# Расширение твинной арифметики и её применение в методах и алгоритмах двустороннего интервального оценивания

#### 03.03.2025

1.2.2. Математическое моделирование, численные методы и комплексы программ

Яворук Татьяна Университет ИТМО

#### План

#### 1. Введение

- Интервальный анализ
- Твины

#### 2. Приложение твинов

- Интервалы в изотопном анализе
- Классификация аминокислот
- Задача линейной регрессии
- Обработка LIDAR данных

#### 3. Заключение

#### Интервальный анализ

• Классическая интервальная арифметика:

$$\mathbb{IR} = \{\mathbf{x} = [\underline{\mathbf{x}}, \overline{\mathbf{x}}] : \underline{\mathbf{x}} < \overline{\mathbf{x}}, \underline{\mathbf{x}} \in \mathbb{R}\}.$$

Результатом любой арифметической операции ★ ∈ {+, -, ·, /} является интервал, такой что

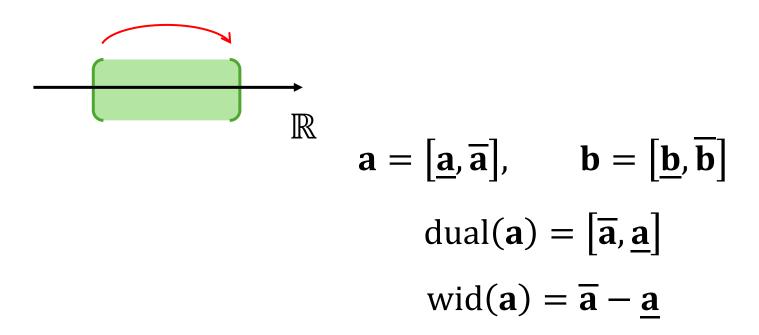
$$\mathbf{x} \star \mathbf{y} = \{ x \star y \mid x \in \mathbf{x}, y \in \mathbf{y} \}, \qquad x, y \in \mathbb{R}.$$

• Полная интервальная арифметика или арифметика Каухера:

$$\mathbb{KR} = \{\mathbf{x} = [\underline{\mathbf{x}}, \overline{\mathbf{x}}], \quad \underline{\mathbf{x}}, \overline{\mathbf{x}} \in \mathbb{R}\}.$$

Появились неправильные интервалы!

#### Операции с интервалами



Для неправильных интервалов ширина отрицательна

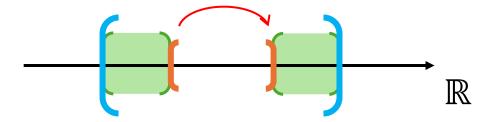
#### Операции с интервалами

Взятие максимума и минимума по включению

$$\wedge(\mathbf{a},\mathbf{b}) = \left[ \max(\underline{\mathbf{a}},\underline{\mathbf{b}}), \min(\overline{\mathbf{a}},\overline{\mathbf{b}}) \right]$$

$$V(\mathbf{a}, \mathbf{b}) = [\min(\underline{\mathbf{a}}, \underline{\mathbf{b}}), \max(\overline{\mathbf{a}}, \overline{\mathbf{b}})]$$

Для непересекающихся интервалов имеем неправильный минимум по включению

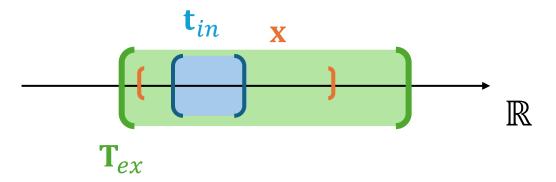


#### Твин

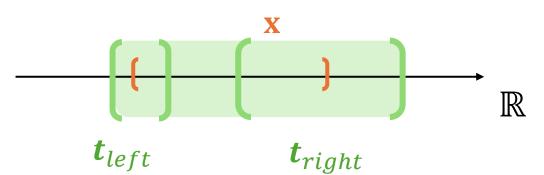
#### Определение

Твином по свойству вложенности ( $\subseteq$ ) называется следующий объект  $\mathbb{T} = [\mathbf{t}_{in}, \mathbf{T}_{ex}] = \{\mathbf{x} \in \mathbb{K}\mathbb{R} \mid \mathbf{t}_{in} \subseteq \mathbf{x} \subseteq \mathbf{T}_{ex}\},$ 

где  $\mathbf{t}_{in}$ ,  $\mathbf{T}_{ex}$  будем называть внутренней и внешней оценками.



