1.2. **Kalman filtering.** The inference aim is to compute $K[X_n \mid Y_{1:n}]$, which we also call the *filter*, recursively in time for the state-space model in (1.4)-(1.5) which is repeated here for convenience:

$$Y_n = g_n X_n + V_n,$$

$$X_{n+1} = f_n X_n + W_n, \qquad n = 1, 2, \ldots \tag{1.11}$$

where $\{V_n\}_n \sim \mathrm{WN}(0, \{r_n\}_n)$, $\{W_n\}_n \sim \mathrm{WN}(0, \{q_n\}_n)$. Furthermore, $X_1$, $\{V_n\}_n$ and $\{W_n\}_n$ are all mutually uncorrelated,

$$\mathrm{Cov}(X_1, W_n) = \mathrm{Cov}(X_1, V_n) = \mathrm{Cov}(W_m, V_n) = 0$$

for all $n \geq 1$ and $m \geq 1$.

The derivation of the Kalman filter has two main components. We commence by assuming we already have calculated $K[X_n \mid Y_{1:n}]$.

- The first step is to convert $K[X_n \mid Y_{1:n}]$ in to $K[X_{n+1} \mid Y_{1:n}]$. This is called the *prediction* step since we are using the observations to infer the value of the state one step into the future.
- The second step is to convert $K[X_{n+1} \mid Y_{1:n}]$ in to $K[X_{n+1} \mid Y_{1:n+1}]$. This is called the *update* step which refines the estimate of $X_{n+1}$ by incorporating the newly observed $Y_{n+1}$.

1.2.1. *The Kalman prediction step.* Using the properties (in Fact 1.1) of the linear prediction $K[\cdot \mid \cdot]$, we can calculate $K[X_{n+1} \mid Y_{1:n}]$ straightforwardly.

Using Fact 1.1 and expanding the definition of $X_{n+1}$,

$$\begin{aligned} K[X_{n+1} \mid Y_{1:n}] &= K[f_n X_n + W_n \mid Y_{1:n}] \\ &= f_n K[X_n \mid Y_{1:n}] + K[W_n \mid Y_{1:n}] \\ &= f_n K[X_n \mid Y_{1:n}] \end{aligned}$$

since $K[W_n \mid Y_{1:n}] = \mathbb{E}(W_n) = 0$ as $\mathrm{Cov}(W_n, Y_i) = 0$ for $i \leq n$.

As a further demonstration, we calculate $K\left[Y_{n+1} \mid Y_{1:n}\right]$ which is an intermediate term we will need later in the Kalman update step. The exact same procedure as above gives

$$
\begin{aligned}
K\left[Y_{n+1} \mid Y_{1:n}\right] &= K\left[g_{n+1}X_{n+1} + V_{n+1} \mid Y_{1:n}\right] \\
&= g_{n+1}K\left[X_{n+1} \mid Y_{1:n}\right] + K\left[V_{n+1} \mid Y_{1:n}\right] \\
&= g_{n+1}f_n K\left[X_n \mid Y_{1:n}\right] \qquad (1.12)
\end{aligned}
$$

since $K\left[V_{n+1} \mid Y_{1:n}\right] = \mathbb{E}(V_{n+1}) = 0$ as $\mathrm{Cov}(V_{n+1}, Y_i) = 0$ for $i \leq n$. Another important quantity we have to calculate sequentially to implement the Kalman filter is the mean square error

$$\sigma_n = \mathbb{E}\left\{\left(X_n - K\left[X_n \mid Y_{1:n}\right]\right)^2\right\}.$$

**Exercise.** Let the mean square error be $\sigma_n = \mathbb{E}\left\{\left(X_n - K\left[X_n \mid Y_{1:n}\right]\right)^2\right\}$. Find $\mathbb{E}\left\{\left(X_{n+1} - K\left[X_{n+1} \mid Y_{1:n}\right]\right)^2\right\}$.

