The Traveling Salesman Problem, Data Parametrization and Multi-resolution Analysis

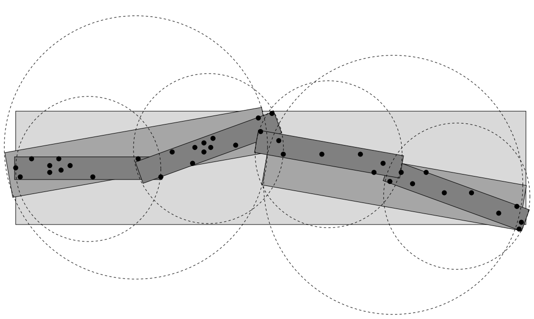
Raanan Schul Stony Brook University, Stony Brook, NY USA
schul@math.sunysb.edu

ABSTRACT

Given a set $K \subset \mathbb{R}^d$, when is it contained in a curve of finite length? How does the answer change if you are willing to capture only part of K? How long are the curves in question? We will discuss some answers to these questions. The set K above can be an infinite set. Furthermore, the ambient dimension d can be taken to be infinite. What is the correct way to generalize the term *curve* to have higher intrinsic dimension? These problems are variants of what is sometimes referred to as the analyst's traveling salesman theorem. We will contrast the analyst's TSP with its CS counterpart in the talk.

Given a Euclidean network of roads, when can we turn it into an *efficient* network by adding *shortcuts*? How much will these shortcuts cost us? By *efficient*, here we mean that the two distances, one being the Euclidean distance, and the other being the distance traveled along the network of roads, are comparable. A recent result says that every Euclidean network is contained inside one of comparable total length, such that the *new network* is itself efficient.

How are these questions related to quantitative or coarse differentiation? Embedding metric spaces into Euclidean space with bounded distortion? Data parametrization and analysis? (In particular, low dimensional data living in a high ambient dimension)



It turns out that there are quantities, which are discrete analogues of the 'average curvature of the set K', or 'the average curvature of a Euclidean network' which help answer and relate all of the above questions. One can obtain these quantities by looking at the principal component analysis of points inside balls at various locations and scales and correctly averaging them. (Or, in the above image, take the eccentricity of each rectangle.) The analysis that is done via a multi-scale and multi-resolution description of K. The above quantities serve as magnitudes of coefficients which are L^2 summable. (In the analyst's TSP case, this sum is, up to constant, the same as the minimal spanning tree length) Indeed, there is also a deep connection between the quantities that appear here and quantities from harmonic analysis, in particular, wavelet analysis. This connection sometimes serves as a guide to proving the results. We will be explaining results from various papers of Jonas Azzam, Guy David, Peter Jones, Gilad Lerman Stephen Semmes and myself.

Mauro Maggioni, another speaker in this session, will speak about other aspects of this connection, as well as how these results are transferred to applications coming from analysis of actual large data sets.

Analysis and CG June 18, 2012, Chapel Hill, North Carolina, USA.

Keywords

Traveling salesman, spanners, shortcuts, curvature

BIO

I did my undergrad at Hebrew University in Jerusalem, I was a graduate student at Yale, and a postdoc at UCLA. I am currently an Assistant Professor at Stony Brook University. My work is on geometric measure theory and harmonic analysis, with some connections to applied mathematics. See `http://www.math.sunysb.edu/~schul`.

References

[1] J. Azzam and R. Schul. A Carleson packing condition for metric differentials Submited. `arxiv:1111.2561`.

[2] J. Azzam and R. Schul. How to take shortcuts in Euclidean space: making a given set into a short quasi-convex set. Proc. London Math. Soc. (2012) `arxiv:0912.1356`.

[3] Raanan Schul. Bi-Lipschitz decomposition of Lipschitz functions into a metric space. *Rev. Mat. Iberoam.*, 25(2):521–531, 2009.

[4] Raanan Schul. Analyst's traveling salesman theorems. A survey. *In the tradition of Ahlfors and Bers, IV*, volume 432 of *Contemp. Math.*, pages 209–220. Amer. Math. Soc., Providence, RI, 2007.

[5] Raanan Schul. Subsets of rectifiable curves in Hilbert space. *Journal d'Analyse Mathématique* 103 (2007), 331-375..