Kalman Filter and Parameter Identification

Florian Herzog

2013

Continuous-time Kalman Filter

In this chapter, we shall use stochastic processes with independent increments $w_1(.)$ and $w_2(.)$ at the input and the output, respectively, of a dynamic system. We shall switch back and forth between the mathematically precise description of these (normalized) Brownian motions by their increments and the sloppy description by the corresponding white noises $\dot{v}(.)$ and $\dot{r}(.)$, respectively, which is preferred in engineering circles

$$
\begin{aligned}
Q^{1/2}(t)dw_1(t) &= [\dot{v}(t) - \overline{u}(t)]dt && (1)\\
R^{1/2}(t)dw_2(t) &= [\dot{r}(t) - \overline{r}(t)]dt\,, && (2)
\end{aligned}
$$

where $Q^{1/2}(t)$ and $R^{1/2}(t)$ are the volatility parameters.
We assume a linear dynamical system, where the inputs $u(t)$ and the measurements $y(t)$ are disturbed by noise.

Continuous-time Kalman Filter

Consider the following linear time-varying dynamic system of order n which is driven by the m-vector-valued white noise $\dot{v}(.)$. Its initial state $x(t_0)$ is a random vector ξ and its p-vector-valued output $y(.)$ is corrupted by the additive white noise $\dot{r}(.)$:

System description in the mathematically precise form:

$$\begin{aligned}
dx(t) &= A(t)x(t)dt + B(t)dv(t) \\
&= [A(t)x(t) + B(t)\overline{u}(t)]dt + B(t)Q^{1/2}(t)dw_1(t) && (3) \\
x(t_0) &= \xi && (4) \\
y(t)dt &= C(t)x(t)dt + dr(t) \\
&= [C(t)x(t) + \overline{r}(t)]dt + R^{1/2}(t)dw_2(t)\,, && (5) \\
& && (6)
\end{aligned}$$

Continuous-time Kalman Filter

System description in the engineer's form:

$$\dot{x}(t) = A(t)x(t) + B(t)\dot{v}(t) \tag{7}$$

$$x(t_0) = \xi \tag{8}$$

$$y(t) = C(t)x(t) + \dot{r}(t)\,, \tag{9}$$

where

$$\xi \;:\; \mathcal{N}(x_0, \Sigma_0) \tag{10}$$

$$\dot{v}(t) \;:\; \mathcal{N}(\overline{u}(t), Q(t)) \tag{11}$$

$$\dot{r}(t) \;:\; \mathcal{N}(\overline{r}(t), R(t))\,. \tag{12}$$

Note that in the last two lines, we have deleted the factor $\frac{1}{dt}$ which ought to multiply $Q(t)$ and $R(t)$.

Continuous-time Kalman Filter

The filtering problem is stated as follows: Find the optimal filter with the state vector $\widehat{x}$ which is optimal in the following sense:

- The state estimation is bias-free:
  $E\{x(t) - \widehat{x}(t)\} \equiv 0$.
- The error of the state estimate has infimal covariance matrix:
  $\Sigma_{opt}(t) \leq \Sigma_{subopt}(t)$.

Continuous-time Kalman Filter

With the above-mentioned assumptions that ξ, v, and r are mutually independent and Gaussian, the optimal filter is the following linear dynamic system:

$$\begin{aligned}
\dot{\widehat{x}}(t) &= A(t)\widehat{x}(t) + B(t)\overline{u}(t) + H(t)[y(t) - \overline{r}(t) - C(t)\widehat{x}(t)] \\
&= [A(t) - H(t)C(t)]\widehat{x}(t) + B(t)\overline{u}(t) + H(t)[y(t) - \overline{r}(t)] \quad (13) \\
\widehat{x}(t_0) &= x_0 \quad (14)
\end{aligned}$$

with

$$H(t) = \Sigma(t)C^T(t)R^{-1}(t) \ , \quad (15)$$

where $\Sigma(t)$ is the covariance matrix of the error $\widehat{x}(t) - x(t)$ of the state estimate $\widehat{x}(t)$.

