TD2 Approximation Algorithms

Wed. Oct. 3, 2012 - **Due on Tue. Oct. 9, 2012**

You are asked to complete the exercises and send me your solutions by email at:
`nicolas.schabanel@liafa.univ-paris-diderot.fr`
(or drop it in my mail box on the 7th floor) on or before **Tuesday Oct. 9**.

■ **Exercise 1 (Duality in linear programming).** We say that a linear program is in primal *normal form* if it is a minimization problem, all the variable are non-negative, all constraints are $\geqslant$.

► ***Question 1.1)*** *Describe an algorithm that convert any linear program in a primal normal form with an increase at most linear in the number of variables and constraints.*

► ***Question 1.2)*** *Give the dual of the following linear program directly without converting it into form normal. Compute the dual of dual and check you get back to the original program.*

$$\begin{cases} \text{Minimize} & 3x_1 + 7x_2 - 2x_3 + 4x_4 \\ \textit{such that} & 2x_1 + 3x_2 \geqslant 10 \\ & x_2 + 5x_3 + x_4 \leqslant 20 \\ & x_1 - 2x_4 = 2 \\ & x_1 \geqslant 0, x_2 \geqslant 0, x_4 \leqslant 0 \end{cases}$$

► ***Question 1.3)*** *Explicit the complementary slackness conditions for this problem and use them to show that the solution $x_1 = 0, x_2 = \frac{10}{3}, x_3 = \frac{53}{15}, x_4 = -1$ is optimal with value $\frac{184}{15}$. What is the corresponding dual optimal solution?*

■ **Exercise 2 (Randomized rounding for Set-Cover).** Consider the following problem (Set-Cover): Given an universe $\mathcal{U} = \{u_1, \ldots, u_m\}$ of m elements, a collection $\mathcal{S} = \{S_1, \ldots, S_n\}$ of n subsets of $\mathcal{U}$ such that $\bigcup_{i=1}^n S_i = \mathcal{U}$, and positive weights (w_i) on each S_i, find a subcollection $C \subset \mathcal{S}$ covering $\mathcal{U}$ (i.e. such that $\bigcup_{S_i \in C} S_i = \mathcal{U}$) with minimum weight $w(C) = \sum_{S_i \in C} w_i$.

We consider the following linear relaxation for Set-Cover:

$$\begin{cases} \text{Minimize} & \sum_{i=1}^n w_i x_i \\ \text{such that} & (1) \displaystyle\sum_{i:u_j \in S_i} x_i \geqslant 1 \quad (\forall j \in \{1, \ldots, m\}) \\ & (2)\ 1 \geqslant x_i \geqslant 0 \qquad (\forall i \in \{1, \ldots, n\}) \end{cases}$$

► ***Question 2.1)*** *Show that this linear program (which can be solved in polynomial time in n and m) is a relaxation of Set-Cover. Show that the constraints $1 \geqslant x_i$ can be waived (i.e. that removing them does not change the optimal solution of the linear program). What is the trivial relationship between* OPT$_{LP}$ *(the optimal value of the linear program) and* OPT *(the weight of an optimal set cover)?*

Now assume that we have obtained an optimal fractional solution (x_i^*) to the linear program. We now want to transform it into a proper set cover. A natural approach is to interpret the x_i^* as the probability that S_i belongs to the aimed optimal subcollection. Let us denote by C a random subcollection obtained by putting each S_i in C independently with probability x_i^*.

▶ ***Question 2.2)*** *What is the expected weight of C?*

▶ ***Question 2.3)*** *Fix some $u_j \in \mathcal{U}$. Show that the probability that u_j is not covered by C (i.e. that $u_j \notin \bigcup_{S_i \in C} S_i$) is at most $1/e$.*
▷ Hint. *Use that $\prod_k (1 - a_k) \leqslant e^{-\sum_k a_k}$ for all $a_k \leqslant 1$.*

We now consider the following algorithm:

Algorithm 1 Randomized rounding for Set-Cover

1: Let x_i^* be an optimal solution to the linear program. Let $\ell := 0$.
2: **repeat**
3: Set $\ell := \ell + 1$ and Draw a random collection C_ℓ by selecting each S_i in C_ℓ independently with probability x_i^*.
4: **until** $\displaystyle\bigcup_{k=1}^{\ell} C_k$ covers $\mathcal{U}$.
5: Set $T := \ell$ and output the collection $\Gamma := \displaystyle\bigcup_{k=1}^{T} C_k$.

Note that the variable T (line 5) stands for the random variable for the number of random collections successively computed by the algorithm. Since all the C_k's are independently and identically distributed as C in question 2.2, and since T only depends on $C_1, \ldots, C_T$, Wald's equation (admitted) ensures that

$$\mathbb{E}[w(\Gamma)] \leqslant \mathbb{E}[w(C_1) + \cdots + w(C_T)] = \mathbb{E}[T] \cdot \mathbb{E}[C].$$

We are thus now left with estimating $\mathbb{E}[T]$.

▶ ***Question 2.4)*** *Show that: for all k,* Pr$\big\{\exists j, u_j$ *is not covered by* $C_1 \cup \cdots \cup C_k\big\} \leqslant \dfrac{m}{e^k}$.

▶ ***Question 2.5)*** *Show that:* Pr$\big\{T > \lceil \ln m \rceil + t\big\} \leqslant e^{-t}$*, for all $t \geqslant 0$. Conclude that:* $\mathbb{E}[T] \leqslant \ln m + O(1)$.
▷ Hint. *Use that $\mathbb{E}[X] = \sum_{n \geqslant 0}$ Pr$\{X > n\}$ for every non-negative integer-valued random variable X. And decompose $\mathbb{E}[T]$ according to the events $\{T \leqslant \ln m\}$ and $\{T > \ln m\}$.*

▶ ***Question 2.6)*** *What is the running time of our algorithm (excluding the resolution of the linear program)? Give an upper bound on the expected weight of its output. What is its approximation ratio?*