Труды Петрозаводского государственного университета
Серия "Математика" Выпуск 18, 2011

УДК 517.954

С. В. Маничева, И. А. Чернов

ГРАДИЕНТНАЯ ИДЕНТИФИКАЦИЯ ЭВОЛЮЦИОННЫХ СЕТОЧНЫХ ЗАДАЧ

В статье предлагается эффективный метод вычисления градиента конечномерного функционала на пространстве большой размерности с ограничениями особой структуры (уравнения нелинейной эволюционной сеточной краевой задачи).

§ 1. Введение

Различные физические процессы, включая процессы тепломассопереноса, описываются эволюционными краевыми задачами с одной пространственной переменной. Достаточно упомянуть модели взаимодействия водорода с металлами [1, 2] и локально-одномерное описание термодинамики морского льда [3]. Физический смысл влечет зависимость уравнений модели от числовых параметров (например, характеристик материала). Представляют интерес эффективные методы оценки этих параметров на основе измерений. Нелинейное описание ряда процессов существенно осложняет аналитическое решение краевых задач. На первое место выходят различные численные методы, в том числе сеточные, включающие в себя конечно-разностные, конечно-элементные, конечно-объемные и т. д. Таким образом, мы приходим к задаче условной минимизации конечномерного функционала на пространстве большой размерности, но с системой ограничений особой структуры, допускающей эффективное вычисление градиента в подпространстве параметров (как правило, небольшой размерности). Методу вычисления градиента с учетом структуры ограничений в форме эволюционной сеточной задачи достаточно общего вида и посвящена работа.

§ 2. Постановка задачи

Введем в прямоугольнике $[0,1]\times[0,1]$ равномерную сетку $\bar{D}$ с шагами τ и h. Ее узлы (n,i) — это точки $(t_n,x_i)=(n\tau,ih)$ при $0\le n\le N$, $0\le i\le I$, $N\tau=1$, $Ih=1$. Пусть D — подмножество внутренних узлов ($0<n<N$, $0<i<I$). Значение сеточной функции u в узле (n,i) обозначим u_n^i. Совокупность значений u_n^i при фиксированном n обозначим $\vec{u}_n$, а объединение $\vec{u}_n$ при $n\le m$ — через $\vec{u}_{n\le m}$. Рассмотрим сеточную задачу (систему алгебраических уравнений)

$$f_n^i(\vec{u}_{m\le n},\vec{\xi})=0,\quad (n,i)\in\bar{D}, \tag{1}$$

где функции f_m^k гладкие, а вектор $\vec{\xi}\in R^M$ — это вектор параметров. Предполагаем, что эта система разрешима при любом $\vec{\xi}\in R^M$.

Подчеркнем «независимость от будущего» — независимость функций f_n^i от u_m^k при $m>n$, которую предполагаем; при трактовке индекса n как времени указанное свойство и означает эволюционность задачи. В самом деле, при $n=0$ функции $f_0^i(\vec{u}_0,\vec{\xi})$ при заданном $\vec{\xi}$ зависят только от $\vec{u}_0$, поэтому, решая систему $f_0^i(\vec{u}_0,\vec{\xi})=0$, определяем u_0^i при $0\le i\le I$. Далее решаем систему $f_1^i(\vec{u}_{m\le 1},\vec{\xi})=0$ относительно $\vec{u}_1^i$, с учетом известных $\vec{\xi}$ и $\vec{u}_0^i$. Продолжая далее, определяем последовательно все u_n^i. Итак, эволюционная задача размерности $(N+1)\times(I+1)$ сводится к $N+1$ задаче размерности $I+1$ без каких-либо дополнительных предположений.

К таким задачам сводятся, в частности, разностные схемы для параболических краевых задач [4]. В этом случае функции f_n^i являются сеточными аппроксимациями уравнения в частных производных и краевых условий. Зависимость f_n^i от n и i связана с зависимостью коэффициентов исходной задачи от времени и точки пространства, а зависимость от предыстории — со внутренними состояниями различной физической природы [5]. Отметим, что перенумерация узлов позволяет записать в виде (1) также и разностные схемы с несколькими пространственными переменными [6].

