JOURNAL OF FORMALIZED MATHEMATICS
Volume 15, Released 2003, Published 2003
Inst. of Computer Science, Univ. 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" by R. Tyrrell Rockafellar. 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 RFUNCT_3:def 13) is established. We obtain another version of the criterion for a set to be convex (see CONVEX2:6 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.

WWW: http://mizar.org/JFM/Vol15/convfun1.html

The articles [24], [28], [25], [8], [17], [9], [3], [26], [14], [4], [29], [11], [6], [7], [18], [23], [21], [15], [5], [10], [20], [16], [2], [12], [27], [13], [1], [19], and [22] provide the notation and terminology for this paper.

1. Product of Two Real Linear Spaces

Let X, Y be non empty RLS structures. The functor AddInProdRLS(X,Y) yields a binary operation on $[:$the carrier of X, the carrier of $Y:]$ and 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 $(\mathrm{AddInProdRLS}(X,Y))(z_1, z_2) = \langle$(the addition of $X)(\langle x_1, x_2 \rangle)$, (the addition of $Y)(\langle y_1, y_2 \rangle)\rangle$.

Let X, Y be non empty RLS structures. The functor MultInProdRLS(X,Y) yields a function from $[:\mathbb{R}, [:$the carrier of X, the carrier of $Y:]:]$ into $[:$the carrier of X, 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 $[:$the carrier of X, the carrier of $Y:]$, x be a vector of X, and y be a vector of Y. Suppose $z = \langle x, y \rangle$. Then $(\mathrm{MultInProdRLS}(X,Y))(\langle a, z \rangle) = \langle$(the external multiplication of $X)(\langle a, x \rangle)$, (the external multiplication of $Y)(\langle a, y \rangle)\rangle$.

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

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

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

The following propositions are true:

(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 $\mathrm{ProdRLS}(X,Y)$, 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 $\mathrm{ProdRLS}(X,Y)$. 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. Observe that $\mathrm{ProdRLS}(X,Y)$ is Abelian.

Let X, Y be add-associative non empty RLS structures. Observe that $\mathrm{ProdRLS}(X,Y)$ is add-associative.

Let X, Y be right zeroed non empty RLS structures. Observe that $\mathrm{ProdRLS}(X,Y)$ is right zeroed.

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

Let X, Y be real linear space-like non empty RLS structures. Note that $\mathrm{ProdRLS}(X,Y)$ is real linear space-like.

The following proposition is true

(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 $\mathrm{ProdRLS}(X,Y)$. Suppose $\mathrm{len}\, x = n$ and $\mathrm{len}\, y = n$ and $\mathrm{len}\, z = n$ and for every natural number i such that $i \in \mathrm{Seg}\, n$ 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 $\mathbb{R}_{\mathrm{RLS}}$ is defined as follows:

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

One can verify that $\mathbb{R}_{\mathrm{RLS}}$ 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 $\sum F$ yielding an extended real number is defined by the condition (Def. 5).

(Def. 5) There exists a function f from $\mathbb{N}$ into $\overline{\mathbb{R}}$ such that $\sum F = f(\mathrm{len}\, F)$ and $f(0) = 0_{\overline{\mathbb{R}}}$ and for every natural number i such that $i < \mathrm{len}\, F$ 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 $-\infty \notin \mathrm{rng}\, F$ and $-\infty \notin \mathrm{rng}\, G$ 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 $\mathrm{dom}\, F$. If $G = F \cdot s$ and $-\infty \notin \mathrm{rng}\, F$, 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 yields a subset of $\mathrm{ProdRLS}(X, \mathbb{R}_{\mathrm{RLS}})$ and is defined by:

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

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.

One can prove the following two propositions:

(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 $f(p \cdot x_1 + (1-p) \cdot x_2) \leq \overline{\mathbb{R}}(p) \cdot f(x_1) + \overline{\mathbb{R}}(1-p) \cdot f(x_2)$.

(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 $0 \leq p$ and $p \leq 1$ holds $f(p \cdot x_1 + (1-p) \cdot x_2) \leq \overline{\mathbb{R}}(p) \cdot f(x_1) + \overline{\mathbb{R}}(1-p) \cdot f(x_2)$.

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

The following proposition is true

(11) Let f be a partial function from $\mathbb{R}$ to $\mathbb{R}$, g be a function from the carrier of $\mathbb{R}_{\mathrm{RLS}}$ into $\overline{\mathbb{R}}$, and X be a subset of $\mathbb{R}_{\mathrm{RLS}}$. 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. CONVEX2:6 in other words

We now state the 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

Next we state 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 $f(\sum z) \leq \sum F$.

(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 $p(i) \geq 0$ and $z(i) = p(i) \cdot y_i$ and $F(i) = \overline{\mathbb{R}}(p(i)) \cdot f(y_i)$ holds $f(\sum z) \leq \sum F$.

ACKNOWLEDGMENTS

I thank Andrzej Trybulec for teaching me Mizar.

REFERENCES

[1] Grzegorz Bancerek. Cardinal numbers. *Journal of Formalized Mathematics*, 1, 1989. `http://mizar.org/JFM/Vol1/card_1.html`.

[2] Grzegorz Bancerek. The fundamental properties of natural numbers. *Journal of Formalized Mathematics*, 1, 1989. `http://mizar.org/JFM/Vol1/nat_1.html`.

[3] Grzegorz Bancerek and Krzysztof Hryniewiecki. Segments of natural numbers and finite sequences. *Journal of Formalized Mathematics*, 1, 1989. `http://mizar.org/JFM/Vol1/finseq_1.html`.

[4] Józef Białas. Infimum and supremum of the set of real numbers. Measure theory. *Journal of Formalized Mathematics*, 2, 1990. `http://mizar.org/JFM/Vol2/supinf_1.html`.

[5] Józef Białas. Series of positive real numbers. Measure theory. *Journal of Formalized Mathematics*, 2, 1990. `http://mizar.org/JFM/Vol2/supinf_2.html`.

[6] Józef Białas. Some properties of the intervals. *Journal of Formalized Mathematics*, 6, 1994. `http://mizar.org/JFM/Vol6/measure6.html`.

[7] Czesław Byliński. Binary operations. *Journal of Formalized Mathematics*, 1, 1989. `http://mizar.org/JFM/Vol1/binop_1.html`.

[8] Czesław Byliński. Functions and their basic properties. *Journal of Formalized Mathematics*, 1, 1989. `http://mizar.org/JFM/Vol1/funct_1.html`.

[9] Czesław Byliński. Functions from a set to a set. *Journal of Formalized Mathematics*, 1, 1989. `http://mizar.org/JFM/Vol1/funct_2.html`.

[10] Czesław Byliński. Partial functions. *Journal of Formalized Mathematics*, 1, 1989. `http://mizar.org/JFM/Vol1/partfun1.html`.

[11] Czesław Byliński. Some basic properties of sets. *Journal of Formalized Mathematics*, 1, 1989. `http://mizar.org/JFM/Vol1/zfmisc_1.html`.

[12] Czesław Byliński. The sum and product of finite sequences of real numbers. *Journal of Formalized Mathematics*, 2, 1990. `http://mizar.org/JFM/Vol2/rvsum_1.html`.

[13] Agata Darmochwał. Finite sets. *Journal of Formalized Mathematics*, 1, 1989. `http://mizar.org/JFM/Vol1/finset_1.html`.

[14] Noboru Endou, Takashi Mitsuishi, and Yasunari Shidama. Convex sets and convex combinations. *Journal of Formalized Mathematics*, 14, 2002. `http://mizar.org/JFM/Vol14/convex1.html`.

[15] Noboru Endou, Katsumi Wasaki, and Yasunari Shidama. Basic properties of extended real numbers. *Journal of Formalized Mathematics*, 12, 2000. `http://mizar.org/JFM/Vol12/extreal1.html`.

[16] Noboru Endou, Katsumi Wasaki, and Yasunari Shidama. Definitions and basic properties of measurable functions. *Journal of Formalized Mathematics*, 12, 2000. `http://mizar.org/JFM/Vol12/mesfunc1.html`.

[17] Krzysztof Hryniewiecki. Basic properties of real numbers. *Journal of Formalized Mathematics*, 1, 1989. `http://mizar.org/JFM/Vol1/real_1.html`.

[18] Jarosław Kotowicz. Real sequences and basic operations on them. *Journal of Formalized Mathematics*, 1, 1989. `http://mizar.org/JFM/Vol1/seq_1.html`.

[19] Jarosław Kotowicz. Functions and finite sequences of real numbers. *Journal of Formalized Mathematics*, 5, 1993. `http://mizar.org/JFM/Vol5/rfinseq.html`.

[20] Jarosław Kotowicz and Yuji Sakai. Properties of partial functions from a domain to the set of real numbers. *Journal of Formalized Mathematics*, 5, 1993. `http://mizar.org/JFM/Vol5/rfunct_3.html`.

[21] Eugeniusz Kusak, Wojciech Leończuk, and Michał Muzalewski. Abelian groups, fields and vector spaces. *Journal of Formalized Mathematics*, 1, 1989. `http://mizar.org/JFM/Vol1/vectsp_1.html`.

[22] Takaya Nishiyama and Yasuho Mizuhara. Binary arithmetics. *Journal of Formalized Mathematics*, 5, 1993. `http://mizar.org/JFM/Vol5/binarith.html`.

[23] Andrzej Trybulec. Domains and their Cartesian products. *Journal of Formalized Mathematics*, 1, 1989. http://mizar.org/JFM/Vol1/domain_1.html.

[24] Andrzej Trybulec. Tarski Grothendieck set theory. *Journal of Formalized Mathematics*, Axiomatics, 1989. http://mizar.org/JFM/Axiomatics/tarski.html.

[25] Andrzej Trybulec. Subsets of real numbers. *Journal of Formalized Mathematics*, Addenda, 2003. http://mizar.org/JFM/Addenda/numbers.html.

[26] Wojciech A. Trybulec. Vectors in real linear space. *Journal of Formalized Mathematics*, 1, 1989. http://mizar.org/JFM/Vol1/rlvect_1.html.

[27] Wojciech A. Trybulec. Pigeon hole principle. *Journal of Formalized Mathematics*, 2, 1990. http://mizar.org/JFM/Vol2/finseq_4.html.

[28] Zinaida Trybulec. Properties of subsets. *Journal of Formalized Mathematics*, 1, 1989. http://mizar.org/JFM/Vol1/subset_1.html.

[29] Edmund Woronowicz. Relations defined on sets. *Journal of Formalized Mathematics*, 1, 1989. http://mizar.org/JFM/Vol1/relset_1.html.

Received July 17, 2003

Published January 2, 2004