Homework 2

1b. Notice that $f'(x) = -e^{-x}$ and $f''(x) = e^{-x} > 0$ hold for all x. By Remark II on page 48 (or Theorem 2.3.7), $f(x)$ is strictly convex on R^1.

1c. We prove that $f(x)$ is convex by using the definition. That is, for any x_1, x_2 in $I = [-1, 1]$ and any λ with $0 \leq \lambda \leq 1$, we need to show that

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2). \tag{1}$$

Equivalently,

$$|\lambda x_1 + (1-\lambda)x_2| \leq \lambda |x_1| + (1-\lambda)|x_2|. \tag{2}$$

We consider two cases. If x_1 and x_2 have the same sign,

$$|\lambda x_1 + (1-\lambda)x_2| = |\lambda x_1| + |(1-\lambda)x_2| = \lambda|x_1| + (1-\lambda)|x_2|,$$

so (2) holds (in fact, with equality). If x_1 and x_2 have different signs,

$$\begin{aligned} &|\lambda x_1 + (1-\lambda)x_2| \\ &= \max\{|\lambda x_1|,\ |(1-\lambda)x_2|\} - \min\{|\lambda x_1|,\ |(1-\lambda)x_2|\} \\ &\leq \max\{|\lambda x_1|,\ |(1-\lambda)x_2|\} \\ &\leq \lambda|x_1| + (1-\lambda)|x_2|, \end{aligned}$$

so (2) holds again. Thus $f(x)$ is convex.

$f(x)$ is not strictly convex because, as we have seen that (1) holds with equality when x_1 and x_2 have the same sign (even when they have different values).

2c. For $x > 1$, let us define

$$f_1(x) = x,\ f_2(x) = 2x+1,\ f_3(x) = x^8, \text{ and } f_4(x) = -\ln x.$$

Then

$$\begin{aligned} f(x_1, x_2) &= (x_1 + 2x_2 + 1)^8 - \ln(x_1 x_2)^2 \\ &= (x_1 + 2x_2 + 1)^8 - 2\ln x_1 - 2\ln x_2 \\ &= f_3(f_1(x_1) + f_2(x_1)) + 2f_4(x_1) + 2f_4(x_2). \end{aligned}$$

Observe that

- $f_1(x)$, $f_2(x)$ are convex, since $f_1''(x) = f_2''(x) = 0 \geq 0$;
- $f_3(x)$ is increasing and convex, as $f_3'(x) = 8x^7 > 0$ (for $x > 1$) and $f_3''(x) = 56x^6 \geq 0$;
- $f_4(x)$ is strictly convex, as $f_2''(x) = 1/x^2 > 0$ (for $x > 1$).

By Theorem 2.3.10, $f(x)$ is strictly convex.

Alternatively, it is straightforward to verify that

$$Hf(x_1, x_2) = \begin{bmatrix} \frac{2}{x_1^2} + z & 2z \\ 2z & \frac{2}{x_2^2} + 4z \end{bmatrix},$$

where $z = 56(x_1 + 2x_2 + 1)^6$. When considering the leading principal minors we have $\Delta_1 = \frac{2}{x_1^2} + z > 0$ and $\Delta_2 = \frac{4}{x_1^2 x_2^2} + \frac{8z}{x_1^2} + \frac{2z}{x_2^2} > 0$, for all (x_1, x_2) in the domain D. It follows that Hf is positive definite, for all points in the domain. Therefore, by Theorem 2.3.7, f is strictly convex on D.

2d. Let $f_1(x) = 3x$, $f_2(x) = -x$, $f_3(x) = x^2$, and $f_4(x) = e^x$. Then

$$\begin{aligned} f(x_1, x_2) &= 4e^{3x_1 - x_2} + 5e^{x_1^2 + x_2^2} \\ &= 4f_4(f_1(x_1) + f_2(x_2)) + 5f_4(f_3(x_1) + f_3(x_2)). \end{aligned}$$

Observe that

- $f_1(x)$, $f_2(x)$ are convex, since $f_1''(x) = f_2''(x) = 0 \geq 0$;
- $f_3(x)$ is convex, as $f_3''(x) = 2 \geq 0$;
- $f_4(x)$ is increasing and convex, as $f_4'(x) = f_4''(x) = e^x > 0$.

By Theorem 2.3.10, $f(x)$ is strictly convex.

15c. Let $f(x, y, z) = 3x + 4y + 12z$. By AG inequality,

$$\frac{f(x, y, x)}{3} \geq \sqrt[3]{3x \cdot 4y \cdot 12z} = \sqrt[3]{144xyz} = \sqrt[3]{144} = 2\sqrt[3]{18}.$$

Therefore, the minimum value of $f(x, y, z)$ is $6\sqrt[3]{18}$, which is obtained when $3x = 4y = 12z = 2\sqrt[3]{18}$, or equivalently, when $x = 2\sqrt[3]{18}/3$, $y = \sqrt[3]{18}/2$, and $z = \sqrt[3]{18}/6$.

