The University of Texas at Austin
Department of Electrical and Computer Engineering

EE381V: Convex Optimization — Fall 2009

PROBLEM SET ONE SOLUTIONS

C. Caramanis — Due: Wednesday, September 9, 2009.

---

1. Let $f_n : \mathbb{R}^k \to \mathbb{R}$ converge pointwise and uniformly to a continuous function f. Let $\boldsymbol{x}_n \in \mathbb{R}^k$ converge to $\boldsymbol{x} \in \mathbb{R}^k$. Show that $f_n(\boldsymbol{x}_n)$ converges to $f(x)$.

   By the triangle inequality, $\|f_n(\boldsymbol{x}_n) - f(\boldsymbol{x})\| \le \|f_n(\boldsymbol{x}_n) - f(\boldsymbol{x}_n)\| + \|f(\boldsymbol{x}_n) - f(\boldsymbol{x})\|$.
   $\|f_n(\boldsymbol{x}_n) - f(\boldsymbol{x}_n)\| \to 0$ because $f_n \to f$ uniformly (stronger than point-wise is needed).
   $\|f(\boldsymbol{x}_n) - f(\boldsymbol{x})\| \to 0$ because f is continuous and $\boldsymbol{x}_n \to \boldsymbol{x}$. Therefore, for any $\varepsilon > 0$, there exists N_1, N_2 such that $n \ge N_1$ implies $\|f_n(\boldsymbol{x}_n) - f(\boldsymbol{x}_n)\| < \varepsilon/2$, and $n \ge N_2$ implies $\|f(\boldsymbol{x}_n) - f(\boldsymbol{x})\| < \varepsilon/2$. Thus for $n \ge \max\{N_1, N_2\}$, we conclude that $\|f_n(\boldsymbol{x}_n) - f(\boldsymbol{x})\| < \varepsilon$.
   Therefore, $\|f_n(\boldsymbol{x}_n) - f(\boldsymbol{x})\| \to 0$ and $f_n(\boldsymbol{x}_n) \to f(\boldsymbol{x})$.

2. If $C \subseteq U \subseteq \mathbb{R}^k$ is a compact set, and $f : U \to \mathbb{R}$ a continuous function, show that $f(C) \subseteq V$ is also compact; then show that f attains its minimum and maximum on C.

   Let $\{N_\alpha\}$ be an arbitrary open cover of $f(C)$.
   Since f is continuous, each of the sets $f^{-1}(N_\alpha)$ is open and forms an open cover of C.
   Since C is compact, this open cover has a finite subcover. i.e. There exists n such that $X \subset f^{-1}(V_{\alpha_1}) \cup \cdots \cup f^{-1}(V_{\alpha_n})$.
   Since $f(f^{-1}(M)) \subset M$ for every $M \subset V$, the image of the finite subcover of C is a finite subcover of $f(C)$.
   Therefore, $f(C) \subseteq \mathbb{R}$ is compact.

   In particular, there is a finite upper and lower bound. Since it is also closed, for any convergent subsequence, $y_n = f(x_n)$, with $y_n \to y$, we must have $y \in f(C)$, and in particular this shows that f attains its maximum and minimum on C.

3. Let $\{C_n\}_{n=1}^{\infty}$ be a countable collection of compact subsets of $\mathbb{R}^k$, endowed with the standard topology. Show that if $C_{n_1} \cap \cdots \cap C_{n_k} \neq \emptyset$, for any finite subcollection $(n_1, \ldots, n_k)$, then in fact
$$\bigcap_{n=1}^{\infty} C_n \neq \emptyset.$$
   Show by example that this need not hold if the sets C_i are not compact.

   Let $A_1 = C_1$, $A_2 = (A_1 \cap C_2)$, $A_n = (A_{n-1} \cap C_n)$, and so on. This gives a sequence of *nested* compact sets. By assumption, each A_n is nonempty. Take $x_n \in A_n$. Since each x_n is contained in A_1, and A_1 is compact, passing to a subsequence if necessary, we have that the sequence converges, $x_n \to x$. By the nested property, we have that each A_k contains all but finitely many of the x_n, and in particular $x \in A_n$ for all n. This is the common point we had to exhibit.

