Exercise 15.6

$$\{x \in \mathbb{R}^2 : x_1 = x_2\}^\circ = \emptyset,$$

$$(\mathbb{R}^2)^\circ = \mathbb{R}^2,$$

$$\emptyset^\circ = \emptyset,$$

$$\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}^\circ = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 < 1\}.$$

Exercise 15.11

(1) We have
$$\frac{\partial f}{\partial x_1} = \frac{a_1 e^{a x_1}}{e^{a_1 x_1} + e^{a_2 x_2}}, \quad \frac{\partial f}{\partial x_2} = \frac{a_2 e^{a_2 x_2}}{e^{a_1 x_1} + e^{a_2 x_2}},$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = -\frac{a_1 e^{a_1 x_1} a_2 e^{a_2 x_2}}{(e^{a_1 x_1} + e^{a_2 x_2})^2},$$

$$\frac{\partial^2 f}{\partial x_1^2} = \frac{a_1^2 e^{a_1 x_1} e^{a_2 x_2}}{(e^{a_1 x_1} + e^{a_2 x_2})^2} \quad \text{and} \quad \frac{\partial^2 f}{\partial x_2^2} = \frac{a_2^2 e^{a_1 x_1} e^{a_2 x_2}}{(e^{a_1 x_1} + e^{a_2 x_2})^2}.$$

Thus
$$F(x) = \frac{e^{a_1 x_1} e^{a_2 x_2}}{(e^{a_1 x_1} + e^{a_2 x_2})^2} \begin{bmatrix} a_1^2 & -a_1 a_2 \\ -a_1 a_2 & a_2^2 \end{bmatrix}.$$

If $d \in \mathbb{R}^2$, then
$$d^T F(x) d = \frac{e^{a_1 x_1} e^{a_2 x_2}}{(e^{a_1 x_1} + e^{a_2 x_2})^2} \cdot (d_1^2 a_1^2 - 2 d_1 d_2 a_1 a_2 + d_2^2 a_2^2)$$
$$= \frac{e^{a_1 x_1} e^{a_2 x_2}}{(e^{a_1 x_1} + e^{a_2 x_2})^2} (d_1 a_1 - d_2 a_2)^2$$
$$\geq 0.$$

So $F(x)$ is positive semidefinite for all $x \in \mathbb{R}^2$.

Hence f is convex.

(2) We have $\dfrac{\partial f}{\partial x_1} = \dfrac{2x_1}{x_2}$, $\dfrac{\partial f}{\partial x_2} = -\dfrac{x_1^2}{x_2^2}$,

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = -\frac{2x_1}{x_2^2}, \quad \frac{\partial^2 f}{\partial x_1^2} = \frac{2}{x_2}, \quad \frac{\partial^2 f}{\partial x_2^2} = +\frac{2x_1^2}{x_2^3}$$

Hence

$$F(x) = \begin{bmatrix} \frac{2}{x_2} & -\frac{2x_1}{x_2^2} \\ -\frac{2x_1}{x_2^2} & \frac{2x_1^2}{x_2^3} \end{bmatrix} = \frac{2}{x_2^3} \begin{bmatrix} x_2^2 & -x_1x_2 \\ -x_1x_2 & x_1^2 \end{bmatrix}$$

and so if $d \in \mathbb{R}^2$, then

$$d^T F(x) d = \frac{2}{x_2^3} (d_1 x_2 - d_2 x_1)^2 \geq 0 .$$

So $F(x)$ is positive semidefinite for all $(x_1, x_2) \in \mathbb{R}^2$ s.t. $x_2 > 0$. Hence f is convex.

(3) We have $\dfrac{\partial f}{\partial x_1} = -\dfrac{1}{2}\sqrt{\dfrac{x_2}{x_1}}$, $\dfrac{\partial f}{\partial x_2} = -\dfrac{1}{2}\sqrt{\dfrac{x_1}{x_2}}$,

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = -\frac{1}{4}\frac{1}{\sqrt{x_1x_2}}, \quad \frac{\partial^2 f}{\partial x_1^2} = \frac{1}{4}\frac{1}{x_1}\sqrt{\frac{x_2}{x_1}}, \quad \frac{\partial^2 f}{\partial x_2^2} = \frac{1}{4}\frac{1}{x_2}\sqrt{\frac{x_1}{x_2}}.$$

