НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЕ УЧРЕЖДЕНИЕ
ИНСТИТУТ ЯДЕРНОЙ ФИЗИКИ им. Г.И.Будкера СО РАН

С.И. Редин

СИСТЕМАТИЧЕСКИЙ СДВИГ ПАРАМЕТРОВ ПРИ АППРОКСИМАЦИИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

ИЯФ 2002-73

Новосибирск
2002

S.I. Redin

Fit of experimental histograms: bias of fit parameters

Abstract

For a multiparameter fit of an experimental histogram we derive equations for the bias of fit parameters $\{x_i\} \equiv \boldsymbol{x}$ in a form of ensamble average $\langle \Delta \boldsymbol{x} \rangle$ of difference between value of $\boldsymbol{x}$ obtained from the fit and the [unknown] "true" value $\boldsymbol{x}_\circ$: $\Delta \boldsymbol{x} \equiv \boldsymbol{x} - \boldsymbol{x}_\circ$. These equations are derived for the likelihood function $\mathcal{L} \equiv \sum_n \left(\mathcal{N}_n \ln \frac{\mathcal{N}_n}{f} + f - \mathcal{N}_n \right)$ fit (section 3) as well as for the $\chi^2 \equiv \sum_n \frac{(f-\mathcal{N}_n)^2}{\sigma_n^2}$ fit with two choices of variance σ_n^2: $\sigma_n^2 = \mathcal{N}_n$ (section 4) and $\sigma_n^2 = f$ (section 5), where $\mathcal{N}_n$ is number of counts in the n^{th} channel of the histogram and f is a fit function value for this channel. The results for the bias are given in eqs.(17), (25) and (31) for multiparameter fit and in eqs.(19), (26) and (32) for one parameter fit.

Equation for bias for $\mathcal{L}$ fit contains only one term, while equations for bias for χ^2 contain the same term as for $\mathcal{L}$ fit plus two extra terms (second and third). Simple qualitative estimates for arbitrary fit function and numerical estimates for the 5-parameter fit used for the muon (g-2) experiment at BNL show that one of these extra terms is enhanced by factor of $N_{ch} \gg 1$, where N_{ch} is number of channels in the histogram, and therefore is expected to be dominant. By this reason the likelihood function fit is expected to have less bias than both considered χ^2 fits.

As evident from eqs.(25) and (31), the second and third terms for bias for $\chi^2_{\sigma_n^2=\mathcal{N}_n}$ fit are 2 times greater (in absolute value) than that for $\chi^2_{\sigma_n^2=f}$ fit and hance the latter in general has somehow less bias. This can be understood from comparison of Taylor expansions for $\chi^2_{\sigma_n^2=f}$ and $\chi^2_{\sigma_n^2=\mathcal{N}_n}$ with that for $\mathcal{L}$: $\chi^2_{\sigma_n^2=f}$ (eq.5) match $2\mathcal{L}$ (eq.3) [in first two terms of Taylor expansion] better than $\chi^2_{\sigma_n^2=\mathcal{N}_n} = \sum_n \frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n}$ does. That, by the way, gives a clue for possible improvement of χ^2 fit: it should be modified in order to coincide with $2\mathcal{L}$ in first two terms of Taylor expansion. In this work we offer two possible modifications: $\chi^2_{\sigma_n^2=\mathcal{N}_n \times (f/\mathcal{N}_n)^{2/3}}$ (section 10) and $\chi^2_{corr} = \sum_n \left(\frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n} - \frac{2}{3} \frac{(f-\mathcal{N}_n)^3}{\mathcal{N}_n^2} \right)$ (section 11). Direct calculations of bias of fit parameters for these modifications give same results as for the likelihood function fit, see eqs.(52,53) and eqs.(58,59).

Содержание

1 Введение

Подгонка (аппроксимация) экспериментальных распределений (гистограмм) некоторой функцией является одной из наиболее распространенных процедур при анализе экспериментальных данных в физике элементарных частиц. Целью подгонки как правило является получение численных значений параметров подгоночной функции, которые в свою очередь прямо или косвенно связаны с измеряемыми в эксперименте величинами.

В настоящей работе изучается систематический сдвиг параметров при подгонке гистограмм методом максимального правдоподобия и методом минимизации χ^2. Для метода χ^2 рассмотрены два варианта определения статистической вариации σ_n^2, связанной с $n-$м каналом гистограммы: вариант с $\sigma_n^2 = \mathcal{N}_n$, где $\mathcal{N}_n$ – число событий в этом канале гистограммы и вариант с $\sigma_n^2 = f$, где f – значение подгоночной функции для этого канала. Показано, что наименьшим сдвигом обладает метод максимального правдоподобия. Что же касается χ^2 с $\sigma_n = \sqrt{\mathcal{N}_n}$ и с $\sigma_n = \sqrt{f}$, то последнее лучше аппроксимирует функцию правдоподобия и, возможно по этой причине, при минимизации дает меньший (в 2 раза) систематический сдвиг параметров подгонки. Тем не менее, минимизация χ^2 с $\sigma_n = \sqrt{\mathcal{N}_n}$, которая на практике используется чаще всего и "по умолчанию", технически проще и, с некоторой осторожностью, вполне может быть использована.

Из анализа методов χ^2 с $\sigma_n = \sqrt{\mathcal{N}_n}$ и с $\sigma_n = \sqrt{f}$ и их сравнения с методом максимального правдоподобия найден способ возможного улучшения качества подгонки по χ^2, а именно уменьшения систематического сдвига параметров подгонки до величины, получаемой в методе максимального правдоподобия. В работе приведено несколько вариантов такого усовершенствования метода χ^2.

Формулы, полученные при вычислении систематических сдвигов, могут быть использованы для численной оценки сдвигов параметров при подгонке экспериментальных распределений, особенно в прецизионных экспериментах.

2 Функция правдоподобия $\mathcal{L}$; функция χ^2 как аппроксимация функции правдоподобия

Вероятность того что заданное экспериментальное распределение соответствует некоторой теоретической модели, описываемой функцией $f(\boldsymbol{x}, t)$, может быть записана следующим образом:

$$
\begin{aligned}
\mathcal{P} &= \prod_n \frac{[f(\boldsymbol{x}, t_n)]^{\mathcal{N}_n}}{\mathcal{N}_n\,!}\ \exp\left[-f(\boldsymbol{x}, t_n)\right] \\
&\approx \prod_n \frac{1}{\sqrt{2\pi\mathcal{N}_n}}\ \left(\frac{f(\boldsymbol{x}, t_n)}{\mathcal{N}_n}\right)^{\mathcal{N}_n} \exp\left[\mathcal{N}_n - f(\boldsymbol{x}, t_n)\right]. \qquad (1)
\end{aligned}
$$

Здесь мы предполагаем Пуассоновское распределение числа событий в каждом канале гистограммы и используем формулу Стирлинга для $(\mathcal{N}_n\,!)$. В уравнении (1) и далее мы используем следующие обозначения:

- $n-$ номер канала гистограммы, $\mathcal{N}_n-$ число событий в $n-$м канале, $N = \sum_n \mathcal{N}_n$ – полное число событий в гистограмме;
- $f(\boldsymbol{x}, t) \equiv f-$ подгоночная функция, $\boldsymbol{x} \equiv \{x_i\}-$ вектор параметров подгонки, $f(\boldsymbol{x}, t_n)$ – значение подгоночной функции для $n-$го канала гистограммы;
- $f'_i \equiv \frac{\partial f}{\partial x_i}$ – частная производная подгоночной функции f по $i-$му параметру.

Определим функцию правдоподобия как $\mathcal{L} \equiv -\ln \mathcal{P}$:

$$
\begin{aligned}
\mathcal{L} &= \sum_n \left(\frac{1}{2}\ln 2\pi + \frac{1}{2}\ln \mathcal{N}_n + \mathcal{N}_n \ln \frac{\mathcal{N}_n}{f} + f - \mathcal{N}_n\right) \\
&= \mathrm{Const} + \sum_n \left(\mathcal{N}_n \ln \frac{\mathcal{N}_n}{f} + f - \mathcal{N}_n\right). \qquad (2)
\end{aligned}
$$

Заметим что выражение (2) может быть получено и без использования формулы Стирлинга для $(\mathcal{N}_n\,!)$. Более того, в отличие от формулы Стирлинга, уравнение (2) верно для произвольных значений $\mathcal{N}_n$, включая $\mathcal{N}_n = 0$. Для простоты в дальнейшем мы будем отбрасывать константу в уравнении (2), таким образом $\mathcal{L} \equiv \sum_n \left(\mathcal{N}_n \ln \frac{\mathcal{N}_n}{f} + f - \mathcal{N}_n\right)$.

