Derivation of the discrete-time Kalman filter

Rodrigo Ventura
Instituto Superior Técnico

Working draft (updated at September 2017)

1 Introduction

The Kalman filter belongs to a family of stochastic filtering algorithms. The purpose of this document is to provide a simple, yet complete, derivation of the time-discrete version of this filter. It was strongly inspired by these sources:

- Sebastian Thrun, Wolfram Burgard, Dieter Fox. *Probabilistic Robotics.* MIT Press, 2005.
- Maria Isabel Ribeiro. *Kalman and Extended Kalman Filters: Concept, Derivation and Properties.* Instituto de Sistemas e Robótica, Instituto Superior Técnico, 2004.

The Kalman filter aims at estimating the hidden state of a linear time-variant system, given its control input, and a (possibly) partial and noisy observation of the state. First, a formal model of such system, for discrete time, is presented in section 2. Then, the Kalman filter for this system is derived in section 3. This filter can no longer be applied to nonlinear systems. However, one common approach is to linearize the system at each time step, and then apply the Kalman filter to this linearized version. This corresponds to the extended Kalman filter (EKF), which is derived in section 4.

2 System model

Consider the following discrete time linear system:

$$\begin{aligned} x_t &= A_t x_{t-1} + B_t u_t + \varepsilon_t \\ z_t &= C_t x_t + \delta_t \end{aligned} \tag{1}$$

where $x_t \in \mathbb{R}^n$ is the state vector[1], $u_t \in \mathbb{R}^m$ is the control input, and $z_t \in \mathbb{R}^k$ is the measurement. The sources of uncertainty are the state transition noise $\varepsilon_t \in \mathbb{R}^n$, and the measurement noise $\delta_t \in \mathbb{R}^k$. Both of these noise signals are assumed to be normally distributed with zero mean, $\varepsilon_t \sim \mathcal{N}(0, R_t)$ and $\delta_t \sim \mathcal{N}(0, Q_t)$, and both uncorrelated with the state vector. The initial state vector is also assumed to be normal:

$$x_0 \sim \mathcal{N}(\mu_0, \Sigma_0) \tag{2}$$

Under these circumstances, the state vector is always normal distributed, meaning that it suffices to compute its mean and covariance matrix to have a complete description of the state distribution.

3 Kalman filter

The Kalman filter is recursive, meaning that the state distribution of the state at time t is computed from the one at time $t-1$. Consider that we know the sequences of all measurements and all control inputs up to time t. The self-evident notation $z_{1:t} = [z_1, z_2, \dots, z_t]$ and $u_{1:t} = [u_1, u_2, \dots, u_t]$ will be used to denote these sequences. The desired (posterior) state distribution at time t, called *belief*, is defined by

$$\mathrm{bel}(x_t) = p(x_t | z_{1:t}, u_{1:t}) \sim \mathcal{N}(\mu_t, \Sigma_t) \tag{3}$$

where μ_t and Σ_t are the parameters of the normal distribution. To compute this distribution, from the one at time $t-1$

$$\mathrm{bel}(x_{t-1}) = p(x_{t-1} | z_{1:t-1}, u_{1:t-1}) \sim \mathcal{N}(\mu_{t-1}, \Sigma_{t-1}) \tag{4}$$

the Kalman filter proceeds in two steps: a (1) *prediction step*, which uses the control input u_t alone to compute an intermediate belief state

$$\overline{\mathrm{bel}}(x_t) = p(x_t | z_{1:t-1}, u_{1:t}) \sim \mathcal{N}(\overline{\mu}_t, \overline{\Sigma}_t) \tag{5}$$

and a (2) *update step* which uses the measurement z_t alone to improve this belief into the desired $\mathrm{bel}(x_t)$. Each one of these steps will be derived below.

3.1 Prediction step

The goal of this step is to compute $\overline{\mathrm{bel}}(x_t)$ from $\mathrm{bel}(x_{t-1})$ and u_t. The state vector is always normal distributed, meaning that it suffices to compute its mean $\overline{\mu}_t$ and covariance $\overline{\Sigma}_t$.

