Exercise 5.7:

Derandomize the rounding alg. from Section 1.7, using the method of conditional expectations.

Hint: Use the following obj. fct. with random variables X_j, $1 \le j \le m$, and Z.

$$f = \sum_{j=1}^{m} \underbrace{X_j}_{\begin{cases} 1, & \text{if } S_j \text{ incl.} \\ 0, & \text{otherwise} \end{cases}} w_j + \underbrace{\lambda}_{n \cdot \ln n \cdot Z^*_{LP}} \underbrace{Z}_{\begin{cases} 0, & \text{if set cover} \\ 1, & \text{otherwise} \end{cases}}$$

With this obj. fct.,

any infeasible sol. has $f \geq \lambda = n \cdot \ln n \cdot Z^*_{LP}$ $\quad (*)$

For $AlgRR_2$,

$$\begin{aligned} E[f] &= E\Big[\sum_{j=1}^{m} X_j w_j\Big] + \lambda E[Z], \quad \text{by lin. of exp.} \\ &\leq 2 \cdot \ln n \cdot Z^*_{LP} + \cancel{n} \cdot \ln n \cdot Z^*_{LP} \cdot \cancel{n}^{-1}, \quad \text{by the analysis in Sec. 1.7} \\ &= 3 \cdot \ln n \cdot Z^*_{LP} \end{aligned}$$

Thus, using the method of cond. exp., we can find a sol with $f \leq E[f] \leq 3 \cdot \ln n \cdot Z^*_{LP}$, and by $(*)$, such a sol is a set cover.

To derandomize the alg. we must be able to calculate conditional exp values, i.e., calculate $E[f]$, given that decisions about $S_1, \dots, S_\ell$ have already been made:

$$E[f \mid \vec{X}_\ell] = \sum_{j=1}^{\ell} X_j w_j + \sum_{j=\ell+1}^{m} x_j w + \lambda E[Z \mid \vec{X}_\ell]$$

where $\vec{X}_\ell = (X_1, X_2, \dots, X_\ell)$, and $E[Z \mid \vec{X}_\ell]$ can be calculated in the following way.

For each element e_i

$$Pr[e_i \text{ covered} \mid \vec{X}_\ell]$$

$$= \begin{cases} 1, & \text{if } e_i \text{ is contained in a set } S_j \text{ s.t. } j \le \ell \text{ and } X_j = 1 \text{ (i.e., } e_i \text{ is covered by one of the sets } S_1, \dots, S_\ell) \\ 1 - \underbrace{\prod_{\substack{j:\, e_i \in S_j \\ \wedge\, j > \ell}} (1 - x_j)}_{\text{prob. that } e_i \text{ will } \underline{\text{not}} \text{ be covered by any of the sets } S_{\ell+1}, \dots, S_m}, & \text{otherwise} \end{cases}$$

$$E[Z \mid \vec{X}_\ell] = 1 - \prod_{i=1}^{n} Pr[e_i \text{ covered} \mid \vec{X}_\ell]$$

DeRR$_2$

Solve LP optimally

For $\ell \leftarrow 1$ to m

If $E[f \mid (x_1, x_2, \dots, x_{\ell-1}, 0)] \leq E[f \mid (x_1, x_2, \dots, x_{\ell-1}, 1)]$

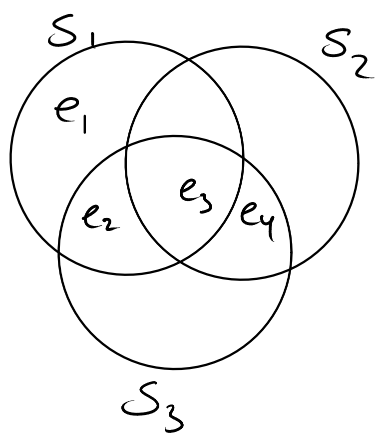
$x_\ell \leftarrow 0$

Else

$x_\ell \leftarrow 1$

$\boxed{\text{Greedy recap.}}$

Ex:



	Price per element 1. it.	2. it.
$w_1 = 12$	4	6
$w_2 = 4$	2	–
$w_3 = 9$	3	9

1. Pick S_2 $\rightarrow$ $\text{price}(e_3) = \text{price}(e_4) = 2$
2. Pick S_1 $\rightarrow$ $\text{price}(e_1) = \text{price}(e_2) = 6$

Total weight

$$
\begin{aligned}
&= w_2 + w_1 \\
&= (\text{price}(e_3) + \text{price}(e_4)) + (\text{price}(e_1) + \text{price}(e_2)) \\
&= (2+2) + (6+6) \\
&= 16
\end{aligned}
$$

Let $q = \max\{\, |S_i| \mid S_i \in \mathcal{S} \,\}$.

