Convex Optimization and applications in Machine Learning **MVA Fall 2019**

Homework 1 — October 14

Instructor: Alexandre d'Aspremont *Student: Antoine Moulin*

The following exercises can be found in the book: **Convex Optimization** by Lieven Vandenberghe and Stephen Boyd.

Exercise 2.12

Which of the following sets are convex?

(a) A ***slab***, *i.e.*, a set of the form $\left\{x \in \mathbb{R}^n \mid \alpha \leq a^T x \leq \beta\right\}$.

(b) A ***rectangle***, *i.e.*, a set of the form $\{x \in \mathbb{R}^n \mid \forall i \in [\![1, n]\!], \alpha_i \leq x_i \leq \beta_i\}$. A rectangle is sometimes called a ***hyperrectangle*** when $n > 2$.

(c) A ***wedge***, *i.e.*, $\left\{x \in \mathbb{R}^n \mid a_1^T x \leq b_1, a_2^T x \leq b_2\right\}$.

(d) The set of points closer to a given point than a given set, *i.e.*,

$$\{x \in \mathbb{R}^n \mid \forall y \in S, ||x - x_0||_2 \leq ||x - y||_2\}$$

where $S \subseteq \mathbb{R}^n$.

(e) The set of points closer to one set than another, *i.e.*,

$$\{x \in \mathbb{R}^n \mid \mathrm{dist}(x, S) \leq \mathrm{dist}(x, T)\}$$

where $S, T \subseteq \mathbb{R}^n$, and

$$\mathrm{dist}(x, S) = \inf\{||x - z||_2 \mid z \in S\}$$

(f) [HUL93, volume 1, page 93] The set $\{x \in \mathbb{R}^n \mid x + S_2 \subseteq S_1\}$, where $S_1, S_2 \subseteq \mathbb{R}^n$ with S_1 convex.

(g) The set of points whose distance to a does not exceed a fixed fraction θ of the distance to b, *i.e.*, the set $\{x \in \mathbb{R}^n \mid ||x - a||_2 \leq \theta\, ||x - b||_2\}$. You can assume $a \neq b$ and $0 \leq \theta \leq 1$.

Solution.

(a) According to the definition, a slab is an intersection of two halfspaces (convex sets). An intersection of convex sets is convex, hence a slab is convex.

(b) A rectangle in dimension n is the intersection of n slabs. For $i \in [\![1, n]\!]$, let denote by e_i the vector whose i-th component is 1 and the others are 0. We have:

$$\{x \in \mathbb{R}^n \mid \forall i \in [\![1, n]\!], \alpha_i \leq x_i \leq \beta_i\} = \bigcap_{i=1}^{n} \left\{x \in \mathbb{R}^n \mid \alpha_i \leq e_i^T x \leq \beta_i\right\}$$

A slab is convex (see (a)) and an intersection of convex sets is convex. Thus, a rectangle is convex.

(c) As in (a), a wedge is the intersection of two halfspaces, thus it is convex.

(d) Let $y \in S$. Let show that $C(y) = \{x \in \mathbb{R}^n \mid ||x - x_0||_2 \leq ||x - y||_2\}$ is a halfspace, hence a convex set. With $n = 2$, we clearly see that the split is made by the perpendicular bisector of the segment $[x_0, y]$. Let $x \in C(y)$.

$$\begin{aligned} ||x - x_0||_2 \leq ||x - y||_2 &\text{ iff } ||x - x_0||_2^2 \leq ||x - y||_2^2 \\ &\text{ iff } (x - x_0)^T (x - x_0) \leq (x - y)^T (x - y) \\ &\text{ iff } -2x_0^T x + ||x_0||_2^2 \leq -2y^T x + ||y||_2^2 \\ &\text{ iff } (y - x_0)^T x \leq \frac{||y||_2^2 - ||x_0||_2^2}{2} \end{aligned}$$

Defining $a = y - x_0$ and $b = \frac{||y||_2^2 - ||x_0||_2^2}{2}$, we see that $C(y)$ is a halfspace, hence convex.

We have $\{x \in \mathbb{R}^n \mid \forall y \in S, ||x - x_0||_2 \leq ||x - y||_2\} = \bigcap_{y \in S} C(y)$. An intersection of convex sets is convex, hence the initial set is convex.

