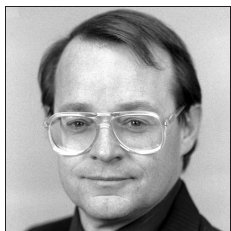# An Underdetermined Linear System for GPS

Dan Kalman

Dan Kalman (kalman@american.edu) joined the mathematics faculty at American University in 1993, following an eight year stint in the aerospace industry and earlier teaching positions in Wisconsin and South Dakota. He has won three MAA writing awards and served a term as Associate Executive Director of the MAA. His interests include matrix algebra, curriculum development, and interactive computer environments for exploring mathematics, especially using Mathwright software.

Finding the general solution to an underdetermined linear system is a standard topic in linear algebra. It contributes to a complete analysis of the behaviors of linear systems, as well as providing a foundation for understanding more abstract topics, including linear transformations, null space, and dimension. But at the first introduction of the topic it would be nice to have a simple, realistic example where the parameterized general solution of an underdetermined system is of practical interest. In this note, I will present such an example connected with the Global Positioning System (GPS) for determining geographical locations.

The basic idea of GPS is a variant on three dimensional triangulation: a point on the surface of the earth is determined by its distances from three other points. Here, the point we wish to determine is the location of the GPS receiver, the other points are satellites, and the distances are computed using the travel times of radio signals. This requires accurate time keeping, prompting a slight modification of the pure spatial triangulation problem. In the modified version, we need four satellites, rather than three, and can then calculate both the location, and the correct time, at the GPS receiver.

Before presenting the example, I should make it clear that the computations that follow are not the same as the methods actually used by GPS. The example assumes exact geometric knowledge, whereas GPS has to deal with real world measurement errors. Thus, GPS typically uses more than four satellites, and a least-squares method to determine the best estimate of the location and time at the receiver. Other refinements in the actual GPS calculations take into account the way a radio signal is impeded by passing through the atmosphere, and the actual encoding of information in the radio signal. For an accurate overview of how the GPS system actually works, see [2]. There is quite a bit of linear algebra involved, and complete details are presented in [3]. Reference [4] provides an additional account of the mathematics involved in GPS computations.

Although the formulation I will present does not represent the methods used by GPS, it does provide a good starting point for understanding the GPS system. For example, the exact geometric model can be perturbed by random errors to simulate the effects of errors in the real system. This idea will be considered briefly below.

The Geometric Model. For concreteness, consider a ship at sea in an unknown location. It has a GPS receiver that obtains simultaneous signals from four satellites. Each signal specifies its time of transmission and the position of the satellite at that time. This allows the GPS receiver to compute its position and time. How does this work?

To begin with, we imagine that there is an xyz-coordinate system with the earth centered at the origin, the positive z axis running through the north pole and fixed relative to the earth. The unknown position of the ship can be expressed as a point (x, y, z), which can later be translated into a latitude and longitude. To simplify things, let us mark off the three axes in units equal to the radius of the earth. Thus, a point at sea level will have $x^2 + y^2 + z^2 = 1$ in this system. Also, we will measure time in units of milliseconds. The GPS system finds distances by knowing how long it takes a radio signal to get from one point to another. For this we need to know the speed of light, approximately equal to .047 (in units of earth radii per millisecond).

Our ship is at an unknown position and has no clock. It receives simultaneous signals from four satellites, giving their positions and times as shown in Table 1. (These numbers were made up for the example; in a real case the satellite positions would not be such simple vectors).

Table 1. Satellite data.

Satellite	Position	Time
1	(1, 2, 0)	19.9
2	(2, 0, 2)	2.4
3	(1, 1, 1)	32.6
4	(2, 1, 0)	19.9

Let (x, y, z) be the ship's position and t the time when the signals arrive. Our goal is to determine the values of these variables. Using the data from the first satellite, we can compute the distance from the ship as follows. The signal was sent at time 19.9 and arrived at time t. Traveling at a speed of .047, that makes the distance

$$d = .047(t - 19.9).$$

This same distance can be expressed in terms of (x, y, z) and the satellite's position $(1, 2, 0)$:

$$d = \sqrt{(x - 1)^2 + (y - 2)^2 + (z - 0)^2}.$$

Combining these results leads to the equation

$$(x - 1)^2 + (y - 2)^2 + z^2 = .047^2(t - 19.9)^2. \tag{1}$$

Similarly, we can derive a corresponding equation for each of the other three satellites. That gives us four equations in four unknowns, and so we can solve for x, y, z and t. These are not linear equations, but we can use algebra to obtain a linear system that we can solve. Before proceeding, however, let us consider the geometry a bit more.

