SI151A
Convex Optimization and its Applications in Information Science, Fall 2024
Homework 2

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

Due on Nov. 25, 2024, 11:59 PM

Read all the instructions below carefully before you start working on the assignment, and before you make a submission.

- You are required to write down all the major steps towards making your conclusions; otherwise you may obtain limited points ($\leq 20\%$) of the problem.
- Write your homework in English; otherwise you will get no points of this homework.
- Do your homework by yourself. Any form of plagiarism will lead to 0 point of this homework. If more than one plagiarisms during the semester are identified, we will prosecute all violations to the fullest extent of the university regulations, including but not limited to failing this course, academic probation, or expulsion from the university.
- If you have any doubts regarding the grading, you need to contact the instructor or the TAs within two days since the grade is announced.

I. **Convex Optimization Problem**

Consider the following compressive sensing problem via ℓ_1-minimization:

$$\begin{aligned} \text{minimize} \quad & \|\boldsymbol{z}\|_1 \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{z} = \boldsymbol{y}, \end{aligned}$$

where $\boldsymbol{A} \in \mathbb{R}^{m \times d}$, $\boldsymbol{z} \in \mathbb{R}^d, \boldsymbol{y} \in \mathbb{R}^m$.

Equivalently reformulate the problem into a linear programming problem. (20 points)

Solution

Let $z_i^+ = \max\{z_i, 0\} \geq 0$, $z_i^- = \max\{-z_i, 0\} \geq 0$, $i = 1, 2, \ldots, d$, where z_i is the i-th component of $\boldsymbol{z}$.
And let $\boldsymbol{z}^+ = \left(z_1^+, \ldots, z_d^+\right)^\top$, $\boldsymbol{z}^- = \left(z_1^-, \ldots, z_d^-\right)^\top$.
Then we have

$$\begin{aligned} z_i = z_i^+ - z_i^-, |z_i| = z_i^+ + z_i^-, \quad & i = 1, 2, \ldots, d \\ z_i^+, z_i^- \geq 0, \quad & i = 1, 2, \ldots, d \end{aligned}$$

So the objective function can be rewritten as

$$\|\boldsymbol{z}\|_1 = \sum_{i=1}^{d} |z_i| = \sum_{i=1}^{d} \left(z_i^+ + z_i^-\right) = \mathbf{1}^\top \left(\boldsymbol{z}^+ + \boldsymbol{z}^-\right) = \mathbf{1}^\top \boldsymbol{z}^+ + \mathbf{1}^\top \boldsymbol{z}^-$$

And the constraint can be rewritten as

$$\boldsymbol{A}\boldsymbol{z} = \boldsymbol{y} \Leftrightarrow \boldsymbol{A}\left(\boldsymbol{z}^+ + \boldsymbol{z}^-\right) = \boldsymbol{y} \Leftrightarrow \boldsymbol{A}\boldsymbol{z}^+ - \boldsymbol{A}\boldsymbol{z}^- = \boldsymbol{y}$$

So above all, the original problem can be equivalently reformulated as a LP problem:

$$\begin{aligned} \min_{\boldsymbol{z}^+, \boldsymbol{z}^- \in \mathbb{R}^d} \quad & \mathbf{1}^\top \boldsymbol{z}^+ + \mathbf{1}^\top \boldsymbol{z}^- \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{z}^+ - \boldsymbol{A}\boldsymbol{z}^- = \boldsymbol{y} \\ & \boldsymbol{z}^+ \succeq 0 \\ & \boldsymbol{z}^- \succeq 0 \end{aligned}$$

II. **Second-order Cone Programming (SOCP)**

Consider the following problem:

$$\min_{x\in\mathbb{R}^n} \|\mathbf{Ax}-\mathbf{b}\|_2+\lambda\|\mathbf{Dx}\|_1,$$

where $\|\cdot\|_p$ is the L_p norm. Equivalently reformulate the problem into a SOCP. (20 points)

Solution

We firstly introduce the auxiliary variable: $y=\|\boldsymbol{Ax}-\boldsymbol{b}\|_2$, to suit the format of SOCP better, it could be relax into $\|\boldsymbol{Ax}-\boldsymbol{b}\|_2\le y$.

