01

MODERN OPTIMIZATION TECHNIQUES - GROUP 01

FIRST TAKE HOME EXAM

SUBMITTED BY: MUHAMMAD INAAM ASHRAF

MATRIKEL NR: 307524

Q2A: Closed form solution of ridge regression

$$f(\hat{\beta};\lambda,D) = \frac{1}{N}\sum_{n=1}^{N}(y_n - \hat{y}_n(x_n;\hat{\beta}))^2 + \lambda\sum_{m=1}^{M}\hat{\beta}_m^2$$

$$= \frac{1}{N}\left(\left(\sum_{n=1}^{N} y_n - \sum_{n=1}^{N}\hat{y}_n(x_n;\hat{\beta})\right)\left(\sum_{n=1}^{N} y_n - \sum_{n=1}^{N}\hat{y}_n(x_n;\hat{\beta})\right)\right) + \lambda\sum_{m=1}^{M}\hat{\beta}_m^2$$

we have $\sum_{n=1}^{N}\hat{y}_n(x_n;\hat{\beta}) = X\hat{\beta}$, $\sum_{n=1}^{N} y_n = Y$, $\sum_{m=1}^{M}\hat{\beta}_m^2 = \langle\hat{\beta},\hat{\beta}\rangle$

So,

$$f(\hat{\beta};\lambda,D) = \frac{1}{N}\left((Y - X\hat{\beta})(Y - X\hat{\beta})\right) + \lambda\langle\hat{\beta},\hat{\beta}\rangle$$

To find close form solution, we take derivative and put it equal to 0.

$$\nabla_{\hat{\beta}} f = \frac{1}{N}\left((Y - X\hat{\beta})(-X^T) + (Y - X\hat{\beta})(-X^T)\right) + 2\lambda\hat{\beta} = 0$$

$$\Rightarrow -\frac{2}{N}X^TY + \frac{2}{N}X^TX\hat{\beta} + 2\lambda\hat{\beta} = 0$$

$$\Rightarrow \frac{1}{N}X^TX\hat{\beta} + \lambda\hat{\beta} = \frac{1}{N}X^TY$$

Sol?

0, 2/

02

M. Inaam Ashraf (307524)

Q2A: Why use SGD?

The closed form solution of linear regression is

$$\beta = (X^T X)^{-1} X^T Y$$

As evident, the computational cost of this solution is way too high as it involves matrix inverse. Therefore, using Gradient Descent is a rational approach to minimize the computational cost.

1/1

Q2B:

$$\mathcal{L}_{ridge}(X, Y, \beta) = \|X\beta - Y\|_2^2 + \lambda \|\beta\|_2$$

a: For update rule, we compute gradient of $\mathcal{L}_{ridge}$.

$$\nabla_\beta \mathcal{L}_{ridge}(X, Y, \beta) = (X\beta - Y)(X^T) + (X\beta - Y)(X^T) + 2\lambda\beta$$

$$= 2X^T(X\beta - Y) + 2\lambda\beta$$

1/1

So, update rule is

$$\beta^{t+1} = \beta^t - \mu\left(2X^T(X\beta^t - Y) + 2\lambda\beta^t\right)$$

Now, we have

$$X = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 3 & 1 \\ 1 & -1 & 2 \end{bmatrix}, \quad \beta^0 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \quad Y = \begin{bmatrix} -1 \\ 5 \\ 7 \end{bmatrix}$$

$\mu = 0.1$ & $\lambda = 0.1$

M. Inaam Ashraf (307524)

Q2B: Iteration 01:

$$\beta^1 = \beta^0 - 2\mu\left(X^T(X\beta^0 - Y) + \lambda\beta^0\right)$$

$$= \begin{bmatrix}-1\\1\\2\end{bmatrix} - 2(0.1)\left(\begin{bmatrix}1&1&1\\2&3&-1\\0&1&2\end{bmatrix}\left(\begin{bmatrix}1&2&0\\1&3&1\\1&-1&2\end{bmatrix}\begin{bmatrix}-1\\1\\2\end{bmatrix} - \begin{bmatrix}-1\\5\\7\end{bmatrix}\right) + (0.1)\begin{bmatrix}-1\\1\\2\end{bmatrix}\right)$$

