ACM 113 Homework 1

October 17, 2007

1 Problem 2.4

We wish to show that the following two sets A and B are equal:

$$A = \left\{ \sum_{i=1}^{k} \theta_i x_i : x_i \in S, \theta_i \geq 0, \sum_{i=1}^{k} \theta_i = 1 \right\}$$

$$B = \bigcap_{C_\alpha \text{ convex}, S \subseteq C_\alpha} C_\alpha$$

First we show $B \subseteq A$. For this, it is enough to show that A is a convex set containing S. Obviously $S \subseteq A$ since every element of S is a convex combination of elements of S. To show convexity of A, note that if $x, y \in A$, we have without loss of generality some set of x_i in S for which $x = \sum \alpha_i x_i$ and $y = \sum \beta_i x_i$ with $\alpha_i, \beta_i \geq 0$ and $\sum \alpha_i = \sum \beta_i = 1$. So then, for $0 \leq \theta \leq 1$, we have

$$\theta x + (1-\theta) y = \sum \gamma_i x_i$$

where $\gamma_i = \theta \alpha_i + (1-\theta)\beta_i$. Since $\gamma_i \geq 0$ and $\sum \gamma_i = 1$, this implies $\theta x + (1-\theta)y \in A$.
Next we show $A \subseteq B$. For this, we must show that $A \subseteq C_\alpha$ for every convex set C_α containing S. But for any $x \in A$, we have $x = \sum \theta_i x_i$ with $\sum \theta_i = 1$ and $\theta_i \geq 0$, where the x_i are in S. Since C_α contains S, it must contain the x_i, and by convexity of C_α, it must contain x as well.
(In the above proof, if we drop the requirement that the coefficients sum to 1, and intersect over conic sets, we get a characterization of the conic hull of S. Whereas if we drop the requirement that the coefficients are nonnegative, and intersect over affine spaces, we get a characterization of the affine hull of S. And if we drop all requirements on coefficients, and intersect over subspaces, we get a characterization of the linear hull, or span, of S.)

2 Problem 2.8

(a) When a_1 and a_2 are linearly independent, this is a parallelogram, hence a polyhedron. It is defined by the equation $Fx = 0$, where the columns of F form a basis for $\text{span}(a_1, a_2)^\perp$ and the inequalities:

$$-\left(a_1 - \frac{\langle a_1, a_2 \rangle}{\|a_2\|^2} a_2\right)^T a_1 \leq \left(a_1 - \frac{\langle a_1, a_2 \rangle}{\|a_2\|^2} a_2\right)^T x \leq \left(a_1 - \frac{\langle a_1, a_2 \rangle}{\|a_2\|^2} a_2\right)^T a_1$$

$$-\left(a_2 - \frac{\langle a_1, a_2 \rangle}{\|a_1\|^2} a_1\right)^T a_2 \leq \left(a_2 - \frac{\langle a_1, a_2 \rangle}{\|a_1\|^2} a_1\right)^T x \leq \left(a_2 - \frac{\langle a_1, a_2 \rangle}{\|a_1\|^2} a_1\right)^T a_2$$

When a_1 and a_2 are linearly dependent, and at least one of them is nonzero, S is a line segment, also a polyhedron. If we assume without loss of generality that $\|a_1\| \geq \|a_2\|$, then S is defined by $Fx = 0$, where the columns of F form a basis for $\text{span}(a_1)^\perp$ and these inequalities :

$$-\|a_1\|^2 \leq a_1^T x \leq \|a_1\|^2$$

When a_1 and a_2 are both zero, S is just a point.
(b) Yes this is a polyhedron. $A = -I, b = 0, F = ((1)^T, a^T, (a^2)^T), g = (0, b_1, b_2)$
(c) Since the ℓ_2 unit ball is self-dual, we can write this as $S = \{x \in R^n : x \geq 0, \|x\| \leq 1\}$. This is the intersection of the nonnegative orthant with the ℓ_2 unit ball. This cannot be represented as a polyhedron due to the curvature of the ball boundary.
(d) This is the unit cube $\|x\|_\infty \leq 1$. This is a polyhedron. $A = (I, -I), b = \mathbf{1}$.

