A Derivation of the Kalman Filter

K.S. Ng

September 26, 2010

1 Some Properties of the Gaussian Density

Definition 1. The (one-dimensional) Gaussian density is defined by

$$\mathcal{N}(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Lemma 1. *Let $f = \mathcal{N}(x; \mu_f, \sigma_f)$ and $g = \mathcal{N}(x; \mu_g, \sigma_g)$. Then the normalisation of the pointwise product $f \times g$ is another Gaussian $\mathcal{N}(x; \mu_{fg}, \sigma_{fg})$, where*

$$\mu_{fg} = \frac{\mu_f \sigma_g^2 + \mu_g \sigma_f^2}{\sigma_f^2 + \sigma_g^2} \quad \textit{and} \quad \sigma_{fg} = \sqrt{\frac{\sigma_f^2 \sigma_g^2}{\sigma_f^2 + \sigma_g^2}}.$$

Proof. Straightforward algebra. □

Definition 2. The convolution of two functions $f(t)$ and $g(t)$ is defined as

$$f \otimes g := \int_{-\infty}^{\infty} f(x-t) g(t) dt.$$

The convolution of two Gaussians is also another Gaussian.

Lemma 2. *Let $f = \mathcal{N}(x; \mu_f, \sigma_f)$ and $g = \mathcal{N}(x; \mu_g, \sigma_g)$. Then*

$$\begin{aligned} f \otimes g &:= \int_{-\infty}^{\infty} \mathcal{N}(x-t; \mu_f, \sigma_f) \mathcal{N}(t; \mu_g, \sigma_g) dt \\ &= \mathcal{N}(x; \mu_f + \mu_g, \sqrt{\sigma_f^2 + \sigma_g^2}). \end{aligned}$$

The following proof of Lemma 2 comes from [Bro03]. The key result we exploit is that the Fourier transform of a convolution is the pointwise product of Fourier transforms:

$$F(f \otimes g) = F(f)F(g),$$

where F is the Fourier transform

$$F(f) = \lambda k. \int_{-\infty}^{\infty} f(x) e^{-2\pi i k x} dx.$$

For a nice and short introduction to the theory of Fourier transforms, see [GBGL08, §III.27]

Proof. (of Lemma 2) We have

$$\begin{aligned}
& F(\mathcal{N}(x;\mu,\sigma)) \\
&= \lambda k. \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} e^{-2\pi i k x} dx && (1) \\
&= \lambda k. \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} e^{-2\pi i k (y+u)} dy && (2) \\
&= \lambda k. \frac{e^{-2\pi i k \mu}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} e^{-2\pi i k y} dy && (3) \\
&= \lambda k. \frac{e^{-2\pi i k \mu}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} (\cos(2\pi k y) - i \sin(2\pi k y)) dy && (4) \\
&= \lambda k. \frac{e^{-2\pi i k \mu}}{\sigma\sqrt{2\pi}} \left[\int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} \cos(2\pi k y) dy - i \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} \sin(2\pi k y)) dy \right] && (5) \\
&= \lambda k. \frac{e^{-2\pi i k \mu}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} \cos(2\pi k y) dy && (6) \\
&= \lambda k. \frac{e^{-2\pi i k \mu}}{\sigma\sqrt{2\pi}} 2 \int_{0}^{\infty} e^{-\frac{y^2}{2\sigma^2}} \cos(2\pi k y) dy && (7) \\
&= \lambda k. \frac{e^{-2\pi i k \mu}}{\sigma\sqrt{2\pi}} \sqrt{\pi 2\sigma^2} e^{-2(\pi k \sigma)^2} && (8) \\
&= \lambda k. e^{-2\pi k \mu i} e^{-2\pi^2\sigma^2 k^2}. && (9)
\end{aligned}$$

The step (2) follows from a change of variable $y = x - \mu$. The step (4) follows from Euler's formula $e^{-i\theta} = \cos\theta - i \sin\theta$. The step (5) follows because the function inside the second integral is an odd function, which integrates to zero. The step (8) follows from the known identity

$$\int_{0}^{\infty} e^{-at^2} \cos(2xt) dt = \frac{1}{2}\sqrt{\frac{\pi}{a}} e^{-\frac{x^2}{a}}.$$

