CS446: Machine Learning, Fall 2017, Homework 3

Name: Lanxiao Bai (lbai5)

Worked individually

Problem 1

Solution:

$$w_t = w_0 + \sum_{i=1}^{t} a_i y_{i-1} x_{i-1} = \sum_{i=1}^{t} a_i y_{i-1} x_{i-1}$$

in which

$$a_i = \begin{cases} 1 & sgn(w_{i-1} \cdot x_i) \neq y_i \\ 0 & \textbf{otherwise} \end{cases}$$

Problem 2

Solution: We notice that in kernelized ridge regression, we have

$$w = \sum_{i=1}^{N} \alpha_i \mathbf{x}_i$$

in which $\alpha_i > 0$ for support vectors.
So the form of two algorithm is the same, while in kernelized perceptron the coefficent a_i is at most 1. Which means that perceptron does not maximize the margin.

Problem 3

Solution:

$$\hat{y} = sgn(w \cdot x) = sgn(\sum_{i=1}^{t} \alpha_i x_{i-1} x) = sgn(\sum_{i=1}^{t} \alpha_i x_{i-1}^T x)$$

in which

$$\alpha_i = \begin{cases} y_i & sgn(w_{i-1} \cdot x_i) \neq y_i \\ 0 & \textbf{otherwise} \end{cases}$$

Problem 4

Solution:

$$\hat{y} = sgn(w \cdot \phi(x)) = sgn(\sum_{i=1}^{t} \alpha_i \phi(x_{i-1}) \phi(x)) = sgn(\sum_{i=1}^{t} \alpha_i \kappa(x, x_{i-1}))$$

Problem 5

Solution: The kernelized algorithm we got is

1. Initialize $\alpha_0 = 0$, where $\alpha_0 \in \mathbb{R}^d$
2. For t in $1 \cdots N$:
   (a) Predict $\hat{y} = sgn(\sum_{i=1}^{t} \alpha_i \kappa(x, x_{i-1}))$
   (b) If $\hat{y}_t \neq y_t$, update the weight:
      i. $\alpha_t = \alpha_{t-1} + y_t \phi(x_t)$

References