高等工程数学

习题解答与提示

（教师内部参考）

教材——

南京理工大学高等工程数学编写组，高等工程数学讲义，2018年7月.

目　录

第 一 章　习题解答与提示

1. 分别证明例1.1-例1.4定义的距离都满足距离的三个条件。

证明：例1.1定义的离散距离满足距离都的三个条件是显然的，在此略去；

例1.2中d_1, d_∞满足非负性、对称性由定义可以直接得到，由于绝对值满足三角不等式，所以d_1, d_∞也满足三角不等式； 例1.2中d_2满足非负性、对称性由定义可以直接得到，取Minkowski不等式（见1.6节）中$p=2$即可证明d_2满足三角不等式;

例1.3中d_p满足非负性、对称性由定义可以直接得到，利用Minkowski不等式即可证明d_p满足三角不等式;

例1.4中d_∞满足非负性、对称性由定义可以直接得到，利用上确界的定义以及绝对值的三角不等式，容易得到d_∞满足三角不等式。

2. 证明极限的性质1.1。

证明：利用距离的三角不等式得：

$$\begin{aligned}\|d(\boldsymbol{x}_n,\boldsymbol{y}_n)-d(\boldsymbol{x}_0,\boldsymbol{y}_0)\| &\le \|d(\boldsymbol{x}_n,\boldsymbol{y}_n)-d(\boldsymbol{y}_n,\boldsymbol{x}_0)\|+\|d(\boldsymbol{y}_n,\boldsymbol{x}_0)-d(\boldsymbol{x}_0,\boldsymbol{y}_0)\| \\ &\le\ \ d(\boldsymbol{x}_n,\boldsymbol{x}_0))+d(\boldsymbol{y}_n,\boldsymbol{y}_0))\longrightarrow 0,\end{aligned}$$

从而

$$\lim_{n\to\infty} d(\boldsymbol{x}_n,\boldsymbol{y}_n)=d(\boldsymbol{x}_0,\boldsymbol{y}_0).$$

若$\boldsymbol{x}_0,\boldsymbol{y}_0\in X$都是点列$\{\boldsymbol{x}_n\}$的极限，则利用三角不等式得：

$$d(\boldsymbol{x}_0,\boldsymbol{y}_0)\le d(\boldsymbol{x}_0,\boldsymbol{x}_n)+d(\boldsymbol{x}_n,\boldsymbol{y}_0)$$

令$n\longrightarrow\infty$有$d(\boldsymbol{x}_n,\boldsymbol{x}_0)\longrightarrow 0, d(\boldsymbol{x}_n,\boldsymbol{y}_0)\longrightarrow 0$,因此 $d(\boldsymbol{x}_0,\boldsymbol{y}_0)=0$即$\boldsymbol{x}_0=\boldsymbol{y}_0$。

3. 证明定理1.1。

证明：(1) $\Longrightarrow$ (2)，由于T是连续的，从而对$\forall\varepsilon>0$，存在$\delta>0$，当 $d_X(\boldsymbol{x},\boldsymbol{x}_0)<\delta$时，有$d_Y(T\boldsymbol{x},T\boldsymbol{x}_0)<\varepsilon$。 由于$\boldsymbol{x}_n\to\boldsymbol{x}_0(n\to\infty)$，对上述$\delta>0$存在自然数$N$，当$n>N$时 $d_X(\boldsymbol{x}_n,\boldsymbol{x}_0)<\delta$，所以$d_Y(T\boldsymbol{x}_n,T\boldsymbol{x}_0)<\varepsilon$,即$T\boldsymbol{x}_n\to T\boldsymbol{x}_0(n\to\infty)$。 $T\boldsymbol{x}_n\to T\boldsymbol{x}_0(n\to\infty)$。

(2) $\Longrightarrow$ (1)，反证法。若T在$\boldsymbol{x}_0$不连续，则存在$\varepsilon_0>0$，使对任意$\delta>0$，存在$\boldsymbol{x}_\delta\in X$， 且$d_X(\boldsymbol{x}_\delta,\boldsymbol{x}_0)<\delta$， 但$d_Y(T\boldsymbol{x}_\delta,T\boldsymbol{x}_0)\ge\varepsilon_0$，特别取$\delta=\frac{1}{n},(n=$

$1,2,\cdots$)， 则有$\boldsymbol{x}_n \in X$，$d_X(\boldsymbol{x}_n, \boldsymbol{x}_0) < \frac{1}{n}$， 但$d_Y(T\boldsymbol{x}_n, T\boldsymbol{x}_0) \geq \varepsilon_0$， 这意味着$\boldsymbol{x}_n \longrightarrow \boldsymbol{x}_0(n \longrightarrow \infty)$，但$T\boldsymbol{x}_n \longrightarrow T\boldsymbol{x}_0(n \longrightarrow \infty)$不成立，矛盾。

4. 证明例1.16。

证明：(1).假设X是有限集，$\{\boldsymbol{x}_n\} \subset X$为一无穷点列，则$X$中至少有一元素在$\{\boldsymbol{x}_n\} \subset X$出现无穷多次，记该元素为$\boldsymbol{x}_0$。 从而在$\{\boldsymbol{x}_n\}$中可以选取子列，使得该子列每项都是$\boldsymbol{x}_0$， 显然该子列是收敛的。

(2).不妨假设$\{x_n\} \subset [a,b]$为一个各元素互异的点列，对$[a,b]$进行二等分成两个闭区间，则必有一个区间中含有$\{x_n\} \subset [a,b]$中无穷多个元素，记该区间为$[a_1, b_1]$。如此反复下去，我们得到一个闭区间序列$\{[a_n, b_n]\}$，该闭区间列中每一个闭区间都含有$\{x_n\} \subset [a,b]$中无穷多个元素，并且

$$[a_n, b_n] \supset [a_{n+1} b_{n+1}], n = 1, 2, \cdots, b_n - a_n \longrightarrow 0(n \longrightarrow \infty).$$

从而存在唯一一点$x_0 \in [a,b]$，使得对任何n有$x_0 \in [a_n, b_n]$(该性质称为闭区间套定理)。根据构造，我们可以在$[a_n, b_n]$中选取$\{x_n\}$ 中互异的子列$\{x_{n_k}\}$，显然该点列以x_0为极限。

(3). 略。

(4). 若$A \subset \boldsymbol{R}^n$为有界闭集，按照(2)的方法，可以同样证明$A$为紧集。若$A$为紧集，下证$A$为有界闭集。首先类似(3)的方法，如果$A$不是有界集，可以构造$A$中一个点列，使之发散到无穷，该点列就不存在收敛子列。如果A不是闭集，设x_0为A的边界点，则对任何自然数n， $B(x_0, \frac{1}{n}) \cap \{A\ \{x_0\}\} \neq \emptyset$。 通过适当选取，可以得到点列$\{x_n\}$使得当$n \neq m$时$x_n \neq x_m$且$x_n \in B(x_0, \frac{1}{n}) \cap \{A\ \{x_0\}\}$。根据构造可知$\{x_n\}$在$A$中不收敛，因为$\{x_n\}$收敛到$A$的边界点$x_0$。

5. 证明定理1.7。

证明：(1) 假设A是距离空间X中的列紧集，$\{\boldsymbol{x}_n\} \in \overline{A}$，且$\{\boldsymbol{x}_n\}$各点互异。 如果$\{\boldsymbol{x}_n\}$中含有无穷多个点在$A$中，记$\{\boldsymbol{x}_n\} \cap A = \{\boldsymbol{x}_{n_k}\}$，则由$A$是列紧以及$\overline{A}$包含$A$的所有聚点，可以得到$\{\boldsymbol{x}_{n_k}\}$在$\overline{A}$中有收敛子列，从而$\{\boldsymbol{x}_n\}$在$\overline{A}$中有收敛子列。 如果$\{\boldsymbol{x}_n\}$中不含有无穷多个点在$A$中，则$\{\boldsymbol{x}_n\}$中含有无穷多个$A$的边界点，记$\{\boldsymbol{x}_n\}$中所有边界点组成的子列为 $\{\boldsymbol{x}_{n_k}\}$，对每个$\boldsymbol{x}_{n_k}$，存在$\boldsymbol{y}_k \in A$使得$d(\boldsymbol{x}_{n_k}, \boldsymbol{y}_k) < \frac{1}{k}$。由于$A$是列紧以及$\overline{A}$ 包含A的所有聚点，所以$\{\boldsymbol{y}_k\}$在存在$\overline{A}$中的收敛子列$\{\boldsymbol{y}_{k_m}\}$，根据构造可以得到$\boldsymbol{x}_{n_{k_m}}$在$\overline{A}$中收敛。 所以$\overline{A}$为紧集。

(2) 假设A是距离空间X中的列紧集，$B \subset A$。若$\{\boldsymbol{x}_n\} \subset B \subset A$，则由于$A$是列紧的，从而$\{\boldsymbol{x}_n\}$有子列 收敛到$X$中，根据定义可知$B$为列紧集。

6. 证明稠密性具有传递性，即A在B中稠密，B在C中稠密，则A在C中稠密。

证明：不妨设A, B, C都是距离空间(X, d)的子集。由B在C中稠密，则对$\forall \boldsymbol{y} \in C, \forall \varepsilon > 0$，存在 $\boldsymbol{z} \in B$ 使得 $d(\boldsymbol{z}, \boldsymbol{y}) < \frac{1}{2}\varepsilon$。同理由$A$在$B$中稠密，存在$\boldsymbol{x} \in A$ 使得 $d(\boldsymbol{z}, \boldsymbol{x}) < \frac{1}{2}\varepsilon$。所以

$$d(\boldsymbol{x}, \boldsymbol{y}) < d(\boldsymbol{x}, \boldsymbol{z}) + d(\boldsymbol{z}, \boldsymbol{y}) < \varepsilon.$$

即由A在C中稠密。

7. 设A是列紧的且是闭的，证明A是紧的。

证明：假设A为距离空间(X, d)中的列紧且闭的集。对任何序列$\{\boldsymbol{x}_n\} \subset A$，由于$A$是列紧的，从而存在子列$\{\boldsymbol{x}_{n_k}\}$在 X中收敛到$\boldsymbol{x}_0$。则$\boldsymbol{x}_0 \in A$或$\boldsymbol{x}_0$为A的边界点。 又由A为闭集，从而$\boldsymbol{x}_0 \in A$，所以A为紧集。

8. 设X是度量空间，$\boldsymbol{x} \in X$，A是X中的紧集，记

$$d(x, A) = \inf\{d(x, y) : y \in A\}$$

证明：当$d(x, A) = 0$时，那么$x \in A$； 若将A换成列紧的，结论是否成立？

证明：若$d(\boldsymbol{x}, A) = 0$， 则根据$d(\boldsymbol{x}, A)$的定义，存在点列$\{\boldsymbol{x}_n\} \subset A$使得 $d(\boldsymbol{x}, \boldsymbol{x}_n) < \frac{1}{n}$。由于$A$是紧集，所以存在$\{\boldsymbol{x}_n\}$的子列$\{\boldsymbol{x}_{n_k}\}$在$A$中收敛到 $\boldsymbol{x}_0$。则

$$d(\boldsymbol{x}, \boldsymbol{x}_0) \le d(x, \boldsymbol{x}_{n_k}) + d(\boldsymbol{x}_{n_k}, x_0) \longrightarrow 0 (n_k \longrightarrow \infty).$$

所以$\boldsymbol{x} = \boldsymbol{x}_0 \in A$。若$A$是列紧集，则上述结论不成立。

9. 记$C[a, b]$为区间$[a, b]$上连续函数空间，对任何$f, g \in C[a, b]$，定义

$$d(f, g) = \max_{a \le x \le b} |f(x) - g(x)|.$$

证明d是$C[a, b]$上的距离，并且$C[a, b]$是完备空间。

证明：根据$d(f, g)$的定义容易得到d满足距离定义的非负性和对称性，下证d满足三角不等式。 对任何$t \in [a, b]$，对任意函数$h \in C[a, b]$，则

$$\begin{aligned} |f(t) - g(t)| &\le |f(t) - h(t)| + |h(t) - g(t)| \\ &\le \max_{a \le t \le b} |f(t) - h(t)| + \max_{a \le t \le b} |h(t) - g(t)| \\ &= d(f, h) + d(h, g). \end{aligned}$$

下证$C[a, b]$为完备空间。设$\{f_n\}$是$C[a, b]$中一Cauchy列，即对$\forall \varepsilon > 0$，存在自然数N，当$n, m > N$时，$d(f_n, f_m) < \varepsilon$， 即对任何$t \in [a, b]$，必有$|f_n(t) -$

$f_m(t)| < \varepsilon$，可知数列$\{f_n(t)\}$是收敛数列。设其极限为$f_0(t)$。在 $|f_n(t) - f_m(t)| < \varepsilon$中令$m \longrightarrow \infty$得$|f_n(t) - f_0(t)| < \varepsilon$，即$\{f_n\}$一致收敛到$f_0$，从而$f_0 \in C[a,b]$且$d(f_n, f_0) \longrightarrow 0(n \longrightarrow \infty)$。所以$C[a,b]$是完备的。

10. 用压缩映射原理证明方程$x = a\sin x$只有唯一解，其中$a \in (0,1)$。

证明：作映射$T : \boldsymbol{R} \longrightarrow \boldsymbol{R}$使得$Tx = a\sin x$。对任何$x, y \in \boldsymbol{R}$（不妨设$x < y$），利用中值定理得存在$\xi \in (x,y)$使得

$$|Tx - Ty| = a|\sin x - \sin y| = a|\cos\xi||x-y| \le a|x-y|.$$

所以T是一压缩映射，从而T有唯一不动点，即$x = a\sin x$有唯一解。

11. 证明下述线性方程组

$$A = \begin{pmatrix} \frac{3}{4} & -\frac{1}{4} & \frac{1}{4} \\ -\frac{1}{7} & \frac{5}{14} & -\frac{1}{5} \\ -\frac{1}{100} & -\frac{1}{120} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{6} \\ \frac{1}{7} \\ \frac{1}{8} \end{pmatrix}$$

有唯一解，并写出方程近似解的迭代序列。

证明：记

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} \frac{3}{4} & -\frac{1}{4} & \frac{1}{4} \\ -\frac{1}{7} & \frac{5}{14} & -\frac{1}{5} \\ -\frac{1}{100} & -\frac{1}{120} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{7} & \frac{9}{14} & \frac{1}{5} \\ \frac{1}{100} & \frac{1}{120} & \frac{1}{2} \end{pmatrix}$$

则容易验证

$$\sum_{j=1}^{3} |c_{ij}| < 1(i = 1,2,3).$$

从而由例1.17结论可得所求线性方程有唯一解，且解可由如下迭代序列

$$\boldsymbol{x}^{(k)} = C\boldsymbol{x}^{(k-1)} + \begin{pmatrix} \frac{1}{6} \\ \frac{1}{7} \\ \frac{1}{8} \end{pmatrix}$$

近似计算得到。

12. 用压缩映射原理构造迭代序列求解下述微分方程

$$\begin{cases} \dfrac{dy}{dx} = 1 + x^2, \\ x(0) = 0 \end{cases}$$

的解。

证明： 可以验证所求微分方程的初值问题满足例1.18的条件，从而有唯一解。由于本题中函数$f(x,y)=1+x^2$是一个和y无关的函数。从而其迭代序列只有一步，即为

$$y(x)=\int_0^x(1+t^2)dt$$

13. 赋范线性空间X是Banach空间的充要条件是对任意序列 $\{\boldsymbol{x}_n\}\subset X$，若 $\sum\limits_{i=1}^{\infty}\|x_i\|$ 收敛，则 $\sum\limits_{i=1}^{\infty}x_i$ 在 X 中收敛。

证明： 若X是Banach空间，且$\sum\limits_{i=1}^{\infty}\|x_i\|$收敛，令$S_n=\boldsymbol{x}_1+\boldsymbol{x}_2+\cdots+\boldsymbol{x}_n$，则

$\|S_{n+p}-S_n\|=\|\boldsymbol{x}_{n+1}+\cdots+\boldsymbol{x}_{n+p}\|\le\|\boldsymbol{x}_{n+1}\|+\cdots+\|\boldsymbol{x}_{n+p}\|\longrightarrow 0\ (n\longrightarrow\infty)$.

