USING ROTATION TRANSFORMATIONS TO MAXIMIZE THE COEFFICIENT OF DETERMINATION IN SIMPLE LINEAR REGRESSION MODELS

Ross Dickinson
University of West Florida
USA

Susan Perry
Pensacola State College
USA

Subhash Bagui
University of West Florida
USA

ABSTRACT

In many instances, excessive variation in observed data limits the utility of simple linear regression. We show that a simple rotation of the coordinate axis about the origin reduces the observed variance in the data, improves estimates of the slope, and increases the coefficient of determination. Furthermore, we provide a modified version of the basic regression model that accommodates a rotation angle and a method for determining the angle that maximizes the coefficient of determination.

Keywords: Linear Regression, Transformation, Rotation Matrix, Maximization, and Coefficient of Determination.

INTRODUCTION

Simple regression, based on the method of least squares, produces the best unbiased estimators for the slope and intercept of fitted lines. However, we've found that the coefficient of determination and the estimate of the slope can be improved by rotating the observed data about the origin. We propose adding a rotation angle to the standard linear regression model and provide a method for computing the value of the angle that maximizes the coefficient of determination (R^2).

The main body of this paper is arranged into four sections. In Section 1, we provide a short overview of the basic regression model. In Section 2, we provide a standard coordinate transformation that rotates the x-y axis through an angle ϕ about the origin then we re-cast the standard model in terms of ϕ. In Section 3, we provide a computational method for determining θ, where θ is the argument of ϕ that maximizes R^2. In Section 4, we use the compounded model to analyze the duration and time interval between Old Faithfull eruptions in Yellowstone National Park. Finally, we close with a summary of the enhanced model's strengths and weaknesses.

DISCUSSION

The Standard Model [3][4]

In statistics, the standard model used to fit a straight line to a set of "n" paired observations $\{(x_i, y_i), i = 1 \cdots n\}$ is given by equation (1) where the individual errors (ε_i's) are assumed to be independent, and normally distributed, with zero mean and constant variance.

$$y_i = \alpha + \beta x_i + \varepsilon_i \tag{1}$$

The coefficient of determination (R^2) measures the degree to which the fitted line explains the variation in the data. Formulas for computing R^2 and estimates of the slope $(\hat{\beta})$ and intercept $(\hat{\alpha})$ are given by equations (2a-c) respectively.

$$R^2 = \frac{COV(x,y)^2}{VAR(x) \cdot VAR(y)} \quad (2a)$$

$$\hat{\beta} = \frac{COV(x,y)}{VAR(x)} \quad (2b)$$

$$\hat{\alpha} = \overline{y} - \overline{x} \cdot \hat{\beta} \quad (2c)$$

Equations (2b-c) produce unbiased estimates of α and β provided that the basic assumptions are satisfied. The resulting "fitted line" is given by equation (2d)

$$\hat{y}_i = \hat{\alpha} + \hat{\beta} x_i \quad (2d)$$

Coordinate Transformations

The following discussion is limited to passive transformations in which a positive fixed angle rotates the x-y axis counter-clockwise about the origin while the relative position and orientation of points and lines remain fixed [1]. Negative angles rotate the x-y axis in the opposite (clockwise) direction [2].

Transforming Data on $\{(x_i, y_i)\ i = 1 \cdots n\}$:

The simple rotation matrix, given by equation (3a) [1], transforms the values of paired observations (x_i, y_i) in the original reference frame into (x_{ir}, y_{ir}), their equivalent values in the rotated reference frame as shown.

$$\begin{bmatrix} x_{ir} \\ y_{ir} \end{bmatrix} = \begin{bmatrix} cos(\emptyset) & sin(\emptyset) \\ -sin(\emptyset) & cos(\emptyset) \end{bmatrix} \cdot \begin{bmatrix} x_i \\ y_i \end{bmatrix} \quad (3a)$$

Theta (ϕ) is a positive rotation angle and the subscript "r" indicates that the assigned values are taken in the rotated reference frame. The corresponding rotation equations [2] for a counter-clockwise transformation are given by

$$x_{ir} = x_i cos(\phi) + y_i sin(\phi) \quad (3b)$$

$$y_{ir} = -x_i sin(\phi) + y_i cos(\phi) \quad (3c)$$

The five key elements of the basic model are derived in three steps: In the first step we apply equation (3) to "n" paired observations $\{(x_{ir}, y_{ir}), i = 1 \cdots n\}$ then treat the results as a linear model with y_{ir} on x_{ir}. In the second step we compute the means, variances, and covariance of the transformed data. Finally, we apply equations (2a-c) to the results. Alternatively, the means, variances, and covariance of the transformed data may be computed directly, for any rotation angle, using equations (4a-e).

