

Elementy Modelowania Matematycznego

Wykład 7

Programowanie nieliniowe i całkowitoliczbowe

Romuald Kotowski

Katedra Informatyki Stosowanej

PJWSTK 2009

Spis treści

- 1 Wstęp
- 2 Programowanie nieliniowe
- 3 Modelowanie całkowiłoczbowe

Spis treści

- 1 Wstęp
- 2 Programowanie nieliniowe
- 3 Modelowanie całkownicze

Spis treści

- 1 Wstęp
- 2 Programowanie nieliniowe
- 3 Modelowanie całkowitoliczbowe

Wstęp

Przypomnijmy: Programowanie matematyczne to problem optymalizacyjny. Zadanie brzmi:
Dokonaj maksymalizacji (minimalizacji) funkcji $f(x)$ przy nałożonych warunkach:

$$g(x) \leq 0, (\geq 0)$$

$$h(x) = 0$$

gdzie $x \in X \subset \mathbb{R}^n$, zaś f , g i h są funkcjami zdefiniowanymi na tym podzbiore.

Nałożone warunki nazywane są warunkami ograniczającymi, natomiast funkcja f to funkcja celu. Rozwiązania tego problemu nazywamy rozwiązaniami optymalnymi.

Wstęp

Nie istnieje jeden efektywny algorytm rozwiązania problemu programowania matematycznego, dlatego problemy należące do różnych klas rozwiązywane są różnymi metodami. Oto najważniejsze z nich:

- programowanie liniowe
- programowanie całkownicze
- programowanie zero-jedynkowe
- programowanie celowe
- programowanie kwadratowe
- programowanie nieliniowe
- programowanie dynamiczne
- programowanie sieciowe

Programowanie matematyczne znalazło szerokie zastosowanie w teorii decyzji, np. przy optymalizacji struktury kosztów produkcji.

Wstęp

Przykład: Do produkcji opakowań potrzebny jest karton i folia aluminiowa, przy czym dostępne są dwie metody produkcji (A i B). W metodzie A zużywamy 0,5 jednostki kartonu i 0,45 jednostki folii. W metodzie B zużywamy odpowiednio 0,6 i 0,5 jednostek produktów. Maksymalna dzienna produkcja jedną i drugą metodą wynosi 200 opakowań. Opakowanie wyprodukowane metodą A przynosi nam zysk w wysokości 1,5 zł, zaś metodą B 1,8 zł. Jednocześnie jesteśmy w stanie dostarczyć dziennie do fabryki 200 jednostek kartonu i 300 jednostek folii. Jaki plan produkcji należy przyjąć, aby zysk z przedsięwzięcia był największy?

Wstęp

Przykład: sformułowanie matematyczne

Formułujemy zadanie programowania matematycznego: Niech x_A i x_B oznaczają odpowiednio liczbę jednostek wyprodukowanych metodą A i B. Zysk można opisać funkcją:

$$f(x) = 1,5\text{zł} * x_A + 1,8\text{zł} * x_B$$

Dziennie zużyjemy

$0,5 * x_A + 0,6 * x_B$ jednostek kartonu i

$0,45 * x_A + 0,5 * x_B$ jednostek folii.

Wstęp

Przykład: sformułowanie matematyczne

Zapisujemy warunki oraz funkcję celu: zmaksymalizować

$$f(x) = 1,5\text{zł} * x_A + 1,8\text{zł} * x_B$$

Warunki ograniczające:

$$0,5 * x_A + 0,6 * x_B \leq 200, \quad 0,45 * x_A + 0,5 * x_B \leq 300$$

$$x_A \leq 200, \quad x_B \leq 200, \quad x_A \geq 0, \quad x_B \geq 0$$

Rozwiązanie optymalne: należy wyprodukować 196 jednostek metodą A i 170 jednostek metodą B. Osiągniemy wtedy maksymalny zysk 600 zł.

Zadanie programowania nieliniowego

są identyczne jak dla programowania liniowego, ale w przeciwieństwie do programowania liniowego, nie istnieje jeden uniwersalny algorytm rozwiązywania zadań programowania nieliniowego. Wynika to z faktu iż funkcje nieliniowe stanowią (w pewnym sensie) dużo bardziej obszerną rodzinę funkcji niż funkcje liniowe – funkcją nieliniową jest każda funkcja, która nie jest liniowa.

