Международная конференция “Обратные и некорректные задачи математической физики”,
посвященная 75-летию академика М.М.Лаврентьева, 20-25 августа 2007 г., Новосибирск, Россия

Регуляризация в параметрических оптимизационных и обратных задачах на основе теории двойственности

М.И. Сумин*

* ННГУ
пр-т Гагарина, 23,
603095 Н.Новгород, Россия
E-mail: msumin@sinn.ru

Работа автора была поддержана РФФИ (гранты 04-01-00460, 07-01-00495)

Введение. Доклад посвящен изложению новых результатов, связанных с регуляризацией в задачах оптимизации и обратных задачах на основе теории двойственности [1] – [5]. В нем рассматривается параметрическая задача минимизации

$$(P_q^\delta) \qquad f^\delta(z) \to \min, \;\; A^\delta z = h^\delta + q, \;\; z \in \mathcal{D} \subset Z, \qquad q \in H- \text{параметр},$$

где верхний индекс $\delta \geq 0$, $\delta \in [0, \delta_0)$, $\delta_0 > 0$ – некоторое число, характеризует ошибку задания исходных данных и случай $\delta = 0$ означает, что исходные данные в задаче заданы точно. Здесь: $\mathcal{D}$ – ограниченное выпуклое замкнутое множество, Z, H – гильбертовы пространства, $f^\delta : Z \to R^1$ – сильно выпуклый функционал с не зависящей от δ постоянной сильной выпуклости, $A^\delta : Z \to H$ – линейный непрерывный оператор, $h^\delta \in H$ – заданный элемент, такие, что

$$|f^\delta(z) - f^0(z)| \leq C\delta(1 + \|z\|), \;\; \|A^\delta - A^0\| \leq C\delta, \;\; \|h^\delta - h^0\| \leq C\delta.$$

Единственное решение задачи (P_q^0), если оно существует, обозначим через z_q^0.

В зависимости от конкретного вида задачи (P_q^0) она может иметь смысл обратной задачи [1], уравнения первого рода в гильбертовом пространстве [2] или задачи оптимального управления [3].

Развитие методов регуляризации на основе теории двойственности позволяет дополнить существующие методы регуляризации оптимизационных задач (см., например, [6]) методами, в которых самым существенным образом используется классическая идея "снятия" ограничений, заложенная в принципе Лагранжа. Известно, что двойственные численные алгоритмы являются одними из самых эффективных для решения оптимизационных задач с ограничениями [6], [7]. Их типичным представителем является классический алгоритм Удзавы (см., например, [7]), заключающийся в непосредственном решении на основе градиентной процедуры задачи, двойственной к исходной оптимизационной задаче. Вопросы его сходимости изучались ранее лишь при двух весьма ограничительных обстоятельствах, одно из которых заключается в предположении точного задания исходных данных оптимизационной задачи, а другое предполагает наличие седловой точки соответствующего функционала Лагранжа (см., в частности, [8], [9]). Формальное же применение алгоритма в его классической форме может приводить и приводит к стандартным эффектам неустойчивости приближенного решения. Простейший пример такой неустойчивости, связанный с задачей минимизации сильно выпуклой квадратичной функции двух переменных на множестве, задаваемом аффинным ограничением типа равенства, эквивалентной задаче поиска нормального решения линейной алгебраической системы двух уравнений с двумя неизвестными можно найти в [3], [4]. Естественно, в случае когда в исходной задаче отсутствует седловая точка, то все классические теоремы сходимости [8], [9] теряют смысл, так как в них предполагается одновременная сходимость по прямой и двойственной переменным.

Рассмотрим два примера задач, в которых для функционала Лагранжа $L_q^0(z, \lambda) \equiv f^0(z) + \langle \lambda, A^0 z - h^0 - q \rangle$ задачи (P_q^0) справедливо равенство

$$\sup_{\lambda \in H} \inf_{u \in \mathcal{D}} L_q^0(z, \lambda) = \inf_{z \in \mathcal{D}} \sup_{\lambda \in H} L_q^0(z, \lambda),$$

но внешний экстремум в левой части не достигается (см. [1], [2], [4]).

