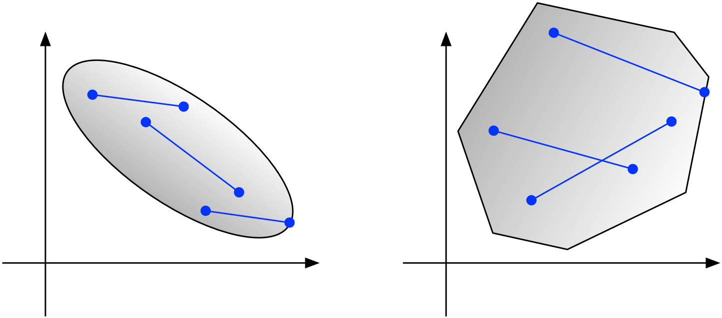
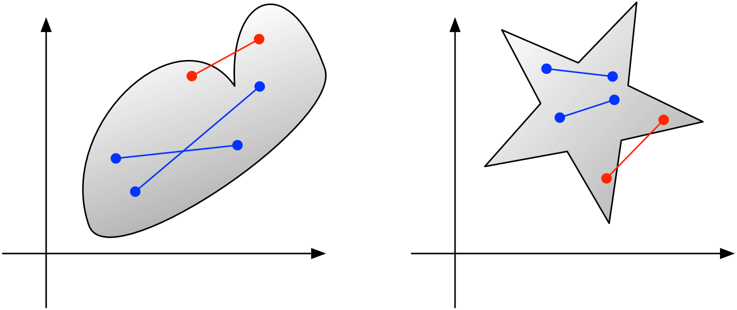
Convex sets

Nonlinear optimization
Instituto Superior Técnico and Carnegie Mellon University PhD course
João Xavier
TA: Hung Tuan

Convex sets



Nonconvex sets



Definition of convex set in a vector space

Important vector spaces

- $\mathbf{R}^n$: set of n-dimensional column vectors
- $\mathbf{R}^{n \times m}$: set of $n \times m$ matrices
- $\mathbf{S}^n$: set of $n \times n$ symmetric matrices

Definition. A set S in a vector space V is convex if

$$(1 - \alpha)x + \alpha y \in S$$

for all $x, y \in S$ and $\alpha \in [0, 1]$.

How do we recognize convex sets?

List of simple ones

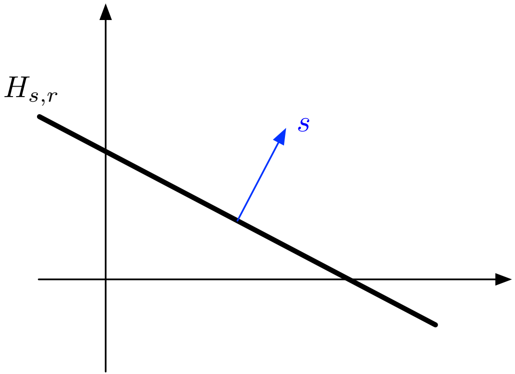
$+$

Apply convexity-preserving operations

Simple convex sets

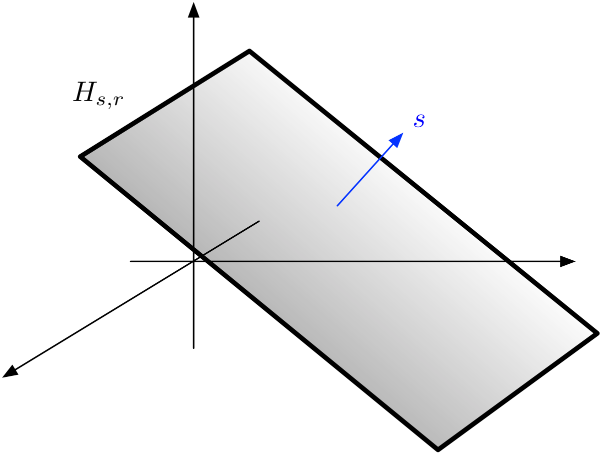
- hyperplanes
- closed half-spaces
- norm balls
- cones

