Quantum Foundations Lecture 3

February 5, 2018

Dr. Matthew Leifer
leifer@chapman.edu

HSC112

Carathéodory's Theorem for Convex Sets

- **Theorem:** Let X be a set of vectors in $\mathbb{R}^n$. Every $\boldsymbol{x} \in \mathrm{conv}(X)$ can be written as a convex combination of vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_m \in X$ with $m \leq n+1$.
- Proof: Let $S = \mathrm{conv}(X)$ and lift S to $\mathbb{R}^{n+1}$ to form a convex cone C.
- Clearly $C = \mathrm{cone}(X')$ where $X' = \{\boldsymbol{x}' | \boldsymbol{x} \in X\}$ and

$$\boldsymbol{x}_j' = \begin{pmatrix} 1 \\ \boldsymbol{x}_j \end{pmatrix}$$

Carathéodory's Theorem for Convex Sets

- Carathéodory's theorem for cones says that every $\boldsymbol{x}' \in C$ can be written as a positive combination of at most $n+1$ vectors in X'

$$\boldsymbol{x}' = \sum_{j=1}^{n+1} \alpha_j \boldsymbol{x}_j' \text{ with } \alpha_j \geq 0.$$

- If $\boldsymbol{x}'$ is in the embedding of S in C then its first component is 1. The first component of every $\boldsymbol{x}_j$ is also 1.
- Therefore $\sum_{j=1}^{n+1} \alpha_j = 1$ and this is a convex combination.

Krein-Milman theorem for convex sets

- **Theorem**: Every closed convex set S in $\mathbb{R}^n$ is the convex hull of its extreme points.
- Let $\mathrm{Ext}(S)$ be the set of extreme points of S. Then, the theorem says
$$S = \mathrm{conv}(\mathrm{Ext}(S))$$
- **Corollary:** Every $x \in S$ can be written as a positive combination of at most $n+1$ extreme points.

Linear and Affine Maps

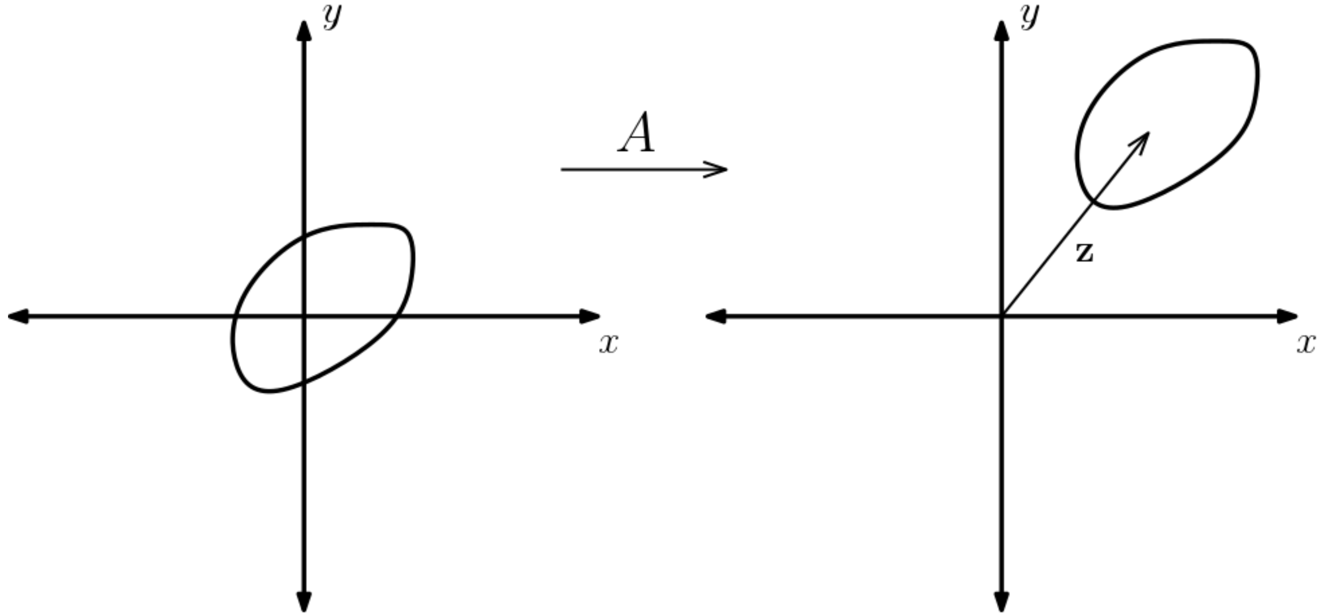
- For vector spaces, convex cones, and convex sets, there are natural classes of maps that preserve their structure.
- Let V be a vector space. A *linear map* $A: V \to V$ is a map that satisfies
$$A(a\boldsymbol{x} + b\boldsymbol{y}) = aA(\boldsymbol{x}) + bA(\boldsymbol{y})$$
for all $\boldsymbol{x}, \boldsymbol{y} \in V$ and scalars a, b.
- Let V be a real vector space. A *No Good Name For It (NGNFI) map* $A: V \to V$ is a map that satisfies
$$A(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \alpha A(\boldsymbol{x}) + \beta A(\boldsymbol{y})$$
for all $\boldsymbol{x}, \boldsymbol{y} \in V$ and $\alpha, \beta \geq 0$.
- Let V be a real vector space. An *affine map* $A: V \to V$ is a map that satisfies
$$A(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \alpha A(\boldsymbol{x}) + \beta A(\boldsymbol{y})$$
for all $\boldsymbol{x}, \boldsymbol{y} \in V$, $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$.

Linear and Affine Maps

- In other words, a NGNFI map is a maps that maps *all* convex cones to convex cones and an affine maps is a map that maps *all* convex sets to convex sets.
- The reason NGNFI maps have no good name is that they are linear maps in disguise.
  - We just define $A(-\alpha\boldsymbol{x}) = -\alpha A(\boldsymbol{x})$ and extend by linearity.
  - Carathéodory's theorem for cones guarantees that this still maps convex cones to convex cones, hence is still an NGNFI map.
- Therefore, we can dispense with NGNFI maps and just use linear maps. This is good news, as linear maps on $\mathbb{R}^n$ are just $n \times n$ matrices.