Система (1) позволяет определить u_n^i для всех $(n,i)\in\bar{D}_N$ при заданном ξ; для вычисления ξ требуется дополнительная информация. Предположим, что вектор ξ должен минимизировать функционал

$$F(\vec{u}_{m\le N},\vec{\xi}). \tag{2}$$

Функцию F предполагаем гладкой и ограниченной снизу на пространстве $R^{N+I+2+M}$. Если речь идет об измерениях, то фактически функция F зависит не от всех значений u_n^i, а только от некоторых, например, только от u_n^0 — значений на поверхности.

§ 3. Метод множителей Лагранжа

Имеем обычную задачу условной оптимизации большой размерности с системой ограничений особой структуры. Поскольку аналитическое решение системы (1) невозможно или очень громоздко, то применим метод множителей Лагранжа, перейдя к задаче безусловной оптимизации функции

$$L(\vec{u}_{m\le N}, \vec{\xi}, \lambda_n^i) = \Lambda F + \sum_{n=0}^{N}\sum_{i=0}^{I}\lambda_n^i f_n^i. \tag{3}$$

Множитель Λ либо равен нулю, либо может быть положен равным единице. Случай $\Lambda = 0$ исключается требованием линейной независимости градиентов $\nabla_u f_n^i$ — векторов частных производных по переменным u_n^i — на решении задачи (1) при любых $\vec{\xi} \in R^M$. В дальнейшем полагаем $\Lambda = 1$.

Частные производные функции Лагранжа L можно разделить на три группы: по переменным u_n^i, по множителям Лагранжа λ_n^i и по параметрам — компонентам ξ_k вектора $\vec{\xi}$. Условие обнуления производных второй группы равносильно системе (1). Найдем и приравняем нулю производные второй группы:

$$\frac{\partial L}{\partial u_n^i} = \frac{\partial F}{\partial u_n^i} + \sum_{m=n}^{N}\sum_{j=0}^{I}\lambda_m^j\frac{\partial f_m^j}{\partial u_n^i} = 0. \tag{4}$$

Отметим, что нижний индекс суммирования равен n в силу эволюционности. Также слагаемое, содержащее производные от F, равно нулю в большинстве узлов, если функционал зависит только от некоторых значений; это позволяет существенно сократить вычисления в прикладных задачах.

Наконец, производные по параметрам имеют вид

$$\frac{\partial L}{\partial \xi_k} = \frac{\partial F}{\partial \xi_k} + \sum_{m=0}^{N}\sum_{i=0}^{I}\lambda_m^i\frac{\partial f_m^i}{\partial \xi_k}. \tag{5}$$

Это — компоненты градиента функции F в подпространстве параметров R^M в любой точке $\vec{\xi} \in R^M$.

§ 4. Декомпозиция

Пусть точка $\vec{\xi} \in R^M$ задана. Выше уже описана декомпозиция системы уравнений (1), сводящейся к последовательному решению подсистем для фиксированного n в ходе возрастания n. Это позволяет эффективно вычислить значения переменных u_n^i. Приравнивая нулю выражения (4), получаем систему линейных уравнений относительно λ_n^i размерности $(N+1) \times (I+1)$, также допускающую декомпозицию. Эта система невырождена в силу предположения о линейной независимости векторов $\nabla_u f$. Параметры ξ_k заданы, а значения u_n^i уже найдены; подсистема (4) при $n = N$ содержит только неизвестные λ_N^i при $0 \le i \le N$ и поэтому может быть решена независимо. Но тогда подсистема (4) при $n = N - 1$ содержит только неизвестные λ_{N-1} и уже вычисленные λ_N^i. Продолжая далее, последовательно находим λ_n^i при фиксированном n в порядке его убывания. Выражения (5) позволяют явно найти градиент ∇F функции F в пространстве параметров.

