Global Focus on Knowledge Lecture (The World of Mathematics #7) 2007-06-07

Mathematics of Optimization
-Viewpoint of Applied Mathematics
Logic of Optimization

Kazuo Murota

Department. of Mathematical Engineering and Information Physics (Faculty of Engineering)

Department. of Mathematical Informatics (Graduate School of Information Science and Tech.)

http://www.misojiro.t.u-tokyo.ac.jp/»murota

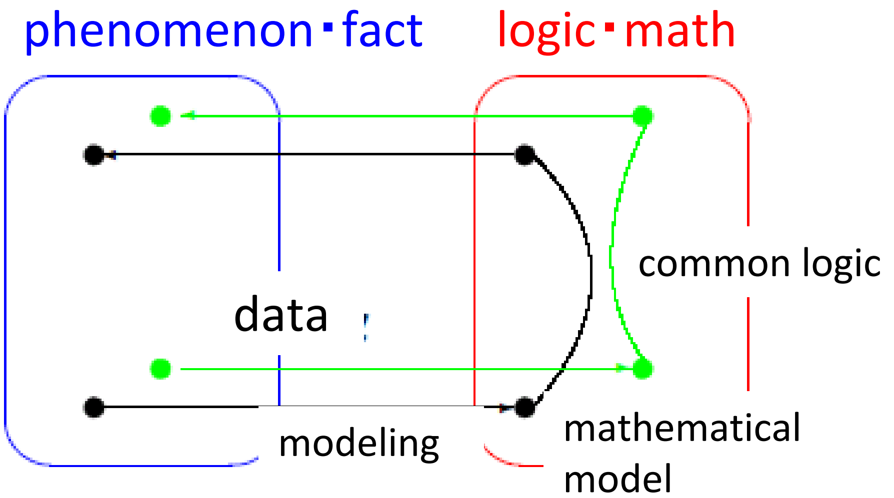
1. Convex function

2. Meaning of convex

3. Discrete convex

The World of Optimization
(Revision)

continuity／dispersion
linear／convex／non-linear



Modeling＋ Logics＋ Algorithm

Beautiful and Useful

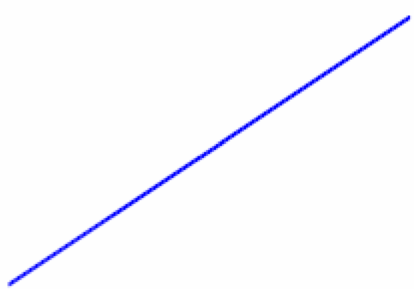
Development of Optimization (Metric Variable)

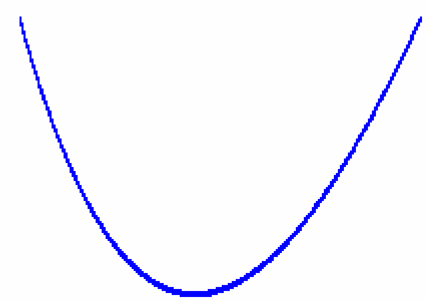
1947 linear planning Dantzig
1960 non-linear planning, Newton method
Powell, Fletcher
1970 $\boxed{\text{convex analysis, dual theorem}}$ Rockafellar
1979 ellipsoid method Khachiyan
1984 interior method Karmarkar
1995 semidefinite program
Alizadeh, Nesterov, Nemirovski

logic： linear／convex／non-linear
environment： enhancement of calculation power

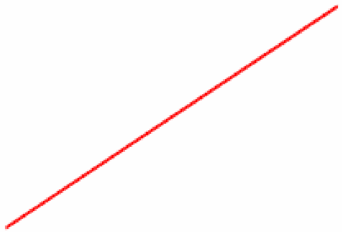
1. Convex Function

Linear to Convex

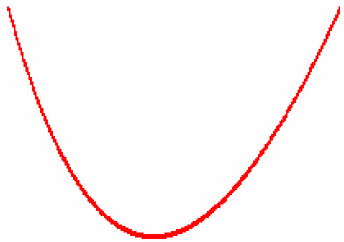




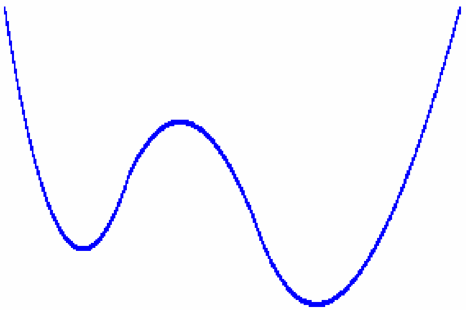
Convex Function (A Feature of Logic)



Linear function



Convex function



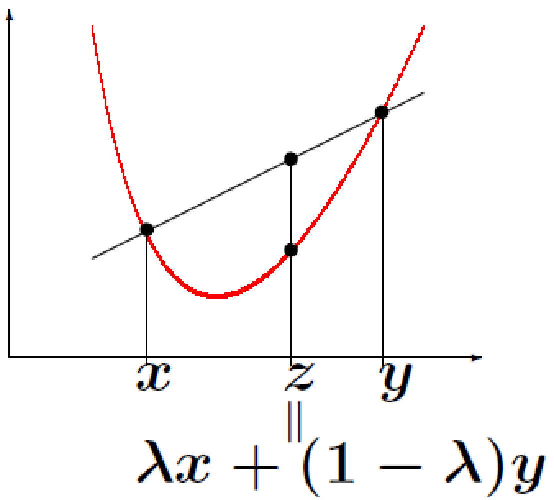
Non-convex function

$\longleftarrow$ Non-linear function $\longrightarrow$

f is convex function $\Longleftrightarrow$

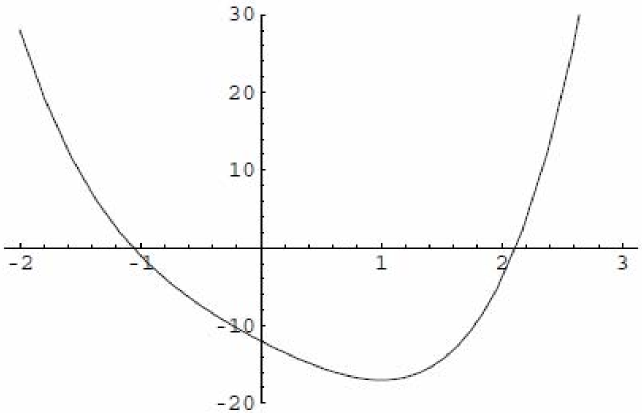
$$\lambda f(x) + (1-\lambda) f(y) \geq f(\lambda x + (1-\lambda) y)$$

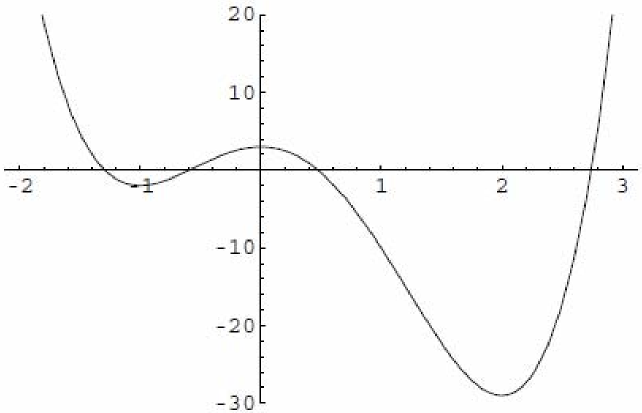
$$(0 < \forall \ \lambda < 1)$$



Convex Function and Non-Convex Function

$$f(x) = x^4 + 2x^2 - 8x - 12 \qquad f(x) = 3x^4 - 4x^3 - 12x^2 + 3$$

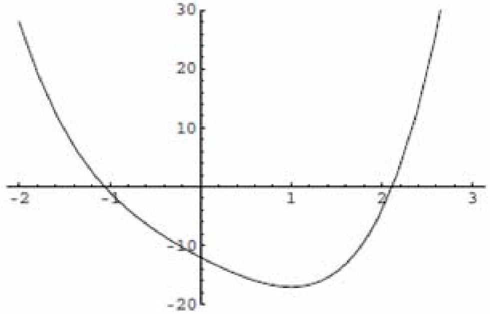




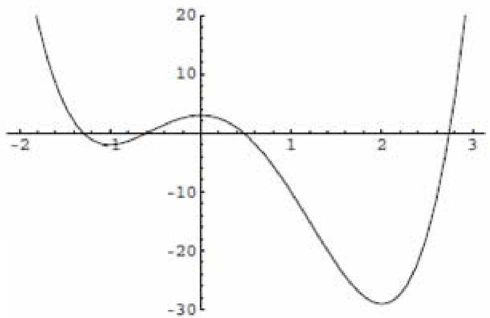
How to Check Convex Function (For a Smooth Function)

