Topic 1: Exercises of Convex Optimization Problems.

Jose M. Barcelo Ordinas

Universidad Politècnica de Catalunya (UPC-BarcelonaTECH),
Computer Architecture Dept.
joseb@ac.upc.edu

February 25, 2021

This part of the TOML-MIRI course is based on the book "Convex Optimization" by Stephen Boyd and Lieven Vandenberghe, Cambridge University Press.

1 Topic 11: Convex Sets.

Exer. 1: Say (justifying it) whether the following sets are convex or not.
a) The set of natural numbers $\mathbb{N}$,
b) B_2 = x$\in$ R^n: $||x||_2 \geq$ r,
c) B_2 = x$\in$ R^n: $||x||_2$ = r,
d) B_2 = x$\in$ R^n: $||x||_2 \leq$ r,

Solution: a) Not convex. Whatever is between two consecutive natural numbers is not in $\mathbb{N}$.
b) Not convex. Draw it.
c) Not convex. Draw it.
d) Convex. Use the definition of convex set and remember that D is a ball and $||x||_2$ is a norm (check the slides).

Exer. 2: Say (justifying it) whether the following sets are convex or not:
a)$\{(x, y) : y = e^x\}$,
b) $\{(x, y) : y \geq e^x\}$,
c) $\{(x, y) : y \leq e^x\}$,
d) $\{(x, y) : xy \geq 1, x > 0, y > 0\}$.

Solution: a) Not convex. See that for two points z,w: $e^{\theta z+(1-\theta)w} \neq \theta z+(1-\theta)w$ (draw a plot to see it). Note: take care not to be confused on whether the function f(x)=e^x is convex or not. As you may see on topic 2, this function is convex.

b) Convex. See that for two points z,w: $e^{\theta z+(1-\theta)w} \leq \theta z + (1-\theta)w$ (draw a plot to see it). Use the epigragh of a convex function is a convex set. Other option: (x=e^z),(u=e^w), then (s=e^t). Since the function is convex, by definition of convex functions:
$e^t = e^{\theta z+(1-\theta)w} \leq \theta e^z + (1-\theta)e^w \leq \theta x + (1-\theta)u \leq s$

c) Not convex. Draw it.

d) Convex. See that if: $xy \geq 1$ and $uv \geq 1$ then $(\theta x + (1-\theta)u)(\theta y + (1-\theta)v) \geq 1$ (draw a plot to see it). Another way:
(s,t)= (θ x + (1-θ) u)(θ y + (1-θ) v) = θ^2 xy + (1-θ)^2 uv + 2θ(1-θ)(xv+yu) $\geq \theta^2$ + (1-θ)^2 + 2 θ(1-θ) = 1

The fact that (xv+yu)$\geq$2 comes from the Arithmetic Mean-Geometric Mean Inequality (AM-GM or AMGM) that says that $\sum_n a_i/n \geq \prod_n a_i^{1/n}$. For example (a+b)/2$\geq \sqrt{ab}$, then (xv+yu)/2$\geq \sqrt{xvyu} \geq$1.

Exer. 3: Proof that the set $H = \{x \in R^n | a_1x_1 + ... + a_nx_n = c\}$ is a convex set. Hint: form a new point as a linear combination of points x and y.

Solution: Let us take points x,y$\in R^n$ and a combination $z = \theta x + (1-\theta)y$ with $0 \leq \theta \leq 1$, then:
$\sum_{i=1}^n a_i z_i = \sum_{i=1}^n a_i(\theta x_i + (1-\theta)y_i)$
$= \theta \sum_{i=1}^n a_i x_i + (1-\theta)\sum_{i=1}^n a_i y_i)$
$= \theta c + (1-\theta)c = c$

Exer. 4: *Show that the following sets are convex.*
a) A slab: $S = \{x \in R_+^n \mid \alpha \leq a^T x \leq \beta\}$
Solution: A slab is an intersection of two halfspaces, hence it is a convex set (and a polyhedron).

b) A Rectangle: $S = \{x \in R_+^n \mid \alpha_i \leq x_i \leq \beta_i\}$
Solution: As in part (a), a rectangle is a convex set and a polyhedron because it is an intersection of halfspaces.

c) The set of positive semi-definite symmetric matrices is convex.
Solution: That is, if M and N are positive semi-definite, for any $0 \leq \theta \leq 1$, then $\theta M + (1-\theta)N$ is also positive semi-definite. For any vector x:
$x^T(\theta M + (1-\theta)N)x = \theta x^T M x + (1-\theta)x^T N x \geq 0$

Exer. 5: Let's have C_1=[1,4], C_2=[2,3] and C_3=[5,7]. Say whether the following sets are convex or not.
a) $C_1 \cup C_2$
b) $C_1 \cup C_3$
c) $C_1 \cap C_2$
d) $C_1 \cap C_3$
e) $C_1 + C_2$

Solution: a) yes, since $C_2 \subset C_1$, $C_1 \cup C_2 = C_1$ = [1,4]
b) no, $C_1 \cup C_3$ = [1,4] $\cup$ [5,7]
c) yes, $C_1 \cap C_2 = C_2$ = [2,3]
d) yes, $C_1 \cap C_3 = \emptyset$
e) yes, $C_1 + C_2$ = [3,7]

Exer. 6: Consider the normal vector a^t = (1,-2,3) and the point x_0 = (2,2,2). Obtain the equation of an hyperplane defined by this normal vector and the point.