Linear and Affine Maps

- Life is not so simple for affine maps. A translation $A(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{z}$ for a fixed vector $\boldsymbol{z}$ is affine.



- This cannot be extended to a linear map because linear maps must satisfy $A(\mathbf{0}) = \mathbf{0}$, but here $A(\mathbf{0}) = \boldsymbol{z}$.

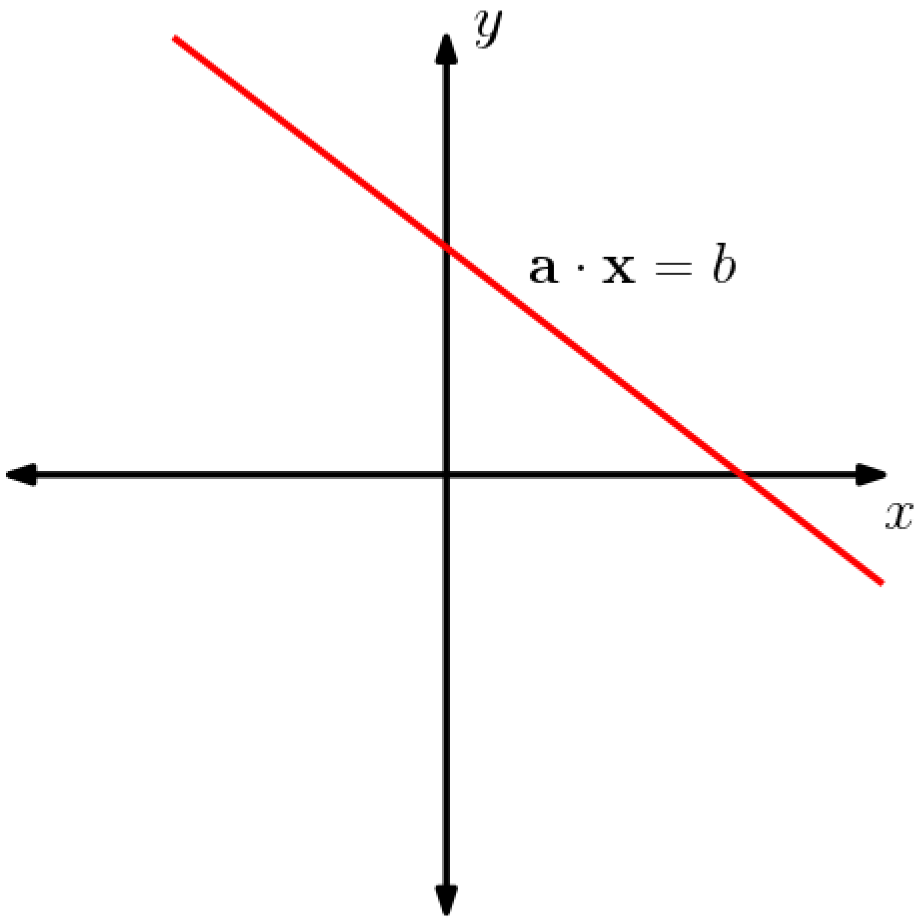
Lifting to the rescue!

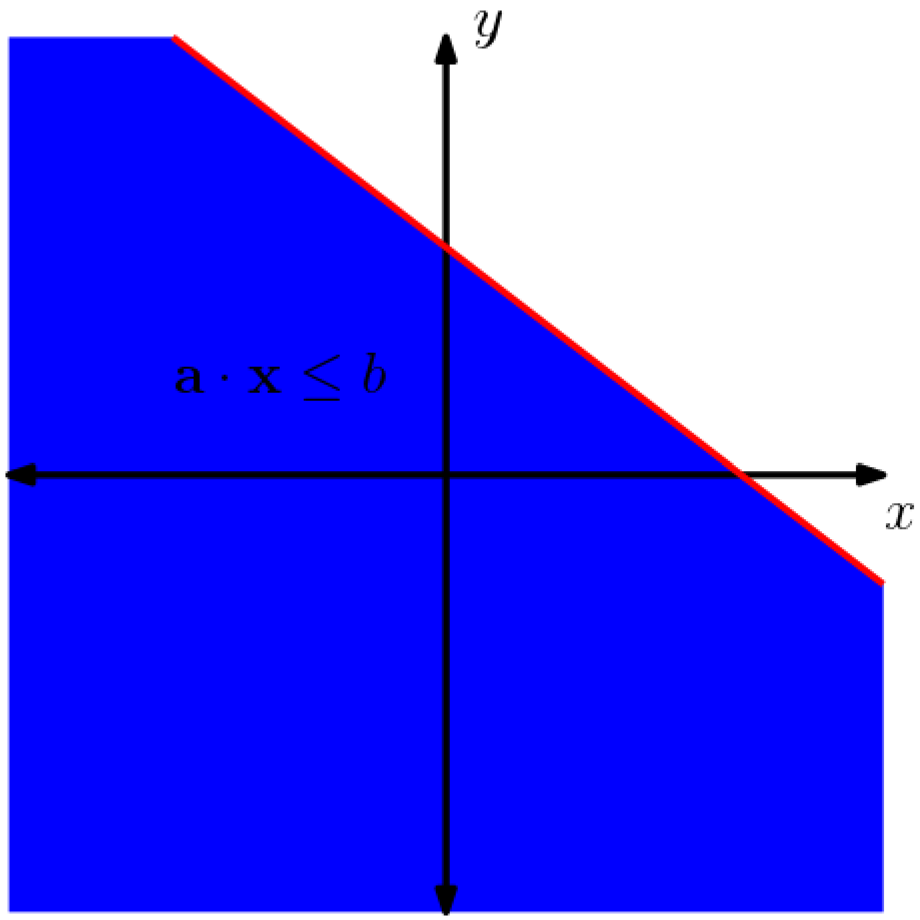
- Given a convex set S that we want to map to another convex set via an affine map A, construct the lifted cone C.
- Define $A'\left(\begin{pmatrix}1\\ \boldsymbol{x}\end{pmatrix}\right) = \begin{pmatrix}1\\ A(\boldsymbol{x})\end{pmatrix}$ for vectors with 1 as their first component.
- Every vector can be written as $a\boldsymbol{x}'$ for some scalar a and vector $\boldsymbol{x}'$ of this form. Defining $A'(a\boldsymbol{x}') = aA'(\boldsymbol{x}')$ then gives us a linear map as before.
- For this reason, it is almost always better to work in the lifted space because we only need to use linear maps.

