Geometric Constructions Relating Various Lines of Regression

Brian J. McCartin†

Applied Mathematics, Kettering University
1700 University Avenue, Flint, MI 48504-6214 USA

Copyright © 2016 Brian J. McCartin. This article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

As a direct consequence of the Galton-Pearson-McCartin Theorem [10, Theorem 2], the concentration ellipse provides a unifying thread to the Euclidean construction of various lines of regression. These include lines of coordinate regression [7], orthogonal regression [13], \( \lambda \)-regression [8] and \((\lambda, \mu)\)-regression [9] whose geometric constructions are afforded a unified treatment in the present paper.

Mathematics Subject Classification: 51M15, 51N20, 62J05

Keywords: geometric construction, linear regression, concentration ellipse, Galton-Pearson-McCartin Theorem

1 Introduction

In fitting a linear relationship [6] to a set of \((x, y)\) data (Figure 1), it is many times assumed that one (‘independent’) variable, say \(x\), is known exactly while the other (‘dependent’) variable, say \(y\), is subject to error. The line \(L\) (Figure 2) is then chosen to minimize the total square vertical deviation \(\Sigma d_y^2\). This is known as the line of regression of \(y\) on \(x\). Reversing the roles of \(x\) and \(y\) minimizes instead the total square horizontal deviation \(\Sigma d_x^2\), thereby yielding the line of regression of \(x\) on \(y\). Either method may be termed coordinate regression.

†Deceased January 29, 2016
This procedure may be generalized to the situation where both variables are subject to independent errors with zero means, albeit with equal variances. In this case, \( L \) is chosen to minimize the total square orthogonal deviation \( \Sigma d^2_\perp \) (Figure 2). This is referred to as orthogonal regression.

This approach may be further extended to embrace the case where the error variances are unequal. Defining \( \sigma^2_u = \text{variance of x-error} \), \( \sigma^2_v = \text{variance of y-error} \) and \( \lambda = \sigma^2_v/\sigma^2_u \), a weighted least squares procedure [1, 16] is employed whereby \( \frac{1+m^2}{\lambda+m^2} \cdot \Sigma d^2_\perp \) is minimized in order to determine \( L \). This is referred to as \( \lambda \)-regression. Finally, if the errors are correlated with \( p_{uv} = \text{covariance of errors} \) and \( \mu = p_{uv}/\sigma^2_u \) then \( \frac{1+m^2}{\lambda-2\mu m+m^2} \cdot \Sigma d^2_\perp \) is minimized in order to determine \( L \). This is referred to as \( (\lambda, \mu) \)-regression.

In the next section, the concept of oblique linear least squares approximation is introduced (Figure 3). In this mode of approximation, one first selects a direction specified by the slope \( s \). One then obliquely projects each data point in this direction onto the line \( L \). Finally, \( L \) is then chosen so as to minimize the total square oblique deviation \( \Sigma d^2 \). It is known that this permits a unified treatment of all of the previously described flavors of linear regression [10].

Galton [7] first provided a geometric interpretation of coordinate regression in terms of the best-fitting concentration ellipse which has the same first moments and second moments about the centroid as the experimental data [3]. Pearson [13] then extended this geometric characterization to orthogonal regression. McCartin [8] further extended this geometric characterization to
λ-regression. Finally, McCartin [9] succeeded in extending this geometric characterization to the general case of $(\lambda, \mu)$-regression. In a similar vein, a geometric characterization of oblique linear least squares approximation [10] further generalizes these Galton-Pearson-McCartin geometric characterizations of linear regression. This result is known as the Galton-Pearson-McCartin Theorem and is reviewed below (Theorem 1).

The primary purpose of the present paper is to utilize the Galton-Pearson-McCartin Theorem to provide a unified approach to the Euclidean construction of these various lines of regression. As such, the concentration ellipse will serve as the adhesive which binds together these varied geometric constructions. However, before this can be accomplished, the basics of oblique least squares, linear regression and the concentration ellipse must be reviewed.

2 Oblique Least Squares

Consider the experimentally ‘observed’ data \( \{(x_i, y_i)\}_{i=1}^{n} \) where \( x_i = X_i + u_i \) and \( y_i = Y_i + v_i \). Here \((X_i,Y_i)\) denote theoretically exact values with corresponding random errors \((u_i, v_i)\). It is assumed that \( E(u_i) = 0 = E(v_i) \), that successive observations are independent, and that \( Var(u_i) = \sigma_u^2, Var(v_i) = \sigma_v^2, Cov(u_i, v_i) = p_{uv} \) irrespective of \( i \).