Для непосредственного сравнения функции правдоподобия с выражением для χ^2, мы домножим $\mathcal{L}$ на 2 и разложим полученное выражение по степеням малой величины $[f(\boldsymbol{x}, t_n) - \mathcal{N}_n]/\mathcal{N}_n$:

$$\begin{aligned}
2\,\mathcal{L} &\equiv 2\sum_n \left(\mathcal{N}_n \ln\frac{\mathcal{N}_n}{f} + f - \mathcal{N}_n\right) \\
&= 2\sum_n \left(-\mathcal{N}_n \ln\left[1 + \frac{f-\mathcal{N}_n}{\mathcal{N}_n}\right] + f - \mathcal{N}_n\right) \\
&\approx 2\sum_n \left(-\mathcal{N}_n \left[\frac{f-\mathcal{N}_n}{\mathcal{N}_n} - \frac{(f-\mathcal{N}_n)^2}{2\,\mathcal{N}_n^2}\right.\right. \\
&+ \left.\left.\frac{(f-\mathcal{N}_n)^3}{3\,\mathcal{N}_n^3} - \frac{(f-\mathcal{N}_n)^4}{4\,\mathcal{N}_n^4} + \ldots\right] + f - \mathcal{N}_n\right) \\
&= \sum_n \left(\frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n} - \frac{2}{3}\,\frac{(f-\mathcal{N}_n)^3}{\mathcal{N}_n^2} + \frac{1}{2}\,\frac{(f-\mathcal{N}_n)^4}{\mathcal{N}_n^3} - \ldots\right) \qquad (3)
\end{aligned}$$

или

$$2\,\mathcal{L} \approx \sum_n \frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n} \qquad (4)$$

Выражение в правой части уравнения (4) является ничем иным как $\chi^2 \equiv \sum_n \frac{(f-\mathcal{N}_n)^2}{\sigma_n^2}$ с $\sigma_n = \sqrt{\mathcal{N}_n}$. С другой стороны, функция χ^2 с $\sigma_n = \sqrt{f(\boldsymbol{x}, t_n)}$ может быть записана как

$$\begin{aligned}
\chi^2_{\sigma_n=\sqrt{f}} &= \sum_n \frac{(f-\mathcal{N}_n)^2}{f} = \sum_n \frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n + (f-\mathcal{N}_n)} \\
&= \sum_n \left(\frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n} - \frac{(f-\mathcal{N}_n)^3}{\mathcal{N}_n^2} + \frac{(f-\mathcal{N}_n)^4}{\mathcal{N}_n^3} - \ldots\right) \qquad (5)
\end{aligned}$$

что ближе к $2\mathcal{L}$ в уравнении (3) чем $\chi^2_{\sigma_n=\sqrt{\mathcal{N}_n}}$. Разумно ожидать, что по этой причине подгонка по $\chi^2_{\sigma_n=\sqrt{f}}$ будет лучше (в частности, будет иметь меньший систематический сдвиг), чем подгонка по $\chi^2_{\sigma_n=\sqrt{\mathcal{N}_n}}$.

3 Систематический сдвиг параметров подгонки для метода максимального правдоподобия

В результате статистических флуктуаций числа событий в индивидуальных каналах гистограммы, подгонка по методу максимального правдоподобия, т.е. оптимизация функции правдоподобия $\mathcal{L} = \sum_n \left(\mathcal{N}_n \ln \frac{\mathcal{N}_n}{f} + f - \mathcal{N}_n\right)$, будет давать значение параметров подгонки $\boldsymbol{x}$, сдвинутое по отношению к “истинному” значению $\boldsymbol{x}_\circ$ на некоторую величину $\Delta\boldsymbol{x}$: $\boldsymbol{x} = \boldsymbol{x}_\circ + \Delta\boldsymbol{x}$. Элементы вектора $\Delta\boldsymbol{x}$ могут быть найдены из условия оптимизации $\mathcal{L}$, т.е. $\partial\mathcal{L}/\partial x_j = 0$:

$$\begin{aligned} 0 &= \frac{\partial\mathcal{L}}{\partial x_j} = \sum_n f'_j \left(-\frac{\mathcal{N}_n}{f} + 1\right) \\ &= \sum_n f'_j \,\frac{f - \mathcal{N}_n}{f} \approx \sum_n f'_j \,\frac{f_\circ + (\sum_i f'_i\,\Delta x_i) - \mathcal{N}_n}{f_\circ} = \\ &= \sum_i \left(\sum_n \frac{f'_i\,f'_j}{f_\circ}\right)\Delta x_i - \left(\sum_n \frac{f'_j}{f_\circ}(\mathcal{N}_n - f_\circ)\right) = 0 \end{aligned} \qquad (6)$$

и следовательно

$$\Delta x_i = \sum_j \left(\mathcal{A}^{-1}\right)_{ij} \sum_n \frac{f'_j}{f_\circ}(\mathcal{N}_n - f_\circ)\,, \qquad \text{где} \quad \mathcal{A}_{ij} = \sum_n \frac{f'_i\,f'_j}{f_\circ} \qquad (7)$$

Здесь мы использовали приближение $f(\boldsymbol{x}) = f(\boldsymbol{x}_\circ + \Delta\boldsymbol{x}) \approx f_\circ + \sum_i f'_i\,\Delta x_i$, где $f_\circ \equiv f(\boldsymbol{x}_\circ, t)$, для того чтобы в явном виде выделить функцию $f_\circ$ и использовать ее свойства: $\langle\,\mathcal{N}_n - f_\circ\,\rangle = 0$, $\langle\,(\mathcal{N}_n - f_\circ\,)^2\rangle = f_\circ$ и $\langle\,(\mathcal{N}_n - f_\circ\,)(\mathcal{N}_m - f_\circ\,)\,\rangle = 0$ для $n \neq m$. Угловыми скобками $\langle ... \rangle$ здесь и далее мы будем обозначать усреднение по ансамблю однотипных распределений (экспериментов).

Уравнение (7), которое в явном виде связывает статистические флуктуации параметров подгонки и статистические флуктуации числа событий в каналах гистограммы, является одним из наиболее важных в статистике. К примеру, его можно использовать для вывода уравнения для такой фундаментальной величины как корреляция параметров подгонки:

$$\begin{aligned}
\langle \Delta x_i \; \Delta x_j \rangle &= \sum_{ab} \left(\mathcal{A}^{-1}\right)_{ia} \left(\mathcal{A}^{-1}\right)_{jb} \sum_{nm} \left(\frac{f'_a}{f_\circ}\right)_n \left(\frac{f'_b}{f_\circ}\right)_m \langle \, (\mathcal{N}_n - f_\circ)(\mathcal{N}_m - f_\circ) \, \rangle \\
&= \sum_{ab} \left(\mathcal{A}^{-1}\right)_{ia} \left(\mathcal{A}^{-1}\right)_{jb} \sum_{n} \frac{f'_a \, f'_b}{f_\circ} = \sum_{ab} \left(\mathcal{A}^{-1}\right)_{ia} \left(\mathcal{A}^{-1}\right)_{jb} \mathcal{A}_{ab} \\
&= \left(\mathcal{A}^{-1}\right)_{ij}
\end{aligned} \tag{8}$$

или

$$\langle \Delta x_i \; \Delta x_j \rangle = \frac{A_{ij}}{det(\mathcal{A})} = (-1)^{i+j} \frac{M_{ij}}{det(\mathcal{A})}. \tag{9}$$

Как частный, но наиболее важный, случай уравнений (8) и (9), можно очень просто получить выражение для статистических ошибок параметров подгонки:

$$\sigma_i^2 \quad \equiv \quad \left\langle \, (\Delta x_i)^2 \right\rangle = \left(\mathcal{A}^{-1}\right)_{i\,i} \tag{10}$$

или

$$\sigma_i^2 = \frac{A_{ii}}{det(\mathcal{A})} = \frac{M_{ii}}{det(\mathcal{A})}. \tag{11}$$

Здесь M_{ij} и A_{ij} – миноры и алгебраические дополнения матричных элементов симметричной матрицы $\mathcal{A}$, а $det(\mathcal{A})$ – ее детерминант.

