量子计算与量子信息 Notes

Flower CA77

目 录

第一部分　基础概念

第一章　简介与概述

1.1 全貌

1.2 量子比特

1.3 量子计算

1.4 量子算法

1.5 实验量子信息处理

1.6 量子信息

第二章　量子力学引论

2.1 线性代数

2.2 量子力学的假设

2.3 应用: 超密编码

2.4 密度算子

2.5 Schmidt 分解与纯化

2.6 EPR 和 Bell 不等式

第三章　计算机科学简介

3.1 计算模型

3.2 计算问题的分析

3.3 关于计算科学的观点

第二部分　量子计算

第四章　量子电路

4.1 量子算法

4.2 单量子比特操作

4.3 受控操作

4.4 测量

4.5 通用量子门

4.6 量子计算电路模型总结

4.7 量子系统的模拟

第五章　量子 Fourier 变换及其应用

5.1 量子 Fourier 变换

5.2 相位估计

5.3 应用: 求阶与因子分解问题

5.4 量子 Fourier 变换的一般应用

第六章　量子搜索算法

6.1 量子搜索算法

6.2 作为量子模拟的量子搜索

6.3 量子计数

6.4 NP 完全问题解的加速

6.5 无结构数据库的量子搜索

6.6 搜索算法的最优性

6.7 黑盒算法的极限

第七章　量子计算机: 物理实现

7.1 指导性原则

7.2 量子计算的条件

7.3 谐振子量子计算机

7.4 光学光量子计算机

7.5 光学腔量子电动力学

7.6 离子阱

7.7 核磁共振

7.8 其他实现方案

第三部分　量子信息

第八章　量子噪声与量子操作

8.1 经典噪声与 Markov 过程

8.2 量子操作

8.3 量子噪声与量子操作的例子

8.4 量子操作的应用

8.5 量子操作形式体系的局限

第九章　量子信息的距离度量

9.1 经典信息的距离度量

9.2 两个量子态有多接近

9.3 量子信道保护信息的效果怎么样

第十章　量子纠错

10.1 背景介绍

10.2 Shor 编码

10.3 量子纠错理论

10.4 构造量子编码

10.5 稳定子编码

10.6 容错量子计算

第十一章　熵与信息

11.1 Shannon 熵

11.2 熵的基本性质

11.3 von Neumann 熵

11.4 强次可加性

第十二章　量子信息论

12.1 量子态的区分与可达信息

12.2 数据压缩

12.3 噪声信道上的经典信息

12.4 有噪声量子信道的量子信息

12.5 作为一种物理资源的纠缠

12.6 量子密码学

附录

附录 A　概率论基础

随机变量 X 取值 x 的概率为 $p(X=x)$:

- 如果 X 是离散型随机变量, 则 $p(X=x)$ 为 $X=x$ 的概率
- 如果 X 是连续型随机变量, 则 $p(X=x)=0$, 因为一个点在一个连续区域内的测度为零, 此时我们引入概率密度 $\rho(x)$, 使得 $p(x\leqslant X<x+dx)=\rho(x)dx$

我们简记 $p(X=x)$ 为 $p_X(x)$, 不引起混淆时进一步简记为 $p(x)$。

常数 c 可以看成取值 $X=c$ 概率为 1 的随机变量 X。

概率密度的概念是始终有效的, 对离散型随机变量 X, 我们可以取 $\rho(x)=\sum\limits_{x'}p(X=x')\delta(x-x')$, 其中求和 x' 取遍 X 的所有值, δ 为 Dirac 的 δ 函数。进一步, 当 x' 不是 X 可取的值时 $p(X=x')=0$, 求和 x' 可以取遍全部值。

一组随机变量 $X_1,X_2,\cdots,X_n$ 组成随机向量 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$, 联合分布 $p(\boldsymbol{X}=\boldsymbol{x})=p(X_1=x_1,X_2=x_2,\cdots,X_n=x_n)$ 也简记为 $p_{\boldsymbol{X}}(\boldsymbol{x})$, 不引起混淆时简单记作 $p(\boldsymbol{x})$。

条件概率 $p(Y=y\mid X=x)=p(X=x,Y=y)/p(X=x)$, 简记为 $p_{Y|X}(y\mid x)=p_{(X,Y)}(x,y)/p_X(x)$, 其中分子是随机向量 (X,Y) 的联合分布。不引起混淆时我们简单记作 $p(y\mid x)$。

注意　在略去概率 p 的角标时 [即简记 $p_X(x)$ 为 $p(x)$] 必须要规范标记变量, 即用单个大写字母 X 表示随机变量, 其小写值 x 表示对应的 X 的取值, 这个规定也是 $p(x)$ 的缺省标准。在表达式比较复杂时, 可以显式写出随机变量 $p(X=x)$。

习题 (教材 A.1)　证明 Bayes 定律 $p(x\mid y)=p(y\mid x)\dfrac{p(x)}{p(y)}$

证明　把待证等式改写为 $p(x\mid y)p(y)=p(y\mid x)p(x)$，可以看到等式两边都是联合分布概率 $p(x,y)$，这就证明了待证方程。

习题 (教材 A.2)　全概率公式 $p(y)=\sum\limits_x p(y\mid x)p(x)$

证明　$\sum\limits_x p(y\mid x)p(x)=\sum\limits_x p(x,y)=p(y)$

期望 $\mathbb{E}X\equiv\sum\limits_x xp(x)$, 方差 $\operatorname{Var}X\equiv\mathbb{E}\left[(X-EX)^2\right]=\mathbb{E}\left(X^2\right)-(\mathbb{E}X)^2$, 标准差 $\Delta X\equiv\sqrt{\operatorname{Var}X}$

