SPEED DISTRIBUTION FROM NORMALLY DISTRIBUTED LOCATION MEASUREMENTS

John Krumm
Microsoft Research
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052
jckrumm@microsoft.com

ABSTRACT

This article describes how to compute a probability distribution of speed from a pair of uncertain location measurements taken at different times. When location measurements are uncertain, the method shows how to propagate this uncertainty to the resulting speed computation. We assume the location measurements are normally distributed. While the resulting speed distribution is not normal, it can be approximated as normal. The article derives the speed distribution, verifies it with a simulation, and gives a numerical example for debugging.

Keywords location measurement · speed distribution · chi distribution

1 Introduction

A location trajectory is a sequence of time-stamped location measurements of a moving object. A common task in analyzing location trajectories is computing speed. This is simple if we ignore the uncertainty in the location measurements, but there are cases where it is important to understand the uncertainty in speed, such as assessing traffic flow and crash risk.

This article presents principled methods for computing probability distributions of speed given time-stamped location measurements. The basic elements of this computation are measurements of time and distance. We assume time can be measured with no uncertainty[1], but that distance is uncertain due to errors in location measurements. Specifically, we begin with two locations represented by two-dimensional vectors $\mathbf{x}_0 = (x_0, y_0)^T$ and $\mathbf{x}_1 = (x_1, y_1)^T$ separated by time T. We will assume that the actual $\mathbf{x_i}$ are distributed as a bivariate normals with mean vector $\mu = (\mu_{i_x}, \mu_{i_y})^T$ and diagonal 2×2 covariance matrix $\sigma^2 I$. For a GPS point this is a workable model [1], although this ignores any correlation between the points, assuming they are independent. The mean μ would represent the actual GPS measurement and a reasonable value for σ could be three meters.

We have $\mathbf{X}_0 \sim \mathcal{N}(\mu_0, \sigma^2 I)$ and $\mathbf{X}_1 \sim \mathcal{N}(\mu_1, \sigma^2 I)$. Assuming the motion between the two points was along a straight line segment, the velocity is distributed as

$$\frac{\mathbf{X}_1 - \mathbf{X}_0}{T} \sim \mathcal{N}\left(\frac{\mu_\mathbf{1} - \mu_\mathbf{0}}{T}, \frac{2\sigma^2}{T^2} I\right) \tag{1}$$

[1] One of the most accurate methods for measuring time is GPS, although there is still some error. According to https://www.atomic-clock.galleon.eu.com/support/gps-time-accuracy.html, the error can be as little as 10 nanoseconds, but can grow to 1 microsecond depending on the GPS receiver. This could make a difference for fast-moving objects. For instance, according to https://en.wikipedia.org/wiki/Ballistic_missile, a ballistic missile may reach a reentry speed of 8000 meters per second. In 10 nanoseconds at that speed, the missile would travel 0.08 millimeters, and in 1 microsecond it would travel 0.8 centimeters.

Although this gives a distribution for the bivariate velocity vector, we want a distribution for the scalar speed, described next.

2 Speed Distribution

From Equation 1, the x and y components of velocity are statistically independent with distributions $V_x \sim \mathcal{N}(\mu_{v_x}, \sigma_v)$ and $V_y \sim \mathcal{N}(\mu_{v_y}, \sigma_v)$, with

$$\begin{aligned}\mu_{v_x} &= (\mu_{1x} - \mu_{0x})/T \\ \mu_{v_y} &= (\mu_{1y} - \mu_{0y})/T \\ \sigma_v^2 &= 2\sigma^2/T^2\end{aligned}$$

The statistic

$$Z = \sqrt{\frac{V_x^2 + V_y^2}{\sigma_v^2}}$$

is distributed according to the noncentral chi distribution [2]. (This is different from the more familiar chi-squared distribution.) The probability distribution function for this statistic is

$$f_z(z) = z e^{-0.5(z^2+\lambda^2)} I_0(\lambda z) \tag{2}$$

for $z \geq 0$, where

$$\lambda = \frac{1}{\sigma_v}\sqrt{\mu_{v_x}^2 + \mu_{v_y}^2}$$

and $I_0(x)$ is a modified Bessel function of the first kind'[3].

The mean and variance of this distribution are

$$\mu_z = \sqrt{\frac{\pi}{2}} L_{1/2}^0\left(\frac{-\lambda^2}{2}\right) \tag{3}$$

$$\sigma_z^2 = 2 + \lambda^2 - \mu_z^2 \tag{4}$$

where $L_{1/2}^0(x)$ is a generalized Laguerre function [4].

Instead of the statistic Z, we are interested in speed $S = \sigma_v Z$. Transforming variables, the distribution of speed is

$$f_s(s) = \frac{1}{\sigma_v} f_z\left(\frac{s}{\sigma_v}\right) \tag{5}$$

for $s \geq 0$.

