Mathematics of Optimization
-Viewpoint of Applied Mathematics
Model and Data

Kazuo Murota

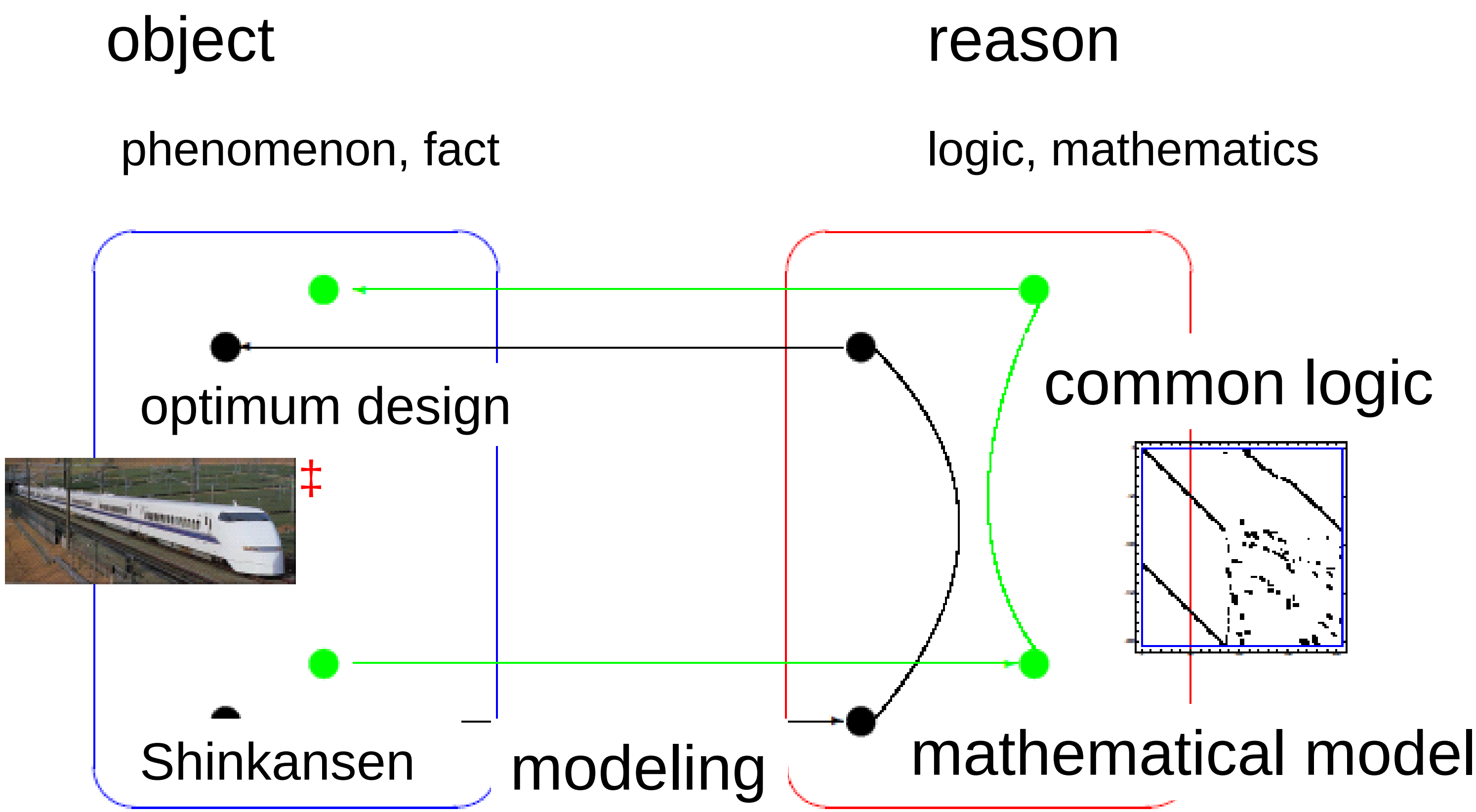
Department. of Mathematical Engineering and Information Physics (Faculty of Engineering)

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

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

- 1. Optimization
- 2. Model

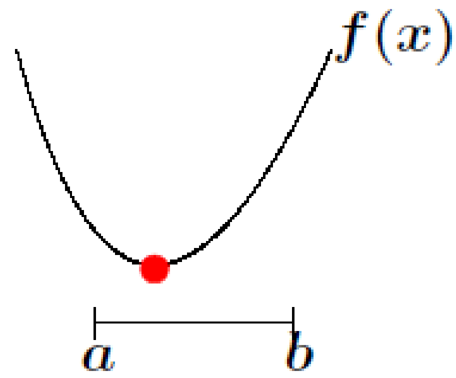
Applied Mathematics･Mathematical Engineering
— Discipline of Methodology



