УДК 517.5

О ВЫБОРЕ ВЕСОВЫХ КОЭФФИЦИЕНТОВ В МЕТОДЕ НАИМЕНЬШИХ КВАДРАТОВ

В.А. Горелик, Н.В. Самсонова

1. Введение

Решение многих научных и практических задач связано с математическим представлением результатов экспериментов - получении тех или иных аналитических выражений, пригодных для изучения, в частности, оптимизации и прогнозирования рассматриваемого процесса. Любая задача, в которой исходная функция задается таблицей своих значений, требует применения того или иного метода аппроксимаций. Наиболее часто в качестве критерия аппроксимации выбирают метод наименьших квадратов.

История развития метода наименьших квадратов началась в 1806 г. с работы французского математика А. М. Лежандра, посвященной вычислению кометных орбит [9], где он впервые ввел понятие и изложил метод наименьших квадратов.

В 1809 г. К. Ф. Гаусс [2] дал первое вероятностное обоснование метода наименьших квадратов, а в 1810 г. он же [3] разработал вычислительную сторону метода для обработки неравноценных наблюдательных данных. В 1821 г. вышла в свет статья К. Ф. Гаусса с изложением этого метода применительно к топографическим задачам [4].

В своих работах К. Ф. Гаусс применял вероятностный подход к обоснованию метода наименьших квадратов. В статье [2] он излагал его как способ оценки параметров по критерию максимизации правдоподобия в современной терминологии. В работе [4], критикуя избранный им ранее путь обоснования и опубликованные к тому времени исследования П. Л. Лапласа по методу наименьших квадратов, К. Ф. Гаусс привел обоснование этого метода в форме минимизации дисперсий. Здесь впервые он рассматривает весовые коэффициенты как величины, обратные дисперсиям.

В 1812 г. П. Л. Лаплас в фундаментальном трактате по теории вероятности [8] получил ряд важных результатов и применил их к методу наименьших квадратов, что позволило находить наивероятнейшие значения измеренных величин и степень достоверности этих расчетов.

Дальнейшие важные результаты были получены в теории метода наименьших квадратов в 1859 г. П. Л. Чебышевым [15], разработавшим теорию интерполирования по методу наименьших квадратов с помощью ортогональных полиномов, носящих его имя.

А. А. Марков в 1898 г. в работах [12], [13] внес в математическую статистику ряд весьма важных идей, прояснивших суть метода наименьших квадратов, и дал, в частности, общепринятое в настоящее время вероятностное обоснование метода.

После работ Маркова А. А. с двадцатых годов нынешнего века метод наименьших квадратов включился в математическую статистику как важная и естественная часть оценивания параметров.

В этой связи ряд интересных и важных результатов получен Ю. Нейманом и Ф. Дэвид, А. Эйткеном [16], С. Рао [14].

В 1946 г. А. И. Колмогоровым [7] дано геометрической изложение метода наименьших квадратов.

На сегодняшний день метод наименьших квадратов в разных научных дисциплинах называется по-разному. Например, математики могут интерпретировать его как метод отыскания для заданной точки функционального пространства ближайшей точки в заданном подпространстве. Специалисты по численному анализу используют этот подход в вопросе об ошибках входной информации. Статистики вводят в свою постановку задачи вероятностные распределения и используют для описания этой области термины регрессионного анализа. Инженеры приходят к нему, занимаясь такими предметами как оценивание параметров или фильтрация.

Главное состоит в следующем: когда эти задачи (сформулированные в любом из заданных контекстов) достигают стадии конкретных расчетов, они содержат в себе одну и ту же центральную проблему, которую можно сформулировать следующим образом.

Обозначим через φ_i - квадрат отклонения в i-ой точке:

$$\varphi_i = (y_i - f(x_i, a_1, a_2, \ldots, a_m))^2 . \tag{1}$$

Согласно методу наименьших квадратов, в качестве оценок неизвестных параметров берутся такие их значения, при которых сумма квадратов отклонений экспериментальных значений y_i от соответствующих "теоретических" $f(x_i, a)$ с учетом весовых коэффициентов будет наименьшей:

$$\Phi = \sum_{i=1}^{n} \lambda_i \varphi_i \to \min . \tag{2}$$

Рассматривая Φ как функцию от $a_1, a_2, \ldots, a_m$, необходимое условие ее экстремума можно записать в виде

$$\frac{\partial \Phi}{\partial a_k} = 0, \qquad k = 1, \ldots, m. \tag{3}$$

Решая полученную систему из m уравнений относительно m неизвестных параметров, можно найти искомые оценки.

Аналогично выглядит метод наименьших квадратов и в случае, когда x-вектор. При этом x_i в формуле (1) следует заменить набором x_{i1}, x_{i2}, $\ldots, x_{is}$, где x_{ij} обозначает значение j-го аргумента функции $y(x_1, x_2, \ldots, x_s)$ в i-м эксперименте.

Систему уравнений (3) называют нормальной системой метода наименьших квадратов.

На практике наиболее часто применяются функции

$$y(x) = a_0 + \sum_{i=1}^{n} a_i x^i, \qquad y(x) = \exp(a_0 + \sum_{i=1}^{n} a_i x^i).$$

Параметры этих функций имеют качественный экономический смысл и легко интерпретируются.

Например, уравнение прямой $y(x) = a_0 + a_1 x$ (полином первой степени) характеризует постоянный прирост, равный a_1 единицам при начальном уровне a_0.

В уравнении параболы второго порядка $y(x) = a_0 + a_1 x + a_2 x^2$ (полином второй степени) - коэффициент a_1 выражает начальную скорость роста, а коэффициент a_2 - постоянную скорость изменения прироста (величина ускорения в среднем за изучаемый период равна $2a_2$ единицам).