Систематические сдвиги параметров подгонки в процессе оптимизации функции правдоподобия могут быть найдены как среднее по ансамблю от величины $\Delta \boldsymbol{x}$, т.е. $\langle \Delta \boldsymbol{x} \rangle$. Однако усреднение величины $\Delta \boldsymbol{x}$, приведенной в уравнении (7), зануляется:

$$\begin{aligned}
\langle \Delta x_i \rangle \quad &= \quad \left\langle \sum_j \left(\mathcal{A}^{-1}\right)_{ij} \sum_n \frac{f'_j}{f_\circ} (\mathcal{N}_n - f_\circ) \right\rangle \\
&= \quad \sum_j \left(\mathcal{A}^{-1}\right)_{ij} \sum_n \frac{f'_j}{f_\circ} \langle \mathcal{N}_n - f_\circ \rangle = 0
\end{aligned} \tag{12}$$

поскольку $\langle \mathcal{N}_n - f_\circ \rangle = 0$. Это означает что приближение, сделанное при выводе уравнения (6), не является достаточным для вычисления систематических сдвигов параметров подгонки. С учетом следующего порядка приближения, вместо уравнения (6) мы получаем:

$$\begin{aligned}
0 &= \frac{\partial \mathcal{L}}{\partial x_j} = \sum_n f'_j \left(-\frac{\mathcal{N}_n}{f} + 1\right) \\
&= \sum_n f'_j \, \frac{f - \mathcal{N}_n}{f} = \sum_n \frac{f'_j + \sum_i f''_{ji} \Delta x_i + ...}{f_\circ + \sum_i f'_i \Delta x_i + ...} \\
&\times \left(f_\circ - \mathcal{N}_n + \sum_i f'_i \, \Delta x_i + \frac{1}{2} \sum_{ik} f''_{ik} \, \Delta x_i \, \Delta x_k + ...\right) \\
&\approx \sum_n \frac{f'_j}{f_\circ} \left(f_\circ - \mathcal{N}_n + \sum_i f'_i \, \Delta x_i + \frac{1}{2} \sum_{ik} f''_{ik} \, \Delta x_i \, \Delta x_k\right) \\
&\times \left(1 - \frac{\sum_i f'_i \Delta x_i}{f_\circ} + \frac{\sum_i f''_{ik} \Delta x_i \Delta x_i}{f_\circ}\right). \qquad (13)
\end{aligned}$$

Мы будем искать решение уравнения (13) в виде последовательных приближений: $\Delta x_i = \Delta x_i^\circ + \Delta x_i^1 + ...$, где Δx_i° – главное приближение, приведенное в уравнении (7):

$$\Delta x_i^\circ = \sum_j \left(\mathcal{A}^{-1}\right)_{ij} \sum_n \frac{f'_j}{f_\circ} (\mathcal{N}_n - f_\circ), \qquad (14)$$

а Δx_i^1 – следующее приближение. Для Δx_i^1 мы получаем уравнение

$$\begin{aligned}
&\sum_n \frac{f'_i}{f_\circ} \left[\sum_p f'_p \, \Delta x_p^1 + \frac{1}{2} \sum_j \sum_k f''_{jk} \, \Delta x_j^\circ \, \Delta x_k^\circ \right. \qquad (15) \\
&\left. + \left(-\frac{\sum_j f'_j \, \Delta x_j^\circ}{f_\circ} + \frac{\sum_j f''_{ij} \, \Delta x_j^\circ}{f'_i}\right) \left((f_\circ - \mathcal{N}_n) + \sum_k f'_k \, \Delta x_k^\circ\right)\right] = 0,
\end{aligned}$$

которое имеет решение

$$\begin{aligned}
\Delta x_p^1 &= \sum_i \left(\mathcal{A}^{-1}\right)_{pi} \left[-\sum_j \sum_n \frac{f'_i \, f'_j}{f_\circ^2} (\mathcal{N}_n - f_\circ) \, \Delta x_j^\circ \right. \\
&+ \sum_{jk} \sum_n \frac{f'_i \, f'_j \, f'_k}{f_\circ^2} \Delta x_j^\circ \, \Delta x_k^\circ + \sum_j \sum_n \frac{f''_{ij}}{f_\circ} (\mathcal{N}_n - f_\circ) \, \Delta x_j^\circ \\
&- \left. \sum_{jk} \sum_n \frac{f''_{ij} \, f'_k}{f_\circ} \Delta x_j^\circ \, \Delta x_k^\circ - \frac{1}{2} \sum_{jk} \sum_n \frac{f'_i \, f''_{jk}}{f_\circ} \Delta x_j^\circ \, \Delta x_k^\circ \right]. \qquad (16)
\end{aligned}$$

Заметим что Δx_p° является линейной, а Δx_p^1 – квадратичной, функциями от "базовых" флуктуаций $(\mathcal{N}_n - f_\circ)$.

Систематический сдвиг $p-$го параметра подгонки по методу максимального правдоподобия находится как среднее по ансамблю от выражения (16):

$$\langle \Delta x_p \rangle \quad = \quad \left\langle \Delta x_p^1 \right\rangle = -\frac{1}{2} \sum_{ijk} \left(\mathcal{A}^{-1}\right)_{pi} \left(\mathcal{A}^{-1}\right)_{jk} \sum_n \frac{f_i' \, f_{jk}''}{f_\circ}. \tag{17}$$

При усреднениях по ансамблю при выводе этого уравнения мы использовали соотношение (8): $\langle\, \Delta x_j^\circ \, \Delta x_k^\circ \,\rangle = \left(\mathcal{A}^{-1}\right)_{jk}$, а также соотношение

$$\begin{aligned} \langle\, (\mathcal{N}_n - f_\circ)\, \Delta x_j^\circ \,\rangle \quad &= \quad \sum_k \left(\mathcal{A}^{-1}\right)_{jk} \times \sum_m \left(\frac{f_k'}{f_\circ}\right)_m \langle\, (\mathcal{N}_n - f_\circ)(\mathcal{N}_m - f_\circ) \,\rangle \\ &= \quad \sum_k \left(\mathcal{A}^{-1}\right)_{jk} (f_k')_n \,. \end{aligned} \tag{18}$$

Отметим что при выводе уравнения (17) вклад от первого слагаемого в квадратных скобках в уравнении (16) сокращает вклад от второго слагаемого, а вклад от третьего сокращает вклад от четвертого. Для случая, когда функция f имеет только один параметр подгонки, уравнение (17) дает:

$$\langle \Delta x \rangle = -\frac{1}{2}\, \sigma^4 \sum_n \frac{f' \, f''}{f_\circ} \,. \tag{19}$$

Здесь мы использовали равенство $\sigma^2 = \left(\mathcal{A}^{-1}\right)$, которое следует из уравнения (10) для одномерного случая.

4 Систематический сдвиг параметров подгонки для метода χ^2 с $\sigma_n = \sqrt{\mathcal{N}_n}$

В результате статистических флуктуаций числа событий в индивидуальных каналах гистограммы, подгонка по $\chi^2 = \sum_n \frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n}$ будет давать значение параметров $\boldsymbol{x}$, сдвинутое по отношению к "истинному" значению $\boldsymbol{x}_\circ$ на некоторую величину $\Delta \boldsymbol{x}$: $\boldsymbol{x} = \boldsymbol{x}_\circ + \Delta \boldsymbol{x}$. Элементы вектора $\Delta \boldsymbol{x}$

могут быть найдены из условия минимизации χ^2, т.е. $\partial\chi^2/\partial x_j = 0$:

$$
\begin{aligned}
0 &= \frac{\partial \chi^2}{\partial x_j} = 2\sum_n f'_j \, \frac{f - \mathcal{N}_n}{\mathcal{N}_n} = 2\sum_n \frac{1}{\mathcal{N}_n}\left(f'_j + \sum_i f''_{ji}\,\Delta x_i + ...\right) \\
&\times \left(f_\circ - \mathcal{N}_n + \sum_i f'_i\,\Delta x_i + \frac{1}{2}\sum_{ik} f''_{ik}\,\Delta x_i\,\Delta x_k + ...\right). \qquad (20)
\end{aligned}
$$

Для упрощения вычислений удобно избавиться от $\mathcal{N}_n$ в знаменателе в правой части уравнения (20):

$$
\begin{aligned}
0 &= \sum_n \frac{1}{f_\circ}\left(1 - \frac{\mathcal{N}_n - f_\circ}{f_\circ} + ...\right)\left(f'_j + \sum_i f''_{ji}\,\Delta x_i + ...\right) \\
&\times \left(f_\circ - \mathcal{N}_n + \sum_i f'_i\,\Delta x_i + \frac{1}{2}\sum_{ik} f''_{ik}\,\Delta x_i\,\Delta x_k + ...\right). \qquad (21)
\end{aligned}
$$

Как и прежде мы будем искать решение в виде последовательных приближений: $\Delta x_i = \Delta x_i^\circ + \Delta x_i^1 + ...$ с главным приближением

$$
\Delta x_i^\circ = \sum_j \left(\mathcal{A}^{-1}\right)_{ij} \sum_n \frac{f'_j}{f_\circ}(\mathcal{N}_n - f_\circ)\,. \qquad (22)
$$

Для следующего приближения Δx_i^1 мы получаем уравнение:

$$
\begin{aligned}
0 &= \sum_n \frac{1}{f_\circ}\left[f'_i\left(\sum_p f'_p\,\Delta x_p^1 + \frac{1}{2}\sum_{jk} f''_{jk}\Delta x_j^\circ\,\Delta x_k^\circ\right)\right. \\
&+ \left.\left(\sum_j f''_{ij}\,\Delta x_j^\circ - \frac{\mathcal{N}_n - f_\circ}{f_\circ}\,f'_i\right)\left((f_\circ - \mathcal{N}_n) + \sum_k f'_k\,\Delta x_k^\circ\right)\right], \qquad (23)
\end{aligned}
$$

