CS787: Advanced Algorithms	Scribe: Jurijs Nazarovs
Lecture 9: Randomized rounding, LP duality.	**Date:** 02/21/2019

9.1 Routing to minimize congestion

We are given a graph $G = (V, E)$ and k "commodities". Each commodity i has a source s_i and a destination t_i . The goal is to find an $(s_i,\ t_i)$ path in the graph G for every commodity i, while minimizing the maximum congestion along any edge. It can be written formally as Minimize congestion $C = \max_{e\in E} |\{i : e \in P_i\}|$, where P_i is the $(s_i,\ t_i)$ path of commodity i, and e is an edge.

Note this problem is similar to a network flow problem. However, this is not a network flow problem because for each commodity i, we need to have a single path between s_i and t_i . We cannot have the flow splitting along multiple branches.

This problem is an NP-hard problem. Thus, we solve it by formulating a corresponding ILP problem, relaxing it to an LP problem, solving the LP, and rounding the solution to an ILP solution.

9.1.1 LP$_1$ (Path LP): LP formulation with exponential number of variables and constraints.

Note, that we do not consider ILP, but LP relaxation from the beginning, by replacing $x_{i,p} \in \{0, 1\}$ with $x_{i,p} \in [0, 1]$.

Let $i = 1, \dots, k$ be the k commodities.
Let $\mathcal{P}_i$ be the set of all paths from s_i to t_i.
We have a variable $x_{i,P}$ for every $P \in \mathcal{P}_i$, $\forall i$.

$$min\ \ t, \text{ s.t.}$$

$$\sum_{p\in\mathcal{P}_i} x_{i,p} = 1, \forall i$$

$$\sum_i \sum_{p\in\mathcal{P}_i : e\in P} x_{i,p} \le t, \forall e$$

$$x_{i,p} \in [0, 1], \forall i, P$$

Note that the actual objective function which is in a min-max form is not linear. So, we employ a trick of introducing a new variable t. We introduce the constraint (second) that congestion on any edge ≤ t, and so all we need to do now is minimize t, which is equivalent to minimizing the maximum congestion. The first makes sure that we select exactly one path for each commodity.

The problem with this LP formulation is that we can have an exponential number of paths between s_i and t_i (for example, in a completely connected graph). We would like a formulation with a polynomial (polynominal in $|V|$, $|E|$, and k) number of variables and constraints. Next we consider an alternative ILP formulation - LP_2.

9.1.2 LP_2 (Flow LP): LP formulation with polynomial number of variables and constraints

Rather than looking at all the paths, we look at an edge granularity and see if a commodity is routed along any edge.

We have variables $y_{e,i}, \forall e \in E, i = 1, \ldots, k$

The LP formulation is following:

$$min \;\; t, \text{ s.t}$$

$$\sum_i y_{e,i} \leq t, \forall e$$

$$\sum_{e \in \delta^-(v)} y_{e,i} = \sum_{e \in \delta^+(v)} y_{e,i}, \forall i, v \neq s_i, t_i$$

$$\sum_{e \in \delta^-(s_i)} y_{e,i} = \sum_{e \in \delta^+(t_i)} y_{e,i} = 1, \forall i$$

$$y_{e,i} \in [0,1], \forall e, i$$

Where

1. δ^+ indicates the flow coming into a vertex,
2. δ^- indicates the flow going out of a vertex

Constraint 3 makes sure that we have a unit flow being routed out of every s_i, and a unit flow being routed into every t_i.
Constraint 2 corresponds to flow conservation at every other vertex.
Congestion along an edge is just the number of commodities being routed along that edge, and constraint 1 ensures that t $\geq$ congestion along any edge. The objective function is to minimize t.

9.1.3 Lemma: LP$_1$ and LP$_2$ are equivalent

Given a solution to either problem, we can obtain a solution to the other problem.

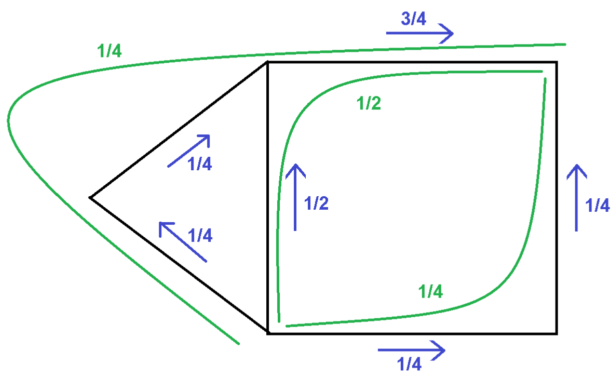

Figure 9.1.1: Example graph for LP$_1$ (green) and LP$_2$ (blue), in which case $\sum_{p\in\mathcal{P}_i} x_{i,P} = \frac{1}{4}+\frac{1}{2}+\frac{1}{4} = 1$

9.1.4 Randomized rounding

Let (x^*, t^*) and (y^*, t^*) are the solution of LP$_1$ and LP$_2$.

For every commodity i, consider the probability distribution on $\mathcal{P}_i$ given by $\{x^*_{i,P}\}$, where $P \in \mathcal{P}_i$. Independently pick one path P of $\mathcal{P}_i$ according to the probability distribution. For example, say we have three paths with values 0.5, 0.3, and 0.2. Then generate a random number between 0 and 1. If the number is between 0 and 0.5, pick the first path. If the number is between 0.5 and 0.8, pick the second path, and if the number is between 0.8 and 1, pick the third path. For any edge e,

$$Pr[\text{commodity i is routed along } e] = x^*_{e,i}$$

Let X_e be a random variable , and let X_e = number of commodities i with $e \in P_i$. In other words, X_e is a random variable which indicates the level of congestion along an edge e. To get the $E[X_e]$, we can use indicator random variables.