Hyperplane in $\mathbf{R}^2$



$$H_{s,r} = \{x \in \mathbf{R}^2 \mid s^T x = r\}$$

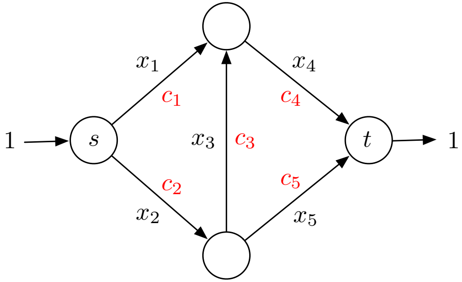
Hyperplane in $\mathbf{R}^3$

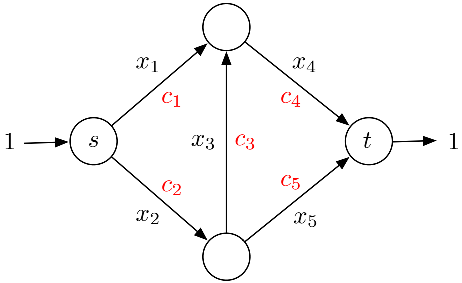


$$H_{s,r} = \{x \in \mathbf{R}^3 \mid s^T x = r\}$$

Example: network flow

- you want to sent one unit of fluid from s to t
- cost per unit flow in arc i is c_i
- how much fluid should be sent through arc i: $x_i =?$





Formulation that minimizes cost:

$$
\begin{aligned}
\underset{x_1,x_2,x_3,x_4,x_5}{\text{minimize}} \quad & c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5 \\
\text{subject to} \quad & 1 = x_1 + x_2 \\
& x_1 + x_3 = x_4 \\
& x_2 = x_3 + x_5 \\
& x_4 + x_5 = 1 \\
& x_1, x_2, x_3, x_4, x_5 \geq 0
\end{aligned}
$$

Hyperplane in a vector space

Default inner-products in important vector spaces

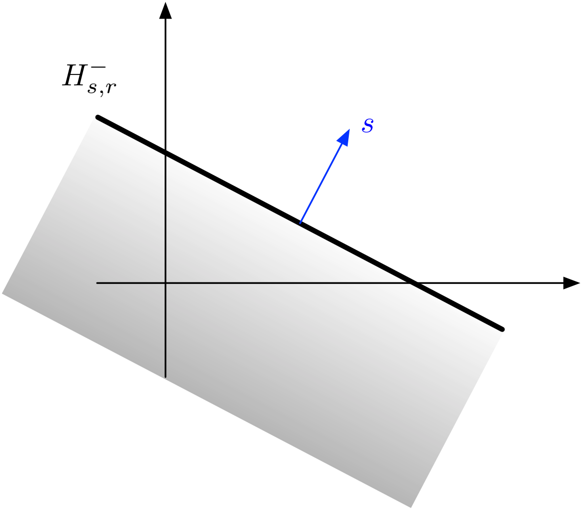
- $\mathbf{R}^n$: $\langle x, y \rangle = x^T y$
- $\mathbf{R}^{n \times m}$: $\langle X, Y \rangle = \mathsf{tr}\left(X^T Y\right)$
- $\mathbf{S}^n$: $\langle X, Y \rangle = \mathsf{tr}\,(XY)$

Definition. A hyperplane in a vector space V with inner product $\langle \cdot, \cdot \rangle$ is a set of the form

$$H_{s,r} = \{x \in V \mid \langle s, x \rangle = r\}$$

for some $s \in V - \{0\}$ and $r \in \mathbf{R}$.

