Solving the Inverse Problem for Measures Using Iterated Function Systems: A New Approach
Author(s): B. Forte and E. R. Vrscay
Source: Advances in Applied Probability, Vol. 27, No. 3 (Sep., 1995), pp. 800-820
Published by: Applied Probability Trust
Stable URL: https://www.jstor.org/stable/1428134
Accessed: 22-02-2019 21:40 UTC

REFERENCES
Linked references are available on JSTOR for this article:
You may need to log in to JSTOR to access the linked references.
SOLVING THE INVERSE PROBLEM FOR MEASURES USING ITERATED FUNCTION SYSTEMS: A NEW APPROACH

B. FORTE* AND E. R. VRSCAY,** University of Waterloo

Abstract

We present a systematic method of approximating, to an arbitrary accuracy, a probability measure \( \mu \) on \( x = [0, 1]^q, q \geq 1 \), with invariant measures for iterated function systems by matching its moments. There are two novel features in our treatment. 1. An infinite set of fixed affine contraction maps on \( X \), \( W = \{w_1, w_2, \cdots \} \), subject to an '\( \varepsilon \)-contractivity' condition, is employed. Thus, only an optimization over the associated probabilities \( p_i \) is required. 2. We prove a collage theorem for moments which reduces the moment matching problem to that of minimizing the collage distance between moment vectors. The minimization procedure is a standard quadratic programming problem in the \( p_i \) which can be solved in a finite number of steps. Some numerical calculations for the approximation of measures on \([0,1]\) are presented.

COLLAGE DISTANCE; QUADRATIC PROGRAMMING; DATA COMPRESSION

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 28A80
SECONDARY 41A20; 58A40

1. Introduction

This paper is concerned with the approximation of probability measures on a compact metric space \( X \) by invariant measures for iterated function systems (IFS): systems of contraction mappings on \( X \), \( w = \{w_1, w_2, \cdots, w_N\} \), with associated probabilities \( p = \{p_1, p_2, \cdots, p_N\} \), introduced in [19] and developed further in [4] and [5]. The approximation of measures and functions with IFS and related methods has received much interest, especially in the context of data compression. As reported in the case of image processing [21], it is desirable to be able to represent a target measure or function with a rather small number of IFS parameters, thus achieving a large 'compression factor'.

The inverse problem of measure construction using IFS may be posed as follows:

Let \( (X, d) \) denote a compact metric space and \( \mathcal{M}(X) \) the set of probability measures on \( \mathcal{B}(X) \), the \( \sigma \)-algebra of Borel subsets of \( X \). Then, given a target...
measure \( v \in \mathcal{M}(X) \) and an \( \epsilon > 0 \), find an IFS \((w, p)\) whose invariant measure \( \mu \) satisfies \( d_H(\mu, v) < \epsilon \), where \( d_H \) denotes the Hutchinson metric (defined in Equation (2.1) below).

Most of the work on the inverse problem with IFS (see the papers cited in the list of references, for example) has been based on a knowledge of the moments of the target measure \( v \). Some form of ‘moment matching’ is applied, roughly as follows. Given a target measure \( v \) (with \( X \subset \mathbb{R} \) for simplicity) with moments \( h_k = \int x^k \, dv \), \( k = 1, 2, 3, \ldots \), find an IFS invariant measure \( \mu \) whose respective moments \( g_k = \int x^k \, d\mu \) are ‘close’ to the \( h_k \). In practical applications, moment matching is performed on a finite sequence of moments. For example, given an \( M > 0 \), one minimizes the distance between the vectors \( (g_1, g_2, \ldots, g_M) \) and \( (h_1, h_2, \ldots, h_M) \).

In the case of an IFS with affine maps, the moments \( g_k \) of its invariant measure \( \mu \) may be computed recursively from the coefficients of the affine contraction maps as well as the associated probabilities. Hence moment matching becomes an optimization problem in terms of the IFS parameters.

The method described in this paper yields a systematic algorithm to approximate measures with IFS invariant measures to arbitrary accuracy. Our method differs from previous efforts in two significant aspects:

1. We first begin with an infinite set \( \mathcal{W} = \{w_1, w_2, \ldots\} \) of fixed affine contraction maps \( w_i: X \to X \) which must satisfy an ‘\( \epsilon \)-contractivity condition’. As such, we consider the \( w_i \) to form a basis for the representation of compact subsets of \( X \). From this set we construct sequences of \( N \)-map IFS with probabilities \((w^N, p^N)\). For each such IFS, only an optimization over the probabilities \( p_i^N \), \( i = 1, 2, \ldots, N \) is required. The probabilities \( p_i \) can be loosely considered as ‘Fourier coefficients’ of the basis functions \( w_i \).

2. The moment matching is accomplished by means of a collage theorem for moments. This is in contrast to a minimization of distances between moment vectors of respectively, target and approximating measures as was done, for example, in [24]–[26]. Since the IFS maps are fixed, the minimization of the collage distance in moment space need only be performed with respect to the probabilities \( p_i \). Moreover, the squared collage distance between moment vectors is quadratic in the \( p_i \) and the minimization becomes a quadratic programming problem which can be numerically solved in a finite number of steps. In many cases, the minimum collage distance is achieved on a boundary point of the simplex \( \Pi^N = \{ (p_1, \ldots, p_N) \mid \sum_{i=1}^N p_i = 1 \} \), which means that one or more of the \( p_i \) are zero. In such cases, superfluous maps \( w_i \) are essentially eliminated from the set \( \mathcal{W} \). A density theorem ensures that as \( N \to \infty \), the collage distance in moment space tends to zero.

The layout of this paper is as follows. In Section 2, after a brief glossary of notation, we discuss affine IFS and introduce the idea of an infinite set of contraction maps which satisfy the \( \epsilon \)-contractivity condition mentioned above. In Section 3, we derive the collage theorem for moments and then prove that the method can be used
to approximate measures to arbitrary accuracy. Section 4 contains some applications
and numerical computations. In Section 5, a few final remarks are made.

2. Iterated function systems and their invariant measures

2.1. Notation. In this paper, the following notation will be employed:

$(X, d)$ a compact metric space. (In applications, where $X$ is the ‘base space’ of
the IFS, $X$ will be a compact subset of $\mathbb{R}^n$, e.g. $[0, 1], [0, 1]^2$.)

$C(X) = \{f : X \rightarrow \mathbb{R}, f$ is continuous$\}.$

$Lip(X) = \{f : X \rightarrow \mathbb{R}, |f(x) - f(y)| \leq d(x, y), x, y \in X\}.$

$Con(X) = \{w : X \rightarrow X, d(w(x), w(y)) \leq cd(x, y), \text{for some } c \in [0, 1), \forall x, y \in X\}$:
the set of contraction maps on $X$. We shall refer to $c$ as the contractivity
factor of $w$.

$\mathcal{H}(X)$ the set of non-empty compact subsets of $X$.

$h$ Hausdorff metric on $\mathcal{H}(X)$: Let the distance between a point $x \in X$ and
a set $A \in \mathcal{H}(X)$ be given by

$$d(x, A) = \inf_{y \in A} d(x, y).$$

Then for $A, B \in \mathcal{H}(X),$

$$h(A, B) = \max \left[\sup_{x \in A} d(x, B), \sup_{y \in B} d(y, A)\right].$$

$(\mathcal{H}(X), h)$ is a complete metric space.

$\mathcal{M}(X)$ the set of probability measures on $\mathcal{H}(X)$, the $\sigma$-algebra of Borel subsets
of $X$.

