Topics in Machine Learning Homework 1

By B00401062 羅文斌

2.1

To prove $f(n) = \theta_1 x_1 + \theta_2 x_2 + \ldots \theta_n x_n \in C$, given $\theta_1 + \theta_2 + \ldots \theta_n = 1$.

When $n = 2, \theta_1 + \theta_2 = 1$. Therefore, $f(2) = \theta_1 x_1 + \theta_2 x_2 \in C$.

Suppose when $n = k-1, \theta_1 + \theta_2 + \ldots \theta_{k-1} = 1$ also satisfies $f(k-1) = \theta_1 x_1 + \theta_2 x_2 + \ldots \theta_{k-1} x_{k-1} \in C$.

Then when $n = k$, $f(k)$ can be expressed as the sum of $\alpha f(k-1) + (1-\alpha) \cdot x_k$, where

$\alpha \geq 0, f(k-1) \in C, x_k \in C$.

Because $f(n) \in C$ holds when $n = 2$, the target function $f(k) = \alpha f(k-1) + (1-\alpha) \cdot x_k \in C$ will also hold.

2.5

Suppose there is a point x_1 in hyperplane $P_1 = \{x \in \mathrm{R}^n | a^{\mathrm{T}} x = b_1\}$, and a point x_2 in hyperplane

$P_2 = \{x \in \mathrm{R}^n | a^{\mathrm{T}} x = b_2\}$.

The distance d between two hyperplanes is the length of the projection of $x_2 - x_1$ on a.

Therefore, $d = |\frac{a^{\mathrm{T}} \cdot (x_2 - x_1)}{||a||_2}| = \frac{|b_2 - b_1|}{||a||_2}$.

2.7

$||x-a||_2 \leq ||x-b||_2$ can be expressed as vector form: $(x-a)^{\mathrm{T}} \cdot (x-a) \leq (x-b)^{\mathrm{T}} \cdot (x-b)$

$x^{\mathrm{T}} \cdot x - x^{\mathrm{T}} \cdot a - a^{\mathrm{T}} \cdot x + a^{\mathrm{T}} \cdot a \leq x^{\mathrm{T}} \cdot x - x^{\mathrm{T}} \cdot b - b^{\mathrm{T}} \cdot x + b^{\mathrm{T}} \cdot b$ ...(1)

Because $x^{\mathrm{T}} \cdot a = a^{\mathrm{T}} \cdot x$ and $x^{\mathrm{T}} \cdot b = b^{\mathrm{T}} \cdot x$, (1) can be simplified as

$-2 \cdot a^{\mathrm{T}} \cdot x + a^{\mathrm{T}} \cdot a \leq -2 \cdot b^{\mathrm{T}} \cdot x + b^{\mathrm{T}} \cdot b$ ...(2)

Move terms with x to one side, and term without x to the other side. (2) becomes

$2 \cdot (b^{\mathrm{T}} - a^{\mathrm{T}}) \cdot x \leq b^{\mathrm{T}} \cdot b - a^{\mathrm{T}} \cdot a$ ...(3)

(3) can be furthur reduced to $(b-a)^{\mathrm{T}} \cdot x \leq \frac{1}{2}(b-a)^{\mathrm{T}} \cdot (b+a)$. Therefore,

$c = (b-a), d = \frac{1}{2} \cdot (b-a)^{\mathrm{T}} \cdot (b+a)$.

Because $\frac{1}{2}(b+a)$ is the midpoint of a and b, the halfspace $(b-a)^{\mathrm{T}} \cdot x \leq \frac{1}{2}(b-a)^{\mathrm{T}} \cdot (b+a)$ contains all the points falling on the left-hand side of the hyperplane $(b-a)^{\mathrm{T}} \cdot x = \frac{1}{2}(b-a)^{\mathrm{T}} \cdot (b+a)$ as shown in the figure.

The hyperplane $(b-a)^{\mathrm{T}} \cdot x = \frac{1}{2}(b-a)^{\mathrm{T}} \cdot (b+a)$ is a plane perpendicular to $(b-a)$ and passing through $\frac{1}{2}(b+a)$.