JOURNAL OF INDUSTRIAL AND
MANAGEMENT OPTIMIZATION
Volume **8**, Number **1**, February **2012**

doi:10.3934/jimo.2012.8.189

pp. **189–227**

CONVEX OPTIMIZATION ON MIXED DOMAINS

Murat Adivar

Izmir University of Economics, Department of Mathematics
35330, Balcova, Izmir, Turkey

Shu-Cherng Fang

Department of Industrial and Systems Engineering
North Carolina State University
Raleigh, NC 27695, USA

(Communicated by Kok Lay Teo)

Abstract. This paper aims to study convex analysis on some "generalized domains," in particular, the domain of the product of closed subsets of reals. We introduce the basic concepts and derive analytic properties regarding convex subsets of mixed domains and convex functions defined on convex sets in mixed domains. The results obtained may open an avenue for modeling and solving a new type of optimization problems that involve both discrete and continuous variables at the same time.

1. Introduction and preliminaries.

Discrete and continuous analyses and optimizations are closely related, yet they are usually treated separately. Stefan Hilger [15] reasoned that there must be an underlying mathematical structure to explain when and why the theories behind the two settings coalesce or differ, and, hence, introduced the theory of time scales in order to unify the seemingly disparate fields of discrete and continuous analyses. The introduction of "Hilger derivative" (delta derivative) on time scales further enables mathematicians to combine differential and difference equations within the framework of dynamic equations on time scales (see for instance [5], [8], [9], [21], [22], and references therein). Many topics of modern mathematics such as the oscillation theory, control theory, and stability theory have been restudied in the new context. Our reference list is by no means complete, but we would like to particularly mention that the oscillation theory on time scales has been studied in [10] and [11]; the control theory and variational problems on time scales can be found in [14], [16], and [17]; some applications of dynamic equations on time scales to economics are documented in [6] and [23]; the stability theory for delay dynamic equations on time scales has been treated in [2], [4], and [5]; the integral equations on time scales are studied in [1], [3] and [18]; and the Fell topology for dynamic equations on time scales is determined in [24]. In our view, the time scale systems might best be employed in engineering applications where both of the discrete time and continuous time systems are used. A simultaneous

2000 *Mathematics Subject Classification.* Primary 90C25, 80M50; Secondary 65K05, 90C30.

Key words and phrases. Continuous optimization, convexity, convex optimization, discrete optimization, generalized convexity, subgradient, time scales.

This work was supported by The Scientific and Technological Research Council of Turkey and The US National Science Foundation Grant #DMI-0553310.

presentation of the two theories under the umbrella of time scales might provide a new perspective and easiness for modeling and solving optimization problems on a general domain.

The notion of convexity for functions of one variable on a time scale has been introduced in [13]. However, to the best of our knowledge, the notions of convexity for functions of several variables and convexity for the subsets in a product space of different time scales have not been seriously investigated. In this paper, we intend to formally study convexity related notions such as convex combination, hyperplane, supporting hyperplane, convex hull, subgradient, epigraph and hypograph for sets and functions in products of time scales. We also intend to introduce a framework of convex optimization on time scales for modeling and solving problems with both discrete and continuous variables. Hopefully, our work is not only a generalization of existing theory but also an initial step for the development of new models and algorithms.

We now introduce some basic definitions and examples in time scale calculus. They can be found in [8] and [9] in which a comprehensive review on time scales is given.

It is important to mention that, throughout this work, we assume that a time scale, denoted $\mathbb{T}$, has the topology inherited from the standard topology on the real numbers $\mathbb{R}$. Moreover, we denote the set of integer numbers by $\mathbb{Z}$, natural numbers by $\mathbb{N}$ and positive natural numbers by $\mathbb{N}_0$.

Definition 1 (Time scale $\mathbb{T}$)**.** An arbitrary nonempty closed subset $\mathbb{T}$ of real numbers is called a time scale.

Definition 2 (Operators on time scales)**.** Let $\mathbb{T}$ be a time scale. The forward jump operator $\sigma : \mathbb{T} \to \mathbb{T}$ and backward jump operator $\rho : \mathbb{T} \to \mathbb{T}$ are defined by

$$\sigma(t) = \inf \{s \in \mathbb{T} : s > t\},$$

and

$$\rho(t) = \sup \{s \in \mathbb{T} : s < t\},$$

respectively. Moreover, the forward step-size function $\mu : \mathbb{T} \to [0, \infty)$ and backward step-size function $\nu : \mathbb{T} \to [0, \infty)$ are defined by

$$\mu(t) = \sigma(t) - t,$$

and

$$\nu(t) = t - \rho(t),$$

respectively.

In the following table, we illustrate the operators σ, ρ, μ, and ν on some particular time scales:

$\mathbb{T}$	$\sigma(t)$	$\rho(t)$	$\mu(t)$	$\nu(t)$
$\mathbb{R}$	t	t	0	0
$\mathbb{Z}$	$t+1$	$t-1$	1	1
$h\mathbb{Z}$	$t+h$	$t-h$	h	h
$q^{\mathbb{N}}$	qt	t/q	$(q-1)t$	$(q-1)\frac{t}{q}$
$2^{\mathbb{N}}$	$2t$	$t/2$	t	$t/2$
$\mathbb{N}_0^2$	$(\sqrt{t}+1)^2$	$(\sqrt{t}-1)^2$	$2\sqrt{t}+1$	$2\sqrt{t}-1$

Table 1

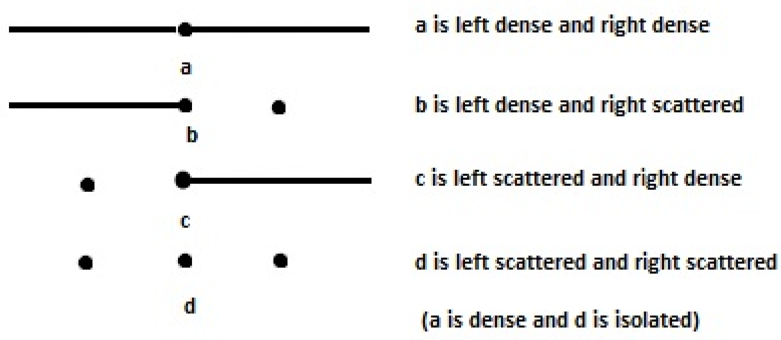

FIGURE 1. Points of time scales

The points of a time scale, $t \in \mathbb{T}$, can be classified in the following manner:

t is right scattered	$if\ t < \sigma(t)$
t is right dense	$if\ t = \sigma(t)$
t is left scattered	$if\ \rho(t) < t$
t is left dense	$if\ \rho(t) = t$
t is isolated	$if\ \rho(t) < t < \sigma(t)$
t is dense	$if\ \rho(t) = t = \sigma(t)$

Table 2

Figure 1 depicts classification of the points of some time scales.

Definition 3 (Time scale interval)**.** Let $\mathbb{T}$ be a time scale, for any $a, b \in \mathbb{R}$ and $a < b$, the time scale interval $[a, b]_{\mathbb{T}}$ is defined by

$$[a, b]_{\mathbb{T}} = [a, b] \cap \mathbb{T}.$$

The intervals $[a, b)_{\mathbb{T}}$, $(a, b]_{\mathbb{T}}$ and $(a, b)_{\mathbb{T}}$ are defined similarly.

The sets $\mathbb{T}^{\kappa}$ and $\mathbb{T}_{\kappa}$ are derived from the time scale $\mathbb{T}$ as follows: If $\mathbb{T}$ has a left-scattered maximum M, then $\mathbb{T}^{\kappa} = \mathbb{T} - \{M\}$. Otherwise $\mathbb{T}^{\kappa} = \mathbb{T}$. If $\mathbb{T}$ has a right-scattered minimum m, then $\mathbb{T}_{\kappa} = \mathbb{T} - \{m\}$. Otherwise $\mathbb{T}_{\kappa} = \mathbb{T}$.

Definition 4 (Delta and nabla derivatives)**.** Let $\mathbb{T}$ be a time scale. The delta derivative f^{Δ} of a function $f : \mathbb{T} \to \mathbb{R}$ is defined at a point $t \in \mathbb{T}^{\kappa}$ by

$$f^{\Delta}(t) := \lim_{s \to t} \frac{f(\sigma(t)) - f(s)}{\sigma(t) - s}, \quad \text{where } s \to t,\ \ s \in \mathbb{T} \backslash \{\sigma(t)\}.$$

The nabla derivative f^{∇} of a function $f : \mathbb{T} \to \mathbb{R}$ is defined at a point $t \in \mathbb{T}_{\kappa}$ by

$$f^{\nabla}(t) := \lim_{s \to t} \frac{f(\rho(t)) - f(s)}{\rho(t) - s}, \quad \text{where } s \to t,\ \ s \in \mathbb{T} \backslash \{\rho(t)\}.$$

Note that the delta and nabla derivatives of a function defined on time scales were first introduced by Hilger [15] with the intention to unify discrete and continuous analyses.

Example 1. *Let* $\overline{q^{\mathbb{Z}}}$ *be the time scale* $\left\{q^k : k \in \mathbb{Z}\right\} \cup \{0\}$ *with* $q > 1$. *The following table shows the derivatives of* f *on some time scales:*

$\mathbb{T}$	$\rho(t)$	$\sigma(t)$	$f^{\nabla}(t)$	$f^{\Delta}(t)$
$\mathbb{R}$	t	t	$f'(t)$	$f'(t)$
$\mathbb{Z}$	$t-1$	$t+1$	$\Delta_{-}f(t)$	$\Delta_{+}f(t)$
$\overline{q^{\mathbb{Z}}}$	t/q	qt	$D_q^{-}f(t)$	$D_q^{+}f(t)$

where $\Delta_{-}f(t) = f(t) - f(t-1)$, $\Delta_{+}f(t) = f(t+1) - f(t)$,

$$D_q^{+}f(t) := \frac{f(qt) - f(t)}{(q-1)t} \tag{1.1}$$

and

$$D_q^{-}f(t) := \frac{f(t) - f(t/q)}{(q-1)t/q}.$$

Based on the above definitions, in the rest of this paper, we study the notion of convexity for the sets in the products of time scales and for the functions defined on the product of time scales. The conventional concepts of convex combination, hyperplane, supporting hyperplane, convex hull, subgradient, epigraph, hypograph and etc. are generalized in later sections. We also extend our findings to treat some convex optimization problems on mixed domains.

2. Convexity on mixed domains.

In this section, we introduce the convexity notion for the sets in products of closed subsets of reals (time scales).

2.1. Right and left convex combinations on $\mathbb{T}$.

Let us start with the convex combination of elements in a time scale.

Definition 5. Let $\mathbb{T}$ be a time scale, $a, b \in \mathbb{T}$ and $\lambda \in [0,1]$. The right and left convex combinations of the elements a and b are defined by

$$K^{+}(a,b;\lambda) := \max\{s \in \mathbb{T} : s \leq a + \lambda(b-a)\}$$

and

$$K^{-}(a,b;\lambda) := \min\{s \in \mathbb{T} : s \geq a + \lambda(b-a)\},$$

respectively. If $x^1, x^2, ..., x^m$ are m elements in $\mathbb{T}$ and $\lambda_1, \lambda_2, ..., \lambda_m \in [0,1]$ with $\Sigma_{i=1}^{m}\lambda_i = 1$, then the right and left convex combinations of $x^1, x^2, ..., x^m$ are defined by

$$K^{+}(x^1, x^2, ..., x^m; \lambda_1, \lambda_2, ..., \lambda_m) = \max\left\{s \in \mathbb{T} : s \leq \textstyle\sum_{i=1}^{m}\lambda_i x^i\right\}$$

and

$$K^{-}(x^1, x^2, ..., x^m; \lambda_1, \lambda_2, ..., \lambda_m) = \min\left\{s \in \mathbb{T} : s \geq \textstyle\sum_{i=1}^{m}\lambda_i x^i\right\},$$

respectively.

The next result follows directly from the definition.

Corollary 1. *Let* $\mathbb{T}$ *be a time scale and* $a,b \in \mathbb{T}$. *Then, the following are true:*

(i)

$$K^{\pm}(a,b;0) = a \text{ and } K^{\pm}(a,b;1) = b;$$

(ii)

$$K^{\pm}(a,b;\lambda) \in \begin{cases} [a,b]_{\mathbb{T}} & \text{if } a < b \\ [b,a]_{\mathbb{T}} & \text{if } a > b \end{cases} \text{ for any } \lambda \in [0,1];$$

(iii) If $a+\lambda(b-a)\in\mathbb{T}$ for some $\lambda\in[0,1]$, then

$$K^{\pm}(a,b;\lambda)=a+\lambda(b-a);$$

(iv) If $K^{+}(a,b;\lambda)=K^{-}(a,b;\lambda)$ for some $\lambda\in[0,1]$, then $a+\lambda(b-a)\in\mathbb{T}$ and

$$K^{\pm}(a,b;\lambda)=a+\lambda(b-a);$$

(v) If $x^{i}=a\in\mathbb{T}$ for all $i=1,2,...,m$, then

$$K^{\pm}(x^{1},x^{2},...,x^{m};\frac{1}{m},\frac{1}{m},...,\frac{1}{m})=a.$$

Example 2. *Let $\mathbb{T}=\mathbb{Z}$, then*

$$K^{+}(a,b;\lambda)=\lfloor a+\lambda(b-a)\rfloor$$

and

$$K^{-}(a,b;\lambda)=\lceil a+\lambda(b-a)\rceil$$

for $a,b\in\mathbb{Z}$ and $\lambda\in[0,1]$, where $\lfloor . \rfloor$ and $\lceil . \rceil$ indicate the greatest integer and the least integer functions, respectively.

Example 3. *If $\mathbb{T}=\mathbb{R}$, then $K^{\pm}(a,b;\lambda)=a+\lambda(b-a)$ for any $a,b\in\mathbb{R}$ and $\lambda\in[0,1]$.*

2.2. **Right and left convex combinations on $\mathbb{T}_1\times\mathbb{T}_2\times...\times\mathbb{T}_n$.** Hereafter, we use the notation Λ^n to denote the product $\mathbb{T}_1\times\mathbb{T}_2\times...\times\mathbb{T}_n$ of the time scales $\mathbb{T}_i$, $i=1,2,...,n$, and e_j to represent the j-th unit vector whose j-th coordinate is 1 and other coordinates are 0. For any two elements

$$x^{1}:=(x_1^1,x_2^1,...,x_n^1)=\sum_{j=1}^{n}x_j^1e_j$$

and

$$x^{2}:=(x_1^2,x_2^2,...,x_n^2)=\sum_{j=1}^{n}x_j^2e_j$$

of Λ^n and $\lambda\in[0,1]$, the right and left convex combinations $K^{\pm}(x^1,x^2;\lambda)$ of x^1 and x^2 are defined by

$$\begin{aligned}K^{\pm}(x^{1},x^{2};\lambda)&=\sum_{j=1}^{n}K_j^{\pm}(x_j^1,x_j^2;\lambda)e_j\\&=\left(K_1^{\pm}(x_1^1,x_1^2;\lambda),K_2^{\pm}(x_2^1,x_2^2;\lambda),...,K_n^{\pm}(x_n^1,x_n^2;\lambda)\right),\end{aligned}$$

where

$$K_j^{+}(x_j^1,x_j^2;\lambda)=\max\left\{s\in\mathbb{T}_j:s\le x_j^1+\lambda(x_j^2-x_j^1)\right\}$$

and

$$K_j^{-}(x_j^1,x_j^2;\lambda)=\min\left\{s\in\mathbb{T}_j:s\ge x_j^1+\lambda(x_j^2-x_j^1)\right\},$$

are the right and left convex combinations on the time scales $\mathbb{T}_j$, $j=1,2,...,n$, respectively. In general, if

$$x^{i}=\sum_{j=1}^{n}x_j^ie_j,\qquad i=1,2,...,m$$

are elements of Λ^n and $\lambda_1, \lambda_2, ..., \lambda_m \in [0,1]$ are scalars such that $\Sigma_{i=1}^m \lambda_i = 1$, then the right and left convex combinations of $x^1, x^2, ..., x^m$ are defined, respectively, by

$$K^+(x^1, x^2, ..., x^m; \lambda_1, \lambda_2, ..., \lambda_m) := \sum_{j=1}^n K_j^+(x_j^1, x_j^2, ..., x_j^m; \lambda_1, \lambda_2, ..., \lambda_m) e_j$$

and

$$K^-(x^1, x^2, ..., x^m; \lambda_1, \lambda_2, ..., \lambda_m) := \sum_{j=1}^n K_j^-(x_j^1, x_j^2, ..., x_j^m; \lambda_1, \lambda_2, ..., \lambda_m) e_j,$$

where

$$K_j^+(x_j^1, x_j^2, ..., x_j^m; \lambda_1, \lambda_2, ..., \lambda_m) := \max\left\{s \in \mathbb{T} : s \leq \textstyle\sum_{i=1}^m \lambda_i x_j^i\right\}$$

and

$$K_j^-(x_j^1, x_j^2, ..., x_j^m; \lambda_1, \lambda_2, ..., \lambda_m) := \min\left\{s \in \mathbb{T} : s \geq \textstyle\sum_{i=1}^m \lambda_i x_j^i\right\}$$

are right and left convex combinations on $\mathbb{T}_j$, respectively, for $j = 1, 2, ..., n$.

