Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute for Operations Research
Department of Mathematics

Convex Optimization

Exam

June 15$^{\text{th}}$, 2022

Student ID:	

General remarks.

- Place your student identity card on the table.
- This is a closed book exam. No extra material is allowed (except for a subject-neutral dictionary).
- Use a pen to write down your solutions. Pencils and red pens are not allowed.
- Turn off all your technical devices and put them away.
- Write your solutions directly after each exercise, if you need more space use the extra sheets at the end, clearly mark where the solution continues. Extra paper will be distributed if needed.
- You have 3 hours to solve the exam. 15 minutes before the end of the exam, no premature submission is allowed anymore.
- For all problems, provide a complete solution in English, including all explanations and implications in a mathematically clear and well-structured way. Please hand in a readable and clean solution. Cross out any invalid solution attempts.
- Unless you are explicitly asked to prove them, you may use results from the lecture without proof (provided that you reference or state them clearly). Results from the problem sets may only be used if you prove them.
- Ask any question that you might have immediately and during the exam.

Good luck!

Problem:	1	2	3	4	5	6	7	Total
Points:	/ 10	/ 6	/ 10	/ 10	/ 8	/ 10	/ 10	/ 64

Exercise 1 (10 Points)

For each of the following statements decide whether they are true (T) or false (F). If a question has multiple statements, points are distributed uniformly. Wrong answers lead to a deduction of the corresponding points while blank answers do not change the score. The minimum number of points achievable for Exercise 1 is zero. No negative points will be carried over.

(2 Points) For each of the following functions decide whether they are convex (T) or not (F).

T	F	
☐	☐	$f_1 \colon \mathbb{R}^{n\times n} \to \mathbb{R}_{>0},\ f_1(A) = \det(A)$ for $n \in \mathbb{N}_{\geq 2}$.
☐	☐	$f_2 \colon \mathbb{R}^n_{\geq 0} \to \mathbb{R},\ f_2(x) = \log\left(\sum_{k=1}^{K} c_k \exp\left(x^\top a_k\right)\right)$ for $K, n \in \mathbb{N}_{\geq 1}$ and $c_k > 0, a_k \in \mathbb{R}^n$ for every $k \in [K]$.

(2 Points) For each of the following statement decide whether they are true (T) or false (F).

T	F	
☐	☐	$f_1 \colon \mathbb{R}_{>0} \to \mathbb{R}_{>0},\ f_1(x) = cx$ is log-log convex for $c > 0$.
☐	☐	$f_2 \colon \mathbb{R}_{>0} \to \mathbb{R}_{>0},\ f_2(x) = x + 1$ is log-log concave for $c > 0$.
☐	☐	$f_3 \colon \mathbb{R}_{>0} \to \mathbb{R}_{>0},\ f_3(x) = \sqrt{x}$ is log-log affine.
☐	☐	$f_4 \colon \mathbb{R}_{>0} \to \mathbb{R}_{>0},\ f_4(x) = e^x$ is log-log affine.

(2 Points) Consider the following minimization problem

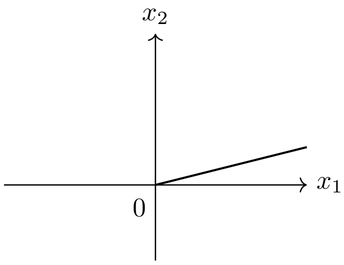
$$\begin{aligned} &\min\ x^2 \\ &\text{s.t. } x \leq 0. \end{aligned}$$

which of the following points are in a central path? Select (T) if there exists a central path containing the point and (F) otherwise.

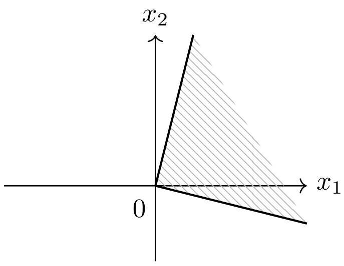
T	F	
☐	☐	$x_1 = -1$
☐	☐	$x_2 = 1/4$
☐	☐	$x_3 = -1/4$
☐	☐	$x_4 = 3$

(2 Points) Consider the following closed sets in $\mathbb{R}^2$. For each of them decide whether the set is a proper cone (T) or not (F). The sets continue in the intuitive way once they are outside the plot.

