Proceedings of the Federated Conference on
Computer Science and Information Systems pp. 463–466

ISBN 978-83-60810-51-4

Broyden restricted class of variable metric methods and oblique projections

Andrzej Stachurski
Institute of Control and Computation Engineering
Warsaw University of Technology
Nowowiejska 15/19, 00-665 Warsaw, Poland
Email: A.Stachurski@ia.pw.edu.pl

***Abstract*—In the paper the new formulation of the Broyden restricted convex class of updates involving oblique projections is introduced. It is a sum of two terms: the first one containing special oblique projection and the second standard term ensuring verification of the quasi-Newton condition (it is also an oblique projection multiplied by appropriate scalar). The applied oblique projection involves vector defined as the convex, linear combination of the difference between consecutive iterative points and the image of the previous inverse hessian approximation on the corresponding difference of derivatives, i.e. gradients. Formula relating coefficient in the convex combination of vectors in the oblique projection with its counterpart in the standard representation of the Broyden convex class is presented.**

Some preliminary numerical experiments results on two twice continuously differentiable strictly convex functions with increasing dimension are included.

I. Introduction

PROBLEM considered in the current paper is the unconstrained minimization of a sufficiently smooth function

$$\min_{\mathbf{x}\in\mathcal{R}^n} f(\mathbf{x}) \tag{1}$$

Solution method assumes that given starting point $\mathbf{x}^0$ every consecutive approximate solution point is generated according to the following iterative formula

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha^k \mathbf{d}^k, \quad \forall k \geq 0 \tag{2}$$

where $\alpha^k > 0$ is the stepsize coefficient found in the directional minimization and the search direction $\mathbf{d}^k$ is equal to

$$\mathbf{d}^k = -\mathbf{H}^k \mathbf{g}^k \tag{3}$$

Matrix $\mathbf{H}^{k+1}$ is calculated at each step with the aid of vectors: $\mathbf{s}^k = \mathbf{x}^{k+1} - \mathbf{x}^k$, $\mathbf{r}^k = \nabla f(\mathbf{x}^{k+1}) - \nabla f(\mathbf{x}^k)$ and the previous matrix $\mathbf{H}^k$.

Problems of unconstrained functions minimization arise first of all as the result of the least squares approach to solve sets of nonlinear equations (see for instance the problem of determining stresses in RC ring sections with openings in Lechman and Stachurski [11] and Stachurski and Lechman [20]) and identification of parameters appearing in the model in a nonlinear way (as for instance in the augmented Gurson model describing the creation and growth of voids in the porous material considered in Nowak and Stachurski in the sequence of publications [12]- [16]).

Broyden convex class of updates is usually expressed in the following way (see for instance Sun and Yuan [25])

$$\begin{aligned}\mathbf{H}^{k+1} \;=\;& \mathbf{H}^k + \left(1 + \Phi \frac{\left(\mathbf{r}^k\right)^T \mathbf{H}^k \mathbf{r}^k}{\left(\mathbf{r}^k\right)^T \mathbf{s}^k}\right) \frac{\mathbf{s}^k \left(\mathbf{s}^k\right)^T}{\left(\mathbf{r}^k\right)^T \mathbf{s}^k} \\ & - (1-\Phi) \frac{\mathbf{H}^k \mathbf{r}^k \left(\mathbf{r}^k\right)^T \mathbf{H}^k}{\left(\mathbf{r}^k\right)^T \mathbf{H}^k \mathbf{r}^k} \\ & - \Phi \frac{\mathbf{s}^k \left(\mathbf{r}^k\right)^T \mathbf{H}^k + \mathbf{H}^k \mathbf{r}^k \left(\mathbf{s}^k\right)^T}{\left(\mathbf{r}^k\right)^T \mathbf{s}^k}\end{aligned} \tag{4}$$

where Φ is a scalar belonging to the interval $[0,1]$.

In the consecutive section we shall show an alternate updating formula of the form

$$\mathbf{H}^{k+1} = \mathbf{P}^T \mathbf{H}^k \mathbf{P} + \beta \bullet \mathbf{Q}$$

where $\mathbf{P}$ and $\mathbf{Q}$ are oblique projections, i.e. $\mathbf{P}$ sets to null any vector collinear with $\mathbf{r}^k$ and $\mathbf{Q}$ nullifies any vector orthogonal to $\mathbf{s}^k$ and $\mathbf{PP} = \mathbf{P}$ and $\mathbf{QQ} = \mathbf{Q}$. Parameter β is a positive scalar changing from one iteration to another. Reader interested in the theory of oblique projections and their properties may find more infromation for instance in Afriat [1] or Szyld [24].

