THE DOMAIN OF ATTRACTION OF A SADDLE IN THE PROCESS OF STEEPEST DESCENT

1970

SovietRxiv

Abstract

Full Text

UDC 518:517.948

Yu. I. LYUBICH

THE DOMAIN OF ATTRACTION OF A SADDLE IN THE PROCESS OF STEEPEST DESCENT

(Presented by Academician L. V. Kantorovich, July 8, 1969)

In the preceding note (^{1}) it was established that the process of steepest descent in the plane can converge to a saddle point only in exceptional situations. For an n-dimensional space R^n the question remains open. It has been possible to obtain only some partial results. They are set out below.

Let $\varphi(x)$ be a function of class C^2 in a domain $G \subset R^n$; $x = 0$ a nondegenerate saddle point; p its Morse index; $\lambda_1, \dots, \lambda_p$ the negative eigenvalues of the Hessian $H_0 = H(0)$; $\{e_j\}_1^n$ the corresponding system of principal axes. Put

$$\rho(x) = \nabla\varphi(x) - H_0 x, \qquad \rho_j(x) = \xi_j[\rho(x)],$$

where $\{\xi_j\}_1^n$ are the coordinates in the basis $\{e_j\}_1^n$.

The **domain of attraction** of the saddle $x = 0$ is the set of those initial approximations for which the process of steepest descent converges to this saddle.

Theorem 1. All points x of the domain of attraction of the saddle satisfy the system of equations

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

Here, as in (^{1}), $\gamma(z)$ is the descent coefficient; Γ is the descent operator:

$$\Gamma z = z - \gamma(z)\nabla\varphi(z).$$

Proof. For brevity put

$$\gamma_k = \gamma(\Gamma^{kx}), \qquad \rho_j^k = \rho_j(\Gamma^{kx}), \qquad \xi_j^k = \xi_j(\Gamma^{kx}).$$

Then

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

Putting

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

we obtain

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

It remains to note that, since $\lim_{m\to\infty} \xi_j^m = 0$, a fortiori $\lim_{m\to\infty} \eta_j^m = 0$, while, on the other hand, $\eta_j^0 = \xi_j^0 = \xi_j$.

Remark. It is not clear whether all points satisfying equations (1) belong to the domain of attraction of the saddle. In any case, the proof given above is not reversible.

From the "physical" point of view, Theorem 1 means that the domain of attraction of the saddle is an $(n - p)$-dimensional manifold. In reality this cannot be proved even locally (i.e., near the saddle). The main difficulty

is contained in the justification of termwise differentiability of the series (1) (differentiability of the individual terms of these series can be proved). We shall establish only that the domain of attraction of the saddle forms an asymptotically zero angle with the plane $\xi_j = 0$ $(j = 1, \dots, p)$.

Suppose that there exists a neighborhood U of zero such that

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

and that zero is not connected with the remaining stationary points lying in U (see (^{1})). Denote by x^- the projection of the vector x onto the subspace corresponding to the eigenvalues $\lambda_1, \dots, \lambda_p$.

Theorem 2. *As $x \to 0$ along the domain of attraction,*

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

In the proof of this theorem, Theorem 1, Lemma 1 (^{1}), and the following are used.

Lemma. *For sufficiently small $\eta > 0$ there exists a neighborhood of zero V_η in which either $\varphi(\Gamma x) < \varphi(0)$, or $\varphi(x) \geq \eta\|x\|^2$.*

Let us note that the study of the series (1) is facilitated by the elementary identity

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

This identity is certainly valid if $\lim_{k\to\infty}\gamma_k > 0$.

But in fact

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

Let us consider the structure of the domain of attraction as a whole. Put

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

Denote by $\mathfrak{M}$ the set of pairs $\{x,y\}$ for which $(x-y)\wedge\nabla\varphi(x) = 0$, $(x-y, \nabla\varphi(y)) = 0$ ($\wedge$ is the symbol of the exterior product).

Theorem 3. *Suppose that on the intersection of the set $\mathfrak{M}$ with the domain $\lambda(x,y) > 0$ the conditions*

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

are fulfilled. If, in a neighborhood of the saddle, its domain of attraction belongs to a smooth q-dimensional manifold, then as a whole it belongs to a countable union of smooth q-dimensional manifolds.

In the case of a hyperbolic saddle (see (^{1})), the conditional Theorem 3 can be replaced by an unconditional assertion:

Theorem 4. *Under the fulfillment of the conditions* (2) *the domain of attraction of a hyperbolic saddle belongs to a countable system of smooth curves.*

This corresponds to the fact that, for a nondegenerate hyperbolic saddle, $p = n-1$. However, in Theorem 4 the saddle may also be degenerate to any degree.

Let us note that Theorems 3 and 4 do not depend on Theorems 1 and 2. Theorem 4 may be regarded as an n-dimensional analogue of Theorem 6 (^{1}). Here, however, it remains unclear whether the conditions (2) are fulfilled in "general position," i.e., for a dense set of functions φ.

Kharkov State University
named after A. M. Gorky

Received
27 VI 1969

REFERENCES

[1] Yu. I. Lyubich, DAN, 179, No. 5 (1968).

---

* As before, on the intersection of $\mathfrak{M}$ with the domain $\lambda(x, y) > 0$.

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.