For an example of bounded but not closed sets where this fails, consider the sets $\{C_n\} = \left(0, \frac{1}{n}\right)$ which are not compact. Any finite intersection of these sets will not be empty, but $\bigcap_{n=1}^{\infty} = \emptyset$. For an example of closed sets where this fails, take $C_n = [n, \infty)$.

4. Dual Cones

   (a) Let $L \subseteq \mathbb{R}^k$ be a subspace. Define

   $$L^* = \{\boldsymbol{u} \,:\, \langle \boldsymbol{u}, \boldsymbol{v} \rangle \geq 0 \quad \forall \boldsymbol{v} \in L\}.$$

   Characterize the set L^*, and show that $(L^*)^* = L$.

   Since L is a subspace, $\boldsymbol{v} \in L$ iff $-\boldsymbol{v} \in L$, therefore we can write, equivalently,

   $$L^* = \{\boldsymbol{u} \,:\, \langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0 \quad \forall \boldsymbol{v} \in L\}.$$

   But this is exactly the definition of the orthogonal complement, so we have: $L^* = L^{\perp}$. Therefore $(L^*)^* = L$, a familiar property from linear algebra.

   (b) Let $\mathbb{S}^n$ denote the set of $n \times n$ symmetric matrices, and $\mathbb{S}^n_+ \subseteq \mathbb{S}^n$ the set of positive semidefinite symmetric matrices. Define an inner product on $\mathbb{S}^n$, by

   $$\langle A, B \rangle = \mathrm{Tr}(AB) = \sum_{i,j} A_{ij} B_{ij}.$$

   Compute $(\mathbb{S}^n_+)^*$, defined as in the first part of the question above.

   By the spectral theorem and the fact that matrices in $\mathbb{S}^n_+$ have nonnegative eigenvalues, we know that $A \in \mathbb{S}^n_+$ iff there exist nonnegative scalars μ_i and vectors $\boldsymbol{x}_i$ such that A is the sum of rank one matrices as follows:

   $$A = \sum_{i=1}^{k} \mu_i \boldsymbol{x}_i \boldsymbol{x}_i^\top.$$

   We thus have:

   $$\begin{aligned}
   (\mathbb{S}^n_+)^* &= \{B \in \mathbb{S}^n \,:\, \mathrm{Tr}(A^\top B) \geq 0 \ \forall A \in \mathbb{S}^n_+\} \\
   &= \{B \in \mathbb{S}^n \,:\, \mathrm{Tr}(B^\top \sum_i \mu_i \boldsymbol{x}_i \boldsymbol{x}_i^\top) \geq 0, \ \forall \mu_i \geq 0, \ \boldsymbol{x}_i \in \mathbb{R}^n\} \\
   &= \{B \in \mathbb{S}^n \,:\, \sum_i \mu_i \boldsymbol{x}_i^\top B \boldsymbol{x}_i \geq 0, \ \forall \mu_i \geq 0, \ \boldsymbol{x}_i \in \mathbb{R}^n\} \\
   &= \{B \in \mathbb{S}^n \,:\, \boldsymbol{x}^\top B \boldsymbol{x} \geq 0, \ \forall \boldsymbol{x} \in \mathbb{R}^n\} \\
   &= \mathbb{S}^n_+.
   \end{aligned}$$

5. Convexity and Topology:

(a) Show that if $S \subseteq \mathbb{R}^d$ is compact, then so is $\operatorname{conv}(S)$.

By Caratheodory's Theorem, $d+1$ points are enough to characterize $\operatorname{conv}(S)$ for $S \subseteq \mathbb{R}^d$. So

$$\operatorname{conv}(S) = \left\{ \sum_{i=1}^{d+1} \alpha_i \boldsymbol{x}_i : \boldsymbol{x}_i \in S, \alpha_i \geq 0, \sum_{i=1}^{d+1} \alpha_i = 1 \right\}$$

Define a set $A = \left\{ \boldsymbol{\alpha} \in \mathbb{R}^{d+1} : \alpha_i \geq 0, \sum_{i=1}^{d+1} \alpha_i = 1 \right\}$
Define also a function $f : S^{d+1} \times A \to \mathbb{R}^d$ such that $f(\boldsymbol{x}, \boldsymbol{\alpha}) = \sum_{i=1}^{d+1} \alpha_i x_i$.
Because f is a continuous function with a compact domain, the image of f is compact. Since the image of f is $\operatorname{conv}(S)$, $\operatorname{conv}(S)$ is compact.

