数学学习笔记

ZHANG Yang

2017 年 12 月 5 日

目录

Part I

高等数学

Chapter 1

函数与极限

初等数学的研究对象基本上是不变的量，而高等数学的研究对象则是变动的量。所谓函数关系就是变量之间的依赖关系，极限方法是研究变量的一种基本方法。本章将介绍映射、函数、极限和函数的连续性等基本概念以及他们的一些性质。

1.1 映射与函数

映射是现代数学中的一个基本概念，而函数是微积分的研究对象，也是映射的一种。本节主要介绍映射、函数及有关概念，函数的性质与运算等。

1.1.1 映射

1.1.1.1 映射概念

定义 1. 设 X、Y 是两个非空集合，如果存在一个法则 f，使得对 X 中每个元素 x，按法则 f，在 Y 中有唯一确定的元素 y 与之对应，那么称 f 为从 X 到 Y 的映射，记作

$$f : X \to Y$$

其中 y 称为元素 x 的像，并记作 $f(x)$，即

$$y = f(x)$$

而元素 x 称为元素 y 的一个原像；集合 X 称为映射 f 的定义域，记作 D_f，即 $D_f = X$；X 中所有元素的像所组成的集合称为映射 f 的值域，记作 R_f 或 $f(x)$，即

$$R_f = f(X) = \{f(x)|x \in X\}$$

从上述映射的定义中，需要注意的是：

(1) 构成一个映射必须具备以下三个要素：(i) 集合 X，即定义域 $D_f = X$；(ii) 集合 Y，即值域的范围：$R_f \in Y$；(iii) 对应法则 f，使对每个 $x \in X$，有唯一确定的 $y = f(x)$ 与之对应 。

(2) 对每个 $x \in X$，元素 x 的像 y 是唯一的；而对每个 $y \in Y$，元素 y 的原像不一定是唯一的；映射 f 的值域 R_f 是 Y 的一个子集，即 $R_f \subset Y$，不一样 $R_f = Y$。

1.1.2 函数

1.1.2.1 函数的概念

定义 2. 设数集 $D \subset R$，则称映射 $f : D \to R$ 为定义在 D 上的函数，通常简记为

$$y = f(x), x \in D$$

其中 x 称为自变量，y 称为因变量，D 称为定义域，记作 D_f，即 $D_f = D$。

函数的定义中，对每个 $x \in D$，按对应法则 f，总有唯一确定的值 y 与之对应，这个值称为函数 f 在 x 处的函数值，记作 $f(x)$，即 $y = f(x)$。因变量 y 与自变量 x 之间的这种依赖关系，通常称为函数关系。函数值 $f(x)$ 的全体所构成的集合称为函数 f 的值域，记作 R_f 或 $f(D)$，即

