SI251 - Convex Optimization, 2024 Spring
Homework 2

Name: **Zhou Shouchen**
Student ID: 2021533042

Due 23:59 (CST), Apr. 10, 2024

1. (50 pts) **Robust quadratic programming**. In the lecture, we have learned about robust linear programming as an application of second-order cone programming. Now we will consider a similar robust variation of the convex quadratic program

$$\begin{array}{ll} \text{minimize} & (1/2)x^T Px + q^T x + r \\ \text{subject to} & Ax \preceq b. \end{array}$$

For simplicity, we assume that only the matrix P is subject to errors, and the other parameters (q, r, A, b) are exactly known. The robust quadratic program is defined as

$$\begin{array}{ll} \text{minimize} & \sup_{P \in \mathcal{E}} \left((1/2)x^T Px + q^T x + r\right) \\ \text{subject to} & Ax \preceq b \end{array}$$

where $\mathcal{E}$ is the set of possible matrices P.

For each of the following sets $\mathcal{E}$, express the robust QP as a convex problem in a standard form (e.g., QP, QCQP, SOCP, SDP).

(a) A finite set of matrices: $\mathcal{E} = \{P_1, \ldots, P_K\}$, where $P_i \in S_+^n, i = 1, \ldots, K$.

(b) A set specified by a nominal value $P_0 \in S_+^n$ plus a bound on the eigenvalues of the deviation $P - P_0$:

$$\mathcal{E} = \{P \in \mathbf{S}^n \mid -\gamma I \preceq P - P_0 \preceq \gamma I\}$$

where $\gamma \in \mathbf{R}$ and $P_0 \in \mathbf{S}_+^n$.

(c) An ellipsoid of matrices:

$$\mathcal{E} = \left\{ P_0 + \sum_{i=1}^{K} P_i u_i \mid \|u\|_2 \leq 1 \right\}.$$

You can assume $P_i \in \mathbf{S}_+^n, i = 0, \ldots, K$.

Solution:

Since $\sup\limits_{P\in\mathcal{E}} \left(\frac{1}{2}x^T Px + q^T x + r\right) = \left(\sup\limits_{P\in\mathcal{E}} \frac{1}{2}x^T Px\right) + q^T x + r$, so we could need to consider the $\sup\limits_{P\in\mathcal{E}} \frac{1}{2}x^T Px$ part of the objective function.

(a) Let

$$t = \sup_{P\in\mathcal{E}} \frac{1}{2}\mathbf{x}^T P\mathbf{x}$$

i.e.

$$\forall i = 1, \cdots, K,\ \frac{1}{2}\mathbf{x}^T P_i\mathbf{x} \leq t$$

So the program can be rewriten as

$$\begin{array}{ll} \min\limits_{\mathbf{x},t} & t + q^T\mathbf{x} + r \\ \text{subject to} & A\mathbf{x} \preceq b \\ & \frac{1}{2}\mathbf{x}^T P_i\mathbf{x} \leq t, \quad i = 1, \ldots, K \end{array}$$

The objective function is linear to the variable $(\mathbf{x}, t)$, and the constrains are in quadratic form.
So above all, the problem is a QCQP.

(b) $\forall P \in \mathcal{E}$, we have

$$-\gamma I \preceq P - P_0 \preceq \gamma I$$

which means that $\forall \mathbf{x} \in \mathbb{R}^n$, we have

$$\mathbf{x}^T(-\gamma I)\mathbf{x} \leq \mathbf{x}^T(P - P_0)\mathbf{x} \leq \mathbf{x}^T(\gamma I)\mathbf{x}$$

i.e.

$$\mathbf{x}^T(P_0 - \gamma I)\mathbf{x} \leq \mathbf{x}^T P\mathbf{x} \leq \mathbf{x}^T(P_0 + \gamma I)\mathbf{x}$$

So the program can be rewriten as

$$\begin{aligned} \min_{\mathbf{x}} \quad & \frac{1}{2}\mathbf{x}^T(P_0 + \gamma I)\mathbf{x} + q^T\mathbf{x} + r \\ \text{subject to} \quad & A\mathbf{x} \preceq b \end{aligned}$$

The objective function is in the quadratic form to the variable $\mathbf{x}$, and the constrains are in linear form. So above all, the problem is a QP.