(e) In general, the set $A = \{x \in \mathbb{R}^n \mid \text{dist}(x, S) \leq \text{dist}(x, T)\}$ is not convex. Here is an example in which it is not convex:

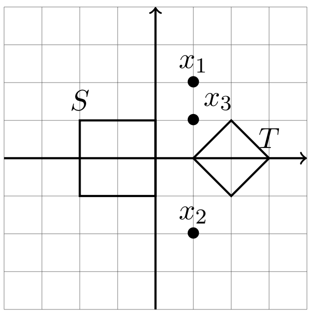

Figure 1: Example in which the set A is not convex.

Here, x_1 and x_2 are in the set A (because for $i \in \{1, 2\}$, $\text{dist}(x_i, S) = \text{dist}(x_i, T)$), but x_3 - which is in the segment $[x_1, x_2]$ - is not. Hence, the set A is not convex.

(f) Let $C = \{x \in \mathbb{R}^n \mid x + S_2 \subseteq S_1\}$, $u, v \in C, \theta \in [0, 1]$ and let show that $(\theta u + (1 - \theta)v) + S_2 \subseteq S_1$. Let $s \in S_2$. We have:

$$\theta u + (1 - \theta)v + s = \theta\underbrace{(u + s)}_{\in S_1} + (1 - \theta)\underbrace{(v + s)}_{\in S_1} \in S_1$$

$u + s \in S_1$ because $u \in C$, $v + s \in S_1$ because $v \in C$ and the convex combination is in S_1 as well because S_1 is convex. Hence, C is convex.

(g) Let $C = \{x \in \mathbb{R}^n \mid ||x - a||_2 \leq \theta \, ||x - b||_2\}$. Let $x \in \mathbb{R}^n$.

$$\begin{aligned} x \in C &\text{ iff } ||x - a||_2 \leq \theta \, ||x - b||_2 \\ &\text{ iff } (x - a)^T (x - a) \leq \theta^2 (x - b)^T (x - b) \\ &\text{ iff } ||x||_2^2 - 2a^T x + ||a||_2^2 \leq \theta^2 \, ||x||_2^2 - 2\theta^2 b^T x + \theta^2 \, ||b||_2^2 \\ &\text{ iff } f(x) \triangleq (1 - \theta^2) \, ||x||_2^2 + 2(\theta^2 b - a)^T x \leq \theta^2 \, ||b||_2^2 - ||a||_2^2 \end{aligned}$$

The function f is convex and C is a sublevel set of f. Hence, C is convex.

Exercise 3.21

Pointwise maximum and supremum. Show that the following functions $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ are convex.

(a) $f(x) = \max_{i \in [\![1,k]\!]} \left|\left|A^{(i)}x - b^{(i)}\right|\right|$, where $A^{(i)} \in \mathbb{R}^{m \times n}, b^{(i)} \in \mathbb{R}^m$ and $||\cdot||$ is a norm on $\mathbb{R}^n$.

(b) $f(x) = \sum_{i=1}^r |x|_{[i]}$ on $\mathbb{R}^n$, where $|x|$ denotes the vector with $|x|_i = |x_i|$ (*i.e.* $|x|$ is the absolute value of x, componentwise), and $|x|_{[i]}$ is the i-th largest component of $|x|$. In other words, $|x|_{[1]}, |x|_{[2]}, \ldots, |x|_{[n]}$ are the absolute values of the components of x, sorted in nonincreasing order.

Solution.

(a) For $i \in [\![1,k]\!]$, the function $x \mapsto \left|\left|A^{(i)}x - b^{(i)}\right|\right|$ is convex as the composition of an affine function and a norm (which is a convex function). The function f is the pointwise maximum of k convex functions, thus f is convex.

(b) It is possible to rewrite the function f as following:

$$\forall x \in \mathbb{R}^n, f(x) = \sum_{i=1}^r |x|_{[i]} = \max_{1 \leq i_1 < \cdots < i_r \leq n} \sum_{k=1}^r |x|_{i_k}$$

f is the pointwise maximum of $\binom{n}{r}$ convex functions, hence f is convex.