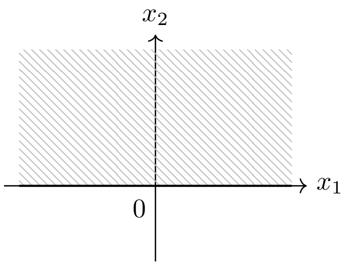
T **F**



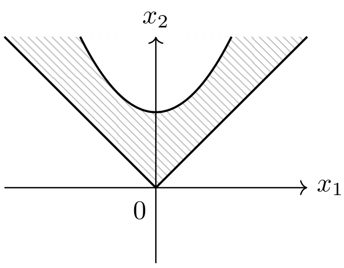
☐ ☐



☐ ☐



☐ ☐



☐ ☐

(2 Points) In the first constraint the constant $+1$ was corrected to -15 to make the problem feasible, true/false questions 1 and 3 were corrected accordingly. Consider the following minimization problem

$$
\begin{aligned}
&\min\ x^2 \\
&\text{s.t. } \tilde{x}^2 - 2\tilde{x} - 15 \leq 0 \\
&\qquad 2x + 5 = 0
\end{aligned}
$$

for which strong duality holds. Let $\tilde{x}$ and $\left(\tilde{\lambda}, \tilde{\nu}\right)$, be optimal primal and dual solutions, respectively. Which of the following conditions are satisfied by the point $(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$.

T	**F**	
☐	☐	$\tilde{x}^2 - 2\tilde{x} - 15 \leq 0$
☐	☐	$\tilde{\lambda} < 0$
☐	☐	$\tilde{\lambda}\left(\tilde{x}^2 - 2\tilde{x} - 15\right) > 0$
☐	☐	$2\tilde{x}\left(1 + \tilde{\lambda}\right) + 2\left(\tilde{\nu} - \tilde{\lambda}\right) = 0$

Exercise 2 (6 Points)

Give a short proof or counterexample for the following statements. If you give a counterexample you also have to show how it contradicts the statement.

(a) **(3 Points)** Let $U, V \subseteq \mathbb{R}$ be convex sets and $f\colon U \to V$ be a bijective, convex and strictly decreasing function. Then $f^{-1}\colon V \to U$ is convex as well.

(b) **(3 Points)** Every convex program has the same optimal value as its dual.

Exercise 3 (10 Points)

A mechanic wants to design a gear system and needs your help to do so. He wants to place a fixed number n of gears into a given box. Any two successive gears have to touch. For the first and the final gear the position is given by the connection of the system to the exterior.

To formalize: We model the gear wheels by circles. You are given a number n and bounds $A, B \in \mathbb{R}_{\geq 0}$, coordinates (a_1, b_1) and (a_n, b_n) for the first and the final circle. You want to place a circle of variable radius r_i for each $i \in \{1, \dots, n\}$ - with two positions given - such that the i-th circle touches the $i+1$st one externally. Furthermore the circles have to lie completely in the axis parallel box defined by $(0,0)$ and (A, B). As a slight relaxation we do not place restrictions on the corresponding circles for i, j with $i + 1 < j$ - else the problem would be non-convex. In particular they might intersect.

The material cost of the gears is proportional to their size so you want to minimize the total area of the circles you placed.

a) **(7 Points)** First consider the relaxation in which successive 'gears' do not need to touch (intersect in one point) but instead are only required to intersect arbitrarily. Solve this using a convex optimization problem. Proof that the problem you give is convex and how it solves the given problem.

b) **(3 Points)** Now show that for an optimal solution to the relaxed problem described in a) the successive circles are indeed touching each other. So you can used the program you designed in a) to solve the original problem.

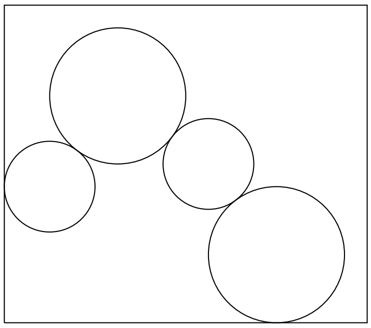


Figure 1: Example of four 'gear wheels' such that two consecutive ones touch and all of them fit into the given box.

Exercise 4 (10 Points)

Consider the following illumination problem: In a room of height h there is a single lamp at the ceiling. Your task is to find the best positions (x_i, y_i) of n work places in the room. For simplicity assume that the places are points on the floor. Additionally you can choose the power p of the lamp which we want to minimize.

The work places should not be to close to each other. To ensure this you are given for some pairs of places (i, j) restriction that x_i has to have to be at least $a_{i,j}$ less then x_j and/or that y_i has to be $b_{i,j}$ less than y_j.