Solution: $x_1 - 2x_2 + 3x_3 = 4$

Exer. 7: Consider the normal surface $x_3=3x_1^2-x_1x_2$. Obtain the equation of the supporting hyperplane at point x_0=(1,2,1).
Solution: the implicit function is $F(x_1,x_2,x_3)=3x_1^2-x_1x_2-x_3=0$. Then, the equation of the supporting hyperplane is $\nabla F(x_0)^t(x-x_0)=0$. Since $\nabla F(x)^t=(6x_1-x_2, -x_1, -1)$, then $\nabla F(x_0)^t=(4, -1, -1)$, and:

$$\nabla F(x_0)^t(x-x_0) = (4,-1,-1)\begin{pmatrix} x_1-1 \\ x_2-2 \\ x_3-1 \end{pmatrix} = 0,$$

implies that the supporting hyperplane is: $4x_1-x_2-x_3=1$.

2 Topic 12: Convex Functions.

Exer. 1: *See whether the following matrix is positive-definite or negative-definite. (i) Use the eigenvalue test and, (ii) use the principal minor test:*

$$P_1 = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 3 & 3 \\ 4 & 0 & 1 \end{pmatrix} \text{ idem with, } P_2 = \begin{pmatrix} -2 & 0 & 0 \\ 1 & -3 & 3 \\ 4 & 0 & -1 \end{pmatrix}$$

For P_1, i) We solve the equation Ax = λ x, it is to say, det(A-λI) = 0 :

$$det(A - \lambda I) = det \begin{pmatrix} 2-\lambda & 0 & 0 \\ 1 & 3-\lambda & 3 \\ 4 & 0 & 1-\lambda \end{pmatrix} = (2-\lambda)(3-\lambda)(1-\lambda) = 0.$$

Then, the eigenvalues are: λ=3, λ=2, λ=1, all are positive, and then the matrix is positive-definite.

ii) Using the principal minors:
1st principal minor: det(A) = 6
2nd principal minor:

$$det \begin{pmatrix} 3 & 3 \\ 0 & 1 \end{pmatrix} = 3$$

3rd principal minor = 1
Then all principal minors are positive and the matrix is positive-definite.

For P_2, you can see that the eigenvalues are (-2, -3, -1) and the principal minors are -6, 3 and -1, negative for the odd minors and positive for the even minor. Then, the matrix is negative-definite.

Exer. 2: *(3.16 Boyd) For each of the following functions determine whether it is convex, concave, quasi-convex, or quasi-concave:*
(a) f(x) = e^x - 1 on R
(b) $f(x_1; x_2) = x_1x_2$ on R^2_{++}
(c) $f(x_1; x_2) = 1/(x_1x_2)$ on R^2_{++}
(d) $f(x_1; x_2) = x_1/x_2$ on R^2_{++}
(e) $f(x_1; x_2) = x_1^2/x_2$ on R^2_{++}
(f) $f(x_1; x_2) = x_1^\alpha x_2^{1-\alpha}$ with $0 \leq \alpha \leq 1$ on R^2_{++}.

Solution:

a) f(x) = $e^x - 1$ on R. Calculate the second derivative and see that is positive. Then, It is strictly convex, and therefore quasiconvex (draw it an obtain the α-sublevels sets). Also quasiconcave but not concave.

b) $f(x_1; x_2) = x_1x_2$ on R^2_{++}.

The Hessian $\nabla^2 f(x)$=(0,1;1,0); which is neither positive semidefinite nor negative semidefinite (see that their eigenvalues are equal to λ_1=1 and λ_1=-1. Therefore, f is neither convex nor concave. It is quasiconcave, since its superlevel sets $\{(x_1, x_2) \in R^2_{++} \mid x_1x_2 \geq \alpha\}$ are convex. It is not quasiconvex.

c) $f(x_1; x_2) = 1/(x_1x_2)$ on R^2_{++},

The Hessian $\nabla^2 f(x)$=$\frac{1}{x_1x_2}$ ($\frac{2}{x_1^2}$,$\frac{1}{x_1x_2}$;$\frac{1}{x_1x_2}$,$\frac{2}{x_2^2}$). Therefore, f is convex and quasiconvex. It is not quasiconcave or concave.

d) $f(x_1; x_2) = x_1/x_2$ on R^2_{++}.

