CS5580: Practice Problems

Note: Work out all problems on your own. If you think your answer is NOT satisfactory/correct, then ask for hints from friend/instructor. You MUST use ONLY the definitions, results/theorems used during the lectures. Do not google for proofs/solutions as they might work with different definitions/axioms. Provide as rigorous proofs as you can.

1 Introduction

1. Show that

$$\begin{array}{cc} \max_{x\in\mathcal{X}} & f(x) \\ \text{s.t.} & g_i(x) \leq 0 \ \forall\ i \end{array} = \begin{array}{cc} -\min_{x\in\mathcal{X}} & -f(x) \\ \text{s.t.} & g_i(x) \leq 0 \ \forall\ i \end{array}$$

2. Prove that

$$h\left(\begin{array}{cc} \min_{x\in\mathcal{X}} & f(x) \\ \text{s.t.} & g_i(x) \leq 0 \ \forall\ i \end{array}\right) = \begin{array}{cc} \min_{x\in\mathcal{X}} & h(f(x)) \\ \text{s.t.} & g_i(x) \leq 0 \ \forall\ i, \end{array}$$

whenever h is a monotonically non-decreasing function.

3. Show that:

$$\begin{array}{ccccc} \min_{y\in\mathcal{Y}} \min_{x\in\mathcal{X}} & f(x,y) & = & \min_{x\in\mathcal{X}} \min_{y\in\mathcal{Y}} & f(x,y) \\ \text{s.t.} & g_i(x,y) \leq 0 \ \forall\ i & & \text{s.t.} & g_i(x,y) \leq 0 \ \forall\ i. \\ & & = & \min_{z\in\mathcal{Z}} & f(z) \\ & & & \text{s.t.} & g_i(z) \leq 0 \ \forall\ i, \end{array}$$

where $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}, z = (x,y)$.

2 Theory: Convex Analysis

2.1 Vector Spaces and subsets in them

1. Consider $\mathbb{R}^2$ (2-dim Euclidean vec) and the 1-norm as the norm. Show that 1-norm is not induced by any inner-product in this space.

2. If S is a subspace show that $S^{\perp}$ is a subspace.

3. If S_1 and S_2 are subspaces of a vector space, show that $S_1 + S_2 = LIN(S_1 \cup S_2)$ and hence is a subspace.

4. Prove the following results which illustrate how limits and lin. comb.; limits and inner-products *distribute*. Assume $\{x_n\} \to x$, $\{y_n\} \to y$ and $\{\alpha_n\} \to \alpha$, $\{\beta_n\} \to \beta$. Here all x_n, y_n, x, y are vectors in some (finite-dim) inner-product space and all $\alpha_n, \beta_n, \alpha, \beta$ are in $\mathbb{R}$.

   (a) $\{\alpha_n x_n + \beta_n y_n\} \to \alpha x + \beta y$

   (b) $\{\langle x_n, y_n\rangle\} \to \langle x, y\rangle$

   (c) $\{||x_n - y_n||\} \to ||x - y||$

5. Show that complement of an open set is closed and vice-versa.

6. Consider the sequence of sets $\{A_n\}$ given by $A_n = [0, 1 + \frac{1}{n})$. What is $\cap_{i=1}^{\infty} A_n$. Justify your answer[1].

7. Write down the outer description for $LIN(X)$ where $X = \{[1\ 1\ 1\ 1]^\top, [1\ 1\ -1\ \ -1]^\top\}$.

8. What is the dimension of the affine set given by the following system of equations: $\sum_{j=1}^{n}(i - j)x_j = i,\ i = 1, \ldots, m$ (Assume $2 \leq m \leq n$).

9. Prove that the cone of psd matrices in the space of all symmetric matrices of a given size is self-dual[2]

[1] This shows that intersection of infinite number of open sets need not be open. By De'Morgan's laws, we also can say that union of infinite number of closed sets need not be closed. Note that finite intersection(union) of open(closed) sets is open(closed).

[2] Note that the same cone in the set of all square matrices of a given size is NOT self-dual. For eg. $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is in the dual cone, which is not a symmetric matrix. However if we started with all matrices M which satisfy $x^\top M x \geq 0$ (and need not be symmetric), then it is easy to see that this set is also a cone and moreover, self-dual.