Example 4. *Let $\Lambda^2 = \mathbb{Z} \times \mathbb{Z}$ and $x^1 = (0,0)$, $x^2 = (-1,2)$ and $x^3 = (1,2)$. Then*

$$\begin{aligned} K^+\left(x^1, x^2; \frac{1}{2}\right) &= \left(\left\lfloor -\frac{1}{2} \right\rfloor, \lfloor 1 \rfloor\right) = (-1, 1), \\ K^-\left(x^1, x^2; \frac{1}{2}\right) &= \left(\left\lceil -\frac{1}{2} \right\rceil, \lceil 1 \rceil\right) = (0, 1), \\ K^+\left(x^1, x^3; \frac{1}{2}\right) &= \left(\left\lfloor \frac{1}{2} \right\rfloor, \lfloor 1 \rfloor\right) = (0, 1), \\ K^-\left(x^1, x^3; \frac{1}{2}\right) &= \left(\left\lceil \frac{1}{2} \right\rceil, \lceil 1 \rceil\right) = (1, 1). \end{aligned}$$

Observe that the right convex combination $K^+\left(x^1, x^2; \frac{1}{2}\right)$ of the points x^1 and x^2 does not lie on the line joining the points x^1 and x^2, i.e. here the line $y = -2x$. Similarly, $K^-\left(x^1, x^3; \frac{1}{2}\right)$ does not lie on the line $y = 2x$ joining x^1 and x^3.

From the above given example and the definition of right and left convex combinations in Λ^n, we have the next result.

Remark 1. The right and left convex combinations $K^\pm(x^1, x^2; \lambda)$ of two points x^1 and x^2 in Λ^n don't have to lie on the straight line joining x^1 and x^2. However, if $x^1 + \lambda(x^2 - x^1) \in \Lambda^n$ for a scalar $\lambda \in [0,1]$, then

$$K^\pm(x^1, x^2; \lambda) = x^1 + \lambda(x^2 - x^1).$$

2.3. **Right-convex set, left-convex set, and convex set.** Using the concepts of convex combinations, we now define convex sets in the product of time scales.

Definition 6 (right/left-convex set)**.** Let S be a subset of Λ^n. The set S is said to be right-convex in Λ^n if and only if $K^+(x^1, x^2, ..., x^m; \lambda_1, \lambda_2, ..., \lambda_m) \in S$ for all $x^1, x^2, ..., x^m \in S$ and $\lambda_1, \lambda_2, ..., \lambda_m \in [0,1]$ with $\sum_{i=1}^m \lambda_i = 1$. The set $S \subset \Lambda^n$ is said to be left-convex in Λ^n if and only if $K^-(x^1, x^2, ..., x^m; \lambda_1, \lambda_2, ..., \lambda_m) \in S$ for all $x^1, x^2, ..., x^m \in S$ and $\lambda_1, \lambda_2, ..., \lambda_m \in [0,1]$ with $\sum_{i=1}^m \lambda_i = 1$.

Definition 7 (Convex set)**.** The set $S \subset \Lambda^n$ is said to be convex in Λ^n if and only if $\sum_{i=1}^m \lambda_i x^i \in S$ for all $x^1, x^2, ..., x^m \in S$ and $\lambda_1, \lambda_2, ..., \lambda_m \in [0,1]$ such that $\sum_{i=1}^m \lambda_i = 1$ and $\sum_{i=1}^m \lambda_i x^i \in \Lambda^n$.

This definition can alternatively be stated as below.

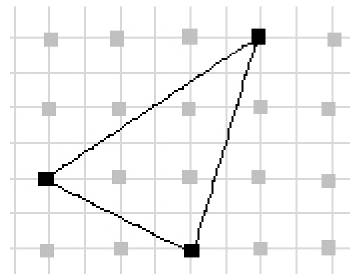

FIGURE 2. Above given three-point set is not convex in $\mathbb{Z}\times\mathbb{Z}$ even though it contains all $\mathbb{Z}\times\mathbb{Z}$ points lying on the lines joining its elements

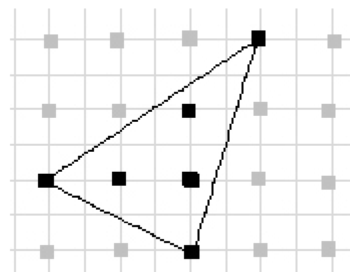

FIGURE 3. The set consisting of black points is a convex set in $\mathbb{Z}\times\mathbb{Z}$

Definition 8. Let S be a subset of Λ^n. Denote by $conv_{\mathbb{R}^n}(S)$ the convex hull of S in $\mathbb{R}^n$. The set S is said to be convex in Λ^n if and only if $a+\lambda(b-a)\in S$ for all $a,b\in conv_{\mathbb{R}^n}(S)\cap\Lambda^n$ and $\lambda\in[0,1]$ such that $a+\lambda(b-a)\in\Lambda^n$.

Hence, we arrive at the following result:

Corollary 2. *Let S be a subset of Λ^n. The set S is convex if and only if $conv_{\mathbb{R}^n}(S)\cap\Lambda^n=S$ (see Figure 2 and Figure 3).*

Lemma 1. *If A is a convex set in $\mathbb{R}^n$, then*

$$conv_{\mathbb{R}^n}(A\cap\Lambda^n)\cap\Lambda^n=A\cap\Lambda^n. \tag{2.1}$$

Proof. Notice that $A\cap\Lambda^n\subset conv_{\mathbb{R}^n}(A\cap\Lambda^n)\cap\Lambda^n$ is always true, since $A\cap\Lambda^n\subset conv_{\mathbb{R}^n}(A\cap\Lambda^n)$. If A is convex in $\mathbb{R}^n$, then

$$A\cap\Lambda^n=conv_{\mathbb{R}^n}(A)\cap\Lambda^n\supset conv_{\mathbb{R}^n}(A\cap\Lambda^n)\cap\Lambda^n.$$

□

This along with the preceding corollary leads to the next result.

Theorem 1. *If A is a convex set in $\mathbb{R}^n$, then $A\cap\Lambda^n$ is convex in Λ^n.*

Corollary 3. *Every right (left)-convex set is convex, but this is not conversely true.*

Proof. Suppose that S is a right convex set and $a,b\in conv_{\mathbb{R}^n}(S)\cap\Lambda^n$. Then S is convex because

$$a+\lambda(b-a)=K^+(a,b;\lambda)\in S$$

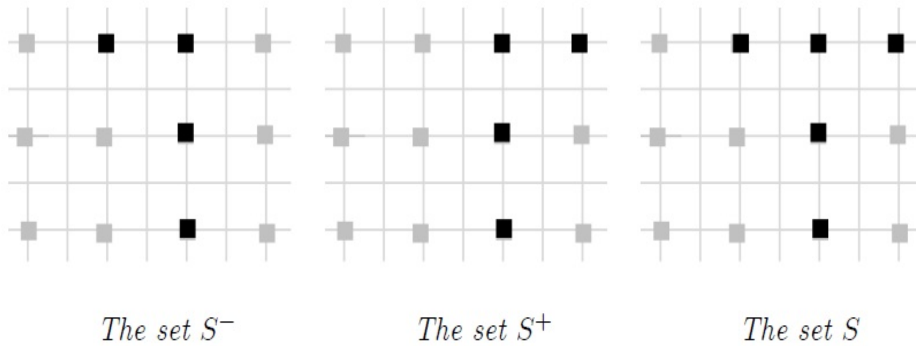

FIGURE 4. The sets S^-, S^+, and S, respectively

for any $\lambda \in [0,1]$ such that $a + \lambda(b-a) \in \Lambda^n$. Similar argument can be applied to a left convex set. Next example shows that convexity may not imply right/left-convexity. □

Example 5. *Let $\Lambda^2 = \mathbb{Z} \times \mathbb{Z}$ and S be the subset of Λ^2 consisting of the points $x^1 = (0,0)$, $x^2 = (-1,2)$, $x^3 = (1,2)$, $x^4 = (0,1)$ and $x^5 = (0,2)$. Define the sets $S^{\pm}$ by*

$$S^+ = \left\{x^1, x^3, x^4, x^5\right\},$$

and

$$S^- = \left\{x^1, x^2, x^4, x^5\right\}.$$

Evidently,

$$S = S^+ \cup S^- = \left\{x^1, x^2, x^3, x^4, x^5\right\}.$$

Then the set S^+ is right-convex but not left-convex, since

$$K^-\left(x^1, x^3; \frac{1}{2}\right) = \left(\left\lceil \frac{1}{2} \right\rceil, \lceil 1 \rceil\right) = (1,1) \notin S^+.$$

The set S^- is left-convex but not right-convex, since

$$K^+\left(x^1, x^2; \frac{1}{2}\right) = \left(\left\lfloor -\frac{1}{2} \right\rfloor, \lfloor 1 \rfloor\right) = (-1,1) \notin S^-.$$

The same arguments show that the set S is neither right-convex nor left-convex. However, the sets $S^{\pm}$ and S are convex, since $conv_{\mathbb{R}^n}(S^{\pm}) \cap \Lambda^n = S^{\pm}$ and $conv_{\mathbb{R}^n}(S) \cap \Lambda^n = S$ (see Figure 4).

2.4. Convex hull.

Definition 9. The convex hull of a set S in Λ^n, denoted by $conv_{\Lambda^n}(S)$, is the collection of all convex combinations of elements of S in Λ^n. In other words, $x \in conv_{\Lambda^n}(S)$ if and only if

$$x = \Sigma_{i=1}^m \lambda_i x^i \in \Lambda^n$$

for some integer $m > 0$, $\lambda_i \in [0,1] \in \mathbb{R}$ such that $\Sigma_{i=1}^m \lambda_i = 1$, and $x^1, x^2, ..., x^m \in S$.

Using Theorem 1 and Definition 9, we have the following results:

Corollary 4. *For any set S in Λ^n,*

$$conv_{\Lambda^n}(S) = conv_{\mathbb{R}^n}(S) \cap \Lambda^n.$$

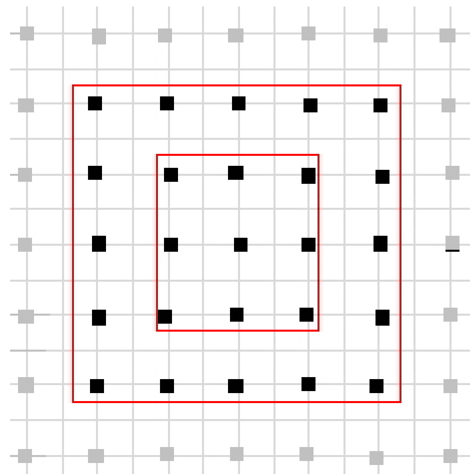

FIGURE 5. The set S consisting of black points surrounded by outer rectangle is a convex set in $\mathbb{Z} \times \mathbb{Z}$. The black points lying between outer and inner rectangles are the points of convex boundary $cbdy_{\Lambda^n}(S)$. The black points surrounded by inner rectangle belong to the convex interior $cint_{\Lambda^n}(S)$ of S

Corollary 5. *Let $S \subset \Lambda^n$, then S is convex in Λ^n if and only if $S = conv_{\Lambda^n}(S)$.*

Lemma 2. *A set S in Λ^n is convex in Λ^n if and only if there is a convex set C in $\mathbb{R}^n$ such that $S = C \cap \Lambda^n$.*

Proof. The necessity part of the proof follows from Theorem 1. For the sufficiency part, if S is assumed to be a convex set in Λ^n, then $S = conv_{\mathbb{R}^n}(S) \cap \Lambda^n$, by Corollary 5. Choosing $C = conv_{\mathbb{R}^n}(S)$ as a convex set, we complete the proof. □

Corollary 6. *$conv_{\Lambda^n}(S)$ is the intersection of all convex sets in Λ^n containing S.*

Corollary 7. *$conv_{\Lambda^n}(S)$ is the smallest convex set in Λ^n containing S.*

2.5. **Convex-closure, convex-interior, convex-boundary.** We now define more details of a convex set in the product of time scales.

Definition 10. Let $S \subset \Lambda^n$ be a convex set in Λ^n. The convex-closure, convex-interior and convex-boundary, denoted by $ccl_{\Lambda^n}(S)$, $cint_{\Lambda^n}(S)$ and $cbdy_{\Lambda^n}(S)$, respectively, are defined by

$$ccl_{\Lambda^n}(S) = \overline{conv_{\mathbb{R}^n}(S)} \cap \Lambda^n,$$

$$cint_{\Lambda^n}(S) = \overset{\circ}{\widehat{conv_{\mathbb{R}^n}}}(S) \cap \Lambda^n,$$

and

$$cbdy_{\Lambda^n}(S) = \partial conv_{\mathbb{R}^n}(S) \cap \Lambda^n,$$

respectively, where $\overline{conv_{\mathbb{R}^n}(S)}$, $\overset{\circ}{\widehat{conv_{\mathbb{R}^n}}}(S)$, and $\partial conv_{\mathbb{R}^n}(S)$ indicate the closure, interior and boundary of $conv_{\mathbb{R}^n}(S)$ in $\mathbb{R}^n$, respectively (see Figure 5).

Remark 2. The convex-interior $cint_{\Lambda^n}(S)$ and convex-closure $ccl_{\Lambda^n}(S)$ of a set S in Λ^n are different from the interior $int_{\Lambda^n}(S)$ and closure $cl_{\Lambda^n}(S)$ of S in Λ^n with the subspace topology inherited from the standard topology on $\mathbb{R}^n$. Moreover,

$$cint_{\Lambda^n}(S) \subset int_{\Lambda^n}(S) \subset S$$

and

$$S \subset cl_{\Lambda^n}(S) \subset ccl_{\Lambda^n}(S).$$

Notice that for a convex set S in $\mathbb{R}^n$ we have

$$ccl_{\mathbb{R}^n}(S) = cl_{\mathbb{R}^n}(S) = \overline{S},$$

$$cint_{\mathbb{R}^n}(S) = int_{\mathbb{R}^n}(S) = \overset{\circ}{S}$$

and

$$cbdy_{\mathbb{R}^n}(S) = \partial S.$$

Corollary 8. *Let $S \subset \Lambda^n$ be a convex set in Λ^n. Then $ccl_{\Lambda^n}(S)$ is closed in Λ^n and $cint_{\Lambda^n}(S)$ is open is Λ^n. Furthermore, $ccl_{\Lambda^n}(S)$ is closed in $\mathbb{R}^n$ (since Λ^n is closed in $\mathbb{R}^n$).*

Corollary 9. *Let $S \subset \Lambda^n$ be a convex set in Λ^n. Then the convex-closure $ccl_{\Lambda^n}(S)$ and the convex-interior $cint_{\Lambda^n}(S)$ are convex sets in Λ^n.*

2.6. **Minimum distance from a point to a convex set.** The concept of minimum distance is introduced in this subsection. Let us start with a well known result:

Theorem 2. *Let S be a nonempty, closed convex set in $\mathbb{R}^n$ and $y \notin S$. Then, there is a unique point $\overline{x} \in S$ with the minimum distance from y. Furthermore, $\overline{x}$ is the minimizing point if and only if*

$$\langle y - \overline{x}, x - \overline{x} \rangle \leq 0 \text{ for all } x \in S. \tag{2.2}$$

However, for convex sets in a time scale, we may not have a unique minimizing point $\overline{x}$ satisfying the inequality (2.2).

Example 6. *Let $\Lambda^2 = \mathbb{Z} \times \mathbb{Z}$ and $S = \{(1,2),(1,3),(1,4),(2,3),(2,4),(3,4)\}$. Choose $Y = (3,1)$. We see that $B = (2,3)$ is the minimizing point in convex set S in Λ^2. However, $B$is not the unique minimizing point and $A = (1,2)$is another minimizing point with $\langle Y - B, A - B \rangle = 1 \geq 0$ (see Figure 6).*

To guarantee the uniqueness of the minimizing point satisfying Ineq. (2.2), we need to impose some conditions on $y \notin S$. The following lemma known as the Weierstrass‘ Theorem which will be used to prove propositions later.

Lemma 3. *Let A be a nonempty and compact set in $\mathbb{R}^n$ and $f : A \to \mathbb{R}$ be continuous on A. Then the problem $\min\{f(x) : x \in A\}$ attains its minimum, that is, there exists a minimizing point to this problem.*

Theorem 3. *Let S be a nonempty closed set in Λ^n and $y \notin S$ be a point in $\mathbb{R}^n$. Then there exists a point $\overline{x} \in S$ with minimum distance from y. Furthermore, if $\overline{x} \in S$ is the minimizing point and $y \notin S$ is a vector satisfying*

$$\langle y - (\lambda x + (1-\lambda)\overline{x}), x - \overline{x} \rangle \neq 0, \text{ for all } \lambda \in (0,1) \text{ and } x \in S - \{\overline{x}\}, \tag{2.3}$$

then the minimizing point $\overline{x} \in S$ is unique and

$$\langle y - \overline{x}, x - \overline{x} \rangle \leq 0 \text{ for all } x \in S. \tag{2.4}$$

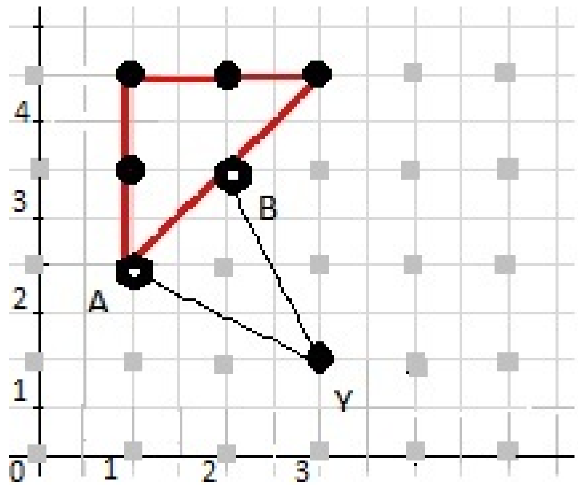

FIGURE 6. A and B are the minimizing points

Conversely, if $y \notin S$ is any vector in $\mathbb{R}^n$ and the inequality (2.4) holds, then $\overline{x}$ is the unique minimizing point.

