УДК 519.677

Ю. Ю. Гончаренко
Севастопольский национальный университет ядерной энергии и промышленности
ул. Курчатова, 7, 99026 Севастополь, Украина

Математическая модель описания некорректно поставленных задач

Идентификация образов в ряде случаев требует перебора множества моделей и созданных на их основе алгоритмов, что относится к классу некорректно поставленных задач. Показано, что математическое описание некорректно поставленных задач определяется семейством элементов, заданных в гильбертовых пространствах, для которых ограниченный линейный оператор определяет приближенное решение, связанное с точным, с последующей регуляризацией, осуществляемой одним из приведенных методов, что на практике требует добавления априорных классификационных признаков.

Ключевые слова: *математическая модель, гильбертово пространство, некорректно поставленная задача, априорные признаки.*

Введение

Большинство задач, возникающих перед исследователями, приходится решать при дефиците исходных данных в процессе принятия решения. К такому типу относятся задачи поиска объектов с использованием радиоэлектронных систем различного назначения [1], определения качества оценивания и прогнозирования ресурсных характеристик сложных технических объектов [2], охраны окружающей среды и рационального использования природных ресурсов [3] и многие другие. Целью этих задач является идентификация образов и определение их параметров в результате набора выполненных измерений. Вид измерений (электромагнитные, акустические, радиационные и др.), результат измерений, их численность в группе, количество групп зависит от множества факторов, учесть которые не представляется возможным. В результате приходиться использовать не один, а достаточно большое число моделей и созданных на их базе алгоритмов. Такого рода задачи часто называют некорректными или некорректно поставленными. Всех их, несмотря на различие предметных областей, может объединять возможность описания с помощью универсального математического аппарата. Известны разработки ряда ученых, таких как Андерсон [4], Адамар [5], Иосида [6], в которых рассматриваются некоторые аспекты математического описания некорректно поставленных задач. Тем не менее, до сих пор поиск оптимального решения с заданной степенью точности не теряет своей актуальности, как в прикладных, так и фундаментальных исследованиях. Одним из направлений решения этих задач — использование приемов регуляризации, которые применяются главным образом тогда, когда обучение алгоритма оказывается неустойчивым.

Постановка цели и задач научного исследования

Целью данной работы является разработка математической модели описания некорректно поставленных задач и определение методов их регуляризации.

Для достижения поставленной цели необходимо решить следующие научные задачи. Во-первых, систематизировать данные и сформулировать определение, привести математическое описание и классификацию некорректно поставленных задач, во-вторых, рассмотреть методы ее регуляризации.

Математическая модель описания некорректно поставленных задач

Известно, что гильбертово пространство — это обобщение эвклидова пространства, допускающее бесконечную размерность. Гильбертово пространство — линейное (векторное) пространство (над полем вещественных или комплексных чисел), в котором для двух любых элементов x и y определено скалярное произведение (x, y) и полное относительно порожденной скалярным произведением метрики $d(x, y) = \|x - y\| = \sqrt{(x - y, x - y)}$.

Проведем анализ линейной некорректно поставленной задачи в гильбертовых пространствах H и K и действующий из H в K линейный ограниченный оператор A, при этом для известного $g \in K$ требуется найти $f \in H$, для которого выполняется условие:

$$Af = g. \tag{1}$$

Задача (1) называется корректно поставленной по Адамару, если для любого $g \in K$ она имеет единственное решение, непрерывно зависящее от g, в противном случае — некорректной. Для некорректно поставленной задачи оператор A^{-1} или не существует, или определен не на всем пространстве K, или не является непрерывным, при этом сложность решения таких задач состоит в том, что решение уравнения (1), если оно существует, не обязательно может быть близким к решению уравнения $Af = g^{\varepsilon}$, где g^{ε} близко к g.

