Владикавказский математический журнал
Октябрь–декабрь, 2004, Том 6, Выпуск 4

УДК **517.5**

ОПТИМАЛЬНАЯ ИНТЕРПОЛЯЦИЯ И ПРИНЦИП ЛАГРАНЖА[1]

Г. Г. Магарил-Ильяев

Дорогому учителю Владимиру Михайловичу
Тихомирову с благодарностью за науку

На примере задачи интерполяции, демонстрируется применение принципа Лагранжа для решения задач оптимального восстановления линейных функционалов.

Задача интерполяции, которая здесь рассматривается, заключается в следующем. Пусть на отрезке $[a,b]$ заданы n точек $a \leqslant t_1 < \ldots < t_n \leqslant b$, и в этих точках известны значения некоторой функции $x(\cdot)$, т. е. известны числа $x(t_1),\ldots,x(t_n)$, а мы хотим определить значение этой функции в другой точке $\tau \in [a,b]$, где $\tau \neq t_i$, $1 \leqslant i \leqslant n$. Стандартный прием состоит в том, что строят полином $p\,(\cdot)$ степени $n-1$ такой, что $p\,(t_i) = x(t_i)$, $1 \leqslant i \leqslant n$ (он определен однозначно и называется *интерполяционным полиномом Лагранжа*) и считают, что $x(\tau) = p\,(\tau)$. Можно поступить проще: соединяя точки $(t_i, x(t_i))$ и $(t_{i+1}, x(t_{i+1}))$, $i = 1,\ldots,n-1$, отрезками прямых, получим кусочно линейную функцию $f(\cdot)$ и будем считать, что $x(\tau) = f(\tau)$. Можно придумать и другие способы восстановления (как мы будем говорить) функции $x(\cdot)$ в точке τ. При этом ясно, что до тех пор пока нет никакой априорной информации о функции $x(\cdot)$, мы не можем отдать предпочтение ни одному из таких способов. Предположим теперь, что такая информация о функции $x(\cdot)$ имеется и она заключается в том, что $x(\cdot)$ принадлежит некоторому фиксированному классу функций C на отрезке $[a,b]$. Такого рода информация может появиться, например, из следующих соображений: пусть функция $x(\cdot)$ имеет некую физическую природу и известно (в силу этой природы), что резкие скачки $x(\cdot)$ невозможны. Формализовать это можно, например, так: производная функции $x(\cdot)$ по модулю не превосходит некоторого положительного числа. Тем самым выделится класс функций C, которому наша функция принадлежит.

Итак, будем исходить из того, что мы наблюдаем значения функции $x(\cdot)$, про которую известно, что она принадлежит некоторому классу функций C. В этой ситуации мы уже можем сравнивать различные способы восстановления функций из C в точке τ. Поступаем следующим образом. Пусть $m(\cdot)$ — произвольная функция n переменных. Для каждого $x(\cdot)$ в качестве оценки $x(\tau)$ возьмем число $m(x(t_1),\ldots,x(t_n))$. Эффективность такого метода восстановления $x(\tau)$ на классе C будем характеризовать величиной

$$e(\tau, C, t_1,\ldots,t_n; m(\cdot)) = \sup_{x(\cdot)\in C} |x(\tau) - m(x(t_1),\ldots,x(t_n))|, \tag{1}$$

[1]Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (гранты №02-01-39012 и №02-01-00386), программ государственной поддержки ведущих научных школ Российской Федерации (НШ-304.2003.1) и «Университеты России» (УР.04.03.067), а также при поддержке U.S. CRDF-R.F. Ministry of Education Award VZ-0100-0.

дающей максимальное (на классе) расхождение между истинным значением функции в точке τ и оценкой этого значения с помощью функции (метода) $m(\cdot)$.

Заметим, что оценка $x(\tau)$ с помощью интерполяционного полинома Лагранжа состоит в том, что мы берем число $p(\tau) = \sum_{i=1}^{n} l_i(\tau)x(t_i)$, где $l_i(\cdot)$, $1 \leqslant i \leqslant n$, — полиномы степени $n-1$ такие, что $l_i(\tau_j) = 1$, если $i = j$, и $l_i(\tau_j) = 0$, если $i \neq j$. Другими словами, в качестве метода восстановления $x(\tau)$ мы рассматриваем линейную функцию $m(\xi_1, \ldots, \xi_n) = \sum_{i=1}^{n} l_i(\tau)\xi_i$.