And let $\boldsymbol{z}=\boldsymbol{Dx}$, $z_i=(Dx)_i$, where $z_i,(Dx)_i$ is the i-th element of $\boldsymbol{z}$ and $\boldsymbol{Dx}$ respectively.
Suppose that $\boldsymbol{D}\in\mathbb{R}^{m\times n}$, then we have $\boldsymbol{z},\boldsymbol{z}^+,\boldsymbol{z}^-\in\mathbb{R}^m$.
Let $z_i^+=\max\{0,z_i\}\ge 0$, $z_i^-=\max\{0,-z_i\}\ge 0$, $\boldsymbol{z}^+=\left(z_1^+,z_2^+,\ldots,z_n^+\right)^\top$, $\boldsymbol{z}^-=\left(z_1^-,z_2^-,\ldots,z_n^-\right)^\top$.
Then we have $\boldsymbol{z}=\boldsymbol{z}^++\boldsymbol{z}^-$, $\boldsymbol{z}^+\succeq 0$, $\boldsymbol{z}^-\succeq 0$.
i.e. we have

$$\begin{aligned}\|\boldsymbol{Dx}\|_1=\sum_{i=1}^m|z_i|=\sum_{i=1}^m\left(z_i^+-z_i^-\right)=\mathbf{1}^\top\left(\boldsymbol{z}^++\boldsymbol{z}^-\right)\\ \boldsymbol{Dx}=\boldsymbol{z}=\boldsymbol{z}^+-\boldsymbol{z}^-\end{aligned}$$

So above all, we can reformulate the problem into a SOCP:

$$\begin{aligned}\min_{\boldsymbol{x}\in\mathbb{R}^n,y\in\mathbb{R},\boldsymbol{z}^+\in\mathbb{R}^m,\boldsymbol{z}^-\in\mathbb{R}^m}\quad & y+\lambda\mathbf{1}^\top\left(\boldsymbol{z}^++\boldsymbol{z}^-\right)\\ \text{subject to}\quad & \|\boldsymbol{Ax}-\boldsymbol{b}\|_2\le y\\ & 0\le z_i^+,\quad i=1,\ldots,m\\ & 0\le z_i^-,\quad i=1,\ldots,m\\ & \boldsymbol{Dx}-\left(\boldsymbol{z}^+-\boldsymbol{z}^-\right)=\mathbf{0}\end{aligned}$$

III. **Semidefinite Programming (SDP)**

Consider the following eigenvalue optimization problem:

$$\min_{\mathbf{w},\mathbf{S}} \ \lambda_{max}(\mathbf{S}) - \lambda_{min}(\mathbf{S}),$$
$$s.t. \ \ \mathbf{S} = \mathbf{B} - \sum_{i=1}^{k} w_i \mathbf{A}_i,$$

where $\mathbf{B} \in \mathbb{R}^{n\times n}$ and $\mathbf{A}_i \in \mathbb{R}^{n\times n}, i = 1,\ldots,k$ are given symmetric data matrices, $w_i \in \mathbb{R}, i = 1,\ldots,k$ are weights, and $\lambda(\cdot)$ means the eigenvalue. Equivalently reformulate the problem into a SDP. (20 points)

Solution

Since $\boldsymbol{B}, \boldsymbol{A}_i$ are symmetric, then $\boldsymbol{S}$ is symmetric, which means that $\boldsymbol{S}$ has n eigenvalues.
Let $u = \lambda_{\max}(\boldsymbol{S}), v = \lambda_{\min}(\boldsymbol{S})$, then we have

$$v\boldsymbol{I} \preceq \boldsymbol{S} = \mathbf{B} - \sum_{i=1}^{k} w_i \mathbf{A}_i \preceq u\boldsymbol{I}$$