(b) Show that if $S \subseteq \mathbb{R}^2$ is closed, then $\operatorname{conv}(S)$ need not be closed.

Consider a set consisting of all the points of a line, plus one other point in $\mathbb{R}^d$. The set is closed. The convex hull of this set is the set of all the points in the plane between the point and the line, including the line and the point, but not including the line parallel to the given line containing the given point. This convex hull is not a closed set.

(c) Show that if $S \subseteq \mathbb{R}^d$ is open, then $\operatorname{conv}(S)$ is also open.

Let $z \in \operatorname{conv}(S)$. $z = \lambda x + (1-\lambda)y$ for $x, y \in S$.
Since S is open, there exist two balls, $B_1(x)$ and $B_2(y)$ with radii ϵ_1 and ϵ_2 respectively, which are contained in S. So $\lambda B_1(x) + (1-\lambda)B_2(y) \subseteq \operatorname{conv}(S)$.
Since the sum and scaling of open sets (defined the usual way) is open, $\lambda B_1(x) + (1-\lambda)B_2(y)$ is open.
Therefore, there exists a ball $B_3(z)$ with radius ϵ_3 such that $B_3(z) \subseteq \lambda B_1(x) + (1-\lambda)B_2(y) \subseteq \operatorname{conv}(S)$, and $\operatorname{conv}(S)$ is open.

6. Let $S \subseteq V$ be a set, and take two points $u, v \notin \operatorname{conv}(S)$. Show that if $u \in \operatorname{conv}(S \cup \{v\})$, and $v \in \operatorname{conv}(S \cup \{u\})$, then $u = v$.

   Suppose $u \neq v$. If $u \in \operatorname{conv}(S \cup \{v\})$ and $u, v \notin \operatorname{conv}(S)$, $\lambda v + (1-\lambda)x = u$ for some $x \in S$ where $0 < \lambda < 1$.
   So $v = \frac{1}{\lambda}u - \frac{(1-\lambda)}{\lambda}x$ and $v \notin \operatorname{conv}(S \cup \{u\})$, which is a contradiction.

7. Consider a univariate degree k polynomial $f(z)$, with $k \geq 1$, $f(z)$. Suppose $f(z)$ has roots $z_1, \ldots, z_k \in \mathbb{C}$ (not necessarily distinct) where we write $z_d = x_d + iy_d$. Show that if $\hat{z} = \hat{x} + i\hat{y}$ is a root of the $(k-1)$-degree polynomial $f'(z)$, then $(\hat{x}, \hat{y}) \in \operatorname{conv}\{(x_1, y_1), \ldots, (x_k, y_k)\}$.[1]

   Let $\hat{z}$ be a root of $f'(z)$. If w is also a root of $f(z)$, the result is trivial. Suppose $\hat{z}$ is not also a root of $f(z)$.

$$\sum_{i=1}^{k} \prod_{j \neq i} (\hat{z} - z_j) = 0$$

   So $\sum_{i=1}^{k} \prod_{j \neq i} \overline{(\hat{z} - z_j)} = 0$ for $\bar{z}$ denoting the complex conjugate of z.

---

[1] This exercise taken from Barvinok's *A Course in Convexity.*

Multiplying both sides by $\prod_{i=1}^{k}(\hat{z}-z_i)$ yields

$$\sum_{i=1}^{k}(\hat{z}-z_i)\prod_{j\neq i}|\hat{z}-z_j|=0$$

Expressing $\hat{z}$ as a combination of $\{z_i\}_{i=1}^{k}$,

$$\hat{z}=\sum_{i=1}^{k}\frac{\frac{1}{|\hat{z}-z_i|}}{\sum_{i=1}^{k}\frac{1}{|\hat{z}-z_j|}}z_i$$

which is a convex combination. Therefore, $\hat{z}$ is is $\operatorname{conv}(\{z_i\}_{i=1}^{k})$.