$$\overline{x}_r = \overline{x} \cdot cos(\phi) + \overline{y} \cdot sin(\phi) \quad (4a)$$

$$\overline{y}_r = -\overline{x} \cdot sin(\phi) + \overline{y} \cdot cos(\phi) \quad (4b)$$

$$VAR(x_r) = \left(VAR(x)cos^2(\emptyset) + 2COV(x,y)sin(\emptyset)cos(\emptyset) + VAR(y)sin^2(\emptyset)\right) \quad (4c)$$

$$VAR(y_r) = \left(VAR(x)sin^2(\emptyset) - 2COV(x,y)sin(\emptyset)cos(\emptyset) + VAR(y)cos^2(\emptyset)\right) \quad (4d)$$

$$COV(x_r, y_r) = sin(\phi)cos(\phi)\left(VAR(y) - VAR(x)\right) + \left(cos^2(\phi) - sin^2(\phi)\right)COV(x,y) \quad (4e)$$

The coefficient of determination and estimates of the slope and intercept can now be stated in-terms of any arbitrary rotation angle using equation (5a-c)

$$R_r^2 = \frac{COV(x_r, y_r)^2}{VAR(x_r) \cdot VAR(y_r)} \quad (5a)$$

$$\hat{\beta}_r = \frac{COV(x_r, y_r)}{VAR(x_r)} \quad (5b)$$

$$\hat{\alpha}_r = \overline{y}_r - \overline{x}_r \cdot \hat{\beta}_r \quad (5c)$$

where the "fitted line" in the rotated reference frame is given by equation (5d).

$$\hat{y}_{ir} = \hat{\alpha}_r + \hat{\beta}_{ir} x_{ir} \quad (5d)$$

Transforming fitted lines:

Similarly equations (6a-b) are used to transform equations of straight lines in one reference frame into equivalent lines in a second rotated reference frame. Negative rotation angles return the inverse transformation of equation (6a-b) and reverse the direction of rotation [2]. The derivation of the basic rotation matrix leading to equations (6a-b) is given in Appendices A and B

$$\beta_t = \frac{-sin(\emptyset) + \beta cos(\emptyset)}{cos(\emptyset) + \beta sin(\emptyset)} \quad (6a)$$

$$\alpha_t = \alpha\left(cos(\emptyset) + \frac{sin^2(\emptyset) - \beta sin(\emptyset)cos(\emptyset)}{cos(\emptyset) + \beta sin(\emptyset)}\right) \quad (6b)$$

Here, α and β are the intercept and slope of a fixed line in the first reference frame and the subscript "t" denotes their equivalent values under rotation.

Applying equation (6a-b) to the fitted line in the original reference frame transforms equation (2d) into equation (7a) the equivalent line in the optimal reference frame where R^2 is at maximum.

$$\hat{y} = \hat{\alpha}_t + \hat{\beta}_t x \quad (7a)$$

Here, $\hat{\beta}_t$ and $\hat{\alpha}_t$ are given by equations (6a-b) respectively and ϕ, the generic rotation angle, is equal to θ, the angle that maximizes R^2.

Applying equation (6a-b) to the fitted line in the optimal reference frame transforms equation (5d) into equation (7b) the equivalent line in the original reference frame.

$$\hat{y} = \hat{\alpha}_{rt} + \hat{\beta}_{rt} x \quad (7a)$$

where

$$\hat{\beta}_{rt} = \frac{-sin(\theta_r)+\hat{\beta}_r cos(\theta_r)}{cos(\theta_r)+\hat{\beta}_r sin(\theta_r)}$$

$$\hat{\alpha}_{rt} = \hat{\alpha}_r \left(cos(\theta_r) + \frac{sin^2(\theta_r)-\hat{\beta}_r cos(\theta_r) sin(\theta_r)}{cos(\theta_r)+\hat{\beta}_r sin(\theta_r)}\right)$$

$$\theta_r = -\theta$$

Characteristics of Coordinate Transformations:

The source data, listed in Table 1, on the next page, are used in the following discussion.

Experimenters don't have explicit knowledge of the sampling errors in their data but, in this case, the errors are included for comparison and test purposes. The errors are independent and normally distributed with zero mean and a constant variance of 25.

Points, straight lines, and the angles between intersecting lines are invariant under rotation, but the slope, intercept, variance, covariance and coefficient of determination are periodic. There are two important points here. First, the angle between the fitted line and the true line, given by equation (8a), is invariant under rotation.

$$tan^{-1}\hat{\beta} - tan^{-1}\beta = tan^{-1}\hat{\beta}_t - tan^{-1}\beta_t \qquad (8a)$$

As we show in Section 4, the estimated slope of $\hat{\beta}_r$ is equal to $\sqrt{R^2}$ when the coefficient of determination is at maximum. Consequently, $\hat{\beta}_r$ is closer to β_t than $\hat{\beta}_t$ and therefor a more accurate estimator of β_t. Conversely $\hat{\beta}_{rt}$ will be a better estimator of β than $\hat{\beta}$. Further more, if $\hat{\beta} - \beta$ is small, we can invoke the small angle approximation to get equation (8b). It is easily shown, using equations (1) and (2a), that $\hat{\beta}_\varepsilon$, in equation (8b), is the estimated slope of the error regressed on x.

$$\hat{\beta} - \hat{\beta}_\varepsilon = \beta \qquad (8b)$$

According to equation (8b) the true slope β can be partitioned into two components; the estimated slopes of y_i on x_i and ε_i on x_i respectively. Any rotated reference frame in which $\hat{\beta}_{r\varepsilon} < \hat{\beta}_\varepsilon$ will necessarily produce a better estimate of β.