Pure spatial triangulation has a simple geometric representation. Specifying the distance from an unknown position P to a known fixed point F restricts P to lie on a sphere centered at F. Given the distances to three points, F_1, F_2, F_3, we conclude that P must lie at a point of intersection of three spheres. The intersection is typically a pair of points. To see this, observe that the intersection of any two of the spheres is a circle, which (except for degenerate cases) meets the third sphere in two points. Moreover, there is a simple characterization of the points. Consider the plane that contains

the intersection circle of the first two spheres. It contains the points where all three spheres meet. These points are similarly contained in the plane of intersection for the second and third spheres. The intersection of the two planes is a line, and the points we seek are the intersections of that line with any of the three spheres. Ultimately we hope to eliminate one of these intersection points. For GPS, knowing that the receiver is on the surface of the earth permits this final step.

For the GPS problem, triangulation takes place in both space and time, complicating the geometry somewhat. As a simplification, let us look at triangulation in a plane, rather than in space. A receiver at an unknown position P obtains simultaneous signals from three sources. The locations of the sources and the times at which the signals were broadcast are known. Locating P can be visualized dynamically. Imagine each signal radiating outward in a circular wave. There are three of these waves, initiated at three different times, as if three pebbles were cast into a still pond. We are trying to locate the point at which the outermost ripples from each stone first meet. This view is captured in several snapshots in Figure 1.
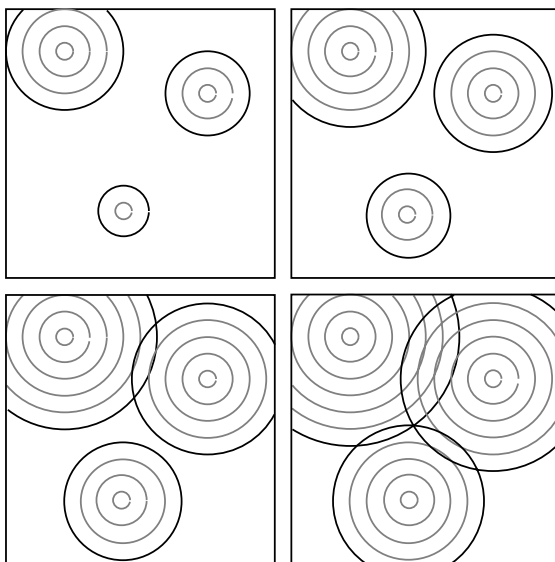

Figure 1. Ripples radiate from three stones.

An alternative to this dynamic visualization is obtained by considering a three dimensional *space-time*. The horizontal plane represents spatial positions, and the vertical axis represents time. Each point (x, y, t) thus identifies a specific location and time. In this setting, the snapshots of Figure 1 are horizontal sections of a three dimensional figure. For each signal there is a cone. A point (x, y, t) is on that cone if the signal arrives at point (x, y) at time t. The vertex of the cone is at (x_0, y_0, t_0), where t_0 indicates when the signal was broadcast, and (x_0, y_0) is the location from which the signal was sent. Our problem is to find a point of intersection of three distinct *light cones*.

As depicted in Figure 2, the intersection of two cones lies in a plane. Arguing as before, with three cones we can form two planes of intersection. These meet in a line
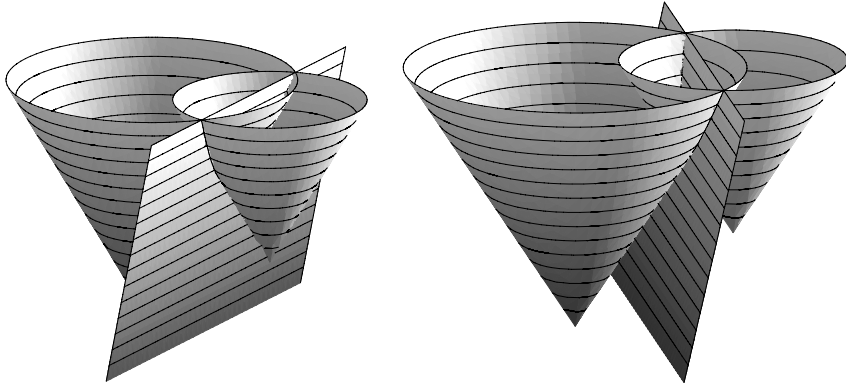
© THE MATHEMATICAL ASSOCIATION OF AMERICA

Figure 2. Two intersecting cones.

which must contain the points common to all three cones. Intersecting the line with any one cone then determines the points we want.

This geometric picture is an exact analog of the GPS problem. Indeed, we may view (1) as a cone in four dimensional space-time. With four satellites there will be a system of four similar equations, and the solutions are points of intersection of four light cones. As in the three dimensional case, we can intersect the cones in pairs to identify (hyper)planes, and the intersection of three of these specifies a line. Algebraically, this corresponds to solving an underdetermined system of linear equations, and the line is the general set of solutions. We will look at the details of the algebra next.

