Computational Geometry 46 (2013) 140–147

Contents lists available at SciVerse ScienceDirect

Computational Geometry: Theory and Applications

www.elsevier.com/locate/comgeo

# Oja centers and centers of gravity [☆]

Dan Chen [a], Olivier Devillers [b], John Iacono [c], Stefan Langerman [d], Pat Morin [a,*]

[a] *School of Computer Science, Carleton University, Canada*
[b] *INRIA Sophia Antipolis — Méditerranée, France*
[c] *Department of Computer and Information Science, Polytechnic University, United States*
[d] *Département d'Informatique, Université Libre de Bruxelles, Belgium*

A R T I C L E I N F O

*Article history:*
Received 7 January 2011
Accepted 23 April 2012
Available online 26 April 2012
Communicated by Stephane Durocher

*Keywords:*
Data depth
Oja depth
Centerpoint theorem

A B S T R A C T

*Oja depth* (Oja 1983) is a generalization of the median to multivariate data that measures the centrality of a point $x$ with respect to a set $S$ of points in such a way that points with smaller Oja depth are more central with respect to $S$. Two relationships involving Oja depth and centers of mass are presented. The first is a form of Centerpoint Theorem which shows that the center of mass of the convex hull of a point set has low Oja depth. The second is an approximation result which shows that the center of mass of a point set approximates a point of minimum Oja depth.



## 1. Introduction

Given a set $S$ of $n$ points in $\mathbb{R}^d$, the *Oja depth* [9] of a point $x \in \mathbb{R}^d$ is

$$\mathrm{d}(x, S) = \sum_{y_1,\ldots,y_d \in \binom{S}{d}} \mathrm{v}(x, y_1, \ldots, y_d),$$

where $\mathrm{v}(p_1, \ldots, p_{d+1})$ denotes the volume of the simplex whose vertices are $p_1, \ldots, p_{d+1}$.[1] A point in $\mathbb{R}^d$ with the minimum Oja depth is called an *Oja center*.

### 1.1. New results

In this paper we consider relationships between centers of mass of certain sets and Oja depth. The *center of mass* of a finite point set $S \subset \mathbb{R}^d$ is the average of those points,

$$\mathrm{c}(S) = |S|^{-1} \sum_{x \in S} x.$$

If $P \subset \mathbb{R}^d$ is a bounded object of non-zero volume, the center of mass of $P$ is

☆ This work was initiated during the Workshop on Computational Geometry 2006 in Caldes de Malavella. The authors wish to thank Ferran Hurtado and the organizers for the opportunity of working on this topic.

* Corresponding author.
*E-mail addresses:* dchen4@connect.carleton.ca (D. Chen), Olivier.Devillers@sophia.inria.fr (O. Devillers), jiacono@poly.edu (J. Iacono), stefan.langerman@ulb.ac.be (S. Langerman), morin@scs.carleton.ca (P. Morin).

[1] In Oja's original definition, the sum is normalized by dividing by $\binom{|S|}{d}$. We omit this here since it changes none of our results and clutters our formulas.

0925-7721/$ – see front matter 
http://dx.doi.org/10.1016/j.comgeo.2012.04.004

$$\mathrm{c}(P) = \frac{\int_{x \in P} x \mathrm{d}x}{\mathrm{v}(P)}.$$

In this paper, we prove the following results about the Oja depth of an $n$ point set $S$, whose convex hull $A$ has unit volume and that has an Oja center $x$:

$$\mathrm{d}\big(\mathrm{c}(A), S\big) \leqslant \binom{n}{d} / (d+1), \tag{1}$$

$$\mathrm{d}\big(\mathrm{c}(S), S\big) \leqslant (d+1)\mathrm{d}(x, S). \tag{2}$$

The bound in (1) is not known to be tight. The bound in (2) is tight, up to a lower-order term, for some point sets $S$.

### 1.2. Related results

Our first result, (1), is a form of *Centerpoint Theorem* that upper-bounds the Oja depth of $\mathrm{c}(A)$, and hence also the Oja depth of $x$, in terms of the volume of the convex hull of $S$. Previously, centerpoint theorems were known for other depth functions such as Tukey depth [7,10,12] and simplicial depth [2,3,6]. To the best of our knowledge, this is the first such result for Oja depth.

