

Numerical Optimization

Convex Sets

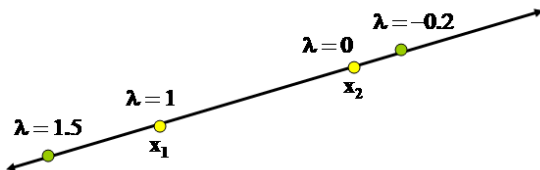
Shirish Shevade

Computer Science and Automation
Indian Institute of Science
Bangalore 560 012, India.

NPTEL Course on Numerical Optimization

Line and line segment

Let $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n, \mathbf{x}_1 \neq \mathbf{x}_2$.



Line passing through \mathbf{x}_1 and \mathbf{x}_2 :

$$\{\mathbf{y} \in \mathbb{R}^n : \mathbf{y} = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2, \quad \lambda \in \mathbb{R}\}$$

Line Segment, $LS[\mathbf{x}_1, \mathbf{x}_2]$:

$$\{\mathbf{y} \in \mathbb{R}^n : \mathbf{y} = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2, \quad \lambda \in [0, 1]\}$$

Affine sets

Definition

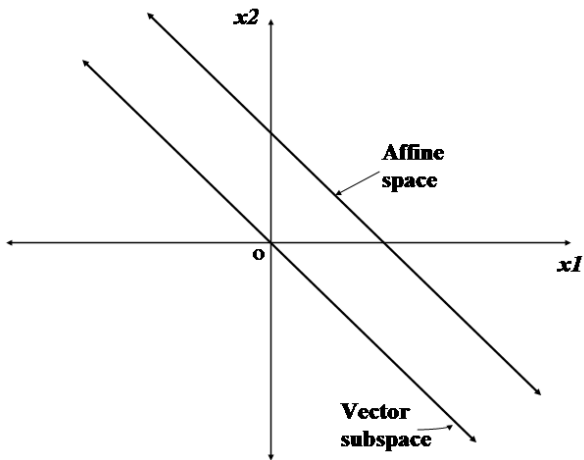
A set $X \subseteq \mathbb{R}^n$ is **affine** if for any $\mathbf{x}_1, \mathbf{x}_2 \in X$, and $\lambda \in \mathbb{R}$,

$$\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \in X.$$

If $X \subseteq \mathbb{R}^n$ is an affine set, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in X$ and $\sum_i \lambda_i = 1$, then the point $\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \dots + \lambda_k \mathbf{x}_k \in X$.

Examples

- Solution set of linear equations: $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b}\}$ where $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$.
- A subspace or a translated subspace



Result

If X is an affine set and $\mathbf{x}_0 \in X$, then $\{\mathbf{x} - \mathbf{x}_0 : \mathbf{x} \in X\}$ forms a subspace.

Proof.

Let $Y = \{\mathbf{x} - \mathbf{x}_0 : \mathbf{x} \in X\}$. To show that Y is a subspace, we need to show, $\alpha\mathbf{y}_1 + \beta\mathbf{y}_2 \in Y$ for any $\mathbf{y}_1, \mathbf{y}_2 \in Y$ and $\alpha, \beta \in \mathbb{R}$.

Let $\mathbf{y}_1, \mathbf{y}_2 \in Y$ and $\alpha, \beta \in \mathbb{R}$.

Therefore, $\mathbf{y}_1 + \mathbf{x}_0 \in X$ and $\mathbf{y}_2 + \mathbf{x}_0 \in X$.

$$\alpha\mathbf{y}_1 + \beta\mathbf{y}_2 = \alpha(\mathbf{y}_1 + \mathbf{x}_0) + \beta(\mathbf{y}_2 + \mathbf{x}_0) - (\alpha + \beta)\mathbf{x}_0.$$

$$\alpha\mathbf{y}_1 + \beta\mathbf{y}_2 + \mathbf{x}_0 = \alpha(\mathbf{y}_1 + \mathbf{x}_0) + \beta(\mathbf{y}_2 + \mathbf{x}_0) + (1 - \alpha - \beta)\mathbf{x}_0.$$

Since X is an affine set, $\alpha\mathbf{y}_1 + \beta\mathbf{y}_2 + \mathbf{x}_0 \in X$.

Thus, $\alpha\mathbf{y}_1 + \beta\mathbf{y}_2 \in Y \Rightarrow Y$ is a subspace.



Definition

Let $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$. A point \mathbf{x} is said to be an **affine combination** of points in X if

$$\mathbf{x} = \sum_{i=1}^k \lambda_i \mathbf{x}_i, \quad \sum_{i=1}^k \lambda_i = 1.$$

Definition

Let $X \subseteq \mathbb{R}^n$. The set of all affine combinations of points in X is called the **affine hull** of X .

$$\text{aff}(X) = \left\{ \sum_{i=1}^k \lambda_i \mathbf{x}_i : \mathbf{x}_1, \dots, \mathbf{x}_k \in X, \sum_i \lambda_i = 1 \right\}$$

Convex Sets

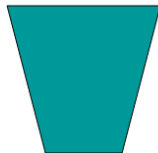
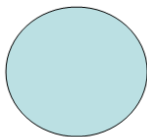
Definition

A set $C \subseteq \mathbb{R}^n$ is **convex** if for any $\mathbf{x}_1, \mathbf{x}_2 \in C$ and any scalar λ with $0 \leq \lambda \leq 1$, we have

