Benjamin Hill

Brother Martin

MATH 411

9 December 2023

Reality Check 4: GPS Errors

The Global Positioning System (GPS) consists of 24 satellites carrying atomic clocks, orbiting the earth at an altitude of 20200 km. At any time, from any point on earth, five to eight satellites are in direct line of sight. Each satellite has a simple mission: to transmit carefully synchronized signals from predetermined positions in space, to be picked up by GPS receivers on earth. The receivers use the information, with some mathematics (described shortly), to determine accurate $(x,\ y,\ z)$ coordinates of the receiver.

At a given instant, the receiver collects the synchronized signal from the i^{th} satellite and determines its transmission time t_i, the difference between the time the signal was transmitted and the time it was received. The nominal speed of the signal is the speed of light, $c = 299792.458$ km/sec. Multiplying transmission time by c gives the distance of the satellite from the receiver, putting the receiver on the surface of a sphere centered at the satellite position and with radius ct_i. If three satellites are available, then three spheres are known, whose intersection consists of two points. One intersection point is the location of the receiver. The other is normally far from the earth's surface and can be safely disregarded. In theory, the problem is reduced to computing this intersection, the common solution of three sphere equations.

However, there is a major problem with this analysis. First, although the transmissions from the satellites are timed nearly to the nanosecond by onboard atomic clocks, the clock in the typical low-cost receiver on earth has relatively poor accuracy. If we solve the three equations with slightly inaccurate timing, the calculated position could be wrong by several kilometers. Fortunately, there is a way to fix this problem. The price to pay is one extra satellite. Define d to be the difference between the synchronized time on the (now four) satellite clocks and the earth-bound receiver clock. Denote the location of satellite i by $(A_i,\ B_i,\ C_i)$. Then the true intersection point $(x,\ y,\ z)$ satisfies the following system (which will be referred to as system (4.37))

$$r_1(x,y,z,d) = \sqrt{(x-A_1)^2 + (y-B_1)^2 + (z-C_1)^2} - c(t_1 - d) = 0$$

$$r_2(x,y,z,d) = \sqrt{(x-A_2)^2 + (y-B_2)^2 + (z-C_2)^2} - c(t_2 - d) = 0$$

$$r_3(x,y,z,d) = \sqrt{(x-A_3)^2 + (y-B_3)^2 + (z-C_3)^2} - c(t_3 - d) = 0$$

$$r_4(x,y,z,d) = \sqrt{(x-A_4)^2 + (y-B_4)^2 + (z-C_4)^2} - c(t_4 - d) = 0$$

to be solved for the unknowns x, y, z, d. Solving the system reveals not only the receiver location, but also the correct time from the satellite clocks, due to knowing d. Therefore, the inaccuracy in the GPS receiver clock can be fixed by using one extra satellite.

The system of equations above can be seen to have two solutions $(x,\ y,\ z,\ d)$. The equations can be equivalently written

$$(x - A_1)^2 + (y - B_1)^2 + (z - C_1)^2 = c^2(t_1 - d)^2$$
$$(x - A_2)^2 + (y - B_2)^2 + (z - C_2)^2 = c^2(t_2 - d)^2$$
$$(x - A_3)^2 + (y - B_3)^2 + (z - C_3)^2 = c^2(t_3 - d)^2$$
$$(x - A_4)^2 + (y - B_4)^2 + (z - C_4)^2 = c^2(t_4 - d)^2.$$

This system will be referred to as system (4.38). Note that by subtracting the last three equations from the first in turn, three linear equations are obtained. Each linear equation can be used to eliminate a variable x, y, z, and by substituting into any of the original equations, a quadratic equation in the single variable d results. Therefore, system (4.37) has at most two real solutions, and they can be found by the quadratic formula.

Two further problems emerge when GPS is deployed. First is the conditioning of the system of equations (4.37). We will find that solving for $(x,\ y,\ z,\ d)$ is ill-conditioned when the satellites are bunched closely in the sky.

The second difficulty is that the transmission speed of the signals is not precisely c. The signals pass through 100 km of ionosphere and 10 km of troposphere, whose electromagnetic properties may affect the transmission speed. Furthermore, the signals may encounter obstacles on earth before reaching the receiver, an effect called multipath interference. To the extent that these obstacles have an equal impact on each satellite path, introducing the time correction d on the right side of (4.37) helps. In general, however, this assumption is not viable and will lead us to add information from more satellites and consider applying Gauss-Newton to solve a least squares problem.

