"Completing the square"

Let $Y_i \mid \theta, \sigma^2 \overset{iid}{\sim} \text{normal}(\theta, \sigma^2)$ *note: $\vec{y} = \{y_1, \dots, y_n\}$

Let $\theta \mid \sigma^2 \sim \text{normal}(\mu_0, \tau_0^2)$

We want to prove that the full conditional posterior $\theta \mid \sigma^2, \vec{y}$ is normal with some posterior mean μ_n and posterior variance τ_n^2. i.e. we want to prove:

$$\boxed{\theta \mid \sigma^2, \vec{y} \sim \text{normal}(\mu_n, \tau_n^2)}$$

PROOF:

STEP 1: Bayes' thm:

$$p(\theta \mid \sigma^2, \vec{y}) \propto \underbrace{p(\vec{y} \mid \theta, \sigma^2)}_{\prod_{i=1}^{n} \text{dnorm}(y_i, \theta, \sigma)} p(\theta \mid \sigma^2)$$

by iid

$$\propto \exp\left\{\frac{-1}{2\sigma^2} \Sigma (y_i - \theta)^2\right\} \cdot \exp\left\{\frac{-1}{2\tau_0^2} (\theta - \mu_0)^2\right\}$$

STEP 2: simplify by expanding quadratic terms and absorbing what we can into the normalizing constant. $\boxed{\text{*Note: } \Sigma y_i = n\bar{y}}$

$$\propto \exp\left\{\frac{-1}{2\sigma^2}(n\theta^2 - 2n\bar{y}\theta) - \frac{1}{2\tau_0^2}(\theta^2 - 2\mu_0\theta)\right\}$$

combine like terms:

$$\propto \exp\left\{-\frac{1}{2}\underbrace{\left(\frac{n}{\sigma^2} + \frac{1}{\tau_0^2}\right)}_{\text{"a"}}\theta^2 + \underbrace{\left(\frac{n\bar{y}}{\sigma^2} + \frac{\mu_0}{\tau_0^2}\right)}_{\text{"b"}}\theta\right\}$$

define "a" & "b":

$$\propto \exp\left\{-\frac{1}{2}a\theta^2 + b\theta\right\}$$

(✱) Notice: this is the kernel of a normal with $\boxed{\text{mean} = \frac{b}{a}}$ and $\boxed{\text{variance} = a^{-1}}$

therefore, $\theta \mid \sigma^2, \vec{y} \sim \text{normal}(\mu_n, \tau_n^2)$ where $\mu_n = \frac{b}{a}$ and $\tau_n^2 = a^{-1}$

and a, b are defined above. □

Proof of (*) on previous page: This answers the question:

How did I recognize $\exp\{-\frac{1}{2}a\theta^2 + b\theta\}$ as the kernel of a normal w/ mean $\frac{b}{a}$ and variance a^{-1}?

Assume $\theta \mid \vec{y}, \sigma^2 \sim \text{normal}\left(\frac{b}{a}, a^{-1}\right)$.

Then
$$
\begin{aligned}
p(\theta \mid \vec{y}, \sigma^2) &\propto \exp\left\{-\frac{1}{2} a \left(\theta - \frac{b}{a}\right)^2\right\} \\
&\propto \exp\Big\{-\frac{1}{2}a\theta^2 + b\theta \underbrace{- \frac{1}{2}\cancel{a}\frac{b^2}{a^{\cancel{2}}}}_{\text{constant in } \theta.}\Big\} \\
&\propto \exp\left\{-\frac{1}{2}a\theta^2 + b\theta\right\}
\end{aligned}
$$

* Today's punchline:

Assume:

1. $Y_i \mid \theta, \sigma^2 \sim N(\theta, \sigma^2)$
2. $\theta \mid \sigma^2 \sim N(\mu_0, \sigma^2/\kappa_0)$
3. $1/\sigma^2 \sim \text{gamma}\left(\frac{\nu_0}{2}, \frac{\nu_0}{2}\sigma_0^2\right)$

then the posterior

$$p(\theta, \sigma^2 \mid y_1, \dots, y_n)$$

$$= \underbrace{p(\theta \mid \sigma^2, y_1, \dots, y_n)}_{\text{dnorm}(\theta;\ \mu_n,\ \tau_n)} \underbrace{p(\sigma^2 \mid y_1, \dots, y_n)}_{\text{dinvgamma}\left(\sigma^2;\ \frac{\nu_n}{2},\ \frac{\nu_n}{2}\sigma_n^2\right)}$$

