Машинное обучение.
Домашнее задание №7

Задача 1. Рассмотрим задачу с двумя классами и линейно разделимой обучающей выборкой. Задача оптимизации, соответствующая построению оптимальной разделяющей гиперплоскости, имеет вид

$$\begin{cases} \frac{1}{2}\|w\|^2 \to \min\limits_{w,b} \\ y_i\left(\langle w, x_i\rangle + b\right) \geqslant 1, \quad i = 1, \ldots, \ell. \end{cases}$$

Докажите, что оптимальная разделяющая гиперплоскость не изменится, если единицы в правых частях всех условий-неравенств одновременно заменить произвольным числом $\gamma > 0$.

Задача 2. Разберите вывод двойственной задачи к SVM. Убедитесь, что можете ответить на следующие вопросы:

1. Почему условия Куна-Таккера для данной задачи являются необходимыми и достаточными?
2. Почему, решив первые три уравнения из условий Куна-Таккера и подставив решения в лагранжиан, мы получим двойственную функцию? Получим ли мы такой же результат, если найдем двойственную функцию по определению?
3. Почему в двойственную задачу добавляют ограничения, полученные из условий Куна-Таккера ($0 \leqslant \lambda_i \leqslant C$ и $\sum_{i=1}^{\ell} \lambda_i y_i = 0$)?
4. Почему про двойственную функцию можно сказать, что в ней стоит L_1-норма штрафов, хотя там стоит сумма без модулей ($\sum_i \xi_i$)?
5. Почему про линейный классификатор, настроенный методом опорных векторов, можно сказать, что он разрежен по объектам?
6. Какую интерпретацию имеют двойственные переменные λ_i?
7. Как, решив двойственную задачу, восстановить решение прямой (w, b, ξ)?

Задача 3. Рассмотрим задачу SVM, в которой сумма штрафов ξ_i заменена на сумму их квадратов:

$$\begin{cases} \frac{1}{2}\|w\|^2 + C\sum\limits_{i=1}^{\ell}\xi_i^2 \to \min\limits_{w,b,\xi} \\ y_i\left(\langle w, x_i\rangle + b\right) \geqslant 1 - \xi_i, \quad i = 1, \ldots, \ell, \\ \xi_i \geqslant 0, \quad i = 1, \ldots, \ell. \end{cases}$$

Постройте двойственную к ней.