Homework Set #5 - Solutions

1. The first thing to note is that maximizing the minimum SINR is equivalent to minimizing the maximum of the reciprocal of the SINR. Then, from the description of the optimization problem and constraints mentioned, the SINR maximization problem is as follows:

$$\begin{array}{ll} \text{minimize} & \displaystyle\max_{i=1,\ldots,n} \frac{\displaystyle\sum_{k\neq i} G_{i,k} p_k + N_i}{G_{i,i} p_i} \\ \text{subject to} & 0 \leq p_i \leq P_i^{\max} \\ & \displaystyle\sum_{k \in \mathcal{K}_l} p_k \leq P_l^{\text{gp}} \,,\ l = 1, \ldots, m \\ & \displaystyle\sum_{k=1}^{n} G_{i,k} p_k \leq P_i^{\text{rc}} \,,\ i = 1, \ldots, n \end{array} .$$

To simplify this into the desired generalized linear-fractional program form, let us first define $\mathbf{p} \triangleq \begin{bmatrix} p_1 & \cdots & p_n \end{bmatrix}^T$ to be the optimization variable. Clearly we have $\mathbf{p} \in \mathbb{R}^n$. For the objective, let us define the following quantities for $i = 1, \ldots, n$.

$$\begin{aligned} [\mathbf{c}_i]_k &\triangleq \begin{cases} G_{i,k} \,, & k \neq i \\ 0 \,, & k = i \end{cases} \,, \\ d_i &\triangleq N_i \,, \\ [\mathbf{e}_i]_k &\triangleq \begin{cases} 0 \,, & k \neq i \\ G_{i,i} \,, & k = i \end{cases} \,, \\ f_i &\triangleq 0 \,. \end{aligned}$$

Here, $\mathbf{c}_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$, $\mathbf{e}_i \in \mathbb{R}^n$, and $f_i \in \mathbb{R}$. For the constraints, let us first define the following vectors.

$$\begin{aligned} \mathbf{p}^{\max} &\triangleq \begin{bmatrix} P_1^{\max} & \cdots & P_n^{\max} \end{bmatrix}^T \,, \\ \mathbf{p}^{\text{gp}} &\triangleq \begin{bmatrix} P_1^{\text{gp}} & \cdots & P_m^{\text{gp}} \end{bmatrix}^T \,, \\ \mathbf{p}^{\text{rc}} &\triangleq \begin{bmatrix} P_1^{\text{rc}} & \cdots & P_n^{\text{rc}} \end{bmatrix}^T \,. \end{aligned}$$

Here, $\mathbf{P}^{\max} \in \mathbb{R}^n$, $\mathbf{P}^{\text{gp}} \in \mathbb{R}^m$, and $\mathbf{P}^{\text{rc}} \in \mathbb{R}^n$. To address the group transmitter power constraints, let us define the matrix $\mathbf{K} \in \mathbb{R}^{m \times n}$ as follows:

$$[\mathbf{K}]_{l,k} = \begin{cases} 1 \,, & k \in \mathcal{K}_l \\ 0 \,, & k \notin \mathcal{K}_l \end{cases} \,,\ l = 1, \ldots, m \,,\ k = 1, \ldots, n \,.$$

Then, we can express the SINR maximization problem in the following form.

$$\begin{array}{ll} \text{minimize} & \max\limits_{i=1,\ldots,n} \dfrac{\mathbf{c}_i^T \mathbf{x} + d_i}{\mathbf{e}_i^T \mathbf{x} + f_i} \\ \text{subject to} & \begin{bmatrix} -\mathbf{I}_n \\ \mathbf{I}_n \\ \mathbf{K} \\ \mathbf{G} \end{bmatrix} \mathbf{p} \preceq \begin{bmatrix} \mathbf{0}_{n\times 1} \\ \mathbf{p}^{\max} \\ \mathbf{p}^{\text{gp}} \\ \mathbf{p}^{\text{rc}} \end{bmatrix} \end{array} .$$

Note that this is a generalized linear-fractional program.

