# Неравенство Крамера –Рао для оценок с ограничениями и потенциальная точность количественного рентгеноспектрального анализа

*Eduard* Shekhter,
*The free researcher*, *Germany*, *Linsingenstr. 34, 30163 Hannover*,
ed.shekhter@gmail.com
www.chekhter.de

*Резюме*

В работе [1] предложено рассматривать метод фундаментальных параметров количественного рентгеноспектрального анализа как статистическую задачу оценки параметров при наличии ограничений. Точность оценки можно характеризовать дисперсионной матрицей оценки. Наилучшей или оптимальной оценкой считается несмещенная оценка минимальной дисперсии. Тогда для этой минимальной дисперсии можно указать потенциально достижимую границу, называемую границей Крамера – Рао. Она не зависит от способа оценивания, а только от количества информации, содержащейся в зарегистрированном сигнале. В данной статье приводится вывод неравенства Крамера – Рао как для безусловных оценок, так и для оценок с с ограничениями. На примере двухкомпонентного образца иллюстрируются его свойства. Полученный результат сравнивается с формулой Зиболда, [4] характеризующей потенциальную точность количественного микроанализа.

## 1 Введение

Если некий прибор, такой как радиолокатор, телефонный кабель или рентгеновский микроанализатор рассматривать как черный ящик, канал передачи информации, свойства которого известны, и на вход которого подан сигнал, а на выходе этот сигнал регистрируется на фоне шумов этого канала, то с помощью системы обработки этот сигнал может быть восстановлен с точностью, не превышающей некоторой границы, определяемой отношением сигнал – шум на входе канала, а более обще - количеством информации, содержащейся в принятом сигнале.

Впрочем, этот результат известен не только каждому грамотному физику, но любому неискушенному. Сколько не крути ручку приемника, настраивая на волну станции или увеличивая громкость, при слабом сигнале не добиться хорошего качества.

Если точность оценки характеризовать дисперсией оценки параметров этого сигнала, то эта граничная дисперсия может быть вычислена при условии знания статистических свойств шума, на фоне которого сигнал зарегистрирован. Граница эта не зависит от способа обработки сигнала. Она называется границей Крамера – Рао.

Если задачу решения уравнений метода фундаментальных параметров количественного рентгеноспектрального анализа рассматривать как задачу статистической оценки параметров, то с помощью этой границы можно оценить теоретически потенциальную точность метода в случае многокомпонентного образца.

В литературе по математической статистике имеются многочисленные описания свойств границы Крамера – Рао, выведенных при достаточно произвольных предположениях об исходных статистических распределениях, и обобщения на оценки со смещением или оценки с ограничениями, как для линейного, так и для нелинейного оценивания, но отсутствует компактное описание их свойств. Поэтому я привожу как ссылки на литературу [5, 7, 8, 10, 11], так и вывод соответствующих выражений.

## 2 Статистическая постановка задачи

В работе [1] предложено рассматривать метод фундаментальных параметров количественного рентгеноспектрального анализа как статистическую задачу оценки параметров при наличии ограничений. Так как аналитический сигнал формируется как отношение интенсивностей пиков характеристического излучения к аналогичной интенсивности, измеренной на образце сравнения (или рассчитанной), с целью исключения приборной константы, то уравнения связи можно записать

$$\frac{c_i}{c_i^{st}} = \frac{I_i}{I_i^{st}} f_i(c_1, \ldots c_n, p_1, \ldots p_m) \tag{1}$$

Здесь $\frac{c_i}{c_i^{st}}$ относительные концентрации, $k_i = \frac{I_i}{I_i^{st}}$ отношение интенсивности излучения характеристической линии , исправленной на фон сплошного излучения, к аналогичному излучению от эталонного образца, $f_i(c_1, \ldots c_n, p_1, \ldots p_m)$ - сложная функция, скорее алгоритм, описывающая механизм взаимодействия элементов. За подробностями я отсылаю к работе – обзору [2].

Пусть $F_i(c) = \frac{c_i}{c_i^{st} f_i(c_1, \ldots c_n, p_1, \ldots p_m)}$ наша расчетная физическая модель ,то уравнение (1) с учетом шумов эксперимента можно переписать

$$\mathbf{k} = \mathbf{F}(\mathbf{c}) + \mathbf{n} \tag{2}$$

где все величины, входящие в формулу, считаются векторными для многокомпонентного случая. Вектор $\mathbf{n}$ обозначает вектор статистических погрешностей, вектор шума с известными статистическими свойствами. Будем считать, что погрешности обусловлены только статистикой счета рентгеновских квантов. Тогда среднее значение шума равно нулю и вектор $\mathbf{k}$ это случайный вектор экспериментальных данных, имеющих нормальное распределение с известной дисперсионной матрицей $\mathbf{D_k}(\mathbf{c})$, и $\mathbf{c}$ - вектор искомых концентраций . Матрица

$\mathbf{D_k}\left(\mathbf{c}\right)$ диагональна, так как погрешности аналитического сигнала для различных компонент статистически независимы.

