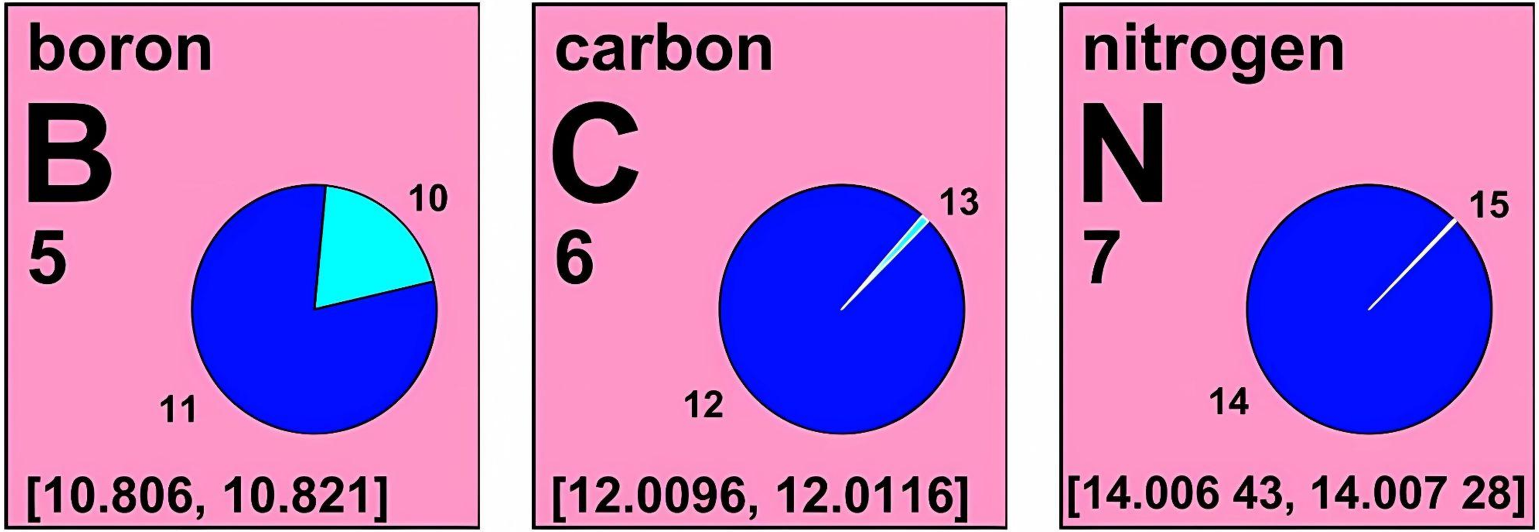
Аффинные и функциональные арифметики: краткий обзор основных идей и разработок

Скорик Дмитрий Александрович

ФИЦ ИВТ, Новосибирск, 2024

Описание неопределенности

Атомные веса

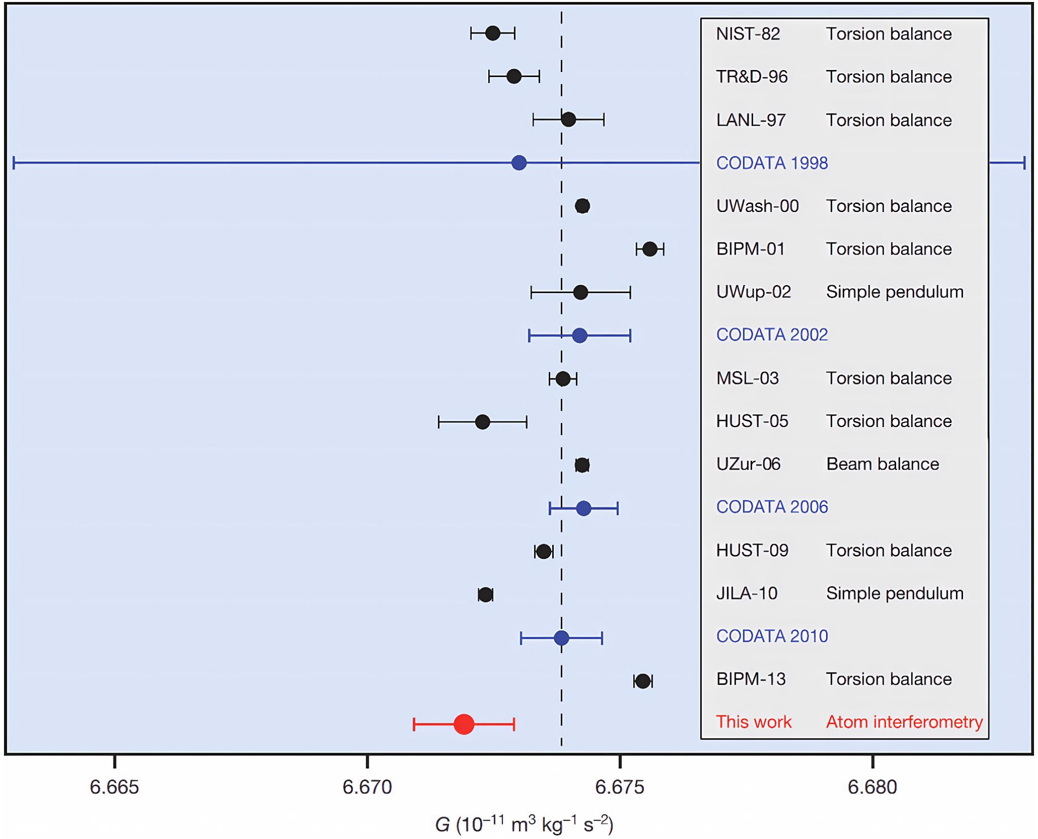


Описание неопределенности

Физические константы,

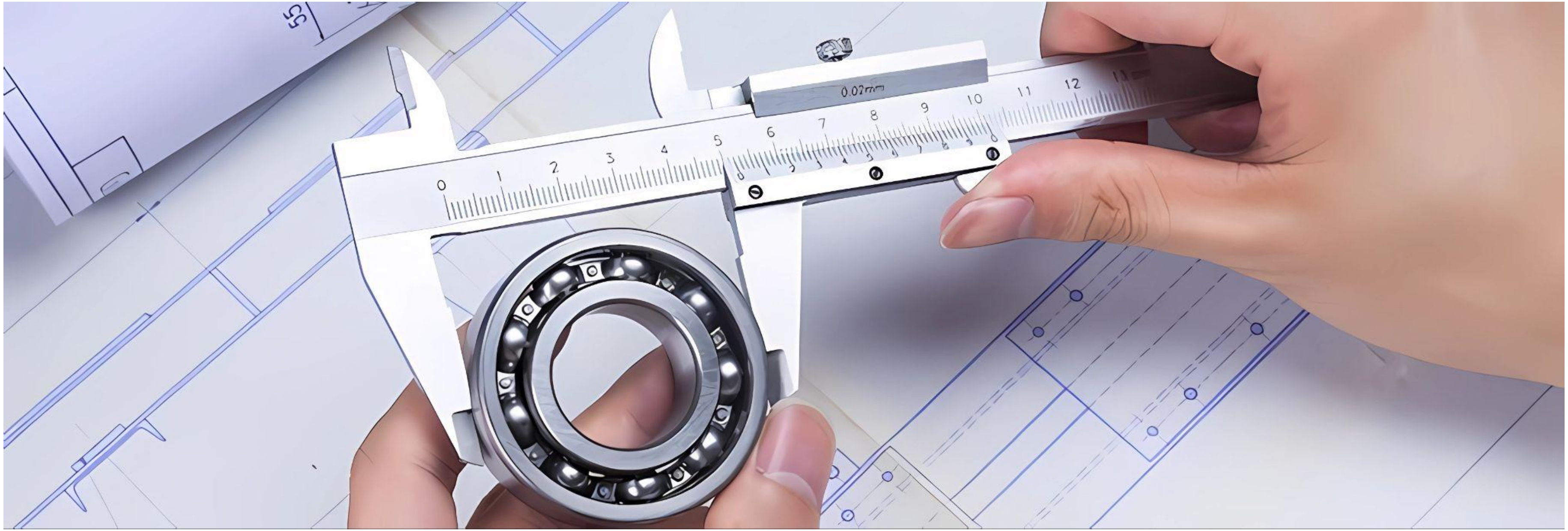
гравитационная константа

$$G \approx 6.6743 \cdot 10^{-11}$$



Описание неопределенности

Погрешность измерений



Описание неопределенности

Задачи, чувствительные к входным данным, встречаются не только в задачах по расчетам космических полетов. Простой пример Мюллера:

