# PRACE NAUKOWE
Uniwersytetu Ekonomicznego we Wrocławiu

# RESEARCH PAPERS
of Wrocław University of Economics

Nr 446


## Metody i zastosowania badań operacyjnych


Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2016

Redakcja wydawnicza: Joanna Świrska-Korłub

Redakcja techniczna: Barbara Łopusiewicz

Korekta: Barbara Cibis

Łamanie: Małgorzata Myszkowska

Projekt okładki: Beata Dębska

Informacje o naborze artykułów i zasadach recenzowania
znajdują się na stronach internetowych
www.pracenaukowe.ue.wroc.pl
www.wydawnictwo.ue.wroc.pl



**ISSN 1899-3192**
**e-ISSN 2392-0041**

**ISBN 978-83-7695-610-7**

Wersja pierwotna: publikacja drukowana
Zamówienia na opublikowane prace należy składać na adres:
Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
ul. Komandorska 118/120, 53-345 Wrocław
tel./fax 71 36 80 602; e-mail:econbook@ue.wroc.pl
www.ksiegarnia.ue.wroc.pl

Druk i oprawa: TOTEM

## Wstęp

# Wstęp

Kolejna, XXXIV Ogólnopolska Konferencja Naukowa im. Profesora Władysława Bukietyńskiego, organizowana corocznie przez najważniejsze ośrodki naukowe zajmujące się dziedziną badań operacyjnych, w roku 2015 odbyła się w pięknym, zabytkowym i świeżo odremontowanym zespole pałacowo-parkowym w Łagowie koło Zgorzelca. Konferencję zrealizowaną pod nazwą *Metody i Zastosowania Badań Operacyjnych* przygotowała Katedra Badań Operacyjnych Uniwersytetu Ekonomicznego we Wrocławiu pod kierownictwem dr. hab. Marka Nowińskiego, prof. UE.

Konferencje te mają już długoletnią tradycję – są to coroczne spotkania pracowników nauki specjalizujących się w badaniach operacyjnych. Głównym celem konferencji było, podobnie jak w latach ubiegłych, stworzenie (przede wszystkim dla młodych teoretyków, a także praktyków dyscypliny) forum wymiany myśli na temat najnowszych osiągnięć dotyczących metod ilościowych wykorzystywanych do wspomagania procesów podejmowania decyzji, a także prezentacja nowoczesnych zastosowań badań operacyjnych w różnych dziedzinach gospodarki. Ten cenny dorobek naukowy nie może być zapomniany i jest publikowany po konferencji w postaci przygotowywanego przez organizatorów zeszytu naukowego zawierającego najlepsze referaty na niej zaprezentowane.

W pracach Komitetu Naukowego Konferencji uczestniczyli czołowi przedstawiciele środowisk naukowych z dziedziny badań operacyjnych w Polsce; byli to: prof. Jan B. Gajda (Uniwersytet Łódzki), prof. Stefan Grzesiak (Uniwersytet Szczeciński), prof. Bogumił Kamiński (SGH w Warszawie), prof. Ewa Konarzewska-Gubała (Uniwersytet Ekonomiczny we Wrocławiu), prof. Donata Kopańska-Bródka, prof. Maciej Nowak i prof. Tadeusz Trzaskalik (Uniwersytet Ekonomiczny w Katowicach), prof. Dorota Kuchta (Politechnika Wrocławska), prof. Krzysztof Piasecki (Uniwersytet w Poznaniu) i prof. Józef Stawicki (Uniwersytet Mikołaja Kopernika w Toruniu).

Zakres tematyczny konferencji obejmował teoretyczne i praktyczne zagadnienia dotyczące przede wszystkim:
- modelowania i optymalizacji procesów gospodarczych,
- metod wspomagających proces negocjacji,
- metod oceny efektywności i ryzyka na rynku kapitałowym i ubezpieczeniowym,
- metod ilościowych w transporcie i zarządzaniu zapasami,
- metod wielokryterialnych,
- optymalizacji w zarządzaniu projektami oraz analizy ryzyka decyzyjnego.

W konferencji wzięło udział 43 przedstawicieli różnych środowisk naukowych, licznie reprezentujących krajowe ośrodki akademickie. W trakcie sześciu sesji plenarnych, w tym dwóch sesji równoległych, przedstawiono 27 referatów, których poziom naukowy w przeważającej części był bardzo wysoki. Zaprezentowane referaty, po pozytywnych recenzjach, zostają dziś opublikowane w Pracach Naukowych Uniwersytetu Ekonomicznego we Wrocławiu w postaci artykułów naukowych w specjalnie wydanym zeszycie konferencyjnym.

Przypominając przebieg konferencji, nie można nie wspomnieć o konkursie zorganizowanym dla autorów referatów niebędących samodzielnymi pracownikami nauki. Dotyczył on prezentacji najciekawszego zastosowania badań operacyjnych w praktyce gospodarczej. Komitet Organizacyjny Konferencji powołał kapitułę konkursu, w której skład weszli: prof. Ewa Konarzewska-Gubała – przewodnicząca, prof. Jan Gajda, prof. Stefan Grzesiak i prof. Donata Kopańska-Bródka. Członkowie Komisji Konkursowej oceniali referaty ze względu na:

- innowacyjność, oryginalność metody będącej przedmiotem zastosowania,
- znaczenie zastosowania dla proponowanego obszaru,
- stopień zaawansowania implementacji metody w praktyce.

Spośród 15 referatów zgłoszonych wyróżniono: 1. miejsce: dr Michał Jakubiak i dr hab. Paweł Hanczar (Uniwersytet Ekonomiczny we Wrocławiu), *Optymalizacja tras zbiórki odpadów komunalnych na przykładzie MPO Kraków*; 2. miejsce: mgr Dagmara Piesiewicz i dr hab. Paweł Hanczar (Uniwersytet Ekonomiczny we Wrocławiu), *Logistyka odzysku – optymalizacja przepływów w systemie gospodarki komunalnej*; 3. miejsce: dr Dorota Górecka i dr Małgorzata Szałucka (Uniwersytet Mikołaja Kopernika w Toruniu), *Wybór sposobu wejścia na rynek zagraniczny – wieloaktorska analiza wielokryterialna a podejście oparte na dominacjach stochastycznych*.