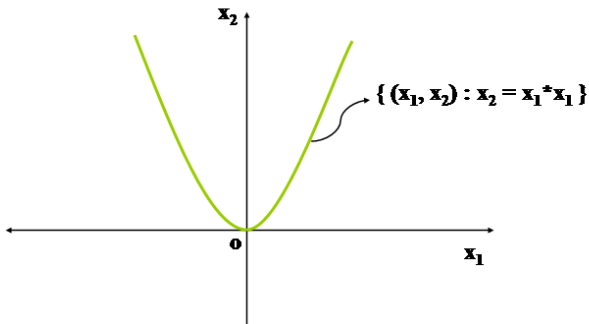
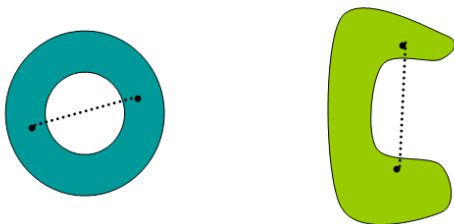
$$\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \in C.$$

Examples

- Some simple convex sets



Some examples of nonconvex sets



Convex sets

Examples:

- The empty set ϕ , any singleton set $\{\mathbf{x}_0\}$ and \mathbb{R}^n are convex subsets of \mathbb{R}^n .
- A closed ball in \mathbb{R}^n , $B[\mathbf{x}_0, r]$ and an open ball in \mathbb{R}^n , $B(\mathbf{x}_0, r)$, are convex sets.
- Any affine set is a convex set.
- A line segment is a convex set, but not an affine set.

- Possible to construct new convex sets using given convex sets

Definition

Let $P \subset \mathbb{R}^n$, $Q \subset \mathbb{R}^n$, $\alpha \in \mathbb{R}$.

- The *scalar multiple* αP of the set P is defined as

$$\alpha P = \{\mathbf{x} : \mathbf{x} = \alpha \mathbf{p}, \mathbf{p} \in P\}$$

- The *sum* of two sets P and Q is the set,

$$P + Q = \{\mathbf{x} : \mathbf{x} = \mathbf{p} + \mathbf{q}, \mathbf{p} \in P, \mathbf{q} \in Q\}.$$

Theorem

If $\{C_i\}, i \in \mathcal{A}$, is any collection of convex sets, then $\cap_{i \in \mathcal{A}} C_i$ is a convex set.

Proof.

Let $C = \cap_{i \in \mathcal{A}} C_i$.

Case 1. If C is empty or singleton, then clearly C is a convex set.

Case 2. To show that, for any $\mathbf{x}_1, \mathbf{x}_2 \in C$ and $\lambda \in [0, 1]$,

$$\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \in C.$$

Let $\mathbf{x}_1, \mathbf{x}_2 \in C$. Clearly, $\mathbf{x}_1, \mathbf{x}_2 \in C_i$ for every $i \in \mathcal{A}$.

Since every C_i is convex,

$$\begin{aligned} & \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \in C_i, \forall i \in \mathcal{A}, \lambda \in [0, 1] \\ \Rightarrow & \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \in \cap_{i \in \mathcal{A}} C_i, \forall \lambda \in [0, 1] \\ \Rightarrow & \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \in C, \forall \lambda \in [0, 1]. \end{aligned}$$



Theorem

If C_1 and C_2 are convex sets, then $C_1 + C_2$ is a convex set.

Proof.

Let $\mathbf{x}_1, \mathbf{y}_1 \in C_1$ and $\mathbf{x}_2, \mathbf{y}_2 \in C_2$.

So, $\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}_1 + \mathbf{y}_2 \in C_1 + C_2$. We need to show,

$$\lambda(\mathbf{x}_1 + \mathbf{x}_2) + (1 - \lambda)(\mathbf{y}_1 + \mathbf{y}_2) \in C_1 + C_2 \quad \forall \lambda \in [0, 1].$$

Since C_1 and C_2 are convex,

$$\mathbf{z}_1 = \lambda\mathbf{x}_1 + (1 - \lambda)\mathbf{y}_1 \in C_1 \quad \forall \lambda \in [0, 1], \text{ and}$$

$$\mathbf{z}_2 = \lambda\mathbf{x}_2 + (1 - \lambda)\mathbf{y}_2 \in C_2 \quad \forall \lambda \in [0, 1].$$

Thus, $\mathbf{z}_1 + \mathbf{z}_2 \in C_1 + C_2 \dots \dots (1)$

Now, $\mathbf{x}_1 + \mathbf{x}_2 \in C_1 + C_2$ and $\mathbf{y}_1 + \mathbf{y}_2 \in C_1 + C_2$.

Therefore, from (1),

$$\mathbf{z}_1 + \mathbf{z}_2 = \lambda(\mathbf{x}_1 + \mathbf{x}_2) + (1 - \lambda)(\mathbf{y}_1 + \mathbf{y}_2) \in C_1 + C_2 \quad \forall \lambda \in [0, 1].$$

Thus, $C_1 + C_2$ is a convex set. □

Theorem

If C is a convex set and $\alpha \in \mathbb{R}$, then αC is a convex set.

Proof.

Let $\mathbf{x}_1, \mathbf{x}_2 \in C$.

Since C is convex, $\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \in C \forall \lambda \in [0, 1]$.

Also, $\alpha \mathbf{x}_1, \alpha \mathbf{x}_2 \in \alpha C$.

Therefore, $\lambda(\alpha \mathbf{x}_1) + (1 - \lambda)(\alpha \mathbf{x}_2) = \alpha(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \in \alpha C$
for any $\lambda \in [0, 1]$.