Таким образом, учет структуры системы ограничений позволяет эффективно вычислить градиент, который может быть использован, например, для методов спуска или тяжелого шарика. Метод тяжелого шарика (движение в пространстве параметров вдоль траектории системы $\mu\ddot{\xi} + \gamma\dot{\xi} + \nabla F = 0$, где точкой обозначено дифференцирование по параметру траектории, а коэффициенты «трения» γ и «инерции» μ подбираются) выглядит предпочтительнее градиентного спуска из-за «инерции», которая позволяет избежать остановки в критических точках, не являющихся минимумами, а при удачном выборе коэффициента «трения» γ — и миновать «ложные минимумы». Разумеется, система дифференциальных уравнений, описывающая движение тяжелого шарика, решается численно.

§ 5. Частные случаи

Рассмотрим несколько частных случаев поставленной задачи. Сначала изучим простейший модельный пример.

Пусть сетка состоит из шести точек (n, i), $n = 0, 1$, $i = 0, 1, 2$.

Функции f_n^i заданы следующим образом:

$$
\begin{aligned}
&f_0^i = u_0^i - 1, \quad f_1^0 = u_1^0 - 1,\\
&f_1^1 = u_1^0 - (2+\xi_1)u_1^1 + u_1^2 + \xi_1 u_0^1,\\
&f_1^2 = u_1^2 - u_1^1 + \xi_1\xi_2(u_1^2)^2.
\end{aligned}
$$

Такой выбор связан с аппроксимацией уравнения диффузии на сетке: при единичных шагах сетки f_1^1 является разностной схемой для уравнения диффузии ($(\xi_1)^{-1}$ — это коэффициент диффузии), f_0^i задают постоянное начальное распределение, f_1^0 — граничное условие I рода, а f_1^2 — нелинейное граничное условие «объемной десорбции» [1].

Решим систему уравнений (1). Очевидно, $u_1^0 = u_0^i = 1$ для всех $i = 0, 1, 2$, u_1^2 является положительным корнем квадратного уравнения

$$
(2+\xi_1)\xi_1\xi_2(u_1^2)^2 + (1+\xi_1)u_1^2 - (1+\xi_1) = 0,
$$

(при $\xi_k > 0$, что подразумевается, единственный вещественный положительный корень существует), а u_1^1 выражается через u_1^2:

$$
u_1^1 = \frac{1+\xi_1+u_1^2}{2+\xi_1}.
$$

Функционал зададим формулой

$$
F = \left(\xi_2(u_1^2)^2 - \frac{4}{9}\right)^2.
$$

Это среднеквадратичное отклонение «квадратичного потока» $\xi_2(u_1^2)^2$ от модельного значения, соответствующего значению $u_1^2 = 1/3$, которое получается при $\xi_1 = 1$, $\xi_2 = 4$.

Решим систему (4). Поскольку u_0^0 и u_0^2 входят только в f_0^0 и в f_0^2 соответственно, то $\lambda_0^0 = \lambda_0^2 = 0$. Величина u_0^1 линейно входит в f_0^1 и f_1^1, так что получаем уравнение $\lambda_0^1 + \xi_1\lambda_1^1 = 0$. Аналогично получаются уравнения $\lambda_1^0 + \lambda_1^1 = 0$ и $(2+\xi_1)\lambda_1^1 + \lambda_1^2 = 0$. Из последнего уравнения λ_1^1 выразим через λ_1^2. Наконец, u_1^2 нелинейно входит в формулы для функционала F и функций f_1^1 и f_1^2, что приводит к уравнению

$$
4\xi_2 u_1^2\left(\xi_2(u_1^2)^2 - \frac{4}{9}\right) + \lambda_1^1 + \lambda_1^2\bigl(1 + 2\xi_1\xi_2 u_1^2\bigr) = 0.
$$