10. Prove that once a cone has a line, then it must have all parallel lines to it through every point in the cone.
11. Consider the cone $K = \{(x, z) \in \mathbb{R}^{n+1} \mid x \in \mathbb{R}^n, z \in \mathbb{R}, ||x||_p \leq z\}$, where $p \in [1, \infty]$. Provide an analytic expression for its dual cone. Hint: Holder's inequality.
12. Consider the cone $K = \{(x, z) \in \mathbb{R}^{n+1} \mid x \in \mathbb{R}^n, z \in \mathbb{R}, ||x||_M \leq z\}$, where[3] $M \succ 0$. Provide an analytic expression for its dual cone.
13. What is the dual norm for the Frobenius norm ?
14. Problems 2.2, 2.8, 2.13 from Boyd's book.
15. Show that intersection of two polyhedral cones is a polyhedral cone[4].
16. Show that any linear combination of convex sets (in a particular inner-product space) is a convex set (in that particular inner-product space). Hence set of all convex sets (in a particular inner-product space) forms a vector space by itself!
17. Show that the tangent cone at any point in a convex set is indeed a cone.

2.2 Functions on vector spaces

1. Show that a function $f : A \mapsto \mathbb{R}$, where A is an affine set in V (the set of all vectors in a vector space), is affine iff $f(x) = \langle a, x \rangle - b \; \forall \; x \in A$ for some $b \in \mathbb{R}$ and some a, a vector in the linear set associated with the affine set A.
2. Show that $f : C \mapsto \mathbb{R}$, where $C \subset V$ is a conic set (convex set), is a conic function (convex function) if and only if $epi(f)$ is a conic set (convex set).
3. Show that f is an affine function if and only if f and $-f$ are both convex functions.

[3] $||x||_M \equiv \sqrt{x^\top M x}$.

[4] Recall that according to our definition a polyhedral cone is that with finite primal description. Also, after proving this result, it follow by induction that intersection of any finite number of polyhedral cones is a polyhedral cone.

4. Suppose $f : X \times Y \mapsto \mathbb{R}$, where $X \subset V$ is an arbitrary set of vectors and $Y \subset V$ is a convex set of vectors, such that $f(x, y)$ is a convex function with fixed x i.e., $f(x, \cdot)$ is a convex function for each x. Then the function $h : Y \mapsto \mathbb{R}$ given by $h(y) = \sup_{x \in X} f(x, y)$ is also a convex function. Additionally, show that if $f(x, y) =< x, y >$, then the resultant h is a conic function.

5. Starting from the Jensen's inequality prove the AM-GM inequality. After proving this try proving the Holder's inequality from Jensen's inequality. If you dont get the second proof, then google. The proofs are quite instructive. Subsequent to this exercise prove the same inequalities starting from the Fenchel's inequality.

6. Find conjugate (f^*) of f where $f(\mathrm{x})$ is given by: i) $f(x) = 3x^2$ ii) $f(\mathrm{x}) = \max_i x_i$ iii) $f(\mathrm{x}) = -(\Pi_i x_i)^{\frac{1}{n}}, \mathrm{x} > 0$.

7. Compute the sub-differential set at all points for the function $h = \max(f, g)$ in terms of sub-differentials for f and g.

8. Show that the supporting hyperplane at any relint point for a closed convex function cannot be vertical. Do this using the Lipschitz continuity result.

9. Show that if there is a real-valued function defined over a convex domain that satisfies all the sub-gradient inequalities centered at points in the domain, then the function must be convex.

10. Show that if there is a real-valued continuous function defined over a convex domain that satisfies all the sub-gradient inequalities centered at points in the rel.int. of the domain, then the function must be convex.

11. Let f be a convex function and L_α is a level-set of it over which the function is not a constant. Show that the normal cone of L_α at some x^* on the rel.bnd. of L_α is the conic-hull of the negative of the sub-differential set at the same point[5].

[5] After proving this, firstly observe the result in the special case when f is differentiable at that point. Realize that this special case result can also be proved starting from the directional derivative definition

12. Show that the inner-product induced norm is always self-dual (in that inner-product space). Hint: Cauchy-Schwartz inequality.

13. Let $x \in \mathbb{R}^n$ and $M \succ 0$ (a pd matrix). Prove that the matrix-fractional function $f(x, M) = x^\top M^{-1}x$ is a convex function. Hint: Use the Schur-complement lemma.

14. Let $\mathcal{V} = (V, +, \cdot, \langle\rangle)$ be an inner-product space and $C \subset V$ be a convex set. Let $f : C \mapsto \mathbb{R}$ be a continuous function. Assume that the sub-gradient at $x_0 \in relint(C)$ is $\nabla f(x_0)$. Then f is strictly convex if and only if $f(x) > f(x_0) + \langle \nabla f(x_0), (x - x_0)\rangle \ \forall x \in dom(f)$. i.e., strict convexity is characterized by sub-gradient strict inequality satisfaction centered at any relint point of the domain.

15. In the context of the above with $V = \mathbb{R}^n$, show that if the Hessian matrix is pd at all points in C, then f is strictly convex.

