Fitting a line of best fit to correlated data of varying precision

R.E. Deakin[1] and M.N. Hunter[2]
[1]Bonbeach VIC, 3196, Australia; [2]Maribyrnong, VIC, 3032, Australia
email: randm.deakin@gmail.com

September 2017

Abstract

Fitting a straight line through (X,Y) data is a common problem in estimation. Using a data plot and a ruler, the problem is solved by slowly moving the ruler to a position that visually minimizes the perpendicular distances between the data points and the ruler. A mathematical solution can be determined using the theory of *least squares* that supposes the most probable 'answer' is one that minimizes the sum of the weighted squares of residuals, where residuals are small corrections to the (X,Y) data and weights are numbers reflecting precision of measurements.

York (1966, 1968) solved this problem for uncorrelated and correlated data, but unfortunately his solution is not well known and scientists and engineers often use inappropriate methods embedded in software products and calculators. This paper will show, in detail, how York solved this problem of estimation.

Introduction

Fitting a straight line of best fit $y = a + bx$ through (X,Y) data is a common problem in estimation. It is also known as Linear Regression and its first use in what would now be called statistics is attributed to the English scientist Sir Francis Galton (1822–1911). In its simplest form the X-values are considered error-free and the Y-values are measurements (subject to error) with equal precision. The solution for the parameters a (intercept on y-axis) and $b = \tan\theta$ (gradient of line) is a relatively simple application of the theory of least squares and can be found as a special function on many scientific calculators.

When the (X,Y) data are both considered as measurements subject to error then the problem is more complicated. The earliest published work on this topic is by R.J. Adcock, an attorney from Monmouth (Illinois, USA) who in three papers in *The Analyst*[1] (Adcock, 1877, 1878a, 1878b) discussed a method of fitting a straight line to n points such that the sum of the squares of the normals is a minimum. Adcock (1878b) gave the function to be minimized and a worked example but unfortunately there were some errors (in theory and calculation) although he did correctly state that the line passes through the centroid of the n points (Finney 1996, Farebrother 1999). Adcock's errors were corrected by Kummell[2] (1879) who extended Adcock's least squares function to incorporate varying precision of (X,Y) data.

Karl Pearson (1901) adopted the approach of Adcock and Kummell of minimizing the sum of squares of the normals from n points to a line or a plane and gave proofs that the best fitting lines or planes pass through the centroid[3] of a system. In the case of n points in a plane Pearson called the line of best fit the *major axis of the correlation ellipsoid* and he provided a worked example of a line of best fit through $n = 10$ points of supposed equal weight which has become a standard test set (see Example 1).

Deming (1943) proposed that the line of best fit is obtained by minimizing a function S = sum of weighted squared residuals, subject to the condition that adjusted points lie on the line $y = a + bx$ – residuals are small (unknown) corrections to measured values such that *measurement* + *residual* = *'true' value* – and

---

[1] First published in Des Moines, Iowa in 1874 but ceased in 1883. It was continued from 1884 as the *Annals of Mathematics.* An earlier journal of the same name was published by Dr Robert Adrian (Pennsylvania USA) in 1808 but ceased after publication of a single volume.
[2] Charles H. Kummell was a surveyor/geodesist with the U.S. Lake Survey, Detroit, Michigan.
[3] Pearson's centroid is defined as the point having mean values of the coordinates. In this paper 'centroid' is taken to mean a point having weighted mean values of the coordinates

Deming used a weight W_k that is a function of the (unknown) gradient b and the weights $w(X_k), w(Y_k)$ of the measured values (X_k, Y_k). In an example, using a simplified method of solution, Deming noted that the solution for b was iterative, since b was contained in W_k. Also, in a further example, where the ratio of weights $w(X_k)/w(Y_k)$ equalled a constant, he obtained the same result for b as Kummell (1879) and Pearson (1901).

York (1966), using the same approach as Deming, (but with a more rigorous analysis) arrives at his *Least Squares Cubic:* $b^3 - 3Ab^2 + 3Bb - C = 0$. The solution of York's cubic can present difficulties because of the implicit dependence of b (via the weight function W_k and the centroidal coordinates U_k, V_k) in the coefficients A, B and C, and it is possible that the roots of the cubic may be complex (Reed 1989). McIntyre et al. (1966), using a slightly different least squares technique, also arrive at York's least squares cubic, although there is no mention of York's earlier paper. Both of these independent determinations of the least squares cubic assume errors in both coordinates but no correlations between the errors.

Williamson (1968) derives York's cubic for the uncorrelated errors case, but noting the implicit dependence of b in the coefficients A, B and C, states that York's choice of a cubic in b is just one of many polynomial representations. Williamson then gives a linear equation in the gradient b.

York (1968), referring to his and McIntyre's et al. earlier papers derives the *Generalized Least Squares Cubic* for the case of correlated errors in the measured values (X_k, Y_k) and following Williamson (1968) reduces the cubic to a linear equation in b.

Since the work of York (1966, 1968), McIntyre et al (1966) and Williamson (1968) there have been many papers on the topic of fitting straight lines to data, some offering variations on the least squares approach, e.g. Krane and Schecter (1981), Lybanon (1984), Neri et al. (1989) and Kiryati and Bruckstein (1992) and others offering comparisons of various published methods, e.g. Macdonald and Thompson (1992), Duer et al. (2008) and Cantrell (2008). York et al. (2004) summarizes the earlier work of York (1966, 1968) and provides a compact set of equations for the determination of the parameters b and a (gradient and intercept), as well as estimates of variances of the parameters that are a simplification of those given in York (1968).

In this paper we have set out, in detail, the methods used to obtain the *Generalized Least Squares Cubic* in b and the linear equation in b given in York (1968) and equations for the computation of residuals and estimates of variances of b and the y-intercept a. Worked examples are provided that cover the general solution as well as special cases. Two appendices are contained in this paper; the first shows an alternative approach to minimisation of a 'least squares' function that leads to the least squares cubic; and the second shows a very detailed derivation of the necessary equations for variance estimation of the parameters a and b. The detailed analysis leading to these equations is different from York (1968) and York et al. (2004) but confirms those published results.

Nomenclature

In this paper we have adopted the notation of York et al. (2004). Also, there is much use of the summation symbol $\sum$. For convenience and clarity $\sum_k A_k$, $\sum A_k$ and $\sum A$ all mean $\sum_{k=1}^{n} A_k$

Symbol	Meaning	Definition
a, b	y-intercept and gradient of line of best fit $y = a + bx$	
$b_0, b_1, b_2, \ldots, b_n, b_{n+1}$	initial and successive values of the gradient b in an iterative solution	
j, k	integer counters	

Symbol	Meaning	Definition
n	integer; number of data points; iteration number	
$\mathbf{Q}_k$	variance matrix of measurements	$\mathbf{Q}_k = \begin{bmatrix} s^2(X_k) & s(X_kY_k) \\ s(X_kY_k) & s^2(Y_k) \end{bmatrix}$
r_k	correlation	$r = \dfrac{s(X_kY_k)}{s(X_k)s(Y_k)}$
S	least squares function	$S = \mathbf{v}_1^T\mathbf{W}_1\mathbf{v}_1 + \mathbf{v}_2^T\mathbf{W}_2\mathbf{v}_2 + \cdots$
$s^2(X_k), s^2(Y_k)$	unbiased estimates of variances	
$s(X_k), s(Y_k)$	unbiased estimates of standard deviations	
$s(X_kY_k)$	covariance	
(U_k, V_k)	centroidal coordinates of P_k	$U_k = X_k - \bar{X},\ \ V_k = Y_k - \bar{Y}$
$\mathbf{v}_k^T$	transpose of vector of residuals	$\mathbf{v}_k^T = \begin{bmatrix} v(X_k) & v(Y_k) \end{bmatrix}$
$v(X_k), v(Y_k)$	residuals (corrections to measurements)	$X_k + v(X_k) = x_k,\ \ Y_k + v(Y_k) = y_k$
$\mathbf{W}_k$	weight matrix	$\mathbf{W}_k = \mathbf{Q}_k^{-1}$
$w(X_k), w(Y_k)$	weights of coordinates (X_k, Y_k)	$w(X_k) = \dfrac{1}{s^2(X_k)},\ \ w(Y_k) = \dfrac{1}{s^2(Y_k)}$
W_k	weight function	$W_k = \dfrac{\alpha_k^2}{b^2 w(Y_k) + w(X_k) - 2br_k\alpha_k}$
(x_k, y_k)	adjusted coordinates of point P_k	
(X_k, Y_k)	measured coordinates of point P_k	
$(\bar{X}, \bar{Y})$	centroid	$\bar{X} = \dfrac{\sum W_k X_k}{\sum W_k},\ \ \bar{Y} = \dfrac{\sum W_k Y_k}{\sum W_k}$
α_k		$\alpha_k = \sqrt{w(X_k)w(Y_k)} = \dfrac{1}{s(X_k)s(Y_k)}$
β_k		$\beta_k = W_k\left(\dfrac{U_k}{w(Y_k)} + \dfrac{bV_k}{w(X_k)} - \dfrac{r_k}{\alpha_k}(bU_k + V_k)\right)$
$\bar{\beta}$		$\bar{\beta} = \dfrac{\sum_k W_k\beta_k}{\sum_k W_k}$
λ_k	*Lagrange multiplier*	$\lambda_k = W_k(a + bX_k - Y_k)$

Note: For the special case $b = 0$, the weight function $W_k = w(Y_k)$ and special results follow for the centroid $(\bar{X}, \bar{Y})$, centroidal coordinates (U_k, V_k) and the factors β_k and $\bar{\beta}$.

The equation of a straight line

The equation of a straight line can be expressed as

$$y = a + bx \tag{1}$$

where $b = \tan\theta$ is the gradient of the line and a is the intercept of the line on the y-axis.

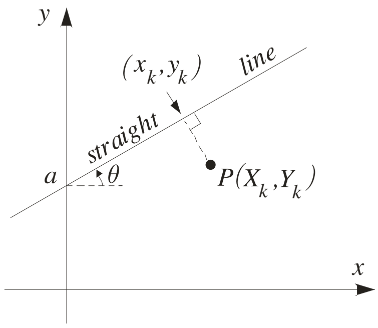

Figure 1. The orthogonal projection of $P\left(X_k, Y_k\right)$ onto the straight line $y = a + bx$

In Figure 1 P_k is one of a number of points $k = 1, \ldots, n$ that lie on or near the straight line. $\left(X_k, Y_k\right)$ are measurements and x_k, y_k are the coordinates of P_k orthogonally projected onto the straight line. We call $\left(x_k, y_k\right)$ the adjusted coordinates of P_k.

Measurements, residuals, weights, variance, covariance and correlation

Defining residuals $v\left(X_k\right), v\left(Y_k\right)$ as small unknown corrections to measurements X_k, Y_k we write

$$\begin{aligned} X_k + v\left(X_k\right) &= x_k \\ Y_k + v\left(Y_k\right) &= y_k \end{aligned} \quad \text{or} \quad \begin{aligned} v\left(X_k\right) &= x_k - X_k \\ v\left(Y_k\right) &= y_k - Y_k \end{aligned} \tag{2}$$

and the vector of residuals $\mathbf{v}$ is

$$\mathbf{v}^T = \left[v\left(X_1\right) \quad v\left(Y_1\right) \,\middle|\, v\left(X_2\right) \quad v\left(Y_2\right) \,\middle|\, \cdots \quad \cdots \,\middle|\, v\left(X_n\right) \quad v\left(Y_n\right) \right] = \left[\mathbf{v}_1^T \quad \mathbf{v}_2^T \quad \cdots \quad \mathbf{v}_n^T \right] \tag{3}$$

where the superscript T denotes transpose and $\mathbf{v}_k^T = \left[v\left(X_k\right) \quad v\left(Y_k\right) \right]$

If measurements $\left(X_k, Y_k\right)$ have associated estimates of variances $s^2\left(X_k\right), s^2\left(Y_k\right)$ and covariances $s\left(X_k Y_k\right)$ then we define the block-diagonal variance matrix of measurements $\mathbf{Q}$ as

$$\mathbf{Q} = \left[\begin{array}{cc:cc:cc:cc} s^2(X_1) & s(X_1Y_1) & 0 & 0 & \cdots & \cdots & 0 & 0 \\ s(X_1Y_1) & s^2(Y_1) & 0 & 0 & \cdots & \cdots & 0 & 0 \\ \hdashline 0 & 0 & s^2(X_2) & s(X_2Y_2) & 0 & \cdots & 0 & 0 \\ 0 & 0 & s(X_2Y_2) & s^2(Y_2) & 0 & \cdots & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & \ddots & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \ddots & & \vdots \\ \hdashline 0 & 0 & 0 & 0 & \cdots & \cdots & s^2(X_n) & s(X_nY_n) \\ 0 & 0 & 0 & 0 & \cdots & \cdots & s(X_nY_n) & s^2(Y_n) \end{array}\right] = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Q}_n \end{bmatrix} \tag{4}$$

where the 2-by-2 sub-matrices $\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_n$ are each symmetric. Here variances are measures of the dispersion from mean values and covariances are measures of independence. If two random quantities are independent then their covariance will be zero. For a finite population of N quantities $X_1, X_2, \ldots, X_N$ the population mean and variance are $\mu = \frac{1}{N}\sum_{j=1}^{N} X_j$ and $\sigma^2 = \frac{1}{N}\sum_{j=1}^{N}\left(X_j - \mu\right)^2$ respectively. For a sample of n quantities from an infinite population the unbiased estimates of the mean and variances are $\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$ and $s^2(X) = \frac{1}{n-1}\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2$ respectively and the standard deviation $s(X)$ is defined to be the positive square-root of the variance. Also, for samples (X_j, Y_j) each of size n an unbiased estimate of the covariance is $s(XY) = \frac{1}{n-1}\sum_{j=1}^{n}\left(X_j - \bar{X}\right)\left(Y_j - \bar{Y}\right)$.

The block-diagonal weight matrix of the measurements denoted by $\mathbf{W}$ is defined to be

$$\mathbf{W} = \mathbf{Q}^{-1} = \begin{bmatrix} \mathbf{W}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{W}_n \end{bmatrix} \tag{5}$$

where the 2-by-2 sub-matrices $\mathbf{W}_1 = \mathbf{Q}_1^{-1}, \mathbf{W}_2 = \mathbf{Q}_2^{-1}, \ldots, \mathbf{W}_n = \mathbf{Q}_n^{-1}$ are symmetric.

The variance matrix of the measurements at P_k is $\mathbf{Q}_k = \begin{bmatrix} s^2(X_k) & s(X_kY_k) \\ s(X_kY_k) & s^2(Y_k) \end{bmatrix}$ and the weight matrix at P_k is

$$\mathbf{W}_k = \mathbf{Q}_k^{-1} = \frac{1}{s^2(X_k)s^2(Y_k) - s^2(X_kY_k)}\begin{bmatrix} s^2(Y_k) & -s(X_kY_k) \\ -s(X_kY_k) & s^2(X_k) \end{bmatrix} \tag{6}$$

Note here we have used a standard result for a 2-by-2 matrix inverse where if $\mathbf{N} = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$ then

$\mathbf{N}^{-1} = \dfrac{1}{N_{11}N_{22} - N_{12}N_{21}} \begin{bmatrix} N_{22} & -N_{12} \\ -N_{21} & N_{11} \end{bmatrix}$

The correlation r_k between the measurements X_k, Y_k is defined as

$$r_k = \frac{s(X_k Y_k)}{s(X_k)s(Y_k)} \tag{7}$$

then

$$\frac{1}{s^2(X_k)s^2(Y_k) - s^2(X_k Y_k)} = \frac{1}{(1 - r_k^2)s^2(X_k)s^2(Y_k)} \tag{8}$$

Substituting (8) into (6) gives

$$\mathbf{W}_k = \frac{1}{1 - r_k^2} \left[\begin{array}{c|c} \dfrac{1}{s^2(X_k)} & -\dfrac{r_k}{s(X_k)s(Y_k)} \\ \hline -\dfrac{r_k}{s(X_k)s(Y_k)} & \dfrac{1}{s^2(Y_k)} \end{array} \right] \tag{9}$$

Defining weights as the inverse of the unbiased estimates of variances, then

$$w(X_k) = \frac{1}{s^2(X_k)} \quad \text{and} \quad w(Y_k) = \frac{1}{s^2(Y_k)} \tag{10}$$

And defining

$$\alpha_k = \sqrt{w(X_k)w(Y_k)} = \frac{1}{s(X_k)s(Y_k)} \tag{11}$$

the weight matrix at P_k can be written as

$$\mathbf{W}_k = \frac{1}{1 - r_k^2} \left[\begin{array}{c|c} w(X_k) & -r_k\,\alpha_k \\ \hline -r_k\,\alpha_k & w(Y_k) \end{array} \right] \tag{12}$$

The least squares function

The *least squares function* S = the sum of the weighted squares of residuals. S is a scalar quantity and can be defined as a matrix product

$$S = \mathbf{v}^T \mathbf{W} \mathbf{v} \tag{13}$$

With (3) and (5) we may write S as

$$S = \mathbf{v}_1^T \mathbf{W}_1 \mathbf{v}_1 + \mathbf{v}_2^T \mathbf{W}_2 \mathbf{v}_2 + \mathbf{v}_3^T \mathbf{W}_3 \mathbf{v}_3 + \cdots + \mathbf{v}_n^T \mathbf{W}_n \mathbf{v}_n$$

and with (12)

$$\begin{aligned}\mathbf{v}_k^T \mathbf{W}_k \mathbf{v}_k &= \frac{1}{1-r_k^2}\begin{bmatrix} v(X_k) & v(Y_k)\end{bmatrix}\begin{bmatrix} w(X_k) & -r_k\,\alpha_k \\ -r_k\,\alpha_k & w(Y_k)\end{bmatrix}\begin{bmatrix} v(X_k) \\ v(Y_k)\end{bmatrix} \\ &= \frac{1}{1-r_k^2}\left\{ w(X_k)\big(v(X_k)\big)^2 - 2r_k\,\alpha_k\, v(X_k)\, v(Y_k) + w(Y_k)\big(v(Y_k)\big)^2 \right\} \\ &= \frac{1}{1-r_k^2}\left\{ w(X_k)\big(x_k - X_k\big)^2 - 2r_k\,\alpha_k\big(x_k - X_k\big)\big(y_k - Y_k\big) + w(Y_k)\big(y_k - Y_k\big)^2 \right\}\end{aligned} \tag{14}$$

we have a least squares function S as (York, 1968, eq. 1 with $\alpha_k = \sqrt{w(X_k)\,w(Y_k)}$)

$$S = \sum_k \frac{1}{1-r_k^2}\left\{ w(X_k)\big(x_k - X_k\big)^2 - 2r_k\,\alpha_k\big(x_k - X_k\big)\big(y_k - Y_k\big) + w(Y_k)\big(y_k - Y_k\big)^2 \right\} \tag{15}$$

Note here that the variables in S are the adjusted measurements (x_k, y_k) that lie on the line $y = a + b$.

