ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 1971

Seria: AUTOMATYKA z. 17 Nr kol. 302

MACIEJ BARGIELSKI
Katedra Kompleksowych Systemów Sterowania

ORGANIZACJA POSZUKIWANIA OPTYMALNYCH PARAMETRÓW MODELU MATEMATYCZNEGO

Streszczenie. W niniejszej pracy przedstawiono w krótkim zarysie niezbędne dla rozwiązania problemu poszukiwania optymalnych parametrów modelu matematycznego cechy charakterystyczne powierzchni wielu zmiennych. Następnie zaproponowano pewien algorytm umożliwiający znajdywanie maksimum (lub minimum) dostatecznie szerokiej klasy powierzchni wielu zmiennych.

1. Wstęp

Identyfikacja obiektów, czyli znalezienie numerycznych wartości współczynników modelu matematycznego danego obiektu jest nader istotnym problemem. Praktyka wykazuje, że otrzymane z procedury identyfikacyjnej współczynniki są nie tylko funkcją parametrów identyfikowanego obiektu (co jest oczywiste), ale i funkcją sposobu przeprowadzania identyfikacji (co jest już mniej oczywiste).

Ściślej rzecz biorąc, przy metodzie identyfikacyjnej polegającej na minimalizacji odległości pomiędzy zbiorem sygnałów pomierzonych a wyliczonych na podstawie modelu, współczynniki tego modelu są zależne od przyjętej odległości, czyli inaczej mówiąc - od metryki wprowadzonej w danej przestrzeni eksperymentu. Dla lepszego unaocznienia tego faktu rozpatrzmy identyfikację statyczną obiektu liniowego jednowymiarowego o modelu matematycznym

$$y = ax + b. \tag{1}$$

Dysponujemy zbiorami sygnałów wejściowych

$$E(x) = (x_1, x_2, \ldots, x_n)$$

oraz sygnałów wyjściowych

$$E(y) = (y_1, y_2, \ldots, y_n).$$

Przyjmijmy jako metrykę

$$d(E(x),E(y)) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2} \tag{2}$$

co po zminimalizowaniu metryki

$$d(E(y^*),E(x)),$$

gdzie $E(y^*)$ jest zbiorem sygnałów obliczonych na podstawie modelu

$$E(y^*) = (y_1^* = ax_1 + b,\ y_2^* = ax_2 + b,\ \ldots,\ y_n^* = ax_n + b)$$

ze względu na nieznane wartości a i b pozwoli na znalezienie ich wartości optymalnych a^* i b^* w sensie tej metryki.
Przyjmując jako metrykę

$$d'(E(x),\ E(y)) = \sum_{i=1}^{n} |y_i - x_i| \tag{3}$$

lub

$$d''(E(x),\ E(y)) = \sum_{i=1}^{n} \operatorname{sign}|y_i - x_i| \tag{4}$$

otrzymamy wartości $a^{*\prime}$ i $b^{*\prime}$ oraz $a^{*\prime\prime}$ i $b^{*\prime\prime}$ różne od otrzymanych poprzednio.

Dodatkową trudnością jest to, że przyjęcie modelu matematycznego w formie równoważnej (1) pod względem matematycznym, a mianowicie

$$x = \frac{1}{a}y - \frac{b}{a} \tag{5}$$

przy tej samej metryce, np. (2) prowadzi do zupełnie różnych wartości a oraz b.

W przypadku modeli matematycznych o większej liczbie zmiennych zbiory E(x) oraz E(y) są, ogólnie biorąc elementami przestrzeni E^n. Zbiór E(y) można więc rozpatrywać jako pewną reprezentantę hiperpowierzchni w tej przestrzeni. Problem minimalizacji odległości d sprowadza się więc do znalezienia minimum tej powierzchni ze względu na nieznane współczynniki modelu przy ograniczeniach wynikających z fizycznych właściwości obiektu, a więc znalezienia minimum znajdującego się w pewnym obszarze zwanym polem eksperymentu.

Najczęściej dotychczas używaną jest metryka euklidesowa (1), dlatego istnieją opracowania szeregu metod poszukiwania ekstremum (maksimum lub minimum) powierzchni wypukłych klasy C^2. Okazuje się jednak, że pojęcie wypukłości funkcji jest pojęciem zbyt wąskim i w wielu przypadkach trzeba znaleźć ekstremum funkcji niewypukłej, kiedy to metody te zawodzą. Przyjęcie innych metryk np. (3) lub (4) prowadzi do innych, specjalnych rodzajów powierzchni posiadających jednoznaczne minimum, a więc rozwiązanie problemu identyfikacji w tych przestrzeniach istnieje.

