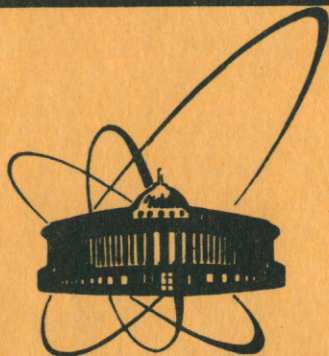


85-884



СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ
ДУБНА

P10-85-884

Б.З.Белашев*, Л.М.Сороко

МЕТОД МАКСИМУМА ЭНТРОПИИ
И ПУАССОНОВСКИЙ ШУМ

* Институт геологии Карельского филиала АН СССР,
Петрозаводск

1985

Описан метод максимума энтропии (ММЭ) решения обратных задач для пуассоновского шума. Метод был использован для устранения приборного уширения при малой статистике счета. Приведены результаты моделирования на ЭВМ метода максимума энтропии для входных сигналов, содержащих слившиеся дублеты, и для квазилинейных входных сигналов. Показано, что метод максимума энтропии дает хорошие оценки при малой статистике счета, $\sim 10 \pm 20$, для широкого класса входных сигналов и не требует каких-либо модификаций даже тогда, когда размытие описывается широкими пространственно-инвариантной или пространственно-неинвариантной функциями. Рассмотрено остаточное размытие в оценке искомого сигнала методом максимума энтропии. Показано, что ММЭ может быть использован для решения обратных задач в физике высоких энергий.

Работа выполнена в Лаборатории ядерных проблем ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1985

Перевод О.С.Виноградовой

The method of maximum entropy (MENT) for solving the inverse problems with Poisson-like noise is described. It has been used for removing of the instrumental spreading at low events statistics. The results of the computer simulation of the MENT for inputs signals, containing unresolved doublets, and for quasi-linear input signals are presented. The MENT to be shown gives good estimations at low events statistics, $\sim 10 \pm 20$, for wide class of input signals and does not require any amelioration even in the case of very broad spreading described by the space-invariant or space-variant functions. The residual spreading of the MENT estimations is analyzed. It is shown that the MENT developed in this paper for the case of Poisson-Like noise can be used for solving the inverse problems in the high energy physics.

The investigation has been performed at the Laboratory of Nuclear Problems, JINR.

ВВЕДЕНИЕ

Методом максимума энтропии (ММЭ) называют нелинейный метод решения обратных задач и задач спектрального анализа, когда оценку искомого сигнала ищут путем устремления к максимуму энтропийного функционала задачи. Впервые ММЭ был использован в спектральном анализе ^{/1/}. При помощи ММЭ функция корреляции искомого сигнала была проэкстраполирована за пределы измеряемого окна, были устранены боковые лепестки функции размытия и достигнуто разрешение в оценке искомого сигнала, которое превышало разрешение классических методов спектрального анализа. Затем ММЭ использовался для решения обратных задач с аддитивным шумом ^{/2/}. При этом обнаружены важные преимущества ММЭ ^{/3/}, а именно: разрешение и устойчивость оценки к шумам были выше, чем у классических методов решения обратных задач с регуляризацией ^{/4/}. После разработки ММЭ для обратных задач с пуассоновским шумом ^{/5/} появилась возможность использования метода в физике частиц высоких энергий.

Успешное использование ММЭ в радиоастрономии ^{/6/}, реконструктивной томографии ^{/7,8/} и в ряде других задач ^{/9/} инициировали попытки создать логически замкнутый фундамент ММЭ и вывести его из класса эвристических методов ^{/10-15/}.

В работе изложен ММЭ решения обратных задач для пуассоновского шума. Разработанный метод был использован для устранения приборного уширения при малой статистике счета. Приведены результаты моделирования на ЭВМ метода максимума энтропии для входных сигналов, содержащих слившиеся дублеты, и для квазилинейных входных сигналов. Показано, что метод максимума энтропии дает хорошие оценки при малой статистике счета, $\sim 10 \pm 20$, для широкого класса входных сигналов, и не требует каких-либо модификаций даже тогда, когда размытие прибора описывается широкими пространственно-инвариантной или пространственно-неинвариантной функциями. Рассмотрено остаточное размытие в оценке искомого сигнала методом максимума энтропии. Показано, что ММЭ может быть успешно использован для решения обратных задач в физике высоких энергий.

