5.3 Approximate Filter/Smoother Methods

This book has been primarily devoted to the principles underlying various state estimation methods, rather than to addressing practical issues of implementation. A few methods were introduced to reduce computation (Lagrange multipliers, and of ensemble methods), avoiding the computation of the covariance matrices in the sequential (smoother) methods). Lagrange multiplier methods are attractive because they do not demand the covariance matrices; but their main weakness is that they therefore do not provide them.

Unsurprisingly therefore, numerous approaches have attempted to approximate the full results of the filter/smoother algorithms, both to reduce the burden of the state estimates themselves and of the corresponding error covariances. We examine some examples of such approaches.[162]

Steady-state Approximation

Consider, as an example, the Kalman filter, Eqs. (4.49-4.53) of Chapter 4. The error covariances, $\mathbf{P}(t, -)$, $\mathbf{P}(t)$ are propagated as,

$$\mathbf{P}(t, -) = \mathbf{A}(t-1)\mathbf{P}(t-1)\mathbf{A}(t-1)^T + \mathbf{\Gamma}\mathbf{Q}(t-1)\mathbf{\Gamma}^T, \tag{5.7} \quad \text{\{riccati1\}}$$

$$\mathbf{P}(t) = \mathbf{P}(t, -) - \mathbf{P}(t, -)\mathbf{E}(t)^T\left[\mathbf{E}(t)\mathbf{P}(t, -)\mathbf{E}(t)^T + \mathbf{R}(t)\right]^{-1}\mathbf{E}(t)\mathbf{P}(t, -), \tag{5.8} \quad \text{\{riccati2\}}$$

and does not involve the actual data (as is true of all linear estimates). These equations can be simply time-stepped from $\mathbf{P}(0)$ to any time t, assuming the availability of $\mathbf{R}(t)$, $\mathbf{E}(t)$ and $\mathbf{P}(0)$. Knowledge of $\mathbf{P}(t)$ then permits the finding of $\mathbf{K}(t)$, and both are determined before any observations actually exist.

Let the model and data stream be time independent, $\mathbf{A}(t) = \mathbf{A}$, $\mathbf{E}(t) = \mathbf{E}$, $\mathbf{Q}(t) = \mathbf{Q}$, $\mathbf{R}(t) = \mathbf{R}$. Substituting for $\mathbf{P}(t, -)$ one has,

$$\begin{aligned} \mathbf{P}(t) = \mathbf{A}\mathbf{P}(t-1)\mathbf{A}^T + \\ \mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^T - \left[\mathbf{A}\mathbf{P}(t-1)\mathbf{A}^T + \mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^T\right]\mathbf{E}^T\left\{\mathbf{E}\left[\mathbf{A}\mathbf{P}(t-1)\mathbf{A}^T + \mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^T\right]\mathbf{E}^T + \mathbf{R}\right\}^{-1} \times \\ \mathbf{E}\left[\mathbf{A}\mathbf{P}(t-1)\mathbf{A}^T + \mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^T\right], \ t = 0, 1, ... \end{aligned} \tag{5.9}$$

Suppose the difference equation (5.9) approaches a steady-state. That is, as $t \to \infty$, $\mathbf{P}(t) = \mathbf{P}(t-1) \equiv \mathbf{P}_\infty$. Then it follows from Eq. (4.52), that $\mathbf{K}(t) = \mathbf{K}_\infty$ also becomes steady. Once $\mathbf{P}$ and $\mathbf{K}$ cease to change, the computational load of the filter is enormously reduced: the
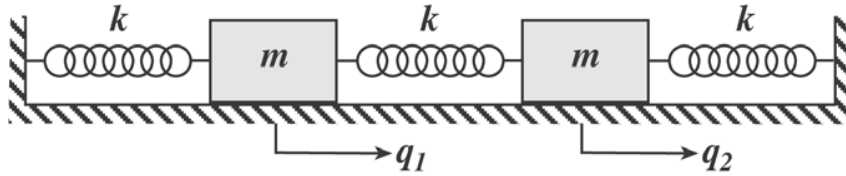
Figure 5.2: Coupled mass-spring oscillator (after McCuskey, 1959). Rest positions of the two masses define the coordinates $q_{1,2}$.

