Appendix A

Kalman Filter

OPTIMAL estimation theory has a very broad range of applications which vary from estimation of river flows to satellite orbit estimation and nuclear reactor parameter identification. In this appendix we present an informal description of the Kalman filter, which is one of the basic tools stemming from estimation theory. We begin with a short definition of the optimal estimation domain to indicate the role of the Kalman filter.

According to (Gelb, 1974) *"an optimal estimator is a computational algorithm that processes measurements to deduce a minimum error (in accordance with some stated criterion of optimality) estimate of the state of a system by utilizing: knowledge of system and measurement dynamics, assumed statistics of system noises and measurement errors, and initial condition information."* The key feature of this formulation is that all measurements and the knowledge about the system are used to evaluate the estimate and that the estimation error is minimized in a well defined statistical sense. There are three main estimation problem classes: filtering, prediction, and smoothing.

The filtering problem corresponds to the cases where an estimate is needed at the moment of the last measurement. In the case of prediction the estimate is required for an instant after the last measurement. When the time of required estimate is between the first and last measurement the problem falls into the smoothing category. As the name indicates, the Kalman filter provides an optimal estimate for the last measurement instant. It is one of the basic filtering techniques which is applicable for estimation of the state of a linear system. It is also a good example of optimal estimators' capabilities and limitations.

The general structure of the filtering process is given in Figure A.1. Here it is assumed that a linear system model and models characterizing system and measurement errors are available. In this case the Kalman filter processes the measurements to provide an optimal estimate of the system state which minimizes the mean square estimation error. It should be underlined that while the Kalman filter provides the information how to process the measured data, it does not indicate the optimal mesurement schedule. Before describing the Kalman filter in more detail, we will introduce some basic concepts and models used in the filtering procedure.

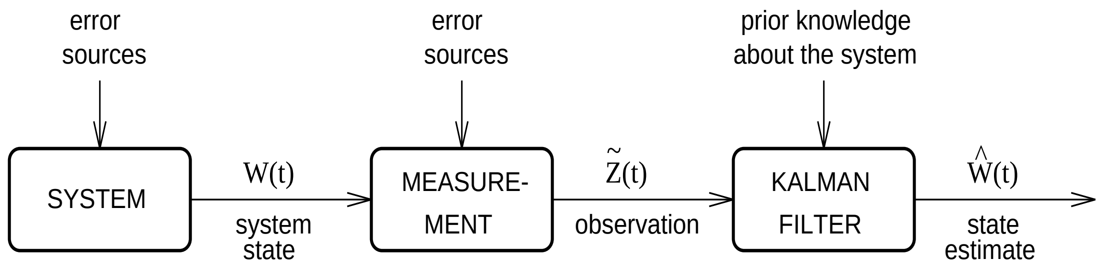

Figure A.1. *General structure of the filtering process.*

A.1 Linear System Model

In general the dynamics of a linear system can be described in either the frequency domain or the time domain. In the following we use the time domain, which is more convenient from mathematical and notational viewpoints. Also, such a description is more natural, which results in a better understanding of the system's behavior.

Let $W(t) = [W^1(t), W^2(t), ..., W^n(t)]^T$ denote the system state vector described by n parameters which are a function of time t. The dynamics of this system can be described by the following first-order differential equation:

$$\dot{W}(t) = F(t)W(t) + G(t)\mathbf{e}(t) + L(t)\mathbf{c}(t) \tag{A.1}$$

where $\mathbf{e}(t)$ is a random forcing function, $\mathbf{c}(t)$ is a control (deterministic) function, and $F(t)$, $G(t)$, $L(t)$ are matrices defining the dynamics of the system. The differential equation determines the system's subsequent behavior assuming that the state vector at a certain point in time and a description of the forcing and control functions are given. A block representation of the linear system dynamics is shown in Figure A.2.

