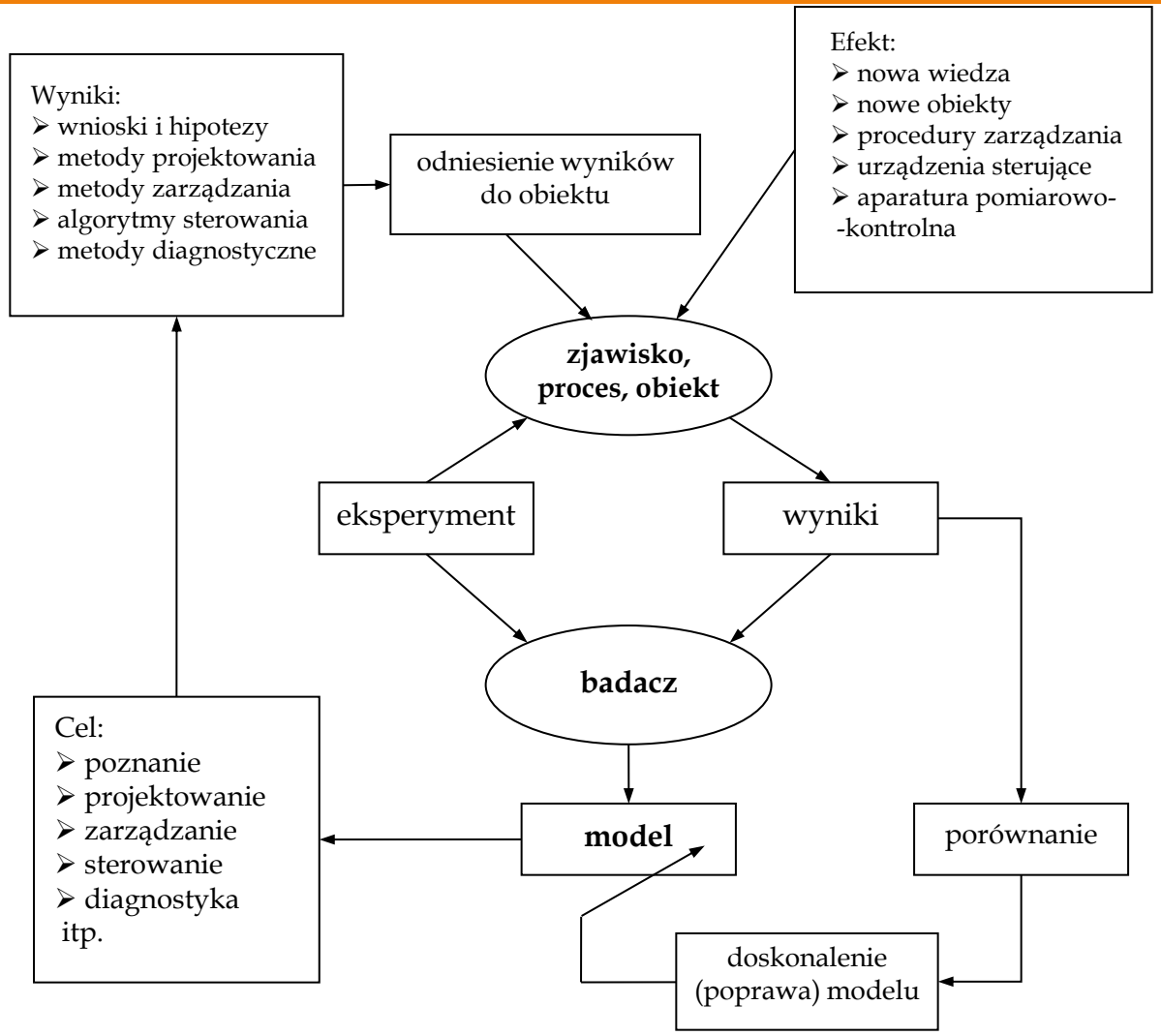
# Optymalizacja systemów



Wykład 16. Podsumowanie

# Model w badaniach systemowych



# Sformułowanie zadania wspomagania decyzji

Decyzja: różne typy elektrowni



Elektrownia wodna

$x^{(1)}$



Elektrownia atomowa

$x^{(2)}$



Elektrownia wiatrowa

$x^{(3)}$

Obrazy:
http://ziemianarozdrozu.pl/encyklopedia/67/hydroenergety
http://kresy24.pl/showNews/news_id/5871/
http://windy-future.info/2009/10/13/large-wind-turbine/

$x^{(1)}, x^{(2)}, x^{(3)}$ - obciążenie elektrowni – zmienne decyzyjne

$c_1, c_2, c_3$ – koszt jednostkowy wytworzenia energii

Cel: minimalizacja kosztów wytworzenia: $F(x^{(1)}, x^{(2)}, x^{(3)}) = c_1 x^{(1)} + c_2 x^{(2)} + c_3 x^{(3)}$

Ograniczenia: – spełnienie zapotrzebowania na energię: $x^{(1)} + x^{(2)} + x^{(3)} \geq \beta$

– możliwości poszczególnych elektrowni: $0 \leq x^{(n)} \leq \alpha_n, \; n = 1,2,3$

# Sformułowanie zadania wspomagania decyzji

**Zmienne decyzyjne:** $x = \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(S)} \end{bmatrix}$ **Funkcja celu:** $y = F(x)$

**Zbiór rozwiązań dopuszczalnych** (zwykle określony przez wskazanie ograniczeń):

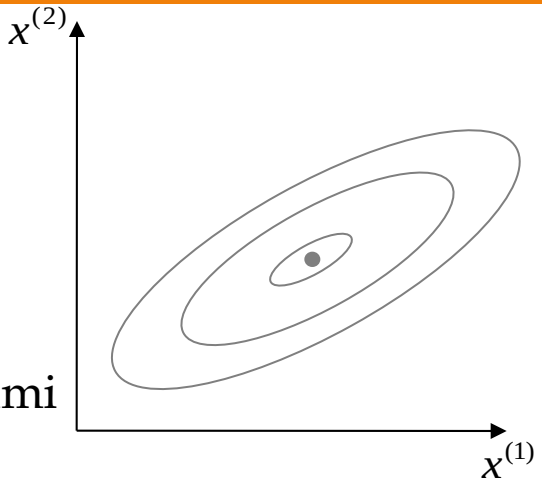
$$x \in \mathscr{D}_x$$

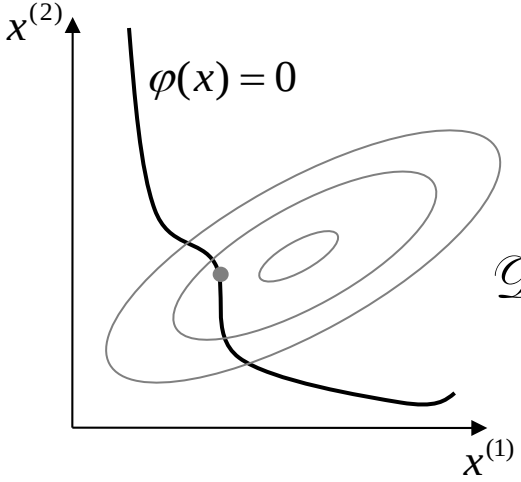
**Decyzja optymalna:** $x^* \rightarrow F(x^*) = \min_{x \in \mathscr{D}_x} F(x), \quad \min F(x) = -\max\left(-F(x)\right)$

**Decyzja zadowalająca:** $\tilde{x} \rightarrow F(\tilde{x}) \leq F(x^*) + \alpha = \gamma \wedge x \in \mathscr{D}_x$

# Typowe zadania decyzyjne

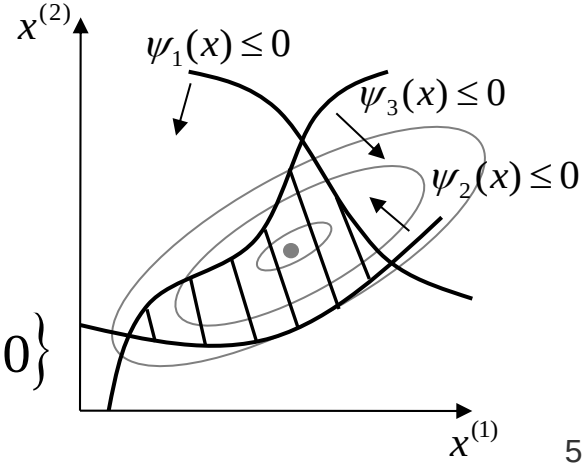
Zadanie decyzyjne bez ograniczeń: $\mathscr{D}_x = \mathscr{R}^S$




Zadanie decyzyjne z ograniczeniami równościowymi:

$$\mathscr{D}_x = \left\{ x \in \mathscr{R}^S : \varphi_1(x) = 0, \varphi_2(x) = 0, \ldots, \varphi_L(x) = 0, L \leq S \right\}$$




Zadanie decyzyjne z ograniczeniami nierównościowymi:

$$\mathscr{D}_x = \left\{ x \in \mathscr{R}^S : \psi_1(x) \leq 0, \psi_2(x) \leq 0, \ldots, \psi_M(x) \leq 0 \right\}$$

# Analityczne metody optymalizacji

- Zadanie optymalizacji bez ograniczeń
- Zadanie optymalizacji z ograniczeniami równościowymi – metoda współczynników Lagrange'a
- Zadanie optymalizacji z ograniczeniami nierównościowymi – metoda Kuhna-Tuckera

# Zadanie optymalizacji bez ograniczeń

Zadanie optymalizacji: $x^* \rightarrow F(x^*) = \min_{x \in \mathscr{D}_x = \mathscr{R}^S} F(x)$

Założenie: $F(x)$ jest funkcją ciągłą i różniczkowalną.

Warunkiem koniecznym aby $x^*$ było minimum lokalnym jest:

$$\nabla_x F(x)\big|_{x^*} = 0_S$$

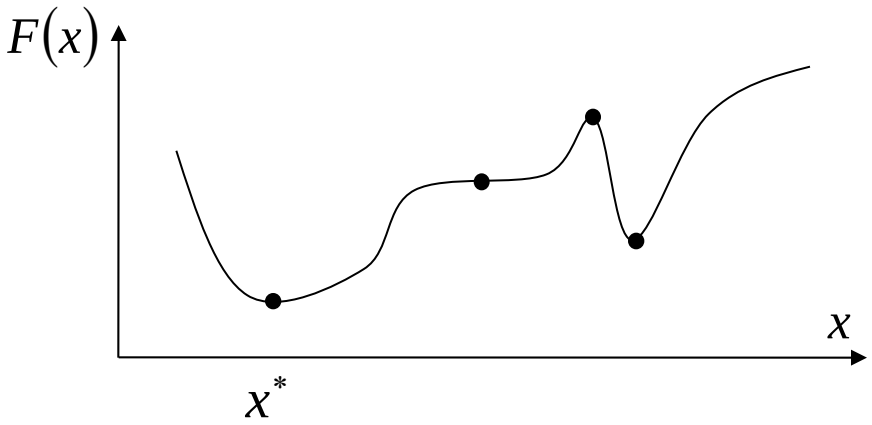
Jeżeli $F(x)$ jest funkcją pseudo - wypukłą, powyższe równanie jest warunkiem koniecznym i wystarczającym aby $x^*$ było minimum globalnym

Jeżeli $H(x)$ jest dodatnio pół określona $\forall\ x \in \mathscr{R}^S$ to rozwiązanie powyższego równania jest minimum globalnym

Jeżeli $H(x)$ jest dodatnio określona $\forall\ x \in \mathscr{R}^S$ powyższe równanie ma jedno rozwiązanie $x^*$ i jest ono minimum globalnym

# Zadanie optymalizacji bez ograniczeń

Równanie $\nabla_x F(x)\big|_{x^*} = 0_S$ może mieć wiele rozwiązań




Warunki optymalności drugiego rzędu:

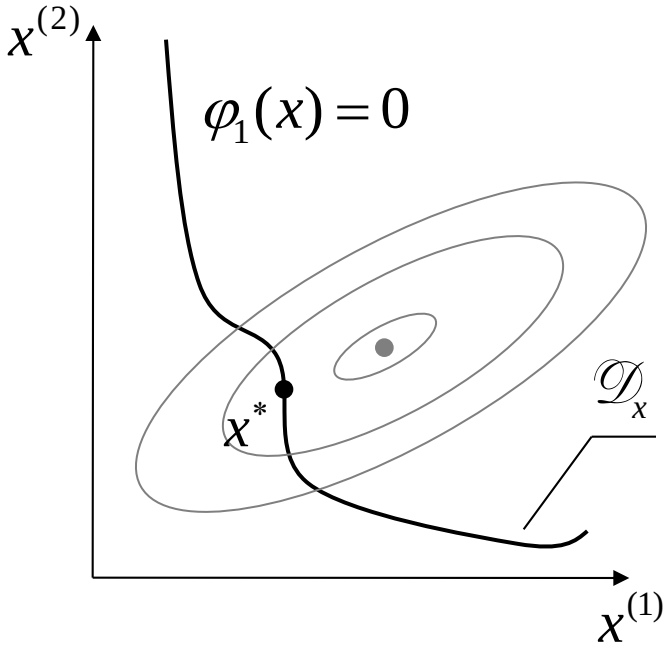
Jeżeli $H(x^*)$ jest dodatnio pół określona w punkcie $x^*$
to $x^*$ jest minimum lokalnym

Jeżeli $H(x^*)$ jest ujemnie pół określona w punkcie $x^*$
to $x^*$ jest maksimum lokalnym

# Zadanie optymalizacji z ograniczeniami równościowymi
## Metoda mnożników Lagrange' a

Zadanie optymalizacji: $x^* \rightarrow F(x^*) = \min\limits_{x^* \in \mathscr{D}_x} F(x)$

$$\mathscr{D}_x = \left\{ x \in \mathscr{R}^S : \varphi_1(x) = 0, \varphi_2(x) = 0, \ldots, \varphi_L(x) = 0, \; L \leq S \right\}$$




$$\mathscr{D}_x = \left\{ x \in \mathscr{R}^S : \varphi(x) = 0_L, \; L \leq S \right\}$$

$$\varphi(x) = \begin{bmatrix} \varphi_1(x) \\ \varphi_2(x) \\ \vdots \\ \varphi_L(x) \end{bmatrix} \quad 0_L = \left.\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right\} L - \text{ zer}$$

# Zadanie optymalizacji z ograniczeniami równościowymi
## Metoda mnożników Lagrange' a

- Metoda współczynników Lagrange' a

Funkcja Lagrange'a:

$$L(x,\lambda) = F(x) + \sum_{l=1}^{L} \lambda_l \varphi_l(x) = F(x) + \lambda^T \varphi(x)$$

gdzie: $\varphi(x) = \begin{bmatrix} \varphi_1(x) \\ \varphi_2(x) \\ \vdots \\ \varphi_L(x) \end{bmatrix}$, $\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_L \end{bmatrix}$ - wektor współczynników Lgrange' a

Warunki konieczne optymalności:

$$\nabla_x L(x,\lambda)\big|_{x^*,\lambda^*} = 0_S$$

$$\nabla_\lambda L(x,\lambda)\big|_{x^*,\lambda^*} = 0_L \quad \Leftrightarrow \quad rank\, G(x) = rank \left[ G(x) \;\vdots\; -\nabla_x F(x) \right],$$

$$G(x) = \left[ \nabla_x \varphi_1(x) \;\vdots\; \nabla_x \varphi_2(x) \;\vdots\; \cdots \;\vdots\; \nabla_x \varphi_L(x) \right]$$

# Zadanie optymalizacji z ograniczeniami równościowymi
## Metoda mnożników Lagrange' a

Powyższy układ równań może mieć wiele rozwiązań

Warunki optymalności drugiego rzędu:

Oznaczmy: $H_L(x) = \nabla_{xx} L(x, \lambda)$

Jeżeli $H_L(x^*)$ jest dodatnio określona w punkcie $x^*$
to $x^*$ jest minimum lokalnym

Jeżeli $H_L(x^*)$ jest ujemnie określona w punkcie
to $x^*$ jest maksimum lokalnym $x^*$

Jeżeli funkcja $F(x)$ jest wypukła, a ograniczenia są liniowe czyli mają postać $\varphi_l(x) = p_l^T x - \alpha_l = 0, \quad l = 1,2,\ldots,L$ to powyższy układ równań ma jedno rozwiązanie i jest ono rozwiązaniem optymalnym

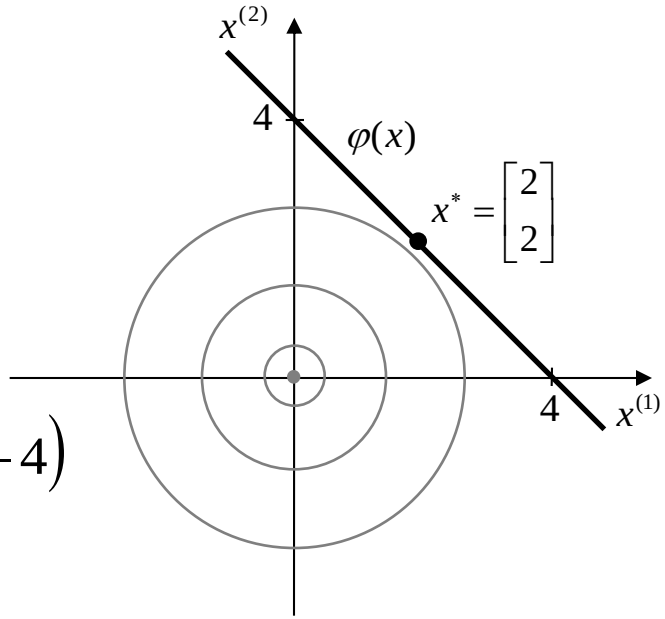
# Zadanie optymalizacji z ograniczeniami równościowymi
## Metoda mnożników Lagrange' a

- Metoda Lagrange' a - Przykład 1

$$F(x) = \left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2$$

$$\varphi(x) = x^{(1)} + x^{(2)} - 4 = 0$$

$$L(x,\lambda) = \left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2 + \lambda\left(x^{(1)} + x^{(2)} - 4\right)$$

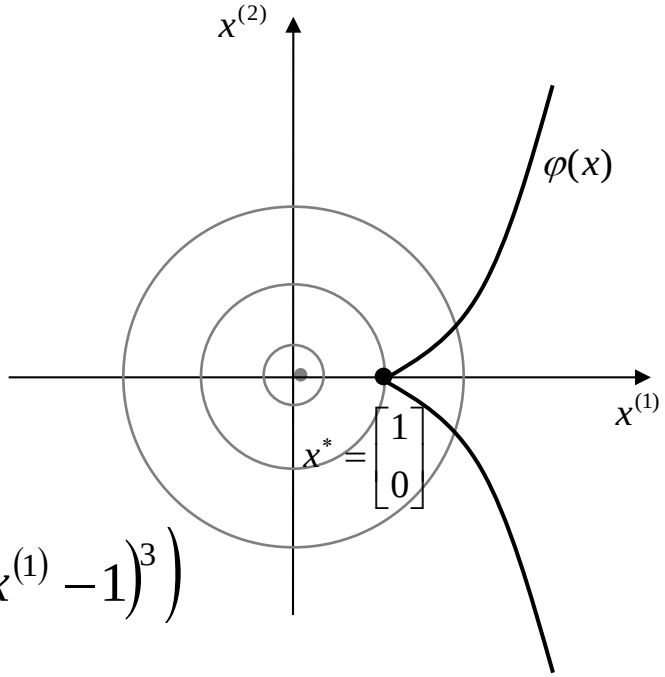


# Zadanie optymalizacji z ograniczeniami równościowymi
## Metoda mnożników Lagrange' a

- Metoda współczynników Lagrange'a – Przykład 2 rozwiązanie nieregularne

$$F(x) = \left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2$$

$$\varphi(x) = \left(x^{(2)}\right)^2 - \left(x^{(1)} - 1\right)^3 = 0$$




$$L(x,\lambda) = \left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2 + \lambda\left(\left(x^{(2)}\right)^2 - \left(x^{(1)} - 1\right)^3\right)$$

# Zadanie optymalizacji z ograniczeniami równościowymi
## Metoda mnożników Lagrange' a

Jeżeli F(x) jest funkcją ciągłą, różniczkowalną i wypukłą oraz ograniczenia $\varphi_1(x), \varphi_2(x), \ldots, \varphi_L(x)$ są liniowe to układ równań:

$$\nabla_x L(x,\lambda)\big|_{x^*,\lambda^*} = 0_S$$

$$\nabla_\lambda L(x,\lambda)\big|_{x^*,\lambda^*} = 0_L$$

ma jedno rozwiązania i jest ono rozwiązaniem zadania optymalizacji z ograniczeniami równościowymi.

Powyższy układ równań w tym przypadku jest warunkiem koniecznym i wystarczającym

# Zadanie optymalizacji z ograniczeniami równościowymi Metoda mnożników Lagrange' a

- Uogólniona metoda współczynników Lagrange' a

$$\nabla_x L(x,\lambda,\lambda_0) = \lambda_0 \nabla_x F(x) + \sum_{l=1}^{L} \lambda_l \nabla_x \varphi_l(x) = 0_S$$

$1^O$ $\lambda_0 \neq 0$ $\quad \nabla_x F(x) + \sum_{l=1}^{L} \frac{\lambda_l}{\lambda_0} \nabla_x \varphi_l(x) = 0_S \Rightarrow \nabla_x F(x) + \sum_{l=1}^{L} \lambda_l' \nabla_x \varphi_l(x) = 0_S$

$\lambda_0 = 1$ $\quad \nabla_x F(x) + \sum_{l=1}^{L} \lambda_l \nabla_x \varphi_l(x) = 0_S$ - z tego warunku otrzymamy rozwiązania regularne

$2^O$ $\lambda_0 = 0$ $\quad \sum_{l=1}^{L} \lambda_l \nabla_x \varphi_l(x) = 0_S$ - z tego warunku otrzymamy rozwiązania nieregularne

Podobnie jak poprzednio otrzymane rozwiązania wymagają zbadania warunków drugiego rzędu czyli zbadania określoności macierzy:

$$H_L(x,\lambda,\lambda_0) = \nabla_{xx}^2 L(x,\lambda,\lambda_0).$$

# Zadanie optymalizacji z ograniczeniami nierównościowymi
## Metoda Kuhna – Tucker'a

Zadanie optymalizacji:

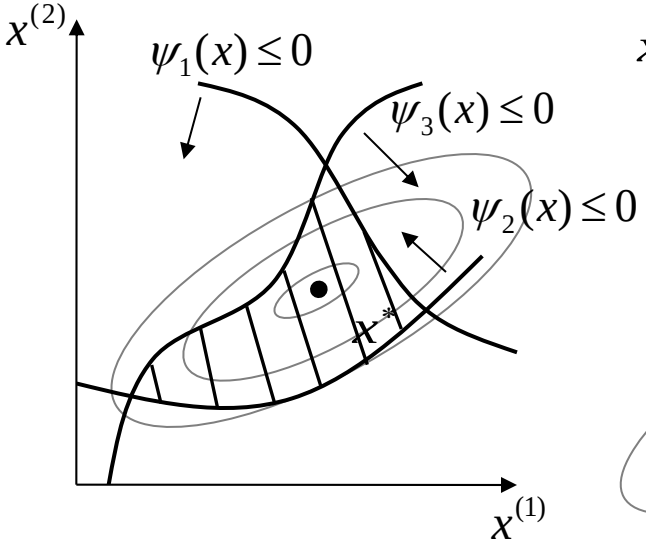
$$x^* \rightarrow F(x^*) = \min_{x \in \mathscr{D}_x} F(x)$$

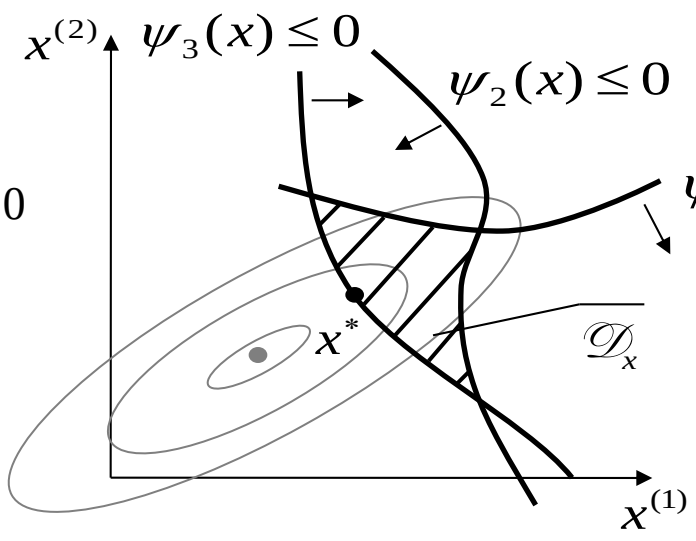
$$\mathscr{D}_x = \left\{ x \in \mathscr{R}^S : \psi_1(x) \le 0, \psi_2(x) \le 0, \ldots, \psi_M(x) \le 0 \right\}$$

$$\mathscr{D}_x = \left\{ x \in \mathscr{R}^S : \psi(x) \le 0_M \right\}$$

gdzie:

$$\psi(x) = \begin{bmatrix} \psi_1(x) \\ \psi_2(x) \\ \vdots \\ \psi_M(x) \end{bmatrix} \qquad 0_M = \left.\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right\} M$$




$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_S \end{bmatrix} \qquad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_S \end{bmatrix}$$

$$\alpha \le \beta \Rightarrow \forall_{s=1,\ldots,S} \alpha_s \le \beta_s$$

# Zadanie optymalizacji z ograniczeniami nierównościowymi
# Metoda Kuhna – Tucker'a

Funkcja Lagrange' a:

$$L(x,\mu) = F(x) + \mu^T \psi(x) \;\Leftrightarrow\; L(x,\mu) = F(x) + \sum_{m=1}^{M} \mu_m \psi_m(x)$$

gdzie: $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_M \end{bmatrix}$ - wektor współczynników Lgrange' a

Warunki konieczne optymalności:

$$\nabla_x L(x,\mu)\Big|_{x^*,\mu^*} = 0_S$$

$$\mu^T \nabla_\mu L(x,\mu)\Big|_{x^*,\mu^*} = 0$$

$$\nabla_\mu L(x,\mu)\Big|_{x^*,\mu^*} \le 0_M$$

$$\mu^* \ge 0_M \qquad \Leftrightarrow \quad \text{gdy rozwiązanie jest regularne}$$

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_S \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_S \end{bmatrix} \qquad \alpha \le \beta \Rightarrow \forall_{s=1,\ldots,S} \alpha_s \le \beta_s$$

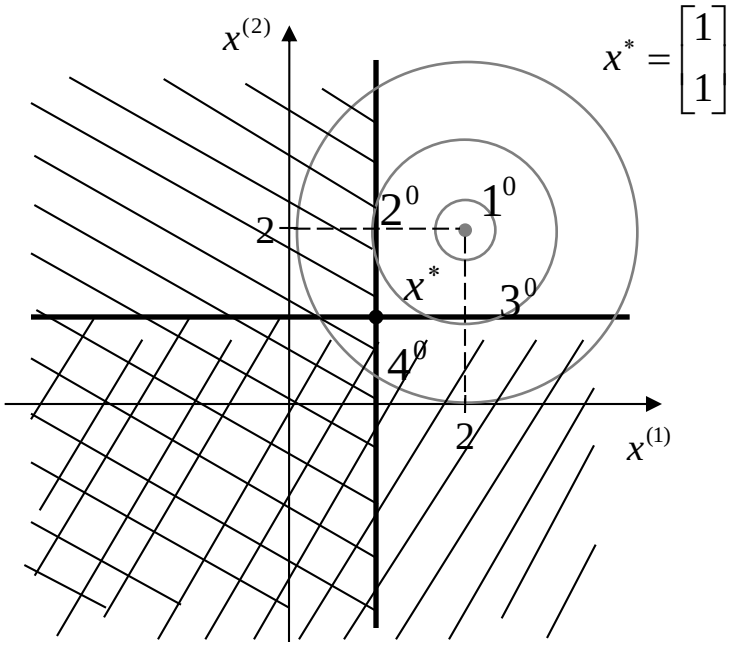
# Zadanie optymalizacji z ograniczeniami nierównościowymi
# Metoda Kuhna – Tucker'a

Przykład 1

$$F(x) = \left(x^{(1)} - 2\right)^2 + \left(x^{(2)} - 2\right)^2$$

$$\psi_1(x) = x^{(1)} - 1 \le 0$$

$$\psi_2(x) = x^{(2)} - 1 \le 0$$




$$L(x, \lambda) = \left(x^{(1)} - 2\right)^2 + \left(x^{(2)} - 2\right)^2 + \mu_1\left(x^{(1)} - 1\right) + \mu_2\left(x^{(2)} - 1\right)$$

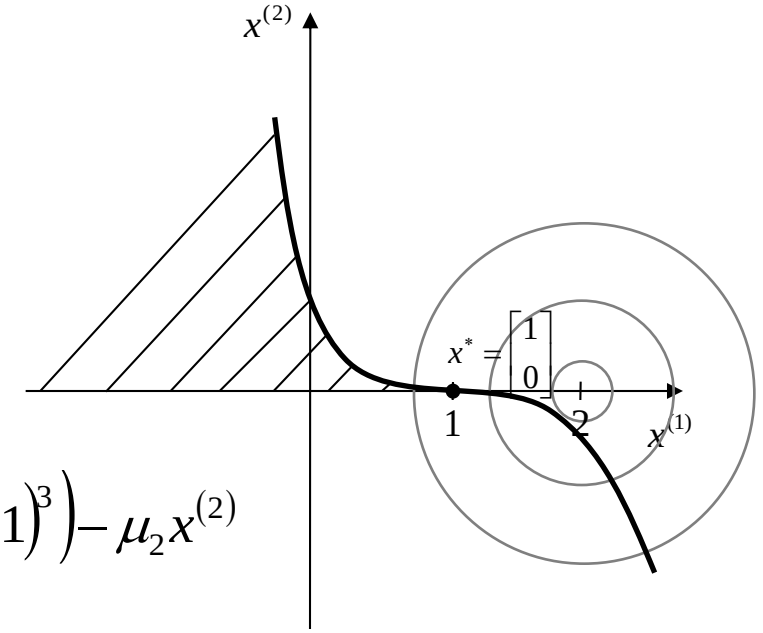
# Zadanie optymalizacji z ograniczeniami nierównościowymi Metoda Kuhna – Tucker'a

Przykład 2 – rozwiązanie nieregularne

$$F(x) = \left(x^{(1)} - 2\right)^2 + \left(x^{(2)}\right)^2$$

$$\psi_1(x) = x^{(2)} + \left(x^{(1)} - 1\right)^3 \leq 0$$

$$\psi_2(x) = -x^{(2)} \leq 0$$




$$L(x,\lambda) = \left(x^{(1)} - 2\right)^2 + \left(x^{(2)}\right)^2 + \mu_1\left(x^{(2)} + \left(x^{(1)} - 1\right)^3\right) - \mu_2 x^{(2)}$$

