Exercises Lecture 19 - Solutions

Ex III.7.3

(i) suppose $\underline{x} = \alpha_1 \underline{z}_1 + (1-\alpha_1)\underline{z}_2$, $\underline{y} = \alpha_2 \underline{z}_1 + (1-\alpha_2)\underline{z}_2$ for some $\alpha_1, \alpha_2 \in [0,1]$ so $\underline{x}, \underline{y} \in L(\underline{z}_1, \underline{z}_2)$. Now let $\alpha \in [0,1]$. Then

$$\alpha \underline{x} + (1-\alpha)\underline{y} = \alpha\alpha_1 \underline{z}_1 + \alpha(1-\alpha_1)\underline{z}_2 + (1-\alpha)\alpha_2 \underline{z}_1 + (1-\alpha)(1-\alpha_2)\underline{z}_2$$

$$= (\alpha\alpha_1 + (1-\alpha)\alpha_2)\underline{z}_1 + (\alpha(1-\alpha_1) + (1-\alpha)(1-\alpha_2))\underline{z}_2 \quad \text{and}$$

$\alpha\alpha_1 + (1-\alpha)\alpha_2 \in [0,1]$ since $0 \le \alpha\alpha_1 \le \alpha$, $(1-\alpha)\alpha_2 \le (1-\alpha)$

and $\alpha(1-\alpha_1) + (1-\alpha)(1-\alpha_2) = \alpha + (1-\alpha) - (\alpha\alpha_1 + (1-\alpha)\alpha_2))$
$= 1 - (\alpha\alpha_1 + (1-\alpha)\alpha_2)$

Therefore $\alpha\underline{x} + (1-\alpha)\underline{y} \in L(\underline{z}_1, \underline{z}_2)$ and $L(\underline{z}_1, \underline{z}_2)$ is convex.

(ii) Let $\underline{x}, \underline{y} \in [\underline{a}, \underline{b}]$ so $a_i \le x_i \le b_i$, $a_i \le y_i \le b_i$ $i = 1, \dots, k$. Then for $\alpha \in [0,1]$
$a_i = \alpha a_i + (1-\alpha)a_i \le \alpha x_i + (1-\alpha)y_i \le \alpha b_i + (1-\alpha)b_i = b_i$ for $i = 1, \dots, k$ which implies $\alpha\underline{x} + (1-\alpha)\underline{y} \in [\underline{a}, \underline{b}]$ so $[\underline{a}, \underline{b}]$ is convex.

(iii) Suppose $\|\underline{x} - \underline{\mu}\|^2 \le r^2$, $\|\underline{y} - \underline{\mu}\|^2 \le r^2$ so $\underline{x}, \underline{y} \in B_r(\underline{\mu})$. Then $\|\alpha\underline{x} + (1-\alpha)\underline{y} - \underline{\mu}\|^2$
$= \|\alpha(\underline{x} - \underline{\mu}) + (1-\alpha)(\underline{y} - \underline{\mu})\|^2$ and consider the function $f(\underline{x}) = \|\underline{x}\|^2 = \sum_{i=1}^{k} x_i^2$. Then
$\left(\frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j}\right) = 2I$ which is p.d. and so

f is a convex function. Therefore

$\|\alpha(\underline{x}-\underline{\mu})+(1-\alpha)(\underline{y}-\underline{\mu})\|^2 \le \alpha\|\underline{x}-\underline{\mu}\|^2+(1-\alpha)\|\underline{y}-\underline{\mu}\|^2$
$\le \alpha r^2+(1-\alpha)r^2 = r^2$ and so $\alpha\underline{x}+(1-\alpha)\underline{y} \in B_r(\underline{\mu})$
and the ball is convex.

