

ОБ ОДНОМ ПОДХОДЕ К УСЛОВНОЙ МИНИМИЗАЦИИ D.C. ФУНКЦИЙ

Е. А. Котельников

Институт вычислительной математики и математической геофизики СО РАН,
630090, Новосибирск, Россия

УДК 519.853.32

В работе предлагается алгоритм поиска глобального минимума функции, представимой в виде разности двух выпуклых функций, на замкнутом ограниченном выпуклом множестве. При построении алгоритма используются элементы многокритериальной оптимизации.

Ключевые слова: *d.c.* функция, многокритериальная оптимизация, Парето-оптимальные решения.

In this paper the algorithm of search of a global minimum of function, representable in the form of a difference of two convex functions, on the closed limited convex set is considered. At algorithm construction elements multi-criteria optimization are used.

Key words: *d.c.* function, multi-criteria optimization, Pareto-optimal solution.

Введение. Рассматривается задача поиска глобального минимума числовой функции f , представимой в виде разности двух выпуклых функций f_1 и f_2 , на ограниченном замкнутом выпуклом множестве X из R^n :

$$\text{найти } \min_{x \in X} f(x). \quad (1)$$

Будем считать, что функции f_1 и f_2 дифференцируемы (f_1 – дважды); функция f_1 строго выпукла, а функция f_2 выпукла, но не постоянна. Функции, которые можно представить в виде разности двух выпуклых функций, называют в литературе *d.c.* функциями (от английского „difference of two convex functions“). Обширный обзор литературы по данной тематике и описание применения задачи (1) можно найти в [1].

В общем случае *d.c.* функция f не является выпуклой функцией, поэтому нам предстоит найти среди точек локального минимума функции f на множестве X точку, в которой значение функции f окажется наименьшим. В помощь решению этой проблемы привлечем некоторые результаты исследования из области многокритериальной (векторной) минимизации [2].

Рассмотрим функции f_1 и f_2 как частные критерии, образующие векторный критерий $F = (f_1, f_2)$ минимизации на множестве X . Каждое решение $x \in X$ характеризуется векторной оценкой $F(x)$, и выбор какого-либо решение из множества X сводится к выбору соответствующей оценки из множества достижимых оценок

$$Y = F(X) = \left\{ y \in R^2 : y_1 = f_1(x), y_2 = f_2(x), x \in X \right\}. \quad (2)$$

В этом случае пространство R^2 называют критериальным. Если оценка y^0 достижима ($y^0 \in Y$), то существует точка $x \in X$, которая является решением системы уравнений $F(x) = y^0$.

В реальных задачах описать множество Y очень сложно, а то и невозможно [2]. В статье предлагается использовать множество Y следующим образом. В критериальном пространстве рассматривается задача минимизации линейной функции $y_1 - y_2$ (y_1 и y_2 — координаты точки $y \in R^2$) на множестве Y , связь которой с исходной состоит в том, что по глобальному решению этой задачи можно восстановить глобальное решение исходной задачи (1). Хотя множество Y задано в неявном виде, используя его некоторые свойства и наглядность изображения в двухмерном пространстве, можно найти способ решения построенной задачи.

В первом разделе статьи приводятся некоторые свойства множества Y , строится задача в критериальном пространстве и описывается стратегия поиска ее решения.

Во втором разделе рассматривается процедура решения задачи (1), в основу которой положена стратегия поиска глобального решения задачи в критериальном пространстве.

1. Решение задачи в критериальном пространстве. Перепишем задачу (1) в следующем виде: минимизировать линейную функцию

$$\varphi(y) = y_1 - y_2 \quad (3)$$

при ограничениях $f_1(x) - y_1 = 0, f_2(x) - y_2 = 0, x \in X$. Очевидно, что ограничения последней задачи описывают множество Y . Тогда совместно с задачей (1) будем рассматривать задачу:

$$\text{найти } \min_{y \in Y} \varphi(y). \quad (4)$$

Лемма. Если y^* — глобальное решение задачи (4), то решение системы уравнений: $f_1(x) = y_1^*, f_2(x) = y_2^*, x \in X$ является глобальным решением задачи (1).

Доказательство. Пусть x^* — решение системы уравнений $F(x) = y^*, x \in X$ и x^* не является точкой глобального минимума функции f на множестве X , тогда существует точка $\bar{x} \in X$, такая, что $f(\bar{x}) < f(x^*)$. Обозначим через \bar{y} точку $F(\bar{x})$, тогда $\bar{y}_1 - \bar{y}_2 = f_1(\bar{x}) - f_2(\bar{x}) < f_1(x^*) - f_2(x^*) = y_1^* - y_2^*$, или $\varphi(\bar{y}) < \varphi(y^*)$, что противоречит нашему допущению. Лемма доказана.

Утверждение данной леммы дает возможность смоделировать в критериальном пространстве процесс решения задачи (1).

Предположим, что мы имеем точки

$$a = \arg \min_{x \in X} f_1(x), b = \arg \max_{x \in X} f_2(x), c = \arg \max_{x \in X} f_1(x), d = \arg \min_{x \in X} f_2(x). \quad (5)$$

Определим точки $A, B, C, D \in R^2$ следующим образом:

$$A = F(a), B = F(b), C = F(c), D = F(d). \quad (6)$$

Из всех перечисленных точек необходимо найти точку a , а остальные используются только для иллюстрации описания алгоритма.

В векторной минимизации интерес представляет юго-западная часть границы множества Y [2], то есть кривая между точками A и D . Для выпуклых задач векторной минимизации, то есть когда множество X выпукло и функции f_1, f_2 выпуклы, эту часть границы называют множеством эффективных оценок относительно множества Y или множеством Парето и обозначают $P(Y)$. Для невыпуклых задач не все точки юго-западной части границы множества Y являются эффективными [2].

В данной статье нас будет интересовать северо-западный участок границы множества Y , то есть кривая, расположенная между точками A и B . Обозначим этот участок границы через U .

Определение. Точка $y^* \in U$ является точкой касания кривой U и прямой $L = \{y \in R^2 : \varphi(y) = \varphi(y^*)\}$, если существует малая δ -окрестность $\{y \in U : \|y - y^*\|_2 < \delta\}$ точки y^* , в которой или $\varphi(y) \geq \varphi(y^*)$, или $\varphi(y) \leq \varphi(y^*)$.

В первом случае y^* — точка локального минимума функции φ на множестве Y , во втором случае y^* — точка локального максимума функции φ на кривой U .

Пусть y^* — произвольная точка локального минимума функции φ на множестве Y и $M(y^*)$ — множество точек v локального минимума φ и Y , таких, что $\varphi(v) = \varphi(y^*)$. Множество $M(y^*)$ может состоять из точки, отрезка прямой или, возможно, из точек и отрезков. Очевидно, что число таких множеств конечно. Это значит, что любая последовательность $\{y^{(k)}\}, k = 1, 2, \dots$, точек локального минимума φ на Y , таких, что $\varphi(y^{(k+1)}) < \varphi(y^{(k)})$, конечна.

1.1. *Описание множества Y и анализ кривой U .* Не умаляя общности, будем считать, что $f_1(x) \geq 0$ и $f_2(x) \geq 0 \forall x \in X$.

Отвлечемся на время от множества Y и рассмотрим его зеркальное отражение относительно оси абсцисс. Рассмотрим невыпуклую двухкритериальную задачу минимизации с частными критериями $g_1(x) = f_1(x)$, $g_2(x) = -f_2(x)$, $x \in X$, у которой множество достижимых оценок $Z = \left\{z \in R^2 : z_1 = g_1(x), z_2 = g_2(x), x \in X\right\}$.

Тогда юго-западная граница множества Z , которую обозначим через V , является зеркальным отражением кривой U относительно оси абсцисс.

