目　　录

1 集合理论　关系　函数

1.1　$x \in A, x \notin B$

元素 x 属于集合 A，但 x 不属于集合 B.

1.2　$A = \{$典型元素：定义特征$\}$

集合定义的一般形式.

1.3　以下几种逻辑运算符号通常在表明 P 和 Q 关系时使用：

- $P \wedge Q$ 指“P 和 Q”
- $P \vee Q$ 指“P 或 Q”
- $P \Rightarrow Q$ 指“如果 P 则 Q”(或者“P 仅当 Q”，或“P 蕴涵 Q”.)
- $P \Leftarrow Q$ 指“如果 Q 则 P”(或“P 如果 Q”)
- $P \Leftrightarrow Q$ 指“P 当且仅当 Q”
- $\neg P$ 指“非 P”

逻辑运算符号.(注意“P 或 Q”是指“或者 P 或者 Q，或者两者都成立”.)

1.4

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
T	T	F	T	T	T	T
T	F	F	F	T	F	F
F	T	T	F	T	T	F
F	F	T	F	F	T	T

逻辑运算的是非表.在此 T 指“是”，F 指“非”.

1.5
- P 是 Q 的一个充分条件：$P \Rightarrow Q$
- Q 是 P 的一个必要条件：$P \Rightarrow Q$
- P 是 Q 的一个充分必要条件：$P \Leftrightarrow Q$

经常用到的术语.

1.6　$A \subset B \Leftrightarrow A$ 的每一个元素也是 B 的元素.

A 是 B 的一个子集.

1.7 $A \cup B = \{x : x \in A \lor x \in B\}$（$A$ 与 B 的并集）

$A \cap B = \{x : x \in A \land x \in B\}$（$A$ 与 B 的交集）

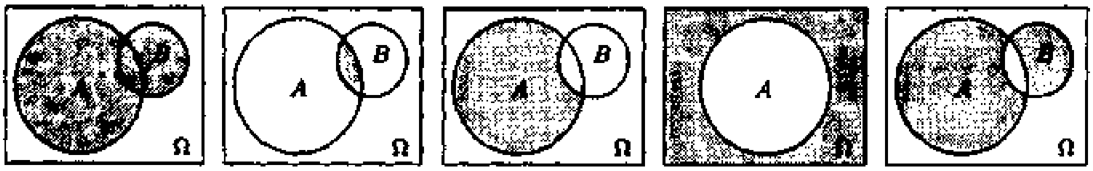
$A \setminus B = \{x : x \in A \land x \notin B\}$（$A$ 与 B 的差集）

$A^c = \{x : x \notin A\}$（A 的补集）

$A \triangle B = (A \setminus B) \cup (B \setminus A)$

（A 与 B 的对称差）

集合理论的基本定义.如果 Ω 是一个全集,则 $A^c = \Omega \setminus A$. A^c 的另一种符号是 $\complement_A$.



$A \cup B$　　$A \cap B$　　$A \setminus B$　　A^c　　$A \triangle B$

1.8
$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$
$$A \triangle B = (A \cup B) \setminus (A \cap B)$$
$$(A \triangle B) \triangle C = A \triangle (B \triangle C)$$

集合理论的重要运算法则.

1.9 $A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \cdots, a_n) : a_i \in A_i,\ i = 1, 2, \cdots, n\}$

集合 $A_1, A_2, \cdots, A_n$ 的笛卡尔乘积.

1.10 $R \subset A \times B$

任何 $A \times B$ 的子集 R 称为从集合 A 到集合 B 的关系.

1.11 $xRy \Leftrightarrow (x, y) \in R$

$x \not R y \Leftrightarrow (x, y) \notin R$

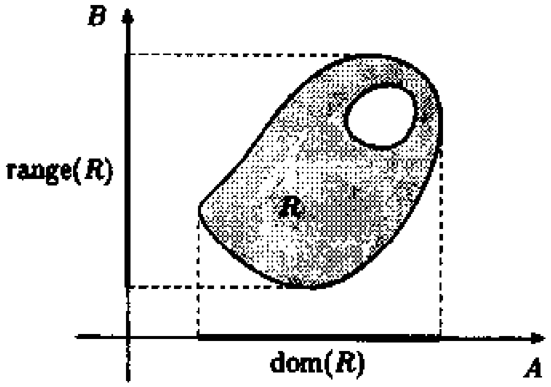
另一种关系和无关系的记法.我们称 x 满足对 y 的 R 关系,如果 $(x, y) \in R$.

1.12 • $\mathrm{dom}(R) = \{a \in A : (a, b) \in R$ 对于 B 中某一 $b\}$
$= \{a \in A : aRb$ 对于 B 中某一 $b\}$

• $\mathrm{range}(R) = \{b \in B : (a, b) \in R$ 对于 A 中某一 $a\}$
$= \{b \in B : aRb$ 对于 A 中某一 $a\}$

关系的区域及范围.

1.13



(1.13)定义的关系 R 的区域及范围,阴影部分集合是关系的图形.

1.14 $R^{-1} = \{(b, a) \in B \times A : (a, b) \in R\}$

从 A 到 B 的关系 R 的逆关系. R^{-1} 是从 B 到 A 的关系.

1.15 让 R 为从 A 到 B 的关系, S 为从 B 到 C 的关系,我们则可定义 R 和 S 的复合 $S \circ R$ 为属于 $A \times C$ 的 (a, c) 的集合,满足属于 B 的元素 b 具有 aRb 及 bSc. $S \circ R$ 是一个从 A 到 C 的关系.

$S \circ R$ 是两个关系 R 及 S 的复合.

1.16 从 A 到 A 自己的关系 R 称为 A 的二元关系. R 的二元关系具有

- 自反性如果 aRa 对于每一在 A 中的 a 成立;
- 非自反性如果 $a \not R a$ 对于每一在 A 中的 a 成立;
- 完整性如果 aRb 或 bRa 对于每一在 A 中的 a 和 b 成立,且 $a \neq b$;
- 传递性如果 aRb 及 bRc 则有 aRc;
- 对称性如果 aRb 则有 bRa;
- 反对称性如果 aRb 及 bRa 则有 $a = b$;
- 非对称性如果 aRb 则有 $b \not R a$.

特殊的关系.

1.17 A 的二元关系 R 被称为

- 先有序的(或拟有序的)关系，如果它是自反性的和传递性的；
- 弱有序的关系，如果它是传递性的和完整的；
- 部分有序的关系，如果它是自反性的，传递性的及反对称性的；
- 线性(或完全)有序的，如果它是自反性的，传递性的，反对称性的及完整的；
- 等价关系，如果它是自反性的，传递性的及对称的.
- 关系 $=$ 在实数中是一个相应关系.
- 关系 $\geq$ 在实数中是一个线性有序关系.
- 关系 $<$ 在实数中是一个弱有序及非自反性，非对称性关系.
- 关系 $\subseteq$ 在给定集合的子集中是一个部分有序关系.

特殊的关系.(这些术语并不是通用的.)注意线性有序与完整的部分有序是一样的.
序列关系通常用符号 $\preccurlyeq$, $\leqslant$, $\ll$ 等标示. 逆关系用 $\succcurlyeq$, $\geqslant$, $\gg$ 等标示.

1.18 • 关系 $x \preccurlyeq y$ (“y 至少与 x 一样好”)在一个商品向量集合里通常假定为完整的先有序关系.

- 关系 $x \prec y$ (“y(严格好于)x”)在一个商品向量集合里通常假定为非自反的，传递的(因此是非对称性).
- 关系 $x \sim y$(“x 与 y 无差别”)在一个商品向量集合里通常假定为等价关系.

关系的例子. 对于关系 $x \preccurlyeq y$, $x \prec y$, 及 $x \sim y$, 见第26章.

1.19 令 $\preccurlyeq$ 为在集合 A 中的一个先有序关系. 在 A 中的一个元素 g 称为在 A 中对于关系 $\preccurlyeq$ 的最大元素，如果 $x \preccurlyeq g$ 对于每一在 A 中的 x 成立. 在 A 中的一个元素 m 称为在 A 中对于关系 $\preccurlyeq$ 的极大元素，如果 $x \in A$ 且 $m \preccurlyeq x$ 意味着 $x \preccurlyeq m$. 对 $\preccurlyeq$ 的最小和极小元素分别是对于 $\preccurlyeq$ 的逆关系 $\succcurlyeq$ 的极大和最大元素.

有序集合中最大元素，极大元素，最小元素和极小元素的定义.

1.20 如果≤是一在 A 中的先有序关系，M 是 A 的一个子集，在 A 中的元素 b 称为 M 的一个上限(对于≤)，如果 $x \leqslant b$ 对于每一在 M 中的 x 成立．一个 M 的下限是一在 A 中的元素 a 使 $a \leqslant x$ 对于所有在 M 中的 x.

上限和下限的定义.

1.21 如果≤是在非空集合 A 中的一个先有序关系，且如果 A 的每一线性有序子集 M 都在 A 中有上限，则对于关系≤在 A 中存在一个极大元素.

Zorn 引理.(通常适用于部分有序，但对于先有序同样适用.)

1.22 一个从 A 到 B 的关系 R 称为函数或映射，如果对于每一在 A 中的 a，在 B 中都存在一个唯一的 b，使 aRb．如果用 f 作为函数符号，则 afb 可写成 $f(a) = b$．f 的图形则定义为
$\operatorname{graph}(f) = \{(a, b) \in A \times B : f(a) = b\}$.

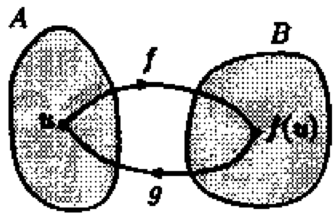
从集合 A 到集合 B 的函数及其图形的定义.

1.23 一个从 A 到 B 的函数 $f(f:A \to B)$ 称为
- 单射的(或一一对应的)如果 $f(x) = f(y)$ 意味着 $x = y$;
- 满射的(或自身映射)如果范围 $\operatorname{dom}(f) = B$;
- 双射的如果它既是单射的又是满射的.

函数的一些重要概念.

1.24 如果 $f:A \to B$ 双射的(即既是一一对应的又是自身映射的)，则它有反函数 $g:B \to A$，定义为 $g(f(u)) = u$ 对于所有 $u \in A$.

反函数的特点．f 的反函数通常记为 f^{-1}.

1.25 

反函数概念的图示.

1.26 如果 f 是一个从 A 到 B 的函数，且 $C \subset A$，$D \subset B$，我们则可使用记法
- $f(C) = \{f(x) : x \in C\}$
- $f^{-1}(D) = \{x \in A : f(x) \in D\}$

$f(C)$ 称为 A 在关系 f 下的像，$f^{-1}(D)$ 称为 D 的逆像.

1.27 如果 f 是从 A 到 B 的一个函数,而且 $S \subset A$, $T \subset A$, $U \subset B$, $V \subset B$, 则

$$f(S \cup T) = f(S) \cup f(T)$$
$$f(S \cap T) \subset f(S) \cap f(T)$$
$$f^{-1}(U \cup V) = f^{-1}(U) \cup f^{-1}(V)$$
$$f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V)$$
$$f^{-1}(U \setminus V) = f^{-1}(U) \setminus f^{-1}(V)$$

重要的事实.(包含符号 $\subset$ 不能由 $=$ 替代)

1.28 如果 $\mathrm{N} = \{1, 2, 3, \cdots\}$ 为自然数的集合,而 $\mathrm{N}_n = \{1, 2, 3, \cdots, n\}$, 则

- 集合 A 是有限的如果它是空的,或对于某一自然数 n,存在一从 A 到 N_n 的一一对应函数.
- 集合 A 是可数性无限的如果存在一从 A 到 N 的一一对应函数.

一个有限的或可数性无限的集合通常称为可数的. 有理数集合是可数性无限的,而实数集合是不可数的.

参考文献

参照 Halmos (1974), Ellickson (1993)及 Hildenbrand (1974).

2 方程式 一元函数 复数

2.1 $ax^2+bx+c=0 \Leftrightarrow x_{1,2}=\dfrac{-b\pm\sqrt{b^2-4ac}}{2a}$

一般一元二次方程的根.($a\neq 0$)如果 $b^2\geqslant 4ac$ 则是实数根.(假设 a, b,和 c 是实数)

2.2 如果 x_1 及 x_2 是方程 $x^2+px+q=0$ 的根,则

$x_1+x_2=-p$, $x_1x_2=q$

Viète 定理.

2.3 $ax^3+bx^2+cx+d=0$

一般一元三次方程.

2.4 $x^3+px+q=0$

如果在(2.3)中的 x 写为 $x-b/3a$,(2.3)可简写为(2.4).

2.5 $x^3+px+q=0$ 且 $\Delta=4p^3+27q^2$, 则有

- 三个不同的根,如果 $\Delta<0$;
- 三个实根,其中至少两个相等,如果 $\Delta=0$;
- 一个实根,两个复根,如果 $\Delta>0$.

(2.4)中根的分类.(假设 p 和 q 是实数)

2.6 方程 $x^3+px+q=0$ 的解为

$x_1=u+v$, $x_2=\omega u+\omega^2 v$, 以及 $x_3=\omega^2 u+\omega v$,其中 $\omega=-\dfrac{1}{2}+\dfrac{i}{2}\sqrt{3}$, $x_3=-\dfrac{u+v}{2}-\dfrac{u-v}{2}\sqrt{-3}$, 且有

$$u=\sqrt[3]{-\frac{q}{2}+\frac{1}{2}\sqrt{\frac{4p^3+27q^2}{27}}}$$

$$v=\sqrt[3]{-\frac{q}{2}-\frac{1}{2}\sqrt{\frac{4p^3+27q^2}{27}}}$$

对于一元三次方程的根的 Cardano 公式. i 是虚数单位(参照(2.72)),而 ω 是 1 的第三个复根(参见(2.85)).(除非必要时,尽量不要使用这一公式!)

2.7	如果 x_1, x_2 和 x_3 是方程 $x^3 + px^2 + qx + r = 0$ 的根,则 $x_1 + x_2 + x_3 = -p$ $x_1x_2 + x_1x_3 + x_2x_3 = q$ $x_1x_2x_3 = -r$	有用的关系.
2.8	$P(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$	n 阶多项式. ($a_n \neq 0$)
2.9	对于(2.8)中的多项式 $P(x)$,存在常数 x_1, x_2, $\cdots$, x_n(实数或复数),使 $P(x) = a_n(x - x_1)\cdots(x - x_n)$	代数基本定理. x_1, $\cdots$, x_n 称为 $P(x)$ 的零点和 $P(x) = 0$ 的根.
2.10	$x_1 + x_2 + \cdots + x_n = -\frac{a_{n-1}}{a_n}$ $x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n = \sum_{i<j} x_ix_j = \frac{a_{n-2}}{a_n}$ $x_1x_2\cdots x_n = (-1)^n \frac{a_0}{a_n}$	$P(x) = 0$ 的根与系数的关系,其中 $P(x)$ 定义在(2.8). ((2.2)与(2.7)的一般化)
2.11	如果(2.8)的系数 a_n, a_{n-1}, $\cdots$, a_1, a_0 均为整数,p 和 q 是没有共同因子的整数,且 $P(p/q) = 0$, 则 p 可整除 a_0, q 可整除 a_n.	多项式的有理零点.
2.12	设 k 为(2.8)中系数数列 a_n, a_{n-1}, $\cdots$, a_1, a_0 改变符号的次数. $P(x) = 0$ 的实正根的个数,包括根的乘积,是 k 或 k 减去一个正偶数. 如果 $k = 1$, 则方程仅有一个正实根.	*Descartes* 符号法则.
2.13	方程 $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$ 的图形为 • 一个椭圆,一点,或空集,如果 $4AC > B^2$; • 一条抛物线,一条直线,两条平行线,或空集,如果 $4AC = B^2$; • 一个双曲线或两条相交线,若 $4AC < B^2$.	二次曲线的分类. A,B,C 不全为 0.

2.14 $x = x'\cos\theta - y'\sin\theta$, $y = x'\sin\theta + y'\cos\theta$ 其中 $\cot 2\theta = (A - C)/B$

将方程(2.13)转换成对于 x', y'的二次方程,其中 $x'y'$ 的系数为0.

2.15 $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

两点 $(x_1,\ y_1)$ 和 $(x_2,\ y_2)$ 之间的(欧几里德)距离.

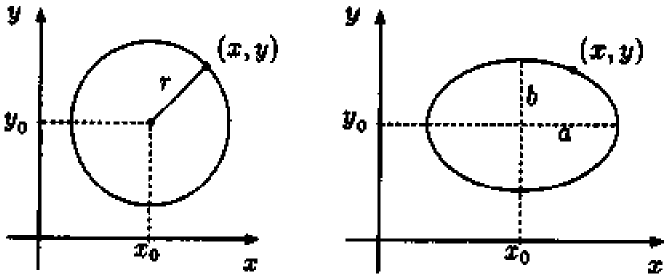
2.16 $(x - x_0)^2 + (y - y_0)^2 = r^2$

以 $(x_0,\ y_0)$ 为圆心,r 为半径的圆.

2.17 $\dfrac{(x - x_0)^2}{a^2} + \dfrac{(y - y_0)^2}{b^2} = 1$

以 $(x_0,\ y_0)$ 为中心,轴与坐标轴平行的椭圆.

2.18



(2.16)与(2.17)的图形.

2.19 $\dfrac{(x - x_0)^2}{a^2} - \dfrac{(y - y_0)^2}{b^2} = \pm 1$

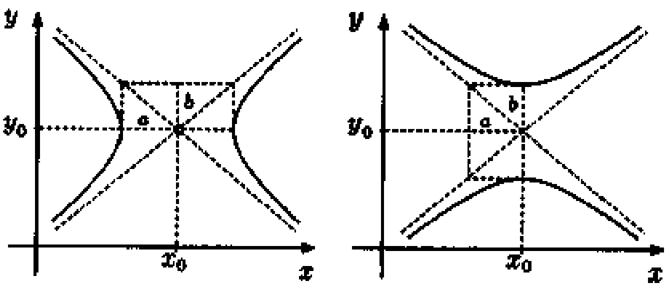
以 $(x_0,\ y_0)$ 为中心,轴与坐标轴平行的双曲线.

2.20 (2.19)的渐近线:

$$y - y_0 = \pm \frac{b}{a}(x - x_0)$$

(2.19)中双曲线的渐近线的公式.

2.21



(2.19)与(2.20)中的双曲线和其渐近线的图形,对应于各自的正负号.

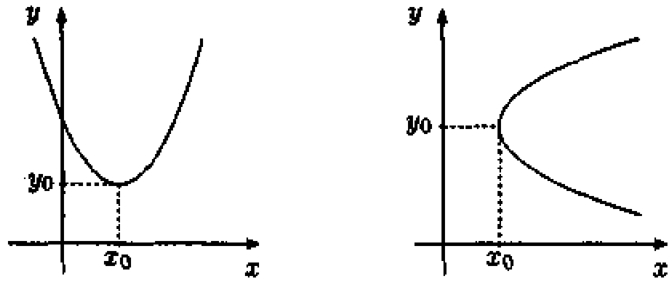
2.22 $y - y_0 = a(x - x_0)^2,\ a \neq 0$

以 $(x_0,\ y_0)$ 为顶点,主轴平行于 y 轴的抛物线.

2.23 $x - x_0 = a(y - y_0)^2,\ a \neq 0$

以 $(x_0,\ y_0)$ 为顶点，主轴平行于 x 轴的抛物线.

2.24



(2.22)与(2.23)中的抛物线的图形，其中 $a > 0$.

2.25 函数 f 是

- 递增的，如果

  $x_1 < x_2 \Rightarrow f(x_1) \leqslant f(x_2)$
- 严格递增的，如果

  $x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$
- 递减的，如果

  $x_1 < x_2 \Rightarrow f(x_1) \geqslant f(x_2)$
- 严格递减的，如果

  $x_1 < x_2 \Rightarrow f(x_1) > f(x_2)$
- 偶函数，如果 $f(x) = f(-x)$ 对于所有 x 成立.
- 奇函数，如果 $f(x) = -f(-x)$ 对于所有 x 成立.
- 关于直线 $x = a$ 对称的，如果

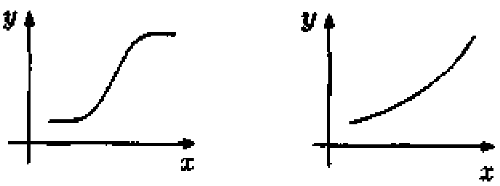
  $f(a + x) = f(a - x)$ 对于所有 x 成立.
- 关于点 $(a, 0)$ 对称的，如果

  $f(a - x) = -f(a + x)$ 对于所有 x 成立.
- 周期性的，其周期为 k，如果存在一个 $k > 0$ 使

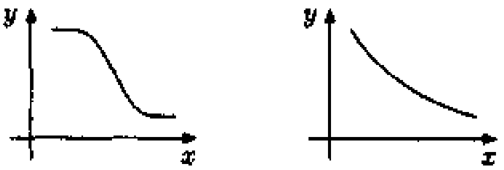
  $f(x + k) = f(x)$ 对于所有 x 成立.

函数的性质.

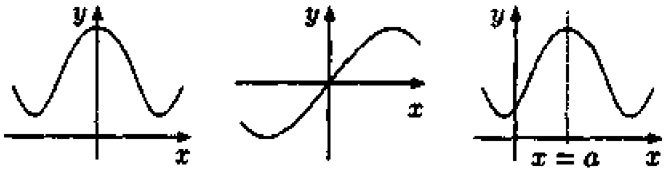
2.26



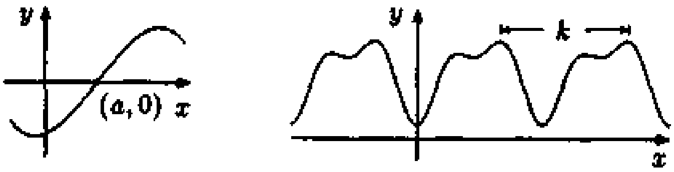
递增及严格递增函数的图形.

2.27 

递减及严格递减函数的图形.

2.28 

偶函数,奇函数以及对称于 $x = a$ 的函数的图形.

2.29 

对称于点$(a, 0)$的函数,及以 k 为周期的函数的图形.

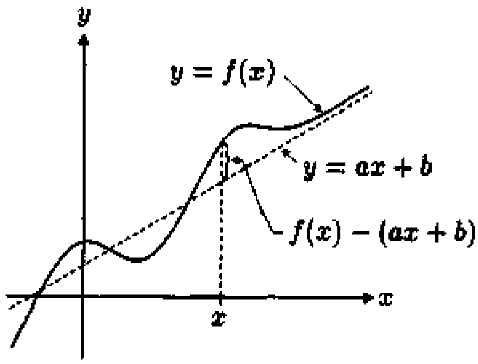
2.30 $y = ax + b$ 是一条曲线$y = f(x)$的非竖直渐近线,如果

$$\lim_{x\to\infty}(f(x) - (ax + b)) = 0$$

或

$$\lim_{x\to-\infty}(f(x) - (ax + b)) = 0$$

非竖直渐近线的定义.

2.31 

$y = ax + b$ 是曲线 $y = f(x)$ 的一条渐近线.

2.32 怎样寻找曲线 $y = f(x)$ 当 $x \to \infty$ 时的渐近线:

- 检查 $\lim\limits_{x\to\infty}(f(x)/x)$. 如果当 $x \to \infty$ 时极限不存在,则没有渐近线.
- 如果 $\lim\limits_{x\to\infty}(f(x)/x) = a$, 检查极限 $\lim\limits_{x\to\infty}(f(x) - ax)$. 如果当 $x \to \infty$ 时极限不存在,则没有渐近线.
- 如果 $\lim\limits_{x\to\infty}(f(x) - ax) = b$, 则 $y = ax + b$ 是曲线 $y = f(x)$ 当 $x \to \infty$ 时的一条渐近线.

寻找曲线 $y = f(x)$ 当 $x \to \infty$ 时的非竖直渐近线的方法. 将 $x \to \infty$ 换为 $x \to -\infty$, 则可求到曲线 $y = f(x)$ 当 $x \to -\infty$ 时的非竖直渐近线.

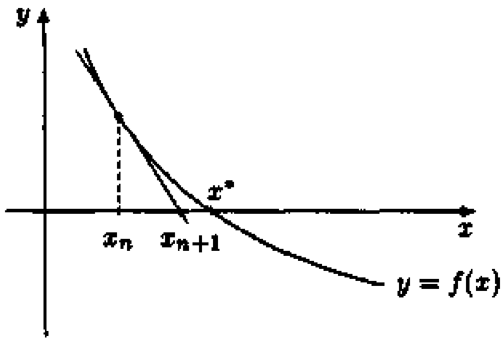
2.33 要求 $f(x)=0$ 的近似根,定义 x_n 对于 $n=1, 2, \cdots$,

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

如果 x_n 接近一个真实根 x^*,则数列 $\{x_n\}$ 将迅速收敛至真实根.

牛顿近似法.

2.34



牛顿近似法的图示.

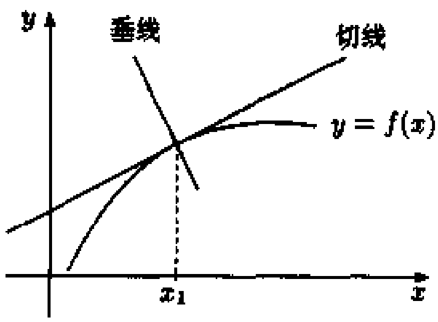
2.35 $y - f(x_1) = f'(x_1)(x - x_1)$

与 $y=f(x)$ 相切于点 $(x_1, f(x_1))$ 的切线方程.

2.36 $y - f(x_1) = -\dfrac{1}{f'(x_1)}(x - x_1)$

与 $y=f(x)$ 垂直于 $(x_1, f(x_1))$ 点的垂线方程.

2.37



$y = f(x)$ 在点 $(x_1, f(x_1))$ 的切线与垂线.

2.38 (i) $a^r \cdot a^s = a^{r+s}$ (ii) $(a^r)^s = a^{rs}$

(iii) $(ab)^r = a^r b^r$ (iv) $a^r / a^s = a^{r-s}$

(v) $\left(\dfrac{a}{b}\right)^r = \dfrac{a^r}{b^r}$ (vi) $a^{-r} = \dfrac{1}{a^r}$

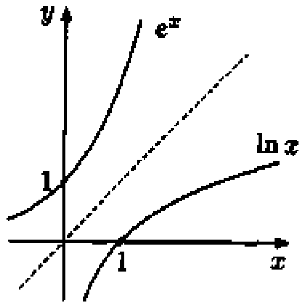
幂法则.(r 与 s 是任意实数,a 与 b 是正实数)

2.39 $\mathrm{e} = \lim\limits_{n\to\infty}\left(1 + \dfrac{1}{n}\right)^n = 2.718281828459\cdots$

e 的定义.(参见(8.20)中 $x=1$)

2.40 $\mathrm{e}^{\ln x} = x$

自然对数的定义.

2.41 

$y = e^x$ 和 $y = \ln x$ 的图形.

2.42 $\ln(xy) = \ln x + \ln y$, $\ln\dfrac{x}{y} = \ln x - \ln y$

$\ln x^p = p\ln x$, $\ln\dfrac{1}{x} = -\ln x$

自然对数函数的法则.(x 与 y 为正)

2.43 $a^{\log_a x} = x\ (a > 0,\ a \neq 1)$

底数为 a 的对数的定义.

2.44 $\log_a x = \dfrac{\ln x}{\ln a}$, $\log_a b \cdot \log_b a = 1$

$\log_e x = \ln x \quad \log_{10} x = \log_{10} e \cdot \ln x$

不同底数的对数.

2.45 $\log_a(xy) = \log_a x + \log_a y$

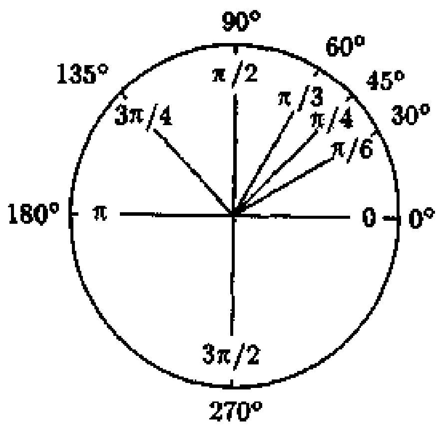
$\log_a \dfrac{x}{y} = \log_a x - \log_a y$

$\log_a x^p = p\log_a x$, $\log_a \dfrac{1}{x} = -\log_a x$

对数法则.(x 与 y 为正)

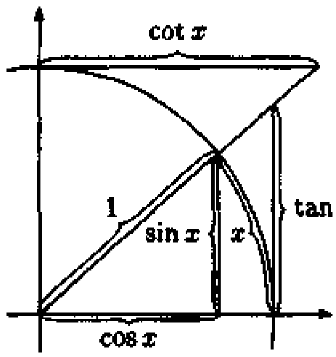
2.46 $1^\circ = \dfrac{\pi}{180}\text{rad}$, $1\ \text{rad} = \left(\dfrac{180}{\pi}\right)^\circ$

角度与弧度(rad)的关系.

2.47 

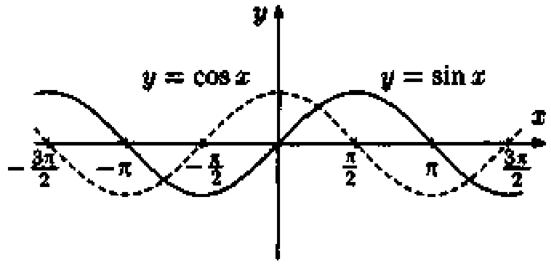
角度与弧度的关系.

2.48



基本三角函数的定义. x 是弧的长度，也是角的弧度.

2.49

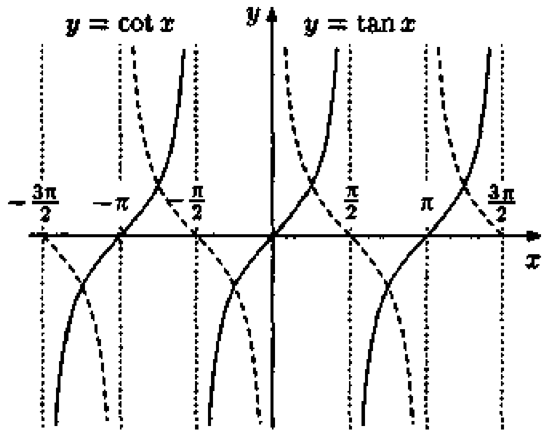


$y = \sin x$ 与 $y = \cos x$ 的图形. 函数 sin 和 cos 是周期性的，周期为 2π:
$\sin(x + 2\pi) = \sin x$,
$\cos(x + 2\pi) = \cos x$.

2.50　$\tan x = \dfrac{\sin x}{\cos x}$, $\cot x = \dfrac{\cos x}{\sin x} = \dfrac{1}{\tan x}$

正切函数与余切函数的定义.

2.51



$y = \tan x$ 与 $y = \cot x$ 的图形. 函数 tan 和 cot 是周期性的，周期为 π:
$\tan(x + \pi) = \tan x$,
$\cot(x + \pi) = \cot x$.

2.52

x	0	$\frac{\pi}{6} = 30°$	$\frac{\pi}{4} = 45°$	$\frac{\pi}{3} = 60°$	$\frac{\pi}{2} = 90°$
$\sin x$	0	$\frac{1}{2}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}\sqrt{3}$	1
$\cos x$	1	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}$	0
$\tan x$	0	$\frac{1}{3}\sqrt{3}$	1	$\sqrt{3}$	*
$\cot x$	*	$\sqrt{3}$	1	$\frac{1}{3}\sqrt{3}$	0

* 无定义

三角函数的特殊值.

2.53

x	$\frac{3\pi}{4}=135^\circ$	$\pi=180^\circ$	$\frac{3\pi}{2}=270^\circ$	$2\pi=360^\circ$
$\sin x$	$\frac{1}{2}\sqrt{2}$	0	-1	0
$\cos x$	$-\frac{1}{2}\sqrt{2}$	-1	0	1
$\tan x$	-1	0	*	0
$\cot x$	-1	*	0	*

*无定义

三角函数公式.(关于三角函数的级数展开,参见第8章)

2.54 $\sin^2 x+\cos^2 x=1$

2.55 $\tan^2 x=\frac{1}{\cos^2 x}-1,\ \cot^2 x=\frac{1}{\sin^2 x}-1$

2.56 $\cos(x+y)=\cos x\cos y-\sin x\sin y$

$\cos(x-y)=\cos x\cos y+\sin x\sin y$

2.57 $\sin(x+y)=\sin x\cos y+\cos x\sin y$

$\sin(x-y)=\sin x\cos y-\cos x\sin y$

2.58 $\tan(x+y)=\frac{\tan x+\tan y}{1-\tan x\tan y}$

$\tan(x-y)=\frac{\tan x-\tan y}{1+\tan x\tan y}$

三角函数公式.

2.59 $\cos 2x=2\cos^2 x-1=1-2\sin^2 x$

$\sin 2x=2\sin x\cos x$

2.60 $\sin^2\frac{x}{2}=\frac{1-\cos x}{2},\ \cos^2\frac{x}{2}=\frac{1+\cos x}{2}$

2.61 $\cos x+\cos y=2\cos\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right)$

$\cos x-\cos y=-2\sin\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right)$

2.62 $\sin x + \sin y = 2\sin\left(\frac{x+y}{2}\right)\cos\left(\frac{x-y}{2}\right)$

$\sin x - \sin y = 2\cos\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right)$

2.63 $y = \arcsin x \Leftrightarrow x = \sin y$, $x \in [-1, 1]$, $y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

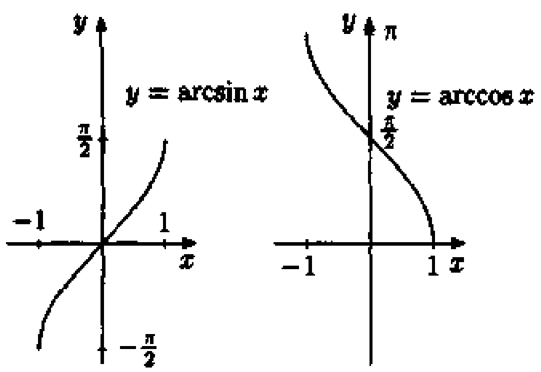
$y = \arccos x \Leftrightarrow x = \cos y$, $x \in [-1, 1]$, $y \in [0, \pi]$

$y = \arctan x \Leftrightarrow x = \tan y$, $x \in \mathrm{R}$, $y \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$

$y = \operatorname{arccot} x \Leftrightarrow x = \cot y$, $x \in \mathrm{R}$, $y \in (0, \pi)$

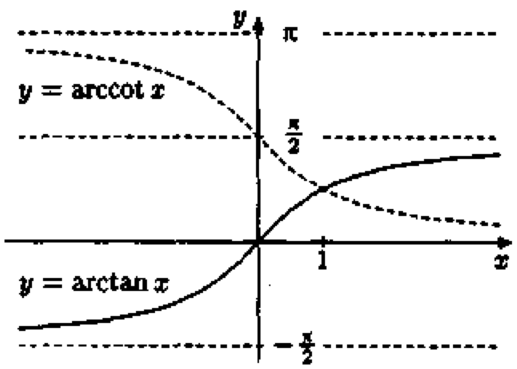
反三角函数的定义.

2.64



反三角函数 $y = \arcsin x$ 和 $y = \arccos x$ 的图形.

2.65



反三角函数 $y = \arctan x$ 和 $y = \operatorname{arccot} x$ 的图形.

2.66 $\arcsin x = \sin^{-1} x$, $\arccos x = \cos^{-1} x$

$\arctan x = \tan^{-1} x$, $\operatorname{arccot} x = \cot^{-1} x$

反三角函数的另一种记法.

2.67 $\arcsin(-x) = -\arcsin x$

$\arccos(-x) = \pi - \arccos x$

$\arctan(-x) = -\arctan x$

$\operatorname{arccot}(-x) = \pi - \operatorname{arccot} x$

$\arcsin x + \arccos x = \frac{\pi}{2}$

$\arctan x + \operatorname{arccot} x = \frac{\pi}{2}$

$\arctan \frac{1}{x} = \frac{\pi}{2} - \arctan x,\ x > 0$

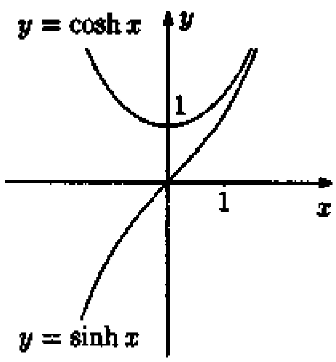
$\arctan \frac{1}{x} = -\frac{\pi}{2} - \arctan x,\ x < 0$

反三角函数的性质.

2.68 $\sinh x = \frac{e^x - e^{-x}}{2},\ \cosh x = \frac{e^x + e^{-x}}{2}$

双曲正弦函数和双曲余弦函数.

2.69



双曲函数 $y = \sinh x$ 与 $y = \cosh x$ 的图形.

2.70 $\cosh^2 x - \sinh^2 x = 1$

$\cosh(x + y) = \cosh x \cosh y + \sinh x \sinh y$

$\cosh 2x = \cosh^2 x + \sinh^2 x$

$\sinh(x + y) = \sinh x \cosh y + \cosh x \sinh y$

$\sinh 2x = 2\sinh x \cosh x$

双曲函数 $y = \sinh x$ 与 $y = \cosh x$ 的性质.

2.71 $y = \operatorname{arcsinh} x \Leftrightarrow x = \sinh y$

$y = \operatorname{arccosh} x,\ x \geqslant 1 \Leftrightarrow x = \cosh y,\ y \geqslant 0$

$\operatorname{arcsinh} x = \ln\left(x + \sqrt{x^2 + 1}\right)$

$\operatorname{arccosh} x = \ln\left(x + \sqrt{x^2 - 1}\right),\ x \geqslant 1$

反双曲函数的定义.

复数

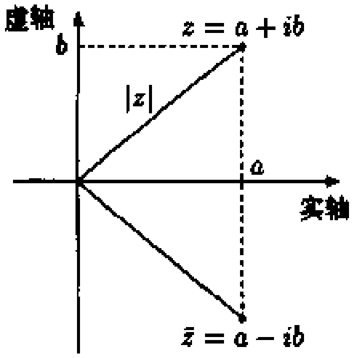
2.72　$z = a + \mathrm{i}b,\ \bar{z} = a - \mathrm{i}b$

复数及其共轭. $a, b \in \mathbf{R}$, 且 $\mathrm{i}^2 = -1$. i 称为虚数单位

2.73　$|z| = \sqrt{a^2 + b^2},\ \mathrm{Re}(z) = a,\ \mathrm{Im}(z) = b$

$|z|$是 $z = a + \mathrm{i}b$ 的模或量值. $\mathrm{Re}(z)$和 $\mathrm{Im}(z)$分别为它的实部和虚部.

2.74



复数及其共轭的几何表达.

2.75
- $(a + \mathrm{i}b) + (c + \mathrm{i}d) = (a + c) + \mathrm{i}(b + d)$
- $(a + \mathrm{i}b) - (c + \mathrm{i}d) = (a - c) + \mathrm{i}(b - d)$
- $(a + \mathrm{i}b)(c + \mathrm{i}d) = (ac - bd) + \mathrm{i}(ad + bc)$
- $\dfrac{a + \mathrm{i}b}{c + \mathrm{i}d} = \dfrac{1}{c^2 + d^2}((ac + bd) + \mathrm{i}(bc - ad))$

复数的加、减、乘和除.

2.76　$|\bar{z}_1| = |z_1|,\ z_1\bar{z}_1 = |z_1|^2,$

$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2,$

$|z_1 z_2| = |z_1||z_2|,$

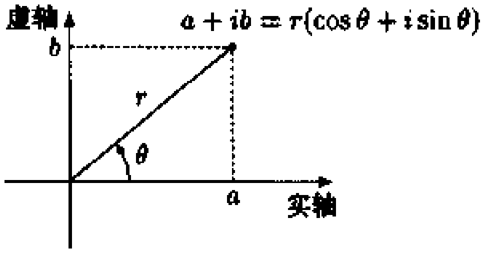
$|z_1 + z_2| \leqslant |z_1| + |z_2|$

基本法则. z_1 和 z_2 为复数.

2.77　$z = a + \mathrm{i}b = r(\cos\theta + \mathrm{i}\sin\theta) = r\mathrm{e}^{\mathrm{i}\theta}$, 其中 $r = |z| = \sqrt{a^2 + b^2},\ \cos\theta = \dfrac{a}{r},\ \sin\theta = \dfrac{b}{r}$

复数的三角函数形式或极坐标形式. 角 θ 称为 z 的自变量. 对于 $\mathrm{e}^{\mathrm{i}\theta}$ 参见(2.81).

2.78



复数的三角函数形式的几何表达.

2.79 如果 $z_k = r_k(\cos\theta_k + \mathrm{i}\sin\theta_k)$, $k = 1, 2$, 则

$$z_1 z_2 = r_1 r_2(\cos(\theta_1 + \theta_2) + \mathrm{i}\sin(\theta_1 + \theta_2))$$

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}(\cos(\theta_1 - \theta_2) + \mathrm{i}\sin(\theta_1 - \theta_2))$$

三角函数形式的乘与除.

2.80 $(\cos\theta + \mathrm{i}\sin\theta)^n = \cos n\theta + \mathrm{i}\sin n\theta$

De Moivre 公式, $n = 0, 1, \cdots$.

2.81 如果 $z = x + \mathrm{i}y$, 则

$$\mathrm{e}^z = \mathrm{e}^{x+\mathrm{i}y} = \mathrm{e}^x \cdot \mathrm{e}^{\mathrm{i}y} = \mathrm{e}^x(\cos y + \mathrm{i}\sin y).$$

复指数函数.

2.82 $\mathrm{e}^{\pi\mathrm{i}} = -1$

非常重要的关系.

2.83 $\mathrm{e}^{\bar z} = \overline{\mathrm{e}^z}$, $\mathrm{e}^{z+2\pi\mathrm{i}} = \mathrm{e}^z$, $\mathrm{e}^{z_1+z_2} = \mathrm{e}^{z_1}\mathrm{e}^{z_2}$, $\mathrm{e}^{z_1-z_2} = \mathrm{e}^{z_1}/\mathrm{e}^{z_2}$

复指数函数的法则.

2.84 $\cos z = \dfrac{\mathrm{e}^{\mathrm{i}z} + \mathrm{e}^{-\mathrm{i}z}}{2}$, $\sin z = \dfrac{\mathrm{e}^{\mathrm{i}z} - \mathrm{e}^{-\mathrm{i}z}}{2\mathrm{i}}$

欧拉公式.

2.85 如果 $a = r(\cos\theta + \mathrm{i}\sin\theta) \neq 0$, 则方程

$$z^n = a$$

确切地有 n 个根,

$$z_k = \sqrt[n]{r}\left(\cos\frac{\theta + 2k\pi}{n} + \mathrm{i}\sin\frac{\theta + 2k\pi}{n}\right)$$

对 $k = 0, 1, \cdots, n-1$.

第 n 个复根数, $n = 1, 2, \cdots$.

参 考 文 献

绝大多数的公式可在任何微积分教材里找到,比如 Edwards & Penney (1998) 或 Sydsaeter & Hammond (1995). 关于(2.3)—(2.12),参照 Turnbull(1952).

3 极限 连续 微分(一元)

3.1	当 x 接近 a 时,$f(x)$接近于极限 A, $\lim\limits_{x\to a}f(x) = A$ 或 $f(x)\to A$ 当 $x\to a$ 如果对于每一个数 $\varepsilon>0$,存在一个数 $\delta>0$,使 $\lvert f(x)-A\rvert<\varepsilon$,当 $x\in D_f$ 且 $0<\lvert x-a\rvert<\delta$	一元函数的极限的定义.D_f 是 f 的定义域.
3.2	如果 $\lim\limits_{x\to a}f(x) = A$ 且 $\lim\limits_{x\to a}g(x) = B$,则 • $\lim\limits_{x\to a}(f(x)\pm g(x)) = A\pm B$ • $\lim\limits_{x\to a}(f(x)\cdot g(x)) = A\cdot B$ • $\lim\limits_{x\to a}\dfrac{f(x)}{g(x)} = \dfrac{A}{B}$ (如果 $B\neq 0$)	极限的法则.
3.3	如果 $\lim\limits_{x\to a}f(x) = f(a)$,则 f 是在点 $x = a$ 连续的,即如果 $a\in D_f$,且对于每一数 $\varepsilon>0$,存在一个数 $\delta>0$,满足 $\lvert f(x)-A\rvert<\varepsilon$,若 $x\in D_f$ 且 $\lvert x-a\rvert<\delta$, f 是在集合 $S\subset D_f$ 里连续的,如果 f 在 S 中的每一点上连续.	连续的定义.
3.4	如果 f 和 g 在 a 点连续,则 • $f\pm g$ 和 $f\cdot g$ 在 a 点连续; • f/g 在 a 点连续,如果 $g(a)\neq 0$.	连续函数的性质.
3.5	如果 g 在 a 点连续,且 f 在 $g(a)$ 连续,则 $f(g(x))$在 a 点连续.	复合函数的连续性.
3.6	连续函数通过加,减,乘,除,复合得到的任何函数都在其定义域里连续.	一个有用的结论.

3.7 如果对于每一 $\varepsilon > 0$，存在一个 $\delta > 0$（取决于 ε 但不取决于 x 和 y），使 $|f(x) - f(y)| < \varepsilon$，若 $x, y \in S$ 且 $|x - y| < \delta$，则 f 在集合 S 里是一致连续的.

一致连续的定义.

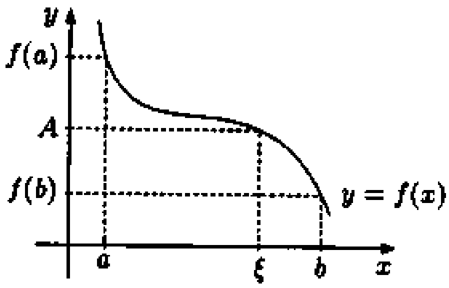
3.8 如果 f 在闭区间 I 里连续，则 f 在 I 上是一致连续的.

在闭区间里的连续函数是一致连续的.

3.9 如果 f 是在包含 a 与 b 的闭区间 I 里连续的，且 A 在 $f(a)$ 与 $f(b)$ 之间，则至少存在 a 与 b 之间的一点 ξ，使 $A = f(\xi)$.

介值定理.

3.10



介值定理的图示.

3.11 $$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

导数的定义. 如果极限存在，f 称为在 x 点可微.

3.12 $y = f(x)$ 的导数的其他记法：

$$f'(x) = y' = \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}f(x)}{\mathrm{d}x} = Df(x)$$

导数的其他记法.

3.13 $y = f(x) \pm g(x) \Rightarrow y' = f'(x) \pm g'(x)$

一般法则.

3.14 $y = f(x)g(x) \Rightarrow$

$y' = f'(x)g(x) + f(x)g'(x)$

3.15 $$y = \frac{f(x)}{g(x)} \Rightarrow y' = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$$

3.16 $y = f(g(x)) \Rightarrow y' = f'(g(x)) \cdot g'(x)$

链法则.

3.17 $y = (f(x))^{g(x)} \Rightarrow$

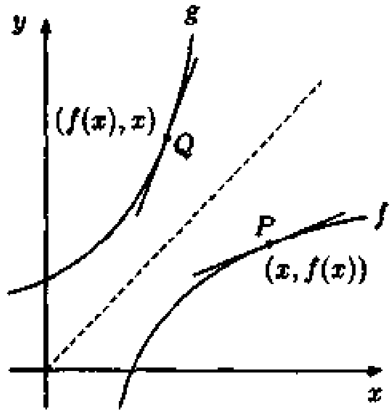
$$y' = (f(x))^{g(x)}\left(g'(x)\ln f(x) + g(x)\frac{f'(x)}{f(x)}\right)$$

3.18 如果 $g = f^{-1}$ 是一一对应函数 f 的反函数，f 在 x 可微，且 $f'(x) \neq 0$，则 g 在 $f(x)$ 可微，且

$$g'(f(x)) = \frac{1}{f'(x)}$$

f^{-1} 是 f 的反函数.

3.19



如果在 P 点的切线的斜率为 $k = f'(x)$，则在 Q 点的切线的斜率 $g'(f(x))$ 等于 $1/k$.

3.20 $y = c \Rightarrow y' = 0$ （c 为常数） 特殊法则.

3.21 $y = x^a \Rightarrow y' = ax^{a-1}$ （a 为常数）

3.22 $y = \dfrac{1}{x} \Rightarrow y' = -\dfrac{1}{x^2}$

3.23 $y = \sqrt{x} \Rightarrow y' = \dfrac{1}{2\sqrt{x}}$

3.24 $y = e^x \Rightarrow y' = e^x$

3.25 $y = a^x \Rightarrow y' = a^x \ln a$ （$a > 0$）

3.26 $y = \ln x \Rightarrow y' = \dfrac{1}{x}$

3.27 $y = \log_a x \Rightarrow y' = \dfrac{1}{x}\log_a e$ （$a > 0,\ a \neq 1$）

3.28 $y = \sin x \Rightarrow y' = \cos x$

3.29 $y = \cos x \Rightarrow y' = -\sin x$

3.30 $y = \tan x \Rightarrow y' = \dfrac{1}{\cos^2 x} = 1 + \tan^2 x$

3.31 $y = \cot x \Rightarrow y' = -\dfrac{1}{\sin^2 x} = -(1 + \cot^2 x)$

3.32 $y=\sin^{-1}x=\arcsin x \Rightarrow y'=\dfrac{1}{\sqrt{1-x^2}}$

3.33 $y=\cos^{-1}x=\arccos x \Rightarrow y'=-\dfrac{1}{\sqrt{1-x^2}}$

3.34 $y=\tan^{-1}x=\arctan x \Rightarrow y'=\dfrac{1}{1+x^2}$ 特殊法则.

3.35 $y=\cot^{-1}x=\operatorname{arccot} x \Rightarrow y'=-\dfrac{1}{1+x^2}$

3.36 $y=\sinh x \Rightarrow y'=\cosh x$

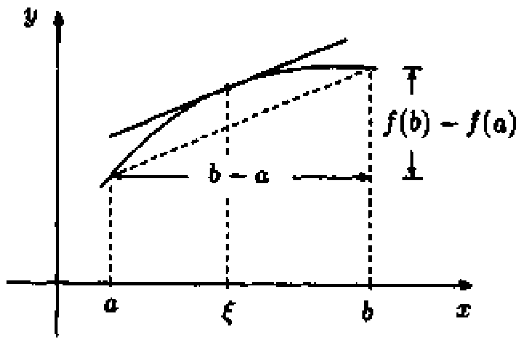
3.37 $y=\cosh x \Rightarrow y'=\sinh x$

3.38 如果 f 在 $[a, b]$ 连续, 在 (a, b) 可微, 则在 (a, b) 至少存在一点 ξ,使

$$f'(\xi)=\frac{f(b)-f(a)}{b-a}$$

中值定理.

3.39



中值定理的图示.

3.40 如果 f 和 g 在 $[a, b]$ 连续,在 (a, b) 可微,则在 (a, b) 至少存在一点 ξ,使

$$[f(b)-f(a)]g'(\xi)=[g(b)-g(a)]f'(\xi)$$

柯西一般中值定理.

3.41 假设 f 和 g 在包含 a 的一个区间 (α, β) 里(除了在点 a)可微,并假设当 x 趋向于 a 时 $f(x)$ 和 $g(x)$ 趋向于 0. 如果 $g'(x)\neq 0$ 对于所有在 (α, β) 中的 $x\neq a$ 成立,且 $\lim\limits_{x\to a} f'(x)/g'(x)=L$ (L 有限, $L=\infty$ 或 $L=-\infty$), 则

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\frac{f'(x)}{g'(x)}=L$$

洛必达法则. 当 $x\to a^+$, $x\to a^-$, $x\to+\infty$, 或 $x\to-\infty$, $f(x)\to\pm\infty$, $g(x)\to\pm\infty$ 时同样成立.

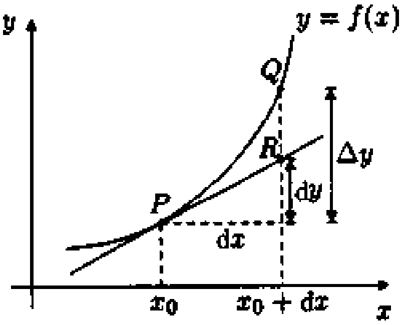
3.42 如果 $y = f(x)$ 且 dx 是任一数，

$$dy = f'(x)dx$$

是 y 的微分.

微分的定义.

3.43



微分的几何图示.

3.44 当 $|dx|$ 很小时 $\Delta y = f(x + dx) - f(x) \approx f'(x)dx$

一个有用的近似，精确表达在(3.45)中.

3.45 $f(x + dx) - f(x) = f'(x)dx + \varepsilon dx$

其中 $\varepsilon \to 0$ 当 $dx \to 0$

可微函数的性质.（如果 dx 非常小，则 ε 很小，且 εdx "非常非常小"）

3.46 $d(af + bg) = adf + bdg$（a 与 b 为常数）

$$d(fg) = gdf + fdg$$

$$d(f/g) = (gdf - fdg)/g^2$$

$$df(u) = f'(u)du$$

微分法则. f 与 g 可微，u 是任一可微函数.

参 考 文 献

所有公式都是标准的，均可在几乎所有的微积分教材里找到，例如 Edwards & Penney (1998)，或 Sydsaeter & Hammond (1995). 对于一致连续，参考 Rudin (1982).

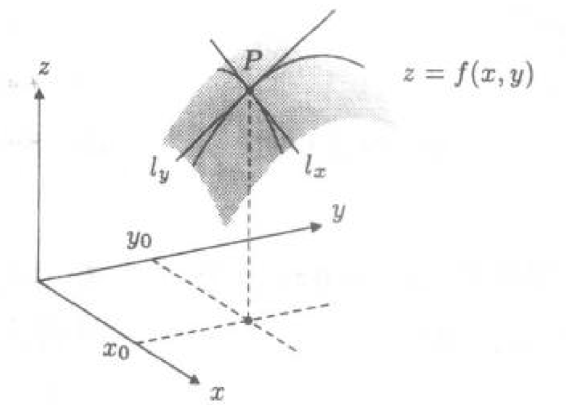
4 偏导数

4.1 如果 $z = f(x_1, \cdots, x_n) = f(\mathbf{x})$，则

$$\frac{\partial z}{\partial x_i} = \frac{\partial f}{\partial x_i} = f_i'(\mathbf{x}) = D_{x_i} f = D_i f = f_i'$$

都表示当其他变量固定时 $f(x_1, \cdots, x_n)$ 对 x_i 的导数.

偏导数的定义.(还有其他的记法)

4.2 

二元函数 $z = f(x, y)$ 的偏导数的几何图示：$f_1'(x_0, y_0)$ 是切线 l_x 的斜率，$f_2'(x_0, y_0)$ 是切线 l_y 的斜率.

4.3 $$\frac{\partial^2 z}{\partial x_j \partial x_i} = f_{ij}''(x_1, \cdots, x_n) = \frac{\partial}{\partial x_j} f_i'(x_1, \cdots, x_n)$$

$z = f(x_1, \cdots, x_n)$ 的二阶偏导数.

4.4 $$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}, \quad i, j = 1, 2, \cdots, n$$

杨氏定理，当两个偏导数连续时成立.

4.5 在集合 $S \subset R^n$ 中，$f(x_1, \cdots, x_n)$ 称为 C^k 级函数，或简称为 C^k，如果 f 的所有 $\leqslant k$ 阶的偏导数都在 S 里连续.

C^k 函数的定义.(对于连续性的定义，参见(12.12))

4.6 $z = F(x, y)$, $x = f(t)$, $y = g(t) \Rightarrow$

$$\frac{dz}{dt} = F_1'(x, y)\frac{dx}{dt} + F_2'(x, y)\frac{dy}{dt}$$

链法则.

4.7 如果 $z = F(x_1, \cdots, x_n)$ 且 $x_i = f_i(t_1, \cdots, t_m)$, $i = 1, \cdots, n$, 则对于所有 $j = 1, \cdots, m$

$$\frac{\partial z}{\partial t_j} = \sum_{i=1}^{n} \frac{\partial F(x_1, \cdots, x_n)}{\partial x_i} \frac{\partial x_i}{\partial t_j}$$

链法则.(一般情况)

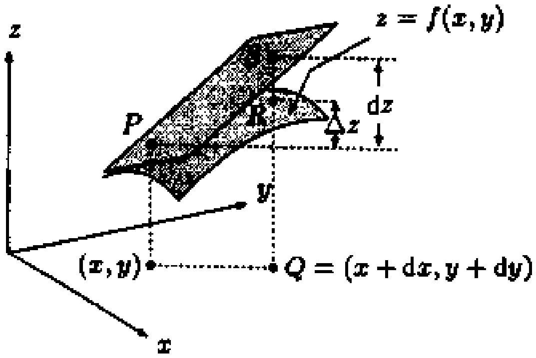
4.8 如果 $z = f(x_1, \cdots, x_n)$ 且 $dx_1, \cdots, dx_n$ 是任意数,则

$$dz = \sum_{i=1}^{n} f_i'(x_1, \cdots, x_n) dx_i$$

被称为 z 的微分.

微分的定义.

4.9



二元函数的微分的几何图示.这也是(4.10)中,近似值 $\Delta z = dz$ 的图示.

4.10 当 $|dx_1|, \cdots, |dx_n|$ 很小时, $\Delta z \approx dz$, 其中

$$\Delta z = f(x_1 + dx_1, \cdots, x_n + dx_n) - f(x_1, \cdots, x_n)$$

有用的近似,对于可微函数的精确表达见(4.11).

4.11 如果 $f_i'(x)$ 存在,且存在函数 $\epsilon_i = \epsilon_i(dx_1, \cdots, dx_n)$, $i = 1, \cdots, n$, 而且当 dx_i 接近 0 时都接近 0,并且 $\Delta z - dz = \epsilon_1 dx_1 + \cdots + \epsilon_n dx_n$, 则 f 在 x 可微.

可微性的定义.

4.12 如果 f 是一个 C^1 函数,即具有一阶连续偏导数,则 f 可微.

重要的事实.

4.13 $d(af + bg) = a df + b dg$ (a 和 b 为常数)

$d(fg) = g df + f dg$

$d(f/g) = (g df - f dg)/g^2$

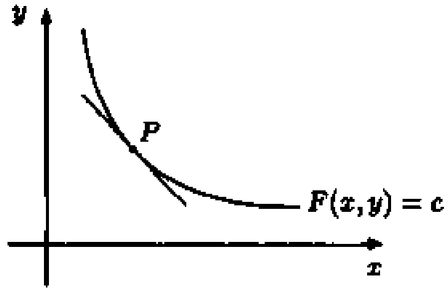
$dF(u) = F'(u) du$

微分法则. f 和 g 是 $x_1, \cdots, x_n$ 的可微函数, F 是一个一元可微函数, u 是 $x_1, \cdots, x_n$ 的任一可微函数.

4.14 $F(x, y)=c \Rightarrow \dfrac{\mathrm{d}y}{\mathrm{d}x}=-\dfrac{F_1'(x, y)}{F_2'(x, y)}$

$z=F(x, y)$ 的阶层曲线的斜率.精确的假设参见(4.17).

4.15



在 P 点切线的斜率是

$$\frac{\mathrm{d}y}{\mathrm{d}x}=-\frac{F_1'(x,y)}{F_2'(x,y)}.$$

4.16 如果 $y=f(x)$ 是一个 C^2 函数,满足 $F(x, y)=c$,则

$$f''(x)=-\frac{[F_{11}''(F_2')^2-2F_{12}''F_1'F_2'+F_{22}''(F_1')^2]}{(F_2')^3}$$

$$=\frac{1}{(F_2')^3}\begin{vmatrix} 0 & F_1' & F_2' \\ F_1' & F_{11}'' & F_{12}'' \\ F_2' & F_{12}'' & F_{22}'' \end{vmatrix}$$

一个有用的结论.所有偏导数取于(x, y).

4.17 如果 $F(x, y)$是在集合 A 里的C^k 函数,(x_0, y_0)是 A 内的一点, $F(x_0, y_0)=c$, 且 $F_2'(x_0, y_0)\neq 0$, 则方程 $F(x, y)=c$ 定义y为在(x_0, y_0)点邻域内 x 的 C^k 函数, $y=\varphi(x)$, 且 y 的导数为

$$\frac{\mathrm{d}y}{\mathrm{d}x}=-\frac{F_1'(x, y)}{F_2'(x, y)}$$

隐函数定理(对于一般结论,参见(6.3)).

4.18 如果 $F(x_1, x_2, \cdots, x_n, z)=c$ (c 为常数),则

$$\frac{\partial z}{\partial x_i}=-\frac{\partial F/\partial x_i}{\partial F/\partial z},\ i=1, 2, \cdots, n\ \left(\frac{\partial F}{\partial z}\neq 0\right)$$

(4.14)的一般化

齐次函数和位似函数

4.19 $f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_n)$ 是 $D \subset \mathbf{R}^n$ 内的 k 次齐次函数,如果

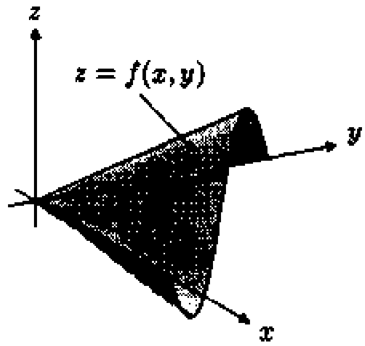
$f(tx_1, tx_2, \cdots, tx_n) = t^k f(x_1, x_2, \cdots, x_n)$

对于所有 $t > 0$ 和所有在 D 中的

$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ 都成立.

齐次函数的定义. D 是一个锥体因为 $\mathbf{x} \in D$ 和 $t > 0$ 意味着 $t\mathbf{x} \in D$.

4.20



1 次齐次函数的图示.

4.21 $f(\boldsymbol{x}) = f(x_1, \cdots, x_n)$ 是在开锥体 D 内的 k 次齐次函数,当且仅当

$\sum_{i=1}^{n} x_i f_i'(\boldsymbol{x}) = kf(\boldsymbol{x})$ 对于所有 $\boldsymbol{x}$ 在 D 内成立.

欧拉定理,对于 C^1 函数成立.

4.22 如果 $f(\boldsymbol{x}) = f(x_1, \cdots, x_n)$ 在开锥体 D 内为 k 次齐次函数,则

- $\partial f / \partial x_i$ 在 D 内齐次度为 $k - 1$;
- $\sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j f_{ij}''(\boldsymbol{x}) = k(k-1) f(\boldsymbol{x})$

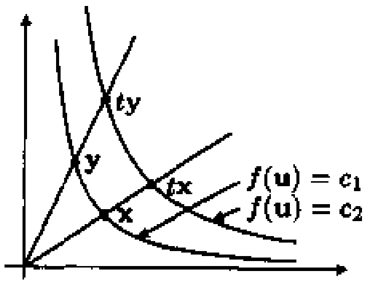
齐次函数的性质.

4.23 $f(\mathbf{x}) = f(x_1, \cdots, x_n)$ 是在锥体 D 内的位似函数,如果对于所有 $\boldsymbol{x}, \boldsymbol{y} \in D$ 及所有 $t > 0$,

$f(\boldsymbol{x}) = f(\boldsymbol{y}) \Rightarrow f(t\boldsymbol{x}) = f(t\boldsymbol{y})$

位似函数的定义.

4.24



位似函数的几何图示. 若 $f(\boldsymbol{u})$是位似的,如果 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 在同一阶层曲线上,则 $t\boldsymbol{x}$ 和 $t\boldsymbol{y}$ 也在同一阶层曲线上(当 $t > 0$).

