Convex set

2.1

According to the definition of convexity, this conclusion holds for $k=2$.

Suppose this conclusion holds for $k(k>=2)$, we have:

With $x_1, x_2, ..., x_k \in C$ and $\theta_1, \theta_2, ..., \theta_k \in \mathbb{R}^n$ satisfy $\theta_i \geq 0$ and $\theta_1+\theta_2+\cdots+\theta_k=1$, $\theta_1 x_1+\cdots+\theta_k x_k \in C$.

Then, for any $x_1, x_2, ..., x_k, x_{k+1} \in C$ and $\theta_1, \theta_2, .., \theta_k, \theta_{k+1}$ satisfy $\theta_i \geq 0$ and $\theta_1+\theta_2+\cdots+\theta_k+\theta_{k+1}=1$, we have $u=\frac{\theta_1 x_1+\cdots+\theta_k x_k}{\theta_1+\cdots+\theta_k} \in C, v=x_{k+1} \in C$. Let $\lambda_1=\theta_1+\cdots+\theta_k, \lambda_2=\theta_{k+1}$, then $\lambda_1+\lambda_2=1$, so $\lambda_1 u+\lambda_2 v \in C$, i.e. $\theta_1 x_1+\cdots+\theta_k x_k+\theta_{k+1} x_{k+1} \in C$. That is, the conclusion holds for $k+1$.

In summary, the conclusion holds for all $k \geq 2$.

2.3

Suppose two points $x_1, x_2 \in C$.

Firstly, let's prove that $\forall p \in [0, 2^k]$, $\frac{p}{2^k} x_1+(1-\frac{p}{2^k}) x_2 \in C$ holds for all $k \geq 1$:

Since C is midpoint convex, $\frac{x_1+x_2}{2} \in C$, which means the conclusion holds for $k=1$.

Suppose the conclusion holds for $k(k \geq 1)$, we have: $\forall p \in [0, 2^k]$, $\frac{p}{2^k} x_1+(1-\frac{p}{2^k}) x_2 \in C$

Then, for $p=0,2,4,6, \cdots, 2^k$, $\frac{p}{2^{k+1}} x_1+(1-\frac{p}{2^{k+1}}) x_2 \in C$.

For $p=1,3,5, \cdots, 2^k-1$, $p=\frac{p-1}{2^{k+1}}+\frac{p+1}{2^{k+1}}$, so $\frac{p}{2^{k+1}} x_1+(1-\frac{p}{2^{k+1}}) x_2=\frac{1}{2}\{[\frac{p-1}{2^{k+1}} x_1+(1-\frac{p-1}{2^{k+1}}) x_2]+[\frac{p+1}{2^{k+1}} x_1+(1-\frac{p+1}{2^{k+1}}) x_2]\} \in C$

That is, the conclusion holds for $k+1$.

In summary, the conclusion holds for all $k \geq 1$.

When k becomes infinity, $\{\frac{p}{2^k}|p \in [0, 2^k]\} = \{x|0 \leq x \leq 1\}$.

Therefore, C is convex.

2.5

The distance between these two hyperplanes is the same as the distance between $x_1 = \frac{b_1}{\|a\|_2^2}a$ and $x_2 = \frac{b_2}{\|a\|_2^2}a$, i.e. $d = \|x_1 - x_2\|_2 = \frac{|b_1 - b_2|}{\|a\|_2}$.

2.8

(a) S is a Hydrohedra. It is an intersection of some hyperplane.

(b) S is a polyhedron. It is the solution of some linear inequalities and some linear equalities.

(c) S is not a polyhedron. It is the intersection of a ball $\{x|\|x\|_2 \leq 1\}$ and $\mathbb{R}_+^n$.

(d) S is a polyhedron. It is the intersection of $\{x \mid |x_k| \leq 1, k = 1, ..., n\}$ and $\mathbb{R}_+^n$.

2.10

(a) Let $f(x) = x^T Ax + b^T x + c$, then $\nabla^2 f(x) \succeq 0$ if $A \succeq 0$ i.e. f is convex. That means C is convex.

(b) Best answer: I don't know.

2.12

(a) A *slab* is convex. Proof:

Let $S = \{x \in \mathbb{R}^n | \alpha \le a^T x \le \beta\}$. For any $x_1, x_2 \in S$ and $\theta \in [0,1]$, $\theta\alpha \le a^T \theta x_1 \le \theta\beta$ and $(1-\theta)\alpha \le a^T(1-\theta)x_2 \le (1-\theta)\beta$ holds. So $\alpha \le a^T[\theta x_1 + (1-\theta)x_2] \le \beta$, which means $\theta x_1 + (1-\theta)x_2 \in S$.

