Homework 1

YACHAOUI Ayman
Convex Optimization
ayman.yachaoui@telecom-paristech.fr

March 29, 2017

A Convex sets

Which of the following sets are convex?

(a) $S = \{x \in \mathbb{R}^n \mid \alpha \leq a^\intercal x \leq \beta\}$
For all $(u, v, \theta) \in S^2 \times [0, 1]$ let $w = \theta u + (1-\theta)v$. Then:
$a^\intercal w = \theta a^\intercal u + (1-\theta) a^\intercal v \geq \theta\alpha + (1-\theta)\alpha = \alpha$
$a^\intercal w = \theta a^\intercal u + (1-\theta) a^\intercal v \leq \theta\beta + (1-\theta)\beta = \beta$
Finally we have $\alpha \leq a^\intercal w \leq \beta$ so $w \in S$ and thus S is a convex set.

(b) $S = \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i, \forall i = 1, ..., n\}$
For all $(u, v, \theta) \in S^2 \times [0, 1]$ let $w = \theta u + (1-\theta)v$. Then for all $i = 1, ..., n$:
$w_i = \theta u_i + (1-\theta)v_i \geq \alpha_i + (1-\theta)\alpha_i = \alpha_i$
$w_i = \theta u_i + (1-\theta)v_i \leq \beta_i + (1-\theta)\beta_i = \beta_i$
Finally we have $\alpha_i \leq w_i \leq \beta_i$ so $w \in S$ and thus S is a convex set.

(c) $S = \{x \in \mathbb{R}^n \mid a_1^\intercal x \leq b_1, a_2^\intercal x \leq b_2\}$
For all $(u, v, \theta) \in S^2 \times [0, 1]$ let $w = \theta u + (1-\theta)v$. Then:
$a_1^\intercal w = \theta a_1^\intercal u + (1-\theta) a_1^\intercal v \leq \theta b_1 + (1-\theta) b_1 = b_1$
$a_2^\intercal w = \theta a_2^\intercal u + (1-\theta) a_2^\intercal v \leq \theta b_2 + (1-\theta) b_2 = b_2$
Finally we have $a_1^\intercal w \leq b_1$ and $a_2^\intercal w \leq b_2$ so $w \in S$ and thus S is a convex set.

(d) $S = \{x \mid \|x - x_0\|_2 \leq \|x - y\|_2, \forall y \in S' \subseteq \mathbb{R}^n\}$
For all $(u, v, y, \theta) \in S^2 \times S' \times [0, 1]$ let $w = \theta u + (1-\theta)v$. Then:

$$\begin{aligned}
\|w - x_0\|_2 &= \|\theta(u - x_0) + (1-\theta)(v - x_0)\|_2 \\
&\leq \theta\|u - x_0\|_2 + (1-\theta)\|v - x_0\|_2 \\
&\leq \theta\|u - y\|_2 + (1-\theta)\|v - y\|_2 \\
&\leq \|x - y\|_2
\end{aligned}$$

Finally $w \in S$ and thus S is a convex set.

(e) $S' = \{x \mid \mathbf{dist}(x, S) \leq \mathbf{dist}(x, T)\}$ where $S, T \subseteq \mathbb{R}^n$
This set is not a convex set, to illustrate this we will exhibit a counter-example in $\mathbb{R}$.
Let $S = \{-1, 1\}$ and $T = \{0\}$, we then chose $u = -2$ and $v = 2$. We observe that $(u, v) \in S'^2$. If we now pick $w = \frac{u}{2} + \frac{v}{2} = 0$ then $w \notin S'$. Thus S' is not convex.

(f) $S = \{x \mid x + S_2 \subseteq S_1\}$ where $S_1, S_2 \subseteq \mathbb{R}^n$ with S_1 convex
For all $(u, v, y, \theta) \in S^2 \times S_2 \times [0, 1]$ let $w = \theta u + (1-\theta)v$. Then:

$$
\begin{aligned}
w + y &= \theta u + (1-\theta)v + y \\
&= \theta(u+y) + (1-\theta)(v+y)
\end{aligned}
$$

By definition of the set S, $u + y \in S_1$ and $v + y \in S_1$. It follows by definition of a convex set that $w + y \in S_1$. As this was demonstrated for every y we have the results $w + S_2 \subseteq S_1$ and thus finally S is a convex set.

(g) $S = \{x \mid \|x-a\|_2 \leq \theta \|x-b\|_2\}$ with $a \neq b$ and $0 \leq \theta \leq 1$
Let $x \in S$, then:

$$
\begin{aligned}
0 &\geq \|x-a\|_2^2 - \theta^2 \|x-b\|_2^2 \\
&\geq (x-a)^\intercal (x-a) - \theta^2 (x-b)^\intercal (x-b) \\
&\geq x^\intercal x - 2a^\intercal x + a^\intercal a - \theta^2 x^\intercal x + 2\theta^2 b^\intercal x - \theta^2 b^\intercal b \\
&\geq (1-\theta^2) x^\intercal x - 2(a^\intercal - \theta^2 b^\intercal) x + a^\intercal a - \theta^2 b^\intercal b \\
&\geq x^\intercal x - 2\frac{(a-\theta^2 b)^\intercal}{1-\theta^2} x + \frac{a^\intercal a - \theta^2 b^\intercal b}{1-\theta^2}
\end{aligned}
$$

By rearranging we have $\|x - x_c\|_2^2 \leq r^2$ with:

$$
x_c = \frac{a - \theta^2 b}{1 - \theta^2}
$$

$$
\begin{aligned}
r^2 &= \frac{(a-\theta^2 b)^\intercal (a - \theta^2 b)}{(1-\theta^2)^2} - \frac{a^\intercal a - \theta^2 b^\intercal b}{1-\theta^2} \\
&= \frac{a^\intercal a + \theta^4 b^\intercal b - 2\theta^2 a^\intercal b - (1-\theta^2) a^\intercal a + (1-\theta^2)\theta^2 b^\intercal b}{(1-\theta^2)^2} \\
&= \frac{\theta^2 a^\intercal a + \theta^2 b^\intercal b - 2\theta^2 a^\intercal b}{(1-\theta^2)^2} \\
&= \frac{\theta^2 (a-b)^\intercal (a-b)}{(1-\theta^2)^2} \\
&= \left(\frac{\theta}{1-\theta^2}\right)^2 \|a-b\|_2^2
\end{aligned}
$$

As a consequence S is a ball and as thus S is a convex set.

B Pointwise maximum and supremum

Show that the following functions $f : \mathbb{R}^n \to \mathbb{R}$ are convex.

(a) $f(x) = \max_{i=1,\ldots,k} \|A^{(i)} x - b^{(i)}\|$, where $A^{(i)} \in \mathbb{R}^{m \times n}$, $b^{(i)} \in \mathbb{R}^m$ and $\|.\|$ is a norm on $\mathbb{R}^m$.
For all $(x, y, \theta) \in (\mathbb{R}^n)^2 \times [0, 1]$, we have:

$$
\begin{aligned}
f(\theta x + (1-\theta) y) &= \max_{i=1,\ldots,k} \|A^{(i)}(\theta x + (1-\theta) y) - b^{(i)}\| \\
&= \max_{i=1,\ldots,k} \|\theta(A^{(i)} x - b^{(i)}) + (1-\theta)(A^{(i)} y - b^{(i)})\| \\
&\leq \theta \max_{i=1,\ldots,k} \|A^{(i)} x - b^{(i)}\| + (1-\theta) \max_{i=1,\ldots,k} \|A^{(i)} y - b^{(i)}\| \\
&\leq \theta f(x) + (1-\theta) f(y)
\end{aligned}
$$

Thus as **dom** $f = \mathbb{R}^n$ is convex, we can deduce that f is convex.