[1] Time dependency on t will be denoted with subscript index $_t$.

The mean follows directly from the system model (1)

$$\begin{aligned}
\overline{\mu}_t &= E[x_t|z_{1:t-1}, u_{1:t}] \\
&= A_t E[x_{t-1}|z_{1:t-1}, u_{1:t}] + B_t u_t + 0 \\
&= A_t \underbrace{E[x_{t-1}|z_{1:t-1}, u_{1:t-1}]}_{\mu_{t-1}} + B_t u_t + 0 \\
&= A_t \mu_{t-1} + B_t u_t
\end{aligned} \tag{6}$$

taking into account that the distribution of x_{t-1} is independent of the future control input u_t. The same line of reasoning applies to the covariance

$$\overline{\Sigma}_t = E[(x_t - \overline{\mu}_t)(x_t - \overline{\mu}_t)^T|z_{1:t-1}, u_{1:t}] \tag{7}$$

Expanding one term of the expression inside the expectation operator, using (1) and (6), we can get

$$\begin{aligned}
x_t - \overline{\mu}_t &= (A_t x_{t-1} + B_t u_t + \varepsilon_t - A_t \mu_{t-1} - B_t u_t) \\
&= A_t(x_{t-1} - \mu_{t-1}) + \varepsilon_t
\end{aligned} \tag{8}$$

Multiplying this expression by its transpose, we obtain

$$\begin{aligned}
(x_t - \overline{\mu}_t)(x_t - \overline{\mu}_t)^T =& A_t(x_{t-1} - \mu_{t-1})(x_{t-1} - \mu_{t-1})^T A_t^T + \\
& A_t(x_{t-1} - \mu_{t-1})\varepsilon_t^T + \\
& \varepsilon_t(x_{t-1} - \mu_{t-1})^T A_t^T + \\
& \varepsilon_t \varepsilon_t^T
\end{aligned} \tag{9}$$

The application of the expectation operator (7) results in

$$\begin{aligned}
\overline{\Sigma}_t &= A_t E[(x_{t-1} - \mu_{t-1})(x_{t-1} - \mu_{t-1})^T|z_{1:t-1}, u_{1:t}]A_t^T + 0 + 0 + R_t \\
&= A_t \underbrace{E[(x_{t-1} - \mu_{t-1})(x_{t-1} - \mu_{t-1})^T|z_{1:t-1}, u_{1:t-1}]}_{\Sigma_{t-1}} A_t^T + R_t \\
&= A_t \Sigma_{t-1} A_t^T + R_t
\end{aligned} \tag{10}$$

noting again that x_{t-1} is independent of the (future) control input u_t.

In summary, the expressions for $\overline{\mu}_t$ and $\overline{\Sigma}_t$, which completely describe the belief $\overline{\text{bel}}(x_t)$, are given by

$$\begin{aligned}
\overline{\mu}_t &= A_t \mu_{t-1} + B_t u_t \\
\overline{\Sigma}_t &= A_t \Sigma_{t-1} A_t^T + R_t
\end{aligned} \tag{11}$$

3.2 Update step

Having the intermediate belief $\overline{\text{bel}}(x_t)$, we want now to update it with the new measurement z_t. To do so we will use the Bayes rule

$$\begin{aligned}\text{bel}(x_t) &= p(x_t|z_{1:t}, u_{1:t}) \\ &= \eta\, p(z_t|x_t, z_{1:t-1}, u_{1:t}) p(x_t|z_{1:t-1}, u_{1:t}) \\ &= \eta\, p(z_t|x_t)\overline{\text{bel}}(x_t)\end{aligned} \tag{12}$$

using the assumption that the measurement z_t is conditionally independent of past measurements and of control, given the knowledge of the state x_t. The term η is a normalizing constant that does not depent on x_t. Given the measurement equation of (1), and knowing the value of x_t, the only source of randomness is δ_t. Thus, it is clear that z_t is normally distributed, with mean $C_t x_t$ (since δ_t has zero mean, and x_t is fixed) and with covariance

$$\begin{aligned}\text{Cov}[z_t|x_t] &= E[(z_t - C_t x_t)(z_t - C_t x_t)^T|x_t] \\ &= E[\delta_t \delta_t^T|x_t] = Q_t\end{aligned} \tag{13}$$