$d_H$ a metric on $\mathcal{M}(X)$, often referred to in the IFS literature as the
Hutchinson metric due to its use in [19]:

$$d_H(\mu, \nu) = \sup_{f \in Lip(X)} \left[\int_X f d\mu - \int_X f d\nu\right], \quad \mu, \nu \in \mathcal{M}(X)$$

$(\mathcal{M}(X), d_H)$ is a complete metric space.

2.2. Affine IFS and infinite sets of contraction maps with $\epsilon$-contractivity. We shall
let $(w, p)$ denote an $N$-map contractive IFS on $X$ with probabilities, that is, a set of
$N$ contraction maps, $w = (w_1, w_2, \ldots, w_N)$, $w_i \in Con(X)$, with associated prob-
babilities $p = (p_1, p_2, \ldots, p_N)$, $p_i \geq 0$, $i = 1, 2, \ldots, N$, and $\sum_{i=1}^N p_i = 1$. The contract-
vity factor of the IFS is defined as

$$c = \max_{1 \leq i \leq N} c_i < 1.$$ 

As usual [4], [5], [19], define a set-valued mapping $\hat{w} : \mathcal{H}(X) \rightarrow \mathcal{H}(X)$ as follows.
For a subset $S \in \mathcal{H}(X)$ denote $w_i(S) = \{w_i(x), x \in S\}, i = 1, 2, \ldots, N$ and let the action of $\hat{w}$ on $\mathcal{H}(X)$ be given by

$$\hat{w}(S) = \bigcup_{i=1}^{N} w_i(S).$$

Then there exists a unique compact set $A \in \mathcal{H}(X)$, the attractor of $w$ (independent of $p$), such that

$$A = \hat{w}(A) = \bigcup_{i=1}^{N} w_i(A)$$

and $h(\hat{w}^n(S), A) \to 0$ as $n \to \infty$ for all $S \in \mathcal{H}(X)$. Now define an operator $M: M(X) \to M(X)$ (often called the 'Markov operator') as follows. For $\mu \in M(X)$, let

$$M\mu = \sum_{i=1}^{N} p_i \mu \circ w_i^{-1}.$$ 

In [19] it was shown that $M$ is a contraction mapping on $(M(X), d_H)$: For all $\mu, \nu \in M(X)$, $d_H(M\mu, M\nu) \leq c d_H(\mu, \nu)$. Thus there exists a unique measure $\bar{\mu} \in M(X)$, the invariant measure of the IFS, for which $M\bar{\mu} = \bar{\mu}$.

In this paper, we consider only IFS with affine maps on $X$. For example, on $\mathbb{R}$ these maps have the general form

$$w_i(x) = s_i x + a_i, \quad |s_i| < 1, \quad s_i, a_i \in \mathbb{R}, \quad i = 1, 2, \ldots, N.$$ 

That there is no loss of generality in this restriction is shown by the following theorem, proved in [10].

**Theorem 2.1.** Let $(X, d)$ denote a compact metric space and $\mathcal{M}_{\text{IFS}}(X) \subset \mathcal{M}(X)$ the subset of invariant measures of affine IFS on $X$. Then $\mathcal{M}_{\text{IFS}}(X)$ is dense in $(\mathcal{M}(X), d_H)$.

Theorem 2.1 is a rather trivial consequence of the following result.

**Proposition 2.2** [23]. Let $(X, d)$ be a compact metric space and $\mathcal{M}_{f}(X) = \{\mu \in \mathcal{M}(X) \mid \mu \text{ has finite support}\}$. Then $\mathcal{M}_{f}(X)$ is dense in $\mathcal{M}(X)$.

In our formal solution to the inverse problem, we shall be constructing sequences of $N$-map IFS with probabilities, denoted as $(w^N, p^N)$, with $N \to \infty$, where the IFS maps in $w^N$ are chosen from a fixed, infinite set $\mathcal{W}$ of contraction maps. A condition must be placed on this set, according to the following definition.

**Definition 2.3.** An infinite set of contraction maps $\mathcal{W} = \{w_1, w_2, \ldots\}$, $w_i \in \text{Con}(X)$ is said to satisfy an $\epsilon$-contractivity condition on $X$ if:

for each $x \in X$ and any $\epsilon > 0$, there exists an $i^* \in \{1, 2, \ldots\}$ such that
$w_i(x) = \frac{1}{2^i} [x + j - 1], \quad i = 1, 2, \cdots, j = 1, 2, \cdots, 2^i.$

For each $i^* \geq 1$, the set of maps $\{w_{i^*j}, j = 1, 2, \cdots, 2^{i^*}\}$ provides $2^{-i^*}$-contractions of $[0, 1]$ in an obvious way. Another possible set of functions is given by

$w_i(x) = \frac{1}{i} [x + j - 1], \quad i = 2, \cdots, j = 2, \cdots, i.$

In either of the above cases, for a given $i^* \geq 1$, the sets $\{w_{i^*j}(x), j = 1, 2, \cdots, j_{\max}\}$ overlap with each other only at single points. This is not necessary and, in special cases, it might be advantageous to let these sets overlap on subintervals of $X$.

2.2. Moment relations for affine IFS. A primary motivation for the use of affine IFS maps is the simplicity of relations involving moments of probability measures.

Given an $N$-map IFS $(w, p)$ with associated Markov operator $M$, let $\mu \in \mathcal{M}(X)$ and $\nu = M\mu$. Then from Equation (2.5), for any continuous function $f : X \to \mathbb{R}$,

$$
\int_X f(x) \, d\nu(x) = \int_X f(x) \, d(M\mu)(x)
= \sum_{i=1}^{N} p_i \int_X (f \circ w_i)(x) \, d\mu(x).
$$

For the remainder of this section, we consider only $X \subset \mathbb{R}$. Let the moments of $\mu$ and $\nu$ be denoted by

$$
\begin{align*}
g_n &= \int_X x^n \, d\mu, \quad h_n = \int_X x^n \, d\nu, \quad n = 0, 1, 2, \cdots,
\end{align*}
$$

where $g_0 = h_0 = 1$. For affine IFS maps on $\mathbb{R}$, set $f(x) = x^n$ in Equation (2.8) to give

$$
\begin{align*}
h_n &= \sum_{k=0}^{n} \binom{n}{k} \left[ \sum_{i=1}^{N} p_i a_i^{n-k} \right] g_k, \quad i = 1, 2, \cdots.
\end{align*}
$$

Now define

$$
D(X) = \left\{ g = (g_0, g_1, g_2, \cdots) \mid g_n = \int_X x^n \, d\mu, \quad n = 0, 1, 2, \cdots, \mu \in \mathcal{M}(X) \right\}.
$$
Solving the inverse problem for measures

i.e. the set of all (infinite) moment vectors for probability measures in $\mathcal{M}(X)$. Then for each $N$-map affine IFS $(w, p)$ on $(X, d)$, i.e. each Markov operator $M: \mathcal{M}(X) \to \mathcal{M}(X)$, there corresponds a linear operator $A: D(X) \to D(X)$. In the standard basis $\{\varepsilon_i = (0, 0, \cdots, 0, 1, 0, \cdots)\}_{i=0}^{\infty}$, $A$ is represented by a lower-triangular matrix.