$$\beta^1 = \begin{bmatrix}-1\\1\\2\end{bmatrix} - 0.2\left(\begin{bmatrix}1&1&1\\2&3&-1\\0&1&2\end{bmatrix}\left(\begin{bmatrix}1\\4\\2\end{bmatrix} - \begin{bmatrix}-1\\5\\7\end{bmatrix}\right) + \begin{bmatrix}-0.1\\0.1\\0.2\end{bmatrix}\right)$$

$$= \begin{bmatrix}-1\\1\\2\end{bmatrix} - 0.2\left(\begin{bmatrix}1&1&1\\2&3&-1\\0&1&2\end{bmatrix}\begin{bmatrix}2\\-1\\-5\end{bmatrix} + \begin{bmatrix}-0.1\\0.1\\0.2\end{bmatrix}\right)$$

$$= \begin{bmatrix}-1\\1\\2\end{bmatrix} - 0.2\left(\begin{bmatrix}-4\\6\\-11\end{bmatrix} + \begin{bmatrix}-0.1\\0.1\\0.2\end{bmatrix}\right)$$

$$= \begin{bmatrix}-1\\1\\2\end{bmatrix} - 0.2\begin{bmatrix}-4.1\\6.1\\-10.8\end{bmatrix} = \begin{bmatrix}-1\\1\\2\end{bmatrix} - \begin{bmatrix}-0.82\\1.22\\-2.16\end{bmatrix}, \quad \beta^1 = \begin{bmatrix}-0.18\\-0.22\\4.16\end{bmatrix}$$

Computing Loss.

$$\mathcal{L}_1^1 = (X\beta^1 - Y)^T(X\beta^1 - Y) + \lambda\beta^{1T}\beta^1$$

$$X\beta^1 - Y = \begin{bmatrix}1&2&0\\1&3&1\\1&-1&2\end{bmatrix}\begin{bmatrix}-0.18\\-0.22\\4.16\end{bmatrix} - \begin{bmatrix}-1\\5\\7\end{bmatrix} = \begin{bmatrix}-0.62\\3.32\\8.36\end{bmatrix} - \begin{bmatrix}-1\\5\\7\end{bmatrix}$$

$$X\beta^1 - Y = \begin{bmatrix}0.38\\-1.68\\1.36\end{bmatrix}$$

$$\ell^1 = \begin{bmatrix} 0.38 & -1.68 & 1.36 \end{bmatrix} \begin{bmatrix} 0.38 \\ -1.68 \\ 1.36 \end{bmatrix} + 0.1 \begin{bmatrix} -0.18 & -0.22 & 4.16 \end{bmatrix} \begin{bmatrix} -0.18 \\ -0.22 \\ 4.16 \end{bmatrix}$$

$$\ell^1 = 4.8164 + 0.1(17.3864), \quad \ell^1 = 6.555$$

V

Iteration 2:

$$\beta^2 = \beta^1 - 0.2\left(\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & -1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0.38 \\ -1.68 \\ 1.36 \end{bmatrix}}_{X\beta^1 - Y} + 0.1 \begin{bmatrix} -0.18 \\ -0.22 \\ 4.16 \end{bmatrix} \right)$$

$$\beta^2 = \beta^1 - 0.2 \left(\begin{bmatrix} 0.06 \\ -5.64 \\ 1.04 \end{bmatrix} + \begin{bmatrix} -0.018 \\ -0.022 \\ 0.416 \end{bmatrix} \right)$$

$$= \beta^1 - 0.2 \begin{bmatrix} 0.042 \\ -5.662 \\ 1.456 \end{bmatrix} = \beta^1 - \begin{bmatrix} 0.00084 \\ -1.1324 \\ 0.2912 \end{bmatrix}$$