Zadanie programowania nieliniowego

Funkcje nieliniowe charakteryzują się następującymi cechami, które mogą utrudniać obliczenia:

- występowanie tzw. ekstremów lokalnych (lokalne minima lub maksima),
- występowanie tzw. punktów siodłowych, czyli takich, dla których funkcja osiąga maksimum dla jednej zmiennej, a minimum dla innej (na wykresie funkcji 2 zmiennych wyglądają one jak przełęcz lub siodło - stąd nazwa),
- nieciągłości ('przerwy' w wykresach),
- osobliwości (funkcja dąży do plus lub minus nieskończoności dla skończonej wartości argumentu).

Wszystko to powoduje, że poszukiwanie rozwiązania konkretnych zadań programowania nieliniowego zależy od szczególnej postaci tego zadania.

Zadanie programowania nieliniowego

Niektóre zadania programowania nieliniowego można rozwiązać:

- przy pomocy specjalnego algorytmu, jeśli zadanie zalicza się do jednego z podtypów, dla których takie algorytmy są znane;
- metodą simpleks, jeżeli istnieje możliwość przekształcenia w zadanie programowania liniowego np. tzw. programowanie ilorazowe - omówione dalej;
- przekształcając do postaci zadania programowania liniowego całkowitoliczbowego – przykładem może być zadanie transportowo-produkcyjne ze stałym kosztem uruchomienia produkcji czy zadanie optymalnej diety ze stałymi kosztami zakupu.

W ogólnym przypadku nie ma niestety żadnej gwarancji, że rozwiązanie podane przez komputer jest choćby zbliżone do poprawnego.

Programowanie ilorazowe

Zadanie programowania ilorazowego jest to maksymalizacja lub minimalizacja ilorazu dwóch funkcji liniowych przy ograniczeniach liniowych. Standardowa postać zadania programowania ilorazowego wygląda następująco:

$$\frac{c_0 + c_1x_1 + \dots + c_nx_n}{d_0 + d_1x_1 + \dots + d_nx_n} \rightarrow \min \quad \max$$

przy ograniczeniach

$$a_{11}x_1 + \dots + a_{1n}x_n \leq b_1$$

.....

$$a_{m1}x_1 + \dots + a_{mn}x_n \leq b_m$$

$$x_1 \geq 0, \dots, x_n \geq 0$$

Jeśli $d_0 + d_1x_1 + \dots + d_nx_n \neq 0$ dla $(x_1, \dots, x_n) \in D$ to zadanie programowania ilorazowego można sprowadzić do zadania programowania liniowego.

Programowanie ilorazowe

Wprowadźmy nowe zmienne:

$$y_1 = \frac{x_1}{d_0 + d_1 x_1 + \dots + d_n x_n}$$

.....

$$y_n = \frac{x_n}{d_0 + d_1 x_1 + \dots + d_n x_n}$$

$$t = \frac{1}{d_0 + d_1 x_1 + \dots + d_n x_n}$$

Wtedy

$$x_i = \frac{y_i}{t}$$

i poszukiwanie rozwiązania zadania programowania ilorazowego sprowadza się do rozwiązania następującego zadania programowania liniowego.

Programowanie ilorazowe

Rozwiązywanie graficzne zadania programowania ilorazowego z dwiema zmiennymi wygląda analogicznie jak rozwiązywanie graficzne zadania programowania liniowego tzn. należy wykreślić w układzie współrzędnych zbiór rozwiązań dopuszczalnych, a następnie sprawdzać wartości funkcji celu dla współrzędnych wierzchołków. Niemniej jednak przekształcenie w zadanie programowania liniowego w podany wyżej sposób nie jest akurat w tym przypadku ułatwieniem, ponieważ przekształcenie to wprowadza dodatkową zmienną t , co prowadziłoby do konieczności sporządzenia wykresu 3-wymiarowego.

