Dual Fitting: Algorithm doesn't use LP, analysis constructs dual solution + bounds gap to ALG

Primal-Dual: Algorithm simultaneously builds primal and dual solutions

Set Cover:

Input:
- universe $\mathcal{U}$
- $\mathcal{S} \subseteq 2^{\mathcal{U}}$
- $c: \mathcal{S} \to \mathbb{R}^+$

Feasible: $\mathcal{I} \subseteq \mathcal{S}$ s.t. $\bigcup_{S \in \mathcal{I}} S = \mathcal{U}$

Objective: $\min \sum_{S \in \mathcal{I}} c(S)$

LP (Primal):

$$\min \sum_{S \in \mathcal{S}} c(S) x_S$$

$$\text{s.t.} \sum_{S \in \mathcal{S} : e \in S} x_S \geq 1 \qquad \forall e \in \mathcal{U}$$

$$x_S \geq 0 \qquad \forall S \in \mathcal{S}$$

Dual: Variable y_e $\forall e \in \mathcal{U}$, constraint for each $S \in \mathcal{S}$

$$\max \sum_{e \in \mathcal{U}} y_e$$

$$\text{s.t.} \sum_{e \in S} y_e \leq c(S) \qquad \forall S \in \mathcal{S}$$

$$y_e \geq 0 \qquad \forall e \in \mathcal{U}$$

Dual fitting: Greedy

- Init $U' = U$, $I = \emptyset$
- While $U' \neq \emptyset$:
  - Let $S = \underset{T \in \mathcal{S}}{\operatorname{argmin}} \dfrac{c(T)}{|T \cap U'|}$
  - $I = I \cup \{S\}$
  - $U' = U' \setminus S$
- return I

Notation: At iteration t, greedy chooses S_t, which covers n_t uncovered elements.

Sps $e \in U$ first covered in iteration t.

$\Rightarrow$ Set $y'_e = \dfrac{c(S_t)}{n_t}$ (split cost of S_t among newly covered elements)

$\Rightarrow$ Set $y_e = \dfrac{y'_e}{H_n}$ $\left(H_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}\right)$

Lemma: y is feasible for dual

Assuming lemma:

Thm: Greedy is an H_n-approximation

Pf: Let I^* optimal solution, $x_S^* = \begin{cases} 1 & \text{if } S \in I^* \\ 0 & \text{otherwise} \end{cases}$

$\Rightarrow c(\text{greedy}) = \sum_{S \in I} c(S)$

Def: $A_S = \{e \in S$ first covered by $S\}$
$\Rightarrow c(S) = \sum_{e \in A_S} y_e'$

$= \sum_{e \in U} y_e'$ (def of y_e')

$= H_n \sum_{e \in U} y_e$ (def of y_e)

$\leq H_n \sum_{S \in \mathcal{S}} c(S) x_S^*$ (weak duality: feasible dual $\leq$ feasible primal)

$= H_n \cdot OPT$

Pf that y is dual feasible:

Let $S \in \mathcal{S}$. Want to show that $\sum_{e \in S} y_e \leq c(S)$

(dual constraint)

- a_k = # elements of S uncovered at beginning of iteration k
- $A_k \subseteq S$ elements of S covered in iteration k

$\Rightarrow |A_k| = a_k - a_{k+1}$

Greedy alg picks set minimizing "average cost"

$\Rightarrow$ Could pick S, average cost $\frac{c(S)}{a_k}$

$$\Rightarrow \quad y'_e \le \frac{c(S)}{a_k} \quad \forall e \in A_k$$

Let ℓ = # iterations

$$\Rightarrow \quad \sum_{e \in S} y_e = \frac{1}{H_n} \sum_{e \in S} y'_e$$

$$= \frac{1}{H_n} \sum_{k=1}^{\ell} \sum_{e \in A_k} y'_e$$

$$\le \frac{1}{H_n} \sum_{k=1}^{\ell} \sum_{e \in A_k} \frac{c(S)}{a_k}$$

$$= \frac{1}{H_n} \sum_{k=1}^{\ell} |A_k| \frac{c(S)}{a_k}$$

$$= \frac{1}{H_n} \sum_{k=1}^{\ell} (a_k - a_{k+1}) \frac{c(S)}{a_k}$$

$$= \frac{c(S)}{H_n} \sum_{k=1}^{\ell} \frac{a_k - a_{k+1}}{a_k}$$

$$\le \frac{c(S)}{H_n} \sum_{i=1}^{|S|} \frac{1}{i}$$

$$= \frac{H_{|S|}}{H_n} c(S) \le c(S)$$

Primal-Dual:

General P-D "schema" (for min problem):

1) Write down primal LP relaxation (min), dual LP (max)
2) Start with $x = \vec{0}$ (primal infeasible), $y = \vec{0}$ (dual feasible)
3) Until x is primal feasible:
   - Increase y until some dual constraint becomes tight (maintaining feasibility of y)
   - Select some of the tight dual constraints, increase corresponding primal variables integrally (intuition: complementary slackness)
   - "Freeze" dual variables in tight dual constraints

Analysis: prove $c^T x \leq \alpha \cdot b^T y$ for some α

$\Rightarrow$ since x, y feasible primal/dual solutions,

$\Rightarrow$ $c(x) \leq \alpha \cdot OPT$

P-D for Set Cover:

- $y = \vec{0}$, $I = \emptyset$ $(x = \vec{0})$
- while $\exists e \in U$ with $e \notin \bigcup_{S \in I} S$
  - Increase y_e until $\exists$ some $S \notin I$ with $e \in S$ s.t. $\sum_{e' \in S} y_{e'} = c(S)$

    $\min_{S \notin I : e \in S} \left(c(S) - \sum_{e' \in S} y_{e'} \right)$

  - Add S to I (set $x_S = 1$)

    $\sum_{e \in S} y_e \le c(S)$ (dual constraint)
- return I (x)

Thm: P-D is an f-approximation $\left(f = \max_{e \in U} |\{S \in \mathcal{S} : e \in S\}| \right)$

Pf: y always dual feasible:

induction:
- true at beginning
- If constraint for S becomes tight, all elements in S covered

  $\Rightarrow$ y_e never increased again for any $e \in S$

$$PD = \sum_{S \in I} c(S) = \sum_{S \in I} \sum_{e \in S} y_e \quad \text{(dual constraint tight } \forall S \in I\text{)}$$

$$= \sum_{e \in S} y_e \cdot |\{S \in I : e \in S\}| \quad \text{(switch order of summations)}$$

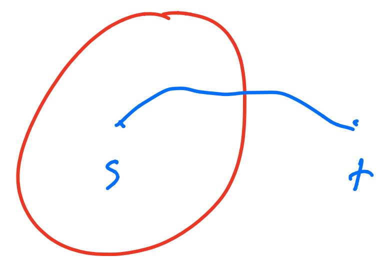
$$\le \sum_{e \in S} f \cdot y_e$$ (def of f)

$$= f \sum_{e \in S} y_e$$

$$\le f \cdot OPT$$ (weak duality)

Shortest s-t path:

- $G=(V,E)$, $c: E \to \mathbb{R}^+$, $s,t \in V$ (can run Dijkstra)
- $\mathcal{S} = \{S \subseteq V : s \in S, t \notin S\}$



Primal LP:

$$\min \sum_{e \in E} c(e)\, x_e$$

$$\text{s.t.} \quad \sum_{e \in \delta(S)} x_e \geq 1 \qquad \forall S \in \mathcal{S}$$

$$x_e \geq 0 \qquad \forall e \in E$$

Dual LP:

$$\max \sum_{S \in \mathcal{S}} y_S$$

$$\text{s.t.} \quad \sum_{S \in \mathcal{S} : e \in \delta(S)} y_S \leq c(e) \qquad \forall e \in E$$

$$y_S \geq 0 \qquad \forall S \in \mathcal{S}$$

PD Alg:

- $y = \vec{0}$, $F = \emptyset$
- while no s-t path in (V, F) {

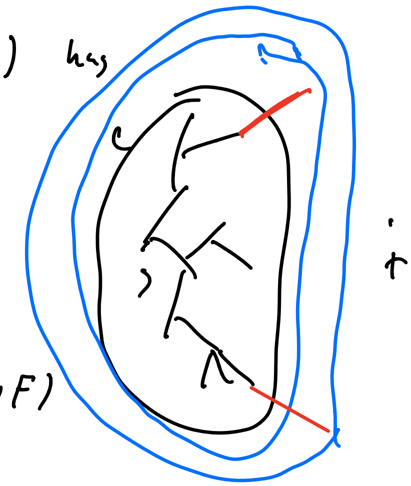
- Let C be connected component of (V, F) containing s
- Increase y_C until some $e \in \delta(C)$ has

$$\sum_{S \in \mathcal{S} : e \in \delta(S)} y_S = c(e)$$

- Add e to F

}

- Return an s-t path P in (V, F)



Fact: this is Dijkstra's Algorithm!

<u>Lemma</u>: Throughout algorithm, F is a tree containing s

<u>Pf</u>: Induction.

Init: $F = \emptyset$. Trivially true

Inductive step:

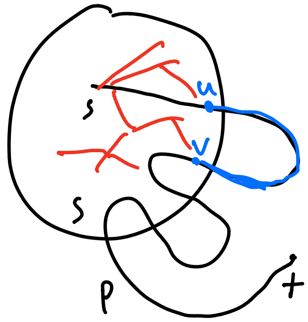
Sps true at some point.

PD adds $e \in \delta(C)$ to F

$\Rightarrow F$ still a tree

<u>Lemma</u>: If $y_S > 0$, then $|P \cap \delta(S)| = 1$

<u>Pf</u>: Sps false: $y_S > 0$ but $|P \cap \delta(S)| > 1$



P': subpath of P with only first and last endpoints in S

when y_S increased, F was a tree spanning S

$\Rightarrow F$ has an s-v path inside S

$\Rightarrow F \cup P'$ has a cycle

$\Rightarrow\Leftarrow$

Thm: PD finds shortest s-t path

Pf:

$$\sum_{e \in P} c(e) = \sum_{e \in P} \sum_{S : e \in \delta(S)} y_S$$ (every edge added has tight dual constraint)

$$= \sum_{S \in \mathcal{S}} \sum_{e \in P \cap \delta(S)} y_S$$ (switch order of summation)

$$= \sum_{S \in \mathcal{S}} y_S \, |P \cap \delta(S)|$$

$$= \sum_{S \in \mathcal{S}} y_S$$ (Lemma)

$$\leq OPT$$ (weak duality)