If the measurements (X_k, Y_k) are assumed to be independent quantities (and therefore uncorrelated) then $r_k = 0$ and (15) becomes (York, 1966, eq. 7)

$$S = \sum_k \left\{ w(X_k)\big(x_k - X_k\big)^2 + w(Y_k)\big(y_k - Y_k\big)^2 \right\} \tag{16}$$

Deming (1943) proposed that the "best" straight line is found by minimizing this S.

The line of best fit

Following York (1966, 1968), the line of best fit $y = a + bx$ is that which minimizes the sum of the weighted squares of residuals. Since we require that the adjusted values (x_k, y_k) of the measurements (X_k, Y_k) lie on the straight line then the parameters a and b can be found by minimizing the expression for S in (15) subject to the constraints that

$$y_k = a + bx_k \quad \text{for} \quad k = 1, \ldots, n \tag{17}$$

To achieve this, York uses a mathematical optimization technique known as the method of *Lagrange multipliers* (Lagrange 1788, Vol. 1, Sect IV) where the function to be minimized is the *Lagrangian* L

$$L = S + 2\sum_k \lambda_k \big(a + bx_k - y_k\big) \tag{18}$$

where λ_k are the (as yet unknown) *Lagrange multipliers.* The 2 is a numerical convenience and it should be noted that $a + bx_k - y_k = 0$ for $k = 1, \ldots, n$.

The variables in L are x_k, y_k, a, b and λ_k and it will have an optimum (minimum or maximum value) when the partial derivatives of L equal zero, i.e.,

$$L = \text{optimum} \quad \text{when} \quad \frac{\partial L}{\partial x_k} = \frac{\partial L}{\partial y_k} = \frac{\partial L}{\partial a} = \frac{\partial L}{\partial b} = \frac{\partial L}{\partial \lambda_k} = 0 \tag{19}$$

The partial derivatives of L equated to zero are

$$
\begin{aligned}
\frac{\partial L}{\partial x_k} &= \sum_k \frac{1}{1-r_k^2}\Big\{2w\big(X_k\big)\big(x_k - X_k\big) - 2r_k\,\alpha_k\,\big(y_k - Y_k\big)\Big\} + 2\sum_k \lambda_k b = 0 \\
\frac{\partial L}{\partial y_k} &= \sum_k \frac{1}{1-r_k^2}\Big\{2w\big(Y_k\big)\big(y_k - Y_k\big) - 2r_k\,\alpha_k\,\big(x_k - X_k\big)\Big\} - 2\sum_k \lambda_k = 0 \\
\frac{\partial L}{\partial a} &= 2\sum_k \lambda_k = 0 \\
\frac{\partial L}{\partial b} &= 2\sum_k \lambda_k x_k = 0 \\
\frac{\partial L}{\partial \lambda_k} &= 2\sum_k \big(a + bx_k - y_k\big) = 0
\end{aligned}
\tag{20}
$$

The first four members of (20) are satisfied when

$$\Big\{w\big(X_k\big)\big(x_k - X_k\big) - r_k\,\alpha_k\,\big(y_k - Y_k\big)\Big\} + \lambda_k\Big(1 - r_k^2\Big)b = 0 \quad \text{for each } k \tag{21}$$

$$\Big\{w\big(Y_k\big)\big(y_k - Y_k\big) - r_k\,\alpha_k\,\big(x_k - X_k\big)\Big\} - \lambda_k\Big(1 - r_k^2\Big) = 0 \quad \text{for each } k \tag{22}$$

$$\sum_k \lambda_k = 0 \tag{23}$$

$$\sum_k \lambda_k x_k = 0 \tag{24}$$

The last member of (20) satisfies the condition that $y_k = a + bx_k$

Equations (21) and (22) can be solved for $x_k - X_k$ and $y_k - Y_k$ yielding

$$
\begin{aligned}
x_k - X_k &= \frac{1}{\alpha_k^2}\Big(-b\,w\big(Y_k\big)\lambda_k + r_k\,\alpha_k\,\lambda_k\Big) \\
y_k - Y_k &= \frac{1}{\alpha_k^2}\Big(-b\,r_k\,\alpha_k\,\lambda_k + w\big(X_k\big)\lambda_k\Big)
\end{aligned}
$$

from which we obtain

$$x_k = X_k + \frac{\lambda_k}{\alpha_k^2}\Big(r_k\,\alpha_k - b\,w\big(Y_k\big)\Big) \tag{25}$$

$$y_k = Y_k + \frac{\lambda_k}{\alpha_k^2}\Big(w\big(X_k\big) - b\,r_k\,\alpha_k\Big) \tag{26}$$

Substituting (25) and (26) into (17) gives

$$Y_k + \lambda_k\left(\frac{w\big(X_k\big)}{\alpha_k^2} - \frac{b\,r_k}{\alpha_k}\right) = a + b\left\{X_k + \lambda_k\left(\frac{r_k}{\alpha_k} - \frac{bw\big(Y_k\big)}{\alpha_k^2}\right)\right\} = a + bX_k + \lambda_k\left(\frac{b\,r_k}{\alpha_k} - \frac{b^2 w\big(Y_k\big)}{\alpha_k^2}\right)$$

and after some algebra we obtain

$$\lambda_k\left(\frac{w\big(X_k\big) - 2b\,r_k\,\alpha_k + b^2 w\big(Y_k\big)}{\alpha_k^2}\right) = a + bX_k - Y_k$$

giving (York 1968)

$$\lambda_k = W_k \left(a + bX_k - Y_k \right) \tag{27}$$

where the weight function W_k is defined as

$$W_k = \frac{\alpha_k^2}{b^2 w\left(Y_k\right) + w\left(X_k\right) - 2b\, r_k\, \alpha_k} \tag{28}$$

Equation (27) is a solution for the *Lagrange multipliers* λ_k and this result can be used in (23) and (24) to give

$$\sum W_k \left(a + bX_k - Y_k \right) = 0 \tag{29}$$

and

$$\sum x_k W_k \left(a + bX_k - Y_k \right) = 0 \tag{30}$$

Expanding (29) gives

$$a\sum W_k + b\sum W_k X_k - \sum W_k Y_k = 0$$

from which we obtain

$$a = \frac{\sum W_k Y_k}{\sum W_k} - b\frac{\sum W_k X_k}{\sum W_k}$$

Defining the *centroid* $\left(\bar{X},\bar{Y}\right)$ as

$$\bar{X} = \frac{\sum W_k X_k}{\sum W_k} \quad , \quad \bar{Y} = \frac{\sum W_k Y_k}{\sum W_k} \tag{31}$$

we obtain

$$a = \bar{Y} - b\bar{X} \quad \text{or} \quad \bar{Y} = a + b\bar{X} \tag{32}$$

Also we define *centroidal coordinates* $\left(U_k, V_k\right)$ as

$$U_k = X_k - \bar{X}, \quad V_k = Y_k - \bar{Y} \tag{33}$$

Deming (1943, Remark 2, Example 4, p. 181) calls $\left(\bar{X},\bar{Y}\right)$ a *quasi center,* although in his example the data are not correlated, i.e.: $r_k = 0$; $W_k = \dfrac{\alpha_k^2}{b^2 w\left(Y_k\right) + w\left(X_k\right)}$; and $\dfrac{1}{W_k} = \dfrac{b^2}{w\left(X_k\right)} + \dfrac{1}{w\left(Y_k\right)}$

Using (33) and (32) we may write

$$\begin{aligned} a + bX_k - Y_k &= a + b\left(U_k + \bar{X}\right) - \left(V_k + \bar{Y}\right) \\ &= a + bU_k + b\bar{X} - V_k - \bar{Y} \\ &= bU_k - V_k \end{aligned} \tag{34}$$

and using this result (27) and (29) become

$$\lambda_k = W_k\left(bU_k - V_k\right) \tag{35}$$

$$\sum W_k\left(bU_k - V_k\right) = 0 \tag{36}$$

[Note that (36) could be expressed as $b\sum WU = \sum WV$ implying a solution for b. But this leads to a division by zero, since by definition of U and V, $\sum WU = \sum WV = 0$. Thus (36) is only used as a means of simplifying the left-hand-side of (30) as shown below.]

Now, using (25), (34) (35) and (36) in the left-hand-side of (30) gives

$$\begin{aligned}\sum x_k W_k\left(a + bX_k - Y_k\right) &= \sum\left\{X_k + \frac{\lambda_k}{\alpha_k^2}\left(r_k\alpha_k - bw\left(Y_k\right)\right)\right\}W_k\left(bU_k - V_k\right)\\ &= \sum\left\{\bar{X} + U_k + W_k\left(bU_k - V_k\right)\left(\frac{r_k}{\alpha_k} - \frac{b}{w\left(X_k\right)}\right)\right\}W_k\left(bU_k - V_k\right)\\ &= \bar{X}\sum W_k\left(bU_k - V_k\right) + \sum W_k U_k\left(bU_k - V_k\right) + \sum W_k^2\left(bU_k - V_k\right)^2\left(\frac{r_k}{\alpha_k} - \frac{b}{w\left(X_k\right)}\right)\\ &= \sum W_k U_k\left(bU_k - V_k\right) + \sum W_k^2\left(bU_k - V_k\right)^2\left(\frac{r_k}{\alpha_k} - \frac{b}{w\left(X_k\right)}\right)\end{aligned}$$

and (30) becomes

$$\sum W_k U_k\left(bU_k - V_k\right) + \sum W_k^2\frac{r_k}{\alpha_k}\left(bU_k - V_k\right)^2 - \sum W_k^2\frac{b}{w\left(X_k\right)}\left(bU_k - V_k\right)^2 = 0 \tag{37}$$

Expanding the terms within the summations then gathering coefficients of b and powers of b we have York's *Generalized Least Squares Cubic* (York, 1968, eq. 3; York et al., 2004, eq. B1)

$$\begin{aligned}&b^3\sum\frac{W_k^2}{w\left(X_k\right)}U_k^2 - b^2\sum W_k^2\left(\frac{r_k}{\alpha_k}U_k + \frac{2}{w\left(X_k\right)}V_k\right)U_k\\ &-b\sum W_k\left(U_k^2 - \frac{2W_k r_k}{\alpha_k}U_k V_k - \frac{W_k}{w\left(X_k\right)}V_k^2\right) + \sum W_k\left(U_k - \frac{W_k\, r_k}{\alpha_k}V_k\right)V_k = 0\end{aligned} \tag{38}$$

Equation (38) is not really a cubic equation in b because of the implicit dependence of the coefficients on the gradient b via the weight function W_k and the centroidal coordinates U_k, V_k (Reed 1989). And the solution for b is iterative and requires special techniques. A far simpler solution is obtained from the following reductions

Reductions of the Generalized Least Squares Cubic

York (1968) and York et al. (2004) show the *Generalized Least Squares Cubic* (38) reduced to quadratic and linear forms. This reduction process is set out below.

Rewrite (38) as

$$\begin{aligned}&b^3\sum\frac{W_k^2}{w\left(X_k\right)}U_k^2 - b^2\sum W_k^2\left(\frac{r_k U_k}{\alpha_k} + \frac{2V_k}{w\left(X_k\right)}\right)U_k\\ &-b\sum W_k^2\left(\frac{U_k^2}{W_k} - \frac{2r_k U_k V_k}{\alpha_k} - \frac{V_k^2}{w\left(X_k\right)}\right) + \sum W_k^2\left(\frac{U_k}{W_k} - \frac{r_k V_k}{\alpha_k}\right)V_k = 0\end{aligned} \tag{39}$$

and using (28) gives

$$
\begin{aligned}
&b^3\sum \frac{W_k^2 U_k^2}{w\left(X_k\right)} - b^2\sum W_k^2\left(\frac{r_k U_k}{\alpha_k} + \frac{2V_k}{w\left(X_k\right)}\right)U_k \\
&-b\sum W_k^2\left(\frac{U_k^2}{\alpha_k^2}\left(b^2 w\left(Y_k\right) + w\left(X_k\right) - 2b\, r_k\, \alpha_k\right) - \frac{2r_k U_k V_k}{\alpha_k} - \frac{V_k^2}{w\left(X_k\right)}\right) \\
&+\sum W_k^2\left(\frac{U_k}{\alpha_k^2}\left(b^2 w\left(Y_k\right) + w\left(X_k\right) - 2b\, r_k\, \alpha_k\right) - \frac{r_k V_k}{\alpha_k}\right)V_k = 0
\end{aligned}
$$

This simplifies to York (1968, eq 4) and York et al. (2004, eq B2)

$$
\begin{aligned}
&b^2\sum W_k^2\left(\frac{1}{w\left(X_k\right)}V_k - \frac{r_k}{\alpha_k}U_k\right)U_k + b\sum W_k^2\left(\frac{1}{w\left(Y_k\right)}U_k^2 - \frac{1}{w\left(X_k\right)}V_k^2\right) \\
&-\sum W_k^2\left(\frac{1}{w\left(Y_k\right)}U_k - \frac{r_k}{\alpha_k}V_k\right)V_k = 0
\end{aligned}
\tag{40}
$$

A linear equation can be obtained by rewriting (40) as

$$
b\sum W_k^2\left(\frac{b\,V_k}{w\left(X_k\right)} - \frac{b\, r_k U_k}{\alpha_k}\right)U_k + b\sum W_k^2\left(\frac{U_k}{w\left(Y_k\right)}\right)U_k - \sum W_k^2\left(\frac{b\,V_k}{w\left(X_k\right)}\right)V_k - \sum W_k^2\left(\frac{U_k}{w\left(Y_k\right)} - \frac{r_k V_k}{\alpha_k}\right)V_k = 0
$$

Gathering terms gives the linear equation (York et al., 2004, eq B3)

$$
b\sum_k W_k^2\left(\frac{1}{w\left(Y_k\right)}U_k + \frac{b}{w\left(X_k\right)}V_k - \frac{b\, r_k}{\alpha_k}U_k\right)U_k - \sum_k W_k^2\left(\frac{1}{w\left(Y_k\right)}U_k + \frac{b}{w\left(X_k\right)}V_k - \frac{r_k}{\alpha_k}V_k\right)V_k = 0 \tag{41}
$$

Solving for the leading coefficient gives York (1968, eq 6) and York et al. (2004, eq B4)

$$
b = \frac{\sum W_k^2\left(\frac{U_k}{w\left(Y_k\right)} + \frac{b\,V_k}{w\left(X_k\right)} - \frac{r_k V_k}{\alpha_k}\right)V_k}{\sum W_k^2\left(\frac{U_k}{w\left(Y_k\right)} + \frac{b\,V_k}{w\left(X_k\right)} - \frac{b\, r_k U_k}{\alpha_k}\right)U_k} \tag{42}
$$

Equation (41) may be rewritten as

$$
\begin{aligned}
b\sum W_k U_k Z_k\left(\frac{U_k}{w\left(Y_k\right)} + \frac{b\,V_k}{w\left(X_k\right)} - \frac{b r_k U_k}{\alpha_k}\right) - b\sum W_k U_k W_k\left(\frac{r_k V_k}{\alpha_k}\right) &= \sum W_k V_k W_k\left(\frac{U_k}{w\left(Y_k\right)} + \frac{b\,V_k}{w\left(X_k\right)} - \frac{r_k V_k}{\alpha_k}\right) \\
&\quad -b\sum W_k U_k W_k\left(\frac{r_k V_k}{\alpha_k}\right)
\end{aligned}
$$

that simplifies to

$$
b\sum W_k U_k W_k\left(\frac{U_k}{w\left(Y_k\right)} + \frac{b\,V_k}{w\left(X_k\right)} - \frac{r_k}{\alpha_k}\left(b\,U_k + V_k\right)\right) = \sum W_k V_k W_k\left(\frac{U_k}{w\left(Y_k\right)} + \frac{b\,V_k}{w\left(X_k\right)} - \frac{r_k}{\alpha_k}\left(b\,U_k + V_k\right)\right)
$$

Defining

$$
\beta_k = W_k\left(\frac{U_k}{w\left(Y_k\right)} + \frac{b\,V_k}{w\left(X_k\right)} - \frac{r_k}{\alpha_k}\left(b\,U_k + V_k\right)\right) \tag{43}
$$

gives (York et al., 2004, eq B6)

$$b\sum W_k\beta_k U_k = \sum W_k\beta_k V_k \quad \text{or} \quad b = \frac{\sum W_k\beta_k V_k}{\sum W_k\beta_k U_k} \tag{44}$$

This linear equation (44) is easiest to solve for b iteratively, as opposed to the quadratic equation (40) or the *Generalized Last Squares Cubic* (38) which require special techniques, especially (38) which requires a Newton-Raphson iterative scheme involving complicated derivatives whereas (40) requires the quadratic formula which contains a choice of $\pm$.

