WEIGHTING OF LINEAR PLOTS IN ENZYME KINETICS

Athel Cornish-Bowden

Department of Biochemistry, University of Birmingham,
P.O. Box 363, Birmingham B15 2TT, U.K.

Received April 22, 1982

Abstract

If the rates of an enzyme-catalysed reaction are equally weight-
ed in least-squares analysis, the corresponding reciprocal rates \(1/v\) should be given weights of \(v^{2-1}\), where \(v\) is the cal-
culated rate, in order to give the same sum of squares as for
the untransformed Michaelis-Menten equation. However, these
weights are functions of the parameter values and therefore
cannot be treated as constants in partial differentiation with
respect to the parameters. Consequently, linear regression
with these weights does not give the same parameter estimates
as non-linear regression of the untransformed equation. The
error can be exactly corrected by using weights of \(v^{2-1}\) in the
linear regression.

Introduction

"It was found upon analysis of the data that equal weighting was to be as-
signed to each measured \(v\) value, which means that the weighting of \(1/v\) is pro-
portional to \(v^4\)."

Dean Burk, 1934 [1]

In the folklore of biochemistry Lineweaver and Burk [2] are given the blame for in-
roducing a method for estimating the parameters of the Michaelis-Menten equation
quite different from the one they actually used. Although their commonly cited pa-
per does not itself give details of how they did their calculations, it contains
references to two others [1, 3] that do. We can see from the statement that I have
quoted at the beginning of this paper that they were aware of three important points
about least-squares analysis: (a) least-squares estimation requires weights; (b) these
should be based on observation, not on surmise; (c) equal weighting of rates implies
grossly unequal weighting of their reciprocals. The first and third of these points
did not begin to penetrate the general biochemical consciousness until a quarter of
a century later, with the publication of papers by Johansen and Lumry [4] and Wilkin-
son [5]. The second point is still little recognized, and even now review articles
continue to be written that advocate weighting assumptions that take no account of
published investigations of actual error behaviour in enzyme kinetics [7-10].
In the present paper I want to examine an inconsistency in the way supposedly equivalent ways of analysing the Michaelis-Menten equation behave. Although it is in most cases a numerically trivial inconsistency it is nonetheless worthy of study, because even a numerically trivial inconsistency casts doubt on the validity of a mathematical method if it cannot be explained.

I. Weighting of the Double-Reciprocal Plot

The Michaelis-Menten equation expresses the dependence of an observed initial rate \( v_i \) on the substrate concentration \( s_i \), and contains two parameters, \( V \) and \( K_m \), that need to be estimated:

\[
v_i = \frac{V_i}{K_m + s_i} + e_i
\]

The error term \( e_i \) in this equation expresses the fact that real experiments cannot be carried out without experimental error, and so \( v_i \) is not in general exactly equal to \( \frac{V_i}{K_m + s_i} \).

The least-squares method of estimating \( V \) and \( K_m \) requires us to define a weighted sum of squares \( SS \),

\[
SS = \sum w_i e_i^2
\]

in which the summation is for \( i = 1 \) to \( n \), i.e., it is carried out over all of \( n \) observations, and the weight \( w_i \) to be assigned to each \( v_i \) should be inversely proportional to the variance of that \( v_i \). In the present paper I shall not consider how these weights should be defined (but see refs. (7-12)), but will treat them as known constants. We now define the best-fit values \( \hat{V} \) and \( \hat{K}_m \) as the values of \( V \) and \( K_m \) respectively that make \( SS \) a minimum. I shall also define

\[
\hat{V}_i = \frac{V_i}{K_m + s_i}
\]

as the calculated rate corresponding to the observed rate \( v_i \).*

Non-linear regression\(^{[4-5]}\) provides one method of minimizing \( SS \). Here, however, I shall discuss the alternative method of using a properly weighted linear regres-

* Although it might seem more logical to use the best-fit parameter values \( \hat{V} \) and \( \hat{K}_m \) on the right-hand side of eqn. (3), it is considerably more useful to define \( \hat{V}_i \) as shown, and so I shall sacrifice strict logical consistency for convenience.
Weighting of linear plots in enzyme kinetics

If correctly carried out this gives exactly the same parameter estimates as non-linear regression and is computationally just as convenient; but it has the added advantages that it is (in my view) easier to understand and it sheds light on the widely used straight-line plots of the Michaelis-Menten equation. The best known of these is the double-reciprocal plot, based on the following form of the Michaelis-Menten equation:

\[ \frac{1}{v} = \left( \frac{K_m}{V} \right) \left( \frac{1}{x} \right) + \frac{1}{V} + e' \]  