Table 1 Example Data

Column 1: X, the independent variable.
Column 2: Y, the dependent variable without error.
Column 3: Independent, randomly generated error.
Column 4: Observed data, sum of elements from columns 2 and 3.

X	Y (No Error)	Error N $(0,5^2)$	Y+ Error Observed
1	3	-5.3245	-2.3245
2	6	-0.762	5.238
3	9	-10.617	-1.6172
4	12	1.8921	13.8921
5	15	-7.1037	7.8963
6	18	5.4398	23.4398
7	21	2.8784	23.8784
8	24	-1.0916	22.9084
9	27	-0.3437	26.6563
10	30	-2.2164	27.7836
11	33	-5.1076	27.8924
12	36	-3.891	32.109
13	39	-5.6252	33.3748
14	42	0.5669	42.5669
15	45	-10.208	34.7916
16	48	-0.7258	47.2742
17	51	-6.9706	44.0294
18	54	2.6189	56.6189
19	57	-4.4368	52.5632
20	60	-9.714	50.286

What is equally important is that the slope, intercept, variance, covariance and coefficient of determination are periodic functions. Figure 1 illustrates the typical behavior of R^2 under rotation. Rotation angles from $-\pi/2$ to $\pi/2$ are plotted on the horizontal axis. The corresponding coefficients of determination, computed using equation (5a) and data from Table 1 are plotted on the vertical axis. First, we observe that the coefficient of determination is periodic, repeating ever $\pi/2$ radians. And second, R^2 has two maximum values on the interval $(-\pi/2, \pi/2)$. Only one is shown here. Recall that θ is the argument of ϕ that maximizes R^2.

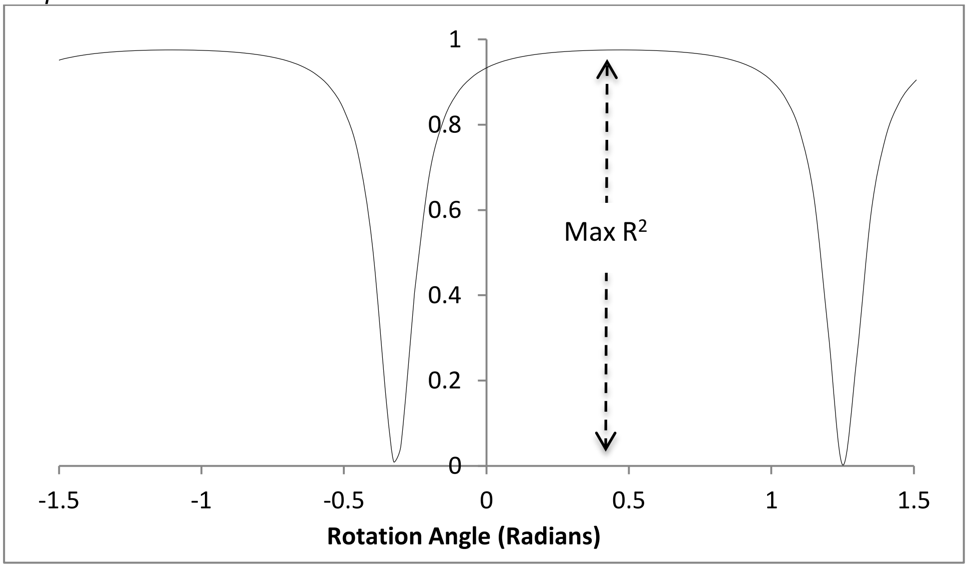

Figure 1 Coefficient of Determination (R^2) as a function of rotation angle (ϕ)

Figure 2 illustrates the effects of rotation on a single data point. Both the primary x-y axis (solid lines) and rotated axis (dashed lines) are shown superimposed over the same origin. The true line $y = \alpha + \beta x$ is also shown (dotted line).

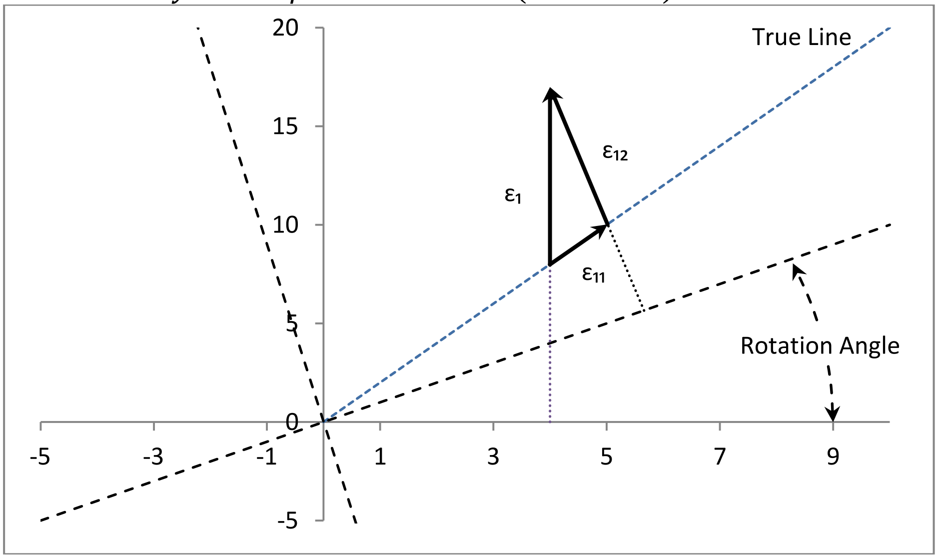

Figure 2 Example of a coordinate transformation partitioning the error

The vector (ε_1), is the error measured in the original (primary) reference frame. Rotating the axis counter clockwise effectively partitions ε_1 into two component vectors, ε_{11} and ε_{12}. The first vector, ε_{11}, lies parallel to the true line and does not contribute to variation in the new reference frame. The second vector, ε_{12}, is the unexplained error and does contribute to variation. As the x-axis approaches the true line, the vector magnitude $|\varepsilon_{12}|$ approaches its minimum value. At this point the unexplained error is minimized, or conversely, R^2 is maximized.

Figure 3 illustrates typical behavior for the mean and variance of the true error (that is the error relative to the true line) before and after rotation. It is fair to say that we don't know the true value of α, β, or the individual ε_i's, but observing their general behavior under rotation is helpful. Two graphs are shown. The solid line shows the sample distribution for the true error in Table 1. The dashed line, which corresponds to R^2 at its maximum, shows the sample distribution for the same data, after rotation. We attribute the collapse of the error about the true line to rotation of the x-y axis as illustrated in Figure 2.

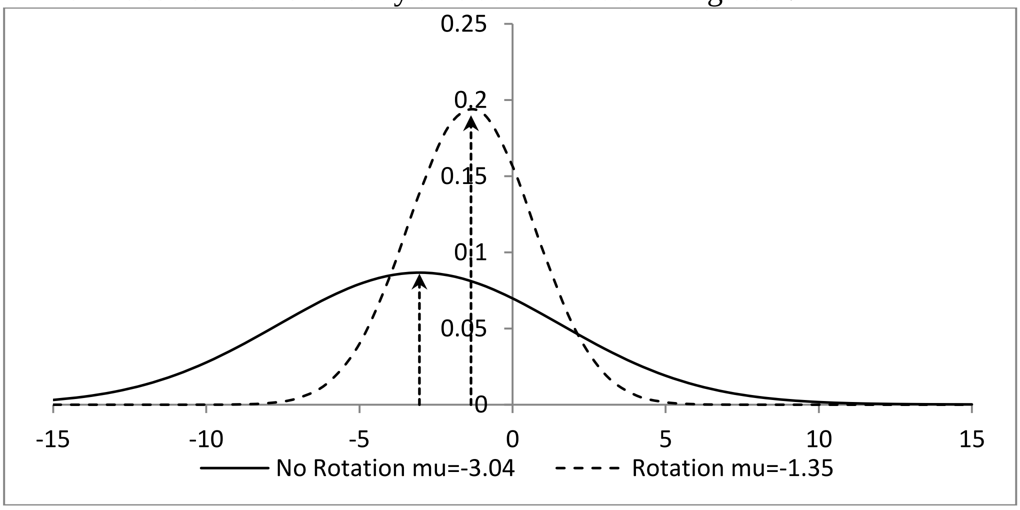

Figure 3 Collapse of the error about the true line due to axis rotation.

We conclude from Figures 1, 2, and 3 that there is an angle θ, on the interval $[-\pi/2, \pi/2]$, that maximizes R^2, which in turn collapses the error about the true line and reduces the difference between the estimated slope $(\hat{\beta}_r)$ and the true slope (β_t). Invoking equation (8b) leads to equation (9), which says that $\hat{\beta}_{rt}$ is a better estimate of β than $\hat{\beta}$.

$$\hat{\beta} - \beta > \hat{\beta}_{rt} - \beta \quad (9)$$

Finding the rotation angle θ

The premise for using this approach rests on finding θ, the argument of ϕ that maximizes R^2 and θ_r the angle that returns $\hat{\beta}_r$ to the original reference frame.

The direct approach is to solve equation (10).

$$\frac{d\left(\frac{Cov(x_r,y_r)^2}{Var(x_r)Var(y_r)}\right)}{d\phi}=0 \quad (10)$$

While technically correct, this method is both difficult to implement and insensitive to small changes in the neighborhood of θ. Figure 4 suggests an easier approach.

