УДК 519.3

ГРАДИЕНТНЫЙ МЕТОД МИНИМИЗАЦИИ БЕЗ ВОЗВРАЩЕНИЙ НА ПОВЕРХНОСТЬ ОГРАНИЧЕНИЙ

В.А.Аберман

Среди методов минимизации функций в задачах с ограничениями типа равенств можно выделить группу градиентно-восстановительных методов, в которых движение происходит по поверхности ограничений в направлении убывания минимизируемой функции. В первом методе этой группы – методе Розена[I] – направление движения из точки на поверхности ограничений определяется проекцией антиградиента минимизируемой функции на касательную плоскость к поверхности ограничений в этой точке. Разработан сопряженно-градиентный восстановительный метод [2]. При построении градиентно-восстановительных методов выбор направления движения можно связывать также с сужением минимизируемой функции на касательную плоскость к допустимому многообразию в текущей точке [3]. Градиентно-восстановительные методы обычно дают глобальную сходимость к локальному минимуму. Однако при их разностной реализации приходится проводить на каждой итерации весьма трудоемкий восстановительный процесс – возвращение на поверхность ограничений.

В настоящей статье разработана модификация градиентно-восстановительных методов, при которой разностная реализация не требует проведения восстановительного процесса. В предлагаемых методах движение происходит, вообще говоря, по недопустимым точкам и описывается системой обыкновенных дифференциаль-

ных уравнений. Правая часть этой системы сочетает в себе слагаемое, связанное с направлением движения соответствующего градиентно-восстановительного метода, и слагаемое, связанное с направлением, приближающим текущую точку траектории к поверхности ограничений. Анализ этой схемы проведен для системы дифференциальных уравнений, в которой первое слагаемое связано с направлением движения в методе Розена, а второе - с направлением спуска на поверхность ограничений методом Ньютона.

I. Основные обозначения

Вектор $u \in R^l$ рассматривается как столбец. Скалярное произведение двух векторов $u \in R^l$ и $v \in R^l$ обозначается через (u, v), а евклидова норма вектора $u \in R^l$ - через $\|u\|$. Через A^T обозначается матрица, транспонированная к матрице A.

Для функции $f: R^n \longrightarrow R$ через $\nabla f(x)$ обозначается ее градиент в точке x. Мы будем рассматривать $\nabla f(x)$ как вектор-столбец, так что транспонированный к нему вектор есть $\frac{df(x)}{dx}$, т.е. $(\nabla f(x), y) = \frac{df(x)}{dx} \cdot y$ для любого $y \in R^n$. Через $F(x)$ обозначается матрица вторых производных функции f в точке x.

Для функции $g = (g_1, \ldots, g_m): R^n \longrightarrow R^m$ через $\nabla g(x)$ обозначается матрица Якоби системы функций $(g_1, \ldots, g_m)$ в точке x, а через $G_i(x)$ - матрицы вторых производных функции g_i в точке x. Если матрица Якоби $\nabla g(x)$ имеет ранг m, то через $L(x)$ обозначим матрицу $F(x) + \sum_{i=1}^{m} \lambda_i(x) G_i(x)$, где строка $(\lambda_1(x), \ldots, \lambda_m(x)) = \nabla f(x)^T \nabla g(x)^T [\nabla g(x) \cdot \nabla g(x)^T]^{-1}$. В этом же случае через $P(x)$ обозначим матрицу $\nabla g(x)^T [\nabla g(x) \nabla g(x)^T]^{-1} \nabla g(x)$, а через $\Pi(x)$ - матрицу $I - P(x)$, где I - единичная матрица порядка n.

2. Постановка задачи и предварительные замечания

Рассматривается задача

$$\min\{f(x): g(x) = 0\},$$

где функции $f: R^n \longrightarrow R$ и $g: R^n \longrightarrow R^m$ принадлежат классу C^2. Множество $\{x \mid g(x) = \mathbb{O}\}$ допустимых точек обозначим через K.

