Аппроксимационный подход к решению задач гравиметрии и магнитометрии. I. Основная вычислительная проблема – регуляризация систем линейных алгебраических уравнений

В. Н. Страхов, А. В. Страхов

Объединенный институт физики Земли им. О. Ю. Шмидта РАН

Это может показаться парадоксальным, но вся наука подчинена идее аппроксимации.

Бертран Рассел

Умный начинает с конца, дурак кончает в начале.

Дьердь Пойа

Введение

Данная работа, состоящая из четырех частей, посвящена центральной, по глубокому убеждению авторов, для гравиметрии и магнитометрии проблеме – разработке современной теории интерпретации потенциальных полей, адекватной реальной вычислительной практике, на базе фундаментальных методообразующих идей аналитической аппроксимации, алгебраизации и распознавания образов. Короче говоря, речь идет об изложении тех фундаментальных основ теории интерпретации, которые должны составить третью парадигму в этой теории, о чем неоднократно писал первый из авторов настоящей статьи, [см. *Страхов,* 1995a, 1995b, 1995c, 1996a, 1996b, 1996c, 1997a, 1997b, 1997c, 1998a, 1998b, 1998c, 1998d, 1998e].

Говоря о единой аппроксимационной основе всех методов, развиваемых в рамках третьей парадигмы, естественно поставить следующие вопросы:

1) каким образом решение задач редуцируется к решению задач аппроксимации?

2) какие типы аппроксимаций имеют для гравиметрии и магнитометрии определяющее значение?

3) какова та вычислительная задача, которая имеет при решении задач построения аппроксимаций центральное значение?

4) каковы методы решения той вычислительной задачи, которая играет центральную роль при построении аппроксимаций?

Статья N RJE99020.

Онлайновая версия этой статьи опубликована 25 июня 1999.
URL: http://eos.wdcb.rssi.ru/rjes/RJE99020/RJE99020.htm

Памятуя изречение знаменитого венгерского математика Д. Пойа, вынесенное в эпиграф, в первых двух частях работы, настоящей и следующей, рассматривается последний из четырех приведенных вопросов, исходя из следующих ответов на второй и третий вопросы:

2) важнейший класс аппроксимаций, используемый при решении задач гравиметрии и магнитометрии – это линейные аппроксимации глобального характера;

3) центральная вычислительная проблема, возникающая при построении линейных аппроксимаций глобального характера – это проблема нахождения устойчивых приближенных решений систем линейных алгебраических уравнений с приближенными данными (или по другому – проблема регуляризации систем линейных алгебраических уравнений с приближенными данными).

При этом очевидными являются следующие факты:

а) смысл термина "глобальная линейная аппроксимация" таков – речь идет о линейной аппроксимации, позволяющей описывать внешнее аномальное поле во всем априорно заданном объеме пространства и находимой по всей экспериментальной информации о поле в заданном объеме пространства;

б) в общем случае системы линейных алгебраических уравнений, устойчивые приближенные решения которых надо отыскивать, имеют приближенно заданные правые части и матрицы;

в) в настоящее время в повестке дня стоит проблема нахождения устойчивых приближенных решений систем линейных алгебраических уравнений с плотно заполненными матрицами, большой и даже сверхбольшой размерности, в которых число подлежащих определению неизвестных имеет порядок 10^4 (большая размерность) и 10^5 (сверхбольшая размерность); это с очевидностью обусловлено тем, что в настоящее время все задачи требуется решать в трехмерной (пространственной) постановке;

г) особо актуальной является проблема редуцирования задач гравиметрии и магнитометрии к решению систем линейных алгебраических уравнений с очень сильно разреженными матрицами, число неизвестных в которых можно брать весьма большим (порядка 10^6–10^7); эта проблема естественно приводит к использованию методов решения задач гравиметрии и магнитометрии, основанных на теориях дискретных физических полей (гравитационного, магнитного);

д) при постановке задач нахождения устойчивых приближенных решений систем линейных алгебраических уравнений с приближенными данными с очевидностью должна использоваться априорная информация об искомом решении и помехах (в задании правой части и матрицы); априорная информация должна быть адекватна реальной геофизической практике.

Как уже сказано выше, детальное рассмотрение первого, а также второго и третьего вопросов будет сделана в третьей части работы. В четвертой части будет приведено описание разработанных компьютерных технологий и примеры их использования.

Что касается настоящей – первой – части работы, то в ней рассматриваются классические постановки задач нахождения устойчивых приближенных решений линейных систем, дается классификация и краткое описание разработанных методов, при этом более подробно рассматриваются: 1) основной вариационный метод А. Н. Тихонова и 2) связанный с ним метод расширенных систем 2-го рода. Далее в первой части работы рассматриваются основы разработанной В. Н. Страховым, частично с соавторами, (см. [*Страхов,* 1990a, 1990b, 1991a, 1991b, 1991c, 1992a, 1992b, 1995b, 1997d, 1997e, 1997f, 1997g, 1997h, 1998d; *Страхов, Страхов,* 1998, 1999a, 1999b, 1999c, 1999d; *Страхов, Тетерин,* 1991a, 1991b, 1991c, 1993; Strakhov et al., 1995]) новой теории регуляризации систем линейных алгебраических уравнений с приближенными данными (третий раздел статьи), и дается классификация созданных в рамках этой теории методов регуляризации, в случае чисто аддитивной помехи в правой части – четвертый раздел первой части.

Во второй части работы приводится более подробное описание методов регуляризации линейных систем, разработанных в рамках новой теории.

1. Краткая характеристика классической теории регуляризации систем линейных алгебраических уравнений с приближенными данными

§ 1. Проблема нахождения приближенных решений систем линейных алгебраических уравнений с приближенными данными имеет, по существу, длительную историю, см., например, [*Альберт,* 1977; *Лоусон, Хенсон,* 1986]. Но громадный прогресс в решении этой проблемы был обусловлен созданием общей теории некорректно поставленных задач, основоположником которой является академик А. Н. Тихонов, и существенный вклад в которую внесли член-корреспондент В. К. Иванов и академик М. М. Лаврентьев и затем их ученики [*Алифанов и др.,* 1998; *Воеводин,* 1969; *Иванов,* 1962, 1966, 1977; *Иванов и др.,* 1978; *Лаврентьев,* 1955, 1959, 1962; *Лисковец,* 1981; *Морозов,* 1973, 1974, 1987; *Танана,* 1981; *Тихонов,* 1943, 1963a, 1963b, 1965; *Тихонов, Арсенин,* 1979; *Тихонов и др.,* 1985]. Ниже излагаются те результаты этой общей теории, которые относятся к проблеме регуляризации линейных систем с аддитивной помехой в задании правой части.

§ 2. Пусть требуется найти “достаточно хорошее” (“разумное”) приближенное решение (псевдорешение) системы линейных алгебраических уравнений

$$\mathrm{A}x = f, \tag{1}$$

по заданной аппроксимации этой системы

$$\mathrm{A}x = f_\delta = f + \delta f. \tag{2}$$

В (1)–(2) A суть (N×M)–матрица с вещественными элементами a_{ij}, $1 \le i \le N$, $1 \le j \le M$, N – число неизвестных (в общем случае N≠M), x есть искомый M-вектор, f_δ – заданный N-вектор, f есть вектор полезного сигнала, δf – вектор помехи (которая предполагается случайной).

В классической теории регуляризации систем линейных алгебраических уравнений (1)[1] всегда принимается, что векторы x и f, для которых выполняется соотношение (1), существуют. Одновременно предполагается, что правая часть в неравенстве

[1] А фактически – линейных операторных уравнений

$$\mathrm{A}x = f_\delta = f + \delta f,$$

где x и f_δ – элементы некоторых банаховых (полных нормированных) пространств X и Y, а A – ограниченный линейный оператор, действующий из X в Y; при этом чаще всего рассматриваются случаи, когда X и Y суть гильбертовы пространства.

$$\left\|\delta f\right\|_{\mathrm{E}} \leq \delta \tag{3}$$

известна[2].

В классической теории регуляризации систем (1) ставится вопрос о построении приближенных решений системы (2), в форме

$$x_\delta = \Re_\delta f_\delta, \tag{4}$$

где $\Re_\delta$ есть (M×M)-матрица, таких, что

$$\left\|x - x_\delta\right\|_{\mathrm{E}} \underset{\delta\to 0}{\to} 0, \tag{5}$$

где δ есть величина, фигурирующая в (3).

Алгоритмы, основанные на нахождении приближенных решений (4) со свойством (5), именуются *регуляризующими* или *регулярными*[3]. Для построения регуляризующих алгоритмов центральное значение имеет оценка абсолютной погрешности приближенного решения, так называемая схема Лаврентьева–Джона [*Лисковец,* 1981; *Морозов,* 1987; *Танана,* 1981]:

$$\begin{aligned} &\left\|x - x_\delta\right\| = \left\|x - \Re_\delta f_\delta\right\|_{\mathrm{E}} \leq \\ &\leq \left\|(\mathrm{E} - \Re_\delta \mathrm{A})x\right\|_{\mathrm{E}} + \delta\left\|\Re_\delta\right\|. \end{aligned} \tag{6}$$

При этом $\|\Re\|$ обозначает произвольную норму матрицы $\Re$, согласованную с эвклидовой нормой векторов, например эвклидова или спектральная.

Из (6) следует, что (5) будет иметь место, если одновременно будут выполняться соотношения

$$\left\|(\mathrm{E} - \Re_\delta \mathrm{A})x\right\|_{\mathrm{E}} \underset{\delta\to 0}{\to} 0, \tag{7}$$

$$\delta\left\|\Re_\delta\right\| \underset{\delta\to 0}{\to} 0\,. \tag{8}$$

§ 3. Исторически первые методы регуляризации задачи нахождения приближенных решений систем (1) по заданным их аппроксимациям (2) были построены для систем с N=M и симметричными положительно полуопределенными матрицами A:

$$\mathrm{A} = \mathrm{A}^{\mathrm{T}} \geq 0; \tag{9}$$

именно, М. М. Лаврентьевым был предложен [*Лаврентьев,* 1959] метод регуляризации, в случае (9) состоящий в переходе от системы (2) к системе

$$(\alpha \mathrm{S} + \mathrm{A})x_\alpha = f_\delta, \tag{10}$$

где $\alpha = \alpha(\delta) > 0$ есть параметр, именуемый параметром регуляризации, а S есть (M×M)-матрица со свойством

$$\mathrm{S} = \mathrm{S}^{\mathrm{T}} > 0, \tag{11}$$

и достаточно хорошо обусловленная. При этом М. М. Лаврентьев [*Лаврентьев,* 1959] установил регулярность метода, основанного на переходе от (2) к (10), на основе чисто спектральных соображений, не апеллируя к какой-либо вариационной постановке. Однако нетрудно найти, что система (10) возникает при рассмотрении условной экстремальной задачи:

$$\begin{aligned} &(\mathrm{S}x, x) = \min_x, \\ &(\mathrm{A}x, x) - 2(f_\delta, x) = \Delta, \end{aligned} \tag{12}$$

где Δ – заданная величина. Действительно, условной экстремальной задаче (12) методом множителей Лагранжа ставится в соответствие семейство безусловных экстремальных задач

$$\alpha(\mathrm{S}x, x) + (\mathrm{A}x, x) - 2(f_\delta, x) = \min_x, \tag{13}$$

где $\alpha > 0$ – параметр (величина, обратная множителю Лагранжа). Экстремаль задачи (13) как раз и удовлетворяет уравнению (10).

Для нахождения значения параметра регуляризации позже было предложено использовать условие невязки

$$\left\|f_\delta - \mathrm{A}x_\alpha\right\|_{\mathrm{E}}^2 = \delta^2, \tag{14}$$

где x_α есть решение уравнения (10). Было доказано, что если

$$\eta^2 = \frac{\delta^2}{\left\|f_\delta\right\|_{\mathrm{E}}^2} < 1, \tag{15}$$

то решение уравнения (14) существует и единственно. Далее А. Н. Тихоновым был предложен метод, ныне именуемый вариационным методом А. Н. Тихонова, который применим к любым системам (2), и который основан на постановке условной вариационной задачи

[2] Тот факт, что в (3), равно как во всех последующих соотношениях, фигурирует именно *эвклидова норма вектора*, не является определяющим. С равным основанием можно пользоваться и другими нормами векторов, а также согласованными с ними нормами матриц. Однако с точки зрения геофизических приложений (прежде всего с точки зрения свойств помех в данных), равно как и с точки зрения использования свойств конечномерных гильбертовых – эвклидовых – пространств, использование именно эвклидовой нормы векторов предпочтительнее всего.

[3] В подавляющем большинстве работ 60–70-х годов А. Н. Тихоновым и его последователями, [см. *Воеводин,* 1969; *Иванов,* 1962, 1966, 1977; *Иванов и др.,* 1978; *Лаврентьев,* 1955, 1959; *Морозов,* 1973, 1974; *Тихонов,* 1943, 1963a, 1963b, 1965; *Тихонов, Арсенин,* 1979], использовался термин *регуляризирующий* алгоритм. Термином регуляризующий впервые воспользовался В. Н. Страхов в работах [*Страхов,* 1968a, 1968b]; постепенно и другие ученые начали его использовать.

$$\begin{aligned} &\Omega(x) = \left\|\mathrm{R}x\right\|_{\mathrm{E}}^{2} = \min_{x}, \\ &\left\|f_\delta - \mathrm{A}x\right\|_{\mathrm{E}}^{2} = \delta^2, \end{aligned} \tag{16}$$

где $\delta^2 = \left\|\delta f\right\|_{\mathrm{E}}^{2}$ – заданная постоянная.

Ясно, что кроме основной экстремальной задачи (16) можно рассматривать и двойственную к ней

$$\begin{aligned} &\left\|f_\delta - \mathrm{A}x\right\|_{\mathrm{E}}^{2} = \min_{x}, \\ &\Omega(x) = \left\|\mathrm{R}x\right\|_{\mathrm{E}}^{2} = \mathrm{C}^2, \end{aligned} \tag{17}$$

где C^2 – заданная постоянная. В (16)–(17) R суть заданная матрица, размера P×M, где в общем случае P≠M, которая именуется матрицей-регуляризатором, соответственно функционал $\Omega(x)$ именуется *функционалом, задающим отношение предпочтения на множестве приближенных решений системы* (2). Последнее означает следующее: если $x^{(1)}$ и $x^{(2)}$ суть приближенные решения (2), для которых

$$\left\|f_\delta - \mathrm{A}x^{(2)}\right\|_{\mathrm{E}}^{2} = \left\|f_\delta - \mathrm{A}x^{(1)}\right\|_{\mathrm{E}}^{2}, \tag{18}$$

то из неравенства

$$\Omega(x^{(2)}) = \left\|\mathrm{R}x^{(2)}\right\|_{\mathrm{E}}^{2} < \left\|\mathrm{R}x^{(1)}\right\|_{\mathrm{E}}^{2} = \Omega(x^{(1)}) \tag{19}$$

следует, что приближенное решение $x^{(2)}$ предпочтительнее приближенного решения $x^{(1)}$. При этом предполагается, что матрица $\mathrm{R}^{\mathrm{T}}\mathrm{R}$ является невырожденной и достаточно хорошо обусловленной. Подчеркнем, что понятие "функционал, задающий отношение предпочтения на множестве приближенных решений системы (2)", введено В. Н. Страховым, основоположник же метода А. Н. Тихонов именовал функционал $\Omega(x)$ *стабилизирующим* [*Тихонов, Арсенин,* 1979]. Соответственно функционал, сопоставляемый любой из условных экстремальных задач (16) и (17),

$$\alpha\left\|\mathrm{R}x\right\|_{\mathrm{E}}^{2} + \left\|f_\delta - \mathrm{A}x\right\|_{\mathrm{E}}^{2}, \tag{20}$$

А. Н. Тихонов именовал "сглаживающим параметрическим функционалом", а параметр $\alpha > 0$, фигурирующий в (20) – параметром регуляризации.

Ясно, что в вариационном методе регуляризации решение $x = x_\alpha$ ищется из условия

$$\alpha\left\|\mathrm{R}x\right\|_{\mathrm{E}}^{2} + \left\|f_\delta - \mathrm{A}x\right\|_{\mathrm{E}}^{2} = \min_{x}; \tag{21}$$

решение x_α существует, единственно и удовлетворяет системе линейных алгебраических уравнений с невырожденной матрицей

$$(\alpha\mathrm{R}^{\mathrm{T}}\mathrm{R} + \mathrm{A}^{\mathrm{T}}\mathrm{A})x_\alpha = \mathrm{A}^{\mathrm{T}} f_\delta = \varphi_\delta. \tag{22}$$

Очевидно, для нахождения значения α и в данном случае может быть использовано уравнение (14), решение которого опять-таки существует и единственно, если выполняется неравенство (15).

Ясно, что величина $\alpha > 0$ является множителем Лагранжа в случае условной вариационной задачи (17), и обратной к множителю Лагранжа – в случае условной вариационной задачи (16).

Заметим, наконец, что уравнение (22) можно рассматривать как регуляризованное по М. М. Лаврентьеву уравнение

$$\mathrm{A}^{\mathrm{T}}\mathrm{A}x_\alpha = \mathrm{A}^{\mathrm{T}} f_\delta = \varphi_\delta, \tag{23}$$

которое возникает в рамках решения классического *метода наименьших квадратов*, основанного на постановке экстремальной задачи:

$$\left\|f_\delta - \mathrm{A}x\right\|_{\mathrm{E}}^{2} = \min_{x}. \tag{24}$$

Действительно, при регуляризации по М. М. Лаврентьеву системы (23) достаточно принять

$$\mathrm{S} = \mathrm{R}^{\mathrm{T}}\mathrm{R}, \tag{25}$$

чтобы получить систему (10).

Отметим еще, что в методе наименьших квадратов (24) *выполняется характеристическое уравнение этого метода*:

$$(\mathrm{A}x, f_\delta - \mathrm{A}x) = 0. \tag{26}$$

Но в случае регуляризованного уравнения (22) данное соотношение не выполняется. Действительно, имеем

$$\begin{aligned} &(\mathrm{A}x_\alpha, f_\delta - \mathrm{A}x_\alpha) = (x_\alpha, \mathrm{A}^{\mathrm{T}} f_\delta - \mathrm{A}^{\mathrm{T}}\mathrm{A}x_\alpha) \\ &= (x_\alpha, \alpha\mathrm{R}^{\mathrm{T}}\mathrm{R}x_\alpha) = \alpha\left\|\mathrm{R}x_\alpha\right\|_{\mathrm{E}}^{2} > 0. \end{aligned} \tag{27}$$

§ 4. В рамках классического вариационного метода А. Н. Тихонова возникает система линейных алгебраических уравнений (22) с матрицей размера (M×M). Ясно, что в случае систем (2) с N>M это является вполне приемлемым. Однако в случае систем (2) с N<M переходить к регуляризованным системам размера (M×M) нежелательно. Желательно было бы иметь регуляризованную систему размерности (N×N). Чтобы лучше уяснить возможность получения подобных регуляризованных систем, вспомним, что система (23), возникающая при

использовании метода наименьших квадратов (24), представляет собой ту, которая возникает при использовании первой трансформации Гаусса исходной системы (2). Следовательно, искомую регуляризованную систему с матрицей размера (N×N) можно получить, сначала переходя к системе, возникающей при второй трансформации Гаусса исходной системы (2):

$$\mathrm{AA^T} z = f_\delta, \quad x = \mathrm{A^T} z, \tag{28}$$

а затем регуляризуя данную систему с симметричной положительно полуопределенной матрицей по М. М. Лаврентьеву:

$$(\alpha\Sigma + \mathrm{AA^T}) z_\alpha = f_\delta; \tag{29}$$

имеем далее

$$x_\alpha = \mathrm{A^T} z_\alpha; \tag{30}$$

при этом в (29)

$$\Sigma = \Sigma^{\mathrm{T}} > 0 \tag{31}$$

есть достаточно хорошо обусловленная матрица размера (N×N). Значение параметра α и в данном случае целесообразно находить из условия (14); решение задачи опять-таки существует и единственно, если выполняется неравенство (15).

