

Uogólniona metoda Bendersa dekompozycji mieszanych, nieliniowych zadań optymalizacji

Streszczenie. Artykuł przedstawia uogólnioną metodę dekompozycji Bendersa, która jest obecnie jednym z podstawowych podejść do rozwiązywania dużych zadań nieliniowej optymalizacji mieszanej (dyskretno-ciągłej), także w przypadku dość szerokiej klasy zadań z niwypukłymi funkcjami celu oraz ograniczeń. Oprócz klasycznych twierdzeń o rzutowaniu i reprezentacji, podane będzie jednolite sformułowanie zadania mastera z cięciami nieliniowymi i liniowymi. Dla tego ostatniego przypadku wskazane będą najbardziej efektywne oraz łatwe w implementacji algorytmy obliczeniowe z rodziny płaszczyzn tnących.

Abstract. The paper presents the Generalized Benders Decomposition method, which is now one of the basic approaches to solve big mixed-integer nonlinear optimization problems, also in the case of quite a broad class of problems with nonconvex objectives and constraints functions. Apart from the classical projection and representation theorems, a unified formulation of the master problem with nonlinear and linear cuts will be given. For the latter case the most effective and, at the same time, easy to implement computational algorithms from the cutting plane family will be pointed out. (**Generalized Benders method of decomposition of mixed-integer nonlinear optimization problems**)

Słowa kluczowe: optymalizacja, algorytmy mieszane, dekompozycja, zadania niwypukłe, zadania biliniowe, programowanie całkowitoliczbowe
Keywords: optimization, mixed-integer algorithms, decomposition, nonconvex problems, bilinear problems, integer programming

Wprowadzenie

Na początku lat 60-tych XX w. holenderski matematyk Jacques F. Benders rozpatrywał zadania o postaci [2]:

$$(1) \quad \max_{x,v} c^T x + f(v)$$

przy ograniczeniach (dalej będzie używany skrót „p.o.”)

$$(2) \quad Ax + F(v) \leq b, \quad b \in \mathbb{R}^m$$

$$(3) \quad x \in X \subseteq \mathbb{R}^n, \quad v \in V \subseteq \mathbb{R}^q$$

Nazywał je zadaniami programowania z mieszanymi zmiennymi (ang. *mixed-variables programming*). W zadaniach tych zarówno wskaźniki jakości, jak i funkcje definiujące ograniczenia, były sumami dwóch składników: liniowego, zależnego od wektora zmiennych $x \in X$ i nieliniowego, zależnego od wektora $v \in V$. Benders nazwał zmienne v zmiennymi komplikującymi, ponieważ przy ich ustaleniu zadanie bardzo się upraszczało - stawało się liniowe. Ponadto zasugerował, że mogą one być dyskretne, czyli zadanie względem samych zmiennych x stawało się także ciągłe i było już bardzo łatwe do rozwiązania za pomocą dowolnego solwera programowania liniowego (np. opartego na metodzie sympleks).

W swojej pracy [2] Benders zaproponował iteracyjną procedurę rozwiązania tego zadania optymalizacji poprzez rozwiązywanie na przemian pomocniczych zadań względem jednego lub drugiego wektora zmiennych. Zadania te były związane z reprezentacją dualną zadania wyjściowego oraz warunkami optymalności. Zadanie względem zmiennych komplikujących v nazwał zadaniem mastera, zadanie względem zmiennych x - zadaniem prymalnym. Zadanie prymalne dostarczało zadaniu mastera w kolejnych jego uruchomieniach ograniczeń związanych z aproksymacją wskaźnika jakości, albo zbioru dopuszczalnego, wyliczanych w różnych punktach zbioru X . Te pierwsze były dodawane w przypadku, gdy dla danego próbnego wektora v istniało rozwiązanie dopuszczalne w zadaniu wyjściowym, te drugie - w przeciwnym razie.

Uogólnienie tego podejścia na zadania nieliniowe przedstawił Arthur Geoffrion [10]. Nazwał je uogólnioną dekompozycją Bendersa (ang. *Generalized Benders Decomposition - GBD*). Tak jak w oryginalnej pracy Bendersa, ograniczenia w zadaniu mastera były zdefiniowane za pomocą lagranżianów odpowiednich zadań prymalnych. Wartości wskaźników w punktach optymalnych uzyskane po każdej iteracji zadania

mastera oraz prymalnego, gdy było dopuszczalne, służyły do oszacowania wskaźnika zadania wyjściowego. W przypadku zadania minimalizacji zadanie mastera dostarczało jego oszacowania od dołu *LBD* (od ang. *lower bound*), zadanie prymalne - od góry - *UBD* (od ang. *upper bound*). Gdy zadanie prymalne było niedopuszczalne, rozwiązywano tzw. zadanie dopuszczalności, polegające na minimalizacji maksymalnego przekroczenia ograniczeń. Test stopu polegał na sprawdzeniu, czy oszacowania górne i dolne wskaźnika jakości są sobie równe zadaną dokładnością.

Podejście Geoffriona w nieznacznie zmodyfikowanej wersji zostało przedstawione w pracach Floudasa [8], [9], w których zakłada się, że zmienne v są binarne oraz uzupełnia ograniczenia zadania o równościowe, mieszane. Niestety, w zadaniu dopuszczalności nie są one relaksowane (tak jak ograniczenia mieszane nierównościowe), co sprawia, że całe podejście może doprowadzić do blokady algorytmu, gdy wybierze się punkt próbny ze zbioru V dla którego któreś z tych ograniczeń jest niespełnione dla wszystkich $x \in X$. Floudas podał kilka wersji algorytmu obliczeniowego, ale dotyczyły one tych samych specyficznych zadań, będąc właściwie jednym algorytmem zapisanym w różny sposób.

Niestety, ani w pracach Floudasa [8], [9], ani w źródłowej pracy Geoffriona [10], nie została przedstawiona wersja o podstawowym znaczeniu w zadaniach wypukłych względem całego wektora (x, v) - z liniowymi cięciami w zadaniu mastera. Jest to także wersja najprostsza w implementacji, wspomina o niej Benders [2], a w nowszych pracach dotyczących programowania mieszanego [17], [19] najczęściej jest przedstawiana jako jedyna.

Niedostatki wymienionych powyżej prezentacji metody Bendersa i brak spójnego i jednolitego opisu tego podejścia przekonały autora o potrzebie uporządkowania wiedzy, czego efektem jest niniejsza praca. A znaczenie tego podejścia rośnie, gdyż rozwiązywane są obecnie coraz większe praktyczne zadania programowania mieszanego, związane np. z energetyką [6], telekomunikacją [14], transportem [11], sieciami gazowymi [20], procesami produkcyjnymi [22], systemami wodnymi [5]. Gdy zadania te przekroczą pewną wielkość, nawet najlepsze solwery komercyjne nie dają sobie rady. Niezbędne jest napisanie własnego specjalizowanego kodu, wykorzystującego własności strukturalne zadania. A do tego właśnie metoda Bendersa świetnie się nadaje.

Sformułowanie zadania

Rozważmy zadanie optymalizacji o następującej postaci:

$$(4) \quad \min_{x,v} f(x, v)$$

p.o.

$$(5) \quad g(x, v) \leq 0$$

$$(6) \quad x \in X \subseteq \mathbb{R}^n$$

$$(7) \quad v \in V \subseteq \mathbb{R}^q$$

gdzie

$$\begin{aligned} f &: \mathbb{R}^n \times \mathbb{R}^q \rightarrow \mathbb{R}, \\ g &: \mathbb{R}^n \times \mathbb{R}^q \rightarrow \mathbb{R}^m \end{aligned}$$

Ewentualne ograniczenia równościowe

$$(8) \quad h_j(x, v) = 0, \quad j = 1, \dots, s$$

można uwzględnić za pomocą standardowych przekształceń, np.:

$$(9) \quad \begin{cases} h_1(x, v) \leq 0 \\ \vdots \\ h_s(x, v) \leq 0 \\ \sum_{j=1}^s h_j(x, v) \geq 0 \end{cases}$$

lub

$$(10) \quad h_j^2(x, v) \leq 0, \quad j = 1, \dots, s$$

Dla uproszczenia ograniczenia równościowe będą dalej pominięte.

