CONVEX AND CONIC OPTIMIZATION
Homework 4

NILAVA METYA
nilava.metya@rutgers.edu
nm8188@princeton.edu

April 4, 2024

Problem 1

The nuclear norm of a matrix $A \in \mathbb{R}^{m\times n}$ is defined by

$$\|A\|_* := \sum_{i=1}^{\min\{m,n\}} \sigma_i(A).$$

1. Show that the dual norm of the spectral norm is the nuclear norm.
2. Show that the problem of minimizing the nuclear norm of a matrix subject to arbitrary affine constraints can be cast as a semidefinite program.

Solution

1. Let $\|\cdot\|$ denote the spectral norm here, and let the nuclear norm be $\|\cdot\|_*$. By definition, $\|A\|_* = \max\limits_{\substack{\|C\|\le 1\\ C\in\mathbb{R}^{m\times n}}} \operatorname{Tr}(C^\top A)$.

   For each matrix $X \in \mathbb{R}^{m\times n}$, denote it's singular values by $\sigma_1(X) \geq \sigma_2(X) \geq \cdots \geq \sigma_{\min\{m,n\}}(X)$ in non-increasing order.

   Lower bound on nuclear norm: Let $A = U\Sigma V^\top$ be the SVD of A. Consider a feasible C_0, namely $C_0 := U\mathcal{I}V^\top$ where $\mathcal{I} = \begin{cases} \begin{bmatrix} I_{m\times m} & \mathbf{0}_{m\times(n-m)} \end{bmatrix} & \text{if } n > m \\ \begin{bmatrix} I_{n\times n} \\ \mathbf{0}_{(m-n)\times n} \end{bmatrix} & \text{otherwise} \end{cases}$. This is feasible because it has dimension $m \times n$ and by it's SVD, it's singular values are all 1, whence $\sigma_1(C_0) = 1$. Since our optimization problem (i.e., the expression of the nuclear norm) is a maximization problem, the objective evaluated at any feasible point (C_0 in this case) is at most the optimal value. This gives

   $$\|A\|_* \geq \operatorname{Tr}(C_0^\top A) = \operatorname{Tr}(V\mathcal{I}^\top\Sigma V^\top) = \operatorname{Tr}(\mathcal{I}^\top\Sigma) = \sum_{i=1}^{\min\{m,n\}} \sigma_i(A).$$

   Upper bound on nuclear norm: Let C be feasible, that is, $C \in \mathbb{R}^{m\times n}$ and $\|C\| \leq 1$. Let $\{u_i\}$ and $\{v_i\}$ be the columns of U and V respectively (so these have norm 1), so that $A = \sum_{i=1}^{\min\{m,n\}} \sigma_i(A) u_i v_i^\top$. Then

$$\begin{aligned}
\mathrm{Tr}(C^\top A) &= \sum_{i=1}^{\min\{m,n\}} \sigma_i(A)\,\mathrm{Tr}(C^\top u_i v_i^\top) = \sum_{i=1}^{\min\{m,n\}} \sigma_i(A)\,\mathrm{Tr}(v_i^\top C^\top u_i) \\
&= \sum_{i=1}^{\min\{m,n\}} \sigma_i(A) v_i^\top C^\top u_i = \sum_{i=1}^{\min\{m,n\}} \sigma_i(A) u_i^\top C v_i \\
&\le \sum_{i=1}^{\min\{m,n\}} \sigma_i(A) \left\| u_i^\top C v_i \right\| \\
&\le \sum_{i=1}^{\min\{m,n\}} \sigma_i(A) \left\| u_i \right\| \left\| C \right\| \left\| v_i \right\| \qquad [\because 2-\text{norm is submultiplicative}] \\
&= \sum_{i=1}^{\min\{m,n\}} \sigma_i(A).
\end{aligned}$$

Since this is true for any feasible C, it must happen that $\|A\|_* \le \sum_{i=1}^{\min\{m,n\}} \sigma_i(A)$.

