International Journal of Difference Equations
ISSN 0973-6069, Volume 12, Number 2, pp. 161–167 (2017)
http://campus.mst.edu/ijde

Discrete Convexity and its Application to Convex Optimization on Discrete Time Scales

Aykut Arslan
Western Kentucky University
Department of Mathematics
Bowling Green, 42101, USA
aykut.arslan858@topper.wku.edu

Abstract

In this paper, we discuss convexity on n-dimensional discrete time scales $\mathbb{T} = \mathbb{T}_1 \times \mathbb{T}_2 \times \cdots \times \mathbb{T}_n$ where $\mathbb{T}_i \subset \mathbb{R}$, $i = 1, 2, \ldots, n$ are discrete time scales where the time points are not necessarily uniformly distributed on a time line. We introduce the discrete analogues of the fundamental concepts of real convex optimization such as convexity of a function, subgradients, and the Karush–Kuhn–Tucker conditions. In the application section we illustrate our result in an example.

AMS Subject Classifications: 52A41, 39A12, 39A70, 26E70.
Keywords: Convex programs, discrete version of topics in analysis, time scales calculus.

1 Introduction

Convex optimization, a branch of mathematical optimization theory, has been developed in two directions. The real convex optimization and the discrete (or combinatorial) convex optimization. Recent developments such as interior point methods, semidefinite programming and robust optimization in convex optimization theory have stimulated new interest by mathematicians and other scientists. It is applied in areas such as automatic control systems, mathematical economics, electronic circuit design and medical imaging etc. [3, 6, 12]. Other applications can be found in combinatorial optimization and global optimization where it has been used to find bounds on the optimal value or to find approximate solutions [7].

On the other hand the discrete convex optimization combines ideas from real convex optimization and combinatorial optimization to provide optimization techniques for

Received August 30, 2016; Accepted April 7, 2017
Communicated by Christopher C. Tisdell

discrete functions with the convexity property. It was first developed for integer valued functions defined on integer lattice points. In [8, 10] the discrete convexity concepts are introduced for real valued functions defined on $\mathbb{Z}^n$. Mozyrska and Torres introduced the convexity of a function defined on a time scale (a nonempty closed subset of R) in their paper [9]. More recently, Adıvar and Fang defined convexity on the product of time scales [1, 2].

Motivated by these pioneers' work, we give a different definition of discrete convexity for functions on domains which are in the product form $\mathbb{T} = \mathbb{T}_1 \times \mathbb{T}_2 \times \cdots \times \mathbb{T}_n$, where $\mathbb{T}_i \subset \mathbb{R}$, $i = 1, 2, \ldots, n$ are discrete time scales where the time points are not necessarily uniformly distributed on a time line (i.e., the graininess is not necessarily constant). This definition and the one given in [1, 2] is compared in Remark 3.5.

There are some advantages of using our definition of discrete convexity. One of the advantage occurs when the objective function and constraint functions are discrete convex but not real convex. In this case one cannot apply real convex optimization methods, but can apply discrete convex optimization instead.

The structure of the paper is as follows: In Section 2, we state the convex optimization problem on various domains and define the Lagrangian function. Next, we state the definition of real convex functions using subgradients. We then give the definitions of partial nabla and partial delta derivatives in n-dimensional time scales. In Section 3, we define discrete convex functions. In Section 4, we modify the Karush–Kuhn–Tucker conditions for the discrete setting and we prove that a saddle point of the Lagrangian gives a solution to the discrete optimization problem.

2 Preliminaries

An optimization problem, or mathematical programming problem, is minimizing the objective function under the given constraints.

$$\text{minimize } f(x) \text{ subject to } g_i(x) \leq 0, x \in X \text{ for } i = 1, 2, \ldots, m. \tag{2.1}$$

Here X could be any of the following sets; $X = \mathbb{R}^n$, $X = \{x \in \mathbb{R}^n | x \geq 0\}$, $X = \mathbb{Z}^n$, or $X = \mathbb{T}_1 \times \mathbb{T}_2 \times \cdots \times \mathbb{T}_n$, where $\mathbb{T}_i \subset \mathbb{R}$, $i = 1, 2, \ldots, n$ are discrete time scales where the time points are not necessarily uniformly distributed on a time line.

