



OMNIS

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.

# Optymalizacja w inżynierii lotniczej i kosmicznej

## Wykład 6

Matematyczne podstawy optymalizacji 5

*Wyznaczanie długości kroku - warunki Armijo i Wolfe*

1130-LK000-MSP-1037



Fundusze Europejskie  
dla Rozwoju Społecznego



Rzeczpospolita  
Polska

Dofinansowane przez  
Unię Europejską



Politechnika Warszawska



# OMNIS

## Zagadnienia

- Poszukiwanie optymalnego kroku w metodach gradientowych
  - wstęp
  - warunek Armijo
  - warunek krzywizny
  - warunek Wolfe'a



# OMNIS Generalna strategia optymalizacji

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;

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





- metody numeryczne głównie zajmują się wyznaczeniem kierunku poszukiwań
- metoda Newtona wyznacza zarówno kierunek jak i wartość skoku ale wymaga funkcji klasy  $C^2$
- dotychczas prezentowane metody wyznaczenia kroku to znalezienie minimum funkcji aproksymującej funkcję celu na kierunku  $S$
- jeżeli jest to z jakiejś przyczyn trudne lub niemożliwe można zastosować krok stały ale nie jest to podejście racjonalne z punktu widzenia kosztów obliczeniowych

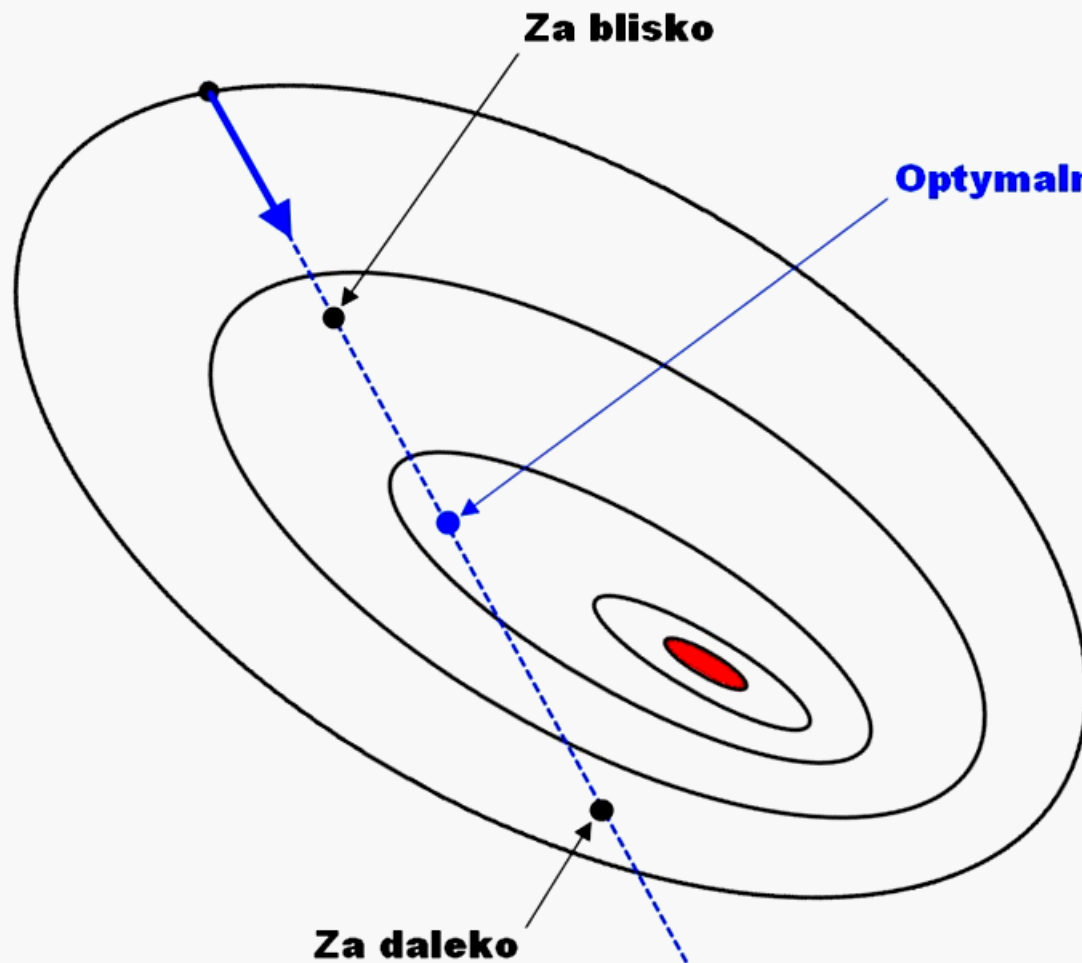




# OMNIS

## Poszukiwanie optymalnego kroku - wstęp

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.