f is convex function $\Longleftrightarrow$ for all x , $f''(x) \geq 0$

$$f(x) = x^4 + 2x^2 - 8x - 12$$
$$\Rightarrow f'(x) = 4x^3 + 4x - 8$$
$$\Rightarrow f''(x) = 12x^2 + 4 > 0 \Rightarrow$$ Convex



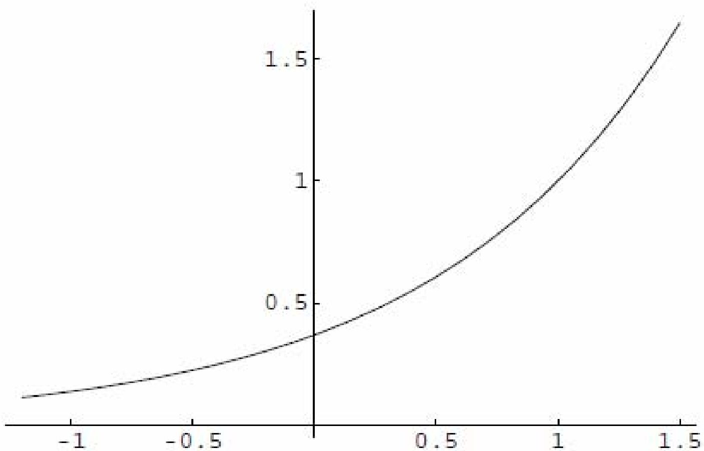
$$f(x) = 3x^4 - 4x^3 - 12x^2 + 3$$
$$\Rightarrow f'(x) = 12x^3 - 12x^2 - 24x$$
$$\Rightarrow f''(x) = 36x^2 - 24x - 24$$
$$\Rightarrow f''(0) = -24 \quad < 0 \quad \Rightarrow$$ Not convex