因此$\{S_n\}$是X中的Cauchy列，由于X是Banach空间，存在$\boldsymbol{x}\in X$ 使得$S_n\to\boldsymbol{x}$ $(n\to\infty)$，即 $\sum\limits_{i=1}^{\infty}\boldsymbol{x}_i=\boldsymbol{x}$.

反之，若$\sum\limits_{i=1}^{\infty}\|\boldsymbol{x}_i\|$收敛，则$\sum\limits_{i=1}^{\infty}\boldsymbol{x}_i$在$X$中收敛。则对于$X$中的Cauchy 列$\{\boldsymbol{x}_n\}$，对$\varepsilon_k=\frac{1}{2^k}$，有$n_1<n_2<\cdots<n_k<\cdots$，使得

$$\|\boldsymbol{x}_{n_{k+1}}-\boldsymbol{x}_{n_k}\|<\frac{1}{2^k}(k=1,2,\cdots)$$

因而$\sum\limits_{n=1}^{\infty}\|\boldsymbol{x}_{n_{k+1}}-\boldsymbol{x}_{n_k}\|<\infty$. 由假设可知$\sum\limits_{n=1}^{\infty}(\boldsymbol{x}_{n_{k+1}}-\boldsymbol{x}_{n_k})$收敛于某个 $\boldsymbol{x}\in X$，即$\{\boldsymbol{x}_{n_k}\}$收敛于$\boldsymbol{x}$，所以$\{\boldsymbol{x}_n\}$也收敛于$\boldsymbol{x}$，从而 X完备，即X为Banach空间。

14. 设X是无限维赋范线性空间，Y是X的有限维子空间，证明：必存在$\boldsymbol{x}_0\in X$且$\|x_0\|=1$，使得$d(x_0,Y)\ge 1$。

证明： 由于Y是X的有限维子空间，从而任取$\overline{\boldsymbol{x}}\in X-Y$，有$d=\inf\limits_{\boldsymbol{y}\in Y}\|\boldsymbol{y}-\overline{\boldsymbol{x}}\|>0$，并且存在$\boldsymbol{y}'\in Y$使得$d=\|\boldsymbol{y}'-\overline{\boldsymbol{x}}\|$。作$\boldsymbol{x}_0=\frac{\boldsymbol{y}'-\overline{\boldsymbol{x}}}{\|\boldsymbol{y}'-\overline{\boldsymbol{x}}\|}$，则$\|\boldsymbol{x}_0\|=1$。任取$y\in Y$，由于$\boldsymbol{y}'+\|\boldsymbol{y}'-\overline{\boldsymbol{x}}\|\boldsymbol{y}\in Y$，则有

$$\|\overline{\boldsymbol{x}}-(\boldsymbol{y}'+\|\boldsymbol{y}'-\overline{\boldsymbol{x}}\|\boldsymbol{y})\|\ge d.$$

因此

$$\begin{aligned}\|\boldsymbol{x}_0-\boldsymbol{y}\| &= \|\tfrac{\boldsymbol{y}'-\overline{\boldsymbol{x}}}{\|\boldsymbol{y}'-\overline{\boldsymbol{x}}\|}-\boldsymbol{y}\|=\tfrac{1}{\|\boldsymbol{y}'-\overline{\boldsymbol{x}}\|}\|\overline{\boldsymbol{x}}-(\boldsymbol{y}'+\|\boldsymbol{y}'-\overline{\boldsymbol{x}}\|y)\|\\ &\ge \tfrac{d}{\|\boldsymbol{y}'-\overline{\boldsymbol{x}}\|}=1.\end{aligned}$$

所以$d(x_0,Y)\ge 1$。

15. 设$\{\boldsymbol{e}_n\}$是内积空间X中一个标准正交系，求证对任何$\boldsymbol{x},\boldsymbol{y}\in X$，有

$$\sum_{n=1}^{\infty}|\langle\boldsymbol{x},\boldsymbol{e}_n\rangle\cdot\langle\boldsymbol{y},\boldsymbol{e}_n\rangle|\le\|\boldsymbol{x}\|\cdot\|\boldsymbol{y}\|.$$

证明： 分别令$\boldsymbol{x}'=\boldsymbol{x}-\sum_{n=1}^{\infty}\langle\boldsymbol{x},\boldsymbol{e}_n\rangle\boldsymbol{e}_n$，$\boldsymbol{y}'=\boldsymbol{x}-\sum_{n=1}^{\infty}\langle\boldsymbol{y},\boldsymbol{e}_n\rangle\boldsymbol{e}_n$，则 $\boldsymbol{x}'$ 与 $\sum_{n=1}^{\infty}\langle\boldsymbol{x},\boldsymbol{e}_n\rangle\boldsymbol{e}_n$ 正交，则由勾股定理得：

$$\|\boldsymbol{x}\|^2=\|\boldsymbol{x}'\|^2+\left\|\sum_{n=1}^{\infty}\langle\boldsymbol{x},\boldsymbol{e}_n\rangle\boldsymbol{e}_n\right\|^2\ge\left\|\sum_{n=1}^{\infty}\langle\boldsymbol{x},\boldsymbol{e}_n\rangle\boldsymbol{e}_n\right\|^2=\sum_{n=1}^{\infty}|\langle\boldsymbol{x},\boldsymbol{e}_n\rangle|^2,$$

$$\|\boldsymbol{y}\|^2=\|\boldsymbol{y}'\|^2+\left\|\sum_{n=1}^{\infty}\langle\boldsymbol{y},\boldsymbol{e}_n\rangle\boldsymbol{e}_n\right\|^2\ge\left\|\sum_{n=1}^{\infty}\langle\boldsymbol{y},\boldsymbol{e}_n\rangle\boldsymbol{e}_n\right\|^2=\sum_{n=1}^{\infty}|\langle\boldsymbol{y},\boldsymbol{e}_n\rangle|^2.$$

于是利用Holder不等式得：

$$\sum_{n=1}^{\infty}|\langle\boldsymbol{x},\boldsymbol{e}_n\rangle\cdot\langle\boldsymbol{y},\boldsymbol{e}_n\rangle|\le\left(\sum_{n=1}^{\infty}|\langle\boldsymbol{x},\boldsymbol{e}_n\rangle|^2\right)^{\frac{1}{2}}\left(\sum_{n=1}^{\infty}|\langle\boldsymbol{y},\boldsymbol{e}_n\rangle|^2\right)^{\frac{1}{2}}\le\|\boldsymbol{x}\|\|\boldsymbol{y}\|.$$

16. 设$\{\boldsymbol{e}_n\}$是Hilbert空间X的一个标准正交系，对任何$\boldsymbol{x}\in X$，求证

$$\boldsymbol{y}=\sum_{n=1}^{\infty}\langle\boldsymbol{x},\boldsymbol{e}_n\rangle\boldsymbol{e}_n$$

在X中更不能收敛，并且$\boldsymbol{x}-\boldsymbol{y}$与每个 $\boldsymbol{e}_n$正交。

证明： 记$S_n=\sum_{k=1}^{n}\langle\boldsymbol{x},\boldsymbol{e}_k\rangle\boldsymbol{e}_k, T_n=\sum_{k=1}^{n}|\langle\boldsymbol{x},\boldsymbol{e}_k\rangle|^2$. 则由习题15可知：$\sum_{k=1}^{\infty}|\langle\boldsymbol{x},\boldsymbol{e}_k\rangle|^2\le\|x\|^2$，从而$T_n$是收敛的。对任何自然数$m$以及$n>m$，有

$$\|S_n-S_m\|^2=\sum_{k=m+1}^{n}|\langle\boldsymbol{x},\boldsymbol{e}_k\rangle|^2=T_n-T_m.$$

从而$\{S_n\}$是一个Cauchy列，由X的完备性，可得$\boldsymbol{y}=\sum_{k=1}^{\infty}\langle\boldsymbol{x},\boldsymbol{e}_k\rangle\boldsymbol{e}_k$在$X$中收敛。 简单计算可以得到$\boldsymbol{x}-\boldsymbol{y}$与每个$\boldsymbol{e}_n$正交。

17. 设 $A\in\boldsymbol{C}^{n\times n}$， 证明： $\dfrac{1}{\sqrt{n}}\|A\|_F\le\|A\|_2\le\|A\|_F$。

证明： 设 A^HA 的特征值为 λ_i, 则 $\lambda_i\ge 0$, $i=1,2,\cdots,n$.

$$\|A\|_2^2=\rho(A^HA)\le\sum_{i=1}^{n}\lambda_i=tr(A^HA)=\|A\|_F^2.$$

另外

$$\|A\|_F^2=tr(A^HA)=\sum_{i=1}^{n}\lambda_i\le n\max_i\lambda_i=n\rho(A^HA)=n\|A\|_2^2.$$

18. 设 $A=\begin{pmatrix}1/2 & 1\\ 0 & 1/2\end{pmatrix}$，对 $k=2,3,\cdots$， 直接计算 A^k 及 $\rho(A^k)$， $\|A^k\|_1$，$\|A^k\|_\infty$， $\|A\|_2$。

解：$A^k=\begin{pmatrix}\frac{1}{2^k} & \frac{k}{2^{k-1}}\\ 0 & \frac{1}{2^k}\end{pmatrix}$, $\rho(A^k)=\frac{1}{2^k}$, $\|A^k\|_1=\frac{2k+1}{2^k}$. 因为

$$A^k(A^k)^T=\begin{pmatrix}\frac{1}{2^k} & \frac{k}{2^{k-1}}\\ 0 & \frac{1}{2^k}\end{pmatrix}\begin{pmatrix}\frac{1}{2^k} & 0\\ \frac{k}{2^{k-1}} & \frac{1}{2^k}\end{pmatrix}=\begin{pmatrix}\frac{4k^2+1}{4^k} & \frac{k}{2^{2k-1}}\\ \frac{k}{2^{2k-1}} & \frac{1}{4^k}\end{pmatrix}$$

所以 $A^k(A^k)^T$ 的特征值为 $\lambda_{1,2}=\frac{1}{4^k}(\sqrt{k^2+1}\pm k)^2$. 从而

$$\|A^k\|_2=\sqrt{\rho(A^k(A^k)^T)}=\frac{1}{2^k}(\sqrt{k^2+1}+k).$$

19. 如果 $\|\cdot\|$ 是 $\boldsymbol{C}^{n\times n}$ 上的矩阵范数，证明：对所有 $c\geq 1$，$c\|\cdot\|$ 是矩阵范数，但是， 对任意 $c<1$，$c\|\cdot\|_1$ 不是矩阵范数。

证明：(1) 对 $c\geq 1$，因为

(i) $c\|A\|\geq 0$, 且 $c\|A\|=0\Longleftrightarrow\|A\|=0\Longleftrightarrow A=0$;

(ii) $c\|\lambda A\|=c|\lambda|\|A\|=\lambda c\|A\|$;

(iii) $c\|A+B\|\leq c\|A\|+c\|B\|$;

(iv) $c\|AB\|\leq c\|A\|\|B\|\leq c\|A\|\cdot c\|B\|$,

所以，$c\|\cdot\|$ 是矩阵范数。

(2) 对 $c<1$,

- 若 $c=0$, 因为当 $A\neq 0$ 时，$c\|A\|_1=0$，所以 $c\|\cdot\|$ 不是矩阵范数。
- 若 $c<0$, 那么 $c\|A+B\|_1\leq c\|A\|_1+c\|B\|_1$ 不一定成立，如设

$$A=\begin{pmatrix}1 & 2\\ 0 & 3\end{pmatrix},B=\begin{pmatrix}1 & -1\\ 0 & -2\end{pmatrix}$$

则 $A+B=\begin{pmatrix}2 & 1\\ 0 & 1\end{pmatrix}$, 且有 $\|A+B\|_1=2,\|A\|_1=5,\|B\|_1=3$, 而 $c\|A+B\|_1=2c>8c=c\|A\|_1+c\|B\|_1$. 所以 $c\|\cdot\|$ 不是矩阵范数。

- 若 $0<c<1$, 那么 $c\|AB\|_1\leq c\|A\|_1\cdot c\|B\|_1$ 不一定成立, 如设

$$A=\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix},B=\begin{pmatrix}1 & 0\\ 0 & 3\end{pmatrix}$$

则 $AB=\begin{pmatrix}1&0\\0&6\end{pmatrix}$, 且有 $\|AB\|_1=6, \|A\|_1=2, \|B\|_1=3$, 而 $c\|AB\|_1=6c>6c^2=c\|A\|_1\cdot c\|B\|_1$. 所以 $c\|\cdot\|$ 不是矩阵范数。

20. 设 $A\in \boldsymbol{C}^{n\times n}$，证明 Hermite 矩阵

$$\hat{A}=\begin{pmatrix}0&A\\A^H&0\end{pmatrix}\in \boldsymbol{C}^{2n\times 2n}.$$

与 A 有相同的谱范数。

证明：因为

$$\hat{A}^H\hat{A}=\begin{pmatrix}0&A\\A^H&0\end{pmatrix}\begin{pmatrix}0&A\\A^H&0\end{pmatrix}=\begin{pmatrix}AA^H&0\\0&A^HA\end{pmatrix}$$

又 AA^H 与 A^HA 的特征值相同，所以 $\hat{A}^H\hat{A}$ 与 A^HA 的特征值相同. 故

$$\|\hat{A}\|_2=\sqrt{\rho(\hat{A}^H\hat{A})}=\sqrt{\rho(A^HA)}=\|A\|_2.$$

21. 设 $A,B\in \boldsymbol{C}^{n\times n}$ 都是 Hermite 矩阵， 证明： $\rho(A+B)\le\rho(A)+\rho(B)$.

