3. **Туз Ю. М.** Структурные методы повышения точности измерительных устройств.— Киев: Высшая школа, 1976.
4. **Алиев Т. М., Сейдель Л. Р.** Автоматическая коррекция погрешности цифровых измерительных устройств.— М.: Энергия, 1975.
5. **Волгин Л. И.** Итерационные алгоритмы повышения точности измерительных устройств.— Автометрия, 1974, № 5, с. 84.
6. **Кэй Б. Дж.** Правильный выбор цифрового вольтметра.— Электроника, 1966, т. 39, № 7, с. 3.
7. **Попов В. П.** Об автоматической коррекции погрешности результатов аналого-цифрового преобразования.— Автометрия, 1976, № 5, с. 62.
8. **Басс Б. П. и др.** Некоторые способы автоматической коррекции нелинейности характеристики цифроаналоговых преобразователей.— Электрон. техника. Сер. Микроэлектроника, 1971, вып. 7 (33), с. 46.
9. **Смолов В. Б. и др.** Микроэлектронные цифроаналоговые и аналого-цифровые преобразователи информации.— Л.: Энергия, 1976, с. 191.

Поступила в редакцию 23 февраля 1981 г.;
окончательный вариант — 12 августа 1981 г.

УДК 517.518.8

Ю. Е. ВОСКОБОЙНИКОВ, А. А. МИЦЕЛЬ
(Новосибирск — Томск)

ПОСТРОЕНИЕ УСТОЙЧИВОГО РЕШЕНИЯ ПЛОХО ОБУСЛОВЛЕННОЙ СИСТЕМЫ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ ПРИ СЛУЧАЙНЫХ ПОГРЕШНОСТЯХ В ИСХОДНЫХ ДАННЫХ

Введение. К решению линейных алгебраических систем уравнений приводят многие задачи параметрической идентификации динамических систем, регрессионного анализа, обработки и интерпретации физического эксперимента. В качестве примера можно привести задачу определения концентрации газов по данным лазерного зондирования. Вычислительные трудности, возникающие при решении таких систем, связаны с плохой обусловленностью (а иногда и вырожденностью) системы и характеризуются крайней неустойчивостью полученного обычными методами решения к погрешностям задания правой части. Для получения устойчивого решения часто пользуются методом регуляризации А. Н. Тихонова [1]. Однако при этом возникает известная проблема выбора параметра регуляризации, для решения которой применяется принцип невязки [2, 3].

В данной статье излагается новый подход к решению этой проблемы, являющийся статистическим развитием принципа невязки и позволяющий оценить значение параметра регуляризации, минимизирующее среднеквадратическую ошибку решения.

Заметим, что частный случай такого подхода к выбору параметра регуляризации при решении интегрального уравнения Вольтерра 1 рода рассмотрен в совместной работе одного из авторов [4].

Итак, дана система линейных алгебраических уравнений

$$Ax = y, \tag{1}$$

где A — матрица размерностью $n \times m$, x — m-мерный вектор искомого решения, y — n-мерный вектор правой части. Предположим, что матрица A системы (1) плохо обусловлена или вырождена; вектор решения системы $Ax = 0$ равен нулевому вектору; вместо точной правой части $\bar{y}$ задан вектор измерений $\tilde{y} = \{\tilde{y}_i = y_i + \xi_i\}$, где ξ_i — шум измерения с характеристиками $M[\xi_i] = 0, M\left[\xi_i^2\right] = \sigma_i^2$, $i = \overline{1, n}$, $M[\cdot]$ — оператор математического ожидания. В дальнейшем рассматривается случай, когда корреляционная матрица $V_\xi = M[\xi\xi^\text{т}]$ вектора $\xi = \{\xi_1, \ldots, \xi_n\}$ диагональна, т. е. $V_\xi = \operatorname{diag}\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2\}$.