Next, define the statistics of the sample data corresponding to the above population statistics. The mean values are given by

\[
\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i; \quad \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i,
\]
the sample variances are given by
\[ \sigma^2_x = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2; \quad \sigma^2_y = \frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^2, \] (2)
the sample covariance is given by
\[ p_{xy} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}) \cdot (y_i - \bar{y}), \] (3)
and (assuming that \( \sigma_x \cdot \sigma_y \neq 0 \)) the sample correlation coefficient is given by
\[ r_{xy} = \frac{p_{xy}}{\sigma_x \cdot \sigma_y}. \] (4)

Note that if \( \sigma^2_x = 0 \) then the data lie along the vertical line \( x = \bar{x} \) while if \( \sigma^2_y = 0 \) then they lie along the horizontal line \( y = \bar{y} \). Hence, we assume without loss of generality that \( \sigma^2_x \cdot \sigma^2_y \neq 0 \) so that \( r_{xy} \) is always well defined. Furthermore, by the Cauchy-Buniakovskiy-Schwarz inequality,
\[ p^2_{xy} \leq \sigma^2_x \cdot \sigma^2_y \quad (\text{i.e.} \quad -1 \leq r_{xy} \leq 1) \] (5)
with equality if and only if \( (y_i - \bar{y}) \propto (x_i - \bar{x}) \), in which case the data lie on the line
\[ y - \bar{y} = \frac{p_{xy}}{\sigma_x^2} (x - \bar{x}) = \frac{\sigma^2_y}{p_{xy}} (x - \bar{x}), \] (6)
since \( p_{xy} \neq 0 \) in this instance. Thus, we may also restrict \(-1 < r_{xy} < 1\).

Turning our attention to Figure 3, the oblique linear least squares approximation, \( L \), corresponding to the specified slope \( s \) may be expressed as
\[ L : y - \bar{y} = m(x - \bar{x}), \] (7)
since it is straightforward to show that it passes through the centroid of the data, \((\bar{x}, \bar{y})\) [10]. \(L\) is to be selected so as to minimize the total square oblique deviation

\[
S := \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} [(x_i - x_*)^2 + (y_i - y_*)^2].
\] (8)

Furthermore, this optimization problem may be reduced [10] to choosing \(m\) to minimize

\[
S(m) = \frac{1 + s^2}{(m-s)^2} \cdot \sum_{i=1}^{n} [m(x_i - \bar{x}) - (y_i - \bar{y})]^2,
\] (9)

thereby producing the extremizing slope [10]

\[
m = \frac{\sigma_y^2 - s \cdot p_{xy}}{p_{xy} - s \cdot \sigma_x^2}.
\] (10)

Equations (7) and (10) together comprise a complete solution to the problem of oblique linear least squares approximation.

### 3 Linear Regression

Some of the implications of oblique linear least squares approximation for linear regression are next considered, beginning with the case of coordinate regression. Setting \(s = 0\) in Equation (10) produces regression of \(x\) on \(y\):

\[
m_x^+ := \frac{\sigma_y^2}{p_{xy}}, \quad m_x^- := 0.
\] (11)

Likewise, letting \(s \to \infty\) in Equation (10) produces regression of \(y\) on \(x\):

\[
m_y^+ := \frac{p_{xy}}{\sigma_x^2}, \quad m_y^- := \infty.
\] (12)

Unsurprisingly, we see that the coordinate regression lines [7] are special cases of oblique regression.

The more substantial case of orthogonal regression is now considered. Setting \(s = -\frac{1}{m}\) in Equation (10) produces

\[
p_{xy} \cdot m^2 - (\sigma_y^2 - \sigma_x^2) \cdot m - p_{xy} = 0,
\] (13)

whose roots provide both the best orthogonal regression line

\[
m_\perp^+ = \frac{(\sigma_y^2 - \sigma_x^2) + \sqrt{(\sigma_y^2 - \sigma_x^2)^2 + 4p_{xy}^2}}{2p_{xy}},
\] (14)
as well as the worst orthogonal regression line

\[ m_{\perp} = \frac{(\sigma_y^2 - \sigma_x^2) - \sqrt{(\sigma_y^2 - \sigma_x^2)^2 + 4p_{xy}^2}}{2p_{xy}}. \]  

(15)

Thus, orthogonal regression [13] is also a special case of oblique regression.

Turning now to the case of \( \lambda \)-regression, set \( s = -\frac{\lambda}{m} \) in Equation (10):

\[ p_{xy} \cdot m^2 - (\sigma_y^2 - \lambda \sigma_x^2) \cdot m - \lambda p_{xy} = 0, \]  

(16)

whose roots provide both the best \( \lambda \)-regression line

\[ m^+_\lambda = \frac{(\sigma_y^2 - \lambda \sigma_x^2) + \sqrt{(\sigma_y^2 - \lambda \sigma_x^2)^2 + 4\lambda p_{xy}^2}}{2p_{xy}}, \]  

as well as the worst \( \lambda \)-regression line

\[ m^-_\lambda = \frac{(\sigma_y^2 - \lambda \sigma_x^2) - \sqrt{(\sigma_y^2 - \lambda \sigma_x^2)^2 + 4\lambda p_{xy}^2}}{2p_{xy}}. \]  

(18)

Thus, \( \lambda \)-regression [8] is also a special case of oblique regression.

