Set Cover Primal

$$\min \sum_{j=1}^{m} x_j w_j$$
$$\text{s.t.} \sum_{j: e_i \in S_j} x_j \geq 1, \quad i = 1, 2, \ldots, n$$
$$x_j \geq 0, \quad j = 1, 2, \ldots, m$$

Covering problem

Set Cover Dual

$$\max \sum_{i=1}^{n} y_i$$
$$\text{s.t.} \sum_{e_i \in S_j} y_i \leq w_j, \quad j = 1, 2, \ldots, m$$
$$y_i \geq 0, \quad i = 1, 2, \ldots, n$$

Packing problem

The primal problem is a covering problem:
Each element has to be covered by at least one set.

The dual problem can be viewed as a packing problem:
Each set S_j has a capacity of w_j.
We interpret y_i as the weight of e_i, and the total weight of elements in S_j must not exceed w_j.

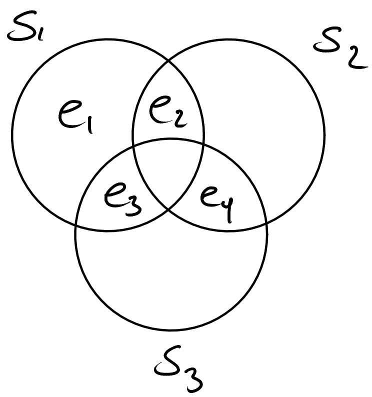
Recall that the dual is constructed such that the value of any solution to the dual is a lower bound on the value of any solution to the primal:

$$Z_{Primal} \geq Z_{Dual} \quad \text{(weak duality property)}$$

In fact,

$$Z^*_{Primal} = Z^*_{Dual} \quad \text{(strong duality property)}$$

Ex:



$w_1 = 1$

$w_2 = 2$

$w_3 = 3$

Primal:

$$\begin{array}{llllll} \min & x_1 & + 2x_2 & + 3x_3 & & \\ \text{s.t.} & x_1 & & & \geq 1 & \\ & x_1 & + x_2 & & \geq 1 & \\ & x_1 & + & x_3 & \geq 1 & \\ & & x_2 & + x_3 & \geq 1 & \\ & x_1, x_2, x_3 \geq 0 & & & & \end{array}$$

OPT = 3:

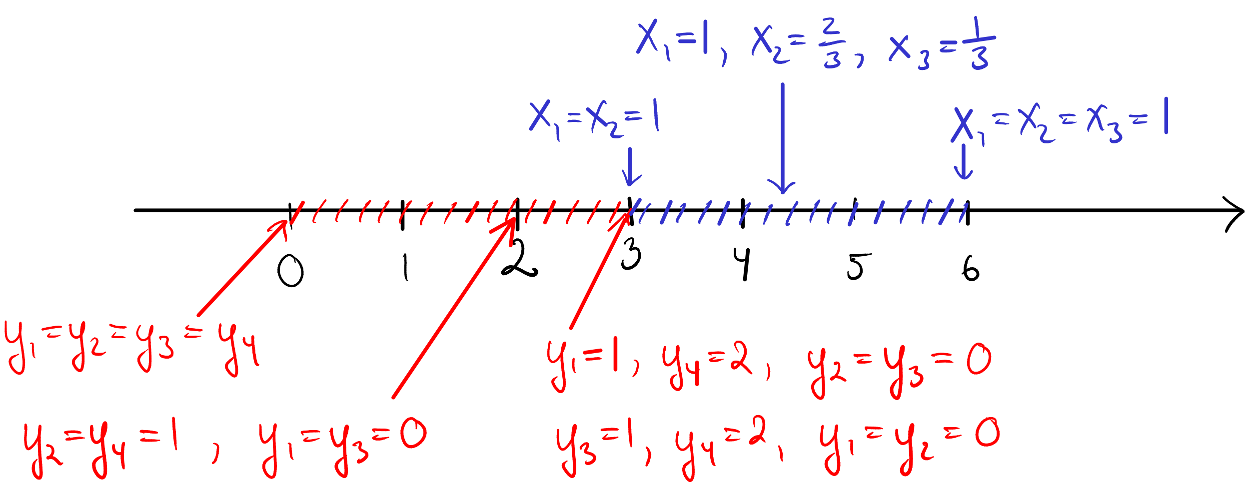
$x_1 = x_2 = 1$

Dual:

$$\begin{array}{llllll} \max & y_1 + y_2 + y_3 + y_4 & \\ \text{s.t.} & y_1 + y_2 + y_3 & \leq 1 \\ & y_2 + y_4 & \leq 2 \\ & y_3 + y_4 & \leq 3 \\ & y_1, y_2, y_3, y_4 \geq 0 & \end{array}$$

OPT = 3:

$y_1 = 1$, $y_4 = 2$ or $y_3 = 1$, $y_4 = 2$



Now back to approximation algorithms for Set Cover.

We have seen an alg. (Alg. 1) which solves the LP relaxation and then does deterministic rounding.

We will now look at an alg. (Alg. 2) which solves the <u>dual</u> of the LP relaxation.
Then each set corresponding to a tight constraint is selected:

Alg. 2 for Set Cover
$\vec{y}^* \leftarrow$ opt. sol. to dual LP
$I' \leftarrow \{ j \mid \sum_{e_i \in S_j} y_i = w_j \}$

Let's see how the alg. would work on the example from before. Recall that there were two optimal solutions to the dual:

$y_1 = 1,\ y_2 = y_3 = 0,\ y_4 = 2$ and $y_1 = y_2 = 0,\ y_3 = 1,\ y_2 = 2.$

In the ex. above:

with $y_1^* = 1$, $y_4^* = 2$, Alg 2 would choose S_1 and S_2 with a total weight of 3.

with $y_3^* = 1$, $y_4^* = 2$, Alg. 2 would choose S_1, S_2, and S_3 with a total weight of 6.

The first solution is optimal, and the latter is a 2-approximation (i.e., an f-approximation).

Alg. 2 is an f-approximation algo. on this example:

If the algo. chooses S_1, S_2, and S_3, the total weight is $W = w_1 + w_2 + w_3$, and

$$w_1 + w_2 + w_3 = (y_1^* + y_2^* + y_3^*) + (y_2^* + y_4^*) + (y_3^* + y_4^*),$$

since the algo. chooses exactly those sets that have LHS = RHS.

Since each y_i is present in at most f constraints,

$$\begin{aligned} W &\le f \cdot (y_1^* + y_2^* + y_3^* + y_4^*) \\ &= f \cdot Z^*_{dual} \\ &\le f \cdot Z^*_{primal}, \text{ by the weak duality property} \\ &= f \cdot OPT \end{aligned}$$

Before giving the general proof that Alg. 2 is an f-approx. alg., we show that it always produces a valid set cover:

Lemma 1.7

Alg. 2 produces a set cover

Proof:

Assume for the sake of contradiction that some element e_k is not covered by $\{S_j \mid j \in I'\}$.

Then $\sum_{e_i \in S_j} y_i < w_j$ for all S_j containing e_k.

These are exactly the constraints involving y_k.

Thus, none of the constraints involving y_k are tight.

This means that y_k can be increased without violating any constraint.

Since this will increase the value $\sum_{i=1}^{n} y_i$ of the sol., we conclude that

the solution $\vec{y}$ was not optimal. $\square$

Ex:

In the ex. above, assume

$$y_1 = y_2 = y_3 = 0$$
$$y_4 = 2$$

Then, only the second constraint is tight, so only S_2 is picked:

$$\begin{aligned} y_1 + y_2 + y_3 \quad &= 0 < 1 \\ y_2 \quad + y_4 &= 2 \\ y_3 + y_4 &= 2 < 3 \end{aligned}$$

e_4 is not covered, since none of the two constraints involving y_4 are tight.

