
Math 173
Problem Set 6
Monday, November 12, 2018

1 Let $\mathcal{V}$ be a real vector space. Prove that a function $f : \mathcal{V} \to \mathbb{R}$ is convex if and only if the epigraph

$$\mathrm{Epi}(f) = \big\{(x, y) : x \in \mathcal{V},\, y \geq f(x)\big\} \subset \mathcal{V} \oplus \mathbb{R}$$

is convex.

$\blacksquare$

2 Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable. Prove that f is convex if and only if $f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle$ for every $x_0 \in \mathbb{R}^n$. That is, convex functions are those who are always bigger than their linear Taylor approximation. Use this to quickly prove that for a differentiable convex function $f : \mathbb{R}^n \to \mathbb{R}$, $\nabla f(x) = 0$ implies that $f(x) = \inf_{y \in \mathbb{R}^n} f(y)$. *Hint:* Write $\langle \nabla f(x_0), x - x_0 \rangle$ as a limit.

$\blacksquare$

3 Recall that the set $\mathbb{S}^n_+$ of positive semi-definite matrices is convex.

(a) Prove the function $A \mapsto \log \det A$ defined on $\mathbb{S}^n_+$ is concave.

(b) Prove the function $A \mapsto \operatorname{tr} \exp A$ defined on $\mathbb{S}^n_+$ is convex.

(c) Prove the function $A \mapsto \det A$ defined on $\mathbb{S}^n_+$ is *not* convex.

$\blacksquare$

4 Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is convex, and that the eigenvalues $\lambda_i(x)$ of the Hessian $\nabla^2 f(x)$ are uniformly bounded between $0 < \ell < \lambda_i(x) < L$ across all of $\mathbb{R}^n$. Consider the following algorithm, known as *gradient descent*, which finds an approximation x_T to a global minimizer x^* of f.

Data: An arbitrary starting point $x_0 \in \mathbb{R}^n$, a step size $0 < \alpha < 2/L$, and a maximum number of iterations $T \in \mathbb{N}$.
Result: An approximate global minimizer $x_T \in \mathbb{R}^n$ to f.
for $t = 1, 2, \ldots, T$ **do**
| $x_t = x_{t-1} - \alpha \nabla f(x_{t-1})$
end

Prove that if $\frac{1-q}{\ell} \leq \alpha \leq \frac{1+q}{L}$ for some $0 < q < 1$ then the global minimizer x^* is unique and

$$\|x_T - x^*\|_2 \leq \frac{\alpha q^T}{1-q} \|\nabla f(x_0)\|_2 \leq \frac{2}{L} \frac{q^T}{1-q} \|\nabla f(x_0)\|_2.$$

That is to say that if we want $\|x_T - x^*\|_2 \leq \epsilon$ then we can just set

$$T \geq \log(\tfrac{1}{q})^{-1} \log\left(\frac{2\|\nabla f(x_0)\|_2}{\epsilon L(1-q)} \right)$$

to guarantee this[a]. *Hint* Define $F(x) = x - \alpha \nabla f(x)$. Prove that the eigenvalues $\gamma_i(x)$ of the Jacobian $\nabla F(x)$ are always bounded above by q: $|\gamma_i(x)| \leq q$. Apply the Banach Fixed Point Theorem from analysis.

[a] It turns out that this is the best convergence rate you could hope for with this class of functions f. [Arjevani2016]

∎