In the case where $\mu = \nu = M\mu$, i.e. $\mu$ is the invariant measure of the IFS, then $h_n = g_n$, $n = 0, 1, 2, \cdots$, and $g \in D(X)$ is the fixed point of $A$. A rearrangement of Equation (2.10) produces the following well-known relation:

\begin{equation}
(2.12) \quad \left[1 - \sum_{i=1}^{N} p_i \delta_i^n\right] g_n = \sum_{k=0}^{n-1} \binom{n}{k} \left[\sum_{i=1}^{N} p_i \delta_i^k a_i^{n-k}\right] g_k, \quad n = 1, 2, \cdots
\end{equation}

which allows the moments $g_n$ to be computed recursively in terms of the IFS parameters $s_i, a_i, p_i$.

Finally, if any of the IFS maps $w_i$ in $w$ are polynomial of degree $\geq 2$, it is not difficult to see that the linear operator $A: D(X) \to D(X)$ is not represented by a lower-triangular matrix. As well, the relations for moments of the IFS invariant measure, unlike Equation (2.12), are not complete, and the moments may not be computed recursively.

3. A collage theorem for moments and the inverse problem

Moment matching for the approximation of measures on $[0, 1]^q$, $q = 1, 2, \cdots$, can be justified by the fact that the convergence of moments is equivalent to the weak convergence of measures. Since we are working on compact spaces, the latter convergence is equivalent to convergence in Hutchinson metric $d_H$. This is summarized in the following theorem which is formally proved in [10].

**Theorem 3.1.** For $X = [0, 1]$, let $\mu$ and $\mu^{(j)} \in \mathcal{M}(X)$, $j = 1, 2, 3, \cdots$, with power moments defined as follows:

\begin{align*}
g_n &= \int_X x^n d\mu, \\
g_n^{(j)} &= \int_X x^n d\mu^{(j)}, \quad n = 0, 1, 2, \cdots
\end{align*}

Then the following are equivalent:

(i) $g_n^{(j)} \to g_n$ as $j \to \infty$, $\forall k$,

(ii) the sequence of measures $\mu^{(j)}$ converges weak* to $\mu$, i.e. for any $f \in C(X)$, $\int f d\mu^{(j)} \to \int f d\mu$, as $j \to \infty$,

(iii) $d_H(\mu^{(j)}, \mu) \to 0$ as $j \to \infty$.

The results of this theorem can easily be extended to $X = [0, 1]^q$, $q \geq 2$.

The idea of using IFS and moment matching for the inverse problem of fractal/measure construction was first suggested in [5], Section 3.3, where the method was applied to a 2-map affine IFS in the complex plane. In the case $X = [0, 1]$ and a target measure $\nu \in \mathcal{M}(X)$ with moments $h_n = \int_X x^n d\nu$, Diaconis...
and Shahshahani [13] proposed the following method. For a fixed number $N > 0$ of affine IFS maps on $\mathbb{R}$ with probabilities, cf. Equation (2.6), and a given number $M > 0$ of moments to be matched, impose the conditions

$$g_n(s, a, p) = h_n, \quad n = 1, 2, \ldots, M,$$

and use the moment recursion relations in Equation (2.12) to solve for the IFS parameters $s_i, a_i, p_i$ directly. However, the $g_n$ are complicated non-linear functions of the IFS parameters and approximation schemes such as the Newton-Kantorovich method are unstable. In [24]–[26] moment matching was performed by minimizing the following truncated $l^2$ distance between target and IFS moment vectors:

$$D^N_M(s, a, p) = \sum_{n=1}^{M} (g_n(s, a, p) - h_n)^2.$$

Since the $g_n$ are differentiable with respect to the IFS parameters $s_i, a_i, p_i$, gradient methods for optimization can be used. This method was also tested for target measures or images in $\mathbb{R}^2$. In the one-dimensional case, the method works reasonably well, although a considerable amount of computation may be required for the optimization of the $3N$ IFS parameters. These difficulties are further enhanced in the two-dimensional case. As well, the graph of the function $D^N_M$ can be very complicated, especially as $N$ or $M$ increases. Local methods such as gradient schemes are not guaranteed to converge to global minima or even reasonable minima.

The modified moment matching approach which we now outline represents a significant improvement because of two major changes:

(i) The IFS maps $w_i$, hence the parameters $s_i$ and $a_i$, $i = 1, 2, \ldots, N$, are fixed. Moment matching is done only with respect to the $p_i$.

(ii) Instead of using Equation (3.2), which involves complicated expressions of the IFS moments $g_n$ in terms of the probabilities $p_i$, we use a collage distance for moments which involves only quadratic terms in the $p_i$.

Let us now recall a simple consequence of the Banach fixed point theorem which, in the IFS literature, is referred to as the collage theorem [4], [7].

**Theorem 3.2 (Collage theorem).** Let $(Y, d_Y)$ be a complete metric space. Given a $y \in Y$, suppose that there exists a map $f \in \text{Con}(Y)$ with contractivity factor $0 \leq c < 1$, so that $d_Y(y, f(y)) < \varepsilon$. If $\bar{y}$ is the fixed point of $f$, i.e. $f(\bar{y}) = \bar{y}$, then $d_Y(\bar{y}, y) < \varepsilon/(1 - c)$.

In other words, suppose there exists a ‘target’ $y$ that we wish to approximate with a fixed point $\bar{y}$ of an unknown mapping $f$. The inverse problem reduces to finding an $f$ which minimizes the collage distance $d_Y(y, f(y))$. This idea was first used for the geometric approximation of sets with IFS attractors [4], [7] as well as for more generalized IFS-type methods used for image representation [21]. The inverse problem for measures using IFS may now be posed as follows:
Solving the inverse problem for measures

given a target measure \( v \in \mathcal{M}(X) \) and a \( \delta > 0 \), find an IFS \((w, p)\) with associated Markov operator \( M \) such that \( d_{\mathcal{H}}(v, Mv) < \delta \).

For ease of presentation, unless otherwise indicated, \( X = [0, 1] \) is the IFS base space for the remainder of this section. The extension to \([0, 1]^q, q \geq 2\) is straightforward.

Now define

\[
P^2(N) = \left\{ \mathbf{c} = (c_0, c_1, c_2, \cdots) \mid c_i \in \mathbb{R}, \right\}
\]

\[
\|\mathbf{c}\|_2^2 = \sum_{k=0}^{\infty} c_k^2 < \infty \}
\]

As well, define the following weighted Banach space of half-infinite sequences:

\[
\widetilde{T}^2(N) = \left\{ \mathbf{c} = (c_0, c_1, c_2, \cdots) \mid c_i \in \mathbb{R}, \right\}
\]

\[
|\mathbf{c}|_{\tilde{H}}^2 = c_0^2 + \sum_{k=1}^{\infty} \frac{1}{k^2} c_k^2 < \infty \}
\]

We shall consider, in particular, the following subset:

\[
\tilde{T}_0^2(N) = \{ \mathbf{c} \in \tilde{T}^2(N) \mid c_0 = 1 \} \subset \tilde{T}^2(N).
\]

Now recall \( D(X) \), the set of moment vectors for all \( \mu \in \mathcal{M}(X) \), defined in Equation (2.11). Clearly, \( D(X) \subset \tilde{T}_0^2(N) \).

**Proposition 3.3.** Let \( X = [0, 1] \) and \( \mu, v^{(n)} \in \mathcal{M}(X), \ n = 1, 2, 3, \cdots \), with associated moment vectors \( g, g^{(n)} \in D(X) \). Then \( \|g - g^{(n)}\|_{\tilde{H}} \to 0 \) as \( n \to \infty \) iff \( d_{\mathcal{H}}(\mu, v^{(n)}) \to 0 \) as \( n \to \infty \).