(4)

This equation cannot be derived from Eq.(1) simply by taking reciprocals of both sides, because the error term \( e' \) in Eq.(1) prevents its right-hand side from having a simple reciprocal resembling the right-hand side of Eq.(4). It follows that \( e' \) is not the same as \( e \), but simple algebra shows that they are related as follows:

\[ e' = -\epsilon_0 (K_0 + x) / v \]  

(5)

Consequently, if we define the sum of squares \( SS' \) for the double-reciprocal plot as follows:

\[ SS' = \sum \frac{w_i e'^2}{2} \]  

(6)

it will be different from the original sum of squares \( SS \) unless the new weights \( w_i \) are defined in a way that takes account of Eq.(5):

\[ w_i = w_i e'/e' \]  

(7)

II. Minimization of the Sum of Squares

For a straight line \( y = a + bx + e' \) the formulae for minimizing the weighted sum of squares \( SS' = \sum w_i e'^2 \) are as follows:

\[ b = \frac{\sum w_i x_i y_i - \sum w_i x_i \sum w_i y_i}{\sum w_i x_i^2 - (\sum w_i x_i)^2} \]  

(8)

\[ a = \frac{\sum w_i x_i - b \sum w_i x_i^2}{\sum w_i} \]  

(9)

We can apply these formulae to Eq.(4) by means of the substitutions \( V = 1/a, K_m = b/a, x_0 = 1/b \), \( y_0 = 1/v \), \( w_i = w_i e'/e' \). Initially there are no values of \( V \) and \( K_m \) for use in calculating \( y_0 \), but as a first approximation we can assume that the measurements are accurate enough to replace \( y_0 \) by \( y_0 \), i.e., to define \( w_i = w_i e'/e' \). Then Eqs.(8-9) allow the calculation of first-approximation values of \( V \) and \( K_m \), which can be used to calculate \( y_0 \), and the cycle of calculations can be repeated until \( V \) and \( K_m \) do not change.
from one cycle to the next.

Now this argument, which is essentially as I have presented it previously\cite{13}, may seem plausible. But plausible or not, it is incorrect, as can be seen from the results shown in Table 1. For the data used as an illustration by Wilkinson\cite{5}, and with \( w_i = 1 \) for all \( i \), \( \hat{V} \) and \( \hat{k_m} \) have been calculated by a direct non-linear fit to Eq. (1), as described by Wilkinson\cite{5}, and also by linear regression of Eq. (4) with weights \( w'_i \) defined in various ways. Not surprisingly, the preliminary results with \( w'_i = \frac{1}{\nu^4} \) are different from those given by non-linear regression, but in addition the results from the iterative approach I have described, with \( w'_i = \frac{\nu^3}{\nu^4} \), are also different, as are those obtained with \( w'_i = \frac{1}{\nu^2} \), as suggested by Cleland\cite{14}. Only a calculation with the apparently arbitrary compromise of \( w'_i = \frac{\nu^3}{\nu^4} \) gives the same parameter estimates as direct non-linear regression. These variations in results are not a consequence of numerical problems, such as rounding error; they reflect the fact that the argument given above for carrying out the linear regression with \( \frac{1}{\nu^2} \) is invalid, for reasons that I shall now discuss.

### Table 1. Results of fitting the same data with different weights.

The data used were those given by Wilkinson\cite{5} for illustration, and consisted of the following six \((s, v)\) pairs: (0.138, 0.148), (0.220, 0.171), (0.291, 0.234), (0.560, 0.324), (0.766, 0.390), (0.146, 0.493).

<table>
<thead>
<tr>
<th>Weighting</th>
<th>( \hat{k_m} )</th>
<th>( \hat{V} )</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight for ( \frac{1}{\nu^4} )</td>
<td>0.59655</td>
<td>0.69040</td>
<td>[5]</td>
</tr>
<tr>
<td>Weight for ( \frac{1}{\nu^2} )</td>
<td>0.57097</td>
<td>0.67986</td>
<td>[1,3-5]</td>
</tr>
<tr>
<td>Weight for ( \frac{1}{\nu^2} )</td>
<td>0.58770</td>
<td>0.68672</td>
<td>[13]</td>
</tr>
<tr>
<td>Weight for ( \frac{1}{\nu^3} )</td>
<td>0.59655</td>
<td>0.69040</td>
<td>This paper</td>
</tr>
<tr>
<td>Weight for ( \frac{1}{\nu^2} )</td>
<td>0.60553</td>
<td>0.69416</td>
<td>[14]</td>
</tr>
</tbody>
</table>

### III. The Correct Weighting Function

The derivation of Eqs. (8-9) for fitting a straight line is made with the assumption that the weights \( w'_i \) are independent of the parameter values, so that they can be treated as constants during partial differentiation. As defined by Eq. (7), however, they are not constants because \( \nu' \) is a function of \( \nu \) and \( k_m \).

