Доклады Международной конференции “Математическая биология и биоинформатика”. Под ред. В.Д. Лахно. Том 8. Пущино: ИМПБ РАН, 2020. Статья № e18. doi: 10.17537/icmbb20.14

Операторный логарифм и вариационные методы в линейных обратных задачах

Дроздовский С.А.

НИЦ Курчатовский институт – ИТЭФ

Stanislav.Drozdovsky@itep.ru

Разработан общий метод решения линейной обратной задачи в рамках вариационного подхода, использующий такие понятия операторного исчисления как логарифм и полугруппа, применимый к широкому классу задач, имеющих приложения в обработке экспериментальных данных. Являясь по сути обобщением метода наискорейшего (градиентного) спуска, в отличие от него новый метод сходится на порядки быстрее. Альтернативные быстро сходящиеся способы как правило обобщают в том или ином виде метод Ньютона. Они вычислительно затратны и предполагают обращение больших нерегулярных матриц на каждой итерации, тогда как предложенный метод использует исключительно прямые вычисления. Кроме того, алгоритм на нижнем уровне сводится к решению дифференциального уравнения с частными производными, что позволяет в численной реализации использовать стандартные модули, решающие дифференциальные уравнения. Полученный результат применим для реконструкции объемной плотности объектов, в частности в протонной и рентгеновской радиографии.

Ключевые слова: линейная обратная задача, регуляризация, градиентный спуск, дробный интеграл, операторные логарифм и полугруппа.

Logarithm of Operator and Variational Methods for Linear Inverse Problems

Drozdovskiy S.A.

NRC Kurchatov Institute – ITEP

A general method is developed to solve the linear inverse problem in the variational framework to apply to a wide range of tasks in experimental data processing. The approach involves such concepts of operator theory as logarithm and semigroup. The method is actually a generalization of the steepest descent method, but converges orders faster. Most of alternative fast ways somehow modify the Newton's method; they are numerically expensive and use the conversion of huge matrixes in each iteration. The new method uses direct calculations only. Besides, the bottom level of the algorithm is reduced to a partial differential equation, which has many standard options available to calculate. This approach can be used in proton and x-ray radiography researches especially in the volumetric density reconstruction.

Key words: linear inverse problem, regularization, steepest descend, fractional calculus, operator logarithm and semigroup.

1. Введение

Обратной задачей называется уравнение $g = K f$ с неизвестным f, где K – оператор, описывающий модель системы и связь между наблюдаемыми данными g и искомыми параметрами системы f. Обычно f и g являются пространственными распределениями некоторых величин.

Даже если формально решение существует и единственно, проблема может заключаться в том, что непосредственное решение неустойчиво, т.е. подвержено патологическим изменениям при небольших отклонениях исходных данных и их производных. Такие задачи называются некорректно поставленными (ill-posed). На практике это означает, что обычные искажения данных, такие как шум, на выходе дают результат, не пригодный к использованию.

Для эффективного решения проблемы используется т.н. *регуляризация*, т. е. вводятся дополнительные ограничения, устраняющие неустойчивость при соблюдении приемлемого правдоподобия.

Вариационный метод [1] заключается в минимизации некоторого функционала $F(K, f, g)$ относительно переменной f.

Обобщенная регуляризация Тихонова использует функционал вида $F_\sigma(K, f, g) = G(K f, g) + \sigma J(f)$, где первое слагаемое отвечает за близость к точному решению, а второе – за регулярность, $\sigma > 0$ является параметром регуляризации. Для f, g определенных на некоторой области в R^n типичный функционал такого рода имеет вид:

$$G(Kf, g) = \int \left| Kf(x) - g(x) \right|^p dx,$$

$$J(f) = \int \left(\sum_i \left(\frac{\partial f}{\partial x_i} \right)^2 \right)^{q/2} dx, \qquad (1)$$

$$1 \le p, q \le 2.$$

Далее мы предполагаем, что оператор K действует в банаховом пространстве, линеен, ограничен и невырожден (не содержит 0 в спектре) [2]. Перечисленные ограничения обусловлены постановкой задачи, а также тем обстоятельством что в реальных вычислениях используется дискретное конечномерное приближение. Кроме того, если оператор вырожден, предполагая дискретность спектра, его можно как угодно близко аппроксимировать невырожденным, поскольку имеет место регуляризованное решение, приближенное по определению.

2. Операторные логарифм и полугруппа

Экспонента оператора K определяется как абсолютно сходящийся ряд

$$\exp K = I + K + K^2 / 2! + \cdots + K^n / n! + \cdots, \qquad (2)$$

где I обозначает тождественное преобразование. Соответственно, *логарифм* $\ln K$ определяется как обратный экспоненте.