Hyperplanes and Half-Spaces

- A *hyperplane* in $\mathbb{R}^n$ is a set of vectors $\boldsymbol{x}$ satisfying an equation of the form
$$\boldsymbol{a} \cdot \boldsymbol{x} = b$$
where $\boldsymbol{a}$ is a fixed vector and b is a scalar.
  - Generalizes the equation for a plane in 3d space.
- A (closed) *half-space* in a real vector space is a set of vectors $\boldsymbol{x}$ satisfying an inequality of the form
$$\boldsymbol{a} \cdot \boldsymbol{x} \leq b$$
where $\boldsymbol{a}$ is a fixed vector and b is a scalar.

Examples





Hyperplane Separation Theorem

- **Theorem**: Given any two disjoint closed convex sets, S and S', there always exists a hyperplane $\boldsymbol{a} \cdot \boldsymbol{x} = b$ such that

$$\boldsymbol{a} \cdot \boldsymbol{x} \leq b \text{ for all } \boldsymbol{x} \in S$$

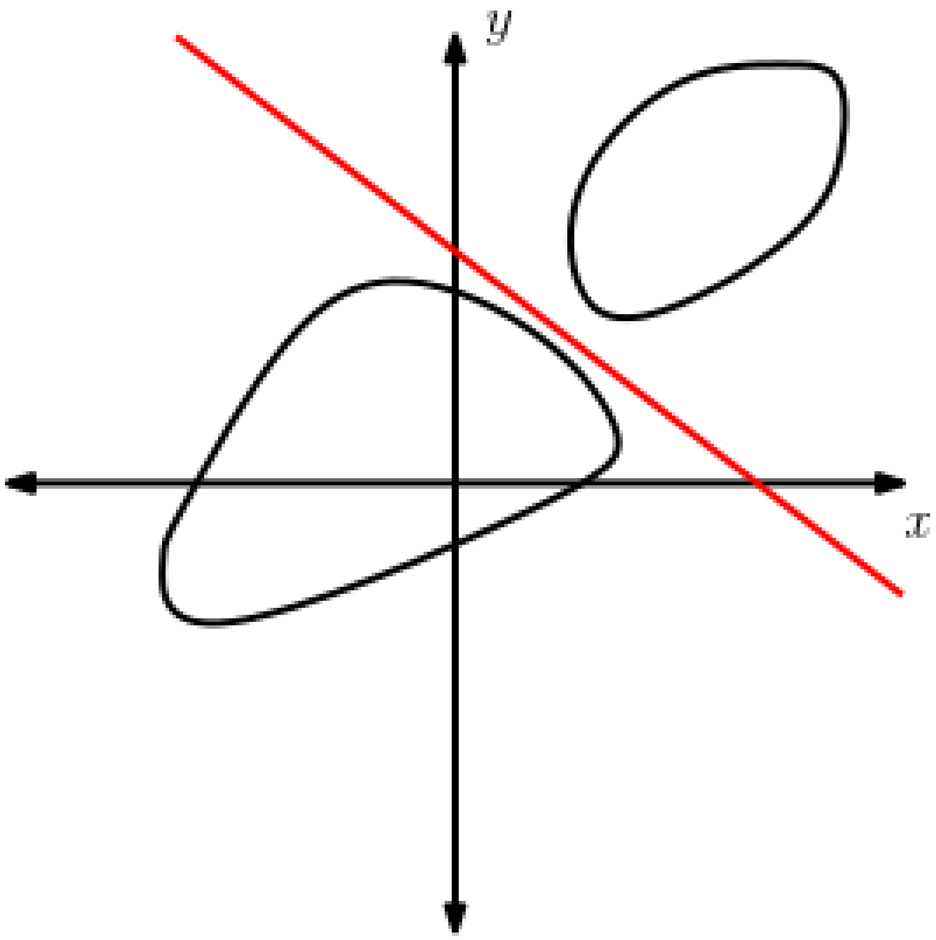
$$\boldsymbol{a} \cdot \boldsymbol{x}' > b \text{ for all } \boldsymbol{x}' \in S'$$

