1 Determination of the accurate location of an aircraft

In this chapter we want to find the most accurate location of an aircraft using information from beacons. This is essentially the same problem that a GPS system or a cellular phone has to solve.

Topics

- Non-linear least squares
- Statistical errors
- Function minimization
- Sensitivity analysis

The chapter is organized as follows:

- in Section 1.1 the problem is described,
- in Section 1.2 we will show how to model this problem mathematically,
- in Section 1.3 we will solve it analytically, and
- in Section 1.4 we will explore methods to analyze the solution we have found.

In Appendix A we explore different minimization methods.

1.1 Introduction

Figure 1.1 illustrates a simplified typical situation of navigation with modern aircraft. The airplane is in an unknown position and receives signals from various beacons. Every signal from the beacons is assumed to contain some error. The main purpose of this problem is to develop a method for computing the most likely position of the aircraft based on all the information available.

We distinguish two kinds of beacons: very high frequency omnirange (VOR) and distance measuring equipment (DME). The VOR beacons allow the airplane to read

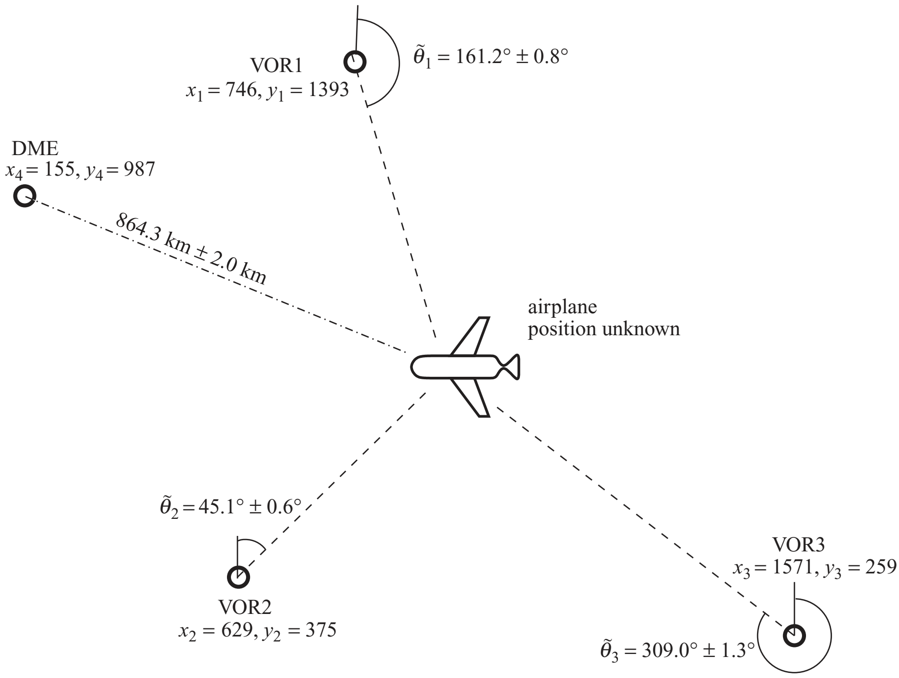

Figure 1.1 **Example of an aircraft and four beacons.**

the angle from which the signal is coming. In other words, θ_1, θ_2 and θ_3 are known to the airplane. The DME beacon, using a signal that is sent and bounced back, allows the distance from the airplane to the beacon to be measured. In this example the distance is 864.3 km $\pm$ 2.0 km.

Each of the measurements is given with an estimate of its error. The standard notation for measurements and errors is $m \pm n$. This means that the true value being measured lies between $m - n$ and $m + n$. Different disciplines have different interpretations for the statement "lies between." It may mean an absolute statement, i.e. the true value is always between the two bounds, or a statistical statement, i.e. the true value lies within the two bounds z% of the time. It is also common to assume that the error has a normal distribution, with average m and standard deviation n. For our analysis, it does not matter which definition of the error range is used, provided that all the measures use the same one.

We will simplify the problem by considering it in two dimensions only. That is, we will not consider the altitude, which could be read from other instruments and would unnecessarily complicate this example. We will denote by x and y the unknown coordinates of the aircraft.