В качестве устойчивых решений системы (1) будем рассматривать регуляризованное (по А. Н. Тихонову) решение (РР) [1]:

$$x_\alpha = (\alpha I + A^\text{т} A)^{-1} A^\text{т} \tilde{y}, \tag{2}$$

статистически регуляризованное решение (СРР):

$$x_\alpha = \left(\alpha H + A^\text{т} V_\xi^{-1} A\right)^{-1} A^\text{т} V_\xi^{-1} \tilde{y}. \tag{3}$$

Здесь I — единичная $(m \times m)$ матрица, H — симметричная положительно-определенная $(m \times m)$ матрица, α — параметр регуляризации, $A^\text{т}$ — транспонированная матрица A. Заметим, что, во-первых, решение (2) — частный случай (3) при $V_\xi = H = \dot{I}$, во-вторых, в решении (3) информация о шуме измерения (матрица V_ξ) учитывается уже на этапе его построения.

При заданном параметре регуляризации α для любой правой части $\tilde{y}$ существуют решения (2), (3). Однако точность полученных решений зависит от величины α, выбор которой представляет основную трудность при использовании регуляризующих методов.

Критерий выбора параметра регуляризации. Одним из объектов, доступных при построении регуляризованного решения, является вектор невязки $e_\alpha = \tilde{y} - Ax_\alpha$. В ряде работ предложены методы выбора параметра регуляризации, основанные на анализе значений определенных функционалов от вектора невязки: принцип невязки [2], принцип обобщенной невязки [3]. Показано, что эти методы позволяют строить оптимальные по порядку [5] решения, однако они не учитывают стохастической природы невязки, обусловленной случайным шумом измерения.

Введем в рассмотрение оператор E_α, позволяющий представить вектор невязки в виде $e_\alpha = E_\alpha \tilde{y}$ (см. приложение), и матрицу вторых моментов вектора невязки $V_e(\alpha) = M\left[e_\alpha e_\alpha^\text{т}\right]$ и попытаемся определить вид этой матрицы, соответствующий наилучшим (относительно принятого критерия) регуляризованным решениям. В качестве критерия, характеризующего точность построенного решения, примем среднеквадратическую ошибку $\varepsilon(\alpha) = M[\|x_\alpha - x\|^2]$, где $\|\cdot\|$ — евклидова норма в конечномерном векторном пространстве. Значение $\alpha_\text{опт}$, доставляющее минимум $\varepsilon(\alpha)$, будем называть оптимальным параметром регуляризации.

Утверждение 1. Выполнение матричного тождества $V_e(\alpha) = V_\xi E_\alpha^\text{т}$ является достаточным условием оптимальности параметра регуляризации.

Доказательство утверждения приведено в приложении.

Следовательно, для построения регуляризованного решения с наименьшей среднеквадратической ошибкой необходимо выбрать значение параметра регуляризации, статистически не противоречащее гипотезе

$$\Gamma: V_e(\alpha) = V_\xi E_\alpha^\text{т} \tag{4}$$

(в дальнейшем будем обозначать его через α_Γ). Заметим, что в отличие от других методов оценивания оптимальных значений параметра регуляризации критерий (4) не требует количественной априорной информации об искомом векторе решения.

Алгоритм выбора параметра регуляризации. При проверке гипотезы (4) возникает следующее затруднение. Для каждого значения α имеется только одна реализация случайного вектора e_α, и оценка матрицы $V_e(\alpha)$, вычисленная по этой одной реализации, оказывается неудовлетворительной для проверки (4). Поэтому для проверки гипотезы (4) привлекаются статистические методы, использующие статистику $\rho(\alpha) = e_\alpha^\text{т} V_*^{-1} e_\alpha$, где V_* — $(n \times n)$-матрица, определенная ниже для каждого регуляризованного решения (2), (3). Для вычисления значений α_Γ, не противоречащих

(4), удобнее оперировать величиной $\gamma=1/\alpha$ и квадратичной формой $R(\gamma)=\rho(1/\alpha)$. Приведем выражения для матрицы V_*, получаемые из (4):

регуляризованное решение (2) — $V_* = V_\xi\left(I+\alpha^{-1}AA^{\mathrm{T}}\right)$;

статистически регуляризованное решение (3) — $V_* = V_\xi\left(V_\xi + \alpha^{-1}AH^{-1}A^{\mathrm{T}}\right)^{-1}V_\xi$.