После подстановки λ_1^1 получаем

$$\lambda_1^2 = -\frac{(2+\xi_1)}{9} \cdot \frac{4\xi_2 u_1^2\left(9\xi_2(u_1^2)^2-4\right)}{1+\xi_1+2(2+\xi_1)\xi_1\xi_2 u_1^2}.$$

Теперь найдем производные по параметрам по формулам (5):

$$\frac{\partial L}{\partial \xi_1} = \lambda_1^2\xi_2(u_1^2)^2 - \lambda_1^1 u_1^1, \quad \frac{\partial L}{\partial \xi_2} = 2(u_1^2)^2\left(\xi_2(u_1^2)^2 - \frac{4}{9}\right) + \lambda_1^2(u_1^2)^2.$$

Очевидно, эти выражения обращаются в ноль в точке $\vec{\xi} = (1,4)$, так как при этом $u_1^2 = 1/3$. Нетрудно убедиться, что $F \equiv 0$ на кривой

$$\xi_2 = 36\left(9 - 4\frac{2+\xi_1}{1+\xi_1}\xi_1\right)^{-2}, \quad 4\frac{2+\xi_1}{1+\xi_1}\xi_1 < 9.$$

Подведем итоги. Предельно упрощенная нелинейная краевая задача допускает аналитическое решение, но формулы весьма громоздки, причем решение задачи определения параметров неединственно; более того, нет изолированных решений.

Теперь рассмотрим частный случай (1): двухслойную разностную схему для эволюционного уравнения, разрешенного относительно производной по времени. Требование «двуслойности» означает, что для аппроксимации производных используются значения сеточного решения только на двух слоях, то есть u_n^i и u_{n-1}^i при фиксированном n и произвольных i. Итак, функции f_n^i теперь имеют вид

$$f_n^i = u_n^i - \nu_n^i u_{n-1}^i - g_n^i(\vec{u}_n, \vec{u}_{n-1}),$$

где ν_n^i — числовой коэффициент (он может понадобиться, например, для задания краевых условий — в этих случаях он равен нулю). Система (4) имеет вид

$$\frac{\partial L}{\partial u_n^i} = \frac{\partial F}{\partial u_n^i} + \lambda_n^i - \nu_{n+1}^i\lambda_{n+1}^i - \sum_{j=0}^{I}\left(\lambda_n^j\frac{\partial g_n^j}{\partial u_n^i} + \lambda_{n+1}^j\frac{\partial g_{n+1}^j}{\partial u_n^i}\right) = 0.$$

Обратим внимание на «обмен знаками»: вместо $u_n^i - \nu_n^i u_{n-1}^i$ в исходной системе (1) имеем $\lambda_n^i - \nu_{n+1}^i\lambda_{n+1}^i$ в сопряженной системе (4). Этот феномен хорошо известен в теории сопряженных задач [7]. Если, например, $\nu_n^i = 1$ при $(n,i) \in D$, а $g_n^i = (u_n^{i-1} - 2u_n^i + u_n^{i+1})\tau h^{-2}$ при

$(n,i) \in D$, то в D задана неявная двухслойная разностная схема для уравнения теплопроводности/диффузии:

$$\frac{u_{n+1}^i - u_n^i}{\tau} = \frac{u_n^{i-1} - 2u_n^i + u_n^{i+1}}{h^2}.$$

Пусть также функционал F не зависит от u_n^i при $(n,i) \in D$ (то есть от значений во внутренних узлах). Тогда система (4) примет вид

$$-\frac{\lambda_{n+1}^i - \lambda_n^i}{\tau} = \frac{\lambda_n^{i-1} - 2\lambda_n^i + \lambda_n^{i+1}}{h^2}, \quad (n,i) \in D, \quad n < N.$$

Эта система также является двухслойной разностной схемой, причем также неявной, но в «обратном времени», то есть при расчете в ходе убывания n. Если формально положить $\lambda_{N+1}^i = 0$, то указанные уравнения окажутся применимы и для $n = N$.

