# Agglomerative Methods in Clustering with Constraints

Vladimir Batagelj
Department of Mathematics
University Edvard Kardelj
Jadranska 19, 61 000 Ljubljana
Yugoslavia

1. december 1985 / TEX: 14. january 1998

**Abstract**

The clustering problem (with constraints) can be posed as an optimization problem determined by the clustering criterion function and the set of feasible clusterings. In the paper an attempt of justification of agglomerative methods for solving the clustering problem and the conditions for their applicability are presented. It is shown that under certain (general) asumptions about the form of the criterion function and about the structure of the set of feasible clusterings:

- the agglomerative methods are "greedy" methods for solving the clustering problem;
- these methods stop in the maximal clustering in the set of feasible clusterings partially ordered by clustering inclusion;
- every feasible clustering can be reached by the method (choosing appropriate dissimilarities between units).

**Math. Subj. Class. (1980)** : 62 H 30, 68 G 10

**Key words** : optimization, greedy methods, clustering inclusion.

Detailed version of the talk presented at the Joint European Meeting of the Psychometric nad Classification Societies, Jouy-en-Josas, France, July, 6-8, 1983.

This work was supported in part by the Research Council of Slovenia, Yugoslavia.

## 1 Preliminaries

The clustering problem (with constraints) can be posed as an optimization problem (Ferligoj and Batagelj, 1982, 1983):

Determine the clustering (= set of clusters) $\mathbf{C}^*$ for which

$$P(\mathbf{C}^*) = \min_{\mathbf{C} \in \mathbf{\Phi}} P(\mathbf{C})$$

where $\mathbf{\Phi}$ is the set of feasible clusterings (determined with constraints) and $P : \mathbf{C} \to \mathbb{R}_0^+$ is the clustering criterion function.

With few exceptions the clustering problem is too hard to be exactly solved efficiently. Therefore approximative/heuristic methods have to be used. Among these agglomerative (hierarchical) and local optimization methods are the most popular. In the paper an attempt of justification of agglomerative clustering methods and the conditions for their applicability are presented, generalizing some ideas from Ferligoj and Batagelj (1983).

Let us introduce some notions which we shall need in the following.

The clustering $\mathbf{C}$ is a *complete* clustering if it is a partition of the set of units $E$. We shall denote by $\mathbf{\Pi}(E)$ the set of all complete clusterings of $E$. Two among them

$$\mathbf{O} \equiv \{\{X\} : X \in E\}$$

and

$$\mathbf{I} \equiv \{E\}$$

deserve to be denoted by special symbols.

The set of feasible clusterings $\mathbf{\Phi}$ can be decomposed into "strata" (layers)

$$\mathbf{\Phi}_k = \{\mathbf{C} \in \mathbf{\Phi} : \mathrm{card}(\mathbf{C}) = k\}$$

It also determines the *feasibility predicate*

$$\Phi(\mathbf{C}) \equiv \mathbf{C} \in \mathbf{\Phi}$$

defined on $\mathcal{P}(\mathcal{P}(E) \setminus \{\emptyset\})$; and conversely

$$\mathbf{\Phi} \equiv \{\mathbf{C} \in \mathcal{P}(\mathcal{P}(E) \setminus \{\emptyset\}) :: \Phi(\mathbf{C})\}$$

In the set of all clusterings the relation of *clustering inclusion* $\sqsubseteq$ can be introduced by

$$\mathbf{C}_1 \sqsubseteq \mathbf{C}_2 \equiv \forall C_1 \in \mathbf{C}_1, C_2 \in \mathbf{C}_2 : C_1 \cap C_2 \in \{\emptyset, C_1\}$$

When $\mathbf{C}_1 \sqsubseteq \mathbf{C}_2$ holds we say that the clustering $\mathbf{C}_1$ is a *refinement* of the clustering $\mathbf{C}_2$.

It is well known that $(\mathbf{\Pi}(E), \sqsubseteq)$ is a partially ordered set (even more, semimodular lattice; Aigner, 1979). Because any subset of partially ordered set is also partially ordered, we have:

L1. Let $\mathbf{\Phi} \subseteq \mathbf{\Pi}(E)$ then $(\mathbf{\Phi}, \sqsubseteq)$ is a partially ordered set.

