Question 2

for Bayesian Linear Regression model, with likelihood $p(y|x,w) = \mathcal{N}(w^T x, \beta^{-1})$
and prior $p(w) = \text{Normal}(0, \lambda^{-1} I)$

The predictive posterior is $p(y_*|x_*) = N(\mu_N^T x_*, \beta^{-1} + x_*^T \Sigma_N x_*)$
$= \mathcal{N}(\mu_N^T x_*, \sigma_N^2(x_*))$

where $\mu_N(x_*) = \Sigma\left(\beta \sum_{n=1}^N y_n x_n\right)$.

and $\sigma_N^2(x_*) = \beta^{-1} + x_*^T \Sigma_N x_*$; $\Sigma_N = \left(\beta \sum_{n=1}^N x_n x_n^T + \lambda I\right)^{-1}$

We have to prove that $\sigma_{N+1}^2(x_*) \leq \sigma_N^2(x_*)$

Observe that $\sigma_N^2 - \sigma_{N+1}^2 = x_*^T\left[\left(\beta \sum_{n=1}^N x_n x_n^T + \lambda I\right)^{-1} - \left(\beta \sum_{n=1}^{N+1} x_n x_n^T + \lambda I\right)^{-1}\right] x_*$

Take $\beta \sum_{n=1}^N x_n x_n^T + \lambda I = M$.

Then, $\sigma_N^2 - \sigma_{N+1}^2 = x_*^T\left[\left(M^{-1}\right) - \left(M + \beta x_{N+1} x_{N+1}^T\right)^{-1}\right] x_*$

We use the matrix identity $(M + vv^T)^{-1} = M^{-1} - \dfrac{(M^{-1}v)(v^T M^{-1})}{1 + v^T M^{-1} v}$

take $v = \sqrt{\beta}\, x_{N+1}$

Then, $\sigma_N^2 - \sigma_{N+1}^2 = x_*^T\left[M^{-1} - (M + vv^T)^{-1}\right] x_*$

$= x_*^T\left[\dfrac{(M^{-1}v)(v^T M^{-1})}{1 + v^T M^{-1} v}\right] x_*$. Since we see that M is a symmetric matrix (Since I is symmetric, $\Sigma x_n x_n^T$ is symmetric so their inverse is symmetric),

$(M^{-1}v)^T = v^T M^{-T} = v^T M^{-1}$

and $v^T M^{-1} v > 0$ [psd matrix M^{-1}].

$\sigma_N^2 - \sigma_{N+1}^2 = \dfrac{(x_*^T (M^{-1}v)).((v^T M^{-1}) x_*)}{1 + v^T M^{-1} v} = \dfrac{\text{non-negative } \cancel{\text{positive}} \text{ value}}{\text{positive value}} \geq 0.$

Thus, $\sigma_N^2 \geq \sigma_{N+1}^2$.

Hence, proved