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



Fundusze Europejskie  
dla Rozwoju Społecznego



Rzeczpospolita  
Polska

Dofinansowane przez  
Unię Europejską



Politechnika Warszawska



Warunek zbieżności:

$$f(x_k + \alpha_k \cdot p_k) \leq f(x_k)$$

Warunek Armijo to zasada doboru długości kroku ( $\alpha_k$ ) w gradientowych metodach optymalizacji numerycznej zapewniająca "wystarczający spadek" wartości funkcji celu ( $f(x)$ ). Gwarantuje on, że krok nie jest zbyt duży, co zapewnia zbieżność algorytmu do minimum. Często stosowany z techniką *backtracking*.

$$f(x_k + \alpha_k \cdot p_k) \leq f(x_k) + c_1 \cdot \alpha_k \cdot \nabla f(x_k)^T \cdot p_k$$





Warunek Armijo: dla ustalonego kierunku poszukiwań  $p_k$  i współczynnika  $c_1 \in (0,1)$ , długość kroku  $\alpha_k > 0$  w każdej iteracji powinna spełniać nierówność:

$$f(x_k + \alpha_k \cdot p_k) - f(x_k) \leq c_1 \cdot \alpha_k \cdot \nabla f(x_k)^T \cdot p_k$$

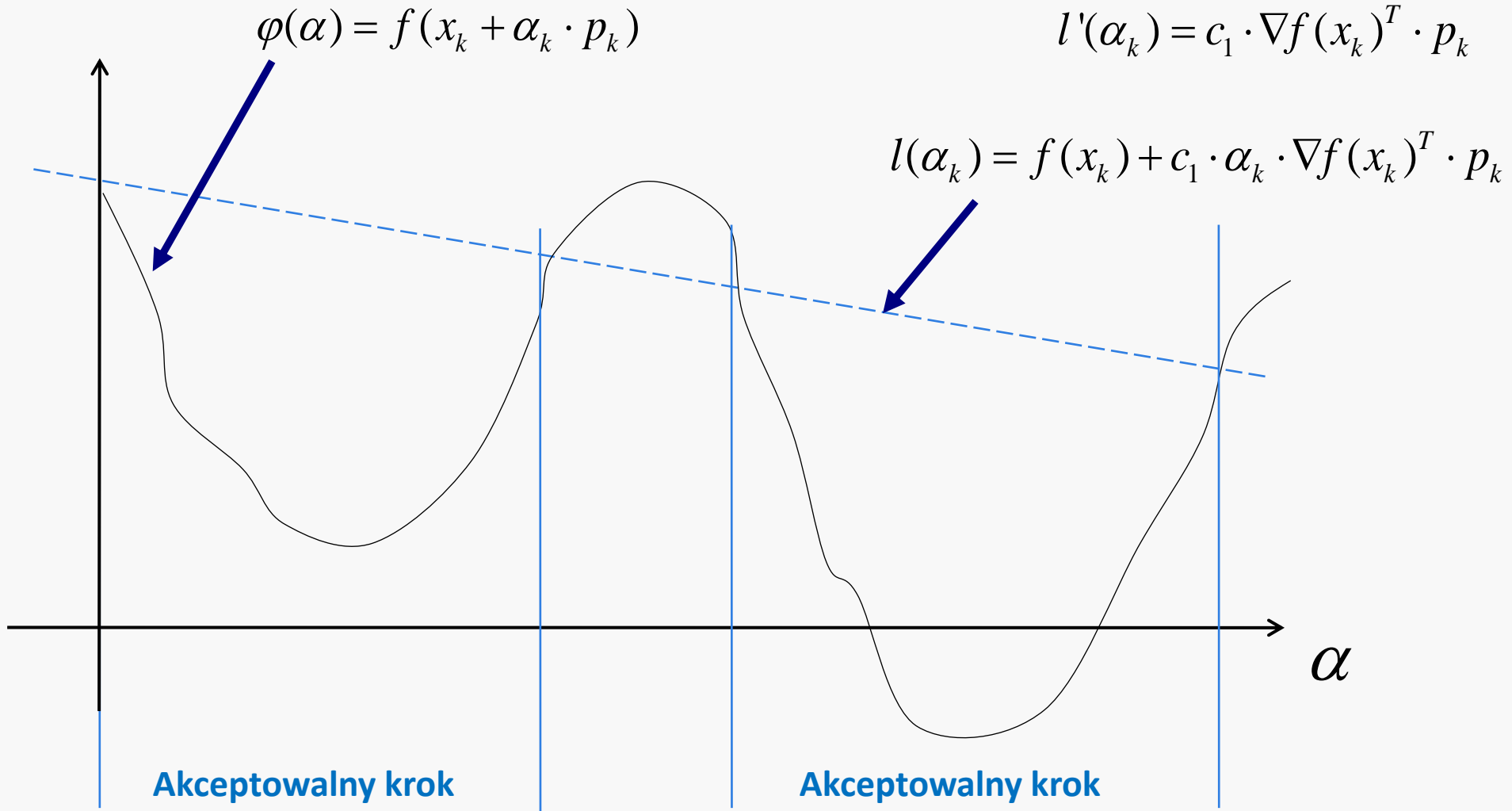
gdzie:

$x_k$  – rozwiązanie z k-tej iteracji,

$p_k$  – kierunek poszukiwań w k-tej iteracji,

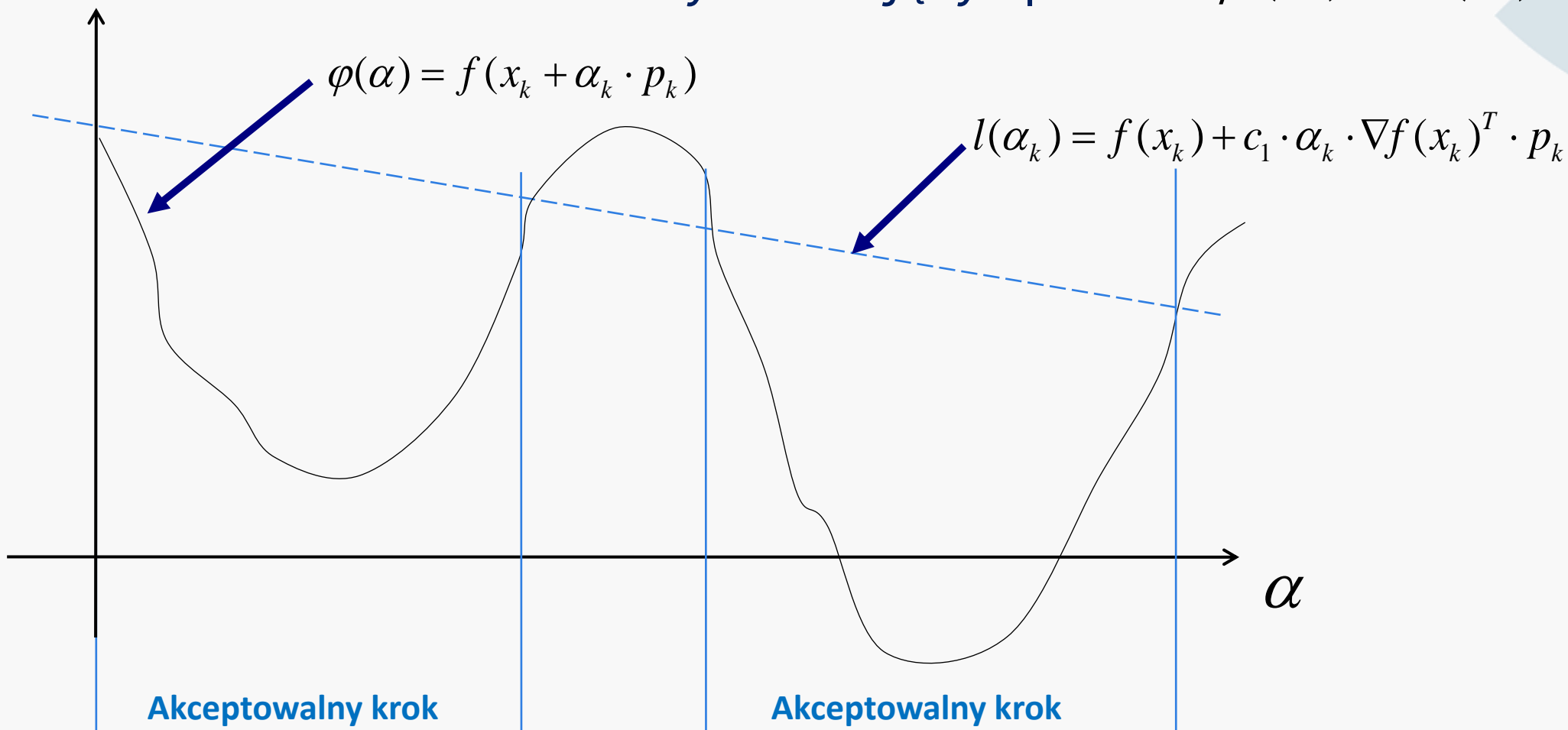
$c_1 \in (0,1)$  – stała, zazwyczaj bardzo mała (np.  $c_1 = 0.0001$ )