Исследуем вопросы, связанные с некорректностью, для чего определим элемент, заменяющий решение задачи в случае его отсутствия. Если $g \in \operatorname{Im} A + (\operatorname{Im} A)^{\perp}$, рассмотрим элемент, минимизирующий норму $\|Af - g\|$, т.е. из всех возможных таких элементов, выберем с минимальной нормой. Обозначим корректно определенный таким образом элемент гильбертова пространства K, называемый обобщенным решением Мура-Пенроуза задачи (1), через $A^{+}g$, и рассмотрим оператор $A^+ : K \to H$, определенный на $\operatorname{Im} A + (\operatorname{Im} A)^\perp$, называемый обобщенным обратным оператором Мура-Пенроуза.

Допустим, что элемент $f = A^+ g$ является решением уравнения

$$A^* A f = A^* g, \tag{2}$$

принадлежащим $\overline{\operatorname{Im} A^*}$. Элемент f минимизирует $\|Af - g\|$ тогда и только тогда, когда

$$(Af - g, Au) = 0 \quad \forall\, u \in H, \tag{3}$$

т.е. при выполнении условия (2), называемого системой нормальных уравнений. Единственный элемент с минимальной нормой из множества решений (3) принадлежит $(Ker\ A)^\perp = \overline{\operatorname{Im} A^*}$.

В общем случае оператор A^+ не является непрерывным. Для перехода к непрерывной задаче введем понятие его регуляризации.

Рассмотрим определенное на пространстве K семейство $(T_\gamma)_{\gamma>0}$ непрерывных линейных операторов $T_\gamma : K \to H$, для которого на всей области определения оператора A^+ выполняется условие:

$$\lim_{\gamma \to 0} T_\gamma g = A^+ g. \tag{4}$$

Если оператор A^+ не ограничен, то $\|T_\gamma\| \to \infty$ при $g \to 0$.

Для отыскания приближенного решения задачи (1) с помощью регуляризации допустим, что $g^\varepsilon \in K$ аппроксимирует g, т.е. $\|g^\varepsilon - g\| \le \varepsilon$. Пусть при этом функция $\gamma(\varepsilon)$ удовлетворяет условию, что при $\varepsilon \to 0$

$$\gamma(\varepsilon) \to 0 \text{ и } \|T_{\gamma(\varepsilon)}\| \varepsilon \to 0. \tag{5}$$

Тогда при $\varepsilon \to 0$:

$$\|T_{\gamma(\varepsilon)} g^\varepsilon - A^+ g\| \le \|T_{\gamma(\varepsilon)}(g^\varepsilon - g)\| + \|T_{\gamma(\varepsilon)} g - A^+ g\| \le \|T_{\gamma(\varepsilon)}\| \varepsilon + \|T_{\gamma(\varepsilon)} g - A^+ g\| \to 0. \tag{6}$$

Значит, если элемент g^ε близок к g, то элемент $T_{\gamma(\varepsilon)} g^\varepsilon$ близок к $A^+ g$. Здесь число γ, от выбора которого зависит качество регуляризации, называется параметром регуляризации. При необходимости его значение можно подобрать эмпирически.

Методы регуляризации некорректно поставленных задач

Рассмотрим различные методы регуляризации: усеченное сингулярное разложение, метод Тихонова-Филлипса, итерационные методы.

Сингулярным разложением оператора A называется выражение

$$Af = \sum_{k=1}^{\infty} \sigma_k \left(f, f_k\right) g_k , \tag{7}$$

где $f_k \in H$; $g_k \in K$ — нормированные ортогональные системы; $\sigma_k > 0$ — сингулярные числа оператора A.

Допустим, что последовательность $\{\sigma_k\}$ ограниченная, тогда A — непрерывный линейный оператор $H \to K$ с сопряженным оператором

$$Ag = \sum_{k=1}^{\infty} \sigma_k \left(g, g_k\right) f_k \tag{8}$$

и самосопряженными операторами в H и K соответственно:

$$A^* Af = \sum_{k=1}^{\infty} \sigma^2{}_k \left(f, f_k\right) f_k , \tag{9}$$

$$AA^* g = \sum_{k=1}^{\infty} \sigma^2{}_k \left(g, g_k\right) g_k . \tag{10}$$

Спектр оператора (9) содержит собственные значения $\sigma^2{}_k$, которым соответствуют собственные элементы f_k и, может быть 0, имеющий бесконечную кратность. Аналогично для оператора (10) с собственными элементами g_k, при этом справедливы соотношения:

$$A^* g_k = \sigma_k f_k ,\ A^* f_k = \sigma_k g_k . \tag{11}$$

Если при этом f_k и g_k — нормированные системы собственных элементов операторов $A^* A$ и AA^*, удовлетворяющие условиям (11), то оператор A представим в виде (7), т.е. допускается сингулярное разложение, тогда:

$$A^+ g = \sum_{k=1}^{\infty} \sigma_k{}^{-1} \left(g, g_k\right) f_k . \tag{12}$$

