[M2, MVA]

Convex Optimization, Algorithms and Applications

Maha ELBAYAD
maha.elbayad@student.ecp.fr

Homework 1

October 25, 2015

Exercise 2.12 - Convex sets

(a) A slab $S = \{x \in \mathbb{R}^n |\ \alpha \leq a^T x \leq \beta\}$

Convex, since $S = \{x \in \mathbb{R}^n |\ \alpha \leq a^T x\} \cap \{x \in \mathbb{R}^n |\ a^T x \leq \beta\}$ is an intersection of two halfspaces.

(b) A rectangle $R = \{x \in \mathbb{R}^n |\ \alpha_i \leq x_i \leq \beta_i,\ i = 1...n\}$

Convex, since $R = \bigcap\limits_{i=1}^{n} [\{x \in \mathbb{R}^n |\ \alpha_i \leq x_i\} \cap \{x \in \mathbb{R}^n |\ x_i \leq \beta_i\}]$ is a finite intersection of halfspaces.

(c) A wedge $W = \{x \in \mathbb{R}^n |\ a_1^T x \leq b_1,\ a_2^T x \leq b_2\}$

Convex, since it's an intersection of two halfspaces.

(d) $C = \{x |\ \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in S\}$ where $S \subset \mathbb{R}^n$

Convex, since:

$$\begin{aligned} C &= \bigcap_{y \in S} \{x |\ \|x - x_0\|_2 \leq \|x - y\|_2\} \\ &= \bigcap_{y \in S} \{x |\ (x - x_0)^T (x - x_0) \leq (x - y)^T (x - y)\} \\ &= \bigcap_{y \in S} \{x |\ (y - x_0)^T x \leq \frac{1}{2}(\|y\|_2^2 - \|x_0\|_2^2)\} \end{aligned}$$

is an intersection of halfspaces.

(e) $C = \{x|\, dist(x,S) \leq dist(x,T)\}$ where $S,T \subset \mathbb{R}^n$

Not convex, in fact:

$$C = \bigcup_{y\in S} \{x|\, \|x-y\|_2 \leq \|x-t\|_2 \,\forall t \in T\}$$

As a reunion of convex sets (d), this set isn't generally convex.

Let us consider in $\mathbb{R}$ the counterexample: $S = \{-a, a\}$ $(a > 0)$ and $T = \{0\}$

$C = \{x|\, x \geq \frac{a}{2}\} \cup \{x|\, x \leq -\frac{a}{2}\}$ is obviously not convex.

(f) The set $S = \{x|\, x + S_2 \subset S_1\}$ where $S_1, S_2 \subset \mathbb{R}^n$ with S_1 convex.

Convex as we can write:

$$S = \bigcap_{y\in S_2} \{x|x+y \in S_1\} = \bigcap_{y\in S_2} [S_1 - y]$$

which is an intersection of the convex sets $[S_1 - y]$, translations of the original convex S_1.

(g) The set $S = \{x|\, \|x-a\|_2 \leq \theta \|x-b\|_2\}$ where $a \neq b$ and $\theta \in [0,1]$

$$S = \{x|\, (1-\theta^2)x^Tx - 2(a-\theta^2 b)^T x + (a^Ta - \theta^2 b^T b) \leq 0\}$$

If $\theta = 1$, S is a **convex** halfspace.

If $\theta \neq 1$, $S = \{x^Tx - 2u^Tx + \lambda \leq 0\}$ where $u = \frac{1}{1-\theta^2}(a - \theta^2 b)$ and $\lambda = \frac{a^Ta - \theta^2 b^T b}{1-\theta^2}$

which is the **convex** ball of center u and radius $R = (u^Tu - \lambda)^{1/2}$

Exercise 2.36 - Euclidean distance matrices

Let $\mathbb{EDM}$ denote the set of Euclidean distance matrices.

$D \in \mathbb{EDM} \iff [D \in \mathbf{S}^n] \wedge [D_{ii} = 0\, (\forall i)] \wedge [x^T D x \leq 0 \,\forall x$ with $\mathbf{1}^T x = 0]$

Let $(e_1, ...e_n)$ be the canonical basis of $\mathbb{R}^n$:

$$\forall i = 1...n,\ D_{ii} = 0 \iff e_i^T D e_i = 0 \iff [e_i^T D e_i \geq 0 \wedge e_i^T D e_i \leq 0]$$

$$\forall x \in (\mathbf{1})^\perp\ x^T D x = \sum_{i,j} x_i x_j D_{ij} \leq 0$$

Solving for D is solving a system of homogenous linear inequalities, which has a convex cone S as a solution space. Hence, $\mathbb{EDM} = S \cap \mathbf{S}^n$ is a convex cone.

