CS 709 Problem-set

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. I can optionally correct and evaluate your problem-set solutions. Please drop by during office hours to get your solutions evaluated.

1 Theory: Convex Analysis

1.1 Domain: Hilbert Space

1. Verify whether the following are linear sets. For cases where the set is linear, provide a basis, a dual basis and determine the dimensionality:
   (a) Set of all bisymmetric matrices of size n
   (b) Set of all $n \times n$ Toeplitz matrices.
   (c) Set of all $n \times n$ diagonally dominant matrices.
   (d) Set of all doubly stochastic matrices of size n.
2. Show that at the end of the procedure described in lectures for reducing a spanning set to smaller sets, one would be left with a linearly independent set.
3. Show that the dimension of a subspace in a vector space is less than or equal to that of the vector space.
4. Show that the set of all functions $f : \mathbb{R} \mapsto \mathbb{R}$ such that $\int f(x)^2 \text{ dx}$ is finite forms a vector space with the usual point-wise $+$ and $\cdot$.

5. Consider a vector $v \in V$ in a inner-product space and a subspace $S \subset V$. Let an orthogonal basis of S be $\{v_1, \ldots, v_m\}$. Compute an expression[1] for $P_S(v)$. Show that $v - P_S(v)$ lies in the orthogonal complement of S.

6. Provide a dual basis for the linear set: $LIN(X)$ where $X = \{[1\ 1\ 1\ 1]^\top, [1\ 1\ -1\ \ -1]^\top\}$.

7. 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) $\{\|x_n - y_n\|\} \to \|x - y\|$

1.2 Subsets of Hilbert spaces

8. If S_1 and S_2 are two linear sets in a vector space $\mathcal{V}$, then show[2] that $S_1 + S_2 = LIN(S_1 \cup S_2)$.

9. Show that complement of an open set is closed and vice-versa[3].

10. Let $\{S_\lambda \mid \lambda \in \Lambda\}$ be a (possibly uncountable) collection of closed sets. Show that $\cap_{\lambda \in \Lambda} S_\lambda$ is a closed set[4]. Also, show that whenever the index set Λ is finite, then $\cup_{\lambda \in \Lambda} S_\lambda$ is a closed set.

11. Show that A^c is not affine whenever A is affine.

[1] Expression involving the basic operations, v and $\{v_1, \ldots, v_m\}$. You should NOT use any optimization theory results other than perhaps school day knowledge about minimizing a quadratic function of a single variable.

[2] This is alternate proof of the fact that sum of two linear sets is linear.

[3] This could have been alternate definition of closed/open-ness.

[4] Through DeMorgan's laws and the above complementarity result of closed and open-ness, we get that (possibly uncountable) union of open sets is open.

12. Consider the set $S = \left\{ \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2\sqrt{3}} \\ \sqrt{\frac{2}{3}} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2\sqrt{3}} \\ \sqrt{\frac{2}{3}} \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{\sqrt{3}} \\ \sqrt{\frac{2}{3}} \end{bmatrix} \right\}$. Let $K = CONIC(S)$. Now compute (simplify) and plot K^*. Also, provide a non-trivial conicly spanning set of K^*.

13. Prove the following assuming K_1, K_2 are cones:

    (a) $(K_1 \cap K_2)^* = K_1^* + K_2^*$ (Dubovitski-Milutin lemma)[5].

    (b) Suppose $K_1 \subset V_1$ is lying in a Hilbert space $\mathcal{V}_1$ and $K_2 \subset V_2$ is lying in another Hilbert space $\mathcal{V}_2$. Show that $K = K_1 \times K_2$ lying in $\mathcal{V}_1 \oplus \mathcal{V}_2$ is a cone.

14. Given a convex set C of dimensionality n, show that it contains a simplex of same dimension in it.

15. Prove that projection of a cone, polyhedral cone, convex set, polytope, polyhedron onto a linear set is a cone, polyhedral cone, convex set, polytope, polyhedron respectively.

