ORTHOGONAL CALIBRATION

By DAVID R. FOX †

University of Colorado, Denver U.S.A.

SUMMARY

An alternative to the inverse and classical methods of calibration is proposed. In the classical approach the calibration data are projected vertically onto the regression line; for the inverse method the projection is horizontal. We examine a compromise, or hybrid approach where the data are projected 'orthogonally' onto the regression line. The parameter estimates so obtained are shown to correspond to the spectral decomposition of the sample covariance matrix of the calibration data. Furthermore, prediction intervals for a calibrated response can be obtained using Fieller's theorem. The methods are illustrated using recently published calibration data.

Keywords: *classical estimator; inverse estimator; principle components regression; general linear model; Fieller's theorem.*

1. Introduction

The problem of linear calibration has received a good deal of attention over the years [Berkson (1969), Kruthkoff (1967,1969), Williams (1969), Lwin and Maritz (1982)] and even after protracted debate over the relative merits of classical versus inverse regression, articles continue appear in relation to this issue [eg. Chow and Shao (1990)].

Much of the early discussion centered on the appropriateness of the different criteria for judging competing estimators (eg. mean square error versus Pitman closeness) and to an extent, the issue remains unresolved. While the "classical" versus "inverse" controversy has subsided, the issue still has important ramifications for for the practitioner. Frequently in my role as consulting statistician I am asked by researchers "should I regress x on y or y on x?" It was Berkson (1950) who formally raised the question of "Are there two regressions" forty years ago, and, for the researcher faced with a practical calibration problem, the question is as valid today as it was then. In their recent article, Chow and Shao (1990) indicate the importance of satisfactorily addressing this issue in pharmaceutical testing.

---

† *Address for correspondence:* Mathematics Department, University of Colorado at Denver, Campus Box 170, PO. Box 173364, Denver, Colorado 80217-3364 U.S.A.

As an example they cite a hypothetical situation in which two companies produce the same drug. Different assessments of assay results may arise not because of inherent product differences but due to the different calibration procedures employed. In tackling this conflict, the authors examined the probability that the ratio of the inverse to classical estimates is within certain prescribed bounds, arguing that if it is then the two methods are "interchangeable". Where a significant difference between the two estimates is observed, the suggestion is that "both companies should use the better method". In making such a recommendation, the authors have done nothing to clarify the situation for we are naturally led to ask what constitutes better or best? It would make more sense to remove the ambiguity in the first place thus avoid being confronted with the dilemma.

In this paper we propose an alternative to the classical and inverse estimation procedures which, in a sense, may be thought of as a compromise between the existing methodologies. Rather than project the calibration data vertically onto the regression line as is done in the classical approach, or horizontally as in the inverse method, we project the data 'orthogonally' as depicted in figure 1 below. The least squares criterion is then applied to the distances, d_i.

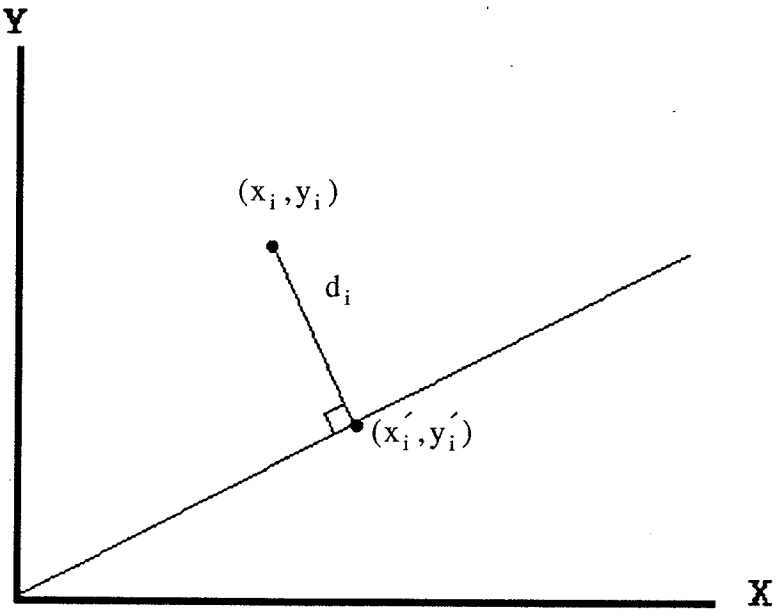

Figure 1. A typical data point (x_i,y_i) and its orthogonal projection (x_i',y_i')

A more general situation corresponding to an arbitrary angle, θ of projection, is depicted in figure 2.