Consider a three-dimensional coordinate system whose origin is the center of the earth (radius = 6370 km). GPS receivers convert these coordinates into latitude, longitude, and elevation data for readout and more sophisticated mapping applications using global information systems (GIS), a process we will not consider here.

Part 1

First, we will solve the system (4.37) by using the Python multivariate root finder, SciPy's "root" function. We will find the receiver position (x, y, z) near earth and time correction d for known, simultaneous satellite positions (15600, 7540, 20140), (18760, 2750, 18610), (17610, 14630, 13480), (19170, 610, 18390) in km, and measured time intervals 0.07074, 0.07220, 0.07690, 0.07242 in seconds, respectively. With the initial vector set to be (x0, y0, z0, d0) = (0, 0, 6370, 0). The SciPy root finder yields the position and time stamp $(x, y, z, d) = (-41.7727096, -16.7891941, 6370.05956, -0.00320156583)$, which is as expected. This is a point very near the earth's geographical north pole.

Part 2

We will write a Python program to carry out the solution via the quadratic formula. A significant chunk of algebra is required to get these equations into the format of a quadratic equation, so we will work them into that format first. As noted in the introduction, we can obtain the necessary quadratic equations by subtracting each equation in turn from the first equation in system (4.37). This results in the following three equations:

$$2(A_2 - A_1)x + 2(B_2 - B_1)y + 2(C_2 - C_1)z + 2c^2(t_1 - t_2)d = \left({A_2}^2 - {A_1}^2\right) + \left({B_2}^2 - {B_1}^2\right) + \left({C_2}^2 - {C_1}^2\right) + c^2({t_1}^2 - {t_2}^2)$$

$$2(A_3 - A_1)x + 2(B_3 - B_1)y + 2(C_3 - C_1)z + 2c^2(t_1 - t_3)d = \left({A_3}^2 - {A_1}^2\right) + \left({B_3}^2 - {B_1}^2\right) + \left({C_3}^2 - {C_1}^2\right) + c^2({t_1}^2 - {t_3}^2)$$

$$2(A_4 - A_1)x + 2(B_4 - B_1)y + 2(C_4 - C_1)z + 2c^2(t_1 - t_4)d = \left({A_4}^2 - {A_1}^2\right) + \left({B_4}^2 - {B_1}^2\right) + \left({C_4}^2 - {C_1}^2\right) + c^2({t_1}^2 - {t_4}^2).$$

Luckily, this eyesore can then be rewritten into the vector equation

$$x\begin{bmatrix} 2(A_2 - A_1) \\ 2(A_3 - A_1) \\ 2(A_4 - A_1) \end{bmatrix} + y\begin{bmatrix} 2(B_2 - B_1) \\ 2(B_3 - B_1) \\ 2(B_4 - B_1) \end{bmatrix} + z\begin{bmatrix} 2(C_2 - C_1) \\ 2(C_3 - C_1) \\ 2(C_4 - C_1) \end{bmatrix} + d\begin{bmatrix} 2c^2(t_1 - t_2) \\ 2c^2(t_1 - t_3) \\ 2c^2(t_1 - t_4) \end{bmatrix} = \begin{bmatrix} ({A_2}^2 - {A_1}^2) + ({B_2}^2 - {B_1}^2) + ({C_2}^2 - {C_1}^2) + c^2({t_1}^2 - {t_2}^2) \\ ({A_3}^2 - {A_1}^2) + ({B_3}^2 - {B_1}^2) + ({C_3}^2 - {C_1}^2) + c^2({t_1}^2 - {t_3}^2) \\ ({A_4}^2 - {A_1}^2) + ({B_4}^2 - {B_1}^2) + ({C_4}^2 - {C_1}^2) + c^2({t_1}^2 - {t_4}^2) \end{bmatrix}$$

which can in turn be rewritten as

$$x\vec{U}_x + y\vec{U}_y + z\vec{U}_z + d\vec{U}_d = \vec{w}$$

with (x, y, z, d) being the only unknowns. Finally, we combine the vectors in this equation into an augmented matrix and row-reduce it to find the solution in terms of d, as follows:

$$\begin{bmatrix} \vec{U}_x & \vec{U}_y & \vec{U}_z & \vec{U}_d & \vec{w} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -10.749 & -41.738 \\ 0 & 1 & 0 & 623.677 & -18.786 \\ 0 & 0 & 1 & 83788.807 & 6101.804 \end{bmatrix}.$$

This result means that we have, in terms of d, $x = -41.738 + 10.749d$, $y = -18.786 - 623.677d$, and $z = 6101.804 - 83788.807d$. Substituting these values into the first equation of (4.37), we find that the quadratic equation we sought, in terms of d, is $4.91198 * 10^7 + 1.50772 * 10^{10}d - 8.28546 * 10^{10}d^2 = 0$. Using the quadratic formula, this means that $d = -0.003202, 0.185173$. The first value of d matches our solution from part 1, so we know that $d = -0.003202$. Substituting this value into the equations above for the receiver coordinates in terms of d, we find that $(x, y, z) = (-41.7727, -16.7892, 6370.0590)$, which is once again the same as the solution found in part 1.

