Convex Optimization

Yunwei Ren

Contents

2 Convex Sets

2.1 Definition of convexity

1.

Proof. For $k = 2$, $\theta_1 x_1 + \theta_2 x_2 \in C$ holds by by definition. We argue by induction on k and assume that the inclusion holds for $k < m$. When $k = m$, denoting $\sum_{i=1}^{m-1} \theta_i$ by s,

$$\sum_{i=1}^{m} \theta_i x_i = s \sum_{i=1}^{m-1} \frac{\theta_i x_i}{s} + \theta_m x_m.$$

Since $\sum_{i=1}^{m-1} \theta_i / s = 1$, by the induction hypothesis, $\sum_{i=1}^{m-1} \theta_i x_i / s \in C$. Meanwhile, as $s + \theta_m = 1$, $\sum_{i=1}^{m} \theta_i x_i \in C$, completing the proof. □

2.

Proof. Clear that the intersection of two convex sets is still convex. Hence, the intersection of $C \subset \mathbb{R}^n$ and any line is convex as long as C is convex.

Now we suppose that the intersection of C and any line is convex. For any $x_1, x_2 \in C$, $C_l = C \cap \{\theta x_1 + (1-\theta) x_2 : \theta \in \mathbb{R}\}$ is convex and therefore $\theta x_1 + (1-\theta) x_2 \in C_l \subset C$ for every $0 \leq \theta \leq 1$. Thus, C is convex.

The above argument, *mutatis mutandis*, gives the second result. □

3.

Proof. For every $\theta \in [0, 1]$, the process of bisecting the interval implies there exists a series $\langle \delta_n \rangle$ whose sum is θ. Hence, for every $a, b \in C$, $x_n = a + (b - a) \sum_{n=1}^{\infty} \delta_n$ converges to $a + \theta(b - a)$. Meanwhile, the midpoint convexity implies $x_n \in C$ for every n. And since C is closed, $a + \theta(b - a) \in C$. Thus, C is convex. □

4.

Proof. Let D be the intersection of all convex sets containing C. If $x \in C$, then its is a convex combination of some points in C. Hence, for every convex set containing C, it contains x. Therefore, $\mathbf{conv}\, C \subset D$. For the converse, since $\mathbf{conv}\, C$ itself is a convex set containing C, $D \subset \mathbf{conv}\, C$. Thus, $\mathbf{conv}\, C = D$. □

2.2 Examples

5.

Solution. $|b_2 - b_1| / \|a\|_2$. □

7.

Proof. $\|x - a\|_2 \leq \|x - b\|_2$ iff $\langle x - a, x - a \rangle \leq \langle x - b, x - b \rangle$ iff $2\langle x, b - a \rangle \leq \langle b, b \rangle - \langle a, a \rangle$. Namely, $2(b - a)^T x \leq \|b\|_2^2 - \|a\|_2^2$. □

2.8

Proof.

(a) It is trivial when a_1 and a_2 are linearly dependent, so we assume that a_1 and a_2 are linearly independent. We first tackle the problem for orthonormal a_1 and a_2 and then reduce the general situation to it.

Suppose that a_1 and a_2 are orthonormal. Let $S_0 = \operatorname{span}(a_1, a_2)$ and $(b_1, \dots, b_{n-2})$ a basis of $S_0^\perp$. Then

$$x \in S_0 \quad \Leftrightarrow \quad \begin{bmatrix} b_1^T \\ \vdots \\ b_{n-2}^T \end{bmatrix} x = Bx = 0.$$

For $y = y_1 a_1 + y_2 a_2 \in S_0$, $y_1 \le 1$ iff $a_1^T y \le 1$ as (a_1, a_2) is an orthonormal basis of S_0. Hence,

$$-1 \le y_1, y_2 \le 1 \quad \Leftrightarrow \quad \begin{bmatrix} a_1^T \\ a_2^T \\ -a_1^T \\ -a_2^T \end{bmatrix} y = Ay \preceq \mathbf{1}.$$

Thus, for orthonormal a_1 and a_2, $S = \{x : \ Bx = 0, \ Ax \preceq \mathbf{1}\}$, a polyhedron.

Now we only assume the liner independence of a_1 and a_2. We know that there exists some invertible n-by-n matrix[1] R such that $[\tilde{a}_1, \tilde{a}_2] = R[a_1, a_2]$ and $\tilde{a}_1$ and $\tilde{a}_2$ are orthonormal. Denoting the set described in the problem with respect to u_1 and u_2 by $S(u_1, u_2)$, $x \in S(a_1, a_2)$ iff $Rx \in S(\tilde{a}_1, \tilde{a}_2)$ iff $Rx \in \{x : \ \tilde{B}x = 0, \ \tilde{A}x \preceq 1\}$ where the meaning of $\tilde{A}$ and $\tilde{B}$ are described in the previous passage. Hence,

$$S(a_1, a_2) = \{x : \ \tilde{B}Rx = 0, \ \tilde{A}Rx \preceq 1\}.$$

(b) Yes, and the provided form has already satisfied the requirement.

