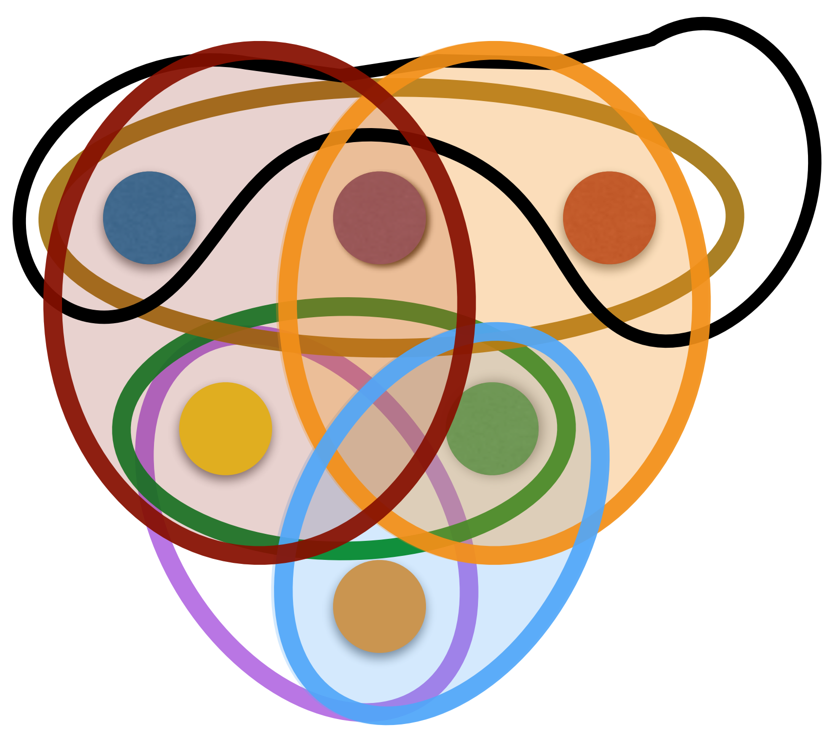
Set cover, linear programming and randomized rounding



Sample-and-iterate algorithm

Repeat

Choose S w.pr. $x_S / \sum_{S'} x(S')$

Put S in cover

(if not there yet)

Until you have a set cover

Repeat

Choose S w.pr. $x_S / \sum_{S'} x(S')$

Put S in cover

(if not there yet)

Until you have a set cover

Analysis

1. Expected cost of output

T = #iterations (stopping time)

C_t = cost of set chosen at iteration t

Cost of output: $\sum_{t=1}^{T} C_t$

Repeat

Choose S w.pr. $x_S / \sum_{S'} x(S')$

Put S in cover

(if not there yet)

Until you have a set cover

Cost of output: $\sum_{t=1}^{T} C_t$

Expected cost of output: $E[\sum_{t=1}^{T} C_t]$

NB: cannot exchange E[] and summation here!

Wald's equation

T stopping time

$\mathrm{X_t}$ random variable

If $\mathrm{X_t}$ bounded from above:

$$\mathbf{E}[\mathbf{X_t}] \leq \mu$$

then $\mathbf{E}[\sum_{\mathbf{t} \leq \mathbf{T}} \mathbf{X_t}] \leq \mu \mathbf{E}[\mathbf{T}]$

NB: can now exchange E[] and summation!

Expected cost of set chosen at iteration t:

$$\mathbf{E}[\mathbf{C_t}] = \sum_{\mathbf{S}} \mathrm{Pr}[\mathbf{S\ chosen}]\mathbf{c_S} = \frac{\sum_{\mathbf{S}} \mathbf{c_S x_S}}{\sum_{\mathbf{S}} \mathbf{x_S}} = \mu$$

Expected cost of output: $\mathbf{E}[\mathbf{T}]\mu = \mathbf{E}[\mathbf{T}]\frac{\sum_{\mathbf{S}} \mathbf{c_S x_S}}{\sum_{\mathbf{S}} \mathbf{x_S}}$

Next problem: What is E[T]?

Repeat

Choose S w.pr. $x_S / \sum_{S'} x(S')$

Put S in cover

(if not there yet)

Until you have a set cover

More notations

2. Expected number of iterations

n_t = #elts not yet covered after t iterations

$$n_0 = n, n_T = 0, n_{T-1} \geq 1$$

Wald's equation for dependent decrements

Consider a random decreasing sequence

$$n_0, n_1, \ldots, n_T,$$

where T is a stopping time. If

$$E[n_t - n_{t+1} | n_t] \geq f(n_t),$$

where f is an increasing function, then

$$E[T] \leq 1 + E[\int_{n_{T-1}}^{n_0} \frac{1}{f(z)dz}].$$

To bound E[T], analyze

#elts not yet covered after t iterations

Analyze $\mathbf{n_t}$

Analyze $\mathbf{E[n_t - n_{t+1} | n_t]}$: elts covered in next iteration

Fix an element $\mathbf{e}$ among the $\mathbf{n_t}$

$$\mathbf{\Pr[e\ covered\ in\ next\ iteration] =}$$
$$\mathbf{\Pr[S\ chosen\ contains\ e] =}$$
$$\mathbf{\textstyle\sum_{S:e\in S} x_S / \sum_{S'} x_{S'} \geq}$$
$$\mathbf{1/\textstyle\sum_{S'} x_{S'}}$$

Sum over $\mathbf{e}$

$$\mathbf{E[n_t - n_{t+1} | n_t] \geq f(n_t)}$$

$$\mathbf{E}[\mathbf{n_t} - \mathbf{n_{t+1}} | \mathbf{n_t}] \geq \mathbf{n_t} / \sum_{\mathbf{S'}} \mathbf{x_{S'}}$$

Wald's equation for dependent decrements

Consider a random decreasing sequence

$$\mathbf{n_0}, \mathbf{n_1}, \ldots, \mathbf{n_T},$$

where T is a stopping time. If

$$\mathbf{E}[\mathbf{n_t} - \mathbf{n_{t+1}} | \mathbf{n_t}] \geq \mathbf{f}(\mathbf{n_t}),$$

where f is an increasing function, then

$$\mathbf{E}[\mathbf{T}] \leq \mathbf{1} + \mathbf{E}[\int_{\mathbf{n_{T-1}}}^{\mathbf{n_0}} \frac{\mathbf{1}}{\mathbf{f(z)dz}}].$$

Application:

Stopping time $\mathbf{E}[\mathbf{T}] \leq \mathbf{1} + \int_{\mathbf{1}}^{\mathbf{n}} \frac{\sum_{\mathbf{S}} \mathbf{x_S}}{\mathbf{z}} \mathbf{dz} = \mathbf{1} + \sum_{\mathbf{S}} \mathbf{x_S} \ln(\mathbf{n})$

Together

Expected cost of output:

$$\left(1 + \left(\textstyle\sum_{S} x_S\right)\ln(n)\right)\frac{\sum_S c_S x_S}{\sum_S x_S} \leq$$
$$(1+\ln(n))\ \mathrm{OPT}$$

Result:

the algorithm gives
a collection of sets that is a set cover
with average cost at most (1+ln(n)) OPT.

Set cover, linear programming and randomized rounding