$$\beta^2 = \begin{bmatrix} -0.18 \\ -0.22 \\ 4.16 \end{bmatrix} - \begin{bmatrix} 0.0008 \\ -1.1324 \\ 0.2912 \end{bmatrix}$$

$$\beta^2 = \begin{bmatrix} -0.1808 \\ 0.9124 \\ 3.8688 \end{bmatrix}$$

V

Computing ℓ^2, first

$$X\beta^2 - Y = \begin{bmatrix} -0.1808 \\ 0.9124 \\ 3.8688 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 1 & 3 & 1 \\ 1 & -1 & 2 \end{bmatrix} - \begin{bmatrix} -1 \\ 5 \\ 7 \end{bmatrix}$$

$$\beta^2 X - Y = \begin{bmatrix} 1.644 \\ 6.4252 \\ 6.6444 \end{bmatrix} - \begin{bmatrix} -1 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} 2.644 \\ 1.4252 \\ -0.3556 \end{bmatrix}$$

M.Inaam Ashraf (~~3~~ 307524) 05

$$\mathcal{L}^2 = \begin{bmatrix} 2.644 & 1.4252 & -0.3556 \end{bmatrix} \begin{bmatrix} 2.644 \\ 1.4252 \\ -0.3556 \end{bmatrix} + 0.1 \begin{bmatrix} -0.1808 & 0.9124 & 3.8688 \end{bmatrix} \begin{bmatrix} -0.1808 \\ 0.9124 \\ 3.8688 \end{bmatrix}$$

$$\mathcal{L}^2 = 9.148 + 0.1(15.83), \quad \mathcal{L}^2 = 10.731$$

✓

Loss is increasing since $\mathcal{L}^2 = 10.731 > \mathcal{L}^1 = 6.555$.

3/3

Q2C. Backtracking condition is given by

$$\mathcal{L}_{ridge}(\beta^0 - \mu \nabla_\beta \mathcal{L}_{ridge}) > \mathcal{L}_{ridge} - a\mu \nabla_{\beta^0} \mathcal{L}_{ridge}^T \cdot \nabla_{\beta^0} \mathcal{L}_{ridge}$$

0,75

Since β^0 is same, I will use the values from previous

Q2B:

Iteration 1: $\beta^1 = \beta^0 - \mu \nabla_{\beta^0} \mathcal{L}$

$$= \beta^0 - 2\mu\left(X^T(X\beta^0 - Y) + \lambda\beta^0\right) = \begin{bmatrix} -0.18 \\ -0.22 \\ 4.16 \end{bmatrix}$$

$$\nabla_{\beta^0} \mathcal{L} = 2\left(X^T(X\beta^0 - Y) - \lambda\beta^0\right) = 2\begin{bmatrix} -4.1 \\ 6.1 \\ -10.8 \end{bmatrix} = \begin{bmatrix} -8.2 \\ 12.2 \\ -21.6 \end{bmatrix}$$

$$\mathcal{L}_{ridge_{\beta^0}} : (X\beta^0 - Y)^T (X\beta^0 - Y) + \lambda \beta^{0T}\beta^0$$

$$X\beta^0 - Y = \begin{bmatrix} 2 \\ -1 \\ -5 \end{bmatrix}$$

$$\mathcal{L}_{\beta^0} = \begin{bmatrix} 2 & -1 & -5 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ -5 \end{bmatrix} + \lambda \begin{bmatrix} -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

06

M. Inaam Ashraf (307524)

$\mathcal{L}_{\beta^0} = 30 + 0.1\,(6) = 30.6$

Checking condition: $\mathcal{L}_{\beta^1}^{1} > \mathcal{L}_{\beta^0} - a\mu\, \nabla_{\beta^0}\mathcal{L}^T \cdot \nabla_{\beta^0}\mathcal{L}$

$$6.555 > 30.6 - (0.1)(0.1)\begin{bmatrix}-8.2 & 12.2 & -21.6\end{bmatrix}\begin{bmatrix}-8.2\\12.2\\-21.6\end{bmatrix}$$

$$6.555 > 30.6 - 6.8264$$

No Change.

$$6.555 \not> 23.77 \rightarrow \text{So } \mu = 0.1$$

Iteration 2:: Again I can use values from 2B.