# Zadanie optymalizacji z ograniczeniami nierównościowymi
# Metoda Kuhna – Tucker'a Warunki regularności rozwiązania

1. Karlina: ograniczenia $\psi_1(x), \psi_2(x), \ldots, \psi_M(x)$ - liniowe
2. Slatera: ograniczenia $\psi_1(x), \psi_2(x), \ldots, \psi_M(x)$ - wypukłe oraz zbiór rozwiązań dopuszczalnych ma niepuste wnętrze
3. Fiacco – Mac Cormica: w punkcie optymalnym gradienty wszystkich ograniczeń aktywnych są liniowo niezależne, czyli:

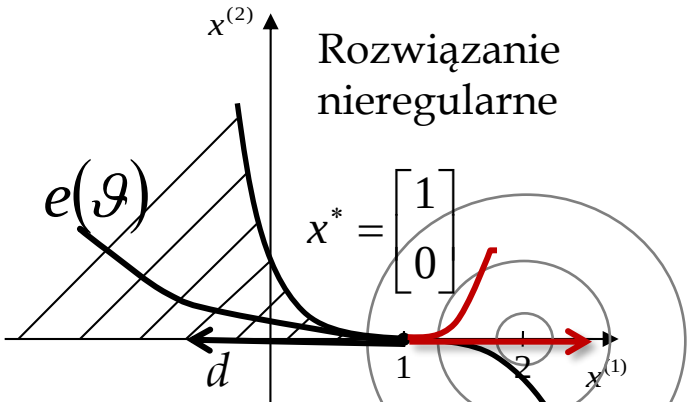
$$\forall\, m \in I(x^*) \quad \nabla_x \psi_m(x^*)\Big|_{x=x^*}$$ są liniowo niezależne

4. Zangwila: $\mathscr{D}(x^*) = \overline{D}(x^*)$
5. Kuhna – Tucker'a: dla każdego kierunku $d \in \mathscr{D}(x^*)$ istnieje krzywa regularna rozpoczynająca się w punkcie $x^*$ i styczna do tego kierunku

$$\forall\, d \in \mathscr{D}(x^*) \quad \exists\, e(\vartheta), \quad \vartheta \in [0,1]$$

- $e(0) = x^*$
- $e(\vartheta) \in D_x \quad \forall\, \vartheta \in [0,1]$
- $\dfrac{de(\vartheta)}{d\vartheta}\Big|_{\vartheta=0} = \tau \cdot d$

$$e(\vartheta) = \begin{bmatrix} e_1(\vartheta) \\ e_2(\vartheta) \\ \vdots \\ e_S(\vartheta) \end{bmatrix}$$



# Zadanie optymalizacji z ograniczeniami nierównościowymi Metoda Kuhna – Tucker'a

Warunki konieczna i wystarczające:

Jeżeli funkcje $F(x), \psi_1(x), \psi_2(x), \ldots, \psi_M(x)$ są ciągłe i różniczkowalne oraz funkcja $F(x)$ jest funkcją pseudo – wypukłą, a ograniczenia $\psi_1(x), \psi_2(x), \ldots, \psi_M(x)$ są funkcjami quasi – wypukłymi to wkład równań i nierówności:

$$\nabla_x L(x,\mu)\Big|_{x^*,\mu^*} = 0_S$$

$$\mu^T \nabla_\mu L(x,\mu)\Big|_{x^*,\mu^*} = 0$$

$$\nabla_\mu L(x,\mu)\Big|_{x^*,\mu^*} \le 0_M$$

$$\mu^* \ge 0_M$$

ma jedno rozwiązanie i jest ono rozwiązaniem zadania optymalizacji z ograniczeniami nierównościowymi

# Szczególny przypadek

$$x^* \rightarrow F(x^*) = \min_{x \in \mathscr{D}_x} F(x)$$

$$\mathscr{D}_x = \left\{ x \in \mathscr{R}^S : x \geq 0_S, \psi(x) \leq 0_M \right\}$$

$$L(x, \mu) = F(x) + \mu^T \psi(x)$$

$$\nabla_x L(x, \mu)\Big|_{x^*, \mu^*} \geq 0_S \qquad \nabla_\mu L(x, \mu)\Big|_{x^*, \mu^*} \leq 0_M$$

$$x^T \nabla_x L(x, \mu)\Big|_{x^*, \mu^*} = 0 \qquad \mu^T \nabla_\mu L(x, \mu)\Big|_{x^*, \mu^*} = 0$$

$$x^* \geq 0_S \qquad \mu^* \geq 0_M$$

# Szczególny przypadek

$$x^* \rightarrow F(x^*) = \min_{x \in \mathscr{D}_x} F(x)$$

$$\mathscr{D}_x = \left\{ x \in \mathscr{R}^S : \varphi(x) = 0_L, \psi(x) \leq 0_M \right\}$$

$$L(x, \lambda, \mu) = F(x) + \lambda^T \varphi(x) + \mu^T \psi(x)$$

$$\nabla_x L(x, \lambda, \mu)\Big|_{x^*, \lambda^*, \mu^*} = 0_S$$

$$\nabla_\lambda L(x, \lambda, \mu)\Big|_{x^*, \lambda^*, \mu^*} = 0_L$$

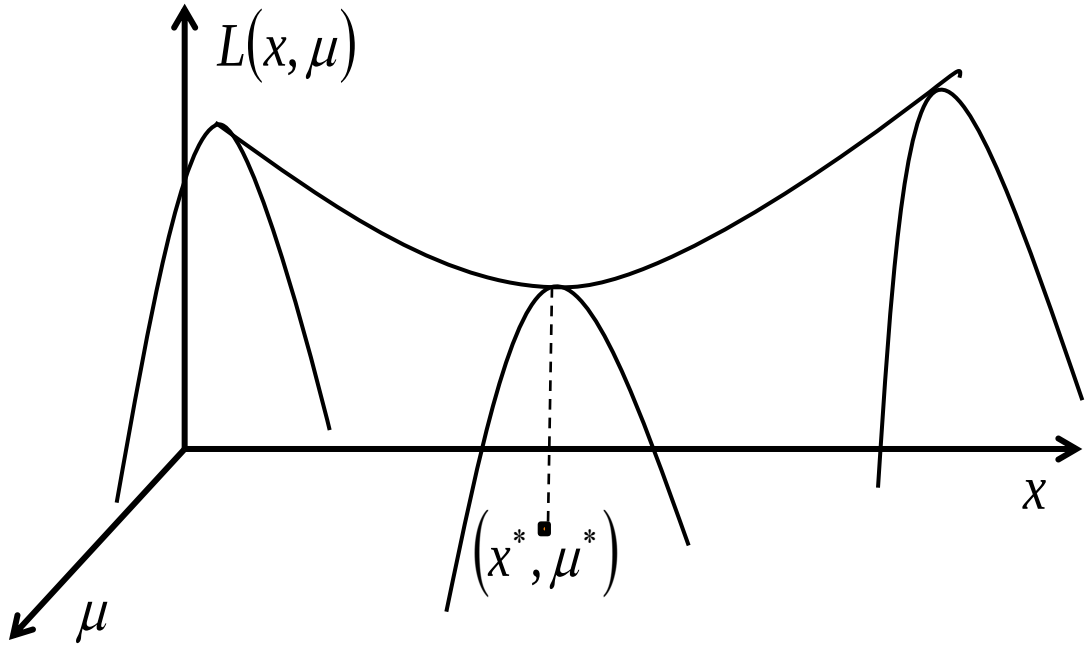
$$\mu^T \nabla_\mu L(x, \lambda, \mu)\Big|_{x^*, \lambda^*, \mu^*} = 0$$

$$\nabla_\mu L(x, \lambda, \mu)\Big|_{x^*, \lambda^*, \mu^*} \leq 0_M$$

$$\mu^* \geq 0_S$$

# Zadanie optymalizacji z ograniczeniami nierównościowymi - Punkt siodłowy

Punkt siodłowy $(x^*, \mu^*)$




$$L(x^*,\mu^*) \le L(x,\mu^*) \quad \forall x \in \mathscr{D}(x) \subseteq \mathscr{R}^S$$

$$L(x^*,\mu) \le L(x^*,\mu^*) \quad \forall \mu \ge 0_M$$

$$L(x^*,\mu^*) = \min_{x \in \mathscr{D}(x)} \max_{\mu \ge 0_M} L(x,\mu)$$

# Zadanie optymalizacji z ograniczeniami nierównościowymi - Punkt siodłowy

Punkt $(x^*, \mu^*)$ jest punktem siodłowym $(x \in \mathscr{D}(x),\ \mu \geq 0_M) \Leftrightarrow$

1. $x^* - \min \text{imalizuje } L(x, \mu)$
2. $\psi_m(x^*) \leq 0 \quad m = 1, 2, \ldots, M$
3. $\mu^* \psi_m(x^*) = 0 \quad m = 1, 2, \ldots, M$

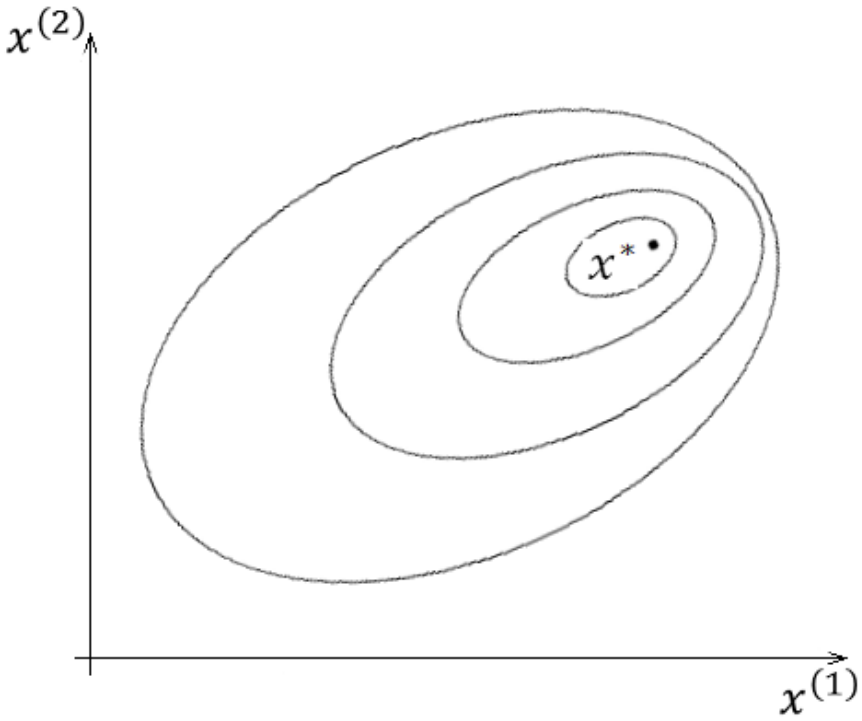
Jeżeli $(x^*, \mu^*)$ jest punktem siodłowym funkcji Lagrange'a $L(x,\mu)$ to jest rozwiązaniem zadania optymalizacji:

$$x^* \rightarrow F(x^*) = \min_{x \in \mathscr{D}_x} F(x)$$

$$\mathscr{D}_x = \left\{ x \in \mathscr{R}^S : \psi_1(x) \leq 0, \psi_2(x) \leq 0, \ldots, \psi_M(x) \leq 0 \right\}$$

# Numeryczne metody optymalizacji

$$x^* \rightarrow F(x^*) = \min_{x \in D_x} F(x)$$




Problemy analityczne:

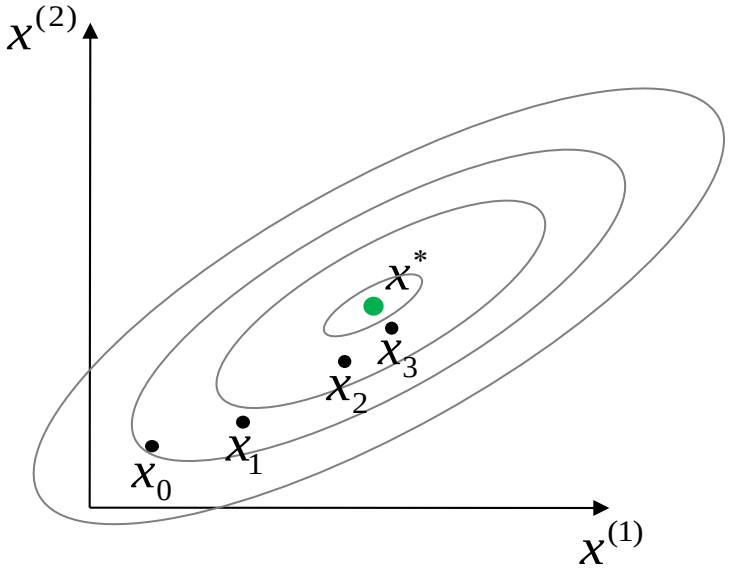
1. Nieliniowa złożona funkcja celu $F$ i ograniczeń $\varphi$ oraz $\psi$.
2. Nieróżniczkowalność funkcji $F$, $\varphi$ oraz $\psi$.
3. Nie jest znana postać analityczna funkcji $F$, $\varphi$ oraz $\psi$, można jedynie „zmierzyć" wartość funkcji
4. Duży wymiar wektora zmiennych decyzyjnych.

# Metody numeryczne

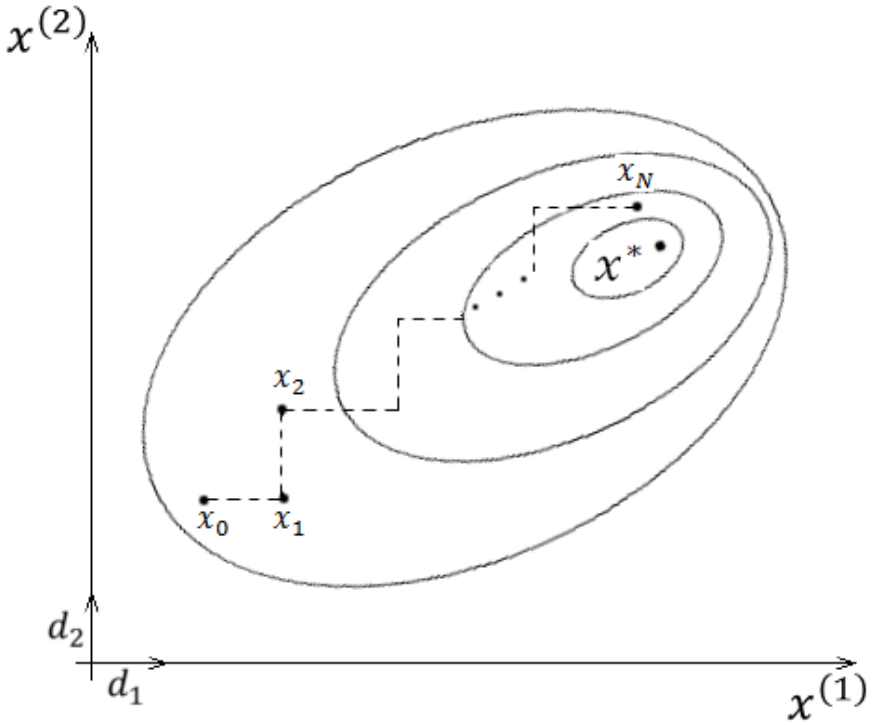
$$x^* \to F(x^*) = \min_{x^* \in \mathcal{D}_x} F(x),$$

Konstruujemy ciąg przybliżeń na podstawie wartości funkcji $F(x)$ w danym punkcie $x$

$$\begin{array}{ccccccccc} x_0 & \rightarrow & x_1 & \rightarrow & x_2 & \rightarrow & \cdots & \rightarrow & x_N \approx x^* \\ \downarrow & & \downarrow & & \downarrow & & & & \downarrow \\ F(x_0) & > & F(x_1) & > & F(x_2) & > & \cdots & > & F(x_N) \approx F(x^*) \end{array}$$



# Numeryczne metody optymalizacji

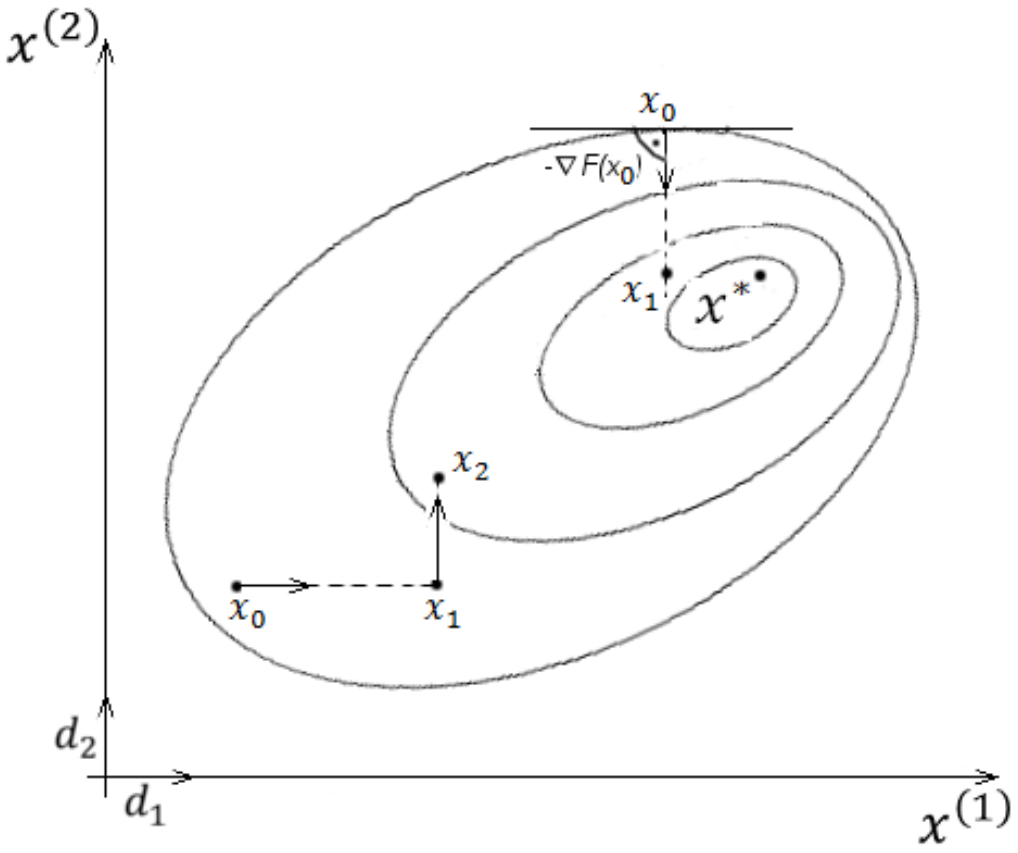



Algorytm
$x_{n+1} = \Psi(x_n), x_0$

- Wybór kierunku poszukiwań.
- Optymalizacja w kierunku.
- Warunki zatrzymania procedury.

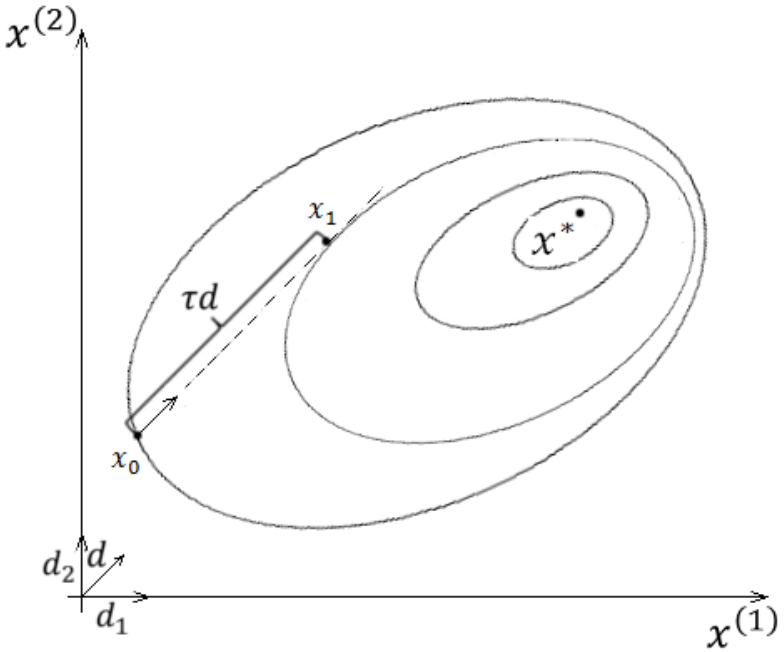
$$x_0, x_1, \dots, x_n, \dots, x_N \approx x^*$$
$$F(x_0) > F(x_1) > \dots > F(x_n) > \dots > F(x_N) \approx F(x^*)$$

# Wybór kierunku poszukiwań




- Kierunki bazowe i ich modyfikacje – metody bez gradientowe.
- Kierunki oparte na gradiencie funkcji – metody gradientowe.
- Inne

# Optymalizacja w kierunku




$x_0$ - punkt początkowy
$x_1$ - punkt końcowy
$d$ - kierunek
$\tau$ - długość kroku w kierunku

$\tau^* \rightarrow F(x_0 + \tau^* d) = \min_{\tau} F(x_0 + \tau d)$

$x_0, d$ – ustalone $F(x_0 + \tau d) \triangleq f(\tau)$

$f(\tau)$ – funkcja jednej zmiennej (długości kroku $\tau$)

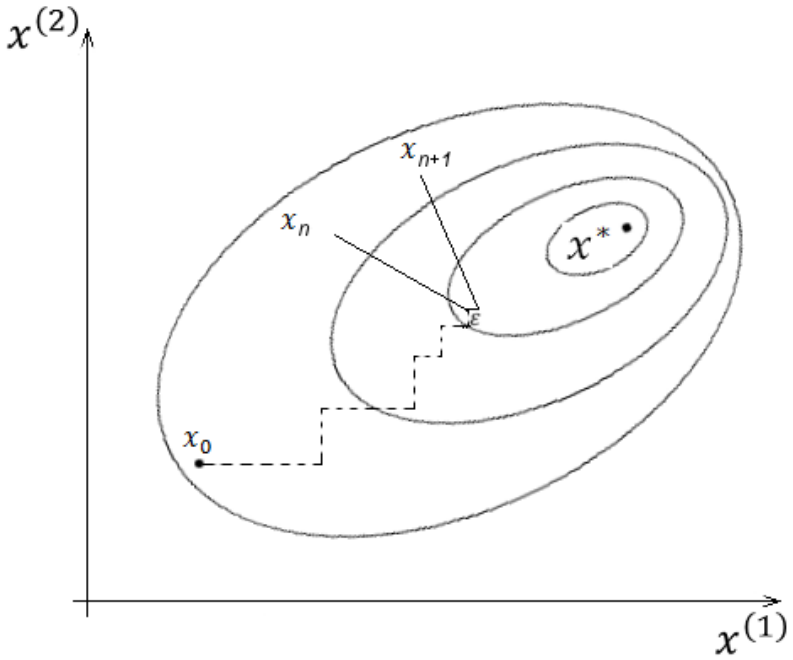
$\tau^* \rightarrow f(\tau^*) = \min_{\tau} f(\tau)$

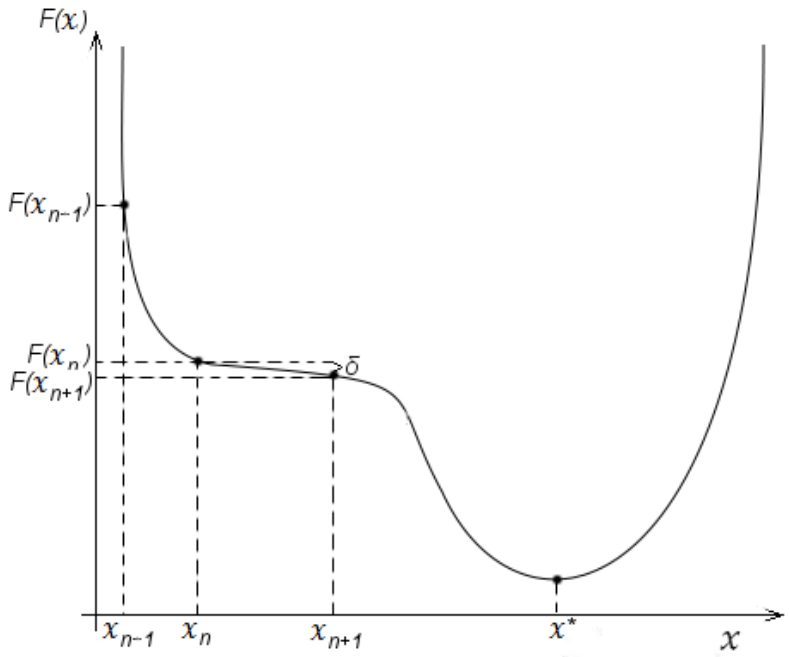
optymalizacja w kierunku ≡ optymalizacja funkcji jednej zmiennej

# Warunki zatrzymania procedury

$$\|x_{n+1} - x_n\| < \varepsilon; \quad |F(x_{n+1}) - F(x_n)| < \delta; \quad \frac{|F(x_{n+1}) - F(x_n)|}{\|x_{n+1} - x_n\|} < \varrho$$

Uwaga!

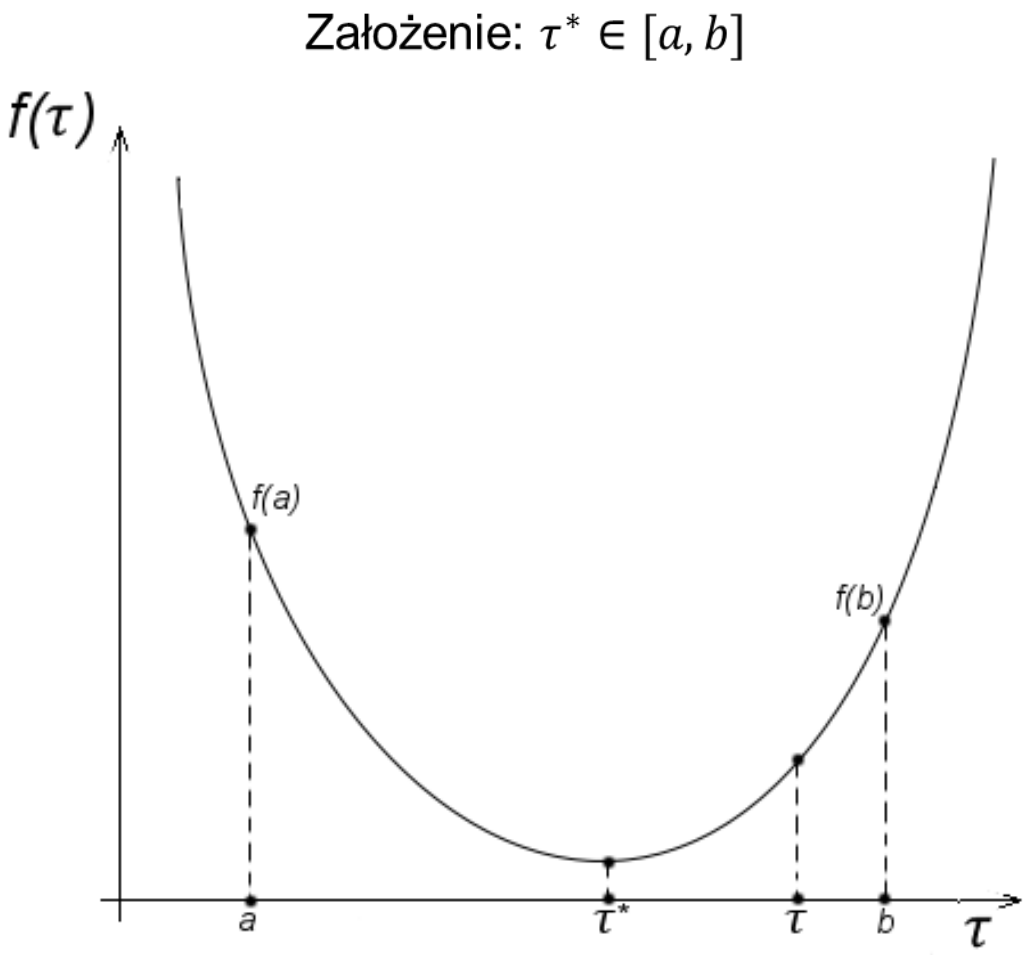


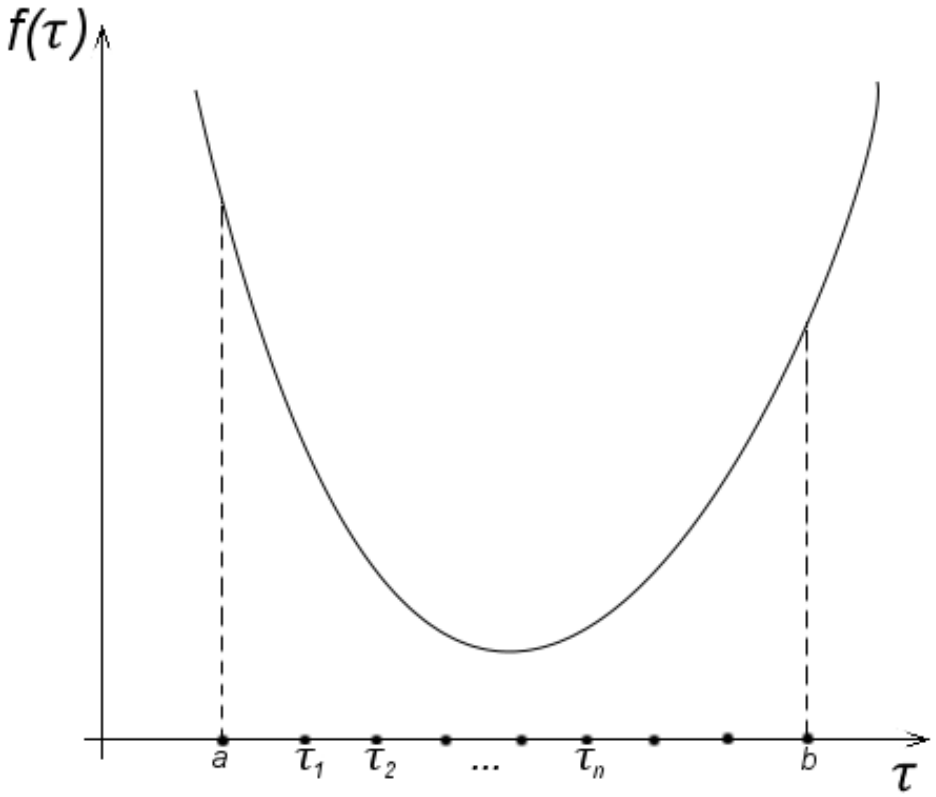
# Metody optymalizacji w kierunku

- Metody optymalizacji w kierunku
  - Podział równomierny
  - Podział na połowę
  - Złoty podział
  - Aproksymacji kwadratowej
  - Metoda pierwszej pochodnej
  - Metoda znaku pochodnej
  - Metoda Newtona
  - Metoda Bolzano