Кроме того, должно априори выполняться

$$\sum_i c_i = 1 \tag{3}$$

Сумма всех весовых концентраций должна составлять 100%. Часто анализируемые образцы в геологии или полупроводниковые составы, равно как сплавы металлов представляют собой соединения с частично известными стехиометрическими формулами. Эти соотношения совместно с уравнением (3) можно записать в виде матричного соотношения

$$\mathbf{Gc} = \mathbf{u} \tag{4}$$

где $\mathbf{G}\left(r \times n\right)$ неквадратная матрица ограничений, $\mathbf{u}$ известный неслучайный вектор. Подробности определений смотри в работе [3].

# 3 Неравенство Крамера – Рао и статистические выводы

## 3.1 Неравенство Крамера – Рао для безусловных оценок

Пусть есть набор экспериментальных данных $x_1,...x_n$, зависящий от параметров $\alpha_1,...\alpha_n$. Пусть $\hat{\alpha}_1,...\hat{\alpha}_n$ несмещенные оценки параметров $\alpha$, $E\left(\hat{\alpha}_i\left(x_1,...x_n\right)\right) = \alpha_i$, с дисперсионной матрицей $\mathbf{D} = \left\{d_{ij}\right\}$, $d_{ij} = E\left[\left(\hat{\alpha}_i - \alpha_i\right)\left(\hat{\alpha}_j - \alpha_j\right)\right]$. Функция плотности вероятности, рассматриваемая как случайная функция набора этих случайных данных $x_1,...x_n$, называется функцией правдоподобия выборки $L$,

$$L(x_1,...x_n,\alpha_1,...\alpha_n) = P\left(x_1,...x_n,\alpha_1,...\alpha_n\right) \tag{5}$$

Матрица $\mathbf{J}$ с элементами

$$\mathbf{J}_{ij} = -E\left(\frac{\partial^2 L}{\partial\alpha_i \partial\alpha_j}\right), \tag{6}$$

вычисленными в точке истинного значения вектора параметров $\boldsymbol{\alpha}$,называется информационной матрицей Фишера.

Справедлива теорема: при условии существования и неособенности матрицы Фишера и еще некоторых условиях, для всяких векторных величин $\boldsymbol{\theta}$ справедливо

$$\boldsymbol{\theta}'\mathbf{D}\boldsymbol{\theta} \geq \boldsymbol{\theta}'\mathbf{J}^{-1}\boldsymbol{\theta} \tag{7}$$

Иначе это неравенство записывают

$$\mathbf{D} - \mathbf{J}^{-1} \geq 0, \tag{8}$$

имея в виду неотрицательную определенность этой матрицы. Равенство в (7), (8) достигается для так называемых эффективных оценок,

$$\mathbf{D}_\theta = \mathbf{J}^{-1} \tag{9}$$

Пусть $\theta_i = \hat{\alpha}_i - \alpha_i$, равно разности между оценкой параметра и его действительным значением. Тогда величина $\boldsymbol{\theta}'\mathbf{D}\boldsymbol{\theta} = const$ называется кореляционным эллипсоидом и определяет область доверительной вероятности при оценке параметров. Таким образом, неравенство (7) определяет нижнюю границу для области доверительной вероятности при оценке параметров, вне зависимости от способа оценивания.

Нам понадобится еще одно утверждение касательно оценок произвольных функций параметров. Пусть $f_1,\dots f_r$ функции от наблюдений, являющиеся несмещенными оценками функций от параметров $\varphi_i(\alpha_1,\dots\alpha_n)$, с корреляционной матрицей $\mathbf{V}$,

$$E(f_i) = \varphi_i \tag{10}$$

$$E\left[(f_i - \varphi_i)(f_j - \varphi_j)\right] = V_{ij} \tag{11}$$

Пусть $L$ функция правдоподобия и $\mathbf{J}$ информационная матрица Фишера, формула (6).Кроме того пусть $\boldsymbol{\Delta}$ матрица с элементами $\Delta_{ij} = \dfrac{\partial \varphi_i}{\partial \alpha_j}$.Тогда в том случае, когда существует матрица

