Итеративная аппроксимация для задач выпуклой оптимизации с операторными ограничениями в гильбертовом пространстве [1]

© 2000 г. М.Р. Давидсон, Н.М. Новикова

117967 ГСП-1, Москва, ул. Вавилова, 40, ВЦ РАН

Поступила в редакцию 20.08.99 г.

Рассмотрена задача поиска значения и реализации минимума интегрального функционала в гильбертовом пространстве при наличии операторных ограничений. Предложен способ итеративной аппроксимации задачи, сводящий ее к последовательности конечномерных задач с бесконечным числом ограничений (т.н. полубесконечной оптимизации), для которых разработан регуляризованный метод агрегирования ограничений. На основе указанного метода, комбинированного с итеративным увеличением порядка аппроксимации, построены алгоритмы решения исходной бесконечномерной задачи в сильно выпуклом и выпуклом случаях. Обоснована их сильная сходимость.

1. Рассматривается задача поиска значения J^0 и реализации u^0 минимума

$$\min_{u\in U} J(u), \quad J(u) = \int\limits_X \Phi(u)[x]dx, \tag{1}$$

$$U = \{u \in U' \mid G(u)[x] \le 0 \text{ п.н. } \forall x \in X, \quad H(u)[y] \le 0 \text{ п.н. } \forall y \in \varGamma\}, \tag{2}$$

где $X \subset \mathbb{R}^n$ — ограниченная область с кусочно-гладкой границей $\varGamma$, U' — выпуклый слабый компакт (например, шар) в гильбертовом пространстве W, непрерывно вложенном в $\mathbb{L}_2(X)$, функция $\Phi(u)[\cdot]$ интегрируема на X $\forall u \in W$, функционалы $\Phi(\cdot)[x]$, $G(\cdot)[x]$ (и $H(\cdot)[y]$) — непрерывные, выпуклые в W п.н. $\forall x \in X$ ($y \in Y$), термин п.н. (почти наверное) здесь и далее понимается в смысле меры Лебега на соответствующем множестве из евклидова пространства.

В этих условиях целевой функционал J является непрерывным, выпуклым, множество ограничений U — выпуклый слабый компакт в W, и минимум в залаче (1),(2) достигается [1]. В случае сильной выпуклости функционала J предположение ограниченности U можно снять, а в дальнейших рассуждениях

[1] Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 99-01-01192) и международной ассоциации INTAS (код проекта INTAS 97-1050).

будем считать, что U' принадлежит шару радиуса R в W.

Нелинейные задачи такого типа возникают, в частности, в математической физике [2], [3], например, квадратичный функционал J (имеющий смысл энергии):

$$\Phi(u) = \sum_{r,s=1}^{n} d_{rs} \frac{\partial u}{\partial x_r} \frac{\partial u}{\partial x_s} + d_0 u^2 - 2fu,$$

где $d_{rs} \in \mathbf{C}^1(X)$, $d_0 \in \mathbf{C}(X)$ — коэффициенты положительно-определенной квадратичной формы, $f \in \mathbb{L}_2(X)$; и нелинейные ограничения U, заданные в операторной форме:

$$U^1 = \{u \in \overset{\circ}{\mathbb{W}}{}_2^1(X) : \ |\nabla u| \le 1 \text{ п.н. } \forall x \in X\}$$

(задача упругопластического кручения) или

$$U^2 = \{u \in \mathbb{W}_2^1(X) : \ u \ge 0 \text{ п.н. } \forall y \in \Gamma\}$$

(задача о течении жидкости через полупроницаемую мембрану). Здесь и далее $\mathbb{W}_2^r(X)$ — пространство Соболева r-го порядка, $\overset{\circ}{\mathbb{W}}{}_2^r(X)$ — пространство функций из $\mathbb{W}_2^r(X)$, след которых на Γ равен нулю. Для негладких u производные понимаются в обобщенном смысле, а выполнение условий, заданных на границе, предполагается для следа (см. [3], [4]). Для упрощения изложения допустим, что условия на границе учитываются путем выбора подходящей аппроксимации пространства W, и будем опускать оператор H в (2). Общий случай рассматривается аналогично.

Пусть известна ортонормированная система функций $\{\xi_j\}_{j=1}^{\infty}$ на X (координатная система, или базис ξ), для которой

$$\forall l = 1, 2, \ldots, \quad \forall a \in \mathbb{R}^l \quad \sum_{j=1}^{l} a_j \xi_j \in W$$

и

$$\forall u \in W \quad \forall \varepsilon > 0 \ \ \exists l^0 : \quad \forall l \ge l^0 \ \ \exists a \in \mathbb{R}^l \quad \| \sum_{j=1}^{l} a_j \xi_j - u \| \le \varepsilon,$$

специальным индексом норму в основном пространстве, для нормы в евклидовом пространстве будем использовать знак модуля, а сумму $\sum_{j=1}^{l} a_j\xi_j$ будем для краткости записывать как $a\xi$, $l = 1, 2, \dots$

Обозначим

$$U_l \stackrel{\mathrm{def}}{=} \{u_l = a\xi |\ a \in \mathbb{R}^l\}, \quad [u]_l \stackrel{\mathrm{def}}{=} \mathrm{pr}_{U_l} u = \sum_{j=1}^{l} a_j\xi_j,$$

где pr — оператор проектирования в соответствующем гильбертовом пространстве, a — вектор коэффициентов Фурье для u.

Считаем, что для решения u^0 задачи (1), (2) заведомо известно $[u^0]_l \in U'\ \ \forall l$, и неоператорные ограничения U' порождают в $\mathbb{R}^l$ достаточно простые множества $A'_l \stackrel{\mathrm{def}}{=} \{a \in \mathbb{R}^l |\ a\xi \in U'\}$, например, $A'_l = \{a \in \mathbb{R}^l\colon\ |a| \le R\}$ (при этом для вектора a^0 коэффициентов Фурье решения u^0 также выполнено включение $(a_1^0, \dots, a_l^0) \in A'_l\ \ \forall l = 1, 2, \dots$). Относительно операторных ограничений предположим, что $\mathrm{int} U \ne \emptyset$. Тогда $U \cap U_l \ne \emptyset\ \ \forall l \ge l_0$ и последовательность $\{u_l^0\}$ реализаций

$$\min_{U \bigcap U_l} J(u_l) \tag{3}$$

является минимизирующей, т.е. $J(u_l^0)$ сходится к J^0. Здесь в принципе возможны две ситуации.

