MAXIMUM CONSISTENCY METHOD
for Data Fitting
under Interval Uncertainty

Sergey P. Shary

Novosibirsk State University,

Institute of Computational Technologies

Novosibirsk, Russia

I. Interval linear systems and their solvability

Interval linear systems of equations

$$\left\{\begin{array}{ccccccccc} \boldsymbol{a}_{11}x_1 &+& \boldsymbol{a}_{12}x_2 &+& \ldots &+& \boldsymbol{a}_{1n}x_n &=& \boldsymbol{b}_1, \\ \boldsymbol{a}_{21}x_1 &+& \boldsymbol{a}_{22}x_2 &+& \ldots &+& \boldsymbol{a}_{2n}x_n &=& \boldsymbol{b}_2, \\ & \vdots & & & \ddots & & & \vdots & \\ \boldsymbol{a}_{m1}x_1 &+& \boldsymbol{a}_{m2}x_2 &+& \ldots &+& \boldsymbol{a}_{mn}x_n &=& \boldsymbol{b}_m, \end{array}\right.$$

or, briefly,

$$\boldsymbol{A}x = \boldsymbol{b}$$

with an interval $m \times n$-matrix $\boldsymbol{A} = (\,\boldsymbol{a}_{ij})$ and m-vector $\boldsymbol{b} = (\,\boldsymbol{b}_i)$.

Interval systems of linear equations

$$\boldsymbol{A}x = \boldsymbol{b}$$

— a family of point linear systems $Ax = b$ with $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$.

Solution set

to the interval system of linear equations is

$$\varXi(\boldsymbol{A}, \boldsymbol{b}) \;=\; \Bigl\{\, x \in \mathbb{R}^n \mid (\exists\, A \in \boldsymbol{A})(\exists\, b \in \boldsymbol{b})(\, Ax = b\,) \,\Bigr\}$$

Also *united solution set* . . .

Solvability of interval equations

= nonemptyness of the solution set, i. e. $\Xi(\boldsymbol{A}, \boldsymbol{b}) \neq \varnothing$

Strictly speaking, there are strong solvability and weak solvability . . .

In general, recongnition of the solvability is NP-hard

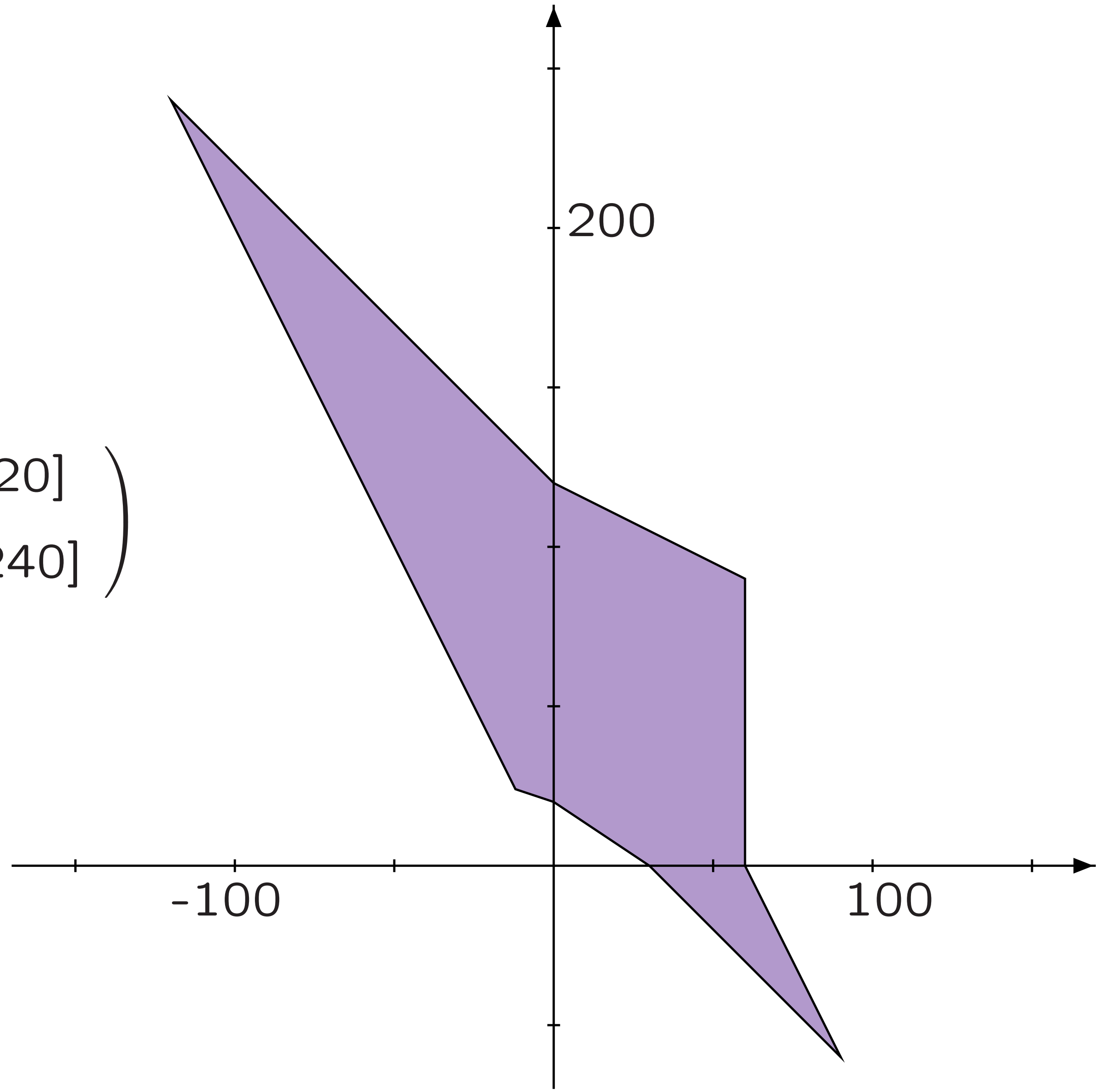
Anatoly V. Lakeyev — 1993

Vladik Kreinovich

Jiři Rohn

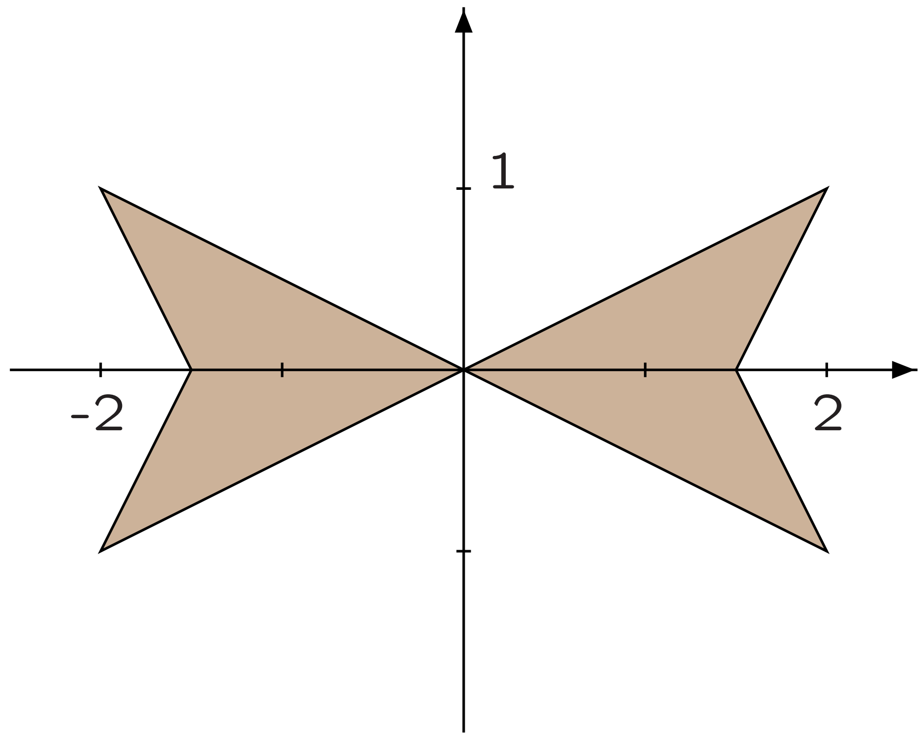
Example: Hansen system

$$\begin{pmatrix} [2,3] & [0,1] \\ [1,2] & [2,3] \end{pmatrix} x = \begin{pmatrix} [0,120] \\ [60,240] \end{pmatrix}$$