Экспонента $y(x) = \exp(a_0 + a_1 x)$ отражает постоянный относительный прирост, равный e^{a_1} единицам, а $y(x) = \exp(a_0 + a_1 x + a_2 x^2)$ характеризует постоянный относительный прирост, равный e^{2a_2} единицам (производные соответствующих логарифмов).

Оценки параметров в модели $\exp(a_0 + \sum_{i=1}^{n} a_i x^i)$ находят путем применения метода наименьших квадратов к логарифмам исходных данных.

2. Метод наименьших квадратов с весовыми коэффициентами

Допустим, $f(x)$ представляет собой линейную относительно аргумента x функцию (полином первой степени):

$$f(x) = ax + b.$$

В результате произведенных экспериментов получены n пар чисел

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n).$$

Таким образом, функция $y(x)$ определена таблицей своих значений y_i при заданных значениях аргументов x_i $(i = \overline{1, n})$. Значения аргумента упорядочены по возрастанию: $x_1 < x_2 < \ldots < x_n$.

Учитывая это, задачу метода наименьших квадратов переформулируем следующим образом: аппроксимировать исходную табличную функцию $y(x)$ линейной функцией $f(x) = ax + b$ исходя из условия минимума суммы квадратов отклонений y_i от $f(x_i)$

$$\Phi = \sum_{i=1}^{n} \lambda_i \varphi_i \to \min, \tag{4}$$

где $\varphi_i = (y_i - ax_i - b)^2$, а λ_i - положительные весовые коэффициенты, выбираемые из эмпирических соображений.

Тогда оценками параметров a и b будут служить такие их значения, при которых Φ принимает наименьшее значение.

Необходимое условие экстремума Φ как функции, зависящей от a и b, можно записать в виде:

$$\begin{aligned} \frac{\partial \Phi}{\partial a} &= 0, \\ \frac{\partial \Phi}{\partial b} &= 0. \end{aligned} \tag{5}$$

Дифференцируя Φ по a и b, получим,

$$\begin{aligned} \frac{\partial \Phi}{\partial a} &= \sum_{i=1}^{n} 2\lambda_i (y_i - a x_i - b)(-x_i) = 2\sum_{i=1}^{n} \lambda_i(-y_i x_i + a x_i^2 + b x_i), \\ \frac{\partial \Phi}{\partial b} &= \sum_{i=1}^{n} 2\lambda_i (y_i - a x_i - b)(-1) = 2\sum_{i=1}^{n} \lambda_i(-y_i + a x_i + b). \end{aligned}$$

Введя обозначения

$$S_1 = \sum_{i=1}^{n} \lambda_i x_i, \qquad S_2 = \sum_{i=1}^{n} \lambda_i x_i^2, \qquad \Lambda = \sum_{i=1}^{n} \lambda_i,$$

$$V_0 = \sum_{i=1}^{n} \lambda_i y_i, \qquad V_1 = \sum_{i=1}^{n} \lambda_i x_i y_i,$$

можно записать

$$\begin{aligned} \frac{\partial \Phi}{\partial a} &= 2(-V_1 + aS_2 + bS_1), \\ \frac{\partial \Phi}{\partial b} &= 2(-V_0 + aS_1 + b\Lambda), \end{aligned}$$

и система (5) принимает вид

$$\begin{cases} aS_2 + bS_1 = V_1, \\ aS_1 + b\Lambda = V_0. \end{cases} \tag{6}$$

Определитель этой системы

$$\Delta = \Lambda S_2 - S_1^2 \tag{7}$$

можно представить в виде

$$\Delta = \Lambda \sum_{i=1}^{n} \lambda_i (x_i - \bar{x})^2, \tag{8}$$

где

$$\bar{x} = \frac{S_1}{\Lambda} = \frac{\sum \lambda_i x_i}{\sum \lambda_i}.$$

В самом деле,

$$\sum_{i=1}^{n} \lambda_i (x_i - \bar{x})^2 = \sum_{i=1}^{n} \lambda_i (x_i^2 - 2x_i\bar{x} + \bar{x}^2) = \sum_{i=1}^{n} \lambda_i x_i^2 - 2\bar{x} \sum_{i=1}^{n} \lambda_i x_i + \bar{x}^2 \sum_{i=1}^{n} \lambda_i =$$

$$= S_2 - 2\frac{S_1}{\Lambda} S_1 + \left(\frac{S_1}{\Lambda}\right)^2 \Lambda = S_2 - \frac{S_1^2}{\Lambda},$$

откуда и следует (8).

Таким образом, исключая случай, когда все $x_i = \bar{x}$, можно утверждать, что $\Delta \neq 0$, и, значит, система (6) имеет единственное решение. Обозначив это решение через $(\tilde{a}, \tilde{b})$, имеем

$$\tilde{a} = \frac{\Lambda V_1 - S_1 V_0}{\Delta}, \qquad \tilde{b} = \frac{S_2 V_0 - S_1 V_1}{\Delta}. \tag{9}$$

Покажем, что при этих значениях a и b сумма (4) принимает наименьшее значение, и найдем его.

Представим Φ в виде

$$\Phi = \sum_{i=1}^{n} \lambda_i \big(y_i - a(x_i - \bar{x}) - a\bar{x} - b\big)^2 = \sum_{i=1}^{n} \lambda_i \big(y_i - a(x_i - \bar{x}) - \beta\big)^2,$$

где $\beta = a\bar{x} + b$.

