plane with a uniform angular velocity ω, which is set as close to $2\pi/T$ as possible, where T is the period of the orbit;

(b) a horizon sensor mounted in the cylinder, which measures the angle $\alpha(t)$ between the line of sight to the Earth's center and a reference line on the rotating cylinder. This reference line is chosen so that $\alpha(T_o) \cong 0$; that is, $\alpha \cong 0$ at perigee. In the absence of measurement error, $\alpha(t)$ is a periodic function, well defined by the parameters e, a, T_o, ω, and $\alpha(T_o)$; e.g., if the orbit is circular, then $\alpha(t) = 0$ if $\omega = 2\pi/T$ exactly. Any deviation of the orbit from a circle will cause the rotating reference line to periodically lead and lag behind the line of sight. The problem here is to improve the estimates of e, a, T_o, ω, and $\alpha(T_o)$, based on noisy measurements of α.

The relationship between α, at a time t, and the parameters a, e, T_o, ω, and $\alpha(T_o)$ is given *implicitly* as follows:†

$$\alpha = \phi - m\,; \qquad \cos\phi = \frac{\cos E - e}{1 - e\cos E}\,; \qquad M = E - e\sin E\,;$$

$$M = \frac{2\pi(t - T_o)}{T}\,; \qquad m = \omega(t - T_o) - \alpha_o\,; \qquad \alpha_o \equiv \alpha(T_o)\,;$$

$$T = \frac{2\pi}{R\sqrt{g}}a^{3/2}\,.$$

Here the angles ϕ, M, and E are known as the true, mean, and eccentric anomalies respectively, g = acceleration of gravity at the Earth's surface, and R = radius of Earth. Note that $m \equiv M$ if $\omega = 2\pi/T$ and $\alpha_o = 0$.

By taking *differentials* of the relationships above and eliminating $d\phi$, dM, dE, dm, and dT, the following relation may be obtained:

$$d\alpha = \frac{\partial\alpha}{\partial a}da + \frac{\partial\alpha}{\partial e}de + \frac{\partial\alpha}{\partial T_o}dT_o + T_o\,d\omega + d\alpha_o\,,$$

where

$$\frac{\partial\alpha}{\partial a} = \frac{3\pi T_o(1 - e^2)\sin E}{aT\sin\phi(1 - e\cos E)^3}\,, \qquad \frac{\partial\alpha}{\partial e} = \frac{(1 - e^2)\sin^2 E}{\sin\phi(1 - e\cos E)^2}$$

$$\frac{\partial\alpha}{\partial T_o} = \omega - \frac{2\pi}{T}\frac{(1 - e^2)\sin E}{\sin\phi(1 - e\cos E)^3}\,,$$

and these partial derivatives are evaluated with the best current estimates of a, e, T_o, ω, α_o at the time of the measurement.

The measurement $z(t)$ is assumed to contain a random error, ...

---

†See, for example, J. M. A. Danby, *Fundamental of Celestial Mechanics*. New York: Macmillan, 1962.

ular velocity ω, which is set ...
s the period of the orbit.
nted in the cylinder, which ...
e of sight to the Earth's center ...
ting cylinder. This reference ...
hat is, $\alpha \cong 0$ at perigee. In the ...
a periodic function, well defined ...
$\alpha(T_o)$; e.g., if the orbit is circular ...
. Any deviation of the orbit from ...
erence line to periodically lead and lag ...
The problem here is to improve the ...
$T_o)$, based on noisy measurements ...
n α, at a time t, and the parameters ...
plicitly as follows:†

$$= \frac{\cos E - e}{1 - e\cos E}; \quad M = E - e \sin E$$

$$= \omega(t - T_o) - \alpha_o ; \quad \alpha_s = \ldots$$

E are known as the true, mean, and ...
g = acceleration of gravity at the Earth ...
Earth. Note that $m \cong M$ if ...

the relationships above and eliminating ...
e following relation may be obtained

$$de + \frac{\partial \alpha}{\partial T_o} dT_o + T_o\, d\omega + d\alpha_o$$

n E ... $(1-e^2)\sin E$ ... $\frac{\partial \alpha}{\partial e} = \frac{(1-e^2)\ldots}{\sin\phi(1-\ldots)}$

are evaluated with the best current ...
e time of the measurement.
assumed to contain a random ...