Finally for each work space there is a target illumination T_i you have to achieve. The illumination at any point is proportional to the power p of the lamp as well as anti-proportional to the square of the distance d_i of the point to the lamp. Note that unlike in the problem set we do not include any angles!

a) **(5 Points)** Formulate the problem of choosing a placement of the work places and the power level, such that all placement and illumination constraints are satisfied, as a convex optimization problem. The objective of the problem is to minimize the needed power p of the lamp. Show that the program is indeed convex and that it models the given task.

b) **(5 Points)** Consider the adaptation in which we are given more than one lamp. The illumination level at a point is then the sum of the illumination we get from every lamp. Show that the set of feasible desk positions and lamp powers (seen as a real vector) is not always convex.

Exercise 5 (8 Points)

Let $n \in \mathbb{N}_{\geq 1}$ and $f\colon \mathbb{R}^n \to \mathbb{R}$. Then the *convex hull* $g\colon \mathbb{R}^n \to \mathbb{R} \cup \{ -\infty \}$ of f is defined by

$$g(x) := \inf \{ t \in \mathbb{R} \mid (x,t) \in \operatorname{conv}(\operatorname{epi} f) \} .$$

You can use the fact that geometrically the epigraph of g is the convex hull of the epigraph of f.
Show that g is convex and if $h\colon \mathbb{R}^n \to \mathbb{R}$ is convex and satisfies $h(x) \leq f(x)$ for all $x \in \mathbb{R}^n$, then $h(x) \leq g(x)$ for all $x \in \mathbb{R}^n$.
Remark. This function g is called the largest convex underestimator of f.

Exercise 6 (10 Points)

Consider a twice differentiable function $f\colon \mathbb{R}^n \to \mathbb{R}$ with $\nabla^2 f(x)$ positive definite for all $x \in \mathbb{R}^n$. Furthermore assume that f has bounded Newton decrement, i.e. there is a positive c such that $\lambda(x)^2 \leq c$ for all $x \in \mathbb{R}^n$.
Show that then the function $g\colon \mathbb{R}^n \to \mathbb{R},\ g(x) = e^{\frac{-f(x)}{c}}$ is concave.

Exercise 7 (10 Points)

For a given matrix $W = W^\top \succ 0$ ($W \in \mathbb{R}^{n\times n}$ is positive definite) consider the problem

$$\begin{aligned} f^* := \min_x \quad & -x^\top W x \\ s.t. \quad & x_i^2 \le 1 \qquad i \in [n] \end{aligned}$$

(a) **(1 Point)** Decide and prove whether this problem is a convex optimization problem or not.

(b) **(2 Points)** Write a Lagrangian for this problem.

(c) **(3 Points)** Show that the dual function can be written as

$$\begin{aligned} \widehat{L}(\lambda) = \max \quad & t \\ & t \le -Tr\left(D_\lambda\right) + x^\top \left(D_\lambda - W\right) x \qquad \text{for all } x \end{aligned}$$

where $D_\lambda := \operatorname{diag}(\lambda)$.

(d) **(4 Points)** Derive an explicit dual program in the form of an SDP

Hint: One line of attack is to rewrite the constraint

$$t \le -Tr\left(D_\lambda\right) + x^\top \left(D_\lambda - W\right) x \qquad \text{for all } x$$

in a matrix positive semidefinite form.

Extra sheets you may need

Exercise 1 Solution (10 Points)

(a) Determinant is not convex as it can be seen by e.g. $A = \operatorname{diag}(1, 1, \ldots, 1, 0), B = \operatorname{diag}(0, 1, 1, \ldots, 1)$. f_2 on the other hand is convex: in the lecture it was shown that any posynomial function is log-log convex and f_2 is the log-log transform of a posynomial. This can also be derived by first noting that the sum of log-convex functions is still log-convex. Therefore it suffices to show that $\log\left(c \exp\left(a^\top x\right)\right)$ is log-convex, which trivially holds as it is affine.

(b) The log-log transform of cx is $x + \log c$ so it is even log-log affine. $x + 1$ transforms to $\log(e^x + 1)$ with first derivative $\frac{e^x}{e^x+1}$ and as second $\frac{1}{(e^x+1)^2}$ so it is log-log convex but not log-log concave. $\sqrt{x}$ is log-log affine as its transform is $0.5x$ and finally e^x has as transform e^x so it is not log-log affine.

