CONVEX ANALYSIS

An Introductory Text

JAN VAN TIEL

Convex Analysis

An Introductory Text

Jan van Tiel

Royal Netherlands Meteorological Institute

JOHN WILEY AND SONS

Chichester • New York • Brisbane • Toronto • Singapore

***Library of Congress Cataloging in Publication Data*:**

Tiel, Jan van.
Convex analysis.

Includes bibliographical references and indexes.
1. Convex functions. 2. Convex sets. 3. Convex programming. I. Title.
QA331.5.T49 1984 515.8'3 83–10176
ISBN 0 471 90263 2
ISBN 0 471 90265 9 (pbk.)

***British Library Cataloguing in Publication Data*:**

Tiel, Jan van
Convex analysis.
1. Convex functions
I. Title
515.8'8 QA331.5

ISBN 0 471 90263 2
ISBN 0 471 90265 9 (pbk.)

Filmset and printed in Northern Ireland at the Universities Press (Belfast) Ltd.
Bound at the Pitman Press Ltd., Bath, Avon.

Convex Analysis

Preface

This little book has evolved from my experience in teaching convex analysis at the University of Utrecht, Holland. In theory and applications, convex analysis is of increasing interest at the present time. This book is primarily an introductory text; therefore I have tried to emphasize the basic concepts and the characteristic methods of this part of mathematics (such as separation, subgradient, conjugate function, convex optimization). A large number of elementary exercises at the ends of the various chapters (with answers and hints at the end of the book) are intended to aid in understanding the concepts employed.

The book is intended for the young student who is interested in convexity and whose mathematical background includes the basic facts of calculus, linear algebra, and general topology; it is also supposed that he is acquainted with the basic concepts of functional analysis (such as normed linear space, Hilbert space, dual).

In order to convey the flavour of the subject and to arouse the student's interest, I have not restricted myself to the finite-dimensional case one usually deals with in practice. But to keep things as simple as possible, of the class of locally convex spaces, the 'natural' domain of convex analysis, only normed spaces appear in this book.

Some historical remarks and additional material are collected in bibliographical notes; of course these are by no means exhaustive.

Chapter 1 summarizes the essentials of the theory of real convex functions on the real line. We also consider some generalizations to functions which can have infinite values.

Chapter 2 studies algebraic properties of convex sets in a linear space. In the case of a linear topological space, we find some topological properties of convex sets.

Chapter 3 develops the theory of separation in a linear space. Applying this theory in the case of a linear topological space yields the Hahn–Banach theorem.

Chapter 4 considers some classical theorems concerning convex subsets of $\mathbb{R}^n$ and some applications to polyhedral cones. Using the notion of relative interior, we study separation in $\mathbb{R}^n$.

Chapter 5 studies convex functions on a linear space which can have infinite values. In a certain sense, locally boundedness turns out to be equivalent to continuity. We study the important concepts of lower semi-continuity and subdifferentiability.

Chapter 6 develops the theory of duality. We find characterizations of the bipolar function and of support functions.

Chapter 7 gives an impression of the meaning of convexity in optimization. It deals mainly with convex programming (Kuhn–Tucker conditions, saddle points and Fenchel's duality theorem).

I am indebted to Professor John Horváth who suggested the writing of an English version of my lecture notes. I wish to thank my colleagues Tineke de Bunje and Leen Roozemond who have read all or part of the manuscript and made many improvements. Finally, my thanks go to Mrs. M. M. Meijer who spent many hours typing the manuscript.

Jan van Tiel

Contents

CHAPTER 1

Convex Functions on $\mathbb{R}$

In this chapter we shall designate by I a (closed, open or half-open, finite or infinite) interval in $\mathbb{R}$.

REAL CONVEX FUNCTIONS

1.1 Definitions

Let f be a function $I \to \mathbb{R}$.