Точку x назовем регулярной точкой, если ранг матрицы $\nabla g(x)$ равен m, и обозначим через S множество всех регулярных точек. Для любой регулярной точки x

$$\nabla g(x)\pi(x) = \mathbb{O}. \tag{2}$$

Действительно, $\nabla g(x)\pi(x) = \nabla g(x)(I - \nabla g(x)^T[\nabla g(x)\nabla g(x)^T]^{-1}\nabla g(x)) =$
$= \nabla g(x) - \nabla g(x)\nabla g(x)^T[\nabla g(x)\nabla g(x)^T]^{-1}\nabla g(x) = \nabla g(x) - \nabla g(x) = \mathbb{O}.$

Если регулярная точка x^* является локальным экстремумом функции f на множестве K, то точка x^* является решением системы уравнений

$$\pi(x)\nabla f(x) = \mathbb{O}, \quad g(x) = \mathbb{O}, \tag{3}$$

и имеет место равенство

$$\frac{d}{dx}\pi(x)\nabla f(x)\Big|_{x^*} = \pi(x^*)L(x^*). \tag{4}$$

Назовем локальный регулярный максимум x^* функции f на множестве K квадратичным, если существует $\alpha_1 > 0$ такое, что для любого $z \in R^n$ выполняется неравенство

$$(\pi(x^*)z)^T L(x^*)(\pi(x^*)z) \leq -\alpha_1 \|\pi(x^*)z\|^2. \tag{5}$$

Назовем локальный регулярный минимум x^* задачи (I) квадратичным, если существует $\alpha_2 > 0$ такое, что для любого $z \in R^n$ выполняется неравенство

$$(\pi(x^*)z)^T L(x^*)(\pi(x^*)z) \geq \alpha_2 \|\pi(x^*)z\|^2. \tag{6}$$

3. Анализ системы дифференциальных уравнений

В множестве S рассмотрим систему дифференциальных уравнений

$$\dot{x} = -\pi(x)\nabla f(x) - \delta\nabla g(x)^T[\nabla g(x)\nabla g(x)^T]^{-1}g(x), \tag{7}$$

где δ - некоторая положительная постоянная. Будем предполагать, что каждая рассматриваемая в дальнейшем траектория лежит в S вместе с некоторой своей ε-окрестностью (ε-окрест-

ностью множества A называется множество $B=\bigcup_{x\in A}\{y: \|x-y\|<\varepsilon\}$. Для краткости в дальнейшем будем полагать $\varphi(x)=\nabla g(x)^T \times [\nabla g(x)\nabla g(x)^T]^{-1} g(x)$ и траекторию системы (7) называть "траектория $x(t)$ ".

ЛЕММА I. Для траектории $x(t)$ существуют такие константы C и $\lambda>0$, что $\|g(x(t))\| \leqslant C e^{-\lambda t}$ (т.е. $\|g(x(t))\| \to 0$ при $t \to \infty$ с экспоненциальной скоростью).

ДОКАЗАТЕЛЬСТВО. Для траектории $x(t)$ в силу соотношения (2) имеем $\frac{d}{dt}\left\{\frac{1}{2} g(x(t))^T g(x(t))\right\} = g(x(t))^T \nabla g(x(t)) \dot{x}(t) =$

$$= -g(x(t))^T \nabla g(x(t))(-\pi(x(t))\nabla f(x(t)) - \delta\varphi(x(t))) =$$

$$= -\delta g(x(t))^T \nabla g(x(t)) \nabla g(x(t))^T [\nabla g(x(t))\nabla g(x(t))^T]^{-1} g(x(t)) =$$

$$= -\delta g(x(t))^T g(x(t)),$$

т.е. $\frac{d}{dt}\left(\frac{1}{2}\|g(x(t))\|^2\right) = -\delta \|g(x(t))\|^2$, откуда $\|g(x(t))\| = \|g(x(0))\| e^{-\delta t}$. Лемма I доказана.

Множество назовем кусочно-гладкосвязным, если для любых двух его точек существует кусочно-гладкий путь, соединяющий эти точки и лежащий в множестве.

В теореме I и ее следствиях будем предполагать, что множество решений системы уравнений (3) состоит из изолированных кусочно-гладкосвязных множеств.

ТЕОРЕМА I. Пусть траектория $x(t)$ ограничена. Тогда все ее ω -предельные точки удовлетворяют системе уравнений (3).

ДОКАЗАТЕЛЬСТВО. Пусть x^* - ω -предельная точка траектории $x(t)$. Тогда, в силу леммы I, $g(x^*)=0$. Докажем, что $\pi(x^*)\nabla f(x^*)=0$.

Предположим противное, т.е. что $\|\pi(x^*)\nabla f(x^*)\|=\beta$, где $\beta>0$. Так как $\pi(x)\nabla f(x)$ и $\varphi(x)$ непрерывно зависят от x в множестве S , то существует замкнутый шар $U(x^*)$ радиуса $\rho>0$ с центром в точке x^* такой, что для всех $x\in U(x^*)$ выполняются неравенства $\|\pi(x)\nabla f(x) + \delta\varphi(x) - \pi(x^*)\nabla f(x^*)\| \leqslant \beta/2$ и $\frac{\beta}{2} \leqslant \|\pi(x)\nabla f(x)\| \leqslant \frac{3\beta}{2}$.

Рассмотрим также замкнутый шар $V(x^*)$ радиуса $\frac{\rho}{2}$ с центром в точке x^*. Заметим, что не существует $\bar{t}$ такого, что $x(t) \in U(x^*)$ при $t > \bar{t}$. В противном случае все ω - предельные точки траектории $x(t)$ принадлежали бы шару $U(x^*)$. С другой стороны, вдоль траектории $x^*(t)$, выходящей из точки x^*, $g(x^*(t)) = 0$ в силу леммы I, и поэтому $\frac{d}{dt} f(x^*(t)) = -\|\pi(x^*(t)) \nabla f(x^*(t))\|^2$, т.е. траектория $x^*(t)$ не может целиком лежать в $U(x^*)$, так как иначе было бы $f(x^*(t)) \leqslant -\frac{\beta}{2} t + f(x^*)$. Но все точки траектории $x^*(t)$ являются ω -предельными для траектории $x(t)$ [5]. Это противоречие показывает, что существуют моменты времени τ_i, i =1,2,... ($\lim\limits_{i \to \infty} \tau_i = \infty$), при которых

$$\|x(\tau_i) - x^*\| = \rho.$$

Поскольку x^* - ω -предельная точка траектории $x(t)$, то найдутся моменты времени σ_i, i =1,2,... ($\lim\limits_{i \to \infty} \sigma_i = \infty$), при которых $\|x(\sigma_i) - x^*\| = \frac{\rho}{2}$. Очевидно, можно считать, что $\sigma_i > \tau_i > \sigma_{i-1}$ и что при всех $t \in (\tau_i, \sigma_i)$ выполняются неравенства $\frac{\rho}{2} \leqslant \|x(t) - x^*\| \leqslant \rho$. Но $x(\sigma_i) - x(\tau_i) = (\sigma_i - \tau_i)\pi(x^*)\nabla f(x^*) + \int_{\tau_i}^{\sigma_i} (\dot{x}(t) - \pi(x^*)\nabla f(x^*))dt$. Поэтому $\frac{\beta}{2}|\sigma_i - \tau_i| \leqslant \|x(\sigma_i) - x(\tau_i)\| \leqslant \frac{3\beta}{2}|\sigma_i - \tau_i|$. Так как $\frac{\rho}{2} \leqslant \|x(\sigma_i) - x(\tau_i)\| \leqslant \frac{3\rho}{2}$, то

$$\frac{\rho}{3\beta} \leqslant \sigma_i - \tau_i \leqslant \frac{3\rho}{\beta}. \tag{8}$$

Далее, найдется $\bar{\bar{t}} \in (\tau_i, \sigma_i)$ такое, что

$$f(x(\sigma_i)) = f(x(\tau_i)) + \frac{df(x(\bar{\bar{t}}))}{dt}(\sigma_i - \tau_i) = f(x(\tau_i)) -$$
$$-\|\pi(x(\bar{\bar{t}}))\nabla f(x(\bar{\bar{t}}))\|^2(\sigma_i - \tau_i) - \delta \nabla f(x(\bar{\bar{t}}))^T \varphi(x(\bar{\bar{t}}))(\sigma_i - \tau_i) \leqslant$$
$$\leqslant f(x(\tau_i)) - \|\pi(x(\bar{\bar{t}}))\nabla f(x(\bar{\bar{t}}))\|^2 \cdot |\sigma_i - \tau_i| +$$
$$+ \delta\|\nabla f(x(\bar{\bar{t}}))\| \cdot \|\varphi(x(\bar{\bar{t}}))\| \cdot |\sigma_i - \tau_i|.$$

Так как $\|g(x(t))\| \to 0$ при $t \to \infty$ и $U(x^*)$ - компакт, то существует t_0 такое, что при $t > t_0$, $t \in [\tau_i, \sigma_i]$, ока-

зывается $\|\varphi(x(t))\| \leq \dfrac{\beta^2}{72M_1\delta}$, где $M_1 = max\{\|\nabla f(x)\|:$ $x \in U(x^*)\}$. Поэтому при $t \geq t_0$ в силу неравенства (8) имеем $f(x(\sigma_i)) \leq f(x(\tau_i)) - \frac{\beta^2}{4}\cdot\frac{\rho}{3\beta} + \delta M \frac{\beta^2}{72M_1\delta}\cdot\frac{3\rho}{\beta} =$ $= f(x(\tau_i)) - \frac{\rho\cdot\beta}{24}$. Таким образом, показано, что за время τ_i, σ_i, где $\tau_i > t_0$, функция f убывает не менее чем на $\frac{\rho\cdot\beta}{24}$.

Выберем ε так, чтобы траектория $x(t)$ входила в множество S вместе со своей ε-окрестностью. Ввиду ограниченности траектории $x(t)$, замыкание ее $\varepsilon/2$-окрестности, обозначаемое в дальнейшем через $\tilde{S}$, ограничено и, ввиду изолированности компонент связности множества решений системы (3), имеет непустое пересечение лишь с конечным числом этих компонент связности. Пусть $\{Y_1, Y_2, \ldots, Y_k\}$ - совокупность таких непустых пересечений. Тогда множество $\bigcup\limits_{i=1}^{k} Y_i$ состоит из всех точек множества $\tilde{S}\cap K$, для которых $\pi(x)\nabla f(x) = 0$. Поскольку $\pi(x)\nabla f(x)$ и $g(x)$ непрерывно зависят от x в множестве S, то каждое множество Y_i замкнуто. Анализ кусочно-гладкого пути, соединяющего две точки множества Y_i и лежащего в множестве решений системы (3), показывает, что функция f постоянна на каждом множестве Y_i.

Так как множества Y_i замкнуты и функция f равномерно-непрерывна в любой их ограниченной окрестности, то существуют такие γ_i -окрестности U_i множеств Y_i, что $U_i \cap U(x^*) = \emptyset$ и $|f(x) - f(y)| < \rho\beta/96k$ при $x \in U_i$ и $y \in Y_i$. Пусть $U = \bigcup\limits_{i=1}^{k} U_i$. Так как множество U открыто, то его дополнение U' замкнуто. Поэтому множество $P = \tilde{S}\cap U'\cap K$ замкнуто и ограничено. Функция $\psi: x \longrightarrow \|\pi(x)\nabla f(x)\|$ непрерывна на множестве P, и так как в множестве P нет точек x, в которых $\psi(x) = 0$, то $\min\limits_{x\in P}\psi(x) = 2\alpha > 0$. Так как функция ψ непрерывна в замыкании $\frac{\varepsilon}{3}$ -окрестности множества $\tilde{S}$, то существует такое $\gamma_0 \in (0, \frac{\varepsilon}{3})$, что для любого x из γ_0 -окрестности A множества P выполняется неравенство $\psi(x) = \|\pi(x)\nabla f(x)\| > \alpha$.