Therefore, αC is a convex set. □

Hyperplane

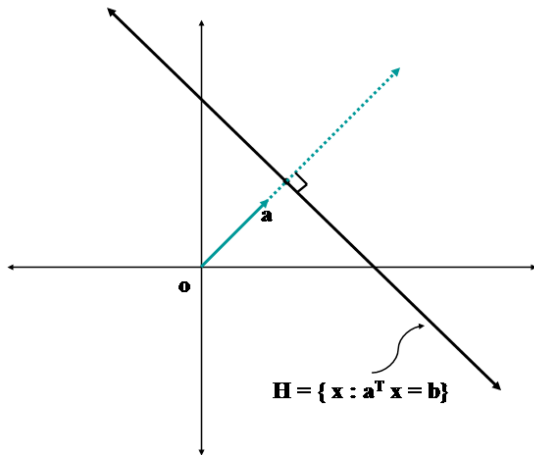
Definition

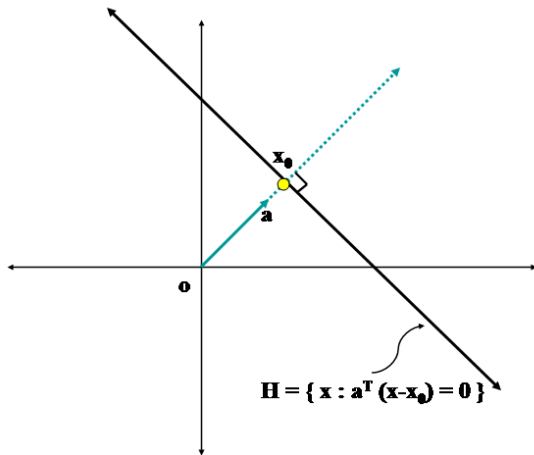
Let $b \in \mathbb{R}$ and $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{a} \neq 0$. Then, the set

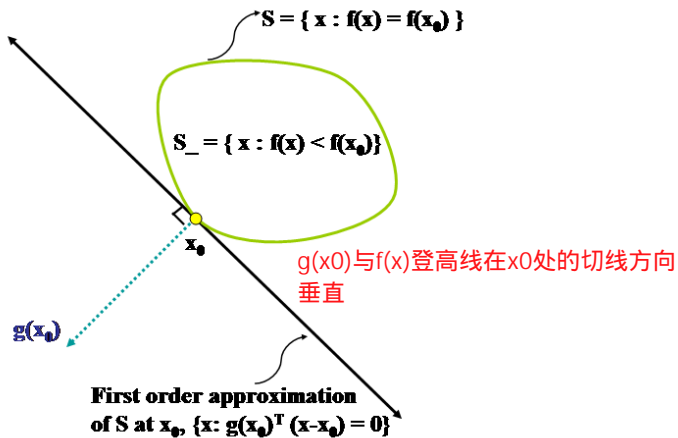
$$H = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} = b\}$$

is said to be a *hyperplane* in \mathbb{R}^n .

- \mathbf{a} denotes the normal to the hyperplane H .
- If $\|\mathbf{a}\| = 1$, then $|b|$ is the distance of H from the origin.
- In \mathbb{R}^2 , hyperplane is a line
- In \mathbb{R}^3 , hyperplane is a plane

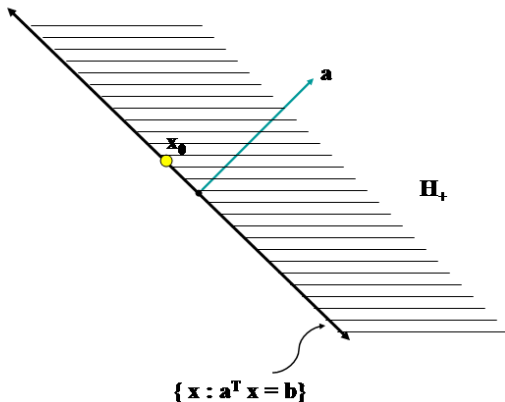


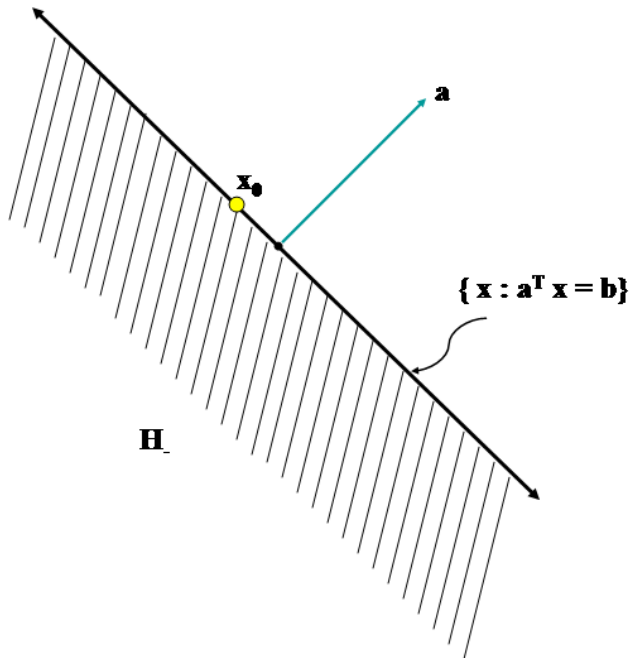




Half-spaces

- The sets $H_+ = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} \geq b\}$ and $H_- = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} \leq b\}$ are called closed positive and negative **half-spaces** generated by H .





