ИЗВЕСТИЯ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

1998 МАТЕМАТИКА № 12 (439)

УДК 519.853

И.Я. ЗАБОТИН

ОБ ОДНОМ ПОДХОДЕ К ПОСТРОЕНИЮ АЛГОРИТМОВ БЕЗУСЛОВНОЙ МИНИМИЗАЦИИ ПСЕВДОВЫПУКЛЫХ ФУНКЦИЙ

Введение

Для безусловной минимизации непрерывно дифференцируемой псевдовыпуклой функции предлагается общий метод, в котором переход от одной итерационной точки к другой производится в некоторых линейных многообразиях. Размерность многообразий задается на каждом шаге произвольно. За счет большой свободы в выборе этих многообразий метод допускает разные реализации, среди которых есть как новые алгоритмы, так и известные методы минимизации выпуклых функций, например, обобщенный метод градиентного спуска, метод сопряженных градиентов, метод Ньютона с регулировкой шага, квазиньютоновские алгоритмы и др. Исследуется сходимость метода, приводятся оценки его скорости сходимости для псевдовыпуклых и сильно выпуклых функций. Отметим, что на основе предложенного подхода можно строить алгоритмы, допускающие распараллеливание процесса решения исходной задачи, получать различные смешанные алгоритмы без дополнительных обоснований их сходимости, исследовать устойчивость некоторых алгоритмов к ошибкам вычисления градиентов.

Решается задача

$$\min_{x\in R_n} f(x), \tag{1}$$

где $f(x)$ — достигающая своего минимального значения в n-мерном евклидовом пространстве R_n непрерывно дифференцируемая псевдовыпуклая функция ([1], [2]), т. е. для любых $x, y \in R_n$ выполняется импликация

$$\langle f'(x), y-x\rangle \ge 0 \;\Rightarrow\; f(y) \ge f(x),$$

$f'(x)$ — градиент функции $f(x)$ в точке x.

Положим $f^* = \min\limits_{x\in R_n} f(x)$, $X^* = \{x \in R_n : f(x) = f^*\}$, $K = \{0, 1, \dots\}$.

1. Общий метод решения задачи

Предлагаемый метод решения задачи (1) вырабатывает последовательность приближений x_k, $k \in K$, и заключается в следующем. Выбирается точка $x_0 \in R_n$ начального приближения. Пусть построена точка x_k, $k \in K$.

1. Если $f'(x_k) = 0$, то $x_k \in X^*$, и процесс прекращается.

2. Выбираются вектор $r_k \in R_n$ и число $\varepsilon_k \ge 0$ так, чтобы

$$\|r_k\| \le \varepsilon_k, \tag{2}$$

и $g_k = f'(x_k) + r_k \ne 0$. Полагается $U(x_k) = \{x \in R_n : x = x_k + \alpha g_k,\ \alpha \in R_1\}$.

3. Задается H_k — конечное множество индексов, выбираются векторы h_i^k, $i \in H_k$, так, чтобы вектор g_k был их линейной комбинацией, строится линейное многообразие

$$M(x_k) = \left\{x \in R_n : x = x_k + \sum_{i\in H_k} \alpha_i h_i^k,\ \alpha_i \in R_1\right\}$$

с условием, что $x_k + g_k \in M(x_k)$.

4. Отыскивается такая точка $y_k \in M(x_k)$, что

$$f(y_k) \leq f_k^* + \delta_k \tag{3}$$

либо

$$f(y_k) \leq (1 - \nu_k) f(x_k) + \nu_k f_k^*, \tag{4}$$

где $f_k^* = \min\limits_{x \in U(x_k)} f(x)$, $\delta_k \geq 0$, $0 < \nu \leq \nu_k \leq 1$.

5. Приближение $x_{k+1} \in R_n$ выбирается из условия

$$f(x_{k+1}) \leq f(y_k).$$

Дополнительные требования к выбору чисел ε_k, δ_k и векторов r_k в методе будут приведены ниже в теоремах сходимости.

Заметим, что выход из цикла на шаге 1 метода связан с выполнением достаточного условия минимума псевдовыпуклой функции ([1], [2]). В число векторов h_i^k, $i \in H_k$, можно включать, например, вектор g_k, тогда условие, высказанное в п. 3, будет выполняться. В методе можно положить $r_k = 0$ $\forall k \in K$, т. е. векторы h_i^k достаточно выбирать так, чтобы их линейной комбинацией был градиент $f'(x_k)$. При построении $M(x_k)$ можно положить $\alpha_i = 0$ для некоторых $i \in H_k$, но при этом важно, чтобы выполнялось включение $x_k + g_k \in M(x_k)$.

Вектор r_k можно интерпретировать, например, как погрешность в вычислении $f'(x_k)$ на k-м шаге. В ([3], с. 96) векторы r_k названы детерминированными помехами в вычислении градиента, причем абсолютными, если $\varepsilon_k = \varepsilon$, и относительными, если $\varepsilon_k = \varepsilon \|f'(x_k)\|$, где $\varepsilon \geq 0$. Там же и, например, в ([4], с. 92) обсуждаются причины возникновения этих видов помех.

Отметим также, что на шаге 5 точку x_{k+1} можно отыскивать любым известным методом или положить

$$x_{k+1} = y_k \quad \forall k \in K. \tag{5}$$

Пусть

$$m_k^* = \min_{x \in M(x_k)} f(x).$$

Так как по выбору векторов h_i^k выполняется включение $U(x_k) \subset M(x_k)$, то $m_k^* \leq f_k^*$. Если значение m_k^* удается оценить, то точку y_k на шаге 4 можно выбирать в $M(x_k)$ из условий $f(y_k) \leq m_k^* + \delta_k$ или $f(y_k) \leq (1 - \nu_k) f(x_k) + \nu_k m_k^*$. Ясно, что в таком случае неравенства (3) или соответственно (4) будут выполняться.

Условия (3), (4) выбора точки y_k из многообразия $M(x_k)$ именно с помощью значения f_k^*, а не m_k^*, позволяют минимизировать одномерную функцию, отыскивая значение f_k^*, и параллельно с помощью любого известного метода проводить спуск из точки x_k в многообразии $M(x_k)$ до выполнения неравенств (3) или (4). Кроме того, отыскание приближения $x_{k+1} \in R_n$ можно вести, не строя многообразия $M(x_k)$, любым известным алгоритмом, начиная с точки x_k, во всем пространстве R_n до выполнения условия $f(x_{k+1}) \leq (1 - \nu_k) f(x_k) + \nu_k f_k^*$, поскольку в таком случае можно считать $M(x_k) = \{x \in R_n : x = \alpha_1 g_k + \alpha_2 (x_{k+1} - x_k),\ \alpha_1, \alpha_2 \in R_1\}$. Последнее замечание позволяет на основе общего метода строить смешанные алгоритмы, в которых можно от шага к шагу менять способы построения подходящих направлений и шаговых множителей.

