

Statistical Calibration by means of Kalman Filter¹

La Taratura Statistica tramite il Filtro di Kalman

S. Salini¹, A. Tiano², A. Zirilli²

¹ Istituto di Statistica

Università Cattolica del Sacro Cuore di Milano

(e-mail: salini.tutor@mi.unicatt.it)

² Dipartimento di Informatica e Sistemistica

Università degli Studi di Pavia

(e-mail: : antonio@control1.unipv.it, zirilli@control1.unipv.it)

Riassunto: Il problema della taratura di uno strumento di misura può essere inquadrato sia in un contesto statistico che in un contesto ingegneristico. Nel primo, il problema viene affrontato distinguendo tra regressione inversa e regressione diretta. I modelli che si ottengono sono modelli statici e sono utilizzati per risalire alle misure ‘esatte’ partendo da misure ‘affette da errore’. In ambito ingegneristico invece, le variabili vengono considerate nell’ordine temporale in cui si osservano. Il filtro di Kalman può essere quindi utilizzato per monitorare l’evolversi delle misure, introducendo così un modello di taratura dinamica.

Keywords: Statistical Calibration, Controlled Calibration, Bootstrap, Kalman filter

1 Introduction

Statistical calibration has some similarities with scientific calibration. The latter is the process whereby the scale of a measuring instrument is determined or adjusted on the basis of a calibration experiment. The former is a kind of inverse prediction, some authors call it inverse regression rather than calibration. Statistical calibration is broadly used in chemistry, engineering, biometrics and is potentially useful in several practical applications. General reference are Osborne (1991), Brown (1993), Sundberg (1999).

In this paper we consider controlled calibration experiments, made to calibrate a measuring instrument. Traditionally, the calibration model obtained from such experiments is a static model and a new training set (or experiment) is necessary to change or update the model. Using Kalman filter theory, introduced by R.E. Kalman (1960) by extending the notion of state variable to the Wiener filter, it is possible both to monitor the statistics of the classical model and to adjust the actual estimation. Section 2 presents the statement of the calibration problem with some details about linear calibration model. Section 3 describes the Kalman algorithm and section 4 applies the algorithm to the classical calibration approach. In this final section some concluding remarks are also highlighted.

¹ Thanks very much to Professor B.V. Frosini and Professor D. Zappa for their precious advice and to G. Brambilla of ‘Istituto di Acustica O.M. Corbino - CNR Roma’ for giving us the data sets.

2 Statement of the problem

Suppose that two different instruments for the measurement of the same phenomenon are considered. The first one (*standard method*) is more laborious, accurate and expensive than the second one (*test method*). The variable associated to the measures obtained by the two instruments are indicated by X and Y respectively. A sample of n units, in which both measures x and y are available, is considered. The set of values (x_i, y_i) $i=1, \dots, n$ is the calibration experiment. The statistical calibration problem arises when only the y_i obtained by the test method are known and the unknown x_i have to be estimated. The solution of this problem depends on the hypotheses on the probabilistic model, supposed to have generated the calibration experiment. In particular, it is assumed that the values y_i are realizations of a random variable (r.v.) Y with known density function.

The assumptions on the values x_i may be of two types: i) the x_i are realizations of a r.v. and therefore (x_i, y_i) are realizations of the bivariate r.v. (*natural calibration*) ii) the x_i are chosen from the experimenter (*controlled calibration*).

In this paper it is supposed that the relation estimated through the data of the calibration experiment between Y and X is:

$$Y = a + bX + e \quad e \sim N(0, s^2) \quad \text{that is} \quad E(Y) = a + bX \quad (2.1)$$

The (2.1) is the univariate linear calibration model. The point estimators proposed in literature for the unknown value of the measure x_i are the classical estimator $\hat{x}_C$ (Eisenhart, 1939) and the inverse estimator $\hat{x}_I$ (Krutchkoff, 1967).

Let

$$\hat{Y} = \hat{a} + \hat{b}X \quad (2.2)$$

where $\hat{a}$ and $\hat{b}$ are the last squared error estimators of a and b , the classical estimator is

$$\hat{x}_C = \frac{y - \hat{a}}{\hat{b}} \quad (2.3)$$

Under normality may be easy shown that $\hat{x}_C$ is the maximum likelihood estimator, but it is not the unbiased estimator with minimum variance. Moreover as $\hat{x}_C$ is the ratio of two independent normal r.v. the variance of $\hat{x}_C$ may be not finite.

The inverse estimators $\hat{x}_I$ is obtained from the estimated inverse regression model

$$\hat{X} = \hat{g} + \hat{d}Y \quad (2.4)$$

by supposing the relation

$$X = g + dY + e \quad e \sim N(0, s^2) \quad \text{that is} \quad E(X) = g + dY \quad (2.5)$$

The estimator $\hat{x}_I$ has a finite variance, but it comes from the equation (2.4) under the not appropriate assumption that the errors e_i and the variable y_i are independent. This is a natural calibration situation.

3 Kalman filter

Given the measure y_k at time t_k , it can be estimate the state x_k at time t_k of a stochastic system. It is assumed that:

i) the state system's dynamics is governed by a linear difference equation of the form:

$$x_{k+1} = Ax_k + h_k \quad (3.1)$$

where A is the state transition matrix and h_k is the state noise;

ii) the measurement is linearly related to the state through an equation of the form:

$$y_k = Hx_k + v_k \quad (3.2)$$

where H is the measurement sensitivity matrix and v_k is the measurement noise.

