Homework 6

Toffalini Leonardo

Exercise 1

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and define $\tilde{f} : \mathbb{R}^n \to \mathbb{R}$ a $\tilde{f}(x) = f(Ax+b)$ where $A \in \mathbb{R}^{n \times n}$ is an invertible matrix and $b \in \mathbb{R}^n$. Prove that if x_0 moves to x_1 by applying one step of Newton's method with respect to $\tilde{f}$, then $y_0 = Ax_0 + b$ moves to $y_1 = Ax_1 + b$ by applying one step of Newton's method with respect to f.

Solution: Given $x_1 = x_0 - H_{\tilde{f}}^{-1}(x_0) \cdot \nabla \tilde{f}(x_0)$, we need to show that

$$Ax_1 + b = Ax_0 + b - H_f^{-1}(Ax_0 + b) \cdot \nabla f(Ax_0 + b).$$

We know that

$$\nabla \tilde{f}(x_0) = A^T \nabla f(Ax_0 + b)$$

by the chain rule, and likewise

$$H_{\tilde{f}}(x_0) = A^T H_f(Ax_0 + b) A.$$

From this we have

$$\begin{aligned}
x_1 &= x_0 - H_{\tilde{f}}^{-1}(x_0) \cdot \nabla \tilde{f}(x_0) \\
x_1 &= x_0 - \left(A^T H_f(Ax_0 + b) A\right)^{-1} \cdot A^T \nabla f(Ax_0 + b) \\
x_1 &= x_0 - A^{-1} H_f^{-1} A^{-T} \cdot A^T \nabla f(Ax_0 + b) \\
x_1 &= x_0 - A^{-1} H_f^{-1} \cdot \nabla f(Ax_0 + b) \\
Ax_1 &= Ax_0 - H_f^{-1} \cdot \nabla f(Ax_0 + b) \\
Ax_1 + b &= Ax_0 + b - H_f^{-1} \cdot \nabla f(Ax_0 + b) \\
y_1 &= y_0 - H_f^{-1} \cdot \nabla f(y_0)
\end{aligned}$$

□

Exercise 2

Consider the following functions $f : K \to \mathbb{R}$. Check if they satisfy the **NL** condition for some constant $0 < \delta_1 < 1$.

1. $f(x) = -\log\cos(x)$ on $K = (-\pi/2, \pi/2)$.
2. $f(x) = x\log(x) + (1-x)\log(1-x)$ on $K = (0,1)$.
3. $f(x) = -\sum_{i=1}^{n} \log(x_i)$ on $K = \mathbb{R}^n_{>0}$.

Solution:

1.

$$H(x) = f''(x) = \left(\frac{1}{\cos} x \sin x\right)' = \operatorname{tg}' x = \frac{1}{\cos^2 x}.$$

$$(1-3\delta)H(x) = (1-3\delta)\frac{1}{\cos^2 x} \overset{?}{\leq} \frac{1}{\cos^2 y}$$

$$(1-3\delta)\frac{1}{\cos^2 x} \leq 1 < \frac{1}{\cos^2 y}$$

$$1 - 3\delta \leq \cos^2 x$$

Let $\delta = 2/3$, then

$$1 - 3\delta = -1 \leq \cos^2 x \quad \forall x.$$

We did not even have to use the fact that $\|y - x\|_x \leq \delta$.

The same reasoning works for the opposite direction aswell.

2.

$$H(x) = f''(x) = (1 + \log x - \log(1-x) - 1)' = (\log x - \log(1-x))' = \frac{1}{x} + \frac{1}{1-x}.$$

$$\frac{(y-x)^2}{x} + \frac{(y-x)^2}{1-x} \leq \delta$$

$$(1-3\delta)\left(\frac{1}{x} + \frac{1}{1-x}\right) \overset{?}{\leq} \frac{1}{y} + \frac{1}{1-y}$$

$$\frac{1}{1} + \frac{1}{1-0} = 2 < \frac{1}{y} + \frac{1}{1-y}$$

It is enough to show

$$(1-3\delta)\left(\frac{1}{x} + \frac{1}{1-x}\right) = (1-3\delta)\frac{1}{x(1-x)} \leq 2$$

$$1 - 3\delta \leq 2x(1-x)$$

Since $0 \leq 2x(1-x)$ when $x \in (0,1)$, we can choose $\delta = 2/3$, thus

$$1 - 3\delta = -1 \leq 2x(1-x) \quad \forall x \in (0,1).$$

A similar argument can be applied for the opposite direction.