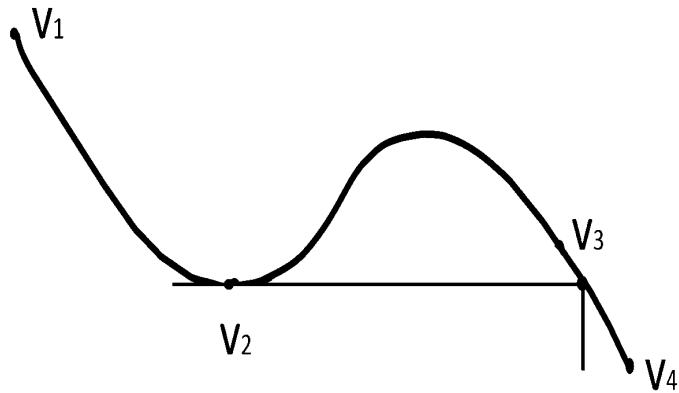
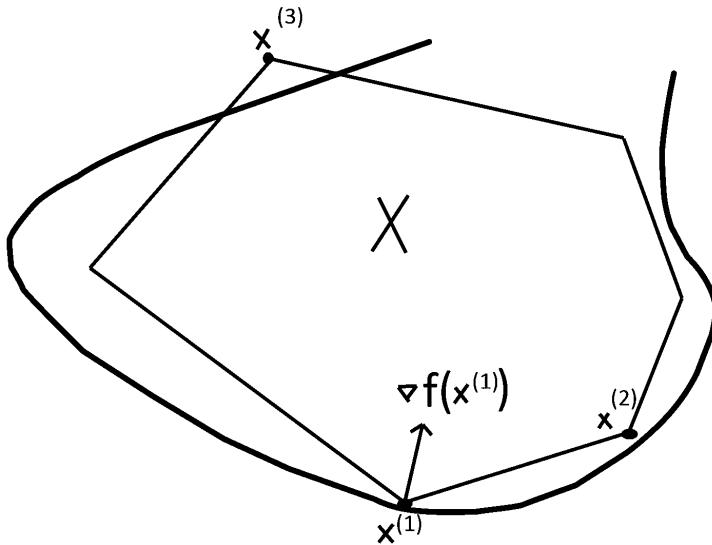
В связи с тем, что последняя задача не является выпуклой, не все точки множества V эффективны. Точка $z^0 \in V$ эффективна, если не существует $z \in Z$, для которой выполнились бы неравенства $z_1 \leq z_1^0, z_2 \leq z_2^0$, причем хотя бы для одного — строго [2]. Геометрически это выглядит следующим образом. Точка z^0 не является эффективной, если прямой угол с вершиной $z^0 \{z \in R^2 : z_1 \leq z_1^0, z_2 \leq z_2^0\}$ содержит хотя бы одну точку из множества Z , отличную от z^0 .

Пример участка кривой V , расположенного между точками v_1 и v_4 , изображен на рис. 1.

Здесь эффективными точками являются точки кривой V между v_1 и v_2 (включая их) и между v_3 и v_4 (включая только v_4). Судя по рис. 1, точка v_2 имеет координаты $v_{21} = g_1(x^*), v_{22} = g_2(x^*)$, где x^* — точка локального минимума функции g_2 на множестве X .

В эффективной точке любое уменьшение значения одного критерия невозможно без увеличения другого, то есть эффективную оценку нельзя улучшить по векторному критерию [2].

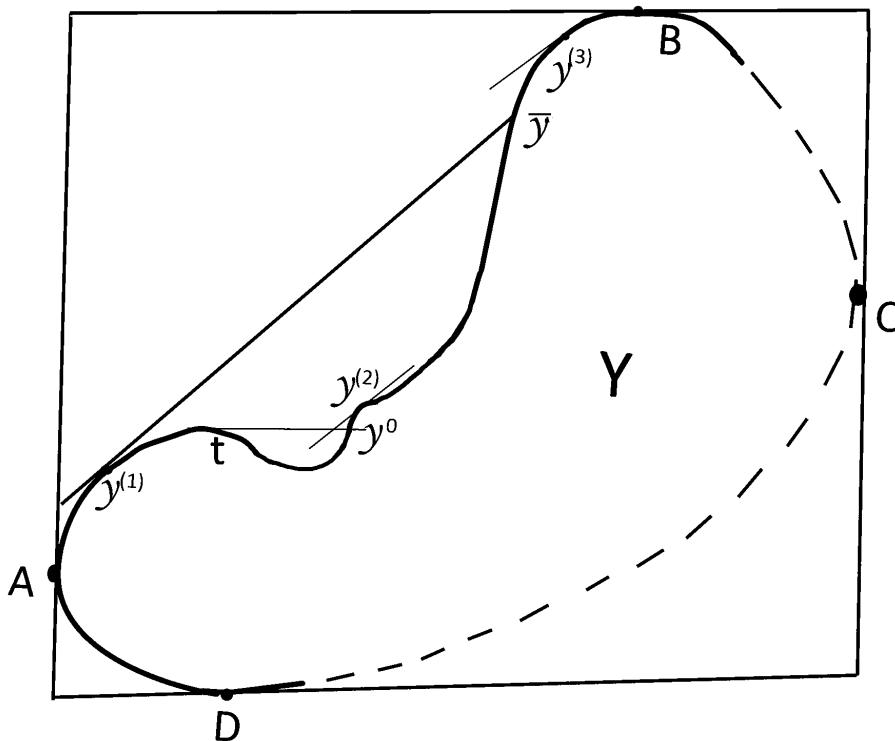
Вернемся теперь к описанию множества Y .

Рис. 1. Пример участка юго-западной границы Z Рис. 2. Пример множества допустимых решений X

Множеству $P(Z)$ зеркально относительно оси абсцисс расположено множество точек из множества U , которое назовем множеством псевдоэффективных оценок относительно множества Y и обозначим через $PP(Y)$. Тогда, если прямой угол $G = \{y \in R^2 : y_1 \leqslant y_1^0, y_2 \geqslant y_2^0\}$ с вершиной в точке $y^0 \in U$ не содержит ни одной точки из множества Y , отличной от y^0 , то $y^0 \in PP(Y)$.

Пример множества Y . Пусть множество X задано в виде шестиугольника, на котором функция f имеет три точки локального минимума $x^{(1)}, x^{(2)}, x^{(3)}$, причем $f(x^{(3)}) < f(x^{(1)}) < f(x^{(2)})$. Точка $x^{(1)}$ найдена одним из подходящих методов минимизации с начальным приближением $x_0 = a$, где точка a определена в (5). Множество X и линия уровня $\{x : f(x) = f(x^{(1)})\}$ представлены на рис. 2.

Множество Y с находящимися на его границе точками $y^{(k)} = F(x^{(k)})$, $k = 1, 2, 3$ изображено на рис. 3. Точки $y^{(k)}$ являются точками локального минимума функции φ , определенной в (3), что соответствует тому, что $y^{(k)}$ — точки касания кривой U (соответственно, множества Y) и прямых

Рис. 3. Пример множества Y

$$L_k = \left\{ y \in R^2 : \varphi(y) = \varphi(y^{(k)}) \right\}. \quad (7)$$

Точки A, B, C, D определены в (6). Фрагмент границы множества Y между точками A и D есть множество эффективных точек $P(Y)$, а между точками A и B — кривая U . В данном примере множества Y все точки кривой U , за исключением участка между точками t и y^0 без точки t , являются псевдоэффективными.

Пройдем по кривой U , отмечая участки псевдоэффективных точек и точек, которые не являются псевдоэффективными.

Для диагностики псевдоэффективности точки y^0 воспользуемся прямым углом G . Если в точке $y^0 \in U$ множество $G \cap Y \setminus \{y^0\}$ пусто, то $y^0 \in PP(Y)$. Поскольку двухкритериальная задача минимизации с частными критериями g_1 и g_2 не выпукла, то множество $P(Z)$ может не быть дугообразно связным [2]. Множество называют дугообразно связным, если любые две его точки можно соединить дугой, целиком лежащей в этом множестве. Поэтому множество $PP(Y)$ может состоять из нескольких не связанных друг с другом фрагментов кривой U . Для любых точек $y', y'' \in PP(Y)$, принадлежащих одному или разным фрагментам, $y'_1 < y''_1$ тогда и только тогда, когда $y'_2 < y''_2$.

Введем обозначения.

1. Для любых точек $y', y'' \in U$ $y' \Rightarrow y''$, если при движении точки y по кривой U точка y' была пройдена раньше, чем пройдена точка y'' .