The input data are summarized in the following table.

	x coordinate	y coordinate	value	error
VOR1	$x_1 = 746$	$y_1 = 1393$	$\tilde{\theta}_1 = 161.2^\circ$	$\tilde{\sigma}_1 = 0.8^\circ$
VOR2	$x_2 = 629$	$y_2 = 375$	$\tilde{\theta}_2 = 45.1^\circ$	$\tilde{\sigma}_2 = 0.6^\circ$
VOR3	$x_3 = 1571$	$y_3 = 259$	$\tilde{\theta}_3 = 309.0^\circ$	$\tilde{\sigma}_3 = 1.3^\circ$
DME	$x_4 = 155$	$y_4 = 987$	$\tilde{d}_4 = 864.3$ km	$\tilde{\sigma}_4 = 2.0$ km
aircraft	x	y		

It is easy to see, that unless we are in a pathological situation, any pair of two VOR/DME readouts will give enough information to compute x and y. If the measurements were exact, the problem would be overdetermined with more than two readouts. Since the measures are not exact, we want to compute x and y using all the information available and, hopefully, obtain a more accurate answer.

The standard measure of angles in aviation is clockwise from North in degrees. This is different from trigonometry, which uses counterclockwise from East in radians. Hence care has to be taken with the conversion from degrees to radians.

We do this step first and get the following results.

	x coordinate	y coordinate	value	error
VOR1	$x_1 = 746$	$y_1 = 1393$	$\tilde{\theta}_1 = 5.0405$ rad	$\sigma_1 = 0.014$ rad
VOR2	$x_2 = 629$	$y_2 = 375$	$\tilde{\theta}_2 = 0.784$ rad	$\sigma_2 = 0.0105$ rad
VOR3	$x_3 = 1571$	$y_3 = 259$	$\tilde{\theta}_3 = 2.461$ rad	$\sigma_3 = 0.023$ rad
DME	$x_4 = 155$	$y_4 = 987$	$\tilde{d}_4 = 864.3$ km	$\sigma_4 = 2.0$ km
aircraft	x	y		

Definition of best approximation Find the aircraft position (x, y) which minimizes the error in the following way.

- Regard the total error ε as a vector of all the measurement errors. This vector of errors contains the error of each measurement for a value of (x, y); in our example it has four components.
- As norm of this vector we use the usual euclidean norm $\|\varepsilon\|_2$ which is defined in our example as $\|\varepsilon\|_2 := \sqrt{\sum_{i=1}^{4} \varepsilon_i^2}$. (We could use other norms instead, for example $\|\varepsilon\|_{\max} := \max(|\varepsilon_i|)$.)
- Find the (x, y) which minimizes the norm of the total error, hence for $\|\varepsilon\|_2$ use the method of least squares (LS).

1.2 Modelling the problem as a least squares problem

Under the assumption that the errors are normally distributed, it is completely appropriate to solve the problem of locating x and y by minimizing the sum of the squares of the errors. On the other hand, if we do not know anything about the distribution of the individual errors, minimizing the sum of their squares has a simple geometrical interpretation, the euclidean norm, which is often a good idea. So, without further discussion, we will pose the problem as a least squares (LS) problem.

We can relate the unknown exact position (x, y) of the airplane with the given VOR positions (x_i, y_i) for $i = 1, \ldots, 3$ by:

$$\tan(\theta_i) = \frac{x - x_i}{y - y_i} \tag{1.1}$$

where θ_i is the angle to the unknown exact airplane position. Using the DME position (x_4, y_4), we get the equation

$$d_4 = \sqrt{(x - x_4)^2 + (y - y_4)^2} \tag{1.2}$$

for d_4, the distance to the unknown exact position of the airplane.

Each of the measures is subject to an error. We will call these errors for the different measures ε_i. Hence $\varepsilon_i = \theta_i - \tilde{\theta}_i$, where θ_i is the real value, $\tilde{\theta}_i$ is the actual measure of the value and ε_i is the measurement error. (If the measure is given as $m \pm n$, $\tilde{\theta}_i = m$.) For example, $\theta_1 = 161.2 + \varepsilon_1$, and we mean θ_1 is the exact angle and ε_1 is the error of the actual measurement.