(b) $f(x) = \sum_{i=1}^{r} |x|_{[i]}$ on $\mathbb{R}^n$.
First let's observe that we can rewrite f as:
$f(x) = \max\{\sum_{k=i_1}^{i_r} |x_k| \mid 1 \leq i_1 \leq i_2 \leq ... \leq i_r \leq n\}$, then for all $(x, y, \theta) \in (\mathbb{R}^n)^2 \times [0, 1]$, we have:

$$
\begin{aligned}
f(\theta x + (1-\theta)y) &= \max\{\sum_{k=i_1}^{i_r} |(\theta x + (1-\theta)y)_k| \mid 1 \leq i_1 \leq i_2 \leq ... \leq i_r \leq n\} \\
&\leq \max\{\sum_{k=i_1}^{i_r} (\theta|x_k| + (1-\theta)|y_k|) \mid 1 \leq i_1 \leq i_2 \leq ... \leq i_r \leq n\} \\
&\leq \max\{\theta\left(\sum_{k=i_1}^{i_r} |x_k|\right) + (1-\theta)\left(\sum_{k=i_1}^{i_r} |y_k|\right) \mid 1 \leq i_1 \leq i_2 \leq ... \leq i_r \leq n\} \\
&\leq \max\{\theta\left(\sum_{k=i_1}^{i_r} |x_k|\right) \mid 1 \leq i_1 \leq i_2 \leq ... \leq i_r \leq n\} \\
&+ \max\{(1-\theta)\left(\sum_{k=i_1}^{i_r} |y_k|\right) \mid 1 \leq i_1 \leq i_2 \leq ... \leq i_r \leq n\} \\
&\leq \theta \max\{\sum_{k=i_1}^{i_r} |x_k| \mid 1 \leq i_1 \leq i_2 \leq ... \leq i_r \leq n\} \\
&+ (1-\theta) \max\{\sum_{k=i_1}^{i_r} |y_k| \mid 1 \leq i_1 \leq i_2 \leq ... \leq i_r \leq n\} \\
&\leq \theta f(x) + (1-\theta) f(y)
\end{aligned}
$$

Thus as **dom** $f = \mathbb{R}^n$ is convex, we can deduce that f is convex.

C Products and ratios of convex functions

In general the product or ratio of two convex functions is not convex. However, there are some results that apply to functions on $\mathbb{R}$. Prove the following.

(a) If f and g are convex, both nondecreasing (or nonincreasing), and positive functions on an interval, then fg is convex.
For all $(x, y, \theta) \in (\textbf{dom } fg)^2 \times [0, 1]$, we have:

$$
\begin{aligned}
fg(\theta x + (1-\theta)y) &= f(\theta x + (1-\theta)y) \times g(\theta x + (1-\theta)y) \\
&\leq (\theta f(x) + (1-\theta) f(y)) \times (\theta g(x) + (1-\theta) g(y)) \\
&\leq \theta^2 f(x)g(x) + (1-\theta)^2 f(y)g(y) + \theta(1-\theta)(f(x)g(y) + f(y)g(x))
\end{aligned}
$$

Besides:

$$
\begin{aligned}
f(x)g(y) + f(y)g(x) &= f(x)g(y) + f(y)g(x) - f(x)g(x) - f(y)g(y) + f(x)g(x) + f(y)g(y) \\
&= (f(x) - f(y))(g(y) - g(x)) + f(x)g(x) + f(y)g(y)
\end{aligned}
$$

Finally as f and g are both nondecreasing (or nonincreasing) then:

$$
\begin{aligned}
&(f(x)-f(y))(g(y)-g(x)) \leq 0 \\
\Longrightarrow\ & f(x)g(y)+f(y)g(x) \leq f(x)g(x)+f(y)g(y) \\
\Longrightarrow\ & fg(\theta x+(1-\theta)y) \leq \theta^2 f(x)g(x)+(1-\theta)^2 f(y)g(y)+\theta(1-\theta)(f(x)g(x)+f(y)g(y)) \\
\Longrightarrow\ & fg(\theta x+(1-\theta)y) \leq \theta f(x)g(x)+(1-\theta)f(y)g(y) \\
\Longrightarrow\ & fg(\theta x+(1-\theta)y) \leq \theta fg(x)+(1-\theta)fg(y)
\end{aligned}
$$

Thus as **dom** fg is an interval it is by definition convex and we can deduce that fg is convex.

(b) If f and g are concave, positive, with one nondecreasing and the other nonincreasing, then fg is concave. For all $(x, y, \theta) \in (\textbf{dom}\ fg)^2 \times [0,1]$, we have in the same spirit as the precedent question:

$$
\begin{aligned}
fg(\theta x+(1-\theta)y) &= f(\theta x+(1-\theta)y) \times g(\theta x+(1-\theta)y) \\
&\geq (\theta f(x)+(1-\theta)f(y)) \times (\theta g(x)+(1-\theta)g(y)) \\
&\geq \theta^2 f(x)g(x)+(1-\theta)^2 f(y)g(y)+\theta(1-\theta)(f(x)g(y)+f(y)g(x))
\end{aligned}
$$

And:

$$
\begin{aligned}
f(x)g(y)+f(y)g(x) &= f(x)g(y)+f(y)g(x)-f(x)g(x)-f(y)g(y)+f(x)g(x)+f(y)g(y) \\
&= (f(x)-f(y))(g(y)-g(x))+f(x)g(x)+f(y)g(y)
\end{aligned}
$$

Finally as f and g are one nondecreasing and the other nonincreasing:

$$
\begin{aligned}
&(f(x)-f(y))(g(y)-g(x)) \geq 0 \\
\Longrightarrow\ & f(x)g(y)+f(y)g(x) \geq f(x)g(x)+f(y)g(y) \\
\Longrightarrow\ & fg(\theta x+(1-\theta)y) \geq \theta^2 f(x)g(x)+(1-\theta)^2 f(y)g(y)+\theta(1-\theta)(f(x)g(x)+f(y)g(y)) \\
\Longrightarrow\ & fg(\theta x+(1-\theta)y) \geq \theta f(x)g(x)+(1-\theta)f(y)g(y) \\
\Longrightarrow\ & fg(\theta x+(1-\theta)y) \geq \theta fg(x)+(1-\theta)fg(y)
\end{aligned}
$$

Thus as **dom** fg is an interval it is by definition convex and we can deduce that fg is concave.

(c) If f is convex, nondecreasing, and positive, and g is concave, nonincreasing, and positive, then f/g is convex. For all $(x, y, \theta) \in (\textbf{dom}\ f/g)^2 \times [0,1]$, we have:

$$
\begin{aligned}
f/g(\theta x+(1-\theta)y) &= \frac{f(\theta x+(1-\theta)y)}{g(\theta x+(1-\theta)y)} \\
&= \frac{f(\theta x+(1-\theta)y)}{g(\theta x+(1-\theta)y)} - \theta\frac{f(x)}{g(x)} - (1-\theta)\frac{f(y)}{g(y)} + \theta\frac{f(x)}{g(x)} + (1-\theta)\frac{f(y)}{g(y)}
\end{aligned}
$$