2. We want to minimize $\|X\|_*$ where $X \in \mathbb{R}^{m\times n}$ such that $\mathrm{Tr}(C_i X) = b_i\ \forall 1 \le i \le r$. But $\|X\|_* = \max_{\substack{\|C\|\le 1 \\ C\in\mathbb{R}^{m\times n}}} \mathrm{Tr}(C^\top X)$. So the problem of our interest is

$$\begin{aligned}
\min_{X\in\mathbb{R}^{m\times n}} \ & \max_{C\in\mathbb{R}^{m\times n}} \mathrm{Tr}(CX) \\
\text{s.t. } & \|C\| \le 1 \\
& \mathrm{Tr}(C_i^\top X) = b_i\ \forall 1 \le i \le r.
\end{aligned} \tag{1}$$

The condition $\|C\| \le 1$ can be written as $I \succeq C^\top C$. So, for each X, we are interested in the subproblem $\max_{\substack{I-C^\top C\succeq 0 \\ C\in\mathbb{R}^{m\times n}}} \mathrm{Tr}(C^\top X) = -\min_{\substack{I-C^\top C\succeq 0 \\ C\in\mathbb{R}^{m\times n}}} \mathrm{Tr}(C^\top X)$. The second expression is valid because the minimum of this expression would simply be the negative of the actual norm. This is an SDP because it can be written as

$$\begin{aligned}
-\min_{C\in\mathbb{R}^{m\times n}} \ & \mathrm{Tr}(C^\top X) \\
\text{s.t. } & \begin{bmatrix} I_m & C \\ C^\top & I_n \end{bmatrix} \succeq 0
\end{aligned} \tag{2}$$

by the theorem on Schur complements. Let's write it in standard SDP form:

$$\begin{aligned}
-\min_{C'\in S^{(m+n)\times(m+n)}} \ & \frac{1}{2}\,\mathrm{Tr}\left(\begin{bmatrix} \mathbf{0}_{m\times m} & X \\ X^\top & \mathbf{0}_{n\times n} \end{bmatrix}^\top C' \right) \\
\text{s.t. } & \mathrm{Tr}\left(\frac{1}{2}(e_i e_j^\top + e_j e_i^\top) C' \right) = \delta_{ij} \text{ for } 1 \le i < j \le m \\
& \mathrm{Tr}\left(\frac{1}{2}(e_i e_j^\top + e_j e_i^\top) C' \right) = \delta_{ij} \text{ for } m+1 \le i < j \le m+n \\
& C' \succeq 0.
\end{aligned}$$

The corresponding dual variable is given by some $y \in \mathbb{R}^{\frac{m(m+1)}{2}+\frac{n(n+1)}{2}}$ corresponding to each pair $1 \leq i < j \leq m$ and each pair $m+1 \leq i < j \leq m+n$. This is same as taking two matrices $P \in S^{m\times m}, Q \in S^{n\times n}$. Thus the dual problem is

$$
\begin{aligned}
- \max_{\substack{P\in S^{m\times m}\\ Q\in S^{n\times n}}} \; & \frac{1}{2}\operatorname{Tr}(I_m^\top P) + \frac{1}{2}\operatorname{Tr}(I_n^\top Q) \\
\text{s.t.} \; & \begin{bmatrix} \mathbf{0}_{m\times m} & X \\ X^\top & \mathbf{0}_{n\times n} \end{bmatrix} \succeq \frac{1}{2}\sum_{1\leq i,j\leq m} P_{ij}(e_ie_j^\top + e_je_i^\top) + \frac{1}{2}\sum_{m+1\leq i,j\leq m+n} Q_{(i-m)(j-m)}(e_ie_j^\top + e_je_i^\top).
\end{aligned}
$$

But $\frac{1}{2}\sum_{1\leq i,j\leq m} P_{ij}(e_ie_j^\top + e_je_i^\top) + \frac{1}{2}\sum_{m+1\leq i,j\leq m+n} Q_{(i-m)(j-m)}(e_ie_j^\top + e_je_i^\top) = \begin{bmatrix} P & \mathbf{0} \\ \mathbf{0} & Q \end{bmatrix}$.