. Fundamental of Celestial Mechanics ...

$$z(t) = \alpha(t) + v \,,$$

where $E(v) = 0$, $E(v^2) = R$, and R is known. Let $\bar{\alpha}(t)$ be the *predicted* measurement, using the best current estimates of a, e, T_o, ω, α_o, at t. Then we have

$$z(t) - \bar{\alpha}(t) \cong d\alpha(t) + v$$

$$z(t) - \bar{\alpha}(t) \cong \left[\frac{\partial \alpha}{\partial a}, \frac{\partial \alpha}{\partial e}, \frac{\partial \alpha}{\partial T_o}, T_o, 1\right] \begin{bmatrix} da \\ de \\ dT_o \\ d\omega \\ d\alpha_o \end{bmatrix} + v \,.$$

This *linear* relation can then be used to estimate da, de, dT_o, $d\omega$, and $d\alpha_o$ from the measurement $z(t) - \bar{\alpha}(t)$.

Problem 1. In Equation (12.2.1), assume that x and v are independent random vectors with gaussian density functions. Show that the joint density function $p(x,v)$ is proportional to $\exp(-J)$, where J is as defined in Equation (12.2.5). Thus, $x = \hat{x}$, $v = z - H\hat{x}$ maximize $p(x,v)$, justifying the name "maximum likelihood estimate."

Problem 2. Establish the relations

$$P = M - MH^T(HMH^T + R)^{-1}HM \,, \qquad \text{(a)}$$

$$PH^TR^{-1} = MH^T(HMH^T + R)^{-1} \,. \qquad \text{(b)}$$

Note that these relations involve inverting matrices of smaller dimension to determine P and K than Equation (12.2.8) if R is of smaller dimension than P; i.e., if $p < n$. Equations (a) and (12.2.8) are known as the "matrix inversion pair" (see Problem 4, Section 1.3).

Problem 3. Complete the square in Equation (12.2.5) and show that

$$J = \tfrac{1}{2}[x - \bar{x} - PH^TR^{-1}(z - H\bar{x})]^T P^{-1}[x - \bar{x} - PH^TR^{-1}(z - H\bar{x})] + \tfrac{1}{2}(z - H\bar{x})^T R^{-1}(R - HPH^T)R^{-1}(z - H\bar{x}) \,.$$

Thus, J is minimized by choosing $x = \hat{x}$, where

$$\hat{x} = \bar{x} + PH^TR^{-1}(z - H\bar{x}) \,,$$

which is in agreement with Equation (12.2.7).

Problem 4. Given two correlated gaussian random vectors x and z, with mean values $\bar{x}, \bar{z}$ and covariance matrices P_{xx}, P_{zz}, respectively, and correlation $E[(x - \bar{x})(z - \bar{z})^T] = P_{xz}$, *show that* the conditional density function $p(x/z)$ is gaussian, with

$$E(x/z) = \bar{x} + P_{xz}P_{zz}^{-1}(z - \bar{z}) = \hat{x},$$
$$E\{[(x - \hat{x})(x - \hat{x})^T]/z\} = P_{xx} - P_{xz}P_{zz}^{-1}P_{xz}^T.$$

Problem 5. In Problem 4, let $z = Hx + v$, where H is a known matrix and v is independent of x, with mean value zero and covariance R. Let $P_{xx} = M$ and *show that*

$$P_{zz} = R + HMH^T, \quad P_{xz} = MH^T, \quad \bar{z} = H\bar{x}.$$

Using these relations in Problem 4, verify Equations (12.2.7) and (12.2.8) (you will also need the results of Problem 2). Note that $K = P_{xz}P_{zz}^{-1}$, a most reasonable result!

Problem 6. In Problem 4, show that the gaussian random vectors $e = E(x/z) - x$ and $z - \bar{z}$ are independent; i.e., we have $E[e(z - \bar{z})^T] = 0$.

Problem 7. Suppose that the number of theoretical relationships among measured variables z and state variables x is less than the number of measured variables; for example, we have

$$Az = Hx + Av,$$

where

$$A \text{ is a } (q \times p)\text{-matrix}, \quad q < p, \quad H \text{ is a } (q \times n)\text{-matrix},$$
$$E(v) = 0, \quad E(vv^T) = R \text{ is a } (p \times p)\text{-matrix}.$$

Show that the estimation procedure of this section applies with z replaced by Az and R by ARA^T.

Problem 8. Consider the usual problem of least square fit, i.e., determining x to minimize

$$J = \tfrac{1}{2}\|z - Hx\|^2.$$

Show that the error of the fit $e = z - H\hat{x}$ is orthogonal to the fit $\hat{z} = H\hat{x}$, in the sense $e^T\hat{z} = 0$.