Hence

$$F(x) = \begin{bmatrix} \frac{1}{4}\frac{1}{x_1}\sqrt{\frac{x_2}{x_1}} & -\frac{1}{4}\frac{1}{\sqrt{x_1x_2}} \\ -\frac{1}{4}\frac{1}{\sqrt{x_1x_2}} & \frac{1}{4}\frac{1}{x_2}\sqrt{\frac{x_1}{x_2}} \end{bmatrix}$$

$$= \frac{1}{4x_1x_2\sqrt{x_1x_2}} \begin{bmatrix} x_2^2 & -x_1x_2 \\ -x_1x_2 & x_1^2 \end{bmatrix}.$$

Thus if $d \in \mathbb{R}^2$, then $d^T F(x) d = \dfrac{1}{4x_1x_2\sqrt{x_1x_2}} (d_1x_2 - d_2x_1)^2 \geq 0$

So $F(x)$ is positive semidefinite whenever $x_1 > 0$ and $x_2 > 0$. Hence f is convex.

Exercise 15.12

(1) (a) If $x, y \in C$ and $t \in (0,1)$, then

$$g(tx + (1-t)y) \leq t g(x) + (1-t) g(y)$$

But f is increasing. Hence

$$f(g(tx + (1-t)y)) \leq f(t g(x) + (1-t) g(y)) \quad ①$$

Finally from the convexity of f,

$$f(t g(x) + (1-t) g(y)) \leq t f(g(x)) + (1-t) f(g(y)) \quad ②$$

From ① and ② it follows that

$$(f \circ g)(tx + (1-t)y) \leq t (f \circ g)(x) + (1-t)(f \circ g)(y).$$

(b) Let $x, y \in C$ and $t \in (0,1)$. Then

$$g(tx + (1-t)y) \geq t g(x) + (1-t) g(y).$$

Since f is decreasing,

$$f(g(tx + (1-t)y)) \leq f(t g(x) + (1-t) g(y)) \quad ③$$

By the convexity of f,

$$f(t g(x) + (1-t) g(y)) \leq t f(g(x)) + (1-t) g(y) \quad ④.$$

From ③ and ④, we have

$$f(g(tx + (1-t)y)) \leq t f(g(x)) + (1-t) f(g(y)),$$

that is $(f \circ g)(tx + (1-t)y) \leq t (f \circ g)(x) + (1-t)(f \circ g)(y)$.

(2) The function $x \overset{f}{\longmapsto} \frac{1}{x} : (0, \infty) \to (0, \infty)$ is convex†. Also it is decreasing. By the preview exercise $f \circ g$ is convex. But $(f \circ g)(x) = f(g(x)) = \dfrac{1}{g(x)}$.

(† For instance since for $x, y \in (0, \infty)$,

$$(\nabla f(y) - \nabla f(x))(y - x) = \left(-\frac{1}{y^2} + \frac{1}{x^2}\right)(y - x) = \frac{(y-x)^2 (y+x)}{y^2 x^2} \geq 0 .)$$

(3). (a) It can be checked that the Hessian is

$$F(x) = \frac{1}{\sum_i e^{a_i x_i}} \begin{bmatrix} a_1^2 e^{a_1 x_1} & & \\ & \ddots & \\ & & a_n^2 e^{a_n x_n} \end{bmatrix} - \frac{1}{\left(\sum_i e^{a_i x_i}\right)^2} \begin{bmatrix} a_1 e^{a_1 x_1} \\ \vdots \\ a_n e^{a_n x_n} \end{bmatrix} \begin{bmatrix} a_1 e^{a_1 x_1} & \dots & a_n e^{a_n x_n} \end{bmatrix}$$

We show this is positive semidefinite as follows:

If $v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$, then

$$v^T F(x) v = \frac{\left(\sum_i a_i^2 v_i^2 e^{a_i x_i}\right)\left(\sum_i e^{a_i x_i}\right) - \left(\sum_i a_i v_i e^{a_i x_i}\right)^2}{\left(\sum_i e^{a_i x_i}\right)^2}$$

But by the Cauchy-Schwarz inequality,

$$\left(\sum_i a_i v_i e^{\frac{1}{2} a_i x_i} \cdot e^{\frac{1}{2} a_i x_i}\right)^2 \le \left(\sum_i a_i^2 v_i^2 e^{a_i x_i}\right)\left(\sum_i e^{a_i x_i}\right).$$

Hence the result follows.

