# DSC 140A - Discussion 04

**Problem 1.**

Using the definition, show that the function $f(\vec{w}) = a\vec{x} \cdot \vec{w} - b$ is convex as a function of $\vec{w}$, where $a, b \in \mathbb{R}$ and $\vec{x}, \vec{w} \in \mathbb{R}^d$.

**Solution:**

$$\nabla f(\vec{w}) = a\vec{x}$$
$$H_f(\vec{w}) = \mathbf{0}_{d,d} \succeq 0$$

In the Hessian, $H_f(\vec{w})$, all values are 0. We can see that the eigenvalues will also be 0. Therefore, $f(\vec{w})$ is convex.

**Problem 2.**

Let $f_1(\vec{w})$ and $f_2(\vec{w})$ be convex functions from $\mathbb{R}^d$ to $\mathbb{R}$. Define

$$f(\vec{w}) = \max\{f_1(\vec{w}), f_2(\vec{w})\}.$$

Show that $f(\vec{w})$ is convex.

**Solution:**

Take $t \in [0, 1]$.

$$\begin{aligned} f(t\vec{w}_1 + (1-t)\vec{w}_2) &= \max\{f_1(t\vec{w}_1 + (1-t)\vec{w}_2), f_2(t\vec{w}_1 + (1-t)\vec{w}_2)\} \\ &\leq \max\{tf_1(\vec{w}_1) + (1-t)f_1(\vec{w}_2), tf_2(\vec{w}_1) + (1-t)f_2(\vec{w}_2)\} \quad \text{(due to convexity of } f_1, f_2\text{)} \\ &\leq \max\{tf_1(\vec{w}_1), tf_2(\vec{w}_1)\} + \max\{(1-t)f_1(\vec{w}_2), (1-t)f_2(\vec{w}_2)\} \\ &= t\max\{f_1(\vec{w}_1), f_2(\vec{w}_1)\} + (1-t)\max\{f_1(\vec{w}_2), f_2(\vec{w}_2)\} \\ &= tf(w_1) + (1-t)f(w_2) \end{aligned}$$

**Problem 3.**

Recall that the Perceptron loss is:

$$L_{\text{perc}}(\vec{w}, \vec{x}, y) = \begin{cases} 0, & \text{if } \operatorname{sign}(\vec{w} \cdot \vec{x}) = y \text{ (correctly classified)}, \\ |\vec{w} \cdot \vec{x}|, & \text{if } \operatorname{sign}(\vec{w} \cdot \vec{x}) \neq y \text{ (misclassified)}. \end{cases}$$

Using the trick that $-y\,\vec{w} \cdot \vec{x} = |\vec{w} \cdot \vec{x}|$ in the case of misclassification, this can be be written in the equivalent form:

$$L_{\text{perc}}(\vec{w}, \vec{x}, y) = \max\{0, -y\,\vec{w} \cdot \vec{x}\}$$

Argue that the perceptron loss is convex as a function of $\vec{w}$.

**Solution:** We'll argue that the loss is the maximum of two convex functions, which (by Problem 2 above) is convex.

Let $f_1(\vec{w}) = 0$ and $f_2(\vec{w}) = -y\,\vec{w} \cdot \vec{x}$. Recognize that

$$L_{\text{perc}}(\vec{w}, \vec{x}, y) = \max\{f_1(\vec{w}),\ f_2(\vec{w})\}$$

$f_2$ is convex by the result of Problem 1 (with $a = y$ and $b = 0$). $f_1$ is constant, and trivially convex. Therefore $L_{\text{perc}}$ is convex by the result of Problem 2.