Now the above Equations (1.1) and (1.2) can be written as:

$$\tan(\theta_i) = \tan(\tilde{\theta}_i + \varepsilon_i) = \frac{x - x_i}{y - y_i} \qquad \text{for } i = 1, \ldots, 3 \tag{1.3}$$

$$d_4 = \tilde{d}_4 + \varepsilon_4 = \sqrt{(x - x_4)^2 + (y - y_4)^2}. \tag{1.4}$$

All the variables are related by this system of four equations in the six unknowns x, y, ε_1, ε_2, ε_3, ε_4. It is normally underdetermined, so we cannot determine the exact position. Instead we are going to determine the solution which minimizes the norm of the total error $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$.

The errors ε_i, when viewed as random errors, have a known average, namely 0 (since the measurement instruments are typically unbiased), and a known variance or standard deviation. For example, ε_1 has average 0 and a standard deviation 0.014 rad. In general for a measure $m_i \pm \sigma_i$ the associated error has average 0 and standard deviation σ_i or variance σ_i^2. If the errors are assumed to be normally distributed, then they have a normal (gaussian) distribution $\mathcal{N}(0, \sigma_i^2)$.

When we minimize the norm of ε it should be done on similarly distributed variables ε_i. (We want to compare apples to apples and not apples to oranges.) To achieve this, we will divide each ε_i by its standard deviation σ_i. The normalized errors ε_i/σ_i have distribution $\mathcal{N}(0, 1)$.

So, since we are using the euclidean norm, that is, we want to minimize the length of the "normalized" error-vector $(\varepsilon_1/\sigma_1, \varepsilon_2/\sigma_2, \varepsilon_3/\sigma_3, \varepsilon_4/\sigma_4)$, we have to minimize $\sum (\varepsilon_i/\sigma_i)^2$. Typically every error ε_i appears in only one equation, and hence it is easy to solve for it, e.g.

$$\varepsilon_4 = \sqrt{(x - x_4)^2 + (y - y_4)^2} - \tilde{d}_4 = \sqrt{(x - 155)^2 + (y - 987)^2} - 864.3.$$

Inverting the equations for $\varepsilon_{1,\ldots,3}$ which contain the tangent function poses a small technical problem because of the periodicity of the tangent function. $\tan(\hat{\theta}_i + \varepsilon_i) = (x - x_i)/(y - y_i)$ for $i = 1, \ldots, 3$ is always correct, but inverting, we obtain $\varepsilon_i = \arctan((x - x_i)/(y - y_i)) - \tilde{\theta}_i + k\pi$ with $k \in \mathbb{Z}$. Inverting trigonometric equations with a computer algebra system like Maple will normally return the principal value, that is a value between $-\pi/2$ and $\pi/2$, which may be in the wrong quadrant.

PRACTICAL NOTE

This brings two problems, one of them trivial, the second one more subtle. The trivial problem is how to convert aviation angles, which after normalization will be in the range from zero to 2π, to the range $-\pi$ to π. This is done by subtracting 2π from the angle if it exceeds π.

The second problem is that arctan returns values between $-\pi/2$ and $\pi/2$. This means that opposite directions, for example the angles 135° and 315°, are indistinguishable. This may result in an equation that cannot be satisfied, or if the angles are reduced to be in the arctan range, then multiple, spurious solutions are possible. To correct this problem we analyze the signs of $x - x_i$ and $y - y_i$ to determine the correct direction, which is called quadrant analysis. This is a well known and common problem, and the function `arctan` with two arguments in Maple (`atan2` in C and Java) does the quadrant analysis and returns a value between $-\pi$ and π, resolving the problem of opposite directions (`arctan` with two arguments also resolves the problem of $y - y_i = 0$ which should return $\pi/2$ or $-\pi/2$ but could cause a division by zero).

