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 $\{\phi^{(k)}(x)\}_1^n$ базисных функций. Требуется найти коэффициенты c_i для $F(x) = \sum\limits_{i=0}^{n} c_i \phi^{(i)}(x)$, 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:

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

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

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 с минимизирующий $\|Ac - b\|_2^2$ является решением системы уравнений $A^T Ax = A^T b$.*

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

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

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 среднеквадратичное приближение.**

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

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

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