Definition 2.1. An optimization problem is called a convex optimization problem or a convex programming problem if f and g_i are real convex functions for $i = 1, 2, \ldots, m$ and $X = \mathbb{R}^n$.

$$\text{minimize } f(x) \text{ subject to } g_i(x) \leq 0, x \in X \text{ for } i = 1, 2, \ldots, m. \tag{2.2}$$

The Lagrangian function corresponding to the objective function $f(x)$ is defined as

$$L(x, u) = f(x) + u^T g(x),$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $g(x) = (g_1(x), g_2(x), \ldots, g_m(x))$.

Definition 2.2. h is called a subgradient for f at $x^0 \in \mathbb{R}^n$ if it satisfies the following inequality

$$f(x) \geq f(x^0) + \langle x - x^0, h \rangle \quad \text{for all} \quad x \in \text{Dom}(f),$$

where $\langle \cdot, \cdot \rangle$ is the dot product.

A function defined on $\mathbb{R}^n$ is called a real convex function if it has a subgradient at each point of its domain. We define discrete convexity in the next section using an analogue of the subgradient property of real convex functions and some techniques from time scales calculus such as partial delta and partial nabla derivatives. Partial delta and partial nabla derivatives are introduced in [4]. For further reading on time scales, we refer the reader to an excellent book on the analysis of time scales [5]. Let μ_i, and ν_i be the graininess functions on $\mathbb{T}_i$ and e_i be the i^{th} basis element of the n-dimensional Euclidean space. The partial delta and nabla derivatives are defined as

$$\Delta_i f(x^0) := \frac{f(x^0 + e_i\mu_i(x^0)) - f(x^0)}{\mu_i(x^0)}, \quad \nabla_i f(x^0) := \frac{f(x^0) - f(x^0 - e_i\nu_i(x^0))}{\nu_i(x^0)}.$$

3 Discrete Convex Functions

Definition 3.1. Let $\mathbb{T} = \mathbb{T}_1 \times \mathbb{T}_2 \times \cdots \times \mathbb{T}_n$, where $\mathbb{T}_i \subset \mathbb{R}$ is a discrete time scale. A function $f : \mathbb{T} \to \mathbb{R}$ is called discrete convex if given any point $a = (a_1, a_2, \ldots, a_n) \in \mathbb{T}$, we have

$$f(x) \geq f(a) + \langle x - a, \nabla^D f(x, a) \rangle \quad \text{for all} \quad x \in \mathbb{T},$$

$$\nabla^D f(x, a) := (f_{x_1}(x, a), f_{x_2}(x, a), \ldots, f_{x_n}(x, a)),$$

$$f_{x_i}(x, a) := \begin{cases} \Delta_i f(a), & \text{if } x_i \geq a_i \\ \nabla_i f(a), & \text{if } x_i \leq a_i. \end{cases}$$

Note that the discrete gradient vector of a function, $(\nabla^D f)(x, y)$, is a function of two vectors, x and y. The definition depends on the difference of the components of these two vectors.

Theorem 3.2. *Any finite sum of discrete convex function is also discrete convex.*

Remark 3.3. Note that discrete convexity is not necessarily a weaker structure than real convexity. In other words real convexity does not imply discrete convexity. For instance, $f(x, y) = 25(2y - x)^2 + 1/4(2 - x)^2$ is real convex however one can show that it does not satisfy the discrete convexity condition.

Remark 3.4. On the other hand, there is a discrete convex function which is not real convex. To construct such a function we assume that the domain of the function is bounded by an interval of length $M > 0$. From the definition of discrete and real convexity one can obtain $f(x, y) = (x + y)^2 - kx^2$ is discrete convex if and only if

$k \leq \dfrac{M^2+M+1}{M(M-1)}$, $f(x,y) = (x+y)^2 - kx^2$ is real convex if and only if $k \leq \dfrac{(M+1)^2}{M^2}$. Therefore, for values of $k \in \left(\dfrac{M^2+M+1}{M(M-1)}, \dfrac{(M+1)^2}{M^2}\right)$, $f(x,y)$ is a discrete convex function, but not real convex.