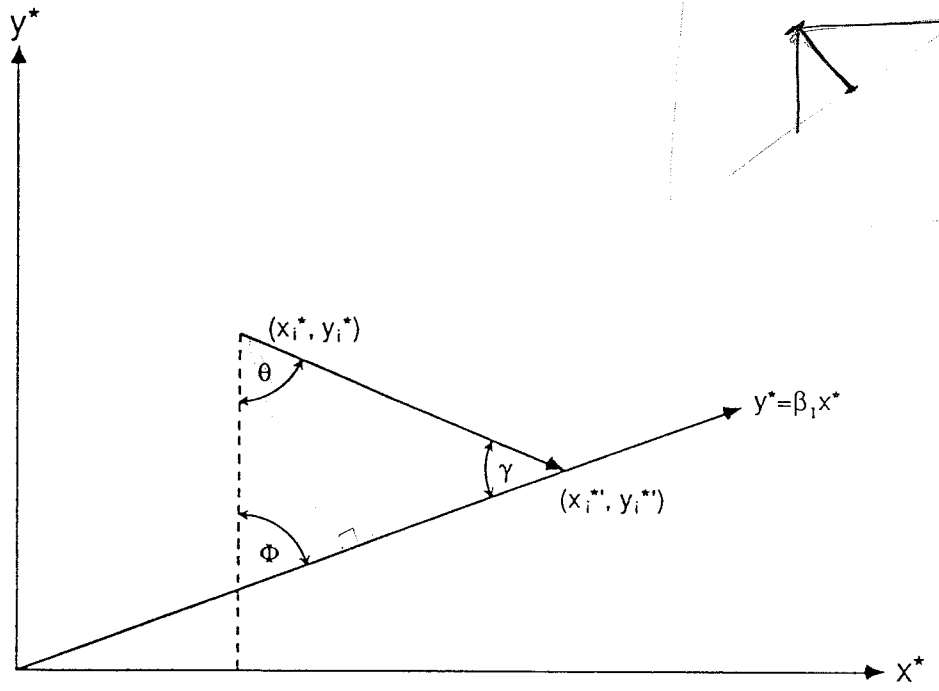

Figure 2. A typical *centered* data point (x_i^*, y_i^*) and its projection $(x_i^{*'}, y_i^{*'})$.

Some clarification of the use of the word orthogonal is necessary since the same terminology is often used in reference to ordinary least squares (OLS). The orthogonality referred to here is in reference to the angle γ in figure 2. In OLS the term is used in connection with the orthogonality of the error vector $(y - \hat{y})$ and the column space spanned by the predictor variables.

One further point to note about the orthogonal projection is that the method takes cognisance of errors in both the x and y data. Fuller (1987) briefly mentions the calibration problem in his comprehensive treatment of measurement error models, although the methods described assume that the magnitudes of the errors in X and Y are known. It is acknowledged that it would be desirable to allow the angle of projection to be dictated by the ratio of these error variances, however for most practical situations this quantity is unknown.

2. The Geometry of Orthogonal Projections

In the following development we assume that X and Y have a bivariate distribution with finite first and second moments. Furthermore X^* and Y^* refer to the centred variables $(X - \mu_X)$ and $(Y - \mu_Y)$ respectively. Our regression model is thus

$$Y_i^* = \beta_1 X_i^* \tag{2.1}$$

Consider a typical data point (x_i^*, y_i^*). Figure 2 shows the point (x_i^*, y_i^*), its projection $(x_i^{*'}, y_i^{*'})$ onto the regression line, and the various angles involved. θ is the angle of projection, γ is the angle the regression line makes with the line segment joining (x_i^*, y_i^*) and $(x_i^{*'}, y_i^{*'})$ and ϕ is the angle the regression line makes with the Y^* axis. Using elementary mensuration and trigonometric relationships we have

$$\tan(\gamma) = \frac{1 + \beta_1 \tan(\theta)}{\beta_1 - \tan(\theta)} \tag{2.2}$$

and

$$\phi = \tan^{-1}\left(\frac{1}{\beta_1}\right) \tag{2.3}$$

Furthermore, the projection equations are given as

$$x_i^{*'} = \frac{y_i^* + x_i^* \cot(\theta)}{\beta_1 + \cot(\theta)} \tag{2.4a}$$

$$y_i^{*'} = \frac{\beta_1 y_i^* + x_i^* \beta_1 \cot(\theta)}{\beta_1 + \cot(\theta)} \tag{2.4b}$$