Exercise 3.32

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.

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

(c) If f is convex, nondecreasing, and positive, and g is concave, nonincreasing, and positive, then f/g is convex.

Solution.

(a) f and g are defined on an interval I. Let $x, y \in I, \theta \in [0,1]$. We have:

$$fg(\theta x + (1-\theta)y) \leq [\theta f(x) + (1-\theta)f(y)]\,[\theta g(x) + (1-\theta)g(y)] \quad \text{by convexity of } f \text{ and } g$$

Besides, if we note $A = [\theta f(x) + (1-\theta)f(y)]\,[\theta g(x) + (1-\theta)g(y)]$, we have:

$$\begin{aligned} A = & \underbrace{\theta^2}_{=\theta(\theta+1-1)} f(x)g(x) + \underbrace{(1-\theta)^2}_{=(1-\theta)-\theta(1-\theta)} f(y)g(y) + \theta(1-\theta)f(x)g(y) + \theta(1-\theta)f(y)g(x) \\ = & \; \theta f(x)g(x) + (1-\theta)f(y)g(y) \\ & + \underbrace{\theta(\theta-1)f(x)g(x) + \theta(1-\theta)f(x)g(y)}_{=\theta(\theta-1)f(x)[g(x)-g(y)]} + \underbrace{\theta(1-\theta)f(y)g(x) - \theta(1-\theta)f(y)g(y)}_{=\theta(1-\theta)f(y)[g(x)-g(y)]} \\ = & \; \theta f(x)g(x) + (1-\theta)f(y)g(y) + \theta(\theta-1)\,[f(x)-f(y)]\,[g(x)-g(y)] \end{aligned}$$

As $\theta \in [0,1]$, we have $\theta(\theta-1) \leq 0$. Besides, because f and g are both nondecreasing or nonincreasing, $[f(x)-f(y)]\,[g(x)-g(y)] \geq 0$. Thus, $\theta(\theta-1)\,[f(x)-f(y)]\,[g(x)-g(y)] \leq 0$. Finally,

$$fg(\theta x + (1-\theta)y) \leq \theta f(x)g(x) + (1-\theta)f(y)g(y)$$

and fg is convex.

(b) As in (a), by convexity of f and g, we can write:

$$fg(\theta x + (1-\theta)y) \geq \theta f(x)g(x) + (1-\theta)f(y)g(y) + \theta(\theta-1)\left[f(x)-f(y)\right]\left[g(x)-g(y)\right]$$

As $\theta \in [0,1]$, we have $\theta(\theta-1) \leq 0$. Besides, because one of the function is nondecreasing and the other is nonincreasing, $\left[f(x)-f(y)\right]\left[g(x)-g(y)\right] \leq 0$. Thus, $\theta(\theta-1)\left[f(x)-f(y)\right]\left[g(x)-g(y)\right] \geq 0$. Finally,

$$fg(\theta x + (1-\theta)y) \geq \theta f(x)g(x) + (1-\theta)f(y)g(y)$$

Thus, fg is concave.

(c) We know that $1/g$ is convex, nondecreasing and positive. Thus, thanks to (a), f/g is convex.

Exercise 3.36

Derive the conjugates of the following functions.

(a) *Max function.* $f(x) = \max_{i\in[\![1,n]\!]} x_i$ on $\mathbb{R}^n$.

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

(c) *Piecewise-linear function* on $\mathbb{R}$. $f(x) = \max_{i\in[\![1,m]\!]}(a_i x + b_i)$ on $\mathbb{R}$. You can assume that the a_i are sorted in increasing order, *i.e.*, $a_1 \leq \cdots \leq a_m$, 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 $f(x) = a_k x + b_k$.

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

(e) *Negative geometric mean.* $f(x) = -\left(\prod_{i=1}^n x_i\right)^{1/n}$ on $\mathbb{R}_{++}$.

(f) *Negative generalized logarithm for second-order cone.*

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

Solution.

(a) First, we remark that for $x, y \in \mathbb{R}^n$, we have:

$$y^T x - f(x) = \sum_{i=1}^n y_i x_i - \max_{i\in[\![1,n]\!]} x_i \leq \max_{i\in[\![1,n]\!]} x_i \left(\sum_{i=1}^n y_i - 1\right)$$

Thus, in order to maximize $y^T x - f(x)$ with respect to the variable x, it seems relevant to take x a constant vector.

