凸 分 析

Jan Van Tiel 著

陈宏基译

（供数学系青年教师进修用）

惠 阳 师 专

译者的话

译这本书是为本系数学分析教研组的教师提供进修用书，作为泛函分析的一个分支，这本书介绍了凸集、凸函数和凸最优化。凸分析的应用很广，此书不但对数学工作者而且对物理学工作者、工程技术人员、企业管理人员、经济学家都是很有用的。

我要感谢何文端副教授，他~~全部~~看完我的全部译稿并且提出宝贵意见，我还要感谢学校教材科全体同志，他们为印出此书付出了劳动，特别要感谢陈宝全同志他为我刻写全部稿文。

译者陈宏基 1987年2月

于惠州西湖

目　录

序 言

这本小册子是根据我在荷兰 Utrech 大学教凸分析的讲义的基础上整理得来的. 目前凸分析在理论和实践上愈来愈有用了. 这本书首先是一本入门读物; 因此我力图把重点放在这门学科的基本概念和特征方法上(比如分离, 次梯度, 共轭函数, 凸最优化). 每一章末尾的大量练习(在书的结尾有答案和提示)是打算用来对有关概念的理解提供帮助.

这本书是为那些对凸分析有兴趣的青年学生写的. 他们具有分析、线代数和一般拓扑学的数学基础知识. 也假定他们已熟悉了泛函分析的基本概念(比如赋范线性空间, 希尔伯特空间, 对偶).

为了突出这门课的特点和引起学生的兴趣, 我没有把自己局限在有限维的情况中, 如常规那样处理, 但为了使局部凸空间, 凸分析"自然域"等有关内容尽量简单化, 这里只提出赋范空间

几个历史事件和补充的资料收集在书目提要的注释里; 当然它决不是详尽的.

第一章扼要地讲实直线上的凸函数的基本理论. 我们也考虑到推广几种可取无限值的函数.

第二章是研究在线性空间中凸函数的代数性质, 在线性拓扑空间的情况, 我们发现凸集的几个拓朴性质.

第三章推导出线性空间的分离定理. 在线性拓朴空间中应用这个定理得出 Hahn-Banach 定理.

第四章考虑关于 $\mathbb{R}^n$ 的凸子集的几个经典定理及其在多胞形锥的几个应用. 根据相对内部原理, 我们在 $\mathbb{R}^n$ 中讨论了分离性.

第五章讨论线性空间中可取无限值的凸函数．在某种意义下可推出局部有界性与连续性是彼此等价的．我们讨论下半连续性和次可微性的一些重要概念．

第六章提出对偶性原理，我们得出双极函数和支撑函数的特征．

第七章给出凸性在最优化中的实质设想．它主要处理有关凸规划问题（Kuhn－Tucker 条件，鞍点和 Fenchel 对偶定理）．

我要感谢 John Horvath 教授他建议我把讲义译成英文．我要感谢我的同事 Tineke de Bunje 和 Leen Roozemond 他们全部或部分读完我的原稿并对它作了许多改正．最后，我感谢 M.M. Meijer 小姐，她花了许多时间打出原稿．

Jan van Tiel 1984

第 一 章

$\mathbb{R}^n$上的凸函数

在这一章里我们将指定 I 表示 $\mathbb{R}$ 中的区间（闭的，开的或半开的，有限的或无限的）.

实的凸函数

1.1 [定义]

设 f 是一个 $I \longrightarrow \mathbb{R}$ 的函数.

(a) f 称为凸的，若对一切的 $a, b \in I$ 及一切的 $\lambda \in \mathbb{R}$ 且 $0 < \lambda < 1$，有

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

图 1 说明凸性的几何意义：以 $(a, f(a))$ 和 $(b, f(b))$ 为端点的弦上点都不在 f 的图象的下方.

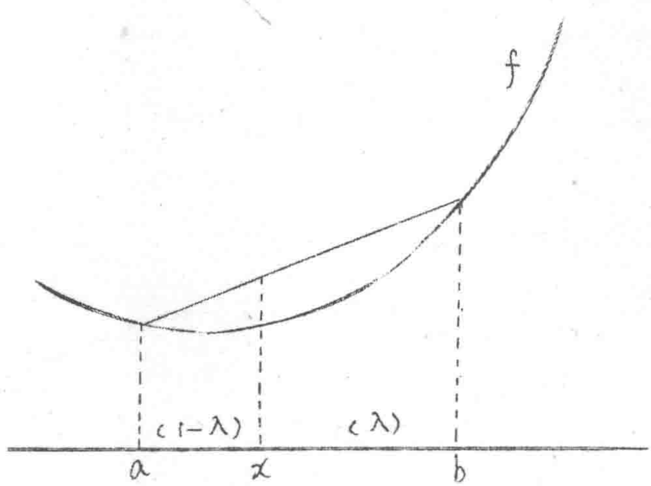

图 1

) f 为严格凸的，若它是凸的且对于 $a \neq b$，不等式 (1) 是严格不等式.