Założenia odnoszące się do funkcji i zbiorów występujących w sformułowaniu będą podawane w kolejnych twierdzeniach. Na razie przyjmujemy tylko, że są one takie, że rozwiązanie zadania (4)-(7) istnieje. Warto jedynie poświęcić kilka słów zbiorowi V . Ma on zawierać się w \mathbb{R}^q . W związku z tym podklasą zadań tej postaci są zadania typu MINLP (od ang. *Mixed Integer NonLinear Programming*), czyli zadania mieszane ciągle-dyskretne, nieliniowe, w których $V \subseteq \mathbb{Z}^q$, gdzie \mathbb{Z} - zbiór liczb całkowitych.

Sugeruje się określenie wektora v - zmiennych komplikujących, w taki sposób, że gdy go ustalimy, to wystąpi jedna z następujących sytuacji:

1. Zadanie może być zdekomponowane na szereg niezależnych podzadań, z których każde wykorzystuje inny podwektor $x_i, i = 1, \dots, p$ wektora

$$(11) \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

Najczęściej wskaźnik łączny ma postać addytywną:

$$(12) \quad \min_{x,v} f(x, v) = \min_{x_1, x_2, \dots, x_p, v} \sum_{i=1}^p f_i(x_i, v) + f_0(v)$$

p.o.

$$(13) \quad g_{ij}(x_i, v) \leq 0, \quad j = 1, \dots, m_i; \quad i = 1, \dots, p$$

$$(14) \quad x_i \in X_i \subseteq \mathbb{R}^{n_i}, \quad i = 1, \dots, p, \quad v \in V \subseteq \mathbb{R}^q,$$

$$\sum_{i=1}^p n_i = n$$

W tym przypadku warto jest zastosować nie tylko dekompozycję, ale także równoległe rozwiązywanie powstałych podzadań prymalnych.

2. Zadanie względem zmiennych x przyjmuje specyficzną strukturę dla której istnieją wydajne algorytmy - np. jest liniowe albo kwadratowe.

3. Zadanie jest wypukłe względem x przy ustalonym v i vice-versa, mimo że jest niewypukłe, gdyby rozważać je dla zmiennej łącznej (x, v) .

Wypukłe względem jednego podwektora przy ustalonym drugim są np. często spotykane w technice zadania biliniowe [1].

Przykłady zadań optymalizacji do rozwiązania których można zastosować metodę Bendersa

1. Zadanie z separowalnym wskaźnikiem jakości oraz separowalną funkcją ograniczeń

$$(15) \quad \min_{y_1, \dots, y_6} [2y_1^2 - y_1 \cdot y_2 + 4y_2^2 + (y_3 - 4)^2 + (y_4 - 3)^2 + 8y_5^2 + y_6^2 - 3y_5]$$

p.o.

$$(16) \quad y_1 \geq 1, \quad y_2 \geq 0$$

$$(17) \quad y_3 \geq 3, \quad y_4 \geq 2$$

$$(18) \quad y_5 \geq 3, \quad y_6 \geq 0$$

$$(19) \quad y_1 + y_2 + y_3^2 + y_4^2 + y_5 + y_6^2 \leq 50$$

W zadaniach takich, gdzie zarówno wskaźnik jakości, jak i funkcje ograniczeń są sumami funkcji zależnych od podwektorów zmiennych decyzyjnych, żeby uzyskać postać standardową (12)-(14) trzeba odpowiednio pogrupować zmienne w podwektory oraz wprowadzić dodatkowe zmienne v rozdzielające ograniczenia.

Przyjmijmy $p = 3$ oraz oznaczmy:

$$(20) \quad x_1 = \begin{bmatrix} x_{1,1} \\ x_{1,2} \end{bmatrix} \triangleq \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad x_2 = \begin{bmatrix} x_{2,1} \\ x_{2,2} \end{bmatrix} \triangleq \begin{bmatrix} y_3 \\ y_4 \end{bmatrix},$$

$$x_3 = \begin{bmatrix} x_{3,1} \\ x_{3,2} \end{bmatrix} \triangleq \begin{bmatrix} y_5 \\ y_6 \end{bmatrix}$$

Teraz:

$$(21) \quad f_1(x_1) = 2x_{1,1}^2 - x_{1,1}x_{1,2} + 4x_{1,2}^2$$

$$(22) \quad X_1 = \{[x_{1,1}, x_{1,2}] \in \mathbb{R}^2 : x_{1,1} \geq 1, x_{1,2} \geq 0\}$$

$$(23) \quad f_2(x_2) = (x_{2,1} - 4)^2 + (x_{2,2} - 3)^2$$

$$(24) \quad X_2 = \{[(x_{2,1}, x_{2,2})] \in \mathbb{R}^2 : x_{2,1} \geq 3, x_{2,2} \geq 2\}$$

$$(25) \quad f_3(x_3) = 8x_{3,1}^2 + x_{3,2}^2 - 3x_{3,1}$$

$$(26) \quad X_3 = \{[x_{3,1}, x_{3,2}] \in \mathbb{R}^2 : x_{3,1} \geq 3, x_{3,2} \geq 0\}$$

$$(27) \quad f_0(v) \equiv 0$$

Wprowadźmy także sztuczne zmienne v_1, v_2, v_3 na ograniczenie składników wiążących zmienne podwektorów, odpowiednio, x_1, x_2 i x_3 , w ograniczeniu łącznym (19).

$$(28) \quad x_{1,1} + x_{1,2} \leq v_1$$

$$(29) \quad x_{2,1}^2 + x_{2,2}^2 \leq v_2$$

$$(30) \quad x_{3,1} + x_{3,2}^2 \leq v_3$$

Otrzymamy następujące funkcje ograniczeń w formacie zadania (12)-(14):

$$(31) \quad g_{11}(x_1, v) = x_{1,1} + x_{1,2} - v_1$$

$$(32) \quad g_{21}(x_2, v) = x_{2,1}^2 + x_{2,2}^2 - v_2$$

$$(33) \quad g_{31}(x_3, v) = x_{3,1} + x_{3,2}^2 - v_3$$

a także zbiór V :

$$(34) \quad V = \{[v_1, v_2, v_3] \in \mathbb{R}^3 : v_1 + v_2 + v_3 \leq 50\}$$

2. Zadanie z łańcuchem (pierścieniem) ograniczeń.

W zadaniach tego typu kolejne ograniczenia wiążą ze sobą kilka kolejnych zmiennych decyzyjnych, tzn. każda ze zmiennych występuje w kilku ograniczeniach (przynajmniej w dwóch) w relacji z sąsiednią zmienną o niższym oraz o wyższym indeksie. Jest to struktura dość typowa dla zadań sterowania optymalnego po zastąpieniu układu s równań nieliniowych (od początku różnicowych lub powstałego w rezultacie dyskretyzacji po czasie układu równań różniczkowych) układem $s + 1$ nierówności przy użyciu przekształcenia (9). Pomysłem na podział zadania i uzyskanie postaci standardowej (12)-(14) jest potraktowanie niektórych zmiennych w wektorze (np. dzielących go na kilka równych części) jako zmiennych komplikujących v_i , a następnie ograniczeń, w których te zmienne występują, jako ograniczeń mieszanych typu (13).