Ясно, что матрицу Σ целесообразно выбирать в виде

$$\Sigma = \Re\Re^{\mathrm{T}}, \tag{32}$$

где $\Re$ есть (N×Q)-матрица. Очевидно, $\Re$ также можно трактовать как матрицу-регуляризатор.

Как это не покажется парадоксальным, но конструкция, описанная в данном параграфе – новая.

§ 5. Методы, описанные в §§ 3–4, являются наиболее типичными для классической теории регуляризации систем линейных алгебраических уравнений. Но в рамках данной теории были предложены и разработаны (правда, с различной степенью детальности) еще по крайней мере четыре метода:

1) метод, основанный на переходе к так называемой "расширенной регуляризованной системе" линейных алгебраических уравнений;

2) метод, основанный на нахождении сингулярного разложения матрицы A системы (2) и "проекционной регуляризации" [*Иванов и др.,* 1978; *Лисковец,* 1981; *Морозов,* 1974];

3) метод *итеративной регуляризации*, основанный на некотором итерационном процессе – для системы (2), либо (что используется гораздо чаще) для систем (23) или (28), получаемых из (2) с помощью первой или второй трансформаций Гаусса; правило останова процесса в рамках итеративной регуляризации – центральная (регуляризующая) процедура, [см. *Алифанов и др.,* 1998; *Иванов и др.,* 1978; *Морозов,* 1974, 1987; *Танана,* 1981; *Тихонов и др.,* 1985];

4) метод, основанный на использовании градиентных процедур поиска минимума квадратичных функционалов, при использовании соответствующего правила останова в процедуре спуска; важнейший частный случай здесь тот, когда применяется метод сопряженных градиентов, см. [*Алифанов и др.,* 1998].

Первый из указанных методов регуляризации будет достаточно подробно описан в следующем разделе настоящей статьи, второй – в следующем параграфе. Что же касается третьего и четвертого методов (точнее – групп методов), то они очень кратко описываются в предпоследнем § 7 данного раздела. Наконец, в заключительном § 8 делаются некоторые дополнительные замечания по основным методам вариационного типа.

§ 6. Суть метода проекционной регуляризации в следующем. Пусть A есть (M×M)-матрица, и пусть существуют такие ортогональные по столбцам матрицы V и U, размером (N×N) и (M×M), что:

А) в случае N>M

$$\mathrm{V^T AU} = \begin{array}{|c|c} \hline \Sigma_{\mathrm{A}} & \updownarrow \mathrm{M} \\ \hline 0 & \mathrm{N} - M \updownarrow \\ \hline \end{array}; \tag{33}$$

Б) в случае N=M

$$\mathrm{V^T AU} = \Sigma_{\mathrm{A}}; \tag{34}$$

В) в случае N<M

$$\mathrm{V^T AU} = \begin{array}{c} \begin{array}{|c|c|} \hline \quad \Sigma_{\mathrm{A}} \quad & \quad 0 \quad \\ \hline \end{array} \\ \leftarrow N \rightarrow\leftarrow M - N \rightarrow \end{array} \tag{35}$$

Во всех трех случаях Σ_{A} есть диагональная матрица:

$$\Sigma_{\mathrm{A}} = \operatorname{diag}\sigma_i, \quad 1 \le i \le \min(N, M), \tag{36}$$

где $\sigma_i = \sigma_i(\mathrm{A})$ суть сингулярные числа матрицы A, упорядоченные по невозрастанию:

$$\sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_n \ge 0, \quad n = \min(N, M). \tag{37}$$

При этом, если A – невырожденная матрица, то $\sigma_{i_{\max}} > 0$, где $i_{\max} = \min(N, M)$, в вырожденном же случае $\sigma_i > 0$ только при всех $i \le i_0$, $1 \le i_0 < \min(N, M)$.

Очевидно, если даже A невырожденная (в общем случае прямоугольная) матрица, она может быть сколь угодно плохо обусловленной, то есть отношение

$$\operatorname{Cond} \mathrm{A} = \frac{\sigma_{\max}}{\sigma_{\min}} \tag{38}$$

может быть сколь угодно велико.

Итак, пусть сингулярное разложение A найдено, то есть получено одно из трех соотношений – (33), (34), (35), где Σ_{A} есть диагональная матрица (36).

Ясно, что в случае (33) получаем систему с диагональной матрицей

$$\Sigma_{\mathrm{A}} w = \widehat{\psi}_\delta, \tag{39}$$

где

$$w = \mathrm{U}^{\mathrm{T}} x, \tag{40}$$

а вектор $\widehat{\psi}_\delta$ определен разложением

$$\psi_\delta = \mathrm{V}^{\mathrm{T}} f_\delta = \begin{array}{|c|c} \widehat{\psi}_\delta & \updownarrow M \\ \hline \breve{\psi}_\delta & \updownarrow N - M \end{array}. \tag{41}$$

Ясно, что вектор $\widehat{\psi}_\delta$ представим в форме

$$\widehat{\psi}_\delta = \breve{\psi} + \delta\widehat{\psi}, \tag{42}$$

и при этом

$$\left\| \delta\widehat{\psi} \right\|_{\mathrm{E}}^2 \leq \delta^2 - \left\| \breve{\psi}_\delta \right\|_{\mathrm{E}}^2 = \Delta^2. \tag{43}$$

Основная процедура вариационного метода состоит в том, что находится наибольший из номеров i (этот номер обозначается $i_{гр}$) сингулярного числа σ_i, такой, что выполняются два неравенства

$$\sum_{i=i_{гр}+1}^{i_{\max}} \widehat{\psi}_{i,\delta}^2 \leq \Delta^2, \qquad \sum_{i=i_{гр}}^{N} \widehat{\psi}_{i,\delta}^2 > \Delta^2. \tag{44}$$

В этом случае принимается

$$w_{рег} = \begin{array}{|c|c} \widehat{w} & \updownarrow i_{гр} \\ \hline 0 & \updownarrow i_{\max} - i_{гр} \end{array}; \tag{45}$$

далее используется соотношение

$$x_{рег} = \mathrm{U} w_{рег}. \tag{46}$$

В случаях же (34) и (35) эквивалентная система с диагональной матрицей имеет вид

$$\Sigma_{\mathrm{A}} w = \psi_\delta, \tag{47}$$

где

$$w = \mathrm{U}^{\mathrm{T}} x, \quad \psi_\delta = \mathrm{V}^{\mathrm{T}} f_\delta. \tag{48}$$

В этом случае опять-таки находится номер $i = i_{гр}$, такой, что имеют место неравенства типа (44) – в последних надо только заменить $\widehat{\psi}_{i,\delta}$ на $\psi_{i,\delta}$, а значение Δ^2 на значение δ^2. Дальнейшее, то есть соотношения (45) и (46), остается в силе.

§ 7. *Метод итеративной регуляризации* по существу был разработан уже на начальном этапе развития теории некорректно поставленных задач, в 60-е годы (см., например, работы [*Алифанов и др.*, 1998; *Алифанов, Румянцев*, 1978, 1980; *Бакушинский*, 1977; *Бакушинский, Страхов*, 1968; *Вайникко*, 1982; *Васильев, Якимович*, 1980; *Гилязов*, 1980; *Емелин, Красносельский*, 1978; *Оганесян, Старостенко*, 1978; *Румянцев*, 1985; *Рязанцева*, 1978; *Савр*, 1984; *Страхов*, 1973a, 1973b]). Его суть очень проста. Пусть задан какой-либо итерационный процесс (процесс построения последовательных приближений $x^{(n)}$, $n = 0, 1, 2, \ldots$), относительно которого известно, что в случае точных данных, т.е. системы (1), он сходится к решению этой системы. Тогда этот процесс может быть применен и к системе (2), но с одним отличием – после того, как для некоторого $x^{(n)}$ впервые оказалось

$$\nu_n = \left\| f_\delta - \mathrm{A} x^{(n)} \right\|_{\mathrm{E}}^2 \leq \delta^2, \tag{49}$$

где δ^2 – постоянная, фигурирующая в неравенстве (3) (см. также (14)), итерационный процесс останавливается, и приближение $x^{(n)}$ принимается за искомое. Далее ясно, что поскольку наиболее полно и глубоко теория итерационных процессов разработана для систем с квадратными матрицами $\mathrm{A} = \mathrm{A}^{\mathrm{T}} \geq 0$, то и регуляризованные итерационные процессы (с правилом останова по невязке) чаще всего используются применительно к системам, получаемым из исходной (2) с помощью первой или второй трансформаций Гаусса. На подробностях здесь нет возможности останавливаться.

Аналогичным образом устроены и регуляризованные алгоритмы, в основу которых положены так называемые методы сопряженных направлений (чаще всего – методы $\mathrm{A}^{\mathrm{T}}\mathrm{A}$- и $\mathrm{A}\mathrm{A}^{\mathrm{T}}$-сопряженных градиентов, [см. *Воеводин, Кузнецов*, 1984; *Марчук, Кузнецов*, 1972; *Уилкинсон, Райнш*, 1976; *Фаддеев, Фаддеева*, 1963]), а также другие градиентные методы минимизации функционалов. И здесь реализуется та же схема:

1) используемый метод должен обеспечивать сходимость к решению (к одному из решений или псевдорешений) в случае системы (1) с точными данными;

2) в случае системы (2) с приближенными данными регуляризация обеспечивается правилом останова по критерию невязки.

§ **8**. В заключение раздела остановимся еще на четырех важных вопросах:

во-первых, на вопросе о порядке величины параметра регуляризации α, находимого из уравнения невязки, в случае вариационных методов;

во-вторых, на вопросе о фактическом решении регуляризованных уравнений, возникающих в вариационных методах;

в-третьих, на вопросе регуляризации в ситуации, когда с погрешностью заданы и правая часть, и матрица системы;

в-четвертых, о трудоемкости вариационных методов и путях ее снижения.

Первый вопрос. Классическое правило А. Н. Тихонова (см. [*Иванов и др.,* 1978; *Лисковец,* 1981; *Морозов,* 1974, 1987; *Танана,* 1981; *Тихонов,* 1963a; *Тихонов, Арсенин,* 1979]) для оптимального значения параметра регуляризации α (при котором выполняется уравнение невязки (14)) таково

$$\alpha = \mathrm{O}(\delta^2). \tag{50}$$

Это правило в дальнейшем было уточнено В. Н. Страховым [*Страхов,* 1973b]

$$\alpha = \mathrm{O}(\eta^2), \quad \eta^2 = \frac{\delta^2}{\left\| f_\delta \right\|_{\mathrm{E}}^2}. \tag{51}$$

Предлагались и иные правила. Но, по-видимому, наибольшее практическое значение имеет правило В. Н. Страхова, относящееся к случаю, когда спектры помехи и полезного сигнала в существенном разделены:

$$\alpha = \mathrm{O}(\lambda_{разд}), \tag{52}$$

где $\lambda_{разд}$ – постоянная разделения спектров полезного сигнала и помехи; при этом рассматриваются все три метода – М. М. Лаврентьева (9)–(14), А. Н. Тихонова (18)–(22), (14) и новый метод (28)–(32), (14). Принимается, что λ – вещественная переменная, λ_i – собственные значения соответствующей симметричной положительно полуопределенной матрицы (A – в методе М. М. Лаврентьева, $\mathrm{A^TA}$ – в методе А. Н. Тихонова, $\mathrm{AA^T}$ – в новом методе), удовлетворяющие неравенствам:

$$\lambda_1 > \lambda_2 > \dots \lambda_{i_{\max}} \geq 0, \tag{53}$$

где $i_{\max} = M$ в случае матриц A и $\mathrm{A^TA}$, и $i_{\max} = N$ в случае матрицы $\mathrm{AA^T}$. Если матрица вырожденная, то отличными от нуля являются лишь λ_i при $i \leq i_0$, $i_0 < i_{\max}$. Пусть $Ker(\bullet)$ – ядро соответствующей матрицы (A, $\mathrm{A^TA}$, $\mathrm{AA^T}$), и пусть e_i суть соответствующие ненулевым λ_i ортонормированные собственные векторы.

Пусть

$$f = \sum_{i=1}^{i_0} c_i(f) e_i \tag{54}$$

и

$$\delta f = \delta_0 f + \sum_{i=1}^{i_0} c_i(\delta f) e_i \tag{55}$$

суть разложения полезного сигнала и помехи по ортонормированным собственным векторам e_i, $\delta_0 f \in Ker(\bullet)$. Спектры полезного сигнала и помехи считаются в существенном разделенными, если существует такое число $\lambda_i = \lambda_{разд}$, что одновременно выполняются два неравенства:

$$\frac{\sum\limits_{i=1}^{\lambda_{разд}} c_i^2(\delta f)}{\sum\limits_{i=1}^{\lambda_{разд}} c_i^2(f)} \ll 1, \tag{56}$$

и

$$\frac{\sum\limits_{i=\lambda_{разд}+1}^{i_0} c_i^2(\delta f)}{\sum\limits_{i=\lambda_{разд}+1}^{i_0} c_i^2(f)} \gg 1. \tag{57}$$

Именно постоянная разделения спектров $\lambda_{разд}$, а не величины δ^2 или η^2, определяют оптимальное значение α в ситуации существенной разделенности спектров полезного сигнала f и помехи δf.

Второй вопрос. Наиболее простой способ найти оптимальное значение $\alpha = \alpha_{opt}$ и соответствующий вектор $x_{a_{opt}}$ во всех трех методах вариационного типа (М. М. Лаврентьева, А. Н. Тихонова и новом) состоит в том, что задается последовательность значений $\alpha = \alpha_k$, $k = 1, 2, \dots, k_{\max}$, и для каждого из них решается соответствующее регуляризованное уравнение – на основе использования алгоритма Холецкого. Например, в случае уравнения (10) имеем

$$\alpha_k \mathrm{S} + \mathrm{A} = \mathrm{L}(\mathrm{S}; \alpha_k)\mathrm{L^T}(\mathrm{S}; \alpha_k), \tag{58}$$

где $\mathrm{L}(\mathrm{S}; \alpha_k)$ суть нижняя треугольная матрица размера M×M. После этого решаются два уравнения с треугольными матрицами

$$\mathrm{L}(\mathrm{S}; \alpha_k) w_{\alpha_k} = f_\delta, \tag{59}$$

$$\mathrm{L^T}(\mathrm{S}; \alpha_k) x_{\alpha_k} = w_{\alpha_k}. \tag{60}$$

Далее вычисляется величина

$$\nu_{\alpha_k} = \left\| f_\delta - \mathrm{A} x_{\alpha_k} \right\|_{\mathrm{E}}^2 \tag{61}$$

и сравнивается с δ^2. Суть поискового алгоритма очевидна. В случае методов А. Н. Тихонова и нового все обстоит совершенно аналогично.

Третий вопрос. Итак, пусть для уравнения (1) задана не аппроксимация (2), в которой правая часть задана с погрешностью, а аппроксимация

$$\mathrm{A}_\Delta x = f_\delta, \tag{62}$$

где

$$\mathrm{A}_\Delta = \mathrm{A} + \Delta\mathrm{A}, \quad f_\delta = f + \delta f, \tag{63}$$

т.е. приближенно заданы и правая часть, и матрица системы.

Ясно, что в этом случае можно использовать все три описанных выше вариационных метода (М. М. Лаврентьева, А. Н. Тихонова и новый), при этом основная проблема состоит в задании оптимального значения параметра α.

С очевидностью имеем неравенство

$$\left\| f_\delta - \mathrm{A} x_\alpha \right\|_{\mathrm{E}} \leq \left\| f_\delta - \mathrm{A}_\Delta x_\alpha \right\|_{\mathrm{E}} + \Delta \left\| x_\alpha \right\|_{\mathrm{E}}, \tag{64}$$

где Δ определена неравенством

$$\left\| \Delta\mathrm{A} \right\|_{\mathrm{E}} \leq \Delta. \tag{65}$$

Поэтому оптимальным следует считать то (наибольшее из возможных) значение α, при котором одновременно выполняются два неравенства

$$\left\| f_\delta - \mathrm{A}_\Delta x_\alpha \right\|_{\mathrm{E}} \leq \delta, \quad \Delta \left\| x_\alpha \right\|_{\mathrm{E}} \leq \delta. \tag{66}$$

Если подобное значение α не существует, то оптимальным является то, для которого

$$\left\| f_\delta - \mathrm{A}_\Delta x_\alpha \right\|_{\mathrm{E}} + \Delta \left\| x_\alpha \right\|_{\mathrm{E}} = \min_\alpha. \tag{67}$$

Четвертый вопрос исключительно важен в связи с тем, что в настоящее время в гравиметрии и магнитометрии остро стоит вопрос о решении систем большой (число неизвестных порядка 10^4) и сверхбольшой (число неизвестных порядка 10^5 и более) размерности.

Ясно, что наиболее трудоемким является проекционный метод (см. § 6), основанный на нахождении сингулярного разложения матриц; его трудоемкость оценивается числами арифметических операций, по порядку равными:

$$\mathrm{O}(\mathrm{NM}^2), \qquad \mathrm{N} \geq \mathrm{M} \tag{68}$$

и

$$\mathrm{O}(\mathrm{N}^2\mathrm{M}), \qquad \mathrm{N} < \mathrm{M}, \tag{69}$$

но при этом постоянные, фигурирующие в символах порядка O, как правило весьма велики – 10 и более.

Естественно, что трудоемкость итерационных методов и методов сопряженных направлений (сопряженных градиентов) оценить наиболее трудно – здесь все зависит от спектров полезного сигнала и помехи, а также от выбора начального приближения. Но в целом это наиболее перспективные, с точки зрения их трудоемкости, методы.

Весьма трудоемкими являются все три вариационных метода, но из них привлекательнее всего выглядит метод М. М. Лаврентьева, в котором отсутствует процедура перемножения матриц. Принимая, что при реализации разложения Холецкого используется аккуратный алгоритм счета скалярного произведения (в котором отдельно суммируются положительные и отрицательные частные произведения) и что используется n различных значений $\alpha = \alpha_k$, найдем, что в главном члене (без учета операций при счете эвклидовых норм невязок) трудоемкость метода составляет

$$\mathrm{P} = \frac{n}{4}\mathrm{M}^3 \tag{70}$$

арифметических операций; при этом обычно n колеблется от 24 до 48.

