Доклады Академии наук СССР
1970. Том 191, № 2

УДК 518:517.948

МАТЕМАТИКА

Ю. И. ЛЮБИЧ

ЗОНА ПРИТЯЖЕНИЯ СЕДЛА В ПРОЦЕССЕ НАИСКОРЕЙШЕГО СПУСКА

(Представлено академиком Л. К. Канторовичем 8 VII 1969)

В предыдущей заметке [1] было установлено, что процесс наискорейшего спуска на плоскости может сходиться к седловой точке лишь в исключительных ситуациях. Для n-мерного пространства R^n вопрос остается открытым. Удалось получить лишь некоторые частичные результаты. Они излагаются ниже.

Пусть $\varphi(x)$ — функция класса C^2 в области $G\subset R^n$; $x=0$ — невырожденная седловая точка; p — ее индекс Морса; $\lambda_1,\ldots,\lambda_p$ — отрицательные собственные значения гессиана $H_0=H(0)$; $\{e_j\}_1{}^n$ — соответствующая система главных осей. Положим

$$\rho(x)=\nabla\varphi(x)-H_0x,\qquad \rho_j(x)=\xi_j[\rho(x)],$$

где $\{\xi_j\}_1{}^n$ — координаты в базисе $\{e_j\}_1{}^n$.

Зоной притяжения седла $x=0$ называется множество тех начальных приближений, для которых процесс наискорейшего спуска сходится к этому седлу.

Теорема 1. *Все точки x зоны притяжения седла удовлетворяют системе уравнений*

$$\xi_j=\sum_{k=0}^{\infty}\frac{\gamma(\Gamma^kx)\,\rho_j(\Gamma^kx)}{\prod\limits_{s=0}^{k}(1+\gamma(\Gamma^sx)\,|\lambda_j|)}\qquad (j=1,\ldots,p). \tag{1}$$

Здесь, как и в [1], $\gamma(z)$ — коэффициент спуска; Γ — оператор спуска: $\Gamma z=z-\gamma(z)\nabla\varphi(z)$.

Доказательство. Положим для краткости

$$\gamma_k=\gamma(\Gamma^kx),\quad \rho_j^k=\rho_j(\Gamma^kx),\quad \xi_j^k=\xi_j(\Gamma^kx).$$

Тогда

$$\xi_j^{k+1}=(1+\gamma_k|\lambda_j|)\,\xi_j^k-\gamma_k\rho_j^k\quad (j=1,\ldots,p).$$

Полагая

$$\eta_j^k=\xi_j^k\prod_{s=0}^{k-1}(1+\gamma_s|\lambda_j|)^{-1},$$

получаем

$$\eta_j^{m+1}=\eta_j^0-\sum_{k=0}^{m}\frac{\gamma_k\rho_j^k}{\prod\limits_{s=0}^{k-1}(1+\gamma_s|\lambda_j|)}.$$

Остается заметить, что, поскольку $\lim\limits_{m\to\infty}\xi_j{}^m=0$, то и подавно $\lim\limits_{m\to\infty}\eta_j{}^m=0$, а с другой стороны, $\eta_j{}^0=\xi_j{}^0=\xi_j$.

Замечание. Неясно, все ли точки, удовлетворяющие уравнениям (1), принадлежат зоне притяжения седла. Во всяком случае приведенное выше доказательство не обратимо.

С «физической» точки зрения теорема 1 означает, что зона притяжения седла есть $(n-p)$-мерное многообразие. В действительности это не удается доказать даже локально (т. е. вблизи седла). Основная трудность заключена в обосновании почленной дифференцируемости рядов (1) (дифференцируемость отдельных членов этих рядов может быть доказана). Мы установим лишь, что зона притяжения седла образует с плоскостью $\xi_j = 0$ ($j = 1, \ldots, p$) асимптотически нулевой угол.

Предположим, что существует такая окрестность U нуля, что

$$\varphi(0) < \varliminf_{x \to \partial U} \varphi(x),$$

и что нуль не связан с остальными стационарными точками, лежащими в U (см. [1]). Обозначим через x^- проекцию вектора x на подпространство, отвечающее собственным значениям $\lambda_1, \ldots, \lambda_p$.

Теорема 2. *При $x \to 0$ вдоль зоны притяжения*

$$\|x^-\| = o(\|x\|).$$

При доказательстве этой теоремы используется теорема 1, лемма 1 [1] и следующая

Лемма. *При достаточно малом $\eta > 0$ существует окрестность нуля V_η, в которой либо $\varphi(\Gamma x) < \varphi(0)$, либо $\varphi(x) \geqslant \eta\|x\|^2$.*

Отметим, что исследование рядов (1) облегчается благодаря элементарному тождеству

$$\sum_{k=0}^{\infty} \frac{\gamma_k}{\prod\limits_{s=0}^{k} (1 + \gamma_s |\lambda|)} = \frac{1}{|\lambda|}.$$

Это тождество заведомо справедливо, если $\varliminf_{k \to \infty} \gamma_k > 0$.

Но в действительности

$$\varliminf_{z \to 0} \gamma(z) \geqslant \| H_0 \|^{-1}.$$

Рассмотрим строение зоны притяжения в целом. Положим

$$\lambda(x, y) = \frac{(\nabla\varphi(x), x - y)}{\| x - y \|^2}, \quad R(x, y) = H(x) - \lambda(x, y) I.$$

Обозначим $\mathfrak{M}$ множество пар $\{x, y\}$, для которых $(x - y) \wedge \nabla\varphi(x) = 0$, $(x - y, \nabla\varphi(y)) = 0$ ($\wedge$ — символ внешнего произведения).

Теорема 3. *Пусть на пересечении множества $\mathfrak{M}$ с областью $\lambda(x, y) > 0$ выполнены условия*

$$\det R(x, y) \neq 0, \quad (R^{-1}(x, y)\nabla\varphi(x), \nabla\varphi(y)) \neq 0 \quad (\nabla\varphi(y) \neq 0) \quad (2)$$

Если в окрестности седла его зона притяжения принадлежит гладкому q-мерному многообразию, то в целом она принадлежит счетному объединению гладких q-мерных многообразий.

В случае гиперболического седла (см. [1]) условную теорему 3 можно заменить безусловным утверждением:

Теорема 4. *При выполнении условий* (2) зона притяжения гиперболического седла принадлежит счетной системе гладких кривых.*

Это соответствует тому, что для невырожденного гиперболического седла $p = n - 1$. Однако в теореме 4 седло может быть и вырожденным любой степени.

Отметим, что теоремы 3, 4 не зависят от теорем 1, 2. Теорему 4 можно рассматривать как n-мерный аналог теоремы 6 [1]. При этом, однако, остается невыясненным, выполняются ли условия (2) в «общем положении», т. е. для плотного множества функций φ.

Харьковский государственный университет
им. А. М. Горького

Поступило
27 VI 1969

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Ю. И. Любич, ДАН, 179, № 5 (1968).

* По-прежнему на пересечении $\mathfrak{M}$ с областью $\lambda(x, y) > 0$.