Lecture notes for Math 647

T. I. Vogel

1/21/09

1 What is a mathematical model?

First off, what is a model? Speaking very roughly, it's something that mimics relevant features of a situation begin studied. For example, if we're thinking of driving around Texas, a road map would be a reasonable model of Texas. On the other hand, if we're interested in aquifers in Texas, a road map would not be a particularly useful model.

A mathematical model is similar. it's a simplified, abstract mathematical construction relating to a real–world situation and used for a specific purpose. In making a model, the world is divided into three parts:

1. Things which may be neglected.
2. Things that affect the model but whose behavior the model is not designed to study (input variables).
3. Things that the model is designed to study (output variables).

In making mathematical models, you must always make simplifying assumptions. Too many, and the model is unrealistic (e.g., a road map of Texas which has only the interstates), but too few, and the model is unworkable (e.g., a map of Texas which has every driveway in the state). The assumptions that you make are a matter of taste and experience.

2 Building a model

Giving rules for building models is something like giving rules for writing poetry. The best I can do is give a framework to help you get started. Here's an outline, with a warning not to take it too seriously.

1. **Formulate the problem.** What is it that you want to know? The nature of the model depends on what you want to do with it (as the example of the two maps of Texas).

2. **Outline the model.** At this point you must (at least mentally) separate the universe into the three parts mentioned in the previous section. The interrelations among the variables must also be specified. This is where most of the work is done.

3. **Can it be used?** Stand back and look at what you've done. Can you obtain the needed data and then use it in the model to make the predictions that you want? If not, you must reformulate the model (step 2) or even the problem (step 1). Note that "useful" does not mean reasonable or accurate; those come in step 4. It means *if* the model fits the situation, will we be able to use it?

4. **Test the model.** Use the model to make predictions that can be checked against data or common sense. It is not advisable to rely entirely on common sense, since that may well be wrong. *Start out with easy predictions* - don't waste time on involved calculations with a model which may be no good. If the predictions are bad and there are no mathematical errors, return to step 2 or 1. If these predictions are acceptable, they should give you some feeling for the accuracy and range of applicability of the model. If they are less accurate than you anticipated, it is a good idea to try to understand why, since this may uncover implicit or false assumptions.

Beware: although "curve-fitting" may be part of the modeling process, it is not all of it. For example, if you propose to model the price of a stock, it's not enough to observe the price on different days, do a least squares fit, and claim that the resulting line models the price of the stock. However, this *would* be valid if you also gave a model supporting your assumption that the price is linear. (Of course, I doubt that any reasonable model of stock price would produce such a result.)

3 An example

Suppose that we are asked to come up with a long term model of the population of the world. A minute on Google can give us a fair amount of data.

For example, we can get estimates of the population every year from 1950 to the present on the US Census Bureau's site. It would be somewhat mindless to attempt to fit a line to these points, unless we already have a model which predicts linear growth. Instead, the simplest assumption about population growth is that if population at time t is $N(t)$, then the rate of change is proportional to $N\left(t\right)$. The reasoning is that without further assumption, the number of families having children in (say) a year is likely to be some fixed number times the total number of families. But the number of children added per year is $\frac{dN}{dt}$ (roughly), so we're assuming that

$$\frac{1}{N}\frac{dN}{dt}=k,$$

for some constant k, called the *net growth rate*. This is easy to solve, and it gives $N\left(t\right)=N_0e^{kt}$. So, the simplest model predicts exponential growth. There are two parameters N_0 and k which we would have to estimate to make predictions. One possibility is to plot the logarithm of the population against time, since if $N\left(t\right)=N_0e^{kt}$, then $\ln\left(N\right)=\ln\left(N_0\right)+kt$, and then try to find the line which best fits the points. This model can certainly be used, but it turns out that it doesn't give very accurate predictions. In particular, as time goes to infinity, so does the population, which is not reasonable. Notice that N_0 and k are input variables: our model is not designed to study them. The population is the output.

One thing which we neglected in the exponential model is the size of the Earth, or how many people the Earth can support. Perhaps a more reasonable model is obtained by assuming that the net growth rate isn't constant, rather it depends on the total population:

$$\frac{1}{N}\frac{dN}{dt}=k\left(N\right).$$

What are reasonable assumptions on $k\left(N\right)$? As crowding increases, we expect it to decrease. If N_∞ is the maximum population which the Earth can sustain (called the *carrying capacity*), then we expect that when the population equals N_∞, the growth rate is zero. We can already use this model to make some predictions even without assuming a specific form of $k\left(N\right)$: from the equations we can determine that if $N_0>0$ then $N\left(t\right)$ will increase an approach N_∞ in the limit. Unfortunately, it's hard to give concrete numbers as a prediction from this.

The usual way to make progress from here is to assume that $k(N)$ is a linear function of N. The simplest way to do this is to assume $k(N) = r\left(1 - \frac{N}{N_\infty}\right)$. This decreases to zero at N_∞, so this $k(N)$ has the form we want. We've come up with the logistic equation:

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{N_\infty}\right).$$

This is a more complicated ODE. It does have an exact solution:

$$N(t) = \frac{N_\infty}{1 - C_1 N_\infty e^{-rt}}.$$

We now have three parameters to try to approximate from our data. This is already pretty nasty to do, and is the sort of thing which we may talk about later.

Between the two models we have a typical trade-off. The first is easy to use, but not very accurate. The second is a little harder to use (i.e., determine reasonable constants), but is probably more reasonable, at least when population is not far from the carrying capacity. On the other hand, for populations which are quite small compared to the carrying capacity, it makes sense to use the exponential model.

Don't assume that logistic growth is the end of population modeling. To come up with more accurate models you could look at individual countries (the growth rate for Denmark is probably smaller than than of Mexico). In that case, you might want to include terms for migration between countries. Here you'd probably get a system of maybe a hundred or so ODE's, so it would be impossible to go forward without using computers. That would be unwieldy, so you might divide into two groups: industrialized countries and developing countries, and get two ODE's, something like

$$\begin{aligned}
\frac{dN_i}{dt} &= r_i N_i\left(1 - \frac{N_i}{N_{i,\infty}}\right) - k_1 N_i + k_2 N_d \\
\frac{dN_d}{dt} &= r_d N_d\left(1 - \frac{N_d}{N_{d,\infty}}\right) + k_1 N_i - k_2 N_d,
\end{aligned}$$

where I've tossed in immigration and emigration terms into the logistic equations. We'd now have 4 constants to estimate, along with initial conditions, so we'd have a lot of work to do to try to use this model. It would be impossible to solve these exactly, but once you have constants, you can solve this numerically. More sophisticated population models will include the age profiles of the population, and so on. There are also models which have a degree of randomness (called *stochastic models*), since the assumption that population is completely deterministic is a shaky one. All of this is just to show that there's a trade-off between sophistication of a model and how easily used it is.

Adapted from "An Introduction to Mathematical Modeling", by E. A. Bender.