

11 Kalman Filtering and Certainty Equivalence

We presents the important concepts of the Kalman filter, certainty-equivalence and the separation principle.

11.1 Preliminaries

Lemma 11.1 *Suppose x and y are jointly normal with zero means and covariance matrix*

$$\text{cov} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} V_{xx} & V_{xy} \\ V_{yx} & V_{yy} \end{bmatrix}.$$

Then the distribution of x conditional on y is Gaussian, with

$$E(x | y) = V_{xy}V_{yy}^{-1}y, \quad (11.1)$$

and

$$\text{cov}(x | y) = V_{xx} - V_{xy}V_{yy}^{-1}V_{yx}. \quad (11.2)$$

Proof. Both y and $x - V_{xy}V_{yy}^{-1}y$ are linear functions of x and y and therefore they are Gaussian. From $E[(x - V_{xy}V_{yy}^{-1}y)y^\top] = 0$ it follows that they are uncorrelated and this implies they are independent. Hence the distribution of $x - V_{xy}V_{yy}^{-1}y$ conditional on y is identical with its unconditional distribution, and this is Gaussian with zero mean and the covariance matrix given by (11.2) ■

The estimate of x in terms of y defined as $\hat{x} = Hy = V_{xy}V_{yy}^{-1}y$ is known as the **linear least squares estimate** of x in terms of y . Even without the assumption that x and y are jointly normal, this linear function of y has a smaller covariance matrix than any other unbiased estimate for x that is a linear function of y . In the Gaussian case, it is also the maximum likelihood estimator.

11.2 The Kalman filter

Let us make the LQG and state-structure assumptions of Section 10.4.

$$x_t = Ax_{t-1} + Bu_{t-1} + \epsilon_t, \quad (11.3)$$

$$y_t = Cx_{t-1} + \eta_t, \quad (11.4)$$

Notice that both x_t and y_t can be written as a linear functions of the unknown noise and the known values of u_0, \dots, u_{t-1} . Thus the distribution of x_t conditional on $W_t = (Y_t, U_{t-1})$ must be normal, with some mean \hat{x}_t and covariance matrix V_t . The following theorem describes recursive updating relations for these two quantities.

Theorem 11.2 (The Kalman filter) *Suppose that conditional on W_0 , the initial state x_0 is distributed $N(\hat{x}_0, V_0)$ and the state and observations obey the recursions of*

the LQG model (11.3)–(11.4). Then conditional on W_t , the current state is distributed $N(\hat{x}_t, V_t)$. The conditional mean and variance obey the updating recursions

$$\hat{x}_t = A\hat{x}_{t-1} + Bu_{t-1} + H_t(y_t - C\hat{x}_{t-1}), \quad (11.5)$$

$$V_t = N + AV_{t-1}A^\top - (L + AV_{t-1}C^\top)(M + CV_{t-1}C^\top)^{-1}(L^\top + CV_{t-1}A^\top), \quad (11.6)$$

where

$$H_t = (L + AV_{t-1}C^\top)(M + CV_{t-1}C^\top)^{-1}. \quad (11.7)$$

Proof. The proof is by induction on t . Consider the moment when u_{t-1} has been determined but y_t has not yet observed. The distribution of (x_t, y_t) conditional on (W_{t-1}, u_{t-1}) is jointly normal with means

$$E(x_t | W_{t-1}, u_{t-1}) = A\hat{x}_{t-1} + Bu_{t-1},$$

$$E(y_t | W_{t-1}, u_{t-1}) = C\hat{x}_{t-1}.$$

Let $\Delta_{t-1} = \hat{x}_{t-1} - x_{t-1}$, which by an inductive hypothesis is $N(0, V_{t-1})$. Consider the **innovations**

$$\xi_t = x_t - E(x_t | W_{t-1}, u_{t-1}) = x_t - (A\hat{x}_{t-1} + Bu_{t-1}) = \epsilon_t - A\Delta_{t-1},$$

$$\zeta_t = y_t - E(y_t | W_{t-1}, u_{t-1}) = y_t - C\hat{x}_{t-1} = \eta_t - C\Delta_{t-1}.$$

Conditional on (W_{t-1}, u_{t-1}) , these quantities are normally distributed with zero means and covariance matrix

$$\text{cov} \begin{bmatrix} \epsilon_t - A\Delta_{t-1} \\ \eta_t - C\Delta_{t-1} \end{bmatrix} = \begin{bmatrix} N + AV_{t-1}A^\top & L + AV_{t-1}C^\top \\ L^\top + CV_{t-1}A^\top & M + CV_{t-1}C^\top \end{bmatrix} = \begin{bmatrix} V_{\xi\xi} & V_{\xi\zeta} \\ V_{\zeta\xi} & V_{\zeta\zeta} \end{bmatrix}.$$

Thus it follows from Lemma 11.1 that the distribution of ξ_t conditional on knowing $(W_{t-1}, u_{t-1}, \zeta_t)$, (which is equivalent to knowing W_t), is normal with mean $V_{\xi\zeta}V_{\zeta\zeta}^{-1}\zeta_t$ and covariance matrix $V_{\xi\xi} - V_{\xi\zeta}V_{\zeta\zeta}^{-1}V_{\zeta\xi}$. These give (11.5)–(11.7). ■

11.3 Certainty equivalence

We say that a quantity a is *policy-independent* if $E_\pi(a | W_0)$ is independent of π .

Theorem 11.3 *Suppose LQG model assumptions hold. Then (i)*

$$F(W_t) = \hat{x}_t^\top \Pi_t \hat{x}_t + \dots \quad (11.8)$$

where \hat{x}_t is the linear least squares estimate of x_t whose evolution is determined by the Kalman filter in Theorem 11.2 and ‘+...’ indicates terms that are policy independent; (ii) the optimal control is given by

$$u_t = K_t \hat{x}_t,$$

where Π_t and K_t are the same matrices as in the full information case of Theorem 7.2.

