Convex Optimization: Fall 2021, Homework #1 (Due: 10/08)

1. (10 points) Positive definite matrix A
   a. Show that the eigenvalues of A are positive
   b. Show that the determinant of A equals the product of its eigenvalues, i.e., $\det(A) = \prod_{i=1}^{n} \lambda_i$.
   c. Show that the positive definite matrix A is invertible.

(sol)

Since A is positive semidefinite, $x^T A x > 0$ for any x. Suppose that λ is the eigenvalue and x is its eigenvector. The, $x^T A x = \lambda \cdot x^T x = \lambda \|x\|_2^2 > 0$, hence, $\lambda > 0$. $\det(A - \lambda I) = \prod_{i=1}^{n}(\lambda_i - \lambda)$. By setting $\lambda = 0$, $\det(A) = \prod_{i=1}^{n} \lambda_i$. Since $\det(A) \neq 0$ for a positive definite matrix A. Hence, A is invertible.

2. (10 points) Show that the matrix $X = \begin{bmatrix} x & y \\ y & z \end{bmatrix}$ (where $x \geq 0,\ z \geq 0,\ xz \geq y^2$) is positive semidefinite. (Hint: Use the fact that the summation of eigenvectors is the same as the summation of diagonal terms, i.e., $\mathrm{tr(X)} = \sum_i X_{ii} = \sum_i \lambda_i$.)

(sol)

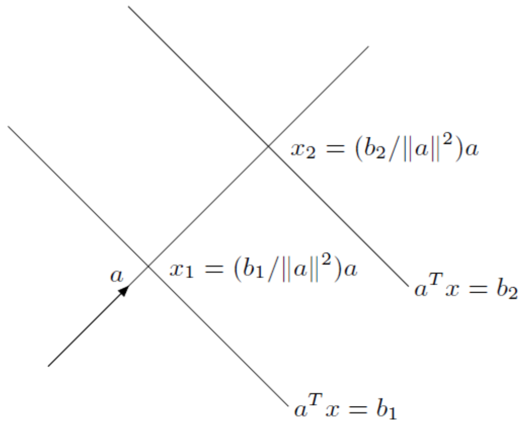
$$\det(X - \lambda I) = \begin{bmatrix} x - \lambda & y \\ y & z - \lambda \end{bmatrix} = (x - \lambda)(z - \lambda) - y^2 = 0$$

Then, $\lambda^2 - (x + z)\lambda + xz - y^2 = 0$. Then, $\lambda_1 + \lambda_2 = x + z$. Also, $\lambda_1 \lambda_2 = \det X = xz - y^2$.

Since $x \geq 0,\ z \geq 0,\ xz \geq y^2$, we observe that $\lambda_1 + \lambda_2 \geq 0$ and $\lambda_1 \lambda_2 \geq 0$, which leads to $\lambda_1 \geq 0, \lambda_2 \geq 0$. Since all the eigenvlaues are nonnegative, X is positive semidefinite.

3. (10 points) Textbook (2.5)

Solution. The distance between the two hyperplanes is $|b_1 - b_2|/\|a\|_2$. To see this, consider the construction in the figure below.



The distance between the two hyperplanes is also the distance between the two points x_1 and x_2 where the hyperplane intersects the line through the origin and parallel to the normal vector a. These points are given by

$$x_1 = (b_1/\|a\|_2^2)a, \qquad x_2 = (b_2/\|a\|_2^2)a,$$

and the distance is

$$\|x_1 - x_2\|_2 = |b_1 - b_2|/\|a\|_2.$$

4. (10 points) Textbook (2.12: a, b, c, d)

(sol)

(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 set. It is also a polyhedron. It is a cone if $b_1 = 0$ and $b_2 = 0$.

(d) 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 2.9).

5. (10 points) Textbook (2.16)

Solution. We consider two points $(\bar{x}, \bar{y}_1 + \bar{y}_2)$, $(\tilde{x}, \tilde{y}_1 + \tilde{y}_2) \in S$, *i.e.*, with

$$(\bar{x}, \bar{y}_1) \in S_1, \qquad (\bar{x}, \bar{y}_2) \in S_2, \qquad (\tilde{x}, \tilde{y}_1) \in S_1, \qquad (\tilde{x}, \tilde{y}_2) \in S_2.$$

