ОБЪЕДИНЕННЫЙ
ИНСТИТУТ
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

ДУБНА

Р11 - 10186

В.Б.Злоказов

ПРИБЛИЖЕННЫЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧ РАЗЛОЖЕНИЯ СМЕСЕЙ

1976

P11 - 10186

В.Б.Злоказов

ПРИБЛИЖЕННЫЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧ РАЗЛОЖЕНИЯ СМЕСЕЙ

Направлено в ЖВМ и МФ

Злоказов В.Б. P11 - 10186

Приближенные методы решения задач разложения смесей

Метод наименьших квадратов в применении к задачам разложения регрессий на компоненты часто уязвим для критики в двух отношениях: 1) грубость моделей компонент из-за нехватки априорных сведений; 2) субъективизм параметризации. В работе изложены методы, основанные на использовании реальных физических процессов в качестве моделей и введении фундаментальных геометрических характеристик функций регрессий в качестве параметров. Исследованы возникающие при этом математические проблемы. Методы были успешно применены при анализе данных спектрометрических измерений.

Работа выполнена в Лаборатории вычислительной техники и автоматизации.

Препринт Объединенного института ядерных исследований

Дубна 1976

Zlokazov V.B. P11 - 10186

Approximate Methods for the Solution of Mixture Decomposition Problems

The application of the least-squares fit technique to the problems of mixture decomposition may be often criticized in two relations: 1) inexactitude of the component models due to the lack of apriori information; 2) parametrization subjectivity. The methods are described, which are based on the use of the measured physical processes as models and of fundamental geometric characteristics of regression functions as parameters. The arising mathematical problems are investigated. The methods have been successfully applied to the analysis of spectrometric data.

Preprint of the Joint Institute for Nuclear Research

Dubna 1976

Введение

Хотя в математической литературе основное внимание уделяется преимущественно задачам разложения смесей распределений /обзор в /1/ /, задачи разложения регрессий на составные компоненты также являются чрезвычайно важными. Эти задачи играют большую роль в экспериментальной ядерной физике, где косвенный характер измерений приводит к тому, что, как правило, регистрируются не только изучаемые, но и всевозможные сопутствующие эффекты, а также различные случайные помехи измерения. Задачи анализа данных таких измерений формулируются в терминах нелинейного регрессионного анализа, и для их решения привлекается обычно техника построения оценок нелинейных параметров регрессий методом наименьших квадратов /м.н.к.-оценок/ /2-3/. Однако очень часто случается так, что даже при анализе выборок большого объема качество получаемых м.н.к.-оценок далеко не соответствует их высокой репутации в математической статистике. Можно вкратце указать на причины такого явления:

1. Систематическая нехватка априорных сведений о механизме изучаемых процессов /особенность большинства экспериментов ядерной физики/, а это ведет к тому, что модели компонент регрессий являются грубыми, т.е. не учитывают многих важных особенностей физического процесса.

2. Неудачная параметризация компонент. Даже если суммарная линия регрессии хорошо соответствует наблюдаемым значениям смеси, распределение компонент внутри такой регрессии может оказаться произвольным из-за сильных корреляций между параметрами.

В результате м.н.к.-оценки оказываются сильно смещенными, причем величина смещения этим методом не определяется и часто никак не контролируется.

Естественно, чисто математическим путем помочь экспериментатору строить точные модели невозможно. Теоретическая физика может оказать также лишь частичную помощь, ибо в любом случае физик-теоретик отразит в модели лишь то, что ему известно об изучаемом процессе. Здесь, однако, возникает идея использовать в качестве моделей реально измеренные физические процессы, которые наблюдались в других экспериментах в изолированном виде. Очевидно, в такую модель автоматически войдут все особенности данного процесса, как известные нам, так и неизвестные, если только он был измерен достаточно точно. Правда, реализация этой идеи сразу же наталкивается на трудность: как увязать величины, оценки которых нам нужны, с такой моделью. Иными словами, возникает проблема параметризации приближенной модели. Способам решения этой проблемы и анализу связанных с ними математических вопросов и посвящены последующие разделы данной статьи.

