Лекция 6. Аппроксимация и интерполяция.

Вычислительная математика

Весенний семестр 2022

Ольга Вячеславовна Перл

Основные определения

Аппроксимация

↓

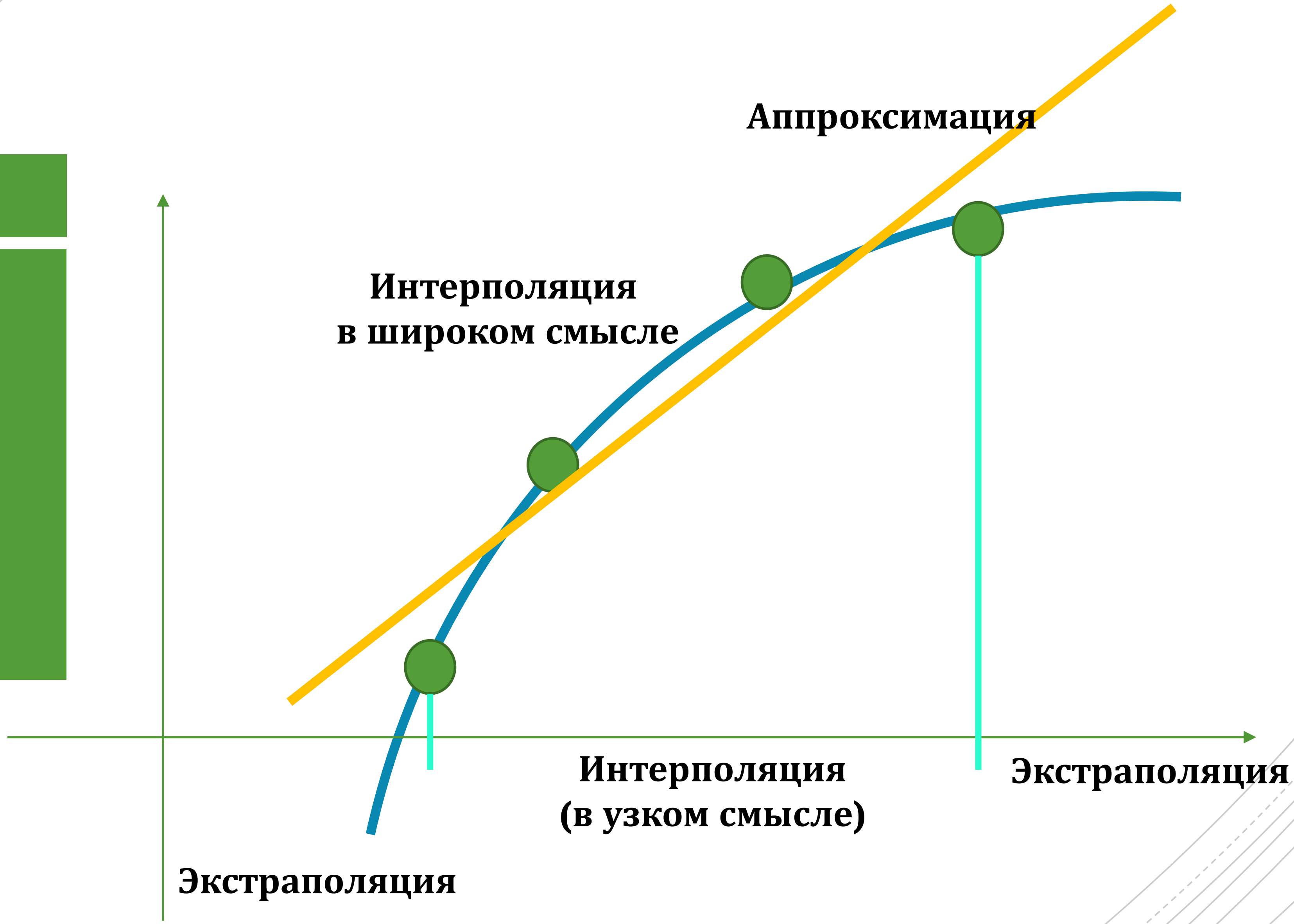
интерполяция (в широком смысле)

↙ ↘

Интерполяция (в узком смысле)

Экстраполяция

Определение на графике



Цель аппроксимации

- Имеется набор данных (или сложная функция, для которой сложно считать значения и которую мы хотим упростить).
- Необходимо получить функцию, которая бы описывала поведение изменения значений в наборе данных (похожая на исходную, но неизвестная функция).
- С помощью новой вычисленной (приближённой) функции становится возможным рассчитать новые значения, которые были пропущены в исходном наборе данных.

