

# Problem Set 3 for EE227C (Spring 2018): Convex Optimization and Approximation

Instructor: Moritz Hardt

Email: `hardt+ee227c@berkeley.edu`

Graduate Instructor: Max Simchowitz

Email: `msimchow+ee227c@berkeley.edu`

April 30, 2018

In this homework we will prove the claims necessary for the convergence bound of the barrier method we saw in Lecture 25.

**Note:** This assignment might seem long, but that's due to lots of hints. So, don't be discouraged.

## Background

We collect some useful background material below. Feel free to consult additional resources on this topic.

**Barrier function.** Recall that we defined the modified objective  $f_\epsilon: \mathbb{R}^n \rightarrow \mathbb{R}$  as

$$f_\epsilon(x) = \frac{1}{\epsilon} c^\top x - \sum_{j=1}^m \ln(A_j^\top x - b_j).$$

Here,  $A_j \in \mathbb{R}^n$  denotes the  $j$ -th row of  $A$  in column form. Let  $s_j(x) = A_j^\top x - b_j$  and define the  $m \times m$  diagonal matrix  $S(x) = \text{diag}(s_1(x), \dots, s_m(x))$ .

Denoting by  $\mathbf{1}$  the all ones vector, we have

$$\begin{aligned}\nabla f_\epsilon(x) &= \frac{c}{\epsilon} - A^\top S(x)^{-1} \mathbf{1} = \frac{c}{\epsilon} - \sum_j \frac{A_j}{S_j(x)} \\ \nabla^2 f_\epsilon(x) &= A^\top S(x)^{-2} A = \sum_j \frac{A_j A_j^\top}{S_j(x)^2},\end{aligned}$$

We will assume that the system  $Ax \geq b$  is feasible, and that the rows of  $A$  span  $\mathbb{R}^n$ . The latter implies that  $f_\epsilon$  is strongly convex. We will denote by

$$x_\epsilon^* = \arg \min_x f_\epsilon(x),$$

the unique minimizer for  $f_\epsilon$ . Further, let  $x^*$  denote a minimizer of the original constrained problem  $\min\{c^\top x \mid Ax \geq b\}$ .

**Newton decrement.** The central quantity  $q(x, \epsilon)$  we needed for the analysis is defined as

$$q(x, \epsilon)^2 = \nabla f_\epsilon(x)^\top \nabla^2 f_\epsilon(x)^{-1} \nabla f_\epsilon(x).$$

This is sometimes called *Newton decrement*.

For the remainder of this problem set, we fix  $\epsilon > 0$  and assume that  $x$  is a point satisfying  $Ax > b$ , i.e., it is in the interior of the polytope  $\{x: Ax \geq b\}$ . Further we let  $\bar{x}$  denote the outcome of taking a single Newton step starting from  $x$ ,

$$\bar{x} = x - \nabla^2 f_\epsilon(x)^{-1} \nabla f_\epsilon(x).$$

We can also express  $q(x, \epsilon)$  as a “local norm” defined by the Hessian applied to the Newton step  $\delta = \nabla^2 f_\epsilon(x)^{-1} \nabla f_\epsilon(x)$ . That is,

$$q(x, \epsilon) = \|\delta\|_{\nabla^2 f_\epsilon(x)} = \sqrt{\delta^\top \nabla^2 f_\epsilon(x) \delta}.$$

## 1 Problem 1 (Optional)

In class, we stated the following proposition.

**Proposition 1.** Assume  $Ax > b$  and  $q(x, \epsilon) \leq 1/2$ . Then,  $c^\top x - c^\top x^* \leq O(\epsilon m)$ .

(A) Prove that

$$c^\top x_\epsilon^* - c^\top x^* \leq \epsilon m.$$

This claim does not need the assumptions on  $x$ .

Hint: Use the fact that the gradient of  $f_\epsilon$  vanishes at  $x_\epsilon^*$  in order to get the upper bound  $\sum_j (s_j(x_\epsilon^*) - s_j(x^*)) / s_j(x_\epsilon^*)$ .

(B) Prove that

$$c^\top x - c^\top x_\epsilon^* \leq O(\epsilon m).$$

## Problem 2

In this problem you will show that Newton's method has quadratic convergence provided that  $q(x, \epsilon)$  is small enough.

**Proposition 2.** Assume  $Ax > b$  and  $q(x, \epsilon)^2 < 1/2$ . Then, the Newton iterate  $\bar{x}$  satisfies  $A\bar{x} > b$  and  $q(\bar{x}, \epsilon)^2 < q(x, \epsilon)$ .

(A) Prove that

$$\nabla f_\epsilon(\bar{x}) = - \sum_{j=1}^m \frac{A_j(A_j^\top(\bar{x} - x))^2}{s_j(x)^2 s_j(\bar{x})}.$$

Use the fact that  $\bar{x}$  minimizes the second order approximation of  $f_\epsilon$  at  $x$  and hence  $\nabla f_\epsilon(x) + \nabla^2 f_\epsilon(x)(\bar{x} - x) = 0$ . Write out what this condition means and use it to derive the expression for the gradient at  $\bar{x}$ .

(B) Show that

$$q(x, \epsilon)^2 = \sum_{j=1}^m \frac{(A_j^\top(\bar{x} - x))^2}{s_j(x)^2}$$

(C) Find a norm  $\|\cdot\|$  of the form  $\|x\| = \|Mx\|_2$ , for some matrix  $M$ , for which you can show for every vector  $z$ ,

$$z^\top \nabla f_\epsilon(\bar{x}) \leq \|z\| \cdot q(x, \epsilon)^2.$$

Use the previous steps and Cauchy-Schwartz.

(D) Complete the proof by relating  $q(\bar{x}, \epsilon)$  and  $\sup_z \frac{z^\top \nabla f_\epsilon(\bar{x})}{\|z\|}$ .

## Problem 3

Your goal is to prove the following proposition.

**Proposition 3.** Assume  $Ax > b$  and  $q(x, \epsilon) \leq 1/2$ . Then, for  $\bar{\epsilon} = \epsilon/(1 + \delta)$  with  $\delta = \frac{1}{5}m^{-1/2}$ , we have  $q(\bar{x}, \bar{\epsilon}) \leq 1/2$ .

(A) Show that

$$\nabla^2 f_{\bar{\epsilon}}(\bar{x})^{-1} \nabla f_{\bar{\epsilon}}(\bar{x}) = \nabla^2 f_\epsilon(\bar{x})^{-1} \nabla f_\epsilon(\bar{x}) + \delta \nabla^2 f_\epsilon(\bar{x})^{-1} c / \epsilon.$$

(B) From the previous step, conclude that

$$q(\bar{x}, \bar{\epsilon}) \leq q(\bar{x}, \epsilon) + \delta \|\nabla^2 f_\epsilon(\bar{x})^{-1} c / \epsilon\|_{\nabla^2 f_\epsilon(\bar{x})}$$

(C) Show that

$$\|\nabla^2 f_\epsilon(\bar{x})^{-1}c/\epsilon\|_{\nabla^2 f_\epsilon(\bar{x})} \leq \sqrt{m} + O(1).$$

Hint: Relate the expression in the left hand side to a projection operator applied to the all ones vector.

(D) Complete the proof by putting together the previous steps and applying Proposition 2.