Convex Optimization - Homework 1

David FAGET CAÑO

Exercise 1 Which of the following sets are convex ?

1) A rectangle, *i.e.*, a set of the form $\{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i,\ i = 1, \ldots, n\}$.

A rectangle is a finite intersection of halfspaces. As a consequence, it is convex (it is what we defined as a polyhedron).

2) The hyperbolic set

$$\{x \in \mathbb{R}^2_+ \mid x_1 x_2 \geq 1\}$$

Let $S = \{x \in \mathbb{R}^2_+ / x_1 x_2 \geq 1\}$, and $x = (x_1, x_2) \in S$, $y = (y_1, y_2) \in S$.

We can prove that S is convex by using the definition of convex set, i.e. showing that S always contain the line segment between x and y.

S is colored in green

We have different cases :

- If $x_1 \geq y_1$ and $x_2 \geq y_2$, then $\ell = \theta x + (1-\theta) y$ $(0 \leq \theta \leq 1)$ and $\ell_1 \geq y_1$, $\ell_2 \geq y_2$. As a consequence, $\ell_1 \ell_2 \geq y_1 y_2 \geq 1$.
- If $x_1 \leq y_1$ and $x_2 \leq y_2$, it is similar.
- If $(y_1 - x_1)(y_2 - x_2) < 0$, we have (with $0 \leq \theta \leq 1$) :

$$(\theta x_1 + (1-\theta) y_1)(\theta x_2 + (1-\theta) y_2) = \theta x_1 x_2 + (1-\theta) y_1 y_2 - \theta(1-\theta)(y_1 - x_1)(y_2 - x_2) \geq 1$$

Conclusion : the hyperbolic set S is convex.

→ 3) The set of points closer to a given point than a given set, *i.e.*,

$$\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in S\}$$

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

This set is convex. Indeed, it can be expressed as the intersection

$$\bigcap_{y \in S} \{x \mid \|x - x_0\|_2 \leq \|x-y\|_2\}.$$

where all the sets of the form $\{x \mid \|x - x_0\|_2 \leq \|x-y\|_2\}$ are halfspaces (for fixed y). Let's prove this by manipulating the expression that defines the set:

$$\|x - x_0\|_2 \leq \|x-y\|_2 \iff (x-x_0)^T(x-x_0) \leq (x-y)^T(x-y)$$

$$\iff x^Tx - 2x_0^Tx + x_0^Tx_0 \leq x^Tx - 2y^Tx + y^Ty$$

$$\iff 2(y-x_0)^Tx \leq y^Ty - x_0^Tx_0$$

This last expression defines a halfspace, in accordance with the definition seen in class.

→ 4) The set of points closer to one set than another, *i.e.*,

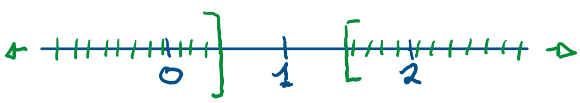
$$\{x \mid \mathbf{dist}(x, S) \leq \mathbf{dist}(x, T)\},$$

where $S, T \subseteq \mathbb{R}^n$, and $\mathbf{dist}(x, S) = \inf\{\|x - z\|_2 \mid z \in S\}$.

This set is not convex in general. Let's give an example, with $n=1$, for which the set isn't convex.
With $S = \{0, 2\}$ and $T = \{1\}$, the set of interest is

$$P := \{x \in \mathbb{R} \mid x \in (-\infty, \tfrac{1}{2}] \cup [\tfrac{3}{2}, +\infty)\}$$

which is not convex.



The set P is colored in green

→ 5) The set

$$\{x \mid x + S_2 \subseteq S_1\},$$

where $S_1, S_2 \subseteq \mathbb{R}^n$ with S_1 convex.

By definition, we have that $x + S_2 \subseteq S_1$ if $x + y \in S_1 \ \forall y \in S_2$.

