Reconciling the Likelihood Formulas in Brown and Sundberg 1987

Brown uses three different expressions for profile likelihood: (2.14), (2.4-2.5) and (3.3, 3.6). To some extent, the expressions appear magically for people not used to the area ad these notes are exercises working out the connections between the different expressions (which reconcile).

Linking 2.14 to 2.4

1987 Equation 2.14, implemented in my profile.lik function, is:

Equation 2.4, implemented in my brown.lik function, is:

These expressions connect in the sense that both expressions yield the same maximum likelihood value for ξ, but it is a bit of an exercise connecting them and the notation can get a bit confusing. The RHS of 2.5 is the (qxq) profile residual matrix for the OLS fit and where is the calibration matrix Y augmented by the information Z to be calibrated and the calibration matrix X is augmented by the profile candidate ξ.

With modern computers, this isn't much work to handle numerically. In older statistics papers, considerable ingenuity is often spent developing methods seemingly to conserve calculation effort and 1987 is perhaps "older" in this respect here. Brown uses some matrix and determinant identities to express the augmented residual matrix in terms of the calibration residual matrix and other terms that require less calculation than re-doing the OLS regression for each ξ in the profile.

At the risk of introducing more inconsistent notation, I'm going to experiment a little here to try to show how things tie together for adding one prediction data vector to the mix. Let's use Y and X to denote the calibration matrices (as with Brown) and y to denote the prediction vector of values (instead of Z).

Through various matrix identities on page 48, Brown and Sundberg re-express the augmented residual matrix (which is a function of ξ), both expressions on the RHS being qxq matrices, the RHS here being (2.13) in slightly different notation. I confirmed that this form of re-expression works.

We'll return in a moment to the expression r(ξ), which is a slight re-expression of .
Because the determinant of the above expressions is used in the log-likelihood expression (2.4), Brown uses a pretty determinant identity to re-express the determinant of the RHS as a scalar (a similar sort of argument occurs in the Wikipedia on ____) , k here just simplifies notation for the scalar in the above expression.

and then applying the determinant identity shown on page 48 using the matrix identity to convert a qxq determinant into a 1x1 determinant. Watch the change in order of the matrices. Slick (and it works).

The expression r(ξ) is defined and is related to as follows (the first step being the definition of r(ξ)):

We can use these two results to get from (2.14) to (2.4). If , according to Brown and Sundberg when only 1 vector is added into the mix, then the (2.14) expression can be written as shown below:

This is 1 divided by the determinant expression already calculated for the augmented residual matrix. The maximum of one expression will coincide with the minimum of the other expression. I haven't experimented with l>1.

The bottom line is that these are both expressions relating to the determinant of the OLS residuals (augmented matrix).

Linking 3.3 to 2.13

Later Brown and Sundberg introduce another expression for profile likelihood in 3.3 which is valid but appears a little magically.

Expression (2.13) was discussed above:

Recall the definition: . Brown uses the notation but I'm going to carry forward the only slightly unsimplified expression. Thus,

Multiplying the scalar inside the curlicue brackets in (2.13) by (using Brown's definition of m), one gets

Brown then does another bit of manipulation by expressing where is the GLS estimate (shown in equation 3.1, which defines H). The purpose here is to link the maximum likelihood residuals to the Inconsistency R, which is an expression based on the residuals to the GLS estimate . Writing things out the long way as an exercise, if one can express the matrix expression on the far right of the most recent expression as:

The last line results from the fact that the middle expressions are identically 0, the first term is the definition of R and the fourth term applies the ||.|| notation as used in Brown and Sundberg.