{sspring.tif}

model must be run only once at each time-step. This reduction in load leads one to understand under what circumstances Eq. (5.9), does asymptote to a steady-state, and then for methods to determine that state. With $\mathbf{K}_\infty$ known, one can, if one chooses, use it in place of $\mathbf{K}(t)$, *even during the period when the steady-state is invalid*. To the extent that the system "forgets" its initial conditions, experience suggests that eventually the estimated state will converge to the correct one, even though the initial transient is not properly computed. (A steady-Kalman filter is a "Wiener filter;" they are usually applied by fast convolution methods (which we omit).[163] Similar considerations apply to the problem of obtaining steady-state solutions (Wiener smoother) to the evolution equation for the RTS smoother; further discussion can be found in the references.

Example. *Consider two masses coupled to each other and the boundaries as indicated in Fig. 5.2. A governing set of differential equations for the position, q_i of each oscillator (not to be confused with the generic control variable), is readily shown to be,*

{mass21}
$$\mathbf{M}\frac{d^2\mathbf{q}}{dt^2} + \mathbf{D}\frac{d\mathbf{q}}{dt} + \mathbf{L}\mathbf{q} = \mathbf{f}. \tag{5.10}$$

Here, $\mathbf{M}, \mathbf{D}, \mathbf{K}$ are matrices, $\mathbf{q}(t) = [q_1(t), q_2(t)]^T$ is the non-equilibrium displacement of the masses, and $\mathbf{f}$ is the forcing vector. To generate the simplest case, take $\mathbf{M} = m\mathbf{I}_2$, so that the masses are identical; $\mathbf{D} = r\mathbf{I}_2$, so that the dissipation is of ordinary Rayleigh type, and

$$L = \left\{ \begin{array}{cc} 2k & -k \\ -k & 2k \end{array} \right\}$$

couples the masses through the connecting springs. Using a simple one-sided discretization of

Eq. (5.10), a canonical statespace approximation is,

$$\mathbf{x}(n+1) = \mathbf{A}\mathbf{x}(n) + \mathbf{f}_d(n)$$

$$\mathbf{A} = \left\{ \begin{array}{cc} 2\mathbf{I}_2 - \mathbf{L}(\Delta t)^2/m & (-1+r\Delta t)\mathbf{I}_2 \\ \mathbf{I}_2 & \mathbf{0} \end{array} \right\}, \quad \mathbf{x}(n) = \left[\begin{array}{cccc} q_1(n) & q_2(n) & q_1(n-1) & q_2(n-1) \end{array} \right]^T,$$

$$\mathbf{f}_d(n) = (\Delta t)^2 \left[\begin{array}{cc} \mathbf{f}(n)^T & \mathbf{0} \end{array} \right]^T.$$

Taking $k = 1, m = 1, r = .01, \Delta t = 0.25$, and f to be a unit variance zero mean forcing of q_1 alone (no forcing applied to q_2), a realization of $q_1(t), q_2(t)$ is shown in Fig. 5.3. Now assume that $\mathbf{E} = \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \end{array} \right]$ so that each time step, only $x_1(n)$, that is the position q_1 is measured. Assume $\mathbf{P}(0) = \mathrm{diag}\left(\begin{array}{cccc} 1 & 1 & 1 & 1 \end{array} \right)$, $\mathrm{diag}(\mathbf{R}) = [1, 0]$, and $\mathbf{Q} = \mathbf{I}_4$. Then time-stepping Eq. (5.9) leads to the results for the diagonal elements of $\mathbf{P}(n)$ as depicted in Fig. 5.3. Both P_{11}, P_{12} (and the off-diagonal elements as well) reach steady-state values before $t = n\Delta t = 10$. At that time, $\mathbf{K}(t)$ has become a constant, and one can cease updating either it or the $\mathbf{P}(t)$. ($\mathbf{P}(t, -)$ has of course, also reached a steady-state.)

How might one find the steady-state of Eq. (5.9)—if it exists? Several methods are known. One of them has been used in the above example: time-step the equation until it asymptotes. Other algorithms exist, including a remarkable one called "doubling." In this algorithm, one time steps the equation from $t = 0$, $\mathbf{P}(0)$, to obtain $\mathbf{P}(1\Delta t)$. One then doubles the time step to compute $\mathbf{P}(3\Delta t)$, doubles again for $\mathbf{P}(6\Delta t)$, etc. With this geometric increase in the time step, convergence, if it occurs, is extremely rapid. A simplified equation is treated this way in the Chapter Appendix.[164]