Proof. First, we show the existence of a minimizing point. Since $S \neq \emptyset$, there exists a point $\widehat{x} \in S$, and we can confine our attention to the set

$$\widehat{S} = S \cap \{x \in S : \|y - x\| \leq \|y - \widehat{x}\|\}$$

in seeking the closest point. In other words, the closest point problem

$$\inf \{\|y - x\| : x \in S\}$$

is equivalent to

$$\inf \left\{\|y - x\| : x \in \widehat{S}\right\}.$$

The compactness of $\widehat{S}$ along with Lemma 3 lead to the existence of the minimizing point $\overline{x} \in S$, which is closest to the point $y \in \mathbb{R}^n - S$.
Suppose that $\overline{x} \in S$ is the minimizing point and $y \notin S$ is a vector satisfying (2.3). Then (2.3) says there is not any $\lambda \in (0, 1)$ such that

$$0 = \langle y - (\lambda x + (1 - \lambda)\overline{x}), x - \overline{x}\rangle = \lambda \langle y - x, x - \overline{x}\rangle + (1 - \lambda) \langle y - \overline{x}, x - \overline{x}\rangle,$$

for all $x \in S - \{\overline{x}\}$. This shows that $\langle y - x, x - \overline{x}\rangle$ and $\langle y - \overline{x}, x - \overline{x}\rangle$ cannot have opposite signs. Since

$$\inf \{\|y - x\| : x \in S\} = \|y - \overline{x}\|$$

and

$$\|y - \overline{x}\|^2 = \|y - x\|^2 + \|x - \overline{x}\|^2 + 2 \langle y - x, x - \overline{x}\rangle,$$

we find

$$\|x - \overline{x}\|^2 + 2 \langle y - x, x - \overline{x}\rangle \leq 0 \text{ for all } x \in S.$$

Therefore,

$$\langle y - x, x - \overline{x}\rangle < 0 \text{ for all } x \in S - \{\overline{x}\}.$$

Consequently,

$$\langle y - \overline{x}, x - \overline{x}\rangle \leq 0 \text{ for all } x \in S - \{\overline{x}\}$$

and (2.4) follows.
Conversely, if $y \in \mathbb{R}^n - S$ is any vector and (2.4) holds, then

$$\|y-x\|^2 = \|y-\overline{x}\|^2 + \|x-\overline{x}\|^2 + 2\langle y-\overline{x}, \overline{x}-x\rangle$$

implies

$$\|y-x\|^2 \geq \|y-\overline{x}\|^2 \text{ for all } x \in S.$$

This shows that $\overline{x}$ is the minimizing point. For the uniqueness of $\overline{x}$, let us assume that there is an $x' \in S$ such that $\|y-x'\| = \|y-\overline{x}\|$. Since

$$\|y-x'\|^2 = \|y-\overline{x}\|^2 + \|x'-\overline{x}\|^2 + 2\langle y-\overline{x}, \overline{x}-x'\rangle,$$

we have

$$\begin{aligned} 0 &= \|\overline{x}-x'\|^2 + 2\langle y-\overline{x}, \overline{x}-x'\rangle \\ &= \langle \overline{x}-x', \overline{x}-x'\rangle + \langle y-\overline{x}, \overline{x}-x'\rangle + \langle y-\overline{x}, \overline{x}-x'\rangle \\ &= \langle y-x', \overline{x}-x'\rangle + \langle y-\overline{x}, \overline{x}-x'\rangle \end{aligned}$$

and hence,

$$\langle y-x', \overline{x}-x'\rangle = -\langle y-\overline{x}, \overline{x}-x'\rangle = \langle y-\overline{x}, x'-\overline{x}\rangle \leq 0.$$

On the other hand, we have

$$\|y-\overline{x}\|^2 = \|y-x'\|^2 + \|x'-\overline{x}\|^2 + 2\langle y-x', x'-\overline{x}\rangle$$

and

$$\|x'-\overline{x}\|^2 = 2\langle y-x', \overline{x}-x'\rangle \leq 0.$$

This is possible only if $x' = \overline{x}$. The proof is complete. □

Different from Theorem 2, in Theorem 3 we ruled out the convexity condition on the set $S \subset \Lambda^n$. Next, we prove that Theorem 2 follows from Theorem 3 as a corollary.

Corollary 10. *If $\Lambda^n = \mathbb{R}^n$, S is a convex set in $\mathbb{R}^n$ and $\overline{x} \in S$ is a point such that*

$$\|y-\overline{x}\| = \min\{\|y-x\| : x \in S\} \text{ for } y \in \mathbb{R}^n - S,$$

then condition (2.3) is automatically satisfied.

Proof. If S is convex in $\mathbb{R}^n$, then $\overline{x}+\lambda(x-\overline{x}) \in S$ for any $x \in S - \{\overline{x}\}$ and $\lambda \in [0,1]$. Since $\overline{x}$ is the minimizing vector we have

$$\|y-\overline{x}+\lambda(x-\overline{x})\|^2 \geq \|y-\overline{x}\|^2. \tag{2.5}$$

Also,

$$\|y-(\overline{x}+\lambda(x-\overline{x}))\|^2 = \|y-\overline{x}\|^2 + \lambda^2\|x-\overline{x}\|^2 - 2\lambda\langle y-\overline{x}, x-\overline{x}\rangle. \tag{2.6}$$

From (2.5) and (2.6), we get

$$2\lambda\langle y-\overline{x}, x-\overline{x}\rangle \leq \lambda^2\|x-\overline{x}\|^2 \text{ for any } 0 \leq \lambda \leq 1. \tag{2.7}$$

Dividing (2.7) by any such $\lambda > 0$ and letting $\lambda \to 0^+$, we have

$$\langle y-\overline{x}, x-\overline{x}\rangle \leq 0. \tag{2.8}$$

Together with

$$\langle y-\overline{x}, x-\overline{x}\rangle = \langle y-x, x-\overline{x}\rangle + \langle x-\overline{x}, x-\overline{x}\rangle, \tag{2.9}$$

we know $\langle y-x, x-\overline{x}\rangle$ cannot be positive. Also, we know

$$\langle y-\overline{x}, x-\overline{x}\rangle^2 + \langle y-x, x-\overline{x}\rangle^2 \neq 0 \text{ for all } x \in S - \{\overline{x}\}. \tag{2.10}$$

Otherwise, $\|x - \overline{x}\| = 0$, i.e. $x = \overline{x}$, which contradicts the assumption of $x \in S - \{\overline{x}\}$. Combining conditions (2.8-2.10), we have

$$\langle y - (\lambda x + (1-\lambda)\overline{x}), x - \overline{x}\rangle = \lambda \langle y - x, x - \overline{x}\rangle + (1-\lambda)\langle y - \overline{x}, x - \overline{x}\rangle < 0$$

for all $\lambda \in (0,1)$ and $x \in S - \{\overline{x}\}$. □

2.7. **Hyperplane, half-space and supporting hyperplane.** We now study the separation properties of convex sets in Λ^n.

Definition 11 (Hyperplane, halfspace)**.** For a given nonzero vector $p \in \mathbb{R}^n$ and a scalar $\alpha \in \mathbb{R}$, when the following set

$$H := \{x \in \Lambda^n : \langle p, x\rangle = \alpha\}$$

is nonempty, we call it a hyperplane in Λ^n. The vector p is called the normal vector of the hyperplane. We also define two half-spaces H^+ and H^- as follows

$$H^+ := \{x \in \Lambda^n : \langle p, x\rangle \geq \alpha\},$$
$$H^- := \{x \in \Lambda^n : \langle p, x\rangle \leq \alpha\}.$$

For a fixed $\overline{x} \in \Lambda^n$, H^+ and H^- are sometimes particularly referred to as:

$$H^+ := \{x \in \Lambda^n : \langle p, x - \overline{x}\rangle \geq \alpha\},$$
$$H^- := \{x \in \Lambda^n : \langle p, x - \overline{x}\rangle \leq \alpha\}.$$

Definition 12 (Supporting hyperplane)**.** Let S be a nonempty convex set in Λ^n and $\overline{x} \in cbdy_{\Lambda^n}(S)$. A hyperplane

$$H := \{x \in \Lambda^n : \langle p, x - \overline{x}\rangle = 0\}$$

is called a supporting hyperplane of S at $\overline{x}$, if either $S \subset H^+$, i.e., $\langle p, x - \overline{x}\rangle \geq 0$ for all $x \in S$, or $S \subset H^-$, i.e., $\langle p, x - \overline{x}\rangle \leq 0$ for all $x \in S$.

Theorem 4. *Let S be a nonempty closed set in Λ^n and $y \in \mathbb{R}^n - S$ an arbitrary vector satisfying (2.3). Then there exists a nonzero vector $p \in \mathbb{R}^n$ and a scalar α such that*

$$\langle p, x\rangle \leq \alpha < \langle p, y\rangle \text{ for all } x \in S.$$

Proof. By Theorem 3, we know the existence of a minimizing vector $\overline{x} \in S$ such that

$$\langle y - \overline{x}, x - \overline{x}\rangle \leq 0 \text{ for all } x \in S.$$

Letting $p = y - \overline{x} \neq 0$ and $\alpha = \langle p, \overline{x}\rangle = \langle y - \overline{x}, \overline{x}\rangle$, we have

$$\langle p, x\rangle = \langle y - \overline{x}, x\rangle \leq \langle y - \overline{x}, \overline{x}\rangle = \alpha.$$

On the other hand,

$$\langle p, y\rangle - \alpha = \langle y - \overline{x}, y\rangle - \langle y - \overline{x}, \overline{x}\rangle = \|y - \overline{x}\| > 0.$$

The proof is complete. □

In the conventional case we have the following result:

Theorem 5. [7, p. 49] *Let A be a nonempty convex set in $\mathbb{R}^n$ and $\overline{x} \in \partial A$. Then, there exists a hyperplane that supports A at $\overline{x}$, i.e., there exists a nonzero vector $p \in \mathbb{R}^n$ such that $p^T(x - \overline{x}) \leq 0$ for every $x \in cl(A)$.*

Analogously, we obtain the following result for Λ^n:

Theorem 6. *Let S be a nonempty convex set in Λ^n and $\overline{x} \in cbdy_{\Lambda^n}(S)$. Then, there exists a hyperplane that supports S at $\overline{x}$, i.e., there exists a nonzero vector $p \in \mathbb{R}^n$ such that $p^T(x - \overline{x}) \leq 0$ for all $x \in ccl_{\Lambda^n}(S)$.*

Proof. Notice that

$$ccl_{\Lambda^n}(S) = \overline{conv_{\mathbb{R}^n}(S)} \cap \Lambda^n$$

and

$$cbdy_{\Lambda^n}(S) = \partial conv_{\mathbb{R}^n}(S) \cap \Lambda^n.$$

If $\overline{x} \in cbdy_{\Lambda^n}(S)$, then $\overline{x} \in \partial conv_{\mathbb{R}^n}(S) \cap \Lambda^n$. Since the set $A = conv_{\mathbb{R}^n}(S)$ is convex in $\mathbb{R}^n$, we know, by Theorem 5, there exists a hyperplane that supports A at $\overline{x}$. This means there is a nonzero vector p in $\mathbb{R}^n$ such that $p^T(x - \overline{x}) \leq 0$ for any $x \in cl(A) = \overline{conv_{\mathbb{R}^n}(S)}$, especially for those $x \in \overline{conv_{\mathbb{R}^n}(S)} \cap \Lambda^n = ccl_{\Lambda^n}(S)$. □

Again, in the conventional case we have the following result:

Theorem 7. [7, p. 49] [7, Theorem 2.4.8] *Let A_1 and A_2 be nonempty convex sets in $\mathbb{R}^n$ such that $A_1 \cap A_2 = \emptyset$. Then there exists a hyperplane that separates A_1 and A_2, i.e., there exists a nonzero vector p in $\mathbb{R}^n$ such that*

$$\inf \{\langle p, x\rangle : x \in A_1\} \geq \sup \{\langle p, x\rangle : x \in A_2\}.$$

Analogously, we have the following result for Λ^n:

Theorem 8. *Let S_1 and S_2 be nonempty convex sets in Λ^n such that $S_1 \cap S_2 = \emptyset$. Then there exists a hyperplane that separates S_1 and S_2, i.e., there exists a nonzero vector p in $\mathbb{R}^n$ such that*

$$\inf \{\langle p, x\rangle : x \in S_1\} \geq \sup \{\langle p, x\rangle : x \in S_2\}. \tag{2.11}$$

Proof. Since $conv_{\mathbb{R}^n}(S_1)$ and $conv_{\mathbb{R}^n}(S_2)$ are convex in $\mathbb{R}^n$, we have a nonzero vector $p \in \mathbb{R}^n$ such that

$$\inf \{\langle p, x\rangle : x \in conv_{\mathbb{R}^n}(S_1)\} \geq \sup \{\langle p, x\rangle : x \in conv_{\mathbb{R}^n}(S_2)\}, \tag{2.12}$$

by Theorem 7. Using the convexity of sets, we have

$$S_1 = conv_{\mathbb{R}^n}(S_1) \cap \Lambda^n \text{ and } S_2 = conv_{\mathbb{R}^n}(S_2) \cap \Lambda^n.$$

Together with (2.12), we have

$$\begin{aligned}\inf \{\langle p, x\rangle : x \in S_1\} &\geq \inf \{\langle p, x\rangle : x \in conv_{\mathbb{R}^n}(S_1)\} \\ &\geq \sup \{\langle p, x\rangle : x \in conv_{\mathbb{R}^n}(S_2)\} \\ &\geq \sup \{\langle p, x\rangle : x \in S_2\}.\end{aligned}$$

This shows (2.11) and completes the proof. □

3. **Convex functions on Λ^n.**

After defining convex sets in Λ^n, we turn our attention to the convex functions defined on Λ^n.

Definition 13. Let S be a convex set in Λ^n. The function $f : S \to \mathbb{R}$ is said to be convex if and only if

$$f\left(\sum_{i=1}^{m} \lambda_i x_i\right) \leq \sum_{i=1}^{m} \lambda_i f(x_i) \tag{3.1}$$

for all $x_i \in S$, $i = 1, 2, ..., m$, and $\lambda_i \in [0, 1]$, , $i = 1, 2, ..., m$, such that $\sum_{i=1}^{m} \lambda_i = 1$ and $\sum_{i=1}^{m} \lambda_i x_i \in \Lambda^n$.

Observe that, for a convex function $f : S \to \mathbb{R}$ the inequality

$$f(a + \lambda(b - a)) \leq f(a) + \lambda(f(b) - f(a)), \tag{3.2}$$

holds for all $a, b \in S$ and $\lambda \in [0, 1]$ such that $a + \lambda(b - a) \in \Lambda^n$. However, the converse of this statement may not be true. To illustrate this we give the following example:

Example 7. *Let $\Lambda^2 = \mathbb{Z} \times \mathbb{Z}$; $S = \{(-1, 0), (0, -1), (0, 0), (0, 1), (1, 0), (1, 1)\}$. Define the function $f : S \to \mathbb{R}$ as follows: $f(-1, 0) = f(0, -1) = f(1, 0) = f(0, 1) = 1$, $f(0, 0) = 0$, and $f(1, 1) = -3$. Then (3.1) does not hold, e.g. in the case when $\lambda_1 = \lambda_2 = \lambda_3 = 1/3$ we have*

$$-1 + \frac{2}{3} < f\left(\frac{1}{3}(-1, 0) + \frac{1}{3}(0, -1) + \frac{1}{3}(1, 1)\right) = f(0, 0) = 0.$$

However, the inequality (3.2) is satisfied for all $a, b \in S$ and $\lambda \in [0, 1]$ such that $a + \lambda(b - a) \in \mathbb{Z} \times \mathbb{Z}$.

Remark 3. In [13, Definition 3.1] the inequality (3.2) is given as a necessary and sufficient condition for the convexity of the function $f : [a, b]_{\mathbb{T}} \to \mathbb{R}$ on a time scale interval $[a, b]_{\mathbb{T}}$. However, Example 7 shows that the inequality (3.2) is only a necessary condition for the convexity of a function defined on the products of time scales. Moreover, in one dimensional case (see [13]) there is no need for defining convexity of the domain of a convex function $f : [a, b]_{\mathbb{T}} \to \mathbb{R}$ since every time scale interval $[a, b]_{\mathbb{T}} = [a, b] \cap \mathbb{T}$ is convex in $\mathbb{T}$. However, in multidimensional case one has to define the notion of convexity of a set before defining convexity of a function.

The next result follows from the definition of convex function immediately.

Corollary 11. *Let S be a convex set in Λ^n. If the function $f : \mathbb{R}^n \to \mathbb{R}$ is convex on $conv_{\mathbb{R}^n}(S)$ then the restricted function $\widehat{f} := f|_S$ is convex on S.*

3.1. **Epigraph and hypograph.**

The convex functions on Λ^n can also be characterized by the epigraphs and hypographs.

Definition 14. Let S be a nonempty set in Λ^n and $f : S \to \mathbb{R}$ be a function. The epigraph of f, denoted by $epif$, is a subset of $\Lambda^n \times \mathbb{R}$ defined by

$$epif := \{(x, y) : x \in S, \quad y \in \mathbb{R}, \quad y \geq f(x)\}.$$

The hypograph of f, denoted by $hypf$, is a subset of $\Lambda^n \times \mathbb{R}$ defined by

$$hypf := \{(x, y) : x \in S, \quad y \in \mathbb{R}, \quad y \leq f(x)\}.$$

Theorem 9. *Let S be a nonempty convex set in Λ^n and $f : S \to \mathbb{R}$ be a function. Then f is convex on S if and only if $epif$ is convex in $\Lambda^n \times \mathbb{R}$.*

Proof. Let f be a convex function on $S \subset \Lambda^n$ and $(x_i, y_i) \in epif$ for all $i = 1, 2, ..., m$. Then, for each $i = 1, 2, ..., m$ we have

$$y_i \geq f(x_i).$$

Suppose $\lambda_i \in (0, 1)$, $i = 1, 2, ..., m$, are scalars such that $\sum_{i=1}^m \lambda_i = 1$ and $\sum_{i=1}^m \lambda_i x_i \in \Lambda^n$. Then

$$\sum_{i=1}^m \lambda_i y_i \geq \sum_{i=1}^m \lambda_i f(x_i) \geq f\left(\sum_{i=1}^m \lambda_i x_i\right)$$

Therefore,

$$\sum_{i=1}^{m} \lambda_i (x_i, y_i) \in epif.$$

This shows that

$$conv_{\mathbb{R}^{n+1}}(epif) \cap (\Lambda^n \times \mathbb{R}) = epif$$

which means $epif$ is convex.