So the (dual) problem is

$$
\begin{aligned}
\min_{\substack{P'\in S^{m\times m}\\ Q'\in S^{n\times n}}} \; & \frac{1}{2}\operatorname{Tr}(P') + \frac{1}{2}\operatorname{Tr}(Q') \\
\text{s.t.} \; & \begin{bmatrix} P' & X \\ X^\top & Q' \end{bmatrix} \succeq 0.
\end{aligned} \tag{3}
$$

with the transformation $P' = -P, Q' = -Q$. This is clearly an SDP (force the required entries to X using affine constraints and take the big matrix to be the variable).

Coming back to our original problem 1 of minimizing nuclear norm subject to affine constraints, we can now rewrite it as

$$
\begin{aligned}
\min_{\substack{X\in\mathbb{R}^{m\times n}\\ P\in S^{m\times m}\\ Q\in S^{n\times n}}} \; & \frac{1}{2}\operatorname{Tr}(P) + \frac{1}{2}\operatorname{Tr}(Q) \\
\text{s.t.} \; & \begin{bmatrix} P & X \\ X^\top & Q \end{bmatrix} \succeq 0 \\
& \operatorname{Tr}(C_i^\top X) = b_i \;\forall 1 \leq i \leq r
\end{aligned}
$$

because of strong duality of SDP (strict feasibility checked later). This is an SDP because the matrix we want to be positive semidefinite is linear in terms of the decision variables X, P, Q.

Checking strong duality: We want to check strict feasibility of both the primal 2 and dual 3 for each $X \in \mathbb{R}^{m\times n}$. Taking $C = \mathbf{0}_{m\times n}$ suffices in the primal 2 because then the required matrix is $I_{m+n} \succ 0$. For the dual 3, we take $P = I_m, Q = \lambda I_n$ where $\lambda \in \mathbb{R}$ is a number strictly more than every eigenvalue of $X^\top X$. Then the matrix of interest becomes $\begin{bmatrix} I_m & X \\ X^\top & \lambda I_n \end{bmatrix}$ and by Schur complements, since $I_m \succ 0$ and $\lambda I_n - X^\top X \succ 0$ (that's how λ was chosen), we conclude that this matrix is $\succ 0$.

Problem 2

You are given a list of distances d_{ij} for $(i,j) \in \{1,\cdots,m\} \times \{1,\cdots,m\}$. You would like to know whether there are points $x_i \in \mathbb{R}^n$, for some value of n, such that $\|x_i - x_j\| = d_{ij}\ \forall i,j$.

1. Show that this problem can be formulated as that of checking whether a fixed matrix whose entries depend on d_{ij} is positive semidefinite. If this test passes, how would you recover n and the points x_i?
2. Give an example of a set of distances that respect the triangle inequality but for which there does not exist an embedding in any dimension.

Solution

1. Start by noting that d_{ij} must be symmetric, that is, $d_{ij} = d_{ji}$, and non-negative. So we'll find the required matrix (whose positive definiteness guarantees the existence of such points) with these assumptions.

   Discussion:
   Observe that if such $x_1, \cdots, x_m \in \mathbb{R}^n$ exists, then $Y^\top Y$ (the sample covariance matrix) is positive semidefinite, where $Y = \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix} \in \mathbb{R}^{n\times m}$. In that case, the $(i,j)^{\text{th}}$ entry of $Y^\top Y$ would look like $x_i^\top x_j$. This $Y^\top Y$ is data dependent, so far from this discussion. Let's try to express it in terms of the d_{ij}'s.
   $d_{ij}^2 = \|x_i\|^2 + \|x_j\|^2 - 2x_i^\top x_j\ \forall i,j$. Replacing x_i by $x_i - \frac{1}{m}\sum_i x_i$ does not change this. So let's just do that, that is, assume that these x_i are centered at 0. Fixing each j we get $c_j := \sum_i d_{ij}^2 = \sum_i \|x_i\|^2 + m\|x_j\|^2$. Similarly fixing i and summing over j gives $e_i := \sum_j d_{ij}^2 = \sum_j \|x_j\|^2 + m\|x_i\|^2$. Let $C = \sum_{i,j} d_{ij}^2$. Thus we have the following

