# Computer Science

Jerzy Świątek

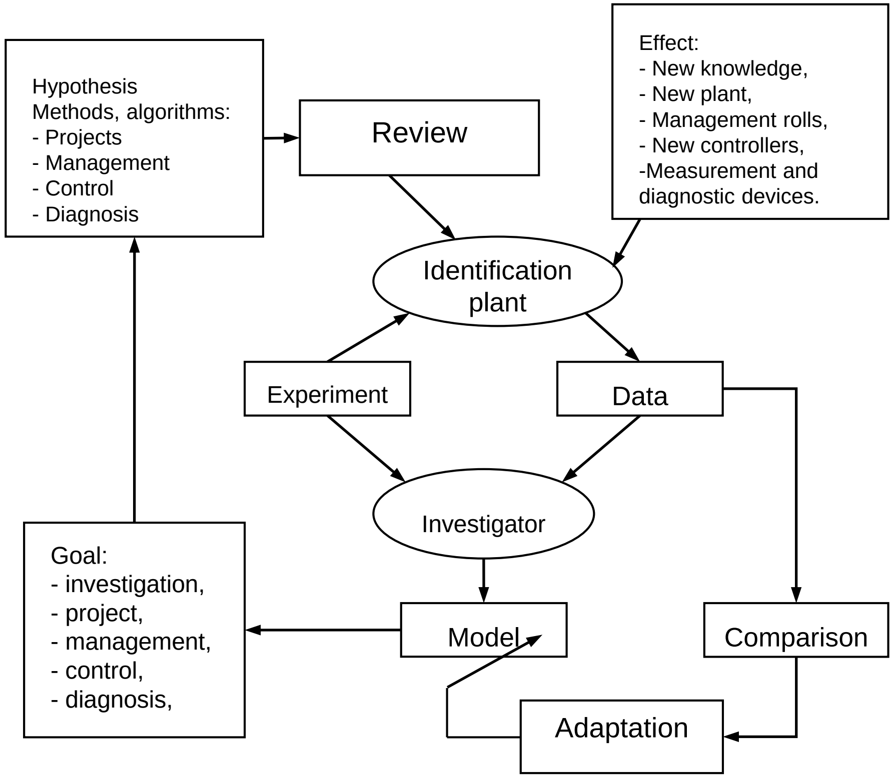
# Systems Modelling and Analysis

Choose yourself and new technologies

## L.16. Model based decision making

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

# Model in the systems research



# Example of decision making

<u>Decision</u>: 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 <u>parameters</u>:

$c_1, c_2, c_3$ – unit costs of workloads

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

<u>Constraints</u>:

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

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

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

# 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):

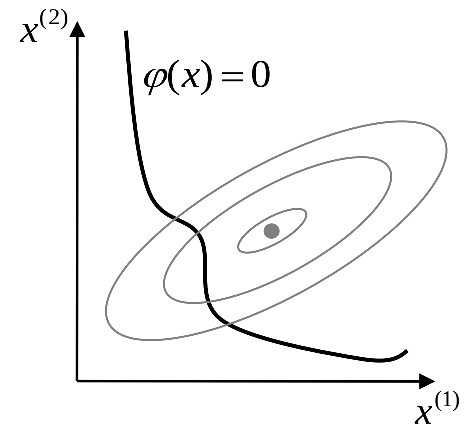
$$x \in \mathrm{D}_x$$

**Optimization task:** $x^* \to F(x^*) = \min\limits_{x^* \in \mathrm{D}_x} F(x),$ $x^*$ – optimal decision

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

# General classification of optimization tasks

Unconstrained optimization: $D_x = R^S$

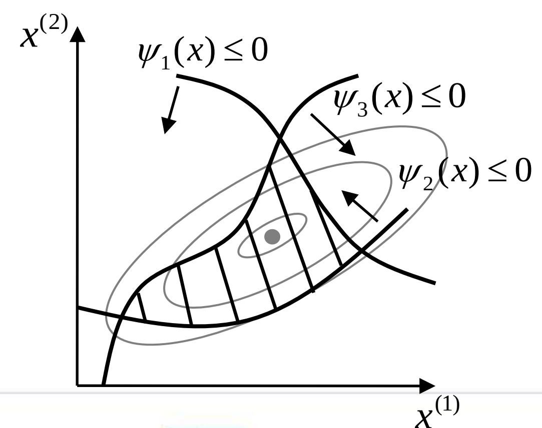



Optimization under equality constraints:

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

Optimization under inequality constraints:

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



# Analytical methods

- Unconstrained optimization
- Lagrange multipliers method – equality constraints
- Kuhn-Tucker conditions – inequality constraints

# Common types of optimization tasks

- Linear programming

Decision variables: $x \in D_x \subseteq R^S$

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:

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

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

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

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

# Common types of optimization tasks

- Quadratic programming

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

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

Constraints:

$$\varphi_l(x) = d_l^T x - \alpha_l = 0, \quad l = 1,2,\ldots,L \qquad d_l = \begin{bmatrix} d_l^{(1)} \\ d_l^{(2)} \\ \vdots \\ d_l^{(S)} \end{bmatrix}$$

$$\psi_m(x) = e_m^T x - \beta_m \le 0, \quad m = 1,2,\ldots,M \qquad e_l = \begin{bmatrix} e_m^{(1)} \\ e_m^{(2)} \\ \vdots \\ e_m^{(S)} \end{bmatrix}$$

# Common types of optimization tasks

- Linear-fractional programming

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

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

Constraints:

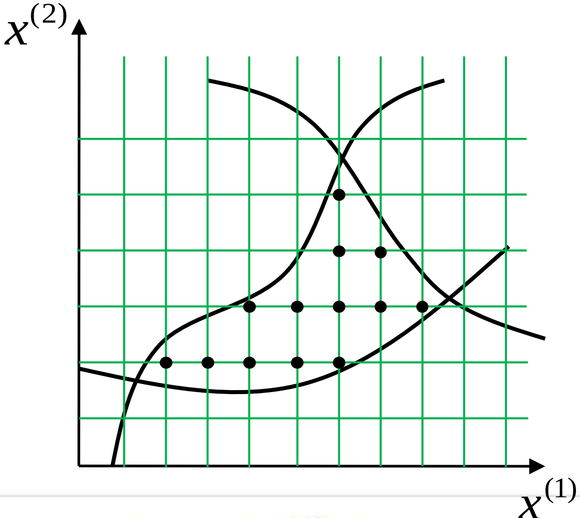
$$\varphi_l(x) = p_l^T x - \alpha_l = 0,\quad l = 1,2,\ldots,L \qquad p_l = \begin{bmatrix} p_l^{(1)} \\ p_l^{(2)} \\ \vdots \\ p_l^{(S)} \end{bmatrix}$$

$$\psi_m(x) = q_m^T x - \beta_m \le 0,\quad m = 1,2,\ldots,M \qquad q_l = \begin{bmatrix} q_m^{(1)} \\ q_m^{(2)} \\ \vdots \\ q_m^{(S)} \end{bmatrix}$$