When does a steady-state exist? In general, uncertainty grows because of errors in initial conditions, and the unknown system perturbations (unknown controls, $\mathbf{u}$). Information that reduces uncertainty is provided by the incoming data stream. Under the right circumstances, one can reach an equilibrium where the new information just balances the new uncertainties. A quantitative answer to the question depends directly upon the discussion in Chapter 4 of the observability of the system. Although we omit the formal derivation, one can understand physically why those requirements must be met. Suppose there is an element of the model which is not observable. Then any error, e.g., in its initial conditions, could grow indefinitely, undetected, without bound. Such growth would mean that the corresponding elements of $\mathbf{P}$ would have to grow, and there would be no steady-state. Suppose to the contrary, that such growth *is* observed. Then if those elements are controllable, one can find controls, $\mathbf{u}(t)$, such
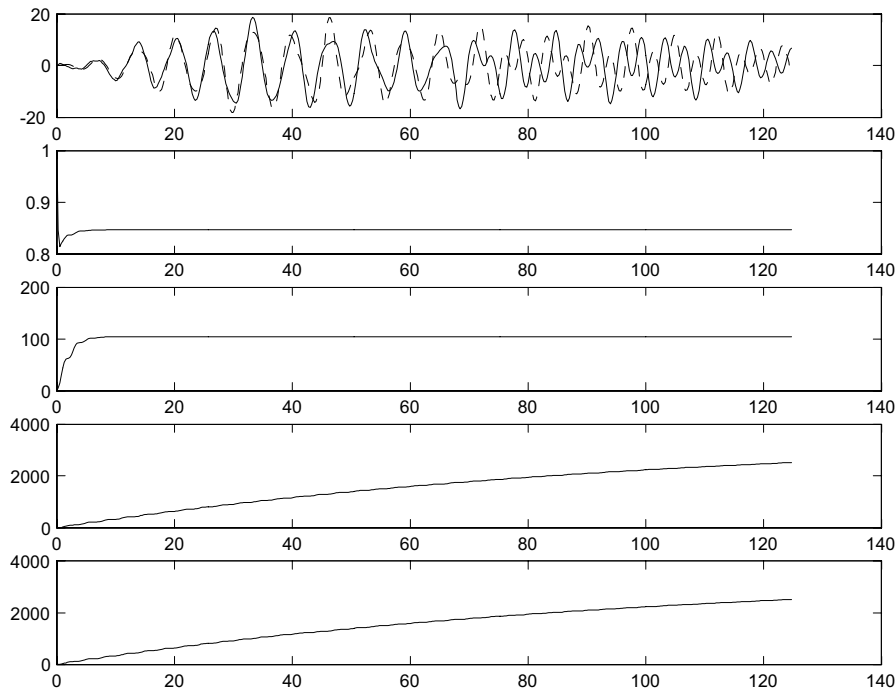
Figure 5.3: Uppermost panel shows the positions, $q_i(t)$, $i = 1, 2$, for the coupled mass-spring oscillator. Second and third panels show $P_{11}(t)$, $P_{22}(t)$ under the assumption that $\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$, that is only $q_1(t)$ is measured (with noise). Both $P_{11}(t)$, $P_{22}(t)$ asymptote to a steady-state, albeit $P_{11}(t) << P_{22}(t)$. Lowest two panels show $P_{11}(t)$, $P_{22}(t)$ when the observation matrix is changed to $\mathbf{E} = \begin{bmatrix} -1 & 1 & 0 & 0 \end{bmatrix}$—that is the observation is of the relative separation of the two masses. In this case, the uncertainty in the absolute positions continues to grow and a steady-state is not reached (there is no dissipation in this example).

{controbserv.e

that the growth is halted; if the elements are not controllable, then they will continue to grow. Note that neither $\mathbf{x}(t)$, nor $\mathbf{u}(t)$ will generally become steady—the state continues to evolve.

In the situation shown in Fig. 5.3, when there is a single measurement of the position, $q_1(t)$, the eigenvalues of the observability matrix, $\mathbf{O}$, range in magnitude from 1.9 down to 0.0067. Albeit there will be considerable uncertainty involved, one can fully determine the initial conditions from the observations. In contrast, when only the relative position, $q_2(t) - q_1(t)$ is measured, two of the eigenvalues of O vanish identically, the system is not completely observable, as as seen in Fig. 5.3, and the uncertainties continue to grow without bound. If one were discussing the smoothing algorithm errors, the structure of $\boldsymbol{\Gamma}$ would enter similarly.[165]