Demonstrate how transformed (forced) linear regression frequently gives a bad fit.
Instructor: Nam Sun Wang

Raw Data:

<table>
<thead>
<tr>
<th>Reactant Conc. (g/L)</th>
<th>Reaction Rate (g/L-h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>i _i _i _i _i _i</td>
<td>i _i _i _i _i _i</td>
</tr>
<tr>
<td>1 2 3 4 5</td>
<td>1 4 5 6 6</td>
</tr>
</tbody>
</table>

Given rate equation:

\[ r = \frac{a \cdot s}{b + s} \]

The above rate equation can be transformed into a linear form:

\[ \frac{1}{r} = \frac{1}{a} + \frac{1}{a} b \cdot \frac{1}{s} \rightarrow \text{intercept} + \text{slope} \cdot \frac{1}{s} \]

Thus, if we plot \(1/r\) vs. \(1/s\), the intercept is \(1/a\) and the slope is \(b/a\).

Regression line in the inverted variable space

\[ \text{line}(S) := \text{A} + \text{B} \cdot S \]

Find the coefficient in the original rate equation

\[ a := \frac{1}{\text{intercept}(S, R)} \quad a = -7.46 \]
\[ b := \text{slope}(S, R) \cdot a \quad b = -8.016 \]

Note that the fit in the inverted variable space seems reasonable overall, although there seems to be a lot of emphasis on the very last point. The fact that the coefficients \(a\) and \(b\) are negative tells us that there may be a problem later.
Rate expression with the best fit parameters $a$ and $b$:

\[
\text{rate}(s) := \frac{a \cdot s}{b + s}
\]

Note that the saturation rate expression is supposed to flatten out instead of rising sharply.

\[ss := 0, 0.1, \ldots, 5\]