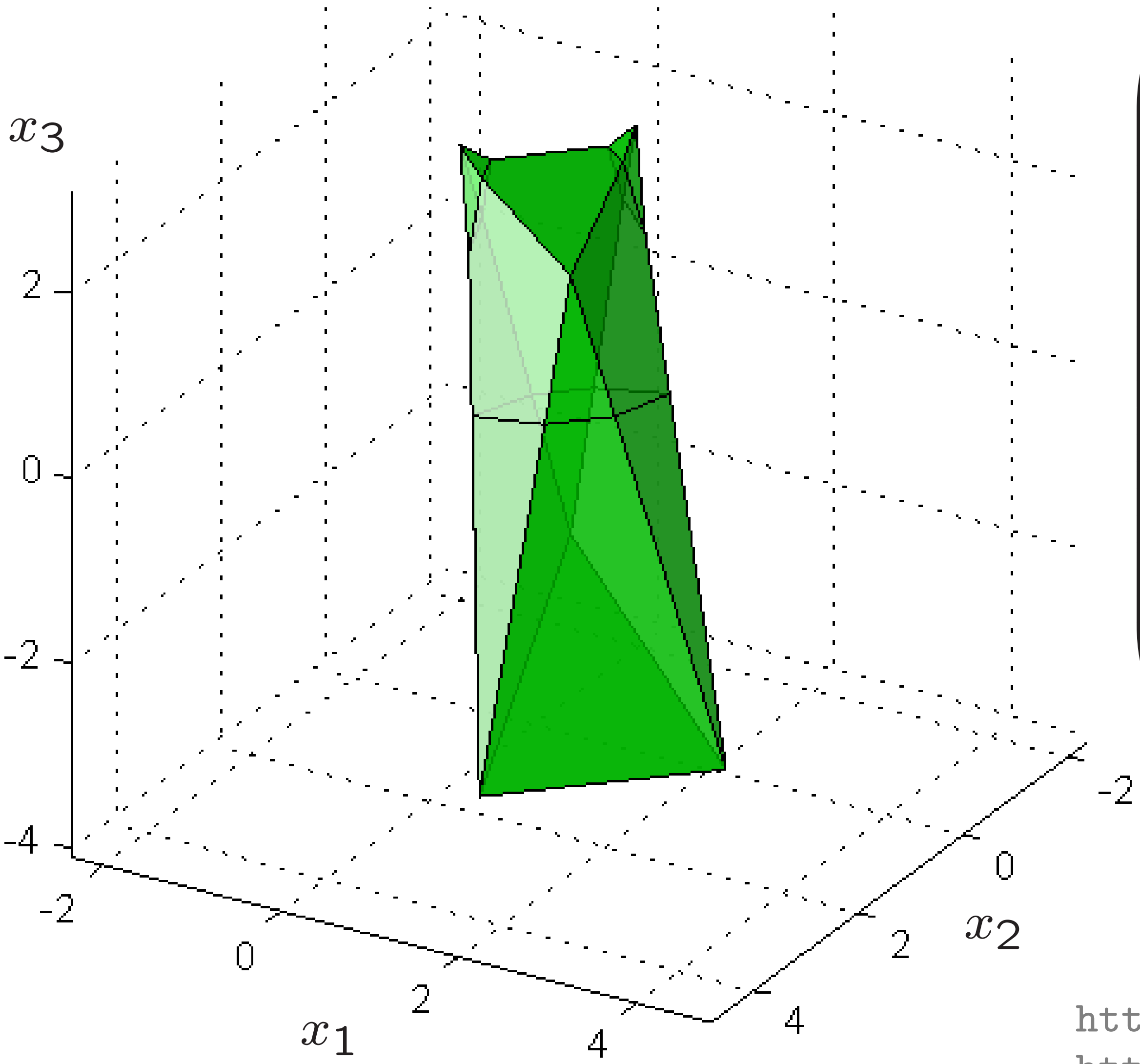


Example: almost disconnected solution set

$$\begin{pmatrix} [2,4] & [-1,1] \\ [-1,1] & [2,4] \end{pmatrix} x = \begin{pmatrix} [-3,3] \\ 0 \end{pmatrix}$$



Example: “bobtail cat”



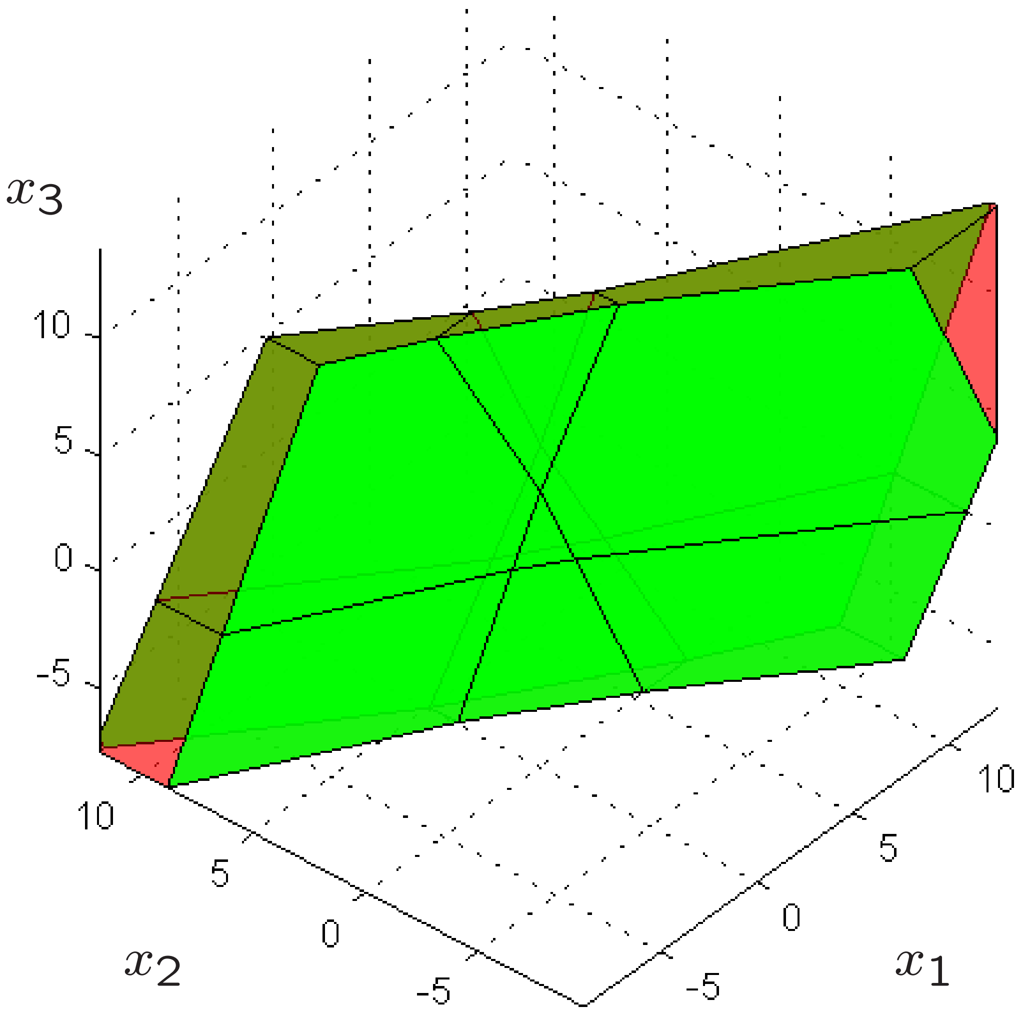
$$\begin{pmatrix} [0.8, 1.2] & [0.8, 1.2] & 1 \\ [0.8, 1.2] & [1.8, 2.2] & 1 \\ [0.8, 1.2] & [2.8, 3.2] & 1 \\ [1.8, 2.2] & [0.8, 1.2] & 1 \\ [1.8, 2.2] & [1.8, 2.2] & 1 \\ [1.8, 2.2] & [2.8, 3.2] & 1 \\ [2.8, 3.2] & [0.8, 1.2] & 1 \\ [2.8, 3.2] & [1.8, 2.2] & 1 \\ [2.8, 3.2] & [2.8, 3.2] & 1 \end{pmatrix} x = \begin{pmatrix} [1, 3] \\ [2, 4] \\ [3, 5] \\ [2, 4] \\ [3, 5] \\ [4, 6] \\ [3, 5] \\ [4, 6] \\ [5, 7] \end{pmatrix}$$

IntLinInc3D package by Irene A. Sharaya

`http://www.nsc.ru/interval/Programing`

`http://www.nsc.ru/interval/sharaya`

Example: one row



$$\big([1.8, 2.2] \quad [2.8, 3.2] \quad 1\big)\, x = \big([4, 6]\big)$$

IntLinInc3D package by Irene A. Sharaya

`http://www.nsc.ru/interval/Programing`

`http://www.nsc.ru/interval/sharaya`

II. Recognizing functionals of the solution sets

Characterization of points from the solution set

$$x \in \Xi(\mathbf{A}, \mathbf{b}) \quad \Leftrightarrow \quad \mathbf{A}x \cap \mathbf{b} \neq \varnothing$$

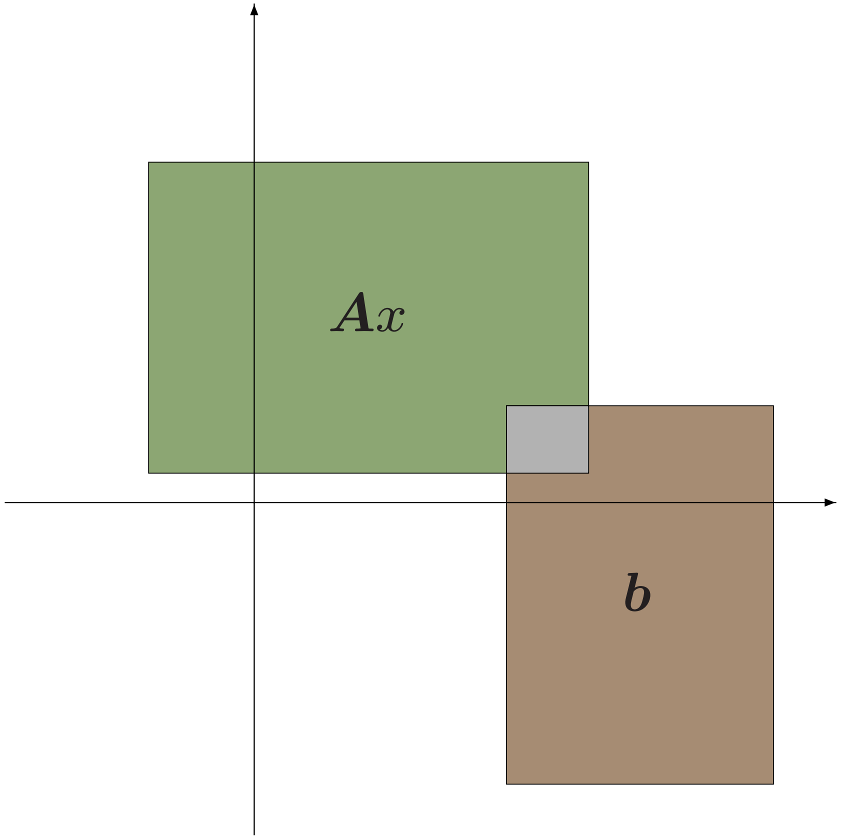
— *Beeck characterization*

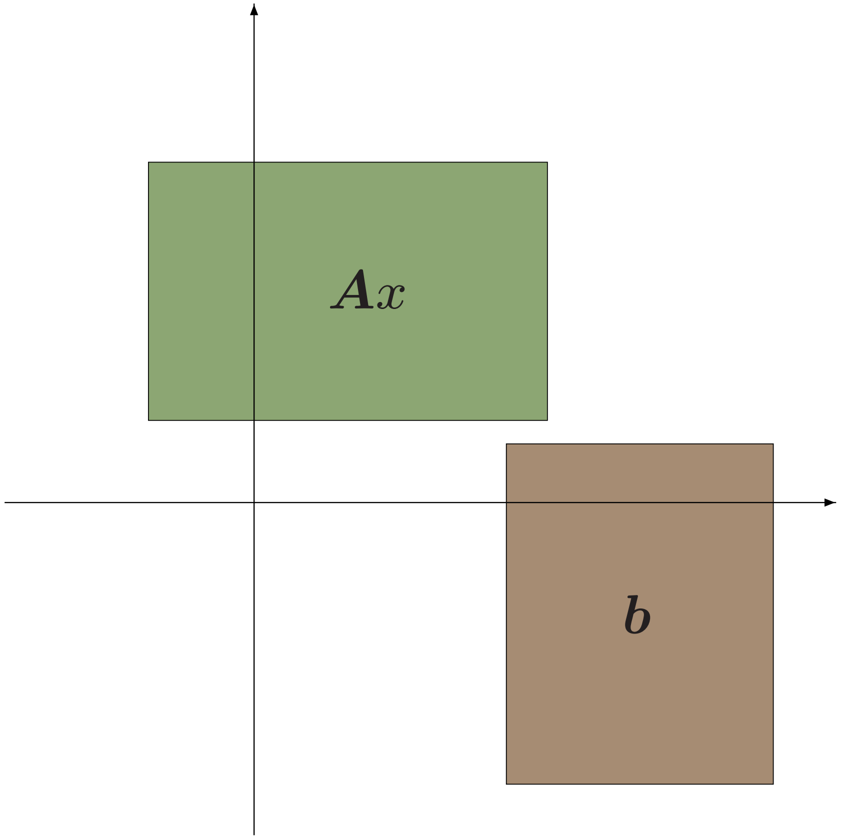
for the solution set to interval linear systems.

