CS526 hw3 P1

I have the LP files online. Payoff matrix:

$r(t,d)$		$d=1$ q_1	$d=2$ q_2	$d=3$ q_3	$d=4$ q_4
		MAX			
MIN "you"	$t=1$ p_1	3	$3/2$	1	1
	$t=2$ p_2	1	2	$4/3$	$4/3$
	$t=3$ p_3	1	1	$5/3$	$5/3$
	$t=4$ p_4	1	1	1	2

(a) An optimal MIN strategy is

$p_1 = 4/19 \simeq 0.21053$,

$p_2 = 6/19 \simeq 0.31579$,

$p_3 = 9/19 \simeq 0.47368$,

$p_4 = 0$

with payoff $P = 27/19 \simeq 1.42105$

(b) An optimal MAX strategy is

$q_1 = 3/19 \simeq 0.15789$,

$q_2 = 4/19 \simeq 0.21053$,

$q_3 = 4/19 \simeq 0.21053$,

$q_4 = 8/19 \simeq 0.42105$

with payoff $Q = 27/19 \simeq 1.42105$

The two LP's (to be used in a solver):

(a) Minimize P

Subject to:

$$p_1 + p_2 + p_3 + p_4 = 1$$
$$P - 3p_1 - 1p_2 - 1p_3 - 1p_4 \geq 0$$
$$P - \tfrac{3}{2}p_1 - 2p_2 - 1p_3 - 1p_4 \geq 0$$
$$P - 1p_1 - \tfrac{4}{3}p_2 - \tfrac{5}{3}p_3 - 1p_4 \geq 0$$
$$P - 1p_1 - \tfrac{4}{3}p_2 - \tfrac{5}{3}p_3 - 2p_4 \geq 0$$
$$p_1, p_2, p_3, p_4 \geq 0$$

(b) Maximize Q

Subject to:

$$q_1 + q_2 + q_3 + q_4 = 1$$
$$Q - 3q_1 - \tfrac{3}{2}q_2 - 1q_3 - 1q_4 \leq 0$$
$$Q - 1q_1 - 2q_2 - \tfrac{4}{3}q_3 - \tfrac{4}{3}q_4 \leq 0$$
$$Q - 1q_1 - 1q_2 - \tfrac{5}{3}q_3 - \tfrac{5}{3}q_4 \leq 0$$
$$Q - 1q_1 - 1q_2 - 1q_3 - 2q_4 \leq 0$$
$$q_1, q_2, q_3, q_4 \geq 0$$

CS526 hw3 P2

Our cover C is the union of D random samples, so $E[\text{cost}(C)] = D \cdot OPT_f$.

C is good enough if we avoid these two bad events:

BAD1: some point is not covered

BAD2: $\text{cost}(C) > (1+\varepsilon) \cdot \ln n \cdot OPT_f$

We choose $D = \ln\left(\frac{4n}{\varepsilon}\right)$, so

$Pr[BAD_1] \le n \cdot (1/e)^D \le \varepsilon/4$.

We suppose n is large enough so $\frac{4}{\varepsilon} \ln \frac{4}{\varepsilon} \le \ln n$, then $D = \ln\left(\frac{4}{\varepsilon}\right) + \ln n \le \left(1 + \frac{\varepsilon}{4}\right) \ln n$, and

$E[\text{cost}(C)] \le \left(1 + \frac{\varepsilon}{4}\right) \ln n \cdot OPT_f$

By Markov's inequality,

$Pr\left[\text{cost}(C) > \left(1+\frac{\varepsilon}{2}\right) E[\text{cost}(C)]\right] < \frac{1}{1+\frac{\varepsilon}{2}} < 1 - \frac{\varepsilon}{3}$. ☆

Since $\left(1+\frac{\varepsilon}{2}\right)\left(1+\frac{\varepsilon}{4}\right) < (1+\varepsilon)$, ☆

$Pr[BAD2] <$ previous $< 1 - \frac{\varepsilon}{3}$.

$\therefore$ $Pr[BAD_1 \vee BAD_2] < 1 - \frac{\varepsilon}{3} + \frac{\varepsilon}{4} = 1 - \frac{\varepsilon}{12}$,

$Pr[C \text{ good enough}] > \varepsilon/12$.

So if we repeat the entire construction of C $12/\varepsilon$ times, then:

$Pr[\text{we see a good enough } C] \ge 1 - \left(1 - \frac{\varepsilon}{12}\right)^{12/\varepsilon} \approx 1 - e^{-1} > 1/2$.

Note we can easily test whether C is good enough in poly time, so total time is $\text{poly}(n, m, 1/\varepsilon)$.

☆ True for sufficiently small $\varepsilon > 0$. Think about Taylor series!

CS526 hw3 P3

<u>Notation</u>: Let $U = \{1, 2, 3, \ldots, m\}$. For $I \subseteq U$, let $S(I) = \bigcup_{i \in I} S_i$, so $f(I) = |S(I)|$.

<u>(a) Show</u> f is submodular on U.

For $X, Y \subseteq U$, we want to show $f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y)$.

First notice $S(X \cup Y) = S(X) \cup S(Y)$, and $S(X \cap Y) \subseteq S(X) \cap S(Y)$.

Then $f(X) + f(Y) = |S(X)| + |S(Y)| = |S(X) \cup S(Y)| + |S(X) \cap S(Y)|$

$\geq |S(X \cup Y)| + |S(X \cap Y)| = f(X \cup Y) + f(X \cap Y)$.

(b)

$I_0 = \phi$

for $j = 1$ to k:

  Pick $x_j \in U$ maximizing $f(I_{j-1} \cup \{x_j\})$

  $I_j = I_{j-1} \cup \{x_j\}$

return $I_k = \{x_1, x_2, \cdots, x_k\}$

(c) For each j we call f $\leq m$ times, so $O(m \cdot k)$ calls.