The clustering inclusion determines two related relations (on $\mathbf{\Phi}$):

$$\mathbf{C}_1 \sqsubset \mathbf{C}_2 \equiv \mathbf{C}_1 \sqsubseteq \mathbf{C}_2 \land \mathbf{C}_1 \neq \mathbf{C}_2$$

and

$$\mathbf{C}_1 \sqsubset\!\!\cdot\, \mathbf{C}_2 \equiv \mathbf{C}_1 \sqsubset \mathbf{C}_2 \land \neg\exists \mathbf{C} \in \mathbf{\Phi} : (\mathbf{C}_1 \sqsubset \mathbf{C} \land \mathbf{C} \sqsubset \mathbf{C}_2)$$

It is not difficult to show that:

L2. $\mathbf{C}_1 \sqsubset \mathbf{C}_2 \Rightarrow \mathrm{card}(\mathbf{C}_1) > \mathrm{card}(\mathbf{C}_2)$

## 2 Conditions on the structure of the set of feasible clusterings

In the following we shall assume that the set of feastible clusterings $\mathbf{\Phi}$ satisfies the following conditions:

**F1.** $\mathbf{\Phi} \subseteq \mathbf{\Pi}(E)$

**F2.** the feasibility predicate $\Phi$ is *local*, i.e., it has the form

$$\Phi(\mathbf{C}) = \bigwedge_{C \in \mathbf{C}} \phi(C)$$

(*) where $\phi(C)$ is a predicate defined on $\mathcal{P}(E) \setminus \{\emptyset\})$ (clusters). The intuitive meaning of $\phi(C)$ is:

$$\phi(C) \equiv \text{ the cluster } C \text{ is "good"}$$

Therefore the locality condition can be read: a "good" clustering $\mathbf{C} \in \mathbf{\Phi}$ consists of "good" clusters.

**F3.** $\mathbf{O} \in \mathbf{\Phi}$

**F4.** the predicate $\Phi$ has the property of *binary heredity* with respect to the *fusibility* predicate $\psi(C_1, C_2)$, i.e.,

$$C_1, C_2 \neq \emptyset \wedge C_1 \cap C_2 = \emptyset \wedge \phi(C_1) \wedge \phi(C_2) \wedge \psi(C_1, C_2) \Rightarrow \phi(C_1 \cup C_2)$$

This condition can be read: fusion of two "related" clusters from "good" clustering results in a "good" clustering.

**F5.** the predicate $\psi$ is *compatible* with clustering inclusion $\sqsubseteq$,i.e.,

$$\forall \mathbf{C}_1, \mathbf{C}_2 \in \mathbf{\Phi} : (\mathbf{C}_1 \sqsubseteq \mathbf{C}_2 \wedge \mathbf{C}_1 \setminus \mathbf{C}_2 = \{C_1, C_2\} \Rightarrow \psi(C_1, C_2) \vee \psi(C_2, C_1))$$

**F6.** the "*interpolation*" property holds in $\mathbf{\Phi}$, i.e.,

$$\forall \mathbf{C}_1, \mathbf{C}_2 \in \mathbf{\Phi} : (\mathbf{C}_1 \sqsubset \mathbf{C}_2 \wedge \mathrm{card}(\mathbf{C}_1) > \mathrm{card}(\mathbf{C}_2) + 1 \Rightarrow \exists \mathbf{C} \in \mathbf{\Phi} : (\mathbf{C}_1 \sqsubseteq \mathbf{C} \wedge \mathbf{C} \sqsubseteq \mathbf{C}_2))$$

**Example 1** It is easy to verify that the sets of feasible clusterings $\mathbf{\Phi}^i(R)$, $i = 1, 2, 4, 5$ from Ferligoj and Batagelj (1983) satisfy the conditions F1 - F6. For instance in the case of $\mathbf{\Phi}^1(R)$ we have:

$$\mathbf{\Phi}(C) \equiv (C, R \cap C \times C) \text{ is a weakly connected graph}$$

and

$$\psi(C_1, C_2) \equiv \exists X \in C_1, Y \in C_2 : (XRY \vee YRX)$$

But, in the case of $\mathbf{\Phi}^3(R)$, with

$$\phi(C) \equiv (C, R \cap C \times C) \text{ is a strongly connected graph}$$

and

$$\psi(C_1, C_2) \equiv \exists X \in C_1, Y \in C_2 : XRY \wedge \exists X \in C_1, Y \in C_2 : YRX$$

the property F6 fails (in general). The counterexample is given in figure 5 in Ferligoj and Batagelj (1983), for which we have

$$\{\{1,2\},\{3,4\},\{5,6\}\} \sqsubset \{\{1,2,3,4,5,6,\}\}$$

□

Let us list some conseqences of conditions F1 – F6. The less trivial proofs are given in the appeddix.