Iterative solutions for b using the linear equation (44) require an initial value b_0 which can be obtained from the special cases below, or more easily from estimating the gradient of a straight line from the extremities of the data set, disregarding any weights. Example 3 below shows the iterative sequence.

***y*-intercept** a, **adjusted values** x_k, y_k **and residuals** $v(X_k)$, $v(Y_k)$

The y-intercept a is obtained from (32) as

$$a = \bar{Y} - b\bar{X} \tag{45}$$

where the centroid $(\bar{X}, \bar{Y})$ is given by (31).

The adjusted values x_k, y_k are the coordinates of the orthogonal projection of $P_k(X_k, Y_k)$ onto the straight line $y = a + bx$ (see Figure 1).

Using (25), (26) and (27) the adjusted values x_k, y_k can be written as

$$x_k = X_k + W_k(a + bX_k - Y_k)\left(\frac{r_k}{\alpha_k} - \frac{b}{w(X_k)}\right)$$
$$y_k = Y_k + W_k(a + bX_k - Y_k)\left(\frac{1}{w(Y_k)} - \frac{r_k b}{\alpha_k}\right)$$

and substituting for the y-intercept a using (45) gives

$$x_k = X_k + W_k\left(b(X_k - \bar{X}) - (Y_k - \bar{Y})\right)\left(\frac{r_k}{\alpha_k} - \frac{b}{w(X_k)}\right)$$
$$y_k = Y_k + W_k\left(b(X_k - \bar{X}) - (Y_k - \bar{Y})\right)\left(\frac{1}{w(Y_k)} - \frac{r_k b}{\alpha_k}\right)$$

but, $U_k = X_k - \bar{X}$, $V_k = Y_k - \bar{Y}$ giving

$$x_k = \bar{X} + U_k + W_k(bU_k - V_k)\left(\frac{r_k}{\alpha_k} - \frac{b}{w(X_k)}\right)$$
$$y_k = \bar{Y} + V_k + W_k(bU_k - V_k)\left(\frac{1}{w(Y_k)} - \frac{r_k b}{\alpha_k}\right)$$

and after some algebra

$$\begin{aligned} x_k &= \bar{X} + W_k\left(\frac{r_k}{\alpha_k}\left(bU_k - V_k\right) - \frac{b^2U_k}{w\left(X_k\right)} + \frac{bV_k}{w\left(X_k\right)} + \frac{U_k}{W_k}\right) \\ y_k &= \bar{Y} + W_k\left(\frac{bU_k}{w\left(Y_k\right)} - \frac{V_k}{w\left(Y_k\right)} - \frac{r_k b}{\alpha_k}\left(bU_k - V_k\right) + \frac{V_k}{W_k}\right) \end{aligned} \tag{46}$$

With the aid of (28)

$$\frac{U_k}{W_k} = \frac{b^2U_k}{w\left(X_k\right)} + \frac{U_k}{w\left(Y_k\right)} - \frac{2br_kU_k}{\alpha_k} \text{ and } \frac{V_k}{W_k} = \frac{b^2V_k}{w\left(X_k\right)} + \frac{V_k}{w\left(Y_k\right)} - \frac{2br_kV_k}{\alpha_k} \tag{47}$$

and substituting (47) into (46) and simplifying gives

$$\begin{aligned} x_k &= \bar{X} + W_k\left(\frac{U_k}{w\left(Y_k\right)} + \frac{bV_k}{w\left(X_k\right)} - \frac{r_k}{\alpha_k}\left(bU_k + V_k\right)\right) \\ y_k &= \bar{Y} + bW_k\left(\frac{U_k}{w\left(Y_k\right)} + \frac{bV_k}{w\left(X_k\right)} - \frac{r_k}{\alpha_k}\left(bU_k + V_k\right)\right) \end{aligned}$$

And finally, using (43) gives

$$\begin{aligned} x_k &= \bar{X} + \beta_k \\ y_k &= \bar{Y} + b\beta_k \end{aligned} \tag{48}$$

The residuals $v\left(X_k\right), v\left(Y_k\right)$ are given by

$$\begin{aligned} v\left(X_k\right) &= x_k - X_k = \bar{X} + \beta_k - X_k = \beta_k - U_k \\ v\left(Y_k\right) &= y_k - Y_k = \bar{Y} + b\beta_k - Y_k = b\beta_k - V_k \end{aligned} \tag{49}$$

Estimates of the Variances of the gradient b and y-intercept a

To derive expressions for estimates of variance it is useful to consider the matrix form of the *Law of Propagation of Variances* (Mikhail 1976) expressed symbolically as:

$$\text{If } \mathbf{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}^T, \ \mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \end{bmatrix}^T \text{ and } \mathbf{y} = \mathbf{y}\left(\mathbf{x}\right) \text{ then } \mathbf{Q}_{yy} = \mathbf{J}_{yx}\mathbf{Q}_{xx}\mathbf{J}_{yx}^T \tag{50}$$

where

$\mathbf{y}$ is an n by 1 vector of computed quantities, $\mathbf{x}$ is an m by 1 vector of variables (having estimates of variances and covariances). The superscript ^{T} denotes transpose.

$\mathbf{y}\left(\mathbf{x}\right)$ represents a vector of linear functions of the variables $\mathbf{x}$.

$\mathbf{Q}_{yy}, \mathbf{Q}_{xx}$ are square variance matrices of orders n and m respectively

$$\mathbf{Q}_{yy} = \begin{bmatrix} s^2\left(y_1\right) & s\left(y_1y_2\right) & \cdots & s\left(y_1y_n\right) \\ s\left(y_2y_1\right) & s^2\left(y_2\right) & \cdots & s\left(y_2y_n\right) \\ \vdots & \vdots & \ddots & \vdots \\ s\left(y_ny_1\right) & s\left(y_ny_2\right) & \cdots & s^2\left(y_n\right) \end{bmatrix} \quad \mathbf{Q}_{xx} = \begin{bmatrix} s^2\left(x_1\right) & s\left(x_1x_2\right) & \cdots & s\left(x_1x_m\right) \\ s\left(x_2x_1\right) & s^2\left(x_2\right) & \cdots & s\left(x_2x_m\right) \\ \vdots & \vdots & \ddots & \vdots \\ s\left(x_mx_1\right) & s\left(x_mx_2\right) & \cdots & s^2\left(x_m\right) \end{bmatrix}$$

and $\mathbf{J}_{yx}$ is the Jacobian matrix of partial derivatives of order n by m

$$\mathbf{J}_{yx} = \begin{bmatrix} \partial y_1/\partial x_1 & \partial y_1/\partial x_2 & \cdots & \partial y_1/\partial x_m \\ \partial y_2/\partial x_1 & \partial y_2/\partial x_2 & \cdots & \partial y_2/\partial x_m \\ \vdots & \vdots & \ddots & \vdots \\ \partial y_n/\partial x_1 & \partial y_n/\partial x_2 & \cdots & \partial y_n/\partial x_m \end{bmatrix}$$

In the case of the gradient b computed from the set of coordinates $\{X_k, Y_k\}$ for $k = 1,2,\ldots,n$ we may write $\mathbf{y} = [b]$ and $\mathbf{x} = \begin{bmatrix} X_1 & Y_1 & X_2 & Y_2 & \cdots & X_n & Y_n \end{bmatrix}^T$ with $\mathbf{Q}_{yy} = \mathbf{Q}_b = [s^2(b)]$, $\mathbf{Q}_{xx} = \mathbf{Q}$ given in (4) and $\mathbf{J}_{yx} = \begin{bmatrix} \frac{\partial b}{\partial X_1} & \frac{\partial b}{\partial Y_1} & \frac{\partial b}{\partial X_2} & \frac{\partial b}{\partial Y_2} & \cdots & \frac{\partial b}{\partial X_n} & \frac{\partial b}{\partial Y_n} \end{bmatrix}$.

Performing the matrix multiplications in (50) gives the estimate of the variance of the gradient b as

$$s^2(b) = \sum \left\{ s^2(X_k)\left(\frac{\partial b}{\partial X_k}\right)^2 + 2\frac{\partial b}{\partial X_k}\frac{\partial b}{\partial Y_k} s(X_k Y_k) + s^2(Y_k)\left(\frac{\partial b}{\partial Y_k}\right)^2 \right\} \tag{51}$$

Using the relationships given in (7), (10) and (11), (51) can be written as

$$s^2(b) = \sum \left\{ \frac{1}{w(X_k)}\left(\frac{\partial b}{\partial X_k}\right)^2 + 2\frac{\partial b}{\partial X_k}\frac{\partial b}{\partial Y_k}\frac{r_k}{\alpha_k} + \frac{1}{w(Y_k)}\left(\frac{\partial b}{\partial Y_k}\right)^2 \right\} \tag{52}$$

Similarly, the estimate of the variance of the y-intercept a is

$$s^2(a) = \sum \left\{ \frac{1}{w(X_k)}\left(\frac{\partial a}{\partial X_k}\right)^2 + 2\frac{\partial a}{\partial X_k}\frac{\partial a}{\partial Y_k}\frac{r_k}{\alpha_k} + \frac{1}{w(Y_k)}\left(\frac{\partial a}{\partial Y_k}\right)^2 \right\} \tag{53}$$

The partial derivatives in (52) and (53) are:

$$\frac{\partial b}{\partial X_k} = \frac{W_k\left\{V_k - 2b(\beta_k - \overline{\beta})\right\}}{D}, \quad \frac{\partial b}{\partial Y_k} = \frac{W_k\left\{2(\beta_k - \overline{\beta}) - U_k\right\}}{D} \tag{54}$$

$$\frac{\partial a}{\partial X_k} = -\frac{\partial b}{\partial X_k}(2\overline{\beta} + \overline{X}) - \frac{bW_k}{\sum W_k}, \quad \frac{\partial a}{\partial Y_k} = -\frac{\partial b}{\partial Y_k}(2\overline{\beta} + \overline{X}) + \frac{W_k}{\sum W_k} \tag{55}$$

where

$$D = \begin{cases} \frac{1}{b}\sum\left\{W_k U_k V_k - W_k^2 \frac{r_k}{\alpha_k}(bU_k - V_k)^2\right\} + 4\sum W_k(\beta_k - U_k)(\beta_k - \overline{\beta}) & \text{if } b \neq 0 \\ \sum w^2(Y_k)\left(\frac{U_k^2}{w(Y_k)} - \frac{V_k^2}{w(X_k)}\right) + 4\sum w(Y_k)(\beta_k - U_k)(\beta_k - \overline{\beta}) & \text{if } b = 0 \end{cases} \tag{56}$$

and

$$\overline{\beta} = \begin{cases} \frac{\sum W_k \beta_k}{\sum W_k} & \text{if } b \neq 0 \\ \frac{\sum w(Y_k)\beta_k}{\sum w(Y_k)} & \text{if } b = 0 \end{cases} \tag{57}$$

Note: In (56) and (57) for the special case $b = 0$, the weight functions $W_k = w(Y_k)$ and special results follow for the centroid $(\overline{X}, \overline{Y})$, centroidal coordinates (U_k, V_k) and factors $\beta_k = U_k - V_k r_k \sqrt{\frac{w(Y_k)}{w(X_k)}}$.

Substituting the partial derivatives into (52) and (53) and simplifying gives [York et al. 2004, eq's (5) and (11) respectively]

$$s^2\left(b\right) = \begin{cases} \dfrac{1}{D^2}\sum W_k^2\left\{\dfrac{U_k^2}{w\left(Y_k\right)} - 2\dfrac{r_k}{\alpha_k}U_kV_k + \dfrac{V_k^2}{w\left(X_k\right)}\right\} & \text{if } b \neq 0 \\ \dfrac{1}{D^2}\sum w^2\left(Y_k\right)\left\{\dfrac{U_k^2}{w\left(Y_k\right)} - 2\dfrac{r_k}{\alpha_k}U_kV_k + \dfrac{V_k^2}{w\left(X_k\right)}\right\} & \text{if } b = 0 \end{cases} \tag{58}$$

$$s^2\left(a\right) = \begin{cases} \left(2\bar{\beta} + \bar{X}\right)^2 s^2\left(b\right) + \dfrac{2}{D}\left(2\bar{\beta} + \bar{X}\right)\bar{\beta} + \dfrac{1}{\sum W_k} & \text{if } b \neq 0 \\ \left(2\bar{\beta} + \bar{X}\right)^2 s^2\left(b\right) + \dfrac{2}{D}\left(2\bar{\beta} + \bar{X}\right)\bar{\beta} + \dfrac{1}{\sum w\left(Y_k\right)} & \text{if } b = 0 \end{cases} \tag{59}$$

Note: In (58) and (59) for the special case $b = 0$, the weight function $W_k = w\left(Y_k\right)$ and special results follow for the centroid $\left(\bar{X},\bar{Y}\right)$, the centroidal coordinates $\left(U_k,V_k\right)$, D and the factors β_k.

The derivations of the partial derivatives $\dfrac{\partial b}{\partial X_k},\dfrac{\partial b}{\partial Y_k},\dfrac{\partial a}{\partial X_k},\dfrac{\partial a}{\partial Y_k}$, the equation for D and the simplified expressions for the variances $s^2(a), s^2(b)$ are given in detail in Appendix B.

Special Cases

[1] Suppose the measurements X_k, Y_k are independent quantities of unit weight; i.e., $r_k = 0$, $w\left(X_k\right) = w\left(Y_k\right) = 1$, $W_k = \dfrac{1}{1+b^2}$ and $\beta_k = \dfrac{1}{1+b^2}\left(U_k + bV_k\right)$. Substituting these results into (44) and simplifying gives

$$b^2\sum U_kV_k + b\sum\left(U_k^2 - V_k^2\right) - \sum U_kV_k = 0$$

Using the standard solution for a quadratic equation gives (York 1966)

$$b = \frac{\sum\left(V_k^2 - U_k^2\right) \pm \sqrt{\left\{\sum\left(V_k^2 - U_k^2\right)\right\}^2 + 4\left\{\sum U_kV_k\right\}^2}}{2\sum U_kV_k} \tag{60}$$

See Example 1.

[2] Suppose the measurements X_k, Y_k are independent, i.e., $r_k = 0$ and

$$b = \frac{\sum W_k\beta_kV_k}{\sum W_k\beta_kU_k}$$

where $$W_k = \frac{\alpha_k^2}{b^2 w\left(Y_k\right) + w\left(X_k\right)},\quad \beta_k = W_k\left(\frac{U_k}{w\left(Y_k\right)} + \frac{bV_k}{w\left(X_k\right)}\right)$$

Solution for b is iterative and requires a starting value. See Example 2.

