Solutions HW 1

Description of convex sets.

Which of the following sets are convex?

1. A *slab*, *i.e.*, a set of the form $\{x \in \mathbf{R}^n \mid \alpha \le a^T x \le \beta\}$.
2. A *rectangle*, *i.e.*, a set of the form $\{x \in \mathbf{R}^n \mid \alpha_i \le x_i \le \beta_i,\ i = 1, \dots, n\}$. A rectangle is sometimes called a *hyperrectangle* when $n > 2$.
3. A *wedge*, *i.e.*, $\{x \in \mathbf{R}^n \mid a_1^T x \le b_1,\ a_2^T x \le b_2\}$.
4. The set of points closer to a given point than a given set, *i.e.*,
$$\{x \mid \|x - x_0\|_2 \le \|x - y\|_2 \text{ for all } y \in S\}$$
where $S \subseteq \mathbf{R}^n$.
5. The set of points closer to one set than another, *i.e.*,
$$\{x \mid \mathbf{dist}(x, S) \le \mathbf{dist}(x, T)\},$$
where $S, T \subseteq \mathbf{R}^n$, and
$$\mathbf{dist}(x, S) = \inf\{\|x - z\|_2 \mid z \in S\}.$$
6. The set $\{x \mid x + S_2 \subseteq S_1\}$, where $S_1, S_2 \subseteq \mathbf{R}^n$ with S_1 convex.
7. The set of points whose distance to a does not exceed a fixed fraction θ of the distance to b, *i.e.*, the set $\{x \mid \|x - a\|_2 \le \theta \|x - b\|_2\}$. You can assume $a \ne b$ and $0 \le \theta \le 1$.

Solution.

1. A slab is an intersection of two halfspaces, hence it is a convex set (and a polyhedron).
2. As in part (a), a rectangle is a convex set and a polyhedron because it is a finite intersection of halfspaces.
3. A wedge is an intersection of two halfspaces, so it is convex set. It is also a polyhedron. It is a cone if $b_1 = 0$ and $b_2 = 0$.
4. This set is convex because it can be expressed as
$$\bigcap_{y \in S} \{x \mid \|x - x_0\|_2 \le \|x - y\|_2\},$$
i.e., an intersection of halfspaces. (For fixed y, the set
$$\{x \mid \|x - x_0\|_2 \le \|x - y\|_2\}$$
is a halfspace; see exercise 2.7).

5. In general this set is not convex, as the following example in $\mathbf{R}$ shows. With $S = \{-1, 1\}$ and $T = \{0\}$, we have
$$\{x \mid \mathbf{dist}(x, S) \leq \mathbf{dist}(x, T)\} = \{x \in \mathbf{R} \mid x \leq -1/2 \text{ or } x \geq 1/2\}$$
which clearly is not convex.

6. This set is convex. $x + S_2 \subseteq S_1$ if $x + y \in S_1$ for all $y \in S_2$. Therefore
$$\{x \mid x + S_2 \subseteq S_1\} = \bigcap_{y \in S_2} \{x \mid x + y \in S_1\} = \bigcap_{y \in S_2} (S_1 - y),$$
the intersection of convex sets $S_1 - y$.

7. The set is convex.
$$\begin{aligned} & \{x \mid \|x - a\|_2 \leq \theta \|x - b\|_2\} \\ = \quad & \{x \mid \|x - a\|_2^2 \leq \theta^2 \|x - b\|_2^2\} \\ = \quad & \{x \mid (1 - \theta^2) x^T x - 2(a - \theta^2 b)^T x + (a^T a - \theta^2 b^T b) \leq 0\} \end{aligned}$$
If $\theta = 1$, this is a halfspace.

Euclidean distance matrices.

Let $x_1, \ldots, x_n \in \mathbf{R}^k$. The matrix $D \in \mathbf{S}^n$ defined by $D_{ij} = \|x_i - x_j\|_2^2$ is called a *Euclidean distance matrix*. It satisfies some obvious properties such as $D_{ij} = D_{ji}$, $D_{ii} = 0$, $D_{ij} \geq 0$, and (from the triangle inequality) $D_{ik}^{1/2} \leq D_{ij}^{1/2} + D_{jk}^{1/2}$. We now pose the question: When is a matrix $D \in \mathbf{S}^n$ a Euclidean distance matrix (for some points in $\mathbf{R}^k$, for some k)? A famous result answers this question: $D \in \mathbf{S}^n$ is a Euclidean distance matrix if and only if $D_{ii} = 0$ and $x^T D x \leq 0$ for all x with $\mathbf{1}^T x = 0$.

