Maestría de Ciencia de Datos

Optimización Convexa

Tarea 10: Funciones Convexas

Estudiante: Daniel Nuño
Profesor: Dr. Juan Diego Sanchez Torres
Fecha entrega: 30 de marzo, 2022

Introduction

When an optimization problem is convex, several properties of its solution can be verified. Therefore, the following exercises review fundamental results for understanding convex optimization problems' solutions. Subsequently, the models' convexity based on the support vector concept that has been worked on up to this point must be verified. Once this property has been verified in the SVMs, the properties of their solutions are deduced.

Activities

Consider the following exercises on convex functions:

Problem 1: Basic Exercises on Convex Functions

Prove (using math) the following results:

1. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable over an open domain. Then, the following are equivalent: *(i)* f is convex. *(ii)* $f(y) \geq f(x) + \nabla f(x)^T(y-x)$, for all $x, y \in \text{dom}(f)$. *(iii)* $\nabla^2 f(x) \succeq 0$, for all $x \in \text{dom}(f)$.

2. Consider an unconstrained optimization problem

$$\begin{aligned} &\min f(x) \\ &\text{s.t. } x \in \mathbb{R}^n \end{aligned}$$

   where f is convex and differentiable. Then, any point $\bar{x}$ that satisfies $\nabla f(\bar{x}) = 0$ is a global minimum.

3. Consider an optimization problem

$$\begin{aligned} &\min f(x) \\ &\text{s.t. } x \in \Omega \end{aligned}$$

   where $f : \mathbb{R}^n \to \mathbb{R}$ is strictly convex on Ω and Ω is a convex set. Then the optimal solution (assuming it exists) must be unique.

Problema 1

30 Marzo, 22
Daniel Niño

1

Definición de la segunda derivada y lo queremos llegar,

$$\frac{f'(y) - f'(x)}{y-x} \geq 0 \quad (1)$$

dadas las ecuaciones que describen el mínimo,

$$f(y) \geq f(x) + f'(x)(y-x) \quad (2)$$

$$f(x) \geq f(y) + f'(y)(x-y) \quad (3)$$

despejando ambas ecuaciones

$$f(y) - f(x) \geq f'(x)(y-x) \quad (2)$$

$$f(x) - f(y) \geq f'(y)(x-y) \quad (3)$$

$$f(y) - f(x) \leq f'(y)(y-x) \quad (3)$$ para que el lado derecho de las desigualdades sea igual

$$f'(x)(y-x) \leq f(y) - f(x) \leq f'(y)(y-x) \quad (4)$$ lado derecho de las 2 y 3 en medio.

$$0 \leq f(y) - f'(x)(y-x) \leq (f'(y) - f'(x))(y-x)$$ En 4, resta $f'(x)(y-x)$

$$0 \leq \frac{f(y) - f(x) - f'(x) f(y-x)}{(y-x)^2} \leq \frac{f'(y) - f'(x)}{y-x}$$ En 4, multiplica $\frac{1}{(y-x)^2}$.

$$f(x) = f(y) + f'(x)(y-x) + f''(z)(y-x)^2, \quad z \in [x, y)$$

$$f(x) \leq f(y) + f'(x)(y-x)$$

En la ecuación 4 queda como

$$0 \leq \frac{f'(y) - f'(x)}{y-x}$$

30 Marzo, 22

2

Recuerda que $\nabla f(\bar{x})=0$ es un minimo, significa que $\bar{x}$ es un punto critico.

Cuando la derivada es 0 cuando $\bar{X}$ porque es un punto critico entonces

$f(y) \geq f(\bar{x})$ $\bar{X}$ es menor que cualquier otro punto.

3

Suponemos que existen X, Y tales que

$f(x) = f(y) \leq f(z)$, para toda $z \in \Omega$

Sea $Z = \dfrac{X+Y}{2}$

$f(z) = f\left(\dfrac{x+y}{2}\right) < \dfrac{1}{2} f(x) + \dfrac{1}{2} f(y)$

$f(z) < f(y)$ Asumiendo que X y Y son igual la ~~Función~~ solución no es única y es estrictamente convexa

Consider the following exercises on SVM:

Problem 2: Basic Exercises on SVM

Prove (using math) the following claims:

1. The L_1 formulation of the SVM for classification is convex.
2. The L_1^ϵ formulation of the SVM for regression is convex.
3. The L_2 formulation of the SVM for classification is strongly convex.
4. The L_2^ϵ formulation of the SVM for regression is strongly convex.

