Toolkit for TCS	**Feb-June 2021**
HW 2: Duality and rounding	
Out: March 16, 2021	**Due:** April 5, 2021

(1) Please take time to write clear and concise solutions. You are STRONGLY *encouraged to submit LaTeXed solutions.*
(2) Collaboration is OK, but please write your answers yourself, and include in your answers the names of EVERYONE *you collaborated with and* ALL *references other than class notes you consulted.*

1. (0 points) *Note: This problem does not carry any points, and you do not need to submit it. However, unless you are comfortable with the (simple) ideas underlying it, the rest of the homework might prove to be difficult.* Throughout this homework, $\mathbb{S}^n$ denotes the set of $n\times$ symmetric matrices with real entries, while $\mathbb{S}^n_+$ denotes the set of symmetric, positive semi-definite matrices with real entries (i.e., symmetric matrices satisfying $v^T Av \geq 0$ for all real vectors v). The inner product $A \cdot B$ of two $n \times n$ matrices with real entries is defined as Trace $\left(AB^T\right)$. Note that $A \cdot (vv^T) = v^T Av$, when v is a vector. Also, for an $n \times n$ matrix A, the notation $A \succeq 0$ means that $A \in \mathbb{S}^n_+$, while the notation $A \preceq 0$ means that $-A \in \mathbb{S}^n_+$.
   - Show that if $A, B \in \mathbb{S}^n_+$, then $A \cdot B \geq 0$.
   - Show that if $A \in \mathbb{S}^n$ is such that $A \cdot X \leq 0$ for all $X \in \mathbb{S}^n_+$, then $A \preceq 0$.

2. (5 points) Fix $A_1, A_2, \ldots, A_k \in \mathbb{S}^n$ and define the set
$$\text{cone}(A_1, A_2, \ldots, A_k) := \{v \in \mathbb{R}^k \mid \text{ there exists } X \in \mathbb{S}^n_+ \text{ such that } v_i = A_i \cdot X \text{ for } 1 \leq i \leq k\}.$$
   - (2 points) Show that $\text{cone}(A_1, A_2, \ldots, A_k)$ is a convex cone with vertex at the origin. It is also closed, and the proof is easy and standard, but you can assume that without proof.
   - (3 points) Show that
$$\text{dual}(\text{cone}(A_1, A_2, \ldots, A_k)) = \{y \in \mathbb{R}^k \mid \sum_{i=1}^{k} y_i A_i \preceq 0\}.$$

3. (5 points) Fix $A_1, A_2, \ldots, A_k \in \mathbb{S}^n$ and $b \in \mathbb{R}^k$ (with entries $b_1, b_2, \ldots, b_k$). Show that *exactly* one of the following two statements is true (which one of them is true depends upon the choice of $A_1, A_2, \ldots A_n$ and b):

   Statement 1 There exists $X \in \mathbb{S}^n_+$ such that
$$A_i \cdot X = b_i, \text{ for } 1 \leq i \leq k.$$
   Statement 2 There exists $y \in \mathbb{R}^k$ such that
$$b^T y > 0, \text{ and}$$
$$\sum_{i=1}^{k} y_i A_i \preceq 0.$$

4. (5 points) Recall the standard form for a semi-definite program:
$$\begin{aligned} \inf \quad & C \cdot X \\ \text{subject to} \quad & A_i \cdot X = b_i, 1 \leq i \leq k \\ & X \succeq 0. \end{aligned}$$
   Here, $C, A_1, A_2, \ldots, A_k$ are given elements of $\mathbb{S}^n$, while b is a vector in $\mathbb{R}^k$. Assume that this data is such that the inf is finite and equal to $p^\star$. Consider now the program
$$\begin{aligned} \sup \quad & \sum_{i=1}^{k} y_i b_i \\ \text{subject to} \quad & \sum_{i=1}^{k} y_i A_i - C \preceq 0, \text{ and} \\ & y_1, y_2, \ldots, y_k \in \mathbb{R}. \end{aligned}$$

Let $d^\star$ denote the optimal value for the above program.

- (2 points) Show that $d^\star \le p^\star$.
- (3 points) Show that $d^\star = p^\star$.