The projection matrix $\mathcal{P}$ corresponding to equations (2.4a) and (2.4b) is thus

$$\mathcal{P} = \frac{1}{1 + \cot(\theta)} \begin{bmatrix} \cot(\theta) & 1 \\ \beta_1 \cot(\theta) & \beta_1 \end{bmatrix} \tag{2.5}$$

For an orthogonal projection we have $\gamma = \frac{\pi}{2}$ and $\theta = \frac{\pi}{2} - \phi$ thus $\tan(\theta) = \tan(\frac{\pi}{2} - \phi)$ $= \cot(\phi)$. Using equation (2.3) we therefore have that $\tan(\theta) = \beta_1$. Substituting into equation (2.5) we obtain the matrix for an orthogonal projection :

$$\mathcal{P} = \frac{1}{1 + \beta_1^2} \begin{bmatrix} 1 & \beta_1 \\ \beta_1 & \beta_1^2 \end{bmatrix} \tag{2.6}$$

3. Least-Squares Estimation

For any given value of θ we wish to estimate β_1 such that the following quantity is minimized

$$Q = \sum_{i=1}^{n} \left\{ \left[\frac{y_i^* - \beta_1 x_i^*}{\beta_1 .+ \cot(\theta)} \right]^2 + \left[\frac{\beta_1 x_i^* - y_i^*}{\beta_1 \tan(\theta) + 1} \right]^2 \right\}$$

$$= \sum_{i=1}^{n} d_i^2 \tag{3.1}$$

Differentiating (3.1) with respect to β_1, equating to zero and solving gives the least-squares estimate :

$$\hat{\beta}_1 = \frac{\sin(\theta) \sum_{i=1}^{n} y_i^{*2} + \cos(\theta) \sum_{i=1}^{n} x_i^* y_i^*}{\cos(\theta) \sum_{i=1}^{n} x_i^{*2} + \sin(\theta) \sum_{i=1}^{n} x_i^* y_i^*} \tag{3.2}$$

Observe that when $\theta = 0$ equation (3.2) reduces to

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i^* y_i^*}{\sum_{i=1}^{n} x_i^{*2}} \tag{3.3}$$

which is the OLS estimate.

4. Relationship with the spectral decomposition of the covariance matrix

Let $\mathcal{C}$ denote the sample covariance matrix :

$$\mathcal{C} = \begin{bmatrix} S_{XX} & S_{XY} \\ S_{XY} & S_{YY} \end{bmatrix}$$

Furthermore, let $\underline{e} = [e_1, e_2]^T$ be an eigenvector of $\mathcal{C}$ and hence $\mathcal{C}\underline{e} = \lambda\underline{e}$ where λ is the corresponding eigenvalue. Now $\lambda\underline{e} = [\lambda e_1, \lambda e_2]^T$ and the gradient of the line from the origin to the point $(\lambda e_1, \lambda e_2)$ is equal to $\tilde{\beta}_1 = \dfrac{\lambda e_2}{\lambda e_1} = \dfrac{e_2}{e_1}$.

But,

$$\mathcal{C}\underline{e} = \begin{bmatrix} S_{XX} & S_{XY} \\ S_{XY} & S_{YY} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} e_1 S_{XX} + e_2 S_{XY} \\ e_1 S_{XY} + e_2 S_{YY} \end{bmatrix}$$

and so

$$\tilde{\beta}_1 = \frac{e_1 S_{XY} + e_2 S_{YY}}{e_1 S_{XX} + e_2 S_{XY}} \qquad (4.1)$$

Equating coefficients in equations (3.2) and (4.1) we see that $e_2 = \sin(\theta)$ and $e_1 = \cos(\theta)$, that is

$$\underline{e} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}$$

However, in the previous section it was shown that for an orthogonal projection $\tilde{\beta}_1 = \tan(\theta)$. We are therefore led to conclude that one of the eigenvectors of $\mathcal{C}$ corresponds to a least-squares fit when the data are projected orthogonally onto the regression line. The other eigenvector is a line at right angles to the first and will not make physical sense in the calibration problem.