8. For A a $n\times n$ square matrix, recall that the characteristic polynomial is defined as:

$$p_A(\lambda)\overset{\triangle}{=}\det(\lambda I-A),$$

where I is the $n\times n$ identity matrix. Prove the Cayley-Hamilton theorem, that states that A satisfies its own characteristic equation, i.e., show that $p_A(A)$ is the all-zeros matrix.

Hint: using the Jordan canonical form should help...

Let $A=U^{-1}JU$ be the Jordan decomposition of A. Let t be the number of jordan blocks and $\{m_i\}_{i=1}^{t}$ be their multiplicities.

$$p_A(A)=(-1)^n\prod_{i=1}^{t}(A-\lambda_i I)^{m_i}=(-1)^n\prod_{i=1}^{t}U^{-1}(J-\lambda_i I)^{m_i}U=U^{-1}p(J)U$$

So $p_A(A)=0$ if and only if $p_A(J)=0$.
Each of the Jordan blocks of $p_A(J)=(-1)^n\prod_{i=1}^{t}(J-\lambda_i I)^{m_i}$ with be $(J_i-\lambda_i I)^{m_i}$, which is equal to zero due to the structure of the Jordan blocks.
Note: $J_i-\lambda_i I$ has zeros along the diagonal and ones on the upper subdiagonal, so $(J_i-\lambda_i I)^{m_i}=0$.

9. (Boyd and Vandenberghe, Ex. 2.10) Consider the set

$$C=\{x\in\mathbb{R}^n\ :\ x^\top Ax+b^\top x+c\leq 0\},$$

where $A\in\mathbb{S}^n$, $b\in\mathbb{R}^n$ and $c\in\mathbb{R}$.

   (a) Show that if $A\in\mathbb{S}_+^n$ (i.e., A is positive semidefinite) then the set C is convex.

   (b) Consider the set obtained by intersecting C with a hyperplane:

$$C_1=C\cap\{x\ :\ g^\top x+h=0\}.$$

   Show that C_1 is convex if there exists $\lambda\in\mathbb{R}$ such that $(A+\lambda gg^\top)\in\mathbb{S}_+^n$.

   (a) If $A\in\mathbb{S}_+^n$, then $x^\top Ax$ is a convex function of x, and hence the $x^\top Ax+b^\top x+c$ is convex, being the sum of convex functions. C is now a sublevel set of a convex function, which we have shown to be convex.

For the details: Let $x = \lambda x_1 + (1-\lambda) x_2$ for $x_1, x_2 \in C$ and $0 \leq \lambda \leq 1$.
For $A \in \mathbb{S}^n_+$, there exists a square root matrix F such that $F^\top F = A$.
$x^\top Ax = \|Fx\|^2$, so $x^\top Ax + b^\top x + c = \|Fx\|^2 + b^\top x + c$.
$\|Fx\|^2 = \|F(\lambda x_1 + (1-\lambda) x_2)\|^2 \leq \lambda \|Fx_1\|^2 + (1-\lambda) \|Fx_2\|^2$ by the properties of norms.
$b^\top x = \lambda b^\top x_1 + (1-\lambda) b^\top x_2$ because $b^\top x$ is linear.
$c = \lambda c + (1-\lambda) c$.
So, $(\lambda x_1 + (1-\lambda) x_2)^\top A (\lambda x_1 + (1-\lambda) x_2) + b^\top (\lambda x_1 + (1-\lambda) x_2) + c \leq \lambda \left(x_1^\top Ax_1 + b^\top x_1 + c\right) + (1-\lambda)\left(x_2^\top Ax_2 + b^\top x_2 + c\right) \leq 0$.
Therefore, $x \in C$ and C is convex.

(b) We have: $x^\top Ax + b^\top x + c = x^\top \left(A + \lambda gg^\top\right) x + b^\top x + c - \lambda x^\top gg^\top x$. Now, on the hyperplane $\{x : g^\top x = -h\}$, we have $\lambda x^\top gg^\top x = c - \lambda h^2$. If $(A + \lambda gg^\top) \in \mathbb{S}^n_+$, then we reduce to the scenario of part (a). The key point here is that while C may not be convex, the condition guarantees that on the given hyperplane, we have convexity.

