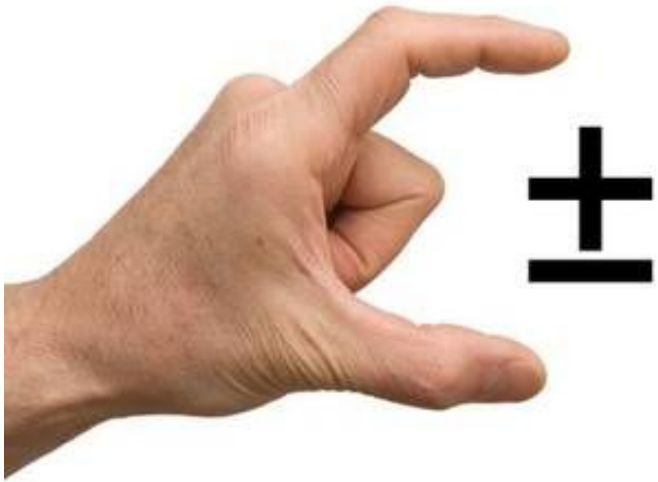
Про функциональные интервалы и их применение к оптимизации методом ветвей и границ

Скорик Дмитрий

Аспирант первого года обучения ФИЦ ИВТ
e-mail: dimakro2010@yandex.ru

Новосибирск, 27.02.2023

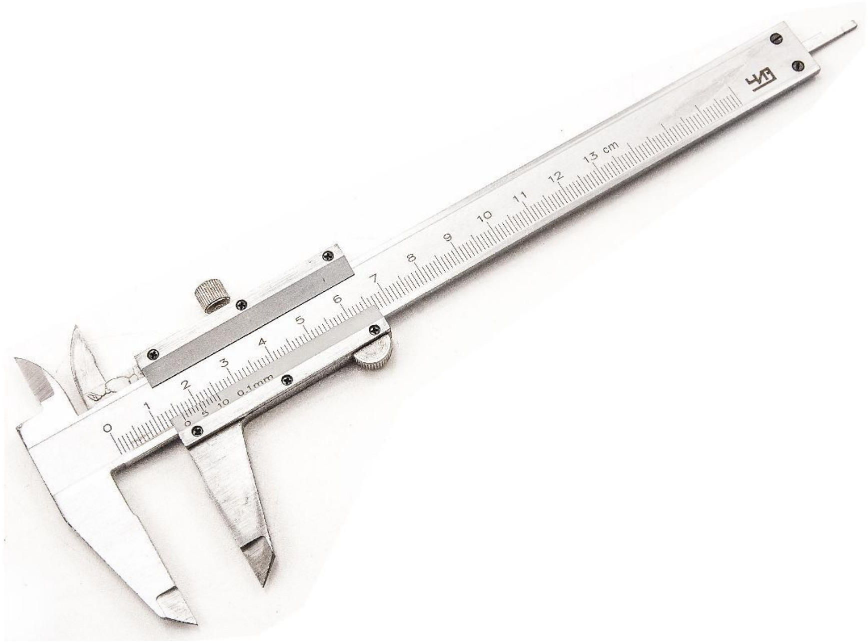
Актуальность использования интервалов



Актуальность использования интервалов

При решении практических задач часто возникают численные неопределенности, имеющие различную природу.

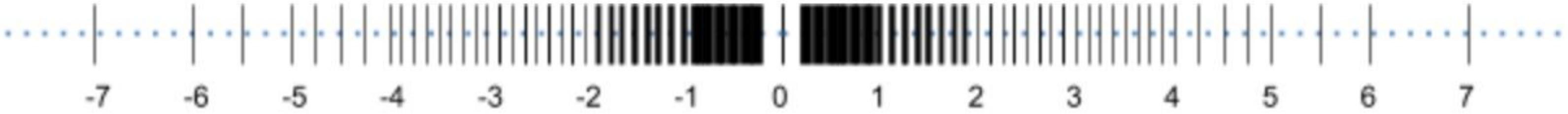
Одни, например, происходят из неточности измерительных приборов или человеческого фактора.



Актуальность использования интервалов

Другие — из-за использования на ЭВМ арифметики с ограниченной точностью
(или арифметики с плавающей точкой).