习题 (教材 A.3)　证明 $\exists x\geqslant\mathbb{E}X,\ \text{s.t.}\ p(x)>0$

证明　反证，只要证明命题 $\forall x\geqslant\mathbb{E}X:p(x)=0$ 是伪命题即可。考虑到

$$\mathbb{E}X=\sum_x xp(x)=\sum_{x<\mathbb{E}X}xp(x)+\sum_{x\geqslant\mathbb{E}X}xp(x)=\sum_{x<\mathbb{E}X}xp(x)<\mathbb{E}X\sum_{x<\mathbb{E}X}p(x)\leqslant\mathbb{E}X\sum_x p(x)=\mathbb{E}X$$

习题 (教材 A.4)　证明 $\mathbb{E}X$ 对 X 是线性的。

证明　$\mathbb{E}(kX)=\sum\limits_x kxp(x)=k\sum\limits_x xp(x)=k\mathbb{E}X$

习题 (教材 A.5)　证明 X,Y 独立时 $\mathbb{E}(XY)=\mathbb{E}X\cdot\mathbb{E}Y$

证明　$\mathbb{E}(XY)=\sum\limits_{x,y}xyp(x,y)\overset{X,Y\text{ 独立}}{=\!=\!=}\sum\limits_{x,y}xyp(x)p(y)=\sum\limits_x xp(x)\sum\limits_y yp(y)=\mathbb{E}X\cdot\mathbb{E}Y$

习题 (教材 A.6, Cheybshev 不等式) $\forall\lambda>0$ 和有限方差的 X，$p(|x-\mathbb{E}X|\geqslant\lambda\Delta X)\leqslant\dfrac{1}{\lambda^2}$

证明 我们设概率密度为 $\rho(x)$，则

$$\begin{aligned}(\Delta X)^2=\operatorname{Var}X=\mathbb{E}\Big[(X-\mathbb{E}X)^2\Big]&=\int(x-\mathbb{E}X)^2\rho(x)dx\\&=\int_{x-\mathbb{E}X\leqslant-\lambda\Delta X}(x-\mathbb{E}X)^2\rho(x)dx+\int_{\mathbb{E}X-\lambda\Delta X}^{\mathbb{E}X+\lambda\Delta X}(x-\mathbb{E}X)^2\rho(x)dx+\int_{x-\mathbb{E}X\geqslant\lambda\Delta X}(x-\mathbb{E}X)^2\rho(x)dx\\&\geqslant\lambda^2\Delta X^2\int_{x-\mathbb{E}X\leqslant-\lambda\Delta X}\rho(x)dx+0+\lambda^2\Delta X^2\int_{x-\mathbb{E}X\geqslant\lambda\Delta X}\rho(x)dx=\lambda^2\Delta X^2\int_{|x-\mathbb{E}X|\geqslant\lambda\Delta X}\rho(x)dx\end{aligned}$$

则 $p(|X-\mathbb{E}X|\geqslant\lambda\Delta X)=\displaystyle\int_{|x-\mathbb{E}X|\geqslant\lambda\Delta X}\rho(x)dx\leqslant\frac{\Delta X^2}{\lambda^2\Delta X^2}=\frac{1}{\lambda^2}$

我们补充一些内容, 首先是关于方差的一些定义。

定义 A.1 (协方差) 两个随机变量 X,Y 的协方差定义为

$$\operatorname{Cov}(X,Y)=\mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)]=\mathbb{E}(XY)-\mathbb{E}X\cdot\mathbb{E}Y$$

容易看出同一个随机变量和其自身的协方差即为其自身的方差 $\operatorname{Cov}(X,X)=\operatorname{Var}X=(\Delta X)^2$, 两个变量的协方差是对称的 $\operatorname{Cov}(X,Y)=\operatorname{Cov}(Y,X)$。

定义 A.2 (相关系数) 两个随机变量 X,Y 的相关系数为其协方差除以两个变量的标准差

$$\rho_{X,Y}=\frac{\operatorname{Cov}(X,Y)}{\Delta X\cdot\Delta Y}=\frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}X}\sqrt{\operatorname{Var}Y}}$$

容易看出同一个随机变量和其自身是完全相关的 $\rho_{X,X}=1$, 两个变量的相关是对称的 $\rho_{X,Y}=\rho_{Y,X}$。

命题 A.1 两个随机变量 X,Y 相关系数 $\rho_{X,Y}=1\iff$ 两者与常数 1 线性相关 $\exists\lambda,\mu\in\mathbb{R}:\lambda X+\mu Y=1$, 此时我们称两个变量 **完全相关** 。若 $\rho_{X,Y}=0$, 我们称 X,Y 完全不相关。根据相关系数的符号 $\rho_{X,Y}>0$ 或 <0 时我们称 X,Y **正相关** 或 **负相关** 。

定义 A.3 对随机向量 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$, 其 **协方差矩阵** $\boldsymbol{C}=\operatorname{Cov}\boldsymbol{X}=[\operatorname{Cov}(X_i,X_j)]_{n\times n}$ 和 **相关系数矩阵** $\boldsymbol{\rho}=\rho_{\boldsymbol{X}}=\left[\rho_{X_i,X_j}\right]_{n\times n}$ 分别为

$$\begin{bmatrix}\operatorname{Var}X_1&\operatorname{Cov}(X_1,X_2)&\cdots&\operatorname{Cov}(X_1,X_n)\\\operatorname{Cov}(X_2,X_1)&\operatorname{Var}X_2&\cdots&\operatorname{Cov}(X_2,X_n)\\\vdots&\vdots&&\vdots\\\operatorname{Cov}(X_n,X_1)&\operatorname{Cov}(X_n,X_2)&\cdots&\operatorname{Var}X_n\end{bmatrix},\quad\begin{bmatrix}1&\rho_{X_1,X_2}&\cdots&\rho_{X_1,X_n}\\\rho_{X_2,X_1}&1&\cdots&\rho_{X_2,X_n}\\\vdots&\vdots&&\vdots\\\rho_{X_n,X_1}&\rho_{X_n,X_2}&\cdots&1\end{bmatrix}$$