Continuous-time Kalman Filter

$\Sigma(t)$ is the covariance matrix of the error $\widehat{x}(t) - x(t)$ of the state estimate $\widehat{x}(t)$ satisfying the following matrix Riccati differential equation:

$$\begin{aligned}\dot{\Sigma}(t) &= A(t)\Sigma(t) + \Sigma(t)A^T(t) \\ &\quad - \Sigma(t)C^T(t)R^{-1}(t)C(t)\Sigma(t) + B(t)Q(t)B^T(t) \qquad (16)\\ \Sigma(t_0) &= \Sigma_0 \; . \qquad (17)\end{aligned}$$

The estimation error $e(t) = x(t) - \widehat{x}(t)$ satisfies the differential equation

$$\begin{aligned}\dot{e} &= \dot{x} - \dot{\widehat{x}} \\ &= Ax + B\overline{u} + B(\dot{v} - \overline{u}) - [A - HC]\widehat{x} - B\overline{u} - HCx - H(\dot{r} - \overline{r}) \\ &= [A - HC]e + B(\dot{v} - \overline{u}) - H(\dot{r} - \overline{r}) \; . \qquad (18)\end{aligned}$$

Continuous-time Kalman Filter

The covariance matrix $\Sigma(t)$ of the state estimation errer $e(t)$ is governed by the matrix differential equation

$$
\begin{aligned}
\dot{\Sigma} &= [A-HC]\Sigma + \Sigma[A-HC]^T + BQB^T + HRH^T && (19)\\
&= A\Sigma + \Sigma A^T + BQB^T - \Sigma C^T R^{-1} C\Sigma \\
&\quad + [H-\Sigma C^T R^{-1}]\, R\, [H-\Sigma C^T R^{-1}]^T \\
&\geq A\Sigma + \Sigma A^T + BQB^T - \Sigma C^T R^{-1} C\Sigma\ . && (20)
\end{aligned}
$$

Obviously, $\dot{\Sigma}(t)$ is infimized at all times t for

$$H(t) = \Sigma(t) C^T(t) R^{-1}(t)\ . \tag{21}$$

Continuous-time Kalman Filter

Given (19) for the covariance matrix $\Sigma(t)$:

$$\dot{\Sigma} = [A - HC]\Sigma + \Sigma[A - HC]^T + BQB^T + HRH^T \ .$$

Its first differential with respect to H, at some value of H, with increment dH can be formulated as follows:

$$\begin{aligned} d\dot{\Sigma}(H, dH) &= \frac{\partial \dot{\Sigma}(H)}{\partial H} dH \\ &= -\, dHC\Sigma - \Sigma C^T dH^T + dHRH^T + HRdH^T \\ &= U[HR - \Sigma C^T] T dH \ , \end{aligned} \tag{22}$$

where T is the matrix transposing operator and U is the operator which adds the transposed matrix to its matrix argument.

Continuous-time Kalman Filter

Consider the following stochastic system of first order with the random initial condition $\xi : \mathcal{N}(x_0, \Sigma_0)$ and the uncorrelated white noises $\dot{v} : \mathcal{N}(0, Q)$ and $\dot{r} : \mathcal{N}(0, R)$:

$$\begin{aligned} \dot{x}(t) &= ax(t) + b[u(t) + \dot{v}(t)] \\ x(0) &= \xi \\ y(t) &= x(t) + \dot{r}(t) \,. \end{aligned}$$

Find the time-invariant Kalman filter! The resulting time-invariant Kalman filter is described by the following equations:

$$\begin{aligned} \dot{\widehat{x}}(t) &= -\sqrt{a^2 + b^2\frac{Q}{R}}\,\widehat{x}(t) + bu(t) + \left(a + \sqrt{a^2 + b^2\frac{Q}{R}}\right) y(t) \\ \widehat{x}(0) &= x_0 \,. \end{aligned}$$