Положим $B = \tilde{S} \cap K$. По построению, множество B содержится в множестве $A \cup U$ вместе со своей τ-окрестностью, где $\tau = min\{\tau_0, \tau_1, \dots, \tau_k\}$. Но в силу леммы I все ω-предельные точки траектории $x(t)$ лежат в B, и поэтому $x(t) \in A \cup U$, начиная с некоторого момента времени T.

Так как $\|\varphi(x(t))\| \longrightarrow 0$ при $t \longrightarrow \infty$ и A - ограниченное множество, то существует такое t_1, что при $t \geqslant t_1$ и $x(t) \in A$ оказывается $\|\varphi(x(t))\| \leqslant \frac{\alpha^2}{2\delta M_2}$, где $M_2 = sup\{\|\nabla f(x)\| : x \in A\}$. Для таких t имеем $\frac{df(x(t))}{dt} =$

$$= - \|\pi(x(t))\nabla f(x(t))\|^2 - \delta\nabla f(x(t))\varphi(x(t)) \leqslant$$

$$\leqslant - \|\pi(x(t))\nabla f(x(t))\|^2 + \delta\|\nabla f(x(t))\| \cdot \|\varphi(x(t))\| \leqslant -\alpha^2 - \delta M_2 \frac{\alpha^2}{2\delta M_2} = -\frac{\alpha^2}{2},$$

т.е. $f(x(t))$ при $t \geqslant t_1$ и $x(t) \in A$ убывает со скоростью не меньшей чем $\frac{\alpha^2}{2}$. С другой стороны, вариация функции f на множестве U меньше $\frac{\rho\beta}{48}$. Поэтому если для некоторого $\bar{t} \geqslant max\{t_0, t_1, T\}$ оказалось, что $x(\bar{t}) \in U$, то, взяв $\tau_i > \bar{t}$, мы найдем, что $x(t) \in U$ при $t \geqslant \tau_i$, так как на промежутке времени $[\tau_i, \sigma_i]$ функция f убыла по крайней мере на $\frac{\rho\beta}{24}$, в множестве A она убывает, а за время пребывания в U не может возрасти более чем на $\frac{\rho\beta}{48}$. Таким образом, начиная с некоторого момента времени, траектория $x(t)$ лежит в A, и, следовательно, функция f вдоль траектории $x(t)$ убывает со скоростью не меньшей чем $\frac{\alpha^2}{2}$, что противоречит конечности функции f в точке x^*. Полученное противоречие показывает, что во всех ω-предельных точках траектории $x(t)$ выполняется соотношение $\pi(x^*)\nabla f(x^*) = \Theta$. Теорема доказана.

СЛЕДСТВИЕ I. Если каждая компонента связности множества решений системы (3) состоит из одной точки, то для ограниченной траектории $x(t)$ существует $\lim\limits_{t \to \infty} x(t) = x^*$, где x^* удовлетворяет системе (3).

ДОКАЗАТЕЛЬСТВО. Так как ω - предельное множество траектории $x(t)$ связно [5] и является в силу теоремы I подмно-

жеством множества решений системы (3), то оно состоит из одной точки. Следствие доказано.

В следствиях 2 и 3 теоремы I будут рассмотрены случаи, когда ограниченность траектории $x(t)$ обеспечивается видом функций f и g.

СЛЕДСТВИЕ 2. Пусть существует такое $a>0$, что множество $C=\{x : \|g(x)\| \leqslant a\}$ ограничено. Тогда ω-предельное множество траектории $x(t)$ непусто и состоит из точек, удовлетворяющих системе (3).

ДОКАЗАТЕЛЬСТВО. В силу леммы I, начиная с некоторого момента времени T, траектория $x(t)$ лежит в множестве C и, следовательно, ограничена. Поэтому имеет место результат теоремы I. Следствие доказано.

СЛЕДСТВИЕ 3. Пусть существуют такие постоянные $L>0$, $a>0$, $l>0$, $l_0>0$ и $\beta>0$, что на множестве $\mathscr{D}=\{x : \|x\| \geqslant L,\ \|g(x)\| \leqslant a,\ x \in S\}$ функция f ограничена снизу и выполняются неравенства $\|\pi(x)\nabla f(x)\| \geqslant l\|\nabla f(x)\|$, $\|\nabla f(x)\| \geqslant l_0$ и $\nabla g(x)^T \nabla g(x)^{T-1} \leqslant \beta$. Тогда ω-предельное множество траектории $x(t)$ непусто и состоит из точек, удовлетворяющих системе (3).