[3] Suppose the measurements X_k, Y_k are independent quantities, but the measurements X_k are regarded as error-free, i.e., $r_k = 0$, $w\left(X_k\right) = \infty$ giving $W_k = w\left(Y_k\right)$ and $\beta_k = U_k$ and

$$b = \frac{\sum w\left(Y_k\right) U_k V_k}{\sum w\left(Y_k\right) U_k^2} \tag{61}$$

[4] Suppose the measurements X_k, Y_k are independent quantities, but the measurements Y_k are regarded as error-free, i.e., $r_k = 0$, $w\left(Y_k\right) = \infty$ giving $W_k = \frac{w\left(X_k\right)}{b^2}$ and $\beta_k = \frac{V_k}{b}$ and

$$b = \frac{\sum w\left(X_k\right) V_k^2}{\sum w\left(X_k\right) U_k V_k} \tag{62}$$

Example 1

Find the best fitting straight line to the following system of points supposed of equal weights (Pearson 1901, p. 569)

k	X_k	$w\left(X_k\right)$	Y_k	$w\left(Y_k\right)$
1	0.0	1	5.9	1
2	0.9	1	5.4	1
3	1.8	1	4.4	1
4	2.6	1	4.6	1
5	3.3	1	3.5	1
6	4.4	1	3.7	1
7	5.2	1	2.8	1
8	6.1	1	2.8	1
9	6.5	1	2.4	1
10	7.4	1	1.5	1

Table 1. Pearson's data (uncorrelated, unit weights)

In this example measurements are uncorrelated with weights equal to unity, i.e., $r_k = 0$, $w\left(X_k\right) = w\left(Y_k\right) = 1$, $W_k = \frac{1}{1+b^2}$, $\beta_k = \frac{1}{1+b^2}\left(U_k + bV_k\right)$ and the result from Special Cases [1] can be used to determine b directly from (60) noting the possibility of two solutions.

$$b = \frac{\sum\left(V_k^2 - U_k^2\right) \pm \sqrt{\left\{\sum\left(V_k^2 - U_k^2\right)\right\}^2 + 4\left\{\sum U_k V_k\right\}^2}}{2\sum U_k V_k}$$

where $$U_k = X_k - \bar{X},\ V_k = Y_k - \bar{Y},\ \bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k,\ \ \bar{Y} = \frac{1}{n}\sum_{k=1}^{n} Y_k$$

Using the data in Table 1, the centroid is $\left(\bar{X} = \frac{38.2}{10} = 3.82,\ \bar{Y} = \frac{37.0}{10} = 3.70\right)$ and the centroidal coordinates U, V (and their squares and products) and the factors β_k are shown in Table 2

k	U_k	V_k	U_k^2	V_k^2	U_kV_k	β_k
1	-3.82	2.20	14.5924	4.8400	-8.4040	-3.868751
2	-2.92	1.70	8.5264	2.8900	-4.9640	-2.964969
3	-2.02	0.70	4.0804	0.4900	-1.4140	-1.850974
4	-1.22	0.90	1.4884	0.8100	-1.0980	-1.318554
5	-0.52	-0.20	0.2704	0.0400	0.1040	-0.316643
6	0.58	0.00	0.3364	0.0000	0.0000	0.446966
7	1.38	-0.90	1.9044	0.8100	-1.2420	1.441856
8	2.28	-0.90	5.1984	0.8100	-2.0520	2.135424
9	2.68	-1.30	7.1824	1.6900	-3.4840	2.611847
10	3.58	-2.20	12.8164	4.8400	-7.8760	3.683800
Sums	0.00	0.00	56.3960	17.2200	-30.4300	

Table 2. A test data set: Pearson's data (unit weights) and centroidal coordinates U, V

$$\sum\left(V_k^2 - U_k^2\right) = \sum V_k^2 - \sum U_k^2 = 17.2200 - 56.3960 = -39.1760$$

$$\sum U_k V_k = -30.4300$$

$$b = \frac{-39.1760 \pm \sqrt{\left\{-39.1760\right\}^2 + 4\left\{-30.4300\right\}^2}}{2\left(-30.4300\right)} = -0.545561 \text{ or } 1.832975$$

Choosing the negative gradient which is supported by the scatter plot of Figure 2,

$$a = \bar{Y} - b\bar{X} = 3.70 - \left(-0.545561\right)\left(3.82\right) = 5.7840$$

The residuals [using (49)] squared residuals and the least squares function S [sum of the squared residuals, see (16) with $w\left(X_k\right) = w\left(Y_k\right) = 1$] are shown in Table 3

	Residuals		
k	$v\left(X_k\right)$	$v\left(Y_k\right)$	squared residuals
1	-0.048751	-0.089360	0.010362
2	-0.044969	-0.082428	0.008817
3	0.169026	0.309820	0.124558
4	-0.098554	-0.180648	0.042347
5	0.203357	0.372748	0.180295
6	-0.133034	-0.243847	0.077160
7	0.061856	0.113380	0.016681
8	-0.144576	-0.265004	0.091130
9	-0.068153	-0.124922	0.020250
10	0.103800	0.190262	0.046974
			$S =$ 0.618573

Table 3. Residuals and least squares function S

The estimates of variances of gradient b and y-intercept a

Using (57) and (56) $\bar{\beta} = 0$, $D = 42.983881$; and from (58) and (59) the variances of the gradient b and y-intercept a are

$$s^2\left(b\right) = 0.023662, \quad s^2\left(a\right) = 0.475052$$

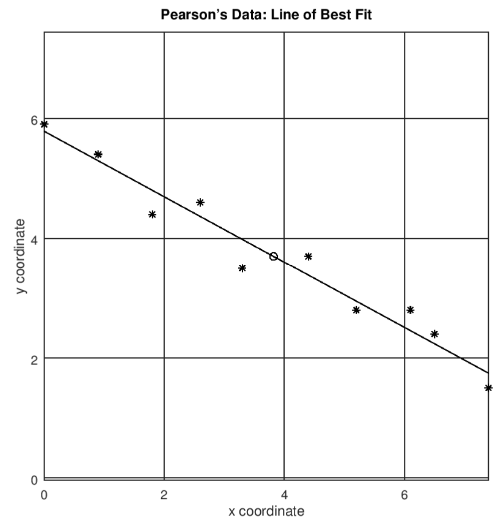

Figure 2. Line of Best Fit for Pearson's data (unit weights). Centroid shown thus ○

Example 2

Find the best fitting straight line to the following system of points of varying weights. Pearson's data with York's weights (Pearson 1901, p. 569, York 1966, Table II, p. 1086). The coordinate values are assumed to be independent, hence the covariance $s\left(X_k Y_k\right) = 0$ and $r_k = 0$

k	X_k	$w\left(X_k\right)$	Y_k	$w\left(Y_k\right)$	r_k
1	0	1000.0	5.9	1.0	0
2	0.9	1000.0	5.4	1.8	0
3	1.8	500.0	4.4	4.0	0
4	2.6	800.0	4.6	8.0	0
5	3.3	200.0	3.5	20.0	0
6	4.4	80.0	3.7	20.0	0
7	5.2	60.0	2.8	70.0	0
8	6.1	20.0	2.8	70.0	0
9	6.5	1.8	2.4	100.0	0
10	7.4	1.0	1.5	500.0	0

Table 4. A test data set: Pearson's data with York's weights and correlation $r_k = 0$

1. Solve for b using (44) and in the following iterative sequence:

   (i) Choose an appropriate starting value of b (for instance, assume the measurements are uncorrelated and of unit weight as in Exercise 1 and $b_0 = -0.545561$).

   (ii) First determine the weight functions W_k and weighted measurements $W_k X_k$, $W_k Y_k$ for each of the 10 points. Second, calculate the centroid $\left(\bar{X}, \bar{Y}\right)$ and finally, for each of the 10 points, calculate the centroidal values U_k, V_k and the factors β_k.

k	W_k	W_kX_k	W_kY_k	U_k	V_k	β_k
1	0.999702	0.000000	5.898244	-4.839249	2.747178	-4.839307
2	1.799036	1.619133	9.714795	-3.939249	2.247178	-3.939345
3	3.990498	7.182897	17.558192	-3.039249	1.247178	-3.037460
4	7.976260	20.738275	36.690795	-2.239249	1.447178	-2.240476
5	19.421932	64.092375	67.976761	-1.539249	0.347178	-1.513153
6	18.614882	81.905479	68.875062	-0.439249	0.547178	-0.478290
7	51.957964	270.181412	145.482299	0.360751	-0.352822	0.434456
8	34.284671	209.136491	95.997078	1.260751	-0.352822	0.947458
9	5.702757	37.067921	13.686617	1.660751	-0.752822	1.395921
10	3.337374	24.696566	5.006061	2.560751	-1.652822	3.026453

Table 5. W_k , W_kX_k , W_kY_k , U_k , V_k and β_k for the 1st iteration

(iii) Use W_k, U_k, V_k and β_k in (44) to calculate an improved estimate of b.

(iv) Use the improved estimate of b and repeat steps (ii) and (iii) until the difference between successive estimates b_{n+1} and b_n reach an acceptably small value.

Table 6 shows successive values of b starting at $b_0 = -0.545561$ and converging to $b = b_4 = -0.480533$

Iteration n	b_n	b_{n+1}
0	-0.545561	-0.479758
1	-0.479758	-0.480557
2	-0.480557	-0.480533
3	-0.480533	-0.480533

Table 6.

2. Using the final value of b, first compute the weight functions W_k and weighted measurements W_kX_k, W_kY_k for each of the 10 points. Then calculate the centroid $\left(\bar{X}, \bar{Y}\right)$; the centroidal values U_k, V_k and the factors β_k .

k	W_k	W_kX_k	W_kY_k	U_k	V_k	β_k
1	0.999769	0.000000	5.898638	-4.910970	2.779975	2.779975
2	1.799252	1.619327	9.715962	-4.010970	2.279975	2.279975
3	3.992624	7.186724	17.567548	-3.110970	1.279975	1.279975
4	7.981570	20.752081	36.715220	-2.310970	1.479975	1.479975
5	19.548599	64.510378	68.420098	-1.610970	0.379975	0.379975
6	18.908453	83.197193	69.961276	-0.510970	0.579975	0.579975
7	55.144280	286.750257	154.403985	0.289030	-0.320025	-0.320025
8	38.712706	236.147505	108.395576	1.189030	-0.320025	-0.320025
9	7.231472	47.004571	17.355534	1.589030	-0.720025	-0.720025
10	4.293468	31.771661	6.440202	2.489030	-1.620025	-1.620025
Sums	158.612194	778.939698	494.874038			

Table 7. W_k , W_kX_k , W_kY_k , U_k , V_k and β_k for the last iteration

$$\overline{X} = \frac{\sum W_k X_k}{\sum W_k} = \frac{778.939698}{158.612194} = 4.910970, \quad \overline{Y} = \frac{\sum W_k Y_k}{\sum W_k} = \frac{494.874038}{158.612194} = 3.120025$$

3. Calculate the intercept a from (45).

$$a = \overline{Y} - b\overline{X} = 3.120025 - \left(-0.480533\right)\left(4.910970\right) = 5.479908$$

4. Calculate the residuals using (49) and the least squares function S using (16).

	Residuals		
k	$v\left(X_k\right)$	$v\left(Y_k\right)$	weighted squared residuals
1	-0.000202	-0.419995	0.176436
2	-0.000305	-0.352425	0.223659
3	0.000825	0.214552	0.184471
4	-0.001771	-0.368626	1.089593
5	0.018513	0.385253	3.036947
6	-0.037984	-0.316184	2.114874
7	0.079998	0.142695	1.809310
8	-0.233783	-0.139002	2.445611
9	-0.084087	-0.003150	0.013719
10	0.874703	0.003641	0.771732
			$S = 11.866353$

Table 8. Residuals and least squares function S

5. Calculate the estimates of variances for gradient b and y-intercept a.

Using (57) and (56) $\overline{\beta} = -0.011723$, $D = 301.704504$; and from (58) and (59) the variances of the gradient b and y-intercept a are

$$s^2\left(b\right) = 0.003320$$
$$s^2\left(a\right) = 0.085225$$

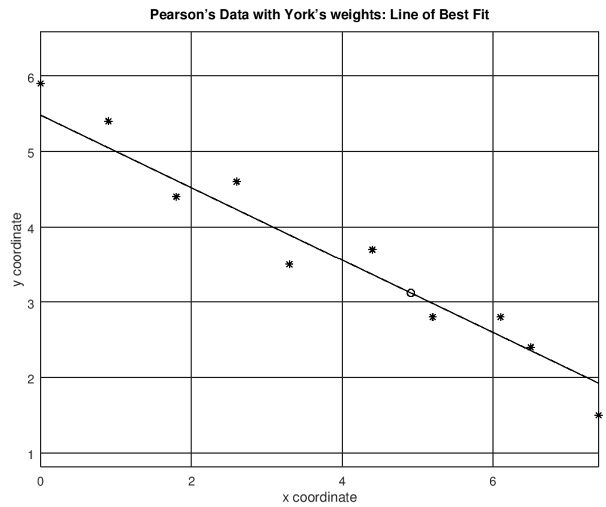

Figure 3. Line of Best Fit for Pearson's data with York's weights. Centroid shown thus ○

Example 3

Find the best fitting straight line to the following system of points of varying weights and correlations. (Pearson's data with York's weights and randomly assigned correlations)

k	X_k	$w(X_k)$	Y_k	$w(Y_k)$	r_k
1	0	1000.0	5.9	1.0	0.989
2	0.9	1000.0	5.4	1.8	-0.870
3	1.8	500.0	4.4	4.0	-0.223
4	2.6	800.0	4.6	8.0	0.099
5	3.3	200.0	3.5	20.0	-0.057
6	4.4	80.0	3.7	20.0	-0.660
7	5.2	60.0	2.8	70.0	0.022
8	6.1	20.0	2.8	70.0	0.741
9	6.5	1.8	2.4	100.0	-0.335
10	7.4	1.0	1.5	500.0	-0.001

Table 9. Pearson's data with York's weights and randomly assigned correlations

1. Solve for b using (44) in the following iterative sequence:

   (i) Choose an appropriate starting value of b (for instance, assume the measurements are uncorrelated and of unit weight as in Exercise 1 and $b_0 = -0.545561$).

   (ii) Determine the weight functions W_k and weighted measurements $W_k X_k$, $W_k Y_k$ for each of the 10 points. Second, calculate the centroid $(\bar{X}, \bar{Y})$ and for each of the 10 points, calculate the centroidal values U_k, V_k. Finally calculate the factors β_k.

k	W_k	$W_k X_k$	$W_k Y_k$	U_k	V_k	β_k
1	0.966723	0.000000	5.703666	-4.686321	2.677183	-4.690028
2	1.874490	1.687041	10.122244	-3.786321	2.177183	-3.782149
3	4.079061	7.342310	17.947868	-2.886321	1.177183	-2.892637
4	7.891270	20.517303	36.299843	-2.086321	1.377183	-2.089941
5	19.800094	65.340310	69.300328	-1.386321	0.277183	-1.368993
6	27.997914	123.190820	103.592281	-0.286321	0.477183	-0.199326
7	50.976901	265.079885	142.735323	0.513679	-0.422817	0.582231
8	19.694267	120.135029	55.143948	1.413679	-0.422817	1.090597
9	6.751780	43.886570	16.204272	1.813679	-0.822817	1.500733
10	3.337917	24.700588	5.006876	2.713679	-1.722817	3.154953

Table 10. W_k, $W_k X_k$, $W_k Y_k$, U_k, V_k and β_k for the 1st iteration

   (iii) Use W_k, U_k, V_k and β_k in (44) to calculate an improved estimate of b.

   (iv) Use the improved estimate of b and repeat steps (ii) and (iii) until the difference between successive estimates b_{n+1} and b_n reach an acceptably small value.

Table 11 shows successive values of b starting at $b_0 = -0.545561$ and converging to $b = b_5 = -0.494346$

Iteration n	b_n	b_{n+1}
0	-0.545561	-0.493216
1	-0.493216	-0.494378
2	-0.494378	-0.494345
3	-0.494345	-0.494346
4	-0.494346	-0.494346

Table 11.