(c) No. Note that $\langle x, y \rangle_2 \le 1$ for all y with 2-norm 1 implies

$$\|x\|_2 = \langle x, x/\|x\| \rangle_2 \le 1.$$

And by the Cauchy-Schwarz inequality, for every $\|x\| \le 1$, $\langle x, y \rangle_2$ holds for every $\|y\|_2 = 1$. Hence, S is the intersection of the unit ball and $\{x : \ x \succeq 0\}$, which is not a polyhedron.

(d) Yes. Let $\tilde{S} = \{x \in \mathbb{R}^n : \ x \succeq 0, \ \|x\|_\infty \le 1\}$, which is clearly a polyhedron since when $x \succeq 0$, $\|x\|_\infty \le 1$ is equivalent to $[e_1, \dots, e_n]x \preceq \mathbf{1}$ where e_i is the i-th vector in the standard basis of $\mathbb{R}^n$.

Now we show that $S = \tilde{S}$. Suppose that $x \succeq 0$. If $\langle x, y \rangle_2 \le 1$ for all y with 1-norm 1, then $x_i = \langle x, e_i \rangle_2 \le 1$. Namely, $\|x\|_\infty \le 1$. Meanwhile, if $\|x\|_\infty \le 1$,

$$\langle x, y \rangle \le \sum_{i=1}^{n} x_i |y_i| \le 1$$

as it is just the weighted average of $x_1, \dots, x_n$. Hence, $S = \tilde{S}$, completing the proof. □

[1] We can use QR factorization to construct the matrix explicitly

2.9

Proof.

(a) By the definition,

$$\begin{aligned}
x \in V &\Leftrightarrow \|x - x_0\|_2^2 - \|x - x_i\|_2^2 \le 0 \\
&\Leftrightarrow 2\langle x, x_i - x_0\rangle \le \langle x_i, x_i\rangle - \langle x_0, x_0\rangle \quad \text{for } i = 1, \dots, K \\
&\Leftrightarrow 2\begin{bmatrix} \langle x, x_1 - x_0\rangle \\ \vdots \\ \langle x, x_K - x_0\rangle \end{bmatrix} \preceq \begin{bmatrix} \|x_1\|_2^2 - \|x_0\|_2^2 \\ \vdots \\ \|x_K\|_2^2 - \|x_0\|_2^2 \end{bmatrix} \\
&\Leftrightarrow 2\begin{bmatrix} (x_1 - x_0)^T \\ \vdots \\ (x_K - x_0)^T \end{bmatrix} x \preceq \begin{bmatrix} \|x_1\|_2^2 - \|x_0\|_2^2 \\ \vdots \\ \|x_K\|_2^2 - \|x_0\|_2^2 \end{bmatrix}
\end{aligned}$$

Hence, V is a polyhedron. Intuitively, the border of a Voronoi set are the lines with the same distances to x_0 and x_i.

(b) Suppose that $P = \{x : \alpha_k^T x \le b_k, k = 1, \dots, K\}$. Let x_0 be any point of P and we construct the other points by reflection. For each k, let $\tilde{x}_k$ be any point of $\{x : \alpha_k^T x = b_k\}$, $U_k = I - 2\alpha_k \alpha_k^T / \|\alpha_k\|_2^2$, the Householder matrix, and

$$R_k(x) = U_k(x - \tilde{x}_k) + \tilde{x}_k = x + 2\frac{\alpha_k}{\|\alpha_k\|_2^2}(b_k - \alpha_k^T x).$$

It is easy to verified that P is the Voronoi region of x_0 with respect to $R_1(x_0), \dots, R_K(x_0)$.

□

10.

Proof.

(a) Suppose $x_1, x_2 \in C$ and $\theta \in (0,1)$. Let $x = \theta x_1 + (1-\theta)x_2$. Since A is symmetric, $x_2^T A x_1 = x_1^T A x_2$. Thus,

$$\begin{aligned}
f(x) &= x^T A x + b^T x + c \\
&= \theta^2 x_1^T A x_1 + 2\theta(1-\theta) x_1^T A x_2 + (1-\theta)^2 x_2^T A x_2 \\
&\quad + \theta b^T x_1 + (1-\theta) b^T x_2 + \theta c + (1-\theta)c.
\end{aligned}$$

Note that

$$\begin{aligned}
\theta^2 x_1^T A x_1 + \theta b_1^T x_1 + \theta c &= \theta(x_1^T A x_1 + b_1^T x_1 + c) - \theta(1-\theta) x_1^T A x_1 \\
&\le -\theta(1-\theta) x_1^T A x_1
\end{aligned}$$

and we can get a similar inequality for x_2. Hence,

$$\begin{aligned}
f(x) &\le -\theta(1-\theta)(x_1^T A x_1 - 2x_1^T A x_2 + x_2^T A x_2) \\
&= -\theta(1-\theta)(x_1 - x_2)^T A (x_1 - x_2) \le 0
\end{aligned}$$

as $A \succeq 0$. Hence, C is convex.

(b) Put $H = \{x : g^T x + h = 0\}$, $B = A + \lambda g g^T$ and

$$C_B = \{x \in \mathbb{R}^n : x^T B x + b^T x + c - \lambda h^2 \le 0\}.$$

By (a), C_B is convex and so does $C_B \cap H$. Suppose $x \in H$, then $x^T B x = x^T A x + \lambda h^2$. Therefore, $C_B \cap H = C$. Thus, C is convex.

□