LE MATEMATICHE
Vol. LX (2005) – Fasc. II, pp. 247–257

SOME OPEN PROBLEMS REGARDING THE DETERMINATION OF A SET FROM ITS COVARIOGRAM

GABRIELE BIANCHI

We present and discuss some open problems related to the determination of a set K from its covariogram, i.e. the function which provides the volumes of the intersections of K with all its possible translates.

1. Introduction and motivation.

Let K be a convex body in $\mathbb{R}^n$. The *covariogram* $g_K(x)$ of K is the function

$$g_K(x) = V_n(K \cap (K + x)),$$

where $x \in \mathbb{R}^n$ and V_n denotes n-dimensional Lebesgue measure in $\mathbb{R}^n$. This functional was introduced by G. Matheron in 1965 in his book on random sets [15]. In the literature the same notion appears also under different names, like *set covariance*, or *autocorrelation function* of a set.

Some properties of the covariogram are immediate: the support of g_K coincide with the *difference set* of K

$$K - K = \{k_1 - k_2 : k_1, k_2 \in K\}.$$

1991 *Mathemathics Subject Classification*: Primary 52A20; Secondary 60D05, 42C99, 52A22

Key words and pharses: Covariogram, Joint covariogram, Set covariance, Autocorrelation, Chord length distribution, Intersection with translates, Phase retrieval

Moreover g_K is unchanged by a translation or a reflection (in a point) of K. A consequence of the Brunn-Minkowski inequality is the concavity of $g_K^{1/n}$ on its support [23], p. 410-411.

Matheron in 1986 [16] asked the following problem and conjectured a positive answer for the case $n = 2$.

Covariogram problem. Does the covariogram determine a convex body, among all convex bodies, up to translation and reflection?

[7] contains a very detailed introduction to this and related problems. In this note we briefly recall the motivation of this problem, present some open problems and update [7] with recent information. Matheron observed [15], p. 86, that, for $r > 0$ and $u \in S^{n-1}$,

$$\frac{\partial}{\partial r} g_K(ru) = -V_{n-1}(\{y \in u^\perp : V_1(K \cap (l_u + y)) \geq r\}), \tag{1.1}$$

where $l_u + y$ denotes the line parallel to u through y and $u^\perp$ denotes the orthogonal complement of u. This formula allows some interpretation of the covariogram problem.

The right hand side of (1.1) coincides with the rearrangement of $X_u K$, the $X-$ray function of K in the direction u, see [9] for definitions. Thus while in convex tomography it is required to determine K from its $X-$ray functions in a finite number of directions, the covariogram problem asks for the determination of K from the knowledge of the rearrangements of $X_u K$, for all u.

The right hand side of (1.1) also gives the distribution of the lengths of the chords of K which are parallel to u. There are many examples in stereology, in statistical shape recognition and in image analysis where one tries to infer some properties of an unknown body from chord length measurements, see [22] and [8] as examples. Blaschke asked whether the chord length distribution characterises a set and Mallows and Clark [17] presented examples which provide a negative answer. Assume now to have the information about the distribution of the chord lengths not in the "completely mixed" form, but separated direction by direction: (1.1) shows that the problem of determining a body from this data, the "orientation-dependent" chord length distributions, is equivalent to the covariogram one, see [19].

The covariogram problem appear independently in other contexts. In the setting of probability theory Adler and Pyke [2] asked whether the distribution of the difference $X - Y$ of two independent random variables X, Y that are uniformly distributed over K determines K up to translation and reflection. Since it is easily proved that

$$g_K = 1_K * 1_{-K}\,, \tag{1.2}$$

and this convolution is, up to a multiplicative factor, the probability density of $X - Y$, this problem is equivalent to the previous one. The same authors [3] find the covariogram problem relevant also in the study of scanning Brownian processes and of the equivalence of measures induced by these processes for different base sets.