Not for $\beta(2)$

$$\beta^2 = \begin{bmatrix}-0.1808\\0.9124\\3.8688\end{bmatrix}, \quad \mathcal{L}_{\beta^2} = 10.731, \quad \mathcal{L}_{\beta^1} = 6.555$$

$$\nabla_{\beta^1}\mathcal{L} = 2\begin{bmatrix}0.042\\-5.662\\1.456\end{bmatrix} = \begin{bmatrix}0.084\\-11.324\\2.912\end{bmatrix}$$

Checking condition.

$$10.731 > 6.555 - (0.1)(0.1)\begin{bmatrix}0.084 & -11.324 & 2.912\end{bmatrix}\begin{bmatrix}0.084\\-11.324\\2.912\end{bmatrix}$$

$$10.731 > 6.555 - 1.367$$

$$10.731 > 5.187 \rightarrow \text{True.}$$

So, μ for next iteration $\mu = b\mu = 0.1(0.1) = 0.001$

Here, Backtracking will help in the next iteration as the loss will be decreased because we have adjusted step size μ.

$\frac{2.5}{3}$

Q3A a) In GD, we compute new β after going through all observations i.e. we update parameters after observing the whole dataset.

In SGD, we pick a sample of the observations, (e.g. one or more) compute the gradient and update the parameters (β) with each sample.

$\frac{1}{1}$

b) In SGD, we do not have access to the new Loss of all observations after each iteration, so we cannot compare the total losses. i.e. $\mathcal{L}_{\beta'}$ for whole data can not be computed after every iteration.

Instead we compare the sample losses i.e.

$$\mathcal{L}_{\beta}(\beta')_{sample} > \mathcal{L}(\beta^{\circ})$$

$\frac{1}{2}$

Q3B update rule is

$$\beta^{t+1} = \beta^{t} - \mu\left(2x^{T}_{t,:}(x_{t,:}\beta^{t} - y_{t}) + 2\lambda\beta^{t}\right)$$

$\frac{2}{1}$

Iteration 1:- $x_{0,:}\beta^{0} - y_{0} = \begin{bmatrix}1 & 2 & 0\end{bmatrix}\begin{bmatrix}-1\\1\\2\end{bmatrix} - (-1) = 2$

$$\beta' = \begin{bmatrix}-1\\1\\2\end{bmatrix} - 0.1(2)\left(\begin{bmatrix}1\\2\\0\end{bmatrix}(2) + (0.1)\begin{bmatrix}-1\\1\\2\end{bmatrix}\right)$$

08

M. Inaam Ashraf (307524)

$$\cancel{\beta^1 = \begin{bmatrix}-1\\1\\2\end{bmatrix} - 0.1(2)\left(\begin{bmatrix}2\\4\\0\end{bmatrix} + \begin{bmatrix}-0.1\\0.1\\0.2\end{bmatrix}\right)}$$

$$\cancel{= \begin{bmatrix}-1\\1\\2\end{bmatrix} - 0.1(2)\begin{bmatrix}1.9\\4.1\\0.2\end{bmatrix} = \begin{bmatrix}-1\\1\\2\end{bmatrix} - \begin{bmatrix}0.42\\0.78\\-0.04\end{bmatrix}, \beta^1 = \begin{bmatrix}-1.42\\0.22\\2.04\end{bmatrix}}$$

$$\cancel{\mathcal{L}_{\beta^1} = (X_{0,:}\beta^1 - Y)^2 + \lambda(\beta^T\beta)}$$

$$\cancel{= \left(\begin{bmatrix}1 & 2 & 0\end{bmatrix}\begin{bmatrix}-1.42\\0.22\\2.04\end{bmatrix} - (-1)\right)^2 + (0.1)\begin{bmatrix}-1.42 & 0.22 & 2.04\end{bmatrix}\begin{bmatrix}-1.42\\0.22\\2.04\end{bmatrix}}$$

