Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software
Yaakov Bar-Shalom, X. Rong Li, Thiagalingam Kirubarajan

ISBNs: 0-471-41655-X (Hardback) 0-471-22127-9 (Electronic)

Chapter 4

LINEAR DYNAMIC SYSTEMS WITH RANDOM INPUTS

4.1 INTRODUCTION

4.1.1 Outline

This chapter deals with the modeling of linear dynamic systems excited by random inputs, called ***noise***.[1] Continuous-time systems are discussed in Section 4.2. Section 4.3 deals with discrete-time systems.

The state-space models for continuous time and discrete time are presented, and it is shown how the latter can be derived from the former by discretization. The state-space model directly defined in discrete time is also discussed.

The Markov property of the state of a linear system driven by white noise is discussed and used to obtain the propagation equations for the mean and covariance of the state.

The power spectral density (the Fourier transform of the autocorrelation function) of the output of a linear system is related to the state space representation via the transfer function, and it is shown how its factorization makes it possible to prewhiten an autocorrelated random process or sequence.

4.1.2 Linear Stochastic Systems — Summary of Objectives

Present the state-space models for

- continuous-time linear stochastic systems
- discrete-time linear stochastic systems

[1] According to a former program manager at a major Federal research agency, noise is beneficial — it lubricates the system. One can add to this that it also provides opportunities for research. More importantly, it was found via psychological experiments that people will lose their sanity without noise, which is particularly true for experts on stochastic systems (especially those dealing with estimation and tracking).

and the connection between them.

Discuss the implications of the Markov property.

Derive the propagation equations for the mean and covariance of the state of a linear system driven by white noise.

Frequency domain approach — connect the power spectral density with the state space representation.

Show how spectral factorization can be used to prewhiten an autocorrelated random process.

4.2 CONTINUOUS-TIME LINEAR STOCHASTIC DYNAMIC SYSTEMS

4.2.1 The Continuous-Time State-Space Model

The ***state-space representation*** of ***continuous-time linear stochastic systems*** can be written as

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + D(t)\tilde{v}(t) \tag{4.2.1-1}$$

where

x is the ***state vector*** of dimension n_x,
u is the ***input vector*** (control) of dimension n_u,
$\tilde{v}$ is the (continuous-time) input disturbance or ***process noise***, also called ***plant noise***, a vector of dimension n_v,
A, B, and D are known matrices of dimensions $n_x \times n_x$, $n_x \times n_u$, and $n_x \times n_v$, respectively;

A is called the ***system matrix***,
B is the (continuous-time) ***input gain***,
D is the (continuous-time) ***noise gain***.

Equation (4.2.1-1) is known as the ***dynamic equation*** or the ***plant equation***. The state $x(t)$ is a (linear) ***diffusion process*** subject to a ***drift*** due to $u(t)$.

The output of the system is, in general, a vector of dimension n_z

$$z(t) = C(t)x(t) + \tilde{w}(t) \tag{4.2.1-2}$$

where

$\tilde{w}$ is the (unknown) output disturbance or ***measurement noise***, and
C is a known $n_z \times n_x$ matrix, called the ***measurement matrix***.

Equation (4.2.1-2) is known as the ***output equation*** or the ***measurement equation***.

In the absence of the disturbances $\tilde{v}$ and $\tilde{w}$, that is, in the deterministic case, given the initial condition $x(t_0)$ and the input ***function in the interval*** $[t_0, t]$ denoted as

$$u_{[t_0,t]} \triangleq \{u(\tau), t_0 \leq \tau \leq t\} \tag{4.2.1-3}$$

one can compute the future output at any time $t > t_0$

$$z(t) = z[x(t_0), u_{[t_0,t]}, t, t_0] \tag{4.2.1-4}$$

The ***state*** of a deterministic system is defined as the smallest vector that *summarizes the past of the system.*

Any linear differential equation that describes an input-output relationship can be put in the form of a first-order vector differential equation as in (4.2.1-1). For example, an nth-order scalar differential equation can be rewritten as a first-order differential equation for an n-vector — that is, n first-order equations — by a suitable definition of state variables.

The initial conditions of the nth-order differential equation can be taken as state variables or any invertible linear transformation of them.

In the stochastic case, as will be discussed in detail in Chapter 10, the pdf of the deterministic state vector of the system *summarizes the past in a probabilistic sense.* This requires that the process noise be *white.* Then the pdf of the state vector is called the ***information state.***

In the stochastic case, the noises are usually assumed to be

1. zero-mean,
2. white, and
3. mutually independent

stochastic processes. If the noise is not zero mean, its mean (if known) can be taken as a known input.

4.2.2 Solution of the Continuous-Time State Equation

The state equation (4.2.1-1) has the following solution:

$$x(t) = F(t, t_0)x(t_0) + \int_{t_0}^{t} F(t, \tau)[B(\tau)u(\tau) + D(\tau)\tilde{v}(\tau)]\, d\tau \tag{4.2.2-1}$$

where $x(t_0)$ is the initial state and $F(t, t_0)$ is the ***state transition matrix*** from t_0 to t.

The transition matrix has the following properties:

$$\frac{dF(t, t_0)}{dt} = A(t)F(t, t_0) \tag{4.2.2-2}$$

$$F(t_2, t_0) = F(t_2, t_1)F(t_1, t_0) \qquad \forall t_1 \tag{4.2.2-3}$$

$$F(t, t) = I \tag{4.2.2-4}$$

The last two imply that

$$F(t, t_0) = F(t_0, t)^{-1} \tag{4.2.2-5}$$

The transition matrix has, in general, no explicit form, unless the following *commutativity property* is satisfied:

$$A(t) \int_{t_0}^{t} A(\tau)\, d\tau = \int_{t_0}^{t} A(\tau)\, d\tau A(t) \tag{4.2.2-6}$$

Then (and only then)

$$F(t, t_0) = e^{\int_{t_0}^{t} A(\tau)\, d\tau} \tag{4.2.2-7}$$

Condition (4.2.2-6) is satisfied for time-invariant systems or diagonal $A(t)$.