Ряд (12) сходится для $g \in \operatorname{Im} A + \left(\operatorname{Im} A\right)^{\perp}$ из области определения оператора A^+. Если $g = Av + u$, где $u \in \left(\operatorname{Im} A\right)^{\perp}$, то $\left(g, g_k\right) = \left(Av + u, g_k\right) = \left(Av, g_k\right) + \left(u, g_k\right) = \left(v, A^* g_k\right) + \left(u, g_k\right)$. Из (11) следует, что $g_k \in \operatorname{Im} A$ и $A^* g_k = \sigma_k f_k$, откуда

$$\left(g, g_k\right) = \sigma_k \left(v, f_k\right), \tag{13}$$

и сходится ряд:

$$f^{+} = \sum_{k=1}^{\infty} \sigma_k^{-1} \left(g, g_k\right) f_k = \sum_{k=1}^{\infty} \left(v, f_k\right) f_k . \tag{14}$$

Применяя оператор $A^* A$ к (14), получим:

$$A^* A f^{+} = \sum_{k=1}^{\infty} \sigma_k^{-1} \left(g, g_k\right) A^* A f_k = \sum_{k=1}^{\infty} \sigma_k \left(g, g_k\right) f_k = A^* g . \tag{15}$$

Так как в соответствии с (11) $f^{+} \in \overline{\operatorname{Im} A^*}$, то $f^{+} = A^{+} g$. Получили, что оператор A^{+} не ограничен, когда $\sigma_{k_j} \to 0$ для некоторой последовательности $k_j \to \infty$. В этом случае оператор A^{+} можно регуляризовать с помощью усеченного сингулярного разложения:

$$T_\gamma g = \sum_{k \le \frac{1}{\gamma}} \sigma_k^{-1} \left(g, g_k\right) f_k . \tag{16}$$

При этом $T_\gamma g \to A^{+} g$ при $\gamma \to 0$ и

$$\left\| T_\gamma \right\| \le \sup_{k \le \frac{1}{\gamma}} \sigma_k^{-1} . \tag{17}$$

В общем виде

$$T_\gamma g = \sum_{k=1}^{\infty} F_\gamma \left(\sigma_k\right) \left(g, g_k\right) f_k , \tag{18}$$

где $F_\gamma(\sigma)$ аппроксимирует σ^{-1} при больших σ и стремится к нулю при $\sigma \to 0$. Здесь функция F_γ называется фильтром, а методы регуляризации по формуле (18) называются цифровой фильтрацией.

Применение сингулярного разложения позволяет лучше понять природу некорректности. Так, если g^ε аппроксимирует g, то $\left\| g - g^\varepsilon \right\| \le \varepsilon$. Если известно g^ε, то для коэффициентов разложения $A^{+} g$ справедливо:

$$\left| \sigma_k^{-1} \left(g, g_k\right) - \sigma_k^{-1} \left(g^\varepsilon, g_k\right) \right| \le \frac{\varepsilon}{\sigma_k} . \tag{19}$$

Отсюда следует, что значение g_k для малых σ_k нельзя вычислить с достаточной степенью точности. В связи с этим, если известны сингулярные числа и соответствующие им элементы g_k, то можно судить о том, какие компоненты решения f уравнения (1) вычисляются по приближению g^ε элемента g, а какие нет.

Метод Тихонова-Филлипса.