4.25 设$f(\mathbf{x})$为定义在开锥体 D 内的连续位似函数. 假设 f 沿D 的每一射线是严格递增的(即对于每一 $\mathbf{x}_0 \in D$, $f(t\mathbf{x}_0)$ 是 t 的严格递增函数), 则存在一齐次函数 g 和一严格递增函数 F,使
$f(\mathbf{x}) = F(g(\mathbf{x}))$ 对于所有在 D 内的$\mathbf{x}$ 成立.

连续位似函数的性质(有时作为位似函数的定义). 可以假设 g 的齐次度为 1.

梯度　方向导数　切平面

4.26 $$\nabla f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \cdots, \frac{\partial f(\mathbf{x})}{\partial x_n}\right)$$

f 在$\mathbf{x} = (x_1, \cdots, x_n)$ 的梯度.

4.27 $$f'_{\mathbf{a}}(\mathbf{x}) = \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{a}) - f(\mathbf{x})}{h}, \ \| \mathbf{a} \| = 1$$

f 在$\mathbf{x}$ 以$\mathbf{a}$ 为方向的方向导数.

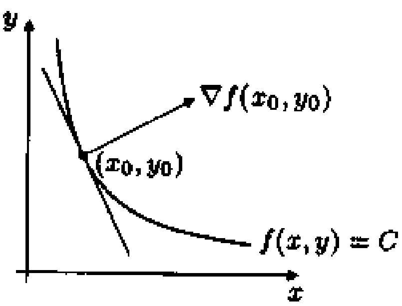
4.28 $$f'_{\mathbf{a}}(\mathbf{x}) = \sum_{i=1}^{n} f'_i(\mathbf{x}) a_i = \nabla f(\mathbf{x}) \cdot \mathbf{a}$$

方向导数和梯度的关系.

4.29
- $\nabla f(\mathbf{x})$垂直于截面 $f(\mathbf{x}) = C$
- $\nabla f(\mathbf{x})$标示 f 的最大增量的方向.
- $\| \nabla f(\mathbf{x}) \|$ 度量 f 在方向 $\nabla f(\mathbf{x})$上的变化率.

梯度的性质.

4.30

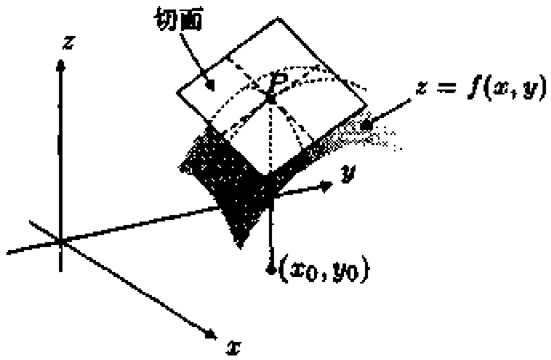


$f(x, y)$在(x_0, y_0)处的$\nabla f(x_0, y_0)$.

4.31 $z = f(x, y)$ 在点 $P = (x_0, y_0, z_0)$ 的切平面,且 $z_0 = f(x_0, y_0)$, 其方程形式为
$$z - z_0 = f'_1(x_0, y_0)(x - x_0) + f'_2(x_0, y_0)(y - y_0)$$

切平面的定义.

4.32



函数及其切平面的图示.

4.33 在点 $\boldsymbol{x}^0 = (x_1^0, \cdots, x_n^0)$ 切于函数的截面

$$F(\boldsymbol{x}) = F(x_1, \cdots, x_n) = C$$

的超切面方程为

$$\nabla F(\boldsymbol{x}^0) \cdot (\boldsymbol{x} - \boldsymbol{x}^0) = 0$$

超切面的定义.向量 $\nabla F(\boldsymbol{x}^0)$ 称为超切面的法线.

参 考 文 献

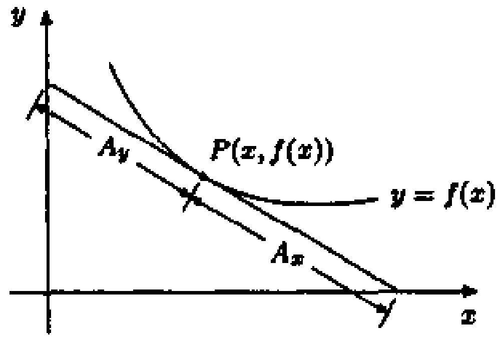
所有这些公式都是标准的并可在几乎所有的微积分教材里找到,例如 Edwards & Penney (1998)或 Sydsaeter & Hammond (1995).关于位似函数的性质,参考 Simon & Blume (1994), Shephard (1970)和 Førsund (1975).

5 弹性 替代弹性

5.1 $\mathrm{El}_x f(x) = \dfrac{x}{f(x)} f'(x) = \dfrac{\mathrm{d}(\ln f(x))}{\mathrm{d}(\ln x)}$

$\mathrm{El}_x f(x)$表示 $f(x)$ 对于 x 的弹性，是当 x 增加百分之一时 $f(x)$ 的相应百分比变化.

5.2



马歇尔法则的图示.

5.3 马歇尔法则：要找到图中 $y = f(x)$ 对于 x 在 P 点的弹性，首先画一条切于曲线 P 点的切线. 量出从 P 到切线与 y 轴交点的距离 A_y，和从 P 到切线与 x 轴交点的距离 A_x. 则 $\mathrm{El}_x f(x) = \pm A_y / A_x$.

马歇尔法则. 距离取正值. 如果曲线在 P 点上升取正号，反之则负号.

5.4 $\mathrm{El}_x(f(x)g(x)) = \mathrm{El}_x f(x) + \mathrm{El}_x g(x)$

弹性计算的一般法则.

5.5 $\mathrm{El}_x\left(\dfrac{f(x)}{g(x)}\right) = \mathrm{El}_x f(x) - \mathrm{El}_x g(x)$

5.6 $\mathrm{El}_x(f(x) \pm g(x)) = \dfrac{f(x)\mathrm{El}_x f(x) \pm g(x)\mathrm{El}_x g(x)}{f(x) \pm g(x)}$

5.7 $\mathrm{El}_x f(g(x)) = \mathrm{El}_u f(u)\mathrm{El}_x u, \ u = g(x)$

5.8 $\mathrm{El}_x A = 0,\ \mathrm{El}_x x^a = a,\ \mathrm{El}_x e^x = x$.（$A$ 和 a 是常数，$A \neq 0$.）

弹性计算的特殊法则.

5.9 $$\mathrm{El}_x \sin x = x\cot x,\ \mathrm{El}_x \cos x = -x\tan x$$

弹性计算的特殊法则.

5.10 $$\mathrm{El}_x \tan x = \frac{x}{\sin x\cos x},\ \mathrm{El}_x \cot x = \frac{-x}{\sin x\cos x}$$

5.11 $$\mathrm{El}_x \ln x = 1/\ln x,\ \mathrm{El}_x \log_a x = 1/\ln x$$

5.12 $$\mathrm{El}_i f(\boldsymbol{x}) = \mathrm{El}_{x_i} f(\boldsymbol{x}) = \frac{x_i}{f(\boldsymbol{x})}\frac{\partial f(\boldsymbol{x})}{\partial x_i}$$

$f(\boldsymbol{x}) = f(x_1, \cdots, x_n)$ 对于 x_i，（$i = 1, \cdots, n$）的偏弹性.

5.13 如果 $z = F(x_1, \cdots, x_n)$ 且 $x_i = f_i(t_1, \cdots, t_m)$ 对于 $i = 1, \cdots, n$ 成立，则对于所有 $j = 1, \cdots, m$，有

$$\mathrm{El}_{t_j} z = \sum_{i=1}^{n} \mathrm{El}_i F(x_1, \cdots, x_n)\mathrm{El}_{t_j} x_i$$

弹性的链法则.

5.14 f 在 $\boldsymbol{x}$ 的方向弹性，在方向 $\boldsymbol{x}/\|\boldsymbol{x}\|$ 上，是

$$\mathrm{El}_a f(\boldsymbol{x}) = \frac{\|\boldsymbol{x}\|}{f(\boldsymbol{x})} f'_a(\boldsymbol{x}) = \frac{1}{f(\boldsymbol{x})}\nabla f(\boldsymbol{x})\cdot \boldsymbol{x}$$

$\mathrm{El}_a f(\boldsymbol{x})$ 是对应于 $\boldsymbol{x}$ 每一元素增加百分之一时 $f(\boldsymbol{x})$ 的近似百分比变化. $f'_a(\boldsymbol{x})$ 的定义见(4.27)—(4.28)

5.15 $$\mathrm{El}_a f(\boldsymbol{x}) = \sum_{i=1}^{n} \mathrm{El}_i f(\boldsymbol{x}),\ \boldsymbol{a} = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$$

一个有用的事实(Passus 方程).

5.16 $$R_{yx} = \frac{f'_1(x, y)}{f'_2(x, y)},\ f(x, y) = c$$

y 与 x 之间的边际替代率（简称为 MRS），R_{yx}，是当每单位 x 减少时，为停留在 f 的相同水平曲线上，y 的近似增加值.

5.17	• 当 f 是一效用函数，x 和 y 是商品时，R_{yx} 称为边际替代率(简称为 MRS). • 当 f 生产函数而 x 和 y 是要素投入时，R_{yx} 称为边际技术替代率(简称为 MRTS). • 当 $f(x, y) = 0$ 是隐含生产函数(对于给定要素投入)，而 x 和 y 是两个产品时，R_{yx} 称为边际产品转换率(简称为 MRPT).	(5.16)的不同形式.参见第 25 章和第 26 章.
5.18	$$\sigma_{yx} = \mathrm{El}_{R_{yx}}\left(\frac{y}{x}\right) = -\frac{\partial\ln\left(\frac{y}{x}\right)}{\partial\ln\left(\frac{f_2'}{f_1'}\right)},\ f(x, y) = c$$	y 与 x 之间的替代弹性 σ_{yx}，是当边际替代率改变百分之一，f 不变时，相对应的比率 y/x 的百分比变化.
5.19	$$\sigma_{yx} = \frac{\frac{1}{xf_1'} + \frac{1}{yf_2'}}{-\frac{f_{11}''}{(f_1')^2} + 2\frac{f_{12}''}{f_1'f_2'} - \frac{f_{22}''}{(f_2')^2}},$$ $f(x, y) = c$	替代弹性的另一种公式.
5.20	如果 $f(x, y)$的齐次度为 1，则 $$\sigma_{yx} = \frac{f_1'f_2'}{ff_{12}''}$$	一个特例.
5.21	$$\sigma_{ij} = -\frac{\partial\ln\left(\frac{x_i}{x_j}\right)}{\partial\ln\left(\frac{f_i'}{f_j'}\right)},\ f(x_1, \cdots, x_n) = c,\ i \neq j$$	n 个变量时的替代弹性.
5.22	$$\sigma_{ij} = \frac{\frac{1}{x_if_i'} + \frac{1}{x_jf_j'}}{-\frac{f_{ii}''}{(f_i')^2} + \frac{2f_{ij}''}{f_i'f_j'} - \frac{f_{jj}''}{(f_j')^2}},\ i \neq j$$	替代弹性，$f(x_1, \cdots, x_n) = c$.

参考文献

这些公式通常在微积分教材里找不到.对于(5.4)—(5.20)参照如 Sydsaeter & Hammond (1995).对于(5.21)—(5.22)参照 Blackorby & Russell (1989)及 Fuss & McFadden (1978).对于生产理论中的替代弹性,参照第 25 章.

6 方程组

6.1 $$\begin{aligned} f_1(x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_m) &= 0 \\ f_2(x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_m) &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ f_m(x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_m) &= 0 \end{aligned}$$

方程组的一般形式,具有 n 个外生变量, $x_1, \cdots, x_n$, 及 m 个内生变量, $y_1, \cdots, y_m$.

6.2 $$\frac{\partial \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})}{\partial \boldsymbol{y}} = \begin{pmatrix} \dfrac{\partial f_1}{\partial y_1} & \cdots & \dfrac{\partial f_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial y_1} & \cdots & \dfrac{\partial f_m}{\partial y_m} \end{pmatrix}$$

$f_1, \cdots, f_m$ 对于 $y_1, \cdots, y_m$ 的雅各比矩阵.

6.3 假设 $f_1, \cdots, f_m$ 是定义在 $\mathbf{R}^{n+m}$ 里的集合 A 内的 C^k 函数, $(\boldsymbol{x}^0, \boldsymbol{y}^0) = (x_1^0, \cdots, x_n^0, y_1^0, \cdots, y_m^0)$ 为(6.1)的在 A 内的一个解. 同时假设(6.2)中的在 $(\boldsymbol{x}^0, \boldsymbol{y}^0)$ 点的雅各比矩阵 $\partial \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})/\partial \boldsymbol{y}$ 的行列式不等于 0. 则(6.1)定义 $y_1, \cdots, y_m$ 为在 $(\boldsymbol{x}^0, \boldsymbol{y}^0)$ 的某一邻域内的 $x_1, \cdots, x_n$ 的 C^k 函数,而且,对于 $j = 1, \cdots, n$

$$\begin{pmatrix} \dfrac{\partial y_1}{\partial x_j} \\ \vdots \\ \dfrac{\partial y_m}{\partial x_j} \end{pmatrix} = - \begin{pmatrix} \dfrac{\partial f_1}{\partial y_1} & \cdots & \dfrac{\partial f_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial y_1} & \cdots & \dfrac{\partial f_m}{\partial y_m} \end{pmatrix}^{-1} \begin{pmatrix} \dfrac{\partial f_1}{\partial x_j} \\ \vdots \\ \dfrac{\partial f_m}{\partial x_j} \end{pmatrix}$$

广义隐函数定理(它是(6.1)中定义内生变量 $y_1, \cdots, y_m$ 为外生变量 $x_1, \cdots, x_n$ 的可微函数的充分条件.)

6.4	$f_1(x_1, x_2, \cdots, x_n) = 0$ $f_2(x_1, x_2, \cdots, x_n) = 0$ …………………… $f_m(x_1, x_2, \cdots, x_n) = 0$	m 个方程,n 个变量的一般方程组.
6.5	如果存在一组 k 个变量可自由选择,并且当 k 个变量给予特定值时,剩余的 $n-k$ 个具有单一解,则方程组(6.4)具有 k 个自由度.如果变量取值限定在 $\mathbf{R}^n$ 中的集合 S 里,则方程组在 S 里有 k 个自由度.	方程组的自由度的定义.
6.6	要寻找方程组的自由度,数一下变量的个数 n,和方程的个数 m.如果 $n>m$,则方程组有 $n-m$ 个自由度.如果 $n<m$,则在一般情况下方程组无解.	“计数法则”.这是一个粗略法则,并不总是成立.
6.7	如果(6.3)中的条件得到满足,则方程组(6.1)有 n 个自由度.	(局部)计数法则.
6.8	$\mathbf{f}'(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial f_1(\mathbf{x})}{\partial x_1} & \cdots & \dfrac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_m(\mathbf{x})}{\partial x_1} & \cdots & \dfrac{\partial f_m(\mathbf{x})}{\partial x_n} \end{pmatrix}$	对于 $x_1, \cdots, x_n$ 的 $f_1, \cdots, f_m$ 的雅各比矩阵,也可写为 $\partial \mathbf{f}(\mathbf{x})/\partial \mathbf{x}$.(参见(6.2))
6.9	如果 $\mathbf{x}^0 = (x_1^0, \cdots, x_n^0)$ 是(6.4)的一个解,$m \leqslant n$,雅各比矩阵的秩 $\mathbf{f}'(\mathbf{x})$ 等于 m,则方程组(6.4)在 $\mathbf{x}^0$ 的某一邻域有 $n-m$ 个自由度.	(局部)计数法则(当 $f_1, \cdots, f_m$ 是 C^1 函数时成立).
6.10	如果存在一个实 C^1 函数 F 定义在包含 $S = \{(f_1(\mathbf{x}), \cdots, f_m(\mathbf{x})) : \mathbf{x} \in A\}$ 的某一开集内,使 $F(f_1(\mathbf{x}), \cdots, f_m(\mathbf{x})) = 0$ 对于所有 $\mathbf{x} \in A$ 成立,其中 $\nabla F \neq \mathbf{0}$ 在 S 内,则函数 $f_1(\mathbf{x}), \cdots, f_m(\mathbf{x})$ 是在 $\mathbf{R}^n$ 里的开集合 A 内函数相关的.	函数相关的定义.(参见 Marsden & Hoffman (1993))
6.11	如果 $f_1(\mathbf{x}), \cdots, f_m(\mathbf{x})$ 在开集 $A \subset \mathbf{R}^n$ 内是函数相关的,则雅各比矩阵 $\mathbf{f}'(\mathbf{x})$ 的秩小于 m.	函数相关的必要条件.

6.12 如果方程组(6.4)有解,且如果 $f_1(\boldsymbol{x}), \cdots, f_m(\boldsymbol{x})$ 是函数相关的,则(6.4)至少有一个多余方程. | 计数法则无效的充分条件.

6.13 $$\det(\boldsymbol{f}'(\boldsymbol{x})) = \begin{vmatrix} \frac{\partial f_1(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial f_1(\boldsymbol{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial f_n(\boldsymbol{x})}{\partial x_n} \end{vmatrix}$$ | $f_1, \cdots, f_n$ 对于 $x_1, \cdots, x_n$ 的雅各比行列式.(行列式请参见第 20 章)

6.14 如果 $f_1(\boldsymbol{x}), \cdots, f_n(\boldsymbol{x})$ 是函数相关的,则行列式 $\det(\boldsymbol{f}(\boldsymbol{x})) \equiv 0$ | (6.11)的一个特例.逆定理一般并不成立.

6.15 $$\left.\begin{array}{l} y_1 = f_1(x_1, \cdots, x_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ y_n = f_n(x_1, \cdots, x_n) \end{array}\right. \Leftrightarrow \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$$ | 从 $\mathbf{R}^n \to \mathbf{R}^n$ 的转换 $\boldsymbol{f}$.

6.16 假设在(6.15)中的转换 $\boldsymbol{f}$ 是在 $\boldsymbol{x}^0$ 的邻域内的 C^1 转换,而且(6.13)中的雅各比行列式在 $\boldsymbol{x}^0$ 非零.则存在一个 C^1 转换 $\boldsymbol{g}$,是 $\boldsymbol{f}$ 的局部逆转换,即 $\boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x})) = \boldsymbol{x}$ 对于所有 $\boldsymbol{x}$ 在 $\boldsymbol{x}^0$ 的某一邻域成立. | 局部逆转换的存在.

6.17 假设 $\boldsymbol{f}: \mathbf{R}^n \to \mathbf{R}^n$ 是 C^1,且存在正数 h 和 k 使 $|\det(\boldsymbol{f}'(\boldsymbol{x}))| \geqslant h > 0$ 且 $|\partial f_i(\boldsymbol{x})/\partial x_j| \leqslant k$ 对于所有 $\boldsymbol{x}$ 和所有 $i, j = 1, \cdots, n$. 则 $\boldsymbol{f}$ 有一个定义在所有 $\mathbf{R}^n$ 内的 C^1 逆. | 广义逆转换的存在.(Hadamard)

6.18 假设 $\boldsymbol{f}: \mathbf{R}^n \to \mathbf{R}^n$ 是 C^1 转换而且对于所有 $\boldsymbol{x}$ 在(6.13)中的行列式 $\neq 0$. 则 $\boldsymbol{f}(\boldsymbol{x})$ 有一个逆是 C^1 函数且定义在所有 $\mathbf{R}^n$,当且仅当 $\inf\{\|\boldsymbol{f}(\boldsymbol{x})\| : \|\boldsymbol{x}\| \geqslant n\} \to \infty$ 当 $n \to \infty$ | 广义反函数定理.

6.19 假设 $\boldsymbol{f}: \mathbf{R}^n \to \mathbf{R}^n$ 是 C^1,让 Ω 为长方形 $\Omega = \{\boldsymbol{x} \in \mathbf{R}^n : \boldsymbol{a} \leqslant \boldsymbol{x} \leqslant \boldsymbol{b}\}$,其中 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 是在 $\mathbf{R}^n$ 中的给定向量.则 $\boldsymbol{f}$ 在 Ω 中一一对应,只要以下条件有一个对于所有 $\boldsymbol{x}$ 得到满足:

- 雅各比矩阵 $\boldsymbol{f}'(\boldsymbol{x})$ 仅有严格为正的主子式.
- 雅各比矩阵 $\boldsymbol{f}'(\boldsymbol{x})$ 仅有严格为负的主子式.

| *Gale-Nikaido* 定理.(对于主子式,参见(20.15))

6.20 一个 $n \times n$ 矩阵 $\mathbf{A}$(不一定是对称的)称为正准定性的，如果对于每一 n 维向量 $\mathbf{x} \neq \mathbf{0}$，$\mathbf{x}'\mathbf{A}\mathbf{x} > 0$.

正准定性的定义.

6.21 假设 $f:\mathbf{R}^n \to \mathbf{R}^n$ 是一个 C^1 函数，且假设雅各比矩阵 $f'(\mathbf{x})$ 是在一个凸集合 Ω 内正准定性的. 则 f 是在 Ω 内的一一对应.

Gale-Nikaido 定理.

6.22 $f:\mathbf{R}^n \to \mathbf{R}^n$ 称为是一个收缩映射如果存在常数 $k \in [0, 1)$ 使

$$\| f(\mathbf{x}) - f(\mathbf{y}) \| \leqslant k \| \mathbf{x} - \mathbf{y} \|$$

对于所有 $\mathbf{x} \in \mathbf{R}^n$ 和 $\mathbf{y}$ 在 $\mathbf{R}^n$ 成立.

收缩映射的定义.

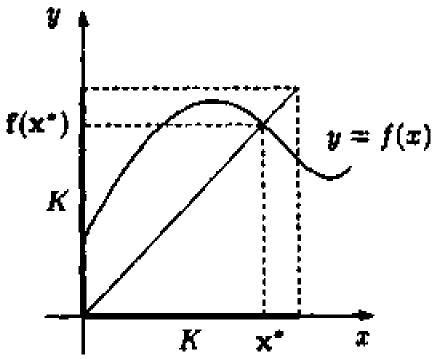
6.23 如果 $f:\mathbf{R}^n \to \mathbf{R}^n$ 是一个收缩映射，则 f 有一个唯一的不动点，即有一点 $\mathbf{x}^* \in \mathbf{R}^n$ 使 $f(\mathbf{x}^*) = \mathbf{x}^*$. 对于任一 $\mathbf{x}_0 \in \mathbf{R}^n$ 我们有 $\mathbf{x}^* = \lim_{n\to\infty} \mathbf{x}_n$，其中 $\mathbf{x}_n = f(\mathbf{x}_{n-1})$ 对于 $n \geqslant 1$.

收缩映射中不动点的存在.(这一结论可一般化到完整的度量空间里. 参见(18.26))

6.24 设 K 为一个在 $\mathbf{R}^n$ 中的非空紧凸集合，f 为一个从 K 到 K 映射的函数，则 f 有一不动点 $\mathbf{x}^* \in K$，即有点 $\mathbf{x}^*$ 使 $f(\mathbf{x}^*) = \mathbf{x}^*$

Brouwer 不动点定理.

6.25



Brouwer 不动点定理当 $n = 1$ 时的图示.

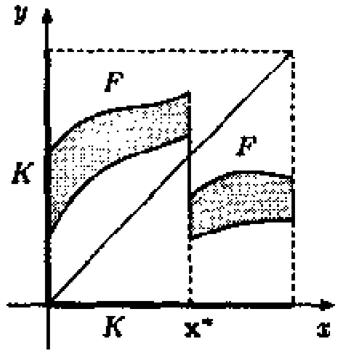
6.26 K 为一个在 $\mathbf{R}^n$ 中的紧凸集合，f 为一个对应，使 K 中每一点 $\mathbf{x}$ 与一个 K 的非空凸子集 $f(\mathbf{x})$ 对应. 假设 f 是一个闭图形，即集合

$$\{(\mathbf{x}, \mathbf{y}) \in \mathbf{R}^{2n} : \mathbf{x} \in K \text{ 及 } \mathbf{y} \in f(\mathbf{x})\}$$

是在 $\mathbf{R}^{2n}$ 内封闭的. 则 f 有一不动点，即一点 $\mathbf{x}^* \in K$，使 $\mathbf{x}^* \in f(\mathbf{x}^*)$.

Kakutani 不动点定理.(参见(12.35)关于对应的定义)

6.27



Kakutani 不动点定理当 $n = 1$ 时的图示.

6.28 如果 $\boldsymbol{x} = (x_1, \cdots, x_n)$ 和 $\boldsymbol{y} = (y_1, \cdots, y_n)$ 是在 $\mathbf{R}^n$ 中的两点,则 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的相交 $\boldsymbol{x} \wedge \boldsymbol{y}$ 和相并 $\boldsymbol{x} \vee \boldsymbol{y}$ 定义为:

$$\boldsymbol{x} \wedge \boldsymbol{y} = (\min\{x_1, y_1\}, \cdots, \min\{x_n, y_n\})$$

$$\boldsymbol{x} \vee \boldsymbol{y} = (\max\{x_1, y_1\}, \cdots, \max\{x_n, y_n\})$$

$\mathbf{R}^n$ 中两个向量相交与相并的定义.

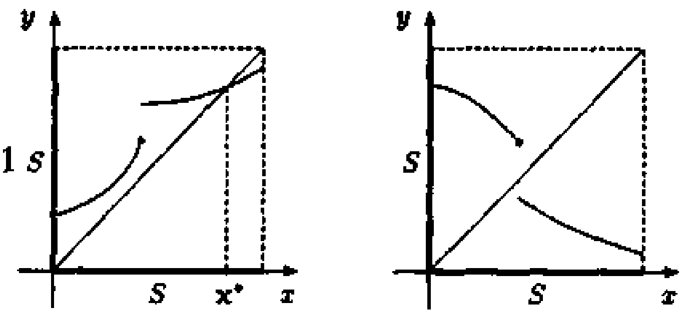
6.29 在 $\mathbf{R}^n$ 内的集合 S 称为 $\mathbf{R}^n$ 的一个子格,如果 S 中的任何两点的相交和相并都在 S 内.如果 S 同时也是一个紧集合,则 S 称为一个紧子格.

$\mathbf{R}^n$ 的(紧)子格的定义.

6.30 设 S 为 $\mathbf{R}^n$ 的一个非空紧子格.若 $\boldsymbol{f}: S \to S$ 为一个递增函数,例如 $\boldsymbol{x}, \boldsymbol{y} \in S$ 当 $\boldsymbol{x} \leqslant \boldsymbol{y}$ 时隐含 $\boldsymbol{f}(\boldsymbol{x}) \leqslant \boldsymbol{f}(\boldsymbol{y})$.则,$\boldsymbol{f}$ 在 S 上有一不动点,即有点 $\boldsymbol{x}^* \in S$ 使 $\boldsymbol{f}(\boldsymbol{x}^*) = \boldsymbol{x}^*$

Tarski 不动点定理.(这一定理不适用于递减函数.参见(6.31))

6.31



x^* 是左图中递增函数的不动点.另一图中的递减函数无不动点.

6.32

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{gathered}$$

m 个方程,n 个未知数的一般线性方程组.

6.33 $$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

$$\mathbf{A}_b = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}$$

$\mathbf{A}$ 是(6.32)的系数矩阵，$\mathbf{A}_b$ 是“扩展系数矩阵”.

6.34
- 当且仅当 $r(\mathbf{A}) = r(\mathbf{A}_b)$ 时，方程组(6.32)至少有一个解.
- 如果 $r(\mathbf{A}) = r(\mathbf{A}_b) = k < m$，则方程组(6.32)有 $m - k$ 个多余方程.
- 如果 $r(\mathbf{A}) = r(\mathbf{A}_b) = k < n$，则方程组(6.32)有 $n - k$ 个自由度.

线性方程组的主要结论，$r(\mathbf{B})$是矩阵 $\mathbf{B}$ 的秩.（参见(19.23)）

6.35
$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

m 个方程，n 个未知数的一般齐次线性方程组.

6.36
- 当且仅当 $r(\mathbf{A}) < n$，齐次方程组(6.35)有一非零解.
- 如果 $n = m$，则当且仅当 $|\mathbf{A}| = 0$ 时，齐次方程组(6.35)有非零解.

齐次线性方程组的重要结论.

参 考 文 献

对于(6.1)—(6.16)和(6.22)—(6.23)，参见 Rudin (1982)，或 Marsden & Hoffman (1993). 对于(6.17)—(6.21)参见 Parthasarathy (1983). 对于 Brouwer 和 Kakutani 的不动点定理，参见 Nikaido (1970)或 Scarf (1973). 对于 Tarski 的不动点定理以及相关材料，参见 Sundaram (1996). (6.34)—(6.36)是线性代数中的一般结论，参见 Fraleigh & Beauregard (1995)或 Lang (1987).

7 不等式

7.1 $\big||a|-|b|\big| \leqslant |a \pm b| \leqslant |a|+|b|$

三角不等式. $a, b \in \mathrm{R}(\text{or } \mathrm{C})$.

7.2 $\dfrac{n}{\sum_{i=1}^{n} 1/a_i} \leqslant \left(\prod_{i=1}^{n} a_i\right)^{1/n} \leqslant \dfrac{\sum_{i=1}^{n} a_i}{n}$, $a_i > 0$

调和平均≤几何平均≤算术平均. 等号当且仅当 $a_1 = \cdots = a_n$ 时成立.

7.3 $\dfrac{2}{1/a_1 + 1/a_2} \leqslant \sqrt{a_1 a_2} \leq \dfrac{a_1 + a_2}{2}$

(7.2)当 $n = 2$.

7.4 $a_1^{\lambda_1} \cdots a_n^{\lambda_n} \leqslant \lambda_1 a_1 + \cdots + \lambda_n a_n$

加权平均的不等式, $a_i \geqslant 0$, $\sum_{i=1}^{n} \lambda_i = 1$, $\lambda_i \geqslant 0$.

7.5 $a_1^{\lambda} a_2^{1-\lambda} \leqslant \lambda a_1 + (1-\lambda) a_2$

(7.4)中 $n = 2$, $a_1 \geqslant 0$, $a_2 \geqslant 0$, $\lambda \in [0, 1]$.

7.6 $\sum_{i=1}^{n} |a_i b_i| \leqslant \left[\sum_{i=1}^{n} |a_i|^p\right]^{1/p} \left[\sum_{i=1}^{n} |b_i|^q\right]^{1/q}$

Hölder 不等式. $p, q > 1$, $1/p + 1/q = 1$. 等号只有当 $|b_i| = c|a_i|^{p-1}$ 对于一个非负数 c 成立时成立.

7.7 $\left[\sum_{i=1}^{n} |a_i b_i|\right]^2 \leqslant \left[\sum_{i=1}^{n} a_i^2\right] \left[\sum_{i=1}^{n} b_i^2\right]$

柯西-施瓦兹不等式. ($p = q = 2$ 代入(7.6))

7.8	$\left[\sum_{i=1}^{n} a_i\right]\left[\sum_{i=1}^{n} b_i\right] \leqslant n \sum_{i=1}^{n} a_i b_i$	切比雪夫不等式. $a_1 \geqslant \cdots \geqslant a_n$ $b_1 \geqslant \cdots \geqslant b_n$
7.9	$\left[\sum_{i=1}^{n} \lvert a_i + b_i \rvert^p\right]^{1/p} \leqslant \left[\sum_{i=1}^{n} \lvert a_i \rvert^p\right]^{1/p} + \left[\sum_{i=1}^{n} \lvert b_i \rvert^p\right]^{1/p}$	*Minkowski* 不等式. $p \geqslant 1$. 等号只有当 $\lvert b_i \rvert = c \lvert a_i \rvert$ 对于一个非负数 c 成立时成立.
7.10	如果 f 是凸的,则 $f\left[\sum_{i=1}^{n} a_i x_i\right] \leqslant \sum_{i=1}^{n} a_i f(x_i)$	*Jensen* 不等式. $\sum_{i=1}^{n} a_i = 1$, $a_i \geqslant 0$, $i = 1, \cdots, n$.
7.11	$\left[\sum_{i=1}^{n} \lvert a_i \rvert^q\right]^{1/q} \leqslant \left[\sum_{i=1}^{n} \lvert a_i \rvert^p\right]^{1/p}$	另一 *Jensen* 不等式; $0 < p < q$.
7.12	$\int_a^b \lvert f(x) g(x) \rvert \mathrm{d}x \leqslant \left[\int_a^b \lvert f(x) \rvert^p \mathrm{d}x\right]^{1/p} \left[\int_a^b \lvert g(x) \rvert^q \mathrm{d}x\right]^{1/q}$	*Hölder* 不等式. $p > 1$, $q > 1$, $1/p + 1/q = 1$. 等号只有当 $\lvert g(x) \rvert = c \lvert f(x) \rvert^{p-1}$ 对于一个非负数 c 成立时成立.
7.13	$\left[\int_a^b f(x) g(x) \mathrm{d}x\right]^2 \leqslant \int_a^b (f(x))^2 \mathrm{d}x \int_a^b (g(x))^2 \mathrm{d}x$	柯西-施瓦兹不等式.
7.14	$\left[\int_a^b \lvert f(x) + g(x) \rvert^p \mathrm{d}x\right]^{1/p} \leqslant \left[\int_a^b \lvert f(x) \rvert^p \mathrm{d}x\right]^{1/p} + \left[\int_a^b \lvert g(x) \rvert^p \mathrm{d}x\right]^{1/p}$	*Minkowski* 不等式. $p \geqslant 1$. 等号只有当 $g(x) = c f(x)$ 对于一个非负数 c 成立时成立.

7.15 如果 f 是凸的,则

$$f\left(\int a(x)g(x)\mathrm{d}x\right)\leqslant\int a(x)f(g(x))\mathrm{d}x$$

Jensen 不等式. $a(x)\geqslant 0$, $f(u)\geqslant 0$, $\int a(x)\mathrm{d}x=1$. f 定义在 g 的值域里.

7.16 如果 f 在 I 区间里是凸的,X 是一个有有限期望值的随机变量,则

$$f(E[X])\leqslant E[f(X)]$$

如果 f 是严格凸的,不等号严格成立,除非 X 是概率为 1 的常数.

Jensen 不等式的特殊情况. E 是期望值符号.

7.17 如果 U 在 I 区间里是凹的,X 是一个有有限期望值的随机变量,则

$$E[U(X)]\leqslant U(E[X])$$

效用理论的一个重要现象.(在(7.16)中代入 $f=-U$ 得到.)

参 考 文 献

关于不等式的较好的参考文献仍为 Hardy, Littlewood, & Pólya (1952).

8 级数 泰勒公式

8.1 $\sum_{i=0}^{n-1}(a+id)=na+\frac{n(n-1)d}{2}$

算术级数的前 n 项之和.

8.2 $a+ak+ak^2+\cdots+ak^{n-1}=a\frac{1-k^n}{1-k},\ k\neq 1$

几何级数的前 n 项之和.

8.3 $a+ak+\cdots+ak^{n-1}+\cdots=\frac{a}{1-k}$

如果 $|k|<1$

无穷几何级数之和.

8.4 $a_1+a_2+\cdots+a_n+\cdots=s$ 指

$$\lim_{n\to\infty}(a_1+a_2+\cdots+a_n)=s$$

无穷级数收敛的定义.如果级数不收敛,则是发散的.

8.5 $a_1+\cdots+a_n+\cdots$ 收敛$\Rightarrow\lim_{n\to\infty}a_n=0$

无穷级数收敛的一个必要(但不是充分)条件.

8.6 $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|<1\Rightarrow\sum_{n=1}^{\infty}a_n$ 收敛

比率测试.

8.7 $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|>1\Rightarrow\sum_{n=1}^{\infty}a_n$ 发散

比率测试.

8.8 如果 $\sum a_n$ 是一个仅有正项的级数,$f(x)$是一个定义在 $x\geqslant 1$ 的正值递减的连续函数,且如果 $f(n)=a_n$ 对于所有整数 $n\geqslant 1$ 成立,则无穷级数和广义积分

$$\sum_{n=1}^{\infty}a_n \text{ 和} \int_1^{\infty}f(x)\mathrm{d}x$$

或都收敛,或都发散.

积分测试.

8.9 如果 $0 \leqslant a_n \leqslant b_n$ 对于所有 n 成立，则 • $\sum a_n$ 收敛，如果 $\sum b_n$ 收敛. • $\sum b_n$ 发散，如果 $\sum a_n$ 发散.	比较测试.
8.10 $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ 是收敛的 $\Leftrightarrow p > 1$	一个重要的结论.
8.11 $f(x) \approx f(a) + f'(a)(x-a)$ （x 在 a 附近）	在 $x = a$ 点的一阶（线性）近似.
8.12 $f(x) \approx f(a) + f'(a)(x-a) + \frac{1}{2} f''(a)(x-a)^2$ （x 在 a 附近）	在 $x = a$ 点的二阶（二次）近似.
8.13 $f(x) = f(0) + f'(0)\frac{x}{1!} + \cdots + f^{(n)}(0)\frac{x^n}{n!} + f^{(n+1)}(\theta x)\frac{x^{n+1}}{(n+1)!},\ 0 < \theta < 1$	Maclaurin 公式. 最后一项为拉格朗日误差项.
8.14 $f(x) = f(0) + f'(0)\frac{x}{1!} + f''(0)\frac{x^2}{2!} + \cdots$	$f(x)$ 的 Maclaurin 级数，对于当 n 接近 ∞，(8.13) 中的误差项趋向于 0 的 x 成立.
8.15 $f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(a+\theta(x-a))}{(n+1)!}(x-a)^{n+1},\ 0 < \theta < 1$	泰勒公式. 最后一项是拉格朗日误差项.
8.16 $f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots$	$f(x)$ 的泰勒级数，对于 (8.15) 中当 n 接近于 ∞ 时误差项接近于 0 的 x 成立.

8.17	$f(x, y) \approx f(a, b) + f_1'(a, b)(x-a) + f_2'(a, b)(y-b)$ （(x, y)接近于(a, b)）	$f(x, y)$在(a, b)附近的一阶(线性)近似.
8.18	$f(x, y) \approx f(a, b) + f_1'(a, b)(x-a) + f_2'(a, b)(y-b) + \frac{1}{2}[f_{11}''(a, b)(x-a)^2 + 2f_{12}''(a, b)(x-a)(y-b) + f_{22}''(a, b)(y-b)^2]$	$f(x, y)$在(a, b)附近的二阶(二次)近似.
8.19	$f(\mathbf{x}) = f(\mathbf{a}) + \sum_{i=1}^{n} f_i'(\mathbf{a})(x_i - a_i) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}''(\mathbf{a} + \theta(\mathbf{x} - \mathbf{a}))(x_i - a_i)(x_j - a_j)$	有 n 个变量的函数的泰勒公式, $\theta \in (0, 1)$.
8.20	$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$	对于所有 x 成立.
8.21	$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$	对于 $-1 < x \leqslant 1$ 成立.
8.22	$(1+x)^m = \binom{m}{0} + \binom{m}{1}x + \binom{m}{2}x^2 + \cdots$	对于 $-1 < x < 1$ 成立. $\binom{m}{k}$的定义参见(8.27).
8.23	$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \cdots$	对于所有 x 成立.
8.24	$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots$	对于所有 x 成立.
8.25	$\arcsin x = x + \frac{1}{2}\frac{x^3}{3} + \frac{1 \cdot 3}{2 \cdot 4}\frac{x^5}{5} + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\frac{x^7}{7} + \cdots$	对于 $\mid x \mid \leqslant 1$ 成立.
8.26	$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$	对于 $\mid x \mid \leqslant 1$ 成立.
8.27	$\binom{m}{k} = \frac{m(m-1)\cdots(m-k+1)}{k!}, \binom{m}{0} = 1$	二项式系数. (m 是一任意实数, k 是一自然数)

8.28	$\sum_{i=0}^{n}\binom{n}{i}a^{n-i}b^{i}=(a+b)^{n}$	Newton 二项式公式.
8.29	$\sum_{i=0}^{n}\binom{n}{i}=(1+1)^{n}=2^{n}$	(8.28)的特殊情况.
8.30	$\sum_{i=r}^{n}\binom{i}{r}=\binom{n+1}{r+1}$	二项式系数的特性.
8.31	$\sum_{i=0}^{j}\binom{n}{i}\binom{m}{j-i}=\binom{n+m}{j}$	
8.32	$\sum_{i=0}^{n}i\binom{n}{i}=n2^{n-1}$	
8.33	$\sum_{i=0}^{n}i^{2}\binom{n}{i}=(n^{2}+n)2^{n-2}$	
8.34	$\sum_{i=0}^{n}\binom{n}{i}^{2}=\binom{2n}{n}$	
8.35	$(a_1+a_2+\cdots+a_m)^n=\sum_{k_1+\cdots+k_m=n}\frac{n!}{k_1!\cdots k_m!}a_1^{k_1}\cdots a_m^{k_m}$	多项式公式.
8.36	$1+2+3+\cdots+n=\frac{n(n+1)}{2}$	求和公式.
8.37	$1+3+5+\cdots+(2n-1)=n^2$	
8.38	$1^2+2^2+3^2+\cdots+n^2=\frac{n(n+1)(2n+1)}{6}$	
8.39	$1^3+2^3+3^3+\cdots+n^3=\left(\frac{n(n+1)}{2}\right)^2$	
8.40	$1^4+2^4+3^4+\cdots+n^4=\frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}$	
8.41	$\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{n^2}+\cdots=\frac{\pi^2}{6}$	一个著名的结论.
8.42	$\lim_{n\to\infty}\left[\left(\frac{1}{1}+\frac{1}{2}+\cdots+\frac{1}{n}\right)-\ln n\right]=\gamma\approx 0.5772\cdots$	常数 γ 称为欧拉常数.

参考文献

所有这些公式都是标准的，通常可在微积分教材里找到，例如 Edwards & Penney (1998). 对于二项式系数，请参考概率理论教材.

9 积分

不定积分

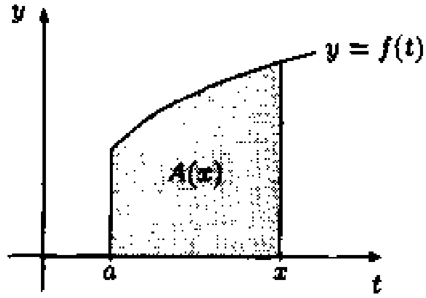
9.1	$\int f(x)\mathrm{d}x = F(x) + C \Leftrightarrow F'(x) = f(x)$	不定积分的定义.
9.2	$\int (af(x) + bg(x))\mathrm{d}x = a\int f(x)\mathrm{d}x + b\int g(x)\mathrm{d}x$	积分的线性性. a 和 b 是常数.
9.3	$\int f(x)g'(x)\mathrm{d}x = f(x)g(x) - \int f'(x)g(x)\mathrm{d}x$	分部积分法.
9.4	$\int f(x)\mathrm{d}x = \int f(g(t))g'(t)\mathrm{d}t,\ x = g(t)$	改变变量(变换积分法).
9.5	$\int x^n\mathrm{d}x = \begin{cases} \dfrac{x^{n+1}}{n+1} + C, & n \neq -1 \\ \ln\|x\| + C, & n = -1 \end{cases}$	特殊积分.
9.6	$\int a^x\mathrm{d}x = \dfrac{1}{\ln a}a^x + C,\ a > 0,\ a \neq 1$	
9.7	$\int \mathrm{e}^x\mathrm{d}x = \mathrm{e}^x + C$	
9.8	$\int x\mathrm{e}^x\mathrm{d}x = x\mathrm{e}^x - \mathrm{e}^x + C$	
9.9	$\int x^n\mathrm{e}^{ax}\mathrm{d}x = \dfrac{x^n}{a}\mathrm{e}^{ax} - \dfrac{n}{a}\int x^{n-1}\mathrm{e}^{ax}\mathrm{d}x,\ a \neq 0$	
9.10	$\int \log_a x\mathrm{d}x = x\log_a x - x\log_a \mathrm{e} + C,\ a > 0,\ a \neq 1$	
9.11	$\int \ln x\mathrm{d}x = x\ln x - x + C$	

9.12 $\int x^n \ln x\,dx = \frac{x^{n+1}((n+1)\ln x - 1)}{(n+1)^2} + C$	特殊积分.（$n \neq -1$）
9.13 $\int \sin x\,dx = -\cos x + C$	
9.14 $\int \cos x\,dx = \sin x + C$	
9.15 $\int \tan x\,dx = -\ln\|\cos x\| + C$	
9.16 $\int \cot x\,dx = \ln\|\sin x\| + C$	
9.17 $\int \frac{1}{\sin x}dx = \ln\left\|\frac{1-\cos x}{\sin x}\right\| + C$	
9.18 $\int \frac{1}{\cos x}dx = \ln\left\|\frac{1+\sin x}{\cos x}\right\| + C$	
9.19 $\int \frac{1}{\sin^2 x}dx = -\cot x + C$	
9.20 $\int \frac{1}{\cos^2 x}dx = \tan x + C$	
9.21 $\int \sin^2 x\,dx = \frac{1}{2}x - \frac{1}{2}\sin x\cos x + C$	
9.22 $\int \cos^2 x\,dx = \frac{1}{2}x + \frac{1}{2}\sin x\cos x + C$	
9.23 $\int \sin^n x\,dx = -\frac{\sin^{n-1}x\cos x}{n} + \frac{n-1}{n}\int \sin^{n-2}x\,dx$	（$n \neq 0$）
9.24 $\int \cos^n x\,dx = \frac{\cos^{n-1}x\sin x}{n} + \frac{n-1}{n}\int \cos^{n-2}x\,dx$	（$n \neq 0$）
9.25 $\int e^{\alpha x}\sin\beta x\,dx = \frac{e^{\alpha x}}{\alpha^2+\beta^2}(\alpha\sin\beta x - \beta\cos\beta x) + C$	（$\alpha^2+\beta^2 \neq 0$）
9.26 $\int e^{\alpha x}\cos\beta x\,dx = \frac{e^{\alpha x}}{\alpha^2+\beta^2}(\beta\sin\beta x + \alpha\cos\beta x) + C$	（$\alpha^2+\beta^2 \neq 0$）

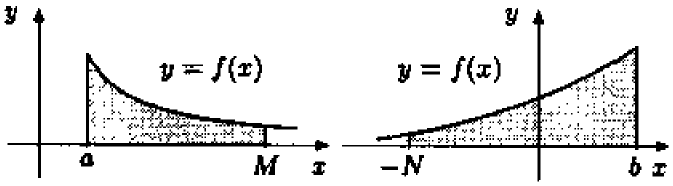
9.27	$\int \frac{1}{x^2-a^2}\mathrm{d}x = \frac{1}{2a}\ln\left\|\frac{x-a}{x+a}\right\| + C$	特殊积分.($a\neq 0$)
9.28	$\int \frac{1}{x^2+a^2}\mathrm{d}x = \frac{1}{a}\arctan\frac{x}{a} + C$	($a\neq 0$)
9.29	$\int \frac{1}{\sqrt{a^2-x^2}}\mathrm{d}x = \arcsin\frac{x}{a} + C$	($a>0$)
9.30	$\int \frac{1}{\sqrt{x^2\pm a^2}}\mathrm{d}x = \ln\left\|x+\sqrt{x^2\pm a^2}\right\| + C$	
9.31	$\int \sqrt{a^2-x^2}\mathrm{d}x = \frac{x}{2}\sqrt{a^2-x^2} + \frac{a^2}{2}\arcsin\frac{x}{a} + C$	($a>0$)
9.32	$\int \sqrt{x^2\pm a^2}\mathrm{d}x = \frac{x}{2}\sqrt{x^2\pm a^2} \pm \frac{a^2}{2}\ln\left\|x+\sqrt{x^2\pm a^2}\right\| + C$	
9.33	$\int \frac{\mathrm{d}x}{ax^2+2bx+c} = \frac{1}{2\sqrt{b^2-ac}}\ln\left\|\frac{ax+b-\sqrt{b^2-ac}}{ax+b+\sqrt{b^2-ac}}\right\| + C$	($b^2>ac$, $a\neq 0$)
9.34	$\int \frac{\mathrm{d}x}{ax^2+2bx+c} = \frac{1}{\sqrt{ac-b^2}}\arctan\frac{ax+b}{\sqrt{ac-b^2}} + C$	($b^2<ac$)
9.35	$\int \frac{\mathrm{d}x}{ax^2+2bx+c} = \frac{-1}{ax+b} + C$	($b^2=ac$, $a\neq 0$)

定积分

9.36	$\int_a^b f(x)\mathrm{d}x = \Big\|_a^b F(x) = F(b)-F(a)$ 如果 $F'(x)=f(x)$ 对于所有在$[a,b]$中的 x 成立.	函数 f 的定积分的定义.

9.37	• $A(x)=\int_a^x f(t)\mathrm{d}t \Rightarrow A'(x)=f(x)$ • $A(x)=\int_x^b f(t)\mathrm{d}t \Rightarrow A'(x)=-f(x)$	重要事实.
9.38		阴影面积为 $A(x)=\int_a^x f(t)\mathrm{d}t$，面积函数 $A(x)$的导数是 $A'(x)=f(x)$.
9.39	$\int_a^b f(x)\mathrm{d}x=-\int_b^a f(x)\mathrm{d}x$ $\int_a^a f(x)\mathrm{d}x=0$ $\int_a^b \alpha f(x)\mathrm{d}x=\alpha\int_a^b f(x)\mathrm{d}x$ $\int_a^b f(x)\mathrm{d}x=\int_a^c f(x)\mathrm{d}x+\int_c^b f(x)\mathrm{d}x$	a, b, c 和 α 是任意实数.
9.40	$\int_a^b f(g(x))g'(x)\mathrm{d}x=\int_{g(a)}^{g(b)} f(u)\mathrm{d}u,\quad u=g(x)$	改变变量(变换积分).
9.41	$\int_a^b f(x)g'(x)\mathrm{d}x=\Big\vert_a^b f(x)g(x)-\int_a^b f'(x)g(x)\mathrm{d}x$	分部积分法.
9.42	$\int_a^\infty f(x)\mathrm{d}x=\lim_{M\to\infty}\int_a^M f(x)\mathrm{d}x$	如果极限存在，积分是收敛的.(相反的情况，积分是发散的.)
9.43	$\int_{-\infty}^b f(x)\mathrm{d}x=\lim_{N\to\infty}\int_{-N}^b f(x)\mathrm{d}x$	如果极限存在，积分是收敛的.(相反的情况，积分是发散的.)

9.44



(9.42)和(9.43)的图示. 左图中阴影面积为$\int_a^M f(x)\mathrm{d}x$，右图中阴影面积为$\int_{-N}^b f(x)\mathrm{d}x$.

9.45 $$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = \int_{-\infty}^{a} f(x)\mathrm{d}x + \int_{a}^{+\infty} f(x)\mathrm{d}x = \lim_{N\to\infty}\int_{-N}^{a} f(x)\mathrm{d}x + \lim_{M\to\infty}\int_{a}^{M} f(x)\mathrm{d}x$$

如果等号右边的极限一定存在,且 a 是任一数,则称积分为收敛的.(如果有一个极限不存在,则积分为发散的)

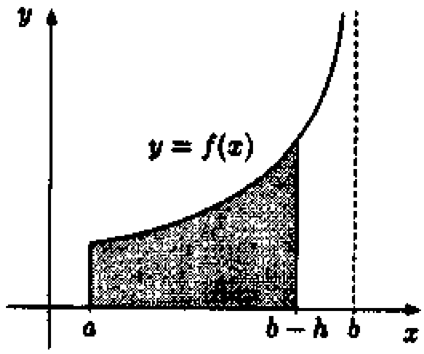
9.46 $$\int_a^b f(x)\mathrm{d}x = \lim_{h\to 0^+}\int_{a+h}^{b} f(x)\mathrm{d}x$$

在区间$(a, b]$内连续函数 f 的积分的定义.

9.47 $$\int_a^b f(x)\mathrm{d}x = \lim_{h\to 0^+}\int_{a}^{b-h} f(x)\mathrm{d}x$$

在区间$[a, b)$内连续函数 f 的积分的定义.

9.48



(9.47)中定义的图示. 阴影面积为$\int_a^{b-h} f(x)\mathrm{d}x$.

9.49 $|f(x)| \leqslant g(x)$ 对于所有 $x \geqslant a \Rightarrow$
$$\left|\int_a^{\infty} f(x)\mathrm{d}x\right| \leqslant \int_a^{\infty} g(x)\mathrm{d}x$$

积分的比较测试. f 和 g 在 $x \geqslant a$ 连续.

9.50 $$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^b f(x, t)\mathrm{d}t = \int_a^b f'_x(x, t)\mathrm{d}t$$

“在积分号内微分.”a 和 b 独立于 x.

9.51 $\frac{\mathrm{d}}{\mathrm{d}x}\int_c^{\infty} f(x, t)\mathrm{d}t = \int_c^{\infty} f_x'(x, t)\mathrm{d}t$

对于(a, b)中的 x 成立，如果$f(x, t)$和 $f_x'(x, t)$对于所有 $t \geqslant c$ 和所有在(a, b)中的 x 连续，且$\int_c^{\infty} f(x, t)\mathrm{d}t$以及$\int_c^{\infty} f_x'(x,t)\mathrm{d}t$在$(a, b)$上单调收敛.

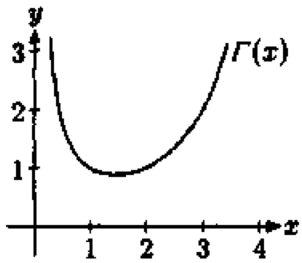
9.52 $\frac{\mathrm{d}}{\mathrm{d}x}\int_{u(x)}^{v(x)} f(x, t)\mathrm{d}t = f(x, v(x))v'(x) - f(x, u(x))u'(x) + \int_{u(x)}^{v(x)} f_x'(x, t)\mathrm{d}t$

莱布尼兹公式.

9.53 $\Gamma(x) = \int_0^{\infty} \mathrm{e}^{-t}t^{x-1}\mathrm{d}t$, $x > 0$

伽马函数.

9.54



伽马函数的图形.极小值 ≈ 0.8856 在 $x \approx 1.4616$.

9.55 $\Gamma(x + 1) = x\Gamma(x)$ 对于所有 $x > 0$

伽马函数的函数方程.

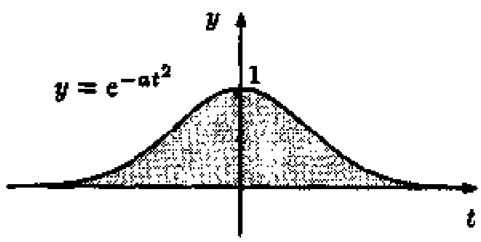
9.56 $\Gamma(n) = (n - 1)!$　当 n 是一个正整数时.

直接从函数方程得到.

9.57 $\int_{-\infty}^{+\infty} \mathrm{e}^{-at^2}\mathrm{d}t = \sqrt{\pi/a}$　$(a > 0)$

一个重要的公式.

9.58



依据(9.57)，阴影部分面积是$\sqrt{\pi/a}$.

9.59	$\int_0^\infty t^k e^{-at^2}dt = \frac{1}{2}a^{-(k+1)/2}\Gamma((k+1)/2)$	对于 $a>0$, $k>-1$ 成立.
9.60	$\Gamma(x)=\sqrt{2\pi}x^{x-\frac{1}{2}}e^{-x}e^{\theta/12x}$, $x>0$, $\theta\in(0,1)$	*Stirling* 公式.
9.61	$B(p,q)=\int_0^1 u^{p-1}(1-u)^{q-1}du$, $p,q>0$	贝塔函数.
9.62	$B(p,q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$	贝塔函数和伽马函数的关系.
9.63	$\int_a^b f(x)dx\approx\frac{b-a}{2n}\left[f(x_0)+2\sum_{i=1}^{n-1}f(x_i)+f(x_n)\right]$	梯形公式. $x_i=a+i\frac{b-a}{n}$, $i=0,\cdots,n$.
9.64	如果 f 是在$[a,b]$中的 C^2,且 $\lvert f''(x)\rvert\leqslant M$ 对于所有$x\in[a,b]$成立,则 $M(b-a)^3/12n^2$ 是(9.63)中近似误差的上限.	梯形误差估计.
9.65	$\int_a^b f(x)dx\approx\frac{b-a}{6n}D$, 其中 $D=f(x_0)+4\sum_{i=1}^{n}f(x_{2i-1})+2\sum_{i=1}^{n-1}f(x_{2i})+f(x_{2n})$	辛普森公式. 点 $x_j=a+j\frac{b-a}{2n}$, $j=0,\cdots,2n$, 将$[a,b]$分为 $2n$ 个相等的子区间.
9.66	如果 f 是在$[a,b]$内的 C^4 函数,且 $\lvert f^{(4)}(x)\rvert\leqslant M$ 对于所有 $x\in[a,b]$成立,则 $M(b-a)^5/180n^4$ 是(9.65)中近似误差的上限.	辛普森误差估计.

多重积分

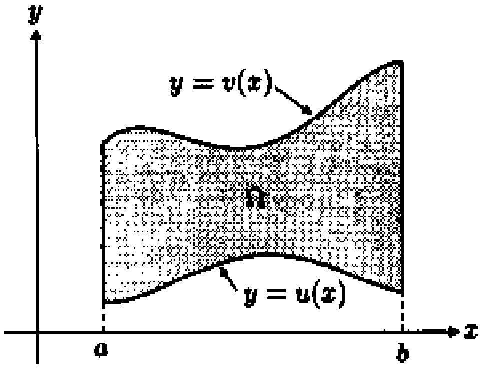
9.67	$\iint_R f(x,y)dxdy=\int_a^b\left(\int_c^d f(x,y)dy\right)dx=\int_c^d\left(\int_a^b f(x,y)dx\right)dy$	$f(x,y)$在长方形 $R=[a,b]\times[c,d]$上的双重积分的定义.(根据 *Fubini* 定理,后两个积分对于连续函数是相等的.)

9.68 $V=\int_a^b\left(\int_{u(x)}^{v(x)}f(x,y)\mathrm{d}y\right)\mathrm{d}x$

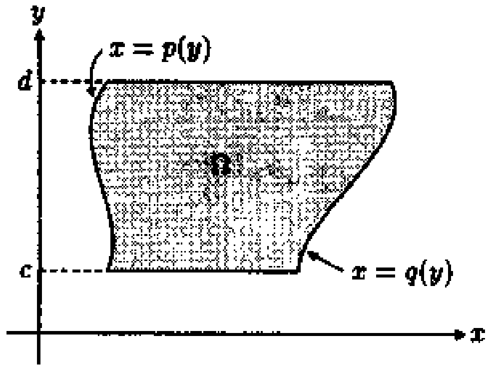
图 A 中函数 $f(x,y)$ 在区域 Ω 上的双重积分.

9.69 $V=\int_c^d\left(\int_{p(y)}^{q(y)}f(x,y)\mathrm{d}x\right)\mathrm{d}y$

图 B 中函数 $f(x,y)$ 在区域 Ω 上的双重积分.



A



B

9.70 $F''_{xy}(x,y)=f(x,y),\ (x,y)\in[a,b]\times[c,d]\Rightarrow\int_c^d\left(\int_a^b f(x,y)\mathrm{d}x\right)\mathrm{d}y=F(b,d)-F(a,d)-F(b,c)+F(a,c)$

一个有趣的结论. $f(x,y)$ 是一个连续函数.

9.71 $\iint_A f(x,y)\mathrm{d}x\mathrm{d}y$
$=\iint_{A'} f(g(u,v),h(u,v))\,|J|\,\mathrm{d}u\mathrm{d}v$

双重积分中的变量改变. $x=g(u,v)$, $y=h(u,v)$ 是从 A' 到 A 的一一对应的 C^1 转换,雅各比行列式 $J=\partial(g,h)/\partial(u,v)$ 在 A' 中不会消失, f 是连续的.

9.72 $\iint\cdots\int_\Omega f(\mathbf{x})\mathrm{d}x_1\cdots\mathrm{d}x_{n-1}\mathrm{d}x_n=$
$\int_{a_n}^{b_n}\left[\int_{a_{n-1}}^{b_{n-1}}\cdots\left(\int_{a_1}^{b_1}f(\mathbf{x})\mathrm{d}x_1\right)\cdots\mathrm{d}x_{n-1}\right]\mathrm{d}x_n$

f 在一个 n 维长方体 Ω 上的 n 重积分. $\mathbf{x}=(x_1,\cdots,x_n)$.

9.73 $\int \cdots \int_{A} f(\boldsymbol{x}) \mathrm{d}x_1 \cdots \mathrm{d}x_n = \int \cdots \int_{A'} f(g_1(\boldsymbol{u}), \cdots, g_n(\boldsymbol{u})) \mid J \mid \mathrm{d}u_1 \cdots \mathrm{d}u_n$

n 重积分的变量改变. $x_i = g_i(\boldsymbol{u})$, $i = 1, \cdots, n$, 是从 A' 到 A 的一一对应的 C^1 转换,雅各比行列式 $J = \dfrac{\partial(g_1, \cdots, g_n)}{\partial(u_1, \cdots, u_n)}$ 在 A' 中不会消失, f 是连续的.

参考文献

大多数公式都可在一般微积分教材里找到,例如 Edwards & Penney (1998). 对于(9.67)—(9.73)参照 Marsden & Hoffman (1993),其中对于多重积分里有精确的描述.(不是所有必要的假设都在多重积分的章节里得到了详细解释.)

10 差分方程

10.1 $x_t = a_t x_{t-1} + b_t,\ t = 1, 2, \cdots$

一阶线性差分方程.

10.2 $x_t = \left(\prod_{s=1}^{t} a_s\right)x_0 + \sum_{k=1}^{t}\left(\prod_{s=k+1}^{t} a_s\right)b_k$

(10.1)的解,如果定义乘积 $\prod_{s=t+1}^{t} a_s$,且其 0 项值为 1.

10.3 $x_t = a^t x_0 + \sum_{k=1}^{t} a^{t-k} b_k,\ t = 1, 2, \cdots$

当 $a_t = a$ 为常数时,(10.1)的解.

10.4
- $x_t = Aa^t + \sum_{s=0}^{\infty} a^s b_{t-s},\ |a| < 1$
- $x_t = Aa^t - \sum_{s=1}^{\infty}\left(\frac{1}{a}\right)^s b_{t+s},\ |a| > 1$

(10.1)的后向和前向解,其中 $a_t = a$,且 A 是一个任意常数.

10.5 $x_t = ax_{t-1} + b \Leftrightarrow$

$$x_t = a^t\left(x_0 - \frac{b}{1-a}\right) + \frac{b}{1-a}$$

当 $a_t = a \neq 1$, $b_t = b$ 时,方程(10.1)及其解.

10.6 (*) $x_t + a_1(t)x_{t-1} + \cdots + a_n(t)x_{t-n} = b_t$

(**) $x_t + a_1(t)x_{t-1} + \cdots + a_n(t)x_{t-n} = 0$

(*)是 n 阶一般线性非齐次差分方程,而(**)是相对应的齐次方程.

10.7 如果 $u_1(t), \cdots, u_n(t)$ 是(10.6)(**)的线性无关的解,u_t^* 是(10.6)(*)的某一特殊解,且 $C_1, \cdots, C_n$ 是任意常数,则(**)的一般解是

$$x_t = C_1u_1(t) + \cdots + C_nu_n(t)$$

而(*)的一般解是

$$x_t = C_1u_1(t) + \cdots + C_nu_n(t) + u_t^*$$

(10.6)解的一般结构.(关于线性无关,参见(11.15))

10.8 当 $b \neq 0$, $x_t + ax_{t-1} + bx_{t-2} = 0$ 有解:

- 对于 $\frac{1}{4}a^2 - b > 0$: $x_t = C_1 m_1^t + C_2 m_2^t$,

其中 $m_{1,2} = -\frac{1}{2}a \pm \sqrt{\frac{1}{4}a^2 - b}$.