Thus, the distribution of z_t given x_t can be written

$$p(z_t|x_t) = |2\pi Q_t|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(z_t - C_t x_t\right)^T Q_t^{-1}\left(z_t - C_t x_t\right)\right] \tag{14}$$

Using (12) we can then write

$$\begin{aligned}\text{bel}(x_t) &= \eta\, p(z_t|x_t)\overline{\text{bel}}(x_t) \\ &= \eta' \exp\left[-\frac{1}{2}\left(z_t - C_t x_t\right)^T Q_t^{-1}\left(z_t - C_t x_t\right)\right] \\ &\qquad \exp\left[-\frac{1}{2}\left(x_t - \overline{\mu}_t\right)^T \overline{\Sigma}_t^{-1}\left(x_t - \overline{\mu}_t\right)\right] \\ &= \eta' \exp\left[-\frac{1}{2}\left[\left(z_t - C_t x_t\right)^T Q_t^{-1}\left(z_t - C_t x_t\right) + \right.\right. \\ &\qquad \left.\left. +\ \left(x_t - \overline{\mu}_t\right)^T \overline{\Sigma}_t^{-1}\left(x_t - \overline{\mu}_t\right)\right]\right]\end{aligned} \tag{15}$$

where both η and η' are normalizing constants. For the following derivations, we will drop the t indices for the sake of clarity. Inside the exponential there is a sum of two quadratic forms, which can be manipulated into a single quadratic form by

$$\begin{aligned}&(z - Cx)^T Q^{-1}(z - Cx) + (x - \overline{\mu})^T \overline{\Sigma}^{-1}(x - \overline{\mu}) = \\ &= x^T(C^T Q^{-1} C + \overline{\Sigma}^{-1})x - 2x^T(C^T Q^{-1} z + \overline{\Sigma}^{-1}\overline{\mu}) + \overline{\mu}^T \overline{\Sigma}^{-1}\overline{\mu} + z^T Q^{-1} z\end{aligned} \tag{16}$$

If $\mathrm{bel}(x_t)$ has a normal distribution, its quadratic form inside the exponential should match the above one. Such quadratic form can be expanded in a way to match the above one:

$$(x-\mu)^T\Sigma^{-1}(x-\mu) = x^T\Sigma^{-1}x - 2x^T\Sigma^{-1}\mu + \mu^T\Sigma^{-1}\mu \tag{17}$$

noting that all these variables are referred to time t, as we are omitting these indices. Comparing (16) with (17), we can immediately derive an expression for the updated covariance

$$\Sigma^{-1} = C^TQ^{-1}C + \overline{\Sigma}^{-1} \tag{18}$$

Concerning now the updated mean, we can use the crossed multiplication, along with (18)

$$\begin{aligned} C^TQ^{-1}z + \overline{\Sigma}^{-1}\overline{\mu} &= \Sigma^{-1}\mu \\ &= C^TQ^{-1}C\mu + \overline{\Sigma}^{-1}\mu \end{aligned} \tag{19}$$

From reordering the terms it follows that

$$\begin{aligned} & C^TQ^{-1}(z - C\mu) = \overline{\Sigma}^{-1}(\mu - \overline{\mu}) \\ \Leftrightarrow\ & C^TQ^{-1}(z - C\mu + C\overline{\mu} - C\overline{\mu}) = \overline{\Sigma}^{-1}(\mu - \overline{\mu}) \\ \Leftrightarrow\ & C^TQ^{-1}(z - C\overline{\mu}) = (\underbrace{C^TQ^{-1}C + \overline{\Sigma}^{-1}}_{\Sigma^{-1}})(\mu - \overline{\mu}) \\ \Leftrightarrow\ & \mu = \overline{\mu} + \Sigma C^TQ^{-1}(z - C\overline{\mu}) \end{aligned} \tag{20}$$

The remaining terms in (16) and (17) shall not raise any concern, since they do not depend on x, and thus any discrepancy can be factored out of the exponential, and captured by the normalizing constant η' in (15).