   $$\begin{aligned} d_{ij}^2 &= \|x_i\|^2 + \|x_j\|^2 - 2x_i^\top x_j \\ \implies \sum_i d_{ij}^2 &= \sum_i \|x_i\|^2 + m\|x_j\|^2 \\ \implies C = \sum_{i,j} d_{ij}^2 &= 2m\sum_i \|x_i\|^2 \\ \implies \sum_i \|x_i\|^2 &= \frac{C}{2m}. \end{aligned}$$

   Plugging this back in the second equation above, $c_j = \dfrac{C}{2m} + m\|x_j\|^2$ and thus $\|x_j\|^2 = \dfrac{c_j}{m} - \dfrac{C}{2m^2}$. Similarly, $\|x_i\|^2 = \dfrac{e_j}{m} - \dfrac{C}{2m^2}$ Plugging this back in gives $d_{ij}^2 = \dfrac{c_j + e_i}{m} - \dfrac{C}{m^2} - 2x_i^\top x_j$ and thus

   $$x_i^\top x_j = \frac{c_j + e_i}{2m} - \frac{C}{2m^2} - \frac{d_{ij}^2}{2}.$$

   So we form a matrix M whose $(i,j)^{\text{th}}$ entry is the above quantity, namely,

   $$\boxed{M_{ij} = \frac{1}{2}\left(-d_{ij}^2 + \frac{\sum_t d_{tj}^2}{m} + \frac{\sum_t d_{it}^2}{m} - \frac{\sum_{s,t} d_{st}^2}{m^2}\right)}.$$

   Wanting it to be of the form $Y^\top Y$ is equivalent to demanding that it's positive semidefinite.

   The actual proof:

If the d_{ij} are not symmetric, then there are no such points. Same if some $d_{ij} < 0$. Same if some $d_{ii} \neq 0$. So we assume $d_{ij} = d_{ji} \geq 0\ \forall i,j$ and $d_{ii} = 0\ \forall i$. So, M is also symmetric.

Claim 1
If there are points $x_1, \cdots, x_m \in \mathbb{R}^n$ such that $d_{ij} = \|x_i - x_j\|$, then M (as constructed above) is positive semidefinite.

Proof. By the above discussion, the existence of such points implies $M = Y^\top Y \succeq 0$. ■

Claim 2
If the above matrix M (obtained from d_{ij}, which are symmetric, non-negative and zero on diagonal) is positive semidefinite, then $\exists n \in \mathbb{Z}_{\geq 1}$ and $0-$centered points $x_1, \cdots, x_m \in \mathbb{R}^n$ such that $d_{ij} = \|x_i - x_j\|\ \forall i, j$.

Proof. The d_{ij} are symmetric, so M is symmetric. This means that $M = U^\top DU$ for some $m \times m$ orthogonal matrix $U^\top$ of eigenvectors with non-negative eigenvalues, and these eigenvalues are the diagonal entries of D. Assume these are arranged in non-increasing order, that is $D = diag(\lambda_1^2, \cdots, \lambda_m^2)$ such that $\lambda_1^2 \geq \cdots \geq \lambda_m^2$. Let n be the number of positive eigenvalues (this is the n we want), that is, $\lambda_1^2 \geq \cdots \geq \lambda_n^2 > 0 = \lambda_{n+1}^2 = \lambda_{n+2}^2 = \cdots = \lambda_m^2$. Here $U^\top = \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$ where u_i are the eigenvectors, that is, $Mu_i = \lambda_i^2 u_i\ \forall i$. Clearly,

$$\begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix} \begin{bmatrix} \lambda_1^2 & & & & \\ & \ddots & & & \\ & & \lambda_n^2 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix} \begin{bmatrix} u_1^\top \\ \\ \vdots \\ \\ u_m^\top \end{bmatrix} = \begin{bmatrix} \lambda_1 u_1 & \cdots & \lambda_n u_n \end{bmatrix}_{m \times n} \underbrace{\begin{bmatrix} \lambda_1 u_1^\top \\ \\ \vdots \\ \\ \lambda_n u_n^\top \end{bmatrix}_{n \times m}}_{Y}.$$