2. (a) Here, we have

$$\begin{array}{ll} \text{minimize} & \max\limits_{i=1,\ldots,K} \left\{(1/2)\,\mathbf{x}^T \mathbf{P}_i \mathbf{x} + \mathbf{q}^T \mathbf{x} + r\right\} \\ \text{subject to} & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array} .$$

Note that the objective is the pointwise maximum of a set of convex functions and as such is convex as well. Starting with an epigraph form of the problem, we can express the above problem as

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & (1/2)\,\mathbf{x}^T \mathbf{P}_i \mathbf{x} + \mathbf{q}^T \mathbf{x} + r \le t\,,\ i = 1, \ldots, K \\ & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array} .$$

which is a QCQP in the joint variable $(\mathbf{x}, t)$.

(b) Note that we can express the problem in the form

$$\begin{array}{ll} \text{minimize} & \sup\limits_{\mathbf{P}\in\mathcal{E}} \left\{(1/2)\,\mathbf{x}^T (\mathbf{P} - \mathbf{P}_0)\,\mathbf{x} + (1/2)\,\mathbf{x}^T \mathbf{P}_0 \mathbf{x} + \mathbf{q}^T \mathbf{x} + r\right\} \\ \text{subject to} & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array} ,$$

which is equivalent to

$$\begin{array}{ll} \text{minimize} & (1/2)\,\mathbf{x}^T \mathbf{P}_0 \mathbf{x} + \mathbf{q}^T \mathbf{x} + r + (1/2) \sup\limits_{\mathbf{P}\in\mathcal{E}} \left\{\mathbf{x}^T (\mathbf{P} - \mathbf{P}_0)\,\mathbf{x}\right\} \\ \text{subject to} & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array} .$$

Defining $\Delta\mathbf{P} \triangleq \mathbf{P} - \mathbf{P}_0$, we have $\mathbf{P} \in \mathcal{E}$ if and only if $-\gamma\mathbf{I} \preceq \Delta\mathbf{P} \preceq \gamma\mathbf{I}$. Thus, the problem is equivalent to the following:

$$\begin{array}{ll} \text{minimize} & (1/2)\,\mathbf{x}^T \mathbf{P}_0 \mathbf{x} + \mathbf{q}^T \mathbf{x} + r + (1/2) \sup\limits_{-\gamma\mathbf{I}\preceq\Delta\mathbf{P}\preceq\gamma\mathbf{I}} \left\{\mathbf{x}^T (\Delta\mathbf{P})\,\mathbf{x}\right\} \\ \text{subject to} & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array} .$$

For given $\mathbf{x}$, it is clear that we have

$$\sup_{-\gamma\mathbf{I}\preceq\Delta\mathbf{P}\preceq\gamma\mathbf{I}} \left\{\mathbf{x}^T (\Delta\mathbf{P})\,\mathbf{x}\right\} = \gamma \mathbf{x}^T \mathbf{x}\,.$$

Hence, the robust QP can be expressed as

$$\begin{array}{ll} \text{minimize} & (1/2)\,\mathbf{x}^T (\mathbf{P}_0 + \gamma\mathbf{I})\,\mathbf{x} + \mathbf{q}^T \mathbf{x} + r \\ \text{subject to} & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array} .$$

which is itself a QP.

(c) Note that we can express the problem in the form

$$\begin{array}{ll} \text{minimize} & (1/2)\,\mathbf{x}^T\mathbf{P}_0\mathbf{x} + (1/2)\left(\sup\limits_{||\mathbf{u}||_2 \le 1}\left\{\sum_{i=1}^K u_i\left(\mathbf{x}^T\mathbf{P}_i\mathbf{x}\right)\right\}\right) + \mathbf{q}^T\mathbf{x} + r \\ \text{subject to} & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array} .$$

But recall from our knowledge of dual norms that

$$||\mathbf{x}||_{\circledast} = \sup_{||\mathbf{y}|| \le 1}\left\{\mathrm{Re}\left[\langle \mathbf{x}, \mathbf{y}\rangle\right]\right\} .$$

For the standard inner product and the ℓ_2-norm, we have

$$||\mathbf{x}||_2 = \sup_{||\mathbf{y}|| \le 1}\left\{\mathbf{y}^T\mathbf{x}\right\} .$$

Thus, we have

$$\sup_{||\mathbf{u}||_2 \le 1}\left\{\sum_{i=1}^K u_i\left(\mathbf{x}^T\mathbf{P}_i\mathbf{x}\right)\right\} = \left|\left|\begin{bmatrix} \mathbf{x}^T\mathbf{P}_1\mathbf{x} \\ \vdots \\ \mathbf{x}^T\mathbf{P}_K\mathbf{x} \end{bmatrix}\right|\right|_2 = \sqrt{\sum_{k=1}^K \left(\mathbf{x}^T\mathbf{P}_k\mathbf{x}\right)^2} .$$