The covariogram problem also arises in Fourier analysis. The phase retrieval problem involves determining a function f in $\mathbb{R}^n$ from the modulus of its Fourier transform $\hat{f}$. It is of great relevance in many applications: as an important example this problem arises in $X-$ray crystallography, where one tries to determine the atomic crystal structure from $X-$ray diffraction images. As Rosenblatt [21] explains "Here the phase retrieval problem arises because the modulus of a Fourier transform is all that can usually be measured after diffraction occurs". We refer to the survey [14] or the book [12] for an introduction to the vast literature on this problem.

Taking Fourier transforms in (1.2) and using the relation $\widehat{1_{-K}}(\xi) = \overline{\widehat{1_K}(\xi)}$, we obtain

$$\widehat{g_K}(\xi) = \widehat{1_K}(\xi)\widehat{1_{-K}}(\xi) = |\widehat{1_K}(\xi)|^2 . \tag{1.3}$$

Thus the phase retrieval problem reduces to the covariogram problem when f is the characteristic function of a convex body. Or equivalently the phase retrieval problem is equivalent to a generalisation of the covariogram problem to functions, where one tries to determine a function f, in a suitable class, from the knowledge of $f(x) * f(-x)$.

We also mention two recent papers on related problems. Gardner, Gronchi and Zong [10] study the covariogram problem in a discrete setting and Jaming and Kolountzakis [13] discuss the problem of determining a function $f : \mathbb{R} \to \mathbb{R}$ from its k correlation function

$$g_{f,k}(x_1, \dots, x_{k-1}) = \int_{\mathbb{R}} f(t)f(t + x_1)\dots f(t + x_{k-1})\,dt ,$$

for an integer $k \geq 3$.

2. Complete solution to Matheron's conjecture in the plane.

Even in $\mathbb{R}^2$ the problem has not yet obtained a complete answer. The most general result regarding plane convex bodies is the following one, proved by this author in [5]. We say that two open arcs on the boundary of a convex body K are *opposite* if there is a point in one arc and a point in the other arc which have opposite outer normals.

Theorem 2.1. ([5]) *Let K be a plane convex body such that either*

(1) *K is not strictly convex or*
(2) *K is not C^1 or*
(3) *∂K contains two opposite open arcs which are C^2.*

The covariogram determines K, up to translations or reflections.

A property of this problem which might be useful for its complete solution is the fact that is "local".

Proposition 2.2. ([5]) *Let H, K be plane convex bodies with equal covariogram and let us assume that $\partial K \cap \partial H$ contains an open arc. Then H is a translation or a reflection of K.*

In view of the last two results solving the following problem would yield Matheron's conjecture in the plane.

Problem 1. Let $f_1, f_2, g_1, g_2 \in C^1([-1,1])$. Assume that $f_1, f_2 \geq 0$ are strictly convex, $g_1, g_2 \leq 0$ are strictly concave and $f_1(0) = f_2(0) = g_1(0) = g_2(0) = 0$. Assume moreover that for each (x_0, y_0) in a neighbourhood of $(0, 0)$ it is

$$\int_{[-1,1]\cap[-1+x_0,1+x_0]} \Big(g_1(x - x_0) + y_0 - f_1(x) \Big)^+ dx = \int_{[-1,1]\cap[-1+x_0,1+x_0]} \Big(g_2(x - x_0) + y_0 - f_2(x) \Big)^+ dx \tag{2.1}$$

Does there exist an open neighbourhood I of 0 such that

$$\text{either} \quad f_1(x) = f_2(x) \quad \text{and} \quad g_1(x) = g_2(x) \quad \forall x \in I$$
$$\text{or else} \quad f_1(x) = -g_2(-x) \quad \text{and} \quad g_1(x) = -f_2(-x) \quad \forall x \in I?$$

Here $f^+(x) = \max(f(x), 0)$ denote the positive part of f. The integrals in (2.1) express the area above the graphs of f_i and below the translated graph of g_i. We remark that in view of Theorem 2.1 it is interesting to study this problem without any assumption on the second derivatives of the functions.

