CS 224 Advanced Algorithms — Spring 2017

Problem Set 7

Due: 11:59pm, Monday, April 17th

Submit solutions to Canvas, one PDF per problem:
https://canvas.harvard.edu/courses/21996

Solution max page limits: 3 pages for problem 1, and 2 pages for problem 2 (not including the bonus problem, which has no page restriction)

See homework policy at http://people.seas.harvard.edu/~cs224/spring17/hmwk.html

Problem 1: In class we didn't discuss the bit complexity in detail for interior point methods. Here you will fill in some of the gaps.

(a) (2 points) For A, b, c having integer entries with $A \in \mathbb{R}^{m \times n}$, consider the linear program

$$\min c^T x$$
$$\text{s.t.} \quad Ax \geq b \tag{1}$$

Define $L = C(1 + \log(1 + d_{max}) + \log(1 + \max\{\|b\|_\infty, \|c\|_\infty\}))$ for sufficiently large constant C. Here d_{max} is the largest absolute value of a determinant of a submatrix of A. Show that if (1) is bounded then there is an optimum solution x^* to this LP s.t. for all i, x_i^* requires $\leq L$ bits of precision. That is, for all i the value x_i^* is rational and can be expressed as α_i / β_i for integers $-m2^L \leq \alpha_i, \beta_i \leq m2^L$.

(b) (5 points) Consider the linear program (1) from (a). Consider the transformation to a new LP as follows:

$$\min c^T x + n^{C'} \cdot 2^{CL} z$$
$$\text{s.t.} \quad Ax + \mathbb{1} z \geq b$$
$$0 \leq z \leq 2^L$$
$$\forall i, \ -2^{L+1} \leq x_i \leq 2^{L+1} \tag{2}$$

where $\mathbb{1}$ is the all-ones vector. We noted that $x = 0, z = \|b\|_\infty$ is an obvious feasible solution of (2). Show that for some constants $C, C' > 0$ (1) is bounded and feasible with optimal solution x^* iff $(x, z) = (x^*, 0)$ is an optimal solution to (2) for some x^* with $\|x^*\|_\infty \leq 2^L$.

(c) (4 points) In class we assumed the matrix $\nabla^2 f_{\lambda_k}(x)$ considered in Newton steps was always invertible. Show that this can be ensured throughout the course of the path-following IPM algorithm presented in class to solve (1), at the expense of increasing both the number of rows and columns of A by an additive $O(m+n)$ in pre-processing.

(d) (4 points) In the Lecture 22 notes, we show if x is *perfectly* central for $\lambda = \Omega(m/\varepsilon)$, then $c^T x \le \mathsf{OPT} + \varepsilon$. However, interior point does not find a perfectly central point for large λ, but rather it finds a *finely* central point. Show that x being finely central for $\lambda = \Omega(m/\varepsilon)$ is sufficient to achieve $c^T x \le \mathsf{OPT} + \varepsilon$. **Hint:** It may help to first prove a generalized Cauchy-Schwarz inequality, namely that $\langle x, y \rangle \le \|x\|_B \cdot \|y\|_{B^{-1}}$ for any real, symmetric positive definite matrix B.

Problem 2: In the "learning from experts" setup, on each day it will either rain or not. n experts each tell us their prediction as to whether it will rain, and we must take their predictions as input to form our own prediction. We then find out at the end of the day whether we were right or wrong (it either rained or didn't). If T is the total number of days, E_i is the number of mistakes of expert i, and M is our number of mistakes, then forming our predictions using the Multiplicative Weights algorithm guarantees

$$M \le \min_{1 \le i \le n} E_i + \underbrace{\eta T + \frac{\ln n}{\eta}}_{R}$$

for any $\eta \le 1/2$. In particular if we choose $\eta = \sqrt{(\ln n)/T}$ when $T \ge 2 \ln n$ (so that $\eta \le 1/2$), then our regret R is at most $2\sqrt{T \ln n}$.

(a) (5 points) Suppose we flip a fair coin T times. Let X be the number of heads, so that $T - X$ is the number of tails. Show that $\mathbb{E} \min\{X, T - X\} \le T/2 - \Omega(\sqrt{T})$. **Hint:** Use Stirling's formula to derive that $\binom{2r}{r} = \Theta(4^r/\sqrt{r})$ as $r \to \infty$.

(b) (2 points) Show that *no randomized algorithm* (not just Multiplicative Weights) can achieve expected regret $\mathbb{E} R = o(\sqrt{T})$ in general. **Hint:** Consider $n = 2$ and use (a).

(c) (8 points) Show that no algorithm can achieve $\mathbb{E} R = o(\sqrt{T \ln n})$ when T is sufficiently large. You may use, without proof, that if X is the number of heads after flipping T independent fair coins, then for any $t \in [0, T/8]$,

$$\mathbb{P}(X \le \frac{T}{2} - t) \ge \frac{1}{15} e^{-16t^2/T}.$$