K

k. *See* QUANTITY THEORY OF MONEY.

Kalman filter. The Kalman filter is a technique to use noisy measurements of an unobservable vector of random variables to make an optimal estimate of the value of that vector, when the vector evolves according to a dynamic linear equation. I first describe some applications, then sketch some algebra.

There are two broad classes of applications. One is in algorithms for maximum likelihood estimation and forecast of time-series models. The Kalman filter is often computationally convenient here; see Harvey (1989) for many illustrations.

The second class of applications is to economic models in which variables are unobservable, and only measured with noise. One possibility is that all the relevant variables are observable to economic agents but not to the research economist. Hamilton (1985), for example, assumes essentially that market expectations of US inflation and the real rate on government bonds follow a vector autoregressive process (the 'transition equation', in the terminology defined below). Economists such as Hamilton observe only the realizations of these variables, which, under rational expectations, differ from the unobserved market expectations by a serially uncorrelated error (the 'measurement equation'). Hamilton uses the Kalman filter to estimate the market expectations.

Another possibility is that some relevant variables are unobservable to economic agents, who use the Kalman filter to estimate the values of these variables. Lin, Engle and Ito (1991), for example, assume that stock returns in Japan and the US respond to both local and global factors (the measurement equation); these factors are unobserved. Investors apply the Kalman filter to extract estimates of the local and global factors, which are used to forecast stock returns and volatility.

Formally, it is assumed that the system in question may be written in *state-space form*, evolving according to a pair of equations,

$$x_{t+1} = Ax_t + Bu_t + \varepsilon_{1t},\ \varepsilon_t \sim \text{iid}\, N(0,V_1), \qquad (1)$$

$$y_t = Cx_t + \varepsilon_{2t},\ \varepsilon_{2t} \sim \text{iid}\, N(0,V_2). \qquad (2)$$

Equation (1) is the *transition equation*, equation (2) is the *measurement equation*, x_t is an $(s\times 1)$ unobserved *state vector* that one wishes to estimate, and y_t is the $(m\times 1)$ noisy, observed measurement of x_t. The unobserved disturbances ε_{1t} $(s\times 1)$ and ε_{2t} $(m\times 1)$ are uncorrelated both with each other and with x_0, the presample value of x_t. The $(c\times 1)$ vector u_t follows a known sequence. In most applications in which $B\neq 0$, u_t is a *control vector* chosen to influence the path of x_t optimally. The analyst starts with a prior on the mean and variance of the presample values of x_0, $x_0 \sim N(x_0|_0, \Sigma_0|_0)$. A, B and C, which obey some technical conditions, are dimensioned commensurately with x_t, y_t and u_t. The algebra below can be generalized to allow A, B and C to vary with time, and to allow correlation between ε_{1t} and ε_{2t}; see Anderson and Moore (1979) and Harvey (1989).

The aim is to use observations on $\{y_j\}$, $j = 1,\ldots,t$, to compute estimates of x_t, $t = 1,\ldots,T$. Let $x_t|_n$ denote $E(x_t|y_n, y_{n-1}, \ldots, y_1, u_n, u_{n-1},\ldots,u_1)$, $\Sigma_t|_{t-1}$ the estimate of $E(x_t - x_t|_{t-1})\ (x_t - x_t|_{t-1})'$ available at time $t-1$, with an analogous definition for $\Sigma_t|_t$. Given that all variables are normally distributed, one could in principle use standard formulas for conditional expectations of normally distributed random variables (equivalently, formulas for signal extraction) to solve for $x_t|_t$. (For example, if $A = B = 0$ and $s = m = 1$, so that all variables are scalars, it is easy to see that $x_t|_t = [CV_1/(C^2V_1 + V_2)]y_t$.) But such computations (illustrated in Bertsekas 1976) quickly get very complicated in the general case. Kalman's contribution (1960, cited in Harvey 1989) was to note that the computational burden is greatly lessened by doing the computations recursively, first using $x_0|_0$ and $\Sigma_{0|0}$ to get $x_{1|0}$ and $\Sigma_{1|0}$, then using these to get $x_1|_1$ and $\Sigma_{1|1}$, and so on.