There are two points in the proof of Theorem 2.1 (3) where the regularity of the boundary has a crucial role. To explain this let us sketch how Problem 1 can be solved under the assumption that f_1, f_2, g_1 and g_2 are C^2.

- Step 1. For $c \in \mathbb{R}$ in a neighbourhood of 0 let $x_{f_i} c = (df_i/dx)^{-1}(c)$ and $x_{g_i} c = (dg_i/dx)^{-1}(c)$. Fix c, consider translation vectors (x_0, y_0) such that

$$\Big(g_1(x - x_0) + y_0 - f_1(x) \Big)^+$$

is supported in a small neighbourhood J of $x_{f_1}c$ and study the asymptotic behaviour of the integrals in (2.1) as the translation vectors are changed in such a way that J shrink to a point. The equality (2.1), due to the C^2 regularity of the functions, implies the equality of the (non ordered) sets

$$\left\{\frac{d^2 f_1}{dx^2}(x_{f_1}c), -\frac{d^2 g_1}{dx^2}(x_{g_1}c)\right\} = \left\{\frac{d^2 f_2}{dx^2}(x_{f_2}c), -\frac{d^2 g_2}{dx^2}(x_{g_2}c)\right\}. \tag{2.2}$$

This result is typical of the covariogram problem, where the asymptotic behaviour of g_K near the point of its support with outer normal u, for a fixed $u \in S^{n-1}$, gives some information about the curvature of ∂K in the points with outer normal u and $-u$. It is not clear how to substitute this step when the regularity is missing.

- Step 2. If

$$\frac{d^2 f_1}{dx^2}(x_{f_1}c) = -\frac{d^2 g_1}{dx^2}(x_{g_1}c) \quad \text{for all } c \text{ in a neighbourhood of } 0 \tag{2.3}$$

then

$$f_1(x) \equiv -g_1(-x) \equiv f_2(x) \equiv -g_2(-x) \quad \text{in a neighbourhood of } 0,$$

and the desired conclusion is obtained. This follows from (2.2) and the fact that the four functions coincide at $x = 0$, together with their first derivatives.

- Step 3. Assume now that in each neighbourhood of 0 there are c_0 such that

$$\frac{d^2 f_1}{dx^2}(x_{f_1}c_0) \neq -\frac{d^2 g_1}{dx^2}(x_{g_1}c_0).$$

For those c_0 there exists a neighbourhood $W = W(c_0)$ of c_0 such that, either

$$\frac{d^2 f_1}{dx^2}(x_{f_1}c) = \frac{d^2 f_2}{dx^2}(x_{f_2}c) \quad \text{and} \tag{2.4}$$

$$\frac{d^2 g_1}{dx^2}(x_{g_1}c) = \frac{d^2 g_2}{dx^2}(x_{g_2}c) \quad \forall c \in W$$

or else

$$\frac{d^2 f_1}{dx^2}(x_{f_1}c) = -\frac{d^2 g_2}{dx^2}(x_{g_2}c) \quad \text{and} \tag{2.5}$$

$$\frac{d^2 f_2}{dx^2}(x_{f_2} c) = -\frac{d^2 g_1}{dx^2}(x_{g_1} c) \quad \forall c \in W.$$

This is the second point where the C^2 regularity of the functions comes into play.

- Step 4. More work is needed to prove that the choice among (2.4) and (2.5) is constant across adjacent different neighbourhoods W, work that we do not describe here and which does not involve the regularity of the functions. Once say (2.4) is established for all c in a neighbourhood of 0, the desired conclusion follows as in Step 2.

3. Algorithms to reconstruct convex polygons.

The first contribution to the question posed by Matheron was made in 1993 by Nagel [19], who gave a positive answer when K is a planar convex polygon, see also [4] for an alternative proof. Schmitt [22] gives an explicit reconstructive procedure under the assumption that the convex polygon does not have any pair of parallel edges.

Problem 2. Find an algorithm to reconstruct a convex polygon from its covariogram, assuming to have either exact data or approximate data.