Weźmy pod uwagę zadanie następujące:

$$\min_{y \in \mathbb{R}^n} y_1^2 + y_2^2 + \dots + y_9^2$$

p.o.

$$y_{k+1} - y_k \leq \sin k, \quad k = 1, \dots, 8$$

$$y_1 - y_9 \leq 0.5$$

Oznaczmy:

$$v_1 \triangleq y_3, \quad v_2 \triangleq y_6, \quad v_3 \triangleq y_9$$

Pozostałe elementy będą tworzyć podwektory o wymiarze 2. Oznaczmy je jako x_1, x_2, x_3 , czyli:

$$x_1 = \begin{bmatrix} x_{1,1} \\ x_{1,2} \end{bmatrix} \triangleq \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad x_2 = \begin{bmatrix} x_{2,1} \\ x_{2,2} \end{bmatrix} \triangleq \begin{bmatrix} y_4 \\ y_5 \end{bmatrix},$$

$$x_3 = \begin{bmatrix} x_{3,1} \\ x_{3,2} \end{bmatrix} \triangleq \begin{bmatrix} y_7 \\ y_8 \end{bmatrix}$$

Zadania to można przedstawić w formacie (12)-(14) przyjmując:

$$(35) \quad f_1(x_1) = x_{1,1}^2 + x_{1,2}^2$$

$$(36) \quad X_1 = \{[x_{1,1}, x_{1,2}] \in \mathbb{R}^2 : x_{1,2} - x_{1,1} \leq \sin 1\}$$

$$(37) \quad g_{11}(x_1, v) = x_{1,1} - v_3 - 0.5$$

$$(38) \quad g_{12}(x_1, v) = v_1 - x_{1,2} - \sin 2$$

$$(39) \quad f_2(x_2) = x_{2,1}^2 + x_{2,2}^2$$

$$(40) \quad X_2 = \{[x_{2,1}, x_{2,2}] \in \mathbb{R}^2 : x_{2,2} - x_{2,1} \leq \sin 4\}$$

$$(41) \quad g_{21}(x_2, v) = x_{2,1} - v_1 - \sin 3$$

$$(42) \quad g_{22}(x_2, v) = v_2 - x_{2,2} - \sin 5$$

$$(43) \quad f_3(x_3) = x_{3,1}^2 + x_{3,2}^2$$

$$(44) \quad X_3 = \{[x_{3,1}, x_{3,2}] \in \mathbb{R}^2 : x_{3,2} - x_{3,1} \leq \sin 7\}$$

$$(45) \quad g_{31}(x_3, v) = x_{3,1} - v_2 - \sin 6$$

$$(46) \quad g_{32}(x_3, v) = v_3 - x_{3,2} - \sin 8$$

$$(47) \quad f_0(v) = \sum_{i=1}^3 v_i^2$$

Dekompozycja.

Zadanie (4)-(7) można przedstawić w następujący sposób [10]:

$$(48) \quad \min_{v \in V} \inf_{x \in X} f(x, v)$$

p.o.

$$(49) \quad g(x, v) \leq 0$$

Infimum w zadaniu wewnętrznym pojawiło się ze względu na to, że dla pewnych v może ono być nieograniczone. Oznaczmy

$$(50) \quad z(v) \triangleq \inf_{x \in X} f(x, v)$$

p.o.

$$(51) \quad g(x, v) \leq 0$$

Zadanie (50)-(51) rozwiązywane dla ustalonego v będziemy nazywać zadaniem prymalnym,

Zadanie (48)-(49) może być teraz zapisane w postaci

$$(52) \quad \min_{v \in V \cap V_0} z(v)$$

gdzie

$$(53) \quad V_0 = \{v : \exists x \in X \quad g(x, v) \leq 0\}$$

i interpretowane jako rzutowanie zadania (48)-(49) na przestrzeń zmiennych v [10]. Sformułowanie (52) nazywamy zadaniem мастера.

Wymaganie na v określone w (52) jako $v \in V \cap V_0$ wynika z konieczności zagwarantowania istnienia rozwiązania - wartości $z(v)$. Zbiór V_0 nosi nazwę zbioru rozwiązywalności (ang. *solvability set*).

Problem polega na tym, że znamy $z(v)$ oraz V_0 jedynie pośrednio poprzez ich definicje.

Obowiązują następujące twierdzenia [10]:

Twierdzenie 1 (o rzutowaniu)

1. Zadanie (4)-(7) nie ma rozwiązania lub jest nieograniczone wtedy i tylko wtedy gdy to samo jest prawdziwe w odniesieniu do zadania (52).
2. Jeżeli (\hat{x}, \hat{v}) jest optymalnym rozwiązaniem zadania (4)-(7), to \hat{v} jest optymalnym rozwiązaniem zadania (52),
3. Jeżeli \hat{v} jest optymalnym rozwiązaniem zadania (52) i \hat{x} osiąga infimum w zadaniu (50)-(51) przy $v = \hat{v}$, to (\hat{x}, \hat{v}) jest optymalnym rozwiązaniem zadania (4)-(7)

Twierdzenie 2 (o reprezentacji V_0)

Założmy że X jest niepustym, wypukłym zbiorem oraz że funkcja g jest wypukła na X dla każdego ustalonego $v \in V$.

Założmy ponadto że zbiór

$$Z_v = \{z \in \mathbb{R}^m : \exists x \in X \quad g(x, v) \leq z\}$$

jest domknięty dla każdego ustalonego $v \in V$. Wówczas punkt $v^* \in V$ należy także do zbioru V_0 wtedy i tylko wtedy gdy:

$$(54) \quad \inf_{x \in X} L_f(x, v^*, \lambda) \leq 0, \quad \forall \lambda \in \Lambda$$

gdzie

$$(55) \quad \Lambda = \left\{ \lambda \in \mathbb{R}^m : \lambda \geq 0, \sum_{j=1}^m \lambda_j = 1 \right\}$$

oraz

$$(56) \quad L_f(x, v, \lambda) = \lambda^T g(x, v)$$

Twierdzenie 3 (o reprezentacji $z(v)$)

Założmy, że X jest niepustym, wypukłym zbiorem, a funkcje f i g wypukłe na X dla każdego ustalonego $v = v^* \in V$. Założmy ponadto, że dla v^* , przynajmniej jeden z poniższych warunków jest spełniony:

1. $z(v^*)$ jest skończone i w zadaniu (50)-(51) istnieje optymalny wektor mnożników Lagrange'a
2. $z(v^*)$ jest skończone, $g(x, v^*)$ i $f(x, v^*)$ są ciągłe na X , zbiór X jest domknięty, zaś zbiór rozwiązań optymalnych z dokładnością do ε zadania (50)-(51) jest niepusty i ograniczony dla pewnego $\varepsilon \geq 0$
3. $z(v^*) = -\infty$

Wówczas

$$(57) \quad z(v) = \sup_{\lambda \geq 0} \inf_{x \in X} L_o(x, v, \lambda), \quad \forall v \in V \cap V_0$$

gdzie

$$(58) \quad L_o(x, v, \lambda) = f(x, v) + \lambda^T g(x, v)$$

Podstawiając w zadaniu (52) zależność (57) za $z(v)$ i (54) za ograniczenie $v \in V_0$ uzyskujemy równoważne zadanie:

$$(59) \quad \min_{v \in V} \sup_{\lambda \geq 0} \inf_{x \in X} L_o(x, v, \lambda)$$

p.o.

$$(60) \quad \inf_{x \in X} L_f(x, v, \lambda) \leq 0, \quad \lambda \in \Lambda$$

Korzystając z definicji supremum jako najmniejszego górnego ograniczenia oraz wprowadzając dodatkową, skalarną

zmienną μ , otrzymujemy następującą postać zadania mastera, równoważną (59)-(60):

$$(61) \quad \min_{v \in V, \mu} \mu$$

p.o.

$$(62) \quad \inf_{x \in X} L_o(x, v, \lambda) \leq \mu, \quad \forall \lambda \geq 0$$

$$(63) \quad \inf_{x \in X} L_f(x, v, \lambda) \leq 0, \quad \forall \lambda \in \Lambda$$

W zadaniach wywodzących się z praktyki można zakładać, że funkcja $z(v)$ (patrz wzór (50)) jest ograniczona dla wszystkich $v \in V$, zbiór X zwarty, a funkcje $f(x, v)$ i $g(x, v)$ ciągłe w całej dziedzinie. W związku z tym możemy infimum zastąpić przez minimum. Wtedy zadanie mastera przyjmie następującą postać:

$$(64) \quad \min_{v \in V, \mu} \mu$$

p.o.

$$(65) \quad \min_{x \in X} L_o(x, v, \lambda) \leq \mu, \quad \forall \lambda \geq 0$$

$$(66) \quad \min_{x \in X} L_f(x, v, \lambda) \leq 0, \quad \forall \lambda \in \Lambda$$

gdzie funkcje L_o i L_f są dane wzorami (58), (56).