For a time-invariant system, assuming $t_0 = 0$, one has

$$F(t) \triangleq F(t, 0) = e^{At} \tag{4.2.2-8}$$

Evaluation of the Transition Matrix

Some of the computational methods for the evaluation of the matrix e^{At} are briefly presented below.

1. Infinite series method:

$$e^{At} = \sum_{k=0}^{\infty} \frac{(At)^k}{k!} = I + At + \frac{A^2 t^2}{2} + \cdots \tag{4.2.2-9}$$

where I is the identity matrix of the same dimension $n \times n$ as A. This is a numerical method and it requires series truncation (unless a closed-form expression can be found for each term).

2. Laplace transform method:

$$e^{At} = \mathcal{L}^{-1}\{(sI - A)^{-1}\} \tag{4.2.2-10}$$

where $\mathcal{L}^{-1}$ is the inverse Laplace transform. This is practical if one can find a closed-form expression of the required matrix inverse above.

3. ***Interpolating polynomial*** method. Compute the eigenvalues λ_i of A, $i = 1, \ldots, n_e$, where n_e is the number of *distinct eigenvalues*, with multiplicities m_i, and

$$\sum_{i=1}^{n_e} m_i = n \tag{4.2.2-11}$$

Then find a polynomial of degree $n - 1$

$$g(\lambda) = \sum_{k=0}^{n-1} g_k \lambda^k \tag{4.2.2-12}$$

which is equal to $e^{\lambda t}$ *on the spectrum of* A, that is,

$$\frac{d^j}{d\lambda^j} g(\lambda)|_{\lambda=\lambda_i} = \frac{d^j}{d\lambda^j} e^{\lambda t}|_{\lambda=\lambda_i} \qquad i = 1, \ldots, n_e, \quad j = 0, \ldots, m_i - 1 \tag{4.2.2-13}$$

Then

$$e^{At} = g(A) \tag{4.2.2-14}$$

Example — Coordinated Turn

Consider an object moving in a plane with ***constant speed*** (the *magnitude* of the velocity vector) and turning with a ***constant angular rate*** (i.e., executing a ***coordinated turn*** in aviation language).

The equations of motion in the plane (ξ, η) in this case are

$$\ddot{\xi} = -\Omega\dot{\eta} \qquad\qquad \ddot{\eta} = \Omega\dot{\xi} \tag{4.2.2-15}$$

where Ω is the constant angular rate ($\Omega > 0$ implies a counterclockwise turn).[2]

The state space representation of the above with the state vector

$$x \triangleq [\xi \;\; \dot{\xi} \;\; \eta \;\; \dot{\eta}]' \tag{4.2.2-16}$$

is

$$\dot{x} = Ax \tag{4.2.2-17}$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\Omega \\ 0 & 0 & 0 & 1 \\ 0 & \Omega & 0 & 0 \end{bmatrix} \tag{4.2.2-18}$$

It can be easily shown that the eigenvalues of A are 0, 0, and $\pm\Omega j$.

It can be shown, using one of the techniques discussed earlier for evaluating the transition matrix, that for A given above one has

$$e^{At} = \begin{bmatrix} 1 & \dfrac{\sin\Omega t}{\Omega} & 0 & -\dfrac{1-\cos\Omega t}{\Omega} \\ 0 & \cos\Omega t & 0 & -\sin\Omega t \\ 0 & \dfrac{1-\cos\Omega t}{\Omega} & 1 & \dfrac{\sin\Omega t}{\Omega} \\ 0 & \sin\Omega t & 0 & \cos\Omega t \end{bmatrix} \tag{4.2.2-19}$$

This allows, among other things, easy generation of state trajectories for such turns (the position evolves along circular arcs). These turns are common for aircraft as well as other flying objects.

4.2.3 The State as a Markov Process

Assume the process noise entering the state equation (4.2.1-1) to be zero mean and *white*, that is, that $\tilde{v}(t)$ is *independent* of $\tilde{v}(\tau)$ for all $t \neq \tau$. In this case, the autocorrelation of $\tilde{v}(t)$ is

$$E[\tilde{v}(t)\tilde{v}(\tau)'] = V(t)\delta(t-\tau) \tag{4.2.3-1}$$

[2] This is in accordance to the trigonometric convention. In the navigation convention, angles are measured clockwise (from North); thus a positive turn rate implies a clockwise turn.

The whiteness property of the process noise allows the preservation of the state's property of summarizing the past in the following sense: the pdf of the state at some time t conditioned on its values up to an earlier time t_1 depends only on the last value $x(t_1)$:

$$p[x(t)|x_{[-\infty,t_1]}, u_{[t_1,t]}] = p[x(t)|x(t_1), u_{[t_1,t]}] \tag{4.2.3-2}$$

This follows from the complete unpredictability of the process noise due to its whiteness. Were the process noise autocorrelated ("colored"), (4.2.3-2) would not hold because states prior to t_1 could be used to predict the process noise $\tilde{v}_{[t_1,t]}$, and thus $x(t)$, in some fashion.