As an alternative to performing a spectral decomposition of $\mathfrak{C}$ in order to determine $\tilde{\beta}_1$ we can obtain an expression in terms of familiar sums of squares by dividing equation (3.2) through by $\cos(\theta)$ and noting that for an orthogonal projection $\beta_1 = \tan(\theta)$. The resulting equation is a quadratic in β_1 which has solutions

$$\tilde{\beta}_1 = \frac{(S_{YY} - S_{XX}) \pm \left\{ \left(S_{YY} - S_{XX} \right)^2 + 4S_{XY}^2 \right\}^{1/2}}{2S_{XY}} \qquad (4.2)$$

Once again, only one of the roots of equation (4.2) will make sense in the context of the problem.

5. Interval estimation for orthogonally-calibrated data

To construct prediction intervals for some calibrated X_0 corresponding to a given value of $Y = y_0$ we make use of a procedure due to Fox (1991) and asymptotic distributional results for the eigenvectors of $\mathfrak{C}$. Our estimated calibration model is $\hat{Y}_i = \tilde{\beta}_0 + \tilde{\beta}_i X_i$ where $\tilde{\beta}_1$ is the orthogonal estimate for the regression slope and $\tilde{\beta}_0$ is estimated in the usual way, that is $\tilde{\beta}_0 = \bar{y} - \tilde{\beta}_1 \bar{x}$. In matrix notation

$$\hat{\mathbf{Y}} = \mathbf{X}\, \underline{\tilde{\beta}} \qquad (5.1)$$

where the elements of $\underline{\tilde{\beta}}$ are $\tilde{\beta}_0$ and $\tilde{\beta}_1$. For a new observation Y_0 the classical estimator of X is $(Y_0 - \tilde{\beta}_0)/\tilde{\beta}_1 = \hat{X}_0$. Fox (1991) showed how Fieller's theorem [Fieller (1944)] may be utilized to provide a $(1-\alpha)100\%$ prediction interval for X_0. The method is now briefly described.

First, we define a new quantity $\underline{\beta}^*$ as the vector obtained by inserting Y_0 in the first row position of $\underline{\beta}$. That is:

$$\underline{\beta}^{*T} = \left[Y_0 \mid \underline{\beta}^T \right] \qquad (5.2)$$

$\tilde{\beta}^*$ is similarly defined by replacing $\underline{\beta}$ in equation (5.2) with $\underline{\tilde{\beta}}$.

Under the assumption that Y_0 is a random variable from the same population as the original calibration data, we have:

$$\mathbf{V} = \mathrm{Cov}\left[\hat{\underline{\beta}}^{*}\right] = \left[\begin{array}{c|c} \sigma_{\varepsilon}^{2} & \mathbf{0} \\ \hline \mathbf{0}^{T} & \Sigma \end{array}\right] \tag{5.3}$$

$$= \sigma_{\varepsilon}^{2}\left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0}^{T} & (X^{T}X)^{-1} \end{array}\right] \tag{5.4}$$

where $\mathbf{0}$ is a (1 x 2) vector of zeros.

The classical estimator can be written as the ratio of linear combinations of the elements of $\tilde{\underline{\beta}}^{*}$. Specifically,

$$\hat{X}_{0C} = \frac{K^{T}\tilde{\beta}^{*}}{L^{T}\tilde{\beta}^{*}} \tag{5.5}$$

where $K^{T} = [1 \quad -1 \quad 0]$
and $L^{T} = [0 \quad 0 \quad 1]$.

