ISSN 1064–5624, Doklady Mathematics, 2007, Vol. 76, No. 2, pp. 790–793. © Pleiades Publishing, Ltd., 2007.
Original Russian Text © K.V. Rudakov, A.A. Cherepnin, Yu.V. Chekhovich, 2007, published in Doklady Akademii Nauk, 2007, Vol. 416, No. 4, pp. 457–460.

COMPUTER SCIENCE

On Metric Properties of Spaces in Classification Problems

Corresponding Member of the Russian Academy of Sciences K. V. Rudakov, A. A. Cherepnin, and Yu. V. Chekhovich

Received May 3, 2007

DOI: 10.1134/S1064562407050377

This work is performed in the framework of the algebraic approach to the problem of synthesizing correct classification algorithms [1–6]. We use the basic definitions and notions from these papers without references. We introduce notions of the solvability and regularity radii of problems and the correctness radii of algorithms. We establish relationships between these notions, which, in particular, can be used to analyze the quality of initial data in applied classification problems.

Consider a set $\mathfrak{Z} = \{Z\}$ of classification problems. It is assumed that all problems from $\mathfrak{Z}$ are subject to a system of universal constraints. Thus, each problem Z from $\mathfrak{Z}$ is completely determined by a pair of matrices $(\hat{I}, \hat{\hat{I}})$ (an information matrix and an informational matrix); moreover, $\hat{I} \in \Im^*$ and $\hat{\hat{I}} \in \Im^*$, i.e., $Z = Z(\hat{I}, \hat{\hat{I}})$.

We also assume that the problem set $\mathfrak{Z}$ is endowed with a metric r and the set $\mathfrak{Z}$ is compact. We denote the set of all regular problems from $\mathfrak{Z}$ by $\mathfrak{Z}_R$ and the set of all solvable problems by $\mathfrak{Z}_S$.

Throughout the paper, we require that the following assumption hold.

General assumption. For any informational matrix $\hat{\hat{I}} \in \tilde{\Im}^*$, there exists an information matrix $\hat{I} \in \Im^*$ such that the problem $Z(\hat{I}, \hat{\hat{I}})$ is unsolvable.

Note that, if the universal constraints for a problem set $\mathfrak{Z}$ are expressed in terms of the category Φ_0, then the general assumption reduces to the requirement that the space $\tilde{\Im}^*$ contains no matrices consisting of pairwise equal elements; obviously, the consideration of such matrices is of no practical interest.

Dorodnicyn Computing Center, Russian Academy of Sciences, ul. Vavilova 40, GSP-1, Moscow, 119991 Russia

Definition 1. Let $Z_0 \in \mathfrak{Z}$. The regularity radius of the problem Z_0 is defined as

$$R(Z_0) = \inf_{z \in \mathfrak{Z} - \mathfrak{Z}_R} r(Z_0, Z),$$

and the solvability radius of the problem Z_0 is defined as

$$S(Z_0) = \inf_{z \in \mathfrak{Z} - \mathfrak{Z}_S} r(Z_0, Z).$$

Note that the inclusion $\mathfrak{Z}_R \subseteq \mathfrak{Z}_S$ directly implies $R(Z_0) \le S(Z_0)$.

In what follows, we assume that $\mathfrak{Z}_R$ and $\mathfrak{Z}_S$ are open sets in the topology determined by the metric r; thus, in Definition 1, we take minima over the closed sets $\mathfrak{Z} - \mathfrak{Z}_R$ and $\mathfrak{Z} - \mathfrak{Z}_S$. This assumption seems to be fairly natural. For example, if the system of universal constraints is described in terms of the category Φ_0, then

$$\mathfrak{Z}_R = \{(\hat{I}, \hat{\hat{I}}) \big| \ \hat{I} = \|I_{ij}\|_{q \times l} \in \mathfrak{L}_{ql}(\Im),$$
$$\hat{\hat{I}} = \|\tilde{I}_{ij}\|_{q \times l} \in \mathfrak{L}_{ql}(\tilde{\Im}), \ \underset{S}{\forall} (i_1, j_1) \ne (i_1, j_1) \colon I_{i_1 j_1} \ne I_{i_2 j_2} \},$$

