Computer Science

Jerzy Świątek

Systems Modelling and Analysis

L.4. Model based decision making

Model in the systems research

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
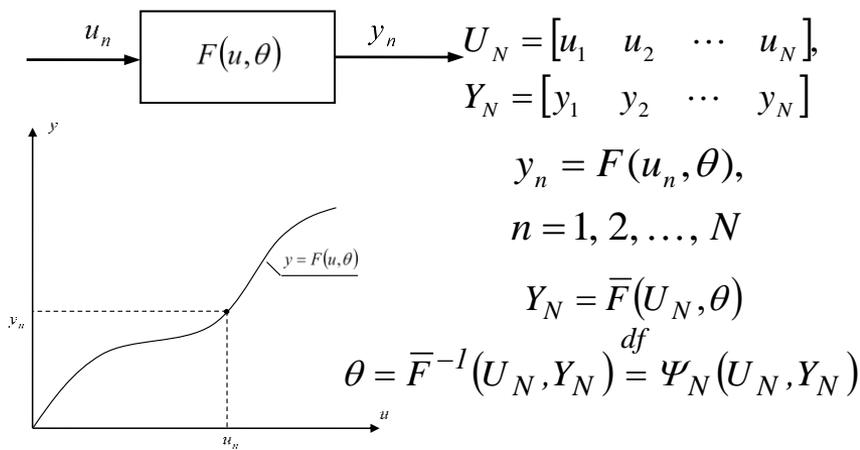
Wrocław University of Technology

EUROPEAN
SOCIAL FUND
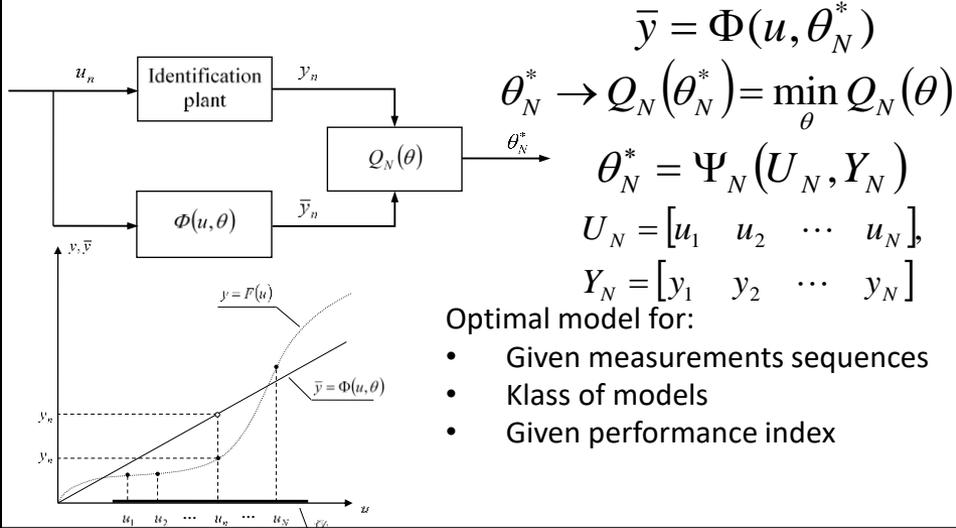
Project co-financed from the EU European Social Fund
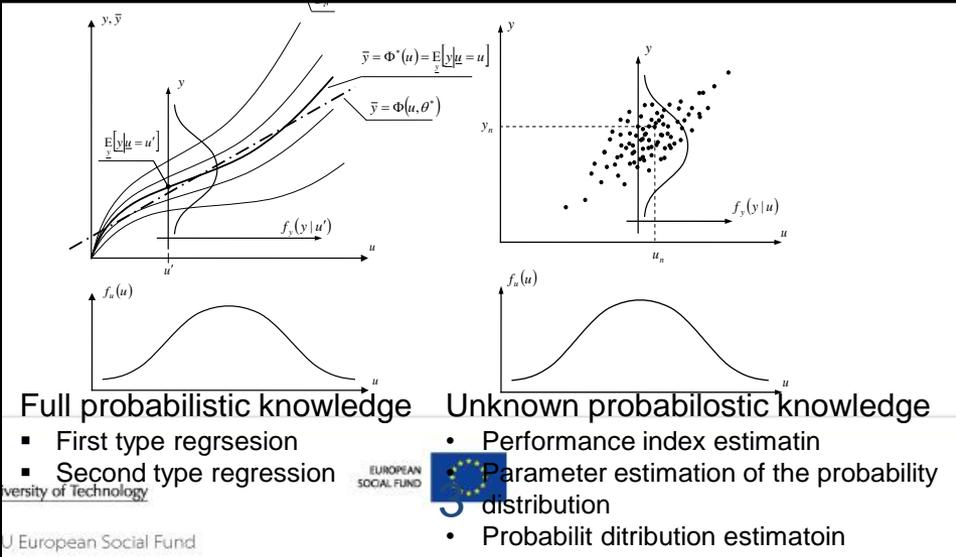
Plant in the class of model

D e t e r m i n i s t y c z n y

$$U_N = \begin{bmatrix} u_1 & u_2 & \cdots & u_N \end{bmatrix},$$
$$Y_N = \begin{bmatrix} y_1 & y_2 & \cdots & y_N \end{bmatrix}$$

$$y_n = F(u_n, \theta),$$
$$n = 1, 2, \ldots, N$$

$$Y_N = \overline{F}(U_N, \theta)$$

$$\theta = \overline{F}^{-1}(U_N, Y_N) \overset{df}{=} \Psi_N(U_N, Y_N)$$

Choice of the best model

$$\overline{y} = \Phi(u, \theta_N^*)$$
$$\theta_N^* \rightarrow Q_N(\theta_N^*) = \min_\theta Q_N(\theta)$$
$$\theta_N^* = \Psi_N(U_N, Y_N)$$
$$U_N = \begin{bmatrix} u_1 & u_2 & \cdots & u_N \end{bmatrix},$$
$$Y_N = \begin{bmatrix} y_1 & y_2 & \cdots & y_N \end{bmatrix}$$

Optimal model for:
- Given measurements sequences
- Klass of models
- Given performance index

Methods:
- Least square
- Maximum likelihood
- Bayes'a

L o s o w y

$$\theta_N = \Psi_N(U_N, W_N)$$

Full probabilistic knowledge
- First type regrsesion
- Second type regression

Unknown probabilostic knowledge
- Performance index estimatin
- Parameter estimation of the probability distribution
- Probabilit ditribution estimatoin

Example of decision making

Decision: workloads of power plants

hydroelectric plant
$x^{(1)}$

nuclear power plant
$x^{(2)}$

wind turbine
$x^{(3)}$

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

Given parameters:

c_1, c_2, c_3 – unit costs of workloads

Objective is to minimize overall costs: $F(x^{(1)}, x^{(2)}, x^{(3)}) = c_1 x^{(1)} + c_2 x^{(2)} + c_3 x^{(3)}$

Constraints: – demand must be met: $x^{(1)} + x^{(2)} + x^{(3)} \geq \beta$

             – energy production capabilities are limited: $0 \leq x^{(n)} \leq \alpha_n, \ n = 1,2,3$

Basic ingredients of optimization task formulation

Decision variables: $x = \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(S)} \end{bmatrix}$

Objective function: $y = F(x)$

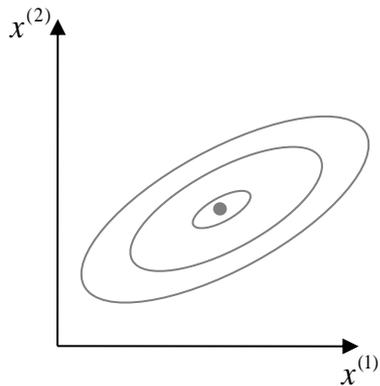**Set of feasible decisions** (commonly defined by variables domain and constraints):

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

Optimization task: $x^* \to F(x^*) = \min_{x^* \in \mathscr{D}_x} F(x), \qquad x^* -$ optimal decision

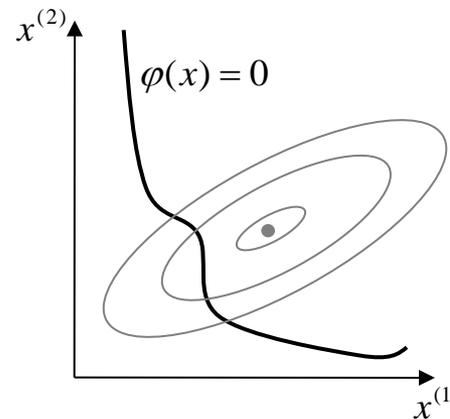$$\min F(x) = -\max\left(-F(x)\right)$$

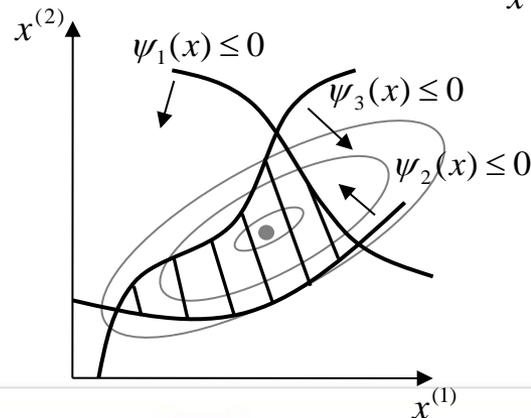# General classification of optimization tasks

Unconstrained optimization:

$$\mathscr{D}_x = \mathscr{R}^S$$

Optimization under equality constraints:

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

Optimization under inequality constraints:

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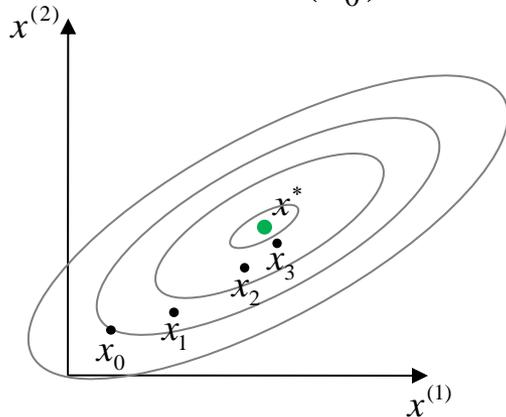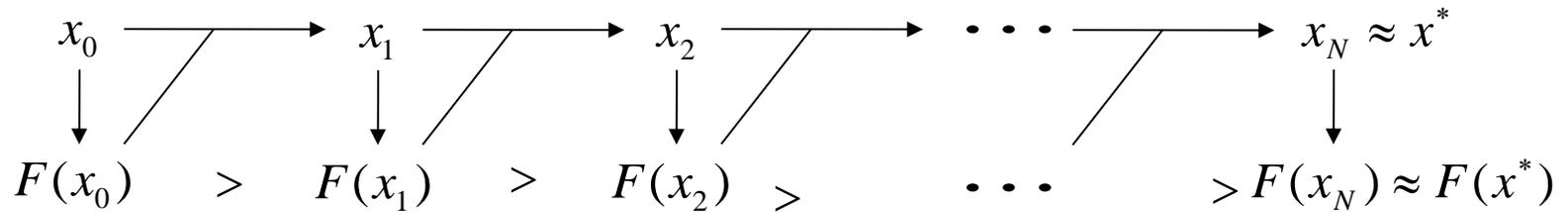
Analytical methods

- Unconstrained optimization

- Lagrange multipliers method – equality constraints

- Kuhn-Tucker conditions – inequality constraints

Numerical methods

We only use information about values of objective function $F(x)$ for a given value of x.

$$x_0 \longrightarrow x_1 \longrightarrow x_2 \longrightarrow \cdots \longrightarrow x_N \approx x^*$$

$$F(x_0) \quad > \quad F(x_1) \quad > \quad F(x_2) \quad > \quad \cdots \quad > F(x_N) \approx F(x^*)$$

The general idea behind numerical methods.

Common types of optimization tasks

- Linear programming

Decision variables: $x \in \mathcal{D}_x \subseteq \mathcal{R}^S$

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

$$c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_S \end{bmatrix}$$

Constraints:

$$\varphi_l(x) = a_l^T x - \alpha_l = 0,$$
$$l = 1, 2, \ldots, L$$

$$a_l = \begin{bmatrix} a_l^{(1)} \\ a_l^{(2)} \\ \vdots \\ a_l^{(S)} \end{bmatrix}$$

$$\psi_m(x) = b_m^T x - \beta_m \leq 0,$$
$$m = 1, 2, \ldots, M$$

$$b_l = \begin{bmatrix} b_m^{(1)} \\ b_m^{(2)} \\ \vdots \\ b_m^{(S)} \end{bmatrix}$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Common types of optimization tasks

- Quadratic programming

Decision variables: $x \in \mathscr{D}_x \subseteq \mathscr{R}^S$

Objective function: $F(x) = x^T A x + b^T x + c$ $A \in \mathscr{R}^{S \times S}$, $b \in \mathscr{R}^S$, $c \in \mathscr{R}$

Constraints:

$$\varphi_l(x) = d_l^T x - \alpha_l = 0,$$
$$l = 1, 2, \ldots, L$$

$$d_l = \begin{bmatrix} d_l^{(1)} \\ d_l^{(2)} \\ \vdots \\ d_l^{(S)} \end{bmatrix}$$

$$\psi_m(x) = e_m^T x - \beta_m \le 0,$$
$$m = 1, 2, \ldots, M$$

$$e_l = \begin{bmatrix} e_m^{(1)} \\ e_m^{(2)} \\ \vdots \\ e_m^{(S)} \end{bmatrix}$$

Common types of optimization tasks

- Linear-fractional programming

Decision variables: $\quad x \in \mathscr{D}_x \subseteq \mathscr{R}^S$

Objective function: $\quad F(x) = \dfrac{a^T x + b}{c^T x + d}$ $\qquad a \in \mathscr{R}^S,\, b \in \mathscr{R},\, c \in \mathscr{R}^S, d \in \mathscr{R}$

Constraints:

$$\varphi_l(x) = p_l^T x - \alpha_l = 0, \\ l = 1,2,\ldots,L \qquad p_l = \begin{bmatrix} p_l^{(1)} \\ p_l^{(2)} \\ \vdots \\ p_l^{(S)} \end{bmatrix} \qquad \psi_m(x) = q_m^T x - \beta_m \le 0, \\ m = 1,2,\ldots,M \qquad q_l = \begin{bmatrix} q_m^{(1)} \\ q_m^{(2)} \\ \vdots \\ q_m^{(S)} \end{bmatrix}$$
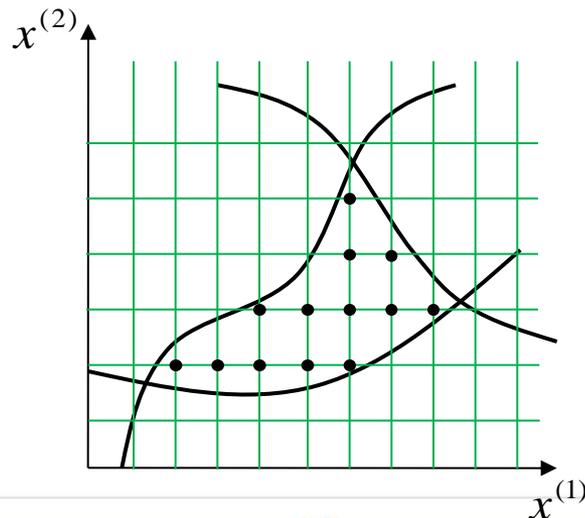
Common types of optimization tasks

- Integer programming

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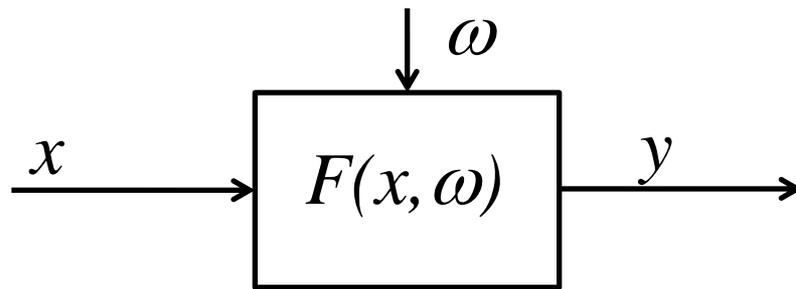

$x^{(2)}$

$x^{(1)}$

Special cases

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

$$x \in \overline{\mathscr{D}_x} = \left\{ x^{(s)} \in \left\{ 0, 1 \right\}, s = 1, 2, \ldots, S \right\}$$

Decision making under uncertainty

$$x \longrightarrow \boxed{\begin{array}{c} \omega \downarrow \\ F(x,\omega) \end{array}} \longrightarrow y$$

$$x^* \rightarrow F\left(x^*,\omega\right) = \min_{x \in D_x(\omega)} F(x,\omega) \quad ???$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Multiobjective optimization

x – vector of decision variables

$F_1(x), F_2(x), \ldots, F_M(x)$ – performance indices

HUMAN CAPITAL
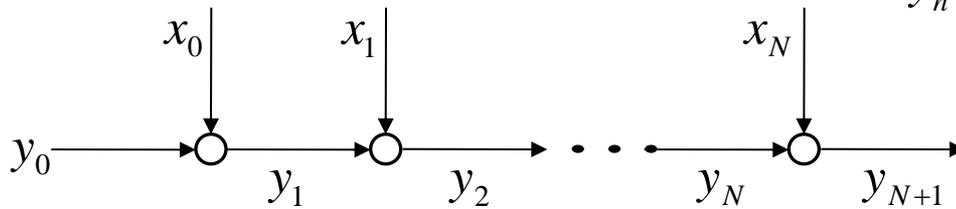HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund
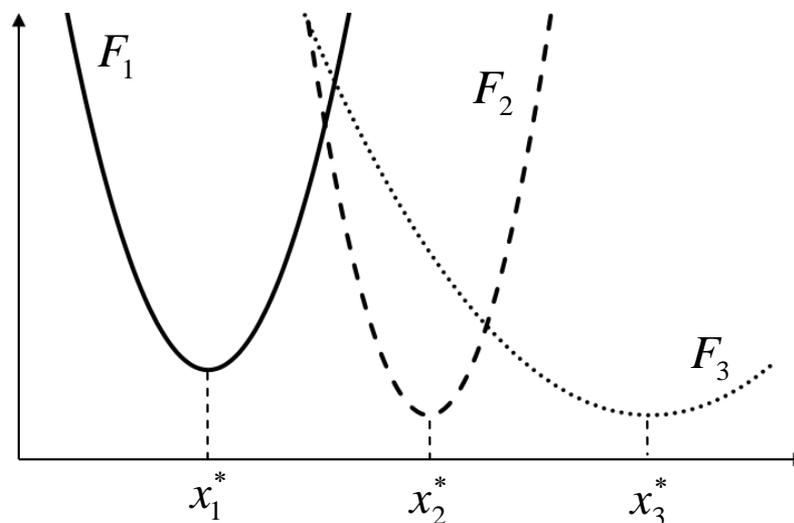
Dynamic optimization

Dynamic process: $\quad y_{n+1} = P(y_n, x_n)$ $\qquad n$ – time step

x_n – decision made at n-th time step

y_n – state of the process at n-th time step

The proble is to find optimal sequence of decisions:

$$x_0^*, x_1^*, \ldots, x_N^*,$$

for which $Q(x_0, x_1, \ldots, x_N)$ is minimal.

