PHYS 512 Problem Set 4

Teophile Lemay, 281081252

1

1.1 a)

Newton's method is a way to perform a non-linear least squares fit to some data by using the gradient of the model in parameter space to update the fit. Fitting the Lorentzian

$$d(t) = \frac{a}{1 + \frac{(t-t_0)^2}{w^2}}$$

requires the following derivatives:

$$\frac{\partial d}{\partial a} = \frac{1}{1 + \frac{(t-t_0)^2}{w^2}}$$

$$\frac{\partial d}{\partial t_0} = \frac{\frac{2a(t-t_0)}{w^2}}{\left(1 + \frac{(t-t_0)^2}{w^2}\right)^2}$$

$$\frac{\partial d}{\partial w} = \frac{\frac{2a(t-t_0)^2}{w^3}}{\left(1 + \frac{(t-t_0)^2}{w^2}\right)^2}$$

which make up the gradient of d in parameter space:

$$\boldsymbol{\nabla} d = \begin{pmatrix} \frac{\partial d}{\partial a} \\ \\ \frac{\partial d}{\partial t_0} \\ \\ \frac{\partial d}{\partial w} \end{pmatrix}$$

For any parameter guess $\mathbf{m}$ evaluating the gradient over an interval and comparing to the data, Newton's method gives a parameter update according to

$$\boldsymbol{\nabla} d(\mathbf{m})^T N^{-1} \boldsymbol{\nabla} d(\mathbf{m}) \delta\mathbf{m} = \boldsymbol{\nabla} d(\mathbf{m})^T N^{-1} (\text{data} - d(\mathbf{m}))$$

$$\delta\mathbf{m} = (\boldsymbol{\nabla} d(\mathbf{m})^T N^{-1} \boldsymbol{\nabla} d(\mathbf{m}))^{-1} \boldsymbol{\nabla} d(\mathbf{m})^T N^{-1} (\text{data} - d(\mathbf{m}))$$

$$\delta\mathbf{m} = (\boldsymbol{\nabla} d(\mathbf{m})^T \boldsymbol{\nabla} d(\mathbf{m}))^{-1} \boldsymbol{\nabla} d(\mathbf{m})^T (\text{data} - d(\mathbf{m}))$$

Where N is a noise matrix which can be omitted by setting it to identity. Iterating Newton's method until the parameter update $\delta\mathbf{m}$ becomes smaller than some threshold gives the best fit parameters.

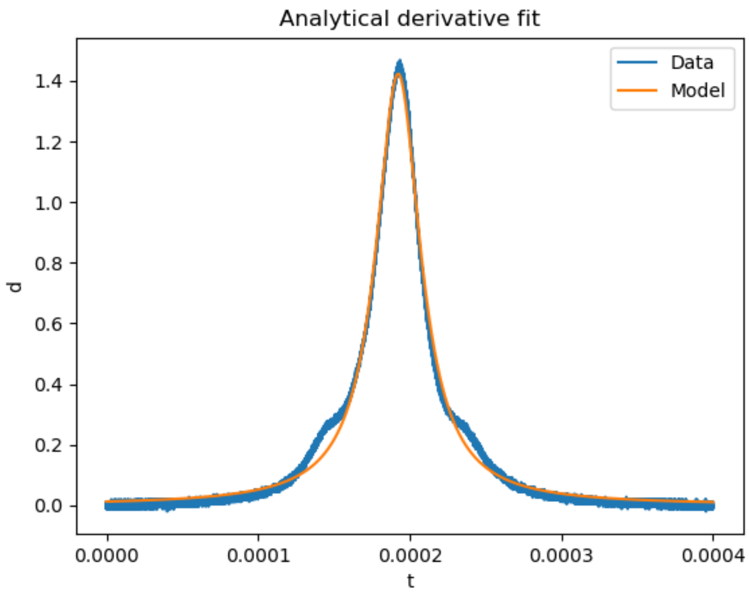
Figure 1 shows the result of a Newton's method fit of the data in "sidebands.npz" to the Lorentzian using the analytical gradient derived above. The iterations were stopped once the magnitude of the parameter update δm reached the threshold of 10^{-10} which took 13 steps. The best fit parameters for the Lorentzian are

$$a = 1.423$$

$$t_0 = 1.924 \cdot 10^{-4}$$

$$w = 1.792 \cdot 10^{-5} \ .$$

Figure 1: Analytical Newton's method fit



1.2 b)

For non-linear least squares fitting, we know

$$\nabla A^T N^{-1} \nabla A \delta m = \nabla A^T N^{-1}(d - A(m))$$

where A is the model, N is a diagonal matrix such that $N_{i,i} = \sigma_i^2$ (assuming random Gaussian noise in data), d is the measurement data, and m are model parameters. Setting $m \rightarrow m_t$ the "true" parameters such that $A(m_t) = d_t$ where d_t is the noiseless data, the equation above can be rearranged to

$$\delta m = (\nabla A^T N^{-1} \nabla A)^{-1} \nabla A^T N^{-1}(d - d_t) \ .$$

Obviously, $d - d_t = n$ is the noise in the data so

$$\delta m = (\nabla A^T N^{-1} \nabla A)^{-1} \nabla A^T N^{-1} n$$

Following the same steps as for linear least squares fitting, the covariance of the parameters $\left\langle (\delta m)^2 \right\rangle$ is

$$\left\langle (\delta m)^2 \right\rangle = (\nabla A^T N^{-1} \nabla A)^{-1}$$

Figure 2 shows a histogram of the difference between the data and the model. Assuming the best fit parameters are close to m_t this also describes the noise in the data. Since the noise is not centered at 0, I chose to estimate the noise at each data point as the root mean square of the difference between the model and the data. Thus, the resulting N matrix is a diagonal matrix with all entries equal to $rms(\text{model} - \text{data})$.
Evaluating the error in each parameter as the square roots of the diagonal elements of the covariance matrix gives error

$$\sigma a = 2.678 \cdot 10^{-3}$$

$$\sigma t_0 = 3.373 \cdot 10^{-8}$$

$$\sigma w = 4.777 \cdot 10^{-8} \ .$$

Figure 2: Noise distribution for analytical Lorentzian fit