Besides:

$$
\begin{aligned}
r &= \frac{f(\theta x+(1-\theta)y)}{g(\theta x+(1-\theta)y)} - \theta\frac{f(x)}{g(x)} - (1-\theta)\frac{f(y)}{g(y)} \\
&= \frac{g(x)g(y)f(\theta x+(1-\theta)y) - g(\theta x+(1-\theta)y)(\theta f(x)g(y)+(1-\theta)f(y)g(x))}{g(x)g(y)g(\theta x+(1-\theta)y)} \\
&\le \frac{g(x)g(y)(\theta f(x)+(1-\theta)f(y)) - (\theta g(x)+(1-\theta)g(y))(\theta f(x)g(y)+(1-\theta)f(y)g(x))}{g(x)g(y)g(\theta x+(1-\theta)y)} \\
&\le \frac{\theta g(x)g(y)f(x)+(1-\theta)g(x)g(y)f(y) - \theta^2 g(x)g(y)f(x) - (1-\theta)^2 g(x)g(y)f(y) - \theta(1-\theta)(g(x)^2f(y)+g(y)^2f(x))}{g(x)g(y)g(\theta x+(1-\theta)y)} \\
&\le \frac{\theta f(x)(g(x)g(y)-\theta g(x)g(y)-(1-\theta)g(y)^2) + (1-\theta)f(y)(g(x)g(y)-(1-\theta)g(x)g(y)-\theta g(x)^2)}{g(x)g(y)g(\theta x+(1-\theta)y)} \\
&\le \frac{\theta(1-\theta)f(x)g(y)(g(x)-g(y)) + \theta(1-\theta)f(y)g(x)(g(y)-g(x))}{g(x)g(y)g(\theta x+(1-\theta)y)} \\
&\le \frac{\theta(1-\theta)(f(x)g(y)-f(y)g(x))(g(x)-g(y))}{g(x)g(y)g(\theta x+(1-\theta)y)}
\end{aligned}
$$

Also we can suppose $x \le y$ without loss of generality, then:

$$
\begin{aligned}
& g(x) \ge g(y) \ge 0, \text{ because } g \text{ is positive nonincreasing.} \\
& f(y) \ge f(x) \ge 0, \text{ because } g \text{ is positive nondecreasing.} \\
\Longrightarrow\ & f(y)g(x) \ge f(x)g(y) \\
\Longrightarrow\ & f(x)g(y) - f(y)g(x) \le 0
\end{aligned}
$$

Finally $r \le 0$ and:

$$
\begin{aligned}
f/g(\theta x+(1-\theta)y) &= r + \theta\frac{f(x)}{g(x)} + (1-\theta)\frac{f(y)}{g(y)} \\
&\le \theta\frac{f(x)}{g(x)} + (1-\theta)\frac{f(y)}{g(y)}
\end{aligned}
$$

Thus as **dom** f/g is an interval it is by definition convex and we can deduce that f/g is convex.

D Conjugate functions

Derive the conjugates of the following functions.

(a) *Max function.* $f(x) = \max_{i=1,\dots,n} x_i$ on $\mathbb{R}^n$.
Then for $(y,\lambda) \in \mathbb{R}^n \times \mathbb{R}_+$, we have:

$$
\begin{aligned}
f^*(y) &= \sup_{x\in\mathbb{R}^n} (y^\intercal x - f(x)) \\
&= \sup_{x\in\mathbb{R}^n} (y^\intercal x - \max_{i=1,\dots,n} x_i) \\
&= \sup_{x\in\mathbb{R}^n} \left(\sum_{i=1}^n y_i x_i - \max_{i=1,\dots,n} x_i \right)
\end{aligned}
$$

- If $\exists i$, such that $y_i < 0$, we take z with $z_i = -\lambda$ and all other coordinates nul. Then:

$$
\begin{aligned}
& y^\intercal z - f(z) = \sum_{k=1}^n y_k z_k - \max_{k=1,\dots,n} z_k = -\lambda y_i \\
\Longrightarrow\ & \lim_{\lambda\to+\infty} y^\intercal z - f(z) = +\infty \\
\Longrightarrow\ & f^*(y) = +\infty
\end{aligned}
$$

- Otherwise $y \succeq 0$, and thus:

$$
\begin{aligned}
y^\intercal x - f(x) &= \sum_{i=1}^{n} y_i x_i - \max_{i=1,\ldots,n} x_i \\
&\leq \left(\sum_{i=1}^{n} y_i\right) x_{\max} - x_{\max} \\
&\leq \left(\sum_{i=1}^{n} y_i - 1\right) x_{\max}
\end{aligned}
$$

  - If $\sum_{i=1}^{n} y_i \neq 1$ we chose $z = -\lambda * \text{sign}(1 - \sum_{i=1}^{n} y_i) * (1, 1, ..., 1)^\intercal$. Then:

$$
\begin{aligned}
y^\intercal z - f(z) &= \sum_{i=1}^{n} y_i z_i - \max_{i=1,\ldots,n} z_i \\
&= -\lambda * \text{sign}\left(1 - \sum_{i=1}^{n} y_i\right) * \left(\sum_{i=1}^{n} y_i\right) + \lambda * \text{sign}\left(1 - \sum_{i=1}^{n} y_i\right) \\
&= \lambda * \text{sign}\left(1 - \sum_{i=1}^{n} y_i\right) * \left(1 - \sum_{i=1}^{n} y_i\right) \\
&= \lambda \left|1 - \sum_{i=1}^{n} y_i\right| \\
\implies & \lim_{\lambda \to +\infty} y^\intercal z - f(z) = +\infty \\
\implies & f^*(y) = +\infty
\end{aligned}
$$

  - Finally if $\sum_{i=1}^{n} y_i = 1$ then $y^\intercal x - f(x) \leq 0$ and the equality holds for z equal to the nul vector. Thus we have: $f^*(y) = 0$.

To summarize:

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

(b) *Sum of largest elements.* $f(x) = \sum_{i=1}^{r} x_{[i]}$ on $\mathbb{R}^n$.
First let's suppose that $r < n$. Then for $(y, \lambda) \in \mathbb{R}^n \times \mathbb{R}_+$, we have:

$$
\begin{aligned}
f^*(y) &= \sup_{x \in \mathbb{R}^n} (y^\intercal x - f(x)) \\
&= \sup_{x \in \mathbb{R}^n} \left(y^\intercal x - \sum_{i=1}^{r} x_{[i]}\right) \\
&= \sup_{x \in \mathbb{R}^n} \left(\sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{r} x_{[i]}\right)
\end{aligned}
$$

- If $\exists i$, such that $y_i < 0$, we take z with $z_i = -\lambda$ and all other coordinates nul. Then:

$$
\begin{aligned}
y^\intercal z - f(z) &= \sum_{k=1}^{n} y_k z_k - \sum_{k=1}^{r} z_{[k]} = -\lambda y_i \\
\implies & \lim_{\lambda \to +\infty} y^\intercal z - f(z) = +\infty \\
\implies & f^*(y) = +\infty
\end{aligned}
$$

- Otherwise $y \succeq 0$, and thus:
  - If $\exists i$, such that $y_i > 1$, we take z with $z_i = \lambda$ and all other coordinates nul. Then:

$$\begin{aligned}
y^\intercal z - f(z) &= \sum_{k=1}^{n} y_k z_k - \sum_{k=1}^{r} z_{[k]} = \lambda(y_i - 1) \\
\implies &\lim_{\lambda\to+\infty} y^\intercal z - f(z) = +\infty \\
\implies &f^*(y) = +\infty
\end{aligned}$$

  - Otherwise $0 \preceq y \preceq 1$, and:

$$\begin{aligned}
y^\intercal x - f(x) &= \sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{r} x_{[i]} \\
&\geq \sum_{i=1}^{n} y_i x_{[n]} - \sum_{i=1}^{r} x_{[1]} \\
&\geq x_{[n]} \left(\sum_{i=1}^{n} y_i \right) - r x_{[1]}
\end{aligned}$$

    - Thus if $\sum_{i=1}^{n} y_i \neq r$, we take $z = -\lambda * \operatorname{sign}(r - \sum_{i=1}^{n} y_i) * (1, 1, ..., 1)^\intercal$:

$$\begin{aligned}
y^\intercal z - f(z) &= \sum_{i=1}^{n} y_i z_i - \sum_{i=1}^{r} z_{[i]} \\
&= \lambda * \operatorname{sign}\left(r - \sum_{i=1}^{n} y_i \right) \left(r - \sum_{i=1}^{n} y_i \right) \\
&= \lambda \left| r - \sum_{i=1}^{n} y_i \right| \\
\implies &\lim_{\lambda\to+\infty} y^\intercal z - f(z) = +\infty \\
\implies &f^*(y) = +\infty
\end{aligned}$$

    - Otherwise finally, $0 \preceq y \preceq 1$, $\sum_{i=1}^{n} y_i = r$ and thus:

$$\begin{aligned}
y^\intercal x - f(x) &= \sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{r} x_{[i]} \\
&\leq \sum_{i=1}^{n} y_{[i]} x_{[i]} - \sum_{i=1}^{r} x_{[i]} \\
&\leq \sum_{i=1}^{r} \left(y_{[i]} - 1 \right) x_{[i]} - \sum_{i=r+1}^{n} y_{[i]} x_{[i]} \\
&\leq \sum_{i=1}^{r} \left(y_{[i]} - 1 \right) x_{[r]} - \sum_{i=r+1}^{n} y_{[i]} x_{[r]} \\
&\leq x_{[r]} \left(\sum_{i=1}^{r} (y_{[i]} - 1) - \sum_{i=r+1}^{n} y_{[i]} \right) \\
&\leq x_{[r]} \left(r - \sum_{i=1}^{n} y_{[i]} \right) \\
&\leq 0
\end{aligned}$$

    Besides the equality holds for z equal to the unity vector. Thus we have: $f^*(y) = 0$.

