1. **请证明Lecture 3，Page 14末尾处介绍的Gaussain regression的beta的分布的均值恰好等价于Ridge regression的解——教材上没有，课堂上讲过。**

   Gaussain regression的beta的分布的均值为：

$$\hat{\beta} = \tau^2 \mathbf{X}^\top \left(\tau^2 \mathbf{X}\mathbf{X}^\top + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{Y}$$

   左乘 $\left(\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I}\right)$ 得到

$$\begin{aligned}
\left(\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I}\right)\hat{\beta} &= \left(\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I}\right)\tau^2 \mathbf{X}^\top \left(\tau^2 \mathbf{X}\mathbf{X}^\top + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{Y} \\
&= \left(\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I}\right)\mathbf{X}^\top \left(\mathbf{X}\mathbf{X}^\top + \frac{\sigma^2}{\tau^2}\mathbf{I}\right)^{-1} \mathbf{Y} \\
&= \left(\mathbf{X}^\top \mathbf{X}\mathbf{X}^\top + \lambda \mathbf{X}^\top\right)\left(\mathbf{X}\mathbf{X}^\top + \frac{\sigma^2}{\tau^2}\mathbf{I}\right)^{-1} \mathbf{Y} \\
&= \mathbf{X}^\top \left(\mathbf{X}\mathbf{X}^\top + \lambda \mathbf{I}\right)\left(\mathbf{X}\mathbf{X}^\top + \lambda \mathbf{I}\right)^{-1} \mathbf{Y} \\
&= \mathbf{X}^\top \mathbf{Y}
\end{aligned}$$

   因此

$$\hat{\beta} = \left(\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I}\right)^{-1} \mathbf{X}^\top \mathbf{Y}$$

   即为Ridge regression的解

2. **请证明Section 1.8所介绍的Bayesian regression，即如何从一些对beta和epsilon的先验假设，证明出ridge regrssion的loss形式（具体证明过程写在了Page 15中的框内）。**

$$\begin{gathered}
\beta \sim N\left(0, \tau^2 I_p\right) \\
Y - X\beta = \epsilon \sim N\left(0, \sigma^2 I_n\right)
\end{gathered}$$

$$\begin{aligned}
p\left(\beta \mid Y, X\right) &\propto p\left(\beta\right) p\left(Y \mid X, \beta\right) \\
&\propto \exp\left(-\frac{1}{2\tau^2}|\beta|^2\right)\exp\left(-\frac{1}{2\sigma^2}|Y - X\beta|^2\right) \\
&= \exp\left(-\frac{1}{2}\left[\frac{1}{\sigma^2}|Y - X\beta|^2 + \frac{1}{\tau^2}|\beta|^2\right]\right) \\
&= \exp\left(-\frac{1}{2\sigma^2}\left[|Y - X\beta|^2 + \lambda|\beta|^2\right]\right)
\end{aligned}$$

   其中

$$\lambda = \frac{\sigma^2}{\tau^2}$$