Тогда

$$\Phi = \sum_{i=1}^{n} \lambda_i (y_i^2 + a^2(x_i - \bar{x})^2 + \beta^2 - 2ay_i(x_i - \bar{x}) - 2\beta y_i + 2a\beta(x_i - \bar{x})).$$

Заметив, что

$$\sum_{i=1}^{n} \lambda_i (x_i - \bar{x}) = \sum_{i=1}^{n} \lambda_i x_i - \bar{x} \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} \lambda_i x_i - \sum_{i=1}^{n} \lambda_i x_i = 0,$$

$$\sum_{i=1}^{n} \lambda_i y_i (x_i - \bar{x}) = \sum_{i=1}^{n} \lambda_i x_i y_i - \bar{x} \sum_{i=1}^{n} \lambda_i y_i = V_1 - \frac{S_1}{\Lambda} V_0 = \frac{\Lambda V_1 - S_1 V_0}{\Lambda} = \frac{\Delta}{\Lambda} \tilde{a},$$

$$\sum_{i=1}^{n} \lambda_i (x_i - \bar{x})^2 = \frac{\Delta}{\Lambda},$$

получим

$$\Phi = \sum_{i=1}^{n} \lambda_i y_i^2 + a^2 \sum_{i=1}^{n} \lambda_i (x_i - \bar{x})^2 + \beta^2 \sum_{i=1}^{n} \lambda_i \; - \; 2a \sum_{i=1}^{n} \lambda_i y_i (x_i - \bar{x}) -$$

$$-2\beta \sum_{i=1}^{n} \lambda_i y_i \; + \; 2a\beta \sum_{i=1}^{n} \lambda_i (x_i - \bar{x}) =$$

$$= \sum_{i=1}^{n} \lambda_i y_i^2 + a^2 \frac{\Delta}{\Lambda} + \Lambda \beta^2 - 2a\tilde{a} \frac{\Delta}{\Lambda} - 2\beta V_0 =$$

$$= \sum_{i=1}^{n} \lambda_i y_i^2 + \frac{\Delta}{\Lambda} (a - \tilde{a})^2 - \frac{\Delta}{\Lambda} \tilde{a}^2 + \Lambda \left(\beta - \frac{V_0}{\Lambda} \right)^2 - \frac{V_0^2}{\Lambda}.$$

Минимальное значение этой суммы достигается, когда

$$a = \tilde{a}, \qquad \beta = \frac{V_0}{\Lambda},$$

а это, в свою очередь, равносильно

$$a = \tilde{a}, \qquad a\bar{x} + b = \frac{V_0}{\Lambda},$$

то есть

$$b = \frac{V_0}{\Lambda} - \tilde{a}\bar{x}.$$

Преобразуя это равенство, получим

$$b = \frac{V_0}{\Lambda} - \frac{\Lambda V_1 - S_1 V_0}{\Delta} \frac{S_1}{\Lambda} = \frac{\Delta V_0 - \Lambda S_1 V_1 + S_1^2 V_0}{\Delta \Lambda} =$$

$$= \frac{\Lambda S_2 V_0 - S_1^2 V_0 - \Lambda S_1 V_1 + S_1^2 V_0}{\Delta \Lambda} = \frac{S_2 V_0 - S_1 V_1}{\Delta} = \tilde{b}$$

Таким образом, в самом деле, при $a = \tilde{a}$, $b = \tilde{b}$, определяемых формулами (9), величина Φ принимает минимальное значение, равное

$$\Phi_{\min} = \sum_{i=1}^{n} \lambda_i y_i^2 - \frac{V_0^2}{\Lambda} - \frac{(\Lambda V_1 - S_1 V_0)^2}{\Lambda S_2 - S_1^2}.$$

Хотя метод наименьших квадратов имеет богатую историю и этому методу уделено большое внимание в научной литературе, нельзя считать, что он исчерпывающе изучен. В частности, мало исследовано влияние на получаемые решения весовых коэффициентов при квадратах отклонений.

В назначении весовых коэффициентов, используемых во многих задачах в качестве характеристик точности или информативности исходных данных, часто есть выбор. Поэтому полезно знать, к чему может приводить варьирование численными значениями весовых коэффициентов, насколько широко могут изменяться решения в результате такого варьирования.

Определение значений параметров λ_i $(i = \overline{1, n})$ основывается на эмпирических соображениях, но так как λ_i отражают лишь соотношения весов, то целесообразно нормировать коэффициенты таким образом, чтобы сумма всех λ_i была равна единице:

$$\Lambda = \sum_{i=1}^{n} \lambda_i = 1. \tag{10}$$

Тогда формулы для определителя системы (5) и "оптимальных параметров" функции $f(x)$ (9) представимы в виде

$$\Delta = S_2 - S_1^2, \quad \tilde{a} = \frac{V_1 - S_1 V_0}{\Delta}, \quad \tilde{b} = \frac{S_2 V_0 - S_1 V_1}{\Delta}$$

(очевидно, что при нормировке значения $\tilde{a}$ и $\tilde{b}$ не меняются).

Значение минимума функции Φ будет равно

$$\Phi_{\min} = \sum_{i=1}^{n} \lambda_i y_i^2 - V_0^2 - \frac{(V_1 - S_1 V_0)^2}{S_2 - S_1^2}$$

(при нормировке оно умножается на величину, обратную сумме коэффициентов λ_i).

3. Выбор весовых коэффициентов

Рассмотрим несколько вариантов выбора весовых коэффициентов.

Случай 1. Пусть λ_i $(i = \overline{1,n})$ есть вероятность того, что в результате проведенного эксперимента получили пару (x_i, y_i).

1.1. Заметим, что "классический вариант метода наименьших квадратов", предполагающий равнозначность полученных исходных данных (x_i, y_i), когда

$$\lambda_i = 1, \qquad i = \overline{1,n}, \tag{11}$$

является частным случаем Случая 1 после нормирования весовых коэффициентов согласно условию (10).