$$\mathfrak{Z}_S = \{(\hat{I}, \hat{\hat{I}}) \big| \ \hat{I} = \|I_{ij}\|_{q \times l} \in \mathfrak{L}_{ql}(\Im),$$
$$\hat{\hat{I}} = \|\tilde{I}_{ij}\|_{q \times l} \in \mathfrak{L}_{ql}(\tilde{\Im}),$$
$$\underset{S}{\forall} (i_1, j_1) \ne (i_1, j_1) \colon (\tilde{I}_{ij} \ne \tilde{I}_{i_2 j_2}) \ne (\tilde{I}_{i_1 j_1} \ne \tilde{I}_{i_2 j_2}) \}.$$

Note also that if a problem Z_0 is unsolvable, then $R(Z_0) = S(Z_0) = 0$, and if a problem Z_0 is solvable but not regular, then $R(Z_0) = 0$.

We refer to problems Z_0 for which $S(Z_0) > R(Z_0)$ as strictly regular and to problems Z_0 for which $R(Z_0) = S(Z_0) > 0$ as nonstrictly regular. It is easy to see that, for nonstrictly regular problems, there is at least one unsolvable problem among the irregular problems nearest to them.

Suppose that

$$\forall_{\mathfrak{I}^* \times \tilde{\mathfrak{I}}^*} (\hat{I}_1, \hat{\hat{I}}_1), (\hat{I}_2, \hat{\hat{I}}_2): \hat{\hat{I}}_1 \neq \hat{\hat{I}}_2: \; r((\hat{I}_1, \hat{\hat{I}}_1), (\hat{I}_2, \hat{\hat{I}}_2)) > r((\hat{I}_1, \hat{\hat{I}}_1), (\hat{I}_2, \hat{\hat{I}}_1)); \quad (1)$$

this condition seems to be fairly natural. Then, the following assertion is valid.

Statement 1. *Let $Z_0(\hat{I}_0, \hat{\hat{I}}_0)$ be a nonstrictly regular problem, and let $Z_1(\hat{I}_1, \hat{\hat{I}}_1)$ be an unsolvable problem nearest to Z_0. Then, $\hat{\hat{I}}_1 = \hat{\hat{I}}_0$.*

Proof. The problem Z_1 is irregular. Since the regularity of a problem is determined only by its information matrix, it follows that the problem $Z_1'(\hat{I}_1, \hat{\hat{I}}_0)$ is irregular as well. If $\hat{\hat{I}}_1 \neq \hat{\hat{I}}_0$, then condition (1) implies $r(Z_0, Z_1') < r(Z_0, Z_1)$. Since $r(Z_0, Z_1') \geq R(Z_0)$, we have $r(Z_0, Z_1) > R(Z_0) = S(Z_0)$. Thus, the problem Z_1 cannot be an unsolvable problem nearest to Z_0.

Note that, for strictly regular problems, Statement 1 is not generally true. Suppose that a problem $Z_0(\hat{I}_0, \hat{\hat{I}}_0)$ is strictly regular and $Z_1(\hat{I}_1, \hat{\hat{I}}_1)$ is an unsolvable problem nearest to Z_0. Clearly, the problem $Z_1'(\hat{I}_1, \hat{\hat{I}}_0)$ is irregular. Under condition (1), we have $r(Z_0, Z_1') \leq r(Z_0, Z_1)$. Obviously, if $\hat{\hat{I}}_1 = \hat{\hat{I}}_0$, then the problems Z_0 and Z_1' coincide; thus, $r(Z_0, Z_1') = r(Z_0, Z_1)$. If $\hat{\hat{I}}_1 \neq \hat{\hat{I}}_0$, then $r(Z_0, Z_1') < r(Z_0, Z_1)$. The question of what conditions ensure the equality $\hat{\hat{I}}_1 = \hat{\hat{I}}_0$ for the unsolvable problems Z_1 nearest to Z_0 is of obvious practical interest. It is intuitively clear that this condition holds if the contribution to the metric r of the difference between informational matrices substantially exceeds that of the difference between information matrices. Rigorous results are as follows.