16. Consider $S \subset \mathbb{R}^n$ that contains all vectors that have an entry 1 at exactly one of the co-ordinates and zero at others. Note that $|S| = n$. Provide the dual description of $CONV(S \cup \{0\})$ (here 0 denotes the zero vector/origin).

17. Let $\mathcal{V} = (V, +_V, \cdot_V, \langle\rangle_V)$ and $\mathcal{W} = (W, +_W, \cdot_W, \langle\rangle_W)$ be two (finite dim.) Hilbert spaces. A mapping $f : V \mapsto W$ is called a linear transformation iff $f(\lambda_1 \cdot_V v_1 +_V \lambda_2 \cdot_V v_2) = \lambda_1 \cdot_W f(v_1) +_W \lambda_2 \cdot_W f(v_2)\ \forall\, v_1, v_2 \in V, \lambda_1, \lambda_2 \in \mathbb{R}$. f is said to be an affine transformation if the above equality holds for any $\lambda_1 + \lambda_2 = 1$. These definitions are natural extensions of the notions of linear, affine functions defined in lectures. Consider the following:

    (a) Let $C \subset V$ be a convex set and $f : V \mapsto W$ be an affine transformation. Show that the affinely transformed set defined as (abuse of notation) $f(C) \equiv \{y = f(x) \mid x \in C\} \subset W$ is itself convex. Also, if C is a cone and f is a linear transformation, then show that $f(C)$ is a cone.

[5] Needless to say, you can use only those results proved/used in lectures.

(b) Let $C \subset W$ be a convex set and $f : V \mapsto W$ be an affine transformation. Show that the affine pre-image set defined as (abuse of notation) $f^{-1}(C) \equiv \{x \in V \mid f(x) \in C\}$ is itself convex. Also, if C is a cone and f is a linear transformation, then show that $f^{-1}(C)$ is a cone.

Using these results, show that the following set (if non-empty) is a convex set[6]: $\{x \in \mathbb{R}^n \mid \|Px + q\| \leq r^\top x + s\}$. Here, P is a $m \times n$ matrix and $q \in \mathbb{R}^m, r \in \mathbb{R}^n, s \in \mathbb{R}$. A convex set of this form is called "second-order cone". Also, using the above results, show that $\{x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i A_i - B \succeq 0\}$ (if non-empty) is a convex set[7]. Here, all A_i and B are symmetric matrices of size m. A convex set of this form is called "cone of a Linear matrix inequality" (or LMI cone in short).

18. Suppose $\|\cdot\|$ is some norm in a Hilbert space that need not be the inner-product induced one. Let $\|\cdot\|_*$ be its dual norm. Then, show that the dual cone of $K = \{(x, y) \mid \|x\| \leq y\}$ is $K^* = \{(x, y) \mid \|x\|_* \leq y\}$.

19. Show that the "hyperbolic set": $\{x \in \mathbb{R}^n \mid x \geq 0, \Pi_{i=1}^n x_i \geq 1\}$ is a convex set. In the special case $n = 2$, write down the dual description of this convex set. Do this exercise using two methods that assume f is the function whose epigraph is the hyperbolic set: i) guess the supporting hyperplane of f, then show it is indeed supporting and subsequently write the dual description of hyperbolic set. ii) write the conjugate of f and thus arrive at a dual description.

20. Let $S_1, S_2 \subset V$ be two disjoint sets. Let $K = \{(a, b) \mid \langle a, x\rangle \leq b \;\forall\; x \in S_1,\; \langle a, x\rangle \geq b \;\forall\; x \in S_2\}$. Show that K is a cone[8].

21. Provide a primal description of the dual cone of $K = \{Ax \mid x \in \mathbb{R}^n, x \geq 0\}$. Here, A is a $m \times n$ matrix.

22. Given a cone K that is pointed (cone is pointed iff it does not contain lines), show that K^* has volume (non-empty interior[9]). Using this, show that if K is a pointed cone with volume, then K^* is also pointed cone with volume.

---

[6] Infact, the set is simply a shifted cone. Thanks to Sachin Pawar for pointing this.
[7] Also a shifted cone.
[8] i.e., the set of all non-strict separators of two disjoint sets forms a cone.
[9] Here, were are not talking about relative interior.

