

# Elementy teorii optymalizacji oraz metody rozwiązywania zadań optymalizacji nieliniowej

## Wprowadzenie

Wiele zagadnień z zakresu ekonomii można przedstawić w postaci zadań optymalizacyjnych. Przykładem mogą być problemy maksymalizacji dochodów w warunkach strategii krótko- i długookresowej [4]. W praktyce powszechnie stosowane są liniowe modele optymalizacyjne, rozwiązywane za pomocą algorytmu sympleks. Znacznie rzadziej natomiast wykorzystywane są modele nieliniowe. Wynika to nie tyle z mniejszej przydatności tych modeli, co z niedostatecznej ich znajomości. Modele nieliniowe są trudniejsze do konstruowania oraz rozwiązywania. W niniejszym artykule przedstawione zostaną pokrótce teoretyczne podstawy oraz algorytmy optymalizacji nieliniowej.

## Elementy teorii optymalizacji

Findeisen, Szymanowski i Wierzbicki [2] podają ogólną definicję optymalizacji: „Optymalizacja jest to postępowanie, polegające na wyborze elementu z danego zbioru w oparciu o relacje, ustalające pewien porządek w tym zbiorze”. Zbiór ten nazywany jest zbiorem rozwiązań dopuszczalnych. Porządek w zbiorze rozwiązań dopuszczalnych może ustalać funkcja rzeczywista, zwana wskaźnikiem jakości lub funkcją celu. Funkcja ta jest określona dla zmiennych zwanych decyzyjnymi. Zadanie optymalizacji polega na takim doborze zmiennych decyzyjnych, aby funkcja celu osiągała wartość maksymalną lub minimalną.

W literaturze teoria optymalizacji formułowana jest zwykle dla zadania minimalizacji i dlatego też tak zostanie przedstawiona w dalszej części tej pracy. Poszukiwanie maksimum funkcji  $f(X)$  zawsze może być sprowadzone do poszukiwania minimum, ponieważ maksymalizacja  $f(X)$  jest równoważna minimalizacji  $-f(X)$ .

Zadanie optymalizacji lub zadanie programowania matematycznego polega więc na poszukiwaniu:

$$\min f(\mathbf{X}) \quad (1)$$

$\mathbf{X} \in Z_R$ , gdzie:

$f$  – funkcja celu. Jest to funkcja  $n$ -zmiennych, przekształcająca  $n$ -wymiarową przestrzeń rzeczywistą  $R^n$  w zbiór liczb rzeczywistych  $R^1$ .

$\mathbf{X}$  – jest  $n$ -wymiarowym wektorem zmiennych decyzyjnych, czyli  $\mathbf{X} \in R^n$ .  $Z_R$  – jest zbiorem rozwiązań dopuszczalnych.

Jeżeli nie ma żadnych ograniczeń narzuconych na wybór rozwiązania (czyli, gdy  $Z_R = R^n$ ), mówimy o zadaniu programowania matematycznego bez ograniczeń zapisywanym jako:

$$\min f(\mathbf{X}) \quad (2)$$

$\mathbf{X} \in R^n$

Jeśli natomiast zmienne decyzyjne muszą spełniać dodatkowe kryteria, tzw. kryteria dopuszczalności (czyli, gdy  $Z_R \subset R^n$ ), zadanie optymalizacji nazywane jest zadaniem programowania matematycznego z ograniczeniami. Ograniczenia formułowane są w postaci równań lub nierówności. Znalezienie rozwiązania optymalnego polega wówczas na takim doborze wartości zmiennych decyzyjnych, aby spełnione były ograniczenia, a funkcja celu osiągała wartość maksymalną lub minimalną. Zadanie optymalizacji z ograniczeniami można zapisać następująco:

$$\min f(\mathbf{X}) \quad (3)$$

$\mathbf{X} \in Z_R = \{ \mathbf{X}: g_i(\mathbf{X}) \leq 0, i = 1, \dots, m \}$

gdzie:

$g_i: R^n \rightarrow R^1$ , dla  $i = 1, \dots, m$  – funkcje ograniczeń.

Wiele problemów ekonomicznych można sprowadzić do zagadnienia poszukiwania wartości ekstremalnych pewnych wielkości. Modele optymalizacyjne bez ograniczeń reprezentują problem maksymalizacji dochodu przedsiębiorstwa w warunkach strategii długookresowej, natomiast modele z ograniczeniami – w warunkach strategii krótkookresowej [4].

