MATEMATISKA INSTITUTIONEN
STOCKHOLMS UNIVERSITET
Avd. Matematik

Skiss of the solutions to Exam 2023-01-10

---

1. (i) If S is a convex set, the intersection of S with a line is convex, since the intersection of two convex sets is convex. Conversely, suppose the intersection of S with any line is convex. Take any two distinct points x_1 and x_2 in S. The intersection of S with the line through x_1 and x_2 is convex. Therefore convex combinations of x_1 and x_2 belong to the intersection, thus also to S.

   (ii) See Solutions manual for BSS Problem 2.53. However S is not a polyhedron. And its extreme points are the set ∂S.

   (iii) Since $\langle y, x\rangle$ is a convex function for $y \in X$, $S_X(x)$ is convex by the property of suprium of convex funtions.

   (iv) Obviously $S_C(x) = S_D(x)$ if $C = D$.
   Next we show $D \subseteq C$. Suppose there is a point y in D, $z \notin C$. Since C is closed, y can be strictly separated from C. In other words, there exists an $x \neq 0$ with $\langle x, y\rangle > b$ and $\langle x, z\rangle < b, \forall z \in C$, implying

$$\sup_{z\in C}\langle x, z\rangle \leq b < \langle x, y\rangle \leq \sup_{z\in D}\langle x, z\rangle,$$

   which implies that $S_C(x) \neq S_D(x)$. The proof of the other direction is similar.

   (v) See Example 1 in lecture notes Day 13. 13 p

2. See BSS Example 4.2.10.
   13 p

3. Note there are several ways to find a dual problem. We provide one here. Let $y_i = A_ix + b_i$. Then we have an equality constrained problem. The Lagrange function

$$L(x, \lambda_1, ..., \lambda_N) = \sum_{i=1}^{N} \|y_i\|_2 + \frac{1}{2}\|x - x_0\|_2^2 + \sum_{i=1}^{N} \lambda_i^t(A_ix + b_i - y_i)$$

   is to be minimized over x and y_i, $i = 1, ..., N$. Clearly this minimization problem can be split to minimizing over x and y_i separately. ¿For fixed i we find

$$\inf_{y_i}(\|y_i\|_2 + \lambda_i^t y_i = \begin{cases} 0 & \|\lambda_i\|_2 \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

   whose reasoning is as follows: if $\|\lambda_i\|_2 \leq 1$, which together with the Cauchy-Schwarz, yields that $\|y_i\|_2 + \lambda_i^t y_i \geq 0$ So the minimum is reached at $y_i = 0$. If $\|\lambda_i\|_2 > 1$, we see that the function is unbounded since $y_i = -t\lambda_i$ tends to $-\infty$ as $t \to \infty$.

   Notice that the function $\frac{1}{2}\|x - x_0\|_2^2 + \sum_{i=1}^{N} \lambda_i^t A_i x$ is a convex function, the necessary and sufficient for optimality is its gradient equal to 0 which is $x = x_0 + \sum_{i=1}^{N} A_i^t\lambda_i$. Evaluating the Lagrange function at the optmia just found we get the dual objective function

$$\phi(\lambda_1, ..., \lambda_N) = \begin{cases} \sum_{i=1}^{N}(A_ix_0 + b_i)^t\lambda_i - \frac{1}{2}\|\sum_{i=1}^{N} A_i^t\lambda_i\|_2^2 & \|\lambda_i\|_2 \leq 1, i = 1, ..., N \\ -\infty & \text{otherwise.} \end{cases}$$

   So the dual problem is

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{N}(A_ix_0 + b_i)^t\lambda_i - \frac{1}{2}\left\|\sum_{i=1}^{N} A_i^t\lambda_i\right\|_2^2 \\ \text{subject to} \quad & \|\lambda_i\|_2 \leq 1, i = 1, ..., N. \end{aligned}$$

   13 p

4. See Solutions Manual to BSS Exercise 4.10

13 p

5. It is not an LP problem but a convex program problem because it is a maximum of function convections, and is equivalent to the LP: Minimize t subject to $a_i^t x + b_i \leq t$, $i = 1, ..., m$.

12 p

You have finished the exam if your homework $p_h \geq 24$. Continue otherwise.

6. Let a_i^t be the rows of the matrix A and introduce $y = Ax + b$. Then (GP) is equivalent to

$$\begin{aligned} \text{minimize} \quad & \log\left(\sum_{i=1}^{m} e^{y_i}\right) \\ \text{subject to} \quad & y = Ax + b \end{aligned}$$

Now we determine the dual objective function:

$$g(\lambda) = \min_{x,y}\left(\log\left(\sum_{i=1}^{m} e^{y_i}\right) + \lambda^t(Ax + b - y)\right) = b^t\lambda + \min_x \lambda^t Ax + \min_y\left(\log\left(\sum_{i=1}^{m} e^{y_i}\right) - \lambda^t y\right)$$

Now $\min_x \lambda^t Ax = \begin{cases} 0 & A^t\lambda = 0 \\ -\infty & \text{otherewise} \end{cases}$.

Note that $\log \sum_{i=1}^{m} e^{y_i}$ is a convex function (prove it!) and so is $\log\left(\sum_{i=1}^{m} e^{y_i}\right) - \lambda^t y$. Then the equations

$$\frac{e^{y_k}}{\sum_{i=1}^{m} e^{y_i}} = \lambda_k, k = 1, ..., m$$

is necessary and sufficient for optimality. Note also that these equations are solvable if and only if $\lambda \geq 0$ and $\sum_{i=1}^{m} \lambda_i = 1$. (Prove it!) By substituting the expression of λ_k into $\log\left(\sum_{i=1}^{m} e^{y_i}\right) - \lambda^t y$ we get

$$\min_y\left(\log\left(\sum_{i=1}^{m} e^{y_i}\right) - \lambda^t y\right) = \begin{cases} -\sum_{i=1}^{m} \lambda_i \log \lambda_i & \lambda \geq 0, \sum_{i=1}^{m} \lambda_i = 1 \\ -\infty & \text{otherwise} \end{cases}.$$

Hence the dual objective function is

$$g(\lambda) = \begin{cases} b^t\lambda - \sum_{i=1}^{m} \lambda_i \log \lambda_i & A^t\lambda = 0, \lambda \geq 0, \sum_{i=1}^{m} \lambda_i = 1 \\ -\infty & \text{otherwise} \end{cases}$$

and thus the resulting dual problem is

$$\begin{aligned} \text{maximize} \quad & b^t\lambda - \sum_{i=1}^{m} \lambda_i \log \lambda_i \\ \text{subject to} \quad & A^t\lambda = 0, \lambda \geq 0, \sum_{i=1}^{m} \lambda_i = 1 \end{aligned}$$

As shown in Problem 5 (PWL) is equivalent to the LP problem: Minimize t subject to $a_i^t x + b_i \leq t$, $i = 1, ..., m$. Its dual is

$$\begin{aligned} \text{maximize} \quad & b^t\lambda \\ \text{subject to} \quad & A^t\lambda = 0, \lambda \geq 0, \sum_{i=1}^{m} \lambda_i = 1 \end{aligned}$$

which is identical to the dual of (PWL) which is obtained as follows. The dual objective function is, following the standard procedure,

$$\phi(\lambda) = \inf_{x,y}\left(\max_{i=1,...,m} y_i + \sum_{i=1}^{m} \lambda_i(a_i^t x + b_i - y_i)\right).$$

To simplify we first observe that the infimum over x is finite only if $\sum_{i=1}^m \lambda_i a_i = 0$. To minimize over y we note that

$$\inf_y \left(\max_i y_i - \lambda^t y \right) = \begin{cases} 0 & \lambda \geq 0, \sum_i \lambda_i = 1 \\ -\infty & \text{otherwise} \end{cases}$$

(Prove it!). So

$$\phi(\lambda) = \begin{cases} b^t \lambda & \sum_{i=1}^m \lambda_i a_i = 0, \lambda \geq 0, \sum_{i=1}^m \lambda_i = 1 \\ -\infty & \text{otherwise.} \end{cases}$$

Hence the resulting dual problem is

$$\begin{array}{ll} \text{mazimize} & b^t \lambda \\ \text{subject to} & A^t \lambda = 0, \sum_{i=1}^m \lambda_i = 1, \lambda \geq 0. \end{array}$$

Assume now that λ^* is dual optimal for dual (GP), then λ^* is also feasible for the dual of (PWP), with objective value

$$b^t \lambda = p_{\text{gp}}^* + \sum_{i=1}^m \lambda_i^* \log \lambda_i^*.$$

This yields

$$p_{\text{pwl}}^* \geq p_{\text{gp}} + \sum_{i=1}^m \lambda_i^* \log \lambda_i^* \geq p_{\text{gp}}^* - \log m$$

The last estimate follows from

$$\inf_{\sum_i \lambda_i} \sum_{i=1}^m \lambda_i \log \lambda_i = -\log m.$$

On the other hand we also have

$$\max_i (a_i^t x + b_i) \leq \log \sum \exp(a_i^t x + b_i), \ \forall x.$$

Therefore $p_{\text{pwl}}^* \leq p_{\text{gp}}^*$. Together with the lower bound we get $0 \leq p_{\text{gp}}^* - p_{\text{pwl}}^* \leq \log m$. 12 p

You have finished the exam if your homework $23 \geq p_h \geq 16$. Continue otherwise.

7. (i) It follows by the Jensen's inequality by taking logarithms on $G(x)$.

(ii) First we show that $G(x)$ is concave on $\mathbb{R}_{++}^n$.

A straightforward (a bit tricky) calculation gives the Hessian $\nabla^2 G(x)$ with components

$$\frac{\partial^2 G(x)}{\partial x_k^2} = -(n-1)\frac{G(x)}{n^2 x_k^2}, \quad \frac{\partial^2 G(x)}{\partial x_k \partial x_l} = \frac{G(x)}{n^2 x_k x_l}, \text{ for } k \neq l.$$

We want to show that this matrix is negative semi-definite. Take any $v \neq 0$ in $\mathbb{R}^n$ we have

$$v^t \nabla^2 G(x) v = -\frac{G(x)}{n^2} \left(n \sum_{i=1}^n \frac{v_i^2}{x_i^2} - \left(\sum_{i=1}^n \frac{v_i}{x_i} \right)^2 \right) \leq 0.$$

The last inequality follows from the Cauchy-Schwarz inequality for the vectors $a = (1, ..., 1)^t$ and $b = (\frac{v_1}{x_1}, ..., \frac{v_n}{x_n})^t$.

Since $G(x)$ is concave and $A(x)$ is convex on $\mathbb{R}_{++}^n$. Then $A(x) - G(x)$ is a convex function thus its level set is a convex set, implying that $C = \{x \in \mathbb{R}_{++}^n : G(x) \geq A(x)\}$ is convex. And this set is a cone since for any point $x \in C$ and any positive number α, $G(\alpha x) = \alpha G(x) \geq \alpha A(x) = A(\alpha x)$ and thus $\alpha x \in C$.

12 p

You have finished the exam if your homework $15 \geq p_h \geq 8$. *Continue otherwise.*

8. (i) By adding constant term $\frac{1}{2}\|c\|_2^2$ to the objective function we have an equivalent optimization problem: Minimize $\frac{1}{2}\|c+x\|_2^2$ subject to $Ax = 0$. So the optimal solution is the projection of $-c$ on to the null space of A, which is

$$x^* = -(I - A^t(AA^t)^{-1}A)c$$

(ii) By changing variable $y = Q^{1/2}(x - \bar{x})$ we can write the optimization problem as follows: Minimize $\frac{1}{2}\|y\|_2^2 + (Q^{-1/2}c)^t y$ subject to $AQ^{-1/2}y = 0$. Now apply the result in the previous problem we get

$$y^* = -(I - Q^{-1/2}A^t(AQ^{-1}A^t)^{-1}AQ^{-1/2})Q^{-1/2}c$$

which gives

$$x^* = \bar{x} - Q^{-1}(c - A^t\lambda)$$

where $\lambda = (AQ^{-1}A^t)^{-1}AQ^{-1}c$.

12 p