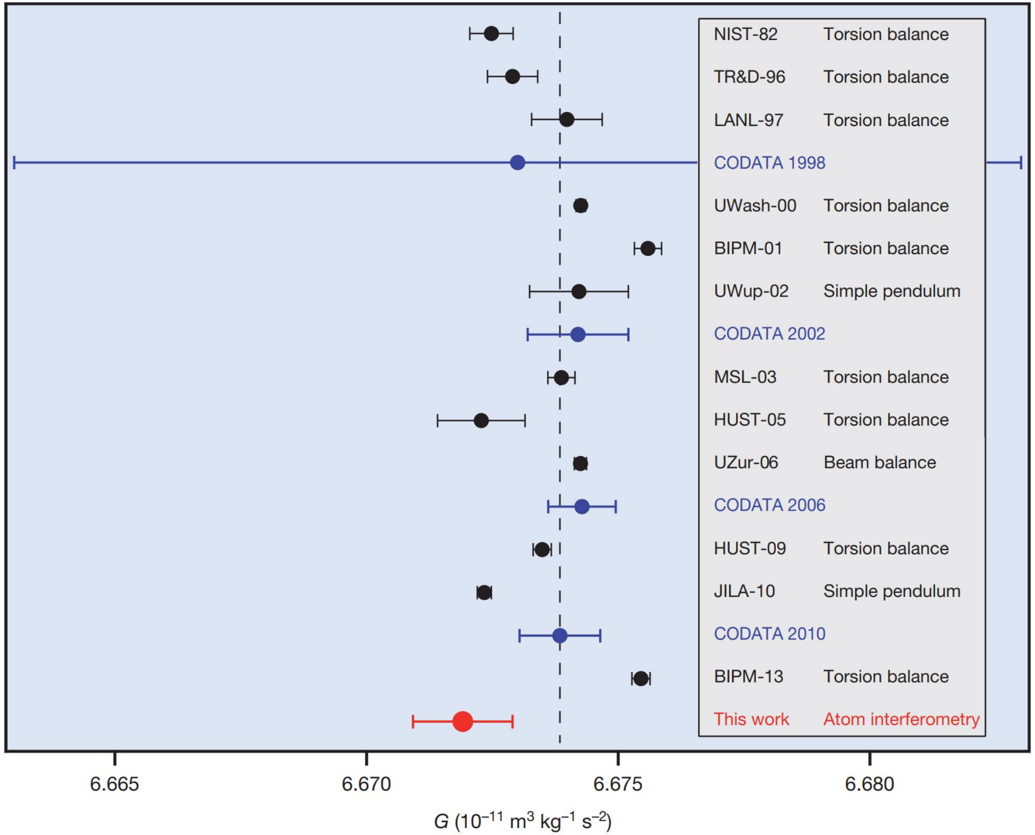
Из свежих новостей: на февральской конференции ISSCC 2023 глава AMD Лиза Су выделила важные аспекты для сохранения темпов наращивания производительности суперкомпьютеров на ближайшие 10 лет. Один из таких аспектов — <u>снижение точности вычислений</u>.



Актуальность использования интервалов

Также существуют неопределенности, которые мы не можем устранить.

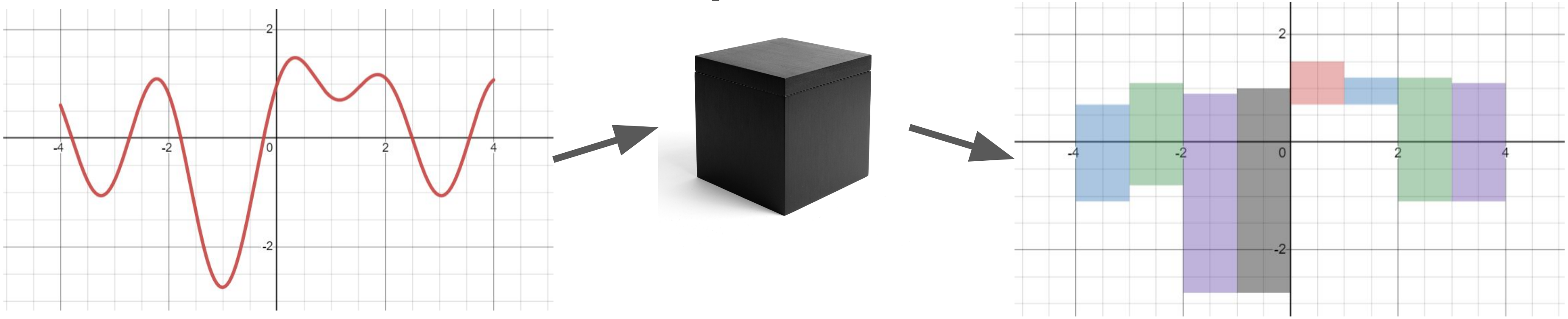
Примером может служить иллюстрация из журнала Nature к истории исследований измерения классической гравитационной константы G.



Актуальность использования интервалов

Еще одним примером неопределенности может служить концепция оракула (или черного ящика). В качестве примера приведен оракул, который умеет давать интервал значений функции по входному интервалу области определения.

истинное значение | черный ящик | наблюдатель (метод оптимизации)



Задача нелинейной оптимизации

Определение нелинейной оптимизации из книги Нестерова Юрия Евгеньевича “Методы выпуклой оптимизации”:

1.1.1. Общая формулировка задачи

Обозначим через x вещественный вектор размерности n:

$$x = \left(x^{(1)}, \ldots, x^{(n)}\right)^T \in \mathbb{R}^n,$$

а через S — некоторое множество из пространства $\mathbb{R}^n$. Пусть $f_0(x), \ldots, f_m(x)$ являются вещественнозначными функциями от x. В этой книге мы будем, как правило, рассматривать один из вариантов следующей общей задачи минимизации:

$$\begin{aligned} &\min f_0(x) \\ &\text{при } f_j(x) \,\&\, 0, \quad j = 1, \ldots, m, \\ &x \in S, \end{aligned} \tag{1.1}$$

где в качестве бинарного отношения & берется $\leqslant$, $\geqslant$ либо $=$.

Задача нелинейной оптимизации

Подавляющее большинство классических “точечных” алгоритмов оптимизации требуют выпуклости для нахождения минимума функций.

Интервальный анализ этого не требует, и позволяет обеспечить двусторонние оценки как на аргумент, доставляющий минимум, так и на значение минимума. Для этого чаще всего используют метод “ветвей и границ”.

