Mariusz Kępa

RÓWNOWAŻNOŚĆ WARUNKÓW KUHNA-TUCKERA I OGRANICZEŃ DUALNYCH ZADANIA PROGRAMOWANIA GEOMETRYCZNEGO

1. Wprowadzenie

W artykule wykazuje się, że jeśli rozwiązanie prymalne zadania programowania geometrycznego spełnia warunki Kuhna-Tuckera, to istnieje związane rozwiązanie dualne spełniające warunki ortogonalności, normalizacji i nieujemności. Twierdzenie to stanowi uzupełnienie pierwszego twierdzenia o dualności programowania geometrycznego, które podaje taką zależność tylko dla prymalnego rozwiązania optymalnego.

W dalszej części pokazuje się jak można wykorzystać warunki ortogonalności dla sformułowania wystarczających warunków optymalności dla zadania dualnego. Pokazuje się również, że pozwala to na nieiteracyjne rozwiązanie szerszej klasy zadań niż zadania o zerowym stopniu trudności, dla których rozwiązania optymalne wyznaczają bezpośrednio ograniczenia dualne.

2. Wybrane pojęcia programowania geometrycznego

Zadanie programowania geometrycznego można sformułować następująco (porównaj 5):

Zadanie prymalne A. Znaleźć minimum posynomianu $g_0(t)$ przy ograniczeniach

$$g_k(t) \leqq 1, \quad k=1,2,\ldots,p$$

$$t_j > 0, \quad k=1,2,\ldots,m$$

gdzie

$$g_k(t) = \sum_{i=1}^{n_k} c_{ki} \prod_{j=1}^{m} t_j^{a_{kij}}, \quad k=0,1,...,p.$$

Symbol n_k oznacza ilość składników posynomianu g_k, współczynniki c_{ki} są dodatnie, a wykładniki a_{kij} są dowolnymi liczbami rzeczywistymi.

Zadanie dualne B. Znaleźć maksimum funkcji dualnej d(y)

$$d(y) = \prod_{k=0}^{p} \prod_{i=1}^{n_k} (c_{ki} y_{ki}^{-1} L_k)^{y_{ki}}$$

gdzie

$$L_k = \sum_{i=1}^{n_k} y_{ki}, \quad k=0,1,...,p.$$

przy ograniczeniach złożonych z warunku normalizacji

$$L_0 = 1$$

warunków ortogonalności

$$\sum_{k=0}^{p} \sum_{i=1}^{n_k} y_{ki} a_{kij} = 0, \quad j=1,2,...,m.$$

i warunków nieujemności

$$y_{ki} \geqq 0.$$

Przystąpimy teraz do zdefiniowania rozwiązań związanych.

Definicja 1.

Rozwiązania prymalne t i dualne y nazywamy rozwiązaniami jeśli

$$y_{ki} = L_k c_{ki} \prod_{j=1}^{m} t_j^{a_{kij}} g_k^{-1}, \quad \begin{matrix} k=0,1,...,p \\ i=1,...,n_k. \end{matrix}$$

Mnożniki L_k są zdefiniowane tak jak w zadaniu dualnym B.

Jednym z podstawowych twierdzeń programowania geometrycznego jest pierwsze twierdzenie o dualności. Dowód twierdzenia znajduje się w pracy 2.

Twierdzenie 1.

Jeśli istnieje rozwiązanie dopuszczalne zadania prymalnego A spełniające warunki $g_k < 1$, k=1,2,...,p i zadanie posiada rozwiązanie optymalne, to:

1) Zadanie dualne B posiada rozwiązanie optymalne.
2) Wartość dualnej funkcji kryterium dla rozwiązania optymalnego równa się wartości prymalnej funkcji kryterium.

3) Dla rozwiązania optymalnego t' zadania prymalnego istnieją nieujemne mnożniki Lagrange'a u_k', $k=1,2,...,p$ takie, że funkcja Lagrange'a

$$L(t,u) = g_0(t) + \sum_{k=1}^{p} u_k [g_k(t) - 1]$$

spełnia zależności

$$L(t',u) \leqq g_0(t') = L(t',u') \leqq L(t,u')$$

dla dowolnych $t_j \geqq 0$ i dowolnych $u_k \geqq 0$. Oprócz tego istnieje optymalne rozwązanie y' zadania dualnego B o składowych

$$y_{ki} = \begin{cases} \dfrac{c_{ki}\, t_1^{a_{ki1}} \dots t_m^{a_{kim}}}{g_0(t)}, & k=0 \\ & i=1,\dots,n_0 \\ \dfrac{u_k\, c_{ki}\, t_1^{a_{ki1}} \dots t_m^{a_{kim}}}{g_0(t)}, & k=1,\dots,p \\ & i=1,\dots,n_k \end{cases} \tag{2}$$