Przy okazji prezentowania opracowania poświęconego XXXIV Konferencji *Metody i Zastosowania Badań Operacyjnych* i jej bardzo wartościowego dorobku nie możemy nie podziękować członkom Komitetu Organizacyjnego Konferencji, w którego skład wchodzili młodzi, acz doświadczeni pracownicy Katedry Badań Operacyjnych Uniwersytetu Ekonomicznego we Wrocławiu: dr Piotr Peternek (sekretarz), dr hab. Marek Kośny, dr Grzegorz Tarczyński oraz mgr Monika Stańczyk (biuro konferencji). Zapewnili oni w sposób profesjonalny sprawne przygotowanie i przeprowadzenie całego przedsięwzięcia oraz zadbali o sprawy administracyjne związane z realizacją konferencji, a także byli odpowiedzialni za dopilnowanie procesu gromadzenia i redakcji naukowych materiałów pokonferencyjnych, które mamy okazję Państwu dziś udostępnić.

Już dzisiaj cieszymy się na nasze kolejne spotkanie w ramach jubileuszowej XXXV Ogólnopolskiej Konferencji Naukowej im. Profesora Władysława Bukietyńskiego, która tym razem będzie organizowana przez naszych przyjaciół z Katedry Badań Operacyjnych Uniwersytetu Ekonomicznego w Poznaniu pod kierownictwem prof. dr. hab. Krzysztofa Piaseckiego.

*Marek Nowiński*

**PRACE NAUKOWE UNIWERSYTETU EKONOMICZNEGO WE WROCŁAWIU**
**RESEARCH PAPERS OF WROCŁAW UNIVERSITY OF ECONOMICS** **nr 446 • 2016**

Metody i zastosowania badań operacyjnych ISSN 1899-3192
e-ISSN 2392-0041

**Marek Nowiński**
Uniwersytet Ekonomiczny we Wrocławiu
e-mail: marek.nowinski@ue.wroc.pl

# TESTOWANIE NIELINIOWYCH ALGORYTMÓW OPTYMALIZACYJNYCH – ZESTAW FUNKCJI TYPU *BENCHMARK*

# TESTING NONLINEAR OPTIMIZATION ALGORITHMS – SET OF BENCHMARK TYPE FUNCTIONS

DOI: 10.15611/pn.2016.446.12
JEL Classification: C61

**Streszczenie:** Praca przedstawia specjalnie dobrany zestaw klasycznych funkcji testujących jakość działania nowo tworzonych lub modyfikowanych algorytmów optymalizacyjnych. Opisywane testy typu benchmark umożliwiają porównywanie efektywności różnych algorytmów minimalizujących funkcje, które są sumami kwadratów funkcji nieliniowych. Ocena ta dotyczy poszukiwania ekstremów lokalnych lub globalnych dla prostych (pod względem liczby zmiennych i składowych funkcji), ale sprawiających duże trudności w procesie znajdowania rozwiązań optymalnych przez swoją silną nieliniowość, wielomodalność, bardzo skomplikowany układ warstwic, brak różniczkowalności itp. Przydatność przedstawionych testów zbadano na przykładzie zaproponowanej przez autora modyfikacji jego własnego algorytmu optymalizacyjnego dla metod najmniejszych kwadratów, który nie wykorzystuje pochodnych funkcji w procesie obliczeniowym. Jest to udoskonalona, hybrydowa wersja metody Marquardta-Levenberga, zapewniająca nieosobliwość macierzy numerycznych oszacowań pochodnych, połączona ze zmodyfikowaną metodą uaktualniającą pierwszego rzędu Broydena do ustalania kierunku i długości kroku poszukiwania. Praca zawiera zestawienie i ocenę uzyskanych wyników numerycznych, które potwierdzają użyteczność zaproponowanego zestawu testów oraz wstępnie pozytywnie weryfikują efektywność zmodyfikowanego algorytmu optymalizacyjnego.

**Słowa kluczowe:** algorytm optymalizacyjny, nieliniowa metoda najmniejszych kwadratów, testy typu benchmark.

**Summary:** The paper presents a specially selected set of classic test functions evaluating the performance of newly created or modified optimization algorithms. Described benchmark tests can be used to compare the effectiveness of different algorithms minimizing the sums of squares of nonlinear functions. The assessment relates to the effectiveness of methods searching for the local or global extremes of simple tests (in terms of number of variables and sub-functions), but causing great difficulties in the process of finding optimal solutions due to their strong nonlinearity, multimodality, very complicated system of contour lines or non-differentiability. Usefulness of this tests set was examined for the modification of the optimization algorithm for least squares methods (proposed by the author), not using any derivatives of functions through a calculation. This is an improved hybrid version of the Levenberg-Marquardt method (ensuring non-singularity of the matrix of numerical estimates of derivatives) combined with modified Broyden's first row updating method for setting the direction and length of searching steps. The paper also includes a summary of the numerical results and its evaluation, which proves the usefulness of the dedicated set of tests and pre-validates the effectiveness of the proposed optimization algorithm.