Show that the set of Euclidean distance matrices is a convex cone. Find the dual cone.

Solution. The set of Euclidean distance matrices in $\mathbf{S}^n$ is a closed convex cone because it is the intersection of (infinitely many) halfspaces defined by the following homogeneous inequalities:
$$e_i^T D e_i \leq 0, \qquad e_i^T D e_i \geq 0, \qquad x^T D x = \sum_{j,k} x_j x_k D_{jk} \leq 0,$$
for all $i = 1, \ldots, n$, and all x with $\mathbf{1}^T x = 1$.

It follows, denoting $\mathbf{Co}$ the conic hull of a set, that the dual cone is given by
$$K^* = \mathbf{Co}(\{-xx^T \mid \mathbf{1}^T x = 1\} \bigcup \{e_1 e_1^T, -e_1 e_1^T, \ldots, e_n e_n^T, -e_n e_n^T\}).$$
This can be made more explicit as follows. Define $V \in \mathbf{R}^{n \times (n-1)}$ as
$$V_{ij} = \begin{cases} 1 - 1/n & i = j \\ -1/n & i \neq j. \end{cases}$$
The columns of V form a basis for the set of vectors orthogonal to $\mathbf{1}$, *i.e.*, a vector x satisfies $\mathbf{1}^T x = 0$ if and only if $x = Vy$ for some y. The dual cone is
$$K^* = \{VWV^T + \mathbf{diag}(u) \mid W \preceq 0, u \in \mathbf{R}^n\}.$$

Pointwise maximum and supremum.

Show that the following functions $f : \mathbf{R}^n \to \mathbf{R}$ are convex.

1. $f(x) = \max_{i=1,\ldots,k} \|A^{(i)}x - b^{(i)}\|$, where $A^{(i)} \in \mathbf{R}^{m\times n}$, $b^{(i)} \in \mathbf{R}^m$ and $\|\cdot\|$ is a norm on $\mathbf{R}^m$.

   Solution. f is the pointwise maximum of k functions $\|A^{(i)}x - b^{(i)}\|$. Each of those functions is convex because it is the composition of an affine transformation and a norm.

2. $f(x) = \sum_{i=1}^r |x|_{[i]}$ on $\mathbf{R}^n$, where $|x|$ denotes the vector with $|x|_i = |x_i|$ (*i.e.*, $|x|$ is the absolute value of x, componentwise), and $|x|_{[i]}$ is the ith largest component of $|x|$. In other words, $|x|_{[1]}, |x|_{[2]}, \ldots, |x|_{[n]}$ are the absolute values of the components of x, sorted in nonincreasing order.

   Solution. Write f as

   $$f(x) = \sum_{i=1}^r |x|_{[i]} = \max_{1\le i_1<i_2<\cdots<i_r\le n} |x_{i_1}| + \cdots + |x_{i_r}|$$

   which is the pointwise maximum of $n!/(r!(n-r)!)$ convex functions.

Products and ratios of convex functions.

In general the product or ratio of two convex functions is not convex. However, there are some results that apply to functions on $\mathbf{R}$. Prove the following.

1. If f and g are convex, both nondecreasing (or nonincreasing), and positive functions on an interval, then fg is convex.

2. If f, g are concave, positive, with one nondecreasing and the other nonincreasing, then fg is concave.

3. If f is convex, nondecreasing, and positive, and g is concave, nonincreasing, and positive, then f/g is convex.

Solution.

1. We prove the result by verifying Jensen's inequality. f and g are positive and convex, hence for $0 \le \theta \le 1$,

   $$\begin{aligned} f(\theta x + (1-\theta)y)\, g(\theta x + (1-\theta)y) &\le (\theta f(x) + (1-\theta)f(y))\,(\theta g(x) + (1-\theta)g(y)) \\ &= \theta f(x)g(x) + (1-\theta)f(y)g(y) \\ &\quad + \theta(1-\theta)(f(y)-f(x))(g(x)-g(y)). \end{aligned}$$

   The third term is less than or equal to zero if f and g are both increasing or both decreasing. Therefore

   $$f(\theta x + (1-\theta)y)\, g(\theta x + (1-\theta)y) \le \theta f(x)g(x) + (1-\theta)f(y)g(y).$$

2. Reverse the inequalities in the solution of part (a).

3. It suffices to note that $1/g$ is convex, positive and increasing, so the result follows from part (a).

Conjugates.

