Аппроксимация данных. Наилучшее равномерное приближение.

Постановка задачи. Пусть задана дискретная функция $y(x_i) = y_i,\ i = 0, 1, 2, \ldots N$. Требуется построить алгебраичекий полиноном $P_n^*(x)$ степени n, являющийся решением следующей задачи

$$\rho = \inf_{P_n(x)} \max_{i\in[0,\ldots,N]} |y_i - P_n(x_i)|$$

Такой полином называется полиномом наилучшего равномерного приближения исходной функции $y(x_i) = y_i$ по узлам $x_i,\ i = 0, 1, 2, \ldots N$.

Решение данной задачи существенно зависит от соотношения между n и N. Так, если $N = n$, т.е. количество узлов совпадает с количеством коэффициентов исходного полинома, то однозначно строится полином, принимающий в узлах x_i значения y_i. Очевидно, что в этом случае $\rho = 0$ и этот полином является искомым $P_n^*(x)$.

Первым нетривиальным (и как оказывается основным) случаем является случай $N = n + 1$. Он приводит к так называемой чебышевской интерполяции. С помощью чебышевской интерполяциии можно построить полином наилучшего приближения и в общем случае $N > n + 1$.

Чебышевская интерполяция.

Имеет место следующая

Теорема. Пусть

$$\rho = \inf_{P_n(x)} \max_{i\in[0,\ldots,n+1]} |y_i - P_n(x_i)|.$$

Полином наилучшего равномерного приближения существует и единствевен. Для того чтобы полином $P_n^*(x)$ был полиномом наилучшего приближения, необходимо и достаточно, чтобы при некотором h выполнялись соотношения

$$(-1)^i h + P_n^*(x_i) = y_i, \quad i \in [0, ..., n+1] \tag{1}$$

При этом имеет место равенство $\rho = |h|$.

Замечание. Из условия теоремы следует, что h и коэффициенты полинома $P^*(x)$ могут быть найдены из решения следующей системы линейных алгебраических уравнений с $(n + 2)$ неизвестными $h, a_0, a_1, ..., a_{n+1}$

$$\begin{cases} h + a_0 + a_1 x_0 + \ldots + a_n x_0^n = y_0 \\ -h + a_0 + a_1 x_1 + \ldots + a_n x_1^n = y_1 \\ h + a_0 + a_1 x_2 + \ldots + a_n x_2^n = y_2 \\ \ldots \\ (-1)^{n+1} h + a_0 + a_1 x_{n+1} \ldots + a_n x_{n+1}^n = y_{n+1} \end{cases}$$

Определителем этой системы является определитель Ван дер Монда, он отличен от нуля если $x_i \neq x_j,\ i \neq j$. Однако, построение полинома $P_n^*(x)$ через явное вычисление его коэффициентов приводит к катастрофической потере точности уже при $n \approx 50$.

Поэтому обычно применяют следующих подход. Имеет место

Утверждение. Величина h из (1) может быть вычислена по формуле

$$h = \sum_{i=0}^{n+1} (-1)^i \alpha_i y_i, \quad \text{где } \alpha_i = \frac{\dfrac{(-1)^i}{\prod\limits_{k\neq i, k=0}^{n+1} (x_i - x_k)}}{\sum\limits_{j=0}^{n+1} \dfrac{(-1)^j}{\prod\limits_{k\neq j, k=0}^{n+1} (x_j - x_k)}} \tag{2}$$

Отметим, что используя разделенные разности

$$f(x_i; x_j) = \frac{f(x_j) - f(x_i)}{x_j - x_i},\ f(x_1; \ldots; x_{k+1}) = \frac{f(x_2; \ldots; x_{k+1}) - f(x_1; \ldots; x_k)}{x_{k+1} - x_1}$$

равенство (2) можно переписать в виде

$$h = \frac{y(x_0; x_1; ...; x_{n+1})}{\varphi(x_0; x_1; ...; x_{n+1})}, \quad \text{где } \varphi_k = \varphi(x_k) = (-1)^k, \tag{3}$$

и записать полином наилучшего равномерного приближения либо в форме Ньютона:

$$P^*(x) = y_0 - h + \sum_{k=1}^{n} \left(y(x_0; \ldots; x_k) - h\varphi(x_0; \ldots; x_k)\right)(x - x_0)\ldots(x - x_{k-1}), \tag{4}$$

либо в форме Лагранжа.

При практическом применении формулы (4) следует построить таблицы разделенных разностей функций y и φ, найти значение h и результат подставить в (4).

Для контроля правильности полученного полинома стоит убедиться в выполнении условия (1) Теоремы.

Общая дискретная задача. Алгоритм Валле-Пуссена.