**Keywords:** optimization algorithm, nonlinear least squares method, benchmarks for algorithm testing.

# 1. Wstęp

Artykuł poświęcono testowaniu algorytmów minimalizujących funkcje, które są sumami kwadratów funkcji nieliniowych. Metody uwzględniające specyfikę tych funkcji podczas procesu minimalizacji nazywane są skrótowo metodami nieliniowych najmniejszych kwadratów i stanowią od wielu lat obiekt szczególnego zainteresowania badaczy. Jest to uzasadnione z wielu powodów. Zagadnienia nieliniowych najmniejszych kwadratów pojawiają się we współczesnej nauce bardzo często i w wielu jej dziedzinach. Matematycy formułują ten problem jako znajdowanie punktu w danej podprzestrzeni, który jest najbliższy pewnemu punktowi w przestrzeni funkcji. Statystycy wprowadzają rozkłady prawdopodobieństwa do swojej koncepcji tego zagadnienia i nazywają je analizą regresji nieliniowej. Inżynierowie, fizycy i chemicy spotykają się z tym zagadnieniem przy podejmowaniu takich tematów, jak: określanie współrzędnych trajektorii lotu, filtracja lub opracowywanie danych eksperymentalnych. Ekonometrycy na co dzień mają do czynienia z metodami nieliniowych najmniejszych kwadratów przy estymacji parametrów nieliniowych modeli ekonometrycznych oraz dopasowywaniu krzywych. W różnych dziedzinach nauk ścisłych można zetknąć się z układami równań różniczkowych oraz ponadokreślonymi układami równań nieliniowych, które mogą być rozwiązywane przy użyciu tych metod.

Z tych powodów zachodzi potrzeba jednolitej oceny przydatności dotychczas znanych metod nieliniowych najmniejszych kwadratów oraz stworzenia na ich bazie prostego i skutecznego algorytmu, który mógłby być stosowany we wszystkich opisanych przypadkach. Widzimy też, że konieczność rozwijania tego rodzaju metod oraz badania ich efektywności jest oczywista. Wynika to z obserwacji, że procedury wykorzystujące charakterystyczną postać minimalizowanej funkcji są, w ogromnej większości przypadków, znacznie bardziej efektywne od ogólnych algorytmów optymalizacyjnych. Metody te są zawsze procedurami iteracyjnymi, polepszającymi na każdym etapie wstępne oszacowanie rozwiązania optymalnego, co umożliwia znajdowanie rozwiązań praktycznych zagadnień optymalizacyjnych, w większości niemożliwych do rozwiązania w sposób analityczny.

## 2. Algorytm optymalizacyjny

W niniejszym artykule zasadniczą uwagę zwrócono na algorytmy optymalizacyjne dla metod najmniejszych kwadratów, które nie wykorzystują pochodnych funkcji w procesie obliczeniowym. Jest to cecha istotnie zwiększająca zakres ich stosowania. W tym przypadku odpada bowiem konieczność założenia różniczkowalności funkcji będącej obiektem minimalizacji. W praktycznych zagadnieniach optymalizacyjnych większość funkcji jest tak skomplikowana, że uniemożliwia to lub w bardzo znacznym stopniu utrudnia zaprogramowanie formuł dla obliczania przybliżeń pochodnych w praktycznych realizacjach algorytmów obliczeniowych. Zawsze jednak wymagają one wykorzystania olbrzymich liczby operacji arytmetycznych i dużej pojemności pamięci maszynowej, nawet przy nieskomplikowanych i niezbyt rozbudowanych zagadnieniach optymalizacyjnych.

Proponowany przez nas algorytm jest udoskonaloną wersją metody Marquardta-Levenberga, służącą do rozwiązywania zagadnień minimalizacyjnych dla sum kwadratów funkcji nieliniowych bez użycia pochodnych[1]. Problem ten możemy zapisać najbardziej ogólnie w postaci:

Niech funkcja $f: R^n \to R^m$ , gdzie $m \geq n$, będzie danym wektorem funkcji $f_i \in R$. Szukamy punktu $x^* = [x_1^*, \dots, x_n^*]$, dla którego funkcja, kryterium będąca sumą kwadratów funkcji $f_i$ ,

$$F(x) = \|f(x)\|^2 = \sum_{i=1}^{m} f_i^2\,(x) \tag{1}$$

osiąga przynajmniej minimum lokalne, a co najmniej jedna z funkcji $f_i$ jest funkcją nieliniową.

Proponowany algorytm na charakter hybrydowy, więc niemożliwe będzie udowodnienie jego zbieżności nawet dla dość ograniczonych klas funkcji. Dowód zbieżności został zastąpiony szeroką gamą badań empirycznych potwierdzających praktyczną przydatność i efektywność algorytmu.

Algorytm skonstruowany został z użyciem techniki siecznej wielu zmiennych (zmodyfikowana metoda uaktualniająca pierwszego rzędu Broydena [Kvaalen 1991]) do aproksymacji jakobianu funkcji celu. Jakobian ten następnie używany jest do otrzymania iteracji typu Marquardta-Levenberga, zapewniającej nieosobliwość macierzy w całym procesie obliczeniowym. Jest on zachowywany w postaci sfaktoryzowanej i uaktualniany podczas każdej iteracji za pomocą techniki dekompozycji ortogonalnych typu QR, a występujące cyklicznie liniowe podzagadnienia najmniejszych kwadratów rozwiązywane są za pomocą metody Goluba [Golub, van Loan 1996]. W procesie obliczeniowym zastosowano również metody zabezpieczające przed niezbieżnością algorytmu w przypadku szczególnie złego uwarunkowania zadania. Są to mianowicie tzw. kroki uboczne, zapobiegające liniowej zależności kolejnych kierunków poszukiwania generowanych przez algorytm, oraz procedura

[1] Opracowano na podstawie [Marquardt 1963] oraz wcześniejszych publikacji autora [Nowiński 1985; 1986; 2007].

restartów, stosowana w przypadku niemożności zmniejszenia wartości funkcji celu na danym kierunku poszukiwania.