ТЕОРИЯ МЕТОДА МАКСИМУМА ЭНТРОПИИ

Запишем исходную обратную задачу в виде:

$$s(x) = \sum_{\xi=1}^{\ell} h(x, \xi) \hat{f}(\xi) + n(x),$$

$$x = 1, 2, \dots, M; \quad \xi = 1, 2, \dots, \ell,$$
(1)

где h — линейный оператор, $\hat{f}(\xi)$ — оценка искомого сигнала, $s(x)$ — измеренные данные, а $n(x)$ — шум. Рассмотрим подкласс обратных задач, когда шум $n(x)$ является пуассоновским. Искомый сигнал $\hat{f}(\xi)$ и шум $n(x)$ связаны между собой функционально-статистически через совместное статистическое распределение

$$P(\hat{f}, n) = P_1(\hat{f}) \cdot P_2(n/\hat{f}),$$
(2)

где функция распределения вероятности $P_1(\hat{f})$ описывает не зависящую от пуассоновского шума статистику искомого сигнала, при этом ее можно выразить через безусловный статистический фактор

$$P_1(\hat{f}) = \left(\frac{N}{\Delta f}\right)! / \prod_{\xi=1}^{\ell} \left[\frac{\hat{f}(\xi)}{\Delta f}\right]!,$$
(3)

где Δf — эффективный инкремент дискретизации сигнала \hat{f} , равный расстоянию между двумя уровнями квантования, а $N = \sum_{\xi=1}^{\ell} \hat{f}(\xi)$.

Функция распределения вероятности $P_2(n/\hat{f})$ описывает только флуктуации, которые зависят от сигнала \hat{f} . Так как пуассоновский шум $n(x)$ связан однозначно в статистическом смысле с измеряемым сигналом $s(x)$ при неизменных значениях $\hat{f}(\xi)$ и $h(x, \xi)$, то

$$P_2(n/\hat{f}) = P_2(s/\hat{f}).$$
(4)

Компоненты измеряемого сигнала $s(x)$ и искомого сигнала $\hat{f}(\xi)$ статистически независимы в различных точках x и ξ . Поэтому $P_2(s/\hat{f})$ можно записать в виде:

$$P_2(s/\hat{f}) = \prod_{x=1}^m [\nu_0(x)]^{\nu(x)} \exp[-\nu_0(x)] / \nu(x)!,$$
(5)

где $\nu(x) = s(x) / \Delta s$, $\nu_0(x) = [s(x) - n(x)] / \Delta x$, Δs — инкремент дискретизации измеряемого сигнала $s(x)$.

Накладываемое условие максимума энтропии в такой задаче эквивалентно условию:

$$\ln P_1(\hat{f}) + \ln P_2(s/\hat{f}) -$$

$$- \sum_{x=0}^m \Lambda(x) \left[\sum_{\xi} h(x, \xi) \hat{f}(\xi) - s(x) \right] \Rightarrow \max,$$
(6)

где $\Lambda(x)$ — множители Лагранжа вариационной задачи на условный экстремум.

Полученное путем решения вариационной задачи (6) явное выражение для оценки искомого сигнала $\hat{f}(\xi)$ имеет вид

$$\hat{f}(\xi) = \Delta f \cdot \exp \left[-1 - \sum_{x=0}^m \lambda(x) h(x, \xi) \right].$$
(7)

Запишем также выражение для оценки измеряемого сигнала $s(x)$:

$$s(x) = \left[\sum_{\xi=1}^{\ell} h(x, \xi) \hat{f}(\xi) \exp \left[-1 - \frac{\lambda(x) + \lambda(0)}{\rho} \right] \right],$$

$$s(0) = \sum_{\xi=1}^{\ell} h(0, \xi) \hat{f}(\xi), \quad x = 0.$$
(8)

