

СЖАТИЕ ДАННЫХ ПРИ АВТОМАТИЗАЦИИ ПРОЦЕССОВ ВОДОПОДГОТОВКИ

А.Б. Кожевников

ООО «Фирменное специализированное предприятие «КРАВТ», г. Калуга, Россия

Рассматривается математическая версия разработки программного обеспечения, позволяющего перевести ручное управление технологией водообработки на автоматический режим.

Перевод технологического оборудования водоподготовки на автоматизированный и автоматический режим работы связан, во-первых, с модернизацией этого оборудования и, во-вторых, с реорганизацией процедур контроля и управления технологическим процессом в сторону переложения их исполнения на вычислительную технику. Второе предполагает не только разработку соответствующего программного обеспечения, но и рациональное представление исходных данных, сопровождающих вычислительные процедуры. Под рациональным представлением понимается, прежде всего, минимизация памяти процессора, необходимая для хранения этой информации, т.е. возникает потребность в сжатии исходных данных, например, путем перевода емких баз экспериментальных данных в менее емкие аналитические зависимости на основе процедур интерполяции и аппроксимации.

Наиболее простым вариантом получения аналитической зависимости $z(x)$ на основе экспериментальных данных является кусочно-аналитическое представление этих данных на основе полигональной функции, все элементы которой есть линейные функции. Такой подход дает следующее выражение для $z(x)$:

$$z(x) = \sum_{k=1}^m \sigma_k(x) \cdot (a_k x_k + b_k), \quad (1)$$

где: $\sigma_k(x) = 0$ при $x \notin [x_k, x_{k+1}]$ и

$$\sigma_k(x) = 1 \text{ при } x \in [x_k, x_{k+1}].$$

Такое представление не дает существенных преимуществ, но является гарантом высокой точности интерполяции, если в пределах каждой подобласти $\sigma_k(x)$, $k=1, \dots, m$ реальная зависимость искомой функции от x линейна. Это является гарантией того, что $z(x)$ достаточно точно определяет искомую зависимость, но не достаточно сжато ее представляет, что позволяет продолжить процедуру ее преобразований в более компактную математическую форму.

Для этих целей умножим $z(x)$ на $x^{s_2} \cos(s_3 x + s_4)$ и преобразуем полученное выражение по Лапласу. Используя таблицы связи оригиналов и изображений, а также с учетом выражения для $z(x)$ будем иметь:

$$\int_{x_1}^{x_{k+1}} e^{-s_1 x} x^{s_2} \cos(s_3 x + s_4) z(x) dx = U(s), \quad (2)$$

где $U(s)$ – известная функция; $s = (s_1, s_2, s_3, s_4)$.

Обозначив $e^{-s_1 x} x^{s_2} \cos(s_3 x + s_4) = H(s, x)$ и через σ – область определения $x \in [x_1, \dots, x_{m+1}]$, получим следующее интегральное уравнение, необходимое для аппроксимации $z(x)$,

$$\int_{\sigma} H(s, x) z(x) dx = U(s). \quad (3)$$

Будем строить приближение для $z(x)$ как непрерывную, удобную для дальнейших аналитических и числовых расчетов аналитическую функцию $\tilde{z}(x)$ в пределах интервала аппроксимации по $x \in [x_1, x_{m+1}]$ следующего вида

$$\tilde{z}(x) = \sum_{i=1}^n D_i H(\alpha_i, x) = \sum_{i=1}^n C_i \Phi(\beta_i, x). \quad (4)$$

В этом выражении $\{\Phi(\beta_i, x)\}$ – система ортонормированных функций (ортонормированный базис), построенных по правилу

$$H(\alpha_i, x) = H(s, x)|_{s=\alpha_i} \text{ при } \beta_i = (\alpha_1, \dots, \alpha_i), \quad i = 1, \dots, n, \quad (5)$$

причем $\{H(\alpha_i, x)\}$ – последовательность линейно независимых функций (неортогональный базис).

При решении рассматриваемого варианта приближения искомой функции $z(x)$ для уверенности в достоверности полученных результатов желательно иметь оценку точности приближения, что возможно, так как в

среднем квадратическом точность приближения δ в процентах определяется по формуле

$$\delta = \sqrt{\frac{\|\tilde{z}\|_2^2 + F(\alpha)}{\|\tilde{z}\|_2^2}} \cdot 100\%, \quad (6)$$

где $\|\tilde{z}\|_2^2 = \int_{t_1}^{t_{m+1}} \left[\sum_{k=1}^m \sigma_k(x)(a_k x + b_k) \right]^2 dx = \sum_{k=1}^m x \left[a_k x \left(\frac{a_k x}{3} + b_k \right) + b_k^2 \right] \Big|_{t_k}^{t_{k+1}}, \quad (7)$

а значения функции $F(\alpha)$ появятся в процессе счета на компьютере, если дальнейшие вычисления строить на основе методологии спектральной оптимизации, которая предполагает следующие математические преобразования и процедуры.