Пример 1. Рассмотрим задачу минимизации

$$\|z\|^2 \to \min, \;\; A^0[z](x) \equiv \int_a^b K(x,s)z(s)\,ds = q(x), \;\; a \le x \le b, \;\; z \in L_2(a,b), \quad q \in L_2(a,b), \tag{1}$$

эквивалентную задаче поиска нормального решения интегрального уравнения Фредгольма первого рода с правой частью $q \in L_2(a,b)$

$$A^0[z](x) \equiv \int_a^b K(x,s)z(s)\,ds = q(x), \;\; a \le x \le b \tag{2}$$

с симметрическим непрерывным на квадрате $\Pi \equiv [a,b] \times [a,b]$ ядром K. Пусть ядро этого уравнения является замкнутым. Пусть далее $z_0 \in L_2(a,b)$ – функция, не являющаяся непрерывной (т.е. соответствующий класс эквивалентности не содержит непрерывной функции), $q_0(x) \equiv A^0[z_0](x)$, $x \in [a,b]$. Тогда рассмотрим уравнение (2) с $q(x) = q_0(x)$, $a \le x \le b$, имеющее по причине замкнутости ядра единственное решение $z^0_{q_0}$, и соответствующую эквивалентную задачу минимизации (1). Покажем, что соответствующая ей двойственная задача

$$V^0_{q_0}(\lambda) \equiv -1/4\langle A^0A^0\lambda, \lambda\rangle - \langle q_0, \lambda\rangle \to \max, \quad \lambda \in L_2(a,b)$$

не имеет решения, т.е. верхняя грань двойственной целевой функции не достигается. Действительно, если бы максимальное значение двойственной целевой функции достигалось в некоторой точке $\lambda \in L_2(a,b)$, то эта точка удовлетворяла бы равенству $-1/2A^0A^0\lambda = q_0$. Но тогда, согласно замкнутости ядра, мы должны были бы иметь и равенство $-1/2A^0\lambda = z_0$. Последнее же равенство противоречиво, так как $z^0_{q_0}$ – функция, не являющаяся непрерывной, в то время как $A^0\lambda$ – непрерывная функция (ядро K непрерывно). Последними рассуждениями, по сути дела, доказано также, что в задаче минимизации (1), эквивалентной уравнению (2) при $q = q_0$, принцип Лагранжа не выполняется.

Пример 2. Рассмотрим обратную задачу наблюдения, в которой требуется найти начальное условие $v(x)$, $x \in (0,1)$ по известному в финальный момент времени T решению $z[v](\cdot,T) = q \in L_2(0,1)$ третьей краевой задачи для уравнения теплопроводности

$$\partial z/\partial t - \partial^2 z/\partial x^2 = 0, \;\; z(x,0) = v(x), \;\; x \in (0,1),$$

$$\partial z(0,t)/\partial x - z(0,t) = 0, \;\; \partial z(1,t)/\partial x + z(1,t) = 0, \;\; t \in [0,T], \;\; v(x) \in V \equiv [-1,1].$$

Эквивалентная задача минимизации имеет здесь вид

$$\int_0^1 v^2(x)\,dx \to \inf, \quad z[v](\cdot,T) = q, \;\; v \in \mathcal{D} \equiv \{v \in L_2(0,1) : \, v(x) \in [-1,1] \text{ п.в. на } (0,1)\}. \tag{3}$$

Определим инъективный оператор $A^0 : L_2(0,1) \to L_2(0,1)$ формулой $A^0[v](\cdot) \equiv z[v](\cdot,T)$ и сопряженный $A^{0*} : L_2(0,1) \to L_2(0,1)$ формулой $A^{0*}[p](\cdot) \equiv \eta[p](\cdot,0)$, $p \in L_2(0,1)$, где $\eta[p]$ - решение сопряженной задачи

$$\partial \eta/\partial t + \partial^2 \eta/\partial x^2 = 0, \;\; \eta(x,1) = p(x), \;\; x \in (0,1),$$

$$\partial \eta(0,t)/\partial x - \eta(0,t) = 0, \;\; \partial \eta(1,t)/\partial x + \eta(1,t) = 0, \;\; t \in [0,T].$$