Derive the conjugates of the following functions.

1. *Max function.* $f(x) = \max_{i=1,\dots,n} x_i$ on $\mathbf{R}^n$.

   Solution. We will show that

$$f^*(y) = \begin{cases} 0 & \text{if } y \succeq 0, \ \ \mathbf{1}^T y = 1 \\ \infty & \text{otherwise.} \end{cases}$$

   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 $x_k = -t$, $x_i = 0$ for $i \neq k$, and let t go to infinity, we see that

$$x^T y - \max_i x_i = -ty_k \to \infty,$$

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

$$x^T y - \max_i x_i = t\mathbf{1}^T y - t$$

   is unbounded above. Similarly, when $y \succeq 0$ and $\mathbf{1}^T y < 1$, we choose $x = -t\mathbf{1}$ and let t go to infinity.

   The remaining case for y is $y \succeq 0$ and $\mathbf{1}^T y = 1$. In this case we have

$$x^T y \leq \max_i x_i$$

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

2. *Sum of largest elements.* $f(x) = \sum_{i=1}^r x_{[i]}$ on $\mathbf{R}^n$.

   Solution. The conjugate is

$$f^*(y) = \begin{cases} 0 & 0 \preceq y \preceq \mathbf{1}, \quad \mathbf{1}^T y = r \\ \infty & \text{otherwise,} \end{cases}$$

   We first verify the domain of f^*. Suppose y has a negative component, say $y_k < 0$. If we choose a vector x with $x_k = -t$, $x_i = 0$ for $i \neq k$, and let t go to infinity, we see that

$$x^T y - f(x) = -ty_k \to \infty,$$

   so y is not in $\mathbf{dom}\, f^*$.

   Next, suppose y has a component greater than 1, say $y_k > 1$. If we choose a vector x with $x_k = t$, $x_i = 0$ for $i \neq k$, and let t go to infinity, we see that

$$x^T y - f(x) = ty_k - t \to \infty,$$

   so y is not in $\mathbf{dom}\, f^*$.

   Finally, assume that $\mathbf{1}^T x \neq r$. We choose $x = t\mathbf{1}$ and find that

$$x^T y - f(x) = t\mathbf{1}^T y - tr$$

is unbounded above, as $t \to \infty$ or $t \to -\infty$.

If y satisfies all the conditions, for $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$ with $x_1 \geq \ldots \geq x_n$,

$$\begin{aligned} x^T y - f(x) &= \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{r} x_i \\ &= \sum_{i=r+1}^{n} x_i y_i + \sum_{i=1}^{r} (y_i - 1) x_i \\ &\leq x_r \left(\sum_{i=r+1}^{n} y_i + \sum_{i=1}^{r} (y_i - 1) \right) \\ &= 0 \end{aligned}$$

with equality for $x = 0$. Therefore $f^*(y) = 0$.

3. *Piecewise-linear function on* $\mathbf{R}$. $f(x) = \max_{i=1,\ldots,m}(a_i x + b_i)$ on $\mathbf{R}$. You can assume that the a_i are sorted in increasing order, *i.e.*, $a_1 \leq \cdots \leq a_m$, and that none of the functions $a_i x + b_i$ is redundant, *i.e.*, for each k there is at least one x with $f(x) = a_k x + b_k$.

   Solution. Under the assumption, the graph of f is a piecewise-linear, with breakpoints $(b_i - b_{i+1})/(a_{i+1} - a_i)$, $i = 1, \ldots, m-1$. We can write f^* as

$$f^*(y) = \sup_x \left(xy - \max_{i=1,\ldots,m} (a_i x + b_i) \right)$$

   We see that $\mathbf{dom}\, f^* = [a_1, a_m]$, since for y outside that range, the expression inside the supremum is unbounded above. For $a_i \leq y \leq a_{i+1}$, the supremum in the definition of f^* is reached at the breakpoint between the segments i and $i+1$, *i.e.*, at the point $(b_{i+1} - b_i)/(a_i - a_{i+1})$, so we obtain

$$f^*(y) = (y - a_i) \frac{b_{i+1} - b_i}{a_i - a_{i+1}} - b_i$$

   where i is defined by $a_i \leq y \leq a_{i+1}$. Hence the graph of f^* is also a piecewise-linear curve.