Пусть вектор r_k при некотором $k \in K$ удовлетворяет помимо (2) условию

$$\langle f'(x_k), r_k \rangle \geq -\gamma \|f'(x_k)\|^2, \tag{6}$$

где $0 \leq \gamma < 1$. Покажем, что тогда точку $y_k \in M(x_k)$ можно выбрать так, что вместе с условием (3) или (4) выполнится неравенство

$$f(y_k) < f(x_k). \tag{7}$$

Действительно, пусть точка $x_k^* \in M(x_k)$ такова, что $f(x_k^*) = m_k^*$. Так как $x_k \in M(x_k)$, то $f(x_k^*) \leq f(x_k)$. Допустим, что последнее соотношение выполняется как равенство. Тогда в силу выбора x_k^* для всех $\alpha_i \in R_1$, $i \in H_k$, выполняется неравенство

$$f(x_k) \leq f\Big(x_k + \sum_{i\in H_k} \alpha_i h_i^k\Big). \tag{8}$$

Поскольку $f'(x_k) \neq 0$ и справедливо (6), то

$$\langle f'(x_k), g_k\rangle \geq (1-\gamma)\|f'(x_k)\|^2 > 0.$$

Значит, $-g_k$ — направление спуска для функции $f(x)$ в точке x_k, и найдется такое число $\tau < 0$, что $f(x_k + \tau g_k) < f(x_k)$. По условию выбора векторов h_i^k существуют такие числа τ_i^k, $i \in H_k$, что $g_k = \sum\limits_{i\in H_k} \tau_i^k h_i^k$. Следовательно, $f\Big(x_k + \sum\limits_{i\in H_k} \alpha_i^k h_i^k\Big) < f(x_k)$, где $\alpha_i^k = \tau\tau_i^k$, $i \in H_k$. Последнее неравенство противоречит утверждению (8) при $\alpha_i = \alpha_i^k$, $i \in H_k$. Таким образом, $f(x_k^*) < f(x_k)$, и в качестве точки y_k, удовлетворяющей условиям (3), (7) или (4), (7), можно взять точку x_k^*, поскольку $m_k^* \leq f_k^*$.

Отметим, что условие (6) предоставляет бóльшую свободу в выборе на k-й итерации векторов r_k, g_k, а значит, и многообразия $M(x_k)$, в котором выбирается вспомогательная точка y_k.

Условия (3), (4) выбора y_k требуют приближенной минимизации одномерной функции $f(x_k + \alpha g_k)$. Удобно иметь критерий задания точек y_k, а значит и x_{k+1}, который не использует оценок значений f_k^*. Пусть $f'(x)$ удовлетворяет в R_n условию Липшица с константой L, и вектор $-g_k$, построенный на шаге 2 метода, является направлением спуска для $f(x)$ в точке x_k. Тогда точку $y_k \in M(x_k)$ можно выбирать так, чтобы

$$f(y_k) \leq f(x_k - \beta_k g_k), \tag{9}$$

где $\beta_k = \rho_k\langle f'(x_k), g_k\rangle/\|g_k\|^2$, $0 < \rho' \leq \rho_k \leq 2(1-\rho'')/L$, $0 < \rho'' < 1$. В частности, если положить $\rho' = 1/L$, $\rho'' = 1/2$, то $\beta_k = \langle f'(x_k), g_k\rangle/L\|g_k\|^2$. Ясно, что такой способ выбора y_k приемлем лишь в случае, когда известна константа L или ее верхняя оценка.

2. Исследование сходимости метода

Всюду в этом параграфе будем считать, что последовательность $\{x_k\}$ построена по предложенному выше методу. Исследуем сначала сходимость релаксационного варианта метода с выбором точек y_k согласно (3) или (4).

Теорема 1. *Пусть*

$$\varepsilon_k \leq \varepsilon < \infty \quad \forall k \in K, \tag{10}$$

$$\delta_k \to 0, \quad k \to \infty, \tag{11}$$

векторы r_k, $k \in K$, удовлетворяют неравенствам (2), (6), *и точки y_k для всех $k \in K$ выбраны с условием* (3), (7) *или* (4), (7). *Если множество*

$$X^0 = \{x \in R_n : f(x) \leq f(x_0)\}$$

ограничено, то для последовательности $\{x_k\}$ справедливо равенство

$$\lim_{k\to\infty} f(x_k) = f^*. \tag{12}$$

Доказательство. Так как согласно (7) последовательность $\{f(x_k)\}$ монотонно убывает и по условию $f^* > -\infty$, то существует

$$\lim_{k\to\infty} f(x_k) \geq f^* \tag{13}$$

и, значит,

$$\lim_{k\to\infty} (f(x_{k+1}) - f(x_k)) = 0. \tag{14}$$

Допустим, что неравенство (13) выполняется как строгое, т. е. утверждение теоремы неверно. Так как $x_k \in X^0$ $\forall k \in K$, а множество X^0 ограничено, то у последовательности $\{x_k\}$, $k \in K$, существует сходящаяся подпоследовательность $\{x_k\}$, $k \in K_1 \subset K$, с некоторой предельной точкой $\overline{x}$. Отметим, что по предположению $\overline{x} \notin X^*$. Согласно (2), (10) последовательность $\{r_k\}$, $k \in K_1$, ограничена. Выделим из нее сходящуюся подпоследовательность $\{r_k\}$, $k \in K_2 \subset K_1$. Пусть $\overline{r}$ — предельная точка этой подпоследовательности. Положим $\overline{g} = f'(\overline{x}) + \overline{r}$. Поскольку $f'(\overline{x}) \neq 0$ и согласно (6) $\langle f'(\overline{x}), \overline{r}\rangle \geq -\gamma\|f'(\overline{x})\|^2$, то $\langle f'(\overline{x}), \overline{g}\rangle \geq (1-\gamma)\|f'(\overline{x})\|^2 > 0$, и $-\overline{g}$ — направление спуска для функции $f(x)$ в точке $\overline{x}$. Следовательно, существует такое число $\tau < 0$, что $f(\overline{x} + \tau\overline{g}) < f(\overline{x})$. Положим $f(\overline{x} + \tau\overline{g}) - f(\overline{x}) = \sigma$. Так как $\sigma < 0$ и функция $f(x)$ непрерывно дифференцируема, то существуют такие окрестности ω_1 точки $\overline{x}$ и ω_2 точки $\overline{r}$, что $f(x+\tau(f'(x)+r)) - f(x) \leq \sigma/2$ $\forall x \in \omega_1$, $r \in \omega_2$. Так как $x_k \to \overline{x}$ и $r_k \to \overline{r}$ при $k \to \infty$, $k \in K_2$, то существует такой номер N, что $x_k \in \omega_1$, $r_k \in \omega_2$ $\forall k \geq N$, $k \in K_2$. Следовательно, $f(x_k + \tau g_k) - f(x_k) \leq \sigma/2$ $\forall k \geq N$, $k \in K_2$, а поскольку $f_k^* \leq f(x_k + \tau g_k)$, то