$$R_f = f(D) = \{y \mid y = f(x)，x \in D\}$$

1.1.2.2 函数的几种特性

1.1.2.2.1 函数的有界性

1.1.2.2.2 函数的单调性

1.1.2.2.3 函数的奇偶性

1.1.2.2.4 函数的周期性

定义 3. 设函数 $f(x)$ 的定义域为 D，如果存在一个正数 l，使得对于任意 $x \in D$ 有 $(x \pm l) \in D$，且

$$f(x + l) = f(x) \tag{1.1}$$

恒成立，那么 $f(x)$ 为周期函数，l 称为 $f(x)$ 的周期，通常我们说周期函数的周期是指最小正周期。

定义 4. 狄利克雷（*Dirichlet*）函数

$$D(x) = \begin{cases} 1, & x \in Q, \\ 0, & x \in Q^C。 \end{cases} \tag{1.2}$$

很容易验证狄利克雷函数（1.2）是一个周期函数，任何正有理数 r 都是它的周期，因为不存在最小的正有理数，所以他没有最小正周期。

1.1.2.3 反函数与复合函数

1.1.2.4 函数的运算

1.1.2.5 初等函数

在初等数学中已经将过下面几类函数：

幂函数：$y = x^\mu$，$(\mu \in \mathrm{R}$ 是常数$)$，

指数函数：$y = a^x$，$(a > 0$且$a \neq 1)$，

对数函数：$y = \log_a x$，$(a > 0$且$a \neq 1$，特别当$a = e$时，记为$y = \ln x)$，

三角函数：如 $y = \sin x$、$y = \cos x$、$y = \tan x$ 等，

反三角函数：如 $y = \arcsin x$、$y = \arccos x$、$y = \arctan x$ 等。

以上这五类函数统称基本初等函数。

由常数和基本初等函数经过有限次的四则运算或有限次的函数复合步骤所构成的、并可用一个式子表示的函数，称为初等函数。

应用上常遇到以 e 为底数的指数函数 $y = e^x$ 和 $y = e^{-x}$ 所产生的双曲函数以及它们的反函数—反双曲函数，它们的定义如下：

双曲正弦：$\sinh x = \dfrac{e^x - e^{-x}}{2}$

双曲余弦：$\cosh x = \dfrac{e^x + e^{-x}}{2}$

双曲正切：$\tanh x = \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$

1.2 数列的极限

一般地，有如下数列极限的定义：

定义 5. 设 $\{x_n\}$ 为一数列，如果存在常数 a，对于任意给定的正数 ε(不论它多么小)，总存在正整数 N，使得当 $n>N$ 时，不等式

$$|x_n - a| < \varepsilon \tag{1.3}$$

都成立，那么就称常数 a 时数列 $\{x_n\}$ 的极限，或者称数列 $\{x_n\}$ 收敛于 a，记为

$$\lim_{n\to\infty} x_n = a \tag{1.4}$$

或

$$x_n \to a \ (n \to \infty) \tag{1.5}$$

1.2.1 数列极限的定义

待补充

1.2.2 收敛数列的性质

待补充

1.3 函数的极限

待补充

1.3.1 函数极限的定义

待补充

1.3.1.1 自变量趋于有限值时函数的极限

待补充

1.3.1.2 自变量趋于无穷大时函数的极限

待补充

1.3.2 函数极限的性质

待补充

1.4 无穷大与无穷小

待补充

1.4.1 无穷小

待补充

1.4.2 无穷大

待补充

1.5 极限运算法则

待补充

1.6 极限存在准则两个重要极限

待补充

1.7 无穷小的比较

待补充

1.8 函数的连续性与间断点

待补充

1.8.1 函数的连续性

待补充

1.8.2 函数的间断点

待补充

1.9 连续函数的运算与初等函数的连续性

待补充

1.9.1 连续函数的和、差、积、商的连续性

待补充

1.9.2 反函数与复合函数的连续性

待补充

1.9.3 初等函数的连续性

待补充

1.10 闭区间上函数的连续性

待补充

1.10.1 有界性与最大值最小值定理

待补充

1.10.2 零点定理与介值定理

待补充

1.10.3 一致连续性

待补充

Chapter 2

导数与微分

Chapter 3

微分中值定理与导数的应用

Chapter 4

积分表

4.1 基本积分表

$$\int x^n \, \mathrm{d}x = \frac{1}{n+1} x^{n+1}, \ \ n \neq -1 \tag{4.1}$$

$$\int \frac{1}{x} \, \mathrm{d}x = \ln |x| \tag{4.2}$$

$$\int u \ \mathrm{d}v = uv - \int v du \tag{4.3}$$

$$\int e^x \, \mathrm{d}x = e^x \tag{4.4}$$

$$\int a^x \, \mathrm{d}x = \frac{1}{\ln a} a^x \tag{4.5}$$

$$\int \ln x \, \mathrm{d}x = x \ln x - x \tag{4.6}$$

$$\int \sin x \, \mathrm{d}x = -\cos x \tag{4.7}$$

$$\int \cos x \, \mathrm{d}x = \sin x \tag{4.8}$$

$$\int \tan x \, \mathrm{d}x = \ln|\sec x| \tag{4.9}$$

$$\int \sec x \, \mathrm{d}x = \ln|\sec x + \tan x| \tag{4.10}$$

$$\int \sec^2 x \, \mathrm{d}x = \tan x \tag{4.11}$$

$$\int \sec x \tan x \, \mathrm{d}x = \sec x \tag{4.12}$$

$$\int \frac{a}{a^2 + x^2} \, \mathrm{d}x = \tan^{-1} \frac{x}{a} \tag{4.13}$$

$$\int \frac{a}{a^2 - x^2} \, \mathrm{d}x = \frac{1}{2} \ln \left| \frac{x + a}{x - a} \right| \tag{4.14}$$

$$\int \frac{1}{\sqrt{a^2 - x^2}} \, \mathrm{d}x = \sin^{-1} \frac{x}{a} \tag{4.15}$$

$$\int \frac{a}{x\sqrt{x^2 - a^2}} \, \mathrm{d}x = \sec^{-1} \frac{x}{a} \tag{4.16}$$

$$\begin{aligned} \int \frac{1}{\sqrt{x^2 - a^2}} \, \mathrm{d}x &= \cosh^{-1} \frac{x}{a} \\ &= \ln(x + \sqrt{x^2 - a^2}) \end{aligned} \tag{4.17}$$

$$\begin{aligned} \int \frac{1}{\sqrt{x^2 + a^2}} \, \mathrm{d}x &= \sinh^{-1} \frac{x}{a} \\ &= \ln(x + \sqrt{x^2 + a^2}) \end{aligned} \tag{4.18}$$

