Norm 1 optimization under unit vectors

Sharon Rotgaizer

April 27, 2024

It is relatively easy to optimize convex functions without any constraints, but in CS, we need many times to solve problems using the constraint of $\|x\| = 1$. We will see how to solve two fundametal problems.

Unlike in l2, we will deal with max and min problems separately, starting with maximinzing.

Let $x \in \mathbb{R}^d, A \in \mathbb{R}^{n\times d}, b \in \mathbb{R}^d$, and $f(x) = \|Ax - b\|_1$. Our problem is:

$$\max_{\|x\|=1} f(x) = \max_{\|x\|=1} \|Ax - b\|_1$$

From the definition, note that:

$$f(x) = \sum_{i=1}^{n} |A_{i*}^\top x - b_i|$$

We now partition $\mathbb{R}^d$ into 2^n subspaces in the following method, by covering all the options.

$$\mathbb{R}^d = \bigcup_{c\in\{-1,1\}^n} \{x : (Ax - b) \circ c \geq 0_d\} = \bigcup_{c\in\{-1,1\}^n} \{x : \forall i : (A_{i*}x - b_i)c_i \geq 0\}$$

For example, in 1d,

$$|2x - 3| = \begin{cases} 2x - 3 & x \geq 1.5 \\ -(2x - 3) & x \leq 1.5 \end{cases} = \bigcup_{c\in\{-1,1\}} c(2x - 3)$$

Denote $S_c := \{x : (Ax - b) \circ c \geq 0_n\}$, So $\mathbb{R}^d = \cup_{c\in\{-1,1\}^d} S_c$.

Note: It might hold that for some c, $S_c = \emptyset$. When we proceed, we can ignore this case, since we can always check the rank of $A'(Ax - b) \circ c$, so if $n > d$ and $r(A') = d$, there won't be any solution, which means $\forall x : x \notin S_c$.

We can express $\mathrm{Im}(f)$ by:

$$f(\mathbb{R}^d) = \bigcup_{c\in\{-1,1\}^d} f(S_c)$$

Now, let as look at $f(S_c)$. Let $x \in S_c$, then from the absolute value definition:

$$f(x) = \sum_{i=1}^{n} |A_{i*}^\top x - b_i| = \sum_{i=1}^{n} (A_{i*}x - b_i)c_i = w^\top x - d$$

Where $w = \sum_i A_{i*}c_i$, $d = \sum_i b_i c_i$. Our problem becomes:

$$\max_{\substack{\|x\|=1 \\ x\in S_c}} f(x) = \max_{\substack{\|x\|=1 \\ x\in S_c}} w^\top x - d$$

We will use two arguments to solve this problem:

Let $x^* \in S_c$ be the solution. $\|x^*\| = 1 \iff \exists x \in S_c : \|x\| = 1$. This is trivial from the problem definition.

If $\exists x \in S_c$ s.t. $\|x\| = 1$, than:

$$x^* = \arg\max_{\|x\|=1} w^\top x = \frac{w}{\|w\|} = \arg\max_{\substack{\|x\|=1\\ x\in S_c}} w^\top x$$

Proof: If $x^* \in S_c$, trivial. I know for sure that it holds, but can't prove it. it might be also given using the Lagrangian.

The algorithm is:

- For each $c \in \{-1, 1\}^d$, save $x_c = \frac{w}{\|w\|}$. if $x_c \notin S_c$, throw it.

So, $\max_{\|x\|=1} \|Ax - b\|_1 = \max_c \{x_c\}$.
Simulation for all suspected points found by that algorithm is here

For the minimum problem, we need to present few observations first.

- Look at specific S_c, then $f(x) = w^\top x - d$. Under the constraint $\|x\| = 1$, this is an ellipsoid (lemma 1).
- $f(S_c) \subseteq E_c$, where $E_c = \{x : w^\top x - d = 0, \|x\| = 1\}$ (the non-constrained ellipsoid).
- $\max_{\|x\|=1} f(S_c) = \max\{f(x) : x \in E_c\}$.

From those observations, $\min_{\|x\|=1} f(S_c)$ will be a point on the ellipsoid (hull) E_c which is a convex shape. We could potentially start at the max, "slide down", but not forever. We are limited to $x \in S_c$. Hence, the feasible point will be the hull of S_c:

$$x^* \in \{x \in S_c : \exists i : A_{i*}x - b_i = 0\} \subseteq \bigcup_c S_c$$

So x^* is on all the subspaces and hence, this constraint (of being on S_c) can be dropped.
This will allow us to reduce the run time from $\Theta(2^d)$ to $\Theta(d)$.

Assume that for some i, $A_{i*}x - b_i = 0$. One can extract x_d (w.l.o.g), substitute it in $\|x\| = 1$, it will be an ellipsoid constraint. Our new problem is:

$$f(x) = \sum_{i=1}^{d} |A_{i*}^\top x - b_i| = \sum_{i=1}^{d-1} |B_{i*}^\top x - d_i|$$

Where B, d are obtained from the substitution of x_d in $A_{i*}^\top x - b_i$. There is only one ellipsoid hull constraint.

Since both the ellipsoid (hull) and the function are now convex, we can seek for $\|x\| \leq 1$.
So, $\min_{\|x\|=1} \|Ax - b\|_1 = \min_{\|x\|=1} \|Ax - b\|_1$ after the substitution.

Look at the simulation here. You can see that the ellipsoid is decreasing, and its minima is on the edges of the ellipsoid, meaning on one the 2 hyperplanes restricting it.

Sharon: Note the magic here: when we handled only $\|x\| = 1$ in d dimensions, f as some complex shape. When reduced to $d - 1$ dimensions, we are not depend on "z" and hence everything becomes nice.

Lemma 1: the hyperplane $f(x) = w^\top x - b$ s.t. $\|x\| = 1$ is an ellipsoid (hull).

Simulation

Proof: missing

Appendix: We will do the full substitution process. From $A_{i*}x - b_i = 0$ we get:

$$x_d = \frac{1}{A_{id}} \left(b_i - \sum_{j=1}^{d-1} A_{ij} x_j \right)$$

Lets look at $A_{k*}x - b_k$ term in f:

$$\begin{aligned} A_{k*}x - b_k &= \sum_{j=1}^{d-1} A_{kj} x_j + A_{kd} x_d - b_k = \\ &= \sum_{j=1}^{d-1} (A_{kj} - \frac{A_{kd}}{A_{id}}) x_j + \frac{bi}{A_{id}} - b_k \end{aligned}$$

Putting x_d into $\|x\| = 1$ yields to:

$$x_1^2 + ... + x_{d-1}^2 + \frac{1}{A_{id}^2} \left(b_i - \sum_{j=1}^{d-1} A_{ij} x_j \right)^2 = 1$$

$$\sum_{j=1}^{d-1} (A_{ij}^2 + A_{id}^2) x_j^2 + \sum_{j \neq k} A_{ij} A_{ik} x_j x_k - 2b_i \sum_{j=1}^{d-1} A_{ij} x_j + b_i^2 = A_{id}^2$$

Sharon: it might be quadratic?