Notes on the Elvis Problem

1 Preliminaries.

The philosophy of so-called *Convex Analysis* is that sets and functions are treated with equal attention. Let me explain.

Recall a set $F \subseteq \mathbb{R}^n$ is convex provided

$$x,\, y \in F,\, 0 \leq \lambda \leq 1 \quad \Longrightarrow \quad \lambda x + (1-\lambda) y \in F. \tag{1}$$

The set of all nonempty, closed and convex sets is denoted by $\mathcal{C}$.

The *epigraph* $\mathrm{epi}(f)$ of an extended-valued function $f : \mathbb{R}^n \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ is defined as

$$\mathrm{epi}(f) := \{(x, r) : r \geq f(x)\},$$

which is a subset in $\mathbb{R}^{n+1}$. Then $f(\cdot)$ is (1) *lower semicontinuous* (lsc) provided $\mathrm{epi}(f)$ is closed, (2) *convex* provided $\mathrm{epi}(f)$ is a convex set (in $\mathbb{R}^{n+1}$), and (3) *proper* if $\mathrm{epi}(f) \neq \emptyset$ and contains no vertical lines (a vertical line at $x \in \mathbb{R}^n$ occurs when $f(x) = -\infty$). The set of all lsc, convex and proper functions is denoted by $\mathcal{F}$.

Exercise 1.1. *Suppose $f : \mathbb{R}^n \to \overline{\mathbb{R}}$.*

(a) The effective domain *of $f(\cdot)$ is defined by $\mathrm{dom}(f) := \{x : f(x) < \infty\}$. Show if $f(\cdot) \in \mathcal{F}$, then $\mathrm{dom}(f)$ is a convex set. Give an example for which $f(\cdot) \in \mathcal{F}$ but $\mathrm{dom}(f)$ is not closed.*

(b) If $\mathrm{dom}(f) = \mathbb{R}^n$, show that $f \in \mathcal{F}$ if and only if

$$f(x_\lambda) \leq (1-\lambda) f(x_0) + \lambda f(x_1) \tag{2}$$

for all $x_0, x_1 \in \mathbb{R}^n$ and $0 \leq \lambda \leq 1$, and where $x_\lambda := (1-\lambda)\, x_0 + \lambda\, x_1$.

(c) Define an arithmetic and an order relation on $\overline{\mathbb{R}}$ so that the property that a lsc $f(\cdot)$ is convex is characterized by (2).

Associated to any set $S \subseteq \mathbb{R}^n$ is the indicator function $I_S : \mathbb{R}^n \to \overline{\mathbb{R}}$ defined by

$$I_S(x) = \begin{cases} 0 & \text{if } x \in S \\ +\infty & \text{if } x \notin S. \end{cases}$$

Exercise 1.2. *Suppose $S \subseteq \mathbb{R}^n$. Show $I_S(\cdot)$*

(a) is lsc if and only if S is closed;

(b) is a convex function if and only if S is a convex set; and

(c) belongs to $\mathcal{F}$ if and only if S belongs to $\mathcal{C}$.

One of the most important properties of convex sets is the fact that there are two ways to characterize them. The first is the direct definition given above in (1). A second, or *dual* way, is through so-called separation as explained in the following theorem.

Theorem 1.1 (Separation Theorem)**.** *Suppose F is closed and convex and $v \notin F$. Then there exists $\vec{n} \in \mathbb{R}^n$ so that*

$$\sup\big\{\langle v', \vec{n}\rangle : \ v' \in F\big\} < \langle v, \vec{n}\rangle \tag{3}$$

Proof. Please see the Primer on Convex Analysis file. □

The (closed) half space (with normal vector $\vec{n} \in \mathbb{R}^n$, $\vec{n} \neq \mathbf{0}$, and at level $r \in \mathbb{R}$) is defined by

$$\mathcal{H}_{\vec{n},r} := \big\{v : \ \langle \vec{n}, v\rangle \leq r\big\}$$

Exercise 1.3. *Show that F is closed and convex if and only if*

$$F = \bigcap\Big\{\mathcal{H}_{\vec{n},r} : \ \vec{n} \in \mathbb{R}^n, \ r \in \mathbb{R} \text{ are such that } F \subseteq \mathcal{H}_{\vec{n},r}\Big\}$$

2 The *polar* of a set

Definition 2.1. *The polar F° of a set $F \in \mathcal{C}$ is the set*

$$F^\circ := \{\zeta \in \mathbb{R}^n \ : \ \langle \zeta, v\rangle \leq 1 \ \forall v \in F\}.$$

A set F is *bounded* if there exists a constant $m \geq 0$ so that $v \in F \Rightarrow |v| \leq m$. Let

$$\mathcal{C}_0 := \big\{F \in \mathcal{C} : \ F \text{ is bounded and } \mathbf{0} \in \text{int}F\big\}.$$

Exercise 2.1. *Show the following.*

(a) For any nonempty set $F \subseteq \mathbb{R}^n$, one has F° belonging to $\mathcal{C}$.