When K is a convex polygon g_K has a peculiar structure, which might be useful to answer Problem 2. In the interior of its support g_K is a piecewise polynomial of second degree in x and

$$\left\{x \in \operatorname{int}(\operatorname{supp} g_K) : g_K \text{ is not } C^2 \text{ at } x\right\} = \bigcup_{p \text{ vertex } of\, K} \Big((\partial K - p) \cup (p - \partial K)\Big)$$

4. Joint covariogram problem for convex sets.

The positive result of Nagel for polygons has been recently much strengthened. Given two convex bodies K and L in $\mathbb{R}^2$ we define their *joint covariogram* $g_{K,L}(x)$ as the function

$$g_{K,L}(x) = V_2(K \cap (L + x)),$$

where $x \in \mathbb{R}^2$. It is evident that this function is invariant under a translation of K and L by the same vector, and that $g_{K,L} = g_{-L,-K}$. If K, L, K' and L' are convex bodies such that, for some $\tau \in \mathbb{R}^2$, either $(K', L') = (K + \tau, L + \tau)$ or $(K', L') = (-L + \tau, -K + \tau)$ we say that (K', L') and (K, L) are trivial associates. A. A. Volčič and R. Gardner a few years ago posed the following question.

Joint covariogram problem. Does the joint covariogram of $K \subset \mathbb{R}^2$ and $L \subset \mathbb{R}^2$ determine the pair (K, L), among all pairs of convex bodies, up to trivial associates?

This problem restricted to the class of convex polygons has been completely solved by this author in [6].

Proposition 4.1. ([6]) *Let K, L be convex polygons, $K', L' \subset \mathbb{R}^2$ be convex bodies with*

$$g_{K,L} = g_{K',L'}.$$

(1) *There exists a family $\mathcal{P}$ of pairs of parallelograms such that if $(K, L) \in \mathcal{P}$ the answer to the joint covariogram problem is negative;*
(2) *Assume that (K, L) and (K', L') do not belong to $\mathcal{P}$. Then (K, L) and (K', L') are trivial associates.*

This result can be interpreted as showing that the information provided by the covariogram for convex polygons is so rich to be able to determine not only one unknown polygon, but two of them, except for the "meager" family $\mathcal{P}$ of exceptions.

In view of the previous results we pose the following problem. Let C^2_+ denote the class of convex bodies whose boundary is C^2 and has strictly positive Gaussian curvature.

Problem 3. Solve the joint covariogram problem in the class C^2_+ of convex plane bodies. Are there negative examples in this class?

It is interesting to understand if the class of C^2_+ bodies and the class of polygons behave differently with respect to this problem or to the ordinary covariogram problem.

The first three steps of the proof of Theorem 2.1 (3) described in Section 2 carry over to the context of Problem 3. The step which corresponds to Step 4 however requires new ideas. It is not clear whether the ideas used to overcome this difficulty in the case of polygons carry over to C^2_+ bodies.

5. Higher dimensional bodies.

Very little is known regarding the covariogram problem when the space dimension is larger than two.

It is known that centrally symmetric convex bodies, in any dimension, are determined by their covariogram, up to translations. This is a consequence of the fact that g_K determines the volume of K $(= g_K(0))$ and its difference body $K - K$ and of the Brunn-Minkowski inequality. This inequality implies that among all convex bodies with the same difference body the centrally symmetric one is the only set of maximal volume, see [9], Theorem 3.2.3 .

Goodey, Schneider and Weil [11] proves that, if P is an $n-$dimensional simplicial polytope, and P and $-P$ are in general relative position, the covariogram determines P. We explain this assumptions in the case of a three-dimensional polytope P: P is simplicial if all its facets are triangles; P and $-P$ are in general relative position if (i) P has no pair of parallel facets, (ii) it has no pair of parallel edges on "opposite sides" of P (i.e. contained in parallel supporting planes), and (iii) has no edge parallel to a facet and "opposite" to it.

Averkov and this author [1] prove that it suffices to know the covariogram in any open neighbourhood of the boundary of its support in order to determine most (in the sense of Baire category) convex bodies in $\mathbb{R}^n$, for any $n \geq 2$, and they also prove a more precise results for $n = 2$.