ДОКАЗАТЕЛЬСТВО. Сделанные предположения, согласно лемме I, обеспечивают равенства $\lim\limits_{t\to\infty}\|g(x(t))\|=0$ и $\lim\limits_{t\to\infty}\|\varphi(x(t))\|=0$. Поэтому существует такое T, что $\|g(x(t))\| \leqslant a$ и $\|\varphi(x(t))\| \leqslant \leqslant \frac{l^2 l_0}{2\delta}$ при $t>T$. При этих t для точек $x(t)$ с $\|x(t)\| \geqslant L$ имеем $\frac{df(x(t))}{dt} = -\|\pi(x(t))\nabla f(x(t))\|^2 - \delta \nabla f(x(t))^T$ и

$$\varphi(x(t)) \leqslant -\|\pi(x(t))\nabla f(x(t))\|^2 + \frac{\delta}{l}\|\pi(x(t))\nabla f(x(t))\| \frac{l^2 l_0}{2\delta} \leqslant$$

$\leqslant -\frac{l l_0}{2}\|\pi(x(t))\nabla f(x(t))\|$. Для этих же точек $\left\|\frac{dx(t)}{dt}\right\| \leqslant$ $\leqslant (1+\frac{l}{2})\|\pi(x(t))\nabla f(x(t))\|$, и, следовательно,

$$\frac{df(x(t))}{dt} \leqslant -l_1\left\|\frac{dx(t)}{dt}\right\|$$

где $l_1 > 0$.

Будем рассматривать траекторию $x(t)$ лишь для $t \geqslant T$ и покажем, что она ограничена. Пусть $L_1 \geqslant L$ и в шаре $\{x : \|x\| \leqslant L_1\}$ имеются точки траектории $x(t)$. Положим $N = \inf\{f(x) : x \in \mathcal{D}\}$, $N_1 = \max\{f(x) : \|x\| = L_1\}$. Пусть $\|x(t_1)\| = L_1$ и $\|x(t)\| > L_1$ на промежутке (t_1, t_2). Тогда согласно (9) имеем $l_1 \|x(t_2) - x(t_1)\| \leqslant f(x(t_1)) - f(x(t_2)) \leqslant N_1 - N$. Поэтому $\|x(t_2)\| \leqslant L_1 + \frac{N_1 - N}{l} = L_2$, и, следовательно, вся траектория $x(t)$ лежит в шаре $\{\|x\| \leqslant L_2\}$. Доказываемое утверждение теперь вытекает из теоремы I.

ЗАМЕЧАНИЕ. Теорема I показывает, что все ω-предельные точки траектории $x(t)$ удовлетворяют необходимым условиям экстремума. Вообще говоря, при выборе начальной точки движения специальным образом могут возникнуть случаи, когда траектория $x(t)$ стремится к локальному максимуму. Однако для большинства траекторий $x(t)$ эта точка окажется неустойчивой. Покажем, например, что локальный регулярный и квадратичный максимум x^* функции f на множестве K окажется неустойчивым для всех таких траекторий $x(t)$, что $x(0) \in K$, кроме одной тривиальной траектории - стационарной точки x^*. Действительно, для таких траекторий, в силу леммы I, $g(x(t)) = 0$ при любом t, и, следовательно, система дифференциальных уравнений (7) для них примет вид $\dot{x} = -\pi(x) \nabla f(x)$. Пусть $\lim\limits_{t \to \infty} x(t) = x^*$. Обозначим через $\hat{x}$ разность $x(t) - x^*$. Тогда, используя соотношения (4) и (6) и разлагая $\pi(x) \nabla f(x)$ по формуле Тейлора в окрестности точки x^*, получим

$$\frac{d}{dt}\left(\frac{1}{2} \hat{x}^T \hat{x}\right) = \hat{x}^T \dot{x}(t) = \hat{x}^T(-\pi(x(t)) \nabla f(x(t))) = \hat{x}^T(-\pi(x^*) \nabla f(x^*) -$$