Введем функционал Лагранжа

$$L^0_q(v,\lambda) \equiv \int_0^1 v^2(x)\,dx + \int_0^1 \lambda(x)(z[v](x,T) - q(x))\,dx, \;\; v \in \mathcal{D}, \;\; \lambda \in L_2(0,1), \quad V^0_q(\lambda) \equiv \min_{v \in \mathcal{D}} L^0_q(v,\lambda).$$

Пусть $\bar{v} \in \mathcal{D}$ разрывная кусочно-непрерывная функция, $\bar{v}(x) \in (-1,1)$ и $q = q_{\bar{v}} = z[\bar{v}](\cdot,T)$. Можно утверждать, что двойственная к (3) задача при $q = q_{\bar{v}}$, имеющая вид $V^0_{q_{\bar{v}}}(\lambda) \to \max$, $\lambda \in L_2(0,1)$,

решения не имеет. Действительно, в случае существования такого решения λ^* в силу единственности элемента $v[\lambda^*]$, минимизирующего на $\mathcal{D}$ сильно выпуклый функционал Лагранжа $L^0_{q_{\bar{v}}}(\cdot,\lambda^*)$, на нем, в соответствии с правилами дифференцирования маргинальных функционалов (см., например, [1],[4]), занулялся бы градиент $\partial V^0_{q_{\bar{v}}}(\lambda^*) = z[v[\lambda^*]](\cdot,T) - q_{\bar{v}}$ и выполнялось бы равенство $z[v[\lambda^*]](\cdot,T) - q_{\bar{v}} = 0$, причем $v[\lambda^*] = \bar{v}$ в силу инъективности оператора A^0. Одновременно было бы справедливым необходимое условие оптимальности для $\bar{v}$

$$2\int_0^1 \bar{v}(x)(v(x)-\bar{v}(x))\,dx + \int_0^1 \eta[\lambda^*](x,0))(v(x)-\bar{v}(x))\,dx \geq 0 \;\forall\, v\in\mathcal{D},$$

т.е. элемент $2\bar{v}(\cdot) + \eta[\lambda^*](\cdot,0)$ являлся бы опорным к выпуклому множеству $\mathcal{D}$ в точке $\bar{v}$. Но тогда $2\bar{v}(\cdot) + \eta[\lambda^*](\cdot,0) = 0$, так как конус нормалей к $\mathcal{D}$ в точке $\bar{v}$ (в смысле выпуклого анализа) состоит лишь из нулевого элемента. Последнее же равенство противоречиво, так как благодаря "заглаженности" решений сопряженной задачи $\bar{v} \notin Im\,A^{0*}$. Очевидно, функции $\bar{v}$ указанного вида лежат всюду плотно в $\mathcal{D}$ и, как следствие, всюду плотно во множестве всех тех q, для которых разрешимо равенство $z[v](\cdot,T) = q,\ v\in\mathcal{D}$, лежат точки вида $q_{\bar{v}}$, для которых функционал $V^0_{q_{\bar{v}}}$ не достигает максимального значения на $L_2(0,1)$.

Подчеркнем, что в этих двух примерах всюду плотно во множестве всех значений параметра, для которых соответствующая задача (P^0_q) разрешима, лежат как значения q, при которых функционал Лагранжа этой задачи имеет седловую точку, так и значения, при которых функционал Лагранжа седловой точки не имеет.

Мы рассматриваем здесь лишь линейно выпуклые возмущения исходной линейно выпуклой задачи математического программирования. Тем не менее, подобные ситуации возникают во многих приложениях и, в частности, в приложениях, связанных в том числе и с нелинейными обратными задачами наблюдения для дифференциальных уравнений, в которых в результате измерений имеется информация о траектории системы "в целом" (см., в частности, [10]).

