$\{x|Ax=b\}$

2. Convex Set: Terms and Definitions

Definitions: I. Affine Set, II. Cone, and III. Convex Hull

Given $u_1, u_2, \cdots, u_k \in R^n$,

function $f(u,\theta) = \theta_1 u_1 + \theta_2 u_2 + \cdots + \theta_k u_k$

and two conditions

1. $\theta_1 + \theta_2 + \cdots + \theta_k = 1$
2. $\theta_i \geq 0\ \forall i$

I. $\{f(u,\theta)|$ condition 1$\}$: Affine set $Ax=b$

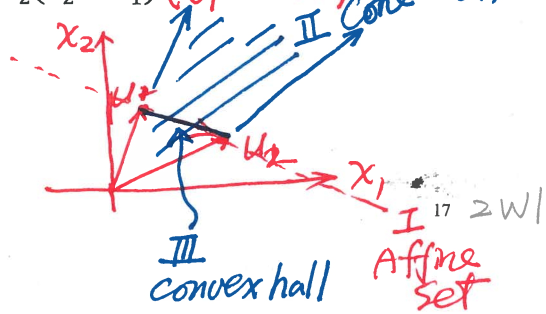
II. $\{f(u,\theta)|$ condition 2$\}$: Cone

III. $\{f(u,\theta)|$ conditions 1 and 2$\}$: Convex hull

$\theta_1 u_1 + \theta_2 u_2$ $\quad \theta_1, \theta_2 \geq 0$

$Ex1: \theta_1 u_1 + \theta_2 u_2 = u_1 + \theta_2(u_2 - u_1)$ $(\theta_1 = 1 - \theta_2)$

$Ex2: \theta_1 u_1 + \theta_2 u_2 + \theta_3 u_3$



2. Sets and Definitions: VI Hyperplane and Half Spaces

Hyperplane $\{x|\ a^T x = b\}, a \in R^n, b \in R$ $\quad a^T x \geq 0,\ a^T x \leq 0$

or $\{x|\ a^T(x - x_0) = 0\}$, for any $x_0 \in R^n, a \in R^n, b \in R$

Half Space $\{x|\ a^T x \leq b\}$ $\ a \in R^n, b \in R$

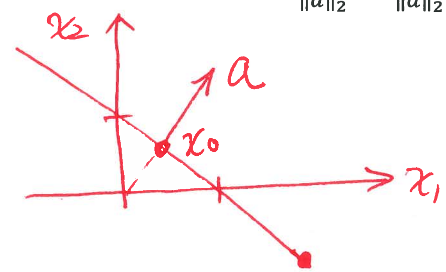
or $\{x|\ a^T(x - x_0) \leq 0\}$

Ex: $\{x|\ x_1 + x_2 = 1\}$ or $\{x|\ [1,1]\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}\right) = 0\}$

or $\{x|\ a^T(x - x_0) = 0\}$, $\ a^T = [1,1], b = 1, x_0 = [2,-1]$

For many applications, we standardize the expression:

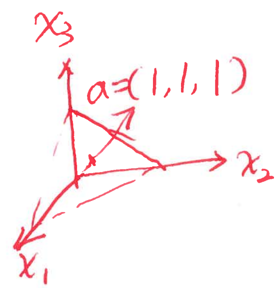
normalize the expression: $\frac{a^T}{\|a\|_2} x = \frac{b}{\|a\|_2}$



2. Sets and Definitions: Hyperplanes

Ex : 3 variables

$\{x|a^T x = b\}$, $a^T = (1,1,1)$, $b = 6$



Ex : 4 variables

$\{x|a^T x = b\}$, $a^T = (1,1,1,1)$, $b = 6$

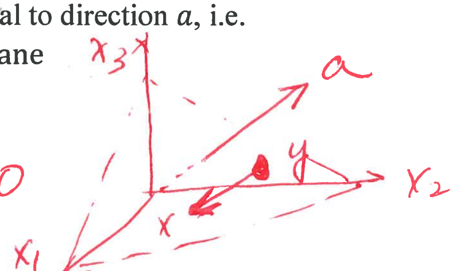
(1) degrees of freedom : $n - 1$ (R^n).

(2) all vectors $(x - y)$ are orthogonal to direction a, i.e.

$a^T(x - y) = 0$, $\forall x, y$ in the hyperplane

Proof: Given $a^T x = b$, $a^T y = b$

Then $a^T(x - y) = b - b = 0$



Exercise: Conceptually (visually) construct hyperplane.