которое имеет решение:

$$
\begin{aligned}
\Delta x_p^1 &= \sum_i \left(\mathcal{A}^{-1}\right)_{pi}\left[-\frac{1}{2}\sum_{jk}\sum_n \frac{f'_i\,f''_{jk}}{f_\circ}\Delta x_j^\circ\,\Delta x_k^\circ\right. \\
&+ \sum_j \sum_n \frac{f''_{ij}}{f_\circ}(\mathcal{N}_n - f_\circ)\,\Delta x_j^\circ - \sum_{jk}\sum_n \frac{f''_{ij}\,f'_k}{f_\circ}\Delta x_j^\circ\,\Delta x_k^\circ \\
&- \left.\sum_n \frac{f'_i}{f_\circ^2}(\mathcal{N}_n - f_\circ)^2 + \sum_j \sum_n \frac{f'_i\,f'_j}{f_\circ^2}(\mathcal{N}_n - f_\circ)\,\Delta x_j^\circ\right]. \qquad (24)
\end{aligned}
$$

Систематический сдвиг p−го параметра подгонки по методу $\chi^2_{\sigma_n=\sqrt{\mathcal{N}_n}}$ можно найти усреднением по ансамблю от выражения (24):

$$\begin{aligned}\langle \Delta x_p \rangle &= \left\langle \Delta x_p^1 \right\rangle = -\frac{1}{2} \sum_{ijk} \left(\mathcal{A}^{-1}\right)_{pi} \left(\mathcal{A}^{-1}\right)_{jk} \sum_n \frac{f_i' \, f_{jk}''}{f_\circ} \\ &- \sum_i \left(\mathcal{A}^{-1}\right)_{pi} \sum_n \frac{f_i'}{f_\circ} + \sum_{ijk} \left(\mathcal{A}^{-1}\right)_{pi} \left(\mathcal{A}^{-1}\right)_{jk} \sum_n \frac{f_i' \, f_j' \, f_k'}{f_\circ^2} \,.\end{aligned} \tag{25}$$

Заметим что вклад второго слагаемого в квадратных скобках в уравнении (24) сокращает вклад от третьего слагаемого. Для случая одного параметра подгонки уравнение (25) дает:

$$\langle \Delta x \rangle = -\frac{1}{2}\, \sigma^4 \sum_n \frac{f' \, f''}{f_\circ} - \sigma^2 \sum_n \frac{f'}{f_\circ} + \sigma^4 \sum_n \frac{f'^{\,3}}{f_\circ^2} \,. \tag{26}$$

5 Систематический сдвиг параметров подгонки для метода χ^2 с $\sigma_n = \sqrt{f}$

В результате статистических флуктуаций числа событий в индивидуальных каналах гистограммы, подгонка по $\chi^2 = \sum_n \frac{(f-\mathcal{N}_n)^2}{f}$ будет давать значение параметров $\boldsymbol{x}$, сдвинутое по отношению к “истинному” значению $\boldsymbol{x}_\circ$ на некоторую величину $\Delta\boldsymbol{x}$: $\boldsymbol{x} = \boldsymbol{x}_\circ + \Delta\boldsymbol{x}$. Элементы вектора $\Delta\boldsymbol{x}$ могут быть найдены из условия минимизации:

$$\begin{aligned} 0 &= \frac{\partial \chi^2}{\partial x_i} = 2 \sum_n \frac{f_i'}{f} \left(f - \mathcal{N}_n - \frac{1}{2} \frac{(f-\mathcal{N}_n)^2}{f} \right) \\ &= 2 \sum_n \left[\frac{f_i' + \sum_j f_{ij}'' \Delta x_j + \ldots}{f_\circ + \sum_j f_j' \Delta x_j + \ldots} \right. \\ &\times \left(f_\circ - \mathcal{N}_n + \sum_j f_j' \Delta x_j \; + \frac{1}{2} \sum_{jk} f_{jk}'' \Delta x_j \Delta x_k + \ldots \right. \\ &- \left. \left. \frac{1}{2} \frac{\left(f_\circ - \mathcal{N}_n + \sum_j f_j' \Delta x_j + \ldots \right)^2}{f_\circ + \ldots} \right) \right] \,. \end{aligned} \tag{27}$$

Мы ищем решение в виде последовательных приближений: $\Delta x_i = \Delta x_i^\circ + \Delta x_i^1 + \ldots$ с главным приближением

$$\Delta x_i^\circ = \sum_j \left(\mathcal{A}^{-1}\right)_{ij} \sum_n \frac{f_j'}{f_\circ}(\mathcal{N}_n - f_\circ)\,. \tag{28}$$

Для следующего приближения Δx_i^1 мы получаем уравнение:

$$\sum_n \frac{f_i'}{f_\circ}\left[\sum_p f_p'\,\Delta x_p^1 + \frac{1}{2}\sum_j\sum_k f_{jk}''\,\Delta x_j^\circ\,\Delta x_k^\circ - \frac{1}{2}\frac{\left(f_\circ - \mathcal{N}_n + \sum_j f_j'\Delta x_j^\circ\right)^2}{f_\circ}\right.$$
$$\left. + \left(-\frac{\sum_j f_j'\,\Delta x_j^\circ}{f_\circ} + \frac{\sum_j f_{ij}''\,\Delta x_j^\circ}{f_i'}\right)\left((f_\circ - \mathcal{N}_n) + \sum_k f_k'\,\Delta x_k^\circ\right)\right] = 0\,, \tag{29}$$

которое имеет решение:

$$\begin{aligned}
\Delta x_p^1 \;=\; & \sum_i \left(\mathcal{A}^{-1}\right)_{pi} \qquad\qquad (30)\\
\times\; & \left[-\sum_j\sum_n \frac{f_i'\,f_j'}{f_\circ^2}(\mathcal{N}_n - f_\circ)\,\Delta x_j^\circ + \sum_{jk}\sum_n \frac{f_i'\,f_j'\,f_k'}{f_\circ^2}\Delta x_j^\circ\,\Delta x_k^\circ\right.\\
& +\sum_j\sum_n \frac{f_{ij}''}{f_\circ}(\mathcal{N}_n - f_\circ)\,\Delta x_j^\circ - \sum_{jk}\sum_n \frac{f_{ij}''\,f_k'}{f_\circ}\Delta x_j^\circ\,\Delta x_k^\circ\\
& -\frac{1}{2}\sum_{jk}\sum_n \frac{f_i'\,f_{jk}''}{f_\circ}\Delta x_j^\circ\,\Delta x_k^\circ + \frac{1}{2}\sum_n \frac{f_i'}{f_\circ^2}(\mathcal{N}_n - f_\circ)^2\\
& \left.-\sum_j\sum_n \frac{f_i'\,f_j'}{f_\circ^2}(\mathcal{N}_n - f_\circ)\,\Delta x_j^\circ + \frac{1}{2}\sum_{jk}\sum_n \frac{f_i'\,f_j'\,f_k'}{f_\circ^2}\,\Delta x_j^\circ\Delta x_k^\circ\right].
\end{aligned}$$

Систематический сдвиг $p-$го параметра подгонки по методу $\chi^2_{\sigma_n=\sqrt{f}}$ можно найти усреднением по ансамблю от выражения (30):

$$\begin{aligned}
\langle \Delta x_p \rangle \;=\; & \left\langle \Delta x_p^1 \right\rangle\\
=\; & -\frac{1}{2}\sum_{ijk}\left(\mathcal{A}^{-1}\right)_{pi}\left(\mathcal{A}^{-1}\right)_{jk}\sum_n \frac{f_i'\,f_{jk}''}{f_\circ} + \frac{1}{2}\sum_i \left(\mathcal{A}^{-1}\right)_{pi}\sum_n \frac{f_i'}{f_\circ}\\
-\; & \frac{1}{2}\sum_{ijk}\left(\mathcal{A}^{-1}\right)_{pi}\left(\mathcal{A}^{-1}\right)_{jk}\sum_n \frac{f_i'\,f_j'\,f_k'}{f_\circ^2}\,. \qquad\qquad (31)
\end{aligned}$$

Заметим что вклад первого слагаемого в квадратных скобках в уравнении (30) сокращает вклад от второго слагаемого, а вклад третьего сокращает вклад от четвертого. Для случая, когда функция f имеет только один параметр подгонки, уравнение (31) дает:

$$\langle \Delta x \rangle = \frac{1}{\omega}\,\frac{\sigma^4}{2}\sum_n \frac{f'\,f''}{f_\circ} + \frac{1}{2}\,\sigma^2 \sum_n \frac{f'}{f_\circ} - \frac{1}{2}\,\sigma^4 \sum_n \frac{f'^{\,3}}{f_\circ^2}\,. \tag{32}$$

6 Оценка величины систематического сдвига параметров подгонки

[Оценка величины систематического сдвига
параметров подгонки]

Для аналитических вычислений и численных оценок удобно заменить в полученных выше соотношениях суммы на интегралы по схеме:

$$\sum_n (...) \approx \frac{1}{b}\int (...)\,dt = \frac{N}{\int f\,dt}\int (...)\,dt\,, \tag{33}$$

