Thinning, Inserting, and Swapping Scattered Data

Michael S. Floater and Armin Iske

Abstract. Thinning, insertion, and swapping algorithms for generating well-distributed hierarchical subsets of scattered data in a given domain in $\mathbb{R}^d$ are proposed. It is found that thinning and insertion yield approximately similar results but that swapping is in general inferior.

§1. Introduction

In two previous papers [1,2], a thinning algorithm, based on Delaunay triangulations, was used to construct hierarchies of well-distributed subsets of scattered data in $\mathbb{R}^2$. These hierarchies were subsequently used for multistep scattered data interpolation using compactly supported radial basis functions [3].

In this paper we consider a more general setting. We suppose we are given a non-empty, simply connected, bounded, open set $\Omega \subset \mathbb{R}^d$ and a set $X \subset \Omega$ of N pairwise distinct scattered data points, $N \geq 1$. We propose and compare three general algorithms: thinning, insertion and swapping, for constructing hierarchical sequences of the form

$$X_1 \subset X_2 \subset \cdots \subset X_N = X, \tag{1}$$

with $\#(X_i) = i$, such that the subsets X_i are as well distributed in Ω as possible according to a measure of uniformity.

The new thinning algorithm is more advanced than the algorithm in [1,2]. The new algorithm treats the boundary of the domain differently and it is greedy, i.e. it maximizes the uniformity at each step. The insertion algorithm is the dual of the thinning algorithm while the swapping algorithm is based on an entirely different strategy.

Proceedings of Chamonix 1996
A. Le Méhauté, C. Rabut, and L. L. Schumaker (eds.), pp. 1–6.

ISBN 1-xxxxx-xxx-x.

§2. Uniformity and Hierarchical Sequences

We will be interested in two ways of measuring the density of the points in X with respect to Ω. Let

$$s(X) = \min_{x_1 \in X} \min_{\substack{x_2 \in X \cup \partial\Omega \\ x_2 \neq x_1}} \|x_1 - x_2\|, \quad \ell(X) = 2 \max_{y \in \Omega} \min_{x \in X \cup \partial\Omega} \|x - y\|, \tag{2}$$

where $\partial\Omega$ denotes the boundary of Ω. We note that the maximum in $\ell(X)$ is attained by some $y \in \Omega$ even though Ω is an open set. In fact we can interpret $\ell(X)$ geometrically by observing that it is the diameter of the largest open ball contained in $\Omega \setminus X$. We can also interpret $s(X)$, namely as the minimum of the distance between X and $\partial\Omega$ and the shortest distance between the points in X.

In the case $d = 1$, Ω will be some interval (a, b) and we denote the points (or values) in X by $x_1, \ldots, x_N$ where $a < x_1 < x_2 < \cdots < x_N < b$. If we then define $x_0 = a$, $x_{N+1} = b$, and $\Delta_j = x_{j+1} - x_j$ for $j = 0, \ldots, N$, one sees that

$$s(X) = \min_{0 \le j \le N} \Delta_j, \qquad \ell(X) = \max_{0 \le j \le N} \Delta_j,$$

that is, s and ℓ are the lengths of the shortest and longest intervals in $\Omega \setminus X$ respectively.

By taking the ratio of s and ℓ in (2) we obtain a measure of the *uniformity* of X with respect to Ω:

$$\rho(X) = s(X)/\ell(X).$$

Proposition 2.1. *For any $d \ge 1$, we have $0 < \rho(X) \le 1$.*

Proof: Obviously $\rho(X) > 0$. Now from (2) we may assume that there exist $x_1 \in X$ and $x_2 \in X \cup \partial\Omega$, $x_2 \neq x_1$, such that $s(X) = \|x_1 - x_2\|$. Then $s(X) = \min_{x \in X \cup \partial\Omega} \|x_1 - x\|$. If we let $y_0 = (x_1 + x_2)/2$ then $y_0 \in \Omega \setminus X$ and

$$\|x - y_0\| \ge \|x - x_1\| - \|x_1 - y_0\| \ge s(X) - s(X)/2 = s(X)/2$$

for all $x \in X \cup \partial\Omega$. It follows from (2) that $\ell(X) \ge s(X)$. ■

Note that when $\Omega = B_1(0)$, the open ball in $\mathbb{R}^d$ with centre 0 and radius 1, and $X = \{0\}$, we see that $\rho(X) = 1$ for any dimension $d \ge 1$ since $s(X) = \ell(X) = 1$. In the case $d = 1$, $\rho(X) = 1$ if and only if Δ_j is constant.

