Kernel methods in machine learning - Penn and paper 3

Christian Segercrantz 481056

April 26, 2022

Task 1

$$
\begin{aligned}
f(\theta x + (\theta - 1)y) \leq & \theta f(x) + (1 - \theta)f(y) && (1)\\
f(z) := & ||z|| && (2)\\
f(\theta x + (\theta - 1)y) = & ||\theta x + (\theta - 1)y|| && (3)\\
\leq & ||\theta x|| + ||(\theta - 1)y||, \qquad |\text{Triangel inequality} && (4)\\
= & \theta ||x|| + (\theta - 1)||y||, \qquad |\text{Absolutely scaleable} && (5)\\
= & \theta f(x) + (1 - \theta)f(y) && (6)\\
\implies f(\theta x + (\theta - 1)y) \leq & \theta f(x) + (1 - \theta)f(y), \qquad f(z) := ||z|| && (7)
\end{aligned}
$$

Task 2

$$
\begin{aligned}
\theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 = & (\theta_1 + \theta_2)\left(\frac{\theta_1}{\theta_1 + \theta_2}x_1 + \frac{\theta_2}{\theta_1 + \theta_2}x_2\right) + \theta_3 x_3 && (8)\\
= & (1 - \theta_3)\left(\frac{\theta_1}{\theta_1 + \theta_2}x_1 + \frac{\theta_2}{\theta_1 + \theta_2}x_2\right) + \theta_3 x_3, \qquad |\theta_1 + \theta_2 = 1 - \theta_3 && (9)\\
x' := & \frac{\theta_1}{\theta_1 + \theta_2}x_1 + \frac{\theta_2}{\theta_1 + \theta_2}x_2 && (10)\\
= & (1 - \theta_3)x' + \theta_3 x_3 && (11)\\
& && (12)
\end{aligned}
$$

We now show that x' is part of the convex set

$$
\begin{aligned}
\theta_1' := & \frac{\theta_1}{\theta_1 + \theta_2} && (13)\\
\theta_2' := & \frac{\theta_2}{\theta_1 + \theta_2} && (14)\\
\theta_1' + \theta_2' = & 1, \quad \theta_1', \theta_2' \geq 0 && (15)\\
\implies \theta_1' x_1 + \theta_2' x_2 = & \theta_1' x_1 + (1 - \theta_1')x_2 && (16)
\end{aligned}
$$

As we can see, x' is part of the covex set C and thus, we can see that the points $(1 - \theta_3)x' + \theta_3 x_3$ are also part of the convex set.

Task 3

$$
\begin{aligned}
\min{}_{w,\xi,b} \quad & \frac{1}{2}||w||^2 + C\sum_{i=1}^{m} \xi_i && (17)\\
\text{subject to:} \quad & y_i(w^\top \phi(x_i) + b) \geq 1 - \xi_i && (18)\\
& \xi_i \geq 0, i = 1, ..., m && (19)\\
& && (20)
\end{aligned}
$$

Task 3a

The Lagrangiang function becomes

$$L(w,\xi,b,\lambda,\nu) = \frac{1}{2}||w||^2 + C\sum_{i=1}^{m}\xi_i + \sum_{i=1}^{m}\lambda_i\left(1-\xi_i-y_i(w^\top\phi(x_i)+b)\right) + \sum_{i=1}^{m}-\nu_i\xi_i \tag{21}$$

$$\tag{22}$$

2

Task 3b

$$\frac{\partial}{\partial\xi}L(w,\xi,b,\lambda,\nu) = \frac{\partial}{\partial\xi}\left(\frac{1}{2}||w||^2 + C\sum_{i=1}^{m}\xi_i + \sum_{i=1}^{m}\lambda_i\left(1-\xi_i-y_i(w^\top\phi(x_i)+b)\right) + \sum_{i=1}^{m}-\nu_i\xi_i\right) \tag{23}$$

$$= Cm - \sum_{i=1}^{m}(\lambda_i+\nu_i) \tag{24}$$

$$Cm = \sum_{i=1}^{m}\lambda_i+\nu_i \tag{25}$$

$$\frac{\partial}{\partial w}L(w,\xi,b,\lambda,\nu) = \frac{\partial}{\partial w}\left(\frac{1}{2}||w||^2 + C\sum_{i=1}^{m}\xi_i + \sum_{i=1}^{m}\lambda_i\left(1-\xi_i-y_i(w^\top\phi(x_i)+b)\right) + \sum_{i=1}^{m}-\nu_i\xi_i\right) \tag{26}$$

$$= w - \sum_{i=1}^{m}\lambda_i y_i\phi(x_i) \tag{27}$$

$$w = \sum_{i=1}^{m}\lambda_i y_i\phi(x_i) \tag{28}$$

$$\frac{\partial}{\partial b}L(w,\xi,b,\lambda,\nu) = \frac{\partial}{\partial b}\left(\frac{1}{2}||w||^2 + C\sum_{i=1}^{m}\xi_i + \sum_{i=1}^{m}\lambda_i\left(1-\xi_i-y_i(w^\top\phi(x_i)+b)\right) + \sum_{i=1}^{m}-\nu_i\xi_i\right) \tag{29}$$

$$= \sum_{i=1}^{m}\lambda_i \tag{30}$$

$$\implies \lambda_i = 0 \tag{31}$$

$$\implies Cm = \sum_{i=1}^{m}\nu_i \tag{32}$$

$$\tag{33}$$