

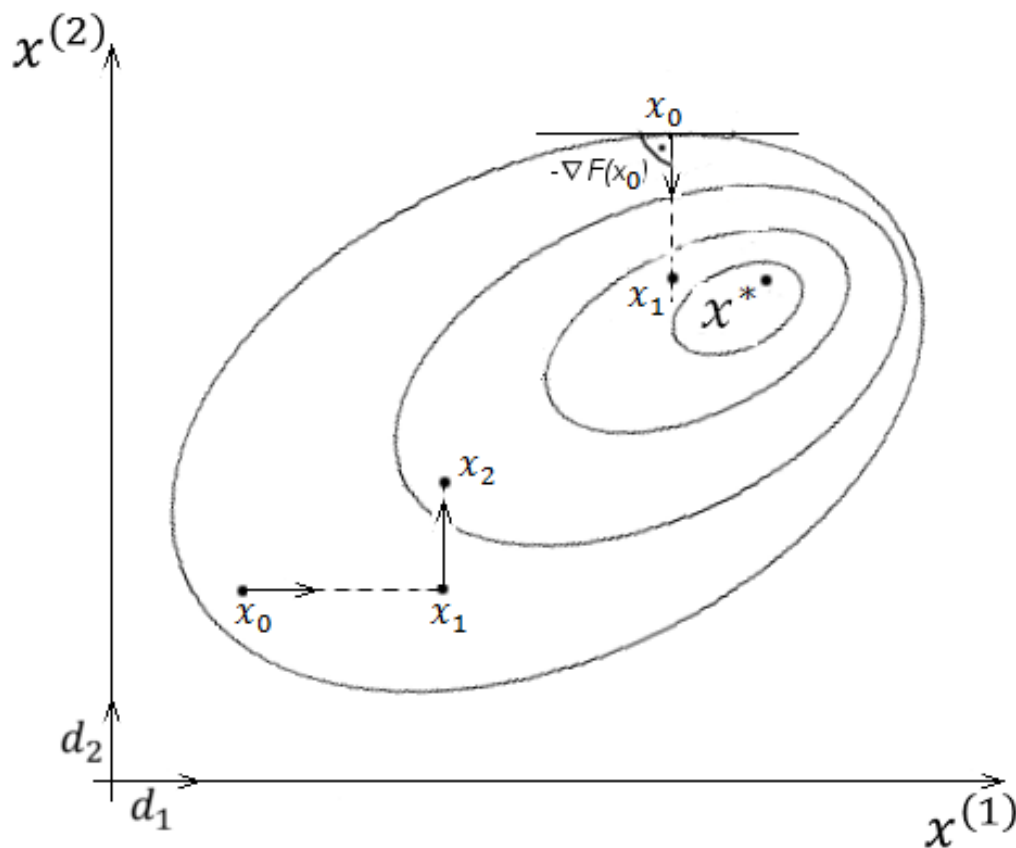


Politechnika Wrocławska

Bezgradientowe metody optymalizacji funkcji wielu zmiennych

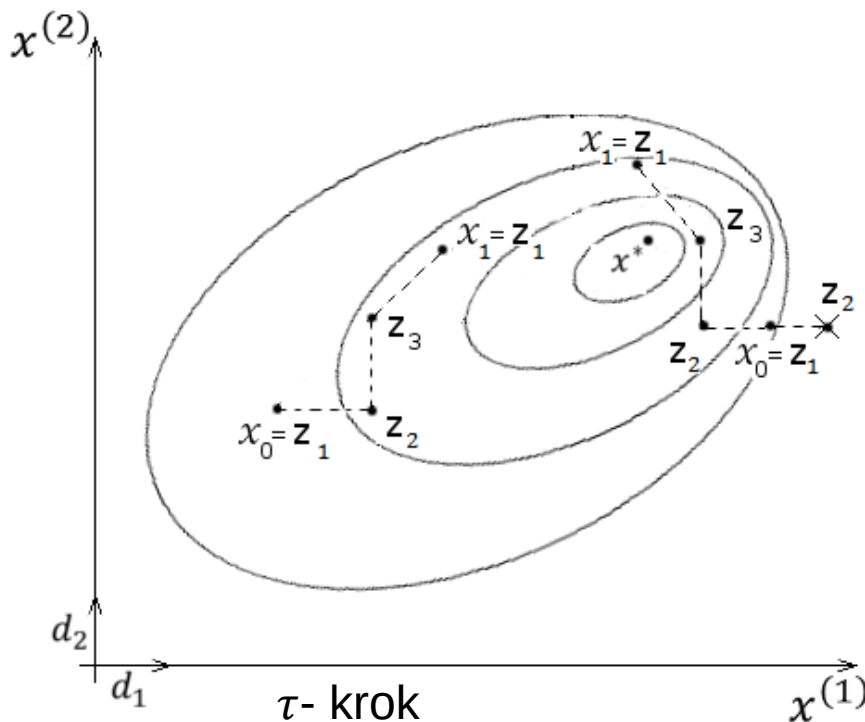
informacje dodatkowe

Wybór kierunku poszukiwań



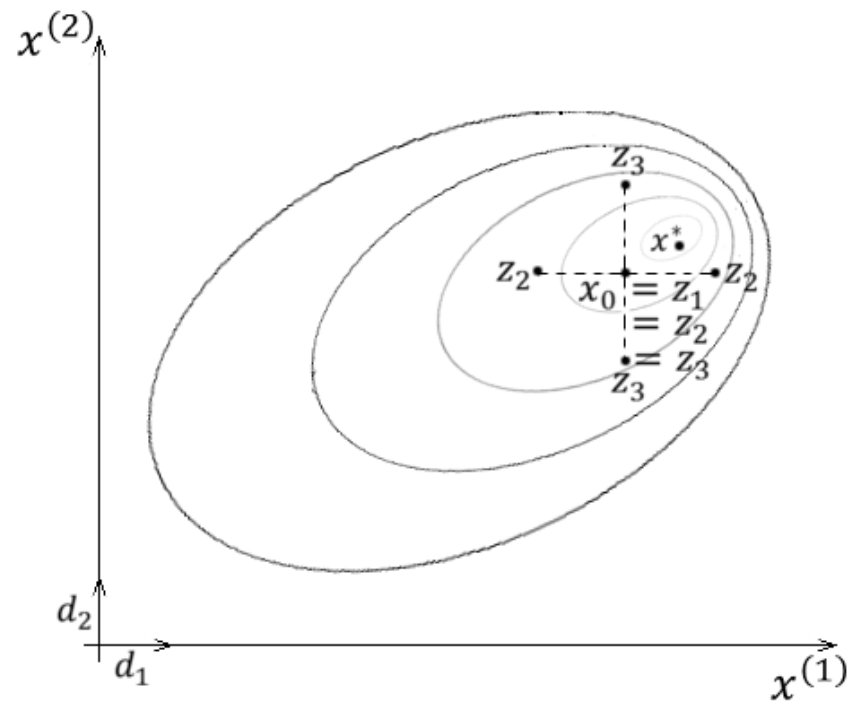
- Kierunki bazowe i ich modyfikacje – metody bezgradientowe.
- Kierunki oparte na gradiencie funkcji – metody gradientowe.