# Common types of optimization tasks

- Integer programming

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




Special cases

$$x \in \overline{\mathrm{D}_x} = \{x_1, x_2, \ldots, x_M\}$$

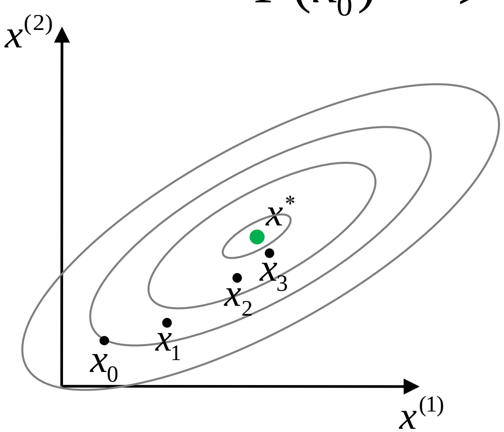
$$x \in \overline{\mathrm{D}_x} = \{x^{(s)} \in \{0, 1\}, s = 1, 2, \ldots, S\}$$

# Numerical methods

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

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

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

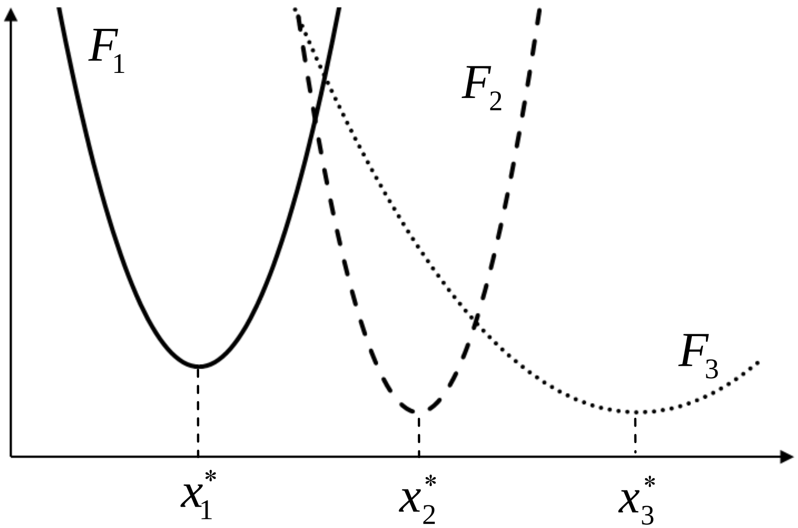



**The general idea** behind numerical methods.

# Multiobjective optimization

$x$ – vector of decision variables

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



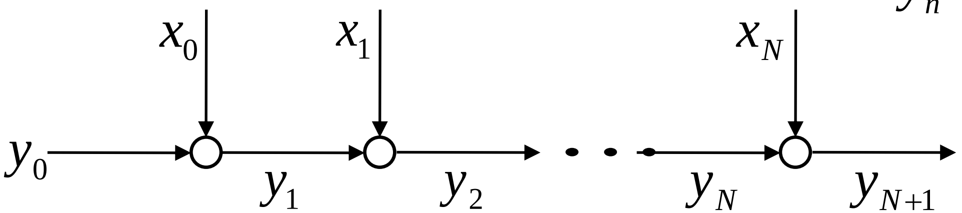
# Dynamic optimization

Dynamic process: $y_{n+1} = P(y_n, x_n)$

$n$ – time step

$x_n$ – decision made at $n$-th time step

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




The task 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.

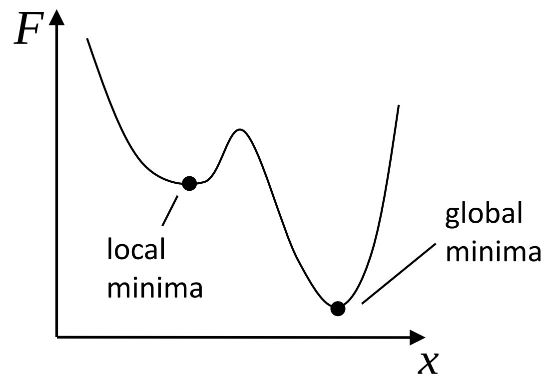

http://www.all-freeware.com/

# Mathematical preliminaries

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

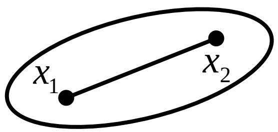
**Local minima:** $\forall_{\varepsilon > 0} \exists_{x \in O(x^*, \varepsilon)} F(x^*) < F(x)$

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

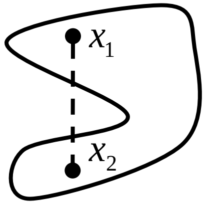


# Mathematical preliminaries

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



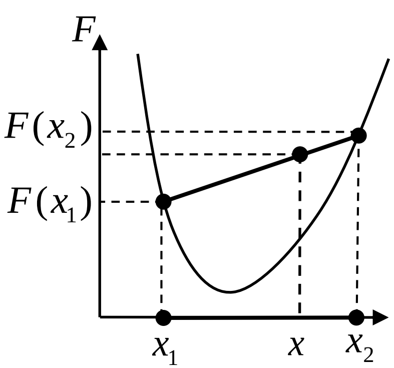

– convex set




– nonconvex set

**Convex function:**

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



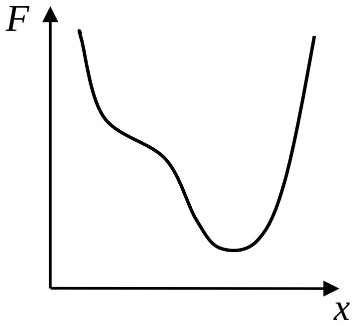
# Mathematical preliminaries

**Pseudo-convex function:**

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

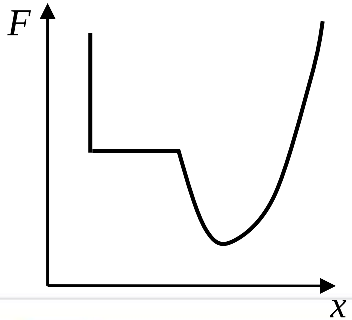
$$F(x) = F(x_0) + (x - x_0)^T \left[\nabla_x F(x_0)\right] + O_2\left(\|x - x_0\|\right)$$

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




**Quasi-convex function:**

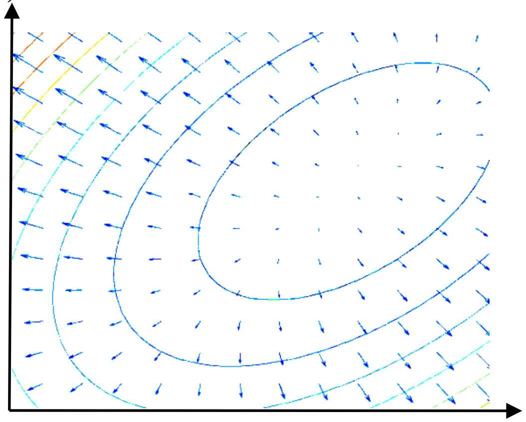
$\mathrm{D}_\alpha = \{x \in \mathrm{D}_x : F(x) \leq \alpha\}$ – convex sets