$$f_k^* - f(x_k) \leq \sigma/2 \quad \forall k \geq N, \quad k \in K_2. \tag{15}$$

Согласно (3), (4) для всех $k \in K$ выполняется $f(x_{k+1}) - \delta_k \leq f_k^*$ или $(f(x_{k+1}) - f(x_k))/\nu_k + f(x_k) \leq f_k^*$. Отсюда и из (15) для всех $k \geq N$, $k \in K_2$, следуют неравенства $f(x_{k+1}) - f(x_k) - \delta_k \leq \sigma/2$ или $f(x_{k+1}) - f(x_k) \leq \nu\sigma/2$, которые в силу (14), (11) при достаточно больших значениях $k \in K_2$ становятся противоречивыми. □

Ниже доказывается сходимость нерелаксационного варианта метода с выбором y_k согласно условию (3). Однако при этом накладываются более жесткие требования на функцию $f(x)$ и выбор последовательности $\{\delta_k\}$.

Лемма. *Пусть $f'(x)$ удовлетворяет в R_n условию Липшица с константой L, подмножества $K' \subset K$ и $K'' \subset K$ таковы, что точки y_k при построении последовательности $\{x_k\}$ выбираются для всех $k \in K'$ согласно* (3), *а для всех $k \in K''$ — согласно* (4). *Тогда*

$$f(x_k) - f(x_{k+1}) \geq \alpha\langle f'(x_k), g_k\rangle - (L\alpha^2/2)\|g_k\|^2 - \delta_k \quad \forall \alpha \in R_1, \ k \in K', \tag{16}$$

$$f(x_k) - f(x_{k+1}) \geq \nu_k\alpha\langle f'(x_k), g_k\rangle - (L\alpha^2/2)\|g_k\|^2 \qquad \forall \alpha \in R_1, \ k \in K''. \tag{17}$$

Доказательство. Пусть $k \in K'$. Тогда согласно (3) $f(x_{k+1}) \leq f_k^* + \delta_k \leq f(x_k - \alpha g_k) + \delta_k$, а значит, $f(x_k) - f(x_{k+1}) \geq f(x_k) - f(x_k - \alpha g_k) - \delta_k$ $\forall \alpha \in R_1$. Отсюда с учетом известного ([5], с. 93) неравенства для функций, градиент которых удовлетворяет условию Липшица, следует оценка (16). Пусть $k \in K''$. Тогда в силу (4) $f(x_k) - f(x_{k+1}) \geq \nu_k(f(x_k) - f(x_k - \alpha g_k))$ $\forall \alpha \in R_1$. Отсюда с учетом неравенства $\nu_k \leq 1$ следует (17). □

Теорема 2. *Пусть для всех $k \in K$ выполняются условия* (6), (10), *точки y_k при построении последовательности $\{x_k\}$ выбраны согласно* (3), *а числа δ_k, $k \in K$, таковы, что*

$$\sum_{k=0}^{\infty} \delta_k = \delta < \infty. \tag{18}$$

Если множество $X_\delta^0 = \{x \in R_n : f(x) \leq f(x_0) + \delta\}$ ограничено, $f'(x)$ удовлетворяет в R_n условию Липшица с константой L, то справедливо равенство (12).

Доказательство. Из (16) с учетом условия (6) и способа задания вектора g_k следует неравенство $f(x_k) - f(x_{k+1}) \geq \alpha(1-\gamma)\|f'(x_k)\|^2 - (L\alpha^2/2)\|g_k\|^2 - \delta_k$ $\forall \alpha \geq 0$, $k \in K$. Тогда

$$f(x_k) - f(x_{k+1}) \geq \max_{\alpha \geq 0}(\alpha(1-\gamma)\|f'(x_k)\|^2 - (L\alpha^2/2)\|g_k\|^2) - \delta_k =$$

$$= (1-\gamma)^2\|f'(x_k)\|^4/(2L\|g_k\|^2) - \delta_k, \ k \in K, \tag{19}$$

и, значит,

$$f(x_{k+1}) \leq f(x_k) + \delta_k, \quad k \in K. \tag{20}$$

Пусть $m \in K \setminus \{0\}$. Суммируя неравенства (20) по k от 0 до $m-1$, получим

$$f(x_m) \le f(x_0) + \sum_{k=0}^{m-1} \delta_k \le f(x_0) + \delta, \quad m = 1, 2, \ldots$$

Таким образом,

$$x_k \in X_\delta^0 \quad \forall k \in K, \tag{21}$$

а поскольку множество X_δ^0 ограничено, то

$$\|f'(x_k)\| \le c_0 < \infty \quad \forall k \in K. \tag{22}$$

Так как функция $f(x)$ достигает в R_n своего минимального значения, то из (20) следует ([5], с. 93) существование предела (13) и равенство (14). В силу (22), (2), (10) последовательность $\{\|g_k\|\}$, $k \in K$, ограничена. Значит, из (19), (14), (18) следует равенство $\lim\limits_{k\to\infty} \|f'(x_k)\| = 0$. Выделим из ограниченной последовательности $\{x_k\}$, $k \in K$, сходящуюся подпоследовательность $\{x_k\}$, $k \in K_1 \subset K$. Пусть z — ее предельная точка. Тогда $\lim\limits_{k\in K_1} \|f'(x_k)\| = \|f'(z)\| = 0$, и z — точка минимума псевдовыпуклой функции $f(x)$. Следовательно, $\lim\limits_{k\in K_1} f(x_k) = f(z) = f^*$, а поскольку последовательность $\{f(x_k)\}$, $k \in K$, сходится, то справедливо (12). $\square$

Обоснуем, наконец, сходимость метода в случае, когда точки y_k выбираются согласно неравенству (9).