Пример 3. Пусть имеется управляемая нелинейная система

$$\dot{y} = f(t,y)u(t) + g(t,y),\ \ 0\leq t\leq T,\ \ y(0)=y_0,\ \ u(t)\in P\ \text{ п.в. на } (0,T),$$

где для определенности $f:\ R^1\times R^n\to R^{n\times r}$, $g:\ R^1\times R^n\to R^n$ – функции, удовлетворяющие глобальному условию Липшица по совокупности переменных, (см. условия теоремы 2.1 из [10, Гл.I]), $P\subset R^r$ – компакт. Если в нашем распоряжении имеется наблюдаемая траектория $y^0(t)$, $0\leq t\leq T$, и нужно определить "вызывающее" эту траекторию минимальное по норме управление $u^0\in\mathcal{D}\equiv\{u\in L_2(0,T):\ u(t)\in P\ \text{ п.в. на } (0,T)\}\subset L_2(0,T)$, то невозмущеннная линейно выпуклая задача (P^0_q) в этом случае имеет вид

$$f^0(u)\equiv\|u\|^2\to\min,$$

$$A^0[u](t)\equiv\int_0^t f(s,y^0(s))u(s)ds = h^0(t)\equiv y^0(t) - \int_0^t g(s,y^0(s))ds - y_0,\ t\in[0,T],\ u\in\mathcal{D}.$$

Соответственно возмущенная линейно выпуклая задача в случае приближенно наблюдаемой траектории $y^\delta(t)$, $|y^\delta(t)-y^0(t)|\leq\delta,\ 0\leq t\leq T$, записывается в том же виде, но с возмущенным аффинным ограничением

$$A^\delta[u](t)\equiv\int_0^t f(s,y^\delta(s))u(s)ds = h^\delta(t)\equiv y^\delta(t) - \int_0^t g(s,y^\delta(s))ds - y_0,\ t\in[0,T].$$

Регуляризация на основе теории двойственности. Двойственная регуляризация для задачи минимизации (P^0_q) заключается в непосредственном решении возмущенной двойственной к (P^0_q) и регуляризованной по Тихонову задачи

$$R^{\delta,\alpha}_q(\lambda)\equiv V^\delta_q(\lambda) - \alpha\|\lambda\|^2\to\sup,\ \ \lambda\in H,\quad V^\delta_q(\lambda)\equiv\min_{z\in\mathcal{D}} L^\delta_q(z,\lambda), \tag{4}$$

$$L_q^\delta(z,\lambda) \equiv f^\delta(z) + \langle \lambda, A^\delta z - h^\delta - q\rangle, \quad z^\delta[\lambda] \equiv \arg\min\{L_q^\delta(z,\lambda) : z \in \mathcal{D}\}, \quad \lambda_q^{\delta,\alpha} \equiv \arg\max\{R_q^{\delta,\alpha}(\lambda) : \lambda \in H\}.$$

При условии согласования $\delta/\alpha(\delta) \to 0$, $\alpha(\delta) \to 0$, $\delta \to 0$ и субдифференцируемости f^0 справедлива сходимость к решению невозмущенной задачи (P_q^0) в метрике Z: $z^\delta[\lambda_q^{\delta,\alpha(\delta)}] \to z_q^0$, $\delta \to 0$. Основой для конструирования двойственного алгоритма решения исходной оптимизационной задачи является наличие явного выражения для совпадающего с производной в смысле Фреше и удовлетворяющего в H условию Липшица супердифференциала вогнутого функционала V_q^δ при $\delta \geq 0$

$$\partial V_q^\delta(\lambda) = A^\delta z^\delta[\lambda] - h^\delta - q.$$

Итеративная регуляризация двойственного алгоритма. С целью конструирования сходящейся по аргументу минимизирующей последовательности в задаче (P_q^0) рассмотрим процедуру итеративной двойственной регуляризации [1] – [4] двойственного алгоритма (4) для задачи (P_q^0)

$$\lambda^{k+1} = \lambda^k + \beta^k \partial V_q^{\delta^k}(\lambda^k) - 2\beta^k\alpha^k\lambda^k, \;\; k = 1,2,\ldots; \;\; \lambda^1 \in H, \tag{5}$$

