



OMNIS

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.

Optymalizacja w inżynierii lotniczej i kosmicznej

Wykład 4

Matematyczne podstawy optymalizacji 3

Funkcje N zmiennych bez więzów

1130-LK000-MSP-1037



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



Politechnika Warszawska



OMNIS

Zagadnienia

- Wstęp do zagadnienia minimalizacji funkcji n -zmiennych bez więzów
- Metody poszukiwania rozwiązania
- Warunki zbieżności





Problematyka analizy zagadnień opisanych funkcjami o n-zmiennych sprowadza się do minimalizacji funkcji

$$\nabla F(\mathbf{X}^*) = \mathbf{0} \quad \mathbf{X}^* - \text{rozwiązanie}$$

gdzie

$$\nabla F(\mathbf{X}) = \begin{Bmatrix} \partial F(\mathbf{X}) / \partial X_1 \\ \partial F(\mathbf{X}) / \partial X_2 \\ \dots \\ \partial F(\mathbf{X}) / \partial X_n \end{Bmatrix}$$

Uwaga !

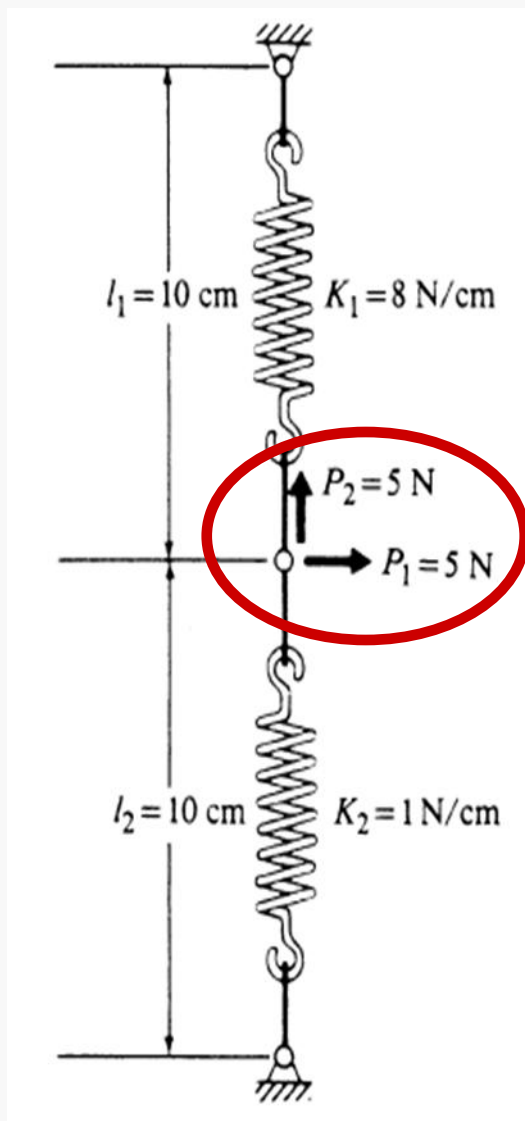
Rzadko jest możliwe rozwiązanie zagadnienia wprost ze względu na silne nieliniowość funkcji.

Dlatego zagadnienia tego typu rozwiązuje się w sposób iteracyjny





OMNIS Przykład (1/6)



Znaleźć warunki równowagi układu sprężyn takie aby energia potencjalna układu była minimalna ?

Wektory sił wymuszających

Geometria układu przed deformacją



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

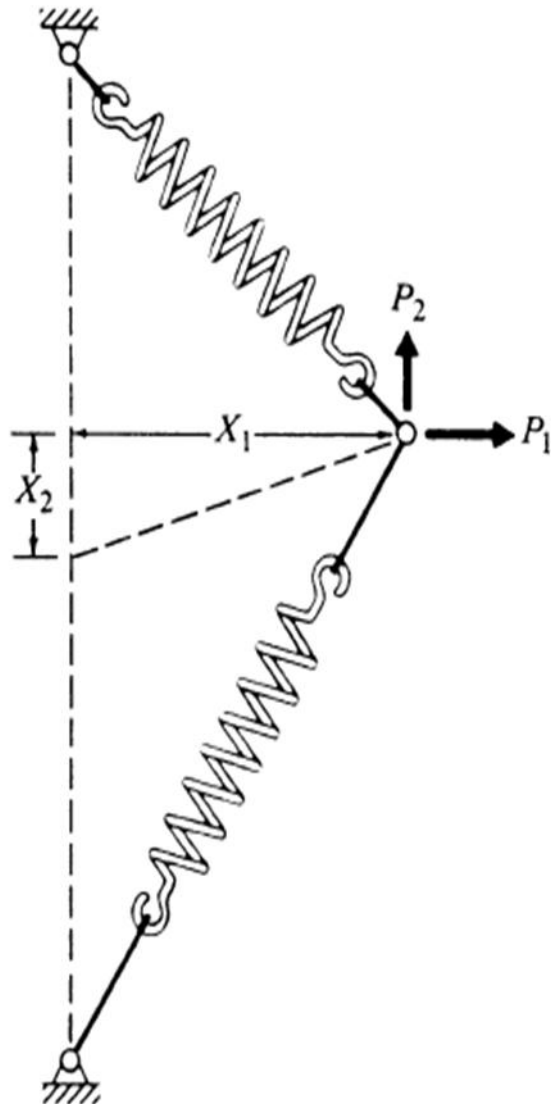
Dofinansowane przez
Unię Europejską



Politechnika Warszawska



OMNIS Przykład (2/6)



Energia potencjalna układu po deformacji:

$$PE = PE_{S1} + PE_{S2} - P_1 X_1 - P_2 X_2$$

gdzie:

$$PE_{S1} = \frac{1}{2} K_1 (\Delta l_1)^2 \quad \text{dla sprężyny 1}$$

$$PE_{S2} = \frac{1}{2} K_2 (\Delta l_2)^2 \quad \text{dla sprężyny 2}$$

Geometria układu po deformacji



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

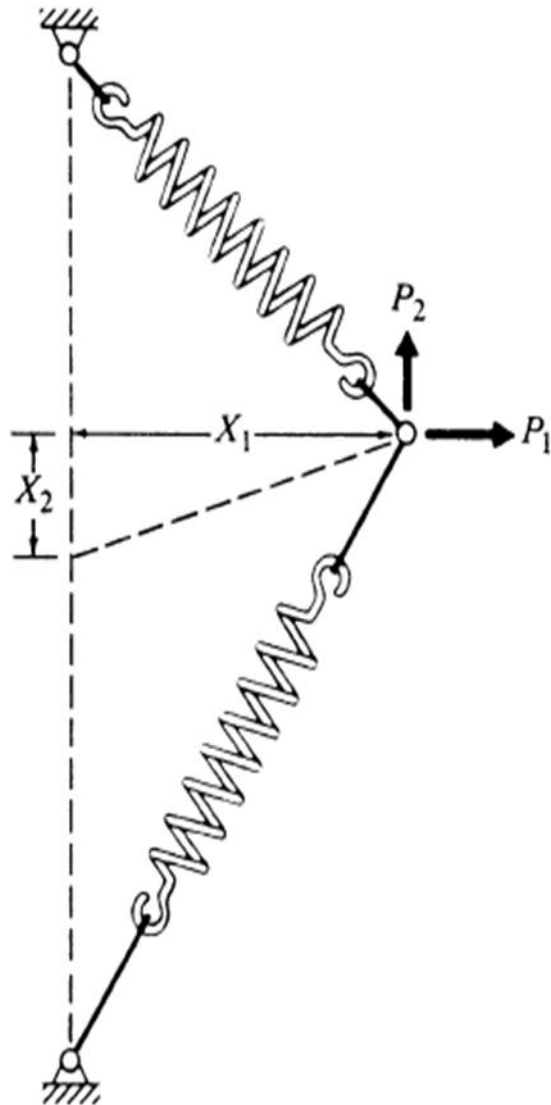
Dofinansowane przez
Unię Europejską



Politechnika Warszawska



OMNIS Przykład (3/6)



Energia potencjalna układu po deformacji:

$$PE = \frac{1}{2} K_1 \left[\sqrt{X_1^2 + (l_1 - X_2)^2} - l_1 \right]^2 + \frac{1}{2} K_2 \left[\sqrt{X_1^2 + (l_2 + X_2)^2} - l_2 \right]^2 - P_1 X_1 - P_2 X_2$$

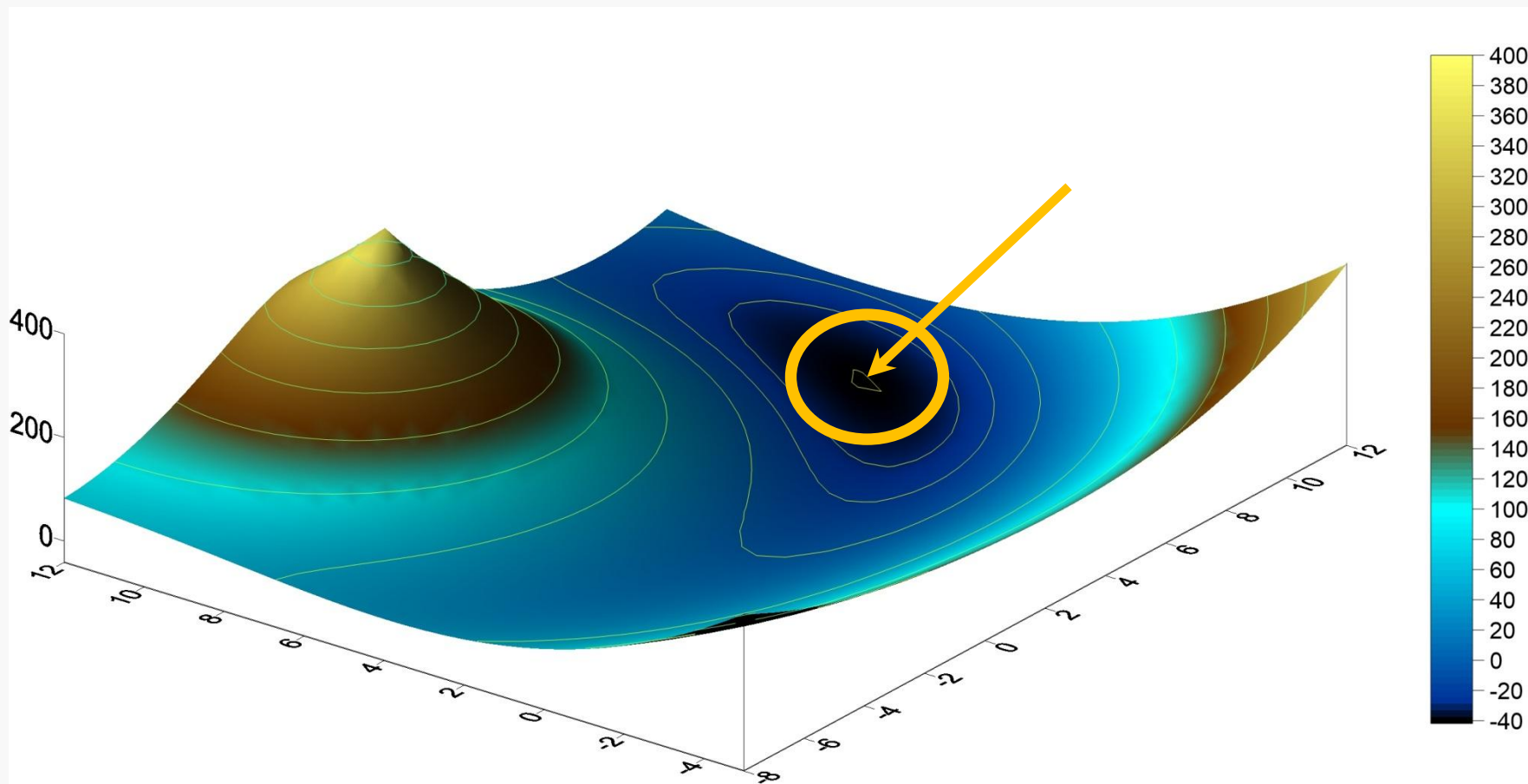
Geometria układu po deformacji





OMNIS Przykład (4/6)

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.