Section III contains some preliminary computational results obtained with the aid of quasi-newton methods with updates defined by the discussed formula with parameter $\Theta = 1,\ 0$ and $\frac{1}{2}$. Testing examples are two strictly convex functions constructed in the way permitting easily increase their dimensions. In the last section IV conclusions and comments following from the numerical experiments and of general theoretical character are presented.

II. Oblique projections in the formula of the Broyden class updates

Broyden convex class may be equivalently represented by the following updating formula

$$\mathbf{H}^{k+1} = \left(\mathbf{P}^k\right)^T \mathbf{H}^k \mathbf{P}^k + \frac{\mathbf{s}^k \left(\mathbf{s}^k\right)^T}{\left(\mathbf{r}^k\right)^T \mathbf{s}^k} \tag{5}$$

where $\mathbf{P}^k$ is the projection matrix defined as follows

$$\mathbf{P}^k = \mathbf{I} - \frac{\mathbf{r}^k \left[\Theta \mathbf{s}^k + (1-\Theta)\mathbf{H}^k \mathbf{r}^k\right]^T}{\left(\mathbf{r}^k\right)^T \left(\Theta \mathbf{s}^k + (1-\Theta)\mathbf{H}^k \mathbf{r}^k\right)} \tag{6}$$

and parameter $\Theta \in [0,1]$. Parameters Φ and Θ are mutually connected by the following formula

$$\Phi = \Theta^2 \frac{\left(\mathbf{r}_k^T \mathbf{s}_k\right)^2}{\left(\mathbf{r}_k^T \mathbf{u}_k\right)^2} \tag{7}$$

Formal prove showing that when equality (7) holds then formulae (4) and (5) are equivalent will be presented in the forthcoming paper.

A. Involved oblique projections

First, let's show that $\mathbf{P}^k$ is the projection matrix transforming vector $\mathbf{r}^k$ to the null vector $\mathbf{0}$

$$\begin{aligned} \mathbf{P}^k\mathbf{r}^k &= \left(\mathbf{I} - \frac{\mathbf{r}^k\left[\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right]^T}{\left(\mathbf{r}^k\right)^T\left(\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right)}\right)\mathbf{r}^k \\ &= \mathbf{r}^k - \mathbf{r}^k\frac{\left[\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right]^T\mathbf{r}^k}{\left(\mathbf{r}^k\right)^T\left(\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right)} \\ &= \mathbf{0} \end{aligned} \tag{8}$$

Matrix $\mathbf{P}^k$ is an oblique projection (definition and properties of such projections may be found for instance in Afriat [1] or Szyld [24]), because

$$\begin{aligned} &\mathbf{P}^k\mathbf{P}^k \\ &= \left(\mathbf{I} - \frac{\mathbf{r}^k\left[\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right]^T}{\left(\mathbf{r}^k\right)^T\left(\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right)}\right) \\ &\quad \bullet \left(\mathbf{I} - \frac{\mathbf{r}^k\left[\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right]^T}{\left(\mathbf{r}^k\right)^T\left(\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right)}\right) \\ &= \mathbf{I} - \frac{\mathbf{r}^k\left[\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right]^T}{\left(\mathbf{r}^k\right)^T\left(\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right)} \\ &\quad + \frac{\mathbf{r}^k\left[\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right]^T}{\left(\mathbf{r}^k\right)^T\left(\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right)} \\ &\quad \bullet \frac{\left(\mathbf{r}^k\right)^T\left(\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right)}{\left(\mathbf{r}^k\right)^T\left(\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right)} \\ &\quad - \frac{\mathbf{r}^k\left[\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right]^T}{\left(\mathbf{r}^k\right)^T\left(\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right)} \\ &= \mathbf{I} - \frac{\mathbf{r}^k\left[\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right]^T}{\left(\mathbf{r}^k\right)^T\left(\Theta\mathbf{s}^k + (1-\Theta)\mathbf{H}^k\mathbf{r}^k\right)} = \mathbf{P}^k \end{aligned}$$

Second term in formula (5) is also an oblique projection

$$\frac{\mathbf{s}^k\left(\mathbf{s}^k\right)^T}{\|\mathbf{s}^k\|^2} \tag{9}$$

multiplied by a scalar

$$\beta = \frac{\|\mathbf{s}^k\|^2}{\left(\mathbf{r}^k\right)^T\mathbf{s}^k} \tag{10}$$

It is not difficult to show that formula (9) defines an oblique projection.