где b – ширина канала гистограммы, а $N = \sum_n \mathcal{N}_n \approx \sum_n f(\boldsymbol{x}, t_n) \approx \frac{1}{b}\int f\,dt$ – полное число событий в гистограмме. Для простоты мы ограничимся случаем одного параметра подгонки. Оценим вклад первого и третьего слагаемых в уравнении (32) для величины систематического сдвига параметров подгонки для метода χ^2 с $\sigma_n = \sqrt{f}$ следующим образом:

$$\frac{1}{2}\,\sigma^4 \sum_n \frac{f'\,f''}{f} \;\sim\; \frac{1}{N^2} \times \frac{N}{\int f\,dt}\int \frac{f'\,f''}{f}\,dt \;\sim\; \frac{1}{N}\,, \tag{34}$$

$$\frac{1}{2}\,\sigma^4 \sum_n \frac{f'^{\,3}}{f^2} \;\sim\; \frac{1}{N^2} \times \frac{N}{\int f\,dt}\int \frac{f'^{\,3}}{f^2}\,dt \;\sim\; \frac{1}{N}\,. \tag{35}$$

Здесь мы использовали $\sigma \sim \frac{1}{\sqrt{N}}$ и

$$\frac{1}{\int f\,dt}\int \frac{f'\,f''}{f}\,dt \;\sim\; \frac{1}{\int f\,dt}\int \frac{f'^{\,3}}{f^2}\,dt \;\sim\; 1\,. \tag{36}$$

Последнее соотношение можно аргументировать следующим образом. Представим функцию f как произведение некоторой “амплитуды” и “нормализованной формы”, тогда в соотношении (36) “амплитуда” сокращается, а то что остается является некоторой функцией “нормализованной формы". После интегрирования это дает некоторые комбинации подгоночных (и фиксированных, если таковые имеются) параметров. Эти комбинации могут содержать сравнительно большие или малые числа, например $(\tau\omega) \sim 100$ для эксперимента g-2, который мы будем обсуждать в следующем разделе, но чаще всего они порядка единицы[1].

В отличие от соотношений (34,35), второе слагаемое в уравнении (32) можно оценить как

$$\frac{1}{2}\,\sigma^2 \sum_n \frac{f'}{f} \;\sim\; \frac{1}{N} \times \sum_n 1 \;\sim\; \frac{N_{ch}}{N}\,, \tag{37}$$

где N_{ch} – число каналов гистограммы. Практически для любой гистограммы $N_{ch} \gg 1$ и поэтому в общем случае второе слагаемое в уравнении (32) значительно больше чем первое и второе слагаемые. Заметим что

- это второе, а также третье, слагаемые отсутствуют в уравнении (19) для систематического сдвига параметра подгонки для метода максимального правдоподобия;
- в уравнении (26) для сдвига для метода $\chi^2_{\sigma_n=\sqrt{N_n}}$ первое слагаемое совпадает с первым слагаемым в уравнении (32) для $\chi^2_{\sigma_n=\sqrt{f}}$, а второе и третье больше по абсолютной величине в 2 раза;
- единственный член в уравнении (19) совпадает с первыми слагаемыми в уравнениях (32) и (26).

Таким образом, с точки зрения минимизации систематического сдвига параметров подгонки:

- наилучшим является метод максимального правдоподобия;
- метод $\chi^2_{\sigma_n=\sqrt{f}}$ лучше чем (но того же порядка что и) метод $\chi^2_{\sigma_n=\sqrt{N_n}}$.

В качестве примера рассмотрим систематический сдвиг частоты ω при подгонке временного распределения распадных электронов в эксперименте по измерению величины $g-2$ мюона в BNL функцией

$$G(t) = N_\circ e^{-t/\tau}[1 + A\,\cos(\omega t + \phi)], \tag{38}$$

[1]заметим что, строго говоря, приведенные в данном разделе оценки правильнее применять к относительному сдвигу $\Delta x/x$ чем к абсолютному сдвигу Δx.

по методу $\chi^2_{\sigma_n=\sqrt{f}}$. Статистические свойства распределения (38) подробно изучались в нашей предыдущей работе [1]. Было получено выражение для статистических ошибок, в частности для частоты ω:

$$\sigma_\omega = \frac{\sqrt{2}}{\tau A \sqrt{N}} \tag{39}$$

Там же было показано, что при особом выборе точки начала отсчета времени, а именно так чтобы время начала гистограммы было $t_{start} = -\tau$, корреляции между параметрами исчезают и все параметры становятся статистически независимыми. Это сильно упрощает вычисления и, в частности, позволяет использовать более простое уравнение (32) для оценки систематического сдвига параметров подгонки.

Вычислим самое большое, а именно второе, слагаемое в уравнении (32) для $\Delta\omega/\omega$ для статистики в 1 миллион ($N = 10^6$) распадных электронов:

$$\begin{aligned}
\frac{\Delta\omega}{\omega} &= \frac{1}{\omega}\frac{\sigma_\omega^2}{2}\sum_n \frac{1}{G}\frac{\partial G}{\partial\omega} \qquad (40)\\
&\approx \frac{1}{2\omega}\left(\frac{\sqrt{2}}{\tau A\sqrt{N}}\right)^2 \frac{1}{b}\int_{-\tau}^{T_{max}} \frac{N_\circ e^{-t/\tau} A\, t\, \sin(\omega t+\phi)}{N_\circ e^{-t/\tau}[1 + A\cos(\omega t+\phi)]}\, dt\\
&\approx \frac{1}{\omega\tau^2 A N b}\int_{-\tau}^{T_{max}} t\, \sin(\omega t+\phi)\, dt\\
&= \frac{1}{\omega^3\tau^2 A N b}\left[-\omega t\, \cos(\omega t+\phi) + \sin(\omega t+\phi)\,\right]\Big|_{-\tau}^{T_{max}}\\
&\sim \frac{T_{max}}{\omega^2\tau^2 A N b} \approx \frac{1}{A}\,\frac{8\tau}{b}\,\epsilon^2\,\frac{1}{N}\\
&= 20 \times 430 \times 1.2\cdot 10^{-4} \times 10^{-6} \approx 10^{-6} = 1 \text{ ppm}\,,
\end{aligned}$$

где $b = 0.15\ \mu$s – ширина канала, $\epsilon = (\tau\omega)^{-1} = 0.0108$ и $A = 0.4$. Верхний предел интегрирования $T_{max} \sim 8\tau$ соответствует максимальному времени набора статистики в гистограмме.

Согласно уравнению (39), статистическая ошибка σ_ω для такой статистики составляет

$$\frac{\sigma_\omega}{\omega} = \frac{1}{\omega}\frac{\sqrt{2}}{\tau A\sqrt{N}} = 1.41 \times 1.08\cdot 10^{-2} \times 2.5 \times 10^{-3} = 3.8\cdot 10^{-5} = 38 \text{ ppm} \tag{41}$$

и значит систематический сдвиг частоты ω существенно меньше чем ее статистическая ошибка, а для большей статистики этот сдвиг будет еще менее заметным, т.к. отношение $\Delta\omega/\sigma_\omega$ падает как $1/\sqrt{N}$. Однако если мы имеем 1 миллиард ($N = 10^{12}$) событий и по каким-либо причинам хотим

- разделить всю статистику на, скажем, 1000 примерно равных частей ($N = 10^6$ в каждой);
- сделать независимую подгонку для каждой из этих частей и затем
- в качестве конечного результата для ω взять средневзвешенное,

то в итоге мы получим тот же систематический сдвиг частоты 1 ppm как и в уравнении (40), но теперь этот сдвиг будет одного порядка со статистической ошибкой взвешенного среднего, которая в данном случае равна $\sigma_\omega/\omega = 38\,\mathrm{ppm} \times \sqrt{10^9/10^{12}} = 1.2$ ppm. Очевидно что такая ситуация неприемлима. В реальности, однако, эти 1000 частей вряд ли будут полностью идентичны. В частности, почти наверняка они будут иметь разные фазы g-2 колебаний и поэтому их систематические сдвиги будут в большой степени взаимно сокращаться при усреднении с весами (или без). Вопрос: можно ли полностью полагаться на подобное сокращение в прецизионном эксперименте ?

Описанная выше ситуация отнюдь не является чисто академическим упражнением. Все в том же эксперименте g-2 статистика делится на 20 независимых частей по числу не вполне идентичных детекторов, в свою очередь эти 20 частей обычно делятся еще на 5-10 частей по числу групп заходов с однотипными условиями проведения эксперимента. В некоторых случаях, в основном в целях изучения систематических эффеков, статистика делится еще дальше, например, по энергии распадных электронов или по временным интервалам с одинаковой фазой когерентных бетатронных колебаний (coherent betatron oscillations, CBO). В результате, число разбиений статистики на “примерно одинаковые” части может легко перевалить за 1000 при полном числе событий в несколько миллиардов.