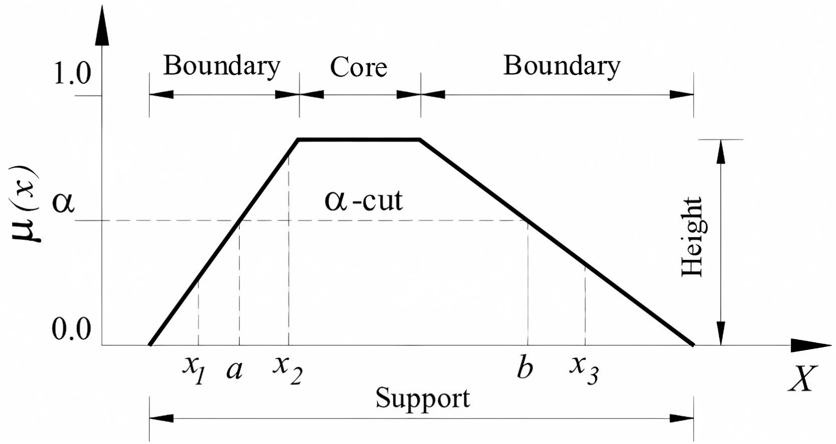
$$f(y, z) = 108 - \frac{815 - 1500/z}{y}$$

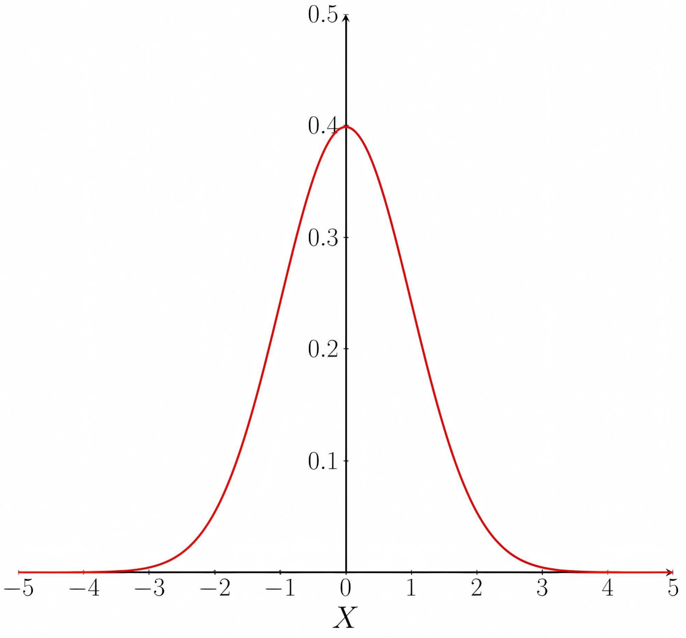
$$x_0 = 4, \quad x_1 = 4.25$$

$$x_i = f(x_{i-1}, x_{i-2})$$

Способы описания неопределенностей

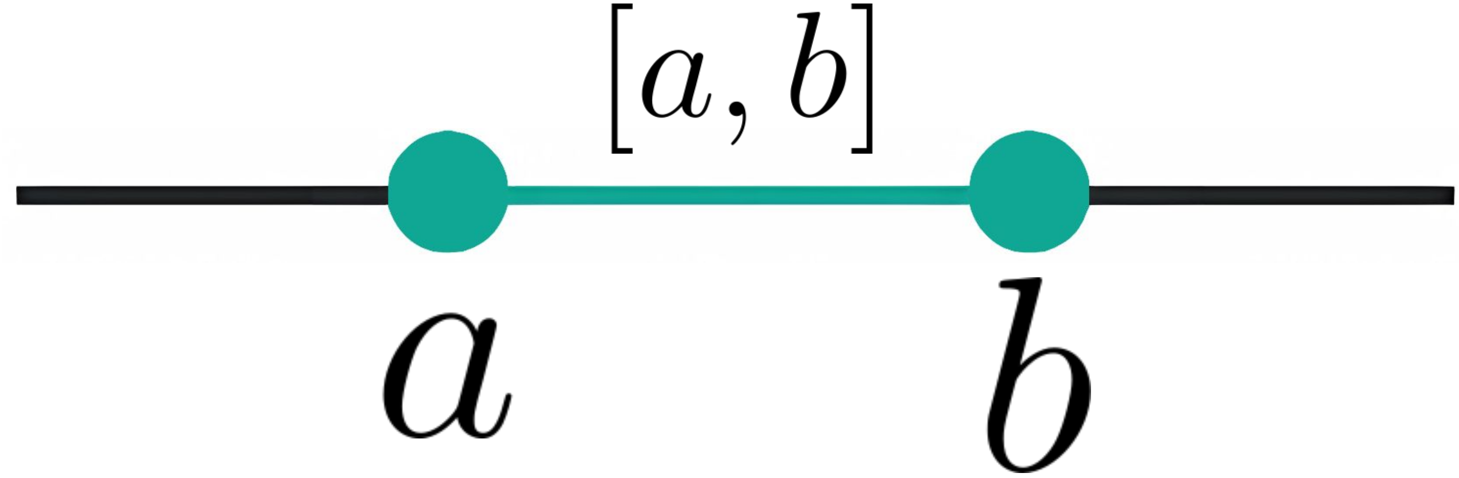
- нечеткие множества
- распределения в теории вероятностей
- интервалы
- …





Интервал

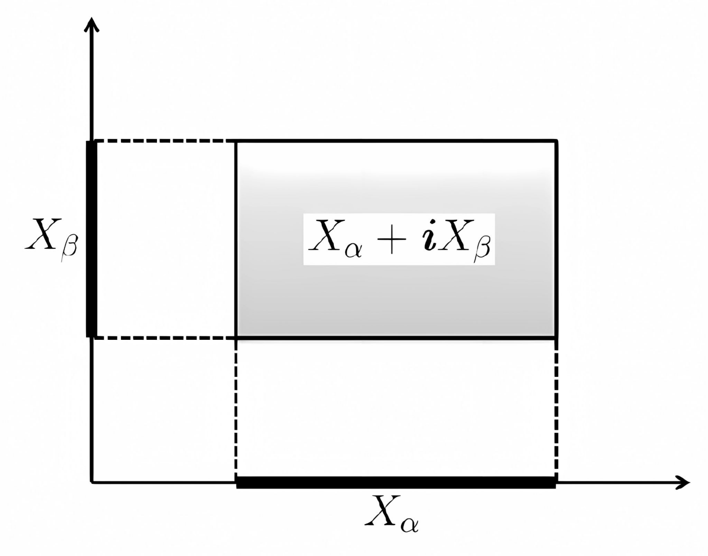
Классический и простой объект для описания разного рода неопределенностей — интервал — замкнутое множество на вещественной прямой, описываемое упорядоченной парой чисел — границами интервала.