3 Problem 2.9

(a) By definition of the ℓ_2 norm, one finds:

$$\|x - x_0\| \leq \|x - x_i\| \Leftrightarrow 2(x_i - x_0)^T x \leq \|x_i\|^2 - \|x_0\|^2$$

In other words, each constraint gives a half-space with separating hyperplane normal to $(x_i - x_0)$ and goes through the midpoint of x_i and x_0. Summarizing this in matrix form, we have $V = \{x : Ax \leq b\}$, where

$$A = 2\left((x_1^T, x_2^T, \ldots, x_K^T) - \mathbf{1}x_0^T\right)$$
$$b = (\|x_1\|^2, \|x_2\|^2, \ldots, \|x_K\|^2) - \|x_0\|^2\mathbf{1}$$

We use the book's notation of using parantheses to form column vectors, and $\mathbf{1}$ for a vector of ones.
(b) Assume that we have a solution x_0 to the strict inequality $Ax < b$. (We can state this as a linear program: $\max_{x,t} t$, subject to $Ax + t \leq b$ and solve with the simplex algorithm, for example.) So if $A = (a_1^T, a_2^T, \ldots, a_K^T)$, then the vectors a_i are equal to the normal vectors $(x_i - x_0)$ multiplied times some positive constant. That is, $\lambda_i a_i = (x_i - x_0)$. The positive constant can be chosen so that the midpoint between x_i and x_0 lies on the boundary. That is, $a_i^T\left(\frac{x_i + x_0}{2}\right) = b_i \Rightarrow a_i^T\left(x_0 + \frac{\lambda_i a_i}{2}\right) = b_i$, which is easily solved to give:

$$x_i = x_0 + 2\left(\frac{b_i - a_i^T x_0}{\|a_i\|^2}\right) a_i$$

(c) In general, not all polyhedral decompositions can be described as Voronoi regions. As a counterexample, consider the case $n = 1$ and the P_i are the intervals with endpoints $(-\infty, -3, -2, 2, 3, +\infty)$. In order for this decomposition to correspond to Voronoi regions, there would need to be numbers $x_1 < \ldots < x_5$ satisfying $\frac{x_i + x_{i+1}}{2} = y_i$ where $y = (-3, -2, 2, 3)$. One can check that this implies

$$x_1 = 2y_1 - 2y_2 + x_3 < y_1 \Rightarrow x_3 < -1$$
$$x_5 = 2y_4 - 2y_3 + x_3 > y_4 \Rightarrow x_3 > 1$$

So this cannot be solved.

4 Problem 2.12

(a) Intersection of 2 half-spaces, hence convex.
(b) Intersection of $2n$ half-spaces, hence convex.
(c) Intersection of 2 half-spaces, hence convex
(d) Expressed as an infinite intersection of half-spaces

$$\bigcap_{y \in S} \{x : \|x - x_0\| \leq \|x - y\|\} = \bigcap_{y \in S} \left\{x : 2(y - x_0)^T x \leq \|y\|^2 - \|x_0\|^2\right\}$$

thus this is convex.
(e) Not necessarily. Counterexample: Let $S = \{+1, -1\}$ and $T = \{0\}$.

(f) This can be expressed as the intersection of shifts of S_1:

$$\{x : x + S_2 \subseteq S_1\} = \{x : x + y \in S_1, \forall y \in S_2\} = \{x : x \in S_1 - y, \forall y \in S_2\} = \bigcap_{y \in S_2} S_1 - y$$