Dla zadania optymalizacji bez ograniczeń i z ograniczeniami formułowane są warunki konieczne oraz konieczne i wystarczające istnienia rozwiązania. W przypadku optymalizacji bez ograniczeń są to warunki istnienia minimum globalnego funkcji  $n$  zmiennych.

Warunkiem koniecznym istnienia minimum funkcji  $f: \mathbb{R}^n \rightarrow \mathbb{R}^1$  w punkcie  $\mathbf{X}^*$  jest to, aby jej gradient (wektor pochodnych cząstkowych) w tym punkcie był równy zeru.

Warunkiem koniecznym i wystarczającym istnienia dokładnie jednego punktu  $\mathbf{X}^*$ , w którym funkcja  $f: \mathbb{R}^n \rightarrow \mathbb{R}^1$  osiąga minimum globalne jest:

- 1) spełnienie warunku koniecznego;
- 2) dodatnia półokreśloność hesjanu (macierzy drugich pochodnych) funkcji  $f$ , dla każdego  $\mathbf{X}$ .

Warunek konieczny (zerowanie gradientu) pozwala znaleźć punkt optymalny jedynie dla prostych zadań. Problem sprowadza się do rozwiązania układu  $n$  równań. Jeśli funkcja  $f$  jest kwadratowa, układ ten jest układem liniowym i wówczas jego rozwiązanie jest możliwe. W ogólnym przypadku jest to jednak układ nieliniowy, a wobec tego zwykle nie da się go rozwiązać analitycznie.

Jeśli nie można znaleźć rozwiązania zadania optymalizacji bez ograniczeń w sposób analityczny, należy posłużyć się jednym z algorytmów optymalizacji nieliniowej bez ograniczeń przedstawionych w dalszej części artykułu.

Warunki konieczne istnienia rozwiązania zadania optymalizacji z ograniczeniami postaci (3) noszą nazwę warunków Kuhna-Tuckera. Zanim zostaną one przedstawione, konieczne jest wprowadzenia pojęcia funkcji Lagrange'a.

Dla zadania (3) funkcja Lagrange'a ma następującą postać:

$$L(\mathbf{X}, \boldsymbol{\lambda}) = f(\mathbf{X}) + \langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{X}) \rangle \quad (4)$$

gdzie:

$\mathbf{g}(\mathbf{X}) = [g_1(\mathbf{X}), g_2(\mathbf{X}), \dots, g_m(\mathbf{X})]$  jest wektorem ograniczeń,

$\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_m]$  jest wektorem mnożników Lagrange'a

Warunkiem koniecznym istnienia minimum lokalnego dla zadania programowania nieliniowego z ograniczeniami postaci (3) w punkcie  $\mathbf{X}^*$  jest spełnienie następujących warunków:

- 1) funkcje  $f$  i  $g_i$  są różniczkowalne;
- 2) istnieje wektor  $\boldsymbol{\lambda}^* \geq \mathbf{0}$ , taki że:

$$\begin{aligned} \nabla_{\mathbf{x}} L(\mathbf{X}^*, \boldsymbol{\lambda}^*) &= \mathbf{0} \\ \langle \boldsymbol{\lambda}^*, \nabla_{\boldsymbol{\lambda}} L(\mathbf{X}^*, \boldsymbol{\lambda}^*) \rangle &= 0 \\ \nabla_{\boldsymbol{\lambda}} L(\mathbf{X}^*, \boldsymbol{\lambda}^*) &\leq \mathbf{0} \end{aligned} \quad (5)$$

Warunek konieczny istnienia rozwiązania podany wyżej stanie się warunkiem wystarczającym, jeśli dodatkowo założona zostanie pseudowypukłość funkcji  $f$  i quasi-wypukłość wszystkich ograniczeń  $g_i$ . (Definicje funkcji pseudowypukłych i quasi-wypukłych podane są w literaturze [1], [2].)

Przedstawione wyżej warunki konieczne i wystarczające istnienia punktu optymalnego obowiązują dla zadania postaci (3), w którym wszystkie ograniczenia są nierównościowe. Jeśli zadanie optymalizacji zawiera również ograniczenia równościowe, należy warunki te nieco zmodyfikować. Odpowiednie twierdzenia można znaleźć w literaturze [2].