Схема метода “ветвей и границ”

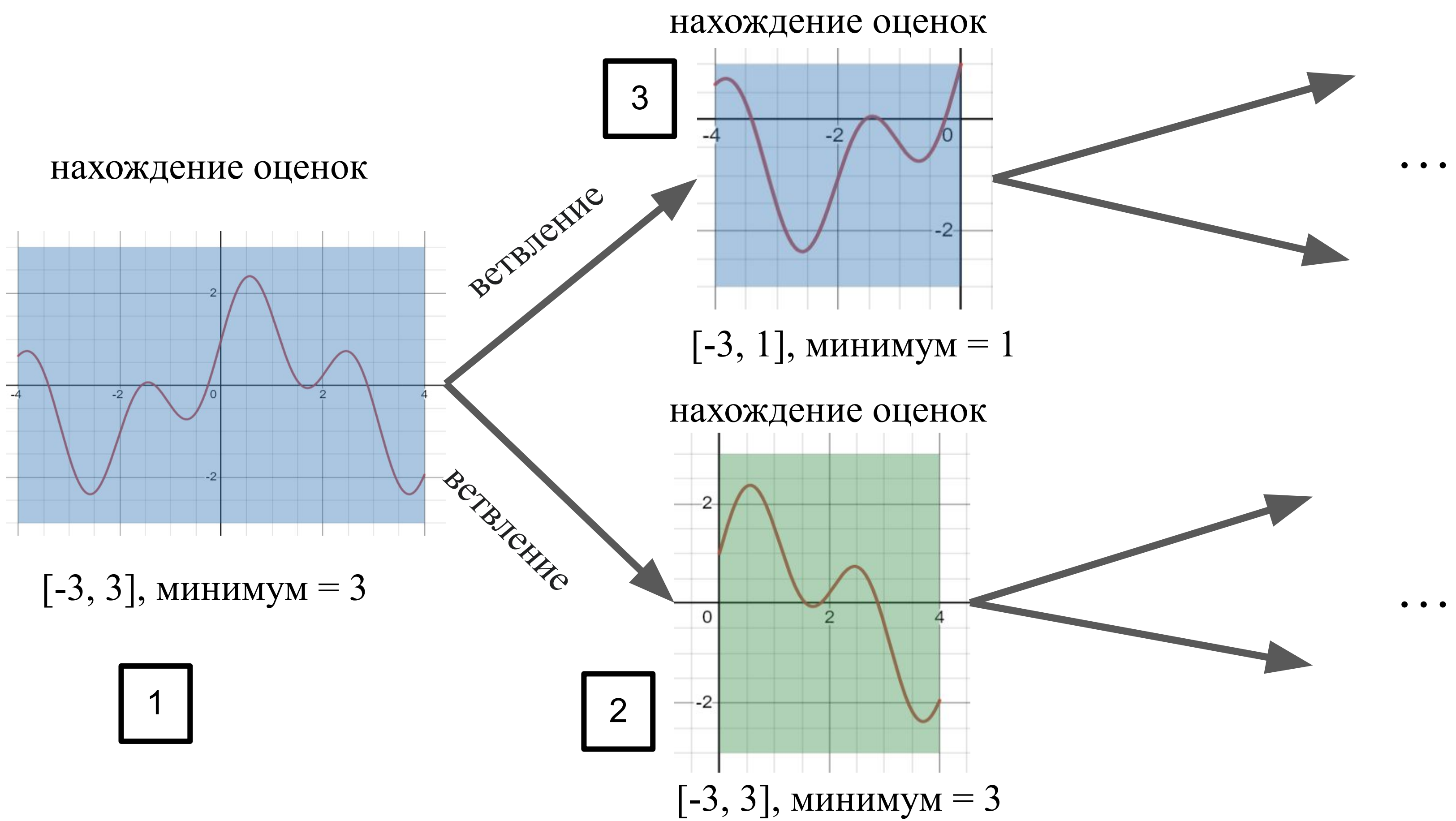
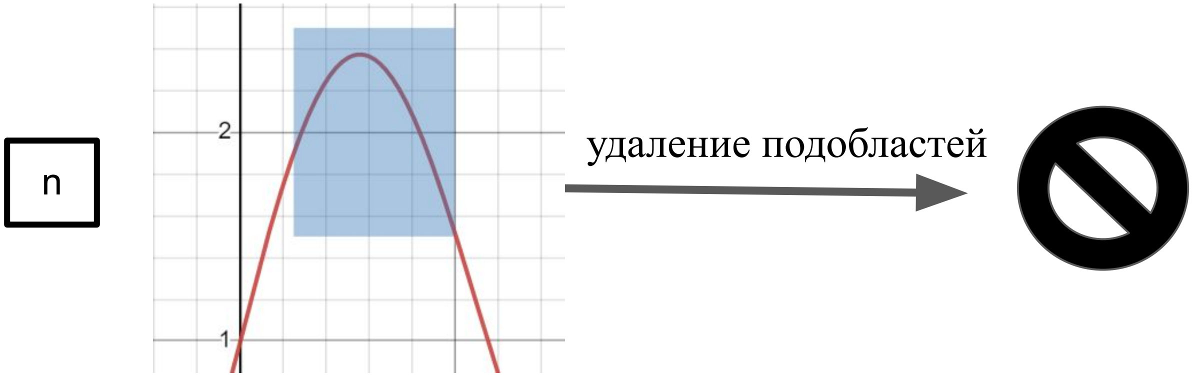


Схема метода “ветвей и границ”