Then, $\{x \mid x + S_2 \subseteq S_1\} = \bigcap_{y \in S_2} \{x \mid x + y \in S_1\}$

But all the sets of the form $\{x \mid x + y \in S_1\}$ are convex (because S_1 is convex) and can be written as $S_1 - y$. As a consequence, their intersection is also convex.

Conclusion: this set is convex.

Exercise 2 For each of the following functions determine whether it is convex or concave or not. *Optional: Determine if they are quasiconvex or quasiconcave.*

1) $f(x_1, x_2) = x_1 x_2$ on $\mathbb{R}^2_{++}$.

This function is twice differentiable in $\mathbb{R}^2_{++}$. We have that

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

The eigenvalues of $\nabla^2 f(x_1, x_2)$ are 1 and -1. As a consequence, $\nabla^2 f(x_1, x_2)$ is not positive semi-definite, and f is not convex. The same argument but with $-f$ shows that f is not concave. Let's prove now that f is not quasiconvex. For example, the sublevel set $S_1 = \{(x_1, x_2) \in \mathbb{R}^2_{++} \mid x_1 x_2 \leq 1\}$ isn't convex (we can easily see it by inspection looking the figure presented in Exercise 1 - 2)). However, $-f$ is quasiconvex, because all the sets $\{(x_1, x_2) \in \mathbb{R}^2_{++} \mid -x_1 x_2 \leq \alpha\}$ are convex (this is proved by reasoning similar than in Exercise 1 - 2)). Then f is quasiconcave.

CONCLUSION: NO convex, NO concave, NO quasiconvex, YES quasiconcave.

2) $f(x_1, x_2) = 1/(x_1 x_2)$ on $\mathbb{R}^2_{++}$.

Again, f is twice differentiable in $\mathbb{R}^2_{++}$. We have that

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{bmatrix}$$

Here, we can say that $\nabla^2 f(x_1, x_2) > 0$ by using Sylvester's criterion. This criterion states that this Hermitian matrix is positive-definite if and only if $\frac{2}{x_1^3 x_2} > 0$ and $\frac{4}{x_1^4 x_2^4} - \frac{1}{x_1^4 x_2^4} = \frac{3}{x_1^4 x_2^4} > 0$. Those conditions are obviously true in $\text{dom } f = \mathbb{R}_{++}$, so we conclude that, in particular, $\nabla^2 f(x_1, x_2) \succeq 0 \; \forall x \in \text{dom } f$.

As a consequence, f is convex (and therefore, quasiconvex).

Let's prove now that f is not quasiconcave (and therefore, not concave). We have that $-f(x_1, x_2) = -\frac{1}{x_1 x_2}$, and the sublevel set:

$S_{-1} = \{(x_1, x_2) \in \mathbb{R}^2_{++} \mid -\frac{1}{x_1 x_2} \leq -1\} = \{(x_1, x_2) \in \mathbb{R}^2_{++} \mid x_1 x_2 \leq 1\}$

is not convex (we easily see it by inspection of the figure in Exercise 1 - 2), below the green line). Then, f is not quasiconcave and not concave.

<u>CONCLUSION</u>: YES convex, YES quasiconvex, NO concave, NO quasiconcave

$\rightarrow$ 3) $f(x_1, x_2) = x_1/x_2$ on $\mathbb{R}^2_{++}$.

Again, f is twice differentiable in $\mathbb{R}^2_{++}$. Its Hessian is:

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & -\frac{1}{x_2^2} \\ -\frac{1}{x_2^2} & \frac{2x_1}{x_2^3} \end{bmatrix}$$

with eigenvalues $\lambda_1 = \frac{x_1 - \sqrt{x_1^2 + x_2^2}}{x_2^3}$ and $\lambda_2 = \frac{x_1 + \sqrt{x_1^2 + x_2^2}}{x_2^3}$.