- 对于 $\frac{1}{4}a^2 - b = 0$:

  $x_t = (C_1 + C_2 t)(-a/2)^t$.

- 对于 $\frac{1}{4}a^2 - b < 0$:

  $x_t = Ar^t \cos(\theta t + \omega)$,

  其中 $r = \sqrt{b}$ 而 $\cos\theta = -\frac{a}{2\sqrt{b}}$, $\theta \in [0, \pi]$.

具有常数系数 a 和 b 的二阶线性齐次差分方程的解. C_1, C_2 和 ω 是任意常数.

10.9 若求

$(*)$ $x_t + ax_{t-1} + bx_{t-2} = c_t$, $b \neq 0$

的一个特殊解,应用下列测试函数,及待定系数法来决定常数:

- 如果 $c_t = c$, 测试 $u_t^* = A$.
- 如果 $c_t = ct + d$, 测试 $u_t^* = At + B$.
- 如果 $c_t = t^n$, 测试 $u_t^* = A_0 + A_1 t + \cdots + A_n t^n$.
- 如果 $c_t = c^t$, 测试 $u_t^* = Ac^t$.
- 如果 $c_t = \alpha \sin ct + \beta \cos ct$, 测试 $u_t^* = A \sin ct + B \cos ct$.

如果函数 c_t 自身是齐次方程的一个解,将测试函数乘以 t. 如果这一新测试函数也满足齐次方程,再次用 t 乘以测试函数.(对于一般程序,参见 Hilde-brand(1968), 1.8 节)

10.10 $(*)$ $x_t + a_1 x_{t-1} + \cdots + a_n x_{t-n} = b_t$

$(**)$ $x_t + a_1 x_{t-1} + \cdots + a_n x_{t-n} = 0$

具有常数系数的线性差分方程.

10.11 $m^n + a_1 m^{n-1} + \cdots + a_{n-1} m + a_n = 0$

(10.10)的特征方程. 它的根称为特征值.

10.12 假设特征方程(10.11)有 n 个不同的根 λ_1, …, λ_n,定义

$$\theta_r = \frac{\lambda_r}{\prod\limits_{\substack{1 \leqslant s \leqslant n \\ s \neq r}} (\lambda_r - \lambda_s)}, \quad r = 1, 2, \cdots, n$$

(10.10)(＊)的一个特殊解为

$$u_t^* = \sum_{r=1}^{n} \theta_r \sum_{i=0}^{\infty} \lambda_r^i b_{t-i}$$

(10.10)(＊)的后向解,当 $|\lambda_r| < 1$ 对 $r = 1, \cdots, n$ 成立时成立.

10.13 要得到(10.10)(＊＊)的 n 个线性无关的解:先找到特征方程(10.11)的所有解.然后:

- 任一实根 m_i 和重数 1 给出解 m_i^t.
- 任一实根 m_j 和重数 $p > 1$ 给出解 m_j^t, tm_j^t, …, $t^{p-1}m_j^t$.
- 任意一对复根 $m_k = \alpha + \mathrm{i}\beta$, $\overline{m_k} = \alpha - \mathrm{i}\beta$ 和重数 1 给出解 $r^t \cos \theta t$, $r^t \sin \theta t$,其中 $r = \sqrt{\alpha^2 + \beta^2}$,且 $\theta \in [0, \pi]$ 满足 $\cos \theta = \frac{\alpha}{r}$, $\sin \theta = \frac{\beta}{r}$.
- 任意一对复根 $m_e = \lambda + \mathrm{i}\mu$, $\overline{m_e} = \lambda - i\mu$ 和因子 $q > 1$ 给出解 u, v, tu, tv, …, $t^{q-1}u$, $t^{q-1}v$,其中 $u = s^t \cos \varphi t$, $v = s^t \sin \varphi t$, $s = \sqrt{\lambda^2 + \mu^2}$,且 $\varphi \in [0, \pi]$ 满足 $\cos \varphi = \frac{\lambda}{s}$ 和 $\sin \varphi = \frac{\mu}{s}$.

求(10.10)(＊＊)的 n 个线性无关解的一般方法.

10.14 如果齐次方程(10.10)(＊＊)的任一解当 $t \to \infty$ 时都趋向于 0,则方程(10.10)称为(整体渐近)稳定的.

常系数的线性方程的稳定性的定义.

10.15 方程(10.10)称为稳定的,当且仅当特征方程(10.11)所有根的模数小于 1.

(10.10)(＊)(或(＊＊))的稳定性的标准.

10.16

$$\left|\begin{array}{c:c} 1 & a_n \\ \hdashline a_n & 1 \end{array}\right| > 0,\quad \left|\begin{array}{cc:cc} 1 & 0 & a_n & a_{n-1} \\ a_1 & 1 & 0 & a_n \\ \hdashline a_n & 0 & 1 & a_1 \\ a_{n-1} & a_n & 0 & 1 \end{array}\right| > 0, \cdots$$

$$\left|\begin{array}{cccc:cccc} 1 & 0 & \cdots & 0 & a_n & a_{n-1} & \cdots & a_1 \\ a_1 & 1 & \cdots & 0 & 0 & a_n & \cdots & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n-1} & a_{n-2} & \cdots & 1 & 0 & 0 & \cdots & a_n \\ \hdashline a_n & 0 & \cdots & 0 & 1 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_n & \cdots & 0 & 0 & 1 & \cdots & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & a_n & 0 & 0 & \cdots & 1 \end{array}\right| > 0$$

(10.11)的所有根具有小于 1 的模数的一个充分必要条件.(*Schur* 定理)

10.17 $x_t + a_1 x_{t-1} = b_t$ 是稳定的$\Leftrightarrow |a_1| < 1$

(10.15)和(10.16)的特例.

10.18 $x_t + a_1 x_{t-1} + a_2 x_{t-2} = b_t$ 是稳定的

$$\Leftrightarrow \begin{cases} 1 - a_2 > 0 \\ 1 - a_1 + a_2 > 0 \\ 1 + a_1 + a_2 > 0 \end{cases}$$

(10.15)和(10.16)的特例.

10.19 $x_t + a_1 x_{t-1} + a_2 x_{t-2} + a_3 x_{t-3} = b_t$ 是稳定的

$$\Leftrightarrow \begin{cases} 3 - a_2 > 0 \\ 1 - a_2 + a_1 a_3 - a_3^2 > 0 \\ 1 + a_2 - |a_1 + a_3| > 0 \end{cases}$$

(10.15)和(10.16)的特例.

10.20 $x_t + a_1 x_{t-1} + a_2 x_{t-2} + a_3 x_{t-3} + a_4 x_{t-4} = b_t$ 是稳定的$\Leftrightarrow$

$$\begin{cases} 1 - a_4 > 0 \\ 3 + 3a_4 - a_2 > 0 \\ 1 + a_2 + a_4 - |a_1 + a_3| > 0 \\ (1 - a_4)^2(1 + a_4 - a_2) > (a_1 - a_3)(a_1 a_4 - a_3) \end{cases}$$

(10.15)和(10.16)的特例.

10.21	$x_1(t) = a_{11}(t)x_1(t-1) + \cdots + a_{1n}(t)x_n(t-1) + b_1(t)$ ………………………… $x_n(t) = a_{n1}(t)x_1(t-1) + \cdots + a_{nn}(t)x_n(t-1) + b_n(t)$	线性差分方程组.
10.22	$\boldsymbol{x}(t) = \mathbf{A}(t-1)\boldsymbol{x}(t) + \boldsymbol{b}(t),\ t = 1, 2, \cdots$	(10.21)的矩阵形式. $\boldsymbol{x}(t)$和$\boldsymbol{b}(t)$是$n \times 1$, $\mathbf{A}(t) = (a_{ij}(t))$是$n \times n$.
10.23	$\boldsymbol{x}(t) = \mathbf{A}^t\boldsymbol{x}(0) + (\mathbf{A}^{t-1} + \mathbf{A}^{t-2} + \cdots + \mathbf{A} + \mathbf{I})\boldsymbol{b}$	当$\mathbf{A}(t) = \mathbf{A}$, $\boldsymbol{b}(t) = \boldsymbol{b}$时(10.22)的解.
10.24	$\boldsymbol{x}(t) = \mathbf{A}\boldsymbol{x}(t-1) \Leftrightarrow \boldsymbol{x}(t) = \mathbf{A}^t\boldsymbol{x}(0)$	(10.23)的一个特例,其中$\boldsymbol{b} = 0$,且$\mathbf{A}^0 = \mathbf{I}$.
10.25	如果$\mathbf{A}$是一个$n \times n$对角线矩阵,其特征值为$\lambda_1, \lambda_2, \cdots, \lambda_n$,则(10.24)的解可写成 $\boldsymbol{x}(t) = \mathbf{P}\begin{pmatrix} \lambda_1^t & 0 & \cdots & 0 \\ 0 & \lambda_2^t & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^t \end{pmatrix}\mathbf{P}^{-1}\boldsymbol{x}(0)$ 其中$\mathbf{P}$是$\mathbf{A}$的相应的线性无关特征向量组成的矩阵.	一个重要的结论.(参见(21.16))
10.26	差分方程(10.22)当$\mathbf{A}(t) = \mathbf{A}$时称为稳定的,如果对于任意选择的向量$\boldsymbol{x}(0)$, $\mathbf{A}^t\boldsymbol{x}(0)$都趋向于向量0.	线性方程组的稳定性的定义.
10.27	差分方程(10.22)当$\mathbf{A}(t) = \mathbf{A}$时称为稳定的,当且仅当$\mathbf{A}$的所有特征值的模数都小于1.	线性方程组稳定性的特点.

10.28 如果 $\mathbf{A}=(a_{ij})_{n\times n}$ 的所有特征值的模数都小于 1，则每一解 $\mathbf{x}(t)$ $$\boldsymbol{x}(t)=\mathbf{A}\boldsymbol{x}(t-1)+\boldsymbol{b},\ t=1,2,\cdots$$ 收敛于向量 $(\mathbf{I}-\mathbf{A})^{-1}\boldsymbol{b}$.	一个重要方程的解.（参见(22.23)）

参 考 文 献

大多数公式和结论都可在 Goldberg(1961)，Gandolfo(1996)和 Hildebrand (1968)中找到. 对于(10.19)和(10.20)，参见 Farebrother(1973).

11 微分方程

一阶方程

11.1 $\dot{x}(t)=f(t)\Leftrightarrow x(t)=x(t_0)+\int_{t_0}^{t}f(\tau)\mathrm{d}\tau$

一个简单的微分方程及其解. $f(t)$是一给定函数,$x(t)$是未知函数.

11.2 $\frac{\mathrm{d}x}{\mathrm{d}t}=f(t)g(x)\Leftrightarrow\int\frac{\mathrm{d}x}{g(x)}=\int f(t)\mathrm{d}t$

计算积分.对由此产生的隐含方程求解 $x=x(t)$.

可分离的微分方程.如果 $g(a)=0$,则 $x(t)\equiv a$ 是一解.

11.3 $\dot{x}=g(x/t)$ 和 $z=x/t\Rightarrow t\frac{\mathrm{d}z}{\mathrm{d}t}=g(z)-z$

射影(或)齐次微分方程.通过变量变换 $z=x/t$ 可得一 z 的可分离方程.

11.4 方程 $\dot{x}=B(x-a)(x-b)$ 有解

$$x\equiv a,\ x\equiv b,\ x=a+\frac{b-a}{1-Ce^{B(b-a)t}}$$

$a\neq b$, $a=0$ 时给出 logistic 方程. C 是一常数.

11.5 • $\dot{x}+ax=b\quad\Leftrightarrow x=Ce^{-at}+\frac{b}{a}$

• $\dot{x}+ax=b(t)\Leftrightarrow x=e^{-at}(C+\int b(t)e^{at}\mathrm{d}t)$

常系数 $a\neq 0$ 一阶线性微分方程. C 是一常数.

11.6 $\dot{x}+a(t)x=b(t)\Leftrightarrow$

$$x=e^{-\int a(t)\mathrm{d}t}\left(C+\int e^{\int a(t)\mathrm{d}t}b(t)\mathrm{d}t\right)$$

一般线性一阶微分方程. $a(t)$和 $b(t)$是给定的. C 是一常数.

11.7 $\dot{x}+a(t)x=b(t)\Leftrightarrow$

$$x(t)=x_0e^{-\int_{t_0}^{t}a(\xi)d\xi}+\int_{t_0}^{t}b(\tau)e^{-\int_{\tau}^{t}a(\xi)d\xi}d\tau$$

对于给定初始值 $x(t_0)=x_0$(11.6)的解.

11.8 $\dot{x}=Q(t)x+R(t)x^n$ 有解

$x(t)=$

$$e^{\frac{P(t)}{1-n}}\left[C+(1-n)\int R(t)e^{-P(t)}dt\right]^{\frac{1}{1-n}}$$

其中 $P(t)=(1-n)\int Q(t)dt$

Bernoulli 方程及其解 ($n\neq 1$). C 是一常数.

11.9 $\dot{x}=P(t)+Q(t)x+R(t)x^2$

Riccati 方程. 一般不能通过分析法解出. 变量变换法 $x=u+1/z$ 仅当知道一特殊解 $u=u(t)$ 时有效.

微分方程

11.10 $P(t,x)+Q(t,x)\dot{x}=0$
称为恰当的,如果存在一 C^1 函数 $\varphi(t,x)$,使 $P(t,x)=\varphi_1'(t,x)$ 及 $Q(t,x)=\varphi_2'(t,x)$. 对于某一常数 C 的解则为 $\varphi(t,x)=C$.

一个恰当方程及其解.

11.11 如果 $P(t,x)$和 $Q(t,x)$是 C^1 函数,则
$P(t,x)+Q(t,x)\dot{x}=0$
在一开长方形 R 内是恰当的,当且仅当 $P_2'(t,x)=Q_1'(t,x)$ 在 R 中.

任一微分方程是恰当方程的充分必要条件.

高阶方程

11.12 $\frac{d^nx}{dt^n}+a_1(t)\frac{d^{n-1}x}{dt^{n-1}}+\cdots+a_{n-1}(t)\frac{dx}{dt}+a_n(t)x=f(t)$

一般线性 n 阶微分方程.

11.13	$\frac{d^n x}{dt^n} + a_1(t)\frac{d^{n-1}x}{dt^{n-1}} + \cdots + a_{n-1}(t)\frac{dx}{dt} + a_n(t)x = 0$	与(11.12)相应的齐次方程.
11.14	函数 $u_1(t),\cdots,u_m(t)$是线性无关的,如果 $C_1u_1(t)+\cdots+C_mu_m(t)=0$ 仅当常数 $C_1,\cdots,C_m$ 都为 0 时对于所有 t 成立. 如果函数不是线性无关的则是线性相关的.	线性相关及线性无关的定义.
11.15	如果 $u_1(t),\cdots,u_n(t)$是齐次方程(11.13)的线性无关的解,而 $u^*(t)$是非齐次方程(11.12)的某一特殊解,则(11.13)的一般解是 $x(t)=C_1u_1(t)+\cdots+C_nu_n(t)$ 而(11.12)的一般解是 $x(t)=C_1u_1(t)+\cdots+C_nu_n(t)+u^*(t)$ 其中 $C_1,\cdots,C_n$ 是任意常数.	(11.13)和(11.12)解的结构.(注意一般来说不可能找到(11.13)的 n 个解 $u_1(t),\cdots,u_n(t)$的分析性表达式)
11.16	如果 $u_1,\cdots,u_n$ 是(11.13)的 n 个线性无关的解,求(11.12)的一个特殊解的方法是:解对于 $\dot{C}_1(t),\cdots,\dot{C}_n(t)$的方程组 $\dot{C}_1(t)u_1+\cdots+\dot{C}_n(t)u_n=0$ $\dot{C}_1(t)\dot{u}_1+\cdots+\dot{C}_n(t)\dot{u}_n=0$ ……………………………………… $\dot{C}_1(t)u_1^{(n-2)}+\cdots+\dot{C}_n(t)u_n^{(n-2)}=0$ $\dot{C}_1(t)u_1^{(n-1)}+\cdots+\dot{C}_n(t)u_n^{(n-1)}=b(t)$ 积分以求得 $C_1(t),\cdots,C_n(t)$.(11.12)的一个特殊解为: $u^*(t)=C_1(t)u_1+\cdots+C_n(t)u_n$.	参数变化法,如果我们知道(11.13)的一般解,总可以求出(11.12)的一个特殊解.在此 $u_j^{(i)}=d^iu_j/dt^i$ 是 u_j 的第 i 阶导数.
11.17	$\ddot{x}+a\dot{x}+bx=0$ 有解 • 对于 $\frac{1}{4}a^2-b>0$: $x=C_1e^{r_1t}+C_2e^{r_2t}$ 其中 $r_{1,2}=-\frac{1}{2}a\pm\sqrt{\frac{1}{4}a^2-b}$. • 对于 $\frac{1}{4}a^2-b=0$: $x=(C_1+C_2t)e^{-at/2}$ • 对于 $\frac{1}{4}a^2-b<0$: $x=Ae^{\alpha t}\cos(\beta t+\omega)$ 其中 $\alpha=-\frac{1}{2}a$, $\beta=\sqrt{b-\frac{1}{4}a^2}$.	有常系数 a 和 b 的二阶齐次微分方程的解, C_1, C_2, A 和 ω 是常数.

11.18	$\ddot{x}+a\dot{x}+bx=f(t)$, $b\neq 0$,有一特殊解 $u^*=u^*(t)$: • $f(t)=A$: $u^*=A/b$ • $f(t)=At+B$: $u^*=\frac{A}{b}t+\frac{bB-aA}{b^2}$ • $f(t)=At^2+Bt+C$: $u^*=\frac{A}{b}t^2+\frac{(bB-2aA)}{b^2}t+\frac{Cb^2-(2A+aB)b+2a^2A}{b^3}$ • $f(t)=pe^{qt}$: $u^*=pe^{qt}/(q^2+aq+b)$ (如果 $q^2+aq+b\neq 0$).	$\ddot{x}+a\dot{x}+bx=f(t)$ 的特殊解. 如果 $f(t)=pe^{qt}$, $q^2+aq+b=0$, 且 $2q+a\neq 0$, 则 $u^*=pte^{qt}/(2q+a)$ 是一个解. 如果 $f(t)=pe^{qt}$, $q^2+aq+b=0$, 且 $2q+a=0$, 则 $u^*=\frac{1}{2}pt^2e^{qt}$ 是一个解.
11.19	$t^2\ddot{x}+at\dot{x}+bx=0$, $t>0$,有解 • 对于 $(a-1)^2>4b$: $x=C_1t^{r_1}+C_2t^{r_2}$, 其中 $r_{1,2}=-\frac{1}{2}\left[(a-1)\pm\sqrt{(a-1)^2-4b}\right]$ • 对于 $(a-1)^2=4b$: $x=(C_1+C_2\ln t)t^{(1-a)/2}$ • 对于 $(a-1)^2<4b$: $x=At^{\lambda}\cos(\mu\ln t+\omega)$ 其中 $\lambda=\frac{1}{2}(1-a)$, $\mu=\frac{1}{2}\sqrt{4b-(a-1)^2}$.	2 阶欧拉方程的解. C_1, C_2, A 和 ω 是任意常数.
11.20	$\frac{d^nx}{dt^n}+a_1\frac{d^{n-1}x}{dt^{n-1}}+\cdots+a_{n-1}\frac{dx}{dt}+a_nx=f(t)$	常系数 n 阶一般线性微分方程.
11.21	$\frac{d^nx}{dt^n}+a_1\frac{d^{n-1}x}{dt^{n-1}}+\cdots+a_{n-1}\frac{dx}{dt}+a_nx=0$	与(11.20)相应的齐次方程.
11.22	$r^n+a_1r^{n-1}+\cdots+a_{n-1}r+a_n=0$	与(11.20)和(11.21)相应的特征方程.

11.23 $x = x(t) = C_1e^{r_1t} + C_2e^{r_2t} + \cdots + C_ne^{r_nt}$

(11.22)的根 $r_1, \cdots, r_n$ 是实根且互不相同时(11.21)的解.

11.24 要求(11.21)的 n 个线性无关的解:找到(11.22)的所有根.

- 一个实根 r_i 及重数 1 给出解 e^{r_it}.
- 一个实根 r_j 及重数 $p > 1$ 给出解 $e^{r_jt}, te^{r_jt}, \cdots, t^{p-1}e^{r_jt}$.
- 一对复根 $r_k = \alpha + i\beta$, $\bar{r}_k = \alpha - i\beta$ 及重数 1 给出解 $e^{\alpha t}\cos\beta t$ 和 $e^{\alpha t}\sin\beta t$.
- 一对复根 $r_e = \lambda + i\mu$, $\bar{r}_e = \lambda - i\mu$ 及重数 $q > 1$,给出解:$u, v, tu, tv, \cdots, t^{q-1}u, t^{q-1}v$,其中 $u = e^{\lambda t}\cos\mu t$ 及 $v = e^{\lambda t}\sin\mu t$.

求(11.21)的 n 个线性无关解的一般方法.

11.25 方程(11.21)(或(11.20))是稳定的(整体渐近稳定的),如果当 $t\to\infty$ 时(11.21)的任一解趋向于 0.

常系数方程的稳定性的定义.

11.26 方程(11.21)是稳定的$\Leftrightarrow$特征方程(11.22)的所有根有负的实部.

(11.21)的稳定条件.

11.27 (11.21)是稳定的$\Rightarrow$对于所有 $i = 1, \cdots, n$, $a_i > 0$

(11.21)的稳定性的必要条件.

11.28 $$\mathbf{A} = \begin{pmatrix} a_1 & a_3 & a_5 & \cdots & 0 & 0 \\ a_0 & a_2 & a_4 & \cdots & 0 & 0 \\ 0 & a_1 & a_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & a_{n-2} & a_n \end{pmatrix}$$

由(11.21)的系数组成的矩阵($a_0 = 1$),此矩阵第 k 列为 $\cdots a_{k+1}a_k a_{k-1}\cdots$ 其中元素 a_k 在主对角线上,元素 a_{k+j} 当 $k+j$ 为负或大于 n 时取 0 值)

11.29	(a_1), $\begin{pmatrix} a_1 & 0 \\ 1 & a_2 \end{pmatrix}$, $\begin{pmatrix} a_1 & a_3 & 0 \\ 1 & a_2 & 0 \\ 0 & a_1 & a_3 \end{pmatrix}$	当 $n = 1, 2, 3$, $a_0 = 1$ 时(11.28)中的矩阵 A.
11.30	(11.21)是稳定的$\Leftrightarrow\begin{cases}\text{(11.28)中 A 的所有} \\ \text{主子式(当 } a_0 = 1) \\ \text{为正}\end{cases}$	Routh-Hurwitz 稳定条件.
11.31	• $\dot{x} + a_1 x = f(t)$ 是稳定的$\Leftrightarrow a_1 > 0$ • $\ddot{x} + a_1\dot{x} + a_2 x = f(t)$ 是稳定的$\Leftrightarrow\begin{cases} a_1 > 0 \\ a_2 > 0\end{cases}$ • $\dddot{x} + a_1\ddot{x} + a_2\dot{x} + a_3 x = f(t)$ 是稳定的 $\Leftrightarrow a_1 > 0,\ a_3 > 0$ 且 $a_1 a_2 > a_3$	(11.30)的特例.(容易看出这相当于要求(11.29)中的矩阵的所有主子式为正)

微分方程组

11.32	$\left.\begin{array}{l} \dfrac{dx_1}{dt} = f_1(t, x_1, \cdots, x_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \dfrac{dx_n}{dt} = f_n(t, x_1, \cdots, x_n) \end{array}\right\} \Leftrightarrow \dot{\boldsymbol{x}} = \boldsymbol{F}(t, \boldsymbol{x})$	正规(非自控)微分方程组. 此处 $\boldsymbol{x} = (x_1, \cdots, x_n)$, $\dot{\boldsymbol{x}} = (\dot{x}_1, \cdots, \dot{x}_n)$, 而 $\boldsymbol{F} = (f_1, \cdots, f_n)$.
11.33	$\dot{x}_1 = a_{11}(t)x_1 + \cdots + a_{1n}(t)x_n + b_1(t)$ $\cdots\cdots\cdots\cdots\cdots\cdots$ $\dot{x}_n = a_{n1}(t)x_1 + \cdots + a_{nn}(t)x_n + b_n(t)$	线性微分方程组.
11.34	$\dot{\boldsymbol{x}} = \mathbf{A}(t)\mathbf{x} + \boldsymbol{b}(t),\ \boldsymbol{x}(t_0) = \boldsymbol{x}^0$	有初始条件的(11.33)的矩阵形式. $\boldsymbol{x}$, $\dot{\boldsymbol{x}}$ 和 $\boldsymbol{b}(t)$ 是列向量且 $\mathbf{A}(t) = (a_{ij}(t))_{n\times n}$.
11.35	$\dot{\boldsymbol{x}} = \mathbf{A}\boldsymbol{x},\ \boldsymbol{x}(t_0) = \boldsymbol{x}^0 \Leftrightarrow \boldsymbol{x} = e^{\mathbf{A}(t-t_0)}\boldsymbol{x}^0$	当 $\mathbf{A}(t) = \mathbf{A}$, $\boldsymbol{b}(t) = 0$ 时(11.34)的解.(矩阵指数参见(19.30))

11.36	如果 $p_j(t)=(p_{1j}(t), \cdots, p_{nj}(t))'$, $j=1, \cdots, n$ 为齐次方程 $\dot{x}=\mathbf{A}(t)x$ 的 n 个线性无关的解，且 $p_j(t_0)=e_j$, $j=1, \cdots, n$，其中 e_j 是 $\mathbf{R}^n$ 中第 j 个标准单位向量. 则方程的分解就是矩阵 $$\mathbf{P}(t, t_0)=\begin{pmatrix} p_{11}(t) & \cdots & p_{1n}(t) \\ \vdots & \ddots & \vdots \\ p_{n1}(t) & \cdots & p_{nn}(t) \end{pmatrix}$$	齐次线性微分方程的分解的定义. 注意 $\mathbf{P}(t_0, t_0)=\mathbf{I}_n$.
11.37	$x=\mathbf{P}(t, t_0)x^0+\int_{t_0}^{t}\mathbf{P}(t, s)b(s)\mathrm{d}s$	(11.34)的解.
11.38	如果 $\mathbf{P}(t, s)$ 是 $\dot{x}=\mathbf{A}(t)x$ 的分解，则 $\mathbf{P}(s, t)'$($\mathbf{P}(s, t)$的转置)是 $\dot{z}=-(\mathbf{A}(t))'z$ 的分解.	一个有用的事实.
11.39	考虑 n 阶微分方程 $$(*)\ \frac{\mathrm{d}^n x}{\mathrm{d}t^n}=F\left(t, x, \frac{\mathrm{d}x}{\mathrm{d}t}, \cdots, \frac{\mathrm{d}^{n-1}x}{\mathrm{d}t^{n-1}}\right)$$ 通过设置新变量, $$y_1=x,\ y_2=\frac{\mathrm{d}x}{\mathrm{d}t}, \cdots, y_n=\frac{\mathrm{d}^{n-1}x}{\mathrm{d}t^{n-1}}$$ (*)可转换成一般方程组 $$\begin{aligned}\dot{y}_1&=y_2\\ \dot{y}_2&=y_3\\ &\cdots\cdots\\ \dot{y}_{n-1}&=y_n\\ \dot{y}_n&=F(t, y_1, y_2, \cdots, y_n)\end{aligned}$$	任一 n 阶微分方程组可通过设置新变量转换成一般方程组(许多高阶微分方程组可用类似方法设置新变量转换成一般方程组)
11.40	考虑初始值问题 $$(*)\ \dot{x}=F(t, x),\ x(t_0)=x^0$$ 其中 $F=(f_1, \cdots, f_n)$ 和其对于 $x_1, \cdots, x_n$ 的一阶偏导数在集合 $$\Gamma=\{(t, x): \vert t-t_0\vert\leqslant a,\ \Vert x-x^0\Vert\leqslant b\}$$ 内连续定义 $$M=\max_{(t, x)\in\Gamma}\Vert F(t, x)\Vert,\ r=\min(a, b/M)$$ 则(*)在开区间 (t_0-r, t_0+r) 内有一唯一解 $x(t)$，且 $\Vert x(t)-x^0\Vert\leqslant b$ 在这区间内.	(局部)存在及唯一性定理.

11.41 考虑初始值问题

(1) $\dot{\boldsymbol{x}} = \boldsymbol{F}(t, \boldsymbol{x}), \boldsymbol{x}(t_0) = \boldsymbol{x}^0$

其中 $\boldsymbol{F} = (f_1, \cdots, f_n)$ 和其对 $x_1, \cdots, x_n$ 的一阶偏导数对于所有$(t, \boldsymbol{x})$连续.进一步假设存在连续函数 $a(t)$和 $b(t)$,使

(2) $\| \boldsymbol{F}(t, \boldsymbol{x}) \| \leqslant a(t) \| \boldsymbol{x} \| + b(t)$ 对于所有$(t, \boldsymbol{x})$

或

(3) $\boldsymbol{x}\boldsymbol{F}(t, \boldsymbol{x}) \leqslant a(t) \| \boldsymbol{x} \|^2 + b(t)$ 对于所有$(t, \boldsymbol{x})$,则给定任意点$(t_0, \boldsymbol{x}^0)$,存在一个(1)的唯一解 $\boldsymbol{x}(t)$定义在$(-\infty, \infty)$

不等式(2)得以满足,特别如果对于所有$(t, \boldsymbol{x})$

(4) $\| \boldsymbol{F}_{\boldsymbol{x}}'(t, \boldsymbol{x}) \| \leqslant c(t)$ 对一连续 $c(t)$

整体存在及唯一性定理.(4)中可用$\boldsymbol{F}_{\boldsymbol{x}}'(t, \boldsymbol{x})$的任意矩阵模数.(对于矩阵的模数参见(19.26))

自控系统

11.42 $\dot{x}_1 = f_1(x_1, \cdots, x_n)$

$\cdots\cdots\cdots\cdots\cdots\cdots$

$\dot{x}_n = f_n(x_1, \cdots, x_n)$

一阶微分方程自控系统的定义.

11.43 如果 $f_i(\boldsymbol{a}) = 0, i = 1, \cdots, n, \boldsymbol{a} = (a_1, \cdots, a_n)$ 是系统(11.42)的均衡点.

(11.42)的均衡点的定义.

11.44 如果 $\boldsymbol{x}(t) = (x_1(t), \cdots, x_n(t))$ 是系统(11.42)在一区间 I 上的一个解,则在 $\mathbf{R}^n$ 内的点$\boldsymbol{x}(t)$的集合描出的 $\mathbf{R}^n$ 内的曲线称为方程组的一条轨线(或一轨道).

轨线(或轨道)的定义,也称为积分曲线.

11.45 如果所有出发点接近 $\mathbf{a}$ 的解都停留在 $\mathbf{a}$ 点附近，(11.42)的均衡点 $\mathbf{a}$ 是(局部)稳定的：对于每一 $\varepsilon > 0$ 存在一个 $\delta > 0$，如果 $\| \mathbf{x} - \mathbf{a} \| < \delta$，则存在(11.42)的一个解 $\varphi(t)$ 定义在 $t \geqslant 0$，且 $\varphi(0) = \mathbf{x}$，并满足

$$\| \varphi(t) - \mathbf{a} \| < \varepsilon \quad \text{对于 } t > 0$$

如果 $\mathbf{a}$ 是稳定的且存在 $\delta' > 0$，使

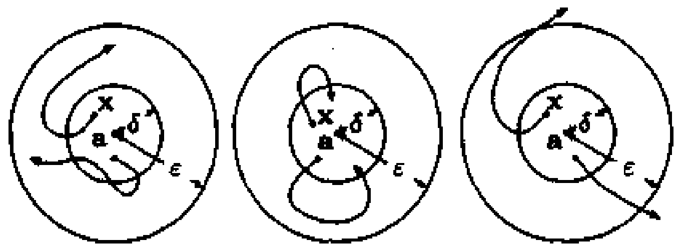
$$\| \mathbf{x} - \mathbf{a} \| < \delta' \Rightarrow \lim_{t \to \infty} \| \varphi(t) - \mathbf{a} \| = 0$$

则 $\mathbf{a}$ 是(局部)渐近稳定的.

如果 $\mathbf{a}$ 是不稳定的，则称为非稳定的.

(局部)稳定性和非稳定性的定义.

11.46



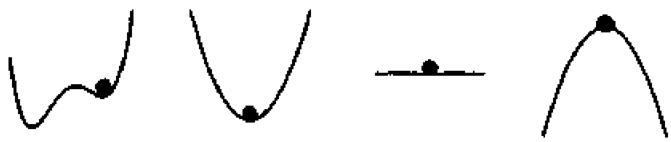
局部稳定　　渐近稳定　　非稳定

稳定性概念的图示. 带箭头的曲线是可能的轨线.

11.47 如果(11.42)的每一解，无论其初始点是什么，都收敛于唯一的均衡点 $\mathbf{a}$，则 $\mathbf{a}$ 是整体渐近稳定的.

整体渐近稳定性.

11.48



局部稳定　　整体稳定　　非稳定　　非稳定

稳定性概念的粗略图示.

11.49 假设 $\mathbf{x}(t)$ 是方程组(11.42)的一个解，其中 $\mathbf{F} = (f_1, \cdots, f_n)$ 是一 C^1 函数，且 $\mathbf{x}(t_0 + T) = \mathbf{x}(t_0)$ 对某一 t_0 和某一 $T > 0$ 成立. 则 $\mathbf{x}(t + T) = \mathbf{x}(t)$ 对所有 t 成立.

如果(11.42)的解经过一段时间 T 回到其初始点，则它必然是周期性的，其周期为 T.

11.50 假设$(x(t), y(t))$是方程组

$$\dot{x} = f(x, y), \ \dot{y} = g(x, y)$$

的解，且驻留在方程组无均衡点的紧区域内．它的轨线必然是一封闭螺旋线，也是方程组的周期性解的轨线．

Poincaré-Bendixson 定理．

11.51 设 $\mathbf{a}$ 为(11.42)的一均衡点，定义

$$\mathbf{A} = \begin{pmatrix} \frac{\partial f_1(\mathbf{a})}{\partial x_1} & \cdots & \frac{\partial f_1(\mathbf{a})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(\mathbf{a})}{\partial x_1} & \cdots & \frac{\partial f_n(\mathbf{a})}{\partial x_n} \end{pmatrix}$$

如果 $\mathbf{A}$ 的所有特征值都有负的实部，则 $\mathbf{a}$ 是(局部)渐近稳定的．

如果至少一个特征值的实部是正的，则 $\mathbf{a}$ 不稳定．

Liapunov 定理．稳定点 $\mathbf{a}$ 称为收点当所有 $\mathbf{A}$ 的特征值均有负实部(称为发点当 $\mathbf{A}$ 的所有特征值均有正实部)．

11.52 $n \times n$ 实矩阵 $\mathbf{A} = (a_{ij})$ 的所有特征值均有负实部的充分必要条件是以下不等式成立

- 对于 $n = 2$：$\mathrm{tr}(\mathbf{A}) < 0$ 且 $|\mathbf{A}| > 0$
- 对于 $n = 3$：$\mathrm{tr}(\mathbf{A}) < 0$，$|\mathbf{A}| < 0$，且

$$\begin{vmatrix} a_{22} + a_{33} & -a_{12} & -a_{13} \\ -a_{21} & a_{11} + a_{33} & -a_{23} \\ -a_{31} & -a_{32} & a_{11} + a_{22} \end{vmatrix} < 0$$

2,3 阶稳定矩阵的特征．($n \times n$ 矩阵经常称为稳定的，如果所有特征值都有负实部．)

11.53 设(a, b)为方程组

$$\dot{x} = f(x, y), \ \dot{y} = g(x, y)$$

的一个稳定点，定义

$$\mathbf{A} = \begin{pmatrix} \frac{\partial f(a, b)}{\partial x} & \frac{\partial f(a, b)}{\partial y} \\ \frac{\partial g(a, b)}{\partial x} & \frac{\partial g(a, b)}{\partial y} \end{pmatrix}$$

则如果 $\mathrm{tr}(\mathbf{A}) < 0$ 且 $\det(\mathbf{A}) > 0$，(a, b) 是局部渐近稳定的．

(11.51)的一个特例．稳定性通过迹和 $\mathbf{A}$ 的行列式的符号来表示，当 $n = 2$ 时成立．

11.54 (11.42)的一个均衡点 $\mathbf{a}$ 称为双曲形的,如果(11.51)中的矩阵 $\mathbf{A}$ 没有实部为 0 的特征值.

双曲形均衡点的定义.

11.55 (11.42)的一个双曲形均衡点或是不稳定的,或是渐近稳定的.

一个重要结论.

11.56 设(a, b)为方程组

$\dot{x}=f(x, y),\ \dot{y}=g(x, y)$

的一个均衡点,定义

$$\mathbf{A}(x, y)=\begin{pmatrix} f_1'(x, y) & f_2'(x, y) \\ g_1'(x, y) & g_2'(x, y) \end{pmatrix}$$

假设对于所有$(x, y)\in \mathbf{R}^2$,以下三个条件均满足:

(a) $\operatorname{tr}(\mathbf{A}(x, y))=f_1'(x, y)+g_2'(x, y)<0$

(b) $\det(\mathbf{A}(x, y))=$

$$\begin{vmatrix} f_1'(x, y) & f_2'(x, y) \\ g_1'(x, y) & g_2'(x, y) \end{vmatrix}>0$$

(c) $f_1'(x, y)g_2'(x, y)$ 和 $f_2'(x, y)g_1'(x, y)$ 中至少有一个在整个平面非零

则(a, b)是整体渐近稳定的.

Olech 定理.

11.57 $V(\mathbf{x})=V(x_1, \cdots, x_n)$ 是(11.42)的一个定义在一个包含均衡点 $\mathbf{a}$ 的开集 Ω 里的 Liapunov 函数,如果

- $V(\mathbf{x})>0$ 对于所有 Ω 内的 $\mathbf{x}\neq\mathbf{a}$ 成立, $V(\mathbf{a})=0$, 且
- $\dot{V}(\mathbf{x})=\sum_{i=1}^{n}\frac{\partial V(\mathbf{x})}{\partial x_i}\frac{\mathrm{d}x_i}{\mathrm{d}t}=\sum_{i=1}^{n}\frac{\partial V(\mathbf{x})}{\partial x_i}f_i(\mathbf{x})\leqslant 0$

对于所有 Ω 内的 $\mathbf{x}\neq\mathbf{a}$ 成立.

Liapunov 函数的定义.

11.58 设 $\mathbf{a}$ 为(11.42)的一个均衡点,假设方程组在包含 $\mathbf{a}$ 的开集 Ω 里存在一 Liapunov 函数 $V(\mathbf{x})$. 则 $\mathbf{a}$ 是一稳定的均衡点. 又如果 $\dot{V}(\mathbf{x})<0$ 对于所有在 Ω 内的 $\mathbf{x}\neq\mathbf{a}$ 成立. 则 $\mathbf{a}$ 局部渐近稳定的.

Liapunov 定理.

11.59 修正的 Lotka-Volterra 模型

$\dot{x}=kx-axy-\varepsilon x^2,\ \dot{y}=-hy+bxy-\delta y^2$

有一渐近稳定均衡点

$$(x_0, y_0)=\left(\frac{ah+k\delta}{ab+\delta\varepsilon}, \frac{bk-h\varepsilon}{ab+\delta\varepsilon}\right)$$

函数 $V(x, y)=H(x, y)-H(x_0, y_0)$,其中 $H(x, y)=b(x-x_0\ln x)+a(y-y_0\ln y)$ 是方程组的 Liapunov 函数,且在除均衡点外有 $\dot{V}(x, y)<0$.

应用(11.58)的一个例子:x 是兔子个数,y 狐狸个数. (a, b, h, k, δ, 及 ε 为正, $bk>h\varepsilon$.) $\varepsilon=\delta=0$ 给出一个经典 Lotka-Volterra 模型,其中 $\dot{V}=0$ 总是成立,而积分曲线则是围绕均衡点的封闭曲线.

11.60 设(a, b)为方程组

$\dot{x}=f(x, y),\ \dot{y}=g(x, y)$

的一个均衡点,定义 $\mathbf{A}$ 为(11.53)中的矩阵. 如果 $|\mathbf{A}|<0$, 则存在(至多 t 的一个转换)确切的两个解 $(x_1(t), y_1(t))$ 和 $(x_2(t), y_2(t))$,定义在区间$[t_0, \infty)$上,且收敛于(a, b). 这些解从相反方向收敛至(a, b),且都切于经过(a, b)点,平行于对应负特征值的直线. 这样的均衡点称为鞍点.

局部鞍点定理. ($|\mathbf{A}|<0$ 当且仅当 $\mathbf{A}$ 的特征值是实的且有相反符号. 关于整体鞍点,参见 Seierstad & Sydsæter (1987), 3.10,定理 9)

偏微分方程

11.61 求下列方程解的方法

$$(*)\ P(x,y,z)\frac{\partial z}{\partial x}+Q(x,y,z)\frac{\partial z}{\partial y}=R(x,y,z)$$

- 找到方程组的解

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{Q}{P},\ \frac{\mathrm{d}z}{\mathrm{d}x}=\frac{R}{P}$$

其中 x 是自由变量.如果解可表示为 $y=\varphi_1(x,C_1,C_2)$ 及 $z=\varphi_2(x,C_1,C_2)$,解出 C_1 和 C_2 得到 $C_1=u(x,y,z)$ 及 $C_2=v(x,y,z)$.

- 如果 Φ 是任意二元 C^1-函数,且函数 u 和 v 至少有一个包含 z,则由方程

$$\Phi(u(x,y,z),v(x,y,z))=0,$$

隐含定义的 $z=z(x,y)$ 是$(*)$的一个解.

一般拟线性一阶偏微分方程及一种求解方法.这种方法一般不能给出$(*)$的所有的解.(详细内容参见 Zachmanoglou & Thoe (1986),第 II 章)

11.62 下列偏微分方程组

$$\frac{\partial z(\boldsymbol{x})}{\partial x_1}=f_1(\boldsymbol{x},z(\boldsymbol{x}))$$

$$\frac{\partial z(\boldsymbol{x})}{\partial x_2}=f_2(\boldsymbol{x},z(\boldsymbol{x}))$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{\partial z(\boldsymbol{x})}{\partial x_n}=f_n(\boldsymbol{x},z(\boldsymbol{x}))$$

对于未知函数 $z(\boldsymbol{x})=z(x_1,\cdots,x_n)$ 有一解,当且仅当 $f_1,\cdots,f_n$ 对 $x_1,\cdots,x_n$ 的 $n\times n$ 阶一阶偏导数矩阵是对称的.

Frobenius 定理.函数 $f_1,\cdots,f_n$ 是 C^1 函数.

参 考 文 献

Braun(1993)是一般微分方程的很好的参考资料.也可参考 Pontryagin (1962).关于(11.28)—(11.31)参考 Takayama (1985)或 Gandolfo (1996).

Beavis & Dobbs (1990)有大多数的数字结论及经济应用,对于(11.61)可参考 Sneddon (1957)或 Zachmanoglou & Thoe (1986),对于(11.62)可参考 Hartman (1982),对于(11.62)的经济应用,请参考 Mas-Colell *et. al*.(1995).

12 欧氏空间拓扑学

12.1 $B(a;\ r) = \{x: \|x - a\| < r\}\ (r > 0)$

在 $\mathbf{R}^n$ 内以 r 为半径，a 为圆心的开 n 球的定义.（$\|\ \|$ 的定义见(18.14)）

12.2 • 一点 $a \in S \subset \mathbf{R}^n$ 是 S 的内点，如果一个以 a 为圆心的 n-球中的所有点都属于 S.

• 一点 $b \in \mathbf{R}^n$(不一定在 S 内)是 S 的边界点，如果每一个以 b 为圆心的 n 球至少包括 S 内的一点，同时至少包括 S 外的一点.

内点与边界点的定义.

12.3 在 $\mathbf{R}^n$ 内的一个集合 S 称为

• **开的** 如果它的所有点均为内点.

• **闭的** 如果 $\mathbf{R}^n \setminus S$ 是开的.

• **有界的** 如果存在一个数 M 使 $\|x\| \leqslant M$ 对于所有 S 内的 x 都成立.

• **紧的** 如果它既是闭的又是有界的.

重要的定义. $\mathbf{R}^n \setminus S = \{x \in \mathbf{R}^n: x \notin S\}$.

12.4 一个集合 $S \subset \mathbf{R}^n$ 是闭的，当且仅当它包含所有边界点. 由 S 和其所有边界点组成的集合 $\bar{S}$ 称为 S 的闭包.

闭集的有用特征，及集合的闭包的定义.

12.5 集合 $S \subset \mathbf{R}^n$ 称为 $\mathbf{R}^n$ 内一点 a 的邻域. 如果 a 是 S 的一个内点.

邻域的定义.

12.6 $\mathbf{R}^n$ 中的一个序列 $\{x_k\}$ 收敛到 x，如果对于每一 $\varepsilon > 0$ 存在一整数 N，使 $\|x_k - x\| < \varepsilon$ 对于所有 $k \geqslant N$ 成立.

$\mathbf{R}^n$ 中序列的收敛. 如果序列不收敛，则称为是发散的

12.7 $\mathbf{R}^n$ 中的序列 $\{x_k\}$ 是一个柯西序列，如果对于每一 $\varepsilon>0$ 存在一整数 N，使 $\|x_j-x_k\|<\varepsilon$ 对于所有 $j,k\geqslant N$ 都成立.

柯西序列的定义.

12.8 $\mathbf{R}^n$ 中的序列 $\{x_k\}$ 是收敛的，当且仅当它是一个柯西序列.

柯西收敛标准.

12.9 $\mathbf{R}^n$ 内的集合 S 是闭的，当且仅当 S 内的点组成的每一收敛序列 $\{x_k\}$ 的极限，$x=\lim\limits_{k\to\infty}x_k$ 也在 S 内.

闭集的特征.

12.10 设 $\{x_k\}$ 为 $\mathbf{R}^n$ 内的一序列，设 $k_1<k_2<k_3<\cdots$ 为一递增整数序列，则 $\{x_{k_j}\}_{j=1}^{\infty}$，称为 $\{x_k\}$ 的子序列.

子序列的定义.

12.11 $\mathbf{R}^n$ 内集合 S 是紧的，当且仅当每一个 S 中点的序列中都有一收敛到 S 中一点的子序列.

紧集的特征.

12.12 $f:M\subset\mathbf{R}^n\to\mathbf{R}$ 是在 M 中的点 a 处连续的，如果对于每一 $\varepsilon>0$，存在一 $\delta>0$，使

$|f(x)-f(a)|<\varepsilon$

对于所有 M 内满足 $\|x-a\|<\delta$ 的 x 都成立.

有 n 个变量的连续函数的定义.

12.13 函数 $f=(f_1,\cdots,f_m):M\subset\mathbf{R}^n\to\mathbf{R}^m$ 是在 M 内的点 a 处连续的，如果对于每一 $\varepsilon>0$ 存在一 $\delta>0$，使

$\|f(x)-f(a)\|<\varepsilon$

对于 M 内所有满足 $\|x-a\|<\delta$ 的 x 都成立.

有 n 个变量的连续向量函数的定义.

12.14 设 $f=(f_1,\cdots,f_m)$ 为一个从 $M\subset\mathbf{R}^n$ 到 $\mathbf{R}^m$ 的函数，且设 a 为 M 内一点，则

- f 是在 a 连续的，当且仅当根据定义(12.12)每一 f_i 在 a 都是连续的.
- f 在 a 是连续的，当且仅当 $f(x_k)\to f(a)$ 对于 M 内每一收敛至 a 的序列 $\{x_k\}$ 都成立.

有 n 个变量的连续向量函数的特征.

12.15	函数 $f:\mathbf{R}^n \to \mathbf{R}^m$ 是在 $\mathbf{R}^n$ 内每一点 x 连续的，当且仅当 $f^{-1}(T)$ 对于 $\mathbf{R}^m$ 内每一开(闭)的集合 T 是开(闭)的.	从 $\mathbf{R}^n$ 至 $\mathbf{R}^m$ 的向量函数的特征.
12.16	如果 f 是从 $\mathbf{R}^n$ 至 $\mathbf{R}^m$ 的连续函数，且 M 是 $\mathbf{R}^n$ 内一紧集，则 $f(M)$ 是紧的.	连续函数从紧集映射至紧集.
12.17	给定 $\mathbf{R}^n$ 内集合 S. 一个围绕点 $\boldsymbol{a} \in S$，r 为半径的相对球 $B^S(\boldsymbol{a};\ r)$，定义为 $B^S(\boldsymbol{a};\ r) = B(\boldsymbol{a};\ r) \cap S$.	相对球的定义.
12.18	相对内点，相对边界点，相对开集和闭集可通过这些概念的通常的方式定义，只要将 $\mathbf{R}^n$ 改为 S，球改为相对球.	相对拓扑概念.
12.19	• $U \subset S$ 在 $S \subset \mathbf{R}^n$ 内是相对开的，当且仅当存在 $\mathbf{R}^n$ 内的一个开集 V，使 $U = V \cap S$. • $F \subset S$ 在 $S \subset \mathbf{R}^n$ 内是相对闭的，当且仅当存在 $\mathbf{R}^n$ 内的一个闭集 H. 使 $F = H \cap S$.	集合 $S \subset \mathbf{R}^n$ 的相对开和相对闭的子集的特征.
12.20	一个从 $S \subset \mathbf{R}^n$ 到 $\mathbf{R}^m$ 的函数 f 是连续的，当且仅当下列条件有一个得到满足： • 对于每一在 $\mathbf{R}^m$ 内的开集合 U，$f^{-1}(U)$ 是在 S 内相对开的. • 对于每一在 $\mathbf{R}^m$ 内的闭集合 T，$f^{-1}(T)$ 是在 S 内相对闭的.	连续性的特征应用于定义域不是整个 $\mathbf{R}^n$ 的函数.
12.21	函数 $f:M \subset \mathbf{R}^n \to \mathbf{R}^m$ 称为在集合 $S \subset M$ 上一致连续的. 如果对于每一 $\varepsilon > 0$，存在一 $\delta > 0$ (依赖于 ε 但不依赖于 $\mathbf{x}$ 和 $\mathbf{y}$)，使 $\Vert f(\boldsymbol{x}) - f(\boldsymbol{y}) \Vert < \varepsilon$ 对于所有在 S 内满足 $\Vert \boldsymbol{x} - \boldsymbol{y} \Vert < \delta$ 的 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 成立.	从 $\mathbf{R}^n$ 至 $\mathbf{R}^m$ 的函数的一致连续性的定义.
12.22	如果 $f:M \subset \mathbf{R}^n \to \mathbf{R}^m$ 是连续的，且集合 $S \subset M$，S 是紧的，则 f 是在 S 上一致连续的.	在紧集上的连续函数是一致连续的.

12.23 设$\{f_n\}$为定义在$S\subset\mathbf{R}^n$上的函数序列，且值域为$\mathbf{R}^m$，序列$\{f_n\}$称为点式收敛至S内一函数f，如果序列$\{f_n(x)\}$(在$\mathbf{R}^m$内)，对于每一$x\in S$收敛至$f(x)$.

函数序列(点式)收敛的定义.

12.24 定义在$S\subset\mathbf{R}^n$上，且值域为$\mathbf{R}^m$的序列$\{f_n\}$被称为一致收敛至S内的一个函数f，如果对于每一$\varepsilon>0$存在一自然数$N(\varepsilon)$(依赖于ε但不依赖于x)，使

$\|f_n(x)-f(x)\|<\varepsilon$

对于所有$n\geqslant N(\varepsilon)$和所有在$S$内的$x$成立.

一致收敛函数序列的定义.

12.25 一个从集合A到集合B的对应F，是一个映射每一$x\in A$到B的一个非空子集$F(x)$的规则，F的图形是集合

$\mathrm{Graph}(F)=\{(a,\ b)\in A\times B:b\in F(a)\}$

对应和其图形的定义.

12.26 对应$F:X\subset\mathbf{R}^n\to\mathbf{R}^m$有一个封闭的图形，如果每一对收敛序列，$X$内的$\{x_k\}$和$\mathbf{R}^m$内的$\{y_k\}$，满足$y_k\in F(x_k)$且$\lim\limits_{k\to\infty}x_k=x\in X$，极限$\lim\limits_{k\to\infty}y_k$属于$F(x)$.

因此F有一封闭的图形，当且仅当Graph(F)是集合$X\times\mathbf{R}^m\subset\mathbf{R}^n\times\mathbf{R}^m$的一个相对闭的子集.

有封闭图形的对应的定义.

12.27 对应$F:X\subset\mathbf{R}^n\to\mathbf{R}^m$称为在$x^0$是下半连续的，如果对于在$F(x^0)$内的每一$y^0$和$y^0$的每一领域$U$，存在$x^0$的一个邻域$N$，使$F(x)\cap U\neq\varnothing$对于所有$x\in N\cap X$成立.

对应的下半连续的定义.

12.28 对应$F:X\subset\mathbf{R}^n\to\mathbf{R}^m$称为在$x^0$上半连续的，如果对每一包含$F(x^0)$的开集$U$，存在一个$x^0$的邻域$N$，使$F(x)\subset U$对于所有$x\in N\cap X$.

对应的上半连续的定义.

12.29	设 $F: X \subset \mathbf{R}^n \to K \subset \mathbf{R}^m$ 为一个对应,其中 K 紧的. 假设对每一 $x \in X$, 集合 $F(x)$是 K 的一个闭子集. 则 F 有一封闭的图形当且仅当 F 是上半连续的.	一个有趣的结论.

下确界和上确界

12.30	• 任何有上限约束的非空实数集合 S 都有一个最小上限 b^*,即 b^* 是 S 的一个上限,且 $b^* \leqslant b$ 对 S 的每一上限 b 成立. b^* 称为 S 的上确界,记作 $b^* = \sup S$. • 任何有下限约束的非空实数集合 S 都有一个最大下限 a^*,即 a^* 是 S 的一个下限,且 $a^* \geqslant a$ 对 S 的每一上限 a 成立. a^* 称为 S 的下确界,记作 $a^* = \inf S$.	实数集合的最小上限和最大下限原理. 如果 S 不受上限约束,记作 $\sup S = \infty$, 如果 S 不受下限约束, 记作 $\inf S = -\infty$. 通常定义 $\sup\varnothing = -\infty$ 和 $\inf\varnothing = \infty$.
12.31	$\inf\limits_{x\in B} f(x) = \inf\{f(x): x \in B\}$ $\sup\limits_{x\in B} f(x) = \sup\{f(x): x \in B\}$	定义在 $\mathbf{R}^n$ 中集合 B 上的实值函数的下确界和上确界的定义.
12.32	$\inf\limits_{x\in B}(f(x) + g(x)) \geqslant \inf\limits_{x\in B} f(x) + \inf\limits_{x\in B} g(x)$ $\sup\limits_{x\in B}(f(x) + g(x)) \leqslant \sup\limits_{x\in B} f(x) + \sup\limits_{x\in B} g(x)$	关于 sup 和 inf 的一些结论.
12.33	$\inf\limits_{x\in B}(\lambda f(x)) = \lambda \inf\limits_{x\in B} f(x)$ 如果 $\lambda > 0$ $\sup\limits_{x\in B}(\lambda f(x)) = \lambda \sup\limits_{x\in B} f(x)$ 如果 $\lambda > 0$	λ 是一实数.
12.34	$\sup\limits_{x\in B}(-f(x)) = -\inf\limits_{x\in B} f(x)$ $\inf\limits_{x\in B}(-f(x)) = -\sup\limits_{x\in B} f(x)$	
12.35	$\sup\limits_{(x,y)\in A\times B} f(x,y) = \sup\limits_{x\in A}(\sup\limits_{y\in B} f(x,y))$	$A \times B = \{(x,y): x \in A \wedge y \in B\}$

12.36	$\underline{\lim}_{x\to x^0} f(x) =$ $\lim_{r\to 0}(\inf\{f(x): 0<\Vert x-x^0\Vert<r,\ x\in M\})$ $\overline{\lim}_{x\to x^0} f(x) =$ $\lim_{r\to 0}(\sup\{f(x): 0<\Vert x-x^0\Vert<r,\ x\in M\})$	$\underline{\lim}=\lim\inf$ 和 $\overline{\lim}=\lim\sup$ 的定义. f 定义在 $M\subset\mathbf{R}^n$, 而 x^0 在 $M\setminus\{x^0\}$ 的闭包中.
12.37	$\underline{\lim}(f+g)\geqslant\underline{\lim}\,f+\underline{\lim}\,g$ $\overline{\lim}(f+g)\leqslant\overline{\lim}\,f+\overline{\lim}\,g$	如果右边部分有定义,则不等式成立.
12.38	$\underline{\lim}\,f\leqslant\overline{\lim}\,f$	lim inf 和 lim sup 的结论.
12.39	$\underline{\lim}\,f=-\overline{\lim}\,(-f)$, $\overline{\lim}\,f=-\underline{\lim}(-f)$	
12.40	设 f 为一定义在区间$[t_0,\infty)$上的实值函数. 则我们定义: • $\underline{\lim}_{t\to\infty} f(t)=\lim_{s\to\infty}\inf\{f(s): s\in[t_0,\infty)\}$ • $\overline{\lim}_{t\to\infty} f(t)=\lim_{s\to\infty}\sup\{f(s): s\in[t_0,\infty)\}$	$\underline{\lim}_{t\to\infty}$ 和 $\overline{\lim}_{t\to\infty}$ 的定义. 公式(12.37)—(12.39)仍然成立.
12.41	• $\underline{\lim}_{t\to\infty} f(t)\geqslant a \Leftrightarrow \begin{cases}\text{对于每一 }\varepsilon>0\text{,存在一}\\ t'\text{,使 }f(t)\geqslant a-\varepsilon\\ \text{对于所有 }t\geqslant t'\text{ 成立.}\end{cases}$ • $\overline{\lim}_{t\to\infty} f(t)\geqslant a \Leftrightarrow \begin{cases}\text{对于每一 }\varepsilon>0\text{ 和每一}\\ t'\text{,存在一 }t\geqslant t'\text{,使}\\ f(t)\geqslant a-\varepsilon\\ \text{对于所有 }t\geqslant t'\text{ 成立.}\end{cases}$	基本事实.

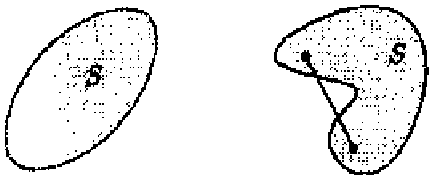
参 考 文 献

Bartle (1982),Marsden & Hoffman (1993),及 Rudin (1982)都是很好的普通拓扑学的参考文献.关于对应及其性质,可参考 Hildenbrand & Kirman (1976) 或 Hildenbrand (1974).

13 凸性

13.1 一个集合 S 在 $\mathbf{R}^n$ 内是凸的，如果
$x, y \in S$ 且 $\lambda \in [0, 1] \Rightarrow \lambda x + (1-\lambda)y \in S$

凸集的定义. 空集定义为凸的.

13.2 

第一个集合是凸的，但第二个不是.

13.3 如果 S 和 T 是 $\mathbf{R}^n$ 内的凸集，则
- $S \cap T = \{x : x \in S$ 及 $x \in T\}$ 是凸的
- $aS + bT = \{as + bt : s \in S, t \in T\}$ 是凸的

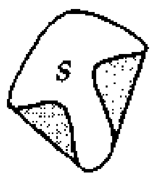
凸集的性质(a 和 b 是实数).

13.4 任何向量 $x = \lambda_1 x_1 + \cdots + \lambda_m x_m$，对 $\lambda_i \geq 0$ 对于 $i = 1, \cdots, m$ 成立，且 $\sum_{i=1}^{m} \lambda_i = 1$，称为一个在 $\mathbf{R}^n$ 内的向量 $\mathbf{x}_1, \cdots, \mathbf{x}_m$ 的凸组合.

向量凸组合的定义.

13.5 $\mathrm{co}(S) =$
{在 S 内有限个数向量的所有凸组合的集合.

$\mathrm{co}(S)$ 是在 $\mathbf{R}^n$ 内集合 S 的凸包.

13.6 

如果 S 是非阴影部分的集合，则 $\mathrm{co}(S)$ 包括加上去的阴影部分.

13.7 $\mathrm{co}(S)$ 是包含 S 的最小凸集.

凸包的一个有用的特征.

13.8 如果 $S \subset \mathbf{R}^n$ 且 $x \in \mathrm{co}(S)$，则 x 是在 S 内至多 $n+1$ 个点的凸组合.

Carathéodory 定理.

13.9 z 是凸集 S 的一个极点,如果 $z \in S$ 且不存在在 S 内的 x 和 y,以及在 $(0, 1)$ 内的 λ,能使 $x \neq y$ 且 $z = \lambda x + (1-\lambda)y$.

极点的定义.

13.10 设 S 为一个在 $\mathbf{R}^n$ 内紧的凸集.则 S 是它的极点的凸包.

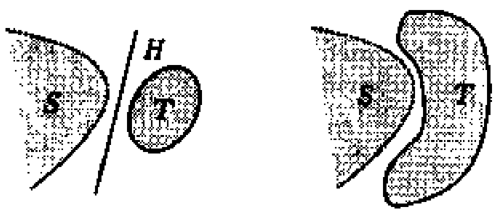
Krein-Milman 定理.

13.11 设 S 和 T 为两个在 $\mathbf{R}^n$ 内的不相连非空凸集,则 S 和 T 可以被一个超平面分开,即存在一个非零向量 $\boldsymbol{a}$,使

$\boldsymbol{a} \cdot \boldsymbol{x} \leqslant \boldsymbol{a} \cdot \boldsymbol{y}$ 对于所有在 S 内的 $\boldsymbol{x}$ 和所有在 T 内的 $\mathbf{y}$ 成立.

Minkowski 分离定理.超平面 $\{\boldsymbol{x}: \boldsymbol{a} \cdot \boldsymbol{x} = A\}$,对 $\boldsymbol{a} \cdot \boldsymbol{x} \leqslant A \leqslant \boldsymbol{a} \cdot \boldsymbol{y}$ 对于所有 $\boldsymbol{x} \in S$ 和所有 $\boldsymbol{y} \in T$ 成立,称为分离.

13.12



第一个图中 S 和 T 是被 H(严格)分离的.第二个图中,S 和 T 不能被超平面分离.

13.13 设 S 为在 $\mathbf{R}^n$ 中有内点的凸集,T 为 $\mathbf{R}^n$ 中的一个凸集,且在 $S \cap T$ 内(如果存在)的点都不是 S 的内点.则 S 和 T 可以被一个超平面分离,即存在一个向量 $\boldsymbol{a} \neq \mathbf{0}$,使

$\boldsymbol{a} \cdot \boldsymbol{x} \leqslant \boldsymbol{a} \cdot \boldsymbol{y}$ 对于所有 $\boldsymbol{x} \in S$ 和所有 $\boldsymbol{y} \in T$ 成立.

在 $\mathbf{R}^n$ 内的一个一般分离定理.

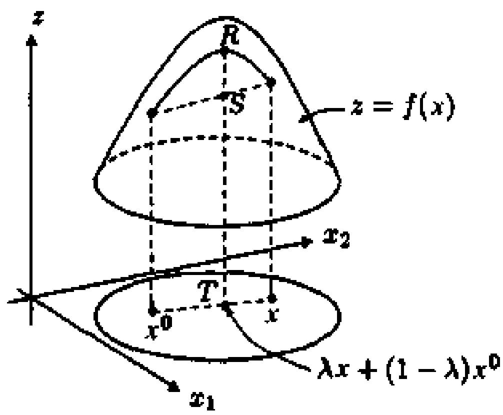
凹函数和凸函数

13.14 定义在 $\mathbf{R}^n$ 内一个凸集 S 上的 $f(\boldsymbol{x}) = f(x_1, \cdots, x_n)$ 在 S 上是凹的,如果

$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{x}^0) \geqslant \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{x}^0)$

对于所有 $\boldsymbol{x}$, $\boldsymbol{x}^0 \in S$ 和所有 $\lambda \in (0, 1)$ 成立.