10. In class we claimed that there are several natural operations on sets that preserve convexity. Convince yourselves that the following all preserve convexity, and then turn in only a proof of the last two.

(a) Cartesian product: If $C_1, \ldots, C_m \subseteq \mathbb{R}^d$ are convex sets, then the set

$$C \triangleq C_1 \times \cdots \times C_m = \{(x_1, \ldots, x_m),\ x_i \in C_i\}$$

is convex.

(b) Affine and inverse maps: For $C \subseteq \mathbb{R}^n$ convex, and $A : \mathbb{R}^n \to \mathbb{R}^m$ a linear operator (i.e., an $m \times n$ matrix) then **show that** the following two sets are convex:

$$\begin{aligned} D_1 &\triangleq \{Ax : x \in C\} \\ D_2 &\triangleq \{x : Ax \in C\}. \end{aligned}$$

(c) Minkowski sum: If $C_1, C_2 \subseteq \mathbb{R}^n$ are convex, **show that**

$$C \triangleq C_1 + C_2 \triangleq \{x = x_1 + x_2 : x_1 \in C_1, x_2 \in C_2\}$$

is convex.

Let $y = \lambda y_1 + (1-\lambda) y_2$ where $y_1, y_2 \in D_1$ and $0 < \lambda < 1$.
There exist $x_1, x_2 \in C$ such that $y_1 = Ax_1$ and $y_2 = Ax_2$. So $y = \lambda Ax_1 + (1-\lambda) Ax_2$.
Because A is linear, $y = A(\lambda x_1 + (1-\lambda) x_2)$.
Because $C \subseteq \mathbb{R}^n$ is convex, $\lambda x_1 + (1-\lambda) x_2 \in C$, so $y \in D_1$.
Therefore, D_1 is convex.

Let $x = \lambda x_1 + (1-\lambda) x_2$ where $x_1, x_2 \in D_2$ and $0 < \lambda < 1$.
Because A is linear, $Ax = \lambda Ax_1 + (1-\lambda) Ax_2$.
There exist $y_1, y_2 \in C$ such that $y_1 = Ax_1$ and $y_2 = Ax_2$. So $Ax = \lambda y_1 + (1-\lambda) y_2$. So $Ax \in C$ because C is convex.
Therefore, $x \in D_2$ and D_2 is convex.

Let $x_A, x_B \in C$. Because $C = C_1 + C_2$, $x_A = x_{A_1} + x_{A_2}$ where $x_{A_1} \in C_1$ and $x_{A_2} \in C_2$. Also $x_B = x_{B_1} + x_{B_2}$ where $x_{B_1} \in C_1$ and $x_{B_2} \in C_2$.
$\lambda x_A + (1-\lambda)\, x_B = \lambda\,(x_{A_1} + x_{A_2}) + (1-\lambda)\,(x_{B_1} + x_{B_2}) = (\lambda x_{A_1} + (1-\lambda)\, x_{B_1}) + (\lambda x_{A_2} + (1-\lambda)\, x_{B_2})$.
Because $\lambda x_{A_1} + (1-\lambda)\, x_{B_1} \in C_1$ and $\lambda x_{A_2} + (1-\lambda)\, x_{B_2} \in C_2$, because C_1 and C_2 are both convex.
Therefore, C is convex.

11. The Asymptotic Cone, also called the Recession Cone.

Unbounded convex sets are not that much more difficult to deal with than compact convex sets. The fundamental reason for this is that convex sets can only be unbounded in very particular ways. This exercise explores this.

Let $C \subseteq \mathbb{R}^d$ be a closed convex set. We define the directions of recession from a particular point $x \in C$ as follows:

$$C_\infty(x) \triangleq \{d \in \mathbb{R}^n \,:\, x + td \in C,\, \forall t > 0\} = \bigcap_{t>0} \frac{C - x}{t}.$$

(a) Show that $C_\infty(x)$ is a closed convex cone.

(b) Now, show that $C_\infty(x)$ does not depend on $x \in C$. One way to do this is to take $x_1, x_2 \in C$, and then show that for $d \in C_\infty(x_1)$ and $t > 0$, we have $x_2 + td \in C$. (Note that C is closed).

