Exercise on Convex Optimization

Songpeng Zu

Start at 2016-01-15

Contents

Abstract

This material is personal answers for the exercise on convex optimization wirtten by stephen boyd and lieven vandenberghe. I only put the main skeletons here.

1 Convex sets

1.1 Definition of convexity

1. $\sum_{i=1}^{k} \theta_k x_k = (\sum_{i=1}^{k-1} \theta_i) \sum_{i=1}^{k-1} \frac{\theta_i}{\sum_{i=1}^{k-1} \theta_i} x_i + \theta_k x_k$ in which, $\sum_{i=1}^{k-1} \frac{\theta_i}{\sum_{i=1}^{k-1} \theta_i} x_i \in \rfloor$ acording to the mathmatical induction.

2. Both are direct, since θ x + (1-θ)y $\in$ a line.

3. We can construct a sequence by the binary search [1], then for any $\theta > 0$, θ x + (1 - θ) y is either in the sequence or the limitation of the sequence.

[1] The binary search means that for two given points x and y, if we want to find a point between them, we can firstly use their middle point. If the middle point is not, then the point must lie in either [x, middle point] or [middle point, y] (we assume that x < y). Then we can follow the same procedure with x or y replaced with the middle point. Then the point will be in found by our procedure, or be the limitation of this procedure, which means we can approximate it as closely as we want.

Note a closed set in the Euclean Space is tight, which means the limitation of the sequence (all the points in the sequences belong to the closed set) is in the set.

4. The affine hull is the minumum affine set that contains $\mathcal{C}$, and the same as the intersection of all the affine sets that contain $\mathcal{C}$.

1.2 Examples

1. |b_1 - b_2|

2. When $a = c\tilde{a}, c \neq 0, d \leq \tilde{d}$, the first halfspace belongs to the second one. When $(a, b) = c(\tilde{a}, \tilde{b}), c \neq 0$, they are equal.

3. $||x - a||_2 \leq ||x - b||_2$ equals to $x^T x - 2a^T x + a^T a \leq x^T x - 2b^T x + b^T b$. Then we can get $(b - a)^T x \leq \frac{b^T b - a^T a}{2}$

4. (a) When both are zeros, yes; when one is zero, then it's a line segment, yes; when both are non-zeros, it forms a parallelogram (specially, $a_1 \perp a_2$, it's a square), so yes.
(b) Yes.
$$S = \{x | -\mathcal{I}x \preceq 0, \begin{bmatrix} 1^T \\ a^T \\ {a^2}^T \end{bmatrix} x = \begin{bmatrix} 1 \\ b_1 \\ b_2 \end{bmatrix}\} \tag{1}$$
(c) Yes. $S = \{x | x \preceq 1, -x \preceq 0\}$. Note, y is on the unit ball. $x^T y$ is the line passing through origin with y as its normal. In fact, it involves the space.
(d) Yes, the same as above.

5. (a) From Exercise 7, we know that V is the intersection of all the halfspaces, then it's polyhedron.
(b) We need to choose a interior point from S (S has nonempty interior). Then for each halfspace expression, we can reversely construct the corresponding equations x_i according to Exercise 7.
(c) Since no specific rules for choosing x_0. We can then say V_k is polyhedon, for k $\in$ 0, ..., K. For each pair of V_i and V_k (i is not equal to k), their intersection is the boundary between them. For any point in $\mathcal{R}$ it must be near to one point from $x_0, \ldots, x_K$. Then the union of V_k is the whole space. *The latter question.*

6. (a) If $A \succeq 0$, i.e., A is a symmetric positive semidefinite matrix. For any x_1, x_2, we have,
$$(x_1 - x_2)^T A(x_1 - x_2) = x_1^T A x_1 + x_2^T A x_2 - 2x_1^T A x_2 \geq 0$$
Then, we get
$$2x_1^T A x_2 \leq x_1^T A x_1 + x_2^T A x_2 \tag{2}$$

Then for any convex combination [2] of x_1 and x_2, y, we have:

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

So C is a convex set when $A \succeq 0$.
(b) For any convex combination of x_1 and x_2, the linear equation is trivial. For the quadratic inequality, we follow the same logic as (a). The different is, instead of use $x_1^T A x_2$ directly, here we use $x_1^T(A+\lambda gg^T)x_2 - \lambda x_1^T gg^T x_2$. The matrix $A + \lambda gg^T$ is positive semidefinite, and we can again use its character as (a). Also note:

$$x^T gg^T x = trace(g^T x x^T g) = trace(g^T x(g^T x)) = g^T x(g^T x)$$

Given the linear equation, we then have

$$x_1^T gg^T x_2 = g^T x_1(g^T x_2) = h^2$$

Then we can directly prove the quadratic inequality. *The latter question.*

7. Change the original inequality as $\log x_1 + \log x_2 \ge 0$, then use the general arithmetic-geometric mean inequality. For $n \ge 3$, we can also generalize the inequality based on the provement of the two-dimensional situation, which uses the $-\log$ function's convexisity.