Proceeding to the case of \( (\lambda, \mu) \)-regression, set \( s = \frac{\mu m - \lambda}{m - \mu} \) in Equation (10):

\[ (p_{xy} - \mu \sigma_x^2) \cdot m^2 - (\sigma_y^2 - \lambda \sigma_x^2) \cdot m - (\lambda p_{xy} - \mu \sigma_y^2) = 0, \]  

(19)

whose roots provide both the best \( (\lambda, \mu) \)-regression line

\[ m^+_\lambda,\mu = \frac{(\sigma_y^2 - \lambda \sigma_x^2) + \sqrt{(\sigma_y^2 - \lambda \sigma_x^2)^2 + 4(p_{xy} - \mu \sigma_x^2)(\lambda p_{xy} - \mu \sigma_y^2)}}{2(p_{xy} - \mu \sigma_x^2)}, \]  

as well as the worst \( (\lambda, \mu) \)-regression line

\[ m^-_\lambda,\mu = \frac{(\sigma_y^2 - \lambda \sigma_x^2) - \sqrt{(\sigma_y^2 - \lambda \sigma_x^2)^2 + 4(p_{xy} - \mu \sigma_x^2)(\lambda p_{xy} - \mu \sigma_y^2)}}{2(p_{xy} - \mu \sigma_x^2)}. \]  

(21)

Thus, \( (\lambda, \mu) \)-regression [9] is also a special case of oblique regression.

4 Concentration Ellipse

As established in the previous section, all of the garden varieties of linear regression are subsumed under the umbrella of oblique linear least squares approximation. The generalization of the Galton-Pearson-McCartin geometric characterizations of linear regression to oblique linear least squares approximation is now considered.
Define the concentration ellipse, $E$, (Figure 4) via

$$
\frac{(x - \bar{x})^2}{\sigma_x^2} - 2 \frac{p_{xy}}{\sigma_x \sigma_y} \cdot \frac{(x - \bar{x})}{\sigma_x} \cdot \frac{(y - \bar{y})}{\sigma_y} + \frac{(y - \bar{y})^2}{\sigma_y^2} = 4(1 - \frac{p_{xy}^2}{\sigma_x^2 \sigma_y^2}),
$$

(22)

which has the same centroid and second moments about that centroid as does the data [3, pp. 283-285] (see [10] for proof). In this sense, it is the ellipse which is most representative of the data points without any a priori statistical assumptions concerning their origin. The concentration ellipse is homothetic to the inertia ellipse with ratio of similitude $4n(\sigma_x^2 \sigma_y^2 - p_{xy}^2)$ [8]. In order to achieve our characterization, the concept of conjugate diameters of an ellipse is required [14, p. 146].

A diameter of an ellipse is a line segment passing through the center connecting two antipodal points on its periphery. The conjugate diameter to a given diameter is that diameter which is parallel to the tangent to the ellipse at either peripheral point of the given diameter (Figure 5).

If the ellipse is described by

$$
ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0 \ (h^2 < ab),
$$

(23)

then, according to [10, p. 146], the slope of the conjugate diameter is

$$
s = -\frac{a + h \cdot m}{h + b \cdot m},
$$

(24)
where $m$ is the slope of the given diameter. Consequently,

$$m = -\frac{a + h \cdot s}{h + b \cdot s},$$  \hspace{1cm} (25)$$

thus establishing that conjugacy is a symmetric relation. As a result, these conjugacy conditions may be rewritten symmetrically as

$$b \ m \cdot s + h \ (m + s) + a = 0.$$  \hspace{1cm} (26)$$

For the concentration ellipse, Equation (22), take

$$a = \sigma_y^2, \ h = -p_{xy}, \ b = \sigma_x^2.$$  \hspace{1cm} (27)$$

so that

$$s = -\frac{\sigma_y^2 - p_{xy} \cdot m}{-p_{xy} + \sigma_x^2 \cdot m}; \ m = -\frac{\sigma_y^2 - p_{xy} \cdot s}{-p_{xy} + \sigma_x^2 \cdot s},$$  \hspace{1cm} (28)$$

which implies that $s$ and $m$ satisfy the symmetric conjugacy condition

$$\sigma_x^2 \cdot (m \cdot s) - p_{xy} \cdot (m + s) + \sigma_y^2 = 0.$$  \hspace{1cm} (29)$$

Direct comparison of Equations (28) and (10) demonstrates that the best and worst slopes of oblique regression satisfy Equation (29) and are thus conjugate to one another. Hence, the following pivotal result has been established [10]:

Figure 5: Conjugate Semidiameters
Theorem 1 (Galton-Pearson-McCartin) The best and worst lines of oblique regression are conjugate diameters of the concentration ellipse whose slopes satisfy the conjugacy relation Equation (29).

Note that $s = \frac{m-\lambda}{m-\mu}$ may be rearranged to reproduce the geometric characterization for $(\lambda, \mu)$-regression of McCartin [9]: $(m - \mu) \cdot (s - \mu) = \mu^2 - \lambda$. Furthermore, subsequently setting $\mu = 0$ reproduces the geometric characterization for $\lambda$-regression of McCartin [8]: $m \cdot s = -\lambda$. Now, setting $\lambda = 1$ reproduces Pearson’s geometric characterization for orthogonal regression [13]: it is the major axis of the concentration ellipse. Finally, setting $\lambda = 0, \infty$ reproduces Galton’s geometric characterization for coordinate regression [7]: they are the lines displayed in Figure 6 connecting the points of horizontal/vertical tangencies, respectively, to the centroid.

5 Geometric Constructions

Previous work on constructing an ellipse from its moments [12] will now be exploited in order to construct various lines of regression. This is possible because, as has been demonstrated above, all of these regression lines are associated with the concentration ellipse.