Для сравнения мы вычислим первое слагаемое для $\Delta\omega$ в уравнении (32), это слагаемое является общим для всех методов подгонки, рассмотренных выше. Для той же самой статистики в $N = 10^6$ распадных электронов мы имеем:

$$
\begin{aligned}
\frac{\Delta\omega}{\omega} &\approx \frac{1}{2\omega}\left(\frac{\sqrt{2}}{\tau A\sqrt{N}}\right)^4 \times \frac{N}{\int_{-\tau}^{\infty} N_\circ e^{-t/\tau}[1+A\cos(\omega t+\phi)]\,dt} \qquad (42)\\
&\times \int_{-\tau}^{\infty} \frac{N_\circ e^{-t/\tau} A\,t\,\sin(\omega t+\phi)\times N_\circ e^{-t/\tau} A\,t^2\,\cos(\omega t+\phi)}{N_\circ e^{-t/\tau}[1+A\cos(\omega t+\phi)]}\,dt\\
&\approx \frac{1}{\omega\tau^4 A^2 N}\times\frac{1}{\int_{-\tau}^{\infty} e^{-t/\tau}\,dt}\times\int_{-\tau}^{\infty} e^{-t/\tau}t^3\sin 2(\omega t+\phi)\,dt\\
&\approx \frac{1}{\omega\tau^4 A^2 N}\times\frac{1}{e\tau}\times\frac{1}{2\omega}\left[-e^{-t/\tau}t^3\cos 2(\omega t+\phi)\right]\Big|_{-\tau}^{\infty} \sim \frac{1}{2\omega^2\tau^2 A^2 N}\\
&= \frac{\epsilon^2}{2A^2N} = \frac{1.2\cdot 10^{-4}}{2\times 0.16\times 10^6} \approx 0.4\cdot 10^{-9} \equiv 0.4\ \mathrm{ppb}\,,
\end{aligned}
$$

что примерно в 2500 раз меньше чем в уравнении (40). Для эксперимента по измерению g-2 это пренебрежимо мало.

Для третьего слагаемого в уравнении (32) для систематического сдвига мы ожидаем тот же порядок величины что и для первого, т.е. ~ 1 ppb.

7 Возможное улучшение качества подгонки по χ^2

Качество подгонки по χ^2 при желании может быть улучшено до уровня подгонки по методу максимального правдоподобия. Наиболее очевидный способ – это сделать подгонку по $\chi^2_{\sigma_n=\sqrt{N_n}}$ и по $\chi^2_{\sigma_n=\sqrt{f}}$ и результаты усреднить с правильными весами. Конкретно для p–го параметра:

$$
x_p = \frac{2}{3}\,(x_p)_{\sigma_n=\sqrt{f}} + \frac{1}{3}\,(x_p)_{\sigma_n=\sqrt{N_n}}\,. \qquad (43)
$$

В этом случае для систематического сдвига получается:

$$\begin{aligned}
\Delta x_p &= \frac{2}{3}\times\left[-\frac{1}{2}\sum_{ijk}\left(\mathcal{A}^{-1}\right)_{pi}\left(\mathcal{A}^{-1}\right)_{jk}\sum_n\frac{f_i'\,f_{jk}''}{f_\circ}\right.\\
&+ \left.\frac{1}{2}\sum_i\left(\mathcal{A}^{-1}\right)_{pi}\sum_n\frac{f_i'}{f_\circ}-\frac{1}{2}\sum_{ijk}\left(\mathcal{A}^{-1}\right)_{pi}\left(\mathcal{A}^{-1}\right)_{jk}\sum_n\frac{f_i'\,f_j'\,f_k'}{f_\circ^2}\right]\\
&+ \frac{1}{3}\times\left[-\frac{1}{2}\sum_{ijk}\left(\mathcal{A}^{-1}\right)_{pi}\left(\mathcal{A}^{-1}\right)_{jk}\sum_n\frac{f_i'\,f_{jk}''}{f_\circ}-\sum_i\left(\mathcal{A}^{-1}\right)_{pi}\sum_n\frac{f_i'}{f_\circ}\right.\\
&+ \left.\sum_{ijk}\left(\mathcal{A}^{-1}\right)_{pi}\left(\mathcal{A}^{-1}\right)_{jk}\sum_n\frac{f_i'\,f_j'\,f_k'}{f_\circ^2}\right]\\
&= -\frac{1}{2}\sum_{ijk}\left(\mathcal{A}^{-1}\right)_{pi}\left(\mathcal{A}^{-1}\right)_{jk}\sum_n\frac{f_i'\,f_{jk}''}{f_\circ}\,, \qquad (44)
\end{aligned}$$

что совпадает со сдвигом при подгонке по методу максимального правдоподобия. Заметим что аналогичная комбинация самих функций $\chi^2_{\sigma_n=\sqrt{\mathcal{N}_n}}$ и $\chi^2_{\sigma_n=\sqrt{f}}$ дает

$$\begin{aligned}
\frac{2}{3}\,\chi^2_{\sigma_n=\sqrt{f}}+\frac{1}{3}\,\chi^2_{\sigma_n=\sqrt{\mathcal{N}_n}} &= \sum_n\frac{2}{3}\,\frac{(f-\mathcal{N}_n)^2}{f}+\frac{1}{3}\,\frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n}\\
&= \sum_n\left[\frac{2}{3}\left(\frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n}-\frac{(f-\mathcal{N}_n)^3}{\mathcal{N}_n^2}+\frac{(f-\mathcal{N}_n)^4}{\mathcal{N}_n^3}-\ldots\right)\right.\\
&+ \left.\frac{1}{3}\,\frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n}\right]\\
&= \sum_n\left[\frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n}-\frac{2}{3}\,\frac{(f-\mathcal{N}_n)^3}{\mathcal{N}_n^2}+\frac{2}{3}\,\frac{(f-\mathcal{N}_n)^4}{\mathcal{N}_n^3}-\ldots\right] \qquad (45)
\end{aligned}$$

что совпадает с разложением в ряд функции $2\,\mathcal{L}$ в уравнении (3) в двух первых, и наиболее важных, слагаемых. Заметим что поотдельности $\chi^2_{\sigma_n=\sqrt{f}}$ и $\chi^2_{\sigma_n=\sqrt{\mathcal{N}_n}}$ совпадают с $2\,\mathcal{L}$ только в одном, первом, слагаемом. Это наводит на мысль о том как можно усовершенствовать "стандартную" подгонку по χ^2: нужно модифицировать функцию χ^2 таким образом, чтобы она совпадала с разложением в ряд функции $2\,\mathcal{L}$ в первых двух (или более) слагаемых. Таким свойством, например, обладает функция, состоящая просто из этих двух слагаемых:

$$\chi^2_{corr} \equiv \sum_n \left(\frac{(f - \mathcal{N}_n)^2}{\mathcal{N}_n} - \frac{2}{3}\, \frac{(f - \mathcal{N}_n)^3}{\mathcal{N}_n^2} \right) \tag{46}$$

Мы обозначили эту функцию как χ^2_{corr}, т.е. "χ^2 corrected". Этот вариант, однако, требует установки новой (по крайней мере для пакета PAW/MINUIT) процедуры оптимизации. Более громоздкий, но, возможно, менее радикальный вариант – это χ^2 с $\sigma_n^2 = \mathcal{N}_n \times (f/\mathcal{N}_n)^{2/3}$:

$$\begin{aligned} \chi^2_{\sigma_n^2=\mathcal{N}_n \times (f/\mathcal{N}_n)^{2/3}} &= \sum_n \frac{(f - \mathcal{N}_n)^2}{\mathcal{N}_n} \left(\frac{f}{\mathcal{N}_n} \right)^{-2/3} \qquad (47) \\ &= \sum_n \frac{(f - \mathcal{N}_n)^2}{\mathcal{N}_n} \left(1 + \frac{f - \mathcal{N}_n}{\mathcal{N}_n} \right)^{-2/3} \\ &\approx \sum_n \left[\frac{(f - \mathcal{N}_n)^2}{\mathcal{N}_n} - \frac{2}{3}\, \frac{(f - \mathcal{N}_n)^3}{\mathcal{N}_n^2} + \frac{5}{6}\, \frac{(f - \mathcal{N}_n)^4}{\mathcal{N}_n^3} - ... \right], \end{aligned}$$

что также совпадает с разложением в ряд функции $2\,\mathcal{L}$ в первых двух слагаемых.

В Приложениях I и II мы покажем что, действительно, систематический сдвиг параметров подгонки для $\chi^2_{\sigma_n^2=\mathcal{N}_n \times (f/\mathcal{N}_n)^{2/3}}$ и для χ^2_{corr} получается таким же как и сдвиг для метода максимального правдоподобия, т.е. наиболее оптимальным.