Первая, когда известна зависимость $l(\varepsilon)$ такая, что

$$\forall \varepsilon > 0, \ \ \forall l \ge l(\varepsilon) \quad \|u^0 - u_l^0\| \le \varepsilon.$$

Если нас интересует только значение J^0, а функцию $u^0(\cdot)$ искать не обязательно, то достаточно знать зависимость $l'(\varepsilon)$:

$$\forall \varepsilon > 0,\ \forall l \ge l'(\varepsilon) \quad |J(u_l^0) - J^0| \le \varepsilon.$$

В этой ситуации можно, задавшись нужной точностью $\varepsilon > 0$, выбрать соответствующее l и задача (1), (2) сведется к решению (3).

Вторая, когда конкретный вид зависимости l, l' от ε не известен. В этом — наиболее распространенном — случае однократное решение задачи (3), вообще (например, удваивая) l и сравнивая между собой полученные решения. Однако с ростом l растет размерность задачи (3) и, следовательно, число итераций, требующееся для поиска приближенного решения, так что указанный способ может не привести к результату без дополнительного согласования параметров аппроксимации с параметрами метода решения аппроксимированной задачи. В данной работе подобное согласование будет проведено. Будет предложена схема, предполагающая итеративное увеличение l в рамках конкретной процедуры вычисления (3), для построения алгоритмов, сходящихся к решению (1), (2) в метрике пространства W.

За счет наличия операторных ограничений (2) в исходной постановке (1), (2) ее аппроксимация (3) оказывается задачей с бесконечным числом ограничений, т.е. задачей *полубесконечной оптимизации* [5],[6]. В настоящее время существует большое количество численных методов полубесконечной оптимизации [7]–[15]. В качестве базовой процедуры для комбинирования с итеративным увеличением размерности пространства будем использовать регуляризованный метод агрегирования ограничений [16], распространяющий на бесконечное число ограничений идею алгоритма агрегирования, предложенного в [17]. В применении к задаче (3) этот метод состоит в следующем.

Выберем произвольное начальное приближение $a^1 \in \mathbb{R}^l$ так, чтобы $a^1\xi \in U'$, и определим $\forall k = 1, 2, \ldots$ приближение a^{k+1} как реализацию

$$\min_{a \in A_k} \{\tau_k J(a\xi) + |a - a^k|^2/2\}, \tag{4}$$

где в случае $a^k\xi \in U$ полагаем $A_k = A'_l \stackrel{\text{def}}{=} \{a \in \mathbb{R}^l |\ a\xi \in U'\}$, а в противном случае A_k — агрегированное множество ограничений. Для его выбора предлагаются три варианта:

$$A_k = A^1_k \stackrel{\text{def}}{=} \{a \in A'_l |\ \int\limits_X G^+(a^k\xi)[x] G(a\xi)[x] dx \le 0\} \tag{5}$$

(верхний индекс “+” обозначает положительную срезку, т.е. максимум из нуля и значения функции), при этом считаем, что интеграл в (5) существует $\forall a \in A'_l$;

$$A_k = A^2_k \stackrel{\text{def}}{=} \{a \in A'_l |\ G(a\xi)[x^{*k}] \le 0\},$$

$$A_k = A_k^3 \stackrel{\text{def}}{=} \{a \in A_l' |\ G(a\xi)[x^k] \le 0\},$$

где x^k — случайная реализация равномерного распределения на X. Выбор τ_k подчинен условиям

$$\tau_k \downarrow 0, \quad \sum \tau_k = \infty. \tag{6}$$

Здесь и далее опускаем индексы суммирования по k от 1 до ∞, а также запись $k \to \infty$ под знаком соответствующего предела.

Очевидно, что все варианты для A_k дают, вообще говоря, существенно более широкие множества, чем определяемое исходными ограничениями $U \cap U_l$

$$A_l^0 \stackrel{\text{def}}{=} \{a \in A_l' |\ \int\limits_X \{G^+(a\xi)[x]\}^2 dx = 0\},$$

и такие, чтобы поиск решения для задачи (4) был проще, чем для задачи (3).

В соответствии с указанными вариантами можно строить три различных алгоритма решения задачи (3) (см. [16],[18]). Для вариационной задачи (1),(2) рассмотрим алгоритм (4),(5), который будем совмещать с увеличением l. А именно, на шаге инициализации алгоритма выберем $l = l_0$. Затем в (4),(5) положим $l = l_k$ и вместо вектора $a^k \in \mathbb{R}^{l_{k-1}}$ будем использовать расширенный вектор $\bar{a}^k \in \mathbb{R}^{l_k}$, дополненный в случае необходимости нулями до требуемой размерности. Скорость роста $l_k \uparrow \infty$ будет далее согласована с выбором управляющей последовательности $\tau_k \downarrow 0$.

Обозначим через v^k проекцию решения u^0 на $U_{l_{k-1}} \cap U$ и через p^k — вектор коэффициентов Фурье v^k, т.е. $v^k = p^k \xi$. Введем также расширенный вектор $\bar{p}^k$:

$$\bar{p}_i^k = p_i^k \quad \forall i \le l_{k-1}, \quad \bar{p}_i^k = 0, \quad l_{k-1} < i \le l_k.$$

Лемма 1. *Имеем* $\sum |p^{k+1} - \bar{p}^k|^2 = \sum ||v^{k+1} - v^k||^2 < \infty$.

Доказательство. Из ортонормальности базиса $|p^{k+1} - \bar{p}^k| = ||v^{k+1} - v^k||$. А поскольку $U_{l_k} \subseteq U_{l_{k+1}}$ и v^{k+1} — ортогональная проекция u^0 на $U_{l_k} \cap U$, то можем записать

$$\langle u^0 - v^{k+1},\ v^k - v^{k+1} \rangle_W \le 0,$$

выводим оценку

$$||v^{k+1}-u^0||^2 \leq ||v^k-u^0||^2 - ||v^{k+1}-v^k||^2,$$

из которой суммированием по k получаем утверждение леммы.

