Completing the square

Several of the explanations on this website require a bit of algebra to understand. "Completing the square" [wikipedia] refers to taking an expression of the form $ax^2 + bx + c$ and, though algebraic manipulation, ending up with an expression of the form $a(x + d)^2 + e$. Completing the square is an useful algebra trick to know since it arises often when multiplying likelihoods by priors, especially in the context of the Normal distribution.

In this reference tutorial, we will demonstrate how to complete the square in both univariate (scalar) and multivariate (matrix) contexts. Incidentally, this article also contains the derivation of the posterior distribution of the mean of the normal distribution when the variance is known in both univariate and multivariate contexts.

Completing the square: univariate

Preliminaries

Consider the case of observing N independent samples from a $\mathrm{Normal}(\mu, \sigma^2)$ distribution. Suppose we know the true value of σ^2, and we are interested in determining the posterior distribution of μ. It is conventional to place a conjugate normal prior on μ. Our model is:

$$y_i \overset{iid}{\sim} \mathrm{Normal}(\mu, \sigma^2)$$
$$\mu \sim \mathrm{Normal}(\mu_0, \tau^2)$$

In Bayesian statistics, the posterior is proportional to the likelihood times the prior. Because all observations are independent, our likelihood is the product of N Normal density functions: one for each y_i. The prior then provides another Normal density function term. After simplifying and dropping terms that are not functions of μ, we end up with a posterior distribution for μ proportional to

$$\exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^N (y_i - \mu)^2 \right\} \exp\left\{ -\frac{1}{2\tau^2} (\mu - \mu_0)^2 \right\}$$

The first term is the likelihood, and the second term is the prior. Because we desire a posterior that is a simple function of μ, we need to gather all the terms that include μ together; as the posterior is written above, μ is scattered across $N + 1$ terms. Completing the square is the trick that will allow us to gather all the μ terms into one.

The first step is to expand the squares containing μ. This yields

$$\exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^N \left(y_i^2 - 2\mu y_i - \mu^2\right) \right\} \exp\left\{ -\frac{1}{2\tau^2} (\mu^2 - 2\mu\mu_0 + \mu_0^2)^2 \right\}$$

We now distribute the sum, yielding

$$\exp\left\{-\tfrac{1}{2\sigma^2}\left(\textcolor{red}{\sum_{i=1}^{N} y_i^2} - 2\mu \sum_{i=1}^{N} y_i + N\mu^2\right)\right\} \exp\left\{-\tfrac{1}{2\tau^2}(\mu^2 - 2\mu\mu_0 + \textcolor{red}{\mu_0^2})^2\right\}$$

There are several terms above that do not involve μ; these are highlighted in red. When we drop those terms, combine all remaining terms within one exponent, and then focus only on what is in the exponent, what remains is

$$-\tfrac{1}{2}\left(\tfrac{N}{\sigma^2}\mu^2 + \tfrac{1}{\tau^2}\mu^2 - 2\mu\tfrac{\sum_{i=1}^{N} y_i}{\sigma^2} - 2\mu\tfrac{1}{\tau^2}\mu_0\right)$$

We can combine the μ^2 terms together, and the μ terms together:

$$-\tfrac{1}{2}\left(\left(\tfrac{N}{\sigma^2} + \tfrac{1}{\tau^2}\right)\mu^2 - 2\left(\tfrac{\sum_{i=1}^{N} y_i}{\sigma^2} + \tfrac{1}{\tau^2}\mu_0\right)\mu\right)$$

Finally, for ease of notation, we can use the fact that $\sum_{i=1}^{N} y_i = N\bar{y}$:

$$-\tfrac{1}{2}\left(\left(\tfrac{N}{\sigma^2} + \tfrac{1}{\tau^2}\right)\mu^2 - 2\left(\tfrac{N}{\sigma^2}\bar{y} + \tfrac{1}{\tau^2}\mu_0\right)\mu\right)$$

After all of this algebraic simplification, inside the parentheses what we have looks something like $ax^2 - 2bx$. We can now "complete the square" to obtain something of the form $(x - c)^2$.

Competing the square

Our first step is to make the notation easier to follow. Let $a = \tfrac{N}{\sigma^2} + \tfrac{1}{\tau^2}$ and $b = \tfrac{N}{\sigma^2}\bar{y} + \tfrac{1}{\tau^2}\mu_0$. Using the new, simplified notation, we have

$$-\tfrac{1}{2}\left(a\mu^2 - 2b\mu\right)$$

We can move the coefficient a on μ outside the parentheses:

$$-\tfrac{a}{2}\left(\mu^2 - 2\tfrac{b}{a}\mu\right)$$

We now add and subtract the same value inside the parentheses. This doesn't change the value at all, since the terms sum to 0:

$$-\tfrac{a}{2}\left(\mu^2 - 2\tfrac{b}{a}\mu + \tfrac{b^2}{a^2} \textcolor{red}{- \tfrac{b^2}{a^2}}\right)$$

In fact, neither term is a function of μ, so we can simply drop the term colored in red.

$$-\tfrac{a}{2}\left(\mu^2 - 2\tfrac{b}{a}\mu + \tfrac{b^2}{a^2}\right)$$

The terms within the parentheses are of the form $x^2 - 2xc + c^2$, which, from the rules learned in algebra, can be simplified to $(x - c)^2$. Applying this to our terms, we obtain:

$$-\frac{a}{2}\left(\mu - \frac{b}{a}\right)^2$$

We have thus completed the square. We are not done, however: this was only the portion of the posterior distribution that was in the exponent. Replacing the terms in the exponent yields:

$$\exp\left\{-\frac{a}{2}\left(\mu - \frac{b}{a}\right)^2\right\}$$

By completing the square, we have revealed that the posterior distribution of μ has the form of a normal distribution with a mean of b/a and a variance of $1/a$, or

$$\mu \mid y \sim \mathrm{Normal}\left(\mu_n, \sigma_n^2\right)$$

where

$$\sigma_n^2 = \frac{1}{a} = \left(\frac{N}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1},$$

$$\mu_n = \frac{b}{a} = \sigma_n^2\left(\frac{N}{\sigma^2}\bar{y} + \frac{1}{\tau^2}\mu_0\right).$$