(c) We can define $y_i = x^T P_i x$, then we have:

$$\begin{aligned} \sup_{P\in\mathcal{E}} \frac{1}{2}x^T P x &= \sup_{\|\mathbf{u}\|\leq 1} \left(\frac{1}{2}x^T P_0 x + \sum_{i=1}^{K} \frac{1}{2} u_i x^T P_i x\right) \\ &= \frac{1}{2}x^T P_0 x + \frac{1}{2}\sup_{\|\mathbf{u}\|\leq 1}\left(\sum_{i=1}^{K} u_i x^T P_i x\right) \\ &= \frac{1}{2}x^T P_0 x + \frac{1}{2}\sup_{\|\mathbf{u}\|\leq 1}\left(\sum_{i=1}^{K} u_i y_i\right) \qquad (y_i = x^T P_i x) \\ &= \frac{1}{2}x^T P_0 x + \frac{1}{2}\sup_{\|\mathbf{u}\|\leq 1} \mathbf{u}^T\mathbf{y} \\ &= \frac{1}{2}x^T P_0 x + \frac{1}{2}\sup_{\|\mathbf{u}\|\leq 1} \|\mathbf{u}\|_2\|\mathbf{y}\|_2 \qquad \text{(Cauchy Inequality)} \\ &= \frac{1}{2}x^T P_0 x + \frac{1}{2}\|\mathbf{y}\|_2 \end{aligned}$$

So the objective function becomes $\frac{1}{2}x^T P_0 x + \frac{1}{2}\|\mathbf{y}\|_2 + q^T x + r$.
Since there are no suitable programmings for existing norm in the objective function, so we can convert it into the constrains.
Let $u = \frac{1}{2}x^T P_0 x, t = \frac{1}{2}\|y\|_2$.
Then the current simplified problem is that

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y},u,t} \quad & u + t + q^T\mathbf{x} + r \\ \text{subject to} \quad & A\mathbf{x} \preceq b \\ & t = \frac{1}{2}\|y\|_2 \\ & u = \frac{1}{2}x^T P_0 x \\ & y_i = x^T P_i x \quad \forall i = 1, \cdots, K \end{aligned}$$

We could find that the closest form for the problem is the SOCP, but has some difference, so we need to do some conversions.
Since u, t are seperated and independent, and the transmissibility of the inequality, to better suit SOCP, we could do the scalings, which would led to the same result as taking minimum:

$$t = \frac{1}{2}\|y\|_2 \Rightarrow t \geq \frac{1}{2}\|y\|_2$$
$$u = \frac{1}{2}x^T P_0 x \Rightarrow u \geq \frac{1}{2}x^T P_0 x$$
$$y_i = x^T P_i x \Rightarrow y_i \geq x^T P_i x \quad \forall i = 1, \cdots, K$$

Since $P_i \in \mathbf{S}_+^n, i = 0, \ldots, K$, so we could do eigenvalue decomposition to each matrix, which are diagonalizable due to the symmetry. $P_i = Q_i^{-1}\Lambda_i Q_i$, and for all eigenvalues in Λ_i is non-negative, so we have $P_i^{\frac{1}{2}} = Q_i^{-1}\Lambda_i^{\frac{1}{2}}Q_i$.
And construct an inequality:

$$\left\| \begin{bmatrix} P_0^{\frac{1}{2}}x \\ u - \frac{1}{2} \end{bmatrix} \right\|_2 \leq u + \frac{1}{2}$$

If we square to the both side, we can get that

$$\|P_0^{\frac{1}{2}}x\|_2^2 + (u - \frac{1}{2})^2 \leq (u + \frac{1}{2})^2$$
$$x^T P_0 x + u^2 - u + \frac{1}{4} \leq u^2 + u + \frac{1}{4}$$
$$\frac{1}{2}x^T P_0 x \leq u$$

Similarly, in the exactly same way, we can construct

$$\left\| \begin{bmatrix} P_i^{\frac{1}{2}}x \\ y_i - \frac{1}{4} \end{bmatrix} \right\|_2 \leq y_i + \frac{1}{4}$$

If we square to the both side, we can get that

$$\|P_i^{\frac{1}{2}}x\|_2^2 + (y_i - \frac{1}{4})^2 \leq (y_i + \frac{1}{4})^2$$
$$x^T P_i x + y_i^2 - \frac{1}{2}y_i + \frac{1}{16} \leq y_i^2 + \frac{1}{2}y_i + \frac{1}{16}$$
$$x^T P_i x \leq y_i$$

So the program can be rewriten as

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y},u,t} \quad & u + t + q^T\mathbf{x} + r \\ \text{subject to} \quad & A\mathbf{x} \preceq b \\ & \frac{1}{2}\|y\|_2 \leq t \\ & \left\| \begin{bmatrix} P_0^{\frac{1}{2}}x \\ u - \frac{1}{2} \end{bmatrix} \right\|_2 \leq u + \frac{1}{2} \\ & \left\| \begin{bmatrix} P_i^{\frac{1}{2}}x \\ y_i - \frac{1}{4} \end{bmatrix} \right\|_2 \leq y_i + \frac{1}{4} \end{aligned}$$

So above all, the problem is a SOCP.