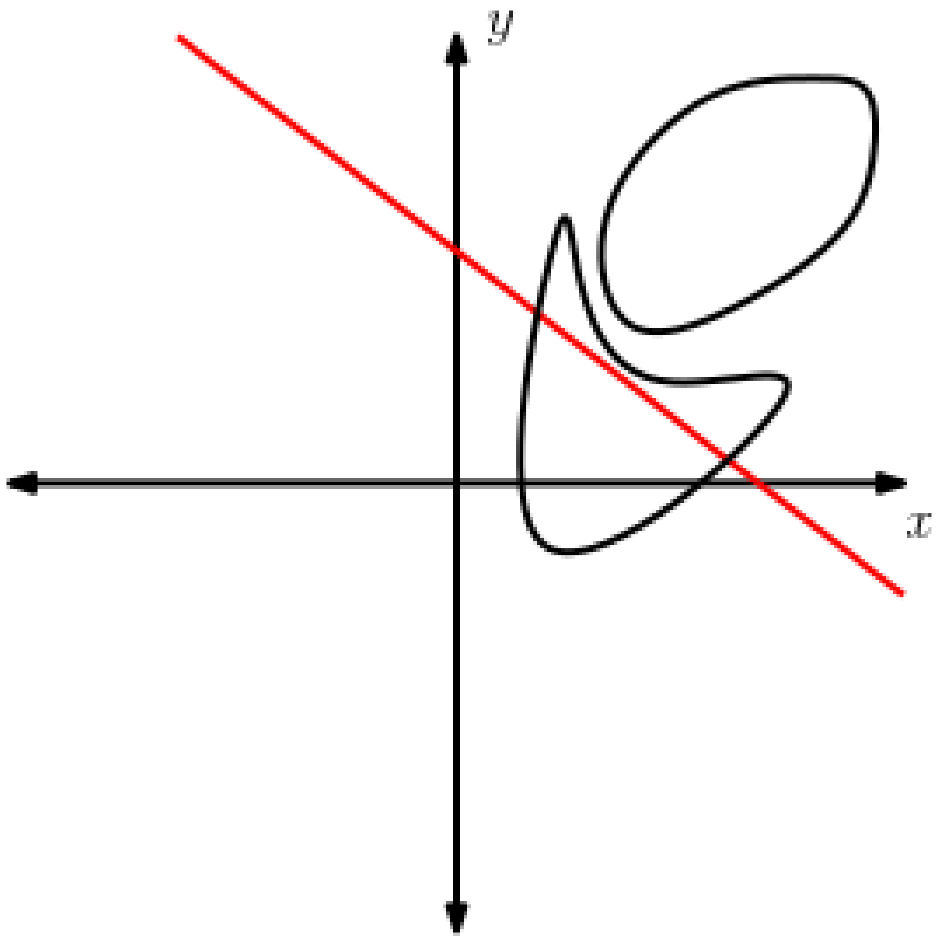
- **Corollary**: Since a single point is a closed convex set, for any point $\boldsymbol{x}'$ outside a closed convex set S**,** there exists a hyperplane such that

$$\boldsymbol{a} \cdot \boldsymbol{x} \leq b \text{ for all } \boldsymbol{x} \in S$$

$$\boldsymbol{a} \cdot \boldsymbol{x}' > b$$

Examples



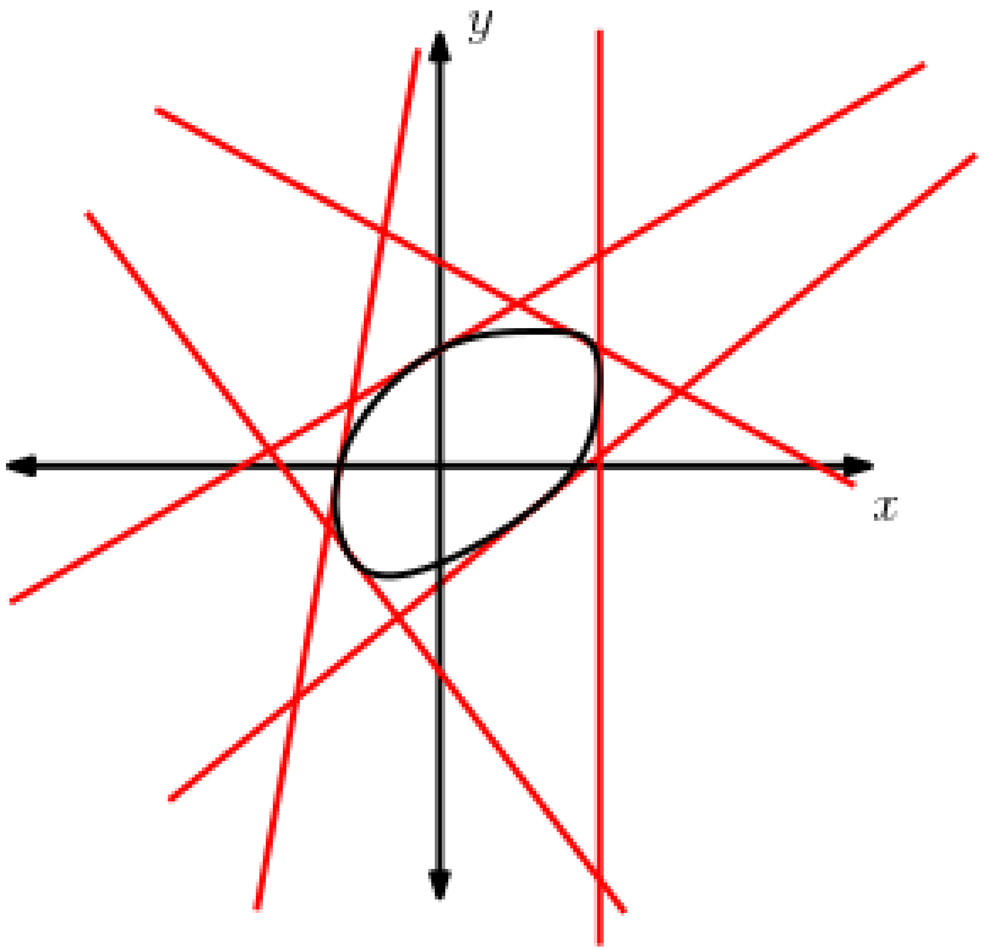


Hyperplane Representation of Convex Sets

- **Theorem**: For a closed convex set S, let H_S be the set of half-spaces that contain S. Then,

$$S = \cap_{h \in H_S} h$$

- Proof: For every point outside S, there exists a half-space h that separates the point from S by the previous corollary. Therefore, every point outside S is excluded by at least one $h \in H_S$.

