УДК 550.8310.17+550.8380.17

П.И. Балк

НЕСКОЛЬКО КОНТРПРИМЕРОВ К СТЕРЕОТИПАМ В ТЕОРИИ ИНТЕРПРЕТАЦИИ ПОТЕНЦИАЛЬНЫХ ПОЛЕЙ

Оценивая современное состояние дел в области решения обратных задач гравиметрии и магнитометрии, В.Н. Страхов в качестве важнейшей проблемы, требующей решения в ближайшие десятилетия, обозначил проблему создания теории интерпретации потенциальных полей, полностью адекватной реалиям геофизической практики [1, 2]. Если говорить об источниках неадекватности в теории интерпретации потенциальных полей, то это, прежде всего, стереотипы мышления, сложившиеся в среде геофизиков-теоретиков. Существует и обратная связь – неадекватность используемых представлений реалиям геофизической практики породила определенные заблуждения, которые впоследствии нашли свое отражение в основных положениях теории интерпретации потенциальных полей.

К самым очевидным, лежащим на поверхности проявлениям неадекватности можно отнести использование двухмерных моделей источников поля и игнорирование факта сферичности Земли. Какие-то признаки неадекватности используемых моделей среды, однажды выпав из поля зрения геофизиков, впоследствии могут оказаться достаточно неожиданными [3]. Но если в практике интерпретации использование различного рода идеализаций, не отвечающих природным условиям, есть, большей частью, следствие скудости программного обеспечения, имеющегося в распоряжении интерпретатора, то в теории решения обратных задач гравиразведки и магниторазведки причины неадекватности лежат глубже и не столь очевидны.

В работах [4-6] основным источником противоречий в современной теории интерпретации потенциальных полей, ставших источником многих заблуждений, распространенных в среде геофизиков, названо столкновение математических и геофизических интересов при выборе актуальных направлений исследований и формулировке критериев состоятельности научных результатов. Механизм зарождения противоречий между геофизическим и математическим видением задачи можно описать так:

а) В практике интерпретации гравитационных (магнитных) аномалий часто возникает некая задача А, для решения которой, исходя из приоритета геофизических ценностей над математическими, уже были разработаны (во многом силами самих геофизиков-теоретиков) определенные методы – возможно, недостаточно формализованные, содержащие элементы эвристики, но вобравшие в себя весь опыт и интуицию интерпретатора, прошедшие серьезную обкатку на модельном и производственном материале.

б) За формализацию постановки задачи и придание ей более строгого вида берется математик, представитель классического направления в теории решения обратных задач, и, переиначив постановку задачи А в соответствии со своими профессиональными интересами и возможностями наработанного в этой области математического аппарата, предлагает геофизикам вместо А другую, якобы «эквивалентную», задачу Б, решение которой, на первый взгляд, обещает нечто большее, чем могли предложить прежние доморощенные методы (в этом смысле чего только может стоить в новой постановке один лишь пассаж «для заданного сколь угодно малого ε обеспечить погрешность решения обратной задачи, не превосходящую это ε»).

в) При столь заманчивых обещаниях, данных, к тому же, математиком, можно не заметить явных противоречий, присущих постановке задачи Б (или попросту, закрыть на них глаза). Так, в математической теории некорректных обратных задач приоритет отдается бесконечномерным моделям изучаемых объектов, что вынуждает вводить в рассмотрение неадекватное реальности условие задания поля на континуальном множестве точек, тогда как на практике, где измерения дискретны, информационно обеспеченными могут быть лишь конечномерные модели. Чтобы обеспечить состоятельность алгоритма решения обратной задачи, в математической теории используется еще одна идеализация – предположение о стремлении нормы помехи к нулю. Естественное при изучении гносеологических проблем принципиальной познаваемости действительности по ее косвенным проявлениям, такое допущение совершенно оторвано от фактических условий геофизического эксперимента. Тем самым, состоятельность алгоритма при решении определенной задачи нам рекомендуется оценивать не по конкретным результатам, а в сослагательном наклонении, по тому, что бы мы могли иметь в неких нереализуемых условиях! Не многим реалистичнее – на это В.Н. Страхов, кстати, неоднократно обращал внимание – выглядит и предпосылка об априори известном (точном!) значении уровня помех, на которой основаны известные методы регуляризации.

г) В математической теории интерпретации потенциальных полей в качестве «правильной» и строго обоснованной (а главное, сулящей наибольшую отдачу) теперь прочно обосновалась постановка Б. Как и принято в математике, это довольно общая постановка, в которой многие важные частности исходной задачи А попросту «растворились». И вообще, любая дополнительная информация, без которой и так сходимость может быть достигнута, в задаче Б перестает играть решающее значение, Если вспомнить основополагающие работы по теории регуляризации, то там делалась ставка лишь на один тип априорной информации – гладкость решения. Надо ли говорить, что это полностью противоречит общей геофизической концепции повышения качества результатов интерпретации за счет привлечения разнородных априорных данных. Желание видеть в обратной задаче изначальную неустойчивость (возможность *сколь угодно* больших ошибок решения) приводит к тому, что в постановке обратной задачи Б игнорируются естественные ограничения на физические и геометрические параметры источников поля, при которых множество допустимых решений изначально является компактным. Если же используются отличные от классического условия гладкости решения, ограничения, то они обязательно должны оставить место для проявления некорректности, а значит, и борьбы с ней (иначе математический интерес к задаче пропадает).