Graficzna interpretacja zagadnienia układu dwóch sprężyn



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską

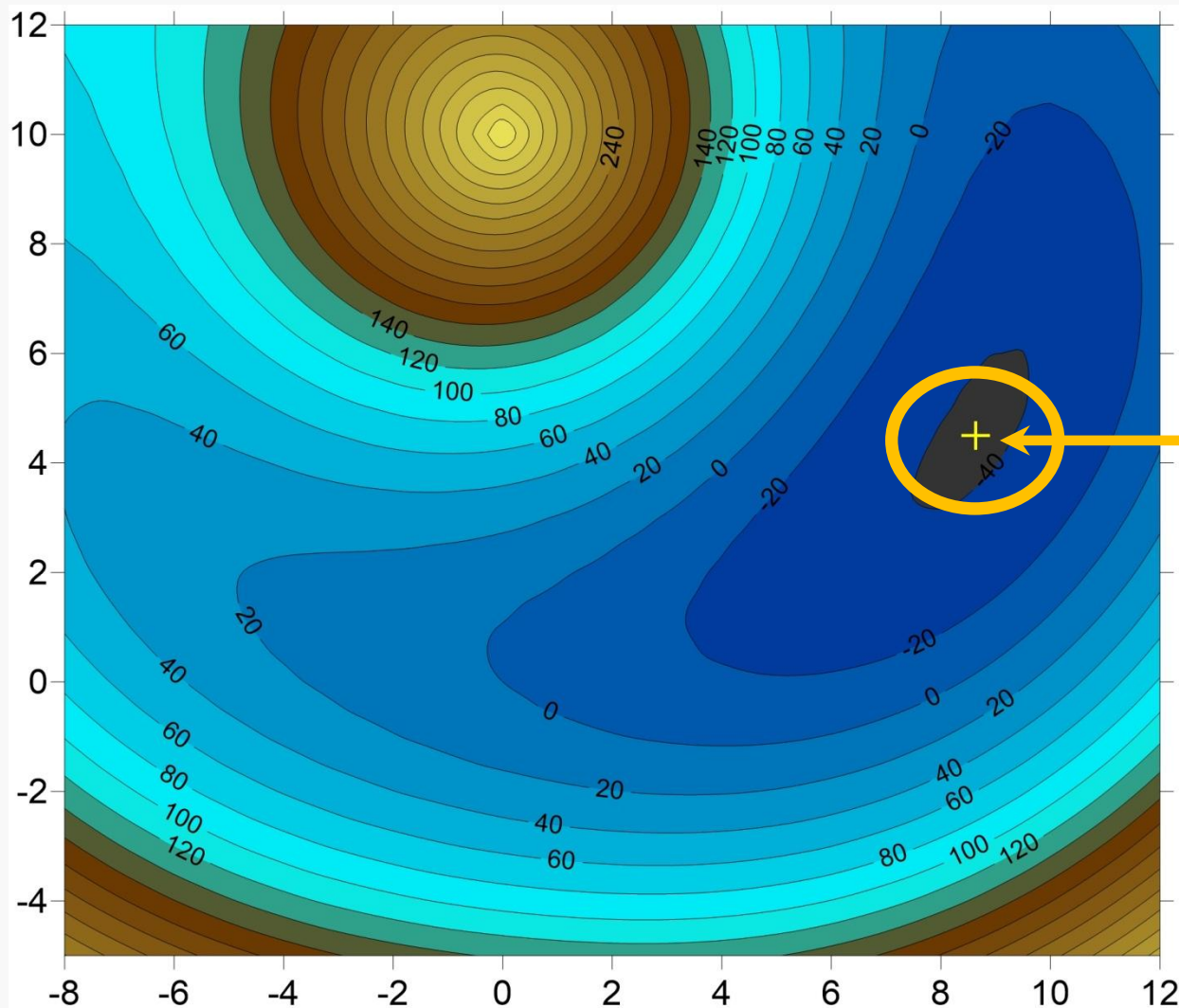


Politechnika Warszawska



OMNIS Przykład (4/6)

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.



Rozwiązanie:

**Minimum energii
potencjalnej jest
osiągalne dla:**

$$X_1 = 8.631 \text{ [cm];}$$

$$X_2 = 4.533 \text{ [cm];}$$

$$PE = -41.81 \text{ [N cm];}$$

**Graficzna interpretacja
zagadnienia układu
dwóch sprężyn**



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską

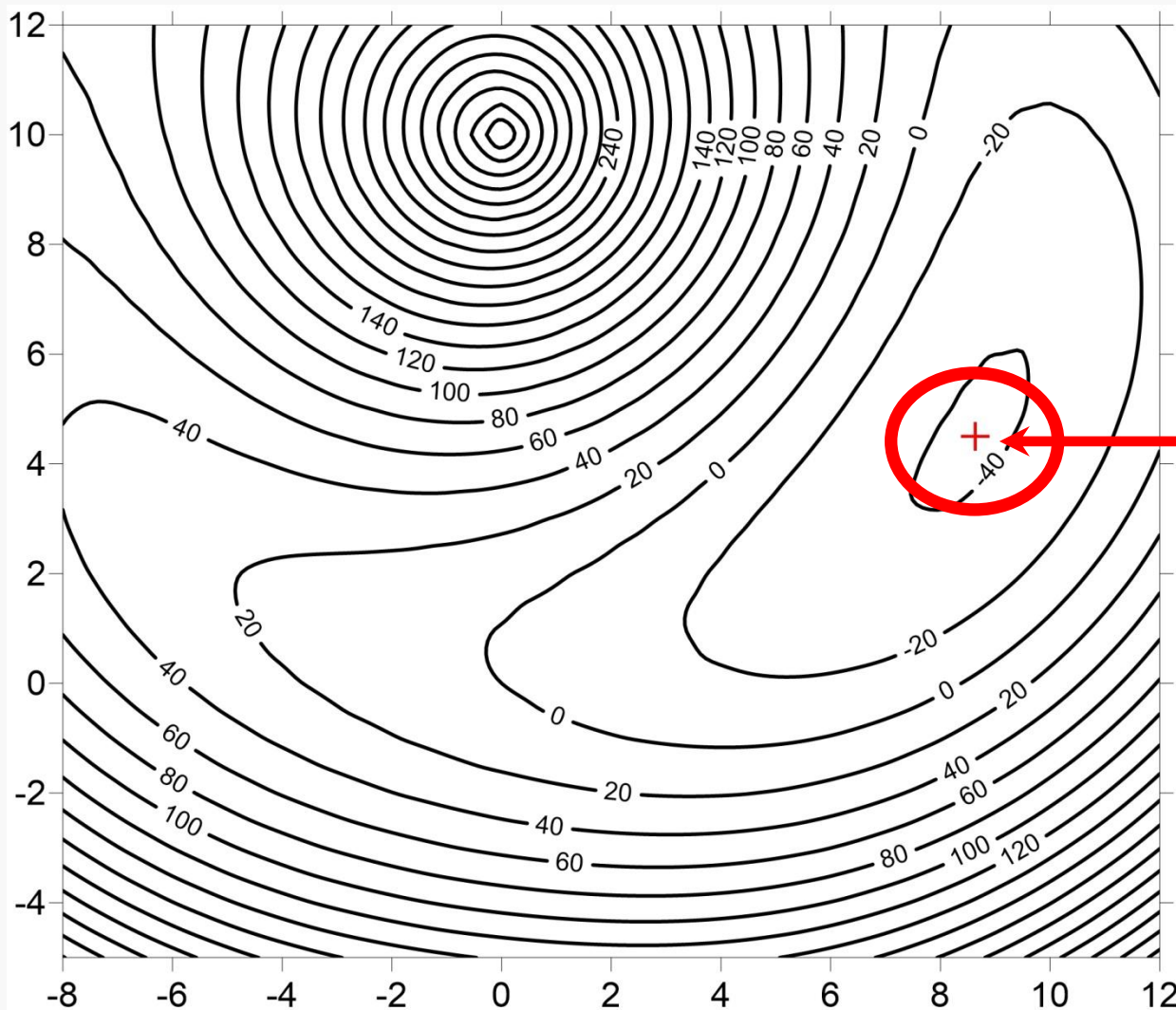


Politechnika Warszawska



OMNIS Przykład (4/6)

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.



Rozwiązanie:

**Minimum energii
potencjalnej jest
osiągalne dla:**

$$X_1 = 8.631 \text{ [cm];}$$

$$X_2 = 4.533 \text{ [cm];}$$

$$PE = -41.81 \text{ [N cm];}$$

**Graficzna interpretacja
zagadnienia układu
dwóch sprężyn**



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



Politechnika Warszawska



OMNIS

Metody poszukiwania minimum

Rodzaje metod omawianych podczas wykładu

- Generalna zasada optymalizacji (strategia ogólna);
- Metody zerowego rzędu (zero-order method);
- Metody pierwszego rzędu (first-order method);
- Metody drugiego rzędu (second-order method);





OMNIS Generalna strategia optymalizacji (1/6)

Traktujemy algorytm optymalizacyjny jako algorytm iteracyjny:

$$\mathbf{X}^q = \mathbf{X}^{q-1} + \alpha^* \mathbf{S}^q$$

gdzie:

$\mathbf{X}^q$ – macierz kolumnowa zmiennych w q -tej iteracji;

$\mathbf{X}^{q-1}$ – macierz kolumnowa zmiennych w $q-1$ -szej iteracji;

q – numer iteracji;

$\mathbf{S}^q$ – wektor kierunku poszukiwań;

α^* – odległość w kierunku wektora $\mathbf{S}$ którą zamierzamy przebyć w q -tej iteracji;



OMNIS Generalna strategia optymalizacji (2/6)

Przykład:

Dla układu sprężyn zakładamy wektor kierunku poszukiwań:

$$S^1 = \begin{Bmatrix} -1.0 \\ -1.0 \end{Bmatrix} \quad \text{oraz punkt początkowy:} \quad X^0 = \begin{Bmatrix} -4.0 \\ 4.0 \end{Bmatrix}$$

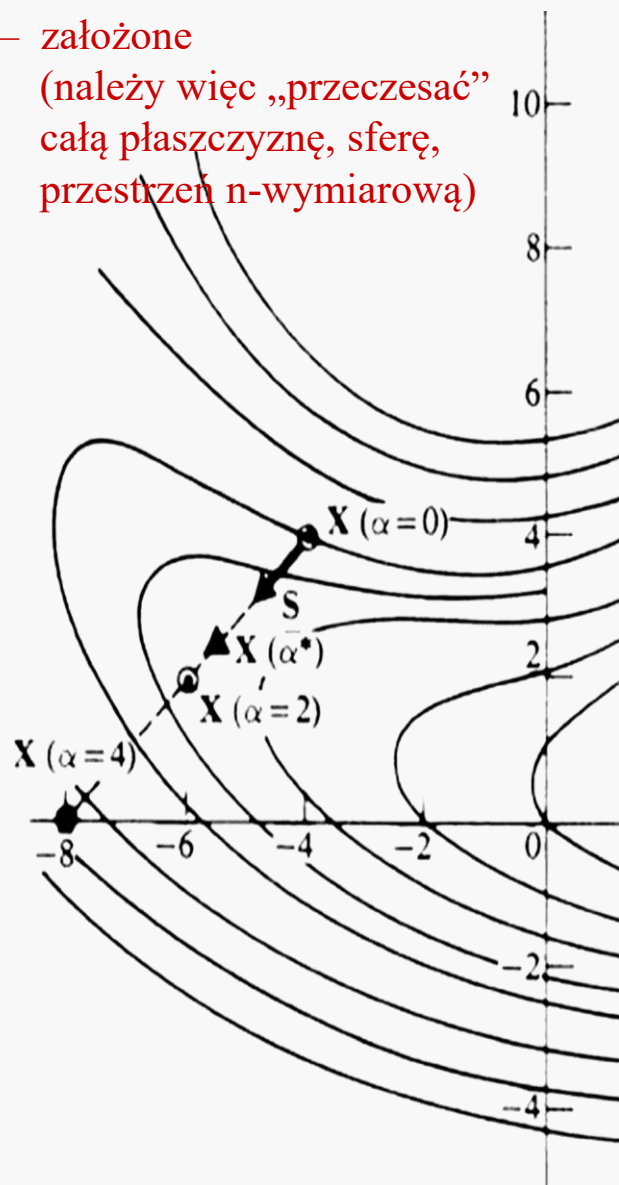
Otrzymujemy zagadnienie na jedną niewiadomą α^* :

$$\text{Dla } \alpha^* = 2.0 \quad X = \begin{Bmatrix} -4.0 \\ 4.0 \end{Bmatrix} + 2.0 \begin{Bmatrix} -1.0 \\ -1.0 \end{Bmatrix} = \begin{Bmatrix} -6.0 \\ 2.0 \end{Bmatrix}$$

$$\text{Dla } \alpha^* = 4.0 \quad X = \begin{Bmatrix} -4.0 \\ 4.0 \end{Bmatrix} + 4.0 \begin{Bmatrix} -1.0 \\ -1.0 \end{Bmatrix} = \begin{Bmatrix} -8.0 \\ 0.0 \end{Bmatrix}$$