**Proof.** The proof follows from the results of Theorem 3.1.

**Proposition 3.4.** Let \( X = [0, 1] \). Define the following metric on \( D(X) \): for \( u, v \in D(X) \), \( d_2(u, v) = \|u - v\|_{\tilde{H}} \). Then \((D(X), d_2)\) is a complete metric space.

The proof of this proposition is given in the Appendix.

Recall that for each Markov operator \( M : \mathcal{M}(X) \to \mathcal{M}(X) \) associated with an \( N \)-map IFS \((w, p)\), there exists a linear operator \( A : D(X) \to D(X) \), whose action is given in Equation (2.10).

**Proposition 3.5.** The linear operator \( A \) is contractive in \((D(X), \tilde{d}_2)\).

**Proof.** In the standard basis \( \{e_i = (0, 0, \cdots, 0, 1, 0, \cdots)\}_{i=0}^{\infty} \), the (infinite) matrix representation of \( A \) is lower triangular. Hence, \( A \) has eigenvalues

\[
\lambda_0 = a_{00} = 1, \quad \lambda_n = a_{nn} = \sum_{i=1}^{N} p_i s_i^p, \quad n \geq 1.
\]
Thus, $|\lambda_n| = |a_{nn}| < c^n < 1$ for $n \geq 1$. A little algebra shows that for any $u, v \in D(X)$, $\|A(u - v)\|_p \leq c \|u - v\|_p$, which implies the contractivity of $A$.

**Corollary 3.6.** The operator $A$ has a unique (attractive) fixed point $\bar{g} \in D(X)$.

The components $g_n$ of $\bar{g}$ are the moments of $\bar{\mu}$, the invariant measure of the IFS $(w, p)$, cf. Equation (2.12). We have now arrived at the major result of this section.

**Corollary 3.7 (collage theorem for moments).** Let $(X, d)$ be a compact metric space and $\mu \in \mathcal{M}(X)$ with moment vector $g \in D(X)$. Let $(w, p)$ be an $N$-map IFS, with contractivity factor $0 \leq c < 1$, such that $\overline{d}_2(g, h) = \|g - h\|_p < \epsilon$, where $h \in D(X)$ is the moment vector corresponding to $\nu = M\mu$. Then

$$\overline{d}_2(g, \bar{g}) < \frac{\epsilon}{1 - c},$$

where $\bar{g}$ is the moment vector corresponding to $\bar{\mu}$, the invariant measure of the IFS $(w, p)$.

Thus, given a target measure $\mu$ with moment vector $g$, the inverse problem becomes one of finding an IFS $(w, p)$ such that the collage distance $\overline{d}_2(g, h)$, where $h = Ag$, is small. Our main result—Theorem 3.9 below—ensures that for IFS constructed from a set $\mathcal{W}$ which satisfies the $\epsilon$-contractivity property the collage distance may be made arbitrarily small. For its proof, we shall make use of Proposition 2.2 and the following result.

**Proposition 3.8 [12].** Let $(X, d)$ be a compact metric space and let $(u, p)$ denote an $N$-map IFS on $(X, d)$ with contractivity factor $c_u$ and invariant measure $\nu \in \mathcal{M}(X)$. Given an $\epsilon > 0$, suppose that there exists another $N$-map IFS, $(v, p)$, with identical probabilities, such that

$$d_H(u, v) = \max_{1 \leq i \leq N} \sup_{x \in X} d(u_i(x), v_i(x)) < (1 - c_u).$$

Then $d_H(\nu, \nu) < \epsilon$, where $\nu \in \mathcal{M}(X)$ is the invariant measure of $(v, p)$.

Now let $\mathcal{W} = \{w_1, w_2, \cdots\}$ be an infinite set of affine contraction maps on $X = [0, 1]$ which satisfies the $\epsilon$-contractivity condition. Let

$$w^N = \{w_1, w_2, \cdots, w_N\}, \quad N = 1, 2, \cdots,$$

denote $N$-map truncations of $\mathcal{W}$. As well, let

$$\Pi^N = \left\{ p^N = (p_1, p_2, \cdots, p_N) \mid p_i \geq 0, \sum_{i=1}^{N} p_i = 1 \right\}$$

denote the set of all probability $N$-vectors for $w^N$. Note that $\Pi^N \subset \mathbb{R}^N$ is compact in the natural topology on $\mathbb{R}^N$. Now let $\mu \in \mathcal{M}(X)$ be a target measure with moment vector $g \in D(X)$. For a $p^N \in \Pi^N$, let $M^N$ be the Markov operator corresponding to
the $N$-map IFS $(\mathbf{w}^N, \mathbf{p}^N)$. Also let $\nu_N = M^N \mu$, with associated moment vector $\mathbf{h}_N \in D(X)$. The collage distance between the moment vectors of $\mu$ and $\nu_N$ will be denoted as

$$\Delta^N(\mathbf{p}^N) = \| \mathbf{g} - \mathbf{h}_N \|_2.$$  

Since $\Delta^N : \Pi^N \rightarrow \mathbb{R}$ is continuous, it attains an absolute minimum value, to be denoted as $\Delta^N_{\text{min}}$, on $\Pi^N$. The following theorem ensures that the collage distance may be made arbitrarily small.

**Theorem 3.9.** $\Delta^N_{\text{min}} \rightarrow 0$ as $N \rightarrow \infty$. 

**Proof.** We first show that $\Delta^N_{\text{min}}$ is non-increasing with respect to $N$, i.e. $\Delta^N_{\text{min}} \leq \Delta^{N+1}_{\text{min}}$ for $n_1 > n_2 \geq 1$.

Let $q^N = (q^N_1, q^N_2, \ldots, q^N_N) \in \Pi^N$ be an absolute minimum point for $\Delta^N$, $N \geq 2$. Also let $\mathbf{p}^N \in \Pi^N$ with the restriction (without loss of generality) $\mathbf{p}^N = 0$. Then $\Delta^N(\mathbf{p}^N) = \Delta^N(q^N) = D^N_{\text{min}}$. Now let $q^{N-1} = (q^{N-1}_1, \ldots, q^{N-1}_{N-1}) \in \Pi^{N-1}$ be an absolute minimum for $\Delta^{N-1}$ and set $p^N_i = q^{N-1}_i$, $i = 1, 2, \ldots, N-1$. Then $\Delta^{N-1}_{\text{min}} \geq \Delta^N(\mathbf{p}^N) \geq \Delta^{N-1}_{\text{min}}$. Thus $\{\Delta^N_{\text{min}}\}_{N=1}^\infty$ is a non-increasing sequence of non-negative numbers. Hence, there exists a limit, $L \geq 0$, of this sequence. We now show that $L = 0$.

If $\lim_{N \rightarrow \infty} d_H(\mu, M^N \mu) = 0$, for one sequence of finite IFS $\{(\mathbf{w}^N, \mathbf{p}^N)\}$, where $\mathbf{p} \in \Pi^N$, $N = 1, 2, \ldots$, then it follows from Theorem 3.1 that $\Delta^N_{\text{min}} \rightarrow 0$. Now, from Proposition 2.2, given any $\epsilon_1 > 0$, there exists a $\mu_f \in M_f(X)$, such that $d_H(\mu, \mu_f) < \epsilon_1$. Then $\mu_f = \sum_{i=1}^{n_f} \alpha_i \delta_{x_i}$ for some $n_f \geq 1$, where $\alpha_i > 0$, $1 \leq i \leq n_f$ and $\sum_{i=1}^{n_f} \alpha_i = 1$: Here, $\delta_y$ denotes a point-mass measure at $y \in X$. Note that $\mu_f$ is the invariant measure for the $n_f$-map IFS $(\mathbf{w}, \mathbf{p})$, where

$$u_i(x) = \bar{x}_i, \quad p_i = \alpha_i, \quad 1 \leq i \leq n_f.$$ 

The contractivity factor of this IFS is $c_u = 0$.