4.2 含有 $ax + b$ 的积分

4.3 含有 $\sqrt{ax + b}$ 的积分

Part II

线性代数

Part III

概率论与数理统计

Chapter 5

随机变量及其分布

5.1 随机变量

现在来讨论如何引入一个法则，将随机实验的每一个结果，即将 S 的每个元素 e 与实数 x 对应起来，从而引入了随机变量的概念。

定义 6. 设随机试验的样本空间为 $S=\{e\}$，$X=X(e)$ 是定义在样本空间 S 上的实值单值函数，称 $X=X(e)$ 为随机变量 [1]。

我们一般以大写的字母如 X、Y、Z、W、$\cdots$ 表示随机变量，而以小写字母 x、y、z、w、$\cdots$ 表示实数。

随机变量的取值随实验的结果而定，而试验的各个结果出现有一定的概率，因而随机变量的取值有一定的概率。

随机变量的取值随试验的结果而定，在试验之前不能预知它取什么值，且它的取值有一定的概率。这些性质显示了随机变量与普通函数有着本质的差异。

随机变量的引入，使我们能用随机变量来描述各种随机现象，并能利用数学分析的方法对随机试验的结果进行深入广泛的研究和讨论。

5.2 离散型随机变量及其分布律

有些随机变量，它全部可能取到的值是有限个或可列无限多个，这种随机变量称为离散型随机变量。

[1] 严格地说，“对于任意实数 x，集合 $\{e|X(e)\leq x\}$（即：使得 $X(e)\leq x$ 的所有样本点 e 所组成的集合）有确定的概率” 这一要求应包括在随机变量的定义之中，一般来说，不满足这一条件的情况，在实际应用中是很少遇到的。因此，我们在定义中未提及这一要求。

容易知道，要掌握一个离散型随机变量 X 的统计规律，必须且只需知道 X 的所有可能取值以及取每一个可能值的概率。

设离散型随机变量 X 所有可能取的值为 $x_k(k=1,2,\cdots)$,X 取各个可能值的概率,即事件 $\{X=x_k\}$，为

$$P\{X=x_k\}=p_k，k=1,2,\cdots。 \tag{5.1}$$

由概率的定义可知，p_k 满足如下两个条件：

1. $p_k \geq 0$，$k=1,2,\cdots$；
2. $\sum\limits_{k=1}^{\infty} p_k = 1$。

下面介绍三种重要的离散型随机变量。

5.2.1 $(0-1)$ 分布

5.2.2 伯努利试验、二项分布

定义 7. 设试验 E 只有两个可能结果：A 及 $\bar{A}$，则 E 为**伯努利（*Bernoulli*）试验**。设 $P(A)=p(0<p<1)$,此时 $P(\bar{A})=1-p$。将 E 独立重复地进行 n 次,则称这一串重复的独立试验为 ***n* 重伯努利试验**。

这里"重复"是指在每次试验中 $P(A)=p$ 保持不变；"独立"是指各次试验的结果互不影响。

以 X 表示 n 重伯努利试验中事件 A 发生的次数，X 是一个随机变量，我们来求它的分布律。X 所有可能取的值为 0，1，2，$\cdots$，n。由于各次试验是相互独立的，因此事件 A 在指定的 $k(0\leq k\leq n)$ 次试验中发生，在其他 $n-k$ 次试验中 A 不发生（例如在前 k 次试验中 A 发生，而后 $n-k$ 次试验中 A 不发生）的概率为

$$\underbrace{p\cdot p\cdot\cdots\cdot p}_{k个}\cdot\underbrace{(1-p)\cdot(1-p)\cdot\cdots\cdot(1-p)}_{n-k个}=p^k(1-p)^{n-k}。 \tag{5.2}$$