Metoda Hooka-Jeveesa - z krokiem dyskretnym



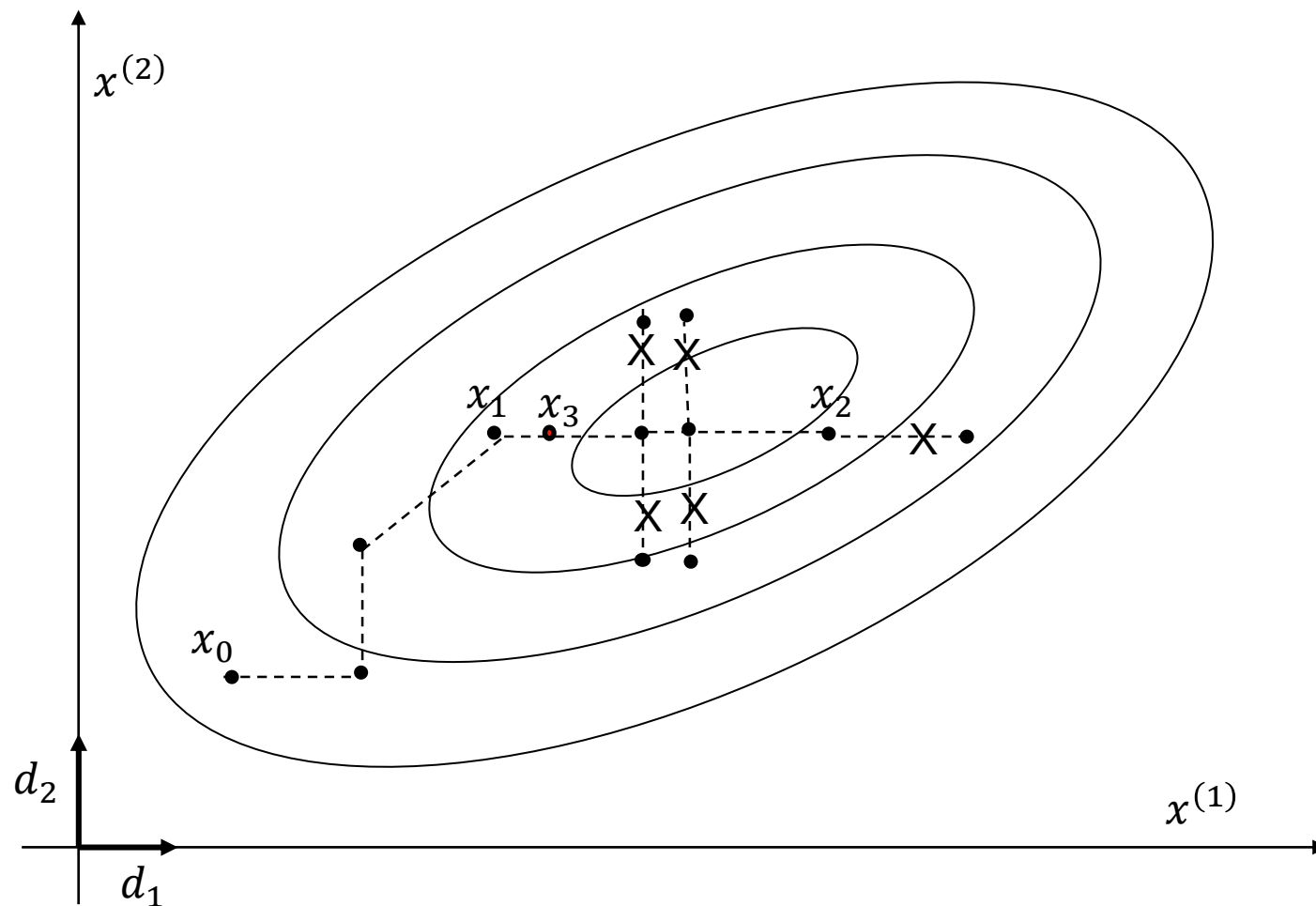
τ - krok

$\alpha > 1$ współczynnik kroku roboczego

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




Metoda Hooka-Jeevesa - z krokiem dyskretnym





Metoda Hooka-Jeveesa - z krokiem dyskretnym

DANE: $d_1, d_2, \dots, d_S, x_0, \tau, \varepsilon, \alpha, \beta$

Krok 0: $z_1 := x_0, s = 1, n = 0$

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

Jeśli $F(z_{s+1}) < F(z_s)$ idź do kroku 2

w przeciwnym razie $z_{s+1} := z_s - \tau d_s$

Jeśli $F(z_{s+1}) < F(z_s)$ idź do kroku 2

w przeciwnym razie $z_{s+1} := z_s$

Krok 2: Jeśli $s < S$, $s := s + 1$ idź do kroku 1

w przeciwnym razie

Jeśli $F(z_{s+1}) < F(z_1)$ idź do kroku 3

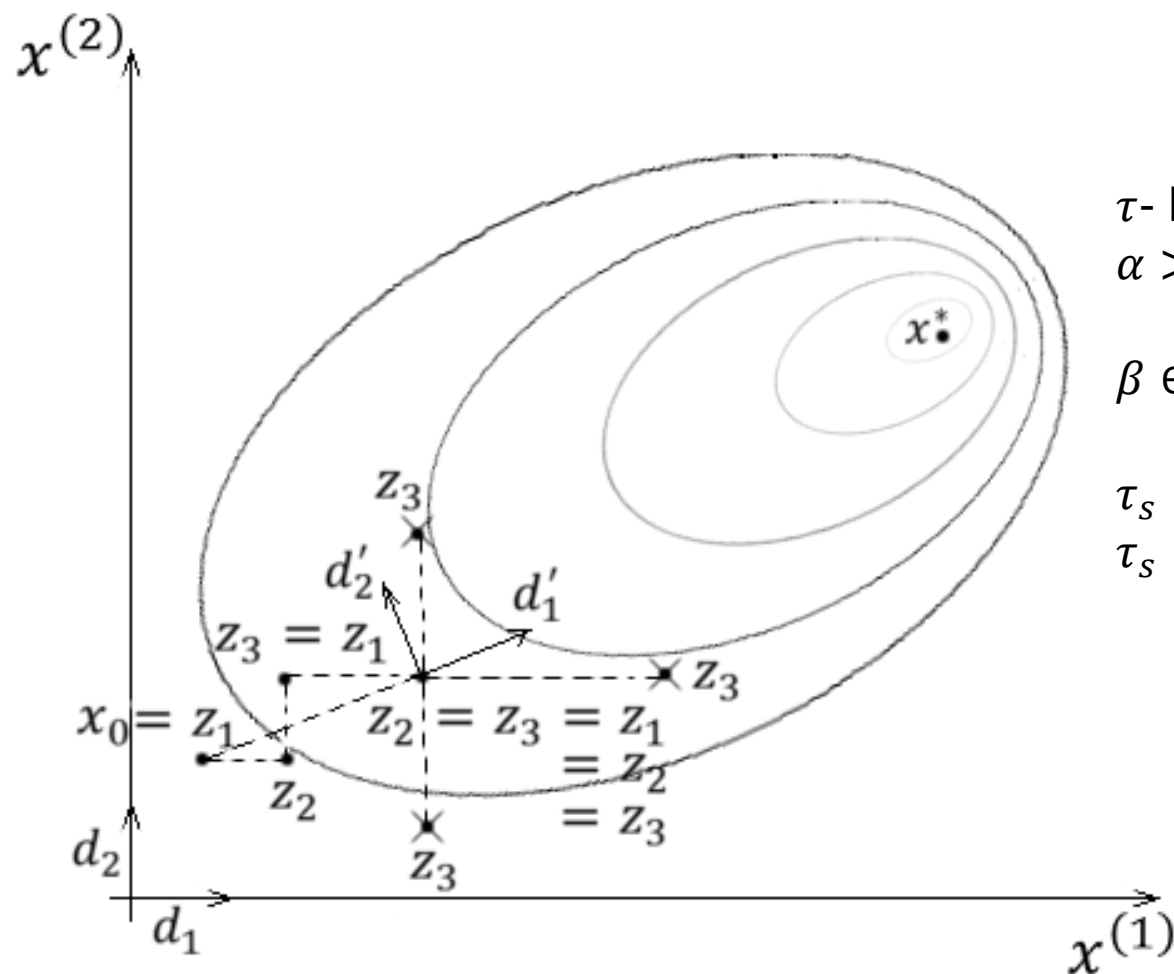
$\tau := \tau\beta, x_{n+1} := x_n, z_s := x_n, n := n + 1, s = 1$ idź do kroku 1