Let $y \in \mathbb{R}^n, \alpha \in \mathbb{R}$. We define $x = \alpha \cdot \mathbf{1}$. We have $y^T x - f(x) = \alpha \sum_{i=1}^n y_i - \alpha = \alpha\left(\sum_{i=1}^n y_i - 1\right)$. Let's first distinguish two cases:

- $\sum_{i=1}^n y_i > 1$. If $\alpha \longrightarrow +\infty$, then $y^T x - f(x) \longrightarrow +\infty$ and $f^*(y) = +\infty$.
- $\sum_{i=1}^n y_i < 1$. Similarly, if $\alpha \longrightarrow -\infty$, then $f^*(y) = +\infty$.

The remaining case is when $\sum_{i=1}^n y_i = 1$. Once again, we distinguish two cases:

- There exists a $j \in [\![1,n]\!]$ such that $y_j < 0$. We redefine the vector x: $x_j = \alpha$ and if $i \neq j, x_i = 0$. We have $y^T x - f(x) = \alpha y_j - \max(0,\alpha)$. If $\alpha \longrightarrow -\infty$, then $y^T x - f(x) \longrightarrow +\infty$ and $f^*(y) = +\infty$.
- For all $i \in [\![1,n]\!]$, $y_i \geq 0$. Then, $y^T x - f(x) \leq 0$, with equality if $x = 0$. Thus, $f^*(y) = 0$.

Finally, we have:

$$\boxed{f^*(y) = \begin{cases} 0 & \text{if } y \succeq \mathbf{0} \text{ and } y^T\mathbf{1} = 1 \\ +\infty & \text{otherwise} \end{cases}}$$

(b) Let $y \in \mathbb{R}^n$, $g_y : x \in \mathbb{R}^n \mapsto \sum_{i=1}^n y_i x_i - \sum_{i=1}^r x_{[i]}$. Because $[\cdot]$ is a permutation, we can write, for all x,

$$g_y(x) = \sum_{i=1}^r (y_{[i]} - 1)x_{[i]} + \sum_{i=r+1}^n y_{[i]}x_{[i]}$$

We distinguish two cases:

- $r < n$. Let suppose there exists a $j \in [\![1, n]\!]$ such that $y_j > 1$. Then, if we take x the vector whose j-th component is $\alpha > 0$ and the other components are 0, we have: $g_y(x) = y_j\alpha - \alpha = (y_j - 1)\,\alpha$. As $y_j > 1$, if α tends to $+\infty$, then $g_y(x)$ tends to $+\infty$ and $f^*(y) = +\infty$.
  Now, we suppose that there exists a $j \in [\![1, n]\!]$ such that $y_j < 0$. With the same vector x but with $\alpha < 0$, we have: $g_y(x) = y_j\alpha \underset{\alpha\to-\infty}{\longrightarrow} +\infty$ and $f^*(y) = +\infty$. Thus, we suppose $0 \preceq y \preceq 1$.
- $r = n$. In this case, for $x \in \mathbb{R}^n$, $g_y(x) = \sum_{i=1}^n (y_i - 1)\,x_i = (y - \mathbf{1})^T x$. If there exists a $j \in [\![1, n]\!]$ such that $y_j \neq 1$, then $f^*(y) = +\infty$. Thus,

$$\boxed{\text{If } r = n,\ f^*(y) = \begin{cases} 0 & \text{if } y = \mathbf{1} \\ +\infty & \text{otherwise} \end{cases}}$$

In the following, we suppose $r < n$.
We can write $g_y(x) = \sum_{i=1}^n (y_i - 1)\,x_i + \sum_{i=1}^n x_i - \sum_{i=1}^r x_{[i]} = \sum_{i=1}^n (y_i - 1)\,x_i + \sum_{i=r+1}^n x_{[i]}$. Let's take x a constant vector $\alpha \cdot \mathbf{1}$. Then, we have:

$$g_y(x) = \alpha \sum_{i=1}^n (y_i - 1) + (n - r)\,\alpha = \alpha\left(y^T\mathbf{1} - r\right)$$