Zadanie (64)-(66) jest bardzo trudne do rozwiązania ze względu na ograniczenia, które muszą być spełnione w nieskończonej, a nawet nieprzeliczalnej, ilości punktów (dla wszystkich λ o współrzędnych nieujemnych albo nieujemnych sumujących się do jedności) oraz istnienie wewnętrznych podzadań optymalizacyjnych (minimalizacji po x).

Trudności te pokonuje się za pomocą relaksacji zadania, a dokładniej rozwiązywania go przy wykorzystaniu kolejnych, coraz dokładniejszych, aproksymacji funkcji występujących po lewej stronie ograniczeń (65) oraz (66). Będą one utworzone z kawałków funkcji L_o albo L_f związanych z optymalnymi rozwiązaniami zadania prymalnego, w sytuacji, odpowiednio, gdy kolejny punkt próbny z zadania mastera v^* należy do V_0 albo nie.

Podstawowe własności zadania prymalnego

Zadanie prymalne to zadanie wyjściowe (4)-(7) rozwiązywane dla ustalonego $v = v^* \in V$

$$(67) \quad \min_{x \in X} f(x, v^*)$$

p.o.

$$(68) \quad g(x, v^*) \leq 0$$

Przy rozwiązywaniu tego zadania mogą wystąpić dwa przypadki: gdy zadanie prymalne jest dopuszczalne, czyli ma rozwiązanie, i gdy go nie ma.

Przypadek, gdy zadanie prymalne ma rozwiązanie

Założmy, że dla v^* uzyskanego ze zrelaksowanego zadania mastera zadanie prymalne (67)-(68) ma rozwiązanie, któremu odpowiada optymalny wektor mnożników Lagrange'a λ^* . Mnożniki nowego ograniczenia typu (65) są tymi właśnie mnożnikami.

Tak więc do zrelaksowanego zadania mastera należy dodać ograniczenie typu

$$(69) \quad \min_{x \in X} L_o(x, v, \lambda^*) \leq \mu$$

Przypadek, gdy zadanie prymalne nie ma rozwiązania

Bardziej skomplikowana sytuacja występuje, gdy zadanie prymalne nie ma rozwiązania. Zgodnie z Twierdzeniem 2, dla ustalonego $v = v^*$ zadanie prymalne ma rozwiązanie, jeżeli spełniony jest następujący warunek

$$(70) \quad \min_{x \in X} [L_f(x, v^*, \lambda) = \lambda^T g(x, v^*)] \leq 0, \quad \forall \lambda \in \Lambda$$

gdzie

$$(71) \quad \Lambda = \left\{ \lambda \in \mathbb{R}^m : \lambda \geq 0, \sum_{j=1}^m \lambda_j = 1 \right\}$$

W przypadku wykrycia przez solver braku rozwiązania zadania prymalnego, należy znaleźć taki wektor $\hat{\lambda} \in \Lambda$, dla którego¹

$$(72) \quad \min_{x \in X} \hat{\lambda}^T g(x, v^*) > 0$$

Następnie przyjmuje się $\lambda^* = \hat{\lambda}$ i do zrelaksowanego zadania mastera dodaje się nierówność

$$(73) \quad \min_{x \in X} L_f(x, v, \lambda^*) \leq 0$$

Twierdzenie 4 W przypadku, gdy zadanie prymalne (67)-(68) jest niedopuszczalne, wektor mnożników Lagrange'a $\hat{\lambda} \in \Lambda$ spełniający warunek (72) można wyznaczyć rozwiązując zadanie pomocnicze (tzw. zadanie dopuszczalności):

$$(74) \quad \min_{x \in X} \max_{j=1, \dots, m} g_j(x, v^*)$$

Dowód:

Zadanie (74) jest równoważne zadaniu:

$$(75) \quad \min_{x \in X, \alpha} \alpha$$

p.o.

$$(76) \quad g_j(x, v^*) \leq \alpha, \quad j = 1, \dots, m$$

Ma ono na pewno rozwiązanie z uwagi na wcześniej przyjęte założenia co do funkcji $g(x, v)$.

Nie zmniejszając ogólności rozważań, przyjmijmy, że zbiór X definiują ograniczenia nierównościowe $r(x)$, czyli mamy zadanie:

$$(77) \quad \min_{x, \alpha} \alpha$$

p.o.

$$(78) \quad g_j(x, v^*) \leq \alpha, \quad j = 1, \dots, m$$

$$(79) \quad r_j(x) \leq 0, \quad j = 1, \dots, t$$

Zakładamy, że wszystkie funkcje ograniczeń g i r są różniczkowalne w sposób ciągły oraz wypukłe, a punkty ich aktywności są regularne. Lagranżian dla tego zadania ma postać:

$$(80) \quad L_{\alpha f}(x, \alpha, v^*, \lambda) = \alpha + \sum_{j=1}^m \lambda_j^g \cdot (g_j(x, v^*) - \alpha) +$$

¹ Współczesne solwery optymalizacyjne przeważnie same podają taki wektor mnożników Lagrange'a w sytuacji, gdy stwierdzają, że nie istnieje rozwiązanie dopuszczalne.

$$+ \sum_{j=1}^t \lambda_j^r \cdot r_j(x)$$

gdzie $\lambda^g \geq 0, \lambda^r \geq 0$ są wektorami mnożników Lagrange'a dla ograniczeń, odpowiednio, (78) i (79). Z warunków koniecznych optymalności Karusha-Kuhna-Tuckera w odniesieniu do zmiennej α otrzymamy równanie:

$$(81) \quad 1 - \sum_{j=1}^m \hat{\lambda}_j^g = 0$$

czyli, porównując z definicją zbioru Λ (71), $\lambda^g \in \Lambda$. Uwzględniając równość (81) we wzorze na lagranżian (80), w punkcie optymalnym, biorąc pod uwagę warunki komplementarności w odniesieniu do ograniczeń (79), otrzymamy:

$$(82) \quad L_{\alpha f}(x^*, \alpha^*, v^*, \hat{\lambda}) = \sum_{j=1}^m \hat{\lambda}_j^g \cdot g_j(x^*, v^*) = L_f(x^*, v^*, \hat{\lambda}^g)$$

A zatem wektor mnożników Lagrange'a dla ograniczeń (78) wyliczonych w punkcie optymalnym w zadaniu (77)-(79) jest szukanym wektorem $\hat{\lambda}$. ■

Podstawowe własności zadania mastera. Cięcia

Najlepiej by było, gdyby z rozwiązań zadania prymalnego dla różnych $v = v^*$ udało się wykorzystać nie tylko mnożniki Lagrange'a w punkcie optymalnym λ^* , ale także optymalne wektory zmiennych prymalnych x^* . Żeby takie podejście prowadziło do rozwiązania zadania podstawowego (4)-(7), muszą być spełnione pewne warunki.