(a) f is said to be *convex* if

$$f(\lambda a+(1-\lambda)b) \leqslant \lambda f(a)+(1-\lambda)f(b) \qquad (1)$$

for all a, $b \in I$ and all $\lambda \in \mathbb{R}$ with $0<\lambda<1$. Figure 1 shows the geometrical meaning of convexity: the chord with endpoints $(a, f(a))$ and $(b, f(b))$ lies nowhere underneath the graph of f.

(b) f is said to be *strictly convex* if it is convex and the strict inequality holds in (1) whenever $a \neq b$.

1.2

We give some other, equivalent, formulations of the convexity of $f: I \to \mathbb{R}$:

(a)

$$f(x) \leqslant \frac{b-x}{b-a} f(a)+\frac{x-a}{b-a} f(b)$$

for all a, b, $x \in I$ with $a<x<b$. Note that the right-hand side of this inequality can be written as

$$f(a)+\frac{f(b)-f(a)}{b-a}(x-a).$$

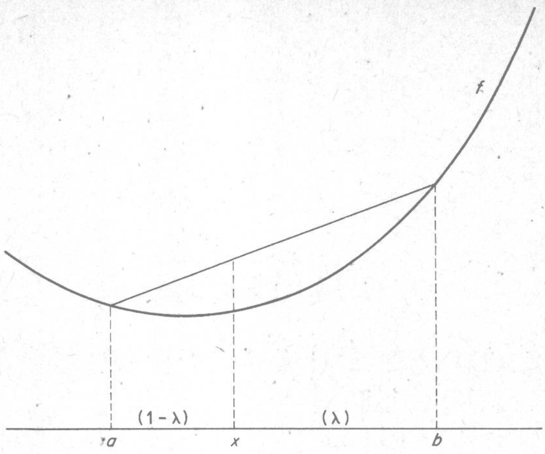

Figure 1

(b)

$$f(\lambda a+\mu b)\leqslant \lambda f(a)+\mu f(b)$$

for all a, $b \in I$ and all λ, $\mu \in \mathbb{R}$ such that $\lambda>0$, $\mu>0$, $\lambda+\mu=1$.

1.3

The proof of the following simple properties is left to the reader.

(a) If f and g are convex functions and $\alpha \geqslant 0$, $\beta \geqslant 0$, then $\alpha f+\beta g$ is convex.

(b) The sum of finitely many convex functions is convex.

(c) The (pointwise) limit of a convergent sequence of convex functions is convex.

(d) Let $f\colon I \to \mathbb{R}$ be convex. Then

$$\sum_{i=1}^{n} \lambda_i x_i \in I \quad \text{and} \quad f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i)$$

whenever $x_i \in I$, $\lambda_i \geqslant 0$ $(1 \leqslant i \leqslant n)$, $\sum_{i=1}^{n} \lambda_i = 1$.

(e) Let f be the pointwise supremum of an arbitrary collection of convex functions $I \to \mathbb{R}$. If f is finite everywhere on I, then f is convex. Does an analogous proposition hold for the infimum?

1.4 Theorem

Let $f\colon I \to \mathbb{R}$ be convex. Then

$$\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(b)-f(a)}{b-a} \leqslant \frac{f(b)-f(x)}{b-x} \tag{2}$$

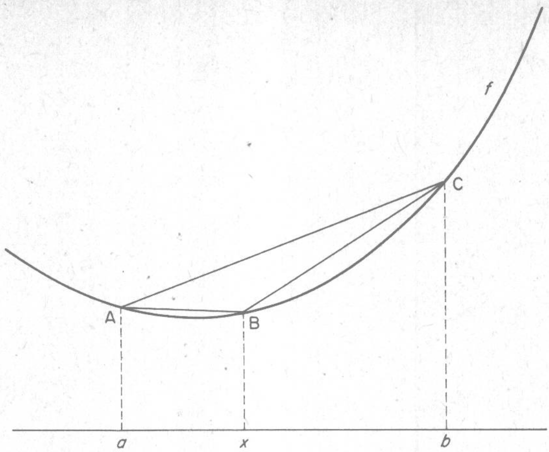

Figure 2

whenever $a, b, x \in I$, $a < x < b$. If f is strictly convex, then the strict inequalities hold in (2).