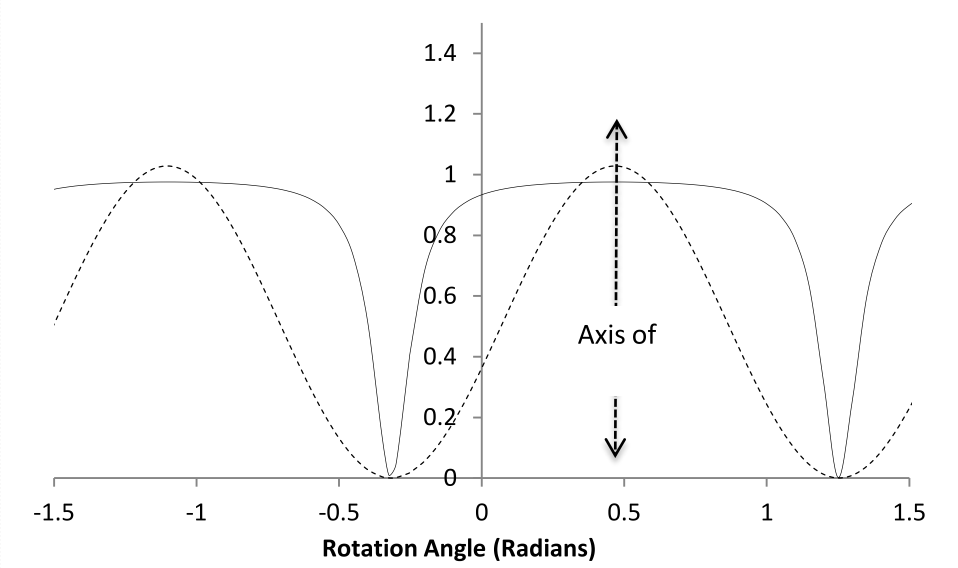

Figure 4 R_r^2 and the $COV(x_r, y_r)^2$ as functions of the rotation angle. Note, the magnitude of the covariance is not shown to scale.

In this approach, the $COV(x_r, y_r)$ is used as a proxy for r_r^2. The coefficient of determination is a function of the covariance squared. Both functions are in phase, have the same period, and share common vertical axis of symmetry that bi-sect their respective maxima. It follows that the value for θ, the angles that maximizes the $COV(x_r, y_r)$ on the interval $[0, \pi/2]$, will also maximizes R^2 and is easily found by solving equation [11a] for ϕ. The general solution is given by equations (11b-c).

$$\frac{d(Cov(x_r,y_r))}{d\phi} = 0 \quad (11a)$$

Expanding and taking the derivative gives:

$$\left(VAR(x_r) - VAR(y_r)\right)tan^2(\phi) - 4COV(x_r, y_r)tan(\phi) + VAR(y_r) - VAR(x_r) \tag{11b}$$

Using the quadratic formula gives:

$$\theta = tan^{-1}\left(\frac{4COV(x_r, y_r) \pm \sqrt{16COV(x_r, y_r)^2 + 4\left(VAR(y_r) - VAR(x_r)\right)^2}}{-2\left(VAR(y_r) - VAR(x_r)\right)}\right) \tag{11c}$$

In truth, equation (11c) gives two values for θ. Use the value that returns a positive covariance when β is equal to or greater than 1, and a negative covariance when β is less than one.

The angle θ has two unique properties in addition to maximizing R^2.

First, the $VAR(x_r)$ equals $VAR(y_r)$ when R^2 is at maximum. It can be shown, with very little manipulation, that equation (11b) is equivalent to equation (12), so they share a common solution.

$$VAR(x_r) - VAR(y_r) = 0 \tag{12}$$

Second, the slope $\hat{\beta}_r$, in the rotated reference frame, is equal to the square root of the coefficient of determination when R^2 is at maximum. This is a direct result of equation (12).

$$R_r^2 = \hat{\beta}_r^2 \frac{VAR(x_r)}{VAR(y_r)} = \hat{\beta}_r^2 \tag{13}$$

Finding the return angle θ_r:

There is an ambiguity in the above solution. The angular difference between $\hat{\beta}_r$ and $\hat{\beta}_t$, the angular error removed from $\hat{\beta}$, is given by equation (13a) and is always positive. Consequently, the return angle must be adjusted to compensate for situations where $\hat{\beta}$ is greater than β. The return angle θ_r is found using equation (13b).

$$\Delta = tan^{-1}(\hat{\beta}_r) - \left(tan^{-1}(\hat{\beta}_t) - \theta\right) \tag{13a}$$

$$\theta_r = \begin{cases} -\theta, \; \hat{\beta} < \beta \\ -(\theta - 2\Delta), \; \hat{\beta} \geq \beta \end{cases} \tag{13b}$$

EXAMPLE – Old Faithful Geyser and the time to its next eruption

Old Faithful geyser in Yellowstone National Park, Wyoming is renowned for both its regular eruptions and the Park Ranger's ability to predict the time of the next event. It is widely accepted that the time-interval between events is linearly dependent on the duration of the preceding eruption.

The Geyser Observation and Study Association (GOSA) is a scientific organization that collects and disseminates information on geysers. The GOSA publishes Table 2, a regression table that has, in the past, been used by park rangers to estimate the time interval between eruptions as a function of the preceding event's duration [6].