Our next result, (2), can be viewed in two ways:

1. The first is a linear-time algorithm to find a point whose depth is a constant factor approximation of the depth of the Oja center. In 1-d, Oja depth is minimized by the median, which can be found in $O(n)$ time. However, in 2-d, the best known algorithm for minimizing Oja depth exactly takes $O(n \log^3 n)$ time [1]. Approximation algorithms for minimizing Oja depth, based on uniform grids and sampling from $\binom{S}{d}$, are given by Ronkainen, Oja, and Orponen [11]; this algorithm and several others are implemented in the R-package, OjaNP [5]. However, in pathological cases, their approximation algorithm is not guaranteed (or even likely) to find a point that closely approximates the Oja center, either in terms of distance or in terms of its Oja depth.[2]
2. Another view of (2) is that it gives insight into the Oja depth function and the Oja center. In some sense, it tells us that the Oja center is not terribly different from the center of mass of $S$, since the center of mass of $S$ minimizes, to within a constant factor, the Oja depth function.

## 2. Oja center and mass center of A

In this section, we relate the Oja depth of the center of mass of the convex hull of $S$ to the volume of the convex hull of $S$. Throughout this section, $A$ denotes the convex hull of $S$ and we assume, without loss of generality, that $\mathrm{v}(A) = 1$.

Our upper-bound is based on the following *central identity*: For any disjoint sets $X, Y \subseteq \mathbb{R}^d$ with $\mathrm{v}(X \cup Y) > 0$,

$$\mathrm{c}(X \cup Y) = \frac{\mathrm{v}(X)\,\mathrm{c}(X) + \mathrm{v}(Y)\,\mathrm{c}(Y)}{\mathrm{v}(X \cup Y)}.$$

We first give an inductive proof of our result for point sets in $\mathbb{R}^2$, and then give a proof for point sets in $\mathbb{R}^d$ that uses tools from convex geometry.

### 2.1. An upper bound in $\mathbb{R}^2$

The following result shows that, for a convex polygon $E$ (e.g., $E = A$), it is not possible to form an overly-large triangle that has $\mathrm{c}(E)$ as one of its vertices:

**Lemma 1.** *Let $A$ be a convex polygon and let $p_1$ and $p_2$ be any two points in $A$. Then* $\mathrm{v}(p_1, p_2, \mathrm{c}(A)) \leqslant \mathrm{v}(A)/3$.

**Proof.** Assume, without loss of generality, that $p_1p_2$ is horizontal and that $\mathrm{c}(A)$ is above $p_1p_2$. We may assume that $p_1p_2$ is edge of $A$ since, otherwise, we can remove the part of $A$ that is below $p_1p_2$. This decreases $\mathrm{v}(A)$ and moves $\mathrm{c}(A)$ further away from the segment $p_1p_2$, which increases $\mathrm{v}(\mathrm{c}(A), p_1, p_2)$.

The proof is by induction on the number of vertices of $A$. If $A$ is a triangle then one can easily verify the result. Therefore, assume $A$ has $n \geqslant 4$ vertices. Consider an edge $ab$ of $A$ where $a \neq p_2$ is adjacent to $p_1$ and let $c \neq a$ be adjacent to $b$ (see Fig. 1). Since $A$ has 4 or more vertices, we may assume that the $y$-coordinate of $b$ is not smaller than the $y$-coordinate of $a$. Otherwise we can reverse the roles of $p_1$ and $p_2$ and redefine $a$ and $b$ with respect to the new $p_1$.

Draw a ray $r$ whose origin is at $p_1$ and such that the triangle $t_1$ supported by $p_1a$, $ab$ and $r$ and the triangle $t_2$ supported by $r$, $ab$, and the line through $bc$ have the same area. Such a ray is guaranteed to exist by a standard continuity argument that starts with $r$ containing $a$ and rotates about $p_1$ until $r$ contains $b$.

[2] This follows from the fact that the value of the Oja depth function and the location of the Oja center can be arbitrarily different for two sets $S_1$ and $S_2$ that differ in only $d$ points [8].