W bardzo prostych przypadkach możliwe jest rozwiązanie zadania na bazie warunków K-T. W większości przypadków jest to niemożliwe i wówczas należy wykorzystać jeden z algorytmów optymalizacji nieliniowej z ograniczeniami.

Podklasą zadań optymalizacji z ograniczeniami jest zadanie programowania liniowego (lub optymalizacji liniowej). Dla takiego zadania funkcja celu i wszystkie ograniczenia muszą być funkcjami liniowymi. Zadanie optymalizacji liniowej rozwiązuje się metodą simpleks. Algorytm simpleks, podobnie jak metody optymalizacji nieliniowej, jest algorytmem iteracyjnym. Oznacza to, że w kolejnych krokach algorytmu otrzymuje się ciąg przybliżeń zbieżny do rozwiązania. Przewaga tego algorytmu nad algorytmami optymalizacji nieliniowej wynika ze sposobu wyznaczania kolejnego rozwiązania. U podstaw algorytmu simpleks leży następujące stwierdzenie: w przypadku liniowego zadania programowania matematycznego  $n$ -wymiarowa liniowa funkcja celu osiąga wartość minimalną lub maksymalną (jeśli ma jedno rozwiązanie) w jednym z wierzchołków  $n$ -wymiarowego wielościanu określonego przez ograniczenia. Aby znaleźć rozwiązanie, wystarczy więc sprawdzać wartość funkcji celu w tych punktach. W przypadku funkcji nieliniowej (bez względu na charakter ograniczeń) rozwiązanie optymalne może znajdować się w dowolnym miejscu, zarówno wewnątrz, jak i na granicy obszaru wyznaczonego ograniczeniami. Z tego też względu dla zadań nieliniowych należy spodziewać się większego nakładu obliczeń niż w przypadku wykorzystania algorytmu simpleks.

## Algorytmy optymalizacji nieliniowej

Metody rozwiązywania zadań optymalizacji nieliniowej formułowane są w postaci algorytmów iteracyjnych.

Algorytm iteracyjny pozwala na wyznaczenie ciągu punktów  $\mathbf{X}^0, \mathbf{X}^1, \mathbf{X}^2, \mathbf{X}^3, \dots$ , takich, że  $\mathbf{X}^k \in \mathbb{R}^n$  (lub  $\mathbf{X}^k \in \mathbb{Z}_R$ , jeśli algorytm ma rozwiązywać zadanie optymalizacji z ograniczeniami). Każdy kolejny punkt  $\mathbf{X}^k$  jest obliczany na

podstawie punktu poprzedniego (proces ten nazywany jest krokiem algorytmu). Algorytm musi być skonstruowany tak, aby generowany ciąg punktów  $\{\mathbf{X}^k\}_{k=1, \dots, \infty}$  coraz dokładniej przybliżał rozwiązanie zadania, lub inaczej mówiąc był zbieżny do tego rozwiązania. Algorytm jest zbieżny do rozwiązania określonego przez punkt  $\mathbf{X}^*$ , jeśli spełniony jest następujący warunek:

$$\lim_{k \rightarrow \infty} \mathbf{X}^k = \mathbf{X}^* \quad (6)$$

Warunek (6) zapewnia jedynie, że ciąg generowany przez algorytm jest zbieżny do rozwiązania w nieskończoności. Oczywiście nie oznacza to, że należy wykonać nieskończoną liczbę kroków. Dokładność obliczeń zakładana jest z góry i wówczas przybliżone rozwiązanie może być wyznaczone w skończonej liczbie kroków.

Schemat działania algorytmu iteracyjnego można przedstawić w następującej postaci:

1. Wyznacz punkt  $\mathbf{X}^0$  (tzw. punkt startowy). Podstaw  $k = 0$ .
2. Oblicz  $\mathbf{X}^{k+1}$  na podstawie  $\mathbf{X}^k$ .
3. Sprawdź, czy punkt  $\mathbf{X}^{k+1}$  spełnia kryterium stopu. Jeśli tak, to koniec obliczeń (znaleziono rozwiązanie). Jeśli nie, to przejdź do kolejnego punktu.
4. Podstaw  $k = k + 1$  i przejdź do punktu 2.

Kryterium stopu jest warunek, którego spełnienie zapewnia rozwiązanie zadania z wystarczającą dokładnością.

## Metody rozwiązywania zadań optymalizacji nieliniowej bez ograniczeń

Metody optymalizacji bez ograniczeń dzielimy na metody poszukiwań prostych i metody poprawy.