Beeck H. Über die Struktur und Abschätzungen der Lösungsmenge von linearen Gleichungssystemen mit Intervallkoeffizienten // *Computing*. –1972. – Vol. 10. – P. 231–244.

Characterization of points from the solution set

Testing Beeck characterization amounts to recognition whether $\boldsymbol{A}x$ and $\boldsymbol{b}$ intersect with each other

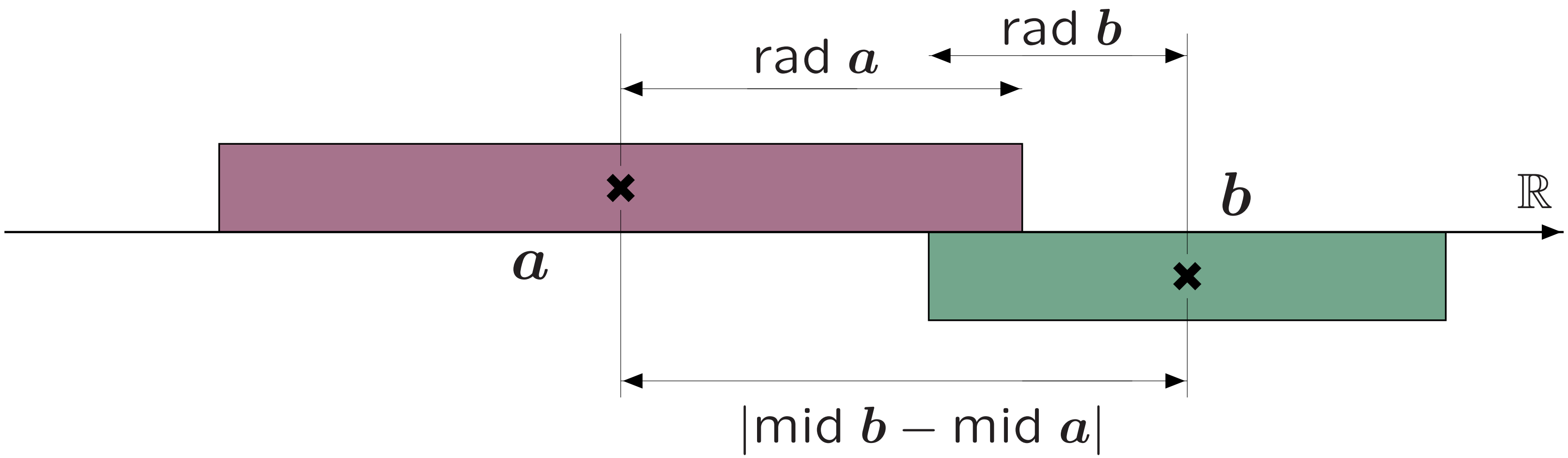




— *intersection measure is an analog of the defect*

Characterization of points from the solution set

$$\boldsymbol{a} \cap \boldsymbol{b} \neq \varnothing \qquad \Leftrightarrow \qquad |\text{mid}\ \boldsymbol{a} - \text{mid}\ \boldsymbol{b}| \leq \text{rad}\ \boldsymbol{a} + \text{rad}\ \boldsymbol{b}$$



This is why

$$\boldsymbol{A}x \cap \boldsymbol{b} \neq \varnothing \qquad \Leftrightarrow \qquad \text{rad}\,(\boldsymbol{A}x)_i + \text{rad}\ \boldsymbol{b}_i - \Big|\,\text{mid}\,(\boldsymbol{A}x)_i - \text{mid}\ \boldsymbol{b}_i\Big| \geq 0,$$
$$i = 1, 2, \ldots, m.$$

Compatibility measure for interval linear systems

As the "compatibility / consistency measure", we can take

$$\min_{1\leq i\leq m} \left\{ \text{rad}\,(\boldsymbol{A}x)_i + \text{rad}\,\boldsymbol{b}_i - \Big|\,\text{mid}\,(\boldsymbol{A}x)_i - \text{mid}\,\boldsymbol{b}_i\Big| \right\}$$

To simplify the expression, we notice that

$$\text{mid}\,(\boldsymbol{A}x) = (\text{mid}\,\boldsymbol{A})\,x \qquad\qquad \text{rad}\,(\boldsymbol{A}x) = (\text{rad}\,\boldsymbol{A})\,|x|,$$

Recognizing functional of the solution set

Theorem

Let $\boldsymbol{A}$ *be an interval* $m\times n$*-matrix and* $\boldsymbol{b}$ *be an interval* m*-vector. Then the expression*

$$\mathrm{Uss}\,(x,\boldsymbol{A},\boldsymbol{b}) \;=\; \min_{1\leq i\leq m}\left\{\,\mathrm{rad}\,\boldsymbol{b}_i + \sum_{j=1}^{n}\,(\mathrm{rad}\,\boldsymbol{a}_{ij})\,|x_j| - \left|\,\mathrm{mid}\,\boldsymbol{b}_i - \sum_{j=1}^{n}\,(\mathrm{mid}\,\boldsymbol{a}_{ij})\,x_j\,\right|\,\right\}$$

defines such a functional Uss $:\mathbb{R}^n\to\mathbb{R}$ *that the membership of a point* $x\in\mathbb{R}^n$ *in the solution set* $\varXi(\boldsymbol{A},\boldsymbol{b})$ *to the interval linear system* $\boldsymbol{A}x=\boldsymbol{b}$ *is equivalent to non-negativity of the functional* Uss *at* x,

$$x\in\varXi(\boldsymbol{A},\boldsymbol{b}) \qquad\Longleftrightarrow\qquad \mathrm{Uss}\,(x,\boldsymbol{A},\boldsymbol{b})\geq 0.$$

Recognizing functional of the solution set

The solution set $\varXi(\boldsymbol{A}, \boldsymbol{b})$ to an interval linear system is a level set

$$\Big\{ \, x \in \mathbb{R}^n \mid \mathsf{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b}) \geq 0 \, \Big\}$$

of the functional Uss .

. . . by the sign of its values, the functional Uss "recognizes" (decides on) the membership of a point in the set $\varXi(\boldsymbol{A}, \boldsymbol{b})$. This is why we use the term "recognizing"

Properties of recognizing functional

Proposition 1

The functional $\operatorname{Uss}(x, \boldsymbol{A}, \boldsymbol{b})$ *is Lipschitz continuous.*

Proposition 2

The functional $\operatorname{Uss}(x, \boldsymbol{A}, \boldsymbol{b})$ *is concave with respect to* x *in each orthant of the space* $\mathbb{R}^n$*.*

If, in the interval matrix $\boldsymbol{A}$*, some columns are entirely non-interval, then* $\operatorname{Uss}(x, \boldsymbol{A}, \boldsymbol{b})$ *is concave within unions of several orthants.*

Proposition 3

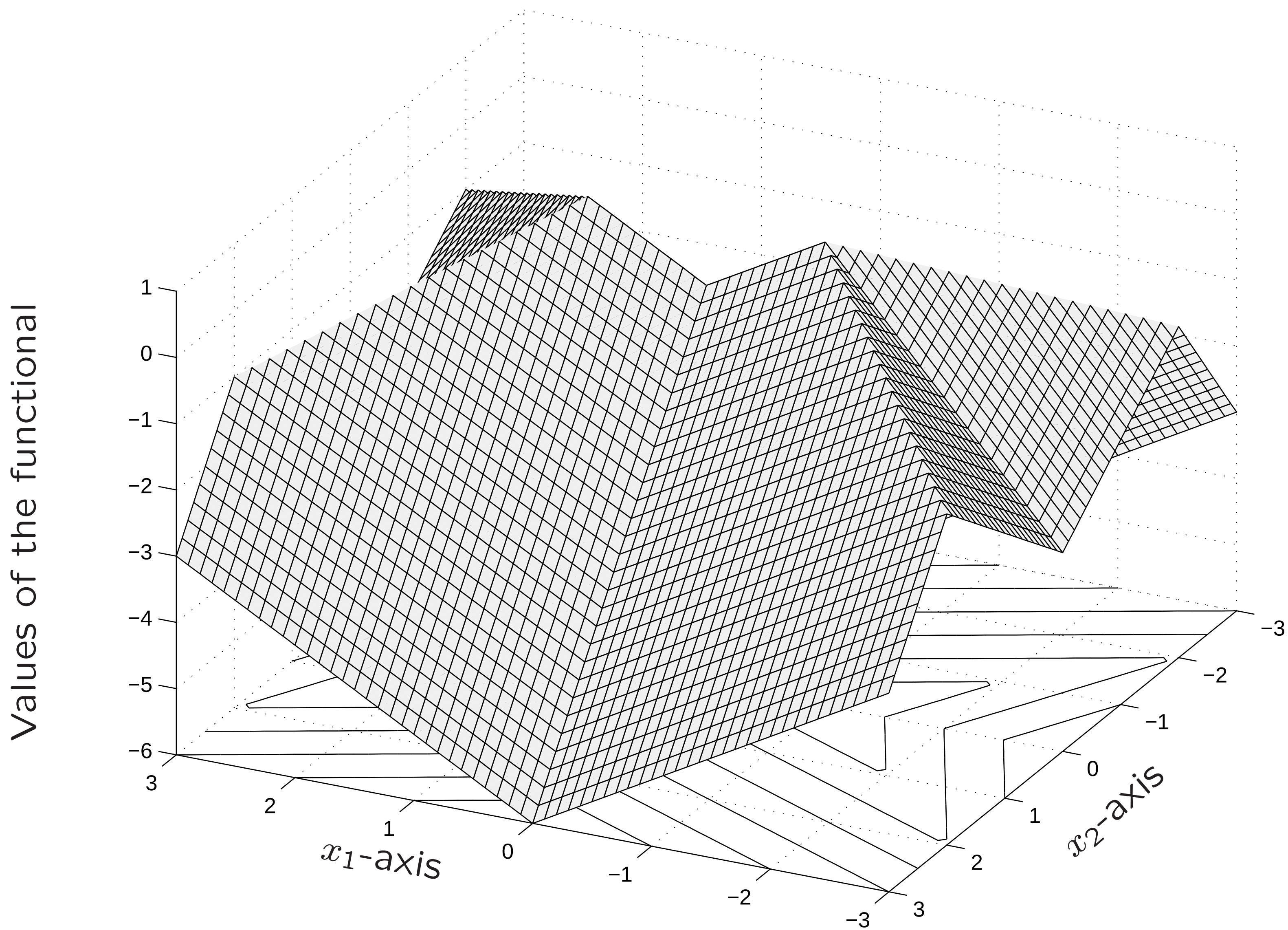
The functional $\operatorname{Uss}(x, \boldsymbol{A}, \boldsymbol{b})$ *is polyhedral, i. e. its hypergraph is a polyhedral set.*