2. Для любых точек $y', y'' \in R^2$ $y' < y''$ тогда и только тогда, когда $y'_1 < y''_1$ и $y'_2 < y''_2$.

Определения.

1. При движении точки $y = (y_1, y_2)^T$ по кривой U первая компонента y_1 растет (убывает) на каком-либо участке U' этой кривой, если для любых точек $y', y'' \in U'$, таких, что $y' \Rightarrow y'', y'_1 < y''_1 (y'_1 > y''_1)$. Такое же определение касается и второй компоненты.

2. При движении точки y по кривой U первая компонента y_1 растет на каком-либо участке U' этой кривой быстрее (медленнее), чем растет вторая компонента y_2 , если для любых точек $y', y'' \in U_1$, таких, что $y' \Rightarrow y''$ и $y' < y''$, верно неравенство $y''_1 - y'_1 > y''_2 - y'_2 (y''_1 - y'_1 < y''_2 - y'_2)$.

3. При движении точки y по кривой U на некотором ее участке U' первая компонента y_1 растет быстрее (медленнее), чем убывает вторая компонента y_2 , если для любых точек $y', y'' \in U'$, таких, что $y' \Rightarrow y'', y'_1 < y''_1$ и $y'_2 > y''_2$, верно неравенство $y''_1 - y'_1 > y'_2 - y''_2 (y''_1 - y'_1 < y'_2 - y''_2)$. Аналогичное определение верно для случая, когда y_1 убывает, а y_2 растет.

Рассмотрим участок кривой U , лежащий между двумя соседними фрагментами множества $PP(Y)$.

Пусть на кривой U существует точка t , которая является образом при отображении F одной из точек локального максимума x^* функции f_2 на множестве X ($t = F(x^*)$). Если известна точка B и $t = B$, то есть x^* — точка глобального максимума f_2 на X , то дальнейшее движение по кривой U прекращаем. Будем пока считать, что точка B нам известна, хотя на практике найти точку B не легче, чем решить исходную задачу (1).

Если $t \neq B$, то справа от точки t находится овраг, так как у текущей точки y кривой U вторая компонента y_2 начинает убывать. Обозначим поверхность оврага через Q . Поставим угол G таким образом, чтобы его вершина y^0 касалась кривой U , а сторона $\{y : y_2 = y^0_2, y_1 < y^0_1\}$ проходила через точку t . Поскольку $t \neq B$, то такая точка y^0 существует. Границами оврага Q будем считать точки t и y^0 . Из того, что $t_1 < y^0_1$ и $t_2 = y^0_2$, следует, что $\varphi(t) < \varphi(y^0)$. Очевидно, что для любой точки $u \in Q$ $\varphi(t) < \varphi(u)$. Таким образом, локальный минимум функции φ на множестве Y нужно искать вне оврагов, а сами овраги, если встречаются на пути движения, необходимо уметь преодолевать.

Предлагается следующий способ преодоления оврага. Заменим временно функцию f_2 на функцию \bar{f}_2 , определенную следующим образом: $\bar{f}_2(x) = \max\{f_2(x), t_2\}$. Тогда множество $\bar{Y} = \{y \in R^2 : y_1 = f_1(x), y_2 = \bar{f}_2(x), x \in X\}$ равно множеству $\{y \in Y : y_2 \geq t_2\} \cup \{y \in R^2 : t_1 \leq y_1 \leq y^0_1, y_2 = t_2\}$. Очевидно, что все точки $y^{(k)}$ ($k > 1$) локального минимума функции φ на множестве Y , в которых значение функции φ не превосходит $\varphi(y^{(1)})$, будут содержаться в множестве \bar{Y} . Это значит, что при замене множества Y на множество \bar{Y} мы не выпустим из вида ни одну точку локального минимума функции φ с меньшими ее значениями величины $\varphi(y^{(1)})$.

1.2. *Поиск глобального минимума функции φ .* Процедуру поиска начинаем с точки A из (6). Используя ее в качестве начального приближения, находим точку $y^{(1)}$ локального минимума функции φ на множестве Y . Для того, чтобы в дальнейшем иметь возможность ссылаться на данную процедуру, назовем ее процедурой $A1(y_0)$ с обязательным указанием начального приближения y_0 .

Далее будем искать очередную точку локального минимума $y^{(2)}$ функции φ на множестве Y , которая удовлетворяет условию $\varphi(y^{(2)}) < \varphi(y^{(1)})$. Для поиска очередной точки локального минимума предлагается сначала найти точку \bar{y} (если такая точка существует) пересечения прямой L_1 из (7) и кривой U , а затем с помощью процедуры $A1(\bar{y})$ — точку $y^{(2)}$. Если такой точки \bar{y} не существует, то необходимо это доказать и прекратить дальнейшие вычисления. В дальнейшем очередную искомую точку локального миниму-

ма функции φ на множестве Y будем обозначать через $y^{(k+1)}$, если последней найденной точкой является $y^{(k)}$.

Таким образом, поиск глобального минимума функции φ на множестве Y сводится к последовательности процедур поиска локальных минимумов $y^{(k)}, k = 1, 2, \dots$, в которых выполнены условия $\varphi(y^{(k+1)}) < \varphi(y^{(k)})$. Как было сказано выше, таких точек $y^{(k)}$ конечное число, то есть процесс не бесконечен.

Зададим малую величину $\varepsilon > 0$, определяющую допуск на равенство значений функции f в разных точках: если $|f(x_1) - f(x_2)| \leq \varepsilon$, то считаем $f(x_1) = f(x_2)$ при $x \neq x_2$.

1.2.1. Алгоритм A1. Алгоритм используется для поиска точек $y^{(k)}, k \geq 1$, локального минимума функции φ на множестве Y и имеет два варианта применения в зависимости от начального приближения y_0 .

В первом варианте находится точка $y^{(1)}$ при $y_0 = A$. При движении по кривой U от точки A до точки $y^{(1)}$ первая компонента y_1 текущей точки $y \in U$ растет, так как $A_1 = \min_{y \in Y} y_1$.

Во втором варианте применения алгоритма находится точка $y^{(k)}, k > 1$, если она существует, при $y_0 = \bar{y}$, где \bar{y} — точка пересечения прямой L_{k-1} , определенной в (7), и кривой U (рис. 3). При движении по кривой U между точками \bar{y} и $y^{(k)}$ для текущей точки выполняются те же требования, что для точки кривой U между A и $y^{(1)}$. Поэтому для простоты изложения опишем движение точки по кривой U только между точками A и $y^{(1)}$.

Рассмотрим все ситуации, которые могут возникнуть при движении текущей точки $y = (y_1, y_2)^T$ по кривой U из точки A .

1. Точка a , определенная в (5), является точкой локального максимума функции f_2 на множестве X . Это значит, что компонента y_1 растет, а компонента y_2 убывает и, следовательно, функция φ возрастает и $y^{(1)} = A, t = A$. Для продолжения поиска очередной точки $y^{(2)}$ будем использовать алгоритм A3 (раздел 1.2.3).

2. Компонента y_2 растет медленнее, чем растет компонента y_1 . Пока это будет продолжаться, $y \in PP(Y)$. В этом случае функция φ растет, а $y^{(1)} = A$, и на очередном этапе поиска $y^{(2)}$ используем алгоритм A2 (раздел 1.2.2).

Далее рассмотрим случаи, когда при движении из точки A компонента y_2 растет быстрее, чем растет компонента y_1 . Такое движение возможно до некоторой точки $s \in U$, при движении после которой скорость компонент относительно друг друга может измениться. Все точки кривой U между A и s являются псевдоэффективными. Обозначим этот участок кривой U через U_1 .

3. Точка s , упомянутая выше, является образом при отображении F из (2) точки локального максимума функции f_1 на множестве X . В этом случае движение по кривой U от точки s возможно по маршруту, который соответствует одновременному уменьшению обеих компонент и росту функции φ , то есть $y^{(1)} = s$ и $t = s$. Для поиска точки $y^{(2)}$ подключаем алгоритм A3.