For any $\epsilon_2 > 0$, let $U_i = N_{\epsilon_2}(\bar{x}_i)$, the $\epsilon_2$-neighbourhood of $\bar{x}_i$, $1 \leq i \leq n_f$. From our refinement assumption, there exist affine maps $v_i = w_{k_i} \in \mathcal{W}$, $1 \leq i \leq n_f$, such that $v_i(X) \subseteq U_i$. Now let $(\mathbf{v}, \mathbf{p})$ be the $n_f$-map IFS,

$$v_i(x) = w_{k_i}(x) = \beta_i x + \gamma_i, \quad p_i = \alpha_i, \quad 1 \leq i \leq n_f.$$ 

The contractivity factor of this IFS is $\beta = \max_{1 \leq i \leq n_f} |\beta_i| < 2\epsilon_2$. Denote its Markov operator as $M_f$ and its invariant measure as $\nu_f$, i.e. $M_f \nu_f = \nu_f$. From the above construction, we have $\sup_{x \in X} d(u_i(x), v_i(x)) < 2\epsilon_2$, $1 \leq i \leq n_f$. Hence, from Proposition 3.8, with $c_u = 0$, we have

$$d_H(\mu_f, \nu_f) < 2\epsilon_2.$$ 

Let $N$ be the smallest integer for which $\{v^N_i\}_{i=1}^{n_f} \in \mathbf{w}^N$. Also let $M^N$ be the Markov
operator for the IFS \((w^N, p^N)\), where \(p_k = \alpha_i, 1 \leq i \leq n_f\), and all other \(p_j = 0\). Then \(M^N = M_f\). Now, given the target measure \(\mu \in \mathcal{M}(X)\), consider the inequality

\[
(3.14) \quad d_H(\mu, M^N\mu) \leq d_H(\mu, \mu_f) + d_H(\mu_f, \nu_f) + d_H(\nu_f, M^N\mu).
\]

Note that

\[
(3.15) \quad d_H(\nu_f, M^N\mu) = d_H(M^N\nu_f, M^N\mu)
\]

\[
\leq \beta d_H(\nu_f, \mu)
\]

\[
\leq 2\epsilon^2[d_H(\nu_f, \mu_f) + d_H(\mu_f, \mu)].
\]

Substitution of inequality (3.15) into inequality (3.14) yields

\[
(1 + 2\epsilon^2)(1 + 2\epsilon^2)
\]

Given \(\epsilon = 1/2^n\), we find \(\epsilon_1, \epsilon_2 > 0\) such that \((1 + 2\epsilon^2)(\epsilon_1 + 2\epsilon_2) < 1/2^n\) and a finite IFS \((w^{N_n}, p^{N_n})\) for which \(d_H(\mu, M^N\mu) < 1/2^n\). Hence \(\lim_{n \to \infty} \inf d_H(\mu, M^N\mu) = 0\). Thus, \(L = 0\) and the theorem is proved.

Remarks

1. Theorem 3.9 is a density result establishing that the set of invariant measures for all \(N\)-map IFS \((w^N, p^N)\), where \(p^N \in \Pi^N, N = 1, 2, \cdots\), is dense in \((\mathcal{M}(X), d_H)\). This result can be extended to \([0, 1]^q, q \geq 2\).

2. Although not explicitly stated in the proof, the collage distances \(\Delta^N\) and, in particular, the sequence \(\Delta^N_{\text{min}}\), are also dependent on the ordering of the \(w_i\) maps in the infinite set \(\mathcal{W}\). However, at this point, we are not interested in any questions about the 'optimal' ordering of the maps in \(\mathcal{W}\) nor how \(N\)-map subsets \(w^N\) should be chosen.

3.1. The inverse problem as a quadratic programming problem. Let us now consider the square of the collage distance, cf. Equation (3.10), between the moment vector \(g\) of the target measure \(\mu \in \mathcal{M}(X)\) and the moment vector \(h_N\) of the measure \(\nu_N = M^N\mu\), where, as above, \(M^N\) is the Markov operator associated with the truncated IFS \((w^N, p^N)\):

\[
(3.16) \quad S^N(p^N) = (\Delta^N)^2(p^N) = \sum_{n=1}^{\infty} \frac{1}{n^2} (h_n - g_n)^2.
\]
Let $A^N: D(X) \to D(X)$ denote the linear operator associated with $M^N$. Then $h_N = A^N g$ and from Equation (2.10),

\begin{equation}
\begin{aligned}
    h_n = \sum_{i=1}^{N} A_{ni}^N p_i^N, \quad n = 1, 2, 3, \ldots ,
\end{aligned}
\end{equation}

where

\begin{equation}
\begin{aligned}
    A_{ni}^N &= \int_{X} (w_i x + a_i)^n d\mu \\
    &= \sum_{k=0}^{n} \binom{n}{k} a_i^n g_k.
\end{aligned}
\end{equation}

The function $S^N(p^N)$ may be written in the following form:

\begin{equation}
\begin{aligned}
    S^N(x) = x^T Q x + b^T x + c, \quad x \in \mathbb{R}^N,
\end{aligned}
\end{equation}

where $x = p^N = (p_1^N, p_2^N, \ldots, p_N^N)$. The elements of the symmetric matrix $Q$ are given by

\begin{equation}
\begin{aligned}
    q_{ij} &= \sum_{n=1}^{\infty} \frac{1}{n^2} A_{ni}^N A_{nj}^N, \quad i, j \in \{1, 2, \ldots, N\}.
\end{aligned}
\end{equation}

As well,

\begin{equation}
\begin{aligned}
    b_i &= -2 \sum_{n=1}^{\infty} \frac{1}{n^2} g_n A_{ni}^N, \quad i = 1, 2, \ldots , N
\end{aligned}
\end{equation}

and

\begin{equation}
\begin{aligned}
    c &= \sum_{n=1}^{M} \frac{g_n^2}{n^2}.
\end{aligned}
\end{equation}

Since $0 \leq A_{ni}^N \leq 1$, the infinite sums in Equations (3.20) and (3.21) converge. The minimization of the moment collage distance $S^N(x)$ becomes the following standard quadratic programming problem with linear constraints:

\begin{equation}
\begin{aligned}
    \text{minimize } S^N(x), \quad \sum_{i=1}^{N} x_i = 1, \quad x_i \geq 0.
\end{aligned}
\end{equation}

This represents a significant simplification of the moment matching problem since quadratic programming problems can be solved computationally in a finite number of steps.
4. Applications and numerical computations

In the applications to be described below, $X = [0, 1]$. The moments of the target measure $\mu$ are denoted by $\{g_n\}_{n=0}^\infty$. The approach begins with the selection of an infinite set $W$ of affine IFS maps, cf. Equation (2.6), satisfying the refinement condition of Definition 2.3. We then truncate this set to produce an $N$-map IFS $(w^N, p^N)$ and solve the quadratic programming problem in Equation (3.23). In practical calculations, it is possible to match only a finite number of moments, so we consider the minimization of the following function:

$$S_N^N(p^N) = (\Delta_M^N)^2(p^N)$$

subject to the linear constraints in Equation (3.23).