1. Optimization

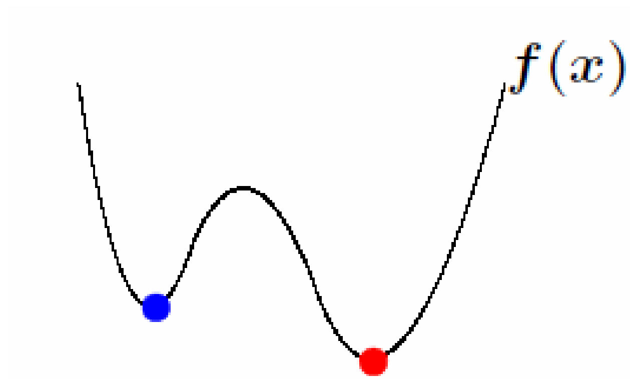
Maximum and Minimum Problem

minimization $f(x)$
constraints $a \cdot x \cdot b$



---

minimization $f(x1; x2; : : :)$
constraints $(x1; x2; : : :)\ 2\ S$

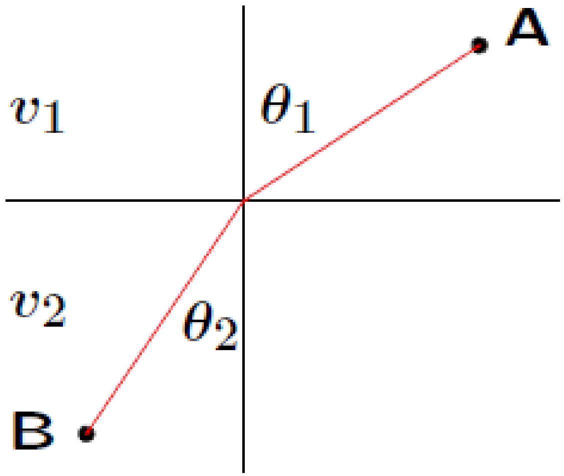


What are *f* and *S* ?

local minimum minimum

Nature Chooses the Best— Variation Principle

Snell's Law: $\frac{\sin\theta_1}{\sin\theta_2} = \frac{v_1}{v_2}$



● Shape of $f'(x) = 0$

Fermat's Principle: the path taken between two points by a ray of light is the path that can be traversed in the least time

● Shape of $f(x) \rightarrow \min$

- least-action principle：motion path＝minimization of action
- energy principle：equilibrium state＝minimization of energy ⇒ variation principle

Man Chooses the Best— Optimum Design

‡

0 Series (1964)

‡

300 Series (1992)

‡

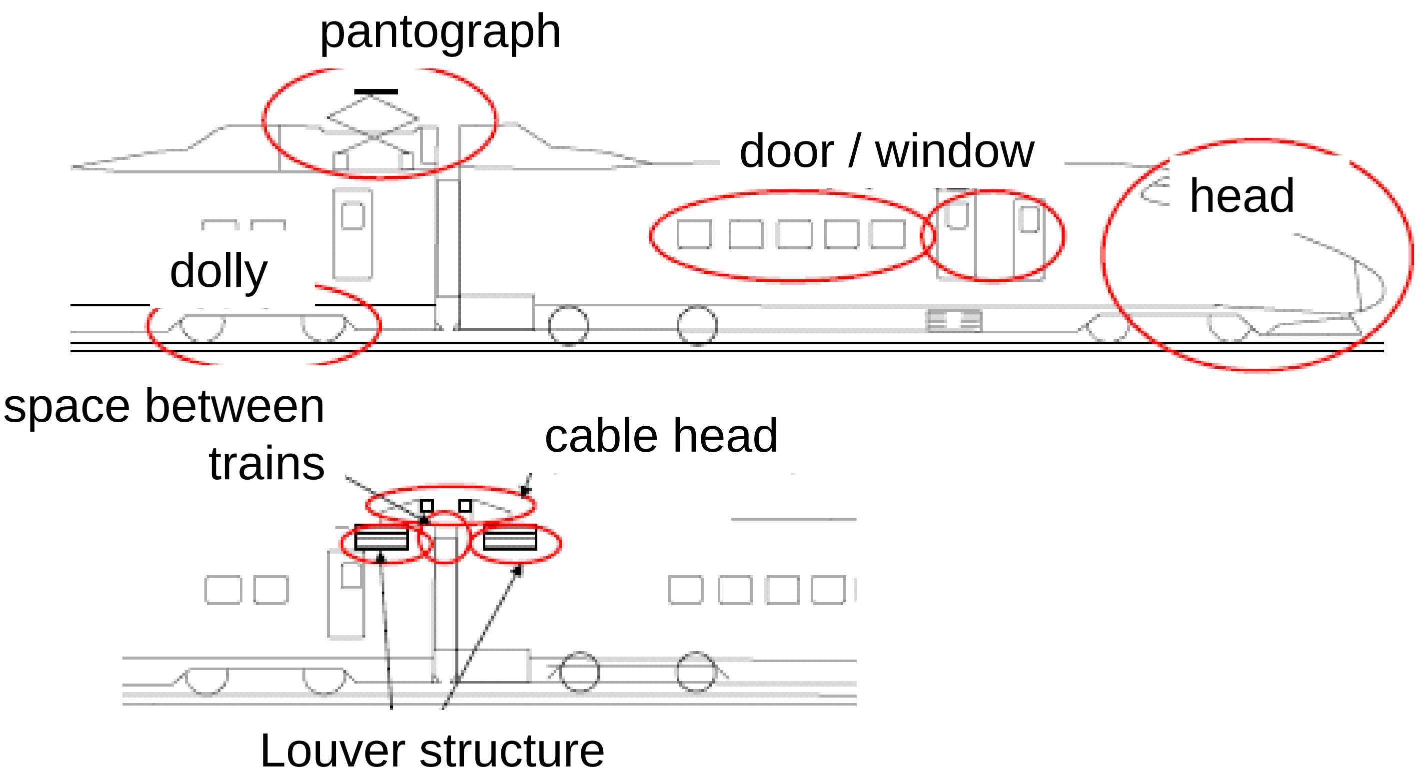
700 Series (1999)