Рассмотрим общий случай $N > n+1$. Положим

$$\rho = \inf_{P(x)} \max_{i \in [0,\dots,N]} |y_i - P(x_i)|.$$

Из всех полиномов $P(x)$ степени n требуется определить полином наилучшего приближения $P^*(x)$. Здесь и далее индекс n у полиномов мы опускаем.

Можно показать, что общая дискретная задача имеет точное решение. Оно может быть получено с помощью конечного числа чебышевских итераций.

Введем следующее определение.

def. Базисом σ называется любая $(n+2)$-точечная подсистема узлов

$$\sigma = \left\{ x_{i_0} < x_{i_1} < \dots < x_{i_{n+1}} \right\}.$$

Очевидно, что на каждом базисе σ можно осуществить чебышевскую интерполяцию, т.е. построить алгебраический полином $P^*(\sigma, x)$, удовлетворяющий соотношениям

$$(-1)^k h + P^*(\sigma, x_{i_k}) = y_k, \quad k \in [0, \dots, n+1].$$

Напомним, что в этом случае

$$\inf_{P(x)} \max_{k \in [0,\dots,n+1]} |y_{i_k} - P(x_{i_k})| = |h|.$$

Положим

$$\begin{aligned} h(\sigma) &= \max_{k \in [0,\dots,n+1]} |y_{i_k} - P^*(\sigma, x_{i_k})|, \\ \varphi(\sigma) &= \max_{i \in [0:N]} |y_i - P^*(\sigma, x_i)|, \end{aligned} \tag{5}$$

т.е. $h(\sigma)$ есть максимальное уклонение на узлах базиса σ, $\varphi(\sigma)$ есть максимальное уклонение на всей системе узлов. Очевидно, что для любого базиса σ выполняется следующее неравенство

$$\varphi(\sigma) \geq h(\sigma)$$

Имеет место

Лемма. *Если для некоторого базиса σ^* оказалось, что*

$$\varphi(\sigma^*) = h(\sigma^*)$$

то полином $P(\sigma^*, x)$ является полиномом $P^*(x)$ наилучшего приближения по всем узлам $x_0, \dots, x_N$.

Таким образом, за конечное число шагов (возможно, осуществив полный перебор базисов σ_i, построенных по узлам $x_0, \dots, x_N$) мы отыщем интересующий нас базис. Однако полный перебор можно значительно сократить.

Опишем некоторое преобразование S базиса σ в

$$\sigma_1 = \left\{ x_{i_0}^{(1)} < x_{i_1}^{(1)} < \dots < x_{i_{n+1}}^{(1)} \right\}$$

удовлетворяющее условию

$$h(\sigma) > h(\sigma_1)$$

Лемма. *Если для базиса σ верно $\varphi(\sigma) > h(\sigma)$, то существует новый базис $\sigma_1 = S\sigma$, такой что $h(\sigma_1) > h(\sigma)$.*

Положим $h^* = \max\limits_{\sigma} h(\sigma)$.

def. Базис σ^* для которого $h(\sigma^*) = h^*$ называется экстремальным базисом.

Поскольку число различных базисов σ конечно, то очевидно экстремальный базис существует, хотя в общем случае он может быть и не единственным.

Теорема. *В общей дискретной задаче полином наилучшего приближения существует и единственен. Для того чтобы полином $P^*(x)$ был полиномом наилучшего приближения, необходимо и достаточно, чтобы он осуществлял чебышевскую интерполяцию на некотором экстремальном базисе σ^*.*

Следствие. *Для того чтобы базис σ был экстремальным, необходимо и достаточно, чтобы выполнялось равенство*

$$\varphi(\sigma) = h(\sigma)$$

Таким образом, алгоритм Валле-Пуссена заключается в переборе базисов по алгоритму $\sigma_i = S\sigma_{i-1}$ до тех пор пока не выполнится равенство $\varphi(\sigma_i) = h(\sigma_i)$.

Алгоритм Валле-Пуссена (S-алгоритм).

Пусть $\Delta_\sigma(x_i) = y_i - P^*(\sigma, x_i)$.

Выберем произвольный базис $\sigma = \left\{ x_{i_0} < x_{i_1} < \dots < x_{i_{n+1}} \right\}$. Вычислим согласно формулам (5) величины $h(\sigma)$ и $\varphi(\sigma)$. Если $h(\sigma) = \varphi(\sigma)$, то согласно Следствию из Теоремы базис σ является экстремальным и построенный по нему чебышевский полином являетя искомым.

