1 Convex Optimization

1.1 Solution

1. Let $C \subseteq \mathbf{R}^n$ be a convex set, with $x_1, \ldots, x_k \in C$, and let $\theta_1, \ldots, \theta_k \in \mathbf{R}$ satisfy $\theta_i \leq 0$, $\theta_1 + \cdots + \theta_k = 1$. Show that $\theta_{1x_1} + \cdots + \theta_{kx_k} \in C$. (The definition of convexity is that this holds for $k = 2$; you must show it for arbitrary k)
Solution This is readily show by induction from the definition of convex set. We illustrate the idea for $k = 3$, leaving the general case to the reader. Suppose that $x_1, x_2, x_3 \in C$, and $\theta_1 + \theta_2 + \theta_3 = 1$ with $\theta_1, \theta_2, \theta_3 \geq 0$. We will show that $y = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 \in C$. At least one of the θ_i is not equal to one; without loss of generality we can assume that $\theta_1 \neq 1$. Then we can write

$$y = \theta_1 x_1 + (1 - \theta_1)(\mu_2 x_2 + \mu_3 x_3) \tag{1}$$
$$\text{where} \quad \mu_2 = \theta_2/(1-\theta_1) \quad \text{and} \quad \mu_2 = \theta_3/(1-\theta_1). \text{Note that} \quad \mu_2, \mu_3 \geq \quad \text{and} \tag{2}$$
$$\mu_1 + \mu_2 = \frac{\theta_2 + \theta_3}{1 - \theta_1} = \frac{1 - \theta_1}{1 - \theta_1} = 1 \tag{3}$$
$$\tag{4}$$

2. Show that a set is convex if and only if its intersection with any line is convex. Show that a set is affine if and only if its intersection with any line is affine.
Solution. We prove the first part. The intersection of two convex sets is convex. Therefore is S is a convex set, the intersection of S with a line is convex.
Conversely, suppose the intersection of S with any line is convex. Take any two distinct points x_1 and $x_2 \in S$. The intersection of S with the line through $x_1 \text{and} x_2$ is convex. Therefore convex combinations of $x_1 \text{and} x_2$ belong to the intersection, hence also to S

3. For each of the following functions determine whether it is convex, concave, quasiconvex, or quasiconcave.
   (a) $f(x_1, x_2) = x_1 x_2 \quad \text{on} \quad \mathbf{R}^2_{++}$.
   Solution. The Hessian of f is
   $$\Delta^2 f(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{5}$$
   $$\tag{6}$$
   Which is neither positive semidefinite nor negative semidefinite. Therefore, f is neither convex and concave. It is quasiconcave, since its superlevel sets are convex. It is not quasiconvex.
   $$\{(x_1, x_2) \in \mathbf{R}^2_{++} \quad | \quad x_1, x_2 \geq \alpha\} \tag{7}$$
   (b) $f(x_1, x_2) = 1/(x_1 x_2) \quad \text{on} \quad \mathbf{R}^2_{++}$.
   Solution. The Hessian of f is
   $$\Delta^2 f(x) = \frac{1}{x_1 x_2} \begin{bmatrix} 2/(x_1^2) & 1/(x_1 x_2) \\ 1/(x_1 x_2) & 2/x_2^2 \end{bmatrix} \succeq 0 \tag{8}$$
   $$\tag{9}$$
   Therefore, f is convex and quasiconvex. It is not quasiconcave or concave.

4. Derive the conjugates of the following functions max function.$f(x) = max_{i=1,\ldots,n} x_1$ on $\mathbf{R}^n$.
Solution. We will show that
$$f^*(y) = \begin{cases} 0, & \text{if } y \succeq 0, 1^T y = 1 \\ \infty, & \text{otherwise} \end{cases} \tag{10}$$
$$\tag{11}$$
We first verify the domain of f^*. First suppose y has a negative component, say $y_k < 0$. If we choose a vector x with $x_k = -t, x_i = 0\text{for } i \neq k$ and let t go to infinity, we see that
$$X^T y - \max_i x_i = -t_{y_k} \to \infty \tag{12}$$
$$\tag{13}$$

so y is not in $\mathbf{dom} f^*$. Next, assume $y \succeq 0$ but $1^T y > 1$. We choose $x = t1$ and let t go to infinity, to show that

$$x^T y - \max_i x_i = t1^T y - 1 \tag{14}$$

is unbounded above. Similarly, when $y \succeq 0$ and $1^T y < 1$, we choose $x = -t1$ and let t go to infinity : The remaining case for y is $y \succeq 0$ and $1^T y = 1$. In this case we have

$$x^T y \leq \max_i x_i \tag{15}$$

for all x, and therefore $x^T y - \max_i x_i \leq 0$ for all$x$with equality for $x = 0$. Therefore $f * (y) = 0$