15 Saving Tips by
RateCatcher.com

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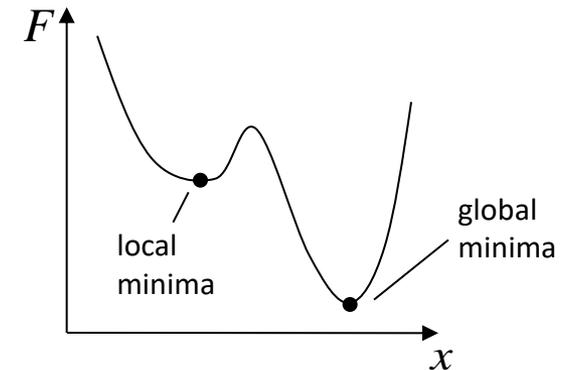
Mathematical preliminaries

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Mathematical preliminaries

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

Local minima: $\quad \forall_{\varepsilon > 0} \exists_{x \in O(x^*, \varepsilon)} F(x^*) < F(x)$

Global minima: $\forall_{x \in \mathcal{D}_x} F(x^*) < F(x)$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND
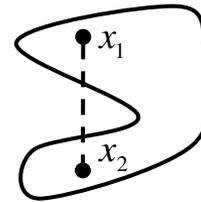
Project co-financed from the EU European Social Fund

Mathematical preliminaries

Convex set: $\forall_{x_1, x_2 \in \mathscr{D}_x} \lambda x_1 + (1-\lambda) x_2 \in \mathscr{D}_x, \quad \lambda \in <0,1>$
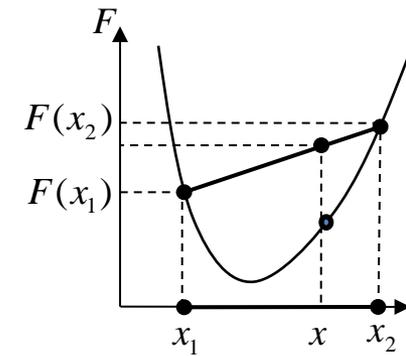

− convex set

− nonconvex set

Convex function:

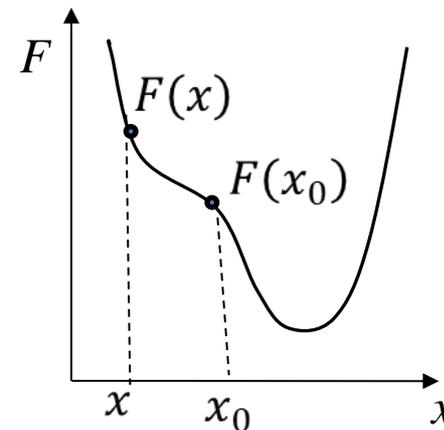$$F(\lambda x_1 + (1-\lambda) x_2) \le \lambda F(x_1) + (1-\lambda) F(x_2), \quad \lambda \in <0,1>$$

Mathematical preliminaries

Pseudo-convex function:

Following the Taylor's expansion of a function, we have:

$$F(x) = F(x_0) + (x - x_0)^T \left[\nabla_x F(x_0) \right] + O_2 \left(\| x - x_0 \| \right)$$

$$(x - x_0)^T \left[\nabla_x F(x_0) \right] \geq 0 \quad \Rightarrow \quad F(x) > F(x_0)$$

Quasi-convex function:

$$\mathscr{D}_\alpha = \left\{ x \in \mathscr{D}_x : F(x) \leq \alpha \right\} \quad \text{– convex sets}$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund
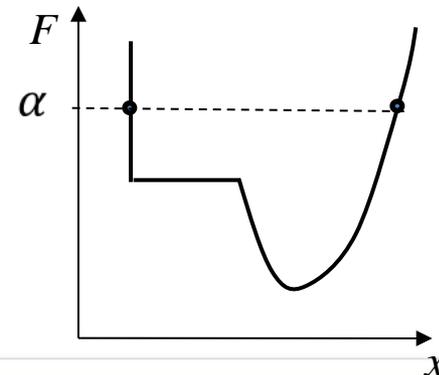
Mathematical preliminaries

Gradient:
$$\nabla_x F(x) = \begin{bmatrix} \dfrac{\partial F}{\partial x^{(1)}} \\ \dfrac{\partial F}{\partial x^{(2)}} \\ \vdots \\ \dfrac{\partial F}{\partial x^{(S)}} \end{bmatrix} = grad\ F(x)$$
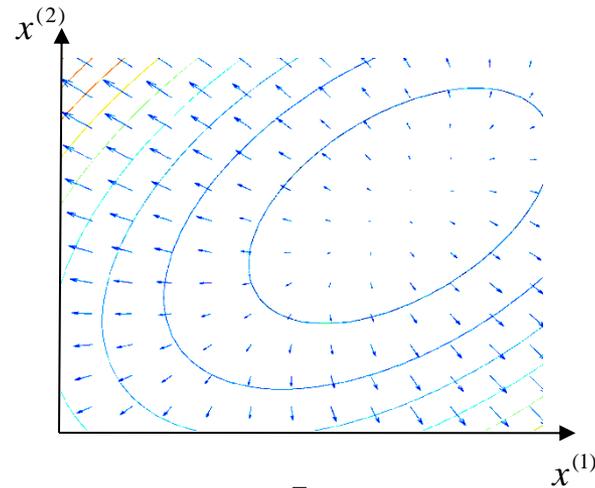
Hessian:
$$H(x) = \nabla_{xx}^2 F(x) = \begin{bmatrix} \dfrac{\partial^2 F}{\partial \left(x^{(1)}\right)^2} & \dfrac{\partial^2 F}{\partial x^{(1)} \partial x^{(2)}} & \cdots & \dfrac{\partial^2 F}{\partial x^{(1)} \partial x^{(S)}} \\ \dfrac{\partial^2 F}{\partial x^{(2)} \partial x^{(1)}} & \dfrac{\partial^2 F}{\partial \left(x^{(2)}\right)^2} & \cdots & \dfrac{\partial^2 F}{\partial x^{(2)} \partial x^{(S)}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 F}{\partial x^{(S)} \partial x^{(1)}} & \dfrac{\partial^2 F}{\partial x^{(S)} \partial x^{(2)}} & \cdots & \dfrac{\partial^2 F}{\partial \left(x^{(S)}\right)^2} \end{bmatrix}$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Mathematical preliminaries

Hessian properties:

$$\frac{\partial^2 F}{\partial x^{(i)} \partial x^{(j)}} = \frac{\partial^2 F}{\partial x^{(j)} \partial x^{(i)}} \Rightarrow H \quad \text{is symmetric matrix}$$

If $\forall_{x \neq 0_S} x^T H x > 0$ then H is positive definite

If $\forall_{x \neq 0_S} x^T H x < 0$ then H is negative definite

If $\forall_{x \neq 0_S} x^T H x \geq 0$ then H is positive semidefinite

If $\forall_{x \neq 0_S} x^T H x \leq 0$ then H is negative semidefinite

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Sylwester criteria:

$$H = \left[h_{ij}\right]_{\substack{i=1,2,\ldots,S \\ j=1,2,\ldots,S}}$$ - Hess matrix

If $\quad \forall\, s = 1, 2, \ldots, S \quad \det(H_{ss}) = \det\left(\left[h_{ij}\right]_{\substack{i=1,2,\ldots,s \\ j=1,2,\ldots,s}}\right) > 0 \quad$ then matrix H is positive definite

if $\quad \forall\, \{i_1, i_2, \cdots, i_s\} \in \{1, 2, \cdots, S\} \quad \det\left(\left[h_{ij}\right]_{\substack{i \in \{i_1,i_2,\cdots,i_s\} \\ j \in \{i_1,i_2,\cdots,i_s\}}}\right) \geq 0 \quad$ then matrix H is semipositive definite

Eigen values of matrix H

$$\det(H - hI) = 0 \quad h_1, h_2, \ldots, h_S \qquad$$ - Eigen values of matrix H

If $\quad \forall\, s = 1, 2, \ldots, S \quad h_s > 0 \quad$ then matrix H is positive definite

If $\quad \forall\, s = 1, 2, \ldots, S \quad h_s \geq 0 \quad$ then matrix H is semipositive definite

$$\forall\, s = 1, 2, \ldots, S \quad \det(H_{ss}) = \det\left(\left[h_{ij}\right]_{\substack{i=1,2,\ldots,s \\ j=1,2,\ldots,s}}\right) > 0$$

$$H = \left[h_{ij}\right]_{\substack{i=1,2,\cdots,S \\ i=1,2,\cdots,S}} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & \cdots & h_{1S} \\ h_{21} & h_{22} & h_{23} & \cdots & h_{2S} \\ h_{31} & h_{32} & h_{33} & \cdots & h_{3S} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{S1} & h_{S2} & h_{S3} & \cdots & h_{SS} \end{bmatrix}$$

$$\forall \left\{i_1, i_2, \cdots, i_s\right\} \in \left\{1, 2, \cdots, S\right\} \quad \det\left(\left[h_{ij}\right]_{\substack{i \in \{i_1, i_2, \cdots, i_s\} \\ j \in \{i_1, i_2, \cdots, i_s\}}}\right) \geq 0$$

For example $\{i_1, i_2, i_3\} = \{1, 3, S\}$

$$det\left(\left[h_{ij}\right]_{\substack{i=1,3,S \\ i=1,3,S}}\right) = det\begin{bmatrix} h_{11} & h_{13} & h_{1S} \\ h_{31} & h_{33} & h_{3S} \\ h_{S1} & h_{S3} & h_{SS} \end{bmatrix} \geq 0$$

$$H = \left[h_{ij}\right]_{\substack{i=1,2,\cdots,S \\ i=1,2,\cdots,S}} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & \cdots & h_{1S} \\ h_{21} & h_{22} & h_{23} & \cdots & h_{2S} \\ h_{31} & h_{32} & h_{33} & \cdots & h_{3S} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{S1} & h_{S2} & h_{S3} & \cdots & h_{SS} \end{bmatrix}$$

$$\begin{bmatrix} \boldsymbol{h_{11}} & h_{12} & \boldsymbol{h_{13}} & \cdots & \boldsymbol{h_{1S}} \\ h_{21} & h_{22} & h_{23} & \cdots & h_{2S} \\ \boldsymbol{h_{31}} & h_{32} & \boldsymbol{h_{33}} & \cdots & \boldsymbol{h_{3S}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{h_{S1}} & h_{S2} & \boldsymbol{h_{S3}} & \cdots & \boldsymbol{h_{SS}} \end{bmatrix}$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

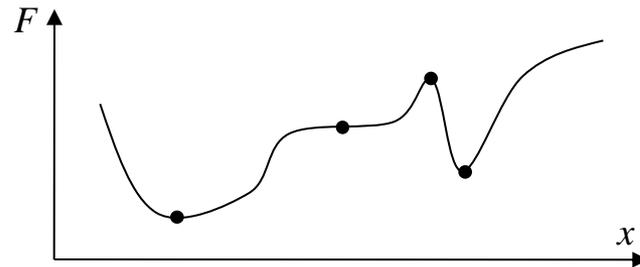Project co-financed from the EU European Social Fund

Unconstrained optimization

Optimization task: $\quad x^* \to F(x^*) = \min\limits_{x^* \in \mathcal{D}_x} F(x)$

Assumption: $F(x)$ is continuous and differentiable.

Necessary condition for x^* to be local minima: $\quad \nabla_x F(x^*) = 0_S$

If $F(x)$ is convex function, then above equation is sufficient condition for x^* to be global minima.

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Unconstrained optimization

Second order conditions of optimality:

If $H(x^*)$ is positive definite at x^* then x^* is local minimum.

If $H(x^*)$ is negative definite at x^* then x^* is local maximum.

If $H(x^*)$ is neither negative semidefinite nor positive semidefinite at x^* then x^* is not optimum.

If $H(x^*)$ is positive (negative) semidefinite and not positive (negative) definite, optimality of x^* cannot be determined.

Example 2.1.1

- $F\big(x^{(1)}, x^{(2)}\big) = 5\big(x^{(1)}\big)^2 + \big(x^{(2)}\big)^2 - 4x^{(1)}x^{(2)} - 2x^{(1)} + 3$

- $\nabla_x F\big(x^{(1)}, x^{(2)}\big)\big|_{x=x^*} = \begin{bmatrix} \dfrac{\partial F(x^{(1)},x^{(2)})}{\partial x^{(1)}} \\ \dfrac{\partial F(x^{(1)},x^{(2)})}{\partial x^{(2)}} \end{bmatrix}\Bigg|_{x=x^*} = \begin{bmatrix} 10x^{(1)*} - 4x^{(2)*} - 2 \\ 2x^{(2)*} - 4x^{(1)*} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ (1) (2)

- z (2) $\rightarrow x^{(2)*} = 2x^{(1)*}$

- z (1) $\rightarrow 10x^{(1)*} - 8x^{(1)*} = 2 \rightarrow x^{(1)*} = 1, x^{(2)*} = 2$ $x^* = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

- $H(x) = \nabla_{xx} F\big(x^{(1)}, x^{(2)}\big) = \begin{bmatrix} 10 & -4 \\ -4 & 2 \end{bmatrix}$

- $det H_{11} = det[10] = 10 > 0$

- $det H_{11} = det \begin{bmatrix} 10 & 4 \\ 4 & 2 \end{bmatrix} = 20 - 16 = 4 > 0$

Matrix   H(x)   is positively defined   then point X* is   minimum

Example 2.1.2

- $F\left(x^{(1)}, x^{(2)}\right) = \alpha\left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2 - 4x^{(1)}x^{(2)} - 2x^{(1)} + 3$

- $\nabla_x F\left(x^{(1)}, x^{(2)}\right)\Big|_{x=x^*} = \begin{bmatrix} \dfrac{\partial F(x^{(1)}, x^{(2)})}{\partial x^{(1)}} \\ \dfrac{\partial F(x^{(1)}, x^{(2)})}{\partial x^{(2)}} \end{bmatrix}\Bigg|_{x=x^*} = \begin{bmatrix} 2\alpha x^{(1)*} - 4x^{(2)*} - 2 \\ 2x^{(2)*} - 4x^{(1)*} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ (1) (2)

- z (2) $\rightarrow x^{(2)*} = 2x^{(1)*}$

- z (1) $\rightarrow 2\alpha x^{(1)*} - 8x^{(1)*} = 2 \rightarrow x^{(1)*} = \dfrac{1}{\alpha-4}, x^{(2)*} = \dfrac{2}{\alpha-4}, \alpha \neq 4$ $x^* = \begin{bmatrix} \dfrac{1}{\alpha-4} \\ \dfrac{2}{\alpha-4} \end{bmatrix}$

- $H(x) = \nabla_{xx} F\left(x^{(1)}, x^{(2)}\right) = \begin{bmatrix} 2\alpha & 4 \\ 4 & 2 \end{bmatrix}$

- $det H_{11} = det[2\alpha] = 2\alpha > 0 \rightarrow \alpha > 0$

- $det H_{11} = det \begin{bmatrix} 2\alpha & 4 \\ 4 & 2 \end{bmatrix} = 4\alpha - 16 > 0 \rightarrow \alpha > 4$

For $\alpha > 4$ matrix $H(x)$ is positively defined then point $\begin{bmatrix} \dfrac{1}{\alpha-4} \\ \dfrac{2}{\alpha-4} \end{bmatrix}$ - minimum

$x^* = \begin{vmatrix} \dfrac{1}{\alpha-4} \\ \dfrac{2}{\alpha-4} \end{vmatrix}$

Example 2.1.2 c.d.

