1. Finite strain EOS fitting: setting up the equations.

To fit the given data, we use the equation from the notes:

$$P = -(1-2\epsilon)^{5/2}\left[C_1\epsilon + C_2\epsilon^2 + C_3\epsilon^3\right] \tag{1}$$

(Note that the definition of the C's here is slightly different from the notes.) A "second order" finite strain fit includes only C_1 whereas a "fourth order" finite strain fit includes all three C's. The definition of strain here is the Eulerian one:

$$\epsilon = \tfrac{1}{2}\left[1 - \left(\frac{V_0}{V}\right)^{2/3}\right] \tag{2}$$

so that given a value for V_0, you can compute a value of ϵ for each $P, V.$ data pair. Depending on the order of fit, you would now construct a matrix $\mathbf{A}$ with 1 to 3 columns. A fourth order fit would require three columns. Inspection of equation (1) shows that the matrix elements are

$$A_{ij} = -(1-2\epsilon_i)^{5/2}\epsilon_i^j \tag{3}$$

where j is the column index and goes from 1 to (possibly) 3 and i is the row index and goes from 1 to N where N is the number of data. Our problem now looks like

$$\mathbf{p} = \mathbf{A}\mathbf{c} \tag{4}$$

where $\mathbf{c}^T = C_1, C_2, C_3$. $\mathbf{p}$ is the column vector of pressure data corresponding to the values of ϵ in the $\mathbf{A}$ matrix. Before inverting equation (4) to get $\mathbf{c}$, it makes sense to divide each row of the $\mathbf{A}$ matrix and the corresponding P datum by the error in that P value and we shall assume this has been done and use the same notation as in (4). If our data are independent, this means that the covariance matrix of $\mathbf{p}$ is just the identity matrix, $\mathbf{I}$. Note that this step downweights data with large errors in the inversion.

2. Finding a solution

If our sytem of equations (4) were well-conditioned, we might just find the least-squares solution:

$$\hat{\mathbf{c}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{p} \tag{5}$$

which minimizes $(\mathbf{A}\cdot\hat{\mathbf{c}} - \mathbf{p})^2$. In reality, $\mathbf{A}$ is usually not well-conditioned and $\mathbf{A}^T\mathbf{A}$ is even worse (the condition number is effectively squared) so the solution (5) is rarely chosen. One way around squaring the condition number is to use a singular value decomposition (SVD) on equation 4. The matrix $\mathbf{A}$ is decomposed into singular values and matrices of left and right eigenvectors:

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T \tag{6}$$

where $\mathbf{U}$ has dimension $N \times N$ (where N is the number of data) and $\mathbf{V}$ has dimension $M \times M$ (where M is the number of model parameters: one to three in our case) and $\mathbf{\Lambda}$ is a $M \times N$ with non-zero diagonal elements. Note that $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^T\mathbf{V} = \mathbf{I}$. The least-squares solution in terms of the SVD is

$$\hat{\mathbf{c}} = \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{U}^T\mathbf{p} = \mathbf{A}^+\mathbf{p} \tag{7}$$

where $\mathbf{A}^+$ can be thought of as the (generalized) inverse of $\mathbf{A}$. If $\mathbf{A}$ is not well-conditioned, it will have some small singular values which will generally lead to some poorly determined contributions to $\hat{\mathbf{c}}$. To see why this is so, consider the covariance matrix of the model. To get the model we are taking a linear combination of data: $\mathbf{A}^+\mathbf{p}$. Now $\mathbf{p}$ has covariance matrix $\mathbf{I}$ so $\hat{\mathbf{c}}$ has covariance matrix $\mathbf{E}$ given by:

$$\mathbf{E} = \mathbf{A}^+\mathbf{I}(\mathbf{A}^+)^T = \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{U}^T\mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{V}^T = \mathbf{V}\mathbf{\Lambda}^{-2}\mathbf{V}^T \tag{8}$$