When the space has dimension $n \geq 4$, this author [5] has given some negative answers.

Theorem 5.1. ([5]) *Let $K \subset \mathbb{R}^n$ and $H \subset \mathbb{R}^k$ be convex bodies and let $\mathcal{L}$ be a nondegenerate linear transformation. Then the two convex bodies $\mathcal{L}(K \times H)$ and $\mathcal{L}(K \times (-H))$ in $\mathbb{R}^{n+k}$ have the same covariogram.*

If neither K nor H are centrally symmetric then $\mathcal{L}(K \times H)$ cannot be obtained from $\mathcal{L}(K \times (-H))$ through a translation or a reflection.

For each $n \geq 4$ there exist pairs of convex polytopes in $\mathbb{R}^n$ which have the same covariogram and which are not a translation or a reflection of each other.

We remark that the previous construction does not provide any example with smooth boundary.

What does it happen for three-dimensional polytopes? The previous construction does not apply because any one-dimensional convex set is centrally symmetric. The next result provide a positive answer for dimension $n = 3$.

Theorem 5.2. ([6]) *Let $P \subset \mathbb{R}^3$ be a convex polytope with nonempty interior. The knowledge of g_P determines P, in the class of convex bodies in $\mathbb{R}^3$, up to translations and reflections.*

Many questions are open here, and their answer might be quite deceived. We mention two natural ones.

Problem 4. Understand which four dimensional polytopes are determined by their covariogram and which are not.

Problem 5. Solve the covariogram problem in the class of three-dimensional C^2_+ convex bodies.

Regarding Problem 5, a difference with it two-dimensional counterpart is that "the proof does not even start" in the sense that it is not clear how to extract from the covariogram the complete information, up to a reflection, about the curvature of the boundary.

To clarify this point, let K be a C^2_+ convex body and, for $u \in S^2$, let $x(u) \in \partial K$ denote the point with outer normal u. The asymptotic behaviour of g_K near the point of $K - K$ with outer normal u determines, for each $t \in u^\perp$,

$$\frac{1}{k(x(u),t)} + \frac{1}{k(x(-u),t)} \quad \text{and} \quad \{G_{\partial K}(x(u)), G_{\partial K}(x(-u))\}\,. \tag{5.1}$$

Here $k(x(u),t)$ and $G_{\partial K}(x(u))$ denote respectively the sectional curvature in direction t and the Gaussian curvature of ∂K at $x(u)$. This information does not fully determine the second order behaviour of the boundary in $x(u)$, up to a reflection.

Some more information about the curvature can be obtained by the second order derivatives of g_K near 0. Assuming higher smoothness of ∂K, starting from the formulas proved in [18] for pure and mixed second order derivatives of g_K, one can compute that

$$\frac{\partial^2}{\partial r^2} g_K(ru) = \frac{r}{4} \int_{w \in S^2 : <w \cdot u> = 0} \frac{< II_{\partial K}(x(w))u, u >^2}{G_{\partial K}(x(w))}\, ds(w) + o(r), \tag{5.2}$$

as $r \to 0^+$. Here $II_{\partial K}(x)$ denotes the second fundamental form of ∂K at x. With similar ideas and using mixed instead of pure derivatives one can compute from g_K

$$\int_{w \in S^2 : <w \cdot u> = 0} \frac{< II_{\partial K}(x(w))u, u >< II_{\partial K}(x(w))t, t >}{G_{\partial K}(x(w))}\, ds(w), \tag{5.3}$$

where $t = t(w) \in S^2 \cap w^\perp \cap u^\perp$.

Rataj [21] also presents a formula which is related to (5.2). It is not clear how to extract information about the curvature of ∂K from the integrals in (5.2) and (5.3).

6. Substitute the volume with more general functionals.

Problem 6. What does it happen to the covariogram problem if we change the definition of covariogram to

$$g_K(x) = \mathcal{F}(K \cap (K + x))$$

where $\mathcal{F}$ is a functional different from the volume?