2. Sets and Definitions: Hyperplanes

Hyperplane : as an Equal potential of cost function

$\min f_0(x) = c^T x$

$e.g.\ [1, 2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$\frac{\partial f_0(x)}{\partial x_1} = 1$

$\frac{\partial f_0(x)}{\partial x_2} = 2$

Vector c is the sensitivity or cost of vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

2. Sets and Definitions

Hyperplane : as a linearized constraint

$$a^T x \le b, x \in R^n$$

$$e.g.\ [1,2]\begin{bmatrix}x_1\\x_2\end{bmatrix} \le 10$$

Remark :

- Hyperplane is one key building block of convex optimization. (theory, algorithms, applications for machine learning, deep learning, …)
- Each hyperplane separates the space into two half spaces. whole space
- If $n \ge p$, p hyperplanes can separate the space into 2^p disjoint regions.

exponential function

2. Sets and Definitions

V. Polyhedra (plural) : Poly (many) Hedron (face)

$$P = \{x | Ax \le b, Cx = d\}$$

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \dots \\ a_m^T \end{bmatrix} \qquad C = \begin{bmatrix} c_1^T \\ c_2^T \\ \dots \\ c_p^T \end{bmatrix}$$

$\bigcap_{i=1}^{m}$ half space; $\bigcap_{j=1}^{p}$ hyperplane;

2. Sets and Definitions

VI. Matrix Space : Positive Semidefinite Cone

① $S^n = \{X \in R^{n\times n} | X = X^T\}$ (Symmetric Matrix the text book)

② $S_+^n = \{X \in S^n | X \succeq 0\}$ $\quad i.e.\ y^T X y \ge 0, \forall y$

$S_{++}^n = \{X \in S^n | X \succ 0\}$ $\quad i.e.\ y^T X y > 0, \forall y \ne 0$

notation

Ex: $X = \begin{bmatrix} x & y \\ y & z \end{bmatrix} \in S_+^2 \Leftrightarrow x \ge 0, z \ge 0, xz \ge y^2$

$$[a\ b] X \begin{bmatrix} a \\ b \end{bmatrix} = a^2 x + b^2 z + 2aby \ge 0, \forall a, b \in \mathbb{R}$$

$$R^T \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} R$$

$$R = \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix}$$

2. Sets and Definitions

$$R^{-1}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \tilde{a} \\ \tilde{b} \end{bmatrix}$$

Ex: $X = \begin{bmatrix} x & y \\ y & z \end{bmatrix} \in S_+^2 \Leftrightarrow x \ge 0, z \ge 0, xz \ge y^2$

$$[a\ b] X \begin{bmatrix} a \\ b \end{bmatrix} = a^2 x + b^2 z + 2aby \ge 0, \forall a, b \in \mathbb{R}$$

Proof : Let $R = \begin{bmatrix} 1 & -x^{-1}y \\ 0 & 1 \end{bmatrix}$

We have $[a\ b] X \begin{bmatrix} a \\ b \end{bmatrix} = [a\ b] R^{-T} R^T X R R^{-1} \begin{bmatrix} a \\ b \end{bmatrix}$

$$= \underline{[a\ b] R^{-T}} \begin{bmatrix} x & 0 \\ 0 & z - x^{-1}y^2 \end{bmatrix} \underline{R^{-1}\begin{bmatrix} a \\ b \end{bmatrix}}$$

$$\underbrace{\begin{bmatrix} 1 & 0 \\ -x^{-1}y & 1 \end{bmatrix}}_{R^T} \begin{bmatrix} x & y \\ y & z \end{bmatrix} \underbrace{\begin{bmatrix} 1 & -x^{-1}y \\ 0 & 1 \end{bmatrix}}_{R} = \begin{bmatrix} x & 0 \\ 0 & z - x^{-1}y^2 \end{bmatrix}$$

$$\begin{bmatrix} x & y \\ -y+y & -x^{-1}y^2+z \end{bmatrix} \begin{bmatrix} 1 & -x^{-1}y \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} x & -y+y \\ 0 & -x^{-1}y^2+z \end{bmatrix}$$

3. Separating Hyperplane

$\{x | a^T x = b\}$ (Classification, Optimization, Duality)