Istotnym problemem tu występującym jest przyporządkowanie odpowiedniej metody poszukiwania ekstremum do przyjętej metryki.

2. Geometria powierzchni wielu zmiennych

W celu przedstawienia trudności, na jakie napotyka się przy przyjęciu różnych metryk rozpatrzmy różne kształty powierzchni rozpiętej nad polem eksperymentu. Powierzchnia ta opisana jest równaniem modelu matematycznego

$$y = f(x_1, x_2, \ldots, x_n) \tag{6}$$

i nosi nazwę powierzchni reakcji (response surface). Cechą charakterystyczną tej powierzchni warunkującą istnienie jednoznacznego ekstremum, jest unimodalność.

Dla zdefiniowania tego pojęcia wprowadźmy wpierw pojęcie trajektorii przechodzącej przez dane dwa punkty $\underline{x}^1$ i $\underline{x}^2$ w przestrzeni eksperymentu

$$\underline{x}^1 = (x_1^1, x_2^1, \ldots, x_n^1)$$

$$\underline{x}^2 = (x_1^2, x_2^2, \ldots, x_n^2).$$

Trajektorią przechodzącą przez dwa punkty $\underline{x}^1$ oraz $\underline{x}^2$ będziemy nazywać każdą krzywą opisaną równaniem

$$\underline{x} = t(\underline{x}^1, \underline{x}^2, \lambda) \tag{7}$$

i taką, że

$$t(\underline{x}^1, \underline{x}^2, 0) = \underline{x}^1$$

$$t(\underline{x}^1, \underline{x}^2, 1) = \underline{x}^2$$

W szczególności, trajektorią może być prosta przechodząca przez punkty $\underline{x}^1$ i $\underline{x}^2$ o równaniu

$$\underline{x} = \underline{x}^1 + \lambda(\underline{x}^2 - \underline{x}^1). \tag{9}$$

Parametr λ określa kierunek ruchu po trajektorii. Możemy więc określić, jak zmienia się powierzchnia (6) wzdłuż danej trajektorii. Jest rzeczą oczywistą, że wartość y jest funkcją parametru λ:

$$y(x_1(\lambda), x_2(\lambda), \ldots, x_n(\lambda)) = y(\lambda). \tag{10}$$

Oznaczmy przez λ^* wartość $\lambda \in [0,1]$, przy której funkcja y osiąga swoje maksimum (minimum)

$$\begin{gathered} y(\lambda^*) = \max_{0 \leq \lambda \leq 1} y(\lambda) \\ (y(\lambda^*) = \min_{0 \leq \lambda \leq 1} y(\lambda)). \end{gathered} \tag{11}$$

Weźmy dwie dowolne wartości λ_1 i λ_2 takie, że $\lambda_1 < \lambda_2$. Mówimy, że funkcja y jest <u>unimodalna na trajektorii</u> od $\underline{x}^1$ do $\underline{x}^2$, jeśli zachodzi:

$$\begin{gathered} \lambda_2 < \lambda^* \Longrightarrow y(\lambda_1) < y(\lambda_2) \quad (y(\lambda_1) > y(\lambda_2)) \\ \wedge \\ \lambda^* < \lambda_1 \Longrightarrow y(\lambda_1) > y(\lambda_2) \quad (y(\lambda_1) < y(\lambda_2)). \end{gathered} \tag{12}$$

Dla zdefiniowania unimodalności funkcji (powierzchni) w obszarze eksperymentu załóżmy, że powierzchnia (6) osiąga swoje maksimum (minimum) w tym obszarze w punkcie $\underline{x}^*$

$$y(\underline{x}^*) = \max_{\underline{x}} y(\underline{x}) \quad (= \min_{\underline{x}} y(\underline{x})). \tag{13}$$

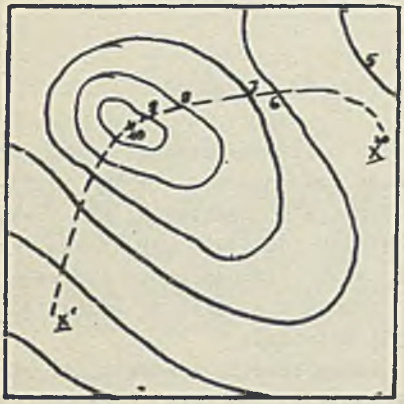
Powierzchnię $y(\underline{x})$ będziemy nazywali <u>unimodalną</u> w obszarze eksperymentu jeśli dla każdych dwóch punktów $\underline{x}^1$ i $\underline{x}^2$ z tego obszaru istnieje trajektoria przechodząca przez $\underline{x}^*$, na której funkcja jest unimodalna.

Definicję powyższą można również podać w inny nieco sposób. Mianowicie mówimy, że trajektoria jest silnie rosnąca (malejąca) jeśli

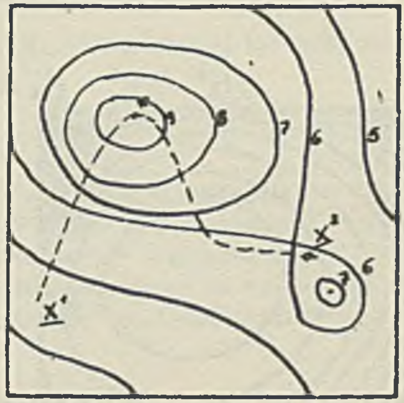
$$\lambda_1 < \lambda_2 \Longrightarrow y(\lambda_1) < y(\lambda_2) \quad (y(\lambda_1) > y(\lambda_2)). \tag{14}$$

Powierzchnia $y(\underline{x})$ jest więc unimodalną w obszarze eksperymentu, jeżeli dla każdego, dowolnego punktu $\underline{x}$ z tego obszaru istnieje silnie rosnąca (malejąca) trajektoria z maksimum (minimum) w punkcie $\underline{x}^*$ (13).

Dla ilustracji powyższych definicji na rys. 1 przedstawiono dwuwymiarową powierzchnię, która jest unimodalna, a na rys. 2 - nie. Powierzchnia z rys. 2 nosi nazwę bimodalnej.



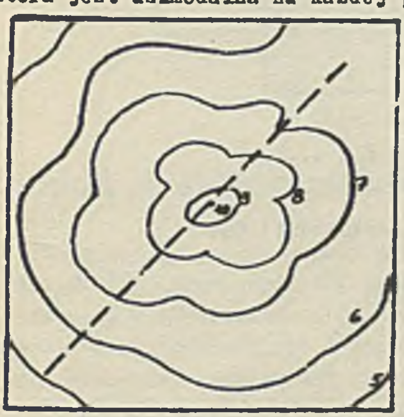
Rys. 1. Powierzchnia unimodalna



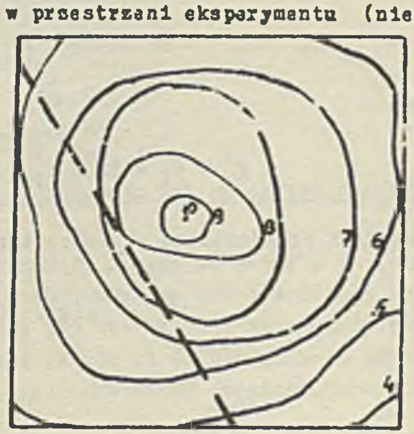
Rys. 2. Powierzchnia bimodalna

Zdefiniowana tu unimodalność jest pojęciem dostatecznie szerokim i zgadza się z określeniem jedyności ekstremum dla dowolnej funkcji niezależnie od przegięć, wąwozów i grzbietów.

Można wprowadzić pewne ograniczenia dla klasy funkcji unimodalnych. I tak, funkcją silnie unimodalną nazywamy funkcję, która jest unimodalna w przestrzeni eksperymentu na każdej prostej przechodzącej przez punkt ekstremalny (rys. 3). Powierzchnią liniowo unimodalną nazywamy powierzchnię, która jest unimodalna na każdej prostej w przestrzeni eksperymentu (nie-



Rys. 3. Powierzchnia silnie-unimodalna



Rys. 4. Powierzchnia liniowo-unimodalna

koniecznie przechodzącej przez punkt ekstremalny $\underline{x}^*$ (rys. 4). W końcu najwęższa klasa unimodalności - wypukłą (wklęsłą) nazywamy powierzchnię spełniającą warunek

$$y(\underline{x}^1 + \lambda(\underline{x}^2 - \underline{x}^1)) \geqslant y(\underline{x}^1) + \lambda(y(\underline{x}^2) - y(\underline{x}^1))$$
$$(y(\underline{x}^1 + \lambda(\underline{x}^2 - \underline{x}^1)) \leqslant y(\underline{x}^1) + \lambda(y(\underline{x}^2) - y(\underline{x}^1))). \tag{15}$$

Rys. 5. Powierzchnia wypukła

Wiele problemów wymagających konkretnego rozwiązania można opisać przy funkcji unimodalnych, nie będących jednak funkcjami wypukłymi. Istnieje więc jednoznaczne ekstremum, które jednak nie jest do osiągnięcia znanymi metodami.