So our data points $x_1, \cdots, x_m$ are the columns of the matrix $\begin{bmatrix} \lambda_1 u_1^\top \\ \\ \vdots \\ \\ \lambda_n u_n^\top \end{bmatrix}$. By the discussion before these two claims, these $x_1, \cdots, x_m \in \mathbb{R}^n$ satisfy $x_i^\top x_j = M_{ij} = \frac{1}{2}\left(-d_{ij}^2 + \frac{\sum_t d_{tj}^2}{m} + \frac{\sum_t d_{it}^2}{m} - \frac{\sum_{s,t} d_{st}^2}{m^2}\right)$. So

$$\|x_i - x_j\|^2 = \|x_i\|^2 + \|x_j\|^2 - 2x_i^\top x_j = -\frac{1}{2}d_{ii}^2 - \frac{1}{2}d_{jj}^2 + \frac{\sum_t d_{it}^2}{m} + \frac{\sum_t d_{jt}^2}{m} - \frac{\sum_{s,t} d_{st}^2}{m^2} + d_{ij}^2 - \frac{\sum_t d_{it}^2}{m} - \frac{\sum_t d_{jt}^2}{m} + \frac{\sum_t d_{it}^2}{m} + \frac{\sum_t d_{jt}^2}{m} = d_{ij}^2.$$

This shows that our constructed $x_1, \cdots, x_m$ indeed satisfy the given distances. Finally, it is easy to see that $2\sum_j M_{ij} = -\sum_j d_{ij}^2 + \frac{1}{m}\sum_{t,j} d_{tj}^2 + \sum_t d_{it}^2 - \frac{m}{m^2}\sum_{s,t} d_{st}^2 = 0$. So $\mathbf{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue 0. This means $\mathbf{1}$ is perpendicular to all eigenvectors $\lambda_1 u_1, \cdots, \lambda_n u_n$ with positive eigenvalues, that is, the column vectors of the matrix with rows $\lambda_i u_i^\top$ $(1 \leq i \leq m)$ sum to 0. In other words, $x_1 + \cdots + x_m = 0$. ■

The above two claims prove that

Theorem 3
Say $d_{ij} \in \mathbb{R}$, for $(i, j) \in [m] \times [m]$, are given. There exists $n \in \mathbb{Z}_{\geq 1}$ and ($0-$centered) points $x_1, \cdots, x_m \in$

$\mathbb{R}^n$ if and only if $d_{ji} = d_{ij} \geq 0, d_{ii} = 0 \; \forall i, j$ and the $m \times m$ matrix M, given by

$$M_{ij} = \frac{1}{2}\left(-d_{ij}^2 + \frac{\sum_t d_{it}^2}{m} + \frac{\sum_t d_{jt}^2}{m} - \frac{\sum_{s,t} d_{st}^2}{m^2}\right),$$

is positive semidefinite.

The above proof shows that if this matrix M is positive semidefinite then we can find points in $\mathbb{R}^n$, where $n = \operatorname{rank} M$, and these points are determined by finding a Cholesky decomposition of M, that is $M = Y^\top Y$ with the columns of Y being the data points.

2. Consider $m = 2$ with $d_{11} = d_{22} = 1, d_{12} = d_{21} = 2$. Then clearly the triangle inequality is satisfied:

   - $1 = d_{11} \leq d_{12} + d_{21} = 4$.
   - $1 = d_{22} \leq d_{21} + d_{12} = 4$.
   - $2 = d_{12} \leq d_{11} + d_{12} = d_{12} + d_{22} = 3$.
   - $2 = d_{21} \leq d_{21} + d_{11} = d_{22} + d_{21} = 3$.