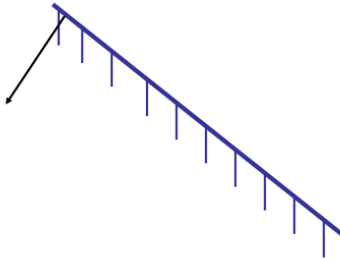
Some more examples of convex sets

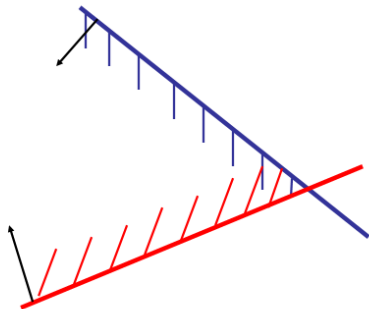
- $H = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} = b\}$ is a convex set
- Let $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m \in \mathbb{R}^n$ and $b_1, b_2, \dots, b_m \in \mathbb{R}$

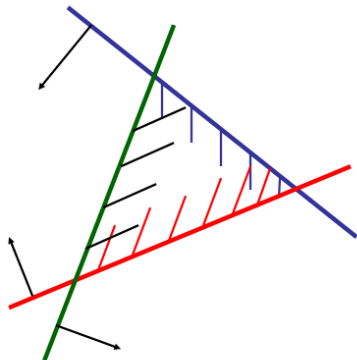
Define, $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$

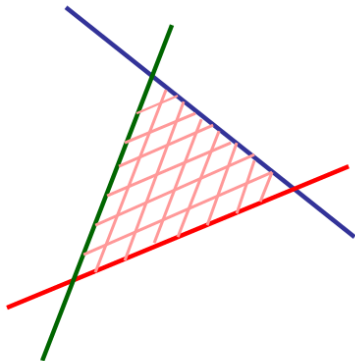
Then, $\{\mathbf{x} : \mathbf{Ax} = \mathbf{b}\}$ is a convex set.

- $H_+ = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} \geq b\}$ and $H_- = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} \leq b\}$ are convex sets.
- $\{\mathbf{x} : \mathbf{Ax} \leq \mathbf{b}\}$ and $\{\mathbf{x} : \mathbf{Ax} \geq \mathbf{b}\}$ are convex sets.









Convex Hull

Definition

The *convex hull* of a set S is the intersection of all the convex sets which contain S and is denoted by $\text{conv}(S)$.

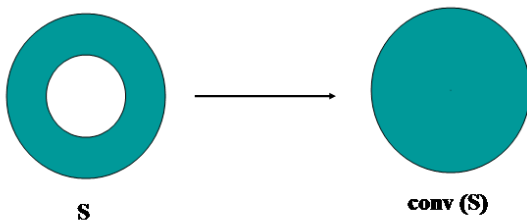
Note: Convex Hull of a set S is the convex set.

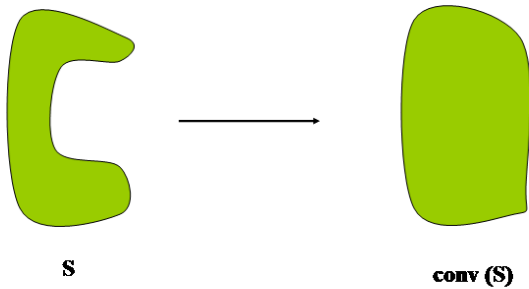
- The smallest convex set that contains S is called the *convex hull* of the set S .

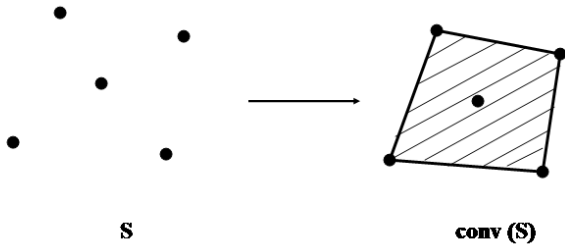
Examples:

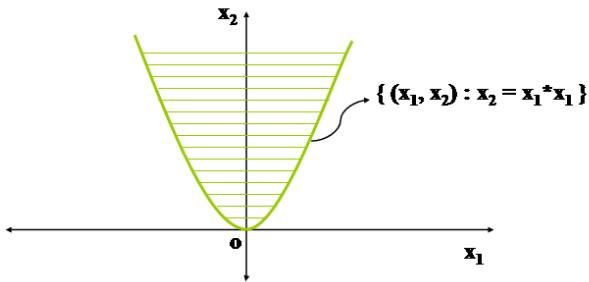
- $\text{conv}(\{x, y\}) = LS[x, y]$ where x and y are two points
- Let $S = \{(x, 0) : x \in \mathbb{R}\} \cup \{(0, y) : y \in \mathbb{R}\}$. Then $\text{conv}(S) = \mathbb{R}^2$.

conv hull of S 就是把 S 不convex的部分补全成convex的了







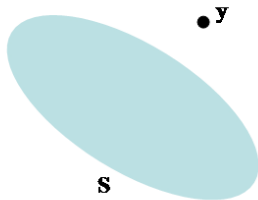


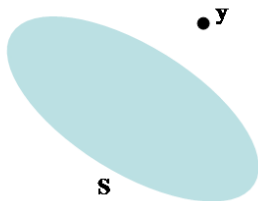
- $S = \{(x_1, x_2) : x_2 = x_1^2\}$
- $\text{conv}(S) = \{(x_1, x_2) : x_2 \geq x_1^2\}$

Theorem CS1

Let $S \subset \mathbb{R}^n$ be a nonempty, closed convex set and $\mathbf{y} \notin S$. Then there exists a unique point $\mathbf{x}_0 \in S$ with minimum distance from \mathbf{y} . Further, \mathbf{x}_0 is the minimizing point iff $(\mathbf{y} - \mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) \leq 0$ for all $\mathbf{x} \in S$.

