Лекция 10

Тема 6. Приближение функций

§ 6.1. Постановка задачи приближения функций

Вычисление значения функции $y = f(x)$ – одна из тех задач, с которой на практике приходится сталкиваться. Естественно, что при решении на ЭВМ серьезных задач желательно иметь быстрые и надежные алгоритмы вычисления значений используемых функций. Для элементарных, а также основных специальных функций такие алгоритмы разработаны, реализованы в виде стандартных программ и включены в математическое обеспечение ЭВМ. Однако в расчетах нередко используются и другие функции, непосредственное вычисление которых затруднено либо приводит к слишком большим затратам машинного времени. Укажем на некоторые типичные ситуации.

1. Функция f задана таблицей своих значений:

$$y_i = f(x_i), \ (i = 0, 1, 2, \dots, n), \tag{6.1}$$

а вычисления проводятся в точках x, не совпадающих с табличными.

2. Непосредственное вычисление значения $y = f(x)$ связано с проведением сложных расчетов и приводит к значительным затратам машинного времени, которые могут оказаться неприемлемыми, если функция f вычисляется многократно.

3. При заданном значении x значение $f(x)$ может быть найденным из эксперимента. В большинстве случаев нахождение значения функции из эксперимента в реальном масштабе времени невозможно. В этой ситуации экспериментальные данные получают до начала вычислений. Нередко они представляют таблицу вида (6.1) с тем отличием, что табличные значения y_i^* отличаются от истинных значений y_i, так как заведомо содержат ошибки эксперимента.

Возникающие проблемы нередко удается решить следующим образом. Функцию $f(x)$ приближенно заменяют другой функцией $g(x)$, вычисляемые значения которой и принимают за приближенное значение функции f. Конечно, такая замена оправдана лишь тогда, когда значения $g(x)$ вычисляются быстро и надежно, а погрешность приближения $f(x) - g(x)$ достаточно мала. Обсудим кратко некоторые вопросы, с которыми в каждом конкретном случае приходится сталкиваться при выборе постановки задачи приближения и метода ее решения.

1). Необходимо решить, какую информацию о функции f можно использовать как входные данные для вычисления приближения g. Например, часто известна или может быть получена таблица значений функции вида (6.1), а иногда – и таблица ее производных. В некоторых случаях можно использовать информацию о значениях функции на всем отрезке $[a, b\,]$.

2). Полезно иметь некоторую дополнительную априорную информацию об аппроксимируемой функции. Часто она бывает качественного характера, например, известно, что функция f «достаточно гладкая» («плавно меняющаяся»), периодическая, монотонная, четная и т.п. Иногда удается получить некоторые количественные характеристики функции f, например, бывают известны верхние оценки для максимума модуля некоторых ее производных, величина периода, оценка уровня погрешности в заданных значениях.

3). Знание свойств функции f позволяет осознанно выбирать класс G аппроксимирующих функций. Часто такой класс представляет собой параметрическое семейство функций вида $y = g\ (x, \boldsymbol{a}) = g\ (x, a_0,\ a_1,\ \dots,\ a_m)$ и выбор конкретной аппроксимирующей функции g осуществляется с помощью параметров $a_0,\ a_1,\ \dots,\ a_m$. Широко используются классы функций вида

$$\Phi_m(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \ldots + a_m\varphi_m(x), \tag{6.2}$$

являющихся линейными комбинациями фиксированного набора некоторых базисных функций $\varphi_0(x),\ \varphi_1(x),\ \ldots,\ \varphi_m(x)$. Функцию $\Phi_m(x)$ часто называют *обобщенным многочленом* по системе функций $\varphi_0(x),\ \varphi_1(x),\ \ldots,\ \varphi_m(x)$, а число m – его *степенью*.

Если в качестве базисных функций берутся степенные функции $\varphi_k(x) = x^k$, то возникает задача приближения алгебраическими многочленами

$$P_m(x) = a_0 + a_1x + \ldots + a_m x^m. \tag{6.3}$$

Отметим, что методы приближения функций алгебраическими многочленам играют важную роль в численном анализе и наиболее глубоко разработаны. Одна из причин этого состоит в том, что многочлены (6.3) легко вычисляются, без труда дифференцируются и интегрируются.

