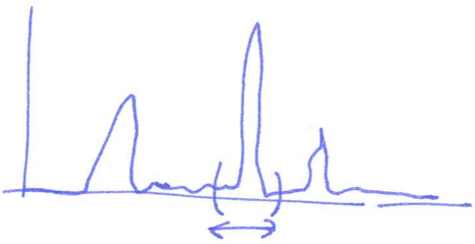
## Smoothed Analysis

Instead of figuring out $\max_{|I|=n} \text{runtime}(I)$

we consider $\max_{|I|=n} \underbrace{\mathop{E}_{I' \sim \text{neighborhood}(I)} \left[\text{runtime}(I')\right]}_{\text{Smooth out the runtime}}$



What is the definition of the neighborhood?
Depends on the problem — but we want it to be big enough so that we can smooth out the performance.

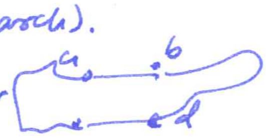
[If smooth complexity is small, means that the worst cases are isolated].

---

- **Major results**: Spielman & Teng showed that a certain pivoting rule for simplex has polynomial smoothed complexity under a certain noise model.
  Improvements by Deshpande & Spielman, and Vershynin.
  Results by Beier & Vöcking on knapsack and other problems.

---

Example: the 2OPT heuristic for TSP. (local search).

- Start with an arbitrary tour. While ∃ a pair  such that the swap $\begin{matrix}(ab)\\(cd)\end{matrix} \to \begin{matrix}(ad)\\(bc)\end{matrix}$ gives a lower cost solution, take it.
- Know that ∃ exponentially long improving paths in some instances.
  But does well in practice.

(Simplified) Smoothed Model: Fix a graph $G=(V,E)$.
For each edge $e$, give a density function $f_e : [0,1] \to [0, \phi]$. (this being small means more "diffuse")
$\Rightarrow$ the edge lengths are drawn independently from $f_e$.

Now run the 2OPT heuristic. (on the graph, not metric completion).

Claim: longest path in the improvement graph is $O(n^6 \log n \cdot \phi)$.

states are tours, directed edge from $T_1$ to $T_2$ if ∃ a swap to get from $T_1 \to T_2$, which is improving.

2

Nb. this means that no matter where we start, after $O(n^6 \log n\, \phi)$ steps we get to a local optimum.

---

Pf: initial tour length $= n$. Suppose the smallest improvement were $\Delta$, then would take $\leq n/\Delta$ steps.

smallest improvement $= \min_{e_1 e_2 e_3 e_4} [\ell(e_1)+\ell(e_2)-\ell(e_3)-\ell(e_4)]$

Claim: $\Pr[\Delta \leq \varepsilon] \leq n^4 \varepsilon \phi$

$\Rightarrow \Pr[T \geq t] \leq \Pr[\Delta \leq n/t] \leq n^5/t \cdot \phi$

$\Rightarrow E[T] = 1 + \sum_{t=1}^{n!} \Pr[T \geq t] \leq n^5 . \phi . \ln(n!) \leq O(n^6 \lg n . \phi)$.

Pf of claim: fix the 4 edges. $n^4$ ways. Now having fixed the lengths of 3 edges,

$\Pr[\cancel{\mathrm{len}(e_4)}\ \Delta_{e_1 e_2 e_3 e_4} \in [0, \varepsilon]] = \Pr[\mathrm{len}(e_4)$ lies in some $\varepsilon$ length interval$]$
$\leq \epsilon\phi$.

$\Rightarrow n^4 . \epsilon\phi$.

$\blacksquare$

---

This was a simplified model of inputs.

in [ERV: Englert Röglin Vöcking] they show that for a smoothed model of choosing ($f_v : [0,1]^d \to [0,\phi]$.) points in ~~the~~ Euclidean space $\mathbb{R}^d$ (for $d$ = constant),

$E[$length of 2OPT improving moves$] \leq O(n^{4+1/3} \log(n\phi) . \phi^{8/3})$

And can improve if choose start point smartly, etc.