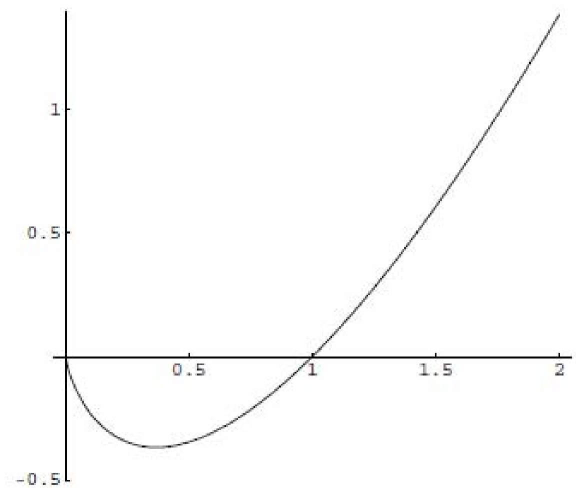


Examples of Convex Function

$$f(x) = \exp(x - 1)$$

$$f(x) = x \log x$$

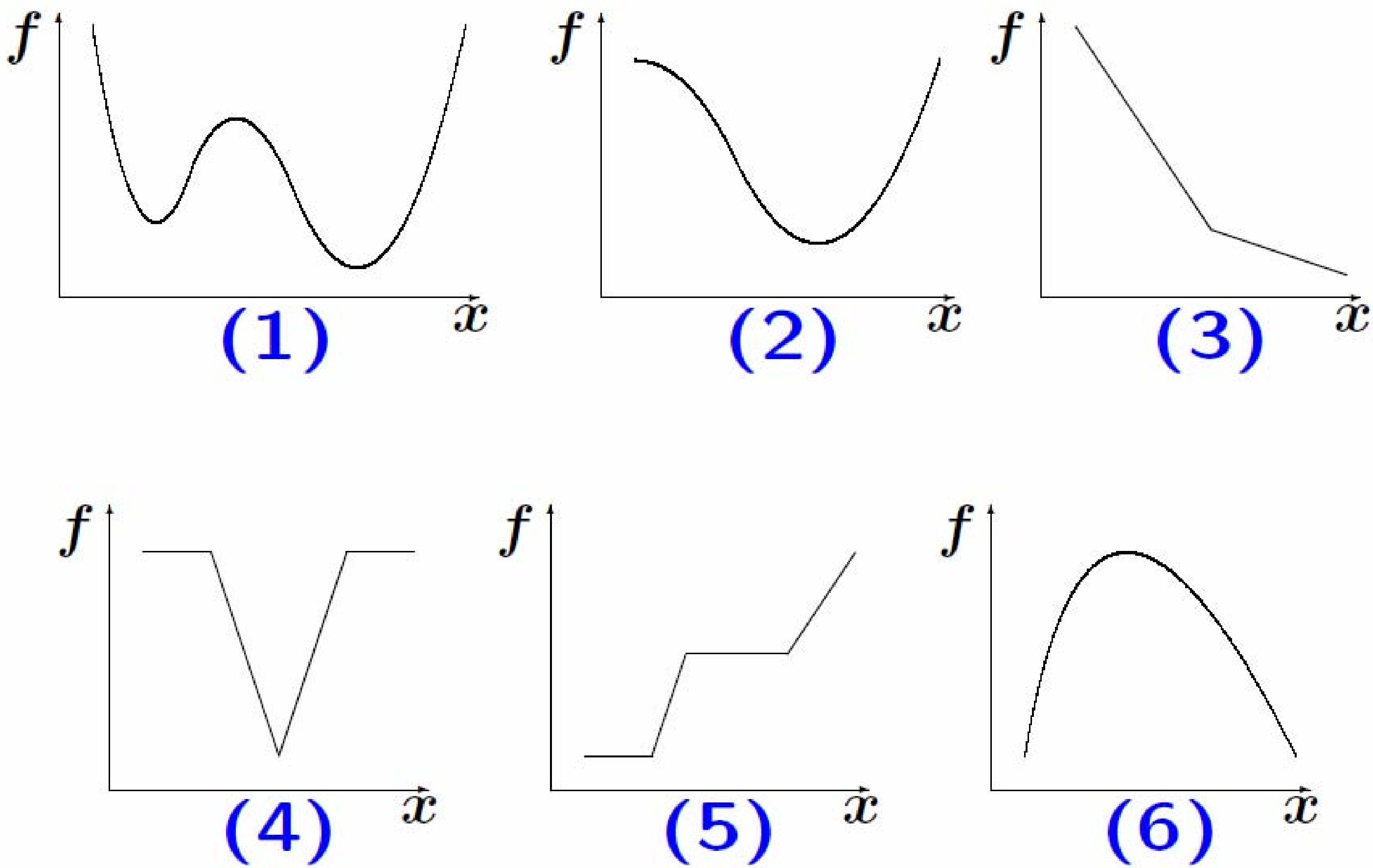




$$f''(x) = \exp(x - 1) \; > 0$$

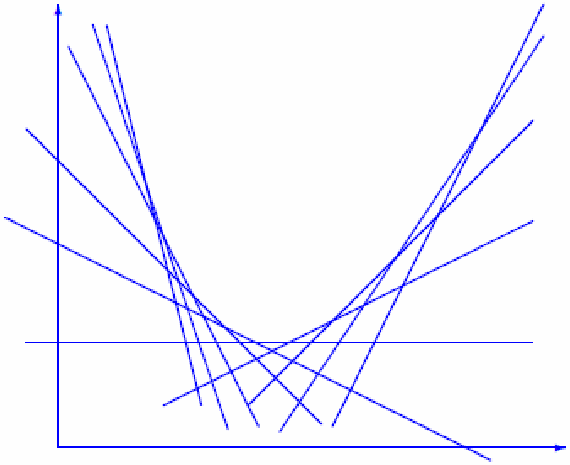
$$f''(x) = 1/x \; > 0$$

Convex Function and Non-convex Function



2. Meaning of Convex

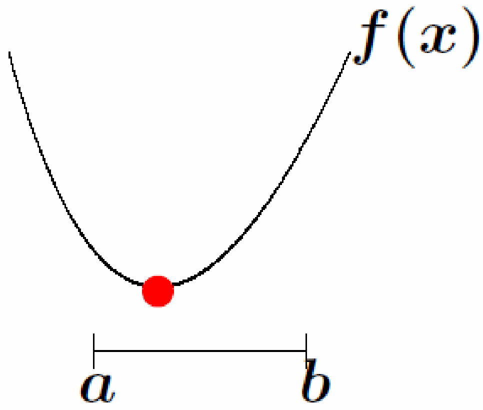
Duality



Minimization Problem

minimization $f(x)$

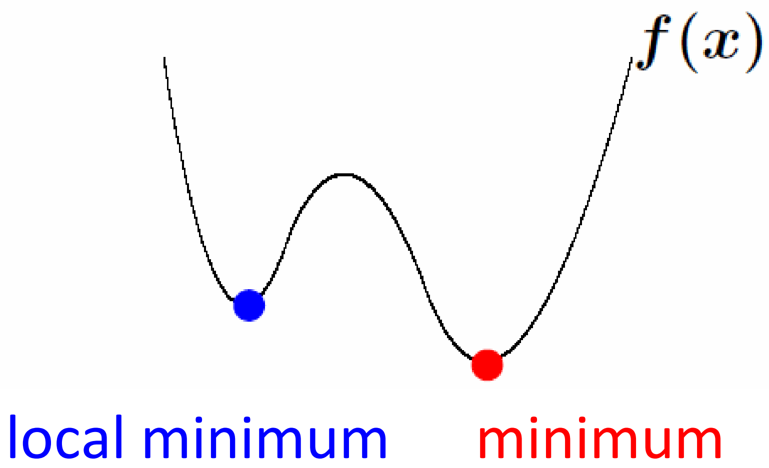
constraints $a \leq x \leq b$



---

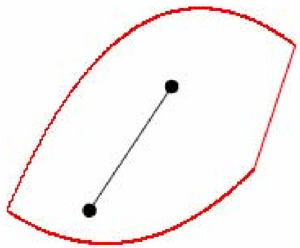
minimization $f(x_1, x_2, \ldots)$

constraints $(x_1, x_2, \ldots) \in S$

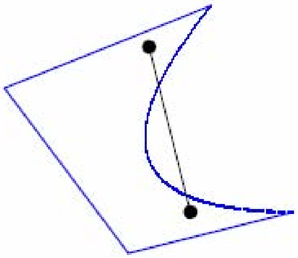


Convex programming problem: minimize convex function among convex set

Convex Set and Convex Function

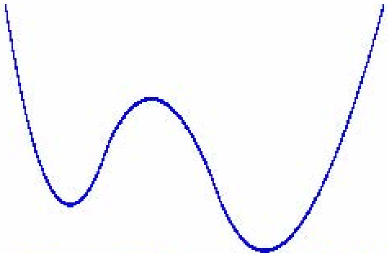


convex set



non-convex set

convex function



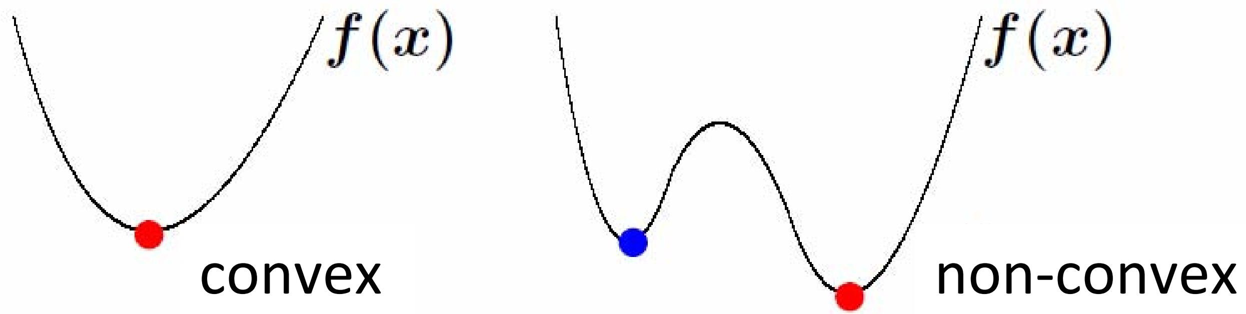
non-convex function

$f : \mathrm{R}^n \rightarrow \mathrm{R}$ is convex function $\Longleftrightarrow$

$$\lambda f(x) + (1-\lambda) f(y) \geq f(\lambda x + (1-\lambda) y) \qquad (0 < \forall \ \lambda < 1)$$

Meaning of Convex Function Upon Optimization

local minimum = global minimum



Various Dualities

- Legendre transformation
- max-min theorem

Duality of Linear Programming

main problem

$$\text{Min. } c^\top x \quad \text{s.t. } Ax = b,\ x \geq 0$$

dual problem

$$\text{Max. } b^\top y \quad \text{s.t. } A^\top y \leq c$$

$$\begin{array}{llrcl}
\textbf{Min.} & 12\,x_1 + 4\,x_2 + 3\,x_3 & & & \\
\textbf{s.t.} & x_1 + x_2 + x_3 & & = & 1 \\
 & 4\,x_1 + x_2 & & = & 2 \\
 & x_1,\ x_2,\ x_3 & & \geq & 0
\end{array}$$

$$\begin{array}{llrcl}
\textbf{Max.} & y_1 + 2\,y_2 & & & \\
\textbf{s.t.} & y_1 + 4\,y_2 & & \leq & 12 \\
 & y_1 + y_2 & & \leq & 4 \\
 & y_1 & & \leq & 3
\end{array}$$

duality theorem:

min. of main problem = max. of dual problem

How to Use Duality Theorem (Accuracy Assured)

$$\alpha = \textbf{Min.}\ c^\top x$$
$$\textbf{s.t.}\quad Ax = b,\ x \geq 0$$

duality theorem for linear programming

$$\boxed{\min\{c^\top x \mid Ax = b, x \geq 0\} = \alpha = \max\{b^\top y \mid A^\top y \leq c\}}$$

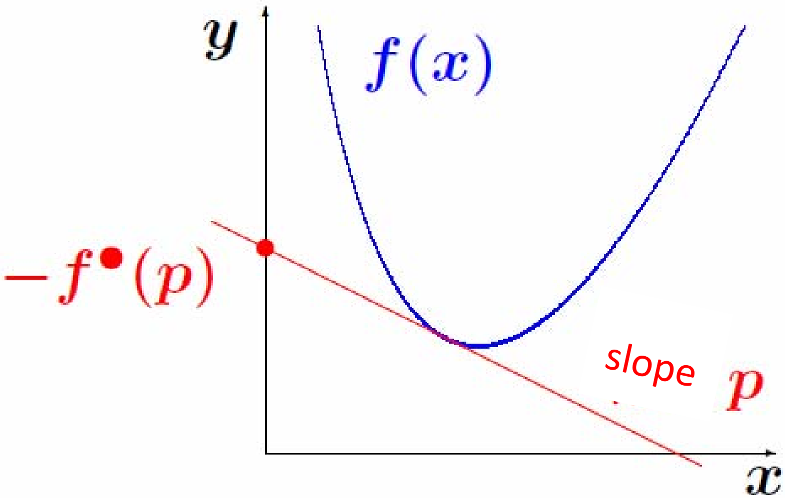
For x that satisfy conditions $0 \leq c^\top x - \alpha \leq \boxed{?}$

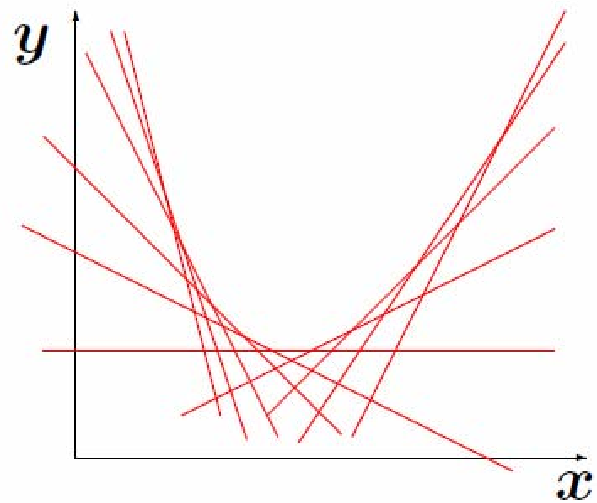
error limit

$\Longrightarrow$ For any y $\quad \boxed{?} = c^\top x - b^\top y$

Legendre Transformation

convex function = envelope curve of a tangent line





Legendre transformation $$f^{\bullet}(p) = \max_{x} \{p \cdot x - f(x)\}$$

$\boxed{\text{theorem: convex func. } f \mapsto f^{\bullet} \mapsto f^{\bullet\bullet} = f}$

Calculation of Legendre Transformation

$$f^{\bullet}(p) = \max_{x} \ \{\underline{p \cdot x - f(x)}\}$$

$$\Rightarrow \ \frac{d}{dx}(\underline{p \cdot x - f(x)}) = p - f'(x) = 0$$

$$\Rightarrow \ p = f'(x)$$

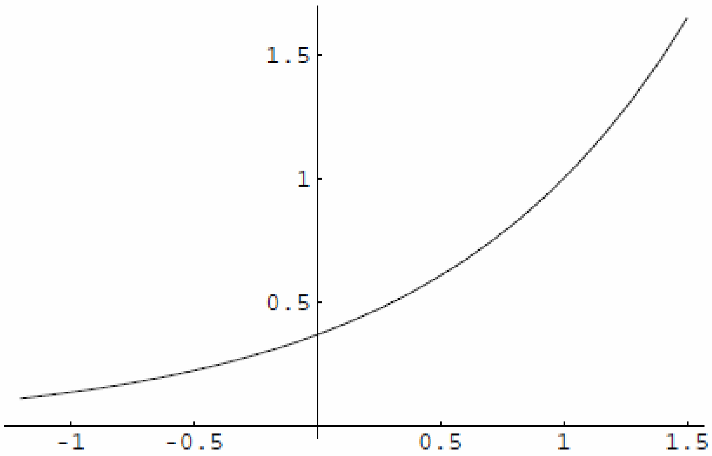
$$\Rightarrow \ x = x(p)$$

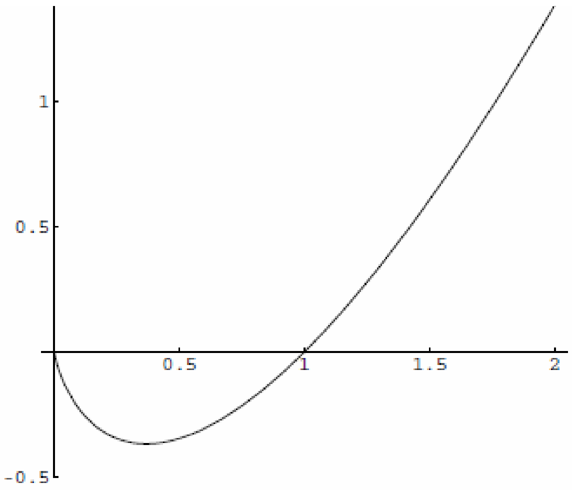
$$\Rightarrow \ f^{\bullet}(p) = p \cdot x(p) - f(x(p))$$

A. M. Legendre
(1752~1833)

Picture from Wikipedia

Examples of Legendre Transformation (1)





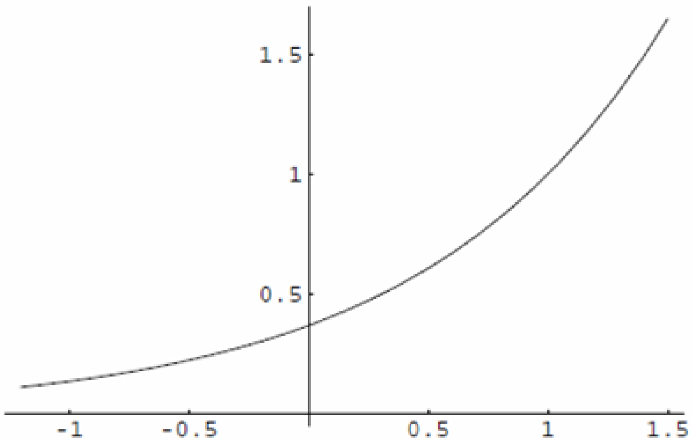
$$f(x) = \exp(x-1) \quad \longrightarrow \quad g(p) = p \log p$$

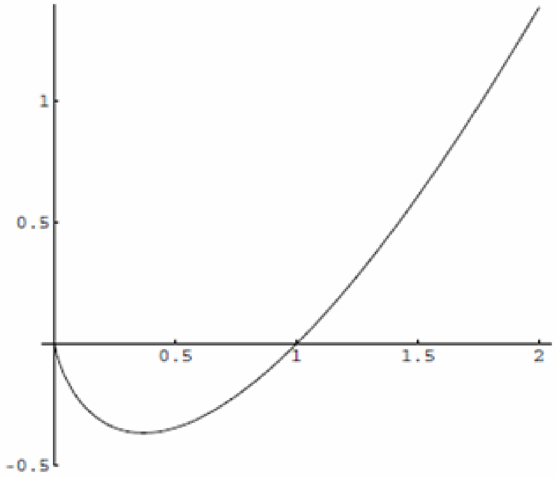
$$f^{\bullet}(p) = g(p)$$

$$p = f'(x) = \exp(x-1) \;\Rightarrow\; x = 1 + \log p$$

$$\Rightarrow\; f^{\bullet}(p) = p \cdot x - f(x) = p \cdot (1 + \log p) - p = p \log p$$

Examples of Legendre Transformation (2)





$$f(x) = \exp(x-1) \qquad \longleftarrow \qquad g(p) = p\log p$$

$$f(x) = g^{\bullet}(x)$$

$$x = g'(p) = 1 + \log p \;\Rightarrow\; p = e^{x-1}$$

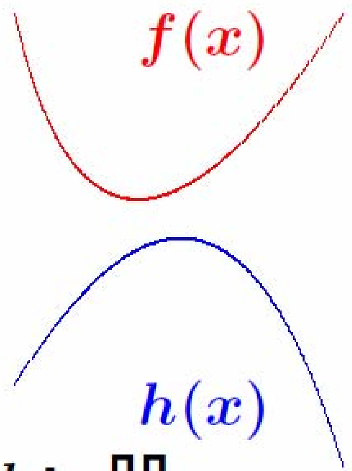
$$\Rightarrow\; g^{\bullet}(x) = p \cdot x - g(p) = xe^{x-1} - (x-1)e^{x-1} = e^{x-1}$$

