A Novel and Simplest Derivation of Measurement Update Equations in the Kalman Filter (Version 3)

Sueo Sugimoto
Ritsumeikan University

April 20, 2014, revised: October 3, 2017, April 29, 2019

Email: sugimoto@se.ritsumei.ac.jp

ABSTRACT

We show the simplest derivation of the Kalman filter, especially, derive the so-called measurement update equation, base on the several easy mathematical concepts such as, conditional expectation, Gaussian conditional probability density function, completing the square, and the matrix inversion lemma.

1 Kalman filter - introduction

The Kalman filter [1]-[5], is the minimum mean square error (minimum error covariance) filter based on the measurement $Y^t := \{y_0, y_1, \cdots, y_t\}$ to estimate the state x_t as follows:

$$\begin{aligned} x_{t+1} &= F_t x_t + G_t w_t, \quad t = 0, 1, \cdots \quad \text{(State equation)}, \quad x_t : (n \times 1), & (1)\\ & \mathrm{E}[\mathrm{x}_0] = \hat{\mathrm{x}}_0 = \hat{\mathrm{x}}_{0|-1}, \quad \mathrm{Var}[\mathrm{x}_0] = \Sigma_0 = \Sigma_{0|-1}, & (2)\\ y_t &= H_t x_t + v_t, \quad t = 0, 1, \cdots \quad \text{(Measurement equation)}, \quad y_t : (m \times 1), & (3)\end{aligned}$$

where w_t and v_t are mutually independent Gaussian white noise such as

$$\begin{aligned} \mathrm{E}[w_t] &= 0, \quad \mathrm{E}[w_t w_\tau] = Q_t \delta_{t-\tau}, & (4)\\ \mathrm{E}[v_t] &= 0, \quad \mathrm{E}[v_t v_\tau] = R_t \delta_{t-\tau}. \quad \text{(where } \delta_t \text{ is the Kronecker's } \delta - \text{function)}. & (5)\end{aligned}$$

It is well known the minimum error variance estimate of x_t based on the measurement Y^t is:

$$\arg\min_{\hat{x}_t} \mathrm{E}\Big[\ ||x_t - \hat{x}_t||^2\ \Big| Y^t\Big] \quad = \quad \mathrm{E}\Big[x_t\ \Big|\ Y^t\Big]. \qquad (6)$$

Namely the minimum variance estimate is the conditional expectation of x_t based on the given mesurements Y^t.

2 CPDF and completing the square

For obtaining the conditiona expectation of x_t, the evaluation of the conditional probability density function (CPDF) of x_t is most important.
Then let us consider the following relations of CPDF:

$$p(x_t|Y^t) \quad = \quad \frac{p(x_t, Y^t)}{p(Y^t)}$$

$$
\begin{aligned}
&= \frac{p(x_t, y_t|Y^{t-1})p(Y^{t-1})}{p(Y^t)} \\
&= \frac{p(y_t|x_t, Y^{t-1})p(x_t|Y^{t-1})p(Y^{t-1})}{p(Y^t)} \\
&= \frac{p(y_t|x_t)p(x_t|Y^{t-1})p(Y^{t-1})}{p(Y^t)} \\
&:= K_0(Y^t)p(y_t|x_t)p(x_t|Y^{t-1}), \qquad (7)
\end{aligned}
$$

where

$$
K_0(Y^t) := \frac{p(Y^{t-1})}{p(Y^t)}.
$$

Then, due to the Gaussian property, we will evaluate $p(y_t|x_t)p(x_t|Y^{t-1})$ as the following formulas:

$$
\begin{aligned}
p(x_t|Y^{t-1})p(y_t|x_t) &= \frac{1}{(2\pi)^{n/2}|\Sigma_{t|t-1}|^{1/2}} \exp\Big\{-\frac{1}{2}[x_t - \hat{x}_{t|t-1}]^{\mathrm{T}}\Sigma_{t|t-1}^{-1}[x - \hat{x}_{t|t-1}]\Big\} \\
&\quad \times\frac{1}{(2\pi)^{m/2}|R_t|^{1/2}} \exp\Big\{-\frac{1}{2}[y_t - H_t x_t]^{\mathrm{T}}R_t^{-1}[y_t - Hx_t]\Big\}. \qquad (8)
\end{aligned}
$$

