FORMALIZED MATHEMATICS
Volume 11, Number 4, 2003
University of Białystok

Definition of Convex Function and Jensen's Inequality

Grigory E. Ivanov
Moscow Institute for Physics and Technology

Summary. Convexity of a function in a real linear space is defined as convexity of its epigraph according to "Convex analysis" [24]. The epigraph of a function is a subset of the product of the function's domain space and the space of real numbers. Therefore, the product of two real linear spaces should be defined. The values of the functions under consideration are extended real numbers. We define the sum of a finite sequence of extended real numbers and get some properties of the sum. The relation between notions "function is convex" and "function is convex on set" (see definition 13 in [21]) is established. We obtain another version of the criterion for a set to be convex (see theorem 6 in [15] to compare) that may be more suitable in some cases. Finally, we prove Jensen's inequality (both strict and not strict) as criteria for functions to be convex.

MML Identifier: `CONVFUN1`.

The terminology and notation used here are introduced in the following articles: [27], [30], [25], [8], [18], [9], [3], [29], [14], [4], [31], [11], [6], [7], [19], [26], [22], [16], [5], [10], [21], [17], [2], [12], [28], [13], [1], [20], and [23].

1. Product of Two Real Linear Spaces

Let X, Y be non empty RLS structures. The functor AddInProdRLS(X,Y) yielding a binary operation on $[\!:$ the carrier of X, the carrier of $Y\,:\!]$ is defined by the condition (Def. 1).

(Def. 1) Let z_1, z_2 be elements of $[\!:$ the carrier of X, the carrier of $Y\,:\!]$, x_1, x_2 be vectors of X, and y_1, y_2 be vectors of Y. Suppose $z_1 = \langle x_1, y_1\rangle$ and $z_2 = \langle x_2, y_2\rangle$. Then (AddInProdRLS$(X,Y))(z_1, z_2) = \langle$(the addition of $X)(\langle x_1, x_2\rangle)$, (the addition of $Y)(\langle y_1, y_2\rangle)\rangle$.

ISSN 1426–2630

Let X, Y be non empty RLS structures. The functor MultInProdRLS(X,Y) yields a function from $[:\mathbb{R}, [:\text{the carrier of } X, \text{ the carrier of } Y:]:]$ into $[:\text{the carrier of } X, \text{ the carrier of } Y:]$ and is defined by the condition (Def. 2).

(Def. 2) Let a be a real number, z be an element of $[:\text{the carrier of } X, \text{ the carrier of } Y:]$, x be a vector of X, and y be a vector of Y. Suppose $z = \langle x, y\rangle$. Then $(\text{MultInProdRLS}(X,Y))(\langle a, z\rangle) = \langle(\text{the external multiplication of } X)(\langle a, x\rangle), (\text{the external multiplication of } Y)(\langle a, y\rangle)\rangle$.

Let X, Y be non empty RLS structures. The functor ProdRLS(X,Y) yields an RLS structure and is defined by:

(Def. 3) $\text{ProdRLS}(X,Y) = \langle[:\text{the carrier of } X, \text{ the carrier of } Y:], \langle 0_X, 0_Y\rangle, \text{AddInProdRLS}(X,Y), \text{MultInProdRLS}(X,Y)\rangle$.

Let X, Y be non empty RLS structures. Note that ProdRLS(X,Y) is non empty.

Next we state two propositions:

(1) Let X, Y be non empty RLS structures, x be a vector of X, y be a vector of Y, u be a vector of ProdRLS(X,Y), and p be a real number. If $u = \langle x, y\rangle$, then $p \cdot u = \langle p \cdot x, p \cdot y\rangle$.

(2) Let X, Y be non empty RLS structures, x_1, x_2 be vectors of X, y_1, y_2 be vectors of Y, and u_1, u_2 be vectors of ProdRLS(X,Y). If $u_1 = \langle x_1, y_1\rangle$ and $u_2 = \langle x_2, y_2\rangle$, then $u_1 + u_2 = \langle x_1 + x_2, y_1 + y_2\rangle$.

Let X, Y be Abelian non empty RLS structures. One can verify that ProdRLS(X,Y) is Abelian.