Hence, the robust QP has the following form:

$$\begin{array}{ll} \text{minimize} & (1/2)\,\mathbf{x}^T\mathbf{P}_0\mathbf{x} + (1/2)\sqrt{\sum_{k=1}^K \left(\mathbf{x}^T\mathbf{P}_k\mathbf{x}\right)^2} + \mathbf{q}^T\mathbf{x} + r \\ \text{subject to} & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array} .$$

Note that the objective is a convex function of $\mathbf{x}$ for the following reasons. First, each of the functions $\mathbf{x}^T\mathbf{P}_i\mathbf{x}$ is convex since $\mathbf{P}_i \succeq \mathbf{0}$. Then, note that the second term of the objective is a composition $h(g_1(\mathbf{x}), \ldots, g_K(\mathbf{x}))$ of $h(\mathbf{y}) = ||\mathbf{y}||_2$ with $g_i(\mathbf{x}) = \mathbf{x}^T\mathbf{P}_i\mathbf{x}$. As the functions g_i are convex and nonnegative, and the function h is convex and nondecreasing in its arguments for $\mathbf{y} \in \mathbb{R}_+^K$, the composition is convex. Thus, the second term of the objective and hence the objective function itself are convex.
The resulting problem can be expressed as

$$\begin{array}{ll} \text{minimize} & (1/2)\,\mathbf{x}^T\mathbf{P}_0\mathbf{x} + ||\mathbf{y}||_2 + \mathbf{q}^T\mathbf{x} + r \\ \text{subject to} & (1/2)\,\mathbf{x}^T\mathbf{P}_i\mathbf{x} \le y_i\,,\ i = 1, \ldots, K \\ & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array} .$$

This can be further simplified to be expressed as an SOCP. To show this, first note that by using an epigraph form of the problem, we can express it as

$$\begin{array}{ll} \text{minimize} & u + t + \mathbf{q}^T\mathbf{x} + r \\ \text{subject to} & (1/2)\,\mathbf{x}^T\mathbf{P}_0\mathbf{x} \le u\,,\ i = 1, \ldots, K \\ & (1/2)\,\mathbf{x}^T\mathbf{P}_i\mathbf{x} \le y_i\,,\ i = 1, \ldots, K \\ & ||\mathbf{y}||_2 \le t \\ & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array} .$$

In turn, this can be expressed as follows:

$$\begin{array}{ll} \text{minimize} & u + t + \mathbf{q}^T\mathbf{x} + r \\ \text{subject to} & \left|\left|\frac{1}{2}\mathbf{P}_0^{1/2}\mathbf{x}\right|\right|_2^2 \leq \frac{u}{2} \\ & \left|\left|\frac{1}{2}\mathbf{P}_i^{1/2}\mathbf{x}\right|\right|_2^2 \leq \frac{y_i}{2}\,,\ i = 1, \ldots, K \\ & ||\mathbf{y}||_2 \leq t \\ & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array},$$

which is possible since $\mathbf{P}_i \succeq \mathbf{0}$ for all i. Now, using the hint, note that we have

$$\left|\left|\frac{1}{2}\mathbf{P}_0^{1/2}\mathbf{x}\right|\right|_2^2 \leq \frac{u}{2} \Longleftrightarrow \left|\left|\begin{bmatrix} \mathbf{P}_0^{1/2}\mathbf{x} \\ u - \frac{1}{2} \end{bmatrix}\right|\right|_2 \leq u + \frac{1}{2}\,.$$

Similarly, we have

$$\left|\left|\frac{1}{2}\mathbf{P}_i^{1/2}\mathbf{x}\right|\right|_2^2 \leq \frac{y_i}{2} \Longleftrightarrow \left|\left|\begin{bmatrix} \mathbf{P}_i^{1/2}\mathbf{x} \\ y_i - \frac{1}{2} \end{bmatrix}\right|\right|_2 \leq y_i + \frac{1}{2}\,.$$