Therefore the power term in (8) can be expressed by

$$
\begin{aligned}
J(x_t) &:= -\frac{1}{2}\Big\{[x_t - \hat{x}_{t|t-1}]^{\mathrm{T}}\Sigma_{t|t-1}^{-1}[x_t - \hat{x}_{t|t-1}] + [y_t - H_t x_t]^{\mathrm{T}}R_t^{-1}[y_t - H_t x_t]\Big\} \\
&= -\frac{1}{2}\Big\{x^{\mathrm{T}}\Sigma^{-1}x - x^{\mathrm{T}}\Sigma^{-1}\hat{x} - \hat{x}^{\mathrm{T}}\Sigma^{-1}x + \hat{x}^{\mathrm{T}}\Sigma^{-1}\hat{x} \\
&\qquad + y^{\mathrm{T}}R^{-1}y - y^{\mathrm{T}}R^{-1}Hx - x^{\mathrm{T}}H^{\mathrm{T}}R^{-1}y + x^{\mathrm{T}}H^{\mathrm{T}}R^{-1}Hx\Big\} \\
&= -\frac{1}{2}\Big\{x^{\mathrm{T}}(\Sigma^{-1} + H^{\mathrm{T}}R^{-1}H)x - x^{\mathrm{T}}(\Sigma^{-1}\hat{x} + H^{\mathrm{T}}R^{-1}y) \\
&\qquad - (\hat{x}^{\mathrm{T}}\Sigma^{-1} + y^{\mathrm{T}}R^{-1}H)x + \hat{x}^{\mathrm{T}}\Sigma^{-1}\hat{x} + y^{\mathrm{T}}R^{-1}y\Big\} \\
&:= -\frac{1}{2}\Big\{\Big[x - (\Sigma^{-1} + H^{\mathrm{T}}R^{-1}H)^{-1}(\Sigma^{-1}\hat{x} + H^{\mathrm{T}}R^{-1}y)\Big]^{\mathrm{T}}(\Sigma^{-1} + H^{\mathrm{T}}R^{-1}H) \\
&\qquad \times\Big[x - (\Sigma^{-1} + H^{\mathrm{T}}R^{-1}H)^{-1}(\Sigma^{-1}\hat{x} + H^{\mathrm{T}}R^{-1}y)\Big] + J_0(\hat{x}, \Sigma, y)\Big\}, \qquad (9)
\end{aligned}
$$

where

$$
\begin{aligned}
J_0(\hat{x}, \Sigma, y) &:= -(\Sigma^{-1}\hat{x} + H^{\mathrm{T}}R^{-1}y)^{\mathrm{T}}(\Sigma^{-1} + H^{\mathrm{T}}R^{-1}H)^{-1}(\Sigma^{-1}\hat{x} + H^{\mathrm{T}}R^{-1}y) \\
&\qquad + \hat{x}^{\mathrm{T}}\Sigma^{-1}\hat{x} + y^{\mathrm{T}}R^{-1}y. \qquad (10)
\end{aligned}
$$

Therefore we have

$$
\begin{aligned}
p(x_t|Y^t) &= K_0(Y^t)p(y_t|x_t)p(x_t|Y^{t-1}) \\
&= K_0(Y^t)\frac{1}{(2\pi)^{n/2+m/2}|\Sigma_{t|t-1}|^{1/2}|R_t|^{1/2}} \exp\{J(x_t)\} \\
&= K(Y^t, \hat{x}_{t|t-1}, \Sigma_{t|t-1}) \exp\{-\frac{1}{2}(x_t - \hat{x}_{t|t})^{\mathrm{T}}\Sigma_{t|t}^{-1}(x_t - \hat{x}_{t|t})\}, \qquad (11)
\end{aligned}
$$

where

$$
\begin{aligned}
\hat{x}_{t|t} &:= (\Sigma_{t|t-1}^{-1} + H_t^{\mathrm{T}} R_t^{-1} H_t)^{-1} (\Sigma_{t|t-1}^{-1} \hat{x}_{t|t-1} + H_t^{\mathrm{T}} R_t^{-1} y_t), & (12)\\
\Sigma_{t|t} &:= (\Sigma_{t|t-1}^{-1} + H_t^{\mathrm{T}} R_t^{-1} H_t)^{-1}, & (13)\\
K(Y^t, \hat{x}_{t|t-1}, \Sigma_{t|t-1}) &:= K_0(Y^t) \frac{1}{(2\pi)^{n/2+m/2} |\Sigma_{t|t-1}|^{1/2} |R_t|^{1/2}} \exp\{-\frac{1}{2} J_0(\hat{x}_{t|t-1}, \Sigma_{t|t-1}, y_t)\}.
\end{aligned}
$$

Upto now, we only apply the mathematical technique of completing the square to the power term in (8) so that we have the relations of (12) and (13). Then we will show the popular measurement update equation of the Kalman filter from the relation (12) and (13).