д) Наступает своего рода развязка – выясняется, что обещанные возможности постановки задачи Б, которые и сделали ее столь привлекательной, требуют идеализаций, оторванных от реалий практики. Иными словами, негативные последствия неадекватности предпосылок, лежащих в фундаменте постановки Б, превосходят достоинства строгих построений, выполненных на этом фундаменте.

е) Но в теории интерпретации потенциальных полей под влиянием идей, заложенных в постановку задачи Б, уже сложились устойчивые стереотипы, не предусматривающие пересмотра основных концептуальных положений. И теперь остается лишь бороться с проблемами, которые привнесла постановка Б и которых могло не быть, если бы постановка задачи А не была отвергнута как математически не вполне обоснованная.

В названных работах [1, 2] критика в адрес неадекватности классического подхода реалиям геофизики служит в главном обоснованию целесообразности регуляризировать не континуальные, а конечномерные постановки обратных задач. По мнению автора настоящей статьи, в действительности имеет место более глубокое проникновение неадекватности в теорию интерпретации потенциальных полей и размежевание формального математического и прикладного геофизического подходов начинается уже на уровне общеметодологических идей и целевых установок интерпретации, на которых и базируются любые алгоритмы решения обратных задач гравиметрии, в том числе и регуляризирующие.

В статье выделены семь, по мнению автора, основных, положений теории интерпретации данных гравиразведки и магниторазведки, на неприемлемость и ошибочность которых в геофизической литературе практически не обращалось внимание. В одних случаях, как нам показалось, будет достаточно доказательной аргументации мнения, противоположного тому, что культивируется в настоящее время в теории, в других – приводятся простые и наглядные контрпримеры.

Стереотип первый: ***точечные (оптимальные) оценки параметров модели источников аномалии являются единственной, безальтернативной формой представления результатов интерпретации.*** При всем многообразии подходов к решению обратных задач для потенциальных полей незыблемой остается форма представления результатов интерпретации. Казалось бы, иного не дано, и если $\hat{\Theta} \in \mathbf{R}^n$ – вектор неизвестных параметров в описании модели источников поля, то и результатом интерпретации обязательно должен быть вектор Θ^* приближенных оценок Θ_j^* параметров Θ_j. Интерпретатора не смущает тот факт, что параметры модели могут даже не иметь содержательной интерпретации (коэффициенты полинома в описании закона изменения избыточной плотности масс и уравнении границы раздела двух сред) – по умолчанию им предполагается, что информацию, хранящуюся в параметрах модели, можно «расшифровать», переформулировав ее в геологически содержательных терминах (по примеру того, как координаты вершин многогранника, описывающего возмущающее тело, можно перевести в геометрическое место точек пространства, занятого аномальными массами).

Все так, будь на практике возможность пренебречь погрешностью $\varepsilon = \|\Theta^* - \widehat{\Theta}\|$ приближенного решения Θ^*, а тем более, возможность организовать построение последовательности таких решений, соответствующих различному уровню помех в измерениях, и обеспечить ее сходимость к точному решению с последующим выбором любого из них в качестве результата интерпретации. Но в ситуации, когда погрешность ε известна весьма приближенно, информативность решения Θ^* (если ее понимать как достоверность данных об объекте исследования) падает до нуля – информации, хранящейся в Θ^*, недостаточно, чтобы указать хотя бы одну точку пространства, гарантированно принадлежащую носителю $\widehat{S}$ аномальных масс (не считая, конечно, тех, что предопределены априорными ограничениями). И это неоспоримый недостаток представления результатов интерпретации в виде единичного решения обратной задачи.

Критика должна быть конструктивной, и еще в работе [7] автор предложил несколько альтернативных форм представления результатов интерпретации данных гравиметрических измерений, объединенных идеей поиска физически содержательных инвариантов на множестве Q допустимых решений обратной задачи. При адекватности всех априорных предпосылок реальным природным соотношениям и условиям проведения геофизического эксперимента (тогда истинное решение принадлежит Q) такие инварианты являются гарантированно достоверной информацией об объекте исследования. Идея была оформлена в виде нового, названного гарантированным, подхода к решению обратных задач для потенциальных полей. В случае обратной задачи рудного типа такой подход предлагает вместо оптимальной (по любому из критериев) оценки S^* носителя $\widehat{S}$ пару $<D_1, D_2>$, где D_1 – отвечающая совокупной информации G минимальная область геологического пространства, гарантированно содержащая носитель $\widehat{S}$, а D_2 – максимальный фрагмент неизвестного носителя, который можно указать по информации G. Были разработаны конструктивные методы построения областей D_1 и D_2, по сути являющихся двухсторонними оценками типа включений для носителя $\widehat{S}$: $D_2 \subset \widehat{S} \subset D_1$.

За период, истекший после выхода работы [7], эффективность альтернативных форм представления результатов интерпретации гравитационных и магнитных аномалий неоднократно подтверждалась при решении практических задач [8]. Тот факт, что новые формы с трудом приживаются на практике, связан, скорее всего, со следующим распространенным заблуждением.

