Recitation 4

Geometric Derivation of SVMs and Complementary Slackness

DS-GA 1003 Machine Learning

Spring 2023

Feb 15, 2023

Logistics

- HW 1 Grades Released tonight (if not already)
- HW 2 Due tonight (late submission until Friday)
- HW 3 Released tonight (if not already)

Agenda: Topics to Discuss

- SVM (Intuition, Derivation)
- Hinge Loss (Relation to SVM)
- Subgradient (Intuition)
- Duality (Intuition, why we care)

SVM:Motivation

Simple Idea:

SVM:Derivation

SVM:Derivation

Subgradient: Intuition

Subgradient: Definition and Properties

A vector $g \in \mathbf{R}^d$ is a subgradient of a convex function f at x if for all z

$$f(z) \geq f(x) + g^T(z - x)$$

Subgradient Descent

- f is differentiable at x iff $\partial f(x) = \{\nabla f(x)\}$
- If f is convex, then subdifferntial is non-empty
- If f is convex, then x is global optimum iff $0 \in \partial f(x)$

For non-convex functions

- Who cares

Subgradient Descent: Overview

- Works in the same way as graidnet descent (almost)
- Special definition for 'gradient' at non-differentiable x
- When landed on non-differentiable x, next step may not decrease function value. But should jump out of it very quickly

Subgradient Descent: Property

- Need to adjust for step size
- Slower than gradient descent

Another approach to solve SVM Optimization

An ancient/classic algorithm called quadratic program can solve convex optimization problems

- Intro to Lagrange Multipliers and Dual variables/problem

Lower Bound Property

Strong VS Weak Duality and KKT

Why we care