TMA4180
Optimisation I
Spring 2018

Norwegian University of Science
and Technology
Department of Mathematical
Sciences

Solutions to exercise set 7

1 We will consider sets $\Omega_1 = \{\, x \in \mathbb{R}^n \mid |x_i| \le 1 \,\} = \{\, x \in \mathbb{R}^n \mid \|x\|_\infty \le 1 \,\}$, $\Omega_2 = \{\, x \in \mathbb{R}^n \mid \|x\|_2 \le 1 \,\}$, and $\Omega_3 = \{\, x \in \mathbb{R}^n \mid \|x\|_2 = 1 \,\}$.

a) Show that the sets Ω_1 and Ω_2 are non-empty, convex and closed.

Solution: $0 \in \Omega_i$, $i = 1, 2$ so the sets are non-empty. Since both sets are unit balls with respect to to some norm, we have for arbitrary $x, y \in \Omega_i$, $i = 1, 2$, $\lambda \in [0, 1]$: $\|\lambda x + (1 - \lambda)y\| \le \|\lambda x\| + \|(1 - \lambda)y\| = \lambda\|x\| + (1 - \lambda)\|y\|$; thereby we have shown convexity.

Closedness follows immediately from the continuity of the norm.

b) In 2D, determine the normal cones $N_{\Omega_i}(x)$ and radial cones (cones of feasible directions) $R_{\Omega_i}(x)$ at $x^1 = (0, 1)$ for $i = 1, 2$ In case of Ω_2 do the same at $x^2 = (1, 1)$.

It may be easier to start with a sketch.

Solution: $(0 + \epsilon p_1, 1 + \epsilon p_2) \in \Omega_1$, $\forall 0 < \epsilon < \hat{\epsilon}$ and some $\hat{\epsilon} > 0$ $\iff$ $|\epsilon p_1| \le 1$ & $|1 + \epsilon p_2| \le 1$, $\forall 0 < \epsilon < \hat{\epsilon}$ and some $\hat{\epsilon} > 0$ $\iff$ $p_2 \le 0$. Therefore, $R_{\Omega_1}(0, 1) = \{\, p \in \mathbb{R}^2 \mid p_2 \le 0 \,\}$.

$(1 + \epsilon p_1, 1 + \epsilon p_2) \in \Omega_1$, $\forall 0 < \epsilon < \hat{\epsilon}$ and some $\hat{\epsilon} > 0$ $\iff$ $|1 + \epsilon p_1| \le 1$ & $|1 + \epsilon p_2| \le 1$, $\forall 0 < \epsilon < \hat{\epsilon}$ and some $\hat{\epsilon} > 0$ $\iff$ $p_1 \le 0$ & $p_2 \le 0$. Therefore, $R_{\Omega_1}(1, 1) = \{\, p \in \mathbb{R}^2 \mid p_1 \le 0$ & $p_2 \le 0 \,\} = -\mathbb{R}_+^2$.

$(0 + \epsilon p_1, 1 + \epsilon p_2) \in \Omega_2$, $\forall 0 < \epsilon < \hat{\epsilon}$ and some $\hat{\epsilon} > 0$ $\iff$ $|\epsilon p_1|^2 + |1 + \epsilon p_2|^2 = 1 + 2\epsilon p_2 + \epsilon^2(p_1^2 + p_2^2) \le 1^2 = 1$, $\forall 0 < \epsilon < \hat{\epsilon}$ and some $\hat{\epsilon} > 0$. Note that in contrast with Ω_1-case, the latter reuirements hold if either $p_1 = 0$ and $p_2 = 0$ or $p_1 \ne 0$ and $p_2 < 0$! Therefore, $R_{\Omega_2}(0, 1) = \{0\} \cup \{\, p \in \mathbb{R}^2 \mid p_2 < 0 \,\}$. This cone is neither open nor closed, and is convex despite the "appendage" 0 "sticking out" (see alst Exercise 2).