1. Свойства приближенных м.н.к.-оценок

Отметим предварительно некоторые особенности использования методов регрессионного анализа в экспериментальной физике. Итак, пусть задано множество точек измерения $\{x_j,\ j=1,\dots,M\}$ /схема, или дизайн, эксперимента/, в которых измерена величина $y(x_j)$, такая, что

$$y(x) = f(x,p) + r(x),$$

где $f(x,p)$ есть суперпозиция истинных физических процессов /как представляющих интерес, так и фоновых/, p – n–мерный вектор параметров, $r(x)$ - помехи измерения, т.е. случайные величины, независимые в разных x_j, имеющие нулевое математическое ожидание и дисперсию $D_1(x)$. Традиционно в регрессионном анализе под объемом выборки /т.е. ее информативностью/ понимается количество точек измерения M. Между тем, в практике современной экспериментальной физики преобладает другой взгляд на понятие информативности измерений, а именно: так как фактически измерение представляет собой сумму однотипных функций /событий/:

$$y(x) = Nf(x,p) + \sum_{i=1}^{N} r_i(x), \qquad /1.1/$$

то и количество информации /иначе, статистика/ считается функцией не количества точек измерения, а количества событий /т.е. N/. Такой взгляд имеет достаточно глубокие физические и математические основания:

1. Добавление дополнительных точек x, лежащих вне определенного интервала значений, лишено физического смысла; добавление же точек, лежащих внутри этого интервала, наталкивается на ограниченность разрешения регистрирующей аппаратуры. Использование приборов со сверхвысоким разрешением, как правило, выявляет тонкую структуру $f(x,p)$, что ведет к замене этой функции другой и, следовательно, переходу к другой задаче.

2. Современная математическая теория эксперимента на многочисленных примерах [4,5] продемонстрировала мысль, что отнюдь не все возможные точки измерения вносят существенный вклад в качество измерения: наибольшую роль играют при этом так называемые "критические" точки, в которых и рекомендуется проводить измерения.

3. Дисперсии м.н.к.-оценок параметров p_i регрессий $\sum_{i=1}^{n} p_i \phi_i(x)$ или $\sum_{i=1}^{n} p_i \frac{\partial f(x,p)}{\partial p_i}$ при $M\to\infty$ асимптотически пропорциональны величинам $1/\sum_{j=1}^{M} \frac{1}{D_i(x_j)}\phi_i^2(x_j)$ [6] или $1/\sum_{j=1}^{M} \frac{1}{D_i(x_j)}\left(\frac{\partial f}{\partial p_i}\right)^2$. Это ведет к парадоксу: функции с интегрируемым в области измерения квадратом /наиболее пригодные для практических целей/ оказываются недопустимыми, так как в этом случае оценки соответствующих параметров являются несостоятельными /их дисперсии не стремятся к нулю при $M\to\infty$/.

Так как в регрессионном анализе такие свойства м.н.к.-оценок, как состоятельность, несмещенность, минимальность дисперсий и нормальность распределения, изучены /6,7/ в условиях возможности неограниченного возрастания числа M, то следует выяснить, как скажется на них фиксация M и возможность варьирования величины N. Запишем предварительно /1.1/ в форме

$$s(x) = \frac{y(x)}{N} = f(x,p) + \frac{\Sigma r_i(x)}{N} = f(x,p) + e(x). \qquad /1.2/$$

По определению, м.н.к.-оценка вектора p_0 есть значения p, минимизирующие выражение

$$F(x,p,e(x)\} = \sum_{j=1}^{M} \frac{1}{D(x_j)} \{s(x_j) - f(x_j,p)\}^2, \qquad /1.3/$$

где $D(x_j) = D_1(x_j)/N$. Пусть впредь p_0 - истинное значение вектора p. Если зависимость f от x и p нелинейна, то вопрос о существовании хотя бы одной состоятельной м.н.к.-оценки является нетривиальным. Действительно, пусть, например, $f(x,p_0) = \exp(-\frac{x}{p_0})$, а $e(x) = a(x) - f(x,p_0)$, где $a(x)$ с ненулевой вероятностью принимает положительные значения. Так как вероятность регистрации, например единиц, во всех точках x_j не равна нулю, а единственный минимум /1.3/ в этом случае будет достигнут, очевидно, при $p=\infty$, то приходим к выводу: м.н.к.-оценка в данной задаче не имеет конечных математического ожидания и дисперсии; следовательно, даже в столь простой /и притом довольно распространенной в экспериментальной физике/ задаче не существует состоятельной м.н.к.-оценки. Но даже если минимум /1.3/ достигается лишь в конечных p /например, если $F(x,p,e(x))$ с вероятностью 1 строго выпукло/, это еще не означает существования состоятельной м.н.к.-оценки, так как точки минимумов F при различных $e(x)$ могут отстоять друг от друга сколь угодно далеко. Не гарантирует существования состоятельной м.н.к.-оценки и малость помехи $e(x)$, так как из сходимости с вероятностью 1 $F(x,p,e(x))$ к $F(x,p,0)$ в общем случае не вытекает сходимость точек минимумов $F(x,p,e(x))$ к p_0.