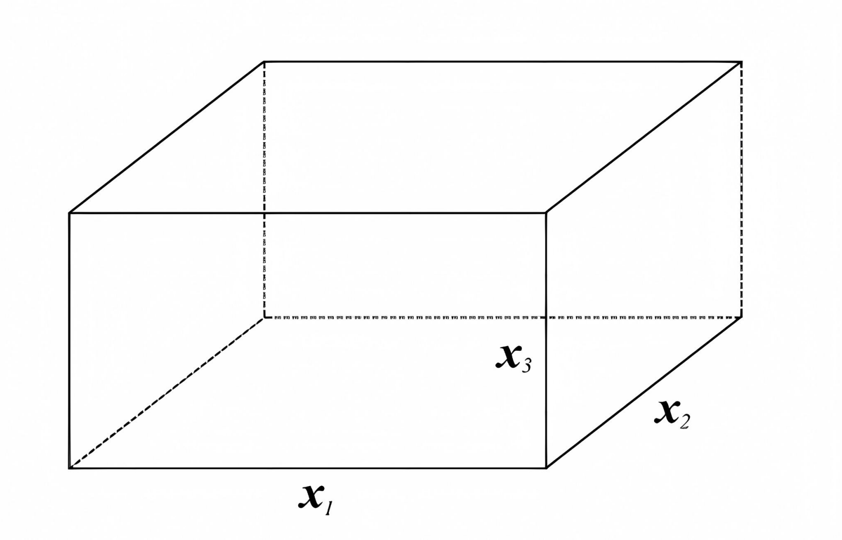


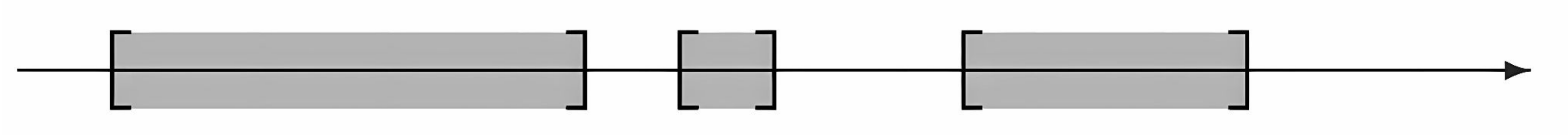
Интервалы бывают разных видов

- одномерный шар в представлении “центр-радиус”
- брус, когда говорим про многомерные интервалы
- комплексные интервалы для описания различного вида множеств на комплексной плоскости
- мультиинтервалы для работы с массивом интервалов, как единым объектом для вычислений
- твины для описания интервала с интервальными концами
- и другие…

Также интервалы имеют разные представления







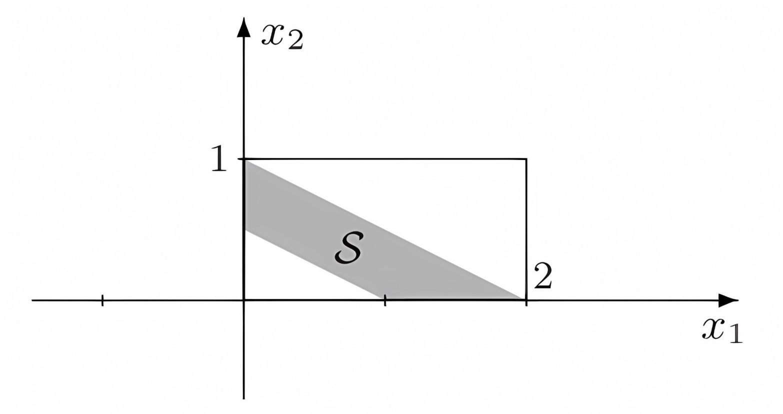
Есть еще причины для новых интервалов

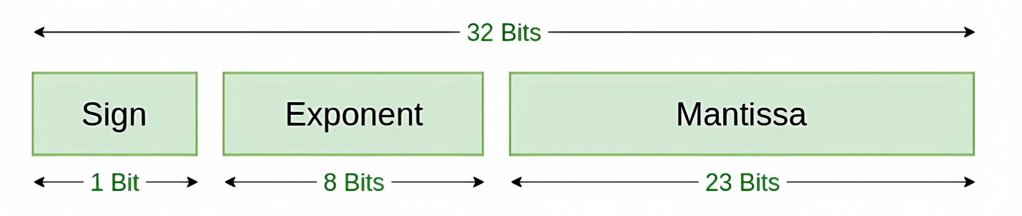
Слабые алгебраические свойства, пример — субдистрибутивность

$$\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) \subseteq \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{c}$$

Есть еще причины для новых интервалов

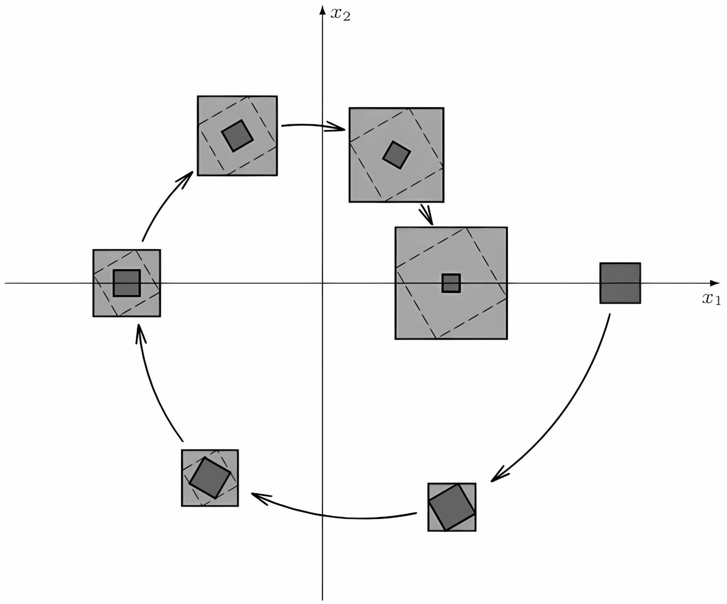
Учет совместной области значений между неизвестными в длинной цепочки вычислений, когда неприменимы методы символьных вычислений или накопление ошибок округлений начинает сильно влиять на итоговый ответ.





Есть еще причины для новых интервалов

Эффект обертывания — эффект, возникающий при несовпадении формы множеств для описания объектов и формы множеств самих описываемых объектов.



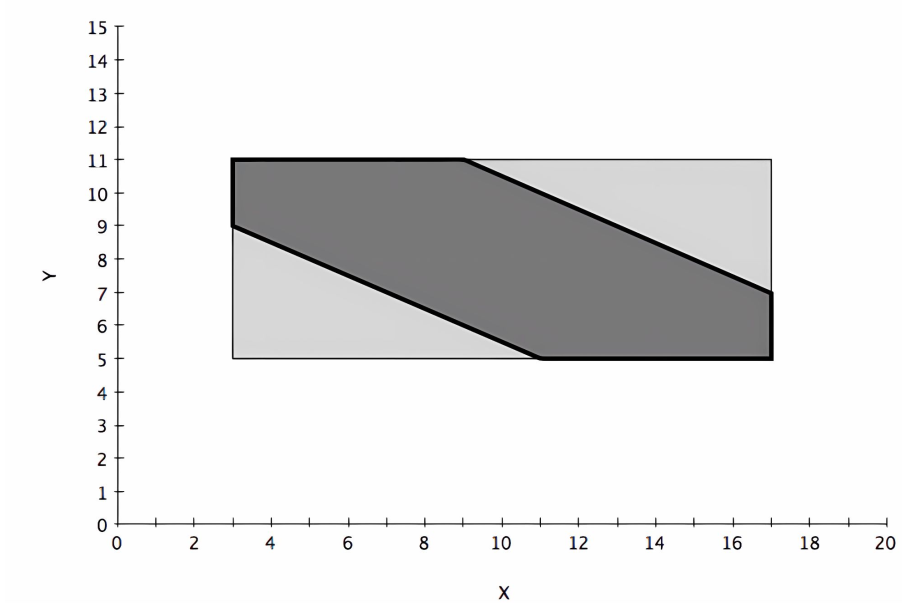
Интервалы бывают разных видов

- одномерный шар в представлении “центр-радиус”
- брус, когда говорим про многомерные интервалы
- комплексные интервалы для описания различного вида множеств в комплексной плоскости
- мультиинтервалы для работы с массивом интервалов, как единым объектом для вычислений
- твины для описания интервала с интервальными концами
- **аффинные интервалы**
- **функциональные интервалы**
- другие…

Аффинная арифметика

Аффинная арифметика

Основная идея — описать совместную область значений более детально — вместо декартового произведения двух областей значений переменных взять вложенный в это произведение выпуклый многогранник.



Описание выпуклых многогранников

Выпуклые многогранники можно описывать по-разному.

Есть частный случай — параллелотопы или зонотопы. Их описывать проще, чем многогранники общего вида. При этом они дают качественное улучшение в описании совместной области значений.

Как правило, зонотопы описываются аффинными формами.

Похожие модели аффинной арифметики были разработаны с небольшой разницей Хансеном и Jorge Stolfi.