证明：因为 $A+B$ 也是 Hermite 矩阵，所以

$$\rho(A+B)=\|A+B\|_2\le\|A\|_2+\|B\|_2=\rho(A)+\rho(B).$$

22. 证明，如果 $A\in \boldsymbol{C}^{n\times n}$ 可逆，则 $\mathrm{cond}(A)=\mathrm{cond}(A^{-1})$。

证明：$\mathrm{cond}(A)=\|A\|\|A^{-1}\|=\|(A^{-1})^{-1}\|\|A^{-1}\|=\mathrm{cond}(A^{-1})$.

23. 设 $A=\begin{pmatrix}2&1&3\\8&2&4\\4&4&10\end{pmatrix}$， 求 A 的条件数 $\mathrm{cond}_\infty(A)$。

证明：因为

$$A^{-1}=\begin{pmatrix}\frac14&\frac18&-\frac18\\-4&\frac12&1\\\frac32&-\frac14&-\frac14\end{pmatrix}$$

所以 $\mathrm{cond}(A)=\|A\|_\infty\|A^{-1}\|_\infty=18\times5.5=99$.

24. 证明：$\mathrm{cond}(AB)\le\mathrm{cond}(A)\mathrm{cond}(B)$。那么 $\mathrm{cond}(\cdot)$ 是 $\boldsymbol{C}^{n\times n}$ 上的矩阵范数吗？

证明：$\mathrm{cond}(AB)=\|AB\|\|(AB)^{-1}\|=\|AB\|\|B^{-1}A^{-1}\|\le\|A\|\|B\|\|B^{-1}\|\|A^{-1}\|=\|A\|\|A^{-1}\|\|B\|\|B^{-1}\|=\mathrm{cond}(A)\mathrm{cond}(B)$。

$\mathrm{cond}(\cdot)$ 不是矩阵范数，因为对 $|\lambda|\ne0,1$, $\mathrm{cond}(\lambda A)=\|\lambda A\|\|(\lambda A)^{-1}\|=|\lambda|\|A\|\|\frac{1}{\lambda}A^{-1}\|=|\lambda|\|A\|\frac{1}{|\lambda|}\|A^{-1}\|=\mathrm{cond}(A)\ne|\lambda|\mathrm{cond}(A)$。

第 二 章　习题解答与提示

1. 设$A \in \boldsymbol{C}^{n\times n}$，若$A$的每一行元素的和(简称为行和)为1，试证：
(1) 1为A的特征值；(2) 若A可逆，则A^{-1}的行和也是1；(3) 给定多项式 $f(x)$. 问 $f(A)$ 的行和是否相等？若相等，等于何值？

证明：(1)取$e = (1, 1, \cdots, 1)^T$，则有$Ae = 1e$，所以1为A的特征值。

(2) 由$Ae = 1e \Rightarrow A^{-1}e = 1 \cdot e$，故$A^{-1}$的行和也是1.

(3) 由第一节性质5知$f(A)\,e = f(1)\,e$，因此$f(A)$的行和相等，且为$f(1)$.

2. 设 $A, B \in \boldsymbol{C}^{n\times n}$， A 的特征值互异。证明：$AB = BA$ 的充分必要条件是 A 与 B 同时可对角化。

证明：(两个可对角化矩阵$A, B \in C^{n\times n}$称为同时可对角化的，如果存在同一个相似变换矩阵$S \in C^{n\times n}$，使得$S^{-1}AS$和$S^{-1}BS$同为对角矩阵。)

充分性 若A和B同时可对角化，则存在可逆矩阵P，使得

$$A = PDP^{-1}, B = P\Lambda P^{-1},$$

其中D和Λ是对角矩阵。于是有

$$AB = \left(PDP^{-1}\right)\left(P\Lambda P^{-1}\right) = PD\Lambda P^{-1} = P\Lambda DP^{-1} = \left(P\Lambda P^{-1}\right)\left(PDP^{-1}\right) = BA,$$

即A与B可交换。

必要性 若$AB = BA$. 设λ为A的特征值，ξ为对应的特征向量，即$A\xi = \lambda\xi$.则当$B\xi \neq 0$时，由$A(B\xi) = AB\xi = BA\xi = \lambda(B\xi)$，知$\xi, B\xi$都是对应$\lambda$的特征向量，而$\lambda$ 是A的单特征值， 所以$\xi, B\xi$线性相关。因此，存在常数μ，使$B\xi = \mu\xi$，即ξ 是B的对应特征值μ的特征向量。

当$B\xi = 0$时，ξ是对应B的0特征值的特征向量。故A的特征向量都是B的特征向量，并且使A对角化的这些特征向量所组成的同一个矩阵也使B对角化。

3. 求下列λ–矩阵的Smith标准形：

(1) $\begin{pmatrix} \lambda^3 - \lambda & 2\lambda^2 \\ \lambda^2 + 5\lambda & 3\lambda \end{pmatrix}$；(2) $\begin{pmatrix} 1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2 \end{pmatrix}$；

(3) $\begin{pmatrix} 0 & 0 & 0 & \lambda^2 \\ 0 & 0 & \lambda^2-\lambda & 0 \\ 0 & (\lambda-1)^2 & 0 & 0 \\ \lambda^2-\lambda & 0 & 0 & 0 \end{pmatrix}$.

解：(1)

$$\begin{pmatrix} \lambda^3-\lambda & 2\lambda^2 \\ \lambda^2+5\lambda & 3\lambda \end{pmatrix} \rightarrow \begin{pmatrix} 3\lambda & \lambda^2+5\lambda \\ 2\lambda^2 & \lambda^3-\lambda \end{pmatrix} \rightarrow \begin{pmatrix} 3\lambda & \lambda^2+5\lambda \\ 0 & \lambda^3-10\lambda^2-3\lambda \end{pmatrix}$$
$$\rightarrow \begin{pmatrix} \lambda & 0 \\ 0 & \lambda(\lambda^2-10\lambda-3) \end{pmatrix} = S(\lambda).$$

(2)

$$\begin{pmatrix} 1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & \lambda^2+\lambda & 0 \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2 \end{pmatrix}$$
$$\rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda(\lambda+1) \end{pmatrix} = S(\lambda).$$

(3) $S(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda(\lambda-1) & 0 & 0 \\ 0 & 0 & \lambda(\lambda-1) & 0 \\ 0 & 0 & 0 & \lambda^2(\lambda-1)^2 \end{pmatrix}$.

4. 判断 λ-矩阵

$A(\lambda) = \begin{pmatrix} 3\lambda+1 & \lambda & 4\lambda-1 \\ 1-\lambda^2 & \lambda-1 & \lambda-\lambda^2 \\ \lambda^2+\lambda+2 & \lambda & \lambda^2+2\lambda \end{pmatrix}$ 与

$B(\lambda) = \begin{pmatrix} \lambda+1 & \lambda-2 & \lambda^2-2\lambda \\ 2\lambda & 2\lambda-3 & \lambda^2-2\lambda \\ -2 & 1 & 1 \end{pmatrix}$ 是否等价。

解：因$A(\lambda)$，$B(\lambda)$有相同的不变因子1，$\lambda-1$，$(\lambda-1)^2$，所以它们等价。

5. 判断矩阵A与 B 是否相似：

(1) $A = \begin{pmatrix} 3 & 2 & -5 \\ 2 & 6 & -10 \\ 1 & 2 & -3 \end{pmatrix}$, $B = \begin{pmatrix} 6 & 20 & -34 \\ 6 & 32 & -51 \\ 4 & 20 & -32 \end{pmatrix}$；

(2) $A=\begin{pmatrix}6&6&-15\\1&5&-5\\1&2&-2\end{pmatrix}$， $B=\begin{pmatrix}37&-20&-4\\34&-17&-4\\119&-70&-11\end{pmatrix}$.

解：(1) 因A与B有相同的初等因子$(\lambda-2),(\lambda-2)^2$，所以$A$与$B$相似。

(2) A与B也有相同的初等因子$(\lambda-3),(\lambda-3)^2$，故$A\tilde{}B$.

6. 设$\varepsilon\neq 0$，证明：

(1) n阶矩阵$A=\begin{pmatrix}a&1&&\\&a&\ddots&\\&&\ddots&1\\&&&a\end{pmatrix}$ 与 $B=\begin{pmatrix}a&\varepsilon&&\\&a&\ddots&\\&&\ddots&\varepsilon\\&&&a\end{pmatrix}$相似；

(2) n阶矩阵$A=\begin{pmatrix}a&1&&\\&a&\ddots&\\&&\ddots&1\\&&&a\end{pmatrix}$ 与 $C=\begin{pmatrix}a&1&&\\&a&\ddots&\\&&\ddots&1\\\varepsilon&&&a\end{pmatrix}$不相似。

证明：(1)A与B的初等因子相同，均为$(\lambda-a)^n$，所以A与B相似。

(2) A的初等因子为$(\lambda-a)^n$，而对

$$(\lambda I-B)=\begin{pmatrix}\lambda-a&-1&&\\&\lambda-a&\ddots&\\&&\ddots&-1\\-\varepsilon&&&\lambda-a\end{pmatrix},$$

有

$$D_n(\lambda)=\det(\lambda I-B)(\lambda-a)^n-\varepsilon,$$

又$(\lambda I-B)$有一个$n-1$阶子式$H_{n-1}(\lambda)=\begin{vmatrix}-1&&&\\\lambda-a&-1&&\\&\ddots&\ddots&\\&&\lambda-a&-1\end{vmatrix}=(-1)^{n-1}$,

由$D_{n-1}(\lambda)\,|D_n(\lambda)\,D_{n-1}(\lambda)|\,H_{n-1}(\lambda)$，知必有$D_{n-1}(\lambda)=1$，进而得$D_1(\lambda)=\cdots=D_{n-1}(\lambda)=1$.于是$B$的不变因子为$d_1(\lambda)=\cdots=d_{n-1}(\lambda)=1,d_n(\lambda)=(\lambda-a)^n-\varepsilon$.所以$B$的初等因子与$A$的初等因子不相同。故$A$与$B$不相似。

7. 求下列矩阵的Jordan标准形：

(1) $\begin{pmatrix}4&-5&2\\5&-7&3\\6&-9&4\end{pmatrix}$； (2) $\begin{pmatrix}1&-3&4\\4&-7&8\\6&-7&7\end{pmatrix}$； (3) $\begin{pmatrix}1&-1&2\\3&-3&6\\2&-2&4\end{pmatrix}$；

(4) $\begin{pmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{pmatrix}$；(5) $\begin{pmatrix} 2 & -1 & 1 & -1 \\ 2 & 2 & -1 & -1 \\ 1 & 2 & -1 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}$.

解：(1) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$；(2) $\begin{pmatrix} 3 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$；(3) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$；

(4) $\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$；(5) $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

8. 求下列矩阵A的Jordan标准形，并求相似变换矩阵P，使$P^{-1}AP = J$：

(1) $A = \begin{pmatrix} -1 & 0 & 0 \\ 1 & 1 & 3 \\ 0 & 2 & 2 \end{pmatrix}$；(2) $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

解：(1) $J = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$.对$\lambda_1 = \lambda_2 = -1$，求解$(\lambda_1 I - A)\,x = 0$，得$p_1 = \begin{pmatrix} 0 \\ -3 \\ 2 \end{pmatrix}$，再求解$(\lambda_1 I - A)\,x = -p_1$，得$p_2 = \begin{pmatrix} -5 \\ 1 \\ 0 \end{pmatrix}$.对$\lambda_3 = 4$，求解$(\lambda_3 I - A)\,x = 0$，得$p_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$，故$P = \begin{pmatrix} 0 & -5 & 0 \\ -3 & 1 & 1 \\ 2 & 0 & 1 \end{pmatrix}$，且$P^{-1}AP = J$.

(2) 因$D_4(\lambda) = \det(\lambda I - A) = (\lambda - 1)^4$，又余子式$D_{41} = -4\lambda(\lambda + 1)$，所以$D_3(\lambda) = 1$，从而$D_2(\lambda) = D_1(\lambda) = 1$，$A$的不变因子为$d_1(\lambda) = d_2(\lambda) = d_3(\lambda) = 1, d_4(\lambda) = (\lambda - 1)^4$，$A$的初等因子为$(\lambda - 1)^4$，故

$$J = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

$\lambda_1 = 1$（4重），下求P.

先求解$(\lambda_1 I - A)\,x = 0$，得$p_1 = (8,0,0,0)^T$,再解

$(\lambda_1 I - A)\,x = -p_1$，得$p_2 = (4,4,0,0)^T$,

$(\lambda_1 I - A)\,x = -p_2$，得$p_3 = (0,-1,2,0)^T$,

$(\lambda_1 I - A)\,x = -p_3$，得$p_4 = (0,1,-2,1)^T$.

因此得　$P = \begin{pmatrix} 8 & 4 & 0 & 0 \\ 0 & 4 & -1 & 1 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$，且$P^{-1}AP = J$.

9. 设$A \in \boldsymbol{C}^{n\times n}$，$\lambda_0$为$A$的$r$重特征值，试用Jordan标准形理论证明

$$\operatorname{rank}(\lambda_0 I - A)^r = n - r.$$

证明：在A的Jordan标准形中，特征值不为λ_0的子块列在一起记为B，特征值为λ_0的子块列在一起记为B_0，即有$J = \begin{pmatrix} B & \\ & B_0 \end{pmatrix}$，$A = P\begin{pmatrix} B & \\ & B_0 \end{pmatrix}P^{-1}$。因$\lambda_0$为$A$的$r$重特征值，因此$B_0$是$r$阶的，$B$为$n-r$阶的，而$B_0$的主对角线上子块都形如$J_i = \begin{pmatrix} \lambda_0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{pmatrix}$，每个子块阶数不超过$r$，所以

$$(\lambda_0 I - A)^r = (\lambda_0 PP^{-1} - PJP^{-1})^r = P(\lambda_0 I - J)^r P^{-1}$$

$$= P\begin{pmatrix} \lambda_0 I_1 - B & \\ & \lambda_0 I_2 - B_0 \end{pmatrix}^r P^{-1} = P\begin{pmatrix} (\lambda_0 I_1 - B)^r & \\ & (\lambda_0 I_2 - B_0)^r \end{pmatrix}P^{-1}$$

因子块$\lambda_0 I_1 - B$的主对角线上元素都不为零，它是$n-r$阶非奇异的，而$\lambda_0 I_2 - B_0$的主对角线上子块都形如$J_i' = \begin{pmatrix} 0 & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ & & & 0 \end{pmatrix}$，每个子块的阶数不超过$r$，所以有$(J_i')^r = 0$，从而$(\lambda_0 I_2 - B_0)^r = 0$。因此，$(\lambda_0 I - A)^r = P\begin{pmatrix} (\lambda_0 I_1 - B)^r & \\ & 0 \end{pmatrix}P^{-1}$，但$\lambda_0 I_1 - B$是$n-r$阶非奇异的，故$rank(\lambda_0 I - A)^r = rank(\lambda_0 I_1 - B)^r = rank(\lambda_0 I_1 - B) = n - r$。

10. 下列矩阵A是否为正规矩阵？若是，试求酉矩阵U，使$U^{-1}AU$为对角矩阵：

(1) $A=\begin{pmatrix}-1&i&0\\-i&0&-i\\0&i&-1\end{pmatrix}$；(2) $A=\begin{pmatrix}0&i&1\\-i&0&0\\1&0&0\end{pmatrix}$.