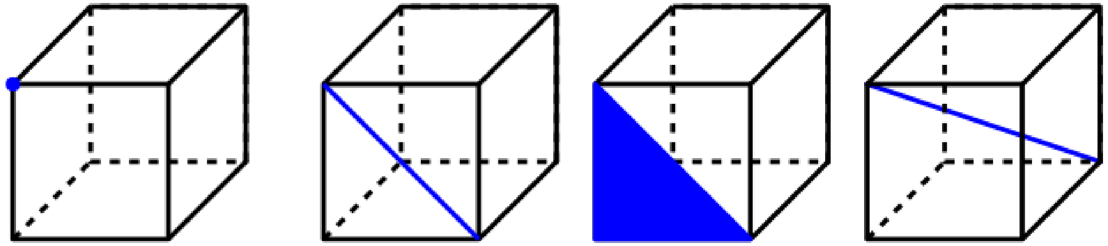


2.iv) Convex Polytopes

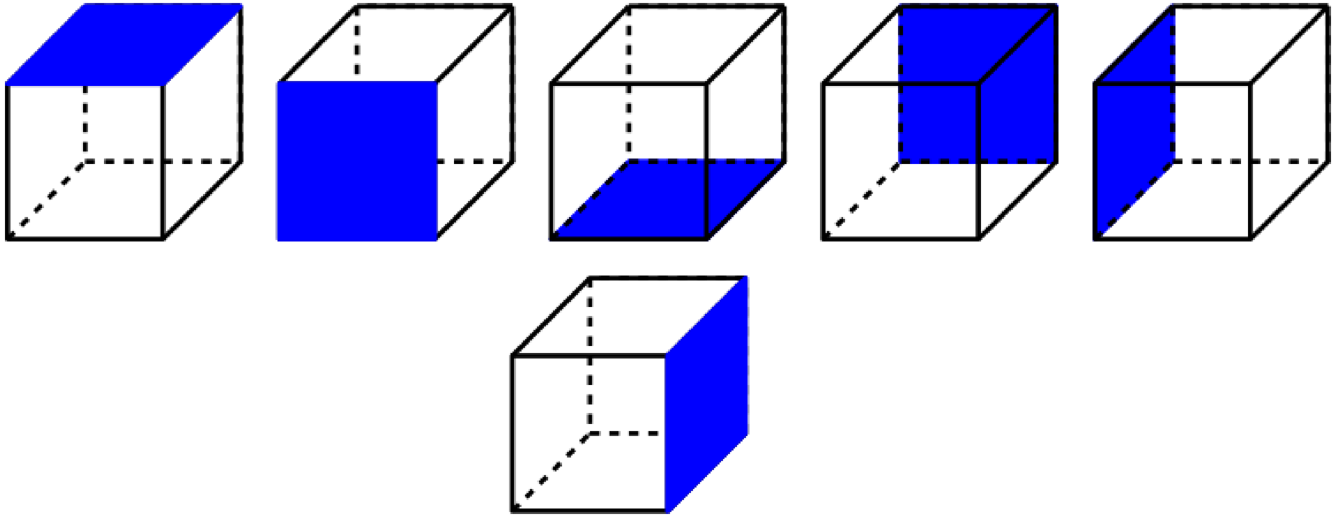
- A *convex polytope* is a convex set that has a finite number of extreme points.
- For a polytope, the extreme points are usually called *vertices*.
- Let X be a subset of the vertices. $\mathrm{conv}(X)$ is called a *face* of the polytope.
- If $\mathrm{conv}(X)$ only contains boundary points of the polytope, and no vertex can be added to X such that this is still the case then $\mathrm{conv}(X)$ is called a *facet.*

Example

Some Faces that are not Facets

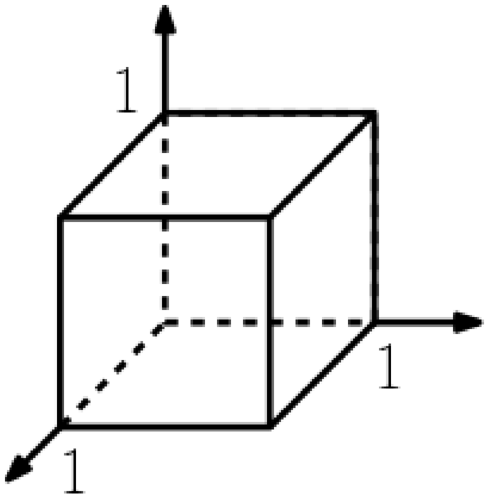


Facets of a cube



V-Representation and H-Representation

- To define a convex polytope, we can give:
  - A list of its vertices – The *V-representation*
  - A list of the half-spaces defining the facets – The *H-representation.*
- Converting between reps is computationally challenging. See ftp://ftp.math.ethz.ch/users/fukudak/reports/polyfaq040618.pdf

