CS 5805: Spring 2025
Homework 1 Solution Sketches
by Ishtiaque Ahmed Khan

Question 1 (5 points):

The hat matrix is defined as :
$H = X(X^T X)^{-1}X^T$
We need to show that H is symmetric, ie $H^T = H$.

Now,
$H = X(X^T X)^{-1}X^T$
$H^T = (X(X^T X)^{-1}X^T)^T$
$= [\{X(X^T X)^{-1}\} X^T]^T$
[Let's apply the property of transpose of Matrix product: $(AB)^T = B^T A^T$]
$= (X^T)^T \{X(X^T X)^{-1}\}^T$
$= X \{(X^T X)^{-1}\}^T X^T$
We can show that $(X^T X)$ is symmetric: $(X^T X)^T = X^T (X^T)^T = X^T X$
And, the inverse of a symmetric matrix is also symmetric. Hence, $\{(X^T X)^{-1}\}^T = (X^T X)^{-1}$
Going back to the previous equation:

$H^T = X (X^T X)^{-1} X^T = H$

Hence, hat matrix H is a symmetric matrix.

Question 2 (5 points):

To prove that the hat matrix is Idempotent, we need to show that HH = H
Now,
$H = X(X^T X)^{-1}X^T$
$H H = \{X(X^T X)^{-1}X^T\} \{ X(X^T X)^{-1}X^T\}$
$= X \{(X^T X)^{-1}X^TX\} \{(X^T X)^{-1}X^T\}$ [property: (AB)(CD) = A(BC)D]
$= X \{(X^T X)^{-1} (X^TX)\} \{(X^T X)^{-1}X^T\}$
$= X I \{(X^T X)^{-1}X^T\}$ [property: $(A^{-1}A) = I$]
$= X(X^T X)^{-1}X^T$
$HH = X(X^T X)^{-1}X^T= H$

Hence, the hat matrix H is idempotent.

Ques 3 (5x2 points):

For each question: 2 points if answer is correct with explanation, zero otherwise

a)

$$y = \sum_{i=0}^{n} a_i x^i$$

We can consider the higher order x-terms such as x^2, x^3 etc. as separate new variables which are transforms of the original variable x. We can represent the equation like this:

$y = a_0 + a_1x^1 + a_2x^2 + a_3x^3 \ldots + a_nx^n$

We can consider these terms as new transformed variables.
$y = a_0 + a_1x + a_2p + a_3q \ldots + a_nr$

This now simply becomes a process of finding the coefficients as normal but with some new extra independent variables. Instead of just having x, we now have variables p, q etc. which are actually x^2, x^3 etc. Hence, we can learn a linear regression model for this.

b)

$$y = ax + b \cdot sin(x) + c \cdot log(x) + d$$

Similar to the logic for (a), we can consider the non-linear x-terms as separate variables. Hence, we can also learn a regression model for this to estimate the values of a, b, c and d.

c)

$$y = ae^{(bx)}$$

Here, we can apply natural logarithm on both sides.
ln(y) = ln(a * exp(bx))

We get,
ln(y) = ln(a) + ln(exp(bx))
ln(y) = ln(a) + bx
Now, we can define new variables,
y' = ln(y)
a' = ln(a)
Then the equation becomes:
y' = a' + bx
This equation is now in a linear form in terms of a’ and b. After doing linear regression, we can estimate the value of b directly and from a’, we can get a = exp(a’).

d)

$$y = ax + bx + c$$

Here the relationship is a simple linear relationship.
We can represent the equation like this: y = (a+b)x + c and then estimate the values of (a+b) and c.
However, we cannot find the values of a and b separately. Considering ax and bx as separate variables will lead to rank deficiency and X^TX not being invertible.

e)

$$y = (ax + b)/(cx + d)$$

Here, the relationship is non-linear and we cannot reorganize it to get a linear relationship in terms of the coefficients. Hence, we cannot accurately estimate this curve using a linear regression model.

Ques 4, 5 and 6 : [Notebook]