

## Некорректные обратные задачи экспериментальной физики

Известно определение некорректно поставленной задачи по Адамару: при бесконечно малом изменении начальных условий решение задачи меняется на конечную величину. Обратные задачи математической физики как правило некорректны.

Поясним, что мы понимаем под обратными задачами и в каком смысле они некорректны.

### прямая задача

Неизвестная функция  $f$  строится из функции  $\varphi$  по ходу причинно-следственных отношений:

$\varphi$ причина	Оператор $\Rightarrow$	$f$ следствие
----------------------	---------------------------	------------------

По известной причине необходимо определить следствия. Примеры прямых задач: по пространственно-временной картине зарядов определить поля или по заданному входному сигналу определить выходной.

### обратная задача

По известным следствиям  $f$  найти причину  $\varphi$ . Примеры: по картине рассеяния определить потенциал или входной сигнал по выходному.

В измерительных задачах важнейшую роль играет вероятностная природа наблюдаемых (“известных”) величин. Речь может идти о “шумах” либо о принципиальных флуктуациях при измерении потока квантовых частиц.

Выясним природу некорректности на типичном примере обратной задачи - уравнении Фредгольма I-го рода для функции  $\varphi(x)$ ,  $x \in [a, b]$

$$\int K(x, y) \varphi(x) dx = f(y) \quad y \in [c, d] \quad (1)$$

или в операторном виде  $\hat{K}\varphi = f$

Ядро уравнения  $K(x, y)$  и свободный член  $f(y)$  - “известные” функции. Интегральный оператор определяет переход  $\varphi \Rightarrow f$ .

Пример физической задачи:

по измеренной зависимости теплоемкости от температуры найти спектр энергетических состояний вещества, если известна вероятность возбуждения состояний при данной температуре:  $c_V(T) = \int g(\varepsilon, T) \omega(\varepsilon) d\varepsilon$

Для разностного ядра  $K(y-x)$  уравнение приобретает вид свертки  $\int_a^b K(y-x) \varphi(x) dx = f(y)$ : например, в задаче восстановления истинного спектра по аппаратурному с известной функцией разрешения.

Формально уравнение (1) не всегда имеет решение вообще и единственное решение в частности. Пример, для ядра вида  $K(x,y) = y+x$  решения нет, если  $f(y)$  - нелинейная функция. Для линейной функции  $f(y) = \lambda_0 + \lambda_1 y$  решение не единственно, т.к. любая функция  $\varphi(x)$ , удовлетворяющая соотношениям

$$\lambda_0 = \int_a^b x \varphi(x) dx \quad \text{и} \quad \lambda_1 = \int_a^b \varphi(x) dx \quad \text{есть решение (1)}$$

В задачах экспериментальной физики чаще всего существование и единственность решения следуют из физической природы  $\varphi(x)$  и ядра. Однако существование единственного закона природы  $\varphi(x)$  не делает корректной задачу его определения из уравнения (1). И проблема заключается в “сглаживающем” действии на функцию  $\varphi$  интегрального оператора  $K$ . Этот известный факт хорошо демонстрируется примером.

Рассмотрим два решения уравнения (1)

$\varphi_1(x)$  - некоторая гладкая функция

$$\varphi_2(x) = \varphi_1(x) + \frac{A}{(b-a)} \sin 2\pi nx$$

Во втором решении добавка к  $\varphi_1(x)$  осциллирующая функция с амплитудой  $A$  и с частотой  $n$ . Очевидно, что для любой сколь угодно большой амплитуды всегда найдется такая большая частота, что разность между “наблюдениями”

$$f_1(y) = \int_a^b K(x,y) \varphi_1(x) dx \quad \text{и} \quad f_2(y) = f_1(y) + \frac{A}{(b-a)} \int_a^b K(x,y) \sin 2\pi nx dx \quad \text{будет сколь}$$

угодно мала.

При каждом фиксированном значении  $y$  “наблюдения” становятся близкими. В физическом смысле это значит, что ядро, как фильтр, сглаживает высокие частоты сколь угодно большой амплитуды.