W pracy ograniczono się jedynie do krótkiego opisu algorytmu, gdyż służy on tylko do zaprezentowania zestawu funkcji testowych (*benchmark*) oceniających efektywność algorytmów minimalizacyjnych dla całej klasy nieliniowych zagadnień bez ograniczeń, czyli tzw. optymalizacji bezwarunkowej. Należy podkreślić, że kierunek poszukiwania $p^k$ w klasycznej metodzie Gaussa-Newtona dany jest wzorem:

$$p^k = -[J(x^k)^T]J(x^k)]^{-1} J(x^k)^T f(x^k), \tag{2}$$

zaś w zmodyfikowanej, w celu zapewnienia odwracalności macierzy, wersji Marquardta i Levenberga:

$$p^k = -\left[J(x^k)^T\right]J(x^k) + \mu^k I]^{-1} J(x^k)^T f(x^k), \tag{3}$$

gdzie wszystkie $\mu^k$ muszą być nieujemne, a symbol $J$ oznacza macierz jakobianu funkcji wektorowej $f$, czyli macierz pierwszych pochodnych cząstkowych tej funkcji $J_{ij} = \{\partial f_i(x)/\partial x_j\}$. W wypadku równania (3) nieosobliwość macierzy w nawiasach kwadratowych jest oczywiście zapewniona, gdy tylko $\mu^k \neq 0$.

Można zauważyć, że równanie (3) reprezentuje równanie normalne dla liniowego zagadnienia najmniejszych kwadratów (gdzie symbol $\cong$ oznacza rozwiązanie danego zagadnienia w sensie najmniejszych kwadratów)

$$\begin{bmatrix} J(x^k) \\ --- \\ (\mu^k)^{1/2} I \end{bmatrix} p^k \cong -\begin{bmatrix} f(x^k) \\ --- \\ 0 \end{bmatrix} \tag{4}$$

ze względu na $p^k$. Zagadnienie to ma rozwiązanie niezależnie od rzędu występujących w nim macierzy, chociaż rozwiązanie to jest jednoznaczne tylko wtedy, gdy rząd jest pełny. Wektor $p^k$ minimalizujący zagadnienie (4) można na przykład uzyskać dzięki metodzie Goluba, wykorzystującej elementarne transformacje ortogonalne [Golub, van Loan 1996]. Stosując proste podstawienia i opuszczając numer iteracji (zastępując rzeczywisty Jakobian przez jego aproksymację uzyskaną przy wykorzystaniu zmodyfikowanej metody Broydena, oznaczoną jako $\widetilde{M}$), możemy zapisać zależność (4) z uwzględnieniem wymiarów macierzy jako:

$$\widetilde{M}_{(m+n)xn} p_{nx1} \cong -f_{(m+n)x1}. \tag{5}$$

Wtedy dla każdej macierzy ortogonalnej $Q$ zachodzi oczywiście zależność

$$\|(\widetilde{M}p + f)\| = \|Q\widetilde{M}p + Qf\|. \tag{6}$$

Można skonstruować taką macierz ortogonalną $Q$ jako produkt innych macierzy ortogonalnych tak, by przetransformować wyrażenie $\widetilde{M}p + f$ do postaci:

$$\left[\frac{R}{0}\right] p \cong -\left[\frac{s}{t}\right], \tag{7}$$

w której $n \times n$ podmacierz $R$ jest górną macierzą trójkątną. Wektor $p$ musi wtedy być wybrany jako taki, który rozwiązuje liniowy układ trójkątny w sensie najmniejszych kwadratów:

$$Rp \cong -s, \tag{8}$$

co jest numerycznie znacznie łatwiejsze do uzyskania niż bezpośrednie rozwiązywanie zagadnienia (5).

# 3. Funkcje testowe

## 3.1. Właściwości funkcji typu *benchmark*

Ocena efektywności i mocy algorytmu przeznaczonego do rozwiązywania zagadnień programowania nieliniowego bez ograniczeń wymaga doboru dość szerokiej klasy różnorodnych problemów testowych oraz rozsądnych kryteriów porównawczych. Algorytmy mogą być badane z teoretycznego punktu widzenia oraz przez testy eksperymentalne. To pierwsze podejście można stosować niestety jedynie do wąskich klas zagadnień i nie do wszystkich algorytmów. Dlatego też w naszym przypadku zajmiemy się badaniem jakości algorytmu oraz porównywaniem z innymi znanymi metodami jedynie przez analizę oraz porównywanie uzyskiwanych wyników numerycznych. Algorytmy mogą być testowane na zagadnieniach o dużej lub małej liczbie zmiennych, różniących się między sobą stopniem nieliniowości i uwarunkowania macierzy charakterystycznej badanej metody. Można je badać na zagadnieniach czysto teoretycznych, zaproponowanych jako testy przez znanych badaczy, oraz na zagadnieniach powstałych przy rozwiązywaniu problemów praktycznych. Pierwsze z nich mają tę zaletę, że są specjalnie dobrane pod kątem badania wrażliwości algorytmu na konkretne i znane niedogodności zadania testowego, takie na przykład, jak skomplikowany układ warstwie funkcji, bardzo strome stoki powierzchni utworzonej przez zbiór wartości funkcji, osobliwości macierzy charakterystycznej metody w punkcie startowym lub optymalnym, duże wartości reszt w punkcie będącym rozwiązaniem itp. Umożliwiają one również szerokie porównania testowanego algorytmu z innymi metodami, gdyż autorzy w swych pracach bardzo często przedstawiają wyniki numeryczne uzyskane przez proponowane metody dla tego typu zagadnień numerycznych. Natomiast grupa druga, zawierająca między innymi problemy rozwiązywania większych układów równań nieliniowych oraz zagadnienia estymacyjne, sprawdza działanie algorytmu na konkretnych zagadnieniach praktycznych, co potwierdza faktyczną moc badanego algorytmu.

W naszym konkretnym przypadku stosowaliśmy nowy algorytm do całej klasy zagadnień pierwszego typu, co pozwala na dokonanie porównań z wynikami numerycznymi osiągniętymi przez inne algorytmy. Badając efektywność algorytmu przez rozwiązywanie całej gamy zagadnień testowych, mamy nadzieję przewidzieć ogólną efektywność tej metody przy problemach podobnego typu.