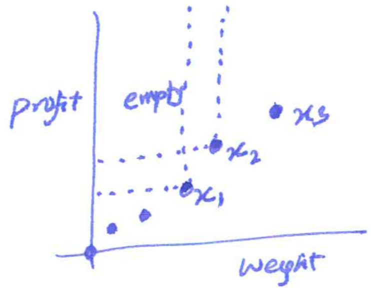
Note: only finds local optima, which may be ~~quite~~ bad. (but not too bad]

[ERV]: the expected approximation ratio for a locally optimal tour chosen from a distributions ~~that is~~ where each point $v$ is picked from $f_v : [0,1]^d \to [0,\phi]$ is $O(\sqrt[d]{\phi})$. [local opt wrt 2opt].

**Knapsack:** sizes/weights $w_i$, profits $p_i$, size = 1 (say).

[Nemhauser/Ullman] Build the pareto curve and pick the best one from it.

- **Pareto curve:** $x \in \{0,1\}^n$ on the pareto curve if $\forall x'$, $w.x' \le w \cdot x \Rightarrow p.x' \le p \cdot x$




- Let $\mathcal{P}(i)$ = pareto curve solutions for items $\{1..i\}$.

  **Fact:** $\mathcal{P}(i+1) \subseteq \mathcal{P}(i) \cup \{x + e_{i+1} \mid x \in \mathcal{P}(i)\}$

  $\Rightarrow$ can compute $\mathcal{P}(n)$ in time $O\left(\sum_{i=1}^{n} |\mathcal{P}(i)|\right)$.

  We'll ignore the size of knapsack, assume we're just looking at unrestricted Pareto curve.

**Smoothing model:** say the weights are chosen randomly, with $w_i : [0,1] \to [0, \phi]$.

$\phi$ = "smoothing parameter". Can generalize this to $w_i : [0 \phi_0] \to [0, \phi]$ etc, but see paper. (~~hmm, why does this matter?~~)

**Claim:** $E[\text{size of pareto curve on } n \text{ items}] = O(n^2 \phi)$

**Pf:** Let's break the weight axis into pieces of width $1/K$, and hence the size of $\mathcal{P}(n)$ can be written as

$$(*) \quad 1 + \lim_{K \to \infty} \sum_{i=0}^{\infty} \mathbb{1}\left(\text{there exists } x \in \mathcal{P}(n) \text{ with } w \cdot x \in \left(\frac{i}{K}, \frac{i+1}{K}\right]\right)$$

all zeros.

N.b. because of smoothing, no two solutions can have same weight (w.p. 1).

N.b. 2: all weights $\in [0,1]$ so total size $\in [0,n]$.

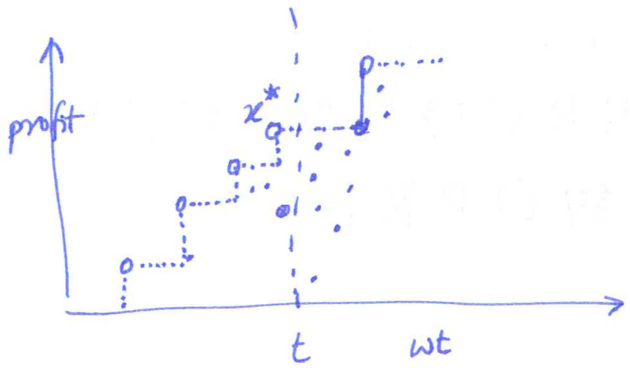
$$(*) = 1 + \lim_{K \to \infty} \sum_{i=0}^{nK} \mathbb{1}\left(\exists\, x \in \mathcal{P}(n) \text{ with } w \cdot x \in \left(\frac{i}{K}, \frac{i+1}{K}\right]\right)$$

For any threshold $t$, define $\Delta(t) = \begin{cases} w \cdot \hat{x} - t & \text{for "least upper bound" } \hat{x} \text{ above } t. \\ \infty & \text{if no such } \hat{x} \text{ exists} \end{cases}$

at $t$:

$x^* = \operatorname{argmax}\{p \cdot x \mid x \in \{0,1\}^n \wedge w \cdot x \le t\}$ ← winner

$\hat{x} = \operatorname{argmin}\{w \cdot x \mid x \in \{0,1\}^n \wedge p \cdot x > p \cdot x^*\}$ ← l.u.b.