Although park rangers have shifted to a bi-modal model, the table is still maintained for historical purposes. The fitted line corresponding to the data in Table 2 is given by equation (14) and is included here as a base line for comparing outputs from the standard model using observed data and the enhanced model using transformed data.

$$Time\ to\ next\ event = 12.375 \cdot Duration\ of\ current\ event + 33.2142\ minutes \tag{14}$$

Table 2 Old Faithfull Eruption Data [5]

Duration min	Interval min
1.5	51
2.0	58
2.5	65
3.0	71
3.5	76
4.0	82
4.5	89
5.0	95

Observed data consists of the 272, randomly selected, paired observation plotted in Figure 5. The horizontal axis shows the duration of each eruption in minutes. The vertical-axis shows the corresponding time-interval to the next event, in minutes. The results of the first regression in the original (primary) reference frame are also shown.

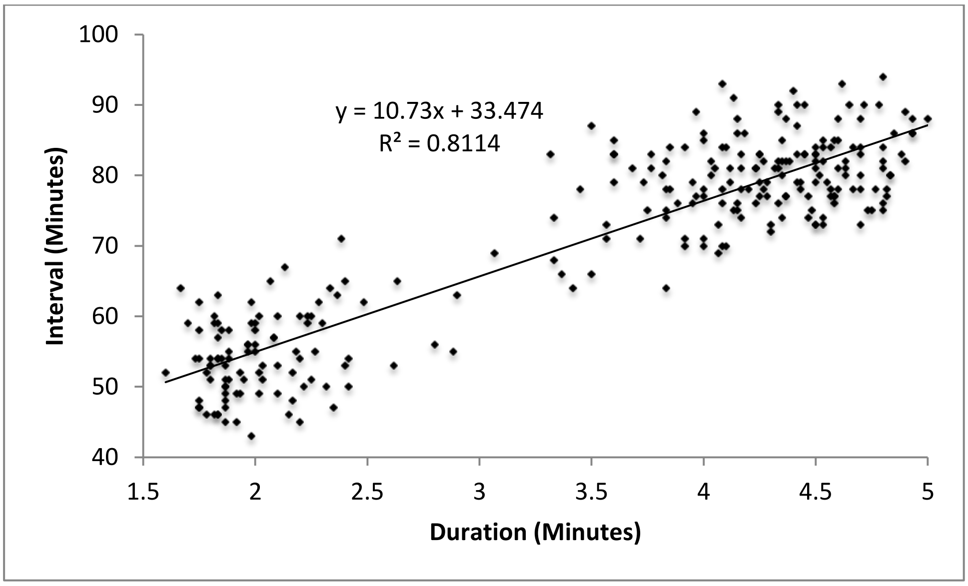

Figure 5 Scatter plot of 272-paired observation of Old Faithfull eruptions in Yellowstone National Park

The data sources are [4] Hardle, W. (1991) Smoothing Techniques with Implementation in S. Oxford University Press, NY and Azzalini, A. and Bowman, A. W. (1990) A look at some data on the Old Faithful Geyser Applied Statistics 39m 357-365.

A rotation angle of .7098 radians was used to transform the data in Figure 5 to the new reference frame shown in Figure 6. Note that R^2 has increased from .8114 in the original reference frame to .9948 in the new reference frame.

The results of the regression, shown in Figure 6, are returned to the original reference frame using an angle of -.7098 radians. This angle is used because $\hat{\beta} = 10.73$ is less than the base line value of 12.357. Figure 7 shows a comparison of the base line and the two regression lines, plotted in the primary reference frame.

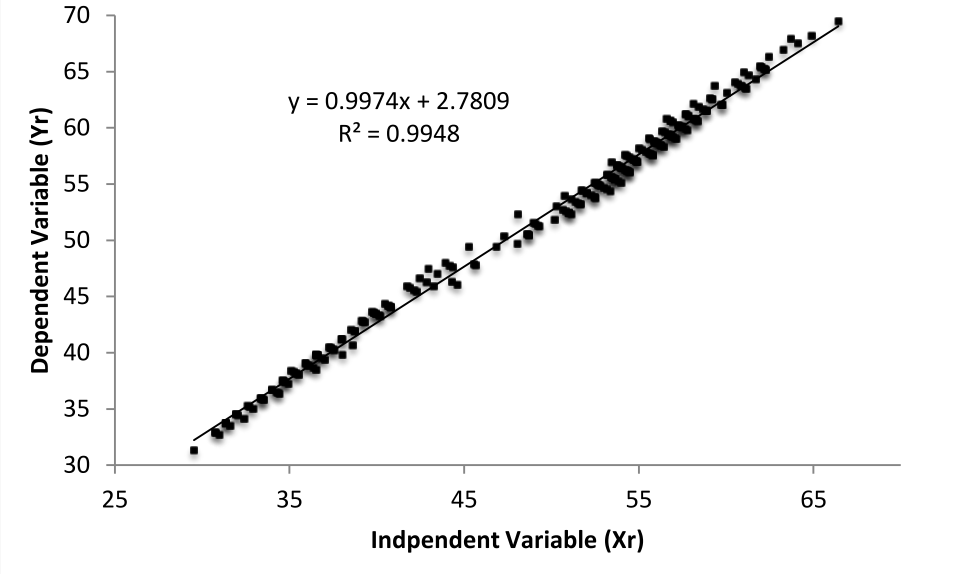

Figure 6 Scatter plot of the 272 paired observation, after rotation.