В случае методов А. Н. Тихонова и нового (см. (28)–(32)) к числу P по (70) должны быть добавлены числа операций, затрачиваемых на нахождение произведений матриц:

$$\mathrm{K}_1 = (\mathrm{P} + \mathrm{N})\mathrm{M}^2, \tag{71}$$

где (P×M) размерность матрицы R, и

$$\mathrm{K}_2 = (\mathrm{Q} + \mathrm{M})\mathrm{N}^2, \tag{72}$$

где (N×Q) – размерность матрицы $\Re$.

Таким образом, трудоемкость всех вариационных методов достаточно велика, при этом главный источник больших объемов арифметических операций – многократное использование разложений Холецкого – при всех значениях $\alpha = \alpha_k$, $k = 1, 2, \ldots, n$.

В связи с этим В. В. Воеводиным [*Воеводин,* 1969; *Воеводин, Кунецов,* 1984] было предложено – для случая, когда в (10) S = E, в (22) R = E, в (28) $\Re$ = E, использовать операцию приведения основной матрицы (A – в методе М. М. Лаврентьева, $\mathrm{A}^\mathrm{T}\mathrm{A}$ – в методе А. Н. Тихонова, $\mathrm{A}\mathrm{A}^\mathrm{T}$ – в новом методе) к трехдиагональной форме.

Именно, в случае (10) при S = E используется ортогональное преобразование

$$\mathrm{V}^\mathrm{T}\mathrm{A}V = \mathrm{T}, \tag{73}$$

где V – ортогональная по столбцам матрица размера M×M, $T = T^T \geq 0$ – трехдиагональная матрица, и система редуцируется к такой:

$$(\alpha E + T) y_\alpha = \mathring{f}_\delta, \qquad y_\alpha = V^T x_\alpha, \quad \mathring{f}_\delta = V^T f_\delta. \tag{74}$$

Соответственно в случае (22) при R=E используется аналогичное ортогональное преобразование

$$U^T A^T A U = T, \tag{75}$$

где U – ортогональная по столбцам (M×M)-матрица, $T = T^T \geq 0$ – трехдиагональная (M×M)-матрица, и система редуцируется к виду

$$(\alpha E + T) z_\alpha = \psi_\delta, \qquad z_\alpha = U^T x_\alpha, \quad \psi_\delta = U^T \varphi_\delta. \tag{76}$$

Наконец, в случае (28), (32), при $\Re = E$ используется ортогональное преобразование

$$\Phi^T A A^T \Phi = \mathring{T}, \tag{77}$$

где Φ – ортогональная по столбцам (N×N)-матрица, $\mathring{T} = \mathring{T}^T \geq 0$ – трехдиагональная (N×N)-матрица, и система редуцируется к виду

$$(\alpha E + \mathring{T}) w_\alpha = \mathring{\varphi}_\delta, \qquad w_\alpha = \Phi z_\alpha, \quad \mathring{\varphi}_\delta = \Phi^T f_\delta. \tag{78}$$

Ясно, что системы с трехдиагональными матрицами решаются очень эффективно, и что числа арифметических операций, затрачиваемых на приведение к трехдиагональной форме, существенно меньше числа P по (70). Однако следует иметь в виду, что при больших значениях N и M, т.е. при большой и сверхбольшой размерности системы, процедуры приведения к трехдиагональному виду не всегда обеспечивают получение необходимой точности – ибо ошибки округления с неизбежностью накапливаются. Поэтому применение приема предварительного приведения к трехдиагональной форме в случае систем большой и сверхбольшой размерности рекомендовано быть не может.

2. Метод регуляризованных расширенных систем линейных алгебраических уравнений 2-го рода и некоторые свойства метода вариационной регуляризации

§ 1. Рассматриваемый в данном разделе первой части работы метод регуляризованных расширенных систем линейных алгебраических уравнений 2 - го рода безусловно имеет определенную историю (см. работы [*Альберт,* 1977; *Лоусон, Хенсон,* 1986]), однако он до самого последнего времени находился как бы в "тени", во всяком случае он не подвергался достаточно глубокому изучению и различным обобщениям.

Отправной пункт в методе – введение вектора невязки

$$r = f_\delta - Ax \tag{79}$$

и объединение уравнений (23) и (79) в одну систему[4]

$$Sw = \left|\begin{array}{c|c} E & A \\ \hline -A^T & 0 \end{array}\right| \times \left|\begin{array}{c} r \\ \hline x \end{array}\right| = \left|\begin{array}{c} f_\delta \\ \hline 0 \end{array}\right| = F_\delta. \tag{80}$$

Тот факт, что из (80) следует и (23) и (79), легко проверяется непосредственно.

Регуляризация системы (80) осуществляется исключительно просто – необходимо нулевой блок в правом нижнем углу матрицы S (которую не следует путать с матрицей-регуляризатором в методе М. М. Лаврентьева, см. (10)–(13)) заменить на матрицу $\alpha R^T R$, где R – матрица-регуляризатор, фигурирующая в методе А. Н. Тихонова, см. (21)–(22). Таким образом регуляризованная система имеет вид:

$$S_\alpha w_\alpha = \left|\begin{array}{c|c} E & A \\ \hline -A^T & \alpha R^T R \end{array}\right| \times \left|\begin{array}{c} r_\alpha \\ \hline x_\alpha \end{array}\right| = \left|\begin{array}{c} f_\delta \\ \hline 0 \end{array}\right| = F_\delta. \tag{81}$$

Из вида матрицы S_α, которая представляет собой сумму симметричной и кососимметричной матриц,

$$S_\alpha = \left|\begin{array}{c|c} E & 0 \\ \hline 0 & \alpha R^T R \end{array}\right| + \left|\begin{array}{c|c} 0 & A \\ \hline -A^T & 0 \end{array}\right|, \tag{82}$$

причем симметричная матрица невырожденная, следует, что при любом $\alpha > 0$ матрица S_α, определенная по (82), является невырожденной.

§ 2. Ясно, что переход к регуляризованной расширенной системе (81) имеет то достоинство, что не нужно находить произведение матриц $A^T A$; ясно также, что в случае R = E имеем $R^T R = E$ и второе произведение матриц находить также не надо. Кроме того, если N≫M, то при любом α система (81) есть система с блочно-стреловидной матрицей и здесь может быть эффективно использована компактная схема разложения матрицы на треугольные множители.

---

[4]Введение расширенной системы может быть осуществлено различными способами; здесь мы придерживаемся способа, предложенного В. Н. Страховым [*Страхов,* 1998d], см. также работы [*Страхов, Страхов,* 1999c, 1999d].

§ **3**. Однако главные достоинства перехода к расширенной регуляризованной системе в другом:

во-первых, возникает возможность более глубокого и детального анализа свойств вариационного метода А. Н. Тихонова;

во-вторых, возникает возможность глубокого обобщения вычислительных алгоритмов на случай наличия дифференцированной информации о векторе помехи.

В данном параграфе осуществляется нахождение ряда принципиально важных соотношений, имеющих место в вариационном методе А. Н. Тихонова.

Прежде всего, если исключить из системы (81) вектор r_α,

$$r_\alpha = f_\delta - \mathrm{A}x_\alpha, \tag{83}$$

то получим определяющее соотношение метода А. Н. Тихонова

$$(\alpha \mathrm{R^T R} + \mathrm{A^T A})x_\alpha = \mathrm{A^T} f_\delta = \varphi_\delta. \tag{84}$$

Далее – если, наоборот, исключать из системы (81) вектор x_α, то получим следующую совокупность формул:

$$x_\alpha = \mu(\mathrm{R^T R})^{-1}\mathrm{A^T} r_\alpha, \tag{85}$$

$$\widetilde{f}_\alpha = \mathrm{A}x_\alpha = \mu \mathrm{A}(\mathrm{R^T R})^{-1}\mathrm{A^T} r_\alpha, \tag{86}$$

$$(\mathrm{E} + \mu \mathrm{A}(\mathrm{R^T R})^{-1}\mathrm{A^T})r_\alpha = f_\delta. \tag{87}$$

При этом в (85)–(87) принято

$$\mu = \frac{1}{\alpha}. \tag{88}$$

Из (85)–(88) следуют такие принципиально важные факты:

1) вектор x_α выражается через r_α, иначе – искомое регуляризованное решение выражается линейно через вектор невязки для регуляризованного решения;

2) вектор $\widetilde{f}_\alpha = \mathrm{A}x_\alpha$, представляющий собой приближенное описание вектора полезного сигнала f, выражается линейно через вектор невязки регуляризованного решения;

3) в выражениях x_α и $\widetilde{f}_\alpha$ через r_α фигурирует величина μ, которая при $\delta \longrightarrow 0$ стремится к бесконечности;

4) вектор невязки регуляризованного решения в методе А. Н. Тихонова удовлетворяет независимому уравнению (22);

5) в выражении $\widetilde{f}_\alpha$ через r_α и в уравнении для r_α фигурирует симметричная положительно определенная матрица

$$\mathrm{W} = \mathrm{A}(\mathrm{R^T R})^{-1}\mathrm{A^T}, \tag{89}$$

которая в случае, когда R суть квадратная (M×M)-матрица, может быть записана в форме

$$\mathrm{W} = (\mathrm{AR}^{-1})(\mathrm{AR}^{-1})^\mathrm{T} = \mathrm{BB^T}. \tag{90}$$

Свойства 1) и 2) решений, получаемых в рамках вариационного метода А. Н. Тихонова, свидетельствуют о возможности возникновения в нем эффектов заглаживания; не имея возможности подробно останавливаться на этом моменте, отметим лишь необходимость проведения спектрального анализа, основанного либо на сингулярном разложении матрицы B,

$$\mathrm{B} = \mathrm{AR}^{-1}, \tag{91}$$

либо на сингулярном разложении пары матриц (A, R), см. [*Лоусон, Хенсон,* 1986].

§ **4**. Свойства 4) и 5) требуют более глубокого проникновения в суть конструкций (85)–(87), что в конечном итоге позволит более глубоко осознать вариационную природу нового метода (28)–(32).

Именно, пусть сначала N<M; рассмотрим условную экстремальную задачу

$$\begin{aligned} &\left\| \mathrm{R}x \right\|_\mathrm{E}^2 = \min_x, \\ &\mathrm{A}x = f_\delta. \end{aligned} \tag{92}$$

Как обычно, используем стандартную процедуру редукции условной экстремальной задачи (92) к семейству безусловных экстремальных задач с помощью множителей Лагранжа λ_i, $1 \le i \le N$:

$$\left\| \mathrm{R}x \right\|_\mathrm{E}^2 - 2\sum_{i=1}^{N} \lambda_i \left(\left(a_{(i)}^\mathrm{T}, x \right) - f_{i,\delta} \right) = \min_x, \tag{93}$$

здесь $a_{(i)}$ суть i-ая вектор-строка матрицы A. Используя классическое правило экстремума, найдем, что возникает соотношение для x :

$$\mathrm{R^T R}x = \mathrm{A^T}\lambda, \tag{94}$$

откуда, вместе с линейными условиями в (92), следует, что вектор $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_N)^\mathrm{T}$ удовлетворяет уравнению

$$\mathrm{W}\lambda = \mathrm{A}(\mathrm{R^T R})^{-1}\mathrm{A^T}\lambda = f_\delta. \tag{95}$$

Осуществим далее регуляризацию методом М. М. Лаврентьева уравнения (95) для вектора λ :

$$(\alpha \mathrm{E} + \mathrm{W})\lambda_\alpha = f_\delta. \tag{96}$$

Соответственно введем регуляризованное решение x_α, сделав очевидные замены в (94):

$$\mathrm{R^T R}x_\alpha = \mathrm{A^T}\lambda_\alpha. \tag{97}$$

Ясно далее, что если принять

$$\alpha\lambda_\alpha = r_\alpha, \tag{98}$$

то соотношение (97) перейдет в соотношение (85), а соотношение (96) – в соотношение (87). Ясно также, что условие N<M не является, вообще говоря, необходимым; конструкция, получаемая в рамках условной экстремальной задачи (92), остается в силе и при N≥M.

Из сказанного вытекают три принципиальных вывода:

во-первых, что условная экстремальная задача (92), в которой ищется минимум квадратичного функционала при линейных ограничениях, оказывается существенно отличной от условной экстремальной задачи (16), в которой фигурирует одно квадратичное ограничение-равенство; именно, во втором случае всегда возникает невырожденная система с симметричной положительно определенной матрицей, а в первом случае система хотя и имеет симметричную положительно полуопределенную матрицу, но последняя может быть и вырожденной;

во-вторых, использование “принудительной регуляризации” системы (95) по М. М. Лаврентьеву (которая, как мы видели в предыдущем разделе статьи, см. § 3, представляет собой процедуру, основанную на решении некоторой условной вариационной задачи) с выбором S=E, приводит к пониманию физического смысла множителей Лагранжа λ_i – это суть приближения к величинам $\lambda_{i,\alpha}$, которые просто равны компонентам вектора $\mu r_\alpha = \frac{r_\alpha}{\alpha}$, r_α – вектор невязки для решения x_α, получаемого в рамках вариационного метода А. Н. Тихонова;

в-третьих, очевидно, что обобщение конструкции “принудительной регуляризации” состоит в том, чтобы вместо (96) использовать

$$(a\Re^{\rm T}\Re + {\rm W})\lambda_\alpha = f_\delta, \tag{99}$$

где $\Re$ – заданная (Q×N)-матрица; суть этого обобщения в том, что здесь возникает конструкция регуляризованной системы с двумя матрицами-регуляризаторами – R (которая “скрыта” в матрице W, см. (95)) и $\Re$.

§ 5. Подход, основанный на введении условной экстремальной задачи (92) с линейными ограничениями, с дальнейшей “принудительной регуляризацией” системы (95), заслуживает более глубокого изучения, которое и проводится в настоящем и следующем параграфах.

Первый принципиальный факт состоит в том, что системой (95) (случай N≤M) обеспечивается *мультипликативная регуляризация* системы (2). Смысл этой конструкции раскрывается, если воспользоваться *сингулярным разложением пары матриц* (A,R) (о понятии сингулярного разложения пары матриц [см. *Лоусон, Хенсон,* 1986]):

$$\begin{aligned}{\rm A} &= {\rm V}\overset{\circ}{\Sigma}_{\rm A}{\rm X},\\ {\rm R} &= {\rm U}\Sigma_{\rm R}{\rm X},\end{aligned} \tag{100}$$

где V и U суть ортогональные по столбцам матрицы размеров (N×N)- и (M×M)-, соответственно, а X есть невырожденная (M×M)-матрица; при этом

$$\overset{\circ}{\Sigma}_{\rm A} = \begin{array}{c}\uparrow\\ N\\ \downarrow\end{array}\underbrace{\begin{array}{|c|c|}\hline \widehat{\Sigma}_{\rm A} & \quad 0, \quad \\ \hline\end{array}}_{\longleftarrow N \longrightarrow\longleftarrow M-N \longrightarrow} \ {\rm N}\le{\rm M} \tag{101}$$

и $\widehat{\Sigma}_{\rm A}$, $\Sigma_{\rm R}$ суть диагональные матрицы с элементами $\widehat{\sigma}_i = \widehat{\sigma}_i({\rm A},{\rm R})$, $i = 1,2,\ldots,N$, и $\sigma_i = \sigma_i({\rm A},{\rm R})$, $i = 1,2,\ldots,M$, соответственно.

Очевидно, из (100) следует

$${\rm AR}^{-1} = \left({\rm V}\overset{\circ}{\Sigma}_{\rm A}{\rm X}\right)\cdot\left({\rm X}^{-1}\Sigma_{\rm R}^{-1}{\rm U}^{\rm T}\right) = {\rm V}\overset{\circ}{\Sigma}_{\rm A}\Sigma_{\rm R}^{-1}{\rm U}^{\rm T} \tag{102}$$

и поэтому

$${\rm W} = \left({\rm AR}^{-1}\right)\left({\rm AR}^{-1}\right)^{\rm T} = {\rm V}\overset{\circ}{\Sigma}_{\rm A}\left(\Sigma_{\rm R}^{-1}\right)^2\overset{\circ}{\Sigma}{}_{\rm A}^{\rm T}\,{\rm V}^{\rm T}. \tag{103}$$

Отсюда следует, что система (95) редуцируется к следующей:

$${\rm D}\nu = \psi_\delta, \tag{104}$$

где

$${\rm D} = \overset{\circ}{\Sigma}_{\rm A}\left({\rm D}_{\rm R}^{-1}\right)^2\overset{\circ}{\Sigma}{}_{\rm A}^{\rm T}, \tag{105}$$

и

$$\nu = {\rm V}^{\rm T}\lambda, \quad \psi_\delta = {\rm V}^{\rm T}f_\delta. \tag{106}$$

Ясно, что D есть диагональная (N×N)-матрица, с элементами

$$d_i = \frac{\left(\widehat{\sigma}_i({\rm A},{\rm R})\right)^2}{\left(\sigma_i({\rm A},{\rm R})\right)^2}, \quad 1\le i\le N. \tag{107}$$

Следовательно, поскольку без ограничения общности можно принять, что

$$1 = \sigma_1({\rm A},{\rm R}) > \sigma_2({\rm A},{\rm R}) > \ldots > \sigma_{\rm N}({\rm A},{\rm R}) > 0, \tag{108}$$

то отсюда вытекают неравенства

$$\begin{aligned} &d_i > \widehat{\sigma}_i({\rm A},{\rm R}), \quad {\rm i} = 2,3,\ldots,\\ &d_1 = \widehat{\sigma}_1({\rm A},{\rm R}).\end{aligned} \tag{109}$$

Отсюда следует, что если матрица $\mathrm{A}A^{\rm T}$ невырожденная, то матрица ${\rm B}B^{\rm T} = ({\rm A}R^{-1})({\rm A}R^{-1})^{\rm T}$ лучше обусловлена, чем $\mathrm{A}A^{\rm T}$. Но если $\mathrm{A}A^{\rm T}$ вырожденная, то ${\rm BB}^{\rm T}$ также вырожденная. Улучшение обусловленности ${\rm B}B^{\rm T}$ по сравнению с $\mathrm{A}A^{\rm T}$, достигаемое за счет умножения A на ${\rm R}^{-1}$, и составляет суть мультипликативной регуляризации, которая однако, полностью проблемы не решает – необходима дополнительная “принудительная регуляризация”, достигаемая переходом от (95) к (96).