If $y^T\mathbf{1} \neq r$, we can make α arbitrarily tend to $-\infty$ or $+\infty$ and show that $f^*(y) = +\infty$. Now, we suppose that $y^T\mathbf{1} = r$. We have:

$$\begin{aligned} g_y(x) &= \sum_{i=1}^r \underbrace{(y_{[i]} - 1)}_{\leq 0}\, \underbrace{x_{[i]}}_{\geq x_{[r]}} + \sum_{i=r+1}^n \underbrace{y_{[i]}}_{\geq 0}\, \underbrace{x_{[i]}}_{\leq x_{[r]}} && (\mathbf{0} \preceq y \preceq \mathbf{1} \text{ and definition of } [\cdot]) \\ &\leq \sum_{i=1}^r (y_{[i]} - 1)x_{[r]} + \sum_{i=r+1}^n y_{[i]}x_{[r]} \\ &= \left(\sum_{i=1}^n y_{[i]} - r\right) x_{[r]} \\ &= 0 && (y^T\mathbf{1} = r) \end{aligned}$$

Then, $g_y(x) \leq 0$. With $x = \alpha \cdot \mathbf{1}$, we have $g_y(x) = 0$. Thus, $f^*(y) = 0$. Finally,

$$\boxed{\text{If } r < n,\ f^*(y) = \begin{cases} 0 & \text{if } 0 \preceq y \preceq 1 \text{ and } y^T\mathbf{1} = r \\ +\infty & \text{otherwise} \end{cases}}$$

(c) Let $y \in \mathbb{R}$. As we consider the supremum of the function $g_y : x \mapsto yx - \max_{i\in[\![1,m]\!]}(a_i x + b_i)$ and that $a_1 \leq \cdots \leq a_m$, the points of interest are the intersection points between $x \mapsto a_i x + b_i$ and $x \mapsto a_{i+1}x + b_{i+1}$, for $i \in [\![1, m-1]\!]$. Those are given by:

$$\forall i \in [\![1, m-1]\!], x_i = \frac{b_i - b_{i+1}}{a_{i+1} - a_i}$$

We also define $x_0 = -\infty, x_m = +\infty$. Thus, we can rewrite, for $x \in \mathbb{R}$,

$$\begin{aligned} g_y(x) &= yx \cdot \mathbb{1}_{\mathbb{R}}(x) - \sum_{i=1}^{m} (a_i x + b_i)\, \mathbb{1}_{[x_{i-1}, x_i]}(x) \\ &= yx \sum_{i=1}^{m} \mathbb{1}_{[x_{i-1}, x_i]}(x) - \sum_{i=1}^{m} (a_i x + b_i)\, \mathbb{1}_{[x_{i-1}, x_i]}(x) \\ &= \sum_{i=1}^{m} [(y - a_i)\, x - b_i]\, \mathbb{1}_{[x_{i-1}, x_i]}(x) \end{aligned}$$

The function we want to maximize is a piecewise-linear function. From the previous equation and the fact that $a_1 \leq \cdots \leq a_m$, we see that:

- If $y < a_1$, then all the slopes are negative and the function is decreasing on $\mathbb{R}$. If x tends to $-\infty$, $g(x)$ tends to $+\infty$, *i.e.* $f^*(y) = +\infty$.
- If $y > a_m$, similarly, all the slopes are positive and the function is increasing. In $+\infty$, g tends to $+\infty$ and $f^*(y) = +\infty$.
- If $j \in [\![1, m-1]\!]$ and $y \in [a_j, a_{j+1}]$, then $y - a_1, \ldots, y - a_j \geq 0$ and $y - a_{j+1}, \ldots, y - a_m \leq 0$. It means that g is first increasing and then decreasing. Hence, it reaches a maximum at x_j, the point that connects the segments i and $i+1$. The value is:

$$\begin{aligned} f^*(y) &= yx_j - a_j x_j - b_j \\ &= y \frac{b_j - b_{j+1}}{a_{j+1} - a_j} - a_j \frac{b_j - b_{j+1}}{a_{j+1} - a_j} - b_j \\ &= (y - a_j) \frac{b_j - b_{j+1}}{a_{j+1} - a_j} - b_j \end{aligned}$$