(b) $f(x) = \|x\|$. Let $x, y \in \mathbb{R}^n$ and $t \in (0,1)$

We have
$$\begin{aligned} f(tx + (1-t)y) &= \| tx + (1-t)y \| \\ &\le \|tx\| + \|(1-t)y\| \\ &= |t| \, \|x\| + (1-t) \, \|y\| \\ &= t\, f(x) + (1-t)\, f(y) \end{aligned}$$

(c) It can be seen that the Hessian is

$$F(x) = \frac{\sqrt[n]{x_1 \cdots x_n}}{n} \left(\begin{bmatrix} \frac{1}{x_1^2} & & \\ & \ddots & \\ & & \frac{1}{x_n^2} \end{bmatrix} - \frac{1}{n} \begin{bmatrix} \frac{1}{x_1} \\ \vdots \\ \frac{1}{x_n} \end{bmatrix} \begin{bmatrix} \frac{1}{x_1} & \cdots & \frac{1}{x_n} \end{bmatrix} \right)$$

So if $v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$, then

$$v^T F(x) v = \frac{\sqrt[n]{x_1 \cdots x_n}}{n} \left(\sum_i \frac{v_i^2}{x_i^2} - \frac{1}{n} \left(\sum_i \frac{v_i}{x_i} \right)^2 \right)$$

$$= \frac{\sqrt[n]{x_1 \cdots x_n}}{n^2} \left(\left(\sum_i \frac{v_i^2}{x_i^2}\right)\left(\sum_i 1^2\right) - \left(\sum_i \frac{v_i}{x_i} \cdot 1\right)^2 \right) \geqslant 0, \text{ by}$$

the Cauchy-Schwarz inequality.

Exercise 15.13

The function $x \overset{f}{\longmapsto} -\log x : (0,\infty) \longrightarrow \mathbb{R}$ is convex
$\left(\frac{d^2}{dx^2}(-\log x) = \frac{1}{x^2} > 0 \quad \forall x \in (0,\infty)\right)$

Hence $\forall \, x_1, \ldots, x_n \in (0,\infty)$,

$$f\left(\frac{x_1 + \ldots + x_n}{n}\right) \leq \frac{f(x_1) + \ldots + f(x_n)}{n}.$$

i.e., $$-\log \frac{x_1 + \ldots + x_n}{n} \leq \frac{(-\log x_1) + \ldots + (-\log x_n)}{n}$$

$$= -\log \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}$$

i.e., $$\log \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} \leq \log \left(\frac{x_1 + \ldots + x_n}{n}\right).$$

Since the exponential is an increasing function, it follows that $$e^{\log \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}} \leq e^{\log \left(\frac{x_1 + \ldots + x_n}{n}\right)}$$

i.e., $$\left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} \leq \frac{x_1 + \ldots + x_n}{n}.$$

Exercise 15.14

(1) If $h : \mathbb{R}^n \to \mathbb{R}$ is a convex function, if $\omega_1, \dots, \omega_K \geq 0$ and $\sum_{k=1}^{K} \omega_k = 1$, then $h\left(\sum_{k=1}^{K} \omega_k x^{(k)}\right) \leq \sum_{k=1}^{K} \omega_k h(x^{(k)})$ for any given $x^{(1)}, \dots, x^{(K)}$ in $\mathbb{R}^n$.

Proof We use induction on K. The base case is clear.

Suppose the claim is true for some K.

Let $\omega_1, \dots, \omega_{K+1} \geq 0$, $\sum_{k=1}^{K+1} \omega_k = 1$ and $x^{(1)}, \dots, x^{(K)}, x^{(K+1)} \in \mathbb{R}^n$.

Then

$$h\left(\sum_{k=1}^{K+1} \omega_k x^{(k)}\right) = h\left((1-\omega_{K+1}) \sum_{k=1}^{K} \frac{\omega_k}{1-\omega_{K+1}} x^{(k)} + \omega_{K+1} x^{(K+1)}\right)$$

$$\leq (1-\omega_{K+1})\, h\left(\sum_{k=1}^{K} \frac{\omega_k}{1-\omega_{K+1}} x^{(k)}\right) + \omega_{K+1} h(x^{(K+1)})$$

(By induction hypothesis, since $\sum_{k=1}^{K} \frac{\omega_k}{1-\omega_{K+1}} = 1$.)

$$\leq (1-\omega_{K+1}) \sum_{k=1}^{K} \frac{\omega_k}{1-\omega_{K+1}} h(x^{(k)}) + \omega_{K+1} h(x^{(K+1)})$$

$$= \sum_{k=1}^{K+1} \omega_k h(x^{(k)}).$$

□

Hence by the convexity of f

$$\sum_{k=1}^{K} \hat{\omega}_k f(x^{(k)}) \geq f\left(\sum_{k=1}^{K} \omega_k x^{(k)}\right). \quad (1)$$

But $g\left(\sum_{k=1}^{K} \omega_k x^{(k)}\right) \leq \sum_{k=1}^{K} \omega_k g(x^{(k)}) \leq 0$

and so $\sum_{k=1}^{K} \omega_k x^{(k)}$ belongs to the feasible set of (P).

