УДК 66.011

А.О. Пушкина, В.В. Макаров

Российский химико-технологический университет им. Д.И.Менделеева, Москва, Россия.

ИССЛЕДОВАНИЕ ОБУСЛОВЛЕННОСТИ КОРРЕЛЯЦИОННОЙ МАТРИЦЫ МЕТОДАМИ СИНГУЛЯРНОГО АНАЛИЗА

Сингулярное разложение корреляционной матрицы применено для определения чувствительности полученного решения к ошибкам входных данных. Практически алгоритм применён для расшифровки оптических спектров и численного решения интегрального уравнения Винера-Хопфа.

Singular value decomposition is used to identify solution sensitivity to input data errors. The technique has been applied to optical spectra identify and to Viner-Hopf equation solution.

Существенная проблема при применении регрессионного анализа состоит в плохой обусловленности системы нормальных уравнений. Аналогичная проблема возникает при исследовании динамики стохастических объектов, состоящей в численном решении уравнения Винера-Хопфа. Проблема заключается в корреляции исходных данных, в том числе и при отсутствии причинной связи между переменными. В связи с этим незначительные погрешности исходной информации приводят к существенным ошибкам результатов решения.

Чувствительность решения системы уравнений

$$AX = B \tag{1}$$

относительно неизвестной величины X приводит к необходимости исследования чувствительности расчета матрицы A^{-1}, обратной матрице A.

Линейный метод наименьших квадратов приводит к необходимости решать систему уравнений:

$$(A^T A)X = (A^T B), \tag{2}$$

откуда

$$X = (A^T A)^{-1}(A^T B), \tag{3}$$

где $A^T A$ – симметричная положительно определённая матрица полного ранга.

При исследовании динамики линейных стохастических систем задача идентификации модели состоит в численном решении уравнения Винера-Хопфа:

$$K_{yx}(\tau) = \int_0^\infty g(t) K_{xx}(t-\tau)d\tau, \qquad t \geq 0 \tag{4}$$

$$g(t) = 0t < 0,$$

где $K_{xx}(t-\tau)$ – корреляционная функция; $K_{yx}(\tau)$ – взаимная корреляционная функция; $g(t)$ – весовая функция; t – время; τ – шаг квантования.

Уравнение (4) можно представить в виде:

$$\frac{1}{T} K_{yx}(t) = \sum_{t=0}^{m} g_i K_{xx}(t - iT), \tag{5}$$

где $g_i = g(iT)$; $i = \overline{1,m}$.

Таким образом, при каждом значении $t = T, 2T, \dots, mT$ получим систему линейных алгебраических уравнений, из которых можно определить m значений ординат весовой функции $g(t)$ в точках $t = T, 2T, \dots, mT$.

Эта система уравнений в матричном виде может быть записана следующим образом:

$$K_{xx}G = K_{yx}. \tag{6}$$

Его решение имеет вид:

$$G = K_{xx}^{-1}K_{yx}. \tag{7}$$

Здесь K_{xx} – корреляционная матрица дискретных значений корреляционной функции:

$$K_{xx} = \begin{Vmatrix} K_{11} & K_{12} & \dots & K_{1m} \\ K_{21} & K_{22} & \dots & K_{2m} \\ \dots & \dots & \dots & \dots \\ K_{m1} & K_{m2} & \dots & K_{mm} \end{Vmatrix}, \tag{8}$$

где $K_{ij} = K_{ji}$; $K_{ij} = K_{xx}(j-i)$; $i, j = \overline{1, m}$, т.е. корреляционная матрица симметрична относительно главной диагонали; K_{yx} – вектор-столбец, элементами которого являются соответствующие значения взаимной корреляционной функции:

$$K_{yx} = \begin{Vmatrix} K_1 \\ K_2 \\ \dots \\ K_m \end{Vmatrix}; \quad K_i = \frac{K_{yx}(iT)}{T}; \tag{9}$$