Интерполяция

- Исходный набор точек называется **узлами интерполяции** / interpolation points, mesh points.

Полином Лагранжа

Полином Лагранжа

Это линейная функция и: $L_1(x_0) = f_0, L_1(x_1) = f_1$

- $L(x_i) = f_i \;\; i = 0, 1, \dots, n$ (1)
- Под степенью полинома мы согласны подразумевать многочлен степени не выше **n**.
- **Теорема:** Существует только 1 интерполяционный полином степени **n** по условию.
- **Доказательство:**

1. Пусть **n=1**, тогда: $L_1(x) = \frac{x - x_1}{x_0 - x_1} f_0 + \frac{x - x_0}{x_1 - x_0} f_1$ (2)
2. Пусть **n=2**, тогда: $L_2(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} f_0 + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} f_1 + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} f_2$
3. И тогда получим: $L_n(x) = \sum_{i=0}^{n} p_{ni}(x) f_i$, где:

Lagrange`s polynomial

$$p_{ni}(x) = \frac{(x - x_0)\dots(x - x_{i-1})(x - x_{i+1})\dots(x - x_n)}{(x_i - x_0)\dots(x_i - x_{i-1})(x_i - x_{i+1})\dots(x_i - x_n)}$$

$$i = 0, 1, \dots, n.$$

Lagrange coefficients

Полином Лагранжа

Построение полинома Лагранжа по данным:

i	0	1	2	3
x_i	0	2	3	5
f_i	1	3	2	5

При **n=3** получим:

$$L_3(x) = \frac{(x-2)(x-3)(x-5)}{(0-2)(0-3)(0-5)} * 1 + \frac{x(x-3)(x-5)}{2(2-3)(2-5)} * 3 + \frac{x(x-2)(x-5)}{3(3-2)(3-2)} * 2 + \frac{x(x-2)(x-3)}{5(5-2)(5-3)} * 5 = 1 + \frac{62}{15}x - \frac{13}{6}x^2 + \frac{3}{10}x^3$$

Полином Лагранжа

PROS

- Может быть легко (без построения нового полинома) использован, когда значения функции меняются без изменения аргументов.
- Требуется меньшее число точек для аппроксимации таких функций как парабола.

CONS

- Требуется строить полином заново при добавлении новых точек.
- Сложность значительно возрастает при возрастании количества точек.
- Шум существенно влияет на весь полином.

Полином Ньютона

Полином Ньютона

- Конечные разности порядка **n** (для равноотстоящих узлов)
- Пусть $y = f(x)$ и $\Delta x = h$ приращение аргумента, тогда:

$$\Delta y = \Delta f(x) = f(x + \Delta x) - f(x)$$

Называется конечной разностью функции **f.**

- Для высших порядков:

$$\Delta^n y = \Delta(\Delta^{n-1} y)$$

- Например:

$$\begin{aligned} \Delta^2 y &= \Delta[f(x + \Delta x) - f(x)] \\ &= [f(x + 2\Delta x) - f(x + \Delta x)] - [f(x + \Delta x) - f(x)] \\ &= f(x + 2x) - 2f(x + \Delta x) + f(x) \end{aligned}$$

Пример

- Построить конечную разность для функции $P(x) = x^3$ если $\Delta x = h = 1$

$$\Delta P(x) = (x+1)^3 - x^3 = 3x^2 + 3x + 1,$$
$$\Delta^2 P(x) = [3(x+1)^2 + 3(x+1) + 1] - (3x^2 + 3x + 1) = 6x + 6$$
$$\Delta^2 P(x) = [6(x+1) + 6] - (6x + 6) = 6$$
$$\Delta^n P(x) = 0, if\ n > 3$$

Конечная разность функции

- Δ - это оператор: $y = f(x) \to \Delta y = f(x + \Delta x) - f(x)$ (где $\Delta x = const$)
- Свойства:
  1. $\Delta(u + v) = \Delta u + \Delta v$;
  2. $\Delta(Cu) = C\Delta u$, where $\mathrm{C} = const$;
  3. $\Delta^m(\Delta^n) = \Delta^{m+n}y$, where $\mathrm{m}, \mathrm{n} \in \mathbb{Z}, \Delta^0 y = y$
- Когда $\Delta x \to 0$ верно:

$$f^{(n)}(x) \approx \frac{\Delta^n f(x)}{(\Delta x)^n}.$$

Полином Ньютона

$$P_n(x) = y_0 + q\Delta y_0 + \frac{q(q^2-1)}{2!}\Delta^2 y_0 + \cdots + \frac{q(q-1)\dots(q-n+1)}{n!}\Delta^n y_0$$