(b) $F \in \mathcal{C}$ is bounded if and only if $\mathbf{0} \in \text{int}(F^\circ)$.

(c) $F \in \mathcal{C}_0$ if and only if $F^\circ \in \mathcal{C}_0$.

(d) If $F = r\overline{\mathbb{B}}$ for some $r > 0$, then $F^\circ = \frac{1}{r}\overline{\mathbb{B}}$.

(e) With $n = 2$ and positive constants a, b, if

$$F = \left\{\begin{pmatrix} x \\ y \end{pmatrix} : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\right\} \quad \text{then} \quad F^\circ = \left\{\begin{pmatrix} \zeta \\ \xi \end{pmatrix} : a^2\zeta^2 + b^2\xi^2 \leq 1\right\}.$$

(f) For $1 \le p < +\infty$, the p - norm is defined on $\mathbb{R}^n$ by

$$\|\mathbf{x}\|_p := \sqrt[p]{\sum_{i=1}^{n} |x_i|^p} \quad \textit{and} \quad \|\mathbf{x}\|_\infty := \max\big\{|x_i| : 1 \le i \le n\big\}, \quad \textit{where } \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

The p unit ball is the set $\overline{\mathbb{B}}_p := \big\{\mathbf{x} : \|\mathbf{x}\|_p \le 1\big\}$, which belongs to $\mathcal{C}_0$. If $F = \overline{\mathbb{B}}_p$, then $F^\circ = \overline{\mathbb{B}}_q$, where $\frac{1}{p} + \frac{1}{q} = 1$.

Actually one should think of F° belonging to the so called *dual* space of $\mathbb{R}^n$, but since $\mathbb{R}^n$ is its own dual, the dual space can be identified with $\mathbb{R}^n$ itself. Nonetheless, one should be cognizant of the three ways elements of $\mathbb{R}^n$ are being used: (1) as so-called state vectors x lying in the ambient space $\mathbb{R}^n$, (2) as "pointers" or velocity vectors $v \in F$ describing perhaps the direction and speed a state vector is moving, and (3) linear functionals that act on the state vectors through an inner product. Let's elaborate on (3): Suppose $\zeta \in \mathbb{R}^n$, and define the map $\ell_\zeta(\cdot) : \mathbb{R}^n \to \mathbb{R}^n$ by $\ell_\zeta(x) = \langle \zeta, x \rangle$. Then $\ell(\cdot) := \ell_\zeta(\cdot)$ satisfies the linearity property

$$\ell(x_1 + r x_2) = \ell(x_1) + r\ell(x_2). \tag{4}$$

Any function $\ell(\cdot)$ satisfying (4) is called a linear functional, and the dual space of $\mathbb{R}^n$ is the set of all linear functionals. The following exercise is the manner in which the dual space can be identified with $\mathbb{R}^n$

Exercise 2.2. *Suppose $\ell(\cdot)$ is a linear functional. Show $\exists$ a unique $\zeta \in \mathbb{R}^n$ with $\ell(\cdot) = \ell_\zeta(\cdot)$.*

In this way, we usually abuse notation by just writing ζ for the linear map $\ell_\zeta(\cdot)$. The point is that elements of $\mathbb{R}^n$ can be viewed in two distinct ways: (1) as the usual coordinate vector with n real components, and (2) in a way that says how it "operates" on the space $\mathbb{R}^n$ by taking an inner product. The definition of F° consists of describing how its elements are acting on elements in $\mathbb{R}^n$, which is why we said it should be thought of as a subset of the dual space. The same applies to the vector ζ in Theorem 1.1.

3 Gauge functions

Suppose $F \in \mathcal{C}$ is given. The gauge function $\gamma_F : X \to [0, \infty]$ is defined by

$$\gamma_F(x) = \inf\left\{ t \ge 0 : \frac{1}{t}\, x \in F \right\}.$$

By convention, if $rx \notin F$ for all $r > 0$, then $\gamma_F(x) = +\infty$. We mainly will be interested in only the case where $F \in \mathcal{C}_0$.

Exercise 3.1. *Let $F \in \mathcal{C}$ with $\mathbf{0} \in F$. Show the following:*

(a) $v \in F$ if and only if $\gamma_F(x) \le 1$.