And it could be reformulate as

$$\begin{aligned} v\boldsymbol{I} + \sum_{i=1}^{k} w_i \boldsymbol{A}_i &\preceq \boldsymbol{B} \\ u(-\boldsymbol{I}) + \sum_{i=1}^{k} w_i(-\boldsymbol{A}_i) &\preceq -\boldsymbol{B} \end{aligned}$$

So above all, the problem could be reformulated as a SDP problem:

$$\begin{aligned} \min_{\mathbf{w}\in\mathbb{R}^n, u\in\mathbb{R}, v\in\mathbb{R}} \quad & u - v \\ \text{s.t.} \quad & v\boldsymbol{I} + \sum_{i=1}^{k} w_i \boldsymbol{A}_i \preceq \boldsymbol{B} \\ & u(-\boldsymbol{I}) + \sum_{i=1}^{k} w_i(-\boldsymbol{A}_i) \preceq -\boldsymbol{B} \end{aligned}$$

IV. **Duality**

Derive a dual for the problem

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y}} \quad & -\mathbf{c}^T\mathbf{x} + \sum_{i=1}^m y_i \log y_i \\ \text{s.t.} \quad & \mathbf{P}\mathbf{x} = \mathbf{y} \\ & \mathbf{x} \succeq 0, \quad \mathbf{1}^T\mathbf{x} = 1, \end{aligned}$$

where $\mathbf{P} \in \mathbb{R}^{m\times n}$ has nonnegative elements, and its columns add up to one (i.e., $\mathbf{P}^T\mathbf{1} = \mathbf{1}$). The variables are $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$. (For $c_j = \sum_{i=1}^m p_{ij} \log p_{ij}$, the optimal value is, up to a factor $\log 2$, the negative of the capacity of a discrete memoryless channel with channel transition probability matrix P; see exercise 4.57 in Boyd S, et al. **Convex optimization** (you can find it in blackboard).)

Simplify the dual problem as much as possible.(20 points)

Solution

Let $\boldsymbol{\lambda} \in \mathbb{R}^m$ be the multiplier of the equality constraint $\boldsymbol{P}\boldsymbol{x} = \boldsymbol{y}$, $\boldsymbol{\mu} \in \mathbb{R}^n$ be the multiplier of the inequality constraint $\boldsymbol{x} \succeq 0$, and $\nu \in \mathbb{R}$ be the multiplier of the equality constraint $\mathbf{1}^\top\boldsymbol{x} = 1$.
Since μ is the multiplier of inequality constrain, so $\boldsymbol{\mu} \succeq 0$.
Then Lagrangian of the problem is

$$\begin{aligned} \mathcal{L}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu},\nu) &= -\boldsymbol{c}^\top\boldsymbol{x} + \sum_{i=1}^m y_i \log y_i + \boldsymbol{\lambda}^\top(\boldsymbol{P}\boldsymbol{x} - \boldsymbol{y}) - \boldsymbol{\mu}^\top\boldsymbol{x} + \nu(\mathbf{1}^\top\boldsymbol{x} - 1) \\ &= \left(\boldsymbol{P}^\top\boldsymbol{\lambda} - \boldsymbol{c} - \boldsymbol{\mu} + \nu\mathbf{1}\right)^\top \boldsymbol{x} + \sum_{i=1}^m y_i\left(\log y_i - \lambda_i\right) - \nu \\ &= \phi(\boldsymbol{x}) + \left(\sum_{i=1}^m h_i(y_i)\right) - \nu \end{aligned}$$

We can find that $\mathcal{L}$ is seperable in $\boldsymbol{x}$ and all y_i:

$$\begin{aligned} \phi(\boldsymbol{x}) &= \left(\boldsymbol{P}^\top\boldsymbol{\lambda} - \boldsymbol{c} - \boldsymbol{\mu} + \nu\mathbf{1}\right)^\top \boldsymbol{x} \\ h_i(y_i) &= \sum_{i=1}^m y_i\left(\log y_i - \lambda_i\right) \end{aligned}$$

