6.S978 Graphs, Linear Algebra, and Optimization – Fall 2015 October 27, 2015

Lecture 11

Lecturer: Aleksander Mądry *Scribe: Slobodan Mitrović*

1 Overview

In this lecture, we continue our study of the oracle based approach to solving the maximum flow problem. We will use electrical flows to develop a new oracle that improves upon the width of the shortest path–based oracle we considered before.

2 Recap of the Last Lecture

Last time, we set up a general framework for approximately solving a decision LP problem $\mathcal{P}$ defined as follows

$$\begin{aligned} \mathcal{P} : Ax &\leq b \\ x &\in \mathcal{K}. \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$ is a constraint matrix, $b \in \mathbb{R}^m$, and $\mathcal{K} \subset \mathbb{R}^n$ is a certain convex set.

This framework assumes the existence of a (θ, ρ)-oracle, for some $0 \leq \theta \leq \rho$, where given a point p on the simplex Δ_m the oracle:

(i) If $\mathcal{P}$ is feasible, outputs a point $\tilde{x} \in \mathcal{K}$ such that: $p^T A\tilde{x} - p^T b \geq 0$; and $A_i\tilde{x} - b_i \in [-\theta, \rho]$ whenever $1 \leq i \leq m$.

(ii) Otherwise, outputs $\perp$ or $\tilde{x}$ as above.

In other words, if $\mathcal{P}$ is feasible, the oracle outputs a point which satisfies the constraints "on average", where the average is taken wrt the input convex combination p. Parameter θ can be seen as the maximal allowed slack and ρ the maximal allowed violation of any individual constraint. Observe that even if $\tilde{x} \notin \mathcal{K}$, it is possible that $\tilde{x}$ satisfies the constraints on average, and gets output by the oracle.

After introducing this notion of (θ, ρ)-oracle we proved the following theorem.

Theorem 1 *Consider a (θ, ρ)-oracle for the feasibility problem $\mathcal{P}$. For any error parameter $0 < \varepsilon \leq \frac{1}{2}$, as long as $\theta \geq \frac{\varepsilon}{2}$, there is an algorithm that after $O(\theta\rho\varepsilon^{-2} \ln m)$ oracle calls outputs a point $\bar{x}$ that satisfies*

$$A\bar{x} \leq b + \varepsilon\vec{1},$$

or $\perp$ if $\mathcal{P}$ is not feasible.

So, whenever we are able to construct a (θ, ρ)-oracle we can use it as a black-box to approximately solve the feasibility problem $\mathcal{P}$.

2.1 Solving the Maximum Flow Problem with (θ, ρ)-oracles

We then showed how to apply this framework to the maximum flow problem. More precisely, given a scalar-parameter F we have defined the following feasibility problem

$$\begin{aligned} \bar{\mathcal{F}}(F) : Fy &\leq \vec{1} \\ y &\in \bar{\mathcal{F}}_{st}, \end{aligned}$$

where $\bar{\mathcal{F}}_{st} := \{\text{abs}(f) \mid Bf = \chi_{st}\}$, and $\text{abs}(f)$ is an m-dimensional vector defined as $\text{abs}(f)_e = |f_e|$. The parameter F is our guess on the value $F^\star$ of a maximum flow. (So, we want to decide if $F \leq F^\star$, to apply the binary search strategy.)

Recall that in our framework the set $\bar{\mathcal{F}}_{st}$ in $\bar{\mathcal{F}}(F)$ that corresponds to the set $\mathcal{K}$ in $\mathcal{P}$ should be convex. One can prove that $\bar{\mathcal{F}}_{st}$ is indeed convex, but this would require some work. However, if we analyze how convexity of $\mathcal{K}$ is used in the proof of Theorem 1, we can see that we only need a weaker guarantee in our case. More precisely, we only need to show that from the fact that $\frac{1}{T}\sum_{t=1}^{T} y^t \leq (1+\varepsilon)\vec{1}$, i.e. that the average of the edge-flow vectors y^t is ε-approximately feasible, where each $y^t = \text{abs}((f^t))$ for some unit s-t flows f^t, it follows that the edge-flow vector $\bar{y} = \text{abs}(\bar{f})$ corresponding to the average $\bar{f}$ of all the flows f^t is also ε-approximately feasible. This is immediately implied by the convexity of the $|\cdot|$ function though. That is, for any coordinate e,

$$\bar{y}_e = \left|\frac{1}{T}\sum_{t=1}^{T} f_e^t\right| \leq \frac{1}{T}\sum_{t=1}^{T} |f_e^t| = \frac{1}{T}\sum_{t=1}^{T} y_e^t \leq 1+\varepsilon.$$

(Note that we implicitly assume here that our oracle provides the corresponding flow vector f^t along with each edge-flow vector y^t.)