5. Approximation width. Let $f_0; \ldots, f_n : \mathbf{R} \to \mathbf{R}$ be given continuous functions. We consider the problem of approximating f_0 as a linear combination of $f_0, \ldots, f_n$. For $x \in \mathbf{R}^n$, we say that $f = x_1 + \cdots + x_n f_n$ approximates f_0 with tolerance $\epsilon > 0$ over the interval $[0, T]$if $|f(t) - f_0(t)| \leq \epsilon$ for $0 \leq t \leq T$. Now we choose a fixed tolerance $\epsilon > 0$ and define the approximation width as the largest T such that f approximates f_0 over the interval $[0, T]$:

$$W(x) = sup\{T||x_1 f_1(t) + \cdots + x_n f_n(t) - f_0(t)| \leq \epsilon \text{ for } 0 \leq t \leq T\} \tag{16}$$

Show that W is quasiconcanve
Solution. To show that W is quasiconcanve we show that the sets $\{x|W(x) \geq \alpha\}$ are convec for all α. We have $W(x) \geq \alpha$ if and only if

$$-\epsilon \leq x_1 f_1(t) + \cdots + x_n f_n(t) - f_0(t) \leq \epsilon \tag{17}$$
$$\tag{18}$$

for all $t \in [0, \alpha)$. Therefore the set $\{x|W(x) \leq \alpha\}$ is an intersection of infinitely many halfspaces (two for each t), hence a convex set.

6. Some simple LPs, give an explicit solution of each of the following LPs.

   (a) Minimizing a linear function over an affine set

$$\text{minimize } c^T x \tag{19}$$
$$\text{subject to } Ax = b \tag{20}$$
$$\tag{21}$$

   Solution We distinguish three possibilities

   i. The problem is infeasible ($b \notin \mathcal{R}(A)$). The optimal value is ∞
   ii. The problem is feasible, and c is orthogonal to the nullspace of A. We can decompose c as

$$c = A^T \lambda + \hat{c}, A\hat{c} = 0 \tag{22}$$
$$\tag{23}$$

   ($\hat{c}$ is the component in the nullspace of A; $A^T \lambda$ is orthogonal to the nullspace.)
   If $\hat{c} = 0$, then on the feasible set the objective function reduces to a constant:

$$c^T x = \lambda^T Ax + \hat{c}^T x = \lambda^T b \tag{24}$$
$$\tag{25}$$

   The optimal value is $\lambda^T b$. All feasible solutions are optimal
   iii. The problem is feasible and c is not in the range of $A^T (\hat{c} \neq 0)$. The problem is unbounded ($p^* = -\infty$). To verfy this, note that $x = x_0 - t\hat{c}$ is feasible for all t; as t goes to infinity, the objective value decreases unboundedly.
   In summary,

$$p^* = \begin{cases} +\infty, & b \notin \mathcal{R}(A) \\ \lambda^T b, & c = A^T \lambda \text{for some } \lambda \\ \infty, & \text{otherwise} \end{cases} \tag{26}$$

7. Minimax rational function fitting. Show that the following problem is quasiconvex :

$$minimize \max_{i=1,\dots,k} \left| \frac{p(t_i)}{q(t_i)} - y_i \right| \tag{27}$$

$$Where \quad p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_m t^m, q(t) = 1 + b_1 t + \cdots + b_n t^n, \tag{28}$$

and the domain of the objective function is defined as

$$D = \{(a, b) \in \mathbf{R}^{m+1} \times \mathbf{R}^n | q(t) > 0, \alpha \leq t \leq \beta\} \tag{29}$$

(30)

In this problem we fit a rational function $p(t)/q(t)$ to given data, while constraining the denominator polynomial to be positive on the interval $[\alpha, \beta]$. The optimization variables are the numerator and denominator coefficients a_i, b_i. The interpolation points $t_i \in [\alpha, \beta]$, and desired function values $y_i, i = 1, \dots, k$, are given.
Solution. Lets show the objective is quasiconvex. Its domain convex. Since $q(t_i) > 0$ for $i = 1, \dots, k$ we have

$$\max_{i=1,\dots,k} |p(t_i)/q(t_i) - y_i| \leq \gamma \tag{31}$$

(32)

if and only if

$$|p(t_i)/q(t_i) - y_i| \leq \gamma q(t_i), i = 1, \dots, k, \tag{33}$$

which defines a convex set in the variables a and b, since the lefthand side is convex and the righthand side is linear. We can further express these inequalities as a set of 2k linear inequalities.

$$-\gamma q(t_i) \leq p(t_i) - y_i q(t_i) \leq \gamma q(t_i), i = 1, \dots, k \tag{34}$$

(35)

8. $\ell_1, \ell_2,$ and ℓ_∞ approximation by a constant vector. What is the solution of the norm approximation problem with one scalar variable $x \in \mathbf{R}$

$$\text{minimize } ||x1 - b|| \tag{36}$$

for the $\ell_1, \ell_2$and $\ell_\infty - norms$ **Solution**.

(a) $\ell_2 - norm$: the average $1^T b/m$
(b) $\ell_1 - norm$: the (or a) median of the coefficeints of b.
(c) $\ell_\infty - norm$: the midrange point(max b_i − maxb_i)/2.