Krok 3: Jeśli $\tau < \varepsilon$, STOP

w przeciwnym razie

$x_{n+1} := z_{s+1} + \alpha(z_{s+1} - z_1), n := n + 1, s := 1$ idź do kroku 1

Metoda Rozenbrocka - z krokiem dyskretnym



τ - krok

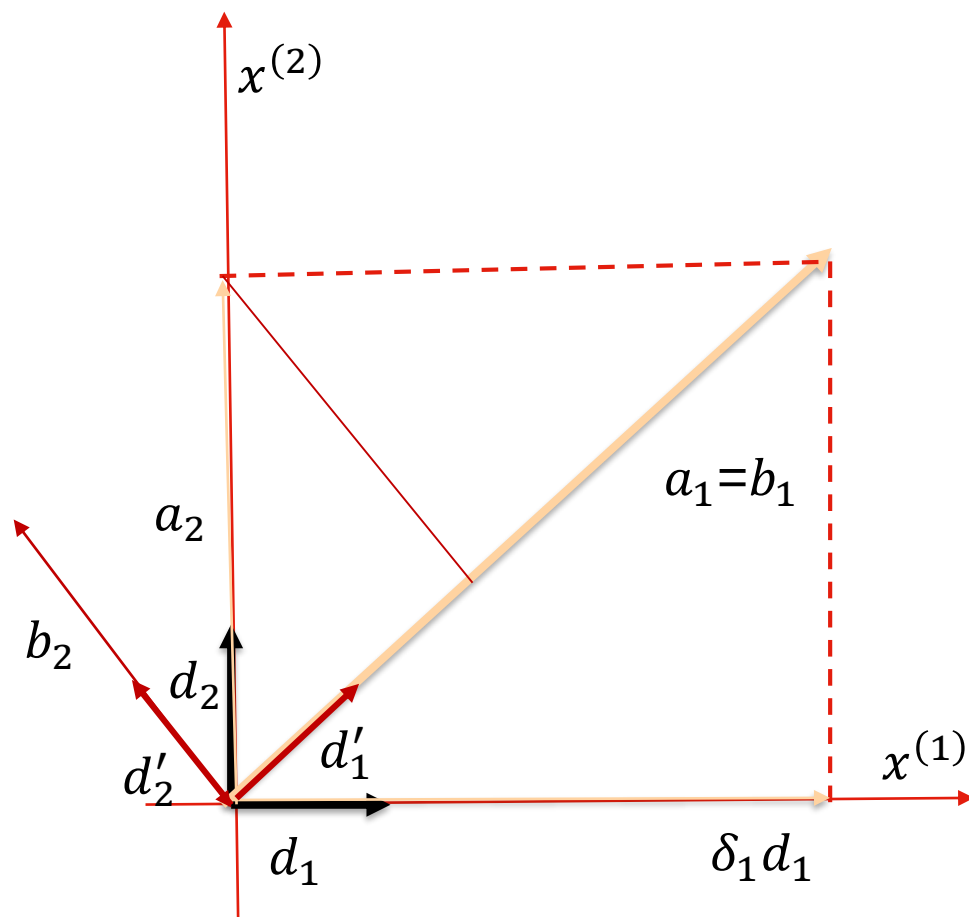
$\alpha > 1$ – współczynnik
korekcji kroku

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

$$\tau_s := \tau_s \alpha$$

$$\tau_s := \tau_s \beta$$

Obrót bazy – ortogonalizacja Grama - Schmitda



$$a_1 = \delta_1 d_1 + \delta_2 d_2$$

$$a_2 = \delta_2 d_2$$

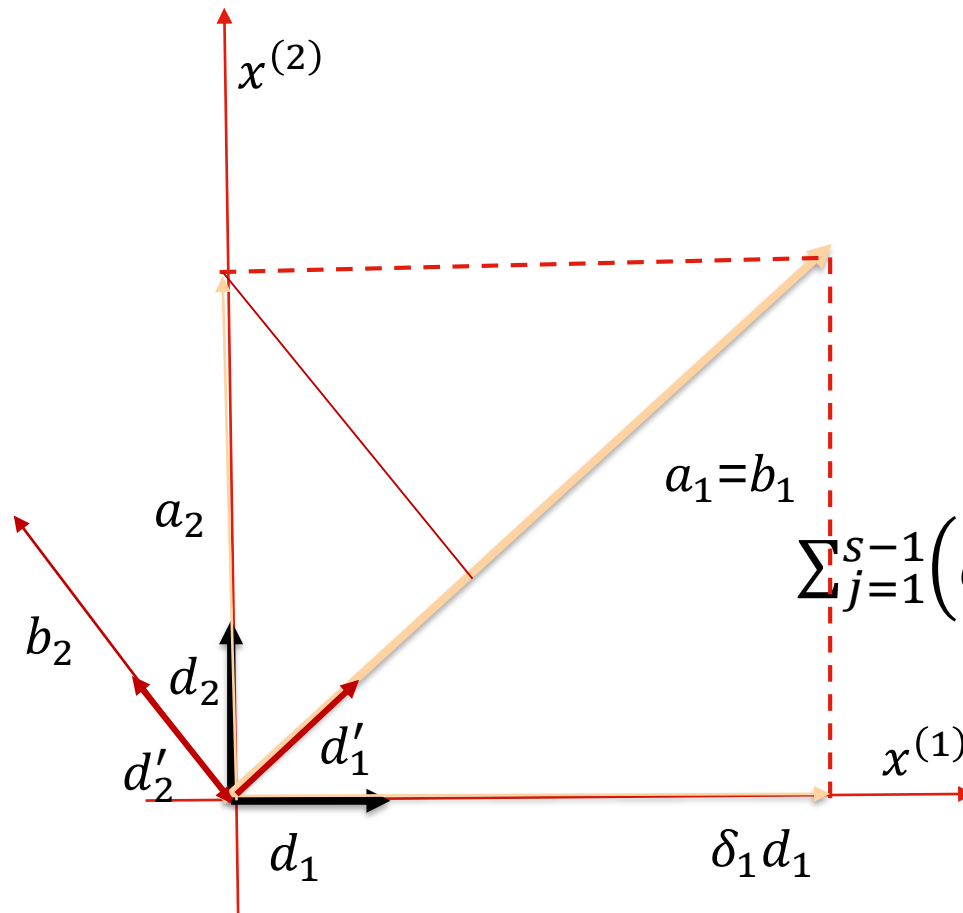
$$b_1 = a_1$$

$$d'_1 = \frac{b_1}{\|b_1\|}$$

$$b_2 = a_2 - (a_1^T d'_1) d'_1$$

$$d'_2 = \frac{b_2}{\|b_2\|}$$

Obrót bazy - ortogonalizacja Grama - Schmitda



$$a_s = \begin{cases} d_s & \delta_s = 0 \\ \sum_{j=1}^s \delta_j d_j & \delta_s \neq 0 \end{cases}$$

$$b_s = \begin{cases} a_s & s = 1 \\ a_s - \sum_{j=1}^{s-1} (a_j^T d'_j) d'_j & s = 2, \dots, S \end{cases}$$

$$d'_s = \frac{b_s}{\|b_s\|} \quad s = 1, 2, \dots, S$$



Metoda Rozenbrocka - z krokiem dyskretnym

DANE: $d_1, d_2, \dots, d_S, x_0, \tau, \varepsilon, \alpha > 1, \beta \in (-1, 0)$

Krok 0: $\tau_1 = \tau_2 = \dots = \tau_S = \tau, \delta_1 = \delta_2 = \dots = \delta_S = 0, z_1 = x_0, n = 0, s = 1$

Krok 1: $z_{s+1} = z_s + \tau_s d_s, \delta_s = \delta_s + \tau_s$
IF $F(z_{s+1}) < F(z_s)$ $\tau_s := \alpha \tau_s$ IDŹ DO KROKU 2
ELSE $F(z_{s+1}) \geq F(z_s)$ $\tau_s := \beta \tau_s$ IDŹ DO KROKU 2

