

AMATH 514 Assignment 8

Tyler Chen

Note: Note that I received a 16/20, however I have not made corrections

Problem 1

Consider the cube $P := [0, 1]^n = \{x : 0 \leq x_i \leq 1, i = 1, \dots, n\}$. Consider a sequence of points $\{x_k\}_{k \geq 0}$ with $x_1 = (\frac{1}{2}, \dots, \frac{1}{2})$ and with the only restriction that $x_{k+1} \in \mathcal{E}(x_k, 1/2)$. Prove that it takes at least $\Omega(\sqrt{n} \log(1/\delta))$ iterations until x_k can be within a $\|\cdot\|_\infty$ -distance of δ from the vertex 0.

We use the interpretation that $s_i(x)$ is the geometric distance of x to the i -th hyperplane. For notational convenience we will start at x_0 .

Observe that,

$$\sum_{i=1}^{2n} \left(\frac{s_i(y) - s_i(x)}{s_i(x)} \right)^2 = \sum_{i=1}^n \left[\left(\frac{y_i - x_i}{x_i} \right)^2 + \left(\frac{y_i - x_i}{1 - x_i} \right)^2 \right] > \sum_{i=1}^n \left(\frac{y_i - x_i}{x_i} \right)^2$$

Suppose $y = d(1, \dots, 1)$ and $x = c(1, \dots, 1)$. Then,

$$R^2 = \sum_{i=1}^n \left(\frac{y_i - x_i}{x_i} \right)^2 = n \left(\frac{d - c}{c} \right)^2 \iff d = c \pm \frac{cR}{\sqrt{n}} = c \left(1 \pm \frac{R}{\sqrt{n}} \right)$$

We define a sequence $\{x_k\}$ where $x_k = c_k(1, \dots, 1)$ for all k and satisfies the relation,

$$c_{k+1} = c_k \left(1 - \frac{R}{\sqrt{n}} \right) = c_0 \left(1 - \frac{R}{\sqrt{n}} \right)^{k+1}$$

Note that x_{k+1} is outside of the ellipsoid $\mathcal{E}(x_k, R)$ since,

$$\sum_{i=1}^{2n} \left(\frac{s_i(x_{k+1}) - s_i(x_k)}{s_i(x_k)} \right)^2 > \sum_{i=1}^n \left(\frac{c_{k+1} - c_k}{c_k} \right)^2 = n \left(\frac{c_{k+1} - c_k}{c_k} \right)^2 = R^2$$

That is, at each step we jump a bit more than allowed by a “valid” sequence.

We now determine how long it takes for $\{x_k\}_{k \geq 0}$ to get within δ of the origin (in the infinity norm). That is, we seek k such that $c_k < \delta$. Thus,

$$c_0 \left(1 - \frac{R}{\sqrt{n}} \right)^k = c_k < \delta$$

Now recall,

$$\log(1 - a) = -a - a^2/2 - \dots < -2a$$

Thus, dividing by c_0 on both sides, and since the logarithm is monotonically increasing,

$$-k \left(\frac{2R}{\sqrt{n}} \right) < k \log \left(1 - \frac{R}{\sqrt{n}} \right) < \log \left(\frac{\delta}{c_0} \right)$$

Therefore

$$k > -\frac{\sqrt{n}}{2R} \log \left(\frac{\delta}{c_0} \right) = \frac{\sqrt{n}}{2R} \log \left(\frac{c_0}{\delta} \right)$$

This proves $\{x_k\}_{k \geq 0}$ takes $\Omega(\sqrt{n} \log(1/\delta))$ steps to be within a $\|\cdot\|_\infty$ -distance of δ from the origin. Therefore, any sequence $\{y_k\}$ where $y_{k+1} \in \mathcal{E}(y_k, 1/2)$ and $y_k = d_k(1, \dots, 1)$ takes at least $\Omega(\sqrt{n} \log(1/\delta))$ to be within a $\|\cdot\|_\infty$ -distance of δ from the origin.

It remains to show that any other sequence z_k satisfying $z_{k+1} \in \mathcal{E}(z_k, 1/2)$ does not converge faster than a sequence along the diagonal between the points 0 and 1.

Let $\{z_k\}_{k \geq 0}$ be any sequence satisfying $z_{k+1} \in \mathcal{E}(z_k, 1/2)$. We claim the sequence $\{y_k\}_{k \geq 0}$ obtained by projecting z_k onto the diagonal between 0 and 1 satisfies $y_{k+1} \in y_k$.

Let u, v be two points in P satisfying $v \in \mathcal{E}(u, 1/2)$. Let x, y be their projections onto the main diagonal.

We have $\|x - y\| \leq \|u - v\|$. Moreover,

$$|s_i(u) - s_i(v)| = |u_i - v_i| \leq |y_i - x_i| = |s_i(y) - s_i(x)|$$

Therefore,

$$\sum_{i=1}^{2n} \left(\frac{s_i(u) - s_i(v)}{s_i(u)} \right)^2 \leq \sum_{i=1}^{2n} \left(\frac{s_i(x) - s_i(y)}{s_i(u)} \right)^2 \leq \sum_{i=1}^{2n} \left(\frac{s_i(x) - s_i(y)}{s_i(x)} \right)^2$$

where the second inequality comes from the fact that for any $s_i(u) < s_i(x)$ there is some j such that $s_j(u) > s_j(x)$. Since x is on the main diagonal it is “further away” from all the boundaries than u . (I know this is kind of handwavy)

This means $y \in \mathcal{E}(x, 1/2)$ as desired.