4. После прохождения кривой U_1 компонента y_1 продолжает расти, а y_2 начинает убывать, а это означает, что точка s есть образ при отображении F точки локального или глобального максимума функции f_2 на множестве X . Тогда при движении после точки s функция φ возрастает; $y^{(1)} = s, t = s$, и для поиска точки $y^{(2)}$ подключаем алгоритм A3. Если s — образ глобального максимума, то с помощью алгоритма A3 можно это подтвердить и остановить процесс вычислений.

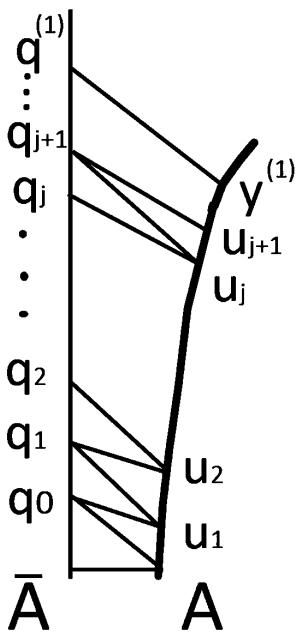


Рис. 4. Геометрическая интерпретация алгоритма A1

5. После прохождения кривой U_1 , компонента y_2 растет медленнее, чем растет компонента y_1 , и, следовательно, функция φ возрастает. Тогда s — точка локального минимума функции φ на Y и $y^{(1)} = s$. Для продолжения поиска точки $y^{(2)}$ используем алгоритм A2.

Таким образом, если двигаться из точки A вдоль кривой U , то есть когда обе компоненты текущей точки y возрастают, а y_2 растет быстрее, чем возрастает y_1 , то обязательно повстречается точка $y^{(1)}$ локального минимума функции φ на множестве Y . Другими словами, если построить минимизирующую последовательность $\{u_j\}$ функции φ на множестве Y , удовлетворяющую условиям: $u_0 = A$, $u_j \in U$, $u_j < u_{j+1}$, $\varphi(u_j) > \varphi(u_{j+1})$, то $u_j \rightarrow y^{(1)}$. В первых четырех случаях точка $y^{(1)}$ достигается за конечное число шагов. В последнем случае последовательность $\{u_j\}$ может быть бесконечна.

Приведем один из способов построения минимизирующей последовательности $\{u_j\}$.

Введем обозначения: $e_1 = (1, 0)^T$, $e_2 = (0, 1)^T$, $e = (1, 1)^T$, $h = (-1, 1)^T$; определим точки $u_0 = A$, $\bar{A} = A - \varepsilon \cdot p \cdot e_1$, где коэффициент p задается пользователем, например, $p = 10$.

В алгоритме, пошаговая схема которого приводится ниже, а его геометрическая интерпретация — на рис. 4, генерируются две последовательности точек $\{q_j\}$ и $\{u_j\}$, $j \geq 0$: если известна точка u_j , $j \geq 0$, то находится q_j — точка пресечения лучей $L'_1 = \{y \in R^2 : y = A' + \gamma e_2, \gamma \geq 0\}$ и $\{y \in R^2 : y = u_j + \gamma \cdot h, \gamma \geq 0\}$ — и вычисляется точка $u_{j+1} \in U$ (см. шаг 1 схемы). Таким образом, все точки q_j расположены на луче L'_1 , а u_j — на кривой U .

Теорема 1. Последовательность $\{u_j\}$, построенная по приведенным выше правилам, является минимизирующей для функции φ на множестве Y , то есть $\varphi(u_j) \rightarrow \varphi(y^{(1)})$ при $j \rightarrow \infty$.

Доказательство. Из того, что u_{j+1} есть решение задачи $\min_{y \in Y} \|q_j - y\|^2$ с начальным приближением $y_n = u_j$, следует $(q_j - u_j)^T(u_{j+1} - u_j) > 0$. Здесь и далее $\|\cdot\|$ — норма евклидова. По построению $q_j = u_j + \gamma \cdot h$ при некотором $\gamma > 0$, тогда $\gamma h^T(u_{j+1} - u_j) > 0$ или $-(u_{j+1,1} - u_{j1}) + (u_{j+1,2} - u_{j2}) > 0$, то есть $\varphi(u_{j+1}) < \varphi(u_j)$. Рассмотрим числовую

последовательность $\{\varphi(u_j)\}$. Она строго убывающая и ограничена снизу числом $\varphi(y^{(1)})$, то есть сходится. Пусть φ^* — ее предел и точка $y^* \in U$, такая, что $\varphi(y^*) = \varphi^*$.

Допустим $\varphi(y^*) > \varphi(y^{(1)})$, то есть $y_1^* - y_2^* - y^{(1)} + y_2^{(1)} > 0$ или $h^T(y^{(1)} - y^*) > 0$. Пусть $q^* = y^* + \gamma h$, где γ выбрано так, чтобы $q^* \in L'_1$, то есть $\gamma = y_1^* - A'_1 > 0$. Тогда $h^T(y^{(1)} - y^*) = \frac{1}{\gamma}(q^* - y^*)^T(y^{(1)} - y^*) > 0$ и, следовательно, задача $\min_{y \in Y} \|q^* - y\|^2$ с начальным приближением $y_n = y^*$ имеет решение $y' \in U$, в котором $\varphi(y') < \varphi^*$. Точка y' является элементом последовательности $\{u_j\}$, так как получена по правилам построения этой последовательности, а это противоречит тому, что φ^* — нижняя граница последовательности $\{\varphi(u_j)\}$. Это противоречие показывает ложность допущения $\varphi(y^*) > \varphi(y^{(1)})$. Теорема доказана.

Теорема 2. Последовательности $\{u_j\}$ и $\{q_j\}$ сходятся, причем, $u_j \rightarrow y^{(1)}$, $q_j \rightarrow q^{(1)}$ при $j \rightarrow \infty$. Здесь $q^{(1)} = y^{(1)} + (y_1^{(1)} - \bar{A}_1)h$.

Доказательство. Для сходимости последовательности точек $\{q_j\}$ достаточно показать сходимость числовой последовательности $\{q_{j2}\}$, так как $q_j = (\bar{A}_1, q_{j2})$ для всех j .

При доказательстве будем исходить из того, что для любого j $u_j < u_{j+1}$. Тогда $q_{j+1,2} - q_{j2} = (u_{j+1,2} - u_{j2}) + (u_{j+1,1} - u_{j1}) > 0$. По правилу построения точек последовательности $\{q_j\}$ имеем $q_{j2} = u_{j2} + u_{j1} - \bar{A}_1 < y_2^{(1)} + y_1^{(1)} - \bar{A}_1 = q_2^{(1)}$, то есть числовая последовательность $\{q_{j2}\}$ монотонно возрастающая и ограничена сверху. Следовательно, она сходится, и пусть q_2^* — ее предел.

Предположим, $q_2^* < q_2^{(1)}$. Обозначим через q^* точку $(\bar{A}_1, q_2^*)^T$ и определим $u^* = q^* - \gamma^* h$, где выберем $\gamma^* > 0$ таким образом, чтобы $u^* \in U$.

Тогда $h^T(y^{(1)} - u^*) = -(y_1^{(1)} - u_1^*) + (y_2^{(1)} - u_2^*) = \varphi(u^*) - \varphi(y^{(1)}) > 0$. Последнее неравенство строгое, так как в противном случае $q^* = q^{(1)}$. Из неравенства $h^T(y^{(1)} - u^*) > 0$ следует, что задача $\min_{y \in Y} \|q^* - y\|^2$ с начальным приближением $y_n = u^*$ имеет решение \bar{u} , отличное от u^* . Определим точку $\bar{q} = (\bar{A}_1, \bar{q}_2)$, где $\bar{q}_2 = \bar{u}_2 + (\bar{u}_1 - \bar{A}_1)$. Тогда $\bar{q}_2 > q_2^*$, так как $u^* < \bar{u}$. Точка \bar{u} получена по правилам построения последовательности $\{q_j\}$, но неравенство $\bar{q}_2 > q_2^*$ противоречит тому, что q_2^* — верхняя граница последовательности $\{q_j\}$. Следовательно, $q_2^* = q_2^{(1)}$ и $q^* = q^{(1)}$.