Итак, значения x_i равнозначны, то есть невозможно отдать предпочтение какому-либо исходу или по причине недостаточной информации, или по причине равновероятности возможных исходов.

$$\lambda_i = \frac{1}{n}, \qquad i = \overline{1,n}.$$

$$S_1 = \sum_{i=1}^{n} \lambda_i x_i = \frac{1}{n} \sum_{i=1}^{n} x_i, \quad S_2 = \sum_{i=1}^{n} \lambda_i x_i^2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2,$$

$$V_0 = \sum_{i=1}^{n} \lambda_i y_i = \frac{1}{n} \sum_{i=1}^{n} y_i, \quad V_1 = \sum_{i=1}^{n} \lambda_i x_i y_i = \frac{1}{n} \sum_{i=1}^{n} x_i y_i.$$

Определитель системы для данного случая можно записать в виде:

$$\Delta = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{1}{n^2}(\sum_{i=1}^{n} x_i)^2.$$

Неизвестные параметры аппроксимационной функции $f(x)$ вычислим по формулам:

$$\tilde{a} = \frac{V_1 - S_1V_0}{\Delta} = \frac{n\sum x_iy_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2},$$
$$\tilde{b} = \frac{S_2V_0 - S_1V_1}{\Delta} = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_iy_i}{n\sum x_i^2 - (\sum x_i)^2}.$$

(суммирование проводится по параметру i, $i = \overline{1, n}$).

Значение минимума функции Φ будет равно:

$$\Phi_{\min} = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \frac{(\sum x_iy_i)^2 - \frac{(n-1)}{n}\sum x_i \sum y_i \sum x_iy_i - \frac{1}{n}\sum x_i^2 (\sum y_i)^2}{n\sum x_i^2 - (\sum x_i)^2}.$$

1.2. Пусть k из исходов (x_i, y_i), $i = \overline{1, n}$, равновероятностны:

$$\begin{gathered} I = \{1, 2, \ldots, n\}, \;\; I = I_1 \cup I_2, \;\; \mid I_1 \mid = k; \\ \lambda_i = 1/(k+1) \;\text{ если } i \in I_1, \;\; \textstyle\sum_{i \in I_2} \lambda_i = 1/(k+1). \end{gathered} \tag{12}$$

$$S_1 = \frac{1}{k+1}\sum_{i \in I_1} x_i + \sum_{i \in I_2} \lambda_i x_i, \quad S_2 = \frac{1}{k+1}\sum_{i \in I_1} x_i^2 + \sum_{i \in I_2} \lambda_i x_i^2,$$

$$V_0 = \frac{1}{k+1}\sum_{i \in I_1} y_i + \sum_{i \in I_2} \lambda_i y_i, \quad V_1 = \frac{1}{k+1}\sum_{i \in I_1} x_iy_i + \sum_{i \in I_2} \lambda_i x_iy_i,$$

$$\Delta = \frac{1}{k+1}\sum_{i \in I_1} x_i^2 + \sum_{i \in I_2} \lambda_i x_i^2 - \Big(\frac{1}{k+1}\sum_{i \in I_1} x_i + \sum_{i \in I_2} \lambda_i x_i\Big)^2 =$$

$$= \frac{(k+1)[\sum_{i \in I_1} x_i^2 + (k+1)\sum_{i \in I_2} \lambda_i x_i^2] - [\sum_{i \in I_1} x_i + (k+1)\sum_{i \in I_2} \lambda_i x_i]^2}{(k+1)^2}.$$

Параметры a и b функции $f(x)$ можно представить в виде:

$$\tilde{a} = \frac{V_1 - S_1V_0}{\Delta} =$$

$$= \frac{(k+1)[\sum_{i\in I_1} x_i y_i - \sum_{i\in I_1} x_i \sum_{i\in I_2} \lambda_i y_i - \sum_{i\in I_2} \lambda_i x_i \sum_{i\in I_1} y_i]+}{(k+1)[\sum_{i\in I_1} x_i^2 + (k+1)\sum_{i\in I_2} \lambda_i x_i^2]-}$$

$$\frac{+(k+1)^2[\sum_{i\in I_2} \lambda_i x_i y_i - \sum_{i\in I_2} \lambda_i x_i \sum_{i\in I_2} \lambda_i y_i] - \sum_{i\in I_1} x_i \sum_{i\in I_1} y_i}{-[\sum_{i\in I_1} x_i + (k+1)\sum_{i\in I_2} \lambda_i x_i]^2},$$

$$\tilde{b} = \frac{S_2V_0 - S_1V_1}{\Delta} =$$

$$= \frac{\sum_{i\in I_1} x_i^2 \sum_{i\in I_1} y_i - \sum_{i\in I_1} x_i \sum_{i\in I_1} x_i y_i + (k+1)[\sum_{i\in I_1} x_i^2 \sum_{i\in I_2} \lambda_i y_i +}{(k+1)\sum_{i\in I_1} x_i^2 +}$$

$$\frac{+\sum_{i\in I_2} \lambda_i x_i^2 \sum_{i\in I_1} y_i - \sum_{i\in I_1} x_i \sum_{i\in I_2} \lambda_i x_i y_i - \sum_{i\in I_2} \lambda_i x_i \sum_{i\in I_1} x_i y_i]+}{+(k+1)^2 \sum_{i\in I_2} \lambda_i x_i^2 -}$$

$$\frac{+(k+1)^2[\sum_{i\in I_2} \lambda_i x_i^2 \sum_{i\in I_2} \lambda_i y_i - \sum_{i\in I_2} \lambda_i x_i \sum_{i\in I_2} \lambda_i x_i y_i]}{-[\sum_{i\in I_1} x_i + (k+1)\sum_{i\in I_2} \lambda_i x_i]^2}.$$