Let $X_{e,i} = 1$, if $e \in P_i$, and 0 otherwise.

$$\mu := E(X_e) = E(\sum_i X_{e,i}) = \sum_i E(X_{e,i}) = \sum_{P\in\mathcal{P}_i : e\in P} x^*_{e,i} = \sum_i y^*_{e,i} \leq t^* \leq \text{OPT}$$

This means that the expected value of congestion along any edge $\leq t^*$, the solution to the LP problem. Now we can show that:-

$$\forall e,\ Pr[X_e \geq \lambda \cdot \text{OPT}] \leq \epsilon := \frac{1}{n^3}$$

$$\implies Pr(\text{max. congestion} > \lambda \cdot \text{OPT}) = Pr(\exists e \text{ s.t. } X_e > \lambda \cdot \text{OPT}) \leq \frac{|E|}{n^3} \leq \frac{1}{n^3} \cdot n^2 \leq \frac{1}{n}$$

and we can get a λ approximation with a high probability.

$$X_e = \sum_i X_{e,i}$$

$$E[X_e] = \mu \leq t^* \leq \text{OPT}$$

$$Pr[X_e > \lambda \cdot \text{OPT}] \leq P(X_e > \lambda\mu)$$

According to Chernoff's bound:

$$Pr[\,X_e > \lambda t]\ \leq (\frac{e^{\lambda-1}}{\lambda^\lambda})^\mu \leq (\frac{\lambda}{e})^{-\lambda\mu}$$

Considering $\lambda \approx \frac{3log(n)}{log(log(n))} \implies \lambda^\lambda \approx n^3$, we have $Pr[X_e > \lambda \cdot \text{OPT}] \leq \frac{1}{n^3}$ for $\mu \leq 1$, which fulfills above. And we have $\lambda = O(\frac{3log(n)}{log(log(n))})$

9.2 LP Duality

The motivation behind using an LP dual is they provide lower bounds on LP solutions. For instance, consider the following LP problem:

$$min\ \ 2x_1 + 3x_2, \text{s.t.}$$

$$x_1 - x_2 \geq 4 \tag{9.2.1}$$

$$x_1 + x_2 \geq 5 \tag{9.2.2}$$

$$x_1 \geq 0$$

$$x_2 \geq 0$$

We would like to come with intuitive approach to find smallest solution to the generic LP mentioned above. The idea is we multiply each constraint with a multiplier and add them, such that the sum of coefficients of any variable is $\leq$ the coefficient of the variable in the objective function. This gives us a lower bound on the LP solution. We want to choose the multipliers such that the lower bound is maximized (giving us the tightest possible lower bound).

9.2.1 From primal to dual

First, let continue to work on the primal LP problem mentioned above:

$$(9.2.1) * y_1 + (9.2.2) * y_2 \iff (y_1 + 3y_2)x_1 + (-y_1 + y_2)x_2 \geq 4y_1 + 5y_2$$

Compare with original problem, we see that $(y_1 + 3y_2)$ corresponds to 2 and $(-y_1 + y_2)$ corresponds to 3. Thus, the resulted dual problem is following, from which we are trying to find best Lower Bound on LP-optimal solution:

$$max \;\; 4y_1 + 5y_2, \text{s.t.}$$

$$y_1 + 3y_2 \leq 2$$

$$-y_1 + y_2 \leq 3$$

$$y_1 \geq 0$$

$$y_2 \geq 0$$

Second, let see the way to perform transformation from primal problem to the dual in general case.

Primal LP

$$\min_x \sum_i c_i x_i \text{ , s.t.}$$

$$\sum_i A_{ij} x_i \geq b_j, \, \forall j$$

$$x_i \geq 0, \, \forall i$$

Corresponding dual problem:

$$\max_y \;\; \sum_j b_j y_j, \text{ s.t.}$$

$$\sum_j A_{ij} y_j \leq c_i, \quad \forall i$$

$$y_j \geq 0, \forall j$$

Expressed in matrix form, the dual problem is, Maximize $b^T y$, such that

$$A^T y \leq c$$

$$y \geq 0$$

Note that the dual of a dual LP is the original primal LP.

9.2.2 Weak LP duality theorem

If x is any primal feasible solution and y is any dual feasible solution, then $Val_P(x) \geq Val_D(y)$

Proof:

$$\sum_i c_i x_i \geq \sum_i (\sum_j A_{ij} y_j) x_i = \sum_j (\sum_i A_{ij} x_i) y_j$$
$$\geq \sum_j b_j y_j$$

The weak duality theorem says that any dual feasible solution is a lower bound to the primal optimal solution. This is a particularly nice result in the context of approximation algorithms. In the previous lectures, we were solving the primal LP exactly and using the LP solution as a lower bound to the optimal ILP solution. By using the weak duality theorem, instead of solving the LP exactly to obtain a lower bound on the optimal value of a problem, we can (more easily) use any dual feasible solution to obtain a lower bound.

9.2.3 Strong LP duality theorem

If the primal has an optimal solution x^* and the dual has an optimal solution y^*, then $c^T x^* = b^T y^*$, i.e, the primal and the dual have the same optimal objective function value.

In general, if the primal is infeasible (there is no feasible point which satisfies all the constraints), the dual is unbounded (the optimal objective function value is unbounded). Similarly, if the dual is infeasible, the primal is unbounded. However, if both the primal and dual are feasible (have at least one feasible point), the strong LP duality theorem says that the optimal solutions to the primal and the dual have the exact same objective function value.