Let X, Y be add-associative non empty RLS structures. Observe that ProdRLS(X,Y) is add-associative.

Let X, Y be right zeroed non empty RLS structures. Observe that ProdRLS(X,Y) is right zeroed.

Let X, Y be right complementable non empty RLS structures. One can check that ProdRLS(X,Y) is right complementable.

Let X, Y be real linear space-like non empty RLS structures. Observe that ProdRLS(X,Y) is real linear space-like.

Next we state the proposition

(3) Let X, Y be real linear spaces, n be a natural number, x be a finite sequence of elements of the carrier of X, y be a finite sequence of elements of the carrier of Y, and z be a finite sequence of elements of the carrier of ProdRLS(X,Y). Suppose $\text{len}\, x = n$ and $\text{len}\, y = n$ and $\text{len}\, z = n$ and for every natural number i such that $i \in \text{Seg}\, n$ holds $z(i) = \langle x(i), y(i)\rangle$. Then $\sum z = \langle \sum x, \sum y\rangle$.

2. Real Linear Space of Real Numbers

The non empty RLS structure $\mathbb{R}_{\rm RLS}$ is defined as follows:

(Def. 4) $\mathbb{R}_{\rm RLS} = \langle \mathbb{R}, 0, +_{\mathbb{R}}, \cdot_{\mathbb{R}} \rangle$.

Let us note that $\mathbb{R}_{\rm RLS}$ is Abelian, add-associative, right zeroed, right complementable, and real linear space-like.

3. Sum of Finite Sequence of Extended Real Numbers

Let F be a finite sequence of elements of $\overline{\mathbb{R}}$. The functor $\sum F$ yields an extended real number and is defined by the condition (Def. 5).

(Def. 5) There exists a function f from $\mathbb{N}$ into $\overline{\mathbb{R}}$ such that $\sum F = f(\operatorname{len} F)$ and $f(0) = 0_{\overline{\mathbb{R}}}$ and for every natural number i such that $i < \operatorname{len} F$ holds $f(i+1) = f(i) + F(i+1)$.

We now state several propositions:

(4) $\sum(\varepsilon_{\overline{\mathbb{R}}}) = 0_{\overline{\mathbb{R}}}$.

(5) For every extended real number a holds $\sum\langle a \rangle = a$.

(6) For all extended real numbers a, b holds $\sum\langle a, b\rangle = a + b$.

(7) For all finite sequences F, G of elements of $\overline{\mathbb{R}}$ such that $-\infty \notin \operatorname{rng} F$ and $-\infty \notin \operatorname{rng} G$ holds $\sum(F \frown G) = \sum F + \sum G$.

(8) Let F, G be finite sequences of elements of $\overline{\mathbb{R}}$ and s be a permutation of $\operatorname{dom} F$. If $G = F \cdot s$ and $-\infty \notin \operatorname{rng} F$, then $\sum F = \sum G$.

4. Definition of Convex Function

Let X be a non empty RLS structure and let f be a function from the carrier of X into $\overline{\mathbb{R}}$. The functor epigraph f yielding a subset of $\operatorname{ProdRLS}(X, \mathbb{R}_{\rm RLS})$ is defined as follows:

(Def. 6) epigraph $f = \{\langle x, y \rangle; x$ ranges over elements of X, y ranges over elements of $\mathbb{R}$: $f(x) \leqslant \overline{\mathbb{R}}(y)\}$.

Let X be a non empty RLS structure and let f be a function from the carrier of X into $\overline{\mathbb{R}}$. We say that f is convex if and only if:

(Def. 7) epigraph f is convex.

The following two propositions are true:

(9) Let X be a non empty RLS structure and f be a function from the carrier of X into $\overline{\mathbb{R}}$. Suppose that for every vector x of X holds $f(x) \neq -\infty$. Then f is convex if and only if for all vectors x_1, x_2 of X and for every real number p such that $0 < p$ and $p < 1$ holds $f(p \cdot x_1 + (1-p) \cdot x_2) \leqslant \overline{\mathbb{R}}(p) \cdot f(x_1) + \overline{\mathbb{R}}(1-p) \cdot f(x_2)$.

