IOE 611: Homework 4

Yulun Zhuang

November 10, 2024

Problem 1

Each of the following `cvx` code fragments describes a convex constraint on the scalar variables x, y, and z, but violates the `cvx` rule set, and so is invalid. Briefly explain why each fragment is invalid. Then, rewrite each one in an equivalent form that conforms to the `cvx` rule set.

(a) `norm([x + 2*y, x - y]) == 0` is invalid because the equality constraints have to be affine for the problem to be convex. Since $\|v\| = 0$ if and only if $v = 0$ element-wise, the constraint can be reformulated as

```
x + 2*y == 0;
x - y == 0;
```

(b) `square(square(x + y)) <= x - y` is invalid because the convexity of a square function of a convex function can not be determined. The square of square is equivalent to the fourth power,

```
power(x + y, 4) <= x - y;
```

(c) `1/x + 1/y <= 1; x >= 0; y >= 0` is invalid because $1/x$ is not convex without restricting the domain to $\mathbb{R}_{++}$. It can be reformulated as

```
inv_pos(x) + inv_pos(y) <= 1;
```

(d) `norm([max(x, 1), max(y, 2)]) <= 3*x + y` is invalid because the convexity of the norm of a convex function is undetermined. Introduce and minimize over additional variables $[u, v]$ which is convex and non-decreasing over its domain

```
norm([u, v]) <= 3*x + y;
u >= max(x, 1)
v >= max(y, 2)
```

(e) `x*y >= 1; x >= 0; y >= 0` if invalid because `x*y` is nonlinear. However, given the variables are both positive from the combination of constraints, it can be reformulated as

```
y >= inv_pos(x)
```

(f) `(x + y)^2 / sqrt(y) <= x - y + 5` is invalid because LHS is convex function over concave function, which has undetermined convexity. It can be reformulated as a quadratic term over a nonincreasing linear term as

```
quad_over_lin(x + y, sqrt(y)) <= (x - y + 5)
```

(g) `x^3 + y^3 <= 1; x >= 0; y >= 0` is invalid because x^3 is not convex. Instead, use `pow_pos(x, p)`

```
pow_pos(x, 3) + pow_pos(y, 3) <= 1;
```

(h) `x + z <= 1 + sqrt(x*y - z^2); x >= 0; y >= 0` is invalid because `x*y` is nonlinear. Observe that

$$\sqrt{xy - z^2} = \left(\det \begin{bmatrix} x & z \\ z & y \end{bmatrix}\right)^{1/2}$$

The constraint is equivalent to

```
x + z <= 1 + det_rootn([x, z; z, y]);
x >= 0; y >= 0;
```

where for $X \in \mathbb{R}^{n\times n}$, `det_rootn(X)` $= \det(X)^{1/n}$.

Problem 2

Consider a two-dimensional bounded object $R \subset \mathbb{R}^2$ that has density $\rho(z)$ at point $z = (x, y) \in \mathbb{R}^2$. Then the mass $m \in \mathbb{R}$, center of gravity $c \in \mathbb{R}^2$, and inertia matrix $M \in \mathbb{R}^{2\times 2}$ of this object are given by, respectively,

$$m = \int_R \rho(z) dxdy,\ c = \frac{1}{m}\int_R \rho(z) z dxdy,\ M = \int_R \rho(z)(z-c)(z-c)^\top dxdy$$

(a) Suppose R is discretized into n pixels, each of area a, with pixel i having constant density ρ_i throughout and location (say, of its center) $z_i \in \mathbb{R}^2$. Replace expressions for m, c, and M with sums and averages using this discretization.

(b) Formulate the following as a convex optimization problem: choose the density vector ρ in order to maximize $\lambda_{\min}(M)$, subject to $0 \le \rho(z) \le \rho_{\max}$ for all $z \in \mathbb{R}$, and a fixed total mass $m = m_{\text{given}}$.

(c) Apply your method to the instance with data in `hwk4p2data.m`.