   But there cannot be a set of points in any dimension satisfying these distances because, for example, $d_{11} \neq 0$.

Problem 3

Recall that the spectral radius of a matrix $A \in \mathbb{R}^{n\times n}$, denoted by $\rho(A)$, is the maximum of the absolute values of its eigenvalues. We call a matrix "stable" if $\rho(A) < 1$. Let us call a pair of real $n\times n$ matrices $\{A_1, A_2\}$ stable if $\rho(\Sigma) < 1$, for any finite product Σ out of A_1 and A_2. (For example, Σ could be $A_2A_1, A_1A_2, A_1A_1A_2A_1$, and so on.)

1. Does stability of A_1 and A_2 imply stability of the pair $\{A_1, A_2\}$?
2. Prove (possibly using optimization) that the pair $\{A_1, A_2\}$ with

$$A_1 = \frac{1}{4}\begin{bmatrix} -1 & -1 \\ -4 & 0 \end{bmatrix}, \quad A_2 = \frac{1}{4}\begin{bmatrix} 3 & 3 \\ -2 & 1 \end{bmatrix}$$

is stable.

Solution

1. **No**.

$\begin{bmatrix} 0.5 & 1 \\ 0 & 0.5 \end{bmatrix}\begin{bmatrix} 0.5 & 0 \\ 1 & 0.5 \end{bmatrix} = \begin{bmatrix} 1.25 & 0.5 \\ 0.5 & 0.25 \end{bmatrix}$ has eigenvalues $\dfrac{3 \pm 2\sqrt{2}}{4}$ and $3 + 2\sqrt{2} > 4$. But the two matrices, that were multiplied, have only eigenvalues $\frac{1}{2} < 1$.

2. We'll use the following lemma.

Lemma 4
Let A, B be two $n\times n$ matrices. If there exists a positive definite matrix $P \in S^{n\times n}$ such that $P - A^\top PA \succ 0$ and $P - B^\top PB \succ 0$ then $\{A, B\}$ is a stable pair of matrices.

Proof. Let $P \succ 0$ be as given. Consider any finite product Σ made out of A, B. It's already known from lecture that $V(x) := x^\top Px$ is a Lyapunov function for both A and B. In particular $V(Ax) < V(x), V(Bx) < V(x)\ \forall x \neq 0$. We will show that V is a Lyapunov function for Σ. $V(x) > 0\ \forall x \neq 0$ and $V(0) = 0$, and V is coercive: because V is Lyapunov for A. It only remains to check that $V(\Sigma x) < V(x)$ for $x \neq 0$. This is clear because if $\Sigma = M_k \cdots M_2M_1$ where $X_i \in \{A, B\}$ then for $x \neq 0$ we have $V(M_k \cdots M_2M_1x) \leq V(M_{k-1} \cdots M_2M_1x) \leq \cdots \leq V(M_1x)) < V(x)$ where the last inequality is strict due to the nonzeroness of x. This shows that the dynamical system given by $x_{k+1} = \Sigma x_k$ is a GAS. In other words, $\rho(\Sigma) < 1$. ■

As a consequence of this lemma, it remains to find a positive definite matrix $P \in S^{2\times 2}$ such that $P \succ A_1^\top PA_1$ and $P \succ A_2^\top PA_2$. Letting `CVXPY` do it does the job. We find a solution to the following SDP (since there is nothing to optimize, it only returns a feasible element):

$$\begin{aligned} \min_{X \in S^{2\times 2}}\ & 0 \\ \text{s.t. } & X - A_1^\top XA_1 \succ 0 \\ & X - A_2^\top XA_2 \succ 0 \\ & X \succ 0. \end{aligned}$$

But `CVXPY` 'wants' psd criteria, not strict ones. So we feed the following SDP to `Python`:

$$\begin{aligned} \min_{X \in S^{2\times 2}} \; & 0 \\ \text{s.t. } & X - A_1^\top X A_1 - \tilde{I} \succeq 0 \\ & X - A_2^\top X A_2 - \tilde{I} \succeq 0 \\ & X - \tilde{I} \succeq 0 \end{aligned}$$

where $\tilde{I} = \begin{bmatrix} 0.0002 & 0 \\ 0 & 0.0002 \end{bmatrix}$. This gives the (approximate) solution $\begin{bmatrix} 93.85890325 & 22.86467265 \\ 22.86467265 & 73.98959241 \end{bmatrix}$.