Если $h(\sigma) < \varphi(\sigma)$, то возьмем точку x_{k_0} для которой

$$\varphi(\sigma) = |\Delta_\sigma(x_{k_0})|.$$

Если таких точек несколько, то выберем произвольную. Возможны три случая:

1. $x_{i_0} < x_{k_0} < x_{i_{n+1}}$.
2. $x_{k_0} < x_{i_0}$.
3. $x_{i_{n+1}} < x_{k_0}$.

В каждом из этих трех случаев преобразование S определяется по-своему.

Первый случай. Выберем ν такое, что $x_{i_\nu} < x_{k_0} < x_{i_{\nu+1}}$. Тогда

$$x_{i_j}^{(1)} = x_{i_j}, \quad j \neq \nu, j \neq \nu+1;$$

Если $sign\Delta_\sigma(x_{k_0}) = sign\Delta_\sigma(x_{i_\nu})$, то

$$x_{i_\nu}^{(1)} = x_{k_0}, \quad x_{i_{\nu+1}}^{(1)} = x_{i_{\nu+1}}.$$

Иначе, т.е. если $sign\Delta_\sigma(x_{k_0}) = sign\Delta_\sigma(x_{i_{\nu+1}})$

$$x_{i_\nu}^{(1)} = x_{i_\nu}, \quad x_{i_{\nu+1}}^{(1)} = x_{k_0}.$$

Второй случай. Если $sign\Delta_\sigma(x_{k_0}) = sign\Delta_\sigma(x_{i_0})$, то

$$x_{i_0}^{(1)} = x_{k_0}, \quad x_{i_j}^{(1)} = x_{i_j},\ j \neq 0.$$

Иначе

$$x_{i_0}^{(1)} = x_{k_0}, \quad x_{i_j}^{(1)} = x_{i_{j-1}},\ j = 1, 2, ..., n+1.$$

т.е. из базиса выбрасывается узел $x_{i_{n+1}}$.

Третий случай. Если $sign\Delta_\sigma(x_{k_0}) = sign\Delta_\sigma(x_{i_{n+1}})$, то

$$x_{i_{n+1}}^{(1)} = x_{k_0}, \quad x_{i_j}^{(1)} = x_{i_j},\ j \neq n+1.$$

Иначе

$$x_{i_{n+1}}^{(1)} = x_{k_0}, \quad x_{i_j}^{(1)} = x_{i_{j+1}},\ j = 0, 1, ..., n.$$

т.е. из базиса выбрасывается узел x_{i_0}.

Из построения базиса и неравенства $\varphi(\sigma) > h(\sigma)$ вытекают следующие свойства базиса σ_1 если $h(\sigma) > 0$:

I. $sign\Delta_\sigma(x_{i_k}^{(1)}) = -sign\Delta_\sigma(x_{i_{k+1}}^{(1)}), \quad k \in [0:n]$.

II. Если обозначить через $x_{i_s}^{(1)}$ узел базиса σ_1, соответствующий x_{k_0}, то

$$\left|\Delta_\sigma(x_{i_k}^{(1)})\right| = \rho(\sigma), \quad k \neq s.$$

$$\left|\Delta_\sigma(x_{i_s}^{(1)})\right| > \rho(\sigma).$$

Данные свойства позволяют доказать основное неравенство

$$\rho(\sigma_1) > \rho(\sigma).$$

Аппроксимация данных. Наилучшее квадратичное приближение.

Постановка задачи. Пусть задана дискретная функция $y(x_i) = y_i,\ i = 0, 1, 2, \ldots N$ и некоторая система $\{\varphi^{(k)}(x)\}_1^n$ базисных функций. Требуется найти коэффициенты c_i для $F(x) = \sum\limits_{i=0}^{n} c_i\varphi^{(i)}(x)$, являющейся решением следующей задачи

$$\rho = \inf_{c_i} \sum_{i=0}^{N} (y_i - F(x_i))^2.$$

Функция $F(x)$ называется наилучшим квадратичным приближением для задачи аппроксимации данных методом наименьших квадратов. В матричном виде нахождение коэффициентов c_i искомого представления сводится к минимизации квадратичного функцилнала:

$$\inf_c \sum_{i=0}^{N} (y_i - \sum_{j=0}^{n} c_j\varphi^{(j)}(x_i))^2 = \inf_c \sum_{i=0}^{N} ([b - Ac]_i)^2,$$

где $c = (c_0, \ldots, c_n)^T$, $(A)_{ij} = \varphi^{(j)}(x_i)$, $b = (y_0, \ldots, y_N)^T$.

В общем случае не все точки x_i могут иметь одинаковую значимость, возможно, некоторые значения известны с погрешностью. Эта информация может быть учтена за счет добавления в минимизационную задачу весовых множителей:

$$\rho = \inf_{c_i} \sum_{i=0}^{N} \alpha_i(y_i - F(x_i))^2$$

В качестве α_i можно взять, например, длину отрезка $x_i - x_{i-1}$, либо $1/\varepsilon_i$, где ε_i – погрешность значения y_i.