An example

Given the interval linear system

$$\begin{pmatrix} [2,4] & [-1,1] \\ [-1,1] & [2,4] \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} [-3,3] \\ 0 \end{pmatrix},$$

we have, for its solution set, . . .



Values of the functional
1
0
−1
−2
−3
−4
−5
−6
3
2
1
0
−1
−2
−3
x_1-axis
−3
−2
−1
0
1
2
3
x_2-axis

Properties of recognizing functional

Proposition 4

If the solution set $\mathit{\Xi}(\boldsymbol{A}, \boldsymbol{b})$ *is bounded, then the functional* $\mathrm{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b})$ *attains a finite maximum over the entire space* $\mathbb{R}^n$*.*

Proposition 5

If $\mathrm{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b}) > 0$*, then* x *is a point from the topological interior* $\mathrm{int}\,\mathit{\Xi}(\boldsymbol{A}, \boldsymbol{b})$ *of the solution set.*

Proposition 6

Let the interval linear system $\boldsymbol{A}x = \boldsymbol{b}$ *be such that its augmented matrix* $(\boldsymbol{A}, \boldsymbol{b})$ *does not contain rows all whose elements have zero endpoints. Then the membership* $x \in \mathrm{intr}\left(\,\mathit{\Xi}(\boldsymbol{A}, \boldsymbol{b}) \cap \mathcal{O}\,\right)$*, where* $\mathcal{O}$ *is an orthant of the space* $\mathbb{R}^n$*, implies the strict inequality* $\mathrm{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b}) > 0$*.*

Solvability examination for interval linear systems of equations

Given an interval linear system $\boldsymbol{A}x = \boldsymbol{b}$, we solve unconstrained maximization problem for the recognizing functional $\mathrm{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b})$.

Suppose $U = \max_{x\in\mathbb{R}^n} \mathrm{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b})$ and it is attained at a point $\tau \in \mathbb{R}^n$. Then

- if $U \geq 0$, then $\tau \in \varXi(\boldsymbol{A}, \boldsymbol{b}) \neq \varnothing$, i. e. the interval linear system $\boldsymbol{A}x = \boldsymbol{b}$ is solvable and τ lies within the solution set;

- if $U > 0$, then $\tau \in \mathrm{int}\,\varXi(\boldsymbol{A}, \boldsymbol{b}) \neq \varnothing$, and the membership of the point τ in the solution set is stable under small perturbations of $\boldsymbol{A}$ and $\boldsymbol{b}$;

- if $U < 0$, then $\varXi(\boldsymbol{A}, \boldsymbol{b}) = \varnothing$, i. e. the interval linear system $\boldsymbol{A}x = \boldsymbol{b}$ is unsolvable.

Correction of interval systems of equations

$$\text{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b}) \;=\; \min_{1\leq i\leq m} \left\{ \text{rad}\,\boldsymbol{b}_i + \sum_{j=1}^{n} (\text{rad}\,\boldsymbol{a}_{ij})\,|x_j| - \left| \text{mid}\,\boldsymbol{b}_i - \sum_{j=1}^{n} (\text{mid}\,\boldsymbol{a}_{ij})\,x_j \right| \right\}$$

— the values rad $\boldsymbol{b}_i$ occur additively in all the generators

Therefore, if

$$\boldsymbol{e} = \Big([-1,1],\ldots,[-1,1]\Big)^{\top},$$

then, for the system $\boldsymbol{A}x = \boldsymbol{b} + C\boldsymbol{e}$ with a widened right-hand side, there holds

$$\text{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b} + C\boldsymbol{e}) \;=\; \text{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b}) + C$$

$$\max_x\ \text{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b} + C\boldsymbol{e}) \;=\; \max_x\ \text{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b}) + C$$

III. Data fitting under interval uncertainty

Data fitting problem

Given an empirical data, we have to construct a functional relationship, of a prescribed form, between "input" and "output" variables

We consider

$$b = x_0 + \sum_{i=1}^{n} a_i x_i$$

with unknown coefficients x_i that should be determined (estimated) from the sets of values

$$\begin{array}{ccccc} a_{11}, & a_{21}, & \ldots, & a_{n1}, & b_1, \\ a_{12}, & a_{22}, & \ldots, & a_{n2}, & b_2, \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{1m}, & a_{2m}, & \ldots, & a_{nm}, & b_m \end{array}$$

Data fitting problem

We get a system of equations

$$\begin{cases} x_0 + a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ x_0 + a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad\qquad \ddots \qquad\qquad \vdots \\ x_0 + a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

or, briefly,

$$Ax = b$$

with an $m \times (n+1)$-matrix $A = (a_{ij})$ and an m-vector $b = (b_i)$.

Its solution, either common or in a generalied sense,
is taken as an estimate of the parameters $x_0, x_1, \ldots, x_n$

Data fitting problem for uncertain data

It is convenient to describe data uncertainty and inaccuracy by intervals

We are given intervals that enclose true values of the quantities under study, i. e. memberships of a_{ij} and b_i in some intervals,

$$a_{ij} \in \boldsymbol{a}_{ij} = [\underline{\boldsymbol{a}}_{ij}, \overline{\boldsymbol{a}}_{ij}] \qquad \text{and} \qquad b_i \in \boldsymbol{b}_i = [\underline{\boldsymbol{b}}_i, \overline{\boldsymbol{b}}_i],$$

and these intervals include both random and systematic errors.

Leonid Kantorovich — 1962

F.C. Schweppe, P.L. Combettes, J.P.Norton,

M. Milanese, G. Belforte, L. Pronzato, E. Walter, L. Jaulin, . . .

M.L. Lidov, A.P. Voshchinin, S.I. Spivak, N.M. Oskorbin, S.I. Zhilin, . . .

СИБИРСКИЙ МАТЕМАТИЧЕСКИЙ ЖУРНАЛ

Том III, № 5 Сентябрь — Октябрь 1962 г.

Л. В. КАНТОРОВИЧ

О НЕКОТОРЫХ НОВЫХ ПОДХОДАХ К ВЫЧИСЛИТЕЛЬНЫМ МЕТОДАМ И ОБРАБОТКЕ НАБЛЮДЕНИЙ *.

Введение

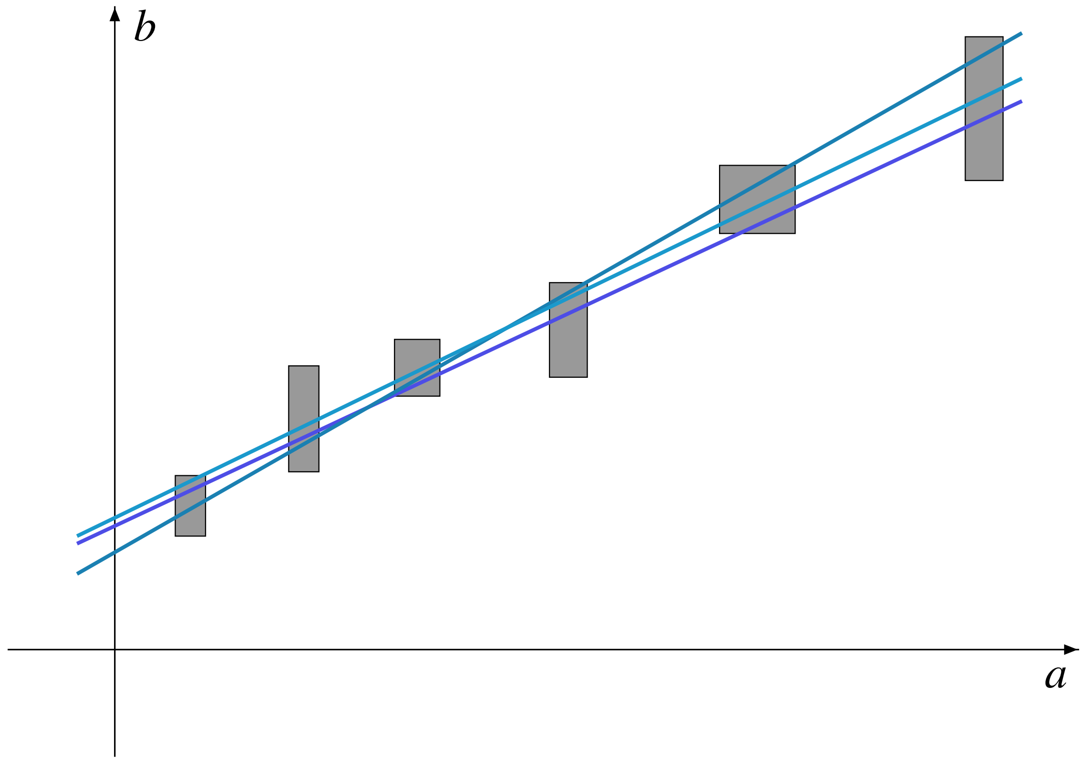
Имевшие место сдвиги в развитии математики и вычислительных средств должны иметь следствием коренные изменения в технике, а возможно и теории численных методов и обработки наблюдений. В той или иной форме отдельные высказываемые ниже соображения встречались в литературе, но не разрабатывались систематически. В частности, мы считаем, что существенное значение имеют следующие моменты:

1. Бóльшая ответственность за результаты расчетов, на которых сейчас нередко базируются решения, касающиеся сложных дорогостоящих объектов современной физики и техники, наличие больших не наблюдаемых этапов при машинных вычислениях повышают требования к надежности окончательных и промежуточных данных, получаемых в процессе применения численных методов и при обработке данных наблюдений. Это обуславливает систематический переход от построения приближенных значений и результатов, к получению точных двухсторонних границ для искомых величин или, если говорить о нечисловых величинах, областей расположения искомых и наблюдаемых величин; иначе говоря возникает задача возможно более точного описания расположения этих величин в соответствующих пространствах их значений. Идеи теорий полуупорядоченных пространств и операций в них, а также некоторых других абстрактных систем объектов дают определенную теоретическую базу для реализации этой точки зрения.