§ 6. В настоящем параграфе будет рассмотрена важная сторона классического вариационного метода регуляризации – наличие так называемых *двойственных соотношений.*

Заметим, прежде всего, что если R – невырожденная (т.е. обратимая) (M×M)-матрица, то система (22) может быть переписана в такой форме:

$$\left(\alpha \mathrm{E} + \mathrm{B}^{\mathrm{T}}\mathrm{B}\right) \mathrm{R} x_{\alpha} = \mathrm{B}^{\mathrm{T}} f_{\delta}, \tag{110}$$

откуда

$$x_{\alpha} = \mathrm{R}^{-1} \left(\alpha \mathrm{E} + \mathrm{B}^{\mathrm{T}}\mathrm{B}\right)^{-1} \mathrm{B}^{\mathrm{T}} f_{\delta}. \tag{111}$$

Одновременно ясно, что из (85) и (88) следует

$$x_{\alpha} = \mathrm{R}^{-1}\mathrm{B}^{\mathrm{T}} \left(\alpha \mathrm{E} + \mathrm{B}^{\mathrm{T}}\mathrm{B}\right)^{-1} f_{\delta}. \tag{112}$$

Но из линейной алгебры известно соотношение (см. [*Воеводин, Кузнецов,* 1984], стр. 214, упр. 29.9), которое в принятых обозначениях записывается так:

$$\mathrm{B}^{\mathrm{T}} \left(\alpha \mathrm{E} + \mathrm{B}^{\mathrm{T}}\mathrm{B}\right)^{-1} = \left(\alpha \mathrm{E} + \mathrm{B}^{\mathrm{T}}\mathrm{B}\right)^{-1} \mathrm{B}^{\mathrm{T}}. \tag{113}$$

Поэтому получаем, что выражения (111) и (112) дают для вектора x_{α} одно и то же значение. В этом и состоит факт наличия для решений x_{α}, получаемых в рамках классического вариационного метода А. Н. Тихонова, двух внешне различных, но фактически тождественных, представлений – получаемых из решения системы (22) и из соотношений (85) и (87), то есть, как мы убедились выше, из “принудительной регуляризации” системы (95), возникающей при решении условной экстремальной задачи (92) с линейными ограничениями.

§ 7. В заключение данного раздела, в котором получены принципиально новые аналитические результаты, относящиеся к классическому вариационному методу А. Н. Тихонова, заметим следующее. В целой серии работ 70-х–80-х годов А. И. Кобрунов развивал так называемый “критериальный подход” к решению задач гравиметрии и магнитометрии, см. [*Кобрунов,* 1978, 1979, 1982, 1983, 1985]. Суть этого подхода состоит в том, чтобы находить решения любых задач, возникающих в гравиметрии и магнитометрии, как решение следующей условной экстремальной задачи:

$$\begin{aligned} &\mathrm{J}(x) = \min_{x}, \\ &\mathrm{A}x = f_{\delta} = f + \delta f. \end{aligned} \tag{114}$$

Здесь $\mathrm{J}(x)$ – некоторый неотрицательный функционал, определенный на элементах x некоторого банахова (полного линейного нормированного) пространства X, f_{δ} – заданный элемент некоторого другого банахова пространства Y, A – линейный ограниченный оператор, действующий из X в Y. По мнению А. И. Кобрунова, использование постановки (114) в случае, когда $\mathrm{J}(x)$ – строго выпуклый функционал, обеспечивает как единственность, так и устойчивость решения – определения x.

Однако разобранный выше частный случай, в котором $\mathrm{X} = \mathrm{R}^{\mathrm{N}}$, $\mathrm{Y} = \mathrm{R}^{\mathrm{M}}$, $\mathrm{N} \leq \mathrm{M}$, т.е. X и Y суть конечномерные эвклидовы пространства, $\mathrm{J}(x) = \left\|\mathrm{R}x\right\|_{\mathrm{E}}^{2}$ и R – невырожденная матрица, т.е. $\mathrm{J}(x)$ есть неотрицательный строго выпуклый функционал, и A есть (N×M)-матрица, опровергает утверждение А. И. Кобрунова. Именно, в этом случае устойчивость решения не обеспечивается, если матрица AA^{T} является вырожденной, следовательно, вырожденной является и матрица $\mathrm{W} = \left(\mathrm{AR}^{-1}\right)\left(\mathrm{AR}^{-1}\right)^{\mathrm{T}}$. Таким образом, как было показано выше, даже в данном простейшем случае необходимо использовать дополнительную “принудительную регуляризацию”, т.е. переход к системе (96). Система же (95) реализует специфическую “мультипликативную регуляризацию”, которая в случае вырожденных систем недостаточна.

3. Новая теория регуляризации систем линейных алгебраических уравнений с приближенными данными

§ 1. Изложенные в двух предыдущих разделах статьи основные положения и конструкции классической теории регуляризации (нахождения устойчивых приближенных решений, согласованных с имеющейся априорной информацией о помехе во входных данных и искомом решении) систем линейных алгебраических уравнений, с однозначностью свидетельствуют о том, что эта теория неадекватна реальной геофизической практике. Это однозначно следует из таких фактов:

1) в классической теории рассматривается только ситуация чисто аддитивных помех (в векторе правой части и матрицы), тогда как на практике достаточно часто встречаются и ситуации более общих мультипликативно-аддитивных помех;

2) в случае чисто аддитивных помех в задании правой части классическая теория основывается на допущении, что известной является величина $\delta^2 = \left\|\delta f\right\|_{\mathrm{E}}^{2}$ квадрата эвклидовой нормы вектора помехи; фактически же эта величина никогда не бывает известна точно, в лучшем же случае бывают известны лишь постоянные в неравенствах

$$0 < \delta_{\min}^{2} \leq \left\|\delta f\right\|_{\mathrm{E}}^{2} \leq \delta_{\max}^{2} < +\infty \tag{115}$$

и неравенство

$$\delta_{\min}^{2} > \inf_{x \in \mathrm{R}^{\mathrm{M}}} \left\|f_{\delta} - \mathrm{A}x\right\|_{\mathrm{E}}^{2}; \tag{116}$$

3) в подавляющем большинстве реальных геофизических ситуаций при наличии чисто аддитивной помехи в векторе правой части даже знание постоянных в неравенствах (115), если только имеет место ситуация, в которой

$$\theta^2 = \frac{\delta^2_{\min}}{\delta^2_{\max}} < \frac{4}{5} \qquad (117)$$

оказывается недостаточным для нахождения приближенного решения удовлетворительного качества;

4) наиболее эффективные алгоритмы, порожденные классической теорией, оказываются весьма трудоемкими и не могут быть использованы для нахождения устойчивых приближенных решений систем большой размерности;

5) приемы уменьшения трудоемкости основных алгоритмов, порождаемых классической теорией, основаны на использовании ортогональных преобразований, построенных на аннулировании основной части элементов симметричных положительно полуопределенных матриц; однако такие преобразования в случае систем большой размерности "отказывают" из-за влияния ошибок округления.

Понимание всей совокупности указанных фактов с очевидностью свидетельствует о том, что нужна новая теория регуляризации (нахождения устойчивых приближенных решений) систем линейных алгебраических уравнений с приближенными данными, существенно более адекватная реальной геофизической практике, которая порождала бы методы (алгоритмы), более эффективные – по показателям точности и быстродействия – в случае систем большой размерности.

Подобная теория была разработана на протяжении последнего десятилетия В. Н. Страховым, [см. *Страхов,* 1991a, 1991b, 1991c, 1992a, 1992b, 1997d, 1997e, 1997f, 1997g, 1997h; *Страхов, Теперин,* 1991a, 1991b, 1991c, 1993; *Strakhov et al.,* 1995]; в последние годы в разработке этой новой теории и ее практической реализации принял участие и А. В. Страхов [*Страхов, Страхов,* 1998, 1999a, 1999b, 1999c, 1999d]. Изложению основных позиций этой новой теории и посвящены настоящий и следующий разделы статьи.

§ **2**. В классической теории регуляризации исходной системе (1) сопоставляется ее аппроксимация (2) (чисто аддитивная помеха только в правой части) и (62)–(63), (65) (аддитивная помеха в правой части и матрице). В рамках новой теории кроме указанных рассматриваются ситуации задания аппроксимаций

$$\mathrm{A}x = f_\delta = (\mathrm{E} + \mathcal{E}) f_{\delta,\mathcal{E}}, \qquad (118)$$

где фактически задан вектор $f_{\delta,\mathcal{E}}$, и

$$\begin{gathered} \mathrm{A}_{\mathbf{\Delta}} x = f_\delta, \\ \mathrm{A}_{\mathbf{\Delta}} = (\mathrm{E} + \mathbf{D}) \mathrm{A}_{\mathbf{\Delta},\mathbf{D}}, \\ f_\delta = (\mathrm{E} + \mathcal{E}) f_{\delta,\mathcal{E}}, \end{gathered} \qquad (119)$$

где фактически заданы матрица $\mathrm{A}_{\mathbf{\Delta},\mathbf{D}}$ и вектор $f_{\delta,\mathcal{E}}$. Здесь E – единичная (N×N)-матрица, а $\mathcal{E}$ и $\mathbf{D}$ – диагональные матрицы,

$$\begin{aligned} \mathcal{E} &= \operatorname{diag} \varepsilon_i, \quad 1 \le i \le N, \\ \mathbf{D} &= \operatorname{diag} \partial_i, \quad 1 \le i \le N, \end{aligned} \qquad (120)$$

при этом

$$\max_i |\varepsilon_i| \le 1, \quad \max_i |\partial_i| \le 1, \qquad (121)$$

так что матрицы $(\mathrm{E} + \mathcal{E})$ и $(\mathrm{E} + \mathbf{D})$ являются невырожденными.

Ниже будет показано, что случаи аппроксимаций (118), (120) и (119), (120) при условиях (121) в определенном смысле редуцируются к случаю аппроксимации (2), т.е. наличию чисто аддитивной помехи в задании правой части системы (1). Поэтому изложение основ новой теории целесообразно начать с рассмотрения основного частного случая задания аппроксимации (2) системы линейных уравнений (1).

§ **3**. Основой новой теории в случае чисто аддитивной помехи в векторе правой части системы являются следующие четыре положения.

1. Основная вычислительная процедура в любом методе – алгоритме – состоит в нахождении последовательности приближенных решений $\widetilde{x}^{(k)}$, $k = 1, 2, \ldots$, и соответствующих им векторов невязок

$$\rho^{(k)} = f_\delta - \mathrm{A}\widetilde{x}^{(k)} \qquad (122)$$

и величин

$$\nu_k = \left\| \rho^{(k)} \right\|^2_{\mathrm{E}}. \qquad (123)$$

При этом каждое приближение $\widetilde{x}^{(k)}$ характеризуется некоторым набором (вектором) *параметров счета*, то есть величин, по которым находится приближенное решение; обычно размерность вектора *параметров счета* – от одного до трех. Желательно (хотя это и не всегда может быть гарантировано), монотонное убывание величин ν_k :

$$\nu_{k+1} < \nu_k, \quad k = 1, 2, \ldots . \qquad (124)$$

Следует подчеркнуть, что нахождение последовательности приближенных решений $\widetilde{x}^{(K)}$ отнюдь не означает, что используется какой-то итерационный процесс. Например, речь может идти о нахождении решений уравнения

$$\left(\alpha_k \mathrm{R}^{\mathrm{T}}\mathrm{R} + \mathrm{A}^{\mathrm{T}}\mathrm{A}\right) x_{\alpha_k} = \mathrm{A}^{\mathrm{T}} f_\delta = \varphi_\delta \tag{125}$$

для последовательности значений α_k, $k = 1, 2, \ldots$; в этом случае $\widetilde{x}^{(k)} = x_{\alpha_k}$ и α_k суть параметр счета (единственный).

2. В каждом методе (алгоритме) присутствует некоторый вектор *показателей качества решения*; в этом векторе выделяются *основные показатели* качества решения (обычно – от одного до трех), и *дополнительные* показатели. С помощью основных показателей качества решений выделяются, среди всех приближенных решений $\widetilde{x}^{(k)}$, так называемые *пробные решения*, с помощью же дополнительных показателей качества решений множество пробных решений сокращается до множества *допустимых решений.*

Ясно, что с помощью основных показателей качества решения формируется и *правило останова* процесса нахождения приближенных решений. Ясно также, что в процессе счета должны запоминаться все пробные решения – и далее, после сужения множества пробных решений до множества допустимых решений, все допустимые решения.

В ситуации, когда известны постоянные в неравенствах (115), используется обычно *один показатель качества* – величина

$$\widehat{\nu}_k = \frac{\left\|\rho^{(k)}\right\|_{\mathrm{E}}^2}{\delta_{\min}^2}; \tag{126}$$

при этом пробными являются все те приближенные решения, для которых

$$1 \le \widehat{\nu}_k \le \nu_{\max} = \frac{\delta_{\max}^2}{\delta_{\min}^2} > 1. \tag{127}$$

Процесс построения приближенных решений останавливается, коль скоро оказывается

$$\widehat{\nu}_k < 1. \tag{128}$$

Кроме основного показателя качества (126) могут использоваться и другие основные показатели, например:

$$\gamma_k = \frac{\left\|\widetilde{x}^{(k)}\right\|_{\mathrm{E}}^2}{\left\|\widetilde{x}^{(k_{\min})}\right\|_{\mathrm{E}}^2}. \tag{129}$$

Здесь $\widetilde{x}^{(k_{\min})}$ – пробное решение с наименьшим значением k; при этом используется правило останова процесса построения последовательных приближений

$$\gamma_k > Crit, \tag{130}$$

где Crit – априорно заданная величина.

Что касается дополнительных показателей качества, то в рассматриваемом случае это могут быть различного рода числовые величины, порождаемые априорной информацией о векторе помехи δf.

Принципиально важно, что если имеет место ситуация, в которой выполняется неравенство (117), то именно последняя информация должна использоваться для задания основных показателей качества решения. Эта ситуация подробнее рассматривается ниже. В ней по существу нахождение множеств пробных и допустимых решений осуществляется с помощью эвристических процедур типа распознавания образов.

3. После того, как найдено множество допустимых приближенных решений системы (2), возникает проблема нахождения искомого – или окончательного, или оптимального – устойчивого приближенного решения. Эта проблема в рамках новой теории решается на основе принципа усреднения допустимых решений:

$$\widetilde{x} = \sum_{k \in \mathrm{J}_k} p_k \widetilde{x}^{(k)}, \tag{131}$$

где весовые коэффициенты p_k имеют свойства

$$p_k > 0, \quad k \in \mathrm{J}_k, \quad \sum_{k \in \mathrm{J}_k} p_k = 1. \tag{132}$$

Здесь J_k – подмножество значений индекса k для допустимых решений.

Ясно, что здесь возникает новая проблема – задания весовых множителей p_k. Эта проблема детально обсуждается ниже, но сразу же подчеркнем тот факт, что и здесь основное значение имеют различного рода эвристические процедуры, и в конечном итоге – технологии распознавания образов.

4. Все три предыдущих положения относились к так называемой *нормальной* ситуации, в которой:

а) в последовательности приближенных решений $\widetilde{x}^{(k)}$ всегда содержатся пробные приближенные решения, а останов процедуры построения приближенных решений осуществляется по значениям основных показателей качества приближенных решений;

б) число находимых допустимых решений, усреднением которых, см. (131)–(132), находится окончательное решение, достаточно велико:

$$\dim \mathrm{J}_k \ge \mathrm{R}_0, \tag{133}$$

где R_0 – заданная величина.

Однако на практике, как это показали достаточно многочисленные расчеты (описание которых будет дано в четвертой части работы), кроме нормальной, может встречаться и *аномальная ситуация*, в которой:

а) либо число найденных допустимых решений меньше R_0 (см. (133));

б) либо вообще не находятся пробные (тем более – допустимые) приближенные решения.

В указанных ситуациях алгоритмы должны обеспечивать:
– в случае а) (число допустимых решений меньше R_0) нахождение дополнительных допустимых решений;
– в случае б) (отсутствие пробных или допустимых решений) нахождение некоторого – единственного! – приближенного решения, которое и принимается за окончательное.

Здесь следует сразу же акцентировать внимание читателя на том обстоятельстве, что в рамках классической теории некорректно поставленных задач (по существу не адекватной реальной геофизической практике) никакие *аномальные ситуации* (обусловленные погрешностями округлений) по существу не принимались во внимание и не обсуждались.

В заключение параграфа подчеркнем, что практическая реализация установок новой теории требует создания принципиально новых компьютерных технологий, в которых в некотором смысле определяющее значение имеют *управляющие* подпрограммы, с помощью которых формируются последовательности векторов параметров счета – для нахождения последовательных приближений, а также осуществляется анализ получаемых приближенных решений – по значениям показателей качества решения, в целях, по существу, ответа на вопрос – "что такое хорошо, и что такое плохо". Эта управляющая программа одновременно должна обеспечивать перестройку вычислительного процесса в случае аномальных ситуаций.

§ **4**. В рамках новой теории регуляризации систем линейных алгебраических уравнений, ориентированной на создание компьютерных технологий, полностью адекватных реальной геофизической практике, и на решений систем линейных алгебраических уравнений большой и сверхбольшой размерности, используется ряд принципиально новых математических объектов и конструктивных идей. Обсуждению конструктивных идей посвящен последний раздел настоящей статьи и вторая часть работы, поэтому здесь остановимся лишь на некоторых новых математических объектах.

Первая группа объектов связано с оценками предельной относительной погрешности любых приближенных решений систем (1), заданных аппроксимациями (2), при использовании представлений

$$\widetilde{x} = \Re f_\delta; \tag{134}$$

здесь $\Re$ суть (M×N)-матрица, заданная либо определяемая некоторыми условиями.

Оценочный функционал в случае аппроксимации (2) имеет вид

$$w(\mathrm{A}, \delta; \Re) = \sup_{\substack{\mathrm{A}x = f, \\ \|\delta f\|_{\mathrm{E}} \leq \delta}} \frac{\|x - \Re f_\delta\|_{\mathrm{E}}}{\|x\|_{\mathrm{E}}}. \tag{135}$$

Ясно, что к числу новых объектов первой группы принадлежит и сама матрица $\Re$, и функционал $w(\mathrm{A}, \delta; \Re)$, и другие подобного рода функционалы, и следующие из оценок предельной относительной погрешности приближенных решений новые постановки безусловных и условных экстремальных задач, и многие другие объекты.

Вторая группа объектов связана с идеей более глубокого понимания метода наименьших квадратов и соответственно – перенесения ряда важных свойств решений, получаемых в методе наименьших квадратов, на регуляризованные решения.

