a. Q: $y = Ax + b + w \qquad w \sim \mathcal{N}(0, Q)$

Find $p(y|x)$.

A: It's Gaussian (linear relationship of Gaussian R.V.'s).

$$\mathbb{E}_w[y|x] = \mathbb{E}[Ax + b + w] = Ax + b + 0$$
$$\mathbb{V}_w[y|x] = \mathbb{V}[Ax + b + w] = 0 + 0 + Q$$
$$\hookrightarrow \underline{p(y|x) = \mathcal{N}(y;\, Ax + b,\, Q)}$$

Tip: Always make sure to write the final answer that is asked for. Otherwise I don't know that you know that a Gaussian is parameterised by the mean and variance...

b. Option I: Find distribution of y after integrating out <u>both</u> x and w.
Still Gaussian (linear ...)

$$\mathbb{E}[y] = \mathbb{E}_{x,w}[Ax + b + w] = A\mathbb{E}(x) + b + 0 = A\mu_x + b$$
$$\mathbb{V}[y] = \mathbb{V}_{x,w}[Ax + b + w]$$
$$= \mathbb{V}[Ax] + \underbrace{\cancel{\mathbb{V}(b)}}_{0,\ \text{const}} + \mathbb{V}[w] \qquad (w, x \text{ independent})$$
$$= A\Sigma_x A^T + Q \qquad (\text{Variance multiplication rule})$$
$$\hookrightarrow p(y) = \mathcal{N}(y;\, A\mu_x + b,\, A\Sigma_x A^T + Q)$$

Option II: Find joint, then marginalise.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} I & 0 \\ A & I \end{bmatrix}\begin{bmatrix} x \\ w \end{bmatrix} + \begin{bmatrix} 0 \\ b \end{bmatrix}$$

$$\hookrightarrow p\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \mathcal{N}\left(\begin{bmatrix} x \\ y \end{bmatrix};\, \begin{bmatrix} \mu_x \\ A\mu_x + b \end{bmatrix},\, \underbrace{\begin{bmatrix} \Sigma_x & \Sigma_x A^T \\ A\Sigma_x & A\Sigma_x A^T + Q \end{bmatrix}}_{\Sigma}\right)$$

$$\Sigma = \begin{bmatrix} I & 0 \\ A & I \end{bmatrix}\begin{bmatrix} \Sigma_x & 0 \\ 0 & Q \end{bmatrix}\begin{bmatrix} I & A \\ 0 & I \end{bmatrix}$$

Marginalisation rule for Gaussians $\Rightarrow$ Choose corresponding mean/variance blocks.

$$\underline{\underline{p(y) = \mathcal{N}(y;\, A\mu_x + b,\, A\Sigma_x A^T + Q)}}$$

Option III: Like option I but different.

$$y = [A\ I]\begin{bmatrix} x \\ w \end{bmatrix} + b$$
$$\mathbb{E}[y] = A\mathbb{E}\left[\begin{bmatrix} x \\ w \end{bmatrix}\right] + b = A\begin{bmatrix} \mu_x \\ 0 \end{bmatrix} + b = A\mu_x + b$$
$$\mathbb{V}[y] = [A\ I]\begin{bmatrix} \Sigma_x & 0 \\ 0 & Q \end{bmatrix}\begin{bmatrix} A^T \\ I \end{bmatrix} = A\Sigma_x A^T + Q$$
$$p(y) = \mathcal{N}(\ldots)$$

c. Same procedure... See lecture.

d. Find joint for $x, y \to$ see option II.

Apply Gaussian conditioning rules.

$$p(x \mid \hat{y}) = \mathcal{N}\big(x;\ \mu_x + \Sigma_x A^T (A\Sigma_x A^T + Q)^{-1} (\hat{y} - (A\mu_x + b)),$$
$$\Sigma_x - \Sigma_x A^T (A\Sigma_x A^T + Q)^{-1} A \Sigma_x \big)$$

Alternative to finding the joint as in option II...

Find covariance $\mathbb{C}(x, y)$ directly.

$$\begin{aligned}
\mathbb{C}[x, y] &= \mathbb{E}_{x,y}[(x-\mu_x)(y - A\mu_x - b)^T] \\
&= \mathbb{E}_{x,w}[(x-\mu_x)(Ax \cancel{+b} - A\mu_x \cancel{-b} + w)^T] \\
&= \mathbb{E}_{x,w}[(x-\mu_x)(x-\mu_x)^T A^T] + \mathbb{E}_{x,w}[(x-\mu_x) w^T] \\
&= \mathbb{V}[x] A^T + \mathbb{E}_x[(x-\mu_x)] \mathbb{E}[w^T] \\
&= \Sigma_x A^T \quad + \quad 0
\end{aligned}$$

N.B: $\mathbb{E}_{x,y}[x \cdot y] = \mathbb{E}_x[x] \cdot \mathbb{E}[y]$ if $p(x,y) = p(x)p(y)$

Proof:
$$\begin{aligned}
\int p(x,y)\, xy\, dx dy &= \int p(x)p(y)\, x\, y\, dx\, dy \\
&= \int p(x)\, x\, dx \cdot \int p(y)\, y\, dy \\
&= \mathbb{E}(x)\, \mathbb{E}[y] \qquad \square
\end{aligned}$$