Problem 6

Set S is convex.

Given:

$$\begin{array}{rcl} u, v & \notin & \text{conv(S)} \\ u & \in & \text{conv}(S \cup \{v\}) \\ v & \in & \text{conv}(S \cup \{u\}) \end{array}$$

TPT (To Prove That) $u = v$.

Observe that since $u \notin S$, and $u \in \text{conv}(S \cup \{v\})$, it must be that $u = av + (1-a)s_1$, where $s_1 \in S$.

Similarly $v = bu + (1-b)s_2$. Substituting for v, we see

$$u = abu + a(1-b)s_2 + (1-a)s_1(1-ab)u = a(1-b)s_2 + (1-a)s_1$$

If $ab \neq 1$, we can divide through by ab, which means u can be written as a convex combination of points in S contradicting the assumption that u was not in conv(S).

Q.E.D.

7

This problem is best understood by first thinking about the case where all the roots of the polynomical $f(x)$ are real—the derivative's correspond to where $f(x)$'s slope is zero, and this has to be in between the largest and smallest zero of $f(x)$. (So all real roots of the derivative can be expressed as a convex combination of the smallest and largest zero of $f(x)$.)

Now we consider the general case. First observe that $f(z) = \prod_{i=1}^{n}(z = z_i)$.

Using the chain rule for derivatives, $(fg)' = f'g + fg'$, $(fgh)' = f'gh + fg'h + fgh'$, etc., we see that

$$\begin{aligned} f'(z) &= \sum_{i=1}^{n} \prod_{j \neq i, 1 \leq j \leq n} (z - z_j) \\ &= \sum_{i=1}^{n} \prod_{1 \leq j \leq n} (z - z_j)/(z - z_i) \end{aligned}$$

So if $\hat{z}$ is a zero of $f'(z)$, but not of $f(z)$, we see that

$$f'(\hat{z}) = \sum_{i=1}^{n} (\hat{z} - z_i)^{-1} = 0$$

Now comes the tricky part—remember $\bar{z}$ is the complex conjugate of z, i.e., $x \bar{+} iy = x - iy$, and $\bar{z} \cdot x$ is the square of the absolute value of z.

$$\sum_{i=1}^{n} (\hat{z} - z_i)^{-1} = \sum_{i=1}^{n} (\bar{\hat{z}} - \bar{z}_i)/((\hat{z} - z_i)(\bar{\hat{z}} - \bar{z}_i)) \quad \sum_{i=1}^{n} (\bar{\hat{z}} - \bar{z}_i)/|(\hat{z} - z_i)|^2$$

Since the derivative is 0 at $\hat{z}$, rewrite the above to get

$$\bar{\hat{z}}(\sum_{i=1}^{n} 1/|(\hat{z} - z_i)|^2) = \sum_{i=1}^{n} \bar{z}_i)/|(\hat{z} - z_i)|^2$$

Divide both sides by $\sum_{i=1}^{n} 1/|(\hat{z} - z_i)|^2)$ and take complex conjugates again, and you see $\hat{z}$ is a weighted sum of the roots of $f(x)$. Furthermore, the weights are positive, and sum to 1.

We're almost done—we need to consider the case where $\hat{z}$ is a zero of $f'(z)$, **and** of $f(z)$. This is trivial, since $\hat{z}$ again a convex combination of the roots of $f(z)$.

Q.E.D.

1 4

I'm having some trouble with this one, even showing that $(L^*)^* = L$; when it says L is a subspace of R^k, does that mean it's an arbitrary subset, or is it specifically a subspace in the sense of linear algebra?

2 9a

$$C = \{x \in R^n | x^T Ax + b^T x + c \leq 0\}$$

TPT if A is psd, then C is convex.

9

The fundamental fact here is that a psd matrix A has a decomposition $A = LDL^T$, where L is lower-diagonal, with the diagonal entries set to 1, and D is diagonal with nonnegative entries on the diagonal. (For pd matrix, the D's entries are strictly positive.)

The proof of this fact comes from Gaussian elimination.

Now set $y = L^T x$. Clearly this is a 1-1 linear transform (since L is invertible), its inverse exists. The image of a convex set under a linear transform is convex, so we'll study the set

$$\{y | (y^T Dy + b(L^T)^{-1} y + c < 0\}$$

This set "decomposes" nicely, it's captured by

$$\sum_{i=1}^{k} (d_i y_i^2 + e_i y + c/k) < 0$$

Note that each d_i is nonnegative.

Refer to $(d_i y_i^2 + e_i y + c/k)$ as s_i.