The Hessian $\nabla^2 f(x)$= (0,$\frac{-1}{x_2^2}$;$\frac{-1}{x_2^2}$,$\frac{2x_1}{x_2^3}$), which is not positive or negative semidefinite. Therefore, f is not convex or concave. It is quasiconvex and quasiconcave (i.e., quasilinear), since the sublevel and superlevel sets are halfspaces.

e) $f(x_1; x_2) = x_1^2/x_2$ on R^2_{++}. f is convex, as mentioned on page 72 (see also figure 3.3) of S. Boyd book. This is easily verified by working out the Hessian:

$\nabla^2 f(x)$= ($\frac{2}{x_2}$,$\frac{-2x_1}{x_2^2}$;$\frac{-2x_1}{x_2^2}$,$\frac{2x_1^2}{x_2^3}$)= $\frac{2}{x_2}$(1;$\frac{-2x_1}{x_2}$)(1;$\frac{-2x_1}{x_2}$)^T). Therefore, f is convex and quasiconvex. It is not concave or quasiconcave.

f) $f(x_1; x_2) = x_1^\alpha x_2^{1-\alpha}$ with $0 \leq \alpha \leq 1$ on R^2_{++}. Concave and quasiconcave. The Hessian is:

$\nabla^2 f(x)$=$\alpha(1-\alpha)x_1^\alpha x_2^{(1-\alpha)}$($\frac{1}{x_1}$;$\frac{-1}{x_2}$)($\frac{1}{x_1}$;$\frac{-1}{x_2}$)^T

Exer. 3: Let $f(x_1, x_2) = x_1^2 - x_1x_2 + x_2^2 + 3x_1 - 2x_2 + 1$ in R. Is f convex, concave, or neither?

Solution: The Hessian matrix of f is: $\nabla^2 f(x)$=(2, -1;-1, 2). This matrix is positive definite, so f is convex.

Exer. 4: Let $f(x_1, x_2) = x_1^3 + 2x_1^2 + 2x_1x_2 + (1/2)x_22 - 8x_1 - 2x_2 - 8$. Find the range of values of (x_1, x_2) for which f is convex, if any.

Solution: The Hessian matrix of f is $\nabla^2 f(x)$=($6x_1 + 4$, 2;2, 1). This matrix is positive semidefinite if $6x_1 + 4 \geq 0$ and $6x_1 \geq 0$, or if $x_1 \geq 0$. Thus f is convex for $x_1 \geq 0$ (and all x_2).

Exer. 5: Let f(x) = ax^α, where a>0 and α are parameters. For what values of α is f (which is twice differentiable) nondecreasing and concave on the interval [0,∞)?

Solution: We have f'(x) = αa$x^{\alpha-1}$ and f''(x) = α(α-1)a$x^{\alpha-2}$. For any value of β we have x^β $\geq$0 for all x$\geq$0, so for f to be nondecreasing and concave we need α $\geq$0 and α(α-1)$\leq$0, or equivalently 0$\leq$ α $\leq$1.

Exer. 6: Determine the values of a (if any) for which the function $2x^2 + 2xz + 2ayz + 2z^2$ is concave and the values for which it is convex.

Solution: The Hessian of the function is $\nabla^2 f(x)$=(4,0,2; 0,0,2a; 2,2a,4). The minors are -4a^2, -4a^2 and 0. Thus for a=0 the Hessian is positive semidefinite, so that the function is convex; for other values of a the Hessian is indefinite, so that the function is neither concave nor convex.

Exer. 7: Say (draw) what are the upper level sets of each of the following functions for the indicated values.

a) f(x,y) = x^2 + y^2 for the value α=a.
b) f(x,y) = -x^2 - y^2 for the value α=a.
c) f(x,y) = xy for the value α=a.

Solution:

a) f(x,y) = x^2 + y^2 for the value α=a. Then the upper level sets are f(x,y) = x^2 + $y^2 \geq a$, then if a<0, $C_a = \emptyset$, and if a>0, C_a={(x,y)$\in$ R^2: x^2 + y^2 $\geq$ a}, then it is the complementary of a disk (if x^2 + y^2 $\leq a$ it would be a disk of radius $a^{1/2}$), then it is not

a convex set. To visualize it, think that x^2+y^2=z is the equation of a Elliptic Paraboloid. Thus, the domain of f is the projection over the x-y axes, a disk of radius $a^{1/2}$.
b) f(x,y) = -x^2 - y^2 for the value α=a. Then the upper level sets are f(x,y) = -x^2 - y^2 $\geq a$. If a>0, $-x^2$ - y^2 $\geq$ a, implies $C_a = \emptyset$. If a<0, C_a={(x,y)$\in$ R^2: -x^2-y^2 $\geq$ a} = {(x,y)$\in$ R^2: x^2+y^2 $\leq$ -a, with $a < 0$}, a disk of radius $a^{1/2}$.
c) f(x,y) = xy for the value α=a. Then the upper level sets are f(x,y) = xy $\geq$a:

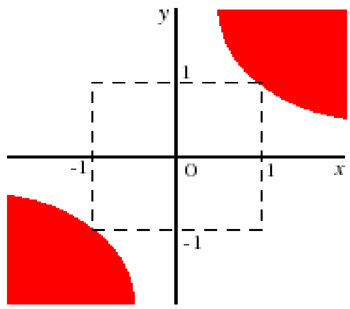

Figure 1: Upper level sets of f(x,y) = xy with a=1.

