Solution to "Normal and Tangent Cones"

In $\mathbb{R}^n$, we consider the set:

$$\Lambda_n = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n : x_i \geq 0 \ \forall i \text{ and } \sum_{i=1}^{n} x_i = 1 \right\}$$

Part (a): Verify that Λ_n is closed and convex.

Convexity

Let $x \in \Lambda_n$ be an arbitrary vector.

It can be decomposed using the canonical basis into $x = (x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i \hat{\mathbf{e}}_i$.
Since $x_i \geq 0$ for all i and $\sum_{i=1}^{n} x_i = 1$ then x is a convex combination of unit vectors $\hat{\mathbf{e}}_i$.

The convex hull of $\{\hat{\mathbf{e}}_1, \ldots, \hat{\mathbf{e}}_n\}$ is the set of all convex combinations of its elements $\hat{\mathbf{e}}_i$, that is, the set of all x.

Thus, $\Lambda_n = \operatorname{conv}\big(\{\hat{\mathbf{e}}_1, \ldots, \hat{\mathbf{e}}_n\}\big)$ is convex.

Closedness

Let $x_k = (x_{1,k}, \ldots, x_{n,k}) \in \Lambda_n$ be an arbitrary vector sequence with $(x_{1,k})_{k\in\mathbb{N}}, \ldots, (x_{n,k})_{k\in\mathbb{N}}$ being n arbitrary sequences in $\mathbb{R}$ such that: $\lim_{k\to\infty} x_{i,k} = x_i$ for all $i \in \{1, \ldots, n\}$.

Since $x_k \in \Lambda_n$ then obtain:

$$x_{i,k} \geq 0 \ \forall i \in \{1, \ldots, n\} \text{ and } \sum_{i=1}^{n} x_{i,k} = 1$$

By the properties of real-valued limits obtain that: $\lim_{k\to\infty} x_{i,k} \geq \lim_{k\to\infty} 0 \implies x_i \geq 0$ for all i.
Moreover, apply the limit to the sum:

$$\lim_{k\to\infty} \sum_{i=1}^{n} x_{i,k} = \sum_{i=1}^{n} \lim_{k\to\infty} x_{i,k} = \sum_{i=1}^{n} x_i = \lim_{k\to\infty} 1 = 1$$

Realize that the vector $x = \lim_{k\to\infty} x_k = (x_1, \ldots, x_n) \in \Lambda_n$ since:

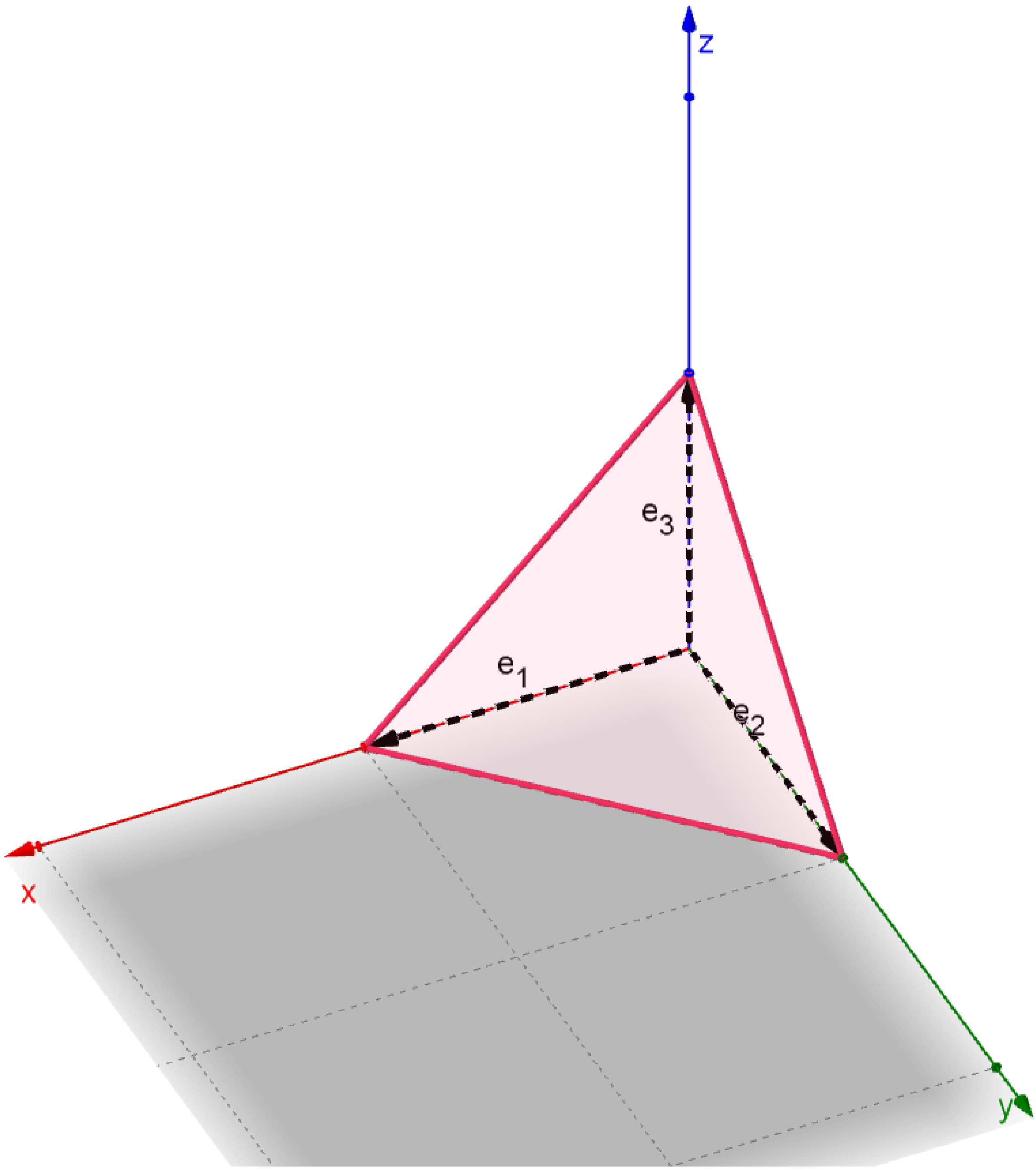
$$x_i \geq 0 \ \forall i \in \{1, \ldots, n\} \text{ and } \sum_{i=1}^{n} x_i = 1$$