Утверждение. *Вектор c минимизирующий $\|Ac - b\|_2^2$ является решением системы уравнений $A^TAx = A^Tb$.*

Данный метод может эффективно применяться, если размерность задачи и число обусловленности A^TA не слишком велики. Для систем большой размерности, а также если матрица A^TA близка к вырожденной (например, система базисных функций почти линейно зависима), тогда рекомендуется применять метод QR–разложения. Предположим, что известна матрица отражений Q такая, что $Q^TQ = I$ и QA равно верхнетреугольной R. Так как искомое решение имеет вид $c = (A^TA)^{-1}A^Tb$, следовательно $Rc = Q^Tb$ и вектор c находится обратным ходом метода Гаусса. При этом

$$\inf_c \|Ac - b\|_2^2 = \inf_c \|Q^TAc - Q^Tb)\|_2^2.$$

Для повышения устойчивости данного алгоритма, преобразуем исходную задачу так, чтобы первые r столбцов матрицы были линейно независимы. Для этого рассмотрим $\tilde{A} = AP$, где P– некоторая матрица перестановок. То есть переставим столбцы в матрице A (как именно определим в процессе вычислений) и решим задачу наименьших квадратов для полученной эквивалентной системы с матрицей $\tilde{A}$, т.е. построим разложение $\tilde{A} = QR$. Наша задача – получение в матрице $R = \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix}$ как можно лучше обусловленный блок R_{11} и как можно меньшие элементы в R_{22}. Отметим, что в приближенных вычислениях блок R_{22} всегда отличен от нуля, хотя исходная задача могла быть неполного ранга.

Численное решение задачи наименьших квадратов методом QR**–разложения с выбором главного столбца.** На k-ом шаге ($1 \le k \le n$) в матрице A выбирается столбец с номером j_k, $k \le j_k \le n$, с наибольшей нормой $\max\limits_{k \le j \le n} \left(\sum\limits_{i=k}^{m} a_{ij}^2 \right)^{1/2}$ в неприведенной части A (подматрицы $A^{(k)}$ с элементами a_{ij} и $k \le i \le m$, $k \le j \le n$). В матрице A столбец j_k переставляется с k-м столбцом. Далее применяется обычное отражение – очередной шаг QR–разложения.

Задание **Наилучшее равномерное приближение.**

Дано: Файл с парами чисел x_i, y_i, $i = 0, \ldots, N$.

Построить: Алгебраичекий полиноном $P_n^*(x)$ степени n, являющийся решением следующей задачи

$$\inf_{P_n(x)} \max_{i \in [0, \ldots, N]} |y_i - P_n(x_i)|$$

Такой полином называется полиномом наилучшего равномерного приближения исходной функции $y(x_i) = y_i$ по узлам x_i, $i = 0, 1, 2, \ldots N$.

Тест: Для функции $f(x) = x^N + x^{N-1}$ сформировать файл с данными $\{x_i, y_i = f(x_i)\}$, $i = 0, \ldots, N$, выбрав в качестве узлов $x_i \in [-1, 1]$ экстремумы многочлена Чебышева $T_N(x)$, и построить численно многочлен $P_n^*(x)$ наилучшего равномерного приближения степени $n = N - 1$. Сравнить с точным аналитическим решением $P_n^*(x) = f(x) - T_N(x) \cdot 2^{1-N}$.

Задание **Наилучшее среднеквадратичное приближение.**

Дано: Файл с парами чисел x_i, y_i, $i = 0, \ldots, N$ и система аналитических функций $\varphi^{(j)}(x)$, $j = 0, \ldots, n$.

Найти: Коэффициенты c_i в разложении функции $F(x) = \sum\limits_{j=0}^{n} c_j \varphi^{(j)}(x)$, являющийся решением задачи

$$\rho = \inf_c \sum_{i=0}^{N} \alpha_i (y_i - F(x_i))^2.$$

Функция $F(x)$ называется наилучшим квадратичным приближением.

Тест: Для матрицы A матрица R из QR–разложения с выбором главного столбца с точностью до знаков имеет вид

$$A = \begin{pmatrix} 1 & -1 & -1 & -1 & -1 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad R = \begin{pmatrix} 2.23 & 0.89 & 0.44 & 0.00 & -0.44 \\ 0.00 & 1.78 & 0.33 & 0.00 & -0.33 \\ 0.00 & 0.00 & -1.63 & 0.00 & 0.41 \\ 0.00 & 0.00 & 0.00 & -1.41 & 0.70 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.10 \end{pmatrix}$$