The dual cone:

Given than[1]:

$$\begin{aligned}\mathbb{EDM} &= \bigcap_{\substack{i=1..n\\ x\in\mathbf{1}^\perp}} \{D \in \mathbf{S}^n|\, e_i^T D e_i \geq 0,\ e_i^T D e_i \leq 0,\ x^T D x \leq 0\}\\ &= \bigcap_{\substack{i=1..n\\ x\in\mathbf{1}^\perp}} \{D \in \mathbf{S}^n|\, \langle e_i e_i^T, D\rangle \geq 0,\ \langle e_i e_i^T, D\rangle \leq 0,\ \langle xx^T, D\rangle \leq 0\}\end{aligned}$$

[1]Using the Frobenius inner product on $\mathbb{R}^{n\times n}$: $\langle A, B\rangle = tr(A^T B)$

And that

$$\mathbb{EDM}* = \{M |\, \langle M, D\rangle \geq 0\, \forall D \in \mathbb{EDM}\}$$

which is the set of every vector inward-normal to a hyperplane supporting the convex cone $\mathbb{EDM}$.

$$\mathbb{EDM}^* = Co\left\{ \bigcup_{\substack{i=1..n \\ x\in\mathbf{1}^\perp}} \{e_i e_i^T, -e_i e_i^T, -xx^T\} \right\}$$

Exercise 3.21 - Pointwise maximum and supremum
(a) $f(x) = max_{i=1...k} \|A^{(i)}x - b^{(i)}\|$ where $A^{(i)} \in \mathbb{R}^{m\times n}$, $b^{(i)} \in \mathbb{R}^m$ and $\|.\|$ is a norm on $\mathbb{R}^m$
f is the pointwise maximum of the fonctions $f_i(x) = \|A^{(i)}x - b^{(i)}\|$ for $i = 1..k$.
f_i is a convex function since it's the composition with an affine function of the norm (always convex).
(b) $f(x) = \sum_{i=1}^{r} |x|_{[i]}$ where $|x|_{[1]}, |x|_{[2]}, .., |x|_{[n]}$ are the absolute values of the components of x, sorted in nonincreasing order.
We can write f as:

$$f(x) = max_{1\leq i_1 <...< i_r \leq n} |x_{i_1}| + |x_{i_2}| + ... + |x_{i_r}|$$

which is a pointwise maximum of $\binom{n}{r}$ convex functions.

Exercise 3.32 - Products and ratios of convex functions
(a) *If f and g are convex, both nondecreasing (or nonincreasing), and positive functions on an interval, then fg is convex.*
Checking Jensen's inequality for $x, y \in \mathbb{R}$ and $t \in [0, 1]$:

$$\begin{aligned}
f(tx + (1-t)y)g(tx + (1-t)y) &\leq (tf(x) + (1-t)f(y))(tg(x) + (1-t)g(y)) \ \textit{(f and g are convex)} \\
&= t^2 f.g(x) + t(1-t)f(x)g(y) + t(1-t)g(x)f(y) + (1-t)^2 f.g(y) \\
&= tf.g(x) + (1-t)f.g(y) \\
&+ t(1-t)[f(x)g(y) + g(x)f(y) - f(x)g(x) - f(y)g(y)] \\
&= tf.g(x) + (1-t)f.g(y) + t(1-t)(g(x) - g(y))(f(y) - f(x))
\end{aligned}$$

Given that f and g have the same monotony, $(g(x) - g(y))(f(y) - f(x))$ is negative.
Thus:

$$f.g(tx + (1-t)y) \leq tf.g(x) + (1-t)f.g(y)$$

(b) *If f, g are concave, positive, with one nondecreasing and the other nonincreasing, then fg is concave.*

$$f(tx+(1-t)y)g(tx+(1-t)y) \geq (tf(x)+(1-t)f(y))(tg(x)+(1-t)g(y)) \text{ (f and g are concave)}$$
$$= tf.g(x)+(1-t)f.g(y)+t(1-t)(g(x)-g(y))(f(y)-f(x))$$

Given that f and g have opposite monotony, $(g(x)-g(y))(f(y)-f(x))$ is positive.
Thus:

$$f.g(tx+(1-t)y) \geq tf.g(x)+(1-t)f.g(y)$$

(c) *If f is convex, nondecreasing, and positive, and g is concave, nonincreasing, and positive, then f/g is convex.*
$\frac{1}{g}$ is positive convex and nondecreasing. According to (a) $f.\frac{1}{g}$ is convex.

