Theorist's Toolkit	**Spring 2019**

Lecture 7: Brunn-Minkowski and Sorting Partial Orders

Scribe: Antares Chen and Jon Hillery 5/20/2019

7.1 Intro

Our goal for today is an extension of efficient sorting to circumstances in which we already have partial information. Recall that the fastest we can sort a completely unsorted set of n items is in $\Theta(n \log n)$ time. We can view sorting with partial information as extending a partial order to a linear (i.e. total) order:

Definition 7.1. *A **partially-ordered set** (a **poset**) is a set X together with a relation $\leq$, i.e. a subset of $X \times X$ satisfying:*

1. $(x, x) \in \leq$ *for all* $x \in X$
2. $(x, y), (y, z) \in \leq \implies (x, z) \in \leq$
3. $(x, y), (y, x) \in \leq \implies x = y$

*We also write $(x, y) \in \leq$ in the more familiar way $x \leq y$. A **linearly ordered set** is a poset for which every pair of items is comparable, i.e. for each $x, y \in X$, either $x \leq y$ or $y \leq x$.*

We can think of sorting given partial information as using an oracle function to find an **extension** of a partial order, i.e. a linear order such that if $x \leq y$ in the partial order, then $x \leq y$ in the total order. We denote $E(\leq)$ as the set of all linear extensions of $\leq$ and $e(\leq)$ as the cardinality of $E(\leq)$.

We know that when the poset $\leq$ is empty, we can sort at best in $\log(e(\leq))$ time ($n \log n$). Our goal today will be to show how we can sort in $\log e(\leq)$ time regardless of the poset. We will do so using the following main theorem:

Theorem 7.2. *For any $\leq$ which is not linear, there exist a, b that are not in the ordering such that*

$$\frac{1}{2e} \leq \frac{e(\leq \cup (a, b))}{e(\leq)} \leq 1 - \frac{1}{2e}$$

where e is the base of the natural logarithm.

This gives efficient sorting because we reduce the size of $e(\leq)$ by a factor of $\frac{1}{2e}$ when we compare a and b. The upper bound is necessary to give this for when $b \leq a$ when we compare them since $e(\leq) = e(\leq \cup (a, b)) + e(\leq \cup (b, a))$, giving by the upper bound of the inequality:

$$\frac{e(\leq \cup (b, a))}{e(\leq)} = 1 - \frac{e(\leq \cup (a, b))}{e(\leq)} \leq 1 - \frac{1}{2e}$$

We can think of this process therefore as somehow finding an a, b that reduces the size of $e(\leq)$ as symmetrically as possible since we don't know beforehand whether $a \leq b$ or $b \leq a$.

7.2 Some Geometry Preliminaries

We can think of this problem geometrically through something called the order polytope.

Definition 7.3. *The* ***order polytope*** *of* $\leq$, $P(\leq)$, *is the set of points* $x = (x_{a_1}, ..., x_{a_n})$ *in* $[0,1]^n$ *(where* n *is the number of items) satisfying* $x_a \leq x_b$ *for all* $(a,b) \in \leq$*. An* ***up-set*** *of these points is a maximal increasing chain in the poset.*

It is not too difficult to see that the vertices of the order polytope are exactly the characteristic vectors of the up-sets in the poset (where a characteristic vector is 1 if the item is in the set and 0 otherwise). This follows because the vertices must be integers (otherwise we could find a further extremal point that is an integer), and they clearly must form this ascending chain because otherwise we could complete the chain to get a "more extremal" point.

Lemma 7.4. *The volume of* $P(\leq)$ *when* $\leq$ *is a linear ordering is* $\frac{1}{n!}$.

Proof. The set of linear orderings partitions the unit cube. This follows since each must be distinct (by definition of the points), and for any point in the unit cube we get a linear ordering by just looking at the value of its coordinates. Since there are $n!$ orderings and the unit cube has unit volume, the volume of each ordering is $\frac{1}{n!}$ (showing the order polytopes are congruent follows roughly from the fact that linear orderings have the same polytope up to a permutation of the standard basis, which is an isometry). □

Therefore, the order polytope of each partial ordering $\leq$ is just $e(\leq) \times \frac{1}{n!}$, and so we can view picking our desired a, b as finding a place to slice the order polytope that equally distributes the volume on both sides of our slice.

