

Mgr Andrzej DOMAŃSKI

Wiele problemów technicznych polega na zaprojektowaniu takiego urządzenia, które spełniając założone warunki charakteryzowałoby się np. możliwie najmniejszym kosztem, ciężarem itp. Osiągnięcie takiego efektu polega na ogół na przeprowadzeniu obliczeń dla tak wielu możliwych przypadków, że jest w praktyce trudne do zrealizowania. Zastosowanie elektronicznej techniki obliczeniowej nierzadko umożliwia uzyskanie rozwiązania.

Zajmuje się tym dziedzina zwana optymalizacją.

Podstawowym problemem stojącym przed rozwiązującym zadanie optymalizacyjne jest zbudowanie odpowiedniego modelu matematycznego. Innymi słowy, chodzi o takie sformułowanie tego problemu, żeby poszukiwane rozwiązanie polegało na znalezieniu ekstremum pewnej funkcji n zmiennych, określonej w pewnym obszarze D .

Dokładniej: w przestrzeni kartezjańskiej C^n dana jest funkcja:

$$f(X) \quad \text{gdzie} \quad X = (x_1, \dots, x_n).$$

Dany jest ponadto obszar D opisany nierównościami

$$g_i(X) \geq 0, \quad i = 1, \dots, K,$$

gdzie g_i — funkcje rzeczywiste.

Funkcja $f(X)$ określona jest w obszarze D . Należy znaleźć maksymalną wartość $f(X)$ w tym obszarze oraz współrzędne punktu, w którym funkcja tę wartość przyjmuje. Funkcję $f(X)$ nazywa się funkcją celu.

Rozwiązanie tego zagadnienia bywa bardzo trudne. Trudność zależy tu od postaci funkcji f i g_i . Jeżeli wszystkie te funkcje są liniowe — mamy do czynienia z problemem tzw. programowania liniowego. W takim przypadku znane są metody dokładnego rozwiązania problemu za pomocą maszyny matematycznej. Często jednak zdarza się w praktyce, że funkcje f i g_i są nieliniowe i skomplikowane tak dalece, że trudno powiedzieć coś o własnościach funkcji f oraz o cechach obszaru D opisanego przez funkcje g_i . W takich przypadkach (tzn. wtedy, kiedy nic nie można z góry założyć o f i g_i) nie istnieją ogólne metody rozwiązania zadania optymalizacyjnego, tzn. nie znamy sposobu gwarantującego znalezienie ekstremum. Stosuje się wówczas metody, które, chociaż matematycznie niedoskonałe, są atrakcyjne z punktu widzenia pracy inżyniera. Punktem wyjścia są tu tzw. metody losowe (zwane też metodami Monte Carlo). Wykorzystują one zdolność maszyny matematycznej do szybkiego sprawdzenia, czy dany punkt \bar{X} należy do D , oraz równie szybkiego obliczenia wartości $f(\bar{X})$. Metoda poszukiwania największej wartości funkcji f polega wówczas na realizacji następującego ciągu czynności:

1. Losuje się punkt \bar{X} (maszyna i to potrafi).
2. Sprawdza się, czy punkt ten należy do D , jeżeli należy — przechodzi się do punktu 3, w przeciwnym przypadku — wraca się do punktu 1.
3. Oblicza się wartość $f(\bar{X})$.
4. Jeżeli $f(\bar{X})$ jest pierwszą uzyskaną wartością — zapamiętuje się ją jako f_{\max} . W przeciwnym przypadku — porównuje się $f(\bar{X})$ z f_{\max} i o ile $f(\bar{X}) > f_{\max}$, podstawia się $f_{\max} = f(\bar{X})$.
5. Wraca się do punktu 1 powtarzając wielokrotnie procedurę 1–5.

Opisana metoda nie gwarantuje, oczywiście, znalezienia globalnego ekstremum funkcji f , jest czasochłonna i mało efektywna. Dokładność wyniku zależy od liczby badanych punktów, a więc od czasu pracy maszyny. Metodę tę stosuje się rzadko, natomiast istnieją jej modyfikacje pozwalające dochodzić do lepszych wyników — aleko szybciej.

Warto tu wspomnieć o tzw. metodzie błędzenia. Polega ona na realizacji ciągu następujących czynności:

1. Losuje się punkt $X^0 \in D$.
2. Otacza się ten punkt n -wymiarowym „prostokątnikiem” o środku X^0 i danych z góry długościach krawędzi (na ogół niewielkich). „Prostokątnik” ten nazywać będziemy komórką K^0 o środku X^0 .
3. Losuje się punkt X^1 należący do K^0 i bada się, czy należy on do obszaru D . Jeżeli należy — oblicza się $f(X^1)$, otacza X^1 komórką i powtarza proces od punktu 2 dla X^1 (zamiast X^0). W przeciwnym przypadku — losuje się w komórce K^0 inny punkt i bada się go tak jak punkt X^1 .
4. Jeżeli przy tym p_0 kolejnych losowań w komórce K^0 nie daje efektu w postaci wylosowania punktu należącego do D — zmniejsza się komórkę dzieląc jej wymiary np. przez 2 i ponawia się próby. Jeżeli natomiast s_0 kolejnych losowań prowadzi do uzyskania punktów należących do D — zwiększa się wymiary komórki mnożąc je np. przez 2 i ponawia się próby.
5. Zapamiętuje się najlepsze wyniki (tzn. wartość funkcji celu i odpowiadający jej punkt).
6. Działania wg punktów 1–5 stanowią I etap metody. W II etapie realizuje się procedurę 1–5,

Przestrzeń kartezjańską n -wymiarową (ozn. C^n) nazywa się zbiór wszystkich n -elementowych ciągów liczb rzeczywistych. Płaszczyzna z układem współrzędnych może być interpretowana jako C^2 (bo każdy punkt możemy uważać za parę liczb), a przestrzeń trójwymiarowa z układem współrzędnych — jako C^3 . Oś liczbowa jest przestrzenią C^1 . Ze względu na geometryczne interpretacje przestrzeni C^1 , C^2 i C^3 przy mówieniu o przestrzeniach C^n stosuje się terminologię geometryczną. W szczególności — elementy C^n nazywa się punktami.

Funkcja rzeczywista — funkcja, której wartości są liczbami rzeczywistymi.

Ekstremum globalne — punkt, w którym funkcja przyjmuje wartość największą lub najmniejszą w całym obszarze.

D. c. zad. F19

Przy ruchu ze zmienną masą całkowity pęd danej części liny nie równa się pędowi środka masy, (tutaj jest dwukrotnie większy). W równaniach (8) powinien występować całkowity pęd danej części liny, czyli v_x i v_z powinny określać prędkość poszczególnych elementów liny, a nie ich środka masy. Ze względu na przybywającą (ubywającą) masę prędkość środka masy jest dwukrotnie mniejsza niż elementów liny,

$$T = R \frac{V\sqrt{2}}{2} \quad (\text{patrz równanie (6)}).$$



z tym że środek komórki przenosi się do nowego punktu tylko wówczas, gdy wartość funkcji f jest w tym punkcie lepsza niż w środku komórki. W II etapie punktami startowymi są najlepsze z punktów uzyskanych w I etapie (np. 10 tych punktów).

Naszkicowana metoda jest efektywniejsza od zwykłych metod losowych. Odpowiednie dobranie wartości p_0 i s_0 pozwala na takie sterowanie procesem, że przeszukiwane jest wnętrze obszaru D bądź jego brzeg. Ponieważ w zadaniach technicznych rozwiązanie (tzn. punkt, w którym funkcja f osiąga ekstremum globalne) leży na ogół na brzegu obszaru, stwarza to dodatkową korzyść dla rozwiązującego. Z drugiej strony stosowanie tej metody za pomocą maszyny jest czasochłonne i w gruncie rzeczy efekty nie są najlepsze.

Krokiem naprzód w stosunku do opisanych metod jest tzw. metoda osiowa (Gaussa-Zeidel). Realizacja jej przebiega następująco:

1. Losuje się punkt startowy $X^0 = (x_1^0, \dots, x_n^0)$ należący do D i oblicza $f(X^0)$.
2. Dla arbitralnie przyjętej liczby $h > 0$ oblicza się $f(x_1^0 + h, \dots, x_n^0)$.
3. Bada się czy $f(x_1^0 + h, \dots, x_n^0) > f(X^0)$. Jeżeli tak — oblicza się $f(x_2^0 + 2h, \dots, x_n^0)$ itd. dopóty, dopóki uzyskuje się wartości lepsze od $f(X^0)$. Jeżeli nie — oblicza się $f(x_1^0 - h, \dots, x_n^0)$ i postępuje analogicznie.
4. „Uzmiennia się” w powyższy sposób kolejno wszystkie współrzędne uzyskując „ścieżki” punktów równoległe odpowiednio do poszczególnych osi układu współrzędnych. Końcowy punkt ostatniej ścieżki zapamiętuje się jako najlepszy i wraca się do punktu 1_0 startując z nowo wylosowanego punktu.

Metoda ta daje niezłe wyniki dla funkcji o niewielkiej liczbie zmiennych. Dla funkcji o dużej liczbie zmiennych jest ona czasochłonna. Nie gwarantuje (podobnie, jak poprzednie metody) znalezienia ekstremum globalnego, a jej efektywność zależy od liczby ścieżek, a więc od czasu pracy maszyny. Na marginesie warto wspomnieć, że istnieje metoda wywodząca się z wyżej opisanej i wzbogacająca ją o obrót układu współrzędnych po określonych cyklach poszukiwań „wzdłuż osi”. W wielu przypadkach jest ona bardzo szybka i efektywna.