Теперь сформулируем следующее условие согласования l_k и τ_k:

$$\tau_k \downarrow 0, \quad l_k - \text{натуральное}, \quad l_k \uparrow \infty, \quad |p^{k+1}-\bar{p}^k| = o(\tau_k). \tag{7}$$

Замечание 1. Как видно из леммы 1, выполнение условия (7) всегда может быть обеспечено за счет выбора достаточно медленного убывания τ_k ($\sum \tau_k = \infty$) в сравнении с ростом l_k. В частности, если известно, что $[u^0]_{l_k} \in U$ (аппроксимированное решение удовлетворяет ограничениям), то условие (7) превращается в $|a^0_{l_k}| = o(\tau_k)$, где a^0 — вектор коэффициентов Фурье решения, следовательно $a^0 \in \mathbb{l}_2$ ($\sum\limits_{i=1}^{\infty} {a_i^0}^2 < \infty$). А если при этом $u^0 \in \mathbb{W}_2^1(X)$, то можно так выбрать базис в $\mathbb{L}_2(X)$, что будет справедливо и $\sum\limits_{i=1}^{\infty}(ia_i^0)^2 < \infty$, откуда $a_i^0 = o(1/i)$, значит, достаточно взять $\tau_k \geq O(1/l_k)$. Упомянутый выбор базиса (подробнее см. в [13]) позволяет получить явную оценку для τ_k, гарантирующую (7) и в случае $[u^0]_{l_k} \notin U, \;\; l_k \to \infty$, если предположить липшицевость G^+ и введенную в [19], [20] ρ-регулярность U с показателем $\theta \geq 1$ (см. ниже в разделе 3), рассматривая в качестве основного пространства $W = \mathbb{L}_2(X)$. Тогда имеем $l\varepsilon_l \to 0$ для $\varepsilon_l \stackrel{\text{def}}{=} \|u^0-[u^0]_l\|, \;\; \|G^+([u^0]_l)\| \leq L_G\|u^0-[u^0]_l\| = L_G\varepsilon_l$ и $\|v^k-[u^0]_{l_k}\|^\theta \leq O(\|G^+([u^0]_{l_k})\|^2)$, в результате

$$|p^{k+1}-\bar{p}^k|^2 \leq ||v^k-u^0||^2 - ||v^{k+1}-u^0||^2 \leq ||v^k-u^0||^2 \leq ||v^k-[u^0]_{l_k}||^2 + \|u^0-[u^0]_{l_k}\|^2 \leq$$

$$\leq O(\varepsilon_{l_k}^{4/\theta}) + \varepsilon_{l_k}^2 = o(l_k^{-4/\theta}+l_k^{-2}),$$

что дает для выполнения (7) неравенство $\tau_k \geq O(l_k^{-2/\theta}+l_k^{-1})$. При этом на выбор l_k условие (7) ограничений не накладывает.

Из выпуклости J, G по a и компактности A'_l при любом l (в силу слабой компактности U' в W) следует, что $\forall l = 1,2,\ldots$ аппроксимации $J(a\xi)$ и $G(a\xi)[x]$, константами $L_J(l)$ и, соответственно, $L'_G(l)[x]$ $\forall x \in X$, зависящими от l. Предположим, что

$$L_G^2(l) \stackrel{\text{def}}{=} \int\limits_X \{L'_G(l)[x]\}^2 dx < \infty.$$

(Указанное свойство обеспечивается достаточной гладкостью базисных функций ξ_i.) Введем также оператор $G^+(\cdot)$ со значениями $G^+(u)[x]$ и будем считать, что $G^+(\cdot)\colon W \to \mathbb{L}_2(X)$.

2. Допустим для начала, что целевой функционал J строго равномерно выпуклый [1, с.56]. Тогда задача (1),(2) корректна по Тихонову и любая минимизирующая последовательность $\{u^k\}$ оказывается сильно сходящейся к u^0 (см. [1, с.57]). Требуемую последовательность будем строить в виде $u^k \stackrel{\text{def}}{=} a^k\xi$, где a^k определяется с помощью описанного выше комбинирования процедуры (4),(5) с итеративной аппроксимацией.

Алгоритм 1

Зададимся начальным приближением $a^1 \in \{a \in \mathbb{R}^{l_0} |\ a\xi \in U'\}$.

Выберем управляющие последовательности $\{\tau_k\}$ и $\{l_k\}$, удовлетворяющие условиям (6),(7) и

$$\tau_k \downarrow 0, \quad l_k - \text{натуральное}, \quad l_k \uparrow \infty, \quad \lim \tau_k L_G(l_k) L_J(l_k) = 0, \tag{8}$$

в частности, при $L_G^2(l)$, $L_J(l)$ не зависящих от l (например, когда G, J — липшицевы в W), можно взять

$$l_k = \lfloor \log_2(k+k_0) \rfloor, \quad \tau_k = \Big[\log_2 \log_2(k+k_1)\Big]^{-1}. \tag{9}$$

Здесь и далее $\lfloor\cdot\rfloor$ — целая часть числа в скобках, $k_0, k_1, \ldots$ — целые > 2.

Положим $\forall k = 1, 2, \ldots$

$$\bar a_i^k = a_i^k \ \ \forall i \le l_{k-1}, \quad \bar a_i^k = 0,\ l_{k-1} < i \le l_k,$$

$$a^{k+1} = \arg\min_{a \in A_k}\Big[\tau_k J(a\xi) + |a - \bar a^k|^2/2\Big], \tag{10}$$

$$A_k = \{a \in \mathbb{R}^{l_k} |\ a\xi \in U',\ \int\limits_X G^+(\bar a^k\xi)[x] G(a\xi)[x] dx \le 0\}. \tag{11}$$

полнения условий (7) и (8) добиться не так просто, поскольку первое требует медленного убывания τ_k в сравнении с ростом l_k, а второе — наоборот. Так, при $L_G^2(l)$, $L_J(l) = O(l^s)$, руководствуясь замечанием 1, получаем $\tau_k \geq O(l_k^{-2/\theta} + l_k^{-1})$, а из (8) следует $\tau_k l_k^{2s} = o(1)$, что предполагает $s < \min(1/\theta, 1/2)$. Поэтому (8) представляется довольно жестким условием. Укажем предположения, гарантирующие возможность конструктивного выбора управляющих последовательностей в алгоритме 1: G и J липшицевы в $W = \mathbb{W}_2^r(X)$, $u^0 \in \mathbb{W}_2^{r+1}(X)$, ограничения (2) ρ-регулярны с некоторым показателем. Тогда выбор гладкого базиса в W, например, тригонометрической системы, обеспечивает независимость L_G^2 и L_J от l, а также справедливость оценки $a_i^0 = o(1/i)$, из которой, как и в замечании 1, выводим достаточность (9) для (7),(8).