Стереотип второй: ***оптимальные оценки параметров модели совместно с «хорошими» оценками их точности являются исчерпывающей информацией об объекте исследования.*** Можно пока оставить вопрос о трудностях построения эффективных оценок $\rho(S^*, \widehat{S}) \le \varepsilon$ $(\rho(\Theta^*, \widehat{\Theta}) \le \varepsilon)$ точности приближенных решений, равно как и то, что когда постановка задачи содержит различные ограничения, конструктивным является не традиционный для математики аналитический, а алгоритмический путь построения оценок точности.

Главное в другом. С тем чтобы обосновать ошибочность приведенного утверждения, возьмем наиболее выгодную для его защиты ситуацию, когда оценка сверху для ошибки приближенного решения обратной задачи известна точно: $\rho(\Theta^*, \widehat{\Theta}) \le \varepsilon$. Тогда $\varepsilon = \sup\rho(\Theta, \Theta^*)$, $\Theta \in Q$. При этом само значение ε может быть построено без того, чтобы использовать явное описание множества Q.

Итак, пусть вся априорная информация G преобразована в пару $<\Theta^*, \varepsilon>$. Ее содержательный смысл: неизвестным вектором $\widehat{\Theta}$ истинных параметров модели источников поля может оказаться любой вектор Θ, такой, что $\rho(\Theta^*, \Theta) \le \varepsilon$. Но если вернуться к исходным данным, то окажется, что некоторые из таких векторов не удовлетворяют информации G. И это неудивительно, поскольку множество $Q_0 = \{\Theta: \rho(\Theta^*, \Theta) \le \varepsilon\}$ в общем случае шире множества допустимых решений Q. Иначе говоря, неопределенность, заключенная в паре $<\Theta^*, \varepsilon>$, выше той, которая заключена в исходной совокупной информации G. При переходе от G к $<\Theta^*, \varepsilon>$ значительная часть информации, содержащейся в G, может быть потеряна. Чем сложнее граница множества Q, тем существеннее различие множеств Q и Q_0, тем заметнее эта потеря. Можно предположить, что стереотип, который мы обсуждаем, во многом обязан тому, что указанное обстоятельство в теории интерпретации осталось незамеченным.

К этому стереотипу мы еще вернемся с более общих позиций и воспользуемся при обсуждении одним контрпримером.

Стереотип третий: ***единственность решения обратной задачи есть основной фактор,***

предопределяющий качество результатов интерпретации. Судя по тому, как в геофизической литературе излагаются последствия, к которым приводит неоднозначность решения обратных задач, определенное заблуждение здесь кроется в автоматическом переносе условий теоретической обратной задачи на прикладные задачи. Неединственность обратной задачи у многих геофизиков прочно ассоциируется с существованием решений, сколь угодно различающихся между собой. Конечно, если обратная задача ставится на паре функциональных пространств без каких-то добавочных условий, когда ядро обратного оператора неограничено, такие решения существуют. Но на практике постановки обратных задач всегда содержат ограничения, продиктованные априорной информацией, и это в корне меняет дело.

Доказательным контраргументом к сложившемуся стереотипу может стать следующий пример, показывающий, что свойство единственности решения нельзя рассматривать вне контекста конкретных условий постановки обратной задачи, что по своему негативному влиянию неоднозначность может уступить другим факторам, способствующим повышению неопределенности выбора, хотя бы той же неадекватности модельных представлений.

Пример. Аномалия $\widehat{U}$ гравитационного потенциала создается шаром $\widehat{S}$ радиуса $R = 1$ массы $\widehat{m}$, центр O которого расположен в заданной точке. Известно, что в сферической системе координат (r, φ, θ) с началом в O плотность $\widehat{\delta}$ шара линейно зависит от r: $\widehat{\delta} = \widehat{a} + \widehat{b}r$ $\delta = a + br$. В рассматриваемом двухпараметрическом модельном классе

$$M = \{\delta = a + br : (a, b) \in \mathbf{R}^2\} \tag{1}$$

обратная задача неоднозначна: при любых a и b, удовлетворяющих условию:

$$\frac{4}{3}\pi a + \pi b = \widehat{m},$$

потенциал шара остается неизменным и аномалии $\widehat{U}$ формально отвечает неограниченное однопараметрическое семейство M' решений

$$\delta_a(r) = a + \left(\frac{\widehat{m}}{\pi} - \frac{4}{3}a\right)r,\ a \in \mathbf{R}. \tag{2}$$

Предположим, однако, что априори известна двухсторонняя оценка $\overline{a} \le \widehat{a} \le \overline{\overline{a}}$ истинного значения $a = \widehat{a}$ плотности масс в центре O шара. Тогда множество эквивалентных решений обратной задачи суживается до отрезка

$$M'' = \{\delta_a(r) \in M' : \overline{a} \le a \le \overline{\overline{a}}\}. \tag{3}$$