Najważniejszy jest następujący:

Założenie 1 Obydwa lagranżiany $L_o(x, v, \lambda)$ oraz $L_f(x, v, \lambda)$ dla dowolnych $x \in X, v \in V$ i $\lambda \geq 0$ mogą być zapisane jako funkcje złożone:

$$(83) \quad L_o(x, v, \lambda) = Q_o(w_o(x, \lambda), v, \lambda)$$

$$(84) \quad L_f(x, v, \lambda) = Q_f(w_f(x, \lambda), v, \lambda)$$

gdzie w_o, w_f są skalarnymi funkcjami x i λ , a Q_o, Q_f są funkcjami rosnącymi względem pierwszego argumentu, wypukłymi względem drugiego.

Twierdzenie 5 Załóżmy, że dla zadania (4)-(7) spełnione jest Założenie 1. Niech x^* będzie optymalnym rozwiązaniem zadania

$$(85) \quad \min_{x \in X} L_o(x, v^*, \lambda^*)$$

Wówczas także dla $v \neq v^*$

$$(86) \quad \min_{x \in X} L_o(x, v, \lambda^*) = L_o(x^*, v, \lambda^*)$$

Dowód:

Zgodnie z Założeniem 1, ze względu na rosnącą zależność funkcji Lagrange'a od pierwszego argumentu, ze wzoru (83) mamy:

$$(87) \quad \begin{aligned} Q_o(w_o(x^*, \lambda^*), v^*, \lambda^*) &= L_o(x^*, v^*, \lambda^*) = \\ &= \min_{x \in X} L_o(x, v^*, \lambda^*) = \min_{x \in X} Q_o(w_o(x, \lambda^*), v^*, \lambda^*) = \\ &= Q_o(\min_{x \in X} w_o(x, \lambda^*), v^*, \lambda^*) \end{aligned}$$

Stąd, biorąc pod uwagę różnowartościowość funkcji ściśle monotonicznych:

$$(88) \quad w_o(x^*, \lambda^*) = \min_{x \in X} w_o(x, \lambda^*)$$

A zatem, dla $v \neq v^*$

$$\begin{aligned} \min_{x \in X} L_o(x, v, \lambda^*) &= \min_{x \in X} Q_o(w_o(x, \lambda^*), v, \lambda^*) = \\ &= Q_o(\min_{x \in X} w_o(x, \lambda^*), v, \lambda^*) = Q_o(w_o(x^*, \lambda^*), v, \lambda^*) = \\ (89) \quad &= L_o(x^*, v, \lambda^*) \end{aligned}$$

Jeśli obowiązuje Założenie 1, ograniczenie (65) w zadaniu мастера w okolicy $v = v^*$ można więc aproksymować za pomocą ograniczenia:

$$(90) \quad \mu \geq L_o(x^*, v, \lambda^*)$$

gdzie wektory x^*, λ^* pochodzą z rozwiązania zadania primalnego (67)-(68).

Ograniczenie to będziemy dalej nazywać cięciem.

Z Założenia 1 wynika jeszcze jedno udogodnienie. Biorąc pod uwagę wypukłość funkcji Lagrange'a L_o względem v , w przypadku, gdy jest ona różniczkowalna względem tego wektora², otrzymamy:

$$(91) \quad L_o(x^*, v, \lambda^*) \geq L_o(x^*, v^*, \lambda^*) + \frac{\partial L_o^T}{\partial v}(x^*, v^*, \lambda^*)(v - v^*)$$

Dzięki temu cięciu (90) może być zrelaksowane i zastąpione w zadaniu мастера cięciem liniowym

$$(92) \quad \mu \geq L_o(x^*, v^*, \lambda^*) + \frac{\partial L_o^T}{\partial v}(x^*, v^*, \lambda^*)(v - v^*)$$

Analogiczne rozumowanie można przeprowadzić dla ograniczenia dopuszczalności (66).

Łatwo sprawdzić, że Założenie 1 jest spełnione dla dwóch klas zadań:

1. Separowalnych - gdy funkcje f, g są sumami składników zależnych od x i v , czyli

$$(93) \quad f(x, v) = f_1(x) + f_2(v)$$

$$(94) \quad g(x, v) = g_1(x) + g_2(v)$$

przy czym składniki f_2, g_2 są wypukłe na V .

2. Zadań programowania ze zmiennymi czynnikami (ang. *variable factor programming*) o postaci:

$$(95) \quad \min_{x, v} \left[f(x, v) = \sum_{i=1}^q v_i f_i(x_i) \right]$$

p.o.

$$(96) \quad g(x, v) = \sum_{i=1}^q v_i x_i - c \leq 0$$

$$(97) \quad Av \leq b, v \geq 0, v \in \mathbb{R}^q$$

$$(98) \quad x_i \geq 0, x_i \in \mathbb{R}^m$$

²W przypadku, gdy funkcja Lagrange'a nie jest różniczkowalna, gradient można zastąpić subgradientem.

gdzie $v_i, i = 1, \dots, q$ są nieujemnymi skalarami, zaś $c, x_i, i = 1, \dots, q$ wektorami o wymiarze m . Szczególnym przypadkiem tego typu zadań jest zadanie biliniowe, które jest niewypukłe [1].

W przypadku, gdy zadanie (4)-(7) nie jest ani separowalne, ani ze zmiennymi czynnikami, ale wypukłe względem łącznego wektora zmiennych (x, v) , można także stosować cięcia (92).

Twierdzenie 6 Dla zadań wypukłych względem łącznego wektora zmiennych (x, v) oszacowanie (92) jest ważne.

Dowód:

Dla zadań wypukłych, z różniczkowalnymi funkcjami² mamy dla ustalonego $\lambda = \lambda^*$ oraz wszystkich x i v :

$$\begin{aligned} L_o(x, v, \lambda^*) &\geq L_o(x^*, v^*, \lambda^*) + \frac{\partial L_o^T}{\partial x}(x^*, v^*, \lambda^*)(x - x^*) + \\ (99) \quad &+ \frac{\partial L_o^T}{\partial v}(x^*, v^*, \lambda^*)(v - v^*) \end{aligned}$$

Nierówność ta będzie zachowana, gdy obliczymy obustronnie minimum względem x :

$$\begin{aligned} \min_{x \in X} L_o(x, v, \lambda^*) &\geq \min_{x \in X} \left\{ L_o(x^*, v^*, \lambda^*) + \right. \\ &+ \left. \frac{\partial L_o^T}{\partial x}(x^*, v^*, \lambda^*)(x - x^*) + \frac{\partial L_o^T}{\partial v}(x^*, v^*, \lambda^*)(v - v^*) \right\} = \\ &= L_o(x^*, v^*, \lambda^*) + \frac{\partial L_o^T}{\partial v}(x^*, v^*, \lambda^*)(v - v^*) + \\ (100) \quad &+ \min_{x \in X} \frac{\partial L_o^T}{\partial x}(x^*, v^*, \lambda^*)(x - x^*) \end{aligned}$$

Zauważmy, że ostatni składnik wyrażenia (100) jest równy zero, jako że punkt x^* jest optymalny dla $v = v^*, \lambda = \lambda^*$ (nie ma w nim dopuszczalnego kierunku poprawy, czyli o pochodnej kierunkowej ujemnej), a zatem nierówność (92) obowiązuje.

Algorytm obliczeniowy

Będziemy teraz rozważać przypadek, gdy w punkcie optymalnym zadania primalnego x^* w wersji zarówno dla $v^* \in V \cap V_0$, czyli (67)-(68), jak i $v^* \in V \setminus V_0$, czyli (75)-(76), istnieje wektor mnożników Lagrange'a λ^* . Zadania te muszą więc spełniać pewne warunki regularności (o czym poniżej).