Continuous/discrete-time Kalman Filter

The continuous-time measurement (9) which is corrupted by the white noise error $\dot{r}(t)$ with infinite covariance,

$$y(t) = C(t)x(t) + \dot{r}(t) \ ,$$

must be replaced by an "averaged" discrete-time measurement

$$y_k = C_k x(t_k) + \widetilde{r}_k \tag{23}$$

$$\widetilde{r}_k \ : \ \mathcal{N}(\overline{r}_k, R_k) \tag{24}$$

$$\text{Cov}(\widetilde{r}_i, \widetilde{r}_j) = R_i \delta_{ij} \ . \tag{25}$$

Continuous/discrete-time Kalman Filter

At any sampling time t_k we denote measurements by $\widehat{x}(t_k|t_{k-1})$ and error covariance matrix $\Sigma(t_k|t_{k-1})$ before the new measurement y_k has been processed and the state estimate $\widehat{x}(t_k|t_k)$ and the corresponding error covariance matrix $\Sigma(t_k|t_k)$ after the new measurement y_k has been processed.
As a consequence, the continuous-time/discrete-time Kalman filter alternatingly performs update steps at the sampling times t_k processing the latest measurement information y_k and open-loop extrapolation steps between the times t_k and t_{K+1} with no measurement information available:

Continuous/discrete-time Kalman Filter

The Continuous-Time/Discrete-Time Kalman Filter

Update at time t_k:

$$\begin{aligned}\widehat{x}(t_k|t_k) &= \widehat{x}(t_k|t_{k-1}) \\ &\quad + \Sigma(t_k|t_{k-1})C_k^T\left[C_k\Sigma(t_k|t_{k-1})C_k^T + R_k\right]^{-1} \\ &\qquad \times [y_k - \overline{r}_k - C_k\widehat{x}(t_k|t_{k-1})] \qquad (26)\\ \Sigma(t_k|t_k) &= \Sigma(t_k|t_{k-1}) \\ &\quad - \Sigma(t_k|t_{k-1})C_k^T\left[C_k\Sigma(t_k|t_{k-1})C_k^T + R_k\right]^{-1} C_k\Sigma(t_k|t_{k-1}) \qquad (27)\end{aligned}$$

or, equivalently:

$$\Sigma^{-1}(t_k|t_k) = \Sigma^{-1}(t_k|t_{k-1}) + C_k^T R_k^{-1} C_k. \qquad (28)$$

Continuous/discrete-time Kalman Filter

Continuous-time extrapolation from t_k to t_{k+1} with the initial conditions $\widehat{x}(t_k|t_k)$ and $\Sigma(t_k|t_k)$:

$$
\begin{aligned}
\dot{\widehat{x}}(t|t_k) &= A(t)\widehat{x}(t|t_k) + B(t)\overline{u}(t) &\quad (29)\\
\dot{\Sigma}(t|t_k) &= A(t)\Sigma(t|t_k) + \Sigma(t|t_k)A^T(t) + B(t)Q(t)B^T(t)\ . &\quad (30)
\end{aligned}
$$

Assuming that the Kalman filter starts with an update at the initial time t_0, this Kalman filter is initialized at t_0 as follows:

$$
\begin{aligned}
\widehat{x}(t_0|t_{-1}) &= x_0 &\quad (31)\\
\Sigma(t_0|t_{-1}) &= \Sigma_0\ . &\quad (32)
\end{aligned}
$$