Если точность d решения (a^*, b^*) теоретической обратной задачи по аномалии $\widehat{U}$ оценивать в метрике $d = (a^* - \widehat{a})^2 + (b^* - \widehat{b})^2$, то при любом выборе его из множества эквивалентности M'' и любом положении точки $\widehat{a}$ на отрезке $\left[\overline{a}, \overline{\overline{a}}\right]$ выполняется неравенство

$$d \le \frac{25}{9}(\overline{\overline{a}} - \overline{a})^2. \tag{4}$$

Пусть теперь альтернативный выбор заключается в том, чтобы искать решение обратной задачи в более узком классе моделей $M_0 \subset M$ – семействе постоянных плотностей $\delta = a$, $a \in \mathbf{R}$ масс, заполняющих $\widehat{S}$, то есть в неадекватном классе, но зато со свойством единственности. Тогда

$$a^* = \widehat{a} + \frac{3}{4}\widehat{b}\ -$$

теперь уже единственное решение обратной задачи, отвечающее аномалии $\widehat{U}$. Его погрешность

$$d = \frac{25}{16}\widehat{b}^2. \tag{5}$$

Из сопоставления правых частей (4) и (5) следует, что при выполнении условия

$$\left|\overline{\overline{a}} - \overline{a}\right| \le \frac{3}{4}\left|\widehat{b}\right| \tag{6}$$

любое эквивалентное решение, взятое из модельного класса, где единственности нет, превзойдет по точности решение обратной задачи, взятое из класса, где единственность имеет место.

Какой же из этого примера следует сделать вывод? Разумеется, никто не станет отрицать, что при прочих равных условиях модельные классы со свойством единственности априори предпочтительней тех, где это условие не соблюдается. Вопрос лишь в недопустимости абсолютизации этого свойства. Сложившееся представление о единственности решения обратной задачи как неотъемлемом атрибуте любого состоятельного алгоритма и, тем более, как некой гарантии качества результатов интерпретации, необходимо переосмыслить. Строгая неоднозначность обратной это лишь крайнее проявление ε-эквивалентности, и совсем не обязательно, что негативные последствия от нее окажутся более значимыми, чем последствия от приближенной эквивалентности.

Стереотип четвертый. ***Основным критерием состоятельности любого алгоритма количественной интерпретации данных гравиразведки и магниторазведки является сходимость последовательности приближенных решений к точному решению.*** Это положение является «краеугольным камнем» классической теории решения некорректных обратных задач и отражает общее представление математика об «изящности» и законченном виде результата. Критерий был бы неоспорим, если бы такую сходимость фактически удавалось организовать на практике. В реальности речь может идти лишь о воображаемой последовательности приближенных решений обратной задачи. При решении конкретной задачи интерпретатору нет никакого дела до виртуального поведения приближенных решений на любом отрезке значений нормы помехи, который не пересекается с интервалом $[\varepsilon_{\min}, \varepsilon_{\max}]$ возможных значений уровня помех.

Пример. Представим следующую ситуацию, возможность возникновения которой не должна вызывать сомнений и потому конкретный числовой материал здесь не понадобится. Итак, пусть для решения обратной задачи гравиразведки (построения приближенной оценки Θ^* вектора $\Theta \subset \mathbf{R}^n$ параметров модели источников поля) при некоторой априорной информации G, включающей оценку $\varepsilon \in [\varepsilon_{\min}, \varepsilon_{\max}]$ значения ε нормы помехи ξ измерений $\Delta\tilde{g}_i = \Delta\hat{g}_i + \xi_i$, $i = 1, 2, \ldots, m$, есть два принципиально различных алгоритма. Первый из них (алгоритм А) с учетом выбора модельного класса гарантирует (теоретически) сходимость. Это означает, что если бы решение обратной задачи можно было выполнить многократно (при $k = 1, 2, \ldots$) по данным $\Delta\tilde{g}_i^{(k)} = \Delta\hat{g}_i + \lambda_k\xi_i$, где (λ_k) – последовательность, сходящаяся к нулю (соответственно, стремится к нулю и норма помехи), то последовательность соответствующих приближенных решений (Θ_k^*) сходилась бы к точному вектору $\hat{\Theta}$, а последовательность $\rho(\Theta_k^*, \hat{\Theta})$ фактических значений ошибок этих решений – к нулю. Второй, математически «не состоятельный», алгоритм Б сходимость не гарантирует, например, из-за сомнений, что какое-то из ограничений может оказаться не вполне адекватным действительности, при том, что использование в целом информации G вполне окупается.

Предположим, что при всех λ_k, таких, что $\|\lambda_k \xi\| \in [\varepsilon_{\min}, \varepsilon_{\max}]$, точность решения $\hat{\Theta}^*$, полученного по алгоритму А, ниже точности решения по алгоритму Б (тогда как вне этого интервала зависимости точности решения от уровня помех по каждому из алгоритмов пусть ведут себя как угодно). Такой исход не удивителен, если, например, алгоритм А проигнорировал часть априорной информации, без которой он и так смог обеспечить сходимость. Ясно, что приоритет в подобных условиях за алгоритмом Б, хотя этот выбор противоречит сложившемуся стереотипу.