Допустим, что

$$T_\gamma = \left(A^* A + \gamma I\right)^{-1} A^* = A^* \left(A^* A + \gamma I\right)^{-1}, \tag{20}$$

что равносильно характеристике $f_\gamma = T_\gamma g$ как элемента, минимизирующего величину $\|Af - g\|^2 + \gamma \|f\|^2$. Покажем, что $f_\gamma = T_\gamma g \to f^+ = A^+ g$ при $\gamma \to 0$. В соответствии с (2)

$$A^* Af = A^* g, \quad f^+ \perp Ker\ A \tag{21}$$

справедливо

$$\left(A^* A + \gamma I\right)\left(f^+ - f_\gamma\right) = \gamma f^+. \tag{22}$$

Пусть E_λ — разложение единицы, связанное с самосопряженным оператором $A^* A$ соотношениями:

$$A^* A = \int \lambda dE_\lambda, \quad \left(A^* A + \gamma I\right)^{-1} = \int (\lambda + \gamma)^{-1} dE_\lambda. \tag{23}$$

Здесь $E_\lambda = 0$ при $\lambda < 0$. В соответствии с (22) получается:

$$f^+ - f_\gamma = \int \frac{\gamma}{\lambda + \gamma} dE_\lambda f^+. \tag{24}$$

Отсюда:

$$\begin{gathered} \left\|f^+ - f_\gamma\right\|^2 = \int \left(\frac{\gamma}{\lambda + \gamma}\right)^2 d\left\|E_\lambda f^+\right\|^2 = \int_{-\infty}^{\sqrt{\gamma}} \left(\frac{\gamma}{\lambda + \gamma}\right)^2 d\left\|E_\lambda f^+\right\|^2 + \int_{\sqrt{\gamma}}^{\infty} \left(\frac{\gamma}{\lambda + \gamma}\right)^2 d\left\|E_\lambda f^+\right\|^2 \le \\ \le \int_{-\infty}^{\sqrt{\gamma}} d\left\|E_\lambda f^+\right\|^2 + \gamma \int_{\sqrt{\gamma}}^{\infty} d\left\|E_\lambda f^+\right\|^2 = \left\|E_{\sqrt{\gamma}} f^+\right\|^2 + \gamma\left(\left\|f^+\right\|^2 - \left\|E_{\sqrt{\gamma}} f^+\right\|^2\right). \end{gathered} \tag{25}$$

Для оператора ортогонального проектирования на $Ker\left(A^* A\right) = Ker\ A$ при условии $f^+ \perp Ker\ A$ выполняется равенство:

$$\lim_{\gamma \to 0} \left\|E_{\sqrt{\gamma}} f^+\right\|^2 = \left\|P f^+\right\|^2 = 0, \tag{26}$$

при этом с учетом (25) получается $f_\gamma \to f^+$ при $\gamma \to 0$.

Если оператор A допускает сингулярное разложение (7), то f_γ имеет вид (18), где:

$$F_\gamma(\sigma) = \frac{1}{1+\frac{\gamma}{\sigma^2}}\frac{1}{\sigma}. \tag{27}$$

Таким образом, метод Тихонова-Филлипса является частным случаем цифровой фильтрации.

Итерационные методы.

Допустим, что имеется итерационный метод решения уравнения (1) в виде

$$f^{k+1} = B_k f^k + C_k g\,,\ \ k = 0,1,..., \tag{28}$$

где B_k, C_k — непрерывные линейные операторы.

Пусть $f^k \to A^+ g$. Обозначим индекс, стремящийся к бесконечности при $\gamma \to 0$ через $k(\gamma)$, $\gamma > 0$, при этом регуляризация оператора A^+ представится в виде:

$$T_\gamma f = f^{k(\gamma)}. \tag{29}$$

В этом случае выбор соответствующего параметра регуляризации γ сводится к выбору числа шагов итерации. Итерационный метод (28) не всегда сходится для g из области определения оператора A^+, иногда возможна не полная сходимость, когда на некотором шаге метод определяет удовлетворительное решение, а потом расходится.

Рассмотренные методы регуляризации дают возможность решения некорректной задачи приближенными методами. Для выяснения степени точности сделаем дополнительные допущения для оператора A и точного решения f с использованием следующей модели.