The following cases seem worth investigation: the perimeter; the $(n-1)-$volume of the projection on a fixed hyperplane; more generally the mixed volume with a fixed convex set.

G. Averkov and this author are investigating some aspects of Problem 6.

REFERENCES

[1] G. Averkov - G. Bianchi, *Retrieving convex bodies from restricted covariogram functions,* preprint.

[2] R.J. Adler - R. Pyke, *Problem 91-3,* The Institute of Mathematical Statistics Bulletin, 20 (1991), p. 409.

[3] R.J. Adler - R. Pyke, *Scanning Brownian processes,* Adv. Appl. Prob. (SGSA), 29 (1997), pp. 295–326.

[4] G. Bianchi, *Determining convex polygons from their covariogram,* Adv. Appl. Prob. (SGSA), 34 (2002), pp. 261–266.

[5] G. Bianchi, *Matheron's conjecture for the covariogram problem,* J. London Math. Soc., 71 (2005), pp. 203–220.

[6] G. Bianchi, *The covariogram determines three dimensional convex polytopes,* preprint.

[7] G. Bianchi - F. Segala - A. Volčič, *The solution of the covariogram problem for plane C^2_+ bodies,* J. Diff. Geom., 60 (2002), pp. 177–198.

[8] A. Cabo - A. Baddeley, *Estimation of mean particle volume using the set covariance function,* Adv. Appl. Probab., 35 (2003), pp. 27–46.

[9] R.J. Gardner, *Geometric Tomography,* Cambridge University Press, New York, 1995.

[10] R.J. Gardner - P. Gronchi - C. Zong, *Sums, projections, and sections of lattice sets, and the discrete covariogram,* Discrete Comput. Geom., 34 (2005), pp. 391–409.

[11] P. Goodey - R. Schneider - W. Weil, *On the determination of convex bodies by projection functions,* Bull. London Math. Soc., 29 (1997), pp. 82–88.

[12] N.E. Hurt, *Phase retrieval and zero crossing,* Kluwer, Dordrecht, 1989.

[13] P. Jaming - M.N. Kolountzakis, *Reconstruction of functions from their triple correlations,* New York J. Math., 9 (2003), pp. 149–164 (electronic).

[14] M.V. Klibanov - P.E. Sacks - A.V. Tikhonravov, *The phase retrieval problem,* Inverse Problems, 11 (1995), pp. 1–28.

[15] G. Matheron, *Random sets and integral geometry,* Wiley, New York, 1975.

[16] G. Matheron, *Le covariogramme géometrique des compacts convexes de* $\mathbb{R}^2$*,* Technical report 2/86, Centre de Géostatistique, Ecole des Mines de Paris, 54 (1986).

[17] C.L. Mallows - J.M. Clark, *Corrections to "Linear-intercept distributions do not characterize plane sets",* J. Appl. Prob., 8 (1971), pp. 208–209.

[18] M. Meyer - S. Reisner - M. Schmuckenschläger, *The volume of the intersection of a convex body with its translates,* Mathematika 40 (1993), pp. 278–289.

[19] W. Nagel, *Orientation-dependent chord length distributions characterise convex polygons,* J. Appl. Prob., 30 (1993), pp. 730–736.

[20] J. Rataj, *On set covariance and three-point test sets,* Czechoslovak Math. J., 54 (2004), pp. 205–214.

[21] J. Rosenblatt, *Phase retrieval,* Comm. Math. Phys., 95 (1984), pp. 317–343.

[22] M. Schmitt, *On two inverse problems in mathematical morphology,* Mathematical Morphology in Image Processing, Dekker, New York, 1993, pp.151–169.

[23] R. Schneider, *Convex bodies: the Brunn-Minkowsky theory,* Cambridge University Press, Cambridge 1993.

Dipartimento di Matematica,
Università di Firenze,
Viale Morgagni 67/A, 50134 Firenze (ITALY)
e-mail: gabriele.bianchi@unifi.it