8 Заключение

В настоящей работе мы рассмотрели систематический сдвиг параметров при подгонке экспериментальных гистограмм по методу максимального правдоподобия и по нескольким разновидностям метода χ^2. Здесь мы приводим краткий обзор полученных результатов для различных методов подгонки и функций оптимизации:

- функция правдоподобия $\mathcal{L} = \sum_n \left(\mathcal{N}_n \ln \frac{\mathcal{N}_n}{f} + f - \mathcal{N}_n \right)$, Раздел 2 : этот метод (метод максимального правдоподобия) дает наименьший сдвиг параметров, но требует вычисления логарифмов, что замедляет процесс оптимизации;

- $\chi^2_{corr} = \sum_n \left(\frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n} - \frac{2}{3}\, \frac{(f-\mathcal{N}_n)^3}{\mathcal{N}_n^2} \right)$, Приложение I :
дает такой же сдвиг параметров как и метод максимального правдоподобия, но требует установки новой процедуры минимизации;

- χ^2 с $\sigma_n^2 = \sigma_n^2(\mathcal{N}_n, f) = \mathcal{N}_n \times (f/\mathcal{N}_n)^{2/3}$, Приложение II :
дает такой же сдвиг параметров как и метод максимального правдоподобия, но требует либо установки новой процедуры минимизации, либо использования следующей итерационный процедуры:

  – сначала минимизировать χ^2 с $\sigma_n^2 = \mathcal{N}_n$ и тем самым найти значение подгоночной функции в первой итерации: $f_1 \equiv f(\boldsymbol{x}_1, t_n)$;
  – затем минимизировать χ^2 с $\sigma_n^2 = \sigma_n^2(\mathcal{N}_n, f_1)$ и найти значение подгоночной функции во второй итерации: $f_2 \equiv f(\boldsymbol{x}_2, t_n)$;
  – затем минимизировать χ^2 с $\sigma_n^2 = \sigma_n^2(\mathcal{N}_n, f_2)$ и найти значение подгоночной функции в третьей итерации: $f_3 \equiv f(\boldsymbol{x}_3, t_n)$;
  – продолжить этот процесс до полной сходимости параметров подгонки.

  Другой возможный недостаток этого метода – вычисление кубических корней, что обычно медленнее (?) чем умножение;

- χ^2 с $\sigma_n = \sqrt{f}$, Раздел 3 :
стандартный метод подгонки для эксперимента по измерению g-2 мюона. По сравнению с методом максимального правдоподобия дает сдвиг параметров больше в $\sim N_{ch}$ раз, где N_{ch} – число каналов гистограммы. Требует итерационной процедуры, описанной выше;

- χ^2 с $\sigma_n = \sqrt{\mathcal{N}_n}$, Раздел 2 :
дает сдвиг параметров в 2 раза больший, чем дает χ^2 с $\sigma_n = \sqrt{f}$. Однако, поскольку это "всего лишь" в 2 раза (т.е. "того же порядка"), в большинстве случаев его можно предпочесть χ^2 с $\sigma_n = \sqrt{f}$ так как он не требует итерационной процедуры;

- Комбинирование результатов минимизации χ^2 с $\sigma_n = \sqrt{f}$ и χ^2 с $\sigma_n = \sqrt{\mathcal{N}_n}$:

$$x_p \;=\; \frac{2}{3}\,(x_p)_{\sigma_n=\sqrt{f}} \;+\; \frac{1}{3}\,(x_p)_{\sigma_n=\sqrt{\mathcal{N}_n}}$$

  дает такой же сдвиг параметров как и метод максимального правдоподобия, требует итерационный процедуры.

9 Приложение I : Подгонка по χ^2 с $\sigma_n^2 = \mathcal{N}_n \times (f/\mathcal{N}_n)^{2/3}$

В результате статистических флуктуаций числа событий в индивидуальных каналах гистограммы, подгонка по $\chi^2 = \sum_n \frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n} \left(\frac{f}{\mathcal{N}_n}\right)^{-2/3}$ будет давать значение параметров $\boldsymbol{x}$, сдвинутое по отношению к "истинному" значению $\boldsymbol{x}_\circ$ на некоторую величину $\Delta\boldsymbol{x}$: $\boldsymbol{x} = \boldsymbol{x}_\circ + \Delta\boldsymbol{x}$. Элементы вектора $\Delta\boldsymbol{x}$ могут быть найдены из условия минимизации:

$$
\begin{aligned}
0 \;=\; & \frac{\partial\chi^2}{\partial x_j} = 2\sum_n f'_j \left[\frac{f-\mathcal{N}_n}{\mathcal{N}_n}\left(\frac{f}{\mathcal{N}_n}\right)^{-2/3} - \frac{1}{3}\frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n}\left(\frac{f}{\mathcal{N}_n}\right)^{-5/3}\right] \\
=\; & 2\sum_n f'_j\, f^{-2/3}\, \mathcal{N}_n^{-1/3} \left[f-\mathcal{N}_n - \frac{1}{3}\frac{(f-\mathcal{N}_n)^2}{f}\right] \\
=\; & 2\sum_n \left(f'_j + \sum_i f''_{ji}\,\Delta x_i + \ldots\right)\left(f_\circ + \sum_i f'_i\,\Delta x_i + \ldots\right)^{-2/3} [f_\circ + (\mathcal{N}_n - f_\circ)]^{-1/3} \\
\times & \left(f_\circ - \mathcal{N}_n + \sum_j f'_j\Delta x_j + \frac{1}{2}\sum_{jk} f''_{jk}\Delta x_j\Delta x_k + \ldots\right. \\
- & \left.\frac{1}{3}\frac{\left(f_\circ - \mathcal{N}_n + \sum_j f'_j\Delta x_j + \ldots\right)^2}{f_\circ + \ldots}\right) \\
=\; & 2\sum_n \frac{1}{f_\circ}\left(f'_j + \sum_i f''_{ji}\,\Delta x_i + \ldots\right)\left(1 - \frac{2}{3}\frac{\sum_i f'_i\,\Delta x_i + \ldots}{f_\circ}\right)\left(1 - \frac{1}{3}\frac{\mathcal{N}_n - f_\circ}{f_\circ}\right) \\
\times & \left(f_\circ - \mathcal{N}_n + \sum_j f'_j\Delta x_j + \frac{1}{2}\sum_{jk} f''_{jk}\Delta x_j\Delta x_k + \ldots\right. \\
- & \left.\frac{1}{3}\frac{\left(f_\circ - \mathcal{N}_n + \sum_j f'_j\Delta x_j + \ldots\right)^2}{f_\circ + \ldots}\right)\,.
\end{aligned}
\tag{48}
$$

Мы ищем решение в виде последовательных приближений: $\Delta x_i = \Delta x_i^\circ + \Delta x_i^1 + \ldots$ с главным приближением

$$
\Delta x_i^\circ = \sum_j \left(\mathcal{A}^{-1}\right)_{ij} \sum_n \frac{f'_j}{f_\circ}(\mathcal{N}_n - f_\circ)
\tag{49}
$$

Для следующего приближения Δx_i^1 мы получаем уравнение:

$$0 \;=\; \sum_n \frac{f_i'}{f_\circ}\left[\sum_p f_p'\,\Delta x_p^1 + \frac{1}{2}\sum_{jk} f_{jk}''\Delta x_j^\circ\,\Delta x_k^\circ - \frac{1}{3}\,\frac{\left(f_\circ - \mathcal{N}_n + \sum_j f_j'\Delta x_j^\circ\right)^2}{f_\circ}\right. \tag{50}$$

$$\left. + \;\left(\frac{\sum_j f_{ij}''\,\Delta x_j^\circ}{f_i'} - \frac{2}{3}\,\frac{\sum_j f_j'\Delta x_j^\circ}{f_\circ} - \frac{1}{3}\,\frac{\mathcal{N}_n - f_\circ}{f_\circ}\right)\left((f_\circ - \mathcal{N}_n) + \sum_k f_k'\,\Delta x_k^\circ\right)\right],$$

которое имеет решение:

$$\begin{aligned}
\Delta x_p^1 \;&=\; \sum_i \left(\mathcal{A}^{-1}\right)_{pi} \times \left[-\frac{1}{2}\sum_{jk}\sum_n \frac{f_i'\,f_{jk}''}{f_\circ}\Delta x_j^\circ\,\Delta x_k^\circ + \frac{1}{3}\sum_n \frac{f_i'}{f_\circ^2}\left(\mathcal{N}_n - f_\circ\right)^2 - \right.\\
&-\; \frac{2}{3}\sum_j\sum_n \frac{f_i'\,f_j'}{f_\circ^2}\left(\mathcal{N}_n - f_\circ\right)\Delta x_j^\circ + \frac{1}{3}\sum_{jk}\sum_n \frac{f_i'\,f_j'\,f_k'}{f_\circ^2}\,\Delta x_j^\circ\,\Delta x_k^\circ + \\
&+\; \sum_j\sum_n \frac{f_{ij}''}{f_\circ}(\mathcal{N}_n - f_\circ)\,\Delta x_j^\circ - \sum_{jk}\sum_n \frac{f_{ij}''\,f_k'}{f_\circ}\Delta x_j^\circ\,\Delta x_k^\circ - \\
&-\; \frac{2}{3}\sum_j\sum_n \frac{f_i'\,f_j'}{f_\circ^2}(\mathcal{N}_n - f_\circ)\,\Delta x_j^\circ + \frac{2}{3}\sum_{jk}\sum_n \frac{f_i'\,f_j'\,f_k'}{f_\circ^2}\Delta x_j^\circ\,\Delta x_k^\circ - \\
&-\; \left.\frac{1}{3}\sum_n \frac{f_i'}{f_\circ^2}(\mathcal{N}_n - f_\circ)^2 + \frac{1}{3}\sum_j\sum_n \frac{f_i'\,f_j'}{f_\circ^2}(\mathcal{N}_n - f_\circ)\,\Delta x_j^\circ\right]
\end{aligned} \tag{51}$$