Conversely, if $epif$ is convex, then for the elements $(x_i, f(x_i)) \in epif$, $i = 1, 2, ..., m$, and for the scalars $\lambda_i \in (0,1)$ $i = 1, 2, ..., m$, such that $\sum_{i=1}^{m} \lambda_i = 1$ and $\sum_{i=1}^{m} \lambda_i x_i \in \Lambda^n$, we have

$$\sum_{i=1}^{m} \lambda_i (x_i, f(x_i)) \in epif.$$

That is,

$$\sum_{i=1}^{m} \lambda_i f(x_i) \geq f\left(\sum_{i=1}^{m} \lambda_i x_i\right).$$

Hence f is convex. □

3.2. **Subgradient.** Once convex functions on Λ^n are defined, we can study the first order information by the subgradients.

Definition 15. Let S be a nonempty convex set in Λ^n and $f : S \to \mathbb{R}$ be a convex function on S. Then $\xi \in \mathbb{R}^n$ is called a subgradient of f at $\overline{x} \in S$ if

$$f(x) \geq f(\overline{x}) + \xi^T (x - \overline{x}) \text{ for all } x \in S.$$

Similarly, let $f : S \to \mathbb{R}$ be a concave function on S (i.e., $-f$ is convex on S). Then $\xi \in \mathbb{R}^n$ is called the subgradient of f at $\overline{x} \in S$ if

$$f(x) \leq f(\overline{x}) + \xi^T (x - \overline{x}) \text{ for all } x \in S.$$

Theorem 10. *Let S be a nonempty convex set in Λ^n and $f : S \to \mathbb{R}$ be a convex function on S. Then, for $\overline{x} \in cint_{\Lambda^n}(S)$, there exists a vector ξ such that the hyperplane*

$$H = \left\{(x, y) : y = f(\overline{x}) + \xi^T (x - \overline{x})\right\}$$

supports $epif$ at $(\overline{x}, f(\overline{x}))$. In particular,

$$f(x) \geq f(\overline{x}) + \xi^T (x - \overline{x}) \textit{ for all } x \in S,$$

i.e., ξ is a subgradient of f at $\overline{x}$.

Proof. Theorem 9 implies that $epif$ is convex. Noting that $(\overline{x}, f(\overline{x}))$ belongs to the boundary of $epif$, by Theorem 6, there exists a nonzero vector $(\xi_0, \eta) \in \mathbb{R}^n \times \mathbb{R}$ such that

$$\xi_0^T (x - \overline{x}) + \eta (y - f(\overline{x}) \leq 0 \text{ for any } (x, y) \in epif. \tag{3.3}$$

Here η must be nonpositive. Otherwise, the above inequality may cause a contradiction by choosing y to be sufficiently large. We now show that $\eta < 0$. If not, i.e., $\eta = 0$, then

$$\xi_0^T (x - \overline{x}) \leq 0 \text{ for all } x \in S. \tag{3.4}$$

Take any $x' \in conv_{\mathbb{R}^n}(S)$, then there exist scalars $\lambda_i \geq 0$, $i = 1, 2, ..., m$, and $x_1, x_2, ..., x_m \in S$ such that $\sum_{i=1}^{m} \lambda_i = 1$ and

$$x' = \sum_{i=1}^{m} \lambda_i x_i.$$

Applying (3.4), we have

$$\xi_0^T(x' - \overline{x}) = \sum_{i=1}^{m} \lambda_i \xi_0^T(x_i - \overline{x}) \leq 0.$$

This means

$$\xi_0^T(x - \overline{x}) \leq 0 \text{ for all } x \in conv_{\mathbb{R}^n}(S). \tag{3.5}$$

Since $\overline{x} \in cint_{\Lambda^n}(S) = \overset{\circ}{\widehat{conv_{\mathbb{R}^n}(S)}} \cap \Lambda^n$, there exists a $\lambda > 0$ such that $\overline{x} + \lambda\xi_0 \in conv_{\mathbb{R}^n}(S)$. That is,

$$\lambda \xi_0^T \xi_0 \leq 0.$$

This leads to $\xi_0 = 0$ and $(\xi_0, \eta) = (0, 0)$ which contradicts the fact that (ξ_0, η) is a nonzero vector. Therefore, $\eta < 0$. Denoting $\xi_0 / |\eta|$ by ξ and dividing the inequality (3.3) by $|\eta|$, we get

$$\xi^T(x - \overline{x}) - y + f(\overline{x}) \leq 0 \text{ for any } (x, y) \in epif. \tag{3.6}$$

In particular, the hyperplane $H = \left\{(x, y) : y = f(\overline{x}) + \xi^T(x - \overline{x})\right\}$ supports $epif$ at $(\overline{x}, f(\overline{x}))$. By letting $y = f(x)$ in (3.6), we get $f(x) \geq f(\overline{x}) + \xi^T(x - \overline{x})$ for all $x \in S$. This completes the proof. □

Theorem 11. *Let S be a nonempty convex set in Λ^n and $f : S \to \mathbb{R}$. If for every point $\overline{x} \in cint_{\Lambda^n}(S)$ there exists a subgradient vector ξ such that*

$$f(x) \geq f(\overline{x}) + \xi^T(x - \overline{x}) \text{ for all } x \in S,$$

then f is convex on $cint_{\Lambda^n}(S)$.

Proof. Let $x_i \in cint_{\Lambda^n}(S)$ for all $i = 1, 2, ..., m$ and $\lambda_i \in (0, 1)$, $i = 1, 2, ..., m$ be scalars such that $\sum_{i=1}^m \lambda_i = 1$ and $\sum_{i=1}^m \lambda_i x_i \in \Lambda^n$. By Corollary 9, we know $cint_{\Lambda^n}(S)$ is convex and

$$\sum_{i=1}^{m} \lambda_i x_i \in cint_{\Lambda^n}(S).$$

By our assumption, there is a subgradient ξ of f at $\sum_{i=1}^m \lambda_i x_i$. In particular, for each $j = 1, 2, ..., m$ we have

$$f(x_j) \geq f(\sum_{i=1}^{m} \lambda_i x_i) + \xi^T(x_1 - \sum_{i=1}^{m} \lambda_i x_i).$$

Since $\sum_{j=1}^m \lambda_j \beta = \beta$ for a $\beta = \sum_{i=1}^m \lambda_i x_i$ we obtain

$$\begin{aligned} \sum_{j=1}^{m} \lambda_j f(x_j) &\geq f(\sum_{i=1}^{m} \lambda_i x_i) + \xi^T(\sum_{j=1}^{m} \lambda_j x_j - \sum_{j=1}^{m} \lambda_j \sum_{i=1}^{m} \lambda_i x_i) \\ &= f(\sum_{i=1}^{m} \lambda_i x_i) + \xi^T(\sum_{j=1}^{m} \lambda_j x_j - \sum_{i=1}^{m} \lambda_i x_i) \\ &= f(\sum_{i=1}^{m} \lambda_i x_i) \end{aligned}$$

and the result follows. □

3.3. **Minima and maxima of a convex function.** For convex functions defined on Λ^n, we now study their minimum or maximum solutions.

Definition 16. For a given function $f : \Lambda^n \to \mathbb{R}$ and a given set $S \subset \Lambda^n$, consider the following problem:

$$\text{minimize } f(x) \text{ subject to } x \in S. \tag{3.7}$$

A point $x \in S$ is called a feasible solution to the problem. If $\overline{x} \in S$ and $f(x) \geq f(\overline{x})$ for all $x \in S$, then $\overline{x}$ is called an optimal solution or a global optimal solution. The collection of optimal solutions are called alternative optimal solutions.

Theorem 12. *Let $f : \Lambda^n \to \mathbb{R}$ be a convex function and S be a nonempty convex set in Λ^n. The point $\overline{x} \in S$ is an optimal solution to the problem (3.7) if and only if f has a subgradient ξ at $\overline{x}$ such that*

$$\xi^T(x - \overline{x}) \geq 0 \textit{ for all } x \in S.$$

Proof. Suppose that $\xi^T(x - \overline{x}) \geq 0$ for all $x \in S$, where ξ is a subgradient of f at $\overline{x}$. By the convexity of f, we have

$$f(x) \geq f(\overline{x}) + \xi^T(x - \overline{x}) \geq f(\overline{x}) \text{ for all } x \in S.$$

Hence $\overline{x}$ is an optimal solution to the problem.
To show the converse, suppose that $\overline{x}$ is an optimal solution to the problem (3.7) and we construct the following two sets in $\Lambda^n \times \mathbb{R}$:

$$D_1 := \{(x - \overline{x}, y) : x \in \Lambda^n \text{ and } y > f(x) - f(\overline{x})\}$$
$$D_2 := \{(x - \overline{x}, y) : x \in S \text{ and } y \leq 0\}.$$

One may easily verify that both of D_1 and D_2 are convex. Moreover, $D_1 \cap D_2 = \emptyset$. Otherwise, there would exist a point (x, y) such that

$$x \in S \quad \text{and} \quad 0 \geq y > f(x) - f(\overline{x}).$$

This contradicts the assumption of $\overline{x}$ being an optimal solution to the problem. By Theorem 8, there is a hyperplane that separates D_1 and D_2, i.e., there exists a nonzero vector (ξ_0, η_0) and a scalar α such that

$$\xi_0^T(x - \overline{x}) + \eta y \leq \alpha, \quad \textit{for } x \in \Lambda^n \text{ and } y > f(x) - f(\overline{x}) \tag{3.8}$$
$$\xi_0^T(x - \overline{x}) + \eta y \geq \alpha, \quad \textit{for } x \in S \text{ and } y \leq 0. \tag{3.9}$$

If we let $x = \overline{x}$ and $y = 0$ in (3.9), it follows that $\alpha \leq 0$. Next, let $x = \overline{x}$ and $y = \varepsilon > 0$ in (3.8). It follows that $\eta\varepsilon \leq \alpha$. Since this is true for every $\varepsilon > 0$, we have $\eta \leq 0$ and $\alpha \geq 0$. Consequently, $\eta \leq 0$ and $\alpha = 0$. If $\eta = 0$, then, from (3.8),

$$\xi_0^T(x - \overline{x}) \leq 0 \text{ for all } x \in \Lambda^n. \tag{3.10}$$

Let $r > 0$ be a sufficiently large real number such that

$$\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}(H_r)}},$$

where the set $H_r := B(\overline{x}, r) \cap \Lambda^n$. Since

$$\xi_0^T(x - \overline{x}) \leq 0 \text{ for all } x \in H_r,$$

we have

$$\xi_0^T(x - \overline{x}) \leq 0 \text{ for all } x \in conv_{\mathbb{R}^n}(H_r). \tag{3.11}$$

On the other hand, there exists a scalar $c > 0$ such that

$$x = \overline{x} + c\xi_0 \in conv_{\mathbb{R}^n}(H_r).$$

Together with (3.11), we have

$$0 \geq \xi_0^T(x - \overline{x}) = c\,\|\xi_0\|^2 .$$

Hence, $\xi_0 = 0$. Since $(\xi_0, \eta) \neq (0, 0)$, we must have $\eta < 0$. Dividing (3.8) and (3.9) by $-\eta$, and denoting $\xi = -\frac{\xi_0}{\eta}$, we get the following inequalities

$$y \geq \xi^T(x - \overline{x}), \quad for\ x \in \Lambda^n \text{ and } y > f(x) - f(\overline{x}), \tag{3.12}$$

$$\xi^T(x - \overline{x}) \geq y, \qquad for\ x \in S \text{ and } y \leq 0. \tag{3.13}$$

By letting $y = 0$ in (3.13), we get $\xi^T(x - \overline{x}) \geq 0$ for all $x \in S$. From (3.12), it is obvious that

$$f(x) - f(\overline{x}) + \varepsilon \geq \xi^T(x - \overline{x}) \text{ for all } x \in \Lambda^n \text{ and } \varepsilon > 0.$$

Taking limit as $\varepsilon \to 0$, we have

$$f(x) \geq f(\overline{x}) + \xi^T(x - \overline{x}) \text{ for all } x \in \Lambda^n.$$

Therefore, ξ is a subgradient of f at $\overline{x}$ with the property that $\xi^T(x - \overline{x}) \geq 0$ for all $x \in S$. This completes the proof. □

Corollary 12. *Under the assumptions of Theorem 12, if $S \subset \Lambda^n$ is convex with*

$$cint_{\Lambda^n}(S) = \overset{\circ}{\widehat{conv_{\mathbb{R}^n}(S)}} \cap \Lambda^n \neq \emptyset, \tag{3.14}$$

then $\overline{x} \in cint_{\Lambda^n}(S)$ is an optimal solution to the problem (3.7) if and only if there exists a zero subgradient of f at $\overline{x}$.

Proof. The previous theorem says that $\overline{x} \in cint_{\Lambda^n}(S) \subset S$ is an optimal solution to the problem (3.7) if and only if f has a subgradient ξ at $\overline{x}$ such that $\xi^T(x - \overline{x}) \geq 0$ for any $x \in S$. Consequently, the linearity of the inner product implies that

$$\xi^T(x - \overline{x}) \geq 0 \text{ for all } x \in conv_{\mathbb{R}^n}(S). \tag{3.15}$$

Since $\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}(S)}}$, there exists a positive real c such that $x = \overline{x} - c\xi \in conv_{\mathbb{R}^n}(S)$. By (3.15), we have $-c\,\|\xi\|^2 \geq 0$. Hence we have $\xi = 0$. □

4. **Differentiable convex functions.**

In this section, we study the differentiability of convex functions on time scales and use it for solving optimization problems.

4.1. **Uniqueness of the subgradient.**

First, we recall the following uniqueness result in the conventional continuous calculus:

Lemma 4. [7, Lemma 3.3.2] *Let A be a nonempty set in $\mathbb{R}^n$ and $f : A \to \mathbb{R}$ be convex. If f is differentiable at $\overline{x} \in \overset{\circ}{A}$, then the set of subgradients of f at $\overline{x}$ is the singleton*

$$\{\text{grad}\, f(\overline{x})\} := \left\{ \left(\frac{\partial f(\overline{x})}{\partial_1 x_1}, ..., \frac{\partial f(\overline{x})}{\partial_n x_n} \right) \right\},$$

i.e., grad $f(\overline{x})$ *is the only vector such that*

$$f(x) \geq f(\overline{x}) + (\text{grad}\, f(\overline{x}))^T(x - \overline{x}),\ for\ all\ x \in A. \tag{4.1}$$

In order to show the above result may not be true in time scales, we first prove the next lemma.

Lemma 5. *Let* $f : [a,b]_{\mathbb{T}} \to \mathbb{R}$ *be a convex function. Then*

$$\frac{f(t)-f(s)}{t-s} \leq \frac{f(u)-f(s)}{u-s} \leq \frac{f(y)-f(s)}{y-s}, \tag{4.2}$$

for any $t, u, s, y \in [a,b]_{\mathbb{T}}$ *satisfying* $t \leq u < s < y$. *If* f *is* $\Delta-$*differentiable and* $\nabla-$*differentiable at* $\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}[a,b]_{\mathbb{T}}}} \cap \mathbb{T}$, *then*

$$f^{\nabla}(\overline{x}) \leq f^{\Delta}(\overline{x}), \tag{4.3}$$

where $f^{\Delta}(\overline{x})$ *and* $f^{\nabla}(\overline{x})$ *are the delta and nabla derivatives of* f, *respectively.*

Proof. Ineq. (4.2) follows directly from the definition of convexity. The equality holds in (4.3), if $\overline{x}$ is right and left dense. In other cases, we can get (4.3) from (4.2) by making the following substitutions:
(i) $t = \rho(\overline{x})$, $s = \overline{x}$, $y = \sigma(\overline{x})$, if $\overline{x}$ is right and left scattered;
(ii) $t = \rho(\overline{x})$, $s = \overline{x}$, and $y = \overline{x} + h$ for $h > 0$ being sufficiently small, if $\overline{x}$ is left scattered and right dense;
(iii) $t = \overline{x} - h$ for $h > 0$ being sufficiently small, $s = \overline{x}$, and $y = \sigma(\overline{x})$, if $\overline{x}$ is right scattered and left dense. □

The next theorem shows that the subgradient may not be unique for convex functions defined on arbitrary time scales.

Theorem 13. *Let* $f : [a,b]_{\mathbb{T}} \to \mathbb{R}$ *be a convex function on a time scale interval* $[a,b]_{\mathbb{T}}$ *with the property that* $\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}[a,b]_{\mathbb{T}}}} \cap \mathbb{T} \neq \varnothing$. *Then the set*

$$\partial f(\overline{x}) := \{\xi \in \mathbb{R} : f(x) \geq f(\overline{x}) + \xi(x - \overline{x}) \text{ for all } x \in [a,b]_{\mathbb{T}}\}$$

contains $f^{\Delta}(\overline{x})$ *and* $f^{\nabla}(\overline{x})$ *for all* $\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}[a,b]_{\mathbb{T}}}} \cap \mathbb{T}$. *Moreover, every function* $\varphi : [a,b]_{\mathbb{T}} \to \mathbb{R}$ *satisfying*

$$f^{\nabla}(\overline{x}) \leq \varphi(\overline{x}) \leq f^{\Delta}(\overline{x}) \tag{4.4}$$

belongs to $\partial f(\overline{x})$ *for any* $\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}[a,b]_{\mathbb{T}}}} \cap \mathbb{T}$, *and vice versa.*

Proof. If $\overline{x} = x$, then the proof is easy. Hence we may assume that $\overline{x} \neq x$. When $x > \overline{x}$, if the point $\overline{x} \in [a,b]_{\mathbb{T}}^{\kappa}$ is right scattered, then we know $x \geq \sigma(\overline{x})$ and

$$\frac{f(x)-f(\overline{x})}{x-\overline{x}} \geq \frac{f(\sigma(\overline{x}))-f(\overline{x})}{\sigma(\overline{x})-\overline{x}} = f^{\Delta}(\overline{x}),$$

by the convexity of f. If the point $\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}[a,b]_{\mathbb{T}}}} \cap \mathbb{T}$ is right dense, then, for sufficiently small $h > 0$, we have $x \geq \overline{x} + h$ and

$$\frac{f(x)-f(\overline{x})}{x-\overline{x}} \geq \frac{f(\overline{x}+h)-f(\overline{x})}{h}.$$

Taking the limit as $h \to 0$ gives the desired result.