W szczególności, powierzchnie wielu zmiennych posiadają grzebienie (grzbiety lub wąwozy). Grzebieniem nazywać będziemy miejsce geometryczne takich punktów, na których proces poszukiwania ekstremum metodą siecznych (Gaussa-Seidela) zatrzymuje się nie osiągnąwszy ekstremum. Dla matematycznej formalizacji tej definicji przyjmijmy, że $\varepsilon_1 > 0$ jest liczbą określającą najmniejszą odległość w kierunku prostopadłym do osi x_i ortogonalnego układu współrzędnych, że punkty

$$\underline{x}^1 = (x_1^1, x_2^1, \ldots, x_i^1, \ldots, x_n^1)$$

oraz

$$\underline{x}^1 + \varepsilon_i = (x_1^1, x_2^1, \ldots, x_i^1 + \varepsilon_i, \ldots, x_n^1)$$

posiadają rozróżnialne wartości funkcji (powierzchni) tj.

$$y(\underline{x}^1) \neq y(\underline{x}^1 + \varepsilon_i).$$

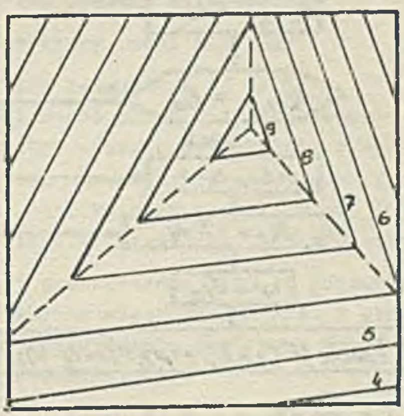
Mówimy, że punkt $\underline{x}^*$ znajduje się na grzbiecie (w wąwozie), jeśli zachodzi

$$\forall_1: \quad y(\underline{x}^*) > \max(y(\underline{x}^* - \varepsilon_1), y(\underline{x}^* + \varepsilon_1))$$
$$(y(\underline{x}^*) < \min(y(\underline{x}^* - \varepsilon_1), y(\underline{x}^* + \varepsilon_1))). \tag{16}$$

Z określenia grzebienia widać, że jego wielkość (powierzchnia) jest zależna od dwóch czynników:

- kształtu samej powierzchni oraz
- kroku poszukiwania ε_1.

W przypadku powierzchni klasy C^2 przez odpowiednie zmniejszenie kroku poszukiwania można czasami "wyjść" z grzebienia, tzn. kontynuować proces poszukiwania aż do osiągnięcia ekstremum. Jednak w przypadku często stosowanej miary modułowej takie rozwiązanie jest niemożliwe ze względu na ostrość występujących tu grzebieni (pochodna nie istnieje) (rys. 6).