Therefore, the robust QP can be expressed as follows:

$$\begin{array}{ll} \text{minimize} & u + t + \mathbf{q}^T\mathbf{x} + r \\ \text{subject to} & \left|\left|\begin{bmatrix} \mathbf{P}_0^{1/2}\mathbf{x} \\ u - \frac{1}{2} \end{bmatrix}\right|\right|_2 \leq u + \frac{1}{2} \\ & \left|\left|\begin{bmatrix} \mathbf{P}_i^{1/2}\mathbf{x} \\ y_i - \frac{1}{2} \end{bmatrix}\right|\right|_2 \leq y_i + \frac{1}{2}\,,\ i = 1, \ldots, K \\ & ||\mathbf{y}||_2 \leq t \\ & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array},$$

which (once r is appropriately removed) is an SOCP in the joint variable $(t, u, \mathbf{x}, \mathbf{y}) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^K$.

3. (a) Let $\mathbf{A} \in \mathbb{S}^k$ and $\mathbf{R} \in \mathbb{R}^{k \times k}$ with $\mathbf{R}$ nonsingular. Note that $\mathbf{A} \succeq \mathbf{0}$ if and only if $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^k$. Hence, with $\mathbf{x} = \mathbf{R}\mathbf{y}$, we have $\mathbf{y}^T\mathbf{R}^T\mathbf{A}\mathbf{R}\mathbf{y} \geq 0$ for all $\mathbf{y} \in \mathbb{R}^k$, which means that $\mathbf{R}^T\mathbf{A}\mathbf{R} \succeq \mathbf{0}$. Thus, $\mathbf{A} \succeq \mathbf{0}$ implies $\mathbf{R}^T\mathbf{A}\mathbf{R} \succeq \mathbf{0}$. Similarly, if $\mathbf{R}^T\mathbf{A}\mathbf{R} \succeq \mathbf{0}$, then $\mathbf{y}^T\mathbf{R}^T\mathbf{A}\mathbf{R}\mathbf{y} \geq \mathbf{0}$ for all $\mathbf{y} \in \mathbb{R}^k$. With $\mathbf{y} = \mathbf{R}^{-1}\mathbf{x}$, this yields $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^k$. As such, $\mathbf{R}^T\mathbf{A}\mathbf{R} \succeq \mathbf{0}$ implies $\mathbf{A} \succeq \mathbf{0}$. Combining both results, it follows that $\mathbf{A} \succeq \mathbf{0}$ if and only if $\mathbf{R}^T\mathbf{A}\mathbf{R} \succeq \mathbf{0}$. Setting $\mathbf{A}$ as

$$\mathbf{A} \triangleq x_1\mathbf{F}_1 + \cdots + x_n\mathbf{F}_n\,,$$

yields the equivalence of the two SDPs.

(b) A diagonal matrix is negative semidefinite if and only if its diagonal elements are nonpositive. As such, the constraint of the equivalent SDP becomes the following in this case.

$$x_1\left[\widetilde{\mathbf{F}}_1\right]_{i,i} + \cdots + x_n\left[\widetilde{\mathbf{F}}_n\right]_{i,i} + \left[\widetilde{\mathbf{G}}\right]_{i,i} \leq 0\,,\ i = 1, \ldots, k\,.$$

Thus, if we define the matrix $\mathbf{A} \in \mathbb{R}^{k \times n}$ and vector $\mathbf{b} \in \mathbb{R}^k$ as follows:

$$\begin{aligned} [\mathbf{A}]_{i,j} &\triangleq \left[\widetilde{\mathbf{F}}_j\right]_{i,i}, \; i = 1, \ldots, k, \; j = 1, \ldots, n, \\ [\mathbf{b}]_i &\triangleq -\left[\widetilde{\mathbf{G}}\right]_{i,i}, \; i = 1 \ldots, k, \end{aligned}$$

then the equivalent SDP can be expressed in the following form.

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array} .$$

Note that this is an LP.

(c) If we define the matrix $\mathbf{A} \in \mathbb{R}^{(k-1) \times n}$ and $\boldsymbol{\alpha} \in \mathbb{R}^n$ as

$$\mathbf{A} \triangleq \left[\begin{array}{ccc} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{array} \right], \; \boldsymbol{\alpha} \triangleq \left[\begin{array}{ccc} \alpha_1 & \cdots & \alpha_n \end{array} \right]^T,$$

then the SDP is equivalent to the following one.