When $x < \overline{x}$, if the point $\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}[a,b]_{\mathbb{T}}}} \cap \mathbb{T}$ is right scattered, then

$$\frac{f(\sigma(\overline{x}))-f(x)}{\sigma(\overline{x})-x} \geq \frac{f(x)-f(\overline{x})}{x-\overline{x}}.$$

Using the formula $f(\sigma(\overline{x})) = f(\overline{x}) + (\sigma(\overline{x}) - \overline{x})f^{\Delta}(\overline{x})$, we get

$$f(x) \geq f(\overline{x}) + f^{\Delta}(\overline{x})(x - \overline{x})$$

as desired. If the point $\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}}} [a,b]_{\mathbb{T}} \cap \mathbb{T}$ is right dense then, for sufficiently small $h > 0$, we have $\overline{x} + h > \overline{x} > x$ and

$$\frac{f(\overline{x}+h) - f(x)}{\overline{x}+h-x} \geq \frac{f(x) - f(\overline{x})}{x - \overline{x}}.$$

Simplifying the last inequality leads to

$$\frac{f(\overline{x}+h) - f(\overline{x})}{h} \geq \frac{f(x) - f(\overline{x})}{x - \overline{x}}.$$

Taking limit on both sides as $h \to 0$ and multiplying by $x - \overline{x}$, we have

$$f(x) \geq f(\overline{x}) + f^{\Delta}(\overline{x})(x - \overline{x}). \tag{4.5}$$

One may similarly show that

$$f(x) \geq f(\overline{x}) + f^{\nabla}(\overline{x})(x - \overline{x}), \tag{4.6}$$

for any $\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}}} [a,b]_{\mathbb{T}} \cap \mathbb{T}$ and $x \in [a,b]_{\mathbb{T}}$. This shows that any function satisfying (4.4) belongs to $\partial f(\overline{x})$ for every $\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}}} [a,b]_{\mathbb{T}} \cap \mathbb{T}$.

Now, suppose we have $\varphi \in \partial f(\overline{x})$ for all $\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}}} [a,b]_{\mathbb{T}} \cap \mathbb{T}$. If $\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}}} [a,b]_{\mathbb{T}} \cap \mathbb{T}$ is right or left dense, then either

$$\varphi(\overline{x}) = \lim_{h \to 0^+} \frac{f(\overline{x}+h) - f(\overline{x})}{h}$$

or

$$\varphi(\overline{x}) = \lim_{h \to 0^-} \frac{f(\overline{x}+h) - f(\overline{x})}{h}.$$

Hence, the proof follows from the fact that

$$f^{\nabla}(\overline{x}) \leq \lim_{h \to 0^{\pm}} \frac{f(\overline{x}+h) - f(\overline{x})}{h} \leq f^{\Delta}(\overline{x}).$$

Suppose that the point $\overline{x} \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}}} [a,b]_{\mathbb{T}} \cap \mathbb{T}$ is isolated. Then substituting $x = \sigma(\overline{x})$ in (4.5) leads to

$$f^{\Delta}(\overline{x})(\sigma(\overline{x}) - \overline{x}) = f(\sigma(\overline{x})) - f(\overline{x}) \geq \varphi(\overline{x})(\sigma(\overline{x}) - \overline{x}).$$

Therefore, $\varphi(\overline{x}) \leq f^{\Delta}(\overline{x})$. On the other hand, by substituting $x = \rho(\overline{x})$ into (4.6), we get

$$f^{\nabla}(\overline{x})(\rho(\overline{x}) - \overline{x}) = f(\rho(\overline{x})) - f(\overline{x}) \geq (\rho(\overline{x}) - \overline{x})\varphi(\overline{x}).$$

Dividing both sides by $\rho(\overline{x}) - \overline{x}$, we have $\varphi(\overline{x}) \geq f^{\nabla}(\overline{x})$. □

4.2. **Subgradient of a convex function on Λ^n.** In this subsection, we characterize convex functions in terms of subgradients on the product of time scales. First, note that the following definitions can be found in [21] and [22].

Definition 17. Let $\mathbb{T}_i$, $i \in \{1, 2, ..., n\}$, be time scales and $f : \Lambda^n \to \mathbb{R}$ be a function. The partial Δ-derivative of f with respect to $x_i \in \mathbb{T}_i^{\kappa}$ is defined by

$$\frac{\partial f(x)}{\Delta_i x_i} := \lim_{\substack{s_i \to x_i \\ s_i \neq \sigma(x_i)}} \frac{f(x_1, x_2, ..., \sigma_i(x_i), x_{i+1}, ..., x_n) - f(x_1, x_2, ..., s_i, x_{i+1}, ..., x_n)}{\sigma_i(x_i) - s_i}, \tag{4.7}$$

where $\sigma_i : \mathbb{T}_i \to \mathbb{T}_i$ is the forward jump operator on the $i-th$ time scale $\mathbb{T}_i$. Similarly, the partial ∇-derivative of $f : \Lambda^n \to \mathbb{R}$ is defined by

$$\frac{\partial f(x)}{\nabla_i x_i} := \lim_{\substack{s_i \to x_i \\ s_i \neq \rho(x_i)}} \frac{f(x_1, x_2, ..., \rho_i(x_i), x_{i+1}, ..., x_n) - f(x_1, x_2, ..., s_i, x_{i+1}, ..., x_n)}{\rho_i(x_i) - s_i}, \tag{4.8}$$

where $\rho_i : \mathbb{T}_i \to \mathbb{T}_i$ is the backward jump operator on the $i-th$ time scale $\mathbb{T}_i$.

In preparation for the next result, let's define the set

$$\bigcup_{i=1}^{n} B_i(x, h_i^{\pm})$$

as a frame at the point $x = (x_1, x_2, ..., x_n)$ in Λ^n, where

$$B_i(x, h_i^{\pm}) = \left\{ se^i + \sum_{\substack{j=1 \\ j \neq i}}^{n} x_j e^j : s \in N_i^{\pm}(x, h_i^{\pm}) \right\}, \tag{4.9}$$

e^i, $i = 1, 2, ..., n$, are the unit vectors whose components are determined by

$$e_j^i = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases};$$

the sets $N_i^{\pm}(x, h_i^{\pm})$, $i = 1, 2, ..., n$, given by

$$N_i^{+}(x, h_i^{+}) := \begin{cases} \{x_i, \sigma_i(x_i)\} & \text{if } \mu_i(x_i) > 0 \\ [x_i, x_i + h_i^{+}) & \text{if } \mu_i(x_i) = 0 \end{cases}, \text{ where } h_i^{+} > 0,$$

and

$$N_i^{-}(x, h_i^{-}) := \begin{cases} \{\rho_i(x_i), x_i\} & \text{if } \nu_i(x_i) > 0 \\ (x_i - h_i^{-}, x_i] & \text{if } \nu_i(x_i) = 0 \end{cases} \text{ where } h_i^{-} > 0,$$

are the sets in the time scale $\mathbb{T}_i$.

In order for the partial derivatives $\frac{\partial f(x)}{\Delta_i x_i}$ and $\frac{\partial f(x)}{\nabla_i x_i}$ given by (4.7) and (4.8) to be well defined at a point $x = (x_1, x_2, ..., x_n) \in S$ one has to assume that

$$\bigcup_{i=1}^{n} B_i(x, h_i^{\pm}) \subset S \tag{4.10}$$

to guarantee that the vectors $(x_1, x_2, ..., \sigma_i(x_i), x_{i+1}, ..., x_n)$ and $(x_1, x_2, ..., s_i, x_{i+1}, ..., x_n)$, $s \in N_i^{\pm}(x, h_i^{\pm})$, $i = 1, 2, ..., n$, are in S. Notice that the condition (4.10) also implies $x \in cint_{\Lambda^n}(S)$.

Then we are ready to state our next result.

Theorem 14. *Let S be a nonempty convex set in Λ^n. Let $f : S \to \mathbb{R}$ be a function such that the partial derivatives $\left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}}$ and $\left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}}$, $i = 1, 2, ..., n$, exist at any point $\overline{x} = (\overline{x}_1, \overline{x}_2, ..., \overline{x}_n) \in cint_{\Lambda^n}(S)$ satisfying (4.10).*

1. *If f is convex on S, then there exist scalars $\lambda_i(\overline{x}) \in [0, 1]$, $i = 1, 2, ..., n$, such that the vector*

$$\xi(\overline{x}) := \sum_{i=1}^{n} \left(\lambda_i(\overline{x}) \left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}} + (1 - \lambda_i(\overline{x})) \left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}} \right) e_i \tag{4.11}$$

is a subgradient of f at any point $\overline{x} \in cint_{\Lambda^n}(S)$ satisfying (4.10), i.e.,

$$f(x) \geq f(\overline{x}) + \xi(\overline{x})^T (x - \overline{x}) \text{ for all } x \in S. \tag{4.12}$$

2. Suppose that

$$cint_{\Lambda^n}(S) = \left\{ \overline{x} \in S : \bigcup_{i=1}^{n} B_i(\overline{x}, h_i^{\pm}) \subset S \right\}. \tag{4.13}$$

Then f is convex on $cint_{\Lambda^n}(S)$ provided that for any $\overline{x} \in cint_{\Lambda^n}(S)$ there exist scalars $\lambda_i(\overline{x}) \in [0,1]$, $i = 1, 2, ..., n$, such that the vector $\xi(\overline{x})$ defined by (4.11) is a subgradient of f at $\overline{x}$.

Proof. By Theorem 11, we know that $f : S \to \mathbb{R}$ is convex on $cint_{\Lambda^n}(S)$ if for each point $\overline{x} \in cint_{\Lambda^n}(S)$ there exists a subgradient vector $\xi(\overline{x})$ such that

$$f(x) \geq f(\overline{x}) + \xi(\overline{x})^T (x - \overline{x}) \text{ for all } x \in S. \tag{4.14}$$

Hence the second part of the proof can be seen from (4.13) and the existence of partial derivatives $\left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}}$ and $\left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}}$, $i = 1, 2, ..., n$.

For the first part, we let f be convex. Then by Theorem 10 we know that for any $\overline{x} \in cint_{\Lambda^n}(S)$ satisfying (4.10) there is a subgradient vector

$$\xi(\overline{x}) = (\xi_1(\overline{x}), \xi_2(\overline{x}), ..., \xi_n(\overline{x})) \tag{4.15}$$

of f at $\overline{x}$. Define functions $g_i : \mathbb{T}_i \to \mathbb{R}$ by

$$g_i(s) = f(\overline{x}_1, \overline{x}_2, ..., \overline{x}_{i-1}, s, \overline{x}_{i+1}, ..., \overline{x}_n), \;\; i = 1, 2, ..., n.$$

Obviously, g_i is convex for every i. Substituting

$$x = (\overline{x}_1, \overline{x}_2, ..., \overline{x}_{i-1}, s, \overline{x}_{i+1}, ..., \overline{x}_n)$$

into (4.14), we have

$$g_i(s) \geq g_i(\overline{x}_i) + \xi_i(\overline{x})(s - \overline{x}_i)$$

for all $s \in I_{\overline{x}_i}$, where

$$I_{\overline{x}_i} := \{ s \in \mathbb{T}_i : (\overline{x}_1, \overline{x}_2, ..., \overline{x}_{i-1}, s, \overline{x}_{i+1}, ..., \overline{x}_n) \in S \}.$$

The condition (4.10) implies that

$$\overline{x}_i \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}(I_{\overline{x}_i})}} \cap \mathbb{T}_i.$$

This means $\xi_i(\overline{x})$ is a subgradient for the convex function g_i at $\overline{x}_i \in \overset{\circ}{\widehat{conv_{\mathbb{R}^n}(I_{\overline{x}_i})}} \cap \mathbb{T}_i$. From Theorem 13, we further know the existence of a scalar $\lambda_i(\overline{x}) \in [0,1]$ such that

$$\xi_i(\overline{x}) = \lambda_i(\overline{x}) \left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}} + (1 - \lambda_i(\overline{x})) \left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}}. \tag{4.16}$$

The proof is complete. □

Notice that if the point $\overline{x} = (\overline{x}_1, \overline{x}_2, ..., \overline{x}_n) \in S$ mentioned in Theorem 14 is a point having dense components, i.e., $\sigma_i(\overline{x}_i) = \rho_i(\overline{x}_i) = 0$ for all $i = 1, 2, ..., n$, then

$$\left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}} = \left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}} = \left.\frac{\partial f(x)}{\partial x_i}\right|_{x=\overline{x}}, \tag{4.17}$$

and the inequality (4.12) turns into (4.1). Therefore, Theorem 14 leads to the following result.

Corollary 13. *Let S be a nonempty convex set in Λ^n and $\overline{x} = (\overline{x}_1, \overline{x}_2, ..., \overline{x}_n) \in cint_{\Lambda^n}(S)$ a point satisfying (4.10) and $\sigma_i(\overline{x}_i) = \rho_i(\overline{x}_i) = 0$ for all $i = 1, 2, ..., n$.. Let $f : S \to \mathbb{R}$ be a function such that the partial derivatives $\left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}}$ and $\left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}}$ exist for all $i = 1, 2, ..., n$. If f is convex on S, then*

$$gradf(\overline{x}) := \sum_{i=1}^{n} \left.\frac{\partial f(x)}{\partial_i x_i}\right|_{x=\overline{x}} e_i$$

is a subgradient for f at $\overline{x}$, i.e.,

$$f(x) \geq f(\overline{x}) + gradf(\overline{x})^T (x - \overline{x}) \text{ for all } x \in S.$$

Remark 4. For a convex function defined on a convex subset S of the product of time scales $\mathbb{T}_1$, $\mathbb{T}_2$, ...,$\mathbb{T}_2$, the vectors

$$grad_{\Delta} f\,(\overline{x}) = \sum_{i=1}^{n} \left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}} e_i \tag{4.18}$$

and

$$grad_{\nabla} f\,(\overline{x}) = \sum_{i=1}^{n} \left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}} e_i \tag{4.19}$$

may not be subgradients at a point satisfying (4.10). To see this, one may consider the convex function $f(x_1, x_2) = (x_1 - x_2 - 1/2)^2$ defined on $\mathbb{Z} \times \mathbb{Z}$ with

$$grad_{\Delta} f\,(x_1, x_2)^T = (2x_1 - 2x_2, 2x_2 - 2x_1 + 2)$$

and

$$grad_{\nabla} f\,(x_1, x_2)^T = (2x_1 - 2x_2 - 2, 2x_2 - 2x_1)\,.$$

It is easy to see that neither $grad_{\Delta} f\,(x_1, x_2)$ nor $grad_{\nabla} f\,(x_1, x_2)$ is subgradient for f at the point $(0, 0)$. However, for $\lambda = 1$ and $\widehat{\lambda} = 0$, the vector

$$\begin{aligned} \xi(x_1, x_2) &= [\lambda\,(2x_1 - 2x_2) + (1 - \lambda)\,(2x_2 - 2x_1 + 2)]\, e_1 \\ &\quad + \left[\widehat{\lambda}\,(2x_1 - 2x_2 - 2) + \left(1 - \widehat{\lambda}\right)(2x_2 - 2x_1)\right] e_2 \end{aligned}$$

is a subgradient of f at any point (x_1, x_2) with $x_1 = x_2$. It turns out that Theorem 13 has no straightforward generalization to the multidimensional case.

4.3. **A necessary and sufficient condition for optimality.** The next theorem provides a necessary and sufficient condition for the existence of an optimal solution to the problem (3.7).

Theorem 15. *Let S be a convex set in Λ^n and $f : S \to \mathbb{R}$ a convex function with the partial derivatives $\left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}}$ and $\left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}}$, $i = 1, 2, ..., n$, at the point $\overline{x} \in cint_{\Lambda^n}(S)$ satisfying (4.10). Then $\overline{x}$ is an optimal solution to the problem (3.7) if and only if there exist scalars $\lambda_i(\overline{x}) \in [0, 1]$ such that*

$$\xi(\overline{x}) := \sum_{i=1}^{n} \left(\lambda_i(\overline{x}) \left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}} + (1 - \lambda_i(\overline{x})) \left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}} \right) e_i \tag{4.20}$$

is a zero subgradient for f at $\overline{x}$.

Proof. Corollary 12 says that $\overline{x} \in cint_{\Lambda^n}(S)$ is an optimal solution to the problem (3.7) if and only if there exists a zero subgradient $\xi(\overline{x})$ of f at $\overline{x}$. By Theorem 14, the existence of the zero subgradient vector $\xi(\overline{x})$ is equivalent to the existence of scalars $\lambda_i(\overline{x}) \in [0,1]$ such that $\xi_i(\overline{x})$ defined by (4.16) is zero for each $i = 1, 2, ..., n$, and $\xi(\overline{x})$ given by (4.20) is a subgradient for f at $\overline{x}$. This completes the proof. □

If $\Lambda^n = \mathbb{R}^n$, then (4.17) holds, and hence, the vector $\xi(\overline{x})$ defined by (4.16) coincides with $gradf$. Thus, we can deduce the conventional condition for optimality from Theorem 15.