gdzie $t=t'$ i $u=u'$. Poza tym

$$L_k = \frac{u_k}{g_0(t')}, \quad k=1,\dots,p.$$

4) Rozwiązanie optymalne t' zadania prymalnego spełnia układ równań

$$c_{ki}\, t_1^{a_{ki1}} \dots t_m^{a_{kim}} = \begin{cases} y_{ki}\, d(y), & k=0, \quad i=1,\dots,n_0 . \\ y_{ki}\, L_k^{-1}, & k=1,\dots,p \\ & i=1,\dots,n_k . \end{cases}$$

W następnym punkcie postaramy się udowodnić twierdzenie uzupełniające pierwsze twierdzenie o dualności.

3. Warunki Kuhna-Tuckera, a ograniczenia dualne programowania geometrycznego

Sformułujemy teraz następujące twierdzenie

Twierdzenie 2.
Jeśli rozwiązanie prymalne zadania programowania geometrycznego spełnia warunki Kuhna-Tuckera, to istnieje związane rozwiązanie dualne spełniające warunki ortogonalności, normalizacji i nieujemności.

Dowód

Spełnienie przez rozwiązanie zadania A warunków Kuhna-Tuckera oznacza, że $\exists\, U \geqq 0 : \nabla_t L(t,u) = 0$ i $[g_k(t) -]\, u_k = 0$, k=1,2,...,p. Wykażemy najpierw, że z warunku $\nabla_t L(t,u) = 0$ wynika, że

$$L_k g_0 = u_k , \quad k=1,2,...,p. \tag{3}$$

Ponieważ

$$L(t,u) = g_0(t) + \sum_{k=1}^{p} u_k [g_k - 1]$$

zatem

$$\nabla_t L(t,u) = \nabla_t g_0(t) + \sum_{k=1}^{p} u_k \nabla_t g_k(t) .$$

Z definicji g_k podanej w zadaniu prymalnym A i założeniu (3) wynika, że

$$L'(t_j) = \sum_{i=1}^{n_k} a_{0ij} c_{0i} \prod_{j=1}^{m} t_j^{a_{0ij}} + \\ + \sum_{k=1}^{p} L_k g_0 \sum_{k=1}^{n_k} a_{kij} c_{ki} \sum_{j=1}^{m} t_j^{a_{kij}} t_j^{-1} \tag{4}$$

dla j=1,...,m. $L'(t_j)$ są pochodnymi cząstkowymi funkcji Lagrange'a L(t,u). Jeśli y jest rozwiązaniem związanym spełniającym warunki def. 1, to z (4) wynika

$$L'(t_j) = g_0 \sum_{i=1}^{n_k} a_{0ij} y_{0i} t_j^{-1} + \sum_{k=1}^{p} g_0 \sum_{i=1}^{n_k} a_{kij} y_{ki} t_j^{-1} \tag{5}$$

dla j=1,2,...,m.

Ponieważ $\nabla_t L(t,u) = 0$, pochodne cząstkowe $L'(t_j)$ równają się zeru. Po pomnożeniu równania (5) przez $g_0^{-1} t_j$ otrzymujemy

$$\sum_{i=1}^{n_k} a_{0ij} y_{ki} + \sum_{k=1}^{p} \sum_{i=1}^{n_k} a_{kij} y_{ki} = 0, \quad j=1,...,m .$$

co stanowi zapis równoważny warunkom ortogonalności zadania B.

Spełnienie warunku normalizacji wynika z definicji rozwiązań, związanych. Wykażemy teraz, że istnieje nieujemne rozwiązanie związane. Jeśli $u_k > 0$ to $L_k > 0$ i wtedy, wobec dodatności posynomianów g_k, stałych c_{ki} i zmiennych t_j, z definicji 1 wynika, że zmienne y_{ki} są dodatnie. Jeśli $u_k = 0$ to $L_k = 0$ i wtedy, również zgodnie z definicją 1, $y_{ki} = 0$, k=1,...,p, i=1,...,n_k. Zatem zmienne y_{ki} są nieujemne.

Twierdzenie 3.
Jeśli istnieje rozwiązanie prymalne t zadania A związane z rozwiązaniem dopuszczalnym zadania dualnego B, to rozwiązanie t spełnia warunki Kuhna-Tuckera.