Transition matrix

Let us consider a system without the forcing and control functions:

$$\dot{W}(t) = F(t)W(t) \tag{A.2}$$

For this system one can define the transition matrix $\Phi(t, t_0)$ which defines the system state at a time t based on the knowledge of the state at t_0:

$$W(t) = \Phi(t, t_0)W(t_0) \tag{A.3}$$

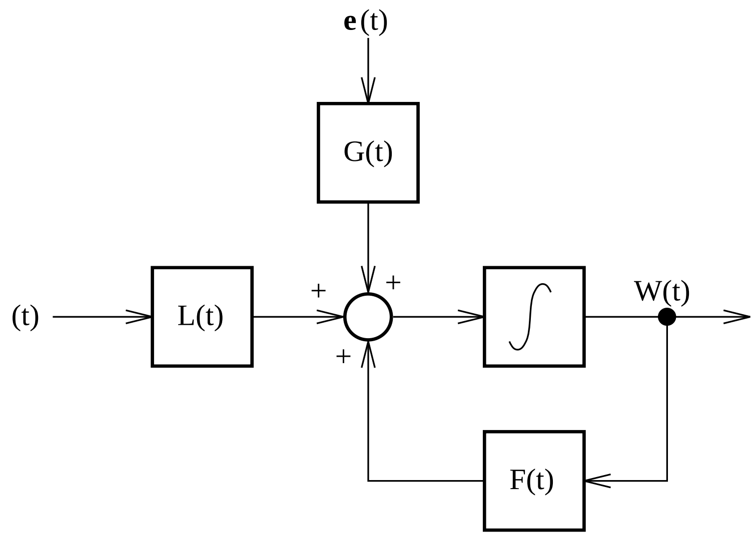

Figure A.2. *Linear system dynamics.*

Obviously the transition matrix is a function of matrix $F(t)$. In particular the following general relations can be derived:

$$\frac{d}{dt}\Phi(t,t_0) = F(t)\Phi(t,t_0) \tag{A.4}$$

$$|\Phi(t,t_0)| = \exp\left[\int_{t_0}^{t} \mathrm{trace}[F(\tau)]d\tau\right] \tag{A.5}$$

In the case of stationary systems the transition matrix only depends on the difference $t-t_0$ and the F matrix is time invariant. This leads to the following definition of the transition matrix:

$$\Phi(t-t_0) = e^{(t-t_0)F} \tag{A.6}$$

Discrete representation

Up to now we considered a continuous time model. Nevertheless, in many cases only discrete points in time, t_k, $k = 1, 2, ...$, are of interest. In this case the system dynamics can be described by the following difference equation:

$$W_{k+1} = \Phi_k W_k + \Gamma_k \mathbf{e}_k + \Lambda_k \mathbf{c}_k \tag{A.7}$$

where

$$\Phi_k = \Phi(t_{k+1}, t_k) \tag{A.8}$$

$$\Gamma_k \mathbf{e}_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1},\tau)G(\tau)\mathbf{e}(\tau)d\tau \tag{A.9}$$

$$\Lambda_k \mathbf{c}_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau) L(\tau) \mathbf{c}(\tau) d\tau \tag{A.10}$$

A block representation of the discrete system dynamics is given in Figure A.3. In the remainder of this appendix we will consider only discrete systems. Obviously, most of the models and features to follow have their corresponding representation in the continuous time domain (Gelb, 1974).

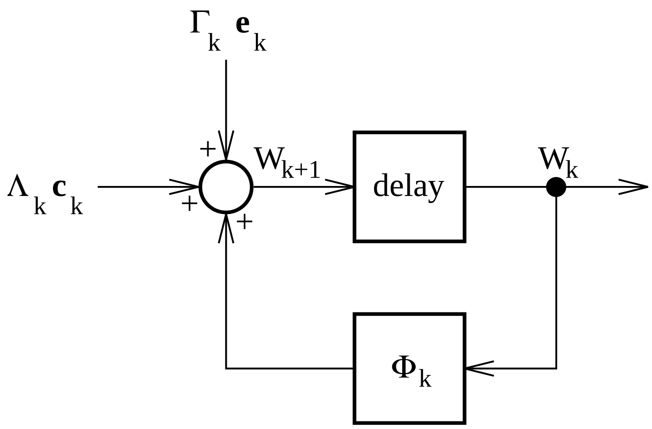

Figure A.3. *Discrete system dynamics.*

A.1.1 Observability and controllability

To discuss observability we introduce the concept of measurements, $\tilde{Z}_k$, $k = 1, 2, ...$. The measurements are assumed to be linearly related to the system state:

$$\tilde{Z}_k = H_k W_k + \mathbf{u}_k \tag{A.11}$$

where H_k is the observation matrix and $\mathbf{u}_k$ is the measurement noise. A system is observable if it is possible to determine $W_1, ..., W_k$ based on corresponding measurements in a noise free environment. A precise observability condition expressed in terms of matrices Φ and H can be found in (Gelb, 1974).

