Available online at www.sciencedirect.com

SCIENCE DIRECT®

Computational Geometry 32 (2005) 223–237

Computational Geometry
Theory and Applications

www.elsevier.com/locate/comgeo

Extremal point queries with lines and line segments and related problems ☆

Ovidiu Daescu [a,*], Robert Serfling [b]

[a] *Department of Computer Science, University of Texas at Dallas, Richardson, TX 75080, USA*
[b] *Department of Mathematical Sciences, University of Texas at Dallas, Richardson, TX 75080, USA*

Received 22 November 2004; received in revised form 1 February 2005; accepted 17 March 2005

Available online 4 May 2005

Communicated by S. Akl

Abstract

We address a number of extremal point query problems when P is a set of n points in $\mathbb{R}^d$, $d \geqslant 3$ a constant, including the computation of the farthest point from a query line and the computation of the farthest point from each of the lines spanned by the points in P. In $\mathbb{R}^3$, we give a data structure of size $O(n^{1+\varepsilon})$, that can be constructed in $O(n^{1+\varepsilon})$ time and can report the farthest point of P from a query line segment in $O(n^{2/3+\varepsilon})$ time, where $\varepsilon > 0$ is an arbitrarily small constant. Applications of our results also include: (1) Sub-cubic time algorithms for fitting a polygonal chain through an indexed set of points in $\mathbb{R}^d$, $d \geqslant 3$ a constant, and (2) A sub-quadratic time and space algorithm that, given P and an anchor point q, computes the minimum (maximum) area triangle defined by q with $P \setminus \{q\}$.

Keywords: Algorithm; Computational geometry; Segment; Query; Farthest point

1. Introduction

Approximating point sets by simple geometric objects is a major topic in geometric optimization. While such problems are well studied in the plane ($\mathbb{R}^2$), there are few results in higher dimensions, most

☆ This work was supported in part by the NSF award CCF-0430366.

* Corresponding author. Tel.: 972-883-4196. Fax: 972-883-2349.
E-mail address: daescu@utdallas.edu (O. Daescu).

0925-7721/$ – see front matter
doi:10.1016/j.comgeo.2005.03.002

related to computing simple geometric objects (smallest ball, minimum radius cylinder, etc.) enclosing a set of points. Computing the smallest enclosing cylinder (ball, cylindrical shell, etc.) is a fundamental problem in data analysis, computational metrology and data compression; see [5,9] and the references within. For example, in data analysis, for a given data set P of n points in $\mathbb{R}^d$, one may want to fit a line L through P that minimizes $\max\{d(L,p)\colon p \in P\}$, where $d(L,p) = \min\{d(q,p)\colon q \in L\}$ and $d(q,p)$ denotes the Euclidean distance between the points q and p. This is equivalent to finding the smallest radius cylinder enclosing P.

In some cases, a line L may be provided, corresponding to some hypothesis on the data set, and one would like to compute $\max\{d(L,p)\colon p \in P\}$ or $\min\{d(L,p)\colon p \in P\}$ (the farthest or closest point from the query line L), to validate or invalidate the hypothesis. In other problems, it is useful to index the points in the input set P. In the indexed version, for example, the index of a point may give an order on the time at which the data was collected, thus "weakly" encoding a $(d+1)$-dimension of the data set. Then, one could make queries *in the past* to extract some features of the data set, such as computing the extent of a subset of P collected in a given time frame. Some other applications of these scenarios include data compression, computational biology (in modeling neurons and molecules) and face recognition. For example, in [24], Li and Lu propose a novel classification method for face recognition that is based on computing the nearest feature line to a query point. A similar approach, when used to identify individuals in crowds, would result in computing closest points from query lines (a line would correspond to two distinct prototype feature points of a *query* individual).

In this paper we discuss a number of query problems on computing extremal points in a point set $P \in \mathbb{R}^d$, $d \geqslant 3$ a constant. In some of these problems the queries are known in advance (off-line queries) while in the others the queries are given on-line. All queries are related to measuring the extent of P. One of our main results is a data structure for answering farthest point queries for line segments in $\mathbb{R}^3$. The solution for answering farthest point from line segment queries relies on a solution for the following problem.

Farthest Point From Line (*FPFL*). Preprocess a set P of n points in $\mathbb{R}^d$, where $d \geqslant 3$ is a constant, such that given a query line L one can efficiently report the farthest point of P from L.

We also show how to use our solution to this problem to solve the following problem on indexed points.

Farthest Point—Indexed Set (*FPIS*). Given a set of indexed points $P = \{p_1, p_2, \ldots, p_n\}$ in $\mathbb{R}^d$, where $d \geqslant 3$ is a constant, for each line $L(p_i, p_j)$ defined by two points $p_i, p_j \in P$, $1 \leqslant i < j \leqslant n$, find the farthest point $p_k \in P$ such that $i < k < j$.

Similarly, one may want to evaluate the extent of a data set by computing the minimum-width cylindrical shell with central axis L enclosing the data set P, where a cylindrical shell is the region enclosed between two co-axial cylinders, and the width of a cylindrical shell is defined as the absolute value of the difference of the radii of the two cylinders. If no axis is specified, then the minimum-width cylindrical shell for a set of n points in $\mathbb{R}^3$ can be computed in $O(n^5)$ time [2]. However, if more complex geometric objects are sought to capture the features of P, the problem may become significantly more difficult. For example, one may want to fit a polygonal chain through P, with the condition that the vertices of the polygonal chain are a subset of the points in P and such that it approximates the extent of P based on