RM-75-64

CREDIBILITY THEORY AND KALMAN FILTERING WITH EXTENSIONS

R. K. Mehra

November 1975

Credibility Theory and Kalman Filtering with Extensions

R. K. Mehra*

Introduction

Credibility theory developed in actuarial mathematics involves the estimating of risk premiums based on the collective and individual risk data. Though relatively old in origin [15], it has only recently been put on a sound mathematical foundation [1,2,10]. A number of extensions of the theory have also appeared recently [3,7] under the titles of "multidimensional credibility theory," "time-inhomogeneous credibility theory" and "evolutionary credibility theory."

In this paper it is shown that the estimate and prediction problems considered in credibility theory are similar to those considered in Kalman filtering theory. The state vector of the risk model consists of the average risk variables such as the number of claims, cost of claims, etc. for a particular risk θ from the collective of risks Θ. The observed data consists of the risk variables for the collective and for the individual risks. It is required to predict the values of risk variables in the next time period based on this data and to adjust the individual premiums based on claims experience in such a way as to converge to their true values.

The application of Kalman filtering, innovation theory and maximum likelihood estimation [12] gives results more general than those obtained before. It also opens up the possibility for considering more general models and testing their validity on real data. On the other hand, credibility theory provides certain specialized results which have not been fully appreciated in the Kalman filtering literature. In particular, it has been shown in credibility theory that the same prediction formulae hold under a variety of measurement noise distribution functions such as exponential, poisson, binomial, etc. [6]. Thus, the connection between credibility theory and Kalman filtering is significant and may be expected to lead to further developments in both areas.

---

*Harvard University, Cambridge, Massachusetts, and IIASA, Laxenburg, Austria.

I. Problem Formulation

Let θ denote the risk parameter that varies from one individual to the other. Let $U(\theta)$ be the prior distribution function of θ over the whole population of individuals.

Let $\xi_1, \xi_2, \ldots, \xi_t, \ldots, \xi_n$ be the experience observation vectors on an individual with a specific θ. Prediction of ξ_{n+1} based on the population statistics and the experience observations $\{\xi_1, \ldots, \xi_n\}$ is required.

Assume that $\xi_t(\theta)$ is normally distributed with mean $m_t(\theta)$ and covariance $C_t(\theta)$ for a specific individual. We may write

$$\xi_t(\theta) = m_t(\theta) + v_t(\theta) \tag{1}$$

where $v_t(\theta)$ is a white noise sequence, normally distributed with zero mean and covariance $C(\theta)$ for a given θ. Thus $\xi_1(\theta), \ldots, \xi_n(\theta)$ are conditionally independent. Further let the expected values of $m_t(\theta)$ and $C_t(\theta)$ over the population of all θ be m_t and E_t:

$$m_t = \underset{\theta}{E}(m_t(\theta))$$

$$E_t = \underset{\theta}{E}[C_t(\theta)]$$

and

$$D_t = \underset{\theta}{E}\{(m_t(\theta) - m_t)(m_t(\theta) - m_t)^T\} \quad .$$

The Kalman filter formulation [15] requires setting up a state equation and a measurement equation. Equation (1) is essentially the measurement equation with $m_t(\theta)$ representing the "signal" part and $v_t(\theta)$ representing the "noise" part. The evolution of the "signal" part is described by the state equation, which in the most general form can be taken to be,

$$x_{t+1} = F_t x_t + G_t u_t \tag{2}$$

$$m_t(\theta) = H_t x_t \tag{3}$$

$$x_o \sim N(\bar{x}_o, P_o) \quad .$$

Here x_t denotes the state of the system which can have dimension greater or less than $m_t(\theta)$ and u_t is a normally distributed white noise sequence with known mean (say zero) and covariance Q_t.

The model (2)-(3) is perhaps too general for the "credibility problem," but it is stated here since the Kalman filter is valid for this general model. The model (2)-(3) may be specialized to various cases considered in Jewell [7] as follows:

i) <u>Time-Homogeneous Risks Submodel</u>: Let

$$F_t = I \quad , \qquad u_t \equiv 0 \quad , \qquad H_t = I$$

then

$$m_{t+1}(\theta) = m_t(\theta)$$

or

$$m_t(\theta) = m_o(\theta) \quad .$$

Also since $m_o(\theta)$ is normally distributed with mean m and covariance D,

$$\bar{x}_o = m \tag{4}$$

$$P_o = D \quad . \tag{5}$$

Furthermore E_t is assumed constant (= E).

ii) <u>Time-Dependent Risk Model</u>: $u_t \equiv 0$, $H_t = I$, so that $m_{t+1}(\theta) = F_t m_t(\theta)$. The other quantities same as above

iii) Evolutionary Risk Model: $F_t = I$, $H_t = I$, $G_t = I$ so that

$$m_{t+1}(\theta) = m_t(\theta) + u_t \tag{6}$$

with x_o, P_o, and E_t the same as in i). Now we state the Kalman filter equations and show how they simplify to the "credibility theory" results.