Algebraic Solution. Let us focus again on the equation for the first satellite:

$$(x - 1)^2 + (y - 2)^2 + z^2 = .047^2(t - 19.9)^2.$$

Expanding all the squares and rearranging leads to this version:

$$2x + 4y - 2(.047^2)(19.9)t = 1^2 + 2^2 - .047^2(19.9)^2 + x^2 + y^2 + z^2 - .047^2t^2.$$

Similar equations can be derived for the three other satellites. Writing all four equations together gives

$$2x + 4y + 0z - 2(.047^2)(19.9)t = 1^2 + 2^2 + 0^2 - .047^2(19.9)^2$$
$$+ x^2 + y^2 + z^2 - .047^2t^2$$
$$4x + 0y + 4z - 2(.047^2)(2.4)t = 2^2 + 0^2 + 2^2 - .047^2(2.4)^2$$
$$+ x^2 + y^2 + z^2 - .047^2t^2$$
$$2x + 2y + 2z - 2(.047^2)(32.6)t = 1^2 + 1^2 + 1^2 - .047^2(32.6)^2$$
$$+ x^2 + y^2 + z^2 - .047^2t^2$$
$$4x + 2y + 0z - 2(.047^2)(19.9)t = 2^2 + 1^2 + 0^2 - .047^2(19.9)^2$$
$$+ x^2 + y^2 + z^2 - .047^2t^2$$

The quadratic terms in all the equations are the same, so by subtracting the first equation from each of the other three, we obtain a system of three linear equations:

$$2x - 4y + 4z + 2(.047^2)(17.5)t = 8 - 5 + .047^2(19.9^2 - 2.4^2)$$
$$0x - 2y + 2z - 2(.047^2)(12.7)t = 3 - 5 + .047^2(19.9^2 - 32.6^2)$$
$$2x - 2y + 0z + 2(.047^2)(0)t = 5 - 5 + .047^2(19.9^2 - 19.9^2)$$

Geometrically, each of these equations represents a hyperplane containing the intersection of two light cones.

Now we know that this system cannot have a unique solution. But if the satellite data are accurate, there must be a solution to the original system of quadratic equations, and this linear system must be consistent. By deriving the general solution, it will be possible to express three of the unknowns in terms of the fourth. Then, substitution in one of the original quadratic equations will produce a quadratic equation in one variable. Solving that will lead, in turn, to values for the other three variables.

So, proceeding according to this plan, we formulate the linear system as an augmented matrix:

$$\begin{bmatrix} 2 & -4 & 4 & .077 & 3.86 \\ 0 & -2 & 2 & -.056 & -3.47 \\ 2 & -2 & 0 & 0 & 0 \end{bmatrix}.$$

In this matrix, the integer entries should be considered to be exact values. The other entries were computed to approximately sixteen-place accuracy (using a similarly accurate value for the speed of light), and then rounded to three decimal digits. This is a convenience for the sake of appearances that will be adhered to through all the following calculations. In general, all of the computations are carried through with full accuracy, but appear with only a few decimal digits in print.

Continuing with the solution of the system, the reduced row echelon form for the augmented matrix is

$$\begin{bmatrix} 1 & 0 & 0 & .095 & 5.41 \\ 0 & 1 & 0 & .095 & 5.41 \\ 0 & 0 & 1 & .067 & 3.67 \end{bmatrix}.$$

Therefore, we have the general solution

$$x = 5.41 - .095t, \quad y = 5.41 - .095t, \quad z = 3.67 - .067t, \quad t \text{ free.}$$

Returning to (1), and substituting the above expressions for x, y, and z, we obtain

$$(5.41 - .095t - 1)^2 + (5.41 - .095t - 2)^2 + (3.67 - .067t)^2 = .047^2(t - 19.9)^2$$

or

$$0.02t^2 - 1.88t + 43.56 = 0,$$

leading to two solutions, 43.1 and 50.0. If we select the first solution, then $(x, y, z) = (1.317, 1.317, 0.790)$, which has a length of about 2. We are using units of earth radii, so this point is around 4000 miles above the surface of the earth. The second value of t leads to $(x, y, z) = (.667, .667, .332)$, with length 0.9997. That places the point on the surface of the earth (to four decimal places) and gives us the location of the ship.

Of course, to use this information, we would most likely want to convert it to a latitude and longitude, but that computation will not be presented here.