2. Using the final value of b, first compute the weight functions W_k and weighted measurements $W_k X_k$, $W_k Y_k$ for each of the 10 points. Then calculate the centroid $(\bar{X}, \bar{Y})$; the centroidal values U_k, V_k and the factors β_k.

k	W_k	$W_k X_k$	$W_k Y_k$	U_k	V_k	β_k
1	0.969776	0.000000	5.721680	-4.746251	2.708954	-4.757425
2	1.867324	1.680592	10.083549	-3.846251	2.208954	-3.834758
3	4.072346	7.330223	17.918322	-2.946251	1.208954	-2.950281
4	7.903328	20.548652	36.355308	-2.146251	1.408954	-2.151354
5	19.868537	65.566172	69.539880	-1.446251	0.308954	-1.433583
6	27.217323	119.756220	100.704094	-0.346251	0.508954	-0.251367
7	53.492215	278.159518	149.778202	0.453749	-0.391046	0.530262
8	21.699163	132.364896	60.757657	1.353749	-0.391046	1.085011
9	8.259107	53.684195	19.821857	1.753749	-0.791046	1.597214
10	4.059533	30.040544	6.089300	2.653749	-1.691046	3.414615
Sums	149.408652	709.131012	476.769848			

Table 12. W_k, $W_k X_k$, $W_k Y_k$, U_k, V_k and β_k for the last iteration

$$\bar{X} = \frac{\sum W_k X_k}{\sum W_k} = \frac{709.131012}{149.408652} = 4.746251, \quad \bar{Y} = \frac{\sum W_k Y_k}{\sum W_k} = \frac{476.769848}{149.408652} = 3.191046$$

3. Calculate the intercept a from (45).

$$a = \bar{Y} - b\bar{X} = 3.191046 - (-0.494346)(4.746251) = 5.537336$$

4. Calculate the residuals using (49) and the least squares function S using (15).

k	Residuals $v(X_k)$	$v(Y_k)$	weighted squared residuals with correlation
1	-0.011173	-0.357140	0.127550
2	0.011494	-0.313257	0.176654
3	-0.004030	0.249505	0.249484
4	-0.005103	-0.345441	0.956924
5	0.012668	0.399732	3.274959
6	0.094885	-0.384692	3.105485
7	0.076513	0.128913	1.487148
8	-0.268738	-0.145325	1.679104
9	-0.156534	0.001469	0.047595
10	0.760866	0.003045	0.583656
			$S = 11.688557$

Table 13. Residuals and least squares function S

5. Calculate the estimates of variances for gradient b and y-intercept a.

Using (57) and (56) $\overline{\beta} = 0.019045$; $D = 277.150131$; and from (58) and (59) the variances of the gradient b and y-intercept a are

$$s^2\left(b\right) = 0.003586$$
$$s^2\left(a\right) = 0.089426$$

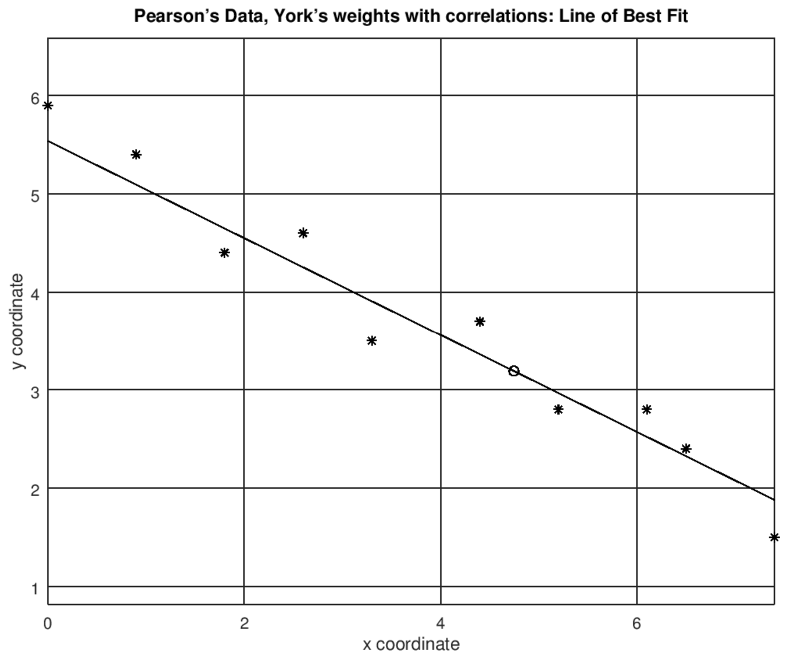

Figure 4. Line of Best Fit for Pearson's data with York's weights and correlations.
Centroid shown thus ○

APPENDIX A: Alternative Derivation of York's Generalized Least Squares Cubic

In York's 1968 paper: *Least Squares Fitting of a Straight Line with Correlated Errors,* he suggests two methods of development of the *Generalized Least Squares Cubic* – first developed for the uncorrelated case in York (1966). A copy of part of the first page is shown below.

> Two equivalent approaches may be adopted, corresponding to the two different points adopted in the uncorrelated-errors case by York [1] and McIntyre et al. [2]. Firstly one may begin by minimizing the expression
>
> $$S = \sum_i \left\{ w\left(X_i\right)\left(x_i - X_i\right)^2 - 2r_i\sqrt{w\left(X_i\right)w\left(Y_i\right)}\left(x_i - X_i\right)\left(y_i - Y_i\right) + w\left(Y_i\right)\left(y_i - Y_i\right)^2 \right\}\frac{1}{1 - r_i^2} \qquad (1)$$
>
> subject to the requirement
>
> $$y_i = a + bx_i, \quad i = 1,\ldots,n\,.$$
>
> X_i, Y_i are the observations, x_i, y_i are the adjusted values of these, $w\left(X_i\right), w\left(Y_i\right)$ are the weights of the various observations, and the r_i are the correlations between the x and y errors. Alternatively one may start by minimizing the expression
>
> $$S = \sum_i Z_i\left(Y_i - bX_i - a\right)^2 \qquad (2)$$
>
> where
>
> $$Z_i = \frac{w\left(X_i\right)w\left(Y_i\right)}{b^2 w\left(Y_i\right) + w\left(X_i\right) - 2br_i\sqrt{w\left(X_i\right)w\left(Y_i\right)}}$$
>
> Pursuing the analysis we find the following generalized versions, for the case of correlated x and y errors, of the least squares cubic and quadratic equations, either of which may be solved for b to yield the best slope:

We have shown how York's equation (1) is derived and have then provided a detailed derivation of the *Generalized Least Squares Cubic* following York's first alternative: minimizing his equation (1) with the conditions that $y_k = a + bx_k$. [Note that York's equation (1) above is our equation (15) with a slight change of notation].

We now show how York's equation (2) is obtained and how this function may be minimized to obtain the *Generalized Least Squares Cubic*. [Note that we have used W_k rather than Z_k].

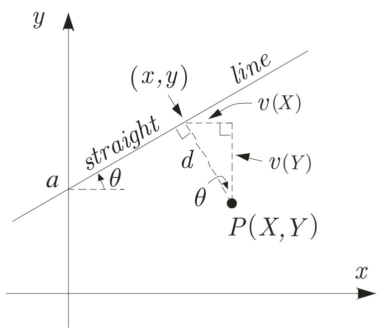

Figure A1. The orthogonal projection of $P\left(X,Y\right)$ onto the straight line $y = a + bx$

In Figure A1, X, Y are the measured coordinates of P, x, y are the adjusted coordinates, d is the length of the normal to the straight line $y = a + bx$ and $v(X), v(Y)$ are residuals such that $X + v(X) = x$ and $Y + v(Y) = y$.

From the diagram

$$v(X) = -v(Y)\tan\theta = -bv(Y) \tag{A1}$$

and $y = a + bx$ can be written as $Y + v(Y) = a + b(X + v(X))$. Using (A1) and simplifying gives

$$v(Y) = \frac{a + bX - Y}{1 + b^2} \tag{A2}$$

Now, from Figure A1, $d^2 = (v(X))^2 + (v(Y))^2$ and using (A1) and (A2) and simplifying gives

$$d^2 = \frac{(a + bX - Y)^2}{1 + b^2} \tag{A3}$$

If measurements X, Y have estimates of variances $s^2(X), s^2(Y)$ and covariances $s(XY)$ then the computed quantity d will have variance $s^2(d)$ obtained from the matrix form of the *Law of Propagation of Variances* (Mikhail 1976) expressed symbolically as equation (50)

In the case of the computed distance d, where d is a function of the variables X and Y, then $\mathbf{y} = [d]$,

$\mathbf{x} = \begin{bmatrix} X & Y \end{bmatrix}^T$, $\mathbf{Q}_{yy} = \mathbf{Q}_{dd} = [s^2(d)]$, $\mathbf{Q}_{xx} = \begin{bmatrix} s^2(X) & s(X,Y) \\ s(X,Y) & s^2(Y) \end{bmatrix} = \frac{1}{\alpha^2}\begin{bmatrix} w(Y) & r\alpha \\ r\alpha & w(X) \end{bmatrix}$ and $\mathbf{J}_{yx} = \begin{bmatrix} \dfrac{\partial d}{\partial X} & \dfrac{\partial d}{\partial Y} \end{bmatrix}$.

The partial derivatives are

$$\frac{\partial d}{\partial X} = \frac{b(a + bX - Y)}{d(1 + b^2)}, \quad \frac{\partial d}{\partial Y} = \frac{-(a + bX - Y)}{d(1 + b^2)}$$

And using these results in (50) and simplifying gives

$$\mathbf{Q}_{dd} = \frac{b^2 w(Y) + w(X) - 2br\alpha}{\alpha^2(1 + b^2)} \tag{A4}$$

Now $\mathbf{Q}_{dd} = [s^2(d)]$ and since $\mathbf{W}_d = [w(d)] = \mathbf{Q}_{dd}^{-1}$ we may write the weight of an orthogonal distance d as

$$w(d) = \frac{\alpha^2(1 + b^2)}{b^2 w(Y) + w(X) - 2br\alpha} \tag{A5}$$

The least squares function S is the sum of the weighted squared distances, i.e.,

$$S = \sum_{k=1}^{n} w(d_k)\, d_k^2 = \sum_{k=1}^{n} \left\{ \left(\frac{\alpha_k^2(1 + b^2)}{b^2 w(Y_k) + w(X_k) - 2br_k\alpha_k} \right) \frac{(a + bX_k - Y_k)^2}{1 + b^2} \right\}$$

giving York [1968, eq. (2)]

$$S = \sum W_k \left(a + bX_k - Y_k\right)^2 \tag{A6}$$

where

$$W_k = \frac{\alpha_k^2}{b^2 w\left(Y_k\right) + w\left(X_k\right) - 2br_k\alpha_k}$$

Alternatively, (A6) may be obtained by incorporating the constraints $y = a + bx$ in (1) and writing

$$\begin{aligned} S &= \sum \frac{1}{1 - r_k^2}\left\{ w\left(X_k\right)\left(x_k - X_k\right)^2 - 2r_k\sqrt{w\left(X_k\right)w\left(Y_k\right)}\left(x_k - X_k\right)\left(y_k - Y_k\right) + w\left(Y_k\right)\left(y_k - Y_k\right)^2 \right\} \\ &= \sum \frac{1}{1 - r_k^2}\left\{ w\left(X\right)\left(x - X\right)^2 - 2r\alpha\left(x - X\right)\left(a + bx - Y\right) + w\left(Y\right)\left(a + bx - Y\right)^2 \right\} \end{aligned}$$

then partial differentiation of both sides with respect to x gives (noting that S is a scalar and $\partial S / \partial x = 0$ for optimum)

$$\begin{aligned} 0 &= \sum \frac{2}{1 - r^2}\left\{ \left(x - X\right)w\left(X\right) - r\alpha\left(a + 2bx - Y - bX\right) + b\left(a + bx - Y\right)w\left(Y\right) \right\} \\ &= \sum \frac{2}{1 - r^2}\left\{ \left[b^2 w\left(Y\right) + w\left(X\right) - 2r\alpha b\right]x - \left[w\left(X\right) - r\alpha b\right]X - \left[r\alpha - bw\left(Y\right)\right]\left(Y - a\right) \right\} \end{aligned}$$

This equation will be satisfied when

$$\left[b^2 w\left(Y\right) + w\left(X\right) - 2r\alpha b\right]x = \left[w\left(X\right) - r\alpha b\right]X - \left[r\alpha - bw\left(Y\right)\right]\left(Y - a\right)$$

and we may simplify this equation by letting

$$\mu = \frac{1}{b^2 w\left(Y\right) + w\left(X\right) - 2br\alpha} = \frac{W^2}{\alpha^2}, \quad \xi = w\left(X\right) - br\alpha \quad \text{and} \quad \zeta = r\alpha - bw\left(Y\right)$$

giving

$$x = \mu\left[\xi X - \zeta\left(Y - a\right)\right]$$

Using this result for x, and noting that $\mu\xi - \mu b\zeta = 1$, then

$$\begin{aligned} &w\left(X\right)\left(x - X\right)^2 - 2r\alpha\left(x - X\right)\left(y - Y\right) + w\left(Y\right)\left(y - Y\right)^2 \\ &= \left\{\left(\mu\xi - 1\right)X - \mu\zeta\left(Y - a\right)\right\}^2 w\left(X\right) \\ &\quad -2r\alpha\left\{\left(\mu\xi - 1\right)X - \mu\zeta\left(Y - a\right)\right\}\left\{\mu b\xi X - \left(\mu b\zeta + 1\right)\left(Y - a\right)\right\} \\ &\quad +\left\{\mu b\xi X - \left(\mu b\zeta + 1\right)\left(Y - a\right)\right\}^2 w\left(Y\right) \\ &= \mu^2\zeta^2\left[bX - \left(Y - a\right)\right]^2 w\left(X\right) - 2r\alpha\mu^2\xi\zeta\left[bX - \left(Y - a\right)\right]\left[bX - \left(Y - a\right)\right] \\ &\quad +\mu^2\xi^2\left[bX - \left(Y - a\right)\right]^2 w\left(Y\right) \\ &= \left(a + bX - Y\right)^2 \mu^2\left[\alpha^2\left\{b^2 w\left(Y\right) + w\left(X\right) - 2r\alpha b\right\} - r^2\alpha^2\left\{b^2 w\left(Y\right) + w\left(X\right) - 2r\alpha b\right\}\right] \\ &= \mu\alpha^2\left(a + bX - Y\right)^2\left(1 - r^2\right) \\ &= W\left(a + bX - Y\right)^2\left(1 - r^2\right) \end{aligned}$$

giving the alternate form of (1) as (A6)

$$S = \sum W_k \left(a + bX_k - Y_k \right)^2 \tag{A6}$$

This demonstrates that S in this form can be obtained without recourse to summing weighted squared distances d.

As York (1968) suggests, the equation of the line of best fit may be obtained by minimizing (A6).

Now the variables in S are a and b and S will be an optimum when the partial derivatives $\frac{\partial S}{\partial a}$ and $\frac{\partial S}{\partial b}$ are equated to zero, and

$$\begin{aligned}\frac{\partial S}{\partial a} &= \sum \left\{ W_k \frac{\partial}{\partial a}\left(a + bX_k - Y_k \right)^2 + \left(a + bX_k - Y_k \right)^2 \frac{\partial}{\partial a} W_k \right\} \\ \frac{\partial S}{\partial b} &= \sum \left\{ W_k \frac{\partial}{\partial b}\left(a + bX_k - Y_k \right)^2 + \left(a + bX_k - Y_k \right)^2 \frac{\partial}{\partial b} W_k \right\}\end{aligned}$$

with partial derivatives

$$\begin{aligned}\frac{\partial}{\partial a}\left(a + bX_k - Y_k \right)^2 &= 2\left(a + bX_k - Y_k \right), \\ \frac{\partial}{\partial b}\left(a + bX_k - Y_k \right)^2 &= 2X_k \left(a + bX_k - Y_k \right), \\ \frac{\partial}{\partial a} W_k &= 0, \\ \frac{\partial}{\partial b} W_k &= \frac{-2W_k^2 \left(b^2 w\left(Y_k\right) + w\left(X_k\right) - 2br_k\alpha_k \right)}{\alpha_k^2}\end{aligned}$$

After some algebra we obtain

$$\begin{aligned}\frac{\partial S}{\partial a} &= 2\sum W_k \left(a + bX_k - Y_k \right) \\ \frac{\partial S}{\partial b} &= 2\sum W_k X_k \left(a + bX_k - Y_k \right) + 2\sum W_k^2 \frac{r_k}{\alpha_k}\left(a + bX_k - Y_k \right)^2 - 2\sum W_k^2 \frac{b}{w\left(X_k\right)}\left(a + bX_k - Y_k \right)^2\end{aligned}$$

and setting the partial derivatives $\frac{\partial S}{\partial a}$ and $\frac{\partial S}{\partial b}$ equal to zero

$$\sum W_k \left(a + bX_k - Y_k \right) = 0 \tag{A7}$$

$$\sum W_k X_k \left(a + bX_k - Y_k \right) + \sum W_k^2 \frac{r_k}{\alpha_k}\left(a + bX_k - Y_k \right)^2 - \sum W_k^2 \frac{b}{w\left(X_k\right)}\left(a + bX_k - Y_k \right)^2 = 0 \tag{A8}$$

Now with (33) and (34) we have $a + bX_k - Y_k = bU_k - V_k$ and $X_k = U_k + \bar{X}$, and using these results in (A7) and (A8) gives

$$\sum W_k \left(bU_k - V_k \right) = 0 \tag{A9}$$

$$\sum W_k \left(U_k + \bar{X} \right)\left(bU_k - V_k \right) + \sum W_k^2 \frac{r_k}{\alpha_k}\left(bU_k - V_k \right)^2 - \sum W_k^2 \frac{b}{w\left(X_k \right)}\left(bU_k - V_k \right)^2 = 0 \tag{A10}$$

The first term of (A10) can be expanded as $\sum W_k U_k \left(bU_k - V_k \right) + \bar{X} \sum W_k \left(bU_k - V_k \right)$ and by virtue of (A9) becomes $\sum W_k U_k \left(bU_k - V_k \right)$. Substituting this result into (A10) gives

$$\sum W_k U_k \left(bU_k - V_k \right) + \sum W_k^2 \frac{r_k}{\alpha_k}\left(bU_k - V_k \right)^2 - \sum W_k^2 \frac{b}{w\left(X_k \right)}\left(bU_k - V_k \right)^2 = 0 \tag{A11}$$

(A11) is our (37) and from this we have demonstrated that the *Generalized Least Squares Cubic* for b can be obtained.

APPENDIX B: Derivation of the Variances of gradient b and y-intercept a.

