A Novel and Simplest derivation of Measurement Update Equations in the Kalman Filter (Version 2)

Sueo Sugimoto
Ritsumeikan University
April 20, 2014, revised: October 20, 2017
Email: sugimoto@se.ritsumei.ac.jp

ABSTRACT

We show the simplest derivation of the Kalman filter, especially, derive the so-called measurement updating equations, 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 Kalamn Filter - Introduction

The Kalman filter [1], [2], is the minimum mean square error (minimum error covariance) filter based on the measurement $Y^t \equiv \{y_0, y_1, \cdots, y_t\}$ to estimate the state x_t as follows:

$$x_{t+1} = F_t x_t + w_t \quad \text{(State equation)}, \quad x_t : \ (n \times 1) \text{ vector} \tag{1}$$

$$\mathrm{E}[\mathrm{x}_0] = \hat{\mathrm{x}}_0 = \hat{\mathrm{x}}_{0|-1}, \quad \mathrm{Var}[\mathrm{x}_0] = \mathrm{P}_0 = \mathrm{P}_{0|-1}, \tag{2}$$

$$y_t = H_t x_t + v_t \quad \text{(Measurement equation)}, \quad y_t : \ (m \times 1) \text{ vector} \tag{3}$$

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

$$\mathrm{E}[w_t] = 0, \quad \mathrm{E}[w_t w_\tau] = Q_t \delta_{t-\tau}, \tag{4}$$

$$\mathrm{E}[v_t] = 0, \quad \mathrm{E}[v_t v_\tau] = R_t \delta_{t-\tau}. \quad \text{(where } \delta_t \text{ is the Kronecker's } \delta - \text{function)} \tag{5}$$

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

$$\arg\min_{\hat{x}_t} \mathrm{E}\Big[\ ||x_t - \hat{x}||^2 \ \Big| Y^t \Big] = \mathrm{E}\Big[x_t \ \Big| \ Y^t \Big] \tag{6}$$

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:

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

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

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

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

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

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

where

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

Therefore we have

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

where

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

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

3 Measurement updating formula

Applying the matrix inversion lemma:

$$(A^{-1} + B^{\mathrm{T}}CB)^{-1} \quad = \quad A - AB^{\mathrm{T}}(BAB^{\mathrm{T}} + C^{-1})^{-1}BA$$

to (12), we have

$$(P^{-1} + H^{\mathrm{T}}R_t^{-1}H)^{-1} \quad = \quad P - PH^{\mathrm{T}}(HPH^{\mathrm{T}} + R)^{-1}HP \tag{14}$$

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

$$\begin{aligned} P_{t|t} &\equiv (P_{t|t-1}^{-1} + H^{\mathrm{T}}R_t^{-1}H)^{-1} \\ &\equiv P_{t|t-1} - P_{t|t-1}H^{\mathrm{T}}(HP_{t|t-1}H^{\mathrm{T}} + R_t)^{-1}HP_{t|t-1}, \end{aligned} \tag{15}$$

and

$$\begin{aligned} \hat{x}_{t|t} &= (P_{t|t-1}^{-1} + H^{\mathrm{T}}R_t^{-1}H)^{-1}(P_{t|t-1}^{-1}\hat{x}_{t|t-1} + H^{\mathrm{T}}R_t^{-1}y_t) \\ &= \Big[P_{t|t-1} - P_{t|t-1}H^{\mathrm{T}}(HP_{t|t-1}H^{\mathrm{T}} + R_t)^{-1}HP_{t|t-1}\Big](P_{t|t-1}^{-1}\hat{x}_{t|t-1} + H^{\mathrm{T}}R_t^{-1}y_t) \\ &= \hat{x}_{t|t-1} + P_{t|t-1}H^{\mathrm{T}}R_t^{-1}y_t - P_{t|t-1}H^{\mathrm{T}}(HP_{t|t-1}H^{\mathrm{T}} + R_t)^{-1}H\hat{x}_{t|t-1} \\ &\quad -P_{t|t-1}H^{\mathrm{T}}(HP_{t|t-1}H^{\mathrm{T}} + R_t)^{-1}HP_{t|t-1}H^{\mathrm{T}}R_t^{-1}y_t \\ &= \quad \hat{x}_{t|t-1} + P_{t|t-1}H^{\mathrm{T}}(HP_{t|t-1}H^{\mathrm{T}} + R_t)^{-1} \\ &\qquad \times\Big[(HP_{t|t-1}H^{\mathrm{T}} + R_t)R_t^{-1}y_t - H\hat{x}_{t|t-1} - HP_{t|t-1}^{-1}H^{\mathrm{T}}R_t^{-1}y_t\Big] \\ &= \hat{x}_{t|t-1} + P_{t|t-1}H^{\mathrm{T}}(HP_{t|t-1}H^{\mathrm{T}} + R_t)^{-1}(y_t - H\hat{x}_{t|t-1}). \end{aligned} \tag{16}$$

Time Updating Formula

The time updating formula are easily obtained by the state equation in [?]:

$$x_{t+1} = F_t x_t + w_t \tag{17}$$

where $\{w_t\}, t = 0, \cdot$ are Gaussin white noises

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

$$\begin{aligned} \mathrm{E}[x_{t+1}|Y^t] &= \mathrm{E}[F_t x_t + w_t|Y^t] \\ &= F_t\mathrm{E}[x_t|Y^t] \end{aligned} \tag{18}$$

therefore

$$\hat{x}_{t+1|t} \quad = \quad F_t\hat{x}_{t|t} \tag{19}$$

Also from (1) and (19)

$$x_{t+1} - \hat{x}_{t+1|t} = F_t(x_t - \hat{x}_{t|t}) + w_t \tag{20}$$

Thus

$$\begin{aligned} \mathrm{E}\Big[(x_{t+1} - \hat{x}_{t+1|t})(x_{t+1} - \hat{x}_{t+1|t})^{\mathrm{T}}\Big] &= \mathrm{E}\Big[(F_t(x_t - \hat{x}_{t|t}) + w_t)(22)(F_t(x_t - \hat{x}_{t|t}) + w_t)^{\mathrm{T}}\Big] \\ P_{t+1|t} &= F_t P_{t|t} F_t^{\mathrm{T}} + Q \qquad (21) \end{aligned}$$

Finally, we have the Kalman Filter, described by Eqs. (16), (!7), (20) and (22) with the intial conditions in Eq. (2).

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.