Completing the square: multivariate

If we are working with multivariate distributions, it is often customary to represent them as functions of vectors and matrices. The density of the p-variate multivariate normal distribution, for instance, is defined as

$$p(\boldsymbol{y} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-\frac{p}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right\}$$

where data $\boldsymbol{y}$ and mean $\boldsymbol{\mu}$ are $p \times 1$ vectors, and $\boldsymbol{\Sigma}$ is a $p \times p$ covariance matrix. Consider the case that we have N observations from a multivariate normal with known covariance, and we wish to find the posterior distribution for the mean vector $\boldsymbol{\mu}$. This case is analogous to the univariate case above, and, analogously, we will use a multivariate Normal prior on $\boldsymbol{\mu}$. Our model is:

$$\boldsymbol{y}_i \overset{iid}{\sim} \mathrm{Multivariate\ Normal}_p\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right)$$
$$\boldsymbol{\mu} \sim \mathrm{Multivariate\ Normal}_p\left(\boldsymbol{\mu}_0, \boldsymbol{V}\right)$$

In the previous section, we detailed a lot of the preliminary algebra in deriving the posterior distribution. Since the algebra is analogous in the multivariate case, there is no reason to reproduce it here. After the preliminary algebra and dropping terms that are irrelevant, the terms inside exponent are

$$-\frac{1}{2}\left(\boldsymbol{\mu}'\left(N\boldsymbol{\Sigma}^{-1} + \boldsymbol{V}^{-1}\right)\boldsymbol{\mu} - \boldsymbol{\mu}'\left(N\boldsymbol{\Sigma}^{-1}\bar{\boldsymbol{y}} + \boldsymbol{V}^{-1}\boldsymbol{\mu}_0\right) - \left(N\boldsymbol{\Sigma}^{-1}\bar{\boldsymbol{y}} + \boldsymbol{V}^{-1}\boldsymbol{\mu}_0\right)'\boldsymbol{\mu}\right)$$

where $\bar{\boldsymbol{y}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{y}$. At this point, we would like to complete the square in a fashion analogous to the way we completed the square in the univariate case. To facilitate this, we introduce analogous notation. Let

$A = N\Sigma^{-1} + V^{-1}$, and let $b = N\Sigma^{-1}\bar{y} + V^{-1}\mu_0$. With this notation, the previous expression is equal to

$$-\tfrac{1}{2}\left(\mu' A \mu - \mu' b - b' \mu\right).$$

In the univariate case, our next two steps were to bring a — analogous to A here — outside the parentheses, and then to introduce new terms to complete the square. However, we cannot do that here. A is between the two vector terms μ' and μ; because matrix multiplication does not commute, we cannot multiply the term by anything to move A. We can, though, introduce a term that is not a function of μ to complete the square:

$$-\tfrac{1}{2}\left(\mu' A \mu - \mu' b - b' \mu + b' A^{-1} b\right).$$

In order to deal with the difficulty with A, we note three things: first, that A is symmetric and invertible, because it is the weighted sum of two symmetric, full-rank covariance matrices; and second, that the identity matrix $I = A^{-1}A = AA^{-1}$. Since multiplying by I does not change a matrix expression as long as the matrix is the proper dimension, we can rewrite our expression above:

$$-\tfrac{1}{2}\left(\mu' A \mu - \mu' A A^{-1} b - b' A^{-1} A \mu + b' A^{-1} A A^{-1} b\right).$$

We can now introduce new notation, analogous to the notation at the end of the univariate competing the square section above. Let $\Sigma_n = A^{-1}$ and $\mu_n = A^{-1}b$. Then the expression above becomes

$$-\tfrac{1}{2}\left(\mu' \Sigma_n^{-1} \mu - \mu' \Sigma_n^{-1} \mu_n - \mu_n' \Sigma_n^{-1} \mu + \mu_n' \Sigma_n^{-1} \mu_n\right).$$

It is now obvious that we can factor the above expression into

$$-\tfrac{1}{2}\left(\mu - \mu_n\right)' \Sigma_n^{-1} \left(\mu - \mu_n\right).$$

If you don't see why, try expanding the expression above. Replacing the terms into the exponential function, we see that the posterior distribution is proportional to

$$\exp\left\{-\tfrac{1}{2}\left(\mu - \mu_n\right)' \Sigma_n^{-1} \left(\mu - \mu_n\right)\right\}.$$

Because this is proportional to the probability density function for a multivariate Normal distribution with mean vector μ_n and covariance matrix Σ_n, we know the exact form of the posterior distribution for the multivariate Normal mean vector μ when the prior is also multivariate Normal:

$$\mu \mid Y \sim \text{Multivariate Normal}_p\left(\mu_n, \Sigma_n\right),$$

where

$$\begin{aligned}
\Sigma_n &= A^{-1} = \left(N\Sigma^{-1} + V^{-1}\right)^{-1}, \\
\mu_n &= A^{-1}b = \Sigma_n\left(N\Sigma^{-1}\bar{y} + V^{-1}\mu_0\right).
\end{aligned}$$

This is analogous to the posterior distribution in the univariate case, as can be easily seen by comparing the parameters of the respective posteriors.