Optimization, AIMS CDT

Francesco Capuano

3rd November, 2025

1 Convex sets, functions and problems

1.1 An Optimization Problem

$$\min_{x \in \mathbb{R}^2} x_1^2 + x_2^2 \quad \text{s.t.} \quad f_1(x) = \frac{x_1}{1 + x_2^2} \leq 0, \;\; h_1(x) = (x_1 + x_2)^2 = 0.$$

(a) Nonconvex formulation. For an optimization problem to be *convex*, the objective function and constraints must be convex as well. The function $f(x_1, x_2) = x_1^2 + x_2^2$ is convex everywhere, as the Hessian of f, $H_f(x) = 2I$). The equality constraint $h_1(x) = (x_1 + x_2)^2$ is a convex function as well, as its Hessian $H_{h_1}(x) = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$ is positive semi-definite ($\lambda_1 = 0, \lambda_2 = 4$). Conversely, $H_{f_1}(x) = \begin{pmatrix} 0 & -\dfrac{2x_2}{(x_2^2+1)^2} \\ -\dfrac{2x_2}{(x_2^2+1)^2} & \dots \end{pmatrix}$ is not positive semi-definite. Indeed, the determinant is

$$\det H_{f_1}(x) = -\frac{4x_2^2}{(x_2^2+1)^2} \leq 0 \, \forall x \in \mathbb{R}^2,$$

which in turn makes the whole problem non-convex.

(b) Feasible set & equivalent convex problem. As $1 + x_2^2 > 0 \, \forall x \in \mathbb{R}^2$, the inequality $\frac{x_1}{(1+x_2^2)} \leq 0$ coincides with $x_1 \leq 0$. Moreover, $(x_1 + x_2)^2 = 0$ is equivalent to the affine equality $x_1 + x_2 = 0$. Thus the feasible set is

$$\mathcal{X} = \{x \in \mathbb{R}^2 \mid x_1 \leq 0, \; x_1 + x_2 = 0\} \equiv \{t \in \mathbb{R}^+ : (-t, t)\}.$$

On $\mathcal{X}$ the objective function coincides with $x_1^2 + x_2^2 = 2t^2$, which is minimized at $t^\star = 0$, i.e. $x^\star = (0, 0)$. A *convex* reformulation that is equivalent to the original problem therefore is:

$$\min_{x \in \mathbb{R}^2} x_1^2 + x_2^2 \quad \text{s.t.} \quad x_1 + x_2 = 0, \;\; x_1 \leq 0,$$

or, in one-dimension only, $\min_{t \geq 0} t^2$. Both formulations are convex.

1.2 Hyperbolic constraints

For $x \in \mathbb{R}^n$, $y, z \in \mathbb{R}$ with $y \geq 0$, $z \geq 0$, prove that:

$$\| \begin{pmatrix} 2x, & y - z \end{pmatrix}^\top \|_2^2 \leq (y+z)^2 \tag{1}$$

Proof. Expanding the right-hand side,

$$\| \begin{pmatrix} 2x, & y - z \end{pmatrix}^\top \|_2^2 \leq (y+z)^2 \tag{2}$$

$$\begin{pmatrix} 2x, & y - z \end{pmatrix} \begin{pmatrix} 2x, \\ y - z \end{pmatrix} \leq (y+z)^2 \tag{3}$$

$$4x^\top x + (y-z)^2 \leq y^2 + 2yz + z^2 \tag{4}$$

$$4x^\top x \leq 4yz \tag{5}$$

$$x^\top x \leq yz \tag{6}$$

□

(a) Maximizing the harmonic mean. Let $t_i \geq (a_i^\top x - b_i)^{-1} > 0$ (from $\mathcal{X} = \{x \in \mathbb{R}^n : Ax > b\}$). Then, the optimization problem can be rewritten as:

$$\begin{aligned} &\min_t \mathbf{1}^\top t \\ &t_i \geq (a_i^\top x - b_i)^{-1}, \quad \forall i \end{aligned} \tag{7}$$

