Convex Analysis, Sect. 2.1

Exercise 1

Show that $N_C\bar{x} = \{\, d \in E \mid \langle d, x - \bar{x}\rangle \le 0, \forall x \in C \,\}$ is a closed set.

Proof: Define $f_x : E \to \mathbb{R}$ by $f_x(d) = \langle d, x - \bar{x}\rangle$. Then f_x is continuous (recall, the norm in E is defined by the inner product $\langle \cdot, \cdot \rangle$) and therefore the preimage $f_x^{-1}([0, +\infty))$ of the closed set $[0, +\infty) \subset \mathbb{R}$ is closed in E. As a result, the intersection of these sets is also closed:

$$N_C(\bar{x}) = \cap_{x \in C} f_x^{-1}([0, +\infty)).$$

□

Exercise 2

Check that C is convex and compute the normal cone $N_C(\bar{x})$, $\bar{x} \in C$.

(a) $C = [a, b] \subset \mathbb{R}$.

Solution: Direct computation shows:

$$N_C(\bar{x}) = \begin{cases} \{0\}, & \text{if } \bar{x} \in (a, b), \\ [0, +\infty), & \text{if } \bar{x} = b, \\ (-\infty, 0], & \text{if } \bar{x} = a. \end{cases}$$

□

(b) C - unit ball in E.

Solution: If $\bar{x} \in \text{int}C$ then for some small $\epsilon > 0$ we have $\bar{x} + \epsilon C \subset C$. Therefore, the inequality $\langle d, x - \bar{x}\rangle \le 0$ for all $x \in C$ implies that $\langle d, \pm\epsilon e_i\rangle \le 0$, where e_i are basis vectors in E, and consequently $d = 0$. Therefore, $N_C(\bar{x}) = \{0\}$ in this case.

If, on the other hand, $\bar{x} = 1$, then $N_C(\bar{x}) = \bar{x}\mathbb{R}_+$.

Indeed, for every $\alpha \ge 0$ and $x \in C$ we have the inequalities

$$\langle \alpha\bar{x}, x - \bar{x}\rangle = \alpha[\underbrace{\langle \bar{x}, x\rangle}_{\le \|\bar{x}\|\|x\| \le 1} - \underbrace{\langle \bar{x}, \bar{x}\rangle}_{=1}] \le \alpha[1 - 1] \le 0,$$

which implies that $N_C(\bar{x}) \supseteq \bar{x}\mathbb{R}_+$.

To obtain the opposite inclusion, let us consider an arbitrary $d \in E = \text{span}(\bar{x}) \oplus [\text{span}(\bar{x})]^\perp$, that is, $d = \alpha\bar{x} + z$, $\alpha \in \mathbb{R}$, $\langle \bar{x}, z\rangle = 0$. Similarly, let us take $x = \beta\bar{x} + \gamma z$, where $\beta, \gamma \in \mathbb{R}$ are selected so that $1 \ge \|x\|^2 = \beta^2 + \gamma^2\|z\|^2$. Then we want the inequality:

$$\langle d, x - \bar{x}\rangle = \alpha(\beta - 1) + \gamma\|z\|^2 \le 0.$$

If $\gamma = 0$ then we are free to select any $\beta \in [-1, 1]$, from which it follows that $\alpha \geq 0$ is a must. If $z \neq 0$, then we can for example take $\beta = \cos(\epsilon)$, $\gamma = \sin(\epsilon)/\|z\|$, this way $\|x\| = 1$ and

$$\langle d, x - \bar{x} \rangle = \alpha(\cos(\epsilon) - 1) + \sin(\epsilon)\|z\| = -\alpha\epsilon^2/2 + \epsilon\|z\| + o(\epsilon^2),$$

which will be positive for small $\epsilon > 0$. Therefore, $N_C(\bar{x}) \subseteq \bar{x}\mathbb{R}_+$. □

(c) C-subspace of E.