The solution is found by substituting the definition of $X_{n+1}$ and $K[X_{n+1} \mid Y_{1:n}]$ and then expanding:

$$\mathbb{E}\left\{(X_{n+1} - K[X_{n+1} \mid Y_{1:n}])^2\right\}$$

$$= \mathbb{E}\left\{(f_n X_n - f_n K[X_n \mid Y_{1:n}] + W_n)^2\right\}$$

$$= f_n^2 \mathbb{E}\left\{(X_n - K[X_n \mid Y_{1:n}])^2\right\} + \mathbb{E}\left\{W_n^2\right\} + 2 f_n \mathbb{E}\left\{(X_n - K[X_n \mid Y_{1:n}]) W_n\right\}$$

$$= f_n^2 \sigma_n + q_n.$$

This expression makes sense since the mean square error $\sigma_n$ of the estimate of $X_n$ using $Y_{1:n}$ is inflated by the term $q_n$ when the next state $X_{n+1}$ is estimated with the same observations. If the state process is static, $X_{n+1} = X_n$, then mean square error is unchanged

Lets consider why $\mathbb{E}\left\{(X_n - K[X_n \mid Y_{1:n}]) W_n\right\} = 0$. Note that $X_n - K[X_n \mid Y_{1:n}]$ is a linear function of $(X_1, W_1, \ldots, W_{n-1}, 1, Y_1, \ldots, Y_n)$

and the expected value of the product of $W_n$ and any one of these terms is zero.

We have just derived the Kalman predictor.

**Algorithm 1** Kalman prediction for a (non-Gaussian) state-space model

1: Given $\hat{X}_n = K\left[X_n \mid Y_{1:n}\right]$ and the mean square error $\sigma_n = \mathbb{E}\left\{\left(\hat{X}_n - X_n\right)^2\right\}$.

2: $\bar{X}_{n+1} = K\left[X_{n+1} \mid Y_{1:n}\right] = f_n \hat{X}_n$.

3: $\bar{\sigma}_{n+1} = \mathbb{E}\left\{\left(\bar{X}_{n+1} - X_{n+1}\right)^2\right\} = f_n^2 \sigma_n + q_n$.

An important point to remember here is that in linear prediction, both the estimate and its mean square error must be computed.

1.2.2. *The Kalman update step.*

**Definition.** Define the *innovations*,

$$I_{n+1} = Y_{n+1} - K\left[Y_{n+1} \mid Y_{1:n}\right].$$

Note that $K\left[Y_{n+1} \mid Y_{1:n}\right]$ has been calculated in (1.12). Think of $I_{n+1}$ as the *unpredictable* part of $Y_{n+1}$. To derive the Kalman filter, use the following key relationship from Fact 1.1,

(1.13)

$$K\left[\cdot \mid Y_{1:n+1}\right] = K\left[\cdot \mid Y_{1:n}, I_{n+1}\right] = K\left[\cdot \mid Y_{1:n}\right] + K\left[\cdot \mid I_{n+1}\right] - \mathbb{E}(\cdot)$$

This equation is stating two important facts:

- That $K\left[\cdot \mid Y_{1:n+1}\right] = K\left[\cdot \mid Y_{1:n}, I_{n+1}\right]$ and there is no gain or loss in estimation when using either $(Y_{1:n}, Y_{n+1})$ or $(Y_{1:n}, I_{n+1})$. We know this to be true from Fact 1.1 because

$$(Y_{1:n}, I_{n+1})^{\mathrm{T}} = C(Y_{1:n}, Y_{n+1})^{\mathrm{T}} + \mathbf{b}$$

through some invertible matrix $C$ and vector $\mathbf{b}$. (Check this.) An invertible linear transformation of the data does not change the best linear estimate it produces.

- $K\left[\cdot \mid Y_{1:n}, I_{n+1}\right]$ is "linear" in its second argument because $I_{n+1}$ is uncorrelated with $Y_i$, $i \leq n$. This again follows from Fact 1.1.

Thus we have the following useful result for sequential estimation.

