ECE 788 - Optimization for wireless networks
Midterm, Fall 2011

Please provide clear and complete answers.

PART I: Questions -

Q.1. (1 point) Calculate the distance between two parallel hyperplanes $\{x \in \mathbb{R}^n | a^T x = b_1\}$ and $\{x \in \mathbb{R}^n | a^T x = b_2\}$ as a function of (a, b_1, b_2). Based on the obtained expression, maximize the distance at hand with respect to vector a and scalars b_1, b_2 under the constraint that a must lie outside the ball $\mathcal{B}_2(0,1)$ (i.e., the unit ball with respect to the ℓ_2 norm) and b_1, b_2 are in the interval $[0,1]$.

Sol.: The line $x = \lambda a/||a||_2$ intersects the two hyperplanes at $\lambda_1 = b_1/||a||_2$ and $\lambda_2 = b_2/||a||_2$, respectively. From this, it follows that the distance between two hyperplanes is $|b_1 - b_2|/|||a||_2$. Therefore, the requested maximum is given by $\sup |b_1 - b_2|/|||a||_2$ under the constraints $||a||_2 > 1$ and $0 \leq b_1, b_2 \leq 1$. We thus get that $\sup |b_1 - b_2|/|||a||_2 = 1$.

Q.2. (1 point) Prove that the function

$$f(x,t) = -\log(t^2 - ||x||_2^2)$$

is convex on $\operatorname{dom} f = \{(x,t) | x \in \mathbb{R}^n,\ t > ||x||_2\}$ (Hint: Show first that $||x||_2^2/t$ is convex on $\{(x,t)|x \in \mathbb{R}^n,\ t > 0\}$).

Sol.: Observe that $f(x,t) = -\log(t^2 - ||x||_2^2) = -\log t - \log(t - ||x||_2^2/t)$. The first term, $-\log t$, is convex on $\mathbb{R}_{++}$. The second term instead can be expressed as the composition

$$-\log(t - ||x||_2^2/t) = h(g(x,t)),$$

where $h(y) = -\log y$ is concave on $\operatorname{dom} h = \mathbb{R}_{++}$ and $\tilde{h}(y)$ is non-increasing; and $g(x,t) = t - ||x||_2^2/t$ is convex on the set $\{(x,t)|x \in \mathbb{R}^n,\ t > 0\}$. This is the case because $||x||_2^2/t$ is the perspective function of $||x||_2^2$. The convexity of $h(g(x,t))$, and thus of $f(x,t)$, follows from the composition rules.

Q.3. (1 point) Find an equation describing a separating hyperplane for sets $\mathcal{S}_1 = \{x \in \mathbb{R}^2_{++} | x_1 x_2 \geq 1\}$ and $\mathcal{S}_2 = \{x \in \mathbb{R}^2 | x_2 \leq 0\}$. Repeat for $\mathcal{S}_1$ and $\mathcal{S}_2 = \{(0,0)\}$. Then, find an equation for the supporting hyperplane of set $\mathcal{S}_1$ at point $x = (1,1)$.

Sol.: The hyperplane $\{x|x_2 = 0\}$ separates the two sets with both $\mathcal{S}_2 = \{x \in \mathbb{R}^2 | x_2 \leq 0\}$ and $\mathcal{S}_2 = \{(0,0)\}$.
From simple geometric considerations $[-1,-1]$ defines the normal to the hyperplane that supports $\mathcal{S}_1$ at point $x = (1,1)$. Moreover, since $x = (1,1)$ is on the hyperplane we have that the desired equation is

$$\begin{aligned} -x_1 - x_2 &= [-1,-1][1,1]^T = -2 \\ \text{or } x_1 + x_2 &= 2. \end{aligned}$$

Q.4. (1 point) Given a non-convex function $f(x)$, show that the function $g(x)$ whose epigraph satisfies epi $g =$ conv epi f is such that $g(x) \geq p(x)$ for any other convex function $p(x)$ that satisfies $p(x) \leq f(x)$.

Sol.: Consider any convex function $p(x)$ that satisfies $p(x) \leq f(x)$. We have that epi p is convex and that epi $p \subseteq$ epi f by the definition of the epigraph of a function. But conv epi f is the smallest convex set that includes epi f and therefore we must have $g(x) \geq p(x)$.

PART II: Problems -

P.1. (2 points) Consider the set of all probability mass functions (pmfs) $\mathbf{p} = (p_1, p_2, ..., p_n)$ (i.e., $p_i \geq 0$ for all $i = 1, .., n$ and $\sum_{i=1}^{n} p_i = 1$). Denote as X the random variable with pmf $\mathbf{p}$ so that $p_i = \Pr[X = x_i]$ for some value $x_i \in \mathbb{R}$, for $i = 1, .., n$. Discuss whether the following sets are convex and, if so, discuss whether they are polyhedra.
a. The set of all pmfs such that $E[X^2] \leq 1$.
b. The set of all pmfs such that $\Pr[X \in \{x_1, x_2\}] = 0.1$.
c. The set of all pmfs whose entropy $-\sum_{i=1}^{n} p_i \log p_i$ is larger than 0.5 (by convention, we define $0 \log 0 = 0$).
d. The set of all pmfs whose entropy $-\sum_{i=1}^{n} p_i \log p_i$ is smaller than 0.5 (by convention, we define $0 \log 0 = 0$).
e. The set of all pmfs such that $var(X) \geq 1$.
f. The set of all pmfs such as the maximum probability of X is less than 0.3.