Квадратичная форма $R(\gamma)$ и ее производная $R'(\gamma)$, используемые при вычислении параметра α, определяются соотношениями

$$R(\gamma)=\tilde{y}^{\mathrm{T}}V_\xi^{-1}e_\gamma;\quad R'(\gamma)=\tilde{y}^{\mathrm{T}}V_\xi^{-1}e'_\gamma, \tag{5}$$

где $e_\gamma = e_{1/\alpha}$, а вектор $e'_\gamma=\frac{\partial}{\partial\gamma}e_\gamma$ находится из следующих систем уравнений:

для регуляризованного решения (2) $\left(I+\gamma AA^{\mathrm{T}}\right)e'_\gamma = -AA^{\mathrm{T}}e_\gamma$;

для статистически регуляризованного решения (3) $\left(V_\xi+\gamma AH^{-1}A^{\mathrm{T}}\right)\times V_\xi^{-1}e'_\gamma = -AH^{-1}A^{\mathrm{T}}V_\xi^{-1}e_\gamma$.

Утверждение 2. Пусть выполнены следующие условия на структуру корреляционной матрицы V_ξ:

для регуляризованного решения (2) матрица $V_\xi=\sigma^2 I$;

для регуляризованного решения (3) матрица V_ξ симметричная и положительно-определенная.

Тогда квадратичная форма $R(\gamma)$ обладает следующими свойствами: 1) является убывающей выпуклой вниз функцией параметра γ, 2) представляет собой сумму квадратов n случайных величин, 3) при принятии гипотезы (4) математическое ожидание этой квадратичной формы равно n.

Доказательство утверждения приведено в приложении.

Указанные свойства функции $R(\gamma)$ позволяют для вычисления значения α, не противоречащего гипотезе (4), построить итерационный процесс

$$\gamma_{k+1}=\gamma_k-(R(\gamma_k)-n)/R'(\gamma_k),\quad 0<\gamma_0\ll 1. \tag{6}$$

В качестве искомого значения α_Γ принимается величина $1/\gamma_k$, где γ_k удовлетворяет условию $R(\gamma_k)\in\Theta_n(\beta)=[\nu_n(\beta/2),\ \nu_n(1-\beta/2)]$. Величина $\nu_n(\beta/2)$ — $\beta/2$-квантиль χ^2-распределения с n степенями свободы, а β — вероятность ошибки 1-го рода при проверке гипотезы (4) (обычно задается равной 0,05) [6].

Замечание. Значение производной $R'(\gamma)$ может быть оценено через разделенную разность $(R(\gamma+\Delta\gamma)-R(\gamma))/\Delta\gamma$ при достаточно малых $\Delta\gamma$.

Утверждение 3. При выполнении условия $\tilde{y}^{\mathrm{T}}V_\xi^{-1}\tilde{y}\geqslant n$ для любого начального значения $\alpha_0=1/\gamma_0$ итерационная процедура (6) вычисляет значение α_Γ, не противоречащее гипотезе (4).

Доказательство утверждения приведено в приложении.

Следовательно, процедура (6) оценивает оптимальное значение параметра регуляризации, минимизирующее среднеквадратическую ошибку решения. Однако необходимым свойством любого алгоритма выбора параметра регуляризации является существование сходимости $x_\alpha\to x$ при стремлении уровня погрешности исходных данных к нулю.

Утверждение 4. Процедура (6) выбора параметра регуляризации гарантирует сходимость $x_\alpha\to x^+$ при $\mathrm{Sp}[V_\xi]\to 0$, где x^+ — точное псевдорешение системы (1) при точно заданной правой части, $\mathrm{Sp}[V_\xi]$ — след матрицы V_ξ.

Доказательство утверждения приведено в приложении.

Статистические характеристики регуляризованного решения. При интерпретации полученного решения достоверная информация о точности регуляризованного решения не менее важна, чем само решение. Поэтому приведем соотношения, позволяющие оценить некоторые характеристики ошибки регуляризованных решений (2), (3).