Аффинные формы

Аффинная форма представляет собой выражение вида:

$$x = x_0 + x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n$$

$$x_i \in \mathbb{R} \qquad \varepsilon_i \in [-1, 1]$$

x_i — частичные отклонения

x_0 — центр аффинной формы

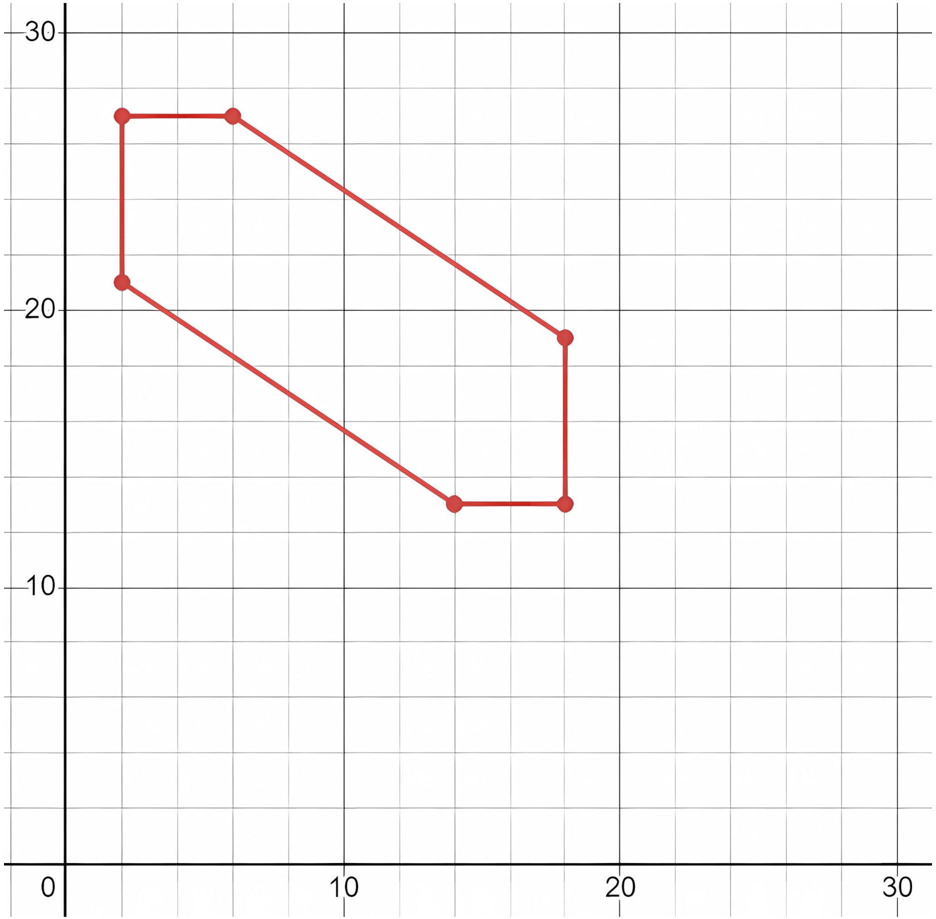
ε_i — символы шума

Пример

$$x = 10 + 2\varepsilon_1 - 6\varepsilon_2$$

$$y = 20 + 4\varepsilon_2 + 3\varepsilon_3$$

Совместная область значений описывается точнее, чем декартово произведение



Арифметика аффинных форм

С аффинными формами достаточно просто определить основные арифметические операции и функции.

Если операция аффинная (сложение, умножение на константу), тогда результат ее действия на аффинные формы будет давать аффинную форму.

$$x = x_0 + \sum_i x_i \varepsilon_i \qquad y = y_0 + \sum_i y_i \varepsilon_i$$

$$x \pm y = x_0 \pm y_0 + \sum_i (x_i \pm y_i) \varepsilon_i$$

$$Nx = Nx_0 + \sum_i N x_i \varepsilon_i$$

Арифметика аффинных форм

Если же операция или функция не аффинная (умножение, деление, sin, …), тогда вычисляют аффинное приближение результата в виде аффинной формы

$$z = G(x, y, \ldots) + z_k \varepsilon_k$$

ε_k — новая символьная переменная, не встречающаяся ранее

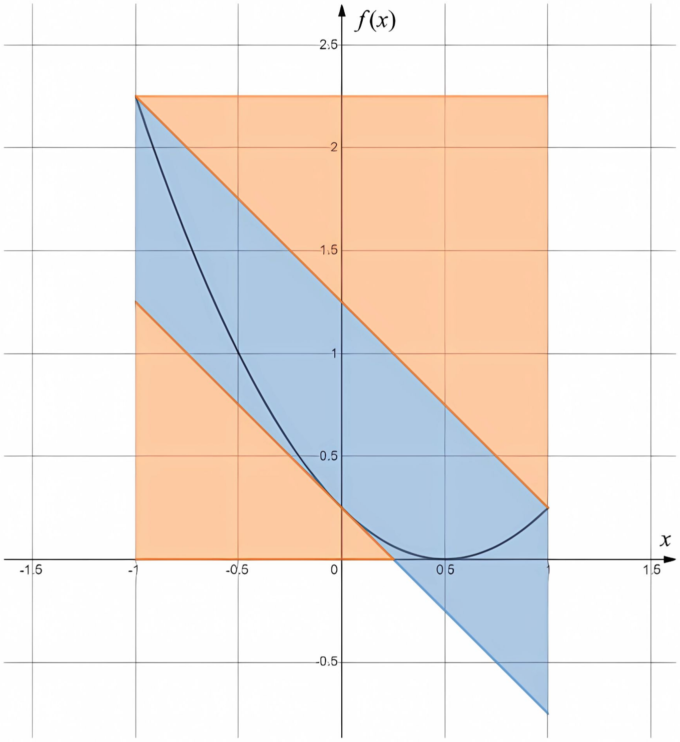
G — приближение рассматриваемой функции в некотором смысле

Про аффинное приближение

Стоит помнить, что нахождение аффинного приближения, в зависимости от алгоритма, может давать худшие результаты по сравнению с классической интервальной арифметикой.

Пример на картинке — поиск аффинного приближения в смысле минимизации Чебышевской метрики для

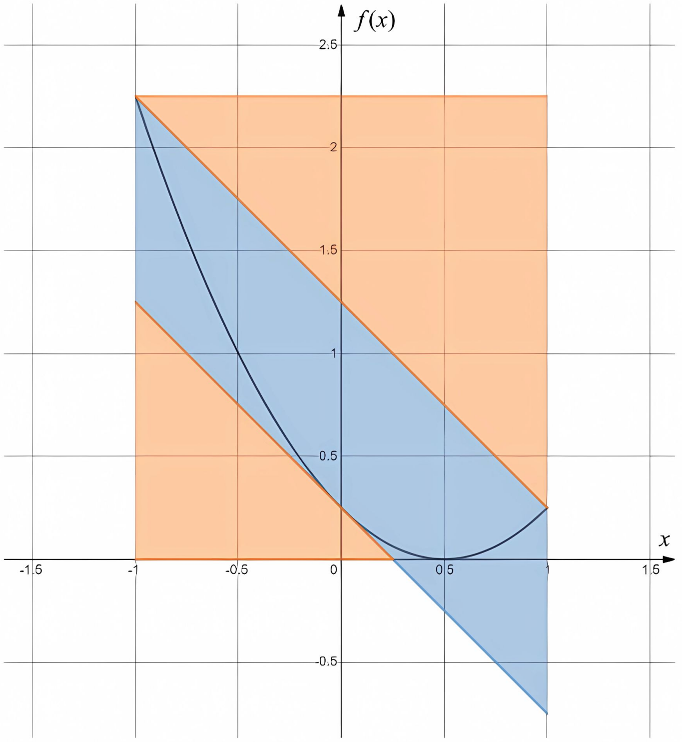
$$(x - 0.5)^2 \qquad x = \varepsilon_1$$



Про аффинное приближение

Идея исправления данного недостатка была предложена Ахмеровым Р. Р. Она состоит в использовании информации от вычислений в классической интервальной арифметике при каждом шаге обработки аффинных форм.

Полученная модификация аффинной арифметики была названа аффинно-интервальной.