**Fact.** *When* $I_{n+1} = Y_{n+1} - K\left[Y_{n+1} \mid Y_{1:n}\right]$ *then*

$$\hat{X}_{n+1} = K\left[X_{n+1} \mid Y_{1:n+1}\right] = K\left[X_{n+1} \mid Y_{1:n}\right] + K\left[X_{n+1} \mid I_{n+1}\right] - \mathbb{E}(X_{n+1}).$$

Having calculated the first term $\bar{X}_{n+1} = K\left[X_{n+1} \mid Y_{1:n}\right]$, we only need to calculate $K\left[X_{n+1} \mid I_{n+1}\right]$ to find $K\left[X_{n+1} \mid Y_{1:n+1}\right]$.

**Fact 1.2.** *The estimate of* $X_{n+1}$ *using* $I_{n+1}$ *is*

$$K\left[X_{n+1} \mid I_{n+1}\right] - \mathbb{E}(X_{n+1}) = \frac{g_{n+1}\left(f_n^2 \sigma_n + q_n\right)}{g_{n+1}^2\left(f_n^2 \sigma_n + q_n\right) + r_{n+1}} I_{n+1}.$$

We will now derive this result. Using the characterisation of the solution in Fact 1.1 and that $\mathbb{E}(I_{n+1}) = 0$,

$$K\left[X_{n+1} \mid I_{n+1}\right] = \mathbb{E}(X_{n+1}) + \frac{\mathbb{E}(X_{n+1} I_{n+1})}{\mathbb{E}(I_{n+1}^2)} I_{n+1}. \tag{1.14}$$

Expand $I_{n+1}$ using the method for equation (1.12),

$$\begin{aligned} I_{n+1} &= g_{n+1} X_{n+1} + V_{n+1} - K\left[g_{n+1} X_{n+1} \mid Y_{1:n}\right] - \underbrace{K\left[V_{n+1} \mid Y_{1:n}\right]}_{=0} \\ &= g_{n+1}\left(X_{n+1} - K\left[X_{n+1} \mid Y_{1:n}\right]\right) + V_{n+1}. \end{aligned}$$

The denominator of (1.14) is

$$\begin{aligned} \mathbb{E}(I_{n+1}^2) &= g_{n+1}^2 \mathbb{E}\left\{\left(X_{n+1} - K\left[X_{n+1} \mid Y_{1:n}\right]\right)^2\right\} + \mathbb{E}\left(V_{n+1}^2\right) + \text{cross terms} \\ &= g_{n+1}^2\left(f_n^2 \sigma_n + q_n\right) + r_{n+1} \end{aligned}$$

where the bracketed expression is given in the derivation of the Kalman predictor and cross-term, $\mathbb{E}\left\{\left(X_{n+1} - K\left[X_{n+1} \mid Y_{1:n}\right]\right) V_{n+1}\right\}$, has zero expectation. Recall that $K\left[X_{n+1} \mid Y_{1:n}\right]$ is a linear function of $(1, Y_1, \ldots, Y_n)$ and the expected value of the product of $V_{n+1}$ and any one of these terms is zero. Likewise, $X_{n+1}$ is a linear function of $(X_1, W_1, \ldots, W_n)$ and the expected value of the product of $V_{n+1}$ and any one of these terms is also zero. These facts imply $\mathbb{E}\left\{\left(X_{n+1} - K\left[X_{n+1} \mid Y_{1:n}\right]\right) V_{n+1}\right\} = 0$.

The numerator of (1.14) is

$$
\begin{aligned}
&\mathbb{E}(X_{n+1} I_{n+1}) \\
&= g_{n+1}\mathbb{E}\left\{X_{n+1}\left(X_{n+1} - K\left[X_{n+1} \mid Y_{1:n}\right]\right)\right\} + \underbrace{\mathbb{E}\left(X_{n+1} V_{n+1}\right)}_{=0} \\
&= g_{n+1}\mathbb{E}\left\{\left(X_{n+1} - K\left[X_{n+1} \mid Y_{1:n}\right]\right)^2\right\} \\
&\quad + g_{n+1}\mathbb{E}\left\{K\left[X_{n+1} \mid Y_{1:n}\right]\left(X_{n+1} - K\left[X_{n+1} \mid Y_{1:n}\right]\right)\right\} \\
&= g_{n+1}\mathbb{E}\left\{\left(X_{n+1} - K\left[X_{n+1} \mid Y_{1:n}\right]\right)^2\right\}
\end{aligned}
$$

