

Convex Optimization Note

Garrick Lin

March 13, 2019

1 Affine Set

1.1 Definition

A set \mathcal{A} is called affine set when it satisfied:

If $\mathbf{x}_1 \in \mathcal{A}$ and $\mathbf{x}_2 \in \mathcal{A}$, then, $\forall \theta \in \mathcal{R}$, $x = \theta\mathbf{x}_1 + (1 - \theta)\mathbf{x}_2$ also belong to \mathcal{A} .

The expression $\theta\mathbf{x}_1 + (1 - \theta)\mathbf{x}_2$ represent a line that cross through \mathbf{x}_1 and \mathbf{x}_2 . Hence, affine set can be explained intuitively as:

Affine set is a set that contain the line which cross through any two point within this set.

1.2 Properties

Assume a affine set $\mathcal{A} \subseteq \mathcal{R}^n$. If \mathcal{A} contain original point($\mathbf{0}$), then, \mathcal{A} is a subspace. In order to prove this, \mathcal{A} must satisfied following three rules:

1. \mathcal{A} contain original point.
2. If $\mathbf{v} \in \mathcal{A}$, then, $\forall \theta \in \mathcal{R}$, $\theta\mathbf{v} \in \mathcal{A}$.
3. If $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{A}$, then, $\mathbf{v}_1 + \mathbf{v}_2 \in \mathcal{A}$.

The first rule is already satisfied. For the second rule, $\forall \theta \in \mathcal{R}$, $\theta\mathbf{v}$ satisfied:

$$\theta\mathbf{v} = \theta\mathbf{v} + (1 - \theta)\mathbf{0}$$

Since \mathcal{A} is affine and $\mathbf{v}, \mathbf{0} \in \mathcal{A}$, $\theta\mathbf{v} \in \mathcal{A}$ always true. For the third rule, there is:

$$\mathbf{v}_1 + \mathbf{v}_2 = 2(\frac{1}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2)$$

$\frac{1}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2$ is in \mathcal{A} as \mathcal{A} is affine and $\frac{1}{2}\mathbf{v}_1, \frac{1}{2}\mathbf{v}_2 \in \mathcal{A}$ (According to the second rule, $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{A}$, then, $\frac{1}{2}\mathbf{v}_1, \frac{1}{2}\mathbf{v}_2 \in \mathcal{A}$). According to the second rule, $2(\frac{1}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2) \in \mathcal{A}$

Therefore, any affine set \mathcal{A}' can be thought as a subspace \mathcal{V} with a transportation \mathbf{v} , which means, $\mathcal{A}' = \mathcal{V} + \mathbf{v}$. On the other hand, an affine set subtract some constant vector ($\mathcal{A}' - \mathbf{v}$), which make the new set contain original point, can form a subspace. Formally, for a affine set $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ and $\mathbf{a}_i \in \mathcal{A}$, the set:

$$\mathcal{A} - \mathbf{a}_i = \{\mathbf{a}_1 - \mathbf{a}_i, \mathbf{a}_2 - \mathbf{a}_i, \dots, \mathbf{a}_{i-1} - \mathbf{a}_i, \mathbf{0}, \mathbf{a}_{i+1} - \mathbf{a}_i, \dots, \mathbf{a}_n - \mathbf{a}_i\} \quad (1)$$

form a subspace. Furthermore, any affine set \mathcal{A} has the form:

$$\mathcal{A} = \{\mathbf{x} | \mathcal{B}\mathbf{x} = \mathbf{c}\} \quad (2)$$

Proof: According to equation 1, $\mathcal{A} - \mathbf{a}_i$ is a subspace. Denote this subspace as $\mathcal{L} = \mathcal{A} - \mathbf{a}_i$. Assume the space that perpendicular to \mathcal{L} , denote as \mathcal{L}^\perp , has basis $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$. Then, \mathcal{L} is the set of vectors that perpendicular to the basis of \mathcal{L}^\perp . As a result, if we write the basis of \mathcal{L}^\perp in matrix form, which is:

$$\mathcal{B} = \begin{bmatrix} \mathbf{b}_1^T \\ \mathbf{b}_2^T \\ \vdots \\ \mathbf{b}_n^T \end{bmatrix}$$