Figure 2 shows the geometrical meaning of this theorem: slope(AB) $\leqslant$ slope(AC) $\leqslant$ slope(BC).

Proof. Since f is convex, we have

$$f(x) \leqslant \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b). \tag{3}$$

From this inequality we can derive

$$f(x) - f(a) \leqslant \frac{a-x}{b-a} f(a) + \frac{x-a}{b-a} f(b) = \frac{x-a}{b-a} [f(b) - f(a)]$$

which proves the first inequality in (2); the second inequality can be proved in a similar way. If f is strictly convex, then the strict inequality holds in (3) and thus also in (2).

1.5

We denote the interior of I by $\operatorname{int}(I)$. Let $f: I \to \mathbb{R}$ be convex, and let $c \in \operatorname{int}(I)$. Let $[a, b] \subset I$ such that $a < c < b$. By Theorem 1.4, we have

$$\frac{f(c) - f(a)}{c - a} \leqslant \frac{f(x) - f(c)}{x - c}$$

whenever $x \in \langle c, b]$. It also follows from Theorem 1.4 that the function

$$x \mapsto \frac{f(x) - f(c)}{x - c}$$

is non-decreasing on $\langle c, b]$. Hence the right derivative

$$f'_+(c) := \lim_{x \downarrow c} \frac{f(x)-f(c)}{x-c}$$

exists. In a similar way it can be proved that the left derivative $f'_-(c)$ exists.

If $a<c<d<b$, then for sufficiently small positive h we have

$$\frac{f(c)-f(c-h)}{h} \leqslant \frac{f(c+h)-f(c)}{h} \leqslant \frac{f(d)-f(d-h)}{h}.$$

Passing to the limit as $h \downarrow 0$, we obtain

$$f'_-(c) \leqslant f'_+(c) \leqslant f'_-(d).$$

We have thus proved the following theorem.

1.6 Theorem

Let $f: I \to \mathbb{R}$ be convex. Then f has a right derivative and a left derivative at every point of $\text{int}(I)$, and f'_- and f'_+ are non-decreasing on $\text{int}(I)$. If $c \in \text{int}(I)$, we have

$$f'_-(c) \leqslant f'_+(c)$$

and

$$f(x) \geqslant f(c) + f'_-(c)(x-c), \quad f(x) \geqslant f(c) + f'_+(c)(x-c)$$

for all $x \in I$ (cf. Figure 3).

Remark. Let $f: [a, b] \to \mathbb{R}$ be convex. The above proof shows that in this case $f'_+(a)$ and $f'_-(b)$ exist if $+\infty$ and $-\infty$ are allowed as limits.

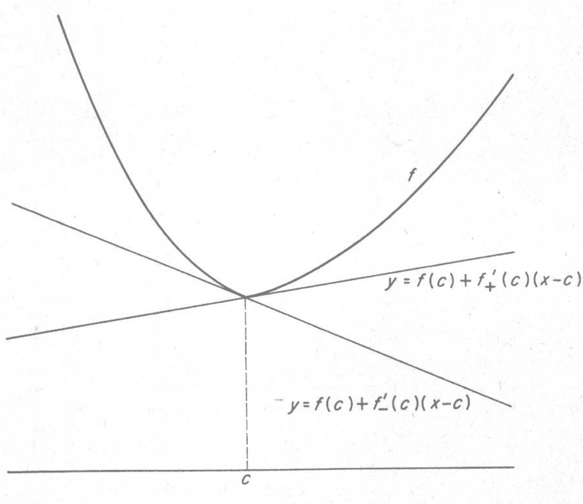

Figure 3

1.7

$f\colon I \to \mathbb{R}$ is called *Lipschitzian* relative to $I_0 \subset I$ if there exists $K > 0$ such that $|f(x) - f(y)| \leqslant K|x - y|$ for all $x, y \in I_0$. This condition implies that f is continuous and even uniformly continuous relative to I_0, and of bounded variation on every closed bounded sub-interval of I_0.