Исследуем сходимость алгоритма 1. Следующая лемма дает оценку невязки ограничений через расстояние между предыдущим и текущим приближениями. Будем использовать обозначение $g_k(a, x)$ для $G(a\xi)[x]$ при $a \in \mathbb{R}^{l_k}$.

Лемма 2. *Для всех* $k = 1, 2, \ldots$ *выполнено*

$$\int\limits_X [g_k^+(a^{k+1}, x)]^2 dx + \int\limits_X [g_k^+(\bar{a}^k, x)]^2 dx \leq L_G^2(l_k)|a^{k+1} - \bar{a}^k|^2.$$

Доказательство. По определению функции $([\cdot]^+)^2$ непосредственной проверкой (при всех вариантах знака аргумента) убеждаемся, что

$$([t + \Delta t]^+)^2 \leq (t^+)^2 + 2t^+\Delta t + (\Delta t)^2.$$

Таким образом, можем записать

$$[g_k^+(a^{k+1}, x)]^2 \leq [g_k^+(\bar{a}^k, x)]^2 + 2g_k^+(\bar{a}^k, x)\Big(g_k(a^{k+1}, x) - g_k(\bar{a}^k, x)\Big) +$$

$$+ \Big(g_k(a^{k+1}, x) - g_k(\bar{a}^k, x)\Big)^2 \quad \forall x \in X.$$

Интегрируя обе части неравенства по x и помня, что a^{k+1} удовлетворяет агрегированным ограничениям, получаем

$$\int\limits_X [g_k^+(a^{k+1}, x)]^2 dx + \int\limits_X [g_k^+(\bar{a}^k, x)]^2 dx \leq \int\limits_X \Big(g(a^{k+1}, x) - g(\bar{a}^k, x)\Big)^2 dx.$$

Теорема 1. *В сделанных предположениях последовательность* $\{u^k = a^k\xi\}$, *построенная по алгоритму* 1, *сильно сходится к решению* u^0 *задачи* (1),(2).

Доказательство. Докажем, что $|a^k - \bar{p}^k| \to 0$, т.е. $||u^k - v^k|| \to 0$, откуда и будет следовать утверждение теоремы. Действительно, $||u^k - u^0|| \leq ||u^k - v^k|| + ||v^k - u^0||$, а $||v^k - u^0|| \to 0$, поскольку $\{v^k\}$ — минимизирующая последовательность в задаче минимизации $||u - u^0||^2$ на U.

Для конечномерной задачи (10),(11) применение теоремы Куна — Таккера дает

$$\langle \tau_k \mathrm{grad} J(a^{k+1}\xi) + a^{k+1} - \bar{a}^k,\ a - a^{k+1}\rangle \geq 0 \quad \forall a \in A_k.$$

Здесь и далее grad обозначает обобщенный градиент по $a \in \mathbb{R}^{l_k}$ выпуклой функции, $\langle \cdot, \cdot \rangle$ — скалярное произведение в $\mathbb{R}^{l_k}$. Положим $a = \bar{p}^k$, тогда

$$\langle \tau_k \mathrm{grad} J(a^{k+1}\xi),\ \bar{p}^k - a^{k+1}\rangle + \langle a^{k+1} - \bar{a}^k,\ \bar{p}^k - a^{k+1}\rangle \geq 0$$

$$\text{и} \quad \langle a^{k+1} - \bar{a}^k,\ a^{k+1} - \bar{p}^k\rangle \leq \tau_k \Big(J(\bar{p}^k\xi) - J(a^{k+1}\xi) \Big)$$

по свойству обобщенных градиентов (субградиентов). Имеем

$$|a^{k+1} - \bar{p}^k|^2 = |a^k - p^k|^2 + 2\langle a^{k+1} - \bar{a}^k,\ \bar{a}^k - \bar{p}^k\rangle + |a^{k+1} - \bar{a}^k|^2. \tag{12}$$

Второе слагаемое в (12) может быть переписано как

$$\langle a^{k+1} - \bar{a}^k,\ \bar{a}^k - \bar{p}^k\rangle = -|a^{k+1} - \bar{a}^k|^2 + \langle a^{k+1} - \bar{a}^k,\ a^{k+1} - \bar{p}^k\rangle,$$

и с учетом предыдущей оценки из (12) выводим

$$|a^{k+1} - \bar{p}^k|^2 = |a^k - p^k|^2 + 2\langle a^{k+1} - \bar{a}^k,\ a^{k+1} - \bar{p}^k\rangle - |a^{k+1} - \bar{a}^k|^2 \leq$$

$$\leq |a^k - p^k|^2 + 2\tau_k \Big(J(\bar{p}^k\xi) - J(a^{k+1}\xi) \Big) - |a^{k+1} - \bar{a}^k|^2.$$

Теперь на основе равенства

$$|a^{k+1} - p^{k+1}|^2 = |a^{k+1} - \bar{p}^k|^2 + |p^{k+1} - \bar{p}^k|^2 + 2\langle a^{k+1} - \bar{p}^k,\ \bar{p}^k - p^{k+1}\rangle$$

окончательно получаем

$$|a^{k+1} - p^{k+1}|^2 \leq |a^k - p^k|^2 + 2\tau_k \Big(J(\bar{p}^k\xi) - J(a^{k+1}\xi) \Big) - |a^{k+1} - \bar{a}^k|^2 + \gamma_k, \tag{13}$$