Здесь прежде всего следует отметить характеристическое уравнение метода наименьших квадратов (26) и тот факт, что в вариационном методе А. Н. Тихонова оно не выполняется. Но легко показать, что если $\widetilde{x}$ есть некоторое приближенное решение, для которого

$$(\mathrm{A}\widetilde{x}, f_\delta - \mathrm{A}\widetilde{x}) \neq 0, \tag{136}$$

то приближенное решение

$$\widetilde{x}_\tau = \tau \widetilde{x}, \tag{137}$$

в котором принято

$$\tau = \frac{(f_\delta, \mathrm{A}\widetilde{x})}{\left\|\mathrm{A}\widetilde{x}\right\|_{\mathrm{E}}^2}, \tag{138}$$

уже удовлетворяет характеристическому уравнению метода наименьших квадратов

$$(\mathrm{A}\widetilde{x}_\tau, f_\delta - \mathrm{A}\widetilde{x}_\tau) = 0. \tag{139}$$

При этом

$$\left\|f_\delta - \mathrm{A}\widetilde{x}_\tau\right\| \leq \left\|f_\delta - \mathrm{A}\widetilde{x}\right\|_{\mathrm{E}}^2. \tag{140}$$

Все эти факты легко получаются, если рассмотреть задачу

$$\left\|f_\delta - \tau\mathrm{A}\widetilde{x}\right\|_{\mathrm{E}}^2 = \min_\tau. \tag{141}$$

Ясно, что соотношение

$$\min_\tau \left\|f_\delta - \tau\mathrm{A}\widetilde{x}\right\|_{\mathrm{E}}^2 = \delta^2, \tag{142}$$

где δ^2 – априорно заданная величина, эквивалентно соотношению

$$\left(\left\|f_\delta\right\|_{\mathrm{E}}^2-\delta^2\right)\left\|\mathrm{A}\widetilde{x}\right\|_{\mathrm{E}}^2-\left(f_\delta,\mathrm{A}\widetilde{x}\right)^2=0. \tag{143}$$

Выполнение же характеристического уравнения метода наименьших квадратов имеет принципиальное значение в тех (весьма часто встречающихся на практике) случаях, когда априорно известно, что полезный сигнал f и помеха δf ортогональны:

$$(f,\delta f)=0. \tag{144}$$

Далее. Из основного уравнения в методе наименьших квадратов следует, что (x_0 – решение системы (23)):

$$\mathrm{A}^{\mathrm{T}}\left(f_\delta-\mathrm{A}x_0\right)=0. \tag{145}$$

Отсюда, как нетрудно найти, следует, что имеют место равенства

$$\left(a^{(i)},f_\delta-\mathrm{A}x_0\right)=0,\quad i=1,2,\ldots,\mathrm{M}, \tag{146}$$

где $a^{(i)}$ суть векторы-столбцы матрицы A, т.е. что вектор невязки для решения, найденного по методу наименьших квадратов, ортогонален ко всем столбцам матрицы A. Далее естественно ввести и использовать квадратичный функционал

$$\begin{aligned}\theta(x)&=\frac{1}{\mathrm{M}}\left\|\mathrm{D_A}\mathrm{A}^{\mathrm{T}}\left(f_\delta-\mathrm{A}x\right)\right\|_{\mathrm{E}}^2\\&=\frac{1}{\mathrm{M}}\sum_{i=1}^{\mathrm{M}}\frac{\left(a^{(i)},\left(f_\delta-\mathrm{A}x\right)\right)^2}{\left\|a^{(i)}\right\|_{\mathrm{E}}^2},\end{aligned} \tag{147}$$

где

$$\mathrm{D_A}=\mathrm{diag}\,\frac{1}{\left\|a^{(i)}\right\|_{\mathrm{E}}}, \tag{148}$$

для которого имеет место оценка

$$\theta(x)\le\left\|f_\delta-\mathrm{A}x\right\|_{\mathrm{E}}^2 \tag{149}$$

и который при $x=x_0$ (решению, получаемому по методу наименьших квадратов) обращается в ноль.

Третья группа объектов, культивируемых в новой теории регуляризации систем линейных алгебраических уравнений, это новые ортогональные преобразования. Точнее, речь идет об ортогональных преобразованиях вида

$$\mathrm{B}=\mathrm{V}^{\mathrm{T}}\mathrm{A}U, \tag{150}$$

в которых V и U суть ортогональные по столбцам матрицы размеров (N×N) и (M×M), соответственно, и которые представимы в форме произведений элементарных матриц плоского вращения, либо вида

$$\begin{gathered}\mathrm{A}=\mathrm{A}_0,\\ \mathrm{A}_k=\mathrm{V}_k^{\mathrm{T}}\mathrm{A}_{k-1}\mathrm{U}_k,\quad k=1,2,\ldots,\end{gathered} \tag{151}$$

в которых V_k и U_k суть ортогональные по столбцам матрицы указанного выше вида. Так вот, речь идет о том, что параметры ($\cos\varphi$ и $\sin\varphi$, где φ имеет соответствующие индексы) матриц элементарных плоских вращений выбираются совсем по-иному, чем в широко используемых в настоящее время процедурах. В последних (приведение к треугольной или двухдиагональной формам, к трехдиагональной или диагональной формам) параметры вращения выбираются из условия аннулирования некоторого элемента матрицы. В новых процедурах, разработанных В. Н. Страховым и А. В. Страховым (см. [*Страхов,* 1997d, 1997e, 1998d; *Страхов, Страхов,* 1999b, 1999c, 1999d; *Страхов, Тетерин,* 1993; *Strakhov et al.,* 1995]), параметры преобразований выбираются из иных условий.

Наиболее важными новыми ортогональными преобразованиями являются:

1) процедура выметания одного столбца в другой;

2) процедура выметания одной строки в другую;

3) процедура раскорреляции данного столбца с данным вектором – за счет преобразования данного столбца с некоторым третьим (опорным);

4) процедура максимизации коэффициента корреляции некоторого (опорного) столбца с заданным – за счет преобразования этого столбца с некоторым третьим,

и ряд других, которые здесь обсуждать не будем, ибо эта тема будет подробно рассмотрена во второй части работы.

Значимость для вычислительной алгебры в целом, а не только для новой теории регуляризации систем линейных алгебраических уравнений, аппарата ортогональных преобразований видна из следующего примера. Пусть ортогональное преобразование (150) при V≡E осуществлено таким образом, что выполняются условия

$$\left(b^{(i)},f_\delta\right)=0,\quad i=2,3,\ldots,\mathrm{M} \tag{152}$$

и

$$\left(b^{(1)},b^{(i)}\right)=0,\quad i=2,3,\ldots,\mathrm{M}, \tag{153}$$

где $b^{(j)}$ – векторы-столбцы матрицы B. В этом случае получаем (для системы (2))

$$\mathrm{C}_z=\mathrm{B}^{\mathrm{T}}\mathrm{B}^z=\left(b^{(1)},f_\delta\right)e,\quad z=\mathrm{U}^{\mathrm{T}}x, \tag{154}$$

где

$$e = \left|\begin{array}{c|c} 1 & \overset{\uparrow}{\underset{\downarrow}{1}} \\ \hline 0 & \mathrm{M} \overset{\uparrow}{\underset{\downarrow}{-}} 1 \end{array}\right. , \qquad (155)$$

и матрица $\mathrm{C} = \mathrm{B}^{\mathrm{T}}\mathrm{B}$ имеет следующую структуру

$$\mathrm{C} = \begin{array}{|c|c|} \hline c_{11} & 0 \\ \hline 0 & \Delta\mathrm{C} \\ \hline \end{array} . \qquad (156)$$

Ясно, что решение системы (154) редуцируется к одному уравнению с одним неизвестным

$$c_{11} z_1 = \left(b^{(1)}, f_\delta\right), \qquad (157)$$

которое элементарным образом регуляризуется

$$\left(c_{11} + \frac{\alpha}{c_{11}}\right) z_{1,\alpha} = \left(b^{(1)}, f_\delta\right), \qquad (158)$$

где $\alpha > 0$ – параметр. Далее используются соотношения

$$z_\alpha = \left|\begin{array}{c|c} z_{1,\alpha} & \overset{\uparrow}{\underset{\downarrow}{1}} \\ \hline 0 & \mathrm{M} \overset{\uparrow}{\underset{\downarrow}{-}} 1 \end{array}\right. , \quad x_\alpha = \mathrm{U} z. \qquad (159)$$

Таким образом, использование подходящих ортогональных преобразований позволяет получить весьма эффективный по быстродействию алгоритм, ибо *считать матрицу* $\Delta\mathrm{C}$ *вовсе не нужно!*

Далее необходимо указать, что в построении целого ряда новых ортогональных преобразований матриц центральную роль играет сформулированная В. Н. Страховым *концепция выметания* матриц [*Страхов,* 1997e]. Ее суть в следующем: пусть в множестве всех пар индексов (i,j), $1 \leq i \leq \mathrm{N}$, $1 \leq j \leq \mathrm{M}$; выделено некоторое подмножество J пар индексов (i,j), которое является односвязным [5]; через CJ обозначается подмножество пар индексов (i,j), получаемое из полного множества удалением подмножества J (ясно, что CJ уже может не быть связным множеством).

Рассматриваются две задачи:

первая – найти ортогональные по столбцам матрицы V и U, размеров (N×N) и (M×M), соответственно, из условия

$$\left\|\mathrm{V}^{\mathrm{T}}\mathrm{A}U\right\|^2_{\mathrm{E(J)}} = \max_{\substack{\mathrm{V} \in \mathrm{O_N},\\ \mathrm{U} \in \mathrm{O_M}}}, \qquad (160)$$

где $\mathrm{O_N}$ и $\mathrm{O_M}$ – множества всех ортогональных по столбцам матриц размеров (N×N) и (M×M), соответственно, $\left\|\cdot\right\|^2_{\mathrm{E(J)}}$ – эвклидова норма матрицы, определенная по элементам, имеющим пары индексов (i,j), принадлежащих подмножеству J;

вторая – найти те же самые ортогональные по столбцам матрицы V и U из условия

$$\left\|\mathrm{V}^{\mathrm{T}}\mathrm{A}U\right\|^2_{\mathrm{E(CJ)}} = \min_{\substack{\mathrm{V} \in \mathrm{O_N},\\ \mathrm{U} \in \mathrm{O_M}}}. \qquad (161)$$

Поскольку, как хорошо известно [см. *Воеводин, Кузнецов,* 1984; *Марчук, Кузнецов,* 1972; *Уилкинсон, Райнш,* 1976; *Фаддеев, Фаддеева,* 1963]);

$$\left\|\mathrm{V}^{\mathrm{T}}\mathrm{A}U\right\|^2_{\mathrm{E}} = \left\|\mathrm{A}\right\|^2_{\mathrm{E}}, \qquad (162)$$

то задачи (160) и (161) эквивалентны; ими определяются такие ортогональные преобразования, которые осуществляют перераспределение эвклидовой нормы матрицы A или *выметание A в J.* Очевидно, наиболее важным является тот случай, когда

$$\left\|\mathrm{V}^{\mathrm{T}}\mathrm{A}U\right\|^2_{\mathrm{E}(J)} = \left\|\mathrm{V}^{\mathrm{T}}\mathrm{A}U\right\|^2_{\mathrm{E}} = \left\|\mathrm{A}\right\|^2_{\mathrm{E}}, \qquad (163)$$

т.е. когда осуществляется *полное выметание* матрицы A на подмножество J, так что если $\mathrm{V}^{\mathrm{T}}\mathrm{AU} = \mathrm{B}$ и $b_{i,j}$ суть элементы B, то

$$b_{i,j} = 0, (i,j) \in \mathrm{C}J. \qquad (164)$$

Ясно, что всегда можно выбирать матрицы V и U в форме

$$\mathrm{V} = \prod_{\mathrm{p}} \mathrm{T}_{i_p,j_p}(\varphi_p), \quad 1 \leq i_p \leq j_p \leq \mathrm{N}, \qquad (165)$$

$$\mathrm{U} = \prod_{\mathrm{q}} \mathrm{T}_{i_q,j_q}(\psi_q), \quad 1 \leq i_q \leq j_q \leq \mathrm{M}, \qquad (166)$$

где $\mathrm{T}_{i_p,j_p}(\varphi_p)$ и $\mathrm{T}_{i_q,j_q}(\psi_q)$ суть матрицы элементарных плоских вращений; при этом параметры вращения у этих матриц определяются в соответствии с основной постановкой – (160) или (161).

Во второй части работы будет подробно рассмотрен ряд конкретных ортогональных преобразований подобного типа.

[5] Множество пар индексов (i,j) можно изображать графически в виде совокупности узлов целочисленной решетки на плоскости. Ясно, что подмножество узлов этой решетки односвязно, если в нем нет изолированных узлов, т.е. таких, что все узлы решетки, смежные к заданному, данному подмножеству не принадлежат.

Четвертая группа объектов – итерационные процессы нового типа, порождаемые последовательностями ортогональных преобразований матриц. Здесь существует большое разнообразие возможностей, которые будут подробно рассмотрены во второй части работы.

Наконец, *пятая группа объектов* – это различного рода функционалы, определяемые:

а) на векторах невязок $\rho^{(k)} = f_\delta - \mathrm{A}\widetilde{x}^{(k)}$;

б) на приближенных решениях $\widetilde{x}^{(k)}$.

§ 5. Обсудим теперь более подробно один из принципиальных элементов новой теории – оценки предельной относительной погрешности приближенных решений $\widetilde{x}$, получаемых по соотношению (134), в котором $\Re$ есть *некоторая заданная матрица* размера (M×N) (в общем случае N≠M). Предполагается при этом, что точное решение x системы (1) существует.

Элементарные выкладки, в которых сначала используется схема Лаврентьева–Джона (6), с заменой $\Re_\delta$ на $\Re$, и далее очевидные неравенства

$$\left\|(\mathrm{E}-\Re\mathrm{A})x\right\|_{\mathrm{E}} \le \left\|\mathrm{E}-\Re\mathrm{A}\right\|\cdot\left\|x\right\|_{\mathrm{E}},$$

$$\frac{\delta}{\|x\|} \le \frac{\|\mathrm{A}\|\delta}{\|f\|_{\mathrm{E}}} \le \frac{\|\mathrm{A}\|\delta}{\|f_\delta\|_{\mathrm{E}}-\delta} = \frac{\|\mathrm{A}\|\eta}{1-\eta}, \tag{167}$$

$$\eta = \frac{\delta}{\|f_\delta\|_{\mathrm{E}}},$$

приводят к следующему результату

$$\begin{gathered} w(\mathrm{A},\delta;\Re) \le \left\|E-\Re\mathrm{A}\right\| + \\ + \nu\|\Re\| \le \sqrt{2}\Big(\left\|E-\Re\mathrm{A}\right\|^2 + \nu^2\|\Re\|^2\Big)^{1/2} \\ \le \sqrt{2}\Big(\left\|E-\Re\mathrm{A}\right\|_{\mathrm{E}}^2 + \nu^2\left\|\Re\right\|_{\mathrm{E}}^2\Big)^{1/2}, \end{gathered} \tag{168}$$

где

$$\nu = \frac{\|\mathrm{A}\|\eta}{1-\eta} \le \nu^* = \frac{\|\mathrm{A}\|_{\mathrm{E}}\eta}{1-\eta}. \tag{169}$$

При этом в (167) и (168) $\|\cdot\|$ суть спектральные нормы матриц, а $\|\cdot\|_{\mathrm{E}}$ – эвклидовы нормы матриц.

Далее – пусть известна лишь аппроксимация (62) уравнения (1), т.е. с погрешностью заданы и правая часть, и матрица системы. В этом случае вводится оценочный функционал

$$\omega(\mathrm{A},\delta,\Delta;\Re) = \sup_{\substack{\mathrm{A}x=f,\\ \|\delta f\|_{\mathrm{E}}\le\delta,\ \|\Delta\mathrm{A}\|_{\mathrm{E}}\le\Delta}} \left(\frac{\|x-\Re f_\delta\|_{\mathrm{E}}}{\|x\|_{\mathrm{E}}}\right). \tag{170}$$

Очевидно, имеем

$$\begin{aligned} \omega(\mathrm{A},\delta,\Delta;\Re) &\le \left\|\mathrm{E}-\Re\mathrm{A}\right\| + \nu\|\Re\| \le \\ &\le \left\|\mathrm{E}-\Re\left(\mathrm{A}_\Delta-\Delta\mathrm{A}\right)\right\| + \nu'\|\Re\| \\ &= \left\|\mathrm{E}-\Re\mathrm{A}_\Delta\right\| + \left(\Delta+\nu'\right)\|\Re\| \\ &= \left\|\mathrm{E}-\Re\mathrm{A}_\Delta\right\| + \nu^*\|\Re\| \\ &\le \sqrt{2}\left(\left\|\mathrm{E}-\Re\mathrm{A}_\Delta\right\|^2 + \left(\nu^*\right)^2\|\Re\|^2\right)^{1/2} \\ &\le \sqrt{2}\left(\left\|\mathrm{E}-\Re\mathrm{A}_\Delta\right\|^2 + \left(\widehat{\nu}^*\right)^2\|\Re\|_{\mathrm{E}}^2\right)^{1/2}, \end{aligned} \tag{171}$$

где положено

$$\begin{gathered} \nu' = \frac{\|\mathrm{A}\|\eta + (1-\eta)}{1-\eta}, \\ \nu^* = \frac{\|\mathrm{A}_\Delta\|\delta+\Delta}{1-\eta} \le \frac{\|\mathrm{A}_\Delta\|_{\mathrm{E}}\,\delta+\Delta}{1-\eta} = \widehat{\nu}^*. \end{gathered} \tag{172}$$

Из неравенства (171) следует, что в случае задания аппроксимации (62) системы (1) можно принять, что матрица A_Δ есть точная матрица, но правая часть f_δ задана с помехой δf большей нормы, а именно

$$\|\delta f\|_{\mathrm{E}} = \delta', \tag{173}$$

где

$$\delta' = \eta'\|f_\delta\|_{\mathrm{E}}, \quad \eta' = \frac{\eta+\chi}{1+\chi} > \eta \tag{174}$$

и

$$\chi = \frac{\Delta}{\|\mathrm{A}_\Delta\|}. \tag{175}$$

Изложенная схема редукции задач нахождения устойчивых приближенных решений систем линейных алгебраических уравнений с приближенно заданными правыми частями и матрицами (помехи чисто аддитивные) к случаю задания аддитивной помехи только в правой части системы безусловно является принципиальной частью новой теории регуляризации систем линейных алгебраических уравнений.

Отметим здесь еще одно важное следствие оценки (171). Ясно, что если матрица A – квадратная, N=M, но вырожденная или очень плохо обусловленная, то тогда разумно перейти от матрицы A к матрице $\mathrm{A}_\Delta = \mathrm{A} + \Delta\mathrm{A}$ с помощью искусственно введенного возмущения $\Delta\mathrm{A}$, введенного таким образом, чтобы матрица $\Delta\mathrm{A}$ была уже разумным образом обусловлена. В таком случае, очевидно, разумно выбрать

$$\Re = \mathrm{A}_\Delta^{-1}; \tag{176}$$

в этом случае оценка предельной относительной погрешности примет вид

$$\omega(\mathrm{A},\delta,\Delta;\Re) \le \mu \operatorname{Cond} A_\Delta, \tag{177}$$

где

$$\operatorname{Cond} A_\Delta = \|\mathrm{A}_\Delta\|\cdot\|\mathrm{A}_\Delta^{-1}\|, \tag{178}$$

и

$$\mu=\frac{\eta+\chi}{1-\eta},\quad \chi=\frac{\Delta}{\left\|\mathrm{A}_{\Delta}\right\|}. \tag{179}$$

Очевидно, в оценке (177)-(179) могут фигурировать либо спектральные, либо эвклидовы норма матриц A_{Δ} и A_{Δ}^{-1}. Ясно, из общих соображений (значение которых будет подробно рассмотрено во второй части работы), что целесообразно ввести следующее условие на вариацию $\Delta\mathrm{A}$ матрицы A (в тех случаях, когда вариация осуществляется в целях регуляризации):

$$\left\|\mathrm{A}_{\Delta}\right\|=\left\|\mathrm{A}\right\|; \tag{180}$$

при этом наиболее важным практически будет случай, когда в (178), (180) фигурирует эвклидова норма матриц.