解：(1) 因$A^H=A$，A为Hermite矩阵，所以为正规矩阵。

A的特征值为$\lambda_1=-1,\lambda_2=1,\lambda_3=-2$，对应的特征向量分别为

$$\xi_1=\begin{pmatrix}1\\0\\-1\end{pmatrix},\quad \xi_2=\begin{pmatrix}1\\-2i\\1\end{pmatrix},\quad \xi_3=\begin{pmatrix}1\\i\\1\end{pmatrix}.$$

将它们单位化得

$$\varepsilon_1=\frac{\xi_1}{\|\xi_1\|}=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\end{pmatrix},\quad \varepsilon_2=\frac{\xi_2}{\|\xi_2\|}=\frac{1}{\sqrt{6}}\begin{pmatrix}1\\-2i\\1\end{pmatrix},\quad \varepsilon_3=\frac{\xi_3}{\|\xi_3\|}=\frac{1}{\sqrt{3}}\begin{pmatrix}1\\i\\1\end{pmatrix}.$$

可得酉矩阵$U=(\varepsilon_1,\varepsilon_2,\varepsilon_3)$，且$U^{-1}AU=diag\,(-1,1,-2)$.

(2) A也为Hermite矩阵，所以也是正规矩阵。

A的特征值为$\lambda_1=0,\lambda_2=\sqrt{2},\lambda_3=-\sqrt{2}$. 对应的特征向量分别为

$$\xi_1=\begin{pmatrix}0\\i\\1\end{pmatrix},\quad \xi_2=\begin{pmatrix}\sqrt{2}\\-i\\1\end{pmatrix},\quad \xi_3=\begin{pmatrix}-\sqrt{2}\\-i\\1\end{pmatrix}.$$

单位化得$\varepsilon_1=\frac{\xi_1}{\|\xi_1\|}=\frac{1}{\sqrt{2}}\begin{pmatrix}0\\i\\1\end{pmatrix}$, $\varepsilon_2=\frac{\xi_2}{\|\xi_2\|}=\frac{1}{2}\begin{pmatrix}\sqrt{2}\\-i\\1\end{pmatrix}$, $\varepsilon_3=\frac{\xi_3}{\|\xi_3\|}=\frac{1}{2}\begin{pmatrix}-\sqrt{2}\\-i\\1\end{pmatrix}$. 故酉矩阵$U=(\varepsilon_1,\varepsilon_2,\varepsilon_3)=\begin{pmatrix}0&\frac{\sqrt{2}}{2}&-\frac{\sqrt{2}}{2}\\\frac{i}{\sqrt{2}}&-\frac{i}{2}&-\frac{i}{2}\\\frac{1}{\sqrt{2}}&\frac{1}{2}&\frac{1}{2}\end{pmatrix}$, $U^{-1}AU=\begin{pmatrix}0&&\\&\sqrt{2}&\\&&-\sqrt{2}\end{pmatrix}$.

11. 设$A=(a_{jk})\in \boldsymbol{C}^{n\times n}$，则$A$是Hermite矩阵的充分必要条件是对任意$x\in\boldsymbol{C}^n$，$x^HAx$ 是实数。

证明：必要性. 如果A是Hermite矩阵，则对任意$x\in C^n$，因x^HAx是数，所以$\overline{(x^HAx)}=\left(x^HAx\right)^H=x^HA^Hx=x^HAx$. 故$x^HAx$是实数。

充分性. 因对任意$x \in C^n$，x^HAx为实数，所以$x^HAx = \left(x^HAx\right)^H = x^HA^Hx$，即$x^H\left(A - A^H\right)x = 0$. 令$B = A - A^H = (b_{kj})$，则对任意$x \in C^n$，$x^HBx = 0$. 特别地，

(1) 取$x = (0, \cdots, 0, \underset{t}{1}, 0, \cdots, 0)^T$，有$x^HBx = b_{tt} = 0,\ t = 1, 2, \cdots, n$.

(2) 取$x = (0, \cdots, 0, \underset{k}{1}, 0, \cdots, 0, \underset{j}{1}, 0, \cdots, 0)^T$，有$x^HBx = b_{kk} + b_{kj} + b_{jk} + b_{jj} = 0$，由(1)知$b_{kj} + b_{jk} = 0$.。

(3) 取$x = (0, \cdots, 0, \underset{k}{1}, 0, \cdots, 0, \underset{j}{i}, 0, \cdots, 0)^T\left(i = \sqrt{-1}\right)$，有$x^HBx = b_{kk} + ib_{kj} - ib_{jk} + b_{jj} = 0$，因此$b_{kj} - b_{jk} = 0$，由(2)得$b_{kj} = b_{jk} = 0$，即$B$=0，所以$A$=$A^H$，即$A$是Hermite矩阵。

12. 设$A \in \boldsymbol{C}^{n\times n}$是Hermite矩阵。证明：$A$是Hermite正定矩阵的充分必要条件是存在Hermite正定矩阵S，使得$A = S^2$.

证明：必要性. 若A是Hermite正定矩阵，则A的特征值$\lambda_i > 0, i = 1, 2, \cdots, n$，且存在酉矩阵$U$，使得$U^HAU = diag\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right)$。于是

$A = Udiag\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n}\right)U^HUdiag\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n}\right)U^H = S^2$，其中

$S = Udiag\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n}\right)U^H$也是Hermite正定矩阵。

充分性. 若$A = S^2$，S为Hermite正定阵，则$A = SS = S^HS$，即得A正定。

13. 若A、B 均为n阶Hermite正定矩阵，且 $AB = BA$，则 AB 为正定矩阵。

证明：因$(AB)^H = B^HA^H = BA = AB$，所以$AB$是Hermite矩阵。又$A$、$B$正定，存在$n$阶可逆矩阵$P$、$Q$，使$A = P^HP,\ B = Q^HQ$.于是

$$Q\left(AB\right)Q^{-1} = QP^HPQ^HQQ^{-1} = \left(PQ^H\right)^H\left(PQ^H\right) = C,$$

其中PQ^H可逆。因此C为Hermite正定矩阵，从而得AB的特征值均大于零，即得AB为正定矩阵。

14. 设$A = (a_{ij}) \in \boldsymbol{C}^{n\times n}$为Hermite正定矩阵，证明：
(1) $\det A \le a_{nn}\det A_{n-1}$，其中$A_{n-1}$为$A$的$n-1$阶顺序主子阵，且等号成立当且仅当$a_{1n} = a_{2n} = \cdots = a_{n-1,n} = 0$；　(2) $\det A \le \prod\limits_{i=1}^{n} a_{ii}$.

证明：(1) 记$A = \begin{pmatrix} A_{n-1} & \alpha \\ \alpha^H & a_{nn} \end{pmatrix}$，其中$\alpha = (a_{1n}, a_{2n}, \cdots, a_{n-1,n})^T$.因$A_{n-1}$正定，所以$A_{n-1}$可逆，且$A_{n-1}^{-1}$也是正定矩阵。由分块矩阵的初等变换可得等式：

$$\begin{pmatrix} I_{n-1} & 0 \\ -\alpha^H A_{n-1}^{-1} & 1 \end{pmatrix}\begin{pmatrix} A_{n-1} & \alpha \\ \alpha^H & a_{nn} \end{pmatrix} = \begin{pmatrix} A_{n-1} & \alpha \\ 0 & a_{nn} - \alpha^H A_{n-1}^{-1}\alpha \end{pmatrix},$$

两边取行列式可得$\det A = \left(a_{nn} - \alpha^H A_{n-1}^{-1}\alpha\right)\det A_{n-1}$，而$\alpha^H A_{n-1}^{-1}\alpha \geq 0$，且等号成立的充要条件是$\alpha = 0$. 所以 $\det A \leq a_{nn}\det A_{n-1}$，且等号成立当且仅当$a_{1n} = a_{2n} = \cdots = a_{n-1,n} = 0$.

由(1)递推下去，有$\det A \in a_{nn}\det A_{n-1} \leq a_{nn}a_{n-1,n-1}\det A_{n-2} \leq \cdots \leq \prod_{i=1}^{n} a_{ii}$.

15. 设 $A = \begin{pmatrix} 0 & 0.8 & -1 \\ 0.1 & 2 & -0.1 \\ 0.2 & 0.7 & 3 \end{pmatrix}$, 用盖尔圆定理证明 A 有 3 个互异实特征值。

证明：

$$D = \begin{pmatrix} 1 & & \\ & \frac{1}{2} & \\ & & 1 \end{pmatrix}, \quad D^{-1}AD = \begin{pmatrix} 0 & 0.4 & -1 \\ 0.2 & 2 & -0.2 \\ 0.2 & 0.35 & 3 \end{pmatrix}$$

盖尔圆

$$G_1 : |z| \leq 1.4 \quad G_2 : |z-2| \leq 0.4 \quad G_3 : |z-3| \leq 0.55$$

它们相互分离。若实矩阵A的特征值λ为复数，则$\overline{\lambda}$也为特征值，但G_1, G_2, G_3关于实轴对称， 则A特征值必皆为实数，故在区间$[-1.4, 1.4], [1.6, 2.4], [2.45, 3.55]$中各有$A$的一个特征值。

16. 若$A = (a_{ij})_{n\times n}$ 按行（列）严格对角占优， 即 $|a_{ii}| > R_i$ $(i = 1, 2, \cdots, n)$, 则$\det A \neq 0$.

证明：$\det A \neq 0$ 的充要条件是 A 的特征值都不等于零。 如 0 是 A 的特征值，则存在盖尔圆 G_{i_0} 使 $0 \in G_{i_0}$, 即

$$|a_{i_0 i_0}| = |0 - a_{i_0 i_0}| \leq R_{i_0}$$

矛盾。

17. 设矩阵$A = \begin{pmatrix} 7 & 3 & -2 \\ 3 & 4 & -1 \\ -2 & -1 & 3 \end{pmatrix}$. 分别试用幂迭代法和逆幂迭代法求矩阵按模最大特征值和按模最小特征值(迭代5步)。

解：(1) 以$(1,1,1)^T$为初始点。

k	v_k^T			m_k
1	1	0.75	0	8
2	1	0.6486	-0.2973	9.25
3	1	0.6176	-0.3711	9.5405
4	1	0.6088	-0.3888	9.5949
5	1	0.6064	-0.3931	9.6041

5步迭代结束后得按模最大特征值为9.6041。

(2) 以$(1,1,1)^T$为初始点。

k	v_k^T			m_k
1	0.36	0.44	1	0.5435
2	0.2768	0.2806	1	0.4617
3	0.2943	0.1855	1	0.4492
4	0.3359	0.1013	1	0.4491
5	0.3843	0.0179	1	0.4518

5步迭代结束后按模最小特征值为 $\frac{1}{0.4518} = 2.2134$.

18. 设矩阵$A = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.

(1) 用幂迭代法求矩阵按模最大的特征值及其特征向量；

(2) 试取平移量$p = \frac{1}{2}(2+(2-\sqrt{3})) \approx 1.134$，用幂法求$A$按模最大特征值。$(\varepsilon = 10^{-4})$

解：(1) 以$(1,1,1)^T$为初始点。

k	v_k^T			m_k
1	1	1	0.5	4
2	1	0.875	0.375	4
3	1	0.8065	0.3226	3.875
4	1	0.7712	0.2966	3.8065
5	1	0.7528	0.2831	3.7712
6	1	0.7431	0.2760	3.7528
7	1	0.7380	0.2723	3.7431
8	1	0.7352	0.2703	3.7380
9	1	0.7337	0.2692	3.7352
10	1	0.733	0.2686	3.7337
11	1	0.7325	0.2683	3.7330
12	1	0.7323	0.2681	3.7325
13	1	0.7322	0.2681	3.7323
14	1	0.7321	0.2680	3.7322

迭代结束后按模最大得特征值为 3.7322, 对应特征向量为

$$(1, 0.7321, 0.2680)^T$$

(2) 以$(1, 1, 1)^T$为初始点。

k	v_k^T			m_k
1	1	1	0.3022	2.8660
2	1	0.7565	0.3348	2.8660
3	1	0.7588	0.2714	2.6225
4	1	0.7347	0.2752	2.6248
5	1	0.7350	0.2683	2.6007
6	1	0.7323	0.2688	2.6010
7	1	0.7324	0.2680	2.5983
8	1	0.7321	0.2680	2.5984
9	1	0.7321	0.2680	2.5981

迭代结束后按模最大得特征值为 $2.5981 + 1.134 = 3.7321$.

19. 试求矩阵$A=\begin{pmatrix} 5 & -4 & -7 \\ -6 & 7 & 11 \\ 6 & -6 & -10 \end{pmatrix}$最接近于2.05的特征值和特征向量。

解：$P(A-2.05I_3)=LU$,其中

$$P=\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad L=\begin{pmatrix} 1 & 0 & 0 \\ -0.4917 & 1 & 0 \\ -1 & 0.6704 & 1 \end{pmatrix}$$

$$U=\begin{pmatrix} -6 & 4.95 & 11 \\ 0 & -1.5662 & -1.5917 \\ 0 & 0 & 0.0170 \end{pmatrix}$$

以$(1,1,1)^T$为初始点。

k	v_k^T			m_k
1	-0.9480	1	-0.9686	-60.5933
2	-0.9972	1	-0.9985	-19.8270
3	-0.9999	1	-0.9999	-19.9783
4	-1	1	-1	-19.9987
5	-1	1	-1	-19.9999

迭代结束后得特征值为 $2.05+\frac{-1}{19.9999}=2$, 对应特征向量为 $(-1,1,-1)^T$.