$For \ \alpha > 4 \ matrix \ H(x) \ is \ positively \ defined$

$$Point \ x^* = \begin{bmatrix} \dfrac{1}{\alpha-4} \\ \dfrac{2}{\alpha-4} \end{bmatrix} \text{ is minimum } (\alpha \neq 4)$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Example 2.1.3

෫ $F(x) = x^T A x + b^T x + c$

A – symmetric matrix, positive define

෫ $A = \begin{bmatrix} a_{ij} \end{bmatrix}_{\substack{i=1,2,\cdots,S \\ i=1,2,\cdots,S}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1S} \\ a_{21} & a_{22} & \cdots & a_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ a_{S1} & a_{S2} & \cdots & a_{SS} \end{bmatrix}$

෫ $x = \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(S)} \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_S \end{bmatrix},$ S – dimensional vectors

෫ $\nabla_x F(x)|_{x=x^*} = \nabla_x (x^T A x + b^T x + c)|_{x=x^*} = 0_S$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Example 2.1.3 c.d.

$$\text{\ding{108}} \quad x^T A x = \begin{bmatrix} x^{(1)} & x^{(2)} & \cdots & x^{(S)} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1S} \\ a_{21} & a_{22} & \cdots & a_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ a_{S1} & a_{S2} & \cdots & a_{SS} \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(S)} \end{bmatrix}$$

$$\text{\ding{108}} \quad x^T A x = \sum_{i=1}^{S} \sum_{j=1}^{S} a_{ij}\, x^{(i)} x^{(j)}$$

$$\text{\ding{108}} \quad \nabla_x (x^T A x) = \begin{bmatrix} \dfrac{\partial (x^T A x)}{\partial x^{(1)}} \\ \dfrac{\partial (x^T A x)}{\partial x^{(2)}} \\ \vdots \\ \dfrac{\partial (x^T A x)}{\partial x^{(S)}} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial}{x^{(1)}} \left(\sum_{i=1}^{S} \sum_{j=1}^{S} a_{ij}\, x^{(i)} x^{(j)} \right) \\ \dfrac{\partial}{x^{(2)}} \left(\sum_{i=1}^{S} \sum_{j=1}^{S} a_{ij}\, x^{(i)} x^{(j)} \right) \\ \vdots \\ \dfrac{\partial}{x^{(S)}} \left(\sum_{i=1}^{S} \sum_{j=1}^{S} a_{ij}\, x^{(i)} x^{(j)} \right) \end{bmatrix}$$

Example 2.1.3 c.d.

$$\mathcal{E}\, \nabla_x(x^T A x) = \begin{bmatrix} \frac{\partial}{\partial x^{(1)}} \left(\sum_{i=1}^{S} \sum_{j=1}^{S} a_{ij}\, x^{(i)} x^{(j)} \right) \\ \frac{\partial}{\partial x^{(2)}} \left(\sum_{i=1}^{S} \sum_{j=1}^{S} a_{ij}\, x^{(i)} x^{(j)} \right) \\ \vdots \\ \frac{\partial}{\partial x^{(S)}} \left(\sum_{i=1}^{S} \sum_{j=1}^{S} a_{ij}\, x^{(i)} x^{(j)} \right) \end{bmatrix}$$

$$\mathcal{E}\, \nabla_x(x^T A x) = \begin{bmatrix} \sum_{j=1}^{S} a_{1j} x^{(j)} + \sum_{i=1}^{S} a_{i1} x^{(i)} \\ \sum_{j=1}^{S} a_{2j} x^{(j)} + \sum_{i=1}^{S} a_{i2} x^{(i)} \\ \vdots \\ \sum^{S} \quad x^{(i)} + \sum^{S} \quad x^{(i)} \end{bmatrix} =$$

Example 2.1.3 c.d.

$$
\text{\Large ☙} \quad \nabla_x(x^T A x) = \begin{bmatrix} \sum_{j=1}^{S} a_{1j} x^{(j)} \\ \sum_{j=1}^{S} a_{2j} x^{(j)} \\ \vdots \\ \sum_{j=1}^{S} a_{Sj} x^{(j)} \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^{S} a_{i1} x^{(i)} \\ \sum_{i=1}^{S} a_{i2} x^{(i)} \\ \vdots \\ \sum_{i=1}^{S} a_{S1} x^{(i)} \end{bmatrix}
$$

$$
\nabla_x(x^T A x) =
$$

$$
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1S} \\ a_{21} & a_{22} & \cdots & a_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ a_{S1} & a_{S2} & \cdots & a_{SS} \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(S)} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{S1} \\ a_{12} & a_{22} & \cdots & a_{S2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{S1} & a_{S2} & \cdots & a_{SS} \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(S)} \end{bmatrix}
$$

$$
\nabla_x(x^T A x) = A x + A^T x
$$

$$
A = A^T \qquad \nabla_x(x^T A x) = 2 A x
$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Example 2.1.3 c.d.

$$\text{\bf{\&}}\quad b^T x = [b_1 \quad b_2 \quad \cdots \quad b_S] \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(S)} \end{bmatrix} = \sum_{i=1}^{S} b_i x^{(j)}$$

$$\text{\bf{\&}}\quad \nabla_x (b^T x) = \begin{bmatrix} \frac{\partial}{\partial x^{(1)}} \left(\sum_{i=1}^{S} b_i x^{(j)} \right) \\ \frac{\partial}{\partial x^{(2)}} \left(\sum_{i=1}^{S} b_i x^{(j)} \right) \\ \vdots \\ \frac{\partial}{\partial x^{(S)}} \left(\sum_{i=1}^{S} b_i x^{(j)} \right) \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_S \end{bmatrix} = b$$

Example 2.1.3 c.d.

- $\nabla_x (x^T A x + b^T x + c)|_{x=x^*} = 2Ax^* + \mathbf{b} = 0_S$

- $x^* = -\frac{1}{2} A^{-1} b$

- $H(x) = \nabla_{xx}(x^T A x + b^T x + c) = \nabla_x(2Ax + \mathbf{b}) = 2A$

Hess matrix is positively defined because
matrix A is assumed to be positively defined

General classification of optimization tasks

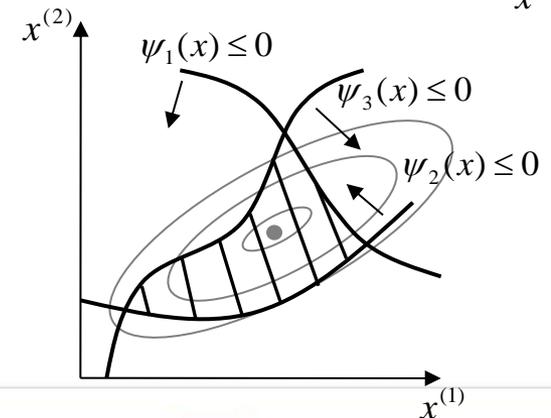Unconstrained optimization:

$$\mathscr{D}_x = \mathscr{R}^S$$

Optimization under equality constraints:

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

Optimization under inequality constraints:

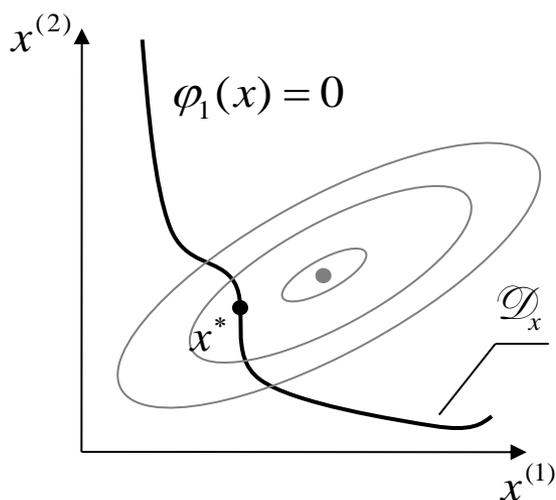$$\mathscr{D}_x = \left\{ x \in \mathscr{R}^S : \psi_1(x) \le 0, \psi_2(x) \le 0, \ldots, \psi_M(x) \le 0 \right\}$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Analytical methods

- Unconstrained optimization

- Lagrange multipliers method – equality constraints

- Kuhn-Tucker conditions – inequality constraints
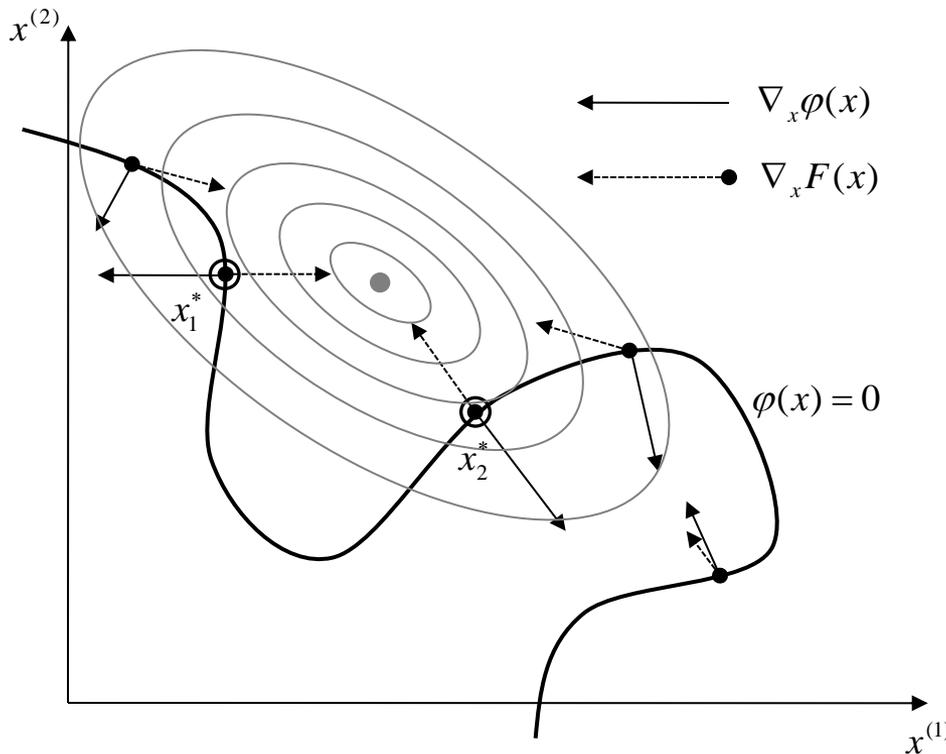
Optimization under equality constraints

Optimization task: $x^* \to F(x^*) = \min\limits_{x^* \in \mathscr{D}_x} F(x)$

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

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Optimization under equality constraints

$$\nabla_x \varphi(x)$$

$$\nabla_x F(x)$$

$$\varphi(x) = 0$$

Locally optimal solution satisfies condition:

$$\nabla_x F(x) + \lambda \nabla_x \varphi(x) = 0_S$$

where

$$\lambda \in \mathscr{R} \quad \text{– Lagrange multiplier}$$

For multiple constraints:

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

Optimization problem under equality constraints
Lagrange' a multiplayers metod

The above system of equation may have several solutions

Second order nesesery conditions:

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

If $\quad H_L(x^*) \quad$ Is positively defined in the point $\quad x^*$

then $\quad x^*$ is local minimum

If $\quad H_L(x^*) \quad$ Is negatively defined in the point $\quad x^*$

then $\quad x^*$ is local minimum

If $F(x)$ is convex function, and constrains are linear one i.e. have the form
$\varphi_l(x) = p_l^T x - \alpha_l = 0, \quad l = 1, 2, \ldots, L$ then the above system of equation
have one solution and it is optimsl point

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Explanation of necessary conditions

❀ $x_1^*, x_2^* \rightarrow F(x_1^*, x_2^*) = \min_{x_1, x_2} F(x_1, x_2)$

With constans

$$\varphi(x_1, x_2) = 0$$

❀ $\varphi(x_1, x_2) = 0 \rightarrow x_2 = \psi(x_1)$

❀ $x_1^* \rightarrow F\big(x_1^*, \psi(x_1^*)\big) = \min_{x_1} F\big(x_1 \psi(x_1)\big)$

❀ $\dfrac{dF(x_1 \psi(x_1))}{\partial x_1} = \dfrac{\partial F(x_1, x_2)}{\partial x_1} + \dfrac{\partial F(x_1, x_2)}{\partial x_2} \dfrac{d\psi(x_1)}{dx_1} = 0$

❀ $\dfrac{d\psi(x_1)}{\partial x_1} = -\dfrac{\frac{\partial \varphi(x_1, x_2)}{\partial x_1}}{\frac{\partial \varphi(x_1, x_2)}{\partial x_2}}$ Derived from the unraveling function

Explanation of necessary conditions

- $$\frac{\partial F(x_1,x_2)}{\partial x_1}+\frac{\partial F(x_1,x_2)}{\partial x_2}\left(-\frac{\frac{\partial \varphi(x_1,x_2)}{\partial x_1}}{\frac{\partial \varphi(x_1,x_2)}{\partial x_2}}\right)=0$$

let
$$\lambda=-\frac{\frac{\partial F(x_1,x_2)}{\partial x_2}}{\frac{\partial \varphi(x_1,x_2)}{\partial x_2}}$$

- $$\frac{\partial F(x_1,x_2)}{\partial x_1}+\lambda\frac{\partial \varphi(x_1,x_2)}{\partial x_1}=0$$

- $$\frac{\partial F(x_1,x_2)}{\partial x_2}+\lambda\frac{\partial \varphi(x_1,x_2)}{\partial x_2}=0$$

- $$x_2=\psi(x_1)\rightarrow\varphi(x_1,x_2)=0$$

Explanation of necessary conditions

- $L(x_1, x_2, \lambda) = F(x_1, x_2) + \lambda\varphi(x_1, x_2)$

- $\dfrac{\partial L(x_1,x_2,\lambda)}{\partial x_1} = 0 \rightarrow \dfrac{\partial F(x_1,x_2)}{\partial x_1} + \lambda\dfrac{\partial \varphi(x_1,x_2)}{\partial x_1} = 0$

- $\dfrac{\partial L(x_1,x_2,\lambda)}{\partial x_2} = 0 \rightarrow \dfrac{\partial F(x_1,x_2)}{\partial x_2} + \lambda\dfrac{\partial \varphi(x_1,x_2)}{\partial x_2} = 0$

- $\dfrac{\partial L(x_1,x_2,\lambda)}{\partial \lambda} = 0 \rightarrow \varphi(x_1, x_2) = 0$

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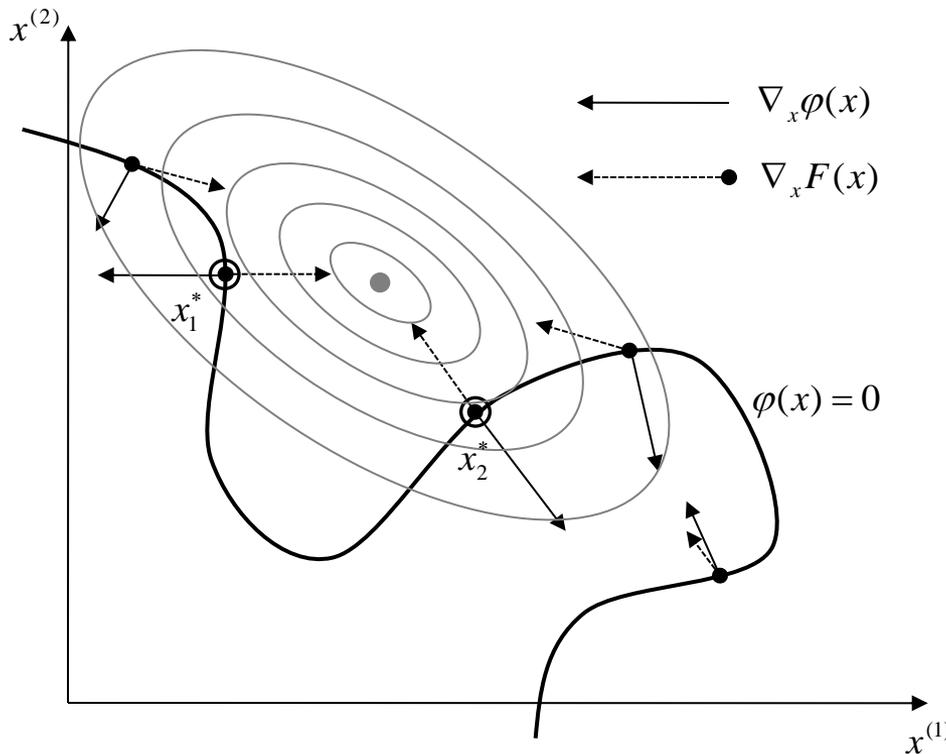
$$\nabla_x L(x, \lambda)\big|_{x^*, \lambda^*} = 0_S$$

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

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Optimization under equality constraints

Locally optimal solution satisfies condition:

$$\nabla_x F(x) + \lambda \nabla_x \varphi(x) = 0_S$$

where

$\lambda \in \mathscr{R}$ – Lagrange multiplier

For multiple constraints:

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

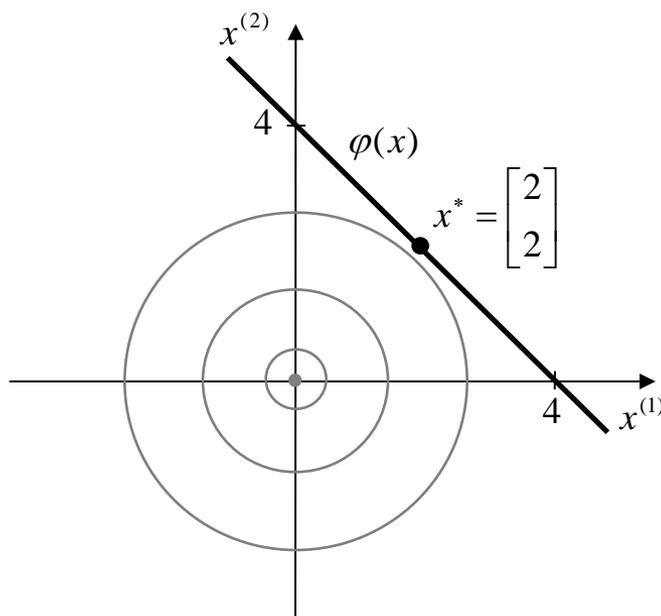$$\nabla_x L(x, \lambda)\big|_{x^*, \lambda^*} = 0_S$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Optimization under equality constraints

- The method of Lagrange multipliers – example 1

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

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

$$x^{(2)}$$

$$4$$

$$\varphi(x)$$

$$x^* = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$4 \quad x^{(1)}$$

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

Example 2.2.1

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

- $\nabla_x L(x,\lambda) = \begin{bmatrix} \frac{\partial L}{\partial x^{(1)}} \\ \frac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2x^{(1)} + \lambda \\ 2x^{(2)} + \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ 　(1) (2)

- $\nabla_\lambda L(x,\lambda) = \frac{\partial L}{\partial \lambda} = x^{(1)} + x^{(2)} - 4 = 0$ 　(3)

- z (1) $\rightarrow x^{(1)} = -\frac{\lambda}{2}$, 　z (2) $\rightarrow x^{(2)} = -\frac{\lambda}{2}$

- z (3) $\rightarrow \left(-\frac{\lambda}{2}\right) + \left(-\frac{\lambda}{2}\right) - 4 = 0$ czyli $\lambda = -4$

- $x^{(1)} = -\frac{\lambda}{2} = -\frac{-4}{2} = 2$, 　$x^{(2)} = -\frac{\lambda}{2} = -\frac{-4}{2} = 2$

Example 2.2.1 c.d.

