1)

1 - 3.1.1 a)

$f(x) = \|x\|^2$, $h(x) = \sum_{i=1}^{n} x_i - 1$

$L(x,\lambda) = \|x\|^2 + \lambda\left(\sum_{i=1}^{n} x_i - 1\right)$

$= x_1^2 + x_2^2 + \cdots + x_n^2 + \lambda(x_1 + x_2 + \cdots + x_n - 1)$

$\frac{\partial L}{\partial x_i} = 2x_i + \lambda = 0$

$\therefore x_i = -\frac{\lambda}{2}, \quad \forall\ 1 \le i \le n$

$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^{n} x_i - 1 = 0$

$\sum_{i=1}^{n} \left(-\frac{\lambda}{2}\right) - 1 = 0 \quad \left(\because x_i = -\frac{\lambda}{2}\right)$

$-\frac{\lambda}{2} n - 1 = 0$

$\therefore \lambda = -\frac{2}{n}$

$\therefore x_i = \frac{1}{n}, \quad \forall\ 1 \le i \le n$

$\therefore x^* = \left(\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right)$

Now, $\nabla^2 f(x) + \lambda \nabla^2 h(x) > 0$

$= 2 + 0$

$= 2 > 0$

Thus, x^* is a strict local minimum of f subject to $h(x) = 0$

x^* is a min point and min value is $\sum_{i=1}^{n} \frac{1}{n^2} = \frac{1}{n}$

1-3.1.1.b)

$f(x) = \sum_{i=1}^{n} x_i$, $h(x) = \|x\|^2 - 1$

$L(x,\lambda) = \sum_{i=1}^{n} x_i + \lambda(\|x\|^2 - 1)$

$= x_1 + x_2 + \cdots + x_n + \lambda(x_1^2 + x_2^2 + \cdots + x_n^2 - 1)$

$\frac{\partial L}{\partial x_i} = 1 + \lambda(2x_i) = 0$

$\therefore x_i = -\frac{1}{2\lambda}$, $\forall\, 1 \le i \le n$

$\frac{\partial L}{\partial \lambda} = \|x\|^2 - 1 = 0$

$\frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} + \cdots + \frac{1}{4\lambda^2} - 1 = 0$

$\frac{n}{4\lambda^2} = 1$ $\quad \lambda^2 = \frac{n}{4}$

$\therefore \lambda = \pm \frac{\sqrt{n}}{2}$

For $\lambda = \frac{\sqrt{n}}{2}$, $x = -\frac{1}{\sqrt{n}}$ or $\lambda = -\frac{\sqrt{n}}{2}$, $x = \frac{1}{\sqrt{n}}$

$\therefore x^* = \left(-\frac{1}{\sqrt{n}}, -\frac{1}{\sqrt{n}}, \cdots, -\frac{1}{\sqrt{n}}\right)$ or $x^* = \left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \cdots, \frac{1}{\sqrt{n}}\right)$

Now, $\nabla^2 f(x) + \lambda \nabla^2 h(x) > 0$

$= 0 + 2\lambda$

$2\lambda > 0$

Therefore, $\lambda = \frac{\sqrt{n}}{2}$, $x = -\frac{1}{\sqrt{n}}$ is a strict local minimum of f subject to $h(x) = 0$

$x^* = \left(-\frac{1}{\sqrt{n}}, -\frac{1}{\sqrt{n}}, \cdots, -\frac{1}{\sqrt{n}}\right)$ is a min point

and min value is $-\frac{n}{\sqrt{n}} = -\sqrt{n}$.