S – założone
(należy więc „przeczesać”
całą płaszczyznę, sferę,
przestrzeń n-wymiarową)



Dla trzech wartości α^* wyznaczamy X_1 i X_2 oraz energię potencjalną układu PE

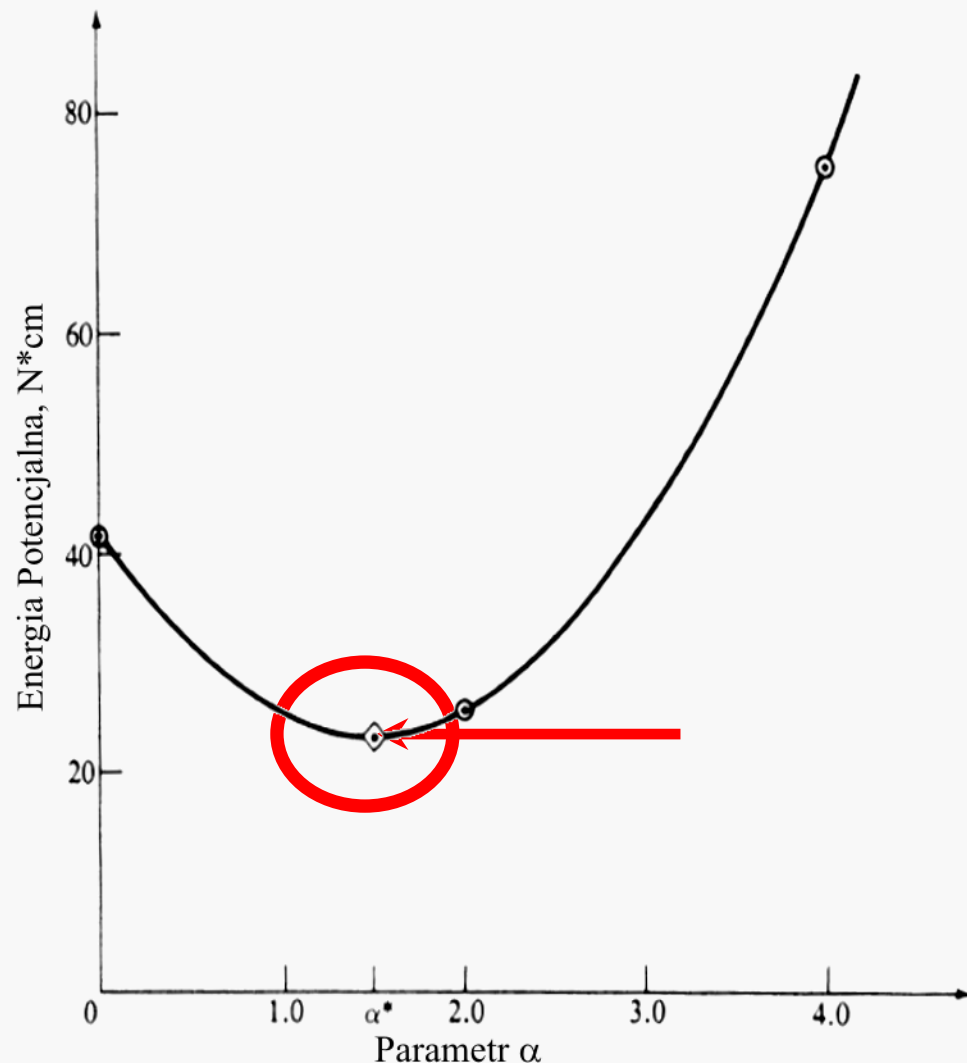
α	X_1	X_2	PE
0.00	-4.00	4.00	41.51
2.00	-6.00	2.00	25.84
4.00	-8.00	0.00	75.44
$\alpha^* =$	1.48	-5.48	23.64



OMNIS

Generalna strategia optymalizacji (4/6)

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.



**Energia potencjalna PE
układu dwóch sprężyn w
funkcji parametru α**

Aproksymacja funkcji w celu
znalezienia wartości
parametru $\alpha = \alpha^*$
odpowiadającego minimum
energii potencjalnej układu



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



Politechnika Warszawska



OMNIS

Generalna strategia optymalizacji (5/6)

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.

Opisany na poprzednich slajdach proces jest jedynie pojedynczą iteracją !

W celu znalezienia rozwiązania należy ponownie wyznaczyć wektor kierunku i powtórzyć proces



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

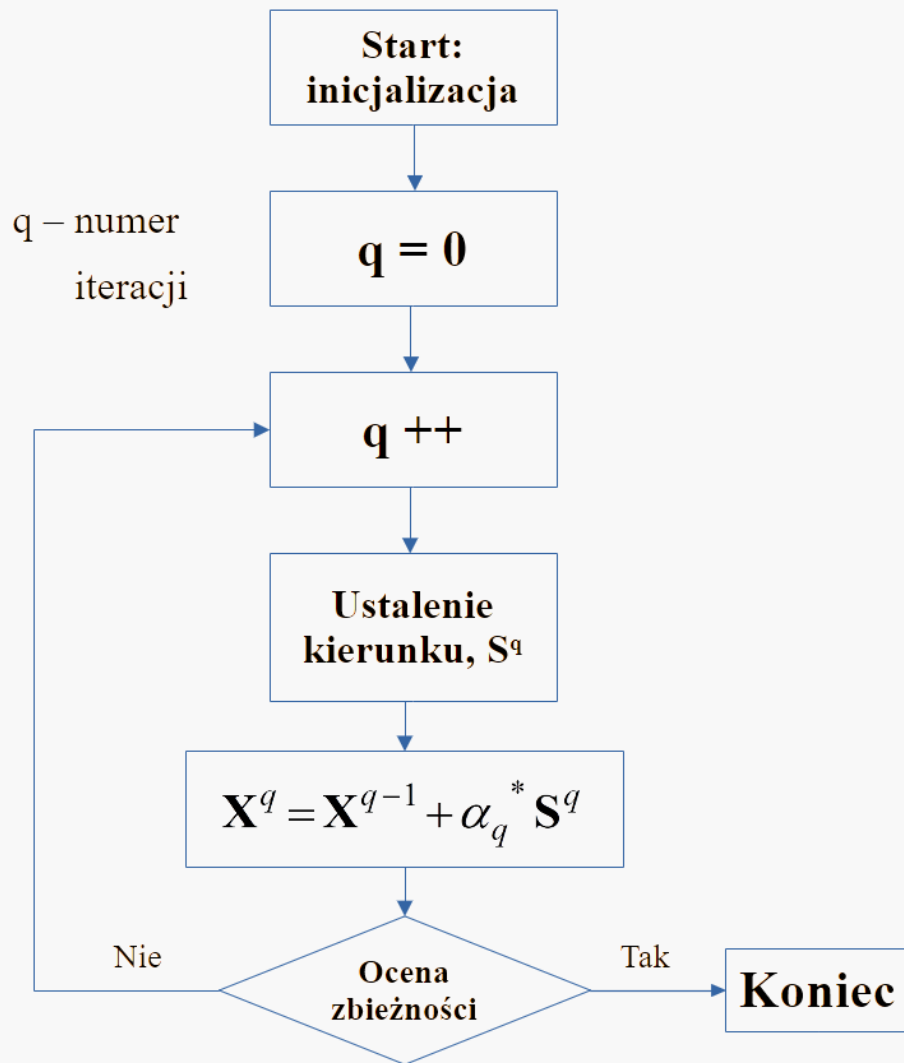
Dofinansowane przez
Unię Europejską



Politechnika Warszawska



OMNIS Generalna strategia optymalizacji (6/6)



Algorytm

Cel:

1. Wyznaczenie kierunku poszukiwań rozwiązania;
2. Poszukiwanie rozwiązania w wyznaczonym kierunku (zagadnienie z jedną niewiadomą – patrz poprzedni wykład);
3. Określenie czy proces jest zbieżny w kierunku akceptowalnego rozwiązania;





OMNIS Metody numeryczne optymalizacji

Cel metody:

- znaleźć optimum w sposób „optymalny”
- szukamy metody, która pozwoli na otrzymanie optymalnego rozwiązania przy minimum kosztów (czasu) obliczeniowego
- metody rzędu:
 - zerowego (bezgradientowe)
 - pierwszego
 - drugiego
 - ...
- czym wyższy rząd, tym zwykle potrzebna mniejsza liczba iteracji ale nie jest to jednoznaczne z kosztem
- czym wyższy rząd metody, tym więcej danych potrzebujemy





OMNIS Metody zerowego rzędu (1/22)

Zalety

metody niezawodne i proste w programowaniu

Wady

wymaga dużej liczby iteracji w celu osiągnięcia rozwiązania, nawet dla prostych zagadnień

Wnioski

metody użyteczne w przypadku zagadnień o małym koszcie obliczeniowym





OMNIS Metody zerowego rzędu (2/22)

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.

Omawiane metody zerowego rzędu

- Metoda dowolnego szukania
- Metoda Hooke'a-Jeevesa
- Metoda sympleks Nelder-Meada
- Metoda Powella



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



Politechnika Warszawska



Metoda Dowolnego Szukania jest najmniej efektywna ale najprostsza w zastosowaniu

Metoda wymaga założenia obszaru poszukiwań rozwiązania w postaci górnej U i dolnej L granicy.

i -tą zmienną X poszukujemy za pomocą wzoru:

$$X_i^q = X_i^L + r(X_i^U - X_i^L)$$

gdzie:

X^U - granica górna; X^L - granica dolna; $X^L \leq X \leq X^U$

r - dowolna liczba rzeczywista z przedziału od 0 do 1;

q - numer iteracji;



Zastosowanie metody Dowolnego Szukania do zagadnienia układu sprężyn (slajd nr 4)

q	X_1	X_2	PE
0	-7.792	13.244	668.46
1	-10.405	18.374	209.69
2	3.893	17.987	62.22
8	-2.001	-0.631	15.92
14	10.050	9.585	-26.01
185	8.088	5.102	-39.33
214	8.989	5.150	-41.51
1737	8.914	5.081	-41.58
4219	8.867	4.630	-41.69

Rozwiązanie:

$X_1 = 8.631$ [cm];