For $0 \le \theta \le 1$,

$$\theta(\bar{x}, \bar{y}_1 + \bar{y}_2) + (1-\theta)(\tilde{x}, \tilde{y}_1 + \tilde{y}_2) = (\theta\bar{x} + (1-\theta)\tilde{x}, (\theta\bar{y}_1 + (1-\theta)\tilde{y}_1) + (\theta\bar{y}_2 + (1-\theta)\tilde{y}_2))$$

is in S because, by convexity of S_1 and S_2,

$$(\theta\bar{x} + (1-\theta)\tilde{x}, \theta\bar{y}_1 + (1-\theta)\tilde{y}_1) \in S_1, \qquad (\theta\bar{x} + (1-\theta)\tilde{x}, \theta\bar{y}_2 + (1-\theta)\tilde{y}_2) \in S_2.$$

6. (10 points) Let X_{opt} be the set of all optimal solutions of the following convex optimization:

$$\begin{aligned} X_{\text{opt}} = \ & \text{argmin} && f_0(x) \\ & \text{subject to} && f_i(x) \le 0, \quad i = 1, \dots, m \\ & && Ax = b \end{aligned}$$

where $f_0(x)$ and $f_i(x)$ are convex functions. Show that X_{opt} is a convex set.

(sol)

Suppose that x, y are optimal solutions. Then, for $0 \le \theta \le 1$,

$f_i(\theta x + (1-\theta)y) \le \theta f_i(x) + (1-\theta) f_i(y) \le 0$ and $A(\theta x + (1-\theta)y) = \theta Ax + (1-\theta)Ay = b$.

$f_0(\theta x + (1-\theta)y) \le \theta f_0(x) + (1-\theta) f_0(y) = f^*$ where the optimal $f^* = f(x) = f(y)$. Hence, $f_0(\theta x + (1-\theta)y) = f^*$, which shows that $\theta x + (1-\theta)y$ is also optimal. Note that $\theta x + (1-\theta)y \in$ $\in X_{\text{opt}}$ for any $x, y \in X_{\text{opt}}$ and $0 \le \theta \le 1$.

7. (10 points) Show that:

   a. Exponential e^{ax} is convex on $\mathbf{R}$, for any $a \in \mathbf{R}$.
   b. Powers x^a is convex on $\mathbf{R}_{++}$ when $a \geq 1$ or $a \leq 0$, and concave for $0 \leq a \leq 1$.
   c. Entropy function (in information theory) $f(x) = -\sum_{i=1}^{n} x_i \log x_i$ with $\mathrm{dom}\, f(x) = \{x \in R_{+t}^n | \sum_{i=1}^{n} x_i = 1\}$ is strictly concave.
   d. Exponential e^{x^2} is convex on $\mathbf{R}$.

(sol)

   a. $f(x) = e^{ax}$, then $f''(x) = a^2 e^{ax} \geq 0$ (convex).
   b. $f(x) = x^a$, then $f''(x) = a(a-1)x^{a-2}$. Since $x \in \mathbf{R}_{++}$, $f''(x) \geq 0$ (convex) for $a \geq 1$ or $a \leq 0$ and $f''(x) \leq 0$ (concave) for $0 \leq a \leq 1$.
   c. $\frac{\partial^2 f(x)}{\partial x_i^2} = -\frac{1}{x_i} < 0$ and $\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = 0$. Hence, $\nabla^2 f(x) \prec 0$, i.e., strictly concave.
   d. $f(x) = e^{x^2} = h(g(x))$ where $g(x) = x^2$ and $h(x) = e^x$. Since h is convex and nondecreasing, and g is convex, $f(x)$ is convex.

8. (10 points) The Kullback-Leibler divergence is given by $D_{KL}(p, q) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i}$ where p_i and q_i represents two distributions. Show that $D_{KL}(p, q) \geq 0$. (Hint: Use Jensen's inequality)

(sol)

Suppose that $S = \{x | p(x) > 0\}$ be the support set of $p(x)$. Then,

$$\begin{aligned} -D_{KL}(p, q) &= -\sum_{x \in S} p \log \frac{p}{q} = \sum_{x \in S} p \log \frac{q}{p} \\ &\leq \log \sum_{x \in S} p \cdot \frac{q}{p} = \log \sum_{x \in S} q \leq 0 \end{aligned}$$

Hence, $D_{KL}(p, q) \geq 0$.