Стереотип пятый. ***Качество результатов интерпретации, выполненной с использованием сразу нескольких взаимонезависимых объемов G_j, $j = 1, 2, \ldots, m$, априорной информации, не уступит качеству результатов интерпретации, выполненной по каждому парциальному объему информации G_j.*** На этом убеждении основана идея комплексирования. Конечно, мы несколько утрировали его; ясно, что если к высокоточным данным G_1 «подмешать» грубые данные G_2, то результаты совместной интерпретации данных G_1 и G_2 скорее всего уступят по точности результатам интерпретации по одним лишь данным G_1. Но то, что решение обратной задачи по совокупным данным $G = \{G_1, G_2, \ldots, G_m\}$ уступит *всем* решениям, построенным по отдельно взятым данным G_j, в общепринятые стандарты не укладывается.

Обсуждаемое заблуждение связано, по мнению автора, с тем, что из поля зрения геофизика выпал факт отсутствия монотонной зависимости качества решения обратной задачи (в форме единичных «оптимальных» оценок параметров модели объекта) от объема и качества априорной информации. При определенном «раскладе» помех в измерениях это способно привести к тому, что некоторый объем априорной информации удается разбить на несколько частей так, что точность решения обратной задачи по каждой отдельно взятой части информации окажется выше точности решения, построенного с использованием всего объема информации. Подобный феномен имеет следующее объяснение.

Пусть $Q_1, Q_2, \ldots, Q_m$ – множества допустимых решений обратной задачи, отвечающие некоторым взаимонесвязанным объемам априорной информации $G_1, G_2, \ldots, G_m$ и $Q_1 \cap Q_2 \cap \ldots \cap Q_m$ – множество допустимых решений, отвечающих всей априорной информации $G = \langle G_1, G_2, \ldots, G_m \rangle$. Если S_j^* – приближенное решение обратной задачи, построенное по информации G_j с помощью какого-то метода, S^* – решение, построенное тем же методом по информации G, и множество

$$Q' = \{S \in Q : \nu(S, S^*) > \nu(S, S_j^*), j = 1, 2, ..., m\}, \quad (7)$$

где $\nu(S_1, S_2)$ – мера близости двух решений (не обязательно метрика) не пусто, то в случае $\hat{S} \in Q'$ решение S^* уступит по точности всем парциальным решениям S_j^*. По сути здесь заложен конструктивный способ моделирования примеров такого рода, после чего исход $Q' \neq \varnothing$ вовсе не выглядит патологическим.

Следует подчеркнуть, что немонотонное поведение зависимости точности результатов интерпретации от качества используемой информации характерно для стандартных форм представления результатов интерпретации. Иное дело, когда за результат интерпретации принимаются инварианты на множестве Q допустимых решений. В этом случае любые дополнительные данные измерений (пусть даже сомнительного качества) не могут привести к уменьшению объема достоверной информации об изучаемом объекте, извлекаемой из интерпретируемых данных.

Пример. Структура оператора прямой задачи здесь не играет принципиального значения. Однако в любом случае этот пример можно рассматривать в аспекте проблемы сглаживания данных измерений. Здесь он удобен тем, что читатель, у которого возникнут сомнения, может самостоятельно с помощью калькулятора все перепроверить. Оцениваются параметры $\hat{c}_1$ и $\hat{c}_2$ линейной функции $\varphi(x) = \hat{c}_1 + \hat{c}_2 x$, $\hat{c}_1 = \hat{c}_2 = 1$, по ее значениям $\tilde{y}_i = \varphi(x)_i + \xi_i$, где ξ_i – случайные помехи. Оценки c_1^* и c_2^* параметров $\hat{c}_1$, $\hat{c}_2$ вычисляются по трем группам данных: по заданным в точках $x_i = i$, $i = 1, 2, \ldots, 5$, значениям $\tilde{y}_i$, где $\xi_1 = 0{,}87$, $\xi_2 = 1{,}59$, $\xi_3 = -0{,}23$, $\xi_4 = 0{,}84$, $\xi_5 = 2{,}40$ (парциальные данные G_1); по заданным в точках $x_i = i$, $i = 6, 7, \ldots, 10$, значениям $\tilde{y}_{i+5}$, где $\xi_6 = -0{,}12$, $\xi_7 = -0{,}89$, $\xi_8 = -0{,}44$, $\xi_9 = 0{,}11$, $\xi_{10} = -0{,}54$ (парциальные данные G_2); по всем десяти измерениям, представленным в данных G_1 и G_2 (совокупные данные G). Оценки c_1^* и c_2^* вычисляются по методу наименьших квадратов, а их качество оценивается по трем функционалам:

$$v_1 = \left((c_1^* - \hat{c}_1)^2 + (c_2^* - \hat{c}_2)^2\right)^{1/2}, \tag{8}$$

$$v_2 = \left|c_1^* - \hat{c}_1\right| + \left|c_2^* - \hat{c}_2\right|, \tag{9}$$

$$v_3 = \max\left\{\left|c_1^* - \hat{c}_1\right|, \left|c_2^* - \hat{c}_2\right|\right\}. \tag{10}$$

Получены следующие результаты.