Покажем теперь сходимость последовательности $\{u_j\}$. По правилам построения последовательности $\{q_j\}$ имеем $q_j = u_j + (u_{j1} - \bar{A}_1)h$. Тогда $h^T q_j = h^T u_j + 2(u_{j1} - \bar{A}_1)$ или $u_{j1} = \frac{1}{2}(h^T q_j - h^T u_j) + \bar{A}_1$. Очевидно, что $h^T u_j = -\varphi(u_j)$. Тогда при $j \rightarrow \infty$ $h^T q_j \rightarrow h^T q^{(1)}$, $h^T u_j \rightarrow -\varphi(y^{(1)})$ и $\lim_{j \rightarrow \infty} u_{j1} = \frac{1}{2}(h^T q^{(1)} + \varphi(y^{(1)}))$. Из определения точки $q^{(1)}$ следует, что $h^T q^{(1)} = h^T y^{(1)} + 2(y_1^{(1)} - \bar{A}_1)$ и, следовательно, $\lim_{j \rightarrow \infty} u_{j1} = y_1^{(1)}$.

Далее имеем $\varphi(u_j) \rightarrow \varphi(y^{(1)})$ при $j \rightarrow \infty$, тогда $\varphi(u_j) - \varphi(y^{(1)}) = (u_{j1} - y_1^{(1)}) + (y_2^{(1)} - u_{j2}) \rightarrow 0$ и $u_{j1} \rightarrow y_1^{(1)}$ при $j \rightarrow \infty$. Следовательно, $u_{j2} \rightarrow y_2^{(1)}$ при $j \rightarrow \infty$. Теорема доказана.

Приведем теперь схему алгоритма.

Шаг 0. Определим точки $u_0 = A$, $\bar{A} = A - \varepsilon \cdot p \cdot e_1$; $q_0 = u_0 + \gamma \cdot h$, где γ находится из условия $q_{01} = \bar{A}_1$, то есть $\gamma = \varepsilon \cdot p$; положим $j = 0$; зададим начальное приближение $y_n = u_0$.

Шаг 1. Найдем локальное решение u_{j+1} задачи $\min_{y \in Y} \|q_j - y\|^2$ с начальным приближением y_n . Если $u_{j+1,2} = u_{1,2}$, то перейдем на шаг 3. Если $\|u_{j+1} - u_j\| \leq \varepsilon$, то перейдем на шаг 4.

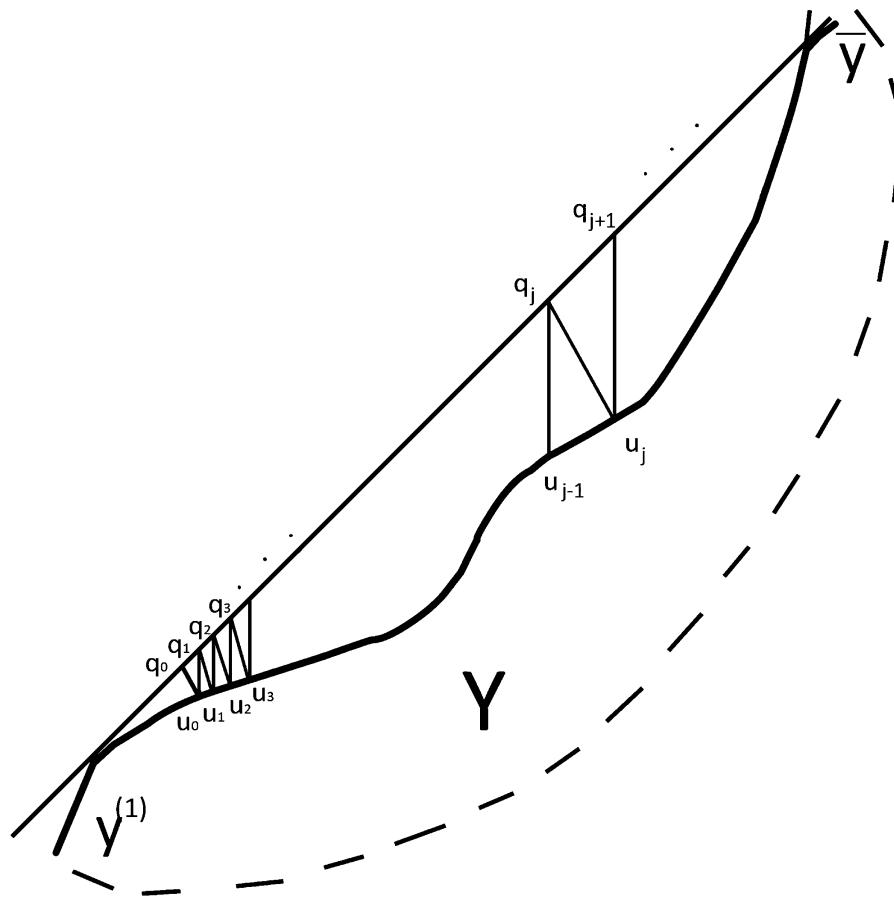


Рис. 5. Геометрическая интерпретация алгоритма A2

Шаг 2. Вычислим точку $q_{j+1} = u_{j+1} + \gamma h$, где γ найдем из условия $q_{j+1,1} = \bar{A}_1$, то есть $\gamma = u_{j+1,1} - \bar{A}_1$; зададим начальное приближение $y_n = u_{j+1}$; переопределим $j = j + 1$ и перейдем на шаг 1.

Шаг 3. Точка u_{j+1} является одновременно точкой $y^{(1)}$ и точкой t , в этом случае переходим на выполнение алгоритма A3.

Шаг 4. Точка u_{j+1} является точкой $y^{(1)}$. Переходим на выполнение алгоритма A2.

1.2.2. Алгоритм поиска на фрагменте множества $PP(Y)$. Алгоритм используется для поиска точки \bar{y} после того, как была найдена очередная точка $y^{(k)}$, $k \geq 1$, локального минимума функции φ на множестве Y (в нашем случае $y^{(1)}$), где \bar{y} — точка пересечения (если она существует) прямой L_1 из (7) и кривой U .

Для идентификации алгоритма назовем его $A2(y_n)$.

Пошаговое описание алгоритма приводится ниже, а геометрическая интерпретация схемы — на рис. 5.

В алгоритме генерируются последовательности точек $\{q_j\}$, $j \geq 0$, расположенных на прямой L_1 , и $\{u_j\}$, $u_j \in U$: если известна точка q_j , $j \geq 0$, то находим точку $u_j \in U$ (шаг 1 схемы) и определяем точку q_{j+1} на прямой L_1 так, чтобы $q_{j+1,1} = u_{j+1}$.

Шаг 0. Зададим начальные значения $y_n = y^{(1)}$, $q_0 = y^{(1)} + p \cdot \varepsilon \cdot e$; $j = 0$.

Шаг 1. Найдем u_j — локальное решение задачи: $\min_{y \in Y} \|q_j - y\|^2$ с начальным приближением y_n . Если представить функцию $\mu(y) = \|q_j - y\|^2$ в виде суммы $(q_{j1} - y_1)^2 + (q_{j2} - y_2)^2$,

то $\mu(y_n) = (q_{j2} - y_{n2})^2$, так как $q_{n1} = u_{j1} = q_{j1}$. Тогда при движении текущей точки y по кривой U влево (y_1 убывает) или вправо (y_1 растет) от точки y_n первое слагаемое растет от нуля. Второе слагаемое при движении точки $y \in U$ влево от точки y_n также растет, так как $y_2 < y_{n2}$ и, следовательно, $q_{j2} - y_2 > q_{j2} - y_{n2} > 0$ или $(q_{j2} - y_2)^2 > (q_{j2} - y_{n2})^2$.

Рассмотрим поведение функции μ и компоненты y_2 текущей точки y при движении вправо от точки y_n . При движении в этом направлении y_2 может как возрастать, так и убывать.

1. Компонента y_2 убывает, что означает $y_n = t$, и функция μ возрастает.

2. Компонента y_2 не убывает, но в точке y_n кривая U и окружность $Q = \{y : \mu(y) = \mu(y_n)\}$ касаются друг друга. Тогда функция μ не убывает при движении по кривой U в малой окрестности точки y_n .