Programowanie ilorazowe

Programowanie ilorazowe jest stosowane przy problemach decyzyjnych wymagających pogodzenia ze sobą dwóch sprzecznych kryteriów optymalności np.

$$\frac{\text{zysk}}{\text{pracochłonność}} \rightarrow \max$$

$$\frac{\text{przychód}}{\text{koszty}} \rightarrow \max$$

$$\frac{\text{koszty paszy}}{\text{dzienny przyrost masy zwierząt}} \rightarrow \min$$

Inny podział postaci zadań programowania nieliniowego

- 1 **Standardowa** — warunki ograniczające w postaci nierówności
- 2 **Kanoniczna** — warunki ograniczające w postaci równości (z wyjątkiem warunków brzegowych)

Program nieliniowy w postaci kanonicznej

Metoda mnożników Lagrange'a

- 1 Sprawdzenie występowania ekstremum bezwarunkowego funkcji celu, spełniającego warunki ograniczające
- 2 Ewentualne przekształcenie funkcji celu do postaci Lagrange'a i poszukiwanie bezwarunkowego ekstremum

Metoda mnożników Lagrange'a

Szukanie ekstremum bezwarunkowego

1. Funkcja celu f osiąga bezwarunkowe ekstremum w punkcie stacjonarnym w przypadku nieujemnej wartości wyznacznika macierzy drugich pochodnych funkcji celu f po poszczególnych zmiennych i ich kombinacjach. Ponadto wszystkie minory główne takiej macierzy muszą być dodatnie.
2. Współrzędne punktu stacjonarnego można otrzymać przyrównując do zera wartości pierwszych pochodnych cząstkowych funkcji celu f po poszczególnych zmiennych.

Metoda mnożników Lagrange'a

Szukanie ekstremum bezwarunkowego

$$\det \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \geq 0$$

$$\frac{\partial f}{\partial x_1} = 0, \quad \frac{\partial f}{\partial x_2} = 0, \dots, \quad \frac{\partial f}{\partial x_n} = 0,$$

Metoda mnożników Lagrange'a

Funkcja Lagrange'a

Funkcja Lagrange'a L wiąże funkcję celu f z funkcjami ograniczeń g_i , dzięki użyciu wektora tzw. nieoznaczonych mnożników Lagrange'a (λ).

$$L(x; \lambda) = f(x) + \lambda g$$

Dzięki wprowadzeniu funkcji L można zastąpić poszukiwania **optymalnej warunkowej wartości funkcji celu f** , poszukiwaniami odpowiadającej jej **bezwarunkowej wartości optymalnej funkcji L** .

Metoda mnożników Lagrange'a

Rozwiązanie optymalne

Rozwiązanie optymalne otrzymuje się rozwiązując następujący układ równań, zawierający $n + r$ równań (n – liczba zmiennych decyzyjnych, r – liczba funkcji ograniczeń g_i):

$$\frac{\partial L}{\partial x} = 0, \quad \frac{\partial L}{\partial \lambda} = 0$$

Metoda mnożników Lagrange'a

Postać standardowa zagadnienia PNL

Do rozwiązania można użyć:

- 1 metody mnożników Lagrange'a;
- 2 twierdzenia Kuhna-Tuckera

Metoda mnożników Lagrange'a

Metoda mnożników Lagrange'a

Sposób postępowania podobny, jak w przypadku kanonicznej postaci zadania PNL. W celu przekształcenia równań w nierówności, wprowadza się zmienne bilansujące, tzw. zmienne nieistotne u^2 .

$$x_1 + x_2 \leq 10 \rightarrow x_1 + x_2 + u^2 = 10$$

$$x_1 + x_2 \geq 10 \rightarrow x_1 + x_2 - u^2 = 10$$

Rozwiązanie optymalne:

$$\frac{\partial L}{\partial x} = 0, \quad \frac{\partial L}{\partial \lambda} = 0, \quad \frac{\partial L}{\partial u} = 0$$

Metoda mnożników Lagrange'a

Twierdzenie Kuhna-Tuckera

$$f(x_1, x_2, \dots, x_n) \rightarrow \min$$

$$g_i((x_1, x_2, \dots, x_n)) \leq 0, \quad i = 1, 2, \dots, r, \quad (x_1, x_2, \dots, x_n) \geq 0$$

$$L(x, \lambda) = f(x) + \lambda g$$

Warunki Kuhna-Tuckera

$$\frac{\partial L}{\partial x} \geq 0, \quad \frac{\partial L}{\partial x} x = 0, \quad g(x) = 0, \quad g\lambda = 0, \quad x \geq 0, \quad \lambda \geq 0$$

