MVA - Convex Optimization
DM1

Ariane ALIX

October 14th, 2019

Exercise 2.12

(a) Let $S_{\alpha,\beta,a}$ be a slab defined by $S_{\alpha,\beta} = \{x \in \mathbb{R}^n \mid \alpha \leq a^T x \leq \beta\}$.

We notice that $S_{\alpha,\beta,a} = \{x \in \mathbb{R}^n \mid \alpha \leq a^T x\} \cap \{x \in \mathbb{R}^n \mid a^T x \leq \beta\}$, and that these two sets are convex because they are halfspaces.

$S_{\alpha,\beta,a}$ being an intersection of two convex sets, it is convex as well.

(b) Let $R_{\alpha,\beta}$ be a rectangle defined by $R_{\alpha,\beta} = \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i,\ i = 1, ..., n\}$.

$$R_{\alpha,\beta} = \bigcap_{i \in [1,n]} \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i\}$$
$$= \bigcap_{i \in [1,n]} (\{x \in \mathbb{R}^n \mid \alpha_i \leq x_i\}) \cap (\{x \in \mathbb{R}^n \mid x_i \leq \beta_i\})$$

Since $R_{\alpha,\beta}$ is an intersection of $2n$ convex sets, it is convex.

(c) Like the two first types of sets, a wedge is an intersection of halfspaces, hence is convex.

(d) Let S_{x_0} be defined as the set of points closer to a given point x_0 than a given set S.

$$S_{x_0} = \{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in S\} \text{ where } S \subseteq \mathbb{R}^n$$
$$= \bigcap_{y \in S} \{x \mid \|x - x_0\|_2 \leq \|x - y\|_2\}$$
$$= \bigcap_{y \in S} \{x \mid \sqrt{(x - x_0)^T (x - x_0)} \leq \sqrt{(x - y)^T (x - y)}\}$$
$$= \bigcap_{y \in S} \{x \mid (x - x_0)^T (x - x_0) \leq (x - y)^T (x - y)\}$$
$$= \bigcap_{y \in S} \{x \mid \|x\|_2^2 + \|x_0\|_2^2 - 2x^T x_0 \leq \|x\|_2^2 + \|y\|_2^2 - 2x^T y\}$$
$$= \bigcap_{y \in S} \{x \mid 2x^T (y - x_0) \leq \|y\|_2^2 - \|x_0\|_2^2\}$$

Once again, S_{x_0} is an intersection of halfspaces and is therfore convex.

(e) We can esaily find a counter-example showing that the set is not convex; in $\mathbb{R}$ by choosing $S = \{0, 3\}$ and $T = \{1, 2\}$, we have :
$\{x \mid dist(x, S) \leq dist(x, T)\} = \{x \geq \frac{3}{2}\} \cup \{x \leq \frac{1}{2}\}$ which is not convex.

(f) Let $S = \{x \mid x + S_2 \subset S_1\}$ where $S_1, S_2 \subset \mathbb{R}^n$ with S_1 convex.
$S = \bigcap_{y \in S_2} \{x \mid x + y \in S_1\}$
Since S_1 is convex and affine transformations preserve the convexity, then the sets of the form $\{x \mid x + y \in S_1\}$ are convex. Therefore, S is convex an an intersection of convex sets.

(g) We can look at two possibilities :