Solution: Clearly $N_C(\bar{x}) \supseteq C^\perp$:

$$\forall x, \bar{x} \in C, d \in C^\perp : \langle d, \underbrace{x - \bar{x}}_{\in C} \rangle = 0.$$

Consider now an arbitrary $d \in E = C \oplus C^\perp$, $d = d_1 + d_2$, $d_1 \in C$, $d_2 \in C^\perp$. Take $x = \bar{x} + d_1 \in C$. Then

$$\langle d, x - \bar{x} \rangle = \underbrace{\langle d_1, d_1 \rangle}_{\geq 0} + \underbrace{\langle d_2, d_1 \rangle}_{=0}.$$

Therefore, in order for $d \in N_C(\bar{x})$it is necessary that $d_1 = 0$, that is, $d \in C^\perp$, implying that $N_C(\bar{x}) \subseteq C^\perp$. □

(d) Closed half-space: $C = \{\, x \in E \mid \langle a, x \rangle \leq b \,\}$, $a \in E \setminus \{0\}$, $b \in \mathbb{R}$.

Solution: As in (b), if $\bar{x} \in \text{int}C$ then $N_C(\bar{x}) = \{0\}$. Let us now consider the case $\langle a, \bar{x} \rangle = b$.

Indeed, let us take any $d \in E = \text{span}(a) \oplus [\text{span}(a)]^\perp$, that is, $d = \alpha a + z$, $\alpha \in \mathbb{R}$, $\langle a, z \rangle = 0$. Similarly, let $x = \bar{x} - \beta a + \gamma z$, $\beta, \gamma \in \mathbb{R}_+$. Clearly $x \in C$ as $\langle a, x \rangle = b - \beta\|a\|^2 \leq b$. Then:

$$\langle d, x - \bar{x} \rangle = -\alpha\beta\|a\|^2 + \gamma\|z\|^2$$

Clearly the necessary condition for $d \in N_C(\bar{x})$ is that $\alpha \geq 0$ and $\|z\| = 0$, that is, $N_C(\bar{x}) \subseteq a\mathbb{R}_+$.

A direct computation shows that these two conditions are also sufficient for $d \in N_C(\bar{x})$, and therefore we conclude that $N_C(\bar{x}) = a\mathbb{R}_+$. □

(e) $C = \{\, x \in \mathbb{R}^n \mid x_j \geq 0, \forall j \in J \,\}$, $J \subseteq \{1, \dots, N\}$.

Solution: $N_C(\bar{x}) = S(\bar{X})$, where

$$S(\bar{x}) = \{\, d \in \mathbb{R}^n \mid d_j \leq 0, \text{if } j \in J \text{ and } \bar{x}_j = 0, d_j = 0, \text{otherwise} \,\}.$$

Let us take any $x \in C$ and $d \in S(x)$. Then

$$\langle d, x - \bar{x} \rangle = \sum_{j \in J \wedge \bar{x}_j = 0} \underbrace{d_j}_{\leq 0} \underbrace{x_j}_{\geq 0} + \sum_{j \notin J \vee \bar{x}_j \neq 0} \underbrace{d_j}_{=0} (x_j - \bar{x}_j) \leq 0,$$

implying $S(\bar{x}) \subseteq N_C(\bar{x})$.

By taking $x \in \mathbb{R}^n$ such that $x_j = \max\{d_j, 0\}$ if $j \in J \wedge \bar{x}_j = 0$, $x_j = \bar{x}_j + \varepsilon d_j$, for $j \notin J \vee \bar{x}_j \neq 0$, where $\varepsilon > 0$ is selected so small that $x \in C$, the same inequality shows that $S(\bar{x}) \supseteq N_C(\bar{x})$. □

Exercise 3

For each of the following cones show that $N_K(0) = -K$.

(a) Non-negative cone: $K = \mathbb{R}^n_+$.