If we take, for example, $x_1 = x_2 = 1$, we obtain $\lambda_1 < 0$ and $\lambda_2 > 0$. As a consequence, $\nabla^2 f(x_1, x_2)$ is not positive semidefinite, and f is not convex. We also see that $\nabla^2 f(x)$ is not negative semidefinite, so f is not concave.

However, it is quasiconvex and quasiconcave. Indeed, the sets $S_\alpha = \{(x_1, x_2) \in \mathbb{R}^2_{++} \mid \frac{x_1}{x_2} \leq \alpha\}$ and $S'_\alpha = \{(x_1, x_2) \in \mathbb{R}^2_{++} \mid \frac{x_1}{x_2} \geq \alpha\}$ are halfspaces (we can easily see it by writing $\frac{x_1}{x_2} \leq \alpha$ as $x_2 \geq \frac{1}{\alpha} x_1$, which is a linear combination). Then f is quasilinear.

CONCLUSION : NO convex , NO concave , YES quasiconvex , YES quasiconcave

→ 4) $f(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$, where $0 \leq \alpha \leq 1$, on $\mathbb{R}^2_{++}$.

The function f is twice differentiable in $\mathbb{R}^2_{++}$. We have :

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = \alpha\, x_1^{\alpha-1} x_2^{1-\alpha} \quad ; \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = (1-\alpha)\, x_1^{\alpha} x_2^{-\alpha}$$

and by deriving these expressions we get the Hessian matrix of f:

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} \alpha(\alpha-1)\, x_1^{\alpha-2} x_2^{1-\alpha} & -\alpha(\alpha-1)\, x_1^{\alpha-1} x_2^{-\alpha} \\ -\alpha(\alpha-1)\, x_1^{\alpha-1} x_2^{-\alpha} & \alpha(\alpha-1)\, x_1^{\alpha} x_2^{-\alpha-1} \end{bmatrix}$$

$$= \underbrace{-\alpha(\alpha-1)\, x_1^{\alpha} x_2^{1-\alpha}}_{\geq 0} \underbrace{\begin{bmatrix} -x_1^{-2} & x_1^{-1} x_2^{-1} \\ x_1^{-1} x_2^{-1} & -x_2^{-2} \end{bmatrix}}_{A}$$

We are going to prove that $A \preceq 0$. Its eigenvalues λ satisfy

$$\left(-\frac{1}{x_1^2} - \lambda\right)\left(-\frac{1}{x_2^2} - \lambda\right) - \frac{1}{x_1^2 x_2^2} = 0 \iff \lambda^2 + \lambda\left(\frac{1}{x_1^2} + \frac{1}{x_2^2}\right) = 0$$

$$\iff \lambda\left(\lambda + \frac{1}{x_1^2} + \frac{1}{x_2^2}\right) = 0$$

Then A has eigenvalues $\lambda_1 = 0$ and $\lambda_2 < 0$ (because $(x_1, x_2) \in \mathbb{R}^2_{++}$)

As a consequence, A is negative semidefinite, and $\nabla^2 f(x_1, x_2)$ is negative semidefinite too. Then f is concave (and quasiconcave). It is not convex.

Finally, with $\alpha = 1/2$, we have $f(x_1, x_2) = \sqrt{x_1 x_2}$, and the sublevel set

$$S_1 = \{(x_1, x_2) \in \mathbb{R}^2_{++} \mid \sqrt{x_1 x_2} \leq 1\}$$
$$= \{(x_1, x_2) \in \mathbb{R}^2_{++} \mid -1 \leq x_1 x_2 \leq 1\}$$
$$= \{(x_1, x_2) \in \mathbb{R}^2_{++} \mid 0 \leq x_1 x_2 \leq 1\}$$

is not convex (as seen before). As a consequence, f is not quasiconvex.

CONCLUSION : YES concave , YES quasiconcave , NO convex , NO quasiconvex.

Exercise 3 Show that following functions are convex

→ 1) $f(X) = \mathbf{Tr}(X^{-1})$ on $\mathbf{dom}\, f = \mathbf{S}^n_{++}$.