По данным G_1:

$c_1^* = 1{,}401$, $c_2^* = 1{,}231$, $v_1 = 0{,}463$, $v_2 = 0{,}632$, $v_3 = 0{,}401$. (11)

По данным G_2:

$c_1^* = 0{,}496$, $c_2^* = 1{,}016$, $v_1 = 0{,}504$, $v_2 = 0{,}520$, $v_3 = 0{,}504$. (12)

По совокупным данным $G = \langle G_1, G_2 \rangle$:

$c_1^* = 2{,}419$, $c_2^* = 0{,}807$, $v_1 = 1{,}433$, $v_2 = 1{,}612$, $v_3 = 1{,}419$. (13)

Таким образом, худший из двух парциальных результатов превосходит по точности результат обработки (тем же методом) совокупных данных более чем в 2 раза, причем по всем трем показателям – v_1, v_2 и v_3. Но и это не все. Расчеты, выполненные с теми же данными, показали: аналогичная картина наблюдается и в случае, когда оценки c_1^* и c_2^* вычисляются по методу наименьших модулей, а также путем минимизации максимального уклонения построенной зависимости от измерений $\tilde{y}_i$ в точках x_i.

С точки зрения концепции повышения качества результатов интерпретации геофизических данных за счет роста объемов используемой информации исход, подобный данному, является достаточно пессимистическим. Однако, настоящий пример ни в коей мере не направлен на подрыв доверия к обсуждаемой идее; он лишь раскрывает еще одну слабую сторону общепринятых форм представления результатов решения обратных задач, в рамках которых данная идея не может быть воплощена в полной мере.

Пример. Аномалия Δg создается двумя шарами $\hat{S}_1$ и $\hat{S}_2$ единичного объема V_1 и V_2 и единичной плотности $\hat{\delta}_1$ и $\hat{\delta}_2$. Центры шаров имеют координаты (−1, 0, 1) и (1, 0, 1). Гравитационная постоянная условно равна единице. Имеются три парциальные группы данных. G_1 – значения $\Delta\tilde{g}_i$ в точках $x_i = -3{,}2 + 0{,}2i$, $i = 1, 2, \ldots, 15$, и помехи ξ_i собраны в массив {−0,53; 1,17; −0,94; 0,80; −0,72; 0,04; −0,26; 0,96; −0,63; 0,02; −2,25; 0,20; −0,91; 0,33; −0,44}; G_2 – значения $\Delta\tilde{g}_{i+15}$ в точках $x_{i+15} = 0{,}2i$, $i = 1, 2, \ldots, 15$, и помехи ξ_{i+15} собраны в массив {−0,75; 1,10; 0,18; −0,61; −0,31; −0,74; 0,81; 0,14; 0,35; −1,62; −0,99; −0,32; 0,19; −0,74; 2,00}; G_3 – значения $\Delta\tilde{g}_{i+30}$ в точках $x_{i+30} = -1{,}7 + 0{,}2i$, $i = 1, 2, \ldots, 15$, и помехи ξ_{i+30} собраны в массив {−0,30; −0,49; 1,33; −0,88; −0,02; −0,41; −0,23; −1,48; 1,20; −0,24; −0,02; −0,42; 0,02; −2,50; 2,59}.

Казалось бы, во всех трех случаях точки измерения аномалии выбраны крайне неудачно: данные G_1 и G_2 характеризуют либо левый, либо правый фланг аномалии, соответственно, а G_3 – ее достаточно узкую эпицентральную часть, тогда как совокупные данные $G = \langle G_1, G_2, G_3 \rangle$ не только полно характеризуют аномалию, но в ее эпицентральной части содержат измерения по более густой сети, чем данные G_1, G_2, G_3. Однако вопреки привычным представлениям решение обратной задачи оказалось наихудшим именно при использовании данных G.

За наилучшие оценки δ_1^*, δ_2^* плотностей шаров $\hat{S}_1$, $\hat{S}_2$ примем те, что минимизируют квадратическую невязку, а качество решения обратной задачи будем оценивать по величине v ошибки определения избыточной массы:

$$v = \left|V_1(\hat{\delta}_1 - \delta_1^*) + V_2(\hat{\delta}_2 - \delta_2^*)\right|. \tag{14}$$

Результаты решения обратной задачи таковы.

По данным G_1:

$$\delta_1^* = 0{,}452,\ \delta_2^* = 1{,}255,\ v = 0{,}293. \tag{15}$$

По данным G_2:

$$\delta_1^* = 1{,}444,\ \delta_2^* = 0{,}766,\ v = 0{,}210. \tag{16}$$

По данным G_3:

$$\delta_1^* = 0{,}836,\ \delta_2^* = 0{,}882,\ v = 0{,}282. \tag{17}$$

По совокупным данным G:

$$\delta_1^* = 0{,}664,\ \delta_2^* = 0{,}891,\ v = 0{,}444. \tag{18}$$

Таким образом, точность определения избыточной массы по совокупности измерений уступает, как минимум, в полтора раза точности ее определения по каждым из трех групп парциальных измерений.

Стереотип шестой. ***Наилучшее решение обратной задачи обязательно является точкой глобального минимума некоторого функционала качества.*** В наиболее характерном случае таким функционалом служит невязка наблюденного и подбираемого полей. Типичная постановка обратной задачи в геофизической литературе выглядит так: осуществляется формализация модели источников поля и модели экспериментальных данных, декларируется набор ограничений, а далее всегда следует стереотипное утверждение: «требуется найти оптимальный вектор параметров модели объекта, минимизирующий функционал качества решения».

