MAT3220: Operation Research
Homework 4

李肖鹏 (116010114)

Due date: March 28, 12 p.m., 2019

Problem 1. Consider the problem of projecting a given point $\vec{v} \in \mathbb{R}^n$ onto an ellipsoid,

$$(P) \quad \min_{\vec{x}} \quad \|\vec{x} - \vec{v}\|_2^2 \qquad s.t. \quad \sum_{i=1}^{n} b_i x_i^2 \leq 1$$

where $b_i > 0$ for $i = 1, 2, \ldots, n$, and $\vec{v}$ is outside the ellipsoid, that is, $\sum_{i=1}^{n} b_i x_i^2 > 1$.

(a) Derive the KKT optimality condition for (P).

The Lagrangian function is defined by

$$L(\vec{x}; \lambda) = \|\vec{x} - \vec{v}\|_2^2 + \lambda \left(\sum_{i=1}^{n} b_i x_i^2 - 1 \right)$$

where $\lambda \geq 0$. For simplicity, define

$$\vec{x}^{\mathrm{T}} D \vec{x} = \sum_{i=1}^{n} b_i x_i^2, \quad \text{where } D = \mathrm{diag}\,(b_1, \ldots, b_n)$$

Thus, the KKT condition is

$$\vec{x}^{\mathrm{T}} D \vec{x} \leq 1, \quad \lambda \geq 0, \quad \lambda \left(\vec{x}^{\mathrm{T}} D \vec{x} - 1 \right) = 0, \quad \vec{x} - \vec{v} + \lambda D \vec{x} = \vec{0}$$

or equivalently,

$$\sum_{i=1}^{n} b_i x_i^2 \leq 1, \quad \lambda \geq 0, \quad \lambda \left(\sum_{i=1}^{n} b_i x_i^2 - 1 \right) = 0, \quad x_i - v_i + \lambda b_i x_i = 0, \ \forall\, i = 1, \ldots, n$$

(b) Derive the Lagrangian dual problem of (P).

First, we derive the dual function. By definition,

$$d(\lambda) = \min_{\vec{x}} L(\vec{x}; \lambda)$$

Notice that $D \succ 0$, and we have

$$L(\vec{x}; \lambda) = \vec{x}^{\mathrm{T}} (I + \lambda D) \vec{x} - 2 \vec{v}^{\mathrm{T}} \vec{x} + \|\vec{v}\|_2^2 - \lambda$$

Since $\lambda \geq 0$, we know $(I + \lambda D) \succ 0$, and the function $L(\vec{x}; \lambda)$ is convex with respect to $\vec{x}$. According to first order sufficient and necessary condition, the global minimum point $\vec{x}^*$ is the point satisfying

$$\nabla_{\vec{x}} L(\vec{x}^*; \lambda) = \vec{0} \Longrightarrow \vec{x}^* - \vec{v} + \lambda D \vec{x}^* = \vec{0}$$

Thus, we can solve the optimal point in closed form, i.e., $\vec{x}^* = (I + \lambda D)^{-1} \vec{v}$. Also, the explicit form of Lagrangian dual function is

$$d(\lambda) = -\vec{v}^{\mathrm{T}} (I + \lambda D)^{-1} \vec{v} + \|\vec{v}\|_2^2 - \lambda = -\sum_{i=1}^{n} \frac{v_i^2}{\lambda b_i + 1} - \lambda + \sum_{i=1}^{n} v_i^2$$

Therefore, the Lagrangian dual problem of (P) is

$$\begin{aligned} \max_{\lambda} \quad & -\sum_{i=1}^{n} \frac{v_i^2}{\lambda b_i + 1} - \lambda + \sum_{i=1}^{n} v_i^2 \\ s.t. \quad & \lambda \geq 0 \end{aligned}$$

(c) Describe a way to solve (P).

Notice that the Lagrangian dual problem is strictly concave and smooth, so it must attain unique global optimal at the point λ^* which satisfies first order necessary (and sufficient) condition, i.e.,

$$\sum_{i=1}^{n} \frac{b_i v_i^2}{(\lambda^* b_i + 1)^2} = 1$$

The function on the left hand side is strictly decreasing for $\lambda \geq 0$, and

$$f(\lambda) = \sum_{i=1}^{n} \frac{b_i v_i^2}{(\lambda b_i + 1)^2}, \quad f(0) = \sum_{i=1}^{n} b_i v_i^2 > 1, \quad \lim_{\lambda \to \infty} f(\lambda) = 0$$

Hence, $f(\lambda) = 1$ has unique solution for $\lambda \geq 0$. In this case, Newton method's can be applied to solve it very efficiently and stably (convergence is ensured by any initial point $\lambda > 0$ because of the convexity of $f(\lambda)$, and $f'(\lambda) \neq 0$).