Duality Theorem (From Linear to Convex)

Duality theorem for linear programming:

$$\min\{c^\top x \mid Ax = b, x \geq 0\} = \max\{b^\top y \mid A^\top y \leq c\}$$

Legendre transformation



$$f^\bullet(p) = \max_x\{p \cdot x - f(x)\}$$

$$h^\circ(p) = \min_x\{p \cdot x - h(x)\}$$

Fenchel duality theorem

f: 凸 h: 凹

$$\min_x\{f(x) - h(x)\} = \max_p\{h^\circ(p) - f^\bullet(p)\}$$

How to Use Fenchel Duality Theorem (Accuracy Assured)

$$\alpha = \min_x \{f(x) - h(x)\}$$

Fenchel duality theorem

$$\boxed{\min_x \{f(x) - h(x)\} = \alpha = \max_p \{h^\circ(p) - f^\bullet(p)\}}$$

For any x , $0 \leq [f(x) - h(x)] - \alpha \leq \boxed{?}$

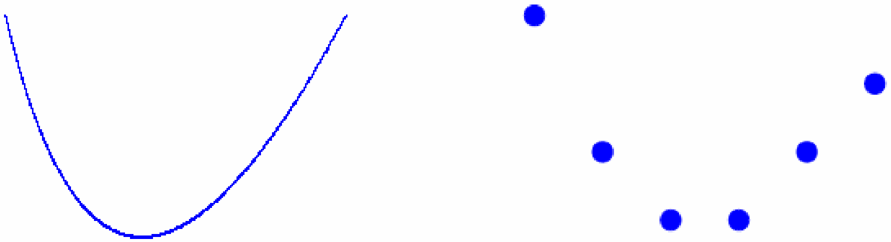
error limit

$\Longrightarrow$ For any p ,

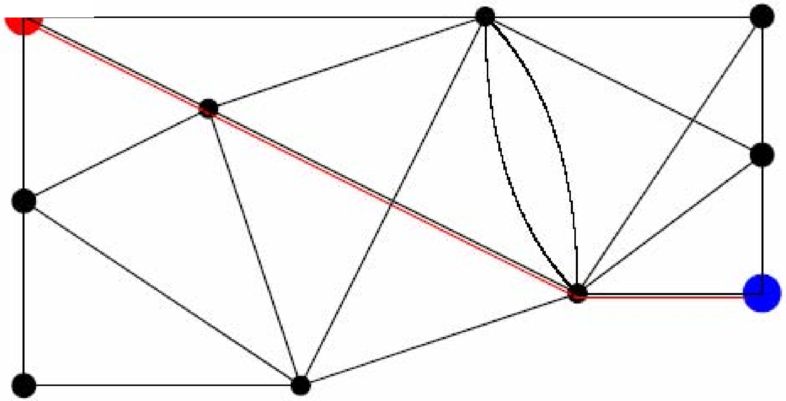
$$\boxed{?} = [f(x) - h(x)] - [h^\circ(p) - f^\bullet(p)]$$

3. Discrete Convex

continuation to discrete



Discrete Optimization



Shortest path problem traveling salesman problem

easily solved (empirical fact) difficult to solve

Dijkstra method (algorithm logic) NP difficulty

Viewpoint of convex analysis

Viewpoint of complexity

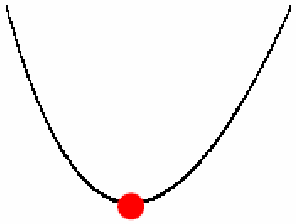
Viewpoint of Discrete Convex Analysis

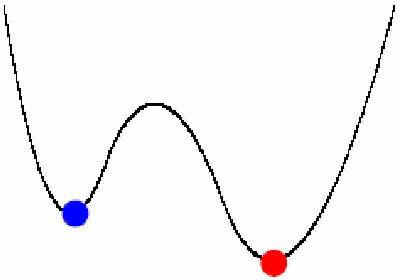
Shortest path problem traveling salesman problem

easily solved (empirical fact) difficult to solve

polynomial time (algorithm logic) NP difficulty

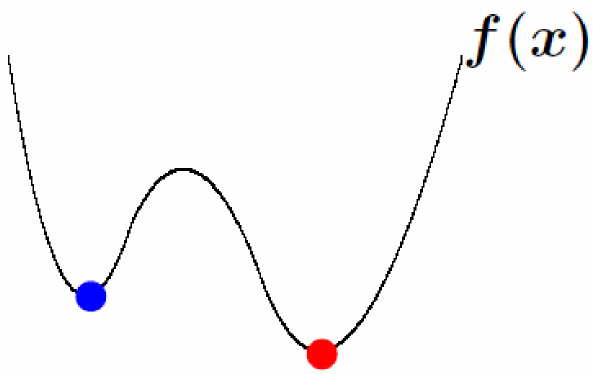
convex not convex

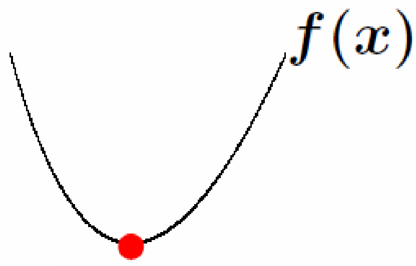




Optimization

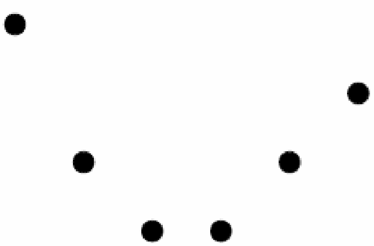
Easy to Solve



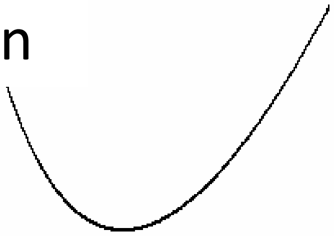


convex function

discrete



$\Longleftrightarrow$ continuation



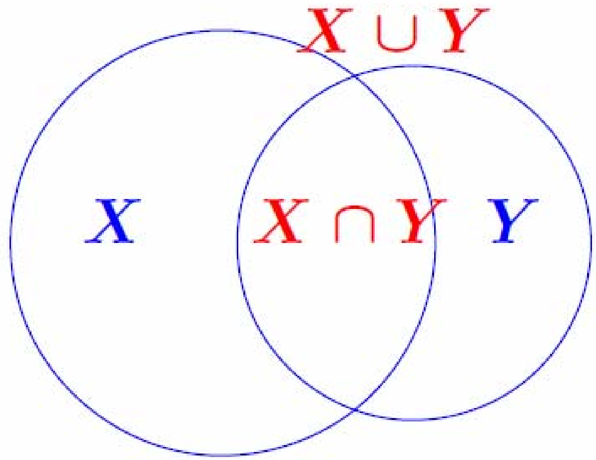
discrete structure
similar to convex function

convex function
with discrete structure

The History of Discrete Convex Analysis

1935	matroid	Whitney
1965	$\boxed{\text{submodullar function}}$	Edmonds
1975	application of matroid	Iri, Tomizawa, Recski
1983	$\boxed{\text{submodullar function}}$ and convexity	Lovász, Frank, Fujishige
1996	propounding of discrete convex analysis	Murota
2000	$\boxed{\text{submodullar function}}$ and minimization algorithm	Iwata, Fujishige, Fleischer, Schrijver

Submodullar Set Function



Number of elements in set X $|X|$:

$$|X| + |Y| = |X \cup Y| + |X \cap Y|$$

The set function $\rho(X)$ is a submodullar :