要定义一个凸函数,改变不等号方向.相应的,f 是凸的当且仅当 $-f$ 是凹的.

13.15		函数 $f(\boldsymbol{x})$是(严格)凹的. $TR = f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{x}^0) \geqslant TS = \lambda f(\boldsymbol{x}) + (1-\lambda)f(\boldsymbol{x}^0)$. ($TR$ 和 TS 是 R 和 S 超过 $\boldsymbol{x}$ 所在的平面的高度.如果所有点在 $\boldsymbol{x}$ 所在的平面以下,高度是负的.)
13.16	$f(\boldsymbol{x})$是严格凹的,如果 $f(\boldsymbol{x})$是凹的,且(13.14)中的不等号≥当 $\boldsymbol{x} \neq \boldsymbol{x}^0$ 时严格成立.	严格凹函数的定义.对于严格凸性,改变不等号方向.
13.17	如果定义在 $\mathbf{R}^n$ 内的凸集 S 上的 $f(\boldsymbol{x})$是凹的(凸的),则 $f(\boldsymbol{x})$在 S 的每一内点上连续.	凹函数和凸函数的连续性.
13.18	• 如果 $f(\boldsymbol{x})$和 $g(\boldsymbol{x})$是凹的(凸的),且 a 和 b 是非负数,则 $af(\boldsymbol{x}) + bg(\boldsymbol{x})$ 是凹的(凸的). • 如果 $f(\boldsymbol{x})$是凹的,$F(u)$是凹且递增的,则 $U(\boldsymbol{x}) = F(f(\boldsymbol{x}))$ 是凹的. • 如果 $f(\boldsymbol{x}) = \boldsymbol{a}\cdot\boldsymbol{x} + b$ 且 $F(u)$是凹的,则 $U(\boldsymbol{x}) = F(f(\boldsymbol{x}))$ 凹的. • 如果 $f(\boldsymbol{x})$是凸的,$F(u)$是凸且递增的,则 $U(\boldsymbol{x}) = F(f(\boldsymbol{x}))$ 是凸的. • 如果 $f(\boldsymbol{x}) = \boldsymbol{a}\cdot\boldsymbol{x} + b$ 且 $F(u)$是凸的,则 $U(\boldsymbol{x}) = F(f(\boldsymbol{x}))$ 是凸的.	凹函数和凸函数的性质.
13.19	一个 C^1 函数 $f(\boldsymbol{x})$是在 $\mathbf{R}^n$ 内的一个开凸集 S 上凹的,当且仅当 $$f(\boldsymbol{x}) - f(\boldsymbol{x}^0) \leqslant \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{x}^0)}{\partial x_i}(x_i - x_i^0)$$ 或相应的 $$f(\boldsymbol{x}) - f(\boldsymbol{x}^0) \leqslant \nabla f(\boldsymbol{x}^0)\cdot(\boldsymbol{x} - \boldsymbol{x}^0)$$ 对于所有 S 中的 $\boldsymbol{x}$ 和 $\boldsymbol{x}^0$ 成立.	C^1 函数的凹性.对于凸性,改变不等号方向.

13.20 一个 C^1 函数 $f(\mathbf{x})$是在 $\mathbf{R}^n$ 内的一个开凸集 S 上严格凹的，当且仅当(13.19)里的不等号对于 $\mathbf{x} \neq \mathbf{x}^0$ 严格成立.

C^1 函数的严格凹性.对于严格凸性，改变不等号方向.

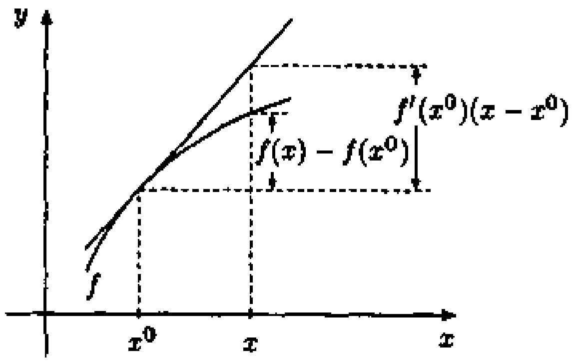
13.21 一个 C^1 函数 $f(x)$在一个开区间 I 上是凹的，当且仅当

$$f(x) - f(x^0) \leqslant f'(x^0)(x - x^0)$$

对于所有 I 中的 x 和 x^0 成立.

(13.19)的一元形式.

13.22



(13.21)的几何诠释. C^1 函数 f 是凹的，当且仅当 f 的图形的所有点在切线下.(图中的 f 实际上是严格凹的.)

13.23 一个 C^1 函数 $f(x, y)$在 xy 平面上的开凸集 S 里是凹的，当且仅当

$$\begin{aligned} f(x, y) - f(x^0, y^0) &\leqslant f_1'(x^0, y^0)(x - x^0) + \\ &\quad f_2'(x^0, y^0)(y - y^0) \end{aligned}$$

对于所有(x, y), $(x^0, y^0) \in S$ 成立.

(13.19)的二元形式.

13.24

$$\mathbf{f}''(\mathbf{x}) = \begin{pmatrix} f_{11}''(\mathbf{x}) & f_{12}''(\mathbf{x}) & \cdots & f_{1n}''(\mathbf{x}) \\ f_{21}''(\mathbf{x}) & f_{22}''(\mathbf{x}) & \cdots & f_{2n}''(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1}''(\mathbf{x}) & f_{n2}''(\mathbf{x}) & \cdots & f_{nn}''(\mathbf{x}) \end{pmatrix}$$

f 在 $\mathbf{x}$ 点的 *Hessian* 矩阵.如果 f 是 C^2 函数，则 Hessian 矩阵是对称的.

13.25 Hessian 矩阵 $\mathbf{f}''(\mathbf{x})$的 r 阶主子式 $\Delta_r(\mathbf{x})$是通过先删除矩阵的任意 $n - r$ 行，再删除相同数目的列所得到的子矩阵的行列式.

Hessian 矩阵的主子式.(参见(20.15))

13.26 一个 C^2 函数 $f(\mathbf{x})$在 $\mathbf{R}^n$ 中开凸集 S 上是凹的，当且仅当对于所有在 S 内的 $\mathbf{x}$ 和所有 Δ_r，

$$(-1)^r \Delta_r(\mathbf{x}) \geqslant 0 \text{ 对 } r = 1, \cdots, n \text{ 成立}.$$

C^2 函数的凹性.

13.27	一个 C^2 函数 $f(\mathbf{x})$ 在 $\mathbf{R}^n$ 中开凸集 S 上是凸的，当且仅当对于所有在 S 内的 $\mathbf{x}$ 和所有 Δ_r， $\Delta_r(\mathbf{x}) \geqslant 0$ 对 $r = 1, \cdots, n$ 成立.	C^2 函数的凸性.
13.28	$$D_r(\mathbf{x}) = \begin{vmatrix} f''_{11}(\mathbf{x}) & f''_{12}(\mathbf{x}) & \cdots & f''_{1r}(\mathbf{x}) \\ f''_{21}(\mathbf{x}) & f''_{22}(\mathbf{x}) & \cdots & f''_{2r}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f''_{r1}(\mathbf{x}) & f''_{r2}(\mathbf{x}) & \cdots & f''_{rr}(\mathbf{x}) \end{vmatrix}$$	f 在 $\mathbf{x}$ 点的 Hessian 矩阵的前主子式，对 $r = 1, 2, \cdots, n$.
13.29	一个 C^2 函数 $f(\mathbf{x})$ 在 $\mathbf{R}^n$ 中的开凸集 S 上是严格凹的，如果对于所有 $\mathbf{x} \in S$， $(-1)^r D_r(\mathbf{x}) > 0$ 对 $r = 1, \cdots, n$ 成立.	C^2 函数凹性的充分(但不是必要)条件.
13.30	一个 C^2 函数 $f(\mathbf{x})$ 在 $\mathbf{R}^n$ 中的开凸集 S 上是严格凸的，如果对于所有 $\mathbf{x} \in S$， $D_r(\mathbf{x}) > 0$ 对 $r = 1, \cdots, n$ 成立.	C^2 函数凸性的充分(但不是必要)条件.
13.31	假设 $f(x)$ 是在开区间 I 上的 C^2 函数，则 • $f(x)$ 在 I 上是凹的 $\Leftrightarrow f''(x) \leqslant 0$ 对于所有 I 中的 x 成立. • $f(x)$ 在 I 上是凸的 $\Leftrightarrow f''(x) \geqslant 0$ 对于所有 I 中的 x 成立. • $f''(x) < 0$ 对于所有 $x \in I$ 成立 $\Rightarrow f(x)$ 在 I 上严格凹的. • $f''(x) > 0$ 对于所有 $x \in I$ 成立 $\Rightarrow f(x)$ 在 I 上严格凸的.	(13.26),(13.27),(13.29)和(13.30)的一元形式. 隐含箭头不能改为等价箭头.($f(x) = -x^4$ 是严格凹的，但 $f''(0) = 0$. $f(x) = x^4$ 是严格凸的，但 $f''(0) = 0$)
13.32	一个 C^2 函数 $f(x, y)$ 在 xy 平面里的开凸集 S 上是凹的，当且仅当 $f''_{11}(x, y) \leqslant 0$, $f''_{22}(x, y) \leqslant 0$ 且 $f''_{11}(x, y) f''_{22}(x, y) - (f''_{12}(x, y))^2 \geqslant 0$ 对于所有 $(x, y) \in S$ 成立.	(13.26)的二元形式. 对于 C^2 函数的凸性，改变头两个不等号的方向.

13.33 一个 C^2 函数 $f(x, y)$ 在 xy 平面里的开凸集 S 上是严格凹的，当(但不是仅当)

$f''_{11}(x, y) < 0$ 且

$f''_{11}(x, y)f''_{22}(x, y) - (f''_{12}(x, y))^2 > 0$

对于所有 $(x, y) \in S$ 成立.

(13.29)的二元形式.(注意两个不等号蕴涵 $f''_{22}(x, y) < 0$) 对于严格凸性，改变第一个不等号方向.

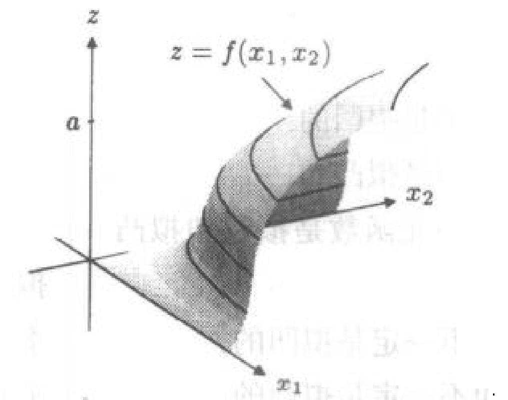
拟凹和拟凸函数

13.34 $f(\mathbf{x})$ 在凸集 $S \subset \mathbf{R}^n$ 上是拟凹的，如果(上)水平集合

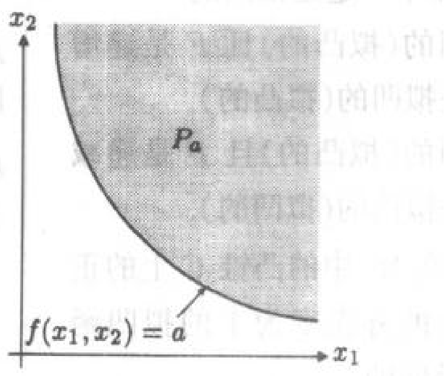
$$P_a = \{\mathbf{x} \in S : f(\mathbf{x}) \geq a\}$$

对于每一实数 a 是凸的.

拟凹函数的定义.(上水平集合也称为上等高线集合.)

13.35 

一个典型的二元拟凹函数，$z = f(x_1, x_2)$.

13.36 

(13.35)中的上水平集合，$P_a = \{(x_1, x_2) \in S : f(x_1, x_2) \geq a\}$.

13.37 $f(\mathbf{x})$ 是在 $\mathbf{R}^n$ 中开凸集 S 上的拟凹函数，当且仅当

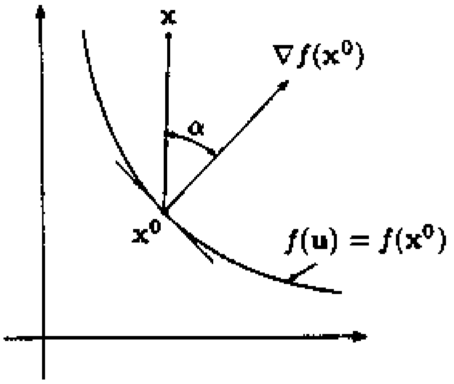
$$f(\mathbf{x}) \geq f(\mathbf{x}^0) \Rightarrow f(\lambda\mathbf{x} + (1-\lambda)\mathbf{x}^0) \geq f(\mathbf{x}^0)$$

对于所有 $\mathbf{x}, \mathbf{x}^0 \in S$ 和所有 $\lambda \in [0, 1]$ 成立.

拟凹函数的特征.

13.38	$f(\boldsymbol{x})$是在 $\mathbf{R}^n$ 中开凸集 S 上的严格拟凹函数，如果 $$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}^0) \Rightarrow f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{x}^0) > f(\boldsymbol{x}^0)$$ 对于所有 $\boldsymbol{x} \neq \boldsymbol{x}^0 \in S$ 和所有 $\lambda \in [0, 1]$成立.	严格拟凹函数的(最普通的)定义.
13.39	$f(\boldsymbol{x})$是在 $S \subset \mathbf{R}^n$ 上(严格)拟凸的，如果 $-f(\boldsymbol{x})$ 是(严格)拟凹的.	(严格)拟凸函数的定义.
13.40	如果 $f_1, \cdots, f_m$ 是定义在 $\mathbf{R}^n$ 中的凸集 S 上的凹函数，g 对于每一 $\boldsymbol{x} \in S$ 定义为 $$g(\boldsymbol{x}) = F(f_1(\boldsymbol{x}), \cdots, f_m(\boldsymbol{x}))$$ 且 $F(u_1, \cdots, u_m)$对每一变量是拟凹和递增的，则 g 是拟凹的.	一个有用的结论.
13.41	(1) $f(\boldsymbol{x})$是凹的 $\Rightarrow$ $f(\boldsymbol{x})$是拟凹的. (2) $f(\boldsymbol{x})$是凸的 $\Rightarrow$ $f(\boldsymbol{x})$是拟凸的. (3) 任意递增或递减的一元函数是拟凹和拟凸的. (4) 一组拟凹函数的和不一定是拟凹的. (5) 一组拟凸函数的和不一定是拟凸的. (6) 如果 $f(\boldsymbol{x})$是拟凹的(拟凸的)且 F 是递增的，则 $F(f(\boldsymbol{x}))$是拟凹的(拟凸的). (7) 如果 $f(\boldsymbol{x})$是拟凹的(拟凸的)且 F 是递减的，则 $F(f(\boldsymbol{x}))$是拟凸的(拟凹的). (8) 如果 $f(\boldsymbol{x})$是定义在 $\mathbf{R}^n$ 中的凸锥 C 上的正函数，f 是在 C 上的齐次度为 1 的拟凹函数，则 f 在 C 上是凹的.	拟凹拟凸函数的基本事实((4)的例子：$f(x) = x^3$ 和 $g(x) = -x$ 都是拟凹的，但 $f(x) + g(x) = x^3 - x$ 不是).
13.42	一个 C^1 函数 $f(\boldsymbol{x})$在 $\mathbf{R}^n$ 中的开凸集 S 上是拟凹的，当且仅当 $$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}^0) \Rightarrow \nabla f(\boldsymbol{x}^0) \cdot (\boldsymbol{x} - \boldsymbol{x}^0) \geqslant 0$$ 对于所有 $\boldsymbol{x}$ 和 $\boldsymbol{x}^0 \in S$ 成立.	C^1 函数的拟凹性.

13.43



(13.42)的几何诠释. 这儿 $\nabla f(\boldsymbol{x}^0)\cdot(\boldsymbol{x}-\boldsymbol{x}^0)\geqslant 0$ 是指角 α 是锐角,即角度小于 90°.

13.44 $$B_r(\boldsymbol{x})=\begin{vmatrix} 0 & f_1'(\boldsymbol{x}) & \cdots & f_r'(\boldsymbol{x}) \\ f_1'(\boldsymbol{x}) & f_{11}''(\boldsymbol{x}) & \cdots & f_{1r}''(\boldsymbol{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_r'(\boldsymbol{x}) & f_{r1}''(\boldsymbol{x}) & \cdots & f_{rr}''(\boldsymbol{x}) \end{vmatrix}$$

对应于 f 在 $\boldsymbol{x}$ 点的有界 *Hessian* 矩阵.

13.45 如果 $f(\boldsymbol{x})$是在 $\mathbf{R}^n$ 中的开凸集 S 上的拟凹函数,则

$(-1)^r B_r(\boldsymbol{x})\geqslant 0$ 对于 $r=1,\cdots,n$

对于所有 $\boldsymbol{x}\in S$ 成立.

C^2 函数拟凹性的必要条件.

13.46 如果 $(-1)^r B_r(\boldsymbol{x})>0$ 且 $r=1,\cdots,n$ 对于所有在 $\mathbf{R}^n$ 中的开凸集 S 上的 $\boldsymbol{x}$ 成立,则 $f(\boldsymbol{x})$在 S 上是拟凹的.

C^2 函数拟凹性的充分条件.

13.47 如果 $f(\boldsymbol{x})$是在 $\mathbf{R}^n$ 中的开凸集 S 上的拟凸函数,则

$B_r(\boldsymbol{x})\leqslant 0,\ r=1,\cdots,n$

对于所有 $\boldsymbol{x}\in S$ 成立.

C^2 函数拟凸性的必要条件.

13.48 如果 $B_r(\boldsymbol{x})<0$ 且 $r=1,\cdots,n$,对于所有在 $\mathbf{R}^n$ 中的开凸集 S 上的 $\boldsymbol{x}$ 成立,则 $f(\boldsymbol{x})$在 S 上是拟凸的.

C^2 函数拟凸性的充分条件.

伪凹函数和伪凸函数

13.49 一个定义在 $\mathbf{R}^n$ 中的凸集 S 上的 C^1 函数 $f(\mathbf{x})$ 在 S 上的点 $\mathbf{x}^0$ 是伪凹的，如果

$(*)$ $f(\mathbf{x}) > f(\mathbf{x}^0) \Rightarrow \nabla f(\mathbf{x}^0)\cdot(\mathbf{x}-\mathbf{x}^0) > 0$

对于所有 $\mathbf{x}\in S$ 成立. $f(\mathbf{x})$是在集合 S 上伪凹的，如果$(*)$对于所有 $\mathbf{x}$ 和 $\mathbf{x}^0\in S$ 成立.

要定义伪凸函数，将$(*)$中第一个不等号逆转.(比较(13.42)中拟凹性的特征.)

13.50 设 $f(\mathbf{x})$为一个定义在 $\mathbf{R}^n$ 中凸集 S 上的 C^1 函数，则

- 如果 f 在 S 是伪凹的，则 f 在 S 上是拟凹的.
- 如果 S 是开的且如果 $\nabla f(\mathbf{x}) \neq 0$ 对于所有在 S 上的 $\mathbf{x}$ 成立，则 f 在 S 上是伪凹的，当且仅当 f 在 S 上是拟凹的.

伪凹函数和拟凹函数的重要关系.

13.51 设 S 为 $\mathbf{R}^n$ 中的开凸集，且设 $f: S\to\mathbf{R}$ 为一个伪凹函数，如果 $\mathbf{x}^0\in S$ 有性质

$\nabla f(\mathbf{x}^0)\cdot(\mathbf{x}-\mathbf{x}^0)\leqslant 0$ 对于所有 $\mathbf{x}\in S$ 成立

(当 $\nabla f(\mathbf{x}^0)=\mathbf{0}$ 时确实成立)，则 $\mathbf{x}^0$ 是 f 在S中的整体最大值点.

为什么引进伪凹性的一个原因.

参 考 文 献

对于(拟)凹和(拟)凸函数，参见 Simone & Blume (1994)或 Sydsaeter & Hammond(1995)，关于伪凹和伪凸函数，参考 Simone & Blume (1994)和他们的参考文献，关于凸集的特殊结论，参见 Nikaido (1968)和 Takayama (1985)，关于凸性的标准参考文献是 Rockafellar (1970).

14 经典最优化理论

14.1 $f(\boldsymbol{x}) = f(x_1, \cdots, x_n)$ 在 $\boldsymbol{x}^0 = (x_1^0, \cdots, x_n^0) \in S$ 有一最大(最小)值,如果

$$f(\boldsymbol{x}^0) - f(\boldsymbol{x}) \geqslant 0 (\leqslant 0) \text{ 对于所有 } \boldsymbol{x} \in S \text{ 成立}$$

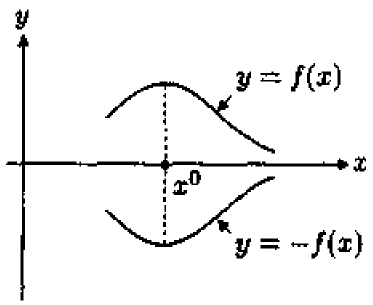
$\boldsymbol{x}^0$ 称为一个最大(最小)值点,$f(\boldsymbol{x}^0)$称为一个最大(最小)值.

n 元函数的(整体)最大(最小)的定义.我们统称其为最优值点和最优值,或极值点和极值.

14.2 $\boldsymbol{x}^0$ 在 S 上将$f(\boldsymbol{x})$最大化,当且仅当 $\boldsymbol{x}^0$ 在 S 上将 $-f(\boldsymbol{x})$最小化.

用来将最小化问题转变为最大化问题.

14.3



(14.2)的图示.x^0 将 $f(x)$最大化,当且仅当 x^0 将 $-f(x)$最小化.

14.4 假设 $f(\boldsymbol{x})$定义在 $S \subset \mathbf{R}^n$ 且 $F(u)$在 f 的值域上严格递增.则 $\boldsymbol{x}^0$ 在 S 上最大化(最小化) $f(\boldsymbol{x})$,当且仅当 $\boldsymbol{x}^0$ 在 S 上最大化(最小化) $F(f(\boldsymbol{x}))$.

一个重要的现象.

14.5 如果 $f: S \to \mathbf{R}$ 在 $\mathbf{R}^n$ 内一个封闭有界集合 S 上是连续的,则在 S 上存在 f 的最大值点和最小值点.

极值定理(或 *Weierstrass* 定理).

14.6 $\boldsymbol{x}^0 = (x_1^0, \cdots, x_n^0)$ 是 $f(\boldsymbol{x})$的一个驻点,如果

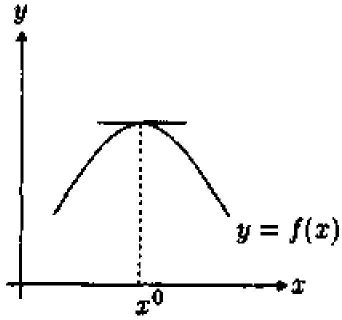
$$f_1'(\boldsymbol{x}^0) = 0,\ f_2'(\boldsymbol{x}^0) = 0,\ \cdots,\ f_n'(\boldsymbol{x}^0) = 0$$

n 元函数的驻点的定义.

14.7 设定义在 $\mathbf{R}^n$ 中一个凸集 S 上的 $f(\boldsymbol{x})$ 为凹的(凸的),并设 $\boldsymbol{x}^0$ 为 S 的一个内点.则 $\boldsymbol{x}^0$ 在 S 上最大化(最小化)$f(\boldsymbol{x})$,当且仅当 $\boldsymbol{x}^0$ 是一个驻点.

凹(凸)函数的最大(最小)化.

14.8



(14.7)的一元函数图示. f 是凹的, $f'(x^0)=0$, x^0 是一个最大值点.

14.9 如果 $f(\boldsymbol{x})$ 在 $S \subset \mathbf{R}^n$ 上有一最大或最小值,则最大值点/最小值点可在下列点找到:

- S 中是驻点的内点,
- 在 S 边界上的极值点,
- 在 S 内的 f 不可微的点.

何处寻找(整体)最大值点或最小值点.

14.10 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}^0$ 有一局部最大(最小)值,如果

$$(*)\ f(\boldsymbol{x}^0) - f(\boldsymbol{x}) \geqslant 0 (\leqslant 0)$$

对于所有在 S 内充分接近 $\boldsymbol{x}^0$ 的 $\boldsymbol{x}$ 成立.更精确地说,存在一 n 球 $B(\boldsymbol{x}^0; r)$ 使$(*)$对于所有在 $B(\boldsymbol{x}^0; r)$内的 $\boldsymbol{x}$ 成立.

n 元函数的局部(或相对)最大(最小)值点的定义.统称为局部极值点.

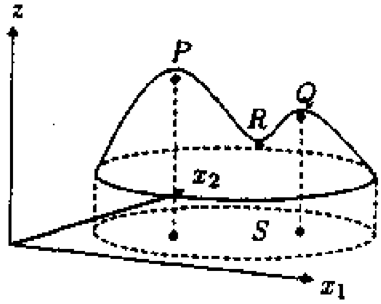
14.11 如果 $f(\boldsymbol{x}) = f(x_1, \cdots, x_n)$ 在 S 的内点 $\boldsymbol{x}^0$ 有一局部最大(最小)值,则 $\boldsymbol{x}^0$ 是 f 的一个驻点.

可微函数的一阶条件.

14.12 $f(\boldsymbol{x}) = f(x_1, \cdots, x_n)$ 的一个驻点 $\boldsymbol{x}^0$ 称为鞍点,如果它既不是局部最大值点也不是局部最小值点,即如果每一 n 球 $B(\boldsymbol{x}^0; r)$包含使 $f(\boldsymbol{x}) < f(\boldsymbol{x}^0)$ 的点 $\boldsymbol{x}$ 和使 $f(\boldsymbol{z}) > f(\boldsymbol{x}^0)$ 的其他点 $\boldsymbol{z}$.

鞍点的定义.

14.13



点 P,Q 和 R 都是驻点. P 是最大值点,Q 是局部最大值点,而 R 是鞍点.

一元函数的特殊结论

14.14 如果 $f(x)$在区间 I 上可微,则

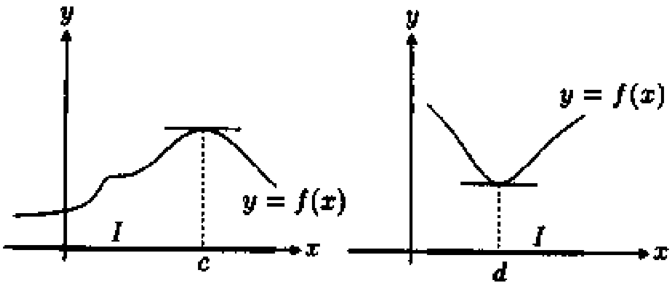
- $f'(x) > 0 \Longrightarrow f(x)$ 是单调递增的
- $f'(x) \geqslant 0 \Longleftrightarrow f(x)$ 是递增的
- $f'(x) = 0 \Longleftrightarrow f(x)$ 是不变的
- $f'(x) \leqslant 0 \Longleftrightarrow f(x)$ 是递减的
- $f'(x) < 0 \Longleftrightarrow f(x)$ 是单调递减的

重要的事实.隐含箭头不能逆转.($f(x) = x^3$ 是单调递增的,但 $f'(0) = 0$. $g(x) = -x^3$ 是单调递减的,但 $g'(0) = 0$)

14.15
- 如果 $f'(x) \geqslant 0$ 对于 $x \leqslant c$ 成立,且 $f'(x) \leqslant 0$ 对于 $x \geqslant c$ 成立,则 $x = c$ 是f 的一个最大值点.
- 如果 $f'(x) \leqslant 0$ 对于 $x \leqslant c$ 成立,且 $f'(x) \geqslant 0$ 对于 $x \geqslant c$ 成立,则 $x = c$ 是f 的一个最小值点.

(整体)最大/最小的一阶导数测试(通常在基本经济数学教材里被忽视).

14.16

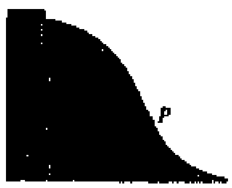


(14.15)的一元函数图示. c 是最大点. d 是最小点.

14.17 c 是$f(x)$的一个拐点如果 $f''(x)$在 c 改变符号.

一元函数拐点的定义.

14.18



拐点的非正统图式.斜率最陡的点 P 是一拐点.

14.19 设 f 为一个在区间 I 上有连续二阶导数的函数,假设 c 是I 的一个内点.则

- c 是f 的一个拐点$\Rightarrow f''(c) = 0$
- $f''(c) = 0$ 且 f''在c 改变符号 $\Rightarrow c$ 是f 的一个拐点

拐点测试.

二阶条件

14.20 如果 $f(\mathbf{x})=f(x_1,\cdots,x_n)$ 在 $\mathbf{x}^0$ 有一局部最大(最小)值,则

$$\sum_{i=1}^{n}\sum_{j=1}^{n}f''_{ij}(\mathbf{x}^0)h_ih_j\leqslant 0\ (\geqslant 0)$$

对于所有 $h_1,\cdots,h_n$ 成立.

局部最大(最小)的一个必要(二阶)条件.

14.21 如果 $\mathbf{x}^0=(x_1^0,\cdots,x_n^0)$ 是 $f(x_1,\cdots,x_n)$的一个驻点,如果 $D_k(\mathbf{x}^0)$是以下行列式,

$$D_k(\mathbf{x}^0)=\begin{vmatrix} f''_{11}(\mathbf{x}^0) & f''_{12}(\mathbf{x}^0) & \cdots & f''_{1k}(\mathbf{x}^0)\\ f''_{21}(\mathbf{x}^0) & f''_{22}(\mathbf{x}^0) & \cdots & f''_{2k}(\mathbf{x}^0)\\ \vdots & \vdots & \ddots & \vdots\\ f''_{k1}(\mathbf{x}^0) & f''_{k2}(\mathbf{x}^0) & \cdots & f''_{kk}(\mathbf{x}^0)\end{vmatrix}$$

则

- 如果 $(-1)^kD_k(\mathbf{x}^0)>0$ 对于 $k=1,\cdots,n$ 成立,则 $\mathbf{x}^0$ 是一局部最大值点.
- 如果 $D_k(\mathbf{x}^0)>0$ 对于 $k=1,\cdots,n$ 成立,则 $\mathbf{x}^0$ 是一局部最小值点.
- 如果 $D_n(\mathbf{x}^0)\neq 0$ 且以上两个条件均不满足,则 $\mathbf{x}^0$ 是一鞍点.

n 元 C^2 函数的驻点的分类. 局部最大/最小的二阶条件.

14.22 $f'(x^0)=0$ 且 $f''(x^0)<0\Rightarrow$ x^0 是 f 的一个局部最大值点.

$f'(x^0)=0$ 且 $f''(x^0)>0\Rightarrow$ x^0 是 f 的一个局部最小值点.

一元函数局部最大/最小的二阶条件.

14.23 如果(x_0,y_0)是 $f(x,y)$的一个驻点且 $D=f''_{11}(x_0,y_0)f''_{22}(x_0,y_0)-(f''_{12}(x_0,y_0))^2$,则

- $f''_{11}(x_0,y_0)>0$ 且 $D>0\Rightarrow$ (x_0,y_0)是 f 的一局部最小值点.
- $f''_{11}(x_0,y_0)<0$ 且 $D>0\Rightarrow$ (x_0,y_0)是 f 的一局部最大值点.
- $D<0\Rightarrow(x_0,y_0)$ 是 f 的一鞍点.

二元函数局部最大/最小的二阶条件. (二元 C^2 函数驻点的分类.)

等式约束下的最优化

14.24 $\max(\min) f(x, y)$ 约束条件为 $g(x, y) = b$

拉格朗日问题. 两个变量,一个约束.

14.25 拉格朗日方法. 解(14.24)的方法:

(1) 引进拉格朗日函数

$$L(x, y) = f(x, y) - \lambda(g(x, y) - c)$$

其中 λ 是一常数.

(2) 将 L 对 x 和 y 微分,令偏导数为0.

(3) (2)中的两个方程与约束条件一起,产生下列三个方程

$$f_1'(x, y) = \lambda g_1'(x, y)$$

$$f_2'(x, y) = \lambda g_2'(x, y)$$

$$g(x, y) = c$$

(4) 解对三个未知变量 x, y 和 λ 的三个方程. 这样求出解决问题的所有可能的 (x, y) 组合.

(14.24)的必要条件. 假设 $g_1'(x, y)$ 和 $g_2'(x, y)$ 并不都消没, 对于更精确的表达, 参见(14.27). λ 称为拉格朗日乘数.

14.26 假设 (x_0, y_0) 满足(14.25)中的条件. 则

(1) 如果 $L(x, y)$ 是凹的,则 (x_0, y_0) 解决了(14.24)中的最大化问题.

(2) 如果 $L(x, y)$ 是凸的,则 (x_0, y_0) 解决了(14.24)中的最小化问题.

解(14.24)中问题的充分条件.

14.27 假设 $f(x, y)$ 和 $g(x, y)$ 在 xy 平面内的区域 S 上都是 C^1 函数,且 (x_0, y_0) 既是 S 的一个内点,也是 $f(x, y)$ 受 $g(x, y) = c$ 约束下的一个局部极值点.

进一步假设 $g_1'(x_0, y_0)$ 和 $g_2'(x_0, y_0)$ 不同时为0,则存在一个唯一的数 λ 使拉格朗日函数

$$L(x, y) = f(x, y) - \lambda(g(x, y) - c)$$

在 (x_0, y_0) 有一驻点.

拉格朗日乘数方法的精确描述(拉格朗日定理).

14.28 考虑问题

局部 max(min)$f(x, y)$ s. t. $g(x, y) = c$

其中(x_0, y_0)满足(14.25)中的一阶条件. 定义有界 Hessian 行列式 $D(x, y)$为

$$D(x, y) \stackrel{!}{=} \begin{vmatrix} 0 & g_1' & g_2' \\ g_1' & f_{11}'' - \lambda g_{11}'' & f_{12}'' - \lambda g_{12}'' \\ g_2' & f_{21}'' - \lambda g_{21}'' & f_{22}'' - \lambda g_{22}'' \end{vmatrix}$$

(1) 如果 $D(x_0, y_0) > 0$,则(x_0, y_0)解决了局部最大化问题.

(2) 如果 $D(x_0, y_0) < 0$,则(x_0, y_0)解决了局部最小化问题.

拉格朗日问题的局部充分条件.

14.29 max(min)$f(x_1, \cdots, x_n)$ s. t.

$$\begin{cases} g_1(x_1, \cdots, x_n) = b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ g_m(x_1, \cdots, x_n) = b_m \end{cases}$$

一般拉格朗日问题. 假设 $m < n$.

14.30 拉格朗日方法. 解(14.29)的方法

(1) 引进拉格朗日函数

$$L(\boldsymbol{x}) = f(\boldsymbol{x}) - \sum_{j=1}^{m} \lambda_j (g_j(\boldsymbol{x}) - b_j)$$

其中 $\lambda_1, \cdots, \lambda_m$ 是常数.

(2) 令 L 的偏导为 0,

$$\frac{\partial L(\boldsymbol{x})}{\partial x_k} = \frac{\partial f(\boldsymbol{x})}{\partial x_k} - \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(\boldsymbol{x})}{\partial x_k} = 0$$

(3) 对 n 个方程和 m 个约束求解 $x_1, \cdots, x_n$ 和 $\lambda_1, \cdots, \lambda_m$.

(14.29)解的必要条件,其中 f 和 $g_1, \cdots, g_m$ 是在 $\mathbf{R}^n$ 内的开集 A 上的 C^1 函数,且 $\boldsymbol{x} = (x_1, \cdots, x_n)$. 假设雅各比矩阵 $(\partial g_j/\partial x_i)_{m\times n}$ 的秩等于 m,(参见(6.8).) $\lambda_1, \cdots, \lambda_m$ 称为拉格朗日乘数.

14.31 如果 $\boldsymbol{x}^0$ 是问题(14.29)的一个解,且梯度 $\nabla g_1(\boldsymbol{x}^0), \cdots, \nabla g_m(\boldsymbol{x}^0)$是线性无关的,则存在唯一的数 $\lambda_1, \cdots, \lambda_m$,使

$$\nabla f(\boldsymbol{x}^0) = \lambda_1 \nabla g_1(\boldsymbol{x}^0) + \cdots + \lambda_m \nabla g_m(\boldsymbol{x}^0)$$

(14.30)的另一种形式.

14.32 假设(14.29)中的 $f(\boldsymbol{x})$和 $g_1(\boldsymbol{x}),\cdots,g_m(\boldsymbol{x})$ 定义在 $\mathbf{R}^n$ 中的开凸集 S 上. 设 $\boldsymbol{x}^0 \in S$ 为拉格朗日函数的一个驻点,且假设 $g_j(\boldsymbol{x}^0) = b_j$, $j = 1, \cdots, m$, 则

$L(\boldsymbol{x})$凹的$\Rightarrow \boldsymbol{x}^0$ 解决了问题(14.29).

(14.29)解的充分条件(对于最小化问题,将"$L(\boldsymbol{x})$凹的"改为"$L(\boldsymbol{x})$凸的".)

14.33

$$B_r = \begin{vmatrix} 0 & \cdots & 0 & \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_r} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_r} \\ \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_1} & F''_{11} & \cdots & F''_{1r} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial x_r} & \cdots & \frac{\partial g_m}{\partial x_r} & F''_{r1} & \cdots & F''_{rr} \end{vmatrix}$$

问题(14.29)的有界 *Hessian* 行列式, $r = 1, \cdots, n$. L 是定义在(14.30)中的拉格朗日函数.

14.34 设 $f(\boldsymbol{x})$和 $g_1(\boldsymbol{x}), \cdots, g_m(\boldsymbol{x})$为 $\mathbf{R}^n$ 中开集S上的 C^2 函数, $\boldsymbol{x}^0 \in S$ 满足问题(14.29)中由(14.30)给出的必要条件,再设 $B_r(\boldsymbol{x}^0)$为(14.33)中在 $\boldsymbol{x}^0$ 点的行列式,则

- 如果 $(-1)^m B_r(\boldsymbol{x}^0) > 0$ 对于 $r = m+1, \cdots, n$ 成立,则 $\boldsymbol{x}^0$ 是(14.29)中的局部最小值点.
- 如果 $(-1)^r B_r(\boldsymbol{x}^0) > 0$ 对于 $r = m+1, \cdots, n$ 成立,则 $\boldsymbol{x}^0$ 是(14.29)中的局部最大值点.

拉格朗日问题的局部充分条件.

值函数及敏感性

14.35 $f^*(\boldsymbol{b}) = \max\limits_{\boldsymbol{x}}\{f(\boldsymbol{x}) : g_j(\boldsymbol{x}) = b_j,\ j = 1, \cdots, m\}$

(14.29)的值函数,其中 $\boldsymbol{b} = (b_1, \cdots, b_m)$,假设最大值存在.

14.36	拉格朗日乘数是"影子价格" $$\frac{\partial f^*(\boldsymbol{b})}{\partial b_i}=\lambda_i(\boldsymbol{b}),\ i=1,\cdots,m$$	$\lambda_1(\boldsymbol{b}),\cdots,\lambda_m(\boldsymbol{b})$ 是(14.30)的唯一的拉格朗日乘数. 可微性的充分条件可从(14.42)中推出.
14.37	$V(\boldsymbol{a})=\max\limits_{x\in X} f(\boldsymbol{x},\boldsymbol{a})$	最大化问题的值函数,假设最大值存在, $X\subset\mathbf{R}^n$, $\boldsymbol{a}\in A$, $A\subset\mathbf{R}^k$.
14.38	如果 $f(\boldsymbol{x},\boldsymbol{a})$在 $X\times A$ 上是连续的且 X 是紧的,则定义在(14.37)的 $V(\boldsymbol{a})$在 A 上是连续的.如果问题(14.37)对于每一 $\boldsymbol{a}\in A$ 有唯一解 $\boldsymbol{x}=\boldsymbol{x}(\boldsymbol{a})$, 则 $\boldsymbol{x}(\boldsymbol{a})$是 $\boldsymbol{a}$ 的一个连续函数.	值函数及其最大化解的连续性.
14.39	假设对 $\boldsymbol{x}\in X$ 求$f(\boldsymbol{x},\boldsymbol{a})$的最大值问题,在 $\boldsymbol{a}=\boldsymbol{a}^0$ 有唯一解 $\boldsymbol{x}^0(\boldsymbol{a}^0)$,且$\partial f/\partial a_i$, ($i=1,\cdots,k$) 存在并在($\boldsymbol{x}^0$, $\boldsymbol{a}^0$)的一个邻域内连续,则对于 $i=1,\cdots,k$, $$\frac{\partial V(\boldsymbol{a}^0)}{\partial a_i}=\frac{\partial f(\boldsymbol{x}^0,\boldsymbol{a}^0)}{\partial a_i}$$	包络定理.
14.40	$\max\limits_{x} f(\boldsymbol{x},\boldsymbol{a})$s.t. $g_j(\boldsymbol{x},\boldsymbol{a})=0,\ j=1,\cdots,m$	有参数 $\boldsymbol{a}=(a_1,\cdots,a_k)$ 的拉格朗日问题.
14.41	$V(\boldsymbol{a})=\sup\{f(\boldsymbol{x},\boldsymbol{a}):g_j(\boldsymbol{x},\boldsymbol{a})=0,\ j=1,\cdots,m\}$	问题(14.40)的值函数.

14.42 考虑问题(14.40),假设

- 对于 $\boldsymbol{a} = \boldsymbol{a}^0$ 问题有唯一解 $\boldsymbol{x}^0 = \mathbf{x}(\boldsymbol{a}^0)$;
- 存在一个球 $B(\boldsymbol{a}^0;\ \alpha)$和一个常数 K,使对于每一 $\boldsymbol{a} \in B(\boldsymbol{a}^0;\ \alpha)$ 存在(14.40)的解 $\boldsymbol{x}'$,使 $\boldsymbol{x}' \in B(\boldsymbol{x}^0;\ K)$;
- f 和 $g_1,\ \cdots,\ g_m$ 在$(\boldsymbol{x}(\boldsymbol{a}^0),\ \boldsymbol{a}^0)$的某一邻域内是 C^1 函数;
- 矩阵$(\partial g_j(\boldsymbol{x}^0)/\partial x_i)_{m\times n}$的秩为 m.

则 $V(\boldsymbol{a})$在 $\boldsymbol{a}^0$ 有偏导数,且

$$\frac{\partial V(\boldsymbol{a}^0)}{\partial a_i} = \frac{\partial L(\boldsymbol{x}(\boldsymbol{a}^0),\ \boldsymbol{a}^0)}{\partial a_i},\ i = 1,\ \cdots,\ k$$

问题(14.40)的一个包络定理.
$L = f - \sum \lambda_j g_j$ 是拉格朗日函数.

参 考 文 献

参考 Simon & Blume (1994), Sydsaeter & Hammond (1995), Intriligator (1971), Luenberger (1973)及 Dixit (1990).

15 线性与非线性规划

线性规划

15.1 $\max z = c_1x_1 + \cdots + c_nx_n$, 约束条件

$$
\begin{aligned}
&a_{11}x_1 + \cdots + a_{1n}x_n \leqslant b_1\\
&a_{21}x_1 + \cdots + a_{2n}x_n \leqslant b_2\\
&\cdots\cdots\\
&a_{m1}x_1 + \cdots + a_{mn}x_n \leqslant b_m\\
&x_1 \geqslant 0, \cdots, x_n \geqslant 0
\end{aligned}
$$

线性规划问题(原型问题). $\sum_{j=1}^{n} c_jx_j$ 称为准则函数. $(x_1, \cdots, x_n)$称可接受的, 如果它满足所有 $m+n$ 个约束.

15.2 $\min Z = b_1\lambda_1 + \cdots + b_m\lambda_m$, 约束条件

$$
\begin{aligned}
&a_{11}\lambda_1 + \cdots + a_{m1}\lambda_m \geqslant c_1\\
&a_{12}\lambda_1 + \cdots + a_{m2}\lambda_m \geqslant c_2\\
&\cdots\cdots\\
&a_{1n}\lambda_1 + \cdots + a_{mn}\lambda_m \geqslant c_n\\
&\lambda_1 \geqslant 0, \cdots, \lambda_m \geqslant 0
\end{aligned}
$$

(15.1)的对偶. $\sum_{i=1}^{m} b_i\lambda_i$ 称为准则函数. $(\lambda_1, \cdots, \lambda_m)$称可接受的, 如果它满足所有 $m+n$ 个约束.

15.3 $\max \boldsymbol{c}'\boldsymbol{x}$ 受约束 $\mathbf{A}\boldsymbol{x} \leqslant \boldsymbol{b}$, $\boldsymbol{x} \geqslant 0$

$\min \boldsymbol{b}'\boldsymbol{\lambda}$ 受约束 $\mathbf{A}'\boldsymbol{\lambda} \geqslant \boldsymbol{c}$, $\boldsymbol{\lambda} \geqslant 0$

(15.1)和(15.2)的矩阵形式. $\mathbf{A} = (a_{ij})_{m\times n}$, $\boldsymbol{x} = (x_j)_{n\times 1}$, $\boldsymbol{\lambda} = (\lambda_i)_{m\times 1}$, $\boldsymbol{c} = (c_j)_{n\times 1}$, $\boldsymbol{b} = (b_i)_{m\times 1}$.

15.4 如果在(15.1)和(15.2)中$(x_1, \cdots, x_n)$和$(\lambda_1, \cdots, \lambda_m)$分别是可接受的,则

$$b_1\lambda_1 + \cdots + b_m\lambda_m \geqslant c_1x_1 + \cdots + c_nx_n$$

对偶中的准则函数的函数值总大于或等于原型中的准则函数的函数值.

15.5 假设$(x_1^*, \cdots, x_n^*)$和$(\lambda_1^*, \cdots, \lambda_m^*)$分别是在(15.1)和(15.2)可接受的,且

$$c_1x_1^* + \cdots + c_nx_n^* = b_1\lambda_1^* + \cdots + b_m\lambda_m^*$$

则$(x_1^*, \cdots, x_n^*)$和$(\lambda_1^*, \cdots, \lambda_m^*)$分别是各自问题的最优解.

一个有趣的结论.

15.6 如果问题(15.1)和(15.2)中的一个有有限个数的最优解,则另一个也有有限个数的最优解,且对应的准则函数值相等.如果一个问题有“非限制最优解”,则另一问题无可接受解.

线性规划的对偶性定理.

15.7 考虑问题(15.1).如果我们将b_i改为$b_i + \Delta b_i$, $i = 1, \cdots, m$,且如果对应的对偶问题有相同的最优解$(\lambda_1^*, \cdots, \lambda_m^*)$,则原型问题的准则函数值的改变为

$$\Delta z^* = \lambda_1^*\Delta b_1 + \cdots + \lambda_m^*\Delta b_m$$

一个重要的敏感性结论.(对偶问题通常会有相同的解,如果$|\Delta b_1|, \cdots, |\Delta b_m|$足够小.)

15.8 第i个对偶变量λ_i^*等于当b_i增加一个单位时,原型问题(15.1)的准则函数值的变化.

λ_i^*可诠释为“影子价格”.(相同条件下的(15.7)的一个特例.)

15.9 假设原型问题(15.1)有最优解$(x_1^*, \cdots, x_n^*)$,而对偶问题(15.2)有最优解$(\lambda_1^*, \cdots, \lambda_m^*)$,则对于$i = 1, \cdots, n$, $j = 1, \cdots, m$:

(1) $x_j^* > 0 \Rightarrow a_{1j}\lambda_1^* + \cdots + a_{mj}\lambda_m^* = c_j$

(2) $\lambda_i^* > 0 \Rightarrow a_{i1}x_1^* + \cdots + a_{in}x_n^* = b_i$

互补宽松性((1)如果最优变量j在原型中是正的,则对偶中对j的约束在最优时是等式,(2)有类似的解释)

15.10 设$\mathbf{A}$为一$m \times n$矩阵且$\boldsymbol{b}$为一n阶向量,则当且仅当不存在$\boldsymbol{x} \geqslant \mathbf{0}$使$\mathbf{A}'\boldsymbol{x} = \boldsymbol{b}$时,存在一向量$\boldsymbol{y}$满足$\mathbf{A}\boldsymbol{y} \geqslant \mathbf{0}$和$\boldsymbol{b}'\boldsymbol{y} < 0$.

Farkas 引理.

非线性规划

15.11 $\max f(x, y)$受约束 $g(x, y) \leqslant b$

非线性规划问题.

15.12 解决问题(15.11)的办法:

(1) 定义拉格朗日函数 L 为

$L(x, y, \lambda) = f(x, y) - \lambda(g(x, y) - b)$

其中 λ 是一与约束 $g(x, y) \leqslant b$ 相联系的拉格朗日乘数.

(2) 令 $L(x, y, \lambda)$对 x 和 y 的偏导数为零:

$L_1'(x, y, \lambda) = f_1'(x, y) - \lambda g_1'(x, y) = 0$

$L_2'(x, y, \lambda) = f_2'(x, y) - \lambda g_2'(x, y) = 0$

(3) 引入互补宽松性条件

$\lambda \geqslant 0 (= 0$ 如果 $g(x, y) < b)$

(4) 寻找满足 $g(x, y) \leqslant b$ 的(x, y).

解问题(15.11)的必要条件,更精确的表达在(15.20)中.如果我们找到所有满足这些条件的(x,y)(以及适当的λ),我们就有了所有解决问题的候选点.如果拉格朗日函数对(x, y)是凹的,则这些条件对最优解是充分的.

15.13 $$\max_{\mathbf{x}} f(\mathbf{x}) \text{ s.t.} \begin{cases} g_1(\mathbf{x}) \leqslant b_1 \\ \cdots\cdots\cdots\cdots \\ g_m(\mathbf{x}) \leqslant b_m \end{cases}$$

非线性规划问题.向量 $\mathbf{x} = (x_1, \cdots, x_n)$是可接受的,如果它满足所有约束.

15.14 $$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{j=1}^{m} \lambda_j (g_j(\mathbf{x}) - b_j)$$

对应(15.13)的拉格朗日函数. $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_m)$是拉格朗日乘数.

15.15 考虑问题(15.13),假设 f 和$g_1, \cdots, g_m$ 是C^1函数.假设存在一向量 $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_m)$ 和一可接受向量 $\mathbf{x}^0 = (x_1^0, \cdots, x_n^0)$,使

(a) $\dfrac{\partial L(\mathbf{x}^0, \boldsymbol{\lambda})}{\partial x_i} = 0, \; i = 1, \cdots, n$

(b) 对于所有 $j = 1, \cdots, m$,

$\lambda_j \geqslant 0 \; (=0$ 如果 $g_j(\mathbf{x}^0) < b_j)$

(c) 拉格朗日函数 $L(\mathbf{x}, \boldsymbol{\lambda})$是 $\mathbf{x}$ 的一个凹函数,则 $\mathbf{x}^0$ 是问题(15.13)的解.

Kuhn-Tucker 充分条件.

15.16	(b′) $\lambda_j \geq 0$ 且 $\lambda_j(g_j(\boldsymbol{x}^0)-b_j)=0$, $j=1$, $\cdots$, m	(15.15)中(b)的另一种表达.
15.17	如果我们将(c)改为 (c′) $f(\boldsymbol{x})$是凹的且 $\lambda_j g_j(\boldsymbol{x})$是拟凸的, $j=1$, $\cdots$, m (15.15)也将成立.	另一种 Kuhn-Tucker 充分条件.
15.18	约束 j 在(15.13)中称为在 $\boldsymbol{x}^0$ 有效的,如果 $g_j(\boldsymbol{x}^0)=b_j$.	有效(或有约束力的)约束的定义.
15.19	以下条件经常附加在问题(15.13)里:函数 g_j ($j=1$, $\cdots$, m)在 $\boldsymbol{x}^0$ 点的梯度,当约束条件在 $\boldsymbol{x}^0$ 是有效的时候,是线性无关的.	问题(15.13)的约束限制.
15.20	假设 $\boldsymbol{x}^0=(x_1^0, \cdots, x_n^0)$是(15.13)的解,且 f 和 $g_1, \cdots, g_m$ 是 C^1 函数.进一步假设约束限制(15.19)在 $\boldsymbol{x}^0$ 点满足,则存在唯一的 λ_1, $\cdots$, λ_m,使 (a) $\dfrac{\partial L(\boldsymbol{x}^0, \boldsymbol{\lambda})}{\partial x_i}=0$, $i=1$, $\cdots$, n (b) $\lambda_j \geq 0$ ($=0$ 如果 $g_j(\boldsymbol{x}^0)<b_j$), 对于所有 $j=1$, $\cdots$, m 成立.	问题(15.13)的 Kuhn-Tucker 必要条件. (注意所有约束限制不满足的可接受点都是最优解的候选者)
15.21	$(\boldsymbol{x}^0, \boldsymbol{\lambda}^0)$是拉格朗日函数 $L(\boldsymbol{x}, \boldsymbol{\lambda})$的一个鞍点,如果 $L(\boldsymbol{x}, \boldsymbol{\lambda}^0) \leq L(\boldsymbol{x}^0, \boldsymbol{\lambda}^0) \leq L(\boldsymbol{x}^0, \boldsymbol{\lambda})$ 对于所有 $\boldsymbol{\lambda} \geq 0$ 和所有 $\boldsymbol{x}$ 成立.	问题(15.13)的鞍点的定义.
15.22	如果 $L(\boldsymbol{x}, \boldsymbol{\lambda})$有一鞍点$(\boldsymbol{x}^0, \boldsymbol{\lambda}^0)$,则$(\boldsymbol{x}^0, \boldsymbol{\lambda}^0)$是(15.13)的解.	问题(15.13)的充分条件(不要求可微性和凹性条件)
15.23	以下条件经常附加在问题(15.13)上:对某一向量 $\boldsymbol{x}'=(x_1', \cdots, x_n')$, $g_j(\boldsymbol{x}')<b_j$ 对 $j=1$, $\cdots$, m.	*Slater* 条件(限制).

15.24	考虑问题(15.13),假设 f 是凹的而 $g_1, \cdots, g_m$ 是凸的.假设 Slater 条件(15.23)得以满足,则使 $\boldsymbol{x}^0$ 为问题的解的一个充分必要条件是存在非负数 $\lambda_1^0, \cdots, \lambda_m^0$,使$(\boldsymbol{x}^0, \boldsymbol{\lambda}^0)$成为拉格朗日函数 $L(\boldsymbol{x}, \boldsymbol{\lambda})$的一个鞍点.	凹性规划中的鞍点结论.
15.25	考虑问题(15.13),假设 f 和 $g_1, \cdots, g_m$ 是 C^1 函数.假设存在数 $\lambda_1, \cdots, \lambda_m$ 和一个向量 $\boldsymbol{x}^0$,使 • $\boldsymbol{x}^0$ 满足(15.15)中的(a)和(b), • $\nabla f(\boldsymbol{x}^0) \neq \boldsymbol{0}$, • $f(\boldsymbol{x})$是拟凹的且 $\lambda_j g_j(\boldsymbol{x})$是拟凸的,对于 $j = 1, \cdots, m$, 则 $\boldsymbol{x}^0$ 是问题(15.13)的解.	拟凹规划的充分条件.
15.26	$V(\boldsymbol{b}) = \max\{f(\boldsymbol{x}): g_j(\boldsymbol{x}) \leqslant b_j, j = 1, \cdots, m\}$	(15.13)的值函数,假设最大值存在,且 $\boldsymbol{b} = (b_1, \cdots, b_m) \in \mathbf{R}^m$.
15.27	(1) $V(\boldsymbol{b})$对每一变量是非递减的. (2) $\partial V(\boldsymbol{b})/\partial b_j = \lambda_j(\boldsymbol{b}), j = 1, \cdots, m$ (3) 如果 $f(\boldsymbol{x})$是凹的且 $g_1(\boldsymbol{x}), \cdots, g_m(\boldsymbol{x})$是凸的,则 $V(\boldsymbol{b})$是凹的.	值函数的性质.((2)的一个精确形式可从(15.30)得到.)
15.28	$\max\limits_{\boldsymbol{x}} f(\boldsymbol{x}, \boldsymbol{a})$ s.t. $g_j(\boldsymbol{x}, \boldsymbol{a}) \leqslant 0, j = 1, \cdots, m$	一个有参数 $\boldsymbol{a} \in \mathbf{R}^k$ 的非线性规划问题.
15.29	$V(\boldsymbol{a}) = \sup\{f(\boldsymbol{x}, \boldsymbol{a}): g_j(\boldsymbol{x}, \boldsymbol{a}) \leqslant 0, j = 1, \cdots, m\}$	问题(15.28)的值函数.

15.30 考虑问题(15.28)且假设

- 对于 $\boldsymbol{a}=\boldsymbol{a}^0$，问题有一唯一解 $\boldsymbol{x}^0$，
- 存在一球 $B(\boldsymbol{a}^0;\ \alpha)$和一常数 K，使对于每一在 $B(\boldsymbol{a}^0;\ \alpha)$内的 $\boldsymbol{a}$ 存在(15.28)的解 $\boldsymbol{x}'$，且 $\boldsymbol{x}'\in B(\boldsymbol{x}^0;\ K)$，
- f 和 $g_1,\ \cdots,\ g_m$ 是在$(\boldsymbol{x}^0,\ \boldsymbol{a}^0)$附近的某一球内的 C^1 函数，
- 对应于在 $\boldsymbol{x}^0$ 点有效的约束 j 的函数 g_j 的在 $\boldsymbol{x}^0$ 的梯度是线性无关的，

则 $V(\boldsymbol{a})$在 $\boldsymbol{a}^0$ 有偏导数，且

$$\frac{\partial V(\boldsymbol{a}^0)}{\partial a_i}=\frac{\partial \mathcal{L}(\boldsymbol{x}^0,\ \boldsymbol{a}^0,\ \boldsymbol{\lambda})}{\partial a_i},\ i=1,\ \cdots,\ k$$

问题(15.28)的包络定理. $\mathcal{L}=f-\sum\lambda_j g_j$ 是拉格朗日函数，约束 j 在$(\boldsymbol{x}^0,\ \boldsymbol{a}^0)$是有效的，如果 $g_j(\boldsymbol{x}^0,\ \boldsymbol{a}^0)=0$.

非负条件下的非线性规划

15.31 $\max\limits_{\boldsymbol{x}} f(\boldsymbol{x})$，受约束 $\begin{cases} g_1(\boldsymbol{x})\leqslant b_1 \\ \cdots\cdots\cdots\cdots,\ \boldsymbol{x}\geqslant 0 \\ g_m(\boldsymbol{x})\leqslant b_m \end{cases}$

如果将非负条件写为 $g_{m+i}(\boldsymbol{x})=-x_i\leqslant 0$. $i=1,\ \cdots,\ n$，(15.31)将简化为(15.13).

15.32 假设在问题(15.31)中，f 和 $g_1,\ \cdots,\ g_m$ 是 C^1 函数，且存在数 $\lambda_1,\ \cdots,\ \lambda_m$ 和一个可接受的向量 $\boldsymbol{x}^0$，使

(a) 对于所有 $i=1,\ \cdots,\ n$，$x_i^0\geqslant 0$ 且

$$\frac{\partial L(\boldsymbol{x}^0,\ \boldsymbol{\lambda})}{\partial x_i}\leqslant 0,\ x_i^0\frac{L(\boldsymbol{x}^0,\ \boldsymbol{\lambda})}{\partial x_i}=0$$

(b) 对于所有 $j=1,\ \cdots,\ m$，

$\lambda_j\geqslant 0$ ($=0$ 如果 $g_j(\boldsymbol{x}^0)<b_j$)

(c) 拉格朗日函数 $L(\boldsymbol{x},\ \boldsymbol{\lambda})$是 $\boldsymbol{x}$ 的一个凹函数，

则 $\boldsymbol{x}^0$ 是问题的解(15.31).

问题(15.31)的 Kuhn-Tucker 充分条件.

$\boldsymbol{\lambda}=(\lambda_i)_{m\times 1}$. $L(\boldsymbol{x}^0,\ \boldsymbol{\lambda})$的定义在(15.14).

15.33 在(15.32)中,(c)可改为

(c′) $f(\boldsymbol{x})$ 是凹的且 $\lambda_j g_j(\boldsymbol{x})$ 对于 $j = 1, \cdots, m$ 是拟凸的.

另一种 Kuhn-Tucker 充分条件.

15.34 假设 $\boldsymbol{x}^0 = (x_1^0, \cdots, x_n^0)$ 是(15.31)的解,其中 f 和 $g_1, \cdots, g_m$ 是 C^1 函数.再假设对应于在 $\boldsymbol{x}^0$ 有效约束的 g_j(包括定义在(15.31)的注解内的函数 $g_{m+1}, \cdots, g_{m+n}$)在 $\boldsymbol{x}^0$ 的梯度是线性无关的,则存在唯一的数 $\lambda_1, \cdots, \lambda_m$,使

(a) 对于所有 $i = 1, \cdots, n$, $x_i^0 \geqslant 0$, 且

$$\frac{\partial L(\boldsymbol{x}^0, \boldsymbol{\lambda})}{\partial x_i} \leqslant 0, \quad x_i^0 \frac{L(\boldsymbol{x}^0, \boldsymbol{\lambda})}{\partial x_i} = 0$$

(b) $\lambda_j \geqslant 0$ ($=0$ 如果 $g_j(\boldsymbol{x}^0) < b_j$), $j = 1, \cdots, m$

问题(15.31)的 Kuhn-Tucker 必要条件.

(注意所有约束限制不满足的可接受点都是最优解的候选者)

参 考 文 献

Gass (1994), Luenberger (1973), Intriligator (1971), Sydsaeter & Hammond (1995), Simon & Blume (1994), Beavis & Dobbs (1990), Dixit (1990).(后三本不包括线性规划.)

16 变分学和最优控制理论

变分

16.1 变分中最简单的问题(t_0, t_1, x^0,和 x^1 是固定的数):

$$\max\int_{t_0}^{t_1} F(t, x, \dot{x})\mathrm{d}t, \quad x(t_0)=x^0, \quad x(t_1)=x^1$$

F 是一 C^2 函数. 未知的 $x = x(t)$ 是可接受的,如果它是 C^1 函数且满足两个边界条件. 对于最小化问题,将 F 改为 $-F$.

16.2 $$\frac{\partial F}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial F}{\partial \dot{x}}\right) = 0$$

欧拉方程. (16.1)有解的一个必要条件.

16.3 $$\frac{\partial^2 F}{\partial \dot{x}\,\partial \dot{x}}\,\ddot{x} + \frac{\partial^2 F}{\partial x \partial \dot{x}}\,\dot{x} + \frac{\partial^2 F}{\partial t \partial \dot{x}} - \frac{\partial F}{\partial x} = 0$$

欧拉方程的另一种形式.

16.4 $F''_{\dot{x}\dot{x}}(t, x(t), \dot{x}(t)) \leqslant 0$ 对于所有 t 在 $[t_0, t_1]$内.

Legendre 条件. (16.1)有解的一个必要条件.

16.5 如果 $F(t, x, \dot{x})$在$(x, \dot{x})$是凹的,一个满足欧拉方程的可接受的函数 $x = x(t)$, 是问题(16.1)的解.

(16.1)有解的充分条件.

16.6 (16.1)中 $x(t_1)$不受限制 $\Rightarrow \left[\frac{\partial F}{\partial \dot{x}}\right]_{t=t_1} = 0$

横截条件. 加上(16.5)给出充分条件.