1 noting that $\mathbb{E}\left(X_{n+1}V_{n+1}\right) = 0$ and

2 $$\mathbb{E}\left\{\left(X_{n+1} - K\left[X_{n+1} \mid Y_{1:n}\right]\right) K\left[X_{n+1} \mid Y_{1:n}\right]\right\} = 0$$

since the error of the prediction has mean zero and is orthogonal to all the random variables from the set $\{Y_1, \ldots, Y_n\}$. (Remember that $K\left[X_{n+1} \mid Y_{1:n}\right]$ is a linear function of $(1, Y_1, \ldots, Y_n)$.) Thus

$$\begin{aligned}\mathbb{E}(X_{n+1}I_{n+1}) &= g_{n+1}\mathbb{E}\left\{\left(X_{n+1} - K\left[X_{n+1} \mid Y_{1:n}\right]\right)^2\right\} \\ &= g_{n+1}\left(f_n^2\sigma_n + q_n\right).\end{aligned}$$

3 This concludes the verification of Fact 1.2.

4 The final step is to calculate the filter's mean square error.

5 **Fact 1.3.** *The mean square error of the updated estimate $\hat{X}_{n+1}$, which*

6 *we denote as*

7 $$\sigma_{n+1} = \mathbb{E}\left\{\left(X_{n+1} - \hat{X}_{n+1}\right)^2\right\}$$

1 *is given by*

2 $$\sigma_{n+1} = \bar{\sigma}_{n+1}\left(1 - \frac{g_{n+1}^2 \bar{\sigma}_{n+1}}{g_{n+1}^2 \bar{\sigma}_{n+1} + r_{n+1}}\right).$$

3 (Verification.) Subtract $X_{n+1}$ from both sides of

$$\hat{X}_{n+1} = K\left[X_{n+1} \mid Y_{1:n}\right] + K\left[X_{n+1} \mid I_{n+1}\right] - \mathbb{E}(X_{n+1})$$

to get

$$X_{n+1} - \hat{X}_{n+1} + \left(K\left[X_{n+1} \mid I_{n+1}\right] - \mathbb{E}(X_{n+1})\right) = X_{n+1} - K\left[X_{n+1} \mid Y_{1:n}\right].$$

4 Now square both sides and take the expectation to get

$$\mathbb{E}\left\{\left(X_{n+1} - \hat{X}_{n+1}\right)^2\right\} + \mathbb{E}\left\{\left(K\left[X_{n+1} \mid I_{n+1}\right] - \mathbb{E}(X_{n+1})\right)^2\right\} + \text{ct}$$
$$= \mathbb{E}\left\{\left(X_{n+1} - K\left[X_{n+1} \mid Y_{1:n}\right]\right)^2\right\}$$

where ct denotes the cross term which is zero since $\mathbb{E}\left\{\left(X_{n+1}-\hat{X}_{n+1}\right) I_{n+1}\right\} = 0$; in fact $X_{n+1}-\hat{X}_{n+1}$ has zero mean and is orthogonal to all the terms $Y_1,\ldots,Y_n,I_{n+1}$ that were used to define $\hat{X}_{n+1}$.