(iv) Suppose $\underline{x} = \underline{\mu} + \Sigma^{1/2}\underline{z}_1$, $\underline{y} = \underline{\mu} + \Sigma^{1/2}\underline{z}_2$ where
$\underline{z}_1, \underline{z}_2 \in B_r(\underline{0})$. Then $\alpha\underline{x}+(1-\alpha)\underline{y}$

$= \alpha\underline{\mu} + (1-\alpha)\underline{\mu} + \Sigma^{1/2}(\alpha\underline{z}_1 + (1-\alpha)\underline{z}_2)$

$= \underline{\mu} + \Sigma^{1/2}(\alpha\underline{z}_1 + (1-\alpha)\underline{z}_2) \in E_r(\underline{\mu}, \Sigma)$ since

$\alpha\underline{z}_1 + (1-\alpha)\underline{z}_2 \in B_r(\underline{0})$ by part (iii).

Therefore $E_r(\underline{\mu}, \Sigma)$ is convex.

(v) $f(\alpha\underline{x}_1 + (1-\alpha)\underline{x}_2) = \underline{a} + \underline{c}'(\alpha\underline{x}_1 + (1-\alpha)\underline{x}_2)$

$= \alpha\underline{a} + (1-\alpha)\underline{a} + \alpha\underline{c}'\underline{x}_1 + (1-\alpha)\underline{c}'\underline{x}_2$

$= \alpha f(\underline{x}_1) + (1-\alpha)f(\underline{x}_2)$ and so f is convex.

(vi) $\dfrac{\partial^2 f(x)}{\partial x^2} = \left(-\dfrac{1}{x}\right)' = \dfrac{1}{x^2} > 0$ when $x > 0$

and so f is convex.

(viii) In (iii) we proved that $f(\underset{\sim}{x}) = \underset{\sim}{x}'\underset{\sim}{x}$ is convex. Now $\underset{\sim}{x}'\Sigma\underset{\sim}{x} = f(\Sigma^{1/2}\underset{\sim}{x})$ and so

$$(\alpha \underset{\sim}{x}_1 + (1-\alpha)\underset{\sim}{x}_2)'\,\Sigma\,(\alpha \underset{\sim}{x}_1 + (1-\alpha)\underset{\sim}{x}_2)$$
$$= f(\Sigma^{1/2}(\alpha \underset{\sim}{x}_1 + (1-\alpha)\underset{\sim}{x}_2)) = f(\alpha \Sigma^{1/2}\underset{\sim}{x}_1 + (1-\alpha)\Sigma^{1/2}\underset{\sim}{x}_2)$$
$$\le \alpha f(\Sigma^{1/2}\underset{\sim}{x}_1) + (1-\alpha) f(\Sigma^{1/2}\underset{\sim}{x}_2)$$
$$= \alpha \underset{\sim}{x}_1'\Sigma\underset{\sim}{x}_1 + (1-\alpha)\underset{\sim}{x}_2'\Sigma\underset{\sim}{x}_2$$

and so the function $f(\underset{\sim}{x}) = \underset{\sim}{x}'\Sigma\underset{\sim}{x}$ is convex.

Exercise IV.7.6 Suppose $\underset{\sim}{x}_1, \underset{\sim}{x}_2 \in C_1 \cap C_2$.

Then $\underset{\sim}{x}_1, \underset{\sim}{x}_2 \in C_i$ and for $\alpha \in [0,1]$ we have $\alpha \underset{\sim}{x}_1 + (1-\alpha)\underset{\sim}{x}_2 \in C_i$ by convexity of C_i for $i = 1, 2$. But this implies that $\alpha \underset{\sim}{x}_1 + (1-\alpha)\underset{\sim}{x}_2 \in C_1 \cap C_2$ and so $C_1 \cap C_2$ is convex.

