# Appendix 2: The Principle of Least Squares

#### Martin Veillette

## Least Squares and Sample Means

When we make multiple measurements of a quantity that contains random fluctuation, we need a method to determine the most probable value. The principle of least squares provides this method by finding the value that minimizes the squared deviations from our measurements.

Let’s say we make measurements, , of a quantity. To find the value whose deviations from our measurements are minimized according to the principle of least squares, we need:

(1)

Let’s denote the mean of the measurements as . We can rewrite the sum of squared deviations as:

(2)

Expanding the squared term:

(3)

The cross-term equals zero by definition of the mean, so:

(4)

This expression clearly reaches its minimum value when , confirming that using the sample mean as the most probable value is consistent with the principle of least squares.

## Fitting a Straight Line Using Least Squares

Consider a set of observations that we wish to fit with a linear relationship:

(5)

We’ll assume that uncertainty exists only in the values, and that all measurements have equal weight (we’ll address weighted least squares later).

For each observation, the deviation from our proposed line is:

(6)

According to the principle of least squares, we want to minimize the sum of the squares of these deviations:

(7)

Expanding this expression:

(8)

Or more compactly:

(9)

Where represents the sum of squared deviations that we want to minimize.

To find the optimal values of and , we take partial derivatives with respect to each parameter and set them equal to zero:

(10)

From the first condition:

(11)

From the second condition:

(12)

Solving these equations simultaneously gives us:

(13)

(14)

Having determined the “best fit” line, we need to quantify the uncertainty in our calculated parameters. Since and are computed from measurements with uncertainty, we can calculate their standard deviations.

For the standard deviation of each value from our fitted line, we use:

(15)

The denominator uses rather than because we’ve estimated two parameters ( and ) from our data, reducing our degrees of freedom.

The standard deviations of the slope and intercept are then:

(16)

(17)

These expressions provide statistical measures of uncertainty in our fitted parameters. When reporting results, we typically state values as and , indicating that the true parameter has about a 68% probability of falling within one standard deviation of our estimate.

## Weighted Least Squares

When measurements have different levels of precision, it makes sense to give more weight to more precise measurements. This approach is called weighted least squares.

### Weighted Mean of Observations

If we have independently measured quantities , each with a standard deviation , the weighted mean is:

(18)

The standard deviation of this weighted mean is:

(19)

### Straight-Line Fitting with Weighted Least Squares

For observations with unequal precision, we modify our least squares approach by assigning weights. If the values have varying precision, but the values are considered exact, the equations for the slope and intercept become:

(20)

(21)

Where represents the weight of each observation, calculated as:

(22)

The weighted standard deviation about the best-fit line is:

(23)

And the standard deviations of the slope and intercept are:

(24)

(25)

Where:

(26)

And is the weighted mean of the values:

(27)

Weighted least squares is particularly valuable when measurements come from different sources with varying precision. By accounting for these differences in precision, we ensure that our fitted parameters are not unduly influenced by less reliable data points.

When reporting results from weighted analyses, it’s important to specify that weighted methods were used and to explain the basis for the weights assigned. This transparency allows others to properly interpret and potentially reproduce your analysis.