Therefore, S is convex.

(b) A *Rectangle* is convex. Proof:

Let $S = \{x \in \mathbb{R}^n | \alpha_i \le x_i \le \beta_i, i = 1, ..., n\}$. For any $x_1, x_2 \in S$ and $\theta \in [0,1]$, $\theta\alpha_i \le \theta x_{1i} \le \theta\beta_i$ and $(1-\theta)\alpha_i \le (1-\theta)x_{2i} \le (1-\theta)\beta_i$ holds. So $\alpha_i \le [\theta x_1 + (1-\theta)x_2]_i \le \beta_i$, which means $\theta x_1 + (1-\theta)x_2 \in S$

Therefore, S is convex.

(c) A *wedge* is convex. It can be proved in the same way as above.

(d) This set is convex. Because $\{x | \|x - x_0\|_2 \le \|x - y\|_2 \text{ for all } y \in S\} = \bigcap\limits_{y \in S} \{x | \|x - x_0\|_2 \le \|x - y\|_2\}$, which is an intersection of norm balls, and norm balls are convex sets.

(e) This set is not convex. Let $n = 1, S = \{2, -2\}, T = \{0\}$, then $\{x | \mathbf{dist}(x, S) \le \mathbf{dist}(x, T)\} = \{x | x \le -1 \wedge x \ge 1\}$, which is clearly not convex.

(f) This set is convex. If $S_2 \subseteq S_1$, then $\{x | x + S_2 \subseteq S_1\} = S_1$, which is convex. Otherwise, $\{x | x + S_2 \subseteq S_1\} = \emptyset$, which is also convex.

(g) This set is convex. $S = \{x | \|x - a\|_2 \le \theta \|x - b\|_2\} = \{x | (1-\theta^2)x^T x - 2(a - \theta^2 b)^T x + (a^T a - \theta^2 b^T b) \le 0\}$. If $\theta = 1$, $S = \{x | -2(a-b)^T x + (a^T a - b^T b) \le 0\}$, which represents a half space. Otherwise, $S = \{x | x^T x - \frac{2(a-\theta^2 b)^T x}{1-\theta^2} + \frac{a^T a - \theta^2 b^T b}{1-\theta^2} \le 0\}$, which represents a ball.

2.16

For any $(x, y_1 + y_2), (u, v_1 + v_2) \in S$ and $\theta \in [0,1]$, there exist $(x, y_1), (u, v_1) \in S_1$, $(x, y_2), (u, v_2) \in S_2$.

Since S_1, S_2 are convex sets,

- $\theta(x, y_1) + (1-\theta)(u, v_1) = (\theta x + (1-\theta)u, \theta y_1 + (1-\theta)v_1) \in S_1$;

- $\theta(x, y_2) + (1-\theta)(u, v_2) = (\theta x + (1-\theta)u, \theta y_2 + (1-\theta)v_2) \in S_2$.

So $(\theta x + (1-\theta)u, \theta(y_1+y_2) + (1-\theta)(v_1+v_2)) = \theta(x, y_1+y_2) + (1-\theta)(u, v_1+v_2) \in S$.

Therefore, S is convex.

2.19

(a) $f^{-1}(C) = \{x \in \mathbf{dom}\, f | g^T f(x) \le h\} = \{x | g^T(Ax+b)/(c^Tx+d) \le h, c^Tx+d > 0\} = \{x | (A^Tg - hc)^T x \le hd - g^Tb, c^Tx + d > 0\}$, which is an intersection of a half space and an open half space.

(b) $f^{-1}(C) = \{x \in \mathbf{dom}\, f | Gf(x) \preceq h\} = \{x | G(Ax+b)/(c^Tx+d) \preceq h, c^Tx+d > 0\} = \{(GA - hc^T)x - hd + Gb \le 0, c^Tx + d > 0\}$, which is an intersection of a polyhedron and an open half space.

(c) $f^{-1}(C) = \{x \in \mathbf{dom}\, f | f(x)^T P^{-1} f(x) \le 1\} = \{x | \frac{(Ax+b)^T}{c^Tx+d} P^{-1} \frac{Ax+b}{c^Tx+d} \le 1, c^Tx+d > 0\}$, which is an intersection of an ellipsoid and an open half space.

(d) $f^{-1}(C) = \{x \in \mathbf{dom}\, f | f_1(x)A_1 + f_2(x)A_2 + \cdots f_n(x)A_n \preceq B\} = \{x | (a_1^Tx + b_1)A_1 + (a_2^Tx + b_2)A_2 + \cdots + (a_n^Tx + b_n)A_n \preceq (c^Tx+d)B, c^Tx+d > 0\}$, which is an intersection of a solution set of a linear matrix inequality and an open half space.