# Metody zawężania odcinka



# Podział równomierny




$\tau_0 = a$

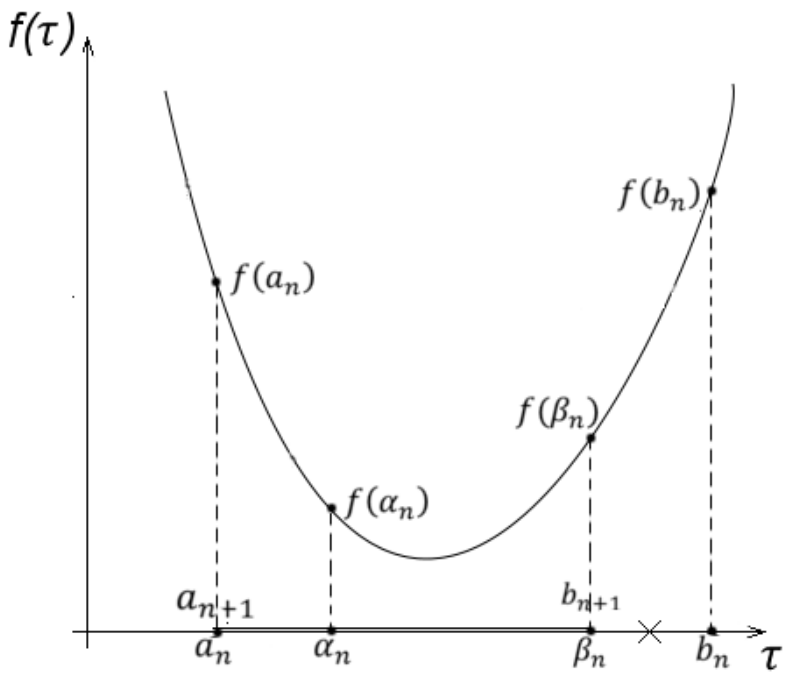
$\tau_n = \tau_0 + n\varepsilon$

$\tau^* \approx \tilde{\tau} \rightarrow f(\tilde{\tau}) = \min_{1 \le n \le N} \{f(\tau_n)\}$

$N = \left[\frac{b-a}{\varepsilon}\right]$ - liczba wyliczeń wartości funkcji

# Zawężanie odcinka

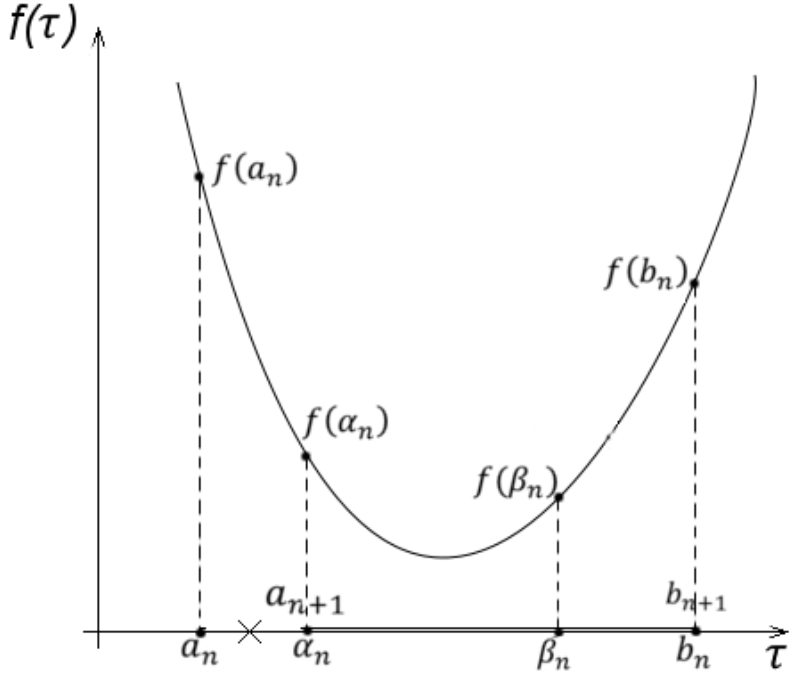
$f(\alpha_n) ? f(\beta_n)$




$f(\alpha_n) \leq f(\beta_n)$

$a_{n+1} := a_n$

$b_{n+1} := \beta_n$

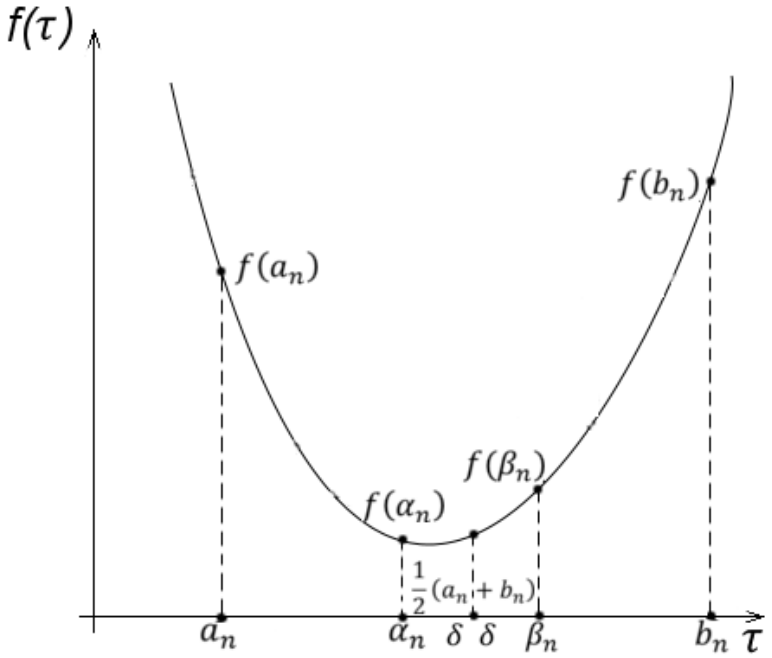



$f(\alpha_n) > f(\beta_n)$

$a_{n+1} := \alpha_n$

$b_{n+1} := b_n$

# Metoda podziału na połowę




$$\alpha_n = \frac{1}{2}(a_n + b_n) - \delta$$

$$\beta_n = \frac{1}{2}(a_n + b_n) + \delta$$

$N = ?$

Dane: $a_0, b_0, \varepsilon, \delta$

Krok 0: $n = 0$

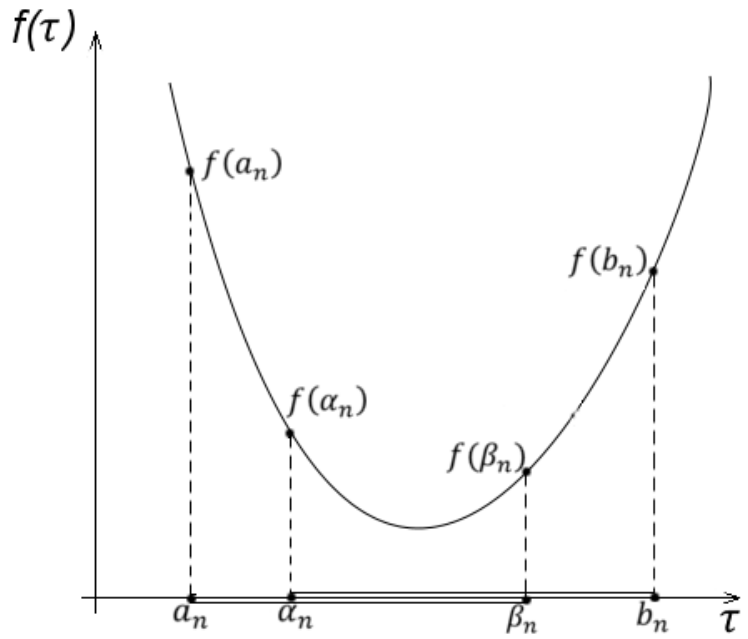
Krok 1: $\alpha_n = \frac{1}{2}(a_n + b_n) - \delta$

$\beta_n = \frac{1}{2}(a_n + b_n) + \delta$

Krok 2: Jeśli $f(\alpha_n) \leq f(\beta_n)$ to
$a_{n+1} := a_n, b_{n+1} := \beta_n,$
w przeciwnym razie
$a_{n+1} := \alpha_n, b_{n+1} := b_n.$

Krok 3: Jeśli $|b_{n+1} - a_{n+1}| \geq \varepsilon$ to
$n := n + 1$, idź do kroku 1,
w przeciwnym razie
$\tilde{\tau} = \frac{1}{2}(a_{n+1} + b_{n+1})$ (STOP)

# Metoda $\gamma$ podziału




$$\frac{b_n - \alpha_n}{b_n - a_n} = \frac{\beta_n - a_n}{b_n - a_n} = \gamma$$
$$\alpha_n = b_n + \gamma(a_n - b_n)$$
$$\beta_n = a_n + \gamma(b_n - a_n)$$

$N = ?$

Dane: $a_0, b_0, \varepsilon, \gamma$

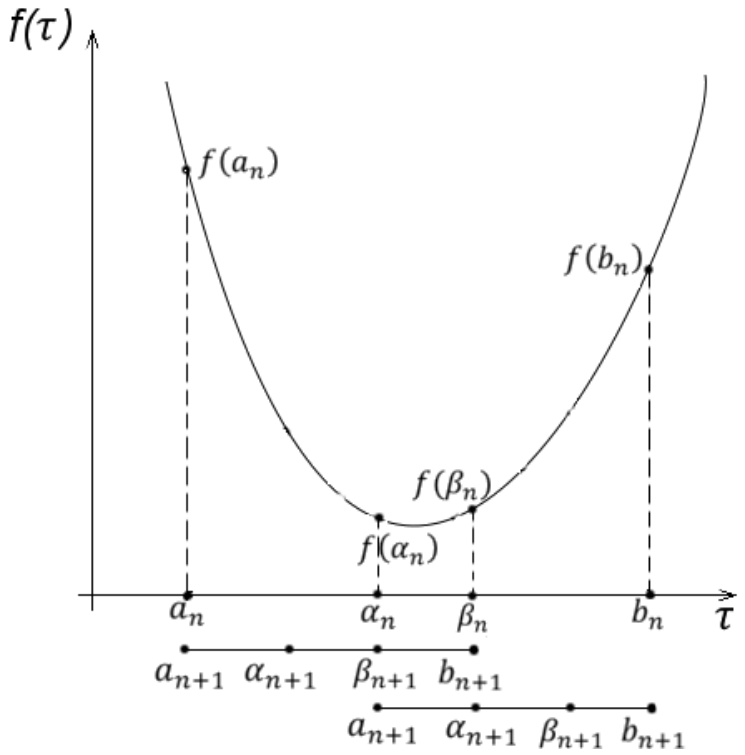
Krok 0: $n = 0$

Krok 1: $\alpha_n = b_n + \gamma(a_n - b_n)$
$\beta_n = a_n + \gamma(b_n - a_n)$

Krok 2: Jeśli $f(\alpha_n) \leq f(\beta_n)$ to
$a_{n+1} := a_n, b_{n+1} := \beta_n,$
w przeciwnym razie
$a_{n+1} := \alpha_n, b_{n+1} := b_n.$

Krok 3: Jeśli $|b_{n+1} - a_{n+1}| \geq \varepsilon$ to
$n := n + 1$, idź do kroku 1,
w przeciwnym razie
$\tilde{\tau} = \frac{1}{2}(a_{n+1} - b_{n+1})$ (STOP)

# Metoda złotego podziału




$$\gamma^2 + \gamma - 1 = 0$$

$$\gamma = \frac{\sqrt{5}-1}{2} \approx 0.618$$

$N = ?$

Dane: $a_0, b_0, \varepsilon, \gamma = \frac{\sqrt{5}-1}{2}$

Krok 0: $n = 0$

$\alpha_0 = b_0 + \gamma(a_0 - b_0)$

$\beta_0 = a_0 + \gamma(b_0 - a_0)$

Krok 1: Jeśli $|b_n - a_n| < \varepsilon$, to

$\tilde{\tau} = \frac{1}{2}(a_n + b_n)$(STOP)

w przeciwnym razie → krok 2

Krok 2: Jeśli $f(\alpha_n) \leq f(\beta_n)$ to

$a_{n+1} := a_n, b_{n+1} := \beta_n,$

$\beta_{n+1} := \alpha_n,$

$\alpha_{n+1} := \beta_n + \gamma(a_n - b_n)$

$n := n + 1$, idź do kroku 1

w przeciwnym razie

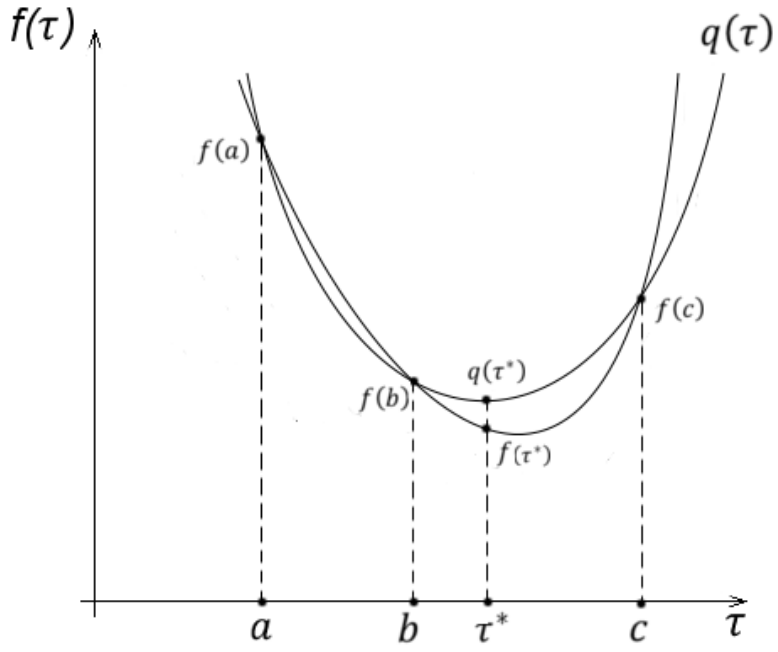
$a_{n+1} := \alpha_n, b_{n+1} := b_n,$

$\alpha_{n+1} := \beta_n,$

$\beta_{n+1} := \alpha_n + \gamma(b_n - \alpha_n)$

$n := n + 1$, idź do kroku 1

# Metoda aproksymacji kwadratowej



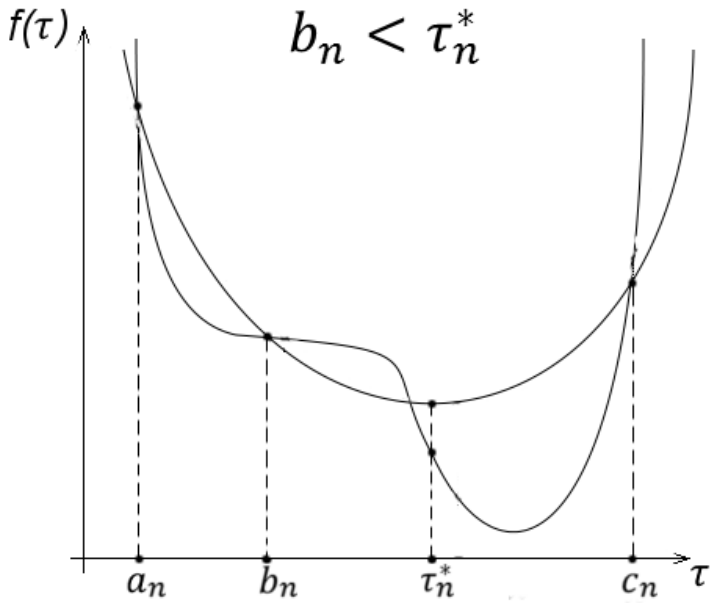

$$a < b < c$$
$$f(a) \geq f(b)$$
$$f(b) \leq f(c)$$

$q(\tau)$ – aproksymacja kwadratowa
$\tau^*$ - minimum funkcji $q(\tau)$

$$q(\tau) = \frac{f(a)(\tau - b)(\tau - c)}{(a-b)(a-c)} + \frac{f(b)(\tau - a)(\tau - c)}{(b-a)(b-c)} + \frac{f(c)(\tau - a)(\tau - b)}{(c-a)(b-c)}$$

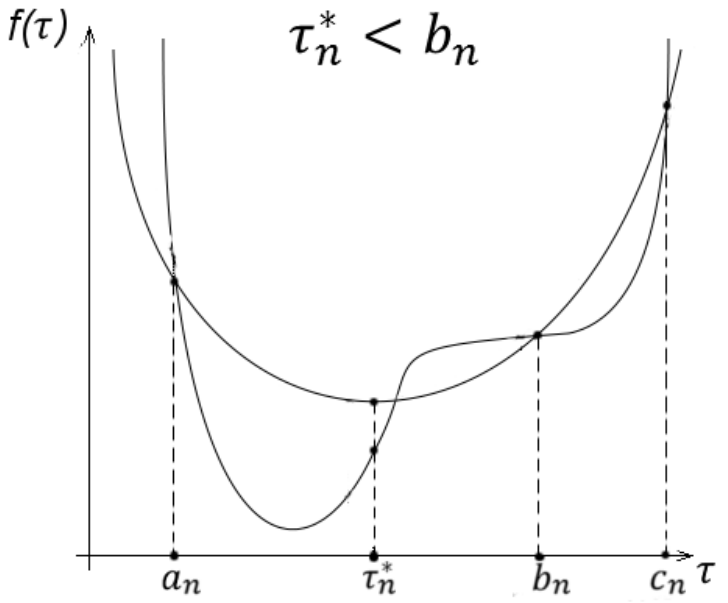
$$\tau^* = \frac{1}{2}\frac{f(a)(b^2 - c^2) + f(b)(c^2 - a^2) + f(c)(a^2 - b^2)}{f(a)(b-c) + f(b)(c-a) + f(c)(a-b)}$$

# Metoda aproksymacji kwadratowej




$$f(b_n) \geq f(\tau_n^*)$$
$$a_{n+1} := b_n$$
$$b_{n+1} := \tau_n^*$$
$$c_{n+1} := c_n$$

$$f(b_n) < f(\tau_n^*)$$
$$a_{n+1} := a_n$$
$$b_{n+1} := b_n$$
$$c_{n+1} := \tau_n^*$$

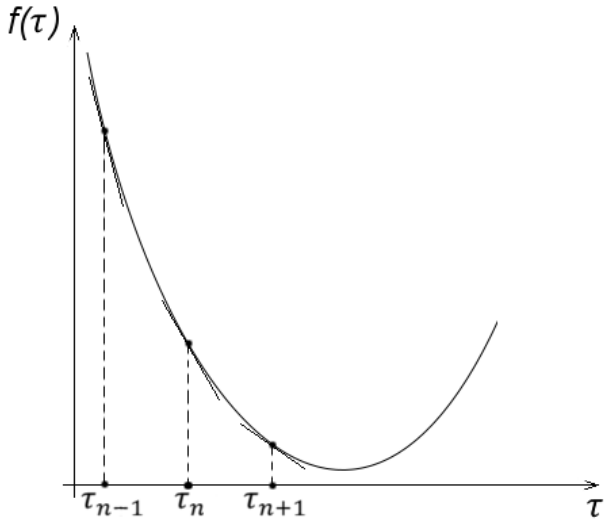



$$f(b_n) \geq f(\tau_n^*)$$
$$a_{n+1} := a_n$$
$$b_{n+1} := \tau_n^*$$
$$c_{n+1} := b_n$$

$$f(b_n) \geq f(\tau_n^*)$$
$$a_{n+1} := \tau_n^*$$
$$b_{n+1} := b_n$$
$$c_{n+1} := c_n$$

$$|c_{n+1} - a_{n+1}| < \varepsilon \qquad \tilde{\tau} = \tau_{n+1}^*$$

# Metoda pierwszej pochodnej




$$\tau_{n+1} = \tau_n - \gamma_n f'(t_n) \qquad \gamma_n > 0, \tau_0$$

$$\lim_{n\to\infty} \gamma_n = \gamma$$

$$\sum_{n=0}^{\infty} \gamma_n = \infty$$

np. $|\tau_{n+1} - \tau_n| < \varepsilon$ (STOP)

$\tau_0$

$\tau_1 = \tau_0 - \gamma_0 f'(\tau_0)$

$\tau_2 = \tau_1 - \gamma_1 f'(\tau_1) = \tau_0 - \gamma_0 f'(\tau_0) - \gamma_1 f'(\tau_1)$

$\tau_{n+1} = \tau_n - \gamma_n f'(\tau_n) = \cdots = \tau_0 - \gamma_0 f'(\tau_0) - \gamma_1 f'(\tau_1) - \cdots - \gamma_n f'(\tau_n)$

$$|\tau_{n+1} - \tau_0| = |\sum_{k=0}^{n} \gamma_k f'(\tau_k)| \le \sum_{k=0}^{n} \gamma_k |f'(\tau_k)| \le \max_{0<k<n} |f'(\tau_k)| \sum_{k=0}^{n} \gamma_k$$

$$|\tau_\infty - \tau_0| \le \sum_{k=0}^{\infty} \gamma_k = \infty$$

# Metoda znaku pierwszej pochodnej

$$\tau_{n+1} = \tau_n - \vartheta_n sign[f'(\tau_n)]$$

$$\gamma_n f'(\tau_n) = \gamma_n |f'(\tau_n)| * sign\, f'(\tau_n) = \vartheta_n\, sign[f'(\tau_n)], \text{ gdzie } \vartheta_n = \gamma_n |f'(\tau_n)|$$

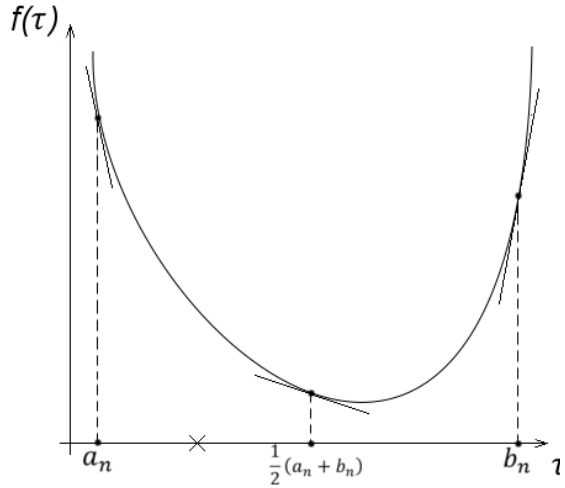
$$\vartheta_n > 0$$

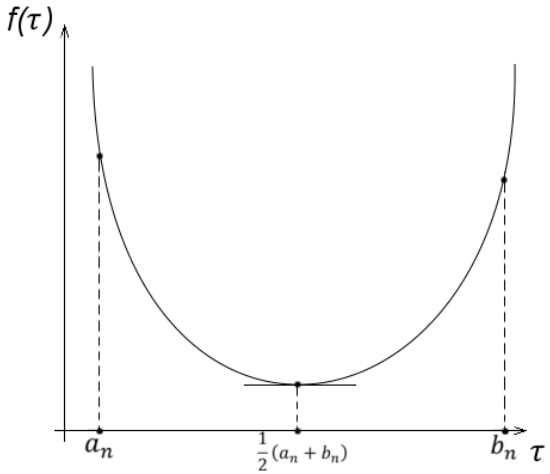
$$\lim_{n\to\infty} \vartheta_n = 0, \text{ bo } \lim_{n\to\infty} |f'(\tau_n)| = 0, \lim_{n\to\infty} \gamma_n = \gamma$$

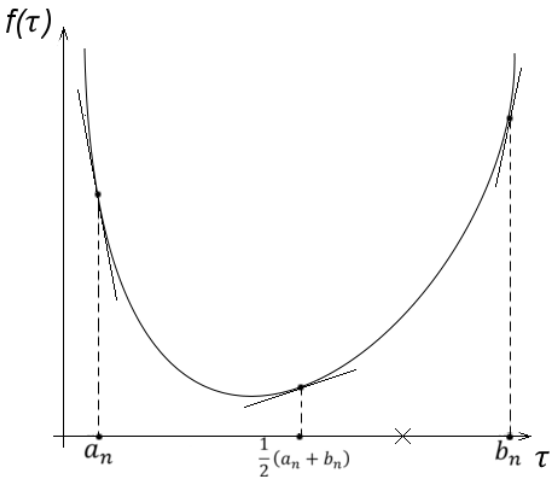
$$\sum_{n=0}^{\infty} \vartheta_n = \infty \qquad \lim_{n\to\infty} \vartheta_n = \lim_{n\to\infty} \gamma_n |f'(\tau_n)| = 0$$

# Metoda Bolzano

$$sign\, a_n \neq sign\, b_n$$




$$sign\, a_n = sign(f'(\frac{1}{2}(a_n + b_n)))$$

$$a_{n+1} := \frac{1}{2}(a_n + b_n)$$

$$b_{n+1} := b_n$$

$$f'\left(\frac{1}{2}(a_n + b_n)\right) = 0$$

$$\tilde{\tau} := \frac{1}{2}(a_n + b_n)$$

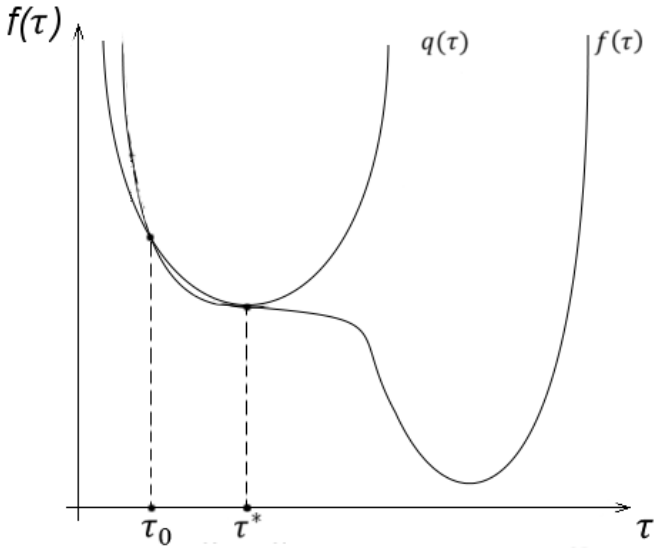
(rzadko)

$$sign\, b_n = sign(f'(\frac{1}{2}(a_n + b_n)))$$

$$a_{n+1} := a_n$$

$$b_{n+1} := \frac{1}{2}(a_n + b_n)$$

# Metoda drugiej pochodnej




$$\tau_0$$

$$\tau_{n+1} = \tau_n - \frac{f'(\tau_n)}{f''(\tau_n)}$$

$$|\tau_{n+1} - \tau_n| < \varepsilon \quad \text{(STOP)}$$

$$f(\tau) = \underbrace{f(\tau_0) + (\tau - \tau_0)f'(\tau_0) + \frac{1}{2}(\tau - \tau_0)^2 f''(\tau_0)}_{q(\tau)} + 0_3(|\tau - \tau_0|)$$

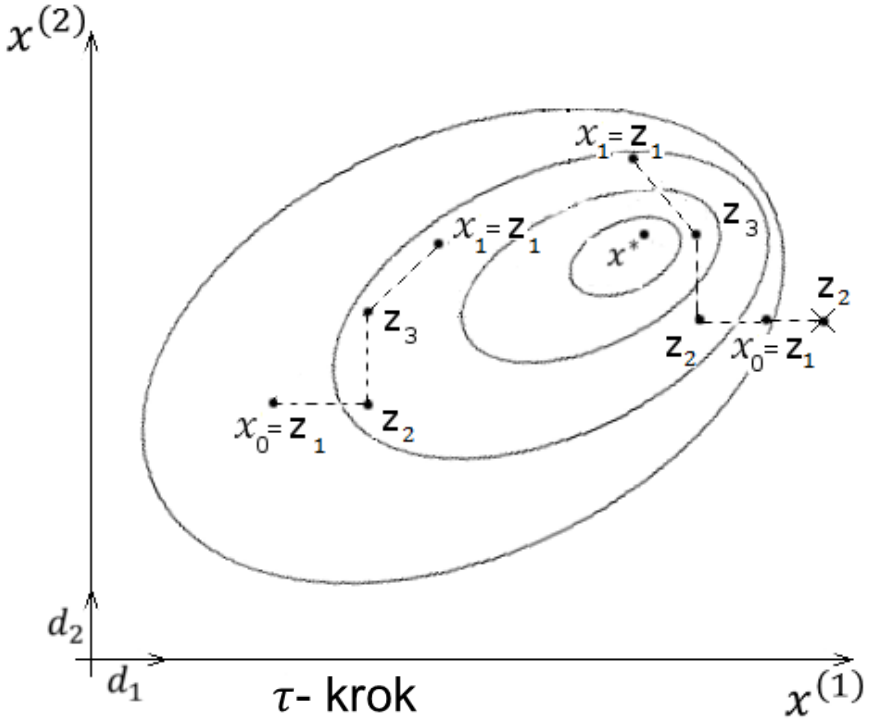
$$q'(\tau) = f'(\tau_0) + (\tau^* - \tau_0)f''(\tau_0) = 0$$

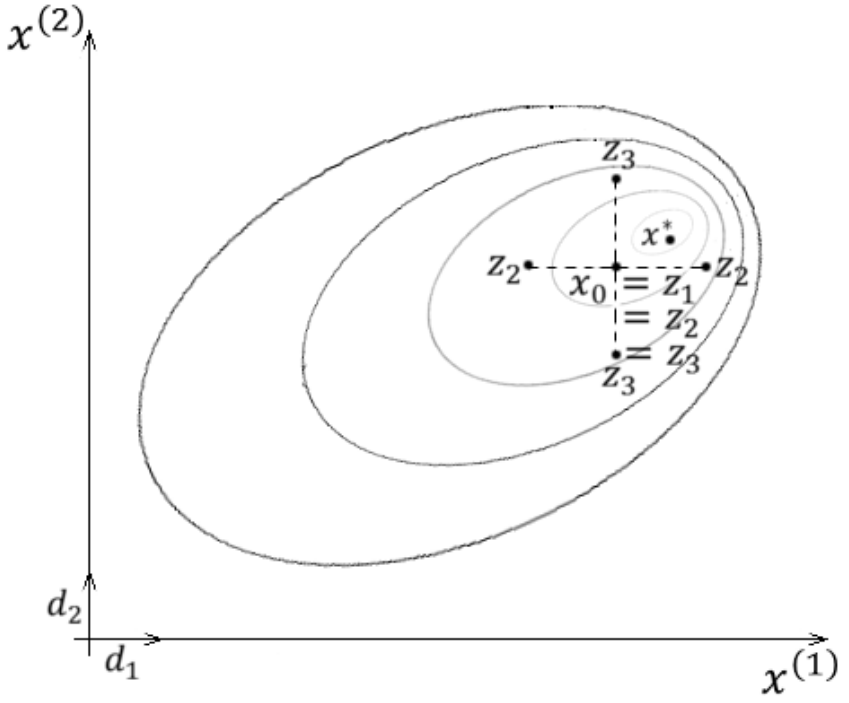
$$\tau^* = \tau_0 - \frac{f'(\tau_0)}{f''(\tau_0)}$$