3.

$$H(x) = \nabla\left(\nabla - \sum_{i=1}^{n} \log x_i\right) = \nabla\left(-\frac{1}{x_1}, ..., -\frac{1}{x_n}\right) = \begin{bmatrix} \frac{1}{x_1^2} & 0 & 0 & ... & 0 \\ 0 & \frac{1}{x_2^2} & 0 & ... & 0 \\ 0 & 0 & \frac{1}{x_3^2} & ... & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & ... & \frac{1}{x_n^2} \end{bmatrix}$$

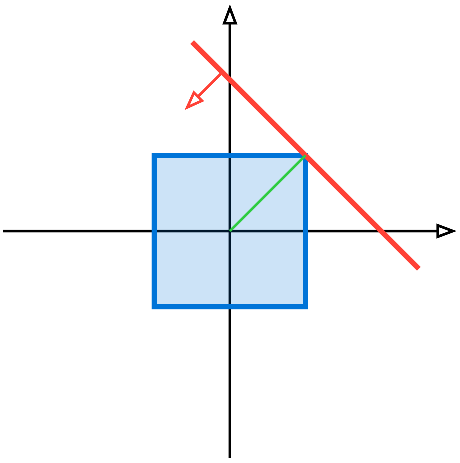
$$\|y - x\|_x = (y-x)^T H(x)(y-x) \leq \delta$$

$\square$

Exercise 3

Let $P = \{x \in \mathbb{R}^2 : |x_1| \leq 1, \quad |x_2| \leq 1\}$ and $c = (1,1) \in \mathbb{R}^2$. Determine the points of the central path Γ_c.

Solution:



The polytope can be given with the following conditions:

$$x_1 \leq 1; \quad -x_1 \leq 1; \quad x_2 \leq 1; \quad -x_2 \leq 1.$$

From this we have the logarithmic barrier function and it's gradient as follows

$$F(x) = -\log(1-x_1) - \log(1+x_1) - \log(1-x_2) - \log(1+x_2)$$

$$\nabla F(x) = \begin{bmatrix} \frac{1}{1-x_1} - \frac{1}{1+x_1} \\ \frac{1}{1-x_2} - \frac{1}{1+x_2} \end{bmatrix}.$$

Let $f_t(x) = t(c \cdot x) + F(x)$, with this we have

$$\nabla f_t(x) = \begin{bmatrix} t + \frac{1}{1-x_1} - \frac{1}{1+x_1} \\ t + \frac{1}{1-x_2} - \frac{1}{1+x_2} \end{bmatrix},$$

and we need this to be the $(0,0)$ vector for all $t > 0$. Since the two equations are the same we only focus on one of them.

$$\begin{aligned} t + \frac{1}{1-x_1} - \frac{1}{1+x_1} &= 0 \\ (1-x_1^2)t + 1 + x_1 - 1 + x_1 &= 0 \\ tx_1^2 - 2x_1 - t &= 0 \\ x_1 = \frac{2 \pm \sqrt{4-4t^2}}{2t} &= \frac{1 \pm \sqrt{1-t^2}}{t}. \end{aligned}$$

From these two solutions only the the one with the minus will be inside the polytope. Thus we have

$$\Gamma_c = \{(\gamma(t), \gamma(t)) : t \geq 0\}, \quad \text{where} \quad \gamma(t) = \frac{1-\sqrt{1-t^2}}{t}.$$

□

Exercise 4

Consider a bounded polyhedron $P = \{x \in \mathbb{R}^n : a_i \cdot x \leq b_i \quad \text{for } i = 1, ..., m\}$ and let $F(x)$ be the logarithmic barrier function on the interior of P, that is $F(x) = -\sum_{i=1}^{m} \log(b_i - a_i \cdot x)$. Write up the gradient and the Hessian of F and prove that F is strictly convex.