3 The measurement update equation

Applying the matrix inversion lemma:

$$
(A^{-1} + B^{\mathrm{T}} C B)^{-1} = A - A B^{\mathrm{T}} (B A B^{\mathrm{T}} + C^{-1})^{-1} B A,
$$

to (13), we have

$$
(\Sigma_{t|t-1}^{-1} + H_t^{\mathrm{T}} R_t^{-1} H_t)^{-1} = \Sigma_{t|t-1} - \Sigma_{t|t-1} H_t^{\mathrm{T}} (H_t \Sigma_{t|t-1} H_t^{\mathrm{T}} + R_t)^{-1} H_t \Sigma_{t|t-1}. \quad (14)
$$

Therefore, from (12) and (13) we have the well-known measurement updating equations in the Kalman filter:

$$
\begin{aligned}
\Sigma_{t|t} &= (\Sigma_{t|t-1}^{-1} + H_t^{\mathrm{T}} R_t^{-1} H_t)^{-1}\\
&= \Sigma_{t|t-1} - \Sigma_{t|t-1} H_t^{\mathrm{T}} (H_t \Sigma_{t|t-1} H_t^{\mathrm{T}} + R_t)^{-1} H_t \Sigma_{t|t-1}, & (15)
\end{aligned}
$$

and

$$
\begin{aligned}
\hat{x}_{t|t} &= (\Sigma_{t|t-1}^{-1} + H_t^{\mathrm{T}} R_t^{-1} H_t)^{-1} (\Sigma_{t|t-1}^{-1} \hat{x}_{t|t-1} + H_t^{\mathrm{T}} R_t^{-1} y_t)\\
&= \Big[\Sigma_{t|t-1} - \Sigma_{t|t-1} H_t^{\mathrm{T}} (H_t \Sigma_{t|t-1} H_t^{\mathrm{T}} + R_t)^{-1} H_t \Sigma_{t|t-1}\Big] (\Sigma_{t|t-1}^{-1} \hat{x}_{t|t-1} + H_t^{\mathrm{T}} R_t^{-1} y_t)\\
&= \hat{x}_{t|t-1} + \Sigma_{t|t-1} H_t^{\mathrm{T}} R_t^{-1} y_t - \Sigma_{t|t-1} H_t^{\mathrm{T}} (H_t \Sigma_{t|t-1} H_t^{\mathrm{T}} + R_t)^{-1} H_t \hat{x}_{t|t-1}\\
&\quad - \Sigma_{t|t-1} H_t^{\mathrm{T}} (H_t \Sigma_{t|t-1} H_t^{\mathrm{T}} + R_t)^{-1} H_t \Sigma_{t|t-1} H_t^{\mathrm{T}} R_t^{-1} y_t\\
&= \quad \hat{x}_{t|t-1} + \Sigma_{t|t-1} H_t^{\mathrm{T}} (H_t \Sigma_{t|t-1} H_t^{\mathrm{T}} + R_t)^{-1}\\
&\qquad \times \Big[(H_t \Sigma_{t|t-1} H_t^{\mathrm{T}} + R_t) R_t^{-1} y_t - H_t \hat{x}_{t|t-1} - H_t \Sigma_{t|t-1}^{-1} H_t^{\mathrm{T}} R_t^{-1} y_t\Big]\\
&= \hat{x}_{t|t-1} + \Sigma_{t|t-1} H_t^{\mathrm{T}} (H_t \Sigma_{t|t-1} H_t^{\mathrm{T}} + R_t)^{-1} (y_t - H_t \hat{x}_{t|t-1})\\
&:= \hat{x}_{t|t-1} + K_t (y_t - H_t \hat{x}_{t|t-1}) & (16)\\
&\quad K_t := \Sigma_{t|t-1} H_t^{\mathrm{T}} (H_t \Sigma_{t|t-1} H_t^{\mathrm{T}} + R_t)^{-1}, \quad \text{(Kalman gain)} & (17)\\
&\quad \nu_t := y_t - H_t \hat{x}_{t|t-1} \quad \text{(Innovation process of } y_t\text{)}. & (18)
\end{aligned}
$$

