Solutions to Part A of Problem Sheet 3

Solution (3.1)

(a) We need to show that the convex combination of points in $B(\boldsymbol{p}, r)$ is again in $B(\boldsymbol{p}, r)$. Let $\boldsymbol{x}, \boldsymbol{y} \in B(\boldsymbol{p}, r)$ and $\lambda \in [0, 1]$. Then

$$\begin{aligned}
\|\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} - \boldsymbol{p}\| &= \|\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} - (\lambda\boldsymbol{p} + (1-\lambda)\boldsymbol{p})\| \\
&= \|\lambda(\boldsymbol{x}-\boldsymbol{p}) + (1-\lambda)(\boldsymbol{y}-\boldsymbol{p})\| \\
&\leq \|\lambda(\boldsymbol{x}-\boldsymbol{p})\| + \|(1-\lambda)(\boldsymbol{y}-\boldsymbol{p})\| \\
&= \lambda\|\boldsymbol{x}-\boldsymbol{p}\| + (1-\lambda)\|\boldsymbol{y}-\boldsymbol{p}\| \\
&\leq \lambda r + (1-\lambda)r = r.
\end{aligned}$$

Therefore, $\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in B(\boldsymbol{p}, r)$.

(b) Let $\boldsymbol{x}, \boldsymbol{y} \in C^*$. Then for all $\boldsymbol{z} \in C$,

$$\langle \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}, \boldsymbol{z}\rangle = \lambda\langle \boldsymbol{x}, \boldsymbol{z}\rangle + (1-\lambda)\langle \boldsymbol{y}, \boldsymbol{z}\rangle \leq (1-\lambda) + \lambda = 1.$$

Therefore, $\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in C^*$.

(c) Denote $B_p = B(\boldsymbol{0}, 1)$ the unit ball with respect to the p-norm, where $p \in \{1, 2, \infty\}$. For the 2-norm, the polar is

$$B_2^* = \{\boldsymbol{y} \in \mathbb{R}^n : \forall \boldsymbol{x}, \|\boldsymbol{x}\|_2 \leq 1 \Rightarrow \langle \boldsymbol{x}, \boldsymbol{y}\rangle \leq 1\}.$$

We claim that $B_2^* = B_2$. To show that $B_2 \subseteq B_2^*$, let $\boldsymbol{y} \in B_2$, so that $\|\boldsymbol{y}\|_2 \leq 1$. By the Cauchy-Schwarz inequality (or the characterization $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2\cos(\theta)$), for all $\boldsymbol{x} \in B_2$,

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle \leq \|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2 \leq 1,$$

so that $\boldsymbol{y} \in B_2^*$. To show the converse inclusion $B_2^* \subseteq B_2$, note that for any $\boldsymbol{y} \in B_2^*$ we have

$$\|\boldsymbol{y}\|_2 = \langle \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|_2}, \boldsymbol{y}\rangle \leq 1,$$

since $\boldsymbol{y}/\|\boldsymbol{y}\|_2 \in B_2$, from which $\boldsymbol{y} \in B_2$ follows.

For the 1-norm,

$$B_1^* = \{\boldsymbol{y} \in \mathbb{R}^n : \forall \boldsymbol{x}, \sum_{i=1}^n |x_i| \leq 1 \Rightarrow \sum_{i=1}^n x_i y_i \leq 1\}.$$

We claim that $B_1^* = B_\infty$. For the inclusion $B_\infty \subseteq B_1^*$, let $\boldsymbol{y} \in B_\infty$, i.e., $\|\boldsymbol{y}\|_\infty = \max_{1\leq i\leq n} |y_i| \leq 1$. Then for all $\boldsymbol{x} \in B_1$, i.e., with $\sum_{i=1}^n |x_i| \leq 1$, we clearly have

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \sum_{i=1}^n x_i y_i \leq \sum_{i=1}^n x_i \leq 1,$$

so that $\boldsymbol{y} \in B_1^*$. Now let $\boldsymbol{y} \in B_1^*$. For every $1 \leq i \leq n$ and $\boldsymbol{x} = \pm \boldsymbol{e}_i = (0, \dots, \pm 1, \dots, 0)^\top$ (± 1 in i-th coordinate) we have that $\|\boldsymbol{x}\|_1 = 1$, and so

$$\pm y_i = \langle \boldsymbol{x}, \boldsymbol{y} \rangle \leq 1,$$

from which $\|\boldsymbol{y}\|_\infty \leq 1$ and $\boldsymbol{y} \in B_\infty$ follows.