Discrete-time Kalman Filter

We consider the following discrete-time stochastic system:

$$
\begin{aligned}
x_{k+1} &= F_k x_k + G_k \widetilde{v}_k && (33)\\
x_0 &= \xi && (34)\\
y_k &= C_k x_k + \widetilde{r}_k && (35)\\
\xi &: \mathcal{N}(x_0, \Sigma_0) && (36)\\
\widetilde{v} &: \mathcal{N}(\overline{u}_k, Q_k) && (37)\\
\mathrm{Cov}(\widetilde{v}_i, \widetilde{v}_j) &= Q_i \delta_{ij} && (38)\\
\widetilde{r} &: \mathcal{N}(\overline{r}_k, R_k) && (39)\\
\mathrm{Cov}(\widetilde{r}_i, \widetilde{r}_j) &= R_i \delta_{ij} && (40)\\
\xi,\ \widetilde{v},\ \widetilde{r} &: \text{mutually independent.} && (41)
\end{aligned}
$$

Discrete-time Kalman Filter

We have the following (approximate) "zero-order-hold-equivalence" correspondences to the continuous-time stochastic system:

$$
\begin{aligned}
F_k &= \Phi(t_{k+1}, t_k) && \text{transition matrix} && (42) \\
G_k &= \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, t) B(t) dt && \text{zero order hold equivalence} && (43) \\
C_k &= C(t_k) && && (44) \\
Q_k &= \frac{Q(t_k)}{t_{k+1} - t_k} && && (45) \\
R_k &= \frac{R(t_k)}{\Delta t_k} && \Delta t_k = \text{averaging time at time } t_k && (46)
\end{aligned}
$$

Discrete-time Kalman Filter

The Discrete-Time Kalman Filter

Update at time t_k:

$$\begin{aligned} \widehat{x}_{k|k} &= \widehat{x}_{k|k-1} \\ &\quad + \Sigma_{k|k-1} C_k^T \left[C_k \Sigma_{k|k-1} C_k^T + R_k \right]^{-1} [y_k - \overline{r}_k - C_k \widehat{x}_{k|k-1}] \quad (47) \\ \Sigma_{k|k} &= \Sigma_{k|k-1} - \Sigma_{k|k-1} C_k^T \left[C_k \Sigma_{k|k-1} C_k^T + R_k \right]^{-1} C_k \Sigma_{k|k-1} \quad (48) \end{aligned}$$

or, equivalently:

$$\Sigma_{k|k}^{-1} = \Sigma_{k-1|k-1}^{-1} + C_k^T R_k^{-1} C_k \quad (49)$$

Discrete-time Kalman Filter

Extrapolation from t_k to t_{k+1}:

$$\widehat{x}_{k+1|k} = F_k \widehat{x}_{k|k} + G_k \overline{u}_k \tag{50}$$

$$\Sigma_{k+1|k} = F_k \Sigma_{k|k} F_k^T + G_k Q_k G_k^T \tag{51}$$

Initialization at time t_0:

$$\widehat{x}_{0|-1} = x_0 \tag{52}$$

$$\Sigma_{0|-1} = \Sigma_0 \ . \tag{53}$$

Extended Kalman Filter

The linearized problem may not be a good approximation we stick to the following philosophy:

- Where it does not hurt: Analyze the nonlinear system.
- Where it hurts: use linearization.

We consider the following nonlinear stochastic system:

$$\dot{x}(t) = f(x(t), u(t), t) + B(t)\dot{v}(t) \tag{54}$$

$$x(t_0) = \xi \tag{55}$$

$$y(t) = g(x(t), t) + \dot{r}(t) \tag{56}$$

$$\xi : \mathcal{N}(x_0, \Sigma_0) \tag{57}$$

$$\dot{v} : \mathcal{N}(\overline{v}(t), Q(t)) \tag{58}$$

$$\dot{r} : \mathcal{N}(\overline{r}(t), R(t))\,. \tag{59}$$

The extended Kalman filter

Dynamics of the state estimation:

$$\dot{\widehat{x}}(t) = f(\widehat{x}(t), u(t), t) + B(t)\overline{v}(t) + H(t)[y(t) - \overline{r}(t) - g(\widehat{x}(t), t)] \quad (60)$$