Solution:

$$\nabla F(x) = -\sum_{i=1}^{m} \frac{1}{b_i - a_i \cdot x} \cdot (-a_i) = \sum_{i=1}^{m} \frac{a_i}{b_i - a_i \cdot x}$$

$$H(x) = \nabla^2 F(x) = \nabla \left(\sum_{i=1}^{m} \frac{a_i}{b_i - a_i \cdot x} \right) = \sum_{i=1}^{m} \frac{a_i}{b_i - a_i \cdot x} \cdot a_i^T = \sum_{i=1}^{m} \frac{a_i \cdot a_i^T}{b_i - a_i \cdot x}.$$

To show that F is strictly convex on the interior of P we need to show that $H(x)$ is positive definite. In other words, that $\forall z \in \mathbb{R}^n$ we have $z^T H(x) z > 0$.

$$z^T H(x) z = z^T = z^T \left(\sum_{i=1}^{m} \frac{a_i \cdot a_i^T}{b_i - a_i \cdot x} \right) z = \sum_{i=1}^{m} \frac{z^T a_i \cdot a_i^T z}{b_i - a_i \cdot x}$$
$$= \sum_{i=1}^{m} \frac{(a_i^T z)^T (a_i^T z)}{b_i - a_i \cdot x} = \sum_{i=1}^{m} \frac{\left\| a_i^T z \right\|^2}{b_i - a_i \cdot x} \geq 0.$$

Since $x \in \text{int}P$, we have $b_i - a_i \cdot x \geq 0$ for any $i = 1, ..., m$, and the norm of any vector is nonnegative.

Now we just need to show that for any $z \neq 0$ we have $z^T H(x) z \neq 0$. This could only happen if $a_i \cdot z = 0$ for all $i = 1, ..., m$. However, since the problem specified that P is bounded, this implies that there are at least n linearly independent constraint vectors a_i, thus such a z cannot exist. $\square$

Exercise 5

Let $P = \{x \in \mathbb{R}^n : a_i \cdot x \leq b_i \quad \text{for } i = 1, ..., m\}$ and $x \in \text{int}(P)$. For a vector $c \in \mathbb{R}^n$, let $c_x = H^{-1}(x)c$, where $H(x)$ is the Hessian of the logarithmic barrier function. Verify that hte point $x - c_x/\|c_x\|_x$ is in P.

Solution: $\square$

Exercise 6

Let $A \in \mathbb{R}^{m \times n}$, and let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a polytope with nonempty interior. Suppose that $x_0 \in P$ satisfies $b_i - a_i \cdot x_0 \geq \delta$ for all $i = 1, \dots, m$ and some $\delta > 0$. Define the logarithmic barrier function $F(x) = -\sum_{i=1}^{m} \log(b_i - a_i \cdot x)$ and let $g(x) = \nabla F(x)$ denote it's gradient. If D is the Euclidian diameter of P, show that for every $x \in P$ and every $i = 1, \dots, m$ we have

$$b_i - a_i \cdot x \geq \frac{\delta}{m + \|g(x)\| D}.$$

Solution: □

Exercise 7

Let $f : \mathbb{R}^n \to \mathbb{R}$ and suppose that there sxists cosntants $0 < m \leq L$ such that for all $z \in \mathbb{R}^n$ the Hessian satisfies $mI \preceq \nabla^2 f(z) \preceq LI$. Let $x \in \mathbb{R}^n$ and let $n(x) = -H^{-1}(x)g(x)$ be the Newton step with $g(x) = \nabla f(x)$ and $H(x) = \nabla^2 f(x)$ Show that

$$f(x + n(x)) - f(x) \leq -\frac{1}{2L}\|g(x)\|^2$$

and

$$f(x + n(x)) - f(x) \geq -\frac{1}{2m}\|g(x)\|^2.$$

(Hint: use the second-order Taylor expansion with the mean-value form of the remainder and the bounds on $\nabla^2 f$.)

Solution:

$\square$