Если функция  $f$  известна точно, то существование близких решений не образуют проблемы: единственность решения, которая имеется в виду, всегда решает проблему. Но при наличии статистической флуктуации в результате наблюдения - функции  $f^*$  (а при экспериментальном ее определении только так и может быть!), ситуация меняется самым кардинальным образом. Если близость экспериментальной функции  $f^*$  к точному образу неизвестной функции  $\varphi$  ограничена (с точностью  $\varepsilon$ ) неравенством  $\max |f^*(y) - f(y)| \leq \varepsilon$ , то всегда найдутся два решения с осциллирующим членом такие, что  $f_1 = \hat{K}\varphi_1$  и  $f_2 = \hat{K}\varphi_2$  будут удовлетворять неравенству с одной и той же экспериментальной функцией  $f^*$ .

Следовательно, если постановка вопроса такова, что необходимо решить уравнение  $\hat{K}\varphi = f^*$  точно, то мы имеем все шансы получить  $\varphi^*(x)$  сколь угодно отличающееся от правильного закона природы  $\varphi$ , породившего  $f^*$  в стохастическом процессе измерений. Собственно в этом и заключается некорректность обратной задачи, а конкретное определение некорректности должно включать рассмотрение конкретной меры уклонения  $f^*$  от  $f$ .

В экспериментальной физике, как правило, функция плотности вероятности для  $f$  (пуассоновская, гауссовская...) такова, что точную грань  $\varepsilon$  указать нельзя, а если можно, то только с “запасом”. Интуитивно очевидно, что адекватным подходом к решению некорректных задач может быть подход вероятностный. В этом подходе необходимо изучить статистические свойства решения  $\varphi$  в зависимости от характера случайного процесса реализации  $f^*(y)$ .

Рассмотрим статистический подход на частном примере инструментальной задачи - преобразование спектра прибором с конечной функцией разрешения - уравнения свертки с случайной помехой

$$f_{exp}(y) = \int_{-\infty}^{\infty} R(y-x)\varphi(x)dx$$

Пусть  $f_{exp}(y)$  есть сумма  $f(y) + \delta(y)$  - точной функции и случайной помехи  $\delta(y)$  со средним равным нулю и корреляционной функцией  $\Delta(\eta) = \langle \delta(y) \cdot \delta(y+\eta) \rangle$ .

(Вообще предполагается, что известна ФПВ для  $\delta(y)$ , но пока достаточно корреляционной функции). Формальное решение уравнения свертки

$$\varphi(x) = \frac{1}{2} \int_{-\infty}^{\infty} \exp(ipx) \frac{\tilde{f}^*(p)}{\tilde{R}(p)} dp,$$

где  $\tilde{f}^*$  и  $\tilde{R}$  - фурье-образы  $f^*$  и  $R$ . При подстановке  $f^* = f + \delta$  показано (С.Г.Раутиан, УФН, 1958, т.66, стр.475), что дисперсия решения есть

$$\sigma_{\varphi}^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{G(p)}{|\tilde{R}(p)|^2} dp, \quad \text{где}$$

спектр мощности случайного шума  $G(p) = \tilde{\Delta}(p)$  - фурье-компонента корреляционной функции  $\delta$ .

Поскольку функция  $R(z=y-x)$  непрерывна -  $|\tilde{R}(p)| \rightarrow 0$  при  $|p| \rightarrow \infty$ . Для конечности дисперсии необходимо, чтобы спектр мощности флуктуаций падал быстрее, чем спектр мощности ядра уравнения. Однако в большинстве практических случаев спектр шума не только не падает, но при больших  $p$  выходит на константу, так как всегда содержит компоненту “белого шума”. Это обстоятельство ведет к тому, что дисперсия решения стремится к бесконечности.

Поясним физическую картину некорректности. Интегральный оператор, действуя на функцию  $\varphi$  (интересующий нас закон природы) в прямой задаче, сглаживает ее высокие гармоники, а оператор, действующий на экспериментальную функцию в обратной задаче увеличивает, как это и требуется, амплитуду истинных гармоник, но при этом раскачивает и ложные гармоники стохастической природы, не имеющие отношения к искомому закону природы. Иными словами, решение обратной задачи неустойчиво по отношению к малым флуктуациям. Заметим, что задача с уравнением Фредгольма II рода

$$f(y) = \varphi(y) + \int_a^b K(x,y)\varphi(x)dx \quad y \in [c,d]$$

корректна из-за наличия искомой функции в виде свободного члена.