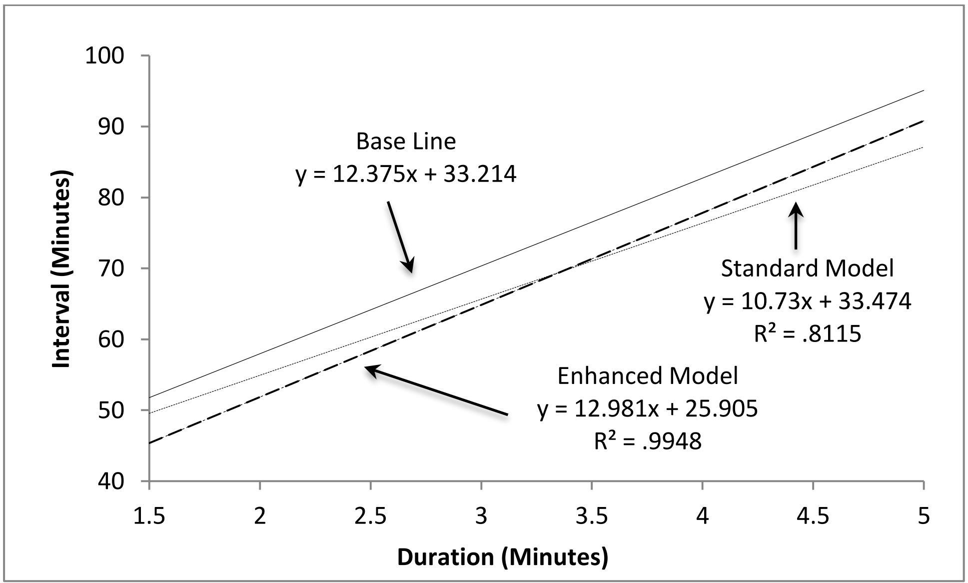

Figure 7 Comparison of the two fitted lines with the base line.

Two observations stand out. First, R^2 increased from 81.2% to 99.5%. Second, the angular error in the slope was reduced by approximately 63%. Overall results are summarized in Table 3.

Table 3 Comparisons of Model Outputs

Parameters	Standard Model	Enhanced Model	Base Line
R^2	.8115	.9948	n/a
$\hat{\beta}$	10.73	12.981	12.375
$\hat{\alpha}$	33.474	25.905	33.214

MODEL STRENGTHS AND WEAKNESSES

There are three issues that should be considered before using this approach.

1. In order to return $\hat{\beta}_r$ and $\hat{\alpha}_r$ to the original reference frame we must determine if $\hat{\beta}$ is greater than or less than β. The model does not provide any information here. The decision to add or subtract the angular error from $\hat{\beta}$ is, for the most part, made on theoretical or empirical grounds.

2. Maximizing R^2 will reduce the variance in the data and increase R^2, but that does not necessarily mean that it will return the best estimate of the slope. Rotation angles that result in R^2 equal to zero may return slightly better estimates.

3. The slopes $\hat{\beta}_{rt}$ and β are approximately equal, but their respective y-intercepts will necessarily differ by a constant. For this reason, some consideration should be given to using the point-slope method with a well-defined boundary condition to estimate alpha.

Rotation transformations are easy to use and, when combined with linear regression, provide a practical way to squeeze more information from limited or poor data than could otherwise be retrieved using simple regression alone. Furthermore, the enhanced model described here lends itself to computer applications, which further enhance its utility.

REFERENCES

[1] Boas M. L. (2006). *Mathematical Methods in the Physical Sciences (3rd ed.).* Hoboken, NJ: John Wiley & Sons, Inc.
[2] Goodman A. (1963). *Analytic Geometry and the Calculus.* New York: Macmillan Company
[3] Larson J. R. & Marx M. L. (2012). *An Introduction to MATHEMATICAL STATISTICS and Its Applications (5th ed.).* Boston, MA: Pearson Education, Inc.
[4] Montgomery, D. C., Peck E. A. & Vining G. G. (2006). *Introduction to Linear Regression Analysis (4th ed.).* Hoboken, NJ: John Wiley & Sons, Inc.
[5] Retrieved April 18 2016 from http://www.stat.cmu.edu/~larry/all-of-statistics/=data/faithful.dat
[6] Retrieved April 18 2016 from http://www.geyserstudy.org/geyser.aspzx?PGeyserNo=OLDFITHULL