B. BFGS update and oblique projections

Representation (5) has appeared for the first time in Stachurski [22]. It was proposed there as a new quasi-newton update. Later the author has realized that it's a new representation of the famous convex class of Broyden proposed for the first time in [3]. Such representation is known for many years for the BFGS update (name derived from the family names of its authors Broyden [3], Fletcher [8], Goldfarb [10] and Shanno [17])

$$\begin{aligned} \mathbf{H}^{k+1}_{BFGS} &= \mathbf{H}^k \\ &+ \left(1 + \frac{\left(\mathbf{r}^k\right)^T\mathbf{H}^k\mathbf{r}^k}{\left(\mathbf{r}^k\right)^T\mathbf{s}^k}\right)\frac{\mathbf{s}^k\left(\mathbf{s}^k\right)^T}{\left(\mathbf{r}^k\right)^T\mathbf{s}^k} \\ &- \frac{\mathbf{s}^k\left(\mathbf{r}^k\right)^T\mathbf{H}^k + \mathbf{H}^k\mathbf{r}^k\left(\mathbf{s}^k\right)^T}{\left(\mathbf{r}^k\right)^T\mathbf{s}^k} \end{aligned} \tag{11}$$

In the limitted memory BFGS method (see for instance Xiao [26]) the following representation of the BFGS update is frequently used

$$\begin{aligned} &\mathbf{H}^{k+1} = \\ &\left(\mathbf{I} - \frac{\mathbf{r}^k\left(\mathbf{s}^k\right)^T}{\left(\mathbf{s}^k\right)^T\mathbf{r}^k}\right)^T\mathbf{H}^k\left(\mathbf{I} - \frac{\mathbf{r}^k\left(\mathbf{s}^k\right)^T}{\left(\mathbf{s}^k\right)^T\mathbf{r}^k}\right) \\ &+ \frac{\mathbf{s}^k\left(\mathbf{s}^k\right)^T}{\left(\mathbf{r}^k\right)^T\mathbf{s}^k} \end{aligned} \tag{12}$$

It is easy to notice that formula (12) is represented by formulae (5) and (6) with $\Theta = 1$.

C. DFP update and oblique projections

The second famous update – DFP (proposed originally by Davidon [5] and [6] and further developed by Fletcher and Powell [7])

$$\mathbf{H}^{k+1}_{DFP} = \mathbf{H}^k - \frac{\mathbf{H}^k\mathbf{r}^k\left(\mathbf{r}^k\right)^T\mathbf{H}^k}{\left(\mathbf{r}^k\right)^T\mathbf{H}^k\mathbf{r}^k} + \frac{\mathbf{s}^k\left(\mathbf{s}^k\right)^T}{\left(\mathbf{r}^k\right)^T\mathbf{s}^k} \tag{13}$$

may be also expressed with the aid of oblique projections as follows

$$\begin{aligned} &\mathbf{H}^{k+1} = \\ &\left(\mathbf{I} - \frac{\mathbf{r}^k\left(\mathbf{H}^k\mathbf{r}^k\right)^T}{\left(\mathbf{H}^k\mathbf{r}^k\right)^T\mathbf{r}^k}\right)^T\mathbf{H}^k\left(\mathbf{I} - \frac{\mathbf{r}^k\left(\mathbf{H}^k\mathbf{r}^k\right)^T}{\left(\mathbf{H}^k\mathbf{r}^k\right)^T\mathbf{r}^k}\right) \\ &+ \frac{\mathbf{s}^k\left(\mathbf{s}^k\right)^T}{\left(\mathbf{r}^k\right)^T\mathbf{s}^k} \end{aligned} \tag{14}$$

It is easy to observe that formula (12) is represented by formulae (5) and (6) with $\Theta = 0$. It was shown in Stachurski [21].

III. Numerical experiments

In the current section the results of numerical calculations are presented. They are realized by means of three variants of updates specified by formulae (5-6) with parameter Θ equal to 1 (corresponding to the BFGS method), 0 (DFP method) and $\frac{1}{2}$. Three variants of directional minimization were tested: Armijo directyional minimization ensuring verification of the Armijo condition (identical with the first Godstein test) (it is denoted below in the results table by A.), directional minimization ensuring verification of the Wolfe conditions (denoted below by W.), directional minimization ensuring verification of both Goldstein conditions (denoted by G. respectively).