Problem 2. Consider the following spectrum management problem. Suppose that there are n frequency tones available for the use of communication. The background noises are assumed to be additive Gaussian, and the noise at tone i is $\sigma_i > 0$, $i = 1, 2, \ldots, n$. Suppose that the user wishes to find an allocation of his/her communication powers over the n frequency tones, and denote x_i to be the power allocated to tone i, $i = 1, 2, \ldots, n$. According to the information theory, the information rate on frequency tone i is $\ln(1 + x_i/\sigma_i)$, $i = 1, 2, \ldots, n$. Therefore, the total information rate is $\sum_{i=1}^{n} \ln(1 + x_i/\sigma_i)$. The problem of maximizing the total information rate is to find the optimal power allocation $x_1, x_2, \ldots, x_n$ for the following optimization problem

$$\begin{aligned} \text{(SMP)} \quad \max_{\vec{x}} \quad & f(\vec{x}) = \sum_{i=1}^{n} \ln\left(1 + \frac{x_i}{\sigma_i}\right) \\ s.t. \quad & \sum_{i=1}^{n} x_i \leq P \\ & x_i \geq 0, \; i = 1, 2, \ldots, n \end{aligned}$$

where $\sigma_i > 0$, $i = 1, 2, \ldots, n$, and $P > 0$ are the parameters of this model, and n is the dimension of the model. Solve (SMP) using the KKT optimality condition.

The Lagrange function is defined as

$$L(\vec{x}; \lambda, \boldsymbol{\lambda_i}) = -\sum_{i=1}^{n} \left(1 + \frac{x_i}{\sigma_i}\right) + \lambda \left(\sum_{i=1}^{n} x_i - P\right) + \sum_{i=1}^{n} \lambda_i(-x_i)$$

where $\lambda, \lambda_i \geq 0$ for all $i = 1, \ldots, n$. The KKT condition of SMP is

$$x_i^* \geq 0, \quad \sum_{i=1}^{n} x_i^* \leq P, \quad \lambda \geq 0, \quad \lambda_i \geq 0, \quad \lambda_i x_i^* = 0, \ \forall\, i = 1, \ldots, n$$

$$\lambda \left(\sum_{i=1}^{n} x_i^* - P\right) = 0, \quad -\frac{1}{\sigma_i + x_i^*} + \lambda - \lambda_i = 0, \ \forall\, i = 1, \ldots, n$$

Eliminate λ_i, we obtain

$$x_i^* \geq 0, \quad \sum_{i=1}^{n} x_i^* \leq P, \quad \lambda \geq 0, \quad \lambda \geq \frac{1}{\sigma_i + x_i^*}$$

$$\left(\lambda - \frac{1}{\sigma_i + x_i^*}\right) x_i^* = 0, \quad \lambda \left(\sum_{i=1}^{n} x_i^* - P\right) = 0$$

Since $\sigma_i > 0, x_i^* \geq 0$, we have $\lambda \geq \frac{1}{\sigma_i + x_i^*} > 0$, and hence $\sum_{i=1}^{n} x_i^* = P$. Therefore, the KKT condition is simplied into

$$x_i^* \geq 0, \quad \sum_{i=1}^{n} x_i^* = P, \quad \lambda \geq \frac{1}{\sigma_i + x_i^*}, \quad \left(\lambda - \frac{1}{\sigma_i + x_i^*}\right) x_i^* = 0$$

If $\lambda < \frac{1}{\sigma_i}$, then $x_i^* > 0$, because if not, then $x_i^* = 0$ implies $\lambda \geq \frac{1}{\sigma_i}$ which is a contradiction. Therefore, $\lambda = \frac{1}{\sigma_i + x_i^*}$, and $x_i^* = \frac{1}{\lambda} - \sigma_i$.
If $\lambda \geq \frac{1}{\sigma_i}$, then $x_i^* = 0$, because if not, $x_i^* > 0$, then $\lambda > \frac{1}{\sigma_i + x_i^*}$, but $\left(\lambda - \frac{1}{\sigma_i + x_i^*}\right) x_i = 0$, so again contradiction shows that $x_i^* = 0$.

