Covariance matrices of self-affine measures

Abstract

In this paper we derive a formula for a covariance matrix of any self-affine measure, i.e. a probability measure $\mu$ satisfying

$$\mu = \sum_{k=1}^{t} p_k \mu \circ S_k^{-1},$$

where $\{S_k(x) = A_k x + b_k\}_{1 \leq k \leq l}$ is a family of affine contractive maps and $\{p_k\}_{1 \leq k \leq l}$ is a set of probability weights. In particular if for every $k$, $A_k = A$ then the formula will be have the following form

$$D^2X = [I \otimes I - A \otimes A]^{-1}D^2B,$$

where $D^2X$ denote the covariance matrix of the measure $\mu$ and $D^2B$ denote a covariance matrix of a discret random variable $B$ with values $b_k$, and corresponding probabilites $p_k$.

2000 Mathematics Subject Classification: 28A80(60A10)

1 Introduction

We will consider a invariant probability measure on $\mathbb{R}^d$

(1) $$\mu = \sum_{k=1}^{t} p_k \mu \circ S_k^{-1},$$

where $\{S_k\}_{1 \leq k \leq l}$ is a family of contractive maps and $\{p_k\}_{1 \leq k \leq l}$ is a set of probability weights. It is often assumed that the maps are similitudes. We will make a more general assumption that the maps are affine contractions on $\mathbb{R}^d$; i.e. $S_k(x) = A_k x + b_k$ and the operator norm $\|A_k\| < 1$ for all $k$.

The definition (1) was introduced by Hutchinson [H]. But an example of such measures has been studied for a long time in the context of Bernoulli convolution, i.e. the example of an invariant measure on the real line

(2) $$\mu = \frac{1}{2} (\mu \circ S_1^{-1} + \mu \circ S_2^{-1}),$$

1
where \( S_1(x) = \beta(x + 1) \), \( S_2(x) = \beta(x - 1) \) for \( \beta \in (0,1) \). It remains difficult open problem to characterize the set of \( \beta \) for which \( \mu \) is absolutely continuous. An another example has been studied in great detail in wavelet theory in connection with the dilatation equation

\[
(3) \quad f(x) = \sum_{k=1}^{l} c_k f(2x - (k - 1)).
\]

The function \( f \) can be considered as the density function of the corresponding absolutely continuous self-affine measure \( \mu \) for \( S_k(x) = \frac{1}{2}(x + (k - 1)) \) and \( p_k = \frac{1}{2}c_k \). In wavelet theory the \( c_k \) may be negative but \( \sum c_k \) must be 2. The invariant measures have a natural connection with fractal geometry [F], because their supports are compact invariant sets

\[
(4) \quad K = \bigcup_{k=1}^{l} S_k(K).
\]

These measures arise also in another areas of mathematics.

A fundamental method for studying these measures is the Fourier transform [S]. Our goal is to show that we can use another probabilistic tools, not only characteristic function, to investigate the self-affine measures. We derive a formula for covariance matrices and give an example of an investigation of measures on Sierpinski triangle.

## 2 Covariance matrices of self-affine measures

The invariant measure \( \mu \) satisfies the following identity

\[
(5) \quad \int_{\mathbb{R}^d} f d\mu = \sum_{k=1}^{l} p_k \int_{\mathbb{R}^d} f \circ S_k d\mu,
\]

where \( f \) is any continuous function on \( \mathbb{R}^d \) [B]. We will denote by \( X \) some \( d \)-dimensional random variable with respect to the probability distribution \( \mu \). We apply the identity (5) to the coordinate functions \( e_i^*(x) = x_i \) of the point \( x = (x_i)_{1 \leq i \leq d} \in \mathbb{R}^d \). By the above a vector of expected values \( EX \) will be equal
The above relation gives

\[
[I - \sum_{k=1}^{l} p_k A_k]EX = \sum_{k=1}^{l} p_k b_k,
\]

where \(I\) is identity matrix on \(\mathbb{R}^d\). The sum \(\sum_{k=1}^{l} p_k b_k\) is a vector of expected values of a \(d\)-dimensional random variable \(B\) with values \(b_k\), and corresponding probabilities \(p_k\). Since \(\|A_k\| < 1\) for all \(k\) then \(\|\sum_{k=1}^{l} p_k A_k\| < 1\). It follows that 1 is not an eigenvalue of the operator \(\sum_{k=1}^{l} p_k A_k\). For this reason the operator \(I - \sum_{k=1}^{l} p_k A_k\) will be invertible. We can rewrite (7) as

\[
EX = [I - \sum_{k=1}^{l} p_k A_k]^{-1}EB.
\]

This means that the expected value of \(X\) linearly depend on the expected value of \(B\).