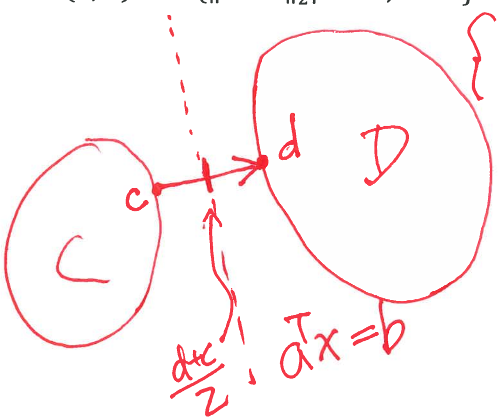
Theorem : Given two convex sets $C \cap D = \emptyset$ in R^n

$$\exists a \in R^n, b \in R, \;\; s.t. \; a^T x \le b, \forall x \in C$$
$$a^T x \ge b, \forall x \in D$$

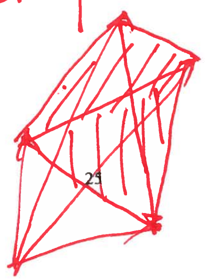
Actually, $a = d - c, b = \frac{\|d\|_2^2 - \|c\|_2^2}{2}$ x_0 $\frac{(d-c)^T(d+c)}{2}$

i.e. $f(x) = a^T x - b = (d-c)^T(x - \frac{d+c}{2})$

For $dist(C, D) = \inf\{\|u - v\|_2 | u \in C, v \in D\}$



$\{U\theta \mid 1^T\theta = 1, \theta \ge 0\}$



3. Separating Hyperplane

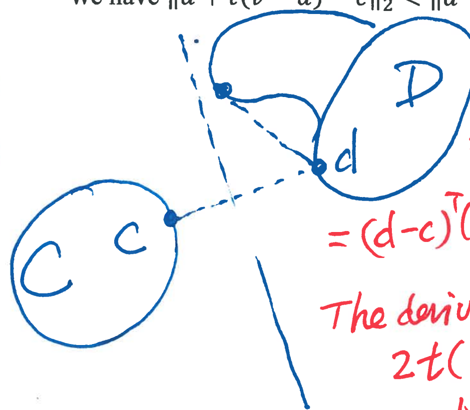
Proof : $\forall v \in D, a^T v \ge a^T d$ should be true

By contradiction, suppose that $a^T v < a^T d$

Then we can show that $d + t(v - d)$ is close to c for $t > 0$

Because $\frac{d}{dt}\|d + t(v-d) - c\|_2^2 = 2(d-c)^T(v-d) < 0$ $(t \to 0)$

We have $\|d + t(v-d) - c\|_2 < \|d - c\|_2$ for tiny $t > 0$



$\|d+t(v-d)-c\|_2^2$

$= \|d-c+t(v-d)\|_2^2$

$= (d-c)^T(d-c) + t^2(v-d)^T(v-d)$

$+ 2t(d-c)^T(v-d)$

The derivative

$2t(v-d)^T(v-d)$

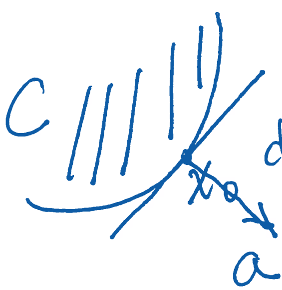
$+ 2(d-c)^T(v-d)$

3. Supporting Hyperplane

Given set $C \in R^n$, and a point x_0 on the boundary of C, the hyperplan $\{x | a^T x = a^T x_0\}$ is called supporting hyperplane of C if $a^T x \le a^T x_0, \forall x \in C$.

Supporting Hyperplane Theorem: For any nonempty convex set C, and a point x_0 on the boundary of C, There exists a support hyperplane to C at x_0.

Proof: A separating hyperplane that separates interior C and $\{x_0\}$ is a supporting hyperplane.



direction that obj decreases

4. Dual Cones

Given Cone K (i.e. $K = \{\sum_{i=1}^{k} \theta_i u_i \,|\theta_i \ge 0, u_i \in R^n, \forall i\}$)

$K^* = \{y | x^T y \ge 0, \forall x \in K\}$

Ex: 1. $K = R_+^n : K^* = R_+^n$

2. $K = S_+^n : K^* = S_+^n$

3. $K = \{(x,t) | \|x\|_2 \le t\} : K^* = \{(x,t) | \|x\|_2 \le t\}$

4. $K = \{(x,t) | \|x\|_1 \le t\} : K^* = \{(x,t) | \|x\|_\infty \le t\}$

Examples