Aby problem testowy był użyteczny, powinien mieć jedno znane ekstremum lub co najwyżej niewielką, ograniczoną ich liczbę. Na aktualnym etapie rozwoju wiedzy nie można się jeszcze spodziewać, że algorytm znajdzie minimum globalne, gdy w badanym zagadnieniu występuje więcej niż jedno minimum, ale powinien on przynajmniej osiągnąć minimum lokalne, aby mógł być uważany za skuteczny. Jednakże jeśli rozpatruje się zagadnienie, w którym w punkcie $x^*$ znajduje się minimum lokalne i oprócz tego funkcja celu ma minimum globalne w $-\infty$, to interesujące jest pytanie, czy jeśli konkretna technika minimalizacyjna spowoduje zbieżność do tego minimum globalnego, to należy to traktować jako sukces czy jako porażkę danej metody.

Eksperymentalne porównywanie algorytmów zależy w znacznym stopniu od sposobu ich zaprogramowania. Drobne szczegóły w procesie programowania mogą często wywierać zasadniczy wpływ na jakość działania algorytmu. Małe zmiany kryteriów zakończenia obliczeń, technik poszukiwania liniowego, testów na osobliwość macierzy i procedur ich odwracania mogą powodować duże zmiany efektywności algorytmów, co istotnie utrudnia ich obiektywne porównywanie.

Generalnie w literaturze akceptuje się pogląd, że podstawowym kryterium oceny algorytmów minimalizacyjnych dla zagadnień bez ograniczeń jest to, czy algorytm potrafi rozwiązać większość postawionych zagadnień testowych, czyli moc badanego algorytmu. Oczywiście każda metoda może zawieść przy specjalnie skonstruowanym (patologicznym) problemie, ale nawet dla zagadnień innych niż patologiczne nie jest celowe pytanie o to, czy algorytm może rozwiązać wszystkie możliwe zagadnienia. Łatwo jest bowiem znaleźć zagadnienia programowania nieliniowego bez ograniczeń, które prowadzą do uzyskania ujemnych argumentów, nieciągłości, dzielenia przez zero i innych podobnych trudności. Ale moc algorytmu jest jednak zawsze sprawą najważniejszą.

Drugim kryterium jest osiągnięcie pożądanego poziomu precyzji rozwiązania, tzn. wartości funkcji celu w punkcie ekstremalnym oraz składowych wektora zmiennych w tym punkcie. Zazwyczaj stopień dokładności rozwiązania zależy od użytych kryteriów zakończenia obliczeń. Aby zapewnić jednolitość tych kryteriów, przyjęto, by ta sama względna precyzja wektora optymalnego $x^*$ lub wartości funkcji celu $F(x)$ była wspólną bazą dla zakończenia poszukiwania w każdym z przedstawionych algorytmów. Zakładając, że ustalono jednolite kryteria zakończenia obliczeń, przyjęto jednocześnie, że następnym miernikiem oceny jakości algorytmu będzie jego efektywność, mierzona za pomocą liczby obliczeń wartości funkcji celu (tzw. oszacowań funkcji) potrzebnych do osiągnięcia pożądanej. Oczywiście kryterium liczby oszacowań funkcji (a w przypadku porównywania metod niewykorzystujących pochodnych funkcji tzw. ekwiwalentnych oszacowań funkcji (*leof*) , kiedy jedno oszacowanie funkcji plus $n$-numerycznych oszacowań pochodnych cząstkowych liczone jest jako $(n + 1)$-ekwiwalentnych oszacowań wartości funkcji) jest znacznie lepsze od porównywania liczby cykli iteracyjnych, gdyż w wielu algorytmach cykl taki znaczy zupełnie coś innego, niż tylko obliczenie nowego kierunku poszukiwania, i zawiera bardzo różną liczbę oszacowań funkcji.

Nie jest to jednakże kryterium doskonałe, gdyż pomija ono zagadnienie wkładu pracy obliczeniowej niezwiązanej z obliczaniem wartości funkcji, a która różni się znacznie w poszczególnych metodach. Poza tym można starać się specjalnie zredukować liczbę koniecznych oszacowań, wartości funkcji przez różnego rodzaju czasochłonne testy, specjalne operacje heurystyczne, działania macierzowe itp. Jednak w naszych warunkach wykorzystanie liczby ekwiwalentnych oszacowań funkcji do ocen i porównań algorytmów uważamy za rozwiązanie najlepsze. Czas obliczeń algorytmu nie znajduje zazwyczaj zastosowania, gdyż porównanie czasów wymaga używania tego samego rodzaju sprzętu obliczeniowego do wykonywania testów numerycznych, a twórcy nowych algorytmów korzystają przede wszystkim z danych zamieszczanych w pracach różnych autorów, co oczywiście wiąże się z użytkowaniem przez nich różnorodnego sprzętu. Dlatego też czas obliczeń może być umieszczony jedynie jako porównanie pracochłonności poszczególnych zagadnień testowych dla nowego algorytmu.

## 3.2. Przykłady funkcji testowych i wyniki numeryczne

W tym podpunkcie zaprezentowano kilka zagadnień testowych używanych do badania nowego algorytmu. Będą one opisane w sposób standardowy: najpierw numer zagadnienia testowego (oznaczony liczbą rzymską) i jego nazwa oraz źródło pochodzenia. W następnej kolejności oznaczono literami:

(a) definicja funkcji testowej $F(x)$ typu (1) i jej krótka charakterystyka,

(b) wymiary zagadnienia, gdzie $m$ to liczba podfunkcji, a $n$ to liczba zmiennych,

(c) punkt startowy $x^0$ , wartość funkcji celu w tym punkcie $F^0$,

(d) punkt minimalny (lub punkty) $x^*$ oraz wartość minimalna funkcji $F^*$,

(e) wyniki numeryczne, gdzie użyto skrótów: *leof* − liczba ekwiwalentnych oszacowań funkcji, $F$ – bieżąca wartość funkcji testowej, zaś gwiazdką (*) oznaczono cykl, w którym spełniony został po raz pierwszy warunek $F(x) - F^* < 10^{-5}$,

(f) wartości składowych wektora zmiennych $x$ odpowiadające uzyskanej, minimalnej sumie kwadratów,

(g) wartość średniego tempa zbieżności, czyli pewnego kryterium efektywności algorytmu zaproponowanego przez Broydena [Kvaalen 1991]

$$R = 1/z \ln [F^0 / ( F^z - F^*),$$

gdzie $F^0$ to początkowa wartość funkcji celu w punkcie startowym, $F^*$ to znana wartość minimalna, a $z$ to całkowita liczba oszacowań funkcji celu koniecznych do uzyskania $F^z$, czyli najmniejszej wartości tej funkcji możliwej do osiągnięcia za pomocą badanej metody.

