Homework 4

Jack Kai Lim

May 7, 2022

Problem 1

a)

To get $H'(x-\bar{x})$ we can use the fact that $b = \bar{y} - a\bar{x}$ where a is the slope of the linear equation and we get,

$$\begin{aligned} y &= ax + \bar{y} - a\bar{x} \\ y - \bar{y} &= a(x - \bar{x}) \end{aligned}$$

Therefore we get that $y - \bar{y} = H'(x - \bar{x}) = a(x - \bar{x})$.
For $G''(y - \bar{y})$ we are able to do something similar, we set $d = \bar{x} - c\bar{y}$ and we can get,

$$\begin{aligned} x &= cy + \bar{x} - c\bar{y} \\ x - \bar{x} &= c(y - \bar{y}) \end{aligned}$$

And finally we get $x - \bar{x} = G''(y - \bar{y}) = c(y - \bar{y})$

b)

Now we try to prove that the inverse of G'' is equal to H', first we get the inverse of G'' which we will label G'

$$\begin{aligned} x - \bar{x} &= c(y - \bar{y}) \\ y - \bar{y} &= \frac{1}{c}(x - \bar{x}) \\ &= G'(x - \bar{x}) \end{aligned}$$

Therefore to prove this conjecture we have to prove that $a = \frac{1}{c}$ we have that $a = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$ and $c = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$

$$\begin{aligned}
a &= \frac{1}{c} \\
ac &= 1 \\
&= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2} \\
&= \frac{\sum_{i=1}^{n} \sum_{j=1}^{n}(x_i - \bar{x})^2(y_i - \bar{y})^2}{\sum_{i=1}^{n} \sum_{j=1}^{n}(x_i - \bar{x})^2(y_i - \bar{y})^2} \\
&= 1
\end{aligned}$$

Therefore the conjecture is true and proven.

Probem 2

a)

There are 2 issues with the existing approach, firstly we will be unable to obtain X^{-1} as the matrix isn't a square matrix and therefore we cannot calculate the determinant and therefore the inverse. Another issue if by some miracle we are able to inverse X is there issue where the dimensions of X do not match up with $\vec{v}$ making matrix multiplication impossible.

b)

As we have that $\vec{w} = X^T\vec{z}$ and $X\vec{w} = \vec{v}$ we can get the equation,

$$XX^T\vec{z} = \vec{v}$$

Here we then have to have that XX^T is full rank in order for it to be invertible, then we can solve equation 1,

$$\begin{aligned}\vec{w} &= X^T\vec{z} \\ &= X^T(XX^T)^{-1}\vec{v}\end{aligned}$$

Now in order to prove if $\vec{w} = X^{-1}\vec{v}$ is equal to $\vec{w} = X^T\vec{z}$ when the matrix is of dimension $d \times d$ it is sufficient to prove that $X^T(XX^T)^{-1} = X^{-1}$

$$\begin{aligned}X^T(XX^T)^{-1} &= \frac{X^T}{XX^T} \\ &= \frac{1}{X} \\ &= X^{-1}\end{aligned}$$

c)

Let $A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 0 & 1 & 1 \end{pmatrix}$,

$$\begin{aligned}
\vec{w} &= A^T (AA^T)^{-1} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \left(\begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\
&= \frac{1}{27} \begin{pmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\
&= \begin{pmatrix} \frac{2}{3} \\ 0 \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}
\end{aligned}$$

Problem 3

First we calculate the value of $\vec{v}^T\vec{w}$,

$$\vec{v}^T\vec{w} = \vec{v} \cdot \vec{w} = \{v_1w_1 + v_2w_2 + ... + v_nw_n\}$$

Then when calculating the derivatives of $\frac{d}{d\vec{w}}$ we have to split it to each of it's individual components, therefore we get,

$$\begin{pmatrix} \frac{d}{d\vec{w_1}}\{v_1w_1 + v_2w_2 + ... + v_nw_n\} \\ \vdots \\ \frac{d}{d\vec{w_n}}\{v_1w_1 + v_2w_2 + ... + v_nw_n\} \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \vec{v}$$