Sol.: a. This is the set $\{\mathbf{p} \in \mathbb{R}^n | \mathbf{p} \succeq \mathbf{0},\ \mathbf{1}^T\mathbf{p} = 1 \text{ and } \mathbf{a}^T\mathbf{p} \leq 1\}$, where $\mathbf{1} = [1, ..., 1]^T$ and $\mathbf{a} = [x_1^2, ..., x_n^2]^T$. It is thus a polyhedron.
b. This is the set $\{\mathbf{p} \in \mathbb{R}^n | \mathbf{p} \succeq \mathbf{0},\ \mathbf{1}^T\mathbf{p} = 1 \text{ and } p_1 + p_2 = 0.1\}$, which is a polyhedron.
c. This is the set $\{\mathbf{p} \in \mathbb{R}^n | \mathbf{p} \succeq \mathbf{0},\ \mathbf{1}^T\mathbf{p} = 1 \text{ and } -\sum_{i=1}^{n} p_i \log p_i \geq 0.5\}$. It is a convex set because it is the intersection of a polyhedron with the set $\{\mathbf{p} \in \mathbb{R}^n | \sum_{i=1}^{n} p_i \log p_i \leq -0.5\}$, which is a sublevel set of function $\sum_{i=1}^{n} p_i \log p_i$, which is convex on $\mathbb{R}_+^n$ (when the domain is extended as mentioned in the text).
d. This is the set $\{\mathbf{p} \in \mathbb{R}^n | \mathbf{p} \succeq \mathbf{0},\ \mathbf{1}^T\mathbf{p} = 1 \text{ and } -\sum_{i=1}^{n} p_i \log p_i \leq 0.5\}$. It is not a convex set. This can be seen by building a simple example with $n = 2$.
e. This is the set $\{\mathbf{p} \in \mathbb{R}^n | \mathbf{p} \succeq \mathbf{0},\ \mathbf{1}^T\mathbf{p} = 1 \text{ and } \mathbf{a}^T\mathbf{p} - (\mathbf{x}^T\mathbf{p})^2 \geq 1\}$, where $\mathbf{x} = [x_1, ..., x_n]^T$, since $var(X) = E[X^2] - E[X]^2$. It is a convex set because it is the intersection of a polyhedron with the set $\{\mathbf{p} \in \mathbb{R}^n | \mathbf{a}^T\mathbf{p} - (\mathbf{x}^T\mathbf{p})^2 \geq 1\}$, which is a superlevel set of function $\mathbf{a}^T\mathbf{p} - (\mathbf{x}^T\mathbf{p})^2$, which is concave ($-(\mathbf{x}^T\mathbf{p})^2$ is a concave quadratic function).
f. This is the set $\{\mathbf{p} \in \mathbb{R}^n | \mathbf{p} \succeq \mathbf{0},\ \mathbf{1}^T\mathbf{p} = 1 \text{ and } \max_i\{p_i\} \leq 0.3\}$. It is a polyhedron, since the inequality $\max_i\{p_i\} \leq 0.3$ is equivalent to the conditions $p_i \leq 0.3$ for all $i = 1, ..., n$.

P.2. (2 points) Consider the unconstrained problem of minimizing the function $f_o(x) = \frac{1}{2}x^TQx + q^Tx$. Vector q is given by $q = [1\ 1]^T$ and matrix Q is $Q = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$.

a. Using **only** the first- and second-order necessary conditions and the second-order sufficient condition for local optimality of general problems (i.e., **not** exploiting convexity-related results), what can be concluded about local minima? (Hint: Recall that to find the eigenvalues λ of a matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we can solve the equation $(a - \lambda)(d - \lambda) - bc = 0$).

b. Following the previous point and using also Weierstrass theorem, what can be concluded about global optimality? What else can be said by using the known properties of convex problems?

c. Repeat the points a. and b. for $Q = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

d. Repeat the points a. and b. for $Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Sol.: a. The first-oder necessary condition for a local minimum is

$$Qx + q = 0,$$

which leads to the unique solution $x = [0\ 1]^T$. This point is thus a candidate for local optimality. Moreover, from the second-order sufficient condition

$$\nabla^2 f_o(x) = Q \succ 0,$$

we conclude that $x = [0\ -1]^T$ is a local minimum. Note that $Q \succ 0$ since its eigenvalues are 0.382 and 2.618.

b. Weierstrass theorem applies, since the function is continuous (and thus lower-semincontinuous), the optimization set is closed and the function is coercive (as it can be easily seen). Therefore, a global minimum exists, and it must be $x = [0\ -1]^T$. The fact that the local minimum is also global and that there is a unique (global) minimum is also supported by the fact that $f_o(x)$ is strictly convex.