Drugi z tych mierników ma postać:

$$R^{*} = 1/r \ln [F^{0} / ( F^{r} - F^{*}),$$

gdzie $r$ to liczba oszacowań funkcji celu niezbędnych do zredukowania wyrażenia:

$$F(x) - F^{*}$$

do wartości < $10^{-5}$, $F^{r}$ to pierwsza wartość funkcji spełniająca ton warunek. Wartość ta odgrywa ważną rolę, gdyż w zależności od końcowego tempa zbieżności może ona być bardzo różna. Wartości tych mierników zawarte są w przedziale (, + ∞), a rezultaty większe od 0,4 uważane są za wyniki bardzo dobre.

W dalszej części tekstu zaprezentowano opis zagadnień testowych wykorzystywanych do badań numerycznych oraz wyniki stosowania nowego algorytmu do tych zagadnień.

**I − funkcja Zangwilla [Zangwill 1967]:**

(a) $F(x) = (x_1 - x_2 + x_3)^2 + (-x_1 + x_2 + x_3)^2 + (x_1 + x_2 - x_3)^2$;

prosty układ równań liniowych rozwiązywany za pomocą redukcji kwadratów reszt;

(b) $m = 3$, $n = 3$;

(c) $x^0 = [100; -1; 2{,}5]^T$, $F^0 = 29\,726{,}75$;

(d) $x^* = [0,0,0]^T$ $F^* = 0$;

(e) wyniki numeryczne:

| Numer cyklu | *leof* | *F* |
|---|---|---|
| 1 | 1 | 29 726,750 |
| … | … | … |
| 12 (*) | 53 | 0,170e-07 |
| 13 | 39 | 0,717e-08 |
| 14 | 42 | 0,611e-16 |
| 15 | 45 | 0,248e-23 |

(f) $x = [0{,}909e - 12; 0; 0]$;

(g) $R = 1.437$ $R^* = 0{,}805$.

**II − funkcja Himmelblau nr 28** [Himmelblau 1972]:

(a) $F(x) = (x_1{}^2 + x_2 - 11)^2 + (x_1 + x_2{}^2 - 7)^2$;

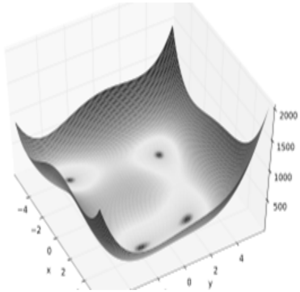
prosty układ równań nieliniowych z czterema minimami lokalnymi, które stanowią jednocześnie minimum globalne; funkcja ta ma również jedno maksimum globalne;

(b) $m = 2$, $n = 2$;

(c) $x^0 = [1;1]^T$, $F^0 = 106$;

(d) $x^* = [3{,}58443, -1{,}84813]^T$ lub $x^* = [3;2]^T$, lub

$x^* = [3{,}77931, -3{,}283186]^T$, lub $x^* = [-2{,}805118; 3{,}131312]^T$ $F^* = 0$




(e) wyniki numeryczne:

| Numer cyklu | *leof* | *F* |
|---|---|---|
| 1 | 1 | 106 |
| … | … | … |
| 4 | 17 | 0,597e-01 |
| 5 | 21 | 0,518e-02 |
| 6(*) | 24 | 0,941e-07 |
| 7 | 39 | 0,940e-07 |

(f) $x = [3{,}0; 1{,}9999256]$;
(g) $R = 0{,}534$ $\quad R^* = 0{,}868$.

**III − funkcja Himmelblaua nr 30** [Himmelblau1972]:

(a) $F(x) = (10[x_3 - (\frac{x_1+x_2}{2})^2]^2 + (1-x_1)^2 + (1-x_2)^2$;
(b) $m = 2$, $n = 3$;
(c) $x^0 = [-1, 2; 2; 0]^T$ , $F^0 = 8{,}40$;
(d) $x^* = [1; 1; 1]^T$ $\quad F^* = 0$;
(e) wyniki numeryczne:

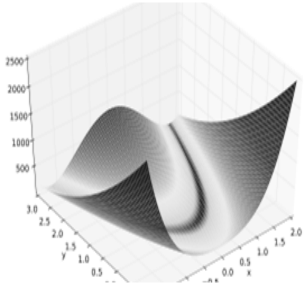
| Numer cyklu | *leof* | *F* |
|---|---|---|
| 1 | 1 | 8,40 |
| … | … | … |
| 13 | 52 | 0,756e-03 |
| 15(*) | 60 | 0,163e-04 |
| … | … | … |
| 19 | 72 | 0,162e-08 |

(f) $x = [1{,}000035; 1.000017; 1{,}000052]$;

(g) $R = 0{,}257$ $\quad R^* = 0{,}295$.