Equation (3.2), known as the measurement equation, describes how a measurement vector y depends on the state vector x and a noise v vector at time k . Equation (1) describes the time evolution of the system's state and in both equations matrix A and H are design matrixes (supposed to be known). When measurement errors v and system noise h are neither auto nor cross-correlated, they do not correlate with the initial state x_0 . Let

$$Q_k = \text{Cov}(h_k) = E\left(h_k, h_k'\right)$$

$$R_k = \text{Cov}(v_k) = E\left(v_k, v_k'\right)$$

the covariance matrixes of h_k and v_k . It is possible to prove that the Kalman state estimator for the system of equations (3.1) and (3.2) is the best linear unbiased estimator (BLUE) and is given by:

$$\hat{x}_k = \tilde{x}_k + K_k n_k \quad (3.3)$$

where

a) $\tilde{x}_k = A\hat{x}_{k-1}$ is an a-priori state estimation based on the knowledge of the state transition matrix A and the state estimation at the previous time;

b) n_k is the innovation defined as:

$$n_k = y_k - H A \hat{x}_{k-1} \quad (3.4)$$

c) K_k is an optimal gain also known as the Kalman filter gain defined as:

$$K_k = \tilde{P}_k H' (H \tilde{P}_k H' + R_k)^{-1} \quad (3.5)$$

where:

$$\tilde{P}_k = \text{Cov}(\tilde{x}_k - x_k) = E\left((\tilde{x}_k - x_k)(\tilde{x}_k - x_k)'\right) = A \hat{P}_{k-1} A' + Q \quad (3.6)$$

and

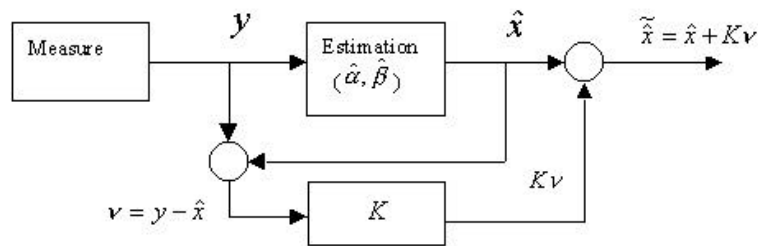
$$\hat{P}_k = \text{Cov}(\hat{x}_k - x_k) = E\left((\hat{x}_k - x_k)(\hat{x}_k - x_k)'\right) = \tilde{P}_k - K_k H \tilde{P}_k \quad (3.7)$$

Kalman filter solves the problem of recursively estimate the state of a dynamic system given its measures. In other words Kalman filter includes the solution of an *inversion* problem, in which one knows how to represent the measurable variables as function of explanatory variables. It inverts this functional relationship and estimates the independent variables as inverted function of the dependent (measurable) variables.

4 Kalman filter for statistical calibration

The “classical” calibration problem defined in section 2 can be studied and solved in the framework of Kalman filter theory where the estimator is proved to be BLUE. Under the hypothesis of error measurement correlation (3.6) and (3.7) we may apply the Kalman filter to upgrading the statistical information relative to the classical estimator. By this upgrading the classical estimation can be adjusted in order to give a posterior estimate. Figure 1 summarises the idea of using the Kalman filter to adjust the prior estimation given by the classical estimator.

Figure 1. Process of upgrading the classical estimator by using Kalman procedure



By substitution of the prior estimation with the classical estimator in (3.6), equation (3.7) provides the upgrading of the a-posterior covariance error of the classical estimator. In other words, (3.7) is an estimation that can be used to update the upcoming estimation. Equations (3.3), (3.4) and (3.5) provides an iterative procedure to compute the upgrading. In this way the estimation process is expected to be more robust with respect the initial statistics used to identify the calibration model. In the final paper a numerical example is presented in which it is shown that the proposed estimator (sketched in figure 1) improves the robustness of the classical estimator. Moreover by observing the behaviour of Kalman and classical estimator for 1000 bootstrap samples it appears that the distribution of residuals of Kalman estimator is nearer to the Gaussian model.

Reference

- Brown, P. J. (1993) *Measurement, Regression and Calibration*, Oxford University Press, Oxford.
- Eisenhart, C. (1939) The interpretation of certain regression methods and their use in biological and industrial research, *Annals of Mathematical Statistics*, 10, 162-186.
- Frosini B.V. (1994) Regressione Diretta e Inversa: Problemi di Taratura e Applicazioni Legali, *Statistica Applicata Vol. 6*, n. 2
- Kalman, R. E. (1960) A new Approach to Linear Filtering and Prediction Problems. *Trans. ASME J. of Basic Engineering*.
- Krutchkoff, R. G. (1967) Classical and inverse regression methods of calibration, *Thecnometrics*, 9, 425-439.
- Osborne, C. (1991) Statistical calibration: a review. *International Statistical Review*, 59, 309-336.
- Sundeberg, R. (1999) Multivariate calibration – direct and indirect regression methodology (with discussion), *Scandinavian Journal of Statistics*, 26, 161-207.