II. Kalman Filtering: [5,8,9]

Let $\hat{x}_{t/\tau}$ denote the best estimate of x_t based on the observation $\{\xi_1,\ldots,\xi_\tau\}$ and let $P_{t/\tau}$ be its covariance. From least squares theory,

$$\hat{x}_{t/\tau} = E\{x_t|\xi_1,\ldots,\xi_\tau\} \tag{7}$$

$$P_{t/\tau} = E\{(x_t - \hat{x}_{t/\tau})(x_t - \hat{x}_{t/\tau})^T|\xi_1,\ldots,\xi_\tau\} \quad . \tag{8}$$

Then the "filtered" estimates $\hat{x}_{t|t}$ and the "predicted" estimates $\hat{x}_{t+1|t}$ can be obtained recursively as follows:

$$\hat{x}_{t+1|t} = F_t \hat{x}_{t|t} \tag{9}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(\xi_t - H_t \hat{x}_{t|t-1}) \tag{10}$$

$$K_t = P_{t|t-1}H_t^T(H_tP_{t|t-1}H_t^T + E_t)^{-1} \tag{11}$$

$$P_{t|t} = (I - K_tH_t)P_{t|t-1} \tag{12}$$

$$P_{t+1|t} = F_tP_{t|t}F_t^T + G_tQ_tG_t^T \tag{13}$$

$$\hat{x}_{o|o} = \bar{x}_o \quad , \qquad P_{o|o} = P_o \quad .$$

Equations (11)-(12) may also be written as

$$K_t = P_{t|t} H_t^T E_t^{-1} \tag{14}$$

$$P_{t|t}^{-1} = P_{t|t-1}^{-1} + H_t^T E_t^{-1} H_t \tag{15}$$

which will be useful in relating to credibility theory. Equations (11)-(13) may be combined into a single recursive equation for $P_{t|t}$ or $P_{t|t-1}$, e.g.

$$P_{t+1|t} = F_t\left\{P_{t|t-1} - P_{t|t-1}H_t^T(H_tP_{t|t-1}H_t^T + E_t)^{-1}H_tP_{t|t-1}\right\} \cdot F_t^T + G_tQ_tG_t^T \quad . \tag{16}$$

Equation (16) is a discrete-time Riccati equation which has been studied extensively in the control literature [5]. The sequence $\xi_t - H_t x_{t\ t-1}$ known as the "innovation sequence" is a Gaussian white noise sequence [8]. We now consider the first special case of credibility theory.

i) Time-Homogeneous Risk Model: Substituting for F_t, H_t etc., in Equation (9)-(16),

$$\begin{aligned} \hat{m}_{t+1|t} &= \hat{m}_{t|t} \\ &= \hat{m}_{t|t-1} + K_t(\xi_t - \hat{m}_{t|t-1}) \end{aligned} \tag{17}$$

$$K_t = P_{t|t}E^{-1}$$

$$P_{t|t}^{-1} = P_{t|t-1}^{-1} + E^{-1} = P_{t-1|t-1}^{-1} + E^{-1}$$

$$P_o = D \quad , \qquad \bar{x}_o = m \quad .$$

Thus

$$P_{n|n}^{-1} = D^{-1} + nE^{-1}$$

$$K_n = (D^{-1} + nE^{-1})^{-1}E^{-1}$$

$$= (ED^{-1} + nI)^{-1} \quad ; \qquad (18)$$

$$\hat{m}_{n+1|n} = \hat{m}_{n\ n-1} + (ED^{-1} + nI)^{-1}(\xi_n - \hat{m}_{n|n-1})$$

$$= (I - (ED^{-1} + nI)^{-1})\ \hat{m}_{n|n-1} + (ED^{-1} + nI)^{-1}\xi_n$$

$$= (ED^{-1} + nI)^{-1}(ED^{-1} + (n-1)I)\ \hat{m}_{n|n-1} + (ED^{-1} + nI)^{-1}\xi_n$$

$$= (ED^{-1} + nI)^{-1}(ED^{-1} + (n-1)I)\{(ED^{-1} + (n-1)I)^{-1}(ED^{-1} + (n-2I)$$