York (1968) and York et al. (2004) give expressions for the variances of the gradient b and y-intercept a that are obtained from the equations

$$s^2\left(b\right) = \left\{\sum\left[\left(\frac{\partial\varphi}{\partial X_k}\right)^2 \frac{1}{w\left(X_k\right)} + \left(\frac{\partial\varphi}{\partial Y_k}\right)^2 \frac{1}{w\left(Y_k\right)} + \frac{2r_k}{\alpha_k}\frac{\partial\varphi}{\partial X_k}\frac{\partial\varphi}{\partial Y_k}\right]\right\} \Big/ \left(\frac{\partial\varphi}{\partial b}\right)^2 \tag{*}$$

where φ is the left-hand-side of (40) – that is the quadratic equation for b – and

$$s^2\left(a\right) = \sum\left\{\frac{1}{w\left(X_k\right)}\left(\frac{\partial a}{\partial X_k}\right)^2 + \frac{1}{w\left(Y_k\right)}\left(\frac{\partial a}{\partial Y_k}\right)^2 + \frac{2r_k}{\alpha_k}\frac{\partial a}{\partial X_k}\frac{\partial a}{\partial Y_k}\right\} \tag{**}$$

(**) is equivalent to our (53) but (*) is only equivalent to our (52) provided that $\frac{\partial b}{\partial X_k} = \frac{\partial\varphi}{\partial X_k}\frac{\partial b}{\partial\varphi} = \frac{\partial\varphi}{\partial X_k}\Big/\frac{\partial\varphi}{\partial b}$ and $\frac{\partial b}{\partial Y_k} = \frac{\partial\varphi}{\partial Y_k}\frac{\partial b}{\partial\varphi} = \frac{\partial\varphi}{\partial Y_k}\Big/\frac{\partial\varphi}{\partial b}$. But, in general, $\frac{\partial b}{\partial\varphi} \neq 1\Big/\frac{\partial\varphi}{\partial b}$ unless the same variables are being held fixed in the partial derivatives. Because of this observation the following analysis is different from York's presentation.

Partial derivatives $\frac{\partial b}{\partial X_k}, \frac{\partial b}{\partial Y_k}$

To obtain an expression for the partial derivative $\frac{\partial b}{\partial X_k}$ we differentiate (40) with respect to X_k keeping $\mathbf{Y} = \left[Y_1 \quad Y_2 \quad \cdots \quad Y_n\right]$ and X_j for $j \neq k$ fixed, giving

$$\begin{aligned}
&2b\frac{\partial b}{\partial X_k}\sum_j W_j^2\left(\frac{V_k}{w\left(X_j\right)} - \frac{U_j r_j}{\alpha_j}\right)U_j + 2b^2\sum_j W_j\frac{\partial W_j}{\partial X_k}\left(\frac{V_j}{w\left(X_j\right)} - \frac{U_j r_j}{\alpha_j}\right)U_j \\
&+ b^2\sum_j W_j^2\left\{\frac{1}{w\left(X_j\right)}\left(U_j\frac{\partial V_j}{\partial X_k} + V_j\frac{\partial U_j}{\partial X_k}\right) - \frac{2U_j r_j}{\alpha_j}\frac{\partial U_j}{\partial X_k}\right\} \\
&+ \frac{\partial b}{\partial X_k}\sum_j W_j^2\left(\frac{U_j^2}{w\left(Y_j\right)} - \frac{V_j^2}{w\left(X_j\right)}\right) + 2b\sum_j W_j\frac{\partial W_j}{\partial X_k}\left(\frac{U_j^2}{w\left(Y_j\right)} - \frac{V_j^2}{w\left(X_j\right)}\right) \\
&+ 2b\sum_j W_j^2\left(\frac{U_j}{w\left(Y_j\right)}\frac{\partial U_j}{\partial X_k} - \frac{V_j}{w\left(X_j\right)}\frac{\partial V_j}{\partial X_k}\right) - 2\sum_j W_j\frac{\partial W_j}{\partial X_k}\left(\frac{U_j}{w\left(Y_j\right)} - \frac{V_j r_j}{\alpha_j}\right)V_j \\
&-\sum_j W_j^2\left\{\frac{1}{w\left(Y_j\right)}\left(U_j\frac{\partial V_j}{\partial X_k} + V_j\frac{\partial U_j}{\partial X_k}\right) - \frac{2V_j r_j}{\alpha_j}\frac{\partial V_j}{\partial X_k}\right\} = 0
\end{aligned} \tag{B1}$$

Differentiating (28) and defining $P_j = \frac{b}{w\left(X_j\right)} - \frac{r_j}{\alpha_j}$ gives

$$\begin{aligned}
\frac{\partial W_j}{\partial X_k} &= -\frac{\alpha_j^2}{\left[b^2 w\left(Y_j\right) + w\left(X_j\right) - 2br_j\alpha_j\right]^2}\left(2bw\left(Y_j\right)\frac{\partial b}{\partial X_k} - 2r_j\alpha_j\frac{\partial b}{\partial X_k}\right) = -2\frac{W_j^2}{\alpha_j^2}\left(bw\left(Y_j\right) - r_j\alpha_j\right)\frac{\partial b}{\partial X_k} \\
&= -2W_j^2\left(\frac{b}{w\left(X_j\right)} - \frac{r_j}{\alpha_j}\right)\frac{\partial b}{\partial X_k} = -2W_j^2 P_j\frac{\partial b}{\partial X_k}
\end{aligned} \tag{B2}$$

So, remembering that $U_k = X_k - \bar{X} = X_k - \dfrac{\sum W_k X_k}{\sum W_k}$

$$\frac{\partial U_j}{\partial X_k} = \delta_{kj} - \frac{W_k}{\sum W_k} - \frac{\sum_j U_j \frac{\partial W_j}{\partial X_k}}{\sum W_k} = \delta_{kj} - \frac{W_k}{\sum W_k} + \frac{2}{\sum W_k}\frac{\partial b}{\partial X_k}\sum_j W_j^2 P_j U_j \tag{B3}$$

where δ_{kj} is the *Kronecker delta* $= \begin{cases} 0 \text{ if } k \neq j \\ 1 \text{ if } k = j \end{cases}$

and similarly

$$\frac{\partial V_j}{\partial X_k} = -\frac{\sum_j V_j \frac{\partial W_j}{\partial X_k}}{\sum W_k} = \frac{2}{\sum W_k}\frac{\partial b}{\partial X_k}\sum_j W_j^2 P_j V_j \tag{B4}$$

And so (B1) with (B2), (B3) and (B4) becomes

$$D\frac{\partial b}{\partial X_k} = -\sum_j W_j^2 \left(\frac{b^2 V_j}{w(X_j)} - \frac{2b^2 U_j r_j}{\alpha_j} + \frac{2bU_j}{w(Y_j)} - \frac{V_j}{w(Y_j)} \right)\left(\delta_{jk} - \frac{W_k}{\sum W_k} \right) \tag{B5}$$

where

$$\begin{aligned} D = {} & 2b\sum_j W_k^2 \left(\frac{V_j}{w(X_j)} - \frac{r_j U_j}{a_j} \right) U_j - 4b^2 \sum_j W_j^3 P_j \left(\frac{V_j}{w(X_j)} - \frac{r_j U_j}{a_j} \right) U_j \\ & + \frac{2b^2}{\sum W_k} \sum_j W_j^2 \frac{U_j}{w(X_j)} \sum_j W_j^2 P_j V_j + \frac{2b^2}{\sum W_k} \sum_j W_j^2 \frac{V_j}{w(X_j)} \sum_j W_j^2 P_j U_j \\ & - \frac{4b^2}{\sum W_k} \sum_j W_j^2 \frac{U_j r_j}{w(X_j)\alpha_j} \sum_j W_j^2 P_j U_j + \sum_j W_j^2 \left(\frac{U_j^2}{w(Y_j)} - \frac{V_j^2}{w(X_j)} \right) \\ & - 4b \sum_j W_j^3 P_j \left(\frac{U_j^2}{w(Y_j)} - \frac{V_j^2}{w(X_j)} \right) + \frac{4b}{\sum W_k} \sum_j W_j^2 \frac{U_j}{w(Y_j)} \sum_j W_j^2 P_j U_j \\ & - \frac{4b}{\sum W_k} \sum_j W_j^2 \frac{V_j}{w(X_j)} \sum_j W_j^2 P_j V_j + 4 \sum_j W_j^3 P_j \left(\frac{U_j}{w(Y_j)} - \frac{r_j V_j}{a_j} \right) V_j \\ & - \frac{2}{\sum W_k} \sum_j W_j^2 \frac{U_j}{w(Y_j)} \sum_j W_j^2 P_j V_j - \frac{2}{\sum W_k} \sum_j W_j^2 \frac{V_j}{w(Y_j)} \sum_j W_j^2 P_j U_j \\ & + \frac{4}{\sum W_k} \sum_j W_j^2 \frac{V_j r_j}{\alpha_j} \sum_j W_j^2 P_j V_j \end{aligned} \tag{B6}$$

Now (B5) can be written as

$$\frac{\partial b}{\partial X_k} = \frac{N(X_k)}{D} \tag{B7}$$

where

$$N\left(X_k\right) = -\sum_j W_j^2 Q_j \left(\delta_{jk} - \frac{W_k}{\sum_k W_k}\right) \tag{B8}$$

and

$$Q_j = \frac{b^2 V_j}{w\left(X_j\right)} - \frac{2b^2 U_j r_j}{\alpha_j} + \frac{2bU_j}{w\left(Y_j\right)} - \frac{V_j}{w\left(Y_j\right)} \tag{B9}$$

The right-hand-side of (B8) is equal to

$$-\left\{W_k^2 Q_k \left(1 - \frac{W_k}{\sum W_k}\right) + \sum_{j \neq k} W_j^2 Q_j \left(0 - \frac{W_j}{\sum W_k}\right)\right\} = -W_k \left\{W_k Q_k - \frac{1}{\sum W_k} \sum_j W_j W_j Q_j\right\}$$

while (B9) may be simplified as follows

$$\begin{aligned}
Q_j &= \frac{b^2 V_j}{w\left(X_j\right)} - \frac{2b^2 U_j r_j}{\alpha_j} + \frac{2bU_j}{w\left(Y_j\right)} - \frac{V_j}{w\left(Y_j\right)} \\
&= 2b\left(\frac{U_j}{w\left(Y_j\right)} + \frac{bV_j}{w\left(X_j\right)} - \frac{r_j}{\alpha_j}\left(bU_j + V_j\right)\right) - V_j\left(\frac{b^2}{w\left(X_j\right)} - \frac{2br_j}{\alpha_j} - \frac{1}{w\left(Y_j\right)}\right) \\
&= 2b\frac{\beta_j}{W_j} - \frac{V_j}{\alpha_j}\left[b^2 w\left(Y_j\right) - 2br_j\alpha_j + w\left(X_j\right)\right] \\
&= 2b\frac{\beta_j}{W_j} - \frac{V_j}{W_j}
\end{aligned}$$

so that

$$W_j Q_j = 2b\beta_j - V_j$$

Using these results and noting that $\sum W_j V_j = 0$ we may write (B8) as

$$N\left(X_k\right) = -W_k\left\{2b\beta_k - V_k - \frac{1}{\sum W_k}\sum_j W_j\left(2b\beta_j - V_j\right)\right\} = W_k\left\{V_k - 2b\left(\beta_k - \overline{\beta}\right)\right\} \tag{B10}$$

where

$$\overline{\beta} = \frac{\sum W_k \beta_k}{\sum W_k} \tag{B11}$$

Similarly

$$\frac{\partial b}{\partial Y_k} = \frac{N\left(Y_k\right)}{D} \tag{B12}$$

where

$$N\left(Y_k\right) = W_k\left\{2\left(\beta_k - \overline{\beta}\right) - U_k\right\} \tag{B13}$$

The variance of the gradient b

Substituting (B7) and (B12) into (52) gives the estimate of the variance of b as

$$s^2\left(b\right) = \sum\left\{\frac{1}{w\left(X_k\right)}\left(\frac{N\left(X_k\right)}{D}\right)^2 + 2\left(\frac{N\left(X_k\right)}{D}\right)\left(\frac{N\left(Y_k\right)}{D}\right)\frac{r_k}{\alpha_k} + \frac{1}{w\left(Y_k\right)}\left(\frac{N\left(Y_k\right)}{D}\right)^2\right\}$$

and substituting (B10) and (B13) and re-arranging gives

$$\begin{aligned} s^2\left(b\right) &= \frac{1}{D^2}\sum W_k^2\left\{\frac{U_k^2}{w\left(Y_k\right)} - 2\frac{r_k}{\alpha_k}U_kV_k + \frac{V_k^2}{w\left(X_k\right)}\right\} \\ &+\frac{1}{D^2}\sum W_k^2\left\{\begin{array}{l}\frac{1}{w\left(X_k\right)}\left[4b^2\left(\beta_k-\overline{\beta}\right)^2 - 4b\left(\beta_k-\overline{\beta}\right)V_k\right] \\ +\left[4b\left(\beta_k-\overline{\beta}\right)U_k - 8b\left(\beta_k-\overline{\beta}\right)^2 + 4\left(\beta_k-\overline{\beta}\right)V_k\right]\frac{r_k}{\alpha_k} \\ +\frac{1}{w\left(Y_k\right)}\left[4\left(\beta_k-\overline{\beta}\right)^2 - 4\left(\beta_k-\overline{\beta}\right)U_k\right]\end{array}\right\}\end{aligned} \tag{B14}$$

Part of (B14) may be simplified as follows

$$\begin{aligned} &\frac{1}{D^2}\sum W^2\left\{\begin{array}{l}\frac{1}{w\left(X\right)}\left[4b^2\left(\beta-\overline{\beta}\right)^2 - 4b\left(\beta-\overline{\beta}\right)V\right] \\ +\left[4b\left(\beta-\overline{\beta}\right)U - 8b\left(\beta-\overline{\beta}\right)^2 + 4\left(\beta-\overline{\beta}\right)V\right]\frac{r}{\alpha} \\ +\frac{1}{w\left(Y\right)}\left[4\left(\beta-\overline{\beta}\right)^2 - 4\left(\beta-\overline{\beta}\right)U\right]\end{array}\right\} \\ &= \frac{1}{D^2}\sum W^2\left\{\begin{array}{l}4\left(\beta-\overline{\beta}\right)^2\left[\frac{b^2}{w\left(X\right)} + \frac{1}{w\left(Y\right)} - \frac{2br}{\alpha}\right] \\ -4\left(\beta-\overline{\beta}\right)\left[\frac{U}{w\left(Y\right)} + \frac{bV}{w\left(X\right)} - \frac{r}{\alpha}\left(bU+V\right)\right]\end{array}\right\} \\ &= \frac{1}{D^2}\sum W^2\left\{4\left(\beta-\overline{\beta}\right)^2\left[\frac{1}{W}\right] - 4\left(\beta-\overline{\beta}\right)\left[\frac{\beta}{W}\right]\right\} \\ &= \frac{1}{D^2}\sum W\left\{4\left(\beta-\overline{\beta}\right)^2 - 4\left(\beta-\overline{\beta}\right)\beta\right\} \\ &= \frac{1}{D^2}\sum W\left\{4\left(\beta-\overline{\beta}\right)^2 - 4\left(\beta-\overline{\beta}\right)\beta\right\} \\ &= \frac{1}{D^2}\sum 4W\left(\beta-\overline{\beta}\right)\left(\beta-\overline{\beta}-\beta\right) \\ &= -\frac{4\overline{\beta}}{D^2}\sum_k W_k\left(\beta_k-\overline{\beta}\right) = 0\end{aligned}$$

since $\sum W_k\left(\beta_k-\overline{\beta}\right) = 0$ by the definition of $\overline{\beta}$ and (B14) becomes

$$s^2\left(b\right) = \begin{cases}\frac{1}{D^2}\sum W_k^2\left\{\frac{U_k^2}{w\left(Y_k\right)} - 2\frac{r_k}{\alpha_k}U_kV_k + \frac{V_k^2}{w\left(X_k\right)}\right\} & \text{if } b \neq 0 \\ \frac{1}{D^2}\sum w^2\left(Y_k\right)\left\{\frac{U_k^2}{w\left(Y_k\right)} - 2\frac{r_k}{\alpha_k}U_kV_k + \frac{V_k^2}{w\left(X_k\right)}\right\} & \text{if } b = 0\end{cases} \tag{B15}$$

noting in (B15) for the special case $b = 0$, the weight function $W_k = w\left(Y_k\right)$ and special results follow for the centroid $\left(\overline{X},\overline{Y}\right)$, the centroidal coordinates $\left(U_k,V_k\right)$ and D.

(B15) is our (58) that is the estimate of the variance of the gradient b.