Once the moments $\langle \bar{x}, \bar{y}, \sigma^2_x, p_{xy}, \sigma^2_y \rangle$ have been computed from the discrete data set $\{(x_i, y_i)\}_{i=1}^{n}$, the equation of the concentration ellipse, $E$, is given by the positive definite quadratic form

$$\begin{bmatrix} x - \bar{x} & y - \bar{y} \end{bmatrix} \begin{bmatrix} \sigma^2_x & p_{xy} \\ p_{xy} & \sigma^2_y \end{bmatrix}^{-1} \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix} = 4,$$

and the elliptical area may be expressed as

$$A_E = 4\pi \sqrt{\sigma^2_x \sigma^2_y - p_{xy}^2}.$$

(31)

The slopes of the principal axes of the concentration ellipse are the roots of the quadratic

$$m^2 + \frac{\sigma^2_x - \sigma^2_y}{p_{xy}} \cdot m - 1 = 0,$$

so that

$$m_+ = \frac{(\sigma^2_y - \sigma^2_x) + \sqrt{(\sigma^2_y - \sigma^2_x)^2 + 4p_{xy}^2}}{2p_{xy}},$$

$$m_- = \frac{(\sigma^2_y - \sigma^2_x) - \sqrt{(\sigma^2_y - \sigma^2_x)^2 + 4p_{xy}^2}}{2p_{xy}},$$

(34)

since $m_+$ has the same sign as $p_{xy}$ [8].
The corresponding squared magnitudes of the principal semiaxes [14, p. 158] are the roots of the quadratic
\[ \Sigma^2 - 4(\sigma_x^2 + \sigma_y^2) \cdot \Sigma + 16(\sigma_x^2 \sigma_y^2 - p_{xy}^2) = 0, \] (35)
so that
\[ (\Sigma_+)^2 = 2 \left[ (\sigma_x^2 + \sigma_y^2) + \sqrt{(\sigma_x^2 - \sigma_y^2)^2 + 4p_{xy}^2} \right], \] (36)
\[ (\Sigma_-)^2 = 2 \left[ (\sigma_x^2 + \sigma_y^2) - \sqrt{(\sigma_x^2 - \sigma_y^2)^2 + 4p_{xy}^2} \right]. \] (37)

Equation (22) may be expanded thereby yielding
\[ \sigma_y^2 \cdot (x - \bar{x})^2 - 2p_{xy} \cdot (x - \bar{x})(y - \bar{y}) + \sigma_x^2 \cdot (y - \bar{y})^2 = 4(\sigma_x^2 \sigma_y^2 - p_{xy}^2) \] (38)
as the equation of the concentration ellipse. Setting \( y = \bar{y} \) yields the corresponding horizontal points of intersection
\[ \left( \bar{x} \pm 2\sqrt{\frac{\sigma_x^2 - p_{xy}^2}{\sigma_y^2}}, \bar{y} \right), \] (39)
while setting \( x = \bar{x} \) yields the corresponding vertical points of intersection
\[ \left( \bar{x}, \bar{y} \pm 2\sqrt{\frac{\sigma_y^2 - p_{xy}^2}{\sigma_x^2}} \right). \] (40)
The points of vertical and horizontal tangency are [10], respectively,
\[ \left( \bar{x} \pm 2\sigma_x, \bar{y} \pm \frac{p_{xy}}{\sigma_x} \right); \quad \left( \bar{x} \pm \frac{p_{xy}}{\sigma_y}, \bar{y} \pm 2\sigma_y \right). \] (41)

With reference to Figure 5, recall that two semidiameters are conjugate to one another if they are parallel to each others tangents [11]. For this to occur, it is necessary and sufficient that their slopes \( m \) and \( s \) satisfy the symmetric conjugacy condition, Equation (29). Thus, the semidiameter to a point of vertical/horizontal tangency is conjugate to the semidiameter to a point of vertical/horizontal intersection, respectively. (See Figure 6.)

In the succeeding subsections, straightedge-compass constructions of the various lines of regression, given the moments, \( \langle \bar{x}, \bar{y}, \sigma_x^2, p_{xy}, \sigma_y^2 \rangle \), will be developed. In this endeavor, certain primitive arithmetic/algebraic Euclidean constructions [2, pp. 121-122], as next described, will be required.

**Proposition 1 (Primitive Constructions)** Given two line segments of non-zero lengths \( a \) and \( b \), one may construct line segments of length \( a + b \), \( a - b \), \( a \cdot b \), \( a/b \) and \( \sqrt{a} \) using only straightedge and compass.
5.1 Coordinate Regression

As previously noted, the semidiameter of a point of vertical intersection is conjugate to the semidiameter to a point of vertical tangency. Thus, by Equations (40) and (41), the vectors
\[
\langle 0, 2\sqrt{\sigma_x^2 - p_{xy}^2/\sigma_y^2} \rangle; \quad \langle 2\sigma_x, 2p_{xy}/\sigma_x \rangle
\]
(42)
define a pair of conjugate semidiameters emanating from the centroid, \((\bar{x}, \bar{y})\).
(See Figure 6, Left.)