Ex IV.7.7 Let $\underset{\sim}{x}, \underset{\sim}{y} \in C_a$ so

$\underset{\sim}{x} = \underset{\sim}{a} + B\underset{\sim}{c}_1$, $\underset{\sim}{y} = \underset{\sim}{a} + B\underset{\sim}{c}_2$ for $\underset{\sim}{c}_1, \underset{\sim}{c}_2 \in C$.
Then for $\alpha \in [0,1]$ $\alpha\underset{\sim}{x} + (1-\alpha)\underset{\sim}{y}$

$$= \alpha \underset{\sim}{a} + (1-\alpha)\underset{\sim}{a} + \alpha B\underset{\sim}{c}_1 + (1-\alpha)B\underset{\sim}{c}_2$$
$$= \underset{\sim}{a} + B(\alpha \underset{\sim}{c}_1 + (1-\alpha)\underset{\sim}{c}_2) \in C_a$$

since $\alpha \underset{\sim}{c}_1 + (1-\alpha)\underset{\sim}{c}_2 \in C$.

Ex IV.7.8 Suppose $\underset{\sim}{x}, \underset{\sim}{y} \in C$. Then $a\underset{\sim}{x} + b\underset{\sim}{y} \in C$ $\forall a, b \in R^1$ and in particular $\alpha\underset{\sim}{x} + (1-\alpha)\underset{\sim}{y} \in C$ $\forall \alpha \in [0,1]$ so C is convex.

Ex. III.7.9 (mislabeled as 7.8)

Suppose $(x_1, y_1), (x_2, y_2) \in S$. Then
$\alpha(x_1, y_1) + (1-\alpha)(x_2, y_2) = (\alpha x_1 + (1-\alpha)x_2, \alpha y_1 + (1-\alpha)y_2)$

and $\alpha x_1 + (1-\alpha)x_2 \in C$ since C is convex
and $f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha)f(x_2)$
$\le \alpha y_1 + (1-\alpha) y_2$
by the convexity of f and the definition of S.

Ex. III.7.10 (mislabeled as 7.9)

$$p(x) = (2\pi\sigma_1^2)^{-1/2} \exp\{-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\}$$

$$q(x) = (2\pi\sigma_2^2)^{-\frac{1}{2}} \exp\{-\frac{1}{2\sigma_2^2}(x-\mu_2)^2\}$$

So $-\log \frac{q(x)}{p(x)} = -\log\left[\left(\frac{\sigma_1^2}{\sigma_2^2}\right)^{1/2} \exp\{-\frac{1}{2\sigma_2^2}(x-\mu_2)^2 + \frac{1}{2\sigma_1^2}(x-\mu_1)^2\}\right]$

$$= \log\frac{\sigma_2}{\sigma_1} + \frac{1}{2\sigma_2^2}(x-\mu_2)^2 - \frac{1}{2\sigma_1^2}(x-\mu_1)^2$$

Therefore $KL(P \| Q) = E_p\left(-\log\frac{q(x)}{p(x)}\right)$

$$= \log\frac{\sigma_2}{\sigma_1} + \frac{1}{2\sigma_2^2}E_p\left((X-\mu_1+\mu_1-\mu_2)^2\right) - \frac{1}{2}$$

$$= \log\frac{\sigma_2}{\sigma_1} + \frac{1}{2\sigma_2^2}E_p\left((x-\mu_1)^2 + 2(x-\mu_1)(\mu_1-\mu_2) + (\mu_1-\mu_2)^2\right) - \frac{1}{2}$$

$$= \log\frac{\sigma_2}{\sigma_1} + \frac{1}{2\sigma_2^2}\left(\sigma_1^2 + 0 + (\mu_1-\mu_2)^2\right) - \frac{1}{2}$$

$$= \log\frac{\sigma_2}{\sigma_1} + \frac{1}{2}\frac{\sigma_1^2}{\sigma_2^2} + \frac{(\mu_1-\mu_2)^2}{2\sigma_2^2} - \frac{1}{2}$$

Ex $\overline{\text{III}}$.7.11 Using P, Q as in Ex. $\overline{\text{III}}$.7.10 we see that generally $KL(\mathbb{P} \| Q) \neq KL(Q \| \mathbb{P})$ and so KL is not symmetric.