To summarize:

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

Now we take a final look at the case $r = n$, then:

$$\begin{aligned} f^*(y) &= \sup_{x \in \mathbb{R}^n} (y^\intercal x - f(x)) \\ &= \sup_{x \in \mathbb{R}^n} \left(y^\intercal x - \sum_{i=1}^n x_{[i]} \right) \\ &= \sup_{x \in \mathbb{R}^n} \left(\sum_{i=1}^n y_i x_i - \sum_{i=1}^n x_i \right) \\ &= \sup_{x \in \mathbb{R}^n} \left(\sum_{i=1}^n (y_i - 1) x_i \right) \\ &= \begin{cases} 0 & \forall i, y_i = 1 \\ +\infty & \text{otherwise} \end{cases} \end{aligned}$$

(c) *Piecewise-linear function.* $f(x) = \max_{i=1,\ldots,m}(a_i x + b_i)$ on $\mathbb{R}$.
Then we have:

$$\begin{aligned} f^*(y) &= \sup_{x \in \mathbb{R}} (yx - f(x)) \\ &= \sup_{x \in \mathbb{R}} \left(yx - \max_{i=1,\ldots,m} (a_i x + b_i) \right) \end{aligned}$$

We can suppose that $a_1 \leq y \leq a_m$ otherwise $yx - f(x) \sim (y - a_{\max})x$ with $(y - a_{\max}) \neq 0$ which is not bounded on one side.
As a consequence there exists k such that $a_k \leq y \leq a_{k+1}$. Besides, $x \mapsto f(x)$ is convex, thus $y \mapsto f^*(y)$ is concave, continuous, derivable almost everywhere (except at the junction points), and so we have that the maximum is attained when the slop of the affine function changes sign (note that it is not neceserraly unique).
As $a_k - y \leq 0$ and $a_{k+1} - y \geq 0$, the maximum is attained at the junction point between a_k and a_{k+1} which is the point x_0 verifying:

$$\begin{aligned} a_k x_0 + b_k &= a_{k+1} x_0 + b_{k+1} \\ x_0 &= -\frac{b_{k+1} - b_k}{a_{k+1} - a_k} \end{aligned}$$

Finally:

$$f^*(y) = \begin{cases} -\frac{(y-a_k)(b_{k+1}-b_k)}{a_{k+1}-a_k} & \exists k \mid a_k \leq y \leq a_{k+1} \\ +\infty & \text{otherwise} \end{cases}$$

(d) *Power function.* $f(x) = x^p$ on $\mathbb{R}_{++}$ with $p > 1$ then $p < 0$.

- $p > 1$, then:

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

Besides $g : x \mapsto yx - x^p$ is continuous, derivable and $\lim_{x \to +\infty} g(x) = -\infty$ thus there is a maximum. If $y < 0$ then g is decreasing and it follows that $f^*(y) = g(0) = 0$ otherwise, there is a maximum which is attained for x_0 such that $g'(x_0) = 0$:

$$g'(x_0) = y - p x_0^{p-1} = 0$$

$$x_0 = \left(\frac{y}{p}\right)^{\frac{1}{p-1}}$$
$$f^*(y) = \left(\frac{y}{p}\right)^{\frac{1}{p-1}} \left(y - \left(\frac{y}{p}\right)^{\frac{1}{p-1}}\right)$$

Therefore:

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

- $p < 0$
  - If $y > 0$ then $\lim_{x \to +\infty} g(x) = +\infty$ thus $f^*(y) = +\infty$.
  - If $y = 0$ then $f^*(y) = 0$
  - Otherwise $y < 0$ then g is continuous $\lim_{x \to 0} g(x) = -\infty$ and $\lim_{x \to +\infty} g(x) = -\infty$ thus there is a maximum which is attained for x_0 such that $g'(x_0) = 0$ and we are in the same situation as before.

  Therefore:

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

(e) *Negative geometric mean.* $f(x) = -\left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}}$ on $\mathbb{R}^n_{++}$.
Then for $(y, \lambda) \in \mathbb{R}^n_{++} \times \mathbb{R}_+$, we have:

$$\begin{aligned} f^*(y) &= \sup_{x \in \mathbb{R}^n_{++}} \left(y^\intercal x - f(x)\right) \\ &= \sup_{x \in \mathbb{R}^n_{++}} \left(\sum_{i=1}^n y_i x_i + \left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}}\right) \end{aligned}$$

- If there exists k such that $y_k \geq 0$ then we chose the vector z equal to 1 everywhere and $z_k = \lambda$, thus:

$$\begin{aligned} y^\intercal z - f(z) &= \sum_{i=1}^n y_i z_i + \left(\prod_{i=1}^n z_i\right)^{\frac{1}{n}} \\ &= \sum_{i \neq k} y_i + \lambda y_k + \lambda^{\frac{1}{n}} \end{aligned}$$

  Thus $\lim_{\lambda \to +\infty} y^\intercal z - f(z) = +\infty$ and $f^*(y) = +\infty$.
- Otherwise $y \prec 0$ and we now chose z such that $(z_i)_{1 \leq i \leq n} = (-\lambda / y_i)_{1 \leq i \leq n} \in \mathbb{R}^n_{++}$ then:

$$\begin{aligned} y^\intercal z - f(z) &= \sum_{i=1}^n y_i z_i + \left(\prod_{i=1}^n z_i\right)^{\frac{1}{n}} \\ &= -\lambda n + \left(\prod_{i=1}^n \frac{-\lambda}{y_i}\right)^{\frac{1}{n}} \\ &= \lambda \left(-n + \left(\prod_{i=1}^n \frac{-1}{y_i}\right)^{\frac{1}{n}}\right) \\ &= \lambda \left(-n + \frac{1}{\left(\prod_{i=1}^n -y_i\right)^{\frac{1}{n}}}\right) \end{aligned}$$

  Therefore if $\left(\prod_{i=1}^n -y_i\right)^{\frac{1}{n}} < \frac{1}{n}$ then $\lim_{\lambda \to +\infty} y^\intercal z - f(z) = +\infty$ and consequently $f^*(y) = +\infty$.

- Finally let's examine the case where $y \prec 0$ and $(\prod_{i=1}^n -y_i)^{\frac{1}{n}} \geq \frac{1}{n}$. Then:

$$
\begin{aligned}
-\sum_{i=1}^n y_i x_i &= \sum_{i=1}^n -y_i x_i \text{ with } -y_i x_i > 0 \\
&\geq n \left(\prod_{i=1}^n -y_i x_i \right)^{\frac{1}{n}}
\end{aligned}
$$

  Then we can write:

$$
\begin{aligned}
y^\intercal x - f(x) &= \sum_{i=1}^n y_i x_i + \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}} \\
&\leq -n \left(\prod_{i=1}^n -y_i x_i \right)^{\frac{1}{n}} + \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}} \\
&\leq \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}} \left(1 - n \left(\prod_{i=1}^n -y_i \right)^{\frac{1}{n}} \right) \\
&\leq 0
\end{aligned}
$$

  Besides the equality holds for the nul vector.