So for $\boldsymbol{x}$:

$$\phi(\boldsymbol{x}) = \begin{cases} 0 & \text{if } \boldsymbol{P}^\top\boldsymbol{\lambda} - \boldsymbol{c} - \boldsymbol{\mu} + \nu\mathbf{1} = \mathbf{0} \\ -\infty & \text{otherwise} \end{cases}$$

And for each component of $\boldsymbol{y}$:

$$\begin{aligned} h_i(y_i) &= y_i\left(\log y_i - \lambda_i\right) \\ h_i'(y_i) &= 1 + \log y_i - \lambda_i \\ h_i''(y_i) &= \frac{1}{y_i} > 0 \end{aligned}$$

So $h_i(y_i)$ is a convex function. To minimize $h_i(y_i)$, we need to find the point where the first derivative is zero:

$$h_i'(y_i^*) = 0 \Rightarrow y_i^* = e^{\lambda_i - 1} \Rightarrow h_i(y_i)_{\min} = h_i(y_i^*) = -e^{\lambda_i - 1}$$

So to make sure the problem is feasible, we have $\boldsymbol{P}^\top\boldsymbol{\lambda} - \boldsymbol{c} - \boldsymbol{\mu} + \nu\mathbf{1} = \mathbf{0}$, and the objective function of the dual problem is

$$\begin{aligned} g(\boldsymbol{\lambda},\boldsymbol{\mu},\nu) &= \inf_{\boldsymbol{x},\boldsymbol{y}} \mathcal{L}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu},\nu) \\ &= \inf_{\boldsymbol{x}} \phi(\boldsymbol{x}) + \left(\sum_{i=1}^m \inf_{y_i} h_i(y_i)\right) - \nu \\ &= 0 + \left(\sum_{i=1}^m -e^{\lambda_i - 1}\right) - \nu \\ &= -\left(\sum_{i=1}^m e^{\lambda_i - 1}\right) - \nu \end{aligned}$$

To achieve the minimum, we have

$$\left.\begin{array}{l}\boldsymbol{P}^\top\boldsymbol{\lambda}-\boldsymbol{c}-\boldsymbol{\mu}+\nu\mathbf{1}=\mathbf{0}\\ \boldsymbol{\mu}\succeq 0\end{array}\right\}\Rightarrow \boldsymbol{P}^\top\boldsymbol{\lambda}+\nu\mathbf{1}\succeq\boldsymbol{c}$$

So the dual problem is:

$$\begin{aligned}\max_{\boldsymbol{\lambda}\in\mathbb{R}^m,\nu\in\mathbb{R}} \quad & -\left(\sum_{i=1}^m e^{\lambda_i-1}\right)-\nu\\ \text{s.t.}\quad & \boldsymbol{P}^\top\boldsymbol{\lambda}+\nu\mathbf{1}\succeq\boldsymbol{c}\end{aligned}$$

To have a better simplification, we can find that for each constrain about w_i and ν, all these constrains uses one inequality to constrain two variables, so we can relax one of the variables by letting $\boldsymbol{w}=\boldsymbol{\lambda}+\nu\boldsymbol{I}$. Since $\boldsymbol{P}^\top\mathbf{1}=\mathbf{1}$, so

$$\boldsymbol{P}^\top\boldsymbol{\lambda}+\nu\mathbf{1}=\boldsymbol{P}^\top\boldsymbol{\lambda}+\nu\boldsymbol{P}^\top\mathbf{1}=\boldsymbol{P}^\top\left(\boldsymbol{\lambda}+\nu\boldsymbol{I}\right)=\boldsymbol{P}^\top\boldsymbol{w}$$

So the dual problem can be further simplified as:

$$\begin{aligned}\max_{\boldsymbol{w}\in\mathbb{R}^m,v\in\mathbb{R}} \quad & -\left(\sum_{i=1}^m e^{w_i-v-1}\right)-v\\ \text{s.t.}\quad & \boldsymbol{P}^\top\boldsymbol{w}\succeq\boldsymbol{c}\end{aligned}$$