8. (a) yes.
(b) yes.
(c) yes.
(d) yes. Suppose x_1, x_2 belongs to the defined set. Then from $||x - x_0||_2 \le ||x - y||_2$, we have that

$$-2x_1^T(x_0 - y) \le y^T y - x_0^T x_0 \tag{3}$$
$$-2x_2^T(x_0 - y) \le y^T y - x_0^T x_0 \tag{4}$$

Then if x is the convex combination of x_1 and x_2, we have

$$
\begin{aligned}
||x - x_0||_2^2 &= ||\theta x_1 + (1-\theta)x_2 - x_0||_2^2 \\
&= \theta^2 x_1^T x_1 + (1-\theta)^2 x_2^T x_2 + 2\theta(1-\theta)x_1^T x_2 + x_0^T x_0 - 2(\theta x_1^T + (1-\theta)x_2^T)x_0
\end{aligned} \tag{5}
$$

Replace x_0 with y, we can have similar formula. Compare the two formulas, only last two terms are different. Considering the previous inequalities, we can see x still satisfies the condition.
(e) yes. The procedure is similar with the one above.
Note we only need to prove x, the convex combination of x_1 and x_2, satisfies:

$$\inf_{x_0}(x_0^T x_0 - 2x^T x_0) \le y^T y - 2x^T y \tag{6}$$

[2] Here we define a convex combination is a mapping from the production of two same regular space to the same space with parameter $\theta \in [0,1]$: $f(x_1, x_2) = \theta x_1 + (1-\theta)x_2$

for any y in the T. And $x_0^T x_0 = \theta x_0^T x_0 + (1-\theta) x_0^T x_0$, x_1 and x_2 satisfies the inequality above.
(f) yes. For any y in S_2, x is the convex combination of x_1 and x_2, and x_1 and x_2 are in the defined set, we have
$x + y = \theta(x_1 + y) + (1-\theta)(x_2 + y), x_1 + y \in S_1, x_2 + y \in S_1$ Note S_1 is convex, so $x + y \in S_1$.
(g) yes. Follow the same procedure as (d).

9. $\{\theta_1 X_1 X_1^T + \theta_2 X_2 X_2^T | \theta_i \geq 0, X_i \in \mathcal{R}^{n \times k}, rank(X_i) = k, i = 1, 2\}$

10. (a) Follow the logic at **12(f)**
(b) Follow the logic above. Note S is convex, then the convex combination of points in S still in S.

11. (a) Yes. Expectation is a linear function for a discrete distribution p.
(b) Yes. $\Pr(x \geq \alpha) = \sum_{i=1}^n p_i \mathcal{I}_{a_i \geq \alpha}$, so it is a linear function for p.
(c) Yes, linear inequality for p.
(d) Yes, the same as above.
(e) Yes, the same as above.
(f) Yes, $Var(X) = EX^2 - (EX)^2$, linear inequality for p.
(g) Yes, the same as above.
(h) $Notsure.$
(i) Yes. for any β so that $\Pr(x \leq \beta) \geq 0.25$, we have that

$$\begin{aligned} Pr(x \leq \beta) &= \sum_{i=1}^n p_i \mathcal{I}_{a_i \leq \beta} \\ &= \sum_{i=1}^n \theta p_i^{(1)} \mathcal{I}_{a_i \leq \beta} + \sum_{i=1}^n (1-\theta) p_i^{(2)} \mathcal{I}_{a_i \leq \beta} \\ &\geq 0.25 \end{aligned} \tag{7}$$

So the inf of β on the convex combination of $p^{(1)}$ and $p^{(2)}$ must less than α.

1.3 Operations that preserve convexity

1. It is direct.

2. (a) $\{\frac{\sum_i^K \theta_i \nu_i}{\sum_i^K \theta_i t_i}; \theta_i \geq 0, \sum_i^K \theta_i = 1\}$
(b) When g is not zero, we have $\{\frac{g\nu}{h - f^T \nu}\}$
$(c)(d)$

3. $f^{-1}(x) = \mathcal{P} \circ \mathcal{Q}^{-1}(x, 1)$, in which $\mathcal{P}$ is the perspective function, and $x \in \{x | c^t(\mathcal{P} \circ (Q)^{-1}(x, 1)) + b > 0\}$

4. (a) $g^T \frac{Ax+b}{c^T x + d} \leq h$, then we can get
$(g^T A - hc^T)x \leq -g^T b + d$
(b) Following the same rule as (a), we can get
$(GA - hc^T)x \preceq -Gb + d$
(c) Again, we can get

$x^T(A^TP^{-1}A - cc^T)x + 2(b^TP^{-1}A - dc^T)x \leq d^2 - b^TP^{-1}b$ (d) We can also get
$\sum_{i=1}^n AxA_i \preceq (c^Tx + d)B - b\sum_{i=1}^n A_i$

2 Convex functions

3 Convex optimization problems

4 Duality

5 Approximation and fitting

6 Statistical estimation

7 Genomic problems