3 Theory: Convex Programs

3.1 Basics and Optimality Conditions

1. Answer whether the following optimization problems can be posed as a convex program. Justify your answers.

   (a) Consider the problem where you need to find the affine function whose sum of squared distances from given m points in an n dimensional space is the least (this is called the least squares problem).

   (b) Suppose you have a discrete random variable X which takes n values. Consider the problem of finding that probability mass function, from the set of all possible probability mass functions for X, which has maximum entropy.

   (c) Consider the standard inner-product space of matrices and let $\mathcal{P}_1$ and $\mathcal{P}_2$ be any two polyhedral sets in it. Consider the problem of finding the distance[6] between the two polyhedra.

[6] Given two sets, the distance between them is the infimum of all distances between points in them.

2. Taking hints from the proof sketch given in the lecture notes, prove theorem 13.0.9 and 13.0.10 in the lecture notes. Note how students oversimplified the proof of 13.0.9 during the lecture.

3. Consider a particle which is constrained to live in the open (standard) second-order cone in $\mathbb{R}^{n+1}$ i.e., the cone: $\left\{ \begin{bmatrix} \mathbf{x} \\ y \end{bmatrix} \mid \mathbf{x} \in \mathbb{R}^n, y \in \mathbb{R}, \|\mathbf{x}\|_2 < y \right\}$. Inside this space, a weird gravitational force f acts on this particle, which depends on its position: $f(\mathbf{x}, y) = \mathbf{a}^\top \mathbf{x} + by + \log(y^2 - \mathbf{x}^\top \mathbf{x})$ where $\mathbf{a} \in \mathbb{R}^n, b \in \mathbb{R}$ are given (i.e., $\mathbf{a}, b$ are parameters of the gravitational field). For what values of these parameters of the field can you find point(s) at which the gravitational force is maximum ? For such parameters, what are the points of maximum gravity and what is the maximum gravity in this space ? Now, can you find the gravitational field setting (i.e., find $\mathbf{a}, b$) such that the maximum gravity in the space is minimized ?

4. Using your knowledge about optimality conditions in problems minimizing quadratic functions of the form $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x}$, prove the Schur complement lemma: the matrix $\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{C}^\top & \mathbf{D} \end{bmatrix}$, where $\mathbf{B}, \mathbf{C}$ are symmetric, is positive definite[7] if and only if both the matrices $\mathbf{B}$ and $\underbrace{\mathbf{D} - \mathbf{C}^\top \mathbf{B}^{-1}\mathbf{C}}_{\text{Schur Complement of } \mathbf{B} \text{ in } \mathbf{A}}$ are positive definite.

5. Using KKT conditions:

   (a) Show that of all possible rectangles inscribed in a given a circle, that with the maximum area is a square.

   (b) Compute $min_x \sum_i a_i x_i^{2r}$ s.t. $\sum_i x_i^{2s} = 1$. Here, all a_i are positive and $0 < s < r$.

   (c) Show that $P_2 = \sqrt{2\rho\sigma P_1}$, where $\mathbf{Q}$ be a positive-definite matrix, ρ, σ are positive reals and

$$P_1 = \max_{\mathbf{x} \in \mathbb{R}^n} \quad \mathbf{c}^\top \mathbf{x} - \tfrac{\rho}{2} \mathbf{x}^\top \mathbf{Q}\mathbf{x}$$
$$\text{s.t.} \qquad \mathbf{A}\mathbf{x} \leq 0$$

[7] $\mathbf{A}$ is positive-definite iff $\mathbf{x}^\top \mathbf{A}\mathbf{x} > 0$ for all non-zero $\mathbf{x}$. Also, a positive definite matrix is always invertible.

$$
\begin{aligned}
P_2 = \max_{\mathbf{x}\in\mathbb{R}^n} & \quad \mathbf{c}^\top\mathbf{x} \\
\text{s.t.} & \quad \mathbf{A}\mathbf{x} \leq 0,\ \mathbf{x}^\top\mathbf{Q}\mathbf{x} \leq \sigma
\end{aligned}
$$