Our goal is to construct from X a hierarchical sequence of N subsets in the form (1) such that the X_i are as uniform as possible. To be precise, let $\mathcal{X} = (X_1, \ldots, X_N)$, define $\rho(\mathcal{X}) = (\rho(X_1), \ldots, \rho(X_N)) \in \mathbb{R}^N$ and denote by $\mathrm{H}(X)$ the set of all hierarchical sequences (1). We will regard $\mathcal{X} \in \mathrm{H}(X)$ as a 'good' hierarchical sequence if, for some p, $1 \le p < \infty$, the l_p norm $||\rho(\mathcal{X})||_p$ is large. Since there are $N!$ hierarchical sequences in $\mathrm{H}(X)$, it would be very costly to compute $||\rho(\mathcal{X})||_p$ for all $\mathcal{X} \in \mathrm{H}(X)$ in order to find a sequence $\mathcal{X}^* \in \mathrm{H}(X)$ for which

$$||\rho(\mathcal{X}^*)||_p = \max_{\mathcal{X} \in \mathrm{H}(X)} ||\rho(\mathcal{X})||_p. \tag{3}$$

Instead we will discuss three different algorithms which generate hierarchical sequences in acceptable computation time.

Note that if $Y \subset X$ and $x \in Y$, then from (2),

$$s(Y \setminus \{x\}) \geq s(Y), \qquad \ell(Y \setminus \{x\}) \geq \ell(Y).$$

Thus if $(X_1, \ldots, X_N) \in \mathrm{H}(X)$ then the sequences of values $s(X_1), \ldots, s(X_N)$ and $\ell(X_1), \ldots, \ell(X_N)$ are monotonically decreasing. However the same will not in general be true of the sequence $\rho(X_1), \ldots, \rho(X_N)$.

§3. Swapping Algorithm

If we are given a hierarchical sequence $\mathcal{X}_0 = (X_1, \ldots, X_N) \in \mathrm{H}(X)$, we can, for any $i = \{1, \ldots, N-1\}$, form a new hierarchical sequence:

$$\mathcal{X}_1 = (X_1, \ldots, X_{i-1}, X_{i-1} \cup (X_{i+1} \setminus X_i), X_{i+1}, \ldots, X_N),$$

in $\mathrm{H}(X)$, where we define $X_0 = \emptyset$. We will refer to this change of sequence as a *swap* for the following reason. We can identify $\mathcal{X}_0$ with the point sequence $S_{\mathcal{X}_0} = (y_1, \ldots, y_N)$, where $X_j \setminus X_{j-1} = \{y_j\}$, for $j = 1, \ldots, N$. Now we see that replacing X_i by $X_{i-1} \cup (X_{i+1} \setminus X_i)$ to form $\mathcal{X}_1$ is equivalent to swapping y_i and y_{i+1} in $S_{\mathcal{X}_0}$ to form $S_{\mathcal{X}_1} = (y_1, \ldots, y_{i-1}, y_{i+1}, y_i, y_{i+2}, \ldots, y_N)$. Since this swap only changes the subset X_i, only the value of $\rho(X_i)$ is changed. This suggests an algorithm for recursively increasing the norm $||\rho(\mathcal{X})||_p$, starting with some arbitrary sequence $\mathcal{X}_0$.

Algorithm 3.1 (Swapping Algorithm).

(1) Let $\mathcal{X}_0 \in \mathrm{H}(X)$ be any hierarchical sequence and let $k = 0$.

