Exercise 1.3.

Consider $y_n = x\theta + \varepsilon_n$

where $x \equiv$ constant $\quad \varepsilon_n \sim N(0, \sigma_\varepsilon^2) \quad n=1,\dots,N$

$\theta \sim N(\mu_p, \sigma_p^2)$

a) Likelihood: $p(y_n|\theta) \sim N(x\theta, \sigma_\varepsilon^2) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left(\frac{-(y_n - x\theta)^2}{2\sigma_\varepsilon^2}\right)$

x is a constant and θ is given (conditioned) so ε_n is the only source of randomness

$p(Y|\theta)$ where $Y = \{y_1, \dots, y_N\}$ can be obtained under the independence assumption: $p(y_1 \dots y_N|\theta) \overset{iid}{=} \prod_{n=1}^{N} p(y_n|\theta) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left(\frac{-(y_n - x\theta)^2}{2\sigma_\varepsilon^2}\right)$

b) We can use Bayes' theorem $p(\theta|Y) = \frac{p(Y|\theta)p(\theta)}{p(Y)}$ to find the posterior. The log posterior is, then

$$\log p(\theta|Y) = \log p(Y|\theta) + \log p(\theta) - \log p(Y)$$

Note that:

* $C_e = -\log p(Y)$ is just a constant wrt θ, called "evidence"
* $\log p(Y|\theta) = \log \prod_{n=1}^{N} p(y_n|\theta) = \sum_{n=1}^{N} \log p(y_n|\theta)$

  and $\log p(y_n|\theta) = \frac{-1}{2}\log 2\pi\sigma_\varepsilon^2 + \frac{-1}{2\sigma_\varepsilon^2}(y_n - x\theta)^2$

  $\sum_{n=1}^{N}(y_n - x\theta)^2 = \sum_{n=1}^{N} y_n^2 - 2y_n x\theta + x^2\theta^2 = \sum_{n=1}^{N} y_n^2 - 2N\bar{y}x\theta + Nx^2\theta^2$
* $\log p(\theta) = \frac{-1}{2}\log 2\pi\sigma_p^2 + \frac{-1}{2\sigma_p^2}(\theta - \mu_p)^2$

Putting everything together:

$$\log p(\theta|Y) = \sum_{n=1}^{N}\left[\frac{-1}{2}\log 2\pi\sigma_\varepsilon^2 + \frac{-1}{2\sigma_\varepsilon^2}(y_n - x\theta)^2\right] + \left[\frac{-1}{2}\log 2\pi\sigma_p^2 + \frac{-1}{2\sigma_p^2}(\theta - \mu_p)^2\right] + C_e$$

$$= \frac{-N}{2}\log 2\pi\sigma_\varepsilon^2 + \frac{-1}{2\sigma_\varepsilon^2}\left(\sum_{n=1}^{N} y_n^2 - 2N\bar{y}x\theta + Nx^2\theta^2\right) + \frac{-1}{2}\log 2\pi\sigma_p^2 + \frac{-1}{2\sigma_p^2}(\theta - \mu_p)^2 + C_e =$$

$$= \frac{-1}{2}\log (2\pi\sigma_\varepsilon^2)^N 2\pi\sigma_p^2 + \frac{-1}{2}\left[\theta^2\left(\frac{1}{\sigma_p^2} + \frac{Nx^2}{\sigma_\varepsilon^2}\right) - 2\theta\left(\frac{N\bar{y}x}{\sigma_\varepsilon^2} + \frac{\mu_p}{\sigma_p^2}\right) + \underbrace{\boxed{\frac{\sum_{n=1}^{N} y_n^2}{\sigma_\varepsilon^2} + \frac{\mu_p^2}{\sigma_p^2}}}_{K}\right] + C_e$$

(this aggrupation of θ terms will be useful for later)

c) Compare with $\log p(\theta|\mu, \sigma^2) = \frac{-1}{2}\log 2\pi\sigma^2 + \frac{-1}{2\sigma^2}(\theta - \mu)^2$

$\Rightarrow$ We can see that the overall structure is the same, but μ and σ^2 have more complicated expressions.