Exercise 3.36 - Conjugate functions

(a) Max function. $f(x) = \max_{i=1,\dots,n} x_i$ on $\mathbb{R}^n$

$$f^*(y) = \sup_{x=(x_1,\dots,x_n)^T \in \mathbb{R}^n} (y^T x - \max_{1\leq i\leq n} x_i)$$

First, let us establish the domain of f^*:
- If y has a strictly negative component $y_j < 0$ then for a vector $x = \alpha e_j$ with $\alpha < 0$, the inner product $y^T x = \alpha y_j$ tends towards $+\infty$ as α decreases.
Hence the supremum is undefined.
- If $\mathbf{1}^T y \neq 1$, for $x = \alpha\mathbf{1}$, $y^T x - \max_i x_i = \alpha(\mathbf{1}^T y - 1)$ which tends towards $+\infty$ as α goes to $+\infty$ (resp. $-\infty$) in case $\mathbf{1}^T y - 1 > 0$ (resp. $\mathbf{1}^T y - 1 < 0$).
- If $y \succeq \mathbf{0}$ and $\mathbf{1}^T y = 1$, for all $x \in \mathbb{R}^n$:

$$y^T x \leq y^T(\max_i x_i \mathbf{1}) = \max_i x_i$$

Hence $x^T y - \max_i x_i \leq 0$.
Furthermore, the upperbound is reached for $x = \mathbf{0}$, thus $f^*(y) = 0$.
The conjugate function is:

$$f^*(y) = \begin{cases} 0,\ y \succeq \mathbf{0} \ \wedge\ \mathbf{1}^T y = 1 \\ +\infty,\ Otherwise \end{cases}$$

(b) Sum of largest elements. $f(x) = \sum_{i=1}^{r} x_{[i]},\ x \in \mathbb{R}^n$

$$f^*(y) = \sup_{x\in\mathbb{R}^n} (y^T x - \sum_{i=1}^{r} |x|_{[i]})$$

- If y has a strictly negative component $y_j < 0$ then for a vector $x = \alpha e_j$ with $\alpha < 0$:

$$y^T x - f(x) = \begin{cases} \alpha y_j, \; if \; r < n \\ \alpha(y_j - 1) \; if \; r = n \end{cases}$$

In both cases $y^T x - f(x)$ tends toward $+\infty$ as $\alpha \to -\infty$.
- If y has a component larger than 1, namely $y_j > 1$ then for a vector $x = \alpha e_j$ with $\alpha > 0$:

$$y^T x - f(x) = \alpha(y_j - 1)$$

which again tends towards $+\infty$ as $\alpha \to +\infty$.
- If $\mathbf{1}^T y \neq r$ then for a vector $x = \alpha \mathbf{1}$

$$y^T x - f(x) = \alpha \mathbf{1}^T y - \alpha r = \alpha(\mathbf{1}^T y - r)$$

which tends toward $+\infty$ if $\alpha \to +\infty$ (in case $\mathbf{1}^T y > r$), or as $\alpha \to -\infty$ (in case $\mathbf{1}^T y < r$).
- If $\mathbf{0} \preceq y \preceq \mathbf{1}$ and $\mathbf{1}^T y = r$:

$$y^T x \leq \sum_{i=1}^{r} x_{[i]}$$

With equality at $x = \mathbf{0}$, hence the conjugate function is:

$$f^*(y) = \begin{cases} 0, \; \mathbf{0} \preceq y \preceq \mathbf{1} \; \wedge \; \mathbf{1}^T y = r \\ +\infty, \; Otherwise \end{cases}$$