Equation (1.3) should be rewritten as:

$$\varepsilon_i = \texttt{arctan}(x - x_i, y - y_i) - \tilde{\theta}_i.$$

Finally the sum of squares that we want to minimize in our example is

$$S(x, y) = \sum_{i=1}^{3} \left(\frac{\texttt{arctan}(x - x_i, y - y_i) - \tilde{\theta}_i}{\sigma_i} \right)^2 + \left(\frac{\sqrt{(x - x_4)^2 + (y - y_4)^2} - \tilde{d}_4}{\sigma_4} \right)^2.$$

This problem is non-linear in its unknowns x and y as x and y are simultaneously arguments of an arctan function and inside a square root function. This means that an explicit solution of the least squares problem is unlikely to exist and we will have to use numerical solutions.

See the interactive exercise "Least squares."

1.2.1 How to solve a mixed problem: equation system and minimization

BASIC

Suppose we want to minimize a function subject to constraints. This can be viewed as solving a minimization jointly with a set (or system) of equations, the constraints. Let $eq_1, eq_2, \ldots, eq_i$ be the equations (constraints), and let $f(x_1, x_2, \ldots, x_k)$ be the function to be minimized.

A simple procedure to solve this problem is by substitution. It consists of the following steps.

(i) Choose an equation eq_j and an unknown in eq_j. (This choice should be made for the equation/unknown which is easiest to solve. The easiest case is solving for an unknown which does not appear in any other equation, in our example the ε_j.)
(ii) Solve eq_j for the unknown and substitute the value for this unknown in f and in all other remaining equations.
(iii) Remove eq_j from the set of equations.

When no more constraints are left we can minimize f in terms of the unknowns that remain. The substituted variables can be computed from the minimal solution by backsubstitution, if desired.

Notice that in our case, normally each error is associated with a measure and each measure gives a constraint (equation). This error will not appear in any other measure/equation. So we have a way of solving all equations easily by substitution.

See the interactive exercise "Measurement."

1.3 Solving the non-linear least squares problem

To solve this problem we will use the computer algebra system Maple, since we need to do some symbolic as well as numerical computations. First we define the input data. We use the vectors `X` and `Y` to store the beacon coordinates and `x` and `y` for the unknown coordinates of the airplane.

```
> theta := array([161.2,45.10,309.0]);
> sigma := array([0.8,0.6,1.3,2.0]);
> X := array([746,629,1571,155]);
> Y := array([1393,375,259,987]);
> d4 := 864.3;
```

$$\theta := [161.2,\ 45.10,\ 309.0]$$
$$\sigma := [0.8,\ 0.6,\ 1.3,\ 2.0]$$
$$X := [746,\ 629,\ 1571,\ 155]$$
$$Y := [1393,\ 375,\ 259,\ 987]$$
$$d_4 := 864.3$$

The angles and the standard deviation of angles have to be converted to radians, as described earlier. These are the calculations which were performed:

```
>  for j from 1 to 3 do
>      theta[j] := evalf(2*Pi*theta[j]/360);
>      if theta[j] > evalf(Pi) then
>         theta[j] := theta[j] - evalf(2*Pi)
>      fi;
>      sigma[j] := evalf(2*Pi*sigma[j] / 360);
>   od:
>   print(theta);
>   print(sigma);
```

$$[2.813470755,\ 0.7871434929,\ -0.890117918]$$
$$[0.01396263402,\ 0.01047197551,\ 0.02268928028,\ 2.0]$$

We are now ready to construct the sum of squares.

```
>  S := sum(((arctan(x-X[i],y-Y[i])-theta[i])/sigma[i])^2, i=1..3)+
>                  +((((x-X[4])^2+(y-Y[4])^2)^(1/2)-d4)/sigma[4])^2;
```

$$\begin{aligned} S := {} & 5129.384919\,(\arctan(x-746,\ y-1393)-2.813470755)^2 \\ & +9118.906531\,(\arctan(x-629,\ y-375)-0.7871434929)^2 \\ & +1942.488964\,(\arctan(x-1571,\ y-259)+0.890117918)^2 \\ & +0.2500000000\,(\sqrt{(x-155)^2+(y-987)^2}-864.3)^2 \end{aligned}$$

Next we solve numerically for the derivatives equated to zero. In Maple, `fsolve` is a basic system function which solves an equation or system of equations numerically.[1]

```
>  sol := fsolve({diff(S,x)=0, diff(S,y)=0},{x=750,y=950});
```

$$\mathrm{sol} := \{x = 978.3070298,\ y = 723.9837773\}$$

[1] ($x = 750$, $y = 950$) is an initial guess to the solver `fsolve`. It will try to find a solution starting from this point. This is useful for two reasons, first it improves the efficiency of the solver and secondly it increases the chances that we converge to a minimum (rather than other places where the derivative is zero like maxima or saddle points).