Construction 1 (Regression of \(y\)-on-\(x\)) Given the first and second moments \((\bar{x}, \bar{y}, \sigma_x^2, \sigma_y^2, p_{xy}, \sigma_y^2)\), the best and worst lines of regression of \(y\)-on-\(x\) are given, respectively, by the pair of conjugate semidiameters
\[
\langle x^+_y, y^+_y \rangle := \langle 2\sigma_x, 2p_{xy}/\sigma_x \rangle; \quad \langle x^-_y, y^-_y \rangle := \langle 0, 2\sqrt{\sigma_x^2 - p_{xy}^2/\sigma_x^2} \rangle
\]
which may be constructed using Proposition 1.

The corresponding best and worst mean-squared vertical deviations are given, respectively, by [15, pp. 220-222]
\[
\frac{1}{n} \cdot (S^+_y)^2 = \frac{4(\sigma_x^2 \sigma_y^2 - p_{xy}^2)}{(x^+_y)^2}, \quad \frac{1}{n} \cdot (S^-_y)^2 = \frac{4(\sigma_x^2 \sigma_y^2 - p_{xy}^2)}{(x^-_y)^2} = \infty.
\]
(43)
Likewise, the semidiameter of a point of horizontal intersection is conjugate to the semidiameter to a point of horizontal tangency. Thus, by Equations (39) and (41), the vectors
\[
\left\langle 2\sqrt{\sigma_y^2 - p_{xy}^2/\sigma_x^2}, \frac{p_{xy}}{\sigma_y} \right\rangle, \left\langle 2\sigma_y, 2\right\rangle (44)
\]
define a pair of conjugate semidiameters emanating from the centroid, \((\bar{x}, \bar{y})\).
(See Figure 6, Right.)

**Construction 2 (Regression of \(x\)-on-\(y\))**

Given the first and second moments \(\langle \bar{x}, \bar{y}, \sigma_x^2, \sigma_y^2, p_{xy} \rangle\), the best and worst lines of regression of \(x\)-on-\(y\) are given, respectively, by the pair of conjugate semidiameters
\[
\langle x^+, y^+ \rangle := \left\langle \frac{p_{xy}}{\sigma_y}, 2\sigma_y \right\rangle; \quad \langle x^-, y^- \rangle := \left\langle 2\sqrt{\sigma_y^2 - p_{xy}^2/\sigma_x^2}, 0 \right\rangle
\]
which may be constructed using Proposition 1.

The corresponding best and worst mean-squared vertical deviations are given, respectively, by [15, pp. 220-222]
\[
\frac{1}{n} \cdot (S_x^+)^2 = 4\frac{(\sigma_x^2\sigma_y^2 - p_{xy}^2)}{(y_+^2)^2} = \frac{\sigma_x^2\sigma_y^2 - p_{xy}^2}{\sigma_y^2}, \quad \frac{1}{n} \cdot (S_x^-)^2 = 4\frac{(\sigma_x^2\sigma_y^2 - p_{xy}^2)}{(y_-^2)^2} = \infty.
\]

### 5.2 Orthogonal Regression

The following beautiful construction due to Carlyle [5, p. 99] and Lill [4, pp. 355-356] will be invoked in order to construct the remaining lines of regression.

**Proposition 2 (Graphical Solution of a Quadratic Equation)**

Draw a circle having as diameter the line segment joining \((0, 1)\) and \((a, b)\). The abscissae of the points of intersection of this circle with the \(x\)-axis are the roots of the quadratic \(x^2 - ax + b = 0\).

A double application of Proposition 2, in concert with Equations (32) and (35), produces the following construction for the best and worst orthogonal regression lines.

**Construction 3 (Orthogonal Regression)**

Given the first and second moments \(\langle \bar{x}, \bar{y}, \sigma_x^2, \sigma_y^2, p_{xy} \rangle\), the best and worst lines of orthogonal regression are given by the pair of principal axes determined as follows:
1. According to Equation (32), apply the Carlyle-Lill construction, Proposition 2, with $(a, b) = (\frac{\sigma^2_y - \sigma^2_x}{p_{xy}}, -1)$ to construct the directions of the principal axes. (See Figure 7, Left.)

2. According to Equation (35), apply the Carlyle-Lill construction, Proposition 2, with $(a, b) = (4(\sigma^2_x + \sigma^2_y), 16(\sigma^2_x \sigma^2_y - p_{xy}^2))$ to construct the squared magnitudes of the principal semiaxes. (See Figure 7, Right.)

3. Apply the \(\sqrt{\cdot}\)-operator, Proposition 1, to these squared magnitudes in order to construct the magnitudes, \(\{\Sigma^+_\perp, \Sigma^-_\perp\}\), of the principal semiaxes emanating from the centroid, \((\bar{x}, \bar{y})\). (See Figure 8.)

4. Denote by \((x^\pm_\perp, y^\pm_\perp)\) the tips of these principal semiaxes lying on the concentration ellipse, \(\mathcal{E}\).