The first directionasl minimization was realized by setting the starting value of the directional stepsize and its consecutive reduction by some constant coefficient belonging to the interval $(0, 1)$ until the first Goldstein condition is met. In the second directional minimization Wolfe conditions were used as the stopping criterion and in the third variant two Goldstein tests served as the stopping criterion for the directional minimization. In the second and third variant of the directional minimization consecutive approximations of the stepsize length were generated as the minimum point of the parabola approximating function $\tilde{f}(\alpha) = f(\mathbf{x}^k + \alpha * \mathbf{d}^k)$.

A. Test functions

Two stricly convex, n-dimensional functions with increasing dimension were used for testing. Dimensions were equal to $n = 2, 10, 50, 100, 500, 1000, 2000$. The first was obtained by raising up to the second power values of a strictly convex quadratic function with positive values (its minimal value was positive)

$$f_1(\mathbf{x}) = \frac{1}{2}\left[f_{qua}(\mathbf{x})\right]^2, \tag{15}$$

where

$$f_{qua}(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{e})^T \mathbf{Q}(\mathbf{x} - \mathbf{e}) + 1.0 \tag{16}$$

Vector $\mathbf{e}^T = [1, 1, \ldots, 1]$ in formula (16) consists of ones, second derivative matrix $\mathbf{Q}$ was generated randomly (some extra operations ensuring its positive definiteness were involved). First, quadratic matrix $\bar{\mathbf{Q}}$ of size $n \times n$ is created by invoking MATLAB function *rand*. Its elements are numbers belonging to the interval $[0, 1]$. Next, lower triangular matrix $\mathbf{L}$ is created on the basis of matrix $\bar{\mathbf{Q}}$. It has entries with 0 values on the main diagonal and its entries below the main diagonal were identical with that of $\bar{\mathbf{Q}}$ matrix, i.e.

$$\begin{array}{ll} L_{ij} = \bar{Q}_{ij}, & \forall i < j, \\ L_{ij} = 0, & \forall i \geq j \end{array}$$

Finally, matrix $\mathbf{Q}$ is defined in the consecutive step by the formula

$$\tilde{\mathbf{Q}} = \frac{1}{2}\left[\mathbf{L} + \mathbf{L}^T\right]$$

to which the diagonal matrix $\mathbf{D}$ defined as

$$D_{ii} = \sum_{j=1}^{n} \tilde{Q}_{ij} + 1, \quad i = 1, \ldots n$$

is added. The resulting matrix $\mathbf{Q} = \tilde{\mathbf{Q}} + \mathbf{D}$ was a diagonally dominated matrix with nonnegative entries. All entries outside the main diagonal belong to the interval $[0, 1]$. Furthermore, it was positive definite. The last property was checked numerically to verify correctness.

Vector $\mathbf{e}$ is the unique, local and global minimum of the function (15) constructed in this way. Its optimal value is $1/2$. Described construction makes use of the random numbers generator, however data defining the generated problem of a given size together with the starting point were stored in the MATLAB data file with extension *.dat* by means of the *save* command. The file has been loaded to the operational memory during the start of any method by means of the *load* command. It ensures compatibility of the computational results for various methods which were tested. For any assumed dimension – 2, 10, 50, 100, 500, 1000, 2000 an independent problem has been generated. In any case, the way of generating the problem was identical with that described above. Similarly, the starting points were the same for any method for the problem of the given dimension. All calculations were run on the 32-bit personal computer with processor *Intel(R) Pentium(R) 4 CPU 3.20GHz*, with RAM memory of 1GB capacity, working under the Windows XP Professional operational system.

In the second example function f_2 is generated similarly. The only difference is that instead of taking the second power of the quadratic function f_{qua} we assume its natural logarithm, i.e.

$$f_2(\mathbf{x}) = \ln(f_{qua}(\mathbf{x})) \tag{17}$$

B. Results obtained by means of selected members of the Broyden convex class

Three variants of Broyden methods belonging to the convex class in version with oblique projections defined by (5) were implemented. The selected three variants are: BFGS method with $\Theta = 1$, DFP method with $\Theta = 0$ and the third version with $\Theta = \frac{1}{2}$. Every method was implemented with three above mentioned directional minimizations: Armijo (A.), Wolfe (W.) and Goldstein (G.). Hence we considered altogether nine variants of methods. Every method variant has been run with the same MATLAB m-function implementing the directional minimization and on the same set of test problems with increasing dimesion, generated as described above. Stopping criterions were also the same - on the derivative norm and on the minimized function value. Let's notice that we know the optimum goal function value. The results for goal function f_1 are collected in table I and for the second goal function f_2 in table II.