Following Zerbe (1978), we define quantities A, B, and C as follows:

$$A = (L^{T}\tilde{\underline{\beta}}^{*})^{2} - t_{\nu,\alpha/2}^{2}\, L^{T}\mathbf{V}L \tag{5.6}$$

$$B = 2\left[t_{\nu,\alpha/2}^{2}\, K^{T}\mathbf{V}L - (K^{T}\tilde{\underline{\beta}}^{*})(L^{T}\tilde{\underline{\beta}}^{*})\right] \tag{5.7}$$

$$C = (K^{T}\tilde{\underline{\beta}}^{*})^{2} - t_{\nu,\alpha/2}^{2}\, K^{T}\mathbf{V}K \tag{5.8}$$

where $t_{\nu,\alpha/2}^{2}$ is the $(1 - \alpha/2)100$ percentile of the T-distribution having ν degrees of freedom. Let a,b, and c be the observed values of the corresponding random variables. Then provided $a > 0$ and $b^{2} - 4ac > 0$, the limits of the $(1 - \alpha)100\%$ prediction interval for X_{0C} are:

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \qquad (5.9)$$

When β_1 is estimated via OLS, then Σ (=Cov[$\underline{\hat{\beta}}$]) in equation (5.3) is $\sigma_\varepsilon^2(X^TX)^{-1}$ and the application of equations (5.5) through (5.9) is straightforward. For the *orthogonal* estimation of β_1 we require Cov[$\underline{\tilde{\beta}}$] before the preceding method of interval estimation can be used. We now derive a first-order approximation to Cov[$\underline{\tilde{\beta}}$].

5.1 A first-order approximation for Cov[$\underline{\tilde{\beta}}$].

The following theorem is due to Anderson (1963) and appears as theorem 8.3.3. in Mardia et.al. (1979):

Theorem Let Σ be a positive definite matrix with distinct eigenvalues. Let $\mathbf{M} \sim W_p(\Sigma,m)$ and set $\mathbf{U} = m^{-1}\mathbf{M}$. Consider the spectral decompositions $\Sigma = \Gamma\Lambda\Gamma^T$ and $\mathbf{U} = \mathbf{GLG}^T$. Then as $m \rightarrow \infty$, $\mathbf{g}_{(i)} \sim N_p(\gamma_i,\mathbf{W}_i/m)$, where

$$\mathbf{W}_i = \lambda_i \sum_{j \neq i} \frac{\lambda_j}{(\lambda_j - \lambda_i)^2} \gamma_{(j)}\gamma_{(j)}^T$$

In otherwords, the eigenvectors of $\mathbf{U}$ are asymptotically normally distributed, unbiased and have the stated asymptotic covariance matrix $\mathbf{W}_i/m$.

$$\text{Now } \underline{\tilde{\beta}} = \begin{bmatrix} \tilde{\beta}_0 \\ \tilde{\beta}_1 \end{bmatrix} = \begin{bmatrix} 1 & -\bar{x} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{Y} \\ \underline{e_2} \\ e_1 \end{bmatrix}.$$

Let $\underline{\varphi} = \begin{bmatrix} \bar{Y} \\ e_2 \\ \hline e_1 \end{bmatrix}$; $\mathbf{T} = \begin{bmatrix} \overline{Y} \\ e_1 \\ e_2 \end{bmatrix}$; $\mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$.

with $\mathrm{Cov}[\mathbf{T}] = \left[\begin{array}{c|c} \sigma_\varepsilon^2/n & \mathbf{0} \\ \hline \mathbf{0}^T & \Psi \end{array}\right]$ (5.10)

where $\Psi = \mathrm{Cov}[\mathbf{e}] = \begin{bmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{bmatrix}$ and $\mathbf{0}$ is a (1 x 2) vector of zeros and the elements of Ψ are computed using the previously stated theorem (i.e. Ψ is either $\mathbf{W}_1$ or $\mathbf{W}_2$ of the theorem, depending on whether $\mathbf{e}$ is the first or second eigenvector of $\mathfrak{C}$). The independence between $\overline{Y}$ and $\mathbf{e}$ implied by equation (5.10) is readily appreciated when one considers that the variance-covariance matrix $\mathfrak{C}$ (and hence $\mathbf{e}$) is unaffected by changes in *location.*

Furthermore, let $\mathbb{E}[\mathbf{T}] = \underline{\Theta} = [\theta_1, \theta_2, \theta_3]^T$

and $g(\underline{\Theta}) = \begin{bmatrix} \theta_1 \\ \hline \theta_3 \\ \theta_2 \end{bmatrix}$.