By Equation (36), the corresponding best mean-squared orthogonal deviation is given by

\[
\frac{1}{n} \cdot (S^+_{\perp})^2 = \frac{4(\sigma^2_x \sigma^2_y - p_{xy}^2)}{(x^+_{\perp})^2 + (y^+_{\perp})^2} = \frac{2(\sigma^2_x \sigma^2_y - p_{xy}^2)}{(\sigma^2_x + \sigma^2_y) + \sqrt{(\sigma^2_x - \sigma^2_y)^2 + 4p_{xy}^2}},
\]

while, by Equation (37), the corresponding worst mean-squared orthogonal deviation is given by

\[
\frac{1}{n} \cdot (S^-_{\perp})^2 = \frac{4(\sigma^2_x \sigma^2_y - p_{xy}^2)}{(x^-_{\perp})^2 + (y^-_{\perp})^2} = \frac{2(\sigma^2_x \sigma^2_y - p_{xy}^2)}{(\sigma^2_x + \sigma^2_y) - \sqrt{(\sigma^2_x - \sigma^2_y)^2 + 4p_{xy}^2}}.
\]
Equations (46) and (47) may be combined and further simplified to read
\[
\frac{1}{n} \cdot (S^\perp) = \frac{1}{2} \cdot \left[ (\sigma_x^2 + \sigma_y^2) \mp \sqrt{(\sigma_x^2 - \sigma_y^2)^2 + 4p_{xy}^2} \right].
\] (48)

6 \text{ λ-Regression}

Turning to the case of λ-regression corresponding to the slope \( s = -\frac{\lambda}{m} \) in Figure 3, it is shown in [8] that the \( x \)-expansion (\( \lambda > 1 \)) / compression (\( \lambda < 1 \))
\[
\xi := \sqrt{\lambda} \cdot x \Rightarrow \sigma_x^2 = \lambda \cdot \sigma_x^2, \quad p_{xy} = \sqrt{\lambda} \cdot p_{xy}, \quad \hat{m} = \frac{m}{\sqrt{\lambda}}
\] (49)
transforms λ-regression into orthogonal regression so that the slopes of the best and worst λ-regression lines are, by Equation (32), roots of the quadratic
\[
m^2 + \frac{\lambda \sigma_x^2 - \sigma_y^2}{p_{xy}} \cdot m - \lambda = 0.
\] (50)
Specifically, by Equation (33), the slope of the best λ-regression line is
\[
m^+ = \frac{(\sigma_y^2 - \lambda \sigma_x^2) + \sqrt{(\sigma_y^2 - \lambda \sigma_x^2)^2 + 4\lambda p_{xy}^2}}{2p_{xy}},
\] (51)
while, by Equation (34), the slope of the worst λ-regression line is
\[
m^- = \frac{(\sigma_y^2 - \lambda \sigma_x^2) - \sqrt{(\sigma_y^2 - \lambda \sigma_x^2)^2 + 4\lambda p_{xy}^2}}{2p_{xy}}.
\] (52)
Figure 9: \( \lambda \)-Regression

By Equation (35), the squared magnitudes of the corresponding conjugate semidiameters are the roots of the quadratic

\[
\Sigma_\lambda^2 - 4(\lambda \sigma_x^2 + \sigma_y^2) \cdot \Sigma_\lambda + 16\lambda(\sigma_x^2 \sigma_y^2 - p_{xy}^2) = 0,
\]

so that

\[
(\Sigma_\lambda^+)^2 = 2 \left[ (\lambda \sigma_x^2 + \sigma_y^2) + \sqrt{(\lambda \sigma_x^2 - \sigma_y^2)^2 + 4\lambda p_{xy}^2} \right], \tag{54}
\]

\[
(\Sigma_\lambda^-)^2 = 2 \left[ (\lambda \sigma_x^2 + \sigma_y^2) - \sqrt{(\lambda \sigma_x^2 - \sigma_y^2)^2 + 4\lambda p_{xy}^2} \right]. \tag{55}
\]

Thus, Construction 3 may be modified to accommodate \( \lambda \)-regression.

**Construction 4 (\( \lambda \)-Regression)** Given the first and second moments \( \langle \bar{x}, \bar{y}, \sigma_x^2, p_{xy}, \sigma_y^2 \rangle \), the best and worst lines of \( \lambda \)-regression are given by the pair of conjugate semidiameters determined as follows:

1. According to Equation (50), apply the Carlyle-Lill construction, Proposition 2, with \((a, b) = (\frac{\sigma_y^2 - \lambda \sigma_x^2}{p_{xy}}, -\lambda)\) to construct the directions of the conjugate semidiameters.

2. According to Equation (53), apply the Carlyle-Lill construction, Proposition 2, with \((a, b) = (4(\lambda \sigma_x^2 + \sigma_y^2), 16\lambda(\sigma_x^2 \sigma_y^2 - p_{xy}^2))\) to construct the squared magnitudes of the conjugate semidiameters.

3. Apply the \( \sqrt{-} \)-operator, Proposition 1, to these squared magnitudes in order to construct the magnitudes, \( \{\Sigma_\lambda^+, \Sigma_\lambda^-\} \), of the conjugate semidiameters emanating from the centroid, \((\bar{x}, \bar{y})\). (See Figure 9.)
4. Denote by \((x^\pm_Y, y^\pm_Y)\) the tips of these conjugate semidiameters lying on the concentration ellipse, \(E\).