Symbol (M) placed instead of the number of iterations means stop due to overcrossing the maximal number of iterations, set by the user. Symbol (P) denotes the user break by pressing simultaneously combinations of keys CTRL-C and (O) stopping the calculations due to the zero value in the denominator in the updating formula (something theoretically impossible in the exact arithmetic, but on the computer we never carry out calculations in the exact arithmetic). Later appropriate safeguards were introduced.

The obtained results prove the BFGS method superiority over the DFP and the third variant with $\Theta = \frac{1}{2}$. Furthermore, they have shown that for problems of larger dimension the Armijo directional minimization (i.e. decreasing the stepsize from given starting value by a constant coefficient until the first Goldstein test is met) is totally unuseful. The directional minimization with the Wolfe stopping conditions proved to be the best one. Number of iterations of the BFGS method

TABLE I
NUMBER OF ITERATIONS OF BROYDEN CONVEX CLASS WITH DIFFERENT DIRECTIONAL MINIMIZATION FOR FUNCTION f_1

Dir.	Problem size $n =$						
min.	2	10	50	100	500	1000	2000
BFGS method ($\Theta = 1$)							
A.	5	(M)	(M)	(M)	(M)	(M)	–
W.	4	27	70	185	533	1080	821
G.	4	47	139	314	809	1728	1762
DFP method ($\Theta = 0$)							
A.	5	(M)	(M)	(M)	(M)	(M)	–
W.	4	29	76	313	5652	(M)	2986(O)
G.	4	68	249	(M)	(M)	(M)	(P)
variant with $\Theta = \frac{1}{2}$							
A.	5	(M)	(M)	(M)	(M)	(P)	–
W.	4	29	69	190	597	1443	1106
G.	4	58	185	582	1331	3254	3819

TABLE II
NUMBER OF ITERATIONS OF BROYDEN CONVEX CLASS WITH DIFFERENT DIRECTIONAL MINIMIZATION FOR FUNCTION f_2

Dir.	Problem size $n =$						
min.	2	10	50	100	500	1000	2000
BFGS method ($\Theta = 1$)							
A.	23	51	314	627	2787	(M)	(M)
W.	4	17	37	48	116	406	768
G.	4	25	38	52	524	428	1655
DFP method ($\Theta = 0$)							
A.	23	51	311	(M)	(M)	(M)	–
W.	4	17	38	48	116	(M)	(M)
G.	4	20	53	67	289	419	1086
variant with $\Theta = \frac{1}{2}$							
A.	23	51	501	(M)	(M)	(M)	–
W.	4	17	40	48	116	(M)	(M)
G.	4	20	60	75	285	534	906

implemented with the directional minimization ensuring verification of the Wolfe conditions was substantially smaller than in all other considered variants. The only exception were the problems of smallest size equal to 2.

IV. CONCLUSIONS AND COMMENTS

Updates representation with oblique projections gives a deeper look into the structure of the existing variable metric updates. It offers new possibilities in convergence analysis of quasi-newton methods for minimization. It would be then possible to exploit the existing rich algebraic theory of oblique projections. Furthermore it opens the possibility to exploit in context of the limitted memory methods any member of the Broyden convex class. We are not restricted to the BFGS as it was up till now.

REFERENCES

[1] S.N. Afriat, "Orthogonal and oblique projectors and the characteristics of pairs of vector spaces.", *Proc. Camb. Philos. Soc.*, vol. 53, pp. 800816, 1957.

[2] M.S. Bazaraa, J. Sherali, and C.M. Shetty, *Nonlinear Programming. Theory and Algorithms*, New York, Chichester, Brisbane, Toronto, John Wiley and Sons, 1993.

[3] C.G. Broyden, "The convergence of a class double-rank minimization algorithms." *Journal of the Institute of Mathematics and its Applications*, vol. 6, pp. 76-90, 1970.