Solution: $\mathbb{R}^n_+$ is the cone from Exercise 2(e) with $J = \{1, \dots, n\}$. With this identification we simply re-use the result: $N_{\mathbb{R}^n_+}(0) = \{\, d \in \mathbb{R}^n \mid d_i \leq 0, i = 1, \dots, n \,\} = -\mathbb{R}^n_+$. $\square$

(b) Semi-definite cone: $K = \mathcal{S}^n_+$.

Solution: On the one hand, for every $X \in \mathcal{S}^n_+$ and every $Y \in -\mathcal{S}^n_+$ we have the inequality

$$\langle Y, X - 0 \rangle \leq \underbrace{\lambda(Y)^T}_{\leq 0} \underbrace{\lambda(X)}_{\geq 0} \leq 0,$$

where we used Fan's theorem. As a result, $-\mathcal{S}^n_+ \subseteq N_{\mathcal{S}^n_+}(0)$.

On the other hand, we can take an arbitrary $Y \in \mathcal{S}^n$ and consider its spectral decomposition, $Y = U \mathrm{Diag} y U^T$, $U \in \mathcal{O}^n$. We then construct $Y^+ = U \mathrm{Diag} y^+ U^T$, where $y^+_i = \max\{y_i, 0\}$ and $y^-_i = \min\{y_i, 0\}$. In particular, $y = y^+ + y^-$, $\langle y^+, y^- \rangle_{\mathbb{R}^n} = 0$, and $Y^+ \in \mathcal{S}^n_+$. Additionally, Y, Y^+ have a simultaneous ordered spectral decomposition (by construction). Therefore, Fan's theorem (with equality) is applicable, and

$$\langle Y, Y^+ \rangle = \lambda(Y)^T \lambda(Y^+) = \langle y, y^+ \rangle_{\mathbb{R}^n} = \langle y^+, y^+ \rangle_{\mathbb{R}^n} + \langle y^-, y^+ \rangle_{\mathbb{R}^n} = \|y^+\|^2_{\mathbb{R}^n}.$$

Therefore, for Y to be in $N_{\mathcal{S}^n_+}(0)$ it is necessary that $y^+ = 0$ - that is, that all eigenvalues of Y are non-positive. Therefore, $-\mathcal{S}^n_+ \supseteq N_{\mathcal{S}^n_+}(0)$.

$\square$

(c) Ice-cream cone: $K = \{\, x \in \mathbb{R}^n \mid x_1 \geq 0, x_1^2 \geq x_2^2 + \cdots + x_n^2 \,\}$.

Solution: We take $x \in K$, $-y \in -K$ and compute

$$\begin{aligned}\langle -y, x - 0 \rangle &= -x_1 y_1 - \sum_{i=2}^n x_i y_i \leq -x_1 y_1 + \left(\sum_{i=2}^n x_i^2 \right)^{1/2} \left(\sum_{i=2}^n y_i^2 \right)^{1/2} \\ &\leq -x_1 y_1 + x_1 y_1 \leq 0,\end{aligned}$$

where we have used Cauchy–Bunyakovsky–Schwarz inequality and the definition of the ice-cream cone. Thus $-K \subseteq N_K(0)$.

To show the opposite inclusion we consider an arbitrary $y \in \mathbb{R}^n$. If $y_1 > 0$ we can define $x_1 = y_1$, $x_2 = \cdots = x_n = 0$ so that $x \in K$ but

$$\langle y, x \rangle = y_1^2 > 0.$$

Therefore for $y \in N_K(0)$ it is necessary that $y_1 \leq 0$.

Further, if $y_1 \le 0$ we define $x_i = y_i$, $i = 2, \dots, n$, and $x_1 = (x_2^2 + \cdots + x_n^2)^{1/2}$. Then $x \in K$ and

$$\langle y, x \rangle = x_1(y_1 + x_1).$$

Thus for $y \in N_K(0)$ it is necessary that $y_1 \le -x_1 = -(y_2^2 + \cdots + y_n^2)^{1/2}$.