Dowód twierdzenia 3 można uzyskać dokonując w odwrotnej kolejności przekształceń przeprowadzonych w dowodzie twierdzenia 2. Założenie o istnieniu prymalnego rozwiązania związanego jest potrzebne dla prawdziwości implikacji: Jeśli (5) to (4).

W następnym punkcie zademonstrujemy przydatność warunków ortogonalności dla formułowania warunków optymalności i rozwiązywania zadań nieliniowych.

4. Warunki optymalności i metody rozwiązywania zadań

Wystarczającym warunkiem optymalności rozwiązań zadań A i B jest równość prymalnej i dualnej funkcji kryterium. Jest to warunek przydatny dla algorytmów iteracyjnych gdyż pozwala na podstawie różnicy kryteriów ocenić jak daleko jesteśmy od rozwiązania optymalnego. Algorytmy iteracyjne pozostawimy jako przedmiot odrębnego opracowania. Poniżej pokażemy, że dla pewnej klasy zadań można w znacznie łatwiejszy sposób sformułować wystarczające warunki optymalności. Wiadomo, że dla tzw. zadań o zerowym stopniu trudności wystarczające warunki optymalności stanowią już ograniczenia dualnego zadania. Możliwości nieiteracyjnego uzyskania rozwiązań nie ograniczają się do zadań tej klasy. Pokażemy to na przykładzie dwóch zadań rozwiązanych w podanej literaturze w sposób iteracyjny.
Zadanie 1.

$$g_0 = 40\, t_1^{-1} t_2^{-1/2} t_3^{-1} + 20\, t_1 t_3 + 20\, t_1 t_2 t_3 \to \min$$

$$g_1 = \frac{1}{3} t_1^{-2} t_2^{-2} + \frac{4}{3} t_2^{1/2} t_3^{-1} \leqq 1$$

$$t_j > 0, \quad i=1,\ldots,5.$$

Wykorzystując warunki ortogonalności do wyrażenia części zmiennych jako funkcje pozostałych można to zadanie zredukować do zadania z jednym ograniczeniem.

$$d(y) = \left(\frac{40}{y_1}\right)^{y_1} \left(\frac{20}{y_2}\right)^{y_2} \left(\frac{20}{y_2}\right)^{y_2} 1^{-0,5y_1+y_2}\, 2^{-y_1+2y_2} \to \max$$

$$y_1 + 2y_2 = 1$$

$$y_1, y_2 \geqq 0.$$

Obliczmy pochodne cząstkowe kryterium

$$d'(y_1) = d(y)\,[\ln 40 - \ln y_1 - 1 - 0{,}5 \ln 1 - \ln 2]$$

$$d'(y_2) = d(y)\,[2 \ln 20 - 2 \ln y_2 - 2 - \ln 1 + 2 \ln 2]$$

po uproszczeniu

$$d'(y_1) = d(y)\left[\ln\left(\frac{20}{y_1}\right) - d(y)\right]$$

$$d'(y_2) = d(y)\left[\ln\left(\frac{1600}{y_2^2}\right) - 2d(y)\right]$$

Rozwiązanie optymalne jest nieujemnym rozwiązaniem układu równań

$$2d'(y_1) = d'(y_2)$$

$$y_1 + 2y_2 = 1$$

Rozwiązaniem tym jest $y_1 = 0{,}2$, $x_2 = 0{,}4$. Pozostałe zmienne dualne przyjmują wartości $y_3 = y_2 = 0{,}4$, $y_4 = -0{,}5y_1 + y_2 = 0{,}3$, $y_5 = y_1 + y_2 = 0{,}6$. Wartość dualnej i prymalnej funkcji kryterium obliczymy podstawiając wartości zmiennych dualnych do funkcji d(y).

$$d(y) = \left(\frac{40}{0{,}2}\right)^{0{,}2}\left(\frac{20}{0{,}4}\right)^{0{,}4}\left(\frac{20}{0{,}4}\right)^{0{,}4} 1^{0{,}3}\, 2^{0{,}6} = 100$$

Przy pomocy zależności podanych w pierwszym twierdzeniu o dualności znajdujemy rozwiązanie prymalne t=(1,1,2).