Альтернатива: определение по свойству ≤:



## Арифметика твинов

$$\mathbb{T}_{k} = \left[ [\underline{\mathbf{t}}_{k}, \overline{\mathbf{t}}_{k}], [\underline{\mathbf{T}}_{k}, \overline{\mathbf{T}}_{k}] \right], k \in \{1, 2\}$$

$$\mathbb{T}_{1} \star \mathbb{T}_{2} = [\mathbf{t}_{in}, \mathbf{T}_{ex}], \qquad \star \in \{+, -, \cdot, /\}$$

$$\mathbf{T}_{ex} = \left[ \underline{\mathbf{T}}_{1}, \overline{\mathbf{T}}_{1} \right] \star \left[ \underline{\mathbf{T}}_{2}, \overline{\mathbf{T}}_{2} \right]$$

Твины Sainz et al.:

$$\mathbf{t}_{\text{in}} = \left[\underline{\mathbf{t}_1}, \overline{\mathbf{t}_1}\right] \star \left[\underline{\mathbf{t}_2}, \overline{\mathbf{t}_2}\right]$$

$$T_1 = ([a^-, a^+], [A^-, A^+]) T_2 = ([b^-, b^+], [B^-, B^+])$$

Define the numbers p and q as follows:

- if  $[a^-, a^+] \neq \emptyset$  and  $[b^-, b^+] \neq \emptyset$   $p = \min (a^- + B^+, b^- + A^+)$  $q = \max (a^+ + B^-, b^+ + A^-)$
- if  $[a^-, a^+] = \emptyset$  and  $[b^-, b^+] \neq \emptyset$   $p = b^- + A^+$   $q = b^+ + A^-$
- if  $[a^-, a^+] \neq \emptyset$  and  $[b^-, b^+] = \emptyset$   $p = a^- + B^+$   $q = a^+ + B^-$
- if  $[a^-, a^+] = \emptyset$  and  $[b^-, b^+] = \emptyset$  p, q is undefined

Define the functions  $\varphi$  and  $\psi$  as follows:

Z is empty set or it have one element,  $I_1, I_2 \in IR$ .

- if  $I_1 \cap I_2 = Z$ :  $\varphi(I_1, I_2) = min_{\subseteq}((c^-, c^+) \mid (c^- \in I_1 \text{ and } c^+ \in I_2) \lor (c^- \in I_2 \text{ and } c^+ \in I_1))$
- else  $\varphi(I_1, I_2) = \emptyset$

$$\psi(I_1, I_2) = \max_{\subseteq} ((c^-, c^+) \mid c^-, c^+ \in I_1 \cup I_2)$$

#### Sum

Let's define inner length of twin as  $|T|_l = |X_l|$  and outer length as |T| = |X|. If  $I \subseteq (\emptyset, X)$  than  $|T|_l = -1$ . It means that only outer estimation exists. Than

- if  $|T_1| \le |T_2|$ ,  $|T_2| \le |T_1|$ ,  $|T_1| + |T_2| = ((p,q), (A-+B^-, A^+ + B^+))$
- else  $T_1 + T_2 = ((\emptyset), (A + B^-, A^+ + B^+))$

#### Multiplication

Multiplication of two twins  $T_1, T_2$ 

- if  $|T_1|_l \neq -1$  and  $|T_2|_l \neq -1$ , then  $T_1 \cdot T_2 = (\psi(\varphi(a^-(B^-, B^+), a^+(B^-, B^+)), \varphi(b^-(A^-, A^+), b^+(A^-, A^+))), (A^-, A^+) \cdot (B^-, B^+))$
- if  $|T_1|_l = -1$  and  $|T_2|_l \neq -1$ , then  $T_1 \cdot T_2 = (\varphi(b^-(A^-, A^+), b^+(A^-, A^+)), (A^-, A^+) \cdot (B^-, B^+))$
- if  $|T_1|_l \neq -1$  and  $|T_2|_l = -1$ , then  $T_1 \cdot T_2 = (\varphi(a^-(B^-, B^+), a^+(B^-, B^+)), (A^-, A^+) \cdot (B^-, B^+))$
- if  $|T_1|_1 = -1$  and  $|T_2|_1 = -1$ , then  $T_1 \cdot T_2 = (\emptyset, (A^-, A^+) \cdot (B^-, B^+))$

Арифметика Нестерова определяется через специальные функции  $\varphi, \psi$ , и некоторые величины p, q.

В случае неправильных интервалов внутренняя оценка превращается в пустое множество

$$T_1 = ([a^-, a^+], [A^-, A^+]) T_2 = ([b^-, b^+], [B^-, B^+])$$

Define the numbers p and q as follows:

- if  $[a^-, a^+] \neq \emptyset$  and  $[b^-, b^+] \neq \emptyset$   $p = \min (a^- + B^+, b^- + A^+)$  $q = \max (a^+ + B^-, b^+ + A^-)$
- if  $[a^-, a^+] = \emptyset$  and  $[b^-, b^+] \neq \emptyset$   $p = b^- + A^+$   $q = b^+ + A^-$
- if  $[a^-, a^+] \neq \emptyset$  and  $[b^-, b^+] = \emptyset$   $p = a^- + B^+$   $q = a^+ + B^-$
- if  $[a^-, a^+] = \emptyset$  and  $[b^-, b^+] = \emptyset$  p, q is undefined

Define the functions  $\varphi$  and  $\psi$  as follows:

Z is empty set or it have one element,  $I_1, I_2 \in IR$ .

• if  $I_1 \cap I_2 = Z$ :  $\varphi(I_1, I_2) = min_{\subseteq}((c^-, c^+) \mid (c^- \in I_1 \text{ and } c^+ \in I_2) \lor (c^- \in I_2 \text{ and } c^+ \in I_1))$ else  $\varphi(I_1, I_2) = \emptyset$ 

$$\psi(I_1, I_2) = max_{\subseteq}((c^-, c^+) \mid c^-, c^+ \in I_1 \cup I_2)$$

#### Sum

Let's define inner length of twin as  $|T|_l = |X_l|$  and outer length as |T| = |X|. If  $I \subseteq (\emptyset, X)$  than  $|T|_l = -1$ . It means that only outer estimation exists. Than

- if  $|T_1| \le |T_2| \lor |T_2| \le |T_1| \lor |T_1| + T_2 = ((p,q), (A-+B^-, A^+ + B^+))$
- else  $T_1 + T_2 = ((\emptyset), (A + B^-, A^+ + B^+))$

#### Multiplication

Multiplication of two twins  $T_1, T_2$ 

- if  $|T_1|_l \neq -1$  and  $|T_2|_l \neq -1$ , then  $T_1 \cdot T_2 = (\psi(\varphi(a^-(B^-, B^+), a^+(B^-, B^+)), \varphi(b^-(A^-, A^+), b^+(A^-, A^+))), (A^-, A^+) \cdot (B^-, B^+))$
- if  $|T_1|_l = -1$  and  $|T_2|_l \neq -1$ , then  $T_1 \cdot T_2 = (\varphi(b^-(A^-, A^+), b^+(A^-, A^+)), (A^-, A^+) \cdot (B^-, B^+))$
- if  $|T_1|_l \neq -1$  and  $|T_2|_l = -1$  then  $T_1 \cdot T_2 = (\varphi(a^-(B^-, B^+), a^+(B^-, B^+)), (A^-, A^+) \cdot (B^-, B^+))$
- if  $|T_1|_l = -1$  and  $|T_2|_l = -1$  then  $T_1 \cdot T_2 = (\emptyset, (A^-, A^+) \cdot (B^-, B^+))$