1.3 Functions over Subsets of Hilbert spaces

23. Let $f : A \mapsto \mathbb{R}$ be a function, where A is an affine set. Let L_A be the linear set associated with A i.e., $L_A = A - A$. Prove the claim made in lecture: f is affine if and only if $\exists\, l \in L_A,\ b \in \mathbb{R} \ \ni\ f(x) = \langle l, x\rangle + b \ \forall\ x \in A$.

24. Consider the function $f : S^n \mapsto \mathbb{R}$ given by $f(M) =$ the absolute value of the maximum eigen value of M. Show that the function is conic using two approaches: i) recalling that $g(M) =$ max. eigen value of M is conic and $h(x) = |x|$ is conic and then $f = h \circ g$ ii) consider the epigraph and write it as an LMI cone.

25. Let K be the cone of psd matrices of size n. Let $g : K \mapsto \mathbb{R}$ be defined by $g(M) =$ max. eigen value of M. Show that $dom(g^*) = K$ and $g^*(M) = trace(M)\ \forall\ M \succeq 0$.

26. Consider the function: $f : \mathbb{R} \times \mathbb{R}^{++} \mapsto \mathbb{R}$ given by $f(x, y) = \frac{x^2}{y}\ \forall\ x, y \in \mathbb{R}, y > 0$. Show that this function is conic. Hint: Use epi-graph definition and write it as LMI cone. Irrespective of this result/proof, show that f is convex using the double derivative criteria[10] for convex functions.

27. Let $f : V \mapsto \mathbb{R}$ be a closed convex function. Let f^* be its conjugate function (for convinienice, assume f^* is well-defined for all $x \in V$)[11]. Show that the sub-differential set of f^* at 0 is exactly the set of all $x \in V$ such that $f(x) = min_{y \in V} f(y)$ i.e., the set of all minimizers of f.

28. This problem tries to show that unlike the fact that polar is extension of concept of dual, in the sense that polar of a cone is its dual cone, the concept of conjugate is not an extension of the notion of dual function. In other words, there are examples of conic functions whose dual function and conjugate function are different. Here is the example: consider the function $f(x) = \|x\|$ (the inner-product induced norm).

[10] i.e., consider 1-d restriction of this function and show it is convex.

[11] or you can use the trick of extending convex functions beyond the domain as explained in the book.

We know that dual function (dual norm) is $f^*(x) = \|x\|$. Show that the conjugate $f'(x) = 0 \ \forall \ \|x\| \leq 1$ (and $f'(x) = \infty$ if $\|x\| > 1$)[12].

29. For a random variable X with mean μ, median m and std. deviation σ, show that $|\mu - m| \leq \sigma$. Hint: Use Jensen's inequality[13].

30. Provide the dual description of the function $f : \mathbb{R}^n \mapsto \mathbb{R}$ given by $f(x) = \frac{1}{2}x^\top Qx$, where $Q \succ 0$; i.e., write f as maximum over a set of affine functions, using two methods: i) write the conjugate of f and then write f in terms of its conjugate. ii) find sub-gradient of f and then write f as maximum of all affine functions on the RHS of all sub-gradient inequalities.

31. Show that the set $C = \{x \in \mathbb{R}^n \mid x^\top Ax + b^\top x + c \leq 0\}$ (assuming it is non-empty), is convex whenever $A \succeq 0$. Hint: consider the function $f(x) = x^\top Ax + b^\top x + c$. This is convex because $g(x) = x^\top Ax$ is convex and $h(x) = b^\top x + c$ is convex. The given set is a level set of this convex function.

32. Show that the following functions are convex[14]:

    - $f : \mathbb{R}^n_{++} \mapsto \mathbb{R}$, such that $f(x) = \sum_{i=1}^n x_i \log(x_i) \ \forall \ x_i > 0$ (negative entropy).
    - $f : \mathbb{R}^n \mapsto \mathbb{R}$, such that $f(x) = \log(\exp x_1 + \ldots + \exp x_n)$
    - $f : S^n_{++} \mapsto \mathbb{R}$, such that $f(X) = \log(\det(X)) \ \forall \ X \succ 0$.

