Information Processing Letters 40 (1991) 113–117
North-Holland

25 October 1991

# An algorithm for the computation of the Hutchinson distance

Jonathan Brandt and Carlos Cabrelli *

*Division of Computer Science, UC Davis, Davis, CA 95616, USA*

Ursula Molter *

*Department of Mathematics, UC Davis, Davis, CA 95616, USA*

Communicated by Q.F. Stout
Received 1 August 1990
Revised 1 April 1991

*Abstract*

Brandt, J., C. Cabrelli, and U. Molter, An algorithm for the computation of the Hutchinson distance, Information Processing Letters 40 (1991) 113–117.

A simple linear-time algorithm for the computation of the Hutchinson metric in the case of finite one-dimensional sequences is presented. The algorithm is derived by way of a proof demonstrating that in this case, the Hutchinson metric can be expressed as the sum of the absolute values of the partial sums of the pointwise difference of the input measures.

*Keywords*: Design of algorithms, similarity measures, Hutchinson distance

## 1. Introduction

There are many settings in which it is desired to compare two probability measures. One currently important setting is in approximation of a measure by a fractal model, for instance an *iterated function system* [1]. Here, one would like to compare the approximation to the desired measure in order to determine the relative fidelity. The Hutchinson metric [2], defined over the space of probability measures, has been suggested to accomplish this comparison. However, practical computation of the Hutchinson distance is difficult because it involves optimization over a large space of functions. We have found that in the case of finite one-dimensional sequences, the optimization reduces to a simple linear-time computation. This algorithm should have immense practical value in assessing the relative similarity of two measures.

## 2. Problem formulation

Let $(X, d)$ be a metric space and $\mu$ and $\nu$ two probability measures on $X$, then the Hutchinson distance between $\mu$ and $\nu$ is defined by

$$d_H(\mu, \nu) = \sup_{f \in Lip_1} \left\{ \int_X f \, d\mu - \int_X f \, d\nu \right\}, \tag{1}$$

0020-0190/91/$03.50 

where

$$Lip_1 = \{f: X \to \mathbb{R}: Lip(f) \leqslant 1\},$$

and

$$Lip(f) = \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)}.$$

It can be shown that $Lip_1$ is a convex set.

The supremum in (1) can be taken over a smaller set by observing that $\int_X f \, d\mu - \int_X f \, d\nu$ has the same value for $f$ as for $f + c$ for any $c \in \mathbb{R}$. Therefore we can consider the equivalence relation

$$f \sim g \quad \Leftrightarrow \quad f - g = \text{const}.$$

and the supremum in (1) will be the same if it is taken over any subset of $Lip_1$ containing at least one function in each equivalence class.

## 3. Hutchinson distance of finite one-dimensional sequences

Let us now consider the discrete one-dimensional case. Suppose $X = \{1, \ldots, n\}$ and $d(i, j) = |i - j|$ and let

$$\mathscr{M} = \left\{\mu = (\mu_1, \ldots, \mu_n): 0 \leqslant \mu_i, \ i = 1, \ldots, n \text{ and } \sum_{i=1}^{n} \mu_i = 1\right\}$$

be the set of probability vectors.

For $\mu, \nu \in \mathscr{M}$,

$$d_H(\mu, \nu) = \sup_f \left\{\sum_{i=1}^{n} f_i \mu_i - \sum_{i=1}^{n} f_i \nu_i: |f_i - f_j| \leqslant |i - j|, \ 1 \leqslant i, j \leqslant n\right\}. \tag{2}$$

One can show that the constraints in (2) are equivalent to

$$|f_i - f_{i+1}| \leqslant 1, \quad 1 \leqslant i \leqslant n - 1.$$

Defining $\eta = \mu - \nu$, (2) can be rewritten as

$$d_H(\mu, \nu) = \sup_f \left\{\sum_{i=1}^{n} f_i \eta_i: |f_i - f_{i+1}| \leqslant 1, \ i = 1, \ldots, n - 1\right\}. \tag{3}$$

Now consider the hyperplane, $P$, defined as

$$P = \left\{x = (x_1, \ldots, x_n): \sum_{i=1}^{n} x_i = 0\right\}.$$

It can be shown that $Lip_1 \cap P$ contains exactly one element of each equivalence class, therefore (3) can be rewritten as

$$d_H(\mu, \nu) = \sup_f \left\{\sum_{i=1}^{n} f_i \eta_i: |f_i - f_{i+1}| \leqslant 1, \ i = 1, \ldots, n - 1, \ \sum_{i=1}^{n} f_i = 0\right\}.$$

* Permanent address: Department of Mathematics FCEyN, UBA Pab. I Cdad. Universitaria (1428) Cap. Fed. Argentina.