3 Topic 13: Convex Optimization Problems.

Exer. 1: a) Consider the function f of a single variable defined by f(x)=-x-1 for x<-1, f(x)=0 for $-1 \leq x \leq 1$, and f(x)=x-1 for x>1. Is the point x=0 a global maximizer and/or a global minimizer and/or a local maximizer and/or a local minimizer of f ?
b) Consider the function f of a single variable defined by f(x)=x+1 for x<-1, f(x)=0 for $-1 \leq x \leq 1$, and f(x)=x-1 for x>1. Is the point x=0 a global maximizer and/or a global minimizer and/or a local maximizer and/or a local minimizer of f ?

Solution: a) The point x=0 is not a global maximizer (f(2)=1>f(0)=0, for example), but is a global minimizer (f(x)$\geq$ f(0)=0 for all x), and is both a local maximizer (f(x)$\leq$ f(0)=0 for all x with $-1 \leq x \leq 1$) and a local minimizer.
b) The point x=0 is not a global maximizer (f(2)=1> f(0)=0, for example), or a global minimizer (f(-2)=-1< f(0)=0, for example). It is both a local maximizer (f(x)$\leq$ f(0)=0 for all x with $-1 \leq x \leq 1$) and a local minimizer.

Exer. 2: *(4.1 Boyd).*
Consider the optimization problem

$$\begin{array}{ll} \text{minimize} & f(x,y) \\ \text{subject to} & 2x+y \geq 1 \\ & x+3y \geq 1 \\ & x \geq 0, y \geq 0 \end{array}$$

Make a sketch of the feasible set. For each of the following objective functions, give the optimal set and the optimal value.
(a) $f(x,y) = x+y$
(b) $f(x,y) = -x-y$
(c) $f(x,y) = x$
(d) $f(x,y) = max\{x,y\}$

Solution: The feasible set is the convex hull above of (0,1) to (0,∞), (0,1) to (2/5,1/5), (2/5,1/5) to (1,0) and (1,0) to (∞,0).

(a) $x^* = (2/5; 1/5)$.
(b) Unbounded below.
(c) $X_{opt} = \{(0, y) \mid y \geq 1\}$.
(d) x^* = (1/3, 1/3). Just, the intersection of line x=y that divides the halfspaces $x > y$ and $x < y$ with the segment 2x+y=1.

Exer. 3: *(4.3 Boyd).* Prove that x^*=(1, 1/2,-1) is optimal for the optimization problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}x^T Px + q^T x + r \\ \text{subject to} & -1 \leq x_i \leq 1, i = 1, 2, 3 \end{array}$$

where

$$P = \begin{pmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{pmatrix}; \quad q = \begin{pmatrix} -22.0 \\ -14.5 \\ 13.0 \end{pmatrix}; \quad r = 1.$$

Hint: test with the optimality condition.

Solution: We verify that x^* satisfies the optimality condition (slide 12 of topic 13). The gradient of the objective function at x^* is: $\nabla f_0(x^*) = (-1, 0, 2)$.
Therefore the optimality condition is that $\nabla f_0(x^*)^T(y - x^*) = -1(y_1 - 1) + 2(y_3 + 1) = -y_1 + 2y_3 + 3 \geq 0$ for all y satisfying $-1 \leq y_i \leq 1$, which is clearly true.

Exer. 4: Express the following GP problem in its convex form.

$$\begin{array}{ll} \text{minimize} & 2x_1^{-1}x_2^2 + x_1^2 x_2^{1/2} x_3^{-1/3} \\ \text{subject to} & 5x_1^{-1}x_2 + 4x_1^3 x_3^{-1} \leq 1 \\ & x_1x_2 + 3x_1^2 x_2^{-1} x_3^2 + x_3 \leq 1 \\ & x_1^{-2} x_2^{-3/2} x_3 = 1 \end{array}$$

Solution: First, write down the vector coefficients:

b_{01}=log 2, b_{02}=log 1 = 0, a_{01}=(-1,2,0), a_{02}=(2,1/2,-1/3)
b_{11}=log 5, b_{12}=log 4, a_{11}=(-1,1,0), a_{12}=(3,0,-1)
b_{21}=log 1 = 0, b_{22}=log 3, b_{23}=log 1 = 0, a_{21}=(1,1,0), a_{22}=(2,-1,2), a_{23}=(0,0,1)
b_{21}=log 1 = 0, a_{31}=(-2,-3/2,1),