One important concept in Kalman filtering is the Kalman gain, corresponding to the matrix $K = \Sigma C^TQ^{-1}$ in (20), multiplying the difference between the observed measurement and the predicted measurement $C\overline{\mu}$. This gain represents how much the predicted state mean $\overline{\mu}$ is updated with the measurement z_t. Moreover, this Kalman gain allows us to get very short expressions for the belief mean and covariance: the mean becomes simply

$$\mu = \overline{\mu} + K(z - C\overline{\mu}) \tag{21}$$

and the covariance requires some additional manipulation

$$\begin{aligned} & K = \Sigma C^TQ^{-1} \\ \Leftrightarrow\ & KC = \Sigma\underbrace{C^TQ^{-1}C}_{=\Sigma^{-1}-\overline{\Sigma}^{-1}} \\ \Leftrightarrow\ & \Sigma = (I - KC)\overline{\Sigma} \end{aligned} \tag{22}$$

The Kalman gain can be rewritten as follows, so that it does not depend on Σ anymore:

$$
\begin{aligned}
K &= \Sigma C^T Q^{-1} \\
&= \Sigma C^T Q^{-1}(C\overline{\Sigma}C^T + Q)(C\overline{\Sigma}C^T + Q)^{-1} \\
&= \Sigma(C^T Q^{-1} C\overline{\Sigma}C^T + C^T)(C\overline{\Sigma}C^T + Q)^{-1} \\
&= \Sigma(C^T Q^{-1} C\overline{\Sigma}C^T + \overline{\Sigma}^{-1}\overline{\Sigma}C^T)(C\overline{\Sigma}C^T + Q)^{-1} \\
&= \Sigma\underbrace{(C^T Q^{-1} C + \overline{\Sigma}^{-1})}_{\Sigma^{-1}}\overline{\Sigma}C^T(C\overline{\Sigma}C^T + Q)^{-1} \\
&= \overline{\Sigma}C^T(C\overline{\Sigma}C^T + Q)^{-1}
\end{aligned}
\tag{23}
$$

After this algebraic *tour de force* we arrive to the update expressions for the mean and covariance of the new updated belief state $\mathrm{bel}(x_t) \sim \mathcal{N}(\mu_t, \Sigma_t)$, after restoring the time indices t

$$
\begin{aligned}
K_t &= \overline{\Sigma}_t C_t^T (C_t \overline{\Sigma}_t C_t^T + Q_t)^{-1} \\
\mu_t &= \overline{\mu}_t + K_t(z_t - C_t\overline{\mu}_t) \\
\Sigma_t &= (I - K_t C_t)\overline{\Sigma}_t
\end{aligned}
\tag{24}
$$

Note also that this step only requires a single matrix inversion. This concludes the derivation of the Kalman filter.

4 Extended Kalman filter

The Kalman filter assumes a linear system, but this dramatically decreases the scope of the application of this filter. The extended Kalman filter (EKF) addresses this problem by linearization of the system, thus allowing the application of the filter framework for nonlinear systems.

We consider the following discrete time nonlinear system:

$$
\begin{aligned}
x_t &= g(x_{t-1}, u_t, \varepsilon_t) \\
z_t &= h(x_t, \delta_t)
\end{aligned}
\tag{25}
$$

where $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$, and $z_t \in \mathbb{R}^k$ are the state, control input, and measurement vectors, as before, and the noise sources $\varepsilon_t \in \mathbb{R}^p$ and $\delta_t \in \mathbb{R}^q$ are both normally distributed with zero mean and identity variance[2], both uncorrelated with the state vector. Note that there is no loss in generality of having unit variance in the noise, since the g and h functions may internally

[2]That is, the covariance matrix is the identity matrix.

scale the noise to arbitrary covariances. In fact, the model is sufficiently generic to include any nonlinear effect of the noise into the system evolution and measurements. Note also that the dimensionality of the dimensions of the noise sources, p and q, are unrelated with the dimensionality of the state and of the measurement.

