satisfies descent property.

Theorem (4.1)

For positive definite quadratic functions the partial Pearson-two (PP2), with inexact line search generates conjugate search directions

i.e

$d_{k+1}^T y_k = -g_{k+1}^T s_k$

Proof:

Consider the search direction defined by equation (13) with $v = s_k$

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

Since for position quadratic function, the equation (15) is true i.e

$$\bar{G}^{-1} y_k = H_k y_k = s_k$$

Therefore

$$d_{k+1} = -H_k g_k + \frac{s_k^T g_{k+1}}{s_k^T y_k} H_k y_k - \frac{s_k^T g_{k+1}}{s_k^T y_k} s_k$$

Multiply both sides by y_k^T

$$d_{k+1} y_k^T = -g_{k+1}^T H_k y_k + \frac{s_k^T g_{k+1}}{s_k^T y_k} y_k^T H_k y_k - s_k^T g_{k+1}$$

$$= -g_{k+1}^T s_k + s_k^T g_{k+1} - s_k^T g_{k+1} = -s_k^T g_{k+1}.$$

Theorem (4.2)

Suppose that α_k satisfies the Wolfe conditions

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

And

$$d_k^T g(x_k + \alpha_k d_k) \ge \sigma d_k^T g_k$$

in the PP2 algorithm, if $-g_{k+1}^T H_k g_k \le g_{k+1}^T H_k g_{k=1}$ then the search directions generated by PP2 algorithm are descent i.e

$$d_{k+1}^T g_k < 0, \ \forall k$$

Proof:

Since $H_1 = I$ and $d_1 = H_1 g_1$ then

$d_1^T g_1 = -\|g_1\|^2 < 0$

Suppose $d_1^T g_k < 0$ or $s_k^T g_k < 0$

Multiplying (13) by g_{k+1}^T with $v = s_k$, we have

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

Note that: $0 < \rho < 1$

By the Wolfe condition

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

$$s_k^T y_k = s_k^T g_{k+1} - s_k^T g_k \ge (\rho - 1) s_k^T g_k > 0$$

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

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

$$g_{k+1}^T H_k y_k = g_{k+1}^T H_k g_{k+1} - g_{k+1}^T H_k g_k \le g_{k+1}^T H_k g_{k+1}$$

Therefore

$$d_{k+1}^T g_{k+1} \le -g_{k+1}^T H_k g_{k+1} + g_{k+1}^T H_k g_{k+1} - \frac{(s_k^T g_{k+1})^2}{s_k^T y_k} < 0 \cdot$$

$$-\frac{(s_k^T g_{k+1})^2}{s_k^T y_k} < 0$$

To prove the global convergence of PP2 algorithm, we use the following algorithm, due to Zoutendijk.

Theorem (Zoutendijk)

Consider any iteration of the form (8) where d_k is a descent direction and α_k satisfies the Wolfe conditions

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

And

$$d_k^T g(x_k + \alpha_k d_k) \ge \sigma d_k^T g_k$$

Suppose that f is bounded below in R^n and assumption (A) hold then

$$\sum_{k=1}^{\infty} \cos^2\theta_k \|g_k\|^2 < \infty \quad \text{.....(19)}$$

Proof (see Zoutendijk) [9].

Inequality (19) implies that

$$\cos^2\theta \|g_k\|^2 \to 0$$

This limit can be used in turn to derive global convergence results for line search algorithms.

5. Numerical experiments.

In this section we report numerical experiments of the proposed method (partial Pearson-two) and classical Pearson-two Quasi-Newton method. Our experiments are performed for 52 non-linear unconstrained optimization problems (functions) in the CUTEr library [10]. Each test problem is made ten experiments with the number of variable 100,200,…, 1000, respectively. In table (1) method examined in our experiments

Table (1) method examined in our experiments

n.	Method name	Description
1	PE	Pearson two QN method
2	PPE	Partial Pearson two QN method

In the line search Procedure, the step-size α_k is chosen so that the Wolfe conditions

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

And

$$d_k^T g(x_k + \alpha_k d_k) \ge \sigma d_k^T g_k$$

Are satisfied with $\rho = 0.1$ and $\sigma = 0.9$. The stopping criterion was $\|g_k\| \le 10^{-6}$.