Есть и технический нюанс

Если цепочка вычислений должна обладать свойством доказательности, то в классической интервальной арифметике мы накапливаем ошибки округления в результирующий интервал при каждой операции, тем самым каждый раз расширяя его.

В такой же ситуации для аффинной арифметики, ошибки округления при каждом вычислении должны быть учтены новым символьным членом даже для аффинных операций.

Про “склейку” символов

Итого, если мы имеем дело с доказательными вычислениями в АА, то количество членов в аффинной форме может быть пропорционально количеству вычислительных операций.

Чтобы уменьшить неконтролируемый рост членов в аффинной форме, часто применяют «склейку символов» — преобразование формы, когда два и более символов заменяются меньшим набором новых символов.

Геометрически это значит, что совместная область значений будет описываться более простым по количеству вершин зонотопом.

Применение на практике

Аффинная арифметика дает преимущество тогда, когда совместная область значений неизвестных существенно отличается от декартового произведения.

Это происходит за счет того, что мы учитываем в цепочках вычислений некоторую информацию о связи между символьными переменными.

Особенно хорошо это проявляется при выполнении аффинных или простых неаффинных операций. Один из примеров целевого применения АА: задача нахождения оценок для объединенного множества решений ИСЛАУ, например, методом Гаусса, где требуется, преимущественно, только складывать и умножать аффинные формы.

Постановка задачи с ИСЛАУ

Классическая постановка задачи

$$\begin{cases} \boldsymbol{a}_{11}x_1 + \boldsymbol{a}_{12}x_2 + \ldots + \boldsymbol{a}_{1n}x_n = \boldsymbol{b}_1, \\ \boldsymbol{a}_{21}x_1 + \boldsymbol{a}_{22}x_2 + \ldots + \boldsymbol{a}_{2n}x_n = \boldsymbol{b}_2, \\ \quad\vdots \qquad\quad \vdots \qquad\quad \ddots \qquad \vdots \qquad\quad \vdots\ , \\ \boldsymbol{a}_{n1}x_1 + \boldsymbol{a}_{n2}x_2 + \ldots + \boldsymbol{a}_{nn}x_n = \boldsymbol{b}_n. \end{cases}$$

$$\Xi_{uni}(F, \boldsymbol{a}, \boldsymbol{b}) = \{\, x \in \mathbb{R}^n \mid (\exists a \in \boldsymbol{a})(\exists\, b \in \boldsymbol{b})(\, F(a, x) = b \,) \,\}$$

Алгоритмическая сложность

Алгоритмическая сложность нахождения оценки объединенного множества решения ИСЛАУ методом Гаусса с использованием классических интервалов

$$\approx O(n^3)$$

Алгоритмическая сложность нахождения оценки объединенного множества решения ИСЛАУ методом Гаусса с использованием через АА

$$\approx O(n^5)$$

Это накладывает ограничение на размерность решаемых задач. В случае независимых интервалов эта размерность будет меньше, чем при наличии априорной информации о связях в матрице — например, матрица симметрична или антисимметрична.

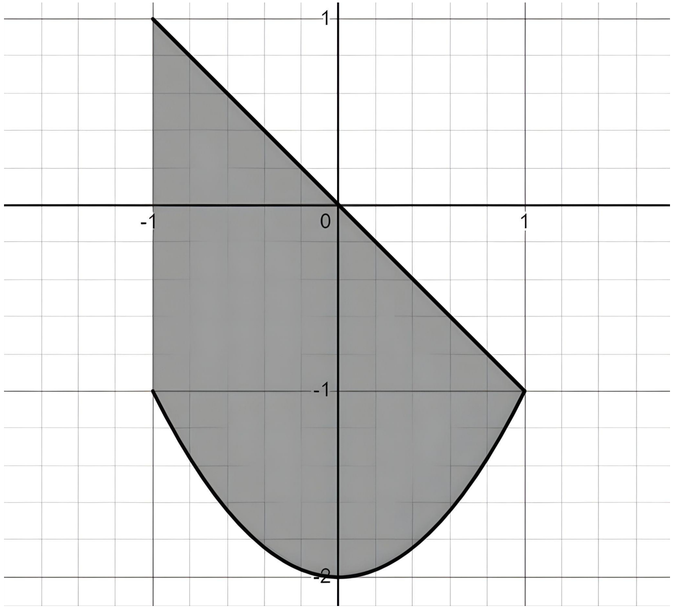
Резюме

- идея АА — учет линейных связей между неизвестными в цепочке вычислений
- каждая операция оперирует массивом классических интервалов вместо одного, но есть возможность “склеивать” символьные переменные для уменьшения этого количества
- в доказательных вычислениях каждая арифметическая операция генерирует новую символьную переменную
- возможна ситуация с неаффинными операциями, когда классическая арифметика дает результаты более точные, но есть интервально-аффинная арифметика авторства Ахмерова Р. Р., которая это учитывает
- если информации о связях между неизвестными нет или мало, цепочка вычислений длинна, то возможна ситуация, когда накопление ошибок округления “съест” преимущество от учета связей

Функциональные интервалы

Функциональные интервалы

Основная идея — описание границ интервала в виде функций от одной или нескольких переменных (параметров) для использования их геометрических свойств при решении вычислительных задач.



Функциональные интервалы

Функциональным интервалом будем называть интервал, границы которого представляются функциями:

$$L : [-1, 1]^n \to \mathbb{R}, \quad U : [-1, 1]^n \to \mathbb{R}$$

которые удовлетворяют свойству:

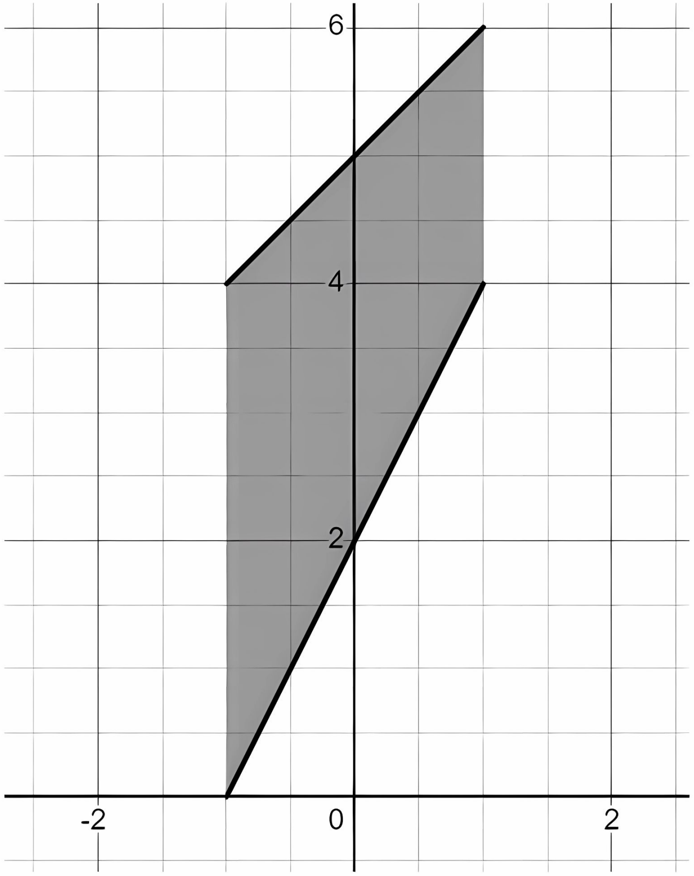
$$\forall x_i \in [-1, 1] \quad L(x_1, \ldots, x_n) \leq U(x_1, \ldots, x_n)$$

Функции L и U будем называть нижней и верхней границей интервала соответственно, а записывать интервал, как [L, U].