Mechanical design＝ optimization that enables ability ↑ cost ↓

Mathematical expression (modeling)
⇒ calculation ⇒ experiment ⇒ manufacturing

Major Sources of Aeroacoustic Noise Around a Train

‡



Provided from A. Sagawa, Railroad Technical Research Institute

The World of Optimization

Operations Research
Mathematical Programming

Modeling＋ Logics＋ Algorithm
#1 #2 #3

Beautiful *and* Useful

The birth of Mathematical Programming

1947 Linear Programming

‡

G. Dantzig

(1914~2005)

‡

J. von Neumann

(1903~1957)

‡

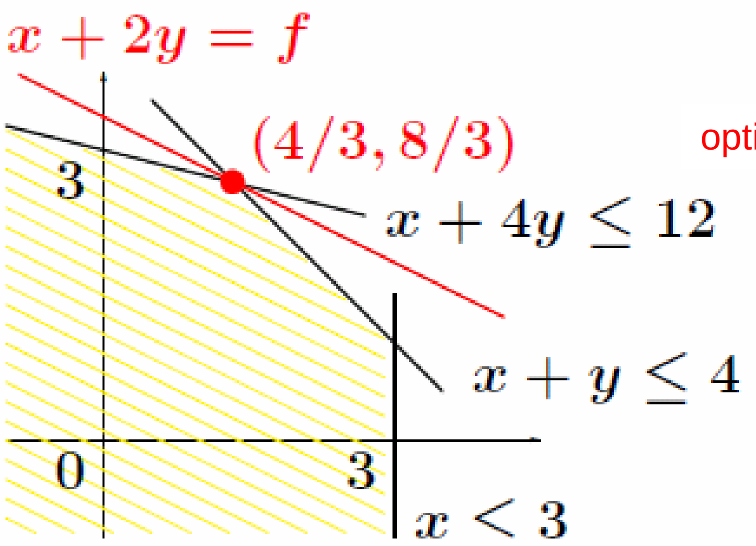
L. V. Kantorovich

(1912~1986)

http://www2.informs.org/Press/GeorgeDantzig.jpg
http://phil.elte.hu/redei/Utrecht/UtrechtNeumann.html
http://www.matematycy.interklasa.pl/images/matematycy/kantorovich.jpg

Linear Programming Problem of 2 Variables

$$\begin{array}{lrcl} \textbf{Maximize} & x + 2y & =: & f \\ \textbf{subject to} & x + 4y & \leq & 12 \\ & x + y & \leq & 4 \\ & x & \leq & 3 \end{array}$$



optimum solution $(x, y) = \left(\frac{4}{3}, \frac{8}{3}\right)$

optimum value $f = \frac{4}{3} + 2 \cdot \frac{8}{3} = \frac{20}{3}$

General Form of Linear Program

Maximize $c^\top x$
subject to $Ax \le b$

$$
\begin{array}{lrcl}
\textbf{Max.} & x_1 + 2\ x_2 & & \\
\textbf{s.t.} & x_1 + 4\ x_2 & \le & 12 \\
 & x_1 + \ \ x_2 & \le & 4 \\
 & x_1 & \le & 3
\end{array}
$$

---

modeling : practically effective in linear approximation
logic : dual theorem
algorithm : simplex method

application to industrial society ⇔ mechanical linear algebra

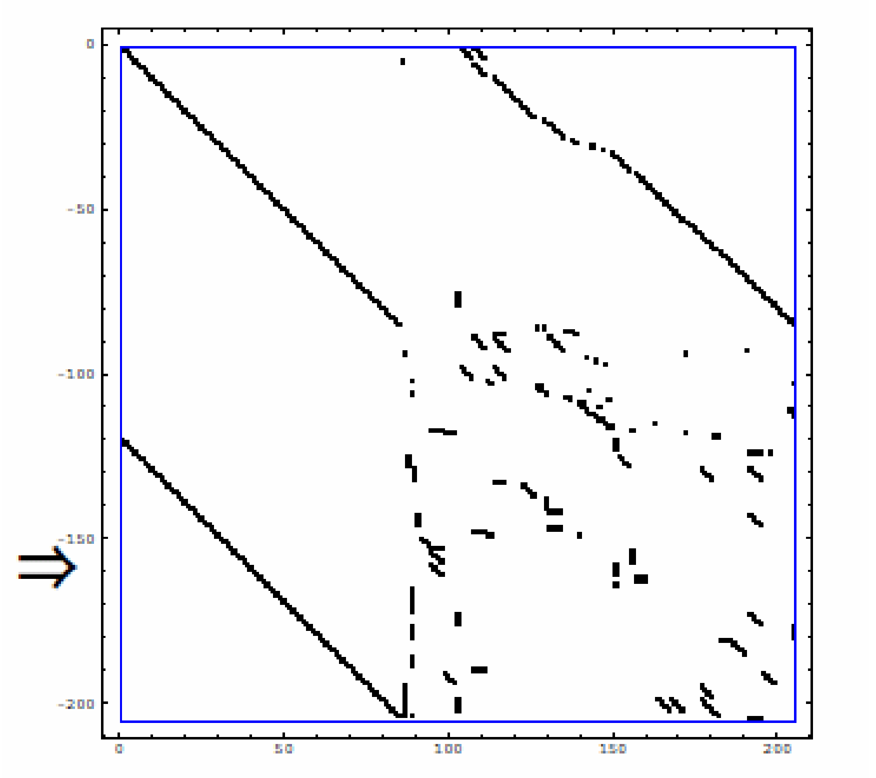
production planning　　　large-scale calculation