More compactly, we can write $x_i^* = \max\{0, \lambda^{-1} - \sigma_i\}$. Subsititute x_i^* into $\sum_{i=1}^{n} x_i^* = P$, we finally obtain

$$\sum_{i=1}^{n} \max\{0, \lambda^{-1} - \sigma_i\} = P$$

Notice that on the left hand side is a piecewise linear increasing function with respect to λ^{-1}, so the equation is easy to solve and has unique solution. After we solve λ, we can solve x_i^*.

Problem 3. Let $\mathcal{X}$ be a closed convex set. Suppose that $\vec{v} \notin \mathcal{X}$. Denote the projection of $\vec{v}$ on $\mathcal{X}$ to be $[\vec{v}]_{\mathcal{X}}$, which is essentially the optimal solution of the following problem,

$$\begin{aligned} \min_{\vec{x}} \quad & \frac{1}{2}\|\vec{x} - \vec{v}\|_2^2 \\ s.t. \quad & \vec{x} \in \mathcal{X} \end{aligned}$$

Suppose that $\vec{v}_1$, $\vec{v}_2$ are two arbitrary points that are not in $\mathcal{X}$. Prove the non-expansiveness property of the projection operation

$$\left\|[\vec{v}_1]_{\mathcal{X}} - [\vec{v}_2]_{\mathcal{X}}\right\|_2 \leq \|\vec{v}_1 - \vec{v}_2\|_2$$

Since the projection problem is convex, the necessary and sufficient optimality condtion of it is

$$\nabla f(\vec{x}^*)^{\mathrm{T}}(\vec{x} - \vec{x}^*) \geq 0 \Longrightarrow (\vec{x}^* - \vec{v})^{\mathrm{T}}(\vec{x} - \vec{x}^*) \geq 0, \quad \forall\, \vec{x} \in \mathcal{X}$$

Therefore, for the projection of any points $\vec{v}_1$ and $\vec{v}_2$, we have

$$(\vec{v}_1 - [\vec{v}_1]_{\mathcal{X}})^{\mathrm{T}}(\vec{x} - [\vec{v}_1]_{\mathcal{X}}) \leq 0 \Longrightarrow (\vec{v}_1 - [\vec{v}_1]_{\mathcal{X}})^{\mathrm{T}}([\vec{v}_2]_{\mathcal{X}} - [\vec{v}_1]_{\mathcal{X}}) \leq 0$$

$$(\vec{v}_2 - [\vec{v}_2]_{\mathcal{X}})^{\mathrm{T}}(\vec{x} - [\vec{v}_2]_{\mathcal{X}}) \leq 0 \Longrightarrow (\vec{v}_2 - [\vec{v}_2]_{\mathcal{X}})^{\mathrm{T}}([\vec{v}_1]_{\mathcal{X}} - [\vec{v}_2]_{\mathcal{X}}) \leq 0$$

Add them up, it yields

$$([\vec{v}_1]_{\mathcal{X}} - [\vec{v}_2]_{\mathcal{X}} - (\vec{v}_1 - \vec{v}_2))^{\mathrm{T}}([\vec{v}_1]_{\mathcal{X}} - [\vec{v}_2]_{\mathcal{X}}) \leq 0$$

Thus, we have

$$\begin{aligned}([\vec{v}_1]_{\mathcal{X}} - [\vec{v}_2]_{\mathcal{X}})^{\mathrm{T}}([\vec{v}_1]_{\mathcal{X}} - [\vec{v}_2]_{\mathcal{X}}) &\leq (\vec{v}_1 - \vec{v}_2)^{\mathrm{T}}([\vec{v}_1]_{\mathcal{X}} - [\vec{v}_2]_{\mathcal{X}}) \\ &\leq \left\|[\vec{v}_1]_{\mathcal{X}} - [\vec{v}_2]_{\mathcal{X}}\right\|_2 \|\vec{v}_1 - \vec{v}_2\|_2\end{aligned}$$

Indeed, the left hand side is just $\left\|[\vec{v}_1]_{\mathcal{X}} - [\vec{v}_2]_{\mathcal{X}}\right\|_2^2$, and we obtain

$$\left\|[\vec{v}_1]_{\mathcal{X}} - [\vec{v}_2]_{\mathcal{X}}\right\|_2 \leq \|\vec{v}_1 - \vec{v}_2\|_2$$