16.7	(16.1)中 $x(t_1) \geqslant x^1 \Rightarrow$ $\left[\frac{\partial F}{\partial \dot{x}}\right]_{t=t_1} \leqslant 0$(= 0 如果 $x(t_1) > x^1$)	横截条件，加上(16.5)给出充分条件.
16.8	(16.1)中 t_1 不受限制$\Rightarrow \left[F - \dot{x}\frac{\partial F}{\partial \dot{x}}\right]_{t=t_1} = 0$	横截条件.
16.9	在(16.1)中 $x(t_1) = g(t_1)$ $\Rightarrow \left[F + (\dot{g} - \dot{x})\frac{\partial F}{\partial \dot{x}}\right]_{t=t_1} = 0$	横截条件. g 是一给定的 C^1 函数.
16.10	$\max\left[\int_{t_0}^{t_1} F(t, x, \dot{x})\mathrm{d}t + S(x(t_1))\right]$, $x(t_0) = x^0$	有附加值函数 S 的变分问题，S 是一 C^1 函数.
16.11	$\left[\frac{\partial F}{\partial \dot{x}}\right]_{t=t_1} + S'(x(t_1)) = 0$	(16.10)的解必须满足(16.2)和这一横截条件.
16.12	如果 $F(t, x, \dot{x})$在$(x, \dot{x})$是凹的且 $S(x)$是凹的，则一个满足(16.11)中欧拉方程的可接受函数是问题(16.10)的解.	(16.10)解的充分条件.
16.13	$\max\int_{t_0}^{t_1} F\left(t, x, \frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}^2x}{\mathrm{d}t^2}, \cdots, \frac{\mathrm{d}^nx}{\mathrm{d}t^n}\right)\mathrm{d}t$	有高阶偏导数的变分问题(边界条件未设定).
16.14	$\frac{\partial F}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial F}{\partial \dot{x}}\right) + \cdots + (-1)^n \frac{\mathrm{d}^n}{\mathrm{d}t^n}\left(\frac{\partial F}{\partial x^{(n)}}\right) = 0$	(16.13)的(一般化)欧拉方程.
16.15	$\max\iint_R F\left(t, s, x, \frac{\partial x}{\partial t}, \frac{\partial x}{\partial s}\right)\mathrm{d}t\mathrm{d}s$	未知的 $x(t, s)$是两个变量函数的变分问题(边界条件未设定).
16.16	$\frac{\partial F}{\partial x} - \frac{\partial}{\partial t}\left(\frac{\partial F}{\partial x'_t}\right) - \frac{\partial}{\partial s}\left(\frac{\partial F}{\partial x'_s}\right) = 0$	(16.15)的(一般化)欧拉方程.

最优控制理论 一种状态和一个变量

16.17 最简单的情况. 固定的时间区间$[t_0, t_1]$和自由的上限：

$$\max\int_{t_0}^{t_1} f(t, x(t), u(t))\mathrm{d}t, \ u(t)\in \mathrm{R}$$

$\dot{x}(t) = g(t, x(t), u(t)), \ x(t_0) = x^0$, $x(t_1)$不受限制.

$(x(t), u(t))$是可接受的，如果它满足微分方程$x(t_0) = x^0$，且$u(t)$是分段连续的. 要解决最小化问题，将f改为$-f$.

16.18 $H(t, x, u, p) = f(t, x, u) + pg(t, x, u)$

与(16.17)相伴的Hamilton函数.

16.19 假设$(x^*(t), u^*(t))$是问题(16.17)的解，则存在一连续函数$p(t)$，使对于所有$t\in[t_0, t_1]$的解：

(1) $H(t, x^*(t), u, p(t)) \leqslant H(t, x^*(t), u^*(t), p(t))$对于所有$u\in\mathrm{R}$成立. 特别地，

$H'_u(t, x^*(t), u^*(t), p(t)) = 0$

(2) 函数$p(t)$满足

$\dot{p}(t) = -H'_x(t, x^*(t), u^*(t), p(t))$,

$p(t_1) = 0$

最大化原理. $p(t)$的微分方程在$u^*(t)$的不连续点不一定成立. 方程$p(t_1) = 0$称为横截条件.

16.20 如果$(x^*(t), u^*(t))$满足(16.19)中的条件，且$H(t, x, u, p(t))$在(x, u)是凹的，则$(x^*(t), u^*(t))$是问题(16.17)的解.

问题(16.17)的*Mangasarian*充分条件.

16.21 $\max\int_{t_0}^{t_1} f(t, x(t), u(t))\mathrm{d}t, \ u(t)\in U\subset\mathrm{R}$

$\dot{x}(t) = g(t, x(t), u(t)), \ x(t_0) = x^0$

(a) $x(t_1) = x^1$或(b) $x(t_1)\geqslant x^1$

有终端条件和固定时间区间的控制问题. U是控制区域. $u(t)$是分段连续的.

16.22 $H(t, x, u, p) = p_0 f(t, x, u) + pg(t, x, u)$

与(16.21)相伴的Hamilton函数.

16.23 假设$(x^*(t), u^*(t))$是问题(16.21)的解，则存在一连续函数$p(t)$和一个数p_0，使对于所有$t \in [t_0, t_1]$：

(1) $p_0 = 0$或1且$(p_0, p(t))$不为$(0, 0)$

(2) $H(t, x^*(t), u, p(t)) \leqslant H(t, x^*(t), u^*(t), p(t))$对所有$u \in U$成立.

(3) $\dot{p}(t) = -H'_x(t, x^*(t), u^*(t), p(t))$

(4) (a′) 对$p(t_1)$没有限制.

(b′) $p(t_1) \geqslant 0$ ($=0$ 如果 $x^*(t_1) > x^1$)

最大化原理. $p(t)$的微分方程在$u^*(t)$的不连续点不一定成立. (b′)称为横截条件. (除了在退化的情况下，我们可设定$p_0 = 1$且忽略(1).)

多种状态和多个变量

16.24 $\max \int_{t_0}^{t_1} f(t, \mathbf{x}(t), \mathbf{u}(t))\mathrm{d}t$

$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t)),\ \mathbf{x}(t_0) = \mathbf{x}^0$

$\mathbf{u}(t) = (u_1(t), \cdots, u_r(t)) \in U \subset \mathbf{R}^r$

(a) $x_i(t_1) = x_i^1, \quad i = 1, \cdots, l$

(b) $x_i(t_1) \geqslant x_i^1, \quad i = l+1, \cdots, q$

(c) $x_i(t_1)$不受限制，$i = q+1, \cdots, n$

有固定时间区间的标准控制问题. U是控制区域，$\mathbf{x}(t) = (x_1(t), \cdots, x_n(t))$，$\mathbf{g} = (g_1, \cdots, g_n)$，$\mathbf{u}(t)$是分段连续的.

16.25 $H(t, \mathbf{x}, \mathbf{u}, \mathbf{p}) = p_0 f(t, \mathbf{x}, \mathbf{u}) + \sum_{i=1}^{n} p_i g_i(t, \mathbf{x}, \mathbf{u})$

Hamilton 函数.

16.26 如果$(\mathbf{x}^*(t), \mathbf{u}^*(t))$是问题(16.24)的解，则存在一常数$p_0$和一连续函数$\mathbf{p}(t) = (p_1(t), \cdots, p_n(t))$，使对于所有$t \in [t_0, t_1]$：

(1) $p_0 = 0$或1且$(p_0, \mathbf{p}(t))$不为$(0, \mathbf{0})$

(2) $H(t, \mathbf{x}^*(t), \mathbf{u}, \mathbf{p}(t)) \leqslant H(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}(t))$对所有$\mathbf{u} \in U$成立.

(3) $\dot{p}_i(t) = -\partial H^* / \partial x_i,\ i = 1, \cdots, n$

(4) (a′) $p_i(t_1)$不受限制，$\quad i = 1, \cdots, l$

(b′) $p_i(t_1) \geqslant 0$ ($= 0$ 如果 $x_i^*(t_1) > x_i^1$) $\quad i = l+1, \cdots, q$

(c′) $p_i(t_1) = 0, \quad i = q+1, \cdots, n$

最大化原理. H^*取值于$(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}(t))$点. $p_i(t)$的微分方程在$\mathbf{u}^*(t)$的不连续点不一定成立. (4)(b′)和(c′)是横截条件. (除了在退化的情况下，我们可以设定$p_0 = 1$且忽略(1).)

16.27 如果对于 $p_0=1$，$(\mathbf{x}^*(t), \mathbf{u}^*(t))$ 满足(16.26)中的所有条件，且 $H(t, \mathbf{x}, \mathbf{u}, \mathbf{p}(t))$在$(\mathbf{x}, \mathbf{u})$是凹的，则$(\mathbf{x}^*(t), \mathbf{u}^*(t))$是问题(16.24)的解.

问题(16.24)的 *Mangasarian* 充分条件.

16.28 (16.27)中的条件 $H(t, \mathbf{x}, \mathbf{u}, \mathbf{p}(t))$在$(\mathbf{x}, \mathbf{u})$是凹的，可改为较弱的条件：最大化的 *Hamilton* 函数

$$\hat{H}(t, \mathbf{x}, \mathbf{p}(t)) = \max_{\mathbf{u}\in U} H(t, \mathbf{x}, \mathbf{u}, \mathbf{p}(t))$$

在 $\mathbf{x}$ 是凹的.

Arrow 充分条件.

16.29 $$V(\mathbf{x}^0, \mathbf{x}^1, t_0, t_1) = \int_{t_0}^{t_1} f(t, \mathbf{x}^*(t), \mathbf{u}^*(t))\mathrm{d}t$$

问题(16.24)的值函数，假设解是$(\mathbf{x}^*(t), \mathbf{u}^*(t))$且 $\mathbf{x}^1=(x_1^1, \cdots, x_q^1)$.

16.30 $$\frac{\partial V}{\partial x_i^0} = p_i(t_0), \quad i=1, \cdots, n$$

$$\frac{\partial V}{\partial x_i^1} = -p_i(t_1), \quad i=1, \cdots, q$$

$$\frac{\partial V}{\partial t_0} = -H^*(t_0), \quad \frac{\partial V}{\partial t_1} = H^*(t_1)$$

值函数的性质.假设 V 是可微的，$H^*(t)=H(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}(t))$.(精确的假设可参考 Seierstad & Sydsaeter (1987)，3.5节)

16.31 如果在问题(16.24)中 t_1 是自由的，且$(\mathbf{x}^*(t), \mathbf{u}^*(t))$是在$[t_0, t_1^*]$对应的问题的解，则所有(16.26)内的条件在$[t_0, t_1^*]$上得到满足，而且

$$H(t_1^*, \mathbf{x}^*(t_1^*), \mathbf{u}^*(t_1^*), \mathbf{p}(t_1^*)) = 0$$

自由终端时间问题的必要条件.(当 t_1 自由时 Hamilton 函数在$(\mathbf{x}, \mathbf{u})$的凹性不是是优解的充分条件.参见 Seierstad & Sydsaeter (1987)，2.9节)

16.32 考虑将问题(16.24)中的终端条件(a), (b)和(c)改为

$$R_k(\boldsymbol{x}(t_1)) = 0,\ k = 1, 2, \cdots, r_1'$$

$$R_k(\boldsymbol{x}(t_1)) \geqslant 0,\ k = r_1'+1, r_1'+2, \cdots, r_1$$

其中 $R_1, \cdots, R_{r_1}$ 是 C^1 函数. 如果$(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t))$是最优解,则(16.26)中的条件将成立,除了将(4)改为:存在数 $a_1, \cdots, a_{r_1}$,使

$$p_j(t_1) = \sum_{k=1}^{r_1} a_k \frac{\partial R_k(\boldsymbol{x}^*(t_1))}{\partial x_j},\ j=1, \cdots, n$$

其中 $a_k \geqslant 0$($=0$ 如果 $R_k(\boldsymbol{x}^*(t_1)) > 0$) 对于 $k = r_1' + 1, \cdots, r_1$ 成立,且将(1)改为

$$p_0 = 0 \text{ 或 } 1,\ (p_0, a_1, \cdots, a_{r_1}) \neq (0, 0, \cdots, 0)$$

如果$\hat{H}(t, \boldsymbol{x}, \boldsymbol{p}(t))$对于 $p_0 = 1$ 在 $\boldsymbol{x}$ 是凹的,且和 $\sum_{k=1}^{r_1} a_k R_k(\boldsymbol{x})$ 在 $\boldsymbol{x}$ 是拟凹的,则$(\boldsymbol{x}^*, \boldsymbol{u}^*)$是最优解.

更一般的终端条件. $\hat{H}(t, \boldsymbol{x}, \boldsymbol{p}(t))$ 的定义在(16.28).

16.33 $$\max\left[\int_{t_0}^{t_1} f(t, \boldsymbol{x}(t), \boldsymbol{u}(t))\mathrm{e}^{-rt}\mathrm{d}t + S(t_1, \boldsymbol{x}(t_1))\mathrm{e}^{-rt_1}\right]$$

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{g}(t, \boldsymbol{x}(t), \boldsymbol{u}(t)),$$
$$\boldsymbol{x}(t_0) = \boldsymbol{x}^0,\ \boldsymbol{u}(t) \in U \in \mathbf{R}^r$$

(a) $x_i(t_1) = x_i^1,\quad i = 1, \cdots, l$

(b) $x_i(t_1) \geqslant x_i^1,\quad i = l+1, \cdots, q$

(c) $x_i(t_1)$ 不受限制, $i = q+1, \cdots, n$

有附加值函数 S 的控制问题. t_0 和 t_1 是固定的.

16.34 $$H^c(t, \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{q}) = q_0 f(t, \boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{q} \cdot \boldsymbol{g}(t, \boldsymbol{x}, \boldsymbol{u})$$

问题(16.33)的现值 Hamilton 函数.

16.35 如果$(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t))$是问题(16.33)的解，则存在一常数 q_0 和一连续函数 $\boldsymbol{q}(t) = (q_1(t), \cdots, q_n(t))$，使对于所有 t 在$[t_0, t_1]$内：

(1) $q_0 = 0$ 或 1，且$(q_0, \boldsymbol{q}(t))$不为$(0, \boldsymbol{0})$

(2) $H^c(t, \boldsymbol{x}^*(t), \boldsymbol{u}, \boldsymbol{q}(t)) \leqslant H^c(t, \boldsymbol{x}^*(t), \boldsymbol{u}^*(t), \boldsymbol{q}(t))$对于所有 $\boldsymbol{u} \in U$ 成立.

(3) $\dot{q}_i - rq_i = -\dfrac{\partial H^c(t, \boldsymbol{x}^*, \boldsymbol{u}^*, \boldsymbol{q})}{\partial x_i}$, $\quad i = 1, \cdots, n$

(4) (a′) $q_i(t_1)$不受限制 $\quad i = 1, \cdots, l$

(b′) $q_i(t_1) \geqslant q_0 \dfrac{\partial S^*(t_1, \mathbf{x}^*(t_1))}{\partial x_i}$

(= 如果 $x_i^*(t_1) > x_i^1$)，$i = l+1, \cdots, m$

(c′) $q_i(t_1) = q_0 \dfrac{\partial S^*(t_1, \mathbf{x}^*(t_1))}{\partial x_i}$, $\quad i = m+1, \cdots, n$

问题(16.33)的最大化原理；现值公式. $q_i = q_i(t)$ 的微分方程在 $\boldsymbol{u}^*(t)$ 的不连续点不一定成立(除了在退化的情况下，我们可以设定 $q_0 = 1$ 且忽略(1)).

16.36 如果$(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t))$对于 $q_0 = 1$ 满足(16.35)中的条件，如果 $H^c(t, \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{q}(t))$在$(\boldsymbol{x}, \boldsymbol{u})$是凹的，且如果 $S(t, \boldsymbol{x})$在 $\boldsymbol{x}$ 是凹的，则$(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t))$是问题的解.

(16.33)解的充分条件.(Mangasarian)

16.37 (16.36)中的条件 $H^c(t, \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{q}(t))$在$(\boldsymbol{x}, \boldsymbol{u})$是凹的可改为较弱的条件：最大化的现值 *Hamilton* 函数

$$\hat{H}^c(t, \boldsymbol{x}, \boldsymbol{q}(t)) = \max_{\boldsymbol{u} \in U} H^c(t, \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{q}(t))$$

在 $\boldsymbol{x}$ 是凹的.

Arrow 充分条件.

16.38 如果 t_1 在(16.33)中是自由的，且如果$(\boldsymbol{x}^*, \boldsymbol{u}^*)$在$[t_0, t_1^*]$上是对应的问题的解，则(16.35)中的所有条件在$[t_0, t_1^*]$上得到满足，而且

$$H^c(t_1^*, \boldsymbol{x}^*(t_1^*), \boldsymbol{u}^*(t_1^*), \boldsymbol{q}(t_1^*)) = q_0 rS(t_1^*, \boldsymbol{x}^*(t_1^*)) - q_0 \frac{\partial S(t_1^*, \boldsymbol{x}^*(t_1^*))}{\partial t_1}$$

当 t_1 不受限制时(16.33)的必要条件(除了在退化的情况下，我们可以设定 $q_0 = 1$).

线性二次型问题

16.39 $\min\left[\int_{t_0}^{t_1}(x'\mathbf{A}x + u'\mathbf{B}u)\mathrm{d}t + (x(t_1))'\mathbf{S}x(t_1)\right]$

$\dot{x} = \mathbf{F}x + \mathbf{G}u$, $x(t_0) = x^0$, $u \in \mathbf{R}^r$

矩阵 $\mathbf{A} = \mathbf{A}(t)_{n\times n}$ 和 $\mathbf{S}_{n\times n}$ 是对称的和半正定性的，$\mathbf{B} = \mathbf{B}(t)_{r\times r}$ 是对称的和正定性的. $\mathbf{F} = \mathbf{F}(t)_{n\times n}$ 且 $\mathbf{G} = \mathbf{G}(t)_{n\times r}$.

线性二次型问题. $\mathbf{A}(t)$, $\mathbf{B}(t)$, $\mathbf{F}(t)$,和 $\mathbf{G}(t)$的元素是 t 的连续函数. $x = x(t)$是 $n\times 1$, $u = u(t)$ 是 $r\times 1$.

16.40 $\dot{\mathbf{R}} = -\mathbf{RF} - \mathbf{F}'\mathbf{R} + \mathbf{RGB}^{-1}\mathbf{G}'\mathbf{R} - \mathbf{A}$

与(16.39)相伴的 *Riccati* 方程.

16.41 假设$(x^*(t), u^*(t))$是问题(16.39)的可接受解，令 $u^* = -(\mathbf{B}(t))^{-1}(\mathbf{G}(t))'\mathbf{R}(t)x^*$，其中 $\mathbf{R} = \mathbf{R}(t)$ 是对称的 $n\times n$ 矩阵，其元素是满足(16.40)的 C^1 函数，且 $\mathbf{R}(t_1) = \mathbf{S}$，则$(x^*(t), u^*(t))$是问题(16.39)的解.

(16.39)的解.

无限时域

16.42 $\max\int_{t_0}^{\infty} f(t, x(t), u(t))\mathrm{e}^{-rt}\mathrm{d}t$

$\dot{x}(t) = g(t, x(t), u(t))$, $x(t_0) = x^0$, $u(t) \in U$

$\lim\limits_{t\to\infty} x(t) \geqslant x^1$ (x^1 是一固定数)

简单的一元无限时域问题，假设积分对于所有可接受解收敛.

16.43 $H^c(t, x, u, q) = q_0 f(t, x, u) + qg(t, x, u)$

问题(16.42)的现值 Hamilton 函数.

16.44 假设问题(16.42)的一对可接受解$(x^*(t), u^*(t))$,当$q_0 = 1$时对所有$t \geqslant t_0$满足以下条件:

(1) $H^c(t, x^*(t), u, q(t)) \leqslant H^c(t, x^*(t), u^*(t), q(t))$对于所有$u \in U$成立.

(2) $\dot{q}(t) - rq = -\partial H^c(t, x^*(t), u^*(t), q(t))/\partial x$

(3) $H^c(t, x, u, q(t))$在(x, u)是凹的.

(4) $\underline{\lim}_{t\to\infty}[q(t)e^{-rt}(x(t) - x^*(t))] \geqslant 0$对于所有可接受的$x(t)$成立.

则$(x^*(t), u^*(t))$是最优解.

Mangasarian 充分条件.(条件(1)和(2)对问题(16.42)是(本质上)必须的,但(4)不是)对必要条件的讨论,参见如 Seierstad & Sydsaeter (1987), 3.7 节)

16.45 $\max \int_{t_0}^{\infty} f(t, \mathbf{x}(t), \mathbf{u}(t))e^{-rt}dt$

$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t)),\ \mathbf{x}(t_0) = \mathbf{x}^0,$

$\mathbf{u}(t) \in U \subset \mathbf{R}^r$

(a) $\lim_{t\to\infty} x_i(t) = x_i^1, \quad i = 1, \cdots, l$

(b) $\underline{\lim}_{t\to\infty} x_i(t) \geqslant x_i^1, \quad i = l+1, \cdots, m$

(c) $t \to \infty$时$x_i(t)$是自由的,

$$i = m+1, \cdots, n$$

多种状态多个控制变量的无限时域问题. 对于 $\underline{\lim}$ 参见(12.40)和(12.41).

16.46 $D(t) = \int_{t_0}^{t}(f^* - f)e^{-r\tau}d\tau$,其中

$f^* = f(\tau, \mathbf{x}^*(\tau), \mathbf{u}^*(\tau)),$

$f = f(\tau, \mathbf{x}(\tau), \mathbf{u}(\tau))$

(16.47)的注解. $(\mathbf{x}^*(t), \mathbf{u}^*(t))$是最优解的候选者,$(\mathbf{x}(t), \mathbf{u}(t))$是任意可接受的对.

16.47 $(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t))$是

- 间断赶上最优(SCU-最优),如果对每一可接受对 $(\boldsymbol{x}(t), \boldsymbol{u}(t))$, $\overline{\lim\limits_{t\to\infty}} D(t) \geqslant 0$

  即对于每一 $\varepsilon > 0$ 和每一 T,存在某一 $t \geqslant T$,使 $D(t) \geqslant -\varepsilon$.
- 赶上最优(CU-最优),如果对每一可接受对 $(\boldsymbol{x}(t), \boldsymbol{u}(t))$, $\underline{\lim\limits_{t\to\infty}} D(t) \geqslant 0$

  即对于每一 $\varepsilon > 0$, 存在某一 T,使 $D(t) \geqslant -\varepsilon$ 对于所有 $t \geqslant T$ 成立.
- 超越最优(OT-最优),如果对一可接受对 $(\boldsymbol{x}(t), \boldsymbol{u}(t))$,存在一个数 T,使 $D(t) \geqslant 0$ 对于所有 $t \geqslant T$ 成立.

无限时域问题的不同最优解标准. 对于 $\underline{\lim}$ 和 $\overline{\lim}$, 参见(12.40)和(12.41). (SCU-最优性又称为弱最优性, 而OT-最优性则称为超越最优性)

16.48 OT-最优性⇒CU-最优性

⇒SCU-最优性

最优性标准之间的关系.

16.49 假设$(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t))$在问题(16.45)中是SCU-, CU-,或OT-最优,则存在一常数 q_0 和一连续函数 $\boldsymbol{q}(t) = (q_1(t), \cdots, q_n(t))$, 使对于所有 $t \geqslant t_0$:

(1) $q_0 = 0$ 或 1 且$(q_0, \boldsymbol{q}(t))$不为$(0, \boldsymbol{0})$

(2) $H^c(t, \boldsymbol{x}^*(t), \boldsymbol{u}, \boldsymbol{q}(t)) \leqslant H^c(t, \boldsymbol{x}^*(t), \boldsymbol{u}^*(t), \boldsymbol{q}(t))$ 对于所有 $\boldsymbol{u} \in U$

(3) $\dot{q}_i - rq_i = -\dfrac{\partial H^c(t, \boldsymbol{x}^*, \boldsymbol{u}^*, \boldsymbol{q})}{\partial x_i}, \quad i = 1, \cdots, n$

最大化原理. 无限时域.(没有横截条件.) $q_i(t)$的微分方程在 $\boldsymbol{u}^*(t)$的不连续点不一定成立.

16.50 在 CU-最优情况下,(16.49)中条件(2)和(3) $(q_0 = 1)$ 对最优解是充分的,如果

(1) $H^c(t, \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{q}(t))$在$(\boldsymbol{x}, \boldsymbol{u})$点是凹的

(2) $\underline{\lim\limits_{t\to\infty}} e^{-rt}\boldsymbol{q}(t) \cdot (\boldsymbol{x}(t) - \boldsymbol{x}^*(t)) \geqslant 0$ 对于所有可接受的 $\boldsymbol{x}(t)$成立.

无限时域问题的充分条件.

16.51 条件(16.50)(2)得以满足,如果对于所有可接受的 $x(t)$,以下条件得以满足:

(1) $\underline{\lim}_{t\to\infty} e^{-rt} q_i(t)(x_i^1 - x_i^*(t)) \geqslant 0,$

$i = 1, \cdots, m$

(2) 存在一常数 M,使

$|e^{-rt} q_i(t)| \leqslant M$ 对于所有 $t \geqslant t_0$, $i = 1, \cdots, m$

(3) 或者存在一个数 $t' \geq t_0$,使 $q_i(t) \geqslant 0$ 对于所有 $t \geqslant t'$ 成立,或存在一个数 P,使 $|x_i(t)| \leqslant P$ 对于所有 $t \geqslant t_0$ 成立,且 $\underline{\lim}_{t\to\infty} q_i(t) \geqslant 0$, $i = l+1, \cdots, m$

(4) 存在一个数 Q,使对于所有 $t \geqslant t_0$, $|x_i(t)| < Q$ 且 $\lim_{t\to\infty} q_i(t) = 0$, $i = m+1, \cdots, n$

(16.50)(2)成立的充分条件. 参见 Seierstad & Sydsæter (1987), 3.7 节,注解 16.

混合约束

16.52 $\max \int_{t_0}^{t_1} f(t, \mathbf{x}(t), \mathbf{u}(t)) \mathrm{d}t$

$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t)),$

$\mathbf{x}(t_0) = \mathbf{x}^0, \ \mathbf{u}(t) \in \mathbf{R}^r$

$h_k(t, \mathbf{x}(t), \mathbf{u}(t)) \geqslant 0, \ k = 1, \cdots, s$

(a) $x_i(t_1) = x_i^1,$ $\quad i = 1, \cdots, l$

(b) $x_i(t_1) \geqslant x_i^1,$ $\quad i = l+1, \cdots, q$

(c) $x_i(t_1)$ 不受限制, $\quad i = q+1, \cdots, n$

混合约束问题. $\mathbf{x}(t) \in \mathbf{R}^n$ (所有对 $\mathbf{u}$ 的约束必须包括在 h_k 约束内)

16.53 $$\mathcal{L}(t, \mathbf{x}, \mathbf{u}, \mathbf{p}, \mathbf{q}) = H(t, \mathbf{x}, \mathbf{u}, \mathbf{p}) + \sum_{k=1}^{s} q_k h_k(t, \mathbf{x}, \mathbf{u})$$

与(16.52)相伴的拉格朗日函数 $H(t, \mathbf{x}, \mathbf{u}, \mathbf{p})$ 是通常的 Hamilton 函数

16.54 假设$(\mathbf{x}^*(t), \mathbf{u}^*(t))$是问题(16.52)的可接受的对，进一步假设存在函数 $\mathbf{p}(t) = (p_1(t), \cdots, p_n(t))$ 和 $\mathbf{q}(t) = (q_1(t), \cdots, q_s(t))$，其中 $\mathbf{p}(t)$是连续的，而$\dot{\mathbf{p}}(t)$和 $\mathbf{q}(t)$是分段连续的，使以下条件在 $p_0 = 1$ 成立：

(1) $\dfrac{\partial \mathcal{L}^*}{\partial u_j} = 0, \qquad j = 1, \cdots, r$

(2) $q_k(t) \geqslant 0 (= 0$ 如果 $h_k(t, \mathbf{x}^*(t), \mathbf{u}^*(t)) > 0), \qquad k = 1, \cdots, s$

(3) $\dot{p}_i(t) = -\dfrac{\partial \mathcal{L}^*}{\partial x_i}, \qquad i = 1, \cdots, n$

(4) (a') $p_i(t_1)$ 不受约束 $\qquad i = 1, \cdots, l$

(b') $p_i(t_1) \geqslant 0 (= 0$ 如果 $x_i^*(t_1) > x_i^1), \qquad i = l+1, \cdots, m$

(c') $p_i(t_1) = 0, \qquad i = m+1, \cdots, n$

(5) $H(t, \mathbf{x}, \mathbf{u}, \mathbf{p}(t))$在$(\mathbf{x}, \mathbf{u})$是凹的

(6) $h_k(t, \mathbf{x}, \mathbf{u})$在$(\mathbf{x}, \mathbf{u})$是拟凹的，$\qquad k = 1, \cdots, s$

则$(\mathbf{x}^*(t), \mathbf{u}^*(t))$是问题的解.

问题(16.52)的 *Mangasarian* 充分条件. $\mathcal{L}^*$ 取值在 $(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \mathbf{p}(t), \mathbf{q}(t))$(标准的最优解的必要条件包括一个约束限制，严格地限制了能够出现在 h_k 约束内的函数的形式.特别地，在最优点每一有效约束必须包含至少一个控制变量作为自变量，详情参见参考文献).

纯状态约束

16.55 $\max \displaystyle\int_{t_0}^{t_1} f(t, \mathbf{x}(t), \mathbf{u}(t))\mathrm{d}t$

$\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t)), \ \mathbf{x}(t_0) = \mathbf{x}^0$

$\mathbf{u}(t) = (u_1(t), \cdots, u_r(t)) \in U \subset \mathbf{R}^r$

$h_k(t, \mathbf{x}(t)) \geqslant 0, \ k = 1, \cdots, s$

(a) $x_i(t_1) = x_i^1, \qquad i = 1, \cdots, l$

(b) $x_i(t_1) \geqslant x_i^1, \qquad i = l+1, \cdots, q$

(c) $x_i(t_1)$ 不受限制，$\quad i = q+1, \cdots, n$

纯状态约束问题. U 是控制区域. $h_1, \cdots, h_s$ 是给定函数.

16.56 $\mathcal{L}(t, x, u, p, q) = H(t, x, u, p) + \sum_{k=1}^{s} q_k h_k(t, x)$

与(16.55)相伴的拉格朗日函数. $H(t, x, u, p)$是通常的 Hamilton 函数.

16.57 假设$(x^*(t), u^*(t))$在问题(16.55)中是可接受的,且存在向量函数 $p(t)$和 $q(t)$,其中 $p(t)$是连续的,而$\dot{p}(t)$和 $q(t)$在$[t_0, t_1]$内是分段连续的,还有数 β_j, $j = 1, \cdots, s$, 使以下条件在 $p_0 = 1$ 成立:

(1) $u = u^*(t)$ 对于 u 在 U 内将 $H(t, x^*(t), u, p(t))$最大化

(2) $q_k(t) \geqslant 0 (= 0$ 如果 $h_k(t, x^*(t)) > 0)$, $k = 1, \cdots, s$

(3) $\dot{p}_i(t) = -\dfrac{\partial \mathcal{L}^*}{\partial x_i}$, $\quad i = 1, \cdots, n$

(4) $p_i(t)$在 t_1 跳跃性不连续,即

$$p_i(t_1^-) - p_i(t_1) = \sum_{j=1}^{s} \beta_j \frac{\partial h_j(t_1, x^*(t_1))}{\partial x_i}, \quad i = 1, \cdots, n$$

(5) $\beta_k \geqslant 0 (= 0$ 如果 $h_k(t_1, x^*(t_1)) > 0)$, $k = 1, \cdots, s$

(6) (a′) $p_i(t_1)$不受限制, $\quad i = 1, \cdots, l$

(b′) $p_i(t_1) \geqslant 0 (= 0$ 如果 $x_i^*(t_1) > x_i^1)$, $\quad i = l+1, \cdots, m$

(c′) $p_i(t_1) = 0$, $\quad i = m+1, \cdots, n$

(7) $\hat{H}(t, x, p(t)) = \max_{u \in U} H(t, x, u, p(t))$ 在 x 是凹的.

(8) $h_k(t, x)$在 x 是拟凹的, $\quad k = 1, \cdots, s$

则$(x^*(t), u^*(t))$是问题的解.

纯状态约束问题(16.55)的充分条件. $p(t) = (p_1(t), \cdots, p_n(t))$, $q(t) = (q_1(t), \cdots, q_s(t))$. $\mathcal{L}^*$ 取值在 $(t, x^*(t), u^*(t), p(t), q(t))$. (定理中的条件多少有限制性. 特别的,有时必须允许 $p(t)$ 在$[t_0, t_1]$内点上有不连续点,详情参见参考文献.)

混合与纯状态约束

16.58 $\max\int_{t_0}^{t_1} f(t, \boldsymbol{x}(t), \boldsymbol{u}(t))\mathrm{d}t$

$\dot{\boldsymbol{x}}(t) = \boldsymbol{g}(t, \boldsymbol{x}(t), \boldsymbol{u}(t)),\ \boldsymbol{x}(t_0) = \boldsymbol{x}^0$

$\boldsymbol{u}(t) = (u_1(t), \cdots, u_r(t)) \in U \subset \mathbf{R}^r$

$h_k(t, \boldsymbol{x}(t), \boldsymbol{u}(t)) \geqslant 0,\quad k = 1, \cdots, s'$

$h_k(t, \boldsymbol{x}(t), \boldsymbol{u}(t)) = \overline{h}_k(t, \boldsymbol{x}(t)) \geq 0,\quad k = s' + 1, \cdots, s$

(a) $x_i(t_1) = x_i^1,\quad i = 1, \cdots, l$

(b) $x_i(t_1) \geqslant x_i^1,\quad i = l + 1, \cdots, q$

(c) $x_i(t_1)$ 不受限制，$i = q + 1, \cdots, n$

一个混合与纯状态约束问题.

16.59 设$(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t))$为在问题(16.58)中可接受的.假设存在向量函数 $\boldsymbol{p}(t)$和 $\boldsymbol{q}(t)$,其中$\boldsymbol{p}(t)$是连续的而$\dot{\boldsymbol{p}}(t)$和 $\boldsymbol{q}(t)$是分段连续的，还有数 β_j, $j = 1, \cdots, s$，使以下条件在 $p_0 = 1$ 时得以满足：

(1) $\left(\dfrac{\partial \mathcal{L}^*}{\partial \boldsymbol{u}}\right)\cdot(\boldsymbol{u} - \boldsymbol{u}^*(t)) \leqslant 0$ 对于所有 $\boldsymbol{u} \in U$ 成立.

(2) $\dot{p}_i(t) = -\dfrac{\partial \mathcal{L}^*}{\partial x_i},\quad i = 1, \cdots, n$

(3) $p_i(t_1) - \sum_{k=1}^{s}\beta_k \dfrac{\partial h_k(t_1, \boldsymbol{x}^*(t_1), \boldsymbol{u}^*(t_1))}{\partial x_i}$

满足

(a′) 不受限制，$i = 1, \cdots, l$

(b′) $\geqslant 0$($= 0$ 如果 $x_i^*(t_1) > x_i^1$)，$i = l + 1, \cdots, m$

(c′) $= 0$，$i = m + 1, \cdots, n$

(4) $\beta_k = 0$，$k = 1, \cdots, s'$

(5) $\beta_k \geqslant 0$($= 0$ 如果$\overline{h}_k(t_1, \boldsymbol{x}^*(t_1)) > 0$)，$k = s' + 1, \cdots, s$

(6) $q_k(t) \geqslant 0$($= 0$ 如果 $h_k(t, \boldsymbol{x}^*(t), \boldsymbol{u}^*(t)) > 0$)，$k = 1, \cdots, s$

(7) $h_k(t, \boldsymbol{x}, \boldsymbol{u})$在$(\boldsymbol{x}, \boldsymbol{u})$是拟凹的，$k = 1, \cdots, s$.

(8) $H(t, \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}(t))$在$(\boldsymbol{x}, \boldsymbol{u})$是凹的.

则$(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t))$是问题的解.

混合与纯状态约束问题的 *Mangasarian* 充分条件($p(t)$是连续的). $\mathcal{L}$ 定义在(16.53)，且 $\mathcal{L}^*$ 取值在$(t, \boldsymbol{x}^*(t), \boldsymbol{u}^*(t), \boldsymbol{p}(t), \boldsymbol{q}(t))$, $\boldsymbol{p}(t) = (p_1(t), \cdots, p_n(t))$, $\boldsymbol{q}(t) = (q_1(t), \cdots, q_s(t))$. 约束认证是不需要的.但通常这些条件不会满足，因为$\boldsymbol{p}(t)$不连续，特别在 t_1.对于允许 $\boldsymbol{p}(t)$ 在$[t_0, t_1]$的内点上有不连续点的充分性结论，参见如 Seierstad & Sydsaeter (1987) 定理 6.2.

参考文献

Kamien & Schwartz (1991), Léonard & Long (1992), Beavis & Dobbs (1990),以及 Intrilligator (1971).对于更详细的内容,参考如 Seierstad & Sydsaeter (1987)或 Feichtinger & Hartl (1986)(德语).

17 离散动态最优化

动态规划

17.1 $\max \sum_{t=0}^{T} f(t, \mathbf{x}_t, \mathbf{u}_t)$

$\mathbf{x}_{t+1} = \mathbf{g}(t, \mathbf{x}_t, \mathbf{u}_t),\ t = 0, \cdots, T-1$

$\mathbf{x}_0 = \mathbf{x}^0,\ \mathbf{x}_t \in \mathbf{R}^n,\ \mathbf{u}_t \in U \subset \mathbf{R}^r,\ t = 0, \cdots, T$

动态规划问题. 此处 $\mathbf{g} = (g_1, \cdots, g_n)$, 且 $\mathbf{x}^0$ 是在 $\mathbf{R}^n$ 内的给定向量, U 是控制区域.

17.2 $J_s(\mathbf{x}) = \max_{\mathbf{u}_s, \cdots, \mathbf{u}_T \in U} \sum_{t=s}^{T} f(t, \mathbf{x}_t, \mathbf{u}_t)$, 其中

$\mathbf{x}_{t+1} = \mathbf{g}(t, \mathbf{x}_t, \mathbf{u}_t),\ t = s, \cdots, T-1,\ \mathbf{x}_s = \mathbf{x}$

问题(17.1)的值函数 $J_s(\mathbf{x})$的定义.

17.3 $J_T(\mathbf{x}) = \max_{\mathbf{u} \in U} f(T, \mathbf{x}, \mathbf{u})$,

$J_s(\mathbf{x}) = \max_{\mathbf{u} \in U} [f(s, \mathbf{x}, \mathbf{u}) + J_{s+1}(\mathbf{g}(s, \mathbf{x}, \mathbf{u}))]$

对于 $s = 0, 1, \cdots, T-1$.

动态规划中的基础方程(Bellman 方程).

17.4 动态规划问题中的“自由控制参数”形式:

$\max \sum_{t=0}^{T} F(t, \mathbf{x}_t, \mathbf{x}_{t+1})$,

$\mathbf{x}_{t+1} \in \Gamma_t(\mathbf{x}_t),\ t = 0, \cdots, T$, $\mathbf{x}_0$ 是给定的

集合 $\Gamma_t(\mathbf{x}_t)$通常定义为, 对于给定的向量函数 $\mathbf{G}$ 和 $\mathbf{H}$ 的向量不等式形式, $\mathbf{G}(t, \mathbf{x}_t) \leqslant \mathbf{x}_{t+1} \leqslant \mathbf{H}(t, \mathbf{x}_t)$.

17.5 $J_s(\mathbf{x}) = \max \sum_{t=s}^{T} F(t, \mathbf{x}_t, \mathbf{x}_{t+1})$, 其中最大化是对所有 $\mathbf{x}_{t+1} \in \Gamma_t(\mathbf{x}_t),\ t = s, \cdots, T$, 且 $\mathbf{x}_s = \mathbf{x}$.

问题(17.4)的值函数, $J_s(\mathbf{x})$.

17.6 $J_T(\boldsymbol{x}) = \max\limits_{\boldsymbol{y}\in \Gamma_T(\boldsymbol{x})} F(T, \boldsymbol{x}, \boldsymbol{y})$

$J_s(\boldsymbol{x}) = \max\limits_{\boldsymbol{y}\in \Gamma_s(\boldsymbol{x})} [F(s, \boldsymbol{x}, \boldsymbol{y}) + J_{s+1}(\boldsymbol{y})]$

对于 $s = 0, 1, \cdots, T$

问题(17.4)的基础方程.

17.7 如果 $\{\boldsymbol{x}_0^*, \cdots, \boldsymbol{x}_{T+1}^*\}$ 是(17.4)的最优解,其中 $\boldsymbol{x}_{t+1}^*$ 对于所有 t 是 $\Gamma_t(\boldsymbol{x}_t^*)$ 的一个内点,且对应 $\boldsymbol{x} \mapsto \Gamma_t(\boldsymbol{x})$ 是半连续的,则 $\{\boldsymbol{x}_0^*, \cdots, \boldsymbol{x}_{T+1}^*\}$ 满足欧拉向量差分方程

$$\boldsymbol{F}_2'(t+1, \boldsymbol{x}_{t+1}, \boldsymbol{x}_{t+2}) + \boldsymbol{F}_3'(t, \boldsymbol{x}_t, \boldsymbol{x}_{t+1}) = \boldsymbol{0}$$

F 是 $1 + n + n$ 个变量的函数, $\boldsymbol{F}_2'$ 记为 F 对第 2, 3, $\cdots$, $n+1$ 个变量的 n 维偏导数向量,而 $\boldsymbol{F}_3'$ 是 F 对第 $n+2$, $n+3$, $\cdots$, $2n+1$ 个变量的 n 维偏导数向量.

无限时域

17.8 $\max \sum\limits_{t=0}^{\infty} \alpha^t f(\boldsymbol{x}_t, \boldsymbol{u}_t)$

$\boldsymbol{x}_{t+1} = \boldsymbol{g}(\boldsymbol{x}_t, \boldsymbol{u}_t)$, $t = 0, 1, 2, \cdots$

$\boldsymbol{x}_0 = \boldsymbol{x}^0$, $\boldsymbol{x}_t \in \mathbf{R}^n$, $\boldsymbol{u}_t \in U \subset \mathbf{R}^r$, $t = 0, 1, 2, \cdots$

一个无限时域问题. $\alpha \in (0, 1)$ 是一固定的贴现因子.

17.9 序列 $\{(\boldsymbol{x}_t, \boldsymbol{u}_t)\}$ 称为可接受的,如果 $\boldsymbol{u}_t \in U$, $\boldsymbol{x}_0 = \boldsymbol{x}^0$, 且差分方程(17.8)对所有 $t = 0, 1, 2, \cdots$ 都满足.

可接受序列的定义.

17.10 (B) $M \leqslant f(\boldsymbol{x}, \boldsymbol{u}) \leqslant N$

(BB) $f(\boldsymbol{x}, \boldsymbol{u}) \geqslant M$

(BA) $f(\boldsymbol{x}, \boldsymbol{u}) \leqslant N$

有界性条件. M 和 N 是给定的数.

17.11 $V(\boldsymbol{x}, \boldsymbol{\pi}, s, \infty) = \sum_{t=s}^{\infty} \alpha^t f(\boldsymbol{x}_t, \boldsymbol{u}_t)$

其中 $\boldsymbol{\pi} = (\boldsymbol{u}_s, \boldsymbol{u}_{s+1}, \cdots)$, 且 $\boldsymbol{u}_{s+k} \in U$ 对 $k = 0, 1, \cdots$, 而且 $\boldsymbol{x}_{t+1} = \boldsymbol{g}(\boldsymbol{x}_t, \boldsymbol{u}_t)$ 对 $t = s, s+1, \cdots$, $\boldsymbol{x}_s = \boldsymbol{x}$ 成立.

给定在 $t = s$ 时的状态向量 $\boldsymbol{x}$,从 s 时期至以后所能获得的总效用.

17.12 $J_s(\boldsymbol{x}) = \sup_{\boldsymbol{\pi}} V(\boldsymbol{x}, \boldsymbol{\pi}, s, \infty)$

其中上确界是对所有向量 $\boldsymbol{\pi} = (\boldsymbol{u}_s, \boldsymbol{u}_{s+1}, \cdots)$ 的，且 $\boldsymbol{u}_{s+k} \in U$，$(\boldsymbol{x}_t, \boldsymbol{u}_t)$ 对于 $t \geqslant s$ 是可接受的，且 $\boldsymbol{x}_s = \boldsymbol{x}$.

问题(17.8)的值函数.

17.13 $J_s(\boldsymbol{x}) = \alpha^s J_0(\boldsymbol{x}),\ s = 1, 2, \cdots$

$$J_0(\boldsymbol{x}) = \max_{\boldsymbol{u} \in U}\{f(\boldsymbol{x}, \boldsymbol{u}) + \alpha J_0(\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{u}))\}$$

值函数的性质，假设(17.10)中至少一个有界性条件是满足的.

离散最优控制理论

17.14 $H = f(t, \boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{p}\boldsymbol{g}(t, \boldsymbol{x}, \boldsymbol{u}),\ t = 0, \cdots, T$

与(17.1)相伴的 Hamilton 函数 $H = H(t, \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p})$，其中 $\boldsymbol{p} = (p^1, \cdots, p^n)$.

17.15 假设$\{(\boldsymbol{x}_t^*, \boldsymbol{u}_t^*)\}$是问题(17.1)的一个最优可接受序列，则在 $\mathbf{R}^n$ 内存在向量 $\boldsymbol{p}_t$，使对于 $t = 0, \cdots, T$：

(1) $H'_{\boldsymbol{u}}(t, \boldsymbol{x}_t^*, \boldsymbol{u}_t^*, \boldsymbol{p}_t) \cdot (\boldsymbol{u} - \boldsymbol{u}_t^*) \leqslant 0$ 对于所有 $\boldsymbol{u} \in U$ 成立

(2) 向量 $\boldsymbol{p}_t = (p_t^1, \cdots, p_t^n)$ 是

$$\boldsymbol{p}_{t-1} = H'_{\boldsymbol{x}}(t, \boldsymbol{x}_t^*, \boldsymbol{u}_t^*, \boldsymbol{p}_t),$$
$$t = 1, \cdots, T$$

的一个解，且 $\boldsymbol{p}_T = \mathbf{0}$

(17.1)的最大化原理，最优解的必要条件，U 是凸的(Hamilton 函数不一定由 $\boldsymbol{u}_t^*$ 最大化).

17.16 (a) $x_T^i = \bar{x}^i$　对于 $i = 1, \cdots, l$

(b) $x_T^i \geqslant \bar{x}^i$　对于 $i = l + 1, \cdots, m$

(c) x_T^i 不受限制　对于 $i = m + 1, \cdots, n$

问题(17.1)的终端条件.

17.17 $$H=\begin{cases} q_0 f(t,\ x,\ u)+pg(t,\ x,\ u),\ t=0, \\ \quad \cdots,\ T-1 \\ f(T,\ x,\ u),\ t=T \end{cases}$$

与(17.1)相伴有终端条件(17.16)的 Hamilton 函数 $H = H(t,\ x,\ u,\ p)$.

17.18 假设$\{(x_t^*,\ u_t^*)\}$是问题(17.1)的一个最优序列,且有终端条件(17.16),则在 $\mathbf{R}^n$ 内存在向量 p_t 和一个数 q_0,且 $(q_0,\ p_T)\neq(0,\ \mathbf{0})$, $q_0=0$ 或 1,使对于 $t=0,\ \cdots,\ T$:

(1) $H_u'(t,\ x_t^*,\ u_t^*,\ p_t)\cdot(u-u_t^*)\leqslant 0$ 对于所有 $u\in U$ 成立

(2) $p_t=(p_t^1,\ \cdots,\ p_t^n)$ 是

$p_{t-1}^i=H_{x^i}'(t,\ x_t^*,\ u_t^*,\ p_t),\ t=1,\ \cdots,\ T-1$ 的一个解

(3) $$p_{T-1}^i=q_0\frac{\partial f(T,\ \mathbf{x}_T^*,\ \mathbf{u}_T^*)}{\partial x_T^i}+p_T^i,$$

其中 p_T^i 满足

(a′) p_T^i 不受限制, $\quad i=1,\ \cdots,\ l$

(b′) $p_T^i\geqslant 0(=0$ 如果 $x_T^{*i}>\bar{x}^i)$, $\quad i=l+1,\ \cdots,\ m$

(c′) $p_T^i=0$, $\quad i=m+1,\ \cdots,\ n$

有终端条件(17.16)的(17.1)的最大化原理,最优解的必要条件.当(17.16)中的(a),(b),或(c)分别成立时,(a′),(b′)或(c′)成立. U 是凸的(除了在退化的情况下,我们可以设定 $q_0=1$).

17.19 假设当 $q_0=1$ 时$\{(x_t^*,\ u_t^*,\ p_t)\}$满足(17.18)中的所有条件,并进一步假设 $H(t,\ x,\ u,\ p_t)$在$(x,\ u)$对每一 $t\geqslant 0$ 是凹的.则序列$\{(x_t^*,\ u_t^*,\ p_t)\}$是最优的.

最优解的充分条件.

无限时域

17.20 $$\max\sum_{t=0}^{\infty}f(t,\ x_t,\ u_t)$$

$$x_{t+1}=g(t,\ x_t,\ u_t),\ t=0,\ 1,\ 2,\ \cdots$$

$$x_0=x^0,\ x_t\in\mathbf{R}^n,\ u_t\in U\subset\mathbf{R}^r,\ t=0,\ 1,\ 2,\ \cdots$$

这里假设每一可接受对无限求和是收敛的.

17.21 序列$\{(x_t^*, u_t^*)\}$是赶上最优的(CU-最优)，如果对每一可接受的序列$\{(x_t, u_t)\}$，

$$\underline{\lim}_{t\to\infty} D(t) \geqslant 0$$

其中 $D(t) = \sum_{\tau=0}^{t}(f(\tau, x_\tau^*, u_\tau^*) - f(\tau, x_\tau, u_\tau))$

"赶上最优"的定义. 关于 $\underline{\lim}$ 参见(12.40)和(12.41).

17.22 假设序列$\{(x_t^*, u_t^*, p_t)\}$满足(17.18)中的条件(1)和(2)且 $q_0 = 1$，进一步假设 Hamilton 函数 $H(t, x, u, p_t)$在(x, u)对于每一 t 是凹的，则$\{(x_t^*, u_t^*\})$是 CU-最优的，如果以下极限条件是满足的：对于所有可接受的序列$\{(x_t, u_t)\}$，

$$\underline{\lim}_{t\to\infty} p_t \cdot (x_t - x_t^*) \geqslant 0$$

无终端条件的无限时域问题的最优解的充分条件.

参 考 文 献

参考 Bellman(1957)以及 Stokey, Lucas, & Prescott (1989).

18 $\mathbf{R}^n$ 中的向量 抽象空间

18.1 $\boldsymbol{a}_1=\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{n1}\end{pmatrix},\ \boldsymbol{a}_2=\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{n2}\end{pmatrix},\ \cdots,\ \boldsymbol{a}_m=\begin{pmatrix}a_{1m}\\a_{2m}\\\vdots\\a_{nm}\end{pmatrix}$

$\mathbf{R}^n$ 中的 m(列)向量.

18.2 如果 $x_1, x_2, \cdots, x_m$ 是实数,则

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots+x_m\boldsymbol{a}_m$$

是 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 的一个线性组合.

向量的线性组合的定义.

18.3 在 $\mathbf{R}^n$ 中的向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 是

- 线性相关的,如果存在不都为零的数 $c_1, c_2, \cdots, c_m$,使

$$c_1\boldsymbol{a}_1+c_2\boldsymbol{a}_2+\cdots+c_m\boldsymbol{a}_m=\mathbf{0}$$

- 线性无关的,如果它们不是线性相关的.

线性相关和线性无关的定义.

18.4 在(18.1)中的向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m$ 是线性无关的,当且仅当矩阵 $(a_{ij})_{n\times m}$ 的秩是 m.

$\mathbf{R}^n$ 中的 m 个向量线性无关的特征.(关于矩阵秩的定义参见(19.23))

18.5 向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 在 $\mathbf{R}^n$ 中是线性无关的,当且仅当

$$\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}\neq 0$$

$\mathbf{R}^n$ 中的 n 个向量线性无关的特征.((18.4)的一个特例)

18.6　在 $\mathbf{R}^n$ 中一个向量的非空子集 V 是 $\mathbf{R}^n$ 的一个子空间，如果 $c_1\boldsymbol{a}_1+c_2\boldsymbol{a}_2\in V$ 对于所有 $\boldsymbol{a}_1, \boldsymbol{a}_2\in V$ 中和所有数 c_1, c_2 成立.

子空间的定义.

18.7　如果 V 是 $\mathbf{R}^n$ 的一个子空间，则 $S[V]$ 是 V 中所有向量线性组合的集合.

$S[V]$ 称为 V 的张成.

18.8　$\mathbf{R}^n$ 的一个子空间 V 中的一组向量 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m$ 称为 V 的基，如果以下两个条件得以满足：

- $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m$ 是线性无关的
- $S[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m]=V$

子空间的基的定义.

18.9　$\mathbf{R}^n$ 的子空间 V 的维数 $\dim V$，是 V 的一个基中向量的个数.（V 的两个基的向量个数总是相同的）

子空间维数的定义.特别的，$\dim\mathbf{R}^n=n$.

18.10　设 V 为 $\mathbf{R}^n$ 的一个 m 维子空间.

- V 中任意 m 个线性无关向量的组合是 V 的一个基.
- V 中任意 m 个能张成 V 的向量组合是 V 的一个基.

关于子空间的重要事实.

18.11　$\boldsymbol{a}=(a_1, \cdots, a_m)$ 和 $\boldsymbol{b}=(b_1, \cdots, b_m)$ 的内积是数

$$\boldsymbol{a}\cdot\boldsymbol{b}=a_1b_1+\cdots+a_mb_m$$

内积的定义，也称为数量积或点积.

18.12

$$\boldsymbol{a}\cdot\boldsymbol{b}=\boldsymbol{b}\cdot\boldsymbol{a}$$
$$\boldsymbol{a}\cdot(\boldsymbol{b}+\boldsymbol{c})=\boldsymbol{a}\cdot\boldsymbol{b}+\boldsymbol{a}\cdot\boldsymbol{c}$$
$$(\alpha\boldsymbol{a})\cdot\boldsymbol{b}=\boldsymbol{a}\cdot(\alpha\boldsymbol{b})=\alpha(\boldsymbol{a}\cdot\boldsymbol{b})$$
$$\boldsymbol{a}\cdot\boldsymbol{a}>0\Leftrightarrow\boldsymbol{a}\neq\boldsymbol{0}$$

内积的性质. α 是一数量（即一实数）.

18.13

$$\|\boldsymbol{a}\|=\sqrt{a_1^2+a_2^2+\cdots+a_n^2}=\sqrt{\mathbf{a}\cdot\mathbf{a}}$$

向量的（*Euclid*）范数（或称模长度）.

18.14　(a) $\|\boldsymbol{a}\|>0$ 对于 $\boldsymbol{a}\neq\boldsymbol{0}$，且 $\|\boldsymbol{0}\|=0$

(b) $\|\alpha\boldsymbol{a}\|=|\alpha|\,\|\boldsymbol{a}\|$

(c) $\|\boldsymbol{a}+\boldsymbol{b}\|\leqslant\|\boldsymbol{a}\|+\|\boldsymbol{b}\|$

(d) $|\boldsymbol{a}\cdot\boldsymbol{b}|\leqslant\|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\|$

范数的性质. $\boldsymbol{a}, \boldsymbol{b}\in\mathbf{R}^n$，$\alpha$ 是一数量.(d)是柯西-施瓦兹不等式. $\|\boldsymbol{a}-\boldsymbol{b}\|$ 是 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 之间的距离.

18.15 两个非零向量 $\mathbf{a}$ 和 $\mathbf{b}$ 之间的角 φ 定义为

$$\cos\varphi = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\|},\ 0\leqslant\varphi<\pi$$

在 $\mathbf{R}^n$ 中两个向量之间的角的定义. 向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 称为正交的,如果 $\boldsymbol{a}\cdot\boldsymbol{b}=0$.

向量空间

18.16 一个向量空间(或线性空间)($\mathbf{R}$ 上的)是一元素为向量的集合 V,其中有两种运算:"加法"($V\times V\to V$)和"数量乘法"($\mathbf{R}\times V\to V$),其对于所有在 V 中的 $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ 和所有实数 α 和 β,满足以下公理:

(a) $(\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z}=\boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z})$,
$\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}$

(b) 有一元素 $\mathbf{0}\in V$ 满足 $\boldsymbol{x}+\mathbf{0}=\boldsymbol{x}$.

(c) 对每一在 V 中的 $\boldsymbol{x}$, V 中有元素 $(-1)\boldsymbol{x}$,使 $\boldsymbol{x}+(-1)\boldsymbol{x}=\mathbf{0}$.

(d) $(\alpha+\beta)\boldsymbol{x}=\alpha\boldsymbol{x}+\beta\boldsymbol{x}$; $\alpha(\beta\boldsymbol{x})=(\alpha\beta)\boldsymbol{x}$;
$\alpha(\boldsymbol{x}+\boldsymbol{y})=\alpha\boldsymbol{x}+\alpha\boldsymbol{y}$; $1\boldsymbol{x}=\boldsymbol{x}$

向量空间的定义. 通过明显的改进,关于线性组合,线性相关和线性无关的向量集合,子空间,以及张成的定义(18.2)(18.3),(18.6)和(18.7),能扩展至向量空间.

18.17 在向量空间 V 中的向量集合 B 是 V 的一个基,如果 B 中的向量是线性无关的,且张成 V, $S[B]=V$.

向量空间的基的定义.

度量空间

18.18 度量空间是一集合 M,具有距离函数 $d: M\times M\to\mathbf{R}$,对于所有在 M 中的 x, y, z 得以满足以下公理:

(a) $d(x, y)\geqslant 0$ 且 $d(x, y)=0\Leftrightarrow x=y$

(b) $d(x, y)=d(y, x)$

(c) $d(x, y)\leqslant d(x, z)+d(z, y)$

度量空间的定义. 距离函数 d 称为 M 上的一个度量, (c)称为三角形不等式.

18.19 在一度量空间中的序列$\{x_n\}$是

- 收敛于极限 x 的，可记为 $\lim_{n\to\infty} x_n = x$（或 $x_n \to x$ 当 $n \to \infty$），如果当 $n \to \infty$ 时 $d(x_n, x) \to 0$；
- 柯西序列，如果对每一 $\varepsilon > 0$，存在一整数 N，使 $d(x_n, x_m) < \varepsilon$ 对所有 $m, n \geqslant N$ 成立.

重要的定义. 一个不收敛的序列称为发散的.

18.20 一个在度量空间 M 内的子集合 S 在 M 中是稠密的，如果 M 中的每一点都是由 S 中的点所构成的序列的极限.

稠密子集的定义.

18.21 一个度量空间 M 是

- 完整的，如果在 M 中的每一柯西序列是收敛的；
- 可分的，如果存在一个可数的 M 的子集 S 在 M 中是稠密的.

完整的和可分的度量空间的定义.

赋范向量空间 Banach 空间

18.22 一个赋范向量空间（在 $\mathbf{R}$ 上）是一向量空间 V，且有一函数 $\|\cdot\|: V \to \mathbf{R}$，使对于所有在 V 上的 $\boldsymbol{x}, \boldsymbol{y}$ 和所有实数 α：

(a) $\|\boldsymbol{x}\| > 0$ 对于 $\boldsymbol{x} \neq \mathbf{0}$ 成立，且 $\|\mathbf{0}\| = 0$

(b) $\|\alpha \boldsymbol{x}\| = |\alpha| \, \|\boldsymbol{x}\|$

(c) $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$

当距离函数为 $d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$ 时，V 称为一个度量空间. 如果这一度量空间是完整的，则 V 称为一 *Banach* 空间.

18.23 • $l^p(n)$: $\mathbf{R}^n$, 且 $\|\boldsymbol{x}\| = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$ $(p \geqslant 1)$(当 $p = 2$ 时,是 Euclid 范数)

• $l^\infty(n)$: $\mathbf{R}^n$, 且 $\|\boldsymbol{x}\| = \max(|x_1|, \cdots, |x_n|)$

• $l^p(p \geqslant 1)$: 所有实数的无限序列 $\boldsymbol{x} = (x_0, x_1, \cdots)$ 的集合,使 $\sum_{i=1}^{\infty} |x_i|^p$ 收敛. $\|\boldsymbol{x}\| = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p}$. 对于 l^p 中的 $\boldsymbol{x} = (x_0, x_1, \cdots)$ 和 $\boldsymbol{y} = (y_0, y_1, \cdots)$, $\boldsymbol{x} + \boldsymbol{y} = (x_0 + y_0, x_1 + y_1, \cdots)$ 且 $\alpha\boldsymbol{x} = (\alpha x_0, \alpha x_1, \cdots)$.

• l^∞: 所有有约束的实数的无限序列 $\boldsymbol{x} = (x_0, x_1, \cdots)$ 的集合,且 $\|\boldsymbol{x}\| = \sup_i |x_i|$. (对 l^p 定义的向量运算.)

• $C(X)$: 所有受约束的连续函数 $f: X \to \mathbf{R}$ 的集合,其中 $X \subset \mathbf{R}^n$ 是一度量空间,且 $\|f\| = \sup_{x \in X} |f(t)|$. 如果 f 和 g 是在 $C(X)$ 中且 $\alpha \in \mathbf{R}$, 则 $f + g$ 和 αf 定义为 $(f+g)(x) = f(x) + g(x)$ 及 $(\alpha f)(x) = \alpha f(x)$.

一些标准的赋范向量空间,同时也是 Banach 空间的实例.

18.24 设 X 为一紧的度量空间,并设 F 为 Banach 空间 $C(X)$ 的一个子集(参见(18.23)),且是

• 一致有界的,即存在一个数 M,使 $|f(x)| \leqslant M$ 对于所有 $f \in F$ 和所有 $x \in X$ 成立.

• 同等连续的,即对于每一 $\varepsilon > 0$, 存在一 $\delta > 0$, 使当 $\|x - x'\| < \delta$ 时, $|f(x) - f(x')| < \varepsilon$ 对于所有 $f \in F$ 成立,

则 F 的闭包是紧的.

Ascoli 定理. (与 Schauder 定理 (18.25) 一起,这一结论在经济动态分析中是很有用的. 参见 Stokey, Lucas, & Prescott (1989))

18.25 如果 K 是一在 Banach 空间 X 中的紧凸集合,则 K 的任意映射入自身的连续函数 f 有一不动点,即在 K 中存在一点 x^*,使 $f(x^*) = x^*$.

Schauder 不动点定理.

18.26 设 $T: X \to X$ 为一完整的度量空间 X 到自己的映射,假设在 $[0, 1)$ 中存在一个数 k,使

(*) $d(T_x, T_y) \leqslant k d(x, y)$ 对于所有 X 中的 x, y 成立,则

(a) T 有一不动点 x^*,即 $T(x^*) = x^*$.

(b) $d(T^n x^0, x^*) \leqslant k^n d(x^0, x^*)$ 对于所有 $x^0 \in X$ 和所有 $n = 0, 1, 2, \cdots$ 成立.

收缩映射中不动点的存在. k 称为收缩映射的模(参见(6.23)对于某一 $k \in [0, 1)$, 满足(*)的映射称为收缩映射.

18.27 设 $C(X)$ 为(18.23)定义的 Banach 空间,而 T 为一 $C(X)$ 到 $C(X)$ 的映射,且满足:

(a) (单调性)如果 $f, g \in C(X)$ 且 $f(x) \leqslant g(x)$ 对于所有 $x \in X$ 成立,则 $(Tf)(x) \leqslant (Tg)(x)$ 对于所有 $x \in X$ 成立.