$$- \frac{d}{dx} \pi(x) \nabla f(x)\Big|_{x^*} \cdot \hat{x} + o(\|\hat{x}\|) + o(\|\hat{x}\|^2) = -(\pi(x^*)\hat{x})^T L(x^*) \times$$

$$\times \pi(x^*)\hat{x} + o(\|\hat{x}\|^2) \geqslant \alpha_1 \|\pi(x^*)\hat{x}\|^2 - |o(\|\hat{x}\|^2)| = \alpha_1 \|\hat{x}\|^2 - |o(\|\hat{x}\|^2)|.$$

Таким образом, в некоторой окрестности точки x^* имеет место неравенство $\frac{d}{dt}\left(\frac{1}{2}\|\hat{x}\|^2\right) \geqslant \frac{\alpha_1}{2} \|\hat{x}\|^2$, что и доказывает утверждение.

Доказанные теорема I и ее следствия не определяют характера сходимости траектории $x(t)$ к ее ω-предельному множеству. Траектория $x(t)$ стремится к своему ω-предельному множеству, вообще говоря, немонотонно. Однако в окрестности локального минимума x^* задачи (I) при некоторых условиях на δ и точку x^* траектория $x(t)$ сходится к точке x^* с экспоненциальной скоростью. Докажем это.

ТЕОРЕМА 2. Пусть x^* - локальный регулярный и квадратичный минимум задачи (I) и в системе дифференциальных уравнений (7) $\delta > \frac{\|L(x^*)\|^2}{4\alpha_2}$ (α_2 взято из соотношения (6) для точки x^*). Тогда траектория $x(t)$ локально сходится к точке x^* с экспоненциальной скоростью.

ДОКАЗАТЕЛЬСТВО. Система дифференциальных уравнений (7) в результате разложения функции, стоящей в ее правой части, по формуле Тейлора в окрестности точки x^* примет вид:

$$\dot{x} = -\pi(x^*)\nabla f(x^*) - \delta\varphi(x^*) - \frac{d}{dx}(\pi(x)\nabla f(x) + \delta\varphi(x))\Big|_{x^*}(x-x^*) + O(\|x-x^*\|).$$

Так как точка x^* удовлетворяет системе (3) и выполняется соотношение (4), получим

$$\dot{x} = -\pi(x^*)L(x^*)(x-x^*) - \delta P(x^*)(x-x^*) + O(\|x-x^*\|). \tag{10}$$

Обозначим через $\hat{x}$ разность $x(t)-x^*$. Тогда, используя соотношение (10) в окрестности точки x^*, получим

$$\frac{d}{dt}\left(\frac{1}{2}\hat{x}^T\hat{x}\right) = -\hat{x}^T\dot{x}(t) = -\hat{x}^T\pi(x^*)L(x^*)\hat{x} - \delta\hat{x}^T P(x^*)\hat{x} + O(\|\hat{x}\|^2) =$$

$$= -\hat{x}^T\pi(x^*)L(x^*)\pi(x^*)\hat{x} - \hat{x}^T\pi(x^*)L(x^*)P(x^*)\hat{x} -$$

$$- \delta(P(x^*)\hat{x})^T(P(x^*)\hat{x}) + O(\|\hat{x}\|^2).$$

Так как по условию теоремы в системе дифференциальных уравнений (7) $\delta > \frac{\|L(x^*)\|^2}{4\alpha_2}$, то существует такое $\hat{t} > 0$, что $\delta \geq \frac{\|L(x^*)\|^2}{4(\alpha_2-\hat{\varepsilon})} + \hat{\varepsilon}$. Поэтому, используя неравенство (6), получим

$$\frac{d}{dt}\left(\frac{1}{2}\hat{x}^T\hat{x}\right) \leqslant -\alpha_2 \|\pi(x^*)\hat{x}\|^2 - \delta\|P(x^*)\hat{x}\|^2 + \|L(x^*)\| \times$$