# Bezgradientowe metody optymalizacji

- Bezgradientowe metody optymalizacji
  - Hooka-Jeevesa (z krokiem dyskretnym i optymalnym)
  - Rosenbrocka (z krokiem dyskretnym i optymalnym)
  - Powella
  - Neldera Meada

# Metoda Hooka-Jeveesa – z krokiem dyskretnym



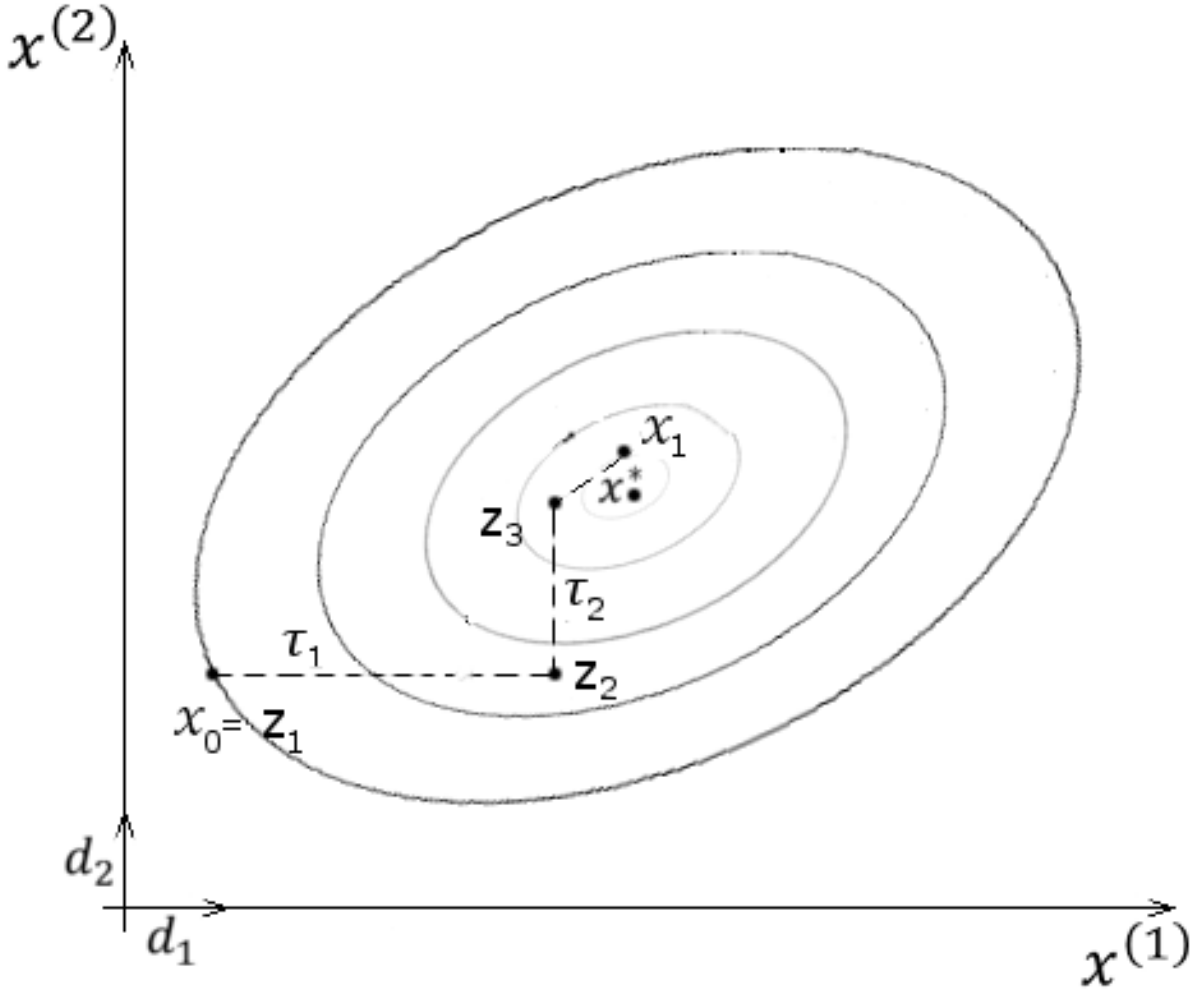

$\tau$- krok

$\alpha > 1$ współczynnik kroku roboczego

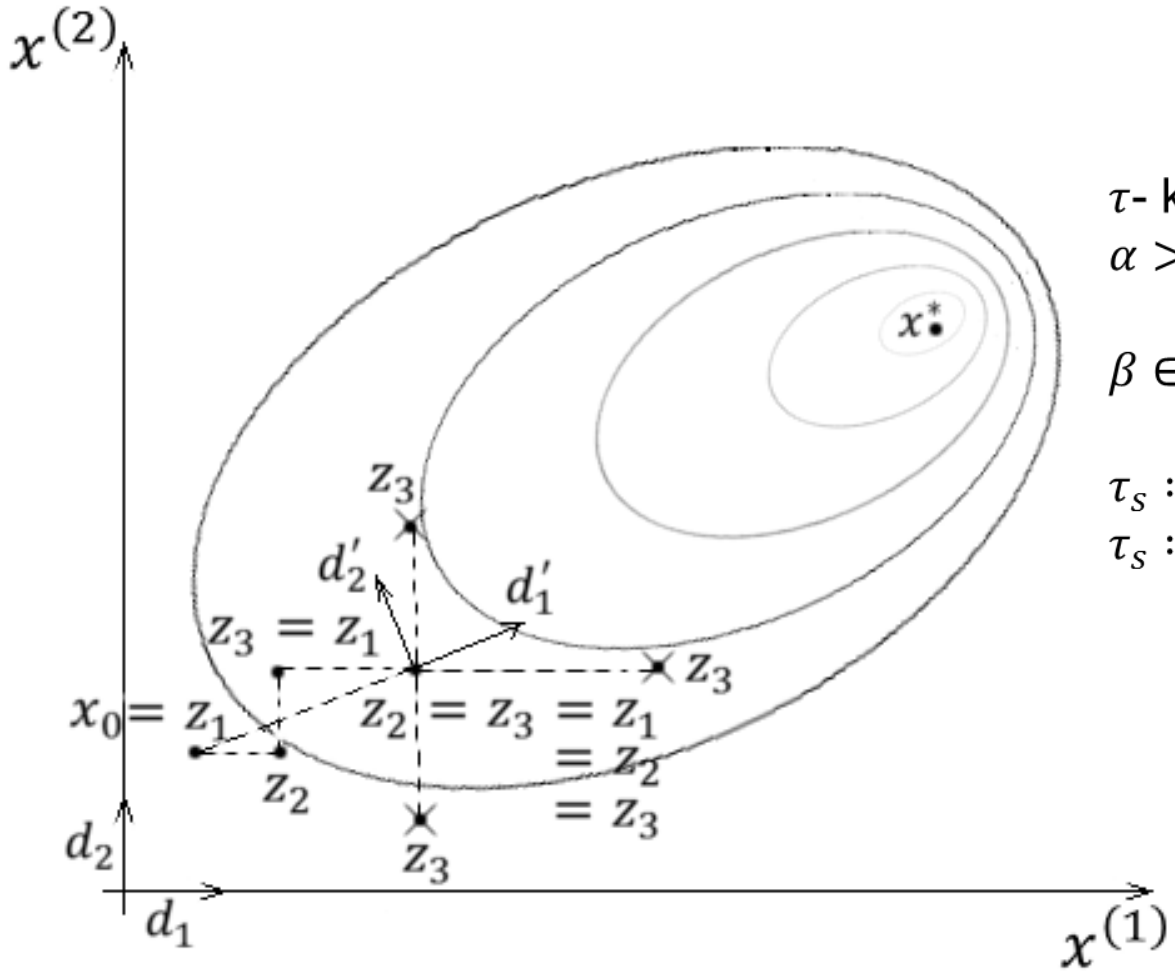
$\beta \in (0,1)$ współczynnika korekcji kroku

$\tau := \tau\beta$

# Metoda Hooka-Jeveesa – z krokiem optymalnym



# Metoda Rozenbrocka – z krokiem dyskretnym




$\tau$- krok

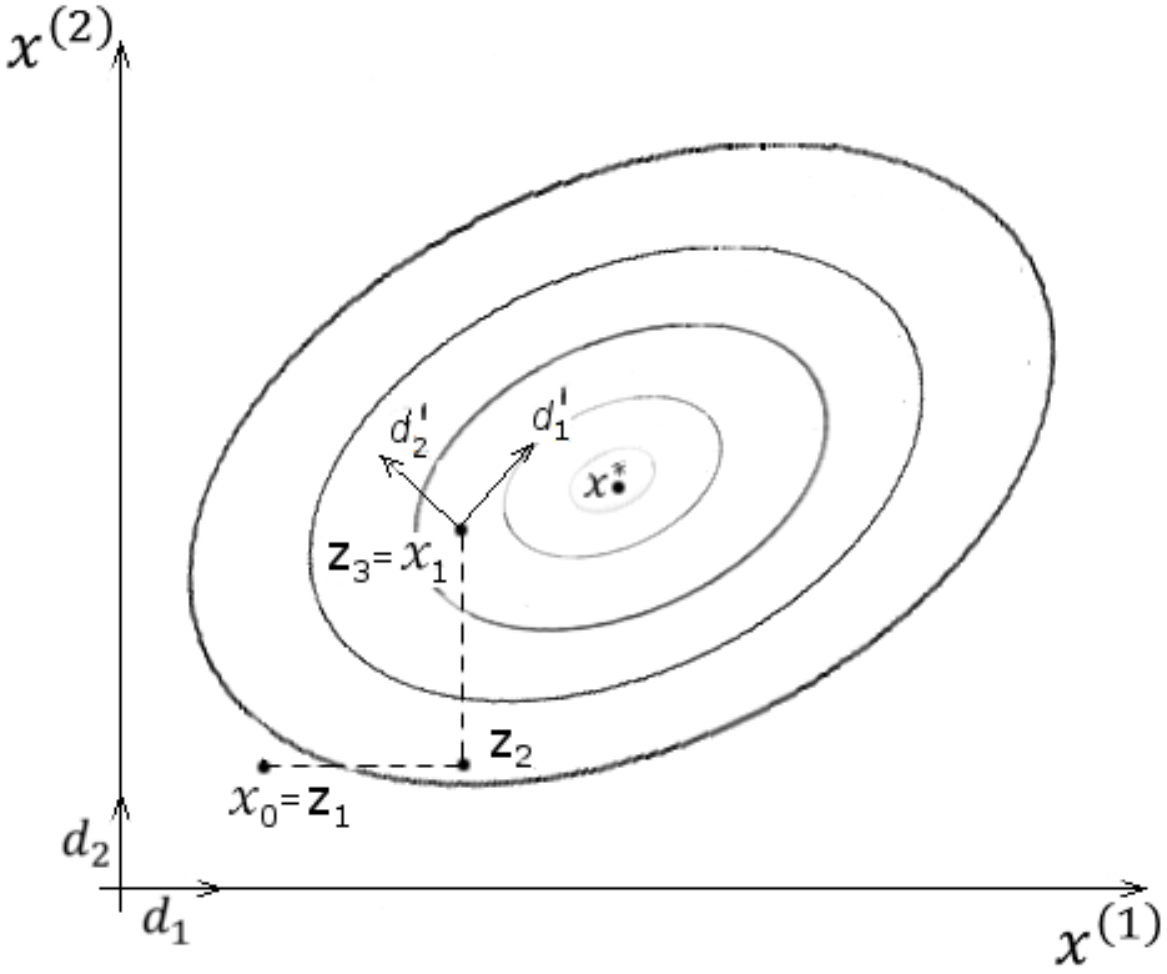
$\alpha > 1$ – współczynnik korekcji kroku

$\beta \in (-1, 0)$ - współczynnika korekcji kroku

$\tau_s := \tau_s \alpha$

$\tau_s := \tau_s \beta$

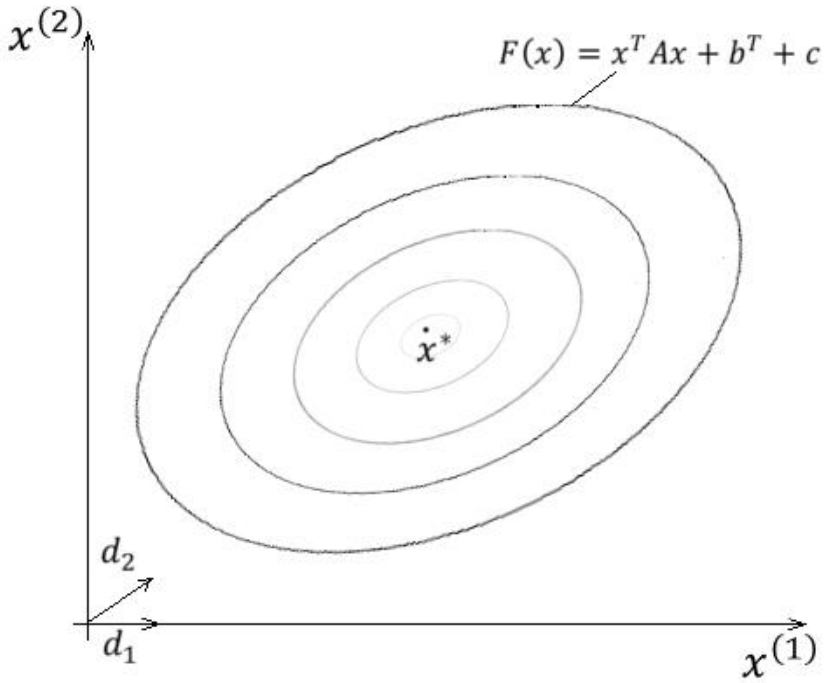
# Metoda Rozenbrocka – z krokiem optymalnym



# Metoda Powella – wektory sprzężone

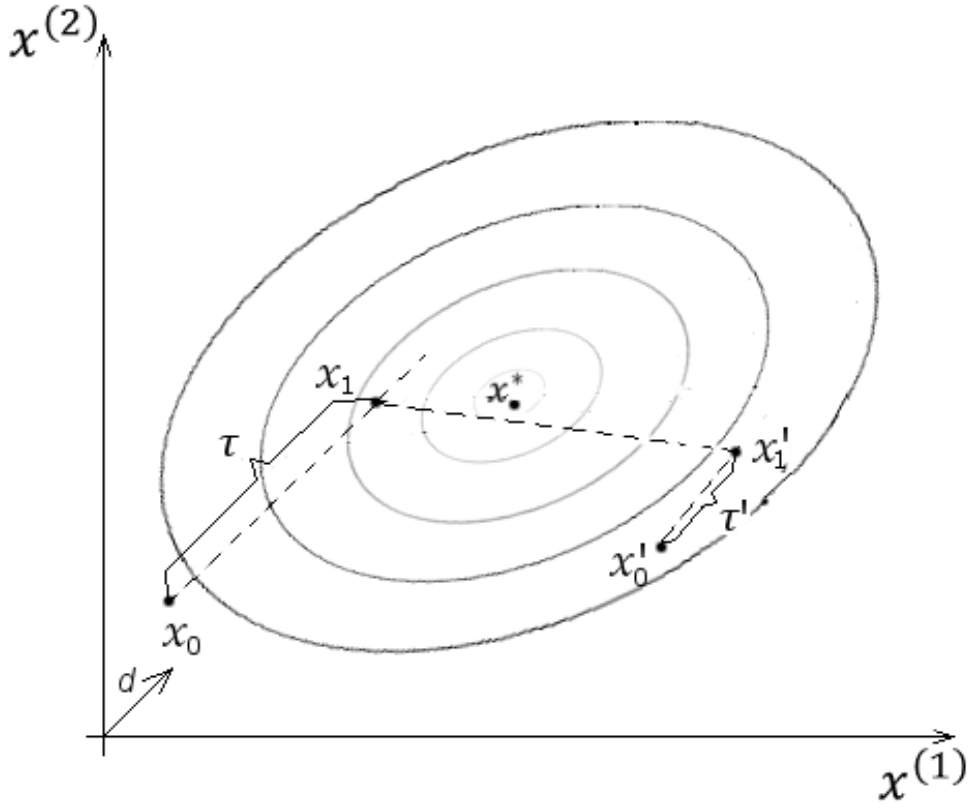
$d_1, d_2, \ldots, d_S$ - sprzężone według macierzy $A$, symetrycznej i dodatnio określonej

$$d_i^T A d_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$



# Metoda Powella

$$F(x) = x^T A x + b^T + c$$

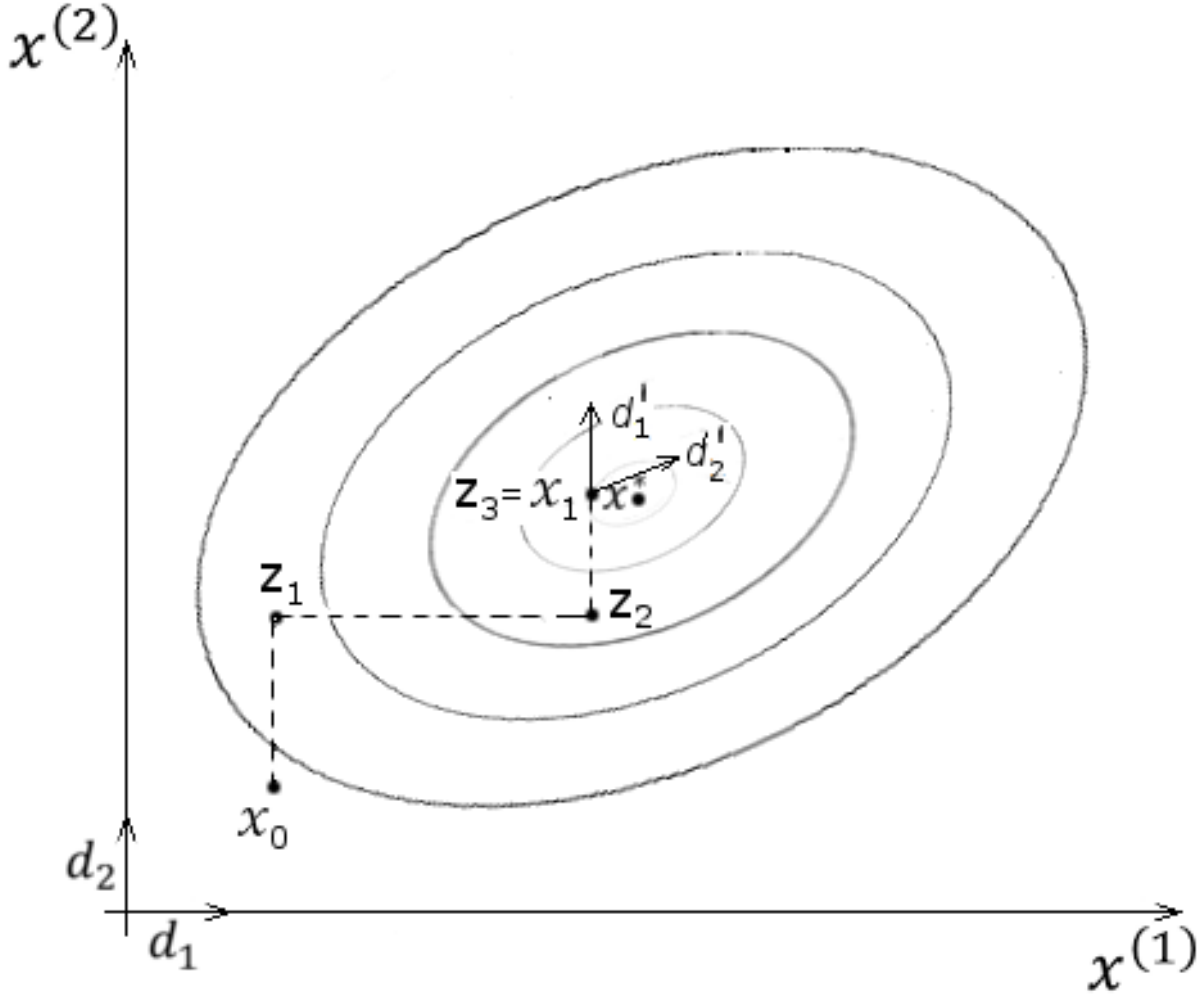



$x_1 = x_0 + x^* d$
$\tau^*$ - min w kierunku $d$ z $x_0$
$x_1' = x_0' + \tau^{*\prime} d$
$\tau^{*\prime}$ - min w kierunku $d$ z $x_0'$

$d^T A d' = 0$
$d, d'$- sprzężone według $A$

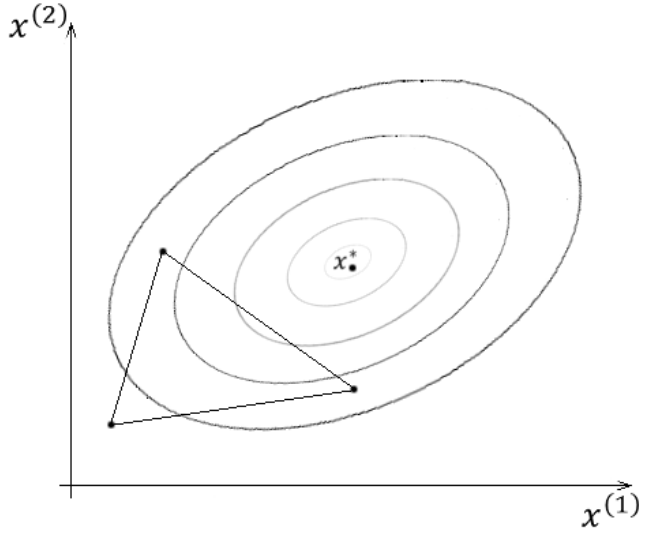
$$d' = \frac{x_1' - x_1}{\|x_1' - x_1\|}$$

# Metoda Powella



# Metoda Neldera-Meada (simplex)

$x_1\ x_2\ \ldots x_{S+1}$ - simplex w przestrzeni $s$-wymiarowej




$$x_H \rightarrow F(x_H) = \max_{1 \le s \le S+1} F(x_s)$$

$$x_L \rightarrow F(x_L) = \min_{1 \le s \le S+1} F(x_s)$$

$$\bar{x} = \frac{1}{s} \sum_{s=1, s \neq H}^{S+1} x_s$$

Generowanie simplexu początkowego

$x_0, c$

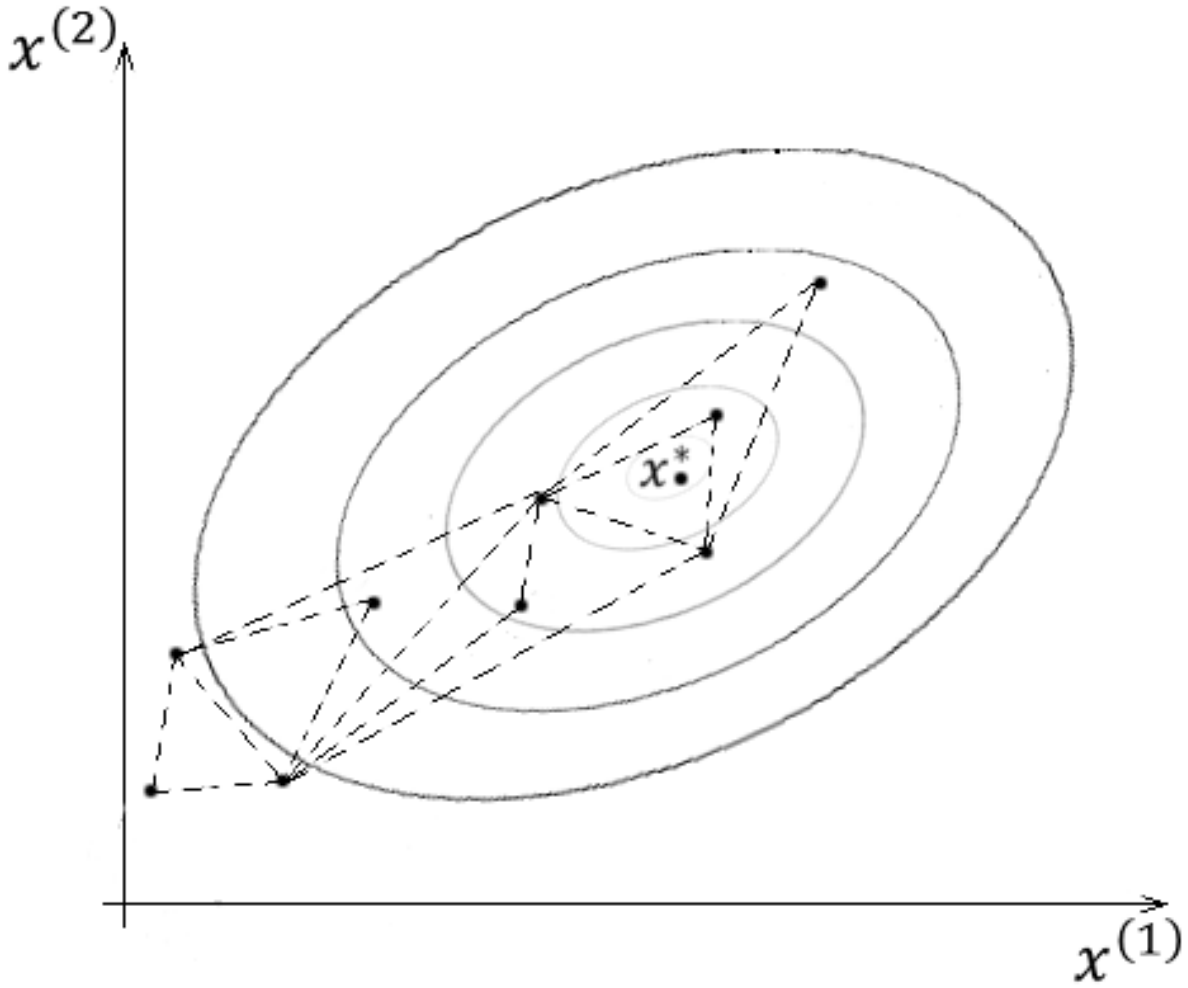
$$a = \frac{c}{S\sqrt{2}}(\sqrt{S+1} + \sqrt{2} - 1)$$

$$\mathrm{b} = \frac{c}{S\sqrt{2}}(\sqrt{S+1} - 1)$$

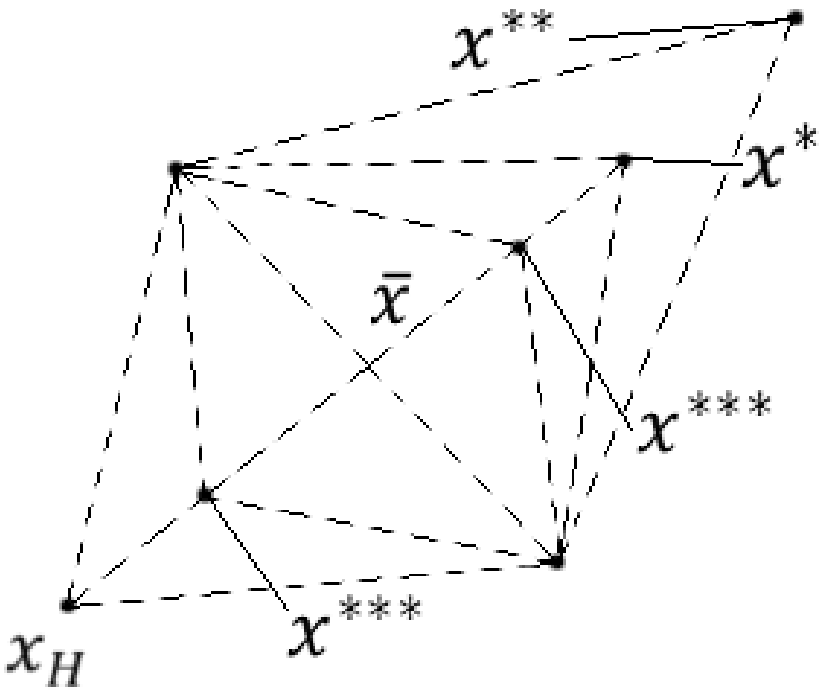
$d_j = [\quad]$

$$x_i = x_0 + d_j, x_{S+1} = x_0$$

# Metoda Neldera-Meada (simplex)



# Metoda Neldera-Mead




Odbicie

$x^* = \bar{x} + \alpha(\bar{x} - x_H)$

$\alpha$ – współczynnik odbicia

Jeśli $\alpha > 0$

$F(x^*) < F(x_L)$

Ekspansja

$x^{**} = \bar{x} + \gamma(x^* - \bar{x}) \qquad \gamma > 1$

$\gamma$ – współczynnik ekspansji

Kontrakcja

Jeśli $F(x^*) > F(x_H)$

$x^{***} = \bar{x} + \beta(x_H - \bar{x})$

Jeśli $F(x^*) > \max\limits_{\substack{1 \le s \le S+1 \\ s \neq H}} F(x_s)$

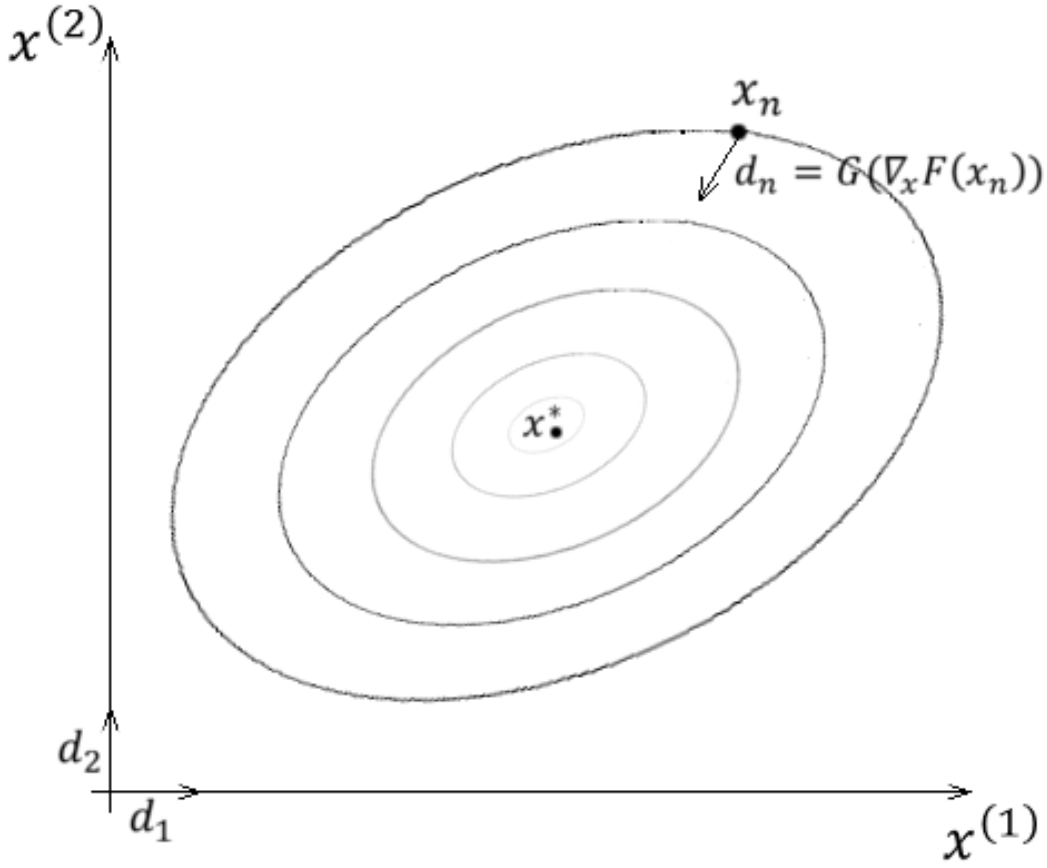
$x^{***} = \bar{x} + \beta(x^* - \bar{x}) \qquad \beta \in (0, 1)$