Krok 2: IF $s < S$ $s := s + 1$ IDŹ DO KROKU 1
IF $F(z_{s+1}) < F(z_1)$ $z_1 := z_{s+1}$ $s := 1$ IDŹ DO KROKU 1
IF $F(z_{s+1}) = F(x_n)$
IF $n = 0$ zmień punkt startowy
ELSE $x_{n+1} = z_{s+1}$

Krok 3: IF STOP? (np.: $\|x_{n+1} - x_n\| < \varepsilon$) STOP

Krok 4: Obrót bazy

ELSE $n := n + 1, s = 1$

$\tau_1 = \tau_2 = \dots = \tau_S = \tau, \delta_1 = \delta_2 = \dots = \delta_S = 0$, IDŹ DO KROKU 1

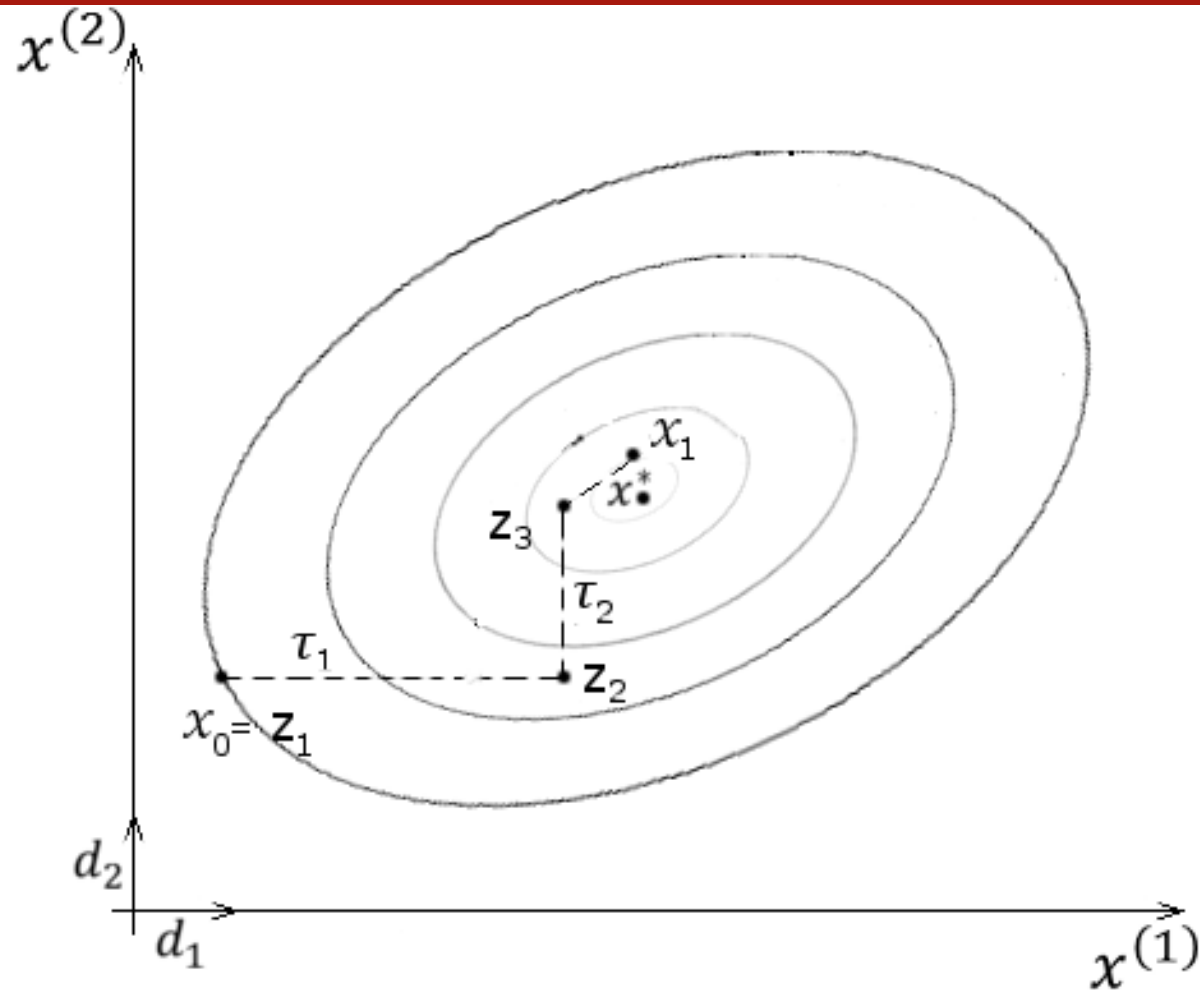
Obrót bazy

Krok 4:

$$a_s = \begin{cases} d_s & \delta_s = 0 \\ \sum_{j=s}^S \delta_j d_j & \delta_s \neq 0 \end{cases}$$
$$b_s = \begin{cases} a_s & s = 1 \\ a_s - \sum_{j=1}^{s-1} (a_j^T d'_j) d'_j \end{cases}$$

