## 2.2 Theoretical Question On Ridge Regression [15%]Lin

Line A could not be produced by ridge regression. Low λ > 0 would mean that our model is based on the ordinary regression, while with large λ regression coefficients grow to 0. This prediction could be possible with negative λ, but since λ > 0 this prediction could not be produced.

Line B seems as if it is a regular linear regression prediction. Therefore, the alpha parameter should be 0, meaning the prediction rule could be produced by ridge regression with a low λ value.

Line C could be produced if the alpha is sufficiently large. With large λ the impact of shrinkage grows and regression coefficients grow towards zero.

Line D could not be produced by ridge regression. We only penalize regression coefficients, so intercept should be the same. With prediction D, it appears that the intercept is also different.

## 3.2 On Regularization in Logistic Regression [15%]

When we try to increase C to regularize θ2 our model relies less on the values of x2 as it shrinks towards 0, thus making the prediction line more vertical.

L2: Could not be produced by regularizing θ2, as it is more horizontal.

L3: Could be produced with a sufficient C, as it tends to discriminate against θ2 and making the line more vertical.

L4: Could not be produced by regularization. While it is vertical, which denotes small weights of θ2, it would not be the most optimal solution. From the looks of it, it would have negative weight for θ2 and since our regularization takes into account square values of the weights, it would most likely choose the line between L3 and X2 axis(as it would produce a smaller error).