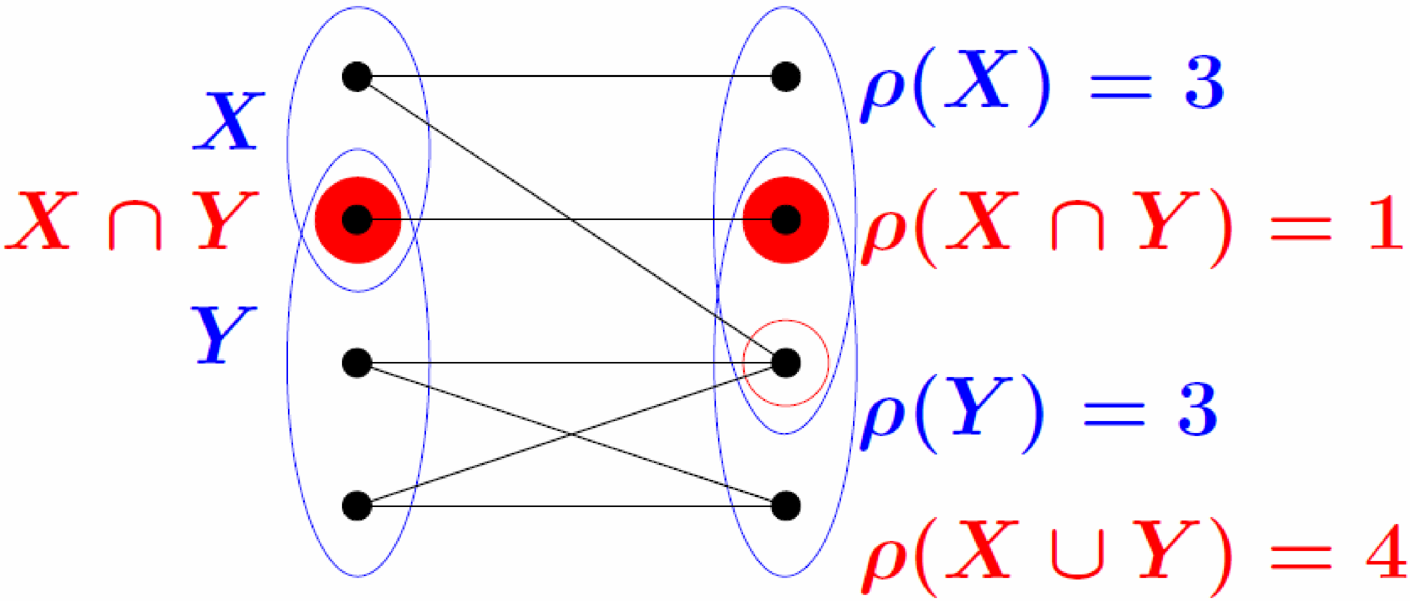
$$\rho(X) + \rho(Y) \geq \rho(X \cup Y) + \rho(X \cap Y)$$

Ex 1 $\rho(X) = \log |X|$

Ex 2 $\rho(X) = \sqrt{|X|}$

An Example of Submodullar Function

Ex 3 $\rho(X) = X$ Number of friends



$$\rho(X) + \rho(Y) \geq \rho(X \cup Y) + \rho(X \cap Y)$$

Submodullar Function and Convex Function (1980s)

$$\rho(X) + \rho(Y) \geq \rho(X \cup Y) + \rho(X \cap Y)$$

- minimization/maximization algorithm

  minimization $\Rightarrow$ polynomial time, maximization $\Rightarrow$ NP difficulty

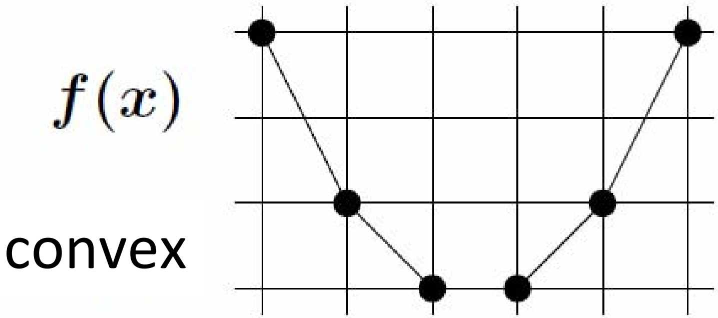
- convex extension (Lovasz)

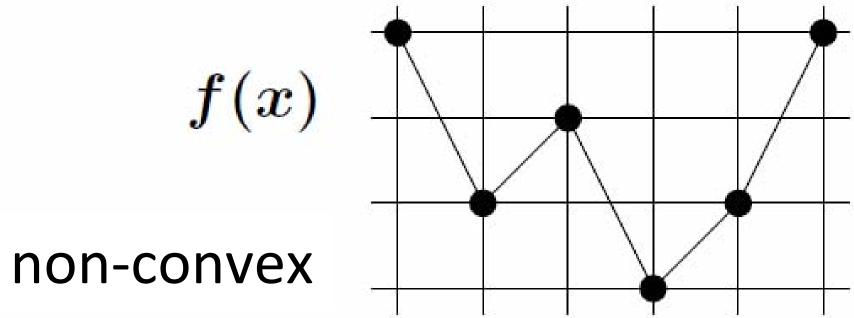
  Set function is submodullar $\Longleftrightarrow$ Lovasz extension is convex

- duality theorem (Edmonds, Frank, Fujishige)

Duality of submodullar function
= convexity + discreteness

Convex Structure in Discrete Optimization





Convex structure = a problem that can be solved

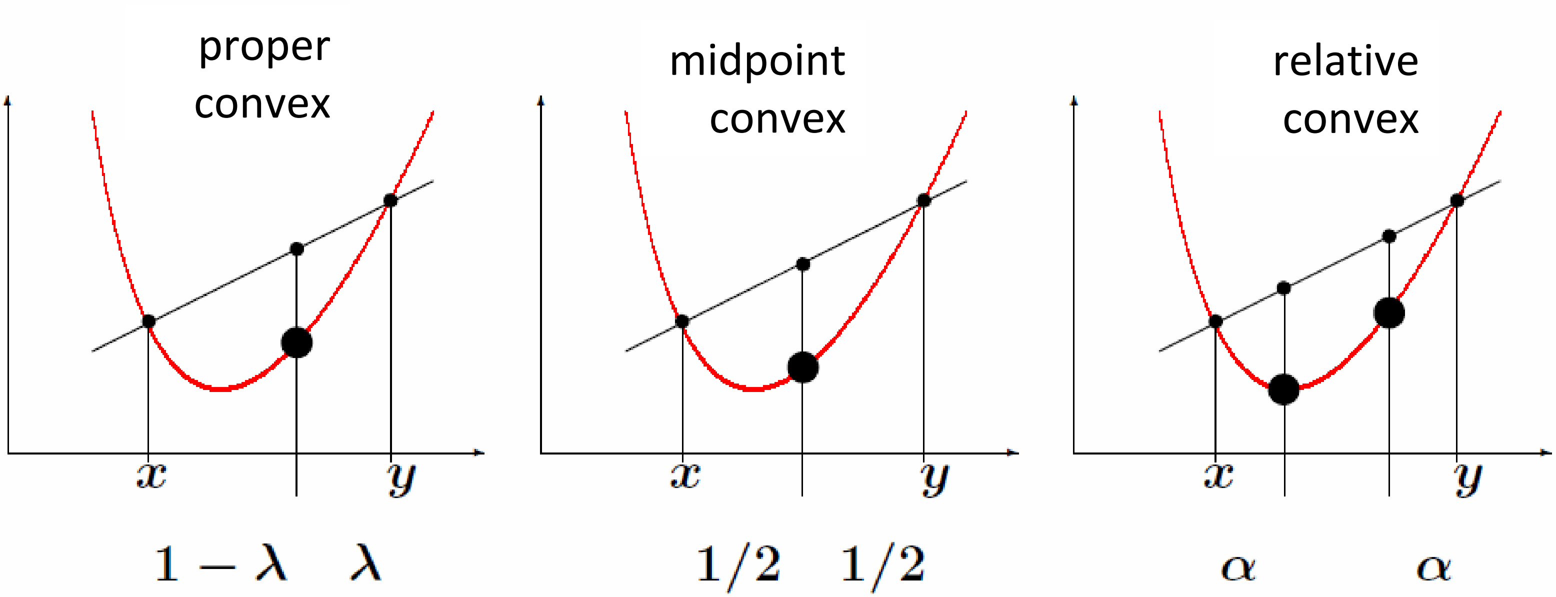
- local/global optimization
- duality

Characteristics of Convex Function – Preparation for Discretization

proper convex $\lambda f(x) + (1-\lambda) f(y) \geq f(\lambda x + (1-\lambda) y)$