Remark 3.5. Adıvar and Fang [1, 2] defined the discrete convex function on $\mathbb{T} = \mathbb{T}_1 \times \mathbb{T}_2 \times \cdots \times \mathbb{T}_n$, where $\mathbb{T}_i \subset \mathbb{R}$, $i = 1, 2, \ldots, n$ are time scales, as a function whose epigraph is convex. Therefore the discrete restriction $f|_{\mathbb{Z}^n}$ of a convex function f on the real domain is convex on $\mathbb{Z}^n$. Conversely, every convex function on a discrete domain can be extended to a convex function on the real domain. However, the discrete convexity in the sense of this paper is not weaker than convexity on the real domain as pointed out in the two abovementioned remarks. Nonetheless, these two definitions match in $\mathbb{T} \subset \mathbb{R}$, a special time scale where the time points are not necessarily uniformly distributed on a time line.

4 Karush–Kuhn–Tucker Conditions on Discrete Time Scales

Definition 4.1. A discrete convex programming problem is an optimization problem with f and g_i are discrete convex functions for $i = 1, 2, \ldots, m$ and $\mathbb{T} = \mathbb{T}_1 \times \mathbb{T}_2 \times \cdots \times \mathbb{T}_n$.

$$\text{minimize } f(x) \text{ subject to } g_i(x) \leq 0, x \in \mathbb{T} \text{ for } i = 1, 2, \ldots, m. \tag{4.1}$$

The set $S = \{x \in \mathbb{T} | g_i(x) \leq 0 \text{ for } i = 1, 2, \ldots, m\}$ is called the feasible set. The Lagrangian associated with this programming problem is a function $L : \mathbb{T} \times \mathbb{R}^m \to \mathbb{R}$ defined as

$$L(x, u) = f(x) + u_1 g_1(x) + \cdots + u_m g_m(x). \tag{4.2}$$

Definition 4.2. A point $(x^0, u^0) \in \mathbb{T} \times \mathbb{R}^m$ is called a saddle point of L if $x^0 \geq 0, u^0 \geq 0$ and $L(x^0, u) \leq L(x^0, u^0) \leq L(x, u^0)$ for all $x \geq 0, u \geq 0$ and $x \in \mathbb{T}$.

Theorem 4.3. *Let (x^0, u^0) be a saddle point of the Lagrangian function L. Then x^0 is a solution to the convex programming problem and $f(x^0) = L(x^0, u^0)$.*

Proof. The condition $L(x^0, u) \leq L(x^0, u^0)$ yields

$$u_1 g_1(x^0) + \ldots + u_m g_m(x^0) \leq u_1^0 g_1(x^0) + \ldots + u_m^0 g_m(x^0).$$

By keeping $u_2, \ldots, u_m$ fixed and taking the limit $u_1 \to \infty$, we infer that $g_1(x^0) \leq 0$. Similarly, one gets $g_2(x^0) \leq 0, \ldots, g_m(x^0) \leq 0$. Thus x^0 belongs to the feasible set S. From $L(x^0, 0) \leq L(x^0, u^0)$ and the definition of S we infer $0 \leq u_1^0 g_1(x^0) + \ldots + u_m^0 g_m(x^0) \leq 0$, hence $u_1^0 g_1(x^0) + \ldots + u_m^0 g_m(x^0) = 0$ and $f(x^0) = L(x^0, u^0)$. Since

$L(x^0, u^0) \leq L(x, u^0)$ for all $x \geq 0$ this implies $f(x^0) \leq f(x) + u_1^0 g_1(x) + \ldots + u_m^0 g_m(x)$. We also have $f(x) + u_1^0 g_1(x) + \ldots + u_m^0 g_m(x) \leq f(x)$ for all $x \in S$.