Пример

Функциональный интервал

$$[2x + 2, x + 5]$$



Арифметика функциональных интервалов

По существу функциональный интервал представляет собой параметрическое семейство классических интервалов. Так, например, функциональный интервал без параметров — классический интервал.

$$[a, b] \Leftrightarrow L(x_1, \ldots, x_n) = a, U(x_1, \ldots, x_n) = b$$

Также аффинные формы — это частный случай функциональных интервалов.

$$x = x_0 + \sum_i x_i \varepsilon_i = [x_0 + \sum_i x_i \varepsilon_i, x_0 + \sum_i x_i \varepsilon_i]$$

Арифметика функциональных интервалов

Из-за того, что вид функции в определении границ функциональных интервалов никак не фиксируется, можно явно описать только те операции с интервалами, которые явно определяются через границы интервала. Например, операции сложения, вычитания, умножения на константу.

$$[L_1, U_1] + [L_2, U_2] = [L_1 + U_1, L_2 + U_2]$$

$$N \cdot [L, U] = [NL, NU], N > 0$$

$$N \cdot [L, U] = [NU, NL], N < 0$$

Арифметика функциональных интервалов

Более сложные операции, например, умножение/деление, нужно описывать в частном порядке, опираясь на выбранный вид функций для границ.

$$F([L_1, U_1], [L_2, U_2]) = \\ [G_L(L_1, U_1, L_2, U_2), G_U(L_1, U_1, L_2, U_2)]$$

G_L — приближение снизу рассматриваемой функции F

G_U — приближение сверху рассматриваемой функции F

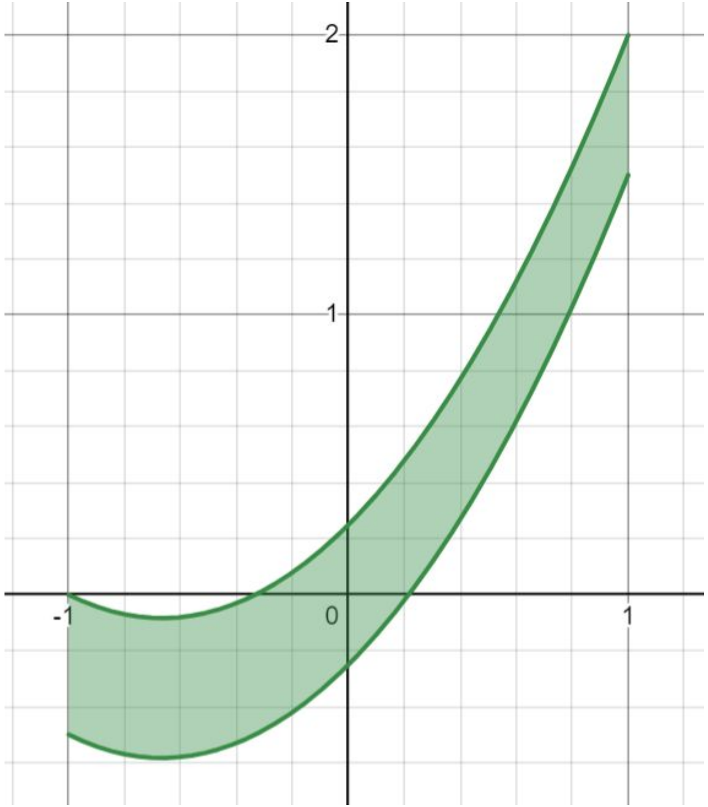
Построение функциональных интервалов

Фиксированного способа для построения функциональных интервалов нет, все зависит от специфики задачи, к которой применяются функциональные интервалы.

Пример: квадратичный функциональный интервал-коридор

Квадратичные функциональные интервалы-коридоры также строятся из разложения функции в ряд Тейлора — специальным образом обрабатывается остаточный член четного порядка. Используются для вычислений в динамических системах.

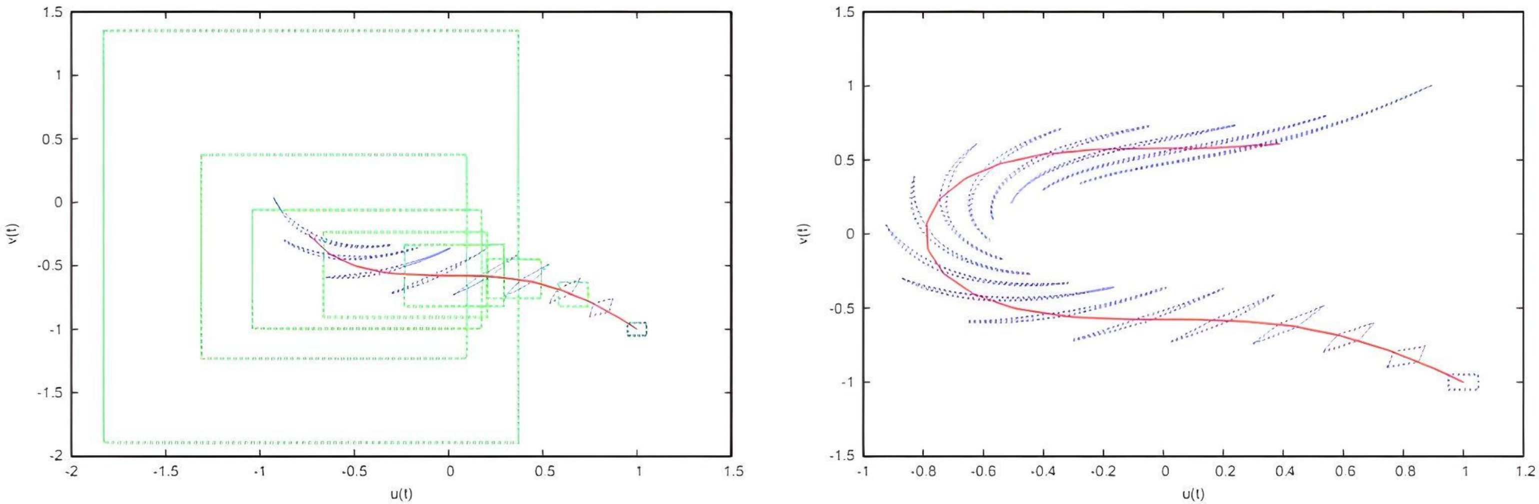
$$x + 0.75x^2 + [-0.25, 0.25]$$



Пример: квадратичный функциональный интервал-коридор

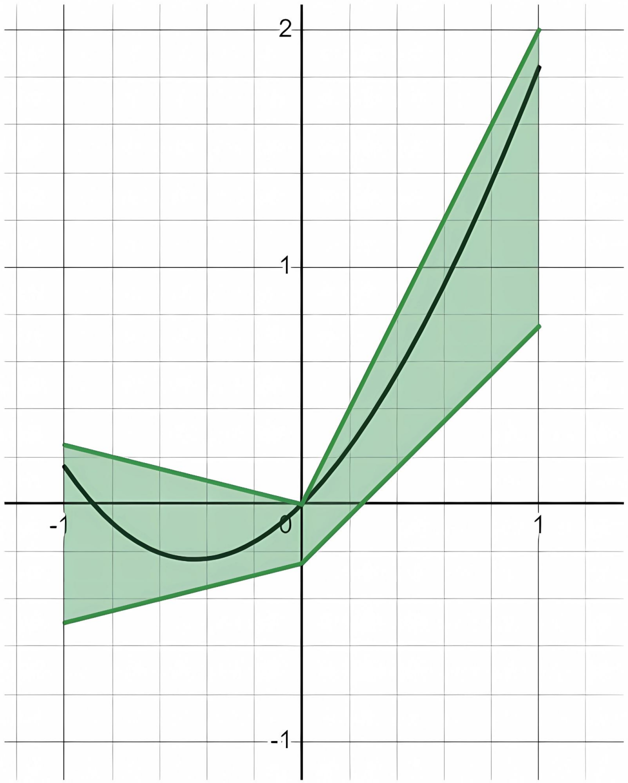
Такие квадратичные функциональные интервалы-коридоры используются при решении систем нелинейных ОДУ.

$$\begin{aligned} u' &= v, \quad u(0) = 1 + a, \\ v' &= u^2, \quad v(0) = -1 + b, \end{aligned}$$