$$\widehat{x}(t_0) = x_0 \quad (61)$$

with $H(t) = \Sigma(t)C^T(t)R^{-1}(t)$. The "error covariance" matrix $\Sigma(t)$ must be calculated in real-time using the folloing matrix Riccati differential equation:

$$\begin{aligned}\dot{\Sigma}(t) &= A(t)\Sigma(t) + \Sigma(t)A^T(t) \\ &\quad - \Sigma(t)C^T(t)R^{-1}(t)C(t)\Sigma(t) + B(t)Q(t)B^T(t)\end{aligned} \quad (62)$$

$$\Sigma(t_0) = \Sigma_0 \ . \quad (63)$$

Extended Kalman Filter

The matrices $A(t)$ and $C(t)$ correspond to the dynamics matrix and the output matrix, respectively, of the linearization of the nonlinear system around the estimated trajectory:

$$A(t) = \frac{\partial f(\widehat{x}(t), u(t), t)}{\partial x} \tag{64}$$

$$C(t) = \frac{\partial g(\widehat{x}(t), t)}{\partial x} \quad . \tag{65}$$

Extended Kalman Filter

Because the dynamics of the "error covariance matrix" $\Sigma(t)$ correspond to the linearized system, we face the following problems:

- This filter is not optimal.
- The state estimate will be biased.
- The reference point for the linearization is questionable.
- The matrix $\Sigma(t)$ is only an approximation of the state error covariance matrix.
- The state vectors $x(t)$ und $\widehat{x}(t)$ are not Gaussian even if ξ, $\dot{v}$, and $\dot{r}$ are.

Extended Kalman Filter

Again, the continous-time measurement corrupted with the white noise $\dot{r}$ with infinite covariance

$$y(t) = g(x(t), t) + \dot{r}(t) \tag{66}$$

must be replaced by an "averaged" dicrete-time measurement

$$\begin{aligned} y_k &= g(x(t_k), t_k) + \widetilde{r}_k &(67)\\ \widetilde{r}_k &: \mathcal{N}(\overline{r}_k, R_k) &(68)\\ \mathrm{Cov}(\widetilde{r}_i, \widetilde{r}_j) &= R_i \delta_{ij}\ . &(69) \end{aligned}$$

Continuous-time/discrete-time extended Kalman filter

Update at time t_k:

$$\begin{aligned}
\widehat{x}(t_k|t_k) &= \widehat{x}(t_k|t_{k-1}) \\
&\quad + \Sigma(t_k|t_{k-1})C_k^T\left[C_k\Sigma(t_k|t_{k-1})C_k^T + R_k\right]^{-1} \\
&\qquad \times [y_k - \overline{r}_k - g(\widehat{x}(t_k|t_{k-1}), t_k)] && (70) \\
\Sigma(t_k|t_k) &= \Sigma(t_k|t_{k-1}) && (71) \\
&\quad - \Sigma(t_k|t_{k-1})C_k^T\left[C_k\Sigma(t_k|t_{k-1})C_k^T + R_k\right]^{-1}C_k\Sigma(t_k|t_{k-1}) && (72)
\end{aligned}$$

Extrapolation between t_k and t_{k+1}:

$$\begin{aligned}
\dot{\widehat{x}}(t|t_k) &= f(\widehat{x}(t|t_k), u(t), t) + B(t)\overline{v}(t) && (73) \\
\dot{\Sigma}(t|t_k) &= A(t)\Sigma(t|t_k) + \Sigma(t|t_k)A^T(t) + B(t)Q(t)B^T(t) && (74)
\end{aligned}$$