第 三 章　习题解答与提示

1. 求$A=\begin{pmatrix}2&1&-3\\2&5&-4\\6&7&3\end{pmatrix}$的Doolittle分解。

解：$A=LU=\begin{pmatrix}1&0&0\\1&1&0\\3&1&1\end{pmatrix}\begin{pmatrix}2&1&-3\\0&4&-1\\0&0&13\end{pmatrix}.$

2. 求$A=\begin{pmatrix}1&2&-1\\2&\frac{2}{3}&\frac{1}{3}\\3&1&1\end{pmatrix}$的选列主元Doolittle分解。

解：$P=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix},\quad L=\begin{pmatrix}1&0&0\\\frac{1}{3}&1&0\\\frac{2}{3}&0&1\end{pmatrix},\quad U=\begin{pmatrix}3&1&1\\0&\frac{5}{3}&-\frac{4}{3}\\0&0&-\frac{1}{3}\end{pmatrix}.$

$$PA=LU=\begin{pmatrix}1&0&0\\\frac{1}{3}&1&0\\\frac{2}{3}&0&1\end{pmatrix}\begin{pmatrix}3&1&1\\0&\frac{5}{3}&-\frac{4}{3}\\0&0&-\frac{1}{3}\end{pmatrix}.$$

3. 求$A=\begin{pmatrix}2.25&-3&4.5\\-3&5&-10\\4.5&-10&34\end{pmatrix}$的Cholesky分解。

解：$A=LL^T=\begin{pmatrix}1.5&0&0\\-2&1&0\\3&-4&3\end{pmatrix}\begin{pmatrix}1.5&0&0\\-2&1&0\\3&-4&3\end{pmatrix}^T.$

4. 求下列矩阵的Hermite标准形和变换矩阵S，并求满秩分解

$$A=\begin{pmatrix}2&4&1&1\\1&2&-1&2\\-1&-2&-2&1\end{pmatrix},\qquad A=\begin{pmatrix}1&2&3&6\\2&4&3&6\\1&2&3&6\\2&4&6&12\end{pmatrix}$$

解：(1) $H=\begin{pmatrix}1&2&0&1\\0&0&1&-1\\0&0&0&0\end{pmatrix}$ $S=\begin{pmatrix}\frac{1}{3}&\frac{1}{3}&0\\\frac{1}{3}&-\frac{2}{3}&0\\1&-1&1\end{pmatrix}$

$A=FG=\begin{pmatrix}2&1\\1&-1\\-1&-2\end{pmatrix}\begin{pmatrix}1&2&0&1\\0&0&1&-1\end{pmatrix}$

(2) $H=\begin{pmatrix}1&2&0&0\\0&0&1&2\\0&0&0&0\\0&0&0&0\end{pmatrix}$ $S=\begin{pmatrix}-1&1&0&0\\\frac{2}{3}&-\frac{1}{3}&0&0\\-1&0&1&0\\-2&0&0&1\end{pmatrix}$

$B=FG=\begin{pmatrix}1&3\\2&3\\1&3\\2&6\end{pmatrix}\begin{pmatrix}1&2&0&0\\0&0&1&2\end{pmatrix}$

5. 设$A=\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix}$,

(1) 求A的奇异值分解；

(2) U为AA^H的正交单位化特征向量组成的酉阵，V为A^HA正交单位化特征向量组成的酉阵， 试验证能否由此构成A的奇异值分解。

解：(1) 因为 $A^TA=\begin{pmatrix}1&0&1\\0&1&1\\1&1&2\end{pmatrix}$ 其特征值为 $\lambda_1=3$， $\lambda_2=1$， $\lambda_3=0$， 对应的特征向量为

$$p_1=\begin{pmatrix}1\\1\\2\end{pmatrix},\quad p_2=\begin{pmatrix}-1\\1\\0\end{pmatrix},\quad p_3=\begin{pmatrix}-1\\-1\\1\end{pmatrix}$$

标准化得

$$V=\begin{pmatrix}\frac{1}{\sqrt{6}}&-\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{3}}\\\frac{2}{\sqrt{6}}&0&\frac{1}{\sqrt{3}}\end{pmatrix}$$

使得 $V^HA^HAV = \begin{pmatrix} 3 & & \\ & 1 & \\ & & 0 \end{pmatrix}$. 计算

$$U_1 = AV_1\Sigma^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 \end{pmatrix}$$

取 $U_2 = (0,0,1)^T$， 则 $U = (U_1, U_2)$ 是酉矩阵。 故 A 的奇异值分解为

$$A = U\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}V^H = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{3} & & \\ & 1 & \\ & & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

(2) 因为 $AA^T = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ 其特征值为 $\lambda_1 = 3$， $\lambda_2 = 1$， $\lambda_3 = 0$， 对应的特征向量为

$$p_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad p_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad p_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

标准化得

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$$

使得 $U^HAA^TU = \begin{pmatrix} 3 & & \\ & 1 & \\ & & 0 \end{pmatrix}$. 但是易计算得$A \neq U\begin{pmatrix} \sqrt{3} & & \\ & 1 & \\ & & 0 \end{pmatrix}V^T$.

6. 设$A \in C^{m\times n}$，证明：

(a) $(A^H)^+ = (A^+)^H$；

(b) $(A^HA)^+ = A^+(A^H)^+$，$(AA^H)^+ = (A^H)^+A^+$.

解：(1) 直接验证$X = (A^+)^H$满足A^H 的四个Penrose方程即可；

(2) 按定义直接验证四个Penrose方程即可。

7. 设 A 是一个正规矩阵，证明：

(a) $AA^+ = A^+A$.

(b) 对于任意自然数 k，有 $(A^k)^+ = (A^+)^k$

解：　由于 $A \in C^{n\times n}$为正规矩阵，因此存在酉矩阵U，使得

$$A = Udiag(\lambda_1, \cdots, \lambda_n)U^H$$

其中$\lambda_1,\cdots,\lambda_n$ 为A的n个特征值。则有

$$A^+ = U^H diag(\lambda_1^+,\cdots,\lambda_n^+)U$$

将上式直接代入验证，可得　$AA^+ = A^+A$, $(A^k)^+ = (A^+)^k$.

8. 设$A = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{bmatrix}$，证明：$A^+ = \begin{bmatrix} A_1^+ & & & \\ & A_2^+ & & \\ & & \ddots & \\ & & & A_s^+ \end{bmatrix}$.

解：　利用分块矩阵乘法，直接验证四个Penrose方程。

9. 利用矩阵的奇异值分解求下面矩阵的 Moore-Penrose逆 A^+:

(1) $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$;　(2) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

(3) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

解：(1) 由于 $A^HA = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$， 特征值 $\lambda_1 = 3, \lambda_2 = 1$, 对应特征向量为 $v_1 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T, v_2 = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})^T$. 从而有

$$\Sigma = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix}, V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

因此有

$$U_1 = AV_1\Sigma^{-1} = \begin{pmatrix} \frac{\sqrt{6}}{3} & 0 \\ \frac{\sqrt{6}}{6} & \frac{1}{\sqrt{2}} \\ \frac{\sqrt{6}}{6} & -\frac{1}{\sqrt{2}} \end{pmatrix}, \ V_1 = V$$

取$U_2 \in C^{3\times 1}$，使得$U = (U_1\ U_2)$ 为3阶酉矩阵，故 $U_2 = (-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})^T$，从而$A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^H$,所以

$$A^+ = V \begin{pmatrix} \Sigma^{-1} & 0 \end{pmatrix} U^H = \frac{1}{3}\begin{pmatrix} 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$$

(2)由于 $A^HA = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}$， 特征值 $\lambda_1 = 6, \lambda_2 = \lambda_3 = 0$, 对应特征向量为

$v_1=(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}})^T, v_2=(\frac{1}{\sqrt{3}},-\frac{2}{\sqrt{3}},\frac{1}{\sqrt{3}})^T, v_3=(\frac{1}{\sqrt{2}},0,-\frac{1}{\sqrt{2}})^T$. 从而有

$$\Sigma=(\sqrt{6}), V=\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

取$V_1=v_1$， 因此有

$$U_1=AV_1\Sigma^{-1}=\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

取$U_2\in C^{2\times1}$，使得$U=(U_1\ U_2)$ 为2阶酉矩阵，故 $U_2=(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}})^T$， 从而$A=U\begin{pmatrix} \Sigma & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}V^H$, 从而

$$A^+=V\begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}U^H=\frac{1}{6}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$$

(3)由于 $A^HA=\begin{pmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{pmatrix}$， 特征值 $\lambda_1=9,\lambda_2=\lambda_3=0$, 对应特征向量为 $v_1=(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}})^T, v_2=(\frac{1}{\sqrt{3}},-\frac{2}{\sqrt{3}},\frac{1}{\sqrt{3}})^T, v_3=(\frac{1}{\sqrt{2}},0,-\frac{1}{\sqrt{2}})^T$. 从而有

$$\Sigma=(3),\ V=\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

因此有

$$U_1=AV_1\Sigma^{-1}=(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}})^T, V_1=v_1,$$

$$U=(U_1,U_2)=\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

使得 $A=U\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}V^H$, 从而

$$A^+=V\begin{pmatrix} \sum^{-1} & 0 \\ 0 & 0 \end{pmatrix}U^H=\frac{1}{9}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

10. 利用矩阵的满秩分解求下面矩阵的 Moore-Penrose逆 A^+:

(1) $A=\begin{bmatrix}1&0&2\\2&1&5\\0&1&-1\\1&3&-1\end{bmatrix}$; (2) $A=\begin{bmatrix}4&-1&-3&-2\\-2&5&-1&-3\\2&13&-9&-5\end{bmatrix}$;

(3) $A=\begin{bmatrix}1&3&3&2\\2&6&9&5\\-1&-3&3&0\end{bmatrix}$; (4) $A=\begin{bmatrix}1&1&0&1\\0&1&1&0\\1&2&1&1\end{bmatrix}$.

解：(1)A 为列满秩矩阵，因此 $A=AI_3=FG$，从而

$$A^+=(A^HA)^{-1}A^H=\begin{pmatrix}6&5&11\\5&11&1\\11&1&31\end{pmatrix}^{-1}\begin{pmatrix}1&2&0&1\\0&1&1&3\\2&5&-1&-1\end{pmatrix}$$

$$=\tfrac{1}{22}\begin{pmatrix}54&-22&14&12\\-23&11&8&1\\-17&11&4&-5\end{pmatrix}$$

(2)A 为列行秩矩阵，因此 $A=I_4A=FG$，从而

$$A^+=A^H(AA^H)^{-1}=A^H\begin{pmatrix}30&-4&32\\-4&39&85\\32&85&279\end{pmatrix}^{-1}$$

$$=\tfrac{1}{1360}\begin{pmatrix}118&-147&41\\-160&-60&100\\-16&164&-92\\-340&-510&170\end{pmatrix}$$

(3)由于 $A=FG=\begin{pmatrix}1&3\\2&9\\-1&3\end{pmatrix}\begin{pmatrix}1&3&0&1\\0&0&1&\frac{1}{3}\end{pmatrix}$，因此

$$A^+=G^H(F^HAG^H)^{-1}F^H=G^H\begin{pmatrix}72&22\\231&116\end{pmatrix}^{-1}F^H$$

$$=\frac{1}{3270}\begin{pmatrix}50&34&-182\\150&102&-546\\-15&186&447\\45&96&-33\end{pmatrix}$$

(4)由于$A=FG=\begin{pmatrix}1&1\\0&1\\1&2\end{pmatrix}\begin{pmatrix}1&0&-1&1\\0&1&1&0\end{pmatrix}$,因此

$$A^+=G^H(F^HAG^H)^{-1}F^H=G^H\begin{pmatrix}3&4\\3&9\end{pmatrix}^{-1}F^H$$

$$=\frac{1}{15}\begin{pmatrix}5&-4&1\\0&3&3\\-5&7&2\\5&-4&1\end{pmatrix}$$

11. 用广义逆矩阵A^+方法判断线性方程组

$$\begin{cases}2x_1+4x_2+x_3+x_4=3\\x_1+2x_2-x_3+2x_4=0\\-x_1-2x_2-2x_3+x_4=3\end{cases}$$

是否有解？如果有解，求通解和极小范数解。

解：令

$$A=\begin{pmatrix}2&4&1&1\\1&2&-1&2\\-1&-2&-2&1\end{pmatrix},\ \boldsymbol{b}=\begin{pmatrix}3\\0\\3\end{pmatrix}.$$

容易求得A的满秩分解为

$$A=FG=\begin{pmatrix}2&1\\1&-1\\-1&-2\end{pmatrix}\begin{pmatrix}1&2&0&1\\0&0&1&-1\end{pmatrix}$$

所以

$$A^+ = G^H(GG^H)^{-1}(F^HF)^{-1}F^H = \frac{1}{33}\begin{pmatrix} 2 & 1 & -1 \\ 4 & 2 & -2 \\ 1 & -5 & -6 \\ 1 & 6 & 5 \end{pmatrix}$$

可以验证

$$AA^+b = (1,2,1)^T \neq \boldsymbol{b},$$

所以原方程组无解。

$$\boldsymbol{x} = A^+\boldsymbol{b} + (I - A^+A)y = \frac{1}{11}\begin{pmatrix} 1 \\ 2 \\ -5 \\ 6 \end{pmatrix} + \frac{1}{11}\begin{pmatrix} 9 & -4 & -1 & -1 \\ -4 & 3 & -2 & -2 \\ -1 & -2 & 5 & 5 \\ -1 & -2 & 5 & 5 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

极小范数最小二乘解为

$$\boldsymbol{x}_0 = A^+\boldsymbol{b} = \frac{1}{11}\begin{pmatrix} 1 \\ 2 \\ -5 \\ 6 \end{pmatrix}$$

第 四 章　习题解答与提示

1. 用 Gauss 消去法和 Doolittle 方法求解

$$\begin{pmatrix} 6 & 2 & 1 & -1 \\ 2 & 4 & 1 & 0 \\ 1 & 1 & 4 & -1 \\ -1 & 0 & -1 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \\ 5 \\ -5 \end{pmatrix}.$$

解： $\boldsymbol{x} = (0.7906, -0.3613, 0.8639, -1.1152)^T$.

2. 用列主元法求解

$$\begin{pmatrix} -0.001 & 1 & 1 \\ 1 & 0.8 & 0 \\ 4 & 5.5 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0.2 \\ 1.4 \\ 7.4 \end{pmatrix}.$$

解： $\boldsymbol{x} = (1.8671, -0.5838, 0.7857)^T$.

3. 用改进平方根法解

$$\begin{pmatrix} 16 & 4 & 8 \\ 4 & 5 & -4 \\ 8 & -4 & 22 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 5 \end{pmatrix}.$$

解： $\boldsymbol{x} = (-1.2743, 2.5417, 1.1528)^T$.