(10) Let X be a real linear space and f be a function from the carrier of X into $\overline{\mathbb{R}}$. Suppose that for every vector x of X holds $f(x) \neq -\infty$. Then f is convex if and only if for all vectors x_1, x_2 of X and for every real number p such that $0 \leqslant p$ and $p \leqslant 1$ holds $f(p \cdot x_1 + (1 - p) \cdot x_2) \leqslant \overline{\mathbb{R}}(p) \cdot f(x_1) + \overline{\mathbb{R}}(1 - p) \cdot f(x_2)$.

5. Relation between Notions "function is convex" and "function is convex on set"

We now state the proposition

(11) Let f be a partial function from $\mathbb{R}$ to $\mathbb{R}$, g be a function from the carrier of $\mathbb{R}_{\rm RLS}$ into $\overline{\mathbb{R}}$, and X be a subset of $\mathbb{R}_{\rm RLS}$. Suppose $X \subseteq \operatorname{dom} f$ and for every real number x holds if $x \in X$, then $g(x) = f(x)$ and if $x \notin X$, then $g(x) = +\infty$. Then g is convex if and only if the following conditions are satisfied:

(i) f is convex on X, and

(ii) X is convex.

6. Theorem 6 from [15] in Other Words

One can prove the following proposition

(12) Let X be a real linear space and M be a subset of X. Then M is convex if and only if for every non empty natural number n and for every finite sequence p of elements of $\mathbb{R}$ and for all finite sequences y, z of elements of the carrier of X such that $\operatorname{len} p = n$ and $\operatorname{len} y = n$ and $\operatorname{len} z = n$ and $\sum p = 1$ and for every natural number i such that $i \in \operatorname{Seg} n$ holds $p(i) > 0$ and $z(i) = p(i) \cdot y_i$ and $y_i \in M$ holds $\sum z \in M$.

7. Jensen's Inequality

One can prove the following two propositions:

(13) Let X be a real linear space and f be a function from the carrier of X into $\overline{\mathbb{R}}$. Suppose that for every vector x of X holds $f(x) \neq -\infty$. Then f is convex if and only if for every non empty natural number n and for every finite sequence p of elements of $\mathbb{R}$ and for every finite sequence F of elements of $\overline{\mathbb{R}}$ and for all finite sequences y, z of elements of the carrier of X such that $\operatorname{len} p = n$ and $\operatorname{len} F = n$ and $\operatorname{len} y = n$ and $\operatorname{len} z = n$ and $\sum p = 1$ and for every natural number i such that $i \in \operatorname{Seg} n$ holds $p(i) > 0$ and $z(i) = p(i) \cdot y_i$ and $F(i) = \overline{\mathbb{R}}(p(i)) \cdot f(y_i)$ holds $f(\sum z) \leqslant \sum F$.

(14) Let X be a real linear space and f be a function from the carrier of X into $\overline{\mathbb{R}}$. Suppose that for every vector x of X holds $f(x) \neq -\infty$. Then f is convex if and only if for every non empty natural number n and for every finite sequence p of elements of $\mathbb{R}$ and for every finite sequence F of elements of $\overline{\mathbb{R}}$ and for all finite sequences y, z of elements of the carrier of X such that $\operatorname{len} p = n$ and $\operatorname{len} F = n$ and $\operatorname{len} y = n$ and $\operatorname{len} z = n$ and $\sum p = 1$ and for every natural number i such that $i \in \operatorname{Seg} n$ holds $p(i) \geqslant 0$ and $z(i) = p(i) \cdot y_i$ and $F(i) = \overline{\mathbb{R}}(p(i)) \cdot f(y_i)$ holds $f(\sum z) \leqslant \sum F$.

Acknowledgments

I thank Andrzej Trybulec for teaching me Mizar.