Consequently $f\left(\sum_{k=1}^{K} \omega_k x^{(k)}\right) \geq f(\hat{x})$ (2), by the optimality of $\hat{x}$ for (P).

From (1) and (2), we obtain

$$\sum_{k=1}^{K} \hat{\omega}_k f(x^{(k)}) \geq f\left(\sum_{k=1}^{K} \omega_k x^{(k)}\right).$$

(2) Let $\widehat{\omega}_k = 0$ if $k \neq k_*$ and $\widehat{\omega}_{k_*} = 1$.

Then $\widehat{\omega} := [\widehat{\omega}_1 \dots \widehat{\omega}_K]^T$ is s.t. $\widehat{\omega}_k \geqslant 0 \quad \forall\, k = 1, \dots, K$.

Also $\sum_{k=1}^{K} \widehat{\omega}_k = 1$.

We have $\sum_{k=1}^{K} \widehat{\omega}_k\, g(x^{(k)}) = g(x^{(k_*)}) = g(\hat{x}) \leq 0$

since $\hat{x}$ is feasible for (P).

Thus $\widehat{\omega}$ is feasible for (LP).

Moreover,
$$\sum_{k=1}^{K} \widehat{\omega}_k\, f(x^{(k)}) = f(x^{(k_*)}) = f(\hat{x}). \quad ③$$

By part (1), we know that if ω is feasible for (LP), then
$$\sum_{k=1}^{K} \omega_k\, f(x^{(k)}) \geqslant f(\hat{x}). \quad ④$$

③ and ④ show that $\hat{\omega}$ is optimal for (LP) and that the optimal values of (P) and (LP) are equal.

Exercise 15.15

We have

$$\frac{\partial f}{\partial x_1}(x) = 2x_1 + 2x_2,$$

$$\frac{\partial f}{\partial x_2}(x) = 10x_2 + 2x_1 + 4x_3 + 4x_2^3, \text{ and}$$

$$\frac{\partial f}{\partial x_3}(x) = 2a x_3 + 4x_2.$$

Hence

$$\frac{\partial^2 f}{\partial x_1^2}(x) = 2, \quad \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) = 2, \quad \frac{\partial^2 f}{\partial x_3 \partial x_1}(x) = 0,$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(x) = 2, \quad \frac{\partial^2 f}{\partial x_2^2}(x) = 10 + 12x_2^2, \quad \frac{\partial^2 f}{\partial x_3 \partial x_2}(x) = 4,$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_3}(x) = 0, \quad \frac{\partial^2 f}{\partial x_2 \partial x_3}(x) = 4, \quad \frac{\partial^2 f}{\partial x_3^2}(x) = 2a.$$

So the Hessian $F(x)$ of f at x is given by

$$F(x) = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 10 + 12x_2^2 & 4 \\ 0 & 4 & 2a \end{bmatrix}.$$

The interior of $C := \mathbb{R}^3$ is not empty, and so we know that f is convex iff $F(x)$ is positive semidefinite at all $x \in C = \mathbb{R}^3$.

For some elementary matrix E_1 we have

$$E_1 F(x) E_1^T = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 8 + 12x_2^2 & 4 \\ 0 & 4 & 2a \end{bmatrix}.$$

As $x_2^2 \geq 0$, we have $8 + 12x_2^2 > 0$. So for an appropriate matrix E_2 we have

$$E_2 E_1 F(x) E_1^T E_2^T = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 8+12x_2^2 & 0 \\ 0 & 0 & 2a - \frac{4}{2+3x_2^2} \end{bmatrix}.$$

So we see that f is convex iff $2a - \frac{4}{2+3x_2^2} \geqslant 0 \quad \forall x_2 \in \mathbb{R}$

(since $2 > 0$ and $8 + 12x_2^2 > 0$.)

We claim that f is convex iff $a \geqslant 1$.

<u>If</u> : $a \geqslant 1$ and $1 \geqslant \frac{2}{2+3x_2^2}$ gives $a \geqslant \frac{2}{2+3x_2^2}$,

and so $2a \geqslant \frac{4}{2+3x_2^2}$ i.e., $2a - \frac{4}{2+3x_2^2} \geqslant 0$.

Hence f is convex.

<u>Only if</u> : If f is convex, then $2a - \frac{4}{2+3x_2^2} \geqslant 0 \quad \forall x_2 \in \mathbb{R}$.

In particular, take $x_2 = 0$. So $2a \geqslant \frac{4}{2+0}$ i.e., $2a \geqslant 2$
i.e., $a \geqslant 1$.

So f is convex iff $a \geqslant 1$.