(c) The central path is defined as the set of minimizers of

$$\min x^2 - \frac{1}{t} \log(-x)$$

for some $t > 0$. Hence they are the solutions of

$$2x - \frac{1}{tx} = 0$$

which are given by $x_t^* = -\sqrt{\frac{1}{2t}}$. This shows that there exists a central path consisting of all negative real numbers. The positive options are not feasible and hence not in a central path.

(d) The first cone is not solid and hence not proper. The second cone is proper. The third cone is not pointed and hence not proper and the last set is not a cone.

(e) The objective and constraints are differentiable, strong duality holds, so every primal and dual optimal solution forms a KKT-point, that satisfied KKT-conditions. Thus the first statement is true (KKT-1), second is false (KKT-2 holds which is $\tilde{\lambda} \geq 0$), third is false (KKT-4 holds which is $\tilde{\lambda}\left(\tilde{x}^2 - 2\tilde{x} + 1\right) = 0$) and the fourth one holds (KKT-5).

Exercise 2 Solution (6 Points)

(a) *Proof.* Since V is convex by definition it suffices to show $f^{-1}(\lambda x + (1-\lambda)\, y) \leq \lambda f^{-1}(x) + (1-\lambda)\, f^{-1}(y)$ for all $\lambda \in [0, 1],\ x, y \in V$. Denote $a = f^{-1}(x), b = f^{-1}(y)$ and calculate

$$\begin{aligned} f\left(f^{-1}\left(\lambda x + (1-\lambda)\, y\right)\right) &= \lambda x + (1-\lambda)\, y \\ &= \lambda f(a) + (1-\lambda)\, f(b) \\ &\geq f\left(\lambda a + (1-\lambda)\, b\right) \end{aligned}$$

. Since f is strictly decreasing this implies $f^{-1}(\lambda x + (1-\lambda)\, y) \leq \lambda f^{-1}(x) + (1-\lambda)\, f^{-1}(y)$.

(b) *Counterexample.* Let $\mathbb{P} := [0, 1], f\colon \mathbb{P} \to \mathbb{R}$ where $f(0) = 1$ and $f(x) = 0$ for $x \in (0, 1]$. This f is clearly convex and the convex problem

$$\begin{aligned} &\min_{x \in \mathbb{P}} f(x) \\ &\text{s.t. } x \leq 0 \end{aligned}$$

has the optimal solution 1. The lagrange dual function is given by

$$\widehat{L}(\lambda) = \inf_{x \in \mathbb{P}} f(x) + \lambda x = 0$$

for $\lambda \geq 0$ and hence strong duality does not hold for this problem .
Full points are given for any other counterexample with proof.

Exercise 3 Solution (10 Points)

a) We model the constraints one by one. First each circle has to lie in the box. This gives for each i four constraints, one per side of the box , which are linear and hence convex :

$$x_i - r_i \geq 0, \quad x_i + r_i \leq A \quad y_i - r_i \geq 0, \quad y_i + r_i \leq B.$$

Next we consider intersection constraints. Two circles intersect iff the distance of their centers is at most the sum of the radii. So we get the constraint

$$\|(x_i, y_i) - (x_{i+1}, y_{i+1})\| \leq r_i + r_{i+1}.$$

This is a second order cone constraint and as such convex . Then finally considering the objective. The area of a circle is proportional to its radius squared so we minimize the sum of r_i^2 which is quadratic and hence convex .

So all together we get the following program:

$$\begin{aligned}
\min \quad & \sum_{i=1}^{n} r_i^2 \\
s.t. \quad & x_i - r_i \geq 0, \quad x_i + r_i \leq A && \forall\ 1 \leq i \leq n, \\
& y_i - r_i \geq 0, \quad y_i + r_i \leq B && \forall\ 1 \leq i \leq n, \\
& \|(x_i, y_i) - (x_{i+1}, y_{i+1})\| \leq r_i + r_{i+1} && \forall\ 1 \leq i \leq n-1, \\
& x_1 = a_1, y_1 = b_1, x_n = a_n, y_n = b_n.
\end{aligned}$$

The final line is optional, the students can also directly use variables a_i, b_i and for 1 and n these are parameter instead but in some kind they have to integrate the given coordinates .

b) Let x_i, y_i, r_i denote an optimal solution, and C_i the corresponding circles. Now assume there was an i such that C_i and C_{i+1} intersect in more than one point. This is equivalent to $\|(x_i, y_i) - (x_{i+1}, y_{i+1})\| < r_i + r_{i+1}$. If $i = 1$ or $i + 1 = n$ then we can reduce the the radius r_i respectively r_{i+1} by an ε and stay feasible. As we are at an optimal solution this is a contradiction.