Введем m-мерный вектор ошибки решения $\Delta_\alpha=x_\alpha-x$, который можно представить в виде

$$\Delta_\alpha = b_\alpha + \xi_\alpha. \tag{7}$$

Случайный вектор ξ_α имеет нулевое среднее, удовлетворяет системе уравнений $\left(A^{\mathrm{T}}V_\xi^{-1}A + \alpha H\right)\xi_\alpha = A^{\mathrm{T}}V_\xi^{-1}\xi$ и определяет влияние шума измерения на случайную

$$\left(A^{\mathrm{T}}V_\xi^{-1}A + \alpha H\right)b_\alpha = -\alpha H x \tag{8}$$

и характеризует систематическую ошибку (смещение) регуляризованного решения.

Представление (7) позволяет оценить нормы $L_\infty = \max\limits_i |\Delta_\alpha(i)|$, $L_2 = (M[\|\Delta_\alpha\|^2])^{1/2}$ ошибки решения Δ_α следующими соотношениями:

$$L_\infty \approx \max_i \left[\,|b_\alpha(i)| + k\{C_\alpha\}_{ii}^{1/2}\right],$$

$$L_2 \approx \left[\sum_{i=1}^{n} b_\alpha^2(i) + \mathrm{Sp}\,[C_\alpha]\right]^{1/2},$$

где $b_\alpha(i)$, $\Delta_\alpha(i)$ — i-е проекции векторов b_α, Δ_α; k — константа, определяющая величину доверительного интервала для $\xi_\alpha(i)$.

Из-за незнания входящего в правую часть (8) точного решения x возможно только определение оценки $\widehat{b}_\alpha$ вектора b_α при подставлении в (8) вместо x регуляризованного решения x_α. Введение нового регуляризованного решения $\overset{\circ}{x}_\alpha = x_\alpha - \widehat{b}_\alpha$ позволяет частично компенсировать смещение регуляризованного решения (см. результаты численных экспериментов).

Численный эксперимент. Рассмотренные алгоритмы выбора параметра регуляризации реализованы в виде комплекса подпрограмм для решения плохо обусловленных систем алгебраических уравнений.

Для сравнения точности регуляризованных решений моделировалась следующая задача. Определялась концентрация атмосферных газов по известным объемным коэффициентам поглощения, рассчитанным для условий «чистой атмосферы». Матричными элементами системы (1) являлись коэффициенты поглощения на длинах волн CO_2-лазера, вычисленные на единицу массы поглощающих излучение газов. Количество искомых газов m равнялось 6. Для расчета правой части (1) использовались средние значения концентрации этих газов, реализуемые в атмосфере. Число используемых частот (каналов измерения) $n = 16$. Континуальное поглощение H_2O считалось известным. Число обусловленности матрицы $A^{\mathrm{T}}A$ составило $\lambda_{\max}/\lambda_{\min} \approx 2 \cdot 10^8$.

В табл. 1 приведено точное решение x и вычисленное псевдорешение при точно заданной правой части $\bar{x}^+$. Из таблицы видно, что в силу плохой обусловленности системы даже при точно известной правой части отдельные компоненты псевдорешения отличаются от точного решения на 4 порядка.

Таблица 1

x_i	11,3	301,0	0,038	0,0509	0,0295	0,48
$\bar{x}_i^+$	12,51	24,467	368,78	34,987	47,789	28,315

Таблица 2

Точность решения	Погрешность правой части			
	РР		СРР	
	δ=1 %	δ=10 %	δ=1 %	δ=10 %
$S\left(x_{\alpha_{\Gamma}}\right)$	0,05355	0,7125	0,05323	0,7122
$S\left(\overset{\circ}{x}_{\alpha_{\Gamma}}\right)$	0,05293	0,6742	0,05213	0,6656
$S\left(x_{\alpha_{W}}\right)$	0,05483	0,7610	0,05486	0,7696

$$S(x_\alpha) = \left\{\frac{1}{l}\sum\left\{\frac{1}{m}\sum_{i=1}^{m}[(x(i)-x_\alpha(i))/x(i)]^2\right\}\right\}^{1/2},$$

полученных по 10 реализациям $\tilde{y}$ ($l = 10$) для регуляризованных решений (2), (3). В этой же таблице приведены значения S для регуляризованных решений с параметром регуляризации α_W, который соответствует матрице $V_* = V_\xi$, что является простым обобщением принципа невязки на случай стохастических погрешностей правой части системы.