$$\hat{m}_{n-1|n-2} + (ED^{-1} + (n-1)I)^{-1}\xi_{n-1}\} + (ED^{-1} + nI)^{-1}\xi_n$$

$$= (ED^{-1} + nI)^{-1}(ED^{-1} + (n-2)I)\ \hat{m}_{n-1|n-2}$$

$$+ (ED^{-1} + nI)^{-1}(\xi_n + \xi_{n-1}) \quad .$$

By induction,

$$m_{n+1\ n} = (ED^{-1} + nI)ED^{-1}\ m + (ED^{-1} + nI)^{-1} \sum_{t=1}^{n} \xi_t$$

$$= (I - Z_n)m + Z_n(\frac{1}{n} \sum_{t=1}^{n} \xi_t) \qquad (19)$$

where

$$Z_n = (\frac{1}{n}\ ED^{-1} + I)^{-1} = \frac{1}{n}\ K_n \quad \text{(Credibility Factor)}. \qquad (20)$$

Equation (19) is the same as Equation (4.26) of Jewell [7]. Kalman filter results for models ii) and iii) also lead to parallel results in credibility theory. We do not pursue this approach here since from a computational view point, the recursive Equations (9)-(13) are easier to solve.

III. Extensions

In credibility theory as well as in several applications of Kalman filtering, the matrices F, G, H, Q and R contain unknown parameters. This problem has been studied extensively in the control literature [11,12,13,14] and computer programs exist for obtaining least squares and maximum likelihood estimates of these parameters. The need for such methods in credibility theory has already been indicated by Bühlmann [4].

An objection to the use of Kalman filtering for credibility theory is the assumption of normality. In the non-normal additive noise case, the Kalman filter still provides optimal linear least squares estimates. Several nonlinear estimating techniques also exist (Jazwinski [5]) and may be used for problems in credibility theory. Another approach is to transform the data in such a way that the transformed variables have normal distributions. For example, log normal distributions can be handled in this fashion.

Conclusions

A connection has been established between problems and results from two different fields viz: Kalman filtering and credibility theory of actuarial mathematics. A number of extensions of credibility theory are indicated based on results from Kalman filtering, innovation theory, and maximum likelihood identification.

References

[1] Bailey, A.L. "A Generalized Theory of Credibility." Proceedings of the Casualty Actuarial Society, Vol. 32. 1945, pp. 13-20.

[2] Bailey, A.L. "Credibility Procedures, Laplace's Generalization of Bayes' Rule and the Combination of Collateral Knowledge with Observed Data." Proceedings of the Casualty Actuarial Society, Vol. 37. 1950, pp. 7-23.

[3] Bühlmann, H. "Credibility Procedures." Sixth Berkeley Symposium. 1971, pp. 515-525.

[4] Bühlmann, H. Mathematical Methods in Risk Theory, New York, Springer-Verlag, 1970.

[5] Jazwinski, A.H. Stochastic Processes and Filtering Theory, New York, Academic Press, 1970.

[6] Jewell, W.S. "Credible Means are Exact Bayesian for Simple Exponential Families." ORC-73-21. Operations Research Center. University of California, Berkeley, October 1973 (to appear in ASTIN Bulletin).

[7] Jewell, W.S. "Model Variations in Credibility Theory." ORC-74-25. Operations Research Center. University of California, Berkeley, August 1974.

[8] Kailath, T. "An Innovations Approach to Least-Squares Estimation, Part 1." IEEE T-AC, AC-13 (December 1968).

[9] Kalman, R.E. "New Methods and Results in Linear Prediction and Filtering Theory." Proc. Symp. on Engg. Appl. of Random Function Theory and Probability. New York, Wiley, 1961.

[10] Mayerson, A.L. "A Bayesian View of Credibility." Proceedings of the Casualty Actuarial Society, Vol. 51. 1964, pp. 85-104.

[11] Mehra, R.K. "Approaches to Adaptive Filtering." IEEE T-AC (October 1972).

[12] Mehra, R.K. "Identification in Control and Econometric Systems, Similarities and Differences." Annals of Economic and Social Measurements (January 1974).

[13] Mehra, R.K. "Identification of Stochastic Linear Dynamic Systems Using Kalman Filter Representation." AIAA Journal (January 1971).

[14] Mehra, R.K. "On Line Identification of Linear Dynamic Systems with Applications to Kalman Filtering." IEEE T-AC, Vol. AC-16 No. 1 (February 1971).

[15] Whitney, A.W. "The Theory of Experience Rating." Proceedings of the Casualty Actuarial Society, Vol. 4. 1918, pp. 274-292.