CFSYSTEMS

Least Squares Residuals

CFS-246

February 23, 2000

David Dunthorn
www.c-f-systems.com

Note

This document derives the details of forming the least squares matrix and extracting the residual sum of squares from the least squares matrix equation. In particular, it includes weighting factors in the development, an element that is usually missing in standard derivations. It was prepared in the process of developing the computer code for Precise Signal Component. It was not given a CFS number at that time, so it has been assigned a current CFS number for release.

This document is viewable only. I believe this will be adequate for people who do not intend to study it. Please contact me through our web site if you need a printable version. I am aware that the no-print can be defeated, but again I ask that you contact me instead. I really need to know if and how people are finding these documents useful, and this seems one of the few ways I have to encourage feedback.

A Note on Least Squares Residuals

The general linear least squares equation is:

$$y = a_0 + a_1 x_1 + a_2 x_2 + ... + a_n x_n \tag{1}$$

The general least requirement is to minimize:

$$\sum_{i=1}^{N} w_i (a_0 + a_1 x_{1i} + a_2 x_{2i} + ... + a_n x_{ni} - y_i)^2 \tag{2}$$

where N is the number of data points being treated and w_i is a weighting factor applied to the i^{th} point. Differentiating this and setting it to zero leads to the so-called *normal equations*:

$$a_0 \sum_i w_i + a_1 \sum_i w_i x_{1i} + a_2 \sum_i w_i x_{2i} + ... + a_n \sum_i w_i x_{ni} = \sum_i w_i y_i$$

$$a_0 \sum_i w_i x_{1i} + a_1 \sum_i w_i x_{1i}^2 + a_2 \sum_i w_i x_{1i} x_{2i} + ... + a_n \sum_i w_i x_{1i} x_{ni} = \sum_i w_i y_i x_{1i}$$

$$a_0 \sum_i w_i x_{2i} + a_1 \sum_i w_i x_{2i} x_{1i} + a_2 \sum_i w_i x_{2i}^2 + ... + a_n \sum_i w_i x_{2i} x_{ni} = \sum_i w_i y_i x_{2i}$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$a_0 \sum_i w_i x_{ni} + a_1 \sum_i w_i x_{ni} x_{1i} + a_2 \sum_i w_i x_{ni} x_{2i} + ... + a_n \sum_i w_i x_{ni}^2 = \sum_i w_i y_i x_{ni}$$

Often written in matrix form as $\mathbb{X}\mathbf{a} = \boldsymbol{Y}$.

It is convenient to eliminate the odd case of a_0. There are two basic ways of doing this. Commonly, the first equation is solved for a_0 and the result substituted into all the remaining equations. This results in many terms of the form $\sum_i w_i x_{ki} x_{li} - \sum_i w_i x_{ki} \sum_i w_i x_{li} / \sum_i w_i$ and these terms are given a shorthand notation, $\sum' x_{ki} x_{li}$, often called the "reduced summation." This is a little more clear when all the $w_i = 1$. Then $\sum' x_{ki} x_{li} \equiv \sum_i x_{ki} x_{li} - \sum_i x_{ki} \sum_i x_{li} / N$. The other approach is to simply drop the a_0 term and let all the $x_{1i} = 1$. Then a_1 becomes the constant term. This approach will be used here, partially because the constant term in the target equation is actually complex, and partially because non-unity weighting factors will be used. Both of these factors tend to make the common practice confusing. The two approaches are algebraically identical, although the numerical computation considerations are different.

So, we have:

$$a_1 \sum_i w_i x_{1i}^2 + a_2 \sum_i w_i x_{1i} x_{2i} + ... + a_n \sum_i w_i x_{1i} x_{ni} = \sum_i w_i y_i x_{1i}$$

$$a_1 \sum_i w_i x_{2i} x_{1i} + a_2 \sum_i w_i x_{2i}^2 + ... + a_n \sum_i w_i x_{2i} x_{ni} = \sum_i w_i y_i x_{2i}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$a_1 \sum_i w_i x_{ni} x_{1i} + a_2 \sum_i w_i x_{ni} x_{2i} + ... + a_n \sum_i w_i x_{ni}^2 = \sum_i w_i y_i x_{ni}$$