Algorytm obliczeniowy oparty na uogólnionej metodzie Bendersa może być sformułowany następująco:

Algorytm GBD

1. Wybierz punkt początkowy $v^0 \in V$; $K_o = \emptyset, K_f := \emptyset, k := 0$, przyjmij tolerancję zbieżności $\varepsilon > 0$.
2. Dla ustalonego $v = v^k$ spróbuj rozwiązać zadanie primalne (67)-(68).
 - Jeżeli udało się $k := k + 1$, zapamiętaj jako x^k i λ^k uzyskany punkt optymalny oraz wektor optymalnych mnożników Lagrange'a, $K_o := K_o \cup \{k\}$. Uaktualnij $UBD = \min\{UBD, z(v^k)\}$. Jeżeli nastąpiło polepszenie oszacowania górnego UBD, to zapamiętaj jako najlepsze dotychczasowe rozwiązanie parę (x^k, v^k) . Jeżeli $UBD - LBD \leq \varepsilon$ to STOP.

- W przeciwnym razie rozwiąż zadanie dopuszczalności (75)-(76)³, $k := k + 1$, zapamiętaj jako x^k i λ^k uzyskany punkt optymalny oraz wektor optymalnych mnożników Lagrange'a odpowiadających ograniczeniom, $K_f := K_f \cup \{k\}$.

3. Rozwiąż zrelaksowane zadanie мастера: w wersji nieliniowej

$$(101) \quad \min_{v \in V, \mu} \mu$$

p.o.

$$(102) \quad L_o(x^l, v, \lambda^l) \leq \mu, \quad l \in K_o$$

$$(103) \quad L_f(x^l, v, \lambda^l) \leq 0, \quad l \in K_f$$

albo zlinearyzowanej

$$(104) \quad \min_{v \in V, \mu} \mu$$

p.o.

$$(105) \quad L_o(x^l, v^l, \lambda^l) + \frac{\partial L_o^T}{\partial v}(x^l, v^l, \lambda^l)(v - v^l) \leq \mu, \quad l \in K_o,$$

(106)

$$L_f(x^l, v^l, \lambda^l) + \frac{\partial L_f^T}{\partial v}(x^l, v^l, \lambda^l)(v - v^l) \leq 0, \quad l \in K_f$$

Niech (v^k, μ^k) będzie optymalnym rozwiązaniem powyższego zadania. Wtedy μ^k jest dolnym oszacowaniem oryginalnego zadania. Niech $LBD = \mu^k$. Jeżeli $UBD - LBD \leq \varepsilon$ to STOP.

4. Skocz do 2.

Zbieżność algorytmu

Zostało udowodnione, że algorytm GBD zbiega w skończonej liczbie kroków dla dowolnego $\varepsilon > 0$, gdy [10]:

1. V jest skończonym zbiorem dyskretnym oraz spełnione są założenia Twierdzenia 2 oraz 3 dla przypadku 1 (wtedy zbieżność zachodzi także dla $\varepsilon = 0$).
2. V jest niepustym, zwartym podzbiorem V_0 , X jest niepustym zwartym zbiorem wypukłym, funkcje f i g są wypukłe na X dla każdego ustalonego $v \in V$ oraz ciągłe na $X \times V$, zbiór optymalnych mnożników Lagrange'a w zadaniu (67)-(68) jest niepusty dla wszystkich $v \in V_0$, a ograniczenia spełniają warunek regularności Slatera: $\exists \bar{x} \in X, \bar{v} \in V \quad g(\bar{x}, \bar{v}) < 0^4$.
3. V nie jest podzbiorem V_0 , funkcja ograniczeń $g(x, v)$ jest separowalna, zbiór X zdefiniowany za pomocą ograniczeń liniowych; pozostałe warunki jak w p. 2. [9].

Niezwykle istotne jest założenie o istnieniu mnożników Lagrange'a dla wszystkich $v \in V \cap V_0$. Niesprawdzenie go może prowadzić do błędów. Rozważmy następujący przykład podany (w kontekście nieco innego podejścia) przez Gro-

theya i in. [13]⁵:

$$(107) \quad \min_{x_1, x_2, v} v^2 - x_2$$

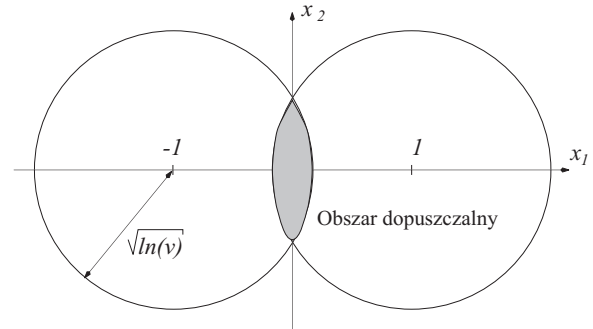
p.o.

$$(108) \quad (x_1 - 1)^2 + x_2^2 \leq \ln v$$

$$(109) \quad (x_1 + 1)^2 + x_2^2 \leq \ln v$$

$$(110) \quad v \geq 1$$

$$(111) \quad -10 \leq x_i \leq 10, \quad i = 1, 2$$



Rys. 1. Przekrój przez zbiór dopuszczalny w zadaniu (107)-(111) dla $v > e$.

Dla $v^* = e$ będziemy mieli zadanie:

$$(112) \quad \min_{x_1, x_2, v} e^2 - x_2$$

p.o.

$$(113) \quad (x_1 - 1)^2 + x_2^2 \leq 1$$

$$(114) \quad (x_1 + 1)^2 + x_2^2 \leq 1$$

$$(115) \quad -10 \leq x_i \leq 10, \quad i = 1, 2$$

Jego rozwiązaniem optymalnym jest jedyny punkt dopuszczalny $(0,0)$. Łatwo sprawdzić, że w tym punkcie mnożniki Lagrange' dla ograniczeń (113)-(114) nie istnieją, a zatem wyrażenia (102) i (105) tracą sens. Wynika to z tego, że w punkcie tym nie jest spełniony warunek regularności Fiacc McCormicka, by gradienty ograniczeń aktywnych były linowo niezależne. Co ciekawe, zadanie pierwotne (107)-(111), względem łącznego wektora (x, v) , jest wypukłe i regularne - warunek ten jest spełniony także w punkcie $(x, v) = (0, 0, e)$.

Algorytm GBD ma dość poważne wady. Najważniejszą z nich jest to, że liczba ograniczeń rośnie w nim z każdym wykonaniem zadania мастера, powodując, że zadanie się komplikuje, a co za tym idzie, maksymalny i średni, w pewnym oknie, czas wykonania tego zadania się wydłuża⁶. Ponadto, nie ma gwarancji, że w kolejnych krokach wartość przybliżenia rozwiązania dopuszczalnego UBD maleje. Dlatego warto zastosować bardziej wyrafinowane algorytmy.

⁵Dla kompletności sformułowania dodany został zbiór X , zdefiniowany jako kostka będąca iloczynem kartezjańskim przedziałów $[-10, 10]$ wzdłuż każdej współrzędnej. Nie ma to znaczenia dla wodu.

⁶Zawsze mogą się zdarzyć krótkie pojedyncze wykonania zadania мастера, np. jeśli kolejny punkt optymalny v^* leży blisko zadane-go początkowego, ale nie jest to rzecz typowa.

³O ile używany solver sam go nie rozwiązuje napotykając niedopuszczalność.

⁴W przypadku, gdy ograniczenia nierównościowe wynikają z zastosowania transformacji (9) lub (10) do ograniczeń równościowych (8), warunek Slatera będzie spełniony, gdy się je zrelaksuje stosując metodę paska epsilonowego.

Najbardziej przydatne, a także najczęściej używane obecnie, są algorytmy oparte na metodzie płaszczyzn tnących, zwłaszcza w powiązaniu z metodą punktu wewnętrznego w wersji centralnej [7].

