

Lecture 16 — March 9, 2023

*Prof. Jelani Nelson**Scribe: Brandon Tran, Laryn Qi*

1 Outline

1. k -server + weighted paging
2. LP Duality recap
3. Online primal-dual

Project Proposal

- Lightweight assignment
- Purpose is primarily for Jelani to catch issues, give advice, provide readings, etc. earlier rather than later
- Max 2 pages

2 k -server problem

The k -server problem can generalize many online problems. The problem is stated as follows: we have an n -point metric space. Each of the k servers lies on a point in the metric space ($k < n$). Points are coming online, and we must move some server to that point if one is not already there. The cost of moving this server is the distance between the server and the destination point. The total cost is the sum of the cost at each step. A metric space is just a set of points with a distance function d defined on them. d satisfies the following properties:

1. The distance from a point x to itself is 0: $d(x, x) = 0$.
2. **Positivity:** $d(x, y) > 0$.
3. **Symmetry:** $d(x, y) = d(y, x)$
4. **Triangle inequality:** $d(x, z) \leq d(x, y) + d(y, z)$

2.1 k -server as a general online problem

1. **Paging** is a special case of k -server. Points \iff pages/memory addresses. $k :=$ size of cache. $n :=$ # of memory addresses (in reality finite, although we often think of it as infinite). $\forall x \neq y : d(x, y) = 1$.

2. **Weighted Paging:** In Weighted Paging, local (in cache) pages cost less than remote pages (e.g. must make a network call). Requesting a page is not always uniform cost. Each page p has cost $c(p)$. $d(x, y) := \frac{1}{2}(c(x) + c(y))$. Initialize all k -servers to the same single NULL point. Every time a page moves to a new point, pay half its cost. Pay the second half when it leaves that point. At the end, move everybody back to NULL to account for remaining costs.

- (maybe) pset5 Q3: Bansal, Buchbinder, Naor showed $O(\lg k)$ competitive for weighted paging [1].

2.2 History of k -server

1. $(2k - 1)$ -competitive "work function" algorithm (deterministic). Shown in [Koutsoupias, Papadimitriou. '95].
2. Conjecture: $O(\lg k)$ randomized is possible. However, [Bubeck, Coester, Rabani '22] (fresh off the press!) showed that you cannot beat $\Omega(\lg^2 k)$.
3. Bubeck et al found $O(\lg^2 k \lg n)$ [4].

3 Recap of Linear Programming and Duality

3.1 Linear Programs

In linear programming, we are given a matrix $A \in \mathbb{R}^{m \times n}$, vectors $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, and we wish to find an $x \in \mathbb{R}^n$ that solves the following optimization problem:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \geq b \\ & x \geq 0 \end{aligned}$$

where the inequalities are taken entrywise. This is called the **primal** linear program. We have a set of m linear constraints, i.e. $A_1^T x \geq b_1$, $A_2^T x \geq b_2$, ... where A_i is the i th row of A . These constraints define a set of half-spaces. The intersection of these half-spaces is always going to be some kind of polytope, with possibly unbounded sides – we call this the feasible region. c can be thought of as "the direction of gravity". Drop a marble in the feasible region, and it follows the direction of c . No matter what, it will eventually end up in a corner.