$\Leftrightarrow$ midpoint convex $f(x) + f(y) \geq 2 f(\frac{x+y}{2})$

$\Leftrightarrow$ relative convex $f(x)+f(y) \geq f(x-\alpha(x-y))+f(y+\alpha(x-y))$



Conceptual Configuration by Discretization

proper convex $\lambda f(x) + (1-\lambda) f(y) \geq f(\lambda x + (1-\lambda) y)$

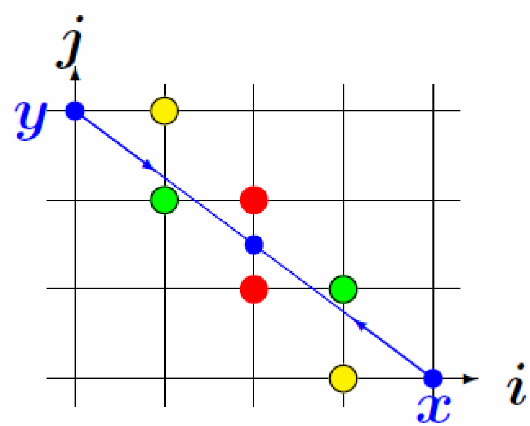
midpoint convex $f(x) + f(y) \geq 2 f\left(\frac{x+y}{2}\right)$

$\longrightarrow$ discrete midpoint convex $\geq f\left(\left\lceil \frac{x+y}{2} \right\rceil\right) + f\left(\left\lfloor \frac{x+y}{2} \right\rfloor\right)$

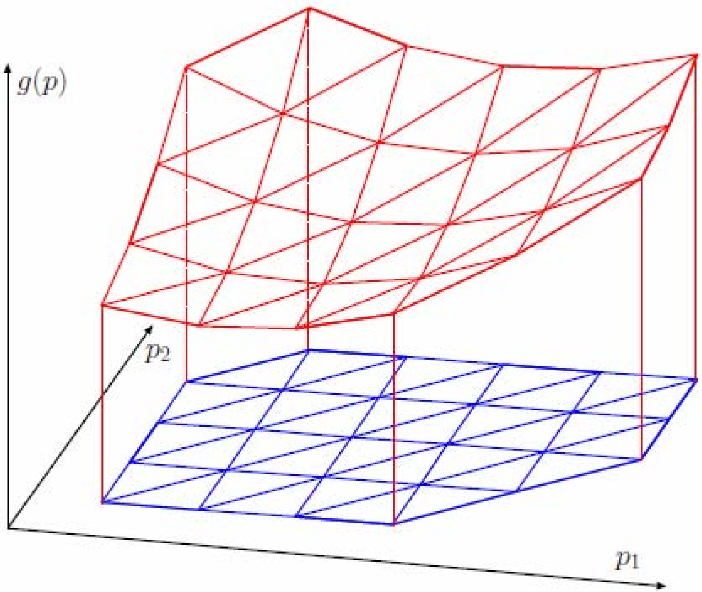
relative convex $f(x) + f(y) \geq f(x - \alpha(x-y)) + f(y + \alpha(x-y))$

$\longrightarrow$ exchanging axiom $\geq \min \left[f(x-e_i) + f(y+e_i), \min_{x_j < y_j} \{ f(x - e_i + e_j) + f(y + e_i - e_j) \} \right]$

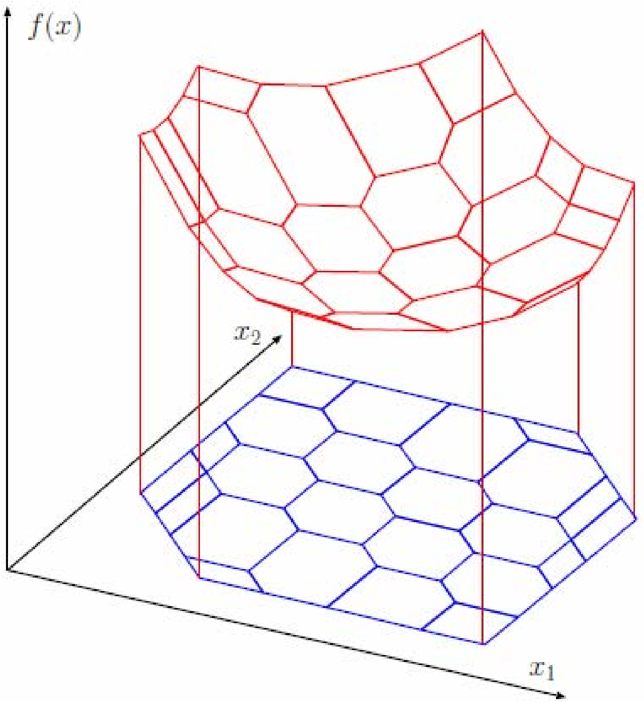
continuation $\mathbf{R}^n \to \mathbf{R}$	discretization	$\mathbf{Z}^n \to \mathbf{R}$
midpoint convexity	$\longrightarrow$	discrete midpoint convex
$\Updownarrow$	discretization	$\mathrm{L}^\natural$ convex function
convexity		
$\Updownarrow$	discretization	$\mathrm{M}^\natural$ convex function
relative convexity	$\longrightarrow$	exchanging axiom



Graphs of Discrete Convex Functions



$\mathrm{L}^\natural$ convex function



$\mathrm{M}^\natural$ convex function

Viewpoint of Convex Analysis is Effective in Discretization

Viewpoint of convex analysis — Viewpoint of complexity

$\Downarrow$ Discrete convex analysis

Viewpoint of convex analysis — Viewpoint of complexity

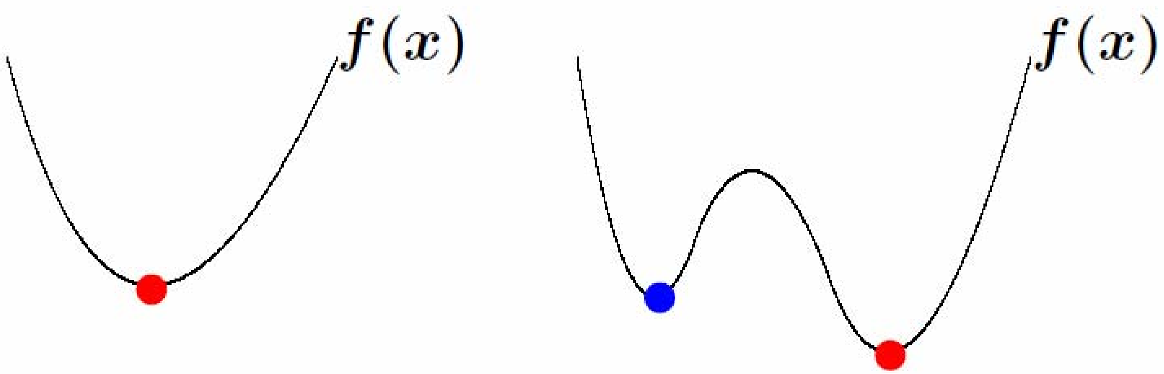
Overview of Discrete Convex Analysis

$\boxed{\text{L convex function} \leftrightarrow \text{M convex function}}$

logical completeness	● local optimum ⇔ global optimum
	● conjugacy: Legendre transformation
	● dual theorem (discretization, Fenchel duality)
	● minimization algorithm
sectoral transversity	● network flow (electric circuit)
	● OR (queue, resource allocation)
	● polynomial matrix
	● utility function, game theory