(b) (贴现性)存在某个 $\alpha \in (0, 1)$, 使对于所有 $f \in C(X)$ 和所有 $a \geqslant 0$, 及所有 $x \in X$, $[T(f + a)](x) \leqslant (Tf)(x) + \alpha a$

则 T 是一模为 α 的收缩映射.

收缩映射的 Blackwell 充分条件. 在此 $(f + a)(x)$ 定义为 $f(x) + a$.

内积空间 Hilbert 空间

18.28 内积空间(在 $\mathbf{R}$ 上)是一向量空间 V,且对于在 V 内的每一有序向量对 $(\boldsymbol{x}, \boldsymbol{y})$ 有一相伴的实数,记为 $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$,对于所有在 V 中的 $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ 和所有实数 α:

(a) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$

(b) $\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle$

(c) $\alpha \langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \alpha \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \alpha \boldsymbol{y} \rangle$

(d) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geqslant 0$ 且 $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0 \Leftrightarrow \boldsymbol{x} = 0$

内积空间的定义. 如果我们定义 $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$, 则 V 成为一赋范向量空间. 如果这一空间是完整的, V 称为一 *Hilbert* 空间.

18.29 • $l^2(n)$, $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i y_i$

• l^2, $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{\infty} x_i y_i$

Hilbert 空间的例子.

18.30 (a) $|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leqslant \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}\sqrt{\langle \boldsymbol{y}, \boldsymbol{y} \rangle}$ 对于所有 $\boldsymbol{x}, \boldsymbol{y} \in V$ 成立

(b) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \frac{1}{4}[\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2]$

(a)是柯西-施瓦兹不等式(等号当且仅当 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是线性相关时成立). 在(b)中的等式表明内积可用范数来表示.

18.31 • 在一内积空间 V 中的两个向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是正交的,如果 $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$.

• 在 V 中的向量集合 S 称为是正交的,如果 $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$ 对于所有 $\boldsymbol{x} \neq \boldsymbol{y} \in S$ 成立.

• 在 V 中的向量集合 S 称为是单位正交的,如果它是正交的且 $\|\boldsymbol{x}\| = 1$ 对于所有 $\boldsymbol{x} \in S$ 成立.

• 在 V 中的标准正交向量集合 S 称为是完整的,如果在 V 中不存在与所有 S 中的向量正交的 $\boldsymbol{x}$.

重要的定义.

18.32 设 U 为一在内积空间 V 中标准正交的向量集合.

(a) 如果 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 是 U 的任意有限个相异元素的组合,则

(*) $\sum_{i=1}^{n} |(\boldsymbol{x}, \boldsymbol{u}_i)|^2 \leqslant \|\boldsymbol{x}\|^2$ 对于所有 V 中的 $\boldsymbol{x}$ 成立.

(b) 如果 V 完整的(一个 Hilbert 空间),且 U 是 V 的一完整的标准正交子集,则

(**) $\sum_{\boldsymbol{u} \in U} |(\boldsymbol{x}, \boldsymbol{u})|^2 = \|\boldsymbol{x}\|^2$ 对于所有在 V 中的 $\boldsymbol{x}$ 成立.

(*)是 *Bessel* 不等式,(**)是 *Parseval* 公式.

参考文献

所有关于在 $\mathbf{R}^n$ 中的向量的结论都是标准的,均可在任意线性代数教材中找到,如 Fraleigh & Beauregard (1995)或 Lang(1987).对于抽象空间,参考 Kolmogorov & Fomin (1970),或 Royden (1968).对于收缩映射及其在动态经济分析中的应用,参考 Stokey, Lucas, & Prescott (1989).

19 矩阵

19.1 $\mathbf{A}=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}=(a_{ij})_{m\times n}$

矩阵的表示法,其中 a_{ij}是在第 i 行第 j 列的元素,矩阵有阶数 $m\times n$,如果 $m=n$,该矩阵是 n 阶方阵.

19.2 $\mathbf{A}=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ 0 & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & a_{nn}\end{pmatrix}$

一个上三角矩阵(所有对角线下的元素都为 0).$\mathbf{A}$ 的转置称为下三角矩阵.

19.3 $\operatorname{diag}(a_1,a_2,\cdots,a_n)=\begin{pmatrix}a_1 & 0 & \cdots & 0\\ 0 & a_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & a_n\end{pmatrix}$

一个对角矩阵.

19.4 $\begin{pmatrix}a & 0 & \cdots & 0\\ 0 & a & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & a\end{pmatrix}_{n\times n}$

一个标量矩阵.

19.5 $\mathbf{I}_n=\begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{pmatrix}_{n\times n}$

单位矩阵或恒等矩阵.

19.6 如果 $\mathbf{A}=(a_{ij})_{m\times n}$，$\mathbf{B}=(b_{ij})_{m\times n}$，且 α 是一数量，我们定义

$$\mathbf{A}+\mathbf{B}=(a_{ij}+b_{ij})_{m\times n}$$
$$\alpha\mathbf{A}=(\alpha a_{ij})_{m\times n}$$
$$\mathbf{A}-\mathbf{B}=\mathbf{A}+(-1)\mathbf{B}=(a_{ij}-b_{ij})_{m\times n}$$

矩阵运算(数量是实数或复数).

19.7
$$(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C})$$
$$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$$
$$\mathbf{A}+\mathbf{0}=\mathbf{A}$$
$$\mathbf{A}+(-\mathbf{A})=\mathbf{0}$$
$$(a+b)\mathbf{A}=a\mathbf{A}+b\mathbf{A}$$
$$a(\mathbf{A}+\mathbf{B})=a\mathbf{A}+a\mathbf{B}$$

矩阵运算的性质.$\mathbf{0}$ 是零矩阵，其所有元素均为零.a 和 b 是数量.

19.8 如果 $\mathbf{A}=(a_{ij})_{m\times n}$ 且 $\mathbf{B}=(b_{ij})_{n\times p}$，我们定义乘积 $\mathbf{C}=\mathbf{AB}$ 为 $m\times p$ 矩阵，$\mathbf{C}=(c_{ij})_{m\times p}$，其中

$$c_{ij}=a_{i1}b_{1j}+\cdots+a_{ik}b_{kj}+\cdots+a_{in}b_{nj}$$

矩阵乘法的定义.

$$\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ \boxed{a_{i1}} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & \cdots & b_{1k} & \cdots & b_{1p} \\ \vdots & & \vdots & & \vdots \\ b_{j1} & \cdots & b_{jk} & \cdots & b_{jp} \\ \vdots & & \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nk} & \cdots & b_{np} \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1k} & \cdots & c_{1p} \\ \vdots & \vdots & \vdots & & \\ c_{i1} & \cdots & \boxed{c_{ik}} & \cdots & c_{ip} \\ \vdots & \vdots & \vdots & \ddots & \\ c_{m1} & \cdots & c_{mk} & \cdots & c_{mp} \end{pmatrix}$$

19.9
$$(\mathbf{AB})\mathbf{C}=\mathbf{A}(\mathbf{BC})$$
$$\mathbf{A}(\mathbf{B}+\mathbf{C})=\mathbf{AB}+\mathbf{AC}$$
$$(\mathbf{A}+\mathbf{B})\mathbf{C}=\mathbf{AC}+\mathbf{BC}$$

矩阵乘法的性质.

19.10
$$\mathbf{AB}\neq\mathbf{BA}$$
$$\mathbf{AB}=\mathbf{0}\nRightarrow\mathbf{A}=\mathbf{0}\text{ 或 }\mathbf{B}=\mathbf{0}$$
$$\mathbf{AB}=\mathbf{AC}\text{ 和 }\mathbf{A}\neq\mathbf{0}\nRightarrow\mathbf{B}=\mathbf{C}$$

矩阵乘法中的重要现象.$\mathbf{0}$ 是零矩阵.$\nRightarrow$ 应读作："不一定隐含".

19.11
$$\mathbf{A}'=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

$\mathbf{A}'$，$\mathbf{A}=(a_{ij})_{m\times n}$ 的转置，是一 $n\times m$ 矩阵，通过对换 $\mathbf{A}$ 的行和列而得到的.

19.12 $(\mathbf{A}')' = \mathbf{A}$

$(\mathbf{A}+\mathbf{B})' = \mathbf{A}' + \mathbf{B}'$

$(\alpha\mathbf{A})' = \alpha\mathbf{A}'$

$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$(注意顺序!)

转置法则.

19.13 $\mathbf{B} = \mathbf{A}^{-1} \Longleftrightarrow \mathbf{AB} = \mathbf{I}_n \Longleftrightarrow \mathbf{BA} = \mathbf{I}_n$

一个 $n \times n$ 矩阵 $\mathbf{A}$ 的逆. $\mathbf{I}_n$ 是单位矩阵.

19.14 $\mathbf{A}^{-1}$存在 $\Longleftrightarrow |\mathbf{A}| \neq 0$

矩阵有逆矩阵,即可逆的一个充分必要条件. $|\mathbf{A}|$表示方阵 $\mathbf{A}$ 的行列式(参见第 20 章).

19.15 $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow \mathbf{A}^{-1} = \dfrac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

当 $|\mathbf{A}| = ad - bc \neq 0$ 时成立.

19.16 如果 $\mathbf{A} = (a_{ij})_{n\times n}$ 是一方阵且 $|\mathbf{A}| \neq 0$, $\mathbf{A}$ 的唯一的逆矩阵为

$\mathbf{A}^{-1} = \dfrac{1}{|\mathbf{A}|}\operatorname{adj}(\mathbf{A})$, 其中

$$\operatorname{adj}(\mathbf{A}) = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

元素 a_{ij}的代数余子式A_{ij}定义为

$$A_{ij} = (-1)^{i+j}\begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & \boxed{a_{ij}} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

方阵的逆矩阵的一般公式. 注意伴随矩阵 adj($\mathbf{A}$)元素标号的顺序! 矩阵 $(A_{ij})_{n\times n}$ 称为代数余子式矩阵,而伴随矩阵是它的转置. 在代数余子式 A_{ij}的公式中,行列式是通过删除$|\mathbf{A}|$的第 i 行和第 j 列得到的.

19.17	$(\mathbf{A}^{-1})^{-1} = \mathbf{A}$ $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$(注意顺序!) $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$ $(c\mathbf{A})^{-1} = c^{-1}\mathbf{A}^{-1}$	逆矩阵的性质.(A 和 B 是可逆的 $n \times n$ 矩阵,$c \neq 0$ 是一数量)
19.18	$(\mathbf{I}_m + \mathbf{AB})^{-1} = \mathbf{I}_m - \mathbf{A}(\mathbf{I}_n + \mathbf{BA})^{-1}\mathbf{B}$	A 是 $m \times n$ 矩阵, B 是 $n \times m$ 矩阵, $\lvert\mathbf{I}_m + \mathbf{AB}\rvert \neq 0$.
19.19	$\mathbf{R}^{-1}\mathbf{A}'(\mathbf{AR}^{-1}\mathbf{A}' + \mathbf{Q}^{-1})^{-1} = (\mathbf{A}'\mathbf{QA} + \mathbf{R})^{-1}\mathbf{A}'\mathbf{Q}$	矩阵的逆阵对.当逆矩阵存在时成立.
19.20	一个 n 阶的方阵 A 称为 • 对称的,如果 $\mathbf{A} = \mathbf{A}'$ • 反对称的,如果 $\mathbf{A} = -\mathbf{A}'$ • 幂等的,如果 $\mathbf{A}^2 = \mathbf{A}$ • 对合的,如果 $\mathbf{A}^2 = \mathbf{I}_n$ • 正交的,如果 $\mathbf{A}'\mathbf{A} = \mathbf{I}_n$ • 奇异的,如果 $\lvert\mathbf{A}\rvert = 0$, 非奇异的如果 $\lvert\mathbf{A}\rvert \neq 0$	一些重要的定义. $\lvert\mathbf{A}\rvert$ 记为方阵 A 的行列式(参见第 20 章).对于幂等矩阵和正交矩阵的性质,参见第 22 章.
19.21	$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$	方阵 $\mathbf{A} = (a_{ij})_{n\times n}$ 的迹,是对角线元素的和.
19.22	$\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$ $\mathrm{tr}(c\mathbf{A}) = c\,\mathrm{tr}(\mathbf{A})$ (c 是一数量) $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$ (如果 AB 是一方阵) $\mathrm{tr}(\mathbf{A}') = \mathrm{tr}(\mathbf{A})$	迹的性质.
19.23	$r(\mathbf{A})$ = A 中线性无关的行的最大个数 = A 中线性无关的列的最大个数 = A 的最大的非零子式的阶数.	矩阵的秩的等价定义.关于子式,参见(20.15).

19.24 (1) $r(\mathbf{A}) = r(\mathbf{A}') = r(\mathbf{A}'\mathbf{A}) = r(\mathbf{A}\mathbf{A}')$

(2) $r(\mathbf{AB}) \leqslant \min(r(\mathbf{A}), r(\mathbf{B}))$

(3) $r(\mathbf{AB}) = r(\mathbf{B})$, 如果 $|\mathbf{A}| \neq 0$

(4) $r(\mathbf{CA}) = r(\mathbf{C})$, 如果 $|\mathbf{A}| \neq 0$

(5) $r(\mathbf{PAQ}) = r(\mathbf{A})$,如果 $|\mathbf{P}| \neq 0$, $|\mathbf{Q}| \neq 0$

(6) $|r(\mathbf{A}) - r(\mathbf{B})| \leqslant r(\mathbf{A}+\mathbf{B}) \leqslant r(\mathbf{A}) + r(\mathbf{B})$

(7) $r(\mathbf{AB}) \geqslant r(\mathbf{A}) + r(\mathbf{B}) - n$

(8) $r(\mathbf{AB}) + r(\mathbf{BC}) \leqslant r(\mathbf{B}) + r(\mathbf{ABC})$

秩的性质.矩阵的阶数满足所需运算的定义.在结论(7)(*Sylvester* 不等式)中,$\mathbf{A}$ 和 $\mathbf{B}$ 是 $n \times n$.(8)称为 *Frobenius* 不等式.

19.25 $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ 对于某一 $\boldsymbol{x} \neq \mathbf{0} \Leftrightarrow r(\mathbf{A}) \leqslant n - 1$

关于齐次方程的有用的结论.$\mathbf{A}$ 是 $m \times n$,$\boldsymbol{x}$ 是 $n \times 1$.

19.26 矩阵的范数是一函数 $\|\cdot\|_\beta$ 使对于每一有一相伴实数 $\|\mathbf{A}\|_\beta$ 的方阵 $\mathbf{A}$,有:

- $\|\mathbf{A}\|_\beta > 0$ 对于 $\mathbf{A} \neq \mathbf{0}$ 成立和 $\|\mathbf{0}\|_\beta = 0$
- $\|c\mathbf{A}\|_\beta = |c|\,\|\mathbf{A}\|_\beta$($c$ 是一数量)
- $\|\mathbf{A}+\mathbf{B}\|_\beta \leqslant \|\mathbf{A}\|_\beta + \|\mathbf{B}\|_\beta$
- $\|\mathbf{AB}\|_\beta \leqslant \|\mathbf{A}\|_\beta \|\mathbf{B}\|_\beta$

矩阵范数的定义.(存在无限个这样的范数,其中一部分在(19.27)给出)

19.27

- $\|\mathbf{A}\|_1 = \max\limits_{j=1,\cdots,n} \sum\limits_{i=1}^{n} |a_{ij}|$
- $\|\mathbf{A}\|_\infty = \max\limits_{i=1,\cdots,n} \sum\limits_{j=1}^{n} |a_{ij}|$
- $\|\mathbf{A}\|_2 = \sqrt{\lambda}$,其中 λ 是 $\mathbf{A}'\mathbf{A}$ 的最大的特征值.
- $\|\mathbf{A}\|_M = n \max\limits_{i,j=1,\cdots,n} |a_{ij}|$
- $\|\mathbf{A}\|_T = \left(\sum\limits_{j=1}^{n} \sum\limits_{i=1}^{n} |a_{ij}|^2\right)^{1/2}$

$\mathbf{A} = (a_{ij})_{n\times n}$ 的一些矩阵范数.(对于特征值,参见第 21 章)

19.28 λ 是 $\mathbf{A} = (a_{ij})_{n\times n}$ 的特征值$\Rightarrow |\lambda| \leqslant \|\mathbf{A}\|_\beta$

$\mathbf{A}$ 的任意特征值的模数小于或等于 $\mathbf{A}$ 的任意矩阵范数.

19.29	$\\|\mathbf{A}\\|_{\beta} < 1 \Rightarrow$ 当 $t \to \infty$ $\mathbf{A}^t \to \mathbf{0}$	当 $t \to \infty$ 时，$\mathbf{A}^t \to \mathbf{0}$ 的充分条件. $\\|\mathbf{A}\\|_{\beta}$是 $\mathbf{A}$ 的任意矩阵范数.
19.30	$\mathrm{e}^{\mathbf{A}} = \sum_{n=0}^{\infty} \frac{1}{n!}\mathbf{A}^n$	方阵 $\mathbf{A}$ 的指数矩阵.
19.31	$\mathrm{e}^{\mathbf{A}+\mathbf{B}} = \mathrm{e}^{\mathbf{A}}\mathrm{e}^{\mathbf{B}}$ 如果 $\mathbf{AB} = \mathbf{BA}$ $(\mathrm{e}^{\mathbf{A}})^{-1} = \mathrm{e}^{-\mathbf{A}}$, $\frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^{x\mathbf{A}}) = \mathbf{A}\mathrm{e}^{x\mathbf{A}}$	指数矩阵的性质.

线性变换

19.32	一函数 $T: \mathbf{R}^n \to \mathbf{R}^m$ 称为是一线性变换(或函数),如果 (1) $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ (2) $T(c\mathbf{x}) = cT(\mathbf{x})$ 对于所有 $\mathbf{R}^n$ 中的 $\mathbf{x}$ 和 $\mathbf{y}$ 和所有数量 c 成立.	线性变换的定义.
19.33	如果 $\mathbf{A}$ 是一 $m \times n$ 矩阵,由 $T_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ 定义的函数 $T_{\mathbf{A}}: \mathbf{R}^n \to \mathbf{R}^m$ 是一线性变换.	一个重要的事实.
19.34	设 $T: \mathbf{R}^n \to \mathbf{R}^m$ 为一线性变换,且 $\mathbf{A}$ 为第 j 列是 $T(\mathbf{e}_j)$的 $m \times n$ 矩阵,其中 $\mathbf{e}_j$ 是在 $\mathbf{R}^n$ 中的第 j 个标准单位向量.则 $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$ 对于所有 $\mathbf{R}^n$ 中的 $\mathbf{x}$ 成立.	矩阵 $\mathbf{A}$ 称为 T 的标准矩阵表示.
19.35	设 $T: \mathbf{R}^n \to \mathbf{R}^m$ 和 $S: \mathbf{R}^m \to \mathbf{R}^k$ 为两个线性变换,且标准矩阵表示分别为 $\mathbf{A}$ 和 $\mathbf{B}$,则两个线性变换的结合 $S \circ T$ 是有标准矩阵表示 $\mathbf{BA}$ 的线性变换.	基本事实.
19.36	设 $\mathbf{A}$ 为一可逆 $n \times n$ 矩阵且有相伴的线性变换 T,与 $\mathbf{A}^{-1}$相对应的变换 T^{-1}是 T 逆变换(函数).	基本事实.

广义逆矩阵

19.37 一个 $n \times m$ 矩阵 $\mathbf{A}^-$ 称为 $m \times n$ 矩阵 $\mathbf{A}$ 的广义逆矩阵，如果它满足

$$\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$$

矩阵的广义逆矩阵的定义.（$\mathbf{A}^-$ 一般并不是唯一的）

19.38 矩阵方程 $\mathbf{A}\mathbf{x} = \mathbf{b}$ 有解的一个充分必要条件是 $\mathbf{A}\mathbf{A}^-\boldsymbol{b} = \boldsymbol{b}$，一般解为 $\boldsymbol{x} = \mathbf{A}^-\boldsymbol{b} + (\mathbf{I} - \mathbf{A}^-\mathbf{A})\boldsymbol{q}$，其中 $\boldsymbol{q}$ 是一有相应阶数的任意向量.

广义逆矩阵的重要应用.

19.39 如果 $\mathbf{A}^-$ 是 $\mathbf{A}$ 的一个广义逆矩阵，则

- $\mathbf{A}\mathbf{A}^-$ 和 $\mathbf{A}^-\mathbf{A}$ 是幂等的
- $r(\mathbf{A}) = r(\mathbf{A}^-\mathbf{A}) = \mathrm{tr}(\mathbf{A}^-\mathbf{A})$
- $(\mathbf{A}^-)'$ 是 $\mathbf{A}'$ 的广义逆矩阵
- $\mathbf{A}$ 是非奇异方阵 $\Rightarrow \mathbf{A}^- = \mathbf{A}^{-1}$

广义逆矩阵的性质.

19.40 $n \times m$ 矩阵 $\mathbf{A}^+$ 称为 $m \times n$ 实矩阵 $\mathbf{A}$ 的 *Moore-Penrose* 逆矩阵，如果它满足以下四个条件：

(i) $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$　　(ii) $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$

(iii) $(\mathbf{A}\mathbf{A}^+)' = \mathbf{A}\mathbf{A}^+$　　(iv) $(\mathbf{A}^+\mathbf{A})' = \mathbf{A}^+\mathbf{A}$

Moore-Penrose 逆矩阵的定义.（$\mathbf{A}^+$ 存在且是唯一的）

19.41 矩阵方程 $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ 有解的一个充分必要条件是 $\mathbf{A}\mathbf{A}^+\boldsymbol{b} = \boldsymbol{b}$. 相应的一般解为 $\boldsymbol{x} = \mathbf{A}^+\boldsymbol{b} + (\mathbf{I} - \mathbf{A}^+\mathbf{A})\boldsymbol{q}$，其中 $\boldsymbol{q}$ 是一有相应阶数的任意向量.

Moore-Penrose 逆矩阵的一个重要应用.

19.42
- $\mathbf{A}$ 是非奇异的方阵 $\Rightarrow \mathbf{A}^+ = \mathbf{A}^{-1}$
- $(\mathbf{A}^+)^+ = \mathbf{A}$, $(\mathbf{A}')^+ = (\mathbf{A}^+)'$
- $\mathbf{A}^+ = \mathbf{A}$，如果 $\mathbf{A}$ 是对称和幂等的.
- $\mathbf{A}^+\mathbf{A}$ 和 $\mathbf{A}\mathbf{A}^+$ 是幂等的.
- $\mathbf{A}$, $\mathbf{A}^+$, $\mathbf{A}\mathbf{A}^+$，和 $\mathbf{A}^+\mathbf{A}$ 有相同的秩.
- $\mathbf{A}'\mathbf{A}\mathbf{A}^+ = \mathbf{A}' = \mathbf{A}^+\mathbf{A}\mathbf{A}'$
- $(\mathbf{A}\mathbf{A}^+)^+ = \mathbf{A}\mathbf{A}^+$
- $(\mathbf{A}'\mathbf{A})^+ = \mathbf{A}^+(\mathbf{A}^+)'$, $(\mathbf{A}\mathbf{A}')^+ = (\mathbf{A}^+)'\mathbf{A}^+$
- $(\mathbf{A} \otimes \mathbf{B})^+ = \mathbf{A}^+ \otimes \mathbf{B}^+$

Moore-Penrose 逆矩阵的性质.（$\otimes$ 是 Kronecker 乘积，参见第 23 章）

分块矩阵

19.43 $\mathbf{P}=\begin{pmatrix}\mathbf{P}_{11} & \mathbf{P}_{12}\\ \mathbf{P}_{21} & \mathbf{P}_{22}\end{pmatrix}$

一个 $(p+q)\times(r+s)$ 阶分块矩阵. ($\mathbf{P}_{11}$是 $p\times r$ 矩阵, $\mathbf{P}_{12}$是 $p\times s$ 矩阵, $\mathbf{P}_{21}$是 $q\times r$ 矩阵, $\mathbf{P}_{22}$是 $q\times s$ 矩阵)

19.44 $\mathbf{P}=\begin{pmatrix}\mathbf{P}_{11} & \mathbf{P}_{12}\\ \mathbf{P}_{21} & \mathbf{P}_{22}\end{pmatrix}, \mathbf{Q}=\begin{pmatrix}\mathbf{Q}_{11} & \mathbf{Q}_{12}\\ \mathbf{Q}_{21} & \mathbf{Q}_{22}\end{pmatrix}\Rightarrow$

$$\mathbf{PQ}=\begin{pmatrix}\mathbf{P}_{11}\mathbf{Q}_{11}+\mathbf{P}_{12}\mathbf{Q}_{21} & \mathbf{P}_{11}\mathbf{Q}_{12}+\mathbf{P}_{12}\mathbf{Q}_{22}\\ \mathbf{P}_{21}\mathbf{Q}_{11}+\mathbf{P}_{22}\mathbf{Q}_{21} & \mathbf{P}_{21}\mathbf{Q}_{12}+\mathbf{P}_{22}\mathbf{Q}_{22}\end{pmatrix}$$

分块矩阵的乘法(我们假设涉及到的乘法是有定义的).

19.45 $\begin{vmatrix}\mathbf{P}_{11} & \mathbf{P}_{12}\\ \mathbf{P}_{21} & \mathbf{P}_{22}\end{vmatrix}=|\mathbf{P}_{11}|\cdot|\mathbf{P}_{22}-\mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{P}_{12}|$

一个分块的 $n\times n$ 矩阵的行列式,假设 $\mathbf{P}_{11}^{-1}$存在.

19.46 $\begin{vmatrix}\mathbf{P}_{11} & \mathbf{P}_{12}\\ \mathbf{P}_{21} & \mathbf{P}_{22}\end{vmatrix}=|\mathbf{P}_{22}|\cdot|\mathbf{P}_{11}-\mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{21}|$

一个分块的 $n\times n$ 矩阵的行列式,假设 $\mathbf{P}_{22}^{-1}$存在.

19.47 $\begin{vmatrix}\mathbf{P}_{11} & \mathbf{P}_{12}\\ \mathbf{0} & \mathbf{P}_{22}\end{vmatrix}=|\mathbf{P}_{11}|\cdot|\mathbf{P}_{22}|$

一个特例.

19.48 $\begin{pmatrix}\mathbf{P}_{11} & \mathbf{P}_{12}\\ \mathbf{P}_{21} & \mathbf{P}_{22}\end{pmatrix}^{-1}=$

$$\begin{pmatrix}\mathbf{P}_{11}^{-1}+\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\boldsymbol{\Delta}^{-1}\mathbf{P}_{21}\mathbf{P}_{11}^{-1} & -\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\boldsymbol{\Delta}^{-1}\\ -\boldsymbol{\Delta}^{-1}\mathbf{P}_{21}\mathbf{P}_{11}^{-1} & \boldsymbol{\Delta}^{-1}\end{pmatrix}$$

其中 $\boldsymbol{\Delta}=\mathbf{P}_{22}-\mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{P}_{12}$.

分块矩阵的逆矩阵,假设 $\mathbf{P}_{11}^{-1}$存在.

19.49 $\begin{pmatrix}\mathbf{P}_{11} & \mathbf{P}_{12}\\ \mathbf{P}_{21} & \mathbf{P}_{22}\end{pmatrix}^{-1}=$

$$\begin{pmatrix}\boldsymbol{\Delta}_1^{-1} & -\boldsymbol{\Delta}_1^{-1}\mathbf{P}_{12}\mathbf{P}_{22}^{-1}\\ -\mathbf{P}_{22}^{-1}\mathbf{P}_{21}\boldsymbol{\Delta}_1^{-1} & \mathbf{P}_{22}^{-1}+\mathbf{P}_{22}^{-1}\mathbf{P}_{21}\boldsymbol{\Delta}_1^{-1}\mathbf{P}_{12}\mathbf{P}_{22}^{-1}\end{pmatrix}$$

其中 $\boldsymbol{\Delta}_1=\mathbf{P}_{11}-\mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{21}$.

分块矩阵的逆矩阵,假设 $\mathbf{P}_{22}^{-1}$存在.

复元素矩阵

19.50 设 $\mathbf{A} = (a_{ij})$ 为一 $m \times n$ 复矩阵(即 $\mathbf{A}$ 的元素为复数).则

- $\overline{\mathbf{A}} = (\bar{a}_{ij})$ 称为 $\mathbf{A}$ 的共轭.($\bar{a}_{ij}$表示a_{ij}的复共轭.)
- $\mathbf{A}^* = \overline{\mathbf{A}}' = (\bar{a}_{ji})$ 称为 $\mathbf{A}$ 的共轭转置.
- $\mathbf{A}$ 称为 Hermitian 矩阵,如果 $\mathbf{A} = \mathbf{A}^*$.
- $\mathbf{A}$ 称为酉矩阵如果 $\mathbf{A}^* = \mathbf{A}^{-1}$.

关于复矩阵的一些有用的定义.

19.51
- $\mathbf{A}$ 是实的 $\Leftrightarrow \mathbf{A} = \overline{\mathbf{A}}$.
- 如果 $\mathbf{A}$ 是实的,则

  $\mathbf{A}$ 是 Hermitian 矩阵$\Leftrightarrow\mathbf{A}$ 是对称的.

根据定义得到的简单推论.

19.52 设 $\mathbf{A}$ 和 $\mathbf{B}$ 为复矩阵,c 为一复数.则

(1) $(\mathbf{A}^*)^* = \mathbf{A}$

(2) $(\mathbf{A}+\mathbf{B})^* = \mathbf{A}^* + \mathbf{B}^*$

(3) $(c\mathbf{A})^* = \bar{c}\,\mathbf{A}^*$

(4) $(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*$

共轭转置的性质.(2)和(4)当矩阵的和与积有定义时成立.

参 考 文 献

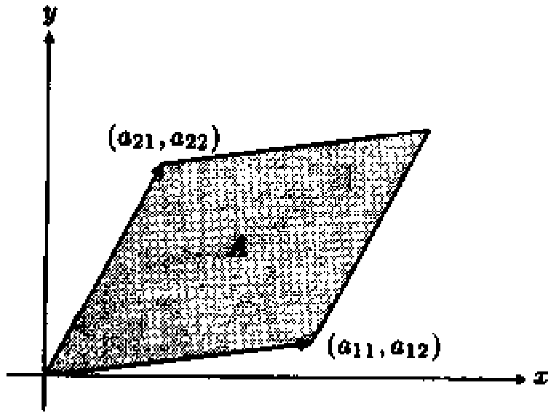
大多数公式都是标准的,并能在任何线性代数的教材里找到,如 Fraleigh & Beauregard (1995)和 Lang (1987).对于(19.26)—(19.29)参见 Faddeeva (1959).对于广义逆矩阵,参见 Magnus & Neudecker (1988).一种标准的参考文献是 Gantmacher (1959).

20 行列式

20.1 $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$

2×2 行列式的定义.

20.2



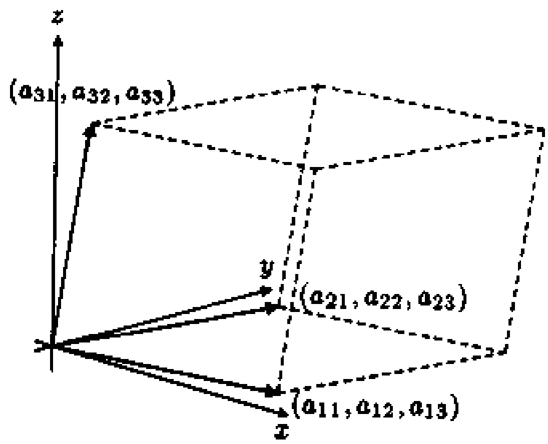
2×2 行列式的几何解释. 面积 A 就是行列式的绝对值

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

20.3 $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{cases} a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} \\ + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} \\ + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \end{cases}$

3×3 行列式的定义.

20.4



3×3 行列式的几何解释. 由三个向量张成的"盒子"的体积是行列式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

的绝对值.

20.5 如果 $\mathbf{A}=(a_{ij})_{n\times n}$ 是一 $n\times n$ 矩阵，$\mathbf{A}$ 的行列式为

$$|\mathbf{A}|=a_{i1}A_{i1}+\cdots+a_{in}A_{in}=\sum_{j=1}^{n}a_{ij}A_{ij}$$

元素 a_{ij} 的代数余子式 A_{ij} 是

$$A_{ij}=(-1)^{i+j}\begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & \boxed{a_{ij}} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

通过沿第 i 行余子式展开而定义的 n 阶行列式. 行列式的值与 i 的选择无关.

20.6 $a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in}=|\mathbf{A}|$

$a_{i1}A_{k1}+a_{i2}A_{k2}+\cdots+a_{in}A_{kn}=0$ 如果 $k\neq i$

$a_{1j}A_{1j}+a_{2j}A_{2j}+\cdots+a_{nj}A_{nj}=|\mathbf{A}|$

$a_{1j}A_{1k}+a_{2j}A_{2k}+\cdots+a_{nj}A_{nk}=0$ 如果 $k\neq j$

余子式沿同一行或同一列展开产生行列式. 余子式沿不同的行或列展开产生的行列式为 0.

20.7
- 如果 $\mathbf{A}$ 的一行(或列)的所有元素都为 0，则 $|\mathbf{A}|=0$.
- 如果交换 $\mathbf{A}$ 的两行(或两列)，行列式改变符号但绝对值不变.
- 如果 $\mathbf{A}$ 一行(或一列)中的所有元素乘以一个数 c，则行列式乘以 c.
- 如果 $\mathbf{A}$ 的两行(或两列)成比例，则 $|\mathbf{A}|=0$.
- 如果 $|\mathbf{A}|$ 的一行(或一列)的倍数加到另一行(或一列)上，行列式的值不变.
- $|\mathbf{A}'|=|\mathbf{A}|$，其中 $\mathbf{A}'$ 是 $\mathbf{A}$ 的转置.

行列式的重要性质. $\mathbf{A}$ 是方阵.

20.8 $|\mathbf{AB}|=|\mathbf{A}|\cdot|\mathbf{B}|$

$|\mathbf{A}+\mathbf{B}|\neq|\mathbf{A}|+|\mathbf{B}|$ (一般)

行列式的性质. $\mathbf{A}$ 和 $\mathbf{B}$ 是 $n\times n$ 矩阵.

20.9	$\begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = (x_2-x_1)(x_3-x_1)(x_3-x_2)$	$n=3$ 时的 *Vandermonde* 行列式.
20.10	$\begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1 \leqslant j < i \leqslant n} (x_i - x_j)$	一般的 Vandermonde 行列式.
20.11	$\begin{vmatrix} a_1 & 1 & \cdots & 1 \\ 1 & a_2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & a_n \end{vmatrix}$ $=(a_1-1)(a_2-1)\cdots(a_n-1)\left[1+\sum_{i=1}^{n} \frac{1}{a_i-1}\right]$	一个特殊行列式. $a_i \neq 1$, $i=1, \cdots, n$.
20.12	$\begin{vmatrix} 0 & p_1 & \cdots & p_n \\ q_1 & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ q_n & a_{n1} & \cdots & a_{nn} \end{vmatrix} = -\sum_{i=1}^{n} \sum_{j=1}^{n} p_i A_{ji} q_j$	一个有用的行列式 ($n \geqslant 2$). A_{ji} 参见 (20.5).
20.13	$\begin{vmatrix} \alpha & p_1 & \cdots & p_n \\ q_1 & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ q_n & a_{n1} & \cdots & a_{nn} \end{vmatrix} = (\alpha - \mathbf{P}'\mathbf{A}^{-1}\mathbf{Q}) \mid \mathbf{A} \mid$	当 $\mathbf{A}^{-1}$ 存在时 (20.12)的一般化. $\mathbf{P}' = (p_1, \cdots, p_n)$, $\mathbf{Q}' = (q_1, \cdots, q_n)$.
20.14	$\mid \mathbf{AB} + \mathbf{I}_m \mid = \mid \mathbf{BA} + \mathbf{I}_n \mid$	一个有用的结论. $\mathbf{A}$ 是 $m \times n$, $\mathbf{B}$ 是 $n \times m$.

20.15 • 一个 $\mathbf{A}$ 的 k 阶子式是一通过删去 $\mathbf{A}$ 的除 k 行和 k 列外所有行和列所得到的 $k \times k$ 矩阵的行列式.

• 一个 $\mathbf{A}$ 的 k 阶主子式是一通过删去除 k 行外所有行以及除相同序号的 k 列外所有列所得到的子式.

• $\mathbf{A}$ 的 k 阶前主子式是通过删去除前 k 行和前 k 列外的所有行和列所得到的主子式.

矩阵的子式,主子式,和前主子式的定义.

20.16 $$D_k = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}, \quad k = 1, 2, \cdots, n$$

$\mathbf{A} = (a_{ij})_{n \times n}$ 的前主子式.

20.17 如果 $|\mathbf{A}| = |(a_{ij})_{n\times n}| \neq 0$, 则 n 个方程 n 个未知数的方程组

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

有唯一解

$$x_j = \frac{|\mathbf{A}_j|}{|\mathbf{A}|}, \quad j = 1, 2, \cdots, n$$

其中

$$|\mathbf{A}_j| = \begin{vmatrix} a_{11} & \cdots & a_{1j-1} & b_1 & a_{1j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j-1} & b_2 & a_{2j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj-1} & b_n & a_{nj+1} & \cdots & a_{nn} \end{vmatrix}$$

Cramer 法则. 注意 $|\mathbf{A}_j|$ 是通过将 $|\mathbf{A}|$ 中的第 j 列替换为以 $b_1, b_2, \cdots, b_n$ 为元素的向量得到的.

参 考 文 献

绝大多数的公式都是标准的,且均可在几乎所有线性代数的教材里找到,如 Fraleigh & Beauregard (1995)或 Lang (1987),一种标准的参考文献是 Gantmacher (1959).

21 特征值 二次型

21.1 数量 λ 称为 $n\times n$ 矩阵 $\mathbf{A}$ 的一个特征值，如果存在一个 n 阶向量 $\boldsymbol{c}\neq\mathbf{0}$ 使

$$\mathbf{A}\boldsymbol{c}=\lambda\boldsymbol{c}$$

向量 $\boldsymbol{c}$ 称为 $\mathbf{A}$ 的特征向量.

特征值和特征向量也称为本征根和本征向量. λ 和 $\boldsymbol{c}$ 可为复数，即使 $\mathbf{A}$ 是实的.

21.2 $$|\mathbf{A}-\lambda\mathbf{I}|=\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}$$

$\mathbf{A}=(a_{ij})_{n\times n}$ 的特征多项式(本征多项式). $\mathbf{I}$ 是 n 阶单位矩阵.

21.3 λ 是 $\mathbf{A}$ 的一个特征值 $\Leftrightarrow p(\lambda)=|\mathbf{A}-\lambda\mathbf{I}|=0$

λ 为 $\mathbf{A}$ 的特征值的一个充分必要条件.

21.4 $$|\mathbf{A}|=\lambda_1\cdot\lambda_2\cdots\lambda_{n-1}\cdot\lambda_n$$
$$\operatorname{tr}(\mathbf{A})=a_{11}+\cdots+a_{nn}=\lambda_1+\cdots+\lambda_n$$

$\lambda_1,\cdots,\lambda_n$ 是 $\mathbf{A}$ 的特征值.

21.5 设 $f(\,)$ 为一多项式. 如果 λ 是 $\mathbf{A}$ 的一个特征值，则 $f(\lambda)$ 是 $f(\mathbf{A})$ 的一个特征值.

矩阵多项式的特征值.

21.6 一个方阵 $\mathbf{A}$ 有一逆矩阵，当且仅当 0 不是 $\mathbf{A}$ 的一个特征值. 如果 $\mathbf{A}$ 有逆矩阵且 λ 是 $\mathbf{A}$ 的一个特征值，则 λ^{-1} 是 $\mathbf{A}^{-1}$ 的一个特征值.

怎样找到方阵的逆矩阵的特征值.

21.7 $\mathbf{A}$ 的所有特征值的模数(严格)小于 1，当且仅当 $t\to\infty$ 时 $\mathbf{A}^t\to\mathbf{0}$.

一个重要的结论.

21.8 $\mathbf{AB}$ 与 $\mathbf{BA}$ 有相同的特征值.

$\mathbf{A}$ 和 $\mathbf{B}$ 是 $n\times n$ 矩阵.

21.9	如果 **A** 是对称的且仅有实元素,则 **A** 的所有特征值均为实的.	
21.10	如果 $p(\lambda)=(-\lambda)^n+b_{n-1}(-\lambda)^{n-1}+\cdots+b_1(-\lambda)+b_0$ 是 **A** 的特征多项式,则 b_k 是 **A** 的所有 $n-k$ 阶主子式的和$\left(\text{共有}\binom{n}{k}\text{个这样的主子式}\right)$.	$n\times n$ 矩阵 **A** 的特征多项式系数的特点(对于主子式,参见(20.15)) $p(\lambda)=0$ 称为 **A** 的特征值方程或特征方程.
21.11	$\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix}=(-\lambda)^2+b_1(-\lambda)+b_0$ 其中 $b_1=a_{11}+a_{22}=\mathrm{tr}(\mathbf{A}),\ b_0=\lvert\mathbf{A}\rvert$	当 $n=2$ 时的(21.10)($\mathrm{tr}(\mathbf{A})$是 **A** 的迹).
21.12	$\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix}=$ $(-\lambda)^3+b_2(-\lambda)^2+b_1(-\lambda)+b_0$ 其中 $b_2=a_{11}+a_{22}+a_{33}=\mathrm{tr}(\mathbf{A})$ $b_1=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}+\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}+\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$ $b_0=\lvert\mathbf{A}\rvert$	当 $n=3$ 时的(21.10).
21.13	**A** 是可对角化的 $\Leftrightarrow \begin{cases} \mathbf{P}^{-1}\mathbf{AP}=\mathbf{D}\ \text{对某一} \\ \text{矩阵}\ \mathbf{P}\ \text{和某一} \\ \text{对角矩阵}\ \mathbf{D}\ \text{成立.} \end{cases}$	定义.
21.14	**A** 和 $\mathbf{P}^{-1}\mathbf{AP}$ 有相同的特征值.	
21.15	如果 $\mathbf{A}=(a_{ij})_{n\times n}$ 有 n 个不同的特征值,则 **A** 是可对角化的.	**A** 为可对角化的充分(但不是必要)条件.

21.16 $\mathbf{A}=(a_{ij})_{n\times n}$ 有 n 个不同的特征值，则 $\mathbf{A}$ 是可对角化的.

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}$$

其中 $\mathbf{P}=(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n)_{n\times n}$.

可对角化矩阵的特点.

21.17 如果 $\mathbf{A}=(a_{ij})_{n\times n}$ 是对称的且有特征值 λ_1, λ_2, $\cdots$, λ_n，则存在一正交矩阵 $\mathbf{U}$，使

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{U}=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}$$

对称矩阵的谱定理.对于正交矩阵的性质，参见第 22 章.

21.18 如果 $\mathbf{A}$ 是一 $n\times n$ 矩阵且有特征值 λ_1, $\cdots$, λ_n(不一定各不相同)，则存在一可逆 $n\times n$ 矩阵 $\mathbf{T}$，使

$$\mathbf{T}^{-1}\mathbf{A}\mathbf{T}=\begin{pmatrix}\mathbf{J}_{k_1}(\lambda_1) & 0 & \cdots & 0\\ 0 & \mathbf{J}_{k_2}(\lambda_2) & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \mathbf{J}_{k_r}(\lambda_r)\end{pmatrix}$$

其中 $k_1+k_2+\cdots+k_r=n$ 且 $\mathbf{J}_k$ 是 $k\times k$ 矩阵

$$\mathbf{J}_k(\lambda)=\begin{pmatrix}\lambda & 1 & 0 & \cdots & 0\\ 0 & \lambda & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & 1\\ 0 & 0 & 0 & \cdots & \lambda\end{pmatrix},\ \mathbf{J}_1(\lambda)=\lambda$$

Jordan 分解定理.

21.19 设 $\mathbf{A}$ 为一复 $n\times n$ 矩阵，则存在酉矩阵 $\mathbf{U}$，使 $\mathbf{U}^{-1}\mathbf{A}\mathbf{U}$ 为上三角形矩阵.

Schur 引理.(对于酉矩阵，参见(19.50))

21.20 设 $\mathbf{A}=(a_{ij})$ 为一 Hermitian 矩阵.则存在酉矩阵 $\mathbf{U}$，使 $\mathbf{U}^{-1}\mathbf{A}\mathbf{U}$ 为一对角矩阵.所有 $\mathbf{A}$ 的特征值则为实的.

Hermitian 矩阵的谱定理.(对于 Hermitian 矩阵，参见(19.50).)

21.21	给定任意矩阵 $\mathbf{A} = (a_{ij})_{n\times n}$，对于每一 $\varepsilon > 0$，存在一矩阵 $\mathbf{B}_\varepsilon = (b_{ij})_{n\times n}$，有 n 个不同的特征值，使 $$\sum_{i,j=1}^{n} \mid a_{ij} - b_{ij} \mid < \varepsilon$$	将矩阵的元素作微小的改变，可得到具有不同特征值的矩阵.
21.22	一个方阵 $\mathbf{A}$ 满足其自身的特征方程：$$p(\mathbf{A}) = (-\mathbf{A})^n + b_{n-1}(-\mathbf{A})^{n-1} + \cdots + b_1(-\mathbf{A}) + b_0\mathbf{I} = \mathbf{0}$$	Cayley-Hamilton 定理. 多项式 $p(\,)$ 的定义见(21.10).
21.23	$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Rightarrow \mathbf{A}^2 - \operatorname{tr}(\mathbf{A})\mathbf{A} + \mid \mathbf{A} \mid \mathbf{I} = \mathbf{0}$	$n = 2$ 时的 Cayley-Hamilton 定理.(参见(21.11))
21.24	$$Q = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = \begin{aligned} & a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n \\ & + \cdots\cdots\cdots\cdots\cdots\cdots \\ & + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 \end{aligned}$$	一个 n 个变量 $x_1, \cdots, x_n$ 的二次型. 我们可在不失一般性的条件下假设 $a_{ij} = a_{ji}$ 对于所有 $i, j = 1, \cdots, n$ 成立.
21.25	$Q = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = \boldsymbol{x}'\mathbf{A}\boldsymbol{x}$，其中 $$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \text{且 } \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$	矩阵形式的二次型. 我们可在不失一般性的条件下，假设 $\mathbf{A}$ 是对称的.
21.26	$\boldsymbol{x}'\mathbf{A}\boldsymbol{x}$ 是 PD$\Leftrightarrow \boldsymbol{x}'\mathbf{A}\boldsymbol{x} > 0$ 对于所有 $\boldsymbol{x} \neq \mathbf{0}$ 成立 $\boldsymbol{x}'\mathbf{A}\boldsymbol{x}$ 是 PSD$\Leftrightarrow \boldsymbol{x}'\mathbf{A}\boldsymbol{x} \geqslant 0$ 对于所有 $\boldsymbol{x}$ 成立 $\boldsymbol{x}'\mathbf{A}\boldsymbol{x}$ 是 ND$\Leftrightarrow \boldsymbol{x}'\mathbf{A}\boldsymbol{x} < 0$ 对于所有 $\boldsymbol{x} \neq \mathbf{0}$ 成立 $\boldsymbol{x}'\mathbf{A}\boldsymbol{x}$ 是 NSD$\Leftrightarrow \boldsymbol{x}'\mathbf{A}\boldsymbol{x} \leqslant 0$ 对于所有 $\boldsymbol{x}$ 成立 $\boldsymbol{x}'\mathbf{A}\boldsymbol{x}$ 是 ID$\Leftrightarrow \boldsymbol{x}'\mathbf{A}\boldsymbol{x}$ 既非 PSD，也非 NSD	二次型($\boldsymbol{x}'\mathbf{A}\boldsymbol{x}$)和对称矩阵($\mathbf{A}$)的定性. 共有五种类型：正定性(PD)，半正定性(PSD)，负定性(ND)，半负定性(NSD)，以及不定性(ID).

21.27	$\mathbf{x}'\mathbf{A}\mathbf{x}$ 是 PD$\Rightarrow a_{ii}>0$ 对于 $i=1,\cdots,n$ $\mathbf{x}'\mathbf{A}\mathbf{x}$ 是 PSD$\Rightarrow a_{ii}\geqslant 0$ 对于 $i=1,\cdots,n$ $\mathbf{x}'\mathbf{A}\mathbf{x}$ 是 ND$\Rightarrow a_{ii}<0$ 对于 $i=1,\cdots,n$ $\mathbf{x}'\mathbf{A}\mathbf{x}$ 是 NSD$\Rightarrow a_{ii}\leqslant 0$ 对于 $i=1,\cdots,n$	令(21.24)中 $x_i=1$ 及 $x_j=0$ 而得到 $j\neq i$.
21.28	$\mathbf{x}'\mathbf{A}\mathbf{x}$ 是 PD$\Leftrightarrow$$\mathbf{A}$ 的所有特征值 >0 $\mathbf{x}'\mathbf{A}\mathbf{x}$ 是 PSD$\Leftrightarrow$$\mathbf{A}$ 的所有特征值 $\geqslant 0$ $\mathbf{x}'\mathbf{A}\mathbf{x}$ 是 ND$\Leftrightarrow$$\mathbf{A}$ 的所有特征值 <0 $\mathbf{x}'\mathbf{A}\mathbf{x}$ 是 NSD$\Leftrightarrow$$\mathbf{A}$ 的所有特征值 $\leqslant 0$	由特征值的符号表述的定性二次型(矩阵)的特点.
21.29	$\mathbf{x}'\mathbf{A}\mathbf{x}$ 是不定性的(ID),当且仅当 $\mathbf{A}$ 至少有一个正的和一个负的特征值.	不定性二次型的一个特点.
21.30	$\mathbf{x}'\mathbf{A}\mathbf{x}$ 是 PD$\Leftrightarrow D_k>0$ 对 $k=1,\cdots,n$ 成立 $\mathbf{x}'\mathbf{A}\mathbf{x}$ 是 ND$\Leftrightarrow(-1)^k D_k>0$ 对 $k=1,\cdots,n$ 成立 其中 $\mathbf{A}$ 的前主子式 D_k 是 $D_k=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k}\\ a_{21} & a_{22} & \cdots & a_{2k}\\ \vdots & \vdots & \ddots & \vdots\\ a_{k1} & a_{k2} & \cdots & a_{kk}\end{vmatrix},\ k=1,2,\cdots,n$	由前主子式表述的定性二次型(矩阵)的特点. 注意将 $>$ 改为 $\geqslant$ 将不会得到关于半定性的类似结论. 例如 $Q=0x_1^2+0x_1x_2-x_2^2$.
21.31	$\mathbf{x}'\mathbf{A}\mathbf{x}$ 是 PSD$\Leftrightarrow\Delta r\geqslant 0$ 对 $r=1,\cdots,n$ 成立, $\mathbf{x}'\mathbf{A}\mathbf{x}$ 是 NSD$\Leftrightarrow(-1)^r\Delta r\geqslant 0$ 对 $r=1,\cdots,n$ 成立,对每个 r, Δr 取遍 $\mathbf{A}$ 的所有 r 阶主子式。	由主子式表述的半正定性和半负定性二次型(矩阵)的特点. (主子式参见(20.15).)
21.32	如果 $\mathbf{A}=(a_{ij})_{n\times n}$ 是正定性的,且 $\mathbf{P}$ 是 $n\times m$ 矩阵并有 $r(\mathbf{P})=m$,则 $\mathbf{P}'\mathbf{A}\mathbf{P}$ 是正定性的.	正定性矩阵的结论.
21.33	如果 $\mathbf{P}$ 是 $n\times m$ 矩阵且 $r(\mathbf{P})=m$,则 $\mathbf{P}\mathbf{P}'$ 是正定性的且秩为 m.	
21.34	如果 $\mathbf{A}$ 是正定性的,则存在一非奇异矩阵 $\mathbf{P}$,使 $\mathbf{P}\mathbf{A}\mathbf{P}'=\mathbf{I}$ 且 $\mathbf{P}'\mathbf{P}=\mathbf{A}^{-1}$.	

21.35 设 $\mathbf{A}$ 为一秩为 $r(\mathbf{A}) = k$ 的 $m \times n$ 矩阵，则存在一 $m \times n$ 酉矩阵 $\mathbf{U}$，一 $n \times n$ 酉矩阵 $\mathbf{V}$，和一个仅有严格为正对角元素的 $k \times k$ 对角矩阵 $\mathbf{D}$，使

$$\mathbf{A} = \mathbf{USV}^*, \text{ 其中 } \mathbf{S} = \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

如果 $k = m = n$，则 $\mathbf{S} = \mathbf{D}$. 如果 $\mathbf{A}$ 是实的，$\mathbf{U}$ 和 $\mathbf{V}$ 可选作正交实矩阵.

奇异值分解定理. $\mathbf{D}$ 的对角元素称为矩阵 $\mathbf{A}$ 的奇异值. 酉矩阵定义见 (19.50)，正交矩阵定义见 (22.8).

21.36 设 $\mathbf{A}$ 和 $\mathbf{B}$ 为对称 $n \times n$ 矩阵，则存在一正交矩阵 $\mathbf{Q}$，使 $\mathbf{Q}'\mathbf{AQ} = \mathbf{D}_1$ 及 $\mathbf{Q}'\mathbf{BQ} = \mathbf{D}_2$，其中 $\mathbf{D}_1$ 和 $\mathbf{D}_2$ 是对角矩阵，当且仅当 $\mathbf{AB} = \mathbf{BA}$.

联合对角化.

二次型

21.37 $(*)\ Q = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j,\ (a_{ij} = a_{ji})$

是在线性约束

$$\begin{aligned} & b_{11}x_1 + \cdots + b_{1n}x_n = 0 \\ (**)\ & \cdots\cdots\cdots\cdots\cdots\cdots \quad (m < n) \\ & b_{m1}x_1 + \cdots + b_{mn}x_n = 0 \end{aligned}$$

下正(负)定的，如果 $Q > 0(< 0)$ 对于所有满足 $(**)$ 的 $(x_1, \cdots, x_n) \neq (0, \cdots, 0)$ 成立.

线性约束下的正(负)定性.

21.38 $$D_r = \begin{vmatrix} 0 & \cdots & 0 & b_{11} & \cdots & b_{1r} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & b_{m1} & \cdots & b_{mr} \\ b_{11} & \cdots & b_{m1} & a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ b_{1r} & \cdots & b_{mr} & a_{r1} & \cdots & a_{rr} \end{vmatrix}$$

与 (21.37) 相伴的加边行列式，$r = 1, \cdots, n$.

21.39 (21.37)中的二次型(*)为受线性约束(* *)的正定性的充分必要条件,假设矩阵 $(b_{ij})_{m\times n}$ 的前 m 列是线性无关的,即

$$(-1)^m D_r > 0, \; r = m+1, \cdots, n$$

(*)为受线性约束(* *)的负定性二次型的对应条件是

$$(-1)^r D_r > 0, \; r = m+1, \cdots, n$$

约束下二次型的定性测试.(假设 $(b_{ij})_{m\times n}$ 的秩为 m 是不够的,例如 $Q(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_3^2$ 约束 $x_3 = 0$)

21.40 二次型 $ax^2 + 2bxy + cy^2$ 对于所有 $(x, y) \neq (0, 0)$,且满足约束 $px + qy = 0$ 时是正的,当且仅当

$$\begin{vmatrix} 0 & p & q \\ p & a & b \\ q & b & c \end{vmatrix} < 0$$

(21.39)的特例,假设 $(p, q) \neq (0, 0)$.

参考文献

绝大多数的公式均可在几乎所有线性代数教材里找到,如 Fraleigh & Beauregard (1995)或 Lang (1987),也可参考 Horn & Johnson (1985),一种标准的参考文献是 Gantmacher (1959).

22 特殊矩阵，Leontief 方程组

幂等矩阵

22.1	$\mathbf{A} = (a_{ij})_{n\times n}$ 是幂等的$\Leftrightarrow \mathbf{A}^2 = \mathbf{A}$	幂等矩阵的定义.
22.2	$\mathbf{A}$ 是幂等的 $\Leftrightarrow \mathbf{I} - \mathbf{A}$ 是幂等的.	幂等矩阵的性质.
22.3	$\mathbf{A}$ 是幂等的 $\Rightarrow$ 0 和 1 是仅有的可能特征值，且 $\mathbf{A}$ 是半正定性的.	
22.4	$\mathbf{A}$ 是幂等的且有 k 个等于 1 的特征值 $\Rightarrow r(\mathbf{A}) = \mathrm{tr}(\mathbf{A}) = k$.	
22.5	$\mathbf{A}$ 是幂等的且 $\mathbf{C}$ 是正交的 $\Rightarrow \mathbf{C}'\mathbf{A}\mathbf{C}$ 是幂等的.	正交距阵定义见 (22.8).
22.6	$\mathbf{A}$ 是幂等的 $\Leftrightarrow$ 与其相关的线性变换是一射影.	一个从 $\mathbf{R}^n$ 到 $\mathbf{R}^n$ 的变换 P 是一射影，如果 $P(P(\boldsymbol{x})) = P(\boldsymbol{x})$ 对于所有在 $\mathbf{R}^n$ 中的 $\boldsymbol{x}$ 成立.
22.7	$\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ 是幂等的.	$\mathbf{X}$ 是 $n \times m$ 矩阵，$\vert\mathbf{X}'\mathbf{X}\vert \neq 0$.

正交矩阵

22.8	$\mathbf{P} = (p_{ij})_{n\times n}$ 是正交的 $\Leftrightarrow \mathbf{P}'\mathbf{P} = \mathbf{P}\mathbf{P}' = \mathbf{I}_n$	正交矩阵的定义.
22.9	$\mathbf{P}$ 是正交的 $\Leftrightarrow \mathbf{P}$ 的列向量是互相正交的单位向量.	正交矩阵的性质.
22.10	$\mathbf{P}$ 和 $\mathbf{Q}$ 都是正交的$\Rightarrow \mathbf{PQ}$ 是正交的.	正交矩阵的性质.

22.11	$\mathbf{P}$是正交的 $\Rightarrow \lvert\mathbf{P}\rvert = \pm 1$，且 1 和 -1 是唯一可能的实特征值.	
22.12	$\mathbf{P}$是正交的 $\Leftrightarrow \lVert \mathbf{P}\boldsymbol{x} \rVert = \lVert \boldsymbol{x} \rVert$ 对于所有 $\mathbf{R}^n$ 中的 $\boldsymbol{x}$ 成立.	正交变换不改变向量的长度.
22.13	如果 $\mathbf{P}$ 是正交的，$\mathbf{P}\boldsymbol{x}$ 与 $\mathbf{P}\boldsymbol{y}$ 之间的角度等于 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 之间的角度.	正交变换不改变角度.

置换矩阵

22.14	$\mathbf{P} = (p_{ij})_{n\times n}$ 是一置换矩阵，如果 $\mathbf{P}$ 的每一行和每一列中都有一个元素等于 1 而其他元素等于 0.	置换矩阵的定义.
22.15	$\mathbf{P}$ 是一置换矩阵 $\Rightarrow \mathbf{P}$ 是非奇异且正交的.	置换矩阵的性质.

非负矩阵

22.16	$\mathbf{A} = (a_{ij})_{m\times n} \geqslant \mathbf{0} \Leftrightarrow a_{ij} \geqslant 0$ 对于所有 i, j 成立. $\mathbf{A} = (a_{ij})_{m\times n} > \mathbf{0} \Leftrightarrow a_{ij} > 0$ 对于所有 i, j 成立.	非负和正矩阵的定义.
22.17	如果 $\mathbf{A} = (a_{ij})_{n\times n} \geqslant \mathbf{0}$，$\mathbf{A}$ 有至少一个非负特征值，最大的非负特征值称为 $\mathbf{A}$ 的 Frobenius 根，记为 $\lambda(\mathbf{A})$. $\mathbf{A}$ 有一对应于 $\lambda(\mathbf{A})$ 的非负特征向量.	非负矩阵的 Frobenius 根（或主导根的定义）
22.18	• μ 是 $\mathbf{A}$ 的一个特征值 $\Rightarrow \lvert\mu\rvert \leqslant \lambda(\mathbf{A})$ • $\mathbf{0} \leqslant \mathbf{A}_1 \leqslant \mathbf{A}_2 \Rightarrow \lambda(\mathbf{A}_1) \leqslant \lambda(\mathbf{A}_2)$ • $\rho > \lambda(\mathbf{A}) \Leftrightarrow (\rho\mathbf{I} - \mathbf{A})^{-1}$ 存在且 $\geqslant \mathbf{0}$ • $\min\limits_{1\leqslant j\leqslant n} \sum\limits_{i=1}^{n} a_{ij} \leqslant \lambda(\mathbf{A}) \leqslant \max\limits_{1\leqslant j\leqslant n} \sum\limits_{i=1}^{n} a_{ij}$	非负矩阵的性质. $\lambda(\mathbf{A})$是 A 的 Frobenius 根

22.19 矩阵 $\mathbf{A}=(a_{ij})_{n\times n}$ 是可分解的或可简约的，如果通过对换某些相同号码的行和列可以将矩阵 $\mathbf{A}$ 转化成

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{pmatrix}$$

其中 $\mathbf{A}_{11}$ 和 $\mathbf{A}_{22}$ 是子方阵.

可分解方阵的定义. 不可分解(简约)的矩阵称为非分解(非简约)型矩阵.

22.20 $\mathbf{A}=(a_{ij})_{n\times n}$ 是可分解的，当且仅当存在一置换矩阵 $\mathbf{P}$，使

$$\mathbf{P}^{-1}\mathbf{AP}=\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{pmatrix}$$

其中 $\mathbf{A}_{11}$ 和 $\mathbf{A}_{22}$ 是子方阵.

可分解矩阵的一个特征.

22.21 如果 $\mathbf{A}=(a_{ij})_{n\times n}\geqslant\mathbf{0}$ 是不可分解的，则

- Frobenius 根 $\lambda(\mathbf{A})>0$ 是特征方程的一个单根，并且存在一与之相伴的特征向量 $\boldsymbol{x}>0$.
- 如果 $\mathbf{A}\boldsymbol{x}=\mu\boldsymbol{x}$ 对于某一 $\mu\geqslant 0$ 和 $\boldsymbol{x}>\mathbf{0}$ 成立，则 $\mu=\lambda(\mathbf{A})$.

可分解矩阵的性质.

22.22 $\mathbf{A}=(a_{ij})_{n\times n}$ 有一主导对角线(d.d.)，如果存在正数 $d_1,\cdots,d_n$，使

$d_j\,|\,a_{jj}\,|>\sum\limits_{i\neq j}d_i\,|\,a_{ij}\,|$ 对 $j=1,\cdots,n$ 成立.

主导性对角矩阵的定义.

22.23 假设 $\mathbf{A}$ 是一主导对角矩阵，则

- $|\,\mathbf{A}\,|\neq 0$.
- 如果对角线元素均为正，则 $\mathbf{A}$ 的所有特征值均有正的实部.

主导对角矩阵的性质.

Leontief 方程组

22.24 如果 $\mathbf{A}=(a_{ij})_{n\times n}\geqslant 0$ 且 $\boldsymbol{c}\geqslant\mathbf{0}$，则

$$\mathbf{A}\boldsymbol{x}+\boldsymbol{c}=\boldsymbol{x}$$

称为一 Leontief 方程组.

Leontief 方程组的定义. $\boldsymbol{x}$ 和 $\boldsymbol{c}$ 是 $n\times 1$ 矩阵.

22.25 如果 $\sum_{i=1}^{n} a_{ij} < 1$ 对于 $j = 1, \cdots, n$ 成立，则 Leontief 方程组有一解 $\mathbf{x} \geqslant \mathbf{0}$.