В (7) и (8) входят переопределенные множители Лагранжа

$$\lambda(x) = \Lambda(x) \cdot \Delta f, \quad x = 0, 1, \dots, m,$$
(9)

$$\rho = \Delta f / \Delta s.$$

Система из $(m+1)$ нелинейных уравнений (8) будет использована для нахождения $(m+1)$ множителей Лагранжа при фиксированном значении параметра ρ , а следовательно, и Δf , так как Δs нам известно. По найденным множителям Лагранжа $\lambda(x)$ и инкременту квантования искомого сигнала Δf находим при помощи (7) оценку искомого сигнала. Выражение для оценки сигнала $\hat{f}(\xi) / \Delta f$ и шума

$$n(x) = \exp \left[-1 - \frac{\lambda(x) + \lambda(0)}{\rho} \right]$$
(10)

имеет такой же вид, как и для аддитивного шума, не зависящего от сигнала¹⁴. Однако в случае обратной задачи с пуассоновским шумом оценка шума входит явным образом в выражение для оценки измеряемого сигнала (8), причем эта связь имеет мультипликативную структуру.

МОДЕЛИРОВАНИЕ НА ЭВМ

Моделирование ММЭ на ЭВМ было проведено для двух классов входных сигналов. Первый класс образовали сигналы, содержащие тесный дублет из резонансов, а второй класс — сигналы, линейно зависящие от независимой переменной. Моделирование велось следующим образом. Сначала исходный сигнал $f(\xi)$ подвергался размытию в приборе с заданной функцией размытия $h(x, \xi)$. Затем к полученному результату добавлялся пуассоновский шум со средним значением нуля

и с дисперсией, определяемой статистическими флуктуациями. Пуассоновский шум ограничивался по амплитуде, равной двум стандартным отклонениям.

Построенная таким образом реализация измеренного сигнала $s(x)$ использовалась при нахождении оценки искомого сигнала $\hat{f}(\xi)$ при помощи ММЭ.

Число отсчетов в исходном и измеренном сигналах равнялось $m = l = 16$, а функция размытия измерительного прибора нормировалась на единичную площадь под соответствующей кривой. Чтобы сжать динамический диапазон множителей Лагранжа, исходные сигналы нормировались на определенные значения, которые зависели от статистики счета. Найденные оценки искомого сигнала затем умножались на это значение. Параметр ρ равнялся 1000. Число итераций, достаточное для получения устойчивой оценки путем решения системы нелинейных уравнений (8) методом Ньютона, составляло 120.

Сначала были исследованы возможности ММЭ давать оценку искомого сигнала в виде тесного дублета при сильном размытии его в измерительном приборе. Число зарегистрированных частиц или квантов излучения было зафиксировано постоянным, варьировались только реализации статистических флуктуаций при указанном ограничении. На рис.1а дана функция размытия $h(x, \xi) = h(x - \xi)$ при $\xi = 10$, которая в этой задаче является пространственно-инвариантной. На рис.1б приведены сигналы на выходе прибора $s(x)$ для трех реализаций пуассоновского шума. Пунктиром показан сигнал, не подверженный воздействию статистических флуктуаций. На рис.1в показаны оценки сигнала $\hat{f}(\xi)$, соответствующие различным реализациям пуассоновского шума. Пунктиром отмечены интенсивности δ -компонент исходного сигнала в виде дублета.

Из рис.1 видно, что ММЭ восстанавливает четко слившиеся компоненты тесного дублета для всех трех реализаций пуассоновского шума. Важно подчеркнуть, что эффективная ширина функции размытия в 2,5 раза превышает расщепление компонент дублета, а число частиц на один счет (или канал) равно 25 в максимуме регистрируемого сигнала $s(x)$. Невязка $[\hat{f}(x) - f(x)]$ равна по порядку величины погрешности измерений, равной стандартному отклонению. Это означает, что ММЭ не создает разбалтывания в решении обратной задачи с пуассоновским шумом.