∞ $\nabla_x \mathrm{L}(x, \lambda) = \begin{bmatrix} \dfrac{\partial L}{\partial x^{(1)}} \\ \dfrac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2x^{(1)} + \lambda \\ 2x^{(2)} + \lambda \end{bmatrix}$

∞ $H_L = \nabla_{xx} \mathrm{L}(x, \lambda) = \begin{bmatrix} \dfrac{\partial^2 L}{\partial^2 x^{(1)}} & \dfrac{\partial^2 L}{\partial x^{(1)} \partial x^{(2)}} \\ \dfrac{\partial^2 L}{\partial x^{(2)} \partial x^{(1)}} & \dfrac{\partial^2 L}{\partial^2 x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

∞ $det\ H_{L11} = det[2] = 2 > 0, \quad det\ H_{L22} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 2 \times 2 = 4 > 0$

Matrix   Is positively defined

$$H_L = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Point   is minimum
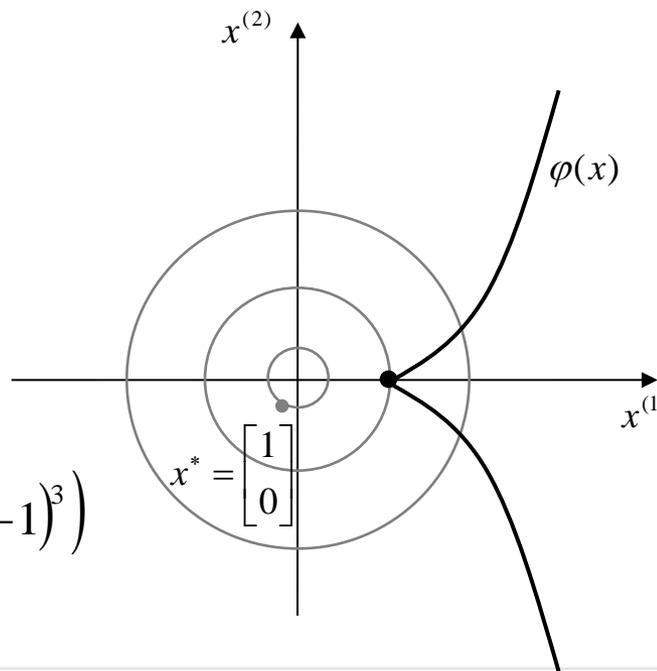
$x = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

Optimization under equality constraints

- The method of Lagrange multipliers – example 2 (irregular)

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

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

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

$$x^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Example 2.2.2

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

ɞ $\nabla_x L(x, \lambda) = \begin{bmatrix} \frac{\partial L}{\partial x^{(1)}} \\ \frac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2x^{(1)} - 3\lambda\left(x^{(1)} - 1\right)^2 \\ 2x^{(2)} + 2\lambda x^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ \quad (1) \quad (2)

ɞ $\nabla_\lambda L(x, \lambda) = \frac{\partial L}{\partial \lambda} = \left(x^{(2)}\right)^2 - \left(x^{(1)} - 1\right)^3 = 0$ \hfill (3)

ɞ z (2) $\rightarrow 2(1 + \lambda)x^{(2)} = 0$ czyli $x^{(2)} = 0$,

ɞ z (3) $\rightarrow (0)^2 - \left(x^{(1)} - 1\right)^3 = 0$, czyli $x^{(1)} = 1$,

ɞ z (1) $\rightarrow 2x^{(1)} - 3\lambda\left(x^{(1)} - 1\right)^2 = 2 \times 1 - 3\lambda(1 - 1)^2 = 2 \neq 0$

Contradiction

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Optimization under equality constraints

- The method of Lagrange multipliers – example 2 explanation

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

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

$$\nabla_x F(x) + G(x)\lambda = 0 \qquad G(x)\lambda = -\nabla_x F(x)$$

Unambiguous solution exists if and only if $\operatorname{rank} G(x) = \operatorname{rank} \left[G(x) \quad \vdots \quad -\nabla_x F(x) \right]$,

which is always true as long as F is convex and φ_l are linear.

How to find irregular solutions?

Optimization under equality constraints
Lagrange' a multipliers method

If F(x) is continuous, differentiable and convex function and constraints
$\varphi_1(x), \varphi_2(x), \ldots, \varphi_L(x)$ are linear then system of equations:

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

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

has one solution and it si solution of optimization task.

The above system of equations is necessary and sufficient condition for optimal solution

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Optimization under equality constraints

- The generalized method of Lagrange multipliers

Generalized Lagrange function:

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

Necessary conditions of optimality:

$$\nabla_x L(x, \lambda, \lambda_0)\big|_{x^*, \lambda^*, \lambda_0} = 0_S$$

$$\nabla_\lambda L(x, \lambda, \lambda_0)\big|_{x^*, \lambda^*, \lambda_0} = 0_L$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Optimization under equality constraints

- The generalized method of Lagrange multipliers

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

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

$\qquad\qquad \lambda_0 = 1 \qquad \nabla_x F(x) + \sum_{l=1}^{L} \lambda_l \nabla_x \varphi_l(x) = 0_S \qquad$ We obtain regular solutions.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ We obtain irregular solutions.

$2^o \qquad \lambda_0 = 0 \qquad \sum_{l=1}^{L} \lambda_l \nabla_x \varphi_l(x) = 0_S$

Second order condition of optimality requires analysis of $H(x, \lambda, \lambda_0) = \nabla_{xx}^2 L(x, \lambda, \lambda_0)$.

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Optimization under equality constraints

- The generalized method of Lagrange multipliers –
  – example 2 once again

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

$$\varphi(x) = \left(x^{(2)}\right)^2 - \left(x^{(1)} - 1\right)^3 = 0$$
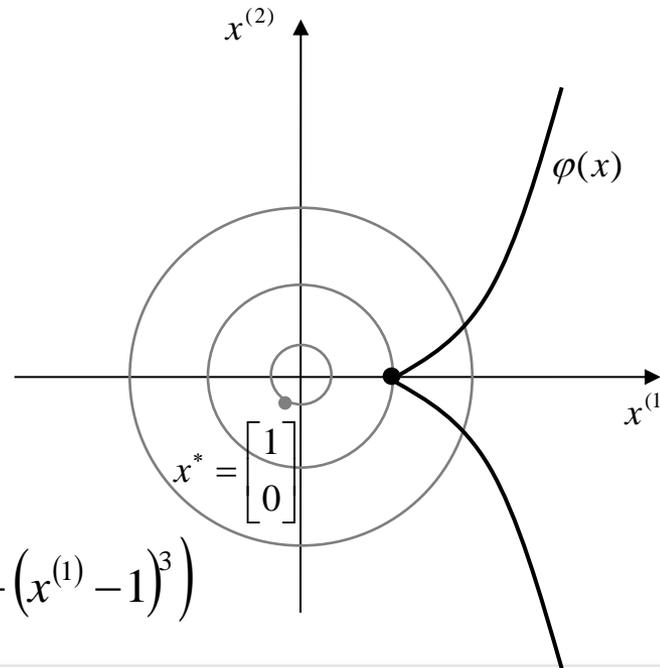
$$L(x, \lambda, \lambda_0) = \lambda_0\left(\left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2\right) + \lambda\left(\left(x^{(2)}\right)^2 - \left(x^{(1)} - 1\right)^3\right)$$

Example 2.2.2

∞ $L(x, \lambda) = \lambda_0 \left(\left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2 \right) + \lambda \left(\left(x^{(2)}\right)^2 - \left(x^{(1)} - 1\right)^3 \right)$

∞ $\nabla_x L(x, \lambda) = \begin{bmatrix} \frac{\partial L}{\partial x^{(1)}} \\ \frac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2\lambda_0 x^{(1)} - 3\lambda \left(x^{(1)} - 1\right)^2 \\ 2\lambda_0 x^{(2)} + 2\lambda x^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ (1)
                                                                                                                                                (2)

∞ $\nabla_\lambda L(x, \lambda) = \frac{\partial L}{\partial \lambda} = \left(x^{(2)}\right)^2 - \left(x^{(1)} - 1\right)^3 = 0$ (3)

∞ Dla $\lambda_0 = 1$

For

∞ $\nabla_x L(x, \lambda) = \begin{bmatrix} \frac{\partial L}{\partial x^{(1)}} \\ \frac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2x^{(1)} - 3\lambda \left(x^{(1)} - 1\right)^2 \\ 2x^{(2)} + 2\lambda x^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ (1)
                                                                                                                                    (2)

As before contradiction

Example 2.2.2

$$\text{L}(x,\lambda) = \lambda_0 \left(\left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2 \right) + \lambda \left(\left(x^{(2)}\right)^2 - \left(x^{(1)} - 1\right)^3 \right)$$

$$\nabla_x \text{L}(x,\lambda) = \begin{bmatrix} \dfrac{\partial L}{\partial x^{(1)}} \\ \dfrac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2\lambda_0 x^{(1)} - 3\lambda\left(x^{(1)} - 1\right)^2 \\ 2\lambda_0 x^{(2)} + 2\lambda x^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \begin{matrix} (1) \\ (2) \end{matrix}$$

$$\nabla_\lambda \text{L}(x,\lambda) = \frac{\partial L}{\partial \lambda} = \left(x^{(2)}\right)^2 - \left(x^{(1)} - 1\right)^3 = 0 \qquad (3)$$

Dla $\lambda_0 = 0$

For

$$\nabla_x \text{L}(x,\lambda) = \begin{bmatrix} \dfrac{\partial L}{\partial x^{(1)}} \\ \dfrac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 3\lambda\left(x^{(1)} - 1\right)^2 \\ \lambda x^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \begin{matrix} (1) \\ (2) \end{matrix}$$

$\text{z }(1) \rightarrow 3\lambda\left(x^{(1)} - 1\right)^2 = 0$ czyli $x^{(1)} = 1$, $\text{z }(2) \rightarrow x^{(2)} = 0$

Example 2.2.2

❧ $\nabla_x \mathrm{L}(x,\lambda) = \begin{bmatrix} \frac{\partial L}{\partial x^{(1)}} \\ \frac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2\lambda_0 x^{(1)} - 3\lambda\left(x^{(1)} - 1\right)^2 \\ 2\lambda_0 x^{(2)} + 2\lambda x^{(2)} \end{bmatrix}$

❧ $H_L = \nabla_{xx}\,\mathrm{L}(x,\lambda) = \begin{bmatrix} \frac{\partial^2 L}{\partial^2 x^{(1)}} & \frac{\partial^2 L}{\partial x^{(1)}\partial x^{(2)}} \\ \frac{\partial^2 L}{\partial x^{(2)}\partial x^{(1)}} & \frac{\partial^2 L}{\partial^2 x^{(2)}} \end{bmatrix} =$

❧ $\begin{bmatrix} 2\lambda_0 - 6\lambda\left(x^{(1)} - 1\right) & 0 \\ 0 & 2\lambda_0 + 2\lambda \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2\lambda \end{bmatrix}$

Matrix H_L is semi positively defined then point $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ - minimum

General classification of optimization tasks

Unconstrained optimization: $\mathscr{D}_x = \mathscr{R}^S$

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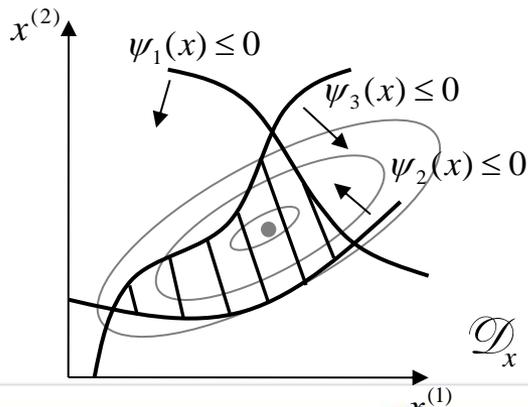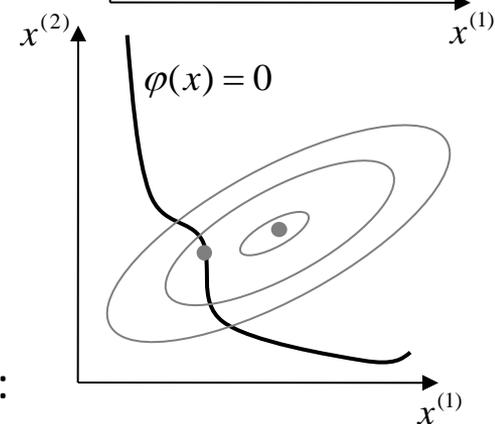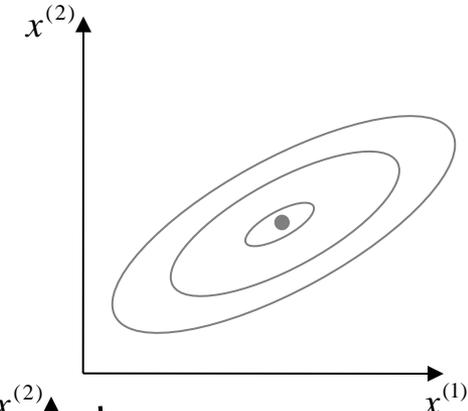
$$\mathscr{D}_x = \left\{ x \in \mathscr{R}^S : \varphi_1(x) = 0, \varphi_2(x) = 0, \ldots, \varphi_L(x) = 0, L \le S \right\}$$

Optimization under inequality constraints:

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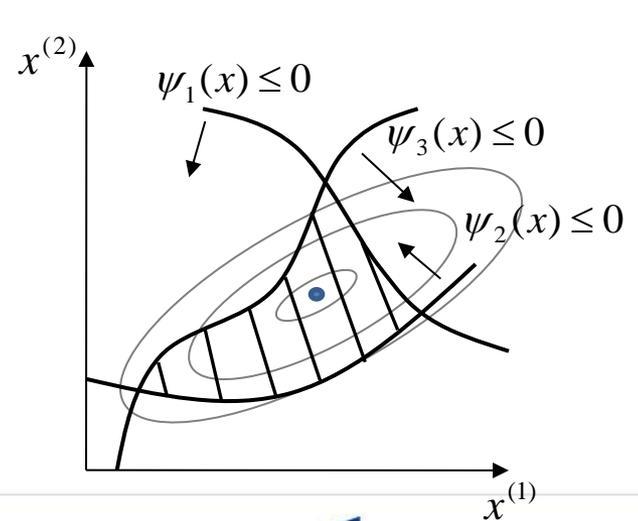
Optimization under inequality constraints

Optimization task: $\quad x^* \rightarrow F(x^*) = \min\limits_{x^* \in \mathcal{D}_x} F(x)$

$$\mathcal{D}_x = \left\{ x \in \mathcal{R}^S : \psi_1(x) \leq 0, \psi_2(x) \leq 0, \ldots, \psi_M(x) \leq 0 \right\}$$
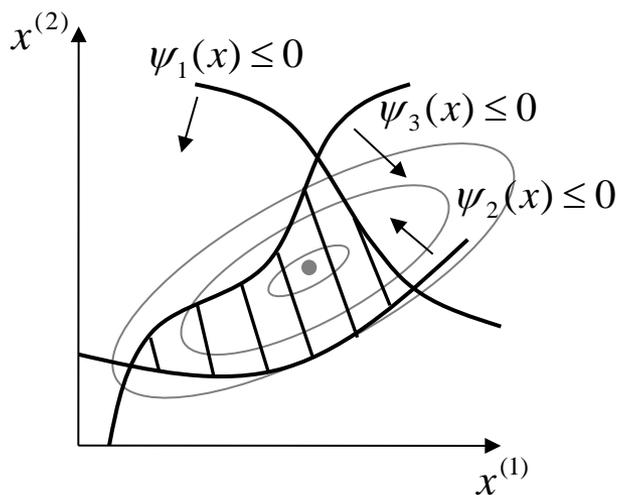
$$\psi(x) = \begin{bmatrix} \psi_1(x) \\ \psi_2(x) \\ \vdots \\ \psi_M(x) \end{bmatrix} \quad 0_M = \left.\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right\} M \qquad \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_S \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_S \end{bmatrix}$$