в которой

$$\|\partial V_q^\delta(\lambda) - \partial V_q^0(\lambda)\| \leq \tau^k(1 + \|\lambda\|), \;\; \tau^k \equiv C\sqrt{\delta^k},$$

а последовательности τ^k, α^k, β^k, $k = 1,2,\ldots$, удовлетворяют условиям согласования

$$\tau^k \geq 0, \; \alpha^k > 0, \; \beta^k > 0, \; \lim_{k\to\infty}(\tau^k + \alpha^k + \beta^k) = 0, \; \frac{\alpha^k}{\alpha^{k+1}} \leq C_0, \tag{6}$$

$$\frac{|\alpha^{k+1} - \alpha^k|}{(\alpha^k)^3\beta^k} \leq C, \;\; \frac{\beta^k}{(\alpha^k)^3} \leq C, \;\; \frac{\tau^k}{(\alpha^k)^3} \leq C, \;\; \sum_{k=1}^{\infty} \alpha^k\beta^k = +\infty.$$

При согласованном в смысле (6) стремлении к нулю величины τ^k, шагового множителя β^k и параметра регуляризации α^k имеет место сильная сходимость в метрике Z приближенных решений $z^{\delta^k}[\lambda^k]$ к решению z_q^0 невозмущенной задачи (P_q^0). Справедлива [1], [4]

Теорема 1. Имеют место предельные соотношения

$$f^0(z^{\delta^k}[\lambda^k]) \to f^0(z_q^0), \quad A^0 z^{\delta^k}[\lambda^k] - h^0 - q \to 0, \;\; k \to \infty,$$

а при дополнительном условии субдифференцируемости f^0 (в смысле выпуклого анализа) в точках $\mathcal{D}$ и предельное соотношение

$$\|z^{\delta^k}[\lambda^k] - z_q^0\| \to 0, \;\; k \to \infty.$$

Другими словами, вне зависимости от того разрешима или нет двойственная задача, итерационная процедура (5), (6) представляет собой регуляризирующий алгоритм.

На основе случая с ограниченным множеством $\mathcal{D}$ конструируется ограниченная по норме минимизирующая последовательность в задаче (P_q^0) и при неограниченном $\mathcal{D}$. Обсуждается как алгоритм двойственной регуляризации с его итеративным вариантом распространяется и на случай задачи (P_q^0), в которой целевой функционал f^0 и его возмущение f^δ являются лишь выпуклыми на $\mathcal{D}$. В этом случае задача (P_q^0) с выпуклым целевым функционалом аппроксимируется по Тихонову задачами с сильно выпуклыми функционалами цели

$$(P_q^{0,\epsilon}) \qquad f^0(z) + \epsilon\|z\|^2 \to \min, \;\; A^0 z = h^0 + q, \;\; z \in \mathcal{D} \subset Z, \;\; q \in H- \text{параметр}, \quad \epsilon > 0,$$

а соответствующая сильная сходимость регуляризованных решений при согласованном стремлении к нулю параметров δ, α, ϵ происходит к нормальному решению задачи (P_q^0).

Обсуждается проблема останова итерационного процесса (6) при фиксированной конечной ошибке задания исходных данных. Обсуждается также возможность распространения идеологии двойственной регуляризации на нелинейные задачи [5].