It turns out that the (close enough) integer matrix $P = \begin{bmatrix} 93 & 22 \\ 22 & 73 \end{bmatrix}$ works. Indeed with this choice of P, the above matrices of interest (which we we want to be positive definite) turn out to be

$$P - A_1^\top P A_1 = \begin{bmatrix} 3.1875 & 10.6875 \\ 10.6875 & 67.1875 \end{bmatrix}, \quad P - A_2^\top P A_2 = \begin{bmatrix} 38.9375 & -17.0625 \\ -17.0625 & 7.875 \end{bmatrix}.$$

Their entry at $(1,1)$ position are positive and their determinants (computed in code) are positive. So these are positive definite. Moreover, it is clear that P has positive determinant because both $93 > 22, 73 > 22$ and top-left entry is positive, whence P is positive definite.

Code included on next page.

[1]:
```
import numpy as np
import cvxpy as cp

n = 2
A1 = np.array([[-1/4, -1/4],[-4/4, 0]])
A2 = np.array([[3/4, 3/4],[-2/4, 1/4]])
id = np.array([[1, 0],[0, 1]])
C = np.array([[0,0],[0,0]])

X = cp.Variable((n,n), symmetric=True)
constraints = [X - A1.T @ X @ A1 - 0.0002*id >> 0, X - A2.T @ X @ A2 - 0.0002*id␣
↪>> 0 , X - 0.0002*id >> 0]
prob = cp.Problem(cp.Minimize(cp.trace(C @ X)),constraints)
prob.solve()
X.value
```

[1]:
```
array([[93.85890325, 22.86467265],
       [22.86467265, 73.98959241]])
```

Let's try it out with $P = \begin{bmatrix} 93 & 22 \\ 22 & 73 \end{bmatrix}$.

[2]:
```
from numpy import linalg as LA
P = np.array([[93, 22],[22, 73]])
B1 = P - (A1.T) @ P @ A1
B2 = P - (A2.T) @ P @ A2
print(LA.det(B1),LA.det(B2))
print("\nP - (A1.T) @ P @ A1 = \n", B1)
print("\nP - (A2.T) @ P @ A2 = \n", B2)
```

```
99.93749999999997 15.503906249999993

P - (A1.T) @ P @ A1 =
 [[ 3.1875 10.6875]
 [10.6875 67.1875]]

P - (A2.T) @ P @ A2 =
 [[ 38.9375 -17.0625]
 [-17.0625   7.875 ]]
```

Problem 4

Give an example of symmetric $n \times n$ matrices $A_1, \cdots, A_m$ with rational entries and rational numbers $b_1, \cdots, b_m$ such that the set

$$S = \left\{ X \in S^{n \times n} \;\middle|\; \mathrm{Tr}(A_i X) = b_i \;\forall 1 \leq i \leq m, X \succeq 0 \right\}$$

is non-empty, but only contains matrices that have at least one irrational entry. Here, $S^{n \times n}$ denotes the set of symmetric $n \times n$ matrices with real entries and Tr stands for the trace operation.

Solution

Recall that if A, B are square matrices then $\begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & B \end{bmatrix} \succeq 0$ iff $A \succeq 0$ and $B \succeq 0$. This is because the eigenvalues of this block matrix is the union of the set of eigenvalues of A and those of B.

