Assignment - 3

Problem 1: Show that the ridge regression estimate is the mean (and mode) of the posterior distribution, under a gaussian prior $\beta \sim N(0, \tau^2 I)$, and Gaussian sampling model $y \sim N(X\beta, \sigma^2 I)$. Find the relationship between the regularization parameter λ in the ridge formula, and the variance τ^2 and σ^2.

Solution: The likelihood : $y \sim N(X\beta, \sigma^2 I)$. Prior: $\beta \sim N(0, \tau^2 I)$
Assume that our input data x is already centered so that we can ignore the intercept term β_0.
Posterior distribution is :

$$p(\beta | y, X) = \frac{1}{Z} p(y|\beta, X)\, p(\beta)$$

where $Z = Z(y, X)$

Log poseterior distribution :

$$p(\beta|y, X) = \frac{1}{Z} p(y|\beta, x)\, p(\beta)$$

$$= \frac{1}{Z} * \frac{1}{(2\pi)^{P/2}\sigma^P} e^{\left\{-\frac{1}{2\sigma^2}(y - X\beta)^t(y - X\beta)\right\}} * \frac{1}{(2\pi)^{P/2}\tau} e^{\left\{-\frac{\beta^T\beta}{2\tau^2}\right\}}$$

$$\log p(\beta|y, X) = -\log Z - k - \frac{1}{2\sigma^2}(y - X\beta)^t(y - X\beta) - \frac{\beta^t\beta}{2\tau^2}$$

Taking derivative & setting to 0, will give

$$\hat{\beta} = \left(X^t X + \frac{\sigma^2}{\tau^2} I\right)^{-1} X^t y$$

Letting $\lambda = \frac{\sigma^2}{\tau^2}$, we see equivalence, & its shows that $P(\beta|y, X)$ is Gaussian and its mean and mode coincide.

Now, $$\Sigma^{-1} = \frac{1}{\sigma^2}\left(X^T X + \frac{\sigma^2}{\tau^2} I\right)$$

which gives $\hat{\beta} = \frac{1}{\sigma^2} \Sigma X^T y$ which is the mean.

Problem 4: (a) Estimate the probability that a student who studies for 40 h and has an undergrad GPA of 3.5 gets an A in class.

Sol:- $$p(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2)}$$

where, x_1 = hours studied x_2 = undergrad gpA

$$p(x) = \frac{e^{-6+0.05x_1+x_2}}{1+e^{-6+0.05x_1+x_2}} = \frac{e^{-6+0.05*40+3.5}}{1+e^{-6+0.05*40+3.5}} = \boxed{37.75}$$

(b) How many hours would student in part (a) need to study to have 50% chance of getting an A in class?

Solution: $p(x) = \frac{e^{-6+0.05x_1+3.5}}{1+e^{-6+0.05x_1+3.5}}$ substituting $P(A) = 0.5$

$x_2 = 3.5$ $x_1 = ?$

$$0.5 = \frac{e^{-6+0.05x_1+3.5}}{1+e^{-6+0.05x_1+3.5}} \Rightarrow \boxed{x_1 = 50 \text{ hours}}$$