It is important to grasp the remarkable fact that (ii) asserts: *the optimal control u_t is exactly the same as it would be if all unknowns were known and took values equal to their linear least square estimates (equivalently, their conditional means) based upon observations up to time t .* This is the idea known as **certainty equivalence**. As we have seen in the previous section, the distribution of the estimation error $\hat{x}_t - x_t$ does not depend on U_{t-1} . The fact that the problems of optimal estimation and optimal control can be decoupled in this way is known as the **separation principle**.

Proof. The proof is by backward induction. Suppose (11.8) holds at t . Recall that

$$\hat{x}_t = A\hat{x}_{t-1} + Bu_{t-1} + H_t\zeta_t, \quad \Delta_{t-1} = \hat{x}_{t-1} - x_{t-1}.$$

Then with a quadratic cost of the form $c(x, u) = x^\top Rx + 2u^\top Sx + u^\top Qu$, we have

$$\begin{aligned} F(W_{t-1}) &= \min_{u_{t-1}} E [c(x_{t-1}, u_{t-1}) + \hat{x}_t^\top \Pi_t \hat{x}_t + \dots \mid W_{t-1}, u_{t-1}] \\ &= \min_{u_{t-1}} E [c(\hat{x}_{t-1} - \Delta_{t-1}, u_{t-1}) \\ &\quad + (A\hat{x}_{t-1} + Bu_{t-1} + H_t\zeta_t)^\top \Pi_t (A\hat{x}_{t-1} + Bu_{t-1} + H_t\zeta_t) \mid W_{t-1}, u_{t-1}] \\ &= \min_{u_{t-1}} [c(\hat{x}_{t-1}, u_{t-1}) + (A\hat{x}_{t-1} + Bu_{t-1})^\top \Pi_t (A\hat{x}_{t-1} + Bu_{t-1})] + \dots, \end{aligned}$$

where we use the fact that conditional on W_{t-1}, u_{t-1} , both Δ_{t-1} and ζ_t have zero means and are policy independent. This ensures that when we expand the quadratics in powers of Δ_{t-1} and $H_t\zeta_t$ the expected value of the linear terms in these quantities are zero and the expected value of the quadratic terms (represented by $+\dots$) are policy independent. ■

11.4 Example: inertialess rocket with noisy position sensing

Consider the scalar case of controlling the position of a rocket by inertialess control of its velocity but in the presence of imperfect position sensing.

$$x_t = x_{t-1} + u_{t-1}, \quad y_t = x_t + \eta_t,$$

where η_t is white noise with variance 1. Suppose it is desired to minimize

$$E \left[\sum_{t=0}^{h-1} u_t^2 + Dx_h^2 \right].$$

Notice that the observational relation differs from the usual model of $y_t = Cx_{t-1} + \eta_t$. To derive a Kalman filter formulae for this variation we argue inductively from scratch. Suppose $\hat{x}_{t-1} - x_{t-1} \sim N(0, V_{t-1})$. Consider a linear estimate of x_t ,

$$\hat{x}_t = \hat{x}_{t-1} + u_{t-1} + H_t(y_t - \hat{x}_{t-1} - u_{t-1}).$$

(The relevant innovation process is now $\tilde{y}_t = y_t - \hat{x}_{t-1} - u_{t-1}$.) Subtracting the plant equation and substituting for x_t and y_t gives

$$\Delta_t = \Delta_{t-1} + H_t(\eta_t - \Delta_{t-1}).$$

The variance of Δ_t is therefore

$$\text{var } \Delta_t = V_{t-1} - 2H_tV_{t-1} + H_t^2(1 + V_{t-1}).$$

Minimizing this with respect to H_t gives $H_t = V_{t-1}(1 + V_{t-1})^{-1}$, so the variance in the least squares estimate of x_t obeys the recursion,

$$V_t = V_{t-1} - V_{t-1}^2(1 + V_{t-1})^{-1} = V_{t-1}/(1 + V_{t-1}).$$

Hence

$$V_t^{-1} = V_{t-1}^{-1} + 1 = \dots = V_0^{-1} + t.$$

If there is complete lack of information at the start, then $V_0^{-1} = 0$, $V_t = 1/t$ and

$$\hat{x}_t = \hat{x}_{t-1} + u_{t-1} + \frac{V_{t-1}(y_t - \hat{x}_{t-1} - u_{t-1})}{1 + V_{t-1}} = \frac{(t-1)(\hat{x}_{t-1} + u_{t-1}) + y_t}{t}.$$

As far as the optimal control is concerned, suppose an inductive hypothesis that $F(W_t) = \hat{x}_t^2 \Pi_t + \dots$, where ‘ \dots ’ denotes policy independent terms. Then

$$\begin{aligned} F(W_{t-1}) &= \inf_u \{u^2 + E[\hat{x}_{t-1} + u + H_t(y_t - \hat{x}_{t-1} - u)]^2 \Pi_t + \dots\} \\ &= \inf_u \{u^2 + (\hat{x}_{t-1} + u)^2 \Pi_t + E[H_t(\eta_t - \Delta_{t-1})]^2 \Pi_t + \dots\}. \end{aligned}$$

Minimizing over u we obtain the usual Riccati recursion of

$$\Pi_{t-1} = \Pi_t - \Pi_t^2/(1 + \Pi_t) = \Pi_t/(1 + \Pi_t).$$

Hence $\Pi_t = D/(1 + D(h-t))$ and the optimal control is the certainty equivalence control $u_t = -D\hat{x}_t/(1 + D(h-t))$. This is the same control as in the deterministic case, but with x_t replaced by \hat{x}_t .