где $q = \frac{x-x_0}{h}$ - это число шагов, которые требуются, чтобы достичь x из x_0

- Если $n = 1$, тогда формула линейной интерполяции:

$$P_1(x) = y_0 + q\Delta y_0$$

- Если $n = 2$, тогда получаем параболическую или квадратичную интерполяцию:

$$P_2(x) = y_0 + q\Delta y_0 + \frac{q(q-1)}{2}\Delta^2 y_0$$

Полином Ньютона для не равноотстоящих узлов

- $[x_i, x_{i+1}] = \frac{y_{i+1}-y_i}{x_{i+1}-x_i}$ **- это разделённая разница.**
- **Тогда интерполяционая формула:**
  $P(x)$
  $= y_0 + [x_0, x_1](x - x_0) + [x_0, x_1, x_2](x - x_0)(x - x_1) + \cdots$
  $+ [x_0, x_1, \ldots, x_n](x - x_0)(x - x_1) \ldots (x - x_{n-1})$

Полином Ньютона

PROS

- Легче добавлять новые точки без перестроения всего полинома.
- Требуется меньшее число точек для приближения таких функций как парабола.

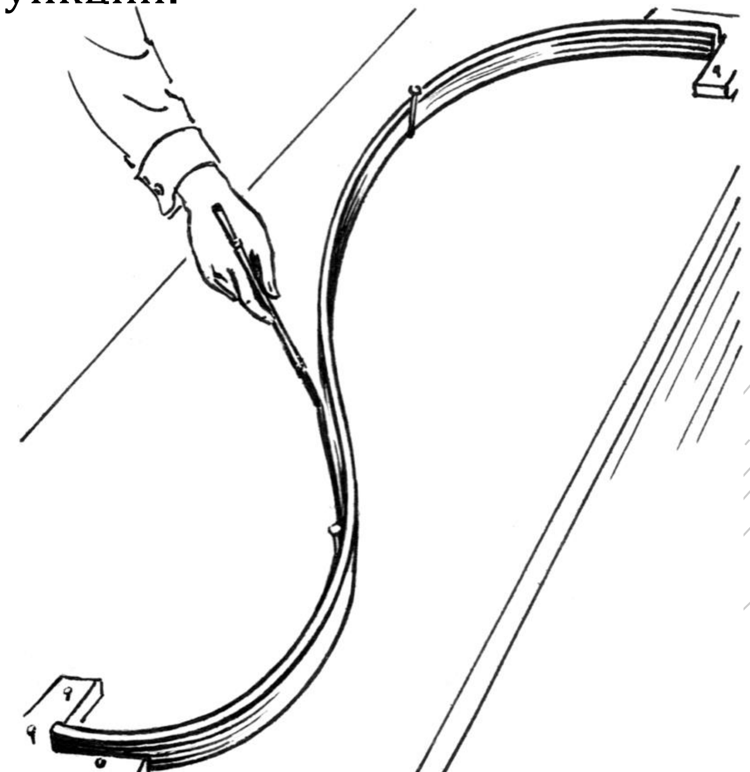
CONS

- Шум существенно влияет на весь полином.

Сплайны

Сплайны

- Термин "сплайн" используется для обозначения широкого класса функций, которые используются в приложениях, требующих интерполяции данных и/или сглаживания. Данные могут быть как одномерными, так и многомерными.
- Сплайн представлял собой гибкую металлическую линейку — универсальный шаблон, который использовался чертежниками для соединения точек на чертеже гладкой кривой, то есть для графического выполнения интерполяции.
- Это общее название для семейства функций.



Сплайны

- Сплайн - это функция, которая вместе с несколькими производными непрерывна на интервале для всего заданного сегмента [a, b] и на каждом частичном сегменте $[x_i, x_{i+1}]$ отдельно - некоторый алгебраический многочлен.
- Максимальная производная по всем частичным сегментам называется степенью сплайна, а разница между степенью сплайна и порядком наибольшей непрерывной на отрезке [a,b] называется дефектом сплайна.
- Например, непрерывная кусочно-заданная функция (ломаная линия) представляет собой сплайн первой степени с дефектом 1, поскольку непрерывна только сама функция, но не ее производная (точки экстремума).