$\mathbf{y} - \mathbf{x}_0$ 一定和 \mathbf{x}_0 处的切线垂直

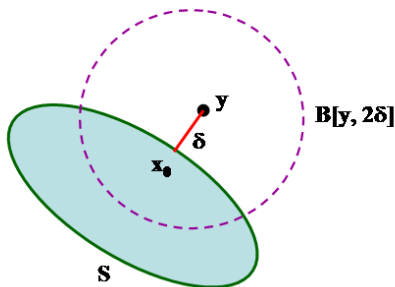




Proof Sketch

Given $S \subset \mathbb{R}^n$, a closed convex set and $y \notin S$.

- First show that there exists a minimizing point \mathbf{x}_0 in S that is closest to y .
- Let $\delta = \inf_{\mathbf{x} \in S} \|\mathbf{x} - y\|$.
- Note that S is not bounded.



Proof Sketch (continued)

- Consider the set $S \cap B[y, 2\delta]$.
- By Weierstrass' Theorem, there exists a minimizing point x_0 in S that is closest to y .

$$x_0 = \arg \min_{x \in S \cap B[y, 2\delta]} \|x - y\|$$

Proof Sketch (continued)

- Show uniqueness using triangle inequality

Assume that there exists $\hat{\mathbf{x}} \in S$ such that

$$\|\mathbf{y} - \mathbf{x}_0\| = \|\mathbf{y} - \hat{\mathbf{x}}\| = \delta.$$

Since S is convex, $\frac{1}{2}(\mathbf{x}_0 + \hat{\mathbf{x}}) \in S$.

Using triangle inequality,

$$2\left\|\mathbf{y} - \frac{(\mathbf{x}_0 + \hat{\mathbf{x}})}{2}\right\| \leq \|\mathbf{y} - \mathbf{x}_0\| + \|\mathbf{y} - \hat{\mathbf{x}}\| = 2\delta$$

We have a contradiction if strict inequality holds.

Proof Sketch (continued)

- To show that \mathbf{x}_0 is the unique minimizing point iff $(\mathbf{y} - \mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) \leq 0$ for all $\mathbf{x} \in S$.

Let $\mathbf{x} \in S$. Assume $(\mathbf{y} - \mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) \leq 0$.

$$\begin{aligned}\|\mathbf{y} - \mathbf{x}\|^2 &= \|\mathbf{y} - \mathbf{x}_0 + \mathbf{x}_0 - \mathbf{x}\|^2 \\ &= \|\mathbf{y} - \mathbf{x}_0\|^2 + \|\mathbf{x}_0 - \mathbf{x}\|^2 + 2(\mathbf{y} - \mathbf{x}_0)^T(\mathbf{x}_0 - \mathbf{x})\end{aligned}$$

- Using the assumption, $(\mathbf{y} - \mathbf{x}_0)^T(\mathbf{x}_0 - \mathbf{x}) \geq 0$, we get

$$\|\mathbf{y} - \mathbf{x}\|^2 \geq \|\mathbf{y} - \mathbf{x}_0\|^2$$

This implies that \mathbf{x}_0 is the minimizing point.

Proof Sketch (continued)

- Assume that \mathbf{x}_0 is the minimizing point, that is,

$$\|\mathbf{y} - \mathbf{x}_0\|^2 \leq \|\mathbf{y} - \mathbf{z}\|^2 \quad \forall \mathbf{z} \in S.$$

Consider any $\mathbf{x} \in S$. Since S is convex,

$$\lambda \mathbf{x} + (1 - \lambda) \mathbf{x}_0 \in S \quad \forall \lambda \in [0, 1].$$

Therefore, $\|\mathbf{y} - \mathbf{x}_0\|^2 \leq \|\mathbf{y} - \mathbf{x}_0 - \lambda(\mathbf{x} - \mathbf{x}_0)\|^2$

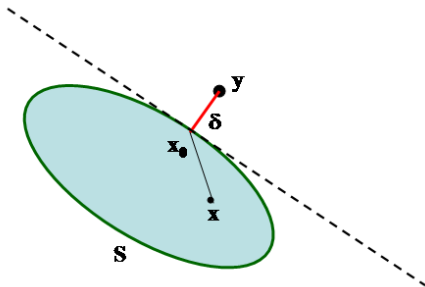
That is,

$$\begin{aligned} \|\mathbf{y} - \mathbf{x}_0\|^2 &\leq \|\mathbf{y} - \mathbf{x}_0\|^2 + \lambda^2 \|\mathbf{x} - \mathbf{x}_0\|^2 - 2\lambda(\mathbf{y} - \mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) \\ 2(\mathbf{y} - \mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) &\leq \lambda \|\mathbf{x} - \mathbf{x}_0\|^2 \end{aligned}$$

Letting $\lambda \rightarrow 0^+$, the result follows. □

Theorem CS1

Let $S \subset \mathbb{R}^n$ be a nonempty, closed convex set and $\mathbf{y} \notin S$. Then there exists a unique point $\mathbf{x}_0 \in S$ with minimum distance from \mathbf{y} . Further, \mathbf{x}_0 is the minimizing point iff $(\mathbf{y} - \mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) \leq 0$ for all $\mathbf{x} \in S$.