(b) $\gamma_F(\cdot)$ *is positively homogeneous: that is,* $\gamma_F(rv) = r\gamma_F(v) \quad \forall v \in \mathbb{R}^n,\ r \geq 0.$

(c) $\gamma_F(\cdot) \in \mathcal{F}$*, and is finite-valued if and only if* $\mathbf{0} \in \text{int}(F)$.

A converse of some of the statements above is given next.

Exercise 3.2. *Suppose* $\gamma(\cdot) : \mathbb{R}^n \to [0, +\infty]$ *belongs to* $\mathcal{C}$ *and is also positively homogeneous. Define* $F := \{x \in \mathbb{R}^n : \gamma(x) \leq 1\}$*. Show* $F \in \mathcal{C}$ *and* $\gamma(\cdot) = \gamma_F(\cdot)$*. Furthermore, show* $\gamma(\cdot)$ *is finite-valued if and only if* $\mathbf{0} \in \text{int}(F)$.

4 Differentiablility concepts of convex objects

For a proper convex function $f : X \to (-\infty, +\infty]$, the subgradient $\partial f(x)$ at a point $x \in \text{dom} f$ is given by

$$\partial f(x) := \{\xi \in X : f(y) \geq f(x) + \langle \xi, y - x \rangle \ \forall y \in X\}. \tag{5}$$

The analogous concept for a set $F \in \mathcal{C}$ is the normal cone: For $v \in F$, the normal cone $N_F(v)$ is given by

$$N_F(v) := \{\zeta : \langle \zeta, v' - v \rangle \leq 0 \quad \forall v' \in F\}.$$

Exercise 4.1.

(a) For $f(\cdot) \in \mathcal{F}$*, show that*

$$\zeta \in \partial f(x) \quad \Longleftrightarrow \quad (\zeta, -1) \in N_{\text{epi}(f)}\big(x, f(x)\big). \tag{6}$$

(b) For $F \in \mathcal{C}$*, show that*

$$\zeta \in \partial I_F(x) \quad \Longleftrightarrow \quad \zeta \in N_F\big(v\big). \tag{7}$$

The general Elvis problem relies on the following relationships of the gauge functions for both the velocity set $F \in \mathcal{C}_0$ and its polar F°.

Exercise 4.2. *Suppose* $F \in \mathcal{C}_0$*. Show that the following statements are equivalent for any* $x, \zeta \in \mathbb{R}^n$.

(a) $\langle \zeta, x \rangle = \gamma_F(x)\gamma_{F^\circ}(\zeta)$.

(b) $\frac{x}{\gamma_F(x)}$ *attains the max over* $v \in F$ *of the map* $v \to \langle \zeta, v \rangle$.

(c) $\zeta \in N_F\left(\frac{x}{\gamma_F(x)}\right)$.

(d) $\frac{\zeta}{\gamma_{F^\circ}(\zeta)}$ *attains the max over* $\xi \in F^\circ$ *of the map* $\xi \to \langle \xi, x \rangle$.

(e) $x \in N_{F^\circ}\left(\frac{\zeta}{\gamma_{F^\circ}(\zeta)}\right)$.

(f) $\frac{\zeta}{\gamma_{F^\circ}(\zeta)} \in \partial\gamma_F(x)$.

(g) $\frac{x}{\gamma_F(x)} \in \partial\gamma_{F^\circ}(\zeta)$.

A convex set $F \subseteq \mathbb{R}^n$ is *strictly* convex if whenever x and y belong to F $(x \neq y)$ and $0 < \lambda < 1$, then $\lambda x + (1-\lambda)y \in \mathrm{int}F$. The so-called dual concept is having a "smooth" boundary. A closed convex set has a smooth boundary if and only if $N_F(x)$ is of the form $\{r\zeta_x : r \geq 0\}$ for some vector ζ_x with $|\zeta_x| = 1$, and the map $x \mapsto \zeta_x$ is continuous on $\mathrm{bdry}F$.

Exercise 4.3. *Show that F is strictly convex if and only if F° has a smooth boundary.*

Given $F \in \mathcal{C}$, the *support* function $\sigma_F(\cdot) : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$\sigma_F(\zeta) = \sup\{\langle \zeta, v\rangle : v \in F\}$$

Exercise 4.4. *Suppose $F \in \mathcal{C}$.*

(a) *Show $\sigma_F(\cdot) \in \mathcal{F}$.*

(b) *If $\sigma_F(\zeta) < +\infty$, show there exists at least one $v \in F$ so that*

$$\sigma_F(\zeta) = \langle \zeta, v\rangle,$$

and is unique if F is strictly convex. Give an example where there may be more than one such v.