.2

我们给出 $f: I \longrightarrow \mathbb{R}$ 凸性的几个等价形式.

1)
$$f(x) \leqslant \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b)$$

对一切 $a, b \in I$ 及 $a < x < b$，注意这个不等式的右边可以写成
$$f(a) + \frac{f(b) - f(a)}{b-a}(x-a).$$

2)
$$f(\lambda a + \mu b) \leqslant \lambda f(a) + \mu f(b)$$

对一切 $a, b \in I$ 及一切的 $\lambda, \mu \in \mathbb{R}$ 且满足 $\lambda > 0$，$\mu > 0$，$\lambda + \mu = 1$.

.3

下面简单性质的证明留给读者.

1) 若 f 和 g 是凸函数且 $\alpha \geqslant 0$，$\beta \geqslant 0$，则 $\alpha f + \beta g$ 也是凸函数.

2) 有限多个凸函数的和是凸函数.

3) 一个收敛的凸函数序列的（点态）极限也是凸函数.

4) 设 $f: I \longrightarrow \mathbb{R}$ 是凸的，则

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

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

5) 设 f 是任意一个 $I \rightarrow \mathbb{R}$ 的凸函数的集合的点态上确界. 若 f 在 I 上处处有限，则 f 是凸的. 对于下确界有类似的命题吗？

4 [定理]

设 $f: I \rightarrow \mathbb{R}$ 是凸的，则

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

这里 $a, b, x \in I$，$a<x<b$．若 f 是严格凸的，则在(2)式中严格不等式成立．图 2 说明了定理的几何意义：斜率（AB）⩽斜率（AC）⩽斜率（BC）．

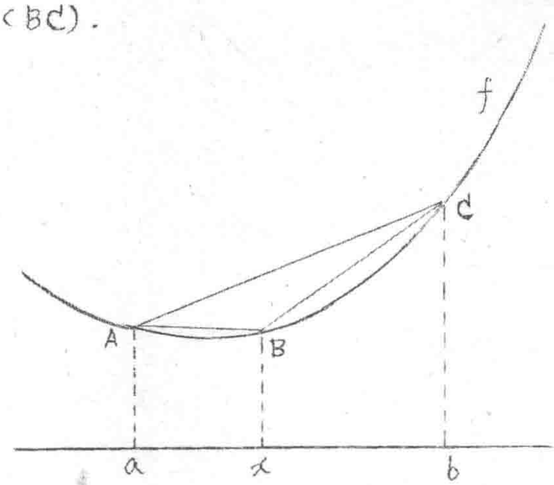

图 2

[证明] 因为 f 是凸的，我们有

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

由这个不等式我们可得

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

证明了(2)式的第一个不等式；第二个不等式可用类似方法获证．若 f 是严格凸的，则在(3)式中严格不等式成立，从而在(2)式中严格不等式也成立．

1.5

我们用 int(I) 表示 I 的内部．设 $f: I \to \mathbb{R}$ 是凸的，设 $c \in \text{int}(I)$．设 $[a,b] \subset I$ 满足 $a<c<b$．由定理 1.4，我们有

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

这里 $x \in \langle c, b]$，根据定理 1.4，函数

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

在 $\langle c, b]$ 上是非递减的．因此右导数

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

存在．用类似方法可证明左导数 $f'_-(c)$ 存在．

若 $a < c < d < b$，则对于充分小的正数 h，我们有

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

当 $h \downarrow 0$ 取极限，我们得到

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

于是我们证明了下面的定理

1.6 [定理]

设 $f: I \to \mathbb{R}$ 是凸的，则 f 在 $\operatorname{int}(I)$ 的每一点上都有右导数和左导数，且 f'_- 和 f'_+ 在 $\operatorname{int}(I)$ 上是非递减的．若 $c \in \operatorname{int}(I)$，我们有

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

及

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

对于一切的 $x \in I$（参见图 3）

[评注] 设 $f: [a, b] \to \mathbb{R}$ 是凸的．若 $+\infty$ 和 $-\infty$ 允许作为极限的话，上面论述说明了在这种情况下 $f'_+(a)$ 和 $f'_-(b)$ 存在．