This can be seen from the solution of the state equation

$$x(t) = F(t,t_1)x(t_1) + \int_{t_1}^{t} F(t,\tau)[B(\tau)u(\tau) + D(\tau)\tilde{v}(\tau)]\, d\tau \tag{4.2.3-3}$$

which indicates that $x(t_1)$ summarizes the past, the input provides the known part of the state's evolution after t_1, and the last term above is the contribution of the process noise, which is *completely unpredictable.*

In other words, the state of a dynamic system driven by white noise is a ***Markov process.***

4.2.4 Propagation of the State's Mean and Covariance

Consider (4.2.1-1) with the known input $u(t)$ and ***nonstationary white process noise*** with nonzero mean

$$E[\tilde{v}(t)] = \bar{v}(t) \tag{4.2.4-1}$$

and autocovariance function

$$E[[\tilde{v}(t) - \bar{v}(t)][\tilde{v}(\tau) - \bar{v}(\tau)]'] = V(t)\delta(t-\tau) \tag{4.2.4-2}$$

The expected value of the state

$$\bar{x}(t) \triangleq E[x(t)] \tag{4.2.4-3}$$

evolves according to the (deterministic) differential equation

$$\boxed{\dot{\bar{x}}(t) = A(t)\bar{x}(t) + B(t)u(t) + D(t)\bar{v}(t)} \tag{4.2.4-4}$$

The above ***propagation equation of the mean*** follows from taking the expected value of (4.2.1-1) or differentiating the expected value of (4.2.2-1) with Leibniz' rule (see problem 4-1).

The covariance of the state

$$P_{xx}(t) \triangleq E[[x(t) - \bar{x}(t)][x(t) - \bar{x}(t)]'] \tag{4.2.4-5}$$

has the expression

$$P_{xx}(t) = F(t,t_0)P_{xx}(t_0)F(t,t_0)' + \int_{t_0}^{t} F(t,\tau)D(\tau)V(\tau)D(\tau)'F(t,\tau)'d\tau \tag{4.2.4-6}$$

and evolves according to the differential equation, known as the ***Lyapunov equation***,

$$\boxed{\dot{P}_{xx}(t) = A(t)P_{xx}(t) + P_{xx}(t)A(t)' + D(t)V(t)D(t)'} \tag{4.2.4-7}$$

This is the ***propagation equation of the covariance***, and it can be proven by evaluating the covariance of (4.2.2-1) using the whiteness of $\tilde{v}(t)$ and differentiating the result (see problem 4-1).

Example

Consider the time-invariant scalar case with $A(t) = -a < 0$ (a stable system), $B(t) = 0$, $D(t) = 1$, $\bar{v}(t) = \bar{v}$, $V(t) = V$. The mean evolves according to

$$\dot{\bar{x}}(t) = -a\bar{x}(t) + \bar{v} \tag{4.2.4-8}$$

which, with initial condition $\bar{x}(0)$, yields

$$\bar{x}(t) = e^{-at}\bar{x}(0) + \frac{\bar{v}}{a}\left(1 - e^{-at}\right) \tag{4.2.4-9}$$

The variance evolution equation is

$$\dot{P}(t) = -2aP(t) + V \tag{4.2.4-10}$$

which, with initial condition $P(0)$, has the solution

$$P(t) = e^{-2at}P(0) + \frac{V}{2a}\left(1 - e^{-2at}\right) \tag{4.2.4-11}$$

Note

In general, the Lyapunov equation has no explicit solution beyond (4.2.4-6), which has to be evaluated numerically. For stable systems the steady-state solution can be obtained with efficient numerical techniques (see, e.g., [Chen84]).

4.2.5 Frequency Domain Approach

Consider the *time-invariant* system driven by noise only

$$\dot{x}(t) = Ax(t) + D\tilde{v}(t) \tag{4.2.5-1}$$

where the noise is zero mean, *stationary*, and white, with autocorrelation function

$$R_{\tilde{v}\tilde{v}}(\tau) = E[\tilde{v}(t+\tau)\tilde{v}(t)'] = V\delta(\tau) \tag{4.2.5-2}$$

and with output

$$z(t) = Cx(t) \tag{4.2.5-3}$$

If the system is *stable* (i.e., all the eigenvalues of the system matrix A are in the left half-plane), then its output becomes a stationary process (when the transient period is over). The autocorrelation of the output is denoted as

$$R_{zz}(\tau) = E[z(t+\tau)z(t)'] \tag{4.2.5-4}$$

The power spectral density — ***power spectrum*** — of the process noise, which is the Fourier transform of its autocorrelation function, is

$$S_{\tilde{v}\tilde{v}}(\omega) = \int_{-\infty}^{\infty} R_{\tilde{v}\tilde{v}}(\tau)e^{-j\omega\tau}d\tau = V \tag{4.2.5-5}$$

It can be shown that the ***power spectral density matrix of the output*** — the Fourier transform of (4.2.5-4) — is

$$S_{zz}(\omega) = H(j\omega)S_{\tilde{v}\tilde{v}}(\omega)H(j\omega)^* = H(j\omega)VH(j\omega)^* \tag{4.2.5-6}$$

where the asterisks denote complex conjugate transpose and

$$H(j\omega) \triangleq C(j\omega I - A)^{-1}D \tag{4.2.5-7}$$

is the transfer function matrix of system (4.2.5-1) from the noise $\tilde{v}$ to the output. Note that $H(j\omega)$ is a ***rational function*** (ratio of polynomials).

Spectral Factorization

Equation (4.2.5-6) leads to the following result: Given a ***rational spectrum***, one can find a linear time-invariant system whose output, when driven by a stationary white noise, will have that spectrum.