The time-update equation

The time-update equation is easily obtained from the state equation in (1):

$$
x_{t+1} = F_t x_t + G_t w_t, \quad (19)
$$

where $\{w_t\},\ t = 0, \cdots$ are Gaussin white noises in (4),

Taking the conditional expectation to both sides in (1) , we have

$$\begin{aligned} \hat{x}_{t+1|t} := \mathrm{E}[x_{t+1}|Y^t] &= \mathrm{E}[F_t x_t + G_t w_t|Y^t] \\ &= F_t \mathrm{E}[x_t|Y^t], \end{aligned} \tag{20}$$

where w_t is an independent (uncorrelated) random variable to Y^t such that $\mathrm{E}[w_t|Y^t] = \mathrm{E}[w_t] = 0$. Therefore

$$\hat{x}_{t+1|t} = F_t \hat{x}_{t|t}. \tag{21}$$

Also from (1) and (21)

$$x_{t+1} - \hat{x}_{t+1|t} = F_t(x_t - \hat{x}_{t|t}) + G_t w_t, \tag{22}$$

we have

$$\begin{aligned} \Sigma_{t+1|t} &:= \mathrm{E}\Big[(x_{t+1} - \hat{x}_{t+1|t})(x_{t+1} - \hat{x}_{t+1|t})^{\mathrm{T}}\Big] \\ &= \mathrm{E}\Big[(F_t(x_t - \hat{x}_{t|t}) + G_t w_t)(F_t(x_t - \hat{x}_{t|t}) + G_t w_t)^{\mathrm{T}}\Big] \\ &= F_t \Sigma_{t|t} F_t^{\mathrm{T}} + G_t Q_t G_t^{\mathrm{T}}. \end{aligned} \tag{23}$$

Finally, we have the Kalman Filter, described by Eqs. (21), (16), (23) and (15) with the intial conditions in Eq. (2). Namely,

- Filtering and prediction estimates:

$$\hat{x}_{t+1|t} = F_t \hat{x}_{t|t}, \tag{24}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t[y_t - H_t \hat{x}_{t|t-1}], \quad t = 0, 1, \cdots \tag{25}$$

- Kalman gain:

$$K_t = \Sigma_{t|t-1} H_t^{\mathrm{T}} (H_t \Sigma_{t|t-1} H_t^{\mathrm{T}} + R_t)^{-1}, \quad t = 0, 1, \cdots \tag{26}$$

- Estimated error covariances:

$$\Sigma_{t+1|t} = F_t \Sigma_{t|t} F_t^{\mathrm{T}} + G_t Q_t G_t^{\mathrm{T}} \tag{27}$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t H_t \Sigma_{t|t-1}, \quad t = 0, 1, \cdots \tag{28}$$

- Initial coditions:

$$\mathrm{E}[\mathrm{x}_0] = \hat{x}_0 = \hat{x}_{0|-1}, \tag{29}$$

$$\mathrm{Var}[\mathrm{x}_0] = \Sigma_0 = \Sigma_{0|-1}. \tag{30}$$

References

[1] R. E. Kalman: A new approach to linear filtering and prediction problems, *Trans. ASME, J. Basic Eng.*, Vol. 82D, No. 1, pp. 34-45, 1960.

[2] Y. C. Ho and R. Lee: A Bayesian approach to problems in stochastic estimation and control, *IEEE Trans. on Automatic Control*, Vol. 9, No. 4, pp. 333-339, 1964.

[3] B. D. O. Anderson and J. B. Moore: *Optimal Filtering*, Prentice-Hall, 1979.

[4] T. Katayama: *Applied Kalman Filtering* (in Japanese), Asakura-Shoten, Tokyo, 1983.

[5] S. Sugimoto, M. Murata and K. Ohnishi(Co-Eds.): *Nonlinear Filters*, Ohm-Sya, Tokyo, (will appear in July 2019).