You can now turn to first principles, and show that if $q \in R^k$ and $r \in R^k$ satisfy the inequalities, then so does $\alpha q + (1-\alpha)r$. You only need to show that for each i, $s_i(\alpha q_i + (1-\alpha)r_i) < \alpha s_i(q_i) + (1-\alpha)s(r_i)$. This is just the test that a quadratic function is convex, since for a convex function f, $f^{-1}(C)$ is convex, when C is convex.

Part b is easy with this decomposition, simply write $x^T Ax + bx + c$ as

$$x^T(A+\lambda gg^T)x - x^T\lambda gg^T x + bx + c \quad = x^T A^* x - \lambda(g^T x)^T(g^T x) + bx + c \text{where } A* = A + \lambda gg^T \quad = x^T A^* x - \lambda(-h)^T($$

Which brings us to the form of part a. (Need some arguing about the regions we're intersecting and why it's ok to substitute for $g^T x + h = 0$.)

11

By defn, $C_\infty(x) = \cap_{t>0} \frac{C-x}{t}$ (Note that $C - x$ is not set difference, it's the application of the affine transform $f(u) = u - x$ to C.)

(a) To show that $C_\infty(x)$ is closed, convex, and a cone.

To show a set is closed, show its complement is open. Apply DeMorgans law to $\cap_{t>0} \frac{C-x}{t}$, it's complement is $\cup t > 0 (\frac{C-x}{t})^c$.

Now we need to show $(\frac{C-x}{t})^c$ is open. We know C is closed (given), this implies $C - x$ is closed (affine transform of closed set) and finally if a sequence of points in R^k $p_1, p_2, \ldots$ converges to say p, then $p_1/t, p_2/t, \ldots$ converges to p/t, so if A is closed so is A/t. So each $\frac{C-x}{t}$) is closed, therefore its complement is open.

Now to show $C_\infty(x)$ is convex. Convex sets are closed under intersection so we just need to show each $\frac{C-x}{t}$ is convex. This is easy, C is convex implies $C - x$ is convex. Scaling is affine, so A convex implies A/t is convex, so each $\frac{C-x}{t}$ is convex.

To show $C_\infty(x)$ is a cone, consider $d \in C_\infty(x) = \cap_{t>0} \frac{C-x}{t}$. Consider αd, we want to show it's in each $\frac{C-x}{t}$ for each t. But $\cap_{t>0} \frac{C-x}{t} = \cap_{t^*>0} \frac{C-x}{t^*/\alpha}$, since the map $t \mapsto t/\alpha$ is 1-1 onto, so we're done.

Q.E.D.

(b) To show that $C_\infty(x_1) = C_\infty(x_2)$ for $x_1, x_2 \in C$.

By symmetry, suffices to show that $C_\infty(x_1) \subset C_\infty(x_2)$. Consider $d \in C_\infty(x_1)$. Want to show that $x_2 + td \in C$ for all $t > 0$.

The way these proofs work is that you "sneak up" on the point you are trying to show is in the desired set through points in the desired set, use closedness of the set to show the point must be in it.

Consider the family of points $x_1 + td + (1-\alpha)(x_2 - x_1)$, for $0 \leq \alpha \leq 1$. Rewrite this expression to

$$\alpha(x_1 + \frac{t}{\alpha}d) + (1-\alpha)x_2$$

All these points are in C, since $x_1 + \frac{t}{\alpha}d$ is in C (since $d \in C_\infty(x_1)$), and C is convex.

Now as $\alpha \to 0$, the family of points converges to $x_2 + td$. Since C is closed, C must contain $x_2 + td$.

Q.E.D.

(c) This one's easy: we know $(0,0) \in C$, so we just compute $\cap_{t>0}(C/t)$.

$$C = \{(x,r)|r \geq x^2\} C/t = \{(x,r)|r/t \geq (x/t)^2 \equiv r \geq x^2/t\}$$

So we want to capture points (x,r) for which $r \geq x^2/t$ for all $t > 0$. This can only happen when $x = 0$, since otherwise we can pick t to be arbitrarily small. So the asymptotic cone is $\{(x,r)|x = 0, r \geq 0\}$.

3 12

Formally, a topological space is compact if for every collection of closed sets, there's a finite sub-collection that includes all the points in the original collection.

As far as R^k goes, a subset of R^k is compact iff it's closed and bounded. (For richer spaces, e.g., R^∞ things start to get complicated, but for R^k it's this simple.)