Since shifts of S_1 are convex, their intersection is as well.
(g) When $\theta = 1$, this is a half-space (as with a Voronoi region with only 2 points). Otherwise, this is a ball, since:

$$\|x - a\| \leq \theta \|x - b\| \Leftrightarrow \left\| x - a + \frac{\theta^2}{1 - \theta^2}(b - a) \right\| \leq \frac{\theta}{1 - \theta^2} \|b - a\|$$

so this is convex. To see this, let $y = x - a$ and $c = b - a$:

$$\|y\|^2 \leq \theta^2 \|y - c\|^2 = \theta^2 \left(\|y\|^2 - 2\langle y, c \rangle + \|c\|^2 \right)$$

$$\Leftrightarrow (1 - \theta^2) \left(\|y\|^2 - \frac{2\theta^2}{1 - \theta^2} \langle y, c \rangle + \frac{\theta^4}{(1 - \theta^2)^2} \|c\|^2 \right) \leq \left(\theta^2 + \frac{\theta^4}{1 - \theta^2} \right) \|c\|^2$$

$$\Leftrightarrow \left\| y + \frac{\theta^2}{1 - \theta^2} c \right\| \leq \frac{\theta}{1 - \theta^2} \|c\|$$

5 Problem 2.13

We claim the following are equivalent:

$$S_1 = \left\{ \sum_{i=1}^{m} \theta X_i X_i^T | X_i \in R^{n \times k}, \text{rank}(X_i) = k, \theta_i \geq 0 \right\}$$

$$S_2 = \left\{ X \in S_n^+ | \text{rank}(X) \geq k \right\} \cup \{\mathbf{0}\}$$

To show $S_1 \subseteq S_2$, suppose $X \in S_1$. Since it's the sum of symmetric matrices, X is symmetric. It is also positive semidefinite, since for $y \in R^n$,

$$\langle Xy, y \rangle = \sum_{i=1}^{m} \theta_i \langle X_i X_i^T y, y \rangle = \sum_{i=1}^{m} \theta_i \langle X_i^T y, X_i^T y \rangle = \sum_{i=1}^{m} \theta_i \|X_i^T y\|^2 \geq 0$$

Now suppose $X \neq \mathbf{0}$ and $\text{rank}(X) < k$. Then there would be an orthnormal set $\{y_j\}$ of $(n - k + 1)$ vectors in the null space of X. Then, as above, we have

$$0 = \langle Xy_j, y_j \rangle = \sum_{i=1}^{m} \theta_i \|X_i^T y_j\|^2$$

Since $\theta_i \geq 0$, all the terms in the above sum are nonnegative. And since $X \neq \mathbf{0}$, for some i', $\theta_{i'} > 0$, thus we have

$$\|X_{i'}^T y_j\|^2 = 0 \Rightarrow X_{i'}^T y_j = 0$$

But this is true for all $(n - k + 1)$ vectors y_j, so this implies $\text{rank}(X_{i'}) < k$, a contradiction. So we must have either $X = \mathbf{0}$ or $\text{rank}(X) \geq k$.
To show $S_2 \subseteq S_1$, suppose $X \in S_2$. The case $X = \mathbf{0}$ is trivial, so assume $\text{rank}(X) \geq k$. Consider a singular value decomposition of it. That is, $X = U \Lambda U^T$, where $UU^T = I$ and Λ is diagonal with at least k positive entries (since $\text{rank}(X) \geq k$). Clearly Λ can be written as a sum of diagonal matrices each of which has exactly k positive entries, with the rest zero.

$$\Lambda = \sum_i \Lambda_i$$

Now take Λ_i and delete the rows and columns with zeros on the diagonal to form $\tilde{\Lambda}_i \in R^{n\times k}$. Also delete the corresponding columns of U to form $\tilde{U}_i \in R^{n\times k}$. Thus

$$X = \sum_i U\Lambda_i U^T = \sum_i \tilde{U}_i\tilde{\Lambda}_i\tilde{U}_i^T$$

Let $X_i = \tilde{U}_i\tilde{\Lambda}_i^{(1/2)}$. We have $X_i \in R^{n\times k}$ with $\text{rank}(X_i) = k$ and $X = \sum_i X_i X_i^T$, so $X \in S_1$.