A solution has been found and it is definitely in the region that we expect it to be. The first measure of success or failure of the approximation is to examine the residues of the least squares approximation. Under the assumption that the errors are normally distributed, this will be the sum of the squares of the four $\mathcal{N}(0, 1)$ variables. Note that

$$E[x_1^2 + \cdots + x_4^2] = 4 \quad \text{if} \quad x_i \sim \mathcal{N}(0, 1)$$

but for the least squares sum $S = \sum_i \varepsilon_i^2$ we choose *optimal* x, y when we minimize S. Therefore we expect

$$E[S_{\min}] = 2$$

since this system has only two degrees of freedom.[2] The norm squared of the error is:

```
>  S0 := evalf(subs(sol, S));
```

$$S_0 := 0.6684712637$$

This value is smaller than 2, and hence it indicates that either we are lucky, or the estimates for the errors were too pessimistic. In either case, this is good news for the quality of the approximation. This together with the eigenvalue analysis from the next section guarantees that we have found the right solution.

1.4 Error analysis/confidence analysis

A plain numerical answer, like the result from `fsolve` above, is not enough. We would like to know more about our result, in particular its confidence. Confidence analysis establishes the relation between a region of values T and the probability that the correct answer lies in this region. In our case we are looking for a range of values which are a "reasonable" answer for our problem, see Figure 1.2:

$$P[\text{given } T\text{, the real answer is outside } T] = \begin{cases} 10^{-6} & \text{very precise} \\ 0.001 & \text{precise} \\ 0.01 & \\ 0.05 & \text{reasonable} \\ p & \text{in general.} \end{cases}$$

[2] The discussion on the degrees of freedom goes beyond the scope of this book. A simple rule of thumb is that if our minimization, after substitution for all the constraints, has k variables left and we have n errors, then the degrees of freedom are $n - k$. In this case: $4 - 2 = 2$.

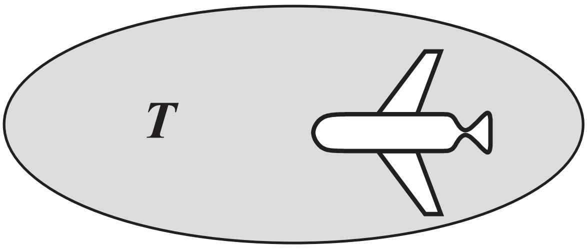

Figure 1.2 **The exact position of the airplane is not known, we can only compute regions T where the airplane is likely to be.**

If the errors are assumed to be normally distributed, the sum of the squares of the errors has a known distribution, called the χ^2-distribution (read: chi-square distribution).

We will work under the assumption that the errors of the measures have a normal distribution. The distribution of $S(x, y)$ is a χ^2-distribution with four degrees of freedom, since it is a sum of squares of four variables which are $\mathcal{N}(0, 1)$ distributed.[3]

For a given confidence level we can bound the value of χ^2 and hence bound the solution (x, y), for example by $S(x, y) \leq v$ with $\Pr(\{\chi_4^2 \leq v\}) =$ confidence. This inequality defines an x, y area T, which is an area where the airplane will be located with the given confidence. (Note: the bigger the confidence, i.e. the lesser the probability of an error, the bigger is the area.)

Knowing its distribution allows us to define a confidence interval for the airplane position. Suppose that we are interested in a 95% confidence interval, then $S(x, y) < v \approx 9.4877$ (Figure 1.3), where this value is obtained from the inverse of the cumulative (`icdf`) of the χ^2-distribution. In Maple this is computed by:

```
>  stats[statevalf,icdf,chisquare[4]](0.95);
```

$$9.487729037$$

The inequality $S(x, y) < 9.4877$ defines an area which contains the true values of x and y with probability 95%. We can draw three areas for three different confidence intervals, e.g. 50%, 95%, 99.9%, all of which are reference values in statistical computations. Notice that the larger the confidence, the larger the ellipse (see Figure 1.4). And see the interactive exercise “Quality control.”