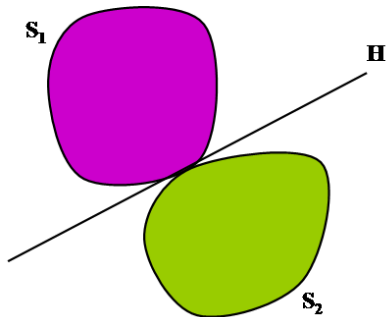


Separating hyperplanes

Definition

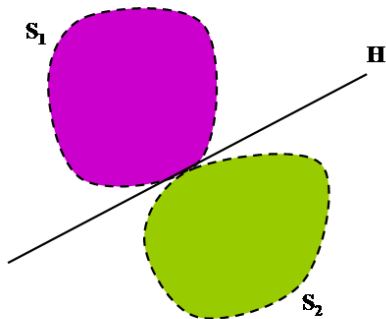
Let S_1 and S_2 be nonempty subsets in \mathbb{R}^n and $H = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} = b\}$ be a hyperplane.

- 1 The hyperplane H is said to *separate* S_1 and S_2 , if $\mathbf{a}^T \mathbf{x} \geq b \forall \mathbf{x} \in S_1$ and $\mathbf{a}^T \mathbf{x} \leq b \forall \mathbf{x} \in S_2$.
- 2 If $\mathbf{a}^T \mathbf{x} > b \forall \mathbf{x} \in S_1$ and $\mathbf{a}^T \mathbf{x} < b \forall \mathbf{x} \in S_2$, then the hyperplane H is said to *strictly separate* S_1 and S_2 .
- 3 The hyperplane H is said to *strongly separate* S_1 and S_2 if $\mathbf{a}^T \mathbf{x} \geq b + \epsilon \forall \mathbf{x} \in S_1$ and $\mathbf{a}^T \mathbf{x} \leq b \forall \mathbf{x} \in S_2$ where ϵ is a positive scalar.



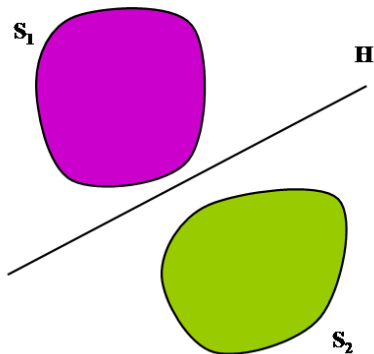
H : Separating hyperplane

The hyperplane $H = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} = b\}$ is said to *separate* S_1 and S_2 , if $\mathbf{a}^T \mathbf{x} \geq b \ \forall \ \mathbf{x} \in S_1$ and $\mathbf{a}^T \mathbf{x} \leq b \ \forall \ \mathbf{x} \in S_2$.



H : Strictly separating hyperplane

If $\mathbf{a}^T \mathbf{x} > b \forall \mathbf{x} \in S_1$ and $\mathbf{a}^T \mathbf{x} < b \forall \mathbf{x} \in S_2$, then the hyperplane $H = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} = b\}$ is said to *strictly separate* S_1 and S_2 .



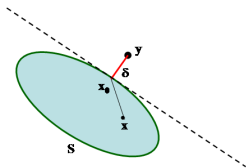
H : Strongly separating hyperplane

The hyperplane $H = \{\mathbf{x} : \mathbf{a}^T \mathbf{x} = b\}$ is said to *strongly separate* S_1 and S_2 if $\mathbf{a}^T \mathbf{x} \geq b + \epsilon \forall \mathbf{x} \in S_1$ and $\mathbf{a}^T \mathbf{x} \leq b \forall \mathbf{x} \in S_2$ where ϵ is a positive scalar.

Separation of a closed convex set and a point

Result

Let S be a nonempty closed convex set in \mathbb{R}^n and $\mathbf{y} \notin S$. Then there exists a nonzero vector \mathbf{a} and a scalar b such that $\mathbf{a}^T \mathbf{y} > b$ and $\mathbf{a}^T \mathbf{x} \leq b \forall \mathbf{x} \in S$.



Proof.

By Theorem CS1, there exists a unique minimizing point $\mathbf{x}_0 \in S$ such that $(\mathbf{x} - \mathbf{x}_0)^T (\mathbf{y} - \mathbf{x}_0) \leq 0$ for each $\mathbf{x} \in S$. Letting $\mathbf{a} = (\mathbf{y} - \mathbf{x}_0)$ and $b = \mathbf{a}^T \mathbf{x}_0$, we get $\mathbf{a}^T \mathbf{x} \leq b$ for each $\mathbf{x} \in S$ and $\mathbf{a}^T \mathbf{y} - b = (\mathbf{y} - \mathbf{x}_0)^T (\mathbf{y} - \mathbf{x}_0) > 0$ (since $\mathbf{y} \neq \mathbf{x}_0$). \square

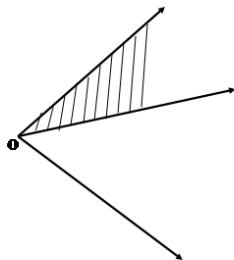
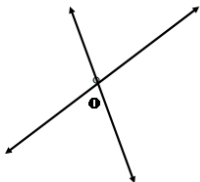
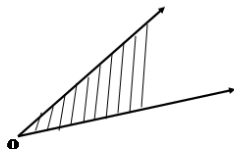
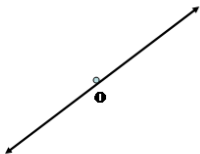
Cone

Definition

A set $K \subseteq \mathbb{R}^n$ is called a *cone* if for every $\mathbf{x} \in K$ and $\lambda \geq 0$, we have $\lambda \mathbf{x} \in K$.

- K is a convex cone if it is convex and a cone.