$$\cancel{= (0.02)^2 + 0.1(6.2264), \quad \mathcal{L}_{\beta^1} = 0.623}$$

~~Iteration 2:~~

$$\cancel{\beta^2 = \begin{bmatrix}1\\3\\1\end{bmatrix} - 0.1(2)\left(\begin{bmatrix}1\\3\\1\end{bmatrix}\left(\begin{bmatrix}1 & 3 & 1\end{bmatrix}\begin{bmatrix}-1.42\\0.22\\2.04\end{bmatrix} - 1\right)\right)}$$

$$\beta^1 = \begin{bmatrix}-1\\1\\2\end{bmatrix} - 0.1(2)\left(\begin{bmatrix}2\\4\\0\end{bmatrix} + \begin{bmatrix}-0.1\\0.1\\0.2\end{bmatrix}\right)$$

$$= \begin{bmatrix}-1\\1\\2\end{bmatrix} - 0.1(2)\begin{bmatrix}1.9\\4.1\\0.2\end{bmatrix} = \begin{bmatrix}-1\\1\\2\end{bmatrix} - \begin{bmatrix}0.38\\0.82\\0.04\end{bmatrix}, \beta^1 = \begin{bmatrix}-1.38\\0.18\\1.96\end{bmatrix}$$

$$\mathcal{L}_{\beta^1} = (X_{0:}\beta^1 - Y)^2 + \lambda\beta^{1T}\beta^1$$

$$= \left(\begin{bmatrix}1 & 2 & 0\end{bmatrix}\begin{bmatrix}-1.38\\0.18\\1.96\end{bmatrix} + 1\right)^2 + 0.1\begin{bmatrix}-1.38 & 0.18 & 1.96\end{bmatrix}\begin{bmatrix}-1.38\\0.18\\1.96\end{bmatrix}$$

$$= (-0.02)^2 + 0.1(5.7784)$$

$$\mathcal{L}_{\beta^1} = \cancel{1.6782}\ 0.5784 \quad ?$$

Q3B. Iteration 2:

$$\beta^2 = \begin{bmatrix}-1.38\\0.18\\1.96\end{bmatrix} - 0.1(2)\left(\begin{bmatrix}1\\3\\1\end{bmatrix}\left(\begin{bmatrix}1 & 3 & 1\end{bmatrix}\begin{bmatrix}-1.38\\0.18\\1.96\end{bmatrix} - 5\right) + 0.1\begin{bmatrix}-1.38\\0.18\\1.96\end{bmatrix}\right)$$

$$= \begin{bmatrix}-1.38\\0.18\\1.96\end{bmatrix} - 0.1(2)\left(\begin{bmatrix}1\\3\\1\end{bmatrix}(-3.88) + \begin{bmatrix}-0.138\\0.018\\0.196\end{bmatrix}\right)$$

$$= \begin{bmatrix}-1.38\\0.18\\1.96\end{bmatrix} - 0.1(2)\begin{bmatrix}-4.018\\-11.622\\-3.684\end{bmatrix}, \quad \beta^2 = \begin{bmatrix}-0.576\\2.5\\2.7\end{bmatrix}$$

$$\ell_{\beta^2} = \left(\begin{bmatrix}1 & 3 & 1\end{bmatrix}\begin{bmatrix}-0.576\\2.5\\2.7\end{bmatrix} - 5\right)^2 + 0.1\, \beta^{2^T} \cdot \beta^2$$

$$= (4.624)^2 + 0.1(13.87), \quad \ell_{\beta^2} = 22.77$$

Iteration 3:

$$\beta^3 = \beta^2 - 0.1(2)\left(\begin{bmatrix}1\\-1\\2\end{bmatrix}\left(\begin{bmatrix}1 & -1 & 2\end{bmatrix}\begin{bmatrix}-0.576\\2.5\\2.7\end{bmatrix} - 7\right) + 0.1\begin{bmatrix}-0.576\\2.5\\2.7\end{bmatrix}\right)$$