Partial differentiation of Eq. (6) with respect to \( k_m \) gives

\[
\frac{\partial SS}{\partial k_m} = 2 \sum w'_i \cdot \frac{\partial \nu'}{\partial k_m} + \sum s_i \cdot \frac{\partial w'_i}{\partial k_m}
\]

(10)
Weighting of linear plots in enzyme kinetics

The second term in this equation is simply a consequence of the dependence of \( \hat{v} \)
and hence \( w_i' \) on \( \bar{K}_m \). Consideration of Eqs. (3), (5) and (7), with some rearrangement, allows this derivative to be written as follows:

\[
\frac{dSS}{dK_m} = 2 \sum \frac{w_i' \hat{e}_i (e_i / v)}{\hat{v}^2} \frac{\partial \hat{v}'}{\partial K_m}
\]

(11)

Comparison with Eq. (10) shows that this is exactly the expression we should have
obtained if we had replaced \( w_i' \) by \( w_i' \hat{v} / \hat{v}_m \), initially but treated it as a constant in
the partial differentiation. The corresponding result may be shown by an analogous
argument to apply to partial differentiation with respect to \( v \). If follows, there-
fore, that the error of ignoring the dependence of the weights on the parameter val-
ues can be exactly corrected by carrying out the linear regression with refined
weights \( w_i' = v - \frac{v}{\hat{v}_m} \hat{v}_m \) instead of \( w_i' = v - \frac{v}{\hat{v}_m} \hat{v} \). Thus the results of Table 1 are explained.

It is perhaps worth adding that the need to replace \( w_i' = \frac{v}{\hat{v}_m} \) with \( w_i' = v - \frac{v}{\hat{v}_m} \hat{v}_m \) ap-
plies only to the iterative process for estimating the parameters \( v \) and \( \bar{K}_m \). For
calculation of the sum of squares at the end of the computation the original weights
\( w_i' = v - \frac{v}{\hat{v}_m} \hat{v}_m \) must be used, because the conclusion that for \( SS \) as defined in Eq. (6) to
be identical with \( SS \) as defined in Eq. (2) requires these weights be independent of
any considerations of partial differentiation.

IV. Discussion

The variation in results given by the various weighting schemes shown in Table 1 is
not great, and is certainly trivial compared with the error incurred by fitting
data to a double-reciprocal plot by eye or by unweighted linear regression**. On
many occasions there is likely to be very little advantage in proceeding beyond the
first-approximation fit obtained with \( w_i' = v - \frac{v}{\hat{v}_m} \hat{v}_m \). If one does decide to go to the ap-
preciable extra computational labour of using refined weights, however, it is ob-
viously better to use the correct formula \( w_i' = v - \frac{v}{\hat{v}_m} \hat{v}_m \), which takes the calculation
100% of the way to the correct solution, rather than either \( w_i' = v - \frac{v}{\hat{v}_m} \hat{v}_m \), which will
take it only 67% of the way, or \( w_i' = v - \frac{v}{\hat{v}_m} \hat{v}_m \), which will take it 33% too far.

I have concentrated in this paper on the weighting of the double-reciprocal
plot, because it is the best known and most widely used plot of the Michaelis-Menten
equation. However, essentially the same kind of analysis can be applied to the al-
ternative plot of \( s_j / v_j \) against \( s_j \), which has the advantage that the weights are
must less steeply dependent on \( \hat{v}_j \). For this plot the first-approximation weights
are \( w_i' = v - \frac{v}{\hat{v}_m} \hat{v}_m \) and the refined weights are \( w_i' = v - \frac{v}{\hat{v}_m} \hat{v}_m \).

** If the same data as in Table 1 are fitted with \( w_i' = 1 \) for each value of \( 1 / v_j \) the
results are \( \bar{K}_m = 0.44062, \bar{v} = 0.38532 \).
References