# Mathematical preliminaries

**Gradient:**

$$\nabla_x F(x) = \begin{bmatrix} \dfrac{\partial F}{\partial x^{(1)}} \\ \dfrac{\partial F}{\partial x^{(2)}} \\ \vdots \\ \dfrac{\partial F}{\partial x^{(S)}} \end{bmatrix}$$




**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:**

$$\frac{\partial^2 F}{\partial x^{(i)} \partial x^{(j)}} = \frac{\partial^2 F}{\partial x^{(j)} \partial x^{(i)}} \Rightarrow H$$ 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

# Podstawy matematyczne

Kryterium Sylwestra:

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

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

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

Eigen values of matrix *H*

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

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

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

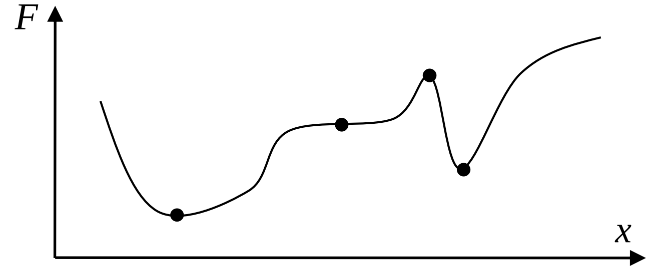
# Unconstrained optimization

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

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

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

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



# 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.

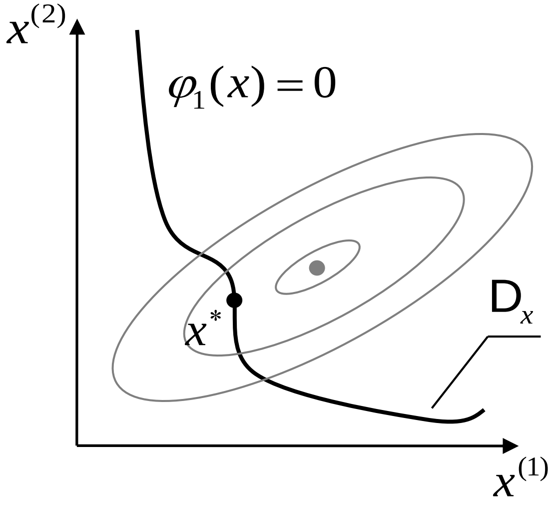
If $H(x^*)$ is neither negative semidefinite nor positive semidefinite at $x^*$ then $x^*$ is not optimum.

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

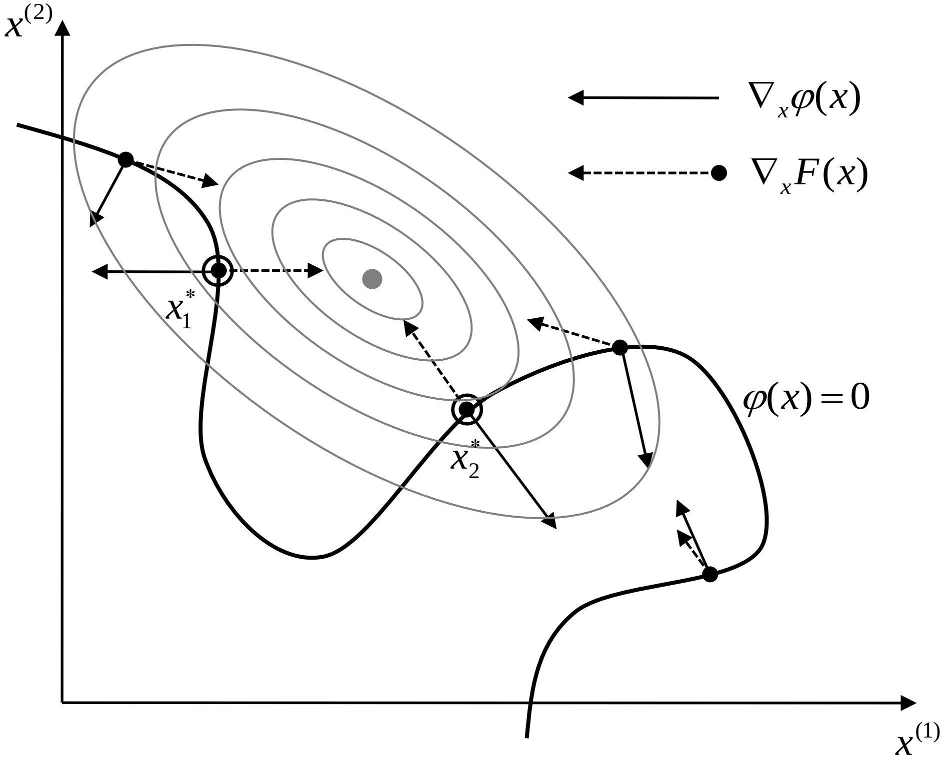
# Optimization under equality constraints

Optimization task: $x^* \to F(x^*) = \min_{x^* \in D_x} F(x)$

$$D_x = \{ x \in R^S : \ \varphi_1(x) = 0, \varphi_2(x) = 0, \ldots, \varphi_L(x) = 0, \ L \le S \}$$



# Optimization under equality constraints




Locally optimal solution satisfies condition:

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

where

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

For multiple constraints:

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

# 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:

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

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

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

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

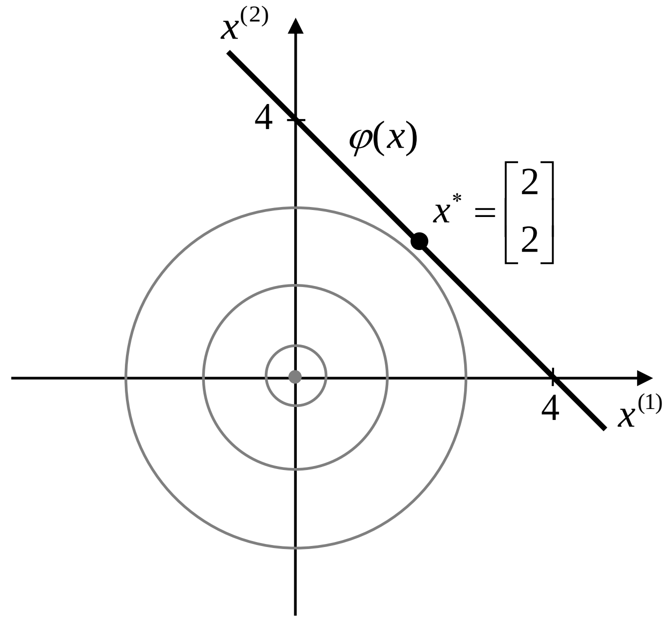
# Optimization under equality constraints