$$\times \|\pi(x^*)\hat{x}\| \cdot \|P(x^*)\hat{x}\| + o(\|\hat{x}\|^2) \leqslant -\alpha_2\|\pi(x^*)\hat{x}\|^2 - \left(\frac{\|L(x^*)\|^2}{4(\alpha_2-\delta)} + \hat{\varepsilon}\right) \times$$

$$\times \|P(x^*)\hat{x}\|^2 + (\alpha_2 - \hat{\varepsilon})\|\pi(x^*)\hat{x}\|^2 + \frac{\|L(x^*)\|^2}{4(\alpha_2-\hat{\varepsilon})} \cdot \|P(x^*)\hat{x}\|^2 + o(\|\hat{x}\|^2) =$$

$$= -\hat{\varepsilon}\|\pi(x^*)\hat{x}\|^2 - \hat{\varepsilon}\|P(x^*)\hat{x}\|^2 + o(\|\hat{x}\|^2) = -\hat{\varepsilon}\|\hat{x}\|^2 + o(\|\hat{x}\|^2).$$

Таким образом, в некоторой окрестности $\hat{U}(x^*)$ точки x^* для траектории $x(t)$ имеет место неравенство $\frac{d}{dt}\left(\frac{1}{2}\|\hat{x}\|^2\right) \leqslant -\frac{\hat{\varepsilon}}{2}\|\hat{x}\|^2$. Если $\hat{t}$ выбрать так, что $x(\hat{t}) \in \hat{U}(x^*)$, то при $t \geqslant \hat{t}$ получим соотношение $\|\hat{x}\|^2 \leqslant |x(\hat{t}) - x^*|^2 e^{-\hat{\varepsilon}(t-\hat{t})}$. Теорема доказана.

СЛЕДСТВИЕ. Пусть в задаче (1) ограничения линейные, а функционал f квадратичный, т.е. $g(x) = Bx + b$, $f(x) = \frac{1}{2}x^TAx + (c,x) + d$. Тогда если матрица B имеет ранг m и существует такое $\alpha_2 > 0$, что на подпространстве $\{x : Bx = 0\}$ выполняется неравенство $(x, Ax) \geqslant \alpha_2(x,x)$, то при $\delta > \frac{\|A\|^2}{4\alpha_2}$ траектория $x(t)$ сходится к решению задачи (I) с экспоненциальной скоростью.

ДОКАЗАТЕЛЬСТВО. Система (3) в рассматриваемом случае имеет единственное решение x^*, которое является решением задачи (I) в силу положительной определенности матрицы $L(x^*) = A$ на подпространстве $\{x : Bx = 0\}$. Система дифференциальных уравнений (7) в результате разложения функции, стоящей в ее правой части, по формуле Тейлора принимает вид (IO) с $o(\|x - x^*\|) = 0$, откуда и вытекает глобальная сходимость траектории $x(t)$ к точке x^* с экспоненциальной скоростью. Следствие доказано.

В заключение отметим, что систему дифференциальных уравнений предложенного типа можно строить на основании градиентно-восстановительных методов, рассмотренных в работах [2] и [3], причем второе слагаемое может быть связано с любым спуском на поверхность ограничений.

Литература

I. Y.Rosen. The gradient projection method for nonlinear programming, part 2, J.Soc. Indust. and Appl. Math., 9(1961), p.514-532.

2. A.Mille, H.Y.Huang, C.Heideman. Sequental Gradient-Restoration algorithm for the minimization of constrained function-ordinary and conjugate gradient version, J. of Optimiration Th. and Appl., 4, 4, 1969, p.213-243.

3. ДАНИЛИН Ю.М. О минимизации функций в задачах с ограничениями типа равенств. - В кн.: Кибернетика, 2, I97I,с.88-95.

4. D.S.Luenberger. The gradient projection method along geodesics. Manag Sci, 18, 11, 1972, p. 620-631.

5. РЕЙССИНГ Р., САНСОНЕ Г, КОНТИ Р. Качественная теория нелинейных дифференциальных уравнений. М., "Наука", I974, с. 49-60.

Поступила в ред.-изд. отд.
20. IУ. I975 г.