4.1 Prediction step

For the prediction step we linearise the first equation in (25) around $x_{t-1} = \mu_{t-1}$ and $\varepsilon_t = 0$

$$\begin{aligned} x_t &\simeq g(\mu_{t-1}, u_t, 0) + G_t^x\,(x_{t-1} - \mu_{t-1}) + G_t^\varepsilon \varepsilon_t \\ &= G_t^x x_{t-1} + \underbrace{[g(\mu_{t-1}, u_t, 0) - G_t^x \mu_{t-1}]}_{\text{new input } u_t'} + G_t^\varepsilon \varepsilon_t \end{aligned} \tag{26}$$

where G_t^x and G_t^ε are the Jacobians of $g(u_t, \mu_{t-1}, \varepsilon_t)$ with respect to the first and third arguments, taken at the linearization point, defined as

$$G_t^x = \frac{\partial g}{\partial x}(\mu_{t-1}, u_t, 0) = \begin{bmatrix} \frac{\partial g^1(x,u,\varepsilon)}{\partial x^1} & \cdots & \frac{\partial g^1(x,u,\varepsilon)}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g^n(x,u,\varepsilon)}{\partial x^1} & \cdots & \frac{\partial g^n(x,u,\varepsilon)}{\partial x^n} \end{bmatrix}_{(\mu_{t-1}, u_t, 0)} \tag{27}$$

where $g(x, u, \varepsilon) = [g^1(x, u, \varepsilon) \cdots g^n(x, u, \varepsilon)]^T$ and $x = [x^1 \cdots x^n]^T$, and as

$$G_t^\varepsilon = \frac{\partial g}{\partial \varepsilon}(\mu_{t-1}, u_t, 0) = \begin{bmatrix} \frac{\partial g^1(x,u,\varepsilon)}{\partial \varepsilon^1} & \cdots & \frac{\partial g^1(x,u,\varepsilon)}{\partial \varepsilon^p} \\ \vdots & \ddots & \vdots \\ \frac{\partial g^n(x,u,\varepsilon)}{\partial \varepsilon^1} & \cdots & \frac{\partial g^n(x,u,\varepsilon)}{\partial \varepsilon^p} \end{bmatrix}_{(\mu_{t-1}, u_t, 0)} \tag{28}$$

where $g(x, u, \varepsilon) = [g^1(x, u, \varepsilon) \cdots g^n(x, u, \varepsilon)]^T$ and $\varepsilon = [\varepsilon^1 \cdots \varepsilon^p]^T$. The above linearization defines a linear system with the state transition model

$$x_t = A_t' x_{t-1} + B_t' u_t' + \varepsilon_t' \tag{29}$$

where $A_t' = G_t^x$, $B_t' = I$, $u_t' = g(\mu_{t-1}, u_t, 0) - G_t^x \mu_{t-1}$, and $\varepsilon_t' = G_t^\varepsilon \varepsilon_t$. The noise term ε_t' depends linearly from a zero mean and identity covariance normal random variable. Thus, it is also normally distributed with zero mean and covariance given by

$$\begin{aligned} R_t' &= E[G_t^\varepsilon \varepsilon_t (G_t^\varepsilon \varepsilon_t)^T] \\ &= G_t^\varepsilon E[\varepsilon_t \varepsilon_t^T] (G_t^\varepsilon)^T \\ &= G_t^\varepsilon (G_t^\varepsilon)^T \end{aligned} \tag{30}$$

since the covariance of ε_t is the identity matrix. Then, according to (11), the predicted state mean is

$$\begin{aligned}\overline{\mu}_t &= A'_t \mu_{t-1} + B'_t u'_t \\ &= G^x_t \mu_{t-1} + g(\mu_{t-1}, u_t, 0) - G^x_t \mu_{t-1} \\ &= g(\mu_{t-1}, u_t, 0)\end{aligned} \tag{31}$$

and the covariance is

$$\begin{aligned}\overline{\Sigma}_t &= A'_t \Sigma_{t-1} A'^T_t + R'_t \\ &= G^x_t \Sigma_{t-1} (G^x_t)^T + G^\varepsilon_t (G^\varepsilon_t)^T\end{aligned} \tag{32}$$