Therefore $\mathbf{G} = \nabla[\, g(\underline{\Theta})] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\dfrac{\theta_3}{\theta_2^2} & \dfrac{1}{\theta_2} \end{bmatrix}$

Now, $\mathrm{Cov}[\, \tilde{\underline{\beta}}\,] = \begin{bmatrix} 1 & -\bar{x} \\ 0 & 1 \end{bmatrix} \mathrm{Cov}[\, \underline{\varphi}\,] \begin{bmatrix} 1 & 0 \\ -\bar{x} & 1 \end{bmatrix}$

and $\mathrm{Cov}[\, \underline{\varphi}\,] \cong \mathbf{G}\, \mathrm{Cov}[\mathbf{T}]\, \mathbf{G}^T$

[Rao (1973),p388].

$$= \begin{bmatrix} \sigma_\varepsilon^2/n & 0 \\ 0 & \omega \end{bmatrix}$$

where $\omega = \frac{\theta_3^2}{\theta_2^4}\psi_{11} - \frac{2\theta_3}{\theta_2^3}\psi_{12} + \frac{1}{\theta_2^2}\psi_{22}$.

Substituting observed values for parameters θ_2 and θ_3 we have:

$$\omega = = \frac{1}{e_1^2}\left[\tilde{\beta}_1^2\,\psi_{11} - 2\tilde{\beta}_1\psi_{12} + \psi_{22}\right]$$

Finally,

$$\mathrm{Cov}\left[\tilde{\underline{\beta}}\right] = \begin{bmatrix} 1 & -\bar{x} \\ 0 & 1 \end{bmatrix} \left[\begin{array}{c|c} \sigma_\varepsilon^2/n & \mathbf{0} \\ \hline \mathbf{0}^T & \Psi \end{array}\right] \begin{bmatrix} 1 & 0 \\ -\bar{x} & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_\varepsilon^2/n + \omega\bar{x}^2 & -\omega\bar{x} \\ -\omega\bar{x} & \omega \end{bmatrix} = \mathbf{M} \qquad (5.11)$$

We now have all the necessary components for obtaining prediction intervals for an orthogonally-calibrated estimate $\hat{X}_{0(\mathrm{orthog})}$ corresponding to a new observation $Y = y_0$. The procedure is summarized below:

1. Form the sample covariance matrix $\mathfrak{C}$ as given in section 4.
2. Obtain the spectral decomposition of $\mathfrak{C}$ and note the principle eigenvector, $\mathbf{e} = [e_1, e_2]^T$.
3. Compute the orthogonal regression estimates:

$$\tilde{\beta}_1 = \frac{e_2}{e_1} \quad \text{and} \quad \tilde{\beta}_0 = \bar{y} - \tilde{\beta}_1\bar{x}$$

and let $\underline{\tilde{\beta}} = \left[\tilde{\beta}_0 , \tilde{\beta}_1 \right]^T$.

4. Augment $\underline{\tilde{\beta}}$ to form $\underline{\tilde{\beta}}^*$ as indicated in equation (5.2).

5. The orthogonally-calibrated estimate $(\hat{X}_{0(orthog)})$ corresponding to y_0 is obtained by computing equation (5.5) using $\underline{\tilde{\beta}}^*$.

6. Obtain matrix **M** using equation (5.11) and substitute this for Σ in equation (5.3) to obtain matrix **V**.

7. Use equations (5.6), (5.7), and (5.8) to compute the sample values a,b,c of A,B, and C for some prescribed level α.

8. The limits of the $(1 - \alpha)100\%$ prediction interval for $X_{0(orthog)}$ are obtained using equation (5.9).

The procedure is not as complex as the steps above may indicate. The matrix calculations involve matrices of order no greater than (3 x 3). The spectral decomposition of $\mathfrak{C}$ is readily achieved using intrinsic functions in matrix-based computer programs such as GAUSS or MATLAB. We now illustrate the procedure by applying it to two recently published calibration problems.

6. Examples

6.1 Pharmaceutical Testing

This first example was used by Chow and Shao (1990) to illustrate the differences between the classical and inverse methods of calibration in a pharmaceutical testing problem. The data consisted of absorbance measurements obtained for 12 preparations of a standard whose concentration was accurately determined. On the basis of the estimated calibration line an estimate was obtained for a new preparation of unknown concentration. The data is reproduced in the table below.