**C1.** $\forall \mathbf{C} \in \mathbf{\Phi} : \mathbf{O} \sqsubseteq \mathbf{C}$

**C2.** $\forall X \in E : \phi(\{X\})$

**C3.** the predicate $\phi$ has the property of *binary divisibility*, i.e.,

$$\phi(C) \wedge \mathrm{card}(C) > 1 \Rightarrow \exists C_1, C_2 \neq : (C_1 \cap C_2 = \emptyset \wedge C_1 \cup C_2 = C \wedge \phi(C_1) \wedge \phi(C_2))$$

In other words: each nontrivial "good" cluster can be split in two "good" clusters.

**C4.** let us define

$$\mathbf{\Psi}(C) \equiv \{(C_1, C_2) : C_1, C_2 \neq \wedge C_1 \cap C_2 = \emptyset \wedge C_1 \cup C_2 = C \wedge \phi(C_1) \wedge \phi(C_2) \wedge \psi(C_1, C_2)\}$$

then we can express a stronger result than C3

$$\phi(C) \wedge \mathrm{card}(C) > 1 \Rightarrow \mathbf{\Psi}(C) \neq \emptyset$$

which implies: each "good" clustering, different from $\mathbf{O}$, can be obtained by fusing two "related" clusters in a "good" clustering.

**C5.** $\mathbf{C}_1 \sqsubset \mathbf{C}_2 \Rightarrow \mathrm{card}(\mathbf{C}_1) = \mathrm{card}(\mathbf{C}_2) + 1$

**C6.** let us define a function $r : \mathbf{\Phi} \to \mathbb{N}$ by the equality

$$r(\mathbf{C}) = card(\mathbf{O}) - card(\mathbf{C})$$

It is easy to verify the propreties

**C6.1.** $r(\mathbf{O}) = 0$

**C6.2.** $\mathbf{C}_1 \sqsubset \mathbf{C}_2 \Rightarrow r(\mathbf{C}_2) = r(\mathbf{C}_1) + 1$

Therefore $r$ is a rank function – the Jordan-Dedekind condition holds in $(\mathbf{\Phi}, \sqsubseteq)$.

**C7.** let us denote with

$$\mathrm{Max}\, \mathbf{\Phi} \equiv \{\mathbf{C} \in \mathbf{\Phi} : \neg\exists \mathbf{C}' \in \mathbf{\Phi} : \mathbf{C} \sqsubset \mathbf{C}'\}$$

the set of all maximal clusterings in $(\mathbf{\Phi}, \sqsubseteq)$, then

$$\forall \mathbf{C} \in \mathbf{\Phi} \setminus \mathrm{Max}\, \mathbf{\Phi} \exists C_1, C_2 \in \mathbf{C} : \psi(C_1, C_2)$$

# 3 Criterion function

In this section we shall slightly generalize some definitions and results from Ferligoj and Batagelj (1982).

We shall call a *dissimilarity between clusters* a function $d : (C_1, C_2) \to \mathbb{R}_0^+$ which is symmetric, i.e., $d(C_1, C_2) = d(C_2, C_1)$.

Let $(\mathbb{R}_0^+, \oplus, 0, \leq)$ be an ordered abelian monoid. Then the criterion function $P$ is *compatible* with dissimilarity $d$ over $\mathbf{\Phi}$ iff:

**(i)** $P(\mathbf{C}) = \bigoplus_{C \in \mathbf{C}} p(C)$

**(ii)** $\forall X \in E : p(\{X\}) = 0$

**(iii)** $\forall C \subseteq E : (\phi(C) \wedge \mathrm{card}(C) > 1 \Rightarrow$
$p(C) = \min_{(C_1,C_2) \in \Psi(C)} (p(C_1) \oplus p(C_2) \oplus d(C_1, C_2))$

Now we are ready to state our first theorem:

**Theorem 1** *Let $P$ be compatible with $d$ over $\mathbf{\Phi}$, $\oplus$ distributes over* min, *and F1 – F5 hold, then*

$$P(\mathbf{C}_k^*) = \min_{\mathbf{C} \in \mathbf{\Phi}_k} P(\mathbf{C}) = \min_{\substack{C_1, C_2 \in \mathbf{C} \in \mathbf{\Phi}_{k+1} \\ \psi(C_1, C_2)}} (P(\mathbf{C}) \oplus d(C_1, C_2))$$

The proof of the theorem is given in the appendix.