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & \left[\begin{array}{cc} \left(\boldsymbol{\alpha}^T \mathbf{x} + \beta\right) \mathbf{I} & (\mathbf{A}\mathbf{x} + \mathbf{b}) \\ (\mathbf{A}\mathbf{x} + \mathbf{b})^T & \left(\boldsymbol{\alpha}^T \mathbf{x} + \beta\right) \end{array} \right] \preceq \mathbf{0} \end{array} .$$

The LMI in this SDP can be equivalently expressed as

$$\left[\begin{array}{cc} \left(-\boldsymbol{\alpha}^T \mathbf{x} - \beta\right) \mathbf{I} & -(\mathbf{A}\mathbf{x} + \mathbf{b}) \\ -(\mathbf{A}\mathbf{x} + \mathbf{b})^T & \left(-\boldsymbol{\alpha}^T \mathbf{x} - \beta\right) \end{array} \right] \succeq \mathbf{0}$$

Using what is known about block matrices and Schur complements, the above condition is true if and only if

$$\left(-\boldsymbol{\alpha}^T \mathbf{x} - \beta\right) \mathbf{I} \succeq \mathbf{0}, \; \left(-\boldsymbol{\alpha}^T \mathbf{x} - \beta\right) - (\mathbf{A}\mathbf{x} + \mathbf{b})^T \left(-\boldsymbol{\alpha}^T \mathbf{x} - \beta\right)^{-1} \mathbf{I} (\mathbf{A}\mathbf{x} + \mathbf{b}) \geq 0.$$

Equivalently, this condition is true if and only if

$$\left(-\boldsymbol{\alpha}^T \mathbf{x} - \beta\right) \geq 0, \; \left(-\boldsymbol{\alpha}^T \mathbf{x} - \beta\right)^2 \geq ||\mathbf{A}\mathbf{x} + \mathbf{b}||_2^2.$$

As the second condition above implies the first, the LMI from above is equivalent to the following:

$$\left(-\boldsymbol{\alpha}^T \mathbf{x} - \beta\right)^2 \geq ||\mathbf{A}\mathbf{x} + \mathbf{b}||_2^2 \iff ||\mathbf{A}\mathbf{x} + \mathbf{b}||_2 \leq -\boldsymbol{\alpha}^T \mathbf{x} - \beta.$$

But this condition is an SOC constraint. Thus, the SDP from above can be expressed as

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & ||\mathbf{A}\mathbf{x} + \mathbf{b}||_2 \leq -\boldsymbol{\alpha}^T \mathbf{x} - \beta \end{array},$$

which is an SOCP with a single SOC constraint.

4. We derive a dual of the problem

$$\begin{array}{ll} \text{minimize} & -\sum_{i=1}^{m} \log y_i \\ \text{subject to} & \mathbf{y} = \mathbf{b} - \mathbf{A}\mathbf{x} \end{array},$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$ has $\mathbf{a}_i$ is its i-th row. The Lagrangian is

$$L(\mathbf{x}, \mathbf{y}, \boldsymbol{\nu}) = -\sum_{i=1}^{m} \log y_i + \boldsymbol{\nu}^T \left(\mathbf{y} - \mathbf{b} + \mathbf{A}\mathbf{x}\right) ,$$

and the dual function is

$$g(\boldsymbol{\nu}) = \inf_{\mathbf{x},\mathbf{y}} \left\{ -\sum_{i=1}^{m} \log y_i + \boldsymbol{\nu}^T \left(\mathbf{y} - \mathbf{b} + \mathbf{A}\mathbf{x}\right) \right\} = \inf_{\mathbf{x},\mathbf{y}} \left\{ \left(\mathbf{A}^T\boldsymbol{\nu}\right)^T \mathbf{x} - \sum_{i=1}^{m} \log y_i + \boldsymbol{\nu}^T\mathbf{y} - \boldsymbol{\nu}^T\mathbf{b} \right\} .$$