Part 3 is essentially the same as Part 2 and is therefore skipped here.

Part 4

Now set up a test of the conditioning of the GPS problem. Define satellite positions (A_i, B_i, C_i) from spherical coordinates $(\rho, \varphi_i, \theta_i)$ as

$$A_i = \rho\cos(\varphi_i)\cos(\theta_i)$$

$$B_i = \rho\cos(\varphi_i)\sin(\theta_i)$$

$$C_i = \rho\sin(\varphi_i)$$

where $\rho = 26570$ km is fixed, while $0 \le \varphi_i \le \pi/2$ and $0 \le \theta_i \le 2\pi$ for $i = 1, \dots, 4$ are chosen arbitrarily. The φ coordinate is restricted so that the four satellites are in the upper hemisphere. We will set $x = 0, y = 0, z = 6370, d = 0.0001$, and calculate the corresponding satellite ranges

$$R_i = \sqrt{{A_i}^2 + {B_i}^2 + (C_i - 6370)^2}$$

and travel times $t_i = d + R_i/c$.

We choose the satellites positions such that the first is over the north pole and the other three are equidistant around the equator, which positions are given by

$$(\varphi_1, \theta_1) = \left(\frac{\pi}{2}, 0\right)$$

$$(\varphi_2, \theta_2) = (0, 0)$$

$$(\varphi_3, \theta_3) = \left(0, \frac{\pi}{3}\right)$$

$$(\varphi_4, \theta_4) = \left(0, \frac{2\pi}{3}\right).$$

We will define an error magnification factor specially tailored to the situation. The atomic clocks aboard the satellites are correct up to about 10^{-8} seconds. Therefore, it is important to study the effect of changes in the transmission time of this magnitude. Let the backward, or input error be the input change in meters. At the speed of light, $\Delta t_i = 10^{-8}$ seconds corresponds to $10^{-8}s \approx 3$ meters. Let the forward, or output error be the change in position $||\Delta x, \Delta y, \Delta z||_\infty$, caused by such a change in t_i, also in meters. Then we can define the dimensionless error magnification factor (EMF) as $EMF = \frac{||\Delta x, \Delta y, \Delta z||_\infty}{c||\Delta t_i, \dots, \Delta t_m||_\infty}$, and the condition number of the problem to be the maximum error magnification factor for all small Δt_i (say, 10^{-8} or less). We will change each t_i defined in the foregoing by $\Delta t_i = \pm 10^{-8}$, not all the same. For this instance, we will define $(\Delta t_1, \Delta t_2, \Delta t_3, \Delta t_4) = (0.5 * 10^{-8}, 10^{-8}, -10^{-8}, -0.5 * 10^{-8})$. We will denote the new solution of the equations (4.37) by (x, y, z, d). Under these parameters, we compute the error in position to be $||\Delta x, \Delta y, \Delta z||_\infty = 0.005995$ and compute the error magnification factor to be $EMF = 0.99999999912$. Trying four other similar variations of the Δt_i values, the maximum position error found is $EMF = 1.000000002$. Based on the error magnification factors we have computed, we estimate the condition number of the system to be approximately 1.

Part 5

Now we will repeat part 4 with a very tightly grouped set of satellites. Let the four satellite positions be given by

$$(\varphi_1, \theta_1) = (0.01, 0)$$

$$(\varphi_2, \theta_2) = (0.03, 0.03)$$

$$(\varphi_3, \theta_3) = (0.05, 0.01)$$

$$(\varphi_4, \theta_4) = (0.05, 0.05).$$

With the same errors Δt_i as in part 4, we find the max position error to be 7060.9 km, which is enormous – that error is larger than the radius of the earth! Likewise, the max EMF is extremely large, being 4710526.5. Therefore the condition number of this system with the satellites so tightly grouped is approximately $4.7 * 10^6$, which is more than six orders of magnitude larger than the condition number for the spread-out satellites. From this comparison we can see clearly why it is necessary to ensure the satellites are spread out far enough. The tighter the grouping of the satellites, the higher the condition number of the system.