WEEKLY PARTICIPATION 3: CONVEXITY OF THE OLS PROBLEM

I made the claim that all the optimization problems we have seen for fitting ML models so far are in fact convex optimization problems. Let's work through the general template for proving this by showing that the OLS problem,

$$\min_{\boldsymbol{\beta} \in \mathbb{R}^d} \frac{1}{n} \|\boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{y}\|_2^2,$$

is convex.

The following properties are basic applications of the definition of convexity: f is convex on its domain if for any $\boldsymbol{x}, \boldsymbol{y} \in \text{dom}(f)$ and any $\alpha \in [0,1]$ it is the case that

$$f(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}).$$

Use this definition to work through the following steps (the proof of each is a line or two of algebra):

(1) Argue that *nonnegative* multiples of convex functions are convex.

(2) Argue that affine functions (functions of the form $f(\boldsymbol{x}) = \langle \boldsymbol{a}, \boldsymbol{x} \rangle + b$) are convex.

(3) Argue that if g is convex and f is affine, then the composition $g(f(\cdot))$ is a convex function.

(4) Argue that the sum of convex functions is convex.

Use the facts just established, along with the fact that x^2 is convex, to argue that the OLS problem is a convex optimization problem.