Completing the Square, step-by-step. **Collins, Sept 1, 2008.**

I went over fairly quickly in class a trick that Bishop (in his PRML book) calls "Completing the Square", for determining what the mean and variance are of a posterior distribution that you *know* should be a Gaussian, because it has the form $\exp\{-1/2(ax^2-2bx+c)\}$. Our result that a Gaussian prior times a Gaussian likelihood yields a Gaussian posterior hinged on this trick, and since it is important to understand this result to understand derivations of the Kalman filter, I thought I should write this note to go over it step-by-step.

To simplify the notation, I am going to write things in terms of a "precision" parameter p that is defined to be one over the variance, that is, $p = 1/\sigma^2$. The formula for a 1D Gaussian (aka Normal distribution) with mean μ and precision p is therefore

$$G(x|\mu,\sigma^2) = G(x|\mu,p^{-1}) = \frac{\sqrt{p}}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}p(x-\mu)^2\right\} \tag{1}$$

This is distribution, so it integrates to 1. Therefore,

$$\int_{-\infty}^{\infty}\frac{\sqrt{p}}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}p(x-\mu)^2\right\}dx \;=\; 1 \tag{2}$$

and therefore

$$\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2}p(x-\mu)^2\right\}dx \;=\; \frac{\sqrt{2\pi}}{\sqrt{p}} \tag{3}$$

We will need this result a bit later.

Let's say that we have just multiplied some prior distribution $p(x)$ times some likelihood function $L(y|x)$ to obtain some function $f(x,y)$ as a result. Further, assume that this function has the form

$$f(x,y) = c_1\exp\left\{-\frac{1}{2}[ax^2-2bx+c]\right\} \tag{4}$$

where c_1 is some constant, where by that I mean any nonzero value at all as long as it doesn't have an x in it. According to Bayes rule, we aren't quite done yet, since we still have to normalize to obtain the posterior $p(x|y)$. We do this by integrating $f(x,y)$ over x from minus infinity to infinity. That is

$$p(x|y) = \frac{f(x,y)}{\int_{-\infty}^{\infty}f(x,y)dx} = \frac{c_1\exp\left\{-\frac{1}{2}[ax^2-2bx+c]\right\}}{\int_{-\infty}^{\infty}c_1\exp\left\{-\frac{1}{2}[ax^2-2bx+c]\right\}} \tag{5}$$

Claim: The yucky looking distribution in equation (5) is in fact just a Gaussian distribution, and it has a mean and precision that we can just "read off by inspection" from equation (4).

Proof: First, let's manipulate equation (4) a bit. Start by taking the a out of the square brackets

$$f(x,y) = c_1\exp\left\{-\frac{1}{2}a\left[x^2-2\frac{b}{a}x+\frac{c}{a}\right]\right\} \tag{6}$$

Now comes the "completing the square part". Inside the square brackets, I am going to add a term $(b^2/a^2-c/a)$ and also subtract the same term, so what I'm really doing is adding 0, and thus not changing anything at all. Why? Because we then will collect terms differently and find that something good happens...

$$
\begin{aligned}
f(x,y) &= c_1 \exp\left\{-\frac{1}{2}a\left[x^2 - 2\frac{b}{a}x + \frac{c}{a}\right]\right\} && (7)\\
&= c_1 \exp\left\{-\frac{1}{2}a\left[x^2 - 2\frac{b}{a}x + \frac{c}{a} + \left(\frac{b^2}{a^2} - \frac{c}{a}\right) - \left(\frac{b^2}{a^2} - \frac{c}{a}\right)\right]\right\} && (8)\\
&= c_1 \exp\left\{-\frac{1}{2}a\left[x^2 - 2\frac{b}{a}x + \left(\frac{c}{a} + \frac{b^2}{a^2} - \frac{c}{a}\right) - \left(\frac{b^2}{a^2} - \frac{c}{a}\right)\right]\right\} && (9)\\
&= c_1 \exp\left\{-\frac{1}{2}a\left[x^2 - 2\frac{b}{a}x + \frac{b^2}{a^2} - \left(\frac{b^2}{a^2} - \frac{c}{a}\right)\right]\right\} && (10)\\
&= c_1 \exp\left\{-\frac{1}{2}a\left[x^2 - 2\frac{b}{a}x + \frac{b^2}{a^2}\right] + \frac{1}{2}a\left[\left(\frac{b^2}{a^2} - \frac{c}{a}\right)\right]\right\} && (11)\\
&= c_1 \exp\left\{-\frac{1}{2}a\left[x^2 - 2\frac{b}{a}x + \frac{b^2}{a^2}\right]\right\} \exp\left\{\frac{1}{2}a\left[\left(\frac{b^2}{a^2} - \frac{c}{a}\right)\right]\right\} && (12)\\
&= c_1\, c_2\, \exp\left\{-\frac{1}{2}a\left[x^2 - 2\frac{b}{a}x + \frac{b^2}{a^2}\right]\right\} && (13)\\
&= c_1\, c_2\, \exp\left\{-\frac{1}{2}a\left(x - \frac{b}{a}\right)^2\right\} && (14)
\end{aligned}
$$

Whew! I told you it was going to be step-by-step. Did you catch what happened between equations (12) and (13)? We took the exp() term that doesn't have any x variables in it, and thus is a constant, and renamed it c_2 to reinforce that it **is** just an unimportant constant. And what was left was a quadratic formula in x that was a perfect square, which we factored!