Convex function

3.3

g is concave.

Because f is convex, for any $f(x_1), f(x_2)$, $\theta f(x_1) + (1-\theta)f(x_2) \ge f(\theta x_1 + (1-\theta)x_2)$.

Because f is increasing, g is increasing. So $g(\theta f(x_1) + (1-\theta)f(x_2)) \ge g(f(\theta x_1 + (1-\theta)x_2)) = \theta x_1 + (1-\theta)x_2 = \theta g(f(x_1)) + (1-\theta)g(f(x_2))$.

Denote $f(x_1)$ as y_1, $f(x_2)$ as y_2, we have $\theta g(y_1) + (1-\theta)g(y_2) \le g(\theta y_1 + (1-\theta)y_2)$ for any y_1, y_2. That is, g is concave.

3.5

$F'(x) = -\frac{1}{x^2}\int_0^x f(t)dt + \frac{f(x)}{x}$

$F''(x) = \frac{2}{x^3}\int_0^x f(t)dt - \frac{2f(x)}{x^2} + \frac{f'(x)}{x} = \frac{2}{x^3}(\int_0^x f(t)dt - xf(x) + \frac{x^2 f'(x)}{2}) = \frac{2}{x^3}(\int_0^x f(t)dt + \int_0^x (-f(t) - tf'(t))dt + \int_0^x (tf'(t) + \frac{t^2}{2}f''(t))dt) = \frac{1}{x^3}\int_0^x t^2 f''(t)dt$

Because f is convex, $f''(t) \geq 0$ for $t \in \mathbf{dom}\ f$, so $F''(x) = \frac{1}{x^3}\int_0^x t^2 f''(t)dt \geq 0$ i.e. $F(x)$ is convex for $x \in \mathbb{R}_{++}$.

3.6

$\mathbf{epi}\ f = \{(x,t) | x \in \mathbf{dom}\ f, f(x) \leq t\}$

(a) The epigraph of f is a halfspace when $f(x)$ satisfy the form $f(x) = a^T x + b$:
$f(x) \leq t \Leftrightarrow a^T x + b - t \leq 0 \Leftrightarrow (a^T, \frac{b}{t} - 1)(x,t)^T \leq 0$

(b) The epigraph of f is a convex cone when the following statement holds: $\forall (x_1, t_1), (x_2, t_2) \in \mathbf{epi}\ f, \forall \theta_1, \theta_2 \geq 0, \theta_1(x_1, t_1) + \theta_2(x_2, t_2) \in \mathbf{epi}\ f$

That is, for all $f(x_1) \leq t_1$ and $f(x_2) \leq t_2$, $f(\theta_1 x_1 + \theta_2 x_2) \leq \theta_1 t_1 + \theta_2 t_2$ holds.

Let $\theta_2 = 0$, we have $f(\theta_1 x_1) \leq \theta_1 f(x_1)$.

Let $\theta_1' = \frac{1}{\theta_1}, x_1 = \theta_1' x_1'$, we have $f(x_1') = f(\theta_1 x_1) \leq \theta_1 f(x_1) = \frac{1}{\theta_1'} f(\theta_1' x_1')$ i.e. $f(\theta_1' x_1') \geq \theta_1' f(x_1')$.

Therefore, $f(\theta x) = \theta f(x)$.

(c) The epigraph of f is a polyhedron $\Leftrightarrow$ $\mathbf{epi}\ f$ is in the form of $\{(x,t) | (A, -c)(x,t)^T \preceq b\} = \{(x,t) | \frac{a_1 x - b_1}{c_1} \leq t, \frac{a_2 x - b_2}{c_2} \leq t, \cdots, \frac{a_i x - b_i}{c_i} \leq t\} \Leftrightarrow f_i(x) = \frac{a_i x - b_i}{c_i}$

3.8

Firstly, prove this conclusion for $n = 1$:

Suppose $f : \mathbb{R} \to \mathbb{R}$ is convex. Let $x, y \in \mathbf{dom}\ f$ and $x < y$, according to the first-order condition,

$f(y) \geq f(x) + f'(x)(y - x), f(x) \geq f(y) + f'(y)(x - y)$

$\Rightarrow f'(x)(y - x) \leq f(y) - f(x) \leq f'(y)(y - x) \geq f'(x)$

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

$\Rightarrow f''(x) \geq 0, \forall x \in \mathbf{dom}\ f$

Suppose $f''(x) \geq 0, \forall x \in \mathbf{dom}\ f$, then for $t \in [x, y]$, $f'(t) = f'(x) + \int_x^t f''(z)dz \geq f'(x)$, so $\int_x^y f'(z)dz \geq \int_x^y f'(x)dz = f'(x)(y - x)$

$\Rightarrow f(y) - f(x) - f'(x)(y - x) \geq \int_x^y f'(z)dz - \int_x^y f'(z)dz = 0$

According to the first-order condition, f is convex.

