Assignment 2 - Inverse Problems 2020

Many types of spectral data consist of several overlapping peaks, each of which has either Gaussian or Lorentzian shape. (Both the Gaussian and the Lorentzian have a single maximum, but the Lorentzian is much longer tailed.) The problem is to determine the location, area, and width of each peak through non-linear least squares inversion. Suppose that we have N data (see Figure 1) with, say,

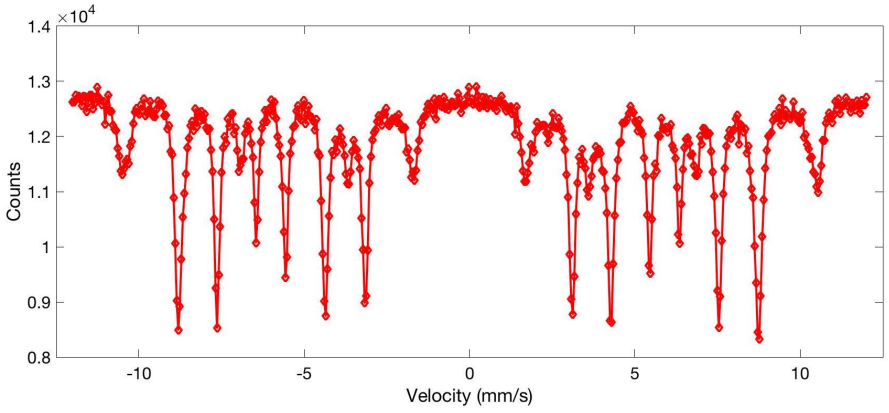

Figure 1: The Mössbauer spectroscopic data to be analyzed in this assignment. For more information about the Mössbauer method, see, e.g., `https://en.wikipedia.org/wiki/Mossbauer_spectroscopy`. Note that the spectral "peaks" are pointing downwards! Data courtesy of NASA and the University of Mainz.

q peaks, each characterized by 3 unknown parameters: its center frequency f_i, area A_i, and width c_i. There are then $M = 3q$ unknown model parameters $m = [A_1, f_1, c_1, ..., A_q, f_q, c_q]^T$. The forward relation is nonlinear and of the form $\mathbf{d} = \mathbf{g}(\mathbf{m})$. The data uncertainties are approximately Gaussian with standard deviation $\sigma = 0.03$ (counts).

In the Gaussian case we have:

$$d_i = \sum_{j=1}^{q} \frac{A_j}{(2\pi)^{1/2} c_j} \exp\left[-\frac{(z_i - f_j)^2}{2c_j^2}\right]$$

In the Lorentzian case we have:

$$d_i = \sum_{j=1}^{q} \frac{A_j c_j^2}{(z_i - f_j)^2 + c_j^2}$$

Note that in the case of the Gaussian, the quantity c_i^2 has the interpretation of variance, but in the case of the Lorentzian (which has infinite variance), it does not. Since the data have Gaussian error, it is appropriate to use Newton's method to solve this problem. Furthermore, the problem will typically be overdetermined, at least if $N > M$ and if the peaks do not overlap completely. The equation is linearized around an initial guess using Taylor's theorem. This linearization involves computing a matrix $\mathbf{G}^{(p)}$ of derivatives (in the p'th iteration) with rows

$$\left[\frac{\partial g_i}{\partial A_1} \; \frac{\partial g_i}{\partial f_1} \; \frac{\partial g_i}{\partial c_1} \cdots \frac{\partial g_i}{\partial A_p} \; \frac{\partial g_i}{\partial f_p} \; \frac{\partial g_i}{\partial c_p}\right]$$

In this problem the Gaussian and Lorentzian models are simple enough for the derivatives to be computed analytically.

1. Derive analytical formulas for $\frac{\partial g_i}{\partial A_p}$, $\frac{\partial g_i}{\partial f_p}$, and $\frac{\partial g_i}{\partial c_p}$ in the Gaussian case.

2. Derive analytical formulas for $\frac{\partial g_i}{\partial A_p}$, $\frac{\partial g_i}{\partial f_p}$, and $\frac{\partial g_i}{\partial c_p}$ in the Lorentzian case.

3. Evaluate these derivatives at the trial solution $\mathbf{m}^{(0)}$ with the starting value chosen from visual inspection of the data. For both the Gaussian and Lorentzian models, find improved estimates of the model parameters by using the recursions $\mathbf{G}^{(p)}\Delta\mathbf{m}^{(p)} = \mathbf{d} - \mathbf{g}(\mathbf{m}^{(p)})$ and $\mathbf{m}^{(p+1)} = \mathbf{m}^{(p)} + \alpha\Delta\mathbf{m}^{(p)}$, where $0 < \alpha < 1$ is an appropriate constant, p is the iteration number, and where the matrix equation is solved with simple least squares. The iterations are terminated when the correction factor $\Delta\mathbf{m}^{(p)}$ becomes negligibly small (for instance, when the absolute value of each component becomes less than some given tolerance).

4. Which one of the model choices (the Gaussian or the Lorentzian) are better able to fit the shape of the curve? Is the misfit (large or small) related to uncertainty (noise) on the data, or to other factors?

One of the drawbacks to these iterative methods is that if the initial guess is too far off, or α is too large, the solution may oscillate wildly or diverge from one iteration to the next. You can avoid this behavior by examining the perturbation on each iteration and, if it is too large, decreasing α. This procedure will prevent the method from wildly "overshooting" the true minimum but will also slow the convergence.