EE290N L. El Ghaoui 9/3/99

Homework Assignment #1 Solutions

1. *Variational characterization of eigenvalues.*

   (a) Introduce the eigenvalue decomposition of A

   $$A = U^T \Lambda U$$

   where U is a rotation ($U^T U = I$) and $\Lambda = \mathbf{diag}(\lambda_1, \ldots, \lambda_n)$ contains the eigenvalues of A. For every x, we have

   $$x^T A x = \sum_{i=1}^n \lambda_i z_i^2$$

   where $z = Ux$, so that

   $$\max \left\{ x^T A x \mid \|x\|_2 = 1 \right\} = \max \left\{ \sum_{i=1}^n \lambda_i z_i^2 \mid \sum_{i=1}^n z_i^2 = 1 \right\}.$$

   It is obvious that the maximum above is bounded above by $\max_i \lambda_i$, and that this number is attained by proper choice of z. Hence the result.

   (b) We have $x^T A x = \mathbf{Tr}\, AX$ where $X = xx^T$, hence we have

   $$\begin{aligned} \lambda_{\max}(A) &= \max \left\{ x^T A x \mid \|x\|_2 = 1 \right\} \\ &= \max \left\{ \mathbf{Tr}\, AX \mid \mathbf{Tr}\, X = 1, \ \ X \succeq 0, \ \ \mathbf{Rank} X = 1 \right\}. \end{aligned}$$

   We obtain an upper bound on $\lambda_{\max}(A)$ by relaxing the rank constraint:

   $$\lambda_{\max}(A) \leq \max \left\{ \mathbf{Tr}\, AX \mid \mathbf{Tr}\, X = 1, \ \ X \succeq 0, \ \ \mathbf{Rank} X = 1 \right\}.$$

   This inequality is attained by proper choice of the matrix X (select X to be of the form xx^T, where x attains the maximum in $\max \left\{ x^T A x \mid \|x\|_2 = 1 \right\}$).

2. *Schur complements.*

   (a) Consider the function

   $$f(x, y) = \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

   When A is non singular, we have:

   $$\min_x \ f(x, y) = y^T (C - B^T A^{-1} B) y = y^T S y.$$

The result hinges on the following simple observation. The function $f(x,y) > 0$ for every $(x,y) \neq (0,0)$ if and only if (a) $f(x,0) > 0$ for every x; (b) $f(x,y) > 0$ for every x and non zero y.

Property (a) is equivalent to $A \succ 0$; (a) implies in particular that A is non singular. Property (b) is equivalent to $\min_x f(x,y) > 0$ for every non zero y; when A is non singular, this is equivalent to $S \succ 0$. This establishes the result.

As a byproduct, we note that if we know $A \succ 0$, then the two nonstrict inequalities: $X \succeq 0$, and $S \succeq 0$, are equivalent.

Another way to prove the result is to first prove that $X \succ 0$ implies $A \succ 0$, and then to pre- and post-multiply matrix X by an invertible matrix, as follows:

$$\begin{bmatrix} A^{-1} & -A^{-1}B \\ 0 & I \end{bmatrix}^T \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} A^{-1} & -A^{-1}B \\ 0 & I \end{bmatrix} = \begin{bmatrix} A^{-1} & 0 \\ 0 & S \end{bmatrix},$$

which proves that if $A \succ 0$, then $X \succeq 0$ if and only if $S \succeq 0$.

(b) When $X \succeq 0$ is not necessarily definite, we must take into account the fact that A can be singular. Let us define the Schur complement as $S = C - B^T A^\dagger B$, where $A^\dagger$ is the *pseudo-inverse* of A.

Let U be an orthonormal matrix such that

$$U^T A U = \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix},$$

where $\Lambda \succ 0$ is diagonal (and invertible). Premultiplying X on the right by $\mathbf{diag}\,(U, I)$ and on the left by $\mathbf{diag}\,(U^T, I)$, we obtain a matrix that has the same "sign" as X:

$$\begin{bmatrix} U^T & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} U & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} \Lambda & 0 & B_1 \\ 0 & 0 & B_2 \\ B_1^T & B_2^T & C \end{bmatrix},$$

where we partitioned $U^T B$ row-wise conformably to A. The above matrix is positive semidefinite if and only if

$$\begin{bmatrix} \Lambda & B_1 \\ B_1^T & C \end{bmatrix} \succeq 0 \text{ and } B_2 = 0.$$