$$d'_s = \frac{b_s}{\|b_s\|} \quad s = 1, 2, \dots, S$$

Metoda Hooka-Jeevesa - z krokiem optymalnym





Metoda Hooka-Jeevesa - z krokiem optymalnym

DANE: $d_1, d_2, \dots, d_S, x_0, \varepsilon$

Krok 0: $z_1 := x_0, n = 0, s = 1$

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

τ_s - min w kierunku d_s

Krok 2: Jeśli $s < S, s = s + 1$ idź do kroku 1

Jeśli $\|z_{s+1} - z_1\| < \varepsilon$ - STOP

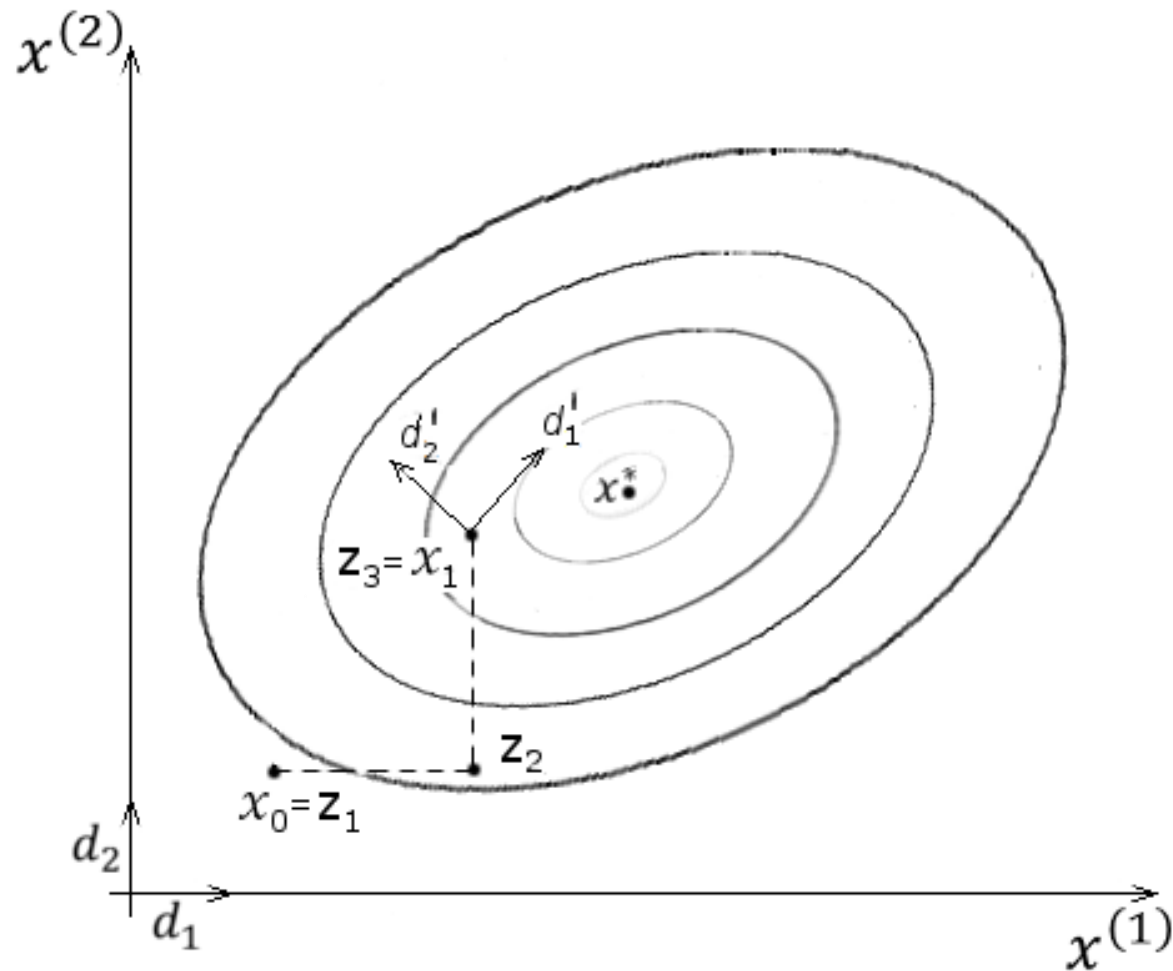
Krok 3: $x_{n+1} := z_s + \tau d$

$\tau \rightarrow$ min w kierunku d

$$d = \frac{z_{s+1} - z_1}{\|z_{s+1} - z_1\|} = \frac{\sum_{s=1}^{S+1} \tau_s d_s}{\|\sum_{s=1}^{S+1} \tau_s d_s\|}$$

$n := n + 1, s = 1, z_1 = x_n$

Metoda Rozenbrocka - z krokiem optymalnym





Metoda Rozenbrocka - z krokiem optymalnym

DANE: $d_1, d_2, \dots, d_S, x_0, \varepsilon$

Krok 0: $z_1 := x_0, n = 0, s = 1$

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

τ_s - min w kierunku d_s

Krok 2: Jeśli $s < S, s := s + 1$ idź do kroku 1

Jeśli $\|z_{s+1} - z_1\| < \varepsilon$ - STOP

Krok 3:

$$a_s = \begin{cases} d_s & \tau_s = 0 \\ \sum_{j=s}^S \tau_j d_j & \tau_s \neq 0 \end{cases}$$
$$b_s = \begin{cases} a_s & s = 1 \\ a_s - \sum_{j=1}^{s-1} (a_j^T d'_j) d'_j \end{cases}$$

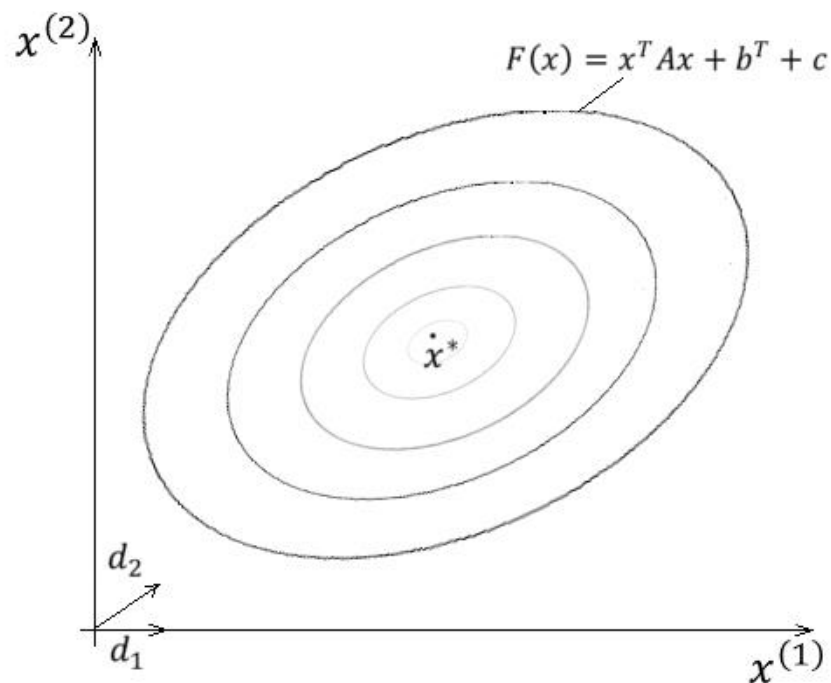
$$d'_s = \frac{b_s}{\|b_s\|} \quad s = 1, 2, \dots, S$$

Krok 4: $d_s := d'_s \quad s = 1, 2, \dots, S, n := n + 1, s = 1$ idź do kroku 1

Metoda Powella - wektory sprzężone

$d_1, d_2, \dots, d_s$ - sprzężone według macierzy A , symetrycznej i dodatnio określonej

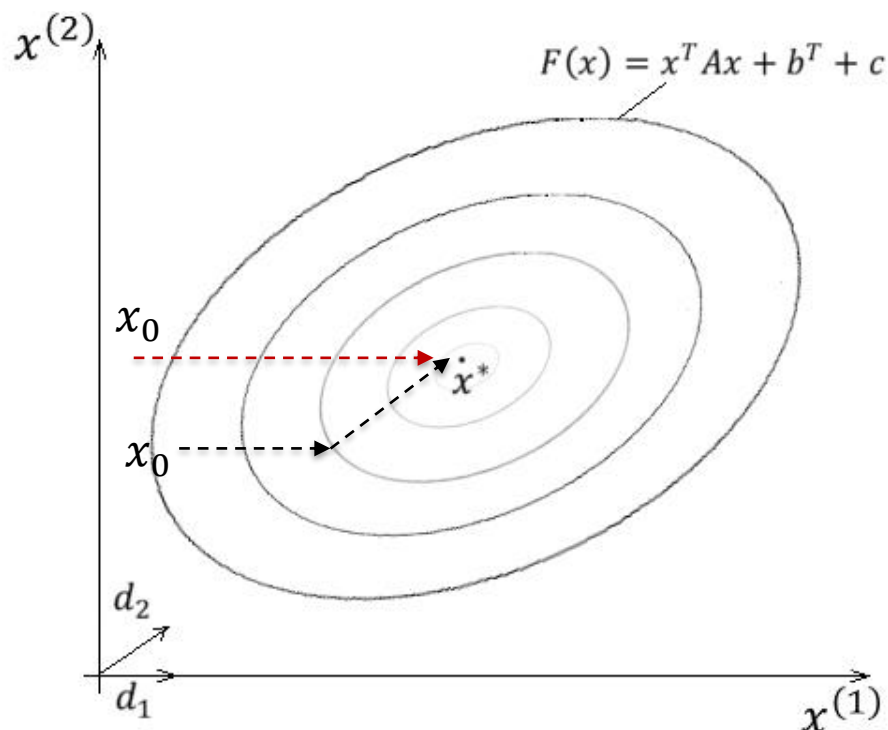
$$d_i^T A d_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$



Metoda Powella - wektory sprzężone

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

Dla tak dobranej funkcji, jeżeli zestaw wektorów $d_1, d_2, \dots, d_S$ jest sprzężony według macierzy A , to minimalizując wzdłuż kierunków sprzężonych otrzymamy rozwiązanie w nie więcej niż S krokach.