The square roots of the diagonal elements of this matrix are the errors on our model parameters ($\sigma_i = \sqrt{E_{ii}}$). Note that that the "correlation matrix" is just $E_{ij}/(\sigma_i\sigma_j)$. From equation (8), we see that small singular values are going to make the errors large. One way to avoid this is to exclude small singular values from the sums implicit in equations 7 and 8 but this will mean that $\mathbf{A}^+\mathbf{A}$ will no longer be $\mathbf{I}$. in fact, substituting 4 into 7 gives

$$\hat{\mathbf{c}} = \mathbf{A}^+\mathbf{A}\mathbf{c} = \mathbf{R}\mathbf{c} \tag{9}$$

and the matrix $\mathbf{R} = \mathbf{A}^+\mathbf{A}$ is sometimes called the "resolution matrix". In a perfectly resolved system, $\mathbf{R} = \mathbf{I}$ but, in general, each model element estimated will be a linear combination of all the model elements. For the truncated SVD approximation to the generalized inverse, $\mathbf{R} = \mathbf{V}\mathbf{V}^T$.

The process of throwing away small singular values is an example of "regularization" of the inverse problem. It is not a commonly used method (because the model we end up with doesn't satisfy any particularly sensible optimization criterion). Usually we seek a model which has some property optimized and still adequately satisfies the data. In our case, we might look for solutions where $\mathbf{c}^T\mathbf{c}$ is minimized and still fits the data adequately (though this is partially what truncating the SVD actually does). An alternative procedure is to supply some additional "a priori" information about the solution $\mathbf{c}$. A full implementation of this requires a discussion of Bayesian inversion which I will forgo here. In our particular case, we often have additional data we can use to constrain the inversion. For example, we may well have an estimate of K_0 from an ultrasonic experiment or some other kind of experiment. The easiest way to incorporate this is to add it as a row to equation (4). Since $C_1 = 3K_0$ we would add the value of K_0 divided by its error σ_K to the $\mathbf{p}$ vector with corresponding matrix row $(1/(3\sigma_K), 0, 0)$. We can manipulate how well this equation is fit by changing σ_K.

3. Error propagation for finite strain fitting

We fit for $\mathbf{c}^T = C_1, C_2, C_3$ where

$$\begin{aligned} C_1 &= 3K_0 \\ C_2 &= \frac{9}{2}K_0(4 - K_0') \\ C_3 &= \frac{9}{2}K_0(K_0K_0" + {K_0'}^2 - 7K_0' + 143/9) \end{aligned} \tag{10}$$

from which it is easy to recover K_0, K_0' and $K_0K_0"$. We have propagated errors to get the (symmetric) covariance matrix on the C's ($\mathbf{E}$ in equation 8) but we are actually interested in the errors in the parameters K_0, K_0' and $K_0K_0"$. We then have to propagate again to get the covariance matrix on $\mathbf{k}^T = K_0, K_0', K_0K_0"$. The relationship between $\mathbf{c}$ and $\mathbf{k}$ in equation (10) is non-linear so we have to linearize. To propagate, we need the $\mathbf{B}$ matrix in the relationship $\delta\mathbf{k} = \mathbf{B}\delta\mathbf{c}$. From equation (10), we have

$$\begin{aligned} &\frac{\partial C_1}{\partial K_0} = 3, \quad \frac{\partial C_2}{\partial K_0} = \frac{C_2}{K_0}, \quad \frac{\partial C_2}{\partial K_0'} = -\frac{9}{2}K_0 \\ &\frac{\partial C_3}{\partial K_0} = \frac{C_3}{K_0}, \quad \frac{\partial C_3}{\partial K_0'} = \frac{9}{2}K_0(2K_0' - 7), \quad \frac{\partial C_3}{\partial K_0K_0"} = \frac{9}{2}K_0 \end{aligned} \tag{11}$$