We have

$$\begin{aligned} x_t|_t &= x_t|_{t-1} + E[(x_t - x_t|_{t-1})|(y_t - y_t|_{t-1})] \\ &= x_t|_{t-1} + D_t(y_t - y_t|_{t-1}) \\ D_t &\equiv \{[E(y_t - y_t|_{t-1})(y_t - y_t|_{t-1})']^{-1} \\ &\quad [E(y_t - y_t|_{t-1})(x_t - x_t|_{t-1})']\}'. \end{aligned} \qquad (3)$$

It follows from (1) and (2) that $x_t|_{t-1} = Ax_{t-1}|_{t-1} + Bu_{t-1}$, $y_t|_{t-1} = Cx_t|_{t-1}$. D_t can be derived with routine but tedious calculations, yielding

$$x_t|_t = Ax_{t-1}|_{t-1} + Bu_{t-1} + \Sigma_t|_{t-1}C'(C\Sigma_t|_{t-1}C' + V_2)^{-1} [y_t - C(Ax_{t-1}|_{t-1} + Bu_{t-1})], \qquad (4a)$$

$$\Sigma_t|_{t-1} = A\Sigma_{t-1}|_{t-1}A' + V_1, \qquad (4b)$$

$$\Sigma_t|_t = \Sigma_t|_{t-1} - \Sigma_t|_{t-1}C'(C\Sigma_t|_{t-1}C' + V_2)^{-1}C\Sigma_t|_{t-1}. \qquad (4c)$$

Equations (4a)–(4c) can be used to solve recursively forward for $x_t|_t$ from $t = 0$. Under suitable conditions, as $t\to\infty$, $\Sigma_t|_{t-1}$ approaches a constant matrix, say, Σ, which is the unique positive semidefinite solution to the 'Riccati equation' $\Sigma = A[\Sigma - \Sigma C'\ (C\Sigma C' + V_2)^{-1}C\Sigma]A' + V_1$; $\Sigma_t|_t$ and D_t approach the constant values implied by (4a) and (4c).

If the normality assumption is dropped in favour of an alternative finite variance distribution, equations (4) describe an estimator of $x_t|_t$ that is optimal from the point of view of a certain mean-squared error criterion (Anderson and Moore 1979).

Equations (4) above give the *filtered* estimates of x_t – those computed using observations on y_s, $s \leq t$. There are analogous formulas for the *smoothed* estimates of x_t, computed from a sample of size $T > t$ using y_s, $1 \leq s \leq T$, or *predicted* values of x_t, computed using y_s, $s < t$ (Anderson and Moore 1979; Harvey 1989).

KENNETH D. WEST

See also BAYESIAN INFERENCE IN TIME SERIES; PREDICTION; SIGNAL EXTRACTION; STATISTICAL INFERENCE IN TIME SERIES; VECTOR AUTOREGRESSION METHODS; VOLATILITY.

BIBLIOGRAPHY

Anderson, B.D.O. and Moore, J.B. 1979. *Optimal Filtering.* Englewood Cliffs, New Jersey: Prentice-Hall.

Bertsekas, D. 1976. *Dynamic Programming and Stochastic Control.* New York: Academic Press.

Hamilton, J.D. 1985. Uncovering financial market expectations of inflation. *Journal of Political Economy* 93: 1224–41.

Harvey, A.C. 1989. *Forecasting, Structural Time Series, and the Kalman Filter.* Cambridge: Cambridge University Press.

Kalman, R.E. 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering, Transactions ASME* Series D, 35–45.

Lin, W., Engle, R.F. and Ito, T. 1991. International transmission of stock volatility – an empirical reappraisal. Manuscript, University of Wisconsin.