So to show C is compact iff $C_\infty = \{0\}$ all we need to do is show $C_\infty = \{0\}$ iff C is bounded and closed.

We don't need to bother with C being closed, that's given.

So first we show C bounded $\Rightarrow C_\infty = \{0\}$. Since C is bounded, it's contained in a large sphere S centered at the origin, with radius α. S is convex, closed, we can talk of S_∞, and furthermore $S_\infty \supset C_\infty$ (need 1 liner for this).

Now consider the diameter of the set $\cap_{\alpha > t > 0} S/t$. It's at most $2/\alpha$, so as α grows unboundedly, the set $\cap_{\alpha > t > 0} S/t$ shrinks and only contains 0, everything else is some finite distance from the origin and gets left out.

Now we show $C_\infty = \{0\}$ implies C is bounded.

We reason by the converse: let C be unbounded. Then we can find a sequence $e_1, e_2, \ldots$ such that the l_2 norm of e_n, $|e_n|$ tends to infinity.

We want to show a point that's not 0 and lies in C_∞. Consider a limit point l of the following (bounded) set of points: $\{q|q = e_i/(|e_i|)\}$. (Such a limit point exists because an infinite family of points in a bounded subset of R^k always has a limit point, from Bolzano Weierstrass.[1])

Let $c_1, c_2, \ldots,$ be a convergent subsequence of $e_1, e_2, \ldots$ that converges to l.

Let $x \in C$ and $t > 0$. Then observe that $x + tl$ is the limit of the following sequence of points: $\{(1 - \frac{t}{|c_i|})x + \frac{t}{|c_i|}c_i|\text{where}|c_i| > t\}$ (since $(1 - \frac{t}{|c_i|})$ tends to 0, and $\frac{t}{|c_i|}c_i$ tends to l)

Each point in this sequence is in C (they are convex combination of points in C), so their limit must be in C (by closure of C), so l is in C_∞.

Q.E.D.

[1] basically binary search of the bounded set

4

(a) This only makes sense if L is a subspace, i.e., is closed under addition and multiplication by scalars.

Every vector space has a basis—this is a deep theorem. For a finite dimensional subspace, you can construct an "orthonormal basis" pretty easily using a technique called Gramm-Schmitt orthnormalization.

An orthonormal basis of a m dimensional subspace $L \subset R^k$ is a set of m vectors $v_1, \ldots, v_m$ such that every pair is orthogonal (their dot product is 0) and each has unit norm.

Let $v_1, \ldots, v_m$ be such a basis for L. Complete this to an ortonormal basis for all of R^k, $V = v_1, \ldots, v_m, v_{m+1}, \ldots, v_k$.

We can write vectors in R^k in terms of their components wrt V.

So any member of L looks like $(a_1, a_2, \ldots, a_m, 0, 0, \ldots, 0)$, where the a_is are arbitrary. This in turn means that if $b = (b_1, \ldots, b_m, b_{m+1}, \ldots, b_k)$ and $\langle b, a \rangle \geq 0$ for all $a \in L$, then so does $b' = (b_1, \ldots, b_m, c_{m+1}, \ldots, c_k)$ for any choice of c_i's.

This in turn means that if a vector $d = (d_1, \ldots, d_m, d_{m+1}, \ldots, d_k)$ in $(L^*)^*$ is to have a nonnegative dot product with every single vector in L^*, it better have entries $d_{m+1}, \ldots, d_k$ set to zero—otherwise consider the element $q = (0, \ldots, 0, -d_{m+1}, \ldots, -d_l)$ which is certainly in L^*. Then the dot product of d with q is negative, contradicting the requirement on d.

So any $d \in (L^*)^*$ must lie in L.

To complete the proof we need to show that $L \subset (L^*)^*$, but this part is trivial.

(b) This is just Ex 2.24 in Boyd, page 52: the answer is not surprising, $(S^n_+)^* = S^n_+$.

13

x_2 lies in the interior, so there's an open ball of radius r centered at x_2 whose intersection with $aff(C)$ is is completely contained within C.

Clearly the closed ball with radius $r/2$ centered at x_2 has this property too.

Claim: the closed ball centered at $\alpha x_1 + (1-\alpha)x_2$ and radius $(1-\alpha)r/2$ has an intersection with $aff(C)$ that completely contained in C.

No time to rigorously prove, it's clear from the geometrical view.

From this it follows that the open ball at each $\alpha x_1 + (1-\alpha)x_2$ with radius $(1-\alpha)r/4$ has has an intersection with $aff(C)$ that completely contained in C.

This means that each such point lies in $ri(C)$.

QED