Для уверенного использования м.н.к.-оценок на практике все же желательно указать достаточные условия, гарантирующие существование состоятельной м.н.к.-оценки.

Рассмотрим линеаризацию выражения /1.3/ в точке p_0:

$$L\{x,p,e(x)\} = \sum_{j=1}^{M} \frac{1}{D(x_j)} \{s(x_j) - f(x_j,p_0) - \sum_{i=1}^{n} \frac{\partial f(x_j,p_0)}{\partial p_i}(p_i - p_{i0})\}^2 \qquad /1.4/$$

и точку p_ℓ ее минимума будем называть линеаризованной м.н.к.-оценкой. Пусть $f(x,p)$ непрерывно дифференцируемы по p_i до 2-го порядка включительно и либо функции $\frac{\partial f(x,p_0)}{\partial p_i}$ образуют систему Чебышева по x, либо существует дизайн $\{x_j\}$, на котором векторы $\frac{\partial f(x_j,p_0)}{\partial p_i}$ линейно независимы. Рассмотрим для данных $\epsilon > 0, N$ усечение распределения помехи $e(x)$:

$$e(x) = \begin{cases} e(x), & \text{если } |e(x)| < \frac{\epsilon}{N} \\ d, & \text{иначе; } d \in [0, \frac{\epsilon}{N}]. \end{cases}$$

Эта процедура соответствует таким действиям экспериментатора: прежде чем измерение анализируется, в нем корректируются те значения, которые слишком сильно отличаются от значений, ожидаемых в данных точках. Впредь будем иметь дело только с такими помехами. Усечение ведет к тому, что $\sup|e(x)| \to 0$ при $N \to \infty$.

Теорема 1. При любом $\epsilon>0$ для достаточно большого N существует выпуклая окрестность P точки p_0, такая, что /1.3/ с вероятностью 1 строго выпукло в данной окрестности, и значения p, обращающие в локальный минимум /1.3/ в P, являются состоятельной м.н.к.-оценкой параметра p_0. Если при этом линеаризованная м.н.к.-оценка имеет асимптотически /при $N \to \infty$/ нормальное распределение, то м.н.к.-оценка нелинейной задачи также будет при $N \to \infty$ иметь нормальное распределение с матрицей ковариаций линеаризованной м.н.к.-оценки.

Доказательство. Рассмотрим некоторую выпуклую окрестность Q точки p_0 и применим формулу Лагранжа к /1.3/:

$$F(x,p,e(x)) = F(x,p_0,e(x)) + (F'_p(x,p_0,e(x)),\Delta p) +$$

$$+ \frac{1}{2}(F''_{pp}(x,p_1,e(x))\Delta p,\Delta p),$$

где $\Delta p = p - p_0$, $\{p_1\} \in Q$.

Условие минимума формально имеет вид:

$$F'_p(x,p_0,e(x)) + (F''_{pp}(x,p_1,e(x)),\Delta p) = 0.$$

Отсюда имеем

$$\Delta p = -(F''_{pp}(x,p_1,e(x)))^{-1} F'_p(x,p_0,e(x)).$$

Пусть

$$A_{ij} = \sum_{t=1}^{M} \frac{1}{D(x_t)} \frac{\partial f(x_t,p_0)}{\partial p_i} \frac{\partial f(x_t,p_0)}{\partial p_j};$$

$$b_i = -\sum_{t=1}^{M} \frac{1}{D(x_t)} (s(x_t) - f(x_t,p_0)) \frac{\partial f(x_t,p_0)}{\partial p_i};$$

$$B_{ij} = \sum_{t=1}^{M} \frac{1}{D(x_t)} \{ (\frac{\partial f_1}{\partial p_i} + \frac{\partial f_1}{\partial p_j}) \sum_{k=1}^{n} \frac{\partial^2 f_2}{\partial p_i \partial p_k} \delta p_k +$$

$$+ (s(x_t) - f_0) \frac{\partial^2 f_1}{\partial p_i \partial p_j} + \sum_{k=1}^{n} \frac{\partial f_3}{\partial p_k} \delta p_k \frac{\partial^2 f_1}{\partial p_i \partial p_j} \}.$$

Здесь мы применили формулу Лагранжа к $f(x,p_1)$ и $\frac{\partial f(x,p_1)}{\partial p_i}$; $f_\ell = f(x_t,p_\ell)$. Тогда имеем

$$\Delta p_i = (A_{ij} + B_{ij})^{-1} b_j.$$

В силу усиленного закона больших чисел для последовательности $s(x) = \frac{y(x)}{N}$ при $N \to \infty$ $s(x)$ сходится с вероятностью 1 к $f(x,p_0)$.