$$= \beta^2 - 0.2\left(\begin{bmatrix}1\\-1\\2\end{bmatrix}(-4.676) + \begin{bmatrix}-0.0576\\0.25\\0.27\end{bmatrix}\right)$$

$$= \beta^2 - 0.2\begin{bmatrix}-4.73\\4.926\\-9.08\end{bmatrix} = \begin{bmatrix}-0.576\\2.5\\2.7\end{bmatrix} - \begin{bmatrix}-0.946\\0.985\\-1.81\end{bmatrix}, \quad \beta^3 = \begin{bmatrix}0.37\\1.515\\4.51\end{bmatrix}$$

$$\ell_{\beta^3} = \left(\begin{bmatrix}1 & -1 & 2\end{bmatrix}\begin{bmatrix}0.37\\1.515\\4.51\end{bmatrix} - 7\right)^2 + 0.1\, \beta^{3^T} \cdot \beta^3$$

$$= (0.875)^2 + 0.1(22.77), \quad \ell_{\beta^3} = 3.043$$

$\frac{2.5}{3}$

10

M. Inaam Ashraf (307524)

Q3B. So the loss first increases and then decreases.

$$\mathcal{L}_{\beta^1} = \cancel{1.6882}\ 0.5784 < \mathcal{L}_{\beta^2} = 22.77 > \mathcal{L}_{\beta^3} = 3.043.$$

Q3C. I will use the values from Q3B.

$$\beta^{t+1} = \beta^t - \frac{\mu_0}{\sqrt{G_n^2 + \epsilon}} \nabla_{\beta^t} \mathcal{L}_{\beta^t}, \quad \text{I assume } \epsilon = 0. \quad G_0^2 = 0$$

Iteration 1:

$$G_1^2 = G_0^2 + \nabla_{\beta^0} \mathcal{L}_{\beta^0} \odot \nabla_{\beta^0} \mathcal{L}_{\beta^0}$$

$$\nabla_{\beta^0} \mathcal{L}_{\beta^0} = 2\begin{bmatrix} 1.9 \\ 4.1 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 3.8 \\ 8.2 \\ 0.4 \end{bmatrix}$$

$$G_1^2 = 0 + \begin{bmatrix} 3.8 \\ 8.2 \\ 0.4 \end{bmatrix} \odot \begin{bmatrix} 3.8 \\ 8.2 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 14.44 \\ 67.24 \\ 0.16 \end{bmatrix}$$

$$\beta^1 = \beta^0 - \frac{0.1}{\sqrt{\begin{bmatrix} 14.44 \\ 67.24 \\ 0.16 \end{bmatrix}}} \begin{bmatrix} 3.8 \\ 8.2 \\ 0.4 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 0.026 \\ 0.012 \\ 0.25 \end{bmatrix} \begin{bmatrix} 3.8 \\ 8.2 \\ 0.4 \end{bmatrix}$$

$$\beta^1 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 0.1 \\ 0.1 \\ 0.1 \end{bmatrix}, \quad \beta^1 = \begin{bmatrix} -1.1 \\ 0.9 \\ 1.9 \end{bmatrix}$$

$$\mathcal{L}_{\beta^1} = \left(\begin{bmatrix} 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} -1.1 \\ 0.9 \\ 1.9 \end{bmatrix} - (-1) \right)^2 + \lambda \beta^{1T} \beta^1$$

$$= (1.7)^2 + 0.1(3.21), \quad \mathcal{L}_{\beta^1} = 3.211$$

1.5 / 3

Iteration 2: $G_2^2 = G_1^2 + \begin{bmatrix} -8.036 \\ -23.24 \\ -7.37 \end{bmatrix} \odot \begin{bmatrix} -8.036 \\ -23.24 \\ -7.37 \end{bmatrix}$

$$G_2^2 = \begin{bmatrix} 14.44 \\ 67.24 \\ 0.16 \end{bmatrix} + \begin{bmatrix} 64.58 \\ 540.1 \\ 54.3 \end{bmatrix}, \quad G_2^2 = \begin{bmatrix} 79 \\ 607.3 \\ 54.46 \end{bmatrix}$$