нахождение оценок



[1.5, 2.5], минимум = 1

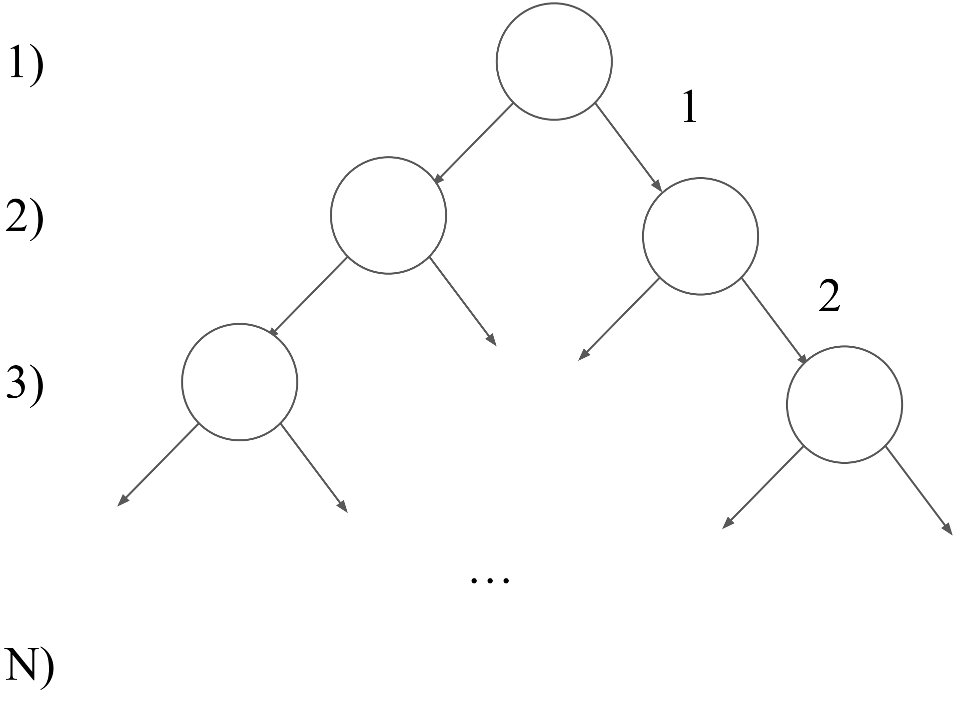
Схема метода “ветвей и границ”

Основные элементы метода “ветвей и границ”

- Ветвление
  - Дихотомия
- Нахождение оценок
  - Естественное интервальное расширение
  - Центрированные формы
  - …
- Удаление подобластей
  - Использование минимума уже пройденных областей
  - Проверка интервальных оценок производных первого или второго порядка

Схема метода “ветвей и границ”

Про ветвление и удаление подобластей



Количество дихотомий на слое *N*:

$$D = 2^{N-1}$$

Количество узлов на слое *N*:

$$D = 2^{N-1}$$

Общее количество дихотомий:

$$D_s = 2^{N-1} - 1$$

Про ветвление и удаление подобластей

Приведем конкретный пример.
Пусть изначально задан одномерный интервал поиска ширины 10.
$\varepsilon_x = 10^{-6}$

Тогда необходимо будет совершить 16777215 дихотомий и в рабочем списке будет содержаться 16777216 интервалов поиска.

Если перед совершением очередной дихотомии мы сможем уменьшать ширину интервала на 1%, то итого количество интервалов в рабочем списке сократится вдвое, а если на 4%, то вчетверо.

При раздувании размера списка становится проблемой потребление памяти, а также оперирование с таким списком. Помимо стандартной проверки с достигнутым уровнем минимума, часто проверяют первую производную на содержание ноля.

Про ветвление и удаление подобластей

Представим, что у нас в распоряжении метода имеется достигнутый уровень минимума:

$$M = 0.5$$

и центрированная форма вида:

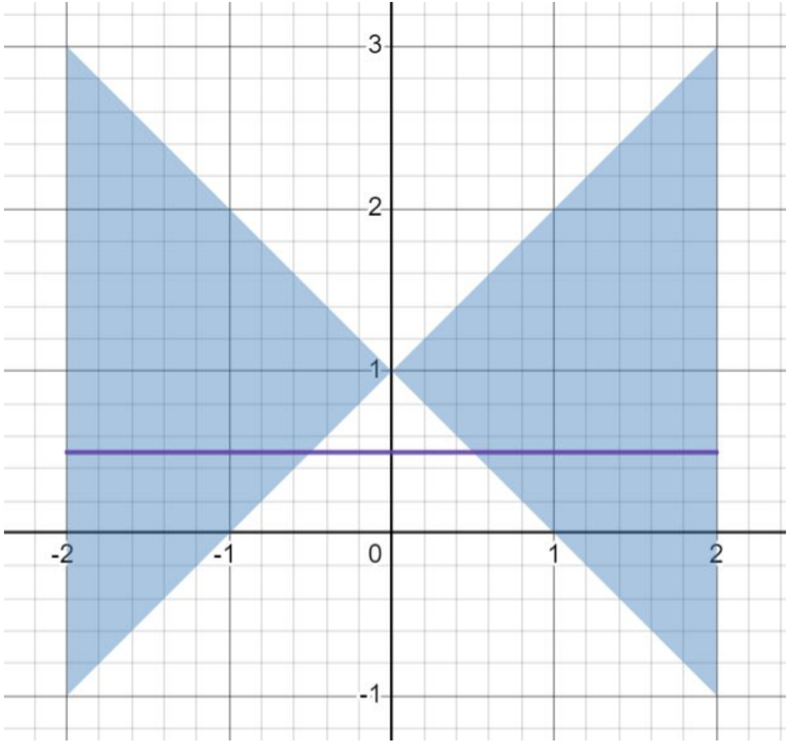
$$1 + [-1, 1]x \qquad\qquad x \in \boldsymbol{x} = [-2, 2]$$

Если мы получили центрированную форму по производной функции, тогда тест на монотонность пройден и мы утверждаем, что интервал нельзя отбросить, поскольку там может содержаться глобальный минимум функции.

