MA3236 AY1819 Sem 1 Answers

Lim Li

November 20, 2020

Question 1

1. A
2. D
3. B
4. B
5. C
6. B

Question 2

(i) From $g_1(x)$, we know that $x_1^2 \leq 5$, so $-\sqrt{5} \leq x_1 \leq \sqrt{5}$. Similarly, we also know $-\sqrt{5} \leq x_2 \leq \sqrt{5}$. Hence, both x_1 and x_2 are bounded by closed sets.

We also know that $f(x)$ is continuous. And since $g_1(x)$ and $g_2(x)$ are closed, hence, the feasible set is closed and bounded.

Hence, the NLP will have an optimal solution.

(ii)

$$\nabla g_1(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \\ 0 \end{pmatrix}$$

$$\nabla g_2(x) = \begin{pmatrix} 0 \\ 3x_2 \\ -1 \end{pmatrix}$$

For the regular condition to not hold, $\nabla g_1(x)$ and $\nabla g_2(x)$ need to be linearly dependent. Since the third coordinate of g_2 is -1 while for g_1 is 0, they cannot be a scalar multiple of each other. Hence, $g_1(x) = 0$. Then, $x_1 = x_2 = 0$, which contradicts g_1. So no irregular points for this case.

Hence, regularity condition hold at every feasible point.

(iii)

$$\nabla f(x) = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$$

There exist unique λ_1, λ_2 such that

$$\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2x_1 \\ 2x_2 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 3x_2 \\ -1 \end{pmatrix} = 0$$

By looking at the third coordinate, we can conclude that $\lambda_2 = 0$.
Hence, $x_1 = -2x_2$.
Substitute this into $g_1(x)$:

$$(-2x_2)^2 + x_2^2 - 5 = 5x_2 - 5 = 0$$
$$\therefore x_2 = \pm 1$$

Hence, the KKT points are $(-2, 1, 1)$ and $(2, -1, -1)$.

(iv)

$$f((-2, 1, 1)) = -5$$
$$f((2, -1, -1)) = 5$$

Min is -5.

Question 3

(i)

$$\nabla f(x) = 2\Sigma x$$
$$\nabla g(x) = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$
$$\nabla h(x) = -v$$

If x^* is a local min, then, if $-v^T x + a = 0$, then there exist unique λ, μ such that $\mu > 0$ and

$$2\Sigma x^* + \lambda \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} - \mu v = 0$$

else if $-v^T x + a \neq 0$, then there exist unique λ such that

$$2\Sigma x^* + \lambda \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 0$$

(ii)

$$\begin{aligned} L(x, \lambda, \mu) &= f(x) + \lambda(g(x) - 1) + \mu h(x) \\ &= x^T \Sigma x + \lambda(\sum_{i=1}^{n} x_i - 1) + \mu(-v^T x + a) \end{aligned}$$

$$\theta(\lambda, \mu) = \inf\{x^T \Sigma x + \lambda(\sum_{i=1}^{n} x_i - 1) + \mu(-v^T x + a)\}$$

(iii)

$$\begin{aligned} f(x) &= x^T \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} x \\ &= 2x_1^2 + 2x_1 x_2 + 2x_2^2 \end{aligned}$$

$$g(x) = x_1 + x_2 = 1$$

$$h(x) = -2x_1 - x_2 + 3 \leq 0$$

Since $x_1 + x_2 = 1$, we substitute $x_2 = 1 - x_1$ into f and h.

$$\begin{aligned} f((x_1, 1-x_2)) &= 2x_1^2 + 2x_1(1-x_1) + 2(1-x_1)^2 \\ &= 2x_1^2 - 2x_1 + 2 \end{aligned}$$

$$\begin{aligned} h((x_1, 1-x_2)) &= -2x_1 - (1-x_1) + 3 \\ &= 2 - x_1 \leq 0 \end{aligned}$$

Hence, we want to minimize $2x_1^2 - 2x_1 + 2$ subject to $x_1 \geq 2$. Since $2x_1^2 - 2x_1 + 2$ is a quadratic with min at $x_1 = 0.5$, hence, with the constraint, the problem minimizes at $x_1 = 2$.

The solution to the problem is $(2, -1)$, and $f((2,-1)) = 6$.

(iv)

$$\theta(\lambda, \mu) = \inf\{2x_1^2 + 2x_1x_2 + 2x_2^2 + \lambda(x_1 + x_2 - 1) + \mu(-2x_1 - x_2 + 3)\}$$

Let $t(x) = 2x_1^2 + 2x_1x_2 + 2x_2^2 + \lambda(x_1 + x_2 - 1) + \mu(-2x_1 - x_2 + 3)$, which is clearly convex.

$$\nabla t(x) = \begin{pmatrix} 4x_1 + 2x_2 + \lambda - 2\mu \\ 2x_1 + 4x_2 + \lambda - \mu \end{pmatrix}$$

We solve for $\nabla t(x) = 0$,

$$4x_1 + 2x_2 + \lambda - 2\mu = 0 \qquad 2x_1 + 4x_2 + \lambda - \mu = 0$$

$$x_1 = \frac{3\mu - \lambda}{6} \qquad x_2 = \frac{-\lambda}{6}$$

Hence, we can substitute these values into $\theta(\lambda, \mu)$

$$\begin{aligned} \theta(\lambda, \mu) &= 2(\frac{3\mu-\lambda}{6})^2 + 2(\frac{3\mu-\lambda}{6})(\frac{-\lambda}{6}) + 2(\frac{-\lambda}{6})^2 + \lambda((\frac{3\mu-\lambda}{6}) + (\frac{-\lambda}{6}) - 1) + \mu(-2(\frac{3\mu-\lambda}{6}) - (\frac{-\lambda}{6}) + 3) \\ &= -\frac{\lambda^2}{6} + \frac{\lambda\mu}{2} - \lambda - \frac{\mu^2}{2} + 3\mu \\ &= \frac{1}{6}(-\lambda^2 + 3\lambda\mu - 6\lambda - 3\mu^2 + 18\mu) \end{aligned}$$