Image of Matrix

math textbook

$$\begin{bmatrix} 2 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 3 & 2 \end{bmatrix}$$

problem of engineering



white part is 0 ⇒

large-scale・sparse matrix

Duality of Linear Program—a Meaning of Transposed Matrix

main problem

$$\text{Max. } c^\top x \quad \text{s.t. } Ax \le b$$

dual problem

$$\text{Min. } b^\top y \quad \text{s.t. } A^\top y = c, \quad y \ge 0$$

dual theorem：maximum of main problem=minimum of dual problem

$$\begin{array}{lrcl} \text{Max.} & x_1 + 2\,x_2 & & \\ \text{s.t.} & x_1 + 4\,x_2 & \le & 12 \\ & x_1 + \quad x_2 & \le & 4 \\ & x_1 & \le & 3 \end{array}$$

$$\begin{array}{lrcl} \text{Min.} & 12\,y_1 + 4\,y_2 + 3\,y_3 & & \\ \text{s.t.} & y_1 + \quad y_2 + \quad y_3 & = & 1 \\ & 4\,y_1 + \quad y_2 & = & 2 \\ & y_1,\ y_2,\ y_3 & \ge & 0 \end{array}$$

dual theorem: =maximum of main problem = minimum of dual problem

Solving Dual Problem

$$(u, v, w) = (y_1, y_2, y_3)$$

$$\begin{array}{llll} \text{Minimize} & 12\,u + 4\,v + 3\,w & =: & g \quad \cdots (1) \\ \text{subject to} & u + v + w & = & 1 \quad \cdots (2) \\ & 4\,u + v & = & 2 \quad \cdots (3) \\ & u,\ v,\ w & \geq & 0 \quad \cdots (4) \end{array}$$

$(3) \rightarrow v = 2 - 4u \qquad \geq 0$

$(2) \rightarrow w = 1 - u - v = -1 + 3u \qquad \geq 0$

$(1) \rightarrow g = 12u + 4(2 - 4u) + 3(-1 + 3u) = \boxed{5 + 5u}$

$(4) \rightarrow \boxed{1/3 \leq u \leq 1/2}$

optimum solution $(u, v, w) = \left(\frac{1}{3}, \frac{2}{3}, 0\right)$ optimum value $g = 5 + 5 \cdot \frac{1}{3} = \frac{20}{3}$

max. of main problem = 20/3 = min. of dual problem

Development of Optimization

(Metric Variable)

1947 linear planning Dantzig

1960 non-linear planning, Newton method

Powell, Fletcher

1970 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

Semidefinite Program

(Modern Linear Program)

linear program(LP)

Min. $c \cdot x$

s.t. $a_i \cdot x = b_i$

$(i = 1, \ldots, m)$

$x \geq 0$

semidefinite program(SDP)