"full cond'l posterior of θ"

Today's agenda:

(1) sketch proof for $p(\sigma^2 \mid y_1, \dots, y_n)$

(2) sample from joint posterior

(3) sample from posterior predictive $p(\tilde{y} \mid y_1, \dots, y_n)$ (time permitting)

Interpretation:

μ_0 : prior guess for θ

σ_0^2 : prior guess for σ^2

κ_0 : prior sample size for θ

ν_0 : prior sample size for σ^2

Assume

① $Y_i \mid \theta, \sigma^2 \overset{iid}{\sim} N(\theta, \sigma^2)$

② $\theta \mid \sigma^2 \sim N(\mu_0, \sigma^2 / \kappa_0)$

③ $1/\sigma^2 \sim \text{gamma}\left(\nu_0/2, \frac{\nu_0 \sigma_0^2}{2}\right)$

then the joint posterior

$$p(\theta, \sigma^2 \mid \vec{y}) = p(\theta \mid \sigma^2, \vec{y}) \, p(\sigma^2 \mid \vec{y})$$

$$= \text{dnorm}(\theta, \mu_n, \tau_n^2) \cdot \text{dinvgamma}\left(\sigma^2, \frac{\nu_n}{2}, \frac{\nu_n \sigma_n^2}{2}\right)$$

✓ shown previously

Goal: to prove $\sigma^2 \mid \vec{y} \sim$ invgamma

Sketch proof:

$$p(\sigma^2 \mid \vec{y}) \propto p(\vec{y} \mid \sigma^2) \, p(\sigma^2)$$

$$\propto p(\sigma^2) \int p(\vec{y}, \theta \mid \sigma^2) \, d\theta$$

$$\propto p(\sigma^2) \int p(\vec{y} \mid \theta, \sigma^2) \, p(\theta \mid \sigma^2) \, d\theta \qquad (\dagger)$$

Known by assumption: ③ ① ②

The integral $(\dagger)$ above reduces to integrating a normal density.

Explicit bookkeeping: attacking $\int p(\vec{y}\mid\theta,\sigma^2)\,p(\theta\mid\sigma^2)\,d\theta$

(1) collect (but **do not** absorb into normalizing const!) terms w/o θ outside integral:

call this "$= C_1(\sigma^2)$"

$$\boxed{(2\pi\sigma^2)^{-n/2}\cdot\left(2\pi\sigma^2/K_0\right)^{-1/2}\cdot\exp\left\{\frac{-1}{2\sigma^2}\left(\sum y_i^2\right)-\frac{K_0}{2\sigma^2}(\mu_0)^2\right\}}\cdot$$

$$\cdot\int\exp\left\{\frac{-1}{2\sigma^2}\left(-2\left(\sum y_i\right)\theta+n\theta^2\right)-\frac{K_0}{2\sigma^2}\left(-2\mu_0\theta+\theta^2\right)\right\}d\theta$$

(2) Compute the integral above using the kernel trick:
hint: it's the kernel of a normal, you must complete the square.

$$\int e^{-\frac{1}{2}a\theta^2+b\theta}\,d\theta=\boxed{(2\pi/a)^{1/2}\exp\left\{\tfrac{1}{2}b^2/a\right\}}$$

call this "$= C_2(\sigma_2)$"

where $a=\left(\frac{n}{\sigma^2}+\frac{K_0}{\sigma^2}\right)$

$b=\dfrac{(n\bar{y}+K_0\mu_0)}{\sigma^2}$

(3) Return to (†) and write down:

$$p(\sigma^2\mid\vec{y})\propto p(\sigma^2)\cdot C_1(\sigma^2)\cdot C_2(\sigma^2)$$