Therefore Λ_n is closed.

Part (b): Sketch Λ_3 and find geometrically (without an analytic proof) $N_{\Lambda_3}(x)$ and $T_{\Lambda_3}(x)$ at the points $x = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, $\left(0, \frac{1}{2}, \frac{1}{2}\right)$ and $(0, 0, 1)$.

For $n = 3$, label $(x_1, x_2, x_3) = (x, y, z) \in \mathbb{R}^3$.
Here $\Lambda_3 = \left\{(x, y, z) \in \mathbb{R}^3 : x + y + z = 1 \text{ and } x, y, z \geq 0\right\} = \operatorname{conv}\left(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\right)$.

Geometrically Λ_3 is the surface bounded equilateral triangle formed by $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$.

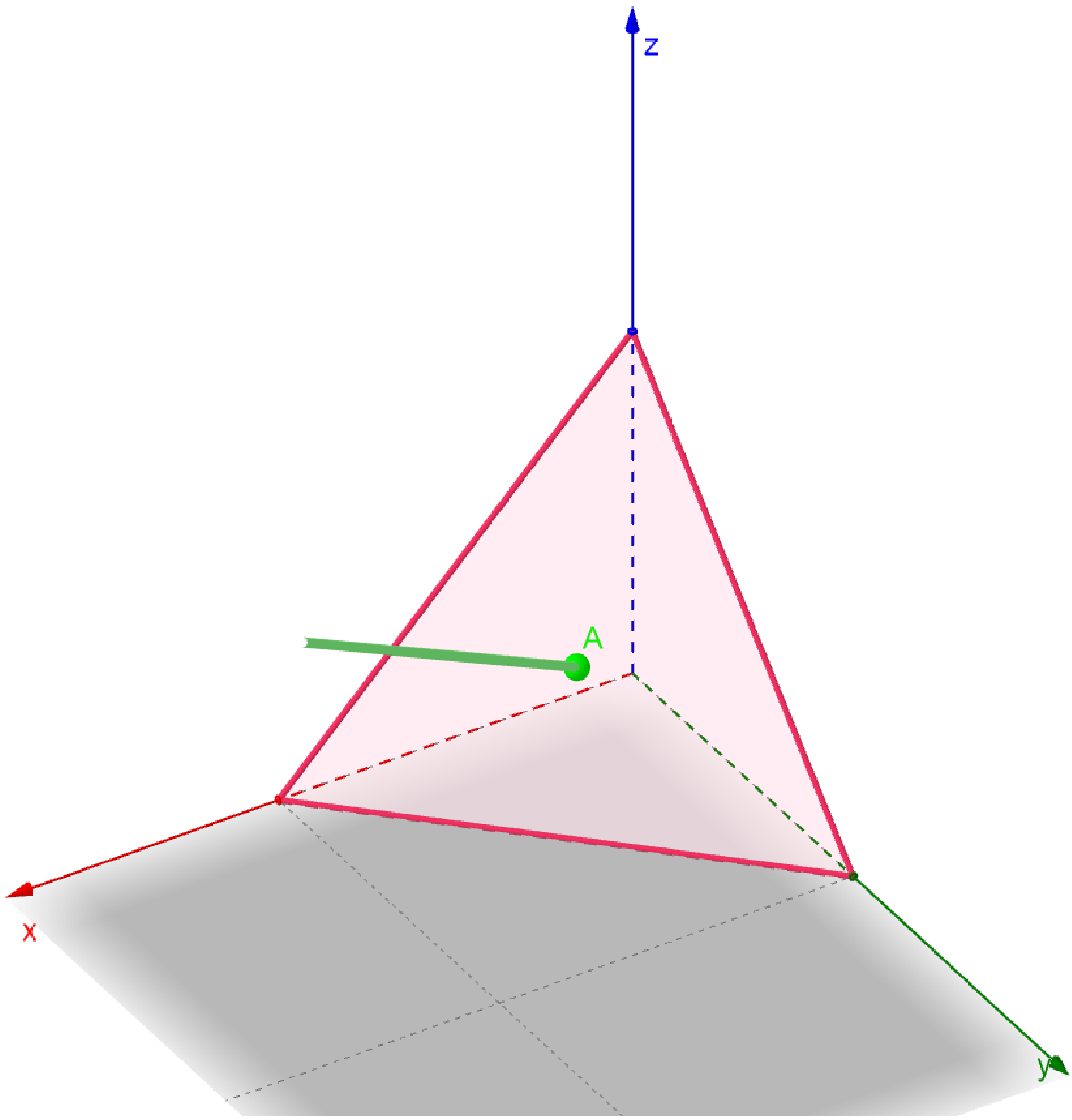


For the interior point $A = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, the normal cone is the orthogonal straight line passing through A.

$$N_{\Lambda_3}(A) = \{(t, t, t) : \forall t \in \mathbb{R}\}.$$

Consequently, the tangent cone at A is the plane containing Λ_3:

$$T_{\Lambda_3}(A) = \{(x, y, z) : x + y + z = 1 \text{ and } x, y, z \in \mathbb{R}\} = \operatorname{aff}\big((\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)\big).$$