Now, use the log-sum-exp terms to express the convex form:

$$\begin{array}{ll} \text{minimize} & log(exp(log(2) - y_1 + 2y_2) + exp(2y_1 + 1/2y_2 - 1/3y_3)) \\ \text{subject to} & log(exp(log(5) - y_1 - y_2) + exp(log(4) + 3y_1 - y_3)) \leq 0 \\ & log(y_1 + y_2) + exp(log(3) + 2y_1 - y_2 + 2y_3) + exp(y_3)) \leq 0 \\ & -2y_1 - 3/2y_2 + y_3 = 0 \end{array}$$

4 Topic 14: Duality.

Exer. 1: Consider the optimization problem and find the maximum using the KKT conditions.

maximize $x^2 + y^2$
subject to $x^2 + xy + y^2 = 3$

Solution: The Lagrangian is L(x,y,λ)=$x^2 + y^2$ + $\lambda(x^2 + xy + y^2 - 3)$. We obtain the gradient:
∇_xL(x,y,λ)=2x+2xλ+yλ=0, then λ=$\frac{-2x}{2x+y}$
∇_yL(x,y,λ)=2y+2yλ+xλ=0, then λ=$\frac{-2y}{2y+x}$
Then, $x^2 = y^2$, then x=$\pm$y.

Using the slackness condition: $x^2 + xy + y^2 - 3$=0 with x=$\pm$y we obtain 4 solutions: (-1,-1), (1,1), (-$\sqrt{3}$,$\sqrt{3}$), ($\sqrt{3}$,-$\sqrt{3}$). The maximum is over the last two solutions: f(-$\sqrt{3}$,$\sqrt{3}$)=f($\sqrt{3}$,-$\sqrt{3}$)=6.

Exer. 2: A convex problem in which strong duality fails. Consider the optimization problem:

minimize e^{-x}
subject to $x^2/y \leq 0$

with variables x,y and the domain D={(x,y) | y>0}. Does the Slater's condition hold for this problem?

Solution:

The Slater's condition is not meet. See that y>0, and also x^2 >0 for any x. Then, x^2/y be $\leq$0, has to be x=0 and then there exists no x such that the constraint f(x,y)=x^2/y <0.

Exer. 3: *(5.1 Boyd). A simple example. Consider the optimization problem:*

minimize $x^2 + 1$
subject to $(x-2)(x-4) \leq 0$

with variable $x \in R$.

(a) Analysis of primal problem. Give the feasible set, the optimal value, and the optimal solution.
(b) Lagrangian and dual function. Plot the objective $x^2 + 1$ versus x. On the same plot, show the feasible set, optimal point and value, and plot the Lagrangian $L(x, \lambda)$ versus x for a few positive values of λ. Verify the lower bound property ($p^* \geq inf_x L(x, \lambda)$ for $\lambda \geq 0$). Derive and sketch the Lagrange dual function g.
(c) Lagrange dual problem. State the dual problem, and verify that it is a concave maximization problem. Find the dual optimal value and dual optimal solution λ^*. Does strong duality hold?
(d) Sensitivity analysis. Let $p^*(u)$ denote the optimal value of the problem.

minimize $x^2 + 1$
subject to $(x-2)(x-4) \leq u$

as a function of the parameter u. Plot $p^*(u)$. Verify that $\frac{\partial p^*(0)}{\partial u} = -\lambda^*$?.

Solution:

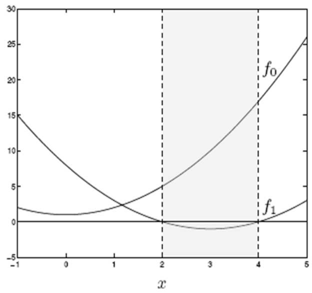

Figure 2: feasible set.

(a) The feasible set is the interval [2,4]. The (unique) optimal point is x^* = 2, and the optimal value is p^* = 5. The plot shows f_0 and f_1.

(b) The Lagrangian is $L(x, \lambda) = (1 + \lambda)x^2 - 6\lambda x + (1 + 8\lambda)$.

The plot shows the Lagrangian $L(x, \lambda) = f_0 + \lambda f_1$ as a function of x for different values of $\lambda \geq 0$. Note that the minimum value of $L(x, \lambda)$ over x (i.e., g(λ)) is always less than p^*. It increases as λ varies from 0 toward 2, reaches its maximum at λ = 2, and then decreases again as λ increases above 2. We have equality $p^* = g(\lambda)$ for λ = 2.

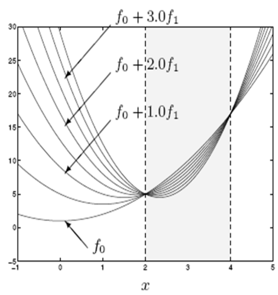

Figure 3: Lagrangian.

For $\lambda > -1$, the Lagrangian reaches its minimum at $x = \frac{3\lambda}{(1+\lambda)}$. For $\lambda \leq -1$ it is unbounded below. Thus, The Lagrange dual function:

$$g(\lambda) = \begin{cases} \frac{-9\lambda^2}{(1+\lambda)} + 1 + 8\lambda & \lambda > -1 \\ -\infty & \lambda \leq -1 \end{cases}$$

which is plotted below.

We can verify that the dual function is concave, that its value is equal to p^* = 5 for λ = 2, and less than p^* for other values of λ.

(c) The Lagrange dual problem is

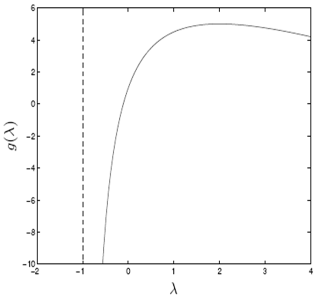

Figure 4: The Lagrange dual function.

$$\begin{array}{ll} \text{maximize} & \dfrac{-9\lambda^2}{(1+\lambda)} + 1 + 8\lambda \\ \text{subject to} & \lambda \geq 0 \end{array}$$

The dual optimum occurs at λ = 2, with d^* = 5. So for this example we can directly observe that strong duality holds (as it must — Slater's constraint qualification is satisfied).

d) The perturbed problem is infeasible for $u < -1$, since $inf_x(x^2 - 6x + 8) = -1$. For $u \geq -1$, the feasible set is the interval $[3 - \sqrt{1+u}, 3 + \sqrt{1+u}]$. given by the two roots of $x^2 - 6x + 8 = u$. For $-1 \leq u \leq 8$ the optimum is $x^*(u) = 3 - \sqrt{1+u}$. For $u \geq 8$, the optimum is the unconstrained minimum of f_0, i.e., $x^*(u) = 0$. In summary,

$$p^*(u) = \begin{cases} -\infty & u < -1 \\ 11 + u - 6\sqrt{1+u} & -1 \leq u \leq 8 \\ 1 & u \geq 8 \end{cases}$$

The Figure shows the optimal value function $p^*(u)$ and its epigraph.

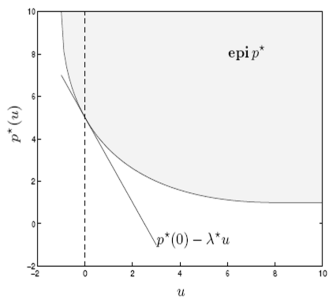

Figure 5: $p^*(u)$.

Finally, we note that $p^*(u)$ is a differentiable function of u, and that $\frac{\partial p^*(0)}{\partial u} = -2 = -\lambda^*$

Exer. 4: *(5.20 Boyd). Dual of channel capacity problem. Derive a dual for the problem*

$$\begin{array}{ll} \textit{minimize} & -c^T x + \sum_{i=1}^{m} (y_i log(y_i)) \\ \textit{subject to} & Px = y \\ & 1^T x = 1 \qquad\qquad\qquad x \geq 0 \end{array}$$

where $P \in R^{mxn}$ has nonnegative elements, and its columns add up to one (i.e., $P^T 1 = 1$). The variables are $x \in R^n$, $y \in R^m$.

Solution:

The Lagrangian is

$L(x, y, \lambda, \nu, \alpha) = -c^T x + \sum_{i=1}^{m} (y_i log(y_i)) - \lambda^T x + \nu(\mathbf{1}^T x - 1) + \alpha^T (Px - y)$=
$(-c - \lambda + \nu\mathbf{1} + P^T \alpha)^T x + \sum_{i=1}^{m} (y_i log(y_i)) - z^T y - \nu.$

The minimum over x is bounded below if and only if: $-c - \lambda + \nu\mathbf{1} + P^T \alpha = 0$

To minimize over y, we set the derivative with respect to y_i equal to zero, which gives $log y_i + 1 - \alpha_i = 0$, and conclude that:
$inf_{y_i \geq 0} \{y_i log(y_i) - \alpha_i y_i)\} = -e^{z_i - 1}$

The dual function is:

$$g(\lambda, \nu, \alpha) = \begin{cases} -\sum_{i=1}^{m} (e^{z_i - 1}) - \nu & \text{for: } (-c - \lambda + \nu\mathbf{1} + P^T \alpha) = 0 \\ -\infty & \text{otherwise} \end{cases}$$