Closed half-space in $\mathbf{R}^2$



$$H_{s,r}^- = \{x \in \mathbf{R}^2 \mid s^T x \le r\}$$

Example: warehouse management

- you can buy two products with costs c_1 and c_2
- you can sell them at prices p_1 and p_2
- demands for the products are d_1 and d_2
- volumes of products are v_1 and v_2
- maximum capacity of your warehouse is volume v
- how much should you order from each product: $x_1 = ?$ and $x_2 = ?$

Formulation that minimizes your net cost:

$$\begin{array}{ll} \underset{x_1 \in \mathbf{R}, x_2 \in \mathbf{R}}{\text{minimize}} & c_1 x_1 + c_2 x_2 - (p_1 \min\{x_1, d_1\} + p_2 \min\{x_2, d_2\}) \\ \text{subject to} & v_1 x_1 + v_2 x_2 \leq v \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{array}$$

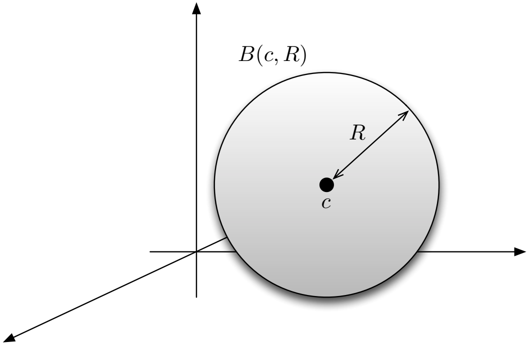
Closed half-space in a vector space

Definition. A closed-half space in a vector space V with inner product $\langle \cdot, \cdot \rangle$ is a set of the form

$$H_{s,r}^{-} = \{x \in V \mid \langle s, x \rangle \leq r\}$$

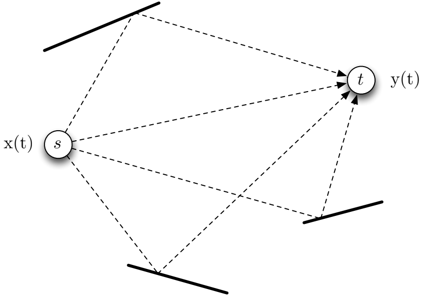
for some $s \in V - \{0\}$ and $r \in \mathbf{R}$.

Norm ball in $\mathbf{R}^3$



$$B(c, R) = \{x \in \mathbf{R}^3 \mid \|x - c\|_2 \leq R\}$$

Example: signal design



- node s transmits $x(1), x(2), x(3), \ldots, x(T)$ to flag an event
- node t receives $y(1), y(2), y(3), \ldots, y(T)$
- multipath channel:

$$y(t) = h_0 x(t) + h_1 x(t-1) + h_2 x(t-2) + h_3 x(t-3)$$

- assume system at rest:

$$\underbrace{\begin{bmatrix} y(1) \\ y(2) \\ y(3) \\ \vdots \\ y(T) \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} h_0 & 0 & \cdots & \cdots & 0 \\ h_1 & h_0 & 0 & \cdots & 0 \\ h_2 & h_1 & h_0 & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & h_2 & h_1 & h_0 \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} x(1) \\ x(2) \\ x(3) \\ \vdots \\ x(T) \end{bmatrix}}_{x}$$

- power constraint: $x(1)^2 + x(2)^2 + \cdots + x(T)^2 \leq p$

- which input x gives the largest output y?

- formulation:

$$\begin{array}{ll} \underset{x \in \mathbf{R}^T}{\text{maximize}} & \|Hx\|_2 \\ \text{subject to} & \|x\|_2 \leq \sqrt{p} \end{array}$$

- a curiosity: the problem above is not convex but it is easily solved, via SVD

Norm ball in a vector space

Default norms in important vector spaces

- euclidean norm in $\mathbf{R}^n$: $\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$
- frobenius norm in $\mathbf{R}^{n\times m}$: $\|X\|_{\mathsf{F}} = \sqrt{\sum_{i,j} X_{ij}^2}$
- frobenius norm in $\mathbf{S}^n$: $\|X\|_{\mathsf{F}} = \sqrt{\sum_{i,j} X_{ij}^2}$

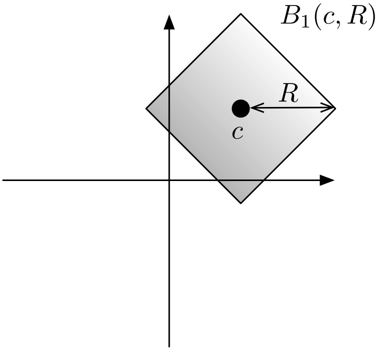
Definition. A norm ball in a vector space V with norm $\|\cdot\|$ is a set of the form

$$B(c, R) = \{x \in V \ : \ \|x - c\| \leq R\}$$

for some $c \in V$ and $R \geq 0$.