Definition 7.5. *The* ***height*** $h_{\leq}(a)$ *of an element* a *in a linear order* $\leq$ *is the number of items which are less than or equal to it. In a partial order, we define the height as the average height of that item over all possible linear extensions of the partial order.*

Lemma 7.6. *The center of gravity of an order polytope* $P(\leq)$ *is the point* $(\frac{1}{n+1}h_{\leq}(a_1), ..., \frac{1}{n+1}h_{\leq}(a_n))$.

Proof. Since each order polytope is a disjoint union of congruent total order polytopes, the center of gravity is just the average of the centers of gravities of those total order polytopes. The lemma holds for a total order polytope since the given coordinates are just the average of its vertices (since they are *maximal* ascending chains). Because the height of an item was defined as the average height over all linear extensions, this completes the proof. □

7.3 The Efficient Comparison Theorem

We are now ready to begin to prove our main theorem. So, suppose we have partial order $\leq$ on X which is not linear.

Lemma 7.7. *There exist $a, b \in X$ such that $|h_{\leq}(a) - h_{\leq}(b)| < 1$.*

Proof. Suppose not. Then, by the pigeonhole principle we must have that the heights of the elements are exactly $1, 2, 3, ..., n$ since each must be between 1 and n. However, if the height of the first element a_1 is 1 then it must be the least element in every linear extension. Then the second item a_2 satisfies $a_1 \leq a_2$ and $a_2 \leq a_2$, and so a_2 must always be the second item in every linear extension. By induction, we already know the total order on all elements, meaning $\leq$ is a linear order, and thus giving a contradiction. □

We claim that these two elements a, b satisfy the desired property of the theorem. Let $P = P(\leq)$, $P_{\leq} = P(\leq \cup(a, b))$, and $P_{\geq} = P(\leq \cup(b, a))$. We again have that P is the disjoint union of $P_{\leq}$ and $P_{\geq}$, and the two are separated by the hyperplane $h = \{x_a = x_b\}$. Since we have shown that volumes correspond to possible linear extensions, it is sufficient to show the following:

$$\frac{vol(P_{\leq})}{vol(P)} \geq \frac{1}{2e}, \frac{vol(P_{\geq})}{vol(P)} \geq \frac{1}{2e}$$

To show this, we first perform a unitary change of coordinates to ones that are easier to work with. This does not affect the proof since it preserves volume. We change to coordinates y_i where $y_1 = x_a - x_b$ and the other y_i form an orthonormal basis with y_1. Our separating hyperplane is now $h = \{y_1 = 0\}$. This change of coordinates gives two nice properties that we claim will be sufficient to complete the proof:

1. The projection of P onto the y_1-axis is exactly $[-1, 1]$. This follows because we can find an up-set which contains a but not b and one which contains b but not a because otherwise (a, b) or (b, a) would be in $\leq$. By convexity, the projection is therefore the whole interval.

2. Let c be the center of gravity of P. Since $|h_{\leq}(a) - h_{\leq}(b)| < 1$, we have that $c_1 = \frac{1}{n+1}(h_{\leq}(a) - h_{\leq}(b))$ satisfies $-\frac{1}{n+1} < c_1 < \frac{1}{n+1}$.

Let P_t be the subset of P such that $y_1 = t$. Define $r(t)$ to be the radius of an $(n-1)$-ball with volume $vol_{n-1}(P_t)$. Then the **Brunn-Minkowski Inequality** gives that $r(t)$ is concave on $[-1, 1]$ (this is the only information we need to know about Brunn-Minkowski). Furthermore, since we can write c_1 as,

$$c_1 = \frac{\int_{-1}^{1} t vol_{n-1}(P_t) dt}{vol(P)}$$

we have that c_1 only depends on $r(t)$! Therefore, we can ignore a lot of the original geometry of our polytope and smooth it out to something nice (which will we make rigorous in the following sections). As discussed previously, the following theorem will complete our proof of the efficient comparison theorem:

Lemma 7.8. $c_1 \geq \frac{-1}{n+1} \implies vol(P_{\geq}) \geq \frac{1}{2e} vol(P)$ *(and a symmetric argument gives the same result for $P_{\leq}$).*

Proof. Define a body K as follows: for each t, let the slice of K with the hyperplane $\{y_1 = t\}$ be the $(n-1)$-ball with volume $k(t)$ where k is the following function:

For $t \geq 0$, define $k(t) = -\frac{r(0)}{u}t + r(0)$ for a u that makes it so $vol(K_{\geq}) = vol(P_{\geq})$. Such a u exists since as $u \to 0$, $vol(K_{\geq}) \to 0$ and as $u \to \infty$, so does $vol(K_{\geq})$ (this can be shown with integration; proof omitted). Therefore, by the intermediate value theorem, there exists some u that works. Since k is linear (where we have defined it) and agrees with r at 0, by concavity of r we have that k is less than r for smaller values of t

and then they intersect and then k is greater than r. In particular, the center of gravity of $K_{\geq}$ is further right (greater y_1-component)than that of $P_{\geq}$.