Odrębną klasę metod optymalizacyjnych stanowią tzw. metody gradientowe. Przykład takiej metody realizuje ciąg następujących czynności:

1. Losuje się punkt startowy $X^0 \in D$.
2. W punkcie X^0 oblicza się grad $f(X^0)$.
3. Określa się punkt $X^1 = X^0 + \text{grad } f(X^0)$.
4. Bada się czy $X^1 \in D$. Jeżeli należy — powtarza się proces od p. 2 dla $X^0 = X^1$ i dalej dla $X^{k-1} = X^k$. Jeżeli nie należy — świadczy to, że wektor gradientu „przeciął” brzeg obszaru. Sprawdza się wówczas,

czy punkt $X^k = X^{k-1} + \frac{1}{s} \text{grad } f(X^{k-1})$ należy do D (s jest daną z góry liczbą większą od 1).

Założona liczba takich badań z s skutkiem negatywnym kończy „ścieżkę” gradientową. Zapamiętuje się wówczas ostatni punkt tej ścieżki należący do D oraz odpowiadającą mu wartość funkcji celu.

5. Powtarza się cały proces dla nowego punktu startowego itd.
6. Jeżeli ekstremum nie znajduje się na brzegu obszaru, proces kończony jest za pomocą osobnego kryterium (np. porównywanie dwóch sąsiednich wartości funkcji celu).

Metody gradientowe bywają szybkie i skuteczne dla wielu funkcji i obszarów. Również i one nie dają gwarancji uzyskania ekstremum globalnego.

Wadą wszystkich naszkicowanych tu (a także innych) znanych metod jest ich niezbieżność. Dlatego ich wartość matematyczna jest zapewne niewielka lub żadna. Są one jednak z powodzeniem stosowane w rozwiązywaniu konkretnych zagadnień technicznych, które nie pozwalają opisać się za pomocą funkcji posiadających określone własności oraz określonych w obszarze spełniających warunki regularności (np. wypukłym). Jeżeli nawet znaleziona za pomocą którejś z tych metod wartość f_{\max} nie jest ekstremum globalnym, to w każdym razie na ogół jest „lepsza” od np. wylosowanej czy odgadniętej. Dlatego są one pożytecznym instrumentem współpracy inżyniera z maszyną matematyczną.

Gradient — wektor, którego współrzędnymi są pochodne cząstkowe funkcji. Dla funkcji dwu zmiennych, na przykład, $\text{grad } f(x, y) =$

$$= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right). \text{ (Proszę sprawdzić, że}$$

$\text{grad } (x \cdot y) = (y, x)$). Kierunek tego wektora interpretuje się jako kierunek najszybszego wzrostu (spadku) wartości funkcji, a jego długość — jako miarę prędkości tego wzrostu (spadku).

Zbieżność metody. Metoda, pozwalająca na znajdowanie rozwiązań przybliżonych nazywa się zbieżną, jeśli można udowodnić, że ciąg otrzymywanych przy zastosowaniu tej metody rozwiązań przybliżonych jest zbieżny do rozwiązania dokładnego.

Od Red.: Nie zawsze zagadnienia optymalizacji polegają na maksymalizacji funkcji celu. Niektórych zagadnień praktycznych nie da się sprowadzić do takiego modelu. Do tej sprawy wrócimy w przyszłości.

Korespondencyjne Towarzystwo Naukowe

Inicjatywa kol. Snopczyńskiego «Delta» 1975,1 — okazała się bardzo cenna. Wielu Czytelników prosi o zamieszczenie ich adresów chcąc nawiązać kontakt z kolegami o podobnych zainteresowaniach. Być może z czasem wypracujemy jakąś formę organizacyjną tych kontaktów — np. Młodzieżowe Towarzystwo Naukowe lub koła Młodzieżowe Polskiego Towarzystwa Fizycznego i Polskiego Towarzystwa Matematycznego. Czekamy na wypowiedzi. Na łamach «Delt» w kąciku Korespondencyjne Towarzystwo Naukowe będziemy zamieszczać adresy Czytelników pragnących nawiązać kontakt listowny z kolegami i koleżankami.

Adresy

Jerzy Brołoń 95-021 Andrzejów k/Łodzi ul. Słowackiego 23
Miroslaw Pstrągowski 93-219 Łódź ul. Tatrzańska 69 m. 33
Artur Kalita 22-460 Szczepieszyn ul. Ogrodowa 17
Stanisław Wróbel 33-101 Tarnów ul. Traugutta 7/1

Tomasz Prokop 15-215 Białystok ul. Konopnickiej 12 m 9
Romuald Mirski 64-920 Piła ul. Bieruta 15 m 1
Tomasz Kasperski 84-300 Łęborg ul. 15 Grudnia 29a m 6
Lidia Czajkowska 26-110 Skarżysko-Kam. ul. Zielona 47