**IV − funkcja Rosenbrocka** [Rosenbrock 1960]:

(a) $F(x) = [(10[x_2 - {x_1}^2])^2]^2 + (1-x_1)^2$ − łukowata dolina o bardzo stromych stokach, co bardzo utrudnia minimalizację;



(b) $m = 2$, $n = 2$;
(c) $x^0 = [-l,2;1,0]^T$ , $F^0 = 24,20$;
(d) $x^* = [1;1]^T$ $F^* = 0$;
(e) wyniki numeryczne:

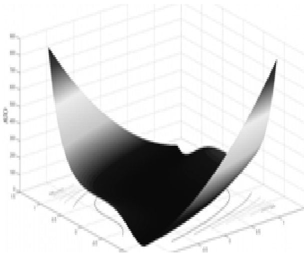
| Numer cyklu | *leof* | *F* |
|---|---|---|
| 1 | 1 | 24,20 |
| … | … | … |
| 20 | 61 | 0,243 |
| … | … | … |
| 27(*) | 64 | 0,176e-09 |
| … | … | … |
| 29 | 96 | 0 |

(f) $x = [1,0; 1,0]^T$ ;

(g) $R = +\infty$ $R^* = 0,195.$

**V − sześcienna, multimodalna funkcja Leona** [Himmelblau1972]:

(a) $F(x) = (10[x_2 - {x_1}^3])^2 + (1 - x_1)^2$ − *bardzo wąska dolina wzdłuż krzywej*



(b) $m = 2$, $n = 2$;
(c) $x^0 = [-l,2;1.0]^T$ , $F^0 = 749,04$;
(d) $x^* = [1;1]^T$ $F^* = 0$;
(e) wyniki numeryczne:

| Numer cyklu | *leof* | *F* |
|---|---|---|
| 1 | 1 | 749,04 |
| … | … | … |
| 20 | 60 | 0,308 |
| … | … | … |
| 35(*) | 107 | 0,260e-07 |
| … | … | … |
| 42 | 131 | 0,355e-12 |

(f) $x = [1{,}0; 0{,}999999]^T$ ;

(g) $R = 0{,}211 \qquad R^* = 0{,}173.$

**VI − kombinacja funkcji Rosenbrocka i Leona** [Himmelblau 1972]:

(a) $F(x) = (10[x_2 - {x_1}^2])^2 + \left(\sqrt{2}(1 - x_1)\right)^2 + (10[x_2 - {x_1}^3])^2$ – *jedno minimum lokalne i jedno globalne;*

(b) $m = 3$, $n = 2$;

(c) $x^0 = [-3;0]^T$ , $F^0 = 81\,031{,}938$;

(d) globalne $x^* = [1;1]^T$ $F^* = 0$ ; lokalne $x^{**} = [0{,}35;0{,}081]^T$ $F^{**} = 1{,}161925$;

(e) wyniki numeryczne:

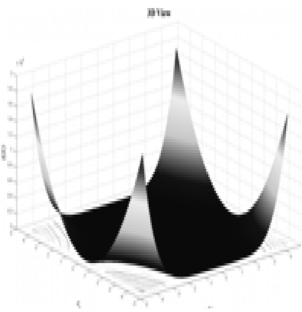
| Numer cyklu | *leof* | *F* |
|---|---|---|
| 1 | 1 | 81 031,938 |
| … | … | … |
| 11 | 42 | 1,250 |
| … | … | … |
| 15(*) | 59 | 1,1619270 |
| … | … | … |
| 17 | 84 | 1,1619264 |

(f) $x = [0{,}346954; 0{,}081069]^T$;

(g) $R = 0{,}299 \qquad R^* = 0{,}427.$

**VII − funkcja Beale'a** [Beale 1958]:

(a) $F(x) = [1{,}5 - x_1(1 - x_2)]^2 + [2{,}25 - x_1(1 - x_2^2)]^2 + [2{,}625 - x_1(1 - x_2^3)]^2$ – *zagadnienie pośrednie między klasycznym przykładem testowym a prostym problemem estymacji parametrów*;



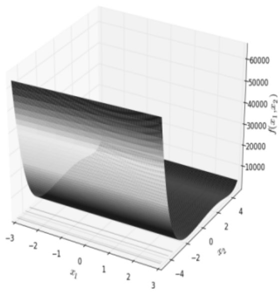
(b) $m = 3$, $n = 2$;
(c) $x^0 = [2;0;2]^T$, $F^0 = 0{,}52978$;
(d) $x^* = [3;\ 0.5]^T$, $F^* = 0$;
(e) wyniki numeryczne:

| Numer cyklu | *leof* | *F* |
|---|---|---|
| 1 | 1 | 0,52978 |
| … | … | … |
| 5 | 18 | 0,347e-01 |
| … | … | … |
| 8(*) | 29 | 0,139e-05 |
| … | … | … |
| 24 | 90 | 0,120e-05 |

(f) $x = [3{,}002739; 0{,}500689]^T$;
(g) $R = 0{,}124 \quad R^* = 0{,}443$.

**VIII − funkcja Freudensteina-Rotha** [Freudenstein, Roth 1963]:

(a) $F(x) = \left[x_1 + \left((5 - x_2)x_2 - 2\right)x_2 - 13\right]^2 + \left[x_1 + \left((x_2 + 1)x_2 - 14\right)x_2 - 29\right]^2$ − układ równań nieliniowych o bardzo skomplikowanym ułożeniu warstwic;




(b) $m = 2$, $n = 2$;
(c) $x^0 = [15; -2]^T$, $F^0 = 1256$;
(d) globalne $x^* = [5;4]^T$, $F^* = 0$; lokalne $x^{**} = [11{,}41\ldots; -0{,}08968\ldots]^T$, $F^{**} =$ 48,9842;
(e) wyniki numeryczne:

| Numer cyklu | *leof* | *F* |
|---|---|---|
| 1 | 1 | 1256,0 |
| … | … | … |
| 6 | 31 | 80,981 |
| … | … | … |
| 156 | 1268 | 48,984192 |

(f) $x = [11{,}429215; -0{,}89574832]^T$;
(g) wartości $R$ oraz $R^*$ nie mogą być wyliczone z powodu zbyt małej dokładności rozwiązań optymalnych podanych przez autora problemu testowego.