Metoda mnożników Lagrange'a

Twierdzenie Kuhna-Tuckera

Oznaczenia:

$$\frac{\partial L}{\partial x} = \nu, \quad g(x) = w$$

Zmodyfikowane warunki Kuhna-Tuckera

$$\frac{\partial L}{\partial x} - \nu = 0, \quad \nu x = 0, \quad g(x) + w = 0, \quad w \lambda = 0, \quad x \geq 0, \quad \lambda \geq 0$$

Metoda mnożników Lagrange'a

Twierdzenie Kuhna-Tuckera

Rozwiązanie optymalne można uzyskać rozpatrując wszystkie możliwe (spełniające ograniczenia) kombinacje wartości składowych wektorów: ν , λ , i w .

W tym celu należy rozwiązać poszczególne układy równań, wynikające z warunków Kuhna-Tuckera.

Modelowanie całkowniczbowe

W modelach programowania matematycznego zmiennych całkowniczbowych używa się m.in.:

- Do reprezentowania wielkości, które w swej naturze są całkowniczbowe, np. liczba produkowanych samochodów, samolotów, liczba budowanych domów, liczba zatrudnionych pracowników itp.

Modelowanie całkowitoliczbowe

- Do modelowania zmiennych decyzyjnych służących do wyboru decyzji ze zbioru możliwych decyzji. Są to najczęściej zmienne binarne. Np.

$$\delta = \begin{cases} 1 & \text{należy zbudować magazyn} \\ 0 & \text{nie budujemy} \end{cases}$$

lub też

$$\gamma = \begin{cases} 0 \geq 0 & \text{nic nie budujemy} \\ 1 & \text{należy zbudować magazyn A} \\ 2 & \text{należy zbudować magazyn A} \end{cases}$$

Modelowanie całkowitoliczbowe

- Do wyrażenia pewnych stanów zmiennych ciągłych w modelach liniowych. Są to binarne zmienne wskaźnikowe.
- Do modelowania warunków logicznych w rzeczywistych zagadnieniach.
- Do modelowania niektórych nieliniowych zależności.

Modelowanie całkowitoliczbowe

Zmienne wskaźnikowe

δ – zmienna wskaźnikowa związana ze zmienną ciągłą x to zmienna binarna, której celem jest rozróżnienie pomiędzy stanem zmiennej $x =$, a stanem $x > 0$.

Modelowanie całkownicze

Zmienne wskaźnikowe

Przykład. (Problem stałych kosztów).

Niech x będzie ilością wytwarzanego produktu po kosztach jednostkowych C_1 , a stałe koszty produkcji niech wynoszą C_2 . Całkowity koszt K_c wynosi zatem:

$$K_c = \begin{cases} 0 & \text{jeśli } x = 0 \\ C_1 x + C_2 & \text{jeśli } x > 0 \end{cases}$$

Koszt całkowity K_c nie jest funkcją liniową. Wprowadzając zmienną wskaźnikową δ taką, że $x > 0 \Rightarrow \delta = 1$ otrzymujemy liniową funkcję celu

$$K_c(x) = C_1 x + C_2 \delta$$

Modelowanie całkownicze

Zmienne wskaźnikowe

Przykład. ((Problem optymalnej mieszanki).

Niech zmienne x_A i x_B reprezentują odpowiednio procentową zawartość składników A i B w mieszance. Dodatkowo, oprócz innych ograniczeń dających wyrazić się w liniowej postaci występuje ograniczenie: *jeśli w mieszance wystąpi składnik A , to musi również w niej wystąpić składnik B .*

Wprowadzamy zmienną wskaźnikową δ taką, że $x_A > 0 \Rightarrow \delta = 1$, czyli ograniczenie

$$x_A - \delta \leq 0$$

Dodatkowo musimy wprowadzić warunek

$$\delta = 1 \Rightarrow x_B > 0$$

który modelujemy nierównościami

Literatura

- [1] H.P. Williams, *Model building in mathematical programming*, John Wiley and Sons, 1993
- [2] F. Plastria, *Formulating logical implications in combinatorial optimization*, European Journal of Operational Research 140 (2002) 338-353

Koniec?

Koniec wykładu 7