3 The Shortest Path–based Oracle for $\bar{\mathcal{F}}(F)$

Our goal now is to design a (θ, ρ)-oracle for the problem $\bar{\mathcal{F}}(F)$. Rewriting the definition of that oracle in the language of our max flow formulation, we want, given an input $p \in \Delta_m$:

(i) If $F \leq F^\star$, output f such that

$$\sum_e (p_e(F|f_e| - 1)) = \sum_e (p_e(Fy_e - 1)) \leq 0, \tag{1}$$

and

$$F|f_e| - 1 = Fy_e - 1 \in [\theta, \rho], \forall e. \tag{2}$$

(ii) Otherwise, i.e. if $F > F^\star$, output $\perp$ or f as above.

Observe that since $y_e = |f_e| \geq 0$ and $F \geq 0$, we have that $Fy_e - 1 \geq -1$ and thus setting $\theta = 1$ satisfies the condition (2) regardless of the designed oracle. So, we only have to analyze the width ρ and to care about the condition (1). Also, the latter condition can be rewritten as

$$p^T y \leq \frac{\|p\|_1}{F} = \frac{1}{F}. \tag{3}$$

Now, last time, we proposed and analyzed an oracle for $\bar{\mathcal{F}}(F)$ that was based on the shortest-path computations. The motivation here was that the optimization problem solved in the shortest path problem is reminiscent of the one that the maximum flow problem solves. More precisely, recall that the (unit-capacity) maximum flow problem corresponds to performing the following optimization.

$$\begin{aligned} \min \|f\|_\infty \\ Bf = \chi_{st} \end{aligned}$$

On the other hand, shortest paths solve the optimization question of the form:

$$\begin{aligned} \min \|Pf\|_1 \\ Bf = \chi_{st}, \end{aligned}$$

where P is a diagonal m-by-m matrix with each diagonal entry $P_{ee} = p_e$ equal to the length p_e of the edge.

So, the feasibility set over which both problems are optimizing are exactly the same. The only difference is the norm they are trying to minimize. For the maximum flow problem, it is the ℓ_∞-norm, while for the shortest path problem it is the ℓ_1-norm.

Building on this similarity, we design an oracle $\mathcal{O}_{SP}$ for the maximum flow problem that simply performs shortest path computations wrt to the lengths given by the convex combination p. More precisely, the oracle $\mathcal{O}_{SP}$, given $p \in \Delta_m$, computes a shortest path $\hat{f} = \arg\min_{Bf=\chi_{st}} \|Pf\|_1$ (note that wlog we can assume that f is supported on a single shortest path) and

(i) if $p^T\hat{y} \leq \frac{1+\varepsilon}{F}$, it returns $\hat{f}$, where $\hat{y} = \mathrm{abs}(\hat{f})$;

(ii) otherwise, it returns $\perp$.

Clearly, if $\bar{\mathcal{F}}(F)$ is not feasible, i.e., if $F > F^\star$ then the oracle is allowed to return a flow or $\perp$ and hence it is always correct. So, we just need to focus on the case of $F \leq F^\star$ and argue that the oracle $\mathcal{O}_{SP}$ always returns a valid solution in that case or, equivalently, that $p^T\hat{y} \leq \frac{1+\varepsilon}{F}$ then. We prove this is the lemma below.

Lemma 2 *If $F \leq F^\star$, then $\mathcal{O}_{SP}$ returns a flow $\hat{f}$ such that $p^T\hat{y} \leq \frac{1+\varepsilon}{F}$.*

Proof Observe that

$$p^T\hat{y} = |P\hat{f}|_1.$$

So, all we need to do here is to argue that the optimal objective value of the shortest path optimization problem is at most $\frac{1+\varepsilon}{F}$. For that it suffices to simply exhibit a flow f' such that $Bf' = \chi_{st}$ and $\|Pf'\|_1 \leq \frac{1+\varepsilon}{F}$.