Минимальное значение функции Φ следовательно будет равно:

$$\Phi_{\min} = \sum_{i=1}^{n} \lambda_i y_i^2 - V_0^2 - \Delta \tilde{a}^2 = \sum_{i\in I_1} y_i^2 + \sum_{i\in I_2} \lambda_i y_i^2 - (\frac{1}{k+1}\sum_{i\in I_1} y_i - \sum_{i\in I_2} \lambda_i y_i)^2 -$$

$$-\frac{(k+1)^2[\sum_{i\in I_2} \lambda_i x_i y_i - \sum_{i\in I_2} \lambda_i x_i \sum_{i\in I_2} \lambda_i y_i]+}{(k+1)\sum_{i\in I_1} x_i^2 +}$$

$$\frac{+(k+1)[\sum_{i\in I_1} x_i y_i - \sum_{i\in I_1} x_i \sum_{i\in I_2} \lambda_i y_i - \sum_{i\in I_1} y_i \sum_{i\in I_2} \lambda_i x_i]-}{+(k+1)^2 \sum_{i\in I_2} \lambda_i x_i^2 -}$$

$$\frac{-\sum_{i\in I_1} x_i \sum_{i\in I_1} y_i}{-[\sum_{i\in I_1} x_i + (k+1)\sum_{i\in I_2} \lambda_i x_i]^2}.$$

Случай 2. Можно предложить вычислять λ_i по формуле:

$$\lambda_i = \frac{2i}{n(n+1)}, \quad (i = \overline{1, n}). \tag{13}$$

Тогда более "поздним наблюдениям" соответствуют большие веса.

Заметим, что условие нормировки (10) выполнено.

$$\Lambda = \sum_{i=1}^{n} \lambda_i = \frac{2}{n(n+1)} \sum_{i=1}^{n} i = \frac{2}{n(n+1)} \frac{n(n+1)}{2} = 1.$$

$$S_1 = \sum_{i=1}^{n} \lambda_i x_i = \frac{2}{n(n+1)} \sum_{i=1}^{n} i x_i, \quad S_2 = \sum_{i=1}^{n} \lambda_i x_i^2 = \frac{2}{n(n+1)} \sum_{i=1}^{n} i x_i^2,$$

$$V_0 = \sum_{i=1}^{n} \lambda_i y_i = \frac{2}{n(n+1)} \sum_{i=1}^{n} i y_i, \quad V_1 = \sum_{i=1}^{n} \lambda_i x_i y_i = \frac{2}{n(n+1)} \sum_{i=1}^{n} i x_i y_i.$$

Определитель равен

$$\Delta = S_2 - S_1^2 = \frac{2}{n(n+1)} \sum_{i=1}^{n} i x_i^2 - \Big(\frac{2}{n(n+1)} \sum_{i=1}^{n} i x_i\Big)^2 =$$

$$= \frac{2}{n(n+1)} \Big[\sum_{i=1}^{n} i x_i^2 - \frac{2}{n(n+1)} \Big(\sum_{i=1}^{n} i x_i\Big)^2 \Big].$$

Неизвестные параметры функции $f(x)$ вычислим следующим образом:

$$\tilde{a} = \frac{V_1 - S_1 V_0}{\Delta} = \frac{n(n+1) \sum i x_i y_i - 2 \sum i x_i \sum i y_i}{n(n+1) \sum i x_i^2 - 2 (\sum i x_i)^2},$$

$$\tilde{b} = \frac{S_2 V_0 - S_1 V_1}{\Delta} = \frac{2 \sum i x_i^2 \sum i y_i - 2 \sum i x_i \sum i x_i y_i}{n(n+1) \sum i x_i^2 - 2 (\sum i x_i)^2}$$

(суммирование производится по параметру i, $i = \overline{1, n}$).

Минимум функции Φ равен

$$\begin{aligned}
\Phi_{\min} &= \sum_{i=1}^{n} \lambda_i y_i^2 - V_1 \tilde{a} - V_0 \tilde{b} = \\
&= \frac{2}{n(n+1)} \sum_{i=1}^{n} i y_i^2 - \frac{2\tilde{a}}{n(n+1)} \sum_{i=1}^{n} i x_i y_i - \frac{2\tilde{b}}{n(n+1)} \sum_{i=1}^{n} i y_i = \\
&= \frac{2}{n(n+1)} \Big[\sum_{i=1}^{n} i y_i^2 - \tilde{a} \sum_{i=1}^{n} i x_i y_i - \tilde{b} \sum_{i=1}^{n} i y_i \Big].
\end{aligned}$$

Случай 3. Пусть λ_i выражает процентное содержание x_i в совокупности:

$$\lambda_i = \frac{x_i}{\sum x_i}, \quad i = \overline{1, n}. \tag{14}$$

Введем обозначение:

$$X = \sum_{i=1}^{n} x_i.$$

Отметим, что условие нормировки выполнено

$$\Lambda = \sum_{i=1}^{n} \frac{x_i}{\sum x_i} = 1.$$

Промежуточные формулы можно переписать следующим образом

$$S_1 = \sum_{i=1}^{n} \lambda_i x_i = \frac{1}{X}\sum_{i=1}^{n} x_i^2, \quad S_2 = \sum_{i=1}^{n} \lambda_i x_i^2 = \frac{1}{X}\sum_{i=1}^{n} x_i^3,$$

$$V_0 = \sum_{i=1}^{n} \lambda_i y_i = \frac{1}{X}\sum_{i=1}^{n} x_i y_i, \quad V_1 = \sum_{i=1}^{n} \lambda_i x_i y_i = \frac{1}{X}\sum_{i=1}^{n} x_i^2 y_i,$$

$$\Delta = S_2 - S_1^2 = \frac{1}{X}\sum_{i=1}^{n} x_i^3 - \frac{1}{X^2}\left(\sum_{i=1}^{n} x_i^2\right)^2 = \frac{\sum x_i \sum x_i^3 - \left(\sum x_i^2\right)^2}{X^2}.$$