- For $\theta = 1$, this case corresponds to question **(d)**, with $x_0 = a$ and $T = \{b\}$, so it is convex.
- For $\theta < 1$, let $S = \{x \mid \|x - a\|_2 \leq \|x - b\|_2\}$
  $S = \{x \mid (x - a)^T(x - a) \leq \theta^2(x - b)^T(x - b)\}$
  $= \{x \mid \|x\|_2^2 + \|a\|_2^2 - 2x^T a \leq \theta^2\|x\|_2^2 + \theta^2\|b\|_2^2 - 2\theta^2 x^T b\}$
  $= \{x \mid (1 - \theta^2)\|x\|_2^2 - 2x^T(a - \theta^2 b) \leq \theta^2\|b\|_2^2 - \|a\|_2^2\}$
  $= \{x \mid (1 - \theta^2)(\|x\|_2^2 - \dfrac{2x^T(a - \theta^2 b)}{(1 - \theta^2)} + \|\dfrac{a - \theta^2 b}{1 - \theta^2}\|_2^2) - \dfrac{\|a - \theta^2 b\|_2^2}{1 - \theta^2} \leq \theta^2\|b\|_2^2 - \|a\|_2^2\}$ which
  $= \{x \mid (1 - \theta^2)(\|(x - \dfrac{(a - \theta^2 b)}{(1 - \theta^2)}\|_2^2) \leq \theta^2\|b\|_2^2 - \|a\|_2^2 + \dfrac{\|a - \theta^2 b\|_2^2}{1 - \theta^2}\}$
  $= \{x \mid \|x - \dfrac{(a - \theta^2 b)}{(1 - \theta^2)}\|_2^2 \leq \dfrac{\theta^2\|b\|_2^2 - \|a\|_2^2}{1 - \theta^2} + \dfrac{\|a - \theta^2 b\|_2^2}{(1 - \theta^2)^2}\}$
  is a bowl, hence convex.

Exercise 3.21

(a) The functions of the form $\|A^{(i)}x - b^{(i)}\|$ are compositions of an affine function with the norm and are therefore convex.

$f(x) = max_{1,\ldots,k}\|A^{(i)}x - b^{(i)}\|$ being a point-wise maximum of those convex functions, it is convex as well.

(b) Let f be defined as $f(x) = \sum_{i=1}^{r} |x|_{[i]}$ where $|x|_{[1]}, |x|_{[2]}, \ldots, |x|_{[n]}$ are the absolute values of the components of x, sorted in decreasing order.

We can also write $f(x)$ as $f(x) = \max_{\{i_1,\ldots,i_r\}}\{\sum_{i_1}^{i_r} |x|_i\}$, where the functions of the form $\sum_{i_1}^{i_r} |x|_i$ are built by choosing r different components of x. Since each of these functions is convex, f is convex too since it is their pointwise-maximum.

Exercise 3.32

(a) Let f and g be convex, nondecreasing (or nonincreasing), and positive functions on an interval of $\mathbb{R}$.

Let α be a real number such as $0 \leq \alpha \leq 1$.

$$\begin{aligned}
f(\alpha x + (1-\alpha)y)g(\alpha x + (1-\alpha)y) &\leq (\alpha f(x) + (1-\alpha)f(y))(\alpha g(x) + (1-\alpha)g(y)) \\
&\quad \text{since f and g are convex and positive} \\
&= \alpha^2 fg(x) + (1-\alpha)^2 fg(y) + \alpha(1-\alpha)f(x)g(y) + \alpha(1-\alpha)g(x)f(y) \\
&= \alpha fg(x) + (1-\alpha)fg(y) - (\alpha-\alpha^2)fg(x) - (\alpha-\alpha^2)fg(y) \\
&\quad + \alpha(1-\alpha)f(x)g(y) + \alpha(1-\alpha)g(x)f(y) \\
&= \alpha fg(x) + (1-\alpha)fg(y) \\
&\quad + \alpha(1-\alpha)(f(x)g(y) + g(x)f(y) - fg(x) - fg(y)) \\
&= \alpha fg(x) + (1-\alpha)fg(y) + \alpha(1-\alpha)(g(x)-g(y))(f(y)-f(x)) \\
&\leq \alpha fg(x) + (1-\alpha)fg(y) \\
&\quad \text{since f and g are both non-decreasing (or non-increasing),} \\
&\quad \text{(g(x)-g(y)) and (f(y)-f(x)) have opposite signs, and their product is less than or equal to 0}
\end{aligned}$$

We verified the Jensen's inequality for fg, showing that it is convex.

(b) Let f, g be concave functions of $\mathbb{R}$, positive, with one nondecreasing and the other nonincreasing.