15d. Let $f(x, y, z) = xy^2z^3$. By AG inequality,

$$\begin{aligned} 39 &= x^3 + y^2 + z \\ &= 13\frac{x^3 + \frac{y^2}{3} + \frac{y^2}{3} + \frac{y^2}{3} + \frac{z}{9} + \frac{z}{9} + \frac{z}{9} + \frac{z}{9} + \frac{z}{9} + \frac{z}{9} + \frac{z}{9} + \frac{z}{9} + \frac{z}{9}}{13} \\ &\geq 13\sqrt[13]{x^3 \cdot \frac{y^2}{3} \cdot \frac{y^2}{3} \cdot \frac{y^2}{3} \cdot \frac{z}{9} \cdot \frac{z}{9} \cdot \frac{z}{9} \cdot \frac{z}{9} \cdot \frac{z}{9} \cdot \frac{z}{9} \cdot \frac{z}{9} \cdot \frac{z}{9} \cdot \frac{z}{9}} \\ &= 13\sqrt[13]{3^{-21}x^3y^6z^9} \\ &= 13 \cdot 3^{-21/13} \cdot f^{3/13}. \end{aligned}$$

Therefore, the maximum value of $f(x, y, z)$ is $3^{34/3}$ (obtained by solving $39 = 13 \cdot 3^{-21/13} \cdot f^{3/13}$), which is achieved when $x^3 = \frac{y^2}{3} = \frac{z}{9} = \frac{39}{13}$, or equivalently, when $x = 3^{1/3}$, $y = 3$, and $z = 3^3$.

24. We first show that $g(\mathbf{x})$ is convex. For any $\mathbf{x}$ and $\mathbf{y}$ in R^n and λ in $[0, 1]$, we have

$$\begin{aligned} g(\lambda\mathbf{x} + [1-\lambda]\mathbf{y}) &= f(\mathbf{a} \cdot (\lambda\mathbf{x} + [1-\lambda]\mathbf{y}) + \alpha) \\ &= f(\lambda(\mathbf{a} \cdot \mathbf{x}) + [1-\lambda](\mathbf{a} \cdot \mathbf{y}) + \alpha) \\ &= f(\lambda(\mathbf{a} \cdot \mathbf{x} + \alpha) + [1-\lambda](\mathbf{a} \cdot \mathbf{y} + \alpha)) \\ &\leq \lambda f(\mathbf{a} \cdot \mathbf{x} + \alpha) + [1-\lambda] f(\mathbf{a} \cdot \mathbf{y} + \alpha) \\ &= \lambda g(\mathbf{x}) + [1-\lambda] g(\mathbf{y}), \end{aligned}$$

where the inequality follows from the assumption that f is convex. By definition, $g(\mathbf{x})$ is convex on R^n.

Next, we show that $g(\mathbf{x})$ is not strictly convex if $n \geq 2$. In this case, $\mathbf{a} \cdot \mathbf{x} = 0$ has a nonzero solution $\mathbf{x}^*$. (Geometrically, $\mathbf{x}^*$ is orthogonal with $\mathbf{a}$. Such a vector $\mathbf{x}^*$ does not exist if $n = 1$ and $\mathbf{a} \neq 0$.) Let us choose $\lambda = 1/2$. Then

$$g(\frac{1}{2}\mathbf{0} + \frac{1}{2}\mathbf{x}^*) = g(\frac{1}{2}\mathbf{x}^*) = f(\frac{1}{2}(\mathbf{a} \cdot \mathbf{x}^*) + \alpha) = f(\alpha).$$

On the other hand,

$$\frac{1}{2}g(\mathbf{0}) + \frac{1}{2}g(\mathbf{x}^*) = \frac{1}{2}f(\mathbf{a} \cdot \mathbf{0} + \alpha) + \frac{1}{2}f(\mathbf{a} \cdot \mathbf{x}^* + \alpha) = \frac{1}{2}f(\alpha) + \frac{1}{2}f(\alpha),$$

which implies $g(\frac{1}{2}\mathbf{0} + \frac{1}{2}\mathbf{x}^*) = \frac{1}{2}g(\mathbf{0}) + \frac{1}{2}g(\mathbf{x}^*)$, so $g(\mathbf{x})$ is not strictly convex.

We now consider $g(x, y, z) = (4x + 5y - 8z + 17)^8$. If we take $\mathbf{a} = (4, 5, -8)$, $\mathbf{x} = (x, y, z)$, $\alpha = 17$, and $f(x) = x^8$, then $g(\mathbf{x}) = f(\mathbf{a} \cdot \mathbf{x} + \alpha)$. Since $f''(x) = 56x^6 \geq 0$, for all x, we know that $f(x)$ is convex. Therefore, the above results imply that $g(x, y, z)$ is convex but not strictly convex on R^3.

Finally, we point out that Theorem 2.3.10(c) is not applicable here. Notice that $g(\mathbf{x}) = f(h(\mathbf{x}))$ is the composition of $f(x) = x^8$ and $h(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} + \alpha$. In order to apply Theorem 2.3.10(c), we have to ensure that $f(x) = x^8$ is increasing (this is the function corresponds to $g(y)$ in the theorem). However, $f(x) = x^8$ is not increasing and thus Theorem 2.3.10(c) cannot be used to make a conclusion.