Min. $C \bullet X$

s.t. $A_i \bullet X = b_i$

$(i = 1, \ldots, m)$

$X \succeq O$

x vector

$c \cdot x$ inner product of vector

$x \geq 0$ non-negative vector

X symmetric matrix

$C \bullet X = \mathrm{Tr}\,(CX)$

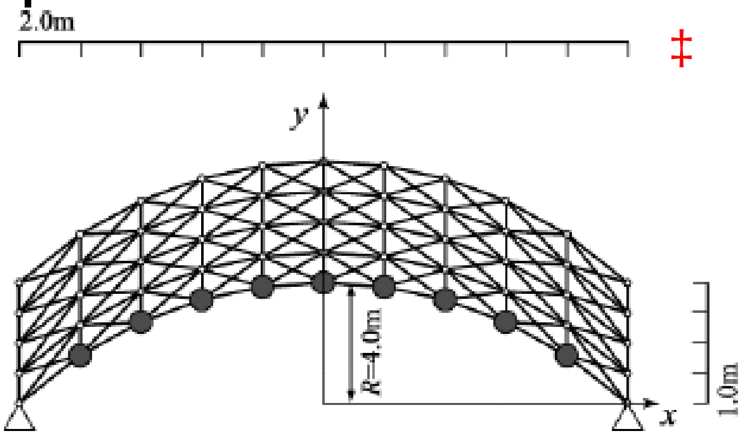
$X \succeq O$ semidefinite value

$(\iff z^\top X z \geq 0, \ \forall z)$

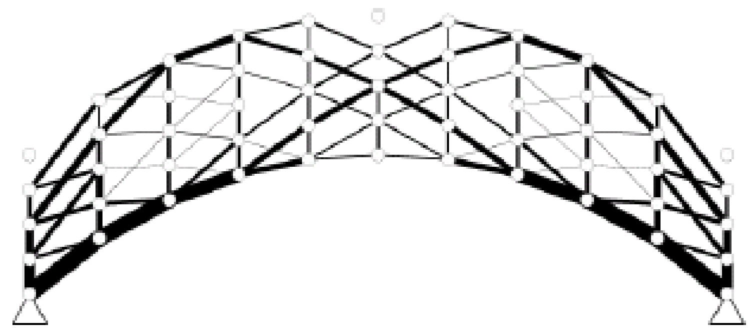
Optimum Design by Semidefinite Program

planar truss

(Dr. Y. Kanno, *Mathematical Engineering*)



‡



problem: minimize weight under「character frequency $\geq$ given α」

(condition not to resonate with earthquake: α = about 40)

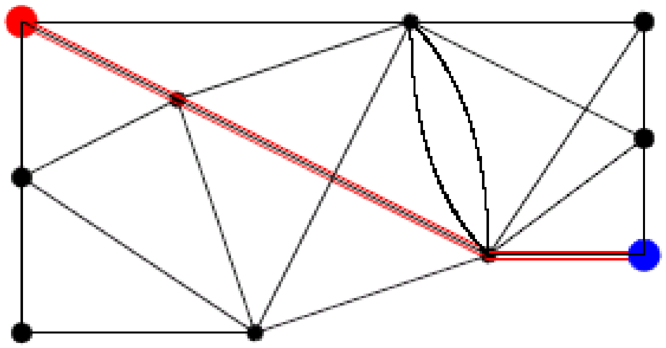
mathematical points :

- character frequency $\Longleftrightarrow$ general proper value of matrix
- any proper value $\lambda \geq \alpha$ $\Longleftrightarrow$ $K - \alpha M$ is a semidefinite value

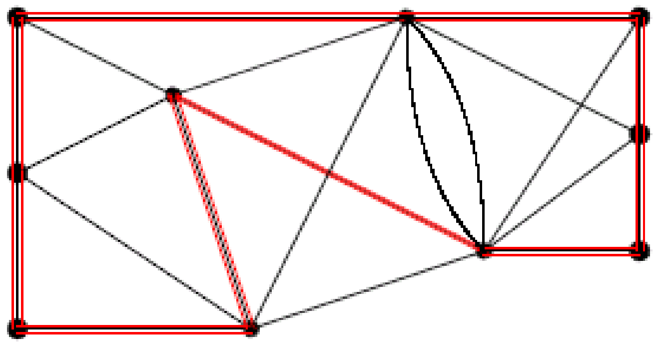
Discrete Optimization

contiguity／discretization

Discrete Optimization



shortest path problem



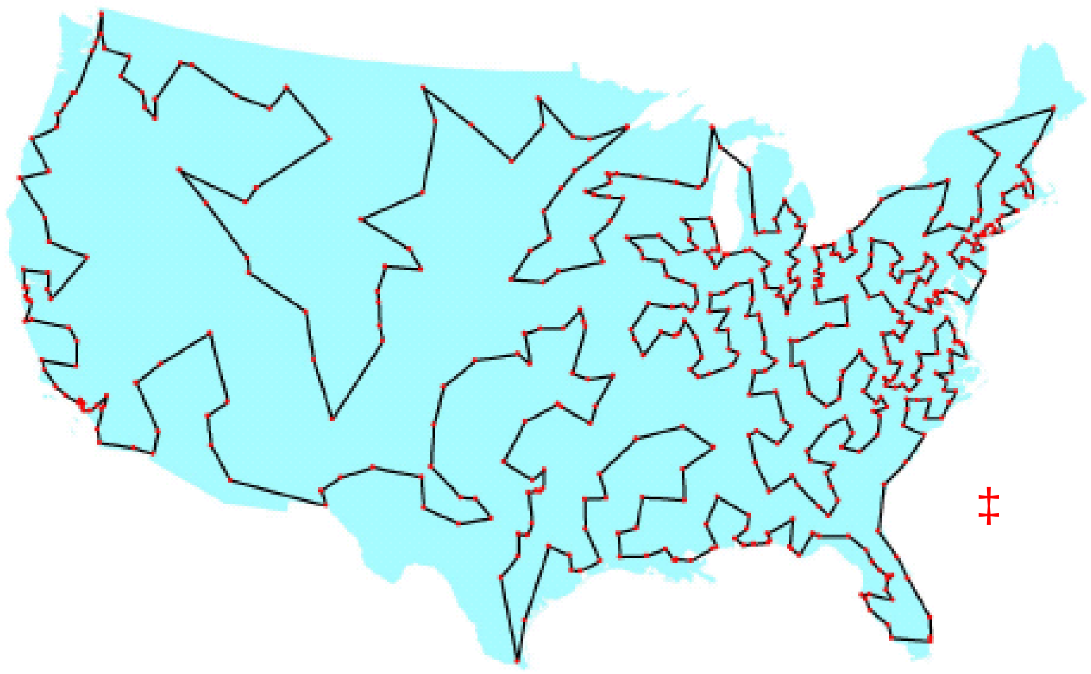
traveling salesman problem

easily solved	(empirical fact)	difficult to solve
Dijkstra method	(algorithm logic)	NP difficulty

