CS 3320 – Numerical Software

Module 10 Homework Key

1. (10 pt) Use least-squares regression to fit a straight line to the data below.

x	0	2	4	6	9	11	12	15	17	19
y	5	6	7	6	9	8	8	10	12	12

Along with the slope and intercept, compute the coefficient of determination, R^2.

	x	y	x*y	x*x	y-yPrection	(y-yPrection)^2	(y-yAve)	(y-yAve)^2
	0	5	0	0	0.11188251	0.012517696	-3.3	10.89
	2	6	12	4	0.39359146	0.154914234	-2.3	5.29
	4	7	28	16	0.6753004	0.456030631	-1.3	1.69
	6	6	36	36	-1.0429907	1.087829505	-2.3	5.29
	9	9	81	81	0.87957276	0.773648247	0.7	0.49
	11	8	88	121	-0.8387183	0.703448372	-0.3	0.09
	12	8	96	144	-1.1978638	1.434877727	-0.3	0.09
	15	10	150	225	-0.2753004	0.075790311	1.7	2.89
	17	12	204	289	1.00640854	1.012858159	3.7	13.69
	19	12	228	361	0.28811749	0.083011688	3.7	13.69
Sum	95	83	923	1277	SSE	5.794926569	SST	54.1
Average		8.3					R2	0.89288491
A	10	95	b	83	Regreesion	4.88811749		
	95	1277		923		0.359145527		

An excel spreadsheet was created for this problem. On the spreadsheet, I calculated $\sum x_i = 95, \sum y_i = 83, \sum x_i^2 = 1277$, and $\sum x_i y_i = 923$. Using this entries, A and b were formulated as:

$$\begin{bmatrix} 10 & 95 \\ 95 & 1277 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 83 \\ 923 \end{bmatrix}.$$

Solving this matrix equation, we have $a_0 = 4.888, a_1 = 0.359$. Our regression result is $y = 0.359x + 4.888$.

Using this model, I can calculate $y - \hat{y}$. $SS_T = \sum (y_i - \bar{y})^2 = 54.1$ and $SS_E = \sum (y_i - \hat{y}_i)^2 = 5.7949$.

$$R^2 = 1 - \frac{SS_E}{SS_T} = 0.8929$$

2. (10 pt) On average the surface area, *A*, of a human is related to weight, *W*, and height, *H*. Measurements for several individuals of height 180 cm and different weights (kg), give values of area (m^2) in the following table.

***W* (kg)**	70	75	77	80	82	84	87	90
***A* (m^2)**	2.10	2.12	2.15	2.20	2.22	2.23	2.26	2.30

Show that a power law, $A = aW^b$, fits these data reasonably well. Present plots of data along with the model line. Predict what the surface area is for a 95-kg person.

Linearization: $\ln(A) = \ln(aW^b)$ or $\ln(A) = \ln(a) + b \cdot \ln(W)$.
Calculating ln(A) and ln(W) before we use the LSE linear regression.

W	A	ln(W)	ln(A)	ln(W)*ln(W)	ln(W)*ln(A)			
70	2.1	4.24849524	0.74193734	18.0497118	3.15211728			
75	2.12	4.31748811	0.75141609	18.6407036	3.24423003			
77	2.15	4.34380542	0.76546784	18.8686455	3.32504336			
80	2.2	4.38202663	0.78845736	19.2021574	3.45504115			
82	2.22	4.40671925	0.7975072	19.4191745	3.51439031			
84	2.23	4.4308168	0.80200159	19.6321375	3.5535221			
87	2.26	4.46590812	0.81536481	19.9443353	3.64134434			
90	2.3	4.49980967	0.83290912	20.2482871	3.74793253			
	sum	35.0950692	6.29506135	154.005153	27.6336211			
	A	8	35.0950692	b	6.29506135	Linearize	-0.8797444	c0
		35.0950692	154.005153		27.6336211	Model	0.3799114	c1
	Original	a=exp(c0)	0.41488893		Weight	Prediction		
	model	b=c1	0.3799114		95	2.34040807		

Linearized model: $\ln(A) = -0.8797 + 0.3799\ln(W)$

Power Law Model: $A = e^{-0.8797}e^{0.3799\ln(W)} = 0.4149W^{0.3799}$

For $W = 95$, the predicted surface area is $A = 0.4149 \cdot 95^{0.3799} = 2.34\ m^2$.

Python Solution 1:
```
import numpy as np
import math

W=np.array([70, 75, 77, 80, 82, 84, 87, 90])
A=np.array([2.1, 2.12, 2.15, 2.2, 2.22, 2.23, 2.26, 2.3])

lnW = np.log(W)
lnA = np.log(A)

lFit = np.polyfit(lnW, lnA, 1)
a=math.exp(lFit[1])
b=lFit[0]

A95 = a*math.pow(95,b)

print(a, b, A95)
a=0.41488892 b=0.37991140 A95=2.3404080703850525
```

Python Solution 2:
```
import numpy as np
import math

W=np.array([70, 75, 77, 80, 82, 84, 87, 90])
A=np.array([2.1, 2.12, 2.15, 2.2, 2.22, 2.23, 2.26, 2.3])

lnW = np.log(W)
lnA = np.log(A)
```

```
matA=np.vstack((lnW, np.power(lnW,0))).transpose()
RHS=np.matmul(matA.transpose(),lnA)

lFit =
np.matmul(np.linalg.inv(np.matmul(matA.transpose(),matA)),RHS)

a=math.exp(lFit[1])
b=lFit[0]

A95 = a*math.pow(95,b)
print(a, b, A95)
a=0.4148889255 b=0.37991140 A95=2.340408070385182
```

3. (10 pt) Fit an exponential model to

x	0.4	0.8	1.2	1.6	2	2.3
y	800	985	1490	1950	2850	3600

Exponential Fit: $y = ae^{bx}$.