Анализируя приведенные результаты численного эксперимента, можно сделать следующие выводы:

1. Предлагаемый алгоритм выбора параметра регуляризации дает возможность оценить оптимальное значение параметра и повысить точность решения по сравнению с регуляризованным решением, параметр которого определяется по принципу невязки.

2. Введение регуляризованного решения $\overset{\circ}{x}_\alpha$ позволяет частично компенсировать смещение и тем самым повышать точность решения плохо обусловленной системы линейных уравнений.

ПРИЛОЖЕНИЕ

Доказательство утверждения 1. Введя в рассмотрение регуляризующий оператор T_α, позволяющий представить регуляризованное решение x_α в виде $x_\alpha = T_\alpha\tilde{y}$, среднеквадратическую ошибку $\varepsilon(\alpha)$ можно выразить таким образом:

$$\varepsilon(\alpha) = \mathrm{Sp}\left[T_\alpha V_{\tilde{y}} T_\alpha^{\mathrm{T}} - 2V A_x^{\mathrm{T}} T_\alpha^{\mathrm{T}} + V_x\right], \tag{П1}$$

где $V_{\tilde{y}} = M\left[\tilde{y}\,\tilde{y}^{\mathrm{T}}\right]$, $V_x = M[xx^{\mathrm{T}}]$ — матрицы вторых моментов векторов $\tilde{y}$, x. Достаточное условие минимума (П1) имеет вид

$$T_\alpha V_{\tilde{y}} = V_x A^{\mathrm{T}}. \tag{П2}$$

Однако из-за незнания V_x это тождество не может быть использовано для вычисления оптимального параметра регуляризации. Преодолеть возникшее затруднение можно следующим путем.

Определим матрицу $V_e(\alpha)$ вторых моментов вектора невязки при $\alpha = \alpha_{\text{опт}}$. Для этого в определение матрицы $V_e(\alpha) = E_\alpha V_{\tilde{y}} E_\alpha^{\mathrm{T}}$ ($E_\alpha = I - AT_\alpha$ — оператор невязки) подставим выражение $E_\alpha V_{\tilde{y}} = V_{\tilde{y}} - AT_\alpha V_{\tilde{y}}$, которое при выполнении (П2) обращается в тождество $E_{\alpha_{\text{опт}}} V_{\tilde{y}} = V_{\tilde{y}} - AV_x A^{\mathrm{T}} = V_\xi$. Тогда следует, что если выполняется $V_e(\alpha) = V_\xi E_\alpha^{\mathrm{T}}$, то $\alpha = \alpha_{\text{опт}}$.

Доказательство утверждения 2. При доказательстве будем пользоваться приемом [7], который для конкретности продемонстрируем на функции $R(\gamma)$ регуляризованного решения (3):

$$R(\gamma)=e_\gamma^{\mathrm{T}}V_\xi^{-1}\left(V_\xi+\gamma AH^{-1}A^{\mathrm{T}}\right)V_\xi^{-1}e_\gamma. \tag{П3}$$

Можно показать, что $e_\gamma=V_\xi(V_\xi+\gamma AH^{-1}A^{\mathrm{T}})^{-1}\tilde{y}$. Матрица $AH^{-1}A^{\mathrm{T}}$ является симметричной $(n\times n)$ положительно-определенной матрицей, и ее можно представить как $V_\xi^{-1/2}AH^{-1}A^{\mathrm{T}}V_\xi^{-1/2}=B\Lambda B^{\mathrm{T}}$, где $\Lambda=\mathrm{diag}\{\lambda_1,\ldots,\lambda_n\}$, $\lambda_i\geqslant 0$, B, B^{T} — ортогональные матрицы соответствующих размерностей. Тогда из (П3) следует:

$$R(\gamma)=\sum_{i=1}^{n}\frac{\{\tilde{f}\}_i^2}{1+\gamma\lambda_i},\quad 0\leqslant\gamma<\infty,\ \text{где}\ \tilde{f}=BV_\xi^{-1/2}\tilde{y}.$$