**IX − funkcja Powella** [Powell 1964]:

(a) $F(x) = [x_1 + 10x_2]^2 + \left[\sqrt{5}(x_2 - x_4)\right]^2 + [(x_2 - 2x_3)^2]^2 + \left[\sqrt{10}(x_1 - x_4)^2\right]^2$ – układ równań nieliniowych z podwójnie osobliwym jakobianem funkcji $F$ w punkcie optymalnym;

(b) $m = 4$, $n = 4$;

(c) $x^0 = [3; -1;0;1]^T$, $F^0 = 215$;

(d) minimum globalne $x^* = [0;0;0;0]^T$, $F^* = 0$;

(e) wyniki numeryczne – jest o jedyna funkcja testowa, dla której nie udało się uzyskać przynajmniej zadowalającej zbieżności do punktu optymalnego, mimo korzystania jedynie z numerycznych oszacowań pochodnych funkcji.

**X − funkcja Himmelblau nr 29** [Himmelblau1972]:

(a) $F(x) = ({x_1}^2 + 12x_2 - 1)^2 + (49x_1^2 + 49x_2^2 + 84x_1 + 2324x_2 - 68)^2$; *przykład skomplikowanego układu równań nieliniowych z minimum lokalnym i globalnym, dla którego macierz $J^TJ$ jest osobliwa w punkcie startowym* $[1;1]^T$;

(b) $m = 2$, $n = 2$;

(c) $x^0 = [l;1]^T$, $F^0 = 3{,}330768e + 06$;

(d) minimum globalne $x^* = [-21{,}026653; -36{,}760090]^T$, $F^* = 0$; lokalne $x^{**} = [0{,}28581;0{,}27936]^T$, $F^{**} = 5{,}9225$;

(e) wyniki numeryczne:

| Numer cyklu | *leof* | *F* |
|---|---|---|
| 1 | 1 | 3,330768e + 06 |
| … | … | … |
| 16 | 56 | 6,674 |
| … | … | … |
| 41 | 149 | 5,92503 |

(f) minimum lokalne $x = [0{,}28581; 0{,}27936]^T$;

(g) $R = 0.114$ , $R^*$ nie jest możliwe do obliczenia ze względu na niewystarczającą dokładność znalezionego rozwiązania $F(x) - F^* \geq 10^{-5}$ .

## 4. Zakończenie

Przeanalizowane przez autora przekłady prostych w zapisie funkcji testowych, które (chociaż niektóre dosyć wiekowe) stanowią dosyć skomplikowane problemy obliczeniowe, mogą być traktowane jako doskonały wstępny test jakości działania algorytmów minimalizujących sumy kwadratów funkcji nieliniowych. Są one również przydatne do badania efektywności innych, bardziej ogólnych metod optymalizacyjnych, gdyż zawierają zagadnienia testowe o niewielkiej liczbie zmiennych, ale powodujące duże trudności obliczeniowe dla wszystkich algorytmów. Spowodowane jest to występowaniem takich specyficznych własności badanych funkcji, jak: silna nieliniowość, wielomodalność, bardzo skomplikowany układ warstwic, osobliwość (lub nawet podwójna osobliwość) w punkcie startowym lub optymalnym, co eliminuje możliwość wykorzystania metod gradientowych oraz bardzo utrudnia numeryczną aproksymację pochodnych.

W tych trudnych warunkach zaproponowany algorytm optymalizacyjny (hybrydowy, bezgradientowy i minimalizujący wykorzystywaną liczbę oszacowań wartości funkcji) okazał się narzędziem bardzo efektywnym i odpornym na prawie wszystkie, opisane pułapki procesu poszukiwania ekstremum. Tylko funkcja Powella (silnie nieliniowa, z podwójnie osobliwym jakobianem funkcji w punkcie optymalnym) okazała się odporna na próby znalezienia jej minimum globalnego. Jednakże inne badania przeprowadzone przez autora[2] potwierdziły efektywność tej metody, a także przydatność przedstawionego w artykule zestawu funkcji testujących do wstępnego określania jakości działania nowotworzonych algorytmów optymalizacyjnych.

## Literatura

Beale E.M.L., 1958, *On an iterative method of finding a local minimum*, Tech. Report, no. 25, STRG, Princeton University.

Freudenstein F., Roth B., 1963, *Numerical solutions of systems of nonlinear equations*, ACM Journal, 10, s. 550-556.

Golub G.H.; van Loan C.F., 1996, *Matrix Computations*, Johns Hopkins University Press.

Himmelblau D.M., 1972, *Applied Nonlinear Programming*, Mc Graw-Hill, New York.

Kvaalen E., 1991, *A faster Broyden method*, BIT Numerical Mathematics (SIAM), 31(2), s. 369-372.

Marquardt D., 1963, *An algorithm for lest squares estimation for nonlinear parameters*, SIAM J. on Appl.Math., 11, s. 431-441.

Nowiński M., 1985, *Zagadnienie minimalizacji sumy kwadratów funkcji nieliniowych bez wykorzystania pochodnych*, rozprawa doktorska, AE, Wrocław.

Nowiński M., 1986, *Modyfikacja algorytmu Marquardta-Levenberga niewykorzystująca analitycznych ani numerycznych pochodnych funkcji*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu, nr 351, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław, s. 129-140.

Nowiński M., 2007, *Nieliniowa dynamika szeregów czasowych w badaniach ekonomicznych*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.

Powell M.J.D., 1964, *An efficient method for finding the minimum of a function of several variables without calculating derivatives*, Computer J., 5, s. 155-162.

Rosenbrock H.H., 1960, *An automatic method for finding the greatest and least value of the function*, Computer J., 3, s. 175-184.

Zangwill W.I., 1967, *Minimizing a function without calculating derivatives*, Computer J., 10, s. 293-296.

[2] Kilkadziesiąt funkcji testowych i praktycznych problemów estymacyjnych o dużych wymiarach (liczbie dopasowywanych punktów oraz zmiennych), które zostaną opisane w kolejnej publikacji autora.