然后我们来补充一些和期望有关的工具。

定义 A.4 (条件期望) 随机变量 X 在取值 x 的情况下随机变量 Y 的条件期望为 Y 在 $X=x$ 下条件概率的期望 $\mathbb{E}(Y\mid X=x)=\displaystyle\sum_y yp(Y=y\mid X=x)$, 简记为 $\mathbb{E}(Y\mid x)=\displaystyle\sum_y yp(y\mid x)$。

定义 A.5 (生成函数) 随机变量 X 的生成函数 $f_X(t)$ 或简记为 $f(X)$, 定义为 $f(X)=\mathbb{E}\big(e^{itX}\big)$, 将其展开可以发现生成函数事实上是概率的 Fourier 变换 $f(X)=\displaystyle\sum_x p(x)e^{itx}$, 当我们把概率写成概率密度的积分时可以更明显的看出 Fourier 变换的形式 $f(X)=\displaystyle\int\rho(x)e^{itx}dx$。

命题 A.2 若 X,Y 独立, 则 $Z=X+Y$ 的生成函数为 $f(Z)=f(X+Y)=f(X)f(Y)$

证明 $f(Z)=\mathbb{E}\big(e^{itZ}\big)=\mathbb{E}\Big(e^{it(X+Y)}\Big)=\mathbb{E}\big(e^{itX}e^{itY}\big)\overset{X,Y\text{ 独立}}{=\!=\!=}\mathbb{E}\big(e^{itX}\big)\mathbb{E}\big(e^{itY}\big)=f(X)f(Y)$

常数 c 的生成函数为 $\mathbb{E}(e^{itc})=e^{itc}$, 是一个转动的相位。

利用生成函数可以证明概率论中最重要的定理:

定理 (中心极限定理)　对均值 μ 标准差 σ 的 i.i.d.(独立同分布) 随机变量 $X_1,X_2,\cdots,X_n$, 当样本数 n 趋向无限大时有极限 $\lim\limits_{n\to\infty}p\left(\dfrac{\sum\limits_{i=1}^{n}X_i-n\mu}{\sigma\sqrt{n}}\leqslant x\right)=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{x}\exp\left(-\frac{1}{2}z^2\right)dz$

证明　设 $Z_n=\dfrac{\sum\limits_{i=1}^{n}X_i-n\mu}{\sigma\sqrt{n}}=\sum\limits_{i=1}^{n}\dfrac{X_i-\mu}{\sigma\sqrt{n}}$, 则

$$f(Z_n)=\prod_{i=1}^{n}f\left(\frac{X_i-\mu}{\sigma\sqrt{n}}\right)=\prod_{i=1}^{n}f\left(\frac{X_i}{\sigma\sqrt{n}}\right)\exp\left(-i\frac{\mu t}{\sigma\sqrt{n}}\right)$$

由于 $X_i\ (i=1,2,\cdots,n)$ i.i.d. , 故上式变成

$$f(Z_n)=\left[f\left(\frac{X}{\sigma\sqrt{n}}\right)\exp\left(-i\frac{\mu t}{\sigma\sqrt{n}}\right)\right]^n=\left[\mathbb{E}\exp\left(i\frac{Xt}{\sigma\sqrt{n}}\right)\exp\left(-i\frac{\mu t}{\sigma\sqrt{n}}\right)\right]^n=\left[\mathbb{E}\exp\left(i\frac{X-\mu}{\sigma\sqrt{n}}t\right)\right]^n$$

作 Taylor 展开得到

$$f(Z_n)=\left[1+i\mathbb{E}\frac{X-\mu}{\sigma\sqrt{n}}t-\mathbb{E}\frac{(X-\mu)^2}{2\sigma^2n}t^2+O\left(n^{-3/2}\right)\right]^n$$

注意到 $n\to\infty$ 时 $\mathbb{E}\overline{X}=\sum\dfrac{\mathbb{E}X}{n}\to\mu$, 也就是说 $n\to\infty$ 时生成函数里面的全部 $\mathbb{E}X$ 换成 μ 后得到的量是原来的等价量, 即

$$f(Z_n)\sim\left[1-\mathbb{E}\frac{(X-\mu)^2}{2\sigma^2n}t^2\right]^n\to\exp\left[-\mathbb{E}\frac{(X-\mu)^2}{2\sigma^2}t^2\right]$$

那么

$$f(Z)=f\left(\lim_{n\to\infty}Z_n\right)\sim\lim_{n\to\infty}f(Z_n)=\exp\left[-\mathbb{E}\frac{(X-\mu)^2}{2\sigma^2}t^2\right]=\exp\left(-\frac{1}{2}t^2\right)$$

正如我们提到的, 生成函数实际上是概率分布的 Fourier 变换, 即 $\displaystyle\int p(z)e^{itz}dz=\exp\left(-\frac{1}{2}t^2\right)$, 那么我们直接将生成函数作 Fourier 逆变换就能得到概率密度 $p(z)=\dfrac{1}{2\pi}\displaystyle\int\exp\left(-\frac{1}{2}t^2\right)e^{-itz}dt=\frac{1}{\sqrt{2\pi}}\exp\left(z^2\right)$, 即标准 Gauss 分布的概率密度。