Тригонометрические многочлены

$$S_m(x) = a_0 + \sum_{1\le k\le m/2}(\alpha_k \cos 2\pi kx + \beta_k \sin 2\pi kx), \tag{6.4}$$

часто используемые для аппроксимации периодических на отрезке [0, 1] функций, также могут быть записаны в виде (6.2), если в качестве базисных функций выбрать функции $\varphi_0(x) = 1$, $\varphi_1(x) = \cos 2\pi x$, $\varphi_2(x) = \sin 2\pi x$, $\varphi_3(x) = \cos 4\pi x$, $\varphi_4(x) = \sin 4\pi x, \dots$.

Используя формулу Эйлера $\exp\{iy\} = \cos y + i \sin y$, можно записать тригонометрический многочлен (6.4) в виде

$$S_m(x) = \sum_{-m/2\le k\le m/2} a_k \exp\{2\pi ikx\}, \tag{6.5}$$

что соответствует выбору базисных функций $\varphi_k(x) = \exp\{2\pi ikx\}$, $-m/2 \le k \le m/2$.

4). Необходим критерий выбора в классе G конкретной функции g, являющейся в этом классе наилучшим приближением к f. Например, требование совпадения функции g с функцией f в некоторых фиксированных точках приводит к задаче интерполяции. Другой распространенный критерий – требование минимизации среднеквадратического отклонения – лежит в основе метода наименьших квадратов.

5). Важно понимать, что решение указанных выше вопросов тесно связано с тем, как мы собираемся использовать приближение g и какая точность нам нужна.

§ 6.2. Интерполяция обобщенными многочленами.

1. Постановка задачи интерполяции. Пусть в точках $x_0, x_1, \dots, x_n$, расположенных на отрезке $[a, b]$ и попарно различных, задана таблица (6.1) значений некоторой функции f. *Задача интерполяции* состоит в построении функции g, удовлетворяющей условию

$$g(x_i) = y_i \quad (i = 0, 1, \dots, n). \tag{6.6}$$

Другими словами, ставится задача построения функции g, график которой проходит через заданные точки (x_i, y_i). Указанный способ приближения функций принято называть *интерполяцией* (или *интерполированием*), а точки x_i – *узлами интерполяции*.

Нетрудно видеть, что выбор функции g неоднозначен, так как по заданной таблице можно построить бесконечно много интерполирующих функций. На практике, как правило, функцию g выбирают из достаточно узкого класса G функций, в котором единственность выбора гарантируется.

2. Экстраполяция. Пусть x_{min} и x_{max} – минимальный и максимальный из узлов интерполяции. В случае, когда интерполяция используется для вычисления приближенного значения функции f в точке x, не принадлежащей отрезку $[x_{min}, x_{max}]$ *(отрезку наблюдения)*, принято говорить о том, что осуществляется *экстраполяция*. Этот метод приближения часто используют с целью прогнозирования характера протекания тех или иных процессов при значениях параметров x, выходящих за пределы отрезка наблюдения. Заметим, что надежность такого прогноза при значениях x, удаленных на значительное расстояние от отрезка $[x_{min}, x_{max}]$, как правило, невелика.

3. Задача интерполяции обобщенными многочленами. Рассмотрим более подробно задачу интерполяции обобщенными многочленами $\Phi_m(x)$ вида (6.2). Назовем обобщенный многочлен $\Phi_m(x)$ *интерполяционным*, если он удовлетворяет условию

$$\Phi_m(x_i) = y_i \quad (i = 0, 1, 2, \dots, n), \tag{6.7}$$

или, что то же самое, системе линейных алгебраических уравнений

$$\begin{gathered} \varphi_0(x_0)a_0 + \varphi_1(x_0)a_1 + \dots + \varphi_m(x_0)a_m = y_0, \\ \varphi_0(x_1)a_0 + \varphi_1(x_1)a_1 + \dots + \varphi_m(x_1)a_m = y_1, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \varphi_0(x_n)a_0 + \varphi_1(x_n)a_1 + \dots + \varphi_m(x_n)a_m = y_n \end{gathered} \tag{6.8}$$

относительно коэффициентов $a_0, a_1, \dots, a_m$.