(2) Suppose $\mathcal{X}_k = (X_1, \ldots, X_N)$. Search for an index $i \in \{1, \ldots, N-1\}$ such that

$$\rho(X_{i-1} \cup (X_{i+1} \setminus X_i)) > \rho(X_i).$$

(3) If such an i exists, define $\mathcal{X}_{k+1}$ from $\mathcal{X}_k$ by replacing X_i by $X_{i-1} \cup (X_{i+1} \setminus X_i)$, let $k = k+1$, and go to step 2.

(4) Else, let $\mathcal{X}^* = \mathcal{X}_k$ and output $\mathcal{X}^*$.

In order to discuss the result of this algorithm, let us say that a hierarchical sequence $\mathcal{X}^* = (X_1, \ldots, X_N) \in \mathrm{H}(X)$ is *globally optimal* with respect to some $p \geq 1$, if it satisfies (3). Meanwhile if

$$\rho(X_{i-1} \cup (X_{i+1} \setminus X_i)) \leq \rho(X_i) \qquad \text{for all } i = 1, \ldots, N-1, \tag{4}$$

then we will say that $\mathcal{X}^*$ is *locally optimal.* We note that since one swap only alters one $\rho(X_i)$, a globally optimal sequence is also locally optimal.

Proposition 3.2. *The swapping algorithm always terminates after finitely many steps and $\mathcal{X}^*$ is locally optimal.*

Proof: In the swapping algorithm we see that

$$||\rho(\mathcal{X}^*)||_p > \cdots > ||\rho(\mathcal{X}_k)||_p > \cdots > ||\rho(\mathcal{X}_1)||_p > ||\rho(\mathcal{X}_0)||_p.$$

Since there are only a finite number of sequences in $\mathrm{H}(X)$ the algorithm has therefore to stop after a finite number of steps. The fact that $\mathcal{X}^*$ is locally optimal follows immediately from the swapping criterion used in Step (2). ■

The following example shows that the swapping algorithm will not always yield a globally optimal sequence.

Example 3.3. *Let $d = 1$, $\Omega = (0,1)$, $N = 2^K - 1$ for some $K > 1$, and $x_i = i/(N+1)$ for $i = 1,\ldots,N$. Let $y_{2k-1} = k/(N+1)$ for $k = 1,\ldots,(N+1)/2$ and $y_{2k} = 1 - k/(N+1)$ for $k = 1,\ldots,(N-1)/2$, and let $\mathcal{X}_0 = (X_1,\ldots,X_N)$ where $X_i = \{y_1,\ldots,y_i\}$. Then one can show that $\mathcal{X}_0$ is a locally optimal sequence. Hence the swapping algorithm yields $\mathcal{X}^* = \mathcal{X}_0$ and we find that $\rho(\mathcal{X}^*) = (1/N, 1/(N-1),\ldots,1/2,1)$. But if $\mathcal{X}$ is defined by $(y_1,\ldots,y_N) = (1/2,1/4,3/4,1/8,3/8,5/8,7/8,1/16,\ldots,1/(N+1),\ldots,N/(N+1))$, then $\rho(\mathcal{X}) = (1,1/2,1,1/2,1/2,1/2,1,\ldots,1/2,1)$ and so $||\rho(\mathcal{X})||_p > ||\rho(\mathcal{X}^*)||_p$ for $1 \le p < \infty$. Thus $\mathcal{X}^*$ is far from globally optimal.*

§4. Thinning and Insertion Algorithms

We now consider two greedy algorithms for generating suitable hierarchical sequences of X. The basic idea in the first algorithm is to let $X_N = X$ and recursively remove points from X until we obtain X_1 containing merely one point.

Definition 4.1. *Let $Y \subset X$. We will say $x \in Y$ is a* `removable point` *if*

$$\rho(Y \setminus \{x\}) = \max_{y \in Y} \rho(Y \setminus \{y\}).$$

Algorithm 4.2 (Thinning Algorithm).
(1) Let $X_N = X$.
(2) For decreasing $i = N,\ldots,2$
  (a) search for a removable point $x \in X_i$,
  (b) let $X_{i-1} = X_i \setminus \{x\}$
(3) Output $\mathcal{X}^* = (X_1,\ldots,X_N)$.