Пусть $H = L_2(\Omega^n)$, где Ω^n — единичный шар в R^n. Допустим, что:

1) $\exists \alpha > 0$:

$$m\|f\|_{H_0^{-\alpha}(\Omega^n)} \le \|Af\|_K \le M\|f\|_{H_0^{-\alpha}(\Omega^n)}, \tag{30}$$

где m и M — положительные константы;

2) точное решение $f \in H_0^\beta(\Omega^n)$, где число $\beta > 0$ и выполняется условие:

$$\|f\|_{H_0^\beta(\Omega^n)} \le \rho\,. \tag{31}$$

В соответствии с (30) существует оператор $A^{-1} : K \to L_2(\Omega^n)$, но не ограничен, т.е. уравнение (1) некорректно для пространств $L_2(\Omega^n)$ и K, а число α определяет меру некорректности. Неравенство (31) — условие гладкости точного решения.

Пусть известен не сам элемент g, а его приближение g^ε : $\left\| g - g^\varepsilon \right\| \le \varepsilon$, тогда вся информация о точном решении описывается неравенствами:

$$\left\| Af - g^\varepsilon \right\|_K \le \varepsilon, \ \ \|f\|_{H_0^\beta(\Omega^n)} \le \rho. \tag{32}$$

Так, если решения f_1, f_2 удовлетворяют (32), то:

$$\left\| A(f_1 - f_2) \right\|_K \le 2\varepsilon, \ \ \|f_1 - f\|_{H_0^\beta(\Omega^n)} \le 2\rho. \tag{33}$$

Введем обозначения

$$d(\varepsilon, \rho) = \sup\left\{ \|f\|_{L_2(\Omega^n)} : \|Af\|_K \le \varepsilon, \quad \|f\|_{H_0^\beta(\Omega^n)} \le \rho \right\} \tag{34}$$

и получим

$$\|f_1 - f_2\|_{L_2(\Omega^n)} \le d(2\varepsilon, 2\rho), \tag{35}$$

где правая часть (35) может рассматриваться как наиболее возможная погрешность восстановления f по неточным данным g^ε при выполнении допущений (30), (31). Сходимость погрешности $d(2\varepsilon, 2\rho)$ при $\varepsilon \to 0$ при условии (30) определяется существованием такой постоянной $c(m, \alpha, \beta)$, что

$$d(\varepsilon, \rho) \le c(m, \alpha, \beta) \varepsilon^{\frac{\beta}{\alpha+\beta}} \rho^{\frac{\alpha}{\alpha+\beta}}. \tag{36}$$

Положив в неравенстве (30) $\gamma = 0$, $\alpha = -\alpha$, получим:

$$\|f\|_{L_2(\Omega^n)} \le \|f\|_{H_0^{-\alpha}(\Omega^n)}^{\frac{\beta}{\beta+\alpha}} \|f\|_{H_0^{\alpha}(\Omega^n)}^{\frac{\alpha}{\beta+\alpha}}. \tag{37}$$

Если при этом $\|Af\| \le \varepsilon$, то с учетом (30)

$$\|f\|_{H_0^{-\alpha}(\Omega^n)} \le \frac{\varepsilon}{m}. \tag{38}$$

Следовательно, если $\|Af\| \le \varepsilon$ и $\|f\|_{H_0^\beta(\Omega^n)} \le \rho$, то

$$\|f\|_{L_2(\Omega^n)} \le \left(\frac{\varepsilon}{m}\right)^{\frac{\beta}{\alpha+\beta}} \rho^{\frac{\alpha}{\alpha+\beta}}. \tag{39}$$

Правое неравенство (30) не является существенным, но оно гарантирует неулучшаемость экспоненциальной оценки. Если в некорректной задаче, удовлетворяющей условиям (30), (31), исходные данные имеют погрешность измерения ε, то решение можно найти с точностью $O\left(\varepsilon^{\frac{\beta}{\alpha+\beta}}\right)$. Так как $\frac{\beta}{\alpha+\beta} < 1$, то погрешность решения всегда больше ε, т.е. происходит потеря точности. В зависимости от соотношений этих величин возможны случаи:

1) если величина $\frac{\beta}{\alpha+\beta}$ близка к 0, то задача считается сильно некорректной, при этом потеря точности катастрофична. Как правило, $\alpha = \infty$, т.е. условие (30) не выполняется;

2) если величина $\frac{\beta}{\alpha+\beta}$ близка к 1, задача считается слабо некорректной, и потеря точности невелика;

3) если величина $\frac{\beta}{\alpha+\beta}$ не близка ни к 0, ни к 1, задача считается умеренно некорректной.