6. Problems 5.26, 5.28, 5.30 (hint: spectral functions) in boyd's book.

3.2 Duality

1. It is believed that two variables $x \in \mathbb{R}^n$ and $y \in \mathbb{R}$ are related by $y = f(x)$, where f is an affine function. An experiment was conducted to obtain $m(> n)$ pairs (x_i, y_i). Though it must happen that all these m points must lie on a hyperplane (because f is affine), because of experimental errors this may not be true. So it is desirable to obtain that f for which $\|e(f)\|_p$ is minimum where $e(f) = [y_1 - f(x_1)\dots y_m - f(x_m)]^\top$ and $\|\cdot\|_p$ is the p-norm with $p \geq 1$ (p might also be ∞). Express this as a convex program (for each value of p). Write a dual of each of these programs. In case $p = 2$ and when it is assumed that the set of all x_i span $\mathbb{R}^n$, one can obtain an analytic expression for the optimal solution using the optimality conditions learned in this course. Please find the expression for the optimal solution[8]. Now, consider another situation where $m < n$ observations were made and suppose that the points (x_i, y_i) do lie on a hyperplane. It is easy to see that in this case there might be multiple f such that $y_i = f(x_i)\ \forall\ i = 1, \dots, m$. In such a case it is desirable to find an $f(x) = a^\top x - b$ with the least $\|a\|_p$. Now express this problem as a convex program and write its dual (for each p). Again in case $p = 2$ and assuming the set of x_i form a linearly independent set, it turns out that the optimal solution has an analytical expression. Find this[9] using your knowledge about optimality conditions. Everywhere in this problem use Conic-duality to write-down duals[10].

[8] This problem is called as least-squares fit and the optimal solution here motivates the definition of the left inverse of a matrix.

[9] This problem is called as min-norm fit and the optimal solution here motivates the definition of the right inverse of a matrix. Infact, in the general case where nothing is known (about m, n and/or the set of x_i), it is desirable to minimize a weighted sum of both the norm of a and least squares error (i.e., $\|e(f)\|_2$) leading to the min-norm least-squares fit and correspondingly the definition of psuedo-inverse of a matrix.

[10] You will later on realize that conic-duality is the easiest way out here.

2. Consider the maximal separation of two convex sets problem discussed in lecture, where we started with ∞-norm bound on the separating vector. Now consider the same problem with a p-norm ($1 \leq p \leq \infty$) bound on the separating vector instead. Is this a conic program? If so write down the conic-dual. Is this the dual you expect? If not, simplify the dual and does it then look intuitive.

3. Write down the problem of finding the minimum radius enclosing sphere for a given set of points in $\mathbb{R}^n$ as an SOCP. Write down its conic-dual. Aditya's intuition was that the dual will be the problem of finding the maximally distant points in the set. Is his intuition correct?

4. Write down the problem of finding the maximally separating hyperplane for two sets of spheres (assume some locations and radii for the spheres in each set) as an SOCP. Write down its conic-dual. Arun's intuition was that the dual will be the problem of minimizing distance between convex-hulls of these sets. Is his intuition correct?

5. Can the epi-graph of the function $f(x, y) = \frac{x^\top x}{y}, y > 0$ be described using SOC constraints? If not, can the closure of this set be expressed using SOC constraints?

6. Can the epi-graph of the function equal to negative of the geometric of two non-negative numbers i.e., $f(x, y) = -\sqrt{xy},\ x \geq 0,\ y \geq 0$ be described using SOC constraints?

7. Quadratically Constrained Quadratic Programs (QCQPs) are OCPs of the following form:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & x^\top A x + b^\top x + c, \\ \text{s.t} \quad & x^\top A_i x + b_i^\top x + c_i \leq 0,\ \forall\ i = 1, \ldots, m. \end{aligned}$$

Ofcourse this OCP is convex if and only if all A, A_i are psd. Show that the convex QCQP can be posed as an SOCP. Also Write down the conic-dual of this (convex) QCQP.

8. Complete the two problems taken-up in the lecture:

(a) Pose the problem of finding the smallest ellipsoid enclosing a set of points as an SDP. Write down its conic-dual. Does dual match the intuitive dual problem here?

(b) Pose the problem of finding the minimum condition number ellipse which seperate two sets each with finite number of points as an SDP. Write down its dual problem.

9. Show that LPs, strictly convex QPs[11], SOCPs are all self-dual. A program with a given form is called self-dual, if a dual is a problem in the form of the original problem. Note that in these cases, one can also write down the dual of the dual, as the dual is itself in primal form, using the same duality scheme as that with which the dual is obtained. In these cases show that the dual of the dual is equal to the primal.

10. Write down the Lagrangian dual and the conic dual of a convex QCQP. Now, both these duals should be equal as they are individually equal to the given primal QCQP. Now, just starting with the duals can you show they are equal (without using the fact that they are individually equal to the primal)?

11. Using Lagrangian duality, show that the problem of maximally separating (use Euclidean distance as distance) two sets, each containing finite number of points is same as that of minimizing (Euclidean) distance between convex hulls of those sets.

12. Show that the SDP-dual of the Lagrangian dual of a (possibly nonconvex) QCQP is the Shor relaxation of the same QCQP.

[11] QP with strictly convex objective is called a strictly convex QP.