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, $\mathbf{v_1} + \mathbf{v_2} \in \mathbf{A}$.

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 $\mathcal{A} = \{\mathbf{a_1}, \mathbf{a_2}, \cdots, \mathbf{a_n}\}$ and $\mathbf{a_i} \in \mathcal{A}$, the set:

$$\mathcal{A} - \mathbf{a_i} = \{\mathbf{a_1} - \mathbf{a_i}, \mathbf{a_2} - \mathbf{a_i}, \cdots, \mathbf{a_{i-1}}, \mathbf{0}, \mathbf{a_{i+1}}, \cdots, \mathbf{a_n} - \mathbf{a_i}\} \tag{1}$$

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

$$\mathcal{A} = \{\mathbf{x} |\, \mathcal{B}\mathbf{x} = \mathbf{c}\} \tag{2}$$

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 $\{\mathbf{b_1}, \mathbf{b_2}, \cdots, \mathbf{b_n}\}$. 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:

$$\mathcal{L} = \{\mathbf{y} |\, \mathcal{B}\mathbf{y} = 0\} \tag{3}$$

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 $\mathbf{b_1}^T\mathbf{v} = 0, \mathbf{b_2}^T\mathbf{v} = 0, \cdots, \mathbf{b_n}^T\mathbf{v} = 0$, then, $(\theta_1\mathbf{b_1} + \theta_2\mathbf{b_2} + \cdots \theta_n\mathbf{b_n})^T\mathbf{v} = 0$, 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:

$$\mathbf{y} = \mathbf{x} - \mathbf{a_i} \tag{4}$$

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

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:

$$\mathcal{H} = \{\mathbf{y} + \mathbf{x_0} |\, \mathbf{w}^T\mathbf{y} = 0\} \tag{5}$$

We can denote $\mathbf{x} = \mathbf{y} + \mathbf{x_0}$, hence, equation 5 become:

$$\mathcal{H} = \{\mathbf{x} |\, \mathbf{w}^T(\mathbf{x} - \mathbf{x_0}) = 0\} \tag{6}$$

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 $\mathbf{w}$.