The mean and variance of speed are

$$\mu_s = \sigma_v \mu_z \tag{6}$$

$$\sigma_s^2 = \sigma_v^2 \sigma_z^2 \tag{7}$$

3 Numerical Example and Normal Approximation

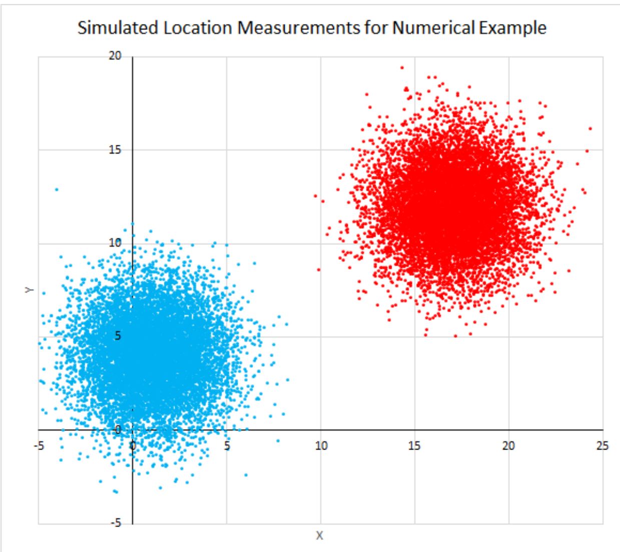
As an example for debugging, we have $\mathbf{x}_0 = (1, 4)^T$, $\mathbf{x}_1 = (17, 12)^T$, $\sigma = 2$, $T = 8$, and $s = 2$. The units are standard: meters for distances and seconds for time. A simulation of 10,000 of both measured points is shown in Figure 1a. We have $\mu_{\mathbf{v}} = (2, 1)$, $\sigma_v^2 = 0.125$, and $\lambda = 6.3246$.

For computing the PDF value $f_s(2)$, we begin with $I_0(\lambda \frac{2}{\sigma_v})$ =$\mathrm{I}_0(35.7771) = 2.3091 \times 10^{14}$. Then $f_s(2) = 0.8570$ from Equations 5 and 2. The full PDF is shown as the thick, gray curve in Figure 1b.

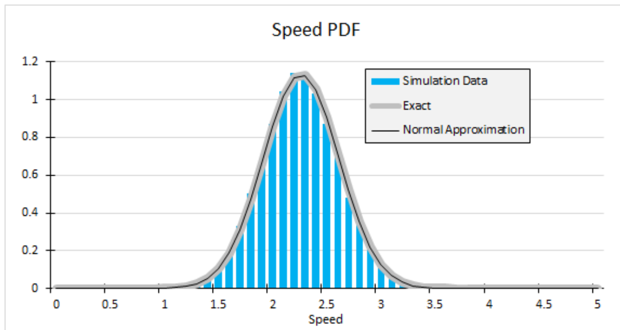
For the mean speed, we have $L^0_{1/2}(-\lambda^2/2) = L^0_{1/2}(-20) = 5.1098$. The mean and variance of the speed distribution, from Equations 6, 7, 3, and 4 are

$$\mu_v = 2.2642$$
$$\sigma_v^2 = 0.1234$$

Naively computing the mean speed gives $||\mathbf{x}_1 - \mathbf{x_0}||/T = 2.2361$, which is close to μ_v.



(a) Simulated location measurements for numerical example



(b) Simulated (bars), exact, and approximate distributions

Figure 1: Simulated data and distributions

The blue bars in Figure 1b shows a histogram of the speeds computed from the 10,000 pairs of points in Figure 1a. The simulated results closely match the exact distribution.

Figure 1b also shows a normal approximation to the exact distribution. The normal approximation uses the same mean and variance as the exact distribution. Visually, the two distributions are indistinguishable. This approximation is valid as long as the mass of the exact distribution does not get too close to zero. The exact distribution is only defined for non-negative speeds, while the normal distribution has support all along the real line. Approximating the exact distribution with a normal can be useful for subsequent calculations that require a normal distribution of speeds, such as a standard Kalman filter.

References

[1] Frank Van Diggelen. Update: Gnss accuracy: Lies, damn lies, and statistics. *GPS world*, 18(1):26–33, 2007.

[2] Wikipedia contributors. Noncentral chi distribution — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Noncentral_chi_distribution&oldid=872282425`, 2018. [Online; accessed 18-April-2019].

[3] Wikipedia contributors. Bessel function — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/wiki/Bessel_function`, 2019. [Online; accessed 5-November-2019].

[4] Wikipedia contributors. Laguerre polynomials — Wikipedia, the free encyclopedia. https://en.wikipedia.org/wiki/Laguerre_polynomial, 2019. [Online; accessed 5-November-2019].