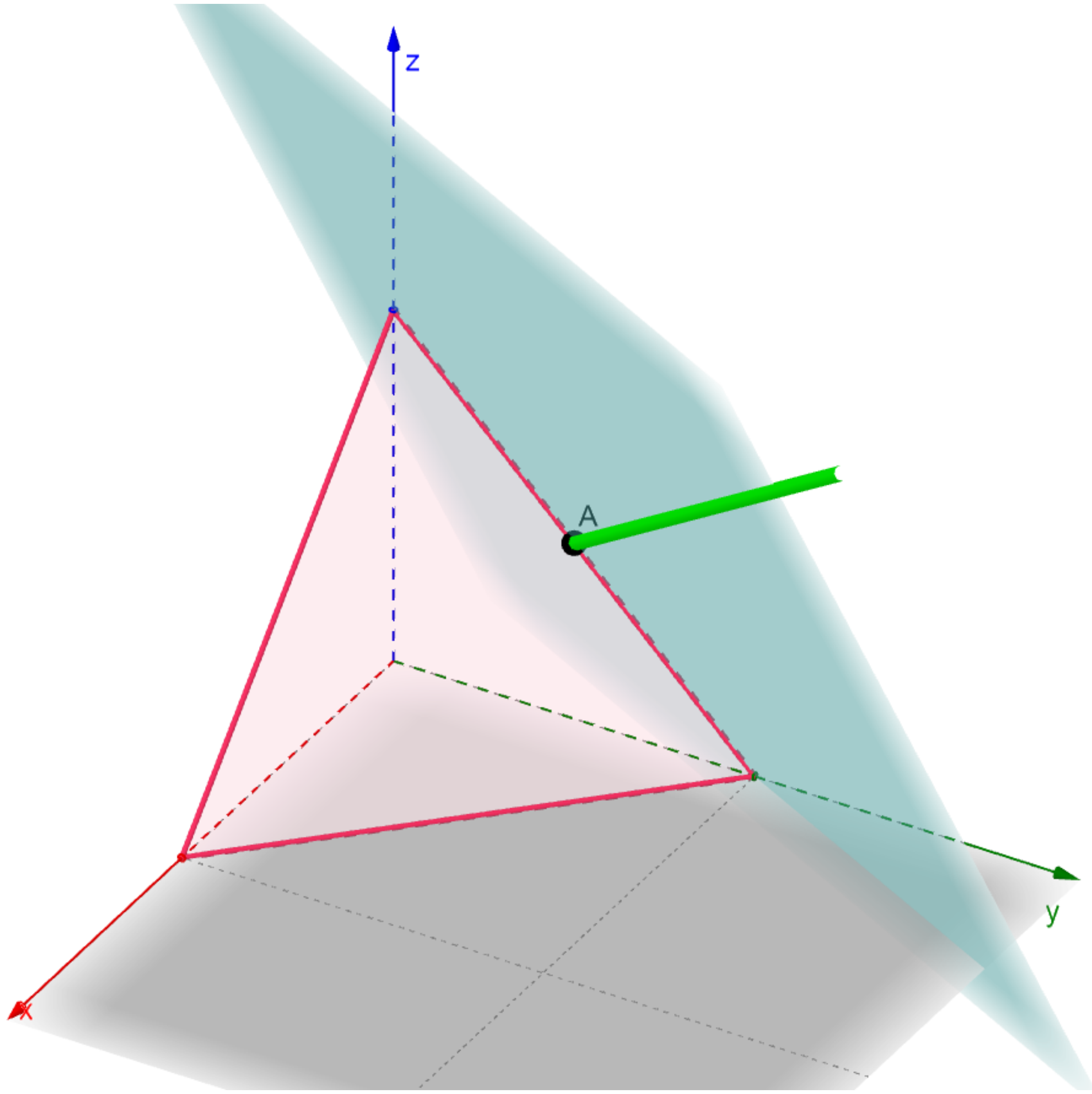


For the edge point $A = \left(0, \frac{1}{2}, \frac{1}{2}\right)$, the normal cone is the plane line passing through A and orthogonal to the edge containing A. This is the plane $y - z = 0$.

$$N_{\Lambda_3}(A) = \{(x, y, z) : y - z = 0 \text{ and } x, y, z \in \mathbb{R}\}.$$

Consequently, the tangent cone at A is also the plane containing Λ_3:

$$T_{\Lambda_3}(A) = \{(x, y, z) : x + y + z = 1 \text{ and } x, y, z \in \mathbb{R}\} = \operatorname{aff}\left(\left(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\right)\right).$$



For the vertex $A = \big(0, 0, 1\big)$, the normal cone is the intersection of the epigraphs of the two planes orthogonal at A to the edges containing A.

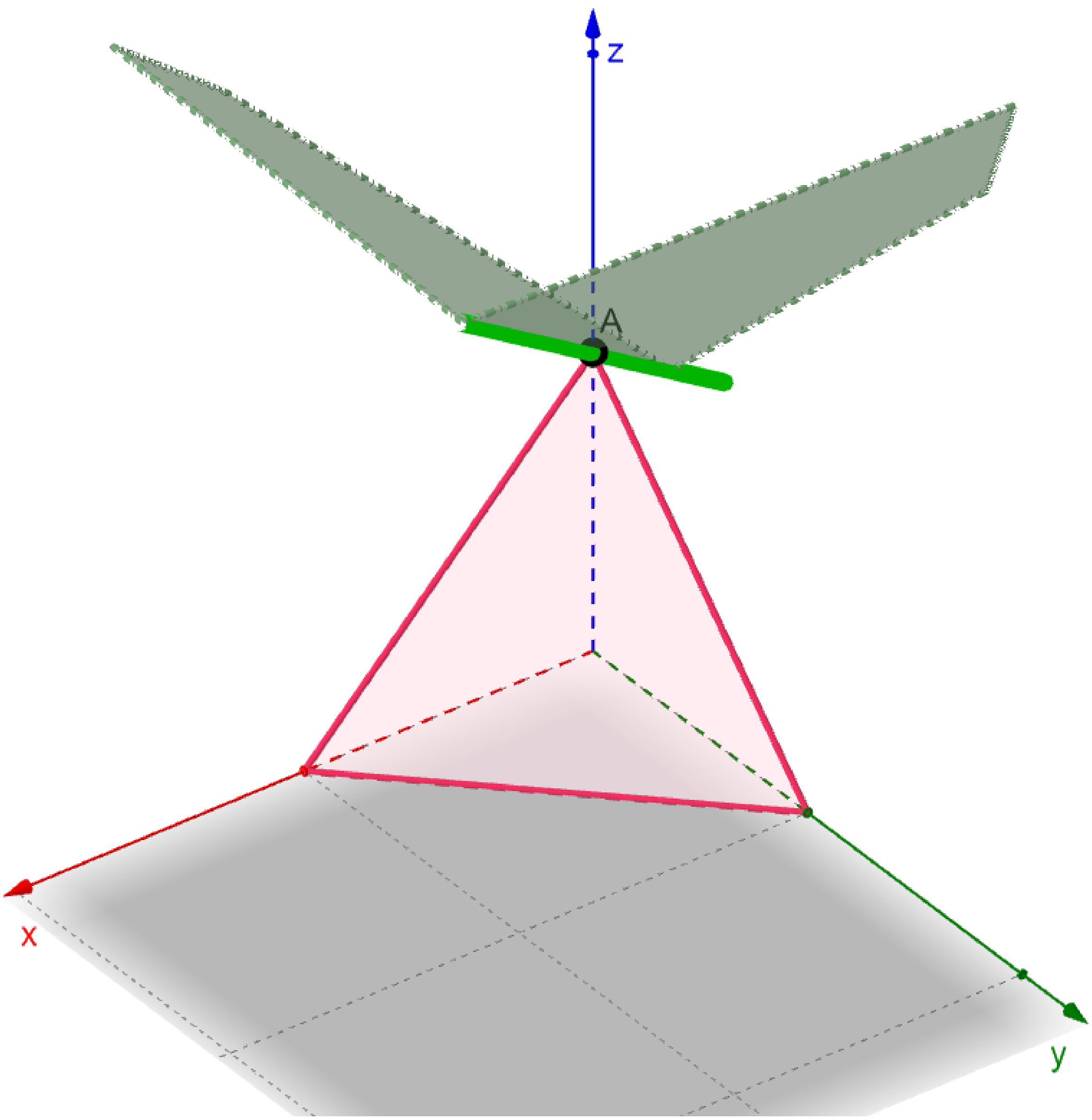
This is the space above both planes $x - z + 1 = 0$ and $y - z + 1 = 0$.