d) Comparing the expression in ⓒ $\log p(\theta|\mu,\sigma^2) = \frac{-1}{2}\log 2\pi\sigma^2 + \frac{-1}{2\sigma^2}\underbrace{(\theta-\mu)^2}_{(\theta^2-2\theta\mu+\mu^2)}$
with the expression in ⓑ:

$$\log p(\theta|Y) = \frac{-1}{2}\log (2\pi\sigma_\varepsilon^2)^N \cdot 2\pi\sigma_\beta^2 + \frac{-1}{2}\left[\theta^2\left(\frac{1}{\sigma_\beta^2} + \frac{Nx^2}{\sigma_\varepsilon^2}\right) - 2\theta\left(\frac{N\bar{y}x}{\sigma_\varepsilon^2} + \frac{\mu_\beta}{\sigma_\beta^2}\right) + K\right] + C_e$$

we can easily see that

$$\circledast\ \frac{1}{\sigma^2} = \frac{1}{\sigma_\beta^2} + \frac{Nx^2}{\sigma_\varepsilon^2} = \frac{\sigma_\varepsilon^2 + \sigma_\beta^2 N x^2}{\sigma_\beta^2\sigma_\varepsilon^2} \Rightarrow \boxed{\sigma^2 = \frac{\sigma_\beta^2\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\beta^2 N x^2}}$$

$$\circledast\ \frac{\mu}{\sigma^2} = \frac{N\bar{y}x}{\sigma_\varepsilon^2} + \frac{\mu_\beta}{\sigma_\beta^2} = \frac{N\bar{y}x\cdot\sigma_\beta^2 + \sigma_\varepsilon^2\mu_\beta}{\sigma_\varepsilon^2\cdot\sigma_\beta^2} \Rightarrow \mu = \frac{N\bar{y}x\cdot\sigma_\beta^2 + \sigma_\varepsilon^2\mu_\beta}{\sigma_\varepsilon^2\sigma_\beta^2}\cdot\sigma^2$$

$$\Rightarrow \boxed{\mu = \frac{N\bar{y}x\,\sigma_\beta^2 + \sigma_\varepsilon^2\mu_\beta}{\sigma_\varepsilon^2 + \sigma_\beta^2 N x^2}}$$

Then, for the equality to verify, we must "complete the square" adding and subtracting μ^2. Also we must add and subtract $\frac{-1}{2}\log 2\pi\sigma^2$ =

$$\log p(\theta|Y) = \frac{-1}{2}\log(2\pi\sigma_\varepsilon^2)^N\cdot 2\pi\sigma_\beta^2 \pm \frac{-1}{2}\log 2\pi \overbrace{\frac{\sigma_\beta^2\sigma_\varepsilon^2}{\sigma_\varepsilon^2+\sigma_\beta^2 N x^2}}^{\sigma^2} +$$

$$+ \frac{-1}{2\underbrace{\left(\frac{\sigma_\beta^2\sigma_\varepsilon^2}{\sigma_\varepsilon^2+\sigma_\beta^2Nx^2}\right)}_{\sigma^2}}\Bigg[\underbrace{\theta^2 - 2\theta\underbrace{\left(\frac{N\bar{y}x\sigma_\beta^2 + \sigma_\varepsilon^2\mu_\beta}{\sigma_\varepsilon^2 + \sigma_\beta^2 N x^2}\right)}_{\mu} + \mu^2}_{(\theta-\mu)^2} - \mu^2 + K\Bigg] + C_e =$$

$$= \underbrace{\frac{-1}{2}\log 2\pi\sigma^2 + \frac{-1}{2\sigma^2}(\theta-\mu)^2}_{\text{same as ⓒ}} + C$$

where C is a proportionality constant formed by all the other elements (none of which depends on θ):

$$C = \frac{-1}{2}\log(2\pi\sigma_\varepsilon^2)^N\cdot \cancel{2\pi\sigma_\beta^2} - \frac{-1}{2}\log 2\pi\frac{\cancel{\sigma_\beta^2}\sigma_\varepsilon^2}{\sigma_\varepsilon^2+\sigma_\beta^2Nx^2} + \frac{-1}{2\sigma^2}\left[K-\mu^2\right] + C_e$$