W metodach poszukiwań prostych w kolejnych krokach algorytmu badane są wartości funkcji celu w otoczeniu aktualnego przybliżenia rozwiązania, czyli w otoczeniu punktu  $\mathbf{X}^k$ . Sposób wyboru badanych punktów z tego otoczenia zależy od zastosowanej metody. Jeśli wartość funkcji celu w którymś z tych punktów jest mniejsza niż w punkcie  $\mathbf{X}^k$ , dokonywana jest zmiana aktualnego przybliżenia rozwiązania. Jeśli nie znaleziono punktu, w którym następuje poprawa, zmniejszane jest przeszukiwane otoczenie, wobec czego poszukiwania odbywają się z większą dokładnością. Do omawianej grupy metod zaliczane są metody Rosenbrocka, Hooke'a-Jeevsa, Nelder-Mead [2]. Metody te są mało efektywne (znalezienie rozwiązania może wymagać wielu iteracji), ale bardzo

niezawodne. Ponieważ nie wymagają znajomości gradientu, mogą być wykorzystane w przypadkach, gdy funkcja celu jest nieróżniczkowalna.

W przypadku metod poprawy każdy krok algorytmu składa się z dwóch etapów. Pierwszy etap polega na wyznaczeniu kierunku poszukiwań, drugi jest minimalizacją funkcji celu wzdłuż tego kierunku. Znalezione minimum staje się kolejnym przybliżeniem rozwiązania.

Metody poprawy dzielone są na gradientowe i bezgradientowe. W przypadku metod gradientowych w kolejnych iteracjach wykorzystywana jest informacja o wartości funkcji celu i jej gradientu, natomiast w metodach bezgradientowych jedynie o wartości funkcji celu. Gradientowe metody poprawy nazywane są metodami kierunków poprawy. Metody gradientowe są zwykle znacznie efektywniejsze niż bezgradientowe. Dlatego też w przypadku zadań optymalizacji z różniczkowalną funkcją celu wskazane jest posługiwanie się metodami kierunków poprawy (czyli metodami gradientowymi). Jako przykłady bezgradientowych metod poprawy można podać metodę Powella czy Davisa-Swanna-Campeya [2], [3].

Wszystkie metody kierunków poprawy działają według następującego schematu:

1. Wyznacz punkt  $\mathbf{X}^0$  (tzw. punkt startowy). Podstaw  $k = 0$ .
2. Wyznacz kierunek poprawy  $\mathbf{D}^k$ .
3. Znajdź punkt  $\mathbf{X}^k$  będący minimum funkcji celu wzdłuż kierunku  $\mathbf{D}^k$ .
4. Sprawdź, czy punkt  $\mathbf{X}^k$  spełnia kryterium stopu. Jeśli tak, to koniec obliczeń (znaleziono rozwiązanie). Jeśli nie, to przejdź do kolejnego punktu.
5. Podstaw  $k = k + 1$  i przejdź do punktu 2.

Poszczególne metody różnią się będą głównie sposobem wyznaczania kierunku poprawy. W konkretnym algorytmie mogą być zastosowane różne metody minimalizacji w kierunku, jak również różne kryteria stopu.

Wszystkie metody poprawy wykorzystują algorytmy pozwalające znaleźć minimum funkcji celu w kierunku  $\mathbf{D} \in \mathbb{R}^n$ . Zadanie to sprowadza się do wyznaczenia wartości parametru  $\tau^* \geq 0$ , gdzie  $\tau^* \in \mathbb{R}^1$ , takiego, że:

$$f(\mathbf{X}_p + \tau^* \cdot \mathbf{D}) = \min_{\tau} f(\mathbf{X}_p + \tau \cdot \mathbf{D}) \quad (7)$$

gdzie:  $\mathbf{X}_p \in \mathbb{R}^n$  jest punktem początkowym minimalizacji w kierunku.

Znając  $\tau^*$  można wyznaczyć punkt  $\mathbf{X}_w \in \mathbb{R}^n$  będący minimum funkcji  $f$  w kierunku  $\mathbf{D}$  według zależności:

$$\mathbf{X}_w = \mathbf{X}_p + \tau^* \cdot \mathbf{D} \quad (8)$$

Metody minimalizacji w kierunku mogą być metodami bezgradientowymi (np. metoda złotego podziału, metoda interpolacji kwadratowej) i gradientowymi (metoda średnich geometrycznych, metoda interpolacyjno-ekstrapolacyjna). Opis wymienionych metod znaleźć można w literaturze [2], [3].