We have:

$$\mathcal{L} = \{\mathbf{y} | \mathcal{B}\mathbf{y} = 0\} \quad (3)$$

Equation 3 always hold as all vector in \mathcal{L}^\perp is the linear combination of the basis, hence, if a vector \mathbf{v} satisfied $\mathbf{b}_1^T \mathbf{v} = 0, \mathbf{b}_2^T \mathbf{v} = 0, \dots, \mathbf{b}_n^T \mathbf{v} = 0$, then, $(\theta_1 \mathbf{b}_1 + \theta_2 \mathbf{b}_2 + \dots \theta_n \mathbf{b}_n)^T \mathbf{v} = 0$, which means, \mathbf{v} is perpendicular to any vector in \mathcal{L}^\perp if \mathbf{v} is perpendicular to the basis of \mathcal{L}^\perp .

Since the vector in \mathcal{L} come from \mathcal{A} subtract \mathbf{a}_i , for all $\mathbf{y} \in \mathcal{L}$, there must exists $\mathbf{x} \in \mathcal{A}$ such that:

$$\mathbf{y} = \mathbf{x} - \mathbf{a}_i \quad (4)$$

By combining equation 4 and equation 3, we have:

$$\mathcal{A} = \{\mathbf{x} | \mathcal{B}(\mathbf{x} - \mathbf{a}_i) = 0\}$$

Which is equivalent to :

$$\mathcal{A} = \{\mathbf{x} | \mathcal{B}\mathbf{x} = \mathcal{B}\mathbf{a}_i\}$$

and has the same form as equation 2, where $\mathbf{c} = \mathcal{B}\mathbf{a}_i$.

2 Hyper-plane

2.1 Definition

In a \mathcal{R}^n space, the hyperplane is defined as a $n-1$ dimension subspace plus a translate \mathbf{x}_0 , which is perpendicular to a one dimension vector \mathbf{w} . The formal mathematical expression of hyper-plane \mathcal{H} is:

$$\mathcal{H} = \{\mathbf{y} + \mathbf{x}_0 | \mathbf{w}^T \mathbf{y} = 0\} \quad (5)$$

We can denote $\mathbf{x} = \mathbf{y} + \mathbf{x}_0$, hence, equation 5 become:

$$\mathcal{H} = \{\mathbf{x} | \mathbf{w}^T (\mathbf{x} - \mathbf{x}_0) = 0\} \quad (6)$$

It is clearly that when $\mathbf{x} = \mathbf{x}_0$, $\mathbf{w}^T (\mathbf{x}_0 - \mathbf{x}_0) = 0$ always hold, hence, \mathbf{x}_0 is a point within hyper-plane. As a result, equation 6 can be explained intuitively as: Hyper-plane is a set of any point \mathbf{x} which satisfied the vector $\mathbf{x} - \mathbf{x}_0$ (start from a fixed point \mathbf{x}_0 which belong to this hyper-plane, end at \mathbf{x}) is perpendicular to a normal vector \mathbf{w} .

Equation 6 give us another way to represent hyper-plane:

$$\mathcal{H} = \{\mathbf{x} | \mathbf{w}^T \mathbf{x} = \mathbf{b}\} \quad (7)$$

Where $\mathbf{b} = \mathbf{w}^T \mathbf{x}_0$.

2.2 Half-space

Intuitively, a hyper-plane split the whole space into two part, each of them is called half-space. The formal mathematical expression for half-space, base on equation 6, is:

$$\{\mathbf{x} | \mathbf{w}^T(\mathbf{x} - \mathbf{x}_0) < 0\}$$

or

$$\{\mathbf{x} | \mathbf{w}^T(\mathbf{x} - \mathbf{x}_0) > 0\}$$

Which corresponding to the two split space. In geometric, two vector \mathbf{v}_1 and \mathbf{v}_2 satisfied: $\mathbf{v}_1^T \mathbf{v}_2 > 0$ if their angle smaller than 90 degree, otherwise, $\mathbf{v}_1^T \mathbf{v}_2 < 0$. As a result, the half-space $\{\mathbf{x} | \mathbf{w}^T(\mathbf{x} - \mathbf{x}_0) < 0\}$ is the set of point \mathbf{x} which satisfied the vector $\mathbf{x} - \mathbf{x}_0$ have a angle greater than 90 degree with a normal vector \mathbf{w} . The half-space $\{\mathbf{x} | \mathbf{w}^T(\mathbf{x} - \mathbf{x}_0) > 0\}$ can be explained in the same way except the angle is smaller than 90 degree.