В этом уравнении из десяти слагаемых в квадратных скобках второе и девятое сокращают друг друга. Затем, при вычислении систематического сдвига p−го параметра подгонки как среднее по ансамблю от уравнения (51), вклад от пятого слагаемого сокращает вклад от шестого, а вклад от суммы третьего, четвертого, седьмого, восьмого и десятого зануляется. Окончательный результат для $\langle \Delta x_p \rangle$:

$$\langle \Delta x_p \rangle \;=\; \left\langle \Delta x_p^1 \right\rangle = -\frac{1}{2}\sum_{ijk}\left(\mathcal{A}^{-1}\right)_{pi}\left(\mathcal{A}^{-1}\right)_{jk}\sum_n \frac{f_i'\,f_{jk}''}{f_\circ} \tag{52}$$

Для случая одного параметра подгонки уравнение (52) дает:

$$\langle \Delta x \rangle = -\frac{1}{2}\,\sigma^4 \sum_n \frac{f'\,f''}{f_\circ} \tag{53}$$

Уравнения (52) и (53) совпадают с уравнениями (17) и (19) для метода максимального правдоподобия и, следовательно, метод $\chi^2_{\sigma_n^2 = \mathcal{N}_n \times (f/\mathcal{N}_n)^{2/3}}$ дает такой же систематический сдвиг параметров подгонки что и метод максимального правдоподобия.

10 Приложение II : Подгонка по $\chi^2_{corr} = \sum_n \left(\frac{(f - \mathcal{N}_n)^2}{\mathcal{N}_n} - \frac{2}{3} \frac{(f - \mathcal{N}_n)^3}{\mathcal{N}_n^2} \right)$

В результате статистических флуктуаций числа событий в индивидуальных каналах гистограммы, подгонка по $\chi^2_{corr} = \sum_n \left[\frac{(f - \mathcal{N}_n)^2}{\mathcal{N}_n} - \frac{2}{3} \frac{(f - \mathcal{N}_n)^3}{\mathcal{N}_n^2} \right]$ будет давать значение параметров $\boldsymbol{x}$, сдвинутое по отношению к "истинному" значению $\boldsymbol{x}_\circ$ на некоторую величину $\Delta \boldsymbol{x}$: $\boldsymbol{x} = \boldsymbol{x}_\circ + \Delta \boldsymbol{x}$. Элементы вектора $\Delta \boldsymbol{x}$ могут быть найдены из условия минимизации:

$$
\begin{aligned}
0 &= \frac{\partial \chi^2_{corr}}{\partial x_j} = 2 \sum_n \frac{f'_j}{\mathcal{N}_n} \left[f - \mathcal{N}_n - \frac{(f - \mathcal{N}_n)^2}{\mathcal{N}_n} \right] \qquad (54) \\
&= 2 \sum_n \frac{1}{f_\circ} \left(1 - \frac{\mathcal{N}_n - f_\circ}{f_\circ} + ... \right) \left(f'_j + \sum_i f''_{ji} \, \Delta x_i + ... \right) \\
&\times \left(f_\circ - \mathcal{N}_n + \sum_j f'_j \Delta x_j \ + \frac{1}{2} \sum_{jk} f''_{jk} \Delta x_j \Delta x_k + ... \right. \\
&- \left. \frac{\left(f_\circ - \mathcal{N}_n + \sum_j f'_j \Delta x_j + ... \right)^2}{f_\circ + ...} \right) .
\end{aligned}
$$

Мы ищем решение в виде последовательных приближений: $\Delta x_i = \Delta x_i^\circ + \Delta x_i^1 + ...$ с главным приближением

$$
\Delta x_i^\circ = \sum_j \left(\mathcal{A}^{-1} \right)_{ij} \sum_n \frac{f'_j}{f_\circ} (\mathcal{N}_n - f_\circ) \, . \qquad (55)
$$

Для следующего приближения Δx_i^1 мы получаем уравнение:

$$
\begin{aligned}
0 \, = \, \sum_n \frac{f'_i}{f_\circ} & \left[\sum_p f'_p \, \Delta x_p^1 + \frac{1}{2} \sum_{jk} f''_{jk} \Delta x_j^\circ \, \Delta x_k^\circ - \frac{\left(f_\circ - \mathcal{N}_n + \sum_j f'_j \Delta x_j^\circ \right)^2}{f_\circ} \right. \\
+ & \left. \left(\frac{\sum_j f''_{ij} \, \Delta x_j^\circ}{f'_i} - \frac{\mathcal{N}_n - f_\circ}{f_\circ} \right) \left((f_\circ - \mathcal{N}_n) + \sum_k f'_k \, \Delta x_k^\circ \right) \right] \qquad (56)
\end{aligned}
$$

которое имеет решение:

$$\begin{aligned}\Delta x_p^1 &= \sum_i \left(\mathcal{A}^{-1}\right)_{pi} \times \left[-\frac{1}{2} \sum_{jk} \sum_n \frac{f'_i\, f''_{jk}}{f_\circ} \Delta x_j^\circ\, \Delta x_k^\circ + \sum_n \frac{f'_i}{f_\circ^2} \left(\mathcal{N}_n - f_\circ\right)^2 \right. \\ &- 2 \sum_j \sum_n \frac{f'_i\, f'_j}{f_\circ^2} \left(\mathcal{N}_n - f_\circ\right) \Delta x_j^\circ + \sum_{jk} \sum_n \frac{f'_i\, f'_j\, f'_k}{f_\circ^2} \Delta x_j^\circ\, \Delta x_k^\circ \\ &+ \sum_j \sum_n \frac{f''_{ij}}{f_\circ} \left(\mathcal{N}_n - f_\circ\right) \Delta x_j^\circ - \sum_{jk} \sum_n \frac{f''_{ij}\, f'_k}{f_\circ} \Delta x_j^\circ\, \Delta x_k^\circ \\ &- \left. \sum_n \frac{f'_i}{f_\circ^2} \left(\mathcal{N}_n - f_\circ\right)^2 + \sum_j \sum_n \frac{f'_i\, f'_j}{f_\circ^2} \left(\mathcal{N}_n - f_\circ\right) \Delta x_j^\circ \right]. \end{aligned} \tag{57}$$

В этом уравнении из восьми слагаемых в квадратных скобках второе и седьмое сокращают друг друга. Затем, при вычислении систематического сдвига $p-$го параметра подгонки как среднее по ансамблю от уравнения (57), вклад от пятого слагаемого сокращает вклад от шестого, а вклад от суммы третьего, четвертого и восьмого зануляется. Окончательный результат для $\langle \Delta x_p \rangle$:

$$\langle \Delta x_p \rangle = \left\langle \Delta x_p^1 \right\rangle = -\frac{1}{2} \sum_{ijk} \left(\mathcal{A}^{-1}\right)_{pi} \left(\mathcal{A}^{-1}\right)_{jk} \sum_n \frac{f'_i\, f''_{jk}}{f_\circ} \tag{58}$$

Для случая одного параметра подгонки уравнение (58) дает:

$$\langle \Delta x \rangle = -\frac{1}{2} \sigma^4 \sum_n \frac{f'\, f''}{f_\circ} \tag{59}$$

Уравнения (58) и (59) совпадают с уравнениями (17) и (19) для метода максимального правдоподобия и, следовательно, метод $\chi^2_{corr} = \sum_n \left[\frac{(f-\mathcal{N}_n)^2}{\mathcal{N}_n} - \frac{2}{3} \frac{(f-\mathcal{N}_n)^3}{\mathcal{N}_n^2} \right]$ дает такой же систематический сдвиг параметров подгонки что и метод максимального правдоподобия.

Список литературы

[1] С.И. Редин, Статистические ошибки и корреляции параметров при аппроксимации экспериментальных данных, Препринт ИЯФ СО РАН 2000-97, Новосибирск, 2000.

С.И. Редин

Систематический сдвиг параметров при аппроксимации экспериментальных данных

S.I. Redin

Fit of experimental histograms: bias of fit parameters

Budker INP 2001-73

Ответственный за выпуск А.М. Кудрявцев
Работа поступила 26.12.2002 г.

Сдано в набор 30.12.2002 г.
Подписано в печать 7.03.2001 г.
Формат бумаги 60×90 1/16 Объем 1.4 печ.л., 1.1 уч.-изд.л.
Тираж 105 экз. Бесплатно. Заказ ь 73

Обработано на IBM PC и отпечатано на
ротапринте ИЯФ им. Г.И. Будкера СО РАН
Новосибирск, 630090, пр. академика Лаврентьева, 11.