MIT 6.854/18.415: Advanced Algorithms	Spring 2016
Lecture 23 – May 2, 2016	
Prof. Ankur Moitra	*Scribe: David Ashpole*

1 Previous Lecture

We talked about using compressed sensing to recover almost k-sparse vectors in $O(klog(n/k))$ measurements.

2 Today: Smoothed Analysis

The worst-case analysis is often too pessimistic, and the average case analysis is sensitive to the distribution.

Smoothed Analysis: [1]

$$\max_x \mathbb{E}_\sigma[\text{time}(Alg(x+\sigma))], \text{ where } \sigma \text{ is a Gaussian Perturbation}$$

We have three approaches to LPs: (1) Simplex, (2) Ellipsoid, (3) Interior Point

Theorem 1. *(The simplex method runs in smoothed polynomial time)*

Smoothed analysis has been applied to Mathematical Programming, Numerical Analysis, Learning, Approximation Algorithms, etc.

Today we will cover knapsack, following [2]
Given Values $v_i \in Values$, and weights $w_i \in Weights$, Find:

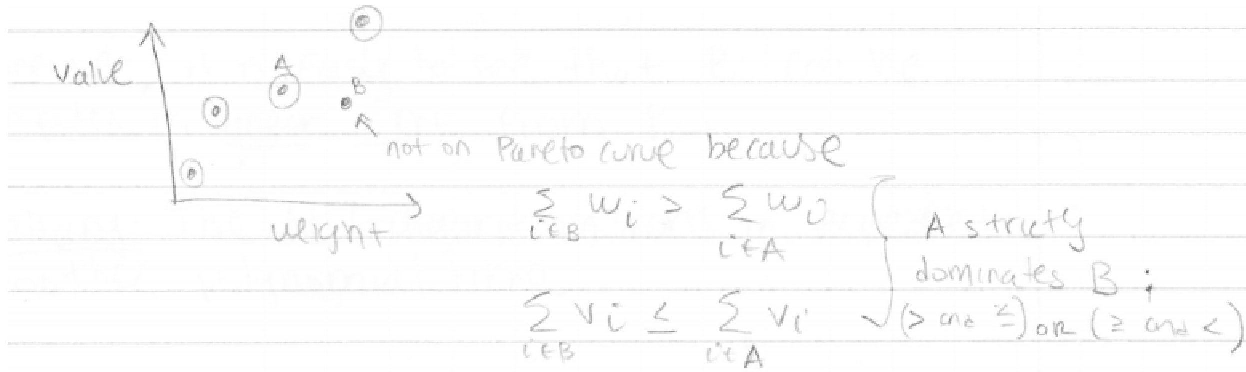
$$\max \sum x_i v_i \text{ s.t. } \sum x_o w_i \leq W;\ x_o \in \{0,1\}$$

Knapsack is NP-Hard, but often easy.

3 Namhauser-Ullman Algorithm

Set $P_o = \emptyset$
For i = 1 to n:
____ Let $T\{A\}_{A\in P_{i-1}} \cup \{A \cup \{i\}\}_{A \in P_{i-1}}$
____ Remove every set from from T if the set is strictly dominated by any other
Find $A \in P_n$ with $\sum_{i\in A} w_i \leq W$, that maximizes $\sum_{i \in A} v_i$

This algorithm constructs Pareto Curves, e.g.



Lemma: Each P_i is the Pareto curve for $2^{[i]}$
Proof: By induction P_{i-1} is the Pareto Curve for $2^{[i]}$
Consider $B \leq [i]$ with $i \in B$. Then if $B \backslash \{i\} \notin P_{i-1}$, B cannot be on P_i because:
If $A \subset [i-1]$ and A strictly dominates $B\backslash\{i\}$, then $A \cup \{i\}$ strictly dominates B.
Thus all feasible candidates for P_i are considered, and
P_i is the Pareto curve for $2^{[i]}$

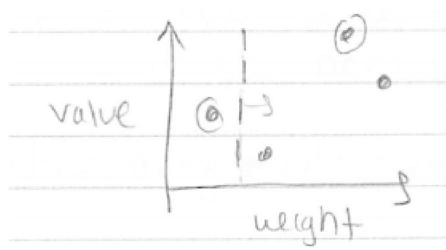
Corrilary: Namhauser-Ullman Algorithm returns the optimal solution
Proof: P_n is the pareto curve for $2^{[n]}$

When is the NU-Algorithm efficient?
Worst Case: $|P_n| \geq C^n$
Theorem: (informal)[2] The expected size of each Pareto Curve P_i is polynomial in the smoothed analysis model.
Moreover, it is easy to see that P_i can be computed in <u>linear time</u> from P_{i-1}

Corollary: The NU-Algorithm runs in expected smoothed polynomial time.
New let's define the relevant smoothed model:

- Let Z_i be independent r.vs whose pdf is bounded by $\emptyset$, supported in [0,1]
- Let $v_i = v_i' + z_i$, $v_i \in [0,1]$, and v_i' is the worst case.
- Let w_i be worst case; arbitrary but distinct.