Оба варианта можно определить при получении точки u_j , равной y_n . В этом случае переопределим $q_j = q_j + \varepsilon \cdot p \cdot e$ и найдем u_j — решение задачи $\min_{y \in Y} \mu(y)$ с начальным приближением y_n . Если опять получим $y_j = y_n$, что означает $y_n = t$, то перейдем на шаг 4. Если $u_j \neq y_n$, то проверим: если $\|q_j - u_j\| \leq \varepsilon$ и $j > 0$, то переходим на шаг 3.

3. Компонента y_2 возрастает, и кривая U пересекается в точке y_n с окружностью Q . Это означает, что функция μ убывает и $u_j \neq y_n$. Тогда, если $\|q_j - u_j\| > \varepsilon$, то переходим на шаг 2; в противном случае — на шаг 3.

Шаг 2. Вычислим точку $q_{j+1} = q_j + \gamma \cdot e$, где величина γ находится из условия $q_{j+1,1} = u_{j1}$, ($\gamma = u_{j1} - q_{j1}$), определим $y_n = u_j$, положим $j = j + 1$ и перейдем на шаг 1.

Шаг 3. Точку u_j берем в качестве начального приближения \bar{y} для алгоритма $A1(\bar{y})$. Далее с помощью алгоритма $A1(\bar{y})$ находим следующую точку $y^{(2)}$.

Шаг 4. Определим функцию $\bar{f}_2(x) = \max\{f_2(x), t_2\}$ и множество $\bar{Y} = \{y \in R^2 : y_1 = f_1(x), y_2 = \bar{f}_2(x), x \in X\}$ и продолжим поиск точки \bar{y} на множестве $PP(\bar{Y})$ (раздел 1.3.3).

1.2.3. Алгоритм поиска на фрагменте множества $PP(\bar{Y})$. Для демонстрации алгоритма возьмем множество \bar{Y} из примера множества Y , изображенного на рис. 3. Тогда множество \bar{Y} выглядит так, как это изображено на рис. 6, а $PP(\bar{Y})$ состоит из точки $A^0 = (A_1, t_2)$ и отрезка кривой между точками y^0 и B , но без точки y^0 .

Для идентификации алгоритма назовем его $A3(y^{(k)}, t)$. Первый параметр — последняя найденная точка локального минимума функции φ на множестве Y (в нашем случае $y^{(1)}$); второй параметр — точка t .

Пошаговое описание алгоритма приводится ниже, а его работу можно проследить по рис. 6.

Шаг 0. Вычислим точку $s = y^{(1)} + \gamma \cdot e$. Величина γ находится из условия $s_1 = t_1$, то есть $\gamma = t_1 - y_1^{(1)}$; определим точку $q_0 = s + (s_2 - t_2)e$, зададим начальное приближение $y_n = t$, положим $j = 0$.

Шаг 1. Найдем u_j — локальное решение задачи: $\min_{y \in \bar{Y}} \|q_j - y\|^2$ с начальным приближением y_n . Если $u_{j2} > t_2$, то переходим на шаг 3. Если $u_{j1} \leq q_{j1}$, то — на шаг 4.

Шаг 2. Вычислим $q_{j+1} = q_j + (q_{j2} - t_2) \cdot e$, определим $y_n = u_j$, положим $j = j + 1$, перейдем на шаг 1.

Шаг 3. При $u_{j2} > t_2$ автоматически множество \bar{Y} заменяется на множество Y . В этом случае необходимо вычислить $q_{j+1} = q_j + \eta \cdot e$, где η вычисляется из условия $q_{j+1,1} = u_{j1}$, то есть $\eta = u_{j1} - q_{j1}$, определим $y_n = u_j$, положим $j = j + 1$ и перейдем на шаг 1 алгоритма $A2$. Таким образом, если $t \neq B$, то выполнение алгоритма $A3$ обязательно должно переходить

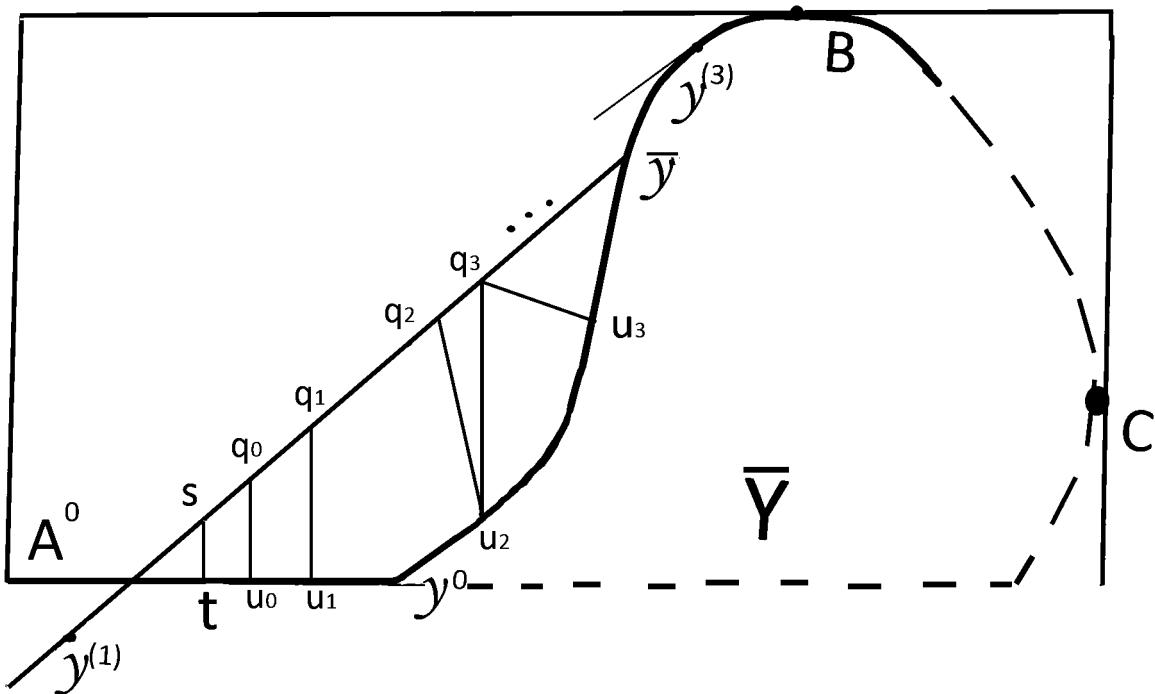


Рис. 6. Геометрическая интерпретация алгоритма A3

к выполнению алгоритма A2. На примере, изображенном на рис. 6, после нахождения точки u_2 начинает работать алгоритм A2.

Шаг 4. Ситуация, когда $u_{j1} \leq q_{j1}$, возможна только при условии, что последняя найденная точка t равна B . В этом случае $u_{j1} = C_1$, где C_1 — первая компонента точки C , определенной в (6). На этом процесс нахождения глобального минимума функции φ на множестве Y прекращаем. Решением задачи (4) является последняя найденная точка $y^{(k)}$ локального минимума функции на Y .

Сходимость алгоритма A2 (возможно совместно с алгоритмом A3) следует из следующего утверждения.

Теорема 3. Если имеется точка $y^{(k)}, k \geq 1$, локального минимума функции φ на множестве Y и существует точка \bar{y} пересечения прямой L_k , определенной в (7), и кривой U , то, стартовав из точки $y^{(k)}$, можно достичь точки \bar{y} с помощью алгоритмов A2 и A3.

Доказательство. В алгоритмах A2 и A3 генерируется последовательность точек $\{q_j\}, j = 0, 1, \dots$, расположенных на отрезке прямой между точками $y^{(k)}$ и \bar{y} и удовлетворяющих неравенствам: для всех $j \geq 0$ $\|q_{j+1} - y^{(k)}\| > \|q_j - y^{(k)}\|$ и $\|q_j - y^{(k)}\| \leq \|\bar{y} - y^{(k)}\|$. Из этих неравенств следует, что числовая последовательность $\{\|q_j - y^{(k)}\|\}, j = 0, 1, \dots$, строго возрастающая и ограниченная сверху, поэтому сходится.

