The problem addressed by the Kalman-Bucy filter theory is the following one. Given a linear dynamical system driven by noise, construct a “real-time” estimate of the state of the system, based on noisy observations of the output of the system.

The historical root of the problem was in the theory of prediction and extrapolation of stationary processes. This was developed during the 1940s by Norbert Wiener, et. al. The need to develop methods applicable to finite observation records, and nonstationary process models came from attempts to improve aircraft flight control systems in the early 1960s.

The subject area has developed far beyond its origins, and now encompasses various classes of nonlinear, finite state, and distributed parameter systems.

The basic problem can be most easily understood in the finite-dimensional, discrete time case. This setup has the smallest number of technical complications, and also happens to be the case most readily adaptable to digital implementations.

We treat the discrete case in this chapter, and the other cases elsewhere.

### 5.1 The Model

The system model is

\[ x_{k+1} = A_k x_k + B_k u_k, \quad y_k = C_k x_k + v_k. \]

Both \( \{u_k\} \) and \( \{v_k\} \) represent noise processes in the model. \( \{u_k\} \) can be thought of as a disturbance acting on the state, while \( \{v_k\} \) models the observation errors for the output measurements.
The basic problem is to form an estimate of

\[ x_n, \]

given the noisy output observation record

\[ \{y_0, y_1, y_2, \ldots, y_{\tau}\}. \]

There are descriptive labels applied to the problem, depending on the relative position of the estimated variable \( x_n \) and the end of the observation record.

- If \( n < \tau \) the problem is called a smoothing problem.
- If \( n = \tau \) the problem is called a filtering problem.
- If \( n > \tau \) the problem is called a prediction problem.

What distinguishes these problems from general statistical estimation problems is in the first case the underlying state variable model for generation of the processes, and secondly the sequential nature of the observation record.

### 5.2 Estimation Criterion

There are various criteria for quality of statistical estimates. There are Bayesian criteria, maximum entropy estimates, and maximum a posteriori estimates. Among least squares estimates there are general least squares (which in general involve nonlinear processing of the observation data), and least squares estimates constrained to be linear in the observation data. The folklore seems to be that all reasonable criteria lead to the same answer in the case of Gaussian problems, which may account in part for the number of available approaches.

The criterion we choose to use is the following one.

**Criterion:** The estimate generated should be a minimum variance estimate of each component of the state vector \( x_n \) being estimated, and it should be linear in the observation data \( \{y_0, y_1, y_2, \ldots, y_{\tau}\} \).

#### 5.2.1 Easy Non-Answer

With the estimation criterion adopted above, the filtering, smoothing, and prediction problems defined above are in principle all already understood, and hence easy on a conceptual basis.

We are given the data \( \{y_0, y_1, y_2, \ldots, y_{\tau}\} \), and assuming that the system output is an \( r \)-dimensional vector, this collection amounts to \( r(\tau + 1) \) random variables (one for each component of each sample) on which to base the estimate. Since we are looking for the linear minimum mean square estimate, the problem is solved by the projection theorem. It is "only" necessary to compute the projection of the