By Equation (46), the corresponding best mean-squared \(\lambda\)-deviation is given by

\[
\frac{1}{n} \cdot \left( S^+_\lambda \right)^2 = \max (1, \lambda^{-1}) \cdot \frac{4\lambda(\sigma^2_x \sigma^2_y - p^2_{xy})}{\lambda (x^+_\lambda)^2 + (y^+_\lambda)^2} \\
= \max (1, \lambda^{-1}) \cdot \frac{2\lambda(\sigma^2_x \sigma^2_y - p^2_{xy})}{(\lambda \sigma^2_x + \sigma^2_y) + \sqrt{(\lambda \sigma^2_x - \sigma^2_y)^2 + 4\lambda p^2_{xy}}} \\
= \max (1, \lambda^{-1}) \cdot \frac{1}{2} \left( \lambda \sigma^2_x + \sigma^2_y \right) - \sqrt{(\lambda \sigma^2_x - \sigma^2_y)^2 + 4\lambda p^2_{xy}}, \quad (56)
\]

while, by Equation (47), the corresponding worst mean-squared \(\lambda\)-deviation is given by

\[
\frac{1}{n} \cdot \left( S^-_\lambda \right)^2 = \max (1, \lambda^{-1}) \cdot \frac{4\lambda(\sigma^2_x \sigma^2_y - p^2_{xy})}{\lambda (x^-_\lambda)^2 + (y^-_\lambda)^2} \\
= \max (1, \lambda^{-1}) \cdot \frac{2\lambda(\sigma^2_x \sigma^2_y - p^2_{xy})}{(\lambda \sigma^2_x + \sigma^2_y) - \sqrt{(\lambda \sigma^2_x - \sigma^2_y)^2 + 4\lambda p^2_{xy}}} \\
= \max (1, \lambda^{-1}) \cdot \frac{1}{2} \left( \lambda \sigma^2_x + \sigma^2_y \right) + \sqrt{(\lambda \sigma^2_x - \sigma^2_y)^2 + 4\lambda p^2_{xy}} \quad (57)
\]

7 \((\lambda, \mu)\)-Regression

Turning to the case of \((\lambda, \mu)\)-regression corresponding to the slope \(s = \frac{\mu m - \lambda}{m - \mu}\) in Figure 3, it is shown in [9] that the linear mapping

\[
\xi := \sqrt{\lambda - \mu^2} \cdot x, \quad \eta := -\mu \cdot x + y \Rightarrow \hat{m} = \frac{m - \mu}{\sqrt{\lambda - \mu^2}}, \quad (58)
\]

\[
\sigma^2_\xi = (\lambda - \mu^2) \cdot \sigma^2_x, \quad \sigma^2_\eta = \mu^2 \cdot \sigma^2_x - 2\mu \cdot p_{xy} + \sigma^2_y, \quad p_{\xi\eta} = \sqrt{\lambda - \mu^2} \cdot (p_{xy} - \mu \sigma^2_x) \quad (59)
\]

transforms \((\lambda, \mu)\)-regression into orthogonal regression so that the slopes of the best and worst \((\lambda, \mu)\)-regression lines are, by Equation (32), roots of the quadratic

\[
m^2 + \frac{\lambda \sigma^2_x - \sigma^2_y}{p_{xy} - \mu \sigma^2_x} \cdot m - \frac{\lambda p_{xy} - \mu \sigma^2_y}{p_{xy} - \mu \sigma^2_x} = 0. \quad (60)
\]

Specifically, by Equation (33), the slope of the best \((\lambda, \mu)\)-regression line is

\[
m^+_{\lambda, \mu} = \frac{(\sigma^2_y - \lambda \sigma^2_x) + \sqrt{(\sigma^2_y - \lambda \sigma^2_x)^2 + 4(p_{xy} - \mu \sigma^2_x)(\lambda p_{xy} - \mu \sigma^2_y)}}{2(p_{xy} - \mu \sigma^2_x)}, \quad (61)
\]
while, by Equation (34), the slope of the worst \((\lambda, \mu)\)-regression line is
\[
m_{\lambda, \mu} = \frac{(\sigma_y^2 - \lambda \sigma_x^2) - \sqrt{(\sigma_y^2 - \lambda \sigma_x^2)^2 + 4(p_{xy} - \mu \sigma_y^2)(\lambda p_{xy} - \mu \sigma_y^2)}}{2(p_{xy} - \mu \sigma_y^2)}. \quad (62)
\]