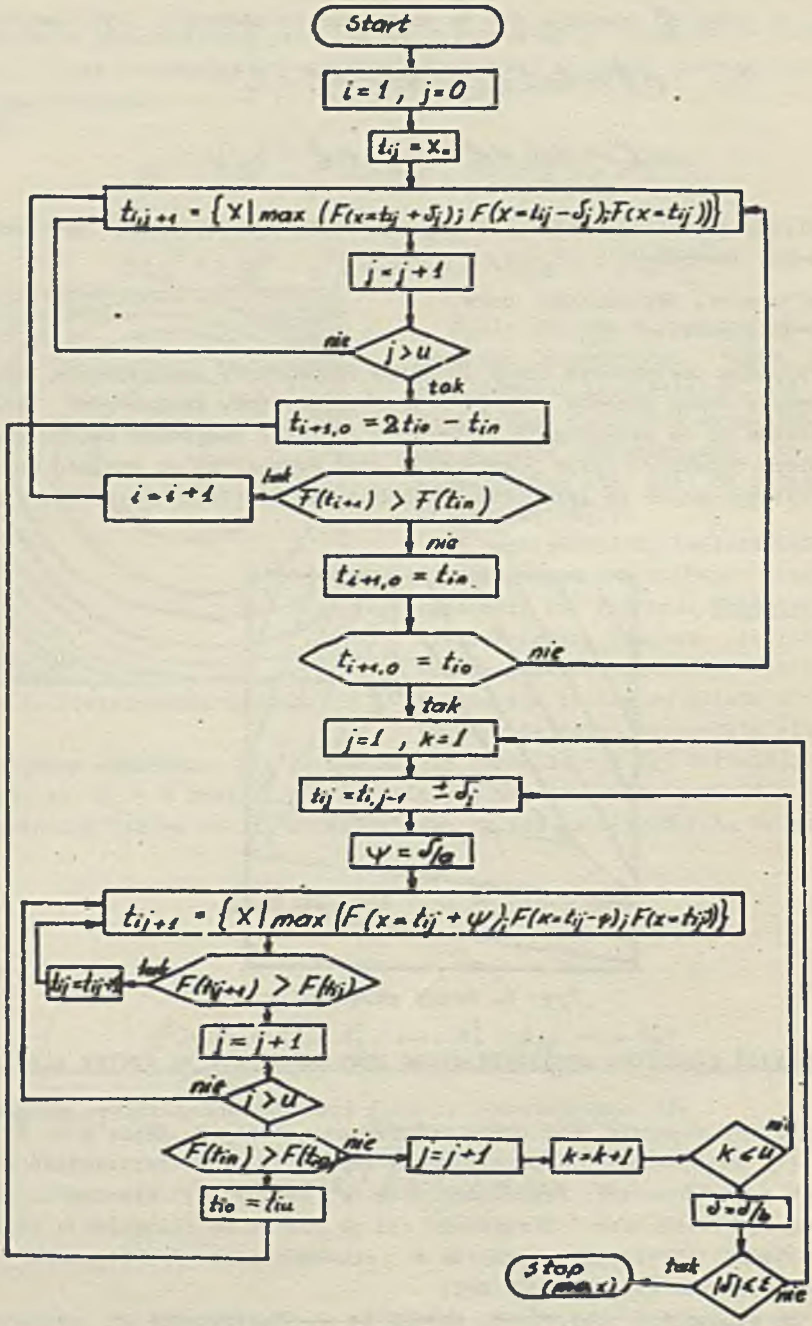


Rys. 6. Ostry grzebień

3. Propozycja algorytmu umożliwiającego poruszanie się po ostrym grzebieniu

Proponowany algorytm jest pewną modyfikacją metody R. Hooke'a i T. A. Jeeves'a [1] opracowanej dla poszukiwania ekstremum na powierzchniach unimodalnych z grzebieniami. Zasadniczą ideą tej metody jest wyszukanie grzebienia danej powierzchni i poruszanie się po nim aż do osiągnięcia ekstremum. Przypadki "utknięcia" algorytmu w nie-ekstremum wyeliminowane są przez zmniejszenie kroku poszukiwań.

Jak już wspomniano poprzednio, metody te są nieefektywne w przypadku ostrych grzebieni i dlatego wprowadzono modyfikację algorytmu. Polega ona


Start
i=1, j=0
j=j+1
nie
j>u
tak
i=i+1
tak
nie
nie
tak
j=1, k=1
j=j+1
nie
j>u
nie
j=j+1
k=k+1
k<u
nie
stop (max)
tak
nie

Rys. 7

na wyeliminowaniu zmniejszania kroku w algorytmie HJ oraz na wyznaczeniu ewentualnego kierunku grzebienia w przypadku "utknięcia" algorytmu HJ i dalszym kontynuowaniu poszukiwania ekstremum wg algorytmu HJ w nowym kierunku, lub na stwierdzeniu, że algorytm osiągnął już ekstremum.

Wyznaczenie kierunku ewentualnego grzebienia polega na wykonaniu skoku określonego punktem zatrzymania się algorytmu HJ i krokiem poszukiwań, a następnie potraktowaniu tego punktu jako nowych warunków początkowych, z których poszukujemy ekstremum w pozostałych kierunkach ortogonalnych. Punkt zatrzymania się algorytmu HJ oraz punkt ekstremalny nowych warunków początkowych (o ile są różne) wyznaczają nowy kierunek poszukiwania ekstremum wg algorytmu HJ.

Jako kryterium zbieżności przyjęto, że algorytm znajduje się w punkcie ekstremalnym z zadaną dokładnością ε, jeżeli po zatrzymaniu się algorytmu skoki na odległość ε we wszystkich kierunkach ortogonalnych dawały punkt ekstremalny pokrywający się z danym. Zbieżność proponowanego algorytmu zagwarantowana jest zbieżnością algorytmu HJ [1].