Сплайны

- Наиболее часто используется кубический сплайн.
- На каждом частичном сегменте:

$$S_3(x) = \frac{(x_{i+1}-x)^2(2(x-x_i)+h)}{h^3} f_i + \frac{(x-x_i)^2(2(x_{i+1}-x)+h)}{h^3} f_{i+1} + \frac{(x_{i+1}-x)^2(x-x_i)}{h^2} m_i + \frac{(x-x_i)^2(x_{i+1}-x)}{h^2} m_{i+1}.$$

Сплайны

PROS

- Намного проще при работе с большим числом точек.
- Лучше обрабатывает шум во входных данных.

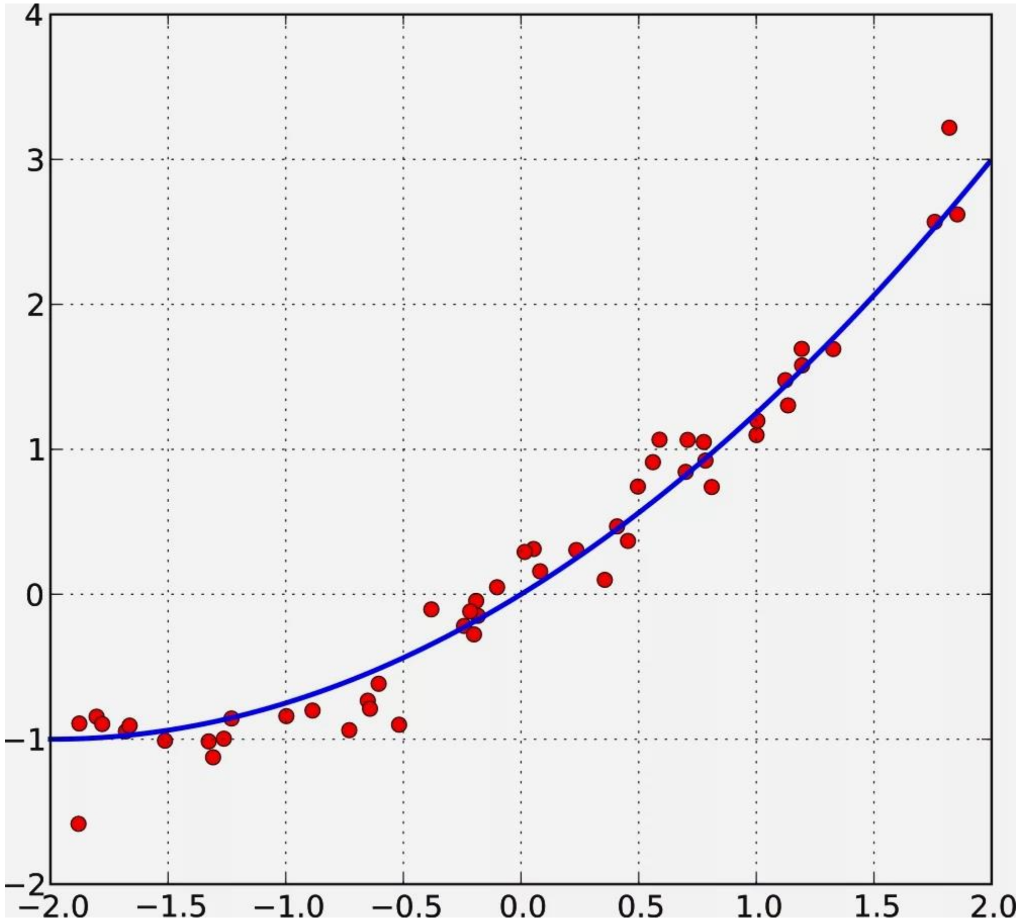
CONS

- Может быть сложно аппроксимировать простые функции (как парабола) в сравнении с полиномиальной интерполяцией.
- Необходимо выбрать правильный метод из семейства для обработки сегментов на границах интерполяции (самые правые и самые левые сегменты).