Но ведь цель интерпретации в извлечении максимума полезной и достоверной информации об объекте исследования, и это не эквивалентно поиску допустимого решения, обладающего некими экстремальными свойствами. Природные объекты, равно как случайные помехи в данных измерений их полей, образуются под влиянием разнообразных факторов (большей частью неконтролируемых) и не обязаны удовлетворять экстремальным принципам, удобным для использования готового математического аппарата. Стремление подогнать постановку задачи под известные математические разработки (тогда как необходимо обратное) – яркий пример проявления неадекватности в теории интерпретации. К тому же, априори все допустимые решения обратной задачи равноправны и любое из них может оказаться неизвестным точным решением. Ясно, что наилучшим допустимым решением является точное решение, но «оптимальное» по минимуму невязки решение, к сожалению, с ним совпадает очень редко.

Парадокс еще заключается в том, что, формулируя обратную задачу как поиск глобального минимума функционала и предлагая соответствующий алгоритм, мы заранее «расписываемся» в том, что алгоритм с этой задачей не справится, поскольку до сих пор проблема глобального минимума нелинейных многоэкстремальных функций не решена [9]. По мнению автора, задача поиска глобального минимума должна рассматриваться как средство, позволяющее выйти на некоторое допустимое решение. Попросту говоря, минимизируя невязку, мы осознанно идем на то, что не достигнем глобального значения, но надеемся, что, возможно, выйдем на какое-то достаточно небольшое (и уже допустимое) ее значение.

Стереотип седьмой. ***Оптимальное решение обратной задачи и неулучшаемая оценка его точности содержат максимальный объем достоверной информации, который может быть извлечен из всей априорной информации.***

Оговоримся сразу, что относить данное заблуждение к стереотипным все же несколько преждевременно, поскольку акцент на достоверность информации в известных автору публикациях детерминистского направления никогда не поднимался, а если возникнет, то, вероятно, лишь в связи с активным внедрением новых форм представления результатов интерпретации.

В терминах нелинейной обратной задачи гравиметрии для распределения масс заданной плотности в некоторой локальной области $\widehat{S} \subset \mathbf{R}^3$ суть вопроса можно прокомментировать следующим образом. Начнем с того очевидного факта, что минимальной областью пространства, гарантированно содержащей искомый носитель масс $\widehat{S} \in Q$, является объединение $D_1 = \cup S_\alpha$ всех допустимых носителей $S_\alpha \in Q$ – приближенных решений обратной задачи, удовлетворяющих априорной информации G, а пересечение $D_2 = \cap S_\alpha$ есть максимальная область, гарантированно являющаяся фрагментом носителя $\widehat{S}$. Области D_1 и D_2 можно рассматривать как инварианты на множестве Q в том смысле, что при любом выборе элемента $S_\alpha \in Q$ в качестве наилучшего решения S^* обратной задачи неизменно область D_2 будет входить (как подмножество) в общий фрагмент $S_\alpha \cap \widehat{S}$ выбранного и истинного носителей, тогда как область D_1 всегда будет связывать оба этих носителя тем, что их содержит. Что же касается объема достоверной информации, которую несет решение S^*, то он равен нулю.

Парадокс в том, мера $\mu(S^* \cap \widehat{S})$ Лебега общего фрагмента носителей S^* и $\widehat{S}$ безусловно превзойдет меру $\mu(D_2)$, но при этом остается неясным, о каком конкретно фрагменте $S^* \cap \widehat{S}$ идет речь – мы вообще не можем указать ни одной точки носителя S^*, достоверно, также принадлежащей и $\widehat{S}$. Мера же области D_2 пусть сравнительно невелика (а при большом уровне неопределенности может, вообще, оказаться, что $D_2 = \varnothing$), но, зато, можно гарантированно утверждать, что при условии адекватности всех модельных представлений все ее точки принадлежат неизвестному носителю $\widehat{S}$.

Но может возникнуть предположение, что если без учета достоверности решение S^* оказалось предпочтительней решения D_2 (в том смысле, что $\mu(S^* \cap S) > \mu(D_2)$, то совместно с оценкой точности прежние формы представления результатов интерпретации окажутся предпочтительней альтернативных теперь уже по величине объема извлекаемой достоверной информации о носителе $\widehat{S}$. Не говоря уже о том, что какие-либо оценки точности в известных детерминистских алгоритмах решения обратных задач обычно не приводятся, преобразование их в достоверную информацию об объекте исследования в этих алгоритмах не предусматривается. Предположим, однако, что такое действие будет выполнено. Покажем, что выделенное жирным шрифтом предположение окажется ошибочным.