To summarize:

$$
f^*(y) = \begin{cases} 0 & y \prec 0 \text{ and } (\prod_{i=1}^n -y_i)^{\frac{1}{n}} \geq \frac{1}{n} \\ +\infty & \text{otherwise} \end{cases}
$$

(f) *Negative generalized logarithm for second order cone.* $f(x,t) = -\log(t^2 - x^\intercal x)$ on $\{(x,t) \in \mathbb{R}^n \times \mathbb{R} \mid \|x\|_2 < t\}$. Then:

$$
\begin{aligned}
f^*(y,u) &= \sup_{(x,t) \in \mathbb{R}^n \times \mathbb{R}} (y^\intercal x + ut - f(x,t)) \\
&= \sup_{(x,t) \in \mathbb{R}^n \times \mathbb{R}} \left(y^\intercal x + ut + \log(t^2 - \|x\|_2^2) \right)
\end{aligned}
$$

Let $g : (x,t) \mapsto y^\intercal x + ut + \log(t^2 - \|x\|_2^2)$.
First we observe that necessarily $t > \|x\|_2 \geq 0$.

- If $u \geq 0$ then $\lim_{t \to +\infty} g(0,t) = +\infty$, thus $f^*(y,u) = +\infty$
- If $u < 0$ then we observe that for x fixed, $\lim_{t \to 0} g(x,t) = -\infty$ and $\lim_{t \to +\infty} g(x,t) = -\infty$, thus as g is continuous and derivable, g admits a maximum in t_x for x fixed and thus:

$$
\begin{aligned}
\frac{\partial g}{\partial t}(x, t_x) &= u + \frac{2t_x}{t_x^2 - \|x\|_2^2} \\
&= 0
\end{aligned}
$$

  Which yields:

$$
\begin{aligned}
0 &= u(t_x^2 - \|x\|_2^2) + 2t_x \\
t_x &= -\frac{1}{u} \pm \sqrt{\|x\|_2^2 + \frac{1}{u^2}} \\
t_x &= -\frac{1}{u} + \sqrt{\|x\|_2^2 + \frac{1}{u^2}} && \text{because of the condition } t_x > 0 \\
t_x &= -\frac{1}{u} \left(1 + \sqrt{1 + u^2 \|x\|_2^2} \right) && \text{because of the condition } u < 0
\end{aligned}
$$

By reinjecting we get:

$$\begin{aligned}
g(x, t_x) &= y^\intercal x + ut_x + \log(t_x^2 - \|x\|_2^2) \\
&= y^\intercal x - \left(1 + \sqrt{1 + u^2\|x\|_2^2}\right) + \log\left(\frac{\left(1 + \sqrt{1 + u^2\|x\|_2^2}\right)^2}{u^2} - \|x\|_2^2\right) \\
&= y^\intercal x - 1 - \sqrt{1 + u^2\|x\|_2^2} + \log\left(\frac{2 + 2\sqrt{1 + u^2\|x\|_2^2}}{u^2}\right) \\
&\geq y^\intercal x - 2 + u\|x\|_2 + \log\left(\frac{2 + 2\sqrt{1 + u^2\|x\|_2^2}}{u^2}\right)
\end{aligned}$$

Now we observe that for every $\lambda > 0$:

$$\begin{aligned}
g(\lambda y, t_x) &\geq \lambda\|y\|_2^2 - 2 + u\lambda\|y\|_2 + \log\left(\frac{2 + 2\sqrt{1 + u^2\lambda^2\|y\|_2^2}}{u^2}\right) \\
&\geq \lambda(\|y\|_2 + u)\|y\|_2 - 2 + \log\left(\frac{2 + 2\sqrt{1 + u^2\lambda^2\|y\|_2^2}}{u^2}\right)
\end{aligned}$$

- If $\|y\|_2 \geq -u$ then $\lim_{\lambda \to +\infty} g(\lambda y, t_x) = +\infty$ and thus $f^*(y, u) = +\infty$.
- Otherwise $\|y\|_2 < -u$ and then we solve for z fixed such that the whole derivative is nul:

$$\begin{aligned}
\frac{\partial g}{\partial z}(z, t_z) &= 0 \\
&= y - \frac{2z}{t_z^2 - \|z\|_2^2} \\
&= y - \frac{2z}{\frac{1}{u^2}\left(1 + \sqrt{1 + u^2\|z\|_2^2}\right)^2 - \|z\|_2^2} \\
z &= \frac{y}{2}\left(\frac{1}{u^2}\left(1 + \sqrt{1 + u^2\|z\|_2^2}\right)^2 - \|z\|_2^2\right) \\
&= y\left(\frac{1 + \sqrt{1 + u^2\|z\|_2^2}}{u^2}\right)
\end{aligned}$$

We can deduce from that there there exists $\mu \in \mathbb{R}^*$ such that $z = \mu y$, then:

$$\begin{aligned}
\mu &= \frac{1 + \sqrt{1 + u^2\mu^2\|y\|_2^2}}{u^2} \\
u^2\mu - 1 &= \sqrt{1 + u^2\mu^2\|y\|_2^2} \\
u^4\mu^2 + 1 - 2u^2\mu &= 1 + u^2\mu^2\|y\|_2^2 \\
0 &= \mu\left(\left(u^2 - \|y\|_2^2\right)\mu - 2\right) \\
\mu &= \frac{2}{u^2 - \|y\|_2^2}
\end{aligned}$$

Therefore we have:

$$\begin{aligned}
x &= \frac{2y}{u^2 - \|y\|_2^2} \\
t &= -\frac{2u}{u^2 - \|y\|_2^2} \\
g(x, t) &= -2 + \log 4 - \log\left(u^2 - \|y\|_2^2\right)
\end{aligned}$$

To summarize:

$$f^*(y) = \begin{cases} -2 + \log 4 - \log\left(u^2 - \|y\|_2^2\right) & \|y\|_2 < -u < 0 \\ +\infty & \text{otherwise} \end{cases}$$

Homework 2

YACHAOUI Ayman
Convex Optimization
ayman.yachaoui@telecom-paristech.fr

March 29, 2017

A Dual of general linear program

Linear programming:

$$\begin{aligned} \text{minimize} \quad & c^\intercal x \\ \text{subject to} \quad & Gx \preceq h \\ & Ax = b \end{aligned}$$

The dual function is:

$$\begin{aligned} g(\lambda, \nu) &= \inf_x (c^\intercal x + \lambda^\intercal (Gx - h) + \nu^\intercal (Ax - b)) \\ &= \inf_x (-\lambda^\intercal h - \nu^\intercal b + (c^\intercal + \lambda^\intercal G + \nu^\intercal A)x) \\ &= \begin{cases} -\lambda^\intercal h - \nu^\intercal b & \text{if } c^\intercal + \lambda^\intercal G + \nu^\intercal A = 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

Thus the dual problem is:

$$\begin{aligned} \text{maximize} \quad & g(\lambda, \nu) \\ \text{subject to} \quad & \lambda \succeq 0 \end{aligned}$$

And by making the implicit equality constraints explicit:

$$\begin{aligned} \text{maximize} \quad & -\lambda^\intercal h - \nu^\intercal b \\ \text{subject to} \quad & c^\intercal + \lambda^\intercal G + \nu^\intercal A = 0 \\ & \lambda \succeq 0 \end{aligned}$$

B Piecewise-linear minimization

Convex piecewise-linear minimization problem:

$$\text{minimize} \quad \max_{1 \le i \le m} (a_i^\intercal x + b_i)$$

Which is equivalent to:

$$\begin{aligned} \text{minimize} \quad & \max_{1 \le i \le m} y_i \\ \text{subject to} \quad & a_i^\intercal x + b_i = y_i, \quad 1 \le i \le m \end{aligned}$$