Leontief 方程组有一非负解的充分条件.

22.26 对于每一 $\mathbf{c} \geqslant \mathbf{0}$, Leontief 方程组 $\mathbf{A}\mathbf{x} + \mathbf{c} = \mathbf{x}$ 有一解 $\mathbf{x} \geqslant \mathbf{0}$, 当且仅当以下条件有一个(从而全部)得以满足：

- 矩阵 $(\mathbf{I} - \mathbf{A})^{-1}$ 存在，是非负的，且等于 $\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \cdots$
- 当 $m \to \infty$ 时 $\mathbf{A}^m \to \mathbf{0}$
- $\mathbf{A}$ 的每一个特征值的模 < 1
- $\begin{vmatrix} 1 - a_{11} & -a_{12} & \cdots & -a_{1k} \\ -a_{21} & 1 - a_{22} & \cdots & -a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{k1} & -a_{k2} & \cdots & 1 - a_{kk} \end{vmatrix} > 0$

对 $k = 1, \cdots, n$ 成立.

Leontief 方程组有一非负解的充分必要条件. 最后的条件是 *Hawkins-Simon* 条件.

22.27 如果 $0 \leqslant a_{ii} < 1$ 对于 $i = 1, \cdots, n$ 成立，且 $a_{ij} \geqslant 0$ 对于所有 $i \neq j$ 成立，则方程组 $\mathbf{A}\mathbf{x} + \mathbf{c} = \mathbf{x}$ 对于每一 $\mathbf{c} \geqslant \mathbf{0}$ 有一解 $\mathbf{x} \geqslant \mathbf{0}$, 当且仅当 $\mathbf{I} - \mathbf{A}$ 有一主导对角线.

Leontief 方程组有非负解的充分必要条件.

参 考 文 献

关于矩阵的结论，参考 Gantmacher (1959)或 Horn & Johnson (1985). 关于 Leontief 方程组，参考 Nikaido (1970)和 Takayama (1985).

23 Kronecker 乘积和 vec 运算　向量和矩阵的微分

23.1 $$\mathbf{A}\otimes\mathbf{B}=\begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B}\\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2n}\mathbf{B}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mn}\mathbf{B}\end{pmatrix}$$

$\mathbf{A}=(a_{ij})_{m\times n}$ 和 $\mathbf{B}=(b_{ij})_{p\times q}$ 的 *Kronecker* 乘积. $\mathbf{A}\otimes\mathbf{B}$ 是 $mp\times nq$ 矩阵. Kronecker 乘积一般不能改变顺序, $\mathbf{A}\otimes\mathbf{B}\neq\mathbf{B}\otimes\mathbf{A}$.

23.2 $$\begin{pmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\otimes\begin{pmatrix} b_{11} & b_{12}\\ b_{21} & b_{22}\end{pmatrix}=$$
$$\begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12}\\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22}\\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12}\\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22}\end{pmatrix}$$

(23.1) 的一个特例.

23.3 $$\mathbf{A}\otimes\mathbf{B}\otimes\mathbf{C}=(\mathbf{A}\otimes\mathbf{B})\otimes\mathbf{C}=\mathbf{A}\otimes(\mathbf{B}\otimes\mathbf{C})$$

一般都成立.

23.4 $$(\mathbf{A}+\mathbf{B})\otimes(\mathbf{C}+\mathbf{D})=\mathbf{A}\otimes\mathbf{C}+\mathbf{A}\otimes\mathbf{D}+\mathbf{B}\otimes\mathbf{C}+\mathbf{B}\otimes\mathbf{D}$$

当 $\mathbf{A}+\mathbf{B}$ 和 $\mathbf{C}+\mathbf{D}$ 有定义时成立.

23.5 $$(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D})=\mathbf{AC}\otimes\mathbf{BD}$$

当 $\mathbf{AC}$ 和 $\mathbf{BD}$ 有定义时成立.

23.6 $$(\mathbf{A}\otimes\mathbf{B})'=\mathbf{A}'\otimes\mathbf{B}'$$

Kronecker 乘积的转置法则.

23.7 $$(\mathbf{A}\otimes\mathbf{B})^{-1}=\mathbf{A}^{-1}\otimes\mathbf{B}^{-1}$$

当 $\mathbf{A}^{-1}$ 和 $\mathbf{B}^{-1}$ 存在时成立.

23.8 $$\mathrm{tr}(\mathbf{A}\otimes\mathbf{B})=\mathrm{tr}(\mathbf{A})\mathrm{tr}(\mathbf{B})$$

$\mathbf{A}$ 和 $\mathbf{B}$ 是方阵,不一定有相同的阶数.

23.9 $\alpha \otimes \mathbf{A} = \alpha\mathbf{A} = \mathbf{A}\alpha = \mathbf{A} \otimes \alpha$

α 是一 1×1 数量矩阵.

23.10 如果 $\lambda_1, \cdots, \lambda_n$ 是 $\mathbf{A}$ 的特征值,且如果 $\mu_1, \cdots, \mu_p$ 是 $\mathbf{B}$ 的特征值,则 $\mathbf{A} \otimes \mathbf{B}$ 的 np 个特征值是 $\lambda_i\mu_j$, $i = 1, \cdots, n$, $j = 1, \cdots, p$.

$\mathbf{A} \otimes \mathbf{B}$ 的特征值,其中 $\mathbf{A}$ 是 $n \times n$ 矩阵而 $\mathbf{B}$ 是 $p \times p$ 矩阵.

23.11 如果 $\mathbf{x}$ 是 $\mathbf{A}$ 的一个特征向量,而 $\mathbf{y}$ 是 $\mathbf{B}$ 的一个特征向量,则 $\mathbf{x} \otimes \mathbf{y}$ 是 $\mathbf{A} \otimes \mathbf{B}$ 的一个特征向量.

注意: $\mathbf{A} \otimes \mathbf{B}$ 的特征向量不一定是 $\mathbf{A}$ 和 $\mathbf{B}$ 的特征向量的 Kronecker 乘积.

23.12 如果 $\mathbf{A}$ 和 $\mathbf{B}$ 是(半)正定的,则 $\mathbf{A} \otimes \mathbf{B}$ 是(半)正定的.

由(23.10)得出.

23.13 $|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{A}|^p \cdot |\mathbf{B}|^n$

$\mathbf{A}$ 是 $n \times n$ 矩阵, $\mathbf{B}$ 是 $p \times p$ 矩阵.

23.14 $r(\mathbf{A} \otimes \mathbf{B}) = r(\mathbf{A})r(\mathbf{B})$

Kronecker 乘积的秩.

23.15 如果 $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n)_{m \times n}$, 则

$$\operatorname{vec}(\mathbf{A}) = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}_{mn \times 1}$$

vec($\mathbf{A}$)是 $\mathbf{A}$ 的列竖叠在一起.

23.16 $$\operatorname{vec}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{12} \\ a_{22} \end{pmatrix}$$

(23.15)的特例.

23.17 $\operatorname{vec}(\mathbf{A} + \mathbf{B}) = \operatorname{vec}(\mathbf{A}) + \operatorname{vec}(\mathbf{B})$

当 $\mathbf{A} + \mathbf{B}$ 有定义时成立.

23.18 $\operatorname{vec}(\mathbf{ABC}) = [\mathbf{C}' \otimes \mathbf{A}]\operatorname{vec}(\mathbf{B})$

当乘积 $\mathbf{ABC}$ 有定义时成立.

23.19 $\operatorname{tr}(\mathbf{AB}) = (\operatorname{vec}(\mathbf{A}'))'\operatorname{vec}(\mathbf{B}) = (\operatorname{vec}(\mathbf{B}'))'\operatorname{vec}(\mathbf{A})$

当运算有定义时成立.

向量和矩阵的微分

23.20 如果 $y = f(x_1, \cdots, x_n) = f(\boldsymbol{x})$，则

$$\frac{\partial y}{\partial \boldsymbol{x}} = \left(\frac{\partial y}{\partial x_1}, \cdots, \frac{\partial y}{\partial x_n}\right)$$

$y = f(\boldsymbol{x})$ 的梯度是一行向量(以向量为变量的数量函数的导数). 梯度的另一种表示是 $\nabla f(\boldsymbol{x})$，参见(4.26).

23.21 $$\begin{array}{l} y_1 = f_1(x_1, \cdots, x_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots \quad \Leftrightarrow \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}) \\ y_m = f_m(x_1, \cdots, x_n) \end{array}$$

从 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 的变换 $\boldsymbol{f}$，我们设 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 为列向量.

23.22 $$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} = \begin{pmatrix} \dfrac{\partial y_1(\boldsymbol{x})}{\partial x_1} & \cdots & \dfrac{\partial y_1(\boldsymbol{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial y_m(\boldsymbol{x})}{\partial x_1} & \cdots & \dfrac{\partial y_m(\boldsymbol{x})}{\partial x_n} \end{pmatrix}$$

变换(23.21)的雅各比矩阵(向量函数对于向量变量的导数).

23.23 $$\frac{\partial^2 y}{\partial \boldsymbol{x} \partial \boldsymbol{x}'} = \frac{\partial}{\partial \boldsymbol{x}} \operatorname{vec}\left[\left(\frac{\partial y}{\partial \boldsymbol{x}}\right)'\right]$$

对于 vec 运算，参见(23.15).

23.24 $$\frac{\partial \mathbf{A}(\boldsymbol{r})}{\partial \boldsymbol{r}} = \frac{\partial}{\partial \boldsymbol{r}} \operatorname{vec}(\mathbf{A}(\boldsymbol{r}))$$

矩阵对向量的导数的一般定义.

23.25 $$\frac{\partial^2 y}{\partial \boldsymbol{x} \partial \boldsymbol{x}'} = \begin{pmatrix} \dfrac{\partial^2 y}{\partial x_1^2} & \cdots & \dfrac{\partial^2 y}{\partial x_n \partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 y}{\partial x_1 \partial x_n} & \cdots & \dfrac{\partial^2 y}{\partial x_n^2} \end{pmatrix}$$

(23.23)的特例. ($\partial^2 y / \partial \boldsymbol{x} \partial \boldsymbol{x}'$ 是(13.24)定义的 Hessian 矩阵.)

23.26 $$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{a}' \cdot \boldsymbol{x}) = \boldsymbol{a}'$$

$\boldsymbol{a}$ 和 $\boldsymbol{x}$ 是 $n \times 1$ 向量.

23.27 $$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}' \mathbf{A} \boldsymbol{x}) = \boldsymbol{x}'(\mathbf{A} + \mathbf{A}')$$

$$\frac{\partial^2}{\partial \boldsymbol{x} \partial \boldsymbol{x}'}(\boldsymbol{x}' \mathbf{A} \boldsymbol{x}) = \mathbf{A} + \mathbf{A}'$$

二次型的微分. $\mathbf{A}$ 是 $n \times n$ 矩阵，$\boldsymbol{x}$ 是 $n \times 1$ 矩阵.

23.28 $$\frac{\partial}{\partial \boldsymbol{x}}(\mathbf{A} \boldsymbol{x}) = \mathbf{A}$$

$\mathbf{A}$ 是 $m \times n$ 矩阵，$\boldsymbol{x}$ 是 $n \times 1$ 矩阵.

23.29 如果 $\mathbf{y} = \mathbf{A}(\mathbf{r})\boldsymbol{x}(\mathbf{r})$, 则

$$\frac{\partial \mathbf{y}}{\partial \mathbf{r}} = (\boldsymbol{x}' \otimes \mathbf{I}_m)\frac{\partial \mathbf{A}}{\partial \mathbf{r}} + \mathbf{A}\frac{\partial \boldsymbol{x}}{\partial \mathbf{r}}$$

$\mathbf{A}(\mathbf{r})$是 $m \times n$ 矩阵, $\boldsymbol{x}(\mathbf{r})$是 $n \times 1$ 矩阵而 $\mathbf{r}$ 是$k \times 1$ 矩阵.

23.30 如果 $y = f(\mathbf{A})$, 则

$$\frac{\partial y}{\partial \mathbf{A}} = \begin{pmatrix} \frac{\partial y}{\partial a_{11}} & \cdots & \frac{\partial y}{\partial a_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial a_{m1}} & \cdots & \frac{\partial y}{\partial a_{mn}} \end{pmatrix}$$

以 $m \times n$ 矩阵 $\mathbf{A} = (a_{ij})$ 为变量的数量函数的导数的定义.

23.31 $$\frac{\partial |\mathbf{A}|}{\partial \mathbf{A}} = (A_{ij}) = |\mathbf{A}|(\mathbf{A}')^{-1}$$

$\mathbf{A}$是 $n \times n$ 矩阵. (A_{ij})是矩阵 $\mathbf{A}$ 的代数余子式. (参见(19.16))最后一个等号当 $\mathbf{A}$ 可逆时成立.

23.32 $$\frac{\partial \mathrm{tr}(\mathbf{A})}{\partial \mathbf{A}} = \mathbf{I}_n, \quad \frac{\partial \mathrm{tr}(\mathbf{A}'\mathbf{A})}{\partial \mathbf{A}} = 2\mathbf{A}$$

$\mathbf{A}$是 $n \times n$ 矩阵. $\mathrm{tr}(\mathbf{A})$是 $\mathbf{A}$ 的迹.

23.33 $$\frac{\partial a^{ij}}{\partial a_{hk}} = -a^{ih}a^{kj};\ i, j, h, k = 1, \cdots, n$$

a^{ij} 是 $\mathbf{A}^{-1}$ 的第(i, j)个元素.

参 考 文 献

以上这些定义在许多经济学教材中都能找到,参考 Dhrymes (1978). Magnus & Neudecker (1988)和 Lütkepohl (1996)发展了一些更一致的记法,且有更多的公式和结论.

24 比较静态

24.1 $$\begin{aligned}E_1(\boldsymbol{p}, \boldsymbol{a}) &= S_1(\boldsymbol{p}, \boldsymbol{a}) - D_1(\boldsymbol{p}, \boldsymbol{a})\\ E_2(\boldsymbol{p}, \boldsymbol{a}) &= S_2(\boldsymbol{p}, \boldsymbol{a}) - D_2(\boldsymbol{p}, \boldsymbol{a})\\ &\cdots\cdots\\ E_n(\boldsymbol{p}, \boldsymbol{a}) &= S_n(\boldsymbol{p}, \boldsymbol{a}) - D_n(\boldsymbol{p}, \boldsymbol{a})\end{aligned}$$

$S_i(\boldsymbol{p}, \boldsymbol{a})$是对商品 i 的供给，$D_i(\boldsymbol{p}, \boldsymbol{a})$是对商品 i 的需求，$E_i(\boldsymbol{p}, \boldsymbol{a})$是超额供给，$\boldsymbol{p} = (p_1, \cdots, p_n)$ 是价格向量，$\boldsymbol{a} = (a_1, \cdots, a_k)$ 是外生变量的向量.

24.2 $$E_1(\boldsymbol{p}, \boldsymbol{a}) = 0, E_2(\boldsymbol{p}, \boldsymbol{a}) = 0, \cdots, E_n(\boldsymbol{p}, \boldsymbol{a}) = 0$$

均衡的条件.

24.3 $$\begin{aligned}E_1(p_1, p_2, a_1, \cdots, a_k) &= 0\\ E_2(p_1, p_2, a_1, \cdots, a_k) &= 0\end{aligned}$$

两种商品时的均衡条件.

24.4 $$\frac{\partial p_1}{\partial a_j} = \frac{\dfrac{\partial E_1}{\partial p_2}\dfrac{\partial E_2}{\partial a_j} - \dfrac{\partial E_2}{\partial p_2}\dfrac{\partial E_1}{\partial a_j}}{\dfrac{\partial E_1}{\partial p_1}\dfrac{\partial E_2}{\partial p_2} - \dfrac{\partial E_1}{\partial p_2}\dfrac{\partial E_2}{\partial p_1}}$$

$$\frac{\partial p_2}{\partial a_j} = \frac{\dfrac{\partial E_2}{\partial p_1}\dfrac{\partial E_1}{\partial a_j} - \dfrac{\partial E_1}{\partial p_1}\dfrac{\partial E_2}{\partial a_j}}{\dfrac{\partial E_1}{\partial p_1}\dfrac{\partial E_2}{\partial p_2} - \dfrac{\partial E_1}{\partial p_2}\dfrac{\partial E_2}{\partial p_1}}$$

两种商品情况下的比较静态结论，$j = 1, \cdots, k$.

24.5 $$\begin{pmatrix}\dfrac{\partial p_1}{\partial a_j}\\ \vdots\\ \dfrac{\partial p_n}{\partial a_j}\end{pmatrix} = -\begin{pmatrix}\dfrac{\partial E_1}{\partial p_1} & \cdots & \dfrac{\partial E_1}{\partial p_n}\\ \vdots & \ddots & \vdots\\ \dfrac{\partial E_n}{\partial p_1} & \cdots & \dfrac{\partial E_n}{\partial p_n}\end{pmatrix}^{-1}\begin{pmatrix}\dfrac{\partial E_1}{\partial a_j}\\ \vdots\\ \dfrac{\partial E_n}{\partial a_j}\end{pmatrix}$$

n 种商品情况下的比较静态结论，$j = 1, \cdots, k$. 关于方阵的逆矩阵，参见(19.16).

24.6 考虑问题

$$\max f(\mathbf{x}, \mathbf{a}) \text{ s.t. } g(\mathbf{x}, \mathbf{a}) = 0$$

其中 f 和 g 是 C^1 函数，且设 $\mathcal{L}$ 为相关的拉格朗日函数，及拉格朗日乘数 λ. 如果 $x_i^* = x_i^*(\mathbf{a})$, $i = 1, \cdots, n$ 是问题的解，则对于 $i, j = 1, \cdots, m$

$$\sum_{k=1}^{n} \mathcal{L}''_{a_i x_k} \frac{\partial x_k^*}{\partial a_j} + g'_{a_i} \frac{\partial \lambda}{\partial a_j} = \sum_{k=1}^{n} \mathcal{L}''_{a_j x_k} \frac{\partial x_k^*}{\partial a_i} + g'_{a_j} \frac{\partial \lambda}{\partial a_i}$$

互反关系. $\mathbf{x} = (x_1, \cdots, x_n)$ 是决策变量，$\mathbf{a} = (a_1, \cdots, a_m)$ 是参数. 对于这些关系的系统化应用，参见 Silberberg (1990).

单调比较静态

24.7 定义在 $\mathbf{R}^m$ 的一个子格 Z 上的函数 $F: Z \to \mathbf{R}$ 称为是上模，如果

$$F(\mathbf{z}) + F(\mathbf{z}') \leqslant F(\mathbf{z} \wedge \mathbf{z}') + F(\mathbf{z} \vee \mathbf{z}')$$

对于所有 $\mathbf{z}$ 和 $\mathbf{z}' \in Z$ 成立. 如果 $\mathbf{z}$ 和 $\mathbf{z}'$ 在预次序关系 $\leqslant$ 下不可比时，不等号严格成立，则 F 称为严格上模.

(严格)上模的定义. 对于子格和格运算 $\wedge$ 和 $\vee$ 的定义，参见(6.28)和(6.29).

24.8 设 S 和 P 分别为 $\mathbf{R}^n$ 和 $\mathbf{R}^l$ 的子格. 函数 $f: S \times P \to \mathbf{R}$ 称为在 $(\mathbf{x}, \mathbf{p})$ 满足递增差别，如果

$$\mathbf{x} \geqslant \mathbf{x}' \text{ 和 } \mathbf{p} \geqslant \mathbf{p}' \Rightarrow$$

$$f(\mathbf{x}, \mathbf{p}) - f(\mathbf{x}', \mathbf{p}) \geqslant f(\mathbf{x}, \mathbf{p}') - f(\mathbf{x}', \mathbf{p}')$$

对于所有对 $(\mathbf{x}, \mathbf{p})$ 和 $(\mathbf{x}', \mathbf{p}') \in S \times P$ 成立. 如果当 $\mathbf{x} > \mathbf{x}'$ 及 $\mathbf{p} > \mathbf{p}'$ 时不等号严格成立，则 f 称为在 $(\mathbf{x}, \mathbf{p})$ 满足严格递增差别.

(严格)递增差别的定义(f 在较大的选择 $\mathbf{x}$ 和较小的选择 $\mathbf{x}'$ 之间的取值差别 $f(\mathbf{x}, \mathbf{p}) - f(\mathbf{x}', \mathbf{p})$ 是参数 $\mathbf{p}$ 的一个(严格)递增函数).

24.9 设 S 和 P 分别为 $\mathbf{R}^n$ 和 $\mathbf{R}^l$ 的子格. 如果 $f: S \times P \to \mathbf{R}$ 是在 $(\mathbf{x}, \mathbf{p})$ 的上模，则

- f 对于给定的 $\mathbf{p}$ 是在 $\mathbf{x}$ 的上模，即对任意给定的 $\mathbf{p}' \in P$，和任意 $\mathbf{x}$ 及 $\mathbf{x}' \in S$，我们有 $f(\mathbf{x}, \mathbf{p}) + f(\mathbf{x}', \mathbf{p}) \leqslant f(\mathbf{x} \wedge \mathbf{x}', \mathbf{p}) + f(\mathbf{x} \vee \mathbf{x}', \mathbf{p})$
- f 满足在 $(\mathbf{x}, \mathbf{p})$ 的递增差别

重要的现象. 注意 $S \times P$ 是 $\mathbf{R}^n \times \mathbf{R}^l = \mathbf{R}^{n+l}$ 的一个子格.

24.10 设 X 为 $\mathbf{R}^m$ 的一个开子格.某一 C^2 函数 $F: X \to \mathbf{R}$ 在 X 是上模,当且仅当对于所有 $\mathbf{x} \in X$,

$$\frac{\partial^2 F}{\partial x_i \partial x_j}(\mathbf{x}) \geqslant 0,\ i,\ j = 1,\ \cdots,\ m,\ i \neq j$$

24.11 假设问题

$$\max F(x,\ p) \quad 受约束\ x \in S \subset \mathbf{R}$$

对每一 $p \in P \subset \mathbf{R}$ 有至少一个解.进一步假设 F 在$(x,\ p)$满足严格递增差别,则最优选择 $x^*(p)$是随参数 p 递增的.

一个特殊的结论,不能扩展到 $S \subset \mathbf{R}^n$, $n \geqslant 2$.

24.12 假设在(24.11)中

$$F(x,\ p) = pf(x) - C(x)$$

其中 S 是紧的且 f 和 C 是连续的,则 $\partial^2 F / \partial x \partial p = f'(x)$,根据(24.10),$F$ 是上模当且仅当 $f(x)$是递增的.因此 $f(x)$递增充分保证最优选择 $x^*(p)$是随 p 递增的.

(24.10)一个重要的推论.

24.13 假设 S 是 $\mathbf{R}^n$ 的一个紧子格,而 P 是 $\mathbf{R}^l$ 的一个子格,且 $f: S \times P \to \mathbf{R}$ 对每一给定的 $\mathbf{p}$ 是 S 的连续函数.假设 f 在$(\mathbf{x},\ \mathbf{p})$满足递增差别,且对每一给定 $\mathbf{p}$ 在 $\mathbf{x}$ 是上模.设对应由 P 至 S 的 Γ 定义为

$$\Gamma(\mathbf{p}) = \operatorname{argmax}\{f(\mathbf{x},\ \mathbf{p}) : \mathbf{x} \in S\}$$

- 对每一 $\mathbf{p} \in \Omega$, $\Gamma(\mathbf{p})$ 是 $\mathbf{R}^n$ 的非空紧子格,且有一最大的元素,记为 $\mathbf{x}^*(\mathbf{p})$
- $\mathbf{p}_1 > \mathbf{p}_2 \Rightarrow \mathbf{x}^*(\mathbf{p}_1) \geqslant \mathbf{x}^*(\mathbf{p}_2)$
- 如果 f 在$(\mathbf{x},\ \mathbf{p})$满足严格递增差别,则 $\mathbf{x}_1 \geqslant \mathbf{x}_2$ 对任意 $x_1 \in \Gamma(\mathbf{p}_1)$和 $x_2 \in \Gamma(\mathbf{p}_2)$成立(当 $\mathbf{p}_1 > \mathbf{p}_2$).

一个主要结论.对于给定 $\mathbf{p}$, $\operatorname{argmax}\{f(\mathbf{x},\ \mathbf{p}) : \mathbf{x} \in S\}$ 是使 $f(\mathbf{x},\ \mathbf{p})$取得最大值的所有 S 中的点 $\mathbf{x}$ 的集合.

参 考 文 献

关于比较静态参考 Varian (1996)或 Silberberg (1990),关于单调比较静态,参考 Sundaram (1996)及 Topkis (1998).

25 成本和利润函数的性质

25.1 $$C(\mathbf{w}, y) = \min_{\mathbf{x}} \sum_{i=1}^{n} w_i x_i \quad 当 \quad f(\mathbf{x}) = y$$

成本最小化,单一产出. f 是生产函数, $\mathbf{w} = (w_1, \cdots, w_n)$ 是要素价格, y 是产出,而 $\mathbf{x} = (x_1, \cdots, x_n)$ 是要素投入, $C(\mathbf{w}, y)$ 是成本函数.

25.2 $$C(\mathbf{w}, y) = \begin{cases} 当要素价格是\ \mathbf{w} = (w_1, \cdots, w_n) \\ 时生产\ y\ 单位商品时的最小成本. \end{cases}$$

成本函数.

25.3
- $C(\mathbf{w}, y)$随 w_i 而增加.
- $C(\mathbf{w}, y)$对 $\mathbf{w}$ 是 1 次齐次的.
- $C(\mathbf{w}, y)$对 $\mathbf{w}$ 是凹的.
- $C(\mathbf{w}, y)$当 $\mathbf{w} > 0$ 时,对 $\mathbf{w}$ 是连续的.

成本函数的性质.

25.4 $$x_i^*(\mathbf{w}, y) = \begin{cases} 使成本最小的第\ i\ 个要素的选择 \\ 是要素价格\ \mathbf{w}\ 和生产水平\ y\ 的函数. \end{cases}$$

条件要素需求函数. $\mathbf{x}^*$ 是解出问题(25.1)的向量 $\mathbf{x}^*$.

25.5
- $x_i^*(\mathbf{w}, y)$随 w_i 递减.
- $x_i^*(\mathbf{w}, y)$对 $\mathbf{w}$ 的齐次度为 0.

条件要素需求函数的性质.

25.6 $$\frac{\partial C(\mathbf{w}, y)}{\partial w_i} = x_i^*(\mathbf{w}, y),\ i = 1, \cdots, n$$

Shephard 引理.

25.7 $$\left(\frac{\partial^2 C(\mathbf{w}, y)}{\partial w_i \partial w_j}\right)_{(n \times n)} = \left(\frac{\partial x_i^*(\mathbf{w}, y)}{\partial w_j}\right)_{(n \times n)}$$

是对称和半负定的.

替代矩阵的性质.

25.8	$\pi(p, w) = \max_{\mathbf{x}}\left(pf(\mathbf{x}) - \sum_{i=1}^{n} w_i x_i\right)$	厂商的利润最大化. p 是产出的价格, $\pi(p, w)$ 是利润函数.
25.9	$\pi(p, w) = \Big\{$ 最大利润是要素价格 w 和产出价格 p 的函数.	利润函数.
25.10	$\pi(p, w) = \max_{y}(py - C(w, y))$	由成本和收益表述的利润函数.
25.11	• $\pi(p, w)$ 随 p 而递增. • $\pi(p, w)$ 对 (p, w) 的齐次度为 1. • $\pi(p, w)$ 在 (p, w) 是凸的. • $\pi(p, w)$ 当 $w > 0$, $p > 0$ 时对 (p, w) 是连续的.	利润函数的性质.
25.12	$x_i(p, w) = \Big\{$ 使利润最大的第 i 个要素的选择是产出价格 p 和要素价格 w 的函数.	要素需求函数. $\mathbf{x}(p, w)$ 是解出问题(25.8)的向量 $\mathbf{x}$.
25.13	• $x_i(p, w)$ 随 w_i 递减. • $x_i(p, w)$ 对 (p, w) 的齐次度为 0. 交叉价格效应是对称的 $\frac{\partial x_i(p, w)}{\partial w_j} = \frac{\partial x_j(p, w)}{\partial w_i}, \ i, j = 1, \cdots, n$	要素需求函数的性质.
25.14	$y(p, w) = \Big\{$ 最大利润的产出是产出价格 p 和要素价格 w 的函数.	供给函数 $y(p, w) = f(\mathbf{x}(p, w))$ 是解出问题(25.10)的 y.
25.15	• $y(p, w)$ 随 p 递增. • $y(p, w)$ 对 (p, w) 的齐次度为 0.	供给函数的性质.
25.16	$\frac{\partial \pi(p, w)}{\partial p} = y(p, w)$ $\frac{\partial \pi(p, w)}{\partial w_i} = -x_i(p, w), \ i = 1, \cdots, n$	*Hotelling* 引理.

25.17	$$\frac{\partial x_j(p,w)}{\partial u_k}=\frac{\partial x_j(w,y)}{\partial u_k}+\frac{\dfrac{\partial x_j(p,w)}{\partial p}\dfrac{\partial y(p,w)}{\partial u_k}}{\dfrac{\partial y(p,w)}{\partial p}}$$	*Puu* 方程，$j,k=1,\cdots,n$，反映替代和要素价格上升的规模效应.

生产理论中的替代弹性

25.18	$$\sigma_{yx}=\mathrm{El}_{R_{yx}}\left(\frac{y}{x}\right)=-\frac{\partial\ln\left(\dfrac{y}{x}\right)}{\partial\ln\left(\dfrac{p_2}{p_1}\right)},\ f(x,y)=c$$	y 和 x 之间的替代弹性，假设市场是自由竞争的.（也参见(5.18)）
25.19	$$\sigma_{ij}=-\frac{\partial\ln\left(\dfrac{C'_i(w,y)}{C'_j(w,y)}\right)}{\partial\ln\left(\dfrac{w_i}{w_j}\right)},\ i\neq j$$ y，C 和 w_k（对于 $k\neq i,j$）是连续的.	要素 i 和要素 j 之间的影子弹性.
25.20	$$\sigma_{ij}=\frac{-\dfrac{C''_{ii}}{(C'_i)^2}+\dfrac{2C''_{ij}}{C'_iC'_j}-\dfrac{C''_{jj}}{(C'_j)^2}}{\dfrac{1}{w_iC'_i}+\dfrac{1}{w_jC'_j}},\ i\neq j$$	(25.19)的另一种形式.
25.21	$$A_{ij}(w,y)=\frac{C(w,y)C''_{ij}(w,y)}{C'_i(w,y)C'_j(w,y)},\ i\neq j$$	*Allen-Uzawa* 替代弹性.
25.22	$$A_{ij}(w,y)=\frac{\epsilon_{ij}(w,y)}{S_j(w,y)},\ i\neq j$$	在此 $\epsilon_{ij}(w,y)$ 是需求的（固定产出）交叉价格弹性，而 $S_j(w,y)=p_jC_j(w,y)/C(w,y)$ 是第 j 个投入在总成本中的份额.
25.23	$$M_{ij}(w,y)=\frac{w_iC''_{ij}(w,y)}{C'_j(w,y)}-\frac{w_iC''_{ii}(w,y)}{C'_i(w,y)}=\epsilon_{ij}(w,y)-\epsilon_{ii}(w,y),\ i\neq j$$	*Morishima* 替代弹性.

25.24 如果 $n > 2$，则 $M_{ij}(\mathbf{w}, y) = M_{ij}(\mathbf{w}, y)$ 对于所有 $i \neq j$ 成立，当且仅当所有 $M_{ij}(\mathbf{w}, y)$ 等于相同的常数.

Morishima 替代弹性的对称性.

特殊的函数形式及其性质

柯柏-道格拉斯函数

25.25 $y = Ax_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}$

定义在 $x_i > 0$，$i = 1, \cdots, n$ 的柯柏-道格拉斯生产函数，$a_1, \cdots, a_n$ 和 A 是正的常数.

25.26 在(25.25)中的柯柏-道格拉斯函数是：

(a) 齐次度为 $a_1 + \cdots + a_n$，

(b) 对于所有 $a_1, \cdots, a_n$ 是拟凹的，

(c) 当 $a_1 + \cdots + a_n \leqslant 1$ 时是凹的，

(d) 当 $a_1 + \cdots + a_n < 1$ 时是严格凹的.

柯柏-道格拉斯函数的性质.
($a_1, \cdots, a_n$，和 A 是正的常数)

25.27 $x_k(\mathbf{w}, y) = A^{-\frac{1}{s}}\left(\frac{a_k}{w_k}\right)\left(\frac{w_1}{a_1}\right)^{\frac{a_1}{s}}\cdots\left(\frac{w_n}{a_n}\right)^{\frac{a_n}{s}}y^{\frac{1}{s}}$

条件要素需求函数，$s = a_1 + \cdots + a_n$.

25.28 $C(\mathbf{w}, y) = sA^{-\frac{1}{s}}\left(\frac{w_1}{a_1}\right)^{\frac{a_1}{s}}\cdots\left(\frac{w_n}{a_n}\right)^{\frac{a_n}{s}}y^{\frac{1}{s}}$

成本函数，$s = a_1 + \cdots + a_n$.

25.29 $\dfrac{w_k x_k}{C(\mathbf{w}, y)} = \dfrac{a_k}{a_1 + \cdots + a_n}$

总成本中要素的份额.

25.30 $x_k(p, \mathbf{w}) = \dfrac{a_k}{w_k}(pA)^{\frac{1}{1-s}}\left(\frac{w_1}{a_1}\right)^{\frac{a_1}{s-1}}\cdots\left(\frac{w_n}{a_n}\right)^{\frac{a_n}{s-1}}$

要素需求函数，$s = a_1 + \cdots + a_n < 1$.

25.31 $\pi(p, \mathbf{w}) = (1 - s)(pA)^{\frac{1}{1-s}}\prod_{i=1}^{n}\left(\frac{w_i}{a_i}\right)^{-\frac{a_i}{1-s}}$

利润函数，$s = a_1 + \cdots + a_n < 1$（如果 $s = a_1 + \cdots + a_n \geqslant 1$，则是递增规模效应，利润最大化问题无解）.

CES(常数替代弹性)函数

25.32 $y=((a_1x_1)^e+(a_2x_2)^e+\cdots+(a_nx_n)^e)^{s/e}$

CES 函数，定义在 $x_i>0$, $i=1,\cdots,n$, $a_1,\cdots,a_n$ 是正的，$e\neq 0$, $e<1$, $s>0$.

25.33 在(25.32)中的 CES 函数是：
(a) 齐次度为 s,
(b) 对 $e\leqslant 1$, $s\in(0,1]$ 是凹的，
(c) 对 $e\leqslant 1$, $s>1$ 是拟凹的，
(d) 对 $e\geqslant 1$, $s\in(0,1]$ 是拟凸的，
(e) 对 $e\geqslant 1$, $s>1$ 是凸的.

CES 函数的性质.

25.34 $x_k^*(w,y)=yw_k^{-1}a_k^{-r}\left[\left(\frac{w_1}{a_1}\right)^r+\cdots+\left(\frac{w_n}{a_n}\right)^r\right]^{-\frac{1}{e}}$

条件要素需求函数，$r=e/(e-1)$.

25.35 $C(w,y)=y^{\frac{1}{s}}\left[\left(\frac{w_1}{a_1}\right)^r+\cdots+\left(\frac{w_n}{a_n}\right)^r\right]^{\frac{1}{r}}$

$r=e/(e-1)$ 时的成本函数.

25.36 $$\frac{w_kx_k^*}{C(w,y)}=\frac{\left(\frac{w_k}{a_k}\right)^r}{\left(\frac{w_1}{a_1}\right)^r+\cdots+\left(\frac{w_n}{a_n}\right)^r}$$

总成本中的要素份额.

最小法则

25.37 $y=\min(a_1+b_1x_1,\cdots,a_n+b_nx_n)$

最小法则. 当 $a_1=\cdots=a_n=0$ 时，这是 *Leontief* 或固定系数函数.

25.38 $x_k^*(w,y)=\frac{y-a_k}{b_k}$, $k=1,\cdots,n$

条件要素需求函数.

25.39 $C(w,y)=\left(\frac{y-a_1}{b_1}\right)w_1+\cdots+\left(\frac{y-a_n}{b_n}\right)w_n$

成本函数.

Diewert(广义 Leontief)成本函数

25.40 $C(w,y)=y\sum_{i,j=1}^{n}b_{ij}\sqrt{w_iw_j}$ 和 $b_{ij}=b_{ji}$

Diewert 成本函数.

25.41 $x_k^*(w, y) = y\sum_{j=1}^{n} b_{kj}\sqrt{w_k / w_j}$

条件要素需求函数.

对数变换成本函数

25.42 $$\ln C(w, y) = a_0 + c_1 \ln y + \sum_{i=1}^{n} a_i \ln w_i + \frac{1}{2}\sum_{i,j=1}^{n} a_{ij} \ln w_i \ln w_j + \sum_{i=1}^{n} b_i \ln w_i \ln y$$

约束：$\sum_{i=1}^{n} a_i = 1$, $\sum_{i=1}^{n} b_i = 0$,

$\sum_{j=1}^{n} a_{ij} = \sum_{i=1}^{n} a_{ij} = 0$, $i = 1, \cdots, n$, $j = 1, \cdots, n$

对数变换成本函数. $a_{ij} = a_{ji}$ 对于所有 i 和 j 成立. 对系数的约束保证 $C(w, y)$的齐次度是1.

25.43 $$\frac{w_k x_k^*}{C(w, y)} = a_k + \sum_{j=1}^{n} a_{kj} \ln w_j + b_i \ln y$$

总成本中的要素份额.

参考文献

一种基本的参考文献是 Varian(1996),对于存在性和可微性的详细讨论,参考 Fuss & McFadden(1978),对于 Puu 方程(25.17)的讨论,参考 Johansen (1972),对于(25.18)—(25.24),参考 Blackorby and Russell(1989). 对于特殊函数形式,参考 Fuss & McFadden(1978).

26 消费者理论

26.1 一个在商品向量 $\mathbf{x} = (x_1, \cdots, x_n)$ 的集合 X 上的偏好关系$\succcurlyeq$是一在 X 上的二元关系，其意义为：

$\mathbf{x} \succcurlyeq \mathbf{y}$ 意指： $\mathbf{x}$ 至少与 $\mathbf{y}$ 一样好.

偏好关系的定义.对于二元关系，参见(1.16).

26.2
- $\mathbf{x} \sim \mathbf{y} \Leftrightarrow \mathbf{x} \succcurlyeq \mathbf{y}$ 且 $\mathbf{y} \succcurlyeq \mathbf{x}$
- $\mathbf{x} \succ \mathbf{y} \Leftrightarrow \mathbf{x} \succcurlyeq \mathbf{y}$ 但 $\mathbf{y} \succcurlyeq \mathbf{x}$ 不成立

由$\succcurlyeq$得出的在 X 中的关系. $\mathbf{x} \sim \mathbf{y}$ 读作“$\mathbf{x}$ 与 $\mathbf{y}$ 无差别”，而 $\mathbf{x} \succ \mathbf{y}$ 读作“$\mathbf{x}$(严格)好于 $\mathbf{y}$”.

26.3
- 一个函数 $u: X \to \mathbf{R}$ 是一代表偏好关系$\succcurlyeq$的效用函数. 如果

  $\mathbf{x} \succcurlyeq \mathbf{y} \Leftrightarrow u(\mathbf{x}) \geqslant u(\mathbf{y})$
- 对于任意单调递增函数 $f: \mathbf{R} \to \mathbf{R}$, $u^*(\mathbf{x}) = f(u(\mathbf{x}))$, 是代表与 $u(\cdot)$相同的偏好关系.

效用函数对于单调递增变换不变的性质，称为次序的. 基数的性质是指那些在单调递增变换中不保留下的性质.

26.4 设$\succcurlyeq$为一完整的，反射的，和可递的偏好关系，且又是连续的，意指对于所有 $\mathbf{x}^0 \in X$, 集合

$\{\mathbf{x}: \mathbf{x} \succcurlyeq \mathbf{x}^0\}$ 和 $\{\mathbf{x}: \mathbf{x}^0 \succcurlyeq \mathbf{x}\}$

都是闭的，则存在一连续的效用函数代表$\succcurlyeq$.

连续效用函数的存在性. 对于关系的性质，参见(1.16).

26.5 $\max_{x} u(x)$ 受约束 $\sum_{i=1}^{n} p_i x_i = m$

受预算约束的效用最大化. $x=(x_1,\cdots,x_n)$是一商品(数量的)向量, $p=(p_1,\cdots,p_n)$是价格向量, m 是收入, 而 u 是效用函数.

26.6 $v(p, m) = \max_{x}\{u(x): p \cdot x = m\}$

间接效用函数. $v(p, m)$作为价格向量 p 和收入 m 的函数的最大效用.

26.7
$$\begin{cases} v(p, m)\text{随 } p \text{ 递减.} \\ v(p, m)\text{随 } m \text{ 递增.} \\ v(p, m)\text{对}(p, m)\text{的齐次度为 } 0. \\ v(p, m)\text{对 } p \text{ 是拟凸的.} \\ v(p, m)\text{对}(p, m),\ p>0,\ m>0\text{ 是连续的.} \end{cases}$$

间接效用函数的性质.

26.8 $\omega = \dfrac{u_1'(x)}{p_1} = \cdots = \dfrac{u_n'(x)}{p_n}$

问题(26.5)的一阶条件, ω 是相应的拉格朗日乘数.

26.9 $\omega = \dfrac{\partial v(p, m)}{\partial m}$

ω 称为货币的边际效用.

26.10 $x_i(p, m) = \begin{cases} \text{对第 } i \text{ 个商品的} \\ \text{最优选择是} \\ \text{价格函数 } p \text{ 和收入 } m \text{ 的函数.} \end{cases}$

消费者需求函数, 或马歇尔需求函数, 由问题(26.5)得出.

26.11 $x(tp, tm) = x(p, m)$, t 是一正数量

需求函数的齐次度为 0.

26.12 $x_i(p, m) = -\dfrac{\dfrac{\partial v(p, m)}{\partial p_i}}{\dfrac{\partial v(p, m)}{\partial m}}$, $i = 1, \cdots, n$

Roy 恒等式.

26.13	$e(\boldsymbol{p}, u) = \min\limits_{x}\{\boldsymbol{p} \cdot \boldsymbol{x} : u(\boldsymbol{x}) \geqslant u\}$	支出函数. $e(\boldsymbol{p}, u)$是在价格为 $\boldsymbol{p}$ 时至少获得效用水平 u 的最小支出.
26.14	$\begin{cases} e(\boldsymbol{p}, u)\text{ 随 }\boldsymbol{p}\text{ 递增.} \\ e(\boldsymbol{p}, u)\text{ 对 }\boldsymbol{p}\text{ 的齐次度为 1.} \\ e(\boldsymbol{p}, u)\text{ 对 }\boldsymbol{p}\text{ 是凹的.} \\ e(\boldsymbol{p}, u)\text{ 对 }\boldsymbol{p}\text{ 当 }\boldsymbol{p} > \boldsymbol{0}\text{ 是连续的.} \end{cases}$	支出函数的性质.
26.15	$\boldsymbol{h}(\boldsymbol{p}, u) = \begin{cases} \text{在价格 }\boldsymbol{p}\text{ 时} \\ \text{获取效用水平 }u\text{ 所} \\ \text{必须的最小支出组合.} \end{cases}$	希克斯(或补充的)需求函数. $\boldsymbol{h}(\boldsymbol{p}, u)$解出问题 $\min\limits_{x}\{\boldsymbol{p} \cdot \boldsymbol{x} : u(\boldsymbol{x}) \geqslant u\}$ 的向量 $\boldsymbol{x}$.
26.16	$\dfrac{\partial e(\boldsymbol{p}, u)}{\partial p_i} = h_i(\boldsymbol{p}, u),\ i = 1, \cdots, n$	支出函数与希克斯需求函数的关系.
26.17	$\dfrac{\partial h_i(\boldsymbol{p}, u)}{\partial p_j} = \dfrac{\partial h_j(\boldsymbol{p}, u)}{\partial p_i},\ i, j = 1, \cdots, n$	Hicks 交叉偏导数的对称性(马歇尔交叉偏导数不一定是对称的.)
26.18	矩阵 $\left(\dfrac{\partial h_i(\boldsymbol{p}, u)}{\partial p_j}\right)_{n\times n}$ 是半负定性的.	由(26.16)和支出函数的凹性得出.
26.19	$e(\boldsymbol{p}, v(\boldsymbol{p}, m)) = m$: $\begin{cases} \text{达到效用 }v(\boldsymbol{p}, m)\text{的} \\ \text{最小支出是 }m. \end{cases}$	除了在非常特殊的情况下都成立的有用恒等式.
26.20	$v(\boldsymbol{p}, e(\boldsymbol{p}, u)) = u$: $\begin{cases} \text{由收入 }e(\boldsymbol{p}, u)\text{而定的} \\ \text{最大效用是 }u. \end{cases}$	
26.21	$x_i(\boldsymbol{p}, m) = h_i(\boldsymbol{p}, v(\boldsymbol{p}, m))$: $\begin{cases} \text{收入为 }m\text{ 时} \\ \text{的马歇尔需} \\ \text{求是效用为} \\ v(\boldsymbol{p}, m)t \\ \text{时的希克斯} \\ \text{需求} \end{cases}$	

26.22 $h_i(\mathbf{p}, u) = x_i(\mathbf{p}, e(\mathbf{p}, u))$: $\begin{cases} \text{效用为 } u \text{ 时} \\ \text{的希克斯需求} \\ \text{等于收入为} \\ e(\mathbf{p}, u)\text{时的} \\ \text{马歇尔需求.} \end{cases}$

26.23 • $e_{ij} = \mathrm{El}_{p_j} x_i = \dfrac{p_j}{x_i} \dfrac{\partial x_i}{\partial p_j}$ (*Cournot* 弹性)

• $E_i = \mathrm{El}_m x_i = \dfrac{m}{x_i} \dfrac{\partial x_i}{\partial m}$ (*Engel* 弹性)

• $S_{ij} = \mathrm{El}_{p_j} h_i = \dfrac{p_j}{x_i} \dfrac{\partial h_i}{\partial p_j}$ (*Slutsky* 弹性)

e_{ij}是需求对价格的弹性，E_i 是需求对收入的弹性，而 S_{ij} 是 Hicks 需求对价格的弹性.

26.24 • $\dfrac{\partial x_i(\mathbf{p}, m)}{\partial p_j} = \dfrac{\partial h_i(\mathbf{p}, u)}{\partial p_j} - x_j(\mathbf{p}, m) \dfrac{\partial x_i(\mathbf{p}, m)}{\partial m}$

• $S_{ij} = e_{ij} + a_j E_i,\ a_j = p_j x_j / m$

两种等价的 *Slutsky* 方程形式.

26.25 以下 $\frac{1}{2}n(n+1)+1$ 个对需求函数的偏导数的约束是线性独立的：

(a) $\displaystyle\sum_{i=1}^{n} p_i \frac{\partial x_i(\mathbf{p}, m)}{\partial m} = 1$

(b) $\displaystyle\sum_{j=1}^{n} p_j \frac{\partial x_i}{\partial p_j} + m \frac{\partial x_i}{\partial m} = 0,\ i = 1, \cdots, n$

(c) $\dfrac{\partial x_i}{\partial p_j} + x_j \dfrac{\partial x_i}{\partial m} = \dfrac{\partial x_j}{\partial p_i} + x_i \dfrac{\partial x_j}{\partial m}$,

$i = 1, \cdots, n-1,\ j = i+1, \cdots, n$

(a)是预算约束对 m 的微分，(b)是应用在消费者需求函数上(对齐次函数)的欧拉方程，(c)是 Slutsky 方程和(26.15)的推论.

26.26 $EV = e(\mathbf{p}^0, v(\mathbf{p}^1, m^1)) - e(\mathbf{p}^0, v(\mathbf{p}^0, m^0))$

EV 是在老(时期 0)价格下取得新(时期 1)效用所需的货币，与在老价格下取得老效用所需的货币之间的差额.

等价变量，$\mathbf{p}^0$, m^0 和 $\mathbf{p}^1$, m^1 分别是时期 0 和时期 1 的价格和收入，$e(\mathbf{p}^0, v(\mathbf{p}^0, m^0)) = m^0_0$.

26.27 $CV=e(\mathbf{p}^1,\ v(\mathbf{p}^1,\ m^1))-e(\mathbf{p}^1,\ v(\mathbf{p}^0,\ m^0))$

CV 是在新(时期1)价格下取得新效用所需的货币,与在新价格下取得老(时期0)效用水平所需的货币之间的差额.

补偿变量,$\mathbf{p}^0$, m^0 和 $\mathbf{p}^1$, m^1 分别是时期0和时期1的价格和收入,$e(\mathbf{p}^1,\ v(\mathbf{p}^1,\ m^1))=m_0^1$.

特殊函数形式及其性质

线性支出系统(LES)

26.28 $$u(\mathbf{x})=\prod_{i=1}^{n}(x_i-c_i)^{\beta_i},\quad \beta_i>0$$

Stone-Geary 效用函数. 如果 $c_i=0$ 对于所有 i 成立,$u(\mathbf{x})$是柯柏-道格拉斯函数.

26.29 $$x_i(\mathbf{p},\ m)=c_i+\frac{1}{p_i}\frac{\beta_i}{\beta}\Big(m-\sum_{i=1}^{n}p_ic_i\Big)$$

需求函数. $\beta=\sum_{i=1}^{n}\beta_i$.

26.30 $$v(\mathbf{p},\ m)=\beta^{-\beta}\Big(m-\sum_{i=1}^{n}p_ic_i\Big)^{\beta}\prod_{i=1}^{n}\Big(\frac{\beta_i}{p_i}\Big)^{\beta_i}$$

间接效用函数.

26.31 $$e(\mathbf{p},\ u)=\sum_{i=1}^{n}p_ic_i+\frac{\beta u^{1/\beta}}{\Big[\prod_{i=1}^{n}\Big(\frac{\beta_i}{p_i}\Big)^{\beta_i}\Big]^{1/\beta}}$$

支出函数.

接近理想的需求系统(AIDS)

26.32 $$\ln(e\ (\mathbf{p},\ u))=a(\mathbf{p})+ub(\mathbf{p})$$

$$a(\mathbf{p})=\alpha_0+\sum_{i=1}^{n}\alpha_i\ln p_i+\frac{1}{2}\sum_{i,j=1}^{n}\gamma_{ij}^{*}\ln p_i\ln p_j$$

$$b(\mathbf{p})=\beta_0\prod_{i=1}^{n}p_i^{\beta_i}$$

约束: $\sum_{i=1}^{n}\alpha_i=1$, $\sum_{i=1}^{n}\beta_i=0$, 以及

$$\sum_{i=1}^{n}\gamma_{ij}^{*}=\sum_{i=1}^{n}\gamma_{ji}^{*}=0.$$

接近理想的需求系统,由支出函数的对数而定义. 约束条件使 $e(\mathbf{p},\ u)$对 $\mathbf{p}$ 的齐次度为1.

26.33 $x_i(p, m)=\frac{m}{p_i}\left(\alpha_i+\sum_{j=1}^{n}\gamma_{ij}\ln p_j+\beta_i\ln\left(\frac{m}{P}\right)\right)$,
其中价格指数 P 由
$$\ln P = \alpha_0 + \sum_{i=1}^{n}\alpha_i \ln p_i + \frac{1}{2}\sum_{i,j=1}^{n}\gamma_{ij}\ln p_i \ln p_j$$
给出,且有 $\gamma_{ij} = \frac{1}{2}(\gamma_{ij}^* + \gamma_{ji}^*) = \gamma_{ji}$

需求函数.

对数变换间接需求函数

26.34 $$\ln v(p, m) = \alpha_0 + \sum_{i=1}^{n}\alpha_i \ln\left(\frac{p_i}{m}\right)+ \frac{1}{2}\sum_{i,j=1}^{n}\beta_{ij}^*\ln\left(\frac{p_i}{m}\right)\ln\left[\frac{p_j}{m}\right]$$

对数变换间接需求函数.

26.35 $$x_i(p, m)=\frac{m}{p_i}\left[\frac{\alpha_i+\sum_{j=1}^{n}\beta_{ij}\ln(p_j/m)}{\sum_{i=1}^{n}\alpha_i+\sum_{i,j=1}^{n}\beta_{ij}^*\ln(p_i/m)}\right]$$
其中 $\beta_{ij} = \frac{1}{2}(\beta_{ij}^* + \beta_{ji}^*)$

需求函数.

价格指数

26.36 考虑一个有 n 种商品的“篮子”.对 $i = 1, \cdots, n$,定义

$q^{(i)}$ = 商品 i 在篮子里的个数

$p_0^{(i)}$ = 商品 i 在0年的单位价格

$p_t^{(i)}$ = 商品 i 在 t 年的单位价格

以0年为基期年,t 年的价格指数 P 定义为
$$P = \frac{\sum_{i=1}^{n}p_t^{(i)}q^{(i)}}{\sum_{i=1}^{n}p_0^{(i)}q^{(i)}}\cdot 100$$

价格指数最普通的定义. P 是100乘以在 t 年商品篮子的成本除以在0年的商品篮子的成本.(更一般的,一种(消费)价格指数可定义为所有价格的函数 $P(p_1, \cdots, p_n)$,齐次度为1,且对每一变量都不递减.)

26.37	• 如果数量 $q^{(i)}$在 P 的公式中是在基期年 0 的消费水平，P 是 *Laspeyres* 价格指数. • 如果数量 $q^{(i)}$是在 t 年的消费水平，P 称为 *Paasche* 价格指数.	两种重要的价格指数.
26.38	$F = \sqrt{(\text{Laspeyres 指数}) \cdot (\text{Paasche 指数})}$	*Fischer* 理想指数.

参 考 文 献

一本基本的参考是 Varian(1996)，对于更深的讨论，参见 Mas-Colell *et. al*. (1995)，对于 AIDS，参见 Deaton & Muellbauer(1980)，对于对数变换，参见 Christensen, Jorgenson & Lau(1975)，也可参见 Phlips(1983).

27 金融和经济增长理论中的问题

27.1 $S_t = S_{t-1} + rS_{t-1} = (1+r)S_{t-1},\ t = 1, 2, \cdots$

在一利率为 r 的账户中,金额 S_{t-1} 经过一个时期递增至 S_t.

27.2 当利率为 r 时,在第 t 期末的本金 S_0 的复利金额 S_t 为

$$S_t = S_0(1+r)^t$$

复利((27.1)中差分方程的解).

27.3 当利率为 r,在每期期末计利时,要得到经过 t 期后复利金额为 S_t 的初始投资额为

$$S_0 = S_t(1+r)^{-t}$$

S_0 称为 S_t 的现值.

27.4 当利息是每年在固定的时间间隔里计复利 n 次,每期利率为 r/n 时,有效年利率为

$$\left(1+\frac{r}{n}\right)^n - 1$$

有效年利率.

27.5 $$A_t = \frac{R}{(1+r)^1} + \frac{R}{(1+r)^2} + \cdots + \frac{R}{(1+r)^t} = R\,\frac{1-(1+r)^{-t}}{r}$$

年金为 R,每期利率为 r 时 t 期的现值 A_t.

27.6 每期年金 R 的无限期年金,每期利率为 r 时的现值 A:

$$A = \frac{R}{(1+r)^1} + \frac{R}{(1+r)^2} + \cdots = \frac{R}{r}$$

无限年金的现值.

27.7 $$T = \frac{\ln\left(\dfrac{R}{R-rA}\right)}{\ln(1+r)}$$

当每期还本为 R,每期利率为 r 时,偿还贷款 A 所需的时期数 T.

27.8 $S_t = (1+r)S_{t-1} + (y_t - x_t),\ t=1, 2, \cdots$

在一利率为 r 的账户里,如果 y_t 是 t 期的存款,x_t 是取款,金额 S_{t-1} 经过一个时期递增至 S_t.

27.9 $S_t = (1+r)^t S_0 + \sum_{k=1}^{t}(1+r)^{t-k}(y_k - x_k)$

方程(27.8)的解.

27.10 $S_t = (1+r_t)S_{t-1} + (y_t - x_t),\ t=1, 2, \cdots$

(27.8)当利率是变量 r_t 时的推广.

27.11 $D_k = \dfrac{1}{\prod_{s=1}^{k}(1+r_s)}$

与(27.10)相伴的贴现因子(从第 k 期贴现至 0 期).

27.12 $R_k = \dfrac{D_k}{D_t} = \prod_{s=k+1}^{t}(1+r_s)$

与(27.10)相伴的利息因子.

27.13 $S_t = R_0 S_0 + \sum_{k=1}^{t} R_k(y_k - x_k)$

(27.10)的解. R_k 的定义在(27.12).((27.9)的推广)

27.14 $a_0 + \dfrac{a_1}{1+r} + \dfrac{a_2}{(1+r)^2} + \cdots + \dfrac{a_n}{(1+r)^n} = 0$

r 是一投资项目的内部报酬率,负的 a_t 代表在时间 t 的支出,正的 a_t 代表收入.

27.15 如果 $a_0 < 0$ 且 $a_1, \cdots, a_n$ 都 $\geqslant 0$,则(27.14)有唯一的解 $1+r^* > 0$,即唯一的内部报酬率 $r^* > -1$. 当 $\sum_{i=0}^{n} a_i > 0$ 时内部报酬率是正的.

Descartes 符号法则(2.12)的应用.

27.16 $A_0 = a_0,\ A_1 = a_0 + a_1,\ A_2 = a_0 + a_1 + a_2,\ \cdots,$ $A_n = a_0 + a_1 + \cdots + a_n$

与(27.14)相伴的累积现金流量.

27.17	如果 $A_n \neq 0$, 且数列 $A_0, A_1, \cdots, A_n$ 只改变一次符号,则(27.14)有一唯一的正的内部报酬率.	*Norstrøm* 法则.
27.18	在连续计利且复利利率为 r 的账户里, K 元经过 t 年的利息为: Ke^{rt}	连续复利.
27.19	利息为 r 时连续复利的有效年利率为 $e^r - 1$	连续复利的有效利率.
27.20	$Ke^{-rt}, \ r = p/100$	在 t 年后到期的金额 K, 当(连续复利)年利率为 $p\%$ 时的现值.
27.21	一个在时间区间 $[0, T]$ 内每年 $K(t)$ 元的连续收入,当连续复利利率为 r 时,在时间 0 的贴现现值是 $$\int_0^T K(t)e^{-rt}dt$$	贴现现值,连续复利.
27.22	一个在时间区间 $[s, T]$ 内每年 $K(t)$ 元的连续收入,当连续复利利率为 r 时,在时间 s 的贴现现值是 $$\int_s^T K(t)e^{-r(t-s)}dt$$	贴现现值,连续复利.
27.23	索洛增长模型 • $X(t) = F(K(t), L(t))$ • $K(t) = sX(t)$ • $L(t) = L_0 e^{\lambda t}$	$X(t)$是国民收入, $K(t)$是在时间 t 的资本, $L(t)$是劳动力, F 是生产方程. s(储蓄率), λ 和 L_0 都是正的常数.
27.24	如果 F 的齐次度为 1, $k(t) = K(t)/L(t)$ 为人均资本量,而 $f(k) = F(k, 1)$, 则(27.23)简化为 $\dot{k} = sf(k) - \lambda k$, $k(0)$是给定的.	(27.23)的一个简化形式.

27.25 如果 $\lambda/s < f'(0) < \infty$，当 $k \to \infty$ 时 $f'(k) \to 0$，且 $f''(k) \leqslant 0$ 对于所有 $k \geqslant 0$ 成立，则(27.24)中的方程在$[0, \infty)$上有一唯一的解.解 k^* 定义为

$$sf(k^*) = \lambda k^*$$

是一稳定的均衡状态.

在$[0, \infty)$上的解的存在性和唯一性由(11.41)得出.

Ramsey 增长模型

27.26
$$\max \int_0^T C(t) e^{-rt} \mathrm{d}t$$
$$C(t) = U(f(K(t)) - \dot{K}(t))$$
$$K(0) = K_0, \; K(T) \geqslant K_1$$

增长理论中的一个标准问题. U 是一效用函数，$K(t)$在时间 t 的资本存量，$f(K)$是生产函数，$C(t)$是消费，r 是贴现率，而 T 是规划的时域.

27.27 $\ddot{K} - f'(K)\dot{K} + \dfrac{U'(C)}{U''(C)}(r - f'(K)) = 0$

问题(27.26)的欧拉方程.

27.28 $\dfrac{\dot{C}}{C} = \dfrac{r - f'(K)}{\check{\omega}}$，其中
$$\check{\omega} = \mathrm{El}_C U'(C) = CU''(C)/U'(C)$$

(27.26)的解的必要条件.

参 考 资 料

对于复利的公式参考 Goldberg(1961)或 Sydsaeter & Hammond(1995)，对于(27.17)参考 Norstrøm(1972)，对于增长理论，参考 Burmeister & Dobell(1970)，Blanchard & Fischer(1989)，或 Barro & Sala-I-Martin(1995).

28 风险和风险规避理论

28.1 $R_A = -\dfrac{u''(y)}{u'(y)}, \quad R_R = yR_A = -\dfrac{yu''(y)}{u'(y)}$

绝对风险规避(R_A)和相对风险规避(R_R). $u(y)$是一效用函数, y是收入或消费.

28.2 • $R_A = \lambda \Leftrightarrow u(y) = A_1 + A_2 e^{-\lambda y}$

• $R_R = k \Leftrightarrow u(y) = \begin{cases} A_1 + A_2 \ln y & \text{如果 } k = 1 \\ A_1 + A_2 y^{1-k} & \text{如果 } k \neq 1 \end{cases}$

固定绝对风险规避和相对风险规避效用函数各自的特征. A_1 和 A_2 是常数, $A_2 \neq 0$.

28.3 • $u(y) = y - \dfrac{1}{2} by^2 \Rightarrow R_A = \dfrac{b}{1 - by}$

• $u(y) = \dfrac{1}{b-1}(a + by)^{1-\frac{1}{b}} \Rightarrow R_A = \dfrac{1}{a + by}$

两种特殊效用函数的风险规避.

28.4 $E(u(y + z + \pi)) = E(u(y))$

$\pi \approx -\dfrac{u''(y)}{u'(y)} \dfrac{\sigma^2}{2} = R_A \dfrac{\sigma^2}{2}$

Arrow-Pratt 风险奖励. π: 风险奖励. z: 均值为零的风险项目. $\sigma^2 = \mathrm{var}[\ ]$: z 的方差. $E[\ ]$是期望. (期望和方差定义在第 31 章)

28.5 如果 F 和 G 是随机收入的累积分布函数(CDF),则

F 一级随机优于 G

$\Leftrightarrow G(Z) \geqslant F(Z)$ 对于所有 $Z \in I$ 成立.

一级随机优于的定义. I 是一闭区间 $[Z_1, Z_2]$,当 $Z \leqslant Z_1$ 时, $F(Z) = G(Z) = 0$,而当 $Z \geqslant Z_2$ 时, $F(Z) = G(Z) = 1$.

28.6 $F \text{ FSD } G \Leftrightarrow \begin{cases} E_F[u(Z)] \geqslant E_G[u(Z)] \\ \text{对于所有递增的 } u(Z) \text{ 成立。} \end{cases}$

重要的结论. FSD 意为"一级随机优于". $E_F[u(Z)]$是当收入 Z 的累积分布函数是 $F(Z)$ 时的期望效用. $E_G[u(Z)]$可类似地定义.

28.7 $T(Z) = \int_{Z_1}^{Z} (G(z) - F(z))\mathrm{d}z$

(28.8)中使用的定义.

28.8 F 二级随机优于 G

$\Leftrightarrow T(Z) \geqslant 0$ 对于所有 $Z \in I$ 成立.