The first of these gives $\delta K_0 = \delta C_1/3 = B_{11}\delta C_1$. From

$$\delta C_2 = \frac{\partial C_2}{\partial K_0}\delta K_0 + \frac{\partial C_2}{\partial K_0'}\delta K_0'$$

we find that

$$\delta K_0' = \frac{2}{9K_0}\frac{C_2}{C_1}\delta C_1 - \frac{2}{9K_0}\delta C_2 = B_{21}\delta C_1 + B_{22}\delta C_2 \tag{12}$$

and from

$$\delta C_3 = \frac{\partial C_3}{\partial K_0}\delta K_0 + \frac{\partial C_3}{\partial K_0'}\delta K_0' + \frac{\partial C_3}{\partial K_0 K_0"}\delta(K_0 K_0")$$

we find that

$$\begin{aligned}\delta(K_0K_0") &= -\frac{2}{9K_0}\left[\frac{C_3}{C_1} + (2K_0' - 7)\frac{C_2}{C_1}\right]\delta C_1 + \frac{2}{9K_0}(2K_0' - 7)\delta C_2 + \frac{2}{9K_0}\delta C_3 \\ &= B_{31}\delta C_1 + B_{32}\delta C_2 + B_{33}\delta C_3\end{aligned} \tag{13}$$

and all other elements of $\mathbf{B}$ are zero (i.e. $\mathbf{B}$ is lower left triangular). The (symmetric) model covariance matrix of the $\mathbf{k}$ vector is just

$$\mathbf{W} = \mathbf{B}\mathbf{E}\mathbf{B}^T \tag{14}$$

Now let $b_{21} = C_2/C_1$, $b_{31} = [C_3 + (2K_0' - 7)C_2]/C_1$ and $b_{32} = 2K_0' - 7$, then

$$\begin{aligned}
W_{11} &= \frac{1}{9}E_{11} \\
W_{12} = W_{21} &= \frac{1}{3}\frac{2}{9K_0}\left[E_{11}b_{21} - E_{12}\right] \\
W_{13} = W_{31} &= \frac{1}{3}\frac{2}{9K_0}\left[-E_{11}b_{31} + E_{12}b_{32} + E_{13}\right] \\
W_{22} &= \left(\frac{2}{9K_0}\right)^2\left[E_{11}b_{21}^2 - 2E_{12}b_{21} + E_{22}\right] \\
W_{23} = W_{32} &= \left(\frac{2}{9K_0}\right)^2\left[-E_{11}b_{21}b_{31} + E_{12}\left(b_{31} + b_{32}b_{21}\right) + E_{13}b_{21} - E_{22}b_{32} - E_{23}\right] \\
W_{33} &= \left(\frac{2}{9K_0}\right)^2\left[E_{11}b_{31}^2 + E_{22}b_{32}^2 + E_{33} - 2E_{12}b_{31}b_{32} - 2E_{13}b_{31} + 2E_{23}b_{32}\right]
\end{aligned} \tag{15}$$

and the standard deviations of $\mathbf{k}$ are the square roots of the diagonal elements of $\mathbf{W}$.

Using (14) instead of simpler versions which ignore covariances is important because, even though the covariance matrix of the original data is diagonal, the covariance matrix of $\mathbf{c}$ can be far from diagonal. You can check this by inspecting the correlation matrix defined above.

4. Did we fit the data?

A useful measure of the fit of the model to the data is χ^2 where

$$\chi^2 = \sum_{i=1}^{N}\left(\frac{P_i - \hat{P}_i}{\sigma_i}\right)^2 \tag{16}$$

where $\hat{P}_i$ is the predicted pressure found by dotting $\hat{\mathbf{c}}$ into $\mathbf{A}$. One would expect that, on average $P_i - \hat{P}_i$ would be on the order of a standard deviation so that χ^2 should be of order N, the number of data. There are quantitative statistical tests for deciding if the χ^2 value implies an adequate fit to the data. In our case, you will find that χ^2/N is of order 16. There can be a couple of reasons for this. One is that a few, very precise data are badly fit since the χ^2 measure is sensitive to outliers – you should check for outliers. Another is that the assigned observational errors are just too small to allow an adequate fit to all the data. If this is the case, then an honest thing to do is to increase the size of the errors. In our case, we would have to multiply all the errors by a factor of 4 to get $\chi^2/N \simeq 1$. This factor will propagate throughout the whole analysis of error so that your final errors on $\mathbf{k}$ will also have to be multiplied by 4.