Then for the case where $f : \mathbb{R}^n \to \mathbb{R}^n$, we can turn the scalars into vectors and get the same conclusion.

3.11

Because f is convex, $f(x) \geq f(y) + \nabla f(y)^T(x - y), f(y) \geq f(x) + \nabla f(x)^T(y - x)$

$\Rightarrow (\nabla f(x) - \nabla f(y))^T(x - y) \geq 0$ i.e. ∇f is monotone.

The converse is not true:

Consider $\psi(x) = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} x$, then $\nabla \psi$ is monotone but never a gradient of any convex function.

3.12

Let $h(x) = a^T x + b$. Because f is convex and g is concave, $S_1 = \{(x,t)|f(x) \le t\}, S_2 = \{(x,t)|g(x) \ge t\}$ are convex. Because $f(x) \ge g(x)$, $S_1 \cap S_2 = \emptyset$. According to the separation hyperplane theorem, there exists a hyperplane $H = \{(x,t)|(a^T, u)(x^T, t)^T = b\} = \{(x,t)|t = \frac{a^T x - b}{u}\}$ separates S_1 and S_2, that is, there exists an affine function $h(x) = \frac{a^T x - b}{u}$ between f and g.

3.15

(a) for $x > 0$, $\lim\limits_{\alpha \to 0} u_\alpha(x) = \lim\limits_{\alpha \to 0} \frac{x^\alpha - 1}{\alpha} = \lim\limits_{\alpha \to 0} \frac{\frac{d(x^\alpha - 1)}{d\alpha}}{\frac{d\alpha}{d\alpha}} = \lim\limits_{\alpha \to 0} x^\alpha \cdot \log x = \log x = u_0(x)$

(b) $u'_\alpha(x) = x^{\alpha-1} > 0, u''_\alpha(x) = (\alpha - 1)x^{\alpha-2} \le 0 \Rightarrow u_\alpha$ is concave and monotone increasing.

$u_\alpha(1) = \frac{1^\alpha - 1}{\alpha} = 0$

3.16

(a) $f(x)$ is convex.

$f''(x) = e^x \ge 0$

(b) $f(x)$ is neither convex nor concave.

$\nabla^2 f(x) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, which is neither positive semidefinite or negative semidefinite.

(c) $f(x)$ is convex.

$$\nabla^2 f(x) = \frac{1}{x_1 x_2} \begin{pmatrix} \frac{2}{x_1^2} & \frac{1}{x_1 x_2} \\ \frac{1}{x_1 x_2} & \frac{2}{x_2^2} \end{pmatrix} \succeq 0$$

(d) $f(x)$ is neither convex or concave.

$\nabla^2 f(x) = \begin{pmatrix} 0 & \frac{-1}{x_2^2} \\ \frac{-1}{x_2^2} & \frac{2x_1}{x_2^3} \end{pmatrix}$, which is neither positive semidefinite or negative semidefinite.

(e) $f(x)$ is convex.

$$\nabla^2 f(x) = \begin{pmatrix} \frac{2}{x_2} & \frac{-2x_1}{x_2^2} \\ \frac{-2x_1}{x_2^2} & \frac{2x_1^2}{x_2^3} \end{pmatrix} = \frac{2}{x_2} \begin{pmatrix} 1 & 0 \\ \frac{-x_1}{x_2} & 0 \end{pmatrix} \begin{pmatrix} 1 & \frac{-x_1}{x_2} \\ 0 & 0 \end{pmatrix} \succeq 0$$

(f) $f(x)$ is concave.

$$\nabla^2 f(x) = \begin{pmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x^{\alpha-1}x_2^{-\alpha} & \alpha(\alpha-1)x_1^{\alpha}x_2^{-\alpha-1} \end{pmatrix} = \alpha(\alpha-1)x_1^{\alpha}x_2^{1-\alpha} \begin{pmatrix} \frac{1}{x_1} & 0 \\ \frac{-1}{x_2} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{x_1} & \frac{-1}{x_2} \\ 0 & 0 \end{pmatrix} \preceq 0$$