Since their has no constrains on ν, so we can simplify the objective function:
Let $\psi(\nu)=-\left(\sum\limits_{i=1}^m e^{w_i-\nu-1}\right)-\nu$, then

$$\begin{aligned}\psi'(\nu)&=\left(\sum_{i=1}^m e^{w_i-1}\right)e^{-\nu}-1\\ \psi''(\nu)&=-\left(\sum_{i=1}^m e^{w_i-1}\right)e^{-\nu}<0\end{aligned}$$

So $\psi(\nu)$ is a concave function, and the maximum of $\psi(\nu)$ is at the point where the first derivative is zero:

$$\psi'(\nu^*)=0\Rightarrow e^{-\nu^*}=\frac{1}{\sum\limits_{i=1}^m e^{w_i-1}}\Rightarrow \nu^*=\log\left(\sum_{i=1}^m e^{w_i-1}\right)$$

So we have:

$$\psi(\nu)_{\max}=\psi(\nu^*)=-\log\left(\sum_{i=1}^m e^{w_i-1}\right)-1=\log\left(\sum_{i=1}^m\frac{e^{w_i}}{e}\right)-1=-\log\left(\sum_{i=1}^m e^{w_i}\right)-2$$

So the objective function is simplified as getting the maximum of $-\log\left(\sum\limits_{i=1}^m e^{w_i}\right)-2$, which also means that getting the minimum of $\log\left(\sum\limits_{i=1}^m e^{w_i}\right)+2$. So the dual problem can be further simplified as:

$$\begin{aligned}\min_{\boldsymbol{w}\in\mathbb{R}^m} \quad & \log\left(\sum_{i=1}^m e^{w_i}\right)+2\\ \text{s.t.}\quad & \boldsymbol{P}^\top\boldsymbol{w}\succeq\boldsymbol{c}\end{aligned}$$

V. **Convex Problem Applications in Power Allocation**

Consider the following power allocation problem

$$\begin{aligned} \underset{p_1,\ldots,p_K}{\text{maximize}} \quad & \sum_{k=1}^{K} \ln\left(1+\frac{p_k\left|h_k\right|^2}{N_0}\right) \\ \text{subject to} \quad & \sum_{k=1}^{K} p_k = P_{\max} \\ & p_k \geq 0, k=1,\ldots,K \end{aligned}$$

where $N_0 > 0$.

1. Determine that this problem is convex or not, and provide your argument. (5 points)
2. Write down the dual problem. (5 points)
3. Derive the KKT conditions. (5 points)
4. Derive the expression of the optimal solution to the problem above. (5 points)

Solution

1. The problem is convex.

Proof:
<1> For objective function:
Let $f(x) = \ln x$, then $f'(x) = \frac{1}{x}, f''(x) = -\frac{1}{x^2} < 0$. So $f(x)$ is a concave function.
And from the property of composition with affaine function, we known that $1+\frac{p_k|h_k|^2}{N_0}$ is an affain function of p_k, so we have $f\left(1+\frac{p_k|h_k|^2}{N_0}\right)$ is a concave function of p_k.
And from the property of non-negative weighted sum, we have $\sum_{k=1}^{K} f\left(1+\frac{p_k|h_k|^2}{N_0}\right) = \sum_{k=1}^{K} \ln\left(1+\frac{p_k\left|h_k\right|^2}{N_0}\right)$ is a concave function of $(p_1,\ldots,p_K)$.
And maximize a concave function is same as minimize a convex function. So the objective function could be regard as convex.
<2> For constrains:
The first constrain is an equality constrain, and it is an affaine function.
The second constrain is an inequality constrain, and it is obviously a convex function.

So above all, the optimization problem has a convex objective function, convex inequality constrains and affain equality constrains. So the problem is a convex optimization problem.