Теорема 3. *Пусть $f'(x)$ удовлетворяет в R_n условию Липшица с константой L, и выполняются условия* (2), (6), (10), (9). *Если множество X^0 ограничено, то справедливо равенство* (12).

Доказательство. Как уже отмечалось, в силу (6) $-g_k$ — направление спуска для $f(x)$ в точке x_k. Поскольку $f(x_{k+1}) \le f(x_k - \beta_k g_k)$, то $f(x_k) - f(x_{k+1}) \ge \beta_k \langle f'(x_k), g_k\rangle - (L/2)\beta_k^2 \|g_k\|^2 = (1 - (L/2)\rho_k)\rho_k \langle f'(x_k), g_k\rangle^2/\|g_k\|^2 \ge \rho'\rho'' \langle f'(x_k), g_k\rangle^2/\|g_k\|^2 > 0$ $\forall k \in K$. Отсюда с учетом ограниченности последовательности $\{f(x_k)\}$ и множества X^0 следует равенство (14) и ограниченность сверху последовательности $\{\|g_k\|\}$. Значит, в силу (6) $\lim\limits_{k\to\infty} \langle f'(x_k), g_k\rangle \ge (1-\gamma) \lim\limits_{k\to\infty} \|f'(x_k)\|^2 = 0$. Как и при доказательстве теоремы 2, отсюда легко следует (12). $\square$

Подчеркнем, что сходимость метода обоснована в теоремах 1–3 без предположения убывания значений ε_k от шага к шагу, т. е. последовательность $\{r_k\}$ может не стремиться к нулю.

Получим теперь оценки скорости сходимости метода при более жестких условиях на функцию $f(x)$ и выбор чисел ε_k. Положим $(f(x) - f^*)/\|f'(x)\| = \theta(x)$ $\forall x \notin X^*$, $a_k = f(x_k) - f^*$.

Теорема 4. *Пусть функция $f(x)$ такова, что*

$$\theta(x) \le \theta < \infty \quad \forall x \in X_\delta^0 \setminus X^*, \tag{23}$$

где величина δ определена в (18), *$f'(x)$ удовлетворяет в R_n условию Липшица с константой L, при построении последовательности $\{x_k\}$ точки y_k выбраны для всех $k \in K$ только одним из трех описанных выше способов,*

$$\varepsilon_k \le \varepsilon \|f'(x_k)\|, \ 0 < \varepsilon < 1, \ k \in K, \tag{24}$$

и $0 \le \delta_k \le \eta k^{-2}$, $\eta > 0$, $k = 1, 2, \ldots$ Тогда справедлива оценка

$$0 \le a_k \le c/k, \quad k = 1, 2, \ldots, \tag{25}$$

где $c > 0$.

Доказательство. В силу (2), (24) $\|r_k\| \le \varepsilon \|f'(x_k)\|$ $\forall k \in K$. Значит,

$$\|g_k\| \le (1+\varepsilon)\|f'(x_k)\| \quad \forall k \in K \tag{26}$$

и

$$\langle f'(x_k), g_k\rangle = \|f'(x_k)\|^2 + \langle f'(x_k), r_k\rangle \ge (1-\varepsilon)\|f'(x_k)\|^2 > 0 \quad \forall k \in K. \tag{27}$$

Докажем (25), считая, что точки y_k выбраны при всех $k \in K$ согласно (3). Из (16), (26) для всех $\alpha \ge 0$, $k \in K$ следуют оценки $f(x_k) - f(x_{k+1}) \ge \alpha\|f'(x_k)\|^2 + \alpha\langle f'(x_k), r_k\rangle - (L\alpha^2/2)\|g_k\|^2 - \delta_k \ge \alpha\|f'(x_k)\|^2 - \alpha\|f'(x_k)\|\|r_k\| - (L\alpha^2/2)(1+\varepsilon)^2\|f'(x_k)\|^2 - \delta_k \ge \alpha(1-\varepsilon)\|f'(x_k)\|^2 - \alpha^2(L(1+\varepsilon)^2/2)\|f'(x_k)\|^2 - \delta_k$. Отсюда

$$a_k - a_{k+1} = f(x_k) - f(x_{k+1}) \ge \max_{\alpha \ge 0}(\alpha(1-\varepsilon) - \alpha^2 L(1+\varepsilon)^2/2)\|f'(x_k)\|^2 - \delta_k =$$
$$= (1-\varepsilon)^2\|f'(x_k)\|^2/(2L(1+\varepsilon)^2) - \delta_k, \tag{28}$$

$k \in K$, т.е. справедливо (20), а значит, и (21). Тогда согласно (23) $a_k \le \theta\|f'(x_k)\|$, $k \in K$, и из (28) имеем

$$a_k - a_{k+1} \ge a_k^2(1-\varepsilon)^2/(2L\theta^2(1+\varepsilon)^2) - \eta k^{-2}, \quad k = 1, 2, \ldots$$

Следовательно, $a_{k+1} \le a_k - a_k^2/\eta' + \eta' k^{-2}$, $k = 1, 2, \ldots$, где $\eta' = \max\{\eta, 2L\theta^2(1+\varepsilon)^2/(1-\varepsilon)^2\}$. Из последнего неравенства и леммы 2.3.5 ([5], с. 95) вытекает (25).

Пусть теперь точки y_k выбраны для всех $k \in K$ согласно (4). Из (17), (26), (27) и неравенства $\nu < \nu_k$ для всех $\alpha \ge 0$, $k \in K$ следуют оценки $f(x_k) - f(x_{k+1}) \ge \nu\alpha\langle f'(x_k), g_k\rangle - (L\alpha^2/2)\|g_k\|^2 \ge \nu\alpha(1-\varepsilon)\|f'(x_k)\|^2 - (L\alpha^2/2)(1+\varepsilon)^2\|f'(x_k)\|^2$. Отсюда

$$a_k - a_{k+1} = f(x_k) - f(x_{k+1}) \ge \max_{\alpha \ge 0}(\alpha\nu(1-\varepsilon) - \alpha^2 L(1+\varepsilon)^2/2)\|f'(x_k)\|^2 =$$
$$= \nu^2(1-\varepsilon)^2\|f'(x_k)\|^2/(2L(1+\varepsilon)^2) > 0, \tag{29}$$

$k \in K$. Значит, $x_k \in X^0 \subset X^0_\delta$ и в силу (23) $\theta(x_k) \le \theta$ для всех $k \in K$. Из (29) и леммы 2.3.4 ([5], с. 94) с учетом последнего неравенства следует оценка (25).