Нас интересует вопрос, на каком методе восстановления $m(\cdot)$ величина (1) достигает своего минимального значения и чему это значение равно? Точнее говоря, нас интересует величина

$$E(\tau, C, t_1, \ldots, t_n) = \inf_{m(\cdot)} e(\tau, C, t_1, \ldots, t_n; m(\cdot)) \tag{2}$$

(где нижняя грань берется по всем функциям n переменных), которая называется *погрешностью оптимального восстановления* и метод $\widehat{m}(\cdot)$, на котором эта нижняя грань достигается, называемый *оптимальным методом восстановления.*

Нахождение оптимальной погрешности восстановления и оптимального метода восстановления будем называть *задачей оптимальной интерполяции на классе* C.

В данной работе мы решаем эту задачу для случая, когда $C = W_\infty^n([a,b])$ (совокупность функций $x(\cdot)$ на $[a,b]$, у которых $(n-1)$-ая производная $x^{(n-1)}(\cdot)$ абсолютно непрерывна на $[a,b]$ и $\|x^{(n)}(\cdot)\|_{L_\infty([a,b])} \leqslant 1$) и показываем, что интерполяционный полином Лагранжа — оптимальный метод восстановления. Основная цель, которую мы при этом преследуем — продемонстрировать применение принципа Лагранжа для решения задачи оптимального восстановления. Этот принцип позволяет получать решения многих экстремальных задач, в некотором смысле, автоматически. Он заключается в том, что надо составить функцию (функцию Лагранжа), которая есть сумма минимизируемого или максимизируемого функционала и функций, задающий различные ограничения с неопределенными множителями (множителями Лагранжа) и тогда решение исходной задачи удовлетворяет необходимым условиям минимума в более простой задаче — в задаче на минимум функции Лагранжа без ограничений (при некотором наборе множителей Лагранжа). Этот прием (который впервые был сформулировал Лагранжем для задач с ограничениями типа равенств) оказался универсальным — он верен (при весьма широких предположениях) для любых экстремальных задач с ограничениями типа равенств, неравенств и включений. Если задача выпукла, то функция Лагранжа на решении достигает минимума и более того, если множитель Лагранжа при минимизируемом (или максимизируемом) функционале отличен от нуля, то функция, которая удовлетворяет необходимым условиям минимума функции Лагранжа, является решением исходной задачи. Подробнее о принципе Лагранжа и его приложениях см. в [1] и [2]. Не бояться применять принцип Лагранжа для решения самых различных задач я научился у В. М. Тихомирова.

Применением принципа Лагранжа, доказывается следующая

Теорема 1. *Справедливо равенство*

$$E(\tau, W_\infty^n([a,b]), t_1, \ldots, t_n) = \frac{1}{n!}\left|\prod_{j=1}^{n}(\tau - t_j)\right|$$

и метод $\widehat{m}(\xi_1, \ldots, \xi_n) = \sum_{i=1}^{n} l_i(\tau)\xi_i$ *(интерполяционный полином Лагранжа) является оптимальным.*

$\lhd$ Свяжем с задачей оптимального восстановления следующую экстремальную задачу

$$x(\tau)\to\max,\quad x(t_i)=0,\quad 1\leqslant i\leqslant 1,\quad x(\cdot)\in W^n_\infty([a,b]) \tag{3}$$

и обозначим через $S(\tau,W^n_\infty([a,b]),t_1,\dots,t_n)$ ее значение. Покажем, что

$$E(\tau,W^n_\infty([a,b]),t_1,\dots,t_n)\geqslant S(\tau,W^n_\infty([a,b]),t_1,\dots,t_n). \tag{4}$$

Действительно, пусть $x(\cdot)$ — допустимая функция в (3). Тогда и $-x(\cdot)$ — допустимая функция в этой задаче и мы имеем для любого метода $m(\cdot)$

$$2x(\tau)\leqslant|x(\tau)-m(0)+x(\tau)+m(0)|\leqslant|x(\tau)-m(0)|+|-x(\tau)-m(0)|$$
$$\leqslant 2\sup_{\substack{x(\cdot)\in W^n_\infty([a,b])\\ x(t_i)=0,\,i=1,\dots,n}}|x(\tau)-m(0)|\leqslant 2\sup_{x(\cdot)\in W^n_\infty([a,b])}|x(\tau)-m(x(t_1),\dots,x(t_n))|.$$