Concen. X_i	Absorbance Y_i
17.65	97.485
17.65	95.406
22.06	121.200
22.06	121.968
26.47	142.346
26.47	145.464
*	90.044
26.47	141.835
26.47	135.625
22.06	113.814
22.06	112.890
17.65	89.872
17.65	90.964

The * in the table above corresponds to the unknown preparation.

The sample covariance matrix $\mathfrak{C}$ is :

$$\mathfrak{C} = \begin{bmatrix} 14.144 & 76.791 \\ 76.791 & 431.328 \end{bmatrix}$$

The principle eigenvalue and eigenvector of $\mathfrak{C}$ are:

$$\lambda = 445.015 \qquad \mathbf{e} = \begin{bmatrix} 0.175459 & 0.984487 \end{bmatrix}^T$$

From which we obtain :

$$\tilde{\beta}_1 = \frac{0.984487}{0.175459} = 5.61093$$

$$\tilde{\beta}_0 = 117.406 - (5.61093)(22.060) = -6.37126$$

The orthogonally-calibrated estimate for the unknown concentration whose absorbance is 90.044 is 17.183. Furthermore,

$$\mathbf{V} = \begin{bmatrix} 16.366 & 0 & 0 \\ 0 & 49.780 & -2.182 \\ 0 & -2.182 & 0.099 \end{bmatrix}$$

and for a 95% prediction interval

$$a = 30.991 \; ; \; b = -1060.286; \; c = 8967.479$$

Using equation (5.9) we obtain the limits of the 95% prediction interval are computed to be (15.299,18.914). For comparison we have summarized the results for all three methods:

Method	**Estimator**	**Point est.**	**95% Prediction limits**
Classical:	$\hat{X}_{0C} = \frac{y_0 + 2.3629}{5.4292}$	17.020	(15.090,18.775)
Inverse:	$\hat{X}_{0I} = 1.158 + 0.17803y_0$	17.189	(15.398,18.979)
Orthogonal :	$\hat{X}_{0(orthog)} = \frac{y_0 + 6.37126}{5.61093}$	17.183	(15.299,18.914)

A couple of points are worth noting:

(i) the orthogonal estimate lies between the classical and inverse estimates.
(ii) the 95% prediction interval widths for all three procedures are almost identical.

6.2 Animal Fat

D'Antuono et. al. (1991) describe a calibration experiment in which instrument readings (Y) were obtained on the carcasses of animals having varying fat content (X). The aim of the experiment was to establish a calibration equation so that future fat levels could be predicted from the reading of a probe inserted into a carcass rather than having to dissect the animal. The data are reproduced below:

X_i	4.06	3.94	2.06	4.86	3.31	3.26	3.17	4.60	2.53	4.02	4.78	4.29	2.62	1.74
Y_i	13	13	7	20	8	11	12	15	9	13	12	18	8	7

X_i	3.21	3.06	2.94	3.34	3.61	2.62	2.55	3.56	3.21	3.44	2.23	2.89	3.86	2.61
Y_i	9	12	13	14	9	60	7	14	10	9	6	8	13	7

X_i	1.60	4.25	2.72	1.26	5.05	3.45	4.86	2.49	2.20	3.11	3.30	3.80
Y_i	3	14	7	4	19	13	15	8	6	9	9	13

We will use the first 39 observations for model-fitting, keeping the last observation (3.80,13) for prediction.

The sample covariance matrix $\mathfrak{C}$ is :

$$\mathfrak{C} = \begin{bmatrix} 0.868 & 3.252 \\ 3.252 & 15.783 \end{bmatrix}$$

The principle eigenvalue and eigenvector of $\mathfrak{C}$ are:

$$\lambda = 16.461036 \qquad \mathbf{e} = \begin{bmatrix} 0.204172 & 0.978935 \end{bmatrix}^T$$