ℓ_1 norm in $\mathbf{R}^n$

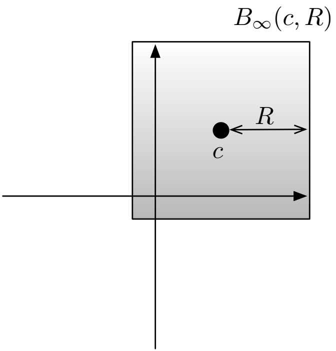
For $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$: $\|x\|_1 = |x_1| + \cdots + |x_n|$



$$B_1(c, R) = \{x \in \mathbf{R}^n \; : \; \|x - c\|_1 \leq R\}$$

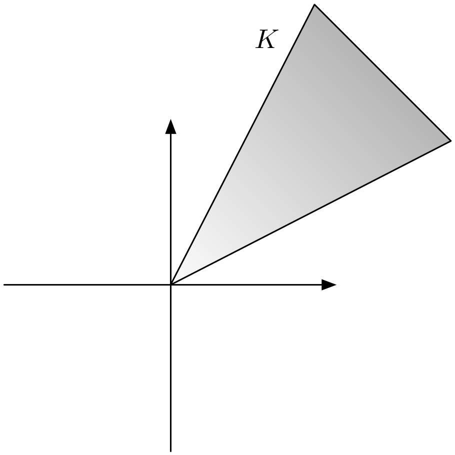
ℓ_∞ norm in $\mathbf{R}^n$

For $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$: $\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$



$$B_\infty(c, R) = \{x \in \mathbf{R}^n \ : \ \|x - c\|_\infty \le R\}$$

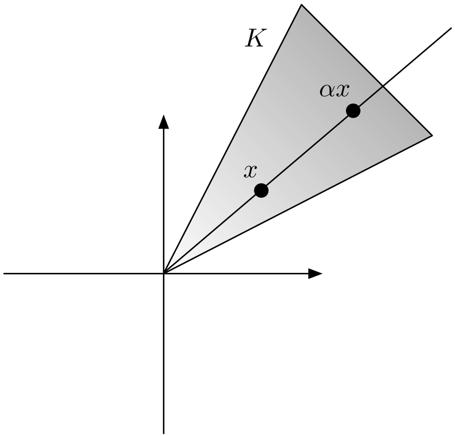
Convex cone in $\mathbf{R}^2$



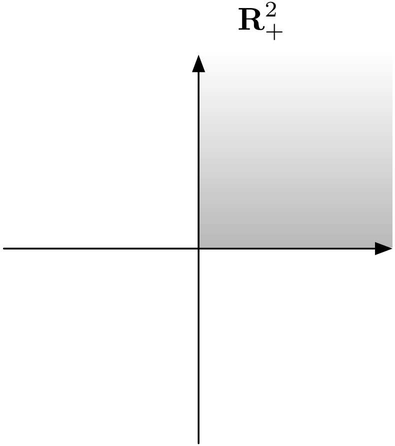
Convex cone in a vector space

Definition. A set K in a vector space V is a convex cone if K is a convex set and