Новое определение в современной нотации с использованием минимума максимума по включению и дуализации:

$$\mathbf{t}_{\text{in}} = \bigvee \left( \text{dual} \left( \bigwedge \left( \underline{\mathbf{t}}_{1} \star \left[ \underline{\mathbf{T}}_{2}, \overline{\mathbf{T}}_{2} \right], \overline{\mathbf{t}}_{1} \star \left[ \underline{\mathbf{T}}_{2}, \overline{\mathbf{T}}_{2} \right] \right) \right), \quad \text{dual} \left( \bigwedge \left( \underline{\mathbf{t}}_{2} \star \left[ \underline{\mathbf{T}}_{1}, \overline{\mathbf{T}}_{1} \right], \overline{\mathbf{t}}_{2} \star \left[ \underline{\mathbf{T}}_{1}, \overline{\mathbf{T}}_{1} \right] \right) \right) \right)$$

$$\mathbf{t}_{\text{in}} = V(\mathbf{t}_1 \star \text{dual } \mathbf{T}_2, \mathbf{t}_2 \star \text{dual } \mathbf{T}_1)$$

Появились неправильные интервалы во внутренней оценке ©

#### Некоторые операции с твинами

• Умножение на скаляр

$$\lambda \mathbb{T} = [\lambda \mathbf{t}_{\text{in}}, \lambda \mathbf{T}_{\text{ex}}]$$

• Индекс (мера, коэффициент) Жаккара – мера сходства двух

множеств 
$$(J(A,B) = \frac{|A \cap B|}{|A \cup B|})$$
:

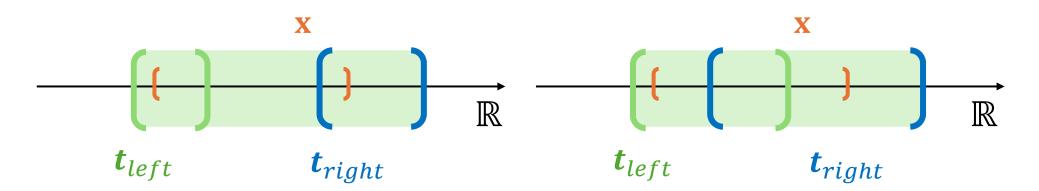
$$Ji(\mathbb{T}) = \frac{wid(\mathbf{t}_{in})}{wid(\mathbf{T}_{ex})}$$
.

 $Ji(\mathbb{T}) < 0$ 
(Может быть!)

## Возникновение неправильных интервалов внутренних оценок

Оценка экспериментальных данных с не одной причиной неопределённости

Пример: измерение температуры (неопределенность явления сопоставима с неопределённостью измерения)



13/31

#### **IUPAC** Periodic Table of Elements and Isotopes

Интервальные значения масс почти всех элементов, образующих органические соединения.

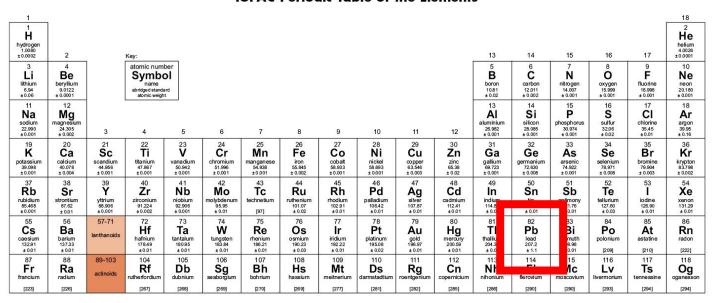
Важнейшие: водород и углерод

https://applets.kcvs.ca/IPTEI/IPTEI.html



#### **IUPAC** Periodic Table of Elements

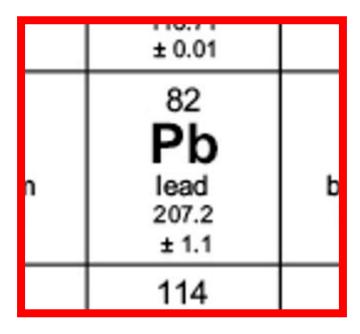
#### **IUPAC Periodic Table of the Elements**