Заметим, что систему уравнений (6.7) можно записать в следующем виде:

$$\boldsymbol{Pa} = \boldsymbol{y}, \tag{6.9}$$

где

$$\boldsymbol{P} = \begin{bmatrix} \varphi_0(x_0) & \varphi_1(x_0) & \dots & \varphi_m(x_0) \\ \varphi_0(x_1) & \varphi_1(x_1) & \dots & \varphi_m(x_1) \\ \dots & \dots & \dots & \dots \\ \varphi_0(x_n) & \varphi_1(x_n) & \dots & \varphi_m(x_n) \end{bmatrix}, \quad \boldsymbol{a} = \begin{bmatrix} a_0 \\ a_1 \\ \dots \\ a_m \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} y_0 \\ y_1 \\ \dots \\ y_n \end{bmatrix}. \tag{6.10}$$

Введем векторы $\boldsymbol{\varphi_j} = (\varphi_j(x_0), \varphi_j(x_1), \dots, \varphi_j(x_n))^{\mathrm{T}}$, $j = 0, 1, \dots, m$. Будем говорить, что *система функций* $\varphi_0(x), \varphi_1(x), \dots \varphi_m(x)$ *линейно зависима в точках* $x_0, x_1, \dots, x_n$, если один из векторов $\boldsymbol{\varphi_j}$ системы $\boldsymbol{\varphi_0}, \boldsymbol{\varphi_1}, \dots, \boldsymbol{\varphi_m}$ может быть представлен в виде линейной комбинации остальных векторов этой системы:

$$\boldsymbol{\varphi_j} = \sum_{k=0, k\neq j}^{m} \boldsymbol{\alpha_k}\, \boldsymbol{\varphi_k}. \tag{6.11}$$

В противном случае систему функций $\varphi_0(x), \varphi_1(x), \dots \varphi_m(x)$ будем называть *линейно независимой в точках* $x_0, x_1, \dots, x_n$.

У т в е р ж д е н и е 6.1. При $m \leq n$ система функций 1, x, x^2, …, x^m линейно независима в точках $x_0, x_1, \dots, x_n$, если они попарно различны.

□ Допустим противное. Тогда справедливо равенство (6.11), которое в данном случае принимает вид

$$x_i^j = \sum_{k=0, k\neq j}^{m} \alpha_k\, x_i^k, \qquad (i = 0, 1, \dots, n) \tag{6.12}$$

Полагая $\alpha_j = -1$, получаем, что многочлен $P_m(x) = \sum_{k=0}^{m} \alpha_k\, x^k$ степени m обращается в нуль в точках $x_0, x_1, \dots, x_n$, число которых равно $(n + 1)$ и, следовательно, больше m. Однако в силу основной теоремы алгебры многочлен степени m, тождественно не равный нулю, не может иметь более m корней. Полученное противоречие доказывает независимость рассматриваемой системы функций. ■

Рассмотрим *матрицу Грама* системы функций $\varphi_0, \varphi_1, \dots, \varphi_m$, имеющую вид

$$\boldsymbol{\Gamma} = \boldsymbol{P}^*\boldsymbol{P} = \begin{bmatrix} (\boldsymbol{\varphi_0}, \boldsymbol{\varphi_0}) & (\boldsymbol{\varphi_1}, \boldsymbol{\varphi_0}) & \dots & (\boldsymbol{\varphi_m}, \boldsymbol{\varphi_0}) \\ (\boldsymbol{\varphi_0}, \boldsymbol{\varphi_1}) & (\boldsymbol{\varphi_1}, \boldsymbol{\varphi_1}) & \dots & (\boldsymbol{\varphi_m}, \boldsymbol{\varphi_1}) \\ \dots & \dots & \dots & \dots \\ (\boldsymbol{\varphi_0}, \boldsymbol{\varphi_m}) & (\boldsymbol{\varphi_1}, \boldsymbol{\varphi_m}) & \dots & (\boldsymbol{\varphi_m}, \boldsymbol{\varphi_m}) \end{bmatrix}. \tag{6.13}$$

