Do (1) and choose one of (2) *or* (3).

1. Show that the expressions for the mean and variance of the posterior distribution for the generalized-least-squares problem

$$d_i = A_{im} s_m + n_i \qquad \text{with} \quad \langle n_i n_j \rangle = N_{ij}$$

(i.e., generalized linear model with arbitrary Gaussian noise) reduces to the simple Gaussian "averaging" operation that we discussed in the first lecture for the appropriate choice of A_{im} and noise correlation matrix N_{ij}.

2. Find a problem in your field of study and figure out how you could apply Bayesian ideas to it. If this is something you can do in a few hours, feel free to do that instead of the rest of this problem sheet. Otherwise, just think about it.

3. On my website (`http://astro.imperial.ac.uk/~jaffe`) you'll find a link to a PDF of these lectures... and a file with some data.

   (a) First, consider only the data with $x > 0$. Model the data as a straight line, $y = ax + b$, and determine a and b, their errors, and their covariance. (It's a worthwhile exercise to code this up yourself if you can, but the basic formulae are present in many software packages such as IDL.)

   (b) Plot the data (with errors), and your fit to it.

   (c) Plot the (two-dimensional) posterior distribution of the parameters.

   (d) Now, do the same thing for the whole data for all x.

   (e) Comment on the difference between the results. Could you think of a way to decorrelate your errors from the first part?

   (f) Is there any information you need to give me so I know exactly what you did?

   (g) In fact, I may have added in a *small* quadratic component to that data. What would you do test this possibility? (I am not asking you to go through the calculation unless you really want to...)