CPSC 440: Advanced Machine Learning

Convex Optimization

Mark Schmidt

University of British Columbia

Winter 2021

Machine Learning and Optimization

- In machine learning, training is typically written as an optimization problem:
  - We optimize parameters w of model, given data.
- There are some exceptions:
  1. Methods based on counting and distances (KNN, random forests).
     - See CPSC 340.
  2. Methods based on averaging and integration (Bayesian learning).
     - Later in course.

  But even these models have parameters to optimize.
- Important class of optimization problems: convex optimization problems.

Convex Optimization

- Consider an optimization problem of the form

$$\min_{w \in \mathcal{C}} f(w).$$

  where we are minimizing a function f subject to w being in the set $\mathcal{C}$.
  - For least squares we have $f(w) = \|Xw - y\|^2$ and $\mathcal{C} \equiv R^d$
  - If we had non-negative constraints, we would have $\mathcal{C} \equiv \{w \mid w \geq 0\}$.
    - Notation: when I write $w \geq 0$ for a vector w I mean inequality holds for each row.

- We say that this is a convex optimization problem if:
  - The set $\mathcal{C}$ is a convex set.
  - The function f is a convex function.

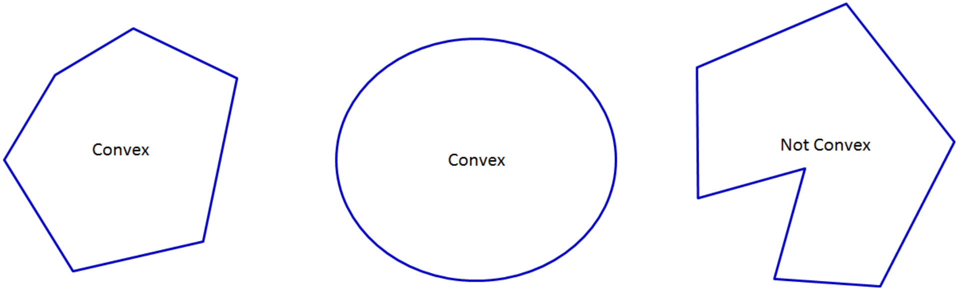
- This lecture is boring, but convexity ideas will show up throughout the course.

Convex Optimization

- Key property of convex optimization problems:
  - All local optima are global optima.
- Convexity is usually a good indicator of tractability:
  - Minimizing convex functions is usually easy.
  - Minimizing non-convex functions is usually hard.
- Off-the-shelf software solves many classes of convex problems (*MathProgBase*).

Definition of Convex Sets

- A set $\mathcal{C}$ is convex if the line between any two points stays also in the set.



Definition of Convex Sets

- To formally define convex sets, we use the notion of convex combination:
  - A convex combination of two variables w and v is given by

    $$\theta w + (1-\theta)v \quad \text{for any} \quad 0 \leq \theta \leq 1,$$

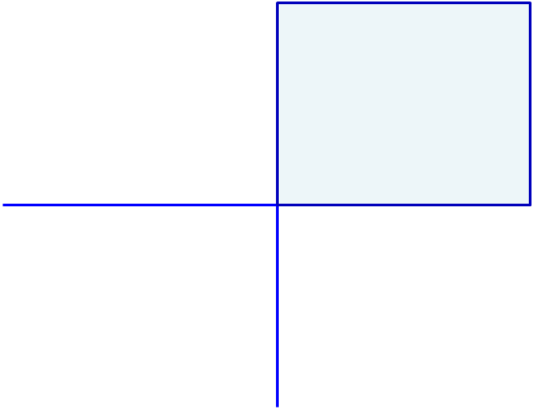
    which characterizes the points on the line between w and v.

- A set $\mathcal{C}$ is convex if convex combinations of points in the set are also in the set:
  - For all $w \in \mathcal{C}$ and $v \in \mathcal{C}$ we have $\underbrace{\theta w + (1-\theta)v}_{\text{convex comb}} \in \mathcal{C}$ for $0 \leq \theta \leq 1$.

- This definition allows us to prove the convexity of many simple sets.

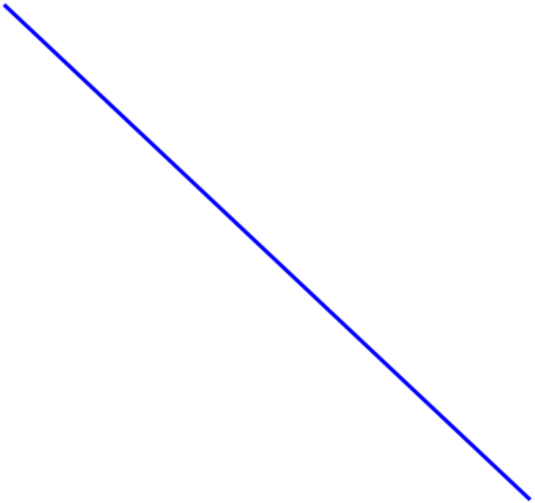
Examples of Simple Convex Sets

- Real space $\mathbb{R}^d$.
- Positive orthant $\mathbb{R}^d_+ : \{w \mid w \geq 0\}$.
- Hyper-plane: $\{w \mid a^\top w = b\}$.
- Half-space: $\{w \mid a^\top w \leq b\}$.
- Norm-ball: $\{w \mid \|w\|_p \leq \tau\}$.
- Norm-cone: $\{(w, \tau) \mid \|w\|_p \leq \tau\}$.