这个定理事实上是说, 从均值 μ 标准差 σ 的统计数据中抽样 X, 当样本容量 n 足够大时样本的均值 $\overline{X}=\dfrac{1}{n}\sum X$ 将趋向于 Gauss 分布 $N\left(\mu,\dfrac{\sigma^2}{n}\right)$, 即 $\dfrac{1}{\sqrt{n}}\overline{X^*}=\dfrac{\sum X-n\mu}{\sigma\sqrt{n}}$ $\left(X^*=\dfrac{X-\mu}{\sigma}\text{ 为 }X\text{ 的标准化变量}\right)$ 的概率密度将趋向于标准 Gauss 分布的概率密度 $\dfrac{1}{\sqrt{2\pi}}\exp\left(-\dfrac{1}{2}t^2\right)$。均值的分布标准差相比样本自身的标准差相差倍数 $\dfrac{1}{\sqrt{n}}$, 这个系数就是所谓的 **经典测量极限** 。

附录 B　群论

B.1 基本定义

定义 B.1 (群)　(1) 封闭性 (2) 结合律 (3) 单位元 (4) 逆元

定义 B.2 (有限群)　若群 G 有限, 则其成员个数 $|G|$ 称为阶数。

定义 B.3 (Abel 群)　运算 **可交换** 的群, 如整数模 n 的加法群 $\mathbb{Z}_n$。

定义 B.4 (阶数)　若 $g \in G$, 使得 $g^r = e$ 的最小正整数 $r \in \mathbb{Z}_{>0}$ 称为其阶数。

定义 B.5 (子群)　$H \leqslant G$ 是指 $H \subset G$ 且 H 在 G 运算下构成群。容易看出单位元 $e \in G$。

习题 (教材 B.1)　证明有限群的成员都有阶数，即 $\forall g \in \text{有限群}, \exists r \in \mathbb{Z}_{>0},\ \text{s.t.}\ g^r = e$。

证明　若某个成员 g 没有阶数，则群 G 有无限大的子集 $\{g^r : r \in \mathbb{Z}_{>0}\}$，矛盾。进一步，我们知道群 G 的任何成员的阶数不超过 $|G|$。

习题 (教材 B.2, Lagrange 定理)　若 $H \leqslant$ 有限群 G，则 $|H|$ 可整除 $|G|$，除数 $[G:H] = \dfrac{|G|}{|H|}$ 称为子群 H 的 Lagrange 指数。

为了证明此定理我们需要引入一些概念:

定义 B.6 (陪集)　设 $H \leqslant g$, 集合 $gH = \{gh : h \in H\}$, $Hg = \{hg : h \in H\}$ 称为 g 对 H 的 **左陪集** 和 **右陪集** 。

命题 B.1　$gH = H \iff g \in H$

证明　$g \in H \implies gH = H$ 是显然的，反过来时注意到 $e \in H$，则 $g = ge \in gH = H$。

命题 B.2　$$g_1H \cap g_2H = \begin{cases} \text{非空集合} & (g_1H = g_2H) \\ \varnothing & (g_1H \neq g_2H) \end{cases}$$

证明　对 $\forall g \in g_1H$ 有 $g = g_1h = g_2\left(g_2^{-1}g_1h\right)$ $(h \in H)$。若 $\exists$成员 $g \in g_1H \cap g_2H$，则有 $h_1, h_2 \in H$ s.t. $g = g_1h_1 = g_2h_2 \implies g_2^{-1}g_1 = h_2h_1^{-1} \in H$，由此 $g = g_2\left(g_2^{-1}g_1h\right) \in g_2H$。类似的 $\forall g \in g_2H \implies g \in g_1H$，这就证明了 $g_1H = g_2H$。

命题 B.3　全部陪集的并 $\bigcup\limits_{g\in G} gH = G$。

证明　由 $gH \subset G$ 可知 $\bigcup\limits_{g\in G} gH \subset G$。

另一方面，由于 $e \in H$，故 $\forall g \in G, g = ge \in gH$，这说明 $\forall g \in G : G \subset gH \implies g \subset \bigcup\limits_{g\in G} gH$。

综合上述所论，$\bigcup\limits_{g\in G} gH \subset G \,\&\&\, g \subset \bigcup\limits_{g\in G} gH \implies G = \bigcup\limits_{g\in G} gH$。

证明 (教材习题 B.2, Lagrange 定理)　我们知道全部陪集 $\{gH : g \in G\}$ 是一组不交的集合，容易看出 $\bigcup \{gH : g \in G\} = G$，即 $G = \bigsqcup\{gH : g \in G\} = \bigsqcup gH$，这说明 $|G| = \sum |gH|$。容易证明 $|gH| = |H|$，这说明 $|G| = \sum |gH| = \sum |H| = |H| \sum 1$，由此命题得证。

习题 (教材 B.3) 证明每个成员 $g \in G$ 的阶数可以整除 $|G|$。

证明 令 $H = \{g^r : 1 \leqslant r \leqslant r_k\} = \{g, g^2, \cdots, g^{r_k-1}, g^{r_k} = e\}$，则容易看出 $H \leqslant G$，根据 Lagrange 定理，H 的指数 $[G:H] = \dfrac{|G|}{|H|} \in \mathbb{Z}_{>0}$，即 $|G| = [G:H] \cdot |H| = [G:H] r_k \implies r_k \mid |G|$。

定义 B.7 若 $\exists g \in G$, s.t. 群成员 $a, b \in G$ 满足 $b = g^{-1}ag$, 则称 a, b 为共轭成员。