Let α be a real number such as $0 \leq \alpha \leq 1$.

$$\begin{aligned}
f(\alpha x + (1-\alpha)y)g(\alpha x + (1-\alpha)y) &\geq (\alpha f(x) + (1-\alpha)f(y))(\alpha g(x) + (1-\alpha)g(y)) \\
&\quad \text{since f and g are concave and positive} \\
&= \alpha^2 fg(x) + (1-\alpha)^2 fg(y) + \alpha(1-\alpha)f(x)g(y) + \alpha(1-\alpha)g(x)f(y) \\
&= \alpha fg(x) + (1-\alpha)fg(y) - (\alpha-\alpha^2)fg(x) - (\alpha-\alpha^2)fg(y) \\
&\quad + \alpha(1-\alpha)f(x)g(y) + \alpha(1-\alpha)g(x)f(y) \\
&= \alpha fg(x) + (1-\alpha)fg(y) \\
&\quad + \alpha(1-\alpha)(f(x)g(y) + g(x)f(y) - fg(x) - fg(y)) \\
&= \alpha fg(x) + (1-\alpha)fg(y) + \alpha(1-\alpha)(g(x)-g(y))(f(y)-f(x)) \\
&\geq \alpha fg(x) + (1-\alpha)fg(y) \\
&\quad \text{since f and g have different monotony, (g(x)-g(y)) and (f(y)-f(x)) have the same sign,} \\
&\quad \text{and their product is greater than or equal to 0}
\end{aligned}$$

Therefore, fg is concave.

(c) Let f be a convex, nondecreasing, and positive function of $\mathbb{R}$, and let g be a concave, nonincreasing, and positive function of $\mathbb{R}$.

Since g is concave, nonincreasing and positive, then $h = \frac{1}{g}$ is convex, nondecreasing and positive. Using the result of question **(a)**, $fh = \frac{f}{g}$ is convex.

Exercise 3.36

(a) Let f be defined on $\mathbb{R}^n$ as $f(x) = \max_{i=1,\dots,n}\ x_i$.

Its conjugate is written as :

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

- If $y_i = 0$ for all i then $f^*(y) = \sup_{x\in\mathbb{R}^n}(-\max_{i=1,\dots,n}\ x_i) = +\infty$ (by choosing $x = u\mathbf{1}$ with $u \to -\infty$)

- If there is one index j such as $y_j < 0$, then by looking at the x of the form $x = u\mathbf{e_j}$, we have : $f^*(y) \geq sup_u(x^T y - \max(0,u)) = \sup_u(uy_j - \max(0,u)) = +\infty$ when $u \to -\infty$.

- Let's admit that all the components of y are positive, and at least one is strictly positive.

  - If $\sum_{i=1}^n y_i > 1$, then by looking at the x of the form $x = u\mathbf{1}$, we have : $f^*(y) \geq sup_u(x^T y - u) = \sup_u((\sum_{i=1}^n y_i - 1)u) = +\infty$ when $u \to +\infty$.
  - If $\sum_{i=1}^n y_i < 1$, then by looking at the x of the form $x = u\mathbf{1}$, we have : $f^*(y) \geq sup_u(x^T y - u) = \sup_u((\sum_{i=1}^n y_i - 1)u) = +\infty$ when $u \to -\infty$.
  - If $\sum_{i=1}^n y_i = 1$, and by writing $X = \max_{i=1,\dots,n} x_i$ we have : $f^*(y) = sup_{x\in\mathbb{R}^n}(x^T y - X) \leq sup_{x\in\mathbb{R}^n}(X(\sum_{i=1}^n y_i - 1)) = 0$

    Since $f^*(\mathbf{0}_n) = 0$, then the conjugate is always 0 in this case.

To conclude, we can define f^* as :

$$f^*(y) = \begin{cases} 0, & \text{if } y \succ 0 \text{ and } \sum_{i=1}^n y_i = 1 \\ +\infty, & \text{otherwise} \end{cases}$$

(b) Let f be defined on $\mathbb{R}^n$ as $f(x) = \sum_{i=1}^r x_{[i]}$ the sum of the r largest components of x.

Its conjugate is written as :

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

- If $y_i = 0$ for all i then $f^*(y) = \sup_{x\in\mathbb{R}^n}(-\sum_{i=1}^r x_{[i]}) = +\infty$ (for example, by choosing $x = u\mathbf{1}$ with $u \to -\infty$)

- If there is one index j such as $y_j < 0$, then by looking at the x of the form $x = u\mathbf{e_j}$, we have : $f^*(y) \geq sup_u(x^T y - \sum_{i=1}^r x_{[i]})$.