Затем вариациям была подвергнута статистика счета. На рис.2а приведены сигналы $s(x)$, зарегистрированные прибором при скоростях счета на отсчет (канал) в максимуме, равных 13, 25, 46 и 90. Пунктиром показаны сигналы, не подверженные воздействию статистических флуктуаций. Реализация пуассоновского шума была одной и той же для всех указанных скоростей счета на отсчет (канал). На рис.2б даны оценки сигнала $\hat{f}(x)$, полученные с помощью ММЭ в этих условиях. Пунктиром отмечены интенсивности компонент дублета в исходном сигнале. Видно, что только при самой низкой скорости счета на отсчет

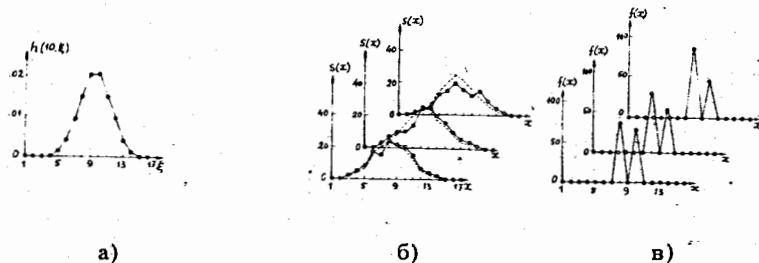


Рис.1. а — функция размытия $h(10, \xi)$; б — сигналы на выходе прибора $s(x)$ для трех реализаций пуассоновского шума и одной и той же статистике счета; в — оценки искомого сигнала $\hat{f}(x)$ в виде тесного дублета, полученные при помощи ММЭ для трех различных реализаций пуассоновского шума.

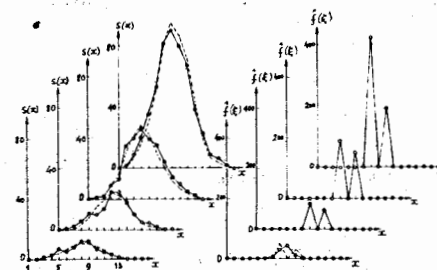


Рис.2. а — сигналы на выходе прибора $s(x)$ при четырех скоростях счета в максимуме: 13, 25, 46 и 90; б — оценки искомого сигнала $\hat{f}(\xi)$, имеющего вид тесного дублета.

(канал) в максимуме, равном 13, ММЭ не позволяет разделить слившиеся компоненты дублета.

Программа ММЭ с пуассоновским шумом была опробована затем на квазилинейных исходных сигналах при очень широких пространственно-неинвариантных функциях размытия. На рис.3а приведено семейство функций размытия $h(x, \xi)$ при различных значениях x , а на рис.3б — семейство из пяти регистрируемых сигналов при статистике счета в максимуме от 60 до 460. На рис.3в показаны полученные оценки искомого сигнала в виде точек с ошибками, которые соединены сплошной кривой, а также истинные сигналы в виде пунктирных линий. Видно, что ММЭ позволяет восстановить квазилинейные сигналы даже тогда, когда функция размытия является пространственно-неинвариантной и когда исходную задачу нельзя свести к свертке. Трудность промоделированной задачи состоит в том, что ширина функции размытия по порядку величины равна ширине области задания функций. С повышением статистики счета достоверность восстановления исходных сигналов возрастает. Данный пример восстановления квазилинейных сигналов при помощи ММЭ наглядно демонстрирует,