- The method of Lagrange multipliers – example 1

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

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

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



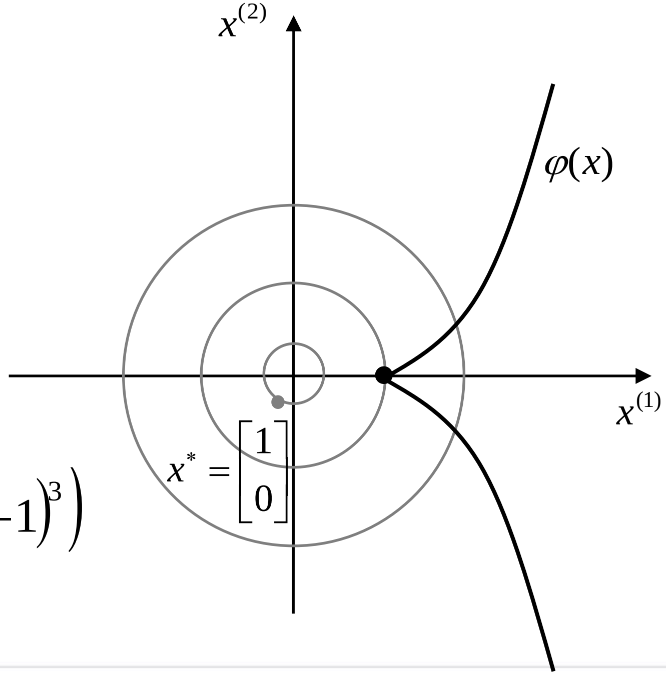
# Optimization under equality constraints

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

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

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

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



# Optimization under equality constraints

- The method of Lagrange multipliers – example 2 explanation

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

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

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

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

which is always true as long as $F$ is convex and $\varphi_l$ are linear.

How to find irregular solutions?

# 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:

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

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

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

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

2° $\lambda_0 = 0$ $\quad \sum_{l=1}^{L} \lambda_l \nabla_x \varphi_l(x) = 0_S$ We obtain irregular solutions.

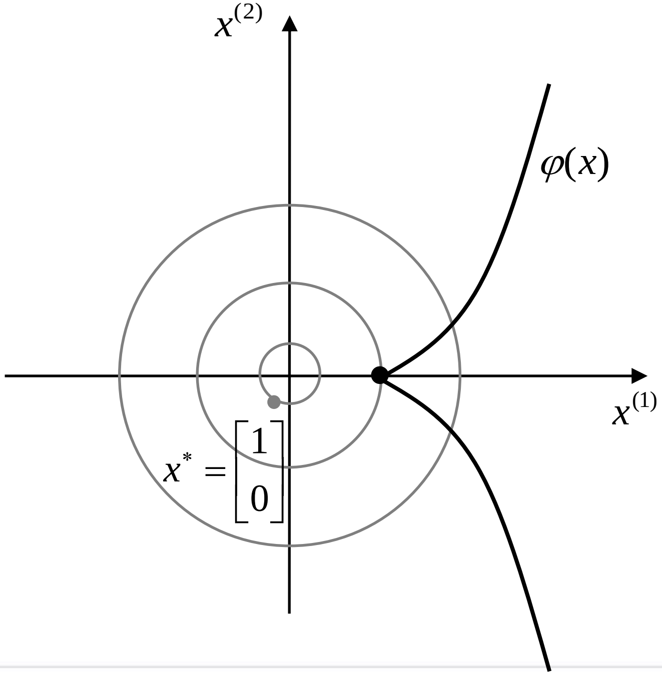
Second order condition of optimality requires analysis of $H(x,\lambda,\lambda_0) = \nabla_{xx}^2 L(x,\lambda,\lambda_0)$.

# Optimization under equality constraints

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

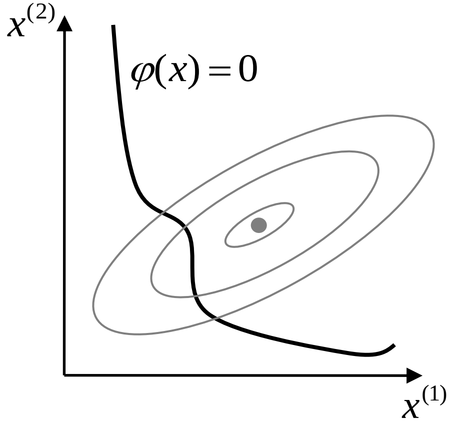
$$F(x) = \left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2$$

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



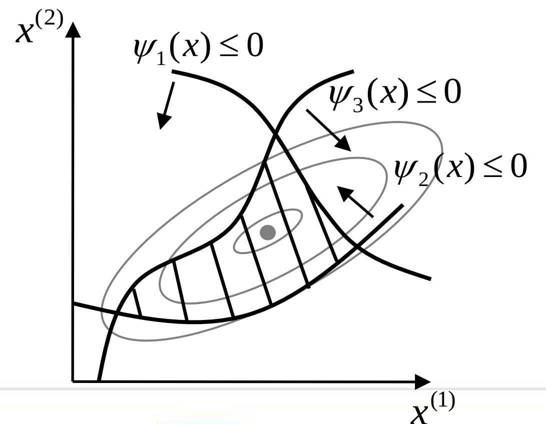
# General classification of optimization tasks

Unconstrained optimization: $D_x = R^S$




Optimization under equality constraints:

$$D_x = \{x \in R^S : \varphi_1(x)=0, \varphi_2(x)=0, \ldots, \varphi_L(x)=0, L \le S\}$$



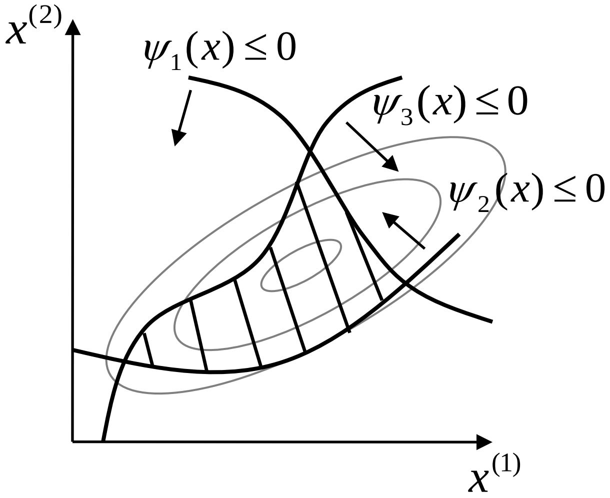

Optimization under inequality constraints:

$$D_x = \{x \in R^S : \psi_1(x)\le 0, \psi_2(x)\le 0, \ldots, \psi_M(x)\le 0\}$$

# Optimization under inequality constraints

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

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

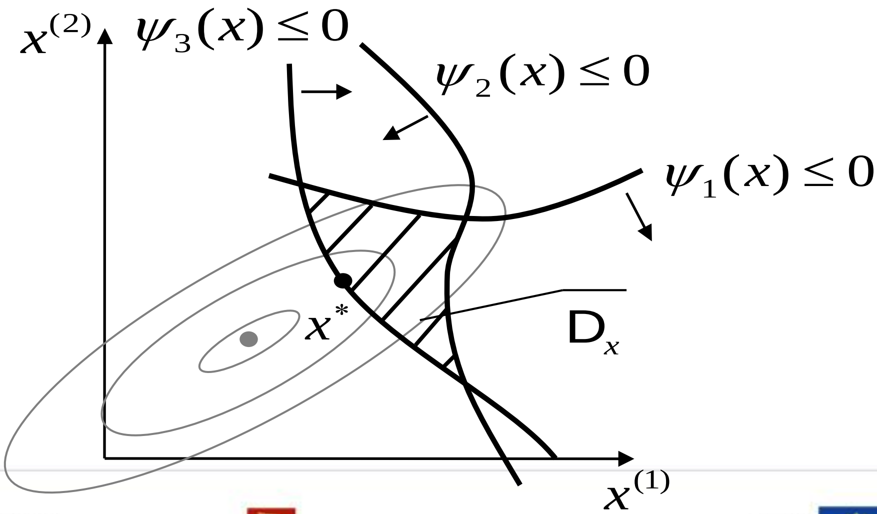


# Optimization under inequality constraints

**Optimization task**

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

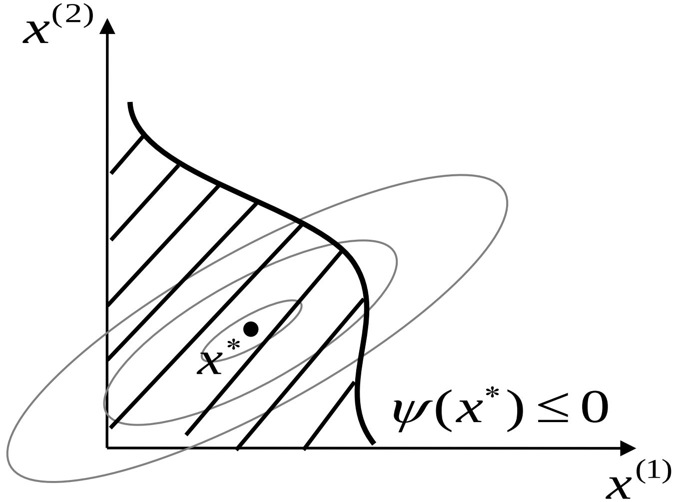
$$D_x = \{ x \in R^S : \psi_1(x) \le 0, \psi_2(x) \le 0, \ldots, \psi_M(x) \le 0, \}$$



# Optimization under inequality constraints

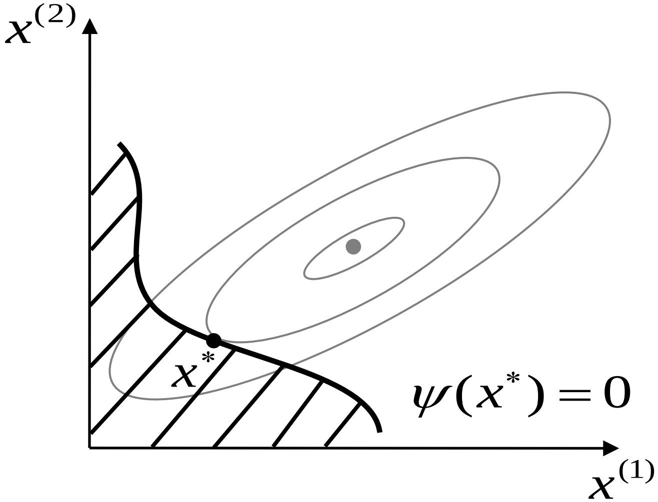
Inactive constraint

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




Active constraint

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




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 conditions

Lagrange function:

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

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

Necessary conditions of optimality:

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

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

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

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

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

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

# Optimization under inequality constraints

Kuhn-Tucker conditions

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

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

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

$$\mu \ge 0_M$$

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

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

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

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

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

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

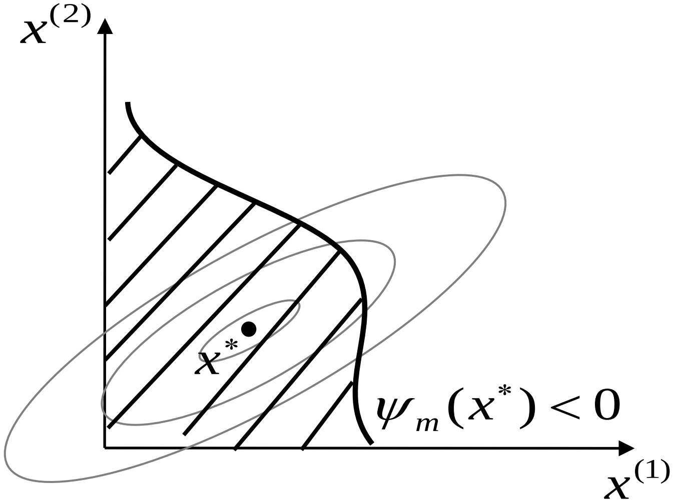
$$\mu_m^* \ge 0 \qquad m = 1, 2, \ldots, M$$