To double check the accuracy of these results, they can be substituted in the original four quadratic equations. For this particular example, however, I concocted the original problem by starting with assumed values for the unknowns, $(2/3, 2/3, 1/3)$ for (x, y, z) and 50 for t, as well as the positions for the satellites. Using these values I determined the length of time it would take the ship to receive a signal from each satellite. That is how I obtained the times in the last column of Table 1. Rounding those times to one decimal place introduced some small errors, so that the computed solution for (x, y, z) is not quite equal to $(2/3, 2/3, 1/3)$. Still, the computed results came out close enough to the *correct* values to provide some confidence in the methods.

Using this approach, it is possible to create other examples for use in a linear algebra class. You can place the satellites and the ship in any positions you wish. Of course, it is important that the satellites be above the horizon as viewed from the ship. (It is a nice exercise to show that a satellite at position **s** is above the horizon from a point **p** on the unit sphere if and only if $\mathbf{s} \cdot \mathbf{p} > 1$.) Alternatively, you can specify values for all of the data in Table 1, but then there is no guarantee that the GPS receiver is on the surface of the earth.

Accuracy Estimates. The purpose of this example has been to present a realistic situation in which finding the general solution of an underdetermined system plays an important role. As noted earlier, GPS receivers do not operate along the lines of the example. While the example is therefore not an accurate model of computations performed by GPS receivers, it nevertheless serves as a practical conceptual tool for analyzing GPS performance. For instance, the example gives insight about the process of spatial triangulation, and the general framework for GPS. It is the sort of idealized computation that one might want to make to get a feel for the relationships, and while there are many ways to solve the original system of quadratic equations, the one presented above is certainly efficient and practical.

In addition, the simple geometric model provides a context in which other aspects of the system can be studied. One important example concerns the relationship between uncertainties in the measured data (satellite positions and times) and the resultant uncertainties in the GPS results. It turns out that this relationship is dependent on the geometry of the satellites and the GPS receiver [1].

As an illustration of this idea, consider again starting with hypothesized positions for the satellites as well as the position and time data for the ship, and computing the appropriate time data for the satellites. This gives us a consistent set of *true* data. Now to simulate the effects of measurement error, add a random perturbation to each of the satellite data. This gives us a perturbed version of Table 1. Based on these perturbed values, solve the equations to compute a predicted position and time for the ship. Comparing these computed values with the prespecified *true* values shows the effects of the simulated errors in the satellite data. For the example above, adding errors on the order of .001 (about 4 miles spatially and 10^{-6} seconds in time) introduces errors on the order of 5 miles in the computed position of the ship. Similarly, perturbations of around 10^{-6} in the satellite data causes errors of around 10 yards for the computed position of the ship.

Let's consider a different geometric configuration, by changing the location of the fourth satellite to $(1.1, .9, 1.2)$, fairly close to satellite 3. This time, perturbing the satellite data by about .001 results in errors around 50 miles in the computed ship position—10 times more than in the first case. A similar tenfold increase in error is observed when the satellite data are perturbed by about 10^{-6}. This illustrates that the

impact of errors in the satellite data depends strongly on the geometric configuration. It also shows how computations in the conceptual geometric model contribute to an understanding of the more complicated real system.

This final discussion would probably not be appropriate in most linear algebra classes. It would require too great a digression from the main ideas of the course, and is not really necessary for understanding the use of linear algebra in the simplified geometric model presented above. Indeed, the computation of an unknown position in that model is the primary emphasis of this paper, as a simple application of solving underdetermined linear systems. On the other hand, it is worthwhile for the instructor to know the context in which applications are really used. In this example, it would be incorrect to suggest to students that GPS receivers use something like the calculations above to determine positions. However, the simple geometric model does have value, for example in modeling performance characteristics of the real system.

References

1. Paul Massatt and Karl Rudnick, Geometric formulas for dilution of precision formulas, *Journal of Navigation* **37**:4 (1990) 379–392.
2. Gilbert Strang, The mathematics of GPS, *SIAM News* **30**:5 (1997). Also online at http://www.siam.org/siamnews/general/gps.htm
3. Gilbert Strang and Kai Borre, *Linear Algebra, Geodesy, and GPS*, Wellesley-Cambridge Press, 1997.
4. Richard B. Thompson, Global positioning system: the mathematics of GPS receivers. *Mathematics Magazine* **71**:4 (1998) 260–269.

The Bipolar Oven

Marc Brodie, of the College of St. Benedict (mbrodie@csbsju.edu), who keeps his eyes open, noticed a sign at a pizza place:

<div align="center">

THE BEST
PIZZA IS
A MATTER OF
DEGREES

CAUTION! OVEN $+/-500°$

</div>

He comments, "While I concede that both temperatures would warrant caution (and ignoring the fact that $-500°$ is below absolute zero), only one will do a reasonable job of cooking a pizza. (Incidentally, this sign was obviously a mass-produced-by-headquarters type sign, not a handwritten scribble on a post-it note.)"