(a) The dual function is:

$$\begin{aligned} g(\lambda) &= \inf_{x,y} \left(\max_{1 \le i \le m} y_i + \sum_{i=1}^m \lambda_i (a_i^\intercal x + b_i - y_i) \right) \\ &= \inf_{x,y} \left(\sum_{i=1}^m \lambda_i b_i + \max_{1 \le i \le m} y_i - \sum_{i=1}^m \lambda_i y_i + \left(\sum_{i=1}^m \lambda_i a_i^\intercal \right) x \right) \end{aligned}$$

- If $\sum_{i=1}^m \lambda_i a_i \neq 0$ then $g(\lambda) = -\infty$
- Otherwise $\sum_{i=1}^m \lambda_i a_i = 0$ and:

$$g(\lambda) = \inf_y \left(\sum_{i=1}^m \lambda_i b_i + \max_{1 \leq i \leq m} y_i - \sum_{i=1}^m \lambda_i y_i \right)$$

  We now chose the vector with each coordinate equal to α, then by reinjecting into the inf:

$$\begin{aligned} g(\lambda) &\leq \sum_{i=1}^m \lambda_i b_i + \alpha - \sum_{i=1}^m \lambda_i \alpha \\ &\leq \sum_{i=1}^m \lambda_i b_i + \alpha \left(1 - \sum_{i=1}^m \lambda_i \right) \end{aligned}$$

  Thus this is unbounded below and:
  - If $\sum_{i=1}^m \lambda_i \neq 1$ then $g(\lambda) = -\infty$
  - Otherwise $\sum_{i=1}^m \lambda_i a_i = 0$ and $\sum_{i=1}^m \lambda_i = 1$. Finally:
    * If $\exists k, \lambda_i < 0$ then we chose the vector with all coordinates equal to zero except the k-th coordinate equal to $\alpha \in \mathbb{R}_-^*$, thus:

$$g(\lambda) \leq \sum_{i=1}^m \lambda_i b_i - \lambda_k \alpha$$

      As $\lambda_k < 0$ and $\alpha < 0$ then $-\lambda_k \alpha < 0$ and the expression is unbounded below therefore $g(\lambda) = -\infty$
    * Otherwise $\sum_{i=1}^m \lambda_i a_i = 0$, $\sum_{i=1}^m \lambda_i = 1$ and $\lambda \succeq 0$ and:

$$\begin{aligned} y_i &\leq \max_{1 \leq i \leq m} y_i \\ \lambda_i y_i &\leq \lambda_i \max_{1 \leq i \leq m} y_i \\ \sum_{i=1}^m \lambda_i y_i &\leq \sum_{i=1}^m \lambda_i \max_{1 \leq i \leq m} y_i \\ 0 &\leq \max_{1 \leq i \leq m} y_i - \sum_{i=1}^m \lambda_i y_i \\ g(\lambda) &\geq \sum_{i=1}^m \lambda_i b_i \end{aligned}$$

      Besides the equality holds for $y = 0$.

To summarize:

$$g(\lambda) = \begin{cases} \sum_{i=1}^m \lambda_i b_i & \text{if } \sum_{i=1}^m \lambda_i a_i = 0,\ \sum_{i=1}^m \lambda_i = 1 \text{ and } \lambda \succeq 0 \\ -\infty & \text{otherwise} \end{cases}$$

Therefore the dual problem is:

$$\begin{aligned} \text{maximize} \quad & g(\lambda) \\ \text{subject to} \quad & \lambda \succeq 0 \end{aligned}$$

And by making the implicit equality constraints explicit:

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^m \lambda_i b_i \\ \text{subject to} \quad & \sum_{i=1}^m \lambda_i a_i = 0 \\ & \sum_{i=1}^m \lambda_i = 1 \\ & \lambda \succeq 0 \end{aligned}$$

(b) The piecewise-linear minimization problem is similar to the linear program:

$$\begin{array}{ll} \text{minimize} & y \\ \text{subject to} & a_i^\intercal x + b_i - y \leq 0, \quad 1 \leq i \leq m \end{array}$$

The dual function is:

$$\begin{aligned} g(\lambda) &= \inf_{x,y}(y + \sum_{i=1}^{m} \lambda_i(a_i^\intercal x + b_i - y)) \\ &= \inf_{x,y}\left(\sum_{i=1}^{m} \lambda_i b_i + \left(\sum_{i=1}^{m} \lambda_i a_i^\intercal\right) x + \left(1 - \sum_{i=1}^{m} \lambda_i\right) y\right) \\ &= \begin{cases} \sum_{i=1}^{m} \lambda_i b_i & \text{if } \sum_{i=1}^{m} \lambda_i = 1 \text{ and } \sum_{i=1}^{m} \lambda_i a_i = 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

Thus the dual problem is:

$$\begin{array}{ll} \text{maximize} & \sum_{i=1}^{m} \lambda_i b_i \\ \text{subject to} & \sum_{i=1}^{m} \lambda_i a_i = 0 \\ & \sum_{i=1}^{m} \lambda_i = 1 \\ & \lambda \succeq 0 \end{array}$$

We observe that the two dual problem from (a) and (b) are identical.

(c) The dual problem (5.62) is:

$$\begin{array}{ll} \text{maximize} & b^\intercal \lambda - \sum_{i=1}^{m} \lambda_i \log \lambda_i \\ \text{subject to} & A^\intercal \lambda = 0 \\ & \mathbf{1}^\intercal \lambda = 1 \\ & \lambda \succeq 0 \end{array}$$

Both problems are convex and feasible, thus we have for all x:

$$p_{gp}^* \leq f_0(x) = \log\left(\sum_{i=1}^{m} \exp\left(a_i^\intercal x + b_i\right)\right) \leq \log\left(\sum_{i=1}^{m} \exp p_{pwl}^*\right) = \log m + p_{pwl}^*$$

$$p_{gp}^* - p_{pwl}^* \leq \log m$$

Also:

$$\max_{1 \leq i \leq m}\left(a_i^\intercal x + b_i\right) = \log\left(\exp\left(\max_{1 \leq i \leq m}\left(a_i^\intercal x + b_i\right)\right)\right) \leq \log\left(\sum_{i=1}^{m} \exp\left(a_i^\intercal x + b_i\right)\right) = f_0(x)$$

$$p_{pwl}^* \leq p_{gp}^*$$

Therefore:

$$0 \leq p_{gp}^* - p_{pwl}^* \leq \log m$$

(d) Following the same reasoning and noting p_γ^* the optimal value of the new problem we have for all x:

$$p_\gamma^* \leq \frac{1}{\gamma} \log\left(\sum_{i=1}^{m} \exp\left(\gamma\left(a_i^\intercal x + b_i\right)\right)\right) \leq \frac{1}{\gamma} \log\left(\sum_{i=1}^{m} \exp\left(\gamma p_{pwl}^*\right)\right) = \frac{1}{\gamma} \log m + p_{pwl}^*$$

$$p_\gamma^* - p_{pwl}^* \leq \frac{1}{\gamma} \log m$$

Also:

$$\max_{1\le i\le m}\left(a_i^\intercal x+b_i\right)=\frac{1}{\gamma}\log\left(\exp\left(\gamma\max_{1\le i\le m}\left(a_i^\intercal x+b_i\right)\right)\right)\le\frac{1}{\gamma}\log\left(\sum_{i=1}^m\exp\left(\gamma\left(a_i^\intercal x+b_i\right)\right)\right)$$

$$p^*_{pwl}\le p^*_\gamma$$

Therefore:

$$0\le p^*_\gamma-p^*_{pwl}\le\frac{1}{\gamma}\log m$$

As a consequence $\lim_{\gamma\to 0}p^*_\gamma=p^*_{pwl}$

C Suboptimality of a simple covering ellipsoid

The problem is:

$$\begin{array}{rl}\text{minimize} & f_0(X)=\log\det X^{-1}\\ \text{subject to} & a_i^\intercal X a_i\le 1,\quad 1\le i\le m\end{array}$$

With $\mathbf{dom} f_0=\mathbf{S}^n_{++}$ and $(a_i)_{1\le i\le m}$ spans $\mathbb{R}^n$.