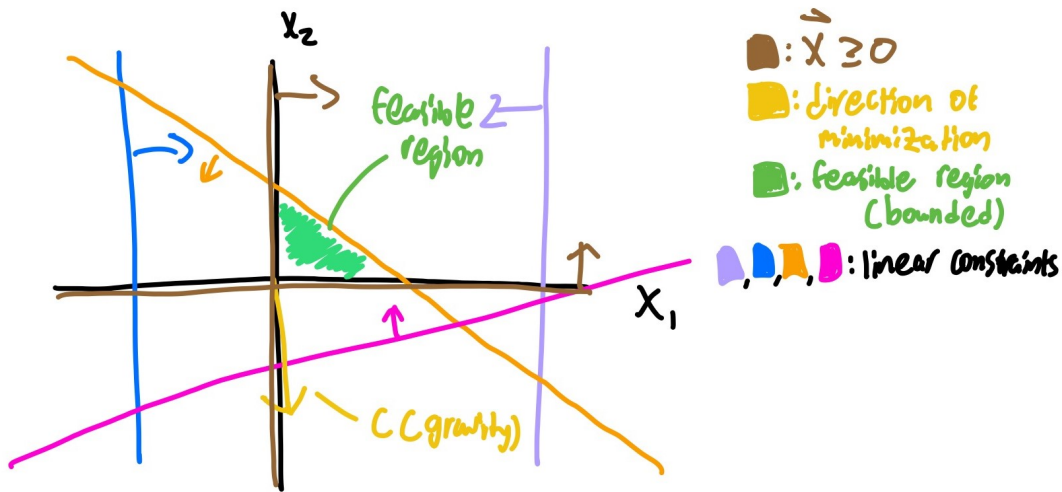


Figure 1: Visualization of an example primal optimization problem

A few notes on the generality of this form of the primal problem:

- We can convert a \geq constraint to a \leq constraint by multiplying by -1 .
- An equality constraint is just a \geq AND \leq constraint. Convert the \geq constraint to \leq as described above.
- If you don't want nonnegative constraints: any number Z can be written as $Z = Z^+ - Z^-$: $Z^+, Z^- \geq 0$.
- If you have maximization problem, negate c to convert to minimization.

3.2 LP Duality

In the primal, we are trying to minimize $c^T x$ subject to $Ax \geq b$ and $x \geq 0$. We would like to provide a lower bound on $c^T x$. This is where duality comes in.

Consider our constraints $Ax \geq b$. For the i th constraint, come up with a new variable y_i and multiply both sides by y_i : $A_i^T x y_i \geq b y_i$. We also require $y_i \geq 0$ so the inequality is preserved. Then, add up all of the constraints. In vector form, this gives us

$$y^T Ax = (A^T y)^T x \geq b^T y$$

It would be great if $A^T y = c$, because this would mean $c^T x \geq b^T y$, and $c^T x$ is what we're trying to lower bound. So $b^T y$ would be a lower bound on the primal LP! It turns out, because $x \geq 0$, it is enough that $A^T y \leq c$, since if we're multiplying the entries of $A^T y$ by nonnegative numbers and adding them up, finding a lower bound on that value implies that $c^T x$ is also lower bounded. Mathematically, $(A^T y)^T x \geq b^T y \implies c^T x \geq b^T y$. Therefore, we define the **dual** linear program as

$$\begin{aligned} \max_{y \in \mathbb{R}^m} \quad & b^T y \\ \text{s.t.} \quad & A^T y \leq c \\ & y \geq 0 \end{aligned}$$

It is possible that the dual is infeasible.

Theorem 3.1 (Weak Duality). $\forall x, y$ feasible for primal/dual respectively. $\text{cost}_{\text{primal}}(x) \geq \text{cost}_{\text{dual}}(y)$. In particular, $\text{OPT}(\text{primal}) \geq \text{OPT}(\text{Dual})$.

Proof. By construction of Primal/Dual. □

Theorem 3.2 (Strong Duality). *Primal bounded (bounded means doesn't go to infinity) and feasible (green region is not empty set) implies dual is as well and $\text{OPT}(\text{primal}) = \text{OPT}(\text{Dual})$*

Proof. In a future lecture. □

4 Online Primal-Dual

[Buchbinder, Naor '09] $x \in \mathbb{R}^n$ starts off as 0. We see constraints (A_i, b_i) come online (from primal) which must be satisfied at each time step. We are allowed to increment entries of x , i.e. the entries of x are only allowed to increase monotonically.

4.1 Approximating OPT

Theorem 4.1 (Approximate Complementary Slackness). *Suppose x, y are feasible solutions to the primal and dual, respectively. The primal has n variables, m constraints, dual is opposite (m variables, n constraints). Variables become constraints and vice versa (e.g. for each primal constraint, we multiply it by a new y_i dual variable).*

1. $x_i > 0 \implies \frac{c_i}{\alpha} \leq (A^T y)_i \leq c_i$ (Second inequality is automatically implied by feasibility of y)
2. $y_i > 0 \implies \beta b_i \geq (Ax)_i \geq b_i$ (Second inequality is automatically implied by feasibility of x)
3. If x, y satisfy 1. and 2., then $c^T x \leq \alpha \beta b^T y$.

