

Solutions to Part A of Problem Sheet 3

Solution (3.1)

- (a) Let $\mathbf{x}, \mathbf{y} \in B(\mathbf{p}, r)$ and $\lambda \in [0, 1]$. Then

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

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

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

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

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

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

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

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

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

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

$$\|\mathbf{y}\|_2 = \left\langle \frac{\mathbf{y}}{\|\mathbf{y}\|_2}, \mathbf{y} \right\rangle \leq 1,$$

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

For the 1-norm,

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

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

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

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

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

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

For the ∞ -norm,

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

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

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

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

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

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

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 $\|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$), and bounded. Therefore there exists a hyperplane H such that $C - D \in \text{int} H_-$ and $\mathbf{0} \in \text{int} H_+$.

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

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

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

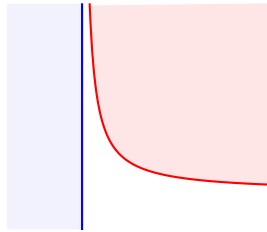


Figure 1: Non-strict separation

$a > 0$ (a vertical line), but any such hyperplane touches D at the point $\mathbf{x} = (a, 1/a)$. However, both sets are clearly disjoint. The only separating hyperplane is $\{\mathbf{x} : x_1 = 0\}$, 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 \mathbf{x} and \mathbf{y} are s horizontal units and t vertical units away, then any path from \mathbf{x} to \mathbf{y} has to move s units to the left and t units up (assuming that \mathbf{y} is to the north-east of \mathbf{x}). The distance in the 1-norm is

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

- (b) The problem here is that the objective function

$$\sum_{i=1}^N \|\mathbf{p}_i - \mathbf{x}\|_1 = \sum_{i=1}^N |p_{i,1} - x_1| + |p_{i,2} - x_2| \quad (1)$$

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 $t \geq 0$. For the 1-norm, we get the equivalence

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

A minimization problem of the form

$$\text{minimize } \|\mathbf{x}\|_1 \quad \text{subject to } \mathbf{Ax} \leq \mathbf{b}$$

can therefore be written as

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

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

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