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

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

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Optimization under inequality constraints

Optimization task

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

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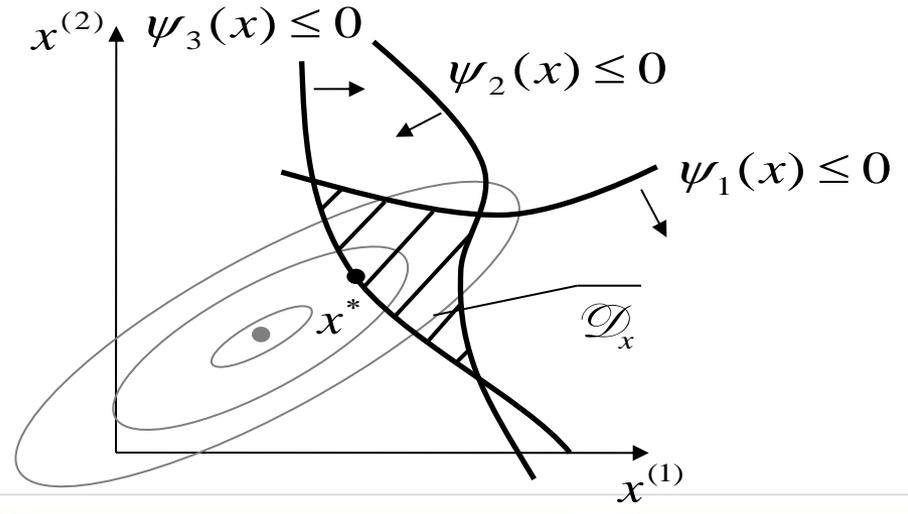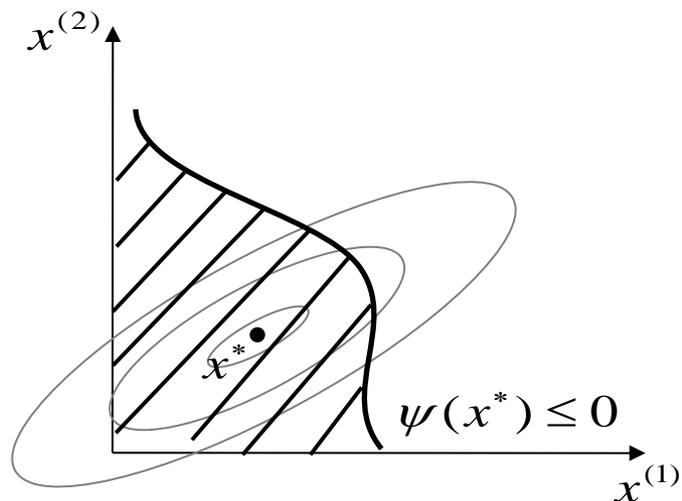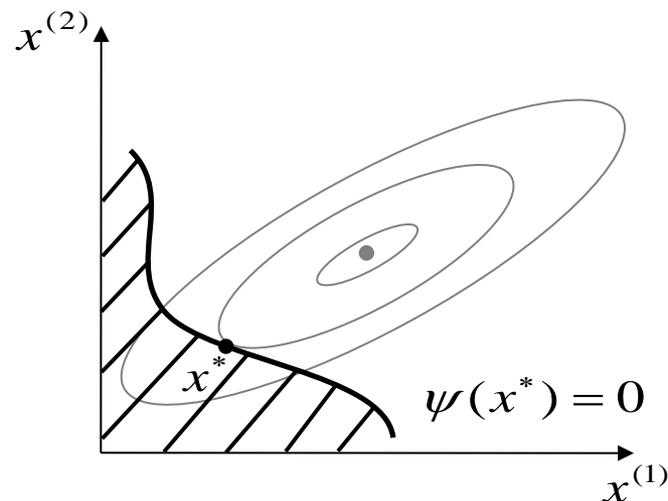
Optimization under inequality constraints

Inactive constraint

$$\psi(x^*) < 0$$

Active constraint

$$\psi(x^*) = 0$$

Optimization under inequality constraints

Lagrange function:

Kuhn-Tucker conditions

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

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_M \end{bmatrix}$$

Necessary conditions of optimality:

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

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

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

$$\mu^* \geq 0_M$$

If solution is regular

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

Optimization under inequality constraints

Kuhn-Tucker conditions

$$\nabla_x L(x,\mu) = \nabla_x F(x) + \sum_{m=1}^{M} \mu_m \nabla_x \psi_m(x) = 0_S$$

$$\mu^T \nabla_\mu L(x,\mu) = \mu^T \psi(x) = \sum_{m=1}^{M} \mu_m \psi_m(x) = 0$$

$$\nabla_\mu L(x,\mu) = \psi(x) \le 0_M$$

$$\mu \ge 0_M$$

$$\mu_1 \psi_1(x) + \mu_2 \psi_2(x) + \ldots + \mu_M \psi_M(x) = 0$$

$$\forall_m \psi_m(x) \le 0 \qquad \forall_m \mu_m \ge 0$$

$$\forall_m \mu_m \psi_m(x) = 0$$

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

$$\mu_m \psi_m(x)\big|_{x^*,\mu^*} = 0 \qquad m = 1, 2, \ldots, M$$

$$\psi_m(x)\big|_{x^*,\mu^*} \le 0 \qquad m = 1, 2, \ldots, M$$

$$\mu_m^* \ge 0 \qquad m = 1, 2, \ldots, M$$

Optimization under inequality constraints

Kuhn-Tucker conditions

$$L(x,\mu) = F(x) + \mu_m \psi_m(x)$$

$$\nabla_x L(x,\mu) = \nabla_x F(x) + \mu_m \nabla_x \psi_m(x) = 0_S$$

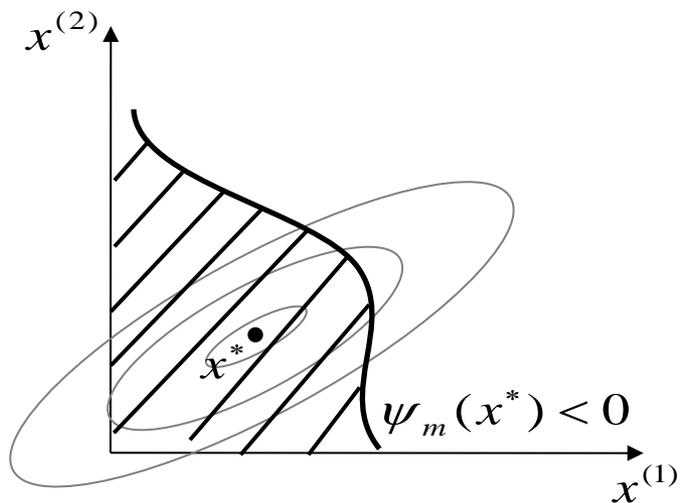$$\mu^T \nabla_\mu L(x,\mu) = \mu_m \psi_m(x) = 0$$

$$\nabla_\mu L(x,\mu) = \psi_m(x) \le 0$$

$$\mu_m \ge 0$$

m - th inactive constraints

$$\mu_m = 0 \quad \psi_m(x) < 0$$

$$\nabla_x L(x,\mu) = \nabla_x F(x) = 0_S$$

$$\nabla_\mu L(x,\mu) = \psi_m(x) < 0$$

Like without constraints

m - th active constraints

$$\mu_m > 0 \quad \psi_m(x) = 0$$

$$\nabla_x L(x,\mu) = \nabla_x F(x) + \mu_m \nabla_x \psi_m(x) = 0_S$$

$$\nabla_\mu L(x,\mu) = \psi_m(x) = 0$$

Like with equality constraints

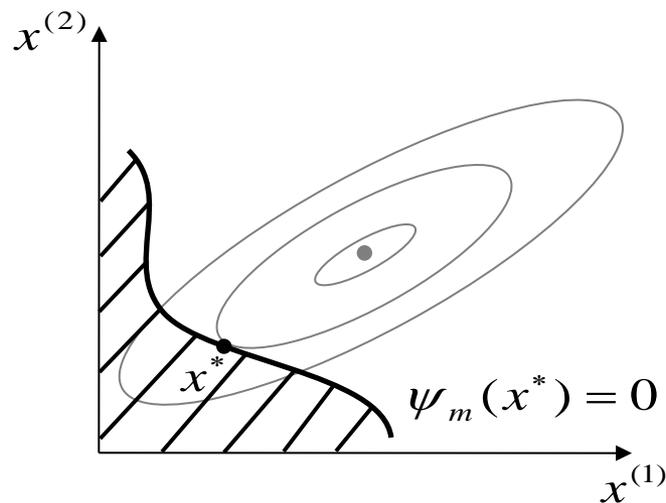# Optimization under inequality constraints

Inactive constraint

$$\psi_m(x^*) < 0$$

Active constraint

$$\psi_m(x^*) = 0$$

Optimization under inequality constraints

Kuhn-Tucker conditions

$$L(x,\mu) = F(x) + \mu_m \psi_m(x)$$

$$\nabla_x L(x,\mu) = \nabla_x F(x) + \mu_m \nabla_x \psi_m(x) = 0_S$$

$$\mu^T \nabla_\mu L(x,\mu) = \mu_m \psi_m(x) = 0$$

$$\nabla_\mu L(x,\mu) = \psi_m(x) \leq 0$$

$$\mu_m \geq 0$$

$$\mu_m = 0 \qquad \psi_m(x) < 0 \qquad m \text{ -th constraint is inactive}$$

$$\mu_m > 0 \qquad \psi_m(x) = 0 \qquad m \text{ -th constraint is active}$$
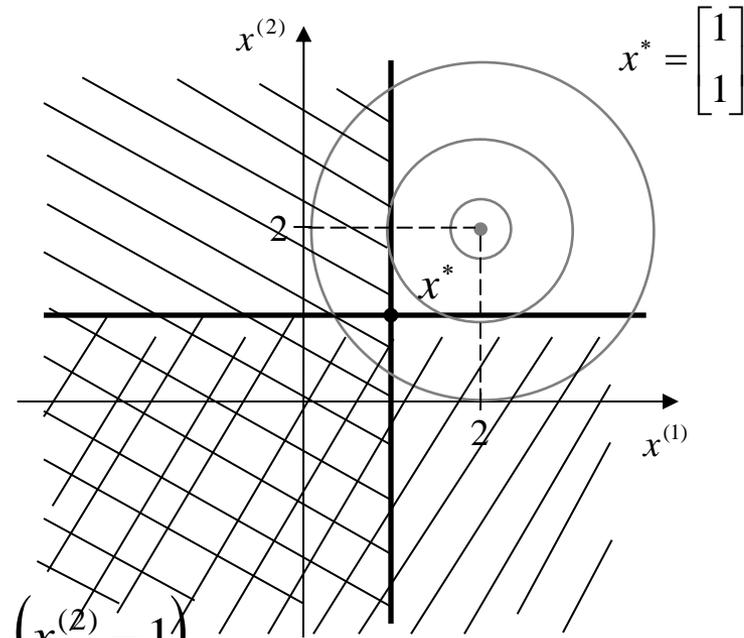
Optimization under inequality constraints

Kuhn-Tucker conditions

Example 1

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

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

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

$$x^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

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

Example 1.

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

- $\nabla_x L(x,\mu\ ) = \begin{bmatrix} 2\left(x^{(1)} - 2\right) + \mu_1 \\ 2\left(x^{(2)} - 2\right) + \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ \quad (1)\ (2)

- $\mu^T \nabla_\mu L(x,\mu\ ) \sim \begin{vmatrix} \mu_1\left(x^{(1)} - 1\right) = 0 \\ \mu_2\left(x^{(2)} - 1\right) = 0 \end{vmatrix}$ \quad (3)\ (4)

- $\nabla_\mu L(x,\mu\ ) = \begin{vmatrix} \left(x^{(1)} - 1\right) \le 0 \\ \left(x^{(2)} - 1\right) \le 0 \end{vmatrix}$ \quad (5)\ (6)

- $\mu = \begin{vmatrix} \mu_1 \ge 0 \\ \mu_2 \ge 0 \end{vmatrix}$ \quad (7)\ (8)

Example 1. c.d.

∞ 1^0 $\mu_1 = 0\ (x^{(1)} - 1 < 0\ ??),\ \mu_2 = 0\ (x^{(2)} - 1 < 0\ ??),$

z (1) → $2(x^{(1)} - 2) = 0 \rightarrow x^{(1)} = 2$

z (2) → $2(x^{(2)} - 2) = 0 \rightarrow x^{(2)} = 2$

z (5) → $(2 - 1) = 1 \geq 0\ sprzczność\ z\ (5)$ contradiction

z (6) → $(2 - 1) = 1 \geq 0\ sprzczność\ z\ (6)$ contradiction

∞ 2^0 $\mu_1 > 0\ (x^{(1)} - 1 = 0\ ??),\ \mu_2 = 0\ (x^{(2)} - 1 < 0\ ??),$

z (3) → $\mu_1(x^{(1)} - 1) = 0/\mu_1 \rightarrow (x^{(1)} - 1) = 0 \rightarrow x^{(1)} = 1$

z (1) → $2(1 - 2) + \mu_1 = 0 \rightarrow \mu_1 = 2$

z (2) → $2(x^{(2)} - 2) = 0 \rightarrow x^{(2)} = 2$

z (6) → $(2 - 1) = 1 \geq 0\ sprzczność\ z\ (6)$ contradiction

Example 1. c.d.

🔊 3^0 $\mu_1 = 0 \left(x^{(1)} - 1 < 0\,?\,?\right), \mu_2 > 0 \left(x^{(2)} - 1 = 0\,?\,?\right),$

z $(1) \rightarrow 2\left(x^{(1)} - 2\right) = 0 \rightarrow x^{(1)} = 2$

z $(5) \rightarrow (2 - 1) = 1 \geq 0\ sprzczność\ z\ (5)$ contradiction

z $(4) \rightarrow \mu_2\left(x^{(2)} - 1\right) = 0/\mu_1 \rightarrow \left(x^{(2)} - 1\right) = 0 \rightarrow x^{(2)} = 1$

z $(2) \rightarrow 2(1 - 2) + \mu_2 = 0 \rightarrow \mu_2 = 2$

🔊 4^0 $\mu_1 > 0 \left(x^{(1)} - 1 = 0\,?\,?\right), \mu_2 > 0 \left(x^{(2)} - 1 = 0\,?\,?\right),$

z $(3) \rightarrow \mu_1\left(x^{(1)} - 1\right) = 0/\mu_1 \rightarrow \left(x^{(1)} - 1\right) = 0 \rightarrow x^{(1)} = 1$
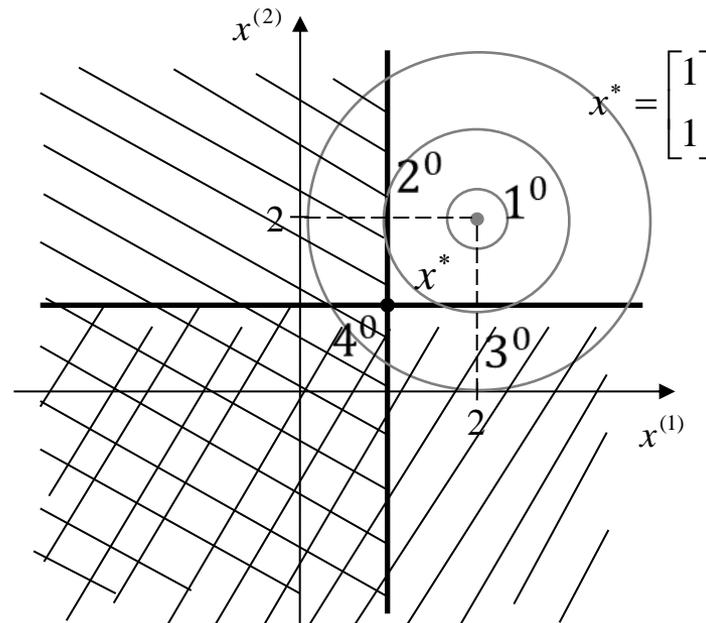
z $(1) \rightarrow 2(1 - 2) + \mu_1 = 0 \rightarrow \mu_1 = 2$

z $(4) \rightarrow \mu_2\left(x^{(2)} - 1\right) = 0/\mu_1 \rightarrow \left(x^{(2)} - 1\right) = 0 \rightarrow x^{(1)} = 1$

z $(2) \rightarrow 2(1 - 2) + \mu_2 = 0 \rightarrow \mu_2 = 2$ Point $x^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ optimal solution

Punkt $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ spełnia równania i jest rozwiązaniem zadania

Optimization under inequality constraints

Kuhn-Tucker conditions

Lagrange' a function :

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

where: $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_M \end{bmatrix}$ - Vector of Lgrange' a multiplayers

Neccesery conditions:

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

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

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

$$\mu^* \geq 0_M \qquad \Leftrightarrow \qquad \text{The solution is regular solution}$$

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

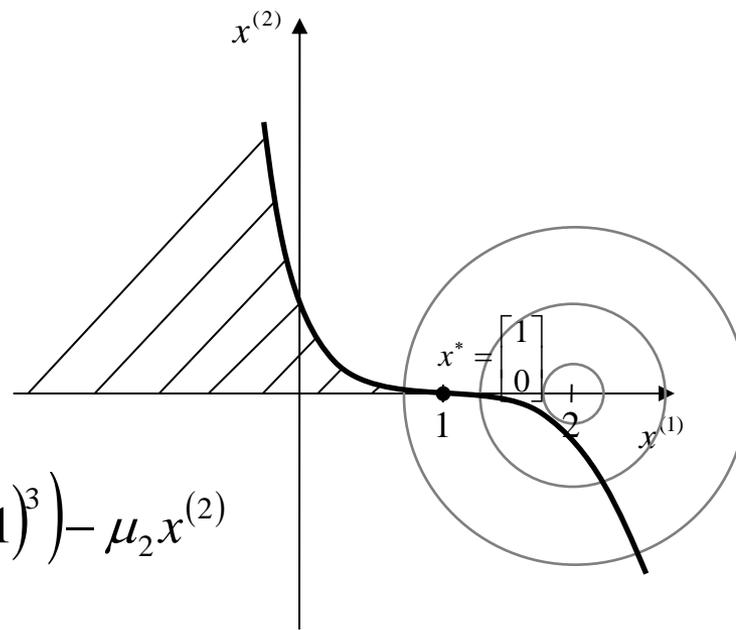# Optimization under inequality constraints

Kuhn-Tucker conditions

Example 2 – irregular

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

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

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

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

Example 2.