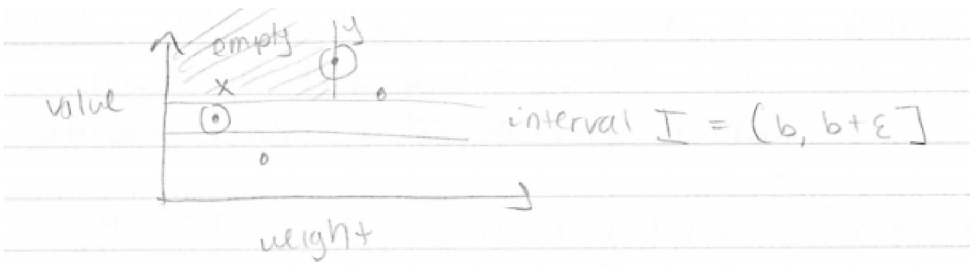
Our Goal is to build up a family of events that will let us bound the size of the Pareto Curve (P_O)
Step 1: A definition of Pareto Optimal, via sweeping
Consider sweeping from <u>low</u> to <u>high</u> weight.



Observation: A point X is Pareto Optimal iff when it arrives, it has strictly largest value.
Aside: If 2^n points had been Gaussian values (average case unstructured rather than smoothed) we immediately have:

$$\mathbb{E}[|PO|] = \sum_{i=1}^{2^n} 1/i \approx n\ln(2)$$

Step 2: Find an event to blame when $x \in PO$
Divide $[0, 2n]$ into intervals of width ϵ
Now if $x \in PO$, we can continue to sweep and find the next point $y \in PO$



Let i be a coordinate s.t. $x_i \neq y_i$. To keep things simple, suppose $x_i = 1$, $y_i = 0$ (other case is basically the same)
New we are ready to define the family of events, E, specified by interval I and index i, and a bit a.

$E \triangleq$ There is an $x \in PO$ with $x \in I$, and if y is the next point on PO, $x_i = a, y_i = \bar{a}$
How many events are there? $(2n/\epsilon)(n)(2) = 4n^2/\epsilon$

Claim: If no two points land in the same interval,

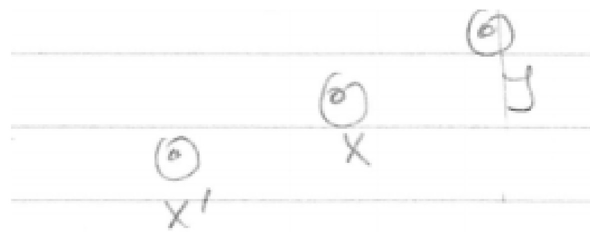
$$|PO| \leq \sum_E \mathbb{1}_E + 1$$

The 1 represents the last point on PO with no y.

Step 3: Bound the probability of each event.
Lets consider the $x_i = 1$, $y_i = 0$ case, and do some backwards reasoning:

Lemma: If $v_1, v_2, ..., v_{i+1}, ..., v_n$ are fixed, there is a unique x that can cause E
Proof: Let $I = (b, b+\epsilon)$. Then the point y must be the point with smallest weight among those with $val(y) > b + \epsilon$.
Note: if $y_i = 1$, we already know E does not occur.

Now x is the point among those with $x_i = 0$, $weight(x) < weight(y)$ that has largest value.
Why? For any other point x' we have:



Furthermore, if $x' \in I$, then $val(x) > b + \epsilon$ (no two points in the same interval), but all other points y' with $val(y') > b + \epsilon$ and $y'_i = 1$ are right of y.
Thus the next PO point after x' cannot have the i^{th} coordinate equal to zero, so E does not happen.

To finish, v_i is still random, so there is at most an $\emptyset\epsilon$ change x lands in i. Thus:
Lemma: $Pr[E] \leq \epsilon\emptyset$
Putting it all together we have: ($\epsilon \to 0$, so no two in same interval a.s.)

$$\mathbb{E}[|PO|] \leq 4n^2\emptyset + 1$$

The exciting takeaway is that the explanatory power of theory is not necessarily limited to the worst case or average case. When faced with a hard problem, explore it in weaker models.

References

[1] Spielman, D. and Teng, S. 2004. Smoothed Analysis of Algorithms. *Journal of the ACM.* 79:385–463.

[2] Beier, R. and Vöcking, B. 2004. An Experimental Study of Random Knapsack Problems. *Springer Berlin Heidelberg.* pp. 616-627.