Optimality Criteria

local optimum and global optimum



convex

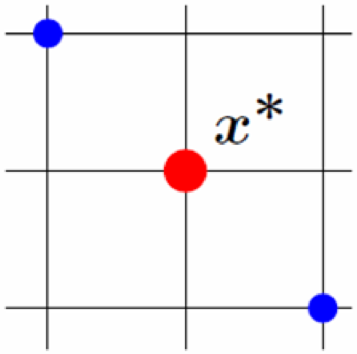
not convex

Optimality Criteria for M Convex Functior (discrete variable)

theorem : $f : \mathrm{Z}^n \to \mathrm{R}$ M convex function

x^*: Global minimum

$\Longleftrightarrow$ Local minimum $f(x^*) \leq f(x^* - e_i + e_j) \qquad (\forall i, j)$

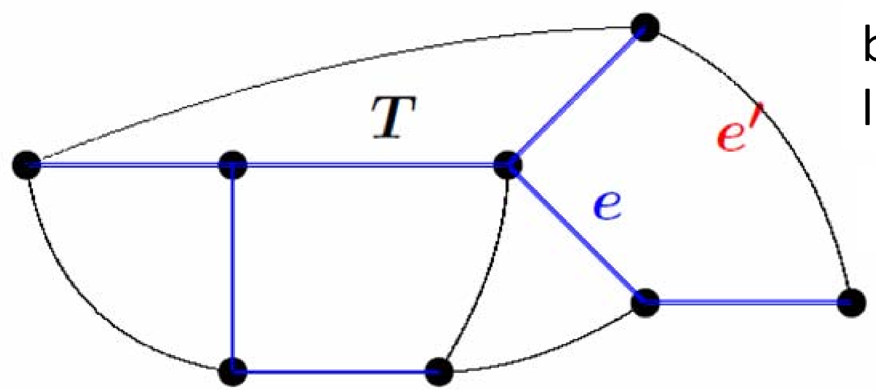


Ex : $x^* + (0, 1, 0, 0, -1, 0, 0, 0)$

By n^2 times function value evaluation, check can be done.

Optimality Criteria for Minimum Spanning Tree Problem

M convex function for an example



branch length function $d : E \to \mathbb{R}$

Tree T 's total length

$$\tilde{d}(T) = \sum_{e \in T} d(e)$$

Theorem：

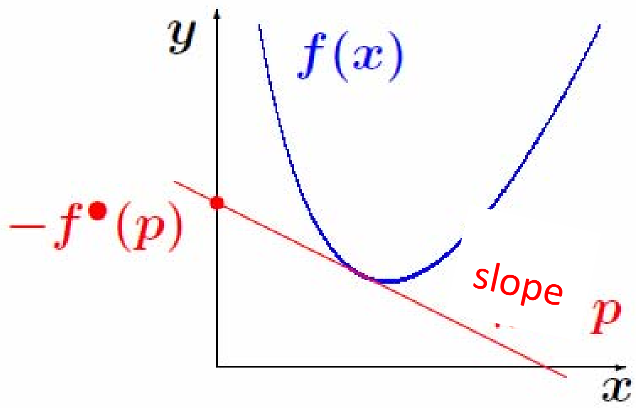
T is the smallest

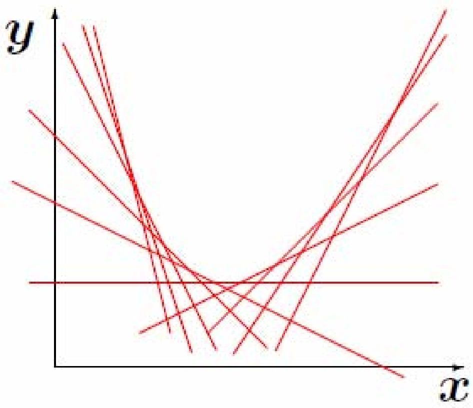
$$\iff \tilde{d}(T) \leq \tilde{d}(T - e + e') \qquad (e' \in C(T \mid e))$$

$$\iff d(e) \leq d(e') \qquad (e' \in C(T \mid e))$$

Discrete Legendre Transformation

convex function = envelope curve of a tangent line





discrete Legendre transformation

$$f^{\bullet}(p) = \max_{x \in \mathbf{Z}^n} \{p \cdot x - f(x)\}$$

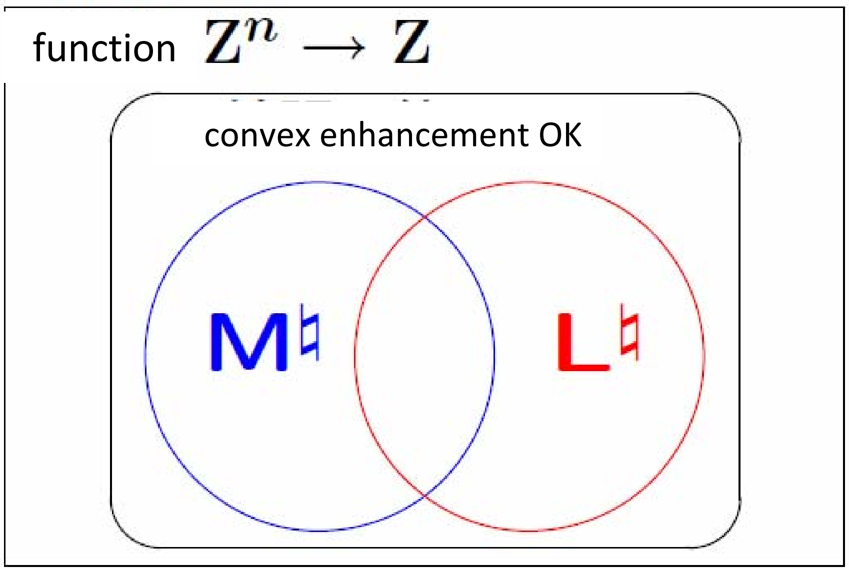
for

$\Rightarrow$ $f : \mathbf{Z}^n \to \mathbf{Z}$, $f^{\bullet} : \mathbf{Z}^n \to \mathbf{Z}$

Discrete Duality (Conjugacy Theorem)

theorem : M convex function $f \mapsto$ L convex function $f^{\bullet} \mapsto f^{\bullet\bullet} = f$

L convex function $f \mapsto$ M convex function $f^{\bullet} \mapsto f^{\bullet\bullet} = f$



Max・Min Theorem

discrete Legendre transformation $f, h : \mathrm{Z}^n \to \mathrm{Z}$

$$f^\bullet(p) = \max\{p \cdot x - f(x) \mid x \in \mathrm{Z}^n\}$$

$$h^\circ(p) = \min\{p \cdot x - h(x) \mid x \in \mathrm{Z}^n\}$$

Fenchel duality theorem $\quad f$: M♮凸 $\quad h$: M♮凹

$$\min_{x \in \mathrm{Z}^n} \{f(x) - h(x)\} = \max_{p \in \mathrm{Z}^n} \{h^\circ(p) - f^\bullet(p)\}$$

$\Uparrow$

Self-conjugating ($f^\bullet$: L♮凸 $\quad h^\circ$: L♮凹)

	【Continuous】	【Discrete】
1947	-----------linear project---------------	
1970	$\boxed{\text{convex analysis}}$	$\boxed{\text{polynomial, NP perfectness}}$
	duality theorem	submodullar function
1980	---------ellipsoid method--------------	
1984	interior method	
1995	semidefinite program	approximation algorithm
2000		discrete convex analysis

convex analysis (duality theorem) - structure

complexity theorem (algorithm) - construction

Summary

Next Time: Algorithm

continuous

discretization

local optimum

dual theorem