In order to prove that f is convex, we are going to use the restriction of a convex function to a line. We define the function $g: \mathbb{R} \to \mathbb{R}$ as $g(t) = f(X + tV)$, with $X \in S^n_{++}$ and $V \in S^n$. Our goal is to prove that g is convex, from which we will conclude that f is also convex. Using similar reasoning as in the example studied in class in slide 7/67 (there, $f(X) = \log\det(X)$), we have:

$$
\begin{aligned}
g(t) &= f(X + tV) \\
&= \mathrm{Tr}\left((X + tV)^{-1}\right) \\
&= \mathrm{Tr}\left(X^{-1}(I + tX^{-1/2} V X^{-1/2})^{-1}\right)
\end{aligned}
$$

Now, the matrix $X^{-1/2} V X^{-1/2}$ is a real symmetric $n \times n$ matrix. We can compute its spectral decomposition:

$$X^{-1/2} V X^{-1/2} = Q \Lambda Q^T$$

where $Q \in \mathbb{R}^{n \times n}$ is orthogonal and $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, with $\lambda_i \in \mathbb{R}$ the eigenvalues of $X^{-1/2} V X^{-1/2}$. We have:

$$
\begin{aligned}
\mathrm{Tr}\left(X^{-1}(I + tX^{-1/2} V X^{-1/2})^{-1}\right) &= \mathrm{Tr}\left(X^{-1}(I + tQ\Lambda Q^T)^{-1}\right) \\
&= \mathrm{Tr}\left(X^{-1} Q (I + t\Lambda)^{-1} Q^T\right) \\
&= \mathrm{Tr}\left(Q^T X^{-1} Q (I + t\Lambda)^{-1}\right)
\end{aligned}
$$

Using the definition of trace, we have:

$$\mathrm{Tr}\left(Q^T X^{-1} Q (I + t\Lambda)^{-1}\right) = \sum_{i=1}^{n} \left(Q^T X^{-1} Q (I + t\Lambda)^{-1}\right)_{ii}$$

But $(I + t\Lambda)^{-1} = \mathrm{diag}\left((1 + t\lambda_i)^{-1}\right)_{i=1,\dots,n}$, then:

$$\sum_{i=1}^{n} \left(Q^T X^{-1} Q (I + t\Lambda)^{-1}\right)_{ii} = \sum_{i=1}^{n} (Q^T X^{-1} Q)_{ii} \, (1 + t\lambda_i)^{-1}$$

But this last expression is a nonnegative weighted sum of $(1 + t\lambda_i)^{-1}$ functions, which are all convex. To conclude, Let's prove that $\frac{1}{1+t\lambda_i}$ is in fact convex $\forall i$. Remember that $t \in \mathbb{R}$ and $\lambda_i \in \mathbb{R}$ $\forall i$. A similar reasoning than in Exercise 1 - 2) proves this. As a consequence, g is convex, and f is convex too.

→ 2) $f(X, y) = y^T X^{-1} y$ on $\mathbf{dom}\, f = \mathbf{S}^n_{++} \times \mathbb{R}^n$ *Hint: express it as a supremum*

In order to prove the convexity of f, we are going to express it as a supremum of a set of convex functions. As seen in class, this implies the convexity of f.
Using the second example given on the slide 24/67 (there, $f(x) = \frac{1}{2} x^T Q x$ with $Q \in S^n_{++}$) we deduce that:

$$y^T X^{-1} y = \sup_{x, Q} \left(2 y^T x - x^T Q x \right)$$

If we denote $z(x, Q) = 2y^T x - x^T Q x$, we have that z is always convex (it is linear in x and linear in Q). As a consequence, f is convex.