Data fitting problem for interval data

A set of parameters x_0, x_1, . . . , x_n of an object *is consistent* with interval experimental data $(\boldsymbol{a}_{i1}, \boldsymbol{a}_{i2}, \ldots, \boldsymbol{a}_{in}, \boldsymbol{b}_i)$, $i = 1, 2, \ldots, m$, if, for every observation i, there exist such representatives $a_{i1} \in \boldsymbol{a}_{i1}$, $a_{i2} \in \boldsymbol{a}_{i2}$, . . . , $a_{in} \in \boldsymbol{a}_{in}$ and $b_i \in \boldsymbol{b}_i$ that

$$x_0 + a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i \,.$$



Data fitting problem for uncertatin data

The set of parameters consistent with the data can be defined formally as

$$\left\{ x \in \mathbb{R}^{n+1} \;\middle|\; \Big(\, \exists (a_{ij}) \in (\boldsymbol{a}_{ij}) \,\Big)\Big(\, \exists (b_i) \in (\boldsymbol{b}_i) \,\Big)\Big(\, Ax = b \,\Big) \right\}$$

where A is an $m \times (n+1)$-matrix having 1's in the first column and a_{ij}'s at the rest places, $b = (b_i)$, i. e., all x's form *solution set* to interval linear system of equations.

In data fitting theory, it is called *parameter uncertainty set*,
set of possible values of the parameters, *information set*, etc.

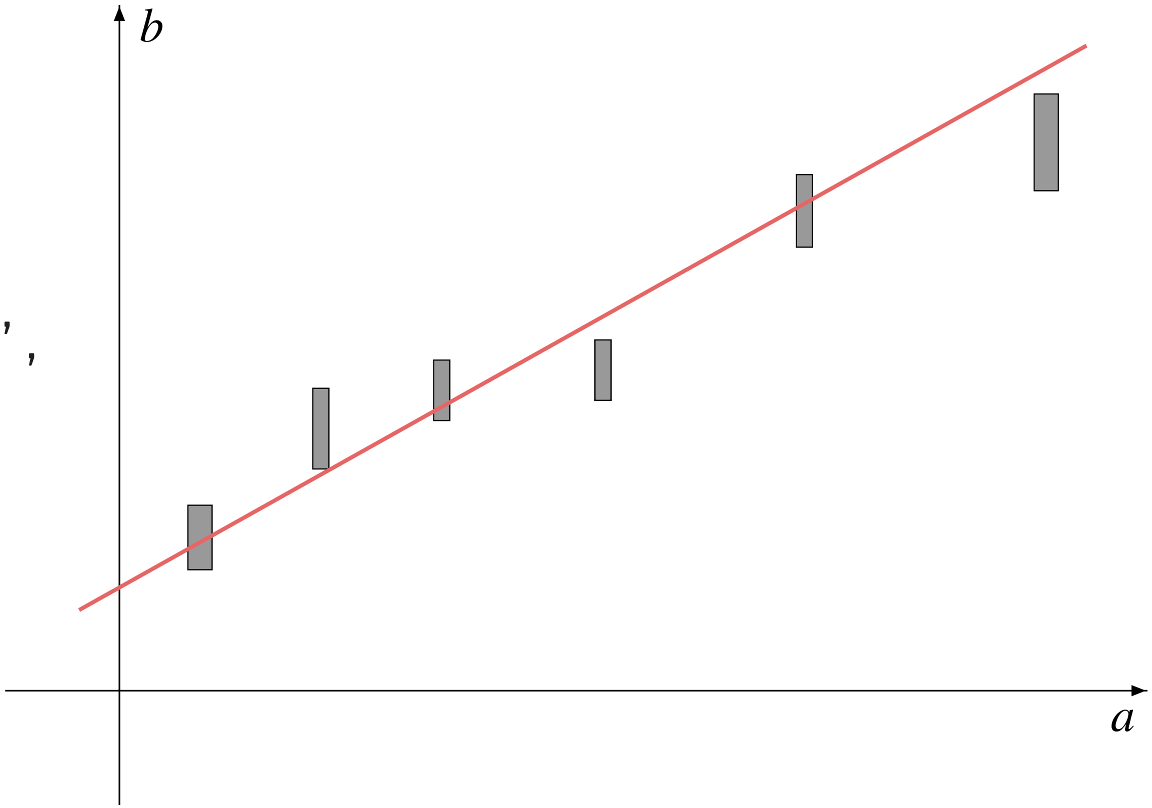
IV. Maximum consistency method

Data fitting under intervally uncertainty

A general way:

1) we assign
   a "consistency measure",

2) we maximize it . . .



An estimate of the parameters is a point
that maximizes the "consistency measure"

Data fitting under intervally uncertainty

What “consistency / inconsistency measure” should we take?

- It must be positive (non-negative) for points from non-empty information set, where the desired “consistency” takes place.
- At the boundary of a non-empty information set, it must be no greater than in its interior.
- Outside the information set, it must be negative, signalling on absence of the “consistency”.

The recognizing functional Uss *suits for our purpose*

Maximum Consistency Method

As an estimate of the parameters, we take a point that provides maximum of the recognizing functional Uss

- If $\max \text{Uss} \geq 0$, the the point lies in the set of parameters consistent with the data (i.e., in the information set).

- If $\max \text{Uss} < 0$, then set of parameters consistent with the data is empty, but the point *minimizes inconsistency*.

Maximum Consistency Method

A practical interpretation:

$\arg\max \text{Uss}$ is the first point that appears in the solution set in the course of uniform widening of the right-hand side vector with respect to its midpoint, since

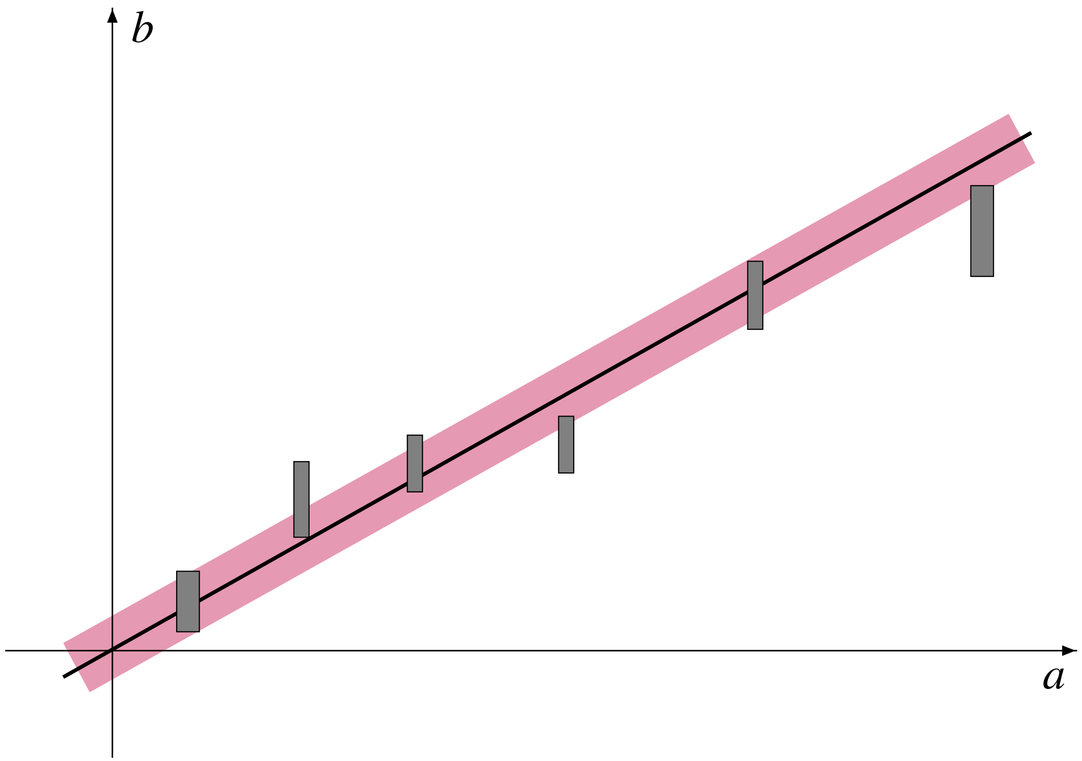
$$\max_x \text{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b} + C\boldsymbol{e}) \;=\; \max_x \text{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b}) + C,$$

where $\boldsymbol{e} = \Big([-1,1], \ldots, [-1,1]\Big)^{\top}$

Maximum Consistency Method

Yet another practical interpretation:

$\arg\max \mathrm{Uss}$ gives parameters of a regression line that should be widened in the smallest possible amount to produce a “regression strip” that intersects all data boxes.