$$\mathbf{R}_+ K = \{\alpha x : \alpha \geq 0, x \in K\} \subset K$$

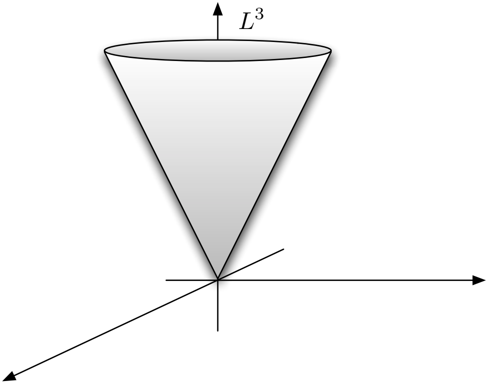


Nonnegative orthant in $\mathbf{R}^n$



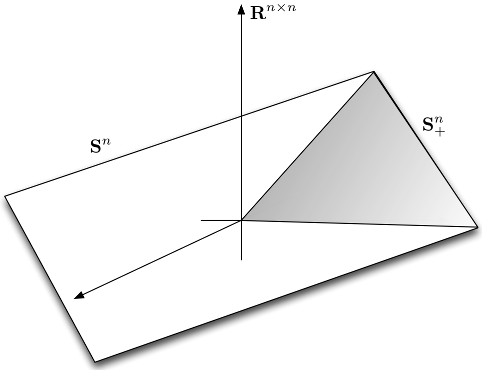
$$\mathbf{R}^n_+ = \{x \in \mathbf{R}^n \,:\, x_i \geq 0, \text{ for } i = 1, \ldots, n\}$$

Lorentz cone or second-order cone in $\mathbf{R}^{n+1}$



$$L^{n+1} = \{(x, x_{n+1}) \in \mathbf{R}^n \times \mathbf{R} \,:\, \|x\|_2 \leq x_{n+1}\}$$

Positive semidefinite cone in $\mathbf{S}^n$



$$\mathbf{S}^n_+ = \{X \in \mathbf{S}^n \mid X \succeq 0\}$$

($X \succeq 0$ means that all eigenvalues of X are nonnegative)

Eigenvalue decomposition

Theorem. Any $X \in \mathbf{S}^n$ can be factored as

$$X = Q\Lambda Q^T$$

where

- $Q = \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix} \in \mathbf{R}^{n\times n}$ is orthogonal: $Q^TQ = QQ^T = I$
- $\Lambda \in \mathbf{R}^{n\times n}$ is diagonal:

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Note: each (λ_i, q_i) is an eigenpair of X,

$$Xq_i = \lambda_i q_i, \quad i = 1, \dots, n.$$

Example: finance

- given an invalid correlation matrix $\widehat{\Sigma}$
- find the closest correlation matrix with unit diagonal

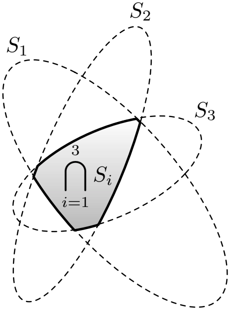
Formulation:

$$\begin{array}{ll} \underset{\Sigma \in \mathbf{S}^n}{\text{minimize}} & \left\| \Sigma - \widehat{\Sigma} \right\|_{\mathsf{F}} \\ \text{subject to} & \Sigma_{ii} = 1, \quad i = 1, \ldots, n, \\ & \Sigma \succeq 0 \end{array}$$

Operations that preserve convexity

- intersection
- push-forward by affine map
- pull-back by affine map

Intersection preserves convexity

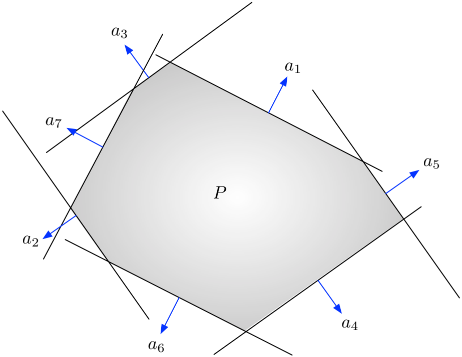


Theorem. If $\{S_i : i \in \mathcal{I}\}$ is a family of convex sets in a vector space V, then their intersection $\bigcap_{i \in \mathcal{I}} S_i$ is a convex set.

Important: the index set $\mathcal{I}$ may be uncountable.