При сделанных относительно K предположениях, логарифм существует, ограничен и невырожден. Его можно вычислить, используя различные формулы.

$$\ln K = \int_0^1 f'(t) \left[K - I \right] \left[f(t) K + \left(1 - f(t) \right) I \right]^{-1} dt,$$

[3], где $f : \mathrm{R} \to \mathrm{C}$ непрерывно дифференцируемая функция, $f(0) = 0 = f^{-1}(0)$, $f(1) = 1$ и область значений функции $(f(t) - 1) / f(t)$ принадлежит резольвентному множеству оператора K. Если спектр оператора K полностью лежит в положительной полуплоскости, то логарифм выражается пределом условно сходящихся рядов:

$$\ln K = \lim_{t \to \infty} -(\gamma + \ln t) I - \sum_{n=1}^{\infty} \frac{(-t)^n}{n! \, n} K^n,$$

где γ – константа Эйлера. На самом деле наличие в резольвентном множестве луча исходящего из 0 гарантирует существование дробных степеней оператора [4], что позволяет возведением в степень и умножением на комплексное число привести его свойства к последнему случаю.

Полугруппой (однопараметрической) [2, 5] называется семейство $\{K_t : 0 \le t < \infty\}$ линейных ограниченных операторов, обладающее свойствами: $K_0 = I$ и $K_{s+t} = K_s K_t$ для всех $s, t \ge 0$. Если функция $K_t x$ непрерывна по t для каждого x, то полугруппа называется *сильно непрерывной*, если K_t непрерывна по t в топологии индуцированной нормой оператора, то полугруппа *равномерно непрерывна*. Оператор $A = \mathrm{d}K_t / \mathrm{d}t \big|_{t=0}$ называется *генератором* полугруппы, для которого выполняется соотношение $K_t = \mathrm{e}^{tA}$ на всюду плотной области определения A, для равномерно непрерывной полугруппы он является логарифмом K_1, ограничен и определен на всем пространстве. Кроме того, $K_t\, g$ является решением дифференциального уравнения

$$dK_t f / dt = AK_t \, f, \; K_0 f = g. \qquad (3)$$

В нашем случае $\{K_t : -\infty < t < \infty\}$ есть равномерно непрерывная группа с генератором $A = \ln K$, таким образом $K = K_1$ включен в группу $K_t = \mathrm{e}^{t \ln K}$.

Дробным интегралом на положительной полуоси [6] называется оператор, определенный на множестве достаточно быстро убывающих на бесконечности функций,

$$\mathcal{J}_\alpha f(x) = \frac{1}{\Gamma(\alpha)} \int_x^{+\infty} \frac{f(t)}{(t-x)^{1-\alpha}} dt, \; \alpha > 0, \qquad (4)$$

который имеет обратный в виде *дробной производной*

$$\mathcal{D}_\alpha f(x) = -\frac{d^{\lceil \alpha \rceil}}{dx^{\lceil \alpha \rceil}} \mathcal{J}_{\lceil \alpha \rceil - \alpha} f(x),$$

$$\mathcal{D}_0 = \mathcal{J}_0 = I = \mathcal{D}_\alpha \mathcal{J}_\alpha.$$

$\mathcal{J}_\alpha$ является сильно непрерывной полугруппой в метрике L_2 с генератором

$$\mathcal{A} f(x) = \gamma f(x) - \frac{d}{dx} \int_x^{+\infty} \ln(t-x) f(t) dt, \qquad (5)$$

где γ – константа Эйлера.

3. Градиентный спуск и его модификация в контексте операторной полугруппы

В дальнейшем, помимо условий, перечисленных в конце раздела 1, мы будем предполагать, что пространство, в котором действует оператор K, гильбертово.

Метод *градиентного* (наискорейшего) *спуска* отыскания минимума [1, 7] основан на движении вдоль градиента функционала с обратным знаком, т. е. в направлении наибольшего убывания. Однако, скорость сходимости может быть очень низкой.

Градиентный спуск допускает реализацию в виде дифференциального уравнения [8]. Соответствующее уравнение для функционала $F(K, f, g)$ имеет вид:

$$df_t / dt = -\nabla F(f_t), \; f_0 = g, \; 0 \le t \le \infty \qquad (6)$$

где t – параметр траектории движения вдоль градиента. Градиент ∇ вычисляется посредством линейной вариации:

$$dF(f+\tau h)/d\tau\,|_{\tau=0} = \langle \nabla F(f), h\rangle .$$

Для упрощения записи здесь и далее K и g исключены из числа аргументов, т. к. они фиксированы. Для функционала (1) соответствующие градиенты имеют вид:

$$\nabla G = pK^*(Kf-g)\,|\,Kf-g\,|^{p-2},$$

$$\nabla J = q\sum_{i=1}^{n}\left(\frac{\partial}{\partial x_i}\right)^*\left[\frac{\partial f}{\partial x_i}\left(\sum_{j=1}^{n}\left(\frac{\partial f}{\partial x_j}\right)^2\right)^{q/2-1}\right]. \qquad (7)$$

Если $p, q < 2$, в этих формулах появляется возможность деления на 0, которая устраняется заменой выражений с отрицательной степенью: $X^{p-2} \to \left[\varepsilon^2 + X^2\right]^{p/2-1}$ и $Y^{q/2-1} \to \left[\varepsilon^2 + Y\right]^{q/2-1}$ для некоторого $\varepsilon > 0$.