Finally,

$$\boxed{f^*(y) = \begin{cases} +\infty & \text{if } y < a_1 \text{ or } y > a_m \\ (y - a_i)\, \frac{b_i - b_{i+1}}{a_{i+1} - a_i} - b_i & \text{if there exists } i \in [\![1, m-1]\!] \text{ s.t. } y \in [a_i, a_{i+1}] \end{cases}}$$

(d) Let $y \in \mathbb{R}, p > 1$. The function $g_y : x \mapsto yx - x^p$ is strictly concave on $\mathbb{R}_{++}$. For $y > 0$, the calculation of the derivative shows that it reaches its maximum at $x_{max} = \left(\frac{y}{p}\right)^{\frac{1}{p-1}}$ and:

$$\begin{aligned} g_y(x_{max}) &= y \left(\frac{y}{p}\right)^{\frac{1}{p-1}} - \left(\frac{y}{p}\right)^{\frac{p}{p-1}} \\ &= y \left(\frac{y}{p}\right)^{\frac{1}{p-1}} - \frac{y}{p} \left(\frac{y}{p}\right)^{\frac{1}{p-1}} \\ &= y \left(\frac{y}{p}\right)^{\frac{1}{p-1}} \left(\frac{p-1}{p}\right) \\ &= (p-1) \left(\frac{y}{p}\right)^{\frac{p}{p-1}} \end{aligned}$$

For $y \leq 0$, the function in nonincreasing and reaches its maximum at $x_{max} = 0$ with the value $g_y(x_{max}) = 0$. Thus,

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

Let $y \in \mathbb{R}, p < 0$. The function g_y is still strictly concave. For $y \geq 0$, the derivative of $g_y : x \mapsto yx - x^p$ is positive on $\mathbb{R}_{++}$, *i.e.*, g_y is increasing on $\mathbb{R}_{++}$. If $x \longrightarrow +\infty$, we have $g_y(x) \longrightarrow +\infty$, hence $f^*(y) = +\infty$.

For $y < 0$, the maximum is reached at $x_{max} = \left(\frac{y}{p}\right)^{\frac{1}{p-1}}$ and the value is $g_y(x_{max}) = (p-1)\left(\frac{y}{p}\right)^{\frac{p}{p-1}}$. Thus,

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

(e) Let $y \in \mathbb{R}^n$ and $g_y : x \in \mathbb{R}_{++} \mapsto \sum_{i=1}^n x_i y_i + \left(\prod_{i=1}^n x_i\right)^{1/n}$.
Suppose there exists $j \in [\![1, n]\!]$ such that $y_j > 0$. If we consider the vector x whose j-th component is $\alpha > 0$ and the others are 1, we have: $g_y(x) = \sum_{i \neq j} y_i + y_j \alpha + \alpha^{1/n} \underset{\alpha \to +\infty}{\longrightarrow} +\infty$. Hence, $f^*(y) = +\infty$.
Now, we suppose that $y \preceq 0$. We have:

$$\begin{aligned} g_y(x) = \left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}} - \sum_{i=1}^n (-y_i)\, x_i &\leq \left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}} - n\left(\prod_{i=1}^n (-y_i)\, x_i\right)^{\frac{1}{n}} && \text{(AM-GM inequality)} \\ &\leq \left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}} \left(1 - n\left(\prod_{i=1}^n (-y_i)\right)^{\frac{1}{n}}\right) \end{aligned}$$

Suppose that $1 - n\left(\prod_{i=1}^n (-y_i)\right)^{\frac{1}{n}} > 0$, *i.e.*, $\frac{1}{n} > \left(\prod_{i=1}^n (-y_i)\right)^{\frac{1}{n}}$. If we take $x_i = -\frac{\alpha}{y_i}$ with $\alpha > 0$, we have:

$$g_y(x) = \left(\prod_{i=1}^n \left(-\frac{1}{y_i}\right)\right)^{\frac{1}{n}} \alpha - n\alpha = \underbrace{\left[\frac{1}{\left(\prod_{i=1}^n (-y_i)\right)^{\frac{1}{n}}} - n\right]}_{>0} \alpha \underset{\alpha \to +\infty}{\longrightarrow} +\infty$$