"Оптимальные" параметры функции $f(x) = ax + b$ равны соответственно:

$$\tilde{a} = \frac{V_1 - S_1 V_0}{\Delta} = \frac{1/X \sum x_i^2 y_i - 1/X \sum x_i^2 1/X \sum x_i y_i}{\Delta} =$$

$$= \frac{\sum x_i \sum x_i^2 y_i - \sum x_i^2 \sum x_i y_i}{X^2} \times \frac{1}{\Delta} = \frac{\sum x_i \sum x_i^2 y_i - \sum x_i^2 \sum x_i y_i}{\sum x_i \sum x_i^3 - (\sum x_i^2)^2},$$

$$\tilde{b} = \frac{S_2 V_0 - S_1 V_1}{\Delta} = \frac{1/X \sum x_i^3 1/X \sum x_i y_i - 1/X \sum x_i^2 1/X \sum x_i^2 y_i}{\Delta} =$$

$$= \frac{\sum x_i^3 \sum x_i y_i - \sum x_i^2 \sum x_i^2 y_i}{X^2} \times \frac{1}{\Delta} = \frac{\sum x_i^3 \sum x_i y_i - \sum x_i^2 \sum x_i^2 y_i}{\sum x_i \sum x_i^3 - (\sum x_i^2)^2}.$$

Минимум суммы квадратов отклонений для построенной функции $f(x)$ равен:

$$\Phi_{\min} = \sum_{i=1}^{n} \lambda_i y_i^2 - \tilde{a} V_1 - \tilde{b} V_0 = \frac{1}{X}\sum_{i=1}^{n} x_i y_i^2 - \frac{\tilde{a}}{X}\sum_{i=1}^{n} x_i^2 y_i - \frac{\tilde{b}}{X}\sum_{i=1}^{n} x_i y_i =$$

$$= \frac{1}{X}\left(\sum_{i=1}^{n} x_i y_i^2 - \tilde{a}\sum_{i=1}^{n} x_i^2 y_i - \tilde{b}\sum_{i=1}^{n} x_i y_i\right).$$

Случай 4. Еще одним из возможных способов определения весовых коэффициентов является выбор их пропорционально величины

$$\bar{y} = \frac{1}{y_i^2}. \tag{15}$$

Заметим, что в данном случае минимизируемая функция Φ несет дополнительную смысловую нагрузку, а именно – она представляет собой сумму квадратов относительных погрешностей приближения таблично заданной функции $y(x)$ функцией $f(x)$, построенной по методу наименьших квадратов.

Пронормировав весовые коэффициенты, получим

$$\lambda_k = \frac{1/y_k^2}{\sum_{i=1}^{n} 1/y_i^2} = \frac{1}{y_k^2 \sum_{i=1}^{n} 1/y_i^2}, \qquad k = \overline{1,n}.$$

или после введения обозначений

$$\bar{Y} = \sum_{i=1}^{n} \frac{1}{y_i^2}; \qquad \lambda_k = \frac{1}{y_k^2 \bar{Y}}, \quad k = \overline{1,n}.$$

Промежуточные формулы примут вид:

$$S_1 = \sum_{i=1}^{n} \lambda_i x_i = \frac{1}{\bar{Y}}\sum_{k=1}^{n} \frac{x_k}{y_k^2}, \quad S_2 = \sum_{i=1}^{n} \lambda_i x_i^2 = \frac{1}{\bar{Y}}\sum_{k=1}^{n}\left(\frac{x_k}{y_k}\right)^2,$$

$$V_0 = \sum_{i=1}^{n} \lambda_i y_i = \frac{1}{\bar{Y}}\sum_{k=1}^{n} \frac{y_k}{y_k^2} = \frac{1}{\bar{Y}}\sum_{k=1}^{n} \frac{1}{y_k},$$

$$V_1 = \sum_{i=1}^{n} \lambda_i x_i y_i = \frac{1}{\bar{Y}}\sum_{k=1}^{n} \frac{x_k y_k}{y_k^2} = \frac{1}{\bar{Y}}\sum_{k=1}^{n} \frac{x_k}{y_k},$$

$$\Delta = S_2 - S_1^2 = \frac{1}{\bar{Y}}\sum_{k=1}^{n}\left(\frac{x_k}{y_k}\right)^2 - \left(\frac{1}{\bar{Y}}\sum_{k=1}^{n}\frac{x_k}{y_k^2}\right)^2 = \frac{\bar{Y}\sum(\frac{x_k}{y_k})^2 - (\sum \frac{x_k}{y_k^2})^2}{\bar{Y}^2}.$$

Неизвестные параметры приближаемой функции $f(x) = ax + b$ можно записать как

$$\tilde{a} = \frac{V_1 - S_1 V_0}{\Delta} = \frac{\bar{Y}\sum(x_k/y_k) - \sum(x_k/y_k^2)\sum(1/y_k)}{\bar{Y}\sum(x_k/y_k)^2 - \big(\sum(x_k/y_k^2)\big)^2},$$

$$\tilde{b} = \frac{S_2 V_0 - S_1 V_1}{\Delta} = \frac{\sum(x_k/y_k)^2\sum(1/y_k) - \sum(x_k/y_k^2)\sum(x_k/y_k)}{\bar{Y}\sum(x_k/y_k)^2 - \big(\sum(x_k/y_k^2)\big)^2}.$$

Заметив, что

$$\sum_{i=1}^{n} \lambda_i y_i^2 = \frac{1}{\bar{Y}} \sum_{k=1}^{n} \frac{y_k^2}{y_k^2} = \frac{n}{\bar{Y}},$$

формулу для вычисления минимального значения функции Φ (квадратов относительных погрешностей) запишем следующим образом

$$\Phi_{\min} = \sum_{i=1}^{n} \lambda_i y_i^2 - V_1 \tilde{a} - V_0 \tilde{b} = \frac{1}{\bar{Y}}(n - \tilde{a}\sum_{k=1}^{n} \frac{x_k}{y_k} - \tilde{b}\sum_{k=1}^{n} \frac{1}{y_k^2}).$$

Случай 5. Пусть

$$\lambda_k = 1 - \delta; \qquad \lambda_i = \frac{\delta}{n-1}, \quad i \neq k. \tag{16}$$

Заметим, что

$$\sum_{i=1}^{n} \lambda_i \equiv const = 1. \tag{17}$$

Для определенных таким образом λ_i найдем пределы параметров a и b приближаемой функции $f(x)$ при $\delta \to 0$.