Theorem. Let $f\colon I \to \mathbb{R}$ be convex and $[a, b] \subset \operatorname{int}(I)$. Then
(a) f is Lipschitzian relative to $[a, b]$.
(b) f is continuous on $\operatorname{int}(I)$.

Proof. There exist $c, d \in I$ such that $c < a < b < d$. By Theorem 1.6, we have

$$f'_+(a) \leqslant f'_+(x) \leqslant \frac{f(x) - f(y)}{x - y} \leqslant f'_-(y) \leqslant f'_-(b)$$

whenever $a \leqslant x < y \leqslant b$. It follows that $|f(x) - f(y)| \leqslant K|x - y|$ where $K := \max(|f'_+(a)|, |f'_-(b)|)$. This proves (a); (b) is an immediate consequence of (a).

Remark. Note that f is not necessarily Lipschitzian relative to I, even if f is bounded, and that f is not necessarily continuous on I, even if I is closed and finite.

1.8

A function which is Lipschitzian relative to an interval $[a, b]$ is absolutely continuous on $[a, b]$; it is a well-known fact that such a function is differentiable almost everywhere. It thus follows from § 1.7 that a convex function is differentiable almost everywhere.

In the sequel we shall prove a still stronger differentiability property of convex functions, without making use of the concept of absolute continuity.

Theorem. Let $f\colon I \to \mathbb{R}$ be convex. Then
(a) On $\operatorname{int}(I)$, f'_- is left-continuous and f'_+ is right-continuous.
(b) There are only countably many points where f is not differentiable.

Proof. (a) In virtue of the continuity of f on $\operatorname{int}(I)$ (§ 1.7) we have, for all x, y, $z \in \operatorname{int}(I)$

$$\frac{f(y) - f(x)}{y - x} = \lim_{z \downarrow x} \frac{f(y) - f(z)}{y - z} \geqslant \lim_{z \downarrow x} f'_+(z)$$

whenever $x < z < y$. Passing to the limit as $y \downarrow x$, we obtain

$$f'_+(x) \geqslant \lim_{z \downarrow x} f'_+(z).$$

Since f'_+ is non-decreasing (Theorem 1.6) we have

$$f'_+(x) \leq \lim_{z \downarrow x} f'_+(z).$$

We conclude that $f'_+(x) = \lim_{z \downarrow x} f'_+(z)$, which proves the right-continuity of f'_+. The left-continuity of f'_- can be proved in a similar way.

(b) By Theorem 1.6, we have

$$f'_+(x) \leq f'_-(y) \leq f'_+(z)$$

for all x, y, $z \in \mathrm{int}(I)$ with $x < y < z$. If f'_+ is continuous at y, we have

$$f'_+(y) = \lim_{x \uparrow y} f'_+(x) = \lim_{z \downarrow y} f'_+(z) = f'_-(y)$$

which means that f is differentiable at y. It follows that the points of $\mathrm{int}(I)$ where f is not differentiable are those where the non-decreasing function f'_+ has a jump. This proves (b), since there are only countably many such jumps.

MIDPOINT CONVEXITY

1.9

Closely related to convexity is the following concept.

Definition. A function $f\colon I \to \mathbb{R}$ is said to be *midpoint convex* if for all a, $b \in I$

$$f\left(\frac{a+b}{2}\right) \leq \tfrac{1}{2}[f(a) + f(b)]. \qquad (4)$$

Figure 4 shows the geometrical meaning of midpoint convexity: the midpoint of the chord connecting two points on the graph of f does not lie underneath the corresponding point on the graph.