Далее ясно, что подобного рода оценки могут быть (и даже обязательно должны быть) использованы и в том случае, когда рассматриваются системы линейных уравнений

$$\mathrm{C}x=\varphi_{\delta},\quad \mathrm{C}=\mathrm{A}^{\mathrm{T}}\mathrm{A},\quad \varphi_{\delta}=\mathrm{A}^{\mathrm{T}}f_{\delta}, \tag{181}$$

и

$$\mathrm{B}z=f_{\delta},\quad \mathrm{B}=\mathrm{A}R^{-1},\quad z=\mathrm{R}x, \tag{182}$$

и т.д. Расписывание соответствующих оценок предоставляется читателю. При этом ясно, что случай систем (181), равно как и систем

$$\overset{\mathrm{O}}{\mathrm{C}}z=f_{\delta},\quad \overset{\mathrm{O}}{\mathrm{C}}=\mathrm{A}A^{\mathrm{T}},\quad x=\mathrm{A}^{\mathrm{T}}z, \tag{183}$$

является с точки зрения практики особенно важным, ибо матрицы C и $\overset{\mathrm{O}}{\mathrm{C}}$ – квадратные, размеров (M×M) и (N×N), соответственно, и $\mathrm{C}=\mathrm{C}^{\mathrm{T}}$, $\overset{\mathrm{O}}{\mathrm{C}}=\overset{\mathrm{O}}{\mathrm{C}}{}^{\mathrm{T}}\geq 0$.

Ясно одновременно, что случай систем (181) вписывается в изложенную выше схему, если просто принять в (134)

$$\Re=\overset{\mathrm{O}}{\Re}\mathrm{A}^{\mathrm{T}}, \tag{184}$$

а случай систем (183) – если принять

$$\Re=\mathrm{A}^{\mathrm{T}}\widetilde{\Re}; \tag{185}$$

в (185) $\overset{\mathrm{O}}{\Re}$ есть (M×M)-матрица, а в (134) $\widetilde{\Re}$ есть (N×N)-матрица.

Забегая вперед (подробно этот вопрос будет рассматриваться во второй части работы), отметим также, что полученные оценки предельной относительной погрешности позволяют ввести в рассмотрение новые постановки безусловных и условных экстремальных задач, например

$$\left\|\mathrm{E}-\Re\mathrm{A}\right\|_{\mathrm{E}}^{2}+\left(\nu^{*}\right)^{2}\left\|\Re\right\|_{\mathrm{E}}^{2}=\min_{\Re}, \tag{186}$$

или

$$\left\|\mathrm{A}_{\Delta}^{-1}\right\|_{\mathrm{E}}^{2}-\min,\ \left\|\mathrm{A}_{\Delta}\right\|_{\mathrm{E}}=\left\|\mathrm{A}\right\|_{\mathrm{E}},\ \left\|\Delta\mathrm{A}\right\|_{\mathrm{E}}^{2}=\Delta^{2}, \tag{187}$$

где Δ^2 – априорно заданное число (см. соотношения (62), (65), и (177)–(179)).

Введение новых постановок безусловных и условных экстремальных задач – одна из отличительных черт новой теории регуляризации систем линейных алгебраических уравнений с приближенными данными.

§ 6. Переходим к рассмотрению еще одной принципиальной стороны новой теории регуляризации систем линейных алгебраических уравнений – постановок задач построения регуляризованных алгоритмов в ситуации мультипликативно-аддитивных помех, в простейшем случае – в задании правой части системы, в более общем – в задании как правой части, так и матрицы системы, см. (118), (120)–(121); в этой связи см. работы [*Страхов,* 1991d, 1991e, 1991f, 1991g, 1991h, 1991i].

Начнем со случая задания аппроксимации (118), (120)–(121) системы (1). Наша основная цель состоит в том, чтобы установить факт принципиальной редукции задачи к нахождению устойчивого приближенного решения системы типа (2), т.е. к случаю чисто аддитивной помехи.

Пусть, как сказано выше, ε_i – суть компоненты диагональной матрицы $\mathcal{E}$. Рассмотрим задачу

$$\gamma\left\|\mathcal{E}\right\|_{\mathrm{E}}^{2}+\left\|\left(\mathrm{E}+\mathcal{E}\right)f_{\delta,\varepsilon}-\mathrm{A}x\right\|_{\mathrm{E}}^{2}=\gamma\sum_{i=1}^{\mathrm{N}}\varepsilon_i^2+ \tag{188}$$
$$+\sum_{i=1}^{\mathrm{N}}\left(\left(a_{(i)}^{\mathrm{T}},x\right)-\left(1+\varepsilon_i\right)f_{i,\delta,\varepsilon}\right)^{2}=\min_{\varepsilon_i},$$

где $\gamma>0$ – параметр (который, вообще говоря, не имеет смысла параметра регуляризации), а $a_{(i)}^{\mathrm{T}}$ – i-ая вектор-строка матрицы A. Выполняя элементарные дифференцирования по ε_i и приравнивая производные нулю, с очевидностью получаем следующие соотношения:

$$\varepsilon_i=\frac{\left(\left(a_{(i)}^{\mathrm{T}},x\right)-f_{i,\delta,\varepsilon}\right)f_{i,\delta,\varepsilon}}{\gamma+f_{i,\delta,\varepsilon}^{2}}, \tag{189}$$

и при этих значениях ε_i

$$\sum_{i=1}^{\mathrm{N}}\left(\left(a_{(i)}^{\mathrm{T}},x\right)-\left(1+\varepsilon_i\right)f_{i,\delta,\varepsilon}\right)^{2} \tag{190}$$
$$=\left\|\mathrm{P}_{\gamma}\left(f_{\delta,\varepsilon}-\mathrm{A}x\right)\right\|_{\mathrm{E}}^{2},$$

$$\sum_{i=1}^{\mathrm{N}} \varepsilon_i^2 = \left\|\mathcal{E}\right\|_{\mathrm{E}}^2 = \left\|\mathrm{Q}_\gamma\left(f_{\delta,\varepsilon} - \mathrm{A}x\right)\right\|_{\mathrm{E}}^2, \qquad (191)$$

где P_γ и Q_γ суть диагональные матрицы с элементами $p_i(\gamma)$ и $q_i(\gamma)$, соответственно равными:

$$p_i(\gamma) = \frac{\gamma}{\gamma + f_{i,\delta,\varepsilon}^2}, \quad i = 1, 2, \ldots, \mathrm{N}, \qquad (192)$$

$$q_i(\gamma) = \frac{|f_{i,\delta}|}{\gamma + f_{i,\delta,\varepsilon}^2}, \quad i = 1, 2, \ldots, \mathrm{N}. \qquad (193)$$

Из (189)–(191) и (192), (193) следует, что если задана аппроксимация (118) системы (1), то фактически имеет место редукция к случаю задания аппроксимации (2), но с одним важным изменением. Именно, и в случае (118) для нахождения устойчивых приближенных решений системы (1) можно использовать алгоритмы, разработанные для аппроксимации (2), но при этом:

а) необходимо иметь априорную информацию как о векторе помехи δf, так и о диагональной матрице помехи $\mathcal{E}$;

б) в числе основных показателей качества приближенных решений $\widetilde{x}^{(k)}$ должны использоваться величины:

$$\left\|\rho^{(k)}\right\|_{\mathrm{E}}^2 = \left\|f_{\delta,\varepsilon} - \mathrm{A}\widetilde{x}^{(k)}\right\|_{\mathrm{E}}^2, \qquad (194)$$

$$\left\|\mathrm{P}_\gamma \rho^{(k)}\right\|_{\mathrm{E}}^2 = \left\|\mathrm{P}_\gamma\left(f_{\delta,\varepsilon} - \mathrm{A}\widetilde{x}^{(k)}\right)\right\|_{\mathrm{E}}^2, \qquad (195)$$

$$\left\|\mathrm{Q}_\gamma \rho^{(k)}\right\|_{\mathrm{E}}^2 = \left\|\mathrm{Q}_\gamma\left(f_{\delta,\varepsilon} - \mathrm{A}\widetilde{x}^{(k)}\right)\right\|_{\mathrm{E}}^2; \qquad (196)$$

при этом, если величина

$$\varepsilon^2 = \left\|\mathcal{E}\right\|_{\mathrm{E}}^2 = \sum_{i=1}^{\mathrm{N}} \varepsilon_i^2 \qquad (197)$$

априорно известна, то по ней должно определяться, в соответствии с (191), значение $\gamma = \gamma_k$ для каждого приближения $\widetilde{x}^{(k)}$. Если же величина ε^2 неизвестна, а известны лишь оценки для нее сверху и снизу

$$0 < \varepsilon_{\min}^2 \le \sum_{i=1}^{\mathrm{N}} \varepsilon_i^2 \le \varepsilon_{\max}^2 < +\infty, \qquad (198)$$

то ситуация сложнее – здесь необходимо использовать технику распознавания образов, основанную на совокупности множества показателей качества решения.

Можно, однако, ставить задачу построения оптимальных решений и в случае, когда вообще отсутствует информация о величинах $\left\|\delta f\right\|_{\mathrm{E}}$ и $\left\|\mathcal{E}\right\|_{\mathrm{E}}$.

Далее. После всего сказанного о случае мультипликативно-аддитивной помехи в задании правой части системы легко понять, что более общий случай задания и правой части, и матрицы системы с мультипликативно-аддитивной помехой редуцируется к случаю чисто аддитивных помех в задании правой части и матрицы. Это следует из простейшего преобразования аппроксимации (119) к виду

$$\mathrm{A}_{\Delta,\,\mathrm{D}}\, x = \left(\mathrm{E} + \overset{\circ}{\mathcal{E}}\right) f_{\delta,\varepsilon}, \qquad (199)$$

где

$$\mathrm{E} + \overset{\circ}{\mathcal{E}} = \left(\mathrm{E} + \mathbf{D}\right)^{-1}\left(\mathrm{E} + \mathcal{E}\right) \qquad (200)$$

и $\overset{\circ}{\mathcal{E}}$ есть диагональная матрица с элементами

$$\overset{\circ}{\mathcal{E}}_i = \frac{1 + \varepsilon_i}{1 + \partial_i} \approx 1 + \varepsilon_i - \partial_i, \left|\overset{\circ}{\mathcal{E}}_i\right| \ll 1. \qquad (201)$$

Все остальное, в свете сказанного в данном параграфе относительно редукции аппроксимации (118) к аппроксимации (2), должно быть читателю достаточно очевидно.

§ **7**. Остается лишь высказать некоторые дополнительные соображения по поводу использования методов (алгоритмов) распознавания образов в процессе нахождения искомого устойчивого приближенного решения системы (1), заданной какой-либо аппроксимацией. Признание того, что без использования процедур распознавания в задаче нахождения устойчивых приближенных решений систем линейных алгебраических уравнений большой (тем более – сверхбольшой) размерности обойтись нельзя – принципиально важный элемент новой теории регуляризации, адекватной реальной геофизической практике.

Выше уже по существу было сказано, что распознавание образов может быть (и практически – должно!) использоваться в трех вариантах:

а) в анализе ситуации – нормальная она или аномальная;

б) при нахождении множеств пробных и допустимых приближенных решений;

в) при нахождении искомых – окончательных – приближенных решений на основе использования “принципа усреднения” допустимых решений, в специальной процедуре нахождения весовых коэффициентов p_k.

Фактически процедуры распознавания образов реализуются в двух формах:

1) в форме рангового распознавания;

2) в форме эвристических процедур проверки некоторой совокупности логических условий, обычно – типа равенств и неравенств.

Ранговое распознавание используется в процедурах назначения весов и выделения множеств допустимых решений. Его суть в следующем. Пусть задано некоторое множество приближенных решений и для него необходимо указать совокупность некоторых характеристических числовых показателей. Например, в задаче сокращения множества пробных решений до множества допустимых решений такими показателями являются числа 1 и 0, 1 – для допустимых решений, 0 – для тех, которые не включаются в число допустимых. Равным образом в задаче нахождения весовых коэффициентов в правиле усреднения (131)–(132) допустимых решений эти коэффициенты и представляют собой характеристические числовые показатели.

Процедура рангового распознавания состоит в том, что вводится так называемое "признаковое пространство"; при этом роль "признаков" выполняют значения различных функционалов на приближенных решениях; обозначим эти функционалы $\omega_q, q = 1, 2, \ldots, \mathrm{Q}$, а значения (неотрицательных) функционалов на приближенных решениях $\widetilde{x}^{(k)}$, $k = k_{\min}$, $k_{\min}+1, \ldots, k_{\max}$, через $\omega_q\left(\widetilde{x}^{(k)}\right) = \omega_{q,k}$. Принимается, что все признаки (функционалы ω_q) введены так, что чем меньше значение $\omega_{q,k} > 0$, тем приближенное решение $\widetilde{x}^{(k)}$ по данному признаку лучше. Совокупность приближенных решений $\left\{\widetilde{x}^{(k)}\right\}_{k=k_{\min}}^{k=k_{\max}}$ ранжируется по каждому признаку; соответствующие ранги $\mathrm{r}g\,\omega_{q,k}$ изменяются от 1 до $k = \left(k_{\max}\right) - k_{\min} + 1$. При этом ранг 1 приписывается тому приближенному решению $\widetilde{x}^{(k)}$, для которого $\omega_{q,k} = \min\limits_k$, ранг 2 – тому, для которого $\omega_{q,k} = \min\limits_k$ из всех остальных приближенных решений, и т.д., смысл процедуры назначения рангов по q-му признаку ясен. Таким образом определяются ранги приближенных решений $\widetilde{x}^{(k)}$, $k_{\min} \leq k \leq k_{\max}$, по всем признакам $q, 1 \leq q \leq \mathrm{Q}$. Далее осуществляется процедура определения для каждого приближенного решения суммарного ранга,

$$\mathrm{R}g\sum(k) = \sum_{q=1}^{\mathrm{Q}} \mathrm{r}g\,\omega_{q,k}. \tag{202}$$

Далее по найденным суммарным рангам и осуществляется нахождение характеристических числовых показателей. В задаче выделения допустимых решений назначается некоторое число Crit, и если выполняется неравенство

$$\mathrm{R}g\sum(k) \leq \mathrm{C}rit, \tag{203}$$

то данному приближенному решению $\widetilde{x}^{(k)}$ сопоставляется характеристический числовой показатель 1, т.е. $\widetilde{x}^{(k)}$ считается допустимым решением; если же неравенство (203) места не имеет, то приближенному решению $\widetilde{x}^{(k)}$ сопоставляется характеристическое число 0, т.е. это приближенное решение допустимым не считается.

В задаче же определения весовых множителей p_k в принципе усреднения пробных решений используется формула

$$p_k = \frac{q_k}{\sum q_k}, \tag{204}$$

в которой

$$q_k = \mathrm{R}g\sum(k). \tag{205}$$

Теперь относительно процедур распознавания образов, основанных на использовании эвристических процедур проверки совокупности логических условий. Эти процедуры поясним на важнейшем примере выделения множеств пробных решений в ситуации, когда либо постоянные в неравенствах (115) неизвестны, либо нарушается условие достаточной близости $\delta^2_{\min}$ к $\delta^2_{\max}$, см. (117). В данных процедурах также используется набор функционалов $\Omega_p(x) \geq 0$, $p = 1, 2, \ldots, \mathrm{P}$, при этом вводятся условия в форме неравенств, выделяющие пробные решения; если

$$0 \leq C^2_{\min}(p) \leq \Omega_p\left(\widetilde{x}^{(k)}\right) \leq C^2_{\max}(p) < +\infty, \tag{206}$$

где $C^2_{\min}(p)$ и $C^2_{\max}(p)$ суть заданные постоянные, то приближенное решение $\widetilde{x}^{(k)}$ является кандидатом в пробные решения по функционалу $\Omega_P(x)$. Принимается, что если данное приближенное решение $\widetilde{x}^{(k)}$ является кандидатом в пробные по заданной доле всех функционалов $\Omega_p(x)$ (например, оно является кандидатом в пробные не менее чем по 3/4 от общего числа функционалов), то оно причисляется к числу пробных. Критерии останова процедуры нахождения пробных решений с помощью проверки логических условий могут быть различными:

а) уже найдено априорно заданное число пробных решений;

б) подряд для заданного числа приближенных решений ни одно из них не оказалось принадлежащим к числу пробных (после того, как некоторое число пробных решений было уже найдено);

– и т.д., и т.п.

Во второй части работы оба основных подхода к использованию методов распознавания образов в компьютерных алгоритмах нахождения устойчивых приближенных решений систем линейных алгебраических уравнений с приближенными данными будут дополнительно обсуждены.

§ **8**. И самое последнее в данном разделе, посвященном изложению основных положений новой теории регуляризации систем линейных алгебраических уравнений с приближенными данными. В § 2 раздела

1 статьи было проведено классическое определение регуляризующего алгоритма по А. Н. Тихонову, [см. *Иванов и др.,* 1978; *Лисковец,* 1981; *Морозов,* 1974, 1987; *Танана,* 1981; *Тихонов, Арсенин,* 1979; *Тихонов и др.,* 1985]. В течение достаточно длинного периода (по крайней мере в 60-е–70-е годы) главным в теории линейных некорректных задач было именно доказательство регулярности различных конкретных алгоритмов, [см. *Алифанов и др.,* 1998; *Воеводин,* 1969; *Иванов,* 1962, 1966, 1977; *Иванов и др.,* 1978; *Лаврентьев,* 1955, 1959, 1962; *Лисковец,* 1981; *Лоусон, Хенсон,* 1986; *Морозов,* 1973, 1974, 1987; *Танана,* 1981]. Однако в дальнейшем было понято, что теоретического обоснования регулярности недостаточно, ибо наличие ошибок округления при реализации регулярных методов на практике может радикально изменять ситуацию. Именно по этой причине в данной статье доказательства регулярности различных обсуждаемых алгоритмов не приводятся[6].

В новой теории регуляризации систем линейных алгебраических уравнений с приближенными данными упор делается на других моментах:

1) нахождения достаточно точных и устойчивых приближенных решений систем в условиях большой априорной неопределенности в информации об искомом решении и полезном сигнале, а также помехах;

2) эффективности алгоритмов по быстродействию (что принципиально важно в случае систем большой и сверхбольшой размерности);

3) использовании распознавания образов для выделения допустимых решений и назначении коэффициентов в принципе усреднения допустимых решений.

По указанным причинам во второй части работы при описании конкретных методов проблема собственно регулярности алгоритмов нахождения допустимых решений *рассматриваться не будет*; этот аспект в рамках новой теории не имеет сколько-нибудь определяющего значения.

4. Классификация основных конструктивных идей, используемых при построении регуляризации систем линейных алгебраических уравнений, разработанных в рамках новой теории

§ 1. В рамках новой теории регуляризации систем линейных алгебраических уравнений с приближенными данными, основные положения которой описаны в предыдущем разделе статьи, было разработано значительное количество конкретных методов (алгоритмов) нахождения приближенных решений систем (1), заданных аппроксимациями (2). При этом разработка методов осуществлялась с учетом следующих основных требований:

1) экономичности, т.е. возможности решения систем большой размерности за приемлемое время на современной персональной вычислительной технике (типа Pentium–2 с оперативной памятью размером 256 мегабайт и тактовой частотой 400 мегагерц);

2) устойчивости, т.е. минимального влияния тех ошибок округления, которые с неизбежностью возникают при проведении вычислений.