Хотя бы потому что функции  $f$  бывает измерена лишь в конечном числе точек, интегральное уравнение (1) так или иначе для численного решения сводится в системе линейных уравнений при помощи подходящей квадратурной формулы - алгебраизуется.

$$\sum_{j=1}^n K_{ij} \varphi_j = f_i \quad i=1, m \quad (1')$$

Числа  $n$  и  $m$  определяют точность этой алгебраизации. Что же происходит при численном решении уравнения (1), т.е. системы (1')? Не станет ли оно корректным? Ответ на этот вопрос, к сожалению, отрицательный из-за плохой обусловленности системы. Как показано Фадеевым, обусловленность связана с собственными числами  $\mu_k$  матрицы  $K^+ K$  и падает с ростом отношения  $\mu_{\max} / \mu_{\min}$ . Решения системы "разбалтываются" в направлении собственных векторов матрицы  $K^+ K$ , которые соответствуют малым  $\mu_k$ . Даже если бы функция  $f$  была задана точно, с ростом порядка матрицы (уменьшением шага по  $x$  и по  $y$ ) падает точность при ее обращении. Иными словами, малое изменение в любом элементе матрицы  $K$  ведет к большим изменениям элементов обратной матрицы.

Таким образом, возникает дилемма:

малые  $n, m$  - плохая точность описания;

большие - неустойчивость

### Нестатистические методы решения некорректных задач

Поскольку при классическом понимании некорректности любая функция  $\varphi$ , удовлетворяющая условию

$$r_\varepsilon(f^*, \hat{K}\varphi) \leq \varepsilon, \text{ где } r - \text{мера уклонения,}$$

есть решение, то удалить "плохие" решения (и тем самым регуляризовать решение), можно доопределив задачу. Заметим, что любой способ доопределения (так или иначе прибегает к априорной информации о  $\varphi$ . Филлипс<sup>1</sup> предложил искать самую "гладкую" среди  $\varphi$  функцию, удовлетворяющую условию на меру уклонения, т.е. ту,

что доставляет минимум функционалу  $\int_a^b \left( \frac{d\varphi}{dx} \right)^2 dx$ . Обычно этот минимум сидит на

границе  $r_\varepsilon(f^*, \hat{K}\varphi) = \varepsilon$  и тогда возникает задача на условный экстремум с множителем Лагранжа

---

<sup>1</sup> (D.L.Phillips, J.Associat.Comput.Machin. 9, 84, 9(1962))

$$r_{\varepsilon}(f^*, \hat{K}\varphi) + \alpha \int_a^b \left( \frac{d\varphi}{dx} \right)^2 dx = \min$$

$\alpha$  - есть параметр регуляризации. Смысл его заключается в том, что при больших  $\alpha$  вес "гладкости" при поиске минимума превалирует над "близостью"  $f^*$  к  $f = \hat{K}\varphi$ . Значение  $\alpha$  находится из условия  $r_{\varepsilon}(f^*, \hat{K}\varphi) = \varepsilon$ . Филлипс обнаружил, что найденное им решение, как правило, сильно заглажено и, таким образом, параметр регуляризации остается не определенным.

Общее понятие регуляризации решения некорректной задачи ввел А.Н.Тихонов. Он доказал теорему о том, что можно построить такое семейство корректных решений, зависящих от  $\alpha$ , что при  $\alpha \rightarrow 0$  и, одновременно,  $\varepsilon \rightarrow 0$ , семейство сходится к истинному  $\varphi$ . При этом уравнение

$$r_{\varepsilon}(f^*, \hat{K}\varphi) + \alpha \int_a^b \left( \frac{d\varphi}{dx} \right)^2 dx = \min$$

постулировано. При практическом применении результаты метода совпадают с методом Филлипса с неопределенным  $\alpha$ .

В литературе существует множество подходов к решению некорректных задач. В духе нашего курса мы займемся в следующем параграфе статистическим методом, идея которого принадлежит В.Турчину. Этот метод основан на байесовском подходе к задачам принятия решения в условиях неопределенности и представляется наиболее ясным и последовательным.

### Метод статистической регуляризации некорректных обратны для уравнения Фредгольма I-го рода <sup>2</sup>

Мы будем рассматривать какое-либо конечно-разностное приближение (алгебраизацию) уравнения Фредгольма I рода

$$\int_a^b K(x, y)\varphi(y)dy = f(x) \quad (1)$$

---

<sup>2</sup> В.Ф.Турчин

В.Ф.Турчин, Козлов, Малкевич, УФН

В.Ф. Турчин, В.З.Нозик, Известия АН Серия Физика атмосферы и океана,

Здесь  $\varphi(x)$  - неизвестное "состояние природы" - закон, который мы хотим узнать,  $K(x,y)$  - известное ядро уравнения, преобразующее закон в измеряемую функцию  $f(y)$ .