To this end, let us set $f' := f^\star$ to be a maximum flow, i.e. a flow such that $\|f^\star\|_\infty \leq \frac{1}{F^\star} \leq \frac{1}{F}$. Then, we have

$$p^T y^\star = \|Pf^\star\|_1 = \sum_e p_e|f_e^\star| \leq \sum_e p_e \frac{1}{F} = \frac{\|p\|_1}{F} = \frac{1}{F} \leq \frac{1+\varepsilon}{F},$$

as desired. ■

To complete the analysis, it remains to bound the width ρ of the oracle $\mathcal{O}_{SP}$. As wlog we can assume that this oracle computes a single shortest s-t path, we have that $|\hat{f}_e| \in \{0, 1\}$, and hence $F|\hat{f}_e| - 1 \in \{-1, F-1\}$. Therefore, $\rho = F - 1$ and, from Theorem 1, we know that we need $O(F\varepsilon^{-2}\ln m)$ oracle calls to compute our approximate solution to the maximum flow problem. Each oracle call can be implemented in $\tilde{O}(m)$ time, so to solve the problem $\bar{\mathcal{F}}(F)$ we need a total $\tilde{O}(mF\varepsilon^{-2})$ time.

4 Beyond the Width of $\Theta(F)$

As we have seen, the shortest path oracle has width of $F-1$ and the resulting running time is $\tilde{O}(mF\varepsilon^{-2})$, which is hardly impressive (and not even polynomial-time). Is there perhaps some other oracle that has asymptotically smaller width?

We have chosen the shortest path oracle as it is ℓ_1-minimization over the same set of constraints as the maximum flow optimization. Electrical flow is yet another problem with this property. Specifically, while the maximum flow corresponds to l_∞ minimization over the space of unit s-t flows, electrical flows corresponds to ℓ_2-minimization over that set. In this section, we will design an oracle based on the electrical flow computation and discuss its efficiency.

4.1 Electrical Flow Based Oracle

For a given vector p and its corresponding diagonal matrix P, the electrical flow minimization is formulated as follows

$$\begin{aligned} min\, & \|P^{1/2}f\|_2^2 \\ & Bf = \chi_{st}, \end{aligned}$$

where as before P is a diagonal matrix such that $P_{e,e} = p_e$. Matrix P represents now the resistances. Let $\mathcal{O}_{EF}$ be an oracle that given a point $p \in \Delta_m$ computes the electrical flow $\tilde{f}$ and returns:

(i) the flow vector $\tilde{f}$, if $\|P\tilde{f}\|_1 \le \frac{1+\varepsilon}{F}$; and

(ii) $\perp$ otherwise.

Now, we would like to argue that such an oracle is correct. More precisely, as in Lemma 2 we are going to show that $\mathcal{O}_{EF}$ returns a flow whenever $F \le F^\star$.

Lemma 3 *If $F \le F^\star$, then $\mathcal{O}_{EF}$ returns a flow $\tilde{f}$ such that $p^T\tilde{y} \le \frac{1+\varepsilon}{F}$.*

Proof Let us first make the following observation. Assume that

$$\|P^{1/2}\tilde{f}\|_2^2 \le \frac{(1+\varepsilon)^2}{F^2}. \tag{4}$$

Then, using Cauchy-Schwarz inequality (4) implies

$$\begin{aligned}
\|P\tilde{f}\|_1 &= \sum_e p_e|\tilde{f}_e| \\
&= \sum_e \sqrt{p_e}\left(\sqrt{p_e}|\tilde{f}_e|\right) \\
&\le \sqrt{\sum_e p_e \sum_e p_e\left(\tilde{f}_e\right)^2} \\
&= \sqrt{\|P^{1/2}\tilde{f}\|_2^2} \\
&\le \frac{1+\varepsilon}{F},
\end{aligned}$$

which is the bound we are aiming to prove. Therefore, as $p^T\tilde{y} = \|P\tilde{f}\|_1$ and as $\mathcal{O}_{EF}$ minimizes the flow energy, if we exhibit a flow f' such that its energy with respect to p is at most $\frac{(1+\varepsilon)^2}{F^2}$ it would imply that (4) holds and we are done. As before, consider f' to be an optimal flow, i.e. $f' := f^\star$. Then, we can bound the energy of $f^\star$ as follows

$$\begin{aligned}
\|P^{1/2}f^\star\|_2^2 &= \sum_e p_e\left(f_e^\star\right)^2 \\
&\le \sum_e \frac{p_e}{(F^\star)^2} \\
&\le \sum_e \frac{p_e}{F^2} \\
&= \frac{1}{F^2} \\
&\le \frac{(1+\varepsilon)^2}{F^2}.
\end{aligned}$$