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

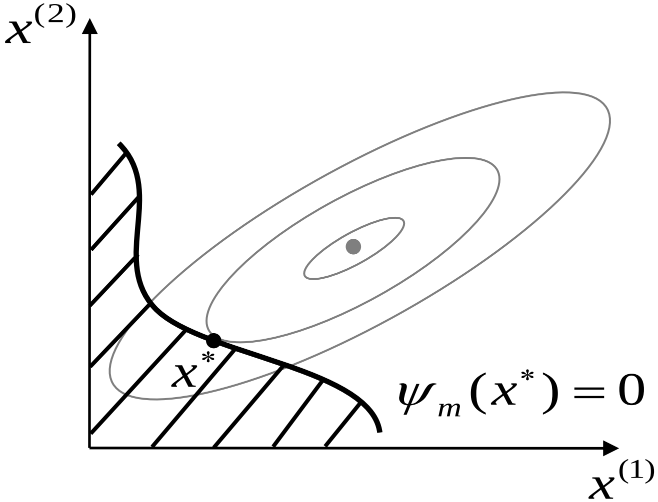
Inactive constraint

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




Active constraint

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



# Optimization under inequality constraints

Kuhn-Tucker conditions

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

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

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

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

$$\mu_m \ge 0$$

$\mu_m = 0 \quad \psi_m(x) < 0$ $\quad m$ -th constraint is inactive

$\mu_m > 0 \quad \psi_m(x) = 0$ $\quad m$ -th constraint is active

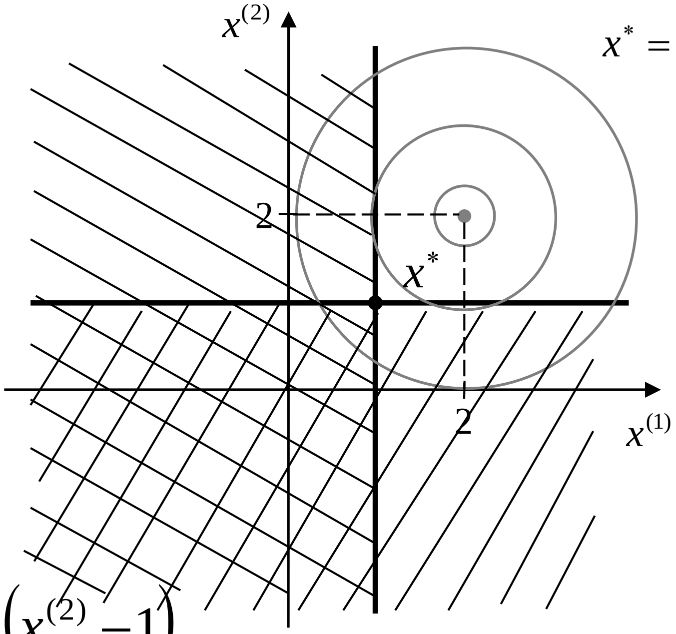
# Optimization under inequality constraints

## Kuhn-Tucker conditions

Example 1

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

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




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

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

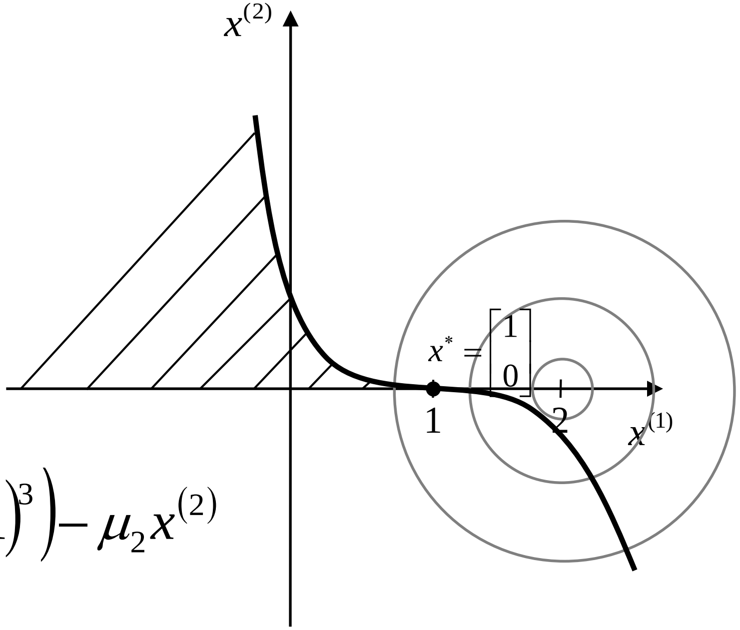
Kuhn-Tucker conditions

Example 2 – irregular

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

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

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

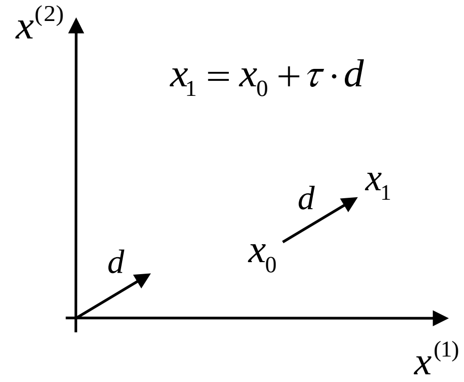



$$L(x,\lambda) = \left(x^{(1)} - 2\right)^2 + \left(x^{(2)}\right)^2 + \mu_1\left(x^{(2)} + \left(x^{(1)} - 1\right)^3\right) - \mu_2 x^{(2)}$$

# Optimization under inequality constraints

## Feasible directions

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

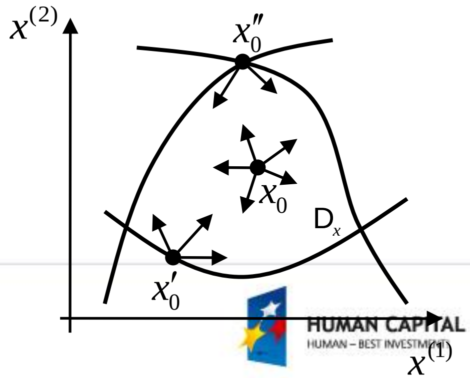


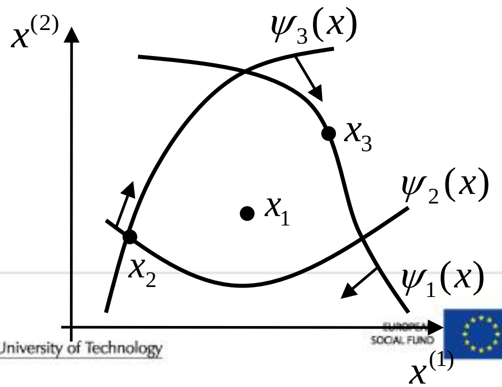

Feasible directions:

$$D(x) = \{d \in \mathbb{R}^s : \exists \tau \quad x + \tau d \in D_x\}$$

Active constraints:

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




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

# Optimization under inequality constraints

# Kuhn – Tucker rolls

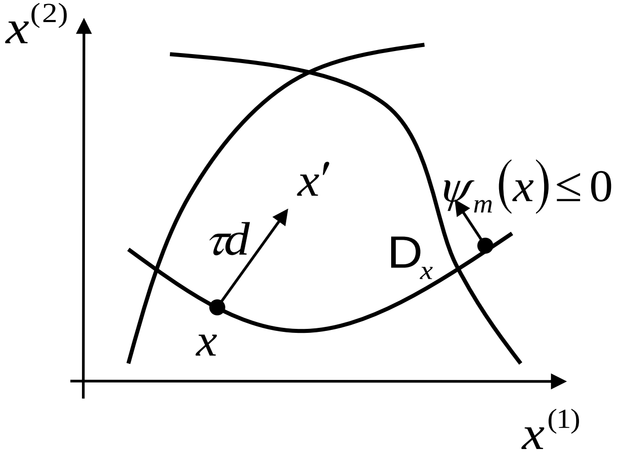
Active constraints – analytical conditions?

$$\forall m \in I(x)$$

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

$$x' = x + \tau d \in D_x, \quad \tau > 0$$

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




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

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

$d^T \nabla_x \psi_m(x) \le 0$ – analytical condition

# Optimization under inequality constraints

## Feasible directions

How to determine the set of feasible directions?