Отметим, что если функция f_n^i не зависит от параметров ξ (что типично для начальных условий или граничных условий I рода), то соответствующий λ_n^i не входит в (5) и может не вычисляться.

Рассмотрим, наконец, еще один частный случай системы (1): зависимость системы от внутренних состояний [4]. Под внутренним состоянием мы понимаем величину, динамика которой определяется текущим состоянием системы, и которая сама влияет на систему. Например, это может быть температура гидрируемого образца, определяющая скорости физико-химических процессов и меняющаяся в силу поглощения тепла в ходе этих реакций. Другой пример — свободная граница, устраненная заменой переменных.

В этом случае систему (1) запишем так:

$$\begin{aligned} f_n^i = G(\vec{u}_{m \le n}, s_n, \vec{\xi}) = 0, \quad (n,i) \in \bar{D}, \\ s_n = s_{n-1} + \Gamma_n(s_{n-1}, \vec{u}_{n-1}, \vec{\xi}) = 0. \end{aligned}$$

Достаточно вычислить производные

$$\frac{\partial f_m^j}{\partial u_n^i} = \frac{\partial G_m^j}{\partial u_n^i} + \frac{\partial G_m^j}{\partial s_m} \cdot \frac{\partial \Gamma_{n+1}}{\partial u_n^i}, \quad \frac{\partial f_n^i}{\partial \xi_k} = \frac{\partial G_n^i}{\partial \xi_k} + \frac{\partial G_n^i}{\partial s_n} \cdot \frac{\partial \Gamma_n}{\partial \xi_k}.$$

Все выкладки выше применимы и в этом частном случае зависимости от «предыстории».

Таким образом, предложенная в статье методика вычисления градиента может быть применена для широкого класса сеточных эволюционных задач и функционалов на решении этих задач и может эффективно использоваться для решения задач параметрической идентификации математических моделей, описываемых эволюционными уравнениями в частных производных.

Résumé

We construct the gradient of finite-dimensional functional in the space of large dimension with constraints. The system of constraints has special structure: it is a nonlinear evolutionary grid boundary-value problem.

Список литературы

[1] Заика Ю. В., Родченкова Н. И. *Моделирование высокотемпературного пика ТДС-спектра дегидрирования* // Математическое моделирование. 2006. Т. 18. № 4. С. 100–112.

[2] Чернов И. А. *Математическая модель экзотермичного формирования гидрида* // Математическое моделирование. 2010. Т. 22. № 1. С. 3–16.

[3] Maykut G. A., Untersteiner N. *Some Results from a Time-Dependent Thermodynamic Model of Sea Ice* // Journal of Geophysical Research. 1971. V. 76. No 6. P. 1550–1575.

[4] Chernov I. A. *Maximum Principle for the Parabolic Boundary-Value Problems with Nonlinear Boundary Conditions and Inner State* // Mathematical Modelling / Christopher R. Brennan. Mathematics Research Development series. 2011. P. 203–246.

[5] Чернов И. А. *Сходимость сеточно-интерполяционных аппроксимаций решения квазилинейной параболической краевой задачи на отрезке* // Труды ПетрГУ. Серия Математика. 2010. Вып. 17. С. 26–37.

[6] Rodchenkova N. I., Zaika Yu. V. *Numerical modelling of hydrogen desorption from cylindrical surface* // International Journal of Hydrogen Energy. Elsevier. 2011. V. 36. P. 1239–1247.

[7] Марчук Г. И., Агошков В. И., Шутяев В. П. *Сопряженные уравнения и методы возмущений в нелинейных задачах математической физики.* М.: Физматлит, 1993. 224 с.

ИПМИ КарНЦ РАН, 185910, Петрозаводск, ул. Пушкинская, 11;
КГПА, 185910, Петрозаводск, ул. Пушкинская, 17
E-mail: IAChernov@yandex.ru, smanicheva@mail.ru