Условие (7) потребовалось для обеспечения $\gamma_k = o(\tau_k)$. Из этого и из условия (6) найдется бесконечно малая $\delta_k > 0$, для которой

$$\delta_k \downarrow 0, \quad \sum \delta_k = +\infty, \quad \delta_k = o(\tau_k), \quad \gamma_k = o(\delta_k).$$

Соответственно такой последовательности введем два индексных множества:

$$K_- \stackrel{\text{def}}{=} \{k = 1, 2, \ldots \mid 2\tau_k(J(a^{k+1}\xi) - J(\bar{p}^k\xi)) + |a^{k+1} - \bar{a}^k|^2 - \gamma_k \geq \delta_k\}$$

$$\text{и} \quad K_+ \stackrel{\text{def}}{=} \{k = 1, 2, \ldots\} \setminus K_-.$$

Последнее множество будет бесконечным. Действительно, предположив, что в K_+ максимальный элемент меньше N, получим

$$\forall k \geq N \quad |a^{k+1} - p^{k+1}|^2 \leq |a^k - p^k|^2 - \delta_k,$$

но это противоречит расходимости ряда из δ_k.

Для индексов $k \in K_+$ выполнено

$$2(J(a^{k+1}\xi) - J(\bar{p}^k\xi)) + |a^{k+1} - \bar{a}^k|^2/\tau_k \leq (\delta_k + \gamma_k)/\tau_k.$$

По лемме 2 и определению δ_k, отсюда заключаем, что

$$J(a^{k+1}\xi) - J(\bar{p}^k\xi) + \left(2L_G^2(l_k)\tau_k\right)^{-1} \int\limits_X [g_k^+(a^{k+1}, x)]^2 dx \leq o(1), \quad k \to \infty, \quad k \in K_+. \tag{14}$$

А с учетом липшицевости J по a и ограниченности $|a^{k+1}|$, $|p^k| \leq R$ приходим для $k \in K_+$ к оценке

$$\left(2L_G^2(l_k)\tau_k\right)^{-1} \int\limits_X [g_k^+(a^{k+1}, x)]^2 dx \leq 2L_J(l_k)|a^{k+1} - \bar{p}^k| + o(1) \leq 4L_J(l_k)R + o(1),$$

из которой, пользуясь (8), выводим неравенство для интеграла

$$\int\limits_X [g_k^+(a^{k+1}, x)]^2 dx \leq o(1), \quad \text{дающее} \quad \lim_{k \in K_+} \int\limits_X [g_k^+(a^{k+1}, x)]^2 dx = 0.$$

Теперь рассмотрим соответствующий функционал

$$S(u) \stackrel{\text{def}}{=} \int\limits_X \{G^+(u)[x]\}^2 dx$$

лунепрерывным снизу на W. Поэтому для любой слабой предельной точки $\hat{u}$ последовательности $\{u^{k+1}\}_{k\in K_+}$ справедливо

$$S(\hat{u}) \le \liminf_{k\in K_+} S(u^{k+1}) = \lim_{k\in K_+} \int_X [g_k^+(a^{k+1}, x)]^2 dx = 0,$$

т.е. $S(\hat{u}) = 0$, $\hat{u} \in U$ и $J(\hat{u}) \ge J^0$. Принимая во внимание слабую полунепрерывность снизу целевого функционала J, далее получаем

$$J^0 \le J(\hat{u}) \le \liminf_{k\in K_+} J(u^{k+1}).$$

Отметим, что $J(v^k) \to J^0$ (ибо $||v^k - u^0|| \to 0$, а J непрерывен); следовательно, полученное неравенство влечет за собой

$$\liminf_{k\in K_+}\{J(u^{k+1}) - J(v^k)\} \ge 0.$$

Но тогда из (14) вытекает

$$J(a^{k+1}\xi) - J(\bar{p}^k\xi) \to 0, \quad \text{т.е.} \quad J(u^{k+1}) \to J^0, \quad k \in K_+.$$

Таким образом, последовательность $\{a^{k+1}\xi\}_{k\in K_+}$ минимизирующая и, как минимизирующая, сходится к u^0. Поэтому и $|a^{k+1} - p^{k+1}|^2 \to 0$, $k \in K_+$.

В результате можем переписать (13) в форме

$$|a^{k+1} - p^{k+1}|^2 \le \max[\kappa(\delta_k),\ |a^k - p^k|^2 - \delta_k]$$

(1-й элемент соответствует $k \in K_+$, а 2-й соответствует $k \in K_-$), где $\kappa(\delta_k) \to 0$ при $\delta_k \to 0$. А это есть известное неравенство Дермана — Сакса [21], из которого следует $|a^k - p^k|^2 \to 0$, $k \to \infty$. Так что $u^k \to u^0$ в W. Теорема полностью доказана.

3. Рассмотрим общий выпуклый случай — без предположения строго равномерной выпуклости J. Тогда задача (1),(2) может иметь не единственное решение u^0, а множество U^0 (выпуклое, замкнутое, ограниченное [1]) решений, и не быть корректной по Тихонову. При этом даже в конечномерном случае алгоритм (4),(5) может не давать сходящейся последовательности (гарантируется лишь бесконечномерных задач), сходимость по функционалу к J^0 не влечет за собой сильной сходимости к множеству решений U^0. Так что результаты о слабой сходимости метода представляют не более, чем теоретический, интерес. Отметим, что без дополнительных предположений о задаче совмещение итеративной аппроксимации с прокс-регуляризацией (см. например, в [22]) гарантирует лишь слабую сходимость соответствующей последовательности. Поэтому далее будет предложена схема итеративной регуляризации по Тихонову алгоритма 1.

Нам потребуются дополнительные предположения регулярности ограничений задачи (1)–(2), а именно, предположение регулярности ограничений задач (3) при всех l. С учетом того, что исходные ограничения — операторные, и в задаче (3) бесконечное число ограничений, нет оснований рассчитывать на обычное условие Лагранжа или на ρ-регулярность, введенную в [19], а будет использоваться ослабленный вариант последней, называемый ρ^θ-регулярностью [20].

