Given a full-rank matrix $X \in \mathbb{R}^{m \times n}$, in addition to finding the optimal D_1 in the conditioning procedure $XD_1^{1/2}$, we will also consider the following two scaling procedures in our paper: $D_2^{1/2}X$ and $D_2^{1/2}XD_1^{1/2}$. We are interested in comparing the performance of these three scaling procedures. In particular, we want to know how much the two-sided scaling improves upon the two one-sided scalings in terms of condition number reduction, and when the two one-sided methods perform differently.

The left-side optimal conditioning problem $D_2^{1/2}X$ can also be cast in DSDP form:

$$\begin{aligned}
&\max_{\tau, D} \tau \\
s.t. X^T DX &\succeq \tau I \\
I &\succeq X^T DX \\
D &\succeq 0
\end{aligned}$$

Note that $X^T DX = \sum_i D_{ii} X_i X_i^T$ where X_i is the i-th row of X, so the problem is in standard DSDP form.

On the other hand, the two-sided optimal scaling problem below,

$$\begin{aligned}
&\min_{\kappa, D_1, D_2} \kappa \\
s.t. X^T D_2 X &\succeq D_1 \\
\kappa D_1 &\succeq X^T D_2 X \\
D_1, D_2 &\succeq 0 \\
Tr(D_1 \cdot I) &\geq 1
\end{aligned}$$

cannot be recast as an SDP. We propose to instead use a bisection procedure to find a solution.

1. Start with some large κ_0 such that the SDP feasibility problem

$$\begin{aligned}
&\min_{D_1, D_2} 0 \\
s.t. X^T D_2 X &\succeq D_1 \\
\kappa_0 D_1 &\succeq X^T D_2 X \\
D_1, D_2 &\succeq 0 \\
Tr(D_1 \cdot I) &\geq 1
\end{aligned}$$

   has a solution. Set the upper bound $\overline{\kappa} = \kappa_0$ and lower bound $\underline{\kappa} = 1$.

2. Solve the DSDP feasibility problem

$$\begin{aligned} \min_{D_1, D_2} \; & 0 \\ s.t. X^T D_2 X & \succeq D_1 \\ \frac{\overline{\kappa} + \underline{\kappa}}{2} D_1 & \succeq X^T D_2 X \\ D_1, D_2 & \succeq 0 \\ Tr(D_1 \cdot I) & \geq 1 \end{aligned}$$

3. If the problem is feasible, set $\frac{\overline{\kappa}+\underline{\kappa}}{2}$ to be the new *upper* bound $\overline{\kappa}$. If the problem is infeasible, set $\frac{\overline{\kappa}+\underline{\kappa}}{2}$ to be new *lower* bound $\underline{\kappa}$.

4. Solve the feasibility problems iteratively, until $\overline{\kappa} - \underline{\kappa} < \epsilon$ for some tolerance parameter ϵ. The final set of feasible solutions D_1, D_2 will provide an ϵ-optimal two-sided preconditioning.

The bisection algorithm for two-sided optimal preconditioning may take significant time for certain problems. If that's the case, please feel free to solve a subset of problems first.

Here are two optional tasks that are preferable if you have time, especially the first one, but can focus on the above first.

- We need a baseline diagonal conditioner. Let's use the simplest one, $D^{-1}A$ (assuming $m > n$), where D is the "diagonal" of A. If we have time, we can also look at matrix equilibriation methods such as Sinkhorn.

- If we iteratively find optimal left and right optimal diagonal scalings, we should expect to converge to the solution to the two-sided optimal scaling problem. If possible, we want to check this for some real matrices.