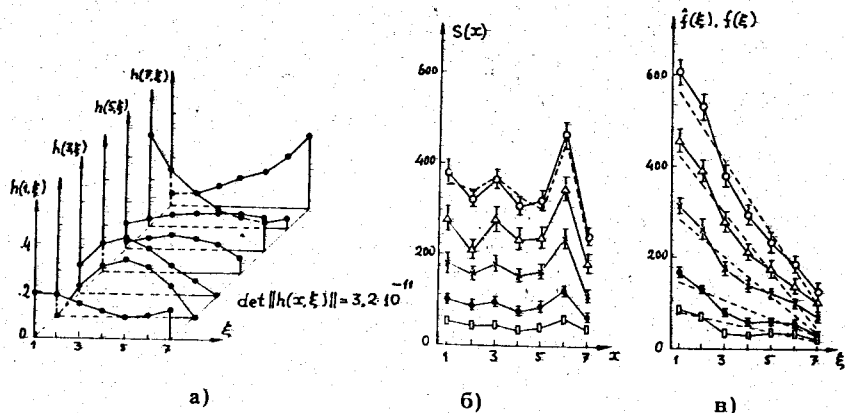


Рис.3. а — семейство функций $h(x, \xi)$ при различных значениях x . б — регистрируемые сигналы $s(x)$ при различной статистике счета, от 60 до 460; в — оценки искомого сигнала $\hat{f}(\xi)$ с ошибочный сигнал $f(\xi)$ (пунктир).

что данный метод можно использовать в физике нейтронов высоких энергий, где спектральный отклик различных детекторов нейтронов охватывает значительную часть всего спектра нейтронов [16].

НЕВЯЗКА И ОСТАТОЧНОЕ РАЗМЫТИЕ

Чтобы более четко обрисовать возможности ММЭ при решении обратных задач с пуассоновским шумом, были введены две количественные меры различия между оценкой сигнала, полученной ММЭ, и исходным сигналом. Первая из них — невязка, равная среднеквадратичному отклонению оценки искомого сигнала $\hat{f}(x)$ от исходного сигнала $f(x)$. Второй мерой является ширина остаточной функции размытия, равная, по определению, ширине оценки для δ -функции на входе. Были исследованы зависимости указанных количественных мер различия от статистики счета, от параметра δ и от ширины функции размытия прибора Γ .

Исследованы две функции размытия:

$$h_1(x, \xi) = \exp\left[-\frac{1}{2}|x - \xi|\right],$$

$$h_2(x, \xi) = \exp\left[-\frac{1}{2,5}|x - \xi|\right] \quad (11)$$

при одной и той же реализации пуассоновских шумов. Исходный сигнал имел вид δ -функции в точке $\xi = 8$ с интенсивностью 1,6 и с различными нормировками: $10^2, 10^3, 10^4, 10^5$ и 10^6 . Сигнал $s(x)$ также

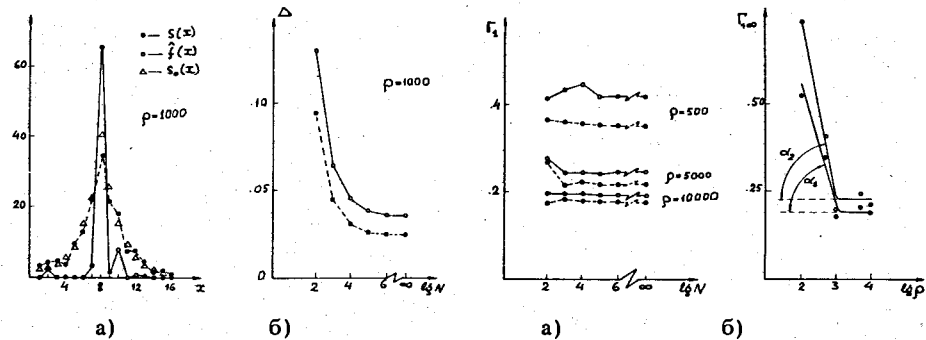


Рис.4. а — функции $s(x)$, $\hat{f}(x)$ и $s_0(x)$ (без шума), соответствующие функции размытия $h_1(x, \xi)$; б — невязка Δ в зависимости от N для двух различных функций размытия: \bullet — h_1 , \circ — h_2 .

Рис.5. а — зависимость ширины Γ_1 функции остаточного размытия при трех значениях параметра ρ для двух различных функций размытия: \bullet — h_1 , \circ — h_2 ; б — зависимость $\Gamma_{1\infty}$ от ρ для двух различных функций размытия: \bullet — h_1 , \circ — h_2 .