The equality from theorem 1 can also be written in the form

$$P(\mathbf{C}_k^*) = \min_{\mathbf{C} \in \mathbf{\Phi}_{k+1}} (P(\mathbf{C}) \oplus \min_{\substack{C_1, C_2 \in \mathbf{C} \\ \psi(C_1, C_2)}} d(C_1, C_2))$$

from where we can see the following "greedy" approximation:

$$P(\mathbf{C}_k^*) = P(\mathbf{C}_{k+1}^*) \oplus \min_{\substack{C_1, C_2 \in \mathbf{C}_{k+1}^* \\ \psi(C_1, C_2)}} d(C_1, C_2)$$

which is the basis for the following agglomerative (binary) procedure for solving the clustering problem :

1. $k := n (= \mathrm{card}(E))$; $\mathbf{C}(k) := \mathbf{O}$;
2. **while** $\exists C_i, C_j \in \mathbf{C}(k)$: $(i \neq j \wedge \psi(C_1, C_2))$ **repeat**
2.1. $(C_p, C_q) := \mathrm{argmin}\{d(C_i, C_j) : i \neq j \wedge \psi(C_i, C_j)\}$;
2.2. $C := C_p \cup C_q$; $k := k - 1$;
2.3. $\mathbf{C}(k) := \mathbf{C}(k+1) \setminus \{C_p, C_q\} \cup \{C\}$;
2.4. determine $d(C, C_s)$ for all $C_s \in \mathbf{C}(k)$
3. $m := k$

Note that, because it is based on an approximation, this procedure is not an exact procedure for solving the clustering problem. But, because of the nature of clustering problem (Garey and Johnson, 1979; Shamos, 1976; Brucker, 1978), it seems that we are forced to search for and to use such procedures.

In other respects this procedure has some nice properties:

**Theorem 2** *All clusterings* $\mathbf{C}(k), k = n, n-1, \ldots, m$ *obtained by the described procedure are feasible. It holds*

$$\mathbf{C}(k) \in \mathbf{\Phi}_k, \quad k = n, n-1, \ldots, m$$

*and*

$$\mathbf{C}(m) \in \text{Max}\,\mathbf{\Phi}$$

An agglomerative procedure is said to be *compatible* with $\mathbf{\Phi}$ iff:

**(i)** every clustering obtained by the procedure is feasible,

**(ii)** every feasible clustering can be obtained by the procedure if we can at each step fuse any pair of "related" clusters.

For our procedure it can be shown:

**Theorem 3** *If* $\mathbf{\Phi}$ *satisfies the conditions F1 – F6 then the described procedure is compatible with* $\mathbf{\Phi}$.

The proofs of theorem 2 and theorem 3 are given in the appendix.

# 4 Conclusion

The results of this paper show that if the set of feasible clusterings $\mathbf{\Phi}$ satisfies the conditions F1 – F6, and the criterion function $P$ is compatible with dissimilarity $d$, then the described agglomerative procedure can be used for solving the clustering problem.

# A Some proofs

## Properties C3 and C4

Let $\mathbf{C} \in \mathcal{P}(E) \setminus \{\emptyset\}$ be a "good" cluster. Then after F2 and C2 the clustering

$$\mathbf{C}(C) \equiv \mathbf{O} \setminus \{\{X\} : X \in C\} \cup \{C\}$$

belongs to the set of feasible clusterings $\mathbf{\Phi}$. Combining this with C1 we have

$$\mathbf{O} \sqsubseteq \mathbf{C}(C)$$

Let $C$ be also nontrivial ($\text{card}(C) > 1$). In this case we can "execute" the following procedure