Пусть, наконец, точки y_k выбраны для всех $k \in K$ из условия (9). Используя оценку величины $a_k - a_{k+1}$, полученную в теореме 3, а также неравенства (6), (26), нетрудно показать, что $a_k - a_{k+1} \ge \rho'\rho''(1-\gamma)^2\|f'(x_k)\|^2/(1+\varepsilon)^2 > 0$ $\forall k \in K$. Значит, $x_k \in X^0$, и в силу (23) $a_k - a_{k+1} \ge \rho'\rho''(1-\gamma)^2 a_k^2/\theta^2(1+\varepsilon)^2$ $\forall k \in K$. Из этого неравенства и леммы 2.3.4 ([5], с. 94) следует (25). $\square$

Заметим, что если функция $f(x)$ выпукла, а множество X^0_δ ограничено, то условие (23) выполняется при $\theta = \operatorname{diam} X^0_\delta$, т.к. из определения выпуклости следует неравенство $f(x) - f^* \le \theta\|f'(x_k)\|$ $\forall x \in X^0_\delta$. Если функция $f(x)$ сильно выпукла, то $\theta(x) \le \|f'(x_k)\|/(4\kappa)$, где κ — константа сильной выпуклости, и множество X^0_δ ограничено при любом $x_0 \in R_n$. Значит, $\|f'(x_k)\| \le c'$ $\forall x \in X^0_\delta$ и в (23) можно положить $\theta = c'/(4\kappa)$.

Теорема 5. *Пусть функция $f(x)$ дважды дифференцируема, $f'(x)$ удовлетворяет в R_n условию Липшица с константой L, существует такое число $\mu > 0$, что*

$$\langle f''(x)\xi, \xi\rangle \ge \mu\|\xi\|^2 \quad \forall x, \xi \in R_n, \tag{30}$$

где $f''(x)$ — матрица вторых производных функции $f(x)$ в точке x. Пусть выполняется (24), точки y_k выбраны для всех $k \in K$ согласно (3) или (4), причем $\delta_k = 0$, $\nu_k = \nu = 1$ $\forall k \in K$. Тогда для последовательности $\{x_k\}$ справедливы оценки

$$0 \le a_k \le a_0 q^k, \tag{31}$$

$$\|x_k - x^*\|^2 \le (2/\mu) a_0 q^k \tag{32}$$

$\forall k \in K$, где $0 < q = 1 - \mu(1-\varepsilon)^2/(L(1+\varepsilon)^2) < 1$, x^* — точка минимума функции $f(x)$ в R_n.

Доказательство. Согласно (30) функция $f(x)$ сильно выпукла в R_n. Значит, X^* состоит из единственной точки, множество $X^0_\delta = X^0$ ограничено, и для функции $f(x)$ верно (23). При указанном способе задания чисел δ_k, ν_k условия (3), (4), а также оценки (28), (29) совпадают. Поэтому

$$a_k - a_{k+1} \geq ((1-\varepsilon)^2/2L(1+\varepsilon)^2)\|f'(x_k)\|^2 \quad \forall k \in K. \tag{33}$$

Известно ([5], с. 184), что $f(x) - f(x^*) \leq \langle f'(x), x - x^*\rangle - \kappa\|x - x^*\|^2$ $\forall x \in R_n$, где κ — константа сильной выпуклости функции $f(x)$. Используя это неравенство при $x = x_k$, как и в ([5], с. 267), нетрудно показать, что $f(x_k) - f(x^*) = a_k \leq \|f'(x_k)\|^2/(4\kappa)$, $k \in K$. Отсюда и из (33) следует оценка $a_k - a_{k+1} \geq (2\kappa(1-\varepsilon)^2/(L(1+\varepsilon)^2))a_k$, $k \in K$, а т. к. $2\kappa = \mu \leq L$ ([5], сс. 185, 186), то $0 \leq a_{k+1} \leq a_k(1-\mu(1-\varepsilon)^2/(L(1+\varepsilon)^2)) = qa_k$, $k \in K$, причем $0 < q < 1$. Значит, $a_k \leq qa_{k-1} \leq \cdots \leq q^k a_0$, $k \in K$, и неравенство (31) доказано. Наконец, поскольку ([5], с. 182) $f(x) - f(x^*) \geq \kappa\|x - x^*\|^2$ $\forall x \in R_n$, то $a_k \geq \kappa\|x_k - x^*\|^2$, $k \in K$, и отсюда с учетом (31) следует (32). $\square$

Нетрудно проверить, что утверждения теорем 1–5 остаются справедливыми, если в методе положить

$$g_k = \lambda^0_k f'(x_k) + r_k, \quad 0 < \lambda' \leq \lambda^0_k \leq \lambda'' < \infty,$$

неравенство (6) заменить неравенством

$$\langle f'(x_k), r_k\rangle \geq -\gamma^0_k\|f'(x_k)\|^2,$$

где $0 \leq \gamma^0_k \leq \theta\lambda^0_k$, $0 < \theta < 1$, а в (24) значение ε выбирать из условия $0 < \varepsilon < \lambda'$.

3. Реализации метода

За счет свободы в выборе векторов r_k, многообразий $M(x_k)$ и точек y_k предложенный метод допускает различные реализации. Приведем и обсудим ниже некоторые из них. Подчеркнем, что теоремы 1–3 сходимости общего метода справедливы при требованиях (2), (6), (10) на выбор векторов r_k, поэтому для обоснования сходимости алгоритмов приходится проверять каждый раз выполнение условий (2), (6), (10).

Пусть

$$H_k = \{1\}, \quad h^k_1 = g_k \quad \forall k \in K. \tag{34}$$

Тогда $M(x_k) = U(x_k)$, и y_k — точка приближенного минимума функции одной переменной $f(x_k + \alpha g_k)$. Если положить $r_k = 0$, $\varepsilon_k = \delta_k = 0$, $\nu_k = 1$, $x_{k+1} = y_k$ $\forall k \in K$, то при условиях (3), (4) алгоритм совпадает с методом наискорейшего спуска для минимизации выпуклых функций, причем векторы r_k удовлетворяют требованиям (2), (6), (10). Поскольку условия (2), (24) выполняются, то оценка (25) скорости сходимости метода наискорейшего спуска остается в силе для псевдовыпуклой функции $f(x)$.