Some examples of cones



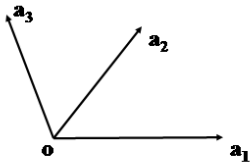
Farkas' Lemma

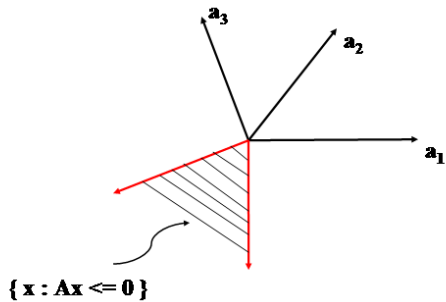
Farkas' Lemma

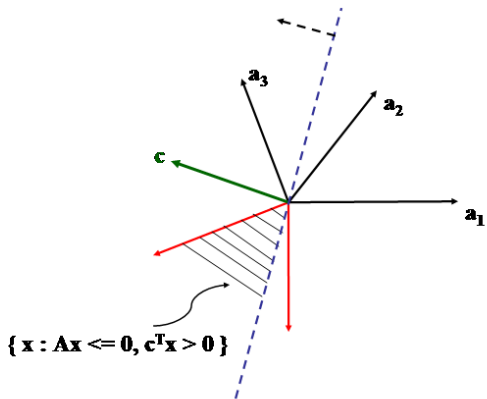
Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{c} \in \mathbb{R}^n$. Then, exactly one of the following two systems has a solution:

- (I) $\mathbf{Ax} \leq \mathbf{0}, \mathbf{c}^T \mathbf{x} > 0$ for some $\mathbf{x} \in \mathbb{R}^n$
- (II) $\mathbf{A}^T \mathbf{y} = \mathbf{c}, \mathbf{y} \geq \mathbf{0}$ for some $\mathbf{y} \in \mathbb{R}^m$.

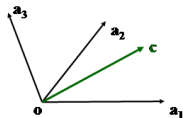
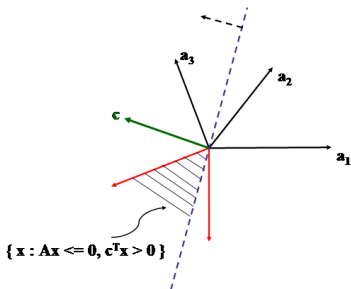
- Let $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m \in \mathbb{R}^n$.
- Define, $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{pmatrix}$.
- For example,







Farkas' Lemma (Geometrical Interpretation)



$$\mathbf{c} = \sum y_i \mathbf{a}_i, y_i \geq 0$$

Exactly one of the following two systems has a solution:

(I) $\mathbf{Ax} \leq \mathbf{0}, \mathbf{c}^T \mathbf{x} > 0$ for some $\mathbf{x} \in \mathbb{R}^n$

(II) $\mathbf{A}^T \mathbf{y} = \mathbf{c}, \mathbf{y} \geq \mathbf{0}$ for some $\mathbf{y} \in \mathbb{R}^m$

Farkas' Lemma

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{c} \in \mathbb{R}^n$. Then, exactly one of the following two systems has a solution:

- (I) $\mathbf{Ax} \leq \mathbf{0}, \mathbf{c}^T \mathbf{x} > 0$ for some $\mathbf{x} \in \mathbb{R}^n$
- (II) $\mathbf{A}^T \mathbf{y} = \mathbf{c}, \mathbf{y} \geq \mathbf{0}$ for some $\mathbf{y} \in \mathbb{R}^m$

Proof of Farkas' Lemma

(a) Suppose, system (II) has a solution.

Therefore, $\exists \mathbf{y} \geq \mathbf{0}$ such that $\mathbf{c} = \mathbf{A}^T \mathbf{y}$.

Consider $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{Ax} \leq \mathbf{0}$.

So, $\mathbf{c}^T \mathbf{x} = \mathbf{y}^T \mathbf{Ax} \leq 0$.

That is, System I has no solution.

Farkas' Lemma

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{c} \in \mathbb{R}^n$. Then, exactly one of the following two systems has a solution:

- (I) $\mathbf{Ax} \leq \mathbf{0}, \mathbf{c}^T \mathbf{x} > 0$ for some $\mathbf{x} \in \mathbb{R}^n$
- (II) $\mathbf{A}^T \mathbf{y} = \mathbf{c}, \mathbf{y} \geq \mathbf{0}$ for some $\mathbf{y} \in \mathbb{R}^m$

Proof of Farkas' Lemma (continued)

(b) Suppose, system (II) has no solution.

Let $S = \{\mathbf{x} : \mathbf{x} = \mathbf{A}^T \mathbf{y}, \mathbf{y} \geq \mathbf{0}\}$, a closed convex set and $\mathbf{c} \notin S$.

Therefore, $\exists \mathbf{p} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ such that

$$\mathbf{p}^T \mathbf{x} \leq \alpha \quad \forall \mathbf{x} \in S \text{ and } \mathbf{c}^T \mathbf{p} > \alpha.$$

This means, $\alpha \geq 0$ (Since $\mathbf{0} \in S$) $\Rightarrow \mathbf{c}^T \mathbf{p} > 0$.

Also, $\alpha \geq \mathbf{p}^T \mathbf{A}^T \mathbf{y} = \mathbf{y}^T \mathbf{A} \mathbf{p} \quad \forall \mathbf{y} \geq \mathbf{0}$.

Since $\mathbf{y} \geq \mathbf{0}, \mathbf{A} \mathbf{p} \leq \mathbf{0}$ (as \mathbf{y} can be made arbitrarily large).

Thus, $\exists \mathbf{p} \in \mathbb{R}^n$ such that $\mathbf{A} \mathbf{p} \leq \mathbf{0}, \mathbf{c}^T \mathbf{p} > 0 \Rightarrow$ System (I) has a solution.