Предметом настоящего исследования является численный метод минимизации функционала F под названием *вариация полугруппового дифференциала*. Идею данного подхода в двух словах можно сформулировать следующим образом. Логарифм оператора генерирует однопараметрическую полугруппу, что позволяет интерпретировать оператор и его обращение как непрерывный процесс, который затем корректируется с учетом задачи оптимизации. Ниже мы даем краткое изложение метода, подробный вывод и доказательство содержатся в [9].

Пусть $A = \ln K$ и $K_t = e^{tA}$. Прямое нерегуляризованное решение задачи $g = Kf$ имеет вид $f = e^{-A}g$, где $f = f_1$ удовлетворяет уравнению (3) в форме

$$df_t/dt = -Af_t,\ f_0 = g,\ 0 \le t \le 1 .$$

Проварьировав правую часть этого уравнения, мы получим направление наибольшего убывания функционала F подобное градиенту, после чего модифицируем уравнение следующим образом:

$$df_t/dt = -Af_t - K^*_{t-1}\nabla F\left(K_{t-1}f_t\right), \qquad (8)$$

начальное условие и интервал интегрирования остаются прежними. Алгоритм регуляризованного решения представляет собой следующую последовательность итераций: $f^1 = f_1$, где f_t есть решение уравнения (8); для $i \ge 1$ $f^{i+1} = f^i_1$, где f^i_t есть решение (8) с начальным условием $f^i_0 = Kf^i$. Окончание процесса фиксируется как обычно, при достижении разницы между соседними элементами последовательности заданной достаточно малой величины. Алгоритм, по сути, аналогичен решению уравнения (6) метода наикратчайшего спуска, но сходится на порядок быстрее.

Чтобы вычислить добавочный член в правой части (8) для функционала (1), следует в выражениях для градиента (7) провести замену $f \to K_{t-1}f_t$ и добавить действие K^*_{t-1}, для G это означает замены $Kf \to K_t f_t$ и $K^* \to K^*_t$. Для вычисления $K_t = e^{tA}$ можно использовать разложение в ряд (2).

4. Пример численной реализации

Продемонстрируем работу нашего метода вариации полугруппового дифференциала ВПД в сравнении с простым градиентным спуском ПГС на примере обращения одномерного оператора $\mathcal{J}_\alpha$ (4), $\alpha = 1/2$, который является элементом сильно непрерывной полугруппы дробного интеграла с генератором (5). В дискретном приближении мы можем их использовать в качестве равномерно непрерывной полугруппы и логарифма соответственно, а для отрицательных степеней применять дробную производную $\mathcal{D}_\alpha$. Для регуляризации используем (1), $p = q = 2$.

Уравнение (8) примет вид:

$$\partial f(x,t)/\partial t = -\mathcal{A}\,f(x,t) + \mathcal{J}^*_t\left(g(x) - \mathcal{J}_t\,f(x,t)\right) + \\ -\sigma\mathcal{J}^*_{\frac{1}{2}+t}\left(\frac{\partial^2}{\partial x^2}\right)^*\frac{\partial^2}{\partial x^2}\mathcal{J}_{\frac{1}{2}+t}\,f(x,t), \\ f(x,0) = g(x),\ 0 \le t \le 1/2 . \qquad (9)$$

Для решения дифференциального уравнения в частных производных используется метод линий [10]. После дискретизации по пространственной переменной x, что в реальности соответствует естественной дискретной структуре исходных данных, и применения разностных схем по x, уравнение приобретает вид системы обыкновенных дифференциальных уравнений, зависящих только от параметра t.

Модельный объект выглядит следующим образом: $K = \mathcal{J}_{1/2}$, f(x) – простая ступенька, $g = Kf$, $\bar{g}(x) = g(x) + \eta(x)$, где η – равномерный гауссов шум (рис. 1), $K\tilde{f} = \bar{g}$ – обратная задача с регуляризацей по схеме (9).