---

[12] Though the notions are different, they are related in the sense that the conjugate tries to capture the y-intercepts correpsonding to the dual function! Infact you can also show: if $f(x) = \|x\|$ (may NOT be the inner-product induced norm), then the conjugate $f'(x) = 0 \ \forall \ \|x\|_* \leq 1$ (and $f'(x) = \infty$ if $\|x\|_* > 1$); where $\|\cdot\|_*$ is the dual norm.

[13] You will also need to recall the fact that median is the best approximation of the random variable by a constant in terms of minimizing the expected absolute error in the approximation

[14] In all these cases, the easiest proof is considering 1-d restriction and showing it is convex using the double derivative criteria.

2 Theory: Convex Programs

2.1 The Basics

33. Pose the following problems as ordinary convex programs:

    (a) Let S_1, S_2 be two sets with finite cardinality in $\mathbb{R}^n$. Consider the problem of finding the $M-$norm i.e., $\|\cdot\|_M$ that minimizes the average distance (with this $M-$norm) between points in the same set while constraining that the distance (with this $M-$norm) between points belonging to different sets is atleast unity.

    (b) Let $S = \{v_1, \ldots, v_m\} \subset \mathbb{R}^n$ be given. Consider the problem of finding the smallest ellipsoid containing this set. Assume that $dim(S) = n$.

    (c) Let S_1, S_2 be two sets with finite cardinality in $\mathbb{R}^n$. Consider the problem of finding the most spherical ellipsoid that seperates the sets, i.e., one of them lies inside and the other outside the ellipsoid.

2.2 Optimality Conditions

34. Consider the system $Ax = b$, where A, b are given matrices of sizes $m \times n$ and $m \times 1$ respectively.

    (a) In the case $rank(A) = n < m$, one knows that the system may not be consistent and one looks for a least square solution. Using KKT conditions show that the least squares solution[15] is $\left(A^\top A\right)^{-1} A^\top b$.

    (b) In the case $rank(A) = m < n$, one knows that the system is consistent/feasible and multiple solutions exist. One may be interested in finding the solution to this system that has minimum length/norm. This is called the min. norm solution. Write this problem as an OCP and then using KKT conditions show that the min.norm solution[16] is $A^\top \left(AA^\top\right)^{-1} b$.

[15] $\bar{A} = \left(A^\top A\right)^{-1} A^\top$ is sometimes called as the left inverse of A as $\bar{A}A = I$.

[16] $\bar{A} = A^\top \left(AA^\top\right)^{-1}$ is sometimes called as the right inverse of A as $A\bar{A} = I$.

35. Consider the above problem of finding the smallest ellipsoid contaning some given points. Write down the KKT conditions and simplify. Then comment why the conditions make sense geometrically.

36. Problem 5.30 in Boyd's book.

2.3 Duality

37. Derive optimality conditions (analogous to KKT) and duals for the following (you may assume appropriate regularity conditions like solvability etc.):

    (a) Minimization of a convex function under conic constraint:

$$\begin{aligned} \min_{x \in \mathcal{X} \subset V} \quad & f(x), \\ \text{s.t.} \quad & b - \mathcal{A}(x) \in K \subset W, \end{aligned}$$

    where $f : \mathcal{X} \mapsto \mathbb{R}$ is a convex function, $\mathcal{A} : V \mapsto W$ is a linear function and K is a cone.

    (b) Minimization of a convex function under generalized conic constraint:

$$\begin{aligned} \min_{x \in \mathcal{X} \subset V} \quad & f(x), \\ \text{s.t.} \quad & \mathcal{G}(x) \in K \subset \mathbb{R}^m, \end{aligned}$$

    where $\mathcal{G}(x) = \begin{bmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{bmatrix}$, f, g_i are all convex functions $\mathcal{X} \mapsto \mathbb{R}$ and K is a cone. First of all, is this program convex?