Вычислим значения промежуточных формул.

$$\begin{aligned} S_1 &= \sum_{i=1}^{n} \lambda_i x_i = (1-\delta)x_k + \frac{\delta}{n-1}\sum_{\substack{i=1 \\ i \neq k}}^{n} x_i = \\ &= (1-\delta)x_k + \frac{\delta}{n-1}\Big(\sum_{i=1}^{n} x_i - x_k\Big) = x_k + \frac{\delta}{n-1}\Big(\sum_{i=1}^{n} x_i - n x_k\Big), \end{aligned}$$

$$
\begin{aligned}
S_2 &= \sum_{i=1}^{n}\lambda_i x_i^2 = (1-\delta)x_k^2 + \frac{\delta}{n-1}\sum_{\substack{i=1\\ i\neq k}}^{n} x_i^2 = \\
&= (1-\delta)x_k^2 + \frac{\delta}{n-1}\Big(\sum_{i=1}^{n} x_i^2 - x_k^2\Big) = x_k^2 + \frac{\delta}{n-1}\Big(\sum_{i=1}^{n} x_i^2 - n x_k^2\Big), \\
V_0 &= \sum_{i=1}^{n}\lambda_i y_i = (1-\delta)y_k + \frac{\delta}{n-1}\sum_{\substack{i=1\\ i\neq k}}^{n} y_i = \\
&= (1-\delta)y_k + \frac{\delta}{n-1}\Big(\sum_{i=1}^{n} y_i - y_k\Big) = y_k + \frac{\delta}{n-1}\Big(\sum_{i=1}^{n} y_i - n y_k\Big), \\
V_1 &= \sum_{i=1}^{n}\lambda_i x_i y_i = (1-\delta)x_k y_k + \frac{\delta}{n-1}\sum_{\substack{i=1\\ i\neq k}}^{n} x_i y_i =
\end{aligned}
$$

$$
= (1-\delta)x_k y_k + \frac{\delta}{n-1}\Big(\sum_{i=1}^{n} x_i y_i - x_k y_k\Big) = x_k y_k + \frac{\delta}{n-1}\Big(\sum_{i=1}^{n} x_i y_i - n x_k y_k\Big).
$$

Введем обозначения

$$
X_1 = \sum_{i=1}^{n} x_i - n x_k, \qquad X_2 = \sum_{i=1}^{n} x_i^2 - n x_k^2,
$$

$$
Y_1 = \sum_{i=1}^{n} y_i - n y_k, \qquad \overline{XY} = \sum_{i=1}^{n} x_i y_i - n x_k y_k.
$$

Тогда

$$
\begin{aligned}
S_1 &= x_k + \frac{\delta}{n-1}X_1, &\qquad S_2 = x_k^2 + \frac{\delta}{n-1}X_2, \\
V_0 &= y_k + \frac{\delta}{n-1}Y, &\qquad V_1 = x_k y_k + \frac{\delta}{n-1}\overline{XY}.
\end{aligned}
$$

Определитель равен

$$
\begin{aligned}
\Delta &= S_2 - S_1^2 = x_k^2 + \frac{\delta}{n-1}X_2 - [x_k + \frac{\delta}{n-1}X_1]^2 = \\
&= x_k^2 + \frac{\delta}{n-1}X_2 - x_k^2 - 2x_k\frac{\delta}{n-1}X_1 - \Big(\frac{\delta}{n-1}\Big)^2 X_1^2 = \\
&= \frac{\delta}{n-1}[X_2 - 2x_k X_1 - \frac{\delta}{n-1}X_1^2].
\end{aligned}
$$

Неизвестные параметры a и b функции $f(x)$ запишем следующим образом:

$$a = \frac{V_1 - S_1 V_0}{\Delta},$$

где

$$V_1 - S_1 V_0 = x_k y_k + \frac{\delta}{n-1}\overline{XY} - (x_k + \frac{\delta}{n-1}X_1)(y_k + \frac{\delta}{n-1}Y) =$$

$$\begin{aligned} &= x_k y_k + \frac{\delta}{n-1}\overline{XY} - x_k y_k - \frac{\delta}{n-1}x_k Y - \frac{\delta}{n-1}y_k X_1 - \Big(\frac{\delta}{n-1}\Big)^2 X_1 Y = \\ &= \frac{\delta}{n-1}[\overline{XY} - x_k Y - y_k X_1 - \frac{\delta}{n-1}X_1 Y], \end{aligned}$$

подставляя, получим

$$\begin{aligned} a &= \frac{V_1 - S_1 V_0}{\Delta} = \frac{\frac{\delta}{n-1}[\overline{XY} - x_k Y - y_k X_1 - \frac{\delta}{n-1}X_1 Y]}{\frac{\delta}{n-1}[X_2 - 2x_k X_1 - \frac{\delta}{n-1}X_1^2]} = \\ &= \frac{\overline{XY} - x_k Y - y_k X_1 - \frac{\delta}{n-1}X_1 Y}{X_2 - 2x_k X_1 - \frac{\delta}{n-1}X_1^2} \end{aligned}$$