2)

2. Industrial Design.

- The volume of can $V=4$

$$V = \pi r^2 h = 4$$

$$h = \frac{4}{\pi r^2}$$

- The surface is

$$A = 2 \cdot \pi r^2 + 2\pi r h$$
$$= 2\pi r^2 + 8\frac{1}{r} \quad \left(\because h = \frac{4}{\pi r^2}\right)$$

- cost of top/bottom = 2* cost of side

$$C = 2(2\pi r^2) + 1\left(\frac{8}{r}\right)$$

- $$\frac{dc}{dr} = 8\pi r - 8r^{-2} = 0$$
$$8\pi r = 8\frac{1}{r^2}$$
$$r^3 = \frac{1}{\pi} \qquad r = \frac{1}{\sqrt[3]{\pi}}$$

- $$h = \frac{4}{\pi r^2} \qquad h = \frac{4}{\sqrt[3]{\pi}}$$

- Radius(r) should be $\frac{1}{\sqrt[3]{\pi}}$ inches

- Height(h) should be $\frac{4}{\sqrt[3]{\pi}}$ inches

3)

3) Prove that 2 LP are dual each other

$\min\ c'x$ $\max\ b'\mu$

s.t $A'x \geq b$ s.t. $A\mu = c,\ \mu \geq 0$

- I'm going to set max is Primal and min is Dual.
- Since max problem has equality form.
  I'll change it to inequality.

$$A\mu = c \;\; \begin{cases} A\mu \leq c \\ A\mu \geq c \rightarrow -A\mu \leq -c. \end{cases}$$

$\boxed{\text{Primal}}$ $\longrightarrow$ $\boxed{\text{Dual}}$

Primal	Dual
$\max\ b'\mu$	$\min\ c'x_1 - c'x_2$
s.t. $A\mu \leq c$	s.t $A'x_1 \geq b$
$-A\mu \leq -c$	$-A'x_2 \geq b$
$\mu \geq 0.$	$x_1 \geq 0 \quad x_2 \geq 0$

- Let $x_1 - x_2 = x$ gives us dual of primal form.

$\min\ c'x$

s.t $A'x \geq b$

- Therefore, the two LP are dual each other.

4)

4-4.2.1.a)

$$\min f(x) = \frac{1}{2}(x_1^2 - x_2^2) - 3x_2$$

$$\text{s.t.} \quad x_2 = 0$$

$$L(x, \lambda) = \frac{1}{2}(x_1^2 - x_2^2) - 3x_2 + \lambda x_2$$

$$\frac{\partial L}{\partial x_1} = x_1 = 0$$

$$\frac{\partial L}{\partial x_2} = -x_2 - 3 + \lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = x_2 = 0$$

$$\therefore x_1 = 0 \quad x_2 = 0 \quad \lambda = 3 \; (\because \lambda = x_2 + 3)$$

Therefore, point (0,0) is the optimal solution since $f(x)$ is convex over constraint.

4-4.2.1.b)

$$\min \; f(x) = \frac{1}{2}(x_1^2 - x_2^2) - 3x_2$$

$$\text{s.t.} \quad x_2 = 0$$

- The augmented Lagrangian is

$$L_c(x,\lambda) = \frac{1}{2}(x_1^2 - x_2^2) - 3x_2 + \ell x_2 + \frac{c}{2}x_2^2$$

$$= \frac{1}{2}x_1^2 - \frac{1}{2}x_2^2 - 3x_2 + \ell x_2 + \frac{c}{2}x_2^2$$

$$= \frac{1}{2}x_1^2 + \left(\frac{c-1}{2}x_2 + \ell - 3\right)x_2$$

- Since $c^k = 10^{k+1}$, $c^k > 1$.

$$\nabla L_{c^k}(x^k, \lambda^k) = \begin{pmatrix} x_1^k \\ (c^k-1)x_2^k + \lambda^k - 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\therefore \; x_1^k = 0, \quad x_2^k = \frac{3-\lambda^k}{c^k-1} \quad \Rightarrow (1)$$