The minimization of the function $S_N^N(x)$ in Equation (4.1) was performed with a quadratic programming (QP) algorithm developed by Best and Ritter [9]. (See the acknowledgements at the end of this paper.) The QP method is superior to gradient projection (GP) methods of minimizing our objective function (such as the Davidon method employed in [24]–[26]) for the following reasons. 1. QP locates the minimum of $S^N$ on the simplex $\Pi^N$ in a finite number of steps, whereas GP converges only to a local minimum of $S^N$ and is sensitive to the initial point from where the search begins. Furthermore, the convergence of GP may be extremely slow, especially when the graph of the objective function is quite flat near a minimum. 2. In many of the problems we have studied, the minimum of $S^N$ is achieved on a boundary point of $\Pi^N$, which implies that one or more probabilities $p_i^N$ are zero. This, in turn, implies that the IFS maps $w_i$ associated with these vanishing probabilities are superfluous. QP essentially discards these maps. In general, we have found that GP rarely converges to such a minimum on the boundary of $\Pi^N$. As a result, many of these superfluous IFS maps are kept in the set.

We show below not only the minimum (truncated) collage distances $\Delta_M^N$ achieved for a particular truncation $(w^N, p^N)$ but also the following (truncated) distances in $D(X)$:

$$\Gamma_M^N = \left[ \sum_{n=1}^M \frac{1}{n^2} (\tilde{g}_n - g_n)^2 \right]^{1/2},$$

where, as above, $g$ denotes the moment vector of the target measure $\mu$ and $\tilde{g}_N$ denotes the moment vector of the invariant measure $\tilde{\mu}_N$ of the IFS $(w^N, p^N)$. In the limit $M \to \infty$, it follows from the collage theorem for moments, cf. Equation (3.6), that

$$\Gamma^N \equiv \bar{d}_2(g, \tilde{g}) < \frac{1}{1 - c} \Delta^N,$$

where $c$ is the contractivity factor of the IFS $(w^N, p^N)$. 

This content downloaded from 129.97.58.73 on Fri, 22 Feb 2019 21:40:13 UTC
All use subject to https://about.jstor.org/terms
In the following calculations, we used the ‘wavelet-type’ affine maps of Equation (2.7). The truncated IFS map vectors \( w^N \) were constructed by arranging the \( w_{ij} \) maps as follows:

\[
W_{1,1}, W_{1,2}, \cdots, W_{2,4}, W_{3,1}, \cdots, W_{3,8}, \cdots
\]

The vector \( w^{N(i*)} \), where \( N(i*) = \sum_{i=1}^{i*} 2^i \), contains affine maps \( w_{ij} \) with contraction factors \( 2^{-i} \), \( i = 1, 2, \cdots, i* \). (For \( i* = 1, 2, 3, 4 \), \( N(i*) = 2, 6, 14, 30 \), respectively.) In all cases, the contractivity factor of \( w^N \) is \( c = \frac{1}{2} \). This appears to be a natural ordering of the maps since the refinement afforded by the IFS \( w^{N(i*)} \) increases with \( i* \).

4.1. A simple target measure with continuous density. We first consider a target measure with continuous probability density function, namely, \( \rho(x) = 6x(1 - x) \). The moments of this measure are

\[
g_n = \int_0^1 x^n \rho(x) \, dx = \frac{6}{(n+2)(n+3)}, \quad n = 0, 1, 2, \cdots.
\]

The convergence of IFS measures to the target measure will be demonstrated not only in terms of moments but also with regard to convergence to the distribution function \( F(x) \), defined as

\[
F(x) = \int_0^x d\mu(t) = \int_0^x \rho(t) \, dt.
\]

In this case, \( F(x) = x^2(3 - 2x) \). The distribution function corresponding to the invariant measure \( \bar{\mu}_N \) of the IFS \( (w^N, p^N) \) will be denoted as \( \bar{F}_N(x) \), i.e.

\[
\bar{F}_N(x) = \int_0^x d\bar{\mu}_N(t).
\]

In order to compute \( \bar{F}_N(x) \), we generate a discrete approximation \( \bar{\mu}_N^K \) to \( \bar{\mu}_N \) on subintervals \( I_k \), \( k = 1, 2, \cdots, K \), formed by the equipartition on \([0, 1]\) generated by the points \( x_i = i/K, i = 0, 1, 2, \cdots, K \). In these calculations \( K = 1000 \). (For further details of this procedure of obtaining a discrete measure, see [25].) The discrete measure is represented by an array \( \bar{M}_k \), \( k = 1, 2, \cdots, K \), where \( \bar{M}_k = \bar{\mu}_N^K(I_k) \). Then

\[
F(x_k) = \sum_{i=1}^k \bar{M}_k.
\]

In Table 1, we summarize the results of our minimization procedure, using \( M = 30 \) moments. The collage distance \( \Delta^N_M \) as well as the distance \( \Gamma^N_M \) between moments of the target measure and the approximating IFS measure are given. In each case, we list \( N \), the number of maps in the truncation \( w^N \) over which the QP optimization was performed, as well as \( N_{\text{QP}} \leq N \), the actual number of non-zero probabilities at the minimum point in \( \Pi^N \). For purposes of comparison, we also list the minimum collage distances \( \Delta^N_M \) yielded by GP as well as \( N_{\text{GP}} \), the number of non-zero probabilities at the minimum. For small values of \( N \), the GP results are consistent with QP. However, as \( N \) increases, the GP method does not converge to minima found by
TABLE 1
Results of moment matching via minimization of the collage distance $S_M^N$ in Equation (4.1) to approximate the measure with probability density function $\rho(x) = 6x(1-x)$. The ‘wavelet-type’ IFS maps of Equation (2.8) were used and the truncated IFS map vectors $w_N$ were constructed according to (4.4) in the text. Columns 2-4 show the results for minimization via quadratic programming (QP). Rows which are missing correspond to values of $N$ for which the QP method produced no decrease in the collage distance $D_M^N$, i.e. the probability assigned to map $w_N$ was zero. The value $N_{QP}$ denotes the number of maps in $w_N$ with non-zero probabilities. The column under $w_{QP}$ lists the indices of those maps. Results obtained from the GP method of minimization are shown in columns 5 and 6, for comparison. $N_{GP}$ denotes the number of maps with non-zero probabilities as determined by the GP method.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$D_M^N$</th>
<th>$\Gamma_M^N$</th>
<th>$N_{QP}$</th>
<th>$w_{QP}$</th>
<th>$D_M^N$</th>
<th>$N_{GP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>$2.13 \times 10^{-2}$</td>
<td>$3.12 \times 10^{-2}$</td>
<td>2</td>
<td>(1, 2)</td>
<td>$2.13 \times 10^{-2}$</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>$3.58 \times 10^{-3}$</td>
<td>$4.08 \times 10^{-3}$</td>
<td>3</td>
<td>(2, 4)</td>
<td>$3.58 \times 10^{-2}$</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>$7.72 \times 10^{-5}$</td>
<td>$7.72 \times 10^{-5}$</td>
<td>5</td>
<td>(1-5)</td>
<td>$7.72 \times 10^{-5}$</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>$7.60 \times 10^{-5}$</td>
<td>$7.59 \times 10^{-5}$</td>
<td>5</td>
<td>(1, 2, 4, 5, 7)</td>
<td>$2.13 \times 10^{-5}$</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>$6.89 \times 10^{-5}$</td>
<td>$6.74 \times 10^{-5}$</td>
<td>6</td>
<td>(2-5, 7, 9)</td>
<td>$6.90 \times 10^{-5}$</td>
<td>7</td>
</tr>
<tr>
<td>12</td>
<td>$3.66 \times 10^{-5}$</td>
<td>$3.52 \times 10^{-5}$</td>
<td>6</td>
<td>(2-4, 8, 10, 12)</td>
<td>$3.66 \times 10^{-5}$</td>
<td>10</td>
</tr>
<tr>
<td>13</td>
<td>$1.05 \times 10^{-6}$</td>
<td>$1.03 \times 10^{-6}$</td>
<td>8</td>
<td>(2-5, 8, 10, 11, 13)</td>
<td>$3.28 \times 10^{-6}$</td>
<td>13</td>
</tr>
</tbody>
</table>

QP to lie on the boundary of $\Pi^N$. As a result, (i) more IFS maps are required for the approximation of the target measure and (ii) the accuracy of the approximation, in terms of moment distance, is poorer, especially as $N$ gets larger.