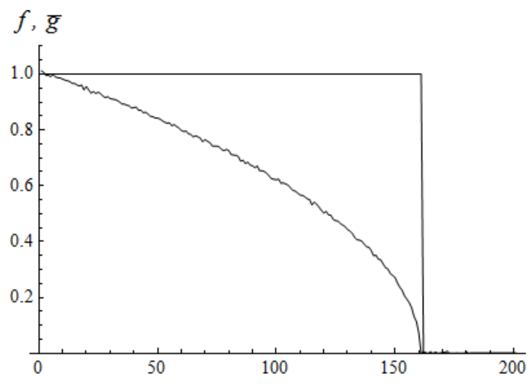

Рис. 1. Оригинальный объект f, подлежащий восстановлению, и его образ $\bar{g}$, искаженный гауссовым шумом.

Значение параметра σ выбрано исходя из приблизительного равенства отношения сигнал/шум на входе и на выходе.

ПГС сходится приблизительно за 100 с условного времени (рис. 2). ВПД сходится за 7 итераций по 0.5 с (рис. 3), а также за 1 итерацию после умножения на 11 добавочного слагаемого.

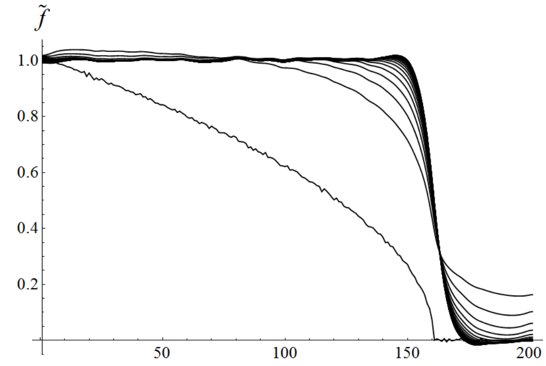

Рис. 2. Последовательность промежуточных значений ПГС решения уравнения (6), $0 \le t \le 100$ с шагом $\Delta t = 5$.

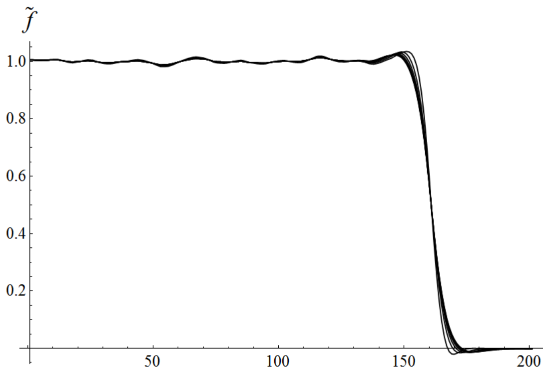

Рис. 3. Последовательность ВПД итераций решения уравнения (8), $k = 1, \dots , 7$.

Очевидно существенное отличие в скоростях сходимости в пользу ВПД.

5. Список литературы

1. Vogel C.R. *Computational Methods for Inverse Problems.* Frontiers in Applied Mathematics. V. 23. Philadelphia: Society for Industrial and Applied Mathematics, 2002. 184 p.
2. Данфорд Н., Шварц Дж. *Линейные операторы. Ч. 1. Общая теория.* М.: ИЛ, 1962. 896 с. (Пер. с англ. Dunford N., Schwartz J.T. *Linear Operators. I. General Theory.* New York: Interscience Publishers, 1958).
3. Wouk A. Integral representation of the logarithm of matrices and operators. *Journal of Mathematical Analysis and Applications.* 1965. V. 11. P. 131–138.
4. Martinez C., Sanz M. *The Theory of Fractional Powers of Operators.* North-Holland Mathematics Studies. V. 187. Amsterdam: North-Holland Publishing Co., 2001. 366 p.
5. Хилле Э., Филлипс Р. *Функциональный анализ и полугруппы.* М.: ИЛ, 1962. 830 с. (Пер. с англ. Hille E., Phillips R.S. *Functional Analysis and Semi-Groups.* Providence: American Mathematical Society, 1957).
6. Самко С.Г., Килбас А.А., Маричев О.И. *Интегралы и производные дробного порядка и некоторые их приложения.* Минск: Наука и техника, 1987. 688 с.
7. Polak El. *Optimization. Algorithms and Consistent Approximations.* Applied Mathematical Sciences. V. 124. New York: Springer-Verlag, 1997. 780 p.
8. Rudin L.I., Osher S., Fatemi E. Nonlinear total variation based noise removal algorithms. *Physica D.* 1992. V. 60. P. 259–268.
9. Дроздовский С.А. *Операторный логарифм и вариационные методы в задачах обращения*: препринт № 05-17 Института теоретической и экспериментальной физики НИЦ Курчатовский институт. Москва: ИТЭФ, 2017. 19 с.
10. Schiesser W.E. *The Numerical Method of Lines: Integration of Partial Differential Equations.* San Diego: Academic Press, 1991. 327 p.