Let \(X \otimes X = [x_i x_j]_{1 \leq i,j \leq d}\) denote the second order tensor build from the coordinates. Using (5) a matrix of second order moments \(E(X \otimes X)\) will be equal

\[
E(X \otimes X) = \int_{\mathbb{R}^d} e_i^*(x) e_j^*(x) d\mu_{1 \leq i,j \leq d}
\]

\[
= \sum_{k=1}^{l} p_k \int_{\mathbb{R}^d} e_i^*(A_k x + b_k) e_j^*(A_k x + b_k) d\mu_{1 \leq i,j \leq d}
\]

\[
= \sum_{k=1}^{l} p_k E((A_k x + b_k) \otimes (A_k x + b_k))
\]

(9) \[
= \sum_{k=1}^{l} p_k [(A_k \otimes A_k) E(X \otimes X) + b_k \otimes A_k EX + A_k EX \otimes b_k + b_k \otimes b_k]
\]
Therefore

\[
[I \otimes I - \sum_{k=1}^{l} p_k(A_k \otimes A_k)]E(X \otimes X) = \sum_{k=1}^{l} p_k(b_k \otimes A_k EX + A_k EX \otimes b_k + b_k \otimes b_k)
\]

(10)

The operator norm of \(A_k \otimes A_k\) is less than 1 on \(\mathbb{R}^d \otimes \mathbb{R}^d\), so, by the same argument as early, we get that the operator \(I \otimes I - \sum_{k=1}^{l} p_k(A_k \otimes A_k)\) is invertible and we obtain

\[
E(X \otimes X) = [I \otimes I - \sum_{k=1}^{l} p_k(A_k \otimes A_k)]^{-1} \times \\
\sum_{k=1}^{l} p_k(b_k \otimes A_k EX + A_k EX \otimes b_k + b_k \otimes b_k).
\]

(11)

Substituting (8) into \(D^2X = E(X \otimes X) - EX \otimes EX\) we can obtain a formula for the covariance matrix of \(X\). But in the general case it will be complicated. This formula takes a surprising simple form when we assumed that all affine maps have the same linear part, i.e. all \(A_k = A\). Under this assumption, using the standard tensor calculus we get

\[
D^2X = [I \otimes I - A \otimes A]^{-1} \left\{ [I \otimes A(I-A)^{-1}] (\mathcal{E}\mathcal{B} \otimes \mathcal{E}\mathcal{B}) + A(I-A)^{-1} \otimes [I \otimes A(I-A)^{-1}] (\mathcal{E}\mathcal{B} \otimes \mathcal{E}\mathcal{B}) \right\} \\
- [I \otimes A(I-A)^{-1} \otimes (I-A)]^{-1} (\mathcal{E}\mathcal{B} \otimes \mathcal{E}\mathcal{B}) \\
= [I \otimes I - A \otimes A]^{-1} (\mathcal{E}(\mathcal{B} \otimes \mathcal{B}) - \mathcal{E}\mathcal{B} \otimes \mathcal{E}\mathcal{B})
\]

(12)

Notice now that the expression \(\mathcal{E}(\mathcal{B} \otimes \mathcal{B}) - \mathcal{E}\mathcal{B} \otimes \mathcal{E}\mathcal{B}\) is the covariance matrix of \(\mathcal{B}\). Thus we obtain

\[
D^2X = [I \otimes I - A \otimes A]^{-1} D^2\mathcal{B}.
\]

(13)

In other words we obtained the following proposition.