APPENDIX A Rotation Transformations

The basic rotation matrix and its inverse are given by equations (A1a) and (A1b) respectively [1] [2] [4].

$$\begin{vmatrix} x' \\ y' \end{vmatrix} = \begin{vmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{vmatrix} \cdot \begin{vmatrix} x \\ y \end{vmatrix} \tag{A1a}$$

where

$$x' = xcos(\theta) + ysin(\theta)$$
$$y' = -xsin(\theta) + ycos(\theta)$$

and

$$\begin{vmatrix} x \\ y \end{vmatrix} = \begin{vmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{vmatrix} \cdot \begin{vmatrix} x' \\ y' \end{vmatrix}$$

(A1b)

where

$$x = x' \cos(\theta) - y'sin(\theta)$$
$$y = x'sin(\theta) + y'cos(\theta)$$

Negative rotation angles change the direction of rotation [2]. To demonstrate this point, we use a negative angle $\theta_r = -\theta$ with equation (A1a):

$$\begin{vmatrix} cos(\theta_r) & sin(\theta_r) \\ -sin(\theta_r) & cos(\theta_r) \end{vmatrix} = \begin{vmatrix} cos(-\theta) & sin(-\theta) \\ -sin(-\theta) & cos(-\theta) \end{vmatrix} = \begin{vmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{vmatrix}$$

Note, using $\theta_r = -\theta$ is equation (A1a) returns the inverse transformation. For simplicity and ease of programing, we use the default rotation matrix (A1a) to transform fitted lines in the original or optimal reference frames to equivalent equations in the optimal or original reference frames respectively and let the sign of the rotation angle dictate the direction of rotation. Using this methodology, equation (A3a) transforms points in the original reference frame into their equivalent values in the optimal reference frame and equation (A3b) reverses the transformation.

(A3a)

$$x_{ir} = x_i cos(\theta) + y_i sin(\theta)$$

$$y_{ir} = -x_i sin(\theta) + y_i cos(\theta)$$

(A3b)

$$x_i = x_{ir} cos(\theta_r) + y_{ir} sin(\theta_r)$$

$$y_i = -x_{ir} sin(\theta_r) + y_{ir} cos(\theta_r)$$
$$\theta_r = -\theta$$

APPENDIX B Returning fitted lines to the original reference frame

x_r and $\hat{y}_r$ are variables in the rotated reference frame and are related by:

$$\hat{y}_r = \hat{\alpha}_r + \hat{\beta}_r x \qquad \text{(B1)}$$

x and $\hat{y}$ are variables in the original reference frame and are related by:

$$\hat{y} = \hat{\alpha}_{rt} + \hat{\beta}_{rt} x \qquad \text{(B2)}$$

The problem is to compute $\hat{\beta}_{rt}$ and $\hat{\alpha}_{rt}$ as functions of $\hat{\beta}_r$ and $\hat{\alpha}_r$.

The following transformation, in equation form, is used

$$x = x_r cos(\theta_r) + \hat{y}_r sin(\theta_r) \qquad \text{(B3a)}$$

$$\hat{y} = -x_r sin(\theta_r) + \hat{y}_r cos(\theta_r) \qquad \text{(B3b)}$$

where $\theta_r = -\theta$

Substituting equation (B1) for $\hat{y}_r$ in equation (B3a) and solving for x_r gives

$$x_r = \frac{x - \hat{\alpha}_r sin(\theta_r)}{cos(\theta_r) + \hat{\beta}_r sin(\theta_r)} \qquad \text{(B3c)}$$

Then substituting x_r into equation (B3b) gives

$$\hat{y} = \frac{x - \hat{\alpha}_r sin(\theta_r)}{cos(\theta_r) + \hat{\beta}_r sin(\theta_r)}\big(-sin(\theta_r)\big) + \hat{\beta}_r cos(\theta_r)\frac{x - \hat{\alpha}_r sin(\theta_r)}{cos(\theta_r) + \hat{\beta}_r sin(\theta_r)} + \hat{\alpha}_r cos(\theta_r)$$

Rearranging terms gives

$$\hat{y} = \hat{\alpha}_r\left(cos(\theta_r) + \frac{sin^2(\theta_r) - \hat{\beta}_r cos(\theta_r) sin(\theta_r)}{cos(\theta_r) + \hat{\beta}_r sin(\theta_r)}\right) + \left(\frac{-sin(\theta_r) + \hat{\beta}_r cos(\theta_r)}{cos(\theta_r) + \hat{\beta}_r sin(\theta_r)}\right)x \qquad \text{(B3d)}$$

where

$$\hat{\alpha}_{rt} = \hat{\alpha}_r\left(cos(\theta_r) + \frac{sin^2(\theta_r) - \hat{\beta}_r cos(\theta_r) sin(\theta_r)}{cos(\theta_r) + \hat{\beta}_r sin(\theta_r)}\right)$$

$$\hat{\beta}_{rt} = \frac{-sin(\theta_r) + \hat{\beta}_r cos(\theta_r)}{cos(\theta_r) + \hat{\beta}_r sin(\theta_r)}$$