1. $\mathbf{C} := \mathbf{O}$;
2. **while** $\text{card}(\mathbf{C}) > \text{card}(\mathbf{C}(C)) + 1$ **repeat**
2.1. $\mathbf{C} := \mathbf{C}'$, where $\mathbf{C}'$ is a feasible clustering satisfying $\mathbf{C} \sqsubset \mathbf{C}' \sqsubset \mathbf{C}(C)$;
3. $\{C_1, C_2\} := \mathbf{C} \setminus \mathbf{C}(C)$

Some comments. The existence of the clustering $\mathbf{C}'$ in the step 2.1. is assured by F6. The termination of the while loop follows from the observation that the difference $\mathrm{card}(\mathbf{C}) - \mathrm{card}(\mathbf{C}(C))$ is positive and is diminshed for at least one after each execution of the body of the loop. At the termination of the while loop the equality $\mathrm{card}(\mathbf{C}) = \mathrm{card}(\mathbf{C}(C)) + 1$ holds. Considering step 3, F5 and the definition of clustering inclusion we obtain C3 and C4.

## Property C7

Let $\mathbf{C} \in \mathbf{\Phi}$ be a nonmaximal clustering. Then there exists a clustering $\mathbf{C}_M \in \mathrm{Max}\,\mathbf{\Phi}$ such that $\mathbf{C} \sqsubset \mathbf{C}_M$.

Again we construct a procedure similar to the previous one :

1. $\mathbf{C}' := \mathbf{C}_M$;
2. **while** $\mathrm{card}(\mathbf{C}) > \mathrm{card}(\mathbf{C}') + 1$ **repeat**

2.1. $\mathbf{C}' := \mathbf{C}''$, where $\mathbf{C}''$ is a feasible clustering satisfying $\mathbf{C} \sqsubset \mathbf{C}'' \sqsubset \mathbf{C}'$;

3. $\{C_1, C_2\} := \mathbf{C} \setminus \mathbf{C}'$

with the same justification. Now, C7 follows directly from step 3 and F5.

## Theorem 1

We start with a sequence of equalities

$$\begin{aligned}
P(\mathbf{C}_k^*) &= \min_{\mathbf{C}\in\mathbf{\Phi}_k} P(\mathbf{C}) = \min_{\mathbf{C}\in\mathbf{\Phi}_k} \bigoplus_{D\in\mathbf{C}} p(D) = \\
&= \min_{C\in\mathbf{C}\in\mathbf{\Phi}_k} \Big( \bigoplus_{D\in\mathbf{C}\setminus\{C\}} p(D) \quad \oplus p(C)\Big) = \\
&= \min_{C\in\mathbf{C}\in\mathbf{\Phi}_k} \Big( \bigoplus_{D\in\mathbf{C}\setminus\{C\}} p(D) \quad \oplus \min_{(C_1,C_2)\in\mathbf{\Psi}(C)} (p(C_1) \oplus p(C_2) \oplus d(C_1, C_2)\Big) = \\
&= \min_{C\in\mathbf{C}\in\mathbf{\Phi}_k} \min_{(C_1,C_2)\in\mathbf{\Psi}(C)} \Big( \bigoplus_{D\in\mathbf{C}\setminus\{C\}} p(D) \quad \oplus (p(C_1) \oplus p(C_2) \oplus d(C_1, C_2)\Big) = \\
&= \min_{C\in\mathbf{C}\in\mathbf{\Phi}_k} \min_{(C_1,C_2)\in\mathbf{\Psi}(C)} \Big(d(C_1, C_2) \oplus \bigoplus_{D\in(\mathbf{C}\setminus\{C\})\cup\{C_1,C_2\}} p(D)\Big) = \\
&= \min_{\substack{C_1,C_2\in\mathbf{C}'\in\mathbf{\Phi}_{k+1} \\ \psi(C_1,C_2)}} \Big(d(C_1, C_2) \oplus \bigoplus_{D\in\mathbf{C}'} p(D)\Big)
\end{aligned}$$

All the equalities, except the last one, are trivial consequences of the theorem assumptions.

To show that the last equality holds we must show that both expressions consist of the same "terms".

Let us first show that

$$\mathbf{C}' = (\mathbf{C} \setminus \{C\}) \cup \{C_1, C_2\}) \in \mathbf{\Phi}_{k+1}$$

Because $\mathbf{C} \in \mathbf{\Phi}_k$ it holds $\forall D \in \mathbf{C} : \phi(D)$. Evidently $\mathbf{C}' \in \mathbf{\Pi}(E)_{k+1}$. From $(C_1, C_2) \in \mathbf{\Psi}(C)$ we see that $\phi(C_1) \wedge \phi(C_2)$ holds. Therefore $\forall D \in \mathbf{C}' : \phi(D)$, and finally $\mathbf{C}' \in \mathbf{\Phi}_{k+1}$.