Как уже замечено, вектор $r_k \neq 0$ можно интерпретировать как погрешность вычисления $f'(x_k)$ в обсуждаемом алгоритме ([3], сс. 97, 98). Если при условиях (34), (5) вектор $r_k \neq 0$ выбирается для всех $k \in K$ согласно (2), (6), (10), то метод наискорейшего спуска при наличии помех остается сходящимся. Если векторы r_k удовлетворяют условиям (2), (24), то оценки (25), (31), (32) скорости сходимости метода не ухудшаются по сравнению с оценками метода наискорейшего спуска с точным вычислением градиента ([5], с. 266), т. е. метод наискорейшего спуска устойчив к помехам в вычислении градиента.

Покажем, что метод покоординатного спуска также является частным случаем предложенного метода. Пусть вектор $r_k = (r^1_k, \ldots, r^n_k)$ в методе для всех $k \in K$ строится следующим образом. Отыскивается такой номер $i_k \in I = \{1, \ldots, n\}$, что

$$\left|\frac{\partial f(x_k)}{\partial \xi_{i_k}}\right| = \max_{i \in I}\left|\frac{\partial f(x_k)}{\partial \xi_i}\right|,$$

и полагается $r_k^i = -\frac{\partial f(x_k)}{\partial \xi_i}$ $\forall i \in I \setminus i_k$, $r_k^{i_k} = 0$. Тогда компоненты g_k^i, $i \in I$, вектора $g_k = f'(x_k) + r_k$ имеют вид $g_k^i = \frac{\partial f(x_k)}{\partial \xi_i}$, если $i = i_k$, и $g_k^i = 0$, если $i \in I \setminus i_k$. Следовательно, при условиях (34), (5) данный алгоритм действительно представляет собой известный вариант метода покоординатного спуска ([6], с. 206). Проверим выполнение условий (2), (6), (10) для векторов r_k в обсуждаемом алгоритме. Пусть $\left(\frac{\partial f(x_k)}{\partial \xi_i}\right)^2 = \theta_i$, $i \in I$. Поскольку $\langle f'(x_k), g_k\rangle = \|f'(x_k)\|^2 + \langle f'(x_k), r_k\rangle$, а с другой стороны $\langle f'(x_k), g_k\rangle = \theta_{i_k}$, то

$$\frac{\|f'(x_k)\|^2 + \langle f'(x_k), r_k\rangle}{\|f'(x_k)\|^2} = \frac{\theta_{i_k}}{\sum\limits_{i\in I}\theta_i} \geq \frac{\theta_{i_k}}{n\theta_{i_k}} = \frac{1}{n}.$$

Отсюда $\langle f'(x_k), r_k\rangle \geq -(1-1/n)\|f'(x_k)\|^2$, и вектор r_k удовлетворяет неравенству (6) при $\gamma = 1 - 1/n$. Кроме того,

$$\|r_k\| = \left(\sum_{i\in I\setminus i_k}\theta_i\right)^{1/2} < \left(\sum_{i\in I}\theta_i\right)^{1/2} = \|f'(x_k)\|,$$

а при условии ограниченности множеств X^0 или X^0_δ последовательность чисел $\varepsilon_k = \|f'(x_k)\|$ ограничена, и требования (2), (10) выполняются.

Пусть в предложенном методе для всех $k \in K$ полагается

$$H_k = \{1\}, \quad r_k = G_k f'(x_k) - f'(x_k), \quad h_1^k = -g_k = -G_k f'(x_k),$$

G_k — симметричная матрица,

$$\|G_k\| \leq c_1, \quad \langle G_k\xi, \xi\rangle \geq \widetilde{\mu}\|\xi\|^2, \;\; \widetilde{\mu} > 0 \;\; \forall \xi \in R_n, \tag{35}$$

точки y_k выбираются из условия (4) или (9) и выполняется (5). Нетрудно видеть, что $-g_k$ есть направление спуска для функции $f(x)$ в точке x_k. Значит, этот алгоритм можно записать в виде

$$x_{k+1} = x_k - \lambda_k G_k f'(x_k), \;\; k \in K, \tag{36}$$

где шаг $\lambda_k > 0$ выбирается так, чтобы точка $x_{k+1} = y_k$ удовлетворяла неравенству (4) или (9). Так как $x_k \in X^0$ $\forall k \in K$, то в случае ограниченности множества X^0 выполняется (22), а значит, и (2), (10), где $\varepsilon = c_0(c_1+1)$. Кроме того, согласно второму из неравенств (35) $\langle f'(x_k), r_k\rangle \geq (\widetilde{\mu} - 1)\|f'(x_k)\|^2$, т. е. справедливо (6) при $\gamma = 1 - \mu'$, где $0 < \mu' < min\{1, \widetilde{\mu}\}$. Таким образом, условия (2), (6), (10) выполнены. Если $\|G_k - E\| \leq \varepsilon < 1$ $\forall k \in K$, то справедливы требования (2), (24), и для (36) можно сформулировать утверждения, аналогичные теоремам 4, 5.

Алгоритм (36) известен для минимизации выпуклых функций как обобщенный метод градиентного спуска ([6], с. 189; [7], с. 61). Отметим, что его сходимость доказана в теореме 1 без традиционного условия Липшица на $f'(x)$. В схему (36) вписываются многие известные методы решения задачи (1). Кроме градиентного алгоритма ($G_k = E$), частным случаем (36) является, например, метод изменения масштабов ([6], с. 190), где в качестве G_k выбираются диагональные матрицы вторых производных функции $f(x)$. Если в (36) положить

$$G_k = (f''(x_k))^{-1} \quad \forall k \in K, \tag{37}$$

то получим метод Ньютона с регулировкой шага ([7], с. 68; [3], с. 66) для минимизации сильно выпуклых дважды непрерывно дифференцируемых функций, сходящийся при любом начальном приближении. Известно ([7], с. 27), что если $f(x)$ удовлетворяет условию (30) и $\langle f''(x)\xi, \xi\rangle \leq M\|\xi\|^2$ $\forall x, \xi \in R_n$, то для матриц G_k, определенных в (37), выполняется (35) при $\widetilde{\mu} = \mu/M^2$, $c_1 = 1/\widetilde{\mu}$. Большой класс квазиньютоновских методов (напр., [3], [8]) также вкладывается в схему (36). Матрицы G_k, используемые, например, в методе Давидона–Флетчера–Пауэлла удовлетворяют ([6], с. 191–193) условиям (35).