Рассмотрим произведение $U_{ij} = A_{ij}^{-1} B_{ij}$. Так как $D(x) = D_1(x)/N$, то мы можем сократить элементы матриц A_{ij}^{-1} и B_{ij} на N, и тогда элементы матрицы A_{ij}^{-1} не будут зависеть от N. Будем теперь уменьшать диаметр окрестности Q и увеличивать N. Так как $\max_k |\delta p_k|$ будут уменьшаться, то с вероятностью 1 будут уменьшаться и элементы матрицы B_{ij}, а также элементы U_{ij}. Возьмем теперь достаточно малую окрестность Q и достаточно большое N, такие, что $\|U_{ij}\| < 1$. Последнее возможно в силу усеченности помехи $e(x)$. Мы имеем:

$$(A_{ij} + B_{ij})^{-1} = (A_{ij}(I_{ij} + A_{ij}^{-1} B_{ij}))^{-1} = (I_{ij} + U_{ij})^{-1} A_{ij}^{-1} =$$

$$= \sum_{n=0}^{\infty} (-1)^n U_{ij}^n A_{ij}^{-1}.$$

Отсюда $\Delta p = \sum_{n=0}^{\infty} (-1)^n U_{ij}^n A_{ij}^{-1} b = A_{ij}^{-1} b + \sum_{n=1}^{\infty} (-1)^n U_{ij}^n A_{ij}^{-1} b$.

Если элементы матрицы B_{ij} достаточно малы по сравнению с элементами A_{ij}, то матрица $(A_{ij} + B_{ij})$ будет положительно-определенной и, следовательно, значения Δp соответствуют локальному минимуму /1.3/. С вероятностью 1

$$\|\Delta p\| \le \|\sum_{n=0}^{\infty} (-1)^n U_{ij}^n\| \cdot \|A_{ij}^{-1}\| \cdot \|b\| = \frac{1}{1 + \|U_{ij}\|} \|A_{ij}^{-1}\| \cdot \|b\| \to 0$$

при $N \to \infty$.

Отсюда на основании теоремы Лебега о мажоранте математическое ожидание Δp стремится к нулю при $N \to \infty$, а это означает асимптотическую несмещенность м.н.к.-оценки p_0. Далее, так как с вероятностью 1 $(\Delta p_i)^2 \to 0$ при

$N\to\infty$, то аналогично получаем, что дисперсия Δp тоже стремится к нулю, т.е. нелинейная м.н.к.-оценка p_0 состоятельна. С вероятностью 1

$$||\sum_{n=1}^{\infty}(-1)^n U_{ij}^n A_{ij}^{-1} b|| \le || \sum_{n=1}^{\infty}(-1)^n U_{ij}^n||\cdot||A_{ij}^{-1}||\cdot||b|| =$$

$$= \frac{||U_{ij}||}{1+||U_{ij}||}||A_{ij}^{-1}||\cdot||b|| \to 0 \text{ при } N\to\infty .$$

Прнменнм теорему статистики /8/:
если $t_n = w_n + z_n$, прн этом w_n сходится по вероятности к w, а z_n - к нулю и распределение w непрерывно, то распределение t_n сходится к распределению w. Отсюда мы вндим, что Δp будет иметь при $N\to\infty$ нормальное распределение. Так как предел Δp с вероятностью 1 совпадает с линеаризованной м.н.к.-оценкой, то A_{ij}^{-1} будет в асимптотике матрицей ковариаций величин Δp_i. Теорема доказана.

Перейдем теперь к приближенным моделям. Мы будем рассматривать только линейные смеси, т.е.:

$$f(x,p) = \sum_{k=1}^{n} m_k(x, q_k),$$

где $m_i(x, q_k)$ в общем случае нелинейно зависит от параметров q_i /размерностью n_i/ и дважды непрерывно дифференцируема по q_i.

Итак, физическую модель $\phi(x,q_i)$ мы можем, очевидно, рассматривать как результат некоторого измерения, давшего

$$\phi_i(x,q_i) = m_i(x,q_i) + \delta_i(x), \qquad /1.5/$$

где случайные величины $\delta_i(x)$ независимы между собой и независимы в разных точках x, имеют нулевое математическое ожидание, статистики N_i и дисперсии $d_i(x)$. Как и выше, мы будем предполагать $\delta_i(x)$ усеченными. Подставим теперь функции ϕ_i в /1.3/ вместо m_i. Для простоты и краткости изложения ограничимся случаем

