Appendix D Maximizing the log likelihood analytically

Consider the case of attempting to find the best estimates $\hat{M}, \hat{R}$ of the magnitude and source-to-station distance of an earthquake given a single amplitude observation at a single channel. As given in Chapter 4, this requires maximizing the posterior density function, $prob(M, R|Y_A)$, which in turn, is equivalent to maximizing the log likelihood. The log likelihood, L, is given by

$$L = constant - \frac{(Y_A - \bar{Y}(M,R))^2}{2\hat{\sigma}^2} \tag{D.1}$$

where $\bar{Y}(M,R)$ is our envelope amplitude attenuation relationship given by Eqn. 2.3, which is repeated here

$$\begin{aligned}
\log_{10} A_i &= a_i M - b_i(R_{1i} + C_i(M)) - d_i \log_{10}(R_{1i} + C_i(M)) + e_i + \epsilon_i \\
A_i &= \text{ground motion envelope amplitude} \\
M &= \text{CISN magnitude } (M_w \text{ for } M > 5.0) \\
R &= \text{ epicentral distance in km for } M < 5 \\
&\quad \text{ closest distance to fault for } M > 5.0 \text{ (when available)} \\
R_1 &= \sqrt{(R^2 + 9)} \\
C_i(M) &= (\arctan(M - 5) + 1.4)(c_{1i} \exp(c_{2i}(M - 5))) \\
e &= \text{constant} + \text{station corrections} \\
\epsilon &= \text{statistical (or prediction) error}, \sim NID(0, \sigma^2) \\
i &= \text{P-wave amplitude, S-wave amplitude}
\end{aligned}$$

Thus, L is a function of M, R. The M, R that maximize L need to satisfy the following conditions (Eqn. 4.72 in Chapter 4)

$$
\begin{aligned}
\left.\frac{\partial\, L}{\partial M}\right|_{\hat{M},\hat{R}} &= 0 && \text{(D.2)}\\
\left.\frac{\partial\, L}{\partial R}\right|_{\hat{M},\hat{R}} &= 0 && \text{(D.3)}
\end{aligned}
$$

The partial derivatives make Eqn. D.2 a fairly complicated system of equations to solve in terms of M, R.

$$
\begin{aligned}
\frac{\partial L}{\partial M} &= -\frac{1}{\sigma^2}\left(Y_A - \bar{Y}(M,R)\right)\left(-a - K\left(b + d\frac{\log_{10}(e)}{R + C(M)}\right)\right) = 0 && \text{(D.4)}\\
\frac{\partial L}{\partial R} &= -\frac{1}{\sigma^2}\left(Y_A - \bar{Y}(M,R)\right)\left(b + d\frac{\log_{10}(e)}{R + C(M)}\right) = 0 && \text{(D.5)}\\
\text{where } K &= c_1\frac{\exp(c_2(M-5))}{1+(M-5)^2} + c_1\, c_2(\arctan(M-5)+1.4)(\exp(c_2(M-5))) \\
&= c_1\frac{\exp(c_2(M-5))}{1+(M-5)^2} + c_2 C(M) && \text{(D.6)}
\end{aligned}
$$

It is quite difficult to solve Eqns. D.4 and D.5 analytically. An analytic solution to this system of equations is perhaps not the appropriate approach for a real-time application such as seismic early warning. As described in Chapter 4, I chose to find the M, R that maximize L via a brute force, direct-search approach. This involves evaluating L for numerous pairs of M, R and directly trying to find the M, R that maximize L.

In Chapter 4, I discuss how the second derivatives of L define the variance and covariances of the estimates of M, R. That is,

$$
\begin{pmatrix} \sigma_M^2 & \sigma_{M,R}^2 \\ \sigma_{M,R}^2 & \sigma_R^2 \end{pmatrix} = \frac{1}{AB - C^2}\begin{pmatrix} -B & C \\ C & -A \end{pmatrix} = -\begin{pmatrix} A & C \\ C & B \end{pmatrix}^{-1} = -H^{-1} \qquad \text{(D.7)}
$$

where

$$A = \left.\frac{\partial^2 L}{\partial M^2}\right|_{\hat{M},\hat{R}} \qquad B = \left.\frac{\partial^2 L}{\partial R^2}\right|_{\hat{M},\hat{R}} \qquad C = \left.\frac{\partial^2 L}{\partial M \partial R}\right|_{\hat{M},\hat{R}} \tag{D.8}$$

As can be expected from Eqns. D.4 and D.5, the expressions for A, B, C in terms of M, R are quite cumbersome.

$$\begin{aligned} \frac{\partial^2 L}{\partial M^2} &= -\frac{1}{\sigma^2}\left(\left(Y_{obs} - \bar{Y}(M,R)\right)\frac{\partial}{\partial M}\left(-a - K\left(b + d\frac{\log_{10}(e)}{R + C(M)}\right)\right)\right) \\ &\quad -\frac{1}{\sigma^2}\left(-a - K\left(b + d\frac{\log_{10}(e)}{R + C(M)}\right)\right)^2 \end{aligned} \tag{D.9}$$

where K is given by Eqn. D.6.

$$\frac{\partial^2 L}{\partial R^2} = -\frac{1}{\sigma^2}\left(-d\frac{\log_{10}(e)(Y_{obs} - \bar{Y}(M,R))}{(R + C(M))^2} + \left(b + d\frac{\log_{10}(e)}{R + C(M)}\right)^2\right) \tag{D.10}$$

$$\frac{\partial^2 L}{\partial M \partial R} = \frac{1}{\sigma^2}\left(d\frac{\log_{10}(e)(Y_{obs} - \bar{Y})}{(R + C(M))^2}\left(c_1\frac{e^{c_2(M-5)}}{1 + (M-5)^2} + c_2 C(M)\right)\right) \tag{D.11}$$

$$+\frac{1}{\sigma^2}\left(1 - a - K\right)\left(b + d\frac{\log_{10}(e)}{R + C(M)}\right) \tag{D.12}$$

The second derivatives of L relative to M, R evaluated at $\hat{M}, \hat{R}$ give information regarding the variances and covariances of the best estimates of magnitude and distance. If we could find M, R that maximize L, then we could do a Taylor series expansion about this maxima $(\hat{M}, \hat{R})$ to get an idea of the spread of the posterior density function, and hence the variances and covariances of the estimates. However, we can expect to run into difficulties with the analytical approach, since L is not a simple function of M and R, and hence, the partial derivatives are complicated. While conceptually simple, it is difficult to solve Eqn. D.2 analytically for the $(\hat{M}, \hat{R})$ that maximize the posterior. We thus choose to approach the problem of maximizing L using a brute force, direct-search method.