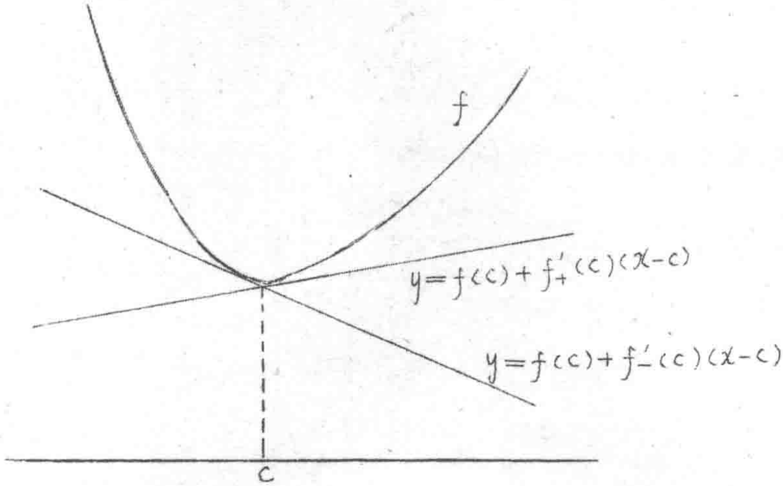

图 3

.7

$f: I \to \mathbb{R}$ 称在 $I_0 \subset I$ 上利普希茨若存在 $K>0$ 使得 $|f(x)-f(y)| \leqslant K|x-y|$ 对于一切的 $x, y \in I_0$．这个条件隐含在 I_0 上是连续的甚至是一致连续的，且在 I_0 的每一个有界闭子区间上是有界变差的.

[定理] 设 $f: I \to \mathbb{R}$ 是凸的及 $[a,b] \subset \operatorname{int}(I)$．则

(a) f 是 $[a,b]$ 上的利普希茨函数．

(b) f 在 $\operatorname{int}(I)$ 上是连续的．

[证明] 存在 $c, d \in I$ 使得 $c<a<b<d$．由定理 1.6 我们有

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

这里 $a \leqslant x < y \leqslant b$，接着有 $|f(x)-f(y)| \leqslant K|x-y|$ 这里 $K:=\max(|f'_+(a)|, |f'_-(b)|)$．这就证明了 (a)；(b) 是 (a) 的一个直接结果．

[评注] 注意尽管 f 是有界的，f 在 I 上利普希茨也不是必要的．尽管 I 是闭的和有限的，f 在 I 上连续同样不是必要的．

1.8

一个在 $[a,b]$ 上的利普希兹函数在 $[a,b]$ 上是绝对连续的。众所周知这样的函数是几乎处处可微分的。由 §1.7 得出凸函数是处处可微分的。

我们将继续证明凸函数一个更强的可微分的性质，而不用到绝对连续性的概念。

【定理】 设 $f: I \to \mathbb{R}$ 是凸的，则

(a) 在 $\mathrm{int}(I)$ 上，f'_- 是左连续的和 f'_+ 是右连续的。

(b) 仅有可数多个点使到 f 不可微。

【证明】 (a) 由 $\mathrm{int}(I)$ 上的 f 的连续性（§1.7），对一切的 $x,y,z \in \mathrm{int}(I)$ 我们有

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

其中 $x<z<y$。当 $y\downarrow x$ 取极限，我们得到

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

因为 f'_+ 是非递减的（定理 1.6），我们有

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

我们得到 $f'_+(x) = \lim\limits_{z\downarrow x} f'_+(z)$ 这就证明了 f'_+ 的右连续性。用同样的方法可证明 f'_- 的左连续性。

(b) 由定理 1.6，我们有

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

对于一切的 $x,y,z \in \mathrm{int}(I)$ 且 $x<y<z$。若 f'_+ 在 y 上连续，我们有

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

这意味着 f 在 y 上是可微的. 由此可知 f 在 $\operatorname{int}(I)$ 上不可微的点正是非递减函数 f'_+ 有一个跳跃的点. 这就证明了 (b). 因为只有可数多次这样的跳跃.