The transfer function of such a system, called ***prewhitening system*** or ***shaping filter***, is obtained by ***factorization*** of the desired spectrum into the product of a function with its complex conjugate. The first factor should correspond to a *causal and stable system.*

Example — Exponentially Decaying Autocorrelation

Consider the scalar stochastic process with ***exponentially decaying autocorrelation*** or ***Ornstein-Uhlenbeck process***

$$R_{zz}(\tau) = \sigma^2 e^{-\alpha|\tau|} \qquad \alpha > 0 \tag{4.2.5-8}$$

The spectrum corresponding to the above is

$$S_{zz}(\omega) = \sigma^2 \frac{2\alpha}{\alpha^2 + \omega^2} \tag{4.2.5-9}$$

Factorization of this spectrum according to (4.2.5-6) yields

$$H(j\omega) = \frac{1}{\alpha + j\omega} \tag{4.2.5-10}$$

$$V = 2\alpha\sigma^2 \tag{4.2.5-11}$$

The state equation corresponding to (4.2.5-10) is

$$\dot{x}(t) = -\alpha x(t) + \tilde{v}(t) \tag{4.2.5-12}$$

with output

$$z(t) = x(t) \tag{4.2.5-13}$$

and process noise autocorrelation

$$E[\tilde{v}(t+\tau)\tilde{v}(t)] = 2\alpha\sigma^2\delta(\tau) \tag{4.2.5-14}$$

Note that the transfer function in (4.2.5-10) corresponds to a stable causal system, while its complex conjugate would represent an unstable noncausal system.

4.3 DISCRETE-TIME LINEAR STOCHASTIC DYNAMIC SYSTEMS

4.3.1 The Discrete-Time State-Space Model

In the ***state space representation*** of ***discrete-time systems***, it is assumed that the input is piecewise constant, that is,

$$u(t) = u(t_k) \qquad t_k \le t < t_{k+1} \tag{4.3.1-1}$$

Then the state at sampling time t_{k+1} can be written, from (4.2.2-1), in terms of the state at t_k as

$$x(t_{k+1}) = F(t_{k+1}, t_k)x(t_k) + G(t_{k+1}, t_k)u(t_k) + v(t_k) \tag{4.3.1-2}$$

where F is the ***(state) transition matrix*** of the system, G is the ***discrete-time gain*** through which the input, assumed to be constant over a sampling period, enters the system and $v(t_k)$ is the ***discrete-time process noise***.

For a *time-invariant* continuous-time system sampled at arbitrary times, the transition matrix is

$$F(t_{k+1}, t_k) = F(t_{k+1} - t_k) = e^{(t_{k+1}-t_k)A} \triangleq F(k) \tag{4.3.1-3}$$

the input gain is

$$G(t_{k+1}, t_k) = \int_{t_k}^{t_{k+1}} e^{(t_{k+1}-\tau)A} B d\tau \triangleq G(k) \tag{4.3.1-4}$$

and the discrete-time process noise relates to the continuous-time noise as

$$v(t_k) = \int_{t_k}^{t_{k+1}} e^{(t_{k+1}-\tau)A} D\tilde{v}(\tau) d\tau \triangleq v(k) \tag{4.3.1-5}$$

Equations (4.3.1-3) to (4.3.1-5) introduce the simplified index-only notation for discrete-time systems, to be used (most of the time) in the sequel.

With the zero-mean and white assumption on $\tilde{v}(t)$, as in (4.2.3-1), it follows that

$$\boxed{E[v(k)] = 0} \tag{4.3.1-6}$$

$$\boxed{E[v(k)v(j)'] = Q(k)\delta_{kj}} \tag{4.3.1-7}$$

where δ_{kj} is the Kronecker delta function. The covariance of the discrete-time process noise is given by

$$Q(k) = \int_{t_k}^{t_{k+1}} e^{(t_{k+1}-\tau)A} DV(\tau)D' e^{(t_{k+1}-\tau)A'} d\tau \tag{4.3.1-8}$$

The proof of (4.3.1-8) is given at the end of this subsection.

The (dynamic) model for ***discrete-time linear stochastic systems*** can be written with the simplified index-only time notation as

$$\boxed{x(k+1) = F(k)x(k) + G(k)u(k) + v(k)} \tag{4.3.1-9}$$

where the input is assumed known along with the matrices $F(k)$ and $G(k)$ and the process noise $v(k)$ is a zero-mean, white random sequence with covariance matrix $Q(k)$. Any (known) nonzero mean of the process noise v can be incorporated into the input.

The ***discrete-time measurement equation*** is, with a similar notation,

$$\boxed{z(k) = H(k)x(k) + w(k)} \tag{4.3.1-10}$$

where $H(k)$ is the ***measurement matrix*** and $w(k)$ is the ***measurement noise*** — a random sequence with moments

$$\boxed{E[w(k)] = 0} \tag{4.3.1-11}$$

$$\boxed{E[w(k)w(j)'] = R(k)\delta_{kj}} \tag{4.3.1-12}$$

The measurement given by (4.3.1-10) represents a "short-term" integration, during which the state is assumed to be constant.

Note that (4.3.1-9) and (4.3.1-10) describe a ***time-varying discrete-time system.***

The process and measurement noise sequences are (usually) assumed uncorrelated, that is,

$$E[v(k)w(j)'] = 0 \qquad \forall k, j \tag{4.3.1-13}$$

In some cases it is convenient to define a ***direct discrete-time model*** rather than a discretized version of a continuous-time model. In such cases the process noise, also modeled as white, enters through a ***noise gain***, denoted as $\Gamma(k)$. Then (4.3.1-9) is replaced by

$$\boxed{x(k+1) = F(k)x(k) + G(k)u(k) + \Gamma(k)v(k)} \tag{4.3.1-14}$$

In this case the process noise covariance $Q(k)$ is defined directly. This will be discussed in more detail in Chapter 6.