Однако все же мы можем сделать некоторое улучшение в дереве поиска, рассмотрим подробнее.

Про ветвление и удаление подобластей

Наглядно представим множество, описываемое центрированной формой и также отметим достигнутый уровень минимума.



Видно, что можно удалить из рассмотрения интервал [-0.5, 0.5], тем самым сократив область поиска на 25%.

Также видно, что данный метод применим, только если достигнутый уровень минимума меньше, чем точечное значение в центрированной форме.

Кроме того, получаем "автоматическую" дихотомию.

Про ветвление и удаление подобластей (пример)

$$\sin(x) + \cos(x) + \sin(3x)\,, x \in [-2 \cdot 10^3, 2 \cdot 10^3]$$

Количество итераций без проверки первой производной: 60311

Количество итераций с проверкой первой производной на содержание нуля: 28280
Количество итераций с методом уменьшения области для центрированных форм: 27556

Количество итераций с проверкой первой производной на содержание нуля, указан начальный минимум “-2”: 28280
Количество итераций с методом уменьшения области для центрированных форм, указан начальный минимум “-2”: 24059

Про функциональные интервалы

Идея использования геометрических свойств интервалов подтолкнула меня к созданию понятия функционального интервала, как обобщения классического интервала.

Определение 1.1. *Функциональным интервалом* назовём интервал, нижняя и верхняя границы которого представляются некоторыми функциями

$$\mathcal{L} : [-1,\, 1]^n \to \mathbb{R} \qquad \text{и} \qquad \mathcal{U} : [-1,\, 1]^n \to \mathbb{R}$$

соответственно, которые удовлетворяют свойству

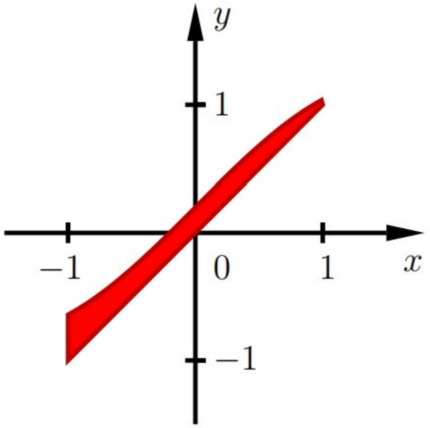
$$\forall x_i \in [-1,\, 1], \qquad i = 1, \ldots, n, \qquad \mathcal{L}(x_1, \ldots, x_n) \le \mathcal{U}(x_1, \ldots, x_n).$$

Функции $\mathcal{L}$ и $\mathcal{U}$ будем называть нижней и верхней *граничной функцией* соответственно. Область значений аргументов этих функций x_i условна. Можно рассматривать любые интервалы изменения значений этих переменных. При этом любой интервал изменения переменной можно линейным преобразованием перевести в интервал $[-1,\, 1]$.

Про функциональные интервалы

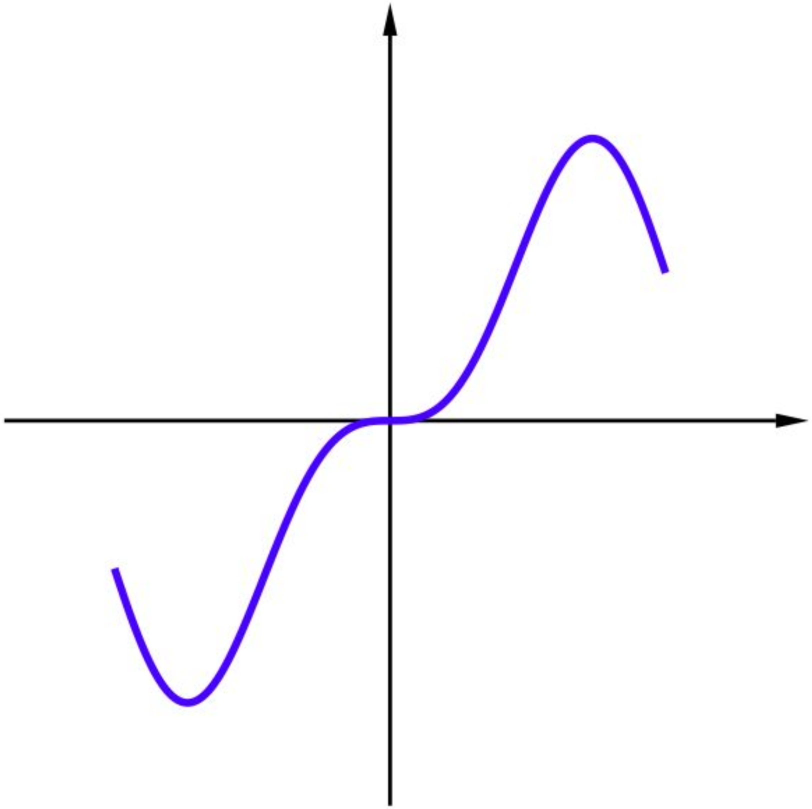
Семейство функциональных интервалов без аргументов будем обозначать, как $\mathbb{FR}_0$. Удобно мыслить это семейство функциональных интервалов, как параметрическое семейство классических интервалов $\left(\mathbb{IR} \equiv \mathbb{FR}_0\right)$.