Solution. (a)

$$
\begin{aligned}
m &= a\sum_{i=1}^n \rho_i \\
c &= \frac{a}{m}\sum_{i=1}^n \rho_i z_i \\
M &= a\sum_{i=1}^n \rho_i (z_i - c)(z_i - c)^\top
\end{aligned}
$$

(b) Given that $\lambda_{\min}(M) \ge t \Leftrightarrow M \succeq tI$, the problem can be formulated as

$$
\begin{aligned}
\max_{\rho, t} \quad & t \\
\text{s.t.} \quad & a\sum_{i=1}^n \rho_i (z_i - c)(z_i - c)^\top \succeq tI \\
& c = \frac{a}{m_{\text{given}}}\sum_{i=1}^n \rho_i z_i \\
& a\sum_{i=1}^n \rho_i = m_{\text{given}} \\
& 0 \le \rho_i \le \rho_{\max},\ i = 1, \dots, n
\end{aligned}
$$

Note that M can be further simplified as

$$\begin{aligned}
M &= a\sum_{i=1}^{n}\rho_i(z_i - c)(z_i - c)^\top \\
&= a\sum_{i=1}^{n}\rho_i(z_i z_i^\top - c z_i^\top - z_i c^\top + cc^\top) \\
&= a\sum_{i=1}^{n}\rho_i z_i z_i^\top - ac\sum_{i=1}^{n}(\rho_i z_i)^\top - a(\sum_{i=1}^{n}\rho_i z_i)c^\top + cc^\top a\sum_{i=1}^{n}\rho_i \\
&= a\sum_{i=1}^{n}\rho_i z_i z_i^\top - m_{\text{given}}cc^\top
\end{aligned}$$

Now we work on the only nonlinear constraint. By using the Schur complement theorem, this constraint can be reformulated as

$$\begin{aligned}
&(a\sum_{i=1}^{n}\rho_i z_i z_i^\top - tI) - m_{\text{given}}cc^\top \succeq 0 \\
\Leftrightarrow\ &\begin{bmatrix} a\sum_{i=1}^{n}\rho_i z_i z_i^\top - tI & c^\top \\ c & 1/m_{\text{given}} \end{bmatrix} \succeq 0
\end{aligned}$$

which is a Linear Matrix Inequality (LMI) in ρ, t.

(c) The optimal inertia matrix is

$$M^* = \begin{bmatrix} 0.6484 & -0.0000 \\ -0.0000 & 0.6484 \end{bmatrix}$$

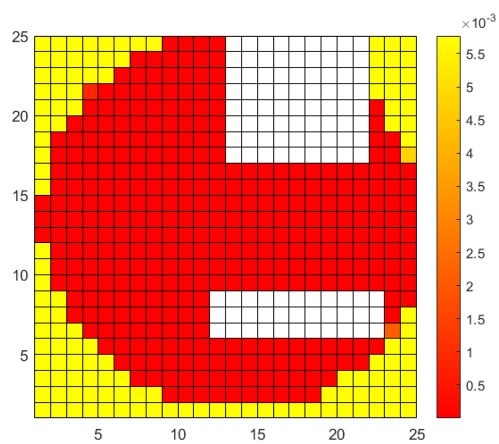

Figure 1: Visualization of the optimal density distribution.

□

Problem 3

Implement the SDP relaxation of the Max-Cut problem.

(a) Use your code to solve an instance of the problem with a weight matrix provided in `hw4p3data1.m`. Plot the distribution of the eigenvalues of Y. Is your SDP relaxation exact? If yes, recover the optimal solution for the Max-Cut problem from Y. If no, explain why.

(b) Use your code to solve an instance of the problem with a weight matrix provided in `hw4p3data2.m`. Plot the distribution of the eigenvalues of Y. Is your SDP relaxation exact? If yes, recover the optimal solution for the Max-Cut problem from Y. If no, explain why.

Solution. (a) The SDP relaxation is not exact since the number of nonzero eigenvalues of Y is more than 1, i.e. the rank of Y is greater than 1.

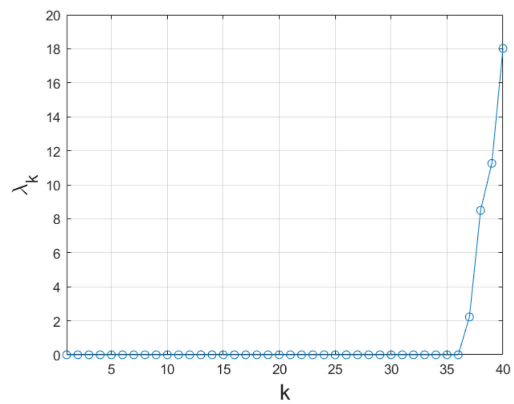

Figure 2: Eigenvalues of Y for `data1`

(b) The SDP relaxation is exact since the number of nonzero eigenvalues of Y is equal to 1 i.e. the rank of Y is equal to 1.