G – вектор-столбец значений ординат весовой функции:

$$G = \begin{Vmatrix} g_1 \\ g_2 \\ \dots \\ g_m \end{Vmatrix}. \tag{10}$$

Чувствительность решения X системы уравнений к ошибкам, содержащихся в элементах матрицы A и вектора B, определяется спектральной нормой матрицы A, равной её максимальному собственному числу. Для вычисления числа обусловленности матрицы A как отношения её максимального собственного числа к минимальному собственному числу выполняется сингулярное разложение матрицы A, которое для симметричной положительно определённой матрицы выполняется по алгоритму:

$$A = QSQ^T, \tag{11}$$

где Q, Q^T – ортогональные матрицы; S – диагональная матрица, элементами которой, образующими главную диагональ, являются собственные числа матрицы A, расположенные в порядке их убывания. Тогда число обусловленности матрицы A определяется как отклонение максимального её собственного числа к минимальному:

$$cond\ A = \frac{\lambda_{max}}{\lambda_{min}}. \tag{12}$$

Тогда относительная ошибка решения определяется неравенствами:

$$\frac{\|\Delta X\|}{\|X+\Delta X\|} \le cond\ A \frac{\|\Delta A\|}{\|A\|} = \frac{\|\Delta A\|}{\lambda_{min}}, \tag{13}$$

$$\frac{\|\Delta X\|}{\|X\|} \le cond\ A \frac{\|\Delta B\|}{\|B\|}. \tag{14}$$

В выражениях (13), (14) все нормы евклидовы.

В качестве примера рассмотрен сингулярный анализ матрицы нормальных уравнений линейного метода наименьших квадратов, применённого для определения концентраций индивидуальных красителей по известным их спектральным коэффициентам поглощения оптического излучения и оптической плотности исходного раствора (см. рис.1.).

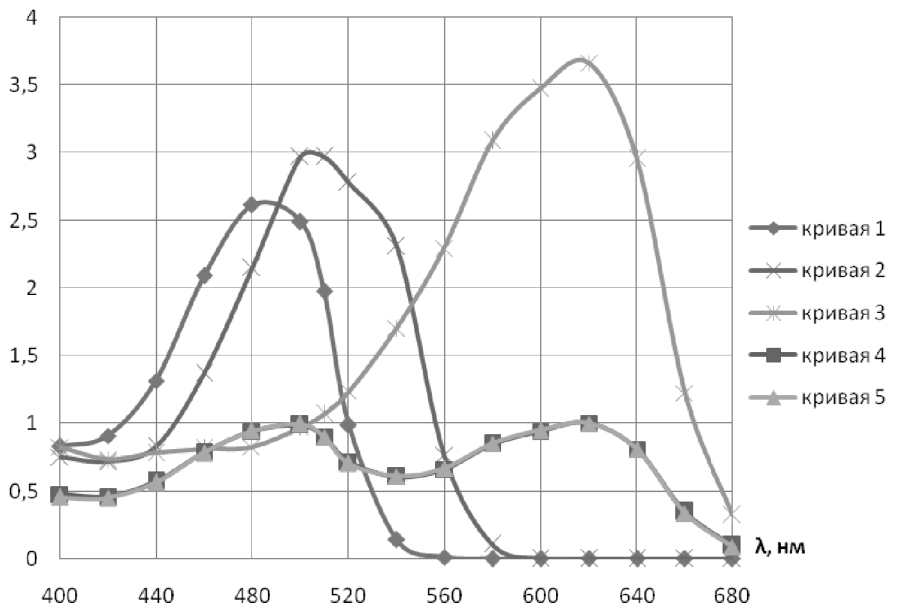

Рис. 1. Спектральные коэффициенты поглощения оптического излучения растворами индивидуальных красителей: кривые 1–3 – заданная оптическая плотность; кривые 4–5 – воспроизведённая оптическая плотность

В результате применения стандартного алгоритма сингулярного разложения симметричной положительно определённой матрицы нормальных уравнений линейного метода наименьших квадратов рассчитана диагональная матрица S

$$S = \begin{Vmatrix} 1{,}47 \cdot 10^9 & 0 & 0 \\ 0 & 6{,}20 \cdot 10^8 & 0 \\ 0 & 0 & 1{,}28 \cdot 10^8 \end{Vmatrix}$$

и определено её число обусловленности

$$cond\ A = \frac{\lambda_{max}}{\lambda_{min}} = \frac{1{,}47 \cdot 10^9}{1{,}28 \cdot 10^8} = 11{,}5$$

Расчёт показал, что число обусловленности матрицы A более, чем на порядок, отличается от единицы – числа обусловленности диагональной матрицы, между элементами которой корреляция отсутствует.

Библиографические ссылки

1. *Воеводин В.В.* Матрицы и вычисления / В. В. Воеводин, Ю. А. Кузнецов. – М.: Наука, 1984. – 320 с.
2. *Лоусон Ч.* Численное решение задач метода наименьших квадратов / Ч. Лоусон, Р. Хенсон. – М.: Наука, 1986. – 232 с.
3. *Дубровский С. А.* Прикладной многомерный статистический анализ. М.: Финансы и статистика, 1982. – 216 с.