[4] R.H. Byrd, J. Nocedal, and Y. Yuan, "Global Convergence of a Class of Variable Metric Algorithms." *SIAM Journal on Numerical Analysis*, vol. 24, pp. 1171-1190, 1987.

[5] W.C. Davidon, "Variable metric method for minimization." *AEC Res. and Dev. Report*, **ANL-5990** (revised) (1959).

[6] W.C. Davidon, "Variable metric method for minimization." *SIAM J. on Optimization*, vol. 1, pp. 1-17, 1991.

[7] R. Fletcher, "A rapid convergent descent method for minimization." *Computer J.*, vol. 6, pp. 163-168, 1963.

[8] R. Fletcher, and M.J.D. Powell, "A new approach to variable metric algorithms." *Computer J.*, vol. 13, pp. 317-322, 1970.

[9] R. Fletcher, *Practical Methods of Optimization*, second edition, Chichester, John Wiley & Sons, 1987.

[10] D. Goldfarb, "A family of variable metric methods derived by variational means." *Mathematics of Computation*, vol. 23, pp. 23-26, 1970.

[11] M. Lechman and A. Stachurski, "Nonlinear Section Model for Analysis of RC Circular Tower Structures Weakened by Openings." *Structural Engineering and Mechanics*, vol. 20, pp. 161-172, 2005.

[12] Z. Nowak, and A. Stachurski, "Nonlinear Regression Problem of Material Functions Identification for Porous Media Plastic Flow." *Engineering Transactions* , vol. 49, pp. 637-661, 2001.

[13] Z. Nowak, and A. Stachurski, "Global Optimization in Material Functions Identification for Voided Media Plastic Flow." *Computer Assited Mechanics and Engineering Sciences* , vol. 9, pp. 205-221, 2002.

[14] Z. Nowak, and A. Stachurski, "Identification of an Augmented Gurson Model Parameters for Plastic Porous Media." *Foundations of Civil and Environmental Engineering* (Publishing House of Pozna University of Technology), No. 2, pp. 171-179, 2002.

[15] Z. Nowak, and A. Stachurski, "Modelling and identification of voids nucleation and growth effects in porous media plastic flow." *Control and Cybernetics* , vol. 32, pp. 820-849, 2003.

[16] Z. Nowak, and A. Stachurski, "Robust Identification of an Augmented Gurson Model for Elasto-plastic Porous Media." *Archives of Mechanics (Archiwum Mechaniki Stosowanej)*, vol. 2, pp. 125-154, 2006.

[17] D.F. Shanno, "Conditioning of quasi-Newton methods for function minimization." *Mathematics of Computation*, vol. 24, pp. 27-30, 1970.

[18] A. Stachurski, "Superlinear Convergence of Broyden's Bounded Θ-Class of Methods." *Mathematical Programming*, vol. 20, pp. 196-212, 1981.

[19] A. Stachurski, and A.P. Wierzbicki, *Introduction to Optimization*, (Podstawy Optymalizacji, in polish), Warszawa, Publishing House of the Warsaw University of Technology, 1999.

[20] A. Stachurski, and M. Lechman, "On Solving a Set of Nonlinear Equations for the Determination of Stresses in RC Ring Sections with Openings." *Communications in Applied Analysis*, vol. 10, pp. 517-536, 2006.

[21] A. Stachurski, "Orthogonal Projections in the Quasi-Newton Variable Metric Updates", paper presented at the International Conference on Modelling and Optimization of Structures, Processes and Systems, held in Durban, 22-24 January 2007, (to appear in IMACS Journal of Mathematics and Computers in Simulation).

[22] A. Stachurski, "On the Structure of Variable Metric Updates" *International Journal of Pure and Applied Mathematics*, vol. 4, pp. 469-476, 2009.

[23] J. Stoer, "On the Convergence Rate of Imperfect Minimization Algorithms in Broyden's β-class." *Mathematical Programming*, vol. 9, pp. 313-335, 1975.

[24] D.B. Szyld, "The many proofs of an identity on the norm of oblique projections." *Numerical Algorithms*, vol. 42, pp. 309323, 2006.

[25] Sun Wenyu, and Yuan Ya-Xiang, *Optimization Theory and Methods. Nonlinear Programming*, Berlin, Springer Verlag, 2006.

[26] Yunhai Xiao, Zengxin Wei, and Zhiguo Wang, "A limited memory BFGS-type method for large-scale unconstrained optimization." *Computers and Mathematics with Applications*, vol. 56, pp. 10011009, 2008.