The constraint:

$$t_i(a_i^\top x + b_i) \geq 1, \qquad t_i \geq 0, \quad a_i^\top x + b_i > 0,$$

is referred to as *hyperbolic* because, if we let $u = a_i^\top x + b_i$, $v = t_i$, the *equality* $uv = 1$ defines a hyperbola in the (u, v)-plane, and the feasible set:

$$\mathcal{X} = \{(u, v) \mid uv \geq 1\}$$

corresponds to the region on (or above) the hyperbola, thus the name *hyperbolic constraint.*

Clearly, $\mathcal{X}$ is *not* convex—the segment between two points on opposite branches do not satisfy $uv \geq 1$. However, one can handle hyperbolic constraints (i.e., hyperbolic feasible sets) within a convex optimization problem by rewriting them in terms of a *second-order cone* (SOC), representing a convex set in a higher-dimensional space.

From eq. 1, by choosing $w = 2$, we obtain the equivalence

$$uv \geq 1 \quad \Longleftrightarrow \quad \left\| \begin{bmatrix} 2 \\ u - v \end{bmatrix} \right\|_2 \leq u + v, \quad u, v \geq 0.$$

Substituting back $u = a_i^\top x + b_i$ and $v = t_i$, the hyperbolic constraint can thus be expressed as the SOC (convex) constraint

$$\left\| \begin{bmatrix} 2 \\ a_i^\top x + b_i - t_i \end{bmatrix} \right\|_2 \leq a_i^\top x + b_i + t_i, \qquad t_i \geq 0, \quad a_i^\top x + b_i \geq 0.$$

This result allows to tackle problem 7 using convex optimization, as both the objective function and constraints are convex.

3. Support functions

Let $S_C(y) := \sup\{y^\top x | x \in C\}$ (possibly $+\infty$).

(a) Convexity of S_C. The support function of a set $C \subseteq \mathbb{R}^n$ is

$$S_C(y) = \sup_{x \in C} y^\top x.$$

Each function $f_x(y) = y^\top x$ is linear (hence convex) in y, and S_C is their pointwise supremum:

$$S_C(y) = \sup_{x \in C} f_x(y).$$

Since the supremum of convex functions is convex (§3.2.3, B&V), S_C is in turn convex for any set C.

(b) $S_C = S_{\text{conv}(C)}$. Clearly, $\text{conv}(C) \supseteq C$, which results in $S_{\text{conv}(C)} \geq S_C$, in keeping with the geometric interpretation of S_C. Conversely, for any given $x \in \text{conv}(C) : x = \sum_i \theta_i x_i$ ($\theta_i \geq 0$, $\sum \theta_i = 1$, $x_i \in C$),

$$y^\top x = \sum_i \theta_i \, y^\top x_i \leq \sum_i \theta_i \, S_C(y) \leq S_C(y).$$

The equality holds for any $x \in \text{conv}(C)$, and thus it must hold for the supremum over $\text{conv}(C)$ too, yielding $S_{\text{conv}(C)} \gtreqless S_C$.

4. Largest-L norm

For $x \in \mathbb{R}^n$, sort $|x|$ in non-increasing order, and define $\|x\|_{[L]} = \sum_{i=1}^{L} |x|_{[i]}$.

(a) Convexity. The largest-L norm can alternatively be represented as:

$$\|x\|_{[L]} = \max_y \Big\{ y^\top x \;:\; \|y\|_\infty \leq 1,\; \|y\|_1 \leq L \Big\}, \tag{8}$$

similarly to how support functions are represented too. The first constraint on y ensures the norm computed solving $\max_y y^\top x$ is not larger than the real norm, while the second ensures it is not smaller. Being a (support) linear function defined on a convex set—intersecting sets preserves convexity, and "all norms on $\mathbb{R}^n$ are convex functions", §3.1.5, B&V)—$\|x\|_{[L]}$ is convex.

(b) Integer programming (IP) formulation. In keeping with the result just presented, one can write the largest-L norm as an integer programming (i.e., combinatorial optimization) problem, where the variables $z_i \in \{0, 1\}$ select which indices in x to use to form the norm. This results in:

$$\|x\|_{[L]} = \max_z \Big\{ \sum_{i=1}^{n} |x_i| \, z_i \;:\; \sum_{i=1}^{n} z_i = L,\; z_i \in \{0,1\} \Big\}.$$

(c) Linear programming formulation. Given in (a).