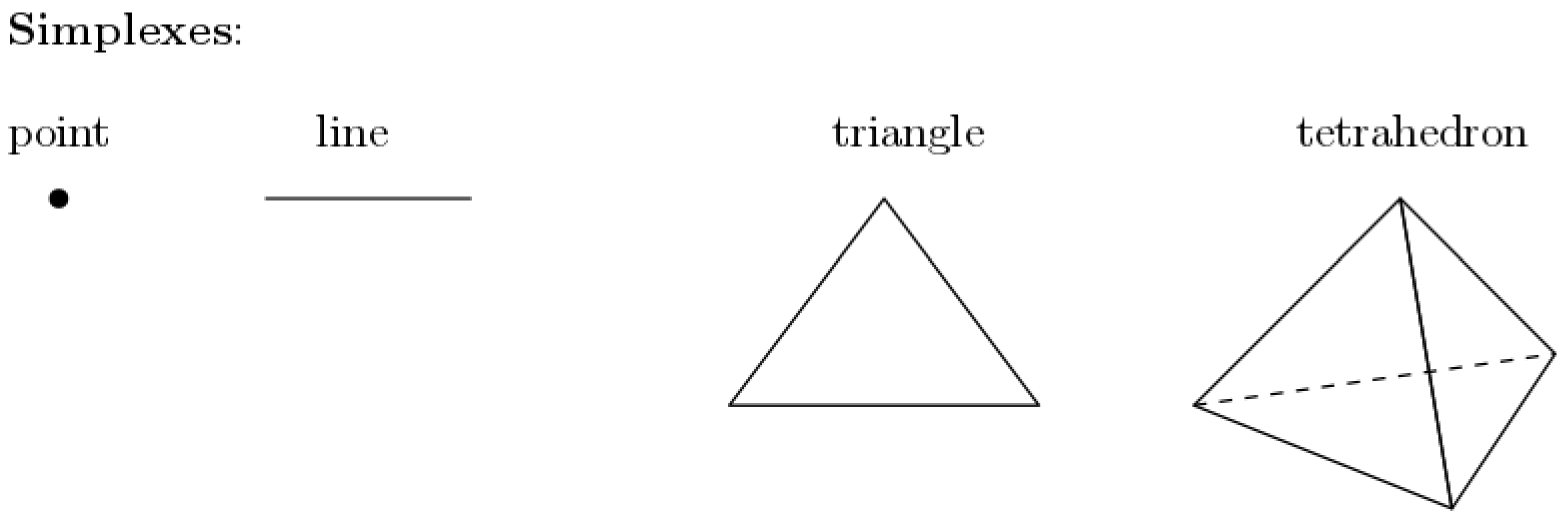


V-rep: $\begin{pmatrix}0\\0\\0\end{pmatrix},\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix},\begin{pmatrix}0\\1\\1\end{pmatrix},\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}1\\1\\0\end{pmatrix},\begin{pmatrix}1\\1\\1\end{pmatrix}$

H-rep: $0 \le x \le 1, \; 0 \le y \le 1, \; 0 \le z \le 1$

Simplexes

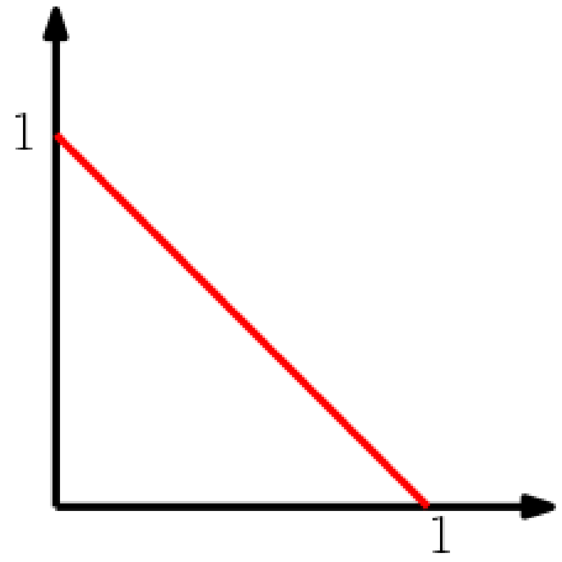
- A *simplex* is a complex polytope such that its vertices are linearly independent in the lifted cone.
- Consequences:
  - A simplex in $\mathbb{R}^n$ can have at most $n + 1$ vertices.
  - Every point inside a simplex can be written *uniquely* as a convex combination of its vertices.



The Probability Simplex

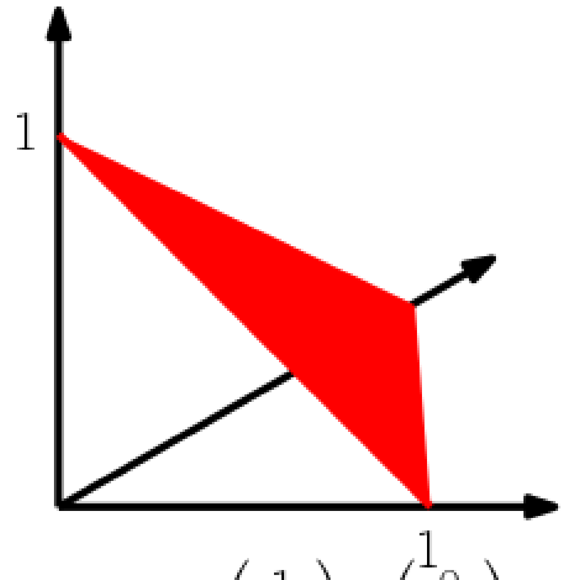
- We now discuss the motivating example for all of this mathematics.
- A *probability vector* in $\mathbb{R}^n$ is a vector $\boldsymbol{p}$ with positive components $p_j \geq 0$ that sum to 1, $\sum_{j=1}^n p_j = 1$.
- The *probability simplex* Δ_n is the set of all probability vectors in $\mathbb{R}^n$.
  - Note that Δ_n actually has dimension $n - 1$ because we could eliminate one component using the equality $\sum_{j=1}^n p_j = 1$.
  - It is convenient to keep representing it in $\mathbb{R}^n$ because that allows us to define the lifted cone on the same space, i.e. we just drop the condition $\sum_{j=1}^n p_j = 1$.

Examples



V-rep: $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

H-rep: $x \geq 0,\ y \geq 0,\ x + y = 1$



V-rep: $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

H-rep: $x \geq 0,\ y \geq 0,\ z \geq 0\ x + y + z = 1$

2.v) Inner Products

- An *inner product* on a vector space is a function $(\mathbf{x}, \mathbf{y})$ from pairs of vectors to scalars, satisfying
  - (Conjugate) symmetry: $(\boldsymbol{x}, \boldsymbol{y})^* = (\boldsymbol{y}, \boldsymbol{x})$
  - Linearity: $(\boldsymbol{z}, a\boldsymbol{x} + b\boldsymbol{y}) = a(\boldsymbol{z}, \boldsymbol{x}) + b$
  - Positive definiteness:

$$(\boldsymbol{x}, \boldsymbol{x}) \geq 0 \text{ and } (\boldsymbol{x}, \boldsymbol{x}) = 0 \text{ iff } \boldsymbol{x} = \mathbf{0}.$$

- Note: $*$ denotes complex conjugate – only relevant for complex vector spaces.
- An *inner product space* is a vector space equipped with an inner product.

Examples

- For $\mathbb{R}^n$ and $\mathbb{C}^n$, the dot product is an inner product
- For $\mathbb{R}^n$: $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix} = x_1 x_1' + x_2 x_2' + \cdots + x_n x_n'$
- For $\mathbb{C}^n$: $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix} = x_1^* x_1' + x_2^* x_2' + \cdots + x_n^* \, x_n'$

Orthonormal bases

- A basis $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ for an n-dimensional vector space is called *orthonormal* if

$$(\boldsymbol{x}_j, \boldsymbol{x}_k) = \delta_{jk} = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k \end{cases}$$

- For an orthonormal basis, if $\boldsymbol{x} = \sum_{j=1}^{n} a_j \boldsymbol{x}_j$ then

$$(\boldsymbol{x}_k, \boldsymbol{x}) = \sum_{j=1}^{n} a_j (\boldsymbol{x}_k, \boldsymbol{x}_j) = \sum_{j=1}^{n} a_j \delta_{jk} = a_k$$

so there is an easy way of finding components.