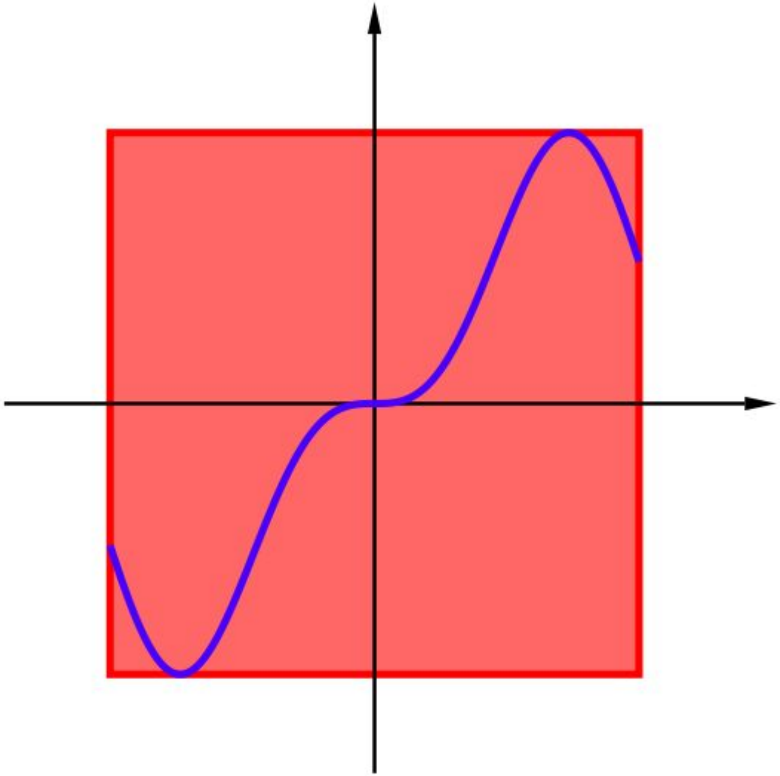
Для записи функциональных интервалов предлагается следующая нотация: вместо числовых границ интервалов будем писать на месте левого конца интервала функцию $\mathfrak{L}$, а вместо правого конца — функцию $\mathfrak{U}$. Например, функциональный интервал $\boldsymbol{x}$, у которого $\mathfrak{L}(x) = x$, а $\mathfrak{U}(x) = \sin(x) + 0.2$, будет записываться как $\boldsymbol{x} = \left[x, \sin(x) + 0.2\right]$.



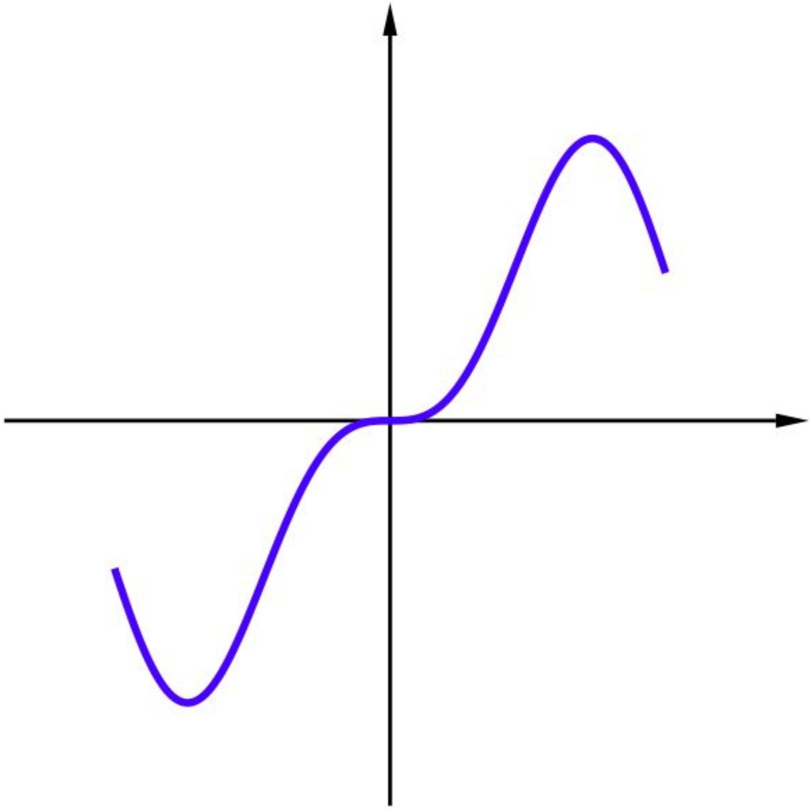
Пример функционального интервала $\boldsymbol{x} = \left[x, \sin(x) + 0.2\right]$.