For the ∞-norm,

$$B_\infty^* = \{\boldsymbol{y} \in \mathbb{R}^n : \forall \boldsymbol{x}, \max_{1 \leq i \leq n} \leq 1 \Rightarrow \sum_{i=1}^n x_i y_i \leq 1\}.$$

For the inclusion $B_1 \subseteq B_\infty$, let $\boldsymbol{y} \in B_1$, i.e., $\sum_{i=1}^n |y_i| \leq 1$. Then for all $\boldsymbol{x}$ with $\max_{1 \leq i \leq n} |x_i| \leq 1$, we have that

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^n x_i y_i \leq \max_{1 \leq i \leq n} |x_i| \sum_{i=1}^n y_i \leq 1,$$

so that $\boldsymbol{y} \in B_\infty^*$. To show $B_\infty^* \subseteq B_1$, let $\boldsymbol{y} \in B_\infty^*$. Let $\boldsymbol{x} = \mathrm{sign}(\boldsymbol{y})$ be the vectors with $x_i = \mathrm{sign}(y_i)$. Then $\|\boldsymbol{x}\|_\infty = 1$, and

$$\sum_{i=1}^n |y_i| = \sum_{i=1}^n x_i y_i = \langle \boldsymbol{x}, \boldsymbol{y} \rangle \leq 1,$$

so that $\boldsymbol{y} \in B_1$.

Solution (3.2) Consider the set $C - D$. Since C and D are disjoint, $\mathbf{0} \notin C - D$. If C and D are bounded and, say, contained in balls of radius r_1 and r_2, then $C - D$ is contained in a ball of radius $r_1 + r_2$ (since $\|\boldsymbol{x} - \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$), and bounded. Therefore there exists a hyperplane H such that $C - D \in \mathrm{int} H_-$ and $\mathbf{0} \in \mathrm{int} H_+$.

For an example where the statement fails if C and D are not bounded, consider

$$C = \{\boldsymbol{x} \in \mathbb{R}^2 : x_1 \leq 0\}, \quad D = \{\boldsymbol{x} \in \mathbb{R}^2 : x_1 x_2 \geq 1\}.$$

Any affine hyperplane that does not touch C has to be of the form $\{\boldsymbol{x} : x_1 = a\}$ for

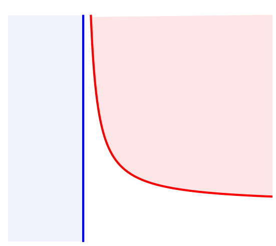

Figure 1: Non-strict separation

$a > 0$ (a vertical line), but any such hyperplane touches D at the point $\boldsymbol{x} = (a, 1/a)$. However, both sets are clearly disjoint. The only separating hyperplane is $\{\boldsymbol{x} : x_1 = 0\}$, but this is not a strict separation.

Solution (3.3)

(a) We work in two dimension, the general case is mathematically the same, and assume grid length $\ell = 1$. If the points $\boldsymbol{x}$ and $\boldsymbol{y}$ are s horizontal units and t vertical units away, then any path from $\boldsymbol{x}$ to $\boldsymbol{y}$ has to move s units to the left and t units up (assuming that $\boldsymbol{y}$ is to the north-east of $\boldsymbol{x}$). The distance in the 1-norm is

$$\|\boldsymbol{y}-\boldsymbol{x}\|_1 = |y_1 - x_1| + |y_2 - x_2| = s + t.$$

(b) The problem here is that the objective function

$$\sum_{i=1}^{N} \|\boldsymbol{p}_i - \boldsymbol{x}\|_1 = \sum_{i=1}^{N} |p_{i,1} - x_1| + |p_{i,2} - x_2| \tag{1}$$

is not a linear function. To get a linear function, we have to find a way to get rid of the absolute values. To do this, we note that

$$|x| \leq t \Leftrightarrow -t \leq x \leq t$$

holds for any numbers x, t with ≥ 0. For the 1-norm, we get the equivalence

$$\|\boldsymbol{x}\|_1 \leq t \Leftrightarrow \exists t_1, \dots, t_n \geq 0, -t_i \leq x_i \leq t_i \text{ and } \sum_{i=1}^{n} t_i \leq t.$$

A minimization problem of the form

$$\text{minimize} \quad \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$$

can therefore be written as

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{n} t_i \\ \text{subject to} \quad & -t_i \leq x_i \leq t_i \\ & t_i \geq 0 \\ & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}. \end{aligned}$$

This is a linear programming problem with twice as many variables as the original problem. We can apply this to the objective function (1), which gives the form

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{n} t_i + \sum_{j=1}^{n} s_j \\ \text{subject to} \quad & t_i, s_j \geq 0 \\ & -t_i \leq p_{i,1} - x_1 \leq t_i \\ & -s_j \leq p_{j,2} - x_2 \leq s_j \\ & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}. \end{aligned}$$