Since

$$K\left[X_{n+1} \mid I_{n+1}\right]-\mathbb{E}(X_{n+1}) = \frac{\mathbb{E}(X_{n+1}I_{n+1})}{\mathbb{E}(I_{n+1}^2)}I_{n+1},$$

$$\mathbb{E}\left\{\left(K\left[X_{n+1} \mid I_{n+1}\right]-\mathbb{E}(X_{n+1})\right)^2\right\} = \frac{\mathbb{E}(X_{n+1}I_{n+1})^2}{\mathbb{E}(I_{n+1}^2)^2}\mathbb{E}(I_{n+1}^2) = \frac{g_{n+1}^2\bar{\sigma}_{n+1}^2}{g_{n+1}^2\bar{\sigma}_{n+1}+r_{n+1}}.$$

Combining gives

$$\sigma_{n+1} = \bar{\sigma}_{n+1}\left(1-\frac{g_{n+1}^2\bar{\sigma}_{n+1}}{g_{n+1}^2\bar{\sigma}_{n+1}+r_{n+1}}\right) = \frac{\bar{\sigma}_{n+1}r_{n+1}}{g_{n+1}^2\bar{\sigma}_{n+1}+r_{n+1}}.$$

All the calculations that have been performed in the Kalman update step are now summarised.

**Algorithm 2** Kalman update for a (non-gaussian) state-space model

1: Given $\bar{X}_{n+1} = K[X_{n+1} \mid Y_{1:n}] = f_n \hat{X}_n$ and $\bar{\sigma}_{n+1} = \mathbb{E}\left\{\left(\bar{X}_{n+1} - X_{n+1}\right)^2\right\}$.

2: $I_{n+1} = Y_{n+1} - g_{n+1}\bar{X}_{n+1}$.

3: $\hat{X}_{n+1} = \bar{X}_{n+1} + \frac{g_{n+1}\bar{\sigma}_{n+1}}{g_{n+1}^2\bar{\sigma}_{n+1}+r_{n+1}} I_{n+1}$.

4: $\sigma_{n+1} = \bar{\sigma}_{n+1}\left(1 - \frac{g_{n+1}^2\bar{\sigma}_{n+1}}{g_{n+1}^2\bar{\sigma}_{n+1}+r_{n+1}}\right)$.

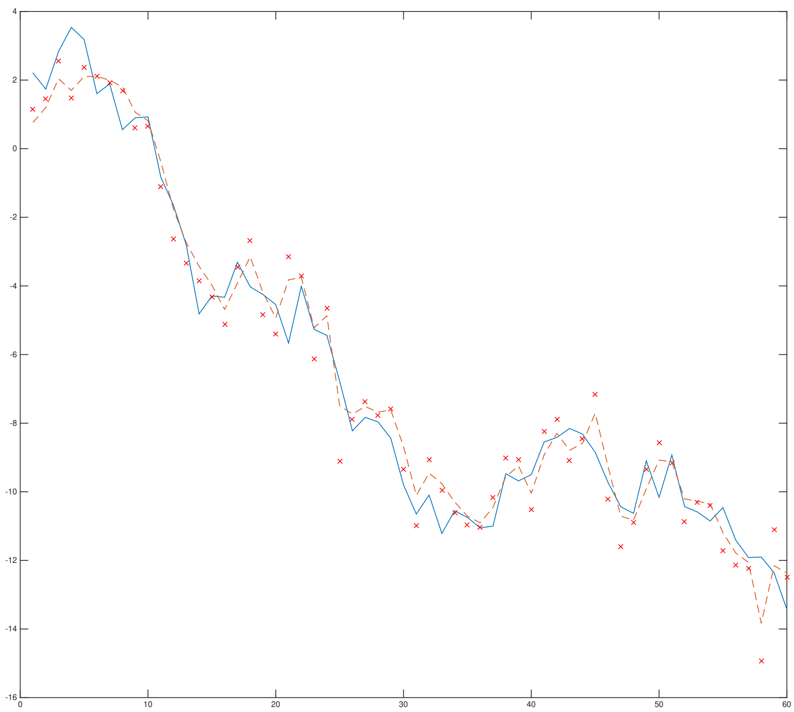


FIGURE 1.1. The Kalman filter for the state-space model in (1.11) with $f_n = g_n = 1$ and noises having unit variance. Solid line is the true state, crosses are the observations and the dashed line is the Kalman estimate of the state using Algorithm 2.