Derivation of the Covariance of the Discretized Process Noise

$$\begin{aligned}
E[v(k)v(j)'] &= E\left\{\int_{t_k}^{t_{k+1}} e^{(t_{k+1}-\tau_1)A} D\tilde{v}(\tau_1)d\tau_1 \left[\int_{t_j}^{t_{j+1}} e^{(t_{j+1}-\tau_2)A} D\tilde{v}(\tau_2)d\tau_2\right]'\right\} \\
&= E\left\{\int_{t_k}^{t_{k+1}}\int_{t_j}^{t_{j+1}} e^{(t_{k+1}-\tau_1)A} D\tilde{v}(\tau_1)\tilde{v}(\tau_2)'D'e^{(t_{j+1}-\tau_2)A'}d\tau_1 d\tau_2\right\} \\
&= \int_{t_k}^{t_{k+1}}\int_{t_j}^{t_{j+1}} e^{(t_{k+1}-\tau_1)A} DE[\tilde{v}(\tau_1)\tilde{v}(\tau_2)']D'e^{(t_{j+1}-\tau_2)A'}d\tau_1 d\tau_2 \\
&= \int_{t_k}^{t_{k+1}}\int_{t_j}^{t_{j+1}} e^{(t_{k+1}-\tau_1)A} DV(\tau_1)\delta(\tau_1-\tau_2)D'e^{(t_{j+1}-\tau_2)A'}d\tau_1 d\tau_2 \\
&= \int_{t_k}^{t_{k+1}} e^{(t_{k+1}-\tau_1)A} DV(\tau_1)D'e^{(t_{k+1}-\tau_1)A'}d\tau_1\, \delta_{kj}
\end{aligned} \tag{4.3.1-15}$$

4.3.2 Solution of the Discrete-Time State Equation

Using (4.3.1-9) for time k and substituting $x(k-1)$ yields

$$\begin{aligned}
x(k) &= F(k-1)x(k-1) + G(k-1)u(k-1) + v(k-1) \\
&= F(k-1)[F(k-2)x(k-2) + G(k-2)u(k-2) + v(k-2)] \\
&\quad + G(k-1)u(k-1) + v(k-1) \\
&= F(k-1)F(k-2)x(k-2) + F(k-1)[G(k-2)u(k-2) \\
&\quad \cdot + v(k-2)] + G(k-1)u(k-1) + v(k-1)
\end{aligned} \tag{4.3.2-1}$$

Repeating the above leads to

$$x(k) = [\prod_{j=0}^{k-1} F(k-1-j)]x(0) + \sum_{i=0}^{k-1}[\prod_{j=0}^{k-i-2} F(k-1-j)][G(i)u(i)+v(i)] \tag{4.3.2-2}$$

The notation for the product of matrices used above is

$$\prod_{j=j_1}^{j_2} F(j) \triangleq F(j_1)F(j_1+1)\ \ldots\ F(j_2) \tag{4.3.2-3}$$

If the upper index in (4.3.2-3) is smaller than the lower index, the result is taken as the identity matrix.

Note that

$$\prod_{j=0}^{k-i-1} F(k-1-j) = F(k-1)F(k-2)\ \ldots\ F(i) = F(t_k, t_i) \tag{4.3.2-4}$$

is the transition matrix from sampling time i to sampling time k.

If the discrete-time system is ***time-invariant***, that is,

$$F(k) = F, \qquad G(k) = G \qquad \forall k \tag{4.3.2-5}$$

then (4.3.2-2) becomes

$$x(k) = F^k x(0) + \sum_{i=0}^{k-1} F^{k-i-1}[Gu(i) + v(i)] \tag{4.3.2-6}$$

4.3.3 The State as a Markov Process

If the discrete-time representation is obtained by discretizing a continuous-time system with white process noise, the resulting discrete-time process noise is a white sequence.

Similarly to the continuous-time case, one has, following (4.3.2-1),

$$x(k) = [\prod_{j=0}^{k-l-1} F(k-1-j)]x(l) + \sum_{i=l}^{k-1}[\prod_{j=0}^{k-i-2} F(k-1-j)][G(i)u(i) + v(i)] \tag{4.3.3-1}$$

Thus, since $v(i)$, $i = l, \ldots, k-1$, are independent of

$$X^l \triangleq \{x(j)\}_{j=0}^{l} \tag{4.3.3-2}$$

which depend only on $v(i)$, $i = 0, \ldots, l-1$, one has

$$p[x(k)|X^l, U^{k-1}] = p[x(k)|x(l), U_l^{k-1}] \qquad \forall k > l \tag{4.3.3-3}$$

where

$$U_l^{k-1} \triangleq \{u(j)\}_{j=l}^{k-1} \tag{4.3.3-4}$$

Thus, the state vector is a ***Markov process***, or, more correctly, a ***Markov sequence***.

As in the continuous-time case, this follows from the complete unpredictability of the process noise due to its whiteness. Were the process noise autocorrelated ("colored"), (4.3.3-3) would not hold because states prior to time l could be used to predict the process noises $v(i)$, $i = l, \ldots, k-1$, and thus $x(k)$, in some fashion.