  - If $r < n$, $\sup_u(uy_j - \max(0,u)) = +\infty$ when $u \to -\infty$.
  - If $r = n$, $\sup_u(uy_j - u) = +\infty$ when $u \to -\infty$.

- If there is one index j such as $y_j > 1$, then by looking at the x of the form $x = u\mathbf{e_j}$ with $u > 0$, we have : $f^*(y) \geq sup_u(x^T y - \sum_{i=1}^r x_{[i]}) = \sup_u(u(y_j - 1)) = +\infty$ when $u \to +\infty$.

- Let's admit that $1 \succeq y \succ 0$.

  - If $\sum_{i=1}^n y_i > r$, then by looking at the x of the form $x = u\mathbf{1}$, we have :
    $f^*(y) \geq sup_u(x^T y - ru) = \sup_u((\sum_{i=1}^n y_i - r)u) = +\infty$ when $u \to +\infty$.
  - If $\sum_{i=1}^n y_i < r$, then by looking at the x of the form $x = u\mathbf{1}$, we have :
    $f^*(y) \geq sup_u(x^T y - ru) = \sup_u((\sum_{i=1}^n y_i - r)u) = +\infty$ when $u \to -\infty$.
  - If $\sum_{i=1}^n y_i = r$, we have :
    $f^*(y) = sup_{x\in\mathbb{R}^n}(x^T y - ru) \leq sup_{x\in\mathbb{R}^n}(u(\sum_{i=1}^n y_i - r)) = 0$
    Since $f^*(\mathbf{0}_n) = 0$, then the conjugate is always 0 in this case.

To conclude, we can define f^* as :

$$f^*(y) = \begin{cases} 0, & \text{if } 1 \succeq y \succ 0 \text{ and } \sum_{i=1}^n y_i = r \\ +\infty, & \text{otherwise} \end{cases}$$

(c) Let f be defined on $\mathbb{R}$ as $f(x) = \max_{i=1,..,m}(a_i x + b_i)$. We assume that the a_i are sorted in increasing order, and that none of the functions $a_i x + b_i$ is redundant, i.e., for each k there is at least one x with $f(x) = a_k x + b_k$.

Its conjugate is written as :

$$\begin{aligned} f^*(y) &= \sup_{x\in dom(f)} (y^T x - f(x)) \\ &= \sup_{x\in\mathbb{R}}(xy - \max_{i=1,..,m}(a_i x + b_i)) \end{aligned}$$

- If $y = 0$ then $f^*(y) = \sup_{x\in\mathbb{R}^n}(-\max_{i=1,..,m}(a_i x + b_i)) = +\infty$ (obtained when $x \to -\infty$ or $x \to +\infty$, depending on the sign of the a_i).

- If $y < a_1$ then $f^*(y) = \sup_{x\in\mathbb{R}^n}(-\max_{i=1,..,m}((a_i - y)x + b_i)) = +\infty$ when $x \to -\infty$.

- If $y > a_m$ then $f^*(y) = \sup_{x\in\mathbb{R}^n}(-\max_{i=1,..,m}((a_i - y)x + b_i)) = +\infty$ when $x \to +\infty$.

- We suppose that $a_1 \leq y \leq a_m$.

  We have :$f^*(y) = \sup_{x\in\mathbb{R}^n}(-\max_{i=1,..,m}((a_i - y)x + b_i))$.

  To distinguish the a_i cases, we look at the break-points x_i of f. They happen at x_i when $a_i x_i + b_i = a_{i+1} x_i + b_{i+1}$,
  therefore $x_i = \frac{b_{i+1} - b_i}{a_i - a_{i+1}}$ is the breakpoint between $f(x) = a_i x + b_i$ and $f(x) = a_{i+1} x + b_{i+1}$.