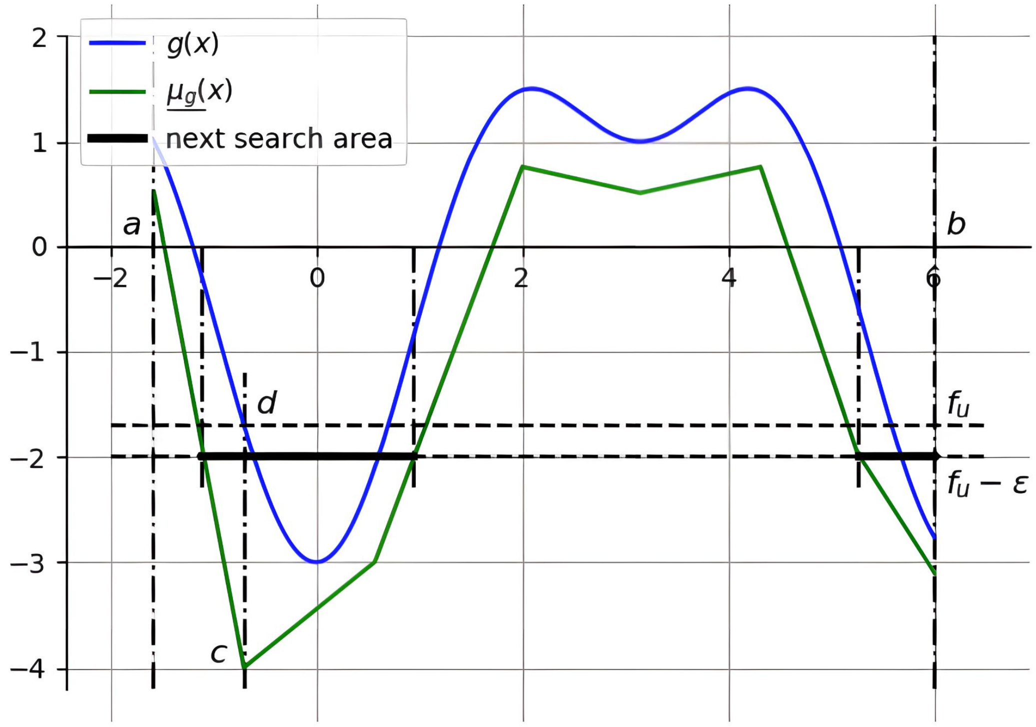
Пример: кусочно-линейный функциональный интервал

Кусочно-линейные функциональные интервалы используются для упрощения вида сложной одномерной функции — сначала априорно строятся приближения для элементарных функций и вводится правило, как обрабатывать композиции функций.



Пример: кусочно-линейный функциональный интервал

Полученное приближение функции используется в алгоритмах типа “ветвей-и-границ” для уменьшения области поиска глобального минимума. Построение приближения не требует дифференцируемости функции.



Пример: кусочно-линейный функциональный интервал

График из диссертации Югая С.А. 1988 года “Разработка и применение функционально-интервального анализа для обеспечения доказательности вычислений на ЭВМ”.

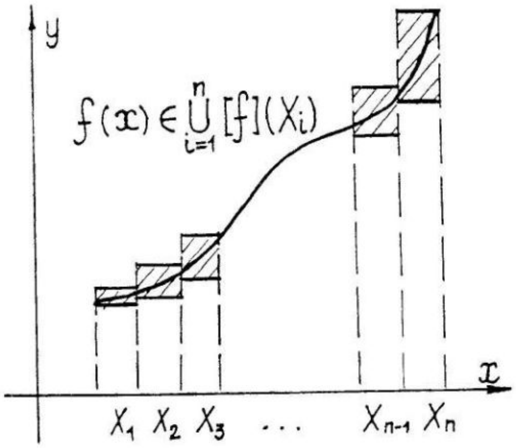


Рис. I.I

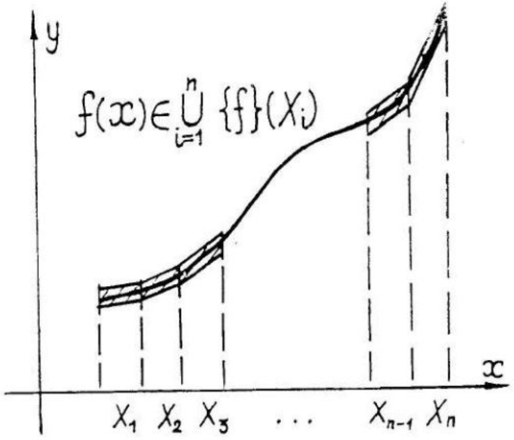
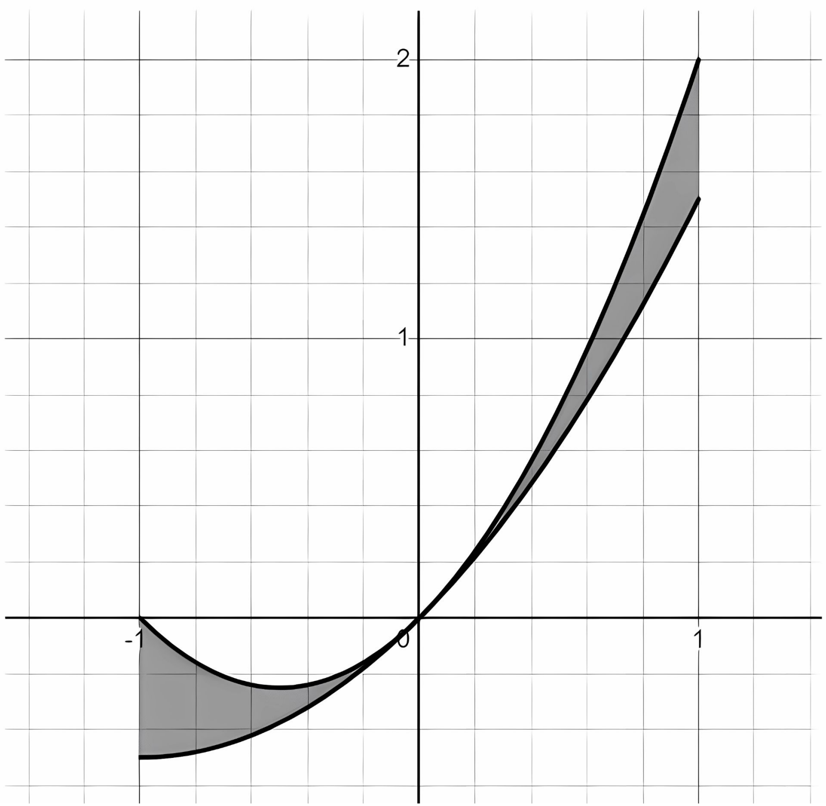


Рис. I.2

Пример: квадратичный функциональный интервал

Квадратичные функциональные интервалы для решения задач одномерной оптимизации строятся из разложения функции в ряд Тейлора — специальным образом обрабатывается остаточный член четного порядка.

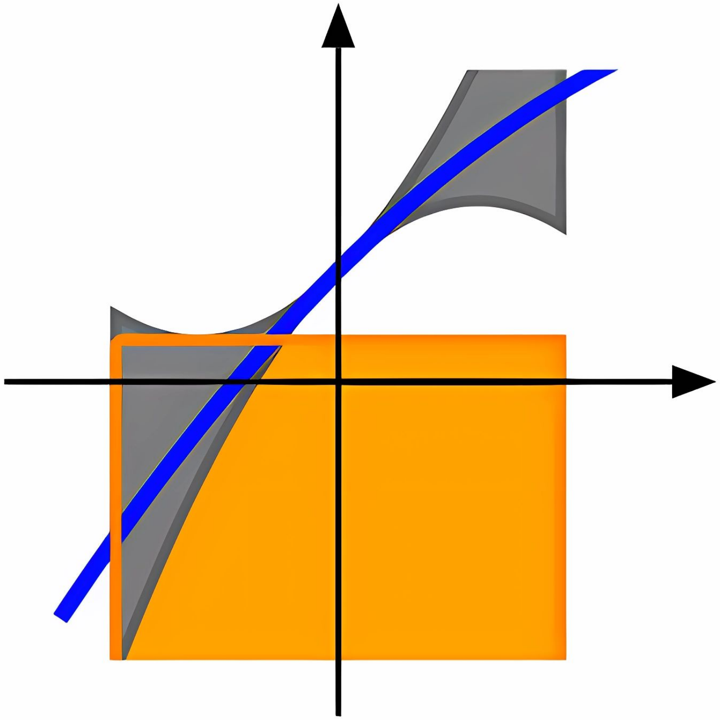
$$[x + 0.5x^2, x + x^2]$$



Пример: квадратичный функциональный интервал

Использование квадратичных функциональных интервалов накладывает ограничение на приближаемую функцию — она должна быть дважды непрерывно дифференцируема.