命题 B.4 群成员间的共轭是等价关系。

证明 1. 任意成员 a 与其自身共轭：

$$a = e^{-1}ae$$

2. 若成员 a 与成员 b 共轭，则成员 b 与成员 a 共轭：

$$\left(b = g^{-1}ag\right) \implies \left[a = \left(g^{-1}\right)^{-1} b \left(g^{-1}\right)\right]$$

3. 若成员 a 与成员 b 共轭，成员 b 与成员 c 共轭，则成员 a 与成员 c 共轭：

$$\left(b = g^{-1}ag\right) \,\&\&\, \left(c = g'^{-1}bg'\right) \implies \left[c = g'^{-1}gagg' = (gg')^{-1}a(gg)\right]$$

定义 B.8 (正规子群) 若 $H \leqslant G$ 且 $\forall g \in G : g^{-1}Hg = H$ (或等价的 $Hg = gH$), 则称 H 为 G 的 **正规子群**，记作 $H \trianglelefteq G$, 记号 $H \triangleleft G$ 表示 $H \neq G$ 的正规子群。

我们指出, Abel 群的任何子群都是正规子群。设 $H \leqslant$ Abel 群 G, 由于子群 H 的成员也是其继承的 Abel 群的成员, 这些成员与 G 中任意成员 g 也是可交换的, 故 $Hg = gH$ 显然成立。

对群 G, $x \in G$ 的共轭类定义为 $G_x \equiv \left\{g^{-1}xg : g \in G\right\}$。容易看出 $y \in G_x \implies x \in G_y$。注意到

$$\begin{aligned} y \in G_x &\implies \exists g \in G : y = g^{-1}xg \\ &\implies \exists g \in G : x = gyg^{-1} = \left(g^{-1}\right)^{-1} y \left(g^{-1}\right) \\ &\xrightarrow{g^{-1} \in G} \exists g \in G : x = g^{-1}yg \implies x \in G_y \end{aligned}$$

习题 (教材 B.4) $y \in G_x \implies G_y = G_x$

证明 由于 $y \in G_x$，$\exists g_0 \in G : y = g_0^{-1}xg_0$ 或 $x = g_0 y g_0^{-1}$。现在设 $t \in G_y$，即 $\exists g \in G : t = g^{-1}yg = g^{-1}g_0^{-1}xg_0g = (g_0g)^{-1}x(g_0g) \implies t \in G_x$，这就证明了 $G_y \subset G_x$。反过来设 $t \in G_x$，即 $\exists g \in G : t = g^{-1}xg = g^{-1}g_0yg_0^{-1}g = \left(g_0^{-1}g\right)^{-1} y \left(g_0^{-1}g\right) \implies t \in G_y$。这就证明了 $G_x \subset G_y$。综合上述，我们有 $G_y = G_x$。

习题 (教材 B.5) $x \in$ Abel 群 $G \implies G_x = \{x\}$

定义 B.9 (生成元) 设 $g_1, g_2, \cdots, g_\ell \in$ 群 G, 则 G 中全部可以写成 $g_1, g_2, \cdots, g_\ell$ 中若干个成员之乘积的群成员构成 G 的一个子集, 叫做 $g_1, g_2, \cdots, g_\ell$ 所生成的子群 H, 记作 $H = \langle g_1, g_2, \cdots, g_\ell \rangle$, 即

$$\langle g_1, g_2, \cdots, g_\ell \rangle = \left\{ g_1^{n_1} g_2^{n_2} \cdots g_\ell^{n_\ell} = \prod_{k=1}^{\ell} g_k^{n_k} : (n_1, n_2, \cdots, n_\ell) \in (\mathbb{Z}_{\geqslant 0})^\ell \right\}$$

$g_1, g_2, \cdots, g_\ell$ 称为子群 $\langle g_1, g_2, \cdots, g_\ell \rangle$ 的生成元。

如果基群 G 是很大的或平凡的, 那么由给定成员生成的继承 G 的子群也称为 (G 的) 一个 **生成群**，我们可以略去 G 的表述。容易看出生成一个阶数为 n 的生成群至少需要 $\log(n)$ 个成员。

定义 B.10 (循环群) 单个成员生成的生成群称为这个成员的**循环群**。容易看出循环群的阶数等于其生成元的阶数。

习题 (教材 B.6) 阶数为素数的群是循环群。

为了证明此命题, 我们需要如下结论:

命题 B.5 生成群的阶数等于其某个生成元的阶数 $\Longleftrightarrow$ 此群为循环群

证明 ($\Longleftarrow$) 是明显的, 现在来证明 ($\Longrightarrow$)。设 $G=\langle g_1,g_2,\cdots,g_n\rangle$, 其中 g_1 的阶数为 $|G|$, 我们的目标是证明 $G=\langle g_1\rangle$。明显 $\langle g_1\rangle\subset\langle g_1,g_2,\cdots,g_n\rangle=G$, 我们注意到由于 g_1 的阶数为 $|G|$, 则 $\langle g_1\rangle$ 必然包含 e, g_1, g_1^2, $\cdots$, $g_1^{|G|-1}$ 这 $|G|$ 个不同的成员, 也就是说 $|\langle g_1\rangle|=|G|$, 这说明 $\langle g_1\rangle=G$。

证明 (教材习题 B.6) 若群 G 的阶数 $|G|$ 为素数, 考虑继承 G 的子群 $H\leqslant G$, 则根据 Lagrange 定理可知其阶数 $|H|$ 应满足 $|G|=[G:H]\cdot|H|$。由于 $|G|$ 是素数, H 的阶数和对基群 G 的指数均为整数, 那么必有 $[G:H]=1,|H|=|G|$ 或 $[G:H]=|G|,|H|=1$。若 $[G:H]=1,|H|=|G|$, 则 $H=G$ 是平凡情况, 另一种情况 $[G:H]=|G|,|H|=1$, 这说明 $H=\{e\}$, 也就是说继承 G 的子群只有 $\{e\}$。对任何生成元个数 >1 的生成群, 其部分生成元单独生成的群都是继承此生成群的子群, 这说明 G 的生成元只有 1 个, 也就是说 G 是循环群。