и

$$b = \frac{S_2 V_0 - S_1 V_1}{\Delta},$$

где

$$S_2 V_0 - S_1 V_1 =$$

$$\begin{aligned} &= (x_k^2 + \frac{\delta}{n-1}X_2)(y_k + \frac{\delta}{n-1}Y) - (x_k + \frac{\delta}{n-1}X_1)(x_k y_k + \frac{\delta}{n-1}\overline{XY}) = \\ &= x_k^2 y_k + \frac{\delta}{n-1}x_k^2 Y + \frac{\delta}{n-1}y_k X_2 + \Big(\frac{\delta}{n-1}\Big)^2 X_2 Y - \\ &\qquad -x_k^2 y_k - \frac{\delta}{n-1}x_k\overline{XY} - \frac{\delta}{n-1}x_k y_k X_1 - \Big(\frac{\delta}{n-1}\Big)^2 X_1\overline{XY} = \\ &= \frac{\delta}{n-1}\Big[x_k^2 Y + y_k X_2 - x_k\overline{XY} - x_k y_k X_1 + \frac{\delta}{n-1}(X_2 Y - X_1\overline{XY})\Big], \end{aligned}$$

получим

$$
\begin{aligned}
b &= \frac{\frac{\delta}{n-1}\Big[x_k^2Y + y_kX_2 - x_k\overline{XY} - x_ky_kX_1 + \frac{\delta}{n-1}(X_2Y - X_1\overline{XY})\Big]}{\frac{\delta}{n-1}[X_2 - 2x_kX_1 - \frac{\delta}{n-1}X_1^2]} = \\
&= \frac{x_k^2Y + y_kX_2 - x_k\overline{XY} - x_ky_kX_1 + \frac{\delta}{n-1}(X_2Y - X_1\overline{XY})}{X_2 - 2x_kX_1 - \frac{\delta}{n-1}X_1^2}.
\end{aligned}
$$

Переходя к пределу при $\delta \to 0$, получаем

$$
\begin{aligned}
\lim_{\delta\to 0} a &= \frac{\overline{XY} - x_kY - y_kX_1}{X_2 - 2x_kX_1} = \tilde{a}, \\
\lim_{\delta\to 0} b &= \frac{x_k^2Y + y_kX_2 - x_k\overline{XY} - x_ky_kX_1}{X_2 - 2x_kX_1} = \tilde{b}.
\end{aligned}
$$

Итак, предельное положение при $\delta \to 0$ занимает прямая $y = \tilde{a}x + \tilde{b}$. Заметим, что

$$
\tilde{a}x_k + \tilde{b} = \frac{\overline{XY} - x_kY - y_kX_1}{X_2 - 2x_kX_1}x_k + \frac{x_k^2Y + y_kX_2 - x_k\overline{XY} - x_ky_kX_1}{X_2 - 2x_kX_1} =
$$

$$
= \frac{x_k\overline{XY} - x_k^2Y - x_ky_kX_1 + x_k^2Y + y_kX_2 - x_k\overline{XY} - x_ky_kX_1}{X_2 - 2x_kX_1} =
$$

$$
= \frac{y_kX_2 - 2x_ky_kX_1}{X_2 - 2x_kX_1} = \frac{y_k(X_2 - 2x_kX_1)}{X_2 - 2x_kX_1} = y_k.
$$

Значит, искомая прямая пройдет через точку (x_k, y_k).

Литература

[1] Воеводин В. В. *Вычислительные основы линейной алгебры.* М.: Наука, 1977. 303 с.

[2] Г а у с с (G a u s s C. F.), *Theoria motus corporum coelestium.* Hamburg, 1809.

[3] Г а у с с (G a u s s C. F.), *Disquisitio de elementis ellipticis Pabladis.* 1810.

[4] Г а у с с (G a u s s C. F.), *Theoria combinationis observationum erroribus minimis obnoxiae.* 1821.

[5] З о р к а л ь ц е в В. И. *Метод наименьших квадратов: геометрические свойства, альтернативные подходы, приложения.* Новосибирск: Наука, 1995. 218 с.

[6] И к р а м о в Х. Д. *Численное решение линейных задач метода наименьших квадратов.* В кн. Математический анализ, 23. Итоги науки и техники, М.: ВИНИТИ, 1985.

[7] К о л м о г о р о в А. Н., *К обоснованию метода наименьших квадратов.* Успехи матем. наук, 1, вып. 1, 1946. С. 57-70.

[8] Л а п л а с (L a p l a c e P. S.) *Théorie analytique des probabilités.* Paris, 1812.

[9] Л е ж а н д р (L e g e n d r e A. M.) *Nouvelles méthodes pour la détermination des orbites des cométes.* Paris, 1806, Appendice sur la méthode des moindres carrés.

[10] Л и н н и к Ю. В. *Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений.* М.:Физматгиз, 1962. 349 с.

[11] Л о у с о н Ч., Х е н с о н Р., *Численное решение линейных задач метода наименьших квадратов.* М.: Наука, 1986. 232 с.

[12] М а р к о в А. А., *Закон больших чисел и метод наименьших квадратов (1898).* Изб труд, Изд АН СССР, 1951. С. 233-251.

[13] М а р к о в А. А., *Исчисление вероятностей.* ГИЗ, изд 4, 1924.

[14] Р а о (R a o C. R.), *On the linear combination of observations and the general theory of least squares.* Sankhya, 7, 1946. C. 237-256.

[15] Ч е б ы ш е в П. Л. *Об интерполировании по способу наименьших квадратов.* Соч., т. 1, 1859.

[16] Э й т к е н (A i t k e n A. C.), *On least squares and linear combination of observations.* Proc. Roy. Soc. Edin, 55, 1935, p. 42-48.