Warunek na wystarczający spadek  $\varphi(\alpha) \leq l(\alpha)$





# OMNIS Poszukiwanie optymalnego kroku – warunek krzywizny

Warunek Armijo (Warunek dostatecznego spadku):

$$f(x_k + \alpha_k \cdot p_k) \leq f(x_k) + c_1 \cdot \alpha_k \cdot \nabla f(x_k)^T \cdot p_k, \quad c_1 \in (0,1)$$

zastosowanie warunku Armijo może powodować zbyt zmniejszenie kroku. W celu wyeliminowania zbyt małego kroku stosuje się tzw. warunek krzywizny/nachylenia (*curvature condition*)

$$\nabla f(x_k + \alpha_k \cdot p_k)^T p_k \leq c_2 \cdot \nabla f(x_k)^T \cdot p_k, \quad c_2 \in (c_1,1)$$

Warunek Armijo + Warunek krzywizny = Warunek Wolfe'a





# OMNIS

## Poszukiwanie optymalnego kroku – warunek krzywizny

Wymagany spadek

$$\nabla f(x_k + \alpha_k \cdot p_k)^T p_k \leq c_2 \cdot \nabla f(x_k)^T \cdot p_k$$

$$\varphi'(\alpha_K) > c_2 \varphi'(0)$$

$f(x_k + \alpha \cdot p_k)$

Zbyt stromy spadek

Akceptowalny krok

Akceptowalny krok

$\alpha$

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.





# OMNIS

## Poszukiwanie optymalnego kroku – warunek Wolfe'a

Dwa warunki: Armijo i krzywizny  
składają się na tzw. warunek Wolfe'a (*Wolfe Condition*)

$$f(x_k + \alpha_k \cdot p_k) \leq f(x_k) + c_1 \cdot \alpha_k \cdot \nabla f(x_k)^T \cdot p_k, \quad c_1 \in (0,1)$$

$$\nabla f(x_k + \alpha_k \cdot p_k)^T p_k \leq c_2 \cdot \nabla f(x_k)^T \cdot p_k, \quad c_2 \in (c_1,1)$$