值得注意的是, 循环群是可以有不平凡的子群的, 例如对 4 阶循环群 $G=\langle e,g,g^2,g^3\rangle$, 子群 $H=\langle e,g^2\rangle$ 就是继承 G 的子循环群, 可以看出 $G=\langle g\rangle$, $H=\langle g^2\rangle$。仅阶数为素数的循环群才会只有平凡的子群, 因为素数是没有 1 和其自身以外的因子的。

习题 (教材 B.7) 每个循环群的子群也是循环群。

证明 继承循环群的子群也能够表示成单个成员的次幂, 因而也是循环群。

习题 (教材 B.8) 若 $g\in G$ 的阶数 r 有限, 则 $g^m=g^n\Longleftrightarrow m=n\pmod r$

证明 $m=n\pmod r$ 说明存在 $k_1,k_2\in\mathbb{Z}_{\geqslant0}$ 以及余数 $s=m\bmod r=n\bmod r\in\mathbb{Z}_{\geqslant0}\cap[0,r[$, 使得 $m=k_1r+s$, $n=k_2r+s$, 那么 $g^m=g^{k_1r+s}=(g^r)^{k_1}g^s=g^s$, 同理 $g^n=g^s$, 这就证明了 ($\Longleftarrow$)。

若 $g^m=g^n$, 则 $g^{m+(r-n)}=g^r=e$, 即 $r\mid(m-n+r)\Longrightarrow r\mid(m-n)$, 也就是 $m=n\pmod r$。

习题 (教材 B.9) 群 G 的成员 g_1,g_2 在继承 G 的同一个子群 H 的陪集中 $\Longleftrightarrow\exists h\in H$ 满足 $g_2=g_1h$。

证明 ($\Longrightarrow$) 是显然的, 现在来证 ($\Longleftarrow$)。容易看出 $g_2\in g_1H$, $g_1=g_2h^{-1}\in g_2H$。$\forall t\in g_1H:\exists h_1\in H$, s.t. $t=g_1h_1=g_2h^{-1}h_1\in g_2H$, 这就证明了 $g_1H\subset g_2H$。类似的可以证明 $g_2H\subset g_1H$, 即 $g_1H=g_2H$。

B.2 表示

一个 n 维矩阵群是一个 $n\times n$ 矩阵构成的集合, 在矩阵乘法下构成群。为了方便我们用 `mathcal` 字母来表示矩阵群。我们记 $F^{m\times n}=\{M_{m\times n}=[m_{ij}]_{m\times n}:m_{ij}\in F\}$ (F 为数域)。单位矩阵记作 $I_n=[\delta_{ij}]_{n\times n}$, 我们已经规定了如下的矩阵群, 称为 **典型群**:

一般线性群
$$\begin{cases}GL(n,F)=\left\{M_n\in F^{n\times n}:\det M_n\neq0\right\}\\ GL_+(n,F)=\left\{M_n\in F^{n\times n}:\det M_n>0\right\},\quad GL_-(n,F)=\left\{M_n\in F^{n\times n}:\det M_n<0\right\}\end{cases}$$

特殊线性群
$$\begin{cases}SL(n,F)=\left\{M_n\in F^{n\times n}:|\det M_n|=1\right\}\\ SL_+(n,F)=SL(n,F)\cap GL_+(n,F),\quad SL_-(n,F)=SL(n,F)\cap GL_-(n,F)\end{cases}$$

正交群
$$\begin{cases}O(n)=\left\{M_n\in\mathbb{R}^{n\times n}:M_n^{\top}M_n=I_n\right\}\\ O_+(n)=O(n)\cap GL_+(n,\mathbb{R}),\quad O_-(n)=O(n)\cap GL_-(n,\mathbb{R})\end{cases}$$

特殊正交群
$$\begin{cases}SO(n)=O(n)\cap SL(n,\mathbb{R})\\ SO_+(n)=SO(n)\cap GL_+(n,\mathbb{R}),\quad SO_-(n)=SO(n)\cap GL_-(n,\mathbb{R})\end{cases}$$

幺正群
$$\begin{cases} U(n)=\left\{M_n\in\mathbb{C}^{n\times n}: M_n^\dagger M_n=I_n\right\} \\ U_+(n)=O(n)\cap GL_+(n,\mathbb{C}),\quad U_-(n)=O(n)\cap GL_-(n,\mathbb{C}) \end{cases}$$

特殊幺正群
$$\begin{cases} SU(n)=U(n)\cap SL(n,\mathbb{C}) \\ SU_+(n)=SU(n)\cap GL_+(n,\mathbb{C}),\quad SU_-(n)=SU(n)\cap GL_-(n,\mathbb{C}) \end{cases}$$

定义 B.11 (表示) 群 G 的 **表示** 是一个保持群乘法的映射 $\rho: G\to$ 矩阵群 $\mathcal{G}$, 其 **维数** 为矩阵群 $\mathcal{G}$ 的维数, 记作 d_ρ。单射表示称为 **同态表示** , 双射表示称为 **同构表示** 。