β – współczynnik kontrakcji

# Gradientowe metody optymalizacji

- Gradientowe metody optymalizacji
  - Gradientu prostego
  - Najszybszego spadku
  - Newtona
  - Gradientu sprzężonego
  - Zmiennej metryki

# Metody gradientowe

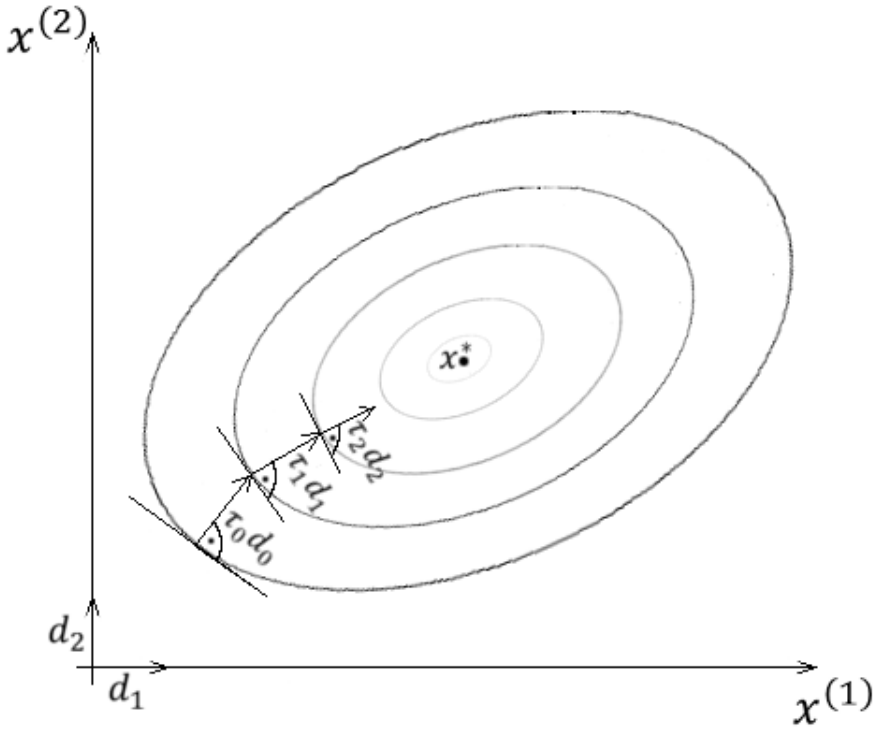



$x_{n+1} = x_n + \tau_n d_n$
$d_n = G(\nabla_x F(x_n))$
$\tau_n$- krok procedury

# Metoda gradientu prostego

$$x_{n+1} = x_n + \tau_n d_n$$

$$d_n = -\nabla_x F(x_n)\,;\tau_n > 0, \lim_{n\to\infty} \tau_n = \tau, \sum_{n=0}^{\infty} \tau_n = \infty$$

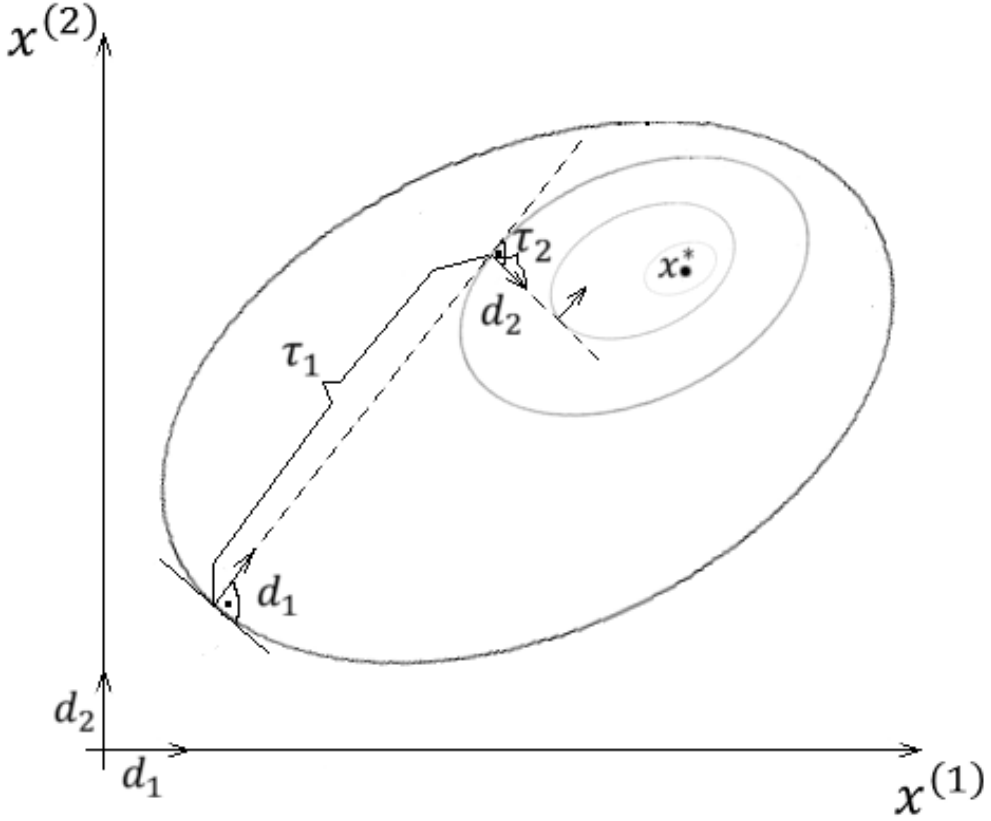



$$\|x_{n+1} - x_n\| = \|\tau_n d_n\| < \varepsilon$$

# Metoda najszybszego spadku

$x_{n+1} = x_n + \tau_n d_n$

$d_n = -\nabla_x F(x_n)$, $\tau_n$- minimum w kierunku $d_n$




$\|x_{n+1} - x_n\| < \varepsilon$

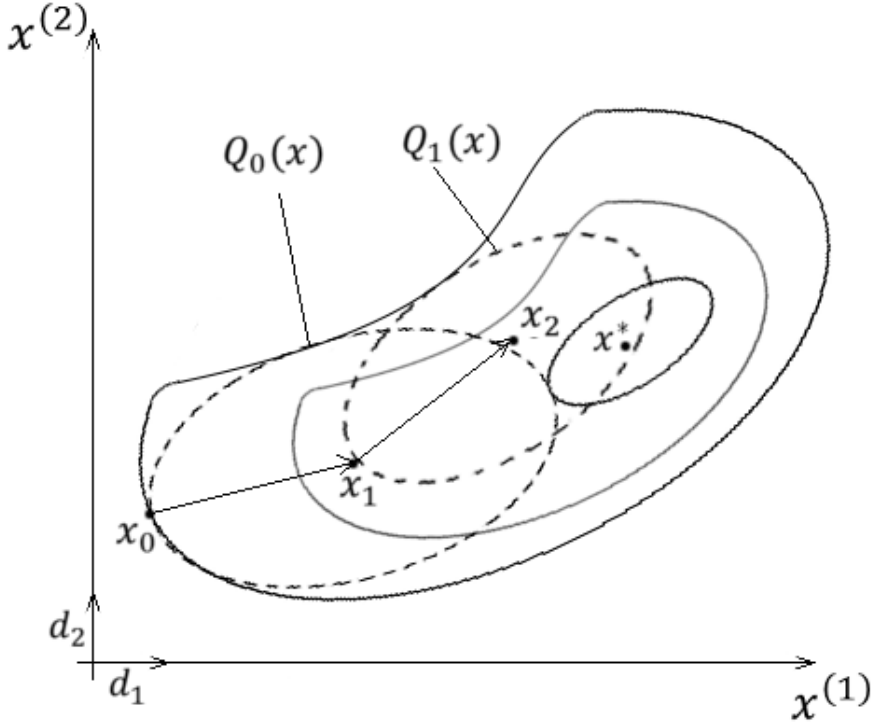
# Metoda Newtona

$$F(x) = \underbrace{F(x_0) + (x - x_0)^T \nabla_x F(x_0) + \frac{1}{2}(x - x_0)^T H(x_0)(x - x_0)}_{Q(x)} + 0_3(\|x - x_0\|)$$

$$\nabla_x Q(x) = \nabla_x F(x_0) + H(x_0)(x^* - x_0) = O_S$$

$$x^* = x_0 - H^{-1}(x_0)\nabla_x F(x_0)$$

$$x_{n+1} = x_n - H^{-1}(x_n)\nabla_x F(x_n)$$



# Metoda Fletchera-Reevesa (gradientu sprzężonego)

Krok 0: $z_1 = x_0,\ s = 1,\ d_1 := -\nabla_x F(z_1)$

Krok 1: $z_{s+1} := z_s + \tau_s d_s$

$\tau_s \rightarrow$ minimum w kierunku $d_s$

Jeśli $\|\tau_s d_s\| < \varepsilon$ (STOP)

w przeciwnym razie idź do kroku 2

Krok 2: $d_{s+1} := -\nabla_x F(z_{s+1}) + \frac{\|\nabla_x F(z_{s+1})\|}{\|\nabla_x F(z_s)\|} d_s$

$s := s + 1,$ idź do kroku 1

$d_1, d_2, \ldots, d_S$ - kierunki sprzężone dla formy kwadratowej

# Metoda zmiennej metryki

Krok 0: $z_1 = x_0,\ d_1 = -D_1 \nabla_x F(z_1), \quad D_1 = I,\ s{=}1$

Krok 1: $z_{s+1} = z_s + \tau_s d_s$ gdzie $\tau_s$ - minimum w kierunku $d_s$

Jeśli $\|\tau_s d_s\| < \varepsilon$ (STOP)

w przeciwnym razie idź do kroku 2.

Krok 2: $d_{s+1} = -D_{s+1} \nabla_x F(z_{s+1})$

$D_{s+1} = D_s + \frac{p_s p_s^T}{p_s^T q_s} - \frac{D_s q_s q_s^T D_s}{q_s^T D_s q_s}$, gdzie

$p_s = \tau_s d_s,\ q_s = \nabla_x F(z_{s+1}) - \nabla_x F(z_s)$

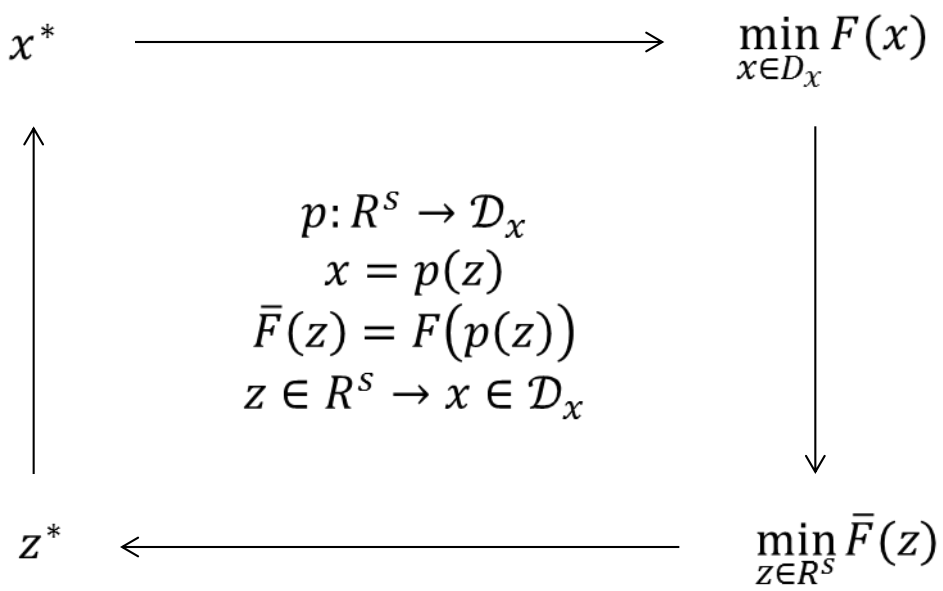
$s := s+1$ idź do kroku 1.

$D_{S+1} \approx H^{-1}(x_{S+1})$

# Metody optymalizacji z ograniczeniami

- Metody optymalizacji z ograniczeniami
  - Transformacji zmiennych
  - Funkcji kary
    - Kary zewnętrznej (kary)
    - Kary wewnętrzne (bariery)
  - Metoda compleks
  - Poszukiwania losowe

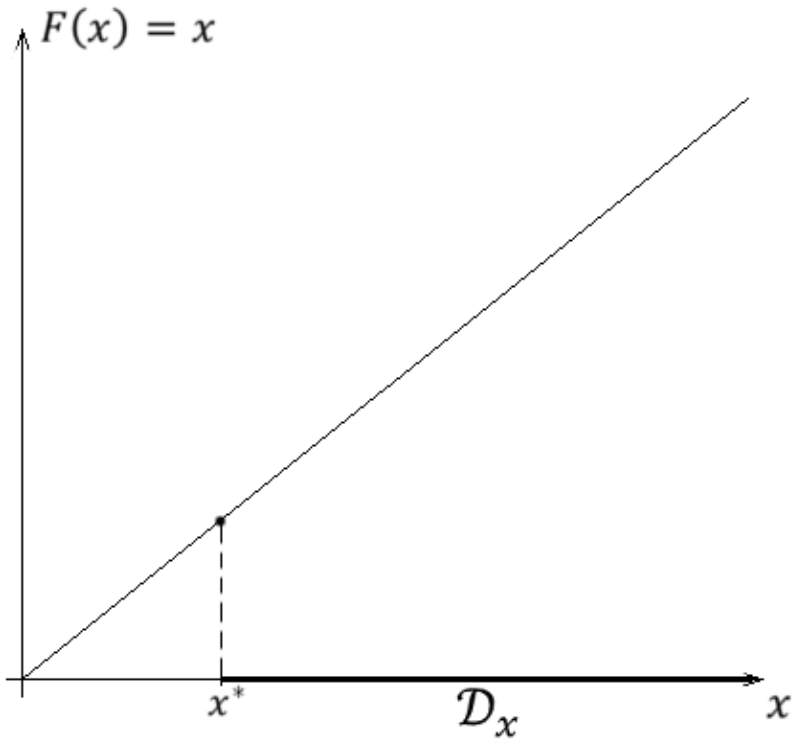
# Metoda transformacji zmiennych



# Przykład

$F(x) = x$

$\mathcal{D}_x = \{x \in R^1 \;\; x \geq 1\}$




$$x^* \rightarrow \min_{x \geq 1} x$$

$$x = z^2 + 1 = p(z)$$

$$z \in R \rightarrow x \in [1, \infty)$$

$$\bar{F}(z) = z^2 + 1$$

$$z^* \rightarrow \min_{z \epsilon R} \bar{F}(z)$$

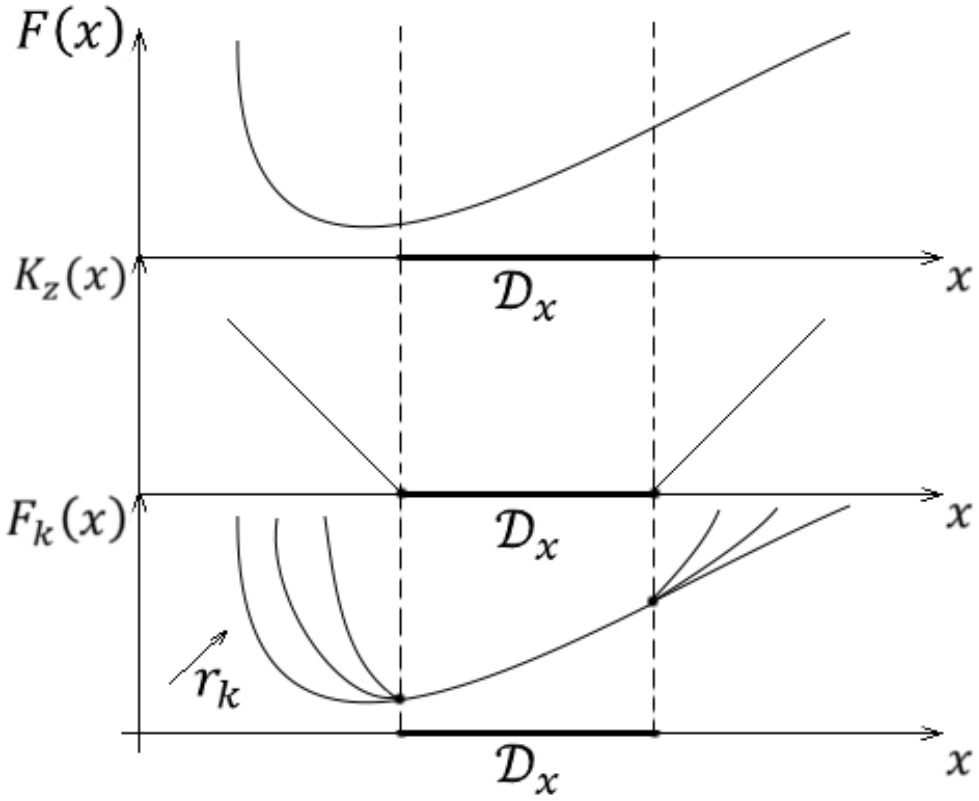
$$z^* \rightarrow \min_{z \epsilon R} (z^2 + 1)$$

$$(z^2 + 1)' = 2z = 0$$

$$z^* = 0$$

$$x^* = p(z^*) = (z^*)^2 + 1 = 1$$

# Metody funkcji kary
# Kary zewnętrzne (kary)




$r_k > 0$

$\lim_{k \to \infty} r_k = \infty$

# Przykłady funkcji kary zewnętrznej

$$\mathcal{D}_x = \{x \in R^S,\ \varphi_l(x) = 0, \qquad l = 1, 2, \dots, L, \qquad \psi_m(x) \leq 0, \qquad m = 1, 2, \dots, M\}$$

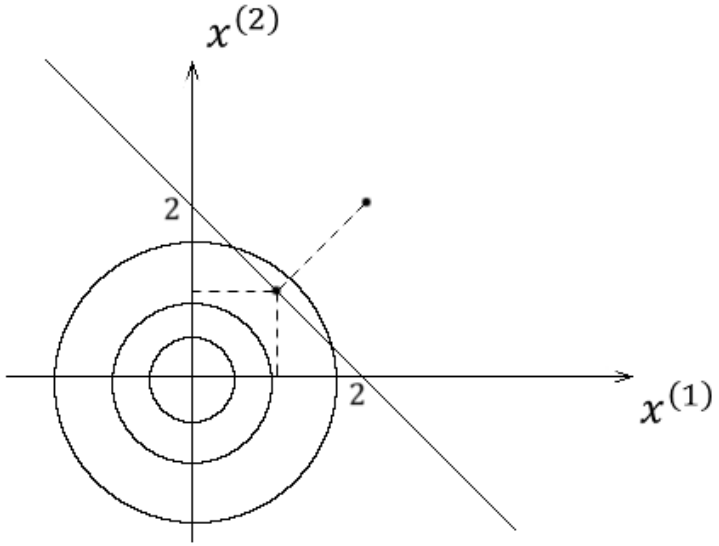
$$\varphi_l(x) \to K_{lZ}(x) = \left(\varphi_l(x)\right)^2,\ l = 1, 2, \dots, L$$

$$\psi_m(x) \to K_{mz}(x) = (\max\{0, \psi_m(x)\})^2, m = 1, 2, \dots, M$$

$$K_z(x) = \sum_{l=1}^{L} r_l\left(\varphi_l(x)\right)^2 + \sum_{m=1}^{M} \rho_m\ \max\{0, \varphi_m(x)\}$$

# Przykład

$$F(x) = \left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2$$
$$\mathcal{D}_x = \{x \in R^2, x^{(1)} + x^{(2)} - 2 = 0\}$$




$$F_k(x) = \left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2 + r_k\left(x^{(1)} + x^{(2)} - 2\right)^2$$

$$F_k'(x) = 0$$

$$2x^{(1)} + 2r_k\left(x^{(1)} + x^{(2)} - 2\right) = 0$$

$$2x^{(2)} + 2r_k\left(x^{(1)} + x^{(2)} - 2\right) = 0$$

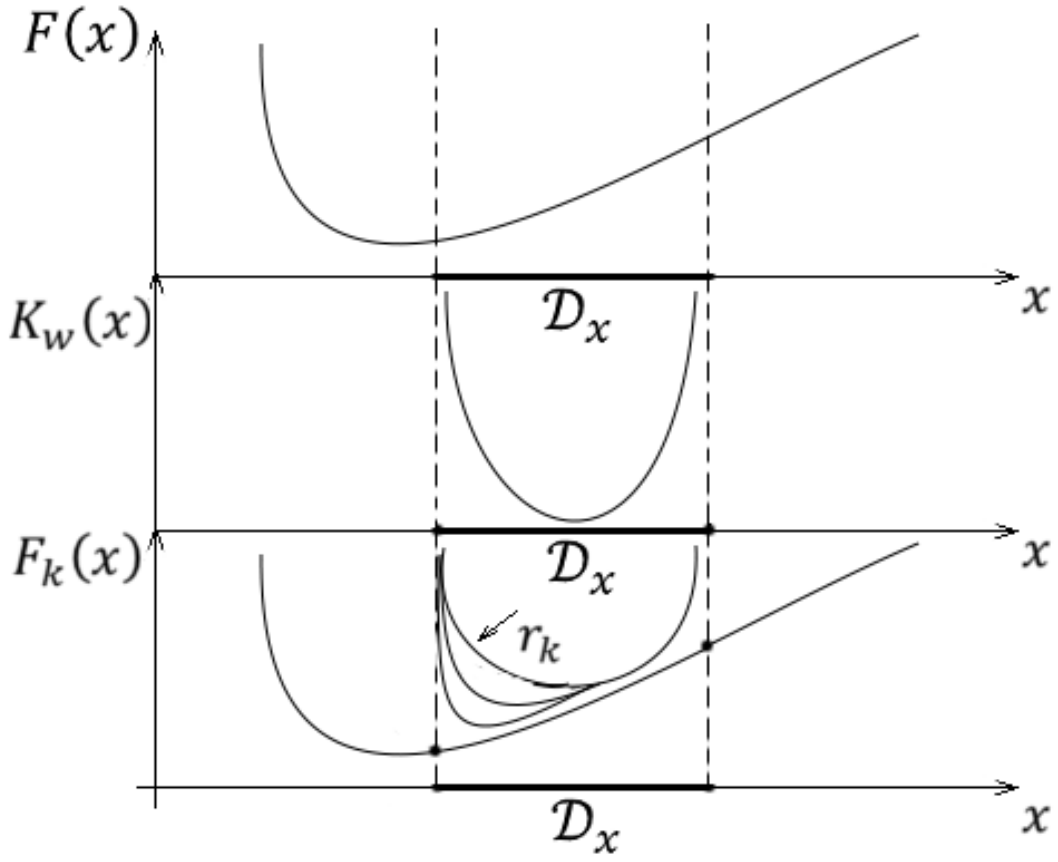
$$x^{(1)} = x^{(2)}$$

$$+2r_k\left(x^{(1)} + x^{(2)} - 2\right) = 0$$

$$x^{(1)} = x^{(2)} = \frac{2r_k}{2r_k + 1}$$

$$x^{(1)} = x^{(2)} = \lim_{r_k \to \infty} \frac{2r_k}{2r_k + 1} = 1, \qquad x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

# Kary wewnętrzne (bariery)




$$r_k > 0 \qquad \lim_{k\to\infty} r_k = 0$$

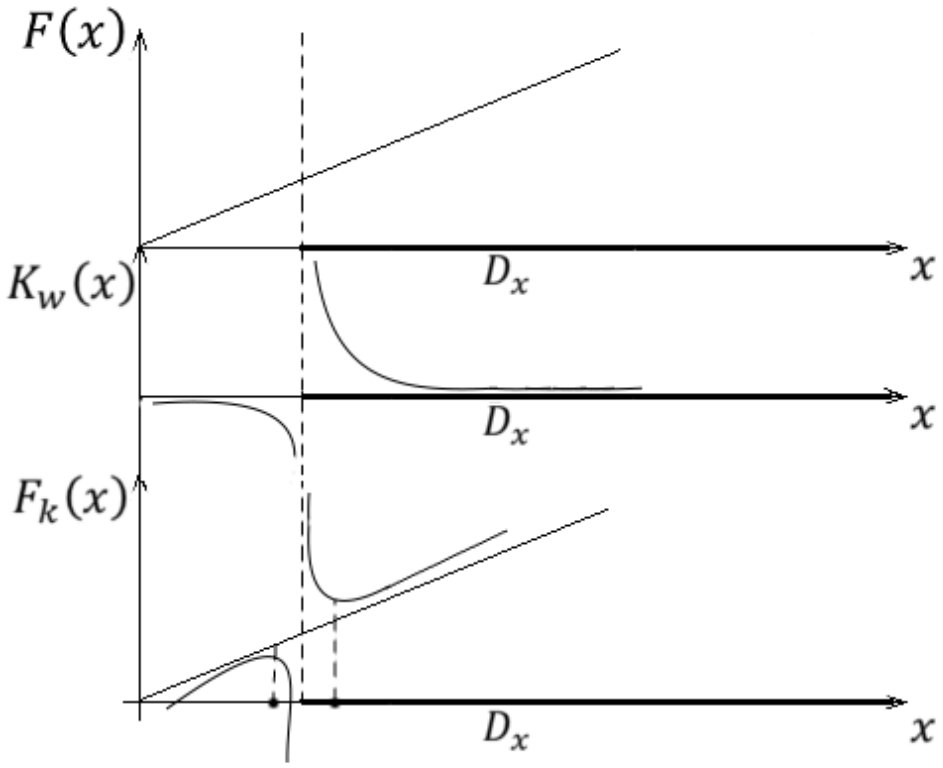
# Kara wewnętrzna (metoda Carrola)

$$\mathcal{D}_x = \{x \in R^s, \qquad \psi_m(x) \leq 0, \qquad m = 1, 2, \dots, M\}$$

$$K_w(x) = \sum_{m=1}^{M} \frac{-r_m}{\psi_m(x)}, \quad m = 1, 2, \dots, M$$

# Przykład

$$F(x) = x;\ \ \mathcal{D}_x = \{x \in R^1, x \geq 1\} \equiv \{x \in R^1, 1 - x \leq 0\}$$




$$K_w(x) = \frac{-1}{1-x} = \frac{1}{x-1}$$

$$F_k(x) = x + r_k \frac{1}{x-1}$$

$$F'(x) = 1 + \frac{-r_k}{(x-1)^2} = 0$$

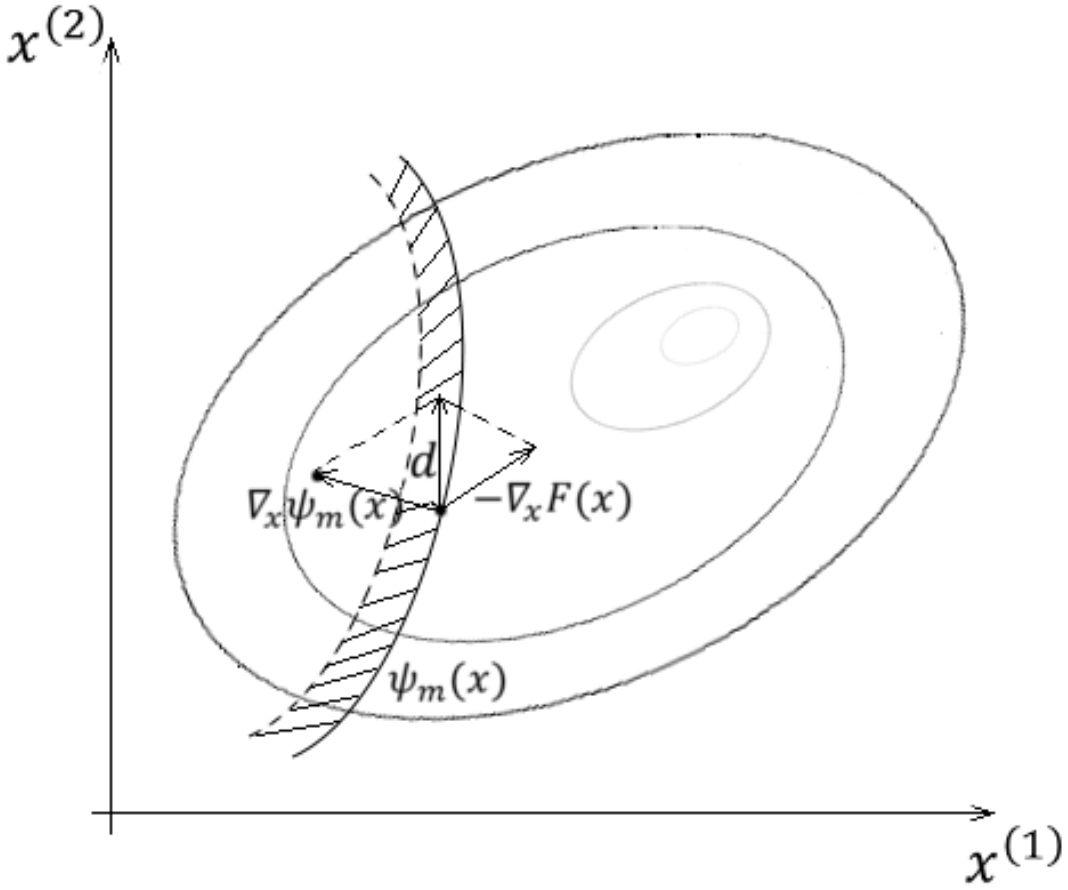
$$x_k = 1 \pm \sqrt{r_k}$$

$$x_k = 1 - \sqrt{r_k} \notin \mathcal{D}_x$$

$$x_k = 1 + \sqrt{r_k} \in \mathcal{D}_x$$

$$\lim_{n \to \infty} x_n = 1$$

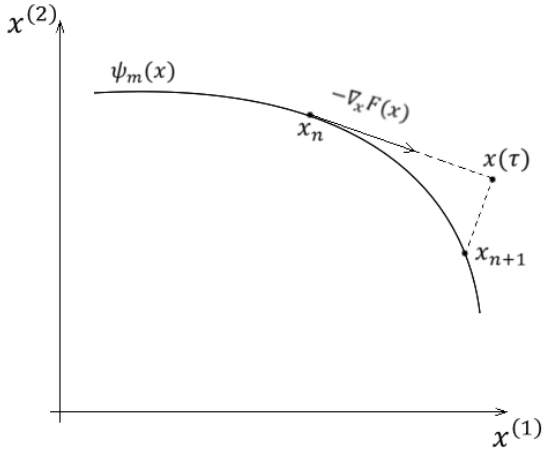
# Metoda modyfikacji kierunków

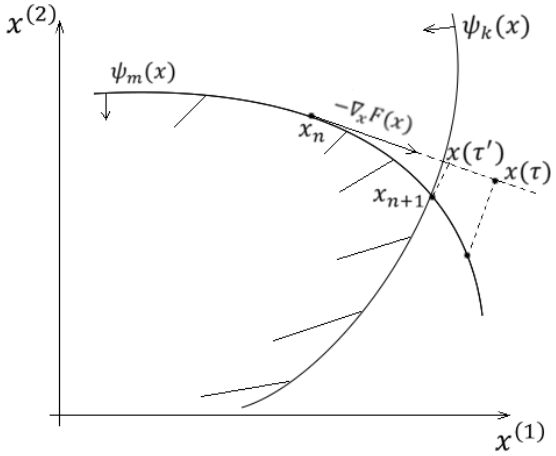


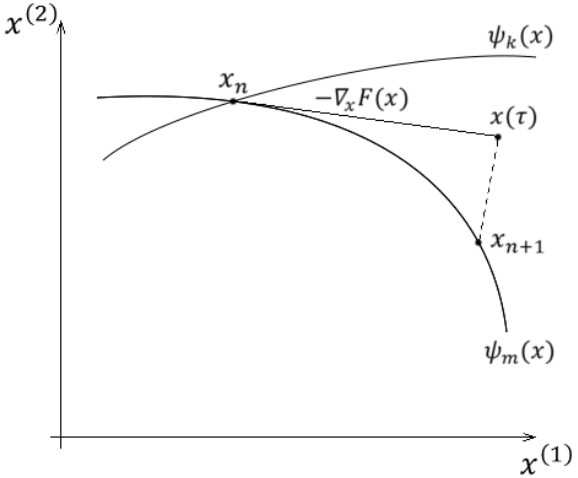

$$d = \frac{\nabla_x \psi_m(x)}{\|\nabla_x \psi_m(x)\|} - \frac{\nabla_x F(x)}{\|\nabla_x F(x)\|}$$