Опишем алгоритм, близкий к (36). Положим в методе

$$H_k = \{1, \ldots, l_k\}, \;\; l_k \geq 1, \;\; r_k = 0, \;\; h_i^k = G_k^i f'(x_k), \;\; i \in H_k, \;\; \forall k \in K,$$

где матрицы G_k^i таковы, что $\sum\limits_{i\in H_k} \gamma_k^i G_k^i = E$ для некоторых чисел γ_k^i. Так как $\sum\limits_{i\in H_k} \gamma_k^i h_i^k = f'(x_k) = g_k$ $\forall k \in K$, то требования к выбору векторов h_i^k в п. 3 метода выполнены. Условия (2), (6), (10) на выбор r_k также выполняются. Значит, алгоритм является сходящимся. Поскольку векторы r_k удовлетворяют и условиям (2), (24), то для алгоритма справедливы оценки, полученные в теоремах 4, 5. В алгоритме можно положить, например, $l_k = 2$, $G_k^1 = (f''(x_k))^{-1}$, $G_k^2 = E$.

В последнее время возобновились публикации по исследованию многошаговых методов минимизации (напр., [9]–[11]). Напомним, что метод называют N-шаговым, $N \geq 1$, если в нем каждая итерационная точка, начиная с x_N, строится на основе N предыдущих приближений. Примером двухшагового алгоритма безусловной минимизации выпуклых функций может служить известный вариант метода сопряженных градиентов ([3], с. 70):

$$
\begin{gathered}
x_{k+1} = x_k - \alpha_1^k f'(x_k) + \alpha_2^k (x_k - x_{k-1}), \quad x_{-1} = x_0, \\
\{\alpha_1^k, \alpha_2^k\} = \arg \min_{\{\alpha_1,\alpha_2\}} f(x_k - \alpha_1 f'(x_k) + \alpha_2(x_k - x_{k-1})) \quad \forall k \in K.
\end{gathered} \tag{38}
$$

Покажем, что алгоритм (38) является частным случаем предложенного метода. Пусть

$$
\begin{gathered}
H_k = \{1, 2\}, \ \ r_k = 0, \ \ \varepsilon_k = \varepsilon = 0, \ \ h_1^k = -f'(x_k), \ \ k \in K, \\
h_2^0 = 0, \ \ h_2^k = x_k - x_{k-1}, \ \ k = 1, 2, \dots
\end{gathered} \tag{39}
$$

Если при этом $x_{k+1} = y_k = \arg \min\limits_{x\in M(x_k)} f(x)$ $\forall k \in K$, то алгоритм (39) предложенного метода совпадает с (38). Поскольку требования (2), (6), (10), (24) на выбор r_k в (39) выполняются, то при условии ограниченности множества X^0 в теореме 1 доказана сходимость метода (38), а в теоремах 4, 5 получены оценки его скорости сходимости для псевдовыпуклой и сильно выпуклой функции. Заметим, что в (38) полагается $f(x_{k+1}) = m_k$. В теоремах 1–4 показано, что метод (38) остается сходящимся и при менее жестких условиях (см. (3), (4), (9)) выбора точек x_{k+1}, при этом оценка (25) его скорости сходимости сохраняется.

В ([5], с. 326; [3], с. 73) рекомендуется включать в метод сопряженных градиентов так называемую процедуру обновления, выбирая некоторое подмножество $K_0 \subset K$ и полагая в (38) $\alpha_2^k = 0$ $\forall k \in K_0$, т. е. время от времени делая шаг, как и в начальной точке, по антиградиенту. Например, предлагается задавать $K_0 = \{0, n, 2n, \dots\}$. В алгоритм (39) можно включать любые процедуры обновления, выбирая различные подмножества $K_0 \subset K$ и полагая $\alpha_2 = 0$ $\forall k \in K_0$ при задании многообразия $M(x_k)$. Утверждения теорем 1–5 для алгоритма (39), а значит, и для метода (38), со всеми процедурами обновления остаются в силе. Отметим, что если в алгоритме (39) выбирать $r_k \neq 0$, $h_1^k = -g_k$ $\forall k \in K$, то получится вариант метода сопряженных градиентов при наличии помех для псевдовыпуклых функций.

Предлагаемый метод дает возможность строить и другие двухшаговые алгоритмы. Опишем некоторые из них. Пусть выполняется (34),

$$
r_0 = 0, \ \ r_k = \gamma_k (x_k - x_{k-1}), \ \ k = 1, 2, \dots, \tag{40}
$$

$0 \leq |\gamma_k| \leq \overline{\gamma}$. Если при этом положить $x_{k+1} = y_k = \arg \min\limits_{x\in U(x_k)} f(x)$, $k \in K$, то алгоритм (34), (40) можно записать в виде

$$
x_{k+1} = x_k - \lambda_k g_k, \quad \lambda_k = \arg \min_{\lambda \geq 0} f(x_k - \lambda g_k), \quad k \in K, \tag{41}
$$

где $g_0 = f'(x_0)$, $g_k = f'(x_k) + \gamma_k (x_k - x_{k-1})$, $k = 1, 2, \dots$ Так как $\langle f'(x_k), x_k - x_{k-1}\rangle = 0$, $k = 1, 2, \dots,$ то независимо от знака числа γ_k

$$
\langle f'(x_k), g_k\rangle = \|f'(x_k)\|^2 + \gamma_k \langle f'(x_k), x_k - x_{k-1}\rangle > 0, \quad k = 1, 2, \dots, \tag{42}
$$

и вектор $-g_k$ для всех $k \in K$ — направление спуска для функции $f(x)$ в точке x_k. Следовательно, процесс (41) релаксационен, и $x_k \in X^0$ $\forall k \in K$. Если $\operatorname{diam} X^0 = d_0 < \infty$, то

$$
\|r_k\| = |\gamma_k| \, \|x_k - x_{k-1}\| \leq \overline{\gamma} d_0, \quad k = 1, 2, \dots, \tag{43}
$$

и условия (2), (10) выполняются. Кроме того, $\langle f'(x_k), r_k\rangle = 0$, $k \in K$, и для векторов r_k справедливо (6). Таким образом, алгоритм (34), (40) является сходящимся для псевдовыпуклой функции $f(x)$.

Покажем, что в алгоритме (41) не обязательно число λ_k выбирать из условия точного минимума. Действительно, пусть в (34), (40) для всех $k \in K$ точки x_{k+1} выбираются согласно неравенству

$$f(x_k - \lambda_k g_k) \le (1-\nu_k) f(x_k) + \nu_k \min_{\lambda} f(x_k - \lambda g_k),$$

и при этом

$$\gamma_{k+1}\langle f'(x_{k+1}), x_{k+1} - x_k\rangle \ge 0. \tag{44}$$

Тогда справедливо (42), а значит, и (43), и $\langle f'(x_k), r_k\rangle \ge 0$, $k \in K$, т. е. выполняется (6). Для обеспечения условия (44) шаг λ_k, больший или меньший полного, должен быть согласован со знаком числа γ_{k+1}. Поскольку в (40) допускается задание $\gamma_k = 0$, то тем самым этот алгоритм, как и (39), предусматривает включение всевозможных процедур обновления без ущерба оценкам скорости сходимости.