Problem 9. In Example 2, suppose that initial estimates of an orbit are $a = 6{,}000$ miles, $\bar{e} = 1/6$, $\bar{T}_o = 0$, $\bar{\omega} = 2\pi/T$, $\bar{\alpha}_o = 0$. Take $R = 3{,}960$ miles, $g = 32.2$ ft sec^{-2}, and make an improved estimate of a, e, T_o, ω, and α_o, using the single measurement

$$z(t) = 16.7 \text{ deg} \quad \text{at} \quad t = 1{,}357 \text{ sec},$$

where

$E(\alpha - \bar{\alpha})^2 = 10^{-2}\,(\text{deg})^2$, $\quad E(a - \bar{a})^2 = 10^{-4}\,(\text{miles})^2$, $\quad E(e - \bar{e})^2 = 10^{-4}$,

$$E(T_o - \overline{T}_o)^2 = 10^2\,(\text{sec})^2, \quad E(\omega - \bar{\omega})^2 = 10^{-10}\,(\text{sec})^{-2},$$

$$E(\alpha_o - \bar{\alpha}_o)^2 = 10^{-2}\,(\text{deg})^2,$$

and all covariances are zero. [HINT: Use Equation (a) of Problem 2, rather than Equation (12.2.8), to find P.]

12.3 Optimal filtering for single-stage linear transitions

Consider a system that makes a discrete transition from state 0 to state 1 according to the linear relation

$$x_1 = \Phi_o x_o + \Gamma_o w_o, \tag{12.3.1}$$

where Φ_o = a known $(n \times n)$ transition matrix, Γ_o = a known $(n \times r)$-matrix, and

$$E(w_o) = \overline{w}_o, \quad E(w_o - \overline{w}_o)(w_o - \overline{w}_o)^T = Q_o. \tag{12.3.2}$$

The forcing vector, w_o, is thus a random vector, with mean $\overline{w}_o$ and covariance Q_o. The state x_o is also a random vector, with mean $\hat{x}_o$ and covariance P_o; that is, we have

$$E(x_o) = \hat{x}_o, \quad E(\hat{x}_o - x_o)(\hat{x}_o - x_o)^T = P_o. \tag{12.3.3}$$

Furthermore, x_o and w_o are independent. From this information, it follows that x_1 is also a random vector, and from Section 11.2, Equations (11.2.7) through (11.2.9), it has a mean value $\bar{x}_1$ and a covariance M_1 given by

$$\bar{x}_1 = \Phi_o \hat{x}_o + \Gamma_o \overline{w}_o, \tag{12.3.4}$$

$$M_1 = \Phi_o P_o \Phi_o^T + \Gamma_o Q_o \Gamma_o^T. \tag{12.3.5}$$

Since $\Gamma_o Q_o \Gamma_o^T$ by definition (12.3.2) is a nonnegative matrix, it is apparent from (12.3.5) that, on the average, the effect of the uncertainty w_o in a transition of the type (12.3.1) is to *increase the uncertainty* in our knowledge of the state x_1.† This is to be contrasted with the result (12.2.8), where it was shown that measurements, on the average, *decrease the uncertainty* in our knowledge of the state.††

Suppose that we make measurements, as in Section 12.2, *after* the transition to state 1. Then, from Equations (12.2.7) and (12.2.8), the best estimate of x_1 is given by $\hat{x}_1$, where

$$\hat{x}_1 = \bar{x}_1 + P_1 H_1^T R_1^{-1}(z_1 - H_1 \bar{x}_1), \tag{12.3.6}$$

---

† Strictly, the uncertainty is increased or left unchanged. The increase in P is, of course, dependent on the term Φ_o.

†† More precisely, the uncertainty is decreased or left unchanged.

$$P_1 = (M_1^{-1} + H_1^T R^{-1} H_1)^{-1} = M_1 - M_1 H_1^T (H_1 M_1 H_1^T + R_1)^{-1} H_1 M_1 .$$

Here $\bar{x}_1$ and M_1 are as given in (12.3.4) and (12.3.5). Note that $\bar{x}_1$ is the estimate of x_1 *before* measurement, whereas $\hat{x}_1$ is the estimate *after* measurement. Similarly, M_1 is the error covariance matrix before measurement and P_1 is the error covariance matrix after measurement. Symbolically, we can describe this process as follows:

$$\begin{array}{llll} & \overline{w}_o & z_1 & \\ & \downarrow & \downarrow & \\ \text{mean:} \quad & \hat{x}_o \longrightarrow & \bar{x}_1 \longrightarrow & \hat{x}_1 , \\ & Q_o & R_1 & \\ & \downarrow & \downarrow & \\ \text{covariance:} \quad & P_o \longrightarrow & M_1 \longrightarrow & P_1 . \end{array}$$