[3] We do not regard S as a function of (x, y) in this case, since (x, y) is kept fixed (it should be the true position of the airplane). Instead we regard S as a random variable, since it still depends on the *four* measurements, which are random variables. Since these measurements contain independent, normally distributed errors $\varepsilon_1, \ldots, \varepsilon_4$, S has a χ^2-distribution with four degrees of freedom.

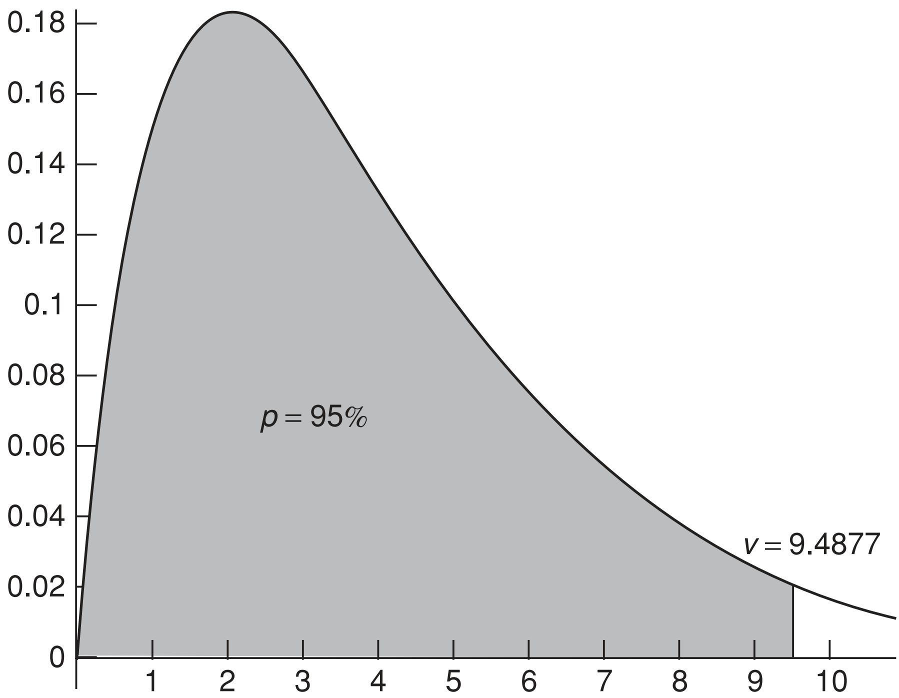

Figure 1.3 The probability density function of a χ^2-distribution with four degrees of freedom.

1.4.1 An approximation for a complicated S

For complicated $S(x, y)$, finding the area may be computationally difficult. So we will show how to use an approximation. We will expand the sum of the squares of the errors as a Taylor series around the minimum and neglect terms of third and higher order. So let $S(x, y) = S(\boldsymbol{p})$ be the sum of squares, which we will define as a function of the position vector $\boldsymbol{p} = (x, y)^T$. Let $\boldsymbol{p}_0$ be the solution of the least squares problem. Then the three-term Taylor series around $\boldsymbol{p}_0$ is

$$S(\boldsymbol{p}) = S(\boldsymbol{p}_0) + S'(\boldsymbol{p}_0)(\boldsymbol{p} - \boldsymbol{p}_0) + \frac{1}{2}(\boldsymbol{p} - \boldsymbol{p}_0)^T S''(\boldsymbol{p}_0)(\boldsymbol{p} - \boldsymbol{p}_0) + \mathcal{O}(\|\boldsymbol{p} - \boldsymbol{p}_0\|^3).$$

The gradient of S, $S'(\boldsymbol{p}) = (S_x, S_y)^T$, is always zero at the minimum, and a numerical check shows that the gradient in our example is indeed within rounding error of (0, 0):

```
> S1 := evalf(subs(sol,linalg[grad](S,[x,y])));
```

$$S_1 := [0.36 \times 10^{-7},\ -0.6 \times 10^{-8}]$$