The equality condition can be written equivalently as $(I - A^\dagger A)B = 0$. The first inequality can be written equivalently as $C - B_1^T \Lambda^{-1} B_1 \succeq 0$, or $C - B^T A^\dagger B \succeq 0$. We have proven the following result.

$$X \succeq 0 \text{ if and only if } A \succeq 0,\ S := C - B^T A^\dagger B \succeq 0 \text{ and } (I - A^\dagger A)B = 0.$$

3. *Dual maps.*

(a) If $\Phi(X) = A^T X + XA$, where $A \in \mathbf{R}^{n\times n}$, and X is symmetric, then the dual map associates a symmetric, $n \times n$ matrix to a symmetric, $n \times n$ matrix S, in such a way that

$$\mathbf{Tr}\, S(A^T X + XA) = \mathbf{Tr}\, X\Phi^*(S)$$

for every $S = S^T$, $X = X^T$. Using the fact that $\mathbf{Tr}\, VW = \mathbf{Tr}\, WV$ for any pair of symmetric matrices, we obtain

$$\begin{aligned} \mathbf{Tr}\, S(A^T X + XA) &= \mathbf{Tr}\, SA^T X + \mathbf{Tr}\, SXA \\ &= \mathbf{Tr}\, X(SA^T) + \mathbf{Tr}\, X(AS) \\ &= \mathbf{Tr}\, X(AS + SA^T) \end{aligned}$$

The dual map is defined by

$$\Phi^*(S) = AS + SA^T.$$

(b) Here, the dual map takes a $r \times r$ symmetric matrix S and makes it a $n \times n$ symmetric matrix such that

$$\mathbf{Tr}\, SX_{11} = \mathbf{Tr}\, X\Phi^*(S)$$

It is straightforward to show that the dual map is

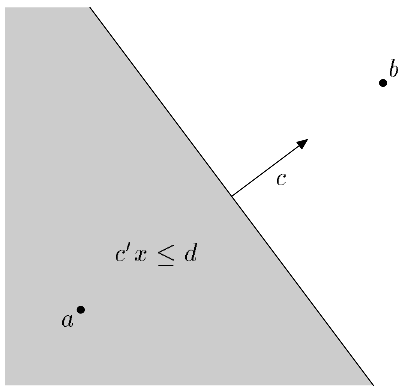
$$\Phi^*(S) = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}$$

where the size of the lower-right zero block is $(n-r) \times (n-r)$.

4. *Voronoi description of halfspace.* Since a norm is always nonnegative, we have $\|x-a\| \leq \|x-b\|$ if and only if $\|x-a\|^2 \leq \|x-b\|^2$, so

$$\begin{aligned} \|x-a\|^2 \leq \|x-b\|^2 &\iff (x-a)^T(x-a) \leq (x-b)^T(x-b) \\ &\iff x^Tx - 2a^Tx + a^Ta \leq x^Tx - 2b^Tx + b^Tb \\ &\iff 2(b-a)^Tx \leq b^Tb - a^Ta \end{aligned}$$

Therefore, we have that the set is indeed a halfspace, with $c = 2(b-a)^T$ and $d = b^Tb - a^Ta$. This makes good geometric sense: the points that are equidistant to a and b are given by a hyperplane whose normal is in the direction $b-a$.



5. (a) A slab is an intersection of two halfspaces, hence it is a convex set and a polyhedron.

   (b) As in part (a), a rectangle is a convex set and a polyhedron because it is a finite intersection of halfspaces.

   (c) A wedge is an intersection of two halfspaces, so it is convex and a polyhedron.

   (d) Let C $= \{ x \mid \|x - x_0\| \leq \|x - y\| \text{ for all } y \in S \}$. Given any $y \in S$, the set of points closer to x_0 than y is $C_y \triangleq \{ x \mid \|x - x_0\| \leq \|x - y\| \}$. In fact, C_y is the halfspace $\{ x \mid (y - x_0)^T x \leq (y^T y - x_0^T x_0)/2 \}$ because

$$\begin{array}{rrcl} \|x - x_0\| \leq \|x - y\| & \Leftrightarrow \quad (x - x_0)^T(x - x_0) & \leq & (x - y)^T(x - y) \\ & \Leftrightarrow \quad x^T x - 2x_0^T x + x_0^T x_0 & \leq & x^T x - 2y^T x + y^T y \\ & \Leftrightarrow \quad (y - x_0)^T x & \leq & (y^T y - x_0^T x_0)/2. \end{array}$$

   Therefore, $C = \bigcap_{y \in S} C_y$ and hence convex.