4. 设 A 是对称正定的三对角矩阵，

$$A = \begin{pmatrix} a_1 & c_1 & & & 0 \\ u_2 & a_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & u_{n-1} & a_{n-1} & c_{n-1} \\ 0 & & & u_n & a_n \end{pmatrix},$$

试给出用 Cholesky 方法解 $A\boldsymbol{x} = \boldsymbol{b}$ 的计算公式。

解： 设 $A = (a_{i,j})$, 因为 A 是对称正定的三对角矩阵，所以

$$a_{i,i-1} = u_i = a_{i-1,i} = c_{i-1},\ (i = 2, 3, \cdots, n),$$

且

$$A=\begin{pmatrix} d_{11} & & & \\ s_{21} & d_{22} & & \\ \vdots & \vdots & \ddots & \\ s_{n1} & s_{n2} & \cdots & d_{nn} \end{pmatrix}\begin{pmatrix} 1 & l_{21} & \cdots & l_{n1} \\ & 1 & \cdots & l_{n2} \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix}$$

式中 $s_{ij}=l_{ij}d_{jj}$。当 $i-j>1$ 时，$s_{ij}=0$，即只有 $s_{i,i-1}\ (i=2,3,\cdots,n)$ 可能不为零。

利用 Cholesky 分解公式得

$$\begin{cases} d_{11}=a_{11}=a_1 \\ s_{i,i-1}=a_{i,i-1}-\sum\limits_{i=1}^{i-2}s_{i,k}l_{i-1,k}=a_{i,i-1}=u_i \\ l_{i,i-1}=s_{i,i-1}/d_{i-1,i-1}=u_i/d_{i-1,i-1}=c_{i-1}/d_{i-1,i-1} \\ d_{ii}=a_{ii}-\sum\limits_{k=1}^{i-1}s_{ik}l_{ik}=a_i-s_{i,i-1}l_{i,i-1}=a_i-u_il_{i,i-1} \end{cases} \qquad i=2,3,\cdots,n.$$

解方程组按如下步骤进行

$$\begin{cases} y_i=b_i-\sum\limits_{k=1}^{i-1}l_{ik}y_k=b_i-l_{i,i-1}y_i \quad (i=1,2,\cdots,n) \\ x_i=y_i/d_{ii}-\sum\limits_{k=i+1}^{n}l_{ki}x_k=y_i/d_{ii}-l_{i+1,i}x_{i+1}=y_i/d_{ii}-l_{i+1,i}x_{i+1} \\ \qquad =(y_i-c_ix_{i+1})/d_{ii} \quad (i=n,n-1,\cdots,1). \end{cases}$$

5. 用追赶法解三对角方程组

$$\begin{bmatrix} 4 & -1 & & \\ -2 & 4 & -1 & \\ & -2 & 4 & -1 \\ & & -2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}=\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

解：$\boldsymbol{x}=(0.2988,0.1951,0.1829,0.3415)^T$.

6. 用广义逆矩阵的方法求矛盾方程组

$$\begin{cases} x_1-x_3+x_4=4 \\ 2x_2+2x_3+2x_4=1 \\ -x_1+4x_2+5x_3+3x_4=2 \end{cases}$$

的全部最小二乘解和极小范数最小二乘解。

解：系数矩阵 A 的满秩分解为：

$$A=FG=\begin{pmatrix} 1 & 0 \\ 0 & 2 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}$$

$$A^+ = \frac{1}{36}\begin{pmatrix} 10 & 4 & -2 \\ 2 & 2 & 2 \\ -8 & -2 & 4 \\ 12 & 6 & 0 \end{pmatrix}.$$

极小范数最小二乘解为

$$\boldsymbol{x}_{ls} = \frac{1}{18}(20, 7, -13, 27)^T.$$

全部最小二乘解为

$$\boldsymbol{x} = \boldsymbol{x}_{ls} + \frac{1}{3}\begin{pmatrix} 2 & 0 & 1 & -1 \\ 0 & 2 & -1 & -1 \\ 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}, \quad (y_1, y_2, y_3, y_4 \in C \text{ 任意})$$

7. 用矩阵 QR 分解的方法求矛盾方程组

$$\begin{cases} 3x_2 + x_3 = 0 \\ 4x_2 - 2x_3 = 0 \\ 2x_1 + x_2 + 2x_3 = 0 \\ 2x_1 + 4x_2 + 3x_3 = 1 \end{cases}$$

的最小二乘解。

解：系数矩阵 A 的QR 分解：

$$A = QR = \begin{pmatrix} 0.0 & 0.55 & 0.6 \\ 0.0 & -0.74 & -0.68 \\ -0.71 & 0.28 & -0.3 \\ -0.71 & -0.28 & 0.3 \end{pmatrix}\begin{pmatrix} -1.4 & -3.54 & -3.54 \\ 0.0 & -5.43 & 0.644 \\ 0.0 & 0.0 & 2.255 \end{pmatrix}.$$

解 $Rx = Q^Tb$ 得最小二乘解 $\boldsymbol{x} = (0, 0.066667, 0.133333)^T$.

8. 用矩阵奇异值分解的方法求矛盾方程组

$$\begin{cases} x_1 + x_3 = 0 \\ x_1 + x_2 = 0 \\ 2x_1 + x_2 + x_3 = 1 \end{cases}$$

的全部最小二乘解和极小范数最小二乘解。

解：$x_{ls} = \frac{1}{9}(2, 1, 1)^T$，全部最小二乘解为

$$x = \frac{1}{9}\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \quad (y_1, y_2, y_3 \in C \text{ 任意}).$$

9. 设 $A\in \boldsymbol{R}^{m\times n}$ 列满秩，令条件数 $\mathrm{cond}_2(A)=\|A\|_2\|A^+\|_2$，证明 $\mathrm{cond}_2(A)=\sqrt{\|A^TA\|_2\|(A^TA)^{-1}\|_2}$.

证明：设 A^TA 的特征值为

$$\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_r>\lambda_{r+1}=\cdots=\lambda_n=0.$$

则 $\|A\|_2=\sqrt{\lambda_1}$，$\left\|A^TA\right\|_2=\lambda_1$，$\left\|\left(A^TA\right)^{-1}\right\|_2=\frac{1}{\lambda_r}$.

利用A的奇异值分解知，$\left(A^+\right)^TA^+$ 与 $\begin{pmatrix} B & 0\\ 0 & 0\end{pmatrix}$ 酉相似

（其中 $B=\mathrm{diag}(\frac{1}{\lambda_1},\frac{1}{\lambda_2},\cdots,\frac{1}{\lambda_r})$），所以 $\|A^+\|_2=\frac{1}{\sqrt{\lambda_r}}$.

10. 设 $A=\begin{pmatrix} 1 & \alpha & \beta\\ 1 & 1 & 1\\ -\beta & \alpha & 1\end{pmatrix}$，其中 α,β 是实参数。对方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 来说，在 (α,β) 平面什么范围内 J 法收敛。

解：迭代矩阵 B_J的特征方程为 $\det[\lambda D+(L+U)]=0$, 即

$$\lambda\left(\lambda^2-2\alpha+\beta^2\right)=0,$$

所以 $\rho\left(B_J\right)=|2\alpha-\beta^2|^{1/2}$，当 $|2\alpha-\beta^2|<1$ 时，J 法收敛。

11. 分别用 J 法，G-S 法和 $\alpha=1.46$ 的 SOR 法解方程组

$$\begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1\\ & & -1 & 2\end{pmatrix}\begin{pmatrix} x_1\\ x_2\\ x_3\\ x_4\end{pmatrix}=\begin{pmatrix} 1\\ 0\\ 1\\ 0\end{pmatrix}$$

并写出相应的迭代矩阵。

解：取近似解的精度为 10^{-4}, 用J法求解，迭代次数为 42, 近似解为

$$x_{42}=(1.199857,1.399813,1.599769,0.799885)^T.$$

用G-S 法 求解，迭代次数为 22, 近似解为

$$x_{22}=(1.199876,1.399838,1.599869,0.799934)^T.$$

用SOR法求解，迭代次数为 14, 近似解为

$$x_{14}=(1.199983,1.399977,1.600001,0.800008)^T.$$

12. 设

$$A=\begin{pmatrix}1&2&-2\\1&1&1\\2&2&1\end{pmatrix},\quad B=\begin{pmatrix}2&-1&2\\1&1&1\\1&1&-2\end{pmatrix}$$

证明：（1）若线性方程组的系数矩阵为 A，则 J 法收敛而 G-S 法不收敛；（2）若线性方程组的系数矩阵为 B，则 J 法不收敛而 G-S 法收敛。

证明：（1）因为 $\rho(B_J)=0,\quad \rho(B_G)=2$ 所以J 法收敛，而 G-S 法发散。

（2）因为 $\rho(B_G)=\frac{1}{2}$ 所以G-S 法收敛。

B_J 的特征方程为 $\lambda^3+\frac{3}{2}\lambda-\frac{1}{4}=0$, 由中值定理， B_J 有一实特征值 λ_1 属于$(0,0.2)$. 设 B_J 的另外两个特征值为 λ_2,λ_3，因 $\lambda_1\lambda_2,\lambda_3=\frac{1}{4}$, 则

$$|\lambda_2|\,|\lambda_3|=\frac{1}{4\,|\lambda_1|}>1$$

所以 $\rho(B_J)>1$， J法发散。

13. 设 $A=\begin{pmatrix}\alpha&1&3\\1&\alpha&2\\-3&2&\alpha\end{pmatrix}$, α 为何值时，J 法收敛？$\alpha=3$ 时怎样?

解：因为 $\rho(B_J)=\frac{2}{|\alpha|}$, 所以当 $|\alpha|>2$ 时J 法收敛。

14. 用迭代公式 $\boldsymbol{x}^{(j+1)}=\boldsymbol{x}^{(j)}-\alpha(A\boldsymbol{x}^{(j)}-\boldsymbol{b})$，求解 $A\boldsymbol{x}=\boldsymbol{b}$.

(1) 若 $A=\begin{pmatrix}3&2\\1&2\end{pmatrix}$，问取什么范围的 α 可使迭代收敛，什么 α 使迭代收敛最快?

(2) 若 $A\in\boldsymbol{R}^{n\times n}$，且 A 对称正定，最大和最小特征值是 λ_1 和 λ_n, 求迭代矩阵 $I-\alpha A$ 的谱半径，证明迭代收敛的充要条件是 $0<\alpha<2\lambda_1^{-1}$. 并问什么参数 α 使 $\rho(I-\alpha A)$ 最小?

证明：迭代矩阵 $B=I-\alpha A$, 若 A 的特征值为 λ，则 B 的特征值为 $1-\alpha\lambda$。

（1）A 的特征值为 1, 4，则 B 的特征值为 $1-\alpha$, $1-4\alpha$，则 $|1-\alpha|<1$, $|1-4\alpha|<1$, 即 $0<\alpha<\frac{1}{2}$ 时迭代收敛，要迭代收敛最快，还需 $|1-\alpha|=|1-4\alpha|$,即 $\alpha=0.4$.

（2）因为 A 对称正定，所以 A 的特征值 λ 大于 0. 若 $\alpha<0$， 则 迭代矩阵 B 的特征值 $1-\alpha\lambda$ 大于 1, $\rho(B)>1$, 迭代发散。所以迭代收敛的充要条件为 $\alpha>0$ 及 A 的任意特征值 λ 满足 $|1-\alpha\lambda|<1$，即 $0<\alpha<2\lambda^{-1}$，所以迭代收敛的充要条件为 $0<\alpha<2\lambda_1^{-1}$. 要迭代收敛最快，还需 $|1-\alpha\lambda_1|=|1-\alpha\lambda_n|$, 即 $\alpha=\frac{2}{\lambda_1+\lambda_2}$.

15. 给定迭代式 $\boldsymbol{x}^{(j+1)} = B\boldsymbol{x}^{(j)} + \boldsymbol{f}$，其中 $B = \begin{pmatrix} \alpha & 4 \\ 0 & \alpha \end{pmatrix}$, $0 < \alpha < 1$，证明在迭代过程中误差的范数 $\|\|_\infty$ 从某次迭代 j_0 开始单调减少（j_0 由充分接近 1 的 α 确定）。

证明： $B^j = \begin{pmatrix} \alpha^j & 4j\alpha^{j+1} \\ 0 & \alpha^j \end{pmatrix}$， 则 $\|B^j\|_\infty = \alpha^j + 4j\alpha^{j-1}$. 用 $e^{(j)}$ 表示第 j 次迭代的误差，则

$$\left\|e^{(j)}\right\|_\infty = \left\|B^j e^{(0)}\right\|_\infty \leq \left\|e^{(0)}\right\|_\infty \left\|B^j\right\|_\infty = \left\|e^{(0)}\right\|_\infty \left(\alpha^j + 4j\alpha^{j-1}\right) = \varepsilon^{(j)}.$$

当 $j > \frac{\alpha^2+3\alpha}{4(1-\alpha)}$时，$\varepsilon^{(j)}$ 单调减少。

16. 若 $\rho(B) = 0$，试证对任意的 $\boldsymbol{x}^{(0)}$，迭代式 $\boldsymbol{x}^{(j+1)} = B\boldsymbol{x}^{(j)} + \boldsymbol{f}$ 最多 n 次迭代就可以得到精确解（提示:考虑 B 及 B^k 的 Jordan 标准形）。

证明： 因为存在约当标准型 J 及可逆矩阵 P，使得 $B = P^{-1}JP$，所以 $B^n = P^{-1}J^nP$. 因为 $\rho(B) = 0$, 所以 B 的特征值均为 0，那么 J 的对角线上的元素均为 0，则 $J^n = 0$。 用 $e^{(n)}$ 表示第 n 次迭代的误差，则 $e^{(n)} = B^n e^{(0)} = 0$.