Similarly, if $P_1$ is defined as

$$P_1 = \{x = (x_1, \dots, x_n) \colon x_1 = 0\},$$

then $Lip_1 \cap P_1$ also contains exactly one element of each equivalence class. In this case, (3) yields

$$d_H(\mu, \nu) = \sup_f \left\{ \sum_{i=1}^{n} f_i \eta_i \colon | f_i - f_{i+1} | \leqslant 1, i = 1, \dots, n-1, f_1 = 0 \right\}.$$

Hence, to compute this distance, one can simply solve the linear programming problem where the cost function, $\sum_{i=1}^{n} \eta_i x_i$, is maximized on the vectors satisfying

$$\sum_{i=1}^{n} x_i = 0; \quad -1 \leqslant x_i - x_{i+1} \leqslant 1, \ 1 \leqslant i \leqslant n-1. \tag{4}$$

It is easy to show that these inequalities describe a convex set (a polyhedron) with $2^{n-1}$ vertices. Therefore, a naive solution such as application of the simplex algorithm would be expected to require exponential time. However, the constraints in (4) contain structure which can be exploited to reduce the search to linear time. To show this, we shall prove that $d_H(\mu, \nu)$ for finite sequences can be expressed in a much simpler form, from which the distance computation algorithm immediately follows.

**Theorem 1.** *Using the notation above, if $\mu$ and $\nu$ are in $\mathscr{M}$, then the Hutchinson distance between $\mu$ and $\nu$ is*

$$d_H(\mu, \nu) = \sum_{k=1}^{n-1} | \mathscr{S}_k |,$$

*where*

$$\mathscr{S}_k = \sum_{i=1}^{k} (\mu_i - \nu_i).$$

**Proof.** Collect the constraints of (4) into an $n \times n$ matrix,

$$V = \begin{bmatrix} 1 & 1 & 1 & . & . & . & 1 \\ 1 & -1 & 0 & . & . & . & 0 \\ 0 & 1 & -1 & 0 & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & . & . & . & 0 & 1 & -1 \end{bmatrix}.$$

$V$ is a nonsingular matrix and if we call its rows $v_0, v_1, \dots, v_{n-1}$, then the set of vectors $\mathscr{B} = \{v_1, \dots, v_{n-1}\}$ form a basis of the hyperplane $P$ because each $v_i$ ($i \neq 0$) is in $P$, they are linearly independent, and the dimension of $P$ is $n-1$.

Suppose a vector $x = (x_1, \dots, x_n)$ is in $P$, then it can be written in the basis $\mathscr{B}$ as

$$x = \sum_{k=1}^{n-1} \sigma_k v_k \quad \text{with } \sigma_k = \sum_{i=1}^{k} x_i.$$

In particular, since $\mu$ and $\nu$ are in $\mathscr{M}$, then $\eta = \mu - \nu$ is in $P$. Therefore, in the basis $\mathscr{B}$,

$$\eta = \sum_{k=1}^{n-1} \mathscr{S}_k v_k,$$

where $\mathscr{S}_k$ is defined above. For each $f$, we then have

$$\langle f, \eta\rangle = \left\langle f, \sum_{k=1}^{n-1} \mathscr{S}_k v_k \right\rangle = \sum_{k=1}^{n-1} \mathscr{S}_k \langle f, v_k\rangle = \sum_{k=1}^{n-1} \mathscr{S}_k (f_k - f_{k+1}),$$

where the symbol $\langle\, ,\rangle$ denotes the usual scalar product.