The issue of controllability is related to the ability of achieving an arbitrary state, in a given number of steps, in a deterministic (noise free) linear dynamic system. In particular a system is controllable in time t_k if for any arbitrary pair of states, W_1, W_k, there is a control, $\mathbf{c}_1, ..., \mathbf{c}_k$ which can drive the system from state W_1 to state W_k. A precise controllability condition expressed in terms of matrixes Φ and Λ can be found in (Gelb, 1974). In the remainder of this appendix we do not consider systems with control and we assume $\mathbf{c}_k = 0$.

Introduction of the measurement model completes the basic linear system model description for estimation purpose which is given by

$$W_{k+1} = \Phi_k W_k + \Gamma_k \mathbf{e}_k \tag{A.12}$$

$$Z_k = H_k W_k + \mathbf{u}_k \tag{A.13}$$

It should be mentioned that this model is not unique in the sense that for given system input and output values, there are many different sets of Φ_k, Γ_k, and H_k which will give the same input-output behavior which corresponds to the choice of a coordinate system (time).

A.1.2 Covariance matrix

The concept of covariance matrix is important in estimation error analysis. Let us begin with the definition of cross-covariance matrix C of two vectors, A and B, whose elements are random variables:

$$C(A, B) = E[(A - E[A])(B - E[B])^T] = E[AB^T] - E[A]\ E[B^T] \tag{A.14}$$

If $A = B$, the covariance matrix C defines the second central moments of the vector elements. In particular the matrix diagonal consists of vector elements' variances while other matrix elements are covariances of two vector's elements identified by the matrix element indices.

In this appendix we consider systems whose forcing functions $\mathbf{e}_k$ are vectors of random variables. It is assumed that any two values of forcing function, $\mathbf{e}_k$, $\mathbf{e}_{k-i}$, $i = 1, 2, ...,$ are uncorrelated, which means that the forcing function generates a white sequence. Observe that once the forcing function is a random variable, the system state is also a random variable. To simplify presentation it is also assumed that the forcing function is unbiased (zero ensemble average values). This does not restrict the generality of the presented models since the bias can be easily removed by subtraction.

Let us define the error in the estimate of the system state as a difference between the estimated value $\hat{W}_k$ and the actual value W_k:

$$\Delta_k = \hat{W}_k - W_k \tag{A.15}$$

Then the estimation error covariance matrix is defined as

$$\mathcal{P}_k = E[\Delta_k \Delta_k^T] \tag{A.16}$$

The covariance matrix $\mathcal{P}_k$ expresses the statistical measure of estimation uncertainty.

A covariance matrix is also used for description of the uncorrelated random sequence $\Gamma_k \mathbf{e}_k$. Here we have

$$E[(\Gamma_k \mathbf{e}_k)(\Gamma_k \mathbf{e}_k)^T] = \Gamma_k \mathcal{Q}_k \Gamma_k^T \tag{A.17}$$

where $\mathcal{Q}_k$ is the covariance matrix of the white sequence.

Estimation error propagation

Based on the transition matrix Φ_k one can define the estimate of the predictable portion of the next state as

$$\hat{W}_{k+1}^e = \Phi_k \hat{W}_k \tag{A.18}$$

Henceforth $\hat{W}_k^e$ is called state estimate extrapolation. By subtracting Equation (A.12) from Equation (A.18) we get

$$\Delta_{k+1} = \Phi_k \Delta_k - \Gamma_k \mathbf{e}_k \tag{A.19}$$

This equation can be used to derive a relation for extrapolation of the error covariance matrix from time t_k to t_{k+1}

$$\mathcal{P}_{k+1}^e = \Phi_k \mathcal{P}_k \Phi_k^T + \Gamma_k \mathcal{Q}_k \Gamma_k^T \tag{A.20}$$

This result indicates that in some cases the error covariance can become unbounded if there are no state measurements.

A.2 Discrete Kalman Filter

In this section we consider a linear discrete system whose dynamics are given by

$$W_{k+1} = \Phi_k W_k + \mathbf{e}_k \tag{A.21}$$

where system state W_k is a n-dimensional vector and $\mathbf{e}_k$ is a white sequence vector with zero mean and covariance matrix $\mathcal{Q}_k$. The system measurements are defined by

$$\tilde{Z}_k = H_k W_k + \mathbf{u}_k \tag{A.22}$$

where measurement $\tilde{Z}_k$ is an l-dimensional vector and $\mathbf{u}_k$ is a white sequence vector with zero mean and covariance matrix $\hat{Y}_k$.