$\mathbf{J}^{-1}$, матрица

$$\mathbf{B} = \mathbf{V} - \boldsymbol{\Delta}\mathbf{J}^{-1}\boldsymbol{\Delta}' \geq \mathbf{0} \tag{12}$$

неотрицательно определенная. Для эффективной оценки $\hat{\boldsymbol{\alpha}}$ функции $\hat{\varphi}_i = \varphi_i(\hat{\boldsymbol{\alpha}})$ будут также эффективными оценками. Следовательно в (11) достигается равенство,

$$\mathbf{V} = \mathbf{D}_\phi = \boldsymbol{\Delta}\mathbf{J}^{-1}\boldsymbol{\Delta}' \tag{13}$$

Пусть $\varphi_i(\alpha_1,\dots\alpha_n)$ линейные функции параметров, то есть

$$\boldsymbol{\varphi} = \mathbf{A}\boldsymbol{\alpha} + \mathbf{b} \tag{14}$$

где $\boldsymbol{\varphi}' = [\varphi_1,\dots\varphi_r]$, $\mathbf{A}[r \times n]$. В этом случае $\boldsymbol{\Delta} = \mathbf{A}$ в (12).и, следовательно,

$$\mathbf{D}_\phi = \boldsymbol{\Delta}\mathbf{J}^{-1}\boldsymbol{\Delta}' = \mathbf{A}\mathbf{J}^{-1}\mathbf{A}' \tag{15}$$

Предположим, что информационная матрица $\mathbf{J}$ вырождена (определитель равен нулю). Это означает, что не существует эффективных оценок для параметров $\alpha$, например, число физических уравнений меньше числа оцениваемых параметров.

Пусть (14) допускающая однозначную эффективную оценку векторная линейная функция параметров. Тогда дисперсионная матрица этих оценок получается путем замены матрицы $\mathbf{J}^{-1}$ в (13) на матрицу $\mathbf{J}^{+}$,

$$\mathbf{D}_\phi = \mathbf{A}\mathbf{J}^{+}\mathbf{A}' \tag{16}$$

Как уже было сказано, равенство в (8), (12) достигается для эффективных оценок. ,

Если заранее неизвестны истинные значения параметров, то значение дисперсионных матриц выражения (9), (13), (15), (16) неизвестно. Тогда при оценке параметров значение, полученное на последнем шаге итерации, принимается за истинное значение параметра, а граница Крамера - Рао, вычисленная в этой точке за истинное значение дисперсионной матрицы, Пусть $\mathbf{x} = \mathbf{k}$, где $\mathbf{k}$ вектор наблюдений пункта 2, формула (2) и вектор параметров $\boldsymbol{\alpha} = \mathbf{c}$. По предположению вектор $\mathbf{k}$ распределен нормально с дисперсионной матрицей $\mathbf{D_k}\left(\mathbf{c}\right)$, $\left(\mathbf{D_k}\right)_{ij} = \sigma_{ii}^2 \delta_{ij}$,

$$\mathbf{k} \sim N\left(\mathbf{F}\left(\mathbf{c}\right), \mathbf{D}_k\right) \tag{17}$$

Тогда логарифм функции правдоподобия

$$\ln L(k) = -\frac{1}{2}\sum_i \frac{\left(k_i - F_i\right)^2}{\sigma_i^2} + const \tag{18}$$

И $E\left[\frac{\partial^2 \ln L}{\partial c_i \partial c_j}\right] = -E\left[\frac{\partial \ln L}{\partial c_i}\frac{\partial \ln L}{\partial c_j}\right]$. Тогда $\frac{\partial \ln L}{\partial c_j} = -\sum_i \frac{\left(F_i - k_i\right)}{\sigma_i^2}\frac{\partial F_i}{\partial c_j}$ и, следовательно, $-E\left[\frac{\partial \ln L}{\partial c_j}\frac{\partial \ln L}{\partial c_k}\right] = -\sum_i \sigma_i^{-2}\frac{\partial F_i}{\partial c_j}\frac{\partial F_i}{\partial c_k}$. Тогда в смысле (9)

$$\mathbf{D}_c = \mathbf{J^{-1}} = \left(\nabla\mathbf{F}'\mathbf{D}_k^{-1}\nabla\mathbf{F}\right)^{-1}, \tag{19}$$

Соответственно (15) и (16) преобразуются

$$\mathbf{D}_\phi = \mathbf{A}\left(\nabla\mathbf{F}'\mathbf{D}_k^{-1}\nabla\mathbf{F}\right)^{-1}\mathbf{A}' \tag{20}$$

$$\mathbf{D}_\phi = \mathbf{A}\left(\nabla\mathbf{F}'\mathbf{D}_k^{-1}\nabla\mathbf{F}\right)^{+}\mathbf{A}' \tag{21}$$

где $\nabla\mathbf{F}$ градиент вектора $\mathbf{F}$.