В (34), (40) можно использовать и другие условия выбора чисел γ_k, например,

$$0 \le |\gamma_k| \le \varepsilon \|f'(x_k)\| / \|x_k - x_{k-1}\|, \quad 0 < \varepsilon < 1, \quad k = 1, 2, \ldots$$

В таком случае для алгоритма (34), (40) выполняется (2), (24), а значит, справедливы утверждения теорем 4, 5. Оба предложенных способа выбора значений γ_k в алгоритме (34), (40) дают более широкие возможности для построения направлений спуска по сравнению со способом задания аналогичного параметра в методе сопряженных градиентов ([3], с. 73; [5], с. 325).

Предложенный метод позволяет строить также N-шаговые алгоритмы для $N \ge 2$. Например, алгоритм (39) можно обобщить, положив

$$H_k = \{1, \ldots, N\}, \quad r_k = 0, \quad \varepsilon_k = 0, \quad x_{k+1} = y_k, \quad h_1^k = -f'(x_k), \quad k \in K,$$
$$h_i^k = 0, \quad k = 0, \ldots, N-2, \quad h_i^k = x_{k-i+2} - x_{k-i+1}, \quad k \ge N-1, \quad i = 2, \ldots, N.$$

Можно получить N-шаговый метод и из алгоритма (34), (40), изменив способ задания векторов r_k. Пусть в (34) r_k выбираются следующим образом:

$$r_k = 0, \quad k = 0, \ldots, N-2; \quad r_k = \sum_{i=2}^{N} \gamma_k^i (x_{k-i+2} - x_{k-i+1}), \quad k \ge N-1, \tag{45}$$

а числа γ_k^i таковы, что $-\overline{\gamma} \le \gamma_k^i \le \overline{\gamma}$, $\overline{\gamma} > 0$, $\langle f'(x_k), r_k\rangle \ge 0$. Тогда алгоритм (34), (45) остается релаксационным, аналогично (43) получается оценка $\|r_k\| \le \overline{\gamma} d_0 (N-1)$, и условия (2), (10), (6) выполняются. Коэффициенты γ_k^i в алгоритме можно выбирать, адаптируя их к процессу минимизации. Пусть

$$0 \le |\gamma_k^i| \le \psi_i \|f'(x_k)\| / \|x_{k-i+2} - x_{k-i+1}\|, \quad i = 2, \ldots, N, \quad k \ge N-1,$$

а значения ψ_i подбираются так, чтобы векторы r_k $\forall k \in K$ удовлетворяли условиям (2), (24). Например, можно положить $\psi_i = \varepsilon_i / N^s$, где $\varepsilon_i \ge 0$, $0 < \sum\limits_{i=2}^{N} \varepsilon_i \le N^{s-1}$, $s \ge 0$, или $\psi_i = 1/(N(N-1))$. Нетрудно видеть, что в обоих случаях условия (2), (24) выполняются при $\varepsilon = 1/N$. Следовательно, утверждения теорем 4, 5 верны. Отметим, что в обоих случаях задания ψ_i увеличение значения N влечет уменьшение величины ε в (24), а значит, и констант c и q в оценках (25), (31), (32) скорости сходимости обсуждаемого алгоритма.

Как и в (34), (40), можно произвольно задавать моменты $k \in K_0 \subset K$ обновления алгоритма (34), (45), полагая $\gamma_k^i = 0$ для всех или некоторых номеров $2 \le i \le N$.

В заключение коротко обсудим возможности предложенного в статье общего подхода к построению алгоритмов безусловной минимизации гладких псевдовыпуклых функций. Он позволяет получать новые алгоритмы, а также модифицировать известные методы выпуклого программирования, расширяя их возможности в выборе направлений итерационного перехода и шаговых множителей. Описанная схема решения задачи (1) позволяет, в частности, строить различные многошаговые алгоритмы с процедурами обновления, по единой методике исследовать их сходимость, получать оценки скорости сходимости для псевдовыпуклых и сильно выпуклых функций. Поскольку на шаге 5 метода процедура отыскания приближения x_{k+1} не конкретизирована, то за счет чередования различных способов построения точек y_k и x_{k+1} можно получать всевозможные смешанные методы на основе известных и новых алгоритмов без дополнительных исследований их сходимости. Считая векторы r_k погрешностью в вычислении градиента, можно исследовать влияние помех на сходимость и оценки скорости сходимости алгоритмов.

Литература

1. Mangasarian O.L. *Pseudo-convex functions* // J. Soc. industr. and appl. math. Ser. A. –1965. – V. 3. – P. 281–290.
2. Заботин Я.И., Кораблев А.И. *Псевдовыпуклые функционалы и их экстремальные свойства* // Изв. вузов. Математика. – 1974. – № 4. – С. 27–31.
3. Поляк Б.Т. *Введение в оптимизацию.* – М.: Наука, 1983. – 384 с.
4. Бахвалов Н.С. *Численные методы. Анализ, алгебра, обыкновенные дифференциальные уравнения.* Учеб. пособие. – М.: Наука, 1973. – 631 с.
5. Васильев Ф.П. *Численные методы решения экстремальных задач.* – М.: Наука, 1988. – 552 с.
6. Карманов В.Г. *Математическое программирование.* Учеб. пособие. – 3-е изд. – М.: Наука, 1986. – 288 с.
7. Пшеничный Б.Н., Данилин Ю.М. *Численные методы в экстремальных задачах.* – М.: Наука, 1975. – 320 с.
8. Сухарев А.Г., Тимохов А.В., Федоров В.В. *Курс методов оптимизации.* – М.: Наука, 1986. – 326 с.
9. Антипин А.С. *Непрерывные и итеративные процессы с операторами проектирования и типа проектирования* // Вопр. кибернетики. Вычисл. вопр. анал. больших систем. – М., 1989. – С. 5–43.
10. Родионов А.В. *О методе тяжелого шарика* // Вопр. кибернетики. Мод. и мет. анал. больших систем. – М., 1991. – С. 96–104.
11. Хотеев С.В. *О многошаговых градиентных методах в задачах оптимизации* // Вопр. кибернетики. Мод. и мет. анал. больших систем. – М., 1991. – С. 104–111.

Казанский государственный университет

Поступила 26.05.1998