From which we obtain :

$$\tilde{\beta}_1 = \frac{0.978935}{0.204172} = 3.749$$

$$\tilde{\beta}_0 = 10.5128 - (3.749)(3.2477) = -1.663$$

The orthogonally-calibrated estimate for the last observation whose reading from the probe was 13 is 3.911. Furthermore,

$$\mathbf{V} = \begin{bmatrix} 4.661 & 0 & 0 \\ 0 & 2.005 & -0.579 \\ 0 & -0.579 & 0.178 \end{bmatrix}$$

and for a 95% prediction interval

$$a = 22.257 \; ; \; b = -168.421; \; c = 298.749$$

Using equation (5.9) we obtain the limits of the 95% prediction interval are computed to be (2.839,4.728). For comparison we have summarized the results for all three methods:

Method	Estimator	Point est.	95% Prediction limits
Classical:	$\hat{X}_{0C} = \frac{y_0 + 1.6639}{3.749}$	3.911	(2.858,5.009)
Inverse:	$\hat{X}_{0I} = 1.081 + 0.20606y_0$	3.760	(2.832,4.689)
Orthogonal :	$\hat{X}_{0(orthog)} = \frac{y_0 + 5.059}{4.795}$	3.766	(2.839,4.728)

The orthogonal and inverse estimates are very similar, although the orthogonal has performed marginally better when compared with the actual value of 3.80.

7. Conclusions

In this paper we have presented a third alternative to the existing methodologies of inverse and classical regression for calibration problems. Simple computational formulae have been provided for orthogonal regression estimates. Furthermore, we have demonstrated the equality of this form of estimation with principal components regression. Procedures for obtaining interval estimates in the orthogonal case have also been provided and these have been illustrated with the use of previously published calibration examples arising from industrial applications.

We believe that the use of orthogonal calibration presents a unified methodology for calibration problems. Furthermore we suggest that orthogonal calibration should become the de facto approach for practitioners grappling with the dilemma of deciding whether to regress X on Y or Y on X. It is argued that if it is legitimate to perform either regression then it is equally valid to use a procedure which guarantees that the resulting estimate will lie between the estimates produced by the classical and inverse methods. The advantage of our procedure is that it finally lays to rest the vexing question of "are there two regressions?".

References

Anderson, T.W. (1963) *"Asymptotic theory for Principal Component Analysis".* Ann. Math. Statist. 34 122-148.

Berkson, J. (1950) *"Are there two regressions?"* Am. Statist. Assoc. 45 164-180.

Berkson, J. (1969) *"Estimation of a Linear Function for a Calibration Line: Consideration of a Recent Proposal"*. Technometrics, 11 649-660.

Chow, S. and Shao, J. (1990) *"On the Difference Between the Classical and Inverse Methods of Calibration"*. Appl. Statist., 39, 2 219-228.

D'Antuono, M.D., Cabassi, P., and Fox, D.R. (1991) *"Estimating Fat Levels in Carcasses or Probing the Probes?"*. Proc. Statcomp/Biostats 1991 Meeting, Queensland, Australia.

Fieller, E.C. (1944) *"A Fundamental Formula in the Statistics of Biological Assay and Some Applications"*. Quarterly Journal of Pharmacy and Pharmacology, 17 117-123.

Fox, D.R. (1990) *"Interval Estimation in Calibration Models: An Application of Fieller's Theorem"*. Commun. Statist.- Theory Meth. 20 , 2 711-720.

Fuller, W.A. (1987) Measurement Error Models. Wiley, New York.

Krutchkoff, R.G. (1967) *"Classical and Inverse Methods of Calibration"*. Technometrics, 9, 425-439.

Krutchkoff, R.G. (1969) *"Classical and Inverse Regression Methods of Calibration in extrapolation"*. Technometrics, 11, 605-608.

Lwin, T. and Maritz, J.S. (1980) *"A Note on the Problem of Statistical Calibration"*. Appl. Statist, 29, (2) 135-141.

Mardia, K.V., Kent, J.T., and Bibby, J.M. (1979) Multivariate Analysis. Academic Press, London.

Rao, C.R. (1973) Linear Statistical Inference and its Applications. 2nd. ed. Wiley, New York.

Williams, E.J. (1969) *"A Statistical Theory of Calibration"*. Ann. Statist. 1, 1-37.

Zerbe, G.O. (1978) *"On Fieller's Theorem and the General Linear Model"*. The Amer. Statist. 32 103-105.