(1) a) (ii) results in a more accurate estimator for $N > 1$. As N becomes large the MMSE estimator weights the data portion more heavily compared to the prior and the accuracy of the estimator is then determined predominantly by σ^2. Since σ^2 is smaller for (ii), this is the scenario which gives a more accurate estimator.

$$\lim_{N \to \infty} \frac{1}{1/\sigma_A^2 + N/\sigma^2} = \sigma^2/N$$

b) $\beta \to \infty$. This results in an uninformative prior $p(A) = \text{constant}$ for all A. Then

$$p(A|x) = \frac{p(x|A)\,p(A)}{\int p(x|A)\,p(A)\,dA} = \frac{p(x|A)\,\cancel{p(A)}}{\cancel{p(A)}\underbrace{\int p(x|A)\,dA}_{\text{indep of}}}$$

since $p(A)$ constant

Picking the max. of $p(A|x)$, therefore is equal to picking the max of $p(x|A)$ which is also the MLE estimator.

Alternative answer: $|\bar{X}| < \alpha/2$

② a) ~~θ~~

$$E[\theta|x] = E(\theta) + \frac{cov(\theta, x)}{Var(x)}\left(x - E(x)\right)$$

$$E[\theta / x = 3] = 1 + \frac{-0.5}{4}\left(3 - (-1)\right)$$

$$= 1 + \frac{-0.5 \times \cancel{4}}{\cancel{4}}$$

$$= 0.5 \text{ (MMSE)}$$

$\hat{\theta}_{MAP} = \hat{\theta}_{MMSE} = 0.5$ since θ, x jointly Gaussian.

b) in this case $E[\theta | x = 3] = 1 + \frac{0}{4} 4$

$$= 1$$

---

③ a) $\hat{A}_{MAP} = \arg\max \; p(x|A)\,p(A)$

$$p(x|A) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\Sigma(x[n] - A)^2}{2\sigma^2}\right]$$

$$p(x|A)\,p(A) = \frac{0.5}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\Sigma(x[n] - A)^2}{2\sigma^2}\right] \times \left(\delta(A+1) + \delta(A-1)\right)$$

$$p(x|A)p(A) = \begin{cases} q \quad \dfrac{0.5}{\sqrt{2\pi\sigma^2}} \exp\left[-\dfrac{\Sigma(x[n]+1)^2}{2\sigma^2}\right], & A=-1 \\ r \quad \dfrac{0.5}{\sqrt{2\pi\sigma^2}} \exp\left[-\dfrac{\Sigma(x[n]-1)^2}{2\sigma^2}\right], & A=1 \\ 0 & \text{, otherwise} \end{cases}$$

therefore $\hat{A}_{MAP} = \begin{cases} 1 & \text{if } r \geq q \\ -1 & \text{if } r < q \end{cases}$

equivalently $\hat{A}_{MAP} = \begin{cases} 1 & \text{if } \Sigma x[n] \geq 0 \\ -1 & \text{if } \Sigma x[n] < 0 \end{cases}$

b) $p(A|y) = \dfrac{p(y|A)\,p(A)}{p(y)}$ Bayes rule

$= \dfrac{p(y)\,p(A)}{p(y)}$ independence of y & A

$= p(A)$

Therefore $\hat{A}_{MMSE} = E[A|y] = E[A]$

$$= \int_{-\infty}^{\infty} A\left(0.5\,\delta(A+1) + 0.5\,\delta(A-1)\right) dA$$

$= -1 \times 0.5 + 1 \times 0.5 = 0$

(4) a) $\hat{\theta}[0] = \dfrac{<\theta, x[0]>}{<x[0], x[0]>^2} x[0]$

$<\theta, x[0]> = E\left[\theta x[0]\right] = E\left[\theta(\theta \times 0 + w[0])\right]$

$= E\left[\theta w[0]\right] = E[\theta] E[w[0]]$

$= 0$

Therefore $\hat{\theta}[0] = 0$

Not surprising since a data point with $n=0$ can't give us any information about the slope of a line passing through the origin.

b) $\tilde{x}[1] = x[1] - \hat{x}[1|0]$

$\hat{x}[1|0] = \dfrac{<x[1], x[0]>}{<x[0], x[0]>^2} x[0]$

$= \dfrac{E\left[x[1]x[0]\right]}{E\left[x^2[0]\right]} x[0] = \dfrac{E\left[(\theta \times 1 + w[1])(\theta \times 0 + w[0])\right]}{E\left[(\theta \times 0 + w[0])^2\right]} x[0]$

using independence & $E[\theta] = 0$, $E[w] = 0 \Rightarrow 0$ $\Rightarrow 0$

$= \dfrac{E\left[\theta w[0]\right] + E\left[w[1]w[0]\right]}{E[\theta^2] + E[w[0]]^2} x[0]$

$= 0$

Therefore $\tilde{x}[1] = x[1]$

$$\hat{\theta}[1] = \hat{\theta}[0] + \Delta\hat{\theta}[1|0] = \Delta\hat{\theta}[1|0]$$

$$\Delta\hat{\theta}[1|0] = \frac{\langle \theta, \tilde{x}[1] \rangle}{\langle \tilde{x}[1], \tilde{x}[1] \rangle^2}\, \tilde{x}[1]$$

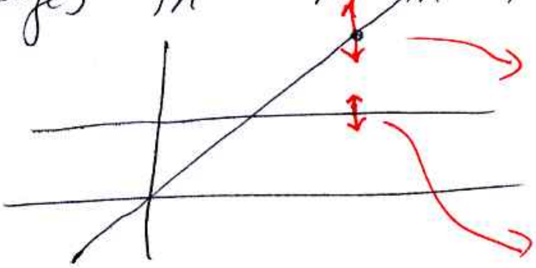
$$= \frac{\langle \theta, x[1] \rangle}{\langle x[1], x[1] \rangle^2}\, x[1]$$

$$= \frac{E(\theta x[1])}{E(x^2[1])}\, x[1] = \frac{E(\theta \cdot (\theta + w[1]))}{E[(\theta + w[1])^2]}\, x$$

$$= \frac{E(\theta^2) + \cancel{E(\theta)E(w[1])}}{E(\theta^2) + E[w^2[1]] + \cancel{2E(\theta)E(w[1])}}\, x[1]$$

$$= \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma^2}\, x[1]$$

c) $x[n]$ is much more sensitive to chang
in θ for large n in the model
$x[n] = \theta n + w[n]$ compared to
the sensitivity of $x[n]$ with respect to
changes in A in the model $x[n] = A + w$



changing the slope by $\Delta\theta$
changes $x[n]$ by $(\Delta\theta)n$

~~At~~ here $x[n]$ changes by ΔA