Which is concave. To find the max, we differentiate and find the stationary point

$$\nabla\theta(\lambda, \mu) = \frac{1}{6}\begin{pmatrix} -2\lambda + 3\mu - 6 \\ 3\lambda - 6\mu + 18 \end{pmatrix} = 0$$

Hence, $\lambda = 6, \mu = 6$, which also satisfies $\mu \geq 0$ condition.

Sub back into $\theta(6, 6) = 6$, which is consistent with part (iii).

Question 4

(i)

$$f(x^0) = 2$$

$$\nabla f(x) = \begin{pmatrix} 2x_1 \\ 4x_2^3 \end{pmatrix}$$

$$\nabla f(x^0) = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

Linear problem LP_1 for x^1:

$$\begin{aligned} \min \quad & z(x) = 2 + \begin{pmatrix} 2 \\ 4 \end{pmatrix}^T x \\ s.t. \quad & h_1(x) \leq 0 \\ & h_2(x) \leq 0 \\ & h_3(x) \leq 0 \end{aligned}$$

(ii)

$$L(x, \mu) = 2 + \begin{pmatrix} 2 \\ 4 \end{pmatrix}^T x + \mu_1 h_1(x) + \mu_2 h_2(x) + \mu_3 h_3(x)$$
$$\theta(\mu) = \inf_x \{L(x, \mu)\}$$

Dual problem:

$$\begin{aligned} \max \quad & \theta(\mu) \\ s.t. \quad & \mu \in \mathbb{R}^3_+ \end{aligned}$$

Question 5

(i) To prove that $g(y)$ is convex, we want to show

$$\lambda g(a) + (1-\lambda) g(b) \geq g(\lambda a + (1-\lambda) b))$$

Note that $\sup f_1(x) + \sup f_2(x) \geq \sup\{f_1(x) + f_2(x)\}$. Hence,

$$\begin{aligned} \lambda g(a) + (1-\lambda) g(b) &= \lambda \sup\{a^T x - f(x)\} + (1-\lambda)\sup\{b^T x - f(x)\} \\ &= \sup\{\lambda a^T x - \lambda f(x)\} + \sup\{(1-\lambda) b^T x - (1-\lambda) f(x)\} \\ &\geq \sup\{\lambda a^T x - \lambda f(x) + (1-\lambda) b^T x - (1-\lambda) f(x)\} \\ &= \sup\{(\lambda a + (1-\lambda) b)^T x - f(x)\} \\ &= g(\lambda a + (1-\lambda) b)) \end{aligned}$$

Hence, $g(x)$ is convex.

(ii) We want to prove that $f(x) = \sup_{y \in \mathbb{R}^n}\{y^T x - g(y)\}$.

The $\geq$ direction:

We are given that

$$\forall y \in \mathbb{R}^n, g(y) = \sup_{x \in \mathbb{R}^n} y^T x - f(x)$$

By sup property,

$$\therefore \forall x, y \in \mathbb{R}^n, g(y) \geq y^T x - f(x)$$

And by rearranging,

$$\therefore \forall x, y \in \mathbb{R}^n, f(x) \geq y^T x - g(y)$$

Hence, by sup property,

$$\therefore \forall x \in \mathbb{R}^n, f(x) \geq \sup_{y \in \mathbb{R}^n}\{y^T x - g(y)\}$$

The $\leq$ direction:

Let x_0 be any arbitrarily fixed x. Since f is convex, we know there exist a plane touching f at x_0 that is always below f. In other words, there exist a $y_0 \in \mathbb{R}^n, c \in \mathbb{R}$ such that

$$\forall x \in \mathbb{R}^n, f(x) \geq y_0^T x + c \qquad f(x_0) = y_0^T x_0 + c$$

To find $g(y_0)$:

$$\begin{aligned} g(y_0) &= \sup_{x\in\mathbb{R}^n} y_0^T x - f(x) \\ &\geq y_0^T x_0 - f(x_0) && \text{by substituting } x_0 \\ &= -c \end{aligned}$$

$$\begin{aligned} g(y_0) &= \sup_{x\in\mathbb{R}^n} y_0^T x - f(x) \\ &\leq \sup_{x\in\mathbb{R}^n} y_0^T x - (y_0^T x + c) && \text{since } f(x) \geq y_0^T x + c \\ &= -c \end{aligned}$$

$$\therefore g(y_0) = -c$$

Now, we substitute y_0 into $\sup_{y\in\mathbb{R}^n}\{y^T x_0 - g(y)\}$:

$$\begin{aligned} \sup_{y\in\mathbb{R}^n}\{y^T x_0 - g(y)\} &\geq y_0^T x_0 - g(y_0) \\ &= y_0^T x_0 + c \\ &= f(x_0) \end{aligned}$$

Hence,

$$f(x_0) \leq \sup_{y\in\mathbb{R}^n}\{y^T x_0 - g(y)\}$$

And since x_0 was arbitrarily chosen,

$$\therefore f(x) \leq \sup_{y\in\mathbb{R}^n}\{y^T x - g(y)\}$$