17. 取 $\boldsymbol{x}^{(0)} = 0$，用 CG 法求解

$$\begin{pmatrix} 6 & 3 & 1 \\ 3 & 5 & 1 \\ 1 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}.$$

解： 取初始近似值 $x_0 = (0, 0, 0)^T$, 则可得

$$x_1 = (0.0000, -0.38460.7692)^T, \qquad x_2 = (0.0472, -0.37010.7874)^T,$$

$$x_3 = (0.06897, -0.3966, 0.7759)^T, \qquad \|Ax_3 - b\|_\infty < 10^{-8}.$$

第 五 章　习题解答与提示

1. 用单纯形方法求解

$$
(1)\quad \begin{aligned} \min\quad & -3x_1 - x_2 \\ s.t.\quad & 3x_1 + 3x_2 + x_3 = 30 \\ & 4x_1 - 4x_2 + x_4 = 16 \\ & 2x_1 - x_2 \le 12 \\ & x_i \ge 0, i = 1, \cdots 4 \end{aligned}
\qquad
(2)\quad \begin{aligned} \min\quad & 3x_1 - 5x_2 - 2x_3 - x_4 \\ s.t.\quad & x_1 + x_2 + x_3 \le 4 \\ & 4x_1 - x_2 + x_3 + 2x_4 \le 6 \\ & -x_1 + x_2 + 2x_3 + 3x_4 \le 12 \\ & x_i \ge 0, i = 1, \cdots 4 \end{aligned}
$$

解： (1) 引入松弛变量x_5,转化为标准型

$$
\begin{aligned} \min\quad & -3x_1 - x_2 \\ s.t.\quad & 3x_1 + 3x_2 + x_3 = 30 \\ & 4x_1 - 4x_2 + x_4 = 16 \\ & 2x_1 - x_2 + x_5 = 12 \\ & x_i \ge 0, i = 1, \cdots 5. \end{aligned}
$$

	x_1	x_2	x_3	x_4	x_5	
x_3	3	3	1	0	0	30
x_4	$\boxed{4}$	-4	0	1	0	16
x_5	2	-1	0	0	1	12
	3	1	0	0	0	0

	x_1	x_2	x_3	x_4	x_5	
x_3	0	$\boxed{6}$	1	$-\frac{3}{4}$	0	18
x_1	1	-1	0	$\frac{1}{4}$	0	4
x_5	0	1	0	$-\frac{1}{2}$	1	4
	0	4	0	$-\frac{3}{4}$	0	-12

	x_1	x_2	x_3	x_4	x_5	
x_3	0	1	$\frac{1}{6}$	$-\frac{1}{8}$	0	3
x_1	1	0	$\frac{1}{6}$	$\frac{1}{8}$	0	7
x_5	0	0	$-\frac{1}{6}$	$\frac{3}{8}$	1	1
	0	0	$-\frac{2}{3}$	$-\frac{1}{4}$	0	-24

所以$x^* = (7,3,0)$，最优目标函数值$f_{min} = -24$.

(2) 引入松弛变量x_5, x_6, x_7,转化为标准型

$$
\begin{aligned}
\min \quad & 3x_1 - 5x_2 - 2x_3 - x_4 \\
s.t. \quad & x_1 + x_2 + x_3 + x_5 = 4 \\
& 4x_1 - x_2 + x_3 + 2x_4 + x_6 = 6 \\
& -x_1 + x_2 + 2x_3 + 3x_4 + x_7 = 12 \\
& x_i \geq 0, i = 1, \cdots 7.
\end{aligned}
$$

单纯法求解过程：

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
x_5	1	$\boxed{1}$	1	0	1	0	0	4
x_6	4	-1	1	2	0	1	0	6
x_7	-1	1	2	3	0	0	1	12
	-3	5	2	1	0	0	0	0

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
x_2	1	1	1	0	1	0	0	4
x_6	5	0	2	2	1	1	0	10
x_7	-2	0	1	$\boxed{3}$	-1	0	1	8
	-8	0	-3	1	-5	0	0	-20

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
x_2	1	1	1	0	1	0	0	4
x_6	$\frac{19}{3}$	0	$\frac{4}{3}$	0	$\frac{5}{3}$	1	$-\frac{2}{3}$	$\frac{14}{3}$
x_4	$-\frac{2}{3}$	0	$\frac{1}{3}$	1	$-\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{8}{3}$
	$-\frac{22}{3}$	0	$-\frac{10}{3}$	0	$-\frac{14}{3}$	0	$-\frac{1}{3}$	$-\frac{68}{3}$

所以$x^* = (0,4,0,\frac{8}{3})$，最优目标函数值$f_{min} = -\frac{68}{3}$.

2. 用两阶段法求解下列问题

$$
\begin{aligned}
\min \quad & x_1 - 3x_2 + x_3 \\
s.t. \quad & 2x_1 - x_2 + x_3 = 8 \\
& 2x_1 + x_2 \geq 2 \\
& x_1 + 2x_2 \leq 10 \\
& x_i \geq 0, i = 1,2,3
\end{aligned}
$$

解：引入剩余变量x_4, 松弛变量x_5， 将原问题转化为标准型

$$\begin{aligned} \min \quad & x_1 - 3x_2 + x_3 \\ s.t. \quad & 2x_1 - x_2 + x_3 = 8 \\ & 2x_1 + x_2 - x_4 = 2 \\ & x_1 + 2x_2 + x_5 = 10 \\ & x_i \geq 0, i = 1,2,3,4,5. \end{aligned}$$

引入人工变量x_6，构造第一阶段问题

$$\begin{aligned} \min \quad & x_6 \\ s.t. \quad & 2x_1 - x_2 + x_3 = 8 \\ & 2x_1 + x_2 - x_4 + x_6 = 2 \\ & x_1 + 2x_2 + x_5 = 10 \\ & x_i \geq 0, i = 1,2,3,4,5,6. \end{aligned}$$

	x_1	x_2	x_3	x_4	x_5	x_6	
x_3	2	-1	1	0	0	0	8
x_6	$\boxed{2}$	1	0	-1	0	1	2
x_5	1	2	0	0	1	0	10
	2	1	0	-1	0	0	2

	x_1	x_2	x_3	x_4	x_5	x_6	
x_3	0	-2	1	1	0	-1	6
x_1	1	$\frac{1}{2}$	0	-$\frac{1}{2}$	0	$\frac{1}{2}$	1
x_5	0	$\frac{3}{2}$	0	$\frac{1}{2}$	1	$-\frac{1}{2}$	9
	0	0	0	0	0	-1	0

则得到原问题的一个基本可行解$x = (1,0,6,0,9)$. 从上表中划去第6列，并按标准型的目标函数系数重新计算判别数及目标函数值，得

	x_1	x_2	x_3	x_4	x_5	
x_3	0	-2	1	1	0	6
x_1	1	$\boxed{\frac{1}{2}}$	0	-$\frac{1}{2}$	0	1
x_5	0	$\frac{3}{2}$	0	$\frac{1}{2}$	1	9
	0	$\frac{3}{2}$	0	$\frac{1}{2}$	0	7

	x_1	x_2	x_3	x_4	x_5	
x_3	4	0	1	-1	0	10
x_2	2	1	0	-1	0	2
x_5	-3	0	0	$\boxed{2}$	1	6
	-3	0	0	2	0	4

	x_1	x_2	x_3	x_4	x_5	
x_3	$\frac{5}{2}$	0	1	0	$\frac{1}{2}$	13
x_2	$\frac{1}{2}$	1	0	0	$\frac{1}{2}$	5
x_4	$-\frac{3}{2}$	0	0	1	$\frac{1}{2}$	3
	0	0	0	0	-1	-2

所以最优解为$x^* = (0, 5, 13)$, 最优目标函数值$f_{min} = -2$.

3. 用大M法求解

$$\begin{aligned} \min \quad & 3x_1 - 2x_2 + x_3 \\ s.t. \quad & 2x_1 + x_2 - x_3 \le 5 \\ & 4x_1 + 3x_2 + x_3 \ge 3 \\ & -x_1 + x_2 + x_3 = 2 \\ & x_1, x_2, x_3 \ge 0 \end{aligned} .$$

解： 构造辅助问题：

$$\begin{aligned} \min \quad & 3x_1 - 2x_2 + x_3 + Mx_6 + Mx_7 \\ s.t. \quad & 2x_1 + x_2 - x_3 + x_4 = 5 \\ & 4x_1 + 3x_2 + x_3 - x_5 + x_6 = 3 \\ & -x_1 + x_2 + x_3 + x_7 = 2 \\ & x_i \ge 0, \ (i = 1, \cdots 7) \end{aligned}$$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
x_4	2	1	-1	1	0	0	0	5
x_6	4	$\boxed{3}$	1	0	-1	1	0	3
x_7	-1	0	1	0	0	0	1	2
	3M-3	4M+2	2M-1	0	-M	0	0	5M

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
x_4	$\frac{2}{3}$	0	-$\frac{4}{3}$	1	$\frac{1}{3}$	-$\frac{1}{3}$	0	4
x_2	$\frac{4}{3}$	1	$\frac{1}{3}$	0	-$\frac{1}{3}$	$\frac{1}{3}$	0	1
x_7	-$\frac{7}{3}$	0	$\boxed{\frac{2}{3}}$	0	$\frac{1}{3}$	-$\frac{1}{3}$	1	1
	-$\frac{1}{3}(7M + 17)$	0	$\frac{2}{3}M - \frac{5}{3}$	0	$\frac{1}{3}M + \frac{2}{3}$	$-\frac{4}{3}M - \frac{2}{3}$	0	M-2

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
x_4	-4	0	0	1	1	-1	2	6
x_2	$\frac{5}{2}$	1	0	0	-$\frac{1}{2}$	$\frac{1}{2}$	-$\frac{1}{2}$	$\frac{1}{2}$
x_3	-$\frac{7}{2}$	0	1	0	$\boxed{\frac{1}{2}}$	-$\frac{1}{2}$	$\frac{3}{2}$	$\frac{3}{2}$
	-$\frac{23}{2}$	0	0	0	$\frac{3}{2}$	$-M-\frac{3}{2}$	$-M+\frac{5}{2}$	$\frac{1}{2}$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
x_4	3	0	-2	1	0	0	-1	3
x_2	-1	1	1	0	0	0	1	2
x_5	-7	0	2	0	1	-1	3	3
	-1	0	-3	0	0	-M	-M-2	-4

4. 求解下列问题的稳定点，并判断是否是问题的局部极小解。

$$\begin{aligned} \min \quad & -3x_1 + x_2 - 4x_3^2 \\ s.t. \quad & -x_1 + x_2 + x_3^2 = 0 \\ & 1 - x_2 - x_3 \geq 0 \end{aligned}$$

解：

$$\nabla f(x) = \begin{pmatrix} -3 \\ 1 \\ 8x_3 \end{pmatrix}, \ \nabla h(x) = \begin{pmatrix} -1 \\ 1 \\ 2x_3 \end{pmatrix}, \nabla g(x) = \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}.$$

所以

$$\begin{aligned} & -3 + w = 0 \\ & 1 - w - v = 0 \\ & 8x_3 - 2wx_3 + v = 0 \\ & v(1 - x_2 - x_3) = 0 \\ & -x_1 + x_2 + x_3^2 = 0 \\ & 1 - x_2 - x_3 \geq 0 \end{aligned}$$

解得： $w = 3,\ v = 2,\ x = (3, 2, -1)$. 且$x$为KKT点。

计算得：

$$G = \{d | d_1 \in R,\ d_2 + d_3 = 0\},$$

则由$d \neq 0$有$d_i \neq 0,\ (i = 1, 2, 3)$. 又因为

$$\nabla_{xx}^2 L = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

则

$$d^T\nabla^2_{xx}Ld = 2d_3^2 > 0.$$

因此，x为严格局部极小解。

5. 用最速下降法求解

$$\min \quad (x_1-1)^2 + x_2^2,$$

初始点$x^{(1)} = (1,1)^T$.

解：

$$\nabla f(x) = \begin{pmatrix} 2(x_1-1) \\ 2x_2 \end{pmatrix}$$

$$d^{(1)} = -\nabla f(x^{(1)}) = \begin{pmatrix} 0 \\ -2 \end{pmatrix},\ x^{(1)} + \alpha d^{(1)} = \begin{pmatrix} 0 \\ 1-2\alpha \end{pmatrix},$$

求解

$$\min f(x^{(1)} + \alpha d^{(1)}) = 1 + (1-2\alpha)^2, \text{得}\alpha = \frac{1}{2}.$$

则

$$x^{(2)} = x^{(1)} + \alpha d^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$d^{(2)} = -\nabla f(x^{(2)}) = \begin{pmatrix} 2 \\ 0 \end{pmatrix},\ x^{(2)} + \alpha d^{(2)} = \begin{pmatrix} 2\alpha \\ 0 \end{pmatrix},$$

求解

$$\min f(x^{(2)} + \alpha d^{(2)}) = (2\alpha-1)^2, \text{得}\alpha = \frac{1}{2}.$$

则

$$x^{(3)} = x^{(2)} + \alpha d^{(2)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

且

$$\nabla f(x^{(3)}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix},\ \|\nabla f(x^{(3)})\| = 0.$$

所以$x^{(3)}$为极小值点。

6. 用牛顿法求解

$$\min \quad (x_1-1)^4 + x_2^2,$$

初始点$x^{(1)} = (0,1)^T$, 迭代两次。

解：

$$\nabla f(x) = \begin{pmatrix} 4(x_1-1)^3 \\ 2x_2 \end{pmatrix},\ \nabla^2 f(x) = \begin{pmatrix} 12(x_1-1)^2 & 0 \\ 0 & 2 \end{pmatrix}$$

第一次迭代：

$$\nabla f(x^{(1)}) = \begin{pmatrix} -4 \\ 2 \end{pmatrix}, \ \nabla^2 f(x^{(1)}) = \begin{pmatrix} 12 & 0 \\ 0 & 2 \end{pmatrix}$$

$$x^{(2)} = x^{(1)} - \nabla^2 f(x^{(1)})^{-1} \nabla f(x^{(1)}) = \begin{pmatrix} \frac{1}{3} \\ 0 \end{pmatrix}.$$

第二次迭代：

$$\nabla f(x^{(2)}) = \begin{pmatrix} -\frac{32}{27} \\ 0 \end{pmatrix}, \ \nabla^2 f(x^{(2)}) = \begin{pmatrix} \frac{48}{9} & 0 \\ 0 & 2 \end{pmatrix}$$

$$x^{(3)} = x^{(2)} - \nabla^2 f(x^{(2)})^{-1} \nabla f(x^{(2)}) = \begin{pmatrix} \frac{5}{9} \\ 0 \end{pmatrix}.$$

7. 用FR共轭梯度法求解

$$\min \quad f(x) = 2(x_1 - x_2^2)^2 + (1 + x_2)^2,$$

初始点仍为$(0,0)^T$。

解：

$$\nabla f(x) = \begin{pmatrix} 4(x_1 - x_2^2) \\ -8(x_1 - x_2^2)x_2 + 2(1 + x_2) \end{pmatrix},$$

$$d^{(1)} = -\nabla f(x^{(1)}) = \begin{pmatrix} 0 \\ -2 \end{pmatrix}, \quad x^{(1)} + \alpha d^{(1)} = \begin{pmatrix} 0 \\ -2\alpha \end{pmatrix}$$

求解

$$\min f(x^{(1)} + \alpha d^{(1)}) = 32\alpha^4 + (1 - 2\alpha)^2, \text{得} \alpha = \frac{1}{4}.$$

则

$$x^{(2)} = x^{(1)} + \alpha d^{(1)} = \begin{pmatrix} 0 \\ -\frac{1}{2} \end{pmatrix}, \ \nabla f(x^{(2)}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

$$\beta = \frac{1}{4}, \quad d^{(2)} = -\nabla f(x^{(2)}) + \beta d^{(1)} = \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix}$$

$$x^{(2)} + \alpha d^{(2)} = \begin{pmatrix} \alpha \\ -\frac{1}{2} - \frac{1}{2}\alpha \end{pmatrix},$$

求解

$$\min f(x^{(2)} + \alpha d^{(2)}) = 2[\alpha - (-\frac{1}{2} - \frac{1}{2}\alpha)^2]^2 + (1 - \frac{1}{2} - \frac{1}{2}\alpha)^2, \text{得} \alpha = 1.$$

则

$$x^{(3)} = x^{(2)} + \alpha d^{(2)} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \nabla f(x^{(3)}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

所以极小值点为$x^{(3)}$.