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

- $\nabla_x L(x, \mu) = \begin{bmatrix} 2(x^{(1)} - 2) + 3\mu_1 (x^{(1)} - 1)^2 \\ 2(x^{(2)} - 2) + \mu_1 - \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ (1)
                         (2)

- $\mu^T \nabla_\mu L(x, \mu) \sim \begin{vmatrix} \mu_1 \left(x^{(2)} + (x^{(1)} - 1)^3 \right) = 0 \\ -\mu_2 x^{(2)} = 0 \end{vmatrix}$ (3)
                         (4)

- $\nabla_\mu L(x, \mu) = \begin{vmatrix} x^{(2)} + (x^{(1)} - 1)^3 \leq 0 \\ -x^{(2)} \leq 0 \end{vmatrix}$ (5)
                         (6)

- $\mu = \begin{vmatrix} \mu_1 \geq 0 \\ \mu_2 \geq 0 \end{vmatrix}$ (7)
                  (8)

contradiction

Example 2.

The above system o equation ought to be solved as before. For each cases it can be shown contradiction.

We will show that solution $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ which can be notice from graphical illustratio does not fulfill system of equations

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Example 2.

For   we obtain

❧ Dla $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ otrzymujemy

contradiction

❧ $\nabla_x L(x, \mu) = \begin{bmatrix} 2(1-2) + 3\mu_1(1-1)^2 = -2 \neq 0 & sprzeczno\acute{c} \\ 2(0-2) + \mu_1 - \mu_2 = 0 \end{bmatrix}$ (1)
   (2)

❧ $\mu^T \nabla_\mu L(x, \mu) \sim \begin{vmatrix} \mu_1(0 + (1-1)^3) = 0 & \quad (3) \\ -\mu_2 0 = 0 & \quad (4) \end{vmatrix}$

❧ $\nabla_\mu L(x, \mu) = \begin{vmatrix} 0 + (1-1)^3 \leq 0 & \quad (5) \\ -0 \leq 0 & \quad (6) \end{vmatrix}$

❧ $\mu = \begin{vmatrix} \mu_1 \geq 0 & \quad (7) \\ \mu_2 \geq 0 & \quad (8) \end{vmatrix}$

Rozwiązanie nieregularne   Irregular solution

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Optimization under inequality constraints

Feasible directions

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_S \end{bmatrix}$$ – direction in $\mathscr{R}^S$

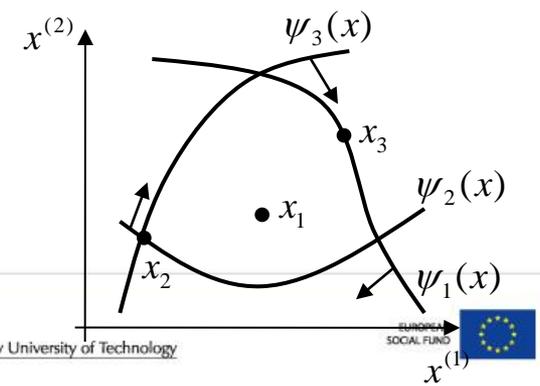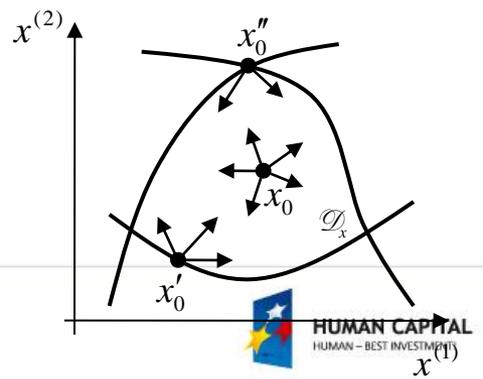$$x_1 = x_0 + \tau \cdot d$$
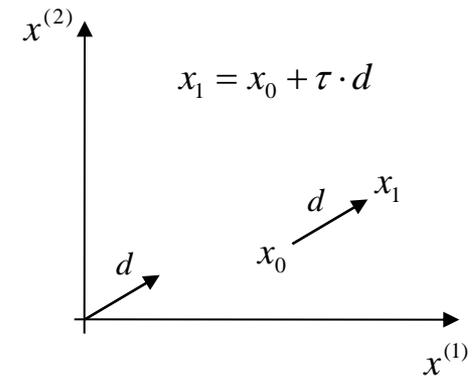
Feasible directions:

$$D(x) = \left\{ d \in \mathscr{R}^s : \exists \tau \quad x + \tau d \in \mathscr{D}_x \right\}$$

Active constraints:

$$I(x) = \left\{ m \in \{1, 2, \ldots, M\} : \psi_m(x) = 0 \right\}$$

$$I(x_1) = \varnothing$$
$$I(x_2) = \{2, 3\}$$
$$I(x_3) = \{1\}$$

Optimization under inequality constraints

Kuhn − Tucker rolls

Active constraints − analytical conditions?

$$D(x) = \left\{ d \in \mathcal{R}^s : \exists \tau \quad x + \tau d \in \mathcal{D}_x \right\}$$
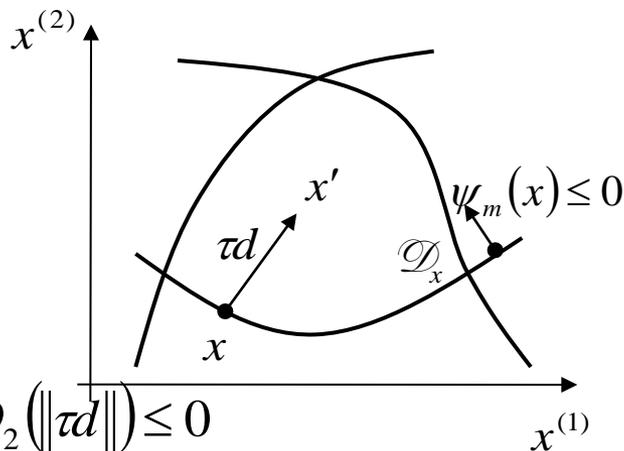
$$\forall m \in I(x)$$

tj.: $\psi_m(x) = 0$

$$x' = x + \tau d \in \mathcal{D}_x, \quad \tau > 0$$

$$\psi_m(x') \le 0$$

$$\psi_m(x') = \psi_m(x + \tau d) = \psi_m(x) + \tau d^T \nabla_x \psi_m(x) + O_2\left(\|\tau d\|\right) \le 0$$

$$\tau d^T \nabla_x \psi_m(x) \le 0 \quad \tau > 0$$

$$d^T \nabla_x \psi_m(x) \le 0 \quad \text{– analytical condition}$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Optimization under inequality constraints

Feasible directions

$$\forall d \in D(x) \wedge \forall m \in I(x)$$

How to determine the set of feasible directions?

Active constraints $\qquad \psi_m(x) = 0$
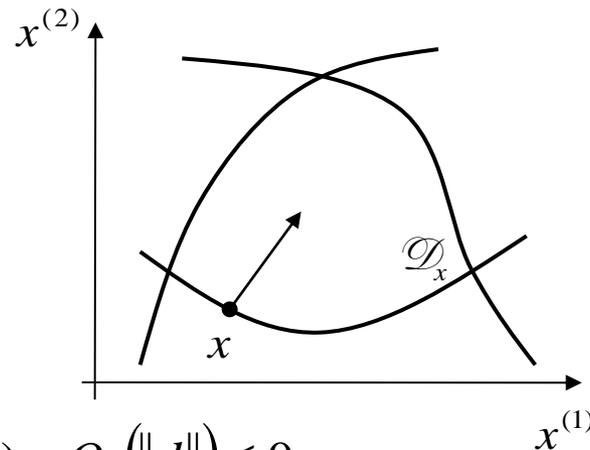
$$x' = x + \tau d \in \mathscr{D}_x$$

$$\psi_m(x') \leq 0$$

$$\psi_m(x') = \psi_m(x + \tau d) = \psi_m(x) + \tau d^T \nabla_x \psi_m(x) + O_2\big(\|d\|\big) \leq 0$$

$$\tau d^T \nabla_x \psi_m(x) \leq 0$$

$$\forall d \in D(x) \wedge \forall m \in I(x) \implies d^T \nabla_x \psi_m(x) \leq 0 \quad - \text{analytical condition}$$

$$\mathscr{D}(x) = \Big\{ d \in \mathscr{R}^s : \forall_m \in I(x),\ d^T \nabla_x \psi_m(x) \leq 0 \Big\} \qquad \mathscr{D}(x) \neq D(x)$$
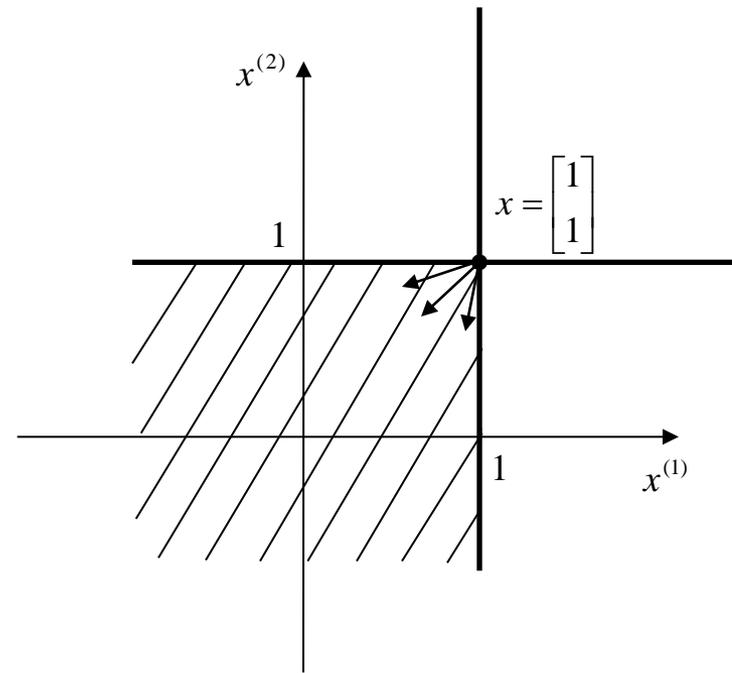
HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Optimization under inequality constraints

Feasible directions

Example

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

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

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Optimization under inequality constraints
Kuhn – Tucker rolls

Example 1

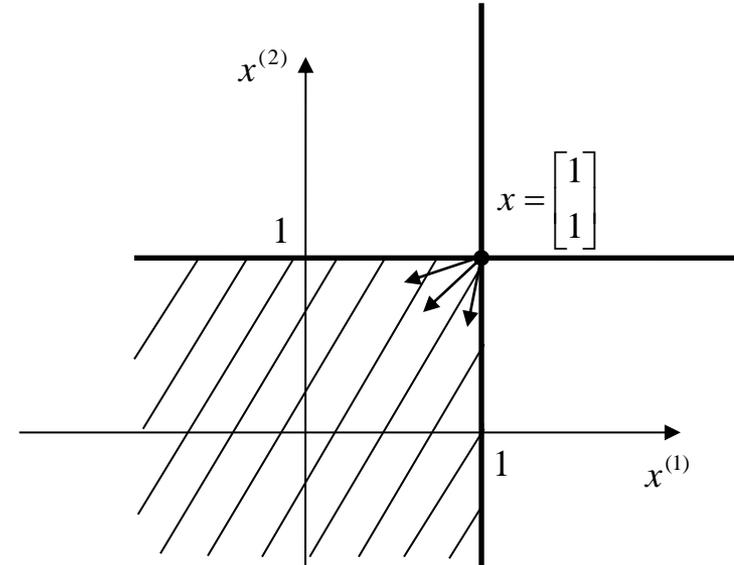$$\psi_1(x) = x^{(1)} - 1 \leq 0$$

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

In the point $\quad x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad I\left(x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = \{1, 2\}$

$$\nabla_x \psi_1(x) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \nabla_x \psi_2(x) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

$$d^T \nabla_x \psi_1(x) = \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 \cdot d_1 + 0 \cdot d_2 \leq 0 \quad \Rightarrow \quad d_1 \leq 0$$

$$d^T \nabla_x \psi_2(x) = \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \cdot d_1 + 1 \cdot d_2 \leq 0 \quad \Rightarrow \quad d_2 \leq 0$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

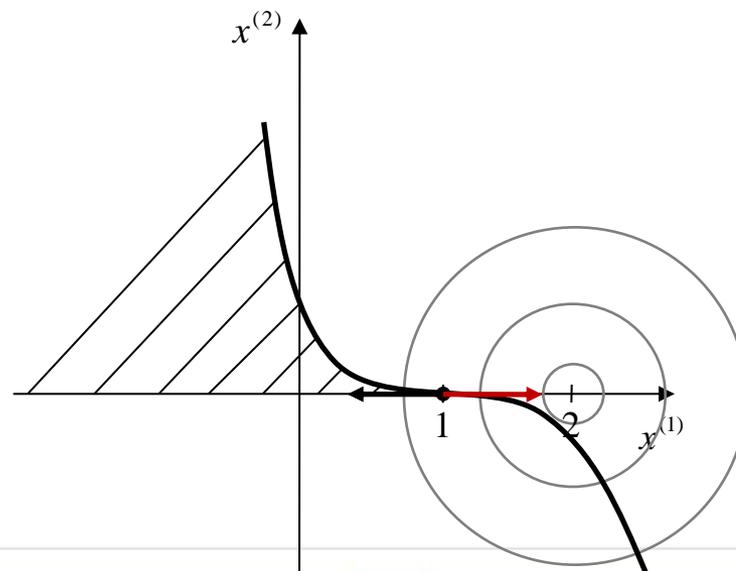# Optimization under inequality constraints

Feasible directions

$$\mathscr{D}(x) = \left\{ d \in \mathscr{R}^s : \forall_m \in I(x),\ d^T \nabla_x \psi_m(x) \le 0 \right\}$$

$$\mathscr{D}(x) \ne D(x) \quad \text{leads to irregular case}$$

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

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

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

Optimization under inequality constraints
Kuhn – Tucker conditions

Attention: Not all direction, which fulfils condition $d^T \nabla_x \psi_m(x) \le 0$ is feasible direction. It mays generate irregular solution $\mathscr{D}(x) \ne D(x)$

$$\mathscr{D}(x) = \left\{ d \in \mathscr{R}^s : \forall_m \in I(x), \ d^T \nabla_x \psi_m(x) \le 0 \right\} \quad D(x) = \left\{ d \in \mathscr{R}^s : \exists \tau \quad x + \tau d \in \mathscr{D}_x \right\}$$
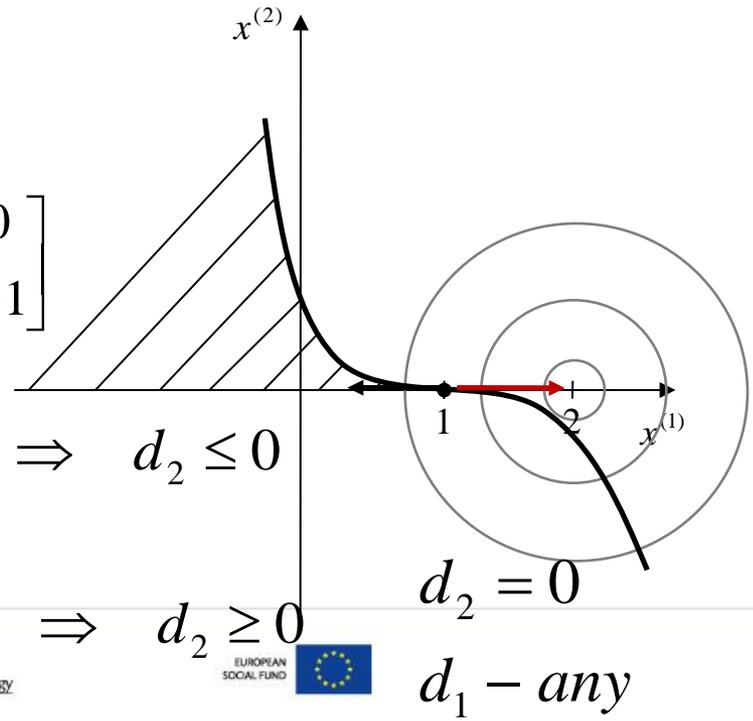
$$\psi_1(x) = x^{(2)} + \left(x^{(1)} - 1 \right)^3 \le 0 \quad \psi_2(x) = -x^{(2)} \le 0 \quad F(x) = \left(x^{(1)} - 2 \right)^2 + \left(x^{(2)} \right)^2$$