Совершенно очевидно, в свете всего сказанного в предыдущем разделе, что понятие *метод* (соответственно – *алгоритм*) нахождения устойчивого приближенного решения системы (1) по заданной ее аппроксимации (2) включает три основные составные части:

A) метод (алгоритм) нахождения приближенных решений системы (2) по заданным значениям параметров счета;

B) метод (алгоритм) управления процессом нахождения последовательности приближенных решений $\widetilde{x}^{(k)}$, формирование множества пробных и допустимых решений в "нормальных" ситуациях, нахождение итогового решения в таких ситуациях;

C) метод (алгоритм) формирования множеств пробных и допустимых решений в "аномальных" ситуациях, нахождение итогового решения в таких ситуациях.

Ясно, что в некотором смысле определяющим (формирующим) является метод (алгоритм) A.

На первый взгляд кажется, что необходимо дать классификацию отдельно: 1) методов A); 2) методов B); 3) методов C). Однако гораздо лучше классифицировать не методы, а *методообразующие идеи*, или *конструктивные*, на что неоднократно указывал В. Н. Страхов, [см. *Страхов,* 1996a, 1996b, 1996c, 1998a, 1998b]. Суть дела здесь в том, что конструктивных (методообразующих) идей существенно меньше, чем методов, ибо последние всегда представляют собой реализацию определенных комбинаций конструктивных идей. Выделение конструктивных идей поэтому можно назвать "анатомией методов". По указанным соображениям ниже приводится именно классификация конструктивных (методообразующих) идей из методов класса A (нахождения приближенных решений при заданных значениях параметров счета).

§ 2. Конструктивные идеи, комбинации которых образуют методы (здесь рассматриваются только

[6] Например, нетрудно получить следующий результат: если все допустимые решения получены с помощью некоторого регулярного алгоритма, и если при $\delta^2_{\max} \longrightarrow 0$ выполняется неравенство

$$\frac{\delta^2_{\min}}{\delta_{\max}} > C > 0,$$

то тогда алгоритм, основанный на нахождении пробных решений и правиле их усреднения, является регулярным.

методы класса А)) можно вводить, исходя их различных соображений. Авторам представляется, что наиболее полезны и понятны три классификации:

I) основанная на степени близости новых конструктивных идей к тем, которые уже давно эксплуатируются в рамках классической теории регуляризации линейных некорректных задач;

II) основанная на оценке степени перспективности (по соображениям полноты учета априорной информации, быстродействия и устойчивости вычислений);

III) основанная на том, используются или не используются ортогональные преобразования, и если используются, то какого типа.

§ **3**. Начнем с краткого изложения классификации I). В ней содержится три раздела: I_1), I_2), I_3).

Конструктивные идеи группы I_1) – наиболее близкие по смыслу к тем, которые используются уже много лет в классической теории линейных некорректных задач (см. раздел 1 настоящей статьи). Это суть следующие идеи:

I_{11}) использования постановок условных экстремальных задач, основанных на оценках предельной относительной погрешности приближенных решений; здесь искомыми величинами являются элементы матрицы $\Re$, фигурирующие в представлениях (134) приближенных решений, см. § 2 предыдущего раздела статьи;

I_{12}) использования расширенных систем линейных алгебраических уравнений 1-го рода; это конструктивная идея кратко описывается ниже, см. следующий параграф;

I_{13}) использования новых итерационных методов классического типа; здесь речь идет о методах, основанных на допущении о существенной разделенности спектров полезного сигнала и помехи, см. § 8 раздела 1-го статьи; эти методы кратко описываются во второй части работы;

I_{14}) использования мультипликативно-аддитивной регуляризации; здесь речь идет о тех конструктивных приемах, которые изложены в § 6 третьего раздела статьи;

I_{15}) использования приема получения решений, удовлетворяющих характеристическому уравнению метода наименьших квадратов; здесь речь идет о приеме, описанном в § 4 предыдущего раздела статьи;

I_{16}) использования приема введения в классические (и новые) конструкции дополнительного параметра с целью получения возможности использования различных “принципов сохранения”, о которых коротко было сказано в § 5 предыдущего раздела статьи; этот прием коротко описывается в следующем параграфе настоящего раздела.

Конструктивные идеи группы I_2) принципиально отличны от идей группы I_1), они предложены В. Н. Страховым или совместно В. Н. Страховым и А. В. Страховым в последние 5–6 лет, см. работы [*Страхов,* 1990a, 1990b, 1991a, 1991b, 1991c, 1992a, 1992b, 1997d, 1997e, 1997f, 1997g, 1997h; *Страхов, Страхов,* 1998, 1999a, 1999b, 1999c, 1999d; *Страхов, Тетерин,* 1991a, 1991b, 1991c, 1993; *Strakhov et al.,* 1995]. Это суть следующие идеи:

I_{21}) использования приема построения последовательности систем малой размерности – на основе процедур ортогональных преобразований исходной системы; это исключительно важная и плодотворная идея подробно описывается во второй части работы;

I_{22}) использования приема авторегуляризации; это также очень важная и плодотворная идея, суть которой в том, чтобы в процессе ортогональных преобразований (по ходу процесса преобразований) осуществлять приближенные процедуры выметания матриц; подробно эта идея рассматривается во второй части работы;

I_{23}) использования фильтрации системы – удаления из нее, по ходу процесса ортогональных преобразований, неинформативных (либо наименее информативных) строк и столбцов; эта очень важная идея подробно рассматривается во второй части работы;

I_{24}) использования так называемой “непараметрической регуляризации” (этот не самый удачный термин введен В. Н. Страховым, впервые предложившим эту идею); в рамках этой конструктивной идеи, в основном относящейся к системам с симметричными положительно полуопределенными и треугольными матрицами, осуществляется процедура уменьшения внедиагональных элементов с одновременным увеличением диагональных элементов; данная идея кратко описывается в следующем параграфе данного раздела статьи;

I_{25}) использования приема регуляризации разложения матрицы на треугольные множители; это одна из важнейших конструктивных идей, она рассматривается во второй части работы;

I_{26}) использования специального приема регуляризации систем с треугольными матрицами, так называемой “треугольной регуляризации”; эта идея также рассматривается во второй части работы;

I_{27}) использования приема ускорения сходимости последовательности приближенных решений; данная конструктивная идея рассматривается во второй части работы;

I_{28}) использования приема нахождения устойчивого приближенного решения системы большой размерности на основе построения устойчивого приближенного решения подсистем существенно меньшей размерности; эта исключительно важная конструктивная идея подробно рассматривается во второй части работы;

I_{29}) использования приближенных линейных аналитических аппроксимаций вектора неизвестных, в которых число параметров существенно меньше размерности искомого вектора; эта конструктивная идея также рассматривается во второй части работы.

Наконец, конструктивные идеи группы I_3) являются в максимальной степени отличными от тех, которые используются в классической теории линейных некорректных задач; эти идеи были предложены В. Н. Страховым и А. В. Страховым в ряде работ, выполненных в 1997–1998 гг., [см. *Страхов, Страхов,* 1998, 1999a, 1999b, 1999c, 1999d]. Это суть следующие идеи, детальное описание которых дается во второй части работы:

I_{31}) использования редукции исходной системы к системе с вектором правой части, содержащим всего одну ненулевую компоненту (первую или последнюю), при этом редуцирование в обязательном порядке включает ортогональное преобразование либо исходной, либо вспомогательной матрицы, на которую потом умножаются обе части исходной системы;

I_{32}) использования редукции исходной системы к последовательности двух уравнений с одним неизвестным; данная редукция также в существенном основывается на использовании ортогональных преобразований;

I_{33}) использование приема редукции исходной системы к одному уравнению с одним неизвестным; эта идея в сжатой форме изложена в § 4 предыдущего раздела данной статьи, и из этого изложения ясно, что здесь также основную роль играют ортогональные преобразования;

I_{34}) использования итерационных процессов нового типа, основанного на корреляционных преобразованиях столбцов матрицы A;

I_{35}) использования итерационных процессов нового типа, основанных на ортогональных преобразованиях, изменяющих некоторое отношение эвклидовых норм парциальных векторов первого столбца матрицы.

Итак, общее число новых конструктивных идей класса A), используемых в рамках новой общей теории регуляризации систем линейных алгебраических уравнений с приближенно заданной правой частью (случай чисто аддитивной помехи) равно 20, из них только 6 имеют достаточно очевидную связь с конструктивными идеями классической теории регуляризации линейных некорректных задач. В следующей части работы будет приведено достаточно детальное описание 14 из этих идей.

§ 4. Итак, приведем краткие описания трех новых конструктивных идей: I_{12}), I_{16}) и I_{24}) по первой классификации.

Конструктивная идея I_{12}) – использование расширенных регуляризованных систем первого рода. Ее суть в том, что система (22), возникающая в классическом вариационном методе регуляризации А. Н. Тихонова, может рассматриваться как первая трансформация Гаусса следующей системы:

$$\Gamma_\alpha x_\alpha = \mathrm{F}_\delta, \tag{207}$$

где

$$\Gamma_\alpha = \left|\begin{array}{c|c} \mathrm{A} & \begin{matrix}\uparrow\\ \mathrm{N}\\ \downarrow\end{matrix} \\ \hline \sqrt{\alpha}\mathrm{R} & \begin{matrix}\uparrow\\ \mathrm{P}\\ \downarrow\end{matrix} \end{array}\right|, \quad \mathrm{F}_\delta = \left|\begin{array}{c|c} f_\delta & \begin{matrix}\uparrow\\ \mathrm{N}\\ \downarrow\end{matrix} \\ \hline 0 & \begin{matrix}\uparrow\\ \mathrm{P}\\ \downarrow\end{matrix} \end{array}\right|. \tag{208}$$

Иначе говоря, система (22) порождается следующей задачей наименьших квадратов:

$$\left\| \mathrm{F}_\delta - \Gamma_\alpha x_\alpha \right\|_{\mathrm{E}}^2 = \min_{x_\alpha} \tag{209}$$

– при фиксированном значении параметра α.

Из сказанного следует два вывода:

во-первых, вариационный метод А. Н. Тихонова действительно представляет собой реализацию некоторого варианта метода наименьших квадратов – для расширенной системы (207), которую естественно назвать *расширенной регуляризованной системой 1-го рода*;

во-вторых, решение системы (207) при заданном $\alpha > 0$ может быть найдено большим числом методов (алгоритмов), как прямых, так и итерационных.

При нахождении решений систем (207) возможно использование различного рода ортогональных преобразований, приводящих ее к максимально простому виду. Читателю рекомендуется рассмотреть возникающие возможности самостоятельно.

Конструктивная идея I_{16}) – введения дополнительных параметров в целях использования тех или иных “законов сохранения” для возмущенных матриц.

Простейший пример реализации этой идеи: переход от уравнения (10) к уравнению

$$\mathrm{A}_{\alpha,\beta} x_{\alpha,\beta} = \Big(\alpha \mathrm{S} + \mathbf{D}_\mathrm{A} + (1-\beta)(\mathrm{A} - \mathbf{D}_\mathrm{A})\Big) x_{\alpha,\beta} = f_\delta. \tag{210}$$

Здесь $\beta, 0 < \beta < 1$, суть новый (дополнительный) параметр, $\mathbf{D}_\mathrm{A}$ – диагональная матрица, элементы которой $\partial_i = a_{ii}$ – диагональным элементам матрицы A. Ясно, что:

1) система (210) эквивалентна более простой, в вычислительном плане, системе

$$\left(\frac{\alpha}{1-\beta}\mathrm{S} + \frac{\beta}{1-\beta}\mathbf{D}_\mathrm{A} + \mathrm{A}\right) x_\alpha = \frac{f_\delta}{1-\beta}; \tag{211}$$

2) возможно использовать следующие соотношения между характеристиками матриц $\mathrm{A}_{\alpha,\beta}$ и A:

$$\left\|\mathrm{A}_{\alpha,\beta}\right\|_{\mathrm{E}}^{2}=\left\|\mathrm{A}\right\|_{\mathrm{E}}^{2} \tag{212}$$

и

$$\mathrm{Sp}\mathrm{A}_{\alpha,\beta}=\sum_{i=1}^{\mathrm{M}} a_{ii}(\alpha,\beta)=\sum_{i=1}^{\mathrm{M}} a_{ii}=\mathrm{Sp}\mathrm{A}, \tag{213}$$

а также некоторых других. Соотношения (212) и (213) и выражают "законы сохранения".

Ясно, что "законы сохранения" позволяют фактически сократить число параметров счета с двух (α,β) до одного – α или β; при этом могут возникнуть некоторые дополнительные условия на допустимые значения параметров α и β.

Ясно, что подобного рода конструкции могут быть использованы и в случае уравнения (22); имеется в виду переход к уравнению

$$\begin{aligned}&\left(\alpha\mathrm{R}^{\mathrm{T}}\mathrm{R}+\mathbf{D}_{\mathrm{A}^{\mathrm{T}}\mathrm{A}}+(1-\beta)\right.\\&\left.\times\left(\mathrm{A}^{\mathrm{T}}\mathrm{A}-\mathbf{D}_{\mathrm{A}^{\mathrm{T}}\mathrm{A}}\right)\right)x_{\alpha,\beta}\\&=\varphi_{\delta}=\mathrm{A}^{\mathrm{T}}f_{\delta},\end{aligned} \tag{214}$$

а также и в других методах. Более подробное рассмотрение этих возможностей предоставляется читателю.

Конструктивная идея I_{24}) – идея использования "непараметрической регуляризации". (Как уже сказано выше, предложенный В. Н. Страховым термин "непараметрическая регуляризация" [*Страхов,* 1997f], которым подчеркивается отличие данной конструкции от классических, в котором вводимый параметр возникает при решении некоторой условной экстремальной задачи методом множителей Лагранжа, не является очень удачным; но он достаточно краток, уже экплуатируется и неясно, каким кратким термином его заменить.)

Отправная позиция идеи "непараметрической регуляризации" в случае систем (1)–(2) с $\mathrm{A}=\mathrm{A}^{\mathrm{T}}\geq 0$ та же, что и в методе М. М. Лаврентьева – необходимо усилить диагональные элементы системы. Поэтому в методе непараметрической регуляризации используются два взаимодополняющих приема:

во-первых, уменьшения внедиагональных элементов;

во-вторых, усиления диагональных элементов.

В случае систем (2) с $\mathrm{A}=\mathrm{A}^{\mathrm{T}}\geq 0$ "непараметрическая регуляризация" состоит в переходе к регуляризованной системе

$$\hat{\hat{\mathrm{A}}}_{\alpha,\beta}x_{\alpha,\beta}=\left((1+\alpha)\mathbf{D}_{\mathrm{A}}+\mathrm{A}_{\beta}\right)x_{\alpha,\beta}=f_{\delta}, \tag{215}$$

где $\alpha>0$ параметр, $\mathbf{D}_{\mathrm{A}}=\mathrm{diag}a_{ii}$,

$$\widehat{\mathrm{A}}_{\beta}=\widehat{\mathrm{A}}_{\beta}^{\mathrm{T}}, \tag{216}$$

и $\widehat{\mathrm{A}}_{\beta}$ есть матрица с нулевой диагональю, элементы которой определены через внедиагональные элементы a_{ij} матрицы A и параметр $\beta, 0<\beta<1$:

$$\widehat{a}_{ij}(\beta)=\begin{cases}(1-\beta)a_{i,j}, & \text{если}\left|a_{i,j}\right|\geq(1-\beta)\max\limits_{i,j,\ i\neq j}\left|a_{i,j}\right|,\\ a_{i,j}, \quad \text{если} \quad i\neq j, & \beta_{\max}\left|a_{i,j}\right|\leq\left|a_{i,j}\right|<(1-\beta)\max\limits_{i,j,\ i\neq j}\left|a_{i,j}\right|,\\ 0, \quad \text{если} \quad i\neq j, & \left|a_{i,j}\right|<\beta\max\limits_{i,j,\ i\neq j}\left|a_{i,j}\right|.\end{cases} \tag{217}$$

Ясно, что и в рамках данной конструкции возможно использование "законов сохранения":

$$\left\|\hat{\hat{\mathrm{A}}}_{\alpha,\beta}\right\|_{\mathrm{E}}^{2}=\left\|\mathrm{A}\right\|_{\mathrm{E}}^{2}, \tag{218}$$

или

$$\mathrm{Sp}\hat{\hat{\mathrm{A}}}_{\alpha,\beta}=\mathrm{Sp}\mathrm{A}, \tag{219}$$

позволяющие перейти от двух параметров счета (α,β) к одному – α или β, с возможным введением дополнительных ограничений на значения параметров.

Практически весьма важно, что описанная конструкция "непараметрической регуляризации" элементарным образом переносится на случай систем линейных уравнений (1) и (2) с треугольными матрицами A – верхней или нижней, при $a_{ii}>0$.

Действительно, соотношения (217) и (218)–(219) и в этом случае (в (217) следует принимать $i-j>0$) могут быть оставлены в силе. При этом мотивация идеи непараметрической регуляризации в случае треугольной A основывается на хорошо известной оценке эвклидовой нормы обратной матрицы, см. [*Лоусон, Хенсон,* 1986], стр. 197. В случае верхней треугольной A эта оценка имеет вид:

$$\left\|\mathrm{A}^{-1}\right\|_{\mathrm{E}}^{2}\leq\sum_{i=1}^{\mathrm{M}}\frac{\mu_i}{a_{ii}^{2}}, \tag{220}$$

где μ_i выражаются соотношениями

$$\begin{aligned}&\mu_1=1,\mu_i=\left(1+c_{i-1}\right)^{2}\mu_{i-1}-2c_{i-1},\\&i=2,3,\ldots,\mathrm{M},\end{aligned} \tag{221}$$

причем

$$c_i = \frac{\max\limits_{j,j>i}|a_{i,j}|}{a_{ii}}. \qquad (222)$$

§ 5. В § 3 данного раздела была приведена первая из трех возможных классификаций 20 новых конструктивных идей. В настоящем, заключительном, параграфе раздела кратко описываются две другие классификации этих 20 конструктивных идей.

Вторая классификация (II). Напомним, что это классификация по степени перспективности (эффективности) конструктивных идей. Она включает подразделение идей также на три группы:

– II_1) наиболее перспективных идей;

– II_2) перспективных идей;

– II_3) наименее перспективных идей.

К группе II_1) относятся идеи (по первой классификации):

I_{13}), I_{21}), I_{22}), I_{25}), I_{33}), I_{35}).

К группе II_2) относятся идеи:

I_{12}), I_{15}), I_{23}), I_{24}), I_{28}), I_{29}), I_{31}), I_{32}), I_{34}).

Наконец, к группе II_3) относятся идеи:

I_{11}), I_{14}), I_{16}), I_{26}), I_{27}).

Третья классификация (III) (основанная на том, используются в реализации этой идеи ортогональные преобразования, или нет). Она, естественно, включает подразделение идей на две группы:

III_1) ортогональные преобразования используются;

III_2) ортогональные преобразования не используются.