- The corresponding optimal value $L_{c^k}(x^k, \lambda^k)$ is

$$L_{c^k}(x^k,\lambda^k) = \left(\frac{c^k-1}{2}\left(\frac{3-\lambda^k}{c^k-1}\right) + \lambda^k - 3\right)\left(\frac{3-\lambda^k}{c^k-1}\right)$$

$$= \left(\frac{3-\lambda^k}{2} + \frac{2\lambda^k-6}{2}\right)\left(\frac{3-\lambda^k}{c^k-1}\right)$$

$$= \frac{\lambda^k-3}{2}\left(\frac{3-\lambda^k}{c^k-1}\right)$$

$$= -\frac{1}{2}\frac{(3-\lambda^k)^2}{c^k-1} \quad \Rightarrow (2)$$

- Quadratic penalty method with $\lambda^k = 0$

k	λ^k	x_1^k	x_2^k	$L_{c^k}(x^k, \lambda^k)$
0	0	0	0.333	-0.5
1	0	0	0.0303	-0.04545
2	0	0	0.003003	0.0045045

- Multiplier method

All are the same with above, the only difference is

$$\lambda^{k+1} = \lambda^k + c^k h(x^k) = \lambda^k + c^k x_2^k, \quad k = 0, 1, 2$$

$$\lambda^0 = 0.$$

k	λ^k	x_1^k	x_2^k	$L_{c^k}(x^k, \lambda^k)$
0	0	0	0.333	-0.5
1	3.333	0	-0.003367	-0.0005611
2	2.996	0	0.0000033	-5.67×10^{-9}

- We can see the multiplier method converges faster than the quadratic penalty method

4-4.2.1.d)

- Augmented Lagrangian to have a min:

$$c+\nabla^2 p(0) > 0$$

- $p(u) = \min_{h(x)=u} f(x)$

$$= \min_{x_2=u} \frac{1}{2}(x_1^2+x_2^2) - 3x_2 = -\frac{1}{2}u^2 - 3u$$

- $c+\nabla^2 p(0) > 0$

$c-1>0$

$\therefore\ c>1$

- Multiplier method with constant c.

$$\lambda^{k+1} = \lambda^k + c\,h(x^k) = \lambda^k + c\left(\frac{3-\lambda^k}{c-1}\right)$$
$$= \frac{c\lambda^k-\lambda^k}{c-1} + \frac{3c-c\lambda^k}{c-1}$$
$$= \frac{-\lambda^k+3c}{c-1}$$

- For $\{\lambda^k\}$ to converse λ^*, $\dfrac{|\lambda^{k+1}-\lambda^*|}{|\lambda^k-\lambda^*|} < 1$

Since λ^* is 3 ($\because$ the result of (a) problem)

$$\frac{|\lambda^{k+1}-\lambda^*|}{|\lambda^k-\lambda^*|} = \frac{\left|\frac{-\lambda^k+3c}{c-1}-3\right|}{|\lambda^k-3|} = \frac{\left|\frac{-\lambda^k+3c-3c+3}{c-1}\right|}{|\lambda^k-3|} = \frac{1}{|c-1|}$$

and $\frac{1}{|c-1|}<1$ makes convergence.

$\therefore\ c>2$

5)

5-1) AMPL model file.