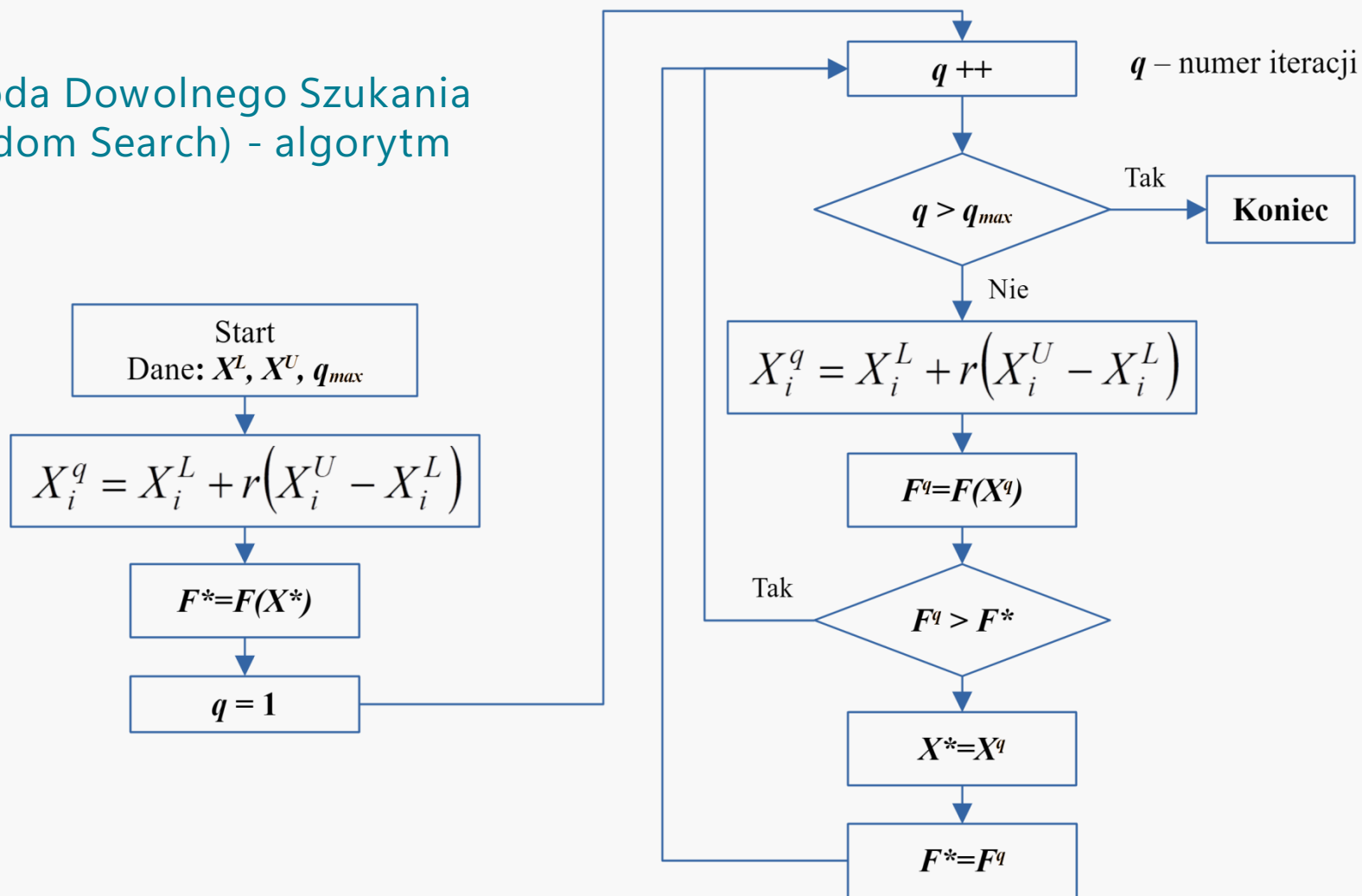
$X_2 = 4.533$ [cm];

$PE = -41.81$ [N cm];



OMNIS Metody zerowego rzędu (5/22)

Metoda Dowolnego Szukania (Random Search) - algorytm





OMNIS

Metody zerowego rzędu (6/22)

Metoda Hooke'a-Jeevesa

Prosta metoda bezpośredniego szukania

Metoda iteracyjna gdzie kierunki poszukiwania są zgodne z kierunkami osi układu współrzędnych

Dwa etapy metody

- etap próbny
- etap roboczy





OMNIS Metody zerowego rzędu (7/22)

Metoda Hooke'a-Jeevesa

Prosta metoda bezpośredniego szukania – etap próbny

1. x_0 – punkt startowy, e – początkowa długość kroku, $j=1$, $F_0 = F(x_0)$
2. wykonaj krok próbny w kierunku x_j : $x_j = x_{j-1} + e$ oraz oblicz $F(x_j)$
3. jeżeli $F < F_0$ to $F_0=F$; przejdź do p. 6
w przeciwnym razie
4. wykonaj krok próbny w przeciwnym kierunku : $x_j = x_{j-1} - 2e$, oblicz $F(x_j)$ i podstaw pod F_0
5. jeżeli krok ten był pomyślny podstaw F w miejsce F_0 . Jeżeli nie, to $j = j+1$ i wróć do (2.)
6. zbadaj, czy $j = n$ (wykonano kroki we wszystkich kierunkach bazy ortogonalnej).
Jeśli nie, to $j = j + 1$ i wróć do (2.). Jeżeli tak to:
7. Zbadaj, czy w wykonanym cyklu wystąpiły kroki pomyślne. Jeżeli tak, podstaw x_j jako punkt bazowy x^B i wykonaj etap roboczy . Jeżeli nie:
8. Jeżeli nie minimum, to: przy pierwszej iteracji zmień punkt startowy, a przy następnych zmniejsz długość kroku i wróć do poprzedniego punktu bazowego i do (1.)



OMNIS

Metody zerowego rzędu (8/22)

Metoda Hooke'a-Jeevesa

Prosta metoda bezpośredniego szukania – etap roboczy

1. wykonaj krok roboczy: $x^0 = x^B + (x^B - x^{B0}) = 2x^B - x^{B0}$
2. podstaw x^B w miejsce x^{B0} i wróć do etapu próbnego

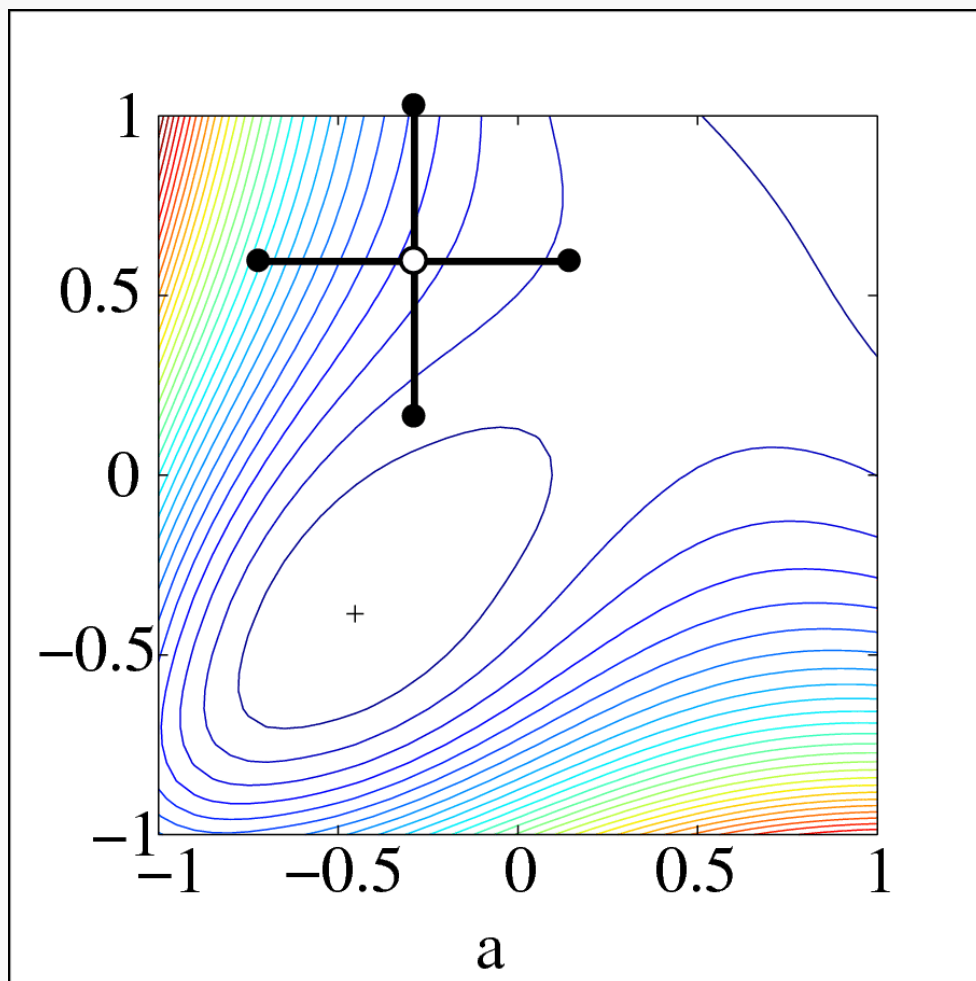




OMNIS

Metody zerowego rzędu (9/22)

Metoda Hooke'a-Jeevesa



Przykład zbiegania się
rozwiązania metodą
Hooke'a-Jeevesa



OMNIS

Metody zerowego rzędu (10/22)

Metoda sympleks Nelder-Meada

Sympleks n -wymiarowy – najmniejszy zbiór wypukły $n+1$ punktów takich, że:

wektory $p_j - p_0$ ($1 \leq j \leq n$) są liniowo niezależne

w przestrzeni 2-wymiarowej – dowolny trójkąt

w przestrzeni 3-wymiarowej – dowolny czworościan



OMNIS

Metody zerowego rzędu (11/22)

Metoda sympleks Nelder-Meada

P_i – i -ty punkt sympleksu ($i=1, 2, \dots, n+1$)

P_h – wybrany punkt sympleksu, w którym wartość funkcji osiąga największą wartość (najgorszy)

P_l – wybrany punkt sympleksu, w którym wartość funkcji osiąga najmniejszą wartość (najlepszy)

P' – środek symetrii sympleksu liczony z wyłączeniem punktu P_h

Operacje:

odbicie punktu P_h względem P'

$$P^* = (1 + \alpha) P' - \alpha P_h$$

ekspansja P^* względem P'

$$P^{**} = (1 + \gamma) P^* - \gamma P'$$

kontrakcja punktu P_h względem P'

$$P^{***} = \beta P_h + (1 - \beta) P'$$

redukcja

$$P_i = (P_i + P_l) / 2, i = 1, 2, \dots, n+1$$





Metoda sympleks Neldera-Meada

1. Utworzenie sympleksu o $n+1$ wierzchołkach
2. Obliczenie wartości funkcji w punktach wierzchołkowych sympleksu
 $F_i = f(P_i), i = 1, 2, \dots, n+1$
3. Wyznaczenie h i l takich, że $f(P_h) = \max$ z wartości F_i , $f(P_l) = \min$ z F_i , dla $i = 1, 2, \dots, n+1$
4. Obliczenie środka symetrii sympleksu z wyłączeniem punktu P_h

$$P' = \frac{\sum_{i=1}^{n+1} P_i}{n}, i \neq h$$

oraz wartości funkcji $F_s = f(P')$

5. Odbicie P^* punktu P_h względem P' i obliczenie wartości funkcji $F_o = f(P^*)$



OMNIS Metody zerowego rzędu (13/22)

Metoda sympleks Neldera-Meada

Jeżeli $F_0 < \min$, to:

6. Obliczenie $P^{**} = (1 + \gamma) P^* - \gamma P'$ i $F_e = f(P^{**})$

- jeżeli $F_e < \min$, podstaw za P_h punkt P^{**} ,
- jeżeli $F_e \geq \min$, podstaw za P_h punkt P^*

7. O ile nie jest spełnione kryterium stopu, idź do kroku 2.

Jeżeli $F_0 > \min$, to:

8. Jeżeli $F_0 \geq f(P_i)$ dla $i = 1, 2, \dots, n+1, i \neq h$

- jeżeli $F_0 \geq \max$, to idź do następnego kroku
- jeżeli $F_0 < \max$, to podstaw za P_h punkt P^*

9. Wykonaj kontrakcję P^{***} punktu P_h względem P' i oblicz $F_k = f(P^{***})$

- jeżeli $F_k \geq \max$, to wykonaj redukcję sympleksu
- jeżeli $F_k < \max$, to podstaw za P_h punkt P^{***} i idź do kroku 11.



Metoda sympleks Neldera-Meada

11. Jeżeli $F_0 < f(P_i)$ dla $i = 1, 2, \dots, n+1, i \neq h$, za P_h punkt P^*

12. Jeżeli warunek stopu nie jest spełniony, idź do kroku 2.

Warunek stopu może być sformułowany na kilka sposobów, np.:

- odległości pomiędzy wszystkimi punktami są mniejsze od ε
- odchylenie standardowe funkcji obliczonej w $n+1$ wierzchołkach sympleksu nie przekracza określonej wartości

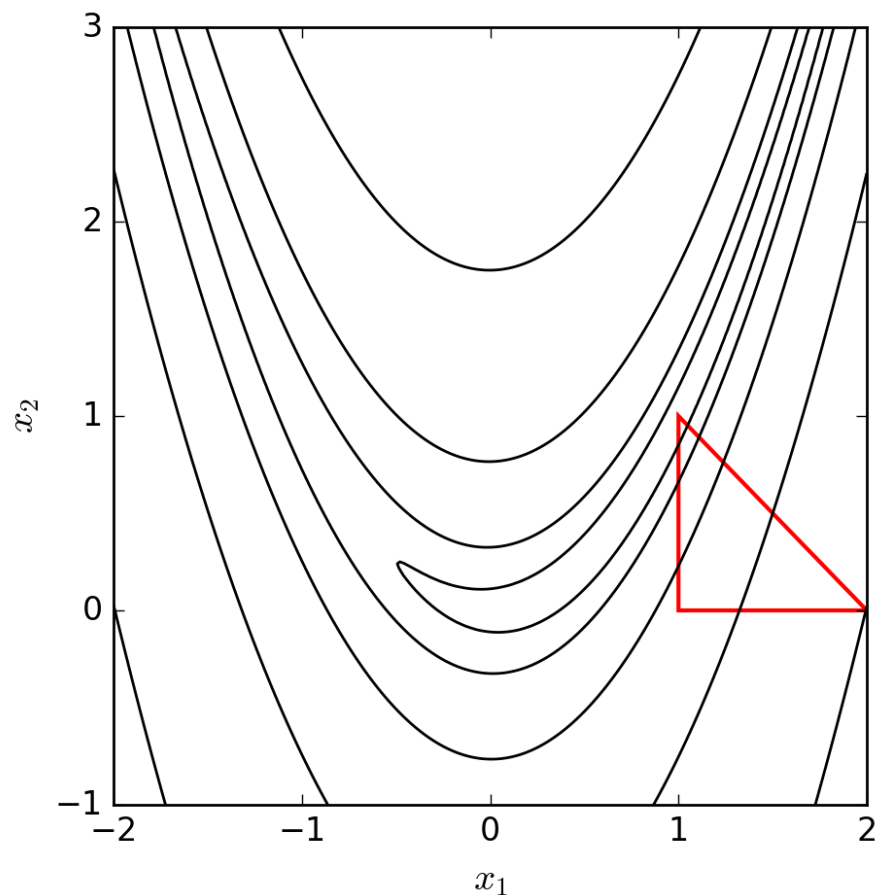
$$Q = \sqrt{\frac{\sum_{i=1}^{n+1} (f(P_i) - f(P_0))^2}{n+1}} \leq \varepsilon$$



OMNIS

Metody zerowego rzędu (15/22)

Metoda sympleks Neldera-Meada



Przykład zbiegania się
rozwiązania metodą
Neldera-Meada

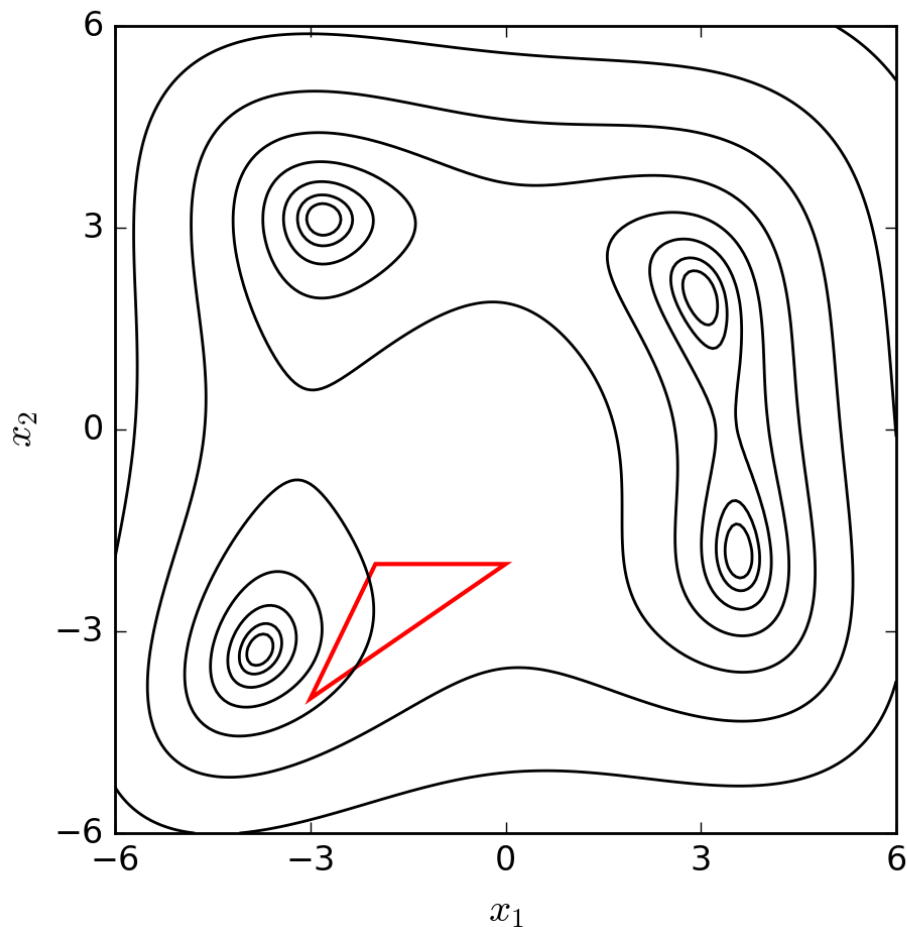


OMNIS

Metody zerowego rzędu (16/22)

Metoda sympleks Neldera-Meada

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.



Przykład zbiegania się
rozwiązania metodą
Neldera-Meada



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



Politechnika Warszawska

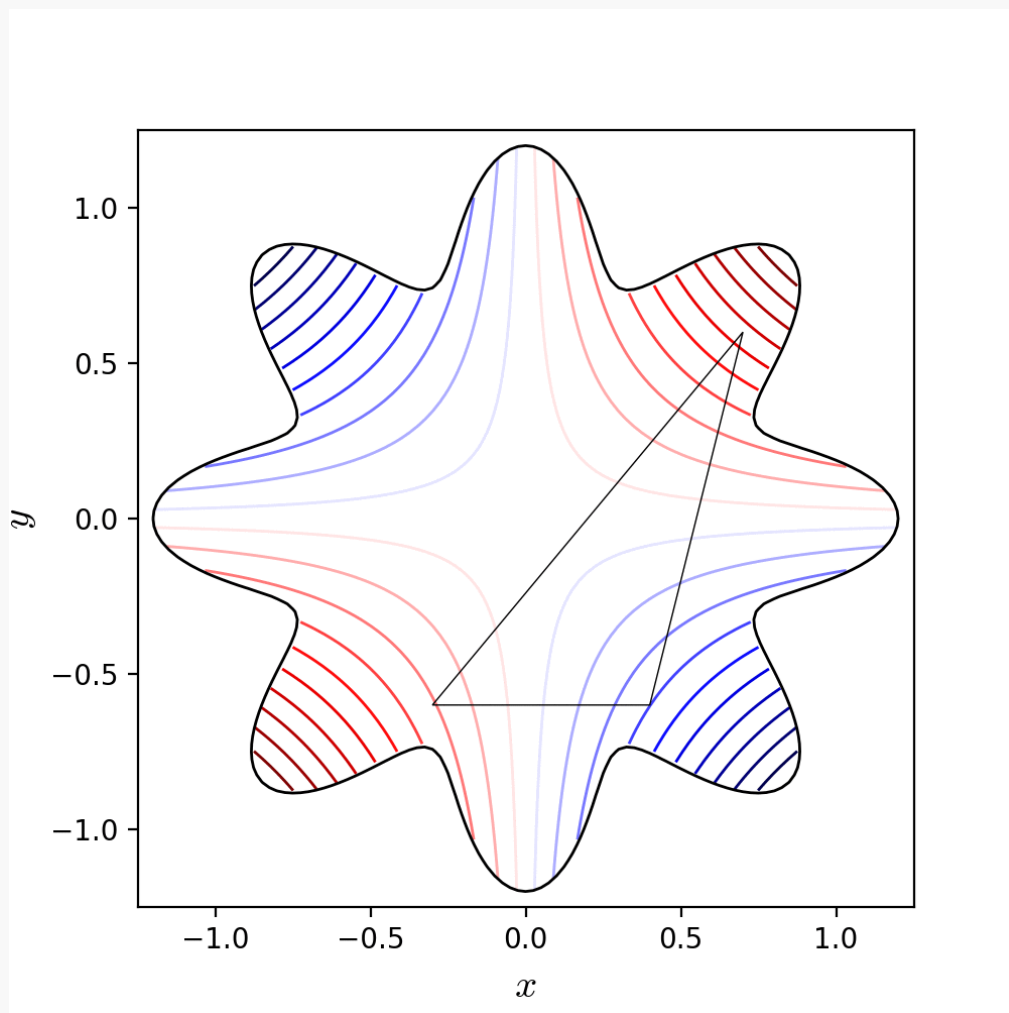


OMNIS

Metody zerowego rzędu (17/22)

Metoda sympleks Nelder-Meada

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.



Przykład zbiegania się
rozwiązania metodą
Nelder-Meada



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



Politechnika Warszawska



Metoda Powell'a jest najefektywniejszą i najpopularniejszą metodą zerowego rzędu.

Metoda bazuje na koncepcji sprzężonych kierunków wg której wektory kierunku S^i i S^j są sprzężone jeśli spełniają warunek:

$$\left(S^i\right)^T H S^j = 0$$

gdzie:

$$H = \begin{bmatrix} \frac{\partial^2 F(\mathbf{X})}{\partial X_1^2} & \frac{\partial^2 F(\mathbf{X})}{\partial X_1 \partial X_2} & \dots & \frac{\partial^2 F(\mathbf{X})}{\partial X_1 \partial X_n} \\ \frac{\partial^2 F(\mathbf{X})}{\partial X_2 \partial X_1} & \frac{\partial^2 F(\mathbf{X})}{\partial X_2^2} & \dots & \frac{\partial^2 F(\mathbf{X})}{\partial X_2 \partial X_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F(\mathbf{X})}{\partial X_n \partial X_1} & \frac{\partial^2 F(\mathbf{X})}{\partial X_n \partial X_2} & \dots & \frac{\partial^2 F(\mathbf{X})}{\partial X_n^2} \end{bmatrix}$$

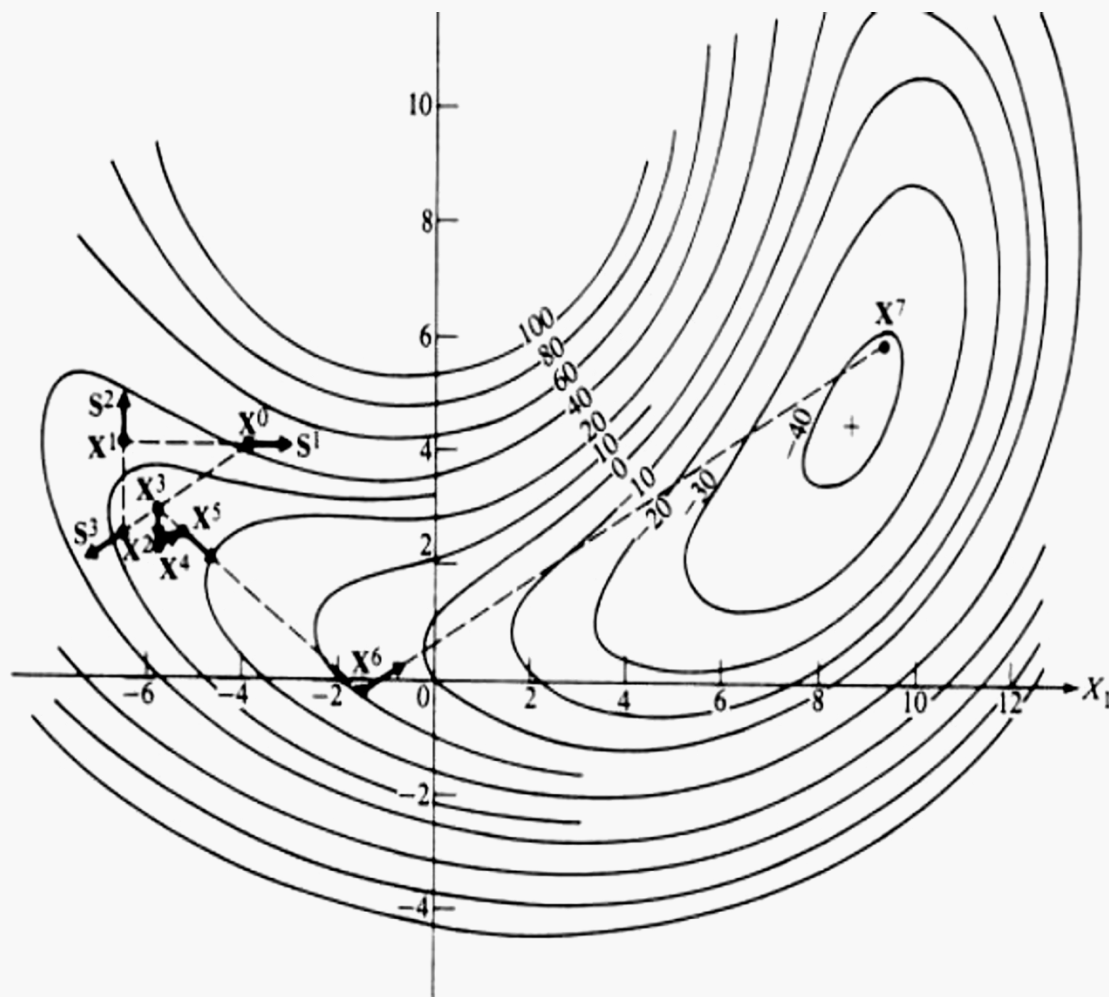
**Metoda Powella jest
metodą bezgradientową !!!**



OMNIS

Metody zerowego rzędu (19/22)

Metoda Powell'a



**Geometryczna
interpretacja
Metody Powell'a**

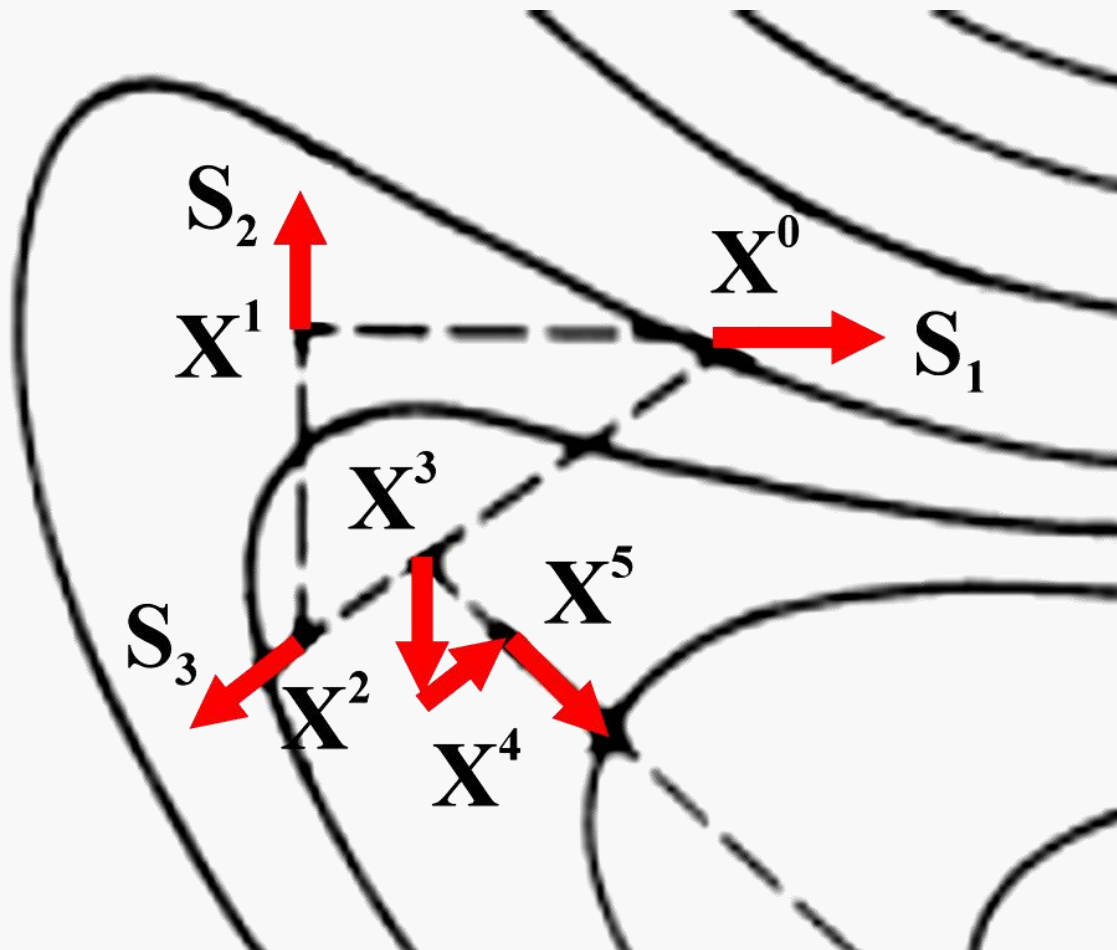




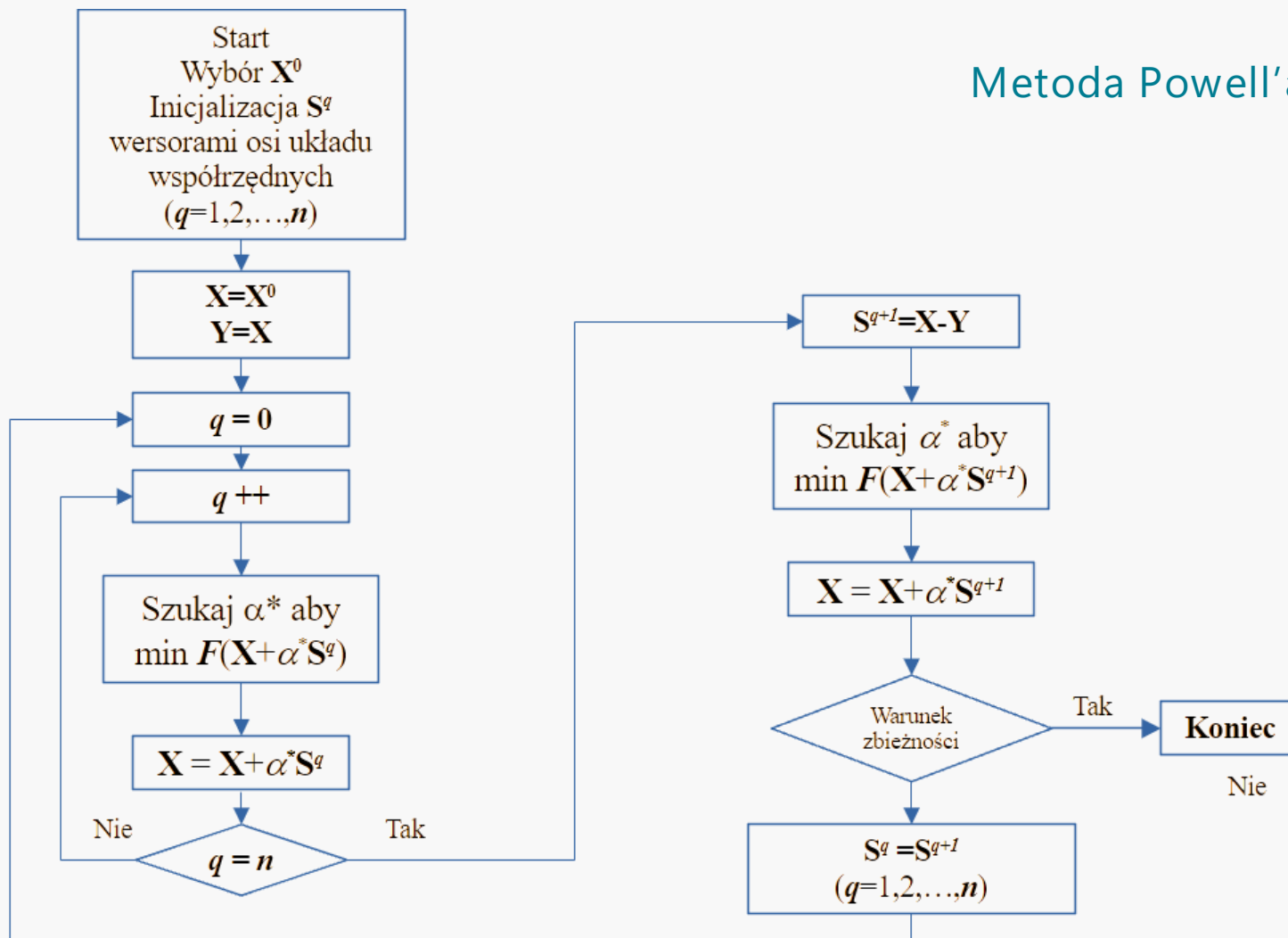
OMNIS

Metody zerowego rzędu (20/22)

Metoda Powell'a



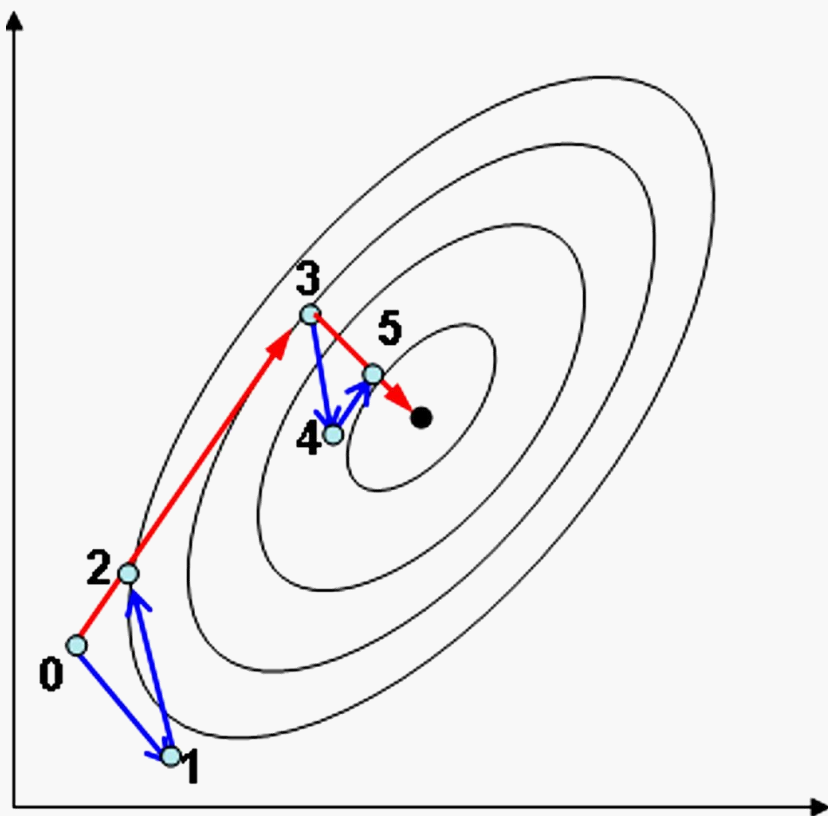
**Geometryczna
interpretacja
Metody Powell'a
szczegóły**





OMNIS Metody zerowego rzędu (22/22)

Metoda Powell'a



Mimo, że metoda Powell'a jest jedną z najefektywniejszych metod zerowego rzędu, może się zdarzyć, że metoda nie pozwoli na uzyskanie dobrego wyniku:

- Jeśli kierunek poszukiwań nie powoduje poprawy, może nie być sprzężenia z kolejnym kierunkiem
- Czasami kierunki poszukiwań mogą stać się równoległe po kilku iteracjach



OMNIS

Metody pierwszego rzędu (1/10)

Zalety

metody efektywniejsze niż metody zerowego rzędu

Wady

większa liczba danych potrzebnych do rozpoczęcia obliczeń (wymagają znajomości gradientów)





OMNIS Metody zerowego rzędu (2/10)

Omawiane metody pierwszego rzędu

- Metoda największego spadku (Steepest Descent)
- Metoda sprzężonych kierunków (Conjugate Direction)
- Metoda zmiennej metryki (Variable Metric Method)
 - Metoda Davidon-Fletcher-Powell (DFP)
 - Broydon-Fletcher-Goldfarb-Shanno (BFGS)



OMNIS

Metody pierwszego rzędu (3/10)

Metoda Największego Spadku (Steepest Descent)

Metoda Największego Spadku jest najbardziej znana ale i najmniej efektywna z metod pierwszego rzędu

W metodzie tej wektor kierunku poszukiwań S^q wyznaczamy za pomocą wzoru:

$$S^q = -\nabla F(X^q)$$

gdzie:

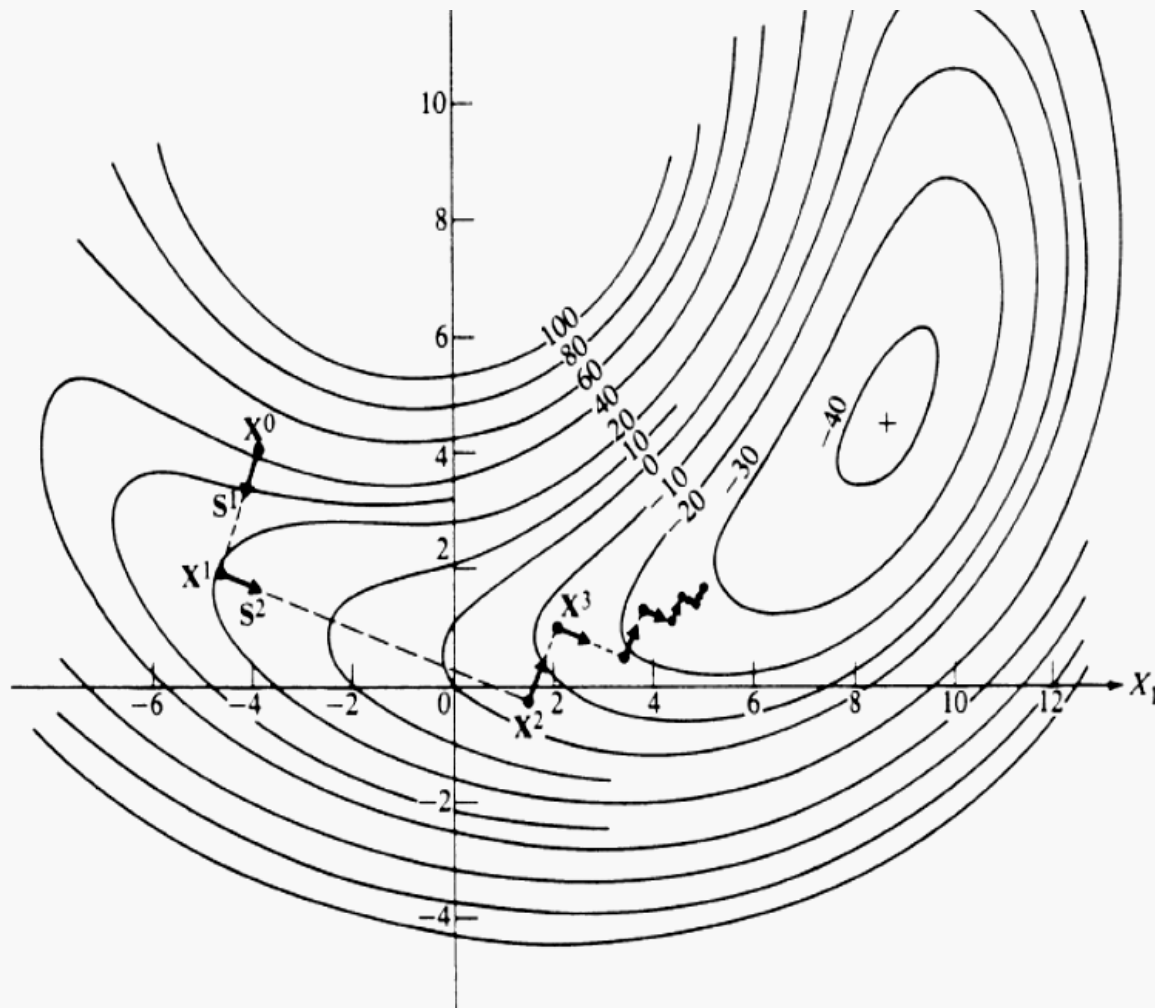
$$\nabla F(X) = \begin{Bmatrix} \partial F(X) / \partial X_1 \\ \partial F(X) / \partial X_2 \\ \dots \\ \partial F(X) / \partial X_n \end{Bmatrix}$$



OMNIS

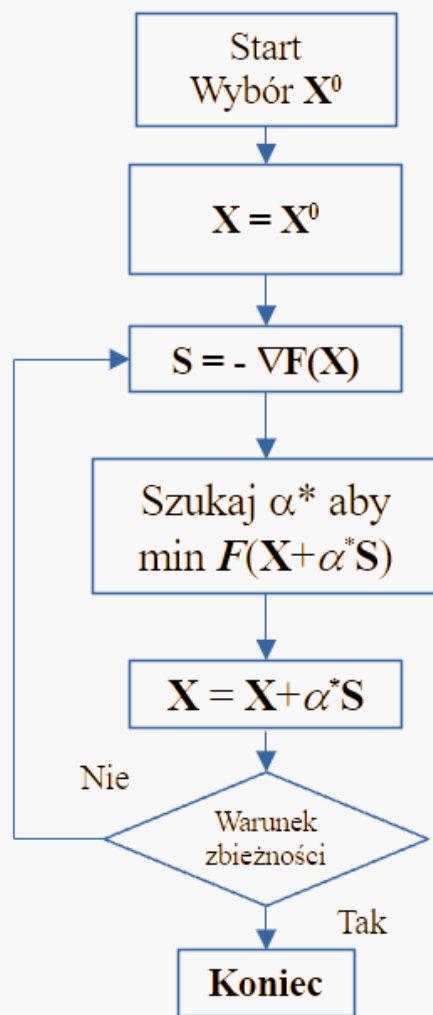
Metody pierwszego rzędu (4/10)

Metoda Największego Spadku (Steepest Descent)



**Geometryczna
interpretacja
Metody
Największego
Spadku**





Metoda Największego Spadku
(Steepest Descent) - algorytm



OMNIS

Metody pierwszego rzędu (6/10)

Metoda sprzężonych kierunków
(The Conjugate Direction method - Fletcher & Reeves)

Metoda wprowadzająca niewielkie modyfikacje do metody Największego Spadku, która znacząco przyspiesza zbieżność procesu obliczeniowego

Wektor kierunku poszukiwań w q -tej iteracji wyznaczamy za pomocą wzoru:

$$S^q = -\nabla F(X^q) + \beta_q S^{q-1}$$

gdzie:
$$\beta_q = \frac{|\nabla F(X^q)|^2}{|\nabla F(X^{q-1})|^2}$$

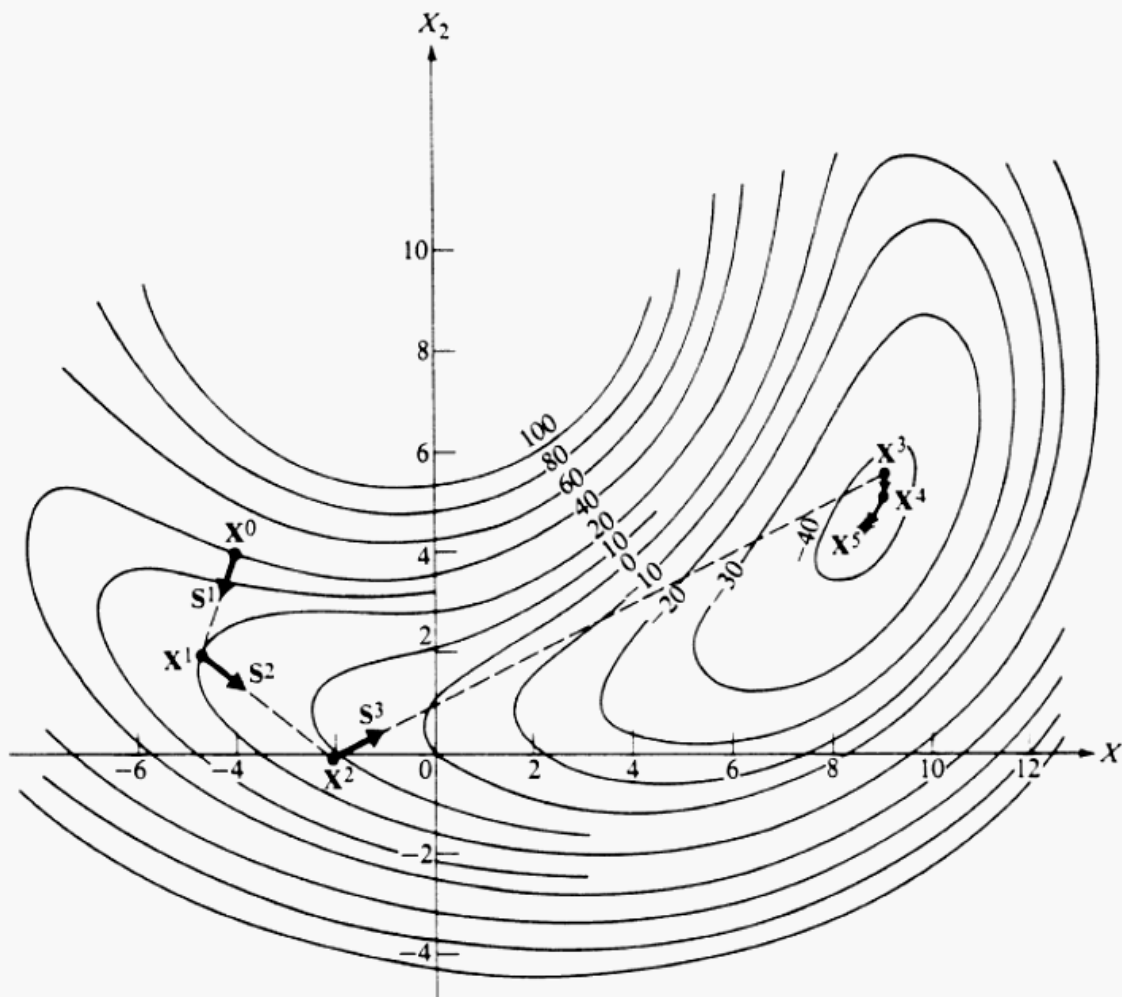
**Element przyspieszający
proces, zawierający
informację o poprzedniej
iteracji**



OMNIS

Metody pierwszego rzędu (7/10)

Metoda sprzężonych kierunków (The Conjugate Direction)



**Geometryczna
interpretacja
Metody
sprzężonych
kierunków**





Metoda zmiennej metryki (**Variable Metric Method**)

Metoda najefektywniejsza z prezentowanych, wykorzystująca do procesu obliczeniowego informacje z poprzednich iteracji w postaci macierzy

H – odwrotna macierz Hessego

Poszukujemy wektora kierunku za pomocą wzoru:

$$S^q = -[H] \nabla F(X^q)$$

Jako warunki początkowe przyjmuje się macierz $H=I$. W następnej iteracji macierz H wyznacza się ze wzoru:

$$[H]^{q+1} = [H]^q + [D]^q$$



Gdzie:

D - macierz symetryczna wyznaczana za pomocą wzoru:

$$D^q = \frac{\sigma + \theta\tau}{\sigma^2} pp^T + \frac{\theta - 1}{\tau} H^q y (H^q y)^T - \frac{\theta}{\sigma} [H^q y p^T + p (H^q y)^T]$$

Wektory: $p = X^q - X^{q-1}$

$$y = \nabla F(X^q) - \nabla F(X^{q-1})$$

Skalary: $\sigma = p \cdot y$

$$\tau = y^T H^q y$$



OMNIS

Metody pierwszego rzędu (10/10)

Metoda zmiennej metryki (**Variable Metric Method**)

Dwie odmiany metody:

Davidon-Fletcher-Powell (DFP) (*starsza*)

$$\theta = 0$$

Broydon-Fletcher-Goldfarb-Shanno (BFGS)
(*najpopularniejsza i najskuteczniejsza*)

$$\theta = 1$$





OMNIS Metody drugiego rzędu (1/4)

Metoda Newton'a

Jest to klasyczna metoda drugiego rzędu wyprowadzana z szeregu Taylor'a:

$$F(X) \approx F(X^q) + \nabla F(X) \cdot \delta X + \frac{1}{2} \delta X^T \cdot H(X^q) \cdot \delta X$$

gdzie:

H - Hesjan

$$\delta X = X^{q+1} - X^q$$

$$\delta X = -[H(X^q)]^{-1} \nabla F(X^q)$$

$$H = \begin{bmatrix} \frac{\partial^2 F(\mathbf{X})}{\partial X_1^2} & \frac{\partial^2 F(\mathbf{X})}{\partial X_1 \partial X_2} & \dots & \frac{\partial^2 F(\mathbf{X})}{\partial X_1 \partial X_n} \\ \frac{\partial^2 F(\mathbf{X})}{\partial X_2 \partial X_1} & \frac{\partial^2 F(\mathbf{X})}{\partial X_2^2} & \dots & \frac{\partial^2 F(\mathbf{X})}{\partial X_2 \partial X_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F(\mathbf{X})}{\partial X_n \partial X_1} & \frac{\partial^2 F(\mathbf{X})}{\partial X_n \partial X_2} & \dots & \frac{\partial^2 F(\mathbf{X})}{\partial X_n^2} \end{bmatrix}$$



OMNIS

Metody drugiego rzędu (2/4)

Metoda Newton'a

Wymagania dla metody są następujące:

- Funkcja $F(X)$ jest klasy C^2 czyli funkcja jest ciągła i podwójnie różniczkowalna,
- $F(X)$ jest ściśle wypukła w badanej dziedzinie





Metoda Newton'a

$$X^{q+1} = X^q + \delta X = X^q - [H(X^q)]^{-1} \nabla F(X^q)$$

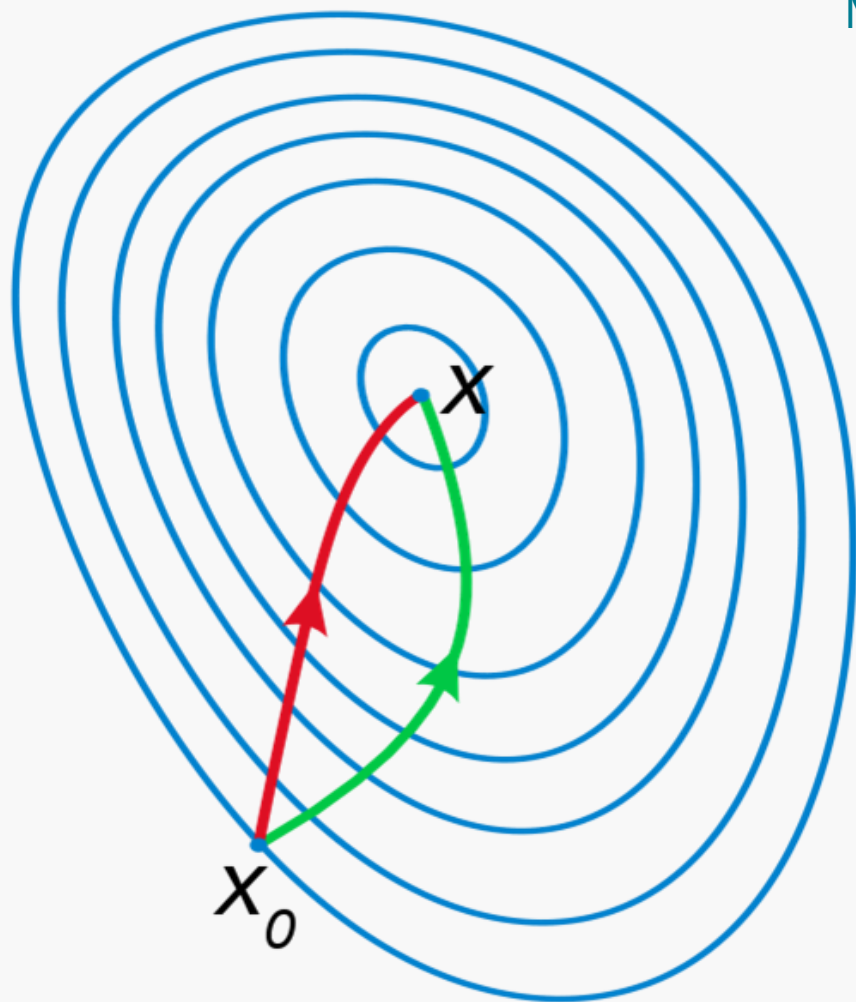
Po porównaniu powyższego równania z generalną strategią optymalizacji $\mathbf{X}^q = \mathbf{X}^{q-1} + \alpha_q^* \mathbf{S}^q$ i założeniu $\alpha^*=1$ otrzymujemy:

$$\mathbf{S}^q = -[H(X^q)]^{-1} \nabla F(X^q)$$



OMNIS Metody drugiego rzędu (4/4)

Metoda Newton'a



Porównanie metody najszybszego spadku (linia zielona) z metodą Newtona (linia czerwona). Na rysunku widać linie poszukiwań minimum dla zadanej funkcji celu. Metoda Newtona używa informacji o krzywiznie w celu zoptymalizowania ścieżki poszukiwań.





OMNIS Porównanie metod optymalizacji (1/5)

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.

Metoda	Pochodne	Koszt iteracji	Zbieżność	Wymagania	Zastosowanie
Newton	Pierwsza i Druga (Hesjan)	Wysoki	Kwadratowa (bardzo szybka)	Funkcja gładka klasy C^2	bardzo małe
Quasi-Newton	Tylko Pierwsza (Gradient)	Średni	Superliniowa (szybka)	Funkcja klasy C^1 różniczkowalna	średnie
Kierunków sprzężonych	Tylko Pierwsza (Gradient)	Średni	Nadliniowa	Funkcja klasy C^1 różniczkowalna	duże
Największy Spadek	Tylko Pierwsza (Gradient)	Średni	Liniowa (wolna)	Funkcja klasy C^1 różniczkowalna	małe-średnie
Powell	Brak	Niski	Liniowa /Superliniowa	Funkcja ciągła	małe



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską

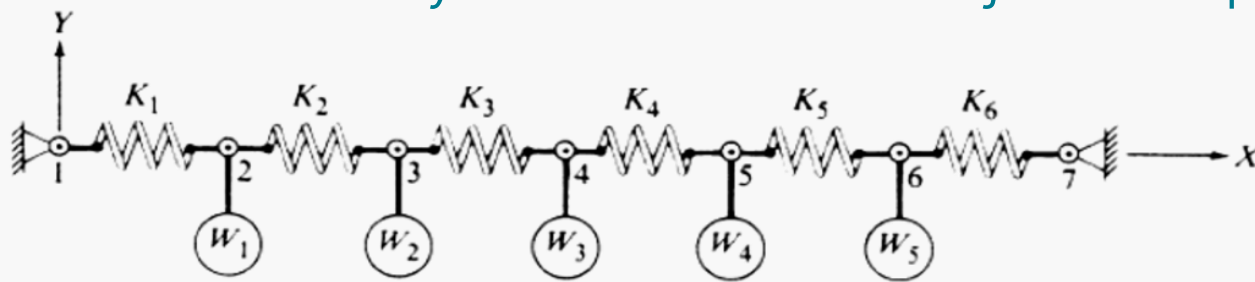


Politechnika Warszawska

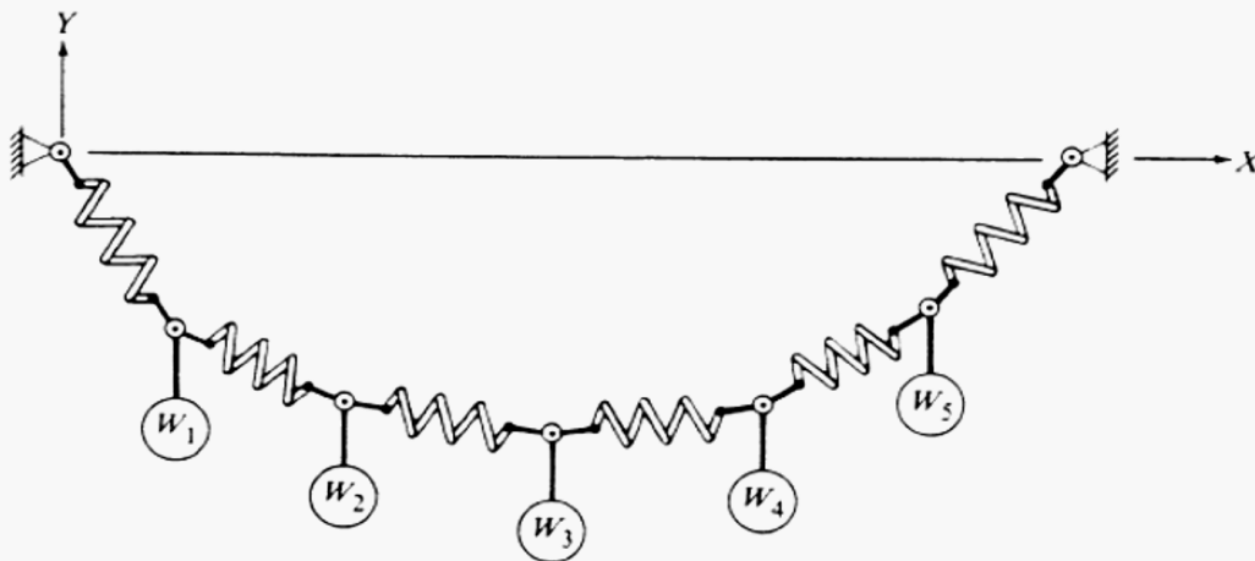


OMNIS Porównanie metod optymalizacji (2/5)

Przykład: Analiza deformacji układu sprężyn i ciężarków



(a) przed deformacją



(b) po deformacji

**Wyznaczyć warunki
równowagi układu
poprzez minimalizację
energii potencjalnej
układu**



OMNIS Porównanie metod optymalizacji (3/5)

Przykład: Analiza deformacji układu sprężyn i ciężarków

Przyrost długości i-tej sprężyny wyrażony jest wzorem:

$$\Delta L_i = \sqrt{(X_{i+1} - X_i)^2 + (Y_{i+1} - Y_i)^2} - L_i^0$$

Energia Potencjalna PE układu wyraża się wzorem:

$$PE = \sum_{i=1}^{N+1} 0.5K_i \Delta L_i^2 + \sum_{j=1}^N W_j Y_j$$

Gdzie:

PE – energia potencjalna układu

K_i – sztywność i-tej sprężyny

W_j – ciężar j-tego ciężarka





OMNIS Porównanie metod optymalizacji (4/5)

Przykład: Analiza deformacji układu sprężyn i ciężarków – porównanie wyników

1. Metoda Powell'a;
2. Metoda Największego Spadku;
3. Metoda sprzężonych kierunków (Fletcher-Reeves);
4. Metoda zmiennej metryki DFP
5. Metoda zmiennej metryki BFGS
6. Metoda Newton'a

Zmienna decyzyjna	Wartość początkowa	Metoda					
		1	2	3	4	5	6
X_1	10.0	10.3	10.3	10.4	10.2	10.2	10.4
X_2	20.0	21.1	20.7	21.1	20.8	20.8	21.1
X_3	30.0	31.7	31.0	31.6	31.4	31.4	31.7
X_4	40.0	42.1	41.3	42.0	41.8	41.7	42.1
X_5	50.0	51.8	51.1	51.6	51.4	51.4	51.8
Y_1	0.0	-4.28	-2.65	-3.96	-4.64	-4.64	-4.28
Y_2	0.0	-7.90	-5.25	-7.77	-8.19	-8.19	-7.90
Y_3	0.0	-9.86	-7.35	-10.2	-10.0	-10.0	-9.86
Y_4	0.0	-9.40	-7.63	-9.52	-9.18	-9.19	-9.40
Y_5	0.0	-6.01	-4.97	-5.79	-5.42	-5.43	-6.01



OMNIS Porównanie metod optymalizacji (5/5)

Przykład: Analiza deformacji układu sprężyn i ciężarków – porównanie iteracji

1. Metoda Powell'a;
2. Metoda Największego Spadku;
3. Metoda sprzężonych kierunków (Fletcher-Reeves);
4. Metoda zmiennej metryki DFP
5. Metoda zmiennej metryki BFGS
6. Metoda Newton'a

Iteracje	Wartość funkcji celu						Uwagi
	1	2	3	4	5	6	
0	0	0	0	0	0	0	
5	0	-223	-1559	-1430	-1432	-2330	
10	-1722	-1958	-3540	-2573	-2572	-3539	
15	-1734	-2355	-4216	-4149	-4145	-4391	
20	-2623	-3964	-4393	-4378	-4378	-4416	wynik po 19 iteracjach
Końcowa	-4416	-3964	-4393	-4378	-4378	-4416	
liczba iteracji	178	40	22	26	26	19	limit liczby iteracji
Liczba obliczeń	465	587	331	383	383	378	



Kryteria zbieżności

- warunek stopu maksymalna liczba iteracji q_{MAX}
- względna i bezwzględna zmiana wartości funkcji F
 - bezwzględna $|F(X^q) - F(X^{q-1})| \leq \varepsilon_A$
 - względna $\frac{|F(X^q) - F(X^{q-1})|}{\max(|F(X^q)|, 10^{-10})} \leq \varepsilon_R$
- warunek Kuhn'a – Tucker'a $|\nabla F(X^q)| \leq \varepsilon_k; \quad \varepsilon_k = 10^{-4}$



OMNIS Optymalizacja w inżynierii lotniczej i kosmicznej

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.

Dziękuję za uwagę



Fundusze Europejskie
dla Rozwoju Społecznego



Rzeczpospolita
Polska

Dofinansowane przez
Unię Europejską



Politechnika Warszawska