4.3.4 Propagation of the State's Mean and Covariance

Consider (4.3.1-14), repeated below for convenience

$$x(k+1) = F(k)x(k) + G(k)u(k) + \Gamma(k)v(k) \tag{4.3.4-1}$$

with the known input $u(k)$ and the process noise $v(k)$ white, but for the sake of generality, ***nonstationary*** with nonzero mean:

$$E[v(k)] = \bar{v}(k) \tag{4.3.4-2}$$

$$\text{cov}[v(k), v(j)] = E[[v(k) - \bar{v}(k)][v(j) - \bar{v}(j)]'] = Q(k)\delta_{kj} \tag{4.3.4-3}$$

Then the expected value of the state

$$\bar{x}(k) \triangleq E[x(k)] \tag{4.3.4-4}$$

evolves according to the difference equation

$$\boxed{\bar{x}(k+1) = F(k)\bar{x}(k) + G(k)u(k) + \Gamma(k)\bar{v}(k)} \tag{4.3.4-5}$$

The above, which is the ***propagation equation of the mean***, follows immediately by applying the expectation operator to (4.3.4-1).

The covariance of the state

$$P_{xx}(k) \triangleq E[[x(k) - \bar{x}(k)][x(k) - \bar{x}(k)]'] \tag{4.3.4-6}$$

evolves according to the difference equation — the ***covariance propagation equation***

$$\boxed{P_{xx}(k+1) = F(k)P_{xx}(k)F(k)' + \Gamma(k)Q(k)\Gamma(k)'} \tag{4.3.4-7}$$

This follows by subtracting (4.3.4-5) from (4.3.4-1), which yields

$$x(k+1) - \bar{x}(k+1) = F(k)[x(k) - \bar{x}(k)] + \Gamma(k)[v(k) - \bar{v}(k)] \tag{4.3.4-8}$$

It can be easily shown that multiplying (4.3.4-8) with its transpose and taking the expectation yields (4.3.4-7). The resulting cross-terms on the right-hand side vanish when the expectation is taken. This is due to the whiteness of the process noise, which causes $x(k)$, being a linear combination of the noises prior to k, to be *independent* of $v(k)$.

Example — Fading Memory Average

Consider the scalar system (also called "first-order Markov")

$$x(k) = \alpha x(k-1) + v(k) \qquad k = 1, \ldots \tag{4.3.4-9}$$

with $x(0) = 0$ and $0 < \alpha < 1$. Note the slight change in the time argument in comparison to (4.3.4-1).

For this system it can be easily shown directly, or, using (4.3.2-2), that its solution is

$$x(k) = \sum_{i=1}^{k} \alpha^{k-i} v(i) \qquad k = 1, \ldots \tag{4.3.4-10}$$

Normalizing the above by the sum of the coefficients yields

$$z(k) \triangleq \frac{x(k)}{\sum_{i=1}^{k} \alpha^{k-i}} \tag{4.3.4-11}$$

which is the ***fading memory average***, or, ***exponentially discounted average*** of the sequence $v(k)$. The term fading memory average is sometimes used for (4.3.4-10), which is really a ***fading memory sum*** — without the normalization. Note that for $\alpha = 1$, (4.3.4-11) becomes the sample average.

If the input has constant mean

$$E[v(k)] = \bar{v} \tag{4.3.4-12}$$

then

$$\bar{x}(k) \triangleq E[x(k)] = \bar{v} \sum_{i=1}^{k} \alpha^{k-i} = \bar{v} \frac{1 - \alpha^k}{1 - \alpha} \tag{4.3.4-13}$$

and

$$\lim_{k \to \infty} \bar{x}(k) = \frac{\bar{v}}{1 - \alpha} \tag{4.3.4-14}$$

It can be easily shown that

$$\bar{z}(k) \triangleq E[z(k)] = \bar{v} \qquad \forall k \tag{4.3.4-15}$$

The use of the fading memory average is for the case of a "slowly" varying mean $\bar{v}(k)$. The larger weightings on the more recent values of $v(k)$ allow "tracking" of its mean at the expense of the accuracy (larger variance) — see also problem 4-4.

4.3.5 Frequency Domain Approach

Consider the time-invariant system driven by noise only,

$$x(k+1) = Fx(k) + \Gamma v(k) \tag{4.3.5-1}$$

where the noise is assumed zero mean, stationary, and white, with autocorrelation

$$R_{vv}(k - l) = E[v(k)v(l)'] = Q\delta_{kl} \tag{4.3.5-2}$$

The output of the system is

$$z(k) = Hx(k) \tag{4.3.5-3}$$

The ***power spectral density (spectrum)*** of the discrete-time process noise is the ***discrete-time Fourier transform (DTFT)***[3]

$$S_{vv}(e^{j\omega T}) = \sum_{n=-\infty}^{\infty} e^{-j\omega T n} R_{vv}(n) = Q \tag{4.3.5-4}$$

where ω is the angular frequency and T the sampling period. Note the constant (flat) spectrum that characterizes a white noise.

If system (4.3.5-1) is ***stable*** (i.e., all the eigenvalues of F are *inside the unit circle*), then the state $x(k)$ will also become a stationary random sequence. The spectrum of the output will be

$$S_{zz}(e^{j\omega T}) = \mathcal{H}(\zeta) S_{vv}(\zeta) \mathcal{H}(\zeta)^* |_{\zeta = e^{j\omega T}} \tag{4.3.5-5}$$

where the asterisk denotes complex conjugate transpose, ζ is the (two-sided) z-transform variable, and

$$\mathcal{H}(\zeta) = H(\zeta I - F)^{-1}\Gamma \tag{4.3.5-6}$$

is the ***discrete-time transfer function*** of the system given by (4.3.5-1) and (4.3.5-3) from the process noise to the output. Note that the transfer function (4.3.5-6) is a ***rational function***.