中点凸性

1.9

以下概念与凸性紧密相关

[定义] 一个函数 $f: I \longrightarrow \mathbb{R}$ 称为是中点凸的若对于一切的 $a, b \in I$, 有

$$f\left(\frac{a+b}{2}\right) \leqslant \frac{1}{2}[f(a)+f(b)]$$

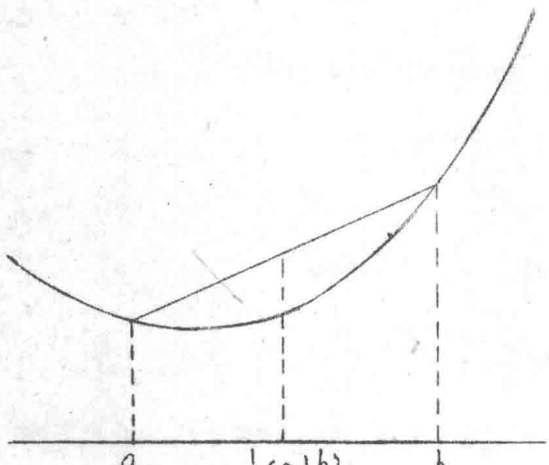

图 4

图 4 说明了中点凸性的几何意义：两端点在 f 的图象上的弦的中点不在图对应点的下方.

1.10 [定理]

设 $f: I \longrightarrow \mathbb{R}$ 是中点凸的和连续的，则 f 是凸的.

[证明] 设(a_k)是I的一个序列；根据(4)式可得

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

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

由归纳法，我们可以证得

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

对于形如2^K的n都成立.

现在假定(5)式当$n=N$成立. 令

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

我们有

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

因此

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

因此可得

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

所以(5)式当$n=N-1$也成立. 我们得到(5)式对于一切$n\in\mathbb{N}$成立.

设$a, b\in I$及$K, n\in\mathbb{N}$且$K<n$，由(5)式得出

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

从而

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

对于一切$\lambda\in\mathbb{Q}$, $0<\lambda<1$. 由f的连续性，我们得到(6)式当$\lambda\in\mathbb{R}$, $0<\lambda<1$也成立.

凸函数的可微性

1.11［定理］

设I是开区间，设 $f: I \longrightarrow \mathbb{R}$ 是二次可微函数．则 f 是凸的当且仅当 $f''(x) \geqslant 0$ 对一切的 $x \in I$.

［证明］"仅当"：由定理 1.6，f' 在I上是非递减的，因此 $f''(x) \geqslant 0$ 对所有的 $x \in I$．

"当"：设 $x, y \in I$，$x<y$ 且 $0<\lambda<1$．由微积分学的中值定理，存在 ξ_1, ξ_2，$x<\xi_1<\lambda x+(1-\lambda)y<\xi_2<y$ 和 ξ_3，$\xi_1<\xi_3<\xi_2$ 使得

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

即 f 是凸的．

［评注］从上面的证明我们可以得出 f 是严格凸的，若 $f''(x)>0$ 对一切 $x \in I$，反之不对：函数 $f: x \mapsto x^4$ 在 $\mathbb{R}$ 上是严格凸的，但我们有 $f''(0)=0$

1.12 不等式

从定理 1.11 可以得到凸函数许多简单例子，由这些函数的意义我们可以得到一些一眼看去并不那么简单的不等式．我们给出一个例子

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

这里 $x>0$，$y>0$，$\lambda>0$，$\mu>0$，$\lambda+\mu=1$．这个不等式可以利用函数 $x \mapsto e^{x}$ 的（严格）凸性到下面形式获得：

$$\exp(\lambda\log x+\mu\log y)\leq\lambda\exp(\log x)+\mu\exp(\log y)$$

叙述(7)式的另一种著名的方法是

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

和

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

这里 $x>0$，$y>0$，$p>1$，$q>1$，$\frac{1}{p}+\frac{1}{q}=1$，当 $p=q=2$ 时，(8)式就是著名的不等式 $\sqrt{xy}\leq\frac{1}{2}(x+y)$.

关于积分定理

1.13 [定理]

设 f 是一个函数 $\langle a,b\rangle\rightarrow\mathbb{R}$. 则 f 是凸的当且仅当 f 可以写成下面的形式

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

这里 $g:\langle a,b\rangle\rightarrow\mathbb{R}$ 是一个非递减的右连续的函数.

[证明] "仅当"：设 f 是凸的及 $c,x\in\langle a,b\rangle$，由定理 1.6 和 §1.8，f'_+ 存在且是非递减的和连续的. 令

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

我们有

$$\lim\frac{1}{\varepsilon}[f(t+\varepsilon)-f(t)]=f'_+(t) \qquad (a<t<b).$$

由 §1.7，存在 $K>0$ 使得

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

对于所有在 c 与 x 之间的 t 及所有充分小的 $\varepsilon>0$，应用勒贝格的控制收敛定理得出