Numerical Astronomy 1 – Part 5
Conclusion

Norman Gray

11 December 1998

1 Modelling

Modelling has the same aim as inverse problem methods – you have a set of data points and you want to get at the physics underneath. In both cases, you know something about the forward problem – in the IP approach you express this in the functional form of the kernel, and in the modelling approach in the functional dependence of the predicted data on a set of model parameters. In the IP approach, what you try to recover is some otherwise unconstrained source function which is distorted by the kernel into the data function; in the modelling approach, you aim to find numerical values for the set of parameters.

In the most typical modelling problem, you will have some physical justification for the dependence of the data on the parameters, which will generally have some direct physical interpretation, as orbital elements, or atomic rate constants, or whatever is appropriate. Sometimes, however, the parameters will be coefficients in an abstract function expansion, for example in terms of gaussians (there is an obvious connection here with the orthogonal function expansion of Sect. 3-1.3).

It is this physical justification for the model which gives modelling both its strength and its weakness. In effect, you are saying 'I really believe the physical system is like *this*, and I just want to know the numbers'. What you are doing here is, in IP terms, packing a huge amount of prior information into the problem. If the physics really *is* like that, and if nothing untoward is happening to your data on its way to your instruments, then you've won – you can read off the numbers and start worrying about what they tell you about astronomy. If the modelling approach is appropriate, but the model is wrong, then even your best fitting set of parameters will show a poor fit to the data (in a suitably quantified sense), and you will be prompted to discard the model. What can go wrong?

The qualifications above are important. A model will be revealed to be wrong *if* the modelling route is appropriate. We have seen, in parts 3 and 4, that prior information, explicit or implicit, can regulate the recovery of an underlying function, avoiding excessive dependence on data noise by pulling it towards some prior assumption such as smoothness or positivity. However, as discussed in Sect. 4-2, too much regulation (that is, choosing too large a regulation parameter λ in Eqn. (4.3)) means that you recover the same source function for just about any data. The same is true in the context of modelling: if the modelling route is inappropriate, you will obtain 'good' fits for just about any data, even if the real physics looks nothing like your model of it. In other words, modelling is bad in such a situation, because it always 'works'. Even if you enter this problem with your eyes open, viewing the assumed model as substantial prior information, the approach can be criticised in the same way that classical IP approaches can, since the prior information enters your technique in an uncontrolled, implicit, fashion.

When, then, is modelling appropriate? If there are so many similarities between modelling and inverse problems, is there a real difference?

Yes. The key is conditioning. In Eqn. (3.23) we see that for an inverse problem $\mathbf{g} = \mathsf{K}\mathbf{u}$

$$\frac{||\delta\mathbf{u}||}{||\mathbf{u}||} \leq C_K \frac{||\delta\mathbf{g}||}{||\mathbf{g}||}, \tag{5.1}$$

and that the characteristic of an inverse problem is that the condition number $C_K = ||\mathsf{K}||||\mathsf{K}^{-1}||$ is large. The characteristic of a modelling problem is that it is *well*-conditioned: the condition number is (speaking roughly) of order 1 and so the errors in the recovered parameters are of order the errors in the data.

If a problem is ill-conditioned, you need to attack it with inverse problem methods, because only these methods are set up to deal with that ill-conditioning. It is unfortunate that inverse problem approaches are generally harder to set up and use than modelling approaches.

Any modelling method should give you parameter estimates; however, it should also give you estimates of both the errors in these estimates, and the quality of the fit, since these can tell you whether the parameters you have obtained are meaningful – large errors or a bad fit tell you that something is wrong. If you do not have these supporting measures, and you do not have independent information about the problem, then be very, very, cautious about drawing any scientific conclusions from the numbers which have appeared in front of you.

There is a great deal more to say about modelling, because it is harder and more subtle than it first appears, and because it is full of elephant traps which are the cause of a good deal of bad science. The technicalities would, however, take us too far afield.

2 Review: Choosing an IP algorithm

In many cases, simply recognising that a problem is ill-conditioned, and so must be treated as an inverse problem, is a significant positive step. This immediately opens up the problem of which IP method to use to approach the problem. In this section, I will briefly review the different methods I have described in the preceding weeks. As well, Craig and Brown give a recommended strategy in their chapter 8.

Quadrature Very simple and straightforward, and so of primarily pedagogical use, I think. Could be useful for the initial exploration of a problem.

Product integration A bit more sophisticated than simple quadrature. Good for relatively undemanding problems.

Polynomial expansion Though it's arguably a type of modelling, this is a flexible IP approach which will work well if you keep your wits about you. It could be very useful in a primarily analytical investigation of an inverse problem. If you choose a bad set of functions, however, things could go badly wrong.

Singular value decomposition Very powerful and stable, but computationally expensive.

Regularisation An extremely flexible technique which will work well if you have a suitable regulating/smoothing functional. It requires a good deal of investigation of the problem before you can make sensible choices about the smoothing parameter.

Backus-Gilbert Computationally expensive, and not generally suitable for actual data processing. It gives good insights into the problem, the soft of data resolution required, and what might constitute an optimal experiment.

Maximum entropy Very well suited to certain types of problems, where discrete features are to be recoved from a noisy background (often used in image processing, for example). The statistics of the recovered information (that is, bias, error estimates) are often not good.

You will typically need to support your investigations by some simulation, such as playing with the source function and using the forward problem to find the effect on the simulated data. Similarly, techniques such as faking data and using your technique to recover the source function will be necessary to help you choose parameters such as the smoothing parameter.

In all cases, the use of a suitable IP method should not, and indeed cannot, be taken as a substitute for fully understanding the problem and its scientific context, and thinking deeply and carefully about it.

Have fun.