Note the following:

- $\begin{bmatrix} 2x & 2 \\ 2 & x \end{bmatrix} \succeq 0 \iff (x \geq 0 \text{ and } 2x^2 \geq 4) \iff (x \geq 0 \text{ and } x^2 \geq 2) \iff x \geq \sqrt{2}$.
- $\begin{bmatrix} 1 & x \\ x & 2 \end{bmatrix} \succeq 0 \iff 2 - x^2 \geq 0 \iff -\sqrt{2} \leq x \leq \sqrt{2}$.

Thus we have the following set if equivalent statements for the matrix $A := \begin{bmatrix} 2x & 2 & 0 & 0 \\ 2 & x & 0 & 0 \\ 0 & 0 & 1 & x \\ 0 & 0 & x & 2 \end{bmatrix} \in S^{4 \times 4}$:

$$\begin{aligned} & A \succeq 0 \\ \iff & \begin{bmatrix} 2x & 2 \\ 2 & x \end{bmatrix} \succeq 0 \text{ and } \begin{bmatrix} 1 & x \\ x & 2 \end{bmatrix} \succeq 0 \\ \iff & x \geq \sqrt{2} \text{ and } -\sqrt{2} \leq x \leq \sqrt{2} \\ \iff & x = \sqrt{2}. \end{aligned}$$

Now it remains to put linear constraints on $X \in S^{4 \times 4}$ so that it takes the form of A with x variable. This is obtained as follows (we denote $\mathrm{Tr}(CX) = C * X$):

1. $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} * X = 0.$
2. $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} * X = 0.$
3. $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} * X = 0.$
4. $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} * X = 0.$
5. $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} * X = 4.$
6. $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} * X = 1.$
7. $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} * X = 2.$
8. $\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} * X = 0.$
9. $\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} * X = 0.$

Step #	What it does	X after this step
$1-4$	Sets the top right 4×4, and hence bottom left, portion to 0	$\begin{bmatrix} a & b & 0 & 0 \\ b & c & 0 & 0 \\ 0 & 0 & d & x \\ 0 & 0 & x & e \end{bmatrix}$
5	Takes the off-diagonals in the top left 4×4 corner (that is, b) and sets their sum to 4. So $b + b = 4 \implies b = 2$.	$\begin{bmatrix} a & 2 & 0 & 0 \\ 2 & c & 0 & 0 \\ 0 & 0 & d & x \\ 0 & 0 & x & e \end{bmatrix}$
6	Sets the top-left entry of the bottom-right 4×4 submatrix to 1.	$\begin{bmatrix} a & 2 & 0 & 0 \\ 2 & c & 0 & 0 \\ 0 & 0 & 1 & x \\ 0 & 0 & x & e \end{bmatrix}$
7	Sets the bottom-right entry to 2.	$\begin{bmatrix} a & 2 & 0 & 0 \\ 2 & c & 0 & 0 \\ 0 & 0 & 1 & x \\ 0 & 0 & x & 2 \end{bmatrix}$
8	Equates $X_{11} = X_{34} + X_{43} = 2x$.	$\begin{bmatrix} 2x & 2 & 0 & 0 \\ 2 & c & 0 & 0 \\ 0 & 0 & 1 & x \\ 0 & 0 & x & 2 \end{bmatrix}$
9	Equates $X_{22} = \frac{1}{2} X_{11}$.	$\begin{bmatrix} 2x & 2 & 0 & 0 \\ 2 & x & 0 & 0 \\ 0 & 0 & 1 & x \\ 0 & 0 & x & 2 \end{bmatrix}$

We make a table to show what X looks like, after the above linear constraints are taken into account.

From the above, it is clear that $n = 4, m = 9$ and the enumeration gives what the $A_1, \cdots, A_m \in S^{n \times n}$ and $b_1, \cdots, b_m \in \mathbb{R}$ are. There is only one matrix $X \in S^{n \times n}$ which satisfies $A_i X = b_i \ \forall 1 \leq i \leq m$ and $X \succeq 0$, namely $\begin{bmatrix} 2\sqrt{2} & 2 & 0 & 0 \\ 2 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 1 & \sqrt{2} \\ 0 & 0 & \sqrt{2} & 2 \end{bmatrix}$.