In this work, we used three codes; where two of the codes are programmer by visual Fortran. The first code was developed by Andrie [11] and improved by Donal and more. The second code developed by Andrei [12] which uses CG algorithms, we improved this code and adapted by using QN algorithms. We developed the third code wing Matlab for results and graphic comparisons.

Table (2) gives the total number of iterations (toit), the total number of function evaluations (tfn) and total time (totime) for solving 520 test problems.

Table (2) comparison between P2 and PP2

n.	Name of Algorithm	Toit	Tfn	Totime
1	P2	118805	416325	134500
2	PP2	108941	376206	122428

In this Figures (1-3) we adopt the performance profiles by Donald and More [13] to compare the performance based on the number of iterations and CPU time. That is, for each method, we plot the fraction ρ of problems for which the method is within a factor tao of the best result. The left side of the figure gives the percentage of the test problems for which a method is the best result, the right side gives the percentage of the test problems that are successfully solved by each of the methods. The top curve is the method that solved the most problems in a result that is within a factor tao of the best results.

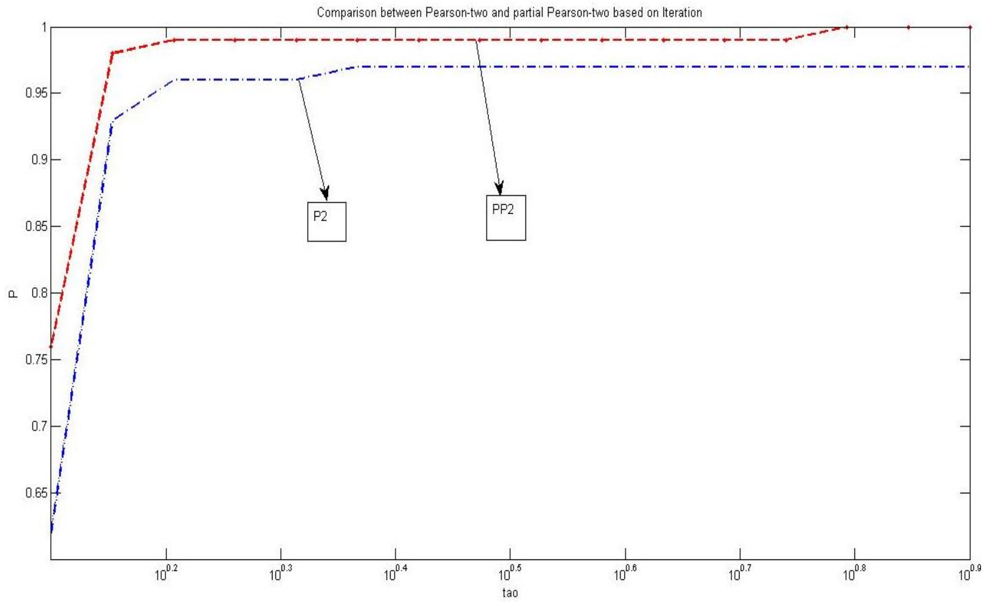

Figure (1) comparison between (P2 and PP2) based on Iteration

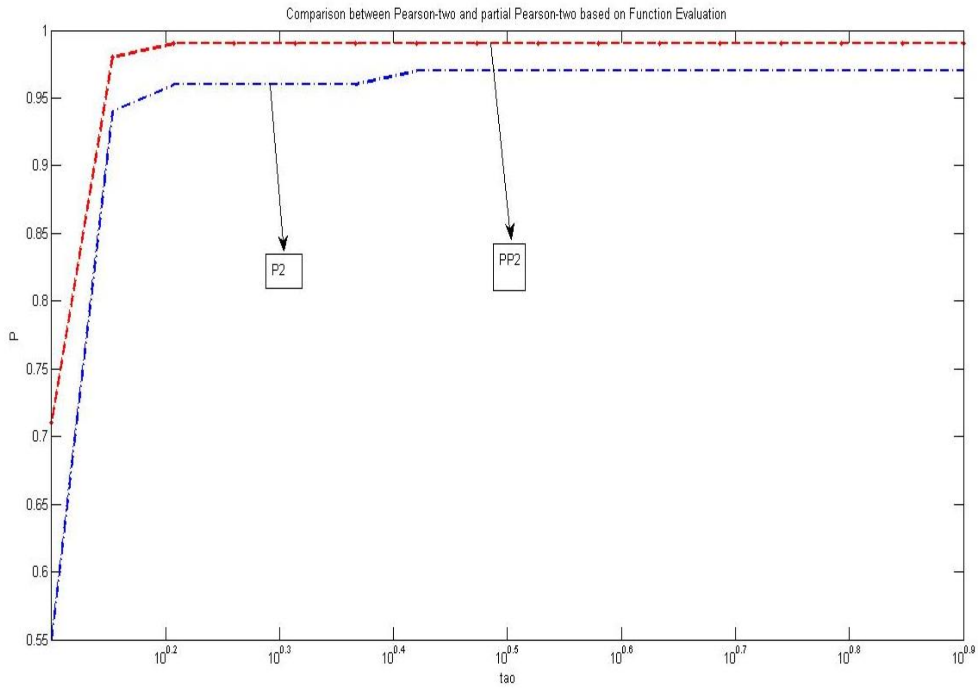

Figure (2) comparison between (P2 and PP2) based on Function evaluation

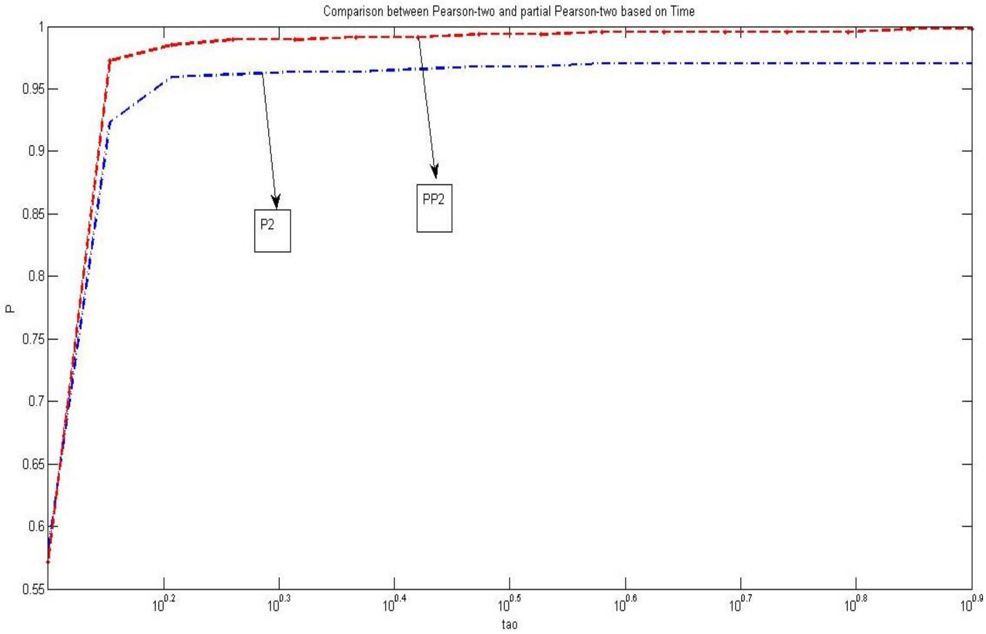

Figure (3) comparison between (P2 and PP2) based on Time

6. Conclusion:

In this study a Partial Pearson-two (PP2) QN method developed for solving large-scale unconstrained optimization problems, in which the Pearson-two (P2) update based on the modified QN equation have applied. An important feature of the proposed method is that it preserves positive definiteness of the updates. The presented method owns global convergence. Numerical results showed that the proposed method is encouraging comparing with the methods Pearson-two (P2) and Partial Pearson-two (PP2).