Let $2\varepsilon < r_i + r_{i+1} - \|(x_i, y_i) - (x_{i+1}, y_{i+1})\|$. Then we claim that moving (x_i, y_i) an ε amount in the direction of (x_{i-1}, y_{i-1}) and reducing r_i by ε gives a feasible solution. Denote the moved circle as B_i and note that B_i and C_{i+1} still intersect as we increase the lhs - triangle inequality - and decrease the rhs of

$$\|(x_i, y_i) - (x_{i+1}, y_{i+1})\| \leq r_i + r_{i+1}.$$

each by at most ε, so by choice of ε the inequality still holds. Next B_i and B_{i-1} still intersect as we decrease both sides by ε. Finally none of the border constraints is violated as we get at most ε closer to it and reduce the radius by the same amount finishing the proof of the claim.

Now we see that also in this case there is another feasible point that agrees in all r_j and has smaller r_i. This contradicts optimality, finishing the proof.

Exercise 4 Solution (10 Points)

a) Assume the lamp to be at $(0, 0)$ else shift the problem. First we model the relative positional requirements. Let L (for left) be the set of ordered pairs i, j for which there is a constraint on the x values and similarly B (below) the set for the y constraints. Then the first constraints are :

$$\begin{aligned}
x_i + a_{i,j} \leq x_j \text{ for } (i, j) \in L \\
y_i + b_{i,j} \leq y_j \text{ for } (i, j) \in B
\end{aligned}$$

These are linear and thus convex constraints . For the illumination constraint we have to compute the distance d_i of (x_i, y_i) to the lamp. This can be done with pythagoras as we get a rectangular triangle between the lamp at $(0, 0, h)$, the origin $(0, 0, 0)$ and the point $(x_i, y_i, 0)$. So $d_i^2 = h^2+x_i^2+y_i^2$. The illumination constraint in its naive form is thus

$$T_i \leq \frac{p}{h^2 + x_i^2 + y_i^2}.$$

To transform it to a convex constraint we multiply with the denominator

$$T_i(h^2 + x_i^2 + y_i^2) \leq p$$

and subtract p

$$T_i(h^2 + x_i^2 + y_i^2) - p \leq 0$$

. The left hand side is convex as it is linear in p and positive quadratic in x_i and y_i. Alternatively we see that its Hessian is diagonal with entries $2T_i$, $2T_i$ and 0.

So alltogether the convex program is

$$\begin{array}{ll} \min & p \\ s.t. & x_i + a_{i,j} \leq x_j \text{ for } (i, j) \in L \\ & y_i + b_{i,j} \leq y_j \text{ for } (i, j) \in B \\ & T_i(h^2 + x_i^2 + y_i^2) - p \leq 0 \end{array}$$

b) Consider two lamps. Without loss of generality we can assume that they are placed at $(0, 0)$ and $(1, 0)$, else we can rotate and scale the situation. Furthemore we consider the special case of a single work place to be positioned and for this one only the positions between the two lamps. Assume it is placed at $(x, 0)$, then its illumination level is

$$f(x) = \frac{p_1}{h^2 + x^2} + \frac{p_2}{h^2 + (1 - x^2)}$$

where p_1 and p_2 are the power levels of the two lamps. For $p_1 = p_2 = 1$ we have

$$f(x) = \frac{1}{h^2 + x^2} + \frac{1}{h^2 + (1 - x^2)}.$$

We want to show that for the right choice of target illumination the two points under the lamps are feasible but the middle point $(\frac{1}{2}, 0)$ is not. For this observe that

$$\begin{aligned} f(0) = f(1) &= \frac{1}{h^2} + \frac{1}{h^2 + 1} \\ f(\frac{1}{2}) = \frac{2}{h^2 + \frac{1}{4}} &= \frac{8}{4h^2 + 1} \leq 8. \end{aligned}$$

So if we choose h small enough, such that $\frac{1}{h^2} \geq 10$, then a target illumination of 9 separates between these points so $(0, 0)$ and $(1, 0)$ are feasible but $(\frac{1}{2}, 0)$ is not. So the set of feasible points is indeed not convex.

Exercise 5 Solution (8 Points)

We first prove that g is convex. This does however immediately follow since $\operatorname{epi} g = \operatorname{conv}(\operatorname{epi} f)$ which, by definition, is convex . Now suppose h is a convex underestimator of f. Then, by definition, we get $\operatorname{epi} f \subseteq \operatorname{epi} h$. Since the convex hull of a set is given by the intersection of all convex sets containing the set **(3 Points if proven)**, we get $\operatorname{conv}(\operatorname{epi} f) \subseteq \operatorname{epi} h$. Since $\operatorname{epi} g = \operatorname{conv}(\operatorname{epi} f)$ we get $\operatorname{epi} g \subseteq \operatorname{epi} h$

and hence $g(x) \geq h(x)$ for all $x \in \mathbb{R}^n$.
Proof of

$$\operatorname{conv}(C) = \bigcap_{\substack{C \subseteq K \\ K \text{ convex}}} K$$

where $C = \{ v_1, \cdot, v_n \}$:
We start with the direction $\subseteq$. Therefore let $K \supseteq C$ be a convex set. Since then $v_1, \ldots, v_n$ are in K, so must be all convex combinations, proving this direction .
The direction $\supseteq$ follows since $\operatorname{conv}(C)$ is convex and contains C, therefore is one of the sets in the intersection .

Exercise 6 Solution (10 Points)

First we compute the gradient and the Hessian:

$$\nabla g(x) = -\frac{e^{\frac{-f(x)}{c}}}{c} \nabla f(x) \qquad \nabla g(x)^2 = -\frac{e^{\frac{-f(x)}{c}}}{c} \nabla^2 f(x) + \frac{e^{\frac{-f(x)}{c}}}{c^2} \nabla f(x) \nabla f(x)^\top$$

.
To show that g is concave it is enough to show that its Hessian is negative semidefinite, so

$$\nabla^2 f(x) - \frac{1}{c} \nabla f(x) \nabla f(x)^\top \succeq 0$$

. To prove this we use the Schur complement. The above holds - using that c is positive - if and only if

$$\begin{pmatrix} \nabla^2 f(x) & \nabla f(x) \\ \nabla f(x)^\top & c \end{pmatrix} \succeq 0$$

. Applying Schur complement another time - this time using that $\nabla^2 f(x)$ is positive definite - gives us that this is equivalent to

$$c - \nabla^\top f(x) \nabla^2 f(x)^{-1} \nabla f(x) \geq 0$$

. By definition the left hand side is $c - \lambda^2(x)$ but this is non-negative by assumption .

Exercise 7 Solution (10 Points)

(a) $x^\top W x$ is a convex function for $W \succeq 0$, so $-x^\top W x$ is concave.

(b) The Lagrangian is

$$L(x, \lambda) = -x^\top W x + \sum_{i=1}^{n} \lambda_i \left(x_i^2 - 1 \right)$$

(c) For $D_\lambda = \operatorname{diag}(\lambda)$, the Lagrangian can be rewritten as

$$L(x, \lambda) = -\operatorname{Tr} D_\lambda + x^\top (D_\lambda - W) x.$$

Thus the dual function can be written as

$$\begin{aligned} \widehat{L}(\lambda) = \max \quad & t \\ & t \leq -Tr(D_\lambda) + x^\top (D_\lambda - W) x \qquad \text{for all } x \in \mathbb{R}^n \end{aligned}$$

(d) The constraints

$$\widehat{L}(\lambda) = \max \ t$$
$$t \leq -Tr\left(D_\lambda\right) + x^\top \left(D_\lambda - W\right) x \qquad \text{for all } x \in \mathbb{R}^n$$

can be formed as

$$\begin{pmatrix} D_\lambda - W & 0 \\ 0 & -t - \mathrm{Tr} D_\lambda \end{pmatrix} \succeq 0$$

Thus the dual function can be formed as

$$\widehat{L}(\lambda) = \max \ t$$
$$\begin{pmatrix} D_\lambda - W & 0 \\ 0 & -t - \mathrm{Tr} D_\lambda \end{pmatrix} \succeq 0$$

Thus the dual function has an explicit form

$$\widehat{L}(\lambda) = \begin{cases} -\mathrm{Tr} \ D_\lambda & \text{if } D_\lambda \succeq W \\ -\infty & \text{o.w.} \end{cases}$$

The dual program over a variable λ can be written as

$$\begin{aligned} \widehat{f}^* := \max_{\lambda} \quad & -\lambda^\top \mathbf{1} \\ s.t. \quad & \mathrm{diag}(\lambda) \succeq W. \end{aligned}$$

The above is an SDP and the constraint $\lambda \geq 0$ is automatically enforced by the PSD constraint.