Итак предположим, что $\forall k = 1, 2, \ldots$

$$\exists \theta \geq 1, \quad \eta(l_k) > 0 : \quad \forall a \in A'_{l_k} \setminus A^0_{l_k} \quad \int\limits_X [g_k^+(a,x)]^2 dx \geq \eta(l_k)\{\rho_k(a, A^0_{l_k})\}^\theta; \tag{15}$$

здесь и далее $\rho_k(\cdot,\cdot)$ служит для обозначения расстояния в пространстве $\mathbb{R}^{l_k}$.

Чтобы обеспечить выполнение условия регулярности (15), для операторных ограничений (2) достаточно предположить следующее:

1) множество X состоит из конечного объединения выпуклых тел в $\mathbb{R}^n$,

2) функции $g_k(a,\cdot)$ липшицевы по x на X с константой $L_x(l_k)$ $\forall a \in A'_{l_k}$ (например, ξ — тригонометрическая система),

3) $\exists \hat{a} \in A'_{l_k}$: $g_k(\hat{a},x) < 0$ $\forall x \in X$.

Тогда, согласно Утверждению из [20], условие (15) справедливо с

$$\eta(l_k) = O\Big(\kappa_k / (2L_x^n(l_k)\mathrm{diam} A^0_{l_k})\Big), \quad \theta = n + 2,$$

где $\kappa_k = -\max\limits_{x\in X} g_k(\hat{a},x)$. Более удобным для проверки, чем 3), является условие

$$\exists \hat{u} \in U' : \quad G(\hat{u})[x] < 0 \;\text{ п.н. } \; \forall x \in X,$$

кручения можно выбрать $\hat{u} \equiv 0$.

В сделанных предположениях для решения задачи (1),(2) в [20] предложен комбинированный метод итеративных штрафов и аппроксимации, регуляризованный с помощью функции Тихонова, и доказана его сильная сходимость. Однако, учет ограничений путем "штрафования" (добавления функционала $C_k S(\cdot)$ к целевому при $C_k \uparrow \infty$) создает определенные вычислительные трудности. Поэтому ниже аналогичная схема регуляризации будет использована при учете ограничений по методу агрегирования (4),(5).

Введем для рассматриваемой задачи функционал Тихонова

$$J_k(u) \stackrel{\text{def}}{=} J(u) + \alpha_k \Omega(u),$$

где $\alpha_k \to 0$ — параметр регуляризации, $\Omega(u)$ — сильно выпуклый, непрерывный функционал (*стабилизатор* [1, с.166]), например $\Omega(u) = ||u||^2$. И будем применять алгоритм 1 при замене J на J_k. При этом нам потребуются дополнительные условия согласования параметров алгоритма с параметром α_k регуляризации.

Прежде всего с $\alpha_k \downarrow 0$ надо согласовать рост размерности аппроксимирующего пространства l_k так, чтобы гарантировать

$$J(v^k) - J^0 = o(\alpha_k). \tag{16}$$

Такие α_k очевидно найдутся при любых l_k. Для конструктивного их задания укажем, что в случае липшицевости J на U (с константой L_J), если $[u^0]_{l_k} \in U$ $\forall k > k^0$, $u^0 \in U^0$, то при условии $U^0 \subset \mathbb{W}^{r+q}(X)$ можно так выбрать базис в $\mathbb{W}^r(X)$, чтобы для $(r+q)$-й производной по x выполнялось $||\xi_i^{(r+q)}|| = O(i^q)$ (например, тригонометрическая система ξ при $r = 0$ или $\xi_i = \sin(ix)/i$ для $X = (0, \pi)$ и нулевых граничных условий при $r = 1$ и т.п.). Тогда для справедливости (16) достаточно обеспечить $\alpha_k \geq O(1/l_{k-1}^q)$.

Действительно, в сделанных предположениях, рассматривая в качестве основного пространства $W = \mathbb{W}^r(X)$, получаем оценку

$$\forall k > k^0 + 1 \quad J(v^k) - J^0 \leq L_J ||v^k - u^0|| \leq L_J ||u^0 - [u^0]_{l_{k-1}}|| =$$

$$= L_J\Big[\sum_{i>l_{k-1}} (a_i^2)\Big]^{1/2} < \frac{L_J}{l_{k-1}^q}\Big[\sum_{i>l_{k-1}} (i^q a_i)^2\Big]^{1/2} = o(L_J/l_{k-1}^q),$$

поскольку последний ряд сходится в силу $u^0 \in \mathbb{W}^{r+q}(X)$ и условия на базис.

Упомянутый выбор базиса позволяет дать явную оценку для α_k, гарантирующую (16) и в случае $[u^0]_l \notin U \ \ \forall l$, если взамен предположить, как в замечании 2, липшицевость G^+ с показателем L_G и ρ^θ-регулярность (15) с $\eta(l) \equiv \eta$. Тогда имеем $l^q \varepsilon_l \to 0$ для $\varepsilon_l \stackrel{\text{def}}{=} \|u^0 - [u^0]_l\|$, $\|G^+([u^0]_l)\| \le L_G\|u^0 - [u^0]_l\| = L_G\varepsilon_l$ и $\|v^k - [u^0]_{l_k}\|^\theta \le O(\|G^+([u^0]_{l_k})\|^2)$, выводя в результате

$$||v^k - u^0|| \le ||v^k - [u^0]_{l_k}|| + \varepsilon_{l_k} \le O(\varepsilon_{l_k}^{2/\theta} + \varepsilon_{l_k}) = o(l_k^{-2q/\theta} + l_k^{-q}),$$

что дает для выполнения (16) неравенство $\alpha_k \ge O(l_k^{-2q/\theta} + l_k^{-q})$. При этом остается произвол в выборе l_k.

Отметим, что при учете ограничений по методу штрафов [20] достаточно либо потребовать, чтобы $\alpha_k \ge O(1/l_{k-1}^q)$ при соответствующих условиях на выбор базиса, либо дополнительно предположить $[u^0]_l \in U$. И в этом смысле метод из [20], как менее чувствительный, может оказаться более удобным. (В обоих случаях сходимость доказывается в метрике пространства W.)