$$x: \psi(x) - \delta \leq 0$$

# Metoda rzutowania - Rozena

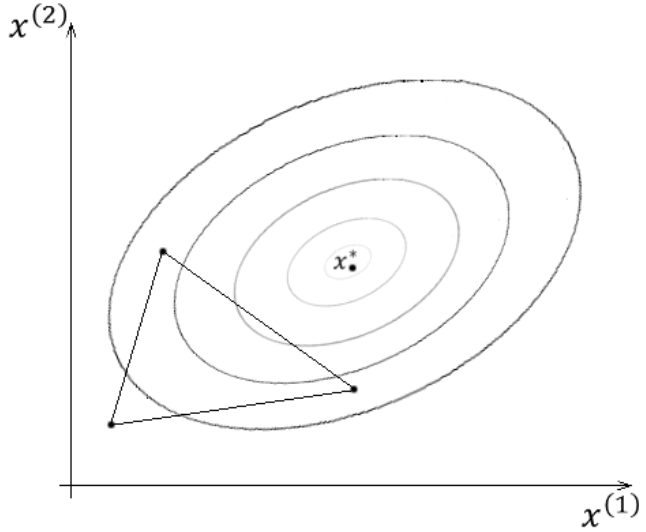


# Metoda Kompleks (Boxa)
# Metoda Neldera-Meada (simplex)

$x_1\ x_2\ \ldots x_K$ - kompleks w przestrzeni $s$-wymiarowej $K \geq S + 1$




$$x_H \rightarrow F(x_H) = \max_{1 \leq s \leq K} F(x_s)$$

$$\bar{x} = \frac{1}{s} \sum_{s=1, s \neq H}^{K} x_s$$

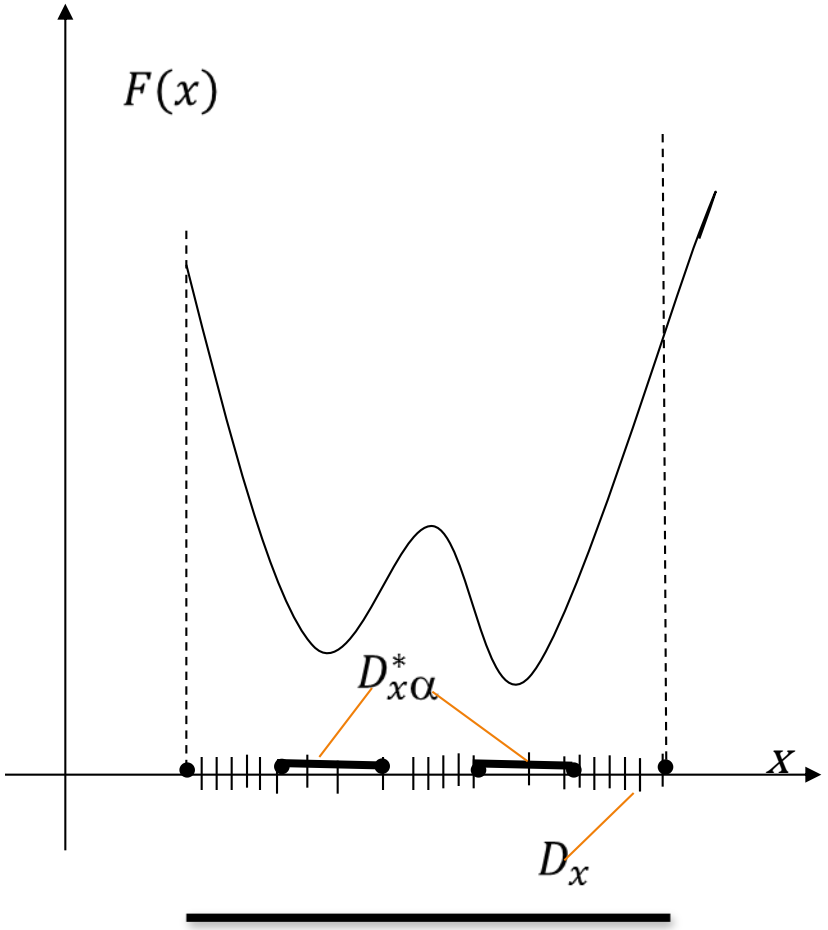
$$x^* = (1 + \alpha)\bar{x}\text{-}x_H$$

$$D_x\colon l_s \leq x^{(s)} \leq u_s \; s = 1, 2, \ldots, S$$
$$l_m \leq x_m(x) \leq u_m \; m = 1, 2, \ldots, M$$

$x_k = l_k + r_k|u_k - l_k|, r_k$ – *liczba losowa* $\in [0, 1], k = 1, 2, \ldots, K$

*Jeśli* $x_k$ *nie należy do* $D_x$ *to przesuń punkt w kierunku centroidu punktów zaakceptowanych*

# Poszukiwania losowe




$$f(x) = \frac{F(x)}{\int_{D_x} F(x)dx}$$

# Zadanie programowania liniowego

$$x^* \to F(x^*) = \min_{x \in \mathcal{D}_x} F(x)$$

$$\mathcal{D}_x = \{x \in R^S, \varphi_l(x) = 0, l = 1, 2, \dots, L, \psi_m(x) \leq 0, m = 1, 2, \dots, M\}$$

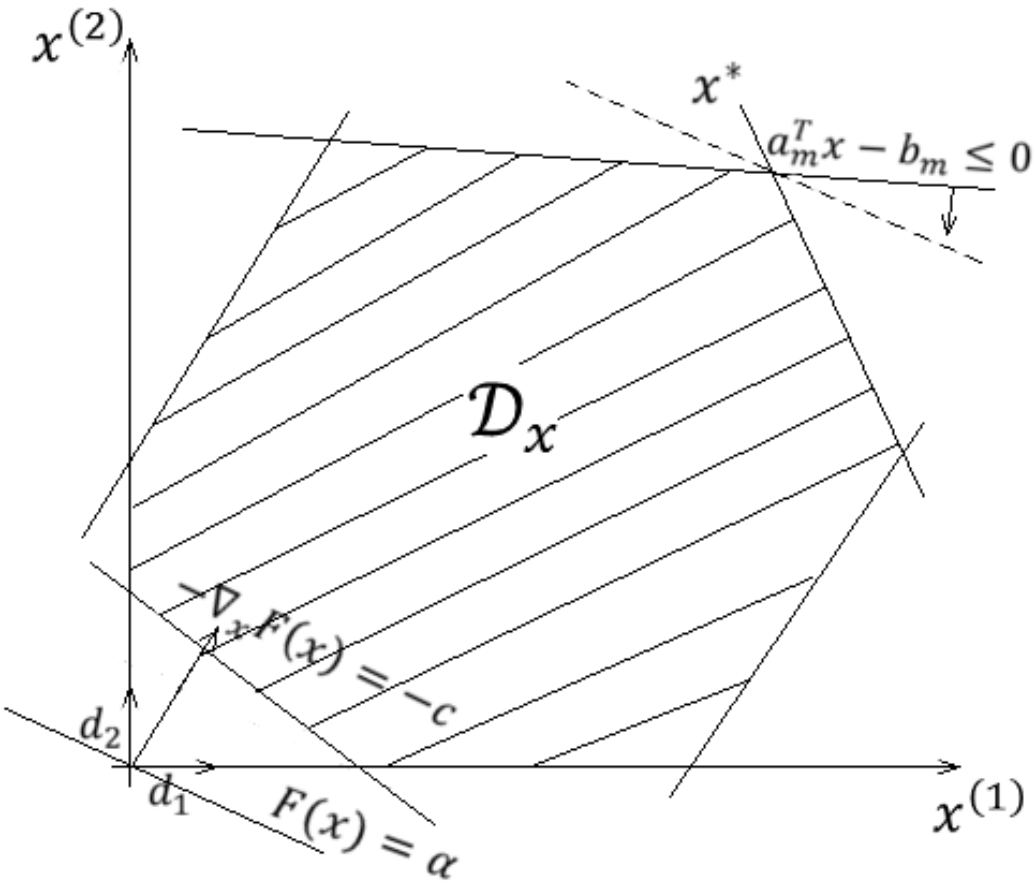
$$F(x) = c^T x = \sum_{s=1}^{S} c_s x^{(s)}$$

$$\varphi_l(x) = a_l^T x - b_l = \sum_{s=1}^{S} a_{ls} x^{(s)} - b_l = 0 \quad l = 1, 2, \dots, L$$

$$\psi_m(x) = a_m^T x - b_m \leq 0 = \sum_{s=1}^{S} a_{ms} x^{(s)} - b_m \leq 0 \quad m = 1, 2, \dots, M$$

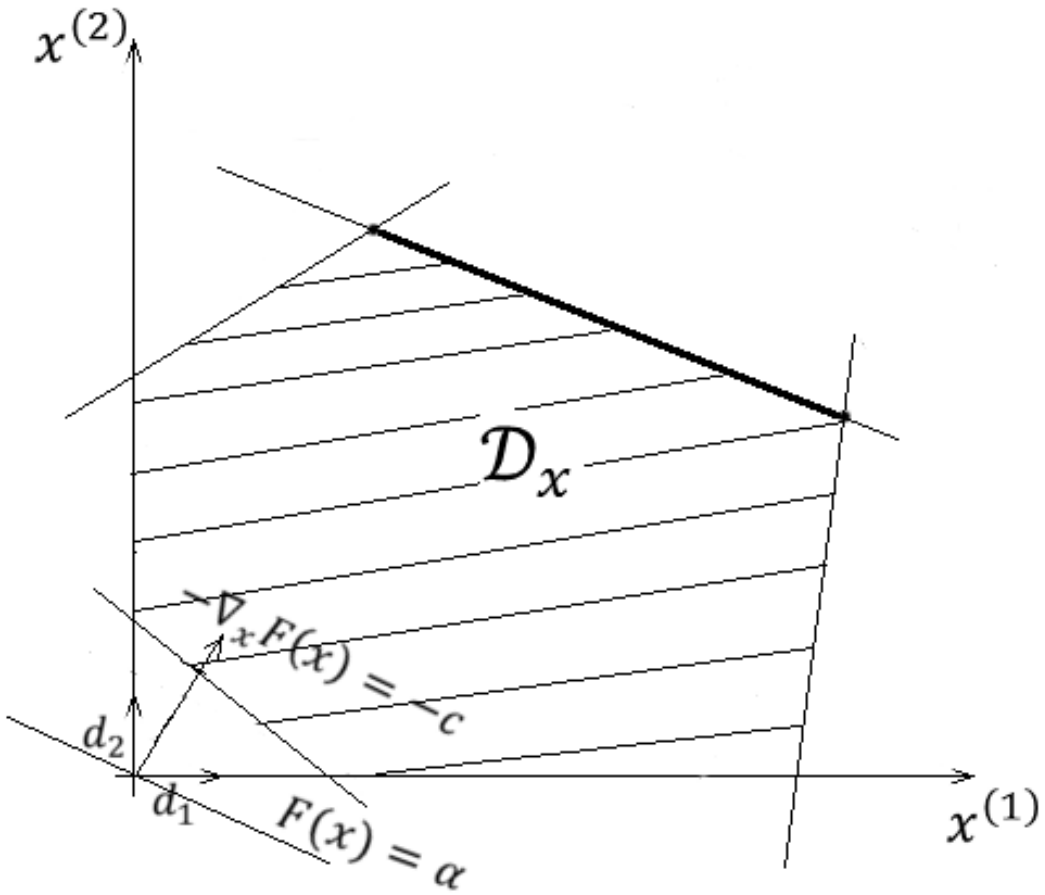
$$x^{(s)} \geq 0 \quad s = 1, 2, \dots, S$$

# Interpretacja graficzna



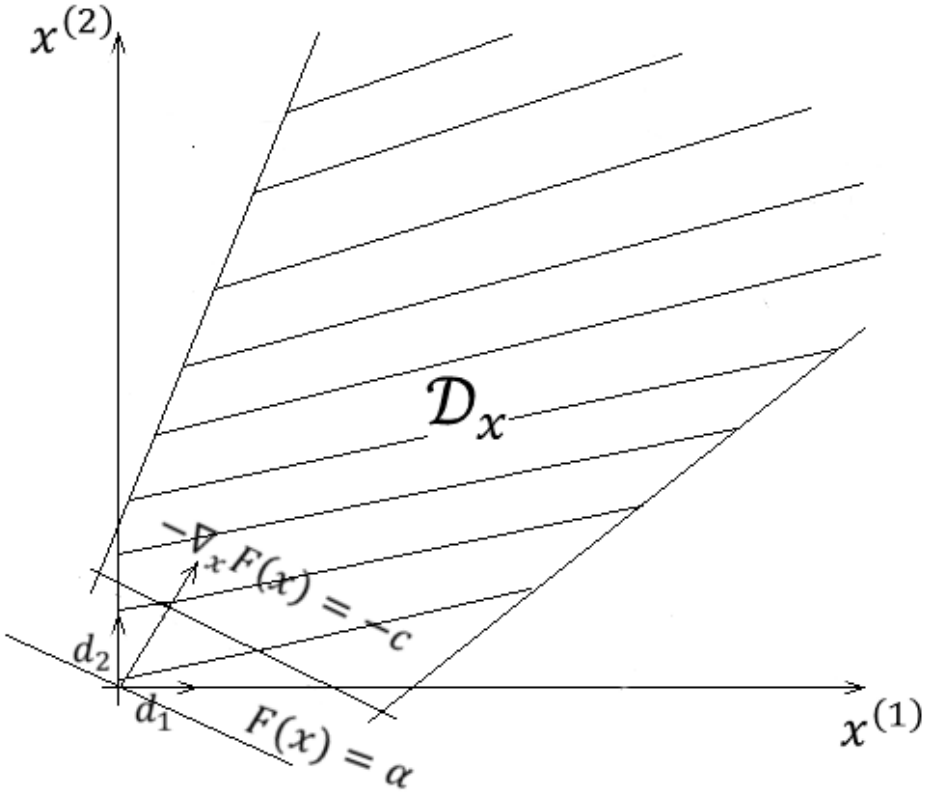

1. Rozwiązanie leży w wierzchołku

# Interpretacja graficzna




2. Rozwiązanie leży na odcinku

# Interpretacja graficzna




3. Rozwiązanie nieograniczone

# Postać kanoniczna

$$F(x) = c^T x$$

A: $\mathcal{D}_X = \{x \in R^S, Ax - b = 0_L, x \geq 0_S\}$

lub

B: $\mathcal{D}_x = \{x \in R^S, Ax - b \leq 0_L, x \geq 0_S\}$

$$c = \begin{bmatrix} c_1 \\ \vdots \\ c_S \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ \vdots \\ b_L \end{bmatrix}, \qquad x = \begin{bmatrix} x^{(1)} \\ \vdots \\ x^{(S)} \end{bmatrix}, \qquad A_{SxL} = \begin{bmatrix} a_{11} & \cdots & a_{1S} \\ \vdots & \ddots & \vdots \\ a_{L1} & \cdots & a_{LS} \end{bmatrix}$$

# Metoda simplex

1. Wyznaczenie rozwiązania początkowego
2. Kryterium zbieżności – zatrzymanie procedury
3. Przechodzenie z jednej bazy do drugiej
4. Postępowanie przy rozwiązaniach zdegenerowanych

# Algorytm simplex

1. Wyznaczamy początkową bazę dopuszczalną
2. Badamy czy $c - c_B B^{-1} A \geq 0_S$. Jeżeli tak to $x_B$ - rozwiązanie problem $x = [x_B \; 0]$
3. Wstawiamy do bazy $k$ – takie, że $c_k - z_k = \min_{1 \leq s \leq S} (c_s - z_s)$
4. Badamy, czy $h_k \leq 0$, jeśli tak – funkcja nieograniczona
5. Wyrzucamy z bazy $l$ – takie, że

$$\frac{h_{l0}}{h_{lk}} = \min_{1 \leq s \leq S} \{\frac{h_{s0}}{h_{sk}}, h_{sk} > 0\}$$

6. $I_B := I_B \setminus \{l\} \cup \{k\}$

   $I_B = \{j \in \{1, 2, \ldots, S\} \quad x^{(j)}$ jest elementem bazy $\}$

7. Jeżeli wskaźnik optymalności zmiennej nie bazowej jest równy 0 to istnieje kolejne rozwiązanie (rozwiązania alternatywne)

# Algorytm simplex

| | | | $c_1$ | ⋯ | $c_k$ | ⋯ | $c_S$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Zmienne bazowe | $c_B$ | $h_0$ | $h_1$ | … | $h_k$ | … | $h_S$ | | $\frac{h_{s0}}{h_{sk}}$ | $h_{sk} \geq 0$ |
| $x_{j1}$ | $c_{j1}$ | $h_{10}$ | $h_{11}$ | ⋯ | $h_{1k}$ | ⋯ | $h_{1s}$ | | | |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | | | |
| ← $x_{jl}$ | $c_{jl}$ | $h_{l0}$ | $h_{l1}$ | ⋯ | $h_{lk}$ | ⋯ | $h_{ls}$ | | | |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | | | |
| $x_{jL}$ | $c_{jL}$ | $h_{L0}$ | $h_{L1}$ | ⋯ | $h_{Lk}$ | ⋯ | $h_{Ls}$ | | | |
| | | | $c_1 - z_1$ | ⋯ | $c_k - z_k$ | ⋯ | $c_s - z_s$ | | | |

↑

$$z_k = \sum_{s \in I_B} c_s h_{sk}$$

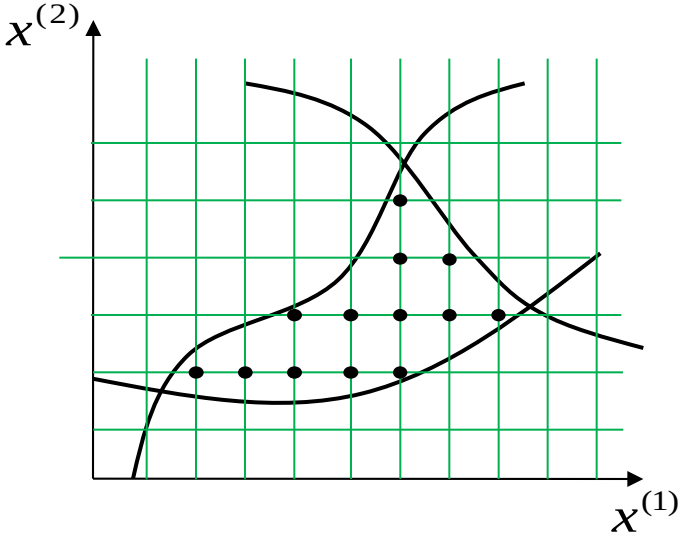
$$h'_{ls} := \frac{h_{ls}}{h_{lk}}; \quad h'_{is} = h_{is} - \frac{h_{ik} h_{ls}}{h_{lk}}$$

$$s = 1, 2, \ldots, S \qquad i = 0, 1, \ldots, S$$
$$s \in I_B \backslash \{l\}$$

# Programowanie całkowitoliczbowe

- Programowanie całkowitoliczbowe

Zmienne decyzyjne: $x \in \overline{\mathscr{D}_x} = \mathscr{D}_x \cap \left\{ x^{(s)} \in \mathscr{C}, s = 1,2,\ldots,S \right\}$




Inne równoważne

$x \in \overline{\mathscr{D}_x} = \{x_1, x_2, \ldots, x_M\}$ - typoszereg

$x \in \overline{\mathscr{D}_x} = \left\{ x^{(s)} \in \{0,1\}, s = 1, 2, \ldots, S \right\}$- programowanie zero – jedynkowe (Boolowskie)

# Programowanie całkowitoliczbowe

Krok 0: $\mathcal{D}_0 = \{\mathcal{D}_{xc} = \mathcal{D}_{01}\},\ n = 0,\ J_0 = 1$

Krok 1: Wyznaczyć taki zbiór $\mathcal{D}^* \in \mathcal{D}_n$

$$\mathcal{F}(\mathcal{D}^*) = \min_{\mathcal{D} \in \mathcal{D}_n} \mathcal{F}(\mathcal{D})$$

Krok 2: Badamy czy $\mathcal{D}^*$ jest zbiorem jednoelementowym ($\{x^*\} = \mathcal{D}^*$) lub $x^* \sim \mathcal{F}(\mathcal{D}^*)$ tj. $\mathcal{F}(\mathcal{D}^*) = F(x^*)$ $x^* \in \mathcal{D}^*$ (oszacowanie jest całkowitoliczbowe) to $x^*$ jest rozwiązaniem optymalnym STOP

Krok 3: $\mathcal{D}^* = \mathcal{D}_{nk}$ dzielimy na $M$ zbiorów rozłącznych

$$\mathcal{D}_{1nk}\mathcal{D}_{2nk} \dots \mathcal{D}_{Mnk} \qquad \mathcal{D}_{nk} = \bigcup_{m=1}^{M} \mathcal{D}_{mnk}$$

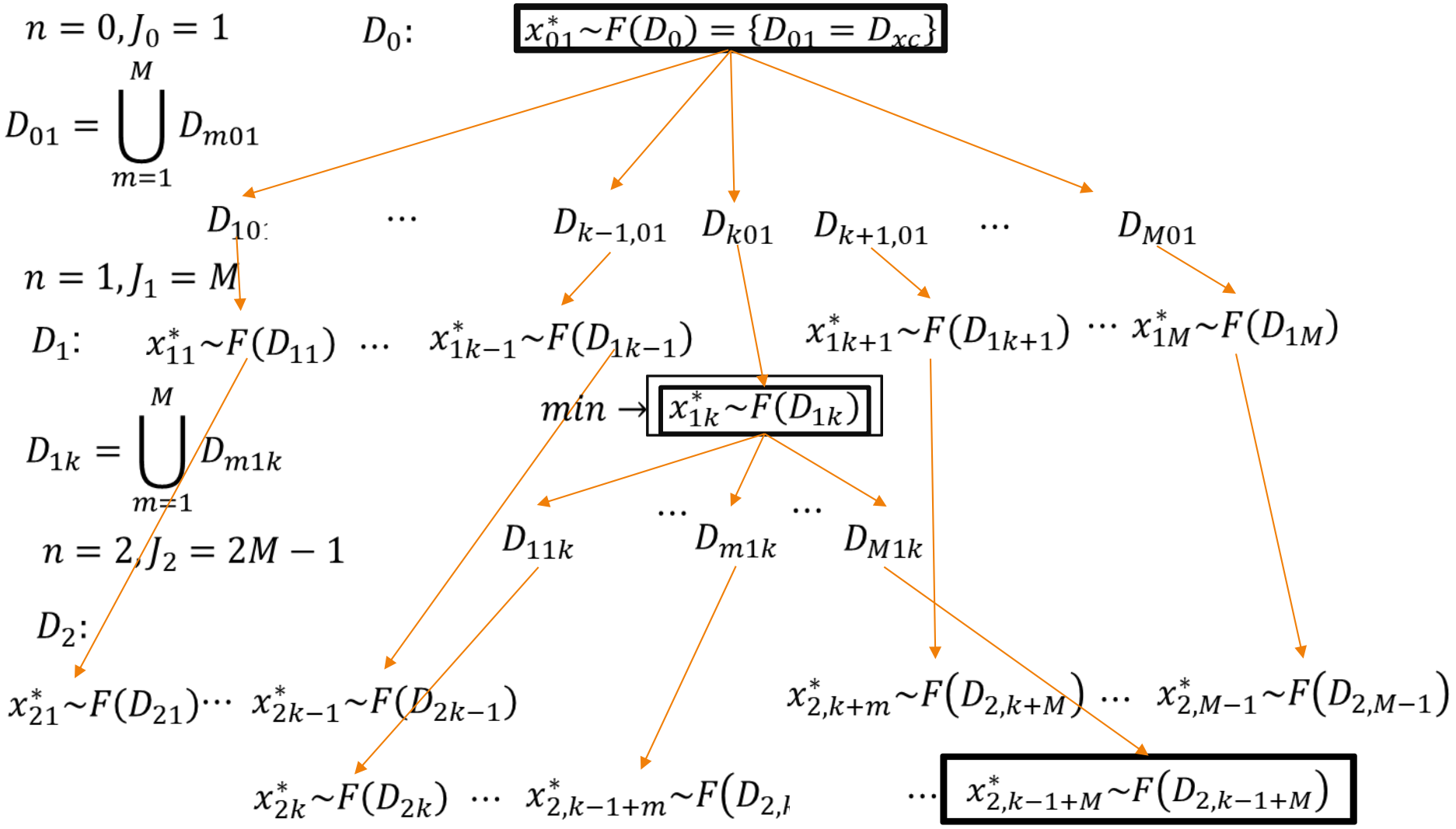
Krok 4: $\mathcal{D}^* = \mathcal{D}_{nk}$

$\mathcal{D}_{n+1} = \mathcal{D}_n \cup \{\mathcal{D}_{1nk}, \mathcal{D}_{2nk}, \dots, \mathcal{D}_{Mnk}\} \backslash \mathcal{D}_{nk}$

$\mathcal{D}_{n+1,j} = \mathcal{D}_{nj} \qquad j = 1, 2, \dots, k-1$

$\mathcal{D}_{n+1,j} = \mathcal{D}_{mnk} \qquad j = k - 1 + m, m = 1, 2, \dots, M$

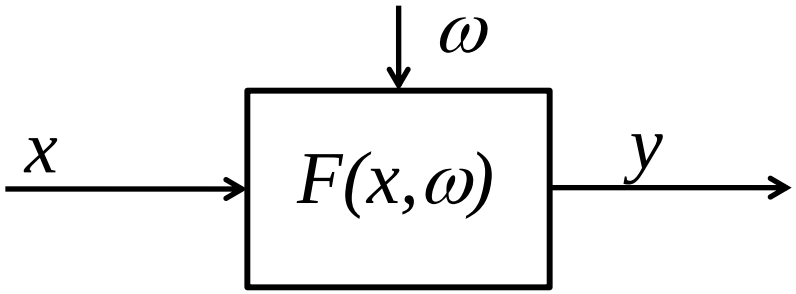
$\mathcal{D}_{n+1,j} = \mathcal{D}_{ni} \qquad j = k + M + i, i = k + 1, \dots, J_n, J_{n+1} = J_n + M - 1$


$n = 0, J_0 = 1$
$D_0$:
$x_{01}^* \sim F(D_0) = \{D_{01} = D_{xc}\}$
$D_{01} = \bigcup_{m=1}^{M} D_{m01}$
$D_{10}$
$\cdots$
$D_{k-1,01}$
$D_{k01}$
$D_{k+1,01}$
$\cdots$
$D_{M01}$
$n = 1, J_1 = M$
$D_1$:
$x_{11}^* \sim F(D_{11})$
$\cdots$
$x_{1k-1}^* \sim F(D_{1k-1})$
$x_{1k+1}^* \sim F(D_{1k+1})$
$\cdots$
$x_{1M}^* \sim F(D_{1M})$
$min \rightarrow$
$x_{1k}^* \sim F(D_{1k})$
$D_{1k} = \bigcup_{m=1}^{M} D_{m1k}$
$n = 2, J_2 = 2M - 1$
$D_{11k}$
$\cdots$
$D_{m1k}$
$\cdots$
$D_{M1k}$
$D_2$:
$x_{21}^* \sim F(D_{21}) \cdots$
$x_{2k-1}^* \sim F(D_{2k-1})$
$x_{2,k+m}^* \sim F(D_{2,k+M})$
$\cdots$
$x_{2,M-1}^* \sim F(D_{2,M-1})$
$x_{2k}^* \sim F(D_{2k})$
$\cdots$
$x_{2,k-1+m}^* \sim F(D_{2,l}$
$\cdots$
$x_{2,k-1+M}^* \sim F(D_{2,k-1+M})$