$n \in N_{\Omega_1}(0, 1)$ $\iff$ $n_1(x_1 - 0) + n_2(x_2 - 1) \le 0$, $\forall -1 \le x_1 \le 1, -1 \le x_2 \le 1$. The last system of inequalities implies $n_1 = 0$ (e.g. take $x_2 = 1$ and $x_1 = \pm 1$ to check the necessity) and $n_2 \le 0$ (e.g. take $x_1 = 0$ and $x_2 = 1$). On the other hand, it is easy to see that any $n = (0, n_1)$ with $n_1 \ge 0$ satisfies the latter system of inequalities (recal that $x_2 - 1 \ge 0$, $\forall x \in \Omega_2$). Therefore $N_{\Omega_1}(0, 1) = \{0\} \times \mathbb{R}_+ \subset \mathbb{R}^2$.

Similar analysis shows that $N_{\Omega_1}(1, 1) = \mathbb{R}_+^2$.

Finally, intutively (or graphically) it should be clear that $N_{\Omega_2}(0, 1) = \{0\} \times \mathbb{R}_+ \subset \mathbb{R}^2$.

Indeed, if $n \in \{0\} \times \mathbb{R}_+$ then $\forall x \in \Omega_2$: $0(x_1 - 0) + n_2(x_2 - 1) \le 0$ because $x_2 \le 1$. Thus $N_{\Omega_2}(0, 1) \supseteq \{0\} \times \mathbb{R}_+$.

To show the opposite inclusion we consider the inequality $\forall x \in \Omega_2$: $n_1(x_1 - 0) + n_2(x_2 - 1) \leq 0$. Taking $x_1 = 0$ and $x_2 = 0$ implies $n_2 \geq 0$. Taking $x_1 = \pm\epsilon$, $x_2 = (1 - \epsilon^2)^{1/2}$, and considering the limit when $\epsilon \downarrow 0$ we get $\pm\epsilon n_1 + ((1 - \epsilon^2)^{1/2} - 1)n_2 = \pm\epsilon n_1 + (1 - \epsilon^2 - 1^2)/((1 - \epsilon^2)^{1/2} + 1)n_2 \leq 0$

$$\lim_{\epsilon \downarrow 0}[\pm n_1 - \frac{\epsilon}{(1 - \epsilon^2)^{1/2} + 1}n_2] = \pm n_1 \leq 0,$$

and therefore $n_2 = 0$.

c) Show that the radial cone $R_{\Omega_3}(x) = \{0\}$ for all $x \in \Omega_3$. Owing to the symmetry it is sufficient to analyse only one such point.

Solution: Indeed, take e.g. $x = (0, 1)$. Then fix an arbitrary $p \in \mathbb{R}^2$. If $x + \epsilon p \in \Omega_3$, $\forall 0 < \epsilon < \hat{\epsilon}$, then $\|x + \epsilon p\|_2^2 = \epsilon^2 p_1^2 + (1 + \epsilon p_2)^2 = 1 + 2\epsilon p_2 + \epsilon^2(p_1^2 + p_2^2) = 1^2 = 1 \iff 2p_2 + \epsilon^2(p_1^2 + p_2^2) = 1$, $\forall 0 < \epsilon < \hat{\epsilon}$. Therefore $2p_2 = 0$ (simply take $\epsilon \to 0$). Dividing by ϵ once again we prove that $p_1^2 = 0$.

d) Consider the projection problem

$$\pi_{\Omega_i}(z) = \arg\min_{y \in \Omega_i} f(y) := \frac{1}{2}\|y - z\|_2^2.$$

Use the necessary and sufficient optimality conditions $\nabla f(\pi_{\Omega_i}(z))^\top(x - \pi_{\Omega_i}(z)) \geq 0$, $\forall x \in \Omega$, to verify the projection formulae $\pi_{\Omega_1}(z) = \min\{1, \max\{-1, z\}\}$, where min and max is applied component-wise to the vector, and $\pi_{\Omega_2}(z) = z/\max\{1, \|z\|_2\}$.

Solution: We will use the notation $x^* = \pi_{\Omega_i}(z)$. In either case $(i = 1, 2)$, the necessary and sufficient (because of convexity) optimality conditions are $x^* \in \Omega_i$ and $\nabla f(x^*)^\top(x - x^*) = (x^* - z)^\top(x - x^*) \geq 0$, $\forall x \in \Omega_i$.
In case of Ω_1 we get:

$$\sum_{i=1}^{n}(x_i^* - z_i) + (x_i - x_i^*) \geq 0, \quad \forall -1 \leq x_i \leq 1.$$