## 3.2 **Неравенство Крамера – Рао для оценок с ограничениями**.

Предположим, что есть система уравнений пункта 2, уравнения (2), (4). Уравнения (2) имеют произвольный ранг, но система уравнений (2) плюс (4) достаточны для получения однозначного решения.

Обратимся к системе уравнений (4). Предполагая, что $r < n$ в (4), запишем общее решение этой системы как

$$\mathbf{c = G^{+}u + (I - G^{+}G)z} \tag{22}$$

где $\mathbf{z}$ произвольный вектор размерности $r$. Если в (18) с помощью (22) выполнить замену $\mathbf{F}\left(\mathbf{c}\right) = \mathbf{F}\left(\mathbf{z}\right)$, то мы получим функцию правдоподобия, зависящую от переменной $\mathbf{z}$ и вектор $\mathbf{c}$ будет допускающей оценку линейной функцией параметров $\mathbf{z}$, случай, описываемый формулой (21). В этом случае

$$\mathbf{A} = \mathbf{I - G^{+}G}, \tag{23}$$

Можно показать путем несложных вычислений, что информационная матрица равна

$$\begin{aligned}\mathbf{J} &= E\left\{\left(\nabla_z \ln L\right)'\left(\nabla_z \ln L\right)\right\} \\ &= \mathbf{(I - G^{+}G)\nabla F' D}_k^{-1}\mathbf{\nabla F(I - G^{+}G)}\end{aligned} \tag{24}$$

С учетом этого, граница Крамера - Рао для оценки с ограничениями равна

$$\mathbf{D}_c = \left[ (\mathbf{I} - \mathbf{G}^+\mathbf{G})\nabla\mathbf{F}'\mathbf{D}_k^{-1}\nabla\mathbf{F}(\mathbf{I} - \mathbf{G}^+\mathbf{G}) \right]^+ \tag{25}$$

Полученный результат справедлив при любых рангах матриц $\nabla\mathbf{F}$ и $\mathbf{G}$, допускающих получение однозначного решения. Полученная дисперсионная матрица (25) является вырожденной с рангом $n-r$, так как имеется $r$ бесшумных наблюдений - связей, уравнения (4).

При оценке любой другой векторной функции $\boldsymbol{\varphi}(\mathbf{c})$, допускающей оценку и зависящей от условного решения, мы получим на основании (21)

$$\mathbf{D}_\varphi = \nabla\boldsymbol{\varphi}'\mathbf{D}_c\nabla\boldsymbol{\varphi} \tag{26}$$

где $\mathbf{D}_c$ определяется (25).

# 4 Двухкомпонентный образец как пример

Предположим, что для двухкомпонентного образца условием является $c_1 + c_2 = 1$ и $\hat{c}_1$ и $\hat{c}_2$ - оценки, полученные при этом условии, тогда, если $\delta c_i = \hat{c}_i - c_i$, то $\delta c_1 = -\delta c_2$, следовательно $\sigma_{c1}^2 = \sigma_{c2}^2 = \sigma^2$ и $<\delta c_1 \delta c_2> = -\sigma^2$. Тогда для всякой оценки для границы Крамера – Рао, формула (25) получим

$$\mathbf{D}_c = \sigma_{\min}^2 \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{27}$$

где $\sigma_{\min}^2$ минимально достижимая дисперсия для оптимальной оценки. Матрица в (28) является вырожденной из-за наличия связи, уравнения нормировки.

Попробуем найти $\sigma_{\min}^2$.

