Доклады Академии наук СССР
1973. Том 212, № 3

УДК 519.24 *МАТЕМАТИКА*

В. Ф. ТУРЧИН, Л. С. ТУРОВЦЕВА

МЕТОД СТАТИСТИЧЕСКОЙ РЕГУЛЯРИЗАЦИИ С АПОСТЕРИОРНОЙ ОЦЕНКОЙ ОШИБКИ ИСХОДНЫХ ДАННЫХ

(Представлено академиком А. Д. Сахаровым 29 V 1972)

Рассмотрим систему уравнений

$$f_j = K_j(\varphi_1, \ldots, \varphi_n), \quad j = 1, \ldots, m, \tag{1}$$

являющуюся конечномерным приближением некорректного уравнения

$$f = \hat{K}\varphi, \tag{2}$$

где f — заданная функция, φ — искомая функция, а $\hat{K}$ — некоторый оператор. Набор n чисел φ_i (вектор $\boldsymbol{\varphi}$) представляет функцию φ, а набор m чисел f_j (вектор $\mathbf{f}$) — функцию f. В одном из наиболее типичных случаев φ и f — функции одного действительного переменного, (2) — уравнение Фредгольма 1-го рода, а (1) — полученная с помощью квадратурных формул система m линейных уравнений с n неизвестными, в которой φ_i и f_j — значения функций φ и f в некоторых (опорных) точках.

Вследствие некорректности (2) точное решение уравнений (1), если оно существует, не представляет интереса, так как дает бессмысленный результат (сильно осциллирующую функцию). Для удовлетворительного решения некорректной задачи надо отказаться от требования, чтобы равенства (1) выполнялись с максимальной возможной точностью, и сделать дополнительное предположение об искомом векторе $\boldsymbol{\varphi}$, исключающее бессмысленные решения.

В методе статистической регуляризации * мы прежде всего формулируем задачу как задачу математической статистики. Будем считать, что имеют место не равенства (1), а равенства

$$f_j = K_j(\varphi_1, \ldots, \varphi_n) + \varepsilon_j, \tag{3}$$

где ε_j — случайные величины. Закон распределения величин ε_j будем считать известным, следовательно, известна и условная вероятность $P(\mathbf{f}|\boldsymbol{\varphi})$ при определенном векторе $\boldsymbol{\varphi}$ (описывающем некоторое состояние природы) получить в результате эксперимента вектор $\mathbf{f}$. Метод статистической регуляризации основан на байесовском подходе, при котором принимается некая априорная плотность вероятности $P(\boldsymbol{\varphi})$ для неизвестного $\boldsymbol{\varphi}$ и по формуле Байеса определяется апостериорная вероятность $P(\boldsymbol{\varphi}|\mathbf{f})$. Математическое ожидание $\boldsymbol{\varphi}$ по апостериорному распределению (байесову оценку $\boldsymbol{\varphi}$) обозначим через $\boldsymbol{\varphi}^{(0)}$ и назовем **регуляризованным решением системы** (1). Корень из дисперсии компоненты φ_i в апостериорном распределении дает среднеквадратичную ошибку решения в i-й точке.

Выбор той или иной конкретной функции $P(\boldsymbol{\varphi})$ всегда содержит элемент произвола, ибо никакой опыт не дает непосредственно $P(\boldsymbol{\varphi})$ и этот выбор оправдывается лишь апостериори успешным использованием метода, полученного на его основе, подобно тому, как любое уравнение, описы-

---

* Обзор предшествующих работ по методу статистической регуляризации и некоторых результатов его применения см. в ([1]).

вающее законы природы, может быть подтверждено лишь апостериори сравнением теоретических и экспериментальных данных.

Выберем $P(\varphi)$ таким образом, чтобы исключить из множества возможных решений негладкие функции φ. Для этого введем неотрицательно определенную симметричную $(n\times n)$-матрицу Ω такую, что квадратичная форма $(\varphi, \Omega\varphi)$ представляет конечно-разностное приближение к какому-либо функционалу, выражающему меру негладкости функции φ (например, к норме p-й производной). В качестве априорной плотности вероятности возьмем такую функцию $P(\varphi)$, которая, во-первых, приводит к заданному значению ω математического ожидания функционала $(\varphi, \Omega\varphi)$:

$$\int (\varphi, \Omega\varphi) P(\varphi)\, d\varphi = \omega,$$

а во-вторых, при условии выполнения указанного равенства содержит минимум информации о φ, т. е. максимизирует функционал

$$H[P(\varphi)] = -\int \ln P(\varphi) P(\varphi)\, d\varphi.$$

Нетрудно показать, что такая функция есть

$$P_\alpha(\varphi) = \mathrm{const}\cdot\alpha^{n/2}\cdot\exp\{-{}^1/{}_2\alpha(\varphi, \Omega\varphi)\}, \tag{4}$$

где $\alpha = n/\omega$. Если из каких-то априорных соображений можно задать определенное значение ω, естественно взять функцию (4) в качестве априорной плотности вероятности для φ.