3.18

(a) Let $g(t) = f(Z + tV)$ with $Z \succ 0$ and $V \in \mathbb{S}^n$, then

$$g(t) = \mathbf{tr}((Z+tV)^{-1}) = \mathbf{tr}(Z^{-1}(I + tZ^{-\frac{1}{2}}VZ^{-\frac{1}{2}})^{-1}) = \mathbf{tr}(Z^{-1}Q(I+t\Lambda)^{-1}Q^T) = \sum_{i=1}^{n}(Q^TZ^{-1}Q)_{ii}(1+t\lambda_i)^{-1}.$$

Since $(Q^TZ^{-1}Q)_{ii} > 0$, $g(t)$ is convex. Therefore f is convex.

(b) Let $g(t) = f(Z + tV)$ with $Z \succ 0$ and $V \in \mathbb{S}^n$, then

$$g(t) = (\det(Z+tV))^{\frac{1}{n}} = (\det Z^{\frac{1}{2}} \det(I + tZ^{-\frac{1}{2}}VZ^{-\frac{1}{2}}) \det Z^{\frac{1}{2}})^{\frac{1}{2}} = (\det Z)^{\frac{1}{n}} \left(\prod_{i=1}^{n}(1+t\lambda_i)\right)^{\frac{1}{n}}.$$

Since $\det Z > 0$, $g(t)$ is concave. Therefore f is concave.

3.20

(a) f is the composition of two convex function, a norm and an affine function.

(b) f is the composition of $g(x) = -(\det X)^{\frac{1}{m}}$ and an affine transformation. And g can be proved to be convex.

(c) f is the composition of $g(x) = \mathbf{tr}(X^{-1})$ and an affine transformation. And g is proved convex in 3.18(a).

3.21

(a) f is the pointwise maximum of k functions, and these functions $\|A^{(i)}x - b^{(i)}\|$ are convex because they are the composition of affine functions and norms.

(b) f is the pointwise maximum of a series of convex functions $\{|x_{i_1}| + \cdots + |x_{i_r}| \mid 1 \le i_1 < i_2 < \cdots < i_r \le n\}$.

3.22

(a) $\log \sum_{i=1}^m e^{a_i^T x + b_i}$ is the composition of a convex function and an affine function, so it's convex. i.e. $-\log \sum_{i=1}^m e^{a_i^T x + b_i}$ is concave.

Let $g(x) = -\log \sum_{i=1}^m e^{a_i^T x + b_i}, h(x) = -\log x$, then $f(x) = h(g(x))$.

Because g is concave while h is convex and decreasing, f is convex.

(b) Let $g_1(u, v, x) = u, g_2(u, v, x) = v - \frac{x^T x}{u}, h(x_1, x_2) = -\sqrt{x_1 x_2}$, then $f(u, v, x) = h(g_1(u, v, x), g_2(u, v, x))$.

Because g is concave while h is convex and decreasing in both direction, f is convex.

(c) Let $g_1(u, v, x) = u, g_2(u, v, x) = v - \frac{x^T x}{u}, h(x_1, x_2) = -\log x_1 x_2$, then $f(u, v, x) = h(g_1(u, v, x), g_2(u, v, x))$.

Because g is concave while h is convex and decreasing in both direction, f is convex.

(d) $f(x, t) = -t^{1-\frac{1}{p}}(t - \frac{\|x\|_p^p}{t^{p-1}})^{\frac{1}{p}}$.

Let $h(x_1, x_2) = -x_1^{\frac{1}{p}} x_2^{1-\frac{1}{p}}, g_1(x, t) = t^{1-\frac{1}{p}}, g_2(x, t) = t - \frac{\|x\|_p^p}{t^{p-1}}$, then $f(x, t) = h(g_1(x, t), g_2(x, t))$.

Because g_1, g_2 are concave while h is convex and decreasing in both direction, f is convex.

(e) $f(x, t) = -(p-1)\log t - \log(t - \frac{\|x\|_p^p}{t^{p-1}})$.

Let $g(x, t) = -(p-1)\log t$ and $h(x, t) = -\log(t - \frac{\|x\|_p^p}{t^{p-1}})$, then $f(x, t) = g(x, t) + h(x, t)$, where g and h are both proved convex. So f is convex.

3.23

(a) $f(x,t)$ is the perspective function of a convex function $||x||_p^p$, so it's convex.

(b) Let $g(x,t)=\frac{x^Tx}{t}$, then g is the perspective function of x^Tx, i.e. g is convex.

Since f is the composition of g and an affine transformation, f is convex.