2. From the analysis above, we can reformulate the problem as

$$\begin{aligned} \underset{p_1,\ldots,p_K}{\text{minimize}} \quad & -\sum_{k=1}^{K} \ln\left(1+\frac{p_k\left|h_k\right|^2}{N_0}\right) \\ \text{subject to} \quad & \sum_{k=1}^{K} p_k = P_{\max} \\ & -p_k \leq 0, k=1,\ldots,K \end{aligned}$$

So let $\lambda \in \mathbb{R}$ be the multiplier of the equality constrain, $\boldsymbol{\mu} \in \mathbb{R}^K$ be the multiplier of the inequality constrains, so $\boldsymbol{\mu} \succeq 0$, and the Lagrangian function is

$$\mathcal{L}(\boldsymbol{p},\lambda,\boldsymbol{\mu}) = -\sum_{k=1}^{K} \ln\left(1+\frac{p_k\left|h_k\right|^2}{N_0}\right) + \lambda\left(\sum_{k=1}^{K} p_k - P_{\max}\right) - \boldsymbol{\mu}^T\boldsymbol{p}$$

The objective function of the dual problem is that

$$g(\lambda,\boldsymbol{\mu}) = \inf_{\boldsymbol{p}} \mathcal{L}(\boldsymbol{p},\lambda,\boldsymbol{\mu})$$

To get $g(\lambda, \boldsymbol{\mu})$, we need to get the minimum of $\mathcal{L}$ over $\boldsymbol{p}$:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial p_k} &= \frac{-\left|h_k\right|^2}{N_0 + p_k \left|h_k\right|^2} + \lambda - \mu_k, \quad k = 1, \dots, K \\
\frac{\partial^2 \mathcal{L}}{\partial p_k^2} &= \frac{\left|h_k\right|^4}{\left(N_0 + p_k \left|h_k\right|^2\right)^2} > 0, \quad k = 1, \dots, K
\end{aligned}$$

So $\mathcal{L}$ is convex in each p_k, and the minimum of $\mathcal{L}$ over $\boldsymbol{p}$ is the point where the derivative of $\mathcal{L}$ over p_k is zero.
So we have

$$\frac{\partial \mathcal{L}}{\partial p_k} = 0 \Rightarrow p_k^* = \frac{1}{\lambda - \mu_k} - \frac{N_0}{\left|h_k\right|^2}, \quad k = 1, \dots, K$$

So

$$g(\lambda, \boldsymbol{\mu}) = \mathcal{L}(p_1^*, \dots, p_K^*, \lambda, \boldsymbol{\mu}) = -\lambda P_{\max} + \sum_{k=1}^{K} \left[-\ln\left(\frac{\left|h_k\right|^2}{N_0(\lambda - \mu_k)} \right) + 1 - \frac{N_0(\lambda - \mu_k)}{\left|h_k\right|^2} \right]$$

So above all, the dual problem is

$$\begin{aligned}
\max_{\lambda \in \mathbb{R}, \boldsymbol{\mu} \in \mathbb{R}^K} \quad & -\lambda P_{\max} + \sum_{k=1}^{K} \left[-\ln\left(\frac{\left|h_k\right|^2}{N_0(\lambda - \mu_k)} \right) + 1 - \frac{N_0(\lambda - \mu_k)}{\left|h_k\right|^2} \right] \\
\text{s.t.} \quad & \boldsymbol{\mu} \succeq 0
\end{aligned}$$