Параметрическая двойственная регуляризация. Сходимость регуляризованного двойственного алгоритма (4) и его итеративного варианта (5), (6) имеет место вне зависимости от того, разрешима или нет, невозмущенная двойственная задача $V_q^0(\lambda) \to \sup$, $\lambda \in H$. Однако, "качество" его сходимости напрямую зависит от дифференциальных свойств функции значений задачи (P_q^0) как функции параметра q [1], [2], [3]. Оказывается, что множество Q всех значений параметра, для которых эта двойственная задача разрешима, плотно во множестве всех тех значений q, для которых разрешима невозмущенная задача (P_q^0) (см., в частности, примеры 1,2). При этом, в случае неразрешимости двойственной задачи, последовательность λ^k, $k = 1, 2, \ldots$, генерируемая двойственным алгоритмом, неограничена по норме. В случае же разрешимости двойственной задачи или, другими словами, в случае существования седловой точки функционала Лагранжа исходной невозмущенной задачи (P_q^0) последовательность λ^k, $k = 1, 2, \ldots$, является ограниченной по норме и условия согласования несколько менее жесткие нежели условия (6). Указанное множество Q можно охарактеризовать, введя в рассмотрение выпуклый функционал значений $\beta^0 : H \to R^1$ невозмущенной задачи (P_q^0): $\beta^0(q) \equiv \{f^0(z_q^0)$, если z_q^0 существует; $+\infty$ в противном случае$\} = \min\limits_{z\in\mathcal{D}} \max\limits_{\lambda\in H} L_q^0(z,\lambda)$. Можно показать, что $\beta^0(q) = \sup\limits_{\lambda\in H} \min\limits_{z\in\mathcal{D}} L_q^0(z,\lambda)$. Тогда можно утверждать, что множество Q совпадает со множеством тех q, для которых не пуст субдифференциал (в смысле выпуклого анализа) $\partial\beta^0(q)$, что, в свою очередь, равносильно выполнимости в задаче (P_q^0) регулярного принципа Лагранжа. Именно в этом и только в этом случае, т.е. при $\lambda \in \partial\beta^0(q)$, имеет место равенство $\partial V_q^0(\lambda) = 0$ или $-q = \partial(-V_0^0)(\lambda)$, которое, в свою очередь, выполняется тогда и только тогда, когда имеет место включение $\lambda \in \partial(-V_0^0)^*(-q)$, где $(-V_0^0)^*$ функция, сопряженная (по Фенхелю) к функции $-V_0^0$.

Список литературы

1. Сумин М. И. Регуляризованный градиентный двойственный метод решения обратной задачи финального наблюдения для параболического уравнения // Ж. вычисл. матем. и матем. физ. 2004. Т. 44. № 11. С. 2001–2019.

2. Сумин М. И. Итеративная регуляризация градиентного двойственного метода для решения интегрального уравнения Фредгольма первого рода // Вестник Нижегородского университета. Серия Математика. 2004. Вып. 1(2). С. 192–208.

3. Сумин М. И. Регуляризованный двойственный алгоритм в задачах оптимального управления для распределенных систем // Вестник Нижегородского университета. Серия Математическое моделирование и оптимальное управление. 2006. Вып. 2(31). С. 82–102.

4. Сумин М.И. Регуляризация в линейно выпуклой задаче математического программирования на основе теории двойственности// Ж. вычисл. матем. и матем. физ. 2007. Т. 47. № 4. С. 602-625.

5. Сумин М.И. Регуляризованный двойственный метод решения нелинейной задачи математического программирования// Ж. вычисл. матем. и матем. физ. 2007. Т. 47. № 5. С. 794-814.

6. Васильев Ф. П. Методы оптимизации. М.: Факториал Пресс, 2002.

7. Мину М. Математическое программирование. Теория и алгоритмы. М.: Наука, 1990.

8. Экланд И., Темам Р. Выпуклый анализ и вариационные проблемы. М.: Мир, 1979.

9. Гловински Р., Лионс Ж.-Л., Тремольер Р. Численное исследование вариационных неравенств. М.: Мир, 1979.

10. Осипов Ю.С., Васильев Ф.П., Потапов М.М. Основы метода динамической регуляризации. М.: Изд-во МГУ, 1999.