Therefore the projection of any “valid” sequence onto the main diagonal is also valid. But no sequence on the diagonal can converge fast enough so the result is proved \square

Problem 2

Recall that the presented interior point method takes $\mathcal{O}(L\sqrt{m})$ iterations to get within an additive 2^{-L} distance to the optimum for a polytope $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ with $A \in \mathbb{R}^{m \times n}$. There is indeed a way of bringing the number of iterations down to $\mathcal{O}(L\sqrt{n})$. A deep result of Nesterov and Nemirovsky says that there is a convex function $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ that is self-concordant which means it satisfies the following properties for some universal constant $C > 0$:

(A) For any $0 \leq R \leq 1/C$ and $x \in \mathcal{E}(x^*, R)$ one has

$$(1 - 2R)\nabla^2\varphi(x) \preceq \nabla^2\varphi(x^*) \preceq (1 + 2R)\nabla^2\varphi(x)$$

where we redefine the ellipsoid $\mathcal{E}(x^*, R) := \{x \in \mathbb{R}^n : (x - x^*)^T [\nabla^2\varphi(x)] (x - x^*) \leq R^2\}$.

(B) One has $\nabla\varphi(x)\nabla\varphi(x)^T \preceq Cn\nabla^2\varphi(x)$ for all $x \in \text{int}(P)$.

(C) If $x \rightarrow \partial P$, then $\varphi(x) \rightarrow \infty$.

For $t \geq 0$ we modify the barrier function to $F_t(x) := t \cdot c^T x + \varphi(x)$. Prove the following:

1. For $x \in \mathcal{E}(x^*, R)$ with $x^* := x^*(t)$ and $x' := x - [\nabla^2 F_t(x)]^{-1} \nabla F_t(x)$ one has $x' \in \mathcal{E}(x^*, \mathcal{O}(R^2))$ for sufficiently small R .
2. $\max\{t \cdot c^T(x - x^*(t)) : x \in \mathcal{E}(x^*(t), R)\} \leq \mathcal{O}(R\sqrt{n})$ for all $t > 0$ and $R > 0$ small enough.

1. Let $y = x^* + h$ for some h with $\|h\| = \mathcal{O}(R)$ (for instance any $y \in \mathcal{E}(x^*, R)$).

Let $F_{ij}(y)$ be an entry of $\nabla^2 F_t(y)$. Then,

$$F_{ij}(y) = F_{ij}(x^*) + h^T \nabla F_{ij}(x^*) + \mathcal{O}(h^2)$$

Since $\nabla F_t(x^*) = 0$, this shows that,

$$\nabla^2 F_t(y) = \nabla^2 F_t(x^*) + \mathcal{O}(h^2)$$

Then, there is some k_1 such that,

$$\nabla^2 F_t(y) \preceq \nabla^2 F_t(x^*) + k_1 R^2$$

Similarly there is some k_2 such that,

$$\nabla^2 F_t(x^*) \preceq \nabla^2 F_t(y) + k_2 R^2$$

Define $k = \max\{|k_1|, |k_2|\}$. Then,

$$\nabla^2 F_t(x^*) - kR^2 \preceq \nabla^2 F_t(y) \preceq \nabla^2 F_t(x^*) + kR^2$$

Therefore, using (A) and the fact that $\nabla^2 F_t(x) = \nabla^2 \varphi(x)$,

$$(1 - 2R)\nabla^2 F_t(x) - kR^2 \preceq \nabla^2 F_t(y) \preceq (1 + 2R)\nabla^2 F_t(x) + kR^2$$

Note that here k depends on x . However, $\sup_{x \in \mathcal{E}} k$ gives a bound for all $x \in \mathcal{E}(x^*, R)$.

Somehow get a bound of the form,

$$(1 - cR)\nabla^2 F_t(x) \leq \nabla^2 F_t(y) \leq (1 + cR)\nabla^2 F_t(x)$$

The result then follows from Lemma 8.3 as the result (claim 3) is proved using claim 2 which is proved using claim 1 shown here.

Sorry for the sloppy work. I was in a bike accident earlier this week and haven't been able to concentrate very well since then. I would have asked for an extension, but at this point I think I need to cut my losses and move on to other work.

2. Suppose $x \in \mathcal{E}(x^*, R)$. That is,

$$(x - x^*)^T [\nabla^2 \varphi(x^*)] (x - x^*) \leq R^2$$

Then, by (B),

$$(x - x^*)^T [\nabla \varphi(x^*) \nabla \varphi(x^*)^T] (x - x^*) \leq Cn(x - x^*)^T [\nabla^2 \varphi(x^*)] (x - x^*) \leq CnR^2$$

But,

$$(x - x^*)^T [\nabla \varphi(x^*) \nabla \varphi(x^*)^T] (x - x^*) = |\nabla \varphi(x^*)^T (x - x^*)|^2$$

Since $\nabla F_t(x^*) = 0$ we have,

$$\nabla \varphi(x^*) = -t \cdot c$$

Thus,

$$|\nabla \varphi(x^*)^T (x - x^*)|^2 = |t \cdot c^T (x - x^*)|^2$$

Combining these statements we have,

$$|t \cdot c^T (x - x^*)|^2 \leq CnR^2$$

This proves $t \cdot c^T (x - x^*) \leq \mathcal{O}(R\sqrt{n})$ for all $x \in \mathcal{E}(x^*, R)$ and sufficiently small R . Then certainly $\max\{t \cdot c^T (x - x^*(t)) : x \in \mathcal{E}(x^*(t), R)\} \leq \mathcal{O}(R\sqrt{n})$ for all $t > 0$ and $R > 0$ small enough. \square