DRAFT: Monday 6th May, 2019, 15:22

Notes on Bounding Box Construction for Filtering Geographic Data

Nick Halmagyi

Laboratoire de Physique Théorique et Hautes Energies,
Université Pierre et Marie Curie, CNRS UMR 7589,
F-75252 Paris Cedex 05, France

halmagyi@lpthe.jussieu.fr

1 Introduction

A common problem is data-science is the following: given a single geographic location S (the source) and a set of geographic locations $\widehat{T}$ (the targets), determine the closest N locations $t_i \in \widehat{T}$ to S. One can achieve this by computing the distance from every target to the source $d(t_i, S)$ but this is prohibitively slow for moderately large set of targets.

A better solution is to make a bounding region $\mathcal{B}$ around S and compute the distance from each $d(t_i, S)$ for each $t_i \in \mathcal{B}$. For well constructed $\mathcal{B}$ one can hope to find the N closest points within $\mathcal{B}$.

Geographic locations are typically given in the longitude-latitude coordinate system

2 Haversine Formula

The Haversine formula gives the distance between two points on the sphere in the latitude-longitude coordinate system:

$$\sin^2\left(\frac{d}{2R}\right) = \sin^2\left(\frac{\varphi_1 - \varphi_2}{2}\right) + \cos\varphi_1 \cos\varphi_2 \sin^2\left(\frac{\lambda_1 - \lambda_2}{2}\right) \tag{1}$$

or more conveniently

$$\cos\frac{d}{R} = \sin\varphi_1\sin\varphi_2 + \cos(\lambda_2-\lambda_1)\cos\varphi_1\cos\varphi_2\,. \tag{2}$$

where R is the radius of the sphere.

2.1 Equal Latitude

When two latitudes are equal

$$\varphi_1 = \varphi_2 = \varphi \tag{3}$$

we have

$$d = 2R\sin^{-1}\left|\cos\varphi\sin\left(\frac{\lambda_1-\lambda_2}{2}\right)\right| \tag{4}$$

$$\Rightarrow \quad |\lambda_1-\lambda_2| = 2\left|\sin^{-1}\left[\frac{\sin\left(\frac{d}{2R}\right)}{\cos(\varphi)}\right]\right|\,. \tag{5}$$

The maximum longitudinal difference between two points is

$$|\lambda_2-\lambda_1|_{\max} = \pi \tag{6}$$

which occurs when

$$\sin\left(\frac{d}{2R}\right) = \cos(\varphi)\,. \tag{7}$$

2.2 Equal Longitudes

When two longitudes are equal

$$\lambda_2 = \lambda_1 \tag{8}$$

the Haversine formula gives

$$\varphi_2 - \varphi_1 = \frac{d}{R} \tag{9}$$

3 Maximum longitude

Given a section of equal distance d from the source φ_1, λ_1 we want to know the maximum of $|\lambda_2 - \lambda_1|$.

Lets first look at an ellipse

$$a^2x^2 + b^2y^2 = 1 \tag{10}$$

solve for x

$$x(y) = \frac{1}{a}\sqrt{(1 - b^2y^2)}\,, \qquad x'(y) = \frac{1}{a}\frac{2b^2y}{(1 - b^2y^2)} \tag{11}$$

so the extremum are at $y = 0$.

So we fix (φ_1, λ_1) and find the extremum of λ_2. We change coordinates to

$$\Lambda = \cos\big(\lambda_1 - \lambda_2(\varphi_2)\big)\,, \qquad \Psi = \sin\varphi_2 \tag{12}$$

and find that the equations to be solved are

$$\csc\varphi_1 \cos\frac{d}{R} = \cot\varphi_1 \Lambda\sqrt{1 - \Psi^2} + \Psi \tag{13}$$

$$\cot\varphi_1 \Lambda\Psi = \sqrt{1 - \Psi^2} \tag{14}$$

The solution is

$$\Lambda = \sec(\varphi_1)\sqrt{\cos^2\big(\frac{d}{R}\big) - \sin^2(\varphi_1)} \tag{15}$$

$$\Psi = \sec\big(\frac{d}{R}\big)\sin\varphi_1 \tag{16}$$

There is a branch point at one of the following points:

$$\text{b.p.} \quad \varphi_1 \pm \frac{d}{R} = \pm\pi/2 \tag{17}$$

which is where the bounding section touching the north(south) pole. At this point the maximum longitude reached is $\lambda_2 - \lambda_1 = \pi/2$ which is half way around the glode (though not necessarily at the equator)

4 Intersection with $\lambda_2 - \lambda_1 = \frac{\pi}{2}$

We want to know, when

$$\varphi_2 - \varphi_1 > \frac{d}{R} \tag{18}$$

so that the bounding circle does not reach the pole. Can the bounding circle extend more than $\pi/2$ in longitude. When $\lambda_2 - \lambda_1 = \pi/2$ we have

$$\cos\frac{d}{R} = \sin\varphi_1 \sin\varphi_2 \tag{19}$$

which will give a solution when

$$\cos\frac{d}{R} < \sin\varphi_1 \tag{20}$$