Farkas' Lemma

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{c} \in \mathbb{R}^n$. Then, exactly one of the following two systems has a solution:

- (I) $\mathbf{Ax} \leq \mathbf{0}, \mathbf{c}^T \mathbf{x} > 0$ for some $\mathbf{x} \in \mathbb{R}^n$
- (II) $\mathbf{A}^T \mathbf{y} = \mathbf{c}, \mathbf{y} \geq \mathbf{0}$ for some $\mathbf{y} \in \mathbb{R}^m$.

Corollary

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Then exactly one of the following systems has a solution:

- (I) $\mathbf{Ax} < \mathbf{0}$ for some $\mathbf{x} \in \mathbb{R}^n$
- (II) $\mathbf{A}^T \mathbf{y} = \mathbf{0}, \mathbf{y} \geq \mathbf{0}$ for some nonzero $\mathbf{y} \in \mathbb{R}^m$.

Corollary

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Then exactly one of the following systems has a solution:

- (I) $\mathbf{Ax} < \mathbf{0}$ for some $\mathbf{x} \in \mathbb{R}^n$
- (II) $\mathbf{A}^T \mathbf{y} = \mathbf{0}, \mathbf{y} \geq \mathbf{0}$ for some nonzero $\mathbf{y} \in \mathbb{R}^m$.

Proof.

We can write system I as

$$\mathbf{Ax} + z\mathbf{e} \leq \mathbf{0} \text{ for some } \mathbf{x} \in \mathbb{R}^n, z > 0$$

where \mathbf{e} is a m -dimensional vector containing all 1's.

That is,

$$(\mathbf{A} \ \mathbf{e}) \begin{pmatrix} \mathbf{x} \\ z \end{pmatrix} \leq \mathbf{0}, \quad (0, \dots, 0, 1) \begin{pmatrix} \mathbf{x} \\ z \end{pmatrix} > 0$$

for some $(\mathbf{x} \ z)^T \in \mathbb{R}^{n+1}$.

Proof. (continued)

Recall Farkas' Lemma.

Therefore, System II is

$$\begin{pmatrix} \mathbf{A}^T \\ \mathbf{e}^T \end{pmatrix} \mathbf{y} = (0, \dots, 0, 1)^T, \mathbf{y} \geq \mathbf{0} \text{ for some } \mathbf{y} \in \mathbb{R}^m.$$

That is $\mathbf{A}^T \mathbf{y} = \mathbf{0}$, $\mathbf{e}^T \mathbf{y} = 1$ and $\mathbf{y} \geq \mathbf{0}$ for some $\mathbf{y} \in \mathbb{R}^m$.

Using Farkas' Lemma, the result follows.

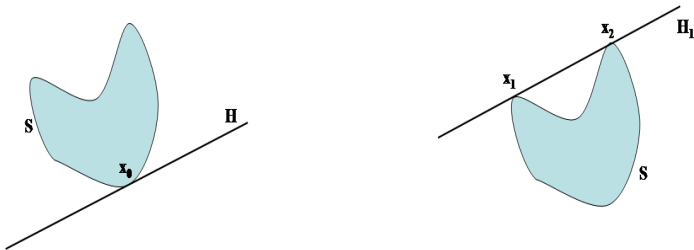
Supporting Hyperplanes of sets at boundary points

Definition

Let $S \subset \mathbb{R}^n$, $S \neq \emptyset$. Let \mathbf{x}_0 be a boundary point of S . A hyperplane $H = \{\mathbf{x} : \mathbf{a}^T(\mathbf{x} - \mathbf{x}_0) = 0\}$ is called a *supporting hyperplane* of S at \mathbf{x}_0 , if either

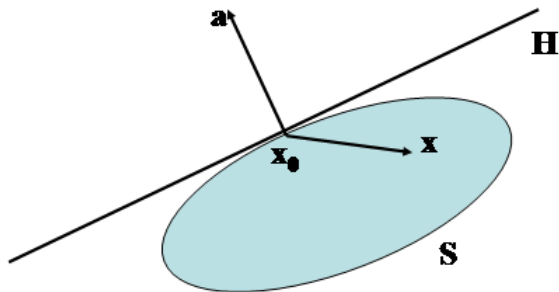
$S \subseteq H_+$ (that is, $\mathbf{a}^T(\mathbf{x} - \mathbf{x}_0) \geq 0 \forall \mathbf{x} \in S$), or

$S \subseteq H_-$ (that is, $\mathbf{a}^T(\mathbf{x} - \mathbf{x}_0) \leq 0 \forall \mathbf{x} \in S$).



Theorem

Let S be a nonempty convex set in \mathbb{R}^n and \mathbf{x}_0 be a boundary point of S . Then, there exists a hyperplane that supports S at \mathbf{x}_0 .



Theorem

Let S_1 and S_2 be two nonempty disjoint convex sets in \mathbb{R}^n . Then, there exists a hyperplane that separates S_1 and S_2 .

Proof.

Consider the set $S = S_1 - S_2 = \{\mathbf{x}_1 - \mathbf{x}_2 : \mathbf{x}_1 \in S_1, \mathbf{x}_2 \in S_2\}$.

Note that S is a convex set and $\mathbf{0} \notin S$.

Thus, we have a convex set and a point not in that set.

So, there exists a vector \mathbf{a} such that $\mathbf{a}^T \mathbf{x} \leq 0 \forall \mathbf{x} \in S$.

In other words, $\mathbf{a}^T \mathbf{x}_1 \leq \mathbf{a}^T \mathbf{x}_2 \forall \mathbf{x}_1 \in S_1, \mathbf{x}_2 \in S_2$.