Problema 2

1 Considera L_1 SVM

$$\min_{w,\xi} \mathcal{P}(w,\xi) = \frac{1}{2} w^T w + C \sum_{k=1}^{N} \xi_k$$

$$\text{s.t. } y_k[w^T \phi(x_k) + b] \geq 1 - \xi_k, \quad k = 1, 2, N$$

$$\xi_k \geq 0, \quad k = 1, \dots, N$$

$\frac{1}{2} w^T w$ es una norma y todas las normas son convexas por regla de la inequalidad triangular y homogeniedad.

Sea $h(x)$ la función que describe la norma

$$h(\theta x + (1-\theta) y) \leq h(\theta x) + h((1-\theta) y)$$

$$= \theta h(x) + (1-\theta) h(y)$$

$C\sum_{k=1}^{N} \xi_k \alpha_k = F(\alpha)$ sea combinación lineal que también es convexa

$\alpha = 1$ $\quad F(\theta x + (1-\theta)y) \leq \sum_k \alpha_k \xi_k (\theta x + (1-\theta)y)$

La restricción también es convexa $\quad = \sum \alpha_k (\xi_k(\theta x) + \xi_k((1-\theta)y))$

$= \sum \alpha_k \theta \xi_k(x) + (1-\theta)\xi_k(y))$

$= \theta F(x) + (1-\theta) F(y)$

La suma de funciones convexas es convexa siempre:

$$P(w, \xi) = \underbrace{\tfrac{1}{2} w^T w}_{f} + C \underbrace{\sum_{k=1}^{N} \xi_k}_{g}$$

$$(f+g)(\theta x + (1-\theta)y) = f(\theta x + (1-\theta)y) + g(\theta x + (1-\theta)y)$$

$$\leq \theta f(x) + (1-\theta) f(y) + \theta g(x) + (1-\theta) g(y)$$

$$= \theta (f+g)(x) + (1-\theta)(f+g)(y)$$

Puedes ver la función $P(w, \xi) = f(w, \xi) + g(w, \xi)$

siendo $f(w, \xi) = w$, $g(w, \xi) = \xi$. Entonces

$f(w, \xi) + g(w, \xi) = w + \xi$.

30 Marzo, 22.

2

Considera la formulación SVR L_1

$$\min P_\epsilon(w,b,\xi,\xi^*) = \frac{1}{2} w^T w + C\sum(\xi_k + \xi_k^*)$$

$$\text{s.t. } y_k - w^T \phi(x_k) - b \leq \epsilon + \xi_k, \; k=1,\dots,N$$

$$w^T\phi(x_k) + b - y_k \leq \epsilon + \xi_k^*, \; k=1,\dots,N$$

$$\xi_k, \xi_k^* \geq 0, \; k=1,\dots,N$$

siguiendo la misma formulación del problema anterior

$\frac{1}{2} W^T W$ es convexa.

La combinación lineal $C\sum(\xi_k + \xi_k^*)$ es convexa.

La suma $P(w,b,\xi,\xi^*) = \frac{1}{2} w^T w + C\sum(\xi_k + \xi_k^*)$ es convexa.

3.- Sea L_2 SVM

$$\min P(w,\xi) = \frac{1}{2} w^T w + \frac{C}{2}\sum_{k=1}^{N} \xi_k^2$$

sujeto a

$$y_k[w^T\phi(x_k) + b] \geq 1 - \xi_k, \; k=1,\dots,N$$

Una función fuertemente convexa implica que es también estrictamente convexa y convexa.

F es convexa si y solo si $f''(x) \geq 0$ para toda x

F es estrictamente convexa si $f''(x) > 0$ para toda x

f es fuertemente convexa si y solo si $f''(x) \geq m > 0$ para toda x

Funciones fuertemente convexas tienen mínima únicas

siendo m un parametro $m > 0$, la siguiente inecualidad para todos los puntos x, y en su dominio

$$(\nabla f(x) - \nabla f(y))^T (x-y) \geq m\|x-y\|_2^2$$

En conclusión, sabemos que las funciones L2 son estrictamente convexas porque solo tienen un mínimo, pero falta determinar si son fuertemente convexas dada la nueva definición.

Table 2.5 Convexity (Concavity) of objective functions. Adapted from [58, p. 92, ©IEEE 2002]

	Hard margin	L1 soft margin	L2 soft margin
Primal	Strictly convex $(\mathbf{w}, b)$	Convex $(\mathbf{w}, b, \boldsymbol{\xi})$	Strictly convex $(\mathbf{w}, b, \boldsymbol{\xi})$
Dual	Concave $(\boldsymbol{\alpha})$	Concave $(\boldsymbol{\alpha})$	Strictly concave $(\boldsymbol{\alpha})$

Shigeo Abe. Support Vector Machines for Pattern Classification, 2 Ed. Springer-Verlag London, 2010. ISBN 978-1-84996-097-7. URL https://www.springer.com/gp/book/9781849960977.