Example: polyhedron

$$P = \{x \in \mathbf{R}^2 : a_i^T x \le b, i = 1, \ldots, 7\}$$



Compact notation:

$$P = \{x \in \mathbf{R}^2 : Ax \le b\} \quad A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_7^T \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_7 \end{bmatrix}$$

Affine map

Definition. An affine map $\mathcal{A} : V \to W$ is a map of the form

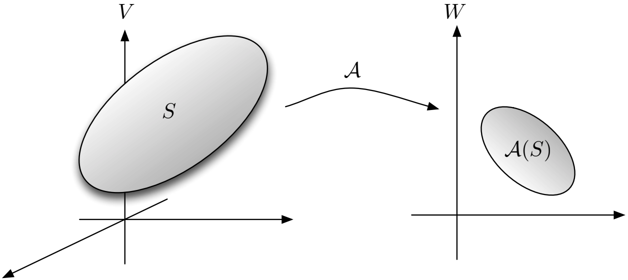
$$\mathcal{A}(v) = \mathcal{L}(v) + w$$

for some linear map $\mathcal{L} : V \to W$ and some $w \in W$.

Examples:

- $\mathcal{A} : \mathbf{R}^m \to \mathbf{R}^n$, $\mathcal{A}(x) = Ax + b$ $(A \in \mathbf{R}^{n\times m}, b \in \mathbf{R}^m)$
  (all affine maps $\mathbf{R}^n \to \mathbf{R}^m$ are of this form)
- $\mathcal{A} : \mathbf{R}^n \to \mathbf{S}^m$, $\mathcal{A}(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ $(A_i \in \mathbf{S}^m)$
  (all affine maps $\mathbf{R}^n \to \mathbf{S}^m$ are of this form)
- $\mathcal{A} : \mathbf{S}^n \to \mathbf{S}^m$, $\mathcal{A}(X) = A^T X A + B$ $(A \in \mathbf{R}^{n\times m}, B \in \mathbf{S}^m)$
  (just an example; not all affine maps $\mathbf{S}^n \to \mathbf{S}^m$ are of this form)

Push-forward by affine map preserves convexity



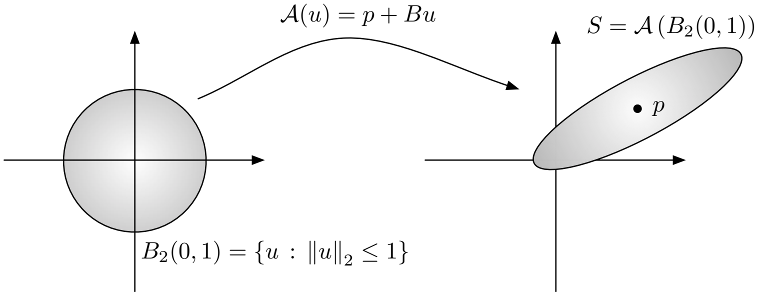
Theorem. If $S \subset V$ is a convex set and $\mathcal{A} : V \to W$ an affine map, then

$$\mathcal{A}(S) = \{\mathcal{A}(v) : v \in S\}$$

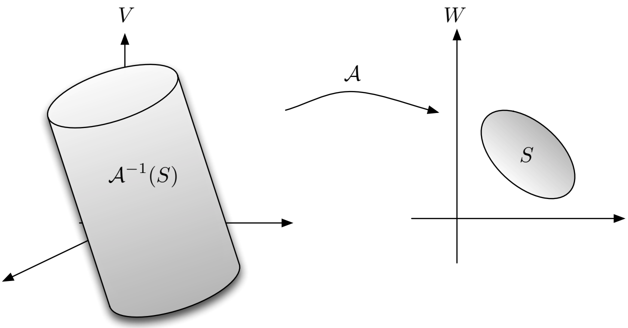
is a convex set.

Example: possible positions of robot after disturbance

- robot is at position $p \in \mathbf{R}^2$
- disturbance $u \in \mathbf{R}^2$ hits the robot
- robot moves to new position $p + Bu$ (B depends on robot mechanic parameters)
- disturbance is unknown but is limited in magnitude: $\|u\|_2 \leq 1$
- is the set S of all possible new positions a convex set?