- Example: For $\mathbb{R}^2$ and $\mathbb{C}^2$, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is orthonormal, but $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is not.

2.vi) Dual Spaces

- A *dual vector* on a vector space V is a linear function from V to the scalars.

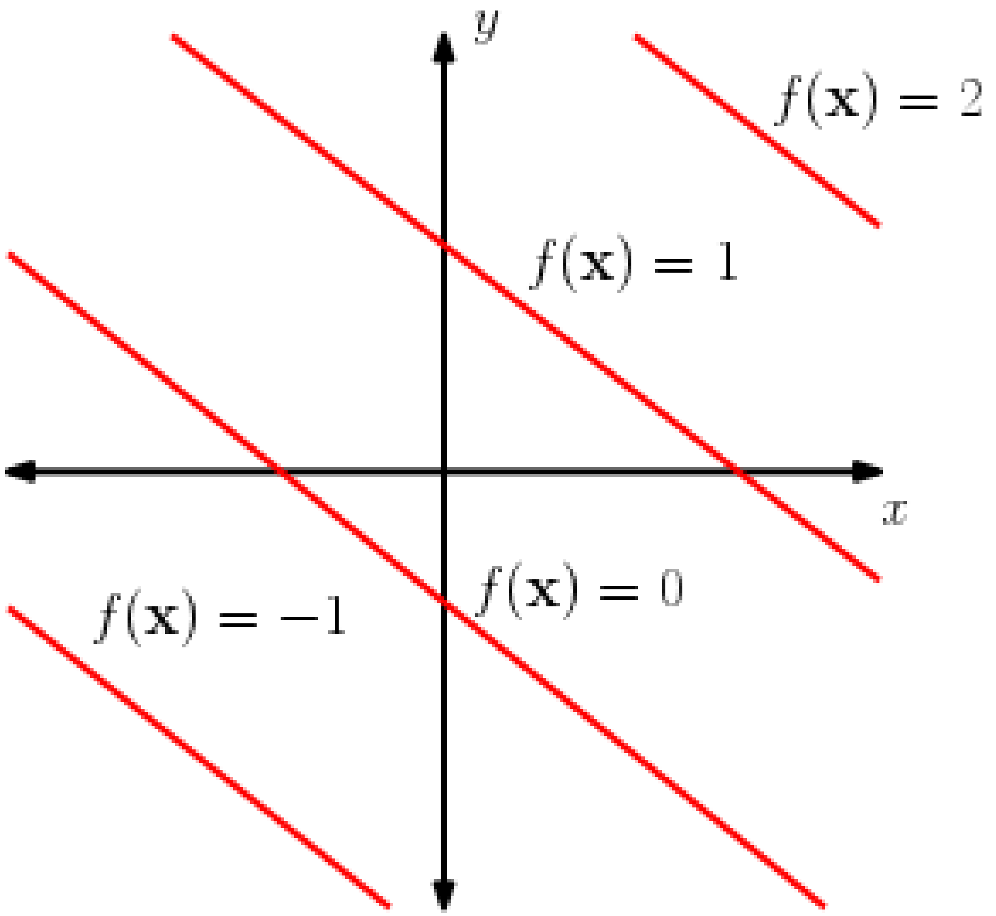
$$f(a\boldsymbol{x} + b\boldsymbol{y}) = af(\boldsymbol{x}) + bf(\boldsymbol{y})$$

- The set of dual vectors, denoted V^* is itself a vector space, called the *dual vector space*.
  - We define addition as: $(f + g)(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x})$
  - And scalar multiplication as: $(af)(\boldsymbol{x}) = a(f(\boldsymbol{x}))$

Dual vectors and Hyperplanes

- Consider the case of $\mathbb{R}^n$ and the set of vectors
$$F_a = \{\boldsymbol{x} \in \mathbb{R}^n | f(\boldsymbol{x}) = a\}$$
for some fixed dual vector f and scalar a.
- If $\boldsymbol{x}, \boldsymbol{y} \in F_a$ then $b\boldsymbol{x} + c\boldsymbol{y} \in F_a$ whenever $b + c = 1$
$$f(b\boldsymbol{x} + c\boldsymbol{y}) = bf(\boldsymbol{x}) + cf(\boldsymbol{y})$$
$$= ba + ca = (b + c)a = a$$
- This means that F_a contains the lines connecting any two points and hence it is a hyperplane.
- We can specify f by giving two hyperplanes on which it takes two different values and then filling in by linearity.

Example



Duals and Inner Products

- It is no accident that a hyperplane in $\mathbb{R}^n$ is given by an equation of the form $\boldsymbol{a} \cdot \boldsymbol{x} = b$.
- If we vary b and keep $\boldsymbol{a}$ fixed then this defines a set of parallel hyperplanes and hence a dual vector
$$f_{\boldsymbol{a}}(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{x}$$
- In any vector space, a vector $\boldsymbol{a} \in V$ determines a dual vector $f_{\boldsymbol{a}} \in V^*$ via
$$f_{\boldsymbol{a}}(\boldsymbol{x}) = (\boldsymbol{a}, \boldsymbol{x})$$

Duals and Inner Products

- The correspondence between vectors and duals is one-to-one. For any dual vector $f \in V^*$, define a vector $\boldsymbol{a}_f \in V$ as follows:
  - Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots$ be an orthonormal basis for V.
  - Define $\boldsymbol{a}_f = \sum_j f(\boldsymbol{x}_j)\boldsymbol{x}_j$.
  - Then $f(\boldsymbol{x}) = (\boldsymbol{a}_f, \boldsymbol{x})$
    - Proof: Since $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots$ is an basis, we can write $\boldsymbol{x} = \sum_j b_j \boldsymbol{x}_j$. Then,

$$(\boldsymbol{a}_f, \boldsymbol{x}) = \left(\boldsymbol{a}_f, \sum_j b_j \boldsymbol{x}_j\right) = \sum_j b_j (\boldsymbol{a}_f, x_j) = \sum_j b_j f(x_j)$$