4. *Power function.* $f(x) = x^p$ on $\mathbf{R}_{++}$, where $p > 1$. Repeat for $p < 0$.

   Solution. We'll use standard notation: we define q by the equation $1/p + 1/q = 1$, *i.e.*, $q = p/(p-1)$.

   We start with the case $p > 1$. For $y \leq 0$ the function $yx - x^p$ is non-negative decreasing and it achieves its maximum for $x \to 0$, so $f^*(y) = 0$. For $y > 0$ the function achieves its maximum at $x = (y/p)^{1/(p-1)}$, where it has value

$$y(y/p)^{1/(p-1)} - (y/p)^{p/(p-1)} = (p-1)(y/p)^q.$$

   Therefore we have

$$f^*(y) = \begin{cases} 0 & y \leq 0 \\ (p-1)(y/p)^q & y > 0. \end{cases}$$

   For $p < 0$ similar arguments show that $\mathbf{dom}\, f^* = -\mathbf{R}_{++}$ and $f^*(y) = (p-1)(y/p)^q$.

5. *Geometric mean.* $f(x) = -(\prod x_i)^{1/n}$ on $\mathbf{R}^n_{++}$.

Solution. The conjugate function is

$$f^*(y) = \begin{cases} 0 & \text{if } y \prec 0,\ \ (\prod_i(-y_i))^{1/n} \geq 1/n \\ \infty & \text{otherwise.} \end{cases}$$

We first verify the domain of f^*. Assume y has a non-negative component, say $y_k \geq 0$. Then we can choose $x_k = t$ and $x_i = 1$, $i \neq k$, to show that

$$x^T y - f(x) = ty_k + \sum_{i \neq k} y_i + t^{1/n}$$

is unbounded above as a function of $t > 0$. Hence the condition $y \prec 0$ is indeed required.

Next assume that $y \prec 0$, but $(\prod_i(-y_i))^{1/n} < 1/n$. We choose $x_i = -t/y_i$, and obtain

$$x^T y - f(x) = -tn + t\left(\prod_i (-\frac{1}{y_i})\right)^{1/n} \to \infty$$

as $t \to \infty$. This demonstrates that the second condition for the domain of f^* is also needed.

Now assume that $y \prec 0$ and $(\prod_i(-y_i))^{1/n} \geq 1/n$, and $x \succ 0$. The arithmetic-geometric mean inequality states that

$$\frac{x^T|y|}{n} \geq \left(\prod_i (-y_i x_i)\right)^{1/n} \geq \frac{1}{n}\left(\prod_i x_i\right)^{1/n},$$

i.e., $-x^T y \geq -f(x)$ with equality for $x \to 0$. Hence, $f^*(y) = 0$.

6. *Negative generalized logarithm for second-order cone.* $f(x,t) = -\log(t^2 - x^T x)$ on $\{(x,t) \in \mathbf{R}^n \times \mathbf{R} \mid \|x\|_2 < t\}$.

Solution.

$$f^*(y,u) = -2 + \log 4 - \log(u^2 - y^T y), \qquad \mathbf{dom}\, f^* = \{(y,u) \mid \|y\|_2 < -u\}.$$

We first verify the domain. Suppose $\|y\|_2 \geq -u$ (so $u \leq 0$). Choose $x = sy$, $t = \|x\|_2 + 1 \geq s\|y\|_2 \geq -su$, with $s \geq 0$. Then

$$\begin{aligned} y^T x + tu &\geq sy^T y - su^2 = s(y^T y - u^2) \geq 0 \\ \log(t^2 - x^T x) &= \log(2s\|y\|_2 + 1) \end{aligned}$$

Therefore

$$y^T x + tu + \log(t^2 - x^T x)$$

is unbounded above.

Next, assume that $\|y\|_2 < -u$. Setting the derivative of

$$y^T x + ut + \log(t^2 - x^T x)$$

with respect to x and t equal to zero, and solving for t and x we see that the maximizer is

$$x = \frac{2y}{u^2 - y^T y}, \qquad t = -\frac{2u}{u^2 - y^T y}.$$

This gives

$$\begin{aligned} f^*(y, u) &= ut + y^T x + \log(t^2 - x^T x) \\ &= -2 + \log 4 - \log(y^T y - u^2). \end{aligned}$$