我们知道矩阵对应线性映射, $n\times n$ 矩阵就对应着 n 维空间上的线性变换。如果我们变换空间的坐标系 (将其基矢变换到另一组基矢), 那么在旧坐标系下表示这个线性变换的矩阵将变成一个新的矩阵。我们设旧坐标系下的基矢为 $\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n$, 在线性变换下变成一组新的基矢 $\vec{f}_1,\vec{f}_2,\cdots,\vec{f}_n$, 不妨设坐标的变换为

$$(x_1,x_2,\cdots,x_n)_{\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n}\longrightarrow(y_1,y_2,\cdots,y_n)_{\vec{f}_1,\vec{f}_2,\cdots,\vec{f}_n},\qquad y_i=\sum_{k=1}^n\lambda_{ik}x_k$$

为了搞清楚变换的本质, 我们在区别于 $\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n$ 和 $\vec{f}_1,\vec{f}_2,\cdots,\vec{f}_n$ 两组基矢以外的第三个固定坐标系下研究问题, 如 图 B.1 所示。设这两组基矢的坐标为 $\vec{e}_k=(e_{k1},e_{k2},\cdots,e_{kn})$ 和 $\vec{f}_i=(f_{i1},f_{i2},\cdots,f_{in})$ $(k,i=1,2,\cdots,n)$, 对一个

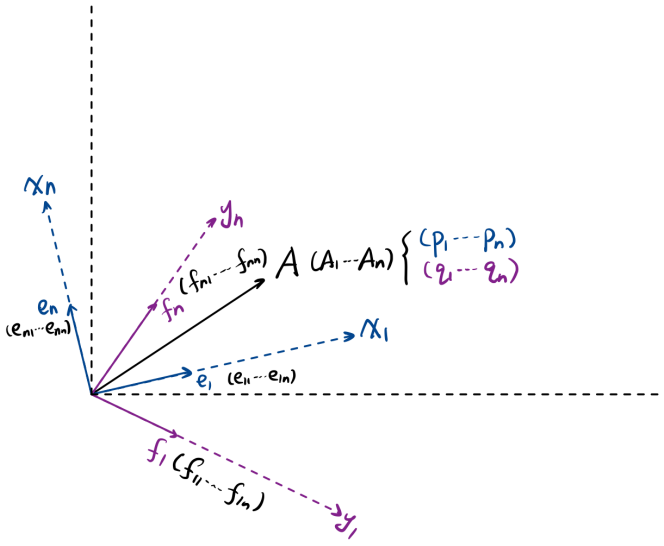

图 B.1 坐标系的变换

固定的矢量 $\vec{A}$ 我们有

$$\vec{A}=(p_1,p_2,\cdots,p_n)_{\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n}=(q_1,q_2,\cdots,q_n)_{\vec{f}_1,\vec{f}_2,\cdots,\vec{f}_n}$$

根据变换规律 $q_i=\sum_{k=1}^n\lambda_{ik}p_k$, 我们显示写出 $\vec{A}$ 的两组基底展开式

$$\vec{A}=\sum_{k=1}^n p_k\vec{e}_k=\sum_{i=1}^n q_i\vec{f}_i=\sum_{i=1}^n\left(\sum_{k=1}^n\lambda_{ik}p_k\right)\vec{f}_i=\sum_{i=1}^n\sum_{k=1}^n\lambda_{ik}p_k\vec{f}_i$$

上式给出了系数 λ_{ik} 的意义: 这个系数同时可以用作坐标的变换和基矢的变换:

$$\vec{e}_k=\sum_{i=1}^n\lambda_{ik}\vec{f}_i,\qquad 或\quad q_i=\sum_{k=1}^n\lambda_{ik}p_k$$

我们将其写成矩阵的形式, 为此引入记号:

$$x=\begin{bmatrix}p_1\\p_2\\\vdots\\p_n\end{bmatrix},\quad y=\begin{bmatrix}q_1\\q_2\\\vdots\\q_n\end{bmatrix},\quad E=\begin{bmatrix}e_{11}&e_{21}&\cdots&e_{n1}\\e_{12}&e_{22}&\cdots&e_{n2}\\\vdots&\vdots&&\vdots\\e_{1n}&e_{2n}&\cdots&e_{nn}\end{bmatrix},\quad F=\begin{bmatrix}f_{11}&f_{21}&\cdots&f_{n1}\\f_{12}&f_{22}&\cdots&f_{n2}\\\vdots&\vdots&&\vdots\\f_{1n}&f_{2n}&\cdots&f_{nn}\end{bmatrix}^{①},\quad P=[\lambda_{ik}]_{n\times n}$$

其中矩阵 E,F 的矩阵元是基矢 $\vec{e}_k,\vec{f}_i$ $(k,i=1,2,\cdots,n)$ 在第三个坐标系下的坐标。注意到

$$\vec{e}_k=\sum_{i=1}^{n}\lambda_{ik}\vec{f}_i\Longrightarrow e_{kj}=\sum_{i=1}^{n}\lambda_{ik}f_{ij}\Longrightarrow E_{jk}=\sum_{i=1}^{n}F_{ji}\lambda_{ik}\Longrightarrow E=FP^{②}$$