And thus, $f^*(y) = +\infty$. Now, we suppose $\left(\prod_{i=1}^n (-y_i)\right)^{\frac{1}{n}} \geq \frac{1}{n}$. Because of the previous inequality and the new hypothesis, we have:

$$g_y(x) \leq \left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}} \left(1 - n\left(\prod_{i=1}^n (-y_i)\right)^{\frac{1}{n}}\right) \leq 0$$

We define the sequence of vectors $(x_k)_{k\in\mathbb{N}} = \left(\frac{1}{k} \cdot \mathbf{1}\right)_{k\in\mathbb{N}}$. We have $g_y(x_k) = \frac{1}{k}\left(\sum_{i=1}^n y_i + 1\right) \underset{k \to +\infty}{\longrightarrow} 0$. Because of the sequential characterization of the supremum, it shows that $f^*(y) = 0$.

Finally,

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

(f) We remark that because of the definition domain (second-order cone, denoted C_2 in the following), we need to have $t > 0$ (if $t \leq 0$, it would be impossible to find x such that $||x||_2 < t$). Let $y \in \mathbb{R}^n, s \in \mathbb{R}$ and

$$g_{y,s} : (x,t) \in C_2 \longmapsto y^T x + st + \log\left(t^2 - x^T x\right)$$

- *Domain of f^**. If $s \geq 0$, then taking $x = 0$ (for all $t > 0$, $(0,t) \in C_2$), we have:

$$g_{y,s}(x,t) = st + 2\log t \underset{t \to +\infty}{\longrightarrow} +\infty, \textit{ i.e.}\,,\ f^*(y) = +\infty$$

  Now, we suppose $s < 0$. Let assume that $||y||_2 \geq -s$.
  Let $\alpha > 1$, $x = \left(\alpha^2 - \frac{1}{\alpha}\right) y$ and $t = \alpha^2 \, ||y||_2$. We have:

$$t - ||x||_2 = \alpha^2 \, ||y||_2 - \left(\alpha^2 - \frac{1}{\alpha}\right) ||y||_2 = \frac{1}{\alpha} \, ||y||_2 > 0$$

  Then, $(x,t) \in C_2$. We compute the value of $g_{y,s}$ at (x,t):

$$\begin{aligned} g_{y,s}(x,t) &= ||y||_2^2 \left(\alpha^2 - \frac{1}{\alpha}\right) + s\,||y||_2\,\alpha^2 + \log\left[\alpha^4\,||y||_2^2 - \left(\alpha^2 - \frac{1}{\alpha}\right)^2 ||y||_2^2\right] \\ &= ||y||_2\,(||y||_2 + s)\,\alpha^2 - \frac{1}{\alpha}\,||y||_2^2 + \log\left[2\alpha - \frac{1}{\alpha^2}\right] + 2\log||y||_2 \end{aligned}$$

  If $||y||_2 > -s$, then $||y||_2 + s > 0$ and $(||y||_2 + s)\,\alpha^2 \underset{\alpha \to +\infty}{\longrightarrow} +\infty$. If $||y||_2 = -s$, then the quadratic term is null but $\log\left[2\alpha - \frac{1}{\alpha^2}\right] \underset{\alpha \to +\infty}{\longrightarrow} +\infty$. Thus,

$$g_{y,s}(x,t) \underset{\alpha \to +\infty}{\longrightarrow} +\infty$$

  and $f^*(y) = +\infty$. Now we suppose $||y||_2 < -s$. The domain of f^* is:

$$\mathrm{dom}\, f^* = \{(y,s) \in \mathbb{R}^n \times \mathbb{R} \mid ||y||_2 < -s\}$$

  ($s < 0$ is included in the constraint.)