In other words, $y \in N_K(0)$ implies that $-y \in K$, or $-K \supseteq N_K(0)$. □

Exercise 4

Given: $A \in \mathcal{L}(E, Y)$, $b \in Y$, $C = \{ x \in E \mid Ax = b \}$. Find $N_C(\bar{x})$, where $\bar{x} \in C$.

Solution: Clearly $x \in C \iff x - \bar{x} \in \ker A$. In addition, for every $x \in C$ we have that $\tilde{x} = 2\bar{x} - x \in C$, and consequently

$$\langle d, x - \bar{x} \rangle \le 0, \forall x \in C \iff d \perp \ker A \iff d \in \mathrm{im} A^* = A^* Y,$$

where the last equivalence is Fredholm alternative.

Fredholm alternative can be shown as follows. $\forall d \in A^*Y$, $\exists y \in Y : d = A^* y$. $\forall z \in \ker A$ we have the equality

$$\langle d, z \rangle = \langle A^* y, z \rangle = \langle y, Az \rangle = \langle y, 0 \rangle = 0,$$

which implies that $A^*Y \subset (\ker A)^\perp$. To show the opposite inclusion we use separation (Theorem 1.1.1). Since we are considering a finite-dimensional situation, the linear subspace $A^*Y \subset E$ is closed. It is also a convex set, and therefore $\forall d \notin A^*Y$ there is $a \in E \setminus \{0\}$ and $b \in \mathbb{R}$, such that $\langle a, A^* y \rangle = \langle Aa, y \rangle \le b < \langle a, d \rangle$, $\forall y \in Y$. By selecting $y = 0 \in Y$ we conclude that $b \in \mathbb{R}_+$, and by chosing $y = \alpha Aa \in Y$, $\alpha \in \mathbb{R}$ we see that $a \in \ker A$. Consequently $d \notin (\ker A)^\perp$, and $A^*Y \supset (\ker A)^\perp$. □

Exercise 7

1. Given: $g : [0, 1] \to \mathbb{R}$ - convex function with $g(0) = 0$. Show: $g(t)/t : (0, 1] \to \mathbb{R}$ - nondecreasing.

Proof: Take $0 < t_1 \le t_2$. Then $t_1 = (t_1/t_2)t_2 + (1 - t_1/t_2)0$, and by convexity of g we get $g(t_1) \le t_1/t_2 g(t_2) + (1 - t_1/t_2)g(0) = t_1/t_2 g(t_2)$. Therefore, $g(t_1)/t_1 \le g(t_2)/t_2$. □

2. Show that for any convex function $f : C \to \mathbb{R}$ the quotient $t \mapsto [f(\bar{x} + t(x - \bar{x})) - f(x)]/t$ is a non-decreasing function on $(0, 1]$.

Proof: We define $g(t) = f(\bar{x} + t(x - \bar{x})) - f(x)$ and use the previous result. □

3. Complete the proof of Proposition 2.1.2 (first order sufficient conditions).

Proof: For all $x \in C$ we have the following string of inequalities:

$$\begin{aligned} 0 \leq f'(\bar{x}; x - \bar{x}) &= \lim_{t \downarrow 0} \frac{f(\bar{x} + t(x - \bar{x})) - f(x)}{t} \\ &\leq \left. \frac{f(\bar{x} + t(x - \bar{x})) - f(x)}{t} \right|_{t=1} = f(x) - f(\bar{x}), \end{aligned}$$

where the inequality on the second line is owing to the monotonicity of the quotient. Thus $\bar{x} \in C$ if the global minimum of f over C. □

Exercise 8

(a) Given: $f : C \to \mathbb{R}$ - strictly convex function; C-convex set. Show: at most one global minimum.