Now, let us show that also each $\mathbf{C}' \in \mathbf{\Phi}_{k+1}$ can be expressed in the form required by the left side of the equality. Let $C_1, C_2 \in \mathbf{C}'$ such that $\psi(C_1, C_2)$ holds and $C = C_1 \cup C_2$. Evidently $(C_1, C_2) \in \mathbf{\Psi}(C)$. Let us denote

$$\mathbf{C} = (\mathbf{C}' \setminus \{C_1, C_2\}) \cup \{C\}$$

Because $\mathbf{C}' \in \mathbf{\Phi}_{k+1}$ it holds $\forall D \in \mathbf{C}' : \phi(D)$. From $(C_1, C_2) \in \mathbf{\Psi}(C)$ and F4 it follows that $\phi(C)$ holds. Therefore $\mathbf{C} \in \mathbf{\Phi}_k$. But, then we can express $\mathbf{C}'$ in the required form

$$\mathbf{C}' = (\mathbf{C}' \setminus \{C_1, C_2\} \cup \{C_1 \cup C_2\}) \setminus \{C_1 \cup C_2\} \cup \{C_1, C_2\} = \mathbf{C} \setminus \{C\} \cup \{C_1, C_2\}$$

We showed that each term from the left side expression can be expressed in the form required by the right side expression and vice-versa. Because of the idempotence of the operation min the question of uniqueness need not to be considered. Therefore the last equality also holds.

## Theorem 2

The step 1 of the clustering procedure combined with F3 implies $\mathbf{C}(n) = \mathbf{O} \in \mathbf{\Phi}_n$.

Let us now show: if $\mathbf{C}(k+1) \in \mathbf{\Phi}_{k+1}$ then, if it exists, $\mathbf{C}(k) \in \mathbf{\Phi}_k$. From the assumption $\mathbf{C}(k+1) \in \mathbf{\Phi}_{k+1}$ it follows $\forall C \in \mathbf{C}(k+1) : \phi(C)$; and from step 2.1 of the procedure $\psi(C/p, C/q)$. Therefore, by F4, $\phi(C_p \cup C_q)$ holds with a desired consequence $\mathbf{C}(k) \in \mathbf{\Phi}_k$.

The proprety $\mathbf{C}(m) \in \text{Max}\,\mathbf{\Phi}$ follows directly from the line 2 of the procedure and C7.

## Theorem 3

The requirement (i) from the definition of the compatibility is already proved in theorem 2.

To prove that also (ii) holds we proceed as follows. Let $\mathbf{C} \in \mathbf{\Phi}$. If $\mathbf{C} \neq \mathbf{O}$ then there exists by F6 a sequence of feasible clusterings

$$\mathbf{O} = \mathbf{C}_0 \sqsubset \mathbf{C}_1 \sqsubset \mathbf{C}_2 \sqsubset \ldots \sqsubset \mathbf{C}_{k-1} \sqsubset \mathbf{C}_k = \mathbf{C}$$

such that, by F5, each $\mathbf{C}_i$ is obtained fusing two clusters of $\mathbf{C}_{i-1}$.

# References


[1] Aigner M.: *Combinatorial theory.* New York: Springer - Verlag, 1979.

[2] Brucker P.: *On the complexity of clustering problems.* (R. Henn, B. Korte, W. Oettli Eds., Optimization and Operations Research, Proceedings, Bonn, 1977. Lecture Notes in Economics and Mathematical Systems 157), Berlin, Springer-Verlag, 1978.

[3] Ferligoj A., Batagelj V.: *Clustering with relational constraint.* Psychometrika, 1982, 47,413-426.

[4] Ferligoj A., Batagelj V.: *Some types of clustering with relational constraints.* Psychometrika, 1983, 48.

[5] Garey M.R., Johnson D.S.: *Computer and intractability.* San Francisco: Freeman, 1979.

[6] Shamos M.I.: *Geometry and statistics: Problems at the interface.* J. Traub (Ed.): Algorithms and Complexity (New directions and recent results). New York: Academic Press, 1976, 251-288.