Remark 5. By Theorem 15, optimality of the point $\overline{x} \in \bigcup_{i=1}^n B_i(\overline{x}, h_i^{\pm}) \subset S$ implies that

$$\left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}} \leq 0 \leq \left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}} \quad \text{for all } i = 1, 2, ..., n. \tag{4.21}$$

If $\overline{x} \in S$ is a point such that $\sigma_i(\overline{x}_i) = \rho_i(\overline{x}_i) = 0$ for all $i = 1, 2, ..., n$, then (4.17) holds for all $i = 1, 2, ..., n$, and hence, the condition (4.21) turns into a sufficient condition guaranteeing optimality of $\overline{x}$. However, in the case when $\sigma_i(\overline{x}_i) = \rho_i(\overline{x}_i) = 0$ is not true for any $i = 1, 2, ..., n$, the condition (4.21) is only a necessary condition for optimality of $\overline{x}$. Using the condition (4.21) one may find the critical points which are candidates to be optimal solution to the problem (3.7). To guarantee optimality of such a point $\overline{x}$ one has to make sure that f has a zero subgradient $\xi(\overline{x})$ of the form (4.20) at $\overline{x}$.

Example 8. *Let* $\Lambda_0^2 = \mathbb{Z} \times \mathbb{Z}$*;* $S_0 = \{(-1,0), (0,-1), (0,0), (0,1), (1,0), (1,1)\}$. *Define the function* $f_0 : S_0 \to \mathbb{R}$ *as follows:* $f_0(-1,0) = f_0(0,-1) = f_0(1,0) = f_0(0,1) = 1$*,* $f_0(0,0) = 0$*, and* $f_0(1,1) = -1$*. Obviously,* f_0 *is convex on the convex set* S_0 *in* $\mathbb{Z} \times \mathbb{Z}$*. Even though, the condition (4.21) holds at the point* $\overline{x} = (0,0)$*,* f_0 *does not attain its minimum at* $\overline{x} = (0,0)$*. This is because* f_0 *has no zero subgradient at the point* $\overline{x} = (0,0)$*.*

Hereafter, we will explore the additional assumptions that makes the condition (4.21) the necessary and sufficient condition for the optimality of the point $\overline{x} \in S$.

Theorem 16. *Let* S *be a convex set in* Λ^n *and* $f : S \to \mathbb{R}$ *a convex function with the partial derivatives* $\left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}}$ *and* $\left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}}$, $i = 1, 2, ..., n$*, at the point* $\overline{x} \in S$*. Suppose that there exists a frame* $\bigcup_{i=1}^n B_i(\overline{x}, h_i^{\pm})$ *at* $\overline{x}$ *satisfying (4.10). Suppose also that* $grad_{\Delta} f$ *and* $grad_{\nabla} f$*, defined by (4.18) and (4.19), are subgradients for* f *at* $\overline{x} \in S$*. Then* $\overline{x}$ *is an optimal solution to the problem (3.7) if and only if (4.21) holds.*

Proof. Since

$$f(x) \geq f(\overline{x}) + grad_{\Delta} f(\overline{x})^T (x - \overline{x}) \text{ for all } x \in S$$

and

$$f(x) \geq f(\overline{x}) + grad_{\nabla} f(\overline{x})^T (x - \overline{x}) \text{ for all } x \in S,$$

we obtain

$$f(x) \geq f(\overline{x}) + \left(\lambda grad_{\Delta} f(\overline{x})^T + (1-\lambda) grad_{\nabla} f(\overline{x})^T\right)(x - \overline{x})$$

for all $x \in S$ and $\lambda \in [0,1]$. This and Corollary 12 shows that the necessary and sufficient condition for $\overline{x}$ to be an optimal solution to the problem (3.7) is that the inequality (4.21) holds. □

Example 9. *Let $\Lambda^2 = \mathbb{Z} \times \mathbb{Z}$ and $f : \Lambda^2 \to \mathbb{R}$ be defined by $f(x_1, x_2) = (x_1 - 1/3)^2 + (x_2 - 1/4)^2$. It is easy to verify that the vectors*

$$grad_{\Delta} f(x_1, x_2)^T = \left(2x_1 + \frac{1}{3}, 2x_2 + \frac{1}{2}\right)$$

and

$$grad_{\nabla} f(x_1, x_2)^T = \left(2x_1 - \frac{5}{3}, 2x_1 - \frac{3}{2}\right)$$

are subgradients for f at the point $(0,0)$. Since the inequality (4.21) holds at the point $(0,0)$, the optimal solution to the problem (3.7) is $(0,0)$ whenever $S = \mathbb{Z} \times \mathbb{Z}$.

In the next theorem we make the following assumptions:

A.1. Let $f : S \subset \Lambda^n \to \mathbb{R}$ be a function on a convex set $S \subset \Lambda^n$ such that f is a restriction of a convex function $F : conv_{\mathbb{R}^n}(S) \to \mathbb{R}^n$ to the set S.

A.2. All partial derivatives

$$\left.\frac{\partial F}{\partial x_i}\right|_{x=\overline{x}}, \quad i = 1, 2, ..., n$$

exist at the points of $\overset{\circ}{\widehat{conv_{\mathbb{R}^n}(S)}}$.

Theorem 17. *Let S be a convex set in Λ^n and $\overline{x} \in S$ a point such that (4.10) and (4.21) hold. Suppose that (A.1-A.2) hold. If $gradF(\overline{x})^T = 0$, then $\overline{x} \in S$ is an optimal solution to the problem (3.7).*

Proof. The proof follows from the classical result Lemma 4 and the fact that f is a restriction of a convex function F. $\square$

Example 10. *Let $\Lambda^2 = \mathbb{R} \times \mathbb{Z}$ and $f : \Lambda^2 \to \mathbb{R}$ be defined by $f(x_1, x_2) = (x_1 - x_2 - 1/3)^2$. The function f is the restriction of convex differentiable function $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $f(x_1, x_2) = (x_1 - x_2 - 1/3)^2$, to $\mathbb{R} \times \mathbb{Z}$. f attains its minimal value 0 over the set $S^* = \{(x_1, x_2) : x_1 - x_2 = 1/3\}$ since $gradF(\overline{x})^T = 0$, for all $\overline{x} \in S^*$.*

In the next result besides A.1 and A.2 we also make the following assumption:

A.3. The vector

$$\Omega f(\overline{x}) := \sum_{i=1}^{n} \omega_i(\overline{x}) e^i \tag{4.22}$$

defined by

$$\omega_i(\overline{x}) := \begin{cases} \left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}} & \text{if } \left.\frac{\partial F}{\partial x_i}\right|_{x=\overline{x}} < 0 \\ 0 & \text{if } \left.\frac{\partial F}{\partial x_i}\right|_{x=\overline{x}} = 0 \\ \left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}} & \text{if } \left.\frac{\partial F}{\partial x_i}\right|_{x=\overline{x}} > 0 \end{cases}, \quad i = 1, 2, ..., n \tag{4.23}$$

is a subgradient for f at $\overline{x} \in S$ whenever $gradF(\overline{x})^T \neq 0$.

Notice that $\Omega f(\overline{x}) = gradf(\overline{x})^T$ if $\Lambda^n = \mathbb{R}^n$.

Theorem 18. *Let S be a convex set in Λ^n and $\overline{x} \in S$ a point such that (4.10) holds. Suppose (A.1-A.3). Then $\overline{x}$ is an optimal solution to the problem (3.7) if and only if (4.21) holds.*

Proof. By Lemma 4 and (A.1-A.3) we have

$$f(x) \geq f(\overline{x}) + \left[\lambda gradF(\overline{x})^T + (1-\lambda)\Omega f(\overline{x})\right](x - \overline{x}) \tag{4.24}$$

for all $x \in S$ and $\lambda \in [0,1]$. (4.24), (4.22-4.23), and Corollary 12 shows that $\overline{x}$ is an optimal solution to the problem (3.7) if and only if (4.21) holds. □

Remark 6. The above theorem offers a procedure for the determination of an optimal solution by solving at most $2n$ inequalities in n variables instead of solving a system of n equations in n variables.

Example 11. *Let the function $f : \mathbb{T}_1 \times \mathbb{T}_2 \to \mathbb{R}$ be defined by*

$$f(x_1, x_2) = (x_1 - x_2 - 1/2)^2.$$

Obviously, f is a restriction of the convex function $F(x_1, x_2) = (x_1 - x_2 - 1/2)^2$ with the partial derivatives

$$\frac{\partial F}{\partial x_1} = -\frac{\partial F}{\partial x_2} = 2(x_1 - x_2 - 1/2).$$

In the following we consider several time scales of $\mathbb{T}_1$ and $\mathbb{T}_2$ such that the function $f(x_1, x_2)$ attains its minimum over the product $\mathbb{T}_1 \times \mathbb{T}_2$

Case 1. *If $\mathbb{T}_1 = \mathbb{T}_2 = \mathbb{R}$, then*

$$\frac{\partial f}{\Delta_1 x_1} = \frac{\partial f}{\nabla_1 x_1} = 2\left(x_1 - x_2 - \frac{1}{2}\right) = -\frac{\partial f}{\Delta_2 x_2} = -\frac{\partial f}{\nabla_2 x_2}.$$

Obviously,

$$\Omega_1 f(\overline{x}) = gradF(\overline{x})$$

is a subgradient for f at any point of $\mathbb{R}^2$. The condition (4.21) turns into conventional condition for optimality and we know that f attains its minimum over the set

$$S_1^* = \left\{(x_1, x_2) \in \mathbb{R} \times \mathbb{R} : x_1 - x_2 - \frac{1}{2} = 0\right\}.$$

Consequently,

$$\min_{(x_1,x_2)\in\mathbb{R}\times\mathbb{R}} f(x_1, x_2) = 0.$$

Case 2. *If $\mathbb{T}_1 = \mathbb{Z}$ and $\mathbb{T}_2 = \mathbb{R}$, then the inequality (4.21) implies that*

$$\frac{\partial f}{\Delta_1 x_1} = 2x_1 - 2x_2 \geq 0 \geq \frac{\partial f}{\nabla_1 x_1} = 2x_1 - 2x_2 - 2$$

and

$$\frac{\partial f}{\Delta_2 x_2} = \frac{\partial f}{\nabla_2 x_2} = -2\left(x_1 - x_2 - \frac{1}{2}\right) = 0.$$

Therefore,

$$\begin{aligned} S_2^* &= \left\{(x_1, x_2) \in \mathbb{Z} \times \mathbb{R} : x_1 - x_2 - 1 \leq 0 \leq x_1 - x_2 \text{ and } x_1 - x_2 - \frac{1}{2} = 0\right\} \\ &= \left\{(x_1, x_2) \in \mathbb{Z} \times \mathbb{R} : x_1 - x_2 - \frac{1}{2} = 0\right\} \end{aligned}$$

is the set of critical points. Since $\left.\frac{\partial F}{\partial x_1}\right|_{\overline{x}} = \left.\frac{\partial F}{\partial x_2}\right|_{\overline{x}} = 0$ for $\overline{x} \in S_2^*$ and

$$\Omega_2 f(\overline{x}) = (0, 0)$$

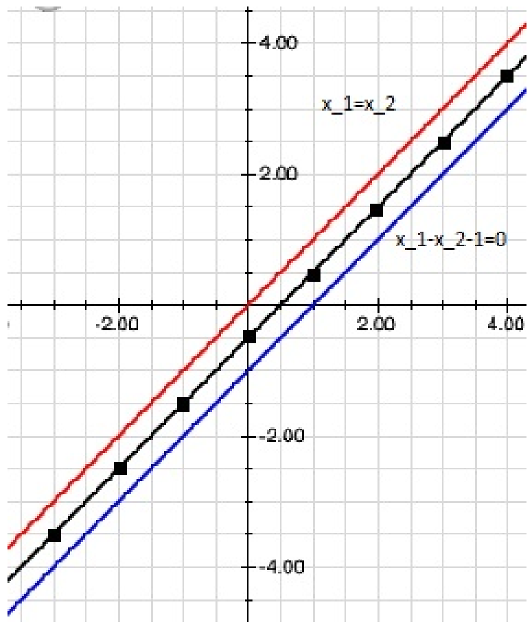

FIGURE 7. Upper line is the line $x_1 = x_2$ and lower line is the line $x_1 - x_2 - 1 = 0$. Black dots are the elements of S_2^*

is a subgradient for f at every point $\overline{x}$ of S_2^*, the set S_2^* becomes the set of all alternative solutions to the problem (3.7) and the minimum value of f over $\mathbb{Z} \times \mathbb{R}$ is

$$\min_{(x_1,x_2)\in\mathbb{Z}\times\mathbb{R}} f(x_1, x_2) = 0$$

(see Figure 7).

Case 3. *If $\mathbb{T}_1 = \mathbb{T}_2 = \mathbb{Z}$, then the condition (4.21) implies that*

$$\frac{\partial f}{\Delta_1 x_1} = 2x_1 - 2x_2 \geq 0 \geq \frac{\partial f}{\nabla_1 x_1} = 2x_1 - 2x_2 - 2$$

and

$$\frac{\partial f}{\Delta_2 x_2} = 2x_2 - 2x_1 + 2 \geq 0 \geq \frac{\partial f}{\nabla_2 x_2} = 2x_2 - 2x_1.$$

Hence we obtain the set

$$\begin{aligned} S_3^* &= \{(x_1, x_2) \in \mathbb{Z} \times \mathbb{Z} : 2x_1 - 2x_2 \geq 0 \geq 2x_1 - 2x_2 - 2\} \\ &= \{(x_1, x_2) \in \mathbb{Z} \times \mathbb{Z} : x_1 = x_2\} \\ &\cup \{(x_1, x_2) \in \mathbb{Z} \times \mathbb{Z} : x_1 - x_2 = 1\} \end{aligned}$$

as the set of critical points. Since $\left.\frac{\partial F}{\partial x_1}\right|_{\overline{x}} = -\left.\frac{\partial F}{\partial x_2}\right|_{\overline{x}} = -1$ for all $\overline{x} \in \{(x_1, x_2) \in \mathbb{Z} \times \mathbb{Z} : x_1 = x_2\}$ we have

$$\begin{aligned} \Omega_3^+ f(\overline{x}) &= \left(\frac{\partial f}{\Delta_1 x_1}, \frac{\partial f}{\nabla_2 x_2}\right) \\ &= (2x_1 - 2x_2, 2x_2 - 2x_1) \end{aligned}$$

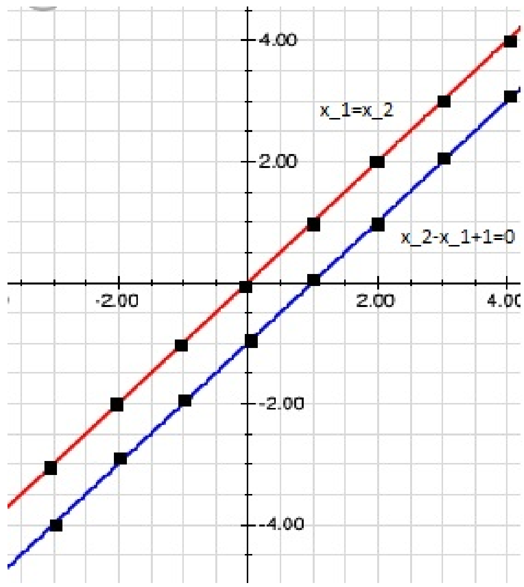

FIGURE 8. Upper line is the line $x_2 = x_1$ and lower line is the line $x_2 - x_1 + 1 = 0$. Black dots are the elements of S_3^*

for $\overline{x} \in \{(x_1, x_2) \in \mathbb{Z}\times\mathbb{Z} : x_1 = x_2\}$. *Obviously* $\Omega_3^+ f(\overline{x})$ *is a subgradient for* f *at each point of the set* $\{(x_1, x_2) \in \mathbb{Z} \times \mathbb{Z} : x_1 = x_2\}$. *Similarly, since* $\frac{\partial F}{\partial x_1}\Big|_{\overline{x}} = -\frac{\partial F}{\partial x_2}\Big|_{\overline{x}} = 1$ *for all* $\overline{x} \in \{(x_1, x_2) \in \mathbb{Z} \times \mathbb{Z} : x_1 - x_2 = 1\}$ *we have*

$$\begin{aligned}\Omega_3^- f(\overline{x}) &= \left(\frac{\partial f}{\nabla_1 x_1}, \frac{\partial f}{\Delta_2 x_2}\right)\\ &= (2x_1 - 2x_2 - 2, 2x_2 - 2x_1 + 2)\end{aligned}$$

for $\overline{x} \in \{(x_1, x_2) \in \mathbb{Z} \times \mathbb{Z} : x_1 - x_2 = 1\}$. *Since* $\Omega_3^- f(\overline{x})$ *is a subgradient for* f *at each point of the set* $(x_1, x_2) \in \{\mathbb{Z} \times \mathbb{Z} : x_1 - x_2 = 1\}$, *the set* S_3^* *becomes the set of alternative solutions and the value*

$$\min_{(x_1,x_2)\in\mathbb{Z}\times\mathbb{Z}} f(x_1, x_2) = \frac{1}{4}$$

is the minimum value of f *over* $\mathbb{Z} \times \mathbb{Z}$ *(see Figure 8).*

Case 4. *If* $\mathbb{T}_1 = \mathbb{Z}$ *and* $\mathbb{T}_2 = 2^{\mathbb{N}} = \{2^n : n \in \mathbb{N}\}$, *then the condition (4.21) implies that*

$$\frac{\partial f}{\Delta_1 x_1} = 2x_1 - 2x_2 \geq 0 \geq \frac{\partial f}{\nabla_1 x_1} = 2x_1 - 2x_2 - 2$$

and

$$\frac{\partial f}{\Delta_2 x_2} = 3x_2 - 2x_1 + 1 \geq 0 \geq \frac{\partial f}{\nabla_2 x_2} = \frac{3}{2}x_2 - 2x_1 + 1.$$