Using (4.3.5-4) in (4.3.5-5) yields the spectrum of the output as

$$S_{zz}(e^{j\omega T}) = \mathcal{H}(e^{j\omega T}) Q \mathcal{H}(e^{j\omega T})^* = \mathcal{H}(e^{j\omega T}) Q \mathcal{H}(e^{-j\omega T})' \tag{4.3.5-7}$$

which, being the product of two rational functions and a constant matrix, is also a rational function.

Spectral Factorization

In view of (4.3.5-7), given a *rational spectrum*, one can find the linear time-invariant system which, driven by white noise, will have the output with the desired spectrum. Based on (4.3.5-7), the transfer function of such a system is obtained by ***factorization*** of the spectrum as follows:

$$S_{zz}(\zeta) = \mathcal{H}(\zeta) Q \mathcal{H}(\zeta^{-1})' \tag{4.3.5-8}$$

since, for

$$\zeta = e^{j\omega T} \tag{4.3.5-9}$$

one has

$$\zeta^* = \zeta^{-1} \tag{4.3.5-10}$$

The resulting transfer function $\mathcal{H}(\zeta)$ specifies the ***prewhitening system*** or ***shaping filter*** for the sequence $z(k)$.

[3]The ***DFT*** is the transform based on a *finite number* of points and evaluated at the *same number* of sampled frequencies; it is usually obtained via the FFT.

Remark

The factor in (4.3.5-8) that corresponds to a *causal* and *stable* system is the one that is to be chosen for the transfer function of the prewhitening system.

Example

Consider the scalar sequence $z(k)$ with mean zero and autocorrelation

$$R_{zz}(n) = \rho\delta_{n,-1} + \delta_{n,0} + \rho\delta_{n,1} \tag{4.3.5-11}$$

Its spectrum — the DTFT written as the two-sided z-transform — is

$$S_{zz}(\zeta) = 1 + \rho\zeta^{-1} + \rho\zeta \tag{4.3.5-12}$$

This can be factorized as

$$S_{zz}(\zeta) = \mathcal{H}(\zeta)\mathcal{H}(\zeta^{-1}) = (\beta_0 + \beta_1\zeta^{-1})(\beta_0 + \beta_1\zeta) \tag{4.3.5-13}$$

where the first factor is the transfer function of a causal system.

The equations for the coefficients β_0 and β_1 are

$$\beta_0^2 + \beta_1^2 = 1 \tag{4.3.5-14}$$

$$\beta_0\beta_1 = \rho \tag{4.3.5-15}$$

with solution

$$\beta_0 = \frac{1}{2}(\sqrt{1+2\rho} + \sqrt{1-2\rho}) \tag{4.3.5-16}$$

$$\beta_1 = \frac{1}{2}(\sqrt{1+2\rho} - \sqrt{1-2\rho}) \tag{4.3.5-17}$$

Thus the transfer function of the system is

$$\mathcal{H}(\zeta) = \beta_0 + \beta_1\zeta^{-1} \tag{4.3.5-18}$$

which corresponds to the following ***moving average (MA)*** or ***finite impulse response (FIR)*** system driven by unity-variance white noise

$$x(k) = \beta_0 v(k) + \beta_1 v(k-1) \tag{4.3.5-19}$$

$$z(k) = x(k) \tag{4.3.5-20}$$

Equations (4.3.5-19) and (4.3.5-20) specify the prewhitening system corresponding to (4.3.5-11).

4.4 SUMMARY

4.4.1 Summary of State Space Representation

State of a deterministic system — the smallest vector that *summarizes in full* its past.

State equation — a first-order differential or difference equation that describes the evolution in time (the dynamics) of the state vector.

Markov process — a stochastic process whose current state contains all the information about the probabilistic description of its future evolution.

A *state equation* driven by *white noise* yields a (vector) *Markov process.*

State of a stochastic system described by a Markov process — *summarizes probabilistically* its past.

All the above statements hold for linear as well as nonlinear systems.

A linear stochastic system's *continuous-time* representation as a differential equation driven by white noise can be written in *discrete-time* as a difference equation driven by a sequence of independent random variables — discrete-time white noise, which can be related to the continuous-time noise.

Alternatively, one can define directly a discrete-time state equation.

Continuous-time white noise: a random process with autocorrelation function a Dirac (impulse) delta function. If it is stationary, it has a spectrum (Fourier transform of the autocorrelation) that is constant ("flat").

Discrete-time white noise: a random sequence with autocorrelation function being a Kronecker delta function. If it is stationary, it has a spectrum (DTFT of the autocorrelation) that is constant ("flat").

The *unconditional mean and covariance* of the state of a linear stochastic system driven by white noise have been shown to evolve ("open loop") according to linear differential or difference equations.

4.4.2 Summary of Prewhitening

For a linear time-invariant system, the frequency domain approach relates the power spectral density of the output to the one of the input via the transfer function.

The spectral density of the output of a stable linear time-invariant system is given by:

> The spectrum of the input premultiplied by the system's transfer function (a rational function) and postmultiplied by its complex conjugate and transpose.

Given a rational spectrum, one can find a linear time-invariant system which, if driven by a stationary white noise, its output will have that spectrum:

The transfer function of such a *prewhitening system* is obtained by factoring the desired spectrum into the product of a function with its complex conjugate — this is called *spectral factorization*.

Due to the facts that

- Estimation results for dynamic systems, to be discussed in the sequel, are for Markov systems; and
- (Most) Markov processes (of interest) can be represented by linear time-invariant systems driven by white noise,

it is very important to find the "prewhitening system" for a given random process or sequence. This can be accomplished via spectral factorization.

4.5 NOTES AND PROBLEMS

4.5.1 Bibliographical Notes

The material on state space representation of stochastic systems is also discussed in, for example, [Sage71]. More on the continuous-time and discrete-time representations can be found, for example, in [Fortmann77] and [Chen84].