In the point $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $I\left(x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = \{1, 2\}$

$$\nabla_x \psi_1(x) = \begin{bmatrix} 3\left(x^{(1)} - 1 \right)^2 \\ 1 \end{bmatrix}_{x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \nabla_x \psi_2(x) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

$$d^T \nabla_x \psi_1(x) = \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \cdot d_1 + 1 \cdot d_2 \le 0 \quad \Rightarrow \quad d_2 \le 0$$

$$d^T \nabla_x \psi_2(x) = \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \cdot d_1 - 1 \cdot d_2 \le 0 \quad \Rightarrow \quad d_2 \ge 0$$

$$d_2 = 0$$
$$d_1 - any$$

Optimization under inequality constraints
Kuhn – Tucker conditions

$$F(x) = F(x_0 + \tau d) = F(x_0) + \tau \left(\nabla_x F(x) \right)^T d + O_3 \left(\| \tau d \| \right)$$

If d such that: $\quad \left(\nabla_x F(x) \right)^T d < 0 \quad to \quad F(x) < F(x_0)$ down ↘

Then

Let us divide set of directions $\quad \mathscr{D}(x) = \left\{ d \in \mathscr{R}^s : \forall_m \in I(x), \ d^T \nabla_x \psi_m(x) \le 0 \right\}$:

$$\mathscr{D}_1(x) = \left\{ d \in \mathscr{R}^s : \forall_m \in I(x), \ d^T \nabla_x \psi_m(x) \le 0 \right\} \wedge \left(\nabla_x F(x) \right)^T d \ge 0 \qquad \text{up} \quad ↗$$

$$\mathscr{D}_2(x) = \left\{ d \in \mathscr{R}^s : \forall_m \in I(x), \ d^T \nabla_x \psi_m(x) \le 0 \right\} \wedge \left(\nabla_x F(x) \right)^T d < 0 \qquad \text{down} \quad ↘$$

Optimization under inequality constraints

Lagrange function:

Kuhn – Tucker conditions

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

Kuhna – Tucker theorem – necessary optimality conditions:

If x^* is local minimum of optimization problem with inequality constraints, functions $:F, \Psi_1, \Psi_2, \ldots, \Psi_M$ are continuous and function F is differentiable then there exists set of Lagrange μ^* such one that together with x^* fulfils

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

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

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

$$\mu^* \geq 0_M$$

Regular solution

$$\Leftrightarrow \quad \mathcal{D}_2(x) = \left\{ d \in \mathcal{R}^s : \forall_m \in I(x),\ d^T \nabla_x \psi_m(x) \leq 0 \right\} \wedge \left(\nabla_x F(x) \right)^T d < 0 = \Phi$$

Irregular solution - Fiacco – Mac Cormica roll

$$\psi_1(x) = x^{(2)} + \left(x^{(1)} - 1\right)^3 \le 0 \qquad \psi_2(x) = -x^{(2)} \le 0 \qquad F(x) = \left(x^{(1)} - 2\right)^2 + \left(x^{(2)}\right)^2$$

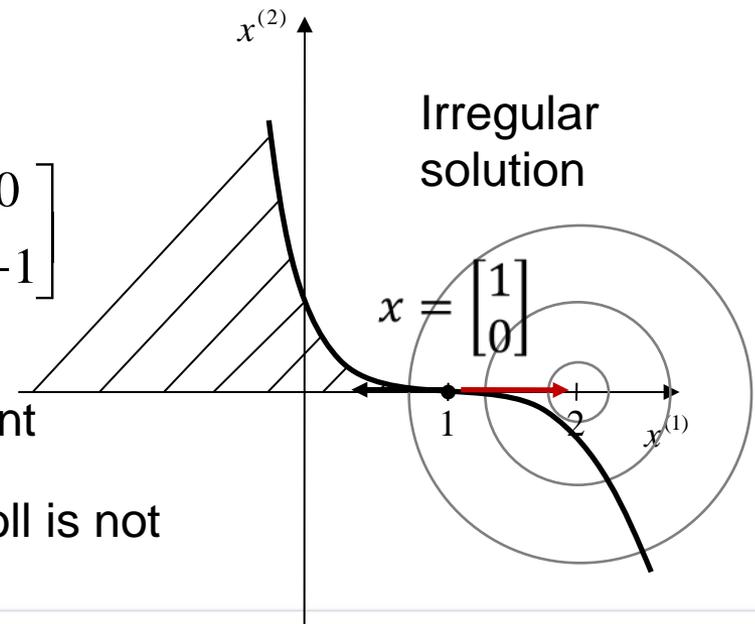In the point $\quad x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad I\left(x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = \{1,\, 2\}$

Consrtaints1 i 2 are active

$$\nabla_x\psi_1(x) = \begin{bmatrix} 3\left(x^{(1)} - 1\right)^2 \\ 1 \end{bmatrix}_{x=\begin{bmatrix}1\\0\end{bmatrix}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \nabla_x\psi_2(x) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

Gradients of constraints are linearly dependent

In the point $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ Fiacco – Mac Cormica roll is not fulfilled

Irregular solution

$x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

Optimization under inequality constraints

Kuhn-Tucker conditions

Sufficient condition of regularity:

$$F, \psi_1, \psi_2, \ldots, \psi_M \quad \text{– continuous and differentiable}$$

$$F \quad \text{– pseudo-convex}$$

$$\psi_1, \psi_2, \ldots, \psi_M \quad \text{– quasi-convex}$$

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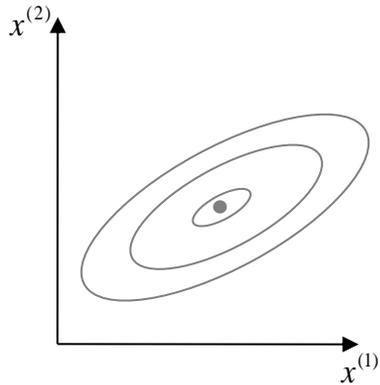
$$\mu^T \nabla_\mu L(x, \mu)\big|_{x^*, \mu^*} = 0$$

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

$$\mu^* \geq 0_M$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

General classification of optimization tasks

Unconstrained optimization:

$$\mathscr{D}_x = \mathscr{R}^S$$

Optimization under equality constraints:

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

Optimization under inequality constraints:

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

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund
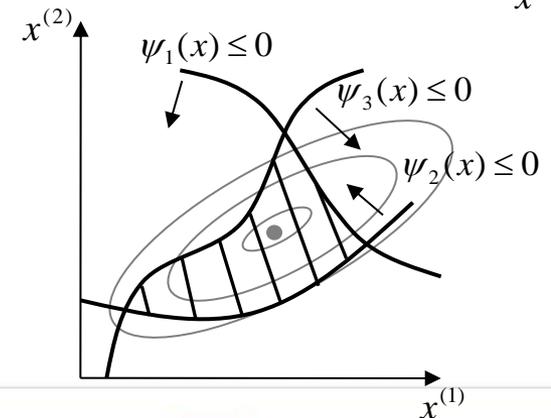
Optimization under inequality constraints
Kuhn – Tucker rolls

Necessary and sufficient conditions :

If functions $F(x), \psi_1(x), \psi_2(x), \ldots, \psi_M(x)$ are continuous and differentiable and function $F(x)$ is pseudo – convex function , and constraints $\psi_1(x), \psi_2(x), \ldots, \psi_M(x)$ are quasi – convex function then system of equations :

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

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

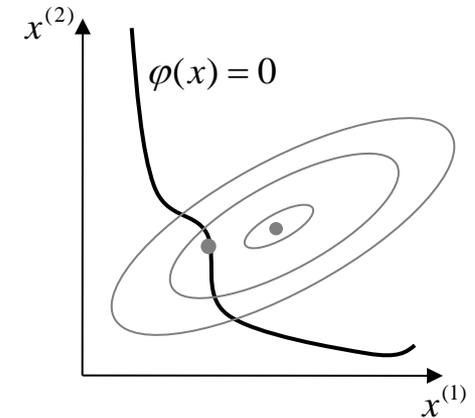$$\nabla_\mu L(x,\mu)\big|_{x^*,\mu^*} \leq 0_M$$

$$\mu^* \geq 0_M$$

Has one solution and it is the solution of the optimisation task with inequality constraints

Optimization under equality constraints

• The method of Lagrange multipliers

Lagrange function:

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

$$\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_L \end{bmatrix}, \quad \varphi(x) = \begin{bmatrix} \varphi_1(x) \\ \varphi_2(x) \\ \vdots \\ \varphi_L(x) \end{bmatrix}$$

Necessary conditions of optimality:

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

$$\nabla_\lambda L(x,\lambda)\big|_{x^*,\lambda^*} = 0_L \qquad \text{If and only if} \qquad \text{rank } G(x) = \text{rank } \big[G(x) \quad \vdots \quad -\nabla_x F(x)\big],$$

Where: $G(x) = \big[\nabla_x \varphi_1(x) \quad \vdots \quad \nabla_x \varphi_2(x) \quad \vdots \quad \cdots \quad \vdots \quad \nabla_x \varphi_L(x)\big]$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Optimization under inequality constraints

Lagrange function:

Kuhn − Tucker conditions

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

Kuhna − Tucker theorem − necessary optimality conditions:

If x^* is local minimum of optimization problem with inequality constraints, functions :$F, \Psi_1, \Psi_2, \ldots, \Psi_M$ are continuous and function F is differentiable then there exists set of Lagrange μ^* such one that together with x^* fulfils

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

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

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

$$\mu^* \geq 0_M$$

Regular solution

$$\Leftrightarrow \quad \mathscr{D}_2(x) = \left\{ d \in \mathscr{R}^s : \forall_m \in I(x), \ d^T \nabla_x \psi_m(x) \leq 0 \right\} \wedge \left(\nabla_x F(x) \right)^T d < 0 = \Phi$$

Saddle point

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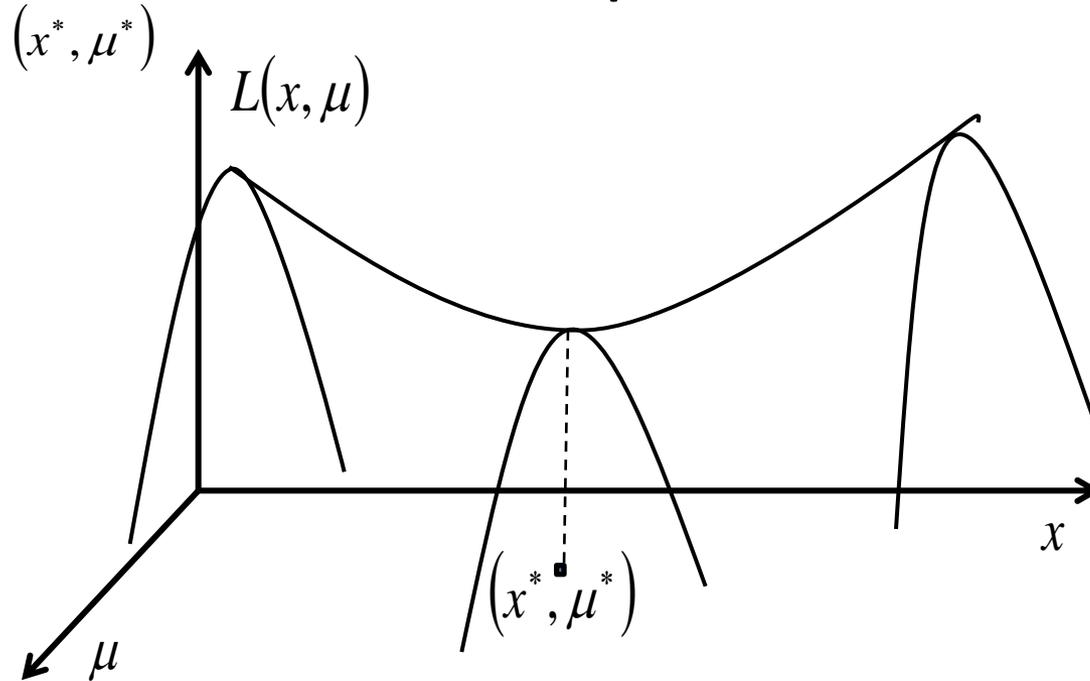
$$L\left(x^*,\mu^*\right)\le L\left(x,\mu^*\right) \quad \forall x \in \mathscr{D}(x) \subseteq \mathscr{R}^S$$

$$L\left(x^*,\mu\right)\le L\left(x^*,\mu^*\right) \qquad\qquad \forall \mu \ge 0_M$$

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

Saddle point

Point $\left(x^*, \mu^*\right)$ is the saddle poin $\qquad \left(x^* \in \mathcal{D}(x), \mu \geq 0_M\right) \quad \Leftrightarrow$

$$
\begin{array}{ll}
1. & x^* - \min imizing \ L(x, \mu) \\
2. & \psi_m\left(x^*\right) \leq 0 \quad m = 1, 2, \ldots, M \\
3. & \mu^* \psi_m\left(x^*\right) = 0 \quad m = 1, 2, \ldots, M
\end{array}
$$

If $\left(x^*, \mu^*\right)$ is the saddle point Lagrange'a function *L(x,μ)* then $\left(x^*, \mu^*\right)$ is the solution of the optimization task:

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

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

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Special case

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

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

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

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

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

$$x^* \geq 0_S$$

$$\nabla_\mu L(x, \mu)\big|_{x^*, \mu^*} \leq 0_S$$

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

$$\mu^* \geq 0_M$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

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

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

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

$$L(x.\square, \mu') = F(x) + \mu^T \psi(x) - \mu'^T x$$

Kuhn-Tucker conditions

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

$$L(x, \mu, \mu') = F(x) + \mu^T \psi(x) - \mu'^T x$$

$$\nabla_x L(x, \mu, \mu') = F(x) + \sum_{m=1}^{M} \mu_m \nabla_x \psi_m(x) - \mu' = 0_S$$

$$\mu^T \nabla_x L(x, \mu, \mu') = \mu^T \psi(x) = 0$$

$$\nabla_\mu L(x, \mu, \mu') = \psi(x) \leq 0_M$$

$$\nabla_{\mu'} L(x, \mu, \mu') = -x \leq 0_S$$

$$\mu \geq 0_M, \mu' \geq 0_S$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

98

Project co-financed from the EU European Social Fund

Special case

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

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

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

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

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

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

Special case

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

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

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

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

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

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

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

$$\mu^* \ge 0_S$$

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

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

$$\varphi(x) = 0_L \equiv \varphi(x) \le 0_L \cap -\varphi(x) \le 0_L$$

$$\mathscr{D}_x = \{ x \in \mathscr{R}^S : \varphi(x) \le 0_L \cap -\varphi(x) \le 0_L, \psi(x) \le 0_M \}$$

$$L(x, \lambda', \lambda, \mu) = F(x) + \lambda^T \varphi(x) - \lambda'^T \varphi(x) + \mu^T \psi(x)$$

Warunki Kuhna-Tuckera

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

$$L(x, \lambda, \lambda', \mu) = F(x) + \lambda^T \varphi(x) - {\lambda'}^T \varphi(x) + \mu^T \psi(x)$$

$$\nabla_x L(x, \lambda, \lambda', \mu) =$$

$$= \nabla_x F(x) + \sum_{l=1}^{L} \lambda_l \nabla_x \varphi_l(x) - \sum_{l=1}^{L} \lambda'_l \nabla_x \varphi_l(x) + \sum_{m=1}^{M} \mu_m \nabla_x \psi_m(x) = 0_S$$

$$\lambda^T \nabla_\lambda L(x, \lambda, \lambda', \mu) = \lambda^T \varphi(x) = 0$$

$${\lambda'}^T \nabla_{\lambda'} L(x, \lambda, \lambda', \mu) = -{\lambda'}^T \varphi(x) = 0$$

$$\mu^T \nabla_\mu L(x, \lambda, \lambda', \mu) = \mu^T \psi(x) = 0$$

$$\nabla_\lambda L(x, \lambda, \lambda', \mu) = \varphi(x) \leq 0_L$$

$$\nabla_{\lambda'} L(x, \lambda, \lambda', \mu) = -\varphi(x) \leq 0_L$$

$$\nabla_\mu L(x, \lambda, \lambda', \mu) = \psi(x) \leq 0_M$$

$$\lambda \geq 0_L, \quad \lambda' \geq 0_L, \quad \mu \geq 0_M$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

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

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

$$\varphi(x) = 0_L \equiv \varphi(x) \le 0_L \cap -\varphi(x) \le 0_L$$

$$\mathcal{D}_x = \left\{ x \in \mathcal{R}^S : \varphi(x) \le 0_L \cap -\varphi(x) \le 0_L, \psi(x) \le 0_M \right\}$$

$$L(x, \lambda', \lambda, \mu) = F(x) + \lambda^T \varphi(x) - {\lambda'}^T \varphi(x) + \mu^T \psi(x)$$