And hence the lemma follows. ■

So, we can use $\mathcal{O}_{EF}$ to solve the maximum flow problem. However, unfortunately, the width of the described oracle is still $F-1$.[1] To see that, consider the following example. Let H be an s-t path, and let p be defined in such a way that $p_e = 0$ for every edge e of H, and $p_e > 0$ otherwise. In such a case, $\mathcal{O}_{EF}$ will put all the flow on H inquiring the width $\rho = F-1$, just like in the case of oracle $\mathcal{O}_{SP}$. However, such a case is very special and we might hope that it will never occur. But, showing that formally could be a quite challenging task. So, instead of hoping nothing bad will happen, we will apply a general approach for handling such undesired behaviors.

[1] Notice that $F^\star$, and hence F as well, can be of the order of n. So, in some cases saying that the oracle width is F is equivalent to saying that the width equals n, or equals m in the sparse graph case.

4.2 Regularization

When dealing with optimization problems, as we are here, an idea widely used in machine learning and optimization is *regularization*. This ideas corresponds to slightly altering the objective function we want to optimize by adding a term to it that penalizes the "unwanted" solutions.

In the case of the electrical flow based oracle, given a point $p \in \Delta_m$ we will apply regularization in the following way. Instead of computing the flow for the resistance vector p, the oracle will compute the electrical flow for the resistances given by $r := p + \frac{\varepsilon}{2m}\vec{1}$. We use $\mathcal{O}_{EFR}$ to refer to that oracle. But, can we still guarantee that $\mathcal{O}_{EFR}$ computes a flow when $F \leq F^\star$? Yes, we can, as the following lemma shows.

Lemma 4 *If $F \leq F^\star$, then given point $p \in \Delta_m$ oracle $\mathcal{O}_{EFR}$ returns a flow $\tilde{f}$ such that $p^T\tilde{y} \leq \frac{1+\varepsilon}{F}$.*

Proof Similarly to the proof of Lemma 3, to prove this lemma we can try to exhibit a flow f' such that $\|P^{1/2}f'\|_2^2 \leq \frac{(1+\varepsilon)^2}{F^2}$. However, $\mathcal{O}_{EFR}$ outputs the electrical flow for the resistance vector r but not for p. Nevertheless, observe that $\|P^{1/2}\tilde{f}\|_2^2 \leq \|R^{1/2}\tilde{f}\|_2^2$, where R is a diagonal matrix such that $R_{e,e} = r_e$. So, if we find a flow f' such that $\|R^{1/2}f'\|_2^2 \leq \frac{(1+\varepsilon)^2}{F^2}$ the lemma will follow for the same reasons as in the proof of Lemma 3.

Yet again, consider $f' := f^\star$. Then, we have

$$\begin{aligned} \|R^{1/2}f^\star\|_2^2 &= \sum_e \left(p_e + \frac{\varepsilon}{2m}\right)(f_e^\star)^2 \\ &\leq \sum_e \frac{p_e + \frac{\varepsilon}{2m}}{F^2} \\ &= \frac{\sum_e p_e + \sum_e \frac{\varepsilon}{2m}}{F^2} \\ &= \frac{1+\frac{\varepsilon}{2}}{F^2}, \end{aligned}$$

which concludes the proof. ■

So, even after applying regularization our oracle is correct (although now only ε-approximately so). However, what is its width? To answer that question, fix any edge e. We will show that $|\tilde{f}_e| \in O(\sqrt{\frac{m}{\varepsilon}})\frac{1}{F}$, which in turn implies that the width of $\mathcal{O}_{EFR}$ is in $O(\sqrt{\frac{m}{\varepsilon}})$. To that end, we upper bound the energy contribution of e by the total flow energy, i.e.