- I put all code into one file. It also includes parameters and commands as well.

```
reset;

param K := 20;
param x1 {1..K};
param x2 {1..K};
param y {1..K};

let x1[1]  := 46;
let x1[2]  := 74;
let x1[3]  := 89;
let x1[4]  := 77;
let x1[5]  := 84;
let x1[6]  := 89;
let x1[7]  := 68;
let x1[8]  := 70;
let x1[9]  := 60;
let x1[10] := 55;
let x1[11] := 35;
let x1[12] := 51;
let x1[13] := 87;
let x1[14] := 83;
let x1[15] := 68;
let x1[16] := 84;
let x1[17] := 74;
let x1[18] := 73;
let x1[19] := 84;
let x1[20] := 91;


let x2[1]  := 0;
let x2[2]  := 0;
let x2[3]  := 16;
let x2[4]  := 16;
let x2[5]  := 21;
let x2[6]  := 15;
let x2[7]  := 14;
let x2[8]  := 6;
let x2[9]  := 13;
let x2[10] := 9;
let x2[11] := 3;
let x2[12] := 7;
let x2[13] := 23;
let x2[14] := 4;
let x2[15] := 0;
let x2[16] := 19;
let x2[17] := 3;
let x2[18] := 0;
let x2[19] := 15;
let x2[20] := 7;


let y[1]  := 1;
let y[2]  := 10;
let y[3]  := 29;
let y[4]  := 25;
let y[5]  := 29;
let y[6]  := 40;
```

```
let y[7]  := 21;
let y[8]  := 0;
let y[9]  := 13;
let y[10] := 4;
let y[11] := 0;
let y[12] := 7;
let y[13] := 21;
let y[14] := 9;
let y[15] := 7;
let y[16] := 22;
let y[17] := 6;
let y[18] := 2;
let y[19] := 29;
let y[20] := 11;


var a1;
var a2;
var b;

minimize mse: sum {i in 1..K} (y[i] - (a1*x1[i] + a2*x2[i] + b))^2;


solve;
display a1;
display a2;
display b;
for {i in 1 .. K} {  display i, x1[i], x2[i], y[i], a1*x1[i] + a2*x2[i] + b, abs(y[i] -
(a1*x1[i] + a2*x2[i] + b)); }
```

5-2) Output of NEOS solver (I chosen CPLEX for the solution of linear programming)

+ The objective function:

$$MSE = \sum_{i=1}^{K} (y[i] - (a_1 * \ x_1[i] + a_2 * \ x_2[i] + b\ )^2$$

+ The result values from NEOS solver for MSE are:

- a1 = 0.270589
- a2 = 0.967714
- b = -14.4511

+ The linear regression function:

$$Linear\ model\ = 0.270589x_1 + \ 0.967714x_2 - 14.4511$$

+ NEOS output is below.

```
*************************************************************

   NEOS Server Version 6.0
   Job#     : 12800824
   Password : KrqfGcvh
   User     :
   Solver   : lp:CPLEX:AMPL
   Start    : 2023-03-07 15:48:48
   End      : 2023-03-07 15:48:49
   Host     : prod-sub-1.neos-server.org
```

```
   Disclaimer:

   This information is provided without any express or
   implied warranty. In particular, there is no warranty
   of any kind concerning the fitness of this
   information for any particular purpose.

   Announcements:
*************************************************************
You are using the solver cplexamp.
Checking ampl.mod for cplex_options...
Executing AMPL.
processing data.
processing commands.
Executing on prod-exec-7.neos-server.org

3 variables, all nonlinear
0 constraints
1 nonlinear objective; 3 nonzeros.

CPLEX 20.1.0.0: threads=4
CPLEX 20.1.0.0: optimal solution; objective 694.0056746
No basis.
a1 = 0.270589

a2 = 0.967714

b = -14.4511

i = 1
x1[i] = 46
x2[i] = 0
y[i] = 1
a1*x1[i] + a2*x2[i] + b = -2.00403
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 3.00403

i = 2
x1[i] = 74
x2[i] = 0
y[i] = 10
a1*x1[i] + a2*x2[i] + b = 5.57245
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 4.42755

i = 3
x1[i] = 89
x2[i] = 16
y[i] = 29
a1*x1[i] + a2*x2[i] + b = 25.1147
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 3.8853

i = 4
x1[i] = 77
x2[i] = 16
y[i] = 25
a1*x1[i] + a2*x2[i] + b = 21.8676
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 3.13236

i = 5
x1[i] = 84
x2[i] = 21
y[i] = 29
a1*x1[i] + a2*x2[i] + b = 28.6003
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 0.399675

i = 6
x1[i] = 89
x2[i] = 15
y[i] = 40
a1*x1[i] + a2*x2[i] + b = 24.147
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 15.853

i = 7
```

```
x1[i] = 68
x2[i] = 14
y[i] = 21
a1*x1[i] + a2*x2[i] + b = 17.4969
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 3.50309

i = 8
x1[i] = 70
x2[i] = 6
y[i] = 0
a1*x1[i] + a2*x2[i] + b = 10.2964
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 10.2964

i = 9
x1[i] = 60
x2[i] = 13
y[i] = 13
a1*x1[i] + a2*x2[i] + b = 14.3645
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 1.36449

i = 10
x1[i] = 55
x2[i] = 9
y[i] = 4
a1*x1[i] + a2*x2[i] + b = 9.14069
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 5.14069

i = 11
x1[i] = 35
x2[i] = 3
y[i] = 0
a1*x1[i] + a2*x2[i] + b = -2.07736
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 2.07736

i = 12
x1[i] = 51
x2[i] = 7
y[i] = 7
a1*x1[i] + a2*x2[i] + b = 6.12291
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 0.877088

i = 13
x1[i] = 87
x2[i] = 23
y[i] = 21
a1*x1[i] + a2*x2[i] + b = 31.3475
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 10.3475

i = 14
x1[i] = 83
x2[i] = 4
y[i] = 9
a1*x1[i] + a2*x2[i] + b = 11.8786
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 2.8786

i = 15
x1[i] = 68
x2[i] = 0
y[i] = 7
a1*x1[i] + a2*x2[i] + b = 3.94892
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 3.05108

i = 16
x1[i] = 84
x2[i] = 19
y[i] = 22
a1*x1[i] + a2*x2[i] + b = 26.6649
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 4.6649

i = 17
x1[i] = 74
x2[i] = 3
y[i] = 6
```

```
a1*x1[i] + a2*x2[i] + b = 8.47559
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 2.47559

i = 18
x1[i] = 73
x2[i] = 0
y[i] = 2
a1*x1[i] + a2*x2[i] + b = 5.30186
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 3.30186

i = 19
x1[i] = 84
x2[i] = 15
y[i] = 29
a1*x1[i] + a2*x2[i] + b = 22.794
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 6.20596

i = 20
x1[i] = 91
x2[i] = 7
y[i] = 11
a1*x1[i] + a2*x2[i] + b = 16.9465
abs(y[i] - (a1*x1[i] + a2*x2[i] + b)) = 5.94645

NEOS Server Home
```

5-3)

index	a1	a2	y	model value	difference
1	46	0	1	-2.004006	3.004006
2	74	0	10	5.572486	4.427514
3	89	16	29	25.114745	3.885255
4	77	16	25	21.867677	3.132323
5	84	21	29	28.60037	0.39963
6	89	15	40	24.147031	15.852969
7	68	14	21	17.496948	3.503052
8	70	6	0	10.296414	10.296414
9	60	13	13	14.364522	1.364522
10	55	9	4	9.140721	5.140721
11	35	3	0	-2.077343	2.077343
12	51	7	7	6.122937	0.877063
13	87	23	21	31.347565	10.347565
14	83	4	9	11.878643	2.878643
15	68	0	7	3.948952	3.051048
16	84	19	22	26.664942	4.664942
17	74	3	6	8.475628	2.475628
18	73	0	2	5.301897	3.301897
19	84	15	29	22.794086	6.205914
20	91	7	11	16.946497	5.946497