Взамен мы получаем как более быстрое уменьшение области поиска глобального минимума, так и оценку множества значения минимума с 3-им порядком точности.



Резюме

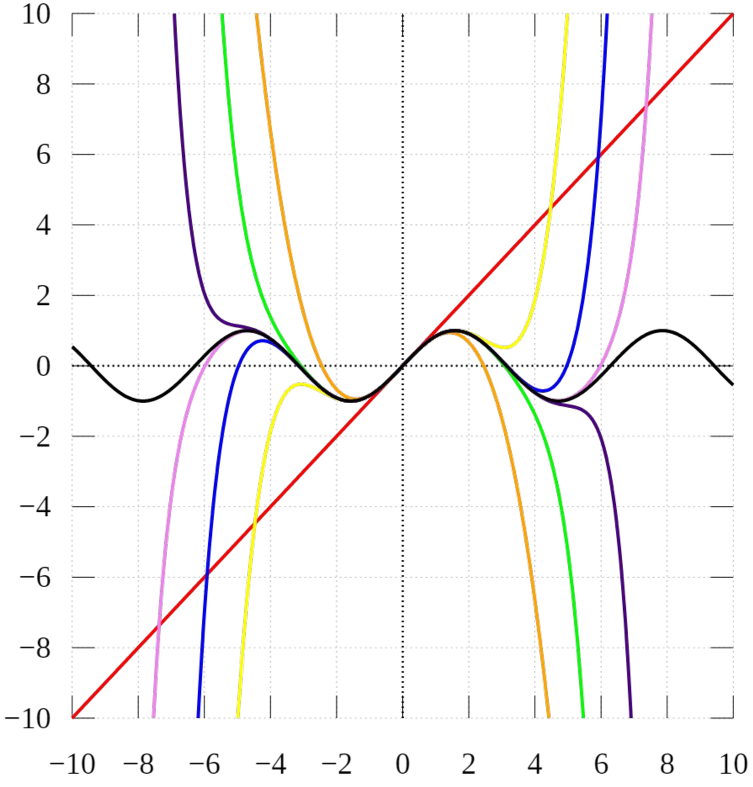
- идея функциональных интервалов — описание границ классического интервала через функции нескольких переменных
- реализация вычислительных операций, кроме тех, что определяются только через границы интервала, требуют уникальной реализации, которая зависит от вида функций границ
- по-прежнему возможна ситуация, когда классическая арифметика может давать результаты более точные, чем при использовании функциональных интервалов
- функциональные интервалы делают упор не на большом количестве рассматриваемых переменных, а на подробном рассмотрении малого количества переменных — например, одномерный полином высокого порядка
- хорошо подходят для итерационных задач малой размерности: динамические системы, методы ветвей-и-границ, …

Немного про Тейлоровские формы

Немного про Тейлоровские формы

Тейлоровские формы или Тейлоровские модели — зарубежное название для методов построения различных интервальных конструкций с использованием разложения Тейлора.

Одним из первых систематизировал и описал Martin Berz.



Про Тейлоровские формы

Многочлен Тейлора

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \ldots$$

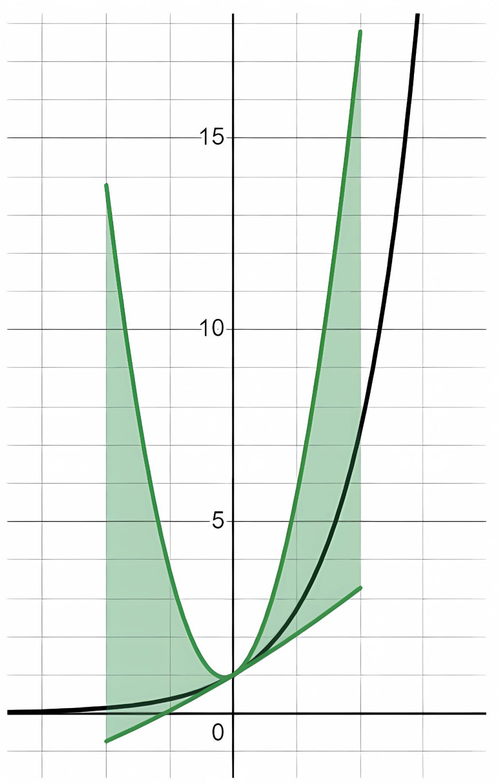
Остаточный член (в данном случае в форме Лагранжа):

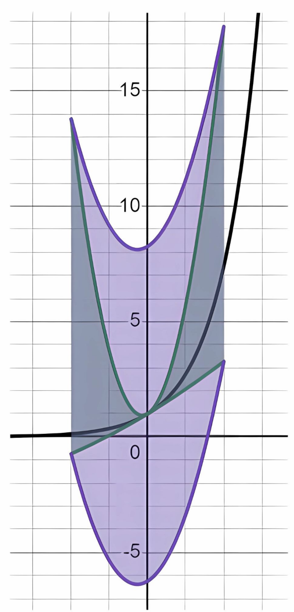
$$R_n(x) = \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}[a + \theta(x-a)] \qquad 0 < \theta < 1$$

Общая схема построения следующая: выбор многочлена Тейлора вплоть до какого-то члена и “интервализация” остаточного члена

Пример: связь с функциональными интервалами

Справа — функциональный интервал с использованием разложения Тейлора до второго порядка (приближение экспоненты на интервале [-2, 2])



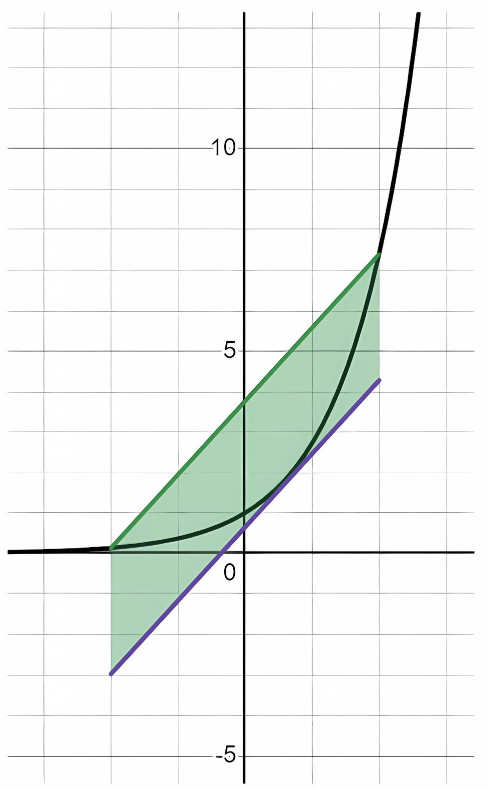


Слева показана разница в подходах между простым “обинтерваливанием” остаточного члена (фиолетовая область) и построением функционального интервала (серая область).

Пример: связь с аффинной арифметикой

В общем виде, определение аффинной формы представляет собой как раз вид разложения Тейлора 1-ого порядка функции многих переменных с остаточными членами, которые “интервализовали”.

На картинке — аффинное чебышевское приближение экспоненты на интервале [-2, 2].



Литература

- Югай С.А. Разработка и применение функционально-интервального анализа для обеспечения доказательности на ЭВМ. 1988.
- Hansen, E.R. (1975). A generalized interval arithmetic. In: Nickel, K. (eds) Interval Mathematics. IMath 1975. Lecture Notes in Computer Science, vol 29. Springer, Berlin, Heidelberg.
- J. L. D. Comba and J. Stolfi. Affine arithmetic and its applications to computer graphics. In Proceedings of VI SIBGRAPI (Brazilian Symposium on Computer Graphics and Image Processing), pages 9-18, 1993.
- Шарый С. П. Конечномерный интервальный анализ. — ФИЦ ИВТ: Новосибирск, 2022. Электронная книга, доступная на http://www.nsc.ru/interval/Library/InteBooks/SharyBook.pdf
- M. Berz. From Taylor series to Taylor models. 1997.

Ваши вопросы