则我们能够给出

$$y=Px,\qquad E=FP$$

我们考虑线性变换 $\mathscr{A}:\vec{v}\to\tilde{\vec{v}}$, 在基矢 $\vec{e}_k$ $(k=1,2,\cdots,n)$ 和 $\vec{f}_i$ $(j=i,2,\cdots,n)$ 下的矩阵分别为 A,B。设 $\vec{v}$ 在基底 $\vec{e}_k$ $(k=1,2,\cdots,n)$ 下的坐标为 x, 在基底 $\vec{f}_i$ $(i=1,2,\cdots,n)$ 下的坐标为 y, 则变换前后的坐标关系应该一致 $(y=Px)\Longrightarrow[(By)=P(Ax)]$, 由此我们得到 $BPx=PAx\Longrightarrow BP=PA\Longrightarrow B=PAP^{-1}$。我们称矩阵 P 是从旧基底 $\vec{e}_k$ $(k=1,2,\cdots,n)$ 到新基底 $\vec{f}_i$ $(i=1,2,\cdots,n)$ 的 **过渡矩阵** [③], 公式 $y=Px$ 称为 **坐标变换公式** , 公式 $E=FP$ 称为 **基底变换公式** , 公式 $B=PAP^{-1}$ 称为 **矩阵变换公式** 。能够通过矩阵变换公式联系起来的两个矩阵称为 **等价** 的矩阵, 明显矩阵的等价是等价关系。

定义 B.12 我们称群 G 的两个同维度的表示 $\rho_1:G\to\mathcal{G}_1$ 和 $\rho_2:G\to\mathcal{G}_2$ 是 **等价** 的表示, 如果存在过渡矩阵 P 使得对任意成员 $g\in G$ 都有 $\rho_2(g)=P\rho_1(g)P^{-1}$ 成立, 也就是说两个表示的矩阵对应着同一个线性变换。

容易看出, 两个等价的表示之间是同构, 这说明我们可以将等价的表示看成是相同的表示。在研究群的表示时, 我们可以将结论直接放到表示矩阵群里研究, 从而摆脱具体的群成员从而仅仅研究这些成员的表示矩阵, 这也是群表示论的目的之一, 即通过表示矩阵来研究群成员, 就像通过坐标和矩阵来研究矢量和线性变换一样。我们从现在开始, 在不做特殊说明的情况下不再区分群 G 的表示和矩阵群 G (即在同一个表示 ρ 下研究时不再显式指出 ρ)。

定义 B.13 (特征标) 对群 G 的表示 $\rho:G\to\mathcal{G}$, 成员 $g\in G$ 对表示 ρ 的 **特征标** 是这个成员的表示矩阵的迹 $\chi_\rho(g)=\mathrm{tr}[\rho(g)]$, 在不引起混淆时也可以简记为 $\chi(g)$。

习题 (教材 B.11) 特征标有如下性质:

1. $\chi(I)=n$
2. $|\chi(g)|\leqslant n$
3. $|\chi(g)|=n\Longrightarrow g=e^{i\theta}I\ \ (\theta\in\mathbb{R})$
4. 对 G 的任意等价类 G_x, χ 在 G_x 上为常数
5. $\chi(g^{-1})=\chi^*(g)$
6. $\forall g\in G$: $\chi(g)$ 为代数数
7. 若 ρ_1,ρ_2 为两个等价的表示, 则 $\chi_{\rho_1}\equiv\chi_{\rho_2}$

习题 (教材 B.12) 证明任何 n 维矩阵群都有等价的 $U(n)$ 表示群。

定义 B.14 如果矩阵群 G 等价于具有相同结构的块下三角矩阵构成的矩阵群, 则称群 G 是 **(部分) 可约** 的, 如果矩阵群 G 等价于具有相同结构的块对角矩阵构成的矩阵群, 则称群 G 是 **完全可约** 的。

---

① 这里矩阵 E,F 的矩阵元貌似进行了转置, 这是因为我们想要把 E,F 写成 $E=(e_1,e_2,\cdots,e_n),F=(f_1,f_2,\cdots,f_n)$ 的形式, 这里 e_k,f_i $(k,i=1,2,\cdots,n)$ 是矢量 $\vec{e}_k,\vec{f}_i$ $(k,i=1,2,\cdots,n)$ 在第三个坐标系下的坐标构成的列矩阵。

② E_{jk} 和 F_{ji} 分别为矩阵 E,F 的矩阵元。

③ 虽然我们说是称 P 为 "从 E 到 F 的" 过渡矩阵, 但实际上从 $E=FP$ 上看 P 是作用到新的基底 F 得到旧的基底 E 的。这么称呼是因为我们是从坐标 $y=Px$ 出发来构造过度矩阵的, 如果从基底 $E=FP$ 出发构造过渡矩阵, 我们的旧基底就变成 "新基底", 新基底就变成 "旧基底"。

群表示论的最基本, 同时也是最重要的结果就是下列的所谓 **Schur 引理**:

定理 (Schur 引理) 对相同阶数的 n_1 维 N 阶矩阵群 G_1 和 n_2 维 N 阶矩阵群 G_2, 如果存在 $n_2 \times n_1$ 矩阵 S 使得对任意遍历 $g_i^{(1)} \in G_1, g_i^{(2)} \in G_2$ $(i = 1, 2, \cdots, N)$ 都有 $Sg_i^{(1)} = g_i^{(2)} S$ 成立, 则成立以下两个结论之一:

1. $S = 0$ 为零矩阵
2. $n_1 = n_2$ 且 $\det S \neq 0$, 即 G_1, G_2 维数相同且 $S \in GL(n)$

命题 B.6 群 G 的表示不可约 $\Longleftrightarrow \sum_{g \in G} |\chi(g)|^2 = |G|$

习题 (教材 B.13) 证明 Abel 群的不可约表示只有 1 维表示, 即总存在等价的幺正 $U(1)$ 表示 (教材习题 B.12 的结论)。

习题 (教材 B.14) 若 ρ 是群 G 的不可约表示, 则其维数 $d_\rho \,|\, |G|$。即群的阶数 $|G|$ 总能被其不可约表示的维数 d_ρ 整除。

B.3 Fourier 变换

附录 C　Solovay-Kitaev 定理

附录 D　数论

附录 E　公钥密码和 RSA 密码系统

附录 F　Lieb 定理的证明