

HW #1 solutions (2026)

① Ex. 2.4,

$$x_1 \sim \mathcal{N}(\mu_1, \sigma_1^2), \quad x_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$$

$$p(y) = ? \quad , \quad \text{where} \quad y = x_1 + x_2$$

$$\begin{aligned} \text{Note that } \mathcal{N}(x|\mu, \sigma^2) &= \mathcal{G}(x-\mu|\sigma) = \\ &= \frac{1}{\sigma} \mathcal{G}\left(\frac{x-\mu}{\sigma}\right) \quad , \quad \text{where} \end{aligned}$$

$$\mathcal{G}(z) = \text{pdf of } \mathcal{N}(0, 1) \text{ standard normal}$$

$$\begin{aligned} \text{Then } p(y) &= \int dx_1 \frac{1}{\sigma_1} \mathcal{G}\left(\frac{x_1 - \mu_1}{\sigma_1}\right) \frac{1}{\sigma_2} \mathcal{G}\left(\frac{y - x_1 - \mu_2}{\sigma_2}\right) = \\ &= \frac{1}{2\pi \sigma_1 \sigma_2} \int dx_1 e^{-\frac{1}{2}\left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2} e^{-\frac{1}{2}\left(\frac{y - x_1 - \mu_2}{\sigma_2}\right)^2} \end{aligned}$$

Complete the square:

$$\begin{aligned} \left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2 + \left(\frac{y - x_1 - \mu_2}{\sigma_2}\right)^2 &= \\ &= (\omega_1 + \omega_2)(x_1 - \bar{x})^2 + \frac{\omega_1 \omega_2}{\omega_1 + \omega_2} (y - (\mu_1 + \mu_2))^2, \end{aligned}$$

$$\text{where } \begin{cases} \omega_1 = \sigma_1^{-2}, \\ \omega_2 = \sigma_2^{-2}. \end{cases} \quad \frac{\omega_1 \omega_2}{\omega_1 + \omega_2} = \frac{1}{\sigma_1^2 + \sigma_2^2}$$

$$\bar{x} = \frac{\omega_1 \mu_1 + \omega_2 (y - \mu_2)}{\omega_1 + \omega_2}$$

Finally,

$$p(y) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2(\sigma_1^2+\sigma_2^2)}(y-(\mu_1+\mu_2))^2} \times$$

$$\times \int dx_1 e^{-\frac{1}{2}(\omega_1+\omega_2)(x_1-\bar{x})^2} \quad \textcircled{=}$$

$$\sqrt{2\pi} \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}}$$

$$\textcircled{=} \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} e^{-\frac{1}{2(\sigma_1^2+\sigma_2^2)}[y-(\mu_1+\mu_2)]^2} =$$

$$= \mathcal{N}(y | \mu_1+\mu_2, \sigma_1^2+\sigma_2^2).$$

2

Ex. 2.6Variance of a sum

$$V[X+Y] = E[(X+Y)^2] - (E[X] + E[Y])^2 =$$

$$= E[X^2 + Y^2 + 2XY] - [E[X]^2 + E[Y]^2 + 2E[X]E[Y]]$$

$$\ominus (E[X^2] - E[X]^2) + (E[Y^2] - E[Y]^2) +$$

$$+ 2(E[XY] - E[X]E[Y]) =$$

$$= \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y].$$

of X & Y are indep., $\text{Cov}[X, Y] = 0$

and $V[X+Y] = V[X] + V[Y].$

3.

(a) Consider $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix}$, then

$$\Delta = \frac{1}{\det \Sigma} \begin{pmatrix} \Sigma_{22} & -\Sigma_{12} \\ -\Sigma_{12} & \Sigma_{11} \end{pmatrix}.$$

" $\Sigma_{11} \Sigma_{22} - \Sigma_{12}^2 \equiv \Delta$

Then $p(y_1, y_2) = \frac{1}{2\pi\sqrt{\Delta}} e^{-\frac{1}{2\Delta} [\Sigma_{22} \underbrace{(y_1 - \mu_1)^2}_{x_1} - 2\Sigma_{12} (y_1 - \mu_1)(y_2 - \mu_2) + \Sigma_{11} \underbrace{(y_2 - \mu_2)^2}_{x_2}]}$.

Next, consider

$$p(y_1) = \int_{-\infty}^{\infty} dy_2 p(y_1, y_2).$$

Complete the square in the exponent:

$$\Sigma_{11} x_2^2 - 2\Sigma_{12} x_1 x_2 = \Sigma_{11} \left(x_2 - \frac{\Sigma_{12}}{\Sigma_{11}} x_1\right)^2 - \frac{\Sigma_{12}^2}{\Sigma_{11}} x_1^2.$$

Then $[...] = \Sigma_{11} \left(x_2 - \frac{\Sigma_{12}}{\Sigma_{11}} x_1\right)^2 + \underbrace{\left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)}_{\frac{\Delta}{\Sigma_{11}}} x_1^2.$

Finally, $p(y_1) = \frac{1}{2\pi\sqrt{\Delta}} e^{-\frac{x_1^2}{2\Sigma_{11}}} \times$

$$\times \underbrace{\int_{-\infty}^{\infty} dx_2 e^{-\frac{\Sigma_{11}}{2\Delta} \left(x_2 - \frac{\Sigma_{12}}{\Sigma_{11}} x_1\right)^2}}_{\sqrt{\frac{2\pi\Delta}{\Sigma_{11}}}} = \frac{1}{\sqrt{2\pi\Sigma_{11}}} e^{-\frac{(y_1 - \mu_1)^2}{2\Sigma_{11}}}$$

as expected.

