Problem 1

Proof. Suppose S is a nonempty subset of V. Then clearly $conv(S)$ is nonempty (a necessary condition for $conv(S)$ to be convex). Choose $x, y \in conv(S)$ and $\lambda \in [0, 1]$. Then we have that:

$$\begin{aligned}\lambda x + (1-\lambda)y &= \lambda(\gamma_1 x_1 + \cdots + \gamma_m x_m) + (1-\lambda)(\delta_1 y_1 + \cdots + \delta_n y_n) \\ &= \lambda\gamma_1 x_1 + \cdots + \lambda\gamma_m x_m + (1-\lambda)\delta_1 y_1 + \cdots + (1-\lambda)\delta_n y_n\end{aligned}$$

Note that the resulting sum is a convex combination of at most $m + n \in \mathbb{N}$ elements of S. Each coefficient is nonnegative (since each γ_i and δ_i is nonnegative, and both λ and $1-\lambda$ are nonnegative). Furthermore, the sum of the coefficients:

$$\lambda\gamma_1 + \cdots + \lambda\gamma_m + (1-\lambda)\delta_1 + \cdots + (1-\lambda)\delta_n = \lambda\sum_{i=1}^{m}\gamma_i + (1-\lambda)\sum_{j=1}^{n}\delta_j = \lambda\cdot 1 + (1-\lambda)\cdot 1 = 1$$

as required. Thus, $\lambda x + (1-\lambda)y \in conv(S)$, so $conv(S)$ is convex. □

Problem 2

Part (i)

Proof. Consider a hyperplane P in V s.t. $P = \{x \in V \mid \langle a, x\rangle = b\}$ with $a \in V$, $a \neq 0$, $b \in \mathbb{R}$. Choose $x, y \in P$ and $\lambda \in [0, 1]$. Then we have that:

$$\begin{aligned}\langle a, \lambda x + (1-\lambda)y\rangle &= \lambda\langle a, x\rangle + (1-\lambda)\langle a, y\rangle \\ &= \lambda b + (1-\lambda)b \\ &= b\end{aligned}$$

Thus, $\lambda x + (1-\lambda)y \in P$ by definition. Thus, P is convex. □

Part (ii)

Proof. Consider a hyperplane H in V s.t. $H = \{x \in V \mid \langle a, x\rangle \leq b\}$ with $a \in V$, $a \neq 0$, $b \in \mathbb{R}$. Choose $x, y \in H$ and $\lambda \in [0, 1]$. Then we have that:

$$\begin{aligned}\langle a, \lambda x + (1-\lambda)y\rangle &= \lambda\langle a, x\rangle + (1-\lambda)\langle a, y\rangle \\ &\leq \lambda b + (1-\lambda)b \\ &\leq b\end{aligned}$$

Thus, $\lambda x + (1-\lambda)y \in H$ by definition. Thus, H is convex. □

Problem 4

Part (i)

Proof.

$$\begin{aligned}
\|x-y\|^2 &= \langle x-y, x-y\rangle \\
&= \langle x-p+p-y, x-p+p-y\rangle \\
&= \langle x-p+p-y, x-p\rangle + \langle x-p+p-y, p-y\rangle \\
&= \langle x-p, x-p+p-y\rangle + \langle p-y, x-p+p-y\rangle \\
&= \langle x-p, x-p\rangle + \langle x-p, p-y\rangle + \langle p-y, x-p\rangle + \langle p-y, p-y\rangle \\
&= \langle x-p, x-p\rangle + \langle p-y, p-y\rangle + 2\langle x-p, p-y\rangle \\
&= \|x-p\|^2 + \|p-y\|^2 + 2\langle x-p, p-y\rangle
\end{aligned}$$

□

Part (ii)

Proof.

$$\begin{aligned}
\|x-y\|^2 &= \|x-p\|^2 + \|p-y\|^2 + 2\langle x-p, p-y\rangle \\
&\geq \|x-p\|^2 + \|p-y\|^2 && \text{by (7.14)} \\
&> \|x-p\|^2 && \text{since } y \neq p
\end{aligned}$$

Thus, since $\|x-y\| \geq 0$ and $\|x-p\| \geq 0$, it follows that $\|x-y\| > \|x-p\|$. □

Part (iii)

Proof.

$$\begin{aligned}
\|x-z\|^2 &= \|x-\lambda y-(1-\lambda)p\|^2 \\
&= \langle x-\lambda y-(1-\lambda)p, x-\lambda y-(1-\lambda)p\rangle \\
&= \langle x-\lambda y-(1-\lambda)p, x-p\rangle + \lambda\langle x-\lambda y-(1-\lambda)p, p-y\rangle \\
&= \langle x-p, x-\lambda y-(1-\lambda)p\rangle + \lambda\langle p-y, x-\lambda y-(1-\lambda)p\rangle \\
&= \langle x-p, x-p\rangle + \lambda\langle x-p, p-y\rangle + \lambda\langle p-y, x-p\rangle + \lambda^2\langle p-y, p-y\rangle \\
&= \|x-p\|^2 + 2\lambda\langle x-p, p-y\rangle + \lambda^2\|y-p\|^2
\end{aligned}$$

□

Part (iv)

Proof. Since $p \in C$ is a projection of x onto C, we know that $\|x-z\|^2 \geq \|x-p\|^2$ for $z \in C$ s.t. $z = \lambda y - (1-\lambda)p$ for some $y \in C$, $\lambda \in [0,1]$. So, we have that:

$$\begin{aligned} 0 &\leq \|x-z\|^2 - \|x-p\|^2 \\ &= 2\lambda\langle x-p, p-y\rangle + \lambda^2\|y-p\|^2 && \text{by part (iii)} \end{aligned}$$

In particular, $0 \leq 2\langle x-p, p-y\rangle + \lambda\|y-p\|^2$ $\forall y \in C, \lambda = 0$, i.e. $0 \leq \langle x-p, p-y\rangle$ $\forall y \in C$. □

Problem 6

Proof. Consider the set $S \coloneqq \{x \in \mathbb{R}^n \mid f(x) \leq c\}$. Choose $x, y \in S$ and $\lambda \in [0,1]$. Then we have that:

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &\leq \lambda f(x) + (1-\lambda)f(y) \\ &\leq \lambda c + (1-\lambda)c \\ &= c \end{aligned}$$

Thus, $\lambda x + (1-\lambda)y \in S$, so S is convex. □

Problem 7

Consider $x, y \in C$ and $\gamma \in [0,1]$. Then we have that:

Proof.

$$\begin{aligned} f(\gamma x + (1-\gamma)y) &= \sum_{i=1}^{k} \lambda_i f_i(\gamma x + (1-\gamma)y) \\ &\leq \sum_{i=1}^{k} \lambda_i(\gamma f_i(x) + (1-\gamma)f_i(y)) \\ &= \gamma \sum_{i=1}^{k} \lambda_i f_i(x) + (1-\gamma)\sum_{i=1}^{k} \lambda_i f_i(y) \\ &= \gamma f(x) + (1-\gamma)f(y) \end{aligned}$$

Thus, f is convex by definition. □

Problem 13

Proof. Suppose not. Then $\exists a, b \in \mathbb{R}^n$ s.t. $f(a) \neq f(b)$. Say that f is bounded above by some constant $M \in \mathbb{R}$. Consider the set $S \subset \mathbb{R}^{n+1}$ with $S := \{x \in \mathbb{R}^{n+1} \mid \exists \lambda \in [0,1]\ s.t.\ \lambda(a, f(a)) + (1-\lambda)x = (b, f(b))$. This is the set of all points on the line between $(a, f(a)) \in \mathbb{R}^{n+1}$ and $(b, f(b)) \in \mathbb{R}^{n+1}$. Since $f(a) \neq f(b)$, we must have that $\exists y \in S$ s.t. the last element of y, call it y_{n+1}, is greater than M. Define $c := (y_1, \ldots, y_n)$. Then we must have that $\exists \lambda \in [0,1]\ s.t.\ \lambda a + (1-\lambda)c = b$. So, $f(b) = f(\lambda a + (1-\lambda)c) \leq \lambda f(a) + (1-\lambda) f(c)$. But we know that $f(b) = \lambda f(a) + (1-\lambda) y_{n+1}$. So, we must have $f(c) \geq y_{n+1} > M$. So, we have reached a contradiction. $\square$

Problem 20

Proof. Choose $x, y \in \mathbb{R}^n$ and $\lambda \in [0,1]$. Then we have that:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \qquad -f(\lambda x + (1-\lambda)y) \leq -\lambda f(x) - (1-\lambda) f(y)$$

It follows then that $f(\lambda x + (1-\lambda)y) = \lambda f(x) + (1-\lambda) f(y)$. Consider the function $g(x) := f(x) - f(0)$. We have that $g(x)$ must be linear since, for $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$:

$$\begin{aligned}
g(x) &= f(x) - f(0) \\
&= f(\frac{1}{\alpha}(\alpha x) + (1 - \frac{1}{\alpha})(0)) - f(0) \\
&= \frac{1}{\alpha} f(\alpha x) + (1 - \frac{1}{\alpha}) f(0) - f(0) \\
&= \frac{1}{\alpha} f(\alpha x) - \frac{1}{\alpha} f(0)
\end{aligned}$$

So, we have that $\alpha g(x) = f(\alpha x) - f(0) = g(\alpha x)$. Also, note that:

$$\begin{aligned}
g(x+y) &= g(2(\frac{1}{2}x + \frac{1}{2}y)) \\
&= 2g(\frac{1}{2}x + \frac{1}{2}y) \\
&= 2(f(\frac{1}{2}x + \frac{1}{2}y) - f(0)) \\
&= 2(\frac{1}{2}f(x) + \frac{1}{2}f(y) - f(0)) \\
&= f(x) - f(0) + f(y) - f(0) \\
&= g(x) + g(y)
\end{aligned}$$

Thus, since $f(x) = g(x) + f(0)$, we have that $f(x)$ affine by definition. $\square$

Problem 21

Proof. Suppose x is a local minimizer of f. Then $\exists \delta > 0$ s.t. $\forall p \in B(x, \delta)$, $f(p) \geq f(x)$. So, $\exists \delta > 0$ s.t. $\forall p \in B(x, \delta)$, $\phi \circ f(p) \geq \phi \circ f(x)$ since ϕ is increasing. So, x is a local minimizer of $\phi \circ f$

Suppose x is a local minimizer of $\phi \circ f$. Then $\exists \delta > 0$ s.t. $\forall p \in B(x, \delta)$, $\phi \circ f(p) \geq \phi \circ f(x)$. So, $\exists \delta > 0$ s.t. $\forall p \in B(x, \delta)$, $\phi^{-1} \circ \phi \circ f(p) \geq \phi^{-1} \circ \phi \circ f(x)$ since ϕ is strictly increasing so ϕ^{-1} is well-defined and increasing. So, x is a local minimizer of f. $\square$