→ 3) $f(X) = \sum_{i=1}^n \sigma_i(X)$ on $\mathbf{dom}\, f = \mathbf{S}^n$, where $\sigma_1(X), \ldots, \sigma_n(X)$ are singular values of a matrix $X \in \mathbb{R}^{n \times n}$. *Hint: express it as a supremum*

One way to prove the convexity of f is showing that it defines a norm (we know that all norms are convex). We are going to check that the conditions of the definition of norm are satisfied (this norm is known as the trace norm).

- $f(X) \geq 0$ and $f(X) = 0 \Leftrightarrow X = 0$.

It is obvious that $f(X) \geq 0$ because singular values are always non-negative. Let's prove the second condition.

$$\Rightarrow) \quad \left. \begin{array}{l} f(X) = 0 \Rightarrow \sum_{i=1}^n \sigma_i(X) = 0 \\ \sigma_i(X) \geq 0 \end{array} \right\} \Rightarrow X = 0$$

$$\Leftarrow) \quad X = 0 \Rightarrow \sigma_i(X) = 0 \ \forall i \Rightarrow f(X) = 0$$

- $f(\alpha X) = |\alpha| f(X) \quad \forall \alpha \in \mathbb{R}, \ X \in \mathbb{S}^n.$

Here, we are going to use that $\sigma_i^2(X) = \lambda_i(XX^*)$ and that $X \in \mathbb{S}^n$. We have:

$$f(\alpha X) = \sum_{i=1}^{n} \sigma_i(\alpha X) = \sum_{i=1}^{n} \sqrt{|\lambda_i(\alpha^2 X X^T)|} = \sum_{i=1}^{n} \sqrt{|\alpha^2| \, |\lambda_i(XX^*)|}$$

$$= |\alpha| \sum_{i=1}^{n} \sqrt{|\lambda_i(XX^*)|} = |\alpha| \sum_{i=1}^{n} \sigma_i(X) = |\alpha| f(X)$$

- $f(X+Y) \leq f(X) + f(Y) \quad \forall x, y \in \mathbb{R}, \ X, Y \in \mathbb{S}^n$

Here, we have that:

$$f(X+Y) = \sum_{i=1}^{n} \sigma_i(X+Y) \underset{\circledast}{\leq} \sum_{i=1}^{n} \sigma_i(X) + \sum_{i=1}^{n} \sigma_i(Y) = f(X) + f(Y)$$

A rigorous proof of $\circledast$ can be found in the paper "Singular value inequalities for matrix sums and minors" written by R.C. Thompson.

In light of the above, f is a norm, thus it is convex.

Exercise 4 We define the *monotone nonnegative cone* as

$$K_{m+} = \{x \in \mathbb{R}^n \mid x_1 \geq x_2 \geq \cdots \geq x_n \geq 0\}.$$

i.e., all nonnegative vectors with components sorted in nonincreasing order.

→ 1. Show that K_{m+} is a proper cone.

Let's prove the 3 conditions which define a proper cone.

- K_{m+} is closed

$$K_{m+} = \{ x \in \mathbb{R}^n \,/\, x_1 \geq x_2 \geq \dots \geq x_n \geq 0 \}$$
$$= \bigcap_{i=1}^{n-1} \{ x_i \geq x_{i+1} \} \cap \{ x_n \geq 0 \}$$

But the above expression is an intersection of closed sets (linear inequality $\geq$). Then K_{m+} is closed.

- K_{m+} is solid

Let's prove that K_{m+} has nonempty interior. For example, $p = (m, m-1, \dots, 1) \in K_{m+}$, then $\text{int}(K_{m+}) \neq \emptyset$.

⊙ K_{m+} is pointed

Let's prove that K_{m+} contains no line.

Suppose that $x = (x_1, x_2, \dots, x_n) \in K_{m+}$.