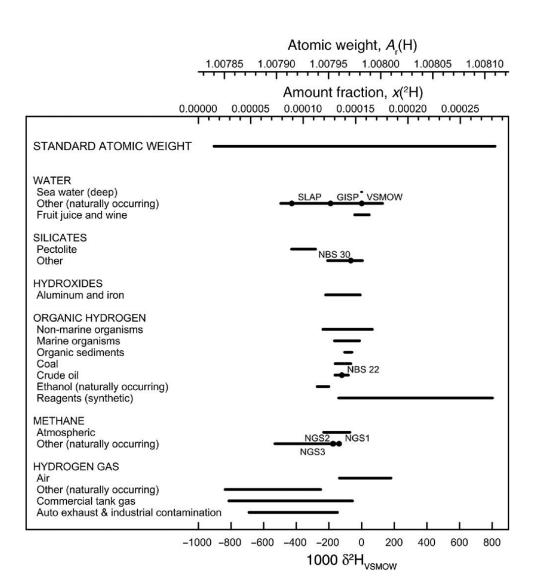
57 <b>La</b> lanthanum 138.91 ± 0.01	58 Ce cerium 140.12 ± 0.01	59 Pr praseodymium 140.91 ±0.01	60 Nd neodymium 144.24 ±0.01	61 Pm promethium	62 <b>Sm</b> samarium 150.36 ± 0.02	63 Eu europium 151.96 ± 0.01	64 <b>Gd</b> gadolinium 157.25 ± 0.03	65 <b>Tb</b> terbium 158.93 ± 0.01	66 <b>Dy</b> dysprosium 162.50 ± 0.01	67 Ho holmium 164.93 ±0.01	68 Er erbium 167.26 ± 0.01	69 Tm thulium 168.93 ± 0.01	70 <b>Yb</b> ytterbium 173.05 ± 0.02	71 Lu lutetium 174.97 ± 0.01
AC actinium	90 <b>Th</b> thorium 232.04 ± 0.01	91 Pa protactinium 231.04 ±0.01	92 U uranium 238.03 ±0.01	93 Np neptunium [237]	94 Pu plutonium [244]	95 Am americium [243]	96 <b>Cm</b> curium [247]	97 <b>Bk</b> berkelium [247]	98 Cf californium	99 Es einsteinium	100 Fm fermium	101 Md mendelevium [258]	102 No nobelium [259]	103 Lr lawrendium

For notes and updates to this table, see www.iupac.org. This version is dated 4 May 2022. Copyright © 2022 IUPAC, the International Union of Pure and Applied Chemistry.



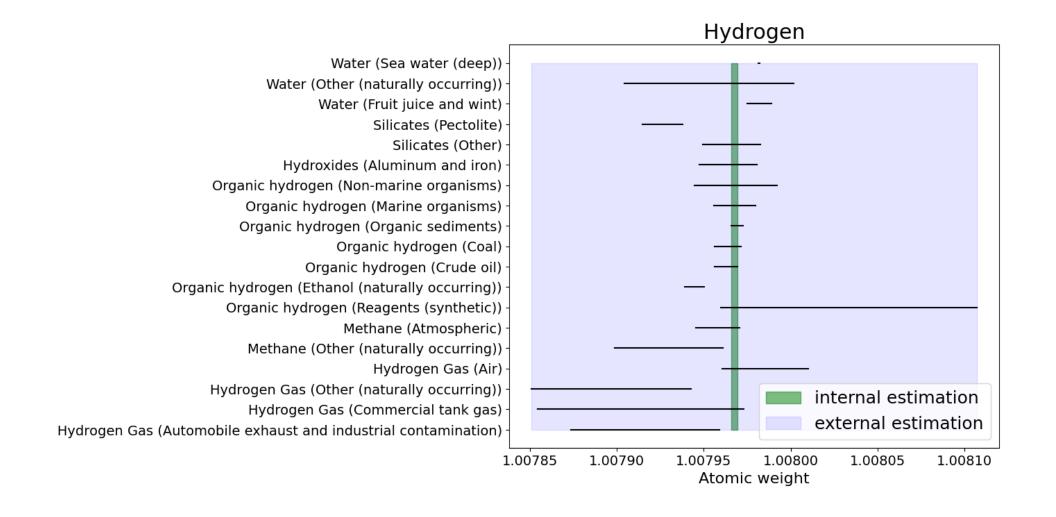
Неопределенность атомной массы свинца превышает единицу атомной массы!