Thm 1.12

Alg. 1.2 is an H_q-approx. alg. for Set Cover

<u>Proof</u>: By Dual Fitting:

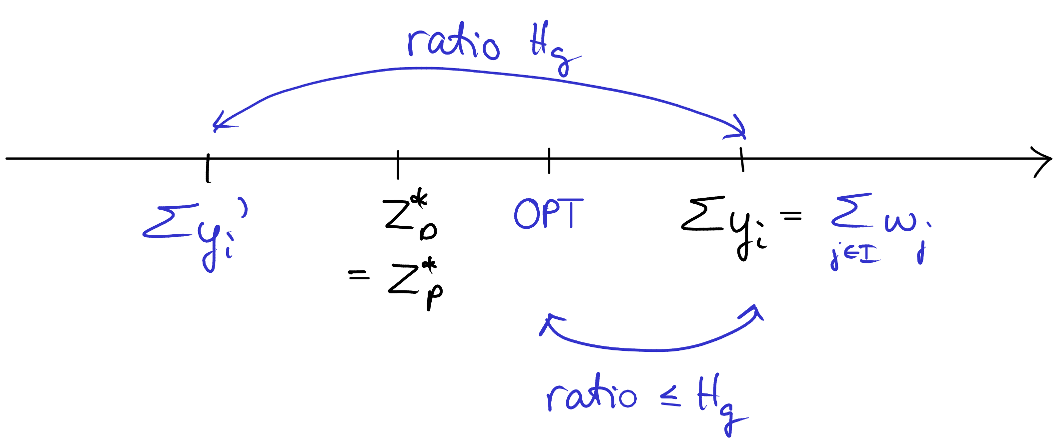
Consider the dual D of the LP for Set Cover. We will construct an infeasible solution $\vec{y}$ and a feasible solution $\vec{y'}$ such that

- $\sum_{i=1}^{n} y_i = \sum_{j \in I} w_j$
- $y_i' = \frac{y_i}{H_q}$, $\quad 1 \le i \le n$

Then,

$$\sum_{j \in I} w_j = \sum_{i=1}^{n} y_i = H_q \sum_{i=1}^{n} y_i' \le H_q Z_D^* \le H_q \cdot OPT,$$

proving the claimed approximation factor.



For $1 \le i \le n$, let $y_i = \text{price}(e_i)$. Then,

$$\sum_{1 \le i \le n} y_i = \sum_{j \in I} w_j$$

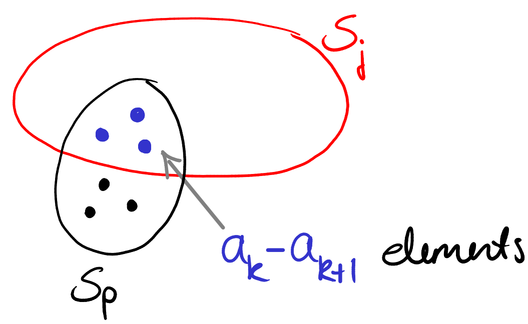
Hence, we just need to show that $\vec{y}$ is feasible:

Consider an arbitrary set S_j.

Let a_k be #uncovered elements in S_j at the beginning of the k'th iteration.

Let S_p be the set chosen by Greedy in the k'th iteration.

S_p covers $a_k - a_{k+1}$ previously uncovered elements in S_j



The price per element in S_j covered in the k'th iteration is at most

$$\frac{w_p}{|\hat{S}_p|} \le \frac{w_j}{a_k}$$

since otherwise S_j would be a more greedy choice. $\lightning$

Thus,

$$\sum_{e_i \in S_j} y_i \leq \underbrace{\sum_{k=1}^{r} (a_k - a_{k+1}) \frac{w_j}{a_k}}^{\text{Total \#terms} = |S_j|, \text{ since } a_1 = |S_j| \text{ and } a_{r+1} = 0}$$

$$\leq w_j \sum_{i=1}^{|S_j|} \frac{1}{i},$$ by the same arguments as in the proof of Thm 1.12.

$$\leq w_j \sum_{i=1}^{g} \frac{1}{i}$$

$$= w_j \cdot H_g$$

Hence,

$$\sum_{e_i \in S_j} y_i' = \frac{1}{H_g} \sum_{e_i \in S_j} y_i \leq w_j$$

□

Compare the proof of Thm 1.12 (dual fitting) to the proof of Thm 1.11:

- Simpler: Compare prices to w_j instead of OPT
- Stronger result: H_g instead of H_n (could also have been obtained with the technique of the proof of Thm 1.11)

Ex from before:

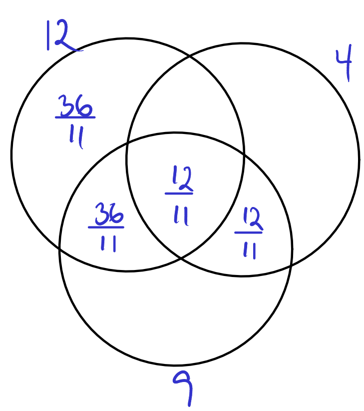
$$y_3 = y_4 = 2$$
$$y_1 = y_2 = 6$$

$$H_3 = 1 + \frac{1}{2} + \frac{1}{3} = \frac{11}{6}$$

$$y_3' = y_4' = \frac{1}{H_3} \cdot 2 = \frac{6}{11} \cdot 2 = \frac{12}{11}$$
$$y_1' = y_2' = \frac{6}{11} \cdot 6 = \frac{36}{11}$$

$\vec{y'}$ is feasible:

$$y_1' + y_2' + y_3' = 2 \cdot \frac{36}{11} + \frac{12}{11} < 8 \le w_1$$
$$y_3' + y_4' = 2 \cdot \frac{12}{11} < 3 \le w_2$$
$$y_2' + y_3' + y_4' = \frac{36}{11} + 2 \cdot \frac{12}{11} < 6 \le w_3$$



Is the upper bound of H_n tight?

If it is, the matching lower bound must come from an instance with

- one set containing all elements
  (follows from the upper bound of H_g)
- only one additional element covered in each it.
  (otherwise, some of the terms in $\frac{1}{n} + \frac{1}{n-1} + \dots + 1$ would be replaced by smaller terms.)

Ex: