Optimization Methods	**Fall 2020**

Homework 3

Instructor: Lijun Zhang — *Name:* Fuyun Wong, *StudentId:* 191300051

Notice

- The submission email is: **opt4grad@163.com**.
- Please use the provided LaTeX file as a template. If you are not familiar with LaTeX, you can also use Word to generate a **PDF** file.

Problem 1: Negative-entropy Regularization

Please show how to compute

$$\operatorname*{argmin}_{x\in\Delta^n} b^\top x + c\cdot\sum_{i=1}^{n} x_i \ln x_i$$

where $\Delta^n = \{x|\sum_{i=1}^n x_i = 1, x_i \geq 0, i = 1,\cdots,n\}$, $b\in\mathbb{R}^n$ and $c\in\mathbb{R}$.

Solution. if $c = 0$, we have

$$\min(b^\top x) = \min\{b_1, b_2, b_3, \ldots, b_n\}$$

.Let's denote the $b_i = \operatorname{argmin}\{b_1, b_2, \ldots, b_n\}$,then it's trivial that the solution of x is $x_i = 1$ and $x_j = 0, \quad j \neq i$.

if $x \neq 0$: Lagrangian:

$$L(x, v, \lambda) = b^\top x + c\cdot\sum_{i=1}^{n} x_i \ln x_i - v^\top x + \lambda(\sum_{i=1}^{n} x_i - 1)$$

because the equality constraint is affine and the inequality constraint is convex. and there do exist a point x_0 (for example $x_i = \frac{1}{n}, \quad i = 1, 2, 3...$). Therefore ,according to the slater's condition the problem has strong duality. then

$$\begin{aligned}
g(v,\lambda) &= \inf((b^\top - v^\top + \lambda 1^T)x - c\sum_{i=1}^{n} x_i \ln x_i) - \lambda \\
&= -c\sup((\frac{v - \lambda\cdot 1 - b}{c})^\top x - \sum_{i=1}^{n} x_i \ln x_i) - \lambda \\
&= -c\sum_{i=1}^{n} e^{\frac{v_i - \lambda - b_i}{c} - 1} - \lambda
\end{aligned}$$

and we have: $v \succeq 0$ and

$$\frac{\partial g}{\partial v} = -e^{\frac{v - \lambda\cdot 1 - b}{c} - 1} \prec 0$$

so, to maximize $g(v, \lambda)$, we let $v = 0$.

Then we have

$$\frac{\partial g}{\partial \lambda} = \sum_{i=1}^{n} e^{\frac{v_i - \lambda - b_i}{c} - 1} - 1 = 0$$

that

$$v^* = 0, \quad \lambda^* = c\ln(\sum_{i=1}^{n} e^{\frac{-b_i - c}{c}})$$

so

$$\frac{\partial(L(x, \lambda^*, v^*)}{\partial x} = b + c(\ln x + 1) - v^* + \lambda^* \cdot 1 = 0$$

.

we have

$$x^* = e^{\frac{v^* - \lambda^* \cdot 1 - b}{c} - 1} = e^{\frac{-\lambda^* \cdot 1 - b}{c} - 1} = e^{-\frac{-c\ln(\sum_{i=1}^{n} e^{\frac{-b_i - c}{c}}) \cdot 1 - b - c}{c}}$$

$\square$

Problem 2: One inequality constraint

(1) With $c \neq 0$, express the dual problem of

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & f(x) \leq 0, \end{aligned}$$

in terms of the conjugate $f^\star$.

(2) Explain why the problem you give is convex. We do not assume f is convex.

Solution. lagrangian:

$$L(\lambda, x) = \inf(c^\top x + \lambda f(x))$$

if $\lambda = 0$

$$g(\lambda) = \inf(c^\top x) = -\infty$$

if $\lambda \neq 0$,

$$\begin{aligned} g(\lambda) &= \inf(L(\lambda, x)) \\ &= \lambda \inf((c/\lambda)^\top x + f(x)) \\ &= -\lambda \inf((-c/\lambda)^\top x - f(x)) \\ &= -\lambda f^*(-c/\lambda) \end{aligned}$$

then the standard form of the problem is

$$\text{minimize} - \lambda f^*(-c/\lambda)$$

$$\text{subjectto} - \lambda \leq 0$$

f^* is always a convex function(pointwise maximum and supremum) and the inequality constrant is convex , and thus the problem is convex.

$\square$

Problem 3: KKT conditions

Consider the problem

$$\begin{aligned}
&\min_{x\in\mathbb{R}^2} x_1^2 + x_2^2 \\
\text{s.t.}\quad &(x_1-1)^2 + (x_2-1)^2 \le 2 \\
&(x_1-1)^2 + (x_2+1)^2 \le 2
\end{aligned}$$

where $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \in \mathbb{R}^2$.

a) Write the Lagrangian for this problem.

Solution.

$$L(x_1, x_2, \lambda_1, \lambda_2) = x_1^2 + x_2^2 + \lambda_1((x_1-1)^2 + (x_2-1)^2 - 2) + \lambda_2((x_1-1)^2 + (x_2+1)^2 - 2)$$

□

b) Does strong duality hold in this problem?

Solution. it's trivial that the objective function is convex and these inequlaity constraint functions are convex, so this problem is covex. The region determined by these two inequalities is the intersection of two disk and we can find a point x_0 in the internal of both of them(not on the border) , which statisfies $(x_1-1)^2 + (x_2-1)^2 - 2 \le 0 \quad (x_1-1)^2 + (x_2+1)^2 - 2 \le 0$ and hence strong duality hold in this problem(according to the slater's condition).

□

c) Write the KKT conditions for this optimization problem.

Solution. the KKT conditions are

$$(x_1-1)^2 + (x_2-1)^2 - 2 \le 0 \quad (x_1-1)^2 + (x_2+1)^2 - 2 \le 0$$

$$\lambda_1 \ge 0 \quad \lambda_2 \ge 0$$

$$2x_1 + 2\lambda_1(x_1-1) + 2\lambda_2(x_1-1) = 0$$

$$2x_2 + 2\lambda_1(x_2-1) + 2\lambda_2(x_2+1) = 0$$

$$\lambda_1((x_1-1)^2 + (x_2-1)^2 - 2) = 0$$

$$\lambda_2((x_1-1)^2 + (x_2+1)^2 - 2) = 0$$

□

Problem 4: Matrix eigenvalues

We denote by $f(A)$ the sum of the largest r eigenvalues of a symmetric matrix $A \in \mathbb{S}^n$ (with $1 \le r \le n$), i.e.,

$$f(A) = \sum_{k=1}^{r} \lambda_k(A),$$

where $\lambda_1(A), \ldots, \lambda_n(A)$ are the eigenvalues of A sorted in decreasing order. Show that the optimal value of the optimization problem

$$\begin{aligned} \max \quad & \boldsymbol{tr}(AX) \\ \text{s.t.} \quad & \boldsymbol{tr} X = r \\ & 0 \preceq X \preceq I, \end{aligned}$$

with variable $X \in \mathbb{S}^n$, is equal to $f(A)$.

Solution. take the dual of the primal problem and we will get

$$\begin{gathered} \min_{Z,s} tr(Z) + ks \\ Z \succeq 0 \\ Z + sI \succeq A \end{gathered}$$

we can prove the optimal value of the problem is $f(A)$ by constructing the optimal points directly.

since A is a real symmetrical problem,take the Eigenvalue composition of A and we can get $A = U\text{diag}(\lambda)U^\top$, where U is an orthonormal matrix of eigenvectors of A and λ is the vector of eigenvalues sorted in decreasing order.