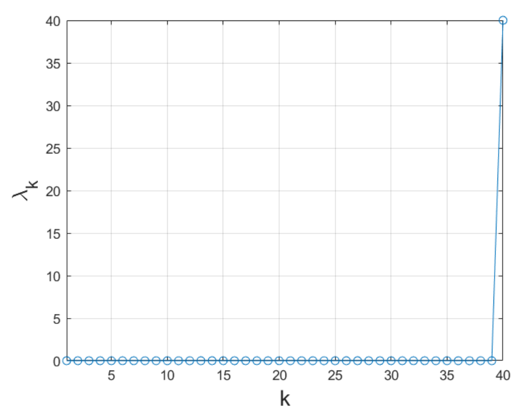

Figure 3: Eigenvalues of Y for `data2`

The recovered solution is

$$x = \sqrt{\lambda_{40}} v_{40} = [-\mathbf{1}_{1\times 20}, \mathbf{1}_{1\times 20}]^\top$$

□

Problem 4

Weak duality for unbounded and infeasible problems. The weak duality inequality, $d^* \leq p^*$, clearly holds when $d^* = -\infty$ or $p^* = \infty$. Show that it holds in the other two cases as well: If $p^* = -\infty$, then we must have $d^* = -\infty$, and also, if $d^* = \infty$, then we must have $p^* = \infty$.

Proof. (a) If $p^* = -\infty$, then the primal problem is unbounded below. In other words, there exists feasible x that makes $f_0(x) \to -\infty$. Then, for the dual function with $\lambda_i \geq 0$ for $i = 1, \ldots, m$,

$$\begin{aligned} g(\lambda, \nu) &= \inf_x f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \\ &= -\infty + \inf_x \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \\ &= -\infty \end{aligned}$$

Since the above holds for all feasible (λ, ν),

$$d^* = \max g(\lambda, \nu) = -\infty$$

(b) To prove by contradiction, assume $p^* \neq \infty$, which implies that p^* is feasible. Therefore there exists $\bar{x}$ such that

$$L(\lambda, \nu, \bar{x}) = f_0(\bar{x}) + \sum_{i=1}^m \lambda_i f_i(\bar{x}) + \sum_{i=1}^p \nu_i h_i(\bar{x})$$

is obtained with $f_i(\bar{x}) \leq 0$ for $i = 1, \ldots, m$ and $h_i(\bar{x}) = 0$ for $i = 1, \ldots, p$. $L(\lambda, \nu, \bar{x})$ can be an upper bound for g, which is the infimum of L,

$$g(\lambda, \nu) = \inf_x f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \leq L(\lambda, \nu, \bar{x})$$

Therefore, if $p^* \neq \infty$, $d^* = \max g(\lambda, \nu) = \infty$ is not true since $g(\lambda, \nu)$ is bounded above. Then, by contradiction, if $g = \infty$, then $p^* = \infty$. $\square$

Problem 5

Suboptimality of a simple covering ellipsoid. Recall the problem of determining the minimum volume ellipsoid, centered at the origin, that contains the points $a_1, \dots, a_m \in \mathbb{R}^n$

$$\begin{array}{ll} \text{minimize} & f_0(X) = \log\det(X^{-1}) \\ \text{subject to} & a_i^\top X a_i \leq 1, \ i = 1, \dots, m, \end{array}$$

with $\mathbf{dom}\, f_0 = \mathbb{S}^n_{++}$. We assume that the vectors $a_1, \dots, m$ span $\mathbb{R}^n$ (which implies that the problem is bounded below).

(a) Show that the matrix

$$X_{sim} = \left(\sum_{k=1}^m a_k a_k^\top \right)^{-1},$$

is feasible.