Если эта функция суть результат измерения, то в конечном числе  $M$  точек  $y_m$  имеется набор случайных величин  $f_m$ . Будем считать, что известна ФПВ для  $f_m$  с конечной дисперсией  $\sigma_m$ . Поскольку измерение осуществляется в конечном числе точек, сведем с помощью той или иной квадратурной формулы интегральное уравнение (1) к системе линейных уравнений

$$\sum_{n=1}^N K_{m,n} \varphi_n = f_m \quad m=1, M \quad (2)$$

или в матричном виде  $\hat{K}\vec{\varphi} = \vec{f}$ , где компоненты матрицы и векторов

$$\hat{K} = K_{m,n} = K(x_m, y_n); \quad \vec{\varphi} = \varphi_n = \varphi(x_n); \quad \vec{f} = f_m = f(y_m)$$

Оценка  $\vec{\varphi}$  по реализации  $\vec{f}$  есть типичная задача математической статистики, которую мы будем рассматривать как задачу принятия решений в условиях неопределенности, вызванной неполным знанием состояния природы  $\varphi$ . Алгоритм, осуществляющий выбор действия для нахождения оценок  $\hat{\vec{\varphi}}$  называется стратегией. Разумная стратегия должна *минимизировать* вред от неполного знания состояния природы. Так как вред можно определить различными способами, существует множество стратегий. Мы будем использовать бейсову стратегию, которая определяется как стратегия минимума средней потери полезности по *апостериорной* плотности вероятности. Напомним, как она строиться.

Если величина оценки  $\hat{\vec{\varphi}}$  не сильно отличается от истинного закона  $\varphi$  (экспериментальная функция содержит достаточно информации для  $\hat{\vec{\varphi}}$ ), то потерю полезности можно определить как

$$U = \sum_n \mu_n (\varphi_n - \hat{\varphi}_n)^2 \quad (3)$$

где  $\hat{\vec{\varphi}} = S(n, m, \vec{f}, \vec{\sigma})$  - искомая стратегия

Естественно говорить о средней потере полезности, указав при этом ансамбль, по которому производится усреднение. В классическом подходе к задаче теории решений это есть ансамбль, образованный гипотетическим многократным повторением

измерения  $f$  при фиксированном  $\varphi$ . Плотность вероятности для такого ансамбля  $P(\vec{f}|\vec{\varphi})$  и усреднение дает

$$\langle U(\vec{\varphi}) \rangle = \int \sum_i \mu_n \left( \varphi_n - S(n, m, \vec{f}, \vec{\sigma}) \right)^2 P(\vec{f}|\vec{\varphi}) d\vec{f}$$

Минимизация такого среднего для определения стратегии  $S$  приводит к бессмыслице.  $S$  явно зависит от  $\varphi$ , т.е.  $S(n, m, \vec{f}, \vec{\sigma}) = \varphi_n$

Таким образом под *средней* потерей полезности надо понимать дополнительное усреднение по  $\varphi$ , что приводит к необходимости введения *априорной* плотности вероятности  $P(\varphi)$ . Такая функция должна выражать сравнительную степень веры в различные состояния природы, предшествующую измерению.

Если это признать, то стратегия  $S$  определится из условия минимизации

$$\min[U] = \int \sum_i \mu_n \left( \varphi_n - S(n, m, \vec{f}, \vec{\sigma}) \right)^2 P(\vec{f}|\vec{\varphi}) P(\vec{\varphi}) d\vec{f} d\vec{\varphi} \quad (4)$$

Введенная функция  $P(\varphi)$  позволяет по теореме Бейеса определить *апостериорную* плотность

$$P(\vec{\varphi}|\vec{f}) = \frac{P(\vec{f}|\vec{\varphi}) P(\vec{\varphi})}{\int P(\vec{f}|\vec{\varphi}) P(\vec{\varphi}) d\vec{\varphi}}$$

вероятность того, что фиксированный результат опыта  $f$  был порожден состоянием  $\varphi$  из множества  $P(\varphi)$ .