V. Practical implementation

Practical implementation

Overall efficiency crucially depends on efficiency of computing max Uss

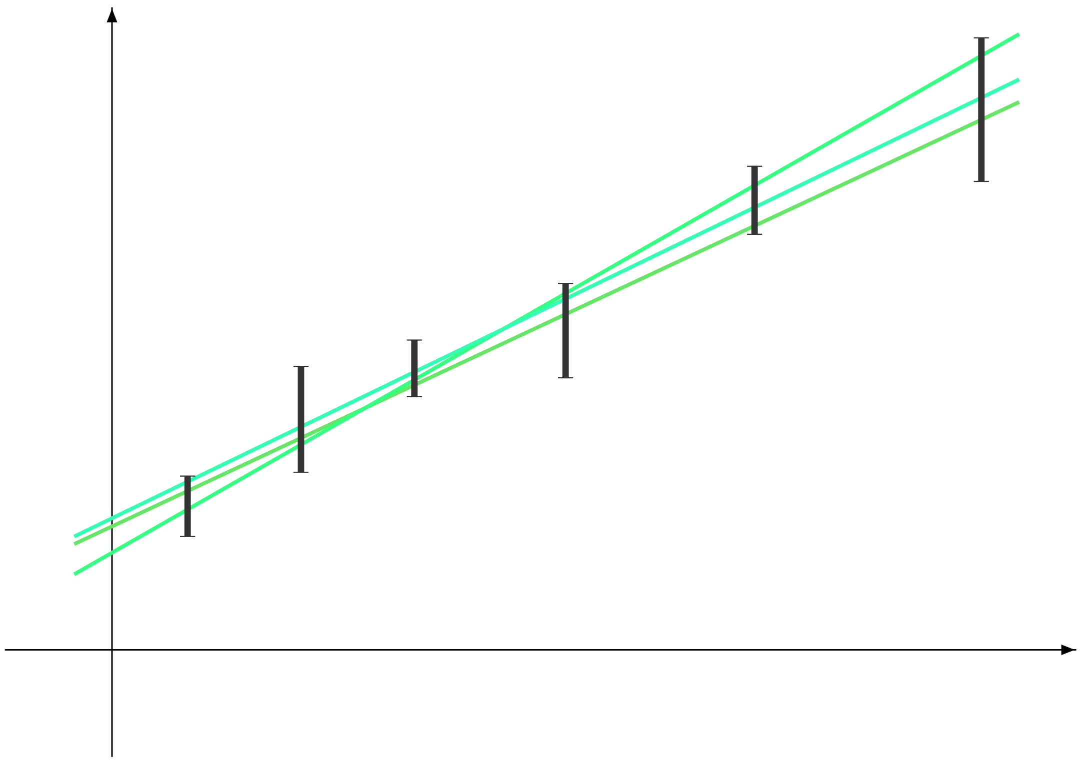
In the general case, it is a global optimization problem
with non-smooth objective function

- global optimization methods for Lipschitz continuous functions
taking into account specificity of the functional Uss

- besides, Uss can be separately maximized in every orthant of $\mathbb{R}^n$

An important particular case

— values of the input variables a are exact,
interval uncertainty is in the output variable b only



An important particular case

— values of the input variables a are exact,
interval uncertainty affects only the output variables b

The interval linear system

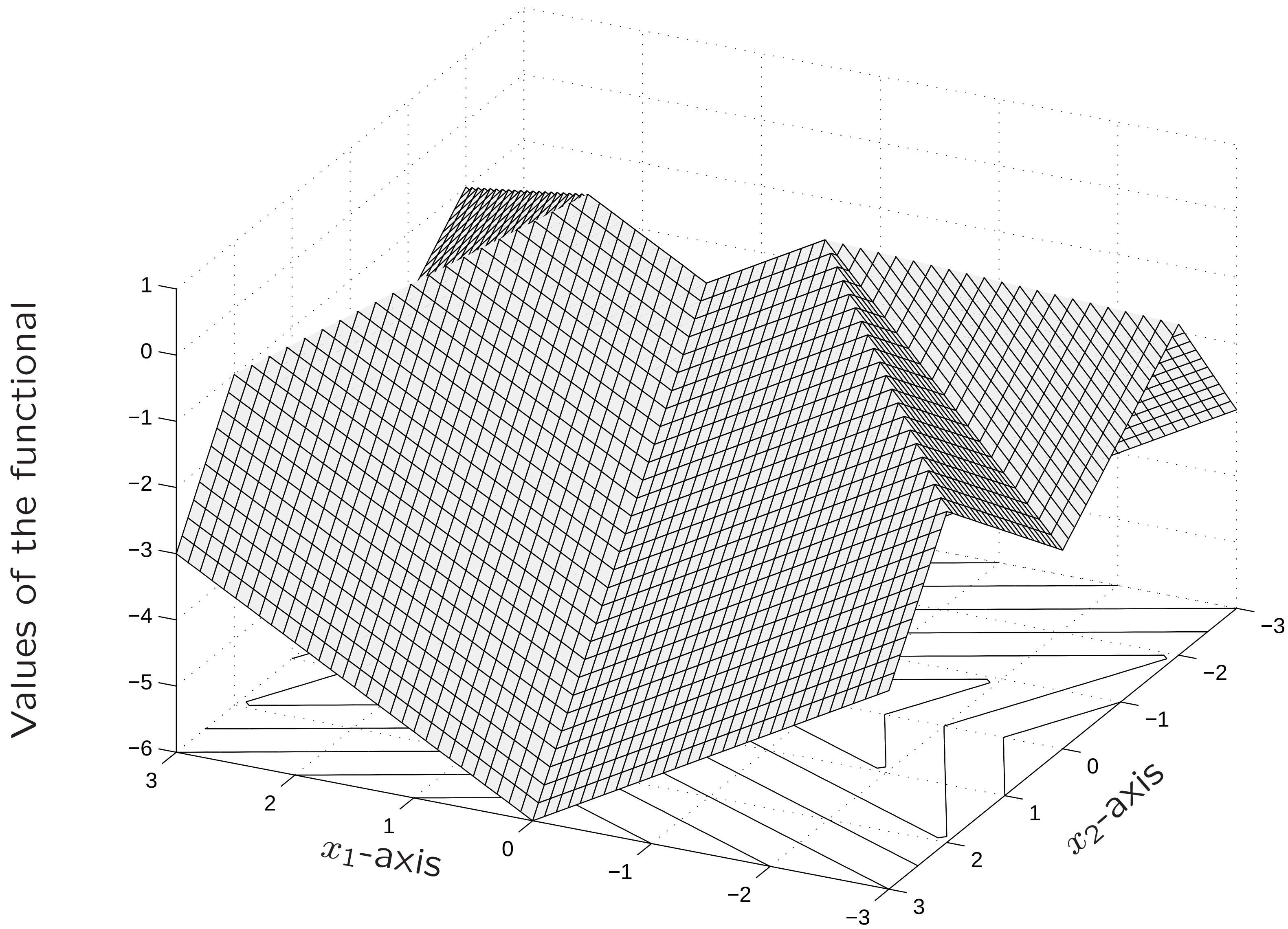
$$Ax = \boldsymbol{b}$$

with a point matrix $A = (a_{ij})$, which leads to

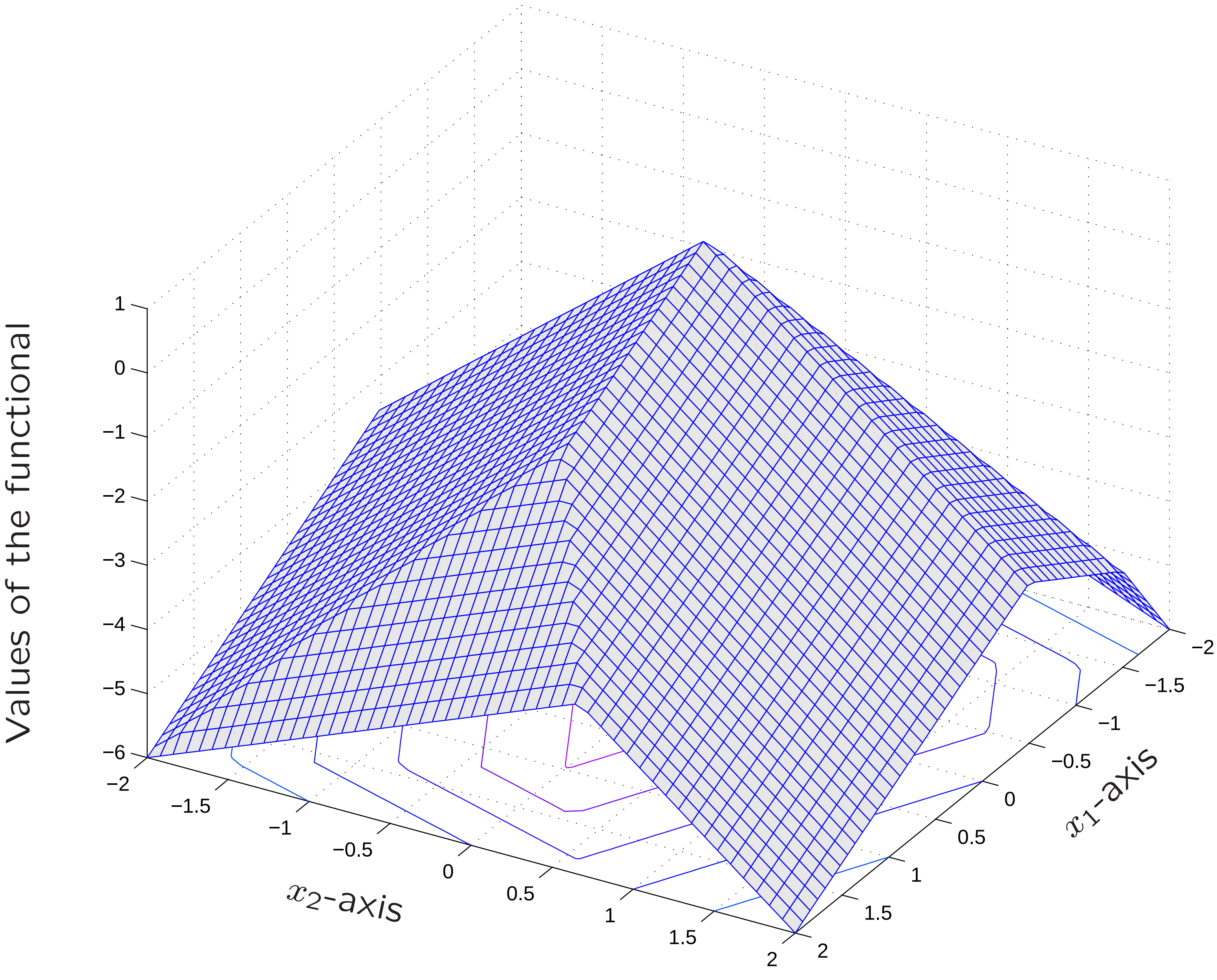
$$\mathrm{Uss}\,(x, \boldsymbol{A}, \boldsymbol{b}) \;=\; \min_{1\leq i\leq m} \left\{ \,\mathrm{rad}\,\boldsymbol{b}_i - \left|\, \mathrm{mid}\,\boldsymbol{b}_i - \sum_{j=1}^{n} a_{ij}\, x_j \,\right| \,\right\}$$

$\Longrightarrow$ the recognizing functional Uss is globally concave

So, instead of



we have


Values of the functional
1
0
−1
−2
−3
−4
−5
−6
−2
−1.5
−1
−0.5
0
0.5
1
1.5
2
x_2-axis
−2
−1.5
−1
−0.5
0
0.5
1
1.5
2
x_1-axis

— graph of the recognizing functional

for the solution set to the interval linear system

$$\begin{pmatrix} 3 & -1 \\ -1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} [-2,2] \\ [0,1] \\ [-1,0] \end{pmatrix}$$

Exact input variables correspond to applicability conditions of the traditional regression analysis, for which the most powerful results on the least squares optimality have been obtained (Gauss-Markov theorem, etc.).