Przykład

$$F(x) = (x^{(1)} - 2)^2 + (x^{(2)} - 1)^2, x_0 = \begin{bmatrix} -2 \\ -2 \end{bmatrix}, d_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, d_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad H(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$d_1^T H d_2 = [1 \quad 0] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0, d_1^T H d_1 = [1 \quad 0] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1, d_2^T H d_2 = [0 \quad 1] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1$$

$$x_1 = x_0 + \tau d_1 = \begin{bmatrix} -2 \\ -2 \end{bmatrix} + \tau \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 + 1 * \tau \\ -2 + 0 * \tau \end{bmatrix} \quad F(x_0 + \tau d) = (-2 + \tau - 2)^2 + (-2 - 1)^2 = (\tau - 4)^2 + 3^2 = \tau^2 - 4\tau + 25 \triangleq f_1(\tau)$$

$$f_1'(\tau) = 2\tau - 8 = 0 \quad \tau^* = 4$$

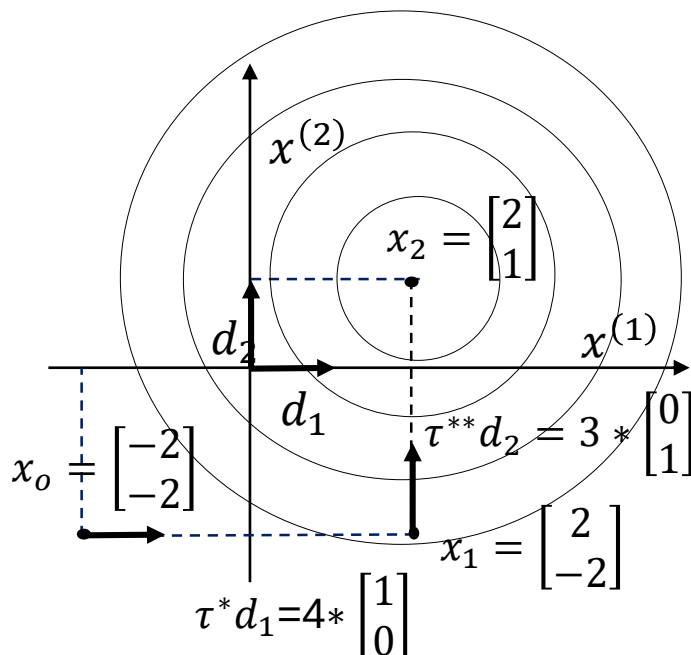
$$x_1 = x_0 + \tau^* d_1 = \begin{bmatrix} -2 \\ -2 \end{bmatrix} + 4 * \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

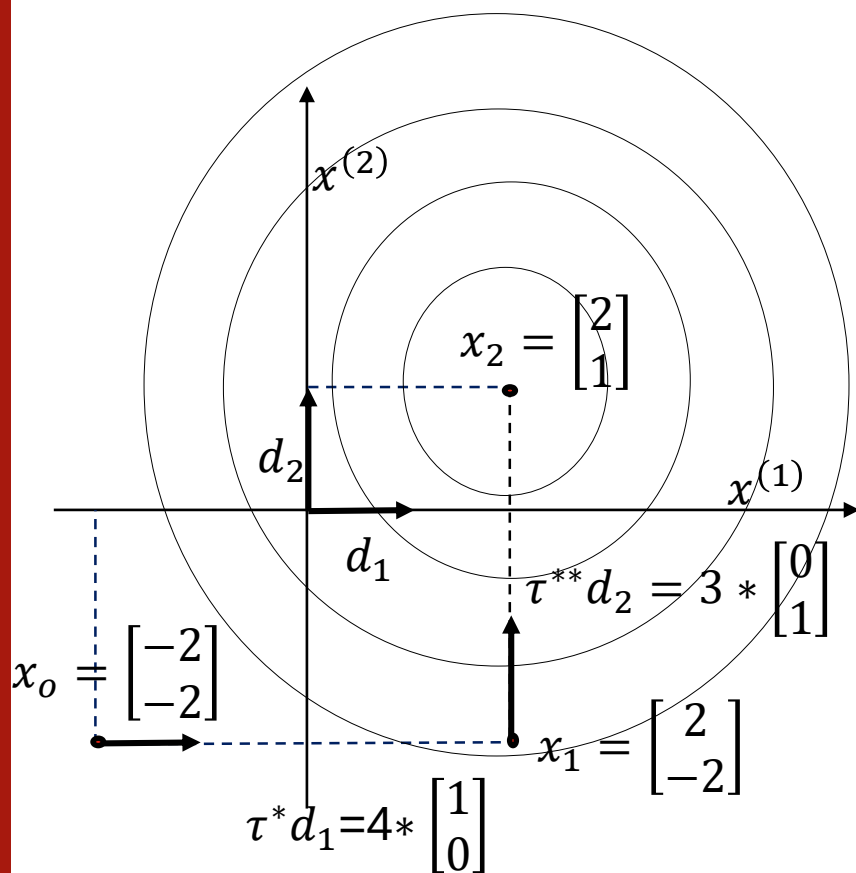
$$x_2 = x_1 + \tau d_2 = \begin{bmatrix} 2 \\ -2 \end{bmatrix} + \tau \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 + 0 * \tau \\ -2 + 1 * \tau \end{bmatrix}$$

$$F(x_1 + \tau d_2) = (2 - 2)^2 + (-2 + \tau - 1)^2 = 4 + (\tau - 3)^2 \triangleq f_2(\tau)$$

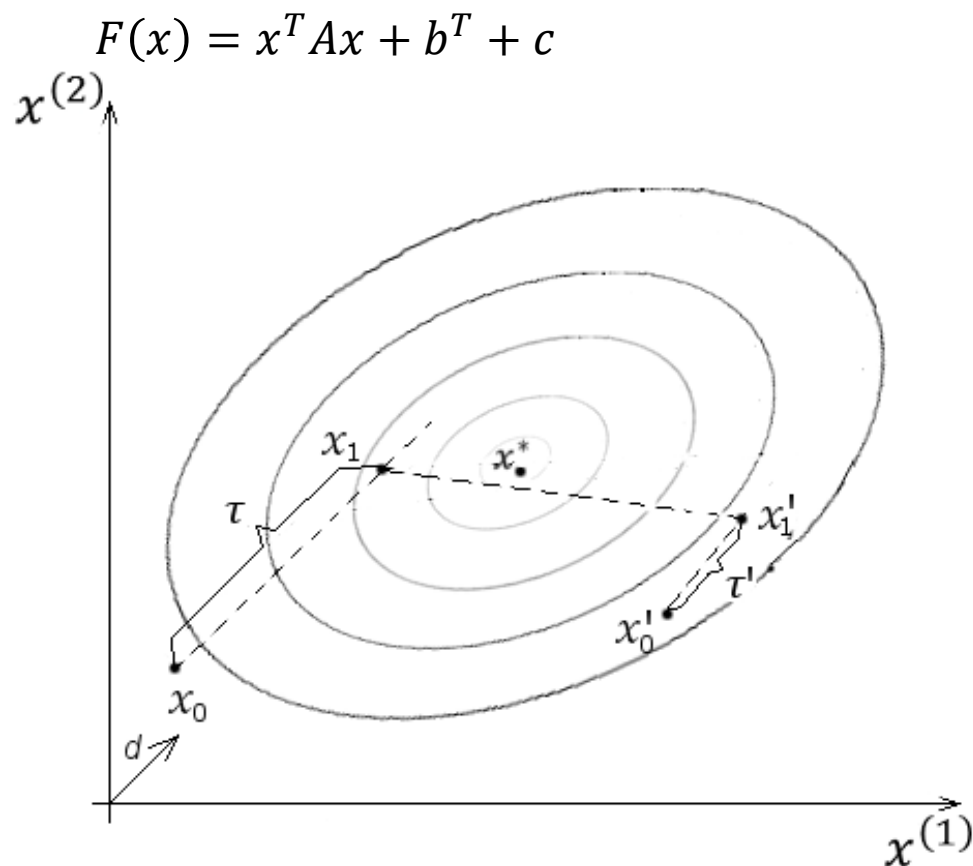
$$f_2'(\tau) = 2\tau - 6 = 0 \quad \tau^{**} = 3$$

$$x_2 = x_1 + \tau d_2 = \begin{bmatrix} 2 \\ -2 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$