First of all, we can analyse this sum term by term, by simply selecting $x = (x_1^*, \ldots, x_{i-1}^*, x_i, x_{i+1}^*, \ldots), x_n^*)$. Then all the terms in the sum above vanish, except for the ith. Obviously the selection of the ith coordinate of x_i^* only affects the ith term in the sum. We will make such a selection, that each term in this sum is guaranteed to be non-negative. *Case 1*: $-1 < z_i < 1$. In this case we put $x_i^* = z_i$ and the ith term in the sum above vanishes (hence non-negative). *Case 2*: $z_i \leq -1$. Then we can put $x_i^* = -1$. As a result $x_i^* - z_i \geq 0$ while $x_i - x_i^* \geq 0$, $\forall -1 \leq x_i \leq 1$. Therefore the product of the two is non-negative. *Case 3*: $z_i \geq 1$. Then we can put $x_i^* = 1$. As a result $x_i^* - z_i \leq 0$ while $x_i - x_i^* \leq 0$, $\forall -1 \leq x_i \leq 1$. Therefore the product of the two is non-negative. Thus we have componentwise computed x_i^* given z_i. The three cases can be summarized as $x_i^* = \min\{1, \max\{-1, z_i\}\}$.

Let us now deal with projections onto Ω_2. Clearly if $\|z\| \leq 1$ then we can put $x^* = z$, and the optimality conditions will be satisfied (left hand side of the inequality will be identically equal to zero). If, on the other hand $\|z\| > 1$, we put $x^* = z/\|z\|$. Then $\nabla f(x^*)^\top(x - x^*) = (1 - \|z\|^{-1})(x^\top x^* - \|x^*\|^2) \geq 0$,

because $1 - \|z\|^{-1} < 0$ by assumption, and Cauchy–Schwarz inequality implies that $x^\top x^* \le \|x\|\|x^*\| = \|x\| \le 1 = \|x^*\|^2$.

In summary, $x^* = z/\max\{1, \|z\|\}$ satisfies the optimality conditions.

e) Consider a two-dimensional situation, and put $f(x) = x_1^2 + (x_2 + 2)^2$. Let $x^{(0)} = (1, -1)$. Find the point of global minimum for this function over Ω_1. (you can do this graphically). Compute one step of the projected gradient method $x^{(1)} = \pi_{\Omega_1}[x^{(0)} - \alpha \nabla f(x^{(0)})]$ using steplength $\alpha = 1$.

Solution: We are looking for the point in Ω_1 which is closest to $z = (0, -2)$. This is evidently $x^* = (0, -1)$, which is a projection of z onto Ω_1.

We now compute the projected gradient step: $\nabla f(x^{(0)}) = (2x_1^{(0)}, 2(x_2^{(0)} + 2))^\top = (2, 2)^\top$. Therefore $\pi_{\Omega_1}[x^{(0)} - \nabla f(x^{(0)})] = \pi_{\Omega_1}[(1, -1) - (2, 2)] = \pi_{\Omega_1}[(-1, -3)] = \max\{-1, \min\{1, (-1, -3)\}\} = (-1, -1)$.

---

2 Consider a convex set $\Omega \subset \mathbb{R}^n$ and a point $x \in \mathbb{R}^n$. Show that the radial cone (cone of feasible directions) $R_\Omega(x)$ is convex.

Solution: Suppose that $p_1, p_2 \in R_\Omega(x)$, and let $p = \lambda p_1 + (1 - \lambda)p_2$, $\lambda \in [0, 1]$. We would like to show that $p \in R_\Omega(x)$.

By definition of $R_\Omega(x)$, $x + \epsilon p_1 \in \Omega$, $\forall 0 < \epsilon < \epsilon_1$, and $x + \epsilon p_2 \in \Omega$, $\forall 0 < \epsilon < \epsilon_2$. Owing to the convexity of Ω we have $x + \epsilon p = \lambda(x + \epsilon p_1) + (1 - \lambda)(x + \epsilon p_2) \in \Omega$, $\forall 0 < \epsilon < \min\{\epsilon_1, \epsilon_2\}$. Therefore, $p \in R_\Omega(x)$.

---

3 **a)** Consider a set Ω described using linear *equality* constraints, $\Omega = \{z \in \mathbb{R}^n \mid a_i^\top z = b_i\}$, $i = 1, \ldots, m$. Express the radial cone $R_\Omega(x)$ and the normal cone $N_\Omega(x)$ using the vectors $a_i \in \mathbb{R}^n$, $i = 1, \ldots, m$ for $x \in \Omega$.