Hence,

$$\begin{aligned}S_4^* &= \left\{(x_1, x_2) \in \mathbb{Z} \times 2^{\mathbb{N}} : x_1 = x_2\right\}\\ &\cup \left\{(x_1, x_2) \in \mathbb{Z} \times 2^{\mathbb{N}} : x_1 - x_2 = 1\right\}\end{aligned}$$

is the set of critical values. Since $\frac{\partial F}{\partial x_1}\Big|_{\overline{x}} = -\frac{\partial F}{\partial x_2}\Big|_{\overline{x}} = -1$ *for all* $\overline{x} \in \{(x_1, x_2) \in \mathbb{Z} \times 2^{\mathbb{N}} : x_1 = x_2\}$ *and* $\frac{\partial F}{\partial x_1}\Big|_{\overline{x}} = -\frac{\partial F}{\partial x_2}\Big|_{\overline{x}} = 1$ *for all* $\overline{x} \in \{(x_1, x_2) \in \mathbb{Z} \times 2^{\mathbb{N}} : x_1 - x_2 = 1\}$ *we obtain*

$$\Omega_4^+ f(\overline{x}) = \left(\frac{\partial f}{\Delta_1 x_1}, \frac{\partial f}{\nabla_2 x_2}\right)$$
$$= (2x_1 - 2x_2, \frac{3}{2}x_2 - 2x_1 + 1)$$

and

$$\Omega_4^- f(\overline{x}) = \left(\frac{\partial f}{\nabla_1 x_1}, \frac{\partial f}{\Delta_2 x_2}\right)$$
$$= (2x_1 - 2x_2 - 2, 3x_2 - 2x_1 + 1)$$

as the subgradients for f *at the elements of the sets* $\{(x_1, x_2) \in \mathbb{Z} \times 2^{\mathbb{N}} : x_1 = x_2\}$ *and* $\{(x_1, x_2) \in \mathbb{Z} \times 2^{\mathbb{N}} : x_1 - x_2 = 1\}$*, respectively. Consequently,* S_4^* *is the set of alternative solutions and*

$$\min_{(x_1,x_2)\in\mathbb{Z}\times 2^{\mathbb{N}}} f(x_1, x_2) = \frac{1}{4}$$

is the minimum value of f *over the set* $\mathbb{Z} \times 2^{\mathbb{N}}$ *(see Figure 9).*

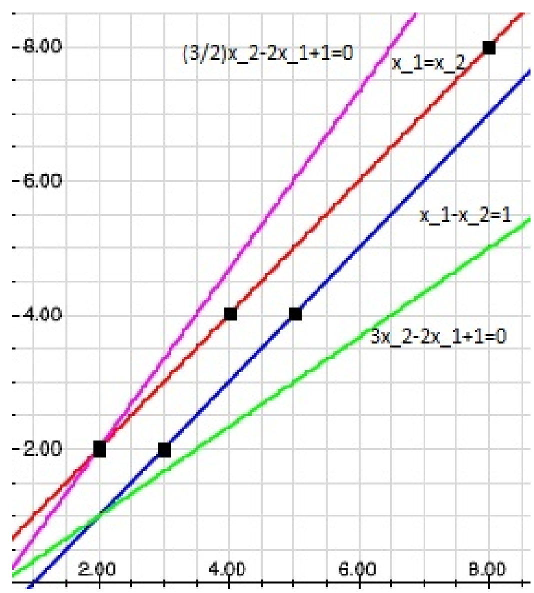

FIGURE 9. Black dots are the elements of S_4^*

Case 5. *If* $\mathbb{T}_1 = \mathbb{Z}$ *and* $\mathbb{T}_2 = \frac{1}{3}\mathbb{Z} = \{n/3 : n \in \mathbb{Z}\}$*, then the condition (4.21) implies that*

$$\frac{\partial f}{\Delta_1 x_1} = 2x_1 - 2x_2 \geq 0 \geq \frac{\partial f}{\nabla_1 x_1} = 2x_1 - 2x_2 - 2$$

and

$$\frac{\partial f}{\Delta_2 x_2} = 2x_2 - 2x_1 + \frac{4}{3} \geq 0 \geq \frac{\partial f}{\nabla_2 x_2} = 2x_2 - 2x_1 + \frac{2}{3}.$$

Solving these inequalities for $(x_1, x_2) \in \mathbb{Z} \times \frac{1}{3}\mathbb{Z}$ *leads to*

$$S_5^* = \left\{(x_1, x_2) \in \mathbb{Z} \times \frac{1}{3}\mathbb{Z} : 2x_2 - 2x_1 + \frac{4}{3} = 0\right\}$$
$$\cup \left\{(x_1, x_2) \in \mathbb{Z} \times \frac{1}{3}\mathbb{Z} : 2x_2 - 2x_1 + \frac{2}{3} = 0\right\}$$

as the set of critical points. Since Since $\left.\frac{\partial F}{\partial x_1}\right|_{\overline{x}} = -\left.\frac{\partial F}{\partial x_2}\right|_{\overline{x}} = 1/6$ *for all* $\overline{x} \in \{(x_1, x_2) \in \mathbb{Z} \times \frac{1}{3}\mathbb{Z} : 2x_2 - 2x_1 + \frac{4}{3} = 0\}$ *and* $\left.\frac{\partial F}{\partial x_1}\right|_{\overline{x}} = -\left.\frac{\partial F}{\partial x_2}\right|_{\overline{x}} = -1/6$ *for all* $\overline{x} \in \{(x_1, x_2) \in \mathbb{Z} \times \frac{1}{3}\mathbb{Z} : 2x_2 - 2x_1 + \frac{2}{3} = 0\}$ *we obtain*

$$\Omega_5^+ f(\overline{x}) = \left(\frac{\partial f}{\nabla_1 x_1}, \frac{\partial f}{\Delta_2 x_2}\right)$$
$$= (2x_1 - 2x_2 - 2, 2x_2 - 2x_1 + \frac{4}{3})$$

and

$$\Omega_5^- f(\overline{x}) = \left(\frac{\partial f}{\Delta_1 x_1}, \frac{\partial f}{\nabla_2 x_2}\right)$$
$$= (2x_1 - 2x_2, 2x_2 - 2x_1 + \frac{2}{3})$$

as the subgradients for f *at the elements of the sets* $\{(x_1, x_2) \in \mathbb{Z} \times \frac{1}{3}\mathbb{Z} : 2x_2 - 2x_1 + \frac{4}{3} = 0\}$ *and* $\{(x_1, x_2) \in \mathbb{Z} \times \frac{1}{3}\mathbb{Z} : 2x_2 - 2x_1 + \frac{2}{3} = 0\}$*, respectively. Consequently,* S_5^* *is the set of alternative solutions and*

$$\min_{(x_1,x_2)\in\mathbb{Z}\times\mathbb{Z}/3} f(x_1, x_2) = \frac{1}{36}$$

is the minimum value of f *over* $\mathbb{Z} \times \frac{1}{3}\mathbb{Z}$.

5. **Optimization on time scales.** In this section we propose the linear programming **(LP)** and the quadratic programming **(QP)** problems on time scales. We start with an extension of LP problem.

5.1. **LP problem on time scales.** Consider the problem

$$\text{minimize} \sum_{i=1}^{n} c_i x_i$$
$$\text{s.t.} \sum_{j=1}^{n} a_{ij} x_j \le b_i \quad \text{for } i = 1, 2, .., n$$
$$x = (x_1, ..., x_n) \in \mathbb{T}_1 \times \mathbb{T}_2 \times ... \times \mathbb{T}_n \tag{5.1}$$

where the coefficients c_i, a_{ij}, b_i, for all $i, j \in \{1, 2, ..., n\}$, are real numbers and $\mathbb{T}_j$, $j = 1, 2, ..n$, are time scales. It is obvious that the problem (5.1) becomes the standard LP problem when $\mathbb{T}_j = \mathbb{R}$, for $i = 1, 2, .., n$, and an integer programming problem when $\mathbb{T}_j = \mathbb{Z}$, for $i = 1, 2, .., n$.

Since there are many time scales other than $\mathbb{R}$ and $\mathbb{Z}$, the problem (5.1) lead to different type of optimization problems.

Given a linear function $f(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} c_i x_i$ defined on a convex region A in $\mathbb{R}^n$, we know by Corollary 11 that the restriction $f|_{\Lambda^n} : A \cap \Lambda^n \to \mathbb{R}$ of f into $A \cap \Lambda^n$ is convex on $A \cap \Lambda^n$. Moreover, we have

$$\Omega g(\overline{x}) = grad_{\Delta} g(\overline{x}) = grad_{\nabla} g(\overline{x}) = gradf(\overline{x}) = \sum_{i=1}^{n} c_i e_i,$$

where $g = f|_{\Lambda^n}$ and $grad_{\Delta} g(\overline{x})$ and $grad_{\nabla} g(\overline{x})$ are defined as in (4.18) and (4.19), respectively. Then Theorem 14 leads to the next result.

Theorem 19. *Let S be a convex set in Λ^n and $f : conv_{\mathbb{R}^n}(S) \to \mathbb{R}$ a linear function. Then, for each $\overline{x} \in cint_{\Lambda^n}(S)$ satisfying (4.10)*

$$grad_{\Delta} f(\overline{x}) = grad_{\nabla} f(\overline{x}) = gradf(\overline{x})$$

holds and $\xi(\overline{x}) = gradf(\overline{x})$ is a subgradient for f at $\overline{x}$, i.e.,

$$f(x) \geq f(\overline{x}) + gradf(\overline{x})^T (x - \overline{x}) \text{ for all } x \in S.$$

Notice that the next result is well known in the literature for conventional continuous optimization.

Theorem 20. [7, p. 103] *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and A be a nonempty convex set in $\mathbb{R}^n$. If f is differentiable, then $\overline{x} \in A$ is an optimal solution to the problem*

$$\text{minimize } f(x) \text{ subject to } x \in A$$

if and only if

$$(gradf(\overline{x}))^T (x - \overline{x}) \geq 0 \text{ for all } x \in A.$$

Furthermore, if A is open, then $\overline{x} \in A$ is an optimal solution if and only if

$$gradf(\overline{x}) = 0.$$

We have a parallel result for linear optimization on time scales:

Remark 7. Let A be a convex set in $\mathbb{R}^n$ such that $cint_{\Lambda^n}(A \cap \Lambda^n) \neq \emptyset$. It is clear that $S := A \cap \Lambda^n$ is convex in Λ^n. Given a linear function $f : \mathbb{R}^n \to \mathbb{R}$, it is evident that $x^* \in S$ is a solution to the problem

$$\text{minimize } f(x) \text{ subject to } x \in S \tag{5.2}$$

if and only if x^* is a solution to the problem

$$\text{minimize } f(x) \text{ subject to } x \in conv_{\mathbb{R}^n}(S).$$

By replacing A with $conv_{\mathbb{R}^n}(S)$ in Theorem 20, we arrive at the conclusion that $\overline{x} \in S$ is an optimal solution to the problem (5.2) if and only if

$$(gradf(\overline{x}))^T (x - \overline{x}) \geq 0 \text{ for all } x \in S.$$

In Figure 10, we illustrate an application of Remark 7 to an integer linear programming problem**(ILP)**.

Example 12. *Consider the following LP on time scales:*

$$\begin{array}{l} \max\ 3x + 5y \\ s.t.\ \ 2x - y \leq 14 \\ \qquad\ x + y \leq 16 \\ \qquad\ x - y \leq 10 \\ x \in \mathbb{R}_+ \cup \{0\} \text{ and } y \in 2^{\mathbb{N}} \cup \{0\} \end{array}.$$

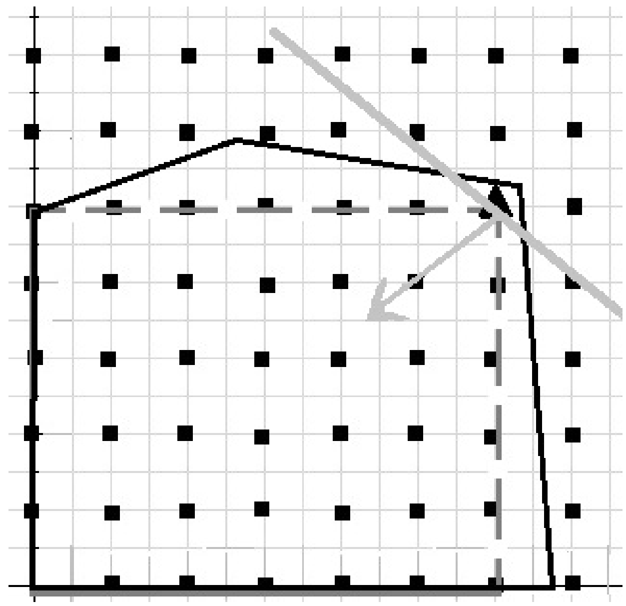

FIGURE 10. The region bounded by the black lines is the polytope A. The black dots surrounded by the dashed rectangle are the elements of the convex set $S = A \cap (\mathbb{Z} \times \mathbb{Z})$ in $\mathbb{Z} \times \mathbb{Z}$. The triangular point is the point $\overline{x}$ at which grad f makes an acute angle with all vectors pointing from $\overline{x}$ towards elements of S. The grey line is the hyperplane whose lower halfplane contains S.

Obviously, the optimal solution (x^*, y^*) to this problem is $(8, 8)$, where the optimal value of $f(x, y) = 3x+5y$ is 64 and the gradf $= (3, 5)$ makes an obtuse angle with all vectors pointing from (x^*, y^*) towards any point of the feasible domain (see Figure 11).

The above example indicates that linear optimization on time scales not only may unify linear programming (LP) and integer programming (IP) but also may provide more general perspective for constructing new models for solving problems.

5.2. **QP problem on time scales.** Consider the following quadratic optimization problem over time scales:

$$\begin{aligned} \text{minimize } & f(x) = \frac{1}{2}x^T Qx + c^T x \\ \text{s.t. } & Ax \leq b \\ & x \in \mathbb{T}_1 \times \mathbb{T}_2 \times ... \times \mathbb{T}_n, \end{aligned} \tag{5.3}$$

where Q is an $n \times n$ symmetric and positive semidefinite matrix, c is an n-dimensional column vector, A is an $m \times n$ constraint matrix, b is an m-dimensional column vector, and $\mathbb{T}_1$, $\mathbb{T}_2$,...,$\mathbb{T}_n$ are time scales. Letting $\mathbb{T}_i = \mathbb{R}$ for all $i = 1, 2, ..., n$, we obtain standard linearly constrained convex quadratic programming problem. Letting $\mathbb{T}_i = \mathbb{Z}$ for all $i = 1, 2, ..., n$, we obtain a convex quadratic integer programming problem. Choosing different time scales in (5.3), we may get hybrid quadratic programming problems.

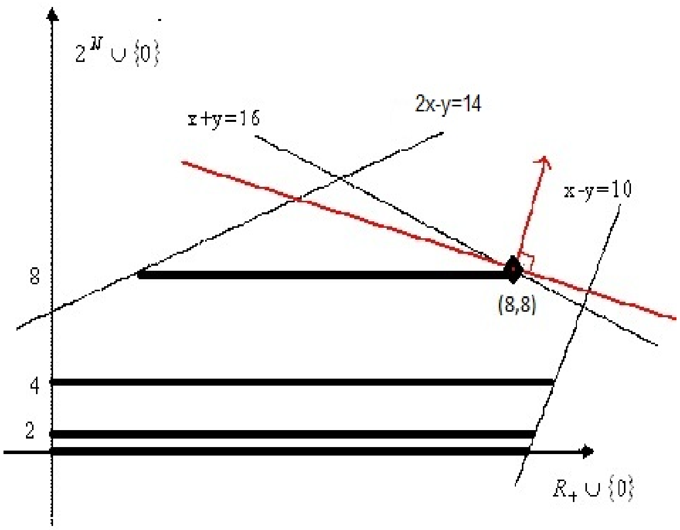

FIGURE 11. The feasible domain for the above given LP consists of four parallel lines bounded by the constraints

We would like to particularly point out that the study of optimization on time scales may provide easiness and new algorithms for solving problems besides the generalization. To illustrate our point, let us consider the following unconstrained optimization problem:

$$\text{minimize } f(x) = \frac{1}{2}x^T Qx + c^T x$$
$$\text{for } x \in \mathbb{T}^n,$$

where Q and c are defined as in (5.3) and $\mathbb{T}$ is a time scale. If we let $\mathbb{T} = \{\rho(0), 0, 1, \sigma(1)\}$, $\rho(0) < 0$, $1 < \sigma(1)$ and $S = \mathbb{T}^n$, then we have $cint_{\mathbb{T}^n}(S) = \{0,1\}^n \neq \emptyset$. It is easy to see that $f(x)$ is convex and S is also convex in $\mathbb{T}^n$. Hence we can apply Theorem 18 to (5.3) to check the existence of an optimal solution. Consequently, we know

$$\overline{x} = (\overline{x}_1, \overline{x}_2, ..., \overline{x}_n) \in cint_{\mathbb{T}^n}(S) = \{0,1\}^n$$

is an optimal solution to the problem (5.3) if and only if $\overline{x}$ is an element of the set

$$S^* = \cap_{i=1}^n \left\{ \overline{x} \in \{0,1\}^n : \left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}} \leq 0 \leq \left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}} \right\},$$

where

$$\left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}} = \begin{cases} f_i(1) - f_i(0) & \text{if } \overline{x}_i = 0 \\ \frac{f_i(\sigma(1)) - f_i(1)}{\sigma(1)-1} & \text{if } \overline{x}_i = 1 \end{cases}, \tag{5.4}$$

$$\left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}} = \begin{cases} \frac{f_i(0) - f_i(\rho(0))}{-\rho(0)} & \text{if } \overline{x}_i = 0 \\ f_i(1) - f_i(0) & \text{if } \overline{x}_i = 1 \end{cases}, \tag{5.5}$$

and

$$f_i(s) = f(\overline{x}_1, ..., \overline{x}_{i-1}, s, \overline{x}_{i+1}, ..., \overline{x}_n),$$

provided $\Omega f(\overline{x})$ is a subgradient for f at any $\overline{x} \in S^*$. On the other hand, if the problem (5.3) has an optimal solution $\overline{x} \in cint_{\mathbb{T}^n}(S) = \{0,1\}^n$, then the problem

$$\begin{aligned} &\text{minimize } f(x) = \frac{1}{2}x^T Q x + c^T x \\ &\qquad\quad \text{for } x \in \{0,1\}^n \end{aligned} \tag{5.6}$$

has the same optimal solution. Hence, we have derived a sufficient condition for the optimal solutions to the problem (5.6).