Kryterium stopu algorytmu optymalizacyjnego ma najczęściej postać:

$$\| \mathbf{X}^{k+1} - \mathbf{X}^k \| \leq \varepsilon \quad (9)$$

W nierówności (9)  $\varepsilon$  jest założoną dokładnością obliczeń, a  $\| \mathbf{X}^{k+1} - \mathbf{X}^k \|$  to odległość przybliżonych rozwiązań uzyskiwanych w kolejnych iteracjach. Ponieważ spełnienie warunku stopu może wymagać bardzo dużego nakładu obliczeń, stosuje się dodatkowe zabezpieczenia. Przykładem może być założenie z góry maksymalnej liczby iteracji bądź obliczeń funkcji celu.

Najbardziej znane metody kierunków poprawy to metody najszybszego spadku, Newtona, gradientu sprzężonego i zmiennej metryki. Krótki opis wymienionych algorytmów zostanie przedstawiony poniżej.

W metodzie najszybszego spadku jako kierunek poszukiwań przyjmowany jest kierunek minus gradientu, czyli:

$$\mathbf{D}^k = -\nabla_x f(\mathbf{X}^k) \quad (10)$$

Następnie wykonywana jest minimalizacja wzdłuż tego kierunku, pozwalająca obliczyć punkt  $\mathbf{X}^{k+1}$ . W punkcie tym wyznaczany jest następnie gradient, który stanowi podstawę do określenia kolejnego kierunku, wzdłuż którego przeprowadzana zostanie minimalizacja.

Najczęściej stosowanym kryterium stopu jest warunek wyznaczony nierównością (9).

Istotną wadą omawianej metody jest wyraźne obniżenie szybkości zbieżności w wypadku funkcji, dla których minimum leży w tzw. wąskiej dolinie.

W metodzie Newtona kierunkiem poszukiwań jest kierunek minus gradientu pomnożony przez odwrotność hesjanu funkcji celu:

$$\mathbf{D}^k = -\mathbf{H}^{-1}(\mathbf{X}^k) \cdot \nabla_x f(\mathbf{X}^k) \quad (11)$$

gdzie:  $\mathbf{H}(\mathbf{X}^k)$  jest wartością hesjanu w punkcie  $\mathbf{X}^k$ .

Metoda Newtona charakteryzuje się szybką zbieżnością do rozwiązania. Istotną jej wadą jest konieczność wyznaczania i odwracania w każdym kroku hesjanu. Jak wiadomo, w przypadku dużego wymiaru macierzy poszukiwanie jej odwrotności sprawia istotne trudności numeryczne.

W literaturze spotkać można wiele odmian metody gradientów sprzężonych. Wspólną ich cechą jest to, że kierunki poszukiwań generowane w kolejnych krokach są kierunkami sprzężonymi. Definicję kierunków sprzężonych można znaleźć w literaturze (2).

W pierwszym kroku algorytmu jako kierunek poszukiwań minimum przyjmowany jest kierunek minus gradientu. Minimum funkcji staje się przybliżeniem rozwiązania w kroku następnym. W  $k$ -tym kroku kierunek poszukiwań określany jest według reguły:

$$\mathbf{D}^k = -\nabla_x f(\mathbf{X}^k) + \beta_k \cdot \mathbf{D}^{k-1} \quad (12)$$

gdzie  $\beta_k$  jest współczynnikiem.

Poszczególne odmiany metod gradientu sprzężonego różnią się sposobem wyznaczania współczynnika  $\beta$ . Odpowiednie wzory podane są w literaturze (2).

W metodzie, a właściwie grupie metod zmiennej metryki w kolejnych krokach algorytmu generowany jest ciąg macierzy  $\{\mathbf{V}^k\}$  będących przybliżeniami odwrotności hesjanu. Ciąg ten wyznaczany jest na podstawie zmian gradientu funkcji celu w poprzednio wykonanych krokach. Metody zmiennej metryki są podobne do metody Newtona i dlatego nazywane są metodami quasi-newtonowskimi.

W każdym kroku algorytmu tworzony jest kierunek poszukiwań według zależności:

$$\mathbf{D}^k = -\mathbf{V}_k \cdot \nabla_x f(\mathbf{X}^k) \quad (13)$$

Następnie w wyniku minimalizacji funkcji w tym kierunku otrzymywany jest punkt  $\mathbf{X}^{k+1}$ .