(c) *Show that $\sigma_F(\zeta) < \infty$ for all $\zeta \in \mathbb{R}^n$ if and only if $\mathbf{0} \in \mathrm{int}(F)$.*

(d) *Suppose $x \notin F$. Show there exists a unique $v \in F$ so that*

$$\|x - v\| = \min\Big\{\|x - v'\| : \ v' \in F\Big\}.$$

Such a v is called the projection of x into F, and is denoted by $\mathrm{proj}_F(x)$.

(e) *Suppose $x \notin F$, and let $v = \mathrm{proj}_F(x)$ and $\zeta := x - v$. Show $\zeta \in N_F(v)$ and $\sigma_F(\zeta) = \langle \zeta, v\rangle$.*

Suppose $f(\cdot) \in \mathcal{F}$. It is obvious from (5) that a point $\bar{x} \in \mathbb{R}^n$ solves the problem

$$\min f(x) \quad \text{over } x \in \mathbb{R}^n \tag{$\mathcal{P}$}$$

if and only if $\mathbf{0} \in \partial f(\bar{x})$. Our Elvis problem will entail a problem like $(\mathcal{P})$ where $f(\cdot)$ is of the form $f(x) = g(x) + \mathcal{I}_\Sigma(x)$ where $g(\cdot) \in \mathcal{F}$ is finite-valued and $\Sigma \in \mathcal{C}$. In this case, problem $(\mathcal{P})$ is equivalent to both of

$$\min g(x) \quad \text{over } x \in \Sigma \tag{$\mathcal{P}'$}$$

$$\min\Big\{g(x) + \mathcal{I}_\Sigma(x)\Big\} \quad \text{over } x \in \mathbb{R}^n. \tag{$\mathcal{P}''$}$$

Considering the form of $(\mathcal{P}'')$, the necessary and sufficient optimality condition that $\bar{x}$ solves $(\mathcal{P}'')$ is that

$$\mathbf{0} \in \partial\big\{g(\cdot) + \mathcal{I}_\Sigma(\cdot)\big\}(\bar{x}). \tag{8}$$

Now it would be nice if we knew that the subgradient of a sum of two elements in $\mathcal{F}$ was the sum of its subgrdients (in some sense). This is akin to the well-known linearity property of ordinary differentiation. It is not true in general:

Exercise 4.5. *With $n = 1$, let*

$$f_1(x) = \begin{cases} -\sqrt{x} & \text{if } x \geq 0 \\ +\infty & \text{if } x < 0 \end{cases} \qquad \text{and} \qquad f_2(x) = \begin{cases} -\sqrt{-x} & \text{if } x \leq 0 \\ +\infty & \text{if } x > 0 \end{cases}$$

Show that $f_1(\cdot)$, $f_2(\cdot) \in \mathcal{F}$, $\partial f_1(\mathbf{0}) = \partial f_2(\mathbf{0}) = \emptyset$ and $\partial\big(f_1(\cdot) + f_2(\cdot)\big)(\mathbf{0}) = \mathbb{R}$,

So a "sum rule" cannot be expected to hold in all cases, but in the case that will be of most interest to us, it does hold. The following is a special case that we will use below.

Theorem 4.1 (Rockafellar, Convex Analysis, Theorem 23.8)**.** *Suppose $f(\cdot)$, $g(\cdot) \in \mathcal{F}$ and $\operatorname{dom}(g) = \mathbb{R}^n$ (that is, $g(\cdot)$ is finite-valued). Then for all $x \in \operatorname{dom}(f)$, we have*

$$\partial\Big(f(\cdot) + g(\cdot)\Big)(x) = \partial f(x) + \partial g(x) := \big\{\zeta + \xi : \zeta \in \partial f(x),\ \xi \in \partial g(x)\big\}$$

Exercise 4.6. *Consider $(\mathcal{P}'')$ with $g(\cdot) \in \mathcal{F}$ and $\Sigma \in \mathcal{C}$. Suppose further that $\operatorname{dom}(g) = \mathbb{R}^n$. Show $\bar{x}$ solves $(\mathcal{P}'')$ if and only if there exists $\zeta \in \partial g(\bar{x})$ satisfying $-\zeta \in N_\Sigma(\bar{x})$.*

Exercise 4.7. *Suppose $f(\cdot) \in \mathcal{F}$ and $g(\cdot)$ is defined by $g(x) = f(-x)$. Show that $\zeta \in \partial g(x)$ if and only if $-\zeta \in \partial f(-x)$. Explain why this is a special case of the Chain Rule for convex functions.*