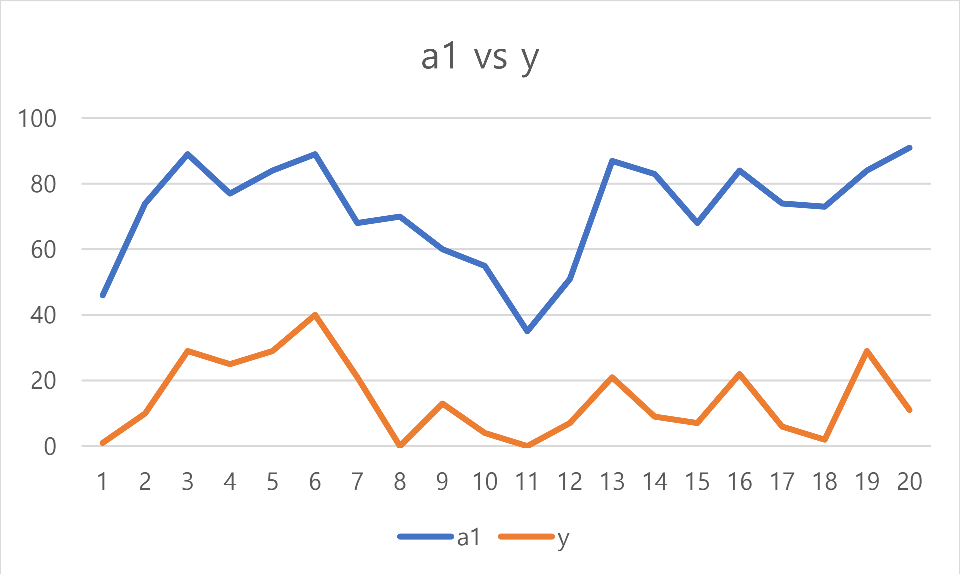
5-4)

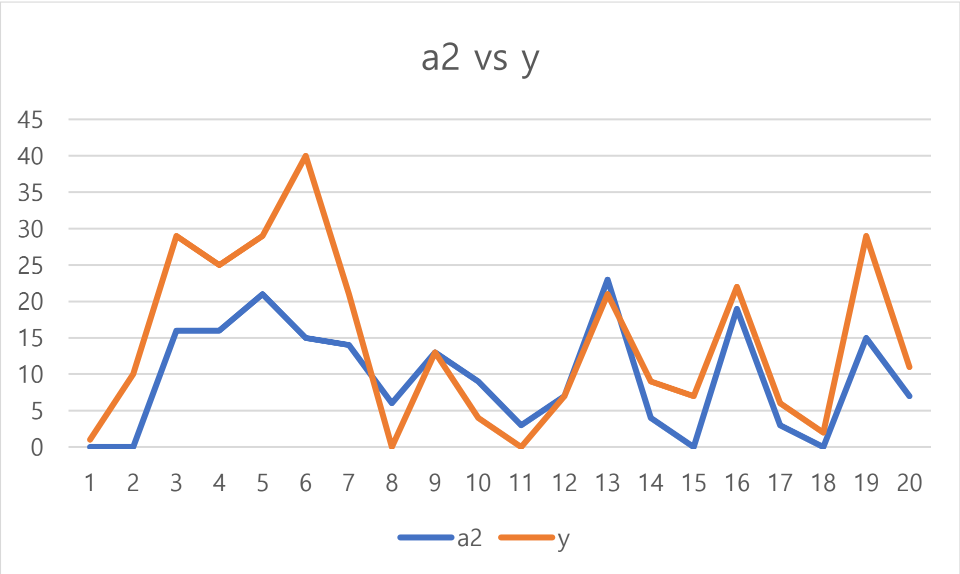
+ How does each attribute influence the change?

  - The dependent variable y has moved to the same directions while a1 and a2 are changing. I can conclude the a1 and a2 have some relationship between dependent variable y.

+ Which attribute seems to have stronger correlation with the change?

  - I think a2 has more influence on the dependent variable y. I plotted a simple graph in Excel, and the graph of ‘a2 vs y’ is showing a much more similar plot to ‘a1 vs y’. This clearly shows that a2 has a stronger correlation with the dependent y variable.





+ Does the linear regression model seem accurate to you?

  - I’ve calculated R-squared value and it was 0.74. It means 74 out of 100 data can be explained by the model I found. As a result, the linear regression model seems to be fairly accurate.