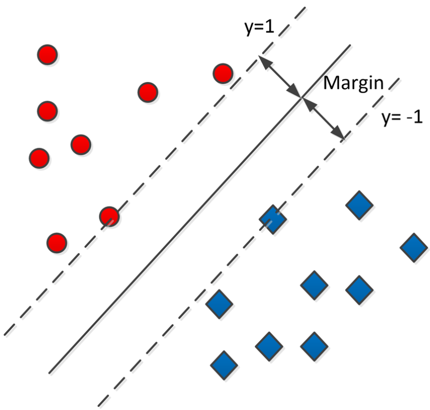
Домашнее задание 9, ККТ

Deadline - 15.11.2024 в 23:59

Основная часть

Задача 1. (2 балла) Рассмотрим задачу о разделении двух подмножеств $\mathbb{R}^d$ гиперплоскостью. Считаем, что даны точки $x_i \in \mathbb{R}^d$ и метки класса $y_i \in \{-1, 1\}$. Мы хотим построить линейную модель с параметрами θ, θ_0, для которой

$$x_i^\top \theta + \theta_0 > 0 \quad \iff \quad y_i > 0$$



Заметим, что наши классы могут быть линейно разделимы с помощью разных гиперплоскостей. Чтобы получить "наилучшую" разделяющую гиперплоскость, потребуем более сильное условие

$$y_i(x_i^\top \theta + \theta_0) \geq 1$$

которое означает, чтобы наша линейная модель не просто верно предсказывает знак, но и имеет некоторую степень робастности, т.е. устойчивости к незначительным изменениям признаков. Но проблема с несколькими плоскостями все еще остается – для любой разделяющей прямой можно раздуть ее коэффициенты, чтобы новое условие выполнялось. Поэтому будем минимизировать ℓ_2 норму параметра θ:

$$\min_{\theta,\theta_0} \frac{1}{2}\|\theta\|_2^2$$
$$\text{s.t. } y_i(x_i^\top \theta + \theta_0) \geq 1$$

Полученная задача подходит для данных, которые полностью линейно отделимы. Теперь перейдем к более общему случаю, для этого ослабим условия, добавив штраф с некоторой константой ρ:

$$\min_{\theta,\theta_0,\xi} \frac{1}{2}\|\theta\|_2^2 + \rho\sum_{i=1}^{n}\xi_i$$
$$\text{s.t. } \xi_i \geq 0,\ y_i(x_i^\top \theta + \theta_0) \geq 1 - \xi_i$$

Выпишите условия ККТ для этой задачи. Выразите параметр θ через оптимальное решение двойственной задачи. Покажите, что оптимальное решение не изменится, если оставить только точки x_i, для которых $y_i(x_i^\top \theta + \theta_0) = 1 - \xi_i$ (*опорные векторы*).

Задача 2. (3 балла) Пусть f – сильно выпуклая дифференцируемая на $\mathbb{R}^n$ функция, матрица $A \in \mathbb{R}^{n\times d}$ такова, что все ее миноры имеют полный ранг (например, это происходит с вероятностью 1, если элементы матрицы как случайные величины имеют совместную плотность).

Докажите, что задача с ℓ_1 регуляризацией

$$\min_{x\in\mathbb{R}^d} \Big[f(Ax) + \|x\|_1 \Big]$$

имеет единственное решение, при чем среди его компонент не более $\min(n, d)$ ненулевых.

Дополнительная часть

Задача 1. (5 баллов) Для задачи

$$\min_{X\in\mathbb{S}^n_{++}} \Big[\mathrm{tr}X - \log\,\det X \Big]$$
$$\text{s.t. } Xs = y,$$

где $\langle s, y \rangle = 1$, с помощью теоремы о достаточности условий ККТ докажите, что решение задачи единственно и записывается в виде

$$X^* = I + yy^\top - \frac{ss^\top}{s^\top s} \ .$$