Ограничимся случаем, когда уравнения (1) линейны, и будем обозначать через K_{ji} элементы соответствующей $(m\times n)$-матрицы. Кроме того предположим, что ошибки измерения независимы и распределены нормально с дисперсией s_j^2 для компоненты f_j. Тогда для условной вероятности получаем

$$P(|\mathbf{f}|\varphi) = \left(\prod_{j=1}^{m} \frac{1}{\sqrt{2\pi s_j^2}}\right) \exp\left\{-\frac{1}{2}\Delta[\mathbf{f}, \varphi]\right\}, \tag{5}$$

где

$$\Delta[\mathbf{f}, \varphi] = \sum_{j=1} \frac{1}{s_j^2}\left(f_j - \sum_{i=1} K_{ji}\varphi_i\right)^2. \tag{6}$$

В работе ([2]) показано, что в этих условиях регуляризованное решение $\varphi^0 = \varphi^\alpha$ удовлетворяет уравнению

$$(K^+WK + \alpha\Omega)\varphi^\alpha = K^+W\mathbf{f}, \tag{7}$$

где W — диагональная матрица с элементами $W_{jj} = s_j^{-2}$. Это уравнение было получено Филлипсом ([3]) из нестатистических (и поэтому, как показано в ([4]), приводящих к неправильному значению α) соображений, и независимо от него постулировано и исследовано А. Н. Тихоновым ([5]) как корректное уравнение, приближенно представляющее некорректное уравнение (2).

Если значение ω априори неизвестно, можно взять априорное распределение для φ в виде «слоистого ансамбля» с плотностью вероятности

$$P(\varphi) = \int_0^\infty P(\alpha) P_\alpha(\varphi)\, d\alpha, \tag{8}$$

где $P(\alpha)$ — априорная плотность вероятности параметра α — принимает постоянное значение в сколь угодно большой области значений α.

Как показано в ([6]), регуляризованное решение $\varphi^{(0)}$ при такой априорной информации о φ получается осреднением решения φ^α по апостериорному распределению для α, имеющему плотность вероятности

$$P(\alpha|\mathbf{f})=\mathrm{const}\cdot P(\alpha)\int P_\alpha(\varphi)P(\mathbf{f}|\varphi)\,d\varphi. \tag{9}$$

Если $P(\alpha|\mathbf{f})$ имеет резкий максимум при некотором $\alpha=\alpha_0$, то вместо осреднения по α можно положить $\varphi^{(0)}=\varphi^{\alpha_0}$. Значение α_0 есть фактически апостериорная оценка параметра α по методу наибольшего правдоподобия.

В ходе проверки метода статистической регуляризации было проведено много математических экспериментов, состоящих в следующем. Некая функция φ подвергалась воздействию интегрального оператора Фредгольма 1-го рода K и к результату добавлялась случайная функция, значения которой во всех точках независимы и распределены по нормальному закону (ошибка измерения). Полученная таким образом функция, имитирующая результат измерения, бралась в качестве правой части уравнения Фредгольма с ядром K и по изложенному методу находилось регуляризованное решение $\varphi^{(0)}$, которое сравнивалось с исходной («истинной») функцией φ.

Эксперименты показали, что при самых различных ядрах K и гладких функциях φ параметр α успешно определялся апостериори, а регуляризованное решение $\varphi^{(0)}$ с точностью до ошибки, указываемой теорией, совпадало с функцией φ. Это свидетельствует о том, что априорная информация, записываемая в виде слоистого ансамбля (8), с одной стороны, достаточна для устранения бессмысленных решений, а с другой стороны, столь невелика по объему и неспецифична, что может быть использована для решения чрезвычайно широкого класса некорректных задач. Существование такого ансамбля — нетривиальный факт, касающийся не собственно математического метода нахождения решения, а статистических свойств объектов, встречаемых и вычислительной практике.

Часто бывает, что набор чисел $s_1, s_2, \ldots, s_m$, характеризующий среднеквадратичную ошибку эксперимента, задан лишь с точностью до некоторого не зависящего от j множителя, который мы обозначим через $1/\sqrt{\beta}$. В этом случае можно попытаться уточнить ошибку апостериори. Подобно тому, как неопределенность параметра α выражалась введением интегрирования по α в априорную плотность вероятности, неопределенность параметра β мы выразим такой же «слоистой» плотностью условной вероятности $P(\mathbf{f}|\varphi)$, т. е. запишем ее в виде интеграла

$$P(\mathbf{f}|\varphi)=\int_0^\infty P(\beta)\left(\frac{\beta}{2\pi}\right)^{m/2}\exp\left\{-\frac{\beta}{2}\Delta[\mathbf{f},\varphi]\right\}d\beta, \tag{10}$$

где $P(\beta)$ будем считать принимающей постоянное значение в сколь угодно большой области значений β. (Величины s_j считаем отнормированными таким образом, что их произведение равно единице.)