Definition 2. The diameter of the compact set of information matrices is defined as

$$D = \max_{(Z_1, Z_2) \in \mathfrak{Z}^2} r(Z_1, Z_2),$$

where the maximum is over pairs of problems with any common informational matrix.

Definition 3. The quantum of variation for informational matrices is defined as

$$\Delta = \max_{(Z_1, Z_2) \in \mathfrak{Z}^2} r(Z_1, Z_2),$$

where the minimum is over all pairs of problems with any common information matrix and different informational matrices.

Statement 2. *If $\Delta > 2D$, a problem $Z_0(\hat{I}_0, \hat{\hat{I}}_0)$ is strictly regular, and $Z_1(\hat{I}_1, \hat{\hat{I}}_1)$ is an unsolvable problem nearest to Z_0, then $\hat{\hat{I}}_1 = \hat{\hat{I}}_0$.*

Proof. Let $\hat{\hat{I}}_1 \neq \hat{\hat{I}}_0$. Suppose that $Z_2(\hat{I}_2, \hat{\hat{I}}_0)$ is the unsolvable problem that exists by the general assumption. Consider the problem $Z_3(\hat{I}_1, \hat{\hat{I}}_0)$. It follows from Definition 2 that $r(Z_0, Z_2) \leq D$ and $r(Z_3, Z_0) \leq D$. By Definition 3, $r(Z_3, Z_1) \geq \Delta$. Thus, $r(Z_1, Z_0) \geq r(Z_1, Z_3) - r(Z_0, Z_3) \geq \Delta - D$. Moreover, $r(Z_1, Z_0) \geq \Delta - D > D \geq r(Z_0, Z_2)$, which contradicts the assumption that Z_1 is an unsolvable problem nearest to Z_0.

Definition 4. A metric r on a problem space $\mathfrak{Z}$ is said to be additive if there exist metrics ρ_1 and ρ_2 on the spaces $\mathfrak{I}^*$ and $\tilde{\mathfrak{I}}^*$, respectively, such that $r(Z_1, Z_2) = \rho_1(\hat{I}_1, \hat{I}_2) + \rho_2(\hat{I}_1, \hat{I}_2)$ for any problems $Z_1(\hat{I}_1, \hat{\hat{I}}_1)$ and $Z_2(\hat{I}_2, \hat{\hat{I}}_2)$ from $\mathfrak{Z}$.

Statement 3. *If the metric r is additive, $\Delta \geq D$, a problem $Z_0(\hat{I}_0, \hat{\hat{I}}_0)$ is strictly regular, and $Z_1(\hat{I}_1, \hat{\hat{I}}_1)$ is an unsolvable problem nearest to Z_0, then $\hat{\hat{I}}_1 = \hat{\hat{I}}_0$.*

Proof. Suppose that $\hat{\hat{I}}_1 \neq \hat{\hat{I}}_0$ and let $Z_2(\hat{I}_2, \hat{\hat{I}}_0)$ be the unsolvable problem that exists by the general assumption. We have $r(Z_0, Z_2) = \rho_1(\hat{I}_0, \hat{I}_2) \leq D$. Moreover, $r(Z_0, Z_1) = \rho_1(\hat{I}_0, \hat{I}_1) + \rho_2(\hat{\hat{I}}_0, \hat{\hat{I}}_1)$, and by Definition 3, $\rho_2(\hat{\hat{I}}_0, \hat{\hat{I}}_1) \geq \Delta$. Since the problem Z_1 is unsolvable and, hence, irregular, while the problem $Z_3(\hat{I}_0, \hat{\hat{I}}_1)$ is regular, it follows that $\hat{I}_1 \neq \hat{I}_0$. Therefore, $\rho_1(\hat{I}_0, \hat{I}_1) > 0$. Thus, $r(Z_0, Z_1) > \Delta \geq D \geq r(Z_0, Z_2)$, and Z_1 cannot be an unsolvable problem nearest to Z_0.