Metoda Powella



$$x_1 = x_0 + \tau^* d$$

τ^* - min w kierunku d z x_0

$$x_1' = x_0' + \tau^{*'} d$$

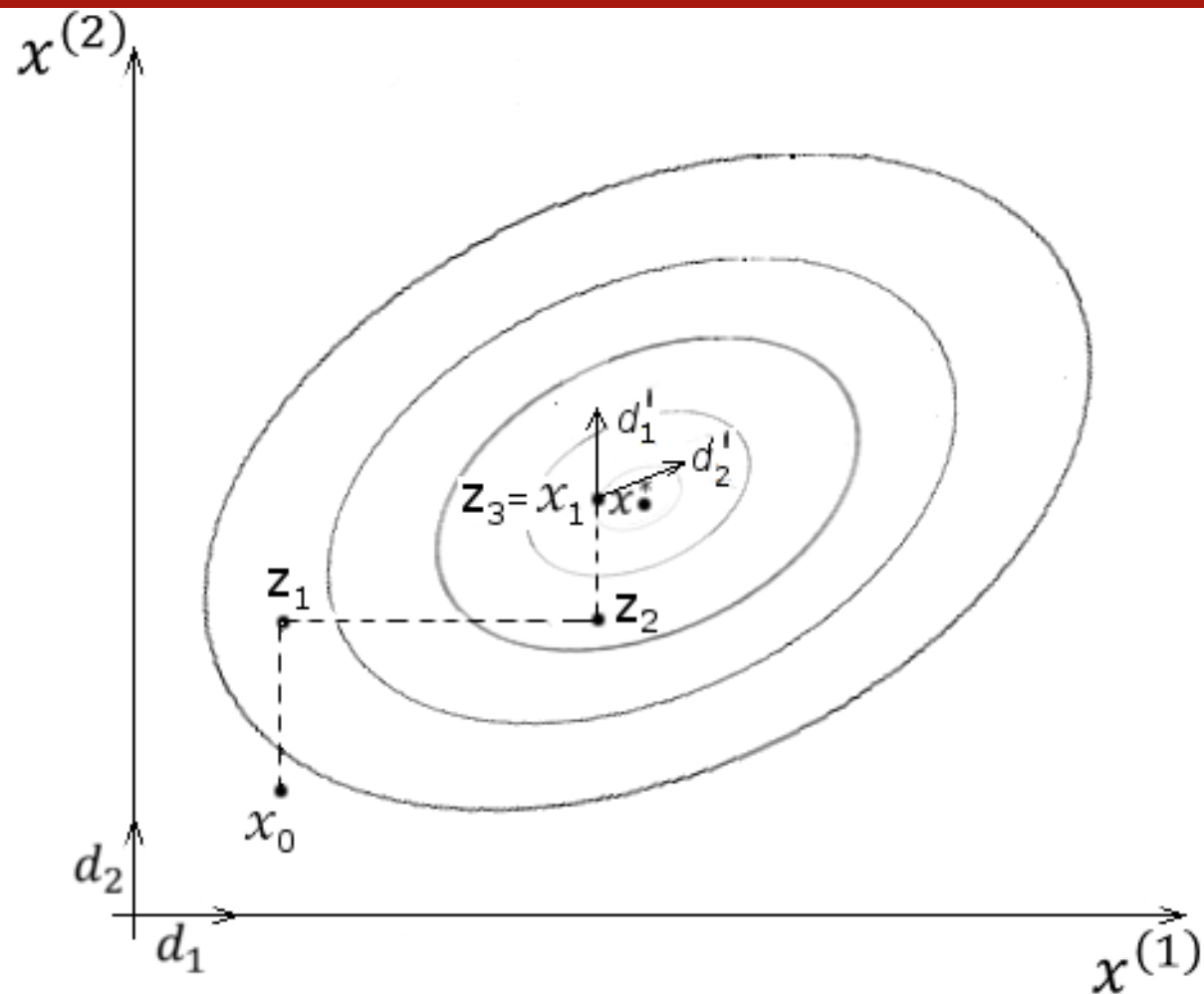
$\tau^{*'}$ - min w kierunku d z x_0'

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

$$d^T A d' = 0$$

d, d' - sprzężone według A

Metoda Powella





Metoda Powella

DANE: $d_1, d_2, \dots, d_S, x_0, \varepsilon$

Krok 0: $z_1 := x_0 + \tau_S d_S, n := 0, s := 1$

τ_S - min w kierunku d_S

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

τ_s - min w kierunku d_s

Krok 2: Jeśli $s < S, s := s + 1$ idź do kroku 1

Jeśli $\|z_{s+1} - z_1\| < \varepsilon$ - STOP

Krok 3: $x_{n+1} := z_{s+1}$

$$d := \frac{z_{s+1} - z_1}{\|z_{s+1} - z_1\|}$$

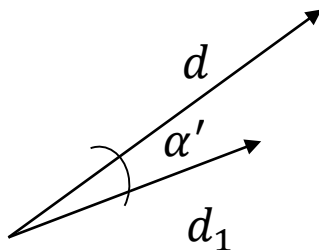
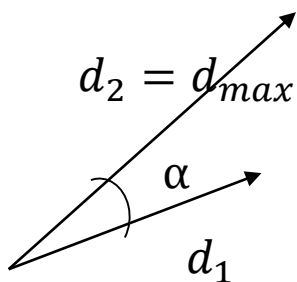
$z_1 := x_{n+1} + \tau d$ τ - min w kierunku d

$d_s := d_{s+1}$ $s = 1, 2, \dots, S - 1$

$d_S := d$

Idź do Kroku 1

Metoda Powella - modyfikacja



$$\Delta = \det[d_1 \ d_2 \ \cdots \ d_S] = \cos \alpha$$

$$\Delta = \det[d_1 \ d_2 \ \cdots \ d_{max} \ \cdots \ d_S] = \cos \alpha$$

$\uparrow d$

$$\Delta' = \det[d_1 \ d_2 \ \cdots \ d \ \cdots \ d_S] = \frac{\tau_{max} \Delta}{\|Z_{S+1} - Z_1\|} = \cos \alpha'$$

Krok 3: $x_{n+1} := z_{S+1}$
 $d := \frac{z_{S+1} - z_1}{\|z_{S+1} - z_1\|}$

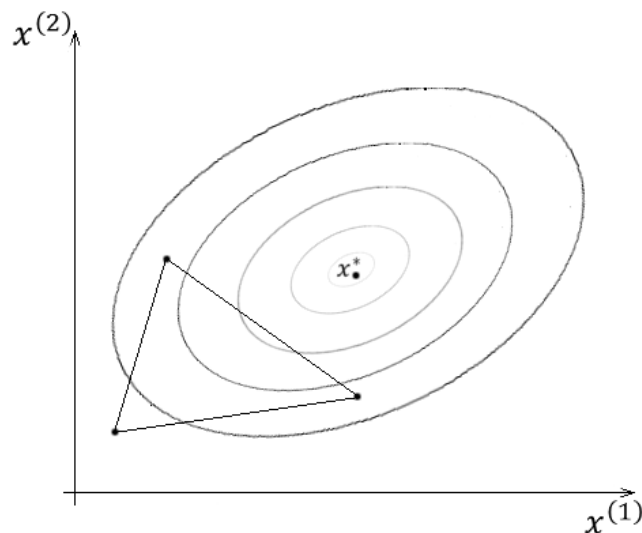
$z_1 := x_{n+1} + \tau d$ τ – min w kierunku d

$$\tau_{max} \rightarrow \max_{1 \leq s \leq S} \|z_{S+1} - z_s\| = \max_{1 \leq s \leq S} \tau_s$$

Jeżeli $\Delta := \frac{\tau_{max} \Delta}{\|z_{S+1} - z_1\|} > 0.8$ $d_{max} = d$ idź do kroku 1