Now, define k for negative t in a somewhat similar fashion: such that $vol(K_{\leq}) = vol(P_{\leq})$, $k(-1) = r(-1) = 0$, and k is a line with the same slope as it had for positive values for some interval with right endpoint 0 and then it is a different line with positive slope. Another intermediate value theorem argument gives that this is possible. Furthermore, k will be greater than r for some interval ending at 0 and then for some leftmost interval it will be less than r, meaning that the center of gravity of $K_{\leq}$ will also be further right than $P_{\leq}$. Consequently, $c_1(K) \geq c_1(P) \geq -\frac{1}{n+1}$.

Let K_1 be the subset of K defined by the interval of length h_1 on which k has positive slope, and K_2 be the rest of K (where k is decreasing), which has length h_2. Let Δ be the difference between h_1 and the length 1 of $K_{\leq}$. Then,

$$c_1(K) = \frac{1}{vol(K_1)}c_1(K_1) + \frac{1}{vol(K_2)}c_1(K_2) = \frac{1}{h_1}c_1(K_1) + \frac{1}{h_2}c_1(K_2)$$

(since the volume integral will scale as a factor of the length of the base). Now, the center of gravity of an $(n-1)$-cone is at $\frac{1}{n+1}$ times its height since similar cones have volumes that are the $n$the power of their similarity ratio and $\int x^n = \frac{x^{n+1}}{n+1}$, which attains the average at $\frac{1}{n+1}$. Therefore,

$$c_1(K) = \frac{h_1(\frac{-h_1}{n+1}) + h_2(\frac{h_2}{n+1})}{h_1 + h_2} = \frac{h_2 - h_1}{n+1} - \Delta$$

and since $h_1 + \Delta = 1$ and $c_1(K) \geq -\frac{1}{n+1}$, we have $h_2 - h_1 - \Delta(n+1) \geq -1$ and so $h_2 + nH_1 \geq n$. Since $h_1 = (u+1) - h_2$, we have that $\frac{u}{h_2} \geq 1 - \frac{1}{n}$. Now, $K_{\geq}$ is similar to K_2 with ratio $\frac{u}{h_2}$, and so $vol(K_{\geq}) = (\frac{u}{h_2})^n vol(K_2) = (\frac{u}{h_2})^n \frac{h_2}{h_1+h_2} vol(K) = \frac{u}{u+1}(\frac{u}{h_2})^{n-1} vol(K)$. Putting this all together, $vol(K_{\geq}) \geq \frac{u}{u+1}(1 - \frac{1}{n})^{n-1} vol(K)$. Since $u \geq 1$, we know that $\frac{u}{u+1} \geq \frac{1}{2}$. We can furthermore see that $(1 - \frac{1}{n})^{n-1} > \frac{1}{e}$ since it is greater for small values of n, is monotonically decreasing, and converges to $\frac{1}{e}$. Therefore, $vol(K_{\geq}) \geq \frac{1}{2e} vol(K)$, completing the proof (since each K part had the same volume as the corresponding P part). $\square$

7.4 Further Results and the $\frac{1}{3} - \frac{2}{3}$ Conjecture

The $\frac{1}{3} - \frac{2}{3}$ **Conjecture** is that we can improve the $\frac{1}{2e}$ bound to $\frac{1}{3}$. We know that this is the best we could do since the ordering $a \leq b$ on $\{a, b, c\}$ satisfies $e(\leq) = 3$ since we could have $c \leq a \leq b$, $a \leq c \leq b$, $a \leq b \leq c$. Therefore, $\frac{e(\leq \cup (a,b))}{e(\leq)} \geq \frac{1}{3}$.

Also, note that the a and b we find was non-constructive. We can actually calculate an a and b in polynomial time using known volume-approximation algorithms, but there is not a general efficient way to find these best choices of comparison; we just know they exist.