这种指定的方式共有 $\binom{n}{k}$ 种。它们是两两互不相容的，故在 n 次试验中 A 发生 k 次的概率为 $\binom{n}{k}p^k(1-p)^{n-k}$，记 $q=1-p$，即有

$$P\{X=k\}=\binom{n}{k}p^kq^{n-k}，k=0,1,2,\cdots,n。 \tag{5.3}$$

显然

$$P\{X=k\}\geq 0，k=0,1,2,\cdots,n; \tag{5.4}$$

$$\sum_{k=0}^{n}P\{X=k\}=\sum_{k=0}^{n}\binom{n}{k}p^kq^{n-k}=(p+q)^n=1。 \tag{5.5}$$

注意到 $\binom{n}{k}p^kq^{n-k}$ 刚好是二项式 $(p+q)^n$ 的展开式中出现 p^k 的那一项，我们称随机变量 X 服从参数 n，p 的二项分布，并记为 $X \sim b(n,p)$。

特别地，当 $n=1$ 时二项分布（5.3）化为

$$P\{X=k\} = p^kq^{1-k}，k=0,1。\tag{5.6}$$

这就是 $(0-1)$ 分布。

5.2.3 泊松分布

定义 8. 设随机变量 X 所有可能取的值为 0，1，2，$\cdots$，而取各个值的概率为

$$P\{X=k\} = \frac{\lambda^k e^{-\lambda}}{k!}，k=0,1,2,\cdots,\tag{5.7}$$

其中，$\lambda>0$ 是常数，则称 X 服从参数为 λ 的泊松分布，记为 $X\sim\pi(\lambda)$。

5.3 随机变量的分布函数

5.4 连续型随机变量及其概率密度

如果对于随机变量 X 的分布函数 $F(x)$，存在非负可积函数 $f(x)$，使对于任意实数 x 有

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t,\tag{5.8}$$

则称 X 为连续型随机变量，$f(x)$ 称为 X 的概率密度函数，简称概率密度。

5.4.1 均匀分布

5.4.2 指数分布

5.4.3 正态分布

定义 9. 若连续型随机变量 X 的概率密度为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}，\ -\infty<x<\infty,\tag{5.9}$$

其中 μ，$\sigma(\sigma>0)$ 为常数，则称 X 服从参数为 μ，σ 的正态分布或高斯（*Gauss*）分布，记为 $X\sim N(\mu,\sigma^2)$。

显然 $f(x) \geqslant 0$，下面来证明 $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$。令 $\dfrac{x-\mu}{\sigma} = t$，得到

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}\,\mathrm{d}t$$

记 $I = \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}\,\mathrm{d}t$，则有 $I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{t^2+u^2}{2}}\,\mathrm{d}t\,\mathrm{d}u$，利用极坐标将它化成累次积分，得到

$$I^2 = \int_0^{2\pi}\int_0^{\infty} r e^{-\frac{r^2}{2}}\,\mathrm{d}r\,\mathrm{d}\theta。$$

而 $I > 0$，故有 $I = \sqrt{2\pi}$，即有

$$\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}\,\mathrm{d}t = \sqrt{2\pi} \tag{5.10}$$

于是 $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}\,\mathrm{d}t = 1$。

$f(x)$ 具有以下性质：

1. 曲线关于 $x = \mu$ 对称。这表明对于任意 $h > 0$ 有

$$P\{\mu - h < X \leqslant \mu\} = P\{\mu < X \leqslant \mu + h\}$$

2. 当 $x = \mu$ 时取得最大值

$$f(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$$

x 离 μ 越远，$f(x)$ 的值越小。这表明对于同样长度的区间，当区间离 μ 越远，X 落在这个区间上的概率越小。

在 $x = \mu \pm \sigma$ 处曲线有拐点，曲线以 Ox 轴为渐进线。

特别，当 $\mu = 0, \sigma = 1$ 时称随机变量 X 服从**标准正态分布**。其概率密度和分布函数分别用 $\varphi(x)$，$\varPhi(x)$ 表示，即有

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \tag{5.11}$$

$$\varPhi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}}\,\mathrm{d}t \tag{5.12}$$

易知

$$\varPhi(-x) = 1 - \varPhi(x) \tag{5.13}$$

引理 1. 若 $X \sim N(\mu, \sigma^2)$，则 $Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$。

证明. $Z = \dfrac{X-\mu}{\sigma}$ 的分布函数为

$$\begin{aligned} P\{Z \le x\} &= P\left\{Z \le \frac{X-\mu}{\sigma}\right\} \\ &= P\{X \le \mu + \sigma x\} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mu+\sigma x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \, \mathrm{d}t, \end{aligned}$$