In Figure 1 are presented some approximations to the distribution function $F(x) = x^2(3 - 2x)$ yielded by the IFS invariant measures $\tilde{\mu}_N$. The convergence of the $\tilde{F}_N(x)$ to $F(x)$ with increasing $N$ is evident.

4.2. IFS reconstruction of the spectral measure of an FCC crystal lattice. We now apply our approximation method to a typical problem from theoretical physics which requires the computation of integrals over a measure which is not explicitly known. In such problems, the moments of these measures are usually available and a standard approach is to employ Padé approximants to reconstruct the measures.

Bessis and Demko [8] first showed that IFS invariant measures could also be used to approximate the measure for a specific problem involving crystal lattices. The problem is to determine thermodynamic averages of crystal models as integrals over distributions. The simple model which they studied was the face-centered cubic (FCC) lattice. The zero-point energy of this lattice is given by the integral

$$u_0 = \frac{1}{2} \int_0^1 \sqrt{x} G(x) \, dx,$$

where $G(x) \, dx$ is the fraction of vibrational modes in the interval $[x, x + dx]$. The function $G(x)$ is not known in closed form. However, the moments $g_n$ of $G(x)$ were
Solving the inverse problem for measures

first calculated to order $n = 34$ by Isenberg [20]. Wheeler and Gordon [27] then constructed Padé approximants from these moments in order to numerically approximate this integral. Using 30 moments, they obtained the bounds

$$0.3408807 < u_0 \leq 0.3408883.$$  

With only 10 moments, their approximation was correct to 1 part in $10^4$. Using additional information about $G(x)$, they were able to improve the bounds in Equation (4.9):

$$0.3408872177 < u_0 \leq 0.3408872204.$$  

Bessis and Demko’s method of polynomial sampling uses homogeneous IFS maps on $[0, 1]$, $w_i(x) = sx + a_i$ with probabilities $p_i$, $i = 1, 2, \ldots, N$. The $2N + 1$ parameters $s, a_i, p_i$ are then determined so that $M$ target moments $g_n$, $n = 1, 2, \ldots, M$, are matched. Bessis and Demko then computed the integral in Equation (4.8) over
the approximating IFS invariant measure. (The method of computing this integral is described below.) For the case $N = 4$ and $M = 9$, they obtained the approximation

$$u_0 = \frac{1}{2} \int_0^1 \sqrt{x} \, d\mu = 0.340899.$$ (4.11)

The relative error of this approximation is $3.5 \times 10^{-5}$, an improvement over the Padé bounds using 10 moments. In [24], these results were improved slightly by minimizing the function in Equation (3.2) with $M = 9$, using gradient optimization. However, the use of a higher number of moments, e.g. $M = 20$ or 30, was not investigated at that time.

In Table 2, we list the results of minimizing the moment collage distance $S_M^N$ in Equation (4.1), using $M = 30$ moments. The 'wavelet-type' IFS functions of Equation (2.7) were again used. In each case, the collage distance $S_M^N$ as well as the moment distance $\Gamma_M^N$ are presented along with the estimate $u_0$ afforded by the IFS invariant measure. The relative errors of these estimates are also shown.

| $N$ | $D_M^N$ | $\Gamma_M^N$ | $u_0$ | $\frac{|u_0 - u|}{u_0}$ | $\tilde{N}$ | $w^\tilde{N}$ |
|-----|---------|-------------|-------|--------------------------|---------|-----------|
| 2   | $1.0 \times 10^{-2}$ | $1.64 \times 10^{-2}$ | 0.328708 | $3.6 \times 10^{-2}$ | 2 (1, 2) |
| 4   | $1.9 \times 10^{-3}$ | $1.29 \times 10^{-3}$ | 0.340433 | $1.3 \times 10^{-3}$ | 3 (1, 2, 4) |
| 5   | $1.6 \times 10^{-3}$ | $1.23 \times 10^{-3}$ | 0.340200 | $2.0 \times 10^{-3}$ | 4 (1, 2, 4, 5) |
| 8   | $9.99 \times 10^{-4}$ | $1.03 \times 10^{-3}$ | 0.341317 | $1.3 \times 10^{-3}$ | 4 (2, 4, 5, 8) |
| 9   | $8.53 \times 10^{-4}$ | $8.64 \times 10^{-4}$ | 0.341763 | $2.6 \times 10^{-3}$ | 4 (2, 5, 8, 9) |
| 12  | $7.67 \times 10^{-4}$ | $7.59 \times 10^{-4}$ | 0.341901 | $3.0 \times 10^{-3}$ | 5 (2, 7, 8, 9, 12) |
| 13  | $1.20 \times 10^{-4}$ | $1.18 \times 10^{-4}$ | 0.340992 | $3.0 \times 10^{-3}$ | 6 (2, 3, 6, 8, 10, 13) |
| 16  | $1.9 \times 10^{-4}$ | $1.18 \times 10^{-4}$ | 0.341074 | $5.5 \times 10^{-4}$ | 6 (2, 6, 8, 10, 13, 16) |
| 17  | $1.17 \times 10^{-4}$ | $1.16 \times 10^{-4}$ | 0.341181 | $8.6 \times 10^{-4}$ | 6 (2, 6, 8, 10, 13, 17) |
| 22  | $1.17 \times 10^{-4}$ | $1.16 \times 10^{-4}$ | 0.341223 | $9.9 \times 10^{-4}$ | 7 (2, 6, 8, 10, 13, 17, 22) |
| 28  | $1.86 \times 10^{-5}$ | $1.77 \times 10^{-5}$ | 0.340711 | $5.1 \times 10^{-4}$ | 7 (2, 14, 15, 18, 19, 22, 28) |
| BD  | $2.33 \times 10^{-5}$ | $2.33 \times 10^{-5}$ | 0.340899 | $3.5 \times 10^{-5}$ | 4 |

This content downloaded from 129.97.58.73 on Fri, 22 Feb 2019 21:40:13 UTC
All use subject to https://about.jstor.org/terms
Solving the inverse problem for measures

The integral in Equation (4.11) was computed as in [8], using the property

\[ T^n f(x_0) \rightarrow \int_X f d\mu, \quad \text{as } n \rightarrow \infty, \quad x_0 \in X. \]

The action of the operator \( T: C(X) \rightarrow C(X) \) is given by [5]

\[ (Tf)(x) = \sum_{i=1}^{N} p_i (f \circ w_i)(x). \]

The iterates in Equation (4.12) are given by the nested sums

\[ (T^nf)(x) = \sum_{i_1=1}^{N} \cdots \sum_{i_n=1}^{N} p_1 \cdots p_n f(w_{i_1} \circ \cdots \circ w_{i_n})(x). \]

The evaluation of this quantity involves the enumeration of an \( N \)-tree to \( n \) generations.