(b) Now we establish a bound on how suboptimal the feasible point X_{sim} is, via the dual problem,

$$\begin{array}{ll} \text{minimize} & \log\det(\sum_{i=1}^m \lambda_i a_i a_i^\top) - \mathbf{1}^\top \lambda + n \\ \text{subject to} & \lambda \succeq 0, \end{array}$$

with the implicit constraint $\sum_{i=1}^m \lambda_i a_i a_i^\top \succ 0$. To derive a bound, we restrict our attention to dual variables of the form $\lambda = t\mathbf{1}$, where $t > 0$. Find (analytically) the optimal value of t, and evaluate the dual objective at this λ. Use this to prove that the volume of the ellipsoid $\{u \mid u^\top X_{sim} u \leq 1\}$ is no more than a factor $(m/n)^{n/2}$ more than the volume of the minimum volume ellipsoid.

Proof. (a)

$$\begin{bmatrix} \sum_{k=1}^m a_k a_k^\top & a_i \\ a_i & 1 \end{bmatrix} = \sum_k \begin{bmatrix} a_k \\ 0 \end{bmatrix} \begin{bmatrix} a_k \\ 0 \end{bmatrix}^\top + \begin{bmatrix} a_i \\ 1 \end{bmatrix} \begin{bmatrix} a_i \\ 1 \end{bmatrix}^\top$$

where $k = 1, \dots, m$ and $k \neq i$. Each of the term is in the form of $V^\top V$ and therefore is PSD. Hence, $Z \succeq 0$ because it is a sum of PSDs.

Let $A = \sum_{k=1}^m a_k a_k^\top, B = a_i, C = 1$, then based on Schur complements, since $Z \succeq 0$,

$$\begin{aligned} C - B^\top A^{-1} B &\geq 0 \\ 1 - a_i^\top \left(\sum_{k=1}^m a_k a_k^\top \right)^{-1} a_i &\geq 0 \\ a_i^\top X_{\text{sim}} a_i &\leq 1 \end{aligned}$$

for $i = 1, \dots, m$. Therefore, X_{sim} is feasible.

(b) Firstly, for $X = X_{\text{sim}}$, the primal objective value is

$$f_0(X_{\text{sim}}) = \log\det(X_{\text{sim}}^{-1}) = \log\det \left(\sum_{k=1}^m a_k a_k^\top \right)$$

Then, to derive the bound via dual with $\lambda = t\mathbf{1}$,

$$\begin{aligned} L(\lambda = t\mathbf{1}) &= \log\det\left(\sum_{i=1}^m t a_i a_i^\top\right) - \mathbf{1}^\top t\mathbf{1} + n \\ &= \log\left(t^n \det\left(\sum_{i=1}^m a_i a_i^\top\right)\right) - mt + n \\ &= \log\det\left(\sum_{i=1}^m a_i a_i^\top\right) + n\log(t) - mt + n \end{aligned}$$

To derive the optimal t^* for the dual problem, find the derivative and set it to zero,

$$\begin{aligned} \frac{dL}{dt} &= \frac{n}{t} - m = 0 \\ t^* &= \frac{n}{m} \end{aligned}$$

Substituting to $L(\lambda = t\mathbf{1})$, the dual objective value is

$$g(t\mathbf{1}) = \log\det\left(\sum_{i=1}^m a_i a_i^\top\right) + n\log\left(\frac{n}{m}\right)$$

The duality gap, $f_0(X_{\text{sim}}) - g(t\mathbf{1})$, is

$$\log\det\left(\sum_{i=1}^m a_i a_i^\top\right) - \log\det\left(\sum_{i=1}^m a_i a_i^\top\right) - n\log\left(\frac{n}{m}\right) = n\log\left(\frac{m}{n}\right)$$

which means X_{sim} is at most $n\log\left(\frac{m}{n}\right)$ larger than the optimal primal objective. The volume of the ellipsoid $\mathcal{E}_x = \{z \mid z^\top X z \leq 1\}$ is proportional to $(\det X^{-1})^{1/2}$, which can be transformed from the primal by multiplying by $1/2$ and taking the exponential,

$$\exp(2f_0) = \exp\left(\frac{1}{2}\log\det(X^{-1})\right) = (\det X^{-1})^{1/2}$$

Applying the same transformation to the duality gap,

$$\exp\left(\frac{n}{2}\log\left(\frac{m}{n}\right)\right) = \left(\frac{m}{n}\right)^{n/2}$$

Therefore, the volume of the ellipsoid with X_{sim} is no more than a factor of $\left(\frac{m}{n}\right)^{n/2}$ more than that of the minimum volume ellipsoid.

$\square$