Переходя справа к нижней грани по всем методам $m(\cdot)$, а затем слева к верхней грани по всем допустимым в (3) функциям $x(\cdot)$, получаем требуемое утверждение. $\rhd$

Задача оптимального восстановления является (в определенном смысле) двойственной к задаче (3) (см. [2]) и поэтому решив последнюю, мы найдем и оптимальный метод и погрешность оптимального восстановления.

Задачу (3) запишем формально следующим образом

$$\begin{gathered}\int_a^b\delta(t-\tau)x(t)\,dt\to\max,\\ \int_a^b\delta(t-t_i)x(t)=0,\quad 1\leqslant i\leqslant 1,\quad x^{(n)}(t)=u(t),\quad |u(t)|\leqslant 1,\end{gathered} \tag{5}$$

где $\delta(\cdot-\xi)$ — δ-функция в точке ξ и последнее неравенство понимается п. в.

Мы выпишем необходимые условия максимума в ней, рассуждая эвристически, не ссылаясь ни на какие точные результаты, а руководствуясь лишь здравым смыслом и верой в то, что принцип Лагранжа верен. Это приведет нас к некоторым соотношениям, справедливость которых мы уже проверим непосредственно. Так как задача (3) выпукла, то эти соотношения позволят нам найти ее решение, а затем доказать все утверждения теоремы.

Отметим, что внешне задача (5) — стандартная задача оптимального управления и можно было бы сразу выписать необходимые условия максимума, но мы их получим, исходя из общих рассуждений.

Выпишем функцию Лагранжа задачи, не включая в нее ограничение $|u(t)|\leqslant 1$ и воспринимая ограничение $x^{(n)}(t)=u(t)$ как континуум равенств, каждое из которых надо умножить на множитель Лагранжа $p\,(t)$, а затем их «просуммировать»:

$$\mathscr{L}(x(\cdot),u(\cdot),\lambda)=\int_{-1}^1((-\delta(t-\tau)+\sum_{i=1}^n\mu_i\delta(t-t_i))x(t)+p\,(t)(x^{(n)}(t)-u(t))\,dt.$$

Здесь $\lambda=(-1,\mu_1,\dots,\mu_n,p\,(\cdot))$ (мы взяли множитель Лагранжа при максимизируемом функционале равным -1, чтобы полученные необходимые условия максимума оказались и достаточными).

Если пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ — решение задачи (5), то согласно принципу Лагранжа существует такое λ, что функция Лагранжа должна достигать минимума в этой точке по $x(\cdot)$ и по $u(\cdot)$, т. е. должны выполняться условия:

$$\mathscr{L}_{x(\cdot)}(\widehat{x}(\cdot), \widehat{u}(\cdot), \lambda) = 0, \qquad \min_{|u(t)|\leqslant 1} \mathscr{L}(\widehat{x}(\cdot), u(\cdot), \lambda) = \mathscr{L}(\widehat{x}(\cdot), \widehat{u}(\cdot), \lambda). \tag{6}$$

Первое соотношение означает, что

$$\int_{-1}^{1} \left(\left(-\delta(t-\tau) + \sum_{i=1}^{n} \mu_i \delta(t-t_i) \right) x(t) + p\,(t) x^{(n)}(t) \right) dt = 0 \quad \forall\, x(\cdot).$$

Интегрируя последнее слагаемое под знаком интеграла n раз по частям, приходим к равенству

$$\int_a^b (-\delta(t-\tau) + \sum_{i=1}^{n} \mu_i \delta(t-t_i) + (-1)^n p^{(n)}(t)) x(t)\, dt \\ + \sum_{k=0}^{n-1} (-1)^k (p^{(k)}(b) x^{(n-k-1)}(b) - p^{(k)}(a) x^{(n-k-1)}(a)) = 0,$$

верному для всех $x(\cdot)$, и тем самым

$$-\delta(t-\tau) + \sum_{i=1}^{n} \mu_i \delta(t-t_i) + (-1)^n p^{(n)}(t) = 0 \tag{7}$$

и

$$p^{(k)}(a) = p^{(k)}(b) = 0, \;\; k = 0, 1, \ldots, n-1. \tag{8}$$

Из второго соотношения в (6) вытекает, что

$$\widehat{x}^{(n)}(t) = \operatorname{sign} p\,(t). \tag{9}$$

Соотношения (7), (8) и (9) и есть необходимые условия максимума в задаче (3), которые можно было сразу выписать.