Shortest Path Problem　　Traveling Salesman Problem

easy to solve　　　difficult to solve

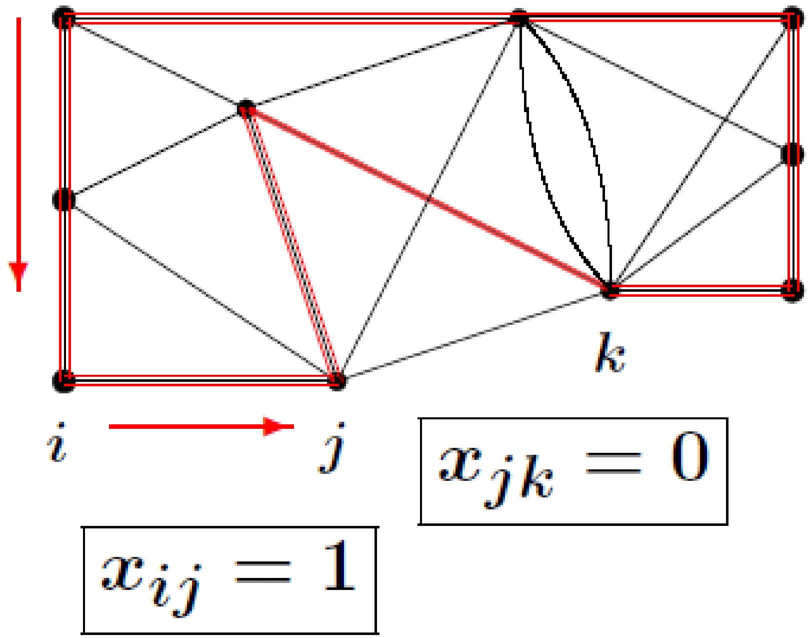
⇒demonstration（by Dr. N. Tsuchimura, Dept. of Math Engineering）



532 cities, M. Padberg–G. Rinaldi (1987)

http://www.tsp.gatech.edu/history/pictorial/att532.html

Mathematical Expression of TSP (Integer Programming)



$$\textbf{Min.} \quad \sum_i \sum_j d_{ij} x_{ij}$$

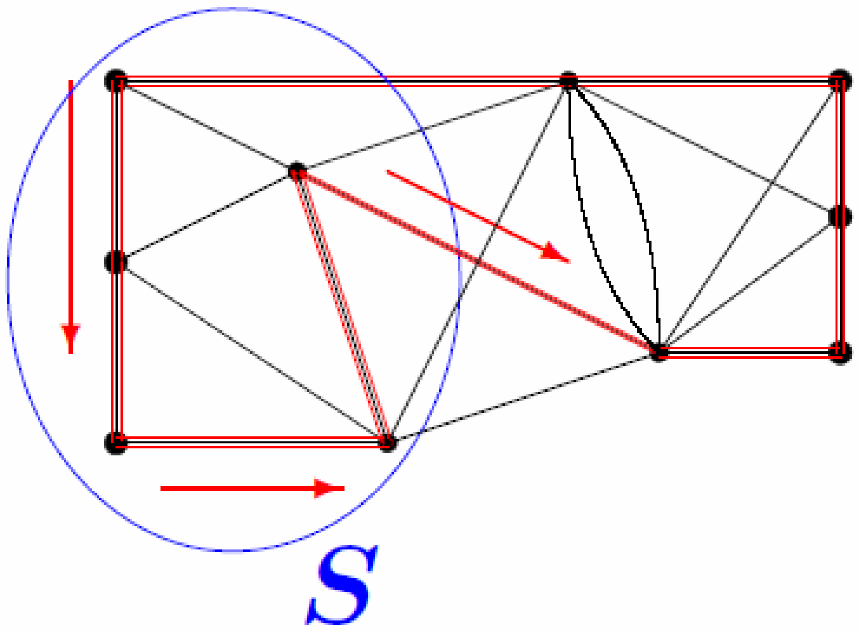
$$\textbf{s.t.} \quad \sum_j x_{ij} = 1$$

$$\sum_i x_{ij} = 1$$

$$\forall S : \quad \sum_{i \in S} \sum_{j \in \overline{S}} x_{ij} \geq 1$$

$$x_{ij} \in \{0, 1\}$$

Mathematical Expression of TSP (Integer Programming)