5. Final remarks

In this paper, we have proved a collage theorem for moments which can greatly simplify the calculations involved in moment matching procedures using IFS. Moment matching becomes the minimization of a moment collage distance. A further simplification is achieved by using a fixed, infinite set of affine IFS maps satisfying an \( \epsilon \)-contractivity condition. In this case, the minimization is performed only with respect to the IFS probabilities \( p_i \). Note from Equation (3.16) that the squared moment collage distance \( S^N \) (as well as its truncation, \( S^N_M \) in Equation (4.1)) is a quadratic function of the \( p_i \). The optimization problem may be performed by quadratic programming (QP) which locates a minimum on the simplex \( \Pi^N \) in a finite number of steps. In many cases, the minimum occurs on the boundary of \( \Pi^N \) so that ‘useless’ IFS maps \( w_i \) are removed. The additional condition that we use an infinite number of IFS maps which satisfy an \( \epsilon \)-contractivity condition ensures that the collage distance \( S^N \rightarrow 0 \) as \( N \rightarrow \infty \).

We have paid little attention to the question of choosing a set of affine IFS maps satisfying the \( \epsilon \)-contractivity condition and which may be ‘optimal’ for a given problem. The ‘wavelet-type’ functions of Equation (2.7) represent a convenient choice of IFS maps. The question of using other maps which may be better suited to particular problems is beyond the scope of this paper. Note that the use of a fixed set of IFS maps has already become a standard tool in image compression methods [21]. Our method differs in that it allows room for increasing degrees of refinement on the base space \( X \), as guaranteed by the \( \epsilon \)-contractivity condition. We have also not given much attention to the question of the ordering of the contraction maps \( w_i \) in the infinite set \( \mathcal{W} \). The ‘wavelet-type’ IFS maps of Equation (2.7) admit a natural ordering. Nevertheless, one may wish to exclude maps representing certain regions.
of $X$ or, alternatively, to insert maps to permit additional refinement in certain regions. From a practical perspective, it will be important to develop optimal algorithms which are based on the problem at hand.

The present work involving IFS with measures was motivated, in part, by an ongoing study of the inverse problem of function approximation using IFS-type methods. Our construction of IFS-type methods over function spaces began with iterated fuzzy set systems (IFZS) [11], [15]: a variation of IFS which is formulated over an appropriate subset of functions from the class of functions $\mathcal{F}(X) = \{u : X \to [0, 1]\}$, often referred to as the class of fuzzy sets on $X$. However, the IFZS approach still employs a Hausdorff metric which is very restrictive from practical as well as theoretical perspectives. By making two modifications to the IFZS approach [16], one arrives at an IFS with ‘grey level maps’ (IFSM) over the space $L^p(X, \mu)$. This, in turn, serves as the motivation to formulate IFS over the general function spaces $L^p(X, \mu)$. Our solution to the inverse problem for function and image approximation in $L^p(X, \mu)$ employs a strategy similar to the one described in this paper—constructing sequences of finite IFSM whose IFS maps $w_i$ are chosen from an infinite set of contraction maps $\mathcal{W}$ which satisfy a refinement condition on $(X, d)$ with respect to a measure $\mu$. The basic aspects of this theory as well as some very encouraging results involving function and image approximation have already been reported [17].

Appendix: Proof of Proposition 3.4

**Proposition 3.4.** Let $X = [0, 1]$. Now define the following metric on $D(X)$ (cf. Equation (2.12)): for $u, v \in D(X)$, $\bar{d}_2(u, v) = \|u - v\|_2$. Then $(D(X), \bar{d}_2)$ is a complete metric space.

**Proof.** Let $g^{(n)} = (g_0^{(n)}, g_1^{(n)}, \cdots) \in D(X)$ for $n = 1, 2, \cdots$ be a Cauchy sequence in $\bar{P}$, that is, for any $\epsilon > 0$, there exists an $N > 0$ such that

\[ \|g^{(n)} - g^{(m)}\|_2 < \epsilon, \quad \forall m, n > N. \]  

Let $\nu^{(n)} \in \mathcal{M}(X)$, $n = 1, 2, \cdots$, be the probability measures whose moments are the components of the $g^{(n)}$, i.e. for $n = 1, 2, \cdots$,

\[ g_k^{(n)} = \int_X x^k \, d\nu^{(n)}, \quad k = 0, 1, 2, \cdots. \]  

Now consider the sequences $a^{(n)} = (a_0^{(n)}, a_1^{(n)}, \cdots)$, $n = 1, 2, \cdots$, where $a_0^{(n)} = g_0^{(n)} = 1$ and

\[ a_k^{(n)} = \frac{1}{k} g_k^{(n)}, \quad k = 1, 2, \cdots. \]
Solving the inverse problem for measures

Since $0 \leq g_k^{(n)} \leq 1$ for $k \geq 0$, it follows that $a^{(n)} \in L^2(N)$ for all $n \geq 1$. Furthermore, from Equation (A.1), $\{a^{(n)}\}_{n=1}^{\infty}$ is a Cauchy sequence in $L^2(N)$. Hence, by the completeness of $L^2(N)$, there exists an $a = (a_0, a_1, \cdots) \in L^2(N)$ such that

(A.4)\[ \|a^{(n)} - a\|_2 \to 0 \quad \text{as } n \to \infty. \]

Now let $g = (g_0, g_1, \cdots)$, where $g_0 = 1$ and $g_k = ka_k$, $k = 1, 2, \cdots$. From Equation (A.4), it follows that for each $k = 1, 2, \cdots$, $|a_k^{(n)} - a^{(n)}| \to 0$ as $n \to \infty$, which, in turn, implies that $|g_k^{(n)} - g_k| \to 0$ as $n \to \infty$. Since $a \in L^2(N)$, then $g$, the limit of the Cauchy sequence $\{g^{(n)}\}$, is an element of $L^2(N)$. We now show that $g \in D(X)$.

A necessary and sufficient condition that an infinite set of real numbers $c = (c_0, c_1, \cdots)$ be the moments of a unique probability measure $\mu \in M(X)$, i.e. $c_n = \int XX^n d\mu$, $n = 0, 1, 2, \cdots$, is that they satisfy the Hausdorff inequalities [3]:

(A.5)\[ H_i(c) = \sum_{m=0}^{l} (-1)^m \binom{i}{m} c_{i+m} \geq 0, \quad i, j \in \{0, 1, 2, \cdots\}. \]

Since for each fixed $n \geq 1$ the $g_k^{(n)}$, $k = 0, 1, 2, \cdots$, are the moments of the measures $\nu^{(n)} \in M(X)$, cf. Equation (A.2), they must satisfy the following relations:

(A.6)\[ H_i(g^{(n)}) = \sum_{m=0}^{l} (-1)^m \binom{i}{m} g_{i+m}^{(n)} \geq 0, \quad i, j \in \{0, 1, 2, \cdots\}. \]

The limit as $n \to \infty$ of each of these inequalities may be now be taken:

(A.7)\[ H_i(g) = \sum_{m=0}^{l} (-1)^m \binom{i}{m} g_{i+m} \geq 0, \quad i, j \in \{0, 1, 2, \cdots\}. \]

The above inequalities are simply the Hausdorff inequalities for the sequence $g$. This implies that $g_k = \int XX^k d\nu$, $k = 0, 1, 2, \cdots$, for a unique measure $\nu \in M(X)$. Thus $g \in D(X)$, which completes the proof.

Acknowledgements

We wish to express our thanks to Professor M. J. Best, Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo for many useful discussions on quadratic programming and for providing the FORTRAN QP subroutines which were used in the computations reported above. This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada, which are gratefully acknowledged.

References