where all the $x_{1i} = 1$. This system of normal equations can be written as:

$$\sum_{l=1}^{n} a_l \sum_{i=1}^{N} w_i x_{ki} x_{li} = \sum_{i=1}^{N} w_i y_i x_{ki} \tag{3}$$

where $k = 1, 2, ...n$. We require that

$$\sum_{i=1}^{N} w_i (a_1 x_{1i} + a_2 x_{2i} + ... + a_n x_{ni} - y_i)^2 \tag{2a}$$

be minimized and we wish to evaluate this residual sum of squares that cannot be further reduced by the regression at hand. First we expand the squared term:

$$(a_1 x_{1i} + a_2 x_{2i} + ... + a_n x_{ni} - y_i)^2 =$$

$$a_1^2 x_{1i} + a_1 a_2 x_{1i} x_{2i} + ... + a_1 a_n x_{1i} x_{ni} - a_1 y_i x_{1i} +$$

$$a_2 a_1 x_{2i} x_{1i} + a_2^2 x_{2i}^2 + ... + a_2 a_n x_{2i} x_{ni} - a_2 y_i x_{2i} +$$

$$...$$

$$a_n a_1 x_{ni} x_{1i} + a_n a_2 x_{ni} x_{2i} + ... + a_n^2 x_{ni}^2 - a_n y_i x_{ni} +$$

$$- a_1 y_i x_{1i} - a_2 y_i x_{2i} - \ldots - a_n y_i x_{ni} + y_i^2$$

So the sum of Equation (2a) can be written as

$$\sum_{i=1}^{N} w_i \sum_{k=1}^{n} a_k x_{ki} \sum_{l=1}^{n} a_l x_{li} - 2\sum_{i=1}^{N} w_i \sum_{k=1}^{n} a_k y_i x_{ki} + \sum_{i=1}^{N} w_i y_i^2 \tag{4}$$

Let us first examine the leftmost sum:

$$\sum_{i=1}^{N} w_i \sum_{k=1}^{n} a_k x_{ki} \sum_{l=1}^{n} a_l x_{li} = \sum_{i=1}^{N} w_i \sum_{k=1}^{n}\sum_{l=1}^{n} a_k a_l x_{ki} x_{li} =$$

$$\sum_{i=1}^{N}\sum_{k=1}^{n}\sum_{l=1}^{n} a_k a_l w_i x_{ki} x_{li} = \sum_{k=1}^{n}\sum_{l=1}^{n} a_k a_l \sum_{i=1}^{N} w_i x_{ki} x_{li} =$$

$$\sum_{k=1}^{n} a_k \sum_{l=1}^{n} a_l \sum_{i=1}^{N} w_i x_{ki} x_{li}$$

The portion of the above equation coming after a_k is exactly the same as the left hand sides of the normal equations as given in Equation (3). Thus:

$$\sum_{k=1}^{n} a_k \sum_{l=1}^{n} a_l \sum_{i=1}^{N} w_i x_{ki} x_{li} = \sum_{k=1}^{n} a_k \sum_{i=1}^{N} w_i y_i x_{ki} \tag{5}$$

Substituting this result into Equation (4), we find that:

$$\sum_{k=1}^{n} a_k \sum_{i=1}^{N} w_i y_i x_{ki} - 2\sum_{i=1}^{N} w_i \sum_{k=1}^{n} a_k y_i x_{ki} + \sum_{i=1}^{N} w_i y_i^2$$

or

$$\sum_{i=1}^{N} w_i y_i^2 - \sum_{k=1}^{n} a_k \sum_{i=1}^{N} w_i y_i x_{ki} \tag{6}$$

Furthermore, the a_k's are computed in the least squares regression and the $\sum_{i=1}^{N} w_i y_i x_{ki}$ terms are the elements of the Y vector used in the least squares calculation, so finally we have:

$$\text{Residual Sum of Squares} = \sum_{i=1}^{N} w_i y_i^2 - \sum_{k=1}^{n} a_k Y_k \tag{7}$$

making the algebraic computation of this value easy, although it is potentially difficult numerically, being the difference between two nearly identical values. This final form is correct regardless of whether the form shown or the "reduced summation" is used.