If we combine the last two inequalities we get $f(x^0) \leq f(x)$ for all x in the feasible set S. Therefore x^0 is a solution to the convex programming problem (4.1). □

Theorem 4.4. *Suppose $f, g_1, \ldots, g_{m-1}$, and g_m are discrete convex functions on $\mathbb{T} = \mathbb{T}_1 \times \mathbb{T}_2 \times \cdots \times \mathbb{T}_n$. Then (x^0, u^0) is a saddle point of the Lagrangian L if and only if*

$$x^0 \geq 0$$
$$\Delta_{x_i} L(x^0, u^0) \geq 0 \textit{ if } x_i^0 = 0$$
$$\Delta_{x_i} L(x^0, u^0) \geq 0,\ \nabla_{x_i} L(x^0, u^0) \leq 0 \textit{ if } x_i^0 > 0$$

and

$$u^0 \geq 0$$
$$\frac{\partial L}{\partial u_j}(x^0, u^0) = g_j(x^0) \leq 0 \textit{ if } u_j^0 = 0$$
$$\frac{\partial L}{\partial u_j}(x^0, u^0) = g_j(x^0) = 0 \textit{ whenever } u_j^0 > 0.$$

Proof. If (x^0, u^0) is a saddle point of L, then clearly we have $x^0, u^0 \geq 0$. If $x_i^0 = 0$, then $\Delta_{x_i} L(x^0, u^0) = \dfrac{L(x^0 + e_i \mu_i(x^0), u^0) - L(x^0, u^0)}{\mu_i(x^0)} \geq 0$ since (x^0, u^0) is saddle point. If $x^0 > 0$, then $\Delta_{x_i} L(x^0, u^0) \geq 0$ and $\nabla_{x_i} L(x^0, u^0) \leq 0$ since $L(x^0, u^0) \leq L(x, u^0)$ for all x. If $u_j^0 = 0$, then $L(x^0, u^0 + te_j) \leq L(x^0, u^0)$ for all $t \geq -u_j^0$. Therefore, $\dfrac{\partial L}{\partial u_j}(x^0, u^0) = \lim\limits_{x \to 0^+} \dfrac{L(x^0, u^0 + te_j) - L(x^0, u^0)}{t} \leq 0$ If $u_j^0 > 0$, then $\dfrac{\partial L}{\partial u_j}(x^0, u^0) = 0$ since (x^0, u^0) is a saddle point. Suppose the conditions in the theorem are satisfied. Since f and g_i are discrete convex functions on $\mathbb{T}$, for a fixed u^0, $L(x, u^0)$ is a discrete convex function too. By convexity of $L(x, u^0)$ we have

$$L(x, u^0) \geq L(x^0, u^0) + \langle (x - x^0), (\nabla^D L)(x, x^0, u^0) \rangle.$$

By the conditions on x and using the definition of discrete gradient we obtain

$$\langle (x - x^0), (\nabla^D L)(x, x^0, u^0) \rangle \geq 0.$$

Therefore, we have $L(x, u^0) \geq L(x^0, u^0)$ for all x.

To show the other side of the inequality, we consider $L(x^0, u)$ as a linear function in

$\mathbb{R}^m$ on variables $u_1, \dots, u_m$. Since it is a linear function on u-coordinates, we have

$$\begin{aligned} L(x^0, u) &= L(x^0, u^0) + \sum_{j=1}^{m} (u_j - u_j^0) \frac{\partial L}{\partial u_j}(x^0, u^0) \\ &\leq L(x^0, u^0) + \sum_{j=1}^{m} u_j \frac{\partial L}{\partial u_j}(x^0, u^0) \\ &\leq L(x^0, u^0). \end{aligned}$$

Hence we have $L(x^0, u) \leq L(x^0, u^0) \leq L(x, u^0)$ for all $x, u \geq 0$. This concludes that (x^0, u^0) is a saddle point. □