(c) Piecewise-linear function on $\mathbb{R}$. $f(x) = \max_{i=1,\ldots,m}(a_i x + b_i)$.
We assume that the scalars a_i are sorted in increasing order ($a_1 \leq ... \leq a_m$) and that:

$$\forall k = 1...m, \; \exists x \in \mathbb{R} : \; f(x) = a_k x + b_k$$

f is piecwise linear, changing its slope from a_k to a_{k+1} at c_k where $a_k c_k + b_k = a_{k+1} c_k + b_{k+1}$ i.e $c_k = -\frac{b_{k+1} - b_k}{a_{k+1} - a_k}$
We have:

$$f^*(y) = \sup_x (xy - \max_{i=1...m}(a_i x + b_i))$$

- If $y \notin [a_1, a_m]$, $xy - \max_{i=1...m}(a_i x + b_i)$ goes to $+\infty$ as $x \to +\infty$ if $y > a_m$ or as $x \to -\infty$ if $y < a_1$.
- If $y \in [a_i, a_{i+1}]$, $i = 1, ...(m-1)$:
For $x \in [c_k, c_{k+1}] : \; f(x) = a_{k+1} x + b_{k+1}$ thus:

$$f^*(y) = \max_k \sup_{x \in [c_k, c_{k+1}]} x(y - a_{k+1}) - b_{k+1}$$

Depending on the sign of the factor $(y - a_{k+1})$ we will have:

$$f^*(y) = \max\big(\max_{k \leq i-1} c_{k+1}(y - a_{k+1}) - b_{k+1}, \max_{k \geq i} c_k(y - a_{k+1}) - b_{k+1}\big) = c_i(y - a_i) - b_i$$

Hence:

$$f^*(y) = \begin{cases} -\frac{b_{i+1}-b_i}{a_{i+1}-a_i}(y - a_i) - b_i, \; if \; y \in [a_1, a_m] \; with \; i \text{ such that } y \in [a_i, a_{i+1}] \\ +\infty, \; if \; y \notin [a_1, a_m] \end{cases}$$

(d) Power function $f(x) = x^p$ on $\mathbb{R}_{++}$ where $p > 1$

$$f^*(y) = sup_{x \in \mathbb{R}_{++}}(xy - x^p)$$

- If $(y \leq 0)$ $f^*(y) = 0$ when $x \to 0^+$.
- If $(y > 0)$ the supremum is reached at $x^* = \left(\frac{y}{p}\right)^{\frac{1}{p-1}}$ with:

$$f^*(y) = (p-1)\left(\frac{y}{p}\right)^{\frac{p}{p-1}}$$

Therefore:

$$f^*(y) = \begin{cases} 0, \; y \leq 0. \\ (p-1)\left(\frac{y}{p}\right)^{\frac{p}{p-1}}, \; y > 0. \end{cases}$$

(d') Power function $f(x) = x^p$ on $\mathbb{R}_{++}$ where $p < 0$
- If $(y > 0)$ $f^*(y) = +\infty$ as $x \to +\infty$
- If $(y \leq 0)$ the supremum is reached at $x^* = -\left(-\frac{y}{p}\right)^{\frac{1}{p-1}}$ with:

$$f^*(y) = (1-p)\left(-\frac{y}{p}\right)^{\frac{p}{p-1}}$$

Therefore:

$$f^*(y) = \begin{cases} (1-p)\left(-\frac{y}{p}\right)^{\frac{p}{p-1}}, \; y \leq 0. \\ +\infty, \; y > 0. \end{cases}$$

(e) Negative geometric mean $f(x) = -(\prod_{i=1..n} x_i)^{1/n}$ on $\mathbb{R}^n_{++}$ (we will assume $n \geq 2$)
- If y has a strictly positive component $y_j > 0$ then for $x = \alpha e_j + \mathbf{1}$ we will have:

$$x^T y + (\prod_{i=1..n} x_i)^{1/n} = \sum_{i \neq j} y_i + (\alpha + 1) y_k + (1 + \alpha)^{1/n}$$

which tends towards $+\infty$ as $\alpha \to +\infty$.
- If $(\prod_i(-y_i))^{1/n} < \frac{1}{n}$ we will choose the vector x such that $x_i = -\frac{\alpha}{y_i}$ for which:

$$\begin{aligned} x^T y + (\prod_{i=1..n} x_i)^{1/n} &= -n\alpha + \frac{\alpha^n}{(\prod_{i=1..n}(-y_i))^{1/n}} \\ &\geq n(\alpha^n - \alpha) \\ &\xrightarrow[\alpha \to +\infty]{} +\infty. \end{aligned}$$