References

[1] Grzegorz Bancerek. Cardinal numbers. *Formalized Mathematics*, 1(**2**):377–382, 1990.
[2] Grzegorz Bancerek. The fundamental properties of natural numbers. *Formalized Mathematics*, 1(**1**):41–46, 1990.
[3] Grzegorz Bancerek and Krzysztof Hryniewiecki. Segments of natural numbers and finite sequences. *Formalized Mathematics*, 1(**1**):107–114, 1990.
[4] Józef Białas. Infimum and supremum of the set of real numbers. Measure theory. *Formalized Mathematics*, 2(**1**):163–171, 1991.
[5] Józef Białas. Series of positive real numbers. Measure theory. *Formalized Mathematics*, 2(**1**):173–183, 1991.
[6] Józef Białas. Some properties of the intervals. *Formalized Mathematics*, 5(**1**):21–26, 1996.
[7] Czesław Byliński. Binary operations. *Formalized Mathematics*, 1(**1**):175–180, 1990.
[8] Czesław Byliński. Functions and their basic properties. *Formalized Mathematics*, 1(**1**):55–65, 1990.
[9] Czesław Byliński. Functions from a set to a set. *Formalized Mathematics*, 1(**1**):153–164, 1990.
[10] Czesław Byliński. Partial functions. *Formalized Mathematics*, 1(**2**):357–367, 1990.
[11] Czesław Byliński. Some basic properties of sets. *Formalized Mathematics*, 1(**1**):47–53, 1990.
[12] Czesław Byliński. The sum and product of finite sequences of real numbers. *Formalized Mathematics*, 1(**4**):661–668, 1990.
[13] Agata Darmochwał. Finite sets. *Formalized Mathematics*, 1(**1**):165–167, 1990.
[14] Noboru Endou, Takashi Mitsuishi, and Yasunari Shidama. Convex sets and convex combinations. *Formalized Mathematics*, 11(**1**):53–58, 2003.
[15] Noboru Endou, Yasumasa Suzuki, and Yasunari Shidama. Some properties for convex combinations. *Formalized Mathematics*, 11(**3**):267–270, 2003.
[16] Noboru Endou, Katsumi Wasaki, and Yasunari Shidama. Basic properties of extended real numbers. *Formalized Mathematics*, 9(**3**):491–494, 2001.
[17] Noboru Endou, Katsumi Wasaki, and Yasunari Shidama. Definitions and basic properties of measurable functions. *Formalized Mathematics*, 9(**3**):495–500, 2001.
[18] Krzysztof Hryniewiecki. Basic properties of real numbers. *Formalized Mathematics*, 1(**1**):35–40, 1990.
[19] Jarosław Kotowicz. Real sequences and basic operations on them. *Formalized Mathematics*, 1(**2**):269–272, 1990.
[20] Jarosław Kotowicz. Functions and finite sequences of real numbers. *Formalized Mathematics*, 3(**2**):275–278, 1992.
[21] Jarosław Kotowicz and Yuji Sakai. Properties of partial functions from a domain to the set of real numbers. *Formalized Mathematics*, 3(**2**):279–288, 1992.
[22] Eugeniusz Kusak, Wojciech Leończuk, and Michał Muzalewski. Abelian groups, fields and vector spaces. *Formalized Mathematics*, 1(**2**):335–342, 1990.

[23] Takaya Nishiyama and Yasuho Mizuhara. Binary arithmetics. *Formalized Mathematics*, 4(**1**):83–86, 1993.
[24] Tyrrell R. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.
[25] Andrzej Trybulec. Subsets of complex numbers. *To appear in Formalized Mathematics*.
[26] Andrzej Trybulec. Domains and their Cartesian products. *Formalized Mathematics*, 1(**1**):115–122, 1990.
[27] Andrzej Trybulec. Tarski Grothendieck set theory. *Formalized Mathematics*, 1(**1**):9–11, 1990.
[28] Wojciech A. Trybulec. Pigeon hole principle. *Formalized Mathematics*, 1(**3**):575–579, 1990.
[29] Wojciech A. Trybulec. Vectors in real linear space. *Formalized Mathematics*, 1(**2**):291–296, 1990.
[30] Zinaida Trybulec. Properties of subsets. *Formalized Mathematics*, 1(**1**):67–71, 1990.
[31] Edmund Woronowicz. Relations defined on sets. *Formalized Mathematics*, 1(**1**):181–186, 1990.

Received July 17, 2003