Since $\langle f, \eta\rangle \leqslant |\langle f, \eta\rangle|$,

$$\langle f, \eta\rangle \leqslant \left| \sum_{k=1}^{n-1} \mathscr{S}_k (f_k - f_{k+1}) \right| \leqslant \sum_{k=1}^{n-1} |\mathscr{S}_k|\,|f_k - f_{k+1}|. \tag{5}$$

For $f \in Lip_1$, $|f_k - f_{k+1}| \leqslant 1$ for $1 \leqslant k \leqslant n-1$, hence (5) yields

$$\langle f, \eta\rangle \leqslant \sum_{k=1}^{n-1} |\mathscr{S}_k|$$

and $d_H(\mu, \nu)$ is therefore bounded as

$$d_H(\mu, \nu) \leqslant \sum_{k=1}^{n-1} |\mathscr{S}_k|.$$

To complete the proof, we now exhibit a function $f^0 \in Lip_1$ such that

$$\langle f^0, \eta\rangle = \sum_{k=1}^{n-1} |\mathscr{S}_k|. \tag{6}$$

The construction of $f^0$ is recursive:

1. Let $f_1^0 = 0$.
2. Assume that $f_i^0$ has been chosen and let $f_{i+1}^0 = f_i^0 - sg(\mathscr{S}_i)$, where $sg(x)$ is the function taking the values 1, 0 or $-1$, when x is positive, zero or negative respectively.

To demonstrate that (6) holds, form the scalar product.

$$\langle f^0, \eta\rangle = \sum_{k=1}^{n-1} \mathscr{S}_k (f_k - f_{k+1}) = \sum_{k=1}^{n-1} \mathscr{S}_k sg(\mathscr{S}_k) = \sum_{k=1}^{n-1} |\mathscr{S}_k|. \qquad \square$$

It follows from the theorem that in order to determine the Hutchinson distance between $\mu$ and $\nu$ we need only compute $\mathscr{S}_k$ for $k = 1, \ldots, n-1$. Also, the proof of the theorem yields a method to calculate all the vectors where this maximum is reached. If $\mathscr{S}_k \neq 0$ for $k = 1, \ldots, n-1$, then only one vector realizes this maximum, and it is the one given by the construction in the theorem. But if any $\mathscr{S}_k = 0$, then there are infinite vectors which realize the maximum and they can be found choosing $f_{k+1}$ to be any value between $f_k - 1$ and $f_k + 1$. These observations are summarized in the following corollary:

**Corollary 2.** *Under the hypothesis of the theorem,*

$$\langle x, \mu - \nu\rangle = \sum_{k=1}^{n-1} |\mathscr{S}_k| = d_H(\mu, \nu),$$

(1) *has unique solution in* $Lip_1 \cap P_1$ *iff* $\mathscr{S}_k \neq 0$ *for* $k = 1, \ldots, n-1$,

(2) *all the solutions in* $Lip_1 \cap P_1$ *are given by*:

- $x_1 = 0$,
- $x_{i+1} = x_i - sg(\mathscr{S}_i) + \delta(\mathscr{S}_i)a_i$ *with* $|a_i| \leqslant 1$. (*Here the symbol* $\delta$ *stands for the delta function*.)

**Note.** Each solution in $Lip_1$ can be decomposed as a solution in $Lip_1 \cap P_1$ plus a vector of the form $(c, \ldots, c)$ with $c \in \mathbb{R}$.

## 4. Conclusions

In the case of finite one-dimensional sequences, the Hutchinson distance can be computed by summing the absolute value of the partial sums of the difference of the two measures and therefore requires linear time. This result does not extend to higher dimensions because the constraints do not form a basis in the hyperplane $P$ (even though they generate $P$, they are not linearly independent). However, there are many potential applications of the one-dimensional Hutchinson metric, not only for fractal approximation, but for more general comparison problems which arise in such fields as signal processing, statistics, and pattern recognition.

## Acknowledgment

We would like to thank Professor Edward R. Vrscay for suggesting the problem and for many fruitful discussions.

One of us (C. Cabrelli) acknowledges support from CONICET, Argentina.

## References

[1] M.F. Barnsley, *Fractals Everywhere* (Academic Press, San Diego, CA, 1988).
[2] J. Hutchinson, Fractals and self-similarity, *Indiana Univ. Math. J.* **30** (1981) 713–747.