$$\psi_m(x) = 0$$

$$x' = x + \tau d \in \mathrm{D}_x$$

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

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

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

$$d^T \nabla_x \psi_m(x) \le 0$$ – analitical condition

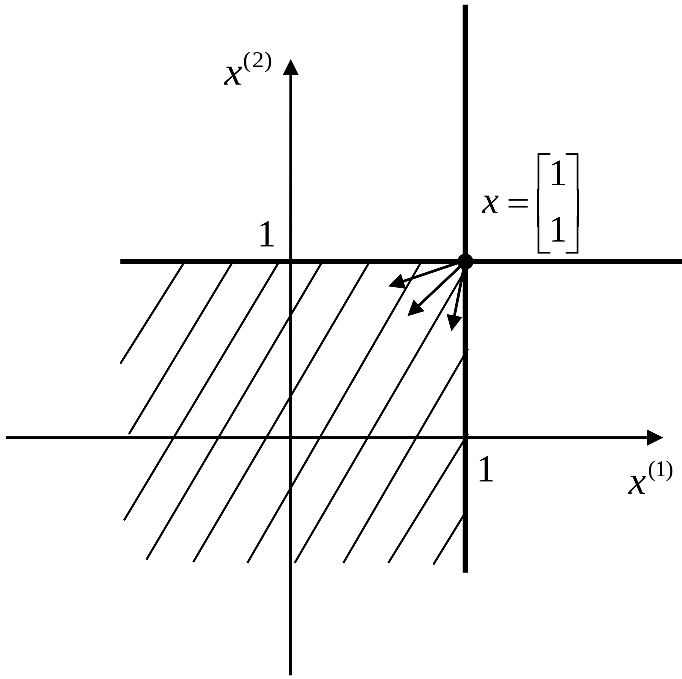
# Optimization under inequality constraints

## Feasible directions

Example

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

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



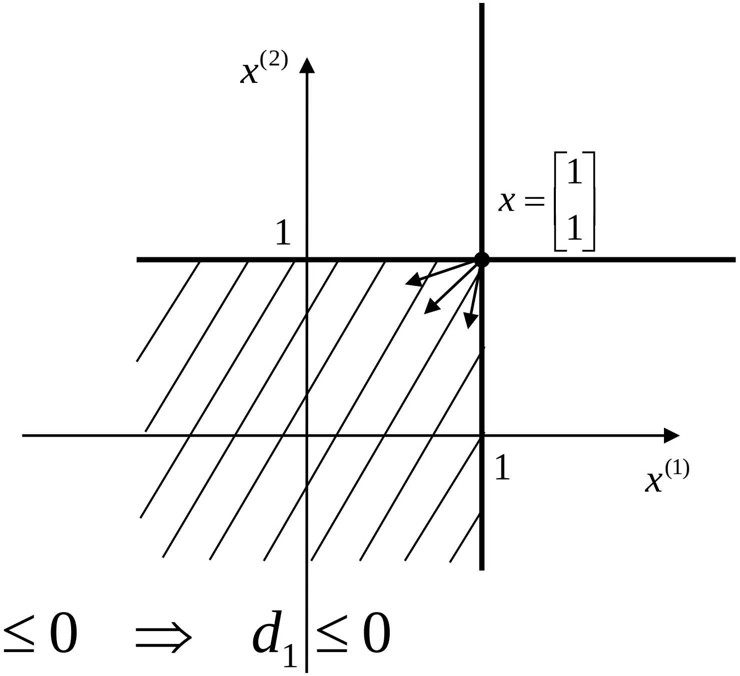
# Optimization under inequality constraints

## Kuhn – Tucker rolls

Example 1

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

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




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

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

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

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

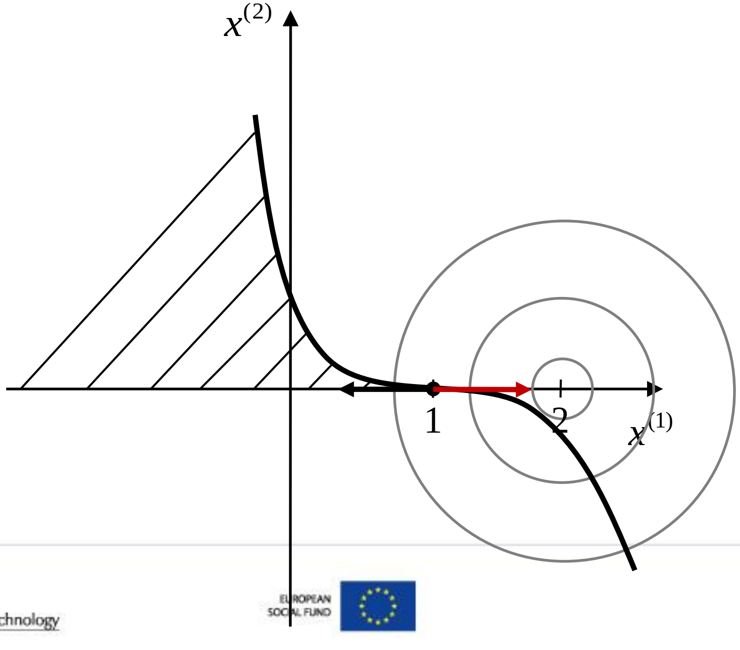
# Optimization under inequality constraints

## Feasible directions

$$\mathcal{D}(x)=\left\{d\in\mathbb{R}^{s}:\forall_{m}\in I(x),\ d^{T}\nabla_{x}\psi_{m}(x)\le 0\right\}$$

$\mathcal{D}(x)\neq D(x)$ leads to irregular case

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




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

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

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

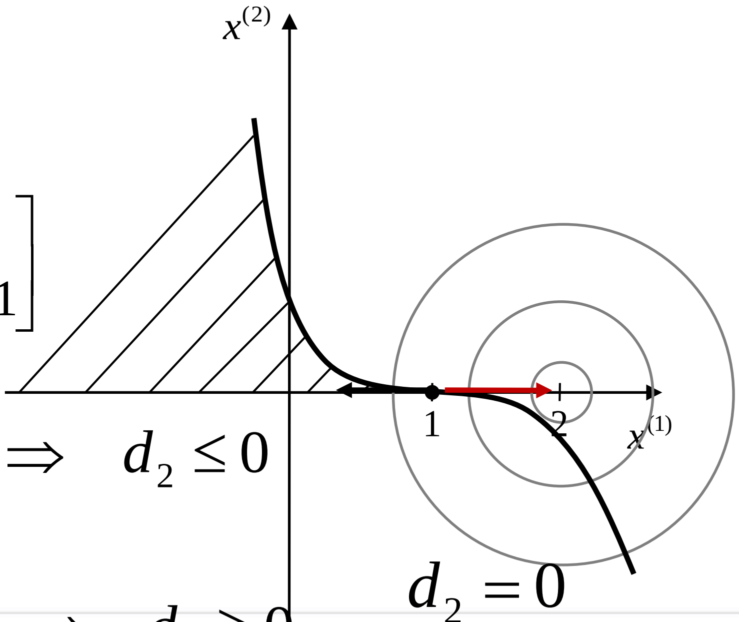
$$\psi_1(x) = x^{(2)} + \left(x^{(1)} - 1\right)^3 \le 0 \qquad \psi_2(x) = -x^{(2)} \le 0 \qquad F(x) = \left(x^{(1)} - 2\right)^2 + \left(x^{(2)}\right)^2$$