5. (4 points) **[Based on a result of Wigderson]** A simple undirected graph $G = (V, E)$ is said to be *3-colourable* if there exists a colouring of its vertices such that no two adjacent vertices are coloured by the same colour. A 3-colourable graph is also guaranteed to contain large independent sets. In this problem and the next, we will partially analyze algorithms that find large independent sets in G whenever G is 3-colourable. We use the notation $n = |V|$, and assume that $n \ge 5$.

   (a) (0 points) Show that there is an independent set of size at least $|V|/3$ in any 3-colourable graph $G = (V, E)$.

   However, finding this linear-sized independent set seems to require knowledge of the 3-colouring of G. In the rest of this problem, we will show how to extract a $\Omega(\sqrt{n})$-sized independent set given the promise that the graph is 3-colourable, but without knowing the 3-colouring itself.

   (b) (2 points) Show that if G has maximum degree at most d, then G has an independent set of size at least $\frac{n}{d+1}$. (You do not need the 3-colourability of G for this part.)

   (c) (1 point) Show that if G is 3-colourable and has a vertex of degree at least d, then G has an independent set of size at least $d/2$. Argue that it is possible to find such an independent set in polynomial time (without knowing a 3-colouring of G, and given only the promise that such a 3-colouring exists).

   (d) (1 point) Show that if G is 3-colourable then it has an independent set of size at least $\sqrt{n}/2$, and that such an independent set can be found in polynomial time (without knowing a 3-colouring of G, and given only the promise that such a 3-colouring exists).

6. (6 points) **[Based on a result of Karger, Motwani and Madhu Sudan]** Now, we analyze a different algorithm for the above problem, which finds a larger independent set ($\tilde{\Omega}(n^{2/3})$ instead of $\Omega(\sqrt{n})$). This algorithm tries to find finite dimensional vectors $\{w_v \mid v \in V\}$ satisfying the following constraints.

$$\|w_v\| = 1 \qquad \text{for all } v \in V.$$
$$w_u^T w_v = -\frac{1}{2} \qquad \text{for all } \{u, v\} \in E.$$

   (a) (0 points) Show that such a set of vectors exists if G is 3-colourable. How might one try to algorithmically find such vectors (without knowing beforehand whether or not the input graph is 3-colorable)? What upper bound can one guarantee *a priori* on the dimension of the space spanned by the vectors w_v?

   For the subsequent parts we assume that we are given the vectors $w_v \in \mathbb{R}^d$ as above, and we try to extract an independent set out of them. For this we use the following randomized algorithm. We sample $g \in \mathbb{R}^d$ distributed according to $\mathcal{N}(0, I_d)$, and try to put into our independent set all the vertices v whose corresponding vectors w_v have a large enough inner product with g, say if $w_v^T g \ge t$, for some threshold $t > 1$ to be chosen later. We call this (random) subset S_t. Of course, S_t might not be an independent set, since it might contain pairs of neighbors. So after the process is completed, we obtain J_t by discarding from S_t any v if one of its neighbors was also in S_t. Clearly, J_t is an independent set. We now try to lower bound its expected size.

   If $X \sim \mathcal{N}(0, 1)$, let $Q(\alpha) := \mathbf{P}[X \ge \alpha]$. Recall that for $\alpha \ge 1$,

$$\frac{1}{2} \le \frac{Q(\alpha)}{\frac{1}{\alpha\sqrt{2\pi}} \exp(-\alpha^2/2)} \le 1.$$

   (b) (1 point) Show that $\mathbf{P}[v \in S_t] = Q(t)$.

   (c) (3 point) Show that in $\{u, v\}$ is an edge in G, then $\mathbf{P}[u \in S_t \mid v \in S_t] \le Q(t\sqrt{3})$.

   (d) (1 point) Show that $\mathbf{P}[v \in S_t$, and none of its neighbors are in $S_t] \ge Q(t)(1 - nQ(t\sqrt{3}))$, and conclude that $E[|J_t|] \ge nQ(t)(1 - nQ(t\sqrt{3}))$.

(e) (1 point) Show that there is a choice of t (in terms only of n) that guarantees that $E[|J_t|] \geq c \cdot \frac{n^{2/3}}{\log n}$ for some positive constant c. (You can assume that n is at least n_0, for some large enough constant n_0).