Efektywne algorytmy rozwiązania zadania мастера w wersji z cięciami liniowymi

Klasyczna metoda płaszczyzn tnących (a ściślej odcinających) Kelley'a [15], do której się sprowadza zadanie мастера w wersji z cięciami zlinearyzowanymi, ma szereg wad. Przede wszystkim jest ona wolno zbieżna. Podawane są szacunki, że dla osiągnięcia rozwiązania z dokładnością $\varepsilon > 0$ metoda ta wymaga $O(\frac{1}{\varepsilon^{q+1}})$ iteracji [12]. Poważnym problemem, wspomnianym wyżej, jest też w którymś momencie liczba cięć, powodująca wydłużenie obliczeń w kolejnych iteracjach. Nie ma, niestety, niezawodnych i prostych reguł na usuwanie starych cięć, nawet tych, które są nieaktywne dla bieżącego rozwiązania zadania (104)-(106) [23]. Można też zaobserwować pewnego rodzaju niestabilność, polegającą na tym, że kolejny punkt wygenerowany przez algorytm może być daleko od poprzedniego, mimo że poprzedni był już blisko optimum, a nawet w nim [3].

Wady te zostały w poważnym stopniu usunięte w modyfikacjach metody Kelleya [7]. Najpopularniejsze i najbardziej przydatne z nich to:

1. Metoda wiązki [18], [25], [16].

Polega ona w najprostszej wersji na dodaniu do aproksymacji liniowej wskaźnika jakości w zadaniu мастера (uzyskanej na podstawie cięć optymalności) kwadratowego składnika proksymalnego, będącego karą za odchylenie od ostatniego istotnego rozwiązania w^k . Słowo "istotne" oznacza to, że przyniosło ono wystarczającą poprawę wyjściowego wskaźnika jakości. Jeżeli to nie miało miejsca, dodawane jest jedynie nowe cięcie. Zadanie (104)-(106) przyjmie zatem postać:

$$(116) \quad \min_{v \in V, \mu} \mu + \frac{1}{2\beta_k} \|v - w^k\|^2$$

p.o.

$$(117) \quad \varphi(v^l) + \nabla\varphi(v^l)^T(v - v^l) \leq \mu, \quad l \in K_o$$

$$(118) \quad \xi(v^l) + \nabla\xi(v^l)^T(v - v^l) \leq 0, \quad l \in K_f$$

gdzie

$$(119) \quad \varphi(v^l) = L_o(x^l, v^l, \lambda^l) = f(x^l, v^l) + (\lambda^l)^T g(x^l, v^l)$$

$$\nabla\varphi(v^l) = \frac{\partial L_o}{\partial v}(x^l, v^l, \lambda^l) =$$

$$(120) \quad = \frac{\partial f}{\partial v}(x^l, v^l) + \frac{\partial g}{\partial v}(x^l, v^l)\lambda^l$$

$$(121) \quad \xi(v^l) = L_f(x^l, v^l, \lambda^l) = (\lambda^l)^T g(x^l, v^l)$$

$$\nabla\xi(v^l) = \frac{\partial L_f}{\partial v}(x^l, v^l, \lambda^l) =$$

$$(122) \quad = \frac{\partial g}{\partial v}(x^l, v^l)\lambda^l$$

Zadanie мастера jest zatem kwadratowe, dość łatwe do rozwiązania za pomocą standardowych solverów.

Zaletą tej metody jest łatwość modyfikacji zbioru cięć (kandydatami do usunięcia są te, dla których mnożniki Lagrange'a są zerowe) i utrzymywania jego liczności na niskim poziomie [23]. Szacuje się, że dla osiągnięcia rozwiązania z dokładnością $\varepsilon > 0$ metoda wiązki wymaga $O(\frac{1}{\varepsilon^2})$ iteracji [12], co oznacza istotną poprawę efektywności w porównaniu z klasyczną metodą Kelleya.

2. Metoda największej kuli wpisanej [4]

W metodzie tej korzysta się z pojęcia zbioru lokalizacyjnego. Dla danego, najlepszego dotychczas (powiedzmy do k -tej iteracji zadania мастера włącznie), oszacowania górnego optymalnej wartości wskaźnika jakości zadania wyjściowego - oznaczymy je jako UBD^k - będzie to zbiór:

$$(123) \quad L_k = \left\{ \begin{array}{l} (v, \mu) \in \mathbb{R}^q \times \mathbb{R} \\ \mu \leq UBD^k \\ \varphi(v^l) + \nabla\varphi(v^l)^T(v - v^l) \leq \mu, \quad l \in K_o \\ \xi(v^l) + \nabla\xi(v^l)^T(v - v^l) \leq 0, \quad l \in K_f \end{array} \right\}$$

Zauważmy, że przyjmując $u = (v, \mu)$ można przedstawić ograniczenia opisujące ten zbiór w sposób następujący:

$$(124) \quad Au \leq b$$

Metoda największej kuli wpisanej jest jedną z metod punktu centralnego (z większej rodziny metod punktu wewnętrznego), które polegają na tym, że poszukuje się w niej w kolejnych krokach k , zamiast punktu minimalizującego wskaźnik (104), punktu, który według pewnej miary leży "w środku" zbioru lokalizacyjnego L_k (123). Środek największej kuli wpisanej dla zbioru $U = \{u \in \mathbb{R}^{q+1} : Au \leq b\}$ (zwany także centrum Czebyszewa), gdzie macierz A ma wymiary $S \times (q+1)$, przy $S = (\bar{K}_o + \bar{K}_f + 1)$, wyznacza się rozwiązując zadanie programowania liniowego:

$$(125) \quad \max_{u \in \mathbb{R}^{q+1}, \sigma \geq 0} \sigma$$

p.o.

$$(126) \quad a_i^T u + \|a_i^T\| \sigma \leq b_i, \quad i = 1, \dots, S$$

Zaletą tej metody jest, poza liniowym sformułowaniem, łatwość eliminacji niepotrzebnych (odległych) cięć na podstawie zerowych wartości mnożników Lagrange'a (nieaktywnych ograniczeń (126)), o ile w kolejnych iteracjach zadania мастера wartość $\hat{\sigma}$ zmalała [4]. Metoda ta potrzebuje $O((q+1) \ln \frac{1}{\varepsilon})$ iteracji dla osiągnięcia rozwiązania optymalnego z dokładnością ε [12].

3. Metoda najmniejszej elipsoidy opisanej [24], [21], [12].

W każdym kroku otrzymujemy elipsoidę

$$(127) \quad E_k = \{u | (u - u^k)^T W_k^{-1} (u - u^k) \leq 1\}$$

Jest ona scharakteryzowana przez dwa parametry: macierz W_k i środek u^k .

Zakładamy, że startujemy od elipsoidy E_0 zawierającej środek zbioru dopuszczalnego U . Kolejne elipsoidy E_k są takie, że E_{k+1} jest elipsoidą o najmniejszej objętości zawierającą $E_k \cap \{u | \langle a_k, u - u^k \rangle \leq 0\}$, gdzie a_k - k -ty wiersz macierzy A we wzorze (124). Jest ona zdefiniowana jako:

$$(128) \quad u^{k+1} = u^k - \frac{1}{n_u + 1} \frac{W_k a_k}{\sqrt{a_k^T W_k a_k}}$$

(129)

$$W_{k+1} = \frac{n_u^2}{n_u^2 - 1} \left(W_k - \frac{2}{n_u + 1} \frac{W_k a_k a_k^T W_k}{a_k^T W_k a_k} \right)$$

gdzie n_u jest wymiarem wektora u - u nas $n_u = q + 1$. Można pokazać [12], że objętość E_{k+1} równa jest objętości E_k razy współczynnik $(1 - 1/(n_u + 1)^2)$.