12.4 Optimal filtering and prediction for linear multistage processes

Consider the linear, stochastic, multistage process described by

$$x_{i+1} = \Phi_i x_i + \Gamma_i w_i , \qquad i = 0, \ldots, N-1, \qquad (12.4.1)$$

where

$$E(x_o) = \bar{x}_o , \qquad (12.4.2)$$

$$E(w_i) = \overline{w}_i , \qquad (12.4.3)$$

$$E(x_o - \bar{x}_o)(x_o - \bar{x}_o)^T = M_o , \qquad (12.4.4)$$

$$E(w_i - \overline{w}_i)(w_j - \overline{w}_j)^T = Q_i \delta_{ij} , \qquad (12.4.5)$$

$$E(w_i - \overline{w}_i)(x_o - \bar{x}_o)^T = 0 . \qquad (12.4.6)$$

Measurements z_i are made while the system is in stage i, and are linearly related to the state x_i by

$$z_i = H_i x_i + v_i , \qquad i = 0, \ldots, N, \qquad (12.4.7)$$

where

$$E(v_i) = 0 , \qquad (12.4.8)$$

$$E(v_i v_j^T) = R_i \delta_{ij} \qquad (12.4.9)$$

$$E(w_i - \overline{w}_i) v_j^T = 0 ,\dagger \qquad \text{and} \qquad E(x_o - \bar{x}_o) v_i^T = 0 . \qquad (12.4.10)$$

It is reasonable to expect (see the derivation in Section 12.7 and Chapter 13) that the weighted-least-square or maximum-likelihood estimate of the state x_k, using only the measurements $z_o, \ldots$

†The case in which w_i and v_j are correlated is considered in Chapter 13.

given by the sequential use of the single-stage estimation procedure of the previous section:

$$\hat{x}_i = \bar{x}_i + K_i(z_i - H_i\bar{x}_i)\,, \qquad (i = 0, \ldots, k, \quad \text{where} \quad k \leq N)\,. \tag{12.4.11}$$

where

$$\bar{x}_{i+1} = \Phi_i\hat{x}_i + \Gamma_i\bar{w}_i\,, \qquad \bar{x}_o \ \text{given}, \tag{12.4.12†}$$

$$K_i = P_i H_i^T R_i^{-1}\,, \tag{12.4.13}$$

$$P_i = (M_i^{-1} + H_i^T R_i^{-1} H_i)^{-1} = M_i - M_i H_i^T (H_i M_i H_i^T + R_i)^{-1} H_i M_i\,, \tag{12.4.14}$$

$$M_{i+1} = \Phi_i P_i \Phi_i^T + \Gamma_i Q_i \Gamma_i^T\,. \tag{12.4.15}$$

This is the *Kalman filter* for linear multistage processes (see Kalman, 1960). Note that the filter (12.4.11) and (12.4.12) is a model of the system (12.4.1), with a correction term proportional to the difference between the actual measurement z_i and the predicted measurement $H_i\bar{x}_i$. The proportionality matrix K_i in (12.4.13) is essentially the ratio between uncertainty in the state P_i and the uncertainty in the measurements R_i; the matrix H_i is simply the state-to-measurement transformation matrix of (12.4.7).

Note that:

1. The propagation of the covariance of the error of the estimate, Equations (12.4.14) and (12.4.15), is *independent* of the measurements z_i. Thus, the covariance matrix can be computed *beforehand* if the parameters of the system and the observation equations are given.
2. The computation of the updated estimate, Equations (12.4.11) and (12.4.12), involves only the *current* measurement and error covariance. Thus, it can easily be carried out in real time.

Prediction of the state beyond the stage where measurements are available, say state m, can be done only by repeated use of (12.4.12); thus,

$$\hat{x}_{i+1} = \bar{x}_{i+1} = \Phi_i\hat{x}_i + \Gamma_i\bar{w}_i\,; \qquad i = m, m+1, \ldots, \tag{12.4.16}$$

where $\hat{x}_m$ is obtained from the filter (12.4.11) through (12.4.15). In other words, the best prediction we can make uses the expected value of w, namely, $\bar{w}_i$, in the transition relations (12.4.1), starting, however, with the filtering estimate of $\hat{x}_m$. Another way of seeing this is to consider $R_i = \infty$ for $i = m, m+1, \ldots$. In this case, (12.4.14) and (12.4.15), (12.4.11) and (12.4.12) reduce to

$$P_{i+1} = \Phi_i P_i \Phi_i^T + \Gamma_i Q_i \Gamma_i^T\,, \tag{12.4.17}$$

---

† $\bar{x}_{i+1}$ in (12.4.12) is, of course, to be understood as $E(x_{i+1}/z_i, \ldots, z_i)$ and not