Kolejnym przykładem niech będzie zadanie 2

$$g_0 = t_1^{-1} t_1^{-1} + 3t_2 \to \min$$

$$3t_1 + 16t_1 t_2 \leqq 1$$

$$t_j > 0, \quad j=1,2.$$

Zadanie 2 pochodzi z pracy [6], gdzie iteracyjnie oszacowano jego rozwiązanie. Pokażemy, że można je rozwiązać nieiteracyjnie. Zredukowane zadanie dualne ma postać

$$d(y) = (1/y_1)^{y_1}\,(3/y_2)^{y_2}\,(3/y_2)^{y_2}\,[16/(y_1-y_2)^{y_1-y_2}\; y_1^{y_1} \to \max$$

$$y_1 + y_2 = 1$$

$$y_1, y_2 \geqq 0.$$

Warunki optymalności $d'(y_1) = d'(y_2)$, $y_1 + y_2 = 1$ prowadzą do układu równań liniowych $16y_2 = 3(y_1 - y_2)$, $y_1 + y_2 = 1$. Rozwiązaniem jest zatem $y_1 = 19/22$, $y_2 = 3/22$. Wartość funkcji kryterium wynosi 22, a rozwiązaniem prymalnym jest $t_1 = 1/19$, $t_2 = 1$.

Uzasadnimy teraz metodę, którą posłużyliśmy się rozwiązując przykładowe zadania. Ograniczenia zadania dualnegoj tworzą układ m+1 równań liniowych z n niewiadomymi, gdzie m jest liczbą zmiennych prymalnych, a $n = n_0 + n_1 + \ldots n_p$ jest liczbą zmiennych dualnych, gdy $n_0 \geqq n-m$, tj. gdy liczba składników prymalnego kryterium jest większa lub równa różnicy liczby zmiennych dualnych jako funkcje pozostałych, można zadanie dualne zredukować do zadania z jednym ograniczeniem.

Dla funkcji dualnej udowodniono w pracy [2] następujące

Twierdzenie 4.

Logarytm funkcji dualnej zadania B, ln d(y) jest funkcją wklęsłą.

Funkcja logarytmiczna jest funkcją ściśle rosnącą. Zatem kryteria maksymalizacji funkcji dualnej i jej logarytmu są równoważne. Niech r(y) = 1 oznacza ograniczenie zredukowanego zadania dualnego. Dla zadania

$$\ln d(y) \to \max$$

$$r(y) = 1,$$

$$y = 0,$$

gdzie y jest wektorem zmiennych zredukowanego zadania dualnego, wystarczające warunki optymalności mają postać

$$r(y) = 1, \quad y = 0 \tag{6}$$

$$\exists\, c: \nabla_y \ln d(y) = c\, \nabla_y r(y), \tag{7}$$

gdzie c jest liczbą rzeczywistą. Ponieważ funkcje ln d(y) i d(y) osiągają maksimum w tym samym punkcie są to również wystarczające warunki optymalności dla zredukowanego zadania dualnego. Ponieważ

$$\nabla_y \ln d(y) = 1/d(y)\, \nabla_y d(y)$$

warunek (7) można zastąpić warunkiem

$$\exists\, a: \nabla_y (y) = a\, \nabla_y r(y) \tag{8}$$

gdzie a jest liczbą rzeczywistą. Warunki (8) i (6) wykorzystaliśmy dla nieiteracyjnego rozwiązania zadań 1 i 2.

5. Zakończenie

W artykule wykazano, przy pomocy pojęcia rozwiązań związanych, zależności pomiędzy warunkami Kuhna-Tuckera, a ograniczeniami dualnymi w programowaniu geometrycznym. Pokazano również dla jakiej klasy zadań można dualne zadanie geometryczne zredukować do zadania z jednym ograniczeniem i jakie korzyści obliczeniowe to przynosi. Dla zadań zredukowanych, korzystając z wklęsłości zlogarytmowanej funkcji dualnej, sformułowano wystarczające warunki optymalności.

Literatura

[1] Advances in Geometric Programming, red. M. Avriel, Plenum Press, New York 1980.

[2] Duffin R.J., Peterson E.L., Zener C.M., Geometric Programming, John Wiley and Sons, New York 1967.

[3] Duffin R.J., Peterson E.L., Geometric Programming with Signomials, Journal of Optimization Theory and Applications, Vol. 1, 1973.

[4] Hanson M.A., On Sufficiency of the Kuhn-Tucker Conditions, Journal of Mathematical Analysis and Applications, Vol. 80, 1981.

[5] Kępa M., Zadanie programowania geometrycznego a ograniczeniami równościowymi, Prace Instytutu Cybernetyki Ekonomicznej, AE Poznań 1984.

[6] Programowanie matematyczne, red. W. Bukietyński, AE Wrocław 1976.

[7] Wilde D.J., Beightler C.S., Foundations of Optimization, Prentice Hall, Englewood Cliffs, New Jersey 1967.

[8] Zangwill W.I., Programowanie nieliniowe, WNT, Warszawa 1974.