Partial derivatives $\dfrac{\partial a}{\partial X_k}, \dfrac{\partial a}{\partial Y_k}$

Expanding (29) and re-arranging gives

$$a\sum W_j = \sum W_j Y_j - b\sum W_j X_j$$

and differentiating with respect to X_k keeping $\mathbf{Y} = \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_n \end{bmatrix}$ and X_j for $j \neq k$ fixed, gives

$$\frac{\partial a}{\partial X_k}\sum_j W_j + a\sum_j \frac{\partial W_j}{\partial X_k} = \sum_j \frac{\partial W_j}{\partial X_k} Y_j - \frac{\partial b}{\partial X_k}\sum_j W_j X_j - b\sum_j \frac{\partial W_j}{\partial X_k} X_j - bW_k$$

that can be re-arranged as

$$\frac{\partial a}{\partial X_k}\sum_j W_j = -\sum_j \left(a + bX_j - Y_j\right)\frac{\partial W_j}{\partial X_k} - \frac{\partial b}{\partial X_k}\sum_j W_j X_j - bW_k$$

Now using (34) and (B2) gives

$$\begin{aligned}\frac{\partial a}{\partial X_k}\sum_j W_j &= -\sum_j \left(bU_j - V_j\right)\left(-2W_j^2 P_j \frac{\partial b}{\partial X_k}\right) - \frac{\partial b}{\partial X_k}\sum_j W_j X_j - bW_k \\ &= \frac{\partial b}{\partial X_k}\left\{-2\sum_j W_j^2 P_j\left(V_j - bU_j\right) - \sum_j W_j X_j\right\} - bW_k \end{aligned} \tag{B16}$$

The term $\sum_j W_j^2 P_j\left(V_j - bU_j\right)$ in (B16) can be simplified with the aid of the definition of P_k and (11), (28) and (36)

$$\begin{aligned}\sum W^2 P\left(V - bU\right) &= \sum W^2\left[\frac{b}{w\left(X\right)} - \frac{r}{\alpha}\right]\left(V - bU\right) \\ &= \frac{1}{b}\sum \frac{W^2}{\alpha^2}\left[b^2 w\left(Y\right) + w\left(X\right) - 2br\alpha + br\alpha - w\left(X\right)\right]\left(V - bU\right) \\ &= \frac{1}{b}\sum \frac{W^2}{\alpha^2}\left[\frac{\alpha^2}{W} + br\alpha - w\left(X\right)\right]\left(V - bU\right) \\ &= \frac{1}{b}\sum W\left(V - bU\right) + \frac{1}{b}\sum \frac{W^2}{\alpha^2}\left[br\alpha - w\left(X\right)\right]\left(V - bU\right) \\ &= \frac{1}{b}\sum \frac{W^2}{\alpha^2}\left[br\alpha - w\left(X\right)\right]\left(V - bU\right) \quad \text{since } \sum W\left(V - bU\right) = 0 \\ &= \frac{1}{b}\sum W^2\left[\frac{1}{w\left(Y\right)} - \frac{br}{\alpha}\right]\left(bU - V\right) \\ &= \sum W^2\left(\frac{U}{w\left(Y\right)} - \frac{V}{bw\left(Y\right)} - \frac{brU}{\alpha} + \frac{rV}{\alpha}\right) \\ &= \sum W^2\left(\frac{U}{w\left(Y\right)} + \frac{bV}{w\left(X\right)} - \frac{r}{\alpha}\left(bU + V\right) - \frac{V}{bw\left(Y\right)} + 2\frac{rV}{\alpha} - \frac{bV}{w\left(X\right)}\right)\end{aligned}$$

Now, from (43) $\dfrac{\beta}{W} = \dfrac{U}{w\left(Y\right)} + \dfrac{bV}{w\left(X\right)} - \dfrac{r}{\alpha}\left(bU + V\right)$ and from (B11) $\overline{\beta}\sum W_k = \sum W_k \beta_k$ so we may write

$$\sum W^2 P\left(V - bU\right) = \sum W^2 \left(\frac{\beta}{W} - \frac{V}{bw\left(Y\right)} + 2\frac{rV}{\alpha} - \frac{bV}{w\left(X\right)}\right)$$
$$= \sum W\beta - \sum W^2 \left(\frac{b}{w\left(X\right)} + \frac{1}{bw\left(Y\right)} - 2\frac{r}{\alpha}\right)V$$
$$= \overline{\beta}\sum W - \sum \frac{W^2}{\alpha^2}\left(bw\left(Y\right) + \frac{w\left(X\right)}{b} - 2r\alpha\right)V$$
$$= \overline{\beta}\sum W - \frac{1}{b}\sum \frac{W^2}{\alpha^2}\left(b^2 w\left(Y\right) + w\left(X\right) - 2br\alpha\right)V$$
$$= \overline{\beta}\sum W - \frac{1}{b}\sum WV$$

Now since $\sum WV = 0$ and provided that $b \neq 0$ we may write $\sum W^2 P\left(V - bU\right) = \overline{\beta}\sum W$ and (B16) becomes

$$\frac{\partial a}{\partial X_k}\sum_j W_j = \frac{\partial b}{\partial X_k}\left\{-2\overline{\beta}\sum_j W_j - \sum_j W_j X_j\right\} - bW_k$$
$$= \frac{\partial b}{\partial X_k}\left\{-2\overline{\beta}\sum_j W_j - \overline{X}\sum_j W_j\right\} - bW_k$$
$$= -\frac{\partial b}{\partial X_k}\left\{\left\{2\overline{\beta} + \overline{X}\right\}\sum_j W_j\right\} - bW_k$$

Therefore, providing that $\sum W \neq 0$

$$\frac{\partial a}{\partial X_k} = -\left\{2\overline{\beta} + \overline{X}\right\}\frac{\partial b}{\partial X_k} - \frac{bW_k}{\sum W_k} = \left\{2\overline{\beta} + \overline{X}\right\}\frac{W_k}{D}\left\{2b\left(\beta_k - \overline{\beta}\right) - V_k\right\} - \frac{bW_k}{\sum W_k} \tag{B17}$$

Likewise, from (29) under the same condition that $\sum W \neq 0$

$$\frac{\partial a}{\partial Y_k} = -\left\{2\overline{\beta} + \overline{X}\right\}\frac{\partial b}{\partial Y_k} + \frac{W_k}{\sum W_k} = \left\{2\overline{\beta} + \overline{X}\right\}\frac{W_k}{D}\left\{U_k - 2\left(\beta_k - \overline{\beta}\right)\right\} + \frac{W_k}{\sum W_k} \tag{B18}$$

The variance of the *y*-intercept a

Substituting (B17) and (B18) into (53) gives the estimate of the variance of a as

$$\begin{aligned} s^2\left(a\right) = \sum\Bigg\{ &\frac{1}{w\left(X_k\right)}\left[\left\{2\overline{\beta} + \overline{X}\right\}\frac{W_k}{D}\left\{2b\left(\beta_k - \overline{\beta}\right) - V_k\right\} - \frac{bW_k}{\sum W_k}\right]^2 \\ &+ \frac{2r_k}{\alpha_k}\left[\left\{2\overline{\beta} + \overline{X}\right\}\frac{W_k}{D}\left\{2b\left(\beta_k - \overline{\beta}\right) - V_k\right\} - \frac{bW_k}{\sum_k W_k}\right]\left[\left\{2\overline{\beta} + \overline{X}\right\}\frac{W_k}{D}\left\{U_k - 2\left(\beta_k - \overline{\beta}\right)\right\} + \frac{W_k}{\sum W_k}\right] \\ &+ \frac{1}{w\left(Y_k\right)}\left[\left\{2\overline{\beta} + \overline{X}\right\}\frac{W_k}{D}\left\{U_k - 2\left(\beta_k - \overline{\beta}\right)\right\} + \frac{W_k}{\sum W_k}\right]^2\Bigg\} \\ = \sum&\left\{A_k + B_k + C_k\right\} \end{aligned} \tag{B19}$$

where

$$
\begin{aligned}
C_k &= \frac{b^2 W_k^2}{w(X_k)\left(\sum W_k\right)^2} - \frac{2r_k b W_k^2}{\alpha_k\left(\sum W_k\right)^2} + \frac{W_k^2}{w(Y_k)\left(\sum W_k\right)^2} \\
&= \frac{W_k^2}{\left(\sum W_k\right)^2}\left\{\frac{b^2}{w(X_k)} + \frac{1}{w(Y_k)} - \frac{2br_k}{\alpha_k}\right\} \\
&= \frac{W_k}{\left(\sum W_k\right)^2}
\end{aligned}
$$

and so

$$
\sum C_k = \frac{1}{\sum W_k} \tag{i}
$$

Further

$$
\begin{aligned}
B_k &= -\frac{2\left(2\overline{\beta} + \overline{X}\right) b W_k^2}{w(X_k) D \sum W_k}\left\{2b\left(\beta_k - \overline{\beta}\right) - V_k\right\} + \frac{2\left(2\overline{\beta} + \overline{X}\right) W_k^2}{w(Y_k) D \sum W_k}\left\{U_k - 2\left(\beta_k - \overline{\beta}\right)\right\} \\
&\quad + \frac{2r_k\left(2\overline{\beta} + \overline{X}\right) W_k^2}{\alpha_k D \sum W_k}\left\{2b\left(\beta_k - \overline{\beta}\right) - V_k\right\} - \frac{2r_k\left(2\overline{\beta} + \overline{X}\right) b W_k^2}{\alpha_k D \sum W_k}\left\{U_k - 2\left(\beta_k - \overline{\beta}\right)\right\} \\
&= \frac{2\left(2\overline{\beta} + \overline{X}\right) W_k^2}{\alpha_k^2 D \sum W_k}\left[-2\left(\beta_k - \overline{\beta}\right)\frac{\alpha_k^2}{W_k} + \frac{\alpha_k^2 \beta_k}{W_k}\right] \\
&= \frac{2\left(2\overline{\beta} + \overline{X}\right) W_k}{D \sum W_k}\left(2\overline{\beta} - \beta_k\right)
\end{aligned}
$$

so that

$$
\sum B_k = \frac{2\left(2\overline{\beta} + \overline{X}\right)}{D}\overline{\beta} \tag{ii}
$$

Finally

$$
\begin{aligned}
A_k &= \frac{\left(2\overline{\beta} + \overline{X}\right)^2 W_k^2}{w(X_k) D^2}\left\{2b\left(\beta_k - \overline{\beta}\right) - V_k\right\}^2 + \frac{\left(2\overline{\beta} + \overline{X}\right)^2 W_k^2}{w(Y_k) D^2}\left\{U_k - 2\left(\beta_k - \overline{\beta}\right)\right\}^2 \\
&\quad + \frac{2r_k\left(2\overline{\beta} + \overline{X}\right)^2 W_k^2}{\alpha_k D}\left\{2b\left(\beta_k - \overline{\beta}\right) - V_k\right\}\left\{U_k - 2\left(\beta_k - \overline{\beta}\right)\right\} \\
&= \frac{\left(2\overline{\beta} + \overline{X}\right)^2 W_k^2}{D^2}\left\{\frac{U_k^2}{w(Y_k)} + \frac{V_k^2}{w(X_k)} - \frac{2r_k U_k V_k}{\alpha_k}\right\} \\
&\quad + \frac{\left(2\overline{\beta} + \overline{X}\right)^2 W_k^2}{D^2}\left\{\begin{aligned}&4\left(\beta_k - \overline{\beta}\right)^2\left[\frac{b^2}{w(X_k)} + \frac{1}{w(Y_k)} - \frac{2br_k}{\alpha_k}\right] \\ &+4\left(\beta_k - \overline{\beta}\right)\left[\frac{U_k}{w(Y_k)} + \frac{bV_k}{w(X_k)} - \frac{r_k}{\alpha_k}\left(bU_k + V_k\right)\right]\end{aligned}\right\}
\end{aligned}
$$

and so

$$\sum A_k = \left(2\overline{\beta} + \overline{X}\right)^2 \underbrace{\frac{1}{D^2}\sum W^2 \left\{\frac{U^2}{w(Y)} + \frac{V^2}{w(X)} - \frac{2rUV}{\alpha}\right\}}_{\text{variance of } b}$$
$$+\left(2\overline{\beta} + \overline{X}\right)^2 \frac{1}{D^2}\sum W^2 \underbrace{\left\{\begin{array}{l} 4\left(\beta - \overline{\beta}\right)^2\left[\frac{b^2}{w(X)} + \frac{1}{w(Y)} - \frac{2br}{\alpha}\right] \\ +4\left(\beta - \overline{\beta}\right)\left[\frac{U}{w(Y)} + \frac{bV}{w(X)} - \frac{r}{\alpha}\left(bU + V\right)\right]\end{array}\right\}}_{=0\text{, see reduction of (B14)}}$$

In this equation, one part is the variance of the gradient b and another has shown to be equal to zero (see the reduction of (B14), hence

$$\sum A_k = \left(2\overline{\beta} + \overline{X}\right)^2 s^2\left(b\right) \tag{iii}$$

Therefore, substituting (i), (ii) and (iii) into (B19) gives

$$s^2\left(a\right) = \begin{cases} \left(2\overline{\beta} + \overline{X}\right)^2 s^2\left(b\right) + \dfrac{2}{D}\left(2\overline{\beta} + \overline{X}\right)\overline{\beta} + \dfrac{1}{\sum W_k} & \text{if } b \neq 0 \\ \left(2\overline{\beta} + \overline{X}\right)^2 s^2\left(b\right) + \dfrac{2}{D}\left(2\overline{\beta} + \overline{X}\right)\overline{\beta} + \dfrac{1}{\sum w\left(Y_k\right)} & \text{if } b = 0 \end{cases} \tag{B20}$$

noting in (B20) for the special case $b = 0$, the weight function $W_k = w\left(Y_k\right)$ and special results follow for the centroid $\left(\overline{X}, \overline{Y}\right)$, the variance $s^2\left(b\right)$, D and $\overline{\beta}$.

(B20) is our (59) that is the estimate of the variance of the y-intercept a

Simplification of the denominator D

D from (B6) can be written as

$$D = A_1 + A_2 + \cdots + A_{12} + A_{13} \tag{B21}$$

Now if $b \neq 0$,

$$\begin{aligned} A_1 + A_6 &= 2b\sum W^2\left(\frac{V}{w(X)} - \frac{rU}{a}\right)U + \sum W^2\left(\frac{U^2}{w(Y)} - \frac{V^2}{w(X)}\right) \\ &= \frac{1}{b}\sum\left[-\frac{W^2 r}{\alpha}\left(bU - V\right)^2 + WUV\right] \end{aligned} \tag{B22}$$

This is because

$$\begin{aligned} &W^2\left\{2b\left(\frac{V}{w(X)} - \frac{rU}{a}\right)U + \frac{U^2}{w(Y)} - \frac{V^2}{w(X)}\right\} \\ &= W^2\left\{\left[\frac{V}{w(X)} + \left(\frac{1}{bw(Y)} - \frac{2r}{\alpha} - \frac{b}{w(X)}\right)U\right]\left(bU - V\right) + \frac{b^2U^2}{w(X)} + \left(\frac{1}{bw(Y)} - \frac{2r}{\alpha}\right)UV\right\} \\ &= \frac{W^2}{b}\left\{\frac{U}{w(Y)} + \frac{bV}{w(X)} - \frac{r}{\alpha}\left(bU - V\right) + \frac{r}{\alpha}\left(V - bU\right) - \frac{b^2U}{w(X)}\right\}\left(bU - V\right) \\ &\quad + \frac{W^2b^2U^2}{w(X)} + W^2\left(\frac{1}{bw(Y)} - \frac{2r}{\alpha}\right)UV \end{aligned}$$

$$
\begin{aligned}
&= \frac{W^2}{b}\left\{\frac{\beta}{W} + \frac{r}{\alpha}\left(V - bU\right) - \frac{b^2 U}{w\left(X\right)}\right\}\left(bU - V\right) + \frac{W^2 b^2 U^2}{w\left(X\right)} + W^2\left(\frac{1}{bw\left(Y\right)} - \frac{2r}{\alpha}\right)UV \\
&= \frac{W\beta}{b}\left(bU - V\right) - \frac{W^2 r}{b\alpha}\left(bU - V\right)^2 + W^2\left(\frac{b}{w\left(X\right)} + \frac{1}{bw\left(Y\right)} - \frac{2r}{\alpha}\right)UV \\
&= \frac{W\beta}{b}\left(bU - V\right) - \frac{W^2 r}{b\alpha}\left(bU - V\right)^2 + \frac{W}{b}UV
\end{aligned}
$$

provided $b \neq 0$, and so, noting that $\sum W\beta\left(bU - V\right) = 0$ from (44)

$$
\sum W^2\left\{2b\left(\frac{V}{w\left(X\right)} - \frac{rU}{a}\right)U + \frac{U^2}{w\left(Y\right)} - \frac{V^2}{w\left(X\right)}\right\} = \frac{1}{b}\sum\left[-\frac{W^2 r}{\alpha}\left(bU - V\right)^2 + WUV\right]
$$

Further

$$
\begin{aligned}
A_2 + A_7 + A_{10} &= -4b^2\sum W^3 P\left(\frac{V}{w\left(X\right)} - \frac{rU}{\alpha}\right)U - 4b\sum W^3 P\left(\frac{U^2}{w\left(Y\right)} - \frac{V^2}{w\left(X\right)}\right) \\
&\quad +4\sum_k W^3 P\left(\frac{U}{w\left(Y\right)} - \frac{rV}{\alpha}\right)V \\
&= 4\sum W\beta\left(\beta - U\right)
\end{aligned}
\tag{B23}
$$