Q1A:

$$S = \{ (x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 \le r^2 \}$$

By def:. $x^2+y^2+z^2 \le r^2 \in S.$

Then $\theta(x^2+y^2+z^2) \le \theta r^2 \le S.$

& $(1-\theta)(x^2+y^2+z^2) \le (1-\theta) r^2 \in S$

Thus.

$$\theta(x^2+y^2+z^2) + (1-\theta)(x^2+y^2+z^2) \le \theta r^2 + (1-\theta) r^2 \in S$$

X No

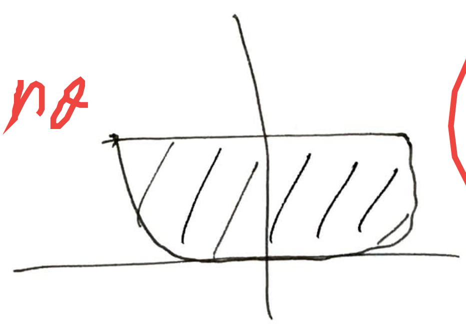
set $u^2 = x^2+y^2+z^2$

Then $\theta u^2 + (1-\theta)u^2 \le \theta r^2 + (1-\theta) r^2 \in S.$

Thus $u^2 \le r^2 \in S$ and it is a convex set.

Shape $\longrightarrow$

no



0/3

Q1C. $h(x) = x^2$

is a convex function becaus.

$$\frac{\partial h(x)}{\partial x} = 2x \quad \text{iso} \quad \Rightarrow \cancel{x = 0} \text{ is}$$

& $$\frac{\partial^2 h}{\partial x^2} = 2 > 0$$ ✓

0.5/1.5

12) M. Inaam Ashraf (307524)

Positive sign of second derivative indicate that the it has a minima and thus is a convex set.

Q1B, obj func. $4x_1 + 3x_2 + 3x_3$

Tableau : $\begin{bmatrix} B & I_{a \times a} & b \\ c & 0_a & a \end{bmatrix}$

Constraints:

$$x_1 + 0 + 2x_3 \geq 2$$
$$3x_1 + x_2 + x_3 \geq 4$$
$$0 + 4x_3 + 0 \geq 1$$
$$x_1 + x_2 + x_3 \geq 1$$

0/1

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 0 & 0 & 0 & 2 \\ 3 & 1 & 1 & 0 & 1 & 0 & 0 & 4 \\ 0 & 4 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 4 & 3 & 3 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{matrix} 2/1 = 2 \\ 4/3 \\ \infty \\ 1 \\ \end{matrix}$$

Pivot is (4, 1)

$$\begin{bmatrix} 0 & -1 & 1 & 1 & 0 & 0 & -1 & 1 \\ 2 & 0 & 0 & 0 & 1 & 0 & -1 & 3 \\ 0 & 4 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & -1 & -1 & 0 & 0 & 0 & -4 & -4 \end{bmatrix}$$

?

Doesn't work.

M. Inaam Ashraf (307524).

Q1B. Obj Function Max. P. $2x_1 + 4x_2 + x_3 + x_4$

Constraints:

$$x_1 + 3x_2 + 0 + x_4 \leq 4$$

$$0 + x_2 + 4x_3 + x_4 \leq 3$$

$$2x_1 + x_2 + 0 + x_4 \leq 3$$

$$\text{Simplex Tableau} = \begin{bmatrix} B & I_{q \times q} & b \\ c & 0_q & 0 \end{bmatrix}$$

Here,

$$B = \begin{bmatrix} 1 & 3 & 0 & 1 \\ 0 & 1 & 4 & 1 \\ 2 & 1 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 3 \\ 3 \end{bmatrix}$$

$$c = \begin{bmatrix} 2 & 4 & 1 & 1 \end{bmatrix}$$

So,

$$\begin{bmatrix} 1 & 3 & 0 & 1 & 1 & 0 & 0 & 4 \\ 0 & 1 & 4 & 1 & 0 & 1 & 0 & 3 \\ 2 & 1 & 0 & 1 & 0 & 0 & 1 & 3 \\ 2 & 4 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{matrix} 4/3 \\ 3/1 \\ 3/1 \\ \end{matrix}$$

