Worksheet 6: Convexity and Gradients

Name: Due November 3, 2023

1. (Properties of convexity) Show that $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) := \log(1 + \exp(x)) = \big(\sigma(-t)\big)^{-1}$ is convex.

2. (Use the definition) Suppose that $f : \mathbb{R}^d \to \mathbb{R}$ decomposes as composition

$$\begin{tikzcd} \mathbb{R}^d \arrow[r, "\eta"] \arrow[rr, bend left, "f"] & \mathbb{R} \arrow[r, "g"] & \mathbb{R} \end{tikzcd}$$

   where $\eta(w) = \langle w, x \rangle + y$. Show that if g is convex, then so is f.

3. (Use the definition) Show that if $f_1, \dots, f_k$ are convex, then $g = \sum_{j=1}^{k} \alpha_j f_j$ is convex for $\alpha_j \geq 0$.

4. (Convexity of LogReg) Conclude that ERM problem for $\mathcal{H} = \big\{\sigma(\alpha x - \beta) : \alpha, \beta \in \mathbb{R}\big\}$ and $c_{\tilde{y}}(x, y) = -y \log(\tilde{y}(x)) - (1 - y)\log(1 - \tilde{y}(x))$ is convex.

5. (Computing Gradients) For $\mathcal{H}_n := \left\{ \sigma(p(x)) : \ p \text{ is degree } n \text{ polynomial, } p(x) = \sum_{j=0}^{n} \alpha_j x^j \right\}$, express $\nabla_\alpha e_S(\tilde{y}_\alpha)$, where $e_S(\tilde{y}) = \frac{1}{m} \sum_{j=1}^{m} c_{\tilde{y}}(x_j, y_j)$ denotes the empirical (risk) expectation (implicit: $S = \big((x_1, y_1), \dots, (x_m, y_m)\big)$) of cost $c_{\tilde{y}}(x, y) = (\tilde{y}(x) - y)^2$.

6. (Computing Gradients) For same $\mathcal{H}_n$ as above, express $\nabla_\alpha e_S(\tilde{y}_\alpha)$, this time $c_{\tilde{y}}(x, y) = -y \log(\tilde{y}(x)) - (1 - y)\log(1 - \tilde{y}(x))$.