5 Application

In this section, we demonstrate our theory on the nonlinear programming problem

$$\begin{aligned} z^* = \min_{x,y} f(x_1, x_2) &= 6(x_1 - 10)^2 + 4(x_2 - 12.5)^2 \\ \text{subject to} \qquad & x_1^2 + (x_2 - 5)^2 \leq 50 \\ & x_1^2 + 3x_2^2 \leq 200 \\ & (x_1 - 6)^2 + x_2^2 \leq 37 \\ & x_i \in \mathbb{Z}^{\geq 0} \;\; \text{for} \;\; i = 1, 2. \end{aligned}$$

Since both $6(x_1 - 10)^2$ and $4(x_2 - 12.5)^2$ are discrete convex, $f(x, y)$ is discrete convex too. For this problem, the Lagrangian is

$$\begin{aligned} L(x_1, x_2, u_1, u_2) =& 6(x_1 - 10)^2 + 4(x_2 - 12.5)^2 + u_1(x_1^2 + (x_2 - 5)^2 - 50) \\ &+ u_2(x_1^2 + 3x_2^2 - 200) + u_3((x_1 - 6)^2 + x_2^2 - 37). \end{aligned}$$

By KKT conditions, we have $u_1(x_1^2 + (x_2 - 5)^2 - 50) = 0$,

$$u_2(x_1^2 + 3x_2^2 - 200) = 0 \quad \text{and} \quad u_3((x_1 - 6)^2 + x_2^2 - 37) = 0.$$

Clearly, we have $x_1 \geq 0$ and $x_2 \geq 0$. From Theorem 4.4, we deduce $\Delta_{x_i} L \geq 0$ and $\nabla_{x_i} L \leq 0$ for $i = 1, 2$. If we combine all these conditions one can reach the optimal solution $(x_1^*, x_2^*) = (7, 6)$. Note that here (u_1, u_2, u_3) are not necessarily unique since the KKT conditions in Theorem 4.4 involves inequalities. Yet, one can choose $u_2 = 0$ and $u_1 = 2, u_3 = 14$ values to justify the above inequalities.

Acknowledgements

We thank the referees for their careful review and constructive comments.

References

[1] M. Adıvar, S. C. Fang, *Convex optimization on mixed domains* J. Ind. Manag. Optim., **8** (1) (2012), 189–227.

[2] M. Adıvar, S. C. Fang, *Convex Analysis and Duality over Discrete Domains* Journal of the Operations Research Society of China, (2017), 1–59.

[3] A. Ben–Tal , A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, SIAM, Philadelphia (2001).

[4] M. Bohner and G. Sh. Guseinov, *Partial differentiation on time scales*, Dynam. Systems Appl., **13**, (3–4), (2004), 35–379.

[5] M. Bohner and A. Peterson, *Dynamic Equations on Time Scales*, Birkhäuser, Boston (2001).

[6] S. Boyd, C. Crusius, and A. Hansson, *Control applications of nonlinear convex programming*, Journal of Process Control, **8**, (5–6), (1998), 313–324.

[7] J. Cochrane and J. Saa–Requejo, *Beyond Arbitrage: Good–Deal Asset Price Bounds in Incomplete Markets*, Journal of Political Economy, **108**, (1), (2000), 79–119.

[8] B.L. Miller, *On minimizing nonseparable functions defined on the integers with an inventory application*, SIAM J. Appl. Math., **21**, (1971), 166–185.

[9] D. Mozyrska and D. F. M. Torres, *The Natural Logarithm on Time Scales*, J. Dyn. Syst. Geom. Theor., **7**, (2009), 41–48.

[10] K. Murota , *Discrete Convex Analysis*, *SIAM Monographs on Discrete Mathematics and Applications*, **10**, Philadelphia (2003).

[11] C. Niculescu, L.E. Persson ,*Convex Functions and Their Applications*, *Springer–Verlag*, New York (2006).

[12] A. Tamura, *Applications of Discrete Convex Analysis to Mathematical Economics*, Publ. RIMS, Kyoto Univ., **40**, (2004), 1015–1037.