Let us define an optimal, unbiased, and consistent estimator. Here optimality is defined as minimization of the mean square estimation error which corresponds to minimization of

$$J_k = E[\Delta_k^T I \Delta_k] = \mathrm{trace}[\mathcal{P}_k] \tag{A.23}$$

where I is identity matrix. An unbiased estimate is defined as the one whose expectation is equal to the expectation of the actual state. A consistent estimate converges to the actual value with the increase in the number of measurements.

The Kalman filter provides an optimal, unbiased, consistent estimate which can be expressed in the linear and recursive form

$$\hat{W}_k = \mathcal{K}'_k \hat{W}_k^e + \mathcal{K}_k \tilde{Z}_k \tag{A.24}$$

where $\mathcal{K}'_k$ and $\mathcal{K}_k$ are weighting matrices. It can be shown (Gelb, 1974) that in order to have the estimate unbiased the following condition must hold:

$$\mathcal{K}'_k = I - \mathcal{K}_k H_k \tag{A.25}$$

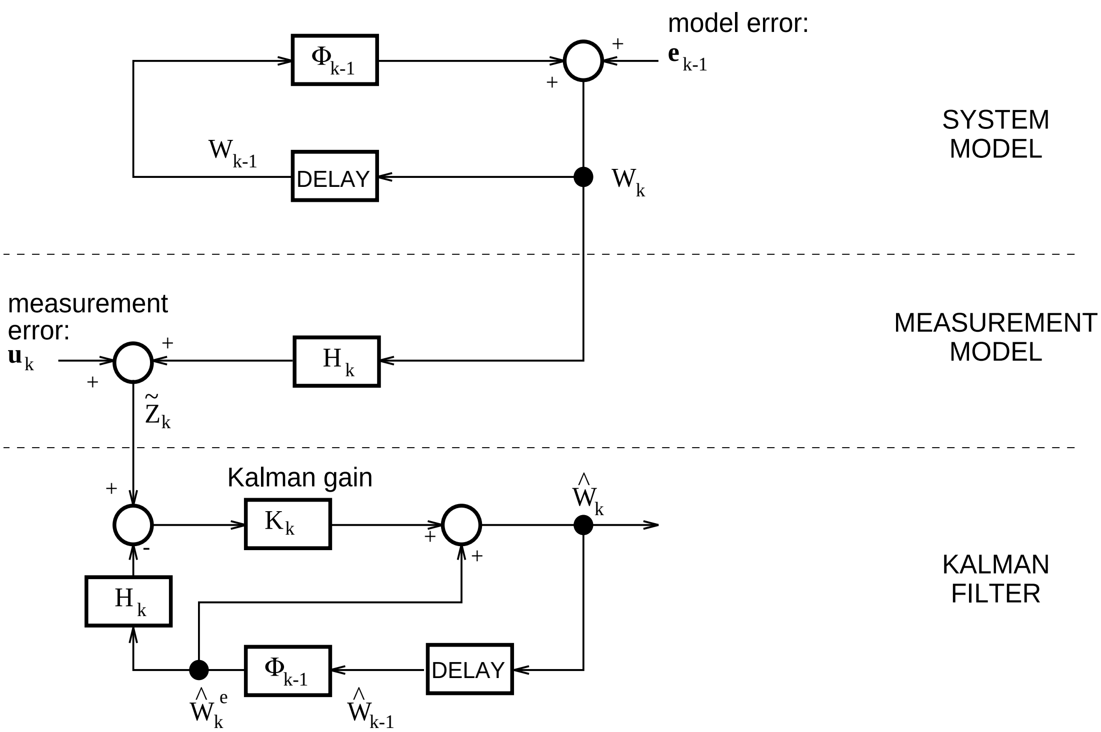

Figure A.4. *System model and Kalman filter.*

Using this relation in Equation (A.24) gives the state estimate update

$$\hat{W}_k = \hat{W}_k^e + \mathcal{K}_k[\tilde{Z}_k - H_k\hat{W}_k^e] \tag{A.26}$$

Based on this relation one can derive the error covariance matrix update

$$\mathcal{P}_k = (I - \mathcal{K}_k H_k)\mathcal{P}_k^e(I - \mathcal{K}_k H_k)^T + \mathcal{K}_k\tilde{Y}_k\mathcal{K}_k^T \tag{A.27}$$