The term $\left(\mathbf{A}^T\boldsymbol{\nu}\right)^T \mathbf{x}$ is unbounded below as a function of $\mathbf{x}$ unless $\mathbf{A}^T\boldsymbol{\nu} = \mathbf{0}$. In addition, the terms in $\mathbf{y}$ are unbounded below if $\boldsymbol{\nu} \not\succ \mathbf{0}$ and achieve their minimum for $y_i = \frac{1}{\nu_i}$ otherwise. Thus, the dual function is given by

$$g(\boldsymbol{\nu}) = \begin{cases} \sum_{i=1}^{m} \log \nu_i + m - \mathbf{b}^T\boldsymbol{\nu} , & \mathbf{A}^T\boldsymbol{\nu} = \mathbf{0} ,\ \boldsymbol{\nu} \succ \mathbf{0} \\ -\infty , & \text{otherwise} \end{cases} .$$

Hence, the dual problem is given by

$$\begin{array}{ll} \text{maximize} & \sum_{i=1}^{m} \log \nu_i - \mathbf{b}^T\boldsymbol{\nu} + m \\ \text{subject to} & \mathbf{A}^T\boldsymbol{\nu} = \mathbf{0} \end{array},$$

with domain $\{\boldsymbol{\nu} : \boldsymbol{\nu} \succ \mathbf{0}\}$.

***5.** The capacity is the optimal value of the problem

$$\begin{array}{ll} \text{maximize} & f_0(\mathbf{x}) \triangleq \sum_{i=1}^{m}\sum_{j=1}^{n} x_j p_{i,j} \log \frac{p_{i,j}}{\sum_{k=1}^{n} x_k p_{i,k}} \\ \text{subject to} & \mathbf{x} \succeq \mathbf{0} ,\ \mathbf{1}^T\mathbf{x} = 1 \end{array},$$

with variable $\mathbf{x}$. It is possible to argue directly that the objective f_0 (which is the mutual information between X and Y) is concave in $\mathbf{x}$. This can be done in several ways, including starting from the convexity of the Kullback-Leibler divergence.

Another related approach is to follow the hint given, and introduce $\mathbf{y} = \mathbf{P}\mathbf{x}$ as another variable. We can express the mutual information in terms of $\mathbf{x}$ and $\mathbf{y}$ as

$$
\begin{aligned}
I(X;Y) &= \sum_i \sum_j x_j p_{i,j} \log_2 \frac{p_{i,j}}{\underbrace{\sum_k x_k p_{i,k}}_{y_i}} \,, \\
&= \sum_i \sum_j x_j p_{i,j} \log_2 p_{i,j} - \sum_i \sum_j x_j p_{i,j} \log_2 y_i \,, \\
&= -\sum_j x_j \underbrace{\left(-\sum_i p_{i,j} \log_2 p_{i,j} \right)}_{c_j} - \sum_i \underbrace{\left(\sum_j x_j p_{i,j} \right)}_{y_i} \log_2 y_i \,, \\
&= -\mathbf{c}^T \mathbf{x} - \sum_i y_i \log_2 y_i \,.
\end{aligned}
$$

Therefore, the channel capacity problem can be expressed as

$$
\begin{array}{ll}
\text{maximize} & I(X;Y) = -\mathbf{c}^T \mathbf{x} - \sum_i y_i \log_2 y_i \\
\text{subject to} & \mathbf{x} \succeq \mathbf{0} \,,\ \mathbf{1}^T \mathbf{x} = 1 \\
& \mathbf{y} = \mathbf{P}\mathbf{x}
\end{array} \,,
$$

with variables $\mathbf{x}$ and $\mathbf{y}$. The objective is an affine function of $\mathbf{x}$ (which is both convex and concave in $\mathbf{x}$) plus the entropy of $\mathbf{y}$ (which is concave in $\mathbf{y}$). Thus, the objective function is concave in $(\mathbf{x}, \mathbf{y})$, the inequality constraints are convex in $(\mathbf{x}, \mathbf{y})$, and the equality constraints are affine in $(\mathbf{x}, \mathbf{y})$. As we are maximizing a concave objective over a convex set, this is a convex optimization problem. By minimizing the negative of the objective, we obtain a problem in standard form. Specifically, the channel capacity problem can be expressed as

$$
\begin{array}{ll}
\text{minimize} & \mathbf{c}^T \mathbf{x} + \sum_i y_i \log_2 y_i \\
\text{subject to} & \mathbf{x} \succeq \mathbf{0} \,,\ \mathbf{1}^T \mathbf{x} = 1 \\
& \mathbf{y} = \mathbf{P}\mathbf{x}
\end{array} \,,
$$

which is a convex optimization problem in standard form.