линейной зависимости $m_i(x,q)$ от q, поскольку мы видели, что для достаточно больших N и достаточно малой окрестности p_0 свойства нелинейной м.н.к.-оценки приближаются к свойствам линеаризованной. Итак,

$$s(x) = \sum_{i=1}^{n} p_i\phi_i(x) + e(x).$$

Формальное вырождение м.н.к.-оценки p_i имеет вид:

$$t_i = \Big(\sum_{k=1}^{M}\frac{1}{D(x_k)}\phi_i(x_k)\phi_j(x_k)\Big)^{-1}\Big(\sum_{k=1}^{M}\frac{1}{D(x_k)}s(x_k)\phi_j(x_k)\Big). \qquad /1.6/$$

Подставим /1.5/ в /1.6/ и введем обозначения:

$$A_{ij} = \sum_{k=1}^{M}\frac{1}{D(x_k)}m_i(x)m_j(x); \quad b_i = \sum_{k=1}^{M}\frac{1}{D(x_k)}s(x_k)m_j(x_k);$$

$$g_j = \sum_{k=1}^{M}\frac{1}{D(x_k)}s(x_k)\delta_j(x_k);$$

$$B_{ij} = \sum_{k=1}^{M}\frac{1}{D(x_k)}(m_i(x_k)\delta_j(x_k) + m_j(x_k)\delta_i(x_k) + \delta_i(x_k)\delta_j(x_k)).$$

Имеем $t_i = (A_{ij} + B_{ij})^{-1}(b_j + g_j)$.

Теорема 2. Пусть N_i - число событий в выборке $\{m_i(x)+\delta_i(x)\}$. Если $\min\limits_i N_i \to\infty$, оценки t_i сходятся с вероятностью 1 к обычным м.н.к.-оценкам p_i.
Доказательство. Так как $\max\limits_i \sup\limits_x|\delta_i(x)|\to 0$, то для достаточно большого $\min\limits_i N_i$ векторы $\{m_i(x_j) + \delta_i(x_j)\}$ линейно независимы и $||U_{ij}|| = ||A_{ij}^{-1}B_{ij}|| < 1$. Аналогично доказательству теоремы 1:

$$t_i = A_{ij}^{-1}b_j + \sum_{n=1}^{\infty}(-1)^n U_{ij}^n A_{ij}^{-1}(b_j + g_j) + A_{ij}^{-1}g_j .$$

Второй член сходится с вероятностью 1 к нулю. Так как с вероятностью 1 $s(x_k)$ ограничена, а $\sup_x|\delta_i(x)|$ для любого i стремится к нулю, то 3-й член тоже сходится к нулю. Теорема доказана.

2. Параметризация компонент смесей

Во избежание субъективизма при параметризации следует в качестве параметров указывать некоторые фундаментальные характеристики функций регрессий, по возможности одинаковые для всей смесей. Реальное измерение дает нам модель в виде $\phi_i(x)$, т.е. мы не знаем, как в нее входят параметры q_i. Так как эти параметры вводятся лишь с целью идентификации компонент в смеси, то мы можем рассматривать проблему параметризации компонент как проблему определения минимального числа геометрических характеристик функций, которые однозначно определяют их удельный вес в смеси. Мы будем считать, что параметры q_i функции $\phi_i(x, q_i)$ для модели имеют некоторые стандартные значения, а для реальных образов этой модели в смеси - какие-то другие значения. Тогда будем считать, что модель $\phi(x)$ и ее образ $s(x)$ связаны некоторым преобразованием T:

$$s(x) = T\phi(x).$$

Будем строить класс операторов $\{T\}$, руководствуясь лишь общелогическими рассуждениями о соотношениях моделей и их образов между собой. Модель и образ обладают свойствами /9/:

1. рефлексивности /любая модель (образ) есть модель (образ) самое (самого) себя/;

2. симметричности /модель и образ можно поменять местами/;

3. транзитивности /модель модели есть модель исходной модели, образ образа есть образ исходного образа/.

Эти свойства означают, что семейство $\{T\}$ должно обладать групповым свойством. Следуя работе /10/,

можно показать, что $\{T\}$ есть непрерывная группа Ли преобразований плоскости (x,y) на самое себя. Такие группы хорошо изучены, и мы можем воспользоваться ими. В данной работе мы ограничимся линейной и проективной группами. Унимодулярная группа является слишком узкой, а импримитивные группы ведут к громоздким выкладкам. Рассмотрим определяющие системы уравнений группы в конечной /алгебраической/ форме /10/. Для линейной группы мы имеем

$$y_2 = a_{11}y_1 + a_{12}x_1 + a_1; \quad x_2 = a_{21}y_1 + a_{22}x_1 + a_2,$$

где x_1, y_1 - старые, x_2, y_2 - новые координаты точки плоскости. Поскольку в нашем случае модель и образ - всегда явные однозначные функции одной переменной, мы можем пренебречь подгруппой вращений, откуда $a_{12} = a_{21} = 0$. Постоянным сдвигом по оси Y мы также можем пренебречь, что дает $a_1 = 0$. Итак,