Теперь проанализируем полученные соотношения. Интегрируя (7), получаем, что

$$p\,(t) = \frac{(-1)^n}{(n-1)!} \left((t-\tau)_+^{n-1} - \sum_{j=1}^{n} \mu_j (t-t_j)_+^{n-1} \right). \tag{10}$$

Покажем, что эта функция (с произвольными μ_i, $1 \leqslant i \leqslant n$) сохраняет знак на $[a,b]$. Действительно, если это не так, то $p\,(\cdot)$ имеет нуль на (a,b). Поскольку $p\,(a) = p\,(b) = 0$, то по теореме Ролля $\dot{p}\,(\cdot)$ имеет на (a,b) не менее двух нулей. Продолжая, получим с учетом (8), что кусочно линейная функция $p^{(n-2)}(\cdot)$ имеет на (a,b) не менее $n-1$ нуля, а значит, снова в силу (8), на отрезке $[a,b]$ она имеет не менее $n+1$ нуля. Тогда по элементарному следствию из теоремы Ролля кусочно постоянная функция $p^{(n-1)}(\cdot)$ имеет на $[a,b]$ не менее n перемен знака. Но эта функция с n интервалами постоянства (она может иметь разрывы только в точках $\tau, t_1, \ldots, t_n$) и поэтому у нее не может быть больше, чем $n-1$ перемен знака. Итак, $p\,(\cdot)$ не меняет знак на $[a,b]$ и тогда согласно (9) функция $\widehat{x}(\cdot)$ есть полином n-го порядка со старшим коэффициентом по модулю равным $1/n!$ По условию $\widehat{x}(t_j)=0$, $1\leqslant j\leqslant n$, и очевидно, $\widehat{x}(\tau)>0$. Этими условиями $\widehat{x}(\cdot)$ определяется однозначно: $\widehat{x}(t)=\alpha(1/n!)\prod_{j=1}^n(t-t_j)$, где $\alpha=\operatorname{sign}\prod_{j=1}^n(\tau-t_j)$.

Определим теперь числа μ_j, $1\leqslant j\leqslant n$. Возьмем любой полином $q(\cdot)$ степени не выше $n-1$, умножим на него обе части равенства (7) и проинтегрируем по отрезку $[a,b]$. Тогда получим $q(\tau)=\sum_{j=1}^n\mu_jq(t_j)$. Подставляя сюда, определенные выше полиномы $l_j(\cdot)\in\mathscr{P}_{n-1}$, найдем, что $\mu_j=l_j(\tau)$, $j=1,\ldots,n$. Таким образом,

$$q(\tau)=\sum_{j=1}^n l_j(\tau)q(t_j). \tag{11}$$

Обозначим через $\mathscr{W}_\infty^n([a,b])$ пространство функций на $[a,b]$, у которых $(n-1)$-ая производная абсолютно непрерывна, а n-ая производная принадлежит $L_\infty([a,b])$ (так что $W_\infty^n([a,b])=\{x(\cdot)\in\mathscr{W}_\infty^n([a,b]) : \|x^{(n)}(\cdot)\|_{L_\infty([a,b])}\leqslant 1\}$). Умножим (7) на функцию $x(\cdot)$ из этого пространства и проинтегрируем по частям. Тогда будем иметь

$$x(\tau)-\sum_{j=1}^n l_j(\tau)x(t_j)=\int_a^b p\,(t)x^{(n)}(t)\,dt. \tag{12}$$

Итак, рассуждая эвристически, мы получили соотношения (11) и (12). Теперь будем рассуждать точно. Элементарная проверка показывает, что равенство (11) действительно имеет место для любого полинома степени не выше $n-1$.