Непосредственным дифференцированием этого выражения убеждаемся в том, что $R'(\gamma)<0$, $R''(\gamma)>0$ для всех γ: $0\leqslant\gamma<\infty$. Для завершения доказательства достаточно в (П3) подставить значение $\gamma=\gamma_{\text{Г}}$, не противоречащее гипотезе (4). Тогда

$$M[R(\gamma)]=\mathrm{Sp}\left[V_\xi^{-1}\left(V_\xi+\gamma_{\text{Г}}AH^{-1}A^{\mathrm{T}}\right)V_\xi^{-1}M\left[e_{\gamma_{\text{Г}}}e_{\gamma_{\text{Г}}}^{\mathrm{T}}\right]\right]=\mathrm{Sp}\,(I)=n.$$

Доказательство утверждения 3. Из определения функции $R(\gamma)$ следуют предельные значения $R(0)=\sum_{i=1}^{n}\tilde{y}_i^2/\sigma_i^2$, $R(\infty)=0$. При выполнении условия утверждения $\tilde{y}^{\mathrm{T}}V_\xi^{-1}\tilde{y}\geqslant n$ в силу монотонности убывания функции $R(\gamma)$ обязательно существует такое $\gamma_{\text{Г}}:0\leqslant\gamma_{\text{Г}}<\infty$, для которого $R(\gamma_{\text{Г}})\in\Theta_n(\beta)$.

Доказательство утверждения 4. В качестве решения системы (1) при точно заданной первой части $\bar{y}$ примем точное псевдорешение [5], определяемое как $x^+=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\bar{y}$. Если матрица $(A^{\mathrm{T}}A)^{-1}$ существует, то $x^+=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}Ax=x$. Так как решения (2), (3) являются регуляризующими, то для сходимости $x_\alpha\to x^+$ при $\mathrm{Sp}[V_\xi]\to 0$ достаточно показать, что $\|Ax_\alpha-Ax\|\to 0$ при $\mathrm{Sp}[V_\xi]\to 0$ [5]. Из определения x^+ следует, что

$$\|Ax^+-Ax_\alpha\|\leqslant\|\bar{y}-\tilde{y}\|+\|\tilde{y}-Ax_\alpha\|=\|\xi\|+\|e_\alpha\|. \tag{П4}$$

Первое слагаемое стремится к нулю при $\mathrm{Sp}[V_\xi]\to 0$. Для анализа поведения второго слагаемого запишем квадратическую форму $R(\gamma)$ в виде

$$R(\gamma)=\sum_{i=1}^{n}e_\alpha(i)\,\tilde{y}_i/\sigma_i^2.$$

Так как в качестве $\alpha_{\text{Г}}$ принимается значение, для которого $R(1/\alpha_{\text{Г}})\in\Theta_n(\beta)$, и этот интервал не зависит от уровня шума, то вектор e_α покоординатно стремится к нулю при $\sigma_i^2\to 0$, $i=\overline{1,n}$. Таким образом, и второе слагаемое в (П4) стремится к нулю при $\mathrm{Sp}[V_\xi]\to 0$.

ЛИТЕРАТУРА

1. **Тихонов А. Н., Арсенин В. Я.** Методы решения некорректных задач.— М.: Наука, 1974.
2. **Морозов В. А.** О принципе невязки при решении операторных уравнений методом регуляризации.— Журн. вычисл. математики и мат. физики, 1968, т. 8, № 2.
3. **Гончарский А. В., Леонов А. С., Ягола А. Г.** Обобщенный принцип невязки.— Журн. вычисл. математики и мат. физики, 1973, т. 13, № 2.
4. **Воскобойников Ю. Е., Томсонс Я. Я.** Выбор параметра регуляризации и ошибки восстановления входного сигнала в методе статистической регуляризации.— Автометрия, 1975, № 4.
5. **Иванов В. К., Васин В. В., Танана В. П.** Теория линейных некорректных задач и ее приложения.— М.: Наука, 1978.
6. **Шметтерер Л.** Введение в математическую статистику.— М.: Наука, 1976.
7. **Гордонова В. И., Морозов В. А.** Численные алгоритмы выбора параметра в методе регуляризации.— Журн. вычисл. математики и мат. физики, 1973, т. 13, № 3.

Поступила в редакцию 24 июля 1980 г.;
окончательный вариант — 26 января 1981 г.