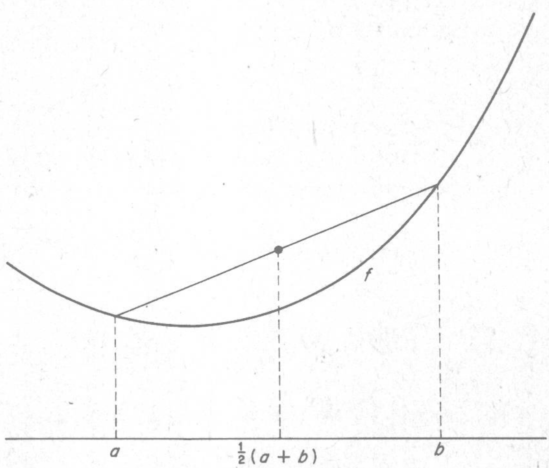

Figure 4

1.10 Theorem

Let $f: I \to \mathbb{R}$ be midpoint convex and continuous. Then f is convex.

Proof. Let (a_k) be a sequence in I. From (4) it follows that

$$f\left(\frac{a_1+a_2+a_3+a_4}{4}\right) \leqslant \tfrac{1}{2} f\left(\frac{a_1+a_2}{2}\right) + \tfrac{1}{2} f\left(\frac{a_3+a_4}{2}\right)$$
$$\leqslant \tfrac{1}{4}[f(a_1)+f(a_2)+f(a_3)+f(a_4)]$$

and by induction, one can prove that

$$f\left(\frac{a_1+a_2+\ldots+a_n}{n}\right) \leqslant \frac{1}{n} \sum_{i=1}^{n} f(a_i) \tag{5}$$

for all n of the form 2^k.

Assume now that (5) holds when $n = N$. Setting

$$a_N = \frac{1}{N-1}(a_1+a_2+\ldots+a_{N-1}),$$

we have

$$a_N = \frac{1}{N}(a_1+\ldots+a_N)$$

and hence

$$f(a_N) = f\left(\frac{a_1+\ldots+a_N}{N}\right) \leqslant \frac{1}{N} \sum_{i=1}^{N-1} f(a_i) + \frac{1}{N} f(a_N).$$

It follows that

$$f(a_N) \leqslant \frac{1}{N-1} \sum_{i=1}^{N-1} f(a_i)$$

so that (5) holds also when $n = N-1$. We conclude that (5) holds for all $n \in \mathbb{N}$.

Let $a, b \in I$ and $k, n \in \mathbb{N}$ such that $k < n$. From (5) it follows that

$$f\left(\frac{k}{n}a + \frac{n-k}{n}b\right) \leqslant \frac{1}{n}[kf(a)+(n-k)f(b)]$$

and hence

$$f(\lambda a + (1-\lambda)b) \geqslant \lambda f(a) + (1-\lambda)f(b) \tag{6}$$

whenever $\lambda \in \mathbb{Q}$, $0 < \lambda < 1$. In virtue of the continuity of f we conclude that (6) also holds whenever $\lambda \in \mathbb{R}$, $0 < \lambda < 1$.

1.11 Theorem

Let I be open, and let $f: I \to \mathbb{R}$ be twice differentiable. Then f is convex if and only if $f''(x) \geqslant 0$ for all $x \in I$.

Proof. 'Only if': by Theorem 1.6, f' is non-decreasing on I. Hence $f''(x) \geqslant 0$ for all $x \in I$.

'If': let $x, y \in I$, $x < y$ and $0 < \lambda < 1$. By the mean value theorem of calculus, there exist ξ_1, ξ_2, $x < \xi_1 < \lambda x + (1-\lambda)y < \xi_2 < y$ and ξ_3, $\xi_1 < \xi_3 < \xi_2$ such that

$$\begin{aligned} &f(\lambda x + (1-\lambda)y) - \lambda f(x) - (1-\lambda)f(y) \\ &\quad = \lambda[f(\lambda x + (1-\lambda)y) - f(x)] + (1-\lambda)[f(\lambda x + (1-\lambda)y) - f(y)] \\ &\quad = \lambda(1-\lambda)(y-x)f'(\xi_1) + (1-\lambda)\lambda(x-y)f'(\xi_2) \\ &\quad = \lambda(1-\lambda)(y-x)(\xi_1 - \xi_2)f''(\xi_3) \leqslant 0. \end{aligned}$$

It follows that f is convex.