Note: many ways to define winner & lub, these definitions are carefully chosen to make proofs work.

$$\Rightarrow \quad (*) = 1 + \lim_{k\to\infty} \sum_{i=1}^{nk} \mathbb{1}\left(\Delta(\tfrac{i}{k}) \in [0, 1/k]\right)$$

$$\boxed{\text{Claim: } \forall t \ \Pr[\Delta(t) \in (0,\epsilon)] \le n\phi\epsilon.}$$

$$\Rightarrow \quad E[|\text{Paretocurve}|] \lessapprox \lim_k \sum_{i=1}^{nk} \Pr[\Delta(i/k) \in (0, 1/k]] \le \lim_k \sum_{i=1} \frac{n\phi}{k} = n^2\phi. \ \checkmark$$

Proof of Claim:

Define
$$x^{*(i)} = \operatorname{argmax}\{p\cdot x \mid x_i = 0 \wedge w\cdot x \le t\}$$
$$\hat{x}^{(i)} = \operatorname{argmin}\{w\cdot x \mid x_i = 1 \wedge p\cdot x > p\cdot x^{*(i)}\}$$

$$\Delta^{(i)}(t) = \begin{cases} w\cdot\hat{x}^{(i)} - t & \text{if } \hat{x}^{(i)} \text{ exists} \\ \infty & \text{otherwise} \end{cases}$$

Subclaim: either $\Delta(t) = \infty$ or $\Delta(t) = \Delta^{i}(t)$ for some $i$.

Pf: if there is some $x^*$ and $\hat{x}^*$ at $t$, then $\exists$ element $i$ in $\hat{x}$ not in $x^*$ (since wts are $\ge 0$)
now this will give $\underset{x^*}{\overset{\shortparallel}{x^{*(i)}}}$ and $\underset{\hat{x}}{\overset{\shortparallel}{\hat{x}^{(i)}}}$. Details easy. ∎

$$\Rightarrow \quad \Pr[\Delta(t) \in (0,\epsilon]] \le \sum_i \Pr[\Delta^{i}(t) \in (0,\epsilon]]$$

Subclaim: $\Pr[\Delta^{(i)}(t) \in (0,\epsilon]] \le \phi\epsilon$.

Pf: fix wts of all except $i$. Now $x^{*(i)}$ fixed. Also identity of $\hat{x}^{(i)}$ fixed, since all items in that min move together. Chance that its weight falls into $\epsilon$ width interval $\le \phi\epsilon$. ∎

∎

Extensions

- Similarly can imagine profit is random, weight is adversarial.
- Now we did not really use that profit was generated by a modular (additive fn) $\text{profit}(x) = \sum_i p_i x_i$, but just that there was a total order of solutions based on profits, and $\hat{x}$ was defined as the ~~elt~~ sol^n higher in this total ordering than $x^*$. We did use that weights were additive
- So can extend to model with general weights, but random profits (note that we now need a ranking of elements according to "weights", and we do need that profits are additive). This extends to any combinatorial optimization problem with solution set $S \subseteq \{0,1\}^n$, profit function $p(x) = \sum_i p_i x_i$ (with $p_i$'s being $\phi$-smooth).
- Roglin & Teng, Moitra & O'Donnell extend to multicriteria approx.

---

Finally: simplex & smoothed analysis.

- shadow vertex pivot rule. [Gass Saaty] project onto plane defined by dirs $c \leftarrow$ objective, $u \leftarrow$ obj that has opt at start point.
- smoothing model: - given $\bar{A}, \bar{b}, c, \sigma$ we ~~the~~ want to solve

$$\max\ c^T x$$
$$\text{st } (\bar{A}+G)x \le (\bar{b}+g)$$

$G, g$ have indep gaussian entries, mean 0, standard deviation $\sigma \cdot \max_i (\|(\bar{b}_i; \bar{a}_i)\|)$

[Like saying adversary can choose means of the random rows & rhs, but not actual values].

- ST showed: let $\vec{c}, \vec{u}$ be vectors $\in \mathbb{R}^d$ (fixed). $a_1 \ldots a_n \in \mathbb{R}^d$ Gaussian vectors (indep.) with means having norms $\le 1$, standard deviation $\sigma$. $\mathbb{R}^d \supseteq P = \{x \mid a_i^T x \le 1 \ \forall i\}$. Then $P$ projected onto plane$(\vec{u}, \vec{c})$ has at most $\text{poly}(n, d, 1/\sigma)$ vertices.

Problem: $u$ is not indep of the $a_i$'s, since it depends on starting point.
How to find starting point anyways?

[ST04] and [Vershynin] handle these. Also Vershynin shows: $\text{poly}(d, \log n, 1/\sigma)$!