Wielką zaletą tej metody jest brak pamięci cięć, dzięki czemu każda iteracja jest bardzo szybka i nie wymaga użycia żadnego solwera na poziomie mastera. Wadą jest nieco wolniejsza zbieżność niż metod przedstawionych powyżej - jej złożoność wynosi $O((q + 1)^2 \ln \frac{1}{\epsilon})$ [12].

Wnioski

Metoda Bendersa, mimo że zaproponowana została już ponad pół wieku temu, jest wciąż bardzo żywotna. Jest ona często stosowana, zwłaszcza w wersji uogólnionej przez Geoffriona, do rozwiązywania wielu dużych zadań praktycznych z dziedziny techniki, przede wszystkim związanych z wszelkiego rodzaju sieciami, np. telekomunikacyjnymi, energetycznymi, gazowymi, transportowymi. Zadania te charakteryzują się mieszanym, dyskretno-ciągłym, charakterem zmiennych decyzyjnych, któremu często towarzyszy niewypukłość funkcji celu oraz ograniczeń. Dla dość szerokiej klasy zadań tego typu, gdy przy ustalonym jednym podwektorze zmiennych zadanie jest wypukłe względem drugiego podwektora i vice-versa (takie są na przykład zadania biliniowe), metoda Bendersa zapewnia zbieżność w skończonej liczbie kroków do dokładnego rozwiązania. Metoda ta radzi sobie z niedopuszczalnością zadań prymalnych - względem wektora x - dla pewnych wartości zmiennych komplikujących v , rozwiązując zadanie dopuszczalności, z którego otrzymywane są nowe ograniczenia - cięcia dopuszczalności. W przypadku, gdy zadanie prymalne jest dopuszczalne, generowane jest cięcie optymalności, eliminujące część zbioru dopuszczalnego zmiennych komplikujących V , dla którego wskaźnik jakości ma gorsze wartości niż uzyskane dotąd.

Występują dwie wersje metody Bendersa: z cięciami nieliniowymi albo liniowymi. I jedno, i drugie można stosować, gdy spełniony jest warunek niezależności rozwiązania zadania minimalizacji lagranżianu w zadaniu prymalnym od wektora zmiennych komplikujących, a także, gdy lagranżian jest wypukły względem zmiennych komplikujących. Cięcia liniowe można stosować również w zadaniach wypukłych względem łącznego wektora (x, v) .

Metoda z cięciami liniowymi (będąca w istocie wersją metody płaszczyzn tnących Kelleya) wydaje się być bardziej praktyczna, ze względu na łatwość rozwiązania zadania mastera, np. stosując solwery liniowe lub kwadratowe. Ma ona wiele zalet, zwłaszcza gdy używa się stosunkowo nowych metod pochodzących z obszaru optymalizacji nieróżniczkowalnej:

- wiązki lub największej kuli wpisanej - łatwo wtedy usunąć zbyteczne cięcia,
- najmniejszej elipsoidy opisanej - w ogóle nie ma takiej potrzeby.

LITERATURA

- [1] Adams W. P., Sherali H. D.: Mixed-integer bilinear programming problems, *Mathematical Programming*, 59(3), str. 279–305, 1993.
- [2] Benders J. F.: Partitioning Procedures for Solving Mixed-Variables Programming Problems, *Numerische Mathematik*, 4(1), str. 238–252, 1962.
- [3] Bertsekas D.P.: *Nonlinear Programming*, 2nd Edition, Athena Scientific, 1999.

- [4] Elzinga J., Moore T. G.: A central cutting plane algorithm for the convex programming problem, *Mathematical Programming*, 8(1), str. 134–145, 1975.
- [5] Cai X., McKinney D.C., Lasdon L.S., Watkins Jr. D.W.: Solving Large Nonconvex Water Resources Management Models Using Generalized Benders Decomposition, *Operations Research*, 49(2), str. 235–245, 2001.
- [6] Chung K.H., Kim B.H., Hur D.: Distributed implementation of generation scheduling algorithm on interconnected power systems, *Energy Conversion and Management*, 52(12), str. 3457–3464, 2011.
- [7] Elhedhli S., Goffin J-L., Vial J.-P.: Nondifferentiable optimization: Cutting plane methods, *Encyclopedia of Optimization*, 2nd Edition (C.A. Floudas and P.M. Pardalos, Eds.), str. 2590–2595, Springer, 2009.
- [8] Floudas C.A.: *Nonlinear and Mixed-Integer Optimization*, Oxford University Press, 1995.
- [9] Floudas C.A.: Generalized Benders Decomposition GBD, *Encyclopedia of Optimization*, 2nd Edition (C.A. Floudas and P.M. Pardalos, Eds.), str. 1162–1175, Springer, 2009.
- [10] Geoffrion A. M.: Generalized Benders Decomposition: *Journal of Optimization Theory and Applications*, 10(4), str. 237–260, 1972.
- [11] Geoffrion A. M., Graves G.W.: Multicommodity Distribution System Design by Benders Decomposition, *Management Science*, 20(5), str. 822–844, 1974.
- [12] Goffin J-L., Vial J.-P.: Convex nondifferentiable optimization: A survey focused on the analytic center cutting plane method, *Optimization Methods and Software*, 17(5), str. 805–867, 2002.
- [13] Grothey A., Leyffer S., McKinnon K. I. M.: A Note on Feasibility in Benders Decomposition, *Numerical Analysis Report NA/188*, Dundee University, 1999.
- [14] Haouari M., Mrad M., Sherali H.D.: Optimum synthesis of discrete capacitated networks with multi-terminal commodity flow requirements, *Optimization Letters*, 1(4), str. 341–354, 2007.
- [15] Kelley J. E.: The Cutting-Plane Method for Solving Convex Programs, *Journal of the Soc. Indust. Appl. Math.*, 8(4), str. 703–712, 1960.
- [16] Kiwiel K.C.: *Methods of Descent for Nondifferentiable Optimization*, Springer, 1985.
- [17] Lee J., Leyffer S.: *Mixed Integer Nonlinear Programming*, Springer, 2012.
- [18] Lemaréchal C.: An extension of Davidon methods to nondifferentiable problems, *Mathematical Programming Study 3* (M.L. Balinski and P. Wolfe, Eds.), str. 95–109, North-Holland, 1975.
- [19] Li D., Sun X.: *Nonlinear Integer Programming*, Springer, 2006.
- [20] Li X.: Parallel nonconvex generalized Benders decomposition for natural gas production network planning under uncertainty, *Computers and Chemical Engineering*, 55(8), str. 97–108, 2013.
- [21] Nemirovski A., Yudin D.: Problem Complexity and Method Efficiency in Optimization, *J.Wiley & Sons*, 1983.
- [22] Osman H., Demirli K.: A bilinear goal programming model and a modified Benders decomposition algorithm for supply chain reconfiguration and supplier selection, *Int. J. Production Economics*, 124(1), str. 97–105, 2010.
- [23] Ruszczyński A.: *Nonlinear Optimization*, Princeton University Press, 2006.
- [24] Shor N.Z.: *Minimization methods for non-differentiable functions*, Springer, Berlin, 1985.
- [25] Wolfe P.: A method of conjugate subgradients for minimizing nondifferentiable functions, *Mathematical Programming Study 3* (M.L. Balinski and P. Wolfe, Eds.), str. 145–173, North-Holland, 1975.

Autor: dr hab. inż. Andrzej Karbowski, Politechnika Warszawska, Wydział Elektroniki i Technik Informacyjnych, Instytut Automatyki i Informatyki Stosowanej, ul. Nowowiejska 15/19, 00-665 Warszawa; NASK (Naukowa i Akademicka Sieć Komputerowa), ul. Wąwozowa 18, 02-796 Warszawa; E-mail: A.Karbowski@elka.pw.edu.pl