# Podejmowanie decyzji w warunkach niepewności

- Niepewność opisana zmienną losową
- Gra z naturą
- Podejście growe

# Niepewność opisana zmienną




$$F(x) = \underset{\underline{\omega}}{E}\left[F(x,\underline{\omega})\right]$$

$$\mathscr{D}_x = \underset{\underline{\omega}}{E}\left[\mathscr{D}_x(\underline{\omega})\right] =$$

$$= \left\{ x \in \mathscr{R}^S ;\ \underset{\underline{\omega}}{E}\left[\varphi_l(x,\underline{\omega})\right] = 0,\ l = 1,\ldots,L,\ \underset{\underline{\omega}}{E}\left[\psi_m(x,\underline{\omega})\right] \le 0,\ m = 1,\ldots,M \right\}$$

$$x^* \rightarrow F(x^*) = \min_{x \in D_x} F(x)$$

# Podejmowanie decyzji w warunkach niepewności – gra z naturą

$a_i$ - minimalna korzyść dla *i* – tego wiersza

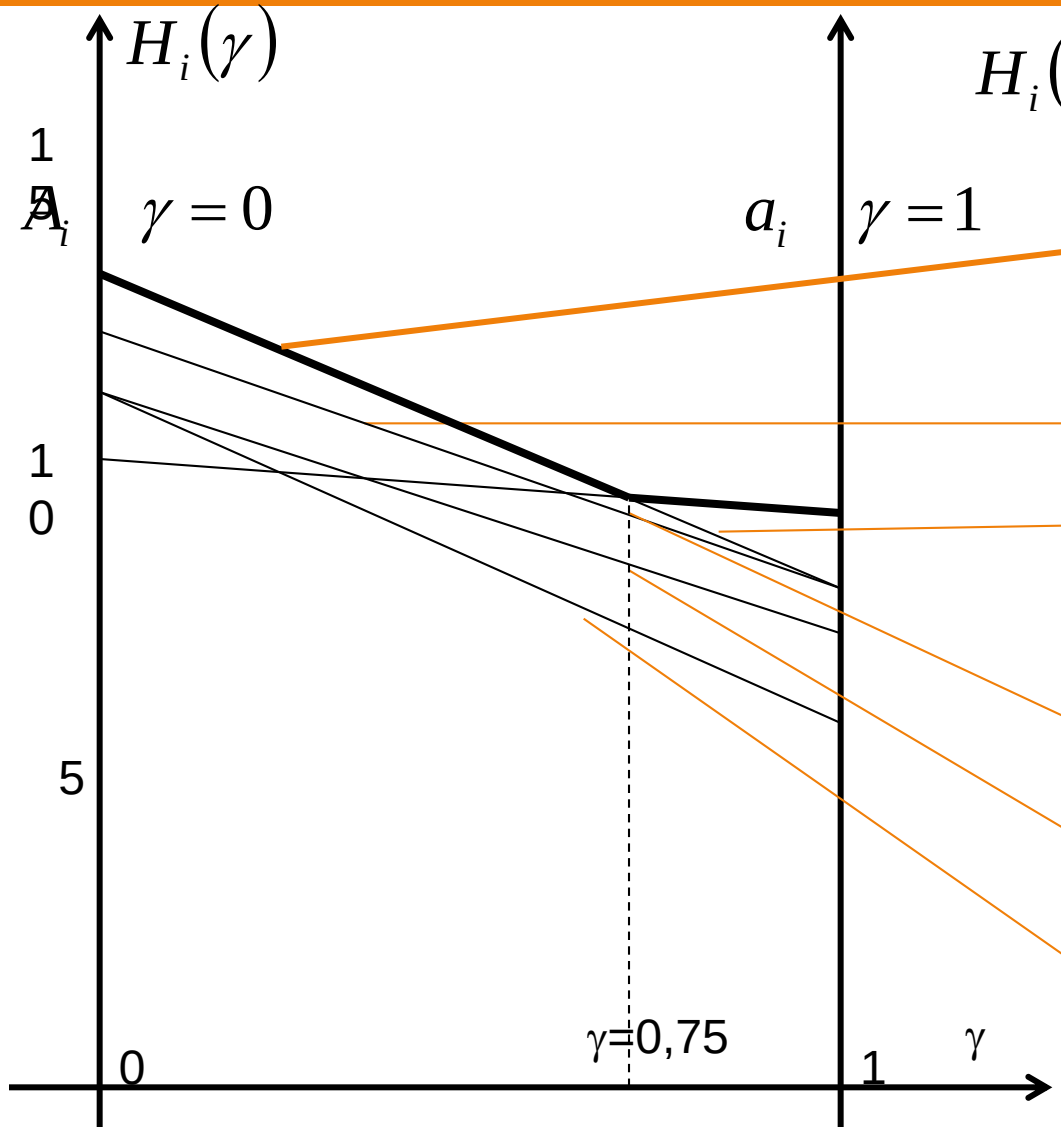
$A_i$ - maksymalna korzyść dla *i* – tego wiersza

$$H_i(\gamma) = a_i\,\gamma + A_i(1-\gamma) \quad \gamma \in [0,1]$$

**Reguła Hurwicza.** Analizując kolejne wiersze macierzy znajdujemy minimalną i maksymalną korzyść, tj..: wartości $a_i$ oraz $A_i$ , a także wartość funkcji $H_i(\gamma)$ dla ustalonego $\gamma$. Wybieramy tę decyzję, dla której wartość funkcji $H_i(\gamma)$ jest największa. W przypadku niejednoznaczności rekomendujemy wszystkie decyzje, dla których powyższy warunek jest spełniony.

| Rodzaj uprawy | Warunki pogodowe | | | *a* min | *A* max | *H(γ)* *γ = 0.5* |
|---|---|---|---|---|---|---|
| | susze | normalne | deszcze | | | |
| 1 | 8 | 10 | 12 | 8 | 12 | 10 |
| 2 | 10 | 11 | 7 | 7 | 11 | 9 |
| 3 | 9 | 13 | 8 | 8 | 13 | 10.5 ← max |
| 4 | 11 | 10 | 6 | 6 | 11 | 8.5 |
| 5 | 10 | 10 | 9 | γ 9 | 10 | 9.5 |

# Podejmowanie decyzji w warunkach niepewności – gra z naturą




$$H_i(\gamma) = a_i\,\gamma + A_i(1-\gamma) \quad \gamma \in [0,1]$$

$$H(\gamma) = \max_{1 \le i \le 5}\{H_i(\gamma)\}$$

$$H_5(\gamma) = 9\gamma + 10(1-\gamma)$$

$$H_3(\gamma) = 8\gamma + 13(1-\gamma)$$

$$H_1(\gamma) = 8\gamma + 12(1-\gamma)$$

$$H_2(\gamma) = 7\gamma + 11(1-\gamma)$$

$$H_4(\gamma) = 6\gamma + 11(1-\gamma)$$

# Gry w podejmowaniu decyzji

Gra dwuosobowa o sumie zerowej

Macierz wypłat dla gracza A:

| A \ B | $B_1$ | $B_2$ | … | $B_m$ | … | $B_M$ |
|---|---|---|---|---|---|---|
| $A_1$ | $a_{11}$ | $a_{12}$ | … | $a_{1m}$ | … | $a_{1M}$ |
| $A_2$ | $a_{21}$ | $a_{22}$ | … | $a_{2m}$ | … | $a_{2M}$ |
| … | … | … | … | … | … | … |
| $A_n$ | $a_{n1}$ | $a_{n2}$ | … | $a_{nm}$ | … | $a_{nM}$ |
| … | … | … | … | … | … | … |
| $A_N$ | $a_{N1}$ | $a_{N2}$ | … | $a_{Nm}$ | … | $a_{NM}$ |

Macierz wypłat dla gracza B:

| A \ B | $B_1$ | $B_2$ | … | $B_m$ | … | $B_M$ |
|---|---|---|---|---|---|---|
| $A_1$ | $-a_{11}$ | $-a_{12}$ | … | $-a_{1m}$ | … | $-a_{1M}$ |
| $A_2$ | $-a_{21}$ | $-a_{22}$ | … | $-a_{2m}$ | … | $-a_{2M}$ |
| … | … | … | … | … | … | … |
| $A_n$ | $-a_{n1}$ | $-a_{n2}$ | … | $-a_{nm}$ | … | $-a_{nM}$ |
| … | … | … | … | … | … | … |
| $A_N$ | $-a_{N1}$ | $-a_{N2}$ | … | $-a_{Nm}$ | … | $-a_{NM}$ |

Gracz A maksymalizuje zyski

Gracz B minimalizuje straty

Zwyczajowo podaje się macierz wypłat dla gracza A

# Gry w podejmowaniu decyzji

- Typowe podejście do rozwiązywania gier
  - Wyznaczenie punktu siodłowego
  - Usunięcie strategii zdominowanych
  - Wyznaczenie strategii mieszanej dla:
    - N=2 i M=2
    - N>2 i M>2

# Gry w podejmowaniu decyzji

Gra dwuosobowa o sumie zerowej

Punkt siodłowy: $\max\limits_{1\le n\le N}\min\limits_{1\le m\le M} a_{nm} = \min\limits_{1\le n\le N}\max\limits_{1\le m\le M} a_{nm}$

| A \ B | | $B_1$ | $B_2$ | … | $B_m$ | … | $B_M$ | min |
|---|---|---|---|---|---|---|---|---|
| | $A_1$ | $a_{11}$ | $a_{12}$ | … | $a_{1m}$ | … | $a_{1M}$ | |
| | $A_2$ | $a_{21}$ | $a_{22}$ | … | $a_{2m}$ | … | $a_{2M}$ | |
| | … | … | … | … | … | … | … | |
| | $A_n$ | $a_{n1}$ | $a_{n2}$ | … | $a_{nm}$ | … | $a_{nM}$ | |
| | … | … | … | … | … | … | … | |
| | $A_N$ | $a_{N1}$ | $a_{N2}$ | … | $a_{Nm}$ | … | $a_{NM}$ | |
| max | | | | | | | | |

$\leftarrow \max\limits_{1\le n\le N}\min\limits_{1\le m\le M} a_{nm}$

$\uparrow \min\limits_{1\le n\le N}\max\limits_{1\le m\le M} a_{nm}$

# Gry w podejmowaniu decyzji

Strategia zdominowana i dominująca

Gracz *A* dysponuje strategiami: $A_1, A_2, \ldots, A_N$

Strategia $A_{n'}$ jest zdominowana przez strategię $A_{n''}$ (dominującą) $\forall\, m = 1, 2, \ldots M \quad a_{n'm} \leq a_{n''m}$ jeżeli

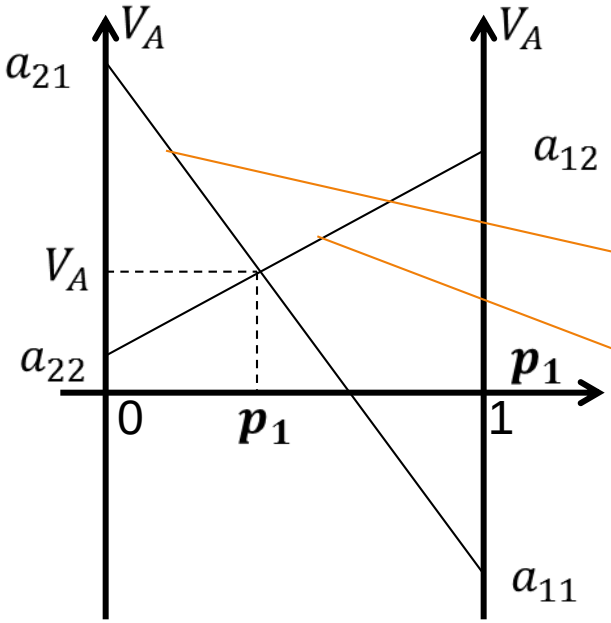
Gracz *B* dysponuje strategiami: $B_1, B_2, \ldots, B_M$

Strategia $B_{m'}$ jest zdominowana przez strategię $B_{m''}$ (dominującą) jeżeli $\forall\, n = 1, 2, \ldots N \quad a_{nm'} \geq a_{nm''}$

# Gry w podejmowaniu decyzji

$N = M = 2$

| A \ B | $q_1$ | $q_2$ |
|---|---|---|
| | $B_1$ | $B_2$ |
| $p_1$ | $A_1$ | $a_{11}$ | $a_{12}$ |
| $p_2$ | $A_2$ | $a_{21}$ | $a_{22}$ |




Równania dla gracza *A*

$$p_1 a_{11} + p_2 a_{21} = V_A \,/\, B_1$$

$$p_1 a_{12} + p_2 a_{22} = V_A \,/\, B_2$$

$$p_1 + p_2 = 1$$

$$p_2 = 1 - p_1$$

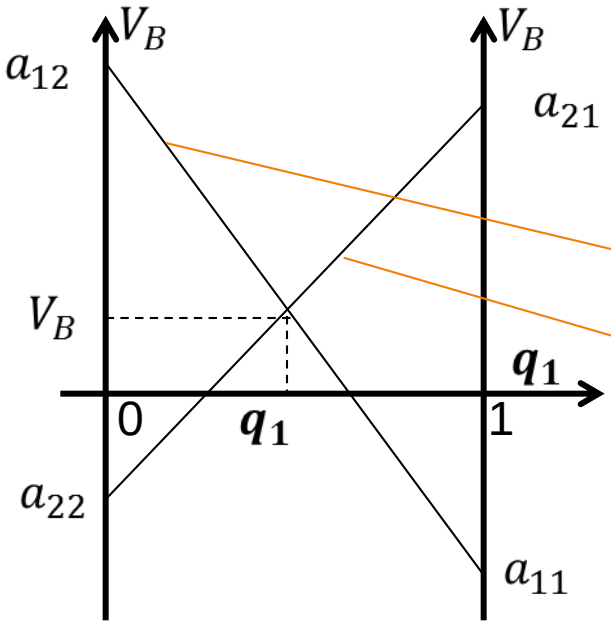
$$p_1 a_{11} + (1 - p_1) a_{21} = V_A \,/\, B_1$$

$$p_1 a_{12} + (1 - p_1) a_{22} = V_A \,/\, B_2$$

# Gry w podejmowaniu decyzji

$N = M = 2$

| A \ B | $q_1$ | $q_2$ |
|---|---|---|
| | $B_1$ | $B_2$ |
| $p_1$ $A_1$ | $a_{11}$ | $a_{12}$ |
| $p_2$ $A_2$ | $a_{21}$ | $a_{22}$ |




Równania dla gracza *B*

$$q_1 a_{11} + q_2 a_{12} = V_B \;/\; A_1$$

$$q_1 a_{21} + q_2 a_{22} = V_B \;/\; A_2$$

$$q_1 + q_2 = 1$$

$$q_2 = 1 - q_1$$

$$q_1 a_{11} + (1 - q_1) a_{12} = V_B \;/\; A_1$$

$$q_1 a_{21} + (1 - q_1) a_{22} = V_B \;/\; A_2$$

# Gry w podejmowaniu decyzji

Gra dwuosobowa o sumie zerowej

Strategie mieszane $N>2$, $M>2$:

| A \ B | | $q_1$ | $q_2$ | | $q_m$ | | $q_M$ |
|---|---|---|---|---|---|---|---|
| | | $B_1$ | $B_2$ | … | $B_m$ | … | $B_M$ |
| $p_1$ | $A_1$ | $a_{11}$ | $a_{12}$ | … | $a_{1m}$ | … | $a_{1M}$ |
| $p_2$ | $A_2$ | $a_{21}$ | $a_{22}$ | … | $a_{2m}$ | … | $a_{2M}$ |
| … | … | … | … | … | … | … | … |
| $p_n$ | $A_n$ | $a_{n1}$ | $a_{n2}$ | … | $a_{nm}$ | … | $a_{nM}$ |
| … | … | … | … | … | … | … | … |
| $p_N$ | $A_N$ | $a_{N1}$ | $a_{N2}$ | … | $a_{Nm}$ | … | $a_{NM}$ |

W tym przypadku rozwiązanie gry sprowadza się do rozwiązania zadania programowania liniowego

# Gry w podejmowaniu decyzji

Zadanie dla gracza A

$$p_1 a_{1m} + p_2 a_{2m} + \cdots + p_N a_{Nm} \geq V_A \;/\; B_m \quad m = 1, 2, \ldots, M$$

$$p_1 + p_2 + \cdots + p_N = 1$$

$$p_n \geq 0 \quad n = 1, 2, \ldots, N$$

$/V_A$

$$\frac{p_1}{V_A} a_{1m} + \frac{p_2}{V_A} a_{2m} + \cdots + \frac{p_N}{V_A} a_{Nm} \geq 1 \;/\; B_m \quad m = 1, 2, \ldots, M$$

$$\frac{p_1}{V_A} + \frac{p_2}{V_A} + \cdots + \frac{p_N}{V_A} = \frac{1}{V_A}$$

$$\frac{p_n}{V_A} \geq 0 \quad n = 1, 2, \ldots, N$$

Niech: $x_n = \dfrac{p_n}{V_A} \quad n = 1, 2, \ldots, N$

# Gry w podejmowaniu decyzji

Zadanie dla gracza A

$$a_{1m}x_1 + a_{2m}x_2 + \cdots + a_{Nm}x_N \geq 1 \,/\, B_m \quad m = 1, 2, \ldots, M$$

$$x_1 + x_2 + \cdots + x_N = \frac{1}{V_A}$$

← Zatem należy minimalizować wyrażenie

$$x_n \geq 0 \quad n = 1, 2, \ldots, N$$

Gracz A maksymalizuje zysk

Ostatecznie zadanie dla gracza A

$$\min_{x_1, x_2, \ldots, x_N} \left(x_1 + x_2 + \cdots + x_N\right) = V_{A\min}$$

Przy ograniczeniach:

$$a_{1m}x_1 + a_{2m}x_2 + \cdots + a_{Nm}x_N \geq 1 \quad m = 1, 2, \ldots, M$$

$$x_n \geq 0 \quad n = 1, 2, \ldots, N$$

$$p_n = x_n V_{A\min}, n = 1, 2, \ldots, N$$

# Gry w podejmowaniu decyzji

Zadanie dla gracza B

$$q_1 a_{n1} + q_2 a_{n2} + \cdots + q_M a_{nM} \le V_B \;/\; A_n \quad n = 1, 2, \ldots, N$$

$$q_1 + q_2 + \cdots + q_M = 1$$

$$q_m \ge 0 \quad m = 1, 2, \ldots, M$$

$/V_B$

$$\frac{q_1}{V_B} a_{n1} + \frac{q_2}{V_B} a_{n2} + \cdots + \frac{q_M}{V_B} a_{nM} \le 1 \;/\; A_n \quad n = 1, 2, \ldots, N$$

$$\frac{q_1}{V_B} + \frac{q_2}{V_B} + \cdots + \frac{q_M}{V_B} = \frac{1}{V_B}$$

$$\frac{q_m}{V_B} \ge 0 \quad m = 1, 2, \ldots, M$$

Niech: $y_m = \dfrac{q_m}{V_B} \quad m = 1, 2, \ldots, M$

# Gry w podejmowaniu decyzji

Zadanie dla gracza B

$$a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nM}y_M \le 1 / A_n \quad n = 1, 2, \ldots, N$$

$$y_1 + y_2 + \cdots + y_M = \frac{1}{V_B}$$

← Zatem należy maksymalizować wyrażenie

$$y_m \ge 0 \quad m = 1, 2, \ldots, M$$

Gracz B minimalizuje straty

Ostatecznie zadanie dla gracza B

$$\max_{y_1, y_2, \ldots, y_M} \left( y_1 + y_2 + \cdots + y_M \right) = V_{B\max}$$

Przy ograniczeniach:

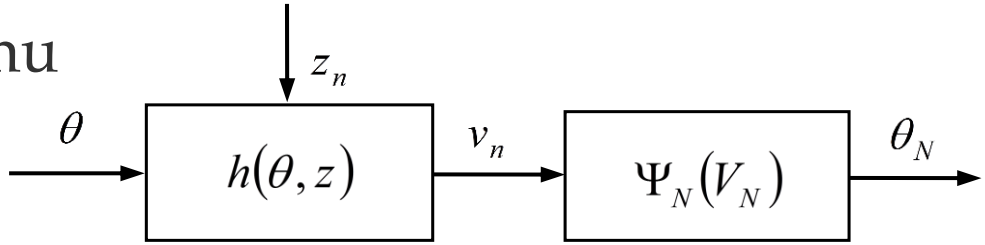
$$a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nM}y_M \le 1 \quad n = 1, 2, \ldots, N$$

$$y_m \ge 0 \quad m = 1, 2, \ldots, M$$

$$q_m = y_n V_{B\max}, m = 1, 2, \ldots, M$$

# Zakłócony pomiar wielkości fizycznej

- Sformułowanie problemu




Opis systemu pomiarowego: $v = h(\theta, z)$

gdzie: $v \in \mathscr{V}$, $h$ – znana funkcja, wzajemnie jednoznaczna względem $z$

$h : \Theta \times \mathscr{Z} \to \mathscr{V}$, $z = h_z^{-1}(\theta, v)$

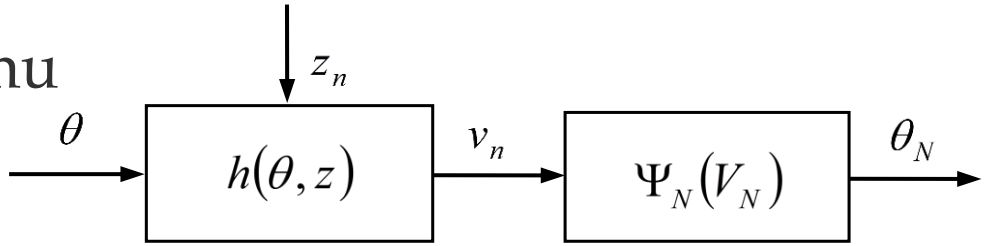
Przykłady funkcji $h$: $v = h(\theta, z) = \theta + z$

$v = h(\theta, z) = \theta \cdot z$

$\mathscr{V}$ – przestrzeń pomiarów( $\dim \theta = \dim z = L$ )

# Zakłócony pomiar wielkości fizycznej

- Sformułowanie problemu




Zakłócenia pomiarowe:

$z_n$ – wartość zmiennej losowej $\underline{z}$ z przestrzeni $\mathcal{Z}$

$f_z(z)$ – funkcja gęstości rozkładu prawdopodobieństwa zmiennej $\underline{z}$

$\theta$ – obserwowany wektor wielkości, wartość zmiennej losowej $\underline{\theta}$, $\theta \in \Theta \subseteq \mathcal{R}^R$

$f_\theta(\theta)$ – funkcja gęstości rozkładu prawdopodobieństwa zmiennej $\underline{\theta}$

Wyniki pomiarów: $V_N = [v_1 \quad v_2 \quad \cdots \quad v_N]$

# Zakłócony pomiar wielkości fizycznej

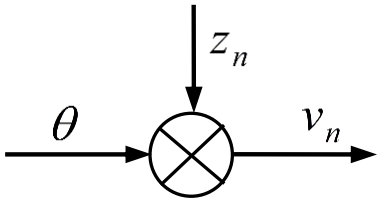
Poszukujemy algorytmu estymacji:

$$\theta_N = \Psi_N(V_N)$$

- Możliwe rozwiązania:
  - Metoda najmniejszych kwadratów
  - Metoda maksymalnej wiarogodności
  - Metody Bayes'a

# Metoda najmniejszych kwadratów

Założenia:




$v = h(\theta, z) = \theta + z$ – zakłócenia addytywne

$\underset{z}{E}[z] = 0$ – wartość oczekiwana zakłóceń wynosi zero

$\underset{z}{Var}[z] < \infty$ – skończona wariancja zakłóceń

Empiryczna wariancja zakłóceń:

$$Var_{zN}(V_N, \theta) = \frac{1}{N}\sum_{n=1}^{N}(v_n - \theta)^2$$

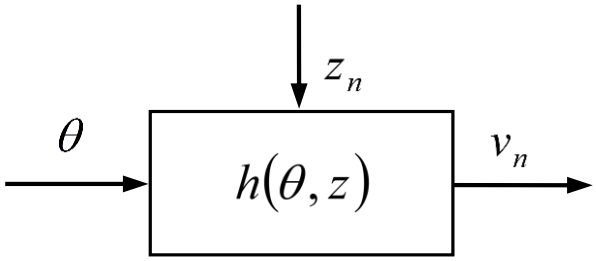
Algorytm estymacji otrzymujemy minimalizując empiryczną wariancje zakłóceń:

$$\theta_N = \Psi_N(V_N) \quad \rightarrow \quad Var_{zN}(V_N, \theta_N) = \min_{\theta \in \Theta} Var_{zN}(V_N, \theta)$$

Algorytm estymacji:

$$\theta_N = \frac{1}{N}\sum_{n=1}^{N} v_n$$

# Metody Bayes'a z obserwacją




Założenia:

$\underline{v} = h(\theta, \underline{z})$ – system pomiarowy opisany jest dowolną funkcją, wzajemnie jednoznaczną względem $z$

$\theta$ – obserwowany wektor wielkości, wartość zmiennej losowej $\underline{\theta}$, $\theta \in \Theta \subseteq \mathscr{R}^R$

Postacie funkcji gęstości rozkładów prawdopodobieństwa $f_z(z)$ oraz $f_\theta(\theta)$ są znane .

Dana jest funkcja strat $L(\theta, \bar{\theta})$, gdzie $\bar{\theta}$ jest wartością estymaty obserwowanej wielkości.

Ryzyko: $$R(\overline{\Psi}) \overset{df}{=} \underset{\underline{\theta}, \underline{V_N}}{E}\left[L\left(\underline{\theta}, \underline{\bar{\theta}} = \overline{\Psi}(\underline{V_N})\right)\right] = \int_{\mathscr{V}_N}\int_{\Theta} L\left(\theta, \overline{\Psi}(V_N)\right) f(\theta, V_N)\, d\theta\, dV_N$$

gdzie: $f(\theta, V_N)$ łączny rozkład prawdopodobieństwa

$$f(\theta, V_N) = f'(\theta | V_N) f''(V_N)$$

gdzie $f'$ jest warunkową, a $f''$ brzegową funkcją gęstości rozkładu prawdopodobieństwa

# Metody Bayes'a

Problem: $\Psi_N \rightarrow R(\Psi_N) = \min_{\overline{\Psi}} R(\overline{\Psi})$

$$R(\overline{\Psi}) = \int_{\mathcal{Y}_N} \int_{\Theta} L(\theta, \overline{\Psi}(V_N)) f'(\theta | V_N) d\theta \; f''(V_N) dV_N$$

$$r(\bar{\theta}, V_N) \overset{\text{df}}{=} \underset{\underline{\theta}}{E}\left[L(\underline{\theta}, \bar{\theta}) | V_N\right] = \int_{\Theta} L(\theta, \bar{\theta} = \overline{\Psi}(V_N)) f'(\theta | V_N) d\theta$$

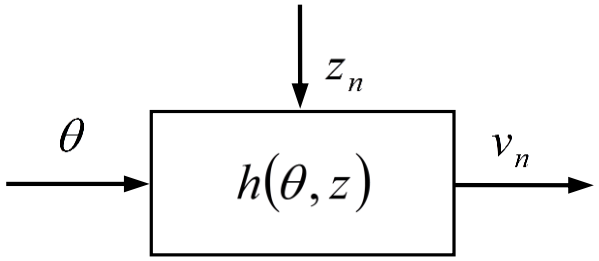
gdzie: $r$ - ryzyko warunkowe

Problem sprowadza się do następującego zadania :

$$\theta_N = \Psi_N(V_N) \rightarrow r(\theta_N, V_N) = \min_{\bar{\theta} \in \Theta} r(\bar{\theta}, V_N)$$

# Metoda maksymalnej wiarogodności

Założenia:




$\underline{v} = h(\theta, \underline{z})$ – system pomiarowy opisany jest dowolną funkcją *h*, wzajemnie jednoznaczną względem *z*

Postać funkcji gęstości rozkładu prawdopodobieństwa zakłóceń $f_z(z)$ jest znana

Funkcja gęstości rozkładu prawdopodobieństwa obserwowanej zmiennej losowej $\underline{v}$

$$f_v(v,\theta) = f_z\left(h_z^{-1}(\theta, v)\right) \cdot |J_h|, \qquad \text{gdzie: } J_h = \frac{\partial h_z^{-1}(\theta, v)}{\partial v}$$

gdzie: $J_h$ - Jacobian, macierz przekształcenia odwrotnego.

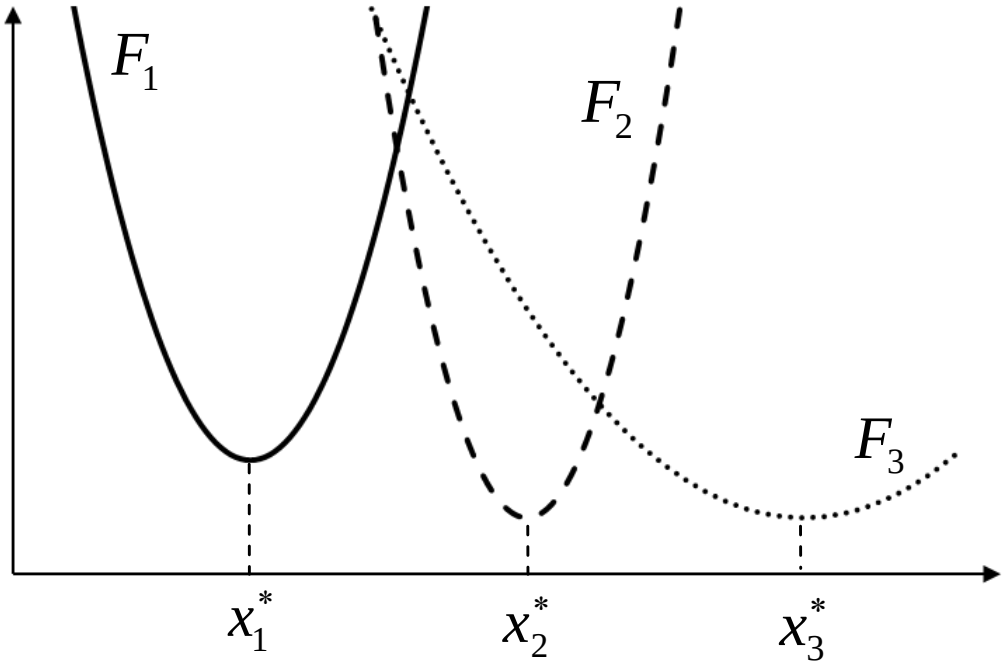
Funkcja wiarogodności ma postać:

$$L_N(V_N, \theta) = \prod_{n=1}^{N} f_v(v_n, \theta) = \prod_{n=1}^{N} f_z\left(h_z^{-1}(\theta, v_n)\right) |J_h|,$$

# Zadanie polioptymalzacji

$x$ – wektor zmiennych decyzyjnych

$F_1(x), F_2(x), \ldots, F_M(x)$ – kilka różnych funkcji celu – ocena wielokryterialna



# Polioptymalizacja

Syntetyczny wskaźnik jakości

$$F(x) = H\left(F_1(x), F_2(x), \ldots, F_K(x)\right)$$

*H(.)* – funkcja monotoniczna ze względu na każdą składową

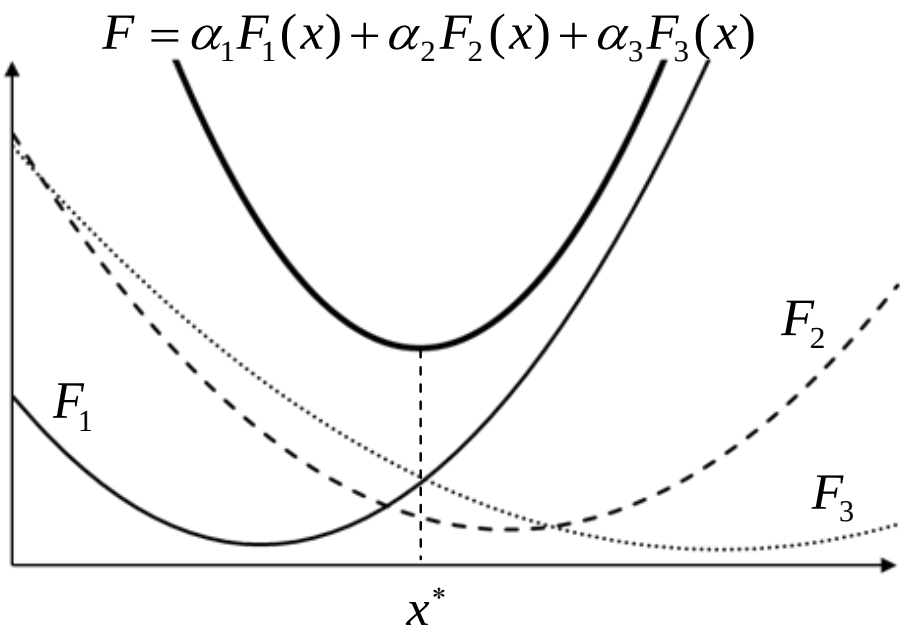
np.: $F(x) = \sum_{k=1}^{K} \alpha_k F_k(x)$

gdzie:

$$\sum_{k=1}^{K} \alpha_k = 1, \quad \alpha_k > 0, \quad k = 1,2,\ldots,K$$

$$F(x) = \prod_{k=1}^{K} F_k(x)$$

$$x^* \rightarrow F(x^*) = \min_{x \in \mathcal{D}_x} F(x)$$



# Polioptymalizacja

Optymalizujemy wybrany wskaźnik,

Pozostałe wskaźniki spełnione są w sposób zadowalający.

