Canad. Math. Bull. Vol. **15** (2), 1972

CONSISTENCY IN THE RECONSTRUCTION OF PATTERNS FROM SAMPLE DATA

BY
MARC MOORE

1. **Introduction and summary.** Let A be a k-dimensional Euclidean region having unit volume. An m-colors pattern is a partition P_A of A into m sets A_i, $i=1,\ldots,m$ with positive volume. P_A is called a random pattern if in addition the partition of A is a realization of a random process with the following stationarity and isotropy properties:

(i) for all points $s\in A$, $P(s\in A_i)=p_i$, $i=1,\ldots,m$

(ii) for all pair of points s, $s'\in A$ with distance $|s-s'|=d$ between them, $P(s'\in A_i \mid s\in A_j)=P_{ij}(d)$, $i,j=1,\ldots,m$.

In [1], Switzer defines a reconstruction rule which produces an estimated reconstruction of the pattern using information obtained from n fixed sample-points $s_1\ldots s_n$ of A. To evaluate the accuracy of a reconstruction rule δ, applied to n sample-points, he proposes the loss function:

$$L_n = 1-\mu[{}_nP_A^\delta \cap P_A]$$

where ${}_nP_A^\delta$ is the partition of A obtained from the reconstruction rule δ applied to the n sample-points; ${}_nP_A^\delta\cap P_A$ is the set of points which have the same color in ${}_nP_A^\delta$ and P_A; μ is the Lebesgue measure.

In this note we define a notion of consistency for a reconstruction rule when the number of fixed sample-points increase and we give sufficient conditions under which the simple nearest-point rule [1] is consistent.

2. Consistency for a Reconstruction rule

Definition. A rule for the reconstruction of a random pattern is consistent if the corresponding sequence of random variables $\langle L_n\rangle$ converges in probability to zero.

Theorem. *A reconstruction rule δ is consistent if and only if*

$$\lim_{n\to\infty} E\mu[{}_nP_A^\delta\cap P_A] = 1.$$

Proof. $\langle 1-\mu[{}_nP_A^\delta\cap P_A]\rangle$ is a sequence of random variables taking all their values on $[0, 1]$; if $\lim_{n\to\infty} E\mu[{}_nP_A^\delta\cap P_A]=1$ then $\lim_{n\to\infty} E(1-\mu[{}_nP_A^\delta\cap P_A])^2=0$ and hence $\langle L_n\rangle$ converges in probability to zero.

Conversely suppose that for all positive ϵ and δ there is an $N_{\varepsilon,\delta}$ such that for all $n>N_{\varepsilon,\delta}$:

$$P(\mu[{}_nP_A^\delta \cap P_A] < 1-\epsilon) < \delta.$$

Since

$$E\mu[{}_nP_A^\delta \cap P_A] = \int_0^{(1-\varepsilon)^-} x\,dQ_n(x) + \int_{(1-\varepsilon)^+}^1 x\,dQ_n(x)$$

where Q_n is the distribution of $\mu[{}_nP_A^\delta \cap P_A]$, then

$$(1-\varepsilon)[1-P(\mu[{}_nP_A^\delta \cap P_A] < 1-\varepsilon)] \le E\mu[{}_nP_A^\delta \cap P_A] \le 1.$$

Therefore for all $n > N_{\varepsilon,\delta}$

$$(1-\varepsilon)(1-\delta) < E\mu[{}_nP_A^\delta \cap P_A] \le 1$$

and

$$\lim_{n\to\infty} E\mu[{}_nP_A^\delta \cap P_A] = 1.$$

Q.E.D.

3. **Application to the simple nearest-point rules** δ'. The color assigned to a point $s \in A$, by the simple nearest-point rule, is the color of the unique sample-point $N(s)$ nearest to s (if $s \in A$ has not a unique nearest-sample point we assign to s an arbitrary color). For that rule Switzer [1, p. 139, Theorem 1] has shown that

$$E\mu[{}_nP_A^{\delta'} \cap P_A] = \sum_{j=1}^m p_j \sum_{i=1}^n \int_{S_i} P_{jj}[|s-s_i|]\,d\mu(s)$$

where $S_i = \{s \in A : N(s) = s_i\}$.