Warunki Kuhna-Tuckera

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Analytical methods

Disadvantages

It is hard to apply them if:

$$F, \varphi, \psi \quad \text{are nonlinear}$$

$$\dim(x) \quad \text{is large}$$

They cannot be applied if:

$$F, \varphi, \psi \quad \text{are not differentiable}$$

$$F \quad \text{is not given by formula and it may only be}$$
$$\text{measured for requested value of} \quad x$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Numerical optimization methods

$$x^* \rightarrow F(x^*) = \min_{x \in D_x} F(x)$$
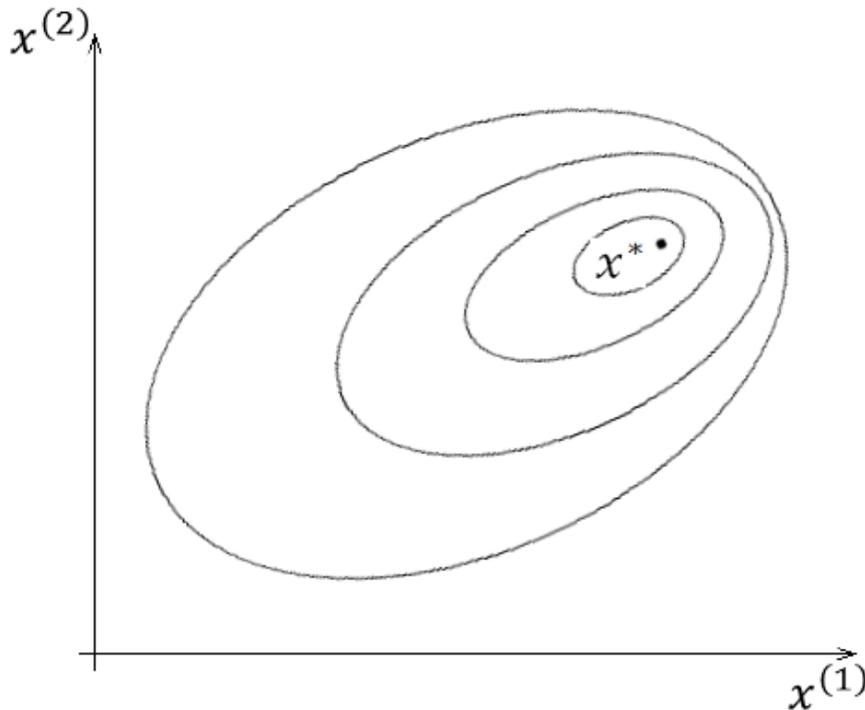
Analytical methods has drawbacks, when:

1. The goal function F and constraints φ, ψ are nonlinear.
2. Functions F, φ and ψ are non-differentiable
3. Mathematical formula describing functions F, φ and ψ is not available, it can only be „measured"
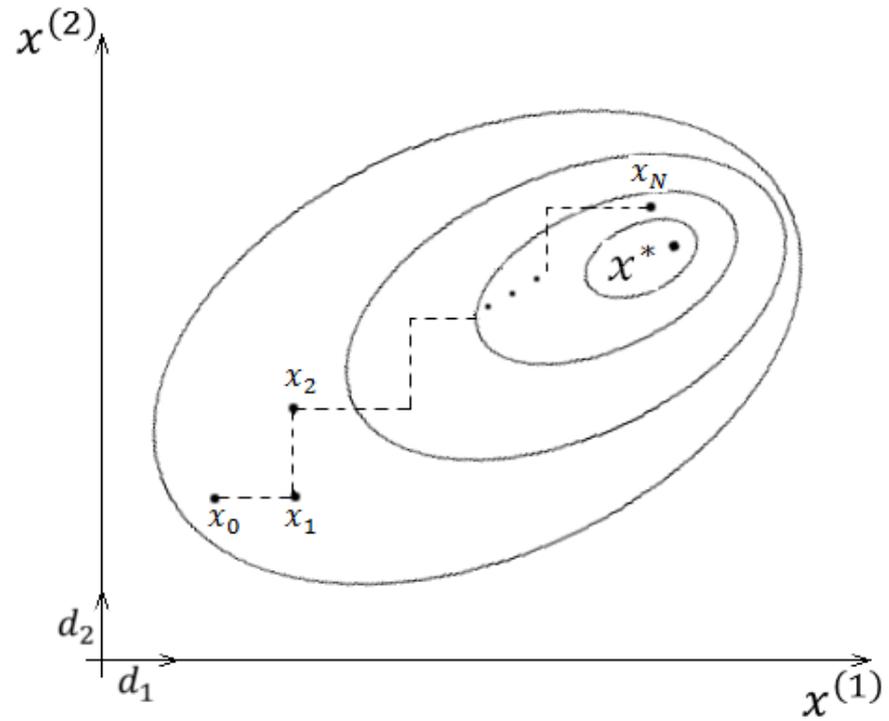4. Large dimension of decision variables vector

Numerical methods

We only use information about values of objective function $F(x)$ for a given value of x.

$$x_0 \longrightarrow x_1 \longrightarrow x_2 \longrightarrow \cdots \longrightarrow x_N \approx x^*$$
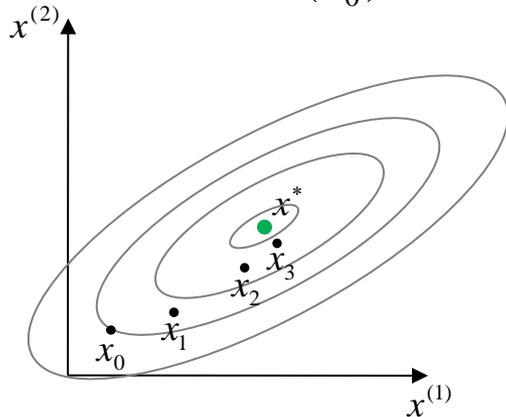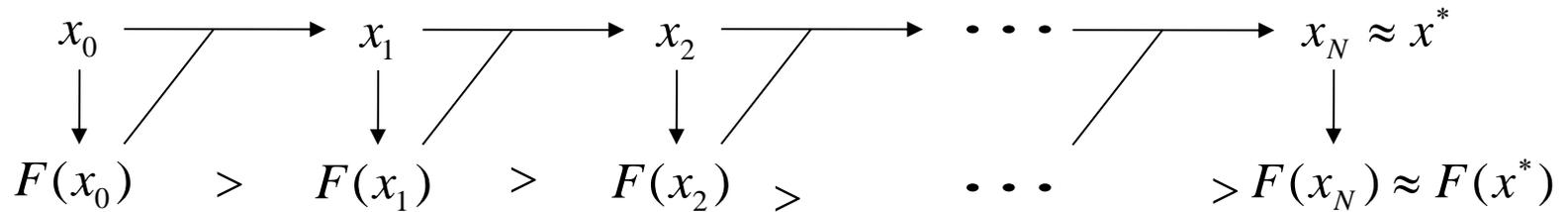
$$F(x_0) \quad > \quad F(x_1) \quad > \quad F(x_2) \quad > \quad \cdots \quad > F(x_N) \approx F(x^*)$$

The general idea behind numerical methods.

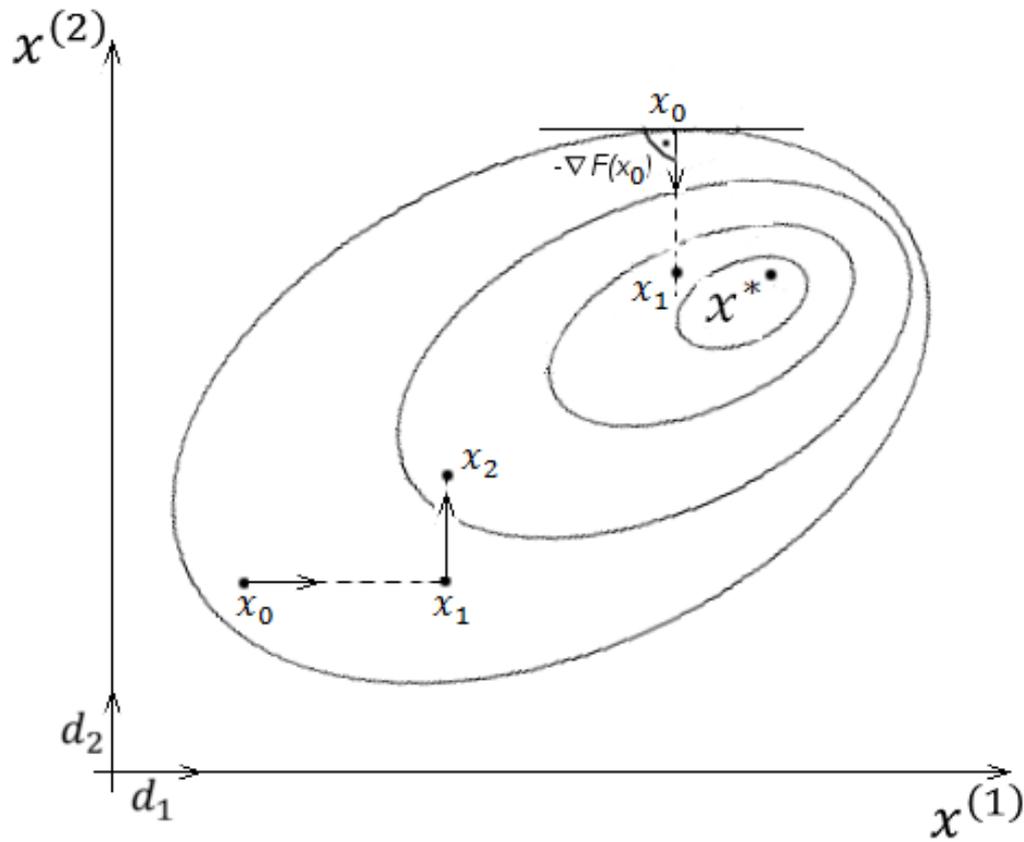# Numerical optimization methods

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

- Choice of the search direction.
- Line search optimization.
- Stopping conditions.

$$x_0, x_1, \ldots, x_n, \ldots, x_N \approx x^*$$
$$F(x_0) > F(x_1) > \ldots > F(x_n) > \ldots > F(x_N) \approx F(x^*)$$
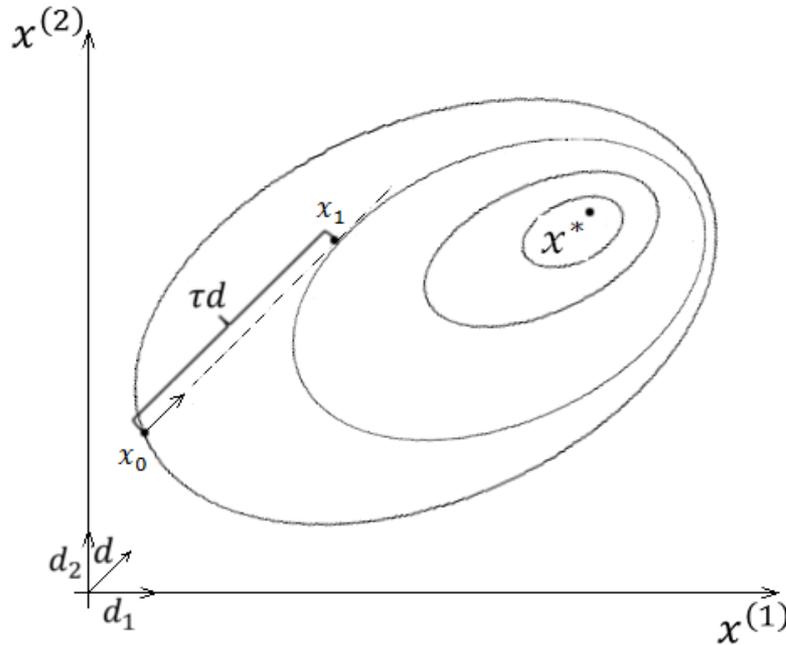
Choice of the search direction

- Basis of search directions – non-gradient methods.

- Search directions based on gradient vectors – gradient-based methods.

Line search optimization

x_0 – initial solution
x_1 – next solution
d – search direction
τ – step size

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

x_0, d – fixed

$$F(x_0 + \tau d) \triangleq f(\tau)$$

$f(\tau)$ – a single variable function (of the step size τ)

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

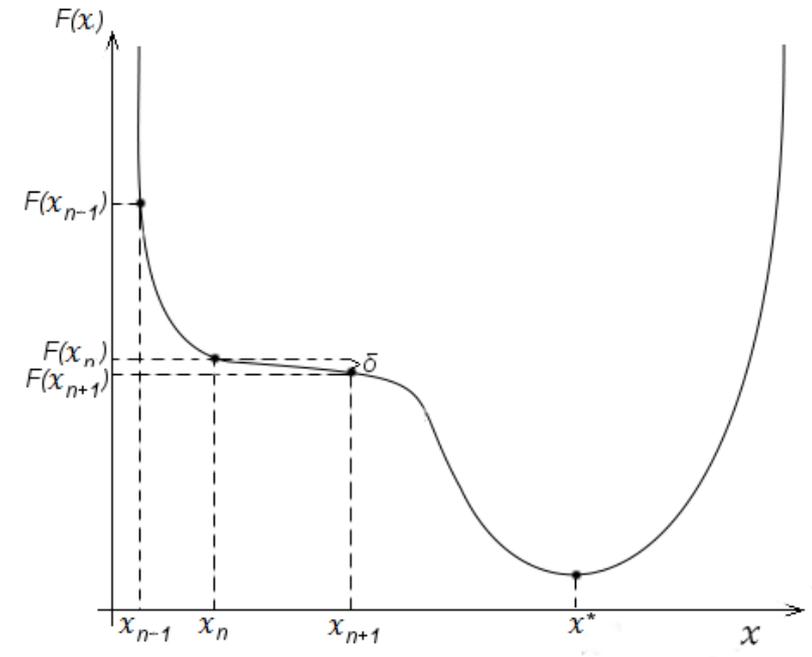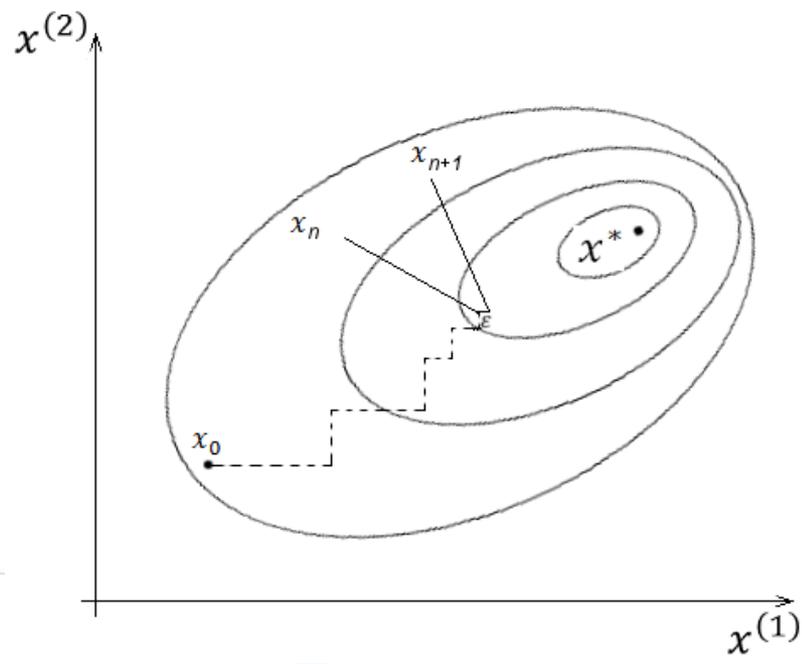line search optimization $\equiv$ optimization of a single variable function

Stopping conditions
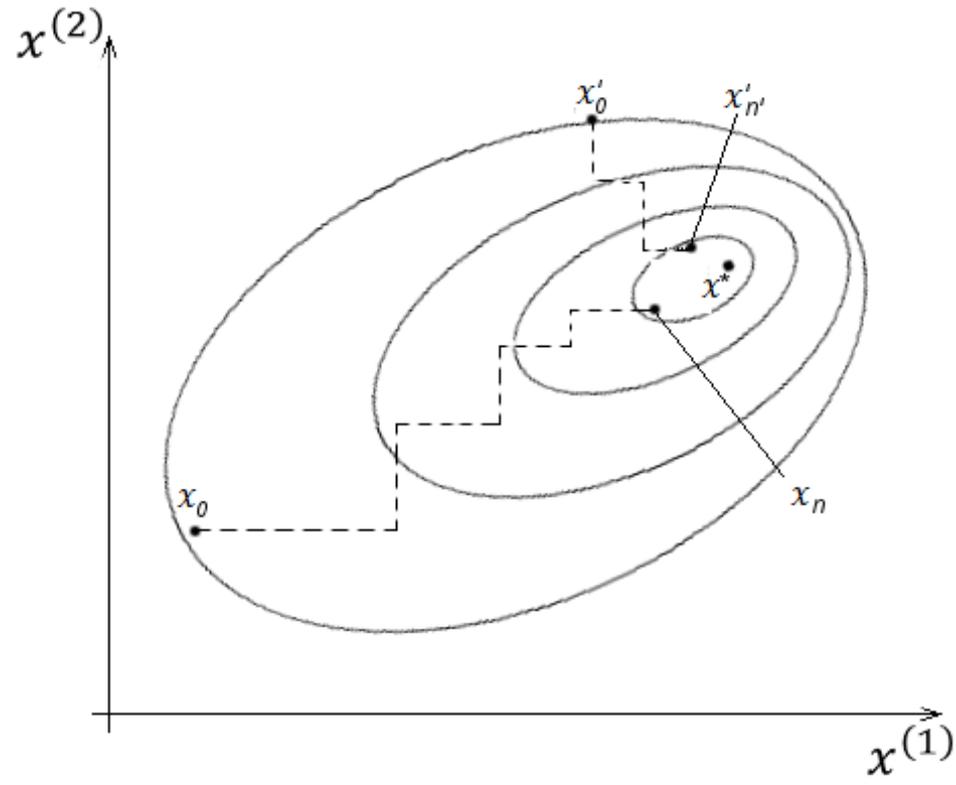
$$\|x_{n+1} - x_n\| < \varepsilon; \qquad\qquad |F(x_{n+1}) - F(x_n)| < \delta;$$

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

Remedy

$$\|x_n - x'_{n'}\| < \varepsilon$$

x_0, x'_0 – different initial solutions
$x_n, x'_{n'}$ – responding final solutions

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN
SOCIAL FUND

Project co-financed from the EU European Social Fund

Thank you for attention