Problem 4

a)

First we can transform $X^T = \{\vec{r_1}, ..., \vec{r_m}\}$ and $X = \{\vec{r_1}^T, ..., \vec{r_m}^T\}$ therefore we get that,

$$\vec{w}^T\{\vec{r_1}, ..., \vec{r_m}\}\{\vec{r_1}^T, ..., \vec{r_m}^T\}\vec{w} = \{\vec{w}^T\vec{r_1}, ..., \vec{w}^T\vec{r_m}\}\{\vec{r_1}^T\vec{w}, ..., \vec{r_m}^T\vec{w}\}$$

$$= \{\vec{w}\cdot\vec{r_1}, ..., \vec{w}\cdot\vec{r_m}\}\{\vec{r_1}\cdot\vec{w}, ..., \vec{r_m}\cdot\vec{w}\} = \sum_{i=1}^{m}(\vec{r_i}\cdot\vec{w})^2 = \sum_{i=1}^{m}(\vec{r_i}^T\vec{w})^2$$

Now we can differentiate the outcome and get the equality we are looking for,

$$\begin{aligned}\frac{d}{d\vec{w}}(\vec{w}^TX^TX\vec{w}) &= \sum_{i=1}^{m}2(\vec{r_i}^T\vec{w})\frac{d}{d\vec{w}}(\vec{r_i}^T\vec{w}) \\ &= \sum_{i=1}^{m}2(\vec{r_i}^T\vec{w})\vec{r_i}\end{aligned}$$

b)

Now we have to prove that, $2X^TX\vec{w} = \sum_{i=1}^{m}2(\vec{r_i}^T\vec{w})\vec{r_i}$ and let $A = X^TX$

$$\begin{aligned}2X^TX\vec{w} &= \sum_{i=1}^{m}2(\vec{r_i}^T\vec{w})\vec{r_i} \\ \sum_{i=1}^{m}\vec{r_i}(\vec{r_i}\cdot\vec{w}) &= \sum_{i=1}^{T}(\vec{r_i}\cdot\vec{w})\vec{r_i}\end{aligned}$$

Now we know that $\vec{r_i}\vec{w}$ gives a scalar, so we let, $a = \vec{r_i}\vec{w}$,

$$\begin{aligned}\sum_{i=1}^{m}\vec{r_i}a &= \sum_{i=1}^{m}a\vec{r_i} \\ \sum_{i=1}^{m}a\vec{r_i} &= \sum_{i=1}^{m}a\vec{r_i}\end{aligned}$$

Therefore this proves that $2X^TX\vec{w} = \sum_{i=1}^{m}2(\vec{r_i}^T\vec{w})\vec{r_i}$ is true and therefore we have proven the given relationship.

Problem 5

a)

With standardizing, I would say that the asking price has the highest correlation to the sale price as the Asking prices value is most closely relate to the Sale price. Where for example, House 1 has the same number of bedrooms and only one less bathroom than House 2. And House 2 also only has 1.33 times more square feet than House 1. And as for year, the year for House 2 is 78 years older than House 1. All of which do not have anything close to a 2 times differences like the Asking price therefore it will have the largest weight.

For no standardizing, I would expect Bedrooms to hold the largest weight as for 1 unit of change in bedroom in the 5 given values, we see the largest average change in Sale Price.

b)

If we use multiple linear regression on the data we will get a linear equation similar to,

$$SalesPrice = \alpha(Year) + \beta(SquareFeet) + \gamma(Bedrooms) + \delta(Bathrooms) + \epsilon(AskingPrice)$$

If we change Year to age, where $Year = 2020 - age$ we can get the weight for age in terms of alpha,

$$\alpha(2020 - age) = \alpha(2020) - \alpha(age)$$

Therefore the weight associated with age will be $\alpha(2020 - age)$

c)

It will not improve the accuracy of the prediction as the number of rooms have already been accounted for therefore adding the new variable room will not make the prediction better. To be accurate it should not change the prediction at all as breaking down rooms into bedroom and bathroom is more deep of a correlation than the total number of rooms.