Example 4.3. *Let $\Omega = (0,1)$ and $x_i = 2^{i-N-1}$, $i = 1,\ldots,N$, for some $N \ge 1$. In this case the hierachical sequence $\mathcal{X}^* = (X_1,\ldots,X_N)$ obtained from the thinning algorithm is unique with $X_i = \{x_{N-i+1},\ldots,x_N\}$ for $i = 1,\ldots,N$. Further, $s(X_i) = 2^{-i}$, $\ell(X_i) = 1/2$ and thus $\rho(X_i) = s(X_i)/\ell(X_i) = 2^{1-i}$. So $\rho(\mathcal{X}^*) = (1,1/2,\ldots,1/2^{N-1})$ which is monotonically decreasing.*

Example 4.4. *Let $\Omega = (0,1)$, $N = 3$, and $x_i = i/4$ for $i = 1,2,3$ and let $\mathcal{X}^*$ be a hierarchical sequence generated by the thinning algorithm. Now the thinning algorithm may remove x_1 or x_3 first, in which case $\rho(\mathcal{X}^*) = (1,1/2,1)$ and $||\rho(\mathcal{X}^*)||_1 = 5/2$. However the algorithm may instead remove x_2 first and then $\rho(\mathcal{X}^*) = (1/3,1/2,1)$ and $||\rho(\mathcal{X}^*)||_1 = 11/6$. So in this example $\mathcal{X}^*$ and $\rho(\mathcal{X}^*)$ are not unique.*

Instead of recursively removing points from $X \subset \Omega$, we now recursively insert them into Ω.

Definition 4.5. *Let $Y \subset X$, $Y \neq X$. Then we will say that $x \in X \setminus Y$ is an* `insertable point` *if*

$$\rho(Y \cup \{x\}) = \max_{y \in X \setminus Y} \rho(Y \cup \{y\}).$$

Algorithm 4.6 (Insertion Algorithm).
(1) Let $X_0 = \emptyset$.
(2) For $i = 0, \ldots, N-1$
(a) search for an insertable point $x \in X \setminus X_i$,
(b) let $X_{i+1} = X_i \cup \{x\}$
(3) Output $\mathcal{X}^* = (X_1, \ldots, X_N)$.

The following examples show that in general either of the thinning and insertion algorithms can be superior in the sense of maximizing $||\rho(\mathcal{X})||_p$.

Example 4.7. *Consider again the uniform data set of Example 4.4 and let $\mathcal{X}^* = (X_1, X_2, X_3)$ be the result of the insertion algorithm. Since $X_1 = \{1/2\}$ we have that $\rho(\mathcal{X}^*) = (1, 1/2, 1)$ and so $||\rho(\mathcal{X}^*)||_1 = 5/2$. Thus for this data set, the insertion algorithm can be superior to the thinning algorithm.*

Example 4.8. *Let $\Omega = (0,1)$ and consider the set $X = \{1/3, 1/2, 2/3\}$. Let $\mathcal{X}_1^*$ and $\mathcal{X}_2^*$ be sequences generated by the thinning algorithm and insertion algorithm respectively. We observe that $\rho(\mathcal{X}_1^*) = (1/2, 1, 1/2)$ while $\rho(\mathcal{X}_2^*) = (1, 1/3, 1/2)$. Therefore $||\rho(\mathcal{X}_1^*)||_p > ||\rho(\mathcal{X}_2^*)||_p$ for $1 \leq p < \infty$ and so in this case the thinning algorithm is superior.*

We now show that with respect to local optimality, thinning and insertion are superior to swapping.

Proposition 4.9. *Every hierarchical sequence generated by either the thinning algorithm or the insertion algorithm is locally optimal.*