Proof. pset5, Q2 □

Significance: In online algorithms, we aren't going to get a competitive ratio of 1, we can only approximately solve the LP. It turns out, approximate complementary slackness gives us a "recipe" for approximating the LP. Essentially what we do is as follows: set up primal and dual LPs, and as points come online, increment x and y such that we maintain feasible solutions to the primal and dual and such that they satisfy the conditions of approximate complementary slackness. At the end, the values of the two problems gives some guarantees about the quality of our solution to the primal with respect to OPT.

$$b^T y \leq \text{OPT}(\text{Dual}) = \text{OPT}(\text{Primal}) \leq c^T x \leq \alpha\beta b^T y \leq \alpha\beta \cdot \text{OPT}$$

which means

$$c^T x \leq \alpha\beta \cdot \text{OPT}$$

So if we maintain primal and dual feasible solutions that satisfy the conditions of approximate complementary slackness, it means that the x we found is “not too far” from OPT.

4.2 Ski rental

Recall the ski rental problem from lecture 14. We formulate the primal LP for this problem as

$$\begin{aligned} \min_x \quad & Bx + \sum_{i=1}^n z_i \\ \text{s.t.} \quad & x + z_i \geq 1 \quad \forall i \\ & x, z \geq 0 \end{aligned}$$

B is the cost of buying the skis, $x \in \{0, 1\}$ represents whether we bought the skis or not. $z_i \in \{0, 1\}$ represents whether or not we rented on day i . On each day, we have to have either bought or rented skis. Below are what the matrices look like for the primal.

$$A = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ \vdots & 0 & 0 & \ddots & \vdots \\ 1 & 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} x \\ z_1 \\ \vdots \\ z_n \end{bmatrix}$$

$$c = \begin{bmatrix} B \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$b = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

The dual is

$$\begin{aligned} \max_y \quad & \sum y_i \\ \text{s.t.} \quad & \sum y_i \leq B \\ & y_i \leq 1 \quad \forall i \\ & y \geq 0 \end{aligned}$$

Algorithm: Maintain primal+dual feasible solutions x, z_i 's y_i 's (initially all 0). Seeing a new constraint is like seeing a new variable y_i . Gradually increase y_i until one of the dual constraints y_i is involved in $(y_i \leq 1, \sum y_i \leq B)$ becomes tight. Set the corresponding primal variable to 1. In terms of approximate complementary slackness: if $x > 0 \implies x = 1 \implies \sum y_i = B \implies \alpha = 1$ or if $z_i > 0 \implies y_i = 1 \implies \alpha = 1$ so in either case $\alpha = 1$. Since $x + z_i \leq 2$ (sum of two 0-1 variables), which means $\beta = 2$. These values of α and β implies $\text{cost}(\text{algorithm solution}) \leq 2OPT(\text{primal}) \leq 2OPT(\text{ski rental})$. Note that our primal problem formulation is not exactly equivalent to ski rental since we're missing integrality constraints. However, adding more constraints to our primal problem could only increase the min, implying the second inequality: $\implies OPT(\text{ski rental}) \geq OPT(\text{primal})$

In online primal-dual, it is possible (depending on the problem) to get a constraint which is impossible to satisfy.

Claim 4.2. With randomization, competitive ratio of roughly $\frac{e}{e-1}$ is possible.

Idea: Will maintain fractional solutions online (i.e. not integer LP, we'll fix this later).

Algorithm: See new day i

- if $x \geq 1$, do nothing (i.e. $x + z_i \geq 1$ is already satisfied)
- set $y_i \leftarrow 1$
- $z_i \leftarrow 1 - x$
- $x \leftarrow (1 + \frac{1}{B})x + \frac{1}{cB}$, c is to be determined.

Recall: $b^T y \leq OPT \leq c^T x$. If $c^T x \leq \gamma b^T y$, then $c^T x \leq \gamma OPT$. Want to upper bound $\frac{c^T x}{b^T y} \leq \gamma$. If we can find this bound, this means we have a γ competitive solution.

Primal solution feasible: $y_i = 1 \geq 0$, $z_i \leftarrow 1 - x$, and x becomes bigger.