(d) Choose $I^* \subseteq U$ so $f(I^*) = OPT_k$ and $|I^*| = k$.

Let $d_j = f(I^*) - f(I_j)$ for $0 \leq j \leq k$. Note $d_0 = OPT_k$.

For $1 \leq j \leq k$, we claim $d_j \leq (1 - \frac{1}{k}) d_{j-1}$. Why?

proof: Let $T = S(I^*) - S(I_{j-1})$, note $|T| \geq d_{j-1}$. Some $y \in I^*$ has $|S_y \cap T| \geq \frac{1}{k}|T|$, and x_k adds at least as much to $f(I_k)$:

$f(I_k) \geq f(I_{k-1}) + \frac{1}{k}|T|$. So

$d_j = f(I^*) - f(I_k) \leq f(I^*) - f(I_{k-1}) - \frac{1}{k}|T| = (1 - \frac{1}{k}) d_{j-1}$. ☒

This implies $d_k \leq (1 - \frac{1}{k})^k d_0$, or

$f(I^*) - f(I_k) \leq (1 - \frac{1}{k})^k f(I^*)$, or

$f(I_k) \geq \left[1 - (1 - \frac{1}{k})^k\right] OPT_k$. ✓

CS526 hw3 P4, "Hollywood"

(a) <u>Variables</u>:

For each actor i, $1 \le i \le n$, $x_i \in \{0,1\}$ indicates whether we hire actor i.

For each investor j, $1 \le j \le m$, $y_j \in \{0,1\}$ indicates whether we get investor j.

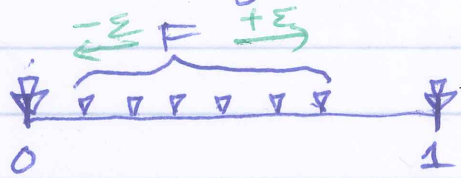
IP: maximize $\sum_j p_j y_j - \sum_i s_i x_i$

subject to: $x_i \ge y_j$ (for all $i \in L_j$, all j)

$x_i, y_j \in \{0,1\}$. ☆

(b) Relax the integer constraints to $0 \le x_i \le 1$, $0 \le y_j \le 1$. Let $z = (x, y)$ be an optimal solution to this LP, with value OPT.

Supposing z is not entirely integer-valued, we argue how to replace it with another optimal z', with <u>more</u> integer values. Repeat until all integer.



Let F be a 0/1 vector indicating which variables in z have <u>fractional</u> values (not 0 or 1), so $z+\varepsilon F$ ($z-\varepsilon F$) is the result of increasing (decreasing) all these by ε.

by our constraints

For small enough ε, note $z+\varepsilon F$ and $z-\varepsilon F$ are both <u>feasible</u>. Since z is their average, they are <u>optimal</u>. Now let $z' = z + \varepsilon F$, where ε is as large as possible for z' to be feasible (some var reached 1). z' is optimal, with fewer fractional values.

CS526 hw3 P5

(b) We build a "fattest path tree" from s to all other reachable vertices.
We need a "Max PQ" data structure.

Binary heap or Fibonacci heap

Initially each vertex v has "fatness" $f[v]=0$ and "parent" $p[v] = \text{NIL}$.
Set $f[s] = \infty$ and $p[s]=s$ (root).
Put all vertices in the PQ, ordered by $f[v]$.

Variant of Dijkstra! Same performance.

```
T = ∅
while PQ not empty:
    Extract the v from PQ
    with maximum f[v].
    IF f[v]==0, break! (all done)
    Add v to tree T, with parent p[v].
    For each edge v→w with capacity c:
        fw = min(c, f[v])
        if fw > f[w]:
            f[w] = fw    //notify PQ
            p[w] = v
```

IF t is in T, then we find the fattest s-t path by tracing back the parent pointers (from t to s).

CS526 hw3 P5 (continued)

(c) Suppose f is an st-flow in G, $val(f) > 0$. Then we can find a simple path P from s to t, using edges where $f(e) > 0$. Let f_P be the "path flow" of value $\min_{e \in P} f(e)$. Then $0 \leq f_P \leq f$, and $f' = f - f_P$ is another st-flow, using a proper subset of the edges used by f.

Repeat until $val(f') = 0$ (a circulation), we get k paths, $val(f) = val(f_{P_1}) + val(f_{P_2}) + \cdots val(f_{P_k})$. Also $k \leq |E|$, since we lose at least one edge per step. So there is some path flow f_{P_i} with $0 \leq f_{P_i} \leq f$ and $val(f_{P_i}) \geq \frac{1}{|E|} val(f)$.

(d) Let f^* be optimal, $val(f^*) = F$. Let f be current Ford-Fulkerson flow (initially $f = 0$) and $f' \equiv f^* - f$ is an optimal flow in the residual graph G_f. By (c), the "fattest path flow" f_P has value $val(f_P) \geq \frac{1}{|E|} val(f') = \frac{1}{|E|}(F - val(f))$.

By induction (and the initial condition $f = 0$) we can show that after t fattest-path iterations,

$$0 \leq (F - val(f)) \leq \left(1 - \frac{1}{|E|}\right)^t \cdot F.$$

Choosing $t = |E| \ln F$, we get

$$0 \leq (F - val(f)) < e^{-\ln F} \cdot F = 1.$$

But this quantity is an integer, so must be 0. $\therefore$ the algorithm terminates after at most $|E| \ln F$ iterations.

Using: $(1 - x) < e^{-x}$ for $x \neq 0$.