Schemat blokowy zmodyfikowanego algorytmu HJ przedstawia rys. 7. Oznaczenia przyjęte na tym rysunku są następujące:

x_0 – arbitralnie wybrany punkt startowy

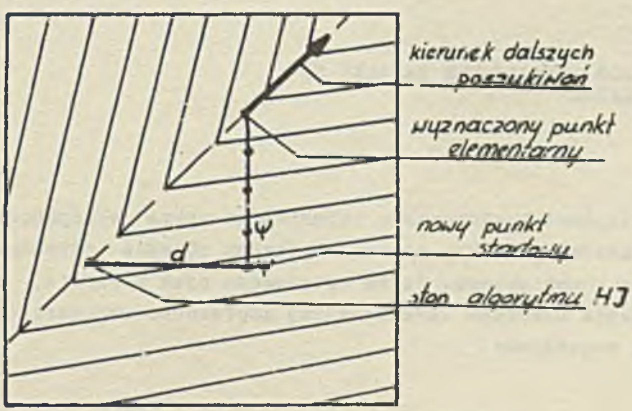
δ – początkowy krok poszukiwań ($\delta > \varepsilon$)

ε – założona dokładność określenia ekstremum

a, b – liczby naturalne wybrane arbitralnie określające dokładność określenie ekstremum procedury poszukiwawczej z nowych warunków początkowych oraz kontrakcję kroku poszukiwań (np. a=5, b=2)

t_{1k} – kolejne współrzędne wyznaczane przez algorytm.

Algorytm z rys. 7 przeznaczony jest do poszukiwania maksimum. Dla lepszej ilustracji działania algorytmu na rys. 8 przedstawiono jego zachowanie się w przypadku poszukiwania ekstremum dwuwymiarowej powierzchni unimodalnej z ostrym grzbietem.



Rys. 8. Zachowanie się algorytmu na grzbiecie

4. Zakończenie

W problemach identyfikacji koniecznym jest przeprowadzenie charakteryzacji i identyfikacji obiektu mając stale na uwadze cel sterowania. Abstrahowanie w identyfikacji od celu sterowania nie daje pożądanych wyników, i dlatego jest rzeczą konieczną dobór odpowiedniej metryki przestrzeni. Z problemem tym wiąże się nierozłącznie wybór odpowiednich metod poszukiwania ekstremum do przyjętej metryki.

Ze względu na duże zalety miary modułowej i związane z nimi coraz szersze zastosowanie tego typu miary do zagadnień praktycznej identyfikacji procesów, koniecznym wydaje się opracowanie coraz efektywniejszych metod poszukiwania ekstremum, działających na niegładkich powierzchniach unimodalnych.

Proponowany algorytm został sprowadzony na przykładach poszukiwania ekstremum powierzchni typu jak na rys. 6, dając za każdym razem wyniki zgodne z rzeczywistym położeniem ekstremum (które było znane) z zadaną dokładnością.

LITERATURA

[1] Hooke R., Jeeves T.A. - Direct search solution of numerical and statistical problems. Jounrn. of Assoc. Comp. Mach. 8,2 (apr. 1961)

[2] Wilde D.J. - Optimum seeking methods. Prentice-Hall Inc. Englewood Cliffs N.J. 1964.

[3] Węgrzyn S. - Podstawy automatyki kompleksowej. Prace IA PAN Warszawa 1969.

Rękopis złożono w Redakcji w dniu 28.I.1970 r.

ОРГАНИЗАЦИЯ ПОИСКОВ ОПТИМАЛЬНЫХ ПАРАМЕТРОВ МАТЕМАТИЧЕСКОЙ МОДЕЛИ

Р е з ю м е

В этом труде вкратце представлено характерные черты поверхности многих переменных, необходимые для решения проблемы поисков оптимальных параметров математической модели. Затем предложено один алгоритм, делающий возможным нахождение максимум (или минимум) достаточно широкого класса поверхности многих переменных.

OPTIMUM MATHEMATICAL MODEL PARAMETER SYSTEM SEEKING

S u m m a r y

In this paper there are given in short a characteristic traits a multidimensional surfaces necessary for solving optimum mathematical model parameter seeking problem. Next some algorithm is propose, in use that is possible to find maximum (or minimum) a sufficiently wide of multidimensional surfaces class.