Proof: Assume that $f(x_1) = f(x_2) = \min_{x \in C} f(x)$. Then, utilizing strict convexity,

$$f((x_1 + x_2)/2) < [f(x_1) + f(x_2)]/2 = \min_{x \in C} f(x),$$

which is a contradiction since $(x_1 + x_2)/2 \in C$ owing to the convexity of C. □

(b) Show that the function $f_y(x) = \|x - y\|^2/2$ is strictly convex.

Solution: Let $x_1 \neq x_2 \in E$, $\lambda \in (0, 1)$ be arbitrary, and let $z = \lambda x_1 + (1 - \lambda)x_2$. We need to show that $f_y(z) < \lambda f_y(x_1) + (1 - \lambda) f_y(x_2)$.

We can expand our function into a Taylor series around z, and we only need to do that up to quadratic terms, since the original function is a quadratic form:

$$f_y(x - z) = \underbrace{f_y(z) + \langle z - y, x - z \rangle}_{=:g(x-z), \text{affine part}} + \frac{1}{2}\|x - z\|^2.$$

The affine part $g_y(x)$ satisfies the "convexity condition" with exact equality: $g_y(\lambda x_1 + (1 - \lambda)x_2) = \lambda g_y(x_1) + (1 - \lambda) g_y(x_2)$, $\forall x_1, x_2 \in E, \lambda \in [0, 1]$. Thus we can concentrate on the quadratic part $h_y(x - z) := f_y(x - z) - g_y(x - z)$. But now the inequality is trivial: $h_y(z - z) = \|z - z\|^2/2 = 0 < \lambda \|x_1 - z\|^2 + (1 - \lambda)\|x_2 - z\|^2 = \lambda h_y(x_1 - z) + (1 - \lambda) h_y(x_2 - z)$, since $z \neq x_1$, $z \neq x_2$, $0 < \lambda < 1$. □

(c) Given: $C \subset E$ - non-empty, closed, convex.

(i) Show that there is a unique closed point $P_C(y) \in C$ to any $y \in E$, which satisfies the conditions

$$\langle y - P_C(y), x - P_C(y)\rangle \leq 0, \qquad \forall x \in C.$$

Proof: Let $f_y(x) = \|x - y\|^2/2$, which is a strictly convex function (cf. (b)). Consider the problem $\min_{x\in C} f_y(x)$. Since C is non-empty and closed, and f_y clearly satisfies the growth condition (1.1.4) and is convex (cf. (b)), Propositions 1.1.5 and 1.1.3 imply that the global minimum is attained. Owing to (a) and (b), this global minimum is unique; we will call it $P_C(y)$.

The first order necessary (Proposition 2.1.1) and sufficient (Proposition 2.1.2) optimality conditions characterize $P_C(y)$ as

$$\begin{gathered} f'_y(P_C(y), x - P_C(y)) \geq 0, \qquad \forall x \in C \\ \Updownarrow \\ \langle P_C(y) - y, x - P_C(y)\rangle \geq 0, \qquad \forall x \in C. \qquad \square \end{gathered}$$

(ii) Show that $d \in N_C(\bar{x})$, $\bar{x} \in C$ if and only if $\bar{x} = P_C(\bar{x} + d)$.

Proof:

$$\begin{gathered} d \in N_C(\bar{x}) \\ \Updownarrow \\ \langle d, x - \bar{x}\rangle \leq 0, \qquad \forall x \in C \\ \Updownarrow \\ \langle \bar{x} + d - \bar{x}, x - \bar{x}\rangle \leq 0, \qquad \forall x \in C \\ \Updownarrow \\ \bar{x} = P_C(\bar{x} + d), \end{gathered}$$

where the last equivalence is owing to the characterization of $P_C(y)$ shown in (i). $\square$

(iii) Show the Lipschitz continuity of the projection: $\forall y, z \in C : \|P_C(y) - P_C(z)\| \leq \|y - z\|$.