$$N_{\Lambda_3}(A) = \{(x, y, z) : x - z + 1 \geq 0 \text{ and } y - z + 1 \geq 0 \text{ and } x, y, z \in \mathbb{R}\}.$$

Consequently, the tangent cone is the polar cone, that is, the intersection of the hypographs of the two planes passing through A and orthogonal to the edges containing A.

This is the space below both planes $x - 2y + z - 1 = 0$ and $-2x + y + z - 1 = 0$.

$$T_{\Lambda_3}(A) = \{(x, y, z) : x - 2y + z - 1 \leq 0 \text{ and } -2z + y + z - 1 \leq 0 \text{ and } x, y, z \in \mathbb{R}\}.$$



Part (c): For $\bar{x} \in \Lambda_n$, we denote by $I(\bar{x})$ the set of i such that $\bar{x}_i = 0$. Note that $I(\bar{x})$ can be an empty set. Prove that:

i. $T_{\Lambda_n}(\bar{x}) = \big\{(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n : \alpha_i \geq 0 \; \forall i \in I(\bar{x}) \text{ and } \sum_{i=1}^n \alpha_i = 0\big\}$.

Since Λ_n is closed and convex, then,

$$T_{\Lambda_n}(\bar{x}) = \operatorname{cl} \underbrace{\big\{\lambda(x - \bar{x}) : x \in \Lambda_n \text{ and } \lambda \geq 0\big\}}_{A}$$

Let $\alpha \in A$ be an arbitrary vector. Obtain $\alpha = \big(\lambda(x_1 - \bar{x}_1), \dots, \lambda(x_n - \bar{x}_n)\big)$.
Therefore, for all $i \in I(\bar{x})$, $\alpha_i = \lambda(x_i - \bar{x}_i) = \lambda x_i \geq 0$ since $\lambda, x_i \geq 0$.

Moreover, since $x, \bar{x} \in \Lambda_n$, then evaluate the sum to obtain:

$$\sum_{i=1}^n \alpha_i = \lambda \left(\sum_{i=1}^n x_i - \sum_{i=1}^n \bar{x}_i \right) = \lambda(1 - 1) = 0$$

Thus $A = \big\{(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n : \alpha_i \geq 0 \; \forall i \in I(\bar{x}) \text{ and } \sum_{i=1}^n \alpha_i = 0\big\}$.

***Lemma*:** A is closed.
Let $\alpha_k = (\alpha_{1,k}, \dots, \alpha_{n,k}) \in A$ be an arbitrary vector sequence with $(\alpha_{1,k})_{k \in \mathbb{N}}, \dots, (\alpha_{n,k})_{k \in \mathbb{N}}$ being n arbitrary sequences in $\mathbb{R}$ such that: $\lim\limits_{k \to \infty} \alpha_{i,k} = \alpha_i$ for all $i \in \{1, \dots, n\}$.

Since $\alpha_k \in A$ then obtain:

$$\alpha_{i,k} \geq 0 \; \forall i \in I(\bar{x}) \text{ and } \sum_{i=1}^n \alpha_{i,k} = 0$$

By the properties of real-valued limits obtain: $\lim\limits_{k \to \infty} \alpha_{i,k} \geq \lim\limits_{k \to \infty} 0 \Longrightarrow \alpha_i \geq 0$ for all $i \in I(\bar{x})$.
Moreover, apply the limit to the sum:

$$\lim_{k \to \infty} \sum_{i=1}^n \alpha_{i,k} = \sum_{i=1}^n \lim_{k \to \infty} \alpha_{i,k} = \sum_{i=1}^n \alpha_i = \lim_{k \to \infty} 0 = 0$$

Realize that the vector $\alpha = \lim\limits_{k \to \infty} \alpha_k = (\alpha_1, \dots, \alpha_n) \in A$ since:

$$\alpha_i \geq 0 \; \forall i \in I(\bar{x}) \text{ and } \sum_{i=1}^n \alpha_i = 0$$

Therefore A is closed.

