Problem 1:  In the matrix completion problem we observe \((M_{i,j})_{(i,j)\in \Omega}\) for some \(m \times n\) rank-\(r\) matrix \(M\) and \(\Omega \subset [m] \times [n]\). We then hope to recover all entries of \(M\) by solving

\[
\begin{align*}
\text{minimize} & \quad \text{rank}(X) \\
\text{s.t.} & \quad \forall (i,j) \in \Omega \quad X_{i,j} = M_{i,j}
\end{align*}
\]

and hope the \(X\) we find equals \(M\). Unfortunately solving the rank minimization problem is NP-hard, so we relax it to

\[
\begin{align*}
\text{minimize} & \quad \|X\|_* \\
\text{s.t.} & \quad \forall (i,j) \in \Omega \quad X_{i,j} = M_{i,j}
\end{align*}
\]

where \(\|X\|_*\) is the nuclear norm of \(X\), i.e. the sum of its singular values. Note the rank is the number of non-zero singular values, and thus this is the matrix analogue of \(\ell_1\) minimization for sparse recovery. In this problem you will show nuclear norm minimization is a semidefinite program, which given known results [1], implies that it can be solved in polynomial time (with some logarithmic dependence on precision).

Before we can show that nuclear norm minimization is a semidefinite program, what is a semidefinite program? It is the problem of finding \(x\) in what follows, given \(c, \{A_i\}_{i=1}^n, B\):

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{s.t.} & \quad A_0 + x_1 A_1 + \ldots + x_n A_n \preceq B
\end{align*}
\]

where \(A_i, B\) are real symmetric matrices of some dimension and \(X \preceq Y\) means \(Y - X\) is positive semidefinite (i.e. \(Y - X\) is symmetric and has all nonnegative eigenvalues).

Now, you will show that in order to solve (2) it suffices to find \(X, Y, Z\) solving

\[
\begin{align*}
\text{minimize} & \quad \text{trace}(Y) + \text{trace}(Z) \\
\text{s.t.} & \quad \forall (i,j) \in \Omega \quad X_{i,j} = M_{i,j} \\
& \quad \begin{pmatrix} Y & X \\ X^T & Z \end{pmatrix} \succeq 0
\end{align*}
\]

(a) (5 points) Show that (4) is indeed a semidefinite program as defined in (3).

Hint: Think of the entries of \(x\) in (3) as being entries in the matrices \(X, Y, Z\).
(b) (5 points) Show that for $X \in \mathbb{R}^{m \times n}$ and $t \in \mathbb{R}$, if $\|X\|_* \leq t$ then there exist matrices $Y \in \mathbb{R}^{m \times m}$ and $Z \in \mathbb{R}^{n \times n}$ such that

$$
\begin{pmatrix}
Y & X \\
X^T & Z
\end{pmatrix} \succeq 0, \quad \text{trace}(Y) + \text{trace}(Z) \leq 2t
$$

**Hint:** Write the SVD $X = U \Sigma V^T$ and define $Y = U \Sigma U^T + \gamma I$ and $Z = V \Sigma V^T + \gamma I$ for some appropriately chosen $\gamma$.

(c) (5 points) Show that if $A, B$ are positive semidefinite, i.e. real and symmetric with nonnegative eigenvalues, then $\text{trace}(AB) \geq 0$.

**Hint:** For $A = U \Sigma U^T$ and a function $f : \mathbb{R}^+ \to \mathbb{R}^+$, define $f(A) = U f(\Sigma) U^T$ where $f(\Sigma)$ replaces each eigenvalue $\sigma$ with $f(\sigma)$. Now write $A = A^{1/2} A^{1/2}$ and use the cyclic property of trace ($\text{trace}(XY) = \text{trace}(YX)$ as long as the dimensions work out).

(d) (5 points) Show the converse of (c). That is, for fixed $X$ if $Y, Z$ exist satisfying the conditions of (c), then $\|X\|_* \leq t$. Conclude that solving (2) and (4) are equivalent.

**Hint:** Write $X = U \Sigma V^T$. Show $\text{trace} \left[ \begin{pmatrix} U U^T & -U V^T \\ -V U^T & V V^T \end{pmatrix} \begin{pmatrix} Y & X \\
X^T & Z \end{pmatrix} \right] \succeq 0$, $\text{trace}(UU^T Y) \leq \text{trace}(Y)$, $\text{trace}(VV^T Z) \leq \text{trace}(Z)$. Combine these with cyclic properties of trace and the fact that $\|X\|_* = \text{trace}(\Sigma)$.

(e) (10 points) **Bonus problem:** Consider the problem of finding $X, Y, Z$ solving

$$
\begin{align*}
\text{minimize} \quad & \text{rank}(\text{diag}(Y, Z)) \\
\text{s.t.} \quad & \forall (i, j) \in \Omega \quad X_{i,j} = M_{i,j} \\
& \begin{pmatrix}
Y & X \\
X^T & Z
\end{pmatrix} \succeq 0
\end{align*}
$$

(5)

where $\text{diag}(Y, Z)$ is the block-diagonal matrix with two blocks: the upper-left block is $Y$, and the bottom-right block is $Z$. Although we do not have efficient algorithms for solving this problem, show that solving (1) and (5) are equivalent. Thus the program (5) can be seen as a relaxation of (4) since $\text{trace}(Y) + \text{trace}(Z) = \text{trace}(\text{diag}(Y, Z))$.

**Hint:** Show $\text{rank}(X) \leq r$ iff there exist $Y = Y^T \in \mathbb{R}^{m \times m}, Z = Z^T \in \mathbb{R}^{n \times n}$ with

$$
\begin{pmatrix}
Y & X \\
X^T & Z
\end{pmatrix} \succeq 0
$$

with $\text{rank}(Y) + \text{rank}(Z) \leq 2r$.

**References**