Continuous-time/discrete-time extended Kalman filter

Again, the matrices $A(t)$ and C_k are the dynamics matrix and the output matrix, respectively, of the nonlinear system linearized around the estimated trajectory:

$$A(t) = \frac{\partial f(\widehat{x}(t), u(t), t)}{\partial x} \tag{75}$$

$$C_k = \frac{\partial g(\widehat{x}(t_k|t_{k-1}), t_k)}{\partial x} \tag{76}$$

Initialization at t_0*:*

$$\widehat{x}(t_0|t_{-1}) = x_0 \tag{77}$$

$$\Sigma(t_0|t_{-1}) = \Sigma_0 \; . \tag{78}$$

Kalman Filter and Parameter Identification

The system and the measurement equations in discrete time are given by

$$\begin{aligned} x_{k+1} &= F_k x_k + G_k u_k + [G_k Q_k^{\frac{1}{2}}] w_k^{(1)} && (79)\\ y_k &= C_k x_k + \overline{r}_k + [R_k^{\frac{1}{2}}] w_k^{(2)} && (80) \end{aligned}$$

The update equations of the Kalman Filter are

$$\begin{aligned} \widehat{x}_{k|k} &= \widehat{x}_{k|k-1} + \Sigma_{k|k-1} C_k^T [C_k \Sigma_{k|k-1} C_k^T + R_k]^{-1} v_k && (81)\\ v_k &= [y_k - \overline{r}_k - C_k \widehat{x}_{k|k-1}] && (82)\\ \Sigma_{k|k} &= \Sigma_{k|k-1} \\ &\quad - \Sigma_{k|k-1} C_k^T [C_k \Sigma_{k|k-1} C_k^T + R_k]^{-1} C_k \Sigma_{k|k-1} && (83) \end{aligned}$$

where v_k denotes the prediction error.

Kalman Filter and Parameter Identification

The prediction equations are given by

$$\widehat{x}_{k|k-1} = F_{k-1}\widehat{x}_{k-1|k-1} + G_{k-1}u_{k-1} \tag{84}$$

$$\Sigma_{k|k-1} = F_{k-1}\Sigma_{k-1|k-1}F_{k-1}^T + G_{k-1}Q_kG_{k-1}^T \tag{85}$$

The algorithm is carried out in two distinct parts:

- Prediction Step
  The first step consists of forming an optimal predictor of the next observation, given all the information available up to time t_{k-1}. This results in a so-called *a priori estimate* for time t.
- Updating Step
  The a priori estimate is updated with the new information arriving at time t_k that is combined with the already available information from time t_{k-1}. The result of this step is the *filtered estimate* or the *a posteriori estimate*

Kalman Filter and Parameter Identification

The prediction error is defined by

$$v_k = y_k - y_{k|k-1} = y_k - E[y_k|\mathcal{F}_{k-1}] = y_k - \overline{r}_k - C_k\widehat{x}_{k|k-1}$$

and its covariance matrix can be calculated by

$$\begin{aligned}
K_{k|k-1} &= Cov(v_k|\mathcal{F}_{k-1}) \\
&= E[v_k v_k^T|\mathcal{F}_{k-1}] \\
&= C_k E[\widehat{x}_{k|k-1}\widehat{x}_{k|k-1}^T|\mathcal{F}_{k-1}]C_k^T + E[w_k^{(2)} {w_k^{(2)}}^T|\mathcal{F}_{k-1}] \\
&= C_k \Sigma_{k|k-1} C_k^T + R_k
\end{aligned}$$

This result is important for forming the log likelihood estimator.

Kalman Filter and Parameter Identification

We start to examine the appropriate *log-likelihood function* with our frame work of the Kalman Filter. Let $\psi \in \mathbb{R}^d$ be the vector of unknown parameters, which belongs to the admissible parameter space Ψ. The various system matrices of the state space models, as well as the variances of stochastic processes, given in the previous section, depend on ψ. The likelihood function of the state space model is given by the joint density of the observational data $y = (y_N, y_{N-1}, \dots, y_1)$

$$l(y, \psi) = p(y_N, y_{N-1}, \dots, y_1; \psi) \tag{86}$$

which reflects how likely it would have been to have observed the date if ψ were the true values of the parameters.