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

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

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

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




$$d_2 = 0$$

$$d_1 - any$$

# Optimization under inequality constraints Kuhn – Tucker rolls

## Regularity Conditions

1. Karlin: constraints $\psi_1(x), \psi_2(x), \ldots, \psi_M(x)$ - linear

2. Slater: constraints $\psi_1(x), \psi_2(x), \ldots, \psi_M(x)$ - convex functions and feasible set is not empty

3.Fiacco – Mac Cormica: in the optimal point gradients of all active constraints are linear independent, i.e.:

$$\forall\, m \in I(x^*) \quad \nabla_x \psi_m(x^*)\big|_{x=x^*}$$ are linear independent

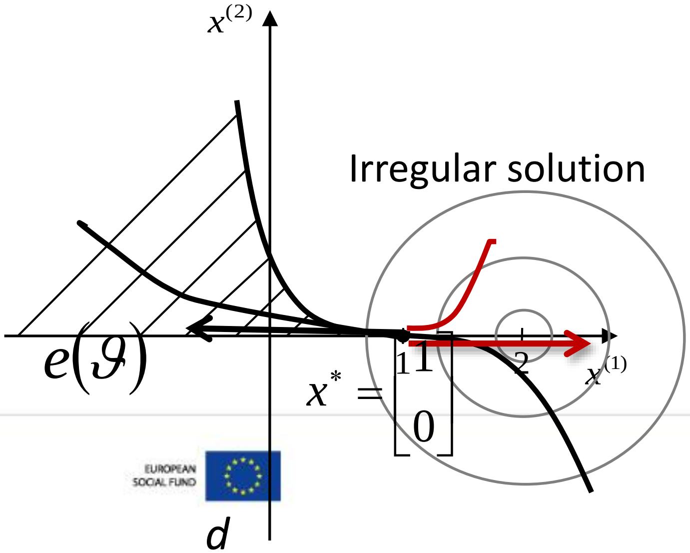
4. Zangwil: $\mathrm{D}(x^*) = \overline{D}(x^*)$

5. Kuhna – Tucker'a: for each direction $d \in \mathrm{D}(x^*)$ there exists regular curve starting in the point $x^*$ tangent to that direction

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

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

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



# Zadanie optymalizacji z ograniczeniami nierównościowymi

# Metoda Kuhna – Tucker'a

Warunki konieczna i wystarczające:

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

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

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

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

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

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

# Optimization under inequality constraints

## Kuhn-Tucker conditions

Sufficient condition of regularity:

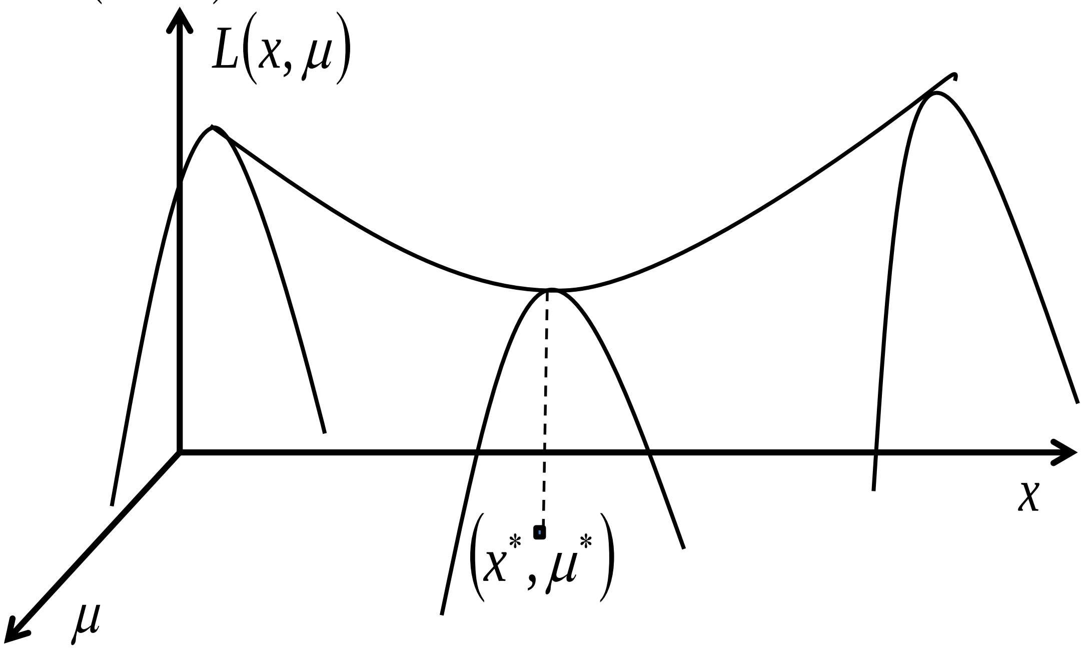
$F, \psi_1, \psi_2, \ldots, \psi_M$ – continuous and differentiable

$F$ – pseudo-convex

$\psi_1, \psi_2, \ldots, \psi_M$ – quasi-convex

# Saddle point

Saddle point $\left(x^*, \mu^*\right)$




$$L\left(x^*, \mu^*\right) \leq L\left(x, \mu^*\right) \quad \forall x \in D(x) \subseteq R^S$$

$$L\left(x^*, \mu\right) \leq L\left(x^*, \mu^*\right) \quad \forall \mu \geq 0_M$$

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

# Saddle point

Point $\left(x^{*},\mu^{*}\right)$ is the saddle poin $\left(x^{*}\in \mathrm{D}\,(x),\ \mu\geq 0_{M}\right)\ \Leftrightarrow$

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

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

$$x^{*}\rightarrow F(x^{*})=\min_{x^{*}\in \mathrm{D}_{x}} F(x)$$

$$\mathrm{D}_{x}=\left\{x\in \mathrm{R}^{S}:\psi_{1}(x)\leq 0,\psi_{2}(x)\leq 0,\ldots,\psi_{M}(x)\leq 0\right\}$$

# Special case

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

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

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

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

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

$$x^* \geq 0_S$$

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

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

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


Project co-financed from the EU European Social Fund

# Special case

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

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

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

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

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

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

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

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

# Analytical methods

## Disadvantages

It is hard to apply them if:

$F, \varphi, \psi$ are nonlinear

$\dim(x)$ is large

They cannot be applied if:

$F, \varphi, \psi$ are not differentiable

$F$ is not given by formula and it may only be measured for requested value of $x$

# Thank you for attention