Furthermore, according to 7, the expression of half-space can also be written as:

$$\{\mathbf{x} | \mathbf{w}^T \mathbf{x} < \mathbf{b}\}$$

Where $\mathbf{b} = \mathbf{w}^T \mathbf{x}_0$.

3 Cone

3.1 Definition

A cone \mathcal{K} is the set of any point \mathbf{x} that satisfied: If $\mathbf{x} \in \mathcal{K}$, then, for any $\theta \geq 0$, $\theta \mathbf{x} \in \mathcal{K}$. A **convex** cone, as its' name suggest, is a cone which is also a convex set.

3.2 Property of Convex Cone

Assume a cone \mathcal{K} is a convex cone, then, it satisfied:

If $\mathbf{x}_1 \in \mathcal{K}$ and $\mathbf{x}_2 \in \mathcal{K}$, then, for any $0 \leq \theta \leq 1$, $\theta \mathbf{x}_1 + (1 - \theta) \mathbf{x}_2 \in \mathcal{K}$.

Furthermore, for any $p, q > 0$, the expression $p\mathbf{x}_1 + q\mathbf{x}_2$ can be written as:

$$p\mathbf{x}_1 + q\mathbf{x}_2 = (p + q) \left(\frac{p}{p+q} \mathbf{x}_1 + \frac{q}{p+q} \mathbf{x}_2 \right) \quad (8)$$

Following the property of convex set, the point $(\frac{p}{p+q} \mathbf{x}_1 + \frac{q}{p+q} \mathbf{x}_2) \in \mathcal{K}$. Additionally, since $(p + q) \in \mathcal{R}$ and $p + q > 0$, the definition of cone indicate $(p + q)(\frac{p}{p+q} \mathbf{x}_1 + \frac{q}{p+q} \mathbf{x}_2) \in \mathcal{K}$. As a result, it can be seen within a convex cone \mathcal{K} :

If $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{K}$, then, for any $\theta_1, \theta_2 > 0$, $\theta_1 \mathbf{x}_1 + \theta_2 \mathbf{x}_2 \in \mathcal{K}$

3.3 Dual Cone

In geometric, a point \mathbf{x} can be defined as the end point of a vector that start with the original point($\mathbf{0}$) and terminate at point \mathbf{x} , which means, any point \mathbf{x} can also be thought as a vector $\mathbf{x} - \mathbf{0}$. Base on this, the dual cone of a cone \mathcal{K} , denote as \mathcal{K}^* , is defined as the set of any point λ which satisfied $\lambda^T \mathbf{x} > 0$ for all $\mathbf{x} \in \mathcal{K}$, which means, the **vector** λ have the angle smaller than 90 degree for any **vector** $\mathbf{x} - \mathbf{0}$ defined by the point \mathbf{x} in \mathcal{K} .

Furthermore, the expression $\lambda^T \mathbf{x} > 0$ define a half-space according to equation 6, where the fixed point \mathbf{x}_0 is the original point. As a result, the dual cone of \mathcal{K} can also be illustrated as:

A point λ is in \mathcal{K}^* if the half-space with the **vector** λ as normal vector, which expressed as $\lambda \mathbf{x} > 0$, contain the original cone \mathcal{K} .

3.4 Generalized Inequality

We can define the generalized inequality based on the concept of cone:

\mathbf{x} is smaller than \mathbf{y} respect to a specific cone \mathcal{K} if $\mathbf{y} - \mathbf{x} \in \mathcal{K}$. Denote as: $\mathbf{x} \preceq_{\mathcal{K}} \mathbf{y}$

Start from the definition of generalized inequality, we can define **minimum** and **minimal** value of a set \mathcal{S} respect to a specific cone \mathcal{K} :

The point $\mathbf{x}_0 \in \mathcal{S}$ is said to be the **minimum** of \mathcal{S} if for any other $\mathbf{y} \in \mathcal{S}$, the expression $\mathbf{x} \preceq_{\mathcal{K}} \mathbf{y}$ always holds.