Max value in last row = 4, Min location in col 2 = 1

So pivot = (1, 2)

Dividing by 3 row 1

$$\begin{bmatrix} 1/3 & 1 & 0 & 1/3 & 1/3 & 0 & 0 & 4/3 \\ 0 & 1 & 4 & 1 & 0 & 1 & 0 & 3 \\ 2 & 1 & 0 & 1 & 0 & 0 & 1 & 3 \\ 2 & 4 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

14) M. Inaam Ashraf (307524).

Performing Row 2 − Row 1

Row 3 − Row 1

Row 4 − 4 × Row 1

$$\begin{bmatrix} 1/3 & 1 & 0 & 1/3 & 1/3 & 0 & 0 & 4/3 \\ -1/3 & 0 & 4 & 2/3 & -1/3 & 1 & 0 & 5/3 \\ 5/3 & 0 & 0 & 2/3 & -1/3 & 0 & 1 & 5/3 \\ -2/3 & 0 & 1 & -1/3 & -4/3 & 0 & 0 & -16/3 \end{bmatrix}$$

$4/3 / 0 = \infty$

$5/3 \div 4 = 5/12$

$5/3 \div 0 = \infty$

Row 2.

Finding next pivot. Max value in last row = 1

So, pivot is (2, 3).

Dividing Row 2 by 4.

$$\begin{bmatrix} 1/3 & 1 & 0 & 1/3 & 1/3 & 0 & 0 & 4/3 \\ -1/12 & 0 & 1 & 1/6 & -1/12 & 1/4 & 0 & 5/12 \\ 5/3 & 0 & 0 & 2/3 & -1/3 & 0 & 1 & 5/3 \\ 2/3 & 0 & 1 & -1/3 & -4/3 & 0 & 0 & -16/3 \end{bmatrix}$$

2/3

Performing Row 4 − Row 2.

$$\begin{bmatrix} 1/3 & 1 & 0 & 1/3 & 1/3 & 0 & 0 & 4/3 \\ -1/12 & 0 & 1 & 1/6 & -1/12 & 1/4 & 0 & 5/12 \\ 5/3 & 0 & 0 & 2/3 & -1/3 & 0 & 1 & 5/3 \\ 3/4 & 0 & 0 & -1/2 & -17/12 & -1/4 & 0 & -23/4 \end{bmatrix}$$

$4/3 \div 1/3 = 4$

$5/12 \div -1/12 = -5$

$5/3 \div 5/3 = 1$

Min value in row 2

Finding next pivot, Max value = 3/4

So, next pivot is (2, 1)

M.Inaam Ashraf (307524)

Dividing Row 2 by $-1/12$ or multiply with -12

$$\begin{bmatrix} 1/3 & 1 & 0 & 1/3 & 1/3 & 0 & 0 & 4/3 \\ 1 & 0 & -12 & -2 & 1 & -3 & 0 & -5 \\ 5/3 & 0 & 0 & 2/3 & -1/3 & 0 & 1 & 5/3 \\ 3/4 & 0 & 0 & -1/2 & -17/12 & -1/4 & 0 & -23/4 \end{bmatrix}$$

Performing. Row 1 $-$ $1/3 \times$ Row 2

Row 3 $-$ $5/3 \times$ Row 2

Row 4 $-$ $3/4 \times$ Row 2

$$\begin{bmatrix} 0 & 1 & 4 & 1 & 0 & 1 & 0 & 3 \\ 1 & 0 & -12 & -2 & 1 & -3 & 0 & -5 \\ 0 & 0 & 20 & 4 & -2 & 5 & 1 & 10 \\ 0 & 0 & 9 & 1 & -13/6 & 2 & 0 & -2 \end{bmatrix}$$

Doesn't solve either ?