**PROPOSITION**

Assume that \(\mu\) is a self-affine measure on \(\mathbb{R}^d\) for a family of linear contractions \(S_k(x) = Ax + b_k\), \(1 \leq k \leq l\). Let \(X\) be some random variable with respect to the probability distribution \(\mu\). Then the covariance matrix of the random variable \(X\)

\[
D^2X = [I \otimes I - A \otimes A]^{-1} D^2\mathcal{B},
\]

(14)
where $D^2B$ denote the covariance matrix of the random variable $B$. □

Remark. When all affine maps $S_k$ have the same linear part then not only the expected value of $X$ linearly depend on expected value of $B$ but also covariance matrix of $X$ linearly depend on the covariance matrix of $B$.

If the matrix $A$ is diagonal then diagonal is the matrix $[I \otimes I - A \otimes A]^{-1}$.

Therefore we get the simple corollary.

COROLLARY

If under the assumptions of Proposition we assume additionaly that the matrix $A$ is diagonal then $X_i = e_i^*(X)$ and $X_j = e_j^*(X)$ are uncorrelated if and only if uncorrelated are $(e_i^*(b_k))_{1 \leq k \leq l}$ and $(e_j^*(b_k))_{1 \leq k \leq l}$. □

In other words we have obtained the following law

$$\int_{\mathbb{R}^d} x_i x_j d\mu = \int_{\mathbb{R}^d} x_i d\mu \int_{\mathbb{R}^d} x_j d\mu \iff \sum_{k=1}^l p_k e_i^*(b_k) e_j^*(b_k) = \sum_{k=1}^l p_k e_i^*(b_k) \sum_{k=1}^l p_k e_j^*(b_k).$$

3 Example

Consider Sierpinski triangle with vertices at the points $(0,0)$, $(1,0)$ and $(\frac{1}{2}, \frac{\sqrt{3}}{2})$. The Sierpinski triangle is an invariant compact set of three contractions on $\mathbb{R}^2$: $S_1(x) = \frac{1}{2}x$, $S_2(x) = \frac{1}{2}x + (\frac{1}{2}, 0)$ and $S_3(x) = \frac{1}{2}x + (\frac{1}{4}, \frac{\sqrt{3}}{4})$. In this case the matrix $A = \frac{1}{2}I$. Let $\mu$ denote an invariant measure for weigths $p_1$, $p_2$ and $p_3$. The expected value $EB = (\frac{1}{2}p_2 + \frac{1}{4}p_3, \frac{\sqrt{3}}{4}p_3)$. The matrix $[I - A]^{-1} = 2I$. By the (8) we get $\int_{\mathbb{R}^2} x_1 d\mu = p_2 + \frac{1}{2}p_3$ and $\int_{\mathbb{R}^2} x_2 d\mu = \frac{\sqrt{3}}{2}p_3$. By the (14) we can obtain terms of the matrix $D^2B$ and $D^2X$. In particular

$$\sum_{k=1}^3 p_k e_1^*(b_k) e_2^*(b_k) - \sum_{k=1}^3 p_k e_1^*(b_k) \sum_{k=1}^3 p_k e_2^*(b_k) = \frac{\sqrt{3}}{16} p_3 (p_1 - p_2).$$

By corollary, if $p_1 = p_2$ then the random variables $X_1$, $X_2$ are uncorrelated and

$$\int_{\mathbb{R}^2} x_1 x_2 d\mu = \int_{\mathbb{R}^2} x_1 d\mu \int_{\mathbb{R}^2} x_2 d\mu = (p_2 + \frac{1}{2}p_3) \frac{\sqrt{3}}{2} p_3 = \frac{\sqrt{3}}{4} p_3.$$

(16)

References
[B] M.F Barnsley, *Fractals everywhere*, 2nd edition, Academic Press, Orlando, FL.1993.

[F] K.J Falconer, *Probabilistic methods in fractal geometry*, Progress in Probability, Vol 37(1995)3-13.

[H] J.E Hutchinson, *Fractals and self similarity*, Indiana Univ. Math. J. 30(1981)713-747.

[S] R.S. Strichartz, *Self-similar measures and their Fourier transform III*, Indiana Univ. Math. J. 42(1993) 367-411.

Krzysztof Zajkowski  
Institute of Mathematics, University of Bialystok  
Akademicka 2, 15-267 Bialystok, Poland  
E-mail:kryza@math.uwb.edu.pl