В нашем случае искомая аналитическая зависимость строится на основе минимизации квадратичного функционала

$$I = \|z(x) - \tilde{z}(x)\|_2^2 = \int_{\sigma} \left[z(x) - \sum_{i=1}^n C_i \Phi(\beta_i, x) \right]^2 dx, \quad (8)$$

Для отыскания искомого базиса выполним в (8) возведение в квадрат и учтем, что $\tilde{z}(x)$ вычисляется по формуле (4). В результате с учетом свойства ортонормированности системы $\{\Phi(\beta_i, x)\}$ получим:

$$I = \int_{\sigma} [z(x)]^2 dx - \sum_{i=1}^n C_i^2. \quad (9)$$

В этом выражении учтено, что если коэффициенты $\{C_i\}$ вычисляются по формуле [2]:

$$C_i = \sum_{j=1}^i b_j \int_{\sigma} H(\alpha_j, x) z(x) dx = \sum_{j=1}^i b_j (\alpha_1, \dots, \alpha_i) Y(\alpha_j), \quad \alpha_i \in Q, \quad i = 1, \dots, n. \quad (10)$$

где $\{b_j = b_j(\alpha_1, \dots, \alpha_i)\}$ – коэффициенты ортонормализации, полученные их численные значения при некоторых фиксированных значениях $\{\alpha_i\}$ доставляют минимум функционалу (8), т.е. дают искомое приближение с некоторой точностью δ . Отсюда следует, что можно подобрать такую последовательность параметров $\{\alpha_i\}$, что минимум квадратичного функционала (8) окажется самым глубоким из всех возможных при вычислении коэффициентов $\{C_i\}$ по формуле (10) и точность приближения δ будет наивысшей. Отыскание таких значений параметров $\{\alpha_i\}$, если проанализировать выражение (9), равнозначно поиску абсолютного минимума функции

$$F(\alpha) = -\sum_{i=1}^n C_i^2 = -\sum_{i=1}^n \left[\sum_{j=1}^i b_j (\alpha_1, \dots, \alpha_i) Y(\alpha_j) \right]^2 \quad (11)$$

по параметрам $\{\alpha_i\}$, а построенный при этом базис следует назвать оптимальным в смысле достижения требуемой точности приближения δ при минимальном значении n .

Это позволяет дать следующую формулировку задаче построения оптимального базиса и элемента $\tilde{z}(x)$.

Задача. Найти такие значения $n^0, \alpha^0 = (\alpha_1^0, \dots, \alpha_{n^0}^0)$, что

$$F(n^0, \alpha^0) = \min_{n \leq N, \alpha \in Q} \left\{ -\sum_{i=1}^n \left[\sum_{j=1}^i b_j(\alpha_1, \dots, \alpha_i) Y(\alpha_j) \right]^2 \right\} \quad (12)$$

при ограничении $\|z - \tilde{z}\| \leq \delta_0$ и минимальном значении n .

Вычислительные процедуры, сопровождающие решение этой задачи, при наличии современной компьютерной техники достаточно весомо подкреплены программным обеспечением и не создают трудностей в получении искомого результата.

Решением этой задачи является базис $\{\Phi_i^0(x)\}$,

где $\Phi_i^0(x) = \Phi(\beta_i^0, x)$, $\beta_i^0 = (\alpha_1^0, \dots, \alpha_i^0)$, $i = 1, \dots, n^0$, и элемент

$$\tilde{z}^0(x) = \sum_{i=1}^{n^0} C_i^0 \Phi(\beta_i^0, x) = \sum_{i=1}^{n^0} D_i^0 H(\alpha_i^0, x). \quad (13)$$

Заметим, что предварительная ортонормализация системы $\{H(\alpha_i, x)\}$, согласно (13), позволяет решить поставленную задачу и определить выражение для $\tilde{z}(x)$ не только в виде ортонормированного ряда, но и в виде последовательности линейно независимых неортогональных функций, что часто оказывается достаточным, так как переход от коэффициентов $\{C_i^0\}$ к коэффициентам $\{D_i^0\}$ не вызывает затруднений

$$D_i^0 = \sum_{j=1}^n C_j^0 b_{ji}, \quad i = 1, \dots, n^0. \quad (14)$$

Часто на практике встречаются двумерные экспериментально построенные графические зависимости. Так, например, зависимость растворимости хлора в воде от pH (кислотности воды) и температуры воды. При водоподготовке оборотной воды плавательных бассейнов и аква-парков существенное значение имеет зависимость концентрации хлора в воде от pH и Redox (окислительно-восстановительного потенциала) воды. Нетрудно адаптировать вышеизложенное и на двумерные, и на многомерные зависимости.

Пусть графически задана двумерная характеристика объекта $z(x, y)$. Разбив ее на m_1 интервалов кусочно-линейной интерполяции по x и m_2 интервалов кусочно-линейной интерполяции по y , можем записать

$$z(x, y) = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \sigma_j(x, y) \cdot [a_j(y) \cdot x + b_j(y)], \quad (15)$$

где

$$a_j(y) = \alpha_j y + b_j;$$

$$b_j(y) = \gamma_j y + \delta_j; \quad i = 1, \dots, m_1; \quad j = 1, \dots, m_2,$$

причем

$$\sigma_j(x, y) = \begin{cases} 0 & \text{при } x \notin [x_i, x_{i+1}] \text{ и } y \notin [y_j, y_{j+1}] \\ 1 & \text{при } x \in [x_i, x_{i+1}] \text{ и } y \in [y_j, y_{j+1}] \end{cases}.$$

Применяя двумерное преобразование Лапласа с конечными пределами к (15), будем иметь

$$\int_{y_1}^{y_{m_2+1}} \int_{x_1}^{x_{m_1+1}} e^{-s_1 x} e^{-p_1 y} z(x, y) dx dy = U(s_1, p_1). \quad (16)$$

Ход дальнейших преобразований для получения приближения в среднем квадратическом для графически заданной двумерной зависимости $z(x, y)$ с требуемой точностью в классе оптимальных базисов очевиден, т.к. опирается на процедуры, изложенные выше для одномерных случаев. Очевидно продолжение этой методики и на случай трехмерных графических экспериментальных зависимостей и представление их в виде достаточно простых аналитических зависимостей с контролем точности аппроксимации.