#### Изотопная информация водорода



Изотопная информация используется в различных науках и индустрии: космологии, климатологии, биологии, науках о земле, криминалистике, палеонтологии и других.

## Атомный вес водорода



#### Сумма твинов

Tyrosine -  $C_9H_{11}NO_3$ Phenylalanine -  $C_9H_{11}NO_2$ Tryptophan -  $C_{11}H_{12}N_2O_2$ 

Условие правильности внутренней оценки у Нестерова

#### Sum

Let's define inner length of twin as  $|T|_l = |X_l|$  and outer length as |T| = |X|. If  $I \subseteq (\emptyset, X)$  than  $|T|_l = -1$ . It means that only outer estimation exists. Than

- if  $|T_1| \le |T_2|_l \lor |T_2| \le |T_1|_l T_1 + T_2 = ((p,q), (A + B^-, A^+ + B^+))$
- else  $T_1 + T_2 = ((\emptyset), (A + B^-, A^+ + B^+))$

## Результаты

Индексы Жаккара

$$Ji(\mathbb{T}) = \frac{wid(\mathbf{t}_{in})}{wid(\mathbf{T}_{ex})}$$

для аминокислот с разным типом твинов.

В фенилаланине прослеживается тип фотосинтеза.

Acid	Type of twin	Jaccard index			
Tyrosine	Nesterov	-0.01497			
Phenylalanine	Nesterov	0.0183			
Tryptophan	Nesterov	-0.0079			
Tyrosine	Sainz et. al.	0.00999			
Phenylalanine	Sainz et. al.	0.0095			
Tryptophan	Sainz et. al.	0.00913			

Tyrosine — 
$$C_9H_{11}NO_3$$
  
Phenylalanine —  $C_9H_{11}NO_2$   
Tryptophan —  $C_{11}H_{12}N_2O_2$ 

## Открытые вопросы

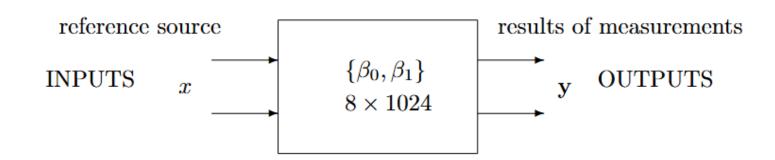
• Как построить твин по данным?

• Другие приложения?

## Задача линейной регрессии. Частный случай

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

$$\mathbf{Y} = \beta_0 + \beta_1 \cdot X$$



## Множества решений ИСЛАУ

$$\Xi_{uni} = \{ x \in R^m | (\exists A \in A) (\exists B \in B) (Ax = b) \},$$
  
$$\Xi_{tol} = \{ x \in R^m | \{ x \in R^m | (\forall A \in A) (\exists B \in B) (Ax = b) \}.$$

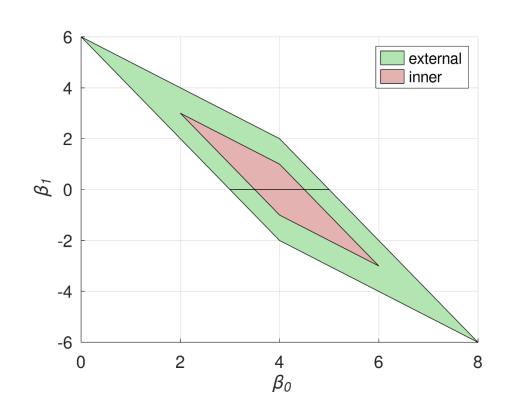
Распознающий функционал

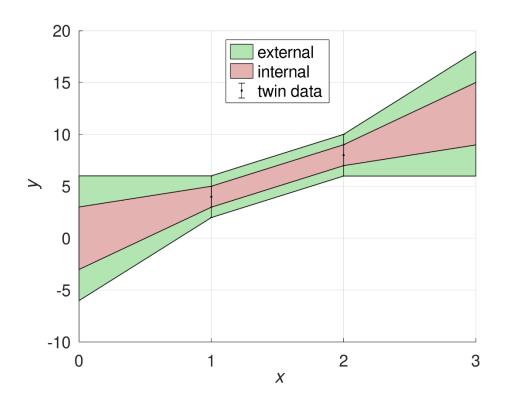
$$Tol(\beta, X, y) = \min_{1 \le i \le n} \{ rad \ y_i - | mid \ y_i - \sum_{j = \overline{1, m}} x_{ij} \beta_j \ | \}$$