$$y_2 = a_{11}y_1; \quad x_2 = a_{22}x_1 + a_2, \tag{/2.1/}$$

или $T\phi(x) = a_{11}\phi(a_{22}x + a_2)$. Обозначив $\frac{a_2}{a_{22}} = -P$, $\frac{1}{a_{22}} = W$, $a_{11} = A$, получим

$$T\phi(x) = A\phi\left(\frac{x-P}{W}\right). \tag{/2.2/}$$

Определим параметры, соответствующие линейной группе, а именно положение P, амплитуду A и ширину W, и будем записывать как модели, так и образы в форме $Ay(x, P, W)$. Определяющая система уравнений для проективной группы имеет вид:

$$y_2 = \frac{a_{11}y_1 + a_{12}x_1 + a_{13}}{b_{11}y_1 + b_{12}x_1 + 1}, \quad x_2 = \frac{a_{21}y_1 + a_{22}x_1 + a_{23}}{b_{21}y_1 + b_{22}x_1 + b_{23}}.$$

Рассуждая аналогично вышеприведенному, мы получим аналог /2.1/:

$$y_2 = a_{11}y_1 \ ; \quad x_2 = \frac{a_{22}x_1 + a_{23}}{b_{22}x_1 + b_{23}},$$

или $\quad T\phi(x) = a_{11}\phi(\frac{a_{22}x_1 + a_{23}}{b_{22}x + b_{23}})$. Обозначив $\frac{a_{23}}{a_{22}} = -P$, $\frac{b_{23}}{a_{22}} = W$, $\frac{b_{22}}{a_{22}} = k$, получим:

$$T\phi(x) = A\phi(\frac{x - P}{kx + W}) . \qquad /2.3/$$

Запись /2.3/ является более общим случаем записи /2.2/, и далее мы будем иметь дело только с ней. Введем запись $Ay(x,P,W,k)$. Тогда смесь мы будем определять как $s(x) = \sum_{i=1}^{n} A_i y_i(x, P_i, W_i, k_i)$ и под разложением смеси будем понимать определение величин A_i, P_i, W_i, k_i. Нормализуем приближенную модель $\phi(x)$ оператором

$$T_1 \phi(x) = \frac{1}{|\phi(x_0)|}\phi(\frac{x - x_0}{x_2 - x_1}), \qquad /2.4/$$

где x_0 - точка максимума, x_1 и x_2 - ближайшие слева и справа к x_0 точки, в которых $\phi(x_1) = \phi(x_2) = 1/2\,\phi(x_0)$. Преобразование /2.4/ принадлежит к семейству $\{T\}$, следовательно, $T_1\phi(x)$ есть модель для образов модели $\phi(x)$. Очевидно, $T_1\phi(x) = 1 \cdot y(x,0,1,0)$. На основе соотношений /2.2/ и /2.3/ можно легко вычислить производные приближенных моделей $y(x)$ по параметрам, необходимые для получения м.н.к.-оценок. Действительно,

$$\frac{ds}{\partial A} = \phi(z); \quad \frac{ds}{\partial P} = -\frac{A}{kx + W}\phi'_x(z); \qquad /2.5/$$

$$\frac{ds}{\partial W} = -\frac{Az}{kx + W}\phi'_x(z); \quad \frac{ds}{\partial k} = -\frac{Azx}{kx + W}\phi'_x(z),$$

где $\quad z = (x - P)/(kx + W)$.

Соотношения /2.5/ позволяют вычислять производные по параметрам через значения нормированной модели и ее производной и значения самих параметров. Аналогично вычисляются производные старших порядков.

3. Понятия чувствительности и разрешения и их влияние на корректность задачи разложения смесей

Немаловажной проблемой является также проблема единственности разложения и его устойчивости к помехам. Очевидно, допустимые смеси образуют только линейно независимые системы функций $\{Ay(x,P,W,k)\}$ и их производных по параметрам. Из /2.5/ видно, что необходимым условием для этого является линейная независимость функций $\phi(x), \phi'_x(x), x\phi'_x(x), x^2\phi'_x(x)$. Рассмотрим дифференциальное уравнение:

$$C_1\phi(x) + (C_2 + C_3x + C_4x^2)\phi'_x(x) = 0. \qquad /3.1/$$

Перебирая различные варианты ненулевых C_i, мы можем сразу указать недопустимые классы функций для анализа смесей: $\quad C\exp(ax)$, $(a+bx)^C$, $C\exp(b \cdot \mathrm{arctg}(\frac{x+a}{d}))$, $C(\frac{a+x}{a-x})^b$, где C,a,b,d - произвольные числа.