Проварьируем каждый член выражения (4) и получим

$$2\mu_n \int d\vec{f} \delta S(n, m, \vec{f}, \vec{\sigma}) \int d\vec{\varphi} P(\vec{\varphi}) P(\vec{f}|\vec{\varphi}) [S(n, m, \vec{f}, \vec{\sigma}) - \varphi_n] = 0$$

или

$$S(n, m, \vec{f}, \vec{\sigma}) = \hat{\varphi} = \int \vec{\varphi} P(\vec{\varphi}|\vec{f}) d\vec{\varphi} \quad (5)$$

Такого результата и следовало ожидать: байесова стратегия определяет решение как среднее  $\varphi$  по *апостериорному* ансамблю  $P(\vec{\varphi}|\vec{f})$

Итак, наилучшая стратегия в байесовском подходе требует введения *априорной* информации, не содержащейся в конкретной реализации данных измерения. Но не подрывает ли в корне это требование доверия к байесовой оценке? Ответ на этот



важный вопрос с точки зрения теории принятия статистических решений состоит в следующем.

При любом подходе к задаче математической статистики **невозможно** пытаться **объективно** судить о том, насколько хорошо в данном **единичном** эксперименте оценка угадывает свойства истинного  $\varphi$ . Задача единичного угадывания вообще не имеет отношения к науке. Возможность принятия каких-либо решений в этом случае надо скорее отнести к области веры. Цель научной теории - разработка общих алгоритмов, предназначенных для **многократного** использования и, в конечном счете, в правильной ориентации во внешнем мире. И ничего удивительного нет в том, что эффективность такого алгоритма должна зависеть от “свойств” этого мира. Эффективные алгоритмы, ориентированные на разные миры, должны быть различны. Таким образом, введение *априорной* плотности  $P(\varphi)$  имеет отношение **не** к угадыванию  $\varphi$  (как это представляется сторонникам чисто классического подхода), а к точной постановке задачи принятия решения в условиях неопределенности знаний о состоянии природы.

Но можем ли мы сказать что-либо существенное о том мире функций  $\varphi$ , который задается *априорной* плотностью?

Прежде чем высказаться по этому пункту заметим, что “наказанием” за отрицательный ответ будет возврат к началу, - если никакой *априорной* информации о сравнительной предпочтительности нет, то единственной возможностью останется принять стратегию равновероятности любых  $\varphi$ , для которых  $P(f|\varphi)$  отлична от нуля. И тогда мы вынуждены будем согласиться с тем, что оценка  $\varphi$  должна быть найдена как единственное решение исходного уравнения (2), т.е. принять нерегуляризованное решение (опять же, если  $\det K$  отличен от нуля!).

Что же *априори* можно сказать о  $P(\varphi)$ ? Если искомый физический закон есть гладкая функция, то и решение нужно искать в ансамбле гладких функций. Гладкость можно характеризовать функционалом

$$\Omega[\varphi] = \int \left[ \frac{d^k}{dx^k} \varphi(x) \right]^2 dx \quad (6)$$

Для алгебраизованного случая в матричном виде

$$\Omega[\varphi] = (\vec{\varphi} \hat{\Omega} \vec{\varphi}) \quad (7)$$

где  $\Omega$  – конечно-разностное приближение интегрального оператора (6)

Пусть для неизвестного нам истинного  $\varphi$  гладкость

$$\Omega[\varphi] = (\vec{\varphi} \hat{\Omega} \vec{\varphi}) = \omega \quad (8)$$

Очевидно, что мы должны так “устроить” ансамбль  $P(\varphi)$ , чтобы средняя гладкость в избранном нами мире равнялась  $\omega$ . Таким образом первое ограничение на “наш мир”

$$\int (\vec{\varphi} \hat{\Omega} \vec{\varphi}) P(\varphi) d\vec{\varphi} = \omega \quad (9)$$

Навязывая миру *априорную* плотность вероятности различных функций, мы хотели бы свести вынужденное насилие к **минимуму**, т.е. потребовать, чтобы информация относительно  $\varphi$ , содержащаяся в  $P(\varphi)$ , была минимальной. Используя определение информации связанное с понятием энтропии (по Шеннону), получаем второе требование в виде

$$I[P(\vec{\varphi})] = \int \lg P(\vec{\varphi}) P(\vec{\varphi}) d\vec{\varphi} = \min \quad (10)$$