Пример. Пусть известно, что $\widehat{S}$ – круг площади $\mu(\widehat{S}) = \pi$ (соответственно, модельный класс M – множество всех кругов меры π) и по информации G построено допустимое приближенное решение $S^* \in Q$. Пусть известно его качество

$$\nu(S^*, \widehat{S}) = \frac{\mu(S^* \cap \widehat{S})}{\mu(\widehat{S})} \qquad (19)$$

($\nu = 1$ при $S^* = \widehat{S}$; $\nu = 0$, если области S^* и $\widehat{S}$ не имеют общих точек). Тогда можно утверждать, что область

$$D = \left\{ \cap S_\alpha : S_\alpha \in M,\ \nu(S_\alpha, S^*) = \nu(S^*, \widehat{S}) \right\} \qquad (20)$$

является фрагментом носителя масс $\widehat{S}$. Если область D непустая, то она также является кругом (пусть r – его радиус), имеющим с $\widehat{S}$ общий центр. Чтобы этот круг имел радиус r, необходимо, чтобы

$$v(S^*,\widehat{S})=\frac{1}{\pi}\left(2\arccos\left(\frac{1-r}{2}\right)-\frac{1}{2}(1-r)\sqrt{(1+r)(3-r)}\right),$$

$$0\le r\le 1. \qquad (21)$$

Значение

$$v_0=\frac{4\pi-3\sqrt{3}}{6\pi}\approx 0{,}4 \qquad (22)$$

является пороговым: при $v<v_0$ оценка (19) не позволяет указать ни одной точки, гарантированно принадлежащей носителю $\widehat{S}$ ($D=\varnothing$); при $v=v_0$ достоверно идентифицировать на предмет принадлежности области $\widehat{S}$ можно лишь ее центр. Для того, например, чтобы из приближенного решения S^* выделить область меры $0{,}5\mu(\widehat{S})$, гарантированно заполненную возмующащими массами, необходимо, чтобы точность $v(S^*, \widehat{S})$ была не ниже 0,81.

Понятно, что при более сложной геометрии носителей масс различие между мерой $\mu(D)$ области, гарантированно принадлежащей носителю $\widehat{S}$, и мерой $\mu(S^*, \widehat{S})$ (неустановленной) области, являющейся общей для приближенного и истинного носителей, будет еще более впечатляющим – пороговое значение $\lambda = \lambda_0$ может приближаться к единице, тогда как мера области D_2, полученной непосредственно из информации G, при определенной геометрии границы множества Q может оказаться значительно больше $\mu(D)$.

Приведенный пример можно рассматривать и в аспекте методообразующей идеи фундаментального уровня, провозглашающей необходимость согласования множества допустимых решений обратной задачи [10] идеи, актуальной не только для геофизики [11]. Но идея построения множеств D_1 и D_2 и представляющих собой неулучшаемые двухсторонние оценки типа включений для изучаемого геометрического места точек $\widehat{S}$, занятого аномалиеобразующими массами, это ведь также не что иное, как один из способов согласования допустимых решений, имеющий, как оказалось, много преимуществ. Отличие лишь в том, что в традиционном понимании результатом такого согласования должно опять-таки стать единичное решение S^* обратной задачи, тогда как при использовании новых форм представления результатов интерпретации это некие общие свойства (инварианты) всех допустимых решений из Q.

Выводы

Из материалов статьи можно извлечь один урок: геофизик не должен отдавать на откуп математику постановку задачи, в особенности ее целевую установку и критерий состоятельности метода ее решения. Неправильное толкование математических результатов, слишком серьезное отношение к заманчивому предложению иметь возможность получения результата с наперед заданной, желаемой точностью (в действительности реализуемое лишь виртуально, на бумаге) не только задвинули на второй план проблемы, решение которых позволило бы существенно повысить качество интерпретации, но и во многом породило многочисленные стереотипные заблуждения, рассмотрение которых стало предметом рассмотрения в настоящей работе.

Ключевые слова: информация, достоверность, адекватность модели, результат интерпретации.

ЛИТЕРАТУРА

1. Страхов В.Н. Геофизика и математика // Физика Земли. – 1995. – № 12. – С. 4-23.
2. Страхов В.Н. Геофизика и математика. Методологические основы математической геофизики // Геофизика. – 2000. – № 1. – С. 3-18.
3. Блох Ю.И. Проблема адекватности интерпретационных моделей в гравиразведке и магниторазведке // Геофизический вестник. – 2004. – № 6. – С. 10-15.
4. Балк П.И. Столкновение геофизических и математических интересов – главный источник противоречий в современной теории интерпретации потенциальных полей // Геофизический журнал. – 2000. – Т. 22. – № 4. – С. 3-20.
5. Балк П.И. Столкновение геофизических и математических интересов – основной источник противоречий в современной теории интерпретации потенциальных полей // Физика Земли. – 2011. – № 3. – С. 85-96.
6. Балк П.И. Математический формализм и невостребованные идеи в теории интерпретации потенциальных полей // Геофизика. – 2002. – № 2. – С. 41-46.
7. Балк П.И. О надежности результатов количественной интерпретации гравитационных аномалий // Изв. АН СССР. Физика Земли. – 1980. – № 6. – С. 65-83.
7. Использование метода конечных элементов при интерпретации данных гравиразведки и магниторазведки / А.С. Долгаль, П.И. Балк [и др.] // Вестник КРАУНЦ. – 2012. – Том. 1. – № 19. – С. 108-127.
8. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы. – М. : Изд-во Бином, 2003. – 632 с.
9. Страхов В.Н. Геофизика и математика // Геофизика. – 2000. – № 1. – С. 3-18.
10. Костомаров Д.П., Зайцев Ф.С., Сучков Е.П. Построение сильно различающихся решений некоторого класса некорректных задач с неточно заданными входными данными // Доклады АН. – 2011. – Т. 437. – № 3. – С. 316-320.