- *Concavity of $g_{y,s}$ on* $\mathrm{dom}\, f^*$. Let $(y,s) \in \mathrm{dom}\, f^*$. For all $(x,t) \in C_2$,

$$g_{y,s}(x,t) = y^T x + st + 2\log t + \log\left(1 - \left(\frac{x}{t}\right)^T \left(\frac{x}{t}\right)\right)$$

  $(x,t) \in C_2 \mapsto y^T x + st$ is an affine function of (x,t), hence concave. $(x,t) \mapsto 2\log t$ is also concave. Let show that $\phi : (x,t) \in C_2 \mapsto \log\left(1 - \left(\frac{x}{t}\right)^T \left(\frac{x}{t}\right)\right)$ is concave. We can write $\phi = h \circ g$ where:

$$\begin{array}{rccl} h: & (0,1) & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \log(1-x) \end{array} \quad \text{and} \quad \begin{array}{rccl} g: & \mathbb{R}^n \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ & (x,t) & \longmapsto & \left(\frac{x}{t}\right)^T \left(\frac{x}{t}\right) \end{array}$$

  Lemma. Let $g : \mathbb{R}^n \to \mathbb{R}, h : \mathbb{R} \to \mathbb{R}$ and $f = h \circ g$. If g is convex, h is concave and nonincreasing, then f is concave.
  Proof. Let $x, y \in \mathbb{R}^n, \theta \in [0,1]$. By convexity of g, we have:

$$g\left(\theta x + (1-\theta)y\right) \leq \theta g(x) + (1-\theta) g(y)$$

  Because h is nonincreasing and concave, we have:

$$f\left(\theta x + (1-\theta)y\right) = h\left(g\left(\theta x + (1-\theta)y\right)\right) \underbrace{\geq}_{h \searrow} h\left(\theta g(x) + (1-\theta) g(y)\right) \underbrace{\geq}_{h \text{ concave}} \theta f(x) + (1-\theta) f(y)$$

  Thus, f is concave. $\square$

In this case, h is twice differentiable. For all x, $h'(x) = -\frac{1}{1-x} < 0$ and $h''(x) = -\frac{1}{(1-x)^2} < 0$, hence h is concave and nonincreasing. Besides, g is convex (as the composition of a convex function nondecreasing in each argument and a convex function). Hence, the lemma shows that ϕ is concave. $g_{y,s}$ is a sum of concave functions, thus it is a concave function. We can compute the gradient and find the maximum.

- *Maximum of $g_{y,s}$.* Let $(y, s) \in \operatorname{dom} f^*$, $(x, t) \in C_2$. The gradients with respect to x and t are given by:

$$\nabla_x g_{y,s}(x,t) = y - \frac{2}{t^2 - x^T x} x \text{ and } \nabla_t g_{y,s}(x,t) = s + \frac{2t}{t^2 - x^T x}$$

$$\begin{aligned}
\left\{\begin{array}{l} \nabla_x g_{y,s}(x,t) = 0 \\ \nabla_t g_{y,s}(x,t) = 0 \end{array}\right. &\text{ iff } \left\{\begin{array}{lcl} 2x &=& y\left(t^2 - x^T x\right) \\ 2t &=& -s\left(t^2 - x^T x\right) \end{array}\right. \\
&\text{ iff } \left\{\begin{array}{lcl} 2x &=& y\left(-\frac{2t}{s}\right) \\ 2t &=& -s\left(t^2 - x^T x\right) \end{array}\right. \\
&\text{ iff } \left\{\begin{array}{lcl} x &=& -\frac{t}{s} y \\ 2t &=& -s\left(t^2 - \frac{t^2}{s^2} y^T y\right) \end{array}\right. \\
&\text{ iff } \left\{\begin{array}{lcl} x &=& -\frac{t}{s} y \\ -2 &=& t\left(s - \frac{1}{s} y^T y\right) \end{array}\right. \quad (t \neq 0) \\
&\text{ iff } \left\{\begin{array}{lcl} x &=& \frac{2y}{s^2 - y^T y} \\ t &=& -\frac{2s}{s^2 - y^T y} \end{array}\right.
\end{aligned}$$

Thus, the maximum is reached at $(x_{max}, t_{max}) = \left(\frac{2y}{s^2 - y^T y}, -\frac{2s}{s^2 - y^T y}\right)$ and its value is:

$$g_{y,s}\left(x_{max}, t_{max}\right) = -2 + \log 4 - \log\left(s^2 - y^T y\right)$$

- *Conclusion.* Finally,

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