Linearization: $\ln(y) = \ln(ae^{bx}) = \ln(a) + bx$

	x	y	ln(y)	x*x	ln(W)*ln(A)		
	0.4	800	6.68461173	0.16	2.67384469		
	0.8	985	6.89264164	0.64	5.51411331		
	1.2	1490	7.3065314	1.44	8.76783768		
	1.6	1950	7.57558465	2.56	12.1209354		
	2	2850	7.95507427	4	15.9101485		
	2.3	3600	8.18868912	5.29	18.833985		
sum	8.3	sum	44.6031328	14.09	63.8208647		
A	6	8.3	b	44.6031328	Linearize	6.30958241	c0
	8.3	14.09		63.8208647	Model	0.81272751	c1
Original	a=exp(c0)	549.815304		x	Prediction		
model	b=c1	0.81272751		2	2793.46331		

The linearized model is $\ln(y) = 6.3096 + 0.81272x$

The exponential model is
$y = e^{6.3096+0.81272x} = e^{6.3096} \cdot e^{0.81272x} = 549.815e^{0.81272x}$

Python Solution:

```
import numpy as np
import math

x=np.array([0.4, 0.8, 1.2, 1.6, 2, 2.3])
y=np.array([800, 985, 1490, 1950,2850, 3600])

lnY = np.log(y)

matA=np.vstack((x, np.power(x,0))).transpose()
```

```
RHS=np.matmul(matA.transpose(),lnY)

lFit =
np.matmul(np.linalg.inv(np.matmul(matA.transpose(),matA)),RHS)

a=math.exp(lFit[1])
b=lFit[0]

y2 = a*math.exp(b*2)

print(a, b, y2)
a=549.815304 b=0.81272751 y(2)=2793.463311758146
```

4. (10 pt) Find a 3rd order polynomial to fit the following data. Plot the data with your model curve. Tabulate the residual of the predict y-value.

x	3	4	5	7	8	9	11	12
y	1.6	3.6	4.4	3.4	2.2	2.8	3.8	4.6

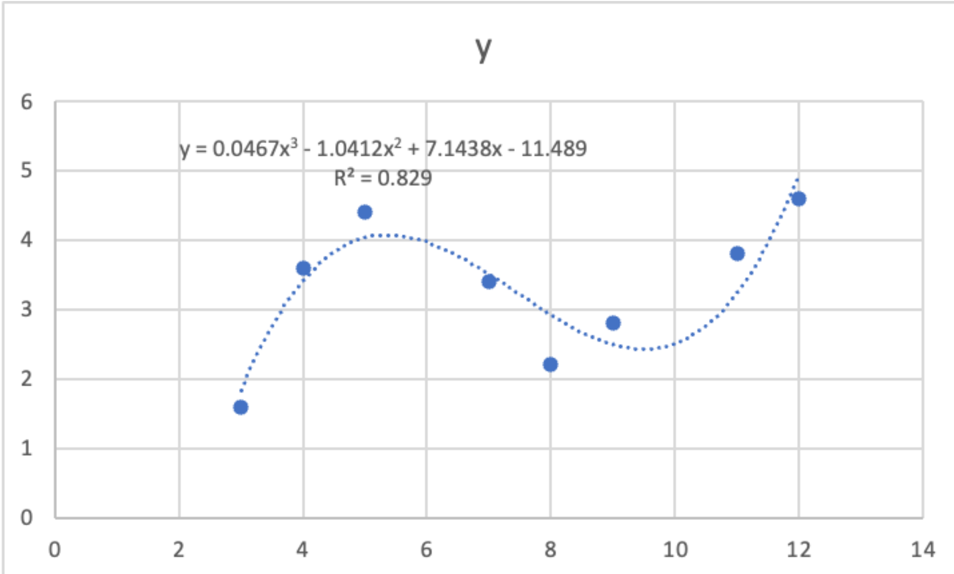
3rd Order Polynomial Fit: $y = a_3x^3 + a_2x^2 + a_1x + a_0$

	x	y	x^2	x^3	x^4	x^5	x^6	xy	x^2*y	x^3*y		yModel	yDiff
	3	1.6	9	27	81	243	729	4.8	14.4	43.2		1.83213464	-0.2321346
	4	3.6	16	64	256	1024	4096	14.4	57.6	230.4		3.41451602	0.18548398
	5	4.4	25	125	625	3125	15625	22	110	550		4.03470796	0.36529204
	7	3.4	49	343	2401	16807	117649	23.8	166.6	1166.2		3.5087479	-0.1087479
	8	2.2	64	512	4096	32768	262144	17.6	140.8	1126.4		2.92270809	-0.7227081
	9	2.8	81	729	6561	59049	531441	25.2	226.8	2041.2		2.49470324	0.30529676
	11	3.8	121	1331	14641	161051	1771561	41.8	459.8	5057.8		3.23302278	0.56697722
	12	4.6	144	1728	20736	248832	2985984	55.2	662.4	7948.8		4.95945937	-0.3594594
sum	59	26.4	509	4859	49397	522899	5689229	204.8	1838.4	18164			
A	8	59	509	4859	b	26.4	Model	-11.488707					
	59	509	4859	49397		204.8		7.14381722					
	509	4859	49397	522899		1838.4		-1.0412069					
	4859	49397	522899	5689229		18164		0.04667602					

According to spreadsheet above: the regression model is:
$y = 0.04668x^3 - 1.0412x^2 + 7.1438x - 11.4887$

The predicted the values and the residual are shown in the last two columns.

The plot:


y
y = 0.0467x³ - 1.0412x² + 7.1438x - 11.489
R² = 0.829
6
5
4
3
2
1
0
0
2
4
6
8
10
12
14