Remark. From the above proof we can conclude that f is strictly convex if $f''(x) > 0$ for all $x \in I$. The converse is not true: the function $f: x \mapsto x^4$ is strictly convex on $\mathbb{R}$, but we have $f''(0) = 0$.

1.12 Inequalities

Many simple examples of convex functions can be obtained from Theorem 1.11, and by means of these functions one can derive inequalities which often are not so simple at first sight. We give an example:

$$x^\lambda y^\mu \leqslant \lambda x + \mu y \tag{7}$$

whenever $x > 0$, $y > 0$, $\lambda > 0$, $\mu > 0$ and $\lambda + \mu = 1$. This inequality can be derived by using the (strict) convexity of the function $x \mapsto e^x$ in the following form:

$$\exp(\lambda \log x + \mu \log y) \leqslant \lambda \exp(\log x) + \mu \exp(\log y).$$

Other well-known ways of presenting (7) are

$$x^{1/p} y^{1/q} \leqslant \frac{1}{p} x + \frac{1}{q} y \tag{8}$$

and

$$xy \leqslant \frac{1}{p} x^p + \frac{1}{q} y^q \tag{9}$$

where $x > 0$, $y > 0$, $p > 1$, $q > 1$ and $1/p + 1/q = 1$. For $p = q = 2$, (8) is the well-known inequality $\sqrt{(xy)} \leqslant \frac{1}{2}(x + y)$.

THEOREMS CONCERNING INTEGRALS

1.13 Theorem

Let f be a function $\langle a, b\rangle \to \mathbb{R}$. Then f is convex if and only if f can be represented in the form

$$f(x) = f(c) + \int_c^x g(t)\,dt \quad (c, x \in \langle a, b\rangle) \tag{10}$$

where g is a non-decreasing right-continuous function $\langle a, b\rangle \to \mathbb{R}$.

Proof. 'Only if': let f be convex and $c, x \in \langle a, b\rangle$. By Theorem 1.6 and § 1.8, f'_+ exists and is non-decreasing and right-continuous. Set

$$h(\varepsilon) := \int_c^x \frac{f(t+\varepsilon) - f(t)}{\varepsilon}\,dt.$$

We have

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon}[f(t+\varepsilon) - f(t)] = f'_+(t) \quad (a < t < b).$$

By § 1.7, there exists $K > 0$ such that

$$\left|\frac{1}{\varepsilon}[f(t+\varepsilon) - f(t)]\right| \leq K$$

for all t between c and x and all sufficiently small $\varepsilon > 0$. Applying Lebesgue's dominated convergence theorem yields

$$\lim_{\varepsilon \downarrow 0} h(\varepsilon) = \int_c^x f'_+(t)\,dt$$

(note that the last integral is a Riemann integral, in virtue of the monotonicity of the integrand). We also have

$$\begin{aligned}\frac{1}{\varepsilon}\int_c^x [f(t+\varepsilon) - f(t)]\,dt &= \frac{1}{\varepsilon}\left[\int_{c+\varepsilon}^{x+\varepsilon} f(t)\,dt - \int_c^x f(t)\,dt\right] \\ &= \frac{1}{\varepsilon}\int_x^{x+\varepsilon} f(t)\,dt - \frac{1}{\varepsilon}\int_c^{c+\varepsilon} f(t)\,dt \to f(x) - f(c) \quad \text{as } \varepsilon \downarrow 0\end{aligned}$$

(in view of the continuity of f). Hence

$$f(x) - f(c) = \int_c^x f'_+(t)\,dt. \tag{11}$$

'If': suppose that (10) holds, where g is non-decreasing. Let x, $y \in \langle a, b\rangle$,