   (e) Let $N = \{ x \mid \mathbf{dist}\,(x, S) \leq \mathbf{dist}\,(x, T) \}$, N is not convex in general. Consider the following counter example in $\mathbf{R}$. Let $S = \{-1, 1\}$ and $T = \{0\}$, then $N = \{ x \in \mathbf{R} \mid x \leq -1/2 \text{ or } x \geq 1/2 \}$ is clearly not convex.

   (f) Let $C_t = \{ a \in \mathbf{R}^k \mid |p(t)| \leq 1 \}$ for a fixed $t \in [\alpha, \beta]$. Since $p(0) = 1$, we have $a_1 = 1$. Let $a^T = [1 \ a_2 \cdots a_k] \in C_t$, $\tilde{a}^T = [1 \ \tilde{a}_2 \cdots \tilde{a}_k] \in C_t$, and define $p(a, t) = 1 + a_2 t + \cdots + a_k t^{k-1}$. For an arbitrary $\lambda \in [0, 1]$,

$$\begin{array}{rcl} |\lambda p(a, t) + (1 - \lambda) p(\tilde{a}, t)| & \leq & \lambda |p(a, t)| + (1 - \lambda)|p(\tilde{a}, t)| \\ & \leq & \lambda + (1 - \lambda) \\ & = & 1. \end{array}$$

   Therefore, $\lambda a + (1 - \lambda)\tilde{a} \in C_t$, *i.e.*, C_t is convex. Since $C = \bigcap_{t \in [\alpha, \beta]} C_t$, C is convex.

6. *Conic hulls of outer products.*

   (a) First we note that xx^T is always symmetric PSD, and that every PSD matrix can be written (by eigenvalue/eigenvector decomposition) as $A = \lambda_1 x_1 x_1^T + \cdots + \lambda_n x_n x_n^T$, *i.e.*, a conic combination of dyads. Hence, the conic hull is the PSD cone. Next we note that if $x \in \mathbf{R}^n$, then we can generate a basis for the set of $n \times n$ symmetric matrices: if $x = e_i$ we have a zero matrix with the ith diagonal element being 1. The other elements of the basis can be generated by letting $x = e_i + e_j$, $i \neq j$. Now that we have a basis with $n(n+1)/2$ elements, it is clear that the dimension of the affine hull is the same.

   (b) Now we have $XX^T \succeq 0$ and $\mathbf{Rank}(XX^T) = k$. A positive combination of such matrices can have rank up to n, but never less than k: let A and B be positive semidefinite matrices with rank k. Suppose that $\mathbf{Rank}(A + B) = \ell < k$. Then we should have $n - \ell$ linearly independent vectors x_i such that $x_i^T (A + B) x_i = 0$. But this means $x_i^T A x_i + x_i^T B x_i = 0$ for all i, that is, $x_i^T A x_i = 0$ and $x_i^T B x_i = 0$

(because $A, B \succeq 0$). But, since $\ell < k$, we have $n-\ell > n-k$, that is, $\textbf{Rank}(A) < k$, a contradiction.

We concluded that $\textbf{Rank}(A+B) \geq k$ for any A, B such that $\textbf{Rank}(A, B) = k$ and $A, B \succeq 0$. If follows that the conic hull of the set is the set of positive semidefinite matrices of rank greater or equal to k (plus the matrix 0).

7. *Expanded and restricted sets.*

   (a) For the first part we consider two points $x_1, x_2 \in S_a$. From the set construction we also have two points $y_1, y_2 \in S$ such that $\|x_1 - y_1\| \leq a$ and $\|x_2 - y_2\| \leq a$. Now, if S is convex, then $\theta y_1 + (1-\theta) y_2 \in S$, and then

$$\begin{aligned} \|\theta x_1 + (1-\theta)x_2 - (\theta y_1 + (1-\theta)y_2)\| &= \|\theta(x_1 - y_1) + (1-\theta)(x_2 - y_2)\| \\ &\leq \theta\|x_1 - y_1\| + (1-\theta)\|x_2 - y_2\| \\ &\leq a, \end{aligned}$$

   so $\theta x_1 + (1-\theta)x_2 \in S_a$, proving convexity.

   For the second part we note that a point x is in S_a if and only if $B(x, a) \subseteq S$. Every point in this ball can be represented as $x_b = x + \alpha t$, where $\|t\| = 1$ and $0 \leq \alpha \leq a$. Consider now two points $x_1, x_2 \in S_a$. We then have

$$\theta x_1 + (1-\theta)x_2 + \alpha t = \theta(x_1 + \alpha t) + (1-\theta)(x_2 + \alpha t)$$

   But $x_1 + \alpha t \in S$ and $(x_2 + \alpha t) \in S$ if S is convex, so we conclude that the ball centered at any convex combination of points in S_a is contained in S – which means that any convex combination of points in S_a is in S_a, and we are done.