$$\begin{aligned} r_e\left(\tilde{f}_e\right)^2 &\leq \sum_e r_e\left(\tilde{f}_e\right)^2 \\ &\leq \frac{1+\frac{\varepsilon}{2}}{F^2}. \end{aligned} \tag{5}$$

Next, observe that

$$\begin{aligned} r_e\left(\tilde{f}_e\right)^2 &= \left(p_e + \tfrac{\varepsilon}{2m}\right)\left(\tilde{f}_e\right)^2 \\ &\geq \tfrac{\varepsilon}{2m}\left(\tilde{f}_e\right)^2. \end{aligned} \tag{6}$$

Putting together (5) and (6) we get

$$|\tilde{f}_e| \leq \sqrt{\frac{2\left(1+\frac{\varepsilon}{2}\right)}{\varepsilon F^2}m} \in O\left(\sqrt{\tfrac{m}{\varepsilon}}\right)\tfrac{1}{F},$$

as advertised. So, we can state the following theorem.

Theorem 5 *For any $\varepsilon > 0$, there is an algorithm that outputs a $(1+\varepsilon)$-approximation of the maximum flow problem after $\tilde{O}\left(\sqrt{\frac{m}{\varepsilon}}\varepsilon^{-2}\right)$ oracle $\mathcal{O}_{EFR}$ calls, and has the total running time of $\tilde{O}\left(m^{3/2}\varepsilon^{-5/2}\right)$.*

4.3 The Width Analysis of $\mathcal{O}_{EFR}$ is Tight

We have just shown that by applying regularization we are able to substantially improve the width of our oracle. However, can we improve it even further by providing a more careful analysis? That is, is the obtained width an artifact of our proofs and in fact the width of $\mathcal{O}_{EFR}$ is $o(\sqrt{\frac{m}{\varepsilon}})$?

Unfortunately, the example in Figure 1 shows that our analysis is tight. Namely, a unit flow in

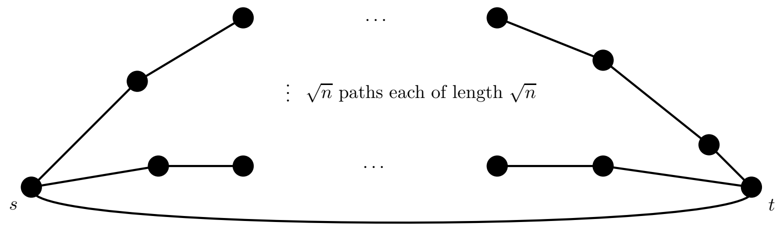

Figure 1: An example that shows our analysis of the width of $\mathcal{O}_{EFR}$ is tight.

the graph in Figure 1 that minimizes the energy sends 1/2 of the flow along the unique s-t edges, and uniformly distributes the remaining 1/2 of the flow over the $\sqrt{n}$ paths. On the other hand, the maximal s-t flow in the graph in the figure is $\sqrt{n}+1$, so for $F = \sqrt{n} \leq F^{\star}$ and uniform p the oracle $\mathcal{O}_{EFR}$ returns a flow that violates the constraints by as much as $\Theta(\sqrt{n})$ and hence its width is $\Theta(\sqrt{n})$. So, from this we can conclude that at least in the case of sparse graphs, i.e. $m = \tilde{O}(n)$, our analysis of the width of $\mathcal{O}_{EFR}$ is tight. As a remark, what happens in our proof is that the energy of the only s-t edge matches most of the energy of the flow (cf. (5)).

4.4 Towards a Reduced Width by Altering the Input Graph

Although the instance in Figure 1 shows the width of $\mathcal{O}_{EFR}$ is $\Theta(\sqrt{m})$, we can very easily change that instance and obtain a much nicer one. Namely, if we remove the shortcutting s-t edge, all of a sudden we get a graph in which the electrical and the maximum flow are exactly the same, in big contrast to the huge difference between these two objects when the s-t edge was present! At the same time, the value of the maximum flow is negligibly affected by the adjustment. Driven by this example and observation, in the next lectures we will develop a more general approach of altering the input graph which in turn will enable us to improve upon the width obtained by $\mathcal{O}_{EFR}$.