Для того чтобы удовлетворить (16), последовательность α_k должна быть достаточно медленно убывающей. Это не противоречит и следующей модификации условий (6),(7):

$$\alpha_k \downarrow 0,\ \tau_k \downarrow 0,\ \sum \alpha_k\tau_k = \infty,\ l_k - \text{натуральное},\ l_k \uparrow \infty,\ \ |p^{k+1} - \bar{p}^k| = o(\alpha_k\tau_k). \quad (17)$$

Условия (8), обеспечивающие оптимальность слабых предельных точек, сохраняются и для рассматриваемого алгоритма, но к ним еще добавятся условия на α_k, зависящие от свойств регулярности ограничений. Запишем их вместе:

$$\lim_{k\to\infty} \tau_k L_G^2(l_k) L_J(l_k) = 0, \quad \tau_k L_G^2(l_k)[L_J(l_k)]^\theta/\eta(l_k) = o(\alpha_k^{\theta-1}). \qquad (18)$$

Условия (18) не противоречат (16), но могут противоречить последней части (17) в случае $L_G^2(l)$, $L_J(l)$ порядка l^s с большим показателем $s \ge 1/\theta$ (см. замечание 2). Поэтому далее предположим, что L_G^2, L_J и η не зависят от l, комендуется учет ограничений по методу штрафов [20]). Тогда для обеспечения (17),(18) можно взять, аналогично (9), $l_k = \lfloor \log_2(k + k_0) \rfloor$ и положить $\alpha_k = \Big[\log_2 \log_2(k + k_1)\Big]^{-1}$, $\tau_k = \Big[\log_2 \log_2(k + k_2)\Big]^{-2n}$.

Теперь формально опишем алгоритм с итеративной регуляризацией.

Алгоритм 2

Для любого начального приближения $a^1 \in \{a \in \mathbb{R}^{l_0} |\ a\xi \in U'\}$ выберем управляющие последовательности $\{\tau_k\}$, $\{\alpha_k\}$ и $\{l_k\}$, удовлетворяющие условиям (16)–(18), и положим $\forall k = 1, 2, \ldots$

$$\bar{a}_i^k = a_i^k \ \ \forall i \le l_{k-1}, \ \ \bar{a}_i^k = 0, \ l_{k-1} < i \le l_k,$$

$$a^{k+1} = \arg\min_{a \in A_k} \Big[\tau_k J_k(a\xi) + |a - \bar{a}^k|^2/2\Big],$$

где A_k определяется (11), как в алгоритме 1.

Сходимость алгоритма 2 обосновывает следующая

Теорема 2. *В сделанных предположениях последовательность* $\{u^k = a^k\xi\}$, *построенная по алгоритму* 2, *сильно сходится к* Ω*-нормальному решению* u^* *задачи* (1),(2), *т.е. к такому* $u^0 \in U^0$, *на котором* (*см.* [1, с.174]) *достигается* $\min\limits_{u^0 \in U^0} \Omega(u^0)$.

Доказательство. Выберем p^k так, чтобы $v^k = p^k\xi$ для $v^k = \mathrm{pr}_{U \bigcap U_{l_{k-1}}} u^*$. Аналогично доказательству теоремы 1 получим, что выполнено (13) при замене J на J_k. Таким образом,

$$|a^{k+1} - p^{k+1}|^2 \le |a^k - p^k|^2 + 2\tau_k\Big(J_k(\bar{p}^k\xi) - J_k(a^{k+1}\xi)\Big) - |a^{k+1} - \bar{a}^k|^2 + \gamma_k, \quad (19)$$

где $\gamma_k = o(\tau_k\alpha_k)$ согласно (17). Рассмотрим бесконечно малую $\delta_k > 0$:

$$\delta_k \downarrow 0, \ \ \sum \delta_k = +\infty, \ \ \delta_k = o(\alpha_k\tau_k), \ \ \gamma_k = o(\delta_k),$$

которая существует с учетом $\sum \tau_k\alpha_k = +\infty$, и введем два индексных множества

$$K_- \stackrel{\mathrm{def}}{=} \{k = 1, 2, \ldots |\ 2\tau_k\Big(J_k(a^{k+1}\xi) - J_k(\bar{p}^k\xi)\Big) + |a^{k+1} - \bar{a}^k|^2 - \gamma_k \ge \delta_k\}$$

$$\text{и} \ \ K_+ \stackrel{\mathrm{def}}{=} \{k = 1, 2, \ldots\} \setminus K_-,$$

имеем

$$J_k(a^{k+1}\xi)-J_k(\bar p^k\xi)+\Big(2L_G^2(l_k)\tau_k\Big)^{-1}\int\limits_X[g_k^+(a^{k+1},x)]^2dx\le(\gamma_k+\delta_k)/\tau_k,\ \ k\in K_+. \quad (20)$$

Отсюда точно так же, поскольку $o(\tau_k\alpha_k)=o(\tau_k)$, заключаем, что

$$\int\limits_X[g_k^+(a^{k+1},x)]^2dx\le O\Big(L_J(l_k)L_G^2(l_k)\tau_k\Big)+o(L_G^2(l_k)\tau_k)\to 0,\quad k\to\infty,\quad k\in K_+,$$

и доказываем $\hat u\in U$ и $J(\hat u)=J^0$, т.е. $\hat u\in U^0$, для любой слабой предельной точки $\hat u$ последовательности $\{u^{k+1}\}_{k\in K_+}$. К сожалению, так как задача (1),(2) не корректна, нельзя сделать вывод о сходимости этой последовательности на основании того, что она является минимизирующей для рассматриваемой задачи.