8. *More sets defined by distances to two points.*

   (a) In fact, the result that the set is not convex in the case $\alpha < 0$ is *not* true!! Two simple counterexamples: $n > 1$ (in this case the set is convex independent of the sign of α) and also $a = b$ (in which case the set is the whole space). I apologize for this.

   The correct statement of the exercise is

   > *Correct statement.* The set $\{x \| \|x - a\| + \alpha \leq \|x - b\|\}$ is convex if and only if $n = 1$, or $\alpha \geq 0$, or $-\alpha \geq \|a - b\|$.

   Let us first prove that, for $\alpha \geq 0$, the set $\{x \mid \|x - a\| + \alpha \leq \|x - b\|\}$ is convex. For a fixed vector δ, the set defined by $\|x - a + \delta\| \leq \|x - b\|$ is a half-space (problem 4 of this homework set). Now, using the triangle inequality, we can easily prove that

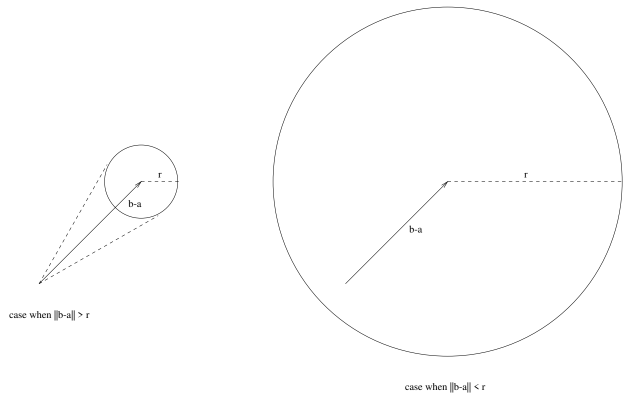
$$\max_{\delta} \{\|x - a + \delta\| \mid \|\delta\| \leq \alpha\} = \|x - a\| + \alpha.$$

   This shows that the set under consideration is the intersection of half-spaces, hence it is convex.

Now, the case $\alpha < 0$. By arguments similar to the case $\alpha \geq 0$, we can prove that the set is the *union* of a family of half-spaces, defined by

$$2(b - a - \delta)^T x \leq \|b - \delta\|^2 - \|a\|^2$$

where δ runs over the ball $\{\delta \mid \|\delta\| \leq -\alpha\}$. The union of half-spaces is convex if and only if it is itself a half-space, or if it is equal to the whole space. The first case means that the vectors $(b - a - \delta)$ are all paralell for every δ, $\|\delta\| \leq -\alpha$, and this can only happen for $n = 1$. The second case can only happen if at least two of these vectors are pointing in opposite directions, and this can only happen if $-\alpha \geq \|b - a\|$.



(b) If $\theta \leq 0$ the set is empty, thus convex. If $\theta > 0$ we square both sides to get the condition

$$(1 - \theta^2)x^T x - 2(a - \theta^2 b)^T x + a^T a - \theta^2 b^T b \leq 0.$$

Writing in matrix form, we get

$$\begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} (1-\theta^2)I & \theta^2 b - a \\ (\theta^2 b - a)^T & a^T a - \theta^2 b^T b \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \leq 0$$

An ellipsoid provided $1 - \theta^2 \geq 0$, that is, $\theta \in [0, 1]$. If $\theta > 1$ we can rewrite the matrix form as

$$\begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} (\theta^2 - 1)I & a - \theta^2 b \\ (a - \theta^2 b)^T & \theta^2 b^T b - a^T a \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \geq 0$$

Which describes the region *outside* an ellipsoid, hence a nonconvex set.

We conclude that the set is convex if and only if $\theta \leq 1$.

(c) Neither set is convex. For the ∞-norm counterexample in $\mathbf{R}^2$ take $a = (0, 0)$, $b = (5, 3)$, $x_1 = (2.5, 2.5)$ and $x_2 = (4.5, -1.5)$. Both x_1 and x_2 belong to the set, but $0.5x_1 + 0.5x_2$ does not.

For the 1-norm case, let $a = (0, 0)$, $b = (2, 4)$, $x_1 = (0, 3)$ and $x_2 = (3, 1)$. Again we have x_1, x_2 in the set, but the midpoint $0.5x_1 + 0.5x_2$ is not.