二级随机优于(SSD)的定义, $I = [Z_1, Z_2]$, 注意 FSD$\Rightarrow$SSD.

28.9 $F \text{ SSD } G \Leftrightarrow \begin{cases} E_F[u(Z)] \geqslant E_G[u(Z)] \\ \text{对于所有递增且} \\ \text{凹的 } u(Z) \text{ 成立.} \end{cases}$

Hadar-Russell 定理. 每一风险规避者偏好 F 超过 G, 当且仅当 F SSD G.

28.10 设 F 和 G 分别为 X 和 Y 的分布函数，$I = [Z_1, Z_2]$，且 $T(Z)$的定义如(28.7)，则以下陈述是等价的：

- $T(Z_2) = 0$ 且 $T(Z) \geqslant 0$ 对于所有 $Z \in I$ 成立.
- 存在一随机变量 ϵ，且 $E[\epsilon \mid X] = 0$ 对于所有 X 成立，使 Y 分布为 $X + \epsilon$.
- F 和 G 有相同的均值，每一风险规避者偏好 F 超过 G.

Rothschild-Stiglitz 定理.

参 考 文 献

参考 Huang & Litzenberger(1988)，Hadar & Russell(1969)，和 Rothschild & Stiglitz(1970).

29 金融和随机微积分

资本资产定价模型

29.1 $E[r_i]=r+\beta_i(E[r_m]-r)$

其中 $\beta_i=\dfrac{\operatorname{corr}(r_i,\ r_m)\sigma_i}{\sigma_m}=\dfrac{\operatorname{cov}(r_i,\ r_m)}{\sigma_m^2}$

r_i 为资产 i 的收益率，$E[r_k]$为 r_k 的期望值. r 为无风险资产的收益率，r_m 为市场收益率，σ_i 为 r_i 的标准差.

单一消费 β 资产定价方程

29.2 $E(r_i)=r+\dfrac{\beta_{ic}}{\beta_{mc}}(E(r_m)-r)$,

其中 $\beta_{jc}=\dfrac{\operatorname{cov}(r_j,\ \mathrm{d}\ln C)}{\operatorname{var}(\mathrm{d}\ln C)}$, $j=i$ 或 m.

C 为消费.
r_m 为任意有价证券的收益.
$\mathrm{d}\ln C$ 是随机对数微分.（参见(29.13)）

29.3 *Black-Scholes* 期权定价模型(欧洲或美洲无股息的买入股票期权)

$$c=c(S,\ K,\ t,\ r,\ \sigma)$$
$$=SN(x)-KN(x-\sigma\sqrt{t})\mathrm{e}^{-rt},$$

其中 $x=\dfrac{\ln(S/K)+(r+\frac{1}{2}\sigma^2)t}{\sigma\sqrt{t}}$,

且 $N(y)=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{y}\mathrm{e}^{-z^2/2}\mathrm{d}z$ 是累积正态分布函数.

c 为对 S 的时间 t 的期权的价值，S 为优先股价格，$\mathrm{d}S/S=\alpha\mathrm{d}t+\sigma\mathrm{d}B$，其中 B 是一(标准)布朗运动，α 为趋势参数，σ 为波动性(测量对平均的偏差). t 为距到期日的时间. r 为利率. K 为交割价格.

29.4 • $\partial c/\partial S = N(x) > 0$

• $\partial c/\partial K = -N(x-\sigma\sqrt{t})e^{-rt} < 0$

• $\partial c/\partial t = \frac{\sigma}{2\sqrt{t}}SN'(x) + re^{-rt}KN(x-\sigma\sqrt{t}) > 0$

• $\partial c/\partial r = tKN(x-\sigma\sqrt{t})e^{-rt} > 0$

• $\partial c/\partial \sigma = SN'(x)\sqrt{t} > 0$

Black-Scholes 模型中有用的敏感性结论.(对(29.5)中的广义 Black-Scholes 模型的相应结论见 Haug(1997),附录 B 中给出)

29.5 广义 *Black-Scholes* 模型,包括持仓成本项 b(用来对欧洲支付连续股息收益资产的,期货期权和货币期权,包括买入期权(c)和卖出期权(p)定价)

$$c = SN(x)e^{(b-r)t} - KN(x-\sigma\sqrt{t})e^{-rt}$$

$$p = KN(\sigma\sqrt{t}-x)e^{-rt} - SN(-x)e^{(b-r)t}$$

其中 $x = \dfrac{\ln(S/K) + (b+\frac{1}{2}\sigma^2)t}{\sigma\sqrt{t}}$.

b 为持有优先证券的持仓成本率. $b=r$ 给出 Black-Scholes 模型, $b=r-q$ 给出连续股息收益为 q 的 Merton 股票期权模型, $b=0$ 给出 Black 期货期权模型.

29.6 $p = c - Se^{(b-r)t} + Ke^{-rt}$

广义 Black-Scholes 模型的买入卖出平价.

29.7 $P(S, K, t, r, b, \sigma) = C(K, S, t, r-b, -b, \sigma)$

给出美国卖出期权 P,根据相应的买入期权 C 的公式的变换.

29.8 当优先资产不支付股息时,美国永久性卖出期权的市场价值:

$$h(x) = \begin{cases} \dfrac{K}{1+\gamma}\left(\dfrac{x}{c}\right)^{-\gamma}, & \text{如果 } x \geqslant c \\ K - x, & \text{如果 } x < c \end{cases}$$

其中 $c = \dfrac{\gamma K}{1+\gamma}$, $\gamma = \dfrac{2r}{\sigma^2}$.

x 为现价.
c 为触发价.
r 为利率.
K 为交割价.
σ 为波动性.

29.9 $X_t = X_0 + \int_0^t u(s, \omega)\mathrm{d}s + \int_0^t v(s, \omega)\mathrm{d}B_s$，其中 $P\left[\int_0^t v(s, \omega)^2 \mathrm{d}s < \infty \text{ 对所有 } t \geq 0\right] = 1$ 成立，且 $P\left[\int_0^t |u(s, \omega)| \mathrm{d}s < \infty \text{ 对所有 } t \geq 0\right] = 1$ 成立. u 和 v 通过滤波 $\{\mathfrak{F}_t\}$ 进行调整，其中 B_t 是一 $\mathfrak{F}_t$ 布朗运动.

X_t 定义为一维随机积分.

29.10 $\mathrm{d}X_t = u\,\mathrm{d}t + v\,\mathrm{d}B_t$

(29.9)的微分形式.

29.11 如果 $\mathrm{d}X_t = u\,\mathrm{d}t + v\,\mathrm{d}B_t$ 且 $Y_t = g(X_t)$，其中 g 是 C^2 函数，则

$$\mathrm{d}Y_t = \left(g'(X_t)u + \frac{1}{2}g''(X_t)v^2\right)\mathrm{d}t + g'(X_t)v\,\mathrm{d}B_t$$

Itô 公式(一维).

29.12 $\mathrm{d}t \cdot \mathrm{d}t = \mathrm{d}t \cdot \mathrm{d}B_t = \mathrm{d}B_t \cdot \mathrm{d}t = 0$，
$\mathrm{d}B_t \cdot \mathrm{d}B_t = \mathrm{d}t$

有用的关系.

29.13 $$\mathrm{d}\ln X_t = \left(\frac{u}{X_t} - \frac{v^2}{2X_t^2}\right)\mathrm{d}t + \frac{v}{X_t}\mathrm{d}B_t$$

$$\mathrm{d}e^{X_t} = \left(e^{X_t}u + \frac{1}{2}e^{X_t}v^2\right)\mathrm{d}t + e^{X_t}v\,\mathrm{d}B_t$$

(29.11)的两个特例.

29.14 $$\begin{pmatrix} \mathrm{d}X_1 \\ \vdots \\ \mathrm{d}X_n \end{pmatrix} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}\mathrm{d}t + \begin{pmatrix} v_{11} & \cdots & v_{1m} \\ \vdots & \ddots & \vdots \\ v_{n1} & \cdots & v_{nm} \end{pmatrix}\begin{pmatrix} \mathrm{d}B_1 \\ \vdots \\ \mathrm{d}B_m \end{pmatrix}$$

(29.10)的向量形式，其中 $B_1, \cdots, B_m$ 是 m 个独立的一维布朗运动.

29.15 如果 $\mathbf{Y} = (Y_1, \cdots, Y_k) = \mathbf{g}(t, \mathbf{X})$，其中 $\mathbf{g} = (g_1, \cdots, g_k)$ 是 C^2 函数，则对于 $r = 1, \cdots, k$

$$\mathrm{d}Y_r = \frac{\partial g_r(t, \mathbf{X})}{\partial t}\mathrm{d}t + \sum_{i=1}^n \frac{\partial g_r(t, \mathbf{X})}{\partial x_i}\mathrm{d}X_i + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2 g_r(t, \mathbf{X})}{\partial x_i \partial x_j}\mathrm{d}X_i\mathrm{d}X_j$$

其中 $\mathrm{d}t \cdot \mathrm{d}t = \mathrm{d}t \cdot \mathrm{d}B_i = 0$ 且 $\mathrm{d}B_i \cdot \mathrm{d}B_j = \mathrm{d}t$，如果 $i = j$；$\mathrm{d}B_i \cdot \mathrm{d}B_j = 0$，如果 $i \neq j$.

n 维 *Itô* 公式.

29.16	$J(t, x) = \max_u E^{t,x}\int_t^T e^{-rs}W(s, X_s, u_s)ds$ 其中 T 是固定的, $u_s \in U$, U 是一固定区间,且 $dX_t = b(t, X_t, u_t)dt + \sigma(t, X_t, u_t)dB_t$	随机控制问题. J 是值函数, u_t 是控制, $E^{t,x}$ 是当初始条件为 $X_t = x$ 的期望.
29.17	$-J_t'(t, x) = \max_{u \in U}[W(t, x, u) + J_x'(t, x)b(t, x, u) + \frac{1}{2}J_{xx}''(t, x)(\sigma(t, x, u))^2]$	*Hamilton-Jacobi-Bellman* 方程. (29.16)的最优解的一个必要条件.

参考文献

对于(29.1)和(29.2)参考 Sharpe(1964),对于(29.3)参考 Black & Scholes (1973),对于(29.5)和许多期权定价公式,参考 Haug(1997),他也提供了详细的参考文献和期权标价公式的计算机码,对于(29.8)参考 Merton(1973),对于随机积分和随机控制理论,参考 Φksendal(1998), Fleming & Rishel(1975) 和 Karatzas & Shreve(1988).

30 非合作博弈论

30.1 在一有 n 个博弈方的博弈中，我们给每一博弈方 $i(i=1,\cdots,n)$ 一个策略集合 S_i 和一纯策略得益函数 u_i，使对于每一策略组合 $\boldsymbol{s}=(s_1,\cdots,s_n)\in S=S_1\times\cdots\times S_n$，每一博弈方 i 有效用 $u_i(\boldsymbol{s})=u_i(s_1,\cdots,s_n)$.

策略（或正规）形式的 n 人博弈，如果所有策略集合 S_i 有有限个数的元素，则称博弈是有限的.

30.2 一 n 人博弈的一个策略组合 $(s_1^*,\cdots,s_n^*)$ 是一纯策略纳什均衡，如果对于所有 $i=1,\cdots,n$ 和所有 $s_i\in S_i$，

$$u_i(s_1^*,\cdots,s_n^*)$$
$$\geqslant u_i(s_1^*,\cdots,s_{i-1}^*,s_i,s_{i+1}^*,\cdots,s_n^*)$$

n 人博弈的纯策略纳什均衡的定义.

30.3 如果对于所有 $i=1,\cdots,n$，策略集合 S_i 是一非空，紧的，且凸的 $\mathbf{R}^m$ 的子集，而 $u_i(s_1,\cdots,s_n)$ 对其第 i 个变量在 $S=S_1\times\cdots\times S_n$ 是连续且拟凹的，则该博弈有一纯策略纳什均衡.

纯策略纳什均衡存在的充分条件（典型地，存在几个纳什均衡）.

30.4 考虑一有限 n 人博弈，其中 S_i 是博弈中 i 的纯策略集合，且记 $S=S_1\times\cdots\times S_n$. Ω_i 为 S_i 的概率分布集合. 一元素 $\sigma_i\in\Omega_i$（σ_i 则成为一函数 $\sigma_i: S_i\to[0,1]$）称为博弈方 i 的一个混合策略，意为如果 i 采用 σ_i，则 i 以概率 $\sigma_i(s_i)$ 选择纯策略 s_i. 如果博弈方选择混合策略组合 $\boldsymbol{\sigma}=(\sigma_1,\cdots,\sigma_n)\in\Omega_1\times\cdots\times\Omega_n$，纯策略组合 $s=(s_1,\cdots,s_n)$ 出现的概率是 $\sigma_1(s_1)\cdots\sigma_n(s_n)$. 博弈方 i 如果选择混合策略组合 σ 的期望得益，则是

$$u_i(\boldsymbol{\sigma})=\sum_{s\in S}\sigma_1(s_1)\cdots\sigma_n(s_n)u_i(s)$$

一个有 n 个博弈方的博弈中混合策略的定义.

30.5 一个混合策略组合 $\sigma^* = (\sigma_1^*, \cdots, \sigma_n^*)$ 是一纳什均衡,如果对于所有 i 和每一 σ_i,

$$u_i(\boldsymbol{\sigma}^*) \geqslant u_i(\sigma_1^*, \cdots, \sigma_{i-1}^*, \sigma_i, \sigma_{i+1}^*, \cdots, \sigma_n^*)$$

一个有 n 个博弈方的博弈中混合策略纳什均衡的定义.

30.6 $\boldsymbol{\sigma}^*$ 是纳什均衡,当且仅当以下两个条件对于所有 $i = 1, 2, \cdots, n$ 成立:

$\sigma_i^*(s_i) > 0 \Rightarrow u_i(\boldsymbol{\sigma}^*) = u_i(s_i, \boldsymbol{\sigma}_{-i}^*)$ 对所有 s_i 成立,

$\sigma_i^*(s_i') = 0 \Rightarrow u_i(\boldsymbol{\sigma}^*) \geqslant u_i(s_i', \boldsymbol{\sigma}_{-i}^*)$ 对所有 s_i' 成立,

其中 $\boldsymbol{\sigma}_{-i}^* = (\sigma_1^*, \cdots, \sigma_{i-1}^*, \sigma_{i+1}^*, \cdots, \sigma_n^*)$,且我们认为 s_i 和 s'_i 是退化的混合策略.

(混合策略)纳什均衡的另一种定义.

30.7 每一有限 n 博弈方的博弈有一混合策略纳什均衡.

一个重要的结论.

30.8 博弈方 i 的纯策略 $s_i \in S_i$ 是被严格主导的,如果存在博弈方 i 的一个混合策略 σ_i,使对于所有其他博弈方的可能的策略组合,i 采用策略 s_i 的得益严格小于采用策略 σ_i 的得益:

$$u_i(s_1, \cdots, s_{i-1}, s_i, s_{i+1}, \cdots, s_n) < u_i(s_1, \cdots, s_{i-1}, \sigma_i, s_{i+1}, \cdots, s_n)$$

对于每一能从其他博弈方的策略集合 $S_1, \cdots, S_{i-1}, S_{i+1}, \cdots, S_n$ 中所得的策略组合 $(s_1, \cdots, s_{i-1}, s_{i+1}, \cdots, s_n)$ 都成立.

严格下策的定义.

30.9 在一 n 人博弈中,以下结论成立:

- 如果反复消去严格下策的过程仅保留了策略 $(s_1^*, \cdots, s_n^*)$,则这些策略是该博弈唯一的纳什均衡.
- 如果混合策略组合 $\sigma^* = (\sigma_1^*, \cdots, \sigma_n^*)$ 是一纳什均衡,且对某一博弈方 i,$\sigma_i^*(s_i) > 0$,则 s_i 在反复消去严格下策的过程中会幸存下来.

有用的结论.反复消去严格下策的过程并不一定消去任何策略.(对反复消去严格下策的过程感兴趣的,可参见有关的文献)

30.10 在有两个博弈方的博弈中,博弈方 1 和 2 分别有 m 和 n 个(纯)策略,可以用两个得益矩阵来表述

$$\mathbf{A}=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},\ \mathbf{B}=\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$$

$a_{ij}(b_{ij})$是当博弈方分别采用纯策略 i 和 j 时博弈方 1(2) 的得益. 如果 $\mathbf{A}=-\mathbf{B}$, 博弈称为是两人零和博弈, 如果 $\mathbf{A}=\mathbf{B}'$, 该博弈称为是两人对称博弈.

30.11 (30.10)中的两人博弈存在一纳什均衡 $(\boldsymbol{p}^*, \boldsymbol{q}^*)$使

- $\boldsymbol{p}\cdot\mathbf{A}\boldsymbol{q}^*\leqslant\boldsymbol{p}^*\cdot\mathbf{A}\boldsymbol{q}^*$ 对于所有 $\boldsymbol{p}$ 在 Δ_m 中成立,
- $\boldsymbol{p}^*\cdot\mathbf{B}\boldsymbol{q}\leqslant\boldsymbol{p}^*\cdot\mathbf{B}\boldsymbol{q}^*$ 对于所有 $\boldsymbol{q}$ 在 Δ_n 中成立,

两人博弈中纳什均衡的存在性. · 表示数量乘积, Δ_k 表示在 $\mathbf{R}^k$ 中包含的所有元素总和为 1 的非负向量的单纯形.

30.12 在一两人零和博弈 ($\mathbf{A}=-\mathbf{B}$) 中,纳什均衡的存在条件相当于条件 $\boldsymbol{p}\cdot\mathbf{A}\boldsymbol{q}$ 有一鞍点 $(\boldsymbol{p}^*, \boldsymbol{q}^*)$,即对于所有 $\boldsymbol{p}$ 在 Δ_m 中和所有 $\boldsymbol{q}$ 在 Δ_n 中,

$$\boldsymbol{p}\cdot\mathbf{A}\boldsymbol{q}^*\leqslant\boldsymbol{p}^*\cdot\mathbf{A}\boldsymbol{q}^*\leqslant\boldsymbol{p}^*\cdot\mathbf{A}\boldsymbol{q}$$

两人零和博弈中纳什均衡的鞍点特征.

30.13 均衡的得益 $v=\boldsymbol{p}^*\cdot\mathbf{A}\boldsymbol{q}^*$ 称为博弈的价值,且

$$v=\min_{\boldsymbol{q}\in\Delta_n}\max_{\boldsymbol{p}\in\Delta_m}\boldsymbol{p}\cdot\mathbf{A}\boldsymbol{q}=\max_{\boldsymbol{p}\in\Delta_m}\min_{\boldsymbol{q}\in\Delta_n}\boldsymbol{p}\cdot\mathbf{A}\boldsymbol{q}$$

两人零和博弈中的经典极小极大定理.

30.14 假设$(\boldsymbol{p}^*, \boldsymbol{q}^*)$和$(\boldsymbol{p}^{**}, \boldsymbol{q}^{**})$是博弈(30.10)的纳什均衡. 则$(\boldsymbol{p}^*, \boldsymbol{q}^{**})$和$(\boldsymbol{p}^{**}, \boldsymbol{q}^*)$也是均衡策略组合.

长方形或互换性特性.

进化博弈论

30.15 在(30.10)中的两人对称博弈中，有 $\mathbf{A} = \mathbf{B}'$，一个策略 p^* 称为是进化稳定策略(ESS)，如果对于所有 $q \neq p^*$ 存在一 $\bar{\varepsilon} > 0$，使

$$q \cdot \mathbf{A}(\varepsilon q + (1-\varepsilon)p^*) < p^* \cdot \mathbf{A}(\varepsilon q + (1-\varepsilon)p^*)$$

对于所有正的 $\varepsilon < \bar{\varepsilon}$ 成立.

$\bar{\varepsilon}$ 的值可能依赖于 q. 生物学角度的诠释：所有动物都被设定采用 p^*，任何尝试入侵的变异 q，都有严格最低的适应度.

30.16 在(30.15)中设定的策略 $\mathbf{p}^*$ 是进化稳定的当且仅当

$q \cdot \mathbf{A}p^* \leqslant p^* \cdot \mathbf{A}p^*$ 对于所有 q 成立

如果 $q \neq p^*$ 和 $q \cdot \mathbf{A}p^* = p^* \cdot \mathbf{A}p^*$ 成立，则 $q \cdot \mathbf{A}q < p^* \cdot \mathbf{A}q$.

第一个条件，(均衡条件)，等价于纳什均衡条件，第二个条件称为稳定性条件.

参 考 文 献

一个标准的参考文献是 Friedman(1986)，也可参考 Gibbons(1992)(最简单的论述)，Kreps(1990)，和 Fudenberg & Tirole(1991)，对于进化博弈论，参考 Weibull(1995).

31 概率和统计

31.1 事件 A 的概率 $P(A)$，满足以下公理：

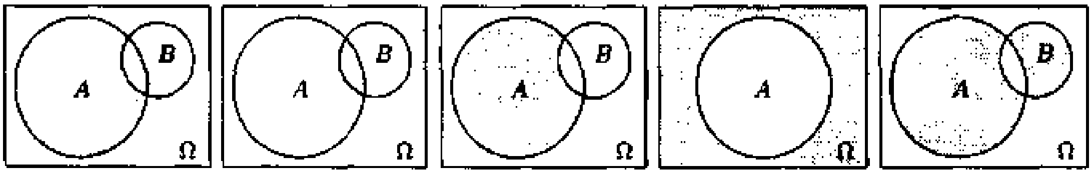
(a) $0 \leqslant P(A) \leqslant 1$

(b) $P(\Omega) = 1$

(c) 如果对于 $i \neq j$，$A_i \cap A_j = \varnothing$，则

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

概率公理．Ω 是包含所有可能结果的样本空间，一个事件是 Ω 的一个子集．



$A \cup B$ A 或 B 发生 | $A \cap B$ A 和 B 都发生 | $A \setminus B$ A 发生，但 B 不发生 | A^c A 不发生 | $A \triangle B$ A 或 B 发生，但不同时

31.2 • $P(A^c) = 1 - P(A)$

• $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

• $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

• $P(A \setminus B) = P(A) - P(A \cap B)$

• $P(A \Delta B) = P(A) + P(B) - 2P(A \cap B)$

概率计算的法则．

31.3 $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$ 是事件 B 已发生时，事件 A 发生的条件概率．

条件概率的定义，$P(B) > 0$．

31.4 A 和 B 是(随机)独立的，如果

$$P(A \cap B) = P(A)P(B)$$

或等价地，如果 $P(B) > 0$ 且

$$P(A \mid B) = P(A)$$

(随机)独立的定义．

31.5 $P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)\cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$

一般概率乘法法则.

31.6 $$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

Bayes 法则. $(P(B) \neq 0.)$

31.7 $$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_{j=1}^{n} P(B \mid A_j) \cdot P(A_j)}$$

广义 Bayes 法则. $A_1, \cdots, A_n$ 是分离的, $\sum_{i=1}^{n} P(A_i) = P(\Omega) = 1$, 其中 $\Omega = \bigcup_{i=1}^{n} A_i$ 是样本空间,B 是一个任意事件.

随机变量(一维)

31.8 • $P(X \in A) = \sum_{x \in A} f(x)$

• $P(X \in A) = \int_A f(x)\mathrm{d}x$

$f(x)$是一离散/连续概率密度函数, X 是一随机变量.

31.9 • $F(x) = P(X \leqslant x) = \sum_{t \leqslant x} f(t)$

• $F(x) = P(X \leqslant x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$

F 是累积离散/连续分布函数,在连续的情况下, $P(X = x) = 0$.

31.10 • $E[X] = \sum_x x f(x)$

• $E[X] = \int_{-\infty}^{\infty} x f(x)\mathrm{d}x$

一个离散/连续概率密度函数 f 的期望.

31.11 • $E[g(X)] = \sum_x g(x)f(x)$

• $E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x$

对于离散/连续概率密度函数 f 的函数 g 的期望.

31.12	$\mathrm{var}[X] = E[(X - E[X])^2]$	一个随机变量的方差，定义为对均值的偏差平方的期望值.
31.13	$\mathrm{var}[X] = E[X^2] - (E[X])^2$	方差的另一种表达.
31.14	$\sigma = \sqrt{\mathrm{var}[X]}$	X 的标准差.
31.15	$\mathrm{var}[aX + b] = a^2\mathrm{var}[X]$	a 和 b 是实数.
31.16	$\mu_k = E[(X - \mu)^k]$	对均值的 k 阶中心矩，$\mu = E[X]$.
31.17	$\eta_3 = \dfrac{\mu_3}{\sigma^3}$，$\eta_4 = \dfrac{\mu_4}{\sigma^4} - 3$	偏斜系数 η_3 和峰态系数 η_4，σ 是标准差，对于正态分布，$\eta_3 = \eta_4 = 0$.
31.18	• $P(\mid X \mid \geqslant \lambda) \leqslant E[X^2]/\lambda^2$ • $P(\mid X - \mu \mid \geqslant \lambda) \leqslant \sigma^2/\lambda^2,\ \lambda > 0$ • $P(\mid X - \mu \mid \geqslant k\sigma) \leqslant 1/k^2,\ k > 0$	不同形式的切比雪夫不等式，σ 是 X 的标准差，$\mu = E[X]$ 是平均值.
31.19	如果 f 在区间 I 上是凸的，且 X 是具有有限期望值的随机变量，则 $f(E[X]) \leqslant E[f(X)]$ 如果 f 是严格凸的，不等号严格成立，除非 X 是一概率为 1 的常数.	Jensen 不等式的特例.
31.20	• $M(t) = E[e^{tX}] = \sum_x e^{tx} f(x)$ • $M(t) = E[e^{tX}] = \int_{-\infty}^{+\infty} e^{tx} f(x)\,dx$	矩生成函数. $M(t)$ 并不总是存在，但如果存在，则 $M(t) = \sum_{k=0}^{\infty} \dfrac{E[X^k]}{k!} t^k$.

31.21 • $C(t)=E[e^{itX}]=\sum_{x} e^{itx}f(x)$

• $C(t)=E[e^{itX}]=\int_{-\infty}^{+\infty} e^{itx}f(x)\mathrm{d}x$

特征函数. $C(t)$ 总是存在，且如果 $E[X^k]$ 对于所有 k 存在，则 $C(t)=\sum_{k=0}^{\infty}\frac{i^kE[X^k]}{k!}t^k$.

随机变量(二维)

31.22 • $P((X,Y)\in A)=\sum_{(x,y)\in A} f(x,y)$

• $P((X,Y)\in A)=\iint_A f(x,y)\mathrm{d}x\,\mathrm{d}y$

$f(x,y)$ 是一二维离散/连续联合概率密度函数，X 和 Y 是随机变量.

31.23 • $F(x,y)=P(X\leqslant x,Y\leqslant y)$

$=\sum_{u\leqslant x}\sum_{v\leqslant y} f(u,v)$(离散情况)

$=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\mathrm{d}u\,\mathrm{d}v$(连续情况)

F 是联合累积离散/连续分布函数.

31.24 • $E[g(X,Y)]=\sum_{x}\sum_{y} g(x,y)f(x,y)$

• $E[g(X,Y)]=$

$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)\mathrm{d}x\,\mathrm{d}y$

$g(X,Y)$ 的期望，其中 X 和 Y 有联合离散/连续概率密度函数 f.

31.25 $\mathrm{cov}[X,Y]=E[(X-E[X])(Y-E[Y])]$ | 协方差的定义.

31.26 $\mathrm{cov}[X,Y]=E[XY]-E[X]E[Y]$ | 有用的事实.

31.27 如果 $\mathrm{cov}[X,Y]=0$，X 和 Y 是不相关的， | 定义.

31.28 $E[XY]=E[X]E[Y]$，如果 X 和 Y 是无关的. | 由 (31.26) 和 (31.27)得出.

31.29 $(E[XY])^2\leqslant E[X^2]E[Y^2]$ | 柯西-施瓦兹不等式.

31.30 如果 X 和 Y 是随机独立的，则 $\mathrm{cov}[X,Y]=0$. | 逆推论不成立.

31.31 $\mathrm{var}[X\pm Y]=\mathrm{var}[X]+\mathrm{var}[Y]\pm 2\,\mathrm{cov}[X,Y]$ | 两个随机变量的和/差的方差.

31.32	$E[a_1X_1+\cdots+a_nX_n+b]=a_1E[X_1]+\cdots+a_nE[X_n]+b$	$X_1, \cdots, X_n$ 是随机变量，而 $a_1, \cdots, a_n, b$ 是实数.
31.33	$\operatorname{var}\left[\sum_{i=1}^{n}a_iX_i\right]=\sum_{i=1}^{n}\sum_{j=1}^{n}a_ia_j\operatorname{cov}[X_i, X_j]$ $=\sum_{i=1}^{n}a_i^2\operatorname{var}[X_i]+2\sum_{j=1}^{n}\sum_{i>j}^{n}a_ia_j\operatorname{cov}[X_i, X_j]$	随机变量线性组合的方差.
31.34	$\operatorname{var}\left[\sum_{i=1}^{n}a_iX_i\right]=\sum_{i=1}^{n}a_i^2\operatorname{var}[X_i]$	当 $X_1, \cdots, X_n$ 不相关的时候的公式(31.33).
31.35	$\operatorname{corr}[X, Y]=\dfrac{\operatorname{cov}[X, Y]}{\sqrt{\operatorname{var}[X]\operatorname{var}[Y]}}\in[-1, 1]$	相关系数定义为标准化的协方差.
31.36	如果 $f(x, y)$ 是 X 和 Y 的联合概率密度函数，则 • $f_X(x)=\sum_y f(x, y)$, $f_Y(y)=\sum_x f(x, y)$ • $f_X(x)=\int_{-\infty}^{+\infty}f(x, y)\mathrm{d}y$, $f_Y(y)=\int_{-\infty}^{+\infty}f(x, y)\mathrm{d}x$ 分别是 X 和 Y 的边际密度.	离散和连续分布的边际密度的定义.
31.37	$f(x \mid y)=\dfrac{f(x, y)}{f_Y(y)}$, $f(y \mid x)=\dfrac{f(x, y)}{f_X(x)}$	条件密度的定义.
31.38	随机变量 X 和 Y 是随机独立的，如果 $f(x, y)=f_X(x)f_Y(y)$. 如果 $f_Y(y)>0$，这相当于 $f(x \mid y)=f_X(x)$.	随机独立性.

31.39 • $E[X \mid y]=\sum_{x} x f(x \mid y)$

• $E[X \mid y]=\int_{-\infty}^{+\infty} x f(x \mid y) \mathrm{d} x$

• $\operatorname{var}[X \mid y]=\sum_{x}(x-E[X \mid y])^{2} f(x \mid y)$

• $\operatorname{var}[X \mid y]=\int_{-\infty}^{+\infty}(x-E[X \mid y])^{2} f(x \mid y) \mathrm{d} x$

离散和连续分布的条件期望和方差的定义.注意 $E[X|y]$ 表示 $E[X|Y=y]$，而 $\operatorname{var}[X|y]$ 表示 $\operatorname{var}[X|Y=y]$.

31.40 $E[Y]=E_{X}[E[Y \mid X]]$

重期望法则. E_X 表示对 X 的期望.

31.41 $E[XY]=E_{X}[XE[Y \mid X]]=E[X_{\mu_{Y \mid X}}]$

对乘积 XY 的期望，等于对 X 和给定 X 时 Y 的条件期望的乘积的期望.

31.42 $\sigma_{Y}^{2}=\operatorname{var}[Y]=E_{X}[\operatorname{var}[Y \mid X]]$
$+\operatorname{var}_{X}[E[Y \mid X]]$
$=E[\sigma_{Y \mid X}^{2}]+\operatorname{var}[\mu_{Y \mid X}]$

Y 的边际方差等于对条件方差的期望，加上条件期望的方差.

31.43 设 $f(x, y)$为一对随机函数(X, Y)的密度函数.假设

$U=\phi_{1}(X, Y), V=\phi_{2}(X, Y)$

定义一个随机变量 X 和 Y 的一一对应 C^1 变换，且逆变换由下式给出

$X=\psi_{1}(U, V), Y=\psi_{2}(U, V)$

则(U, V)的密度函数 $g(u, v)$由下式给出

$g(u, v)=f(\psi_{1}(u, v),$
$\psi_{2}(u, v)) \cdot|J(u, v)|$

给定 Jacobi 行列式

$$J(u, v)=\left|\begin{array}{cc}\frac{\partial \psi_{1}(u, v)}{\partial u} & \frac{\partial \psi_{1}(u, v)}{\partial v} \\ \frac{\partial \psi_{2}(u, v)}{\partial u} & \frac{\partial \psi_{2}(u, v)}{\partial v}\end{array}\right|$$

不等于 0.

怎样找到随机变量变换的密度函数(该公式能简单推广到任意个数的随机变量，所需的约束条件在此并没有完全阐明，可参见参考文献).

统 计 推 断

31.44	如果 $E[\hat{\theta}]=\theta$ 对于所有 $\theta\in\Theta$，则 $\hat{\theta}$ 称为 θ 的无偏估计.	无偏估计的定义. Θ 是参数空间.
31.45	如果 $\hat{\theta}$ 不是无偏的，则 $b=E[\hat{\theta}]-\theta$ 称为是 $\hat{\theta}$ 的偏差.	偏差的定义.
31.46	$\mathrm{MSE}(\hat{\theta})=E[\hat{\theta}-\theta]^2=\mathrm{var}[\hat{\theta}]+b^2$	平均平方误差的定义，MSE.
31.47	$p\lim\hat{\theta}_T=\theta$ 意指对每一 $\varepsilon>0$，$\lim\limits_{T\to\infty}P(\mid\hat{\theta}_T-\theta\mid<\varepsilon)=1$	概率极限的定义. 估计 $\hat{\theta}_T$ 是 T 个观察值的函数.
31.48	$\hat{\theta}$ 是 θ 的一个一致估计，如果对于所有 $\theta\in\Theta$，$p\lim\hat{\theta}_T=\theta$	一致性的定义.
31.49	$\hat{\theta}$ 是渐近无偏的，如果 $\lim\limits_{T\to\infty}E[\hat{\theta}_T]=\theta$ 对于所有 $\theta\in\Theta$ 成立.	渐近无偏估计的定义.
31.50	H_0：零假设(如 $\theta\leqslant 0$) H_1：备择假设(如 $\theta>0$) T：检验统计量 C：临界区域 θ：未知参数	统计检验的一些定义.
31.51	检验：如果 $T\in C$ 拒绝 H_0 而接受 H_1.	检验.
31.52	检验的力度函数是 $\pi(\theta)=P(\text{拒绝 } H_0\mid\theta)$，$\theta\in\Theta$.	检验力度的定义.
31.53	当 H_0 正确时拒绝 H_0 称为第一类错误. 当 H_1 正确时不拒绝 H_0 称为第二类错误.	第一类和第二类错误.
31.54	α 显著水平：最小的 α 使 $P(\text{第一类错误})\leqslant\alpha$ 对于所有 θ 满足 H_0.	检验的 α 显著水平.
31.55	对于给定的数据和检验，显著性概率(或 P 值)是拒绝 H_0 的最小显著水平.	一个重要的概念.

渐 近 结 论

31.56 设$\{X_i\}$为由独立同分布的随机变量组成的数列,每一变量有有限的均值 $E[X_i]=\mu$. 设 $S_n=X_1+\cdots+X_n$, 则

(1) 对每一 $\varepsilon>0$,

$$P\left\{\left|\frac{S_n}{n}-\mu\right|\geqslant\varepsilon\right\}\to 0, \text{ 当 } n\to\infty$$

(2) 以概率 1, $\dfrac{S_n}{n}\to\mu$, 当 $n\to\infty$

(1)是弱大数法则, S_n/n 是 μ 的一个一致估计. (2)是强大数法则.

31.57 设$\{X_i\}$为由独立同分布的随机变量组成的数列,每一变量有有限的均值 $E[X_i]=\mu$, 方差 $\operatorname{var}[X]=\sigma^2$. 令 $S_n=X_1+\cdots+X_n$, 则 $\dfrac{S_n-n\mu}{\sigma\sqrt{n}}$ 的分布当 $n\to\infty$ 时接近标准正态分布. 即

$$P\left\{\frac{S_n-n\mu}{\sigma\sqrt{n}}\leqslant a\right\}\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a}e^{-x^2/2}dx$$

当 $n\to\infty$

中心极限定理.

参 考 文 献

参考 Johnson & Bhattacharyya(1996), Larsen & Marx(1986), Griffiths *et al*. (1993)和 Hogg & Craig(1995).

32 概率分布 最小二乘法

32.1 $f(x)=\begin{cases}\dfrac{x^{p-1}(1-x)^{q-1}}{B(p,q)}, & x\in(0,1)\\ 0, & \text{其他}\end{cases}$

$p>0,\ q>0.$

均值：$E[X]=\dfrac{p}{p+q}$

方差：$\mathrm{var}[X]=\dfrac{pq}{(p+q)^2(p+q+1)}$

k 阶矩：$E[X^k]=\dfrac{B(p+k,q)}{B(p,q)}$

贝塔分布. B 是(9.61)定义的贝塔函数.

32.2 $f(x)=\dbinom{n}{x}p^x(1-p)^{n-x}$

$x=0,1,\cdots,n;\ n=1,2,\cdots\ p\in(0,1)$

均值：$E[X]=np$

方差：$\mathrm{var}[X]=np(1-p)$

矩生成函数：$[pe^t+(1-p)]^n$

特征函数：$[pe^{it}+(1-p)]^n$

二项分布. $f(x)$是当事件对每一观察值发生的概率为 p 时,事件在独立的 n 个观察值中正好发生 x 次的概率.关于$\dbinom{n}{x}$,参见(8.27).

32.3 $f(x,y)=\dfrac{e^{-Q}}{2\pi\sigma\tau\sqrt{1-\rho^2}}$,其中

$$Q=\frac{\left(\frac{x-\mu}{\sigma}\right)^2-2\rho\frac{(x-\mu)(y-\eta)}{\sigma\tau}+\left(\frac{y-\eta}{\tau}\right)^2}{2(1-\rho^2)}$$

$x,y,\mu,\eta\in(-\infty,+\infty),$

$\sigma>0,\ \tau>0,\ |\rho|<1$

均值：$E[X]=\mu,\ E[Y]=\eta$

方差：$\mathrm{var}[X]=\sigma^2,\ \mathrm{var}[Y]=\tau^2$

协方差：$\mathrm{cov}[X,Y]=\rho\sigma\tau$

二重正态分布.(关于矩生成函数和特征函数,参见(32.15)中一般多重正态分布)

32.4 $f(x)=\begin{cases}\dfrac{x^{\frac{1}{2}\nu-1}e^{-\frac{1}{2}x}}{2^{\frac{1}{2}\nu}\Gamma\left(\frac{1}{2}\nu\right)}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\quad \nu=1, 2, \cdots$

均值: $E[X]=\nu$

方差: $\operatorname{var}[X]=2\nu$

矩生成函数: $(1-2t)^{-\frac{1}{2}\nu}$, $t<\frac{1}{2}$

特征函数: $(1-2it)^{-\frac{1}{2}\nu}$

有 ν 自由度的 *chi* 平方分布, Γ 是定义在(9.53)中的伽马函数.

32.5 $f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\quad \lambda>0$

均值: $E[X]=1/\lambda$

方差: $\operatorname{var}[X]=1/\lambda^2$

矩生成函数: $\lambda/(\lambda-t)$, $t<\lambda$

特征函数: $\lambda/(\lambda-it)$

指数分布.

32.6 $f(x)=\dfrac{1}{\beta}e^{-z}e^{-e^{-z}}$, $z=\dfrac{x-\alpha}{\beta}$, $x\in\mathbb{R}$, $\beta>0$

均值: $E[X]=\alpha-\beta\Gamma'(1)$

方差: $\operatorname{var}[X]=\beta^2\pi^2/6$

矩生成函数: $e^{\alpha t}\Gamma(1-\beta t)$, $t<1/\beta$

特征函数: $e^{i\alpha t}\Gamma(1-i\beta t)$

极值(*Gumbel*)分布. $\Gamma'(1)$是伽马函数在1处的导数(参见(9.53)), $\Gamma'(1)=-\gamma$, 其中 $\gamma\approx 0.5772$ 是欧拉常数,参见(8.42).

32.7 $f(x)=\begin{cases}\dfrac{\nu_1^{\frac{1}{2}\nu_1}\nu_2^{\frac{1}{2}\nu_2}x^{\frac{1}{2}\nu_1-1}}{B\left(\frac{1}{2}\nu_1, \frac{1}{2}\nu_2\right)(\nu_2+\nu_1x)^{\frac{1}{2}(\nu_1+\nu_2)}}, & x>0\\ 0, & x\leqslant 0\end{cases}$

$\nu_1, \nu_2=1, 2, \cdots$

均值: $E[X]=\nu_2/(\nu_2-2)$ 对于 $\nu_2>2$ (对于 $\nu_2=1, 2$ 不存在).

方差: $\operatorname{var}[X]=\dfrac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}$ 对于 $\nu_2>4$ (对于 $\nu_2\leqslant 4$ 不存在).

k 阶矩:

$$E[X^k]=\frac{\Gamma\left(\frac{1}{2}\nu_1+k\right)\Gamma\left(\frac{1}{2}\nu_2-k\right)}{\Gamma\left(\frac{1}{2}\nu_1\right)\Gamma\left(\frac{1}{2}\nu_2\right)}\left(\frac{\nu_2}{\nu_1}\right)^k$$

$2k<\nu_2$

F 分布, B 是定义在(9.61)中的贝塔函数. ν_1, ν_2 分别是分子和分母的自由度.

32.8 $f(x)=\begin{cases}\dfrac{\lambda^n x^{n-1}e^{-\lambda x}}{\Gamma(n)}, & x>0,\\ 0 & x\leqslant 0,\end{cases}\quad n,\lambda>0$

均值：$E[X]=n/\lambda$

方差：$\mathrm{var}[X]=n/\lambda^2$

矩生成函数：$[\lambda/(\lambda-t)]^n,\ t<\lambda$

特征函数：$[\lambda/(\lambda-it)]^n$

伽马分布，Γ 是定义在(9.53)中的伽马函数.

32.9 $f(x)=p(1-p)^x,\ x=0,1,2,\cdots;$

$p\in(0,1)$

均值：$E[X]=1/p$

方差：$\mathrm{var}[X]=(1-p)/p^2$

矩生成函数：$p/[1-(1-p)e^t],\ t<-\ln(1-p)$

特征函数：$pe^{it}/[1-(1-p)e^{it}]$

几何分布.

32.10 $f(x)=\dfrac{\dbinom{M}{x}\dbinom{N-M}{n-x}}{\dbinom{N}{n}}$

$x=0,1,\cdots,n;\ n=1,2,\cdots,N$

均值：$E[X]=nM/N$

方差：$\mathrm{var}[X]=np(1-p)(N-n)/(N-1)$，其中 $p=M/N$

超几何分布. 给定 N 个物体组合，其中 M 个物体有某一特定特征，而 $N-M$ 个物体不具有. 从组合中随机挑出 n 个物体，$f(x)$ 则是 x 个物体有该特征，而 $n-x$ 个物体不具有该特征的概率.

32.11 $f(x)=\dfrac{1}{2\beta}e^{-|x-\alpha|/\beta},\ x\in\mathbb{R},\ \beta>0$

均值：$E[X]=\alpha$

方差：$\mathrm{var}[X]=2\beta^2$

矩生成函数：$\dfrac{e^{\alpha t}}{1-\beta^2t^2},\ |t|<1/\beta$

特征函数：$\dfrac{e^{i\alpha t}}{1+\beta^2t^2}$

拉普拉斯分布.

32.12 $f(x) = \dfrac{e^{-z}}{\beta(1+e^{-z})^2}$, $z = \dfrac{x-\alpha}{\beta}$,

$x \in \mathrm{R}$, $\beta > 0$

均值：$E[X] = \alpha$

方差：$\mathrm{var}[X] = \pi^2\beta^2/3$

矩生成函数：

$e^{\alpha t}\Gamma(1-\beta t)\Gamma(1+\beta t) = \pi\beta t e^{\alpha t}/\sin(\pi\beta t)$

特征函数：$i\pi\beta t e^{i\alpha t}/\sin(i\pi\beta t)$

逻辑斯蒂分布.

32.13 $f(x) = \begin{cases} \dfrac{e^{-(\ln x-\mu)^2/2\sigma^2}}{\sigma x\sqrt{2\pi}}, & x > 0, \\ 0, & x \leqslant 0, \end{cases} \quad \sigma > 0$

均值：$E[X] = e^{\mu+(\sigma^2/2)}$

方差：$\mathrm{var}[X] = e^{2\mu}(e^{2\sigma^2} - e^{\sigma^2})$

k 阶矩：$\mathrm{E}[X^k] = e^{k\mu+\frac{1}{2}k^2\sigma^2}$

对数正态分布.

32.14 $f(\boldsymbol{x}) = \dfrac{n!}{x_1!\cdots x_k!}p_1^{x_1}\cdots p_k^{x_k}$

$x_1 + \cdots + x_k = n$, $p_1 + \cdots + p_k = 1$,

$x_j \in \{0, 1, \cdots, n\}$, $p_j \in (0, 1)$, $j = 1, \cdots, k$

X_j 的均值：$E[X_j] = np_j$

X_j 的方差：$\mathrm{var}[X_j] = np_j(1-p_j)$

协方差：$\mathrm{cov}[X_j, X_r] = -np_jp_r$,

$j, r = 1, \cdots, n, j \neq r$

矩生成函数：$\left[\sum_{j=1}^{k} p_j\ e^{t_j}\right]^n$

特征函数：$\left[\sum_{j=1}^{k} p_j\ e^{it_j}\right]^n$

多项分布. $f(\mathbf{x})$是 k 个事件$A_1, \cdots, A_k$ 在 n 个独立观察值中，正好发生 $x_1, \cdots, x_k$ 次的概率，当这些事件发生的概率是 $p_1, \cdots, p_k$ 时.

32.15 $f(\boldsymbol{x}) = \dfrac{1}{(2\pi)^{k/2}\sqrt{|\sum|}}e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\sum^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}$

$\sum = (\sigma_{ij})$ 是对称和正定的，

$\boldsymbol{x} = (x_1, \cdots, x_k)'$, $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_k)'$

均值：$E[X_i] = \mu_i$

方差：$\mathrm{var}[X_i] = \sigma_{ii}$

协方差：$\mathrm{cov}[X_i, X_j] = \sigma_{ij}$

矩生成函数：$e^{\boldsymbol{\mu}'\boldsymbol{t}+\frac{1}{2}\boldsymbol{t}'\sum\boldsymbol{t}}$

特征函数：$e^{-\frac{1}{2}\boldsymbol{t}'\sum\boldsymbol{t}}e^{i\boldsymbol{t}'\boldsymbol{\mu}}$

多重正态分布. $|\sum|$ 是方差-协方差矩阵 $\sum$的行列式，$\boldsymbol{x} = (x_1, \cdots, x_k)'$, $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_k)'$.

32.16 $f(x)=\binom{x-1}{r-1}p^r(1-p)^{x-r}$, $x=r,\ r+1,\ \cdots;\ r=1,\ 2,\ \cdots;\ p\in(0,1)$ 均值: $E[X]=r/p$ 方差: $\text{var}[X]=r(1-p)/p^2$ 矩生成函数: $p^r(1-(1-p)e^t)^{-r}$ 特征函数: $p^r e^{it}(1-(1-p)e^{it})^{-r}$	负二项分布.
32.17 $f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$, $x\in\mathrm{R}$, $\sigma>0$ 均值: $E[X]=\mu$ 方差: $\text{var}[X]=\sigma^2$ 矩生成函数: $e^{\mu t+\frac{1}{2}\sigma^2t^2}$ 特征函数: $e^{i\mu t-\frac{1}{2}\sigma^2t^2}$	正态分布. 如果 $\mu=0$ 且 $\sigma=1$, 就是标准正态分布.
32.18 $f(x)=\begin{cases}\frac{ca^c}{x^{c+1}}, & x>a,\\ 0, & x\leqslant a,\end{cases}$ $a>0,\ c>0$ 均值: $E[X]=ac/(c-1)$, $c>1$ 方差: $\text{var}[X]=a^2c/(c-1)^2(c-2)$, $c>2$ k 阶矩: $E[X^k]=a^kc/c-k$, $c>k$	*Pareto* 分布.
32.19 $f(x)=e^{-\lambda}\frac{\lambda^x}{x!}$, $x=0,\ 1,\ 2,\ \cdots,\ \lambda>0$. 均值: $E[X]=\lambda$ 方差: $\text{var}[X]=\lambda$ 矩生成函数: $e^{\lambda(e^t-1)}$ 特征函数: $e^{\lambda(e^{it}-1)}$	*Poisson* 分布.

32.20 $f(x)=\dfrac{\Gamma\left(\frac{1}{2}(\nu+1)\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{1}{2}\nu\right)}\left(1+\dfrac{x^2}{\nu}\right)^{-\frac{1}{2}(\nu+1)}$,

$x\in\mathrm{R},\ \nu=1,2,\cdots$

均值：$E[X]=0$ 对于 $\nu>1$

(对于 $\nu=1$ 不存在).

方差：$\mathrm{var}[X]=\dfrac{\nu}{\nu-2}$ 对于 $\nu>2$

(对于 $\nu=1,2$ 不存在).

k 阶矩(仅当 $k<\nu$ 时存在)：

$$E[X^k]=\begin{cases}\dfrac{\Gamma\left(\frac{1}{2}(k+1)\right)\Gamma\left(\frac{1}{2}(\nu-k)\right)}{\sqrt{\pi}\,\Gamma\left(\frac{1}{2}\nu\right)}\nu^{\frac{1}{2}k}, & k\ \text{偶数},\\ 0, & k\ \text{奇数}.\end{cases}$$

具有自由度 ν 的学生 t 分布.

32.21 $f(x)=\begin{cases}\dfrac{1}{\beta-\alpha}, & \alpha\leqslant x\leqslant\beta,\\ 0, & \text{其他},\end{cases}\quad \alpha<\beta$

均值：$E[X]=(\alpha+\beta)/2$

方差：$\mathrm{var}[X]=(\beta-\alpha)^2/12$

矩生成函数：$[e^{\beta t}-e^{\alpha t}]/t(\beta-\alpha)$

特征函数：$[e^{i\beta t}-e^{i\alpha t}]/it(\beta-\alpha)$

标准分布.

32.22 $f(x)=\begin{cases}\beta\lambda^{\beta}x^{\beta-1}e^{-(\lambda x)^{\beta}}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\quad \beta,\ \lambda>0$

均值：$E[X]=\dfrac{1}{\lambda}\Gamma\left(1+\dfrac{1}{\beta}\right)$

方差：$\mathrm{var}[X]=\dfrac{1}{\lambda^2}\left[\Gamma\left(1+\dfrac{2}{\beta}\right)-\left(\Gamma\left(1+\dfrac{1}{\beta}\right)\right)^2\right]$

k 阶矩：$E[X^k]=\dfrac{1}{\lambda^k}\Gamma(1+k/\beta)$

Weibull 分布. 当 $\beta=1$ 时就是指数分布.

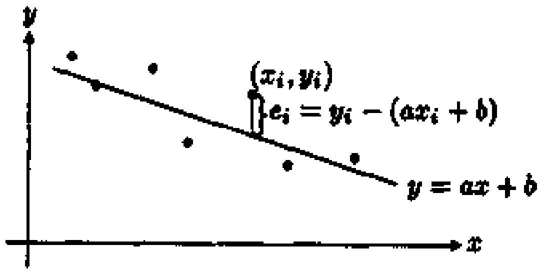
32.23 拟合 n 个数据点 $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ 最好的直线 $y = ax + b$，在使偏差平方和

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} [y_i - (ax_i + b)]^2$$

最小的意义由下式给出

$$y - \overline{y} = \hat{a}(x - \overline{x}),\ \hat{a} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

最小二乘法，这里 $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$，$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$.

32.24 

最小二乘法的图示.

32.25 给定 k 个数量 $x_1, \cdots, x_k$ 的 n 个观察值 $(x_{i1}, \cdots, x_{ik})$，$i = 1, \cdots, n$，和另一数量 y 的 n 个观察值 $y_1, \cdots, y_n$. 定义

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} \text{和}$$

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},\ \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$$

能最好拟合给定观察值的超平面 $y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$ 的系数向量 $\boldsymbol{\beta} = (\beta_0, \cdots, \beta_k)'$，在使偏差平方和，

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

最小的意义上由下式给出

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

最小二乘法，多元回归.

参 考 文 献

参考 Evans, Hasting & Peacock(1993), Johnson, Kotz, & Kemp(1993), Johnson, Kotz, & Balakrishnan(1995), (1997)和 Hogg & Craig(1995).

参考文献

Barro, R. J. and X. Sala-i-Martin: *Economic Growth*, New York: McGraw-Hill(1995).

Bartle, R. G.: *Introduction to Real Analysis*, New York: John Wiley & Sons (1982).

Beavis, B. and I. M. Dobbs: *Optimization and Stability Theory for Economic Analysis*, Cambridge: Cambridge University Press(1990).

Bellman, R.: *Dynamic Programming*, Princeton: Princeton University Press (1957).

Black, F. and M. Scholes: "The pricing of options and corporate liabilities", *Journal of Political Economy*, Vol. 81, 637-654(1973).

Blackorby, C. and R. R. Russell: "Will the real elasticity of substitution please stand up? (A comparison of the Allen/Uzawa and Morishima elasticities)", *American Economic Review*, Vol. 79, 882-888(1989).

Blanchard, O. and S. Fischer: *Lectures on Macroeconomics*, Cambridge: MIT Press(1989).

Braun, M.: *Differential Equations and their Applications*, 4th ed., New York: Springer(1993).

Burmeister, E. and A. R. Dobell: *Mathematical Theories of Economic Growth*, London: Macmillan(1970).

Christensen, L. R., D. Jorgenson and L. J. Lau: "Transcendental logarithmic utility functions", *American Economic Review*, Vol. 65, 367-383(1975).

Deaton, A. and J. Muellbauer: *Economics and Consumer Behaviour*, Cambridge: Cambridge University Press(1980).

Dhrymes, P. J.: *Mathematics for Econometrics*, New York: Springer(1978).

Dixit, A. K.: *Optimization in Economic Theory*, 2nd ed., Oxford: Oxford University Press(1990).

Edwards, C. H. and D. E. Penney: *Calculus with Analytic Geometry*, 5th ed., Englewood Cliffs, N.J.: Prentice-Hall(1998).

Ellickson, B.: *Competitive Equilibrium. Theory and Applications*, Cambridge: Cambridge University Press(1993).

Evans, M., N. Hastings, and B. Peacock: *Statistical Distributions*, 2nd ed., New York: John Wiley & Sons(1993).

Faddeeva, V. N.: *Computational Methods of Linear Algebra*, New York: Dover Publications, Inc. (1959).

Farebrother, R. W.: "Simplified Samuelson conditions for cubic and quartic equations", *The Manchester School of Economic and Social Studies*, Vol. 41, 396-400(1973).

Feichtinger, G. and R. F. Hartl: *Optimale Kontrolle Okonomischer Prozesse*, Berlin: Walter de Gruyter(1986).

Fleming, W. H. and R. W. Rishel: *Deterministic and Stochastic Optimal Control*, Berlin: Springer(1975).

Førsund, F.: "The homothetic production function", *The Swedish Journal of Economics*, Vol. 77, 234-244(1975).

Fraleigh, J. B. and R. A. Beauregard: *Linear Algebra*, 3rd ed., Reading, Mass.: Addison-Wesley(1995).

Friedman, J. W.: *Game Theory with Applications to Economics*, Oxford: Oxford University Press(1986).

Fudenberg, D. and J. Tirole: *Game Theory*, Cambridge: MIT Press(1991).

Fuss, M. and D. McFadden (eds.): *Production Economics: A Dual Approach to Theory and Applications*, Vol. I, Amsterdam: North-Holland(1978).

Gandolfo, G.: *Economic Dynamics*, 3rd ed., Berlin: Springer(1996).

Gantmacher, F. R.: *The Theory of Matrices*, New York: Chelsea Publishing Co. (1959). Reprinted by the American Mathematical Society, Providence, R.I.: AMS Chelsea Publishing(1998).

Gass, S. I.: *Linear Programming. Methods and Applications*, 5th ed., New York: McGraw-Hill(1994).

Gibbons, R.: *A Primer in Game Theory*, New York: Harvester and Wheatsheaf(1992).

Goldberg, S.: *Introduction to Difference Equations*, New York: John Wiley

& Sons(1961).

Griffiths, W. E., R. Carter Hill and G. G. Judge: *Learning and Practicing Econometrics*, New York: John Wiley & Sons(1993).

Hadar, J. and W. R. Russell: "Rules for ordering uncertain prospects", *American Economic Review*, Vol. 59, 25-34(1969).

Halmos, P. R.: *Naive Set Theory*, New York: Springer(1974).

Hardy, G. H., J. E. Littlewood, and G. Pólya: *Inequalities*, Cambridge: Cambridge University Press(1952).

Hartman, P.: *Ordinary Differential Equations*, Boston: Birkhäuser(1982).

Haug, E. G.: *The Complete Guide to Option Pricing Formulas*, New York: McGraw-Hill(1997).

Hildebrand, F. B.: *Finite-Difference Equations and Simulations*, Englewood Cliffs, N.J.: Prentice-Hall(1968).

Hildenbrand, W.: *Core and Equilibria of a Large Economy*, Princeton: Princeton University Press(1974).

Hildenbrand, W. and A. P. Kirman: *Introduction to Equilibrium Analysis*, Amsterdam: North-Holland(1976).

Hogg, R. V. and A. T. Craig: *Introduction to Mathematical Statistics*, 5th ed., Englewood Cliffs, N.J.: Prentice-Hall(1995).

Horn, R. A. and C. R. Johnson: *Matrix Analysis*, Cambridge: Cambridge University Press(1985).

Huang, Chi-fu and R. H. Litzenberger: *Foundations for Financial Economics*, Amsterdam: North-Holland(1988).

Intriligator, M. D.: *Mathematical Optimization and Economic Theory*, Englewood Cliffs, N.J.: Prentice-Hall(1971).

Johansen, L.: *Production Functions*, Amsterdam: North-Holland(1972).

Johnson, N. L., S. Kotz, and S. Kemp: *Univariate Discrete Distributions*, 2nd ed., New York: John Wiley & Sons(1993).

Johnson, N. L., S. Kotz, and N. Balakrishnan: *Continuous Univariate Discrete Distributions*, New York: John Wiley & Sons(1995).

Johnson, N. L., S. Kotz, and N. Balakrishnan: *Discrete Multivariate Distributions*, New York: John Wiley & Sons(1997).

Johnson, R. A. and G. K. Bhattacharyya: *Statistics: Principles and Meth-*

ods, 3rd ed., New York: John Wiley & Sons(1996).

Kamien, M. I. and N. I. Schwartz: *Dynamic Optimization: the Calculus of Variations and Optimal Control in Economics and Management*, 2nd ed., Amsterdam: North-Holland(1991).

Karatzas, I. and S. E. Shreve: *Brownian Motion and Stochastic Calculus*, 2nd ed., New York: Springer(1991).

Kolmogorov, A. N. and S. V. Fomin: *Introductory Real Analysis*, New York: Dover Publications(1975).

Kreps, D. M.: *A Course in Microeconomic Theory*, Princeton: Princeton University Press(1990).

Lang, S.: *Linear Algebra*, 3rd ed., New York: Springer(1987).

Larsen, R. J. and M. L. Marx: *An Introduction to Mathematical Statistics and its Applications*, Englewood Cliffs, N.J.: Prentice-Hall(1986).

Léonard, D. and N. Van Long: *Optimal Control Theory and Static Optimization in Economics*, Cambridge: Cambridge University Press(1992).

Luenberger, D. G.: *Introduction to Linear and Nonlinear Programming*, 2nd ed., Reading, Mass.: Addison-Wesley(1984).

Lütkepohl, H.: *Handbook of Matrices*, New York: John Wiley & Sons (1996).

Magnus, J. R. and H. Neudecker: *Matrix Differential Calculus with Applications in Statistics and Econometrics*, New York: John Wiley & Sons (1988).

Marsden, J. E. and M. J. Hoffman: *Elementary Classical Analysis*, 2nd ed., San Francisco: W. H. Freeman and Company(1993).

Mas-Colell, A., M. D. Whinston, and J. R. Green: *Microeconomic Theory*, Oxford: Oxford University Press(1995).

Merton, R. C.: "Theory of rational option pricing", *Bell Journal of Economics and Management Science*, Vol. 4, 141-183(1973).

Nikaido, H.: *Convex Structures and Economic Theory*, London: Academic Press(1968).

Nikaido, H.: *Introduction to Sets and Mappings in Modern Economics*, Amsterdam: North-Holland(1970).

Norstrøm, C. J.: "A sufficient condition for a unique nonnegative internal rate of return", *Journal of Financial and Quantitative Analysis*, Vol. 7, 1835-1839(1972).

Φksendal, B.: *Stochastic Differential Equations, an Introduction with Applications*, 5th ed., Berlin: Springer(1998).

Parthasarathy, T.: *On Global Univalence Theorems*. Lecture Notes in Mathematics. No.977, Berlin: Springer(1983).

Phlips, L.: *Applied Consumption Analysis*, Amsterdam: North-Holland (1983).

Pontryagin, L.S.: *Ordinary Differential Equations*, Reading, Mass.: Addison-Wesley(1962).

Rockafellar, T.: *Convex Analysis*, Princeton: Princeton University Press (1970).

Rothschild, M. and J. Stiglitz: "Increasing risk: (1) A definition", *Journal of Economic Theory*, Vol. 2, 225-243(1970).

Royden, H. L.: *Real Analysis*, 3rd ed., New York: Macmillan(1968).

Rudin, W.: *Principles of Mathematical Analysis*, 2nd ed., New York: McGraw-Hill(1982).

Scarf, H. (with the collaboration of T. Hansen): *The Computation of Economic Equilibria*. Cowles Foundation Monograph, 24, New Haven: Yale University Press(1973).

Seierstad, A. and K. Sydsæter: *Optimal Control Theory with Economic Applications*, Amsterdam: North-Holland(1987).

Sharpe, W. F.: "Capital asset prices: A theory of market equilibrium under conditions of risk", *Journal of Finance*, Vol. 19, 425-442(1964).

Shephard, R. W.: *Cost and Production Functions*, Princeton: Princeton University Press(1970).

Silberberg, E.: *The Structure of Economics. A Mathematical Analysis*, 2nd ed., New York: McGraw-Hill(1990).

Simon, C.P. and L. Blume: *Mathematics for Economists*, New York: Norton (1994).

Sneddon, I. N.: *Elements of Partial Differential Equations*, New York: McGraw-Hill(1957).

Stokey, N. L. and R. E. Lucas, with E. C. Prescott: *Recursive Methods in*

Economic Dynamics, Cambridge, Mass.: Harvard University Press(1989).

Sundaram, R. K.: *A First Course in Optimization Theory*, Cambridge: Cambridge University Press(1996).

Sydsæter, K. and P. J. Hammond: *Mathematics for Economic Analysis*, Englewood Cliffs, N.J.: Prentice-Hall(1995).

Takayama, A.: *Mathematical Economics*, 2nd ed., Cambridge: Cambridge University Press(1985).

Topkis, Donald M.: *Supermodularity and Complementarity*, Princeton: Princeton University Press(1998).

Turnbull, H. W.: *Theory of Equations*, 5th ed., Edinburgh: Oliver & Boyd (1952).

Varian, H.: *Microeconomic Analysis*, 3rd ed., New York: Norton(1992).

Weibull, J. W.: *Evolutionary Game Theory*, Cambridge: MIT Press(1995).

Zachmanoglou, E. C. and D. W. Thoe: *Introduction to Partial Differential Equations with Applications*, New York: Dover Publications(1986).