***Conclusion*:** $T_{\Lambda_n}(\bar{x}) = A = \big\{(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n : \alpha_i \geq 0 \; \forall i \in I(\bar{x}) \text{ and } \sum_{i=1}^n \alpha_i = 0\big\}$.
As desired.

ii. $N_{\Lambda_n}(\bar{x}) = \{(\beta_0, \ldots, \beta_0) : \beta_0 \in \mathbb{R}\} + \{(\beta_1, \ldots, \beta_n) : \beta_i \leq 0 \; \forall i \in I(\bar{x}) \text{ and } 0 \text{ otherwise}\}$.

Since Λ_n is closed and convex, then its normal and tangent cones are closed and convex. Thus $N_{\Lambda_n}(\bar{x}) = \big(T_{\Lambda_n}(\bar{x})\big)^{\circ}$.
Define $S = \{(\beta_0, \ldots, \beta_0) : \beta_0 \in \mathbb{R}\} + \{(\beta_1, \ldots, \beta_n) : \beta_i \leq 0 \; \forall i \in I(\bar{x}) \text{ and } 0 \text{ otherwise}\}$.

Inclusion 1: $N_{\Lambda_n}(\bar{x}) \subset S$

***Case 1*:** $I(\bar{x}) = \emptyset$
Let $\zeta \in N_{\Lambda_n}(\bar{x})$. Since $I(\bar{x}) = \emptyset$ then $\beta_i = 0$ for all i, and thus: $S = \{(\beta_0, \ldots, \beta_0) : \beta_0 \in \mathbb{R}\}$.

Prove by contradiction: Assume $\zeta \notin S$.
Therefore, there exists $J = \big\{j \in \{1, \ldots, n\} : \zeta_j \neq \zeta_i \text{ for all } i \notin J\big\} \neq \emptyset$. This means there is at least one index j for which the component ζ_j is unique among all other ζ_i for all $i \notin J$.

This implies that for all $i \notin J$, the components $\zeta_i = z$ are identical, and that for all $j \in J$, the components $\zeta_j \neq z$.

Since $\zeta \in N_{\Lambda_n}(\bar{x})$ then $\langle \zeta, \alpha \rangle \leq 0$ for all $\alpha \in T_{\Lambda_n}(\bar{x})$. Evaluate the inner product:

$$\sum_{i=1}^{n} \zeta_i \alpha_i = \sum_{j \in J} \zeta_j \alpha_j + \sum_{i \notin J} \zeta_i \alpha_i = \sum_{j \in J} \zeta_j \alpha_j + z \sum_{i \notin J} \alpha_i = \sum_{j \in J} \zeta_j \alpha_j + z \left(0 - \sum_{j \in J} \alpha_j\right) = \sum_{j \in J} (\zeta_j - z) \alpha_j \leq 0$$

Choose a particular α such that $\operatorname{sgn}(\alpha_j) = \operatorname{sgn}(\zeta_j - z)$, $\forall j \in J$ and that at least one $\alpha_j \neq 0$. This gives $\langle \zeta, \alpha \rangle = \sum_{j \in J} (\zeta_j - z) \alpha_j > 0$. ($\operatorname{sgn}(\cdot)$ denotes the sign)
Contradiction.

***Case 2*: $I(\bar{x}) \neq \emptyset$**
For ease of notation let $I = I(\bar{x})$ and $J = I(\bar{x})^{\mathrm{C}}$.
Let $\zeta \in N_{\Lambda_n}(\bar{x})$. Since $I \neq \emptyset$ then for any $s \in S$ its component: $s_i = \begin{cases} \beta_0 & \forall i \in J \\ \beta_0 + \beta_i & \forall i \in I \end{cases}$

Prove by contradiction: Assume $\zeta \notin S$.
Therefore, there exists $J_0 \subset J : \zeta_{j_0} \neq \beta_0,\ \forall j_0 \in J_0 \neq \emptyset$.
Likewise, there exists $I_0 \subset I : \zeta_{i_0} \neq \beta_0 + \beta_i,\ \forall i_0 \in I_0 \neq \emptyset$.

This simply means that: $\begin{cases} \zeta_i = \beta_0 & \forall i \in J - J_0 \\ \zeta_i = \beta_0 + \beta_i & \forall i \in I - I_0 \\ \zeta_i \neq \beta_0 & \forall i \in J_0 \\ \zeta_i \neq \beta_0 + \beta_i & \forall i \in I_0 \end{cases}$