С помощью формулы Бейеса снова получим апостериорное распределение $P(\varphi|\mathbf{f})$. Нетрудно показать, что найденное отсюда регуляризованное решение $\varphi^{(0)}$ будет результатом осреднения решения $\varphi^{\alpha\beta}$, соответствующего определенной паре значений (α, β) по апостериорной плотности вероятности $P(\alpha,\beta|\mathbf{f})$, которая выражается совершенно аналогично (9). Вычисляя интеграл, находим

$$P(\alpha,\beta|\mathbf{f})=\mathrm{const}\cdot\beta^{m/2}\times$$
$$\times\exp\left\{-\frac{\beta}{2}(\mathbf{f},[W-WKB^{-1}K^+W]\mathbf{f})\right\}\left(\frac{\alpha}{\beta}\right)^{n/2}\cdot(\mathrm{Det}\,B)^{-1/2}, \tag{11}$$

где

$$B\equiv K^+WK+\frac{\alpha}{\beta}\Omega.$$

Выражение, входящее в экспоненту со знаком минус, представляет собой минимум квадратичной формы

$$\frac{\beta}{2}\Delta[\mathbf{f}, \varphi]+\frac{\alpha}{2}(\varphi, \Omega\varphi),$$

поэтому оно заведомо не отрицательно.

Вопрос о том, можно ли заменить осреднение по α и β на вычисление $\varphi^{(0)}$ при некоторых $\alpha=\alpha_0$ и $\beta=\beta_0$, решается в зависимости от поведения функции $P(\alpha, \beta|\mathbf{f})$ в плоскости α, β. Если она имеет резкий максимум в точке (α_0, β_0), такая замена возможна. Значение β_0 дает статистическую оценку ошибки измерения. Серия математических экспериментов, аналогичных упомянутым выше, показала, что в типичных случаях решения уравнения Фредгольма 1-го рода при $m\geq 15$, функция $P(\alpha,\beta|\mathbf{f})$ содержит достаточно четкий максимум, приводящий к удовлетворительной оценке ошибки измерения. Пример функции $P(\alpha,\beta|\mathbf{f})$ приведен на рис. 1.

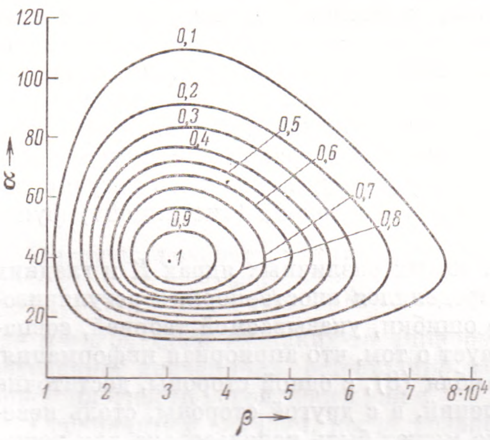

Рис. 1. Линии равных значений апостериорной вероятности $P(\alpha,\beta|\mathbf{f})$ в плоскости α, β. Значения приводятся в отношении к значению в максимуме

Возможность апостериорной оценки ошибки измерения является следствием введения априорной информации о функции φ. Из (11) видно, что на линии постоянного отношения α/β апостериорная вероятность $P(\alpha,\beta|\mathbf{f})$ является простой функцией β, имеющей максимум. Однако при $\alpha=0$ (предполагая наличие нерегуляризованного решения) максимум исчезает.

Авторы выражают глубокую благодарность чл.-корр. АН СССР И. М. Гельфанду, прочитавшему статью в рукописи и сделавшему ряд замечаний, которые позволили существенно улучшить изложение.

Институт прикладной математики
Академии наук СССР
Москва

Поступило
29 V 1972

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] В. Ф. Турчин, В. П. Козлов, М. С. Малкевич, УФН, **102**, в. 3, 345 (1970). [2] В. Ф. Турчин, Журн. вычислит. матем. и матем. физ., **7**, 1270 (1967). [3] D. L. Phillips, J. Assoc. Comp. Mach., **9**, № 1, 84 (1962). [4] В. Ф. Турчин, Журн. вычислит. матем. и матем. физ., 8, 230 (1968). [5] А. Н. Тихонов, ДАН, **151**, 501 (1963). [6] В. Ф. Турчин, В. З. Нозик, Изв. АН СССР, сер. физ. атмосферы и океана, 5, 29 (1969).