Pull-back by affine map preserves convexity



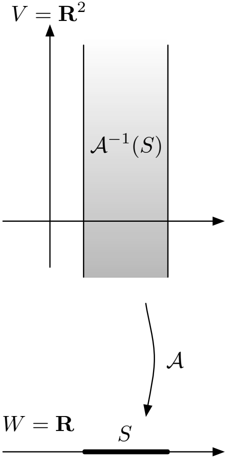
Theorem. If $S \subset W$ is a convex set and $\mathcal{A} : V \to W$ an affine map, then

$$\mathcal{A}^{-1}(S) = \{v \in V \, : \, \mathcal{A}(v) \in S\}$$

is a convex set.

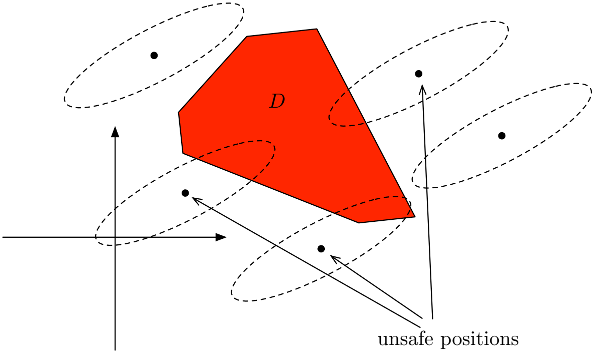
$\mathcal{A}$ does not have to be invertible:

- $\mathcal{A} : \mathbf{R}^2 \to \mathbf{R}$, $\mathcal{A}(x_1, x_2) = x_1$ and $S = [1, 4]$
- $\mathcal{A}^{-1}(S) = \{(x_1, x_2) \in \mathbf{R}^2 : 1 \le x_1 \le 4\}$



Example: unsafe robot positions

- D is a polyhedral danger zone that the robot must not enter
- an unsafe position may enter D after a disturbance
- is the set S of all unsafe positions a convex set?

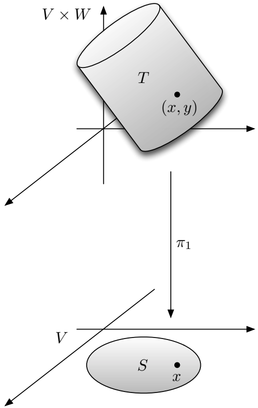


A trivial but useful fact

If $S = \{x \in V \,:\, (x,y) \in T \text{ for some } y \in W\}$ then

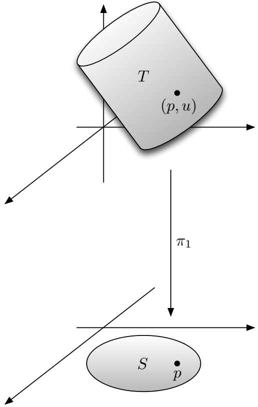
$$S = \pi_1(T)$$

where $\pi_1 : V \times W \to V$ the projection map $\pi_1(x,y) = x$



$$
\begin{aligned}
S &= \left\{p \in \mathbf{R}^2 \,:\, p + Bu \in D, \text{ for some } u \in B_2(0,1)\right\} \\
&= \pi_1\,(\,\underbrace{\left\{(p,u) \in \mathbf{R}^2 \times \mathbf{R}^2 \,:\, p + Bu \in D, u \in B_2(0,1)\right\}}_{T}\,)
\end{aligned}
$$

where $\pi_1 : \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}^2$ the projection map $\pi_1(p,u) = p$



- write $T = \underbrace{\{(p,u) \,:\, p + Bu \in D\}}_{T_1} \cap \underbrace{\{(p,u) \,:\, u \in B_2(0,1)\}}_{T_2}$
- T_1 is convex because it is an affine pull-back of convex set D:

$$T_1 = \mathcal{B}^{-1}(D)$$

  where $\mathcal{B} \,:\, \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}^2$, $\mathcal{B}(p,u) = p + Bu$
- T_2 is convex because it is an affine pull-back of convex set $B_2(0,1)$:

$$T_2 = {\pi_2}^{-1}(B_2(0,1))$$

  where $\pi_2 \,:\, \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}^2$ the projection map $\pi_2(p,u) = u$
- T is convex because it is intersection of two convex sets

$$T = T_1 \cap T_2$$

- S is convex because it is affine push-foward of convex set T:

$$S = \pi_1\,(T)$$