Приведенная выше классификация значительно упрощает оценку степени некорректности задач, но она зависит не только от свойств оператора A, но и от гладкости точного решения. Кроме того, к некорректно поставленным задачам возможен и статистический подход.

Рассмотрим f и g как совокупности случайных величин с совместным нормальным распределением, средними $\overline{f}$ и $\overline{g}$ и ковариациями F и G соответственно, т.е.

$$F(x,x') = E\left(f(x) - \overline{f}(x)\right)\left(f(x') - \overline{f}(x')\right), \tag{40}$$

где E — математическое ожидание. Ковариация G задается аналогично.

Допустим, что f и g связаны линейным соотношением

$$g = Af + n, \tag{41}$$

где A — линейный оператор; n — нормально распределенная совокупность случайных величин (ошибок измерений) со средним 0 и ковариацией Σ.

Пусть f и n независимы, тогда (f, g) имеет среднее $\left(\overline{f}, \overline{g}\right)$ и ковариацию

$$\begin{pmatrix} F & AF^* \\ AF & AFA^* + \Sigma \end{pmatrix}. \tag{42}$$

Примем решением f_B уравнения (41) условное математическое ожидание f при заданном g. Аналогично с конечномерным вариантом имеем:

$$f_B = E(f \mid g) = \overline{f} + FA^*\left(AFA^* + \Sigma\right)^{-1}\left(g - A\overline{f}\right). \qquad (43)$$

Здесь (43) представляет собой наилучшую линейную среднеквадратичную оценку для f. При $\overline{f} = 0$, $F = I$, $\Sigma = \gamma I$ выражение (43) совпадает с формулой (20) регуляризации Тихонова-Филлипса. Выбор значения Σ соответствует «белому шуму», адекватно описывающему ошибки измерений. Выбор значения F отражает свойства f. В цифровой обработке изображений для плотности изображения f используется изотропная экспоненциальная модель

$$F\left(x, x'\right) = \sigma e^{-\lambda|x-x'|}, \qquad (44)$$

где $\sigma, \lambda > 0$. При этом допускается, что соотношение уровней окрашенности изображения в точках x и x', характеризующееся ковариацией распределений $f(x)$ и $f(x')$, зависит исключительно от расстояния между точками x и x' и экспоненциально убывает с увеличением этого расстояния.

Выводы

Таким образом, анализ представленной математической модели показывает, что при решении некорректно поставленных задач идентификации объектов (акустических сигналов, человеческой речи и др.) необходимо к множеству регистрируемых параметров добавить априорные классификационные признаки (хотя бы один), характеризующие объект поиска, которые позволят однозначно его опознать.

1. *Гринченко В.Т.* Волновые задачи акустики / В.Т. Гринченко, И.В. Вовк, В.Т. Маципура. — К.: Интерсервис, 2013. — 572 с.

2. *Маловик К.Н.* Развитие научных основ повышения качества оценивания и прогнозирования ресурсных характеристик сложных объектов / К.Н. Маловик. — Севастополь: СНУЯЭиП, 2013. — 332 с.

3. *Кудрик И.Д.* Экологический мониторинг курортно-туристических ресурсов Крыма / И.Д. Кудрик, Н.И. Ковалев, С.Г. Белявский [и др.]. — Севастополь: Изд-во «Черкасский ЦНТЭИ», 2013. — 257 с.

4. *Андерсон Т.У.* Введение в многомерный статистический анализ / Т.У. Андерсон. — М.: Физматгиз, 1963. — 526 с.

5. *Адамар Ж.* Задача Коши для линейных уравнений с частными производными гиперболического типа / Ж. Адамар. — М.: Наука, 1978. — 320 с.

6. *Иосида К.* Функциональный анализ / К. Иосида. — М.: Мир, 1967. — 620 с.

Поступила в редакцию 27.05.2014