Функция плотности вероятности для двухкомпонентного образца, формула (15) после замены в ней $c_2 = 1 - c_1$ запишется как функция одной переменной $c_1$,

$$P(k_1, k_2 \mid c_1 + c_2 = 1) = const \exp\left\{ -1/2 \left[ \left[ k_1 - F_1(c_1) \right]^2 / \sigma_{k_1}^2 + \left[ k_2 - F_2(c_1) \right]^2 / \sigma_{k_2}^2 \right] \right\} \tag{28}$$

В этом одномерном случае неравенство Крамера - Рао запишется

$$\sigma_c^2 \geq \sigma_{\min}^2 = \left[ \frac{\left(\partial F_1(c_1)/\partial c_1\right)^2}{\sigma_{k_1}^2} + \frac{\left(\partial F_2(c_1)/\partial c_1\right)^2}{\sigma_{k_2}^2} \right]^{-1} \tag{29}$$

Мы имеем как бы два независимых канала измерений для одной переменной $c_1$. Если считать, что для каждого канала $E(k_i) = F_i(c)$ и $\delta k_i = \frac{\partial F_i}{\partial c}\delta c$, тогда $\left(\sigma_c^2\right)_i = \left(\frac{\partial F_i}{\partial c}\right)^{-2} \left(\sigma_k^2\right)_i$. Тогда (29) перепишется

$$\sigma_c^2 \geq \sigma_{\min}^2 = \frac{1}{\left(\sigma_c^{-2}\right)_1 + \left(\sigma_c^{-2}\right)_2} = \left(\sigma_c^{-2}\right)_1 \frac{1}{1+\frac{\left(\sigma_c^{-2}\right)_1}{\left(\sigma_c^{-2}\right)_2}} \tag{30}$$

Мы видим, что граничная дисперсия для оптимальной оценки меньше любой из дисперсий, полученных с использованием только одного из каналов наблюдения. Если мы вычислим $\left(\sigma_c^2\right)_1 = \left(\frac{\partial F_1}{\partial c}\right)^{-2} \sigma_{k_1}^2$ для случая линейной поправки, рассмотренной Зиболдом [3], $\left(c_1\right)_1 = \frac{k_1\alpha_{12}}{1-k_1(1-\alpha_{12})}$, то мы получим в случае предельно малых концентраций этой компоненты $\left(\sigma_c\right)_1 = \alpha_{12}\sigma_{k_1}$. Учтя, что содержание второй компоненты при этом стремится к единице, получим $\left(\sigma_c\right)_2 = \frac{\sigma_{k_2}}{\alpha_{21}}$. Устремив в (30) оцениваемую концентрацию к нулю, получим

$$\sigma_{\min}^2 = \left(\sigma_{k_1}\alpha_{12}\right)^2 \frac{1}{1+\left(\frac{\sigma_{k_1}}{\sigma_{k_2}}\right)^2 (\alpha_{21}\alpha_{12})^2} \tag{31}$$

и тогда (31) является обобщением результата Зиболда при использовании информации от двух каналов наблюдения. Вычисляемый им предел не является минимально достижимым пределом обнаружения.

## Литература


1. Э. Шехтер, Метод фундаментальных параметров в рентгеновской спектроскопии как задача статистической оценки параметров при наличии ограничений. Критический анализ ситуации, www.chekhter.de
2. Steinbrecher, Stefan. A unified Monte Carlo approach for quantitative standardless x-ray fluorescence and electron probe microanalysis inside the scanning electron microscope. Dissertation, 2004, http://nbn-resolving.de/urn/resolver.pl?urn=urn:nbn:de:bsz:21-opus-13007
3. Э. Шехтер, Матричная форма стехиометрических соотношений, www.chekhter.de
4. Ziebold T.O., Anal. Chem. 1967, v.39, Nr. 8, p.86
5. G. A. F. Seber, A.J. Lee, Linear Regression Analysis, 2. ed., J.W. & S. (2003)1
6. A. Albert, Regression and the Moore-Penrose Pseudoinverse, A.P., N.Y.(1972)
7. C.R. Rao, Linear Statistical Inference and Its Application, 2nd ed., J.W., N.Y., (1973)
8. Juan M. Rodriguwz-Poo, Computer-aided Introduction to Econometrics, edited by Springer (2002), 350 p., SBN:354044114X, (Part Multivariate Linear Regression Model)
9. D. Kahaner, C.B. Moller, S. Nash, Numerical methods and software, Prentice-Hall, (1989)
10. G. A. F. Seber, C.J. Wild, Nonlinear Regression, J.W., N.Y., (1989)
11. D.C. Montgomery, E.A. Peck, G.G. Vining, Introduction to Linear Regression Analysis, 3 ed., J.W.&S., N.Y., 2001
12. Ziebold T.O., Ogilvie R.E., Anal. Chem. 1964, v.36, p.322-327
13. E. Shekhter, Industrial Laboratory, 1986, v. 52, Nr. 5, p.p. 445-449
14. E. Shekhter, G. Edelstein, Industrial Laboratory, 1991, v. 57, Nr. 5, p.p. 536-545