Варьируя (10) с условием (9), получаем

$$P_{\alpha}(\vec{\varphi}) = c_{\alpha} \exp \left\{ -\frac{\alpha}{2} (\vec{\varphi} \hat{\Omega} \vec{\varphi}) \right\} \quad (11)$$

где  $\alpha = N/\omega$ ;  $c_{\alpha} = \frac{\sqrt{\det(\Omega)}}{(2\pi)^N} \alpha^{N/2} = c \alpha^{N/2}$  (12)

Плотность (11) для фиксированного  $\alpha$  гауссова - *априори* предпочтительны более гладкие (с меньшими значениями  $(\varphi \Omega \varphi)$ ) функции. Пока мы ничего не можем сказать о величине параметра распределения  $\alpha$ . Важно другое: неопределенность параметра  $\alpha$  отнюдь не влечет предположения  $P(\varphi) = \text{const}$ . Последнее означало бы полное отсутствие корреляций между компонентами  $\varphi$ , в то время как для физического закона, представимого гладкой функцией, справедливо обратное: компоненты  $\varphi$  с близкими значениями  $n$  коррелированы и глубина этих корреляций определяется параметром  $\alpha$ , значение которого пока произвольно.

Полная априорная вероятность  $\varphi$  при любых  $\alpha$  есть

$$P(\vec{\varphi}) = \int P(\alpha) P_{\alpha}(\vec{\varphi}) d\alpha \quad (13)$$

Здесь  $P(\alpha)$  суть *априорная* плотность различных параметров гладкости  $\alpha$  и вот ее то, если нет никакой дополнительной информации, необходимо считать постоянной в сколь угодно большой области  $\alpha > 0$ .

Таким образом, апостериорная плотность, по которой байесова стратегия определяет оценку  $\varphi$ , принимает вид

$$P(\bar{\varphi}|\bar{f}) = \frac{\int d\alpha P(\alpha) P_{\alpha}(\bar{\varphi}) P(\bar{f}|\bar{\varphi})}{\int d\alpha P(\alpha) \int d\bar{\varphi} P_{\alpha}(\bar{\varphi}) P(\bar{f}|\bar{\varphi})} \quad (14)$$

Введем обозначение

$$P(f|\alpha) = \int d\bar{\varphi} P_{\alpha}(\bar{\varphi}) P(\bar{f}|\bar{\varphi}) \quad (15)$$

и выясним смысл этого выражения.

Функция плотности (13) задает “слоистый” ансамбль для  $\varphi$  : множество  $\varphi$  разбивается на  $\alpha$ -слои. Но каждое  $\varphi$  не принадлежит в точности какому-то  $\alpha$ -слою, а может быть его элементом с вероятностью (13). Но если мы зададимся некоторым  $\varphi$ , то один из методов оценок в математической статистики поможет нам оценить значение  $\alpha$ , указывающее из какого  $\alpha$ -слоя по всей вероятности это  $\varphi$  было взято. Вместе с тем реализация  $f$  позволит в свою очередь оценить  $\varphi$ , т.к. известна вероятностная характеристика, порождающая  $f$  из  $\varphi$ . Символически двухступенчатая схема реализации  $f$  выглядит так

$$\begin{array}{ccccc} P(\alpha) & & P_{\alpha}(\varphi) & & P(\varphi|f) \\ \text{-----} \rightarrow & \alpha & \text{-----} \rightarrow & \varphi & \text{-----} \rightarrow f \end{array}$$

Если  $P(f|\alpha)$  есть условная вероятность порождения  $f$  из  $\alpha$ , то по теореме Байеса

$$P(\alpha|\bar{f}) = \frac{P(\alpha) P(\bar{f}|\alpha)}{\int P(\alpha) P(\bar{f}|\alpha) d\alpha} \quad (16)$$

условная плотность того, что данная реализация была рождена из некоторого слоя с параметром  $\alpha$ .

Теперь рассмотрим случай, когда можно аналитически проинтегрировать выражение (5) и получить конечные формулы для байесова решения обратной задачи и ошибки решения. Это можно сделать, когда значения экспериментальной функции  $f_m$  независимы и распределены по нормальному закону.