$$0 < c_1 < c_2 < 1$$

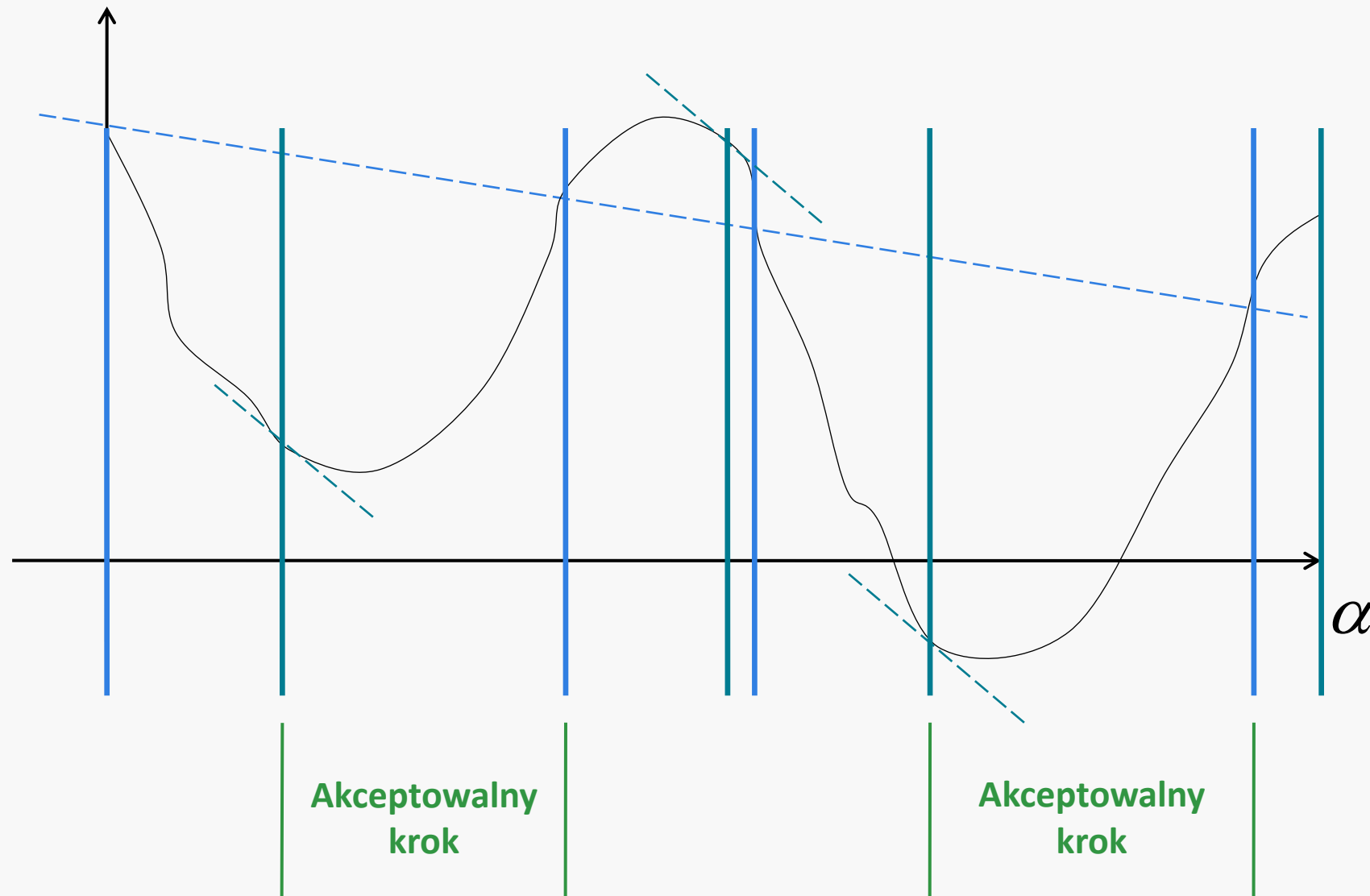




# OMNIS

## Poszukiwanie optymalnego kroku – warunek Wolfe'a

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.



Fundusze Europejskie dla Rozwoju Społecznego



Rzeczpospolita Polska

Dofinansowane przez Unię Europejską



Politechnika Warszawska



# OMNIS Poszukiwanie optymalnego kroku – warunek Wolfe’a

## Silne warunki Wolfe’a (Strong Wolfe Conditions)

Są one bardziej restrykcyjne i częściej stosowane w praktyce (np. w algorytmie L-BFGS), ponieważ ograniczają pochodną kierunkową z obu stron, zmuszając algorytm do lądowania blisko punktu stacjonarnego wzdłuż linii poszukiwań

$$f(x_k + \alpha_k \cdot p_k) \leq f(x_k) + c_1 \cdot \alpha_k \cdot \nabla f(x_k)^T \cdot p_k, \quad c_1 \in (0, 1)$$

$$\left| \nabla f(x_k + \alpha_k \cdot p_k)^T p_k \right| \leq c_2 \cdot \left| \nabla f(x_k)^T \cdot p_k \right|, \quad c_2 \in (c_1, 1)$$

$$0 < c_1 < c_2 < 1$$





- Warunek Armijo odpowiada za maksymalny możliwy krok (wzmocnienie)
- Warunek krzywizny zapobiega zbytniemu zmniejszeniu kroku, jest to szczególnie istotne w ostatniej fazie optymalizacji tuż przed rozwiązaniem optymalnym
- Warunki Armijo i krzywizny mogą być stosowane rozdzielnie
- Warunek Armijo jest słabszy niż warunek Wolfe'a
- W znaczący sposób przyspieszają proces, zmniejszając liczbę iteracji a tym samym koszt obliczeń





### Podsumowanie różnic

| Kryterium | Cel                    | Zapobiega           |
|-----------|------------------------|---------------------|
| Armijo    | Dostateczny spadek     | Zbyt długim krokom  |
| Wolfe     | Odpowiednie nachylenie | Zbyt krótkim krokom |

Warunki Armijo i Wolfe'a to matematyczne kryteria wyboru długości kroku  $\alpha$  w gradientowych algorytmach optymalizacji. Warunek Armijo wymaga, aby spadek wartości funkcji był co najmniej proporcjonalny do kroku i nachylenia gradientu, natomiast warunki Wolfe'a dodatkowo wymuszają, aby krok był na tyle duży, by nachylenie funkcji w nowym punkcie było wyraźnie łagodniejsze niż w punkcie startowym.



OMNIS Optymalizacja w inżynierii lotniczej i kosmicznej

Otwartość. Modernizacja. Nowoczesność. Integracja. Społeczność.

Dziękuję za uwagę

