Delft University of Technology
Faculty of Electrical Engineering, Mathematics, and Computer Science
Circuits and Systems Group

# EE 4530 APPLIED CONVEX OPTIMIZATION

09 April 2021, 09:00–12:20
Block 1 (09:00-10:30)
Open book, strictly timed take-home exam. (Electronic) copies of the book and the course slides allowed. No other tools except a basic pocket calculator permitted.
Upload answers during 10:25–10:35

This block consists of three questions (25 points); more than usual, and this will be taken into account during grading. Answer in English. Make clear in your answer how you reach the final result; the road to the answer is very important. Write your name and student number on each sheet.

*Hint:* Avoid losing too much time on detailed calculations, write down the general approach first.

## Question 1 (8 Points)

For each of the following set or function, establish if it is convex, concave or neither convex nor concave.

(a) $f(X) = \|\text{tr}(AX)Bz - z\|_2^2$, where $\text{tr}(\cdot)$ denotes the trace operator, $A, B, X \in \mathbb{R}^n$ and $z \in \mathbb{R}^n$

(b) $f(x) = e^{\alpha x^\top Ax}$, with $A$ positive definite and $\alpha > 0$.

(c) $f(x) = h(Ax + b)$ where $h(\cdot) : \mathbb{R}^n \to \mathbb{R}$ is a convex function, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$.

(d) The set $\mathcal{C} = \{X \in \mathbb{R}^n | X = \sum_{i=1}^M u_i u_i^\top$, for some $u_i \in \mathbb{R}^n$ and $M \geq 0\}$. Do you recognize which set is it?

(e) The function $f(y) = \sup_x \,(y^\top x - f(x))$, with $f : \mathbb{R}^n \to \mathbb{R}$.

(f) The set $\mathcal{C} = \{x = [x_1, x_2]^\top \in \mathbb{R}^2 | \ \|x - 1\|_2^2 \leq 1, x_2 \geq |x_1|\}$

(g) $f(x) = \min_i \{e^{x_i}\}$

(h) The set $\mathcal{C} = \{x \in \mathbb{R}^n | \sqrt{\|x\|_2} \leq 1\}$.

### Solution

(a) It is convex, since it is the composition of a norm with an affine function on $X$. Indeed $\text{tr}(AX) = \sum_{i,j} A_{ij}X_{ij}$.

(b) It is convex. It is the composition of the function $g(z) = e^{\alpha z}$, which is convex and monotonically increasing over $z \in \mathbb{R}$ and the function $h(x) = \alpha x^\top A x$ is convex over $x \in \mathbb{R}^n$.

(c) By definition, $h(\theta x + (1-\theta)y) \leq \theta h(x) + (1-\theta)h(y)\ \forall\ x, y$ and $\theta \in [0,1]$. Then:

$$\begin{aligned} f(\theta x + (1-\theta)y) &= h(A(\theta x + (1-\theta)y) + b) \\ &= h(\theta A x + (1-\theta)Ay + b) \\ &= h(\theta(Ax+b) + (1-\theta)(Ay+b)) \\ &\leq \theta h(Ax+b) + (1-\theta)h(Ay+b) \\ &= \theta f(x) + (1-\theta) f(y) \end{aligned}$$

Simply saying that a convex function of an affine function is convex is also a correct answer.

(d) It is the set of positive semidefinite matrices, hence it is a convex set.

(e) It is the conjugate function, which is convex, since for a fixed $x$, the function $y^\top x - f(x)$ is convex (affine) in $y$. Thus, for the pointwise supremum operation, the overall function is convex.

(e) It is convex, since it is the intersection between the unit ball centered in the point $(1,1)$ and the values of $x_2$ that are higher than the absolute value function.

(f) Neither convex nor concave. It is the pointwise minimum of convex function.

(g) It is convex. It is the $\alpha$-sublevel set of the quasi-convex function $\sqrt{\|x\|_2}$, for $\alpha = 1$. Alternatively, it is also easy to see that the set is just a 2-norm ball of radius 1.

## Question 2 (9 points)

Consider the following graph filter design problem:

$$\underset{h_0,\dots,h_K}{\operatorname{argmin}} \ \max_n |\hat{h}_n - \sum_{k=0}^{K} h_k \lambda_n^k| \tag{1}$$

where $\hat{h}_0, \dots, \hat{h}_N$ are *given* scalar coefficients, $\lambda_0, \dots, \lambda_N$ are the eigenvalues of a given graph, and $h_0, \dots, h_K$ are the scalar *unknown* filter coefficients.
Define the vectors $\hat{h} := [\hat{h}_1, \dots, \hat{h}_N]^\top$ and $h := [h_0, \dots, h_K]^\top$ and consider the $N \times (K+1)$ Vandermonde matrix

$$V = \begin{pmatrix} \lambda_1^0 & \lambda_1^1 & \cdots & \lambda_1^K \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_N^0 & \lambda_N^1 & \cdots & \lambda_N^K \end{pmatrix}$$

(a) Rewrite problem (1) in vector form in an $\infty$-norm formulation, using $V$, $\hat{h}$ and $h$. Detail the steps to reach such formulation.

(b) Is the problem convex? Why or why not?

(c) Consider solving problem (1) subject to the constraint $\sum_{k=0}^{K} h_k^2 = 1$. Is the problem convex? Why or why not? If not, relax the problem to a convex problem.

(d) Write problem (1) in epigraph form including the relaxed version of the constraint in question (c).

(e) Implement an iterative algorithm to solve it.

### Solution

(a) It can be equivalently rewritten as:

$$\underset{h}{\operatorname{argmin}} \ \|\hat{h} - Vh\|_\infty \tag{2}$$

(b) It is convex, since it is the composition of a norm and an affine function of the unknown vector $h$.

(c) It is not convex, since the equality constraint is not affine. It can be relaxed to $\sum_{k=0}^{K} h_k^2 \leq 1$

(d) Problem (1) can be rewritten as

$$\begin{aligned} \underset{t,h}{\operatorname{argmin}} \quad & t \\ \text{s.t.} \quad & -t\mathbb{1} \preceq \hat{h} - Vh \preceq t\mathbb{1} \\ & \|h\|_2^2 \leq 1. \end{aligned}$$

(e) Note that both $h$ and $t$ are unknown variables and they can be grouped into one unknown $x = [h^T, t]^T$. The problem can then be rewritten as

$$\begin{aligned} \underset{x}{\operatorname{argmin}} \quad & c^T x \\ \text{s.t.} \quad & a_i^T x + \hat{h}_i \preceq 0, \; i = 0, \ldots, N \\ & b_i^T x - \hat{h}_i \preceq 0, \; i = 0, \ldots, N \\ & x^T P x \leq 1, \end{aligned}$$

where $c = [0, \ldots, 0, 1]^T$, $a_i^T$ is the $i$th row of $A = [-V, -1]$, $b_i^T$ is the $i$th row of $B = [V, -1]$, and

$$P = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

A gradient descent algorithm is then easy to derive as

$$x^{(k+1)} = x^{(k)} - \alpha_k g^{(k)},$$

with $\alpha_k$ an appropriate stepsize. As long as the constraints are satisfied, we take $g^{(k)} = c$. If not, pick randomly one of the constraints that is not satisfied and perform a gradient descent step. If the constraint with $a_i$ is not satisfied, take $g^{(k)} = a_i$. If the constraint with $b_i$ is not satisfied, take $g^{(k)} = b_i$. Finally, if the last constraint is not satisfied, you can either project or take $g^{(k)} = 2Px$.

## Question 3 (8 points)

Consider the convex optimization problem

$$\begin{array}{ll} \text{minimize} & a^T x + \sum_{i=1}^{m} \exp y_i \\ \text{subject to} & x \preceq 0 \\ & Bx = y \end{array}$$

with variables $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. The matrix $B \in \mathbb{R}^{m \times n}$ is a wide matrix, i.e., $m < n$, and has full row rank.

(a) Derive the Lagrangian of this problem. Include the domain of the different primal and dual variables.

(b) Give the KKT conditions of this problem.

(c) Derive the dual function of this problem.

(d) What is the dual problem in its simplest form?

(e) Does strong duality hold or not? Explain.

### Solution

(a) Assuming functions are taken element wise, the Lagrangian of this problem is

$$L(x, y, \lambda, \nu) = a^T x + 1^T \exp y + \lambda^T x + \nu^T (Bx - y),$$

where $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $\lambda \in \mathbb{R}^n_+$ and $\nu \in \mathbb{R}^m$.

(b) The first set of KKT conditions corresponds to the primal constraints. The second one is $\lambda \succeq 0$. The third set of KKT conditions corresponds to the complementary slackness conditions given by $\lambda_i x_i = 0$ for $i = 1, 2, \ldots, n$. And finally, the fourth set of KKT conditions equates the gradient of the Lagrangian with respect to $x$ and $y$ to zero:

$$\frac{\partial L}{\partial x} = a + \lambda + B^T \nu = 0 \;\Rightarrow\; \lambda = -a - B^T \nu, \tag{3}$$

$$\frac{\partial L}{\partial y} = \exp y - \nu = 0 \;\Rightarrow\; y = \log \nu. \tag{4}$$

(c) It is clear that the minimum of the Lagrangian is $-\infty$ when (3) is not satisfied. If (3) is satisfied, the linear term in $x$ of the Lagrangian is always 0 and the minimum for $y$ is reached at (4). This gives the following dual function:

$$g(\lambda, \nu) = \begin{cases} 1^T \nu - \nu^T \log \nu & \text{if } \lambda = -a - B^T \nu \\ -\infty & \text{otherwise} \end{cases}$$

(d) For the dual problem we have to maximize this dual function under the constraint that $\lambda \succeq 0$. So we can easily eliminate $\lambda$ in the dual problem, leading to

$$\underset{\nu}{\text{maximize}} \quad g(\nu),$$

where

$$g(\nu) = \begin{cases} 1^T\nu - \nu^T \log \nu & \text{if } a + B^T\nu \preceq 0 \\ -\infty & \text{otherwise} \end{cases}$$

So we obtain the dual problem

$$\begin{aligned} \underset{\nu}{\text{maximize}} \quad & 1^T\nu - \nu^T \log \nu \\ \text{subject to} \quad & a + B^T\nu \preceq 0 \end{aligned}$$

(e) Since the problem is convex, we can check Slater's condition. If that condition is satisfied, there is strong duality. This means that there should exist an $x \prec 0$ that is feasible. Since the cost function has no domain restrictions, and since for any $x$ we can always find a $y$ for which $Bx = y$, any $x \prec 0$ is feasible and hence the problem always has strong duality.