The relationship between the spectra of the output and input of a linear time-invariant system is proven in standard probability texts — for example, in [Papoulis84]. Spectral factorization is treated in more detail in [Anderson79].

4.5.2 Problems

4-1 Mean and covariance of the state of a linear system.

1. Prove (4.2.4-4) by differentiating the expected value of (4.2.2-1) with Leibniz' rule

$$\frac{d}{dt}\int_{a(t)}^{b(t)} f(t,\tau)\,d\tau = \int_{a(t)}^{b(t)} \frac{\partial f(t,\tau)}{\partial t}\,d\tau + \frac{db(t)}{dt} f[t,b(t)] - \frac{da(t)}{dt} f[t,a(t)]$$

2. Prove (4.2.4-6).

3. Prove (4.2.4-7).

4-2 Moments of the white noise response of a scalar linear system. Consider the system

$$\dot{x}(t) = \alpha x(t) + \tilde{v}(t)$$

with $\tilde{v}(t)$ white, with mean $\bar{v}$ and variance q.

1. Find α such that the mean of the resulting stationary process is $\bar{x} = c$. (*Hint*: This is the steady state of the differential equation of the mean.)

2. Using a time domain approach, find the autocorrelation of the resulting stationary process.

4-3 Autocovariance of the state of a discrete-time system.

1. Find the autocovariance

$$V_{xx}(k,j) \triangleq E\Big[[x(k)-\bar{x}(k)][x(j)-\bar{x}(j)]'\Big]$$

in terms of $P_{xx}(j)$ for system (4.3.1-14). Assume $k > j$.

2. Indicate what happens for a stable linear time-invariant system for $k \gg j$. Justify.

4-4 Fading memory average and effective window length. Consider

$$y(k) = \alpha y(k-1) + (1-\alpha)v(k)$$

with $y(0) = 0$ and $0 < \alpha < 1$.

1. Write the solution for $y(k)$.
2. Find the mean and variance of $y(k)$ if $v(k)$ is white with mean $\bar{v}$ and variance σ^2.
3. How does $y(k)$ differ from $z(k)$ in (4.3.4-11)?
4. What is the "effective memory" (window length) N_e as a function of α? Determine α for $N_e = 10$.

4-5 Spectral factorization for prewhitening (shaping). Given the scalar zero-mean random sequence $x(k)$ with autocorrelation

$$R_{xx}(n) = E[x(k)x(k+n)] = \sigma^2 a^{|n|} \qquad 0 < a < 1$$

1. Find its spectrum.
2. Factorize it to find the causal and stable linear system driven by white noise (prewhitening system or shaping filter) whose output has this autocorrelation.

4-6 Autocovariance of the state of a system driven by nonstationary white noise. Consider

$$\dot{x}(t) = Ax(t) + D\tilde{v}(t)$$

$$E[x(t_0)] = \bar{x}(t_0)$$

$$E\Big[[x(t_0)-\bar{x}(t_0)][x(t_0)-\bar{x}(t_0)]'\Big] = P_{xx}(t_0)$$

$$E[\tilde{v}(t)] = \bar{v}(t)$$

$$E\Big[[\tilde{v}(t)-\bar{v}(t)][\tilde{v}(\tau)-\bar{v}(\tau)]'\Big] = Q(t)\delta(t-\tau)$$

Let

$$E[x(t)] \triangleq \bar{x}(t)$$

Find

$$V_{xx}(t,\tau) \triangleq E\Big[[x(t)-\bar{x}(t)][x(\tau)-\bar{x}(\tau)]'\Big]$$

4-7 State prediction in a linear time-invariant discrete time system.

1. Simplify (4.3.3-1) for a time-invariant system, that is, $F(i) = F$, $G(i) = G$.

2. Find a closed-form solution, similar to (4.3.3-1), for the covariance (4.3.4-7) assuming $F(k) = F$, $\Gamma(k) = \Gamma$, $Q(k) = Q$.

4-8 Coordinated turn transition matrix.

1. Derive (4.2.2-15).

2. Derive (4.2.2-19).

4-9 State equation for fading memory average.

1. Find the time-invariant linear state equation that yields, for $n \gg 1$, the fading memory average

$$x(n) = \frac{\sum_{i=0}^{n} \alpha^{n-i} u(i)}{\sum_{i=0}^{n} \alpha^{n-i}} \qquad 0 < \alpha < 1$$

2. With $x(0) = 0$, $E[u(i)] = \overline{u}$, $\mathrm{cov}[u(i), u(j)] = \sigma^2 \delta_{ij}$, find the mean and variance of $x(n)$

3. Verify that for $n \gg 1$, one has $E[x(n)] = \overline{u}$

4. Find the asymptotic value of var$[x(n)]$. Is $x(n)$ a consistent estimator of the parameter $\overline{u}$?

5. Determine α such that the standard deviation of the estimate of the parameter $\overline{u}$ is 0.1σ.

4-10 Independence of measurements.

1. Are the measurements in a standard noisy linear dynamic system independent?

2. Is there a way in which they are independent?

4-11 Spectral factorization for a discrete-time system. Consider the scalar zero-mean random sequence $x(k)$ with autocorrelation

$$R_{xx}(n) = E[x(k)x(k+n)] = \sigma^2 a^{|n|} \qquad 0 < a < 1 \qquad n = \ldots, -1, 0, 1, \ldots$$

1. Find its spectrum.

2. Factorize it to find the causal and stable linear system driven by white noise (prewhitening system or shaping filter) whose output has this autocorrelation.

3. Write the state space equation of this system.