2. (50 pts) **Water-filling**. Please consider the convex optimization problem and calculate its solution

$$\begin{array}{ll} \text{minimize} & -\sum_{i=1}^{n} \log\left(\alpha_i + x_i\right) \\ \text{subject to} & \mathbf{x} \succeq 0, \quad \mathbf{1}^T\mathbf{x} = 1 \end{array}$$

Solution:

Since $\log x$ is a concave function, so $-\log x$ is a convex funciton, so the objective function is a convex function.
And the constrains are affain constrains.
So we can use $\lambda \in \mathbb{R}^n$ as multipliers for the inequality constrains, and $\nu \in \mathbb{R}$ as multiplier for equality constrain.
So the Lagrangian function is

$$\begin{aligned} L(\mathbf{x}, \lambda, \nu) &= -\sum_{i=1}^{n} \log(\alpha_i + x_i) - \lambda^T\mathbf{x} + \nu(\mathbf{1}^T\mathbf{x} - 1) \\ \nabla_{\mathbf{x}} L(\mathbf{x}, \lambda, \nu) &= -\sum_{i=1}^{n} \frac{1}{\alpha_i + x_i} - \lambda + \nu\mathbf{1} \end{aligned}$$

Since we have the convex objective function, and affain constrains, we the optimal solutions must suit the KKT condition:

$$\begin{cases} x \succeq 0, \quad \mathbf{1}^T x = 1 & (1) & \text{primal feasibility} \\ \lambda \succeq 0 & (2) & \text{dual feasibility} \\ \lambda_i x_i = 0 \quad \forall i = 1, \cdots, n & (3) & \text{complementary slackness} \\ \nabla_{\mathbf{x}} L(\mathbf{x}, \lambda, \nu) = 0 & (4) & \text{zero gradiant of Lagrangian with respect to } \mathbf{x} \end{cases}$$

From (4), we can get that:

$$\forall i = 1, 2, \cdots, n \qquad -\frac{1}{\alpha_i + x_i} - \lambda_i + \nu = 0$$

i.e.

$$x_i = -\alpha_i - \frac{1}{\lambda_i - \nu}$$

From (3), we can get that:
1. from (2), we have $\lambda_i \geq 0$, so

$$x_i = 0 \Leftrightarrow \lambda_i = \nu - \frac{1}{\alpha_i} \geq 0 \Leftrightarrow \nu \geq \frac{1}{\alpha_i}$$

2. from (1), we have $x_i \geq 0$, so

$$x_i \neq 0 \Leftrightarrow \lambda_i = 0 \Leftrightarrow \frac{1}{\nu} = x_i + \alpha_i \geq \alpha_i \qquad \text{<1>}$$

From the domain of the log function, we could get that

$$\alpha_i + x_i > 0 \Leftrightarrow \frac{1}{\nu} > 0 \Leftrightarrow \nu > 0 \qquad \text{<2>}$$

Combine <1> and <2>, we can get that i. if $\alpha_i \leq 0$, then $\nu \geq \frac{1}{\alpha_i}$ always holds, with is the same situation with 1.
ii. if $\alpha_i > 0$, then

$$x_i \neq 0 \Leftrightarrow \nu \leq \frac{1}{\alpha_i}$$

So conclude the information we get from (3), we know that:
1. if $\nu \geq \frac{1}{\alpha_i}$, then $x_i = 0$
2. if $\nu < \frac{1}{\alpha_i}$, then $x_i = \frac{1}{\nu} - \alpha_i \geq 0$
So we could see that $x_i = \max\{\frac{1}{\nu} - \alpha_i, 0\}$
From (1), we could get that

$$\mathbf{1}^T \mathbf{x} = \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \max\{\frac{1}{\nu} - \alpha_i, 0\} = 1$$

Since α_i are fixed constants, so we could calculate ν with the above formula.
So above all, after getting the ν, we could get that the variables to make the optimal solution is that

$$x_i == \max\{\frac{1}{\nu} - \alpha_i, 0\}, i = 1, \cdots, n$$

and the optimal solution for the objective function is that

$$\min \mathrm{obj} = -\sum_{i=1}^{n} \log\left(\alpha_i + \max\{0, \frac{1}{\nu} - \alpha_i\}\right)$$