The notions of the diameter of a compact set of information matrices and the quantum of variation of informational matrices can be localized for particular problems.

Definition 5. The depth of a problem $Z_0(\hat{I}_0, \hat{\hat{I}}_0)$ with respect to its information matrix is defined as

$$D(Z_0) = \max_{Z \in \mathfrak{Z}} r(Z_0, Z),$$

where the maximum is over all problems $Z(\hat{I}, \hat{\hat{I}})$ for which $\hat{\hat{I}} = \hat{\hat{I}}_0$.

Note that the depth of any problem Z does not exceed the diameter of the compact set of information matrices.

Definition 6. The quantum of variation for the informational matrix of a problem $Z_0(\hat{I}_0, \hat{\hat{I}}_0)$ is defined as

$$\Delta(Z_0) = \max_{Z \in \mathfrak{Z}} r(Z_0, Z),$$

where the minimum is over all problems $Z(\hat{I}, \hat{\hat{I}})$ such that $\hat{\hat{I}} \neq \hat{\hat{I}}_0$.

Note that the quantum of variation for the informational matrix of any problem Z is no less than the quantum of variation of informational matrices.

Statement 4. *Suppose that a problem $Z_0(\hat{I}_0, \hat{\hat{I}}_0)$ is strictly regular and $Z_1(\hat{I}_1, \hat{\hat{I}}_1)$ is an unsolvable problem nearest to Z_0. If $\Delta(Z_0) > D(Z_0)$, then $\hat{\hat{I}}_1 = \hat{\hat{I}}_0$.*

Proof. Suppose that $\hat{\hat{I}} \neq \hat{\hat{I}}_0$ and $Z_2(\hat{I}_2, \hat{\hat{I}}_2)$ is an unsolvable problem with $\hat{\hat{I}}_2 = \hat{\hat{I}}_0$, which exists by the general assumption. It follows from Definition 5 that $r(Z_2, Z_0) \le D(Z_0)$. Definition 6 and the assumption $\hat{\hat{I}} \neq \hat{\hat{I}}_0$ imply $r(Z_1, Z_0) \ge \Delta(Z_0)$. Since $\Delta(Z_0) > D(Z_0)$, we have $r(Z_1, Z_0) > r(Z_2, Z_0)$, which contradicts the assumption that Z_1 is an unsolvable problem nearest to Z_0.

For strictly regular problems, there exist nearest irregular solvable problems whose neighborhoods contain unsolvable problems.

Definition 7. The depth of a problem $Z_0(\hat{I}_0, \hat{\hat{I}}_0)$ with respect to the informational matrix is defined as

$$\tilde{D}(Z_0) = \max_{Z \in \mathfrak{Z}} r(Z_0, Z),$$

where the maximum is over all problems $Z(\hat{I}, \hat{\hat{I}})$ with $\hat{I} = \hat{I}_0$.

Statement 5. *If $Z_0(\hat{I}_0, \hat{\hat{I}}_0)$ is an irregular solvable problem, then the $\tilde{D}(Z_0)$-neighborhood of Z_0 contains an unsolvable problem.*

Proof. The regularity of a problem in the sense of Zhuravlev [2] is completely determined by the information matrix of this problem; a problem is said to be regular if and only if all problems with the same information matrix are solvable. By assumption, $Z_0(\hat{I}_0, \hat{\hat{I}}_0)$ is an irregular problem. Hence, there exists an unsolvable problem $Z_1(\hat{I}_1, \hat{\hat{I}}_1)$ for which $\hat{I}_1 = \hat{I}_0$. It follows from Definition 7 that $r(Z_0, Z_1) \le \tilde{D}(Z_0)$, as required.