Poszczególne odmiany metody zmiennej metryki różnią się sposobem wyznaczania macierzy  $\mathbf{V}_k$ . Wzory te są dość złożone i z tego względu nie zostaną tu przedstawione – można znaleźć je w literaturze (2).

## Metody rozwiązywania zadań optymalizacji nieliniowej z ograniczeniami

Metody optymalizacji nieliniowej z ograniczeniami zaliczane są do metod aproksymacyjnych. Zamiast zadania wyjściowego określonego zależnością (3) rozwiązywany jest ciąg zadań zastępczych. Ciąg ten tworzony jest w taki sposób, aby zadania zastępcze w kolejnych iteracjach coraz lepiej przybliżały zadanie wyjściowe.



Najbardziej znanymi grupami metod optymalizacji nieliniowej z ograniczeniami są metody funkcji kary, metody modyfikacji kierunku i metoda aproksymacji kwadratowej.

W metodach funkcji kary dokonywana jest zamiana zadania optymalizacji z ograniczeniami na zadanie bez ograniczeń. Polega ona na zmodyfikowaniu funkcji celu przez wprowadzenie do niej kary za przekroczenie ograniczeń. Poszczególne odmiany metod różnią się sposobem wprowadzenia kary.

W metodzie przesuwanej funkcji kary opracowanej przez Powella funkcja celu modyfikowana jest w następujący sposób:

$$P(\mathbf{X}, \mathbf{A}, \mathbf{U}) = f(\mathbf{X}) + \frac{1}{2} \cdot \sum_{i=1}^m \alpha_i \cdot [\max(0; g_i(\mathbf{X}) + u_i)]^2 \quad (14)$$

Zadanie zastępcze polegać będzie na minimalizacji funkcji  $P(\mathbf{X}, \mathbf{A}, \mathbf{U})$  względem  $\mathbf{X}$  zamiast funkcji  $f(\mathbf{X})$ . Wektor  $\mathbf{A} = [\alpha_1, \alpha_2, \dots, \alpha_m]$  nazywany jest wektorem współczynników funkcji kary. Współczynniki te muszą być dodatnie. Wektor  $\mathbf{U} = [u_1, u_2, \dots, u_m]$  natomiast jest wektorem przesunięć kary. Wymiar wektorów  $\mathbf{A}$  i  $\mathbf{U}$  równy jest liczbie ograniczeń. Wektory te mają ustalone wartości w każdym kroku, po czym są tak modyfikowane, aby ciąg  $\{\mathbf{X}^k\}$  dążył przy  $k \rightarrow \infty$  do rozwiązania zadania (3).

Wprowadzenie kary za przekroczenie ograniczeń do funkcji celu pozwala zastąpić zadanie optymalizacji z ograniczeniami zadaniem bez ograniczeń. Zmodyfikowane zadanie jest następnie rozwiązywane metodami optymalizacji nieliniowej bez ograniczeń.

Metody z modyfikacją kierunku dzielą się na metody kierunków dopuszczalnych i metody rzutu ortogonalnego ([2], [3]). Różnią się one sposobem wyznaczania kierunku, w którym minimalizowana jest funkcja celu. W pierwszym przypadku w otoczeniu ograniczeń generowane są kierunki dopuszczalne (przynajmniej częściowo mieszczące się w obszarze rozwiązań dopuszczalnych), w drugim natomiast kierunek gradientu rzutowany jest na powierzchnię styczną do ograniczeń.

W metodzie aproksymacji kwadratowej zadanie optymalizacji nieliniowej z ograniczeniami zastępowane jest w poszczególnych krokach odpowiednio dobranym zadaniem programowania kwadratowego (kwadratowa funkcja celu i liniowe graniczenia). Reguły tworzenia zadania zastępczego znaleźć można w literaturze ([2], [3]).

## Problemy związane ze stosowaniem algorytmów optymalizacji nieliniowej

Zastosowanie konkretnego algorytmu optymalizacji nieliniowej wymaga od użytkownika podania wartości parametrów sterujących oraz punktu startowego. Do grupy parametrów sterujących zaliczamy takie parametry, jak dokładność obliczeń czy maksymalna liczba wyznaczeń wartości funkcji celu. W metodach kierunków poprawy należy dodatkowo podać dokładność wyznaczania minimum w kierunku.