We can increase y_4 from 0 to 1 without violating any constraints
(Then two other constrainsts become tight.)
This increases the sol. value from 2 to 3.
Thus, the sol. above was not optimal.

Or we could increase y_1 from 0 to 1.
Then only the first constraint becomes tight, resulting in an optimal solution.

This illustrates the idea of the primal-dual alg of Section 1.5.

We now give a more formal proof that Alg 2 is an f-approximation algo.

Thm 1.8

Alg. 2 is an f-approx. algo.

<u>Proof</u>:

The correctness follows from Lemma 1.7.

Approx. guarantee:

$$\sum_{j\in I'} w_j = \overbrace{\sum_{j\in I'} \sum_{e_i\in S_j}}^{y_i^* \text{ appears once for each set in the sol. containing } e_i} y_i^*$$

$$= \sum_{i=1}^{n} \underbrace{|\{ j\in I' \mid e_i \in S_j\}|}_{\#\text{sets in the sol. containing } e_i} \cdot y_i^*$$

$$\leq \sum_{i=1}^{n} \underbrace{f_{e_i}}_{\#\text{sets containing } e_i} \cdot y_i^*$$

$$\leq \sum_{i=1}^{n} f \cdot y_i^*$$

$$= f \cdot Z^*_{dual}$$

$$\leq f \cdot Z^*_{primal}, \quad \text{by the weak duality property}$$

$$\leq f \cdot OPT$$

□

Note that for proving the above theorem, we could also use the relaxed c.s.c. (with $b=1$, $c=f$), since

- $\sum_{j: e_i \in S_j} x_j \le f$, for all $i = 1, 2, \ldots, n$, by the def. of f.
- $x_j = 1 \Rightarrow \sum_{e_i \in S_j} y_i = w_j$, by the def. of the alg.

Both Alg. 1 and Alg. 2 rely on solving an LP (optimally). In Section 1.5, we will study a more time efficient alg.

The key observation is that in the proof of Thm 1.8, we did not need the fact that $\vec{y}^*$ is optimal, since $Z_{dual} \le Z^*_{primal}$, for any feasible dual solution.

Thus, the crux is to obtain an index set I'' s.t.

- $\{S_j \mid j \in I''\}$ is a set cover
- $\sum_{j \in I''} w_j = \sum_{j \in I''} \sum_{e_i \in S_j} y_i$, for some feasible sol. $\vec{y}$ to the dual LP

without solving an LP optimally.

Section 1.5 : A Primal-Dual Alg. for Set Cover

Alg. 1.1 for Set Cover: Primal-Dual

$I'' \leftarrow \emptyset$

$\vec{y} \leftarrow \vec{0}$

While $\exists e_k \notin \bigcup_{j \in I''} S_j$

  Increase y_k until some constraint, ℓ, becomes tight, i.e., $\sum_{e_i \in S_\ell} y_i = w_\ell$

  $I'' \leftarrow I'' \cup \{\ell\}$

Note that $e_k \in S_\ell$

Thm 1.9

Alg. 1.1 is an f-approx. alg. for Set Cover

Proof:

Alg. 3 produces a set cover, since as long as some element is not covered, the corresponding dual constraints are non-tight.

The approx. guarantee follows from the same calculations as in the proof of Thm. 1.8, since

$$\sum_{j \in I''} w_j = \sum_{j \in I''} \sum_{e_i \in S_j} y_i \leq f \cdot Z_{dual} \leq f \cdot Z^*_{dual}$$

$\square$

In contrast to Alg. 2 from Section 1.4, Alg. 1.1 does not necessarily produce an optimal dual solution:

In the example above, it might do the following.

$y_2 \leftarrow 1$ (S_1 is picked, e_4 still uncovered)

$y_4 \leftarrow 1$ (S_2 is picked)

(This is fine, since the proof of Thm. 1.8 does not use that $\sum y_i = OPT$, only that $\sum y_i \leq OPT$, which is true for <u>any</u> feasible sol. to the dual.)