Then $x_1 \geq x_2 \geq \dots \geq x_n \geq 0 \Rightarrow x_i \geq 0 \quad \forall i = 1, \dots, n$

If K_{m+} contained a line, we would have that $-x \in K_{m+}$

But $-x \in K_{m+}$ iff $-x_1 \geq -x_2 \geq \dots \geq -x_n \geq 0$, which implies that $x_i \leq 0 \quad \forall i = 1, \dots, n$. We would have $x = 0$, which indicates that the cone contains no entire line.

→ 2. Find the dual cone K^*_{m+}.

We want to find $K^*_{m+} = \{ y \mid y^T x \geq 0 \quad \forall x \in K_{m+} \}$

$= \{ y \mid \sum_{i=1}^{n} y_i x_i \geq 0 \quad \forall x \in K_{m+} \}$

As $x \in K_{m+}$, we can intuit that

$$K^*_{m+} = \{ y \mid \{y_1 \geq 0\} \cap \{y_1 + y_2 \geq 0\} \cap \dots \cap \{y_1 + \dots + y_n \geq 0\} \}$$

Let's prove this, by first proving the following expression using mathematical induction:

$$\sum_{i=1}^{n} y_i x_i = (x_1 - x_2) y_1 + (x_2 - x_3)(y_1 + y_2) + (x_3 - x_4)(y_1 + y_2 + y_3) + \dots + (x_{n-1} - x_n)(y_1 + \dots + y_{n-1}) + x_n (y_1 + \dots + y_n)$$

For $n = 1$, it is obvious. Now suppose that the identity is true for n, and let's see that it is also true for $n+1$.

$$\sum_{i=1}^{n+1} y_i x_i = \sum_{i=1}^{n} y_i x_i + y_{n+1} x_{n+1}$$

$$= (x_1 - x_2) y_1 + \dots + x_n (y_1 + \dots + y_n) + y_{n+1} x_{n+1}$$

$$= (x_1 - x_2) y_1 + \dots + x_n \sum_{i=1}^{n} y_i + x_{n+1} \left(-\sum_{i=1}^{n} y_i + \sum_{i=1}^{n} y_i \right) + x_{n+1} y_{n+1}$$

$$= (x_1 - x_2) y_1 + \dots + (x_n - x_{n+1})(y_1 + \dots + y_n) + x_{n+1} (y_1 + \dots + y_{n+1})$$

Then the formula is true $\forall n \in \mathbb{N}$.

As $x \in K_{m+}$, we have that $(x_{i+1} - x_i) \geq 0 \quad \forall i = 1, \dots, n-1$

and $x_n \geq 0$. As a consequence,

$$K_{m+} = \{ y \mid \{y_1 \geq 0\} \cap \{y_1 + y_2 \geq 0\} \cap \dots \cap \{y_1 + y_2 + \dots + y_n \geq 0\} \}.$$

▷ **Exercise 5** Derive the conjugates of the following functions.

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

We have $f(x) = \max_{i=1,\dots,n} x_i$, and we want to find

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

Suppose that $y \geq 0$. Then $y^T x - \max_{i=1,\dots,n} x_i \leq \max_{i=1,\dots,n} x_i \left(\sum_{i=1}^{n} y_i - 1 \right)$

- If $\sum_{i=1}^{n} y_i > 1$, we can take $x = (\ell, \ell, \dots, \ell)^T$ and doing $\ell \to +\infty$, we obtain that $y^T x - \ell \xrightarrow[\ell \to +\infty]{} +\infty$
- If $\sum_{i=1}^{n} y_i < 1$, we do the same with $x = (-\ell, -\ell, \dots, -\ell)^T$ to obtain that $y^T x + \ell \xrightarrow[\ell \to +\infty]{} +\infty$
- If $\sum_{i=1}^{n} y_i = 1$, we have $y^T x - \max_{i=1,\dots,n} x_i \leq 0$.