Покажем теперь, что тождество (12) справедливо для любой функции $x(\cdot)\in\mathscr{W}_\infty^n([a,b])$ с функцией $p\,(\cdot)$, определяемой формулой (10), где $\mu_j=l_j(\tau)$, $1\leqslant j\leqslant n$.

Из определения $p\,(\cdot)$ сразу следует, что $p^{(k)}(a)=0$, $k=0,1,\ldots,n-1$. Далее $p^{(k)}(b)=0$, $k=0,1,\ldots,n-1$, поскольку полиномы $q_k(t)=(1-t)^k$, $k=0,1,\ldots,n-1$, удовлетворяют (11). Используя это обстоятельство и интегрируя $n-1$ раз по частям в интеграле справа в (12), убеждаемся после несложных преобразований, что равенство (12) справедливо.

Покажем, что функция $\widehat{x}(\cdot)$ — решение задачи (3). Подставляя в (12) любую допустимую в (3) функцию $x(\cdot)$ и оценивая по модулю, будем иметь

$$x(\tau)=\int_a^b p\,(t)x^{(n)}(t)\,dt\leqslant\int_a^b|p\,(t)|dt.$$

Выше было показано, что $p\,(\cdot)$ не меняет знака на $[a,b]$ и поэтому функция $\widehat{x}^{(n)}(\cdot)$ (которая есть константа) совпадает с точностью до знака с $\operatorname{sign}p\,(\cdot)$. Подставим $\widehat{x}(\cdot)$ в (12) и учитывая, что $\widehat{x}(\tau)>0$, получим

$$\widehat{x}(\tau)=\int_a^b p\,(t)\widehat{x}^{(n)}(t)\,dt=\left|\int_a^b p\,(t)\operatorname{sign}p\,(t)\,dt\right|=\int_a^b|p\,(t)|dt. \tag{13}$$

Таким образом, $\widehat{x}(\cdot)$ — решение задачи (3). Тогда вследствие (4) верно

$$E(\tau,W_\infty^n([a,b]),t_1,\ldots,t_n)\geqslant\widehat{x}(\tau)=\frac{1}{n!}\left|\prod_{j=1}^n(\tau-t_j)\right|. \tag{14}$$

Из (12) и (13) для любого $x(\cdot)\in W_\infty^n([a,b])$ имеем

$$\left|x(\tau)-\sum_{j=1}^n l_j(\tau)x(t_j)\right|\leqslant\int_a^b|p\,(t)|dt=\widehat{x}(\tau)=\frac{1}{n!}\left|\prod_{j=1}^n(\tau-t_j)\right|,$$

т. е.

$$E(\tau, W^n_\infty([a,b]), t_1, \ldots, t_n) \leqslant \frac{1}{n!}\left|\prod_{j=1}^{n}(\tau - t_j)\right|.$$

Отсюда и (14) следует нужная оценка для $E(\tau, W^n_\infty([a,b]), t_1, \ldots, t_n)$ и оптимальность метода $\widehat{m}(\xi_1, \ldots, \xi_n) = \sum_{j=1}^n l_j(\tau)\xi_j$. $\triangleright$

Отметим, что если рассматривать задачу оптимальной интерполяции для случая, когда заданы n точек $a \leqslant t_1 < \ldots < t_n \leqslant b$, а класс $C = W^r_\infty([a,b])$, где $n \neq r$, то при $n < r$ задача не имеет смысла, так как можно построить, например, полином степени n (он, очевидно, будет принадлежать данному классу), который в точке τ принимает сколь угодно большие значения. Если же $n > r$, то задача осмысленна и здесь оптимальным методом будет интерполяционный сплайн. Доказательство получается на том же пути, что и в рассмотренном случае, но нужно воспользоваться еще утверждениями о существовании разного рода сплайнов.

Литература

1. *Алесеев В. М., Тихомиров В. М., Фомин С. В.* Оптимальное управление.—М.: Наука, 1979.—429 с.
2. *Магарил-Ильяев Г. Г., Тихомиров В. М.* Выпуклый анализ и его приложения.—М.: Эдиториал УРСС, 2003.—176 с.

Статья поступила 15 ноября 2004 г.

Магарил-Ильяев Георгий Георгиевич, д. ф.-м. н.
г. Москва, Московский государственный институт радиотехники,
электроники и автоматики (Технический университет)
E-mail: `georg@magaril.mccme.ru`