Допустим, α — ее предел и $\alpha < \|\bar{y} - y^{(k)}\|$. Определим точку $q' = y^{(k)} + \frac{\alpha}{\sqrt{2}}e$, $e = (1, 1)^T$, которая лежит на отрезке прямой между точками $y^{(k)}, \bar{y}$ и $\|q' - y^{(k)}\| = \alpha$. Тогда $\|q' - \bar{y}\| = \|\bar{y} - y^{(k)}\| - \alpha > 0$ и $\min_{y \in Y} \|q' - y\|^2 > 0$. Последнее неравенство верно, так как в случае равенства точка q' являлась бы точкой пересечения прямой L_k и кривой U .

Если $y' = \arg \min_{y \in Y} \|q' - y\|^2$, то по определенным в алгоритме A2 правилам возможно продолжение генерации точек: $q'' = q' + \eta e$, где η находится из условия $q'' = y'_1$, то есть

$\eta = y'_1 - q'_1$. Точка q'' лежит на отрезке прямой между точками $y^{(k)}$ и \bar{y} ; $\|q'' - y^{(k)}\| > \|q_j - y^{(k)}\|$ для всех $j > 0$ и $\|q'' - y^{(k)}\| > \alpha$, что противоречит тому, что α — верхняя граница последовательности $\{\|q_j - y^{(k)}\|\}$ точек, сгенерированных в алгоритме $A2$.

Следовательно, $\lim_{j \rightarrow \infty} \|q_j - y^{(k)}\| = \|\bar{y} - y^{(k)}\|$. Точки $y^{(k)}, q_j, \bar{y}$ лежат на одной прямой, и точки q_j расположены между $y^{(k)}$ и \bar{y} , тогда $\|q_j - \bar{y}\| = \|\bar{y} - y^{(k)}\| - \|q_j - y^{(k)}\|$, следовательно $\|q_j - \bar{y}\| \rightarrow 0$ при $j \rightarrow \infty$. Теорема доказана.

Сходимость процесса поиска глобального минимума функции φ на множестве Y следует из следующих соображений.

Данный процесс можно представить в виде последовательности процедур поиска точек $y^{(k)}$, $k > 0$, локального минимума функции φ на Y , таких, что $\varphi(y^{(k+1)}) < \varphi(y^{(k)})$, ($y^{(0)} = A$). Как было сказано выше, таких точек конечное число, а процедура поиска каждой из них состоит из полного или неполного набора алгоритмов $A1, A2, A3$, каждый из которых сходится.

2. Решение задачи в пространстве исходных переменных. Перепишем алгоритмы, описанные в предыдущем разделе, для работы в пространстве исходных переменных. Имена всех алгоритмов сохранены. Оптимизационные задачи, которые решаются в этих алгоритмах, являются задачами невыпуклого программирования, исключение составляет задача $\min_{x \in X} f_1(x)$. Для каждой из них задается точка начального приближения, стартуя из которой находится точка локального минимума целевой функции. Метод минимизации можно выбирать исходя из характеристики целевой функции и способа задания допустимого множества.

2.1. Алгоритм $A1$. Алгоритм предназначен для нахождения локального минимума задачи $\min_{x \in X} f(x)$ с начальным приближением x_n , которое определяется двумя вариантами применения: а) x_n есть точка a из (5); б) x_n есть \bar{x} , найденная алгоритмом $A2$. В обоих вариантах работа алгоритма одинакова, поэтому для простоты изложения рассмотрим один из вариантов, например, первый.

Определим множество $X_1 = \{x \in X : F(x) = y, \text{ для любого } y \in U_1\}$, где U_1 — кривая, определенная в разделе 1.2.1. В этом же разделе приводятся правила построения минимизирующей последовательности точек $\{u_j\}$ функции φ , элементы которой принадлежат U_1 . Поэтому в множестве X_1 существует траектория, которая начинается в точке a и заканчивается в точке локального минимума функции f , и вдоль которой функция f убывает. Наша задача состоит в том, чтобы, опираясь на схему алгоритма $A1$ в критериальном пространстве, построить для функции f минимизирующую последовательность точек $\{x_j\}$, принадлежащих этой траектории.

Схема алгоритма $A1$ в пространстве исходных переменных мало отличается от схемы алгоритма в критериальном пространстве. Основное отличие касается только шага 1.

Шаг 0. Определим точки $u_0, A, \bar{A}, q_0 \in R^2 : A = F(a), u_0 = A, \bar{A} = A - \varepsilon \cdot p \cdot e_1, q_0 = u_0 + \varepsilon \cdot p \cdot h$; зададим начальное приближение $x_n = a$; положим $j = 0$.

Шаг 1. Определим функцию $\psi_j(x) = ((q_{j1} - f_1(x))^2 + (q_{j2} - f_2(x))^2)/2$ и найдем локальное решение x_{j+1} задачи $\min_{x \in X} \psi_j(x)$ с начальным приближением x_n . Вычислим $u_{j+1} = F(x_{j+1})$. Если $u_{j+1,2} \leq u_{j2}$, то перейдем на шаг 3. Если $\|u_{j+1} - u_j\| \leq \varepsilon$, то перейдем на шаг 4.

Шаг 2. Вычислим $q_{j+1} = u_{j+1} + (u_{j+1,1} - \bar{A}_1)h$, зададим начальное приближение $x_n = x_{j+1}$, переопределим $j = j + 1$ и перейдем на шаг 1.

Шаг 3. Точка x_j является точкой локального максимума функции f_2 на множестве X ($F(x_j) = t$) и одновременно точкой $x^{(1)}$ локального минимума функции f на X ($F(x_j) = y^{(1)}$). Зададим точку $s_0 = x_j$ и переходим на выполнение алгоритма A3.

Шаг 4. Точку x_{j+1} можно взять в качестве точки локального минимума функции f на X . Это следует из следующих соображений.

1. Последовательность $\{u_j\}$ сходится к $y^{(1)}$, и малость нормы $\|u_{j+1} - u_j\|$ можно рассматривать как признак близости по норме точки u_{j+1} к $y^{(1)}$.

2. Для любого малого $\delta > 0$ существует номер j_0 — такой, что при $j > j_0$ $\|u_j - y^{(1)}\| < \delta$.

Пусть $x^{(1)} \in X$ точка, такая, что $F(x^{(1)}) = y^{(1)}$. Тогда при $j > j_0$ $\|u_j - y^{(1)}\|^2 = (f_1(x_j) - f_1(x^{(1)}))^2 + (f_2(x_j) - f_2(x^{(1)}))^2 < \delta^2$ или $(f_1(x_j) - f_1(x^{(1)}))^2 < \delta_1^2$ и $(f_2(x_j) - f_2(x^{(1)}))^2 < \delta_2^2$ для некоторых $\delta_1, \delta_2 > 0$, таких, что $\delta_1^2 + \delta_2^2 = \delta^2$.

Из строгой выпуклости f_1 следует неравенство $f_1(x^{(1)}) - f_1(x_j) > \nabla f_1(x_j)^T(x^{(1)} - x_j) + \frac{1}{2}(x^{(1)} - x_j)^T H_j(x^{(1)} - x_j)$, где H_j — матрица Гессе функции f_1 в точке x_j ($H_j = H(x_j)$).

Последовательность $\{f_1(x_j)\}$ строго возрастающая, и $f_1(x_j) \rightarrow y_1^{(1)}$ при $j \rightarrow \infty$, а $y_1^{(1)} = f_1(x^{(1)})$. Тогда $\nabla f_1(x_j)^T(x^{(1)} - x_j) > 0$.

Пусть $l(x)$ — минимальное собственное число матрицы $H(x)$ ($l(x) > 0$) и $l_0 = \min_{x \in X} l(x)$. Тогда $f_1(x^{(1)}) - f_1(x_j) > \frac{1}{2}l_0\|x^{(1)} - x_j\|^2$, то есть $\|x^{(1)} - x_j\|^2 < \frac{2\delta_1}{l_0}$. Это значит, что при больших j x_j близко по норме к $x^{(1)}$.

Итак, точку x_{j+1} можно взять в качестве точки $x^{(1)}$ и перейти на выполнение алгоритма A2. На этом описание схемы алгоритма A1 закончим.