Suppose now that it $\exists t \mid y_t < 0$. We can choose x such as

$$x = (0, \dots, 0, \underset{\substack{\uparrow \\ \text{index } t \ (x_t = \ell)}}{-\ell}, 0, \dots, 0)$$

By doing $\ell \to +\infty$, we have $y^T x - \max_{i=1,\dots,n} x_i = -\ell y_t \xrightarrow[\ell \to +\infty]{} +\infty$

We conclude that:

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

→ 2) *Sum of largest elements.* $f(x) = \sum_{i=1}^{r} x_{[i]}$ on $\mathbb{R}^n$.

Here, we proceed in a similar way to the previous case.

We first suppose that $y \geq 0$.

- If it $\exists t \mid y_t > 1$ and we choose $x = (0, \dots, \underset{\substack{\uparrow \\ \text{index } t}}{\ell}, \dots, 0)$ we have $y^T x - \sum_{i=1}^{r} x_{[i]} = \ell (y_t - 1) \xrightarrow[\ell \to +\infty]{} +\infty$

- If $0 \leq y \leq 1$, $y^T x - \sum_{i=1}^{r} x_{[i]} = \sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{r} x_{[i]}$
  - $\to$ If $\sum_{i=1}^{n} y_i \neq r$, then we can choose $x = (\ell, \dots, \ell)$ to see that $y^T x - \sum_{i=1}^{r} x_{[i]} = \ell \left(\sum_{i=1}^{n} y_i - r \right) \xrightarrow[\ell \to +\infty]{} +\infty$
  - $\to$ If $\sum_{i=1}^{n} y_i = r$, we have $y^T x - \sum_{i=1}^{r} x_{[i]} \leq 0$ for all x. Then $f^*(y) = 0$ in this case.

Now, if we suppose that it $\exists t \mid y_t < 0$, we can choose $x = (0, \dots, \underset{\substack{\uparrow \\ \text{index } t}}{-\ell}, \dots, 0)$ and by doing $\ell \to +\infty$, we obtain $y^T x - f(x) \xrightarrow[\ell \to +\infty]{} +\infty$.

In light of the above, the conclusion is that we have

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

3) *Piecewise-linear function on* $\mathbb{R}$. $f(x) = \max_{i=1,\dots,m}(a_i x + b_i)$ 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 $f(x) = a_k x + b_k$.

The two assumptions tell us that f is a piecewise linear function with changepoints $p_i = \dfrac{b_i - b_{i+1}}{a_{i+1} - a_i}$.

We want to determine

$$f^*(y) = \sup_{x \in \mathbb{R}} \left\{ y^T x - \max_{i=1,\dots,m} (a_i x - b_i) \right\}$$

- If $y < a_1$ or $y > a_m$, we have that
$$y^T x - \max_{i=1,\dots,m} (a_i x - b_i)$$
does not have a finite supremum.
- Suppose now that $a_1 \leq y \leq a_m$. Then, as we suppose that the a_i are sorted in increasing order, we have that it $\exists k \mid a_k \leq y \leq a_{k+1}$. Then $y - a_i \geq 0 \ \forall i = 1, \dots, k$ and $y - a_i \leq 0 \ \forall i = k+1, \dots, m$. As a consequence, the supremum of $y^T x - \max_{i=1,\dots,m} (a_i x - b_i)$ is reached at p_k.

We obtain
$$f^*(y) = y\,\rho_k - a_k\,\rho_k - b_k$$
$$= y\,\frac{b_k - b_{k+1}}{a_{k+1} - b_k} - a_k\,\frac{b_k - b_{k+1}}{a_{k+1} - a_k} - b_k$$
$$= (y - a_k)\left(\frac{b_k - b_{k+1}}{a_{k+1} - a_k}\right) - b_k$$

In light of the above, we conclude that:
$$f^*(y) = \begin{cases} (y-a_k)\left(\dfrac{b_k - b_{k+1}}{a_{k+1} - a_k}\right) - b_k & \text{if } \begin{array}{l} a_1 \leq y \leq a_m \\ a_k \leq y \leq a_{k+1} \end{array} \\ +\infty & \text{otherwise.} \end{cases}$$