Parametry związane z dokładnością w istotny sposób wpływają na przebieg procesów obliczeniowych. Zbyt duża ich wartość powoduje wydłużenie czasu obliczeń, jak również z powodu błędów numerycznych może w pewnych wypadkach uniemożliwić rozwiązanie. Za mała natomiast może prowadzić do błędnego rozwiązania. Dokładność powinna być odpowiednio dobrana do stosowanej metody i do konkretnego zadania. Zwykle doboru dokonuje się metodą prób i błędów. W pewnych przypadkach może wystąpić konieczność uruchamiania algorytmu kilkakrotnie z różnymi wartościami parametrów. Wówczas obowiązuje następująca zasada: im bliżej rozwiązania wykonywane są obliczenia, tym większa powinna być dokładność. Wyznaczone punkty – najlepiej przybliżające rozwiązanie – stają się punktami startowymi w kolejnych uruchomieniach algorytmu.

Dobór punktu startowego ma również bardzo duże znaczenie. Im punkt ten znajduje się bliżej rozwiązania, tym większa gwarancja pomyślnego zakończenia obliczeń. Wiele metod wymaga poza podaniem punktu startowego oszacowania odległości od rozwiązania. Należy podać wielkość promienia kuli o środku w punkcie startowym, w której powinno znajdować się rozwiązanie. Jest to szczególnie trudne, ponieważ zwykle nie znamy rozwiązania zadania nawet w sposób przybliżony. Promień należy wówczas określić przeprowadzając eksperymenty polegające na wielokrotnym uruchamianiu algorytmu z różnymi jego wartościami. Obserwacja wyników obliczeń pozwoli ocenić, jaka wartość jest najbardziej odpowiednia. W przypadku zadań z ograniczeniami należy ponadto zadbać, aby punkt startowy znajdował się w obszarze rozwiązań dopuszczalnych.

Kolejna grupa problemów może wynikać ze złego uwarunkowania zadania optymalizacji. Zadanie optymalizacji bez ograniczeń jest źle uwarunkowane, jeśli różnica między najmniejszą a największą wartością hesjanu funkcji celu w pobliżu rozwiązania jest duża. Konsekwencją tego są bardzo duże różnice w przyrostach funkcji celu spowodowanych jednakowymi przyrostami zmien-

nych decyzyjnych. Przyrost wartości funkcji w wyniku zmiany wartości jednej zmiennej może być o kilka rzędów wielkości większy niż spowodowany taką samą zmianą innej zmiennej. W wypadku złego uwarunkowania należy przeprowadzić skalowanie zmiennych. Skalowanie polega na takim doborze jednostek, w jakich wyrażane są zmienne decyzyjne, aby ich wpływ na wartość funkcji celu był podobny.

W przypadku zadań optymalizacji z ograniczeniami o złym uwarunkowaniu zadania mówi się o Hessjan funkcji Lagrange'a. Złe uwarunkowanie może w tym przypadku przejawiać się dużymi różnicami we wpływie poszczególnych zmiennych decyzyjnych na wartości funkcji celu, jak również na ograniczenia. W takich wypadkach należy poza skalowaniem zmiennych zastosować skalowanie ograniczeń, które polega na pomnożeniu ograniczeń przez właściwie dobrane wielkości.

## Literatura

1. CHIANG A., 1994: *Podstawy ekonomii matematycznej*. PWE, Warszawa.
2. FINDEISEN W., SZYMANOWSKI J., WIERZBICKI A., 1977: *Teoria i metody obliczeniowe optymalizacji*. PWN, Warszawa.
3. KRĘGLEWSKI T., ROGOWSKI T., RUSZCZYŃSKI A., SZYMANOWSKI J., 1984: *Metody optymalizacji w języku FORTRAN*. PWN, Warszawa.
4. PANEK E., 1993: *Elementy ekonomii matematycznej. Statyka*. Wydawnictwo Naukowe PWN, Warszawa.

## The elements of optimisation theory and methods of solving non-linear optimisation problems

### Abstract

This article contains the elements of optimisation theory – the mathematical forms of nonlinear optimisation problems without constraints and with constraints and preconditions and sufficient conditions for existence of the optimal solution.

Next the author presents methods of non-linear optimisation. This presentation includes methods of solving problems without constraints and with constraints. Difficulties occurred when using these methods are discussed at the end.