Здесь в случае, когда функции $\varphi_j(x)$ могут принимать комплексные значения, под $\boldsymbol{P}^*$ понимается сопряженная к $\boldsymbol{P}$ матрица, а элементы γ_{jk} матрицы Грама вычисляются по формуле

$$\gamma_{jk} = (\boldsymbol{\varphi_k}, \boldsymbol{\varphi_j}) = \sum_{i=0}^{n} \varphi_k(x_i)\overline{\varphi_j(x_i)}, \tag{6.14}$$

Если же функции $\varphi_j(x)$ принимают только вещественные значения, то $\boldsymbol{P}^*= \boldsymbol{P}^{\text{т}}$ и элементы матрицы Грама вычисляются по формуле

$$\gamma_{jk} = (\boldsymbol{\varphi_k}, \boldsymbol{\varphi_j}) = \sum_{i=0}^{n} \varphi_k(x_i)\varphi_j(x_i). \tag{6.15}$$

Определитель матрицы Грама det $\boldsymbol{\Gamma}$ принято называть *определителем Грама*. Как следует из курса линейной алгебры, справедлив следующий результат.

Т е о р е м а 6.1. *Система функций* $\varphi_0(x), \varphi_1(x), \dots \varphi_n(x)$ *является линейно независимой в точках* $x_0, x_1, \dots, x_n$ *тогда и только тогда, когда* $m \leq n$ *и определитель Грама* det $\boldsymbol{\Gamma}$ *отличен от нуля.*

Известно, что при $m > n$ система функций $\varphi_0(x), \varphi_1(x), \dots \varphi_m(x)$ линейно зависима в точках $x_0, x_1, \dots, x_n$. Отсюда вытекает неединственность решения $\boldsymbol{a}$ системы (6.9) (если оно существует). Действительно, в этом случае справедливо представление (6.11) и вместе с вектором $\boldsymbol{a}$ решением системы (6.9) является вектор $\boldsymbol{a}' = \boldsymbol{a} + t\,\Delta\boldsymbol{a}$, где $\Delta\boldsymbol{a} = (\alpha_0, \alpha_0, \dots, \alpha_0, -1, \alpha_0, \dots \alpha_0)^{\text{т}}$, а t – любое число. Если же $m < n$, то решение системы (6.9) существует не для всякой правой части $\boldsymbol{y}$. В силу указанных причин при интерполяции обобщенными многочленами число параметров $m+1$ обычно берут равным числу $n+1$ заданных точек. В этом случае $\boldsymbol{P}$ – квадратная матрица и для того, чтобы система (6.9) была однозначно разрешима при любой правой части $\boldsymbol{y}$, необходимо и достаточно, чтобы определитель матрицы $\boldsymbol{P}$ был отличен от нуля. В свою очередь при $m = n$ это условие в силу равенства det $\boldsymbol{\Gamma}$ = det $\boldsymbol{P}^*$ det $\boldsymbol{P}$ и теоремы 6.1 дает следующий результат.

Т е о р е м а 6.2. *Если* $m = n$, *то решение задачи интерполяции обобщенным многочленом* (6.2) *существует и единственно при любом наборе данных* $y_0, y_1, \dots, y_n$ *тогда и только тогда, когда система функций* $\varphi_0(x), \varphi_1(x), \dots \varphi_n(x)$ *линейно независима в точках* $x_0, x_1, \dots, x_n$.