Примечание 1. Если исключить из функций компонент параметр P, анализ смесей экспонент становится возможным. Роль P в этом случае будет играть параметр W. В дальнейшем случай смесей экспонент также будет иметься в виду с учетом данного замечания, т.е. для экспонент $y(x) = A\phi(x,W)$. Рассмотрим теперь достаточные условия. Легко видеть, что функции $\phi, \phi'_x, x\phi'_x, x^2\phi'_x$ в общем не образуют системы Чебышева на X. Действительно, пусть $\phi(x)$ имеет на X не менее 2 экстремалей. Тогда линейная форма левой части /3.1/ может иметь не менее 5 нулей на X, а сумма n таких форм - не менее

5n – нулей. Рассмотрим следующий частный случай. Пусть функции $A_i y_i(x,P_i,W_i,k_i), i=1,...,n,$ таковы, что отрезок X можно разбить на n непересекающихся подмножеств X_i, таких, что $\phi_i(x)$ отлична от нуля лишь на i-ом подмножестве. В этом случае справедлива теорема:

Теорема 1. Существует дизайн $\{x_j\}$, на котором векторы $\phi(x_j), \phi'_x(x_j), x_j\phi'_x(x_j), x_j^2\phi'_x(x_j)$ линейно независимы.

Доказательство. Рассмотрим множество $\{x_j\}$ - объединение экстремальных точек функций $\phi_i(x)$ и их производных. Множество $\{x_j\}$ является искомым. Заметим, что число экстремальных точек у производной на 1 больше, чем у функции. Далее, векторы $\phi'_x(x_j)$, $x_j\phi'_x(x_j), x_j^2\phi'_x(x_j)$ имеют, по крайней мере, для трех j ненулевые компоненты, никакая линейная комбинация которых в нуль не обращается /так как степени x образуют систему Чебышева/; вектор $\phi(x_j)$ имеет ненулевые компоненты для тех j, в которых $\phi'_x(x_j), x_j\phi'_x(x_j)$, $x_j^2\phi'_x(x_j)$ равны нулю, и, следовательно, не может быть линейно выражен через них. Линейная независимость системы таких функций следует из того, что X_i не пересекаются между собой. Теорема доказана.

Примечание 2. На практике теорема будет справедлива и в тех случаях, когда $\phi_i(x)$ лишь несущественно отличны от нуля на X_j при $j \neq i$.

Рассмотрим теперь важнейшие случаи неединственности минимума функционала /1.3/ при построении м.н.к.-оценок.

1. Перемена индексов функций $\phi_i(x)$. Задача физически не изменяется, однако математически это означает преобразование координат пространства значений параметров R^{4n}. Следовательно, выражение /1.3/ имеет в R^{4n}, по крайней мере, n! минимумов, между которыми могут находиться локальные максимумы, седловины и т.д.

2. Более важный случай неединственности мы получим, если ограничимся частым случаем, когда функции $\phi_i(x)$ отличаются друг от друга лишь значениями параметров.

Теорема 2. Для любой $e(x)$ гиперповерхность в R^{4n}

$$F(x,p,e(x)) = \sum_{j=1}^{M} \frac{1}{D(x_j)} \{ s(x_j) - \sum_{i=1}^{n} A_i y(\frac{x_j - P_i}{k_i x_j + W_i}) \}^2$$ вырождается

в гиперцилиндр, если для какого-либо $i(j) A_i \to 0$ и /или/ $P_i \to P_j$.

Доказательство. Первое очевидно, поскольку при неограниченном уменьшении A_i изменения P_i, W_i, k_i не оказывают влияния на значения F. Вырождение становится полным на гиперплоскостях $A_i = 0$. Пусть теперь n=2, а P_1 и P_2 близко подходят друг к другу. Вдоль гиперплоскостей $k_i x + W_i = const$ разложим $\phi(x-P_i)$ в ряд Тейлора в окрестности истинных P_{i0}:

$$\phi(x-P_i) = \phi(x-P_{i0}) + (P_i - P_{i0})\phi'_x(x-P_{i0}) + R_i .$$

Тогда $F(x,p,e(x))$ имеет вид:

$$F = \sum_{j=1}^{M} \frac{1}{D(x_j)} \{ e(x_j) + (A_1 - A_{10})c_1 + (A_2 - A_{20})c_2 + A_1 \Delta P_1 c_3 + A_2 \Delta P_2 c_4 + R_1 + R_2 \}^2 .$$

Если $\|R_1 + R_2\| \ll D(x)$, то при $|P_2 - P_1| \to 0$ F локально превращается в поверхность почти постоянного уровня вдоль гиперплоскостей $b_1 A_1 + b_2 A_2 = const$. Вырождение становится почти полным при $P_1 = P_2$.