令 $\dfrac{t-\mu}{\sigma} = u$，得

$$P\{Z \le x\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} \, \mathrm{d}u = \Phi(x),$$

由此知 $Z = \frac{X-u}{\sigma} \sim N(0,1)$。 □

于是，若 $X \sim N(\mu, \sigma^2)$，则它的分布函数 F(x) 可写成

$$F(x) = P\{X \leqslant x\} = P\left\{\frac{X-\mu}{\sigma} \leqslant \frac{x-\mu}{\sigma}\right\} = \Phi(\frac{x-\mu}{\sigma}) \tag{5.14}$$

对于任意区间 $(x_1, x_2]$，有

$$P\{x_1 < X \leqslant x_2\} = P\left\{\frac{x_1-\mu}{\sigma} < \frac{X-\mu}{\sigma} \leqslant \frac{x_2-\mu}{\sigma}\right\} \tag{5.15}$$

$$= \Phi(\frac{x_2-\mu}{\sigma}) - \Phi(\frac{x_1-\mu}{\sigma}) \tag{5.16}$$

尽管正态变量的取值范围是 $(-\infty, \infty)$，但它的值落在 $(\mu - 3\sigma, \mu + 3\sigma)$ 内几乎是肯定的事，这就是人们所说的 “3σ” 法则。

为了便于今后在数理统计中的应用，对于标准正态随机变量，我们引入了上 α 分位点的定义。

定义 10. 设 $X \sim N(0,1)$，若 z_a 满足条件

$$P\{X > z_a\} = a, \ 0 < a < 1, \tag{5.17}$$

则称点 z_a 为标准正态分布的**上 α 分位点**。

5.5 随机变量的函数的分布

定理 1. 设随机变量 X 具有密度函数 $f_X(x), -\infty < x < \infty$，又设函数 $g(x)$ 处处可导且恒有 $g'(x) > 0$（或恒有 $g'(x) < 0$），则 $Y = g(X)$ 是连续型随机变量，其概率密度为

$$f_Y(y) = \begin{cases} f_X[h(y)]\,|h'(y)|, & \alpha < y < \beta \\ 0, & \text{其他} \end{cases} \tag{5.18}$$

其中 $\alpha = \min\{g(-\infty), g(\infty)\}$，$\beta = \max\{g(-\infty), g(\infty)\}$，$h(y)$ 是 $g(x)$ 的反函数。

Chapter 6

多维随机变量及其分布

Chapter 7

大数定律及中心极限定理

极限定理是概率论的基本理论，在理论研究和应用中起着重要的作用，其中最重要的是称为“大叔定律”与“中心极限定理”的一些定理。大数定律是叙述随机变量序列的前一些项的算术平均值在某种条件下收敛到这些项的均值的算术平均值；中心极限定理则是确定在什么条件下，大量随机变量之和的分布逼近于正态分布。

7.1 大数定律

大量试验证实，随机事件 A 的频率 $f_n(A)$ 当重复试验的次数 n 增大时总呈现出稳定性，稳定在某一个常数的附件。频率的稳定性是概率定义的客观基础。

定理 2 (弱大数定理，辛钦大数定理)**.** 设 X_1，X_2，$\cdots$ 是相互独立，服从同一分布的随机变量序列，且具有数学期望 $E(X_k)=\mu\ (k=1,2,\cdots)$。作前 n 个变量的算术平均 $\dfrac{1}{n}\sum\limits_{k=1}^{n}X_k$，则对于任意 $\epsilon>0$，有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu| < \epsilon\right|\right\}=1 \tag{7.1}$$

定理 3 (伯努利大数定理)**.** 设 f_A 是 n 次独立重复试验中事件 A 发生的次数，p 是事件 A 在每次试验中发生的概率，则对于任意正数 $\epsilon>0$，有

$$\lim_{n\to\infty}P\left\{\left|\frac{f_A}{n}-p\right|<\epsilon\right\}=1 \tag{7.2}$$

或

$$\lim_{n\to\infty}P\left\{\left|\frac{f_A}{n}-p\right|\geqslant\epsilon\right\}=0 \tag{7.3}$$

7.2 中心极限定理

定理 4 (独立同分布的中心极限定理). 设随机变量

定理 5 (李雅普诺夫（Lyapunov）定理). 设随机变量

定理 6 (棣莫弗—拉普拉斯（De Moivre—Laplace）定理). 设随机变量