9. (10 points) Derive the conjugates of the following functions:

   a. $f(x) = |x|$
   b. $f(x) = \exp(x)$

(sol)

   a. $f^*(y) = \sup_{x \in \mathrm{dom}\, f} yx - |x|$. $yx - |x| = (y-1)x$ for $x \geq 0$, which is unbounded if $y > 1$. Then, $\sup_{x \geq 0}(y-1)x = 0$ for $y \leq 1$. Similarly, $yx - |x| = (y+1)x$ for $x \leq 0$. $(y+1)x$ is unbounded if $y < -1$ for $x \leq 0$, i.e., $\sup_{x \leq 0}(y+1)x = 0$ for $y \geq -1$. Hence, $f^*(y) = 0$ for $|y| \leq 1$ and $f^*(y) = \infty$ otherwise.
   b. $f^*(y) = \sup_{x \in \mathrm{dom}\, f} yx - e^x$. $yx - e^x$ is unbounded if $y < 0$. For $y = 0$, $f^*(y) = \sup_{x \in \mathrm{dom}\, f} -e^x = 0$. For $y > 0$, $yx - e^x$ reaches its maximum when $y - e^x = 0$ (i.e., $x = \log y$). Hence, $f^*(y) = y \log y - y$.

10. (10 points) Textbook (3.8)

Solution. We first assume $n = 1$. Suppose $f : \mathbf{R} \to \mathbf{R}$ is convex. Let $x, y \in \mathbf{dom}\, f$ with $y > x$. By the first-order condition,

$$f'(x)(y-x) \leq f(y) - f(x) \leq f'(y)(y-x).$$

Subtracting the righthand side from the lefthand side and dividing by $(y-x)^2$ gives

$$\frac{f'(y) - f'(x)}{y - x} \geq 0.$$

Taking the limit for $y \to x$ yields $f''(x) \geq 0$.
Conversely, suppose $f''(z) \geq 0$ for all $z \in \mathbf{dom}\, f$. Consider two arbitrary points $x, y \in \mathbf{dom}\, f$ with $x < y$. We have

$$\begin{aligned} 0 &\leq \int_x^y f''(z)(y-z)\, dz \\ &= \left(f'(z)(y-z)\right)\Big|_{z=x}^{z=y} + \int_x^y f'(z)\, dz \\ &= -f'(x)(y-x) + f(y) - f(x), \end{aligned}$$

i.e., $f(y) \geq f(x) + f'(x)(y-x)$. This shows that f is convex.
To generalize to $n > 1$, we note that a function is convex if and only if it is convex on all lines, *i.e.*, the function $g(t) = f(x_0 + tv)$ is convex in t for all $x_0 \in \mathbf{dom}\, f$ and all v. Therefore f is convex if and only if

$$g''(t) = v^T \nabla^2 f(x_0 + tv) v \geq 0$$

for all $x_0 \in \mathbf{dom}\, f$, $v \in \mathbf{R}^n$, and t satisfying $x_0 + tv \in \mathbf{dom}\, f$. In other words it is necessary and sufficient that $\nabla^2 f(x) \succeq 0$ for all $x \in \mathbf{dom}\, f$.

11. (10 points) Textbook (3.19)

(a)

Solution. We can express f as

$$\begin{aligned} f(x) &= \alpha_r(x_{[1]} + x_{[2]} + \cdots + x_{[r]}) + (\alpha_{r-1} - \alpha_r)(x_{[1]} + x_{[2]} + \cdots + x_{[r-1]}) \\ &\quad +(\alpha_{r-2} - \alpha_{r-1})(x_{[1]} + x_{[2]} + \cdots + x_{[r-2]}) + \cdots + (\alpha_1 - \alpha_2)x_{[1]}, \end{aligned}$$

which is a nonnegative sum of the convex functions

$$x_{[1]}, \qquad x_{[1]} + x_{[2]}, \qquad x_{[1]} + x_{[2]} + x_{[3]}, \qquad \ldots, \qquad x_{[1]} + x_{[2]} + \cdots + x_{[r]}.$$

Solution. The function

$$g(x, \omega) = -\log(x_1 + x_2 \cos\omega + x_3 \cos 2\omega + \cdots + +x_n \cos(n-1)\omega)$$

is convex in x for fixed ω. Therefore

$$f(x) = \int_0^{2\pi} g(x, \omega) d\omega$$

is convex in x.

(a) $f(x) = \max_{i=1,\ldots,k} \|A^{(i)}x - b^{(i)}\|$, 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.

(b) $f(x) = \sum_{i=1}^r |x|_{[i]}$ 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.