Назовем систему функций $\varphi_0(x), \varphi_1(x), \dots \varphi_m(x)$ *ортогональной на множестве точек* $x_0, x_1, \dots, x_n$, если $(\boldsymbol{\varphi_k}, \boldsymbol{\varphi_j}) = 0$ при $k \neq j$ и $(\boldsymbol{\varphi_k}, \boldsymbol{\varphi_j}) \neq 0$ при $k = j$ для всех $k = 0,1, \dots ,m$; $j = 0, 1, \dots ,m$. Очевидно, что для ортогональной на множестве точек $x_0, x_1, \dots, x_n$ системы функций матрица Грама диагональна, а определитель Грама отличен от нуля. Поэтому всякая ортогональная на множестве точек $x_0, x_1, \dots, x_n$ система функций заведомо является линейно независимой в этих точках.

У т в е р ж д е н и е 6.2. Система функций $\varphi_0(x), \varphi_1(x), \ldots \varphi_{N-1}(x)$, где $\varphi_k(x) = \exp\{2\pi ikx\}$, ортогональна на множестве точек $x_l = {}^l/_N$, где $l = 0, 1, \ldots, N-1$. Здесь i – мнимая единица.

□ Для доказательства ортогональности рассматриваемой системы функций достаточно установить справедливость равенства

$$(\boldsymbol{\varphi_k}, \boldsymbol{\varphi_j}) = N\delta_{kj} \quad (k = 0,1, \ldots, N-1; j = 0, 1, \ldots, N-1), \tag{6.16}$$

где $\delta_{kj} = 0$ при $k \neq j$ и $\delta_{kj} = 1$ при $k = j$. Введем обозначение $\omega = \exp\{2\pi i/N\}$. Тогда $\varphi_k(x_l) = \exp\{2\pi ikl/N\} = \omega^{kl}$ и согласно формуле (6.14) имеем

$$(\boldsymbol{\varphi_k}, \boldsymbol{\varphi_j}) = \sum_{l=0}^{N-1} \omega^{kl}\omega^{-jl} = \sum_{l=0}^{N-1} \omega^{(k-j)l}. \tag{6.17}$$

При $k = j$ правая часть равенства (6.17), очевидно, равна N. При $k \neq j$, используя формулу суммы членов геометрической прогрессии и равенство $\omega^{(k-j)N} = \exp\{2\pi i(k\text{-}j)\} = 1$, имеем

$$(\boldsymbol{\varphi_k}, \boldsymbol{\varphi_j}) = (1 - \omega^{(k-j)N})/(1 - \omega^{k-j}) = 0.$$

Таким образом, равенство (6.16), а вместе с ним и ортогональность системы функций $\varphi_k(x) = \exp\{2\pi ikx\}$ доказаны. ■

В случае, когда система функций $\varphi_0(x), \varphi_1(x), \ldots \varphi_n(x)$ ортогональна на множестве точек $x_0, x_1, \ldots, x_n$, решение задачи интерполяции не представляет затруднений. Действительно, система уравнений (6.9) после умножения на матрицу $\boldsymbol{P}^*$ преобразуется к виду

$$\boldsymbol{\Gamma a} = \boldsymbol{b}, \quad \boldsymbol{b} = \boldsymbol{P}^*\boldsymbol{y}. \tag{6.18}$$

Заметим, что элементы вектора $\boldsymbol{b} = (b_0, b_1, \ldots, b_m)^{\text{T}}$ вычисляются по формуле

$$b_j = (\boldsymbol{y}, \boldsymbol{\varphi_j}) = \sum_{l=0}^{n} y_l\, \overline{\varphi_{j(x_l)}}, \quad j = 0, 1, \ldots, m. \tag{6.19}$$

Так как матрица $\boldsymbol{\Gamma}$ диагональна, то решение системы (6.18) находится в явном виде:

$$a_j = \frac{(\boldsymbol{y}, \boldsymbol{\varphi_j})}{(\boldsymbol{\varphi_j}, \boldsymbol{\varphi_j})}, \quad j = 0, 1, \ldots, m. \tag{6.20}$$