Thus, $y_1 \sim \mathcal{N}(\mu_1, \Sigma_{11}).$

Similarly, $y_2 \sim \mathcal{N}(\mu_2, \Sigma_{22})$;
the calculation is the same with
 $1 \leftrightarrow 2$ everywhere.

(b) Recall that

$$p(y_1, y_2) \sim e^{-\frac{1}{2}(\vec{y} - \vec{\mu})^T \Lambda (\vec{y} - \vec{\mu})} =$$

$$= e^{-\frac{1}{2}[\Lambda_{11}x_1^2 + 2\Lambda_{12}x_1x_2 + \Lambda_{22}x_2^2]}.$$

Then $p(y_1|y_2) \sim e^{-\frac{1}{2}[\underbrace{\Lambda_{11}x_1^2 + 2\Lambda_{12}x_1x_2}_{\text{fixed } y_2}]} \quad \textcircled{\sim}$

$$\Lambda_{11}\left(x_1 + \frac{\Lambda_{12}}{\Lambda_{11}}x_2\right)^2 + \text{const}(x_1)$$

$$\textcircled{\sim} e^{-\frac{1}{2}\Lambda_{11}\left(x_1 + \frac{\Lambda_{12}}{\Lambda_{11}}x_2\right)^2}.$$

Therefore, $p(y_1|y_2) = \mathcal{N}\left(y_1 \mid \mu_1 - \frac{\Lambda_{12}}{\Lambda_{11}} \underbrace{x_2}_{y_2 - \mu_2}, \Lambda_{11}^{-1}\right)$

Moreover, $-\frac{\Lambda_{12}}{\Lambda_{11}} = \frac{\Sigma_{12}}{\Sigma_{22}},$

$$\Lambda_{11}^{-1} = \Sigma_{11} - \frac{\Sigma_{12}^2}{\Sigma_{22}} = \frac{\Delta}{\Sigma_{22}}.$$

Thus, $p(y_1|y_2) = \mathcal{N}\left(y_1 \mid \mu_1 + \frac{\Sigma_{12}}{\Sigma_{22}}(y_2 - \mu_2), \frac{\Delta}{\Sigma_{22}}\right)$

4. (a) Find $p(z|y)$.

$$p(z|y) \sim p(y|z)p(z) \sim e^{-\frac{(z-\mu_z)^2}{2\sigma_z^2}} e^{-\frac{(y-\omega z-b)^2}{2\sigma_y^2}} \quad \textcircled{\sim}$$

$$\textcircled{\sim} e^{-\frac{1}{2} \left[\frac{z^2}{\sigma_z^2} - \frac{2\mu_z z}{\sigma_z^2} + \frac{\omega^2 z^2}{\sigma_y^2} - \frac{2\omega(y-b)z}{\sigma_y^2} \right]} \quad \textcircled{=}$$

↑
only need
terms involving z

$$\textcircled{=} e^{-\frac{1}{2} \left[\underbrace{\left(\frac{1}{\sigma_z^2} + \frac{\omega^2}{\sigma_y^2} \right)}_{\frac{1}{\sigma_{z|y}^2}} z^2 - 2 \underbrace{\left(\frac{\mu_z}{\sigma_z^2} + \frac{\omega(y-b)}{\sigma_y^2} \right)}_{\frac{\mu_{z|y}}{\sigma_{z|y}^2}} z \right]}$$

Hence, $p(z|y) = \mathcal{N}(z | \mu_{z|y}, \sigma_{z|y}^2)$, where

$$\begin{cases} \sigma_{z|y}^2 = \frac{\sigma_z^2 \sigma_y^2}{\sigma_y^2 + \omega^2 \sigma_z^2}, \\ \mu_{z|y} = \sigma_{z|y}^2 \left[\frac{\mu_z}{\sigma_z^2} + \frac{\omega(y-b)}{\sigma_y^2} \right]. \end{cases}$$

alternatively,

$$\mu_{z|y} = \underbrace{\mu_z}_{\text{prior mean}} + \underbrace{\frac{\omega \sigma_z^2}{\sigma_y^2 + \omega^2 \sigma_z^2} [y - (\omega \mu_z + b)]}_{\text{correction to the mean} \sim y - (\omega \mu_z + b)}$$

(b) Find $p(y)$.

Note that

$$\begin{cases} z \sim \mathcal{N}(\mu_z, \sigma_z^2), \\ y = \omega z + b + \epsilon, \text{ where } \epsilon \sim \mathcal{N}(0, \sigma_y^2). \end{cases}$$

Since z & ϵ are independent, we ~~can~~ expect $\omega z + \epsilon$ and therefore y to be also gaussian.

$$\text{Then } E[y] = \omega \underbrace{E[z]}_{\mu_z} + b + \underbrace{E[\epsilon]}_{0} = \omega \mu_z + b,$$

$$\text{Var}[y] = \omega^2 \underbrace{\text{Var}[z]}_{\sigma_z^2} + \underbrace{\text{Var}[\epsilon]}_{\sigma_y^2} = \omega^2 \sigma_z^2 + \sigma_y^2.$$

$$\text{Thus, } p(y) = \mathcal{N}(y | \omega \mu_z + b, \sigma_y^2 + \omega^2 \sigma_z^2).$$