$$\textbf{Min.} \sum_i \sum_j d_{ij} x_{ij}$$

$$\textbf{s.t.} \sum_j x_{ij} = 1$$

$$\sum_i x_{ij} = 1$$

$$\forall S: \sum_{i \in S} \sum_{j \in \overline{S}} x_{ij} \geq 1$$

$$x_{ij} \in \{0, 1\}$$

Development of Discrete Optimization Theory

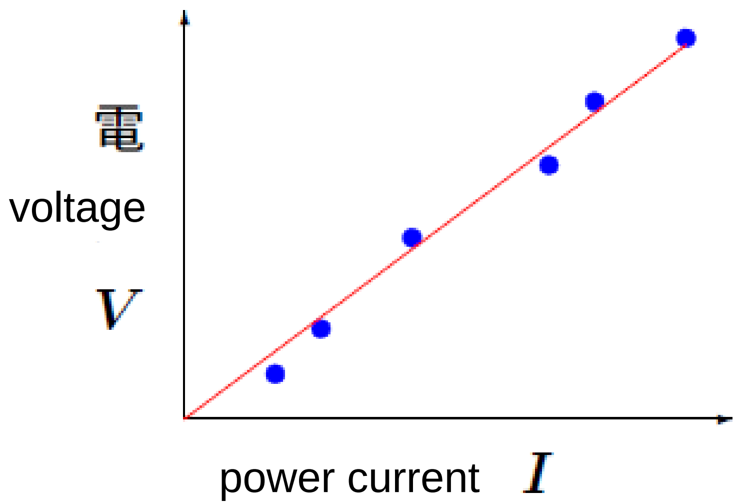
1935	● matroid	Whitney
1947	linear programming (simplex method)	Dantzig
1956	network flow	Ford, Fulkerson
1960	integer programming (cut method)	Gomory
1970	● NP integrity	Cook, Levin
	● graph algorithm	Tarjan
	submodular function	Edmonds
1980	ellipsoid method	Khachiyan
1995	approximate algorithm	Goemans,Williamson
2000	discrete convex analysis	Murota

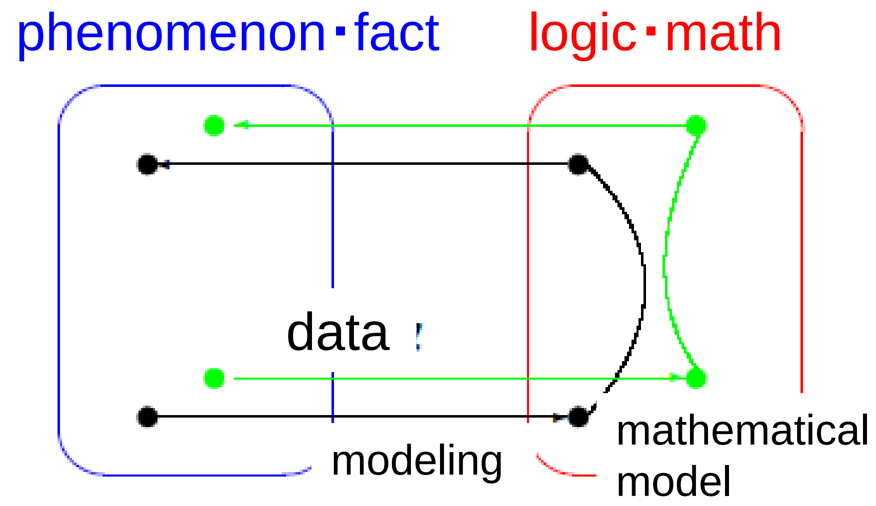
2. Mathematical Model

Data and Model

experimental formula

$$V = RI \quad \text{or} \quad V = RI^2$$



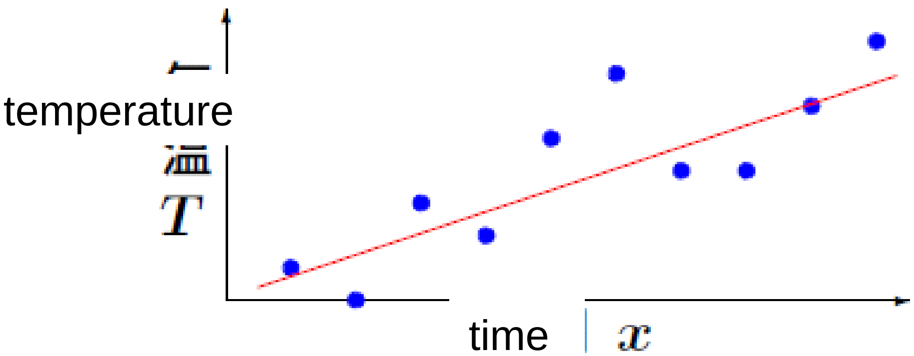


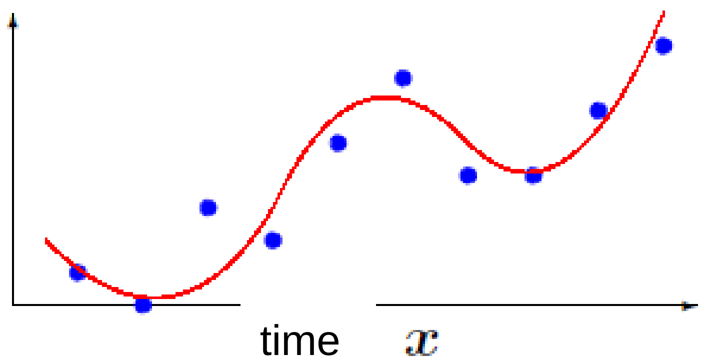
Use observational data to obtain genuine model

least-square method: $f(R) = \sum_i (V_i - RI_i)^2 \to \min$

Adequacy of Model

primary expression $T = ax + b$ quartic expression $T = ax^4 + bx^3 + \cdots$





Complex model
- comply with data(explanation of the past)
- vulnerable to errors and unstable(prediction of the future)

compatibility to data —— simplicity of model

▲

information criterion

Akaike Information Criterion

(compatibility to data) (simplicity of model)