This is because

$$
\begin{aligned}
&-b^2 W^3 P\left(\frac{UV}{w\left(X\right)} - \frac{rU^2}{\alpha}\right) - bW^3 P\left(\frac{U^2}{w\left(Y\right)} - \frac{V^2}{w\left(X\right)}\right) + W^3 P\left(\frac{UV}{w\left(Y\right)} - \frac{rV^2}{\alpha}\right) \\
&= W^3 P\left\{\frac{UV}{w\left(Y\right)} - \frac{bU^2}{w\left(Y\right)} + \frac{bV^2}{w\left(X\right)} - \frac{b^2 UV}{w\left(X\right)} - \frac{r}{\alpha}V^2 + \frac{r}{\alpha}b^2 U^2\right\} \\
&= W^3 P\left\{\frac{U}{w\left(Y\right)}\left(V - bU\right) + \frac{bV}{w\left(X\right)}\left(V - bU\right) - \frac{r}{\alpha}\left(bU + V\right)\left(V - bU\right)\right\} \\
&= W^3 P\frac{\beta}{W}\left(V - bU\right) \\
&= W^2\beta\left(\frac{b}{w\left(X\right)} - \frac{r}{\alpha}\right)\left(V - bU\right) \\
&= W^2\beta\left(\frac{bV}{w\left(X\right)} - \frac{b^2 U}{w\left(X\right)} - \frac{r}{\alpha}\left(V - bU\right)\right) \\
&= W^2\beta\left(\frac{U}{w\left(Y\right)} + \frac{bV}{w\left(X\right)} - \frac{r}{\alpha}\left(bU + V\right) - \frac{U}{w\left(Y\right)} - \frac{b^2 U}{w\left(X\right)} - \frac{r}{\alpha}\left[\left(V - bU\right) - \left(bU + V\right)\right]\right) \\
&= W^2\beta\left(\frac{U}{w\left(Y\right)} + \frac{bV}{w\left(X\right)} - \frac{r}{\alpha}\left(bU + V\right) - U\left[\frac{1}{w\left(Y\right)} + \frac{b^2}{w\left(X\right)} - 2b\frac{r}{\alpha}\right]\right) \\
&= W^2\beta\left(\frac{\beta}{W} - \frac{U}{W}\right) \\
&= W\beta\left(\beta - U\right)
\end{aligned}
$$

Finally

$$
\begin{aligned}
&A_3 + A_4 + A_5 + A_8 + A_9 + A_{11} + A_{12} + A_{13} \\
&= \left(A_3 + A_{11} + A_9 + A_{13}\right) + \left(A_4 + A_{12} + A_8 + A_5\right) \\
&= \left(\begin{aligned} &\frac{2b^2}{\sum W_k}\sum_j W_j^2 \frac{U_j}{w\left(X_j\right)}\sum_j W_j^2 P_j V_j - \frac{2}{\sum W_k}\sum_j W_j^2 \frac{U_j}{w\left(Y_j\right)}\sum_j W_j^2 P_j V_j \\ &-\frac{4b}{\sum W_k}\sum_j W_j^2 \frac{V_j}{w\left(X_j\right)}\sum_j W_j^2 P_j V_j + \frac{4}{\sum W_k}\sum_j W_j^2 \frac{V_j r_j}{\alpha_j}\sum_j W_j^2 P_j V_j \end{aligned}\right) \\
&\quad + \left(\begin{aligned} &\frac{2b^2}{\sum W_k}\sum_j W_j^2 \frac{V_j}{w\left(X_j\right)}\sum_j W_j^2 P_j U_j - \frac{2}{\sum W_k}\sum_j W_j^2 \frac{V_j}{w\left(Y_j\right)}\sum_j W_j^2 P_j U_j \\ &+\frac{4b}{\sum W_k}\sum_j W_j^2 \frac{U_j}{w\left(Y_j\right)}\sum_j W_j^2 P_j U_j - \frac{4b^2}{\sum W_k}\sum_j W_j^2 \frac{U_j r_j}{w\left(X_j\right)\alpha_j}\sum_j W_j^2 P_j U_j \end{aligned}\right) \\
&= -4\bar{\beta}^2 \sum W_k
\end{aligned} \tag{B24}
$$

This is because

$$
\begin{aligned}
&\left(A_3 + A_{11} + A_9 + A_{13}\right) + \left(A_4 + A_{12} + A_8 + A_5\right) \\
&= \frac{2}{\sum W}\sum W^2 PV \sum W^2 \left[\left(\frac{b^2}{w\left(X\right)} - \frac{1}{w\left(Y\right)}\right)U - 2\left(\frac{b}{w\left(X\right)} - \frac{r}{\alpha}\right)V\right] \\
&\quad + \frac{2}{\sum W}\sum W^2 PU \sum W^2 \left[\left(\frac{b^2}{w\left(X\right)} - \frac{1}{w\left(Y\right)}\right)V + 2b\left(\frac{1}{w\left(Y\right)} - \frac{br}{\alpha}\right)U\right]
\end{aligned}
$$

Now with $\frac{1}{W} = \frac{b^2}{w\left(X\right)} + \frac{1}{w\left(Y\right)} - \frac{2br}{\alpha}$ and $P = \frac{b}{w\left(X\right)} - \frac{r}{\alpha}$ we may write $\frac{b^2}{w\left(X\right)} - \frac{1}{w\left(Y\right)} = 2bP - \frac{1}{W}$ and $\frac{1}{w\left(Y\right)} - \frac{br}{\alpha} = \frac{1}{W} - bP$, and

$$
\begin{aligned}
&\left(A_3 + A_{11} + A_9 + A_{13}\right) + \left(A_4 + A_{12} + A_8 + A_5\right) \\
&= \frac{2}{\sum W}\sum W^2 PV \sum W^2 \left(2bPU - \frac{U}{W} - 2PV\right) \\
&\quad + \frac{2}{\sum W}\sum W^2 PU \sum W^2 \left(2bPV - \frac{V}{W} + 2b\frac{U}{W} - 2b^2 PU\right) \\
&= \frac{2}{\sum W}\sum W^2 PV \sum W^2 \left(2bPU - 2PV\right) + \frac{2}{\sum W}\sum W^2 PU \sum W^2 \left(2bPV - 2b^2 PU\right) \\
&\quad - \frac{2}{\sum W}\sum W^2 PV \sum WU - \frac{2}{\sum W}\sum W^2 PU \sum WV + \frac{4b}{\sum W}\sum W^2 PU \sum WU
\end{aligned}
$$

Noting that $\sum WU = \sum WV = 0$ we have

$$
\begin{aligned}
&\left(A_3 + A_{11} + A_9 + A_{13}\right) + \left(A_4 + A_{12} + A_8 + A_5\right) \\
&= \frac{2}{\sum W}\sum W^2PV\sum W^2\left(2bPU - 2PV\right) + \frac{2}{\sum W}\sum W^2PU\sum W^2\left(2bPV - 2b^2PU\right) \\
&= \frac{4}{\sum W}\left\{\sum W^2PV\sum W^2\left(bPU - PV\right) + \sum W^2PU\sum W^2\left(bPV - b^2PU\right)\right\} \\
&= \frac{4}{\sum W}\begin{Bmatrix}\sum W^2PV\sum W^2PbU - \sum W^2PV\sum W^2PV \\ +\sum W^2PU\sum W^2bPV - \sum W^2PU\sum W^2PbbU\end{Bmatrix} \\
&= -\frac{4}{\sum W}\left\{\left(\sum W^2PbU\right)^2 - 2\sum W^2PbU\sum W^2PV + \left(\sum W^2PV\right)^2\right\} \\
&= -\frac{4}{\sum W}\left\{\sum W^2PbU - \sum W^2PV\right\}^2 \\
&= -\frac{4}{\sum W}\left\{\sum W^2P\left(bU - V\right)\right\}^2 \\
&= -\frac{4}{\sum W}\left\{\sum W^2\left(\frac{b}{w\left(X\right)} - \frac{r}{\alpha}\right)\left(bU - V\right)\right\}^2 \\
&= -\frac{4}{\sum W}\left\{\sum W^2\left(\frac{\beta}{W} - \frac{U}{W}\right)\right\}^2 \\
&= -\frac{4}{\sum W}\left\{\sum W\beta - \sum WU\right\}^2 \\
&= -\frac{4}{\sum W}\left\{\sum W\beta\right\}^2 \\
&= -4\overline{\beta}^2\sum W
\end{aligned}
$$

D in (B21) is the summation of the three parts above, (B22), (B23) and (B24)

$$
\begin{aligned}
D &= A_1 + A_2 + \cdots + A_{12} + A_{13} \\
&= \left(A_1 + A_6\right) + \left(A_2 + A_7 + A_{10}\right) + \left[\left(A_3 + A_{11} + A_9 + A_{13}\right) + \left(A_4 + A_{12} + A_8 + A_5\right)\right] \\
&= \frac{1}{b}\sum\left[-\frac{W^2r}{\alpha}\left(bU - V\right)^2 + WUV\right] + 4\sum W\beta\left(\beta - U\right) - 4\overline{\beta}^2\sum W
\end{aligned}
$$

The last two terms (B23) and (B24) can be simplified as follows

$$
\begin{aligned}
4\sum W\beta\left(\beta - U\right) - 4\overline{\beta}^2\sum W &= 4\left\{\sum W\beta\left(\beta - U\right) - \overline{\beta}\sum W\beta\right\} \\
&= 4\left\{\sum W\beta^2 - \sum W\beta U - \sum W\beta\overline{\beta}\right\} \\
&= 4\sum W\left(\beta^2 - \beta U - \beta\overline{\beta}\right) \\
&= 4\sum W\left(\beta - U\right)\left(\beta - \overline{\beta}\right)
\end{aligned}
$$

since $\sum WU = 0$ and D can be written as

$$
D = \frac{1}{b}\sum\left[-\frac{W_k^2r_k}{\alpha_k}\left(bU_k - V_k\right)^2 + W_kU_kV_k\right] + 4\sum W_k\left(\beta_k - U_k\right)\left(\beta_k - \overline{\beta}\right) \quad \text{if } b \neq 0 \qquad \text{(B25a)}
$$

(B25a) is the first member of (56) given in York et. al (2004, eq. 6)

Now if $b = 0$, the terms $A_1, A_2, A_3, A_4, A_5, A_7, A_8$ and A_9 in (B21) are all equal to zero [see also (B6)] and $D = A_6 + A_{10} + A_{11} + A_{12} + A_{13}$. Inspection of the derivation of (B25a) above leads to

$$D = \sum W_k^2 \left(\frac{U_k^2}{w(Y_k)} - \frac{V_k^2}{w(X_k)} \right) + 4\sum W_k \left(\beta_k - U_k\right)\left(\beta_k - \bar{\beta}\right)$$

and

$$D = \sum w^2\left(Y_k\right) \left(\frac{U_k^2}{w(Y_k)} - \frac{V_k^2}{w(X_k)} \right) + 4\sum w\left(Y_k\right)\left(\beta_k - U_k\right)\left(\beta_k - \bar{\beta}\right) \quad \text{if } b = 0 \qquad \text{(B25b)}$$

noting in (B25b) for the special case $b = 0$, the weight function $W_k = w\left(Y_k\right)$ and special results follow for the centroid $\left(\bar{X}, \bar{Y}\right)$, centroidal coordinates $\left(U_k, V_k\right)$ and the factors $\beta_k = U_k - V_k r_k \sqrt{\frac{w(Y_k)}{w(X_k)}}$ and $\bar{\beta} = \frac{\sum w\left(Y_k\right)\beta_k}{\sum w\left(Y_k\right)}$.

(B25b) is the second member of (56).

References

Adcock, R.J. 1877, 'Note on the method of least squares', *The Analyst,* Vol. 4, No. 6, pp. 183-184, Nov., 1877.

Adcock, R.J., 1878a, 'Least squares', *The Analyst,* Vol. 5, No. 1, pp. 21-22, Jan., 1878.

Adcock, R.J., 1878b, 'A problem in least squares', *The Analyst,* Vol. 5, No. 2, pp. 53-54, Mar., 1878.

Cantrell, C.A., 2008, 'Technical Note: Review of methods for linear least-squares fitting of data and application to atmospheric chemistry problems', *Atmospheric Chemistry and Physics,* Vol. 8, pp. 5477-5487.

Deming, W.E., 1943, *Statistical adjustment of data,* John Wiley and Sons, New York (Dover Publications edition, 1985).

Farebrother, R.W., 1999, *Fitting Linear Relationships – A History of the Calculus of observations 1750-1900,* Springer-Verlag, New York.

Finney, D.J., 1996, 'A note on the history of regression', *Journal of Applied Statistics,* Vol. 23, No. 5, pp. 555-558.

Kiryati, N. and Bruckstein, A.M., 1992, 'What's in a Set of Points?', *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 14, No. 4, pp. 496-500, April 1992.

Kummell, C.H., 1879, 'Reduction of observation equations which contain more than one observed quantity', *The Analyst,* Vol. 6, No. 4, pp. 97-105, Jul., 1879.

Krane, K.S. and Schecter, L., 1981, 'Regression line analysis', *American Journal of Physics,* Vol. 50, No. 1, pp. 82-84, January 1982.

Krystek, M. and Anton, M., 2007, 'A weighted total least-squares algorithm for fitting a straight line', *Measurement Science and Technology,* Vol. 18, pp.3438-3442

Lagrange, Joseph-Louis, 1788, *Méchanique Analytique,* 2 Vols, Paris.

https://www.irphe.fr/~clanet/otherpaperfile/articles/Lagrange/N0029071_PDF_1_530.pdf, accessed 27-Mar-2017

Lybanon, M., 1984, 'A better least-squares method when both variables have uncertainties', *American Journal of Physics,* Vol. 52, No. 1, pp. 22-26, January 1984.

Macdonald, J.R. and Thompson, W.J., 1992, 'Least-squares fitting when both variables contain errors: Pitfalls and possibilities', *American Journal of Physics,* Vol. 60, No. 1, pp. 66-73, January 1992.

McIntyre, G.A., Brooks, C., Compston, W. and Turek, A., 1966, 'The statistical assessment of Rb-Sr isochrones', *Journal of Geophysical Research,* Vol. 71, No. 22, pp. 5459-5468, November 15, 1966.

Mikhail, E.M., 1976, *Observations and least Squares,* IEP—A Dun-Donnelley, New York.

Neri, F., Saitta, G. and Chiofalo, S., 1989, 'An accurate and straightforward approach to line regression analysis of error-affected experimental data', *Journal of Physics E: Scientific Instruments,* Vol. 22, No. 4, pp. 215-217.

Pearson, K., 1901, 'On lines and planes of closest fit to systems of points in space', *Philosophical Magazine,* Vol. 2, pp. 559-572.

Reed, C.B., 1989, 'Linear least-squares fits with errors in both coordinates', *American Journal of Physics,* Vol. 57, No. 7, pp. 642-646, July 1989.

Williamson, J.H., 1968, 'Least-squares fitting of a straight line', *Canadian Journal of Physics,* Vol. 46, pp. 1845-1847

York, D., 1966, 'Least-squares fitting of a straight line', *Canadian Journal of Physics,* Vol. 44, pp. 1079-1086, 1966.

York, D., 1968, 'Least squares fitting of a straight line with correlated errors', *Earth and Planetary Science Letters (1969),* pp. 320-324, North-Holland Publishing Company, Amsterdam.

York, D., Evensen, N.M., Lopez Martinez, M. and De Basabe Delgado, J., 2004, 'Unified equations for the slope, intercept, and standard errors of the best straight line', *American Journal of Physics,* Vol. 72, No. 3, pp. 367-375.

York, Derek

Derek York, geophysics professor, science writer (b at Normanton, Yorkshire, Eng 12 Aug 1936; d at Toronto, 9 Aug 2007). A leader in the field of potassium-argon dating of rock, Derek York was a foreign principal investigator for NASA during the Apollo missions to the moon. He came to the University of Toronto in 1960 from Oxford, where he had earned a doctorate of philosophy in earth sciences. His first book, *The Earth's Age and Geochronology*, written with Ronald M. Farquhar (1972), was translated into Chinese, and a subsequent book, Planet Earth (1976), into Japanese and Italian. In 1985 he won the past president's medal of the Geological Association of Canada for research in earth sciences and was elected a fellow of the Royal Society of Canada. He was chair of the physics department at the University of Toronto from 1992-1997, and was J. Tuzo Wilson Professor of Geophysics from 1995-2000.

He wrote science articles for the *Globe and Mail* from 1980-1992 and in 1986 won the Royal Society's Bancroft award for his contribution to public understanding of the earth sciences. In 1996 York was awarded the Sandford Fleming Medal of the Royal Canadian Institute. His last book, *In Search of Lost Time*, was published in 1997.

Canadian Encyclopaedia, article by Judith Knelman published 02-Oct-2009