  As a consequence, for $a_k \leq y \leq a_{k+1}$, and by distinguishing the cases $x \in [x_i, x_{i+1}]$, we can write :

$$
\begin{aligned}
f^*(y) &= \max_i \Big(\sup_{x \in [x_i, x_{i+1}]} \big(-\max_{j=1,..,m}((a_j - y)x + b_j)\big)\Big) \\
&= \max_i \Big(\sup_{x \in [x_i, x_{i+1}]} ((y - a_{i+1})x - b_{i+1})\Big) \\
&= (y - a_k)x_k - b_k
\end{aligned}
$$

since $a_k x_k + b_k = a_{k+1}x_k + b_{k+1}$

and x_{i+1} maximizes the sup for when $k \geq i + 1$ (so we choose $i = k - 1$ for the max);

and x_i when $k \leq i$ (so we choose $i = k$ for the max).

To conclude, we can define f^* as :

$$
f^*(y) = \begin{cases} (y - a_k)\frac{b_{k+1} - b_k}{a_k - a_{k+1}} - b_k, & \text{if } y \in [a_k, a_{k+1}] \\ +\infty, & \text{otherwise} \end{cases}
$$

(d) $\mathbf{p > 1}$

Let f be defined on $\mathbb{R}_{++}$ as $f(x) = x^p$ where $p > 1$.

Its conjugate is written as :

$$
\begin{aligned}
f^*(y) &= \sup_{x \in dom(f)} (y^T x - f(x)) \\
&= \sup_{x \in \mathbb{R}_{++}} (xy - x^p)
\end{aligned}
$$

- If $y \leq 0$, $f^*(y) = \sup_{x \in \mathbb{R}_{++}} (x(y - x^{p-1}))$ with $(y - x^{p-1}) \leq 0$. Therefore, the value of the sup is 0, reached when $x \to 0$.

- If $y > 0$, let's define a function g_y by $g_y(x) = xy - x^p$.
  g_y reaches an extremum for $g_y'(x) = 0$, i.e $y - px^{p-1} = 0$, i.e $x = (\frac{y}{p})^{\frac{1}{p-1}}$ on $\mathbb{R}_{++}$. Since g_y is concave ($g_y''(x) = -p(p-1)x^{p-2} \leq 0$) on $\mathbb{R}_{++}$, its sup is reached at this point.
  As a consequence,
  $$
  \begin{aligned}
  f^*(y) &= y(\frac{y}{p})^{\frac{1}{p-1}} - (\frac{y}{p})^{\frac{p}{p-1}} \\
  &= p(\frac{y}{p})^{\frac{1}{p-1}+1} - (\frac{y}{p})^{\frac{p}{p-1}}. \\
  &= (p-1)\left(\frac{y}{p}\right)^{\frac{p}{p-1}}
  \end{aligned}
  $$

To conclude, we can define f^* as :

$$
f^*(y) = \begin{cases} (p-1)\left(\frac{y}{p}\right)^{\frac{p}{p-1}}, & \text{if } y > 0 \\ 0, & \text{otherwise} \end{cases}
$$

(d) $\mathbf{p < 0}$

Let f be defined on $\mathbb{R}_{++}$ as $f(x) = x^p$ where $p < 0$.

Its conjugate is written as :

$$\begin{aligned} f^*(y) &= \sup_{x \in dom(f)} (y^T x - f(x)) \\ &= \sup_{x \in \mathbb{R}_{++}} (xy - x^p) \end{aligned}$$

- If $y > 0$, $f^*(y) = \sup_{x \in \mathbb{R}_{++}} (x(y - x^{p-1}))$ with $(y - x^{p-1}) \to y$ when $x \to +\infty$. Therefore, the value of the sup is $+\infty$, reached when $x \to +\infty$.
- If $y \leq 0$, let's define a function g_y by $g_y(x) = xy - x^p$.
  g_y reaches an extremum for $g_y'(x) = 0$, i.e $y - px^{p-1} = 0$, i.e $x = -(\frac{-y}{p})^{\frac{1}{p-1}}$ on $\mathbb{R}_{++}$. Since g_y is concave on $\mathbb{R}_{++}$, its sup is reached at this point.
  As a consequence, $f^*(y) = -(p-1)\left(\frac{-y}{p}\right)^{\frac{p}{p-1}}$.

To conclude, we can define f^* as :

$$f^*(y) = \begin{cases} -(p-1)\left(\frac{-y}{p}\right)^{\frac{p}{p-1}}, & \text{if } y \leq 0 \\ +\infty, & \text{otherwise} \end{cases}$$