Metoda Nelder-Meada (simplex)

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



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

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

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

Generowanie simplexu początkowego

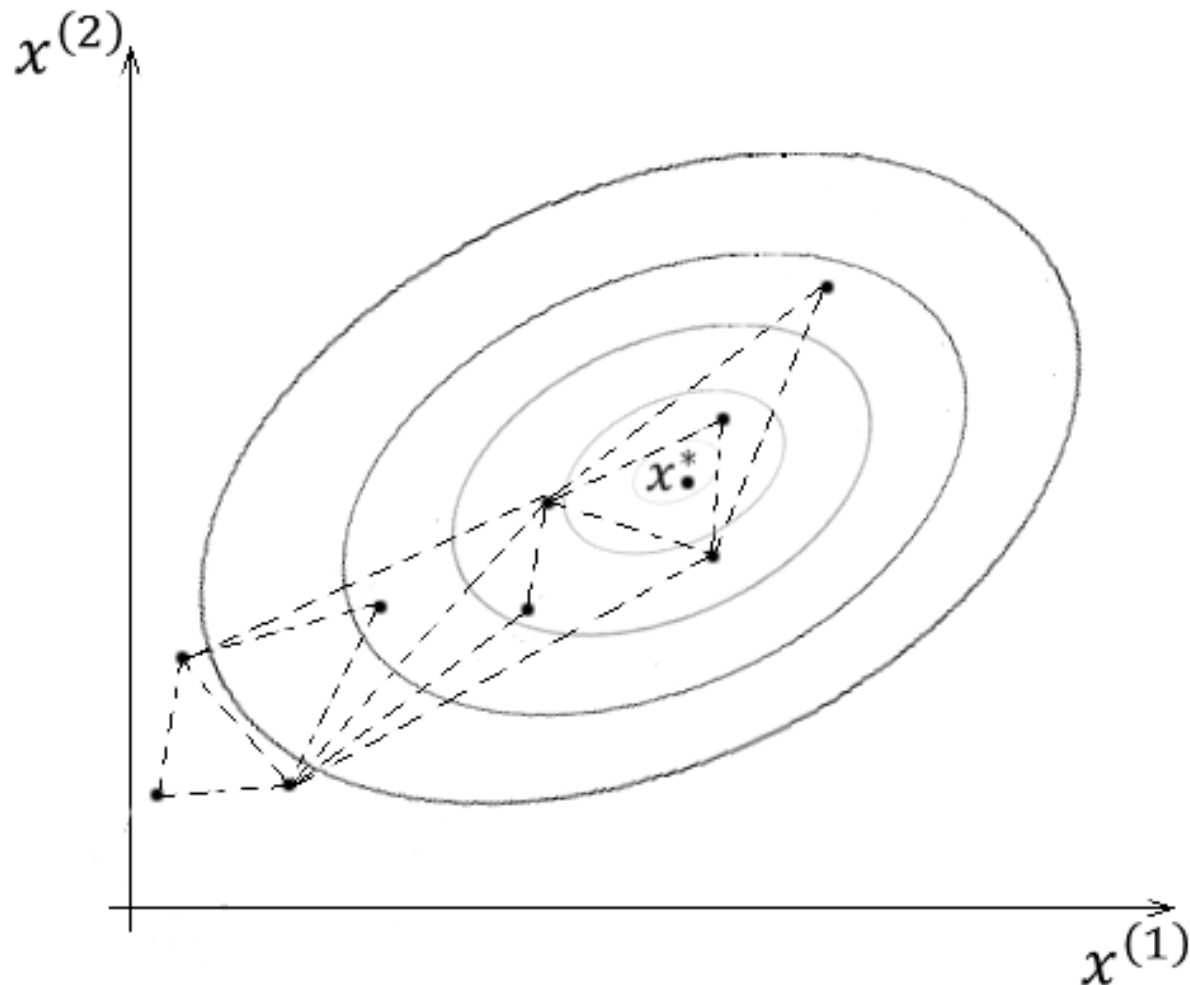
$$x_0, c \quad a = \frac{c}{s\sqrt{2}} (\sqrt{S+1} + \sqrt{2} - 1)$$

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

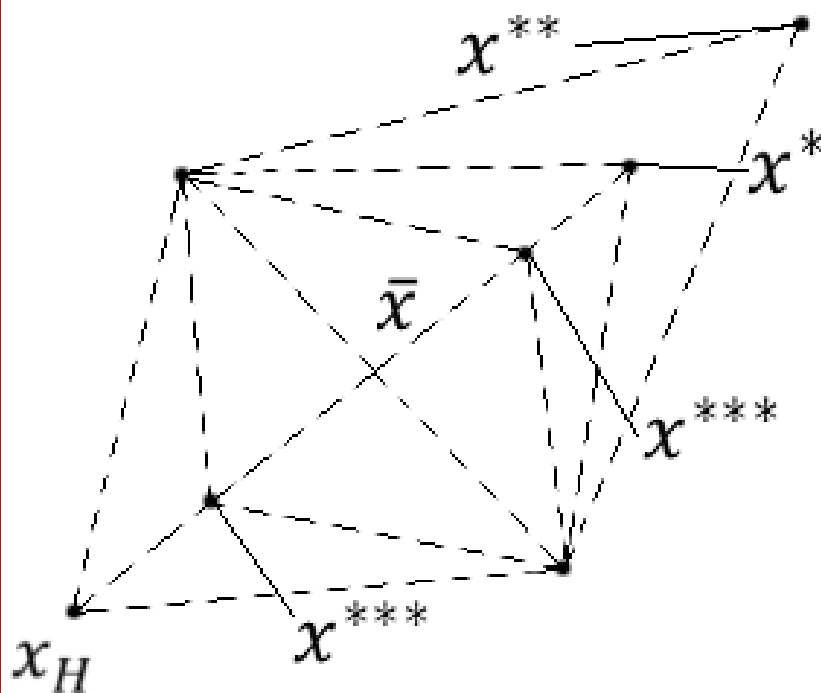
$$d_j^T = [a \ a \ \dots \ a \ b \ a \ \dots \ a]$$

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

Metoda Nelder-Meada (simplex)



Metoda Nelder-Meada



Odbicie

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

α – współczynnik odbicia

Jeśli $\alpha > 0$

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

Ekspansja

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

γ – współczynnik ekspansji

Kontrakcja

Jeśli $F(x^*) > F(x_H)$

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

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

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

β – współczynnik kontrakcji



Metoda Nelder-Meada

DANE: x_0, c, ε

Krok 0: $x_1 x_2 \dots x_{S+1}$ - simplex początkowy, $n = 0$

Krok 1: $x_H \rightarrow F(x_H) = \max_{1 \leq s \leq S+1} F(x_s), x_L \rightarrow F(x_L) = \min_{1 \leq s \leq S+1} F(x_s)$

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

Krok 2: $x^* = \bar{x} + \alpha(\bar{x} - x_H)$

Jeśli $F(x^*) < F(x_L)$ $x^{**} = \bar{x} + \gamma(x^* - \bar{x})$ idź do kroku 3
w przeciwnym razie idź do kroku 4

Krok 3: Jeśli $F(x^{**}) < F(x^*)$ $x_H = x^{**}, n = n + 1$ idź do kroku 1
w przeciwnym razie $x_H = x^*, n = n + 1$ idź do kroku 1

Krok 4: Jeśli $F(x^*) < \max_{\substack{1 \leq s \leq S+1 \\ s \neq H}} F(x_s)$ $x_H = x^*, n = n + 1$

Krok 5: $x' \leftarrow F(x') = \min\{F(x^*), F(x_H)\}$

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

Jeśli $F(x^{***}) > F(x')$ $x_j = x_j + \frac{1}{2}(x_L - x_j), j = 1, 2, \dots, S + 1$ idź do kroku 1
 $x_H = x^{***}, n = n + 1$ idź do kroku 1

Dziękuję za uwagę