A practical implementation

In the case of point matrix $\boldsymbol{A}$, maximization of Uss can rely on the developed convex nonsmooth optimization techniques

(N.Z. Shor's subgradient algorithms, etc.)

A Matlab code `lintreg` that implements maximum consistency method based on the nonsmooth optimization algorithm `ralgb5` by Dr. P. Stetsyuk (Institute of Cybernetics, Kiev, Ukraine) is freely downloadable from

`http://www.nsc.ru/interval`

Russian web-site "**Interval Analysis and its Applications**"

Results and conclusions

- For interval linear systems, introduction of the recognizing functional reduces the problem of solvability recognition to a convenient analytical form.

- *Maximum Consistency Method* is a new and promising technique for data processing under interval uncertainty based on maximization of the recognizing functional.

  It is going to be a good alternative to the traditional Least Squares Method.

I appreciate your attention

VI. Maximum Consistency vs Least Squares

An example of the least squares failure

An example of the least squares failure

. . . an example by Irene A. Sharaya
where the least squares estimate
does not lie in the information set

Let a variable $y \in \mathbb{R}$ depends linearly on a variable $x \in \mathbb{R}$, so that

$$y = \alpha x + \beta.$$

The unknown values of α and β should be determined from the results of the following measurements

Measurement	1	2	3
x	0	1	2
y	1	2	-0.5

An example of the least squares failure

In the experiments,

- the variable x is measured without errors,
- for the variable y, the measurements produce intervals such that
  - their centers are given in the table,
  - all their radii are equal to 1,
  - the true value of y may be any number from the interval (no probabilistic assumptions!)

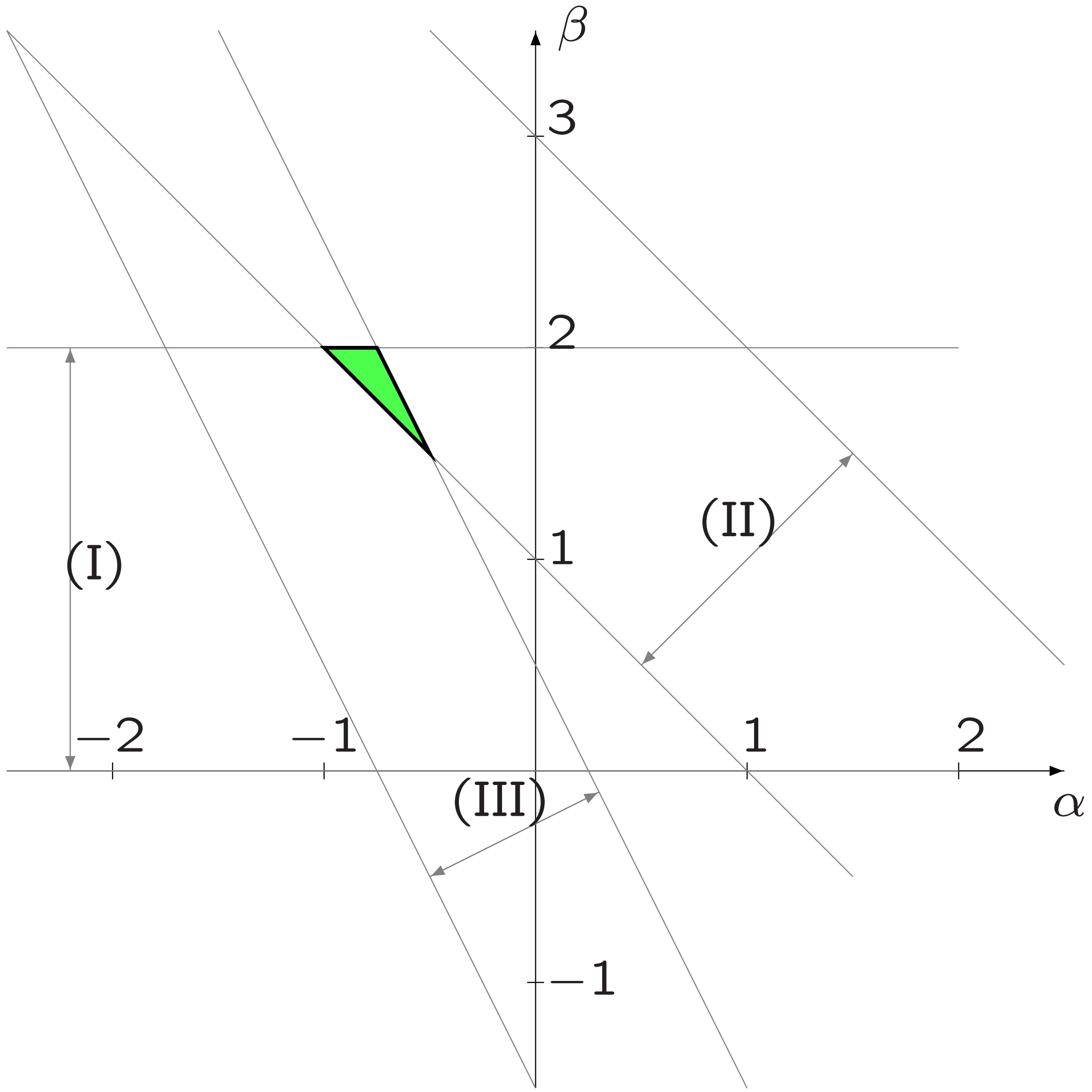
An example of the least squares failure

Information set, i. e. the set of all the pairs α and β, consistent with the measurements is described by the system

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in \begin{pmatrix} 1 + [-1, 1] \\ 2 + [-1, 1] \\ -0.5 + [-1, 1] \end{pmatrix},$$

being intersection of three stripes in $\mathbb{R}^2$:

$$\begin{aligned} &\text{(I)} && \beta \in [0, 2], \\ &\text{(II)} && \beta \in -\alpha + [1, 3], \\ &\text{(III)} && \beta \in -2\alpha + [-1.5, 0.5]. \end{aligned}$$



— information set is marked in green.

This is a triangle with the vertices $(-1, 2)$, $(-0.5, 1.5)$ and $(-0.75, 2)$

An example of the least squares failure

The least squares estimate for α and β can be computed from the normal equations system

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \alpha^\star \\ \beta^\star \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -0.5 \end{pmatrix}.$$

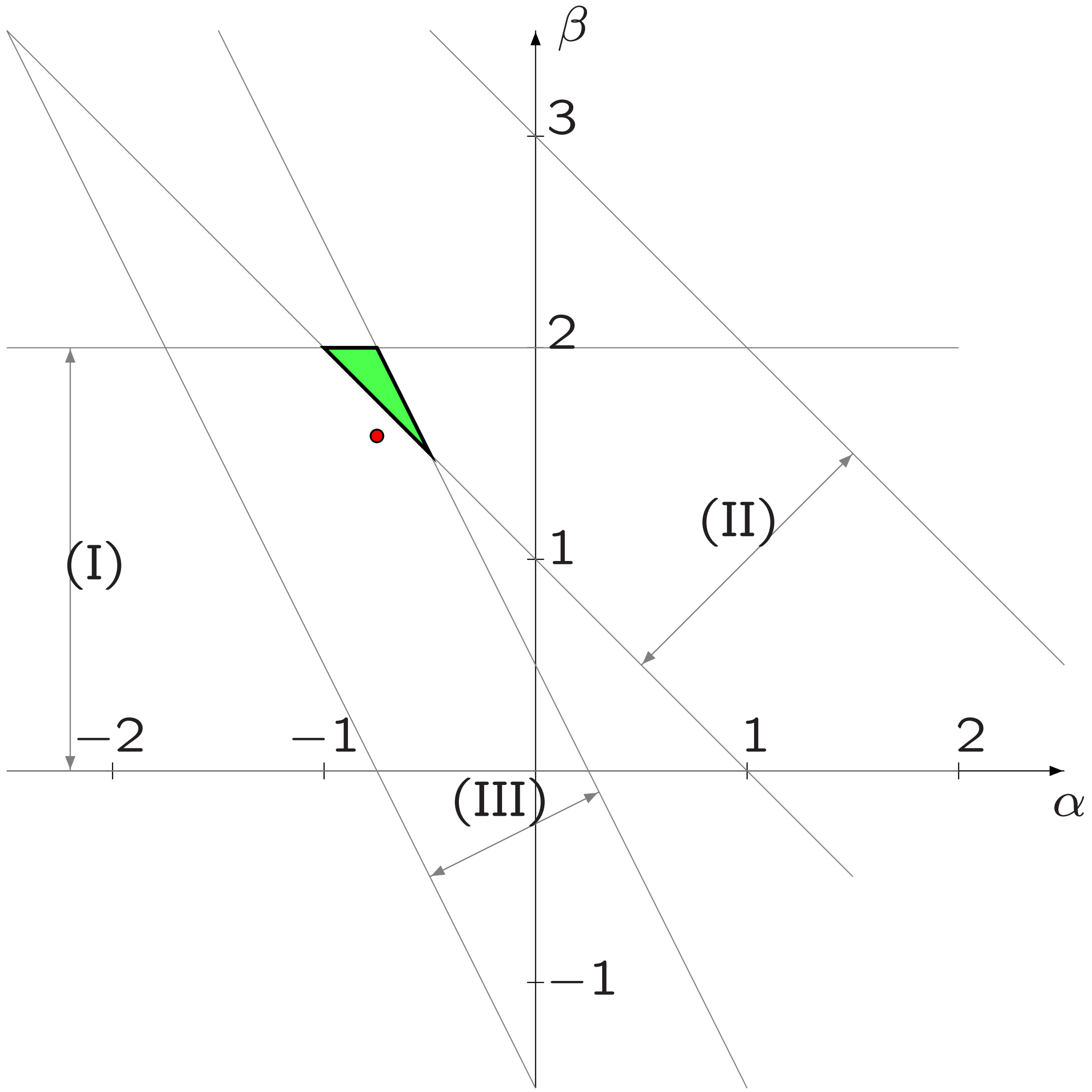
We have

$$\begin{pmatrix} 5 & 3 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} \alpha^\star \\ \beta^\star \end{pmatrix} = \begin{pmatrix} 1 \\ 2.5 \end{pmatrix}.$$