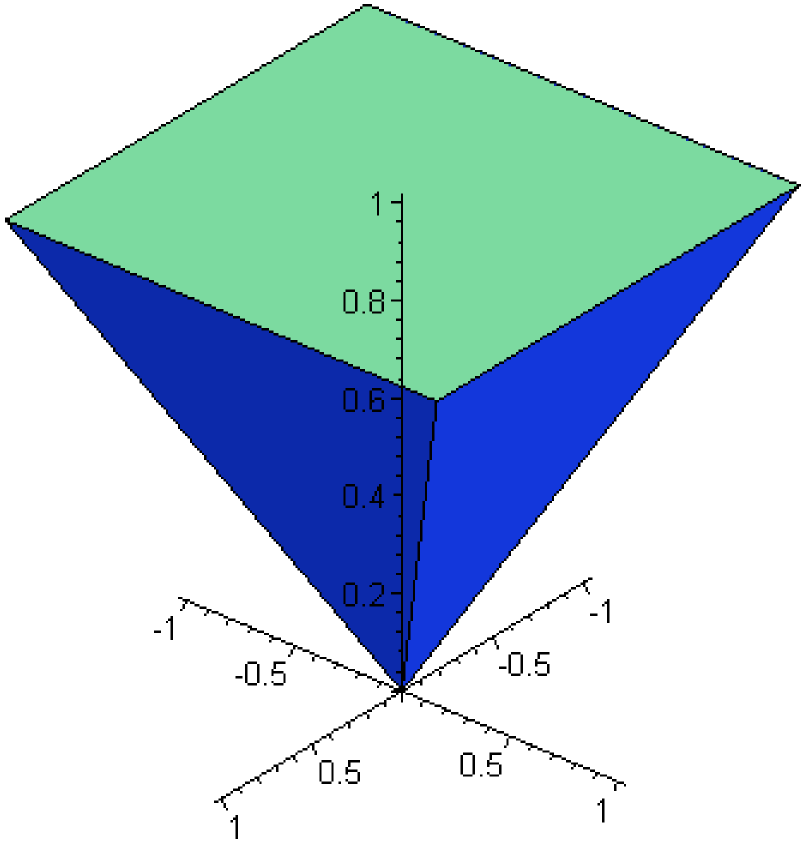
$$= f\left(\sum_j b_j \boldsymbol{x}_j\right) = f(\boldsymbol{x})$$

- A dual vector can be represented either as a set of parallel hyperplanes in V or as a vector in V.

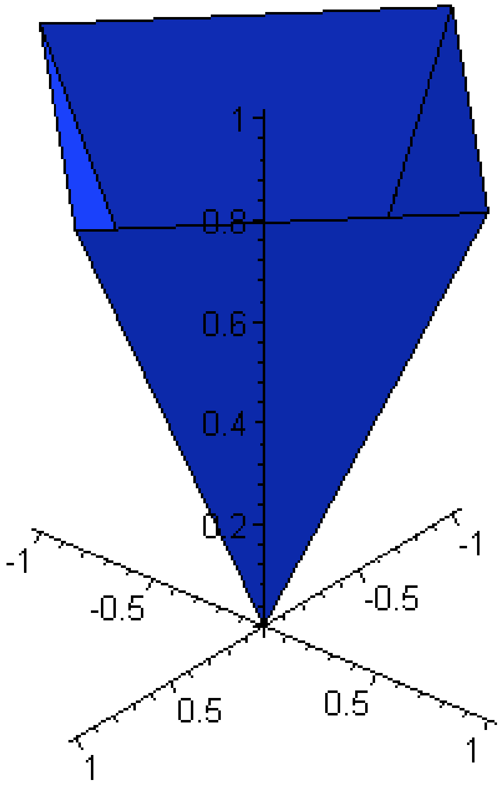
Dual Cones

- Given a convex cone C, the *dual cone* C^* is the set of all dual vectors f such that $f(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in C$.
- Equivalently, we can think of it as the set of all vectors $\boldsymbol{a}$ such that $(\boldsymbol{a}, \boldsymbol{x}) \geq 0$ $\boldsymbol{x} \in C$.
  - This makes it possible to draw the cone and its dual in the same space.
- The dual cone is itself a convex cone.
- Because dual vectors are linear, to check if $f \in C^*$ it suffices to check that $f(\boldsymbol{x}) \geq 0$ for the extremal rays.

Example



Cone C



Dual cone C^*

Self-Duality

- A cone C is called *self-dual* if $C = C^*$ (when we are representing dual vectors as inner products with vectors in the original space).
- Examples:
  - The cone generated by the probability simplex is self-dual.
  - The cone generated by a hypersphere (sphere in arbitrary dimension) is self-dual.

Defining a Closed Convex Set from a Convex Cone

- Last lecture, we explained how to lift a convex set to a convex cone by adding a dimension.
- You can reduce the dimension again by looking at points that lie on a specified hyperplane.
  - This will impose an equation $\boldsymbol{a} \cdot \boldsymbol{x} = b$, which can be used to eliminate one of the components.
- The intersection of a convex cone with a hyperplane will be a closed convex set if it intersects every extremal ray.
- More formally, given a cone C, we choose a dual vector u called the *unit* such that $0 < u(\boldsymbol{x}) < \infty$ on all extremal rays and define

$$S = \{\boldsymbol{x} \in C | u(\boldsymbol{x}) = 1\}$$

Defining a Closed Convex Set from a Convex Cone

- If we are using the standard lifting

$$\boldsymbol{x} \to \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix}$$

then choosing $u(\boldsymbol{x}) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \cdot \boldsymbol{x}$ will give us back our original convex set.