The optimum value of $\mathcal{K}_k$ can be found from minimization of expression A.23 which corresponds to minimization of the length of the estimation error vector. This can be done by evaluating the partial derivative of J_k with respect to $\mathcal{K}_k$ and solving it for zero value. Based on the general relation for the partial derivative of the trace of the product of two matrices the solution gives

$$\mathcal{K}_k = \mathcal{P}_k^e H_k^T[H_k\mathcal{P}_k^e H_k^T + \tilde{Y}_k]^{-1} \tag{A.28}$$

which defines the Kalman gain matrix. Using this matrix in Equation (A.27) defines, after some transformations, the optimized value of the updated error covariance matrix

$$\mathcal{P}_k = (I - \mathcal{K}_k H_k)\mathcal{P}_k^e \tag{A.29}$$

Equations (A.26), (A.28), and (A.29), together with initial conditions, $W_0, \mathcal{P}_0$, define the discrete Kalman filter which is illustrated in Figure A.4. From a practical point of view it is important that the Kalman filter provides its own error analysis by means of the estimation error covariance matrix, $\mathcal{P}_k$. It has been also shown (Weiss, 1970) that, despite its simple recursive nature and linearity, the Kalman filter is the optimal filter if $\mathbf{e}_k$, $\mathbf{u}_k$ are Gaussian (in other words a non-linear filter cannot be better). Otherwise the Kalman filter is the optimal linear filter.

A.3 Discussion and Bibliographic Notes

The first significant contribution to the estimation theory can be traced back to Gauss (circa 1800) who used the technique of deterministic least-squares in simple measurement problems (Mehra, 1970). Fisher (circa 1910) invented the maximum likelihood estimation which is based on probability density function (Weiss, 1970). The design of statistically optimal filters in the frequency domain is due to Wiener (circa 1940) who addressed the continuous time problem using correlation functions and the continuous filter impulse response (Mehra, 1971; Abramson, 1968). The Kalman filter, an optimal linear filter designed in time domain, was developed by Kalman and others; see e.g. (Kalman and Bucy, 1961; Uttam, 1971; Aoki and Huddla, 1967; Tse and Athans, 1970). It is interesting to note that the Kalman filter basically constitutes a recursive solution to the original least-squares problem formulated by Gauss.

In this appendix the Kalman filter presentation follows in principle a description of this technique given in (Gelb, 1974) which provides a simple and interesting picture of the central issues underlying estimation theory and practice. Moreover, it fits very well into the estimation problem treated in Chapter 4. There are many other works dealing with both estimation theory in general and the Kalman filter in particular, e.g. (Papoulis, 1991; Nahi, 1969; Proakis, 1989).

References

Abramson, P.D. Jr. 1968. Simultaneous estimation of the state and noise statistics in linear dynamic systems. Ph.D. Thesis, M.I.T., TE-25.

Aoki, M., and Huddla, J.R. 1967. Estimation of state vector of a linear stochastic system with a constrained estimator. *IEEE Transactions on Automatic Control,* AC-12(4).

Gelb, A. 1974. *Applied Optimal Estimation.* The M.I.T. Press.

Kalman, R.E., and Bucy, R.S. 1961. *Trans. ASME Ser. D.J. Basic Eng.*, 83:95–107. December.

Kalman, R.E. 1962. Fundamental study of adaptive control systems. Technical Report No.ASD-TR-61-27, Vol.I, Flight Control Laboratory, Wright-Patterson Air Force Base, Ohio.

Mehra, R.K. 1970. On the identification of variances and adaptive Kalman filtering. *IEEE Transactions on Automatic Control,* AC-15(2):175–184.

Mehra, R.K. 1971. On-line identification of linear dynamic systems with applications to Kalman filtering. *IEEE Transactions on Automatic Control,* AC-16(1):12–21.

Nahi, N.E. 1969. *Estimation Theory and Applications.* Wiley.

Papoulis, A. 1991. *Probability, Random Variables, and Stochastic Processes.* 3rd ed. McGraw-Hill.

Proakis, J.G. 1989. *Digital Communications.* 2nd ed. McGraw-Hill.

Tse, E., and Athans, M. 1970. Optimal minimal-order observer estimators for discrete linear time-varying systems. *IEEE Transactions on Automatic Control,* AC-15(4).

Uttam, B.J. 1971. On the stability of time-varying systems using an observer for feedback control. *International Journal of Control*, December.

Weiss I.M. 1970. A survey of discrete Kalman-Bucy filtering with unknown noise covariances. In *Proceedings of AIAA Guidance, Control and Flight Mechanics Conference*, Paper No.70–955, Santa Barbara, Calfornia.