(c) Consider the convex set

$$C = \{(x, r) \in \mathbb{R}^2 \,:\, r \geq x^2\}.$$

Find C_∞.

Therefore, thanks to the result of part (b) of this exercise, we can simply define: The set C_∞ is defined as the *asymptotic cone* of the closed convex set C.

(a) Since C is closed and convex, and closedness and convexity are preserved under translation, scaling, and intersection, $\bigcap_{t>0} \frac{C-x}{t}$ is closed and C_∞ is closed and convex.
Let $d \in C_\infty(x)$ and $\lambda > 0$. $x + t(\lambda d) \in C$ because $\lambda t > 0$.
Therefore, $C_\infty(x)$ is a closed convex cone.

(b) Let $x_1,\ x_2 \in C$ and $d \in C_\infty(x_1)$. Then by definition, $x_1 + td \in C \quad \forall t > 0$. We will exhibit a sequence of points in C that converges to $x_2 + td$. Since C is closed, this implies that $x_2 + td \in C$, and hence $d \in C_\infty(x_2)$, which is what we want to show.

To this end, note that $x_n \triangleq x_1 + td + (1 - 1/n)(x_2 - x_1) \in C$, and that as $n \to \infty$, $x_n \to x_2 + td$, and $C_\infty(x)$ is independent of $x \in C$

(c) For $C = \{(x, r) \in \mathbb{R}^2 : r \geq x^2\}$,
$C_\infty = C_\infty((0,0)) = \{\alpha(0,1) : \forall \alpha \geq 0\}$

12. Show that a closed convex set C is compact if and only if $C_\infty = \{0\}$.

One direction is easy: if C is compact, then $C_\infty = \{0\}$. For the converse, suppose that C is not compact. Then there exists a sequence $x_k \in C$ with $||x_k|| \to \infty$. Let $d_k = x_k/||x_k||$. Since $||d_k|| = 1$, and the unit sphere is a compact set, there exists a subsequence of the d_k that converges. For notational convenience, again call this $\{d_k\}$, and let d be the limit point. Then we claim that $d \in C_\infty$. To see this, take any $x \in C$, $t > 0$. We must show that $x + td \in C$. Taking k large enough so that $||x_k|| > t$, we have:

$$x + td = \lim_k \{(1 - t/||x_k||)x + (t/||x_k||)x_k\},$$

and every element in the right is in the set C by convexity, hence $x + td \in C$ since C is closed.

13. The Relative Interior of a convex set $C \subseteq \mathbb{R}^d$.

    (Read the handout on relative interiors). Prove the following important refinement to Problem 3 above (Proposition 2 in the handout):

    Let $C \subseteq \mathbb{R}^k$ be a convex set. For $x_1 \in \operatorname{cl} C$, and $x_2 \in \operatorname{ri} C$, then

    $$\{\alpha x_1 + (1 - \alpha)x_2 \,:\, 0 \leq \alpha < 1\} \subseteq \operatorname{ri} C.$$

    Let $C \subseteq \mathbb{R}^k$ be a convex set. Let $x_1 \in \text{cl}(C)$ and $x_2 \in \text{ri}(C)$
    Let d be the dimension of C, the dimension of $\text{aff}(C)$
    Since $x_1 \in \text{cl}(C)$, there exists $x \in C$ such that $x \in B_{\epsilon_1}(x_1) \cap C \quad \forall \epsilon_1 > 0$. So $x \in B_{\epsilon_1}(x_1) \cap \text{aff}(C)$.
    Since $x_2 \in \text{ri}(C)$, there exists $\epsilon_2 > 0$ such that $B_{\epsilon_2}(x_2) \cap \text{aff}(C) \subseteq C$.
    $\text{conv}(\{x\} \cup B_{\epsilon_2}(x_2))$ is relatively open and contained in C
    $\{\alpha x_1 + (1 - \alpha)x_2 : 0 \leq \alpha < 1\} \subseteq \text{conv}(\{x\} \cup B_{\epsilon_2}(x_2))$ as $x \to x_1$
    Therefore, $\{\alpha x_1 + (1 - \alpha)x_2 : 0 \leq \alpha < 1\} \subseteq \text{ri}(C)$

14. (Optional) Fill in the missing proofs in the handout on the relative interior.