Продолжим доказательство. Раскроем J_k в (20):

$$\begin{aligned}&\alpha_k[\Omega(a^{k+1}\xi)-\Omega(\bar p^k\xi)]\le\\ &\le J(\bar p^k\xi)-J(a^{k+1}\xi)-\Big(2L_G^2(l_k)\tau_k\Big)^{-1}\int\limits_X[g_k^+(a^{k+1},x)]^2dx+(\gamma_k+\delta_k)/\tau_k= \qquad (21)\\ &=(\gamma_k+\delta_k)/\tau_k+J(\bar p^k\xi)-J^0+\Big\{J^0-J(a^{k+1}\xi)-\Big(2L_G^2(l_k)\tau_k\Big)^{-1}\int\limits_X[g_k^+(a^{k+1},x)]^2dx\Big\}.\end{aligned}$$

Оценим выражение в фигурных скобках в (21). Для случая $J^0\le J(a^{k+1}\xi)$ все выражение не положительно. В противном случае $a^{k+1}\in A'_{l_k}\setminus A^0_{l_k}$ и для $a=a^{k+1}$ выполнено неравенство в (15). Обозначим через b^k проекцию a^{k+1} на $A^0_{l_k}$ и будем использовать краткую запись ρ_k для расстояния между a^{k+1} и $A^0_{l_k}$. Тогда

$$J^0-J(a^{k+1}\xi)\le J(b^k\xi)-J(a^{k+1}\xi)\le L_J(l_k)\rho_k(a^{k+1},A^0_{l_k})$$

и для исследуемого выражения в фигурных скобках получим оценку сверху

$$L_J(l_k)\rho_k-\frac{\eta(l_k)\rho_k^\theta}{2L_G^2(l_k)\tau_k}.$$

Полученная разность либо отрицательна, либо не превосходит $L_J(l_k)\rho_k$ при

$$2L_J(l_k)L_G^2(l_k)\tau_k/\eta(l_k)\ge\rho_k^{\theta-1},$$

т.е. не превосходит $O\big([E_f(t_k)] \quad [E_G(t_k)] \quad \big)$ и в силу (18) равна $o(\alpha_k)$.

В результате правая часть в (21) не превосходит $o(\alpha_k)$ из условий (16), (17) и выбора δ_k. Разделив обе части на α_k и переходя к пределу по $k \in K_+$ при $k \to \infty$, установим для любой слабой предельной точки $\hat{u}$, что

$$\Omega(\hat{u}) \le \liminf_{k \in K_+} \Omega(u^{k+1}) \le \lim_{k \in K_+} \Omega(v^k) = \Omega(u^*)$$

за счет изначального выбора p^k. Но $\Omega(u^*) \le \Omega(\hat{u})$ по определению u^*, поскольку $\hat{u} \in U^0$. Таким образом, последовательность $\{u^{k+1}\}_{k \in K_+}$ является минимизирующей в задаче на

$$\min_{u^0 \in U^0} \Omega(u^0),$$

которая уже (в отличие от исходной) будет корректной по Тихонову. Отсюда выводим сильную сходимость $\{u^{k+1}\}_{k \in K_+}$ к u^*.

Теперь, повторяя последний абзац доказательства теоремы 1, можем записать (19) как $|a^{k+1} - p^{k+1}|^2 \le \max[\kappa(\delta_k),\ |a^k - p^k|^2 - \delta_k]$ и придем к утверждению теоремы 2.

1. Васильев Ф.П. Численные методы решения экстремальных задач. М.: Наука, 1981.

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

3. Дюво Г., Лионс Ж.Л. Неравенства в механике и физике. М.: Наука, 1980.

4. Никольский С.М. Приближение функций многих переменных и теоремы вложения. М.: Наука, 1977.

5. Semi-infinite programming / Ed. by R. Hettich. New York: Springer, 1979. Lect. Notes in Control and Inform. Sci. V.15.

6. Semi-infinite programming and applications / Eds. by A. Fiacco, K. Kortanek. New York: Springer, 1983. Lect. Notes in Econom. and Math. Systems. V. 215.

7. Hettich R. A review of numerical methods for semi-infinite optimization // Semi-Infinite Program. and Applic. New York: Springer, 1983. Lect. Notes in Econom. and Math. Systems. V. 215.

8. Hettich R. An implementation of a discretization method for semi-infinite programming // Math. Program. 1986. V. 34. P. 354-361.

9. Hettich R., Kortanek K. Semi-infinite programming: theory, methods, and applications // SIAM Rev. 1993. V.35. P. 380-429.

10. Wardi Y. A stochastic algorithm for optimization problems with continua of inequalities // J. Optimisat. Theory and Appl. 1988. V. 56. P. 285-311.

11. Wardi Y. Stochastic approximation algorithm for minimax problems // J. Optimizat. Theory and Appl. 1990. V. 64. P. 615-640.

оптимизации для задачи с бесконечным числом ограничений // Вычисл. матем. и моделирование. М.: Изд-во МГУ, 1990. С. 383-394.

13. Novikova N. Iterative stochastic methods for solving variational problems of mathematical physics and operations research // J. Math. Sci. (Contemprorary Math. and Its Appl. V. 3). New-York — London: Plenum Publ. Corp. 1994. N1. P.1-125.

14. Volkov Yu.V., Zavriev S.K. A general stochastic outer approximations method // SIAM J. Control Optimizat. 1997. 35. P.1387-1421.

15. Reemtsen R., Görner S. Numerical methods for semi-infinite programming: a survey // Semi-Infinite Programming. Eds. by R. Reemtsen, J.-J. Rückmann. Boston — London — Dordrecht: Kluwer Acad. Publs. 1998. P.195-275.

16. Давидсон М.Р. Регуляризованный метод агрегирования ограничений для задачи полубесконечной оптимизации // Ж. вычисл. матем. и матем. физ. 1998. Т.38. N.5. С.770-777.

17. Ermoliev Y., Kryazhimskii A., Ruszczynski A. Constraint aggregation principle in convex optimization: Working paper WP-95-015. Laxenburg: IIASA, 1995.

18. Davidson M. Stochastic constraint aggregation method for convex semi-infinite problems: Working paper 97-17. Trier: Univ. Trier, 1997.

19. Федоров В.В. Численные методы максимина. М.: Наука, 1979.

20. Новикова Н.М. Об итеративной регуляризации аппроксимации метода штрафов в гильбертовом пространстве // Ж. вычисл. матем. и матем. физ. 1989. Т.29. N.6. С. 949-954.

21. Вазан М. Стохастическая аппроксимация. М.: Мир, 1972.

22. Поляк Б.Т. Введение в оптимизацию. М.: Наука, 1983.