Then let

$$\begin{gathered} X^* = U\text{diag}(1_k, 0_{m-k})U^\top \\ Z^* = U\text{diag}(\lambda - \lambda_k \cdot 1)U^\top \\ s^* = \lambda \end{gathered}$$

where $(\cdot)_+$ is the positive part of a vector.

then we can get the value of the primal problem is

$$\begin{aligned} tr(A^*X) =& tr(U\text{diag}(\lambda)U^\top U\text{diag}(1_k, 0_{m-k})U^\top) \\ =& tr(\text{diag}(diag(1_k, 0_{m-k}))) \\ =& \sum_{i=1}^{k} \lambda_i \\ =& f(A) \end{aligned}$$

and the value of the dual problem is

$$\begin{aligned} tr(Z^*) + ks^* =& k\lambda_k + \sum_{i=1}^{m} (\lambda_i - \lambda_k) \\ =& \sum_{i=1}^{k} ((\lambda_i - \lambda_k) + \lambda_k) \\ =& \sum_{i=1}^{k} \lambda_i \\ =& f(A) \end{aligned}$$

Note that all dual feasible points give a bound for the primal problem and all primal feasible points give a bound for the dual problem. since we have find the points give the same value of dual problem and primal problem, they must be optimal.

so the optimal value of the problem is exact $f(A)$.

$\square$

To solve this problem ,I take the reference:`https://math.stackexchange.com/questions/1044092/sum-of-k-largest-eigenvalues-of-a-symmetric-matrix-as-an-sdp`

Problem 5: Determinant optimization

Derive the dual problem of the following problem

$$\begin{aligned} \min \quad & \log\det X^{-1} \\ \text{s.t.} \quad & A_i^T X A_i \preceq B_i,\ i = 1, \ldots, m \end{aligned}$$

where $X \in \mathbb{S}^n_{++}$, $A_i \in \mathbb{R}^{n\times k_i}$, $B_i \in \mathbb{S}^{k_i}_{++}$, $k_i \in \mathbb{N}_+$, $i = 1, \ldots, m$.

Solution. largangian:

$$L(x, v) = \log\det X^{-1} + \sum_{i=1}^m v_i^\top A_i^\top X A_i v_i - \sum_{i=1}^m v_i^\top B v_i$$

the conjugate function of $f(X) = \log\det X^{-1}$ is $f^*(Y) = \log\det -Y^{-1} - n$

First, we have $\forall y \in \mathbb{R}^n$,we have $y^\top (A_i x_i)(A_i x_i)^\top y = (y^\top A_i x_i)(y^\top A_i x_i) \geq 0$.

so, we know that the matrix$(A_i v_i)(A_i v_i)^\top$ is either a positive definite matrix or a semi-positive definite matrix.

Hence,$when (A_i v_i)(A_i v_i)^\top$ is a positive definite matrix, we have

$$\begin{aligned} g(v) &= \inf(\log\det X^{-1} + \sum_{i=1}^m tr((A_i v_i)(A_i v_i)^\top X)) - \sum_{i=1}^m v_i^\top B v_i \\ &= -\sup(-\sum_{i=1}^m tr((A_i v_i)(A_i v_i)^\top X) - \log\det X^{-1}) - \sum_{i=1}^m v_i^\top B v_i \\ &= -\log\det(\sum_{i=1}^m (A_i v_i)(A_i v_i)^\top)^{-1} - \sum_{i=1}^m v_i^T B v_i + n \\ &= -\log\det(\sum_{i=1}^m (A_i v_i)(A_i v_i)^\top) - \sum_{i=1}^m v_i^T B v_i + n \end{aligned}$$

in the other conditions, $g(y) = 0$. In conclusion, we have

$$g(v) = \begin{cases} -\log\det(\sum_{i=1}^m (A_i v_i)(A_i v_i)^\top) - \sum_{i=1}^m v_i^T B v_i + n, & (A_i v_i)(A_i v_i)^\top \text{ is a positive definite matrix,} \\ -\infty & \text{otherwise} \end{cases}$$

$\square$