Since $\zeta \in N_{\Lambda_n}(\bar{x})$ then $\langle \zeta, \alpha \rangle \leq 0$ for all $\alpha \in T_{\Lambda_n}(\bar{x})$. Evaluate the inner product:

$$\begin{aligned} \sum_{i=1}^{n} \zeta_i \alpha_i &= \sum_{i_0 \in I_0} \zeta_{i_0} \alpha_{i_0} + \sum_{j_0 \in J_0} \zeta_{j_0} \alpha_{j_0} + \sum_{i \in I - I_0} \zeta_i \alpha_i + \sum_{j \in J - J_0} \zeta_j \alpha_j \\ &= \sum_{i_0 \in I_0} \zeta_{i_0} \alpha_{i_0} + \sum_{j_0 \in J_0} \zeta_{j_0} \alpha_{j_0} + \sum_{i \in I - I_0} \beta_i \alpha_i + \beta_0 \left(\sum_{i \in I - I_0} \alpha_i + \sum_{j \in J - J_0} \alpha_j \right) \\ &= \sum_{i_0 \in I_0} \zeta_{i_0} \alpha_{i_0} + \sum_{j_0 \in J_0} \zeta_{j_0} \alpha_{j_0} + \sum_{i \in I - I_0} \beta_i \alpha_i + \beta_0 \left(0 - \sum_{i \in I_0} \alpha_{i_0} - \sum_{j \in J_0} \alpha_{j_0} \right) \\ &= \sum_{i_0 \in I_0} (\zeta_{i_0} - \beta_0) \alpha_{i_0} + \sum_{j_0 \in J_0} (\zeta_{j_0} - \beta_0) \alpha_{j_0} + \sum_{i \in I - I_0} \beta_i \alpha_i \leq 0 \end{aligned}$$

Before choosing an instance of α realize that $(\zeta_{i_0} - \beta_0), (\zeta_{j_0} - \beta_0), \alpha_{j_0} \in \mathbb{R}$ and that $\alpha_{i_0}, \alpha_i \geq 0$. (Refer to sums to identify which indices belong to which set)

Choose a particular α such that $\alpha_i = 0,\ \forall i \in I$, and that $\operatorname{sgn}(\alpha_{j_0}) = \operatorname{sgn}(\zeta_{j_0} - z),\ \forall j_0 \in J_0$ and that at least one $\alpha_{j_0} \neq 0$. This gives $\langle \zeta, \alpha \rangle = \sum_{j_0 \in J_0} (\zeta_j - z) \alpha_{j_0} > 0$.
Contradiction.

Inclusion 2: $S \subset N_{\Lambda_n}(\bar{x})$
Let $s \in S$ be arbitrary. As such $s = (\beta_0 + \beta_1, \dots, \beta_0 + \beta_n)$. For all $\alpha \in T_{\Lambda_n}(\bar{x})$, obtain:

$$\langle s, \alpha \rangle = \sum_{i=1}^{n} (\beta_0 + \beta_i) \alpha_i = \beta_0 \sum_{i=1}^{n} \alpha_i + \sum_{i \in I(\bar{x})} \beta_i \alpha_i + \sum_{i \notin I(\bar{x})} \beta_i \alpha_i = \sum_{i \in I(\bar{x})} \beta_i \alpha_i \leq 0,$$

since $\beta_i \leq 0$ and $\alpha_i \geq 0$ for all $i \in I(\bar{x})$ and $\beta_i = 0$ for all $i \notin I(\bar{x})$ and $\sum_{i=1}^{n} \alpha_i = 0$.

Observe that $\langle s, a \rangle \leq 0$ for all $\alpha \in T_{\Lambda_n}(\bar{x})$ and conclude that $s \in N_{\Lambda_n}(\bar{x})$, as desired.

Part (d): Now let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be convex and differentiable. Prove that $\bar{x} \in \Lambda_n$ is a global minimum of f over Λ_n iff there is a constant $\bar{c}$ such that

$$\begin{cases} f_{x_i}(\bar{x}) = \bar{c} & \forall i \notin I(\bar{x}) \\ f_{x_i}(\bar{x}) \geq \bar{c} & \forall i \in I(\bar{x}) \end{cases}$$

Direction 1 ($\Longrightarrow$): Assume $\bar{x} \in \Lambda_n$ is a global minimum of f over Λ_n.
Since f is convex and differentiable, and that Λ_n is closed and convex, then:
Having $0 \in \nabla f(\bar{x}) + N_{\Lambda_n}(\bar{x})$, then there exists a vector $s \in N_{\Lambda_n}(\bar{x})$ such that $\nabla f(\bar{x}) = -s$.

$$f_{x_i}(\bar{x}) = -s_i = \begin{cases} -\beta_0 & \forall i \notin I(\bar{x}) \\ -\beta_0 - \beta_i & \forall i \in I(\bar{x}) \end{cases}$$

It is equivalent to state the existence of the constant $\bar{c} = -\beta_0$.
Knowing that $\beta_i \leq 0$ for all $i \in I(\bar{x})$ then $-\beta_0 - \beta_i = \bar{c} - \beta_i \geq \bar{c}$.

The desired result is thus obtained.