LS - Least Squares

Метод наименьших квадратов

Пример: LS - Least Squares



LS - Least Squares

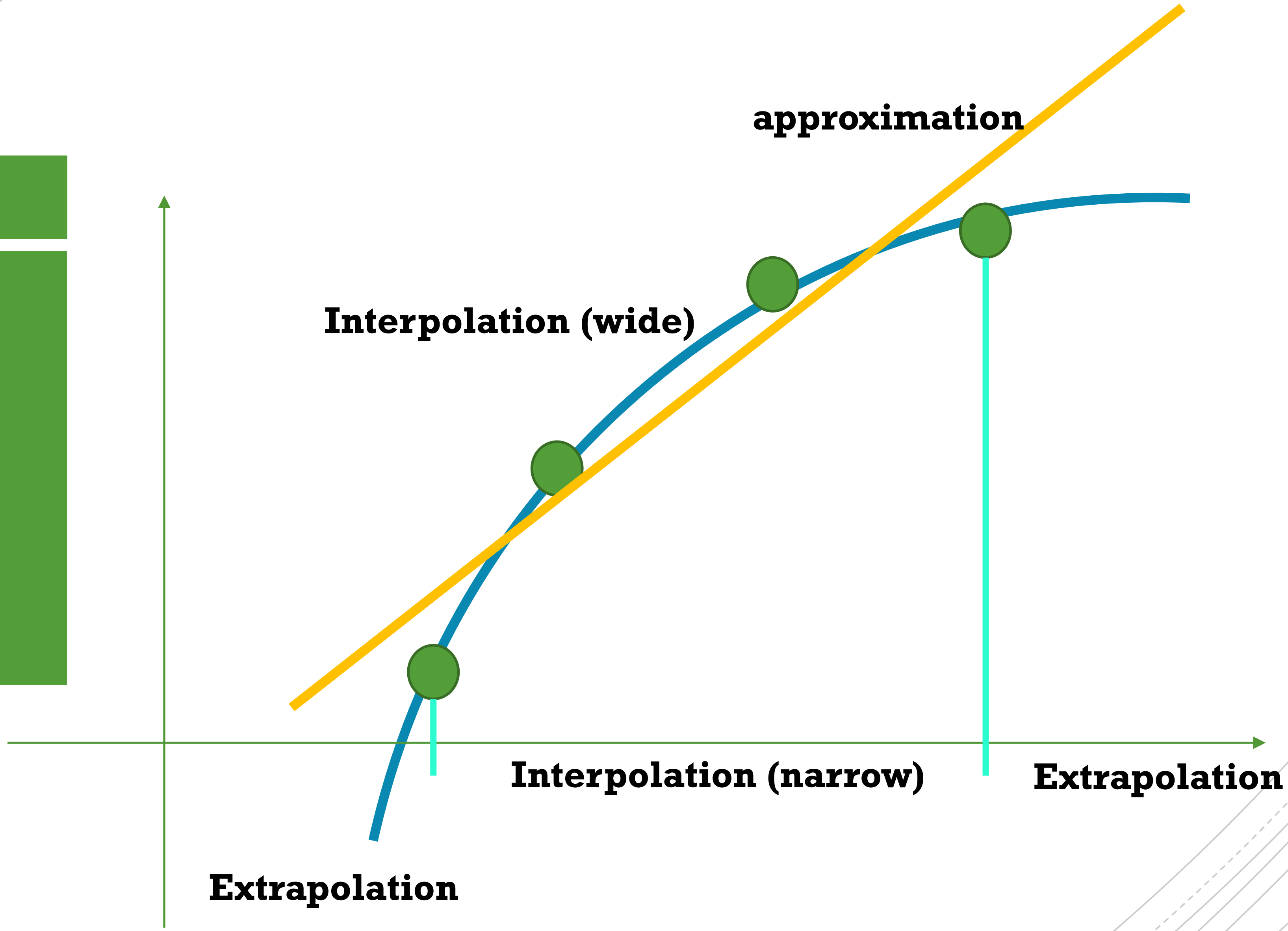
- Самое важное приложение - это подгонка данных.
- Наилучшая подгонка в смысле наименьших квадратов сводит к минимуму сумму квадратов остатков (остаток - это разница между наблюдаемым значением и установленным значением, предоставленным моделью).
- Простой набор данных состоит из **n** точек (пар данных) (x_i, y_i), где x_i независимая переменная,а y_i зависимая переменная, значение которой определяется путем наблюдения. Модельная функция имеет вид $f(x, \beta)$, где m регулируемые параметры хранятся в векторе β.
- Цель состоит в том, чтобы найти значения параметров для модели, которая "наилучшим образом" соответствует данным. Соответствие модели точке данных измеряется ее остаточным значением, определяемым как разница между фактическим значением зависимой переменной и значением, предсказанным моделью:

$$r_i = y_i - f(x_i, \beta)$$

- Метод наименьших квадратов находит оптимальные значения параметров путем минимизации суммы, S, квадратов остатков (residuals):

$$S = \sum_{i=1}^{n} r_i^2$$

Definitions in plot



Спасибо за внимание!

В случае вопросов по лекции задавайте их через форму:

https://forms.yandex.ru/u/61ffab0425b437e0e3410e9b/

Мы обязательно обсудим их на следующем занятии.