Corollary 14. *Let f be a function such that the set*

$$S^* = \cap_{i=1}^n \left\{ \overline{x} \in \{0,1\}^n : \left.\frac{\partial f(x)}{\nabla_i x_i}\right|_{x=\overline{x}} \le 0 \le \left.\frac{\partial f(x)}{\Delta_i x_i}\right|_{x=\overline{x}} \right\}, \tag{5.7}$$

is non-empty and Ωf defined by (4.22) is a subgradient for f at every $\overline{x} \in S^*$. Then every element in S^* is an optimal solution to the problem (5.6), and vice versa.

Example 13. *Define $\rho(0) = -1$, $\sigma(1) = 2$ to obtain the time scale $\mathbb{T} = \{-1,0,1,2\}$. Let*

$$Q = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix} \text{ and } c = \begin{bmatrix} -1 \\ -4 \\ -4 \end{bmatrix}.$$

Evidently, Q is a symmetric and positive definite matrix and f is convex on the convex set $S = \mathbb{T}^n$. Consider the problem

$$\begin{aligned} &\textit{minimize } f(x) = -x_1 - 4x_2 - 4x_3 + x_1^2 + 3x_2^2 + 3x_3^2 \\ &\quad \textit{s.t. } x_1, x_2, x_3 \in \{0,1\}. \end{aligned} \tag{5.8}$$

By (5.4-5.5), we have

$$\frac{\partial f(x)}{\nabla_1 x_1} = 2x_1 - 2, \quad \frac{\partial f(x)}{\Delta_1 x_1} = 2x_1,$$

$$\frac{\partial f(x)}{\nabla_2 x_2} = 6x_2 - 7, \quad \frac{\partial f(x)}{\Delta_2 x_2} = 6x_2 - 1,$$

and

$$\frac{\partial f(x)}{\nabla_3 x_3} = 6x_3 - 7, \quad \frac{\partial f(x)}{\Delta_3 x_3} = 6x_3 - 1.$$

Since

$$\begin{aligned} S_1^* &= \left\{ (x_1,x_2,x_3) \in \{0,1\}^3 : \frac{\partial f(x)}{\nabla_1 x_1} \le 0 \le \frac{\partial f(x)}{\Delta_1 x_1} \right\} \\ &= \{(0,0,0),(0,0,1),(0,1,0),(0,1,1),(1,0,0),(1,0,1),(1,1,0),(1,1,1)\}, \end{aligned}$$

$$\begin{aligned} S_2^* &= \left\{ (x_1,x_2,x_3) \in \{0,1\}^3 : \frac{\partial f(x)}{\nabla_2 x_2} \le 0 \le \frac{\partial f(x)}{\Delta_2 x_2} \right\} \\ &= \{(0,1,0),(1,1,0),(0,1,1),(1,1,1)\}, \end{aligned}$$

and

$$\begin{aligned} S_3^* &= \left\{ (x_1,x_2,x_3) \in \{0,1\}^3 : \frac{\partial f(x)}{\nabla_3 x_3} \le 0 \le \frac{\partial f(x)}{\Delta_3 x_3} \right\} \\ &= \{(1,0,1),(0,0,1),(0,1,1),(1,1,1)\}, \end{aligned}$$

we obtain

$$S^* = S_1^* \cap S_2^* \cap S_3^* = \{(0,1,1),(1,1,1)\}.$$

Since

$$gradf(0,1,1)^T = (-1,2,2)$$

and

$$gradf(1,1,1)^T = (1,2,2)$$

we have

$$\begin{aligned}\Omega f(0,1,1) &= \left(\frac{\partial f(x)}{\Delta_1 x_1}, \frac{\partial f(x)}{\nabla_2 x_2}, \frac{\partial f(x)}{\nabla_3 x_3}\right)_{(0,1,1)} \\ &= (2x_1, 6x_2 - 7, 6x_3 - 7)_{(0,1,1)} \\ &= (0,-1,-1)\end{aligned}$$

and

$$\begin{aligned}\Omega f(1,1,1) &= \left(\frac{\partial f(x)}{\nabla_1 x_1}, \frac{\partial f(x)}{\nabla_2 x_2}, \frac{\partial f(x)}{\nabla_3 x_3}\right)_{(1,1,1)} \\ &= (2x_1 - 2, 6x_2 - 7, 6x_3 - 7)_{(1,1,1)} \\ &= (0,-1,-1)\end{aligned}$$

which are the subgradients of f *at* $(0,1,1)$ *and* $(1,1,1)$*, respectively. By Corollary 14, we can deduce that* $(0,1,1)$ *and* $(1,1,1)$ *are optimal solutions to the problem (5.8) at which the function* $f(x) = -x_1 - 4x_2 - 4x_3 + x_1^2 + 3x_2^2 + 3x_3^2$ *attains a minimum value of* -2.

Note that the existence of an optimal solution to the problem (5.6) is guaranteed by the existence of an element of the set S^* defined by (5.7). However, one may easily find a convex function $f(x) = \frac{1}{2}x^TQx + c^Tx$ such that the setting $\mathbb{T} = \{-1,0,1,2\}$ in Example 13 does not work for finding at least one element of S^*. In such cases, we can redefine the time scale $\mathbb{T} = \{\rho(0), 0, 1, \sigma(1)\}$ to show that $S^* \neq \emptyset$.

Example 14. *Let*

$$Q = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{bmatrix} \text{ and } c = \begin{bmatrix} -1 \\ -16 \\ -16 \end{bmatrix}.$$

for the following quadratic optimization problem:

$$\begin{aligned}&\text{minimize } f(x) = -x_1 - 16x_2 - 16x_3 + x_1^2 + 4x_2^2 + 4x_3^2 \\ &\text{s.t. } x_1, x_2, x_3 \in \{0,1\}. \end{aligned} \tag{5.9}$$

We consider two cases:
Case 1. *(*$S^* = \emptyset$ *) Let* $\mathbb{T} = \{-1,0,1,2\}$*. From (5.4-5.5), we have*

$$\frac{\partial f(x)}{\nabla_2 x_2} = 8x_2 - 20 \quad \text{and} \quad \frac{\partial f(x)}{\Delta_2 x_2} = 8x_2 - 12.$$

This implies that

$$S_2^* = \left\{(x_1,x_2,x_3) \in \{0,1\}^3 : \frac{\partial f(x)}{\nabla_2 x_2} \leq 0 \leq \frac{\partial f(x)}{\Delta_2 x_2}\right\} = \varnothing.$$

Hence we have $S^* = \varnothing$*. In this case, Corollary 14 yields no optimal solution for (5.9).*

Case 2. *($S^* \neq \emptyset$) Set $\rho(0) = -4$ and $\sigma(1) = 5$ to define the time scale $\mathbb{T} = \{-4, 0, 1, 5\}$. From (5.4-5.5), we have*

$$\frac{\partial f(x)}{\nabla_1 x_1} = \begin{cases} \frac{1}{4}\left(x_1^2 - (x_1-4)^2\right) - 1 & \text{if } x_1 = 0 \\ 2x_1 - 2 & \text{if } x_1 = 1 \end{cases},$$

$$\frac{\partial f(x)}{\Delta_1 x_1} = \begin{cases} 2x_1 & \text{if } x_1 = 0 \\ \frac{1}{4}\left((x_1+4)^2 - x_1^2\right) - 1 & \text{if } x_1 = 1 \end{cases},$$

$$\frac{\partial f(x)}{\nabla_2 x_2} = \begin{cases} \left(x_2^2 - (x_2-4)^2\right) - 16 & \text{if } x_2 = 0 \\ 8x_2 - 20 & \text{if } x_2 = 1 \end{cases},$$

$$\frac{\partial f(x)}{\Delta_2 x_2} = \begin{cases} 8x_2 - 12 & \text{if } x_2 = 0 \\ \left((x_2+4)^2 - x_2^2\right) - 16 & \text{if } x_2 = 1 \end{cases},$$

$$\frac{\partial f(x)}{\nabla_3 x_3} = \begin{cases} \left(x_3^2 - (x_3-4)^2\right) - 16 & \text{if } x_3 = 0 \\ 8x_3 - 20 & \text{if } x_3 = 1 \end{cases},$$

$$\frac{\partial f(x)}{\Delta_2 x_2} = \begin{cases} 8x_3 - 12 & \text{if } x_3 = 0 \\ \left((x_3+4)^2 - x_3^2\right) - 16 & \text{if } x_3 = 1 \end{cases},$$

and, consequently,

$$\frac{\partial f(x)}{\nabla_1 x_1} = \begin{cases} -5 & \text{if } x_1 = 0 \\ 0 & \text{if } x_1 = 1 \end{cases}, \frac{\partial f(x)}{\Delta_1 x_1} = \begin{cases} 0 & \text{if } x_1 = 0 \\ 5 & \text{if } x_1 = 1 \end{cases},$$

$$\frac{\partial f(x)}{\nabla_2 x_2} = \begin{cases} -32 & \text{if } x_2 = 0 \\ -12 & \text{if } x_2 = 1 \end{cases}, \frac{\partial f(x)}{\Delta_2 x_2} = \begin{cases} -12 & \text{if } x_2 = 0 \\ 8 & \text{if } x_2 = 1 \end{cases},$$

$$\frac{\partial f(x)}{\nabla_3 x_3} = \begin{cases} -32 & \text{if } x_3 = 0 \\ -12 & \text{if } x_3 = 1 \end{cases}, \frac{\partial f(x)}{\Delta_2 x_2} = \begin{cases} -12 & \text{if } x_3 = 0 \\ 8 & \text{if } x_3 = 1 \end{cases}.$$

Therefore, we have

$$\begin{aligned} S_1^* &= \left\{(x_1, x_2, x_3) \in \{0,1\}^3 : \frac{\partial f(x)}{\nabla_1 x_1} \leq 0 \leq \frac{\partial f(x)}{\Delta_1 x_1}\right\} \\ &= \{(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1)\}, \end{aligned}$$

$$\begin{aligned} S_2^* &= \left\{(x_1, x_2, x_3) \in \{0,1\}^3 : \frac{\partial f(x)}{\nabla_2 x_2} \leq 0 \leq \frac{\partial f(x)}{\Delta_2 x_2}\right\} \\ &= \{(0,1,0), (1,1,0), (0,1,1), (1,1,1)\}, \end{aligned}$$

and

$$\begin{aligned} S_3^* &= \left\{(x_1, x_2, x_3) \in \{0,1\}^3 : \frac{\partial f(x)}{\nabla_3 x_3} \leq 0 \leq \frac{\partial f(x)}{\Delta_3 x_3}\right\} \\ &= \{(1,0,1), (0,0,1), (0,1,1), (1,1,1)\}. \end{aligned}$$

Consequently, we obtain the optimal set

$$S^* = S_1^* \cap S_2^* \cap S_3^* = \{(0,1,1), (1,1,1)\}$$

at which the function $f(x) = -x_1 - 16x_2 - 16x_3 + x_1^2 + 4x_2^2 + 4x_3^2$ attains a minimum value of -24.

The previous examples lead to an open problem to be answered.

Open problem. Given an $n \times n$ symmetric and positive semi definite matrix Q and an $n \times 1$ column vector c. Is there always a time scale $\mathbb{T} = \{\rho(0), 0, 1, \sigma(1)\}$ such that the set S^* defined by (5.7) is non-empty?

6. **Concluding remarks.** The purpose of this paper is two fold: First, we introduce the fundamental concepts of convexity of sets and functions defined on time scales and derive their analytic properties for further analysis. Second, we propose a new problem of optimization on time scales for system modeling with both continuous and discrete variables. Before our work, time scale systems have been considered only for the generalization of differential, difference, q-difference and h-difference equations by using dynamic equations containing Δ-derivative of the functions. This paper is the first work in which sets in the product of time scales are investigated in terms of their convexity properties. Different from the convexity notion in continuous calculus, several researchers (e.g., [12] and [19]) have defined the concept of discrete convexity to explain the convexity of a subset in $\mathbb{Z} \times \mathbb{Z} \times \ldots \times \mathbb{Z}$. Our notion of convexity describes a much more general mathematical structure in the product of time scales. Theorem 15 provides a necessary and sufficient condition for optimality which enables us to handle different type of optimization problems on time scales. It may open an avenue for future research on time scales with analogues results of Lagrange multipliers and Karush-Kuhn-Tucker conditions.

REFERENCES

[1] M. Adıvar and Y. N. Raffoul, *Existence of resolvent for Volterra integral equations on time scales*, Bull. of Aust. Math. Soc., **82** (2010), 139–155.

[2] M. Adıvar and Y. N. Raffoul, *Stability and periodicity in dynamic delay equations*, Computers and Mathematics with Applications, **58** (2009), 264–272.

[3] M. Adıvar and Y. N. Raffoul, *Existence results for periodic solutions of integro-dynamic equations on time scales*, Annali di Matematica Pura ed Applicata (4), **188** (2009), 543–559.

[4] M. Adıvar and E. A. Bohner, *Halanay type inequalities on time scales with applications*, Nonlinear Analysis: Theory, Methods and Applications, **74** (2011), 7519–7531.

[5] D. R. Anderson, R. J. Krueger and A. C. Peterson, *Delay dynamic equations with stability*, Advances in Difference Equations, **2006**, Article ID 94051, 1–19.

[6] F. M. Atici, D. C. Biles and A. Lebedinsky, *An application of time scales to economics*, Mathematical and Computer Modelling, **43** (2006), 718–726.

[7] M. S. Bazaraa, H. D. Sherali and C. M. Shetty, "Nonlinear Programming Theory and Algorithms," 3^{rd} edition, Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2006.

[8] M. Bohner and A. C. Peterson, "Dynamic Equations on Time Scales. An Introduction with Applications," Birkhäuser Boston, Inc., Boston, MA, 2001.

[9] M. Bohner and A. C. Peterson, "Advances in Dynamic Equations on Time Scales," Birkhäuser Boston, Inc., Boston, MA, 2003.

[10] M. Bohner and R. P. Agarwal, *Oscillation and boundedness of solutions to first and second order forced dynamic equations with mixed nonlinearities*, Australian Journal of Mathematical Analysis and Applications, **5** (2008), 12 pp.

[11] M. Bohner, L. Erbe and A. C. Peterson, *Oscillation for nonlinear second order dynamic equations on time scales*, J. Math. Anal. Appl., **301** (2005), 491–507.

[12] V. I. Danilov and G. A. Koshevoĭ, *Discrete convexity*, Journal of Mathematical Sciences, **133** (2006), 1418–1421.

[13] C. Dinu, *Convex functions on time scales*, Annals of the University of Craiova Ser. Math. Inform., **35** (2008), 87–96.

[14] Y. Gong and X. Xiang, *A class of optimal control problems of systems governed by the first order linear dynamic equations on time scales*, Journal of Industrial and Management Optimization, **5** (2009), 1–10.

[15] S. Hilger, *Analysis on measure chains—a unified approach to continuous and discrete calculus*, Results Math., **18** (1990), 18–56.

[16] R. Hilscher, "Optimality Conditions for Time Scale Variational Problems," DSc dissertation, Masaryk University, Brno, 2008.

[17] R. Hilscher and V. Zeidan, *Weak maximum principle and accessory problem for control problems on time scales*, Nonlinear Analysis, **70** (2009), 3209–3226.

[18] T. Kulik and C. C. Tisdell, *Volterra integral equations on time scales: Basic qualitative and quantitative results with applications to initial value problems on unbounded domains*, Int. J. Difference Equ., **3** (2008), 103–133.

[19] K. Murota, *Discrete convex analysis*, Mathematical Programming, **83** (1998), 313–371.

[20] R. T. Rockafellar, "Convex Analysis," Reprint of the 1970 original, Princeton Landmarks in Mathematics, Princeton Paperbacks, Princeton University Press, Princeton, NJ, 1997.

[21] M. Z. Sarıkaya, N. Aktan, H. Yıldırım, and K. İlarslan, *Partial Δ-differentiation for multivariable functions on n-dimensional time scales*, Journal of Mathematical Inequalities, **3** (2009), 277–291.

[22] M. Z. Sarıkaya, N. Aktan, H. Yıldırım, and K. İlarslan, *Directional ∇-derivative and curves on n-dimensional time scales*, Acta. Appl. Math, **105** (2009), 45–63.

[23] C. C. Tisdell and A. Zaidi, *Basic qualitative and quantitative results for solutions to nonlinear, dynamic equations on time scales with an application to economic modelling*, Nonlinear Anal., **68** (2008), 3504–3524.

[24] R. Oberste-Vorth, *The Fell topology for dynamic equations on time scales*, Nonlinear Dyn. Syst. Theory, **9** (2009), 407–414.

Received April 2011; 1st revision August 2011; final revision August 2011.

E-mail address: murat.adivar@ieu.edu.tr
E-mail address: fang@ncsu.edu