(a) First we note that for all k and i:

$$(a_k a_k^\intercal)^\intercal=a_k a_k^\intercal$$
$$a_i^\intercal\left(a_k a_k^\intercal\right)a_i=\left(a_k^\intercal a_i\right)^\intercal\left(a_k^\intercal a_i\right)\ge 0$$

Thus $a_k a_k^\intercal\in\mathbf{S}^n_+$ and $\sum_{k=1}^m a_k a_k^\intercal=X_{sim}^{-1}\in\mathbf{S}^n_+$ and as it is inversible we can add $X_{sim}^{-1}\in\mathbf{S}^n_{++}$.
Finally for all $z\neq\mathbf{0}$ we have: $0<\left(X_{sim}z\right)^\intercal X_{sim}^{-1}\left(X_{sim}z\right)=z^\intercal X_{sim}z$ which yields $X_{sim}\in\mathbf{S}^n_{++}$.
We now need to show that $a_i^\intercal X_{sim}a_i\le 1$ for all $1\le i\le m$. For that we use Schur complement with:

$$X=\left[\begin{array}{cc}A & B\\ B^\intercal & C\end{array}\right]$$
$$A=X_{sim}^{-1}$$
$$B=a_i$$
$$C=1$$
$$S=C-B^\intercal A^{-1}B=1-a_i^\intercal X_{sim}a_i$$

Then $A\in\mathbf{S}^n_{++}$ thus:

$$\det X=\det A\det S$$
$$\det S=1-a_i^\intercal X_{sim}a_i=\det X\det X_{sim}$$

We already know that $\det X_{sim}>0$ thus we focus on $\det X$. Let $(u,v)\in\mathbb{R}^n\times\mathbb{R}$ then:

$$\begin{aligned}\left(\begin{array}{c}u\\ v\end{array}\right)^\intercal X\left(\begin{array}{c}u\\ v\end{array}\right)&=u^\intercal\left(\sum_{k=1}^m a_k a_k^\intercal\right)u+v^2+2va_i^\intercal u\\ &=u^\intercal\left(\sum_{\substack{k=1\\k\neq i}}^m a_k a_k^\intercal\right)u+u^\intercal a_i a_i^\intercal u+v^2+2va_i^\intercal u\\ &=u^\intercal\left(\sum_{\substack{k=1\\k\neq i}}^m a_k a_k^\intercal\right)u+\left(a_i^\intercal u+v\right)^2\\ &\ge 0\end{aligned}$$

As a consequence $X \in \mathbf{S}_+^{n+1}$ and $\det X \geq 0$, therefore:

$$a_i^\intercal X_{sim} a_i \leq 1$$

Which ends the proof that X_{sim} is feasible.

(b) The suboptimal dual problem is:

$$\begin{array}{ll} \text{maximize} & \log\det X_{sim}^{-1} + n\log t - mt + n \\ \text{subject to} & t > 0 \end{array}$$

An analysis of the function $h : t \mapsto n\log t - mt$ reveals that the maximum is obtained for $t_{\max} = \frac{n}{m}$ and $h(t_{\max}) = n\log\frac{n}{m} - n$.
As the problem is convex and feasible we have strong duality thus:

$$f_0(X_{sim}) - p^* \leq f_0(X_{sim}) - g(t_{\max}\mathbf{1}) = n\log\frac{m}{n}$$

Finally the volume of the ellipsoid is proportional to $\left(\det X^{-1}\right)^{\frac{1}{2}}$. If we note X^* the optimal solution of the primal problem we have with the bound above:

$$\begin{aligned} \log\det X_{sim}^{-1} - \log\det X^{*-1} &\leq n\log\frac{m}{n} \\ \frac{\det X_{sim}^{-1}}{\det X^{*-1}} &\leq \left(\frac{m}{n}\right)^n \\ \text{vol}(X_{sim}) &\leq \text{vol}(X^*)\left(\frac{m}{n}\right)^{\frac{n}{2}} \end{aligned}$$

Which ends the proof.

D Deriving a dual problem

The primal problem is:

$$\text{minimize} \quad \sum_{i=1}^N \|A_i x + b_i\|_2 + \frac{1}{2}\|x - x_0\|_2^2$$

With $A_i \in \mathbb{R}^{m_i \times n}$, $b_i \in \mathbb{R}^{m_i}$ and $x_0 \in \mathbb{R}^n$. We start by reformulating the problem:

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{i=1}^N \|y_i\|_2 + \frac{1}{2}\|x - x_0\|_2^2 \\ \text{subject to} & y_i = A_i x + b_i, \quad 1 \leq i \leq N \end{array}$$

Thus the dual function is:

$$\begin{aligned} g(\lambda) &= \inf_{x,(y_i)_{1\leq i\leq N}} \left(\sum_{i=1}^N \|y_i\|_2 + \frac{1}{2}\|x - x_0\|_2^2 + \sum_{i=1}^N \lambda_i^\intercal\left(y_i - A_i x - b_i\right)\right) \\ &= \inf_{x,(y_i)_{1\leq i\leq N}} \left(-\sum_{i=1}^N (\lambda_i^\intercal b_i) + \sum_{i=1}^N (\|y_i\|_2 + \lambda_i^\intercal y_i) + \frac{1}{2}\|x - x_0\|_2^2 - \left(\sum_{i=1}^N \lambda_i^\intercal A_i\right) x\right) \end{aligned}$$

Let $(i, \gamma) \in [\![1, N]\!] \times \mathbb{R}$, then by choosing $y_i = -\gamma\lambda_i$ we have:

$$\begin{aligned} \|y_i\|_2 + \lambda_i^\intercal y_i &= \gamma\|\lambda_i\|_2 - \gamma\|\lambda_i\|_2^2 \\ &= \gamma\|\lambda_i\|_2\left(1 - \|\lambda_i\|_2\right) \end{aligned}$$

If $\|\lambda_i\|_2 > 1$ then $\lim_{\gamma\to+\infty} \gamma\|\lambda_i\|_2\left(1 - \|\lambda_i\|_2\right) = -\infty$. Otherwise for all i, $\|\lambda_i\|_2 \leq 1$ and we have:

$$\|y_i\|_2 + \lambda_i^\intercal y_i \geq \|y_i\|_2 - \|\lambda_i\|_2\|y_i\|_2 = \left(1 - \|\lambda_i\|_2\right)\|y_i\|_2 \geq 0$$

And the equality holds for $y_i = 0$. We now focus on the minimization on x. We note $x_1 = \left(\sum_{i=1}^{N} \lambda_i^\intercal A_i\right)^\intercal$, then:

$$\begin{aligned}
\frac{1}{2}\|x - x_0\|_2^2 - \left(\sum_{i=1}^{N} \lambda_i^\intercal A_i\right) x &= \frac{1}{2}\|x - x_0\|_2^2 - x_1^\intercal x \\
&= \frac{1}{2}x^\intercal x + \frac{1}{2}x_0^\intercal x_0 - x_0^\intercal x - x_1^\intercal x \\
&= \frac{1}{2}\|x - x_0 - x_1\|_2^2 - \frac{1}{2}x_1^\intercal x_1 - x_0^\intercal x_1 \\
&= \frac{1}{2}\|x - x_0 - x_1\|_2^2 - \frac{1}{2}\|x_1 + x_0\|_2^2 + \frac{1}{2}\|x_0\|_2^2
\end{aligned}$$