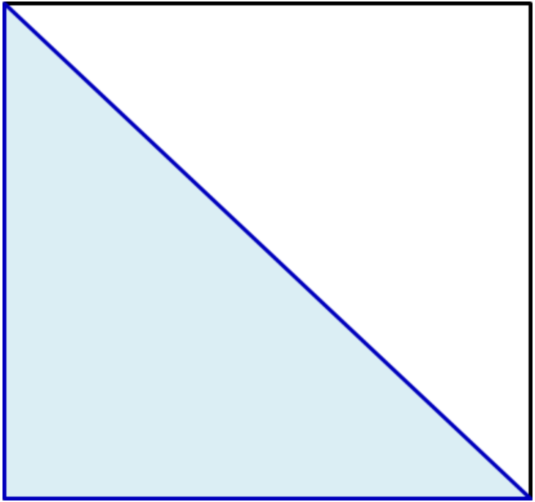
Examples of Simple Convex Sets

- Real space $\mathbb{R}^d$.
- Positive orthant $\mathbb{R}^d_+ : \{w \mid w \geq 0\}$.
- Hyper-plane: $\{w \mid a^\top w = b\}$.
- Half-space: $\{w \mid a^\top w \leq b\}$.
- Norm-ball: $\{w \mid \|w\|_p \leq \tau\}$.
- Norm-cone: $\{(w, \tau) \mid \|w\|_p \leq \tau\}$.



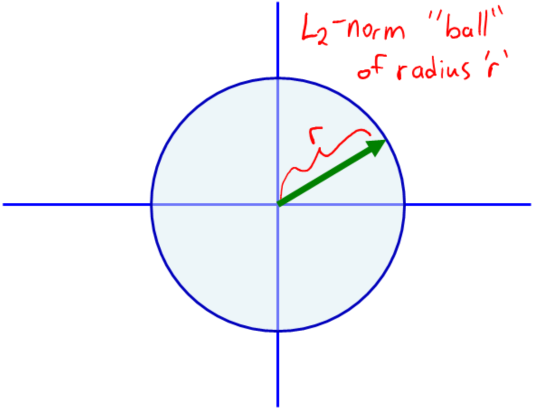
Examples of Simple Convex Sets

- Real space $\mathbb{R}^d$.
- Positive orthant $\mathbb{R}^d_+ : \{w \mid w \geq 0\}$.
- Hyper-plane: $\{w \mid a^\top w = b\}$.
- Half-space: $\{w \mid a^\top w \leq b\}$.
- Norm-ball: $\{w \mid \|w\|_p \leq \tau\}$.
- Norm-cone: $\{(w, \tau) \mid \|w\|_p \leq \tau\}$.



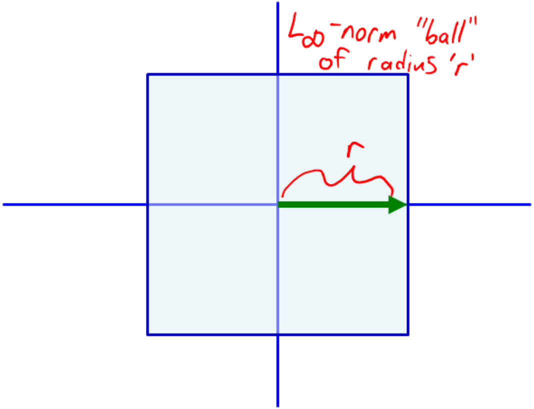
Examples of Simple Convex Sets

- Real space $\mathbb{R}^d$.
- Positive orthant $\mathbb{R}^d_+ : \{w \mid w \geq 0\}$.
- Hyper-plane: $\{w \mid a^\top w = b\}$.
- Half-space: $\{w \mid a^\top w \leq b\}$.
- Norm-ball: $\{w \mid \|w\|_p \leq \tau\}$.
- Norm-cone: $\{(w, \tau) \mid \|w\|_p \leq \tau\}$.



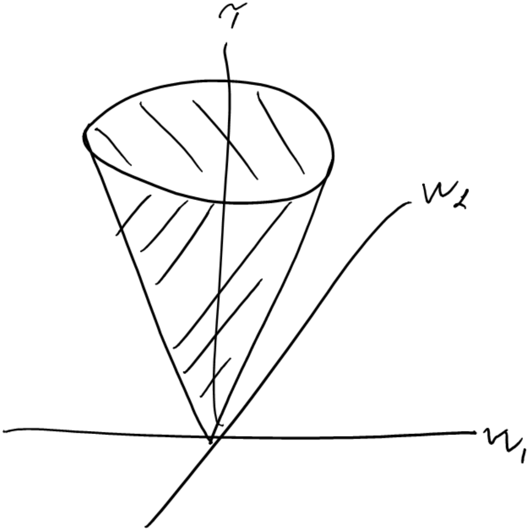
Examples of Simple Convex Sets

- Real space $\mathbb{R}^d$.
- Positive orthant $\mathbb{R}^d_+ : \{w \mid w \geq 0\}$.
- Hyper-plane: $\{w \mid a^\top w = b\}$.
- Half-space: $\{w \mid a^\top w \leq b\}$.
- Norm-ball: $\{w \mid \|w\|_p \leq \tau\}$.
- Norm-cone: $\{(w, \tau) \mid \|w\|_p \leq \tau\}$.



Examples of Simple Convex Sets

- Real space $\mathbb{R}^d$.
- Positive orthant $\mathbb{R}^d_+ : \{w \mid w \geq 0\}$.
- Hyper-plane: $\{w \mid a^\top w = b\}$.
- Half-space: $\{w \mid a^\top w \leq b\}$.
- Norm-ball: $\{w \mid \|w\|_p \leq \tau\}$.
- Norm-cone: $\{(w, \tau) \mid \|w\|_p \leq \tau\}$.



Summary

- Convex optimization problems are a class that we can usually efficiently solve.
- Next time: more about convexity than you ever wanted to know.