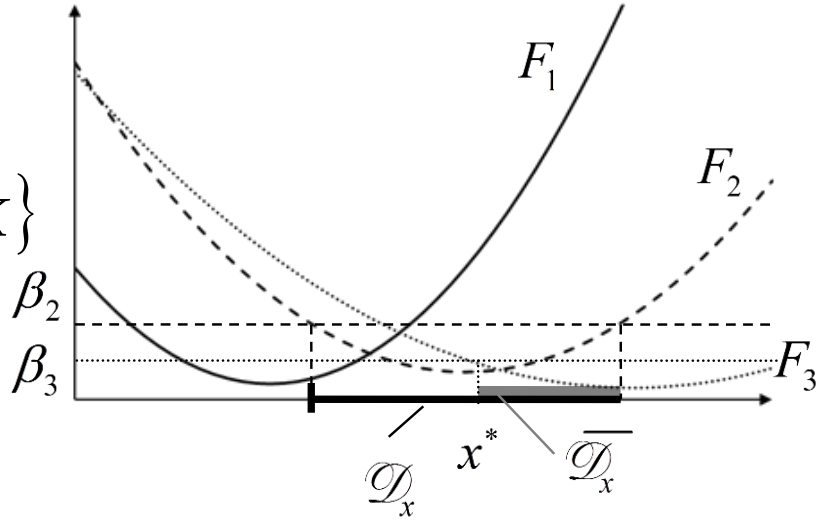
Niech $F_1(x)$ - wybrany wskaźnik

$$F_k(x) \le \beta_k, \quad k = 2,3,\ldots,K$$

Wskaźniki spełnione w sposób zadowalający

$$\overline{\mathscr{D}_x} = \mathscr{D}_x \cap \left\{ x \in \mathscr{R}^S : F_k(x) \le \beta_k,\ k = 2,\ldots,K \right\}$$

$$x^* \to F_1(x^*) = \min_{x \in \overline{\mathscr{D}_x}} F_1(x)$$



# Polioptymalizacja

Uszeregowane wskaźniki jakości – priorytety wśród wskaźników

$$F_1(x) \succ F_2(x) \succ \ldots \succ F_K(x) \quad x \in \mathscr{D}_x$$

Krok 1. $\mathscr{D}_{x1} = \mathscr{D}_x$

$$x_1^* \rightarrow F_1(x_1^*) = \min_{x \in \mathscr{D}_{x1}} F_1(x)$$

Krok 2. $\mathscr{D}_{x2} = \mathscr{D}_{x1} \cap \left\{ x \in \mathscr{R}^S : F_1(x) \leq F_1\left(x_1^*\right) + \gamma_1 \right\}$

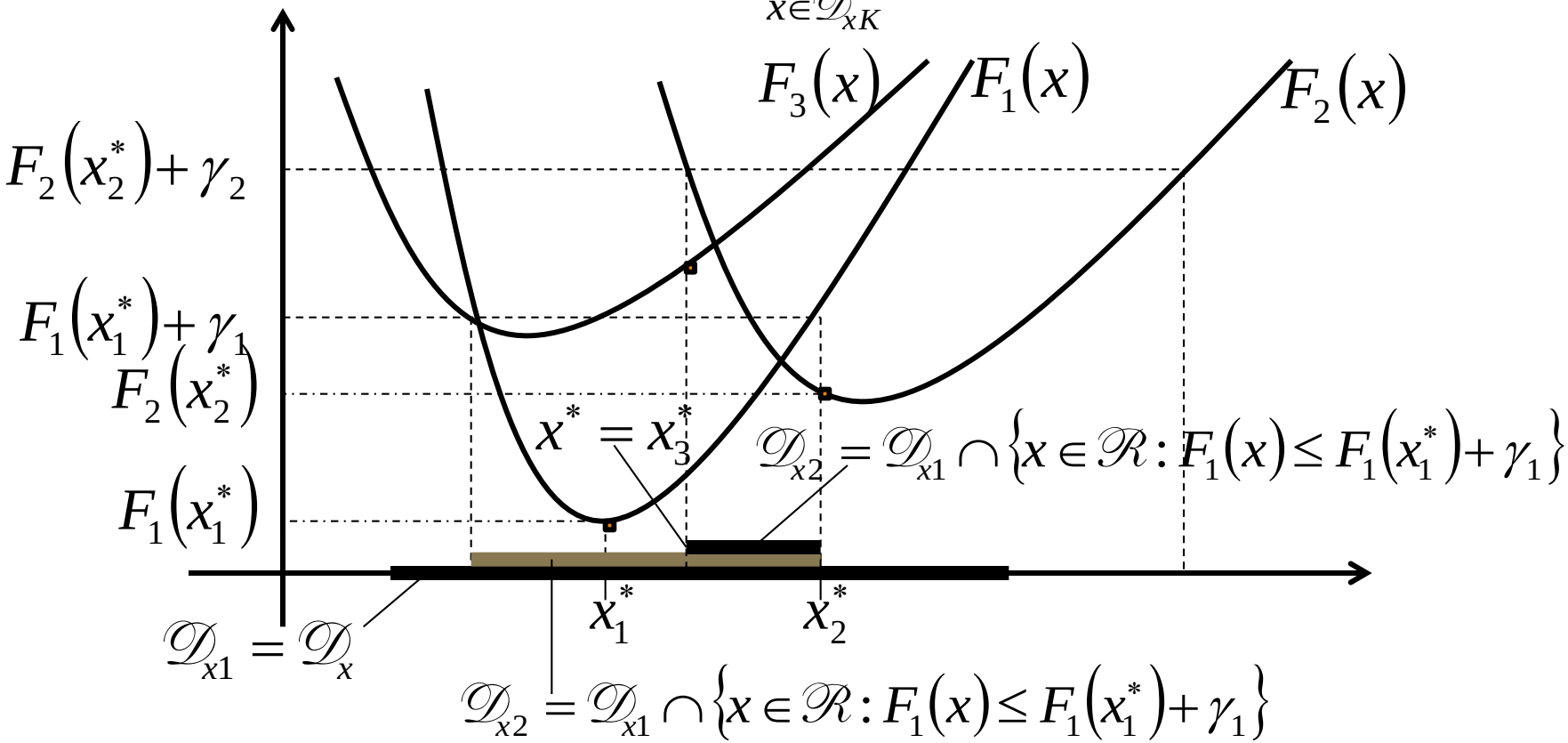
$$x_2^* \rightarrow F_2(x_2^*) = \min_{x \in \mathscr{D}_{x2}} F_2(x)$$

$\vdots$

# Polioptymalizacja

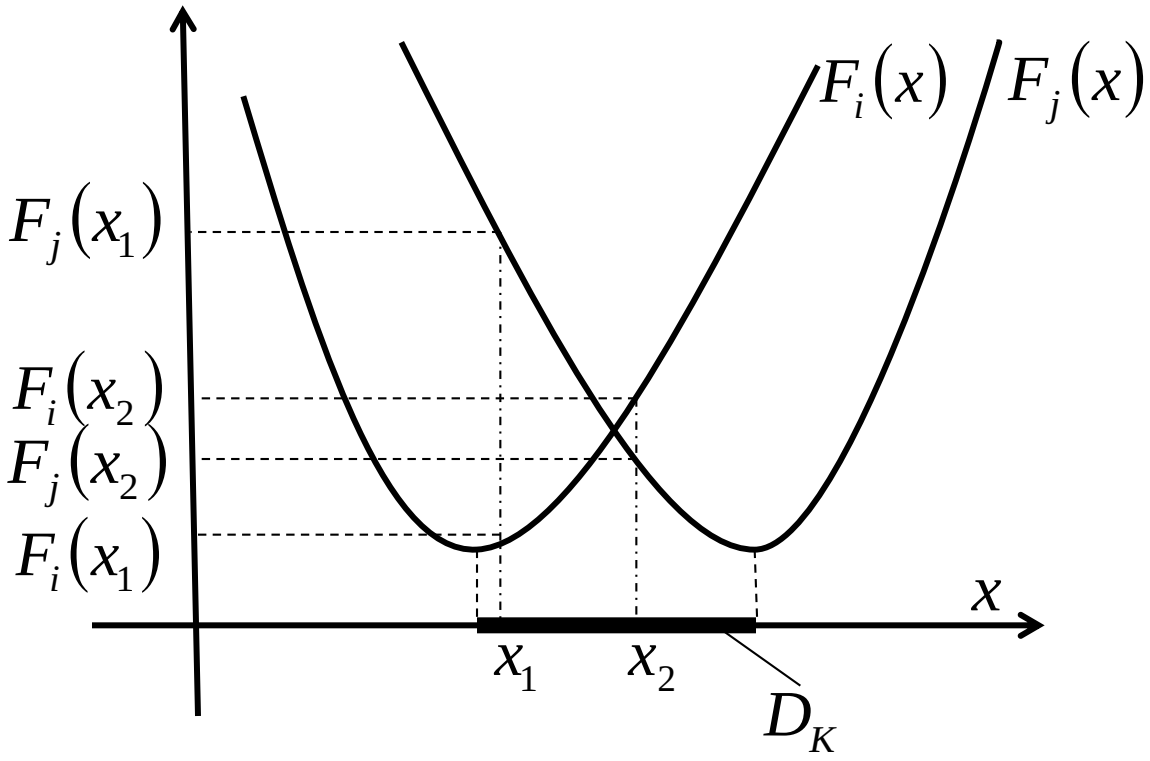
Krok K. $\mathscr{D}_{xK} = \mathscr{D}_{xK-1} \cap \left\{ x \in \mathscr{R}^S : F_{K-1}(x) \le F_1\left(x^*_{K-1}\right) + \gamma_{K-1} \right\}$

$$x^* = x^*_K \rightarrow F_K(x^*_K) = \min_{x \in \mathscr{D}_{xK}} F_K(x)$$



# Zbiór punktów kompromisowych

$$x_1, x_2 \in D_K \Leftrightarrow \forall j \in \{1, 2, \ldots, K\} \exists\, i \in \{1, 2, \ldots, K\}$$
$$F_j(x_1) > F_j(x_2) \Rightarrow F_i(x_1) < F_i(x_2)$$



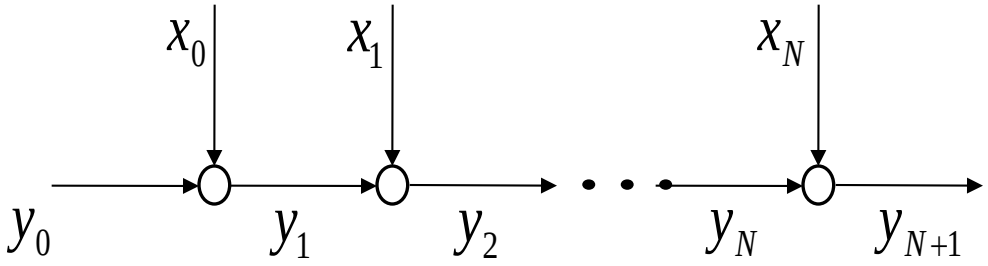
# Programowanie dynamiczne

Proces dynamiczny: $y_{n+1} = P(y_n, x_n)$

$n$ – takt

$x_n$ – decyzja w $n$ – tym takcie

$y_n$ – stan procesu w $n$ – tym takcie







http://www.all-freeware.com/

Zadanie decyzyjne: znaleźć ciąg: $x_0^*, x_1^*, \ldots, x_N^*,$

dla których wskaźnik $Q(x_0, x_1, \ldots, x_N)$ przyjmuje wartość minimalną

# Optymalizacja wieloetapowa

$$x_1^*, x_2^*, \ldots, x_S^* \rightarrow F(x_1^*, x_2^*, \ldots, x_S^*) = \min_{x_1, x_2, \ldots, x_S \in \mathscr{D}_x} F(x_1, x_2, \ldots, x_S)$$

$$\mathscr{D}_x = (x_1, x_2, \cdots, x_S)$$

$$\left\{ [x_1\ x_2 \cdots x_S]^T \in \mathscr{R}^S : \varphi_l(x_1, x_2, \cdots, x_S) = 0, l = 1, 2, \ldots, L, \right.$$
$$\left. \psi_m(x_1, x_2, \cdots, x_S) \leq 0, m = 1, 2, \ldots, M \right\}$$

Powyższe zadanie można rozwiązać krok po kroku optymalizując po jednej zmiennej i uzależniając od pozostałych.

Przyjmijmy oznaczenia:

$$F \equiv F_S, \quad \varphi_l \equiv \varphi_{lS}, l = 1, 2, \ldots, L, \quad \psi_m = \psi_{mS}, m = 1, 2, \ldots, M \quad \mathscr{D}_x \equiv \mathscr{D}_{x_S}$$

# Optymalizacja wieloetapowa

Krok 1.

$$x_S^* = G_S(x_1, \ldots, x_{S-1}) \rightarrow F_S(x_1, x_2, \ldots, x_S^*) = \min_{x_S \in \mathscr{D}_{xS}} F_S(x_1, x_2, \ldots, x_S)$$

Wartość funkcji celu w punkcie optymalnym:

$$F_{S-1}(x_1, x_2, \ldots, x_{S-1}) \stackrel{\Delta}{=} F_S(x_1, x_2, \ldots, x_S^*) = F_S(x_1, x_2, \ldots, G_S(x_1, \ldots, x_{S-1}))$$

Ograniczenia w punkcie optymalnym:

$$\mathscr{D}_{xS-1}(x_1, \ldots, x_{S-1}) \stackrel{\Delta}{=} \mathscr{D}_{xS}(x_1, \ldots, x_{S-1}, x_S^* = G_S(x_1, \ldots, x_{S-1})) =$$

$$\left\{ \begin{array}{l} [x_1 \; x_2 \cdots x_{S-1}]^T \in \mathscr{R}^{S-1} : \\ \varphi_{lS}(x_1, x_2, \cdots, G_S(x_1, \ldots, x_{S-1})) = \varphi_{lS-1}(x_1, x_2, \cdots, x_{S-1}) = 0, l = 1, 2, \ldots, L, \\ \psi_{mS}(x_1, x_2, \cdots, G_S(x_1, \ldots, x_{S-1})) = \psi_{mS-1}(x_1, x_2, \cdots, x_{S-1}) \leq 0, m = 1, 2, \ldots, M \end{array} \right\}$$

# Optymalizacja wieloetapowa

⋮

Krok S-1.

$$x_2^* = G_2(x_1) \rightarrow F_2(x_1, x_2^*) = \min_{x_2 \in \mathscr{D}_{x2}} F_2(x_1, x_2)$$

Wartość funkcji celu w punkcie optymalnym:

$$F_1(x_1) \stackrel{\Delta}{=} F_2(x_1 x_2^*) = F_2(x_1, G_2(x_1))$$

Ograniczenia w punkcie optymalnym:

$$\mathscr{D}_{x1}(x_1) \stackrel{\Delta}{=} \mathscr{D}_{xS-1}(x_1, x_2^* = G_1(x_1)) = \left\{ \begin{array}{l} x_1 \in \mathscr{R}: \\ \varphi_{l2}(x_1, G_1(x_1)) = \varphi_{l1}(x_1) = 0, l = 1, 2, \ldots, L, \\ \psi_{m2}(x_1, G_{S-1}(x_1)) = \psi_{m1}(x_1) \leq 0, m = 1, 2, \ldots, M \end{array} \right\}$$

# Optymalizacja wieloetapowa

Krok S.

$$x_1^* \rightarrow F_1\left(x_1^*\right) = \min_{x_1 \in \mathcal{D}_{x1}} F_1\left(x_1\right)$$

Teraz możemy powrócić do zależności „*G*” wyznaczonych w poprzednich krokach

$$x_1^*$$

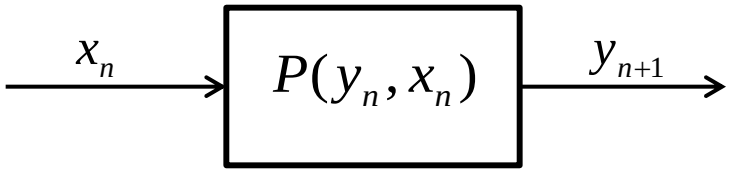
$$x_2^* = G_2\left(x_1^*\right)$$

$$\vdots$$

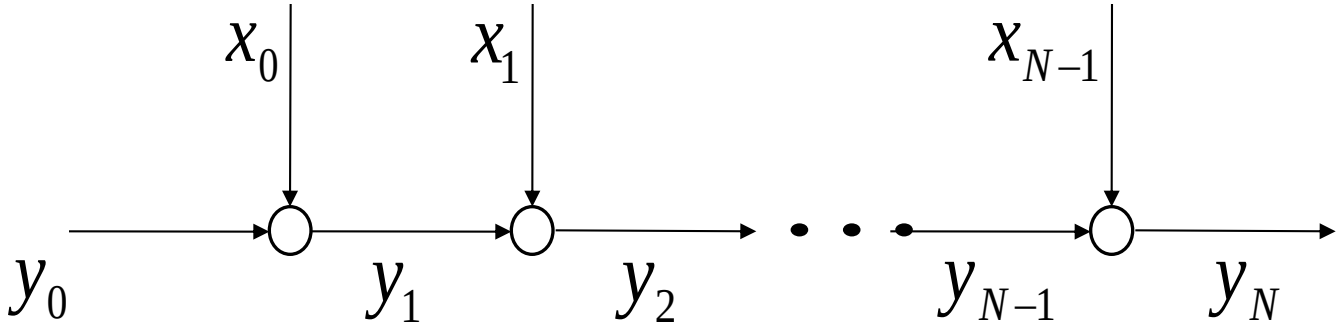
$$x_{S-1}^* = G_{S-1}\left(x_1^*, x_2^* \ldots, x_{S-1}^*\right)$$

$$x_S^* = G_S\left(x_1^*, x^*, \ldots, x_{S-1}^*\right)$$

# Programowanie dynamiczne




$$y_{n+1} = P(y_n, x_n),\ y_o$$




$$Q(x_0, x_1, \ldots, x_{N-1}, y_1, y_2, \ldots, y_N) = \sum_{n=0}^{N-1} A_{n+1}(x_n, y_{n+1}) \stackrel{\Delta}{=} F(y_0, x_0, x_1, \ldots, x_{N-1})$$

# Programowanie dynamiczne

$$Q(x_0, x_1, \ldots, x_{N-1}, y_1, y_2, \ldots, y_N) = \sum_{n=0}^{N-1} A_{n+1}(x_n, y_{n+1})$$

Krok 1. $$x_{N-1}^* \to \min_{x_{N-1}} A_N(x_{N-1}, y_N) = \min_{x_{N-1}} A_N(x_{N-1}, P(y_{N-1}, x_{N-1}))$$

Wiemy że: $$y_N = P(y_{N-1}, x_{N-1})$$

$$x_{N-1}^* = G_{N-1}(y_{N-1}) \to \min_{x_{N-1}} A_N(x_{N-1}, P(y_{N-1}, x_{N-1}))$$

$$V_{N-1}(y_{N-1}) \stackrel{\Delta}{=} \min_{x_{N-1}} A_N(x_{N-1}, P(y_{N-1}, x_{N-1})) =$$

$$= A_N(x_{N-1}^*, P(y_{N-1}, x_{N-1}^*)) = A_N(G_{N-1}(y_{N-1}), P(y_{N-1}, G_{N-1}(y_{N-1})))$$

# Programowanie dynamiczne

Krok N.

$$x_0^* \rightarrow \min_{x_0}\left\{A_1(x_0, y_1) + V_1(y_1)\right\}$$

Wiemy że: $y_1 = P(y_0, x_0)$

$$x_0^* = G_0(y_0) \rightarrow \min_{x_0}\left\{A_1(x_0, P(y_0, x_0)) + V_1(P(y_0, x_0))\right\}$$

$y_0$ jest znaną wartością i teraz możemy wyliczyć kolejne wartości decyzji $x_0^*, x_1^*, \ldots, x_{N-1}^*$,

$$x_0^* = G_0(y_0) \rightarrow y_1 = P(y_0, x_0^*)$$

$$x_1^* = G_1(y_1) \rightarrow y_2 = P(y_2, x_2^*)$$

$$\vdots$$

$$x_{N-2}^* = G_{N-2}(y_{N-2}) \rightarrow y_{N-1} = P(y_{N-2}, x_{N-2}^*)$$

$$x_{N-1}^* = G_{N-1}(y_{N-1}) \rightarrow y_N = P(y_{N-1}, x_{N-1}^*)$$

# Dekompozycja i koordynacja zadania optymalizacji

$$x^* \to F(x^*) = \min_{x \in \mathscr{D}_x} F(x) \quad F(x) = H\left(F_I(x^I, w), F_{II}(x^{II}, w), \ldots, F_K(x^K, w)\right)$$

$$x \in \mathscr{D}_x \subset \mathscr{R}^S, \mathscr{D}_x = \left\{x \in \mathscr{R}^S; \varphi(x) = 0_L, \psi(x) \le 0_M\right\}$$

$$\mathscr{D}_x = \left\{ x = \begin{bmatrix} x^I \\ x^{II} \\ \vdots \\ x^K \\ w \end{bmatrix} \in \mathscr{R}^S, \varphi(x) = \begin{bmatrix} \varphi^I(x^I, w) \\ \varphi^{II}(x^{II}, w) \\ \vdots \\ \varphi^K(x^K, w) \\ \varphi^w(w) \end{bmatrix} = \begin{bmatrix} 0_{L_I} \\ 0_{L_{II}} \\ \vdots \\ 0_{L_K} \\ 0_{L_w} \end{bmatrix} = 0_L, \psi(x) = \begin{bmatrix} \psi^I(x^I, w) \\ \psi^{II}(x^{II}, w) \\ \vdots \\ \psi^K(x^K, w) \\ \psi^w(w) \end{bmatrix} \le \begin{bmatrix} 0_{M_I} \\ 0_{M_{II}} \\ \vdots \\ 0_{M_K} \\ 0_{M_w} \end{bmatrix} = 0_M \right\}$$

$$\sum_{k=I}^{K} L_k + L_w = L, \sum_{k=I}^{K} M_k + M_w = M$$

# Dekompozycja i koordynacja zadania optymalizacji

Zadanie rozwiązujemy w dwóch etapach.
Etap I: Dla ustalonej zmiennej koordynacyjnej $w$ rozwiązujemy $K$ zadań

$$x^{k*}(w) \to F_k\left(x^{k*}(w), w\right) = \min_{x^k \in \mathcal{D}_{x^k}(w)} F_k\left(x^k, w\right),\ k = 1, 2, \ldots, K$$

Następnie wyznaczamy wartość:

$$\overline{F}(w) \stackrel{\Delta}{=} H\left(F_I\left(x^{I*}(w), w\right), F_{II}\left(x^{II*}(w), w\right), \ldots, F_K\left(x^{K*}(w), w\right)\right)$$

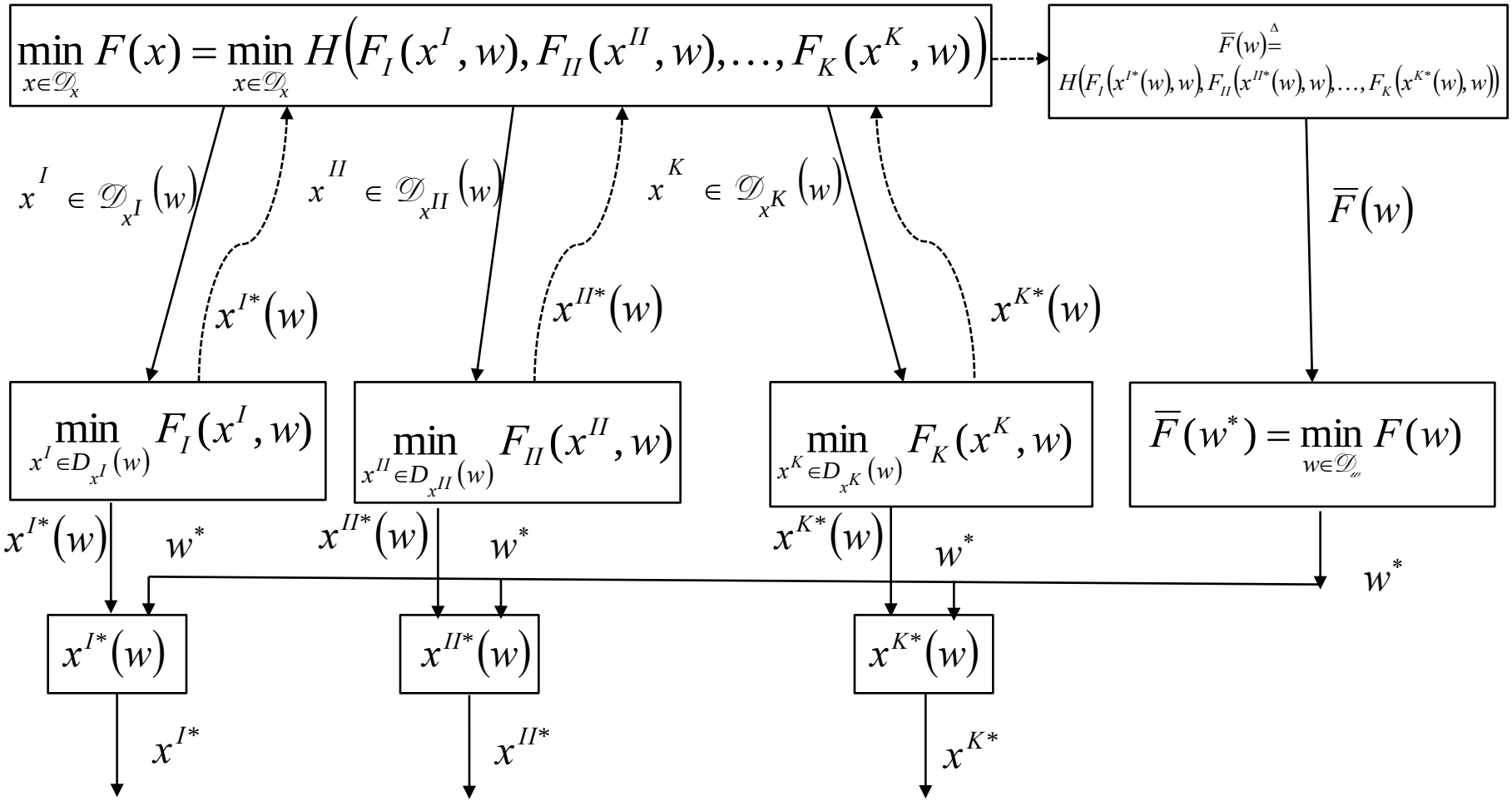
Etap II: wyznaczamy optymalną wartość zmiennej koordynacyjnej rozwiązując zadanie:

$$w^* \to \overline{F}(w^*) = \min_{w \in \mathcal{D}_w} F(w)$$

ostatecznie

$$x^{k*} = x^{k*}\left(w^*\right),\ k = 1, 2, \ldots, K$$

# Dekompozycja i koordynacja zadania optymalizacji



# Optymalizacja systemów

prof. dr hab. inż. Jerzy Świątek

# Egzamin

- Optymalizacja systemów
- Termin 0 – ostatni wykład:
- 3.02.2027 (środa) sala 23 budynek C-3, godz. 13:15 -15:00
- Warunki terminu 0:
  - Pozytywne aktualne (tegoroczne) zaliczenie z ćwiczeń ≥ 3.5
  - Propozycja oceny: ocena z ćwiczeń (+premia za obecność)
  - Obecność na ostatnim wykładzie
  - Nieobecni (bez względu na przyczynę) rezygnują ze zwolnienia
  - Można wcześniej ale nie po terminie 0. Należy przedstawić potwierdzoną ocenę.
- Termin 1:
- 10.02.2027 (środa) sala 23, budynek C-3, godz. 13:15 -15:00

- Termin 2:
- 17.02.2027 (środa) sala 23, budynek C-3, godz. 13:15 -15:00

# Literatura

- **LITERATURA PODSTAWOWA:**
- Findeisen W., Szymanowski J., Wierzbicki A., *Teoria i metody obliczeniowe optymalizacji*, PWN, Warszawa, 1980
- Seidler J., Badach A., Molisz W., *Metody rozwiązywania zadań optymalizacji*, WNT, Warszawa, 1980
- Kusiak J., Danielewska-Tułecka A., Oprocha P. *Optymalizacja. Wybrane metody z przykładami zastosowań*, PWN, Warszawa, 2009.
- Edwin Chong, Stanisław Żak, *An Introduction to Optimization*, John Wiley & Sons, 2013
- https://byes.pl/wyklady-prof-swiatka/

- 
- **LITERATURA UZUPEŁNIAJĄCA:**
- Bazara M.S., Shetty C.M., *Nonlinear Programming, Theory and Algorithms*, John Wiley and Sons, New York 1979
- Brdyś M., Ruszczyński A., *Metody optymalizacji w zadaniach*, WNT, Warszawa 1985
- De Groot M.H., *Optymalne decyzje statystyczne*, PWN, Warszawa 1981
- Zieliński R., Neuman P., *Stochastyczne metody poszukiwania minimum funkcji*, WNT, Warszawa 1985

# Dziękuję za uwagę

