CS257 Linear and Convex Optimization
Homework 7

Due: November 9, 2020

November 2, 2020

1. Consider the quadratic function $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{x}$, where $\boldsymbol{A} \succ \boldsymbol{O}$ is positive definite. Let $\boldsymbol{x}^*$ be the global minimum of f. Suppose that we can only compute the gradient up to some error. We implement the following gradient descent algorithm with constant step size t,

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - t\left(\nabla f(\boldsymbol{x}_k) + \boldsymbol{e}_k\right),$$

where $\boldsymbol{e}_k \in \mathbb{R}^d$ is some additive error term (not to be confused with the k-th unit vector), satisfying a bound of the form $\|\boldsymbol{e}_k\|_2 \leq E$.

(a). Let the eigenvalues of $\boldsymbol{A}$ be $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n > 0$. If there is no error, i.e. $\boldsymbol{e}_k = \boldsymbol{0}$, what condition should t satisfy to ensure that $\boldsymbol{x}_k \to x^*$?

(b). Show $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}^*$

(c). Show $\boldsymbol{x}_{k+1} - \boldsymbol{x}^* = (\boldsymbol{I} - t\boldsymbol{A})(\boldsymbol{x}_k - \boldsymbol{x}^*) - t\boldsymbol{e}_k$

(d). Assume $0 < t < 1/\lambda_1$ for (d)-(f). Show

$$\|\boldsymbol{x}_{k+1} - \boldsymbol{x}^*\|_2 \leq (1 - \lambda_n t) \cdot \|\boldsymbol{x}_k - \boldsymbol{x}^*\|_2 + tE.$$

Hint: Use $\|\boldsymbol{Q}\boldsymbol{x}\| \leq \lambda_{\max}(\boldsymbol{Q})\|\boldsymbol{x}\|$ for $\boldsymbol{Q} \succeq \boldsymbol{O}$ (see slide 14 of Lecture 8).

(e). Show by induction that

$$\|\boldsymbol{x}_k - \boldsymbol{x}^*\|_2 \leq (1 - \lambda_n t)^k \|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_2 + \frac{1 - (1 - \lambda_n t)^k}{\lambda_n} E$$

(f). Conclude

$$\limsup_{k \to \infty} \|\boldsymbol{x}_k - \boldsymbol{x}^*\|_2 \leq \frac{E}{\lambda_n}$$

i.e. $\boldsymbol{x}_k$ "converges" to a ball of radius E/λ_n centered at $\boldsymbol{x}^*$.

2. Suppose $f(\boldsymbol{x})$ is differentiable and α-strongly convex, and $g(\boldsymbol{x})$ is β-smooth. Prove that the function $h(\boldsymbol{x}) = f(\boldsymbol{x}) - g(\boldsymbol{x})$ is convex if $\alpha \geq \beta$.

3. Consider $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}_1^2 + (1 - \epsilon)\boldsymbol{x}_1\boldsymbol{x}_2 + \frac{1}{2}\boldsymbol{x}_2^2$, where $\epsilon \in (0, 1)$. Find the condition number of the Hessian of f. What happens to the condition number as $\epsilon \to 0$?