Thus the minimum is attained for $x = x_0 + x_1$. Finally we have:

$$g\left((\lambda_i)_{1\le i\le N}\right) = \begin{cases} \frac{1}{2}\|x_0\|_2^2 - \sum_{i=1}^{N} b_i^\intercal \lambda_i - \frac{1}{2}\left\|x_0 + \sum_{i=1}^{N} A_i^\intercal \lambda_i\right\|_2^2 & \text{if } \|\lambda_i\|_2 \le 1, \quad \forall i \in [\![1, N]\!] \\ -\infty & \text{otherwise} \end{cases}$$

Thus the dual problem is:

$$\begin{aligned}
\text{maximize} \quad & g\left((\lambda_i)_{1\le i\le N}\right) \\
\text{subject to} \quad & \lambda_i \succeq 0 \quad \forall i \in [\![1, N]\!]
\end{aligned}$$

And by making the implicit equality constraints explicit:

$$\begin{aligned}
\text{maximize} \quad & \frac{1}{2}\|x_0\|_2^2 - \sum_{i=1}^{N} b_i^\intercal \lambda_i - \frac{1}{2}\left\|x_0 + \sum_{i=1}^{N} A_i^\intercal \lambda_i\right\|_2^2 \\
\text{subject to} \quad & \|\lambda_i\|_2 \le 1, \qquad \forall i \in [\![1, N]\!] \\
& \lambda_i \succeq 0, \qquad \forall i \in [\![1, N]\!]
\end{aligned}$$

E Lagrangian relaxation of Boolean LP

(a) The Boolean LP problem can be reformulated as:

$$\begin{aligned}
\text{minimize} \quad & c^\intercal x \\
\text{subject to} \quad & Ax \preceq b \\
& \underbrace{(\mathbf{1} - x) \circ x}_{\text{Hadamard product}} = 0
\end{aligned}$$

The dual function is:

$$\begin{aligned}
g(\lambda, \nu) &= \inf_x \left(c^\intercal x + \lambda^\intercal (Ax - b) + \nu^\intercal ((\mathbf{1} - x) \circ x)\right) \\
&= \inf_x \left(-\lambda^\intercal b + (c^\intercal + \lambda^\intercal A + \nu^\intercal)x - \nu^\intercal (x \circ x)\right)
\end{aligned}$$

- If $\nu \not\preceq 0$ then there exists a k such that $\nu_k > 0$ and by choosing the vector $x_\gamma$with all coordinates nul except for the k-th equal to γ then $\lim_{\gamma\to+\infty} \left(-\lambda^\intercal b + (c^\intercal + \lambda^\intercal A + \nu^\intercal)x_\gamma - \nu^\intercal x_\gamma \circ x_\gamma\right) = -\infty$ and g is unbounded below.
- Otherwise $\nu \preceq 0$ and we have:

$$(c^\intercal + \lambda^\intercal A + \nu^\intercal)x - \nu^\intercal (x \circ x) = \sum_{i=1}^{n} \left(\left(c_i + \nu_i + \sum_{k=1}^{m} \lambda_k A_{ki}\right) x_i - \nu_i x_i^2\right)$$

  - First we observe that if $\nu_i = 0$ then necessarily $c_i + \sum_{k=1}^{m} \lambda_k A_{ki} = 0$ or g is unbounded below.

– Otherwise minimizing $h_i : t \mapsto \left(c_i + \nu_i + \sum_{k=1}^m \lambda_k A_{ki}\right) t - \nu_i t^2$for each i yields:

$$\lim_{t\to-\infty} h(t) = \lim_{t\to+\infty} h(t) = +\infty$$
$$t_{\min} = \frac{c_i + \nu_i + \sum_{k=1}^m \lambda_k A_{ki}}{2\nu_i}$$
$$h(t_{\min}) = \frac{\left(c_i + \nu_i + \sum_{k=1}^m \lambda_k A_{ki}\right)^2}{4\nu_i}$$

Finally:

$$g(\lambda, \nu) = \begin{cases} -\lambda^\intercal b + \sum_{\substack{i=1\\ \nu_i\neq 0}}^n \frac{\left(c_i+\nu_i+\sum_{k=1}^m \lambda_k A_{ki}\right)^2}{4\nu_i} & \text{if } \nu \preceq 0 \\ -\infty & \text{otherwise} \end{cases}$$

Thus the dual problem is:

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

And by making the implicit equality constraints explicit:

$$\begin{array}{ll} \text{maximize} & -\lambda^\intercal b + \displaystyle\sum_{\substack{i=1\\ \nu_i\neq 0}}^n \frac{\left(c_i + \nu_i + \sum_{k=1}^m \lambda_k A_{ki}\right)^2}{4\nu_i} \\ \text{subject to} & \nu \preceq 0 \\ & \lambda \succeq 0 \end{array}$$

By noting $\Delta_i : t < 0 \mapsto \frac{(t+\beta_i)^2}{t}$ with $\beta_i = c_i + \sum_{k=1}^m \lambda_k A_{ki} > 0$ then:

$$\lim_{\nu_i\to 0} \Delta_i(t) = \lim_{\nu_i\to-\infty} \Delta_i(t) = -\infty$$
$$t_{\max} = -|\beta_i|$$
$$\Delta_i(t_{\max}) = \begin{cases} 0 & \beta_i > 0 \\ 4\beta_i & \beta_i < 0 \end{cases}$$

Finally by observing that $\Delta_i(t_{\max}) = \min(0, 4\beta_i)$ we can write the dual problem as:

$$\begin{array}{ll} \text{maximize} & -\lambda^\intercal b + \displaystyle\sum_{\substack{i=1\\ \nu_i\neq 0}}^n \min\left(0, c_i + \sum_{k=1}^m \lambda_k A_{ki}\right) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

(b) The LP relaxation problem is:

$$\begin{array}{ll} \text{minimize} & c^\intercal x \\ \text{subject to} & Ax \preceq b \\ & 0 \leq x_i \leq 1, \quad 1 \leq i \leq n \end{array}$$

The dual function is:

$$\begin{aligned} g(\lambda_0, \lambda_1, \lambda_2) &= \inf_x (c^\intercal x + \lambda_0^\intercal (Ax - b) - \lambda_1^\intercal x + \lambda_2^\intercal (x - \mathbf{1})) \\ &= \inf_x (-\lambda_0^\intercal b - \lambda_2^\intercal \mathbf{1} + (c^\intercal + \lambda_0^\intercal A - \lambda_1^\intercal + \lambda_2^\intercal) x) \\ &= \begin{cases} -b^\intercal \lambda_0 - \mathbf{1}^\intercal \lambda_2 & \text{if } c + A^\intercal \lambda_0 - \lambda_1 + \lambda_2 = 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

Thus the dual problem is:

$$\begin{array}{ll} \text{maximize} & g(\lambda_0, \lambda_1, \lambda_2) \\ \text{subject to} & \lambda_0 \succeq 0 \\ & \lambda_1 \succeq 0 \\ & \lambda_2 \succeq 0 \end{array}$$

And by making the implicit equality constraints explicit:

$$\begin{array}{ll} \text{maximize} & -b^\intercal \lambda_0 - \mathbf{1}^\intercal \lambda_2 \\ \text{subject to} & c + A^\intercal \lambda_0 - \lambda_1 + \lambda_2 = 0 \\ & \lambda_0 \succeq 0 \\ & \lambda_1 \succeq 0 \\ & \lambda_2 \succeq 0 \end{array}$$

We observe that $-\lambda_2 = c + A^\intercal \lambda_0 - \lambda_1 \preceq 0$.
Thus we can simplify the maximization by writing $-\lambda_2 = \min(0, c + A^\intercal \lambda_0)$ and finally the dual problem with explicit constraints becomes:

$$\begin{array}{ll} \text{maximize} & -b^\intercal \lambda_0 + \mathbf{1}^\intercal \min(0, c + A^\intercal \lambda_0) \\ \text{subject to} & \lambda_0 \succeq 0 \end{array}$$

Which is the same dual problem as for the Lagrangian relaxation of Boolean LP.