Proof: Owing to (ii), we have the inequality

$$\langle y - P_C(y), x - P_C(y)\rangle \leq 0, \quad \forall x \in C \implies \langle y - P_C(y), P_C(z) - P_C(y)\rangle \leq 0.$$

Similarly

$$\langle z - P_C(z), P_C(y) - P_C(z)\rangle \leq 0.$$

Therefore we can write

$$\begin{aligned}\|P_C(y) - P_C(z)\|^2 &= \langle P_C(y) - P_C(z), P_C(y) - P_C(z)\rangle \\ &= \underbrace{\langle z - P_C(z), P_C(y) - P_C(z)\rangle}_{\leq 0} - \langle z, P_C(y) - P_C(z)\rangle \\ &+ \underbrace{\langle P_C(y) - y, P_C(y) - P_C(z)\rangle}_{\leq 0} + \langle y, P_C(y) - P_C(z)\rangle \\ &\leq \langle y - z, P_C(y) - P_C(z)\rangle \leq \|y - z\|\|P_C(y) - P_C(z)\|,\end{aligned}$$

where the last inequality is owing to Cauchy, Bunyakovsky, and Schwarz. □

(d) Given $a \in E \setminus \{0\}$, compute the nearest element $P_C(y)$ in $C = \{\, x \in E \mid \langle a, x\rangle = 0 \,\}$ to $y \in E$.

Proof: Owing to (ii), we need to find $x \in C$ such that $d = y - x \in N_C(x)$, where the latter cone equals to $a\mathbb{R}$ as computed in Exercise 4. Thus we need to find a scalar $\alpha \in \mathbb{R}$ such that $x = y - \alpha a \in C$.

$$0 = \langle a, x\rangle = \langle y, a\rangle - \alpha\langle a, a\rangle \implies \alpha = \frac{\langle y, a\rangle}{\|a\|^2} \implies x = P_C(y) = y - \frac{\langle y, a\rangle}{\|a\|^2}a.$$

□

(e) Projection on $\mathbb{R}^n_+$ and $\mathcal{S}^n_+$.

1. Prove that $y^+ = P_{\mathbb{R}^n_+}(y)$, where $y_i^+ = \max\{y_i, 0\}$, $y \in \mathbb{R}^n$.

Proof: Let us define $y^- = y - y^+$. Note that $-y^- \in \mathbb{R}^n_+$ and $\langle y^-, y^+\rangle = 0$. To prove the claim in view of (i), it is sufficient to notice that $\forall x \in \mathbb{R}^n_+$ we have the inequality

$$\langle y - y^+, x - y^+\rangle = \underbrace{\langle y^-, x\rangle}_{\leq 0} - \underbrace{\langle y^-, y^+\rangle}_{=0} \leq 0.$$

□

2. Prove that $Y^+ = P_{\mathcal{S}^n_+}(Y)$, where $Y = U^T \mathrm{Diag} y U$, and $Y^+ = U^T \mathrm{Diag} y^+ U$, $y \in \mathbb{R}^n$, $U \in \mathcal{O}^n$.

Proof: We utilize Fan's theorem (Theorem 1.2.1) twice:

$$\langle Y - Y^+, X\rangle \leq \lambda(Y - Y^+)^T \lambda(X) = \underbrace{\lambda(U^T \mathrm{Diag} y^- U)^T}_{\leq 0} \underbrace{\lambda(X)}_{\geq 0} \leq 0,$$

and (since $Y - Y^+$ and Y^+ admit a simultaneous ordered spectral decomposition)

$$\langle Y - Y^+, Y^+\rangle = \lambda(Y - Y^+)^T \lambda(Y^+) = \langle y^-, y^+\rangle_{\mathbb{R}^n} = 0.$$

As a consequence, $\forall X \in \mathcal{S}^n_+$

$$\langle Y - Y^+, X - Y^+\rangle \leq 0,$$

and therefore $Y^+ = P_{\mathcal{S}^n_+}(Y)$ owing to (i). □