К группе III_1) относятся идеи (по первой классификации):

I_{14}), I_{21}), I_{22}), I_{23}), I_{31}), I_{32}), I_{33}), I_{34}), I_{35}).

К группе III_2) относятся идеи:

I_{11}), I_{12}), I_{13}), I_{15}), I_{16}), I_{24}), I_{25}), I_{26}), I_{27}), I_{28}), I_{29}).

Читатель безусловно обратит внимание на тот факт, что в группе III_1) число наиболее перспективных и перспективных идей больше, чем в группе III_2), хотя их общие численности одинаковы. Здесь четко просматривается тот уже упоминавшийся факт, что в новой теории регуляризации систем линейных алгебраических уравнений ортогональные преобразования играют существенно большую роль, чем в классической.

Заключение

Главная цель настоящей статьи (первой части работы) состоит в том, чтобы изложить основы классической теории регуляризации систем линейных алгебраических уравнений с приближенными данными – с одной стороны, и основы новой теории регуляризации линейных систем (созданной трудами В. Н. Страхова, частично в соавторстве, см. [*Страхов,* 1990a, 1990b, 1991a, 1991b, 1991c, 1992a, 1992b, 1997c, 1997d, 1997e, 1997f, 1997g, 1997h, 1998d, 1998e; *Страхов, Страхов,* 1998, 1999a, 1999b, 1999c, 1999d; *Страхов, Тетерин,* 1991a, 1991b, 1991c, 1993; *Strakhov et al.,* 1995]) – с другой, и продемонстрировать тот факт, что новая теория существенно более адекватна реальной геофизической практике, что она гибче и богаче классической теории. Вместе с тем материал данной статьи является сугубо вводным, только в следующих частях будет в полном объеме продемонстрирована суть основных новых методов (алгоритмов) построения пробных и допустимых решений, а также важнейшая, с точки зрения гравиметрии и магнитометрии, общая идея использования аппроксимационного подхода в целом, и конкретно – методов редуцирования аппроксимационных постановок к нахождению устойчивых приближенных решений систем линейных алгебраических уравнений с приближенными данными, и к демонстрации эффективности аппроксимационного подхода на модельных и практических примерах.

Литература

Алифанов О. М., Румянцев С. В., Об устойчивости итерационных методов решения линейных некорректных задач, *Докл. АН СССР, 242,* (6), 1289–1291, 1978.

Алифанов О. М., Румянцев С. В., Регуляризирующие итерационные алгоритмы для решения обратных задач теплопроводности, *Инж.-физ. журн., 39,* (2), 253–258, 1980.

Алифанов О. М., Артюхин Е. А., Румянцев С. В., *Экстремальные методы решения некорректных задач,* 286 с., Наука, Москва, 1998.

Альберт А., *Регрессия, псевдоинверсия и рекуррентное оценивание,* Наука, Москва, 1977.

Бакушинский А. Б., Методы решения монотонных вариационных неравенств, основанные на принципе итерационной регуляризации, *Журн. выч. математики и мат. физики, 176,* 1350–1362, 1977.

Бакушинский А. Б., Страхов В. Н., О решении некоторых интегральных уравнений 1-го рода методом последовательных приближений, *Журн. выч. математики и мат. физики, 8,* (1), 181–185, 1968.

Вайникко Г. М., *Методы решений линейных некорректных задач в гильбертовых пространствах,* 111 с., Изд-во Тартус. ун-та, Тарту, 1982.

Васильев Ф. П., Якимович М. Д., Об итеративной регуляризации метода условного градиента и метода Ньютона при неточно заданных исходных данных, *Докл. АН СССР, 250,* (2), 165–269, 1980.

Воеводин В. В., О методе регуляризации, *Журн. выч. математики и мат. физики, 9,* (3), 673–675, 1969.

Воеводин В. В., Кузнецов Ю. А., *Матрицы и вычисления,* 318 с., Наука, Москва, 1984.

Гилязов С. Ф., Об устойчивом решении линейных операторных уравнений 2-го рода методом наискорейшего спуска, *Вестн. МГУ, сер. Выч. математика и кибернетика, 4,* (3), 26–37, 1980.

Емелин И. В., Красносельский М. А., Правило останова в итерационных процедурах решения некорректных задач, *Автоматика и телемеханика,* (2), 59–63, 1978.

Иванов В. К., О линейных некорректных задачах, *Докл. АН СССР, 145,* (2), 270–272, 1962.

Иванов В. К., О приближенном решении операторных уравнений первого рода, *Журн. выч. математики и мат. физики, 6,* (6), 1089–1094, 1966.

Иванов В. К., Некоторые вопросы оценок устойчивости при решении некорректных задач, *Неклассические методы в геофизике,* с. 45–53, Новосибирск, 1977.

Иванов В. К., Васин В. В., Танана В. П., *Теория линейных некорректных задач и ее приложения,* 206 с., Наука, Москва, 1978.

Кобрунов А. И., О методе оптимизации при решении обратной задачи гравиразведки, *Изв. АН СССР, Физика Земли,* (8), 73–78, 1978.

Кобрунов А. И., К вопросу об интерпретации аномальных гравитационных полей методом оптимизации, *Изв. АН СССР, Физика Земли,* (10), 67–78, 1979.

Кобрунов А. И., О классах оптимальности решения обратной задачи гравиразведки, *Изв. АН СССР, Физика Земли,* (2), 100–107, 1982.

Кобрунов А. И., Разрешимость и эквивалентность в обратной задаче гравиразведки для нескольких плотностных границ, *Изв. АН СССР, Физика Земли,* (5), 67–75, 1983.

Кобрунов А. И., *Экстремальные классы в задачах гравиметрии и их использование при построении плотностных моделей геологических сред,* Дис. ... д-ра физ.-мат. наук, 439 с., Ивано-Франковск, 1985.

Лаврентьев М. М., О задаче Коши для уравнения Лапласа, *Докл. АН СССР, 102,* (2), 205–206, 1955.

Лаврентьев М. М., Об интегральных уравнениях первого рода, *Докл. АН СССР, 127,* (1), 31–33, 1959.

Лаврентьев М. М., *О некоторых некорректных задачах математической физики,* Наука, Новосибирск, 1962, 91с.

Лисковец О. А., *Вариационные методы решения неустойчивых задач,* 343 с., Наука и техника, Минск, 1981.

Лоусон Ч., Хенсон Р., *Численное решение задач метода наименьших квадратов,* 232 с., Наука, Москва, 1986.

Марчук Г. И., Кузнецов Ю. А., *Итерационные методы и квадратичные функционалы,* Наука, Новосибирск, 1972.

Морозов В. А., Об устойчивых методах решения систем линейных алгебраических уравнений, *Вычислительные методы линейной алгебры,* с. 57–62, Изд-во СО АН СССР, Новосибирск, 1973.

Морозов В. А., *Регулярные методы решения некорректно поставленных задач,* 360 с., МГУ, Москва, 1974.

Морозов В. А., *Регулярные методы решения некорректно поставленных задач,* 239 с., Наука, Москва, 1987.

Оганесян С. М., Старостенко В. И., О корректности постановок задач геофизики, представленных в виде систем уравнений, и итерационных методах их решения, *Изв. АН СССР, Физика Земли,* (8), 54–64, 1978.

Румянцев С. В., Способы учета априорной информации в регуляризирующих градиентных алгоритмах, *Инж.-физ. журн., 49,* (6), 932–936, 1985.

Рязанцева И. П., Регуляризация уравнений с аккретивными операторами методом последовательного приближения, *Сиб. мат. журн., 21,* (1), 223–226, 1978.

Савр Л. Э., Класс итерационных методов для линейных некорректных самосопряженных задач в гильбертовом пространстве, *Докл. АН СССР, 278,* (5), 1066–1070, 1984.

Страхов В. Н., Теория приближенного решения линейных некорректных задач в гильбертовом пространстве и ее использование в разведочной геофизике, I, *Изв АН СССР, Физика Земли,* (8), 1968a.

Страхов В. Н., Теория приближенного решения линейных некорректных задач в гильбертовом пространстве и ее использование в разведочной геофизике, II, *Изв АН СССР, Физика Земли,* (9), 1968b.

Страхов В. Н., О методе последовательных приближений для линейных уравнений в гильбертовом пространстве, *Журн. выч. математики и мат. физики, 13,* (4), 1041–1044, 1973a.

Страхов В. Н., К вопросу о скорости сходимости в методе простой итерации, *Журн. выч. математики и мат. физики, 13,* (6), 1602–1606, 1973b.

Страхов В. Н., Об устойчивых методах решения линейных задач геофизики, Ч. I, Постановки и основные конструктивные идеи, *Изв. АН СССР, Физика Земли,* (8), 1990a.

Страхов В. Н., Об устойчивых методах решения линейных задач геофизики, Ч. II, Основные алгоритмы, *Изв. АН СССР, Физика Земли,* (9), 1990b.

Страхов В. Н., Алгебраические методы в линейных задачах гравиметрии и магнитометрии. Основные идеи, *Докл. АН СССР, 319,* (1), 143–146, 1991a.

Страхов В. Н., Алгебраические методы в линейных задачах гравиметрии и магнитометрии. Постановки экстремальных задач, *Докл. АН СССР, 319,* (2), 342–346, 1991b.

Страхов В. Н., Метод фильтрации систем линейных алгебраических уравнений – основа для решения линейных задач гравиметрии и магнитометрии, *Докл. АН СССР, 320,* (3), 595–600, 1991c.

Страхов В. Н., Обработка геофизической информации при мультипликативно-аддитивных помехах, *Докл. АН СССР, 319,* (3), 599–603, 1991d.

Страхов В. Н., Решение линейных задач геофизики при мультипликативно-аддитивных помехах, *Докл. АН СССР, 319,* (4), 845–848, 1991e.

Страхов В. Н., Решение линейных задач гравиметрии и магнитометрии вариационным и структурно-параметрическим методами при мультипликативно-аддитивных помехах в экспериментальной информации, *Докл. АН СССР, 319,* (6), 1361–1365, 1991f.

Страхов В. Н., Обработка геофизической информации в случае мультипликативно-аддитивных помех при неизвестной дисперсии мультипликативной компоненты, *Докл. АН СССР, 320,* (1), 82–84, 1991g.

Страхов В. Н., Решение задач геофизики редуцируемых к линейным интегральным уравнениям первого рода при мультипликативно-аддитивных помехах во входных данных, *Докл. АН СССР, 320,* (2), 307–310, 1991h.

Страхов В. Н., Решение линейных задач гравиметрии и магнитометрии при мультипликативно-аддитивных помехах, *Докл. АН СССР, 320,* (5), 1114–1116, 1991i.

Страхов В. Н., О линейных некорректных задачах гравиметрии и магнитометрии, *Условно-некорректные задачи математической физики и анализа,* с. 176–204, Наука, Новосибирск, 1992a.

Страхов В. Н., Алгоритмы редуцирования и трансформации аномалий силы тяжести, заданных на физической поверхности Земли, *Интерпретация гравитационных и магнитных полей,* с. 4–82, Наукова думка, Киев, 1992b.

Страхов В. Н., Основные направления развития теории и методологии интерпретации геофизических данных на рубеже XXI столетия, Ч. 1, *Геофизика,* (3), 9–18, 1995a.

Страхов В. Н., Основные направления развития теории и методологии интерпретации геофизических данных на рубеже XXI столетия, Ч. 2, *Геофизика,* (4), 10–20, 1995b.

Страхов В. Н., Геофизика и математика, *Изв. РАН, Физика Земли,* (12), 4–23, 1995c.

Страхов В. Н., Методологические проблемы теории и практики интерпретации данных в прикладной геофизике, *Вопросы методологии интерпретации геофизических данных в прикладной геофизике,* с. 4–20, ОИФЗ РАН, Москва, 1996a.

Страхов В. Н., Методологические особенности интерпретации данных гравиразведки и магниторазведки, *Вопросы методологии интерпретации геофизических данных в прикладной геофизике,* с. 110–123, ОИФЗ РАН, Москва, 1996b.

Страхов В. Н., Развитие общей теории интерпретации геолого-геофизических данных, *Основные достижения ОИФЗ РАН за 1992–1996 гг.,* Т 1, с. 48–53, ОИФЗ РАН, Москва, 1996c.

Страхов В. Н., Общая схема и основные итоги развития теории и практики интерпретации потенциальных полей в СССР и России в XX веке, *Развитие гравиметрии и магнитометрии в XX веке,* с 98–120, ОИФЗ РАН, Москва, 1997a.

Страхов В. Н., Третья парадигма в теории и практике интерпретации потенциальных полей (гравитационных и магнитных аномалий), Ч. I, *Вестн. ОГГГГН РАН,* #1(1), с. 163–198, 1997b.

Страхов В. Н., Третья парадигма в теории и практике интерпретации потенциальных полей (гравитационных и магнитных аномалий), Ч. II, *Вестн. ОГГГГН РАН,* #2(2), с. 55–82, 1997c.

Страхов В. Н., Общая теория нахождения устойчивых приближенных решений систем линейных алгебраических уравнений с приближенно заданными правыми частями и матрицами, возникающих при решении задач геофизики, *Вопросы теории и практики геологической интерпретации гравитационных и магнитных и электрических полей,* с. 38–42, ОИФЗ РАН, Москва, 1997d.

Страхов В. Н., Математический аппарат, используемый при конструировании алгоритмов нахождений устойчивых приближенных решений систем линейных алгебраических уравнений, возникающих в задачах гравиметрии и магнитометрии, *Вопросы теории и практики геологической интерпретации гравитационных и магнитных и электрических полей,* с. 43–75, ОИФЗ РАН, Москва, 1997e.

Страхов В. Н., Экстремальные задачи, непараметрическая регуляризация и фильтрация теории нахождения устойчивых приближенных решений систем линейных алгебраических уравнений с приближенно заданными правыми частями и матрицами, *Вопросы теории и практики геологической интерпретации гравитационных и магнитных и электрических полей,* с. 76–86, ОИФЗ РАН, Москва, 1997f.

Страхов В. Н., Обобщенные QR-алгоритмы нахождения устойчивых приближенных решений систем линейных алгебраических уравнений с приближенно заданной правой частью, возникающие при решении линейных задач гравиметрии и магнитометрии, *Вопросы теории и практики геологической интерпретации гравитационных и магнитных и электрических полей,* с. 87–89, ОИФЗ РАН, Москва, 1997g.

Страхов В. Н., Линейные задачи гравиметрии и магнитометрии в постановках, адекватных геофизической практике, *Вопросы теории и практики геологической интерпретации гравитационных и магнитных и электрических полей,* с. 89–104, ОИФЗ РАН, Москва, 1997h.

Страхов В. Н., Современное состояние и перспективы развития теории интерпретации гравитационных и магнитных аномалий, *Вопросы теории и практики геологической интерпретации гравитационных, магнитных и электрических полей,* с. 4–35, Воронеж, 1998a.

Страхов В. Н., Взгляд в будущее (перспективы развития теории интерпретации геофизических данных в начале XXI века), *III Междунар. конф. "Новые идеи в науках о Земле",* Избр. докл., МГГА, с 39–50, Москва, 1998b.

Страхов В. Н., Г. А. Гамбурцев – один из основоположников гравиразведки, *Григорий Александрович Гамбурцев: Воспоминания, очерки, статьи,* с 153–176, ОИФЗ РАН, 1998c.

Страхов В. Н., Третья парадигма в теории и практике интерпретации потенциальных полей (гравитационных и магнитных аномалий), Ч. III, *Вестн. ОГГГГН РАН,* #1(3), с. 100–152, 1998d.

Страхов В. Н., *Что делать? (О развитии гравиметрии и магнитометрии в России в начале XXI века),* 24 с., ОИФЗ РАН, Москва, 1998g.

Страхов В. Н., Страхов А. В., О регуляризации метода наименьших квадратов, *Геофиз. журн.,* (6), 18–38, 1998.

Страхов В. Н., Страхов А. В., О нахождении устойчивых приближенных решений систем линейных алгебраических уравнений большой и сверхбольшой размерности, возникающих в линейных задачах гравиметрии и магнитометрии, с. 194–213, *Наст. сб.*, 1999a, (в печати).

Страхов В. Н., Страхов А. В., О решении линейных задач гравиметрии и магнитометрии, *Геофиз. журн.*, 1999b, (в печати).

Страхов В. Н., Страхов А. В., Обобщения метода наименьших квадратов и регуляризованные алгоритмы нахождения устойчивых приближенных решений систем линейных алгебраических уравнений, возникающих при решении задач геофизики, I, *Геофиз. журн.*, 1999c, (в печати).

Страхов В. Н., Страхов А. В., Обобщения метода наименьших квадратов и регуляризованные алгоритмы нахождения устойчивых приближенных решений систем линейных алгебраических уравнений, возникающих при решении задач геофизики, II, *Геофиз. журн.*, 1999d, (в печати).

Страхов В. Н., Тетерин Д. Е., Линейные трансформации гравитационных и магнитных аномалий в случае многоэлементных съемок при произвольных сетях наблюдений, *Докл. АН СССР, 318,* (3), 572–576, 1991a.

Страхов В. Н., Тетерин Д. Е., Метод авторегуляризации при решении задач линейных трансформаций гравитационных и магнитных аномалий, *Докл. АН СССР, 318,* (4), 867–871, 1991b.

Страхов В. Н, Тетерин Д. Е., О методе авторегуляризации для решения линейных задач гравиметрии и магнитометрии, *Докл. АН СССР, 318,* (4), 871–874, 1991c.

Страхов В. Н., Тетерин Д. Е., *О методах сингулярных разложений матриц,* Деп. в ВИНИТИ, # 1915–93, 1993.

Танана В. П., *Методы решения операторных уравнений,* 155 с., Наука, Москва, 1981.

Тихонов А. Н., Об устойчивости обратных задач, *Докл. АН СССР, 39,* (5), 195–198, 1943.

Тихонов А. Н., О решении некорректно поставленных задач и методе регуляризации, *Докл. АН СССР, 151,* (3), 501–504, 1963a.

Тихонов А. Н., О решении некорректно поставленных задач и методе регуляризации, *Докл. АН СССР, 153,* (1), 49–52, 1963b.

Тихонов А. Н., О некорректных задачах линейной алгебры и устойчивом методе их решения, *Докл. АН СССР, 163,* (3), 591–594, 1965.

Тихонов А. Н., Арсенин В. Я., *Методы решения некорректных задач,* 286 с., Наука, Москва, 1979.

Тихонов А. Н., Гончарский А. В., Степанов В. В., Ягола А. Г., *Регуляризующие алгоритмы и априорная информация,* 198 с., Наука, Москва, 1985.

Уилкинсон Дж., Райнш К., *Справочник алгоритмов на языке Алгол, Линейная алгебра,* 248 с., Машиностроение, Москва, 1976.

Фаддеев Д. К., Фаддеева В. Н., *Вычислительные методы линейной алгебры,* Изд. 2-е, 734 с., Физматгиз, М.;Л., 1963.

Strakhov V. N., Vorontsov S. V., Seku T., Linear inv ersions in helio-seismology: testing new regularisation techniques for linear algebraic equations, *Cong'94 Helion- and Astero-Seismology, ASP Conf. Ser.,* Vol. 76, 1995.

(Поступила в редакцию 15 апреля 1999.)