By Equation (35), the squared magnitudes of the corresponding conjugate semidiameters are the roots of the quadratic
\[
\Sigma_{\lambda, \mu}^2 - 4 \left[ (\lambda - \mu^2) \sigma_x^2 + (\mu^2 \sigma_x^2 - 2 \mu \sigma_{xy} + \sigma_y^2) \right] \cdot \Sigma_{\lambda, \mu}
\]
\[
+ 16(\lambda - \mu^2) \left[ \sigma_x^2(\mu^2 \sigma_x^2 - 2 \mu \sigma_{xy} + \sigma_y^2) - (p_{xy} - \mu \sigma_x^2)^2 \right] = 0, \quad (63)
\]
so that
\[
(\Sigma_{\lambda, \mu}^+)^2 = 2 \left[ (\lambda - \mu^2) \sigma_x^2 + (\mu^2 \sigma_x^2 - 2 \mu \sigma_{xy} + \sigma_y^2) \right]
\]
\[
+ 2\sqrt{\left[ (\lambda - \mu^2) \sigma_x^2 - (\mu^2 \sigma_x^2 - 2 \mu \sigma_{xy} + \sigma_y^2) \right]^2 + 4(\lambda - \mu^2)(p_{xy} - \mu \sigma_y^2)^2}, \quad (64)
\]
\[
(\Sigma_{\lambda, \mu}^-)^2 = 2 \left[ (\lambda - \mu^2) \sigma_x^2 + (\mu^2 \sigma_x^2 - 2 \mu \sigma_{xy} + \sigma_y^2) \right]
\]
\[
- 2\sqrt{\left[ (\lambda - \mu^2) \sigma_x^2 - (\mu^2 \sigma_x^2 - 2 \mu \sigma_{xy} + \sigma_y^2) \right]^2 + 4(\lambda - \mu^2)(p_{xy} - \mu \sigma_y^2)^2}. \quad (65)
\]
Thus, Construction 3 may be modified to accommodate \((\lambda, \mu)\)-regression.
Construction 5 \((\lambda, \mu)\)-Regression\) Given the first and second moments \(\langle \bar{x}, \bar{y}, \sigma_x^2, p_{xy}, \sigma_y^2 \rangle\), the best and worst lines of \((\lambda, \mu)\)-regression are given by the pair of conjugate semidiameters determined as follows:

1. According to Equation (60), apply the Carlyle-Lill construction, Proposition 2, with \((a, b) = \left( \frac{\sigma_x^2 - \lambda \sigma_x^2}{p_{xy} - \mu \sigma_x^2}, \frac{\mu \sigma_x^2 - \lambda \sigma_x^2}{p_{xy} - \mu \sigma_x^2} \right)\) to construct the directions of the conjugate semidiameters.

2. According to Equation (63), apply the Carlyle-Lill construction, Proposition 2, with \((a, b) = \left( 4 \left( (\lambda - \mu^2) \sigma_x^2 + (\mu^2 \sigma_x^2 - 2 \mu \sigma_{xy} + \sigma_y^2) \right) \right),\)

\[
16(\lambda - \mu^2) \left[ \sigma_x^2 (\mu^2 \sigma_x^2 - 2 \mu \sigma_{xy} + \sigma_y^2) - (p_{xy} - \mu \sigma_x^2)^2 \right] \]

to construct the squared magnitudes of the conjugate semidiameters.

3. Apply the \(\sqrt{\cdot}\)-operator, Proposition 1, to these squared magnitudes in order to construct the magnitudes, \(\{\Sigma_{\lambda, \mu}^+, \Sigma_{\lambda, \mu}^-\}\), of the conjugate semidiameters emanating from the centroid, \((\bar{x}, \bar{y})\). (See Figure 10.)

4. Denote by \(\langle x_{\lambda, \mu}^+, y_{\lambda, \mu}^+ \rangle\) the tips of these conjugate semidiameters lying on the concentration ellipse, \(E\).

By Equation (46), the corresponding best mean-squared \((\lambda, \mu)\)-deviation is given by

\[
\frac{1}{n} \left( S_{\lambda, \mu}^+ \right)^2 = \max (1, \lambda^{-1}) \cdot \frac{4(\lambda - \mu^2) \left[ \sigma_x^2 (\mu^2 \sigma_x^2 - 2 \mu \sigma_{xy} + \sigma_y^2) - (p_{xy} - \mu \sigma_x^2)^2 \right]}{\lambda (x_{\lambda, \mu}^+)^2 - 2 \mu x_{\lambda, \mu}^+ y_{\lambda, \mu}^+ + (y_{\lambda, \mu}^+)^2}
\]

while, by Equation (47), the corresponding worst mean-squared \((\lambda, \mu)\)-deviation is given by

\[
\frac{1}{n} \left( S_{\lambda, \mu}^- \right)^2 = \max (1, \lambda^{-1}) \cdot \frac{4(\lambda - \mu^2) \left[ \sigma_x^2 (\mu^2 \sigma_x^2 - 2 \mu \sigma_{xy} + \sigma_y^2) - (p_{xy} - \mu \sigma_x^2)^2 \right]}{\lambda (x_{\lambda, \mu}^-)^2 - 2 \mu x_{\lambda, \mu}^- y_{\lambda, \mu}^- + (y_{\lambda, \mu}^-)^2}
\]
8 Conclusion

In the foregoing, Euclidean constructions were supplied for various lines of regression. These included lines of coordinate regression [7], orthogonal regression [13], \( \lambda \)-regression [8] and \((\lambda, \mu)\)-regression [9]. The adhesive which bound together these geometric constructions was provided by the concentration ellipse [3]. This unified approach was, in turn, a direct consequence of the Galton-Pearson-McCartin Theorem [10, Theorem 2] which establishes that, as special cases of oblique regression [10], all of the associated best and worst regression lines form pairs of conjugate diameters of the concentration ellipse.

Acknowledgements. I would like to thank my devoted wife Mrs. Barbara A. McCartin for her indispensable aid in constructing the figures.

References


Received: January 1, 2016; Published: March 30, 2016