The dual problem is:

$$\begin{array}{ll} \text{maximize} & -\sum_{i=1}^{m} (e^{z_i - 1}) - \nu \\ \text{subject to} & -c + \nu\mathbf{1} + P^T \alpha \geq 0 \end{array}$$

This can be simplified by introducing a variable $w = z + \nu\mathbf{1}$ (and using the fact that $1 = P^T \mathbf{1}$), which gives:

$$\begin{array}{ll} \text{maximize} & -\sum_{i=1}^{m} (e^{w_i - \nu - 1}) - \nu \\ \text{subject to} & P^T w \geq c \end{array}$$

Finally we can easily maximize the objective function over ν by setting the derivative equal to zero (the optimal value is $\nu = -log(\sum_{i=1}^{m} e^{1 - w_i})$ which leads to:

$$\begin{array}{ll} \text{maximize} & -log(\sum_{i=1}^{m} e^{w_i}) - 1 \\ \text{subject to} & P^T w \geq c \end{array}$$

This is a geometric program, in convex form, with linear inequality constraints (i.e., monomial inequality constraints in the associated geometric program).

Exer. 5: *(5.26 Boyd). Consider the QCQP:*

$$\begin{array}{ll} \textit{minimize} & x_1^2 + x_2^2 \\ \textit{subject to} & (x_1-1)^2 + (x_2-1)^2 \le 1 \\ & (x_1-1)^2 + (x_2+1)^2 \le 1 \end{array}$$

with variable $x \in R^2$. (a) Sketch the feasible set and level sets of the objective. Find the optimal point x^* and optimal value p^*. (b) Give the KKT conditions. Do there exist Lagrange multipliers λ_1^*, λ_2^* that prove that x^* is optimal? (c) Derive and solve the Lagrange dual problem. Does strong duality hold?

Solution:

(a) The figure shows the feasible set (the intersection of the two shaded disks) and some contour lines of the objective function. There is only one feasible point, (1, 0), so it is optimal for the primal problem, and we have p^* = 1.

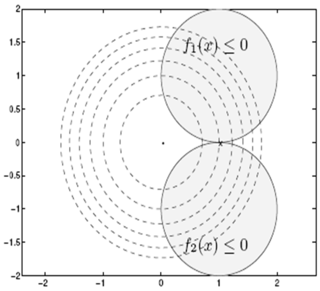

Figure 6: feasible sets.

(b) The KKT conditions are:

$(x_1-1)^2 + (x_2-1)^2 \le 1$
$(x_1-1)^2 + (x_2+1)^2 \le 1$
$\lambda_1 \ge 0, \lambda_2 \ge 0,$
$2x_1 + 2\lambda_1(x_1-1) + 2\lambda_2(x_1-1) = 0$
$2x_2 + 2\lambda_1(x_2-1) + 2\lambda_2(x_2+1) = 0$
$\lambda_1((x_1-1)^2 + (x_2-1)^2 - 1) = 0$
$\lambda_2((x_1-1)^2 + (x_2+1)^2 - 1) = 0$

At x=(1, 0), these conditions reduce to:

$\lambda_1 \ge 0, \lambda_2 \ge 0,$
$2 = 0,$
$-2\lambda_1 + 2\lambda_2 = 0$

which (clearly, in view of the third equation) have no solution.

(c) The Lagrange dual function is given by:

$g(\lambda_1, \lambda_2) = inf_{x_1,x_2} L(x_1, x_2, \lambda_1, \lambda_2)$

where:

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

L reaches its minimum for:
$x_1 = \frac{\lambda_1+\lambda_2}{1+\lambda_1+\lambda_2}$
$x_2 = \frac{\lambda_1-\lambda_2}{1+\lambda_1+\lambda_2}$

and we find:

$$g(\lambda_1, \lambda_2) = \begin{cases} -\frac{(\lambda_1+\lambda_2)^2+(\lambda_1-\lambda_2)^2}{1+\lambda_1+\lambda_2} + \lambda_1 + \lambda_2 & 1 + \lambda_1 + \lambda_2 \\ -\infty & \text{otherwise} \end{cases}$$

where we interpret $a/0 = 0$ if a = 0 and as $-\infty$ if $a < 0$. The Lagrange dual problem is given by:

$$\begin{array}{ll} \text{maximize} & \dfrac{\lambda_1 + \lambda_2 - (\lambda_1 - \lambda_2)^2}{1 + \lambda_1 + \lambda_2} \\ \text{subject to} & \lambda_1 \geq 0, \lambda_2 \geq 0 \end{array}$$

Since g is symmetric, the optimum (if it exists) occurs with $\lambda_1 = \lambda_2$. The dual function then simplifies to

$g(\lambda_1, \lambda_2) = \frac{2\lambda_1}{2\lambda_1+1}$

We see that $g(\lambda_1, \lambda_2)$ tends to 1 as $\lambda_1 \to \infty$. We have $d^* = p^* = 1$, but the dual optimum is not attained.