Выпишем окончательные формулы для этого случая

$$P(\vec{f}|\vec{\phi}) = \prod_{m=1}^M \frac{1}{\sqrt{2\pi\sigma_m^2}} \exp\left\{-\frac{(K\phi - f)_m^2}{2\sigma_m^2}\right\} = , \text{ где} \quad (17)$$

$$= c_1 \exp\left\{-\frac{1}{2}(\vec{\phi}^+ \hat{B} \vec{\phi}) + \vec{b} \vec{\phi}\right\}$$

$$W_{i,j} = \frac{1}{\sigma_i^2} \delta_{i,j} \quad ; \quad \hat{B} \equiv \hat{K}^+ W \hat{K} \quad ; \quad \vec{b} \equiv \hat{K}^+ W \vec{f}$$

$$P_\alpha(\vec{\phi}|\vec{f}) = c_2 \alpha^{N/2} \exp\left\{-\frac{1}{2}(\vec{\phi}[B + \alpha\Omega]\vec{\phi}) + \vec{b} \vec{\phi}\right\} \quad (18)$$

Решение и его ошибка при фиксированном  $\alpha$

$$\hat{\phi}_\alpha = (B + \alpha\Omega)^{-1} \vec{b} \quad \bar{s}_{\hat{\phi}_\alpha} = (B + \alpha\Omega)^{-1}_{nn} \quad (19)$$

Окончательное решение и ошибка получаются усреднением по апостериорной плотности для  $\alpha$

$$\hat{\phi} = \int \hat{\phi}_\alpha P(\alpha|\vec{f}) d\alpha; \quad \bar{s}_{\hat{\phi}}^2 = \int (\hat{\phi}_\alpha - \hat{\phi})^2 P(\alpha|\vec{f}) d\alpha \quad (20)$$

$$P(\alpha|\vec{f}) = c_3 \alpha^{N/2} \frac{1}{\sqrt{\det(B + \alpha\Omega)}} \exp\left\{\frac{1}{2} \vec{b}^+ (B + \alpha\Omega)^{-1} \vec{b}\right\} \quad (21)$$

Это усреднение громоздко и часто не имеет смысла. Действительно, интуитивно ясно, что приступать к решению обратной задачи имеет смысл только, если точность  $f$  и свойства ядра уравнения обеспечивают не слишком большую ошибку решения. Можно показать, что это достигается только в том случае, если апостериорная плотность  $P(\alpha|f)$  обладает достаточно малой дисперсией (ширина  $P(\alpha|f)$  меньше наивероятного значения  $\alpha_{\text{наив}}$ ). Тогда хорошей оценкой (оценкой максимального правдоподобия) будет  $\varphi$  и  $s$  при  $\alpha_{\text{наив}}$ , которое легко найти, как решение уравнения  $dP(\alpha|f)/d\alpha = 0$ . Если же  $P(\alpha|f)$  очень широкая функция, то ошибка  $s$  будет настолько велика, что само решение исходной задачи теряет смысл, хотя формально может быть получено и является правильным.

Такое заключение должно быть правильно понято. Взявшись решать статистическую проблему (в байесовском или классическом подходе), мы должны быть заранее готовы к тому, что в эксперименте недостаточно информации об искомой функции, хотя препятствий к формальному применению метода нет. Более того, планирую эксперимент, предполагающий последующее решение обратной задачи для извлечения

полезной информации, полезно моделированием “испытать” его на  $P(\alpha|f)$  - будет ли при предполагаемой ошибке  $f$  достаточно информации.

Часто мы имеем гораздо больше *априорной* информации, чем в соотношениях (11-13). Например, заранее известно, что восстанавливаемая функция положительна. Естественнo пытаться использовать все сведения об искомом законе для уменьшения ошибки  $S$ , но введение даже простой ступенчатой функции (для обеспечения положительности решения) уже не позволяет получить формулы в аналитическом виде.

Простейшей возможностью является следующая. При найденном наивероятном значении  $\alpha$  можно найти ту грань гиперкуба, на которой лежит положительное решение. (В теории динамического программирования доказана теорема о том, что оптимальное решение при ограничивающих условиях всегда лежит на грани). Используя решение можно несколько уточнить значение  $\alpha$  и само решение. Конечно, это палеативное решение проблемы. Но можно поступить и более последовательно: современные компьютеры позволяют взять многомерные интегралы (13-16) методом Монте-Карло, который позволяет вводить априорную информацию любой природы.

Таким образом, метод статистической регуляризации дает решение обратной задачи, используя всю доступную априорную информацию в духе байесовской стратегии.