Apparently, for the first time, metric properties of spaces in classification problems were studied in [1] in relation to the stability of classification algorithms and their models. Below, we establish a relationship between stability in the sense of Zhuravlev and the notion of radii introduced in this paper.

Definition 8. Suppose that $Z_0(\hat{I}_0, \hat{\hat{I}}_0)$ is a problem and A is a classification algorithm implementing a mapping $A: \mathfrak{I}^* \to \tilde{\mathfrak{I}}^*$. The stability radius of the algorithm A at the point Z_0 of the space $\mathfrak{Z}$ is defined as

$$R_S(A, Z_0) = \min_{Z \in \mathfrak{Z}} r(Z_0, Z),$$

where $Z(\hat{I}, \hat{\hat{I}})$ is any problem for which $A(\hat{I}) \neq \hat{\hat{I}}$.

Definition 9. Suppose that $Z_0(\hat{I}_0, \hat{\hat{I}}_0)$ is a problem and A is a classification algorithm implementing a mapping $A: \mathfrak{I}^* \to \tilde{\mathfrak{I}}^*$. The correctness radius of the algorithm A at the point Z_0 of the space $\mathfrak{Z}$ is defined as

$$R_C(A, Z_0) = \min_{Z \in \mathfrak{Z}} r(Z_0, Z),$$

where $Z(\hat{I}, \hat{\hat{I}})$ is any problem for which $A(\hat{I}) \neq A(\hat{\hat{I}})$.

Obviously, if an algorithm A is incorrect for some problem Z_0, then $R_C(A, Z_0) = 0$.

Note that, generally, any relations between the stability and correctness radii are possible.

It directly follows from the definition of the quantum of variation of informational matrices Δ and the stability radius $R_S(A, Z_0)$ that, if an algorithm A is correct for a problem Z_0 and $R_S(A, Z_0) < \Delta$, then the algorithm A is correct for all problems from the $R_S(A, Z_0)$-neighborhood of the problem Z_0. On the other hand, the algorithm A is surely incorrect for any problem Z at which the minimum in Definition 8 is attained; this, in this case, we have $R_C(A, Z_0) = R_S(A, Z_0)$.

Obviously, the relation $R_C(A, Z_0) < S(Z_0)$ always holds. If $\hat{\hat{I}}_1 \neq \hat{\hat{I}}_0$ for a problem $Z_1(\hat{I}_1, \hat{\hat{I}}_1)$ nearest to $Z_0(\hat{I}_0, \hat{\hat{I}}_0)$, then $R_S(A, Z_0) < S(Z_0)$. Indeed, if $R_S(A, Z_0) \ge S(Z_0)$, then $A(\hat{I}_1) = \hat{\hat{I}}_0 = \hat{\hat{I}}_1$, i.e., A is correct for the problem Z_1, which contradicts the assumption that this problem is unsolvable.

Thus, in this paper, we have introduced and studied the notions of the solvability and regularity radii of classification problems. We demonstrated the special role of the solvability radius; calculating it during an analysis of data in applied problems, we can reveal situations in which the problem under consideration is solvable (objects from different classes are discernible) at the expense of, e.g., a inadequately exact measurement of features.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 07-01-00711 and 05-01-00718) and the program for support of young candidates of sciences (project no. 5266.2007.9).

REFERENCES

1. Yu. I. Zhuravlev, Kibernetika, No. 2, 35–43 (1978).
2. Yu. I. Zhuravlev, Probl. Kibern., No. 33, 5–68 (1978).
3. K. V. Rudakov, Zh. Vychisl. Mat. Mat. Fiz. **26**, 1719–1729 (1988).
4. K. V. Rudakov, Kibernetika, No. 1, 1–5 (1988).
5. K. V. Rudakov, Dokl. Akad. Nauk SSSR **297**, 43–46 (1987).
6. K. V. Rudakov, in *Recognition, Classification, Prediction* (Nauka, Moscow, 1989), pp. 176–201 [in Russian].