We've now simplified $f(x,y)$ considerably. And this was the numerator of the posterior distribution $p(x|y)$, as was seen in equation (5). To compute the denominator of that equation, we have to integrate $f(x,y)$ from minus infinity to infinity. But comparing carefully the form of equation (14) and the integral in equation (3), we see immediately that

$$\int_{-\infty}^{\infty} f(x,y)dx = \int_{-\infty}^{\infty} c_1\, c_2\, \exp\left\{-\frac{1}{2}a\left(x - \frac{b}{a}\right)^2\right\}dx = c_1\, c_2\, \frac{\sqrt{2\pi}}{\sqrt{a}}\ . \qquad (15)$$

We now know the numerator and denominator of equation (5), which we plug in to find the normalized posterior distribution

$$p(x|y) = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)dx} = \frac{c_1\, c_2\, \exp\left\{-\frac{1}{2}a\left(x - \frac{b}{a}\right)^2\right\}}{c_1\, c_2\, \frac{\sqrt{2\pi}}{\sqrt{a}}} = \frac{\sqrt{a}}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}a\left(x - \frac{b}{a}\right)^2\right\} \qquad (16)$$

Which by comparison with equation (1) is seen to be a Gaussian with mean b/a and precision a. **QED**

Application: Don't you just hate it when a math text says "application" when what really follows is just more math? Heh heh. So let's use the previous result to rederive what happens when you combine a general Gaussian prior and a general Gaussian likelihood function. Let's let the prior have mean u and variance σ^2, and let the likelihood have mean v and variance s^2. Further, define their respective precision parameters to be $p = 1/\sigma^2$ and $q = 1/s^2$. We thus have

$$p(x) \;=\; G(x|u, p^{-1}) \;=\; \frac{\sqrt{p}}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\, p(x-u)^2\right\} \tag{17}$$

$$L(y|x) \;=\; G(x|v, q^{-1}) \;=\; \frac{\sqrt{q}}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\, q(x-v)^2\right\} \tag{18}$$

Multiplying these together via Bayes rule, but not yet normalizing, we get

$$\begin{aligned}
f(x,y) &= p(x)L(y|x) &(19)\\
&= \frac{\sqrt{p}}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\, p(x-u)^2\right\} \frac{\sqrt{q}}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\, q(x-v)^2\right\} &(20)\\
&= \frac{\sqrt{pq}}{2\pi} \exp\left\{-\frac{1}{2}\,[px^2 - 2pux + u^2 + qx^2 - 2qvx + v^2]\right\} &(21)\\
&= c_1 \exp\left\{-\frac{1}{2}\,[(p+q)x^2 - 2(pu+qv)x + (u^2+v^2)]\right\} &(22)
\end{aligned}$$

By inspection, this is just an instantiation of equation (4) with

$$\begin{aligned}
a &= p+q &(23)\\
b &= (pu+qb) &(24)\\
c &= (u^2+v^2) &(25)
\end{aligned}$$

and therefore our previous results regarding completing the square apply, so that if we were to normalize this (as required by Bayes rule) to get a posterior distribution that integrates to one, we would find that the posterior is a Gaussian distribution with

$$\begin{aligned}
\text{precision} &= (p+q) &(26)\\
\text{mean} &= (pu+qv)/(p+q) &(27)
\end{aligned}$$

If we wanted the results written in terms of variance, rather than precision, we could plug in $p = 1/\sigma^2$ and $q = 1/s^2$ to find

$$\begin{aligned}
\text{variance} &= \frac{1}{\text{precision}} = 1/(\frac{1}{\sigma^2} + \frac{1}{s^2}) \;=\; \frac{s^2\sigma^2}{s^2+\sigma^2} &(28)\\
\text{mean} &= (\frac{1}{\sigma^2}u + \frac{1}{s^2}v)/(\frac{1}{\sigma^2} + \frac{1}{s^2}) = \frac{s^2u + \sigma^2 v}{s^2+\sigma^2} &(29)
\end{aligned}$$

which is what we had derived in the lecture notes the other day.