3. Suppose that $f_0(\boldsymbol{p}) = -\sum\limits_{k=1}^{K} \ln\left(1 + \frac{p_k \left|h_k\right|^2}{N_0}\right)$, $f_i(\boldsymbol{p}) = -p_i$, $i = 1, \dots, K$, $h(\boldsymbol{p}) = \left(\sum\limits_{k=1}^{K} p_k\right) - P_{\max}$.
To get the strong duality condition, we have:

$$\begin{aligned}
f_0(\boldsymbol{p}^*) = g(\lambda^*, \boldsymbol{\mu}^*) &= \inf_{\boldsymbol{p}} \mathcal{L}(\boldsymbol{p}, \lambda^*, \boldsymbol{\mu}^*) \\
&= \mathcal{L}(\boldsymbol{p}^*, \lambda^*, \boldsymbol{\mu}^*) \\
&= f_0(\boldsymbol{p}^*) + \lambda^* h(\boldsymbol{p}^*) - \boldsymbol{\mu}^{*T} \boldsymbol{p}^* \\
&\leq f_0(\boldsymbol{p}^*) \qquad (\text{since } h(\boldsymbol{p}^*) = 0,\ f_i(\boldsymbol{p}^*) \leq 0)
\end{aligned}$$

So it must have $\boldsymbol{\mu}^{*T} \boldsymbol{p}^* = 0$.
And since we also get that $\boldsymbol{\mu} \geq 0$, so we have $\mu_k^* p_k^* = 0$, $k = 1, \dots, K$.
So from the analysis above, we can get the KKT conditions:

$$\begin{aligned}
\text{Primal feasibility:} \quad & \begin{cases} \sum\limits_{k=1}^{K} p_k^* = P_{\max} \\ p_k^* \geq 0, \quad k = 1, \dots, K \end{cases} \\
\text{Dual feasibility:} \quad & \boldsymbol{\mu}^* \succeq 0 \\
\text{Complementary slackness:} \quad & \mu_k^* p_k^* = 0, \quad k = 1, \dots, K \\
\text{Stationarity:} \quad & p_k^* = \frac{1}{\lambda^* - \mu_k^*} - \frac{N_0}{\left|h_k\right|^2}, \quad k = 1, \dots, K
\end{aligned}$$

4. With KKT condition, we can get the optimal solution:

<1> If $p_k^* = 0$, then from the stationarity condition, we have $\mu_k^* = \lambda^* - \frac{\left|h_k\right|^2}{N_0}$.

And from the dual feasibility condition, we have $\mu_k^* \geq 0$, so we have $\lambda^* \geq \frac{\left|h_k\right|^2}{N_0}$.

<2> If $p_k^* \neq 0$, then from the complementary slackness condition, we have $\mu_k^* = 0$.

And from the stationarity condition, we have $p_k^* = \frac{1}{\lambda^*} - \frac{N_0}{\left|h_k\right|^2}$.

And from the primal feasibility condition, we have $p_k^* \geq 0$, and since $p_k^* \neq 0$, so we have

$$p_k^* = \frac{1}{\lambda^*} - \frac{N_0}{\left|h_k\right|^2} > 0 \Leftrightarrow \lambda^* < \frac{\left|h_k\right|^2}{N_0}$$

So above all, we can get that

$$p_k^* = \begin{cases} 0, & \lambda^* \geq \dfrac{|h_k|^2}{N_0} \\ \dfrac{1}{\lambda^*} - \dfrac{N_0}{|h_k|^2}, & \lambda^* < \dfrac{|h_k|^2}{N_0} \end{cases}$$

Which means that

$$p_k^* = \max\left\{0, \frac{1}{\lambda^*} - \frac{N_0}{|h_k|^2}\right\}$$

And from the primal feasibility condition, we have

$$\sum_{k=1}^{K} p_k^* = P_{\max}$$

So with some methods, such as the bisection method, we can get the value of λ^* with the above equation.
And after calculating λ^*, we can get the optimal solution of p_k^*.
And the optimal value of the origin problem is

$$\begin{aligned} \text{obj} &= \sum_{k=1}^{K} \ln\left(1 + \frac{p_k^* |h_k|^2}{N_0}\right) \\ &= \sum_{k=1}^{K} \ln\left(1 + \frac{\max\left\{0, \frac{1}{\lambda^*} - \frac{N_0}{|h_k|^2}\right\} |h_k|^2}{N_0}\right) \end{aligned}$$