8. 用外点罚函数法和倒数障碍罚函数法求解如下问题

$$\begin{aligned} \min \quad & \tfrac{1}{12}(x_1+1)^2 + x_2 \\ s.t. \quad & x_1 + 1 \geq 0 \\ & x_2 \geq 0 \end{aligned}$$

解： 外点罚函数法：

$$\begin{aligned} F(x,\sigma) \quad &= \tfrac{1}{12}(x_1+1)^2 + x_2 + \sigma[\max\{-1-x_1, 0\}]^2 + \sigma[\max\{-x_2, 0\}]^2 \\ &= \begin{cases} \tfrac{1}{12}(x_1+1)^2 + x_2 + \sigma(1+x_1)^2 + \sigma x_2^2 & x_1 \leq -1, x_2 \leq 0 \\ \tfrac{1}{12}(x_1+1)^2 + x_2 + \sigma(1+x_1)^2 & x_1 \leq -1, x_2 > 0 \\ \tfrac{1}{12}(x_1+1)^2 + x_2 + \sigma x_2^2 & x_1 > -1, x_2 \leq 0 \\ \tfrac{1}{12}(x_1+1)^2 + x_2 & x_1 > -1, x_2 > 0 \end{cases} \end{aligned}$$

则

$$\begin{cases} \frac{\partial F}{\partial x_1} = \frac{1}{6}(x_1+1) + 2\sigma(1+x_1) = 0 \\ \frac{\partial F}{\partial x_2} = 1 + 2\sigma x_2 = 0 \end{cases}$$

得

$$\begin{cases} x_1 = \frac{-\frac{1}{6}-2\sigma}{\frac{1}{6}+2\sigma} \\ x_2 = -\frac{1}{2\sigma} \end{cases}$$

令$\sigma \to \infty$, 得

$$\begin{cases} x_1^* = -1 \\ x_2^* = 0 \end{cases}$$

倒数障碍罚函数法：

$$F(x,\mu) = \frac{1}{12}(x_1+1)^2 + x_2 + \mu(\frac{1}{x_1+1} + \frac{1}{x_2}),$$

令

$$\begin{cases} \frac{\partial F}{\partial x_1} = \frac{1}{6}(x_1+1) - 2\frac{\mu}{(1+x_1)^2} = 0 \\ \frac{\partial F}{\partial x_2} = 1 - \frac{\mu}{x_2^2} = 0 \end{cases}$$

得

$$\begin{cases} x_1 = \sqrt[3]{12\mu} - 1 \\ x_2 = \sqrt{\mu} \end{cases}$$

令$\mu \to 0$, 有

$$\begin{cases} x_1^* = -1 \\ x_2^* = 0 \end{cases}$$

9. 用广义乘子法求解

$$\begin{aligned} \min \quad & x_1^2 + 2x_2^2 \\ s.t. \quad & x_1 + x_2 \geq 1 \end{aligned}$$

解： 定义广义乘子函数

$$\begin{aligned} \phi(x,v,\sigma) &= x_1^2 + 2x_2^2 + \tfrac{1}{2\sigma}\{[\max(0, v-\sigma(x_1+x_2-1))]^2 - v^2\} \\ &= \begin{cases} x_1^2 + 2x_2^2 + \frac{1}{2\sigma}\{[v-\sigma(x_1+x_2-1)]^2 - v^2\} & x_1+x_2-1 \leq \frac{v}{\sigma} \\ x_1^2 + 2x_2^2 - \frac{v^2}{2\sigma} & x_1+x_2-1 > \frac{v}{\sigma} \end{cases} \end{aligned}$$

令

$$\begin{cases} \frac{\partial \phi}{\partial x_1} = 2x_1 - [v - \sigma(x_1+x_2-1)] = 0 \\ \frac{\partial \phi}{\partial x_2} = 4x_2 - [v - \sigma(x_1+x_2-1)] = 0 \end{cases}$$

得

$$\begin{cases} x_1^{(k)} = \frac{2(v^{(k)}+\sigma)}{4+3\sigma} \\ x_2^{(k)} = \frac{v^{(k)}+\sigma}{4+3\sigma} \end{cases}$$

$$v^{(k+1)} = \max(0, v^{(k)} - \sigma(x_1^{(k)} + x_2^{(k)} - 1)) = \frac{4v^{(k)} + 4\sigma}{4+3\sigma}$$

当$v^{(k)} < \frac{4}{3}$时，可知$\{v^k\}$为单调递增有上界的数列，必有极限。则当$k \to \infty$,有

$$v^{(k)} \to \frac{4}{3}, \quad \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \end{pmatrix} \to \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}$$

10. 什么是 P 问题和 NP 问题?

解：对给定的问题，若存在多项式时间算法，则该问题称为多项式时间可解问题，或简称为多项式问题。 所有多项式问题的集合记为 P。 若存在一个多项式函数 $g(x)$ 和一个验证算法 A, 对一类判定问题的任何一个“是”的判定实例 I 都存在一个字符串 S 是 I 的“是”回答, 其规模满足 $l(S) = O(g(l(I)))$, 且算法 A 验证 S 为实例 I 的“是”答案的计算量为 $O(g(l(I)))$, 则称这个判定问题是非确定多项式的, 简记为 NP。

11. 受热金属物体分子运动有什么规律，对求解组合优化问题有什么启发?

解：受热金属物体分子将依概率呈一定的规律性运动： 温度 t 很高时, 金属物体的分子停留在任何状态的概率近似相等; 在同一温度 t 下, 金属物体的分子停留在能量低的状态比在能量高的状态的概率大； 分子在能量最低状态的概率关于温度 t 下降； 分子停留在最低能量状态的概率随温度降低趋于 1； 分子在非能量最低状态的概率随温度降低趋于 0。 将优化问题的可行解对应于分子的状态, 优化问题中的目标函数 $f(x)$ 对应于分子的状态能量， 那么通过构造模拟的降温过程， 可将金属物体分子状态概率分布规律应用于求解优化问题。

12. 什么是 Metropolis 接受准则?

解：Metropolis 接受准则：在状态 s_i时, 产生的状态 s_j 被接受的概率为

$$A_{ij}(t) = \begin{cases} 1, & \text{如果 } f(s_i) \geq f(s_j), \\ \exp\left(-\frac{\Delta f_{ij}}{t}\right), & \text{如果 } f(s_i) < f(s_j). \end{cases}$$

13. 遗传算法中交叉运算如何进行，变异运算如何进行?

解：种群中的染色体以一定的概率 P_c （常称为交叉概率） 被选来作交叉运算产生新的染色体， 被选中作交叉运算的染色体称为父代， 可见大约有 $P_c \cdot pop\_size$ 个被选到的父代， 将被选到的父代两两配对。 即从种群中第 $i = 1$ 个染色体到第 pop_size 个染色体逐个地如下操作: 从 $[0, 1]$ 中随机产生 r, 如 $r < P_c$， 则染色体 V_i 被选为父代。 记父代为 $V_1', V_2', V_3', \cdots$， 将他们配成对 $(V_1', V_2'), (V_3', V_4'), (V_5', V_6'), \cdots$。 两个父代通过交叉运算产生的新染色体称为其后代， 在原种群中用后代替换父代形成新的种群或群体。 种群中的染色体以一定的概率 P_m （常称为变异概率） 被选来作变异运算产生新的染色体，可见大约有 $P_m \cdot pop\_size$ 个被选到的父代。 即从种群中第 $i = 1$ 个染色体到第 pop_size 个染色体逐个地如下操作: 从 $[0, 1]$ 中随机产生 r, 如 $r < P_m$， 则染色体 V_i 被选为父代。 将被选到的父代通过变异运算产生其后代，在原种群中用后代替换父代形成新的种群或群体。

14. 轮盘赌如何实施?

解：轮盘赌的选择过程:

1: 计算所有染色体 V_i 的累积概率 q_i,

$$q_0 = 0, \quad q_i = \sum_{j=1}^{i} Eval(V_j), \quad i = 1, 2, \cdots, pop\_size.$$

2: 在 $(0, q_{pop\_size}]$ 中产生一个随机数 r.

3: 若 $q_{i-1} < r \le q_i$, 则选择染色体 V_i.

4: 重复第 2 至第 3 步 pop_size 次以获得 pop_size 个染色体.

第 六 章　习题解答与提示

1. 分别用Lagrange插值公式表示经过点 $(4,10),(5,5.25),(6,1)$的二次代数多项式。

解：由Lagrange插值公式，得

$$
\begin{aligned}
L_3(x) &= \frac{(x-5)(x-6)}{(4-5)(4-6)}\times 10+\frac{(x-4)(x-6)}{(5-4)(5-6)}\times 5.25+\frac{(x-4)(x-5)}{(6-4)(6-5)}\times 1\\
&= \frac{1}{4}(x^2-28x+136)
\end{aligned}
$$

2. 已知 $f(x)=3^x$ 数据表

x_i	-1	0	1	2
$f(x_i)$	$\frac{1}{3}$	1	3	9

分别用Newton二次、三次插值多项式求$f(\frac{1}{3})$的近似值。

解：计算插商表

x_i	f_i	一阶差商	二阶差商	三阶差商
-1	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{4}{9}$
0	1	2	2	
1	3	6		
2	9			

当取二次插值时，有

$$N_2(x)=\frac{1}{3}+\frac{2}{3}(x+1)+\frac{2}{3}x(x+1)=\frac{2}{3}x^2+\frac{4}{3}x+1,$$

所以$f(\frac{1}{3})=\frac{41}{27}\approx 1.5185185.$

而当取三次插值时，有

$$N_3(x)=\frac{1}{3}+\frac{2}{3}(x+1)+\frac{2}{3}x(x+1)+\frac{4}{9}x(x+1)(x-1)=\frac{4}{9}x^3+\frac{2}{3}x^2+\frac{8}{9}x+1.$$

则$f(x)=1.386831.$

3. 已知 $f(x)=\ln x$ 数据表

x_i	0.4	0.5	0.6	0.7	0.8
$\ln x$	-0.916291	-0.693147	-0.510826	-0.356675	-0.223144

作出差分表，用线性插值和二次插值，计算$\ln 0.54$的近似值。

解：采用线性插值计算时，$\because 0.5 < 0.54 < 0.6, \therefore$ 采用节点 $0.5, 0.6$进行插值。

$$
\begin{aligned}
L_1(x) &= f(x_1)l_1(x) + f(x_2)l_2(x) \\
&= (-0.693147)\frac{x-0.6}{0.5-0.6} + (-0.510826)\frac{x-0.5}{0.6-0.5} \approx -0.6202
\end{aligned}
$$

当采用二次插值计算时取插值节点为$0.4, 0.5, 0.6$,得到

$$
\begin{aligned}
L_2(x) &= f(x_0)l_0(x) + f(x_1)l_1(x) + f(x_2)l_2(x) \\
&= (-0.916291)\frac{(x-0.5)(x-0.6)}{(0.4-0.5)(0.4-0.6)} + (-0.693147)\frac{(x-0.4((x-0.5)}{(0.5-0.4)(0.5-0.6)} \\
&+ (-0.510826)\frac{(x-0.4)(x-0.5)}{(0.6-0.4)(0.6-0.5)} \\
&\approx -0.6153
\end{aligned}
$$

4. 求一个次数不高于四次的多项式$P(x)$,使其满足

$$P(0) = P'(0) = 0, \qquad P(1) = P'(1) = 1, \qquad P(2) = 1.$$

并估计其误差。

解：根据插值条件为$P(0) = P'(0) = 0, \qquad P(1) = P'(1) = 1$，先计算次数不高于4次的Hermite插值多项式

$$H_3(x) = \sum_{i=0}^{1} h_i(x)f(x_i) + \sum_{i=0}^{1} \bar{h}_i(x)f'(x_i)$$

其中

$$
\begin{aligned}
h_0(x) = (1 - 2\frac{x-x_0}{x_0-x_1})\left(\frac{x-x_1}{x_0-x_1}\right)^2 &= (1+2x)(x-1)^2 \\
h_1(x) = (1 - 2\frac{x-x_1}{x_1-x_0})\left(\frac{x-x_0}{x_1-x_0}\right)^2 &= (3-2x)x^2 \\
\bar{h}_0(x) = x(x-1)^2 \qquad \bar{h}_1(x) &= (x-1)x^2
\end{aligned}
$$

所以

$$H_3(x) = (3-2x)x^2 + (x-1)x^2 = -x^3 + 2x^2.$$

设$P(x) = H_3(x) + A(x-x_0)^2(x-x_1)^2$,其中$A$为待定常数。将$P(2) = 1$代入$P(x) = -x^3 - 2x^2 + x^2(x-1)^2$中，得到$A = \frac{1}{4}$,从而

$$P(x) = \frac{1}{4}x^2(x-3)^2.$$

5. 证明:

$$(1)\sum_{i=0}^{n-1}\Delta^2 y_i = \Delta y_n - \Delta y_0; \qquad (2)\Delta(f_i g_i) = f_i\Delta f_i + g_{i+1}\Delta f_i.$$

证明：(1).

$$\begin{aligned}\sum_{i=0}^{n-1}\Delta^2 y_i &= \Delta^2 y_0 + \Delta^2 y_1 + ... + \Delta^2 y_{n-1}\\ &= \Delta y_1 - y_0 + \Delta y_2 - y_1 + ... + \Delta y_n - y_{n-1}\\ &= \Delta y_n - \Delta y_0.\end{aligned}$$

(2).

$$\begin{aligned}\Delta(f_i g_i) &= f_{i+1}g_{i+1} - f_i g_i\\ &= g_{i+1}(f_{i+1} - f_i) + f_i(g_{i=1}g_i)\\ &= f_i\Delta f_i + g_{i+1}\Delta f_i.\end{aligned}$$

6. 已知实验数据

x_i	0	1	2	3	4
$f(x_i)$	2	2.05	3	9.6	34

已知其经验公式为$y = a + bx^2$,用最小二乘方法确定参数a, b.

解：根据最小二