**Name: Lianhan Huang 3700459; Qu Wang 3700666**

1. **Simple Linear Regression**

**1.1**

Mean x = 2

Mean y = 170/3

Cov(x,y) = (1-2)\*(50-170/3) + (2-2)\*(55-170/3) + (3-2)\*(65-170/3) = 15

Var(x) = 2

β1 = Cov / Var = 7.5

β0+ϵ = 41.67

^y = 41.6 + 7.5x

**1.2**

X=5 ^y = 79.1

X=10 ^y = 116.6

**1.3**

RSS = 0.83\*\*2 + 1.67\*\*2 + 0.83\*\*2 = 2.7889

**1.4**

L1: RSS = 0.83\*\*2 + λ \* 7.5\*\*2+ 1.67\*\*2 + λ \* 7.5\*\*2 + 0.83\*\*2 + λ \* 7.5\*\*2

L2: RSS = 0.83\*\*2 + λ \* 7.5+ 1.67\*\*2 + λ \* 7.5 + 0.83\*\*2 + λ \* 7.5

**1.5**

Regularization is to avoid overfitting by punishing extreme value of β which occurs where there are extreme noise and avoid the slope coefficient to be too large.

1. **Multiple linear regression**

**2.1**

X=[[1,1,1], [1,4,1], [1,6,0], [1,8,2], [1,10,1]]

XT=[[1,1,1,1,1], [1,4,6,8,10], [1,1,0,2,1]]

Y=[52,63,62,91,75]

XT \* X = [[ 5 29 5]

[ 29 217 31]

[ 5 31 7]]

XT \* y = [ 343 2154 372]

(XT\*X)-1 = [[ 1.19230769 -0.1025641 -0.3974359 ]

[-0.1025641 0.02136752 -0.02136752]

[-0.3974359 -0.02136752 0.52136752]]

β = [40.19230769 2.8974359 11.6025641 ]

y = 40.19 + 2.897x1 + 11.6x2

**2.2**

x1 = 5, x2 = 1 y = 66.275

**2.3**

^y = [54.687, 63.378, 57.572, 86.566, 80.76]

RSS = 2.687\*\*2 + 0.378\*\*2 + 4.428\*\*2 + 4.434\*\*2 + 5.76\*\*2 = 79.8

**2.4**

RSS = (2.687\*\*2 + λ \* (2.897\*\*2 + 11:6\*\*2)) +

(0.378\*\*2 + λ \* (2.897\*\*2 + 11:6\*\*2)) +

(4.428\*\*2 + λ \* (2.897\*\*2 + 11:6\*\*2)) +

(4.434\*\*2 + λ \* (2.897\*\*2 + 11:6\*\*2)) +

(5.76\*\*2 + λ \* (2.897\*\*2 + 11:6\*\*2))

Large λ:

Benefit - strong regularization and reduce the magnitude of the coefficients and the complexity of the model.

Drawbacks - too much punishment may lead to under fitting.

Small λ:

Benefit - allow the model to realize the complex relation between variables.

Drawbacks - may not work so efficiently and can not avoid overfitting or simplify the model efficiently.