Про функциональные интервалы





Про функциональные интервалы



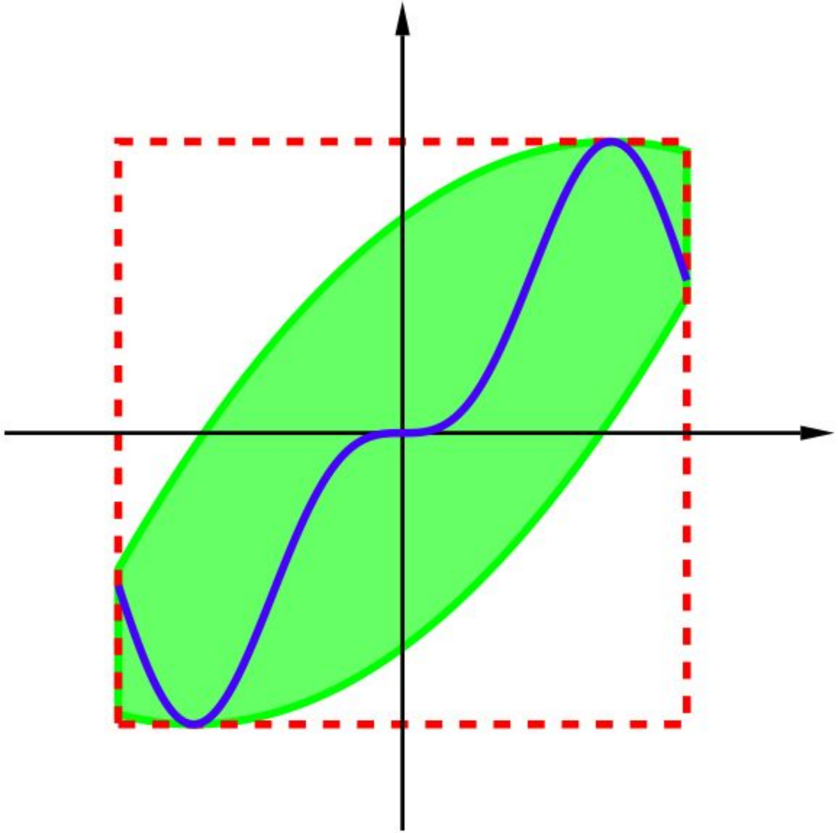


Схема метода “ветвей и границ”

Основные элементы метода “ветвей и границ”

Ветвление

Дихотомия

Нахождение оценок

Естественное интервальное расширение

Центрированные формы

…

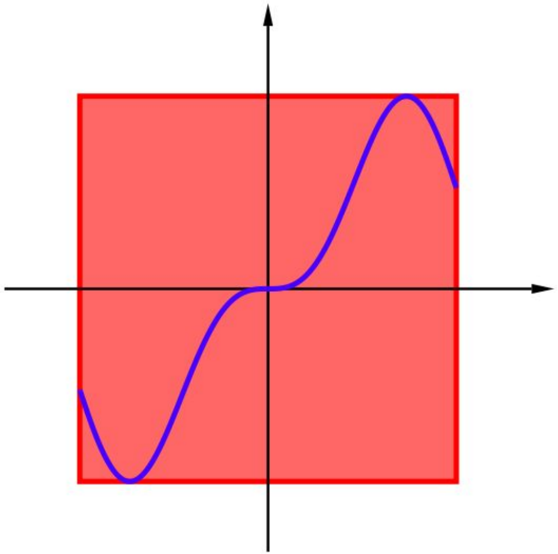
Удаление подобластей

Использование минимума уже пройденных областей

Проверка интервальных оценок производных первого или второго порядка

Про нахождение интервальных оценок

На практике чаще всего используют центрированные формы оценивания в силу их простоты и эффективности (2-й порядок оценивания). Оценка минимума на интервале совпадает с оценкой значений функции на интервале с учетом значения в центре.



При этом построение методов для получения интервальной оценки порядка выше 2-ого значений функций на интервале достаточно сложно, и возможно, излишне.

Идея — оценивать с высоким порядком только минимум.

Квадратичный функциональный интервал

Пусть N — четное число, а f — достаточно гладкая функция, тогда по частичному разложению в ряд Тейлора можно построить функциональный интервал следующего вида:

$$f(x) \in \sum_{k=0}^{N-1} \frac{f^{(k)}(\operatorname{mid} \boldsymbol{x})}{k!} \cdot (x - \operatorname{mid} \boldsymbol{x})^k + \frac{\boldsymbol{f}_{\natural}^{(N)}(\boldsymbol{x})}{N!} \cdot (x - \operatorname{mid} \boldsymbol{x})^N$$

Обозначим

$$\underline{f(x)} = \sum_{k=0}^{N-1} \frac{f^{(k)}(\operatorname{mid} \boldsymbol{x})}{k!} \cdot (x - \operatorname{mid} \boldsymbol{x})^k + \frac{\underline{\boldsymbol{f}_{\natural}^{(N)}(\boldsymbol{x})}}{N!} \cdot (x - \operatorname{mid} \boldsymbol{x})^N,$$

$$\overline{f(x)} = \sum_{k=0}^{N-1} \frac{f^{(k)}(\operatorname{mid} \boldsymbol{x})}{k!} \cdot (x - \operatorname{mid} \boldsymbol{x})^k + \frac{\overline{\boldsymbol{f}_{\natural}^{(N)}(\boldsymbol{x})}}{N!} \cdot (x - \operatorname{mid} \boldsymbol{x})^N.$$

Квадратичный функциональный интервал

Обозначим интервал $\left[\,\underline{f(x)},\, \overline{f(x)}\,\right]$ через $\boldsymbol{F}(x, \boldsymbol{x})$. Так как $f(x) \in \boldsymbol{F}(x, \boldsymbol{x})$, то при этом

$$\operatorname{wid} \boldsymbol{F}(x, \boldsymbol{x}) = \frac{1}{N!} \cdot \operatorname{wid} \boldsymbol{f}_{\natural}^{(N)}(\boldsymbol{x}) \cdot (x - \operatorname{mid} \boldsymbol{x})^N = O\big((\operatorname{wid} \boldsymbol{x})^{N+1}\big) \quad \text{при } \operatorname{wid} \boldsymbol{x} \to 0.$$

Таким образом, для любого x из $\boldsymbol{x}$ верно

$$\operatorname{wid} \boldsymbol{F}(x, \boldsymbol{x}) = O\big((\operatorname{wid} \boldsymbol{x})^{N+1}\big) \qquad \text{при} \quad \operatorname{wid} \boldsymbol{x} \to 0. \tag{1}$$

Квадратичный функциональный интервал

Пусть минимум функции $f(x)$ на интервале $\boldsymbol{x}$ достигается в точке x^*. Очевидно, что

$$f(x^*) \in \Big[\, \min_{\boldsymbol{x}} \underline{\boldsymbol{F}(x, \boldsymbol{x})},\ \min_{\boldsymbol{x}} \overline{\boldsymbol{F}(x, \boldsymbol{x})} \,\Big].$$