Using the definition of condition probability and employing *Bayes's theorem* recursively, we now write the joint density as product of conditional densities

$$l(y, \psi) = p(y_N | y_{N-1}, \ldots, y_1; \psi) \cdot \ldots \cdot p(y_k | y_{k-1}, \ldots, y_1; \psi) \cdot \ldots \cdot p(y_1; \psi) \quad (87)$$

where we approximate the initial density function $p(y_1; \psi)$ by $p(y_1 | y_0; \psi)$. Furthermore, since we deal with a Markovian system, future values of y_l with $l > k$ only depend on $(y_k, y_{k-1}, \ldots, y_1)$ through the current values of y_k, the expression in (87) can be rewritten to depend on only the last observed values, thus it is reduced to

$$l(y, \psi) = p(y_N | y_{N-1}; \psi) \cdot \ldots \cdot p(y_k | y_{k-1}; \psi) \cdot \ldots \cdot p(y_1; \psi) \quad (88)$$

For the purpose of estimating the parameter vector ψ we express the likelihood function in terms of the prediction error. Since the variance of the prediction error v_k is the same as the conditional variance of y_k, i.e.

$$Cov[v_k|\mathcal{F}_{k-1}] = Cov[y_k|\mathcal{F}_{k-1}] \tag{89}$$

we are able to state the density function. The density function $p(y_k|y_{k-1};\psi)$ is a Gauss Normal distribution with conditional mean

$$E[y_k|\mathcal{F}_{k-1}] = \overline{r}_k + C_k\widehat{x}_{k|k-1} \tag{90}$$

with conditional covariance matrix

$$Cov[y_k|\mathcal{F}_{k-1}] = K_{k|k-1}\,. \tag{91}$$

The density function of a n-dimensional Normal distribution can be written in matrix notation as

$$\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\Sigma|}}\, e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}$$

where μ denotes the mean value and Σ covariance matrix. In our case $x-\mu$ is replaced by $y_k - \overline{r}_k - C_k\widehat{x}_{k|k-1} = v_k$ and the covariance matrix is $K_{k|k-1}$. The probability density function thus takes the form

$$p(y_k|y_{k-1};\psi) = \frac{1}{(2\pi)^{\frac{p}{2}}\sqrt{|K_{k|k-1}|}}\, e^{-\frac{1}{2}v_k^T K_{k|k-1}^{-1} v_k}\,, \tag{92}$$

where $v_k \in \mathbb{R}^p$.

Kalman Filter and Parameter Identification

Taking the logarithm of (92) yields

$$\ln(p(y_k|y_{k-1};\psi)) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln|K_{k|k-1}| - \frac{1}{2}v_k^T K_{k|k-1}^{-1} v_k \tag{93}$$

which gives us the whole log-likelihood function as the sum

$$L(y,\psi) = -\frac{1}{2}\sum_{k=1}^{N} n\ln(2\pi) + \ln|K_{k|k-1}| + v_k^T K_{k|k-1}^{-1} v_k \tag{94}$$

In order to estimate the unknown parameters from the log-likelihood function in (94) we employ an appropriate optimization routine which maximizes $L(y,\psi)$ with respect to ψ.

Kalman Filter and Parameter Identification

iven the simplicity of the innovation form of the likelihood function, we can find the values for the conditional expectation of the prediction error and the conditional covariance matrix for every given parameter vector ψ using the Kalman Filter algorithm. Hence, we are able to calculate the value of the conditional log likelihood function numerically. The parameter are estimated based on the optimization

$$\widehat{\psi}_{ML} = \arg\max_{\psi \in \Psi} L(y, \psi) \tag{95}$$

where Ψ denotes the parameter space. The optimization is either an unconstrained optimization problem if $\psi \in \mathbb{R}^d$ or a constrained optimization if the parameter space $\Psi \subset \mathbb{R}^d$.

Kalman Filter and Parameter Identification