From the above, we have

$$
\begin{aligned}
F(f \otimes g) &= F(f)F(g) \\
&= \lambda k.e^{-2\pi k \mu_f i} e^{-2\pi^2 \sigma_f^2 k^2} e^{-2\pi k \mu_g i} e^{-2\pi^2 \sigma_g^2 k^2} \\
&= \lambda k.e^{-2\pi k(\mu_f + \mu_g) i} e^{-2\pi^2 (\sigma_f^2 + \sigma_g^2) k^2}.
\end{aligned} \tag{10}
$$

Equations (9) and (10) now yields

$$
f \otimes g = \mathcal{N}(x; \mu_f + \mu_g, \sqrt{\sigma_f^2 + \sigma_g^2}).
$$

□

2 Derivation of the Kalman Filter

Suppose we have the following models:

$$
\begin{aligned}
&\mathit{obsModel} : \mathit{State} \to \mathit{Density\ Observation} \\
&(\mathit{obsModel}\ s\ o) = (\mathcal{N}\ s\ \sigma_o\ o) \\
&\mathit{transModel} : \mathit{Action} \to \mathit{State} \to (\mathit{Density\ State}) \\
&(\mathit{transModel}\ a\ s\ s') = (\mathcal{N}\ (s + a)\ \sigma_t\ s').
\end{aligned}
$$

Then we have

$$
\begin{aligned}
&\mathit{obsUpdate} : \mathit{Observation} \to (\mathit{Density\ State}) \to (\mathit{Density\ State}) \\
&(\mathit{obsUpdate}\ o\ (\mathcal{N}\ \mu\ \sigma)) = (\mathit{normalise}\ \lambda s.((\mathcal{N}\ \mu\ \sigma\ s) \times (\mathit{obsModel}\ s\ o))) \\
&\qquad = (\mathit{normalise}\ \lambda s.((\mathcal{N}\ \mu\ \sigma\ s) \times (\mathcal{N}\ s\ \sigma_o\ o))) \\
&\qquad = (\mathit{normalise}\ \lambda s.((\mathcal{N}\ \mu\ \sigma\ s) \times (\mathcal{N}\ o\ \sigma_o\ s))) \\
&\qquad = \left(\mathcal{N}\ \frac{\mu\sigma_o^2 + o\sigma^2}{\sigma_o^2 + \sigma^2}\ \sqrt{\frac{\sigma_o^2\sigma^2}{\sigma_o^2 + \sigma^2}}\right),
\end{aligned}
$$

where the last step follows from Lemma 1. We also have

$$
\begin{aligned}
&\mathit{transUpdate} : \mathit{Action} \to (\mathit{Density\ State}) \to (\mathit{Density\ State}) \\
&(\mathit{transUpdate}\ a\ (\mathcal{N}\ \mu\ \sigma)) = \lambda x. \int_{-\infty}^{\infty} (\mathcal{N}\ u\ \sigma\ s) \times (\mathit{transModel}\ a\ s\ x) ds \\
&\qquad = \lambda x. \int_{-\infty}^{\infty} (\mathcal{N}\ u\ \sigma\ s) \times (\mathcal{N}\ (s + a)\ \sigma_t\ x) ds \\
&\qquad = \lambda x. \int_{-\infty}^{\infty} (\mathcal{N}\ u\ \sigma\ s) \times (\mathcal{N}\ a\ \sigma_t\ (x - s)) ds \\
&\qquad = (\mathcal{N}\ (\mu + a)\ \sqrt{\sigma^2 + \sigma_t^2}),
\end{aligned}
$$

where the last step follows from Lemma 2.

References

[Bro03] P.A. Bromiley. Products and convolutions of Gaussian distributions. Technical Report 2003-003, Imaging Science and Biomedical Engineering Division, Medical School, University of Manchester, 2003.

[GBGL08] Timothy Gowers, June Barrow-Green, and Imre Leader, editors. *The Princeton Companion to Mathematics*. Princeton University Press, 2008.