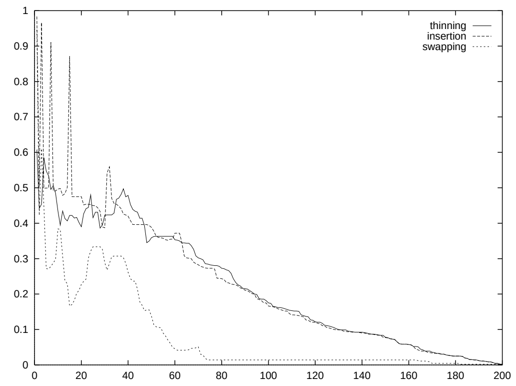
Proof: Let $\mathcal{X} = (X_1, \ldots, X_N)$ be a sequence generated by the thinning algorithm. For $j = 1, \ldots, N$, let $y_j \in X$ denote the point such that $X_j \setminus X_{j-1} = \{y_j\}$. Now noting that $X_{j+1} \setminus \{y_j\} = X_{j-1} \cup (X_{j+1} \setminus X_j)$, we find from Definition 4.1 that for all $j = 1, \ldots, N-1$,

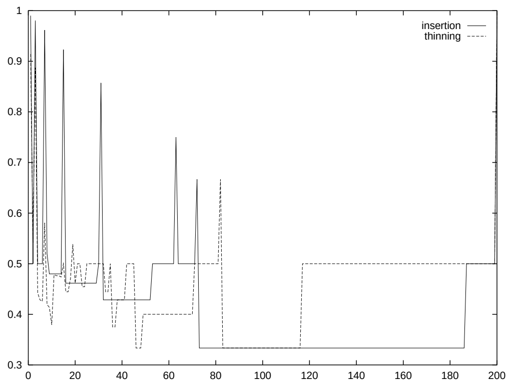
$$\rho(X_j) = \rho(X_{j+1} \setminus \{y_{j+1}\}) \geq \rho(X_{j+1} \setminus \{y_j\}) = \rho(X_{j-1} \cup (X_{j+1} \setminus X_j))$$

and so from (4) $\mathcal{X}$ is locally optimal. Now suppose $\mathcal{X}$ is generated by the insertion algorithm. Then due to Definition 4.5 we obtain

$$\rho(X_j) = \rho(X_{j-1} \cup \{y_j\}) \geq \rho(X_{j-1} \cup \{y_{j+1}\}) = \rho(X_{j-1} \cup (X_{j+1} \setminus X_j)),$$

and so $\mathcal{X}$ is again locally optimal. ■





§5. Numerical Examples

Two numerical examples were computed with $d = 1$, $\Omega = (0, 1)$, $N = 200$ and the points $x_1, \ldots, x_N \in X$ were chosen (i) randomly and (ii) uniformly, that is $x_i = i/(N + 1)$, $i = 1, \ldots, N$. The figure above shows the graphs of the sequences $\rho(\mathcal{X}^*)$ generated by thinning, insertion, and swapping for data set (i) and thinning and insertion for data set (ii).

The table below shows the corresponding values of the norms $||\rho(\mathcal{X}^*)||_1$ and $||\rho(\mathcal{X}^*)||_2$ for data sets (i) and (ii) respectively.

(i)	thin	insert	swap
$\|\|\rho(\mathcal{X}^*)\|\|_1$	43.8	45.4	17.2
$\|\|\rho(\mathcal{X}^*)\|\|_2$	3.9	4.2	2.2

(ii)	thin	insert
$\|\|\rho(\mathcal{X}^*)\|\|_1$	91.3	81.5
$\|\|\rho(\mathcal{X}^*)\|\|_2$	6.6	6.0

References

1. Floater, M. S. and A. Iske, Multistep Scattered Data Interpolation using Compactly Supported Radial Basis Functions, to appear in *J. Comp. Appl. Math.*
2. Floater, M. S. and A. Iske, Thinning and Approximation of Large Sets of Scattered Data, to appear in *Advances in Multivariate Approximation, F. Fontanella, K. Jetter, P.-J. Laurent (eds.)*, World Scientific, 1996.
3. Schaback, R., Creating Surfaces from Scattered Data using Radial Basis Functions, in *Mathematical Methods for Curve and Surfaces*, M. Dæhlen, T. Lyche, and L. L. Schumaker (eds.), Vanderbilt University Press, Nashville, 1995, 477–496.

Michael S. Floater, Armin Iske
SINTEF
Postbox 124, Blindern
N-0314 Oslo, NORWAY
`Michael.Floater@si.sintef.no`
`Armin.Iske@si.sintef.no`