6 Problem 3.7

Suppose f is convex with $f(x) \leq B$ for all $x \in R^n$ and for some $x_1 \neq x_2$, $f(x_1) < f(x_2)$. Then let

$$x_3 = x_1 + \frac{2(B - f(x_1))}{f(x_2) - f(x_1)}(x_2 - x_1)$$

This means $x_2 = \theta x_3 + (1-\theta)x_1$, where

$$\theta = \frac{f(x_2) - f(x_1)}{2(B - f(x_1))}$$

Since f is convex and $0 < \theta < 1$, we have

$$f(x_2) - f(x_1) \leq \theta(f(x_3) - f(x_1))$$

Noting that $f(x_3) \leq B$ and using the definition of θ, we get

$$f(x_2) - f(x_1) \leq \frac{f(x_2) - f(x_1)}{2(B - f(x_1))}(B - f(x_1))$$

which implies $1 \leq \frac{1}{2}$, a contradiction. So in fact f must be constant.

7 Problem 3.16

(a) $f'' = e^x > 0$ so f is convex. $f' = e^x \neq 0$, so f is both quasiconvex and quasiconcave.
(b-f) Noting that cases (b-f) are all special cases of a general function (except the domains are chosen to guarantee that f is nonnegative and continuous.) First we compute the first and second derivatives, for the general case:

$$f(x_1, x_2) = x_1^p x_2^q$$

$$\nabla f(x) = (px_1^{-1}, qx_2^{-1})x_1^p x_2^q$$

$$\nabla^2 f(x) = \begin{bmatrix} p(p-1)x_1^{-2} & pqx_1^{-1}x_2^{-1} \\ pqx_1^{-1}x_2^{-1} & q(q-1)x_2^{-2} \end{bmatrix} x_1^p x_2^q$$

We can just check the sign of the trace and determinent to determine whether the eigenvalues of $\nabla^2 f(x)$ are nonnegative (implying convexity) or nonpositive (implying concavity) or neither.

$$\text{tr}\nabla^2 f(x) = \left(p(p-1)x_1^{-2} + q(q-1)x_2^{-2}\right) x_1^p x_2^q$$

$$\det \nabla^2 f(x) = pq(1-p-q)x_1^{2p}x_2^{2q}$$

For quasiconvexity, we also need to check:

$$y^T\nabla f(x) = 0 \Rightarrow y = \lambda(qx_1, -px_2) \Rightarrow y^T\nabla^2 f(x)y = -pq(p+q)\lambda^2 x_1^{2p}x_2^{2q}$$

This means that all of these functions are at least either quasiconcave or quasiconvex, depending on the sign of $-pq(p+q)$.

	p	q	$p(p-1)$	$q(q-1)$	$pq(1-p-q)$	$-pq(p+q)$
(b)	1	1	0	0	-1	-2
(c)	-1	-1	2	2	3	2
(d)	1	-1	0	2	-1	0
(e)	2	-1	2	2	0	2
(f)	α	$1-\alpha$	$-\alpha(1-\alpha)$	$-\alpha(1-\alpha)$	0	$-\alpha(1-\alpha)$

(b) $\det \nabla^2 f(x) < 0$, so $\nabla^2 f(x)$ has eigenvalues of opposite sign, hence f is neither convex nor concave. But it is quasiconcave.
(c) $\nabla^2 f(x)$ has positive trace and determinent, hence positive eigenvalues. So f is convex and quasiconvex.
(d) $\det \nabla^2 f(x) < 0$, so $\nabla^2 f(x)$ has eigenvalues of opposite sign, hence f is neither convex nor concave. However f is both quasiconcave and quasiconvex.
(e) $\nabla^2 f(x)$ has nonnegative trace and determinent, hence nonnegative eigenvalues. So f is convex and quasiconvex.
(f) $\nabla^2 f(x)$ has nonpositive trace and zero determinent, hence nonpositive eigenvalues. So f is concave and quasiconcave. (In the extreme cases $\alpha = 0$ or $\alpha = 1$ it is also both convex and quasiconvex.)