To summarize, the prediction step of the EKF uses these expressions:

$$\begin{aligned}G^x_t &= \frac{\partial g}{\partial x}(\mu_{t-1}, u_t, 0) \\ G^\varepsilon_t &= \frac{\partial g}{\partial \varepsilon}(\mu_{t-1}, u_t, 0) \\ \overline{\mu}_t &= g(\mu_{t-1}, u_t, 0) \\ \overline{\Sigma}_t &= G^x_t \Sigma_{t-1} (G^x_t)^T + G^\varepsilon_t (G^\varepsilon_t)^T\end{aligned} \tag{33}$$

4.2 Update step

Since the measurement model in EKF is also nonlinear, we will also linearize it. To to so, we linearize (25) around $x_t = \overline{\mu}_t$ and $\delta_t = 0$:

$$\begin{aligned}z_t &\simeq h(\overline{\mu}_t, 0) + H^x_t\,(x_t - \overline{\mu}_t) + H^\delta_t \delta_t \\ &= H^x_t x_t + [h(\mu_t, 0) - H^x_t \overline{\mu}_t] + H^\delta_t \delta_t\end{aligned} \tag{34}$$

where H^x_t and H^δ_t are the Jacobians of $h(\overline{\mu}_t, \delta_t)$ with respect to each one of its arguments

$$\begin{aligned}H^x_t &= \frac{\partial h}{\partial x}(\overline{\mu}_t, 0) \\ H^\delta_t &= \frac{\partial h}{\partial \delta}(\overline{\mu}_t, 0)\end{aligned} \tag{35}$$

Note that in the each step, the relevant model equation is linearized at the latest best estimate of the state: μ_{t-1} for the prediction step, and $\overline{\mu}_t$ for the update one. Under this approximation, expression (34) can be seen as a linear observation model

$$z'_t = C'_t x_t + \delta'_t \tag{36}$$

where $z'_t = z_t - h(\overline{\mu}_t, 0) + H^x_t\overline{\mu}_t$, $C'_t = H^x_t$, and $\delta'_t = H^\delta_t\delta_t$. The Kalman gain is computed as in (24), with the appropriate substitutions:

$$K_t = \overline{\Sigma}_t(H^x_t)^T(H^x_t\overline{\Sigma}_t(H^x_t)^T + H^\delta_t(H^\delta_t)^T)^{-1} \tag{37}$$

noting that δ'_t is normal with zero mean and with covariance $H^\delta_t(H^\delta_t)^T$, obtained similarly to (30). The state mean and covariance follows (24), using now z'_t

$$\begin{aligned}\mu_t &= \overline{\mu}_t + K_t[z_t - h(\overline{\mu}_t, 0) + H^x_t\overline{\mu}_t - H^x_t\overline{\mu}_t] \\ &= \overline{\mu}_t + K_t[z_t - h(\overline{\mu}_t, 0)]\end{aligned} \tag{38}$$

$$\Sigma_t = (I - K_tH_t)\overline{\Sigma}_t \tag{39}$$

In summary, these are the expressions for the update step of the EKF:

$$\begin{aligned}H^x_t &= \frac{\partial h}{\partial x}(\overline{\mu}_t, 0) \\ H^\delta_t &= \frac{\partial h}{\partial \delta}(\overline{\mu}_t, 0) \\ K_t &= \overline{\Sigma}_tH^T_t(H_t\overline{\Sigma}_tH^T_t + H^\delta_t(H^\delta_t)^T)^{-1} \\ \mu_t &= \overline{\mu}_t + K_t[z_t - h(\overline{\mu}_t, 0)] \\ \Sigma_t &= (I - K_tH^x_t)\overline{\Sigma}_t\end{aligned} \tag{40}$$

This concludes the derivation of the Extended Kalman Filter (EKF).

Acknowledgments

The author wishes to thank the useful comments provided by Pedro Lima to an early draft of this document, as well as to the students that helped the author to get rid of some typos.