- $z_i \geq 0$ since $x \leq 1$ and $z_i := 1 - x$.
- $x \geq 0$ since x starts at 0 and can only increase.
- $x + z_i \geq 1$ since $x + z_i = x + (1 - x) = 1 \geq 1$.

Dual solution feasible: Need to show $x = 1$ after $\leq B$ days. $\frac{c^T x}{b^T y}$. Look at $\frac{\Delta c^T x}{\Delta b^T y}$ each day. If these are bounded by γ , then the end result is also bounded by γ . If x isn't already 1: $\Delta \text{dual} = 1$. $\Delta \text{primal} = B\Delta x + 1 - x = B((1 + \frac{1}{B})x + \frac{1}{cB} - x) + 1 - x = Bx + x + \frac{1}{c} - Bx + 1 - x = 1 + \frac{1}{c}$. This means $\frac{\Delta \text{primal}}{\Delta \text{dual}} \leq 1 + \frac{1}{c}$.

Now we'll show that $x = 1$ after $\leq B$ days. Let $r := \frac{1}{cB}$, $q := 1 + \frac{1}{B}$, and $x_i :=$ value of x at day i .

$$x_{i+1} \leftarrow \left(1 + \frac{1}{B}\right)x_i + \frac{1}{cB} = qx_i + r;$$

$$x_0 := 0$$

$$x_1 = r$$

$$x_2 = qr + r$$

$$x_3 = q^2r + qr + r$$

...

$$x_k = r \sum_{i=0}^{k-1} q^i$$

Want: when $k = B$, $x_k = 1$ (otherwise, this ruins dual feasibility) \implies want $r \sum_{i=0}^{B-1} q^i = 1$.

$$r \frac{q^B - 1}{q - 1} = \frac{1}{cB} \frac{\left(1 + \frac{1}{B}\right)^B - 1}{\frac{1}{B}}$$

To make this 1, set $c := \left(1 + \frac{1}{B}\right)^B - 1$. $\left(1 + \frac{1}{B}\right)^B \approx e$, so $c \approx e - 1$. Which means $\frac{\Delta_{\text{primal}}}{\Delta_{\text{dual}}} \leq 1 + \frac{1}{e-1} \approx \frac{e}{e-1}$. This proves the dual solution is feasible. Now, we just need to fix this fractional solution by making it integer.

On first day added x_1 to 0, then added x_2 to that, then x_3 to the result of that, etc. Pick $u \in [0, 1]$ unif. at random.

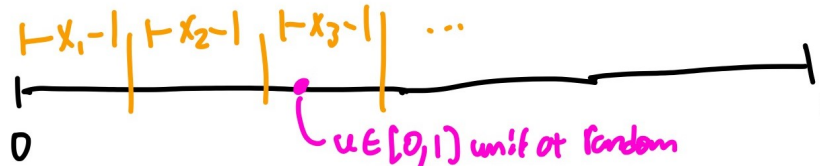


Figure 2: Visualization of fixing fractional solution

Buy on the first day that x exceeds u . The expected cost is $B \cdot P(\text{buy}) + \sum_i P(z_i = 1)$. $P(\text{buy}) = P(u \leq x) = x$, $P(z_i = 1) = z_i \implies E[\text{cost}(\text{primal})] = Bx + \sum_i z_i = \text{cost}(\text{primal})$.

References

- [1] Nikhil Bansal, Niv Buchbinder, Joseph (Seffi) Naor. A Primal-Dual Randomized Algorithm for Weighted Paging. *Foundations of Computer Science (FOCS)*, IEEE Annual Symposium on (2007)
- [2] Elias Koutsoupias and Christos H Papadimitriou. On the k-server conjecture. *Journal of the ACM*, 42(5):971–983, 1995.
- [3] Sébastien Bubeck, Christian Coester, Yuval Rabani. Shortest Paths without a Map, but with an Entropic Regularizer. *arXiv*, 2022
- [4] Bubeck et al. k-server via multiscale entropic regularization. *arXiv*, 2017, STOC 2018
- [5] Niv Buchbinder, Joseph (Seffi) Naor. The Design of Competitive Online Algorithms via a Primal-Dual Approach. *Foundations and Trends in Theoretical Computer Science*, Vol.3, Nos. 2-3, 2009