$$\det \begin{pmatrix} 5 & 3 \\ 3 & 3 \end{pmatrix} = 6, \qquad \begin{pmatrix} 5 & 3 \\ 3 & 3 \end{pmatrix}^{-1} = \frac{1}{6} \begin{pmatrix} 3 & -3 \\ -3 & 5 \end{pmatrix},$$

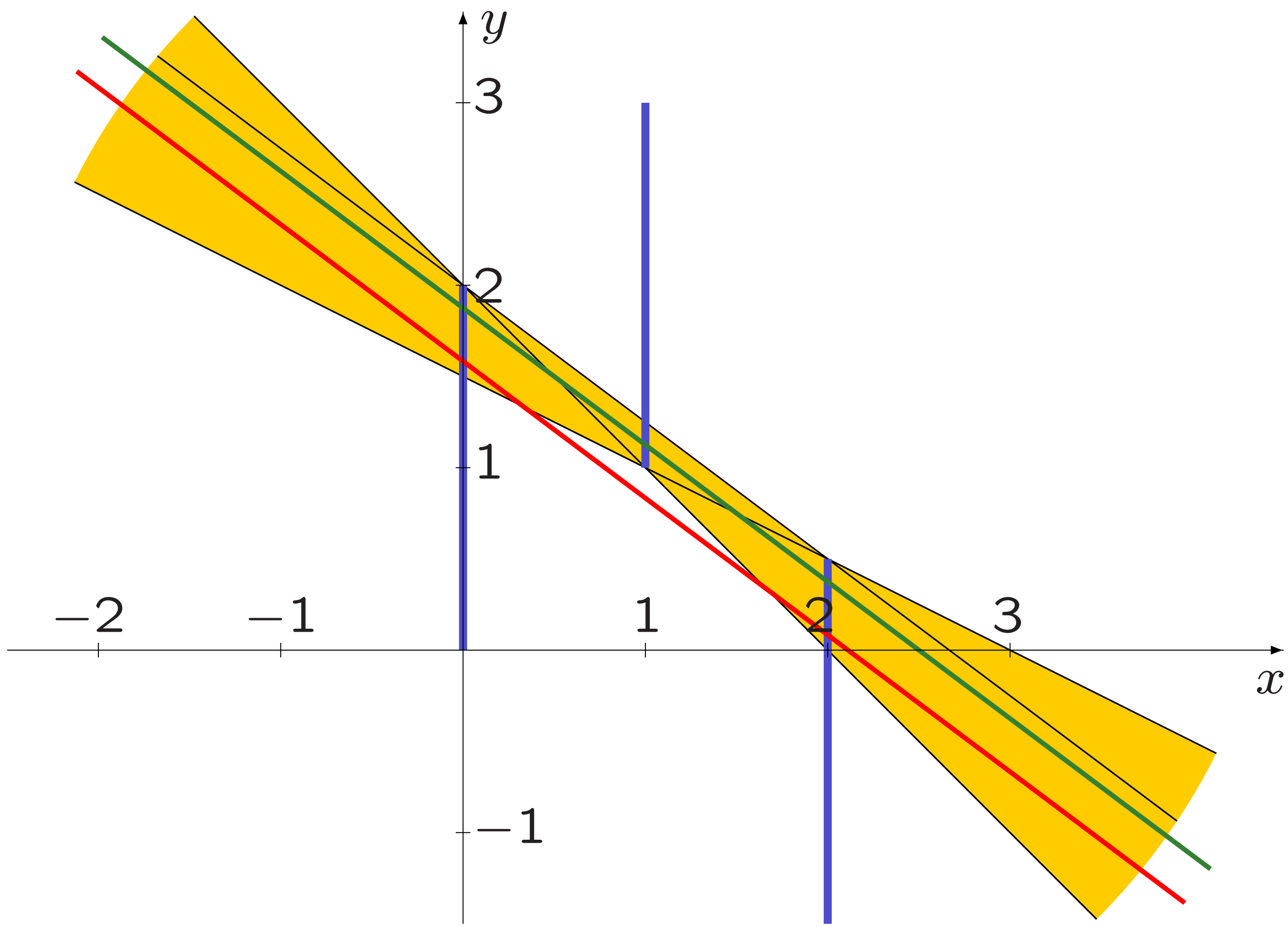
so that the estimate is equal to

$$\begin{pmatrix} \alpha^\star \\ \beta^\star \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 3 & -3 \\ -3 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ 2.5 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} -4.5 \\ 9.5 \end{pmatrix} = \begin{pmatrix} -3/4 \\ 19/12 \end{pmatrix} = \begin{pmatrix} -0.75 \\ 1.5833\ldots \end{pmatrix}.$$



In the space of variables α and β, the LSQ estimate (red point) does not lie in the information set (green triangle)

Comparison of the LSQ estimate with the set of regression lines consistent with the data



In the space of pairs (x, y), the straight line $y = \alpha^{\star}x + \beta^{\star}$ does not lie in the set of all the lines passing through the data intervals

Maximal consistency estimate

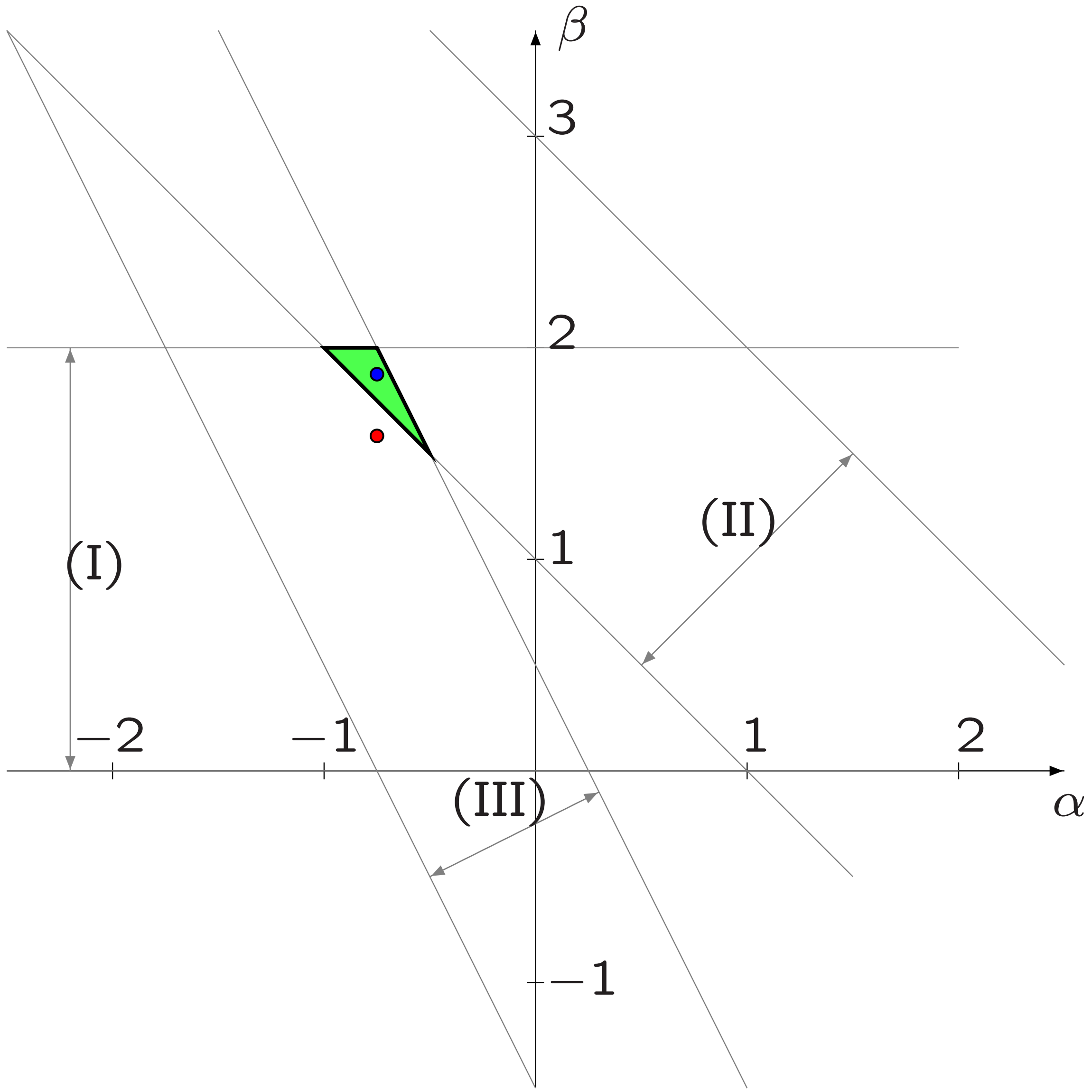
$$\max \text{Uss} = 0.125,$$

which means that the set of parameters
consistent with the data is not empty

The values of the parameters

$$\arg\max \text{Uss} = \begin{pmatrix} -0.75 \\ 1.875 \end{pmatrix}$$

correspond to a green line inside the yellow tube at the picture



. . . maximum consistency estimate

lies within the information set