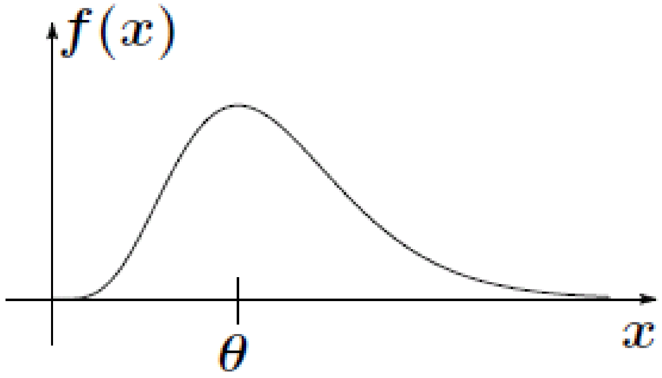
AIC= -2(log likelihood)+2(number of parameters)

1974 Koji Akaike

Log Likelihood

probability density: $f(x;\theta)$

data: $x_1, x_2, \ldots, x_n$



likelihood: $L(\theta) = f(x_1;\theta) \times f(x_2;\theta) \times \cdots \times f(x_n;\theta)$

log likelihood: $\log L(\theta)$

maximization of likelihood $L(\theta) \rightarrow$ estimate value $\hat{\theta}$ of parameter θ

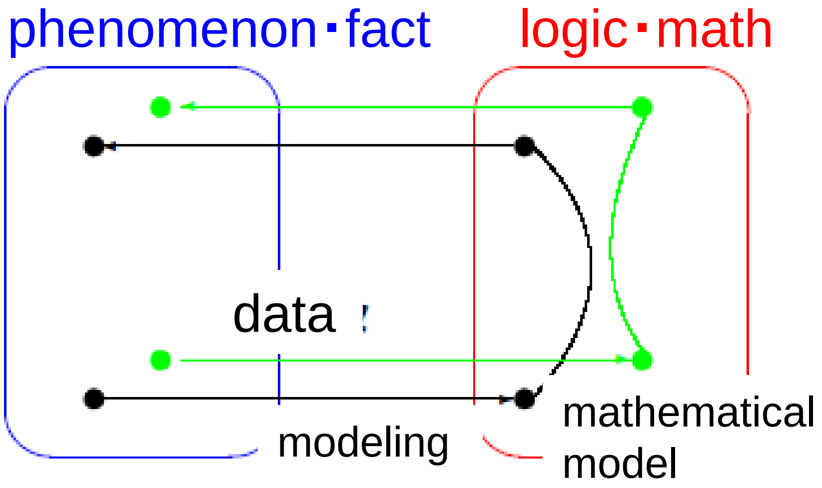
AIC = −2(log likelihood) + 2(number of parameters)

= $-2 \log L(\hat{\theta}) + 2 \dim \theta$

Meaning of Model

Model is

- a bridge between reality and math
- underspecified



Natural science (fundamental science)：

genuine model　precise description of real nature

Technology (engineering)：

appropriate model　tool to achieve a purpose

⇒ science of modeling

Qualitative Change of Data

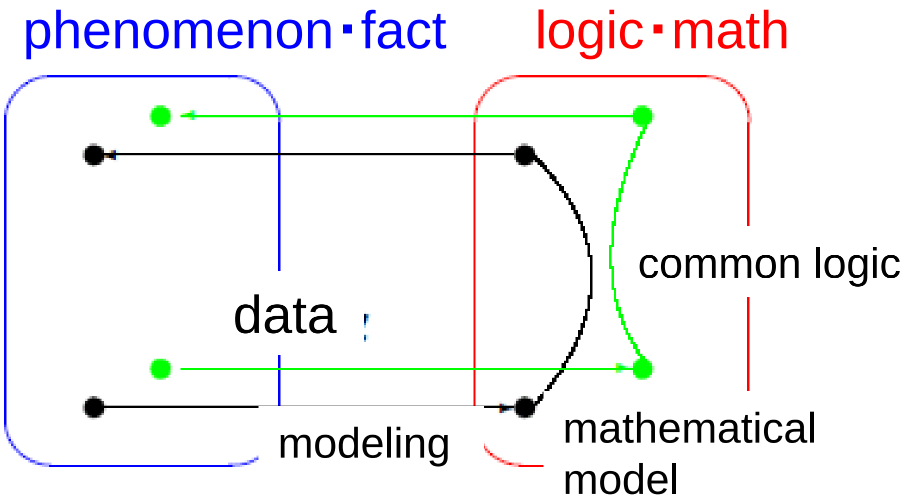
Up till now	Now（Internet age）
Create from experiment and observation	Collect from somewhere
agricultural experimentation	search engine（Yahoo! Google）
experimental design（how to obtain data） variance analysis （how to analyze data）	dispersive accumulation（information explosion） data mining

Basic math：linear algebra（proper value）, idea of probability…

Summary of Today's Lecture　Next Lecture：Logics of Optimization

The world of optimization

continuity／dispersion
linear／convex／non-linear



Modeling＋ Logics＋ Algorithm

Beautiful *and* Useful