References:

[1] B.T. Polyak. Introduction to Optimization. Optimization Software, New York, 1987.

[2] Chong K.P.E and Zak, H.S., " An Introduction to optimization", John Wiley & Sons, INC. New york / chichester/ Weinheim/ Brisbane/ Singapore / Toronto, USA. (2001).

[3] W.C. DAVIDON, Variable metric methods for minimization, Atomic Energy Commission Research, (1959). And Development Report AWL-5990, Argonne National Laboratory, Argonne, IL.

[4] Z. X. Wei, G. H. Yu, G. L. Yuan and Z. G. Lian, The superlinear convergence of a modified BFGS-Type method for unconstrained optimization, Computational optimization and applications 29 (2004) 315-332.

[5] J.D. Pearson. "Variable metric methods of minimization" Computer Journal. 12 (1969), 171-178.

[6] Nocedal. J., "Theory of Algorithms for Unconstrained Optimization", 199-242, Acta numarica , USA. (1992).

[7] Sugiki K. Narushima Y. and Yube H. ' Globally convergence three- term conjugate gradient methods that uses secant conditions and generate descent search direction, for unconstrained optimization. J. of optimization theory and Applications 153. (2012).

[8] Byrd H., Nocedal J. and Yuan Y.. Global convergence of a class of Quasi-Newton methods on convex problems SIAM J-NUMERANAL vol. (24), No.(5), (1987).

[9] G. Zoutendijk. Nonlinear programming, computational methods. In J. Abadie, editor, Integer and Nonlinear Programming, pages 37{86. North-Holland, Amsterdam,1970.

[10] Bongartz, I., A.R. Conn, N.I.M. Gould and Ph.L. Toint,. CUTE: Constrained and unconstrained testing environment: ACM Trans. Math. (1995) Software, 21:123-160. http://portal.acm.org/citation.cfm?doid=200979.201043.

[11] Andrei, N., Scaled conjugate gradient algorithms for unconstrained optimization. Computational Optimization and Applications 38, 401-416 (2007).

[12] Andrei, N., A hybrid conjugate gradient algorithm with modified secant condition for unconstrained optimization. ICI Technical Report, February 6, 2008.

[13] Dolan, E.D. and J.J. Moré,. Benchmarking optimization software with performance profiles. Math. Program., 91: 201-203, ((2002). DOI: 10.1007/ s101070100263.

Tikrit Journal of Pure Science 21 (3) 2016 ISSN: 1813 – 1662 (Print)
E-ISSN: 2415 – 1726 (On Line)

صيغة (PP2) الجزئية لطريقة شبه نيوتن في الامثلية الغير مقيدة

بشير محمد صالح خلف[1] ، خليل خضر عبو[2] ، زياد محمد عبد الله[3]
[1] قسم الرياضيات ، كلية التربية للعلوم الصرفة ، جامعة الموصل ، الموصل ، العراق
[2] قسم الرياضيات ، كلية التربية الاساسية ، جامعة تلعفر ، تلعفر ، العراق
[3] قسم الرياضيات ، كلية علوم الحاسوب والرياضيات ، جامعة الموصل ، الموصل ، العراق

الملخص:

في هذا البحث تم تطوير طريقة جديدة من طرق شبيهة نيوتن (P2) واسميناها بطريقة (PP2) الجزئية ، تعتبر طرق شبيهة نيوتن من اكثر الطرق انتشارا لحل مسائل الامثلية غير المقيدة. ، ولان اغلب طرق شبيهة نيوتن لاتولد دائما شرط الانحدار ولذلك فان خاصية الانحدار والانحدار الكافي تفرض عند تحليل وتمثيل هذه الخوارزميات. تم اثبات خاصية الانحدار في الطريقة المقترحة. والنتائج العددية تبين ان الطريقة المقترحة هي ايضا فعالة جداً وممتازة بالمقارنة مع طريقة (P2) الاصلية.