c. From the first-order necessary condition, we get that the points on the hyperplane $x_1+x_2 = -1$ can all be local minima. These points also satisfy the second-order necessary condition, since Q is semipositive definite (eigenvalues 0 and 2). However, they do not satisfy the second-order sufficient condition. So, we cannot conclude on the local (or global) optimality of these points. Notice also that Weirstrass is not satisfied as the function is not coercive (try with $x = b[-1, 1]^T$ and $b \to \infty$). Using the properties of convex optimization problems, however, we are able to conclude that all the points on the hyperplane $x_1 + x_2 = 1$ are local and global minima.

d. In this case, the system of first-order necessary condition leads to the unique solution $x = -[1\ 1]^T$. But this point does not satisfy the second-order necessary condition, since Q is not semipositive definite (the eigenvalues are 1 and -1). We conclude that there are no optimal points. In fact, the problem is unbounded below (to see this, take $x_1 \to -\infty$ and $x_2 = 0$).

P.3. (2 points) Consider the two objective functions

$$\begin{aligned} f_1(x_1, x_2) &= x_1 - x_2 + 1 \\ \text{and } f_2(x_1, x_2) &= x_2 - x_1 + 1, \end{aligned}$$

to be maximized with the constraints $0 \le x_1 \le 1$ and $0 \le x_2 \le 1$.

a. Show that the set $\mathcal{A}$ of all achievable pairs $(f_1(x_1, x_2), f_2(x_1, x_2))$ is convex using known properties of convex sets, and plot the region $\mathcal{A}$.

b. Describe how you would obtain the boundary of this set by scalarization.

c. Find the Pareto optimal points. Which one of the Pareto optimal points is a Nash equilibrium?
d. Repeat the points above with the objective functions

$$\begin{aligned} f_1(x_1, x_2) &= x_1 - x_2 + 1 \\ \text{and } f_2(x_1, x_2) &= 2x_1 - 2x_2 + 3. \end{aligned}$$

Sol.:
a. The set $\mathcal{A}$ is the image of the convex set $\{(x_1, x_2)|0 \leq x_1 \leq 1 \text{ and } 0 \leq x_2 \leq 1\}$ under the linear transformations $f_1(x_1, x_2)$ and $f_2(x_1, x_2)$, and is therefore convex.
We need to plot the set $\{(f_1, f_2) \in \mathbb{R}^2 | f_1 = x_1 - x_2 + 1 \text{ and } f_2 = x_2 - x_1 + 1, \text{ with } 0 \leq x_1 \leq 1 \text{ and } 0 \leq x_2 \leq 1\}$. We observe that this set is equal to $\{(f_1, f_2) \in \mathbb{R}^2 | f_1 + f_2 = 2, \text{ with } 0 \leq f_1 \leq 2 \text{ and } 0 \leq f_2 \leq 2\}$, as it can be easly shown by checking that either set includes the other. So the feasible set is the segment $f_1 + f_2 = 2$ in the positive quadrant.
b. We can use scalarization since the set $\mathcal{A}$ is convex. Scalarization entails solving the problem

$$\begin{aligned} &\text{maximize } \lambda_1 f_1(x_1, x_2) + \lambda_2 f_2(x_1, x_2) \\ &\text{s.t. } 0 \leq x_1 \leq 1 \text{ and } 0 \leq x_2 \leq 1 \end{aligned}$$

for all $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$, or equivalently

$$\begin{aligned} &\text{maximize } \lambda_1 (x_1 - x_2) + \lambda_2 (x_2 - x_1) \\ &\text{s.t. } 0 \leq x_1 \leq 1 \text{ and } 0 \leq x_2 \leq 1 \end{aligned} .$$

It is easy to see that for $\lambda_1 > \lambda_2$, the solution leads to $x = (1, 0)$, that is, to $f_1 = 2$, $f_2 = 0$, while for $\lambda_2 > \lambda_1$, the solution leads to $x = (0, 1)$, that is, to $f_1 = 0$, $f_2 = 2$. Instead, with $\lambda_1 = \lambda_2$, any feasible x is optimal, and thus the entire set $\mathcal{A}$ lies on its boundary.
c. All points in $\mathcal{A}$ are Pareto optimal. The only Nash equilibrium is $x = (1, 1)$, which leads to $f_1 = 1$, $f_2 = 1$.
d. The set $\mathcal{A}$ is convex for the same reason as above. Now, note that we have

$$\begin{aligned} f_1(x_1, x_2) &= x_1 - x_2 + 1 \\ \text{and } f_2(x_1, x_2) &= 2f_1(x_1, x_2) + 1. \end{aligned}$$

Therefore, we have $\mathcal{A} = \{(f_1, f_2) \in \mathbb{R}^2 | 0 \leq f_1 \leq 2 \text{ and } f_2 = 2f_1 + 1\}$, which is a segment with a positive slope.
Scalarization always yields the point $(f_1, f_2) = (2, 5)$, that is, $x = (1, 0)$. This point is also the only Pareto optimal point and also the only Nash Equilibrium.