Замечание. Рассмотрим градиент функции ψ_j в точке начального приближения $x_n = x_j$: $\nabla \psi_j(x_j) = (f_1(x_j) - q_{j1})\nabla f_1(x_j) + (f_2(x_j) - q_{j2})\nabla f_2(x_j)$.

По правилам построения последовательности $\{q_j\}$ имеем $f_1(x_j) - q_{j1} = q_{j2} - f_2(x_j) > 0$, то есть $\nabla \psi_j(x_j) = (f_1(x_j) - q_{j1})(\nabla f_1(x_j) - \nabla f_2(x_j)) = (f_1(x_j) - q_{j1})\nabla f(x_j)$.

Это означает, что конусы векторов спуска функции ψ_j и f из точки x_j , совпадают. Однако заменить задачу $\min_{x \in X} \psi_j(x)$ с $x_n = x_j$ на задачу $\min_{x \in X} f(x)$ с $x_n = x_j$ нельзя, так как очередная точка x_{j+1} минимизирующей последовательности должна удовлетворять следующим требованиям: а) $\nabla f_1(x_j)^T(x_{j+1} - x_j) > 0$; б) $\nabla f(x_j)^T(x_{j+1} - x_j) < 0$; в) не существует точки $x \in X$, такой, чтобы одновременно выполнялись неравенства $f_1(x) < f_1(x_{j+1})$, $f_2(x) > f_2(x_{j+1})$. Последнее условие означает, что точка $y_{j+1} = F(x_{j+1})$ должна принадлежать $PP(Y)$.

При решении первой задачи эти требования выполняются автоматически, а при решении второй задачи учесть первые два условия легко, а третье требование выполнить достаточно трудно.

Решать задачу $\min_{x \in X} \psi_j(x)$ на начальном этапе (при малых j) можно с малой точностью, и увеличивать точность необходимо по мере приближения к точке локального минимума функции f на множестве X .

2.2. Алгоритм A2. Данный алгоритм предназначен для поиска точки \bar{x} — прообраза точки \bar{y} (рис. 5) при отображении F .

Шаг 0. Зададим начальные значения $x_n = x^{(k)}$, ($k > 1$), $q_0 = F(x^{(k)}) + p \cdot \varepsilon \cdot e$, где p, ε, e — величины, определенные в разделе 1.2.2; положим $j = 0$, $ip = 0$.

Шаг 1. Определим функцию $\psi_j(x) = ((q_{j1} - f_1(x))^2 + (q_{j2} - f_2(x))^2)/2$ и найдем s_j — локальное решение задачи: $\min_{x \in X} \psi_j(x)$ с начальным приближением x_n . Если $s_j = x_n$ и $ip = 0$,

то переопределим $q_j = q_j + \varepsilon \cdot p \cdot e$, положим $ip = 1$ и перейдем на шаг 1. Если $s_j = x_n$ и $ip = 1$, то перейдем на шаг 4.

Если $\psi_j(s_j) \leq \varepsilon$ и $j > 0$, то перейдем на шаг 3.

Шаг 2. Вычислим $q_{j+1} = q_j + (f_1(s_j) - q_{j1}) \cdot e$, определим $x_n = s_j$, положим $j = j + 1$ и перейдем на шаг 1.

Шаг 3. Точку s_j берем в качестве \bar{x} — начального приближения для алгоритма A1. Работа алгоритма A2 прекращается.

Шаг 4. Точка s_j является точкой локального максимума функции f_2 на множестве X $t = F(s_j)$. В этом случае определим функцию $\bar{f}_2(x) = \max\{f_2(x), f_2(s_j)\}$, зададим точку $s_0 = s_j$ и переходим на выполнение алгоритма A3.

Замечание. Решать с высокой точностью задачу $\min_{x \in X} \psi_j(x)$ на начальном этапе (при малых значениях j) нет необходимости, но при относительно малых значениях $\psi_j(s_j)$ точность следует увеличивать.

2.3. Алгоритм A3. В пространстве исходных переменных данный алгоритм позволяет преодолевать участок границы множества X вокруг точки s_0 — локального максимума функции f_2 , на котором $f_2(x) \leq f_2(s_0)$, то есть выполняет ту же функцию, что одноименный алгоритм в критериальном пространстве, позволяющий преодолевать овраги. Хотя в основу этих алгоритмов заложена одна и та же идея преодоления препятствий, их схемы немного отличаются друг от друга. Это связано с тем, что автоматически заменить задачу $\min_{y \in \bar{Y}} \|q_j - y\|^2$ в шаге 2 из раздела 1.2.3 на задачу $\min_{x \in X} \psi_j(x)$, где $\psi_j(x) = (f_1(x) - q_{j1})^2 + (\bar{f}_2(x) - q_{j2})^2$, нельзя. В первом алгоритме переменная y_1 монотонно возрастает на отрезке прямой между точками t и y^0 (рис. 6). В пространстве исходных переменных на множестве $\{x \in X : f_2(x) \leq t_2\}$ может существовать точка x' локального максимума функции f_1 на множестве X , а $\psi_j(x) = (f_1(x) - q_{j1})^2 + (t_2 - q_{j2})^2$. При возникновении такой ситуации возможно существование номера j_0 , такого, что при $j > j_0$ все задачи $\min_{x \in X} \psi_j(x)$ имеют одно и то же решение x' и процесс зацикливается.

Чтобы избежать возникновения подобной ситуации, предлагается сначала минимизировать функцию ψ_j на множестве $\{x \in R^n : f_2(x) \leq t_2\}$ с заданным начальным приближением и в зависимости от того, принадлежит или не принадлежит полученное решение множеству X , принимать решение о дальнейших действиях.

На входе алгоритма два параметра: s_0 — точка локального максимума функции f_2 на множестве X ; $x^{(k)}$ — последняя найденная точка локального максимума функции f на X .

Шаг 0. Вычислим точки $t = F(s_0)$, $y^{(k)} = F(x^{(k)})$, $s = y^{(k)} + (t_1 - y_1^{(k)})e$, определим $q_0 = s + (s_2 - t_2)e$, зададим начальное приближение $x_n = s_0$, положим $j = 0$.

Шаг 1. Найдем локальное решение s_j задачи $\min(q_{j1} - f_1(x))^2$ на множестве $\{x \in R^n : f_2(x) \leq t_2\}$ с начальным приближением x_n . Если $s_j \in X$, то перейдем на шаг 3. Если $f_1(s_j) \leq q_{j1}$, то перейдем на шаг 4.

Шаг 2. Вычислим $q_{j+1} = q_j + (q_{j2} - t_2) \cdot e$, определим $x_n = s_j$, положим $j = j + 1$, перейдем на шаг 1.

Шаг 3. Решим задачу $\min_{x \in X} ((f_1(x) - q_{j1})^2 + (f_2(x) - q_{j2})^2)$ с начальным приближением s_j . Если x_j — решение этой задачи, то точка $u_j = F(x_j)$ является псевдоэффективной. Определим $q_{j+1} = y^{(k)} + (u_{j1} - y_1^{(k)})e$, назначим $x_n = x_j$ и перейдем на шаг 1 алгоритма A2.

Шаг 4. Обнаружена ситуация, когда последняя найденная точка t равна B . Процесс поиска глобального решения задачи (1) заканчиваем, а ее решением является последняя найденная точка $x^{(k)}$.

Заключение. Алгоритм нахождения глобального минимума задачи (1) предлагается представить в виде последовательности задач поиска локального минимума невыпуклых функций на выпуклом множестве при заданном начальном приближении, причем в большинстве задач достаточно найти несколько точек минимизирующей последовательности. Выбор метода минимизации зависит от характеристики целевой функции и способа задания допустимого множества.

Список литературы

1. СТРЕКАЛОВСКИЙ А. С. Элементы невыпуклой оптимизации. Новосибирск: Наука, 2003.
2. Подиновский В. В., Ногин В. Д. Парето-оптимальные решения многокритериальных задач. М.: Наука, 1982.

*Котельников Евгений Алексеевич —
старш. науч. сотрудник Ин-та вычислительной математики
и математической геофизики СО РАН,
тел.: (393) 330-60-66.*

Дата поступления — 15.01.2016