Direction 2 ($\Longleftarrow$): There exists a constant $\bar{c}$ such that $f_{x_i}(\bar{x})$ is given as above.
The function f is convex and differentiable over $\mathbb{R}$, and $\bar{x} \in \Lambda_n$.
Evaluate the inner product $\langle x - \bar{x}, \nabla f(\bar{x}) \rangle$, $\forall x \in \Lambda_n$.

Realize that $f_{x_i} = \bar{c}$ or $f_{x_i} \geq \bar{c}$, therefore $f_{x_i} \geq \bar{c}$.
Thus, $\langle x - \bar{x}, \nabla f(\bar{x}) \rangle = \sum_{i=1}^{n} (x_i - \bar{x}_i) f_{x_i}(\bar{x}) \geq \bar{x} \sum_{i=1}^{n} (x_i - \bar{x}_i) = \bar{c}(1-1) = 0$, $\forall x \in \Lambda_n$.

Conclude that $\bar{x}$ is a local minimum over Λ_n since $\langle x - \bar{x}, \nabla f(\bar{x}) \rangle \geq 0$, $\forall x \in \Lambda_n$.
Since Λ_n is closed and also convex, then obtain the desired result.

Part (e): Minimize $f(x,y,z) = x^2 + 2y^2 + z^2$ and $g(x,y,z) = x + 2y + z$ over Λ_3.
Here f and g are both convex and clearly differentiable.

The Hessian $\nabla^2 f = \text{diag}(2,4,2)$ is positive-definite since all eigenvalues are strictly positive, thus f is strictly convex. The function g is convex because it is linear.

Obtain from previous results that $\Lambda_3 = \{(x,y,z) \in \mathbb{R}^3 : x+y+z = 1 \text{ and } x,y,z \geq 0\}$.
Also proven earlier was that Λ_3 is closed and convex.

Let $(\bar{x},\bar{y},\bar{z})$ be the global minimum of f over Λ_3 and (x^*,y^*,z^*) be that of g.
Compute the gradients $\nabla f = (2x, 4y, 2z)$ and $\nabla g = (1,2,1)$.

Minimize f:
Assume that the optimal point is not at the boundaries, thus $I(\bar{x},\bar{y},\bar{z}) = \emptyset$.
From Part (d) obtain that $\nabla f(\bar{x},\bar{y},\bar{z}) = (\bar{c},\bar{c},\bar{c}) = (2\bar{x}, 4\bar{y}, 2\bar{z})$.

The point $(\bar{x},\bar{y},\bar{z}) = \frac{1}{4}(2\bar{c},\bar{c},2\bar{c})$. And since $\bar{x} \in \Lambda_3$, then: $\bar{x}+\bar{y}+\bar{z} = \frac{5}{4}\bar{c} = 1 \iff \bar{c} = \frac{4}{5}$.

Check the boundaries; at the vertices $f(1,0,0) = f(0,0,1) = 1$ and $f(0,1,0) = 2$.
Edge 1: $f(\lambda, 0, 1-\lambda) = 2\lambda^2 - 2\lambda + 1$. Set $f'(\lambda) = 0 \Longrightarrow \lambda = \frac{1}{2} \Longrightarrow f = \frac{1}{2}$.
Edge 2: $f(0, \lambda, 1-\lambda) = 3\lambda^2 - 2\lambda + 1$. Set $f'(\lambda) = 0 \Longrightarrow \lambda = \frac{1}{3} \Longrightarrow f = \frac{2}{3}$.
Edge 3: $f(\lambda, 1-\lambda, 0) = 3\lambda^2 - 4\lambda + 2$. Set $f'(\lambda) = 0 \Longrightarrow \lambda = \frac{2}{3} \Longrightarrow f = \frac{2}{3}$.
Here $\lambda \in]0,1[$. These are to be compared against the assumed optimal point.

The optimal point is thus $(\bar{x},\bar{y},\bar{z}) = (\frac{2}{5},\frac{1}{5},\frac{2}{5})$ and the optimum is $f^* = \frac{2}{5}$.

Minimize g:
Since g is linear then its minimal point $(x^*,y^*,z^*) \in \text{bdry}(\Lambda_3)$. This is further demonstrated by the fact that $\nabla g(x^*,y^*,z^*) = (\bar{c},\bar{c},\bar{c}) = (1,2,1)$ is a contradiction.

Check the boundaries; at the vertices $g(1,0,0) = g(0,0,1) = 1$ and $g(0,1,0) = 2$.
Edge 1: $g(\lambda, 0, 1-\lambda) = \lambda + 1 - \lambda = 1$.
Edge 2: $g(0, \lambda, 1-\lambda) = \lambda + 1 \geq 1$.
Edge 3: $g(\lambda, 1-\lambda, 0) = 2 - \lambda \geq 1$.
Here $\lambda \in]0,1[$.

Conclude that $(x^*,y^*,z^*) = (\lambda, 0, 1-\lambda)$ for all $\lambda \in [0,1]$ are the global minimal points and that $g^* = 1$ is the global minimum.