$$Tol_{i}(\beta, X, y) = rad y_{i} - |mid y_{i} - \sum_{j=\overline{1,m}} x_{ij}\beta_{j}|\} < 0 \longrightarrow i$$

$$22/3$$

## Задача линейной регрессии





А что если система с внутренними оценками имеет пустое множество решений? 23/31

## Метод I

1. Посчитать

$$Tol_{i}(\beta, X, y) = rad y_{i} - |mid y_{i} - \sum_{j=\overline{1,m}} x_{ij}\beta_{j}|\}$$

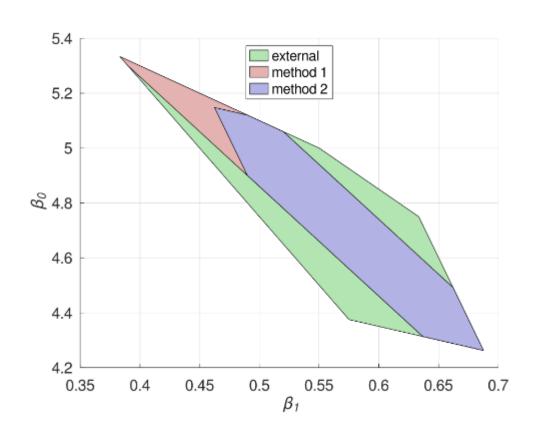
- 2. Найти индексы отрицательных  $Tol_i$  .
- 3. Взять внешние интервалы для найденных индексов.

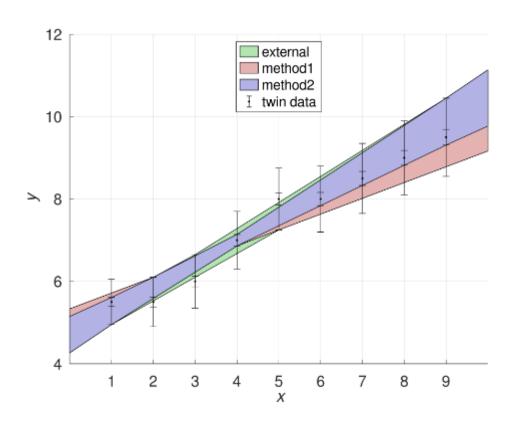
## Метод II

Добавить в систему новые уравнения, соответствующие другим правым частям:



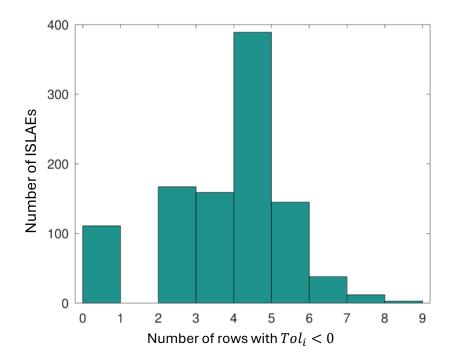
## Пример. Множества решений





#### Приложение

Рассмотрим данные по калибровке 1024 приборов по линейной регрессии по 9 уровням. Для 100 с небольшим случаев все данные совместны. Для 400 случаев прямая проходит через половину измерений (каждая точка - 100 измерений)



#### Преимущества и недостатки алгоритмов

- Уменьшение множества решений
- Увеличение размера системы
- Как выбрать интервалы для двойного оценивания?

#### Заключение

- В работе расширена твинная арифметика Нестерова (добавлена полная интервальная арифметика)
- Рассмотрен способ использования твинов в изотопном анализе данных
- Приведены два алгоритма решения Твинных СЛАУ

#### To-do

- Развитие численных методов с использованием твинов
- Рассмотрение различных способов составления твина из изотопных данных
- Рассмотрение других постановок задач с двусторонним оцениванием

## Спасибо за внимание!

tatianaiavoruk@gmail.com

github.com/Tatiana655/Twins