Предположим также, что минимум нижнего конца интервальной оценки $\boldsymbol{F}(x, \boldsymbol{x})$, т. е. функции $\underline{\boldsymbol{F}(x, \boldsymbol{x})}$ достигается в $\breve{x}$. Тогда имеет место

$$\Big[\, \min_{\boldsymbol{x}} \underline{\boldsymbol{F}(x, \boldsymbol{x})},\ \min_{\boldsymbol{x}} \overline{\boldsymbol{F}(x, \boldsymbol{x})} \,\Big] \subseteq \Big[\, \underline{\boldsymbol{F}(\breve{x}, \boldsymbol{x})},\ \overline{\boldsymbol{F}(\breve{x}, \boldsymbol{x})} \,\Big] = \boldsymbol{F}(\breve{x},\, \boldsymbol{x}).$$

А поскольку для любой точки $x \in \boldsymbol{x}$ верна асимптотическая оценка (1), то она верна и для точки $\breve{x}$. Значит, $\operatorname{wid} \boldsymbol{F}(\breve{x}, \boldsymbol{x}) = O\big((\operatorname{wid} \boldsymbol{x})^{N+1}\big)$ при $\operatorname{wid} \boldsymbol{x} \to 0$.

Таким образом,

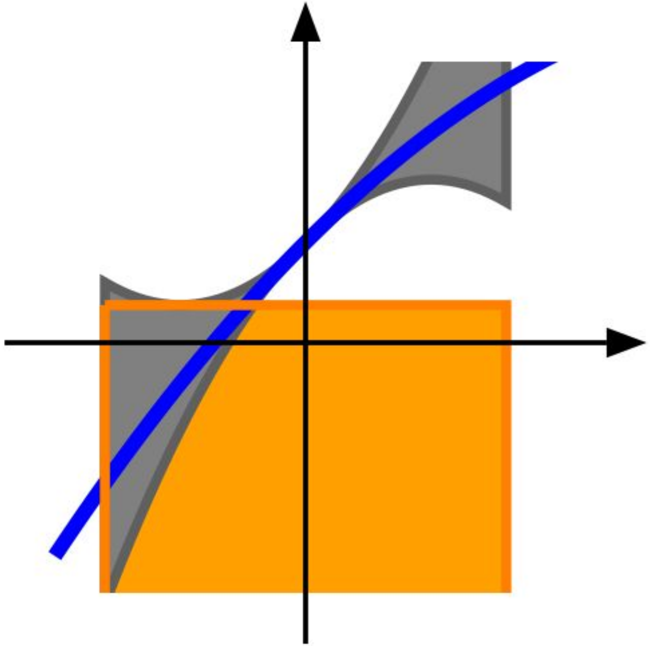
$$\operatorname{wid} \Big[\, \min_{\boldsymbol{x}} \underline{\boldsymbol{F}(x, \boldsymbol{x})},\ \min_{\boldsymbol{x}} \overline{\boldsymbol{F}(x, \boldsymbol{x})} \,\Big] = O\big((\operatorname{wid} \boldsymbol{x})^{N+1}\big) \text{ при } \operatorname{wid} \boldsymbol{x} \to 0.$$

Значит, с помощью интервала такого вида можем оценивать минимум на интервале с теоретически любым нечетным порядком. Далее будем рассматривать квадратичный интервал такого вида (оценка минимума 3-го порядка).

Квадратичный функциональный интервал

Ранее рассматривали, что метод уменьшения области для центрированных форм работает только для случая, когда достигнутый уровень минимума меньше, чем значение функции в центре интервала.

Для функционального квадратичного интервала было получено, что даже при отсутствии этого условия сжатие области поиска будет происходить. Единственное требование для этого - линейный член квадратичного функционального интервала должен быть ненулевым.



Про ветвление и удаление подобластей (пример)

$$\sin(x) + \cos(x) + \sin(3x), x \in [-2 \cdot 10^3, 2 \cdot 10^3]$$

Количество итераций с проверкой первой производной на содержание нуля: 28280 (0.5с)
Количество итераций с методом уменьшения области для центрированных форм: 27556 (0.5с)
Количество итераций с уменьшения области для квадратичного функционального интервала: 5442 (0.3с)

Количество итераций с проверкой первой производной на содержание нуля, указан начальный минимум “-2”: 28280 (0.5с)
Количество итераций с методом уменьшения области для центрированных форм, указан начальный минимум “-2”: 24059 (0.5с)
Количество итераций с уменьшения области для квадратичного функционального интервала, указан начальный минимум “-2”: 5126 (0.3с)

Перспективы и дальнейшие исследования

Квадратичный функциональный интервал имеет следующие преимущества по сравнению с центрированными формами:

1) третий порядок оценки минимума
2) для уменьшения области поиска минимума не обязательно отслеживать положение достигнутого уровня минимума относительно значения в центре
3) можно внедрить “бесплатный” тест на знак второй производной, чтобы можно было, например, подключать быстрые “точечные” методы для локальной оптимизации

Сейчас в разработке находится статья, в которой будут подробно разобраны алгоритмы для реализации описанных методов уменьшения области поиска, а также проведены сравнения различных методов поиска глобального минимума функции для тестовых функций.

Наиболее остро вопрос дробления стоит в многомерном случае, поэтому в настоящее время ведется исследование по разработке похожих методов для оптимизации функций нескольких переменных.

Ваши вопросы