As a consequence of our theorem we obtain:

COROLLARY 1. *δ' is consistent if and only if for all j such that $p_j > 0$,*

$$\lim_{n\to\infty} \sum_{i=1}^n \int_{S_i} P_{jj}[|s-s_i|]\,d\mu(s) = 1.$$

From Corollary 1, the following corollary can be easily proved.

COROLLARY 2. *If for all j such that $p_j > 0$, P_{jj} is a nonincreasing function, strictly decreasing in a neighbourhood of zero, and such that $P_{jj}(0) = 1$ then δ' is consistent if*

(1) $$\lim_{n\to\infty} \max_{1\le i\le n} \sup_{S_i} |s-s_i| = 0.$$

In the case where $k=2$, $m=2$ and

(2) $$\begin{aligned} P_{11}(d) &= p_1 + (1-p_1)\,e^{-cd} \\ P_{22}(d) &= (1-p_1) + p_1\,e^{-cd} \end{aligned} \quad c > 0.$$

Switzer [1, p. 142] considers certain types of arrangements for the sample-points. Because the functions given by (2) satisfy the conditions of Corollary 2 we can make the following observations about some arrangements suggested by Switzer:

(1) If the sample-points are placed at the vertices of a square grid then δ' is

consistent ($n \to \infty$ means here that the number of squares in the grid goes to infinity).

(2) If the sample-points are placed at the vertices of an equilateral triangular grid then δ' is consistent ($n \to \infty$ means here that the number of triangles in the grid goes to infinity).

(3) In the case called "line sampling with spacing σ" the condition (1) of Corollary 2 is not satisfied. Moreover, we have

$$\lim_{n\to\infty} 1-E\mu[{}_nP_A^{\delta'} \cap P_A] = 2p_1(1-p_1)[1-2(1-e^{-\frac{1}{2}c\sigma})(c\sigma)^{-1}]$$

which is not zero for any $\sigma > 0$, then in that case δ' is not consistent.

REMARK. We have supposed that the color observed at each sample-point is the color of that point (i.e. there is no "measurement error"). The following example shows that this assumption is important for the validity of Corollary 2. Denote by Y_{s_i} the observation made at the point s_i; suppose that Y_{s_i} is a random variable with values in $\{1, \ldots, m\}$ (the event $Y_{s_i} = j$ is interpreted as "the point s_i is observed to be in A_j"). In addition, suppose that for each $i = 1, \ldots n$,

$$P[Y_{s_i} = j \mid s_i \in A_k] = f_{jk}, \quad k, j = 1, \ldots, m,$$

and that the distribution of Y_{s_i} does not depend on the pattern in any other way. It is easy (cf. [2]) to prove that

$$E\mu[{}_nP_A^{\delta'} \cap P_A] = \sum_{j=1}^{m} p_j \sum_{i=1}^{n} \int_{S_i} \sum_{k=1}^{m} f_{jk} P_{kj}(|s-s_i|)\, d\mu(s).$$

Then, under the conditions of Corollary 2, if there exists a j such that $p_j > 0$ and $f_{jj} < 1$, δ' is not consistent even if (1) is satisfied.

ACKNOWLEDGEMENT. I wish to thank Professor Paul Switzer for having read the first draft of this note and for having called [2] to my attention.

REFERENCES

1. P. Switzer, *Reconstructing patterns from sample data*, Ann. Math. Statist. **38** (1967), 138–154.

2. ———, *Mapping a geographically correlated environment*, Technical Report No. 145, Stanford University, 1969.

ECOLE POLYTECHNIQUE,
MONTRÉAL, QUÉBEC

https://doi.org/10.4153/CMB-1972-057-5 Published online by Cambridge University Press