- If $(\prod_i(-y_i))^{1/n} > \frac{1}{n}$ for the same vector x, $x_i = -\frac{\alpha}{y_i}$:

$$\begin{aligned} x^T y + (\prod_{i=1..n} x_i)^{1/n} &= -n\alpha + \frac{\alpha^n}{(\prod_{i=1..n}(-y_i))^{1/n}} \\ &\xrightarrow[\alpha \to +\infty]{} +\infty. \end{aligned}$$

- If $y \preceq \mathbf{0}$ and $(\prod_i(-y_i))^{1/n} = \frac{1}{n}$:
For $x \in \mathbb{R}^n_{++}$, we use the AM-GM inequality:

$$\begin{aligned} \frac{-x^T y}{n} &\geq (\prod_{i=1..n}(-x_i y_i))^{1/n} \\ &= (\prod_{i=1..n} x_i)^{1/n} (\prod_{i=1..n}(-y_i))^{1/n} \\ &= \frac{1}{n}(\prod_{i=1..n} x_i)^{1/n} \\ x^T y &\leq f(x). \end{aligned}$$

The AM-GM equality is satisfied for the vector x such that $x_i \propto -\frac{1}{y_i}$
Hence $f^*(y) = 0$
Therefore:

$$f^*(y) = \begin{cases} 0, \; if \; y \preceq \mathbf{0} \; \wedge \; (\prod_i(-y_i))^{1/n} = \frac{1}{n} \\ +\infty \; otherwise \end{cases}$$

For the case $n = 1$:

$$f^*(y) = \sup_{x \in \mathbb{R}_{++}} x(y+1) = \begin{cases} +\infty, \; if \; y + 1 > 0 \\ 0, \; otherwise \end{cases}$$

(f) Negative generalized logarithm for second-order cone:

$$f(x,t) = -\log(t^2 - x^T x) \;\; on \; E = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} \mid \|x\|_2 < t\}$$

$$f^*(y,s) = sup_{(x,t)\in E}((x,t)^T(y,s) - f(x,t))$$

We derive the expression above ($A = (x,t)^T(y,s) - f(x,t)$) with respect to x and t:

$$\begin{aligned}\frac{\partial A}{\partial x_i} = 0 \iff x_i &= \frac{1}{2}y_i(t^2 - x^T x)\\ \Rightarrow x^T x &= \frac{1}{4}(t^2 - x^T x)^2 y^T y\\ \Rightarrow t^2 - x^T x &= 2\frac{\|x\|}{\|y\|}\end{aligned}$$

And for t:

$$\frac{\partial A}{\partial t} = 0 \iff \frac{2t}{t^2 - x^T x} = -s$$

The two conditions combined give us:

$$-\frac{t}{s} = \frac{\|x\|}{\|y\|} \tag{1}$$

Which we rewrite as follows:

$$t = \frac{2s}{s^2 - y^T y},\ x = \frac{2}{s^2 - y^T y}y$$

The domain of f^*:

(1) implies that $\|y\| < -s$, $s < 0$ for $(x,t) \in E$.

- If $s > 0$: we consider $(x,t) = (\mathbf{0}, \alpha)$, $\alpha > 0$ then:

$$x^T y + ts - log(t^2 - x^T x) = \alpha s - 2\log(\alpha) \xrightarrow[\alpha\to+\infty]{} +\infty$$

- If $\|y\| \geq -s,\ s < 0$ we choose the vector $x = \alpha y$, $t = \alpha(\|x\| + 1)$ for a scalar $\alpha > 1$:

$$y^T x + ts = \alpha y^T y + \alpha s(\alpha\|y\| + 1) \geq \alpha(y^T y - s^2) \xrightarrow[\alpha\to+\infty]{} +\infty$$

- If $\|y\| < -s$ and $s < 0$ then we have shown that the supremum is reached at:

$$t = \frac{2s}{s^2 - y^T y},\ x = \frac{2}{s^2 - y^T y}y$$

Hence:

$$f^*(y,s) = \begin{cases} -2 + \log(4) - \log(s^2 - y^T y),\ \|y\| < -s \ \wedge\ s < 0 \\ +\infty,\ otherwise \end{cases}$$