Примечание 3. Практически теорема будет справедлива и при некоторых различиях в виде функций $\phi_i(x)$.

Примечание 4. Если $P_1 \to P_2$, более надежно оценивается сумма $A_1 + A_2$, а не A_i в отдельности. Так как F при этом имеет почти одинаковое распределение для любой комбинации A_i, дающей в сумме некоторое число, то отличить хорошие оценки A_i от плохих методами проверки гипотез с помощью χ^2-критерия практически нет никаких шансов.

Следствие. Задачу минимизации /1.3/ следует рассматривать не на всем R^{4n}, а на симплексе, задаваемом неравенствами: $P_{i+1} > P_i$, $A_i > 0$ /для экспонент $W_{i+1} > W_i$/.

Рассмотрим для вышеупомянутых смесей вопрос об устойчивости оценок по отношению к усеченной помехе e(x). Из рассмотрения рельефа $F(x,p,e(x))$ следует, что угроза выхода на несостоятельную м.н.к.-оценку существует в двух ситуациях: $A_i \to 0$ и $P_i \to P_j$.

Теорема 3. Для устойчивости м.н.к.-оценок в малой окрестности p_0 необходимо существование таких чисел $s_i > 0$ и $r_i > 0$, согласованных с распределением $e(x)$, что с вероятностью 1 м.н.к.-оценки A_i и P_i удовлетворяют условию:

$$A_i \geq s_i; \quad P_{i+1} \geq P_i + r_i. \qquad /3.2/$$

Доказательство. Доказательство очевидно.

Примечание 5. В случае линейной задачи /P_i, W_i, k_i - фиксированы/ для устойчивости м.н.к.-оценки амплитуд необходимо и достаточно, чтобы $A_{i0} \geq s_i$; $P_{i+1,0} \geq P_{i0} + r_i$.

Разумеется, для произвольной помехи $e(x)$ таких чисел /s_i и r_i/ может не существовать, и задача анализа смеси окажется в этом случае некорректно поставленной. Величины s_i и r_i имеют прозрачный физический смысл: это чувствительность и разрешение нашего метода /аналогичные чувствительности и разрешению приборов/. То, что они не равны нулю, отражает очевидный факт, что при наличии помех чувствительность и разрешение метода /как и приборов/ не являются идеальными. Некорректность, таким образом, возникает тогда, когда мы пытаемся разделить смесь, для разделения которой у нашего метода не хватает чувствительности и /или/ разрешения.

Методы, изложенные в работе, были реализованы в программе UPEAK, написанной на языке ФОРТРАН-4 и ориентированной для использования на ЭВМ типа БЭСМ-6. Программа была успешно применена к анализу данных спектрометрических измерений. Модели компонент измерялись экспериментально с большой статистикой и вводились в программу в виде гистограмм.

В заключение автор выражает благодарность проф. Н.Н.Говоруну за поддержку в работе и проф. Е.П.Жидкову за полезные обсуждения.

Литература

1. *А.В.Миленький. Классификация сигналов в условиях неопределенности. М., "Советское радио", 1975.*
2. *V.B.Zlokazov. Method for Processing Discrete Energy Spectra with Complex Peak Shape. Nucl. Instr. and Meth., 1975, 130 (2), p. 543-49.*
3. *В.Б.Злоказов, Л.П.Кулькина, О.Д.Маслов. Инструментальный нейтронно-активационный анализ геологических и биологических объектов с использованием ЭВМ. АЭ, 1975, 39, 286-88.*
4. *В.В.Налимов. Теория эксперимента. М., "Наука", 1971.*
5. *В.В.Федоров. Теория оптимальных экспериментов. М., "Наука", 1971.*
6. *E.Malinvaud. The Consistency of Nonlinear Regressions. Ann. Math. Statist., 1970, 41 (3), 956-969.*
7. *R.I.Jennrich. Asymptotic Properties of Non-Linear Least Squares Estimators. Ann.Math.Statist., 1969, 40(2), p. 633-643.*
8. *С.В.Нагаев. Математическая статистика. Новосибирск, НГУ, 1973, с. 133.*
9. *Чжао Юань-жень. Модели в лингвистике и модели вообще. В сб. Математическая логика и ее применение. М., 1965, с. 281-292.*
10. *В.С.Файн. Распознавание изображений. М., "Наука", 1970.*

Рукопись поступила в издательский отдел
25 октября 1976 года.