нормировался на соответствующие множители, а полученная ММЭ оценка сигнала сравнивалась с указанной δ -функцией. Значения параметра ρ равнялись: $10^2; 5 \cdot 10^2; 10^3; 5 \cdot 10^3; 10^4$.

На рис.4а показаны функции $s(x)$, $\hat{f}(x)$ и $s_0(x)$ (без шума) для функции размытия $h_1(x, \xi)$. На рис.4б даны значения среднеквадратичной оценки Δ при различном числе зарегистрированных частиц N для двух функций размытия в (11). Видно, что невязка Δ уменьшается с увеличением статистики счета. Однако при бесконечно большой статистике и нулевом уровне флуктуаций невязка Δ_∞ не обращается в нуль. Чтобы понять это явление, необходимо учесть то, что параметр ρ имеет конечное значение, а флуктуационная компонента входит мультипликативно в оценку данных наблюдений в модели ММЭ с пуассоновским шумом. При этом множители Лагранжа не могут одновременно обращаться в нуль. Поэтому оценка данных наблюдений

$$\left[\sum_{\xi=1}^{\ell} h(x, \xi) \hat{f}(\xi) \right] \exp\left[-1 - \frac{\lambda(x) + \lambda(0)}{\rho}\right].$$

при конечном значении параметра ρ отличается от истинной функции $s(x)$. При одной и той же статистике счета увеличение ширины размытия набора приводит к росту невязки.

На рис.5а показана зависимость ширины Γ_1 функции остаточного размытия в оценке δ -функции, полученная при помощи ММЭ, от статистики счета N для двух функций размытия (11) и для трех значений параметра ρ . Видно, что ширина Γ_1 остаточной функции размытия в оценке сигнала почти не зависит от статистики счета при больших N . На рис.5б приведена зависимость ширины $\Gamma_{1\infty}$ остаточной функции раз-

мытия при бесконечно большой статистике счета от параметра ρ . Видно, что $\Gamma_{1\infty}$ уменьшается с увеличением ρ и что при больших значениях параметра ρ определение малых ширин $\Gamma_{1\infty}$ затруднено.

ЛИТЕРАТУРА

1. Burg J.P. In: 37-th Meeting Soc. Explor. Geophys., 1967, Preprint.
2. Frieden B.R. J.Opt.Soc.Amer., 1972, 62, p.511.
3. Фриден Б. В кн.: Обработка изображений и цифровая фильтрация (под ред. Т.М.Хуанга). "Мир", М., 1979.
4. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. "Наука", М., 1979, 2-е изд.
5. Frieden B.R., Wells D.C. J.Opt.Soc.Amer., 1978, 68, p.93.
6. Wernecke S.J., D'Addario L.R. IEEE Trans., 1977, 26, p.351.
7. Minerbo G. Comp.Graph.Image Proc., 1979, 10, p.48.
8. Коттрел Г.А. Томография нейтральных пучков. Приборы для научных исследований, 1984, № 9, с.35.
9. Сороко Л.М. ЭЧАЯ, 1981, 12, вып.3, с.754.
10. Ulrich T.J. et al. J.Geophys.Res., 1973, 78, p.4959.
11. Haykin S. (Ed.). Nonlinear Methods of Spectral Analysis. Springer-Verlag, Berlin, 1979. (Topics in Applied Phys., vol.34).
12. Жуковский Е.Л. ДАН СССР, 1979, 246, № 5, с.1041.
13. Берсенев С.М., Жуковский Е.Л., Киселев В.М. Препринт ВЦ СО АН СССР, № 14, Красноярск, 1982.
14. Белашев Б.З., Сороко Л.М. ОИЯИ, P10-80-680, Дубна, 1980.
15. Белашев Б.З. и др. ОИЯИ, P10-82-101, Дубна, 1982.
16. Бекуру К., Вирту К. Нейтронная физика. Атомиздат, М., 1968, с.456.

Рукопись поступила в издательский отдел
9 декабря 1985 года.