Solution: If $x + \epsilon p \in \Omega$ for some $\epsilon \neq 0$, then $a_i^\top p = 0$, $i = 1, \ldots, m$, and vice versa. Therefore $R_\Omega(x) = \mathrm{span}[a_1, \ldots, a_m]^\perp = \cap_{i=1}^m \ker a_i^\top \equiv R_\Omega$.

Note that $\Omega = x + R_\Omega$: this is the usual superposition principle in linear algebra (all solutions to the non-homogeneous linear system are expressable as a sum of some solution to the non-homogeneous linear system plus all solutions to the homogeneous linear system), and that R_Ω is a linear subspace of $\mathbb{R}^n$. Therefore, $p^\top(y - x) \le 0$, $\forall y \in \Omega$ iff $p^\top d \le 0$, $\forall d \in R_\Omega$. But if $d \in R_\Omega$, so is $-d$, meaning that $p^\top(y - x) \le 0$, $\forall y \in \Omega$ iff $p \perp R_\Omega = \mathrm{span}[a_1, \ldots, a_m]^\perp$. Therefore p must be expressible as a linear combination of a_i, $i = 1, \ldots m$; in other words $\exists \lambda \in \mathbb{R}^m : p = \sum_{i=1}^m \lambda_i a_i$. Thus $N_\Omega(x) \equiv N_\Omega = \mathrm{span}[a_1, \ldots, a_m]$.

b) Consider a set Ω described using linear *inequality* constraints, $\Omega = \{z \in \mathbb{R}^n \mid a_i^\top z \le b_i\}$, $i = 1, \ldots, m$. Given $x \in \Omega$ let $\mathcal{A}(x) \subset \{1, \ldots, m\}$ be the set of active, or binding constraints: $a_i^\top x = b_i$, $i \in \mathcal{A}(x)$, $a_i^\top x < b_i$, $i \notin \mathcal{A}(x)$. Express the radial cone $R_\Omega(x)$ using the vectors $a_i \in \mathbb{R}^n$, $i = 1, \ldots, m$, and $\mathcal{A}(x)$. Show that $N_\Omega(x) \supseteq \{\sum_{i \in \mathcal{A}} \lambda_i a_i \mid \lambda_i \ge 0\}$. (In fact the last inclusion is equality, but it is a bit more difficult to prove.)

Solution: Note first that for an arbitrary $p \in \mathbb{R}^n$ and $\forall i \notin \mathcal{A}(x)$ there is $\epsilon_{p,i} > 0$ such that the inequality $a_i^\top(x + \epsilon p) < b_i$, holds $\forall |\epsilon| < \epsilon_{i,p}$: this is owing

to the continuity of the function $x \mapsto a_i^\top x$. Thus as far as the cone of feasible directions is concerned, inactive constraints can be safely ignored.

For the active constraints, on the other hand, if $p \in R_\Omega(x)$ then we must have the inequality $a_i^\top(x + \epsilon p) = a_i^\top x + \epsilon a_i^\top p = b_i + \epsilon a_i^\top p \leq b_i$, $\forall 0 < \epsilon < \epsilon_{i,p}$, which happens iff $a_i^\top p \leq 0$.

Therefore, the cone of feasible directions is found as $R_\Omega(x) = \{\, p \in \mathbb{R}^n \mid a_i^\top p \leq 0, i \in \mathcal{A}(x) \,\}$.

To determine the normal cone we first observe that $\Omega - x \subseteq R_\Omega(x)$, owing to the convexity of Ω: indeed for any $y \in \Omega$ the line segment between x and y is in Ω, and therefore $x + \epsilon(y - x) \in \Omega$, $\forall 0 < \epsilon < 1$. Per definition, $y - x \in R_\Omega(x)$.

Now consider $p = \sum_{i \in \mathcal{A}} \lambda_i a_i$, where $\lambda_i \geq 0$. Then $\forall d \in R_\Omega(x) \supseteq \Omega - x$: $p^\top d = \sum_{i \in \mathcal{A}} \lambda_i a_i^\top d \leq 0$ and therefore $N_\Omega(x) \supseteq \{\, \sum_{i \in \mathcal{A}} \lambda_i a_i \mid \lambda_i \geq 0 \,\}$.