Recall that the KKT conditions only hold if (1) strong duality holds, (2) the primal optimum is attained, and (3) the dual optimum is attained. In this example, the KKT conditions fail because the dual optimum is not attained.

5 Topic 15: Methods.

Exer. 1: *Describe the first iteration of the Gradient Descent method with Line Search for the problem:*

$$\textit{minimize} \quad f(z) = f(x, y) = (y + x^2 - 1)^2 + (x + y^2 - 1)^2$$

with initial point $z^0 = (x^0, y^0) = (0, 0)$.

Solution:

In order to find the initial direction we calculate the gradient of the function $f(z) = f(x,y) = (y+x^2-1)^2 + (x+y^2-1)^2$:

$\nabla f(z) = (4x^3+2y^2+2xy-2x-2, 4y^3+2x^2+2xy-2y-2)$ and then the initial direction is $d^0 = (d_1^0, d_2^0) = -\nabla f(0,0) = (2,2)$.

Now, the step for the iteration k=0 is: $h(s) = argmin_{s>0}(f(z^0+sd^0))$=$argmin_{s>0}(f(x^0+sd_1^0, y^0+sd_2^0))$= $argmin_{s>0}(f(2s,2s))$=$argmin_{s>0}(2s+4s^2-1)^2+(2s+4s^2-1)^2 = argmin_{s>0}2(4s^2+2s-1)^2$.

Calculating the minimum of this function over s results: $\frac{\partial h(s)}{\partial s} = 4s^2+2s-1 = 0$, with $s > 0$. The solution of this equation is: $s = \frac{-1\pm\sqrt{5}}{4}$. Then, $s^0 = \frac{-1+\sqrt{5}}{4} > 0$, and the first point of the iteration will be:

$z^1 = (x^1, y^1) = (x^0+s^0d_1^0, y^0+s^0d_2^0) = (0+2s^0, 0+2s^0) = (\frac{-1+\sqrt{5}}{2}, \frac{-1+\sqrt{5}}{2})$

Exer. 2: *(11.1 Boyd).* Barrier method example. Consider the simple problem

$$\begin{array}{ll} \text{minimize} & x^2+1 \\ \text{subject to} & 2 \le x \le 4 \end{array}$$

which has feasible set [2, 4], and optimal point x*=2. Plot f_0 , and $tf_0 0 + \phi$, for several values of t$\geq$ 0, versus x. Label x*(t).

Solution: The figure shows the function f_0 + (1/t)I for $f_0(x)=x^2+1$, with barrier function I(x)= -log(x-2) - log(4-x), for t= 10^{-1} , $10^{-0.8}$, $10^{-0.6}$,..., $10^{0.6}$, $10^{0.8}$, 10. The inner curve corresponds to t=0.1, and the outer curve corresponds to t=10. The objective function is shown as a dashed curve.

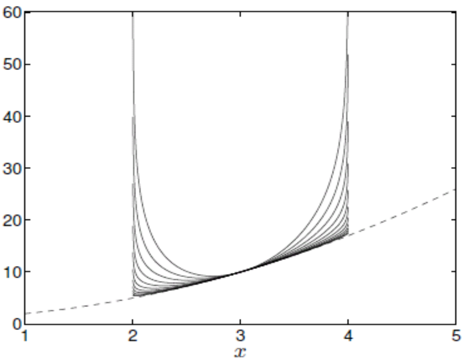

Figure 7: Barrier Log method.

Exer. 3: Consider the following problem:

$$\begin{array}{ll} \text{minimize} & x_1 - 2x_2 \\ \text{subject to} & 1 + x_1 - x_2^2 \ge 0 \\ & x_2 \ge 0 \end{array}$$

Write the Log Barrier function and give a solution as a function of the log barrier parameter t.

Solution: The log barrier function is given by (calling $\mu = 1/t$:

$\phi(\mu)$=-μ log(1+x_1-x_2^2) - μ log(x_2), and the problem to be solved is:

$$\text{minimize} \quad x_1 - 2x_2 - \mu\, log(1+x_1-x_2^2) - \mu\, log(x_2)$$

Applying now the first order conditions:
1 - μ/$(1 + x_1 - x_2^2)$=0
-2 + $2\mu x_2$/$(1 + x_1 - x_2^2)$ - μ/x_2=0
If the constrains are strictly (¿) satisfied, then from the second equation we may get an expression for x_2: $x_2^2 - x_2 - \mu/2 = 0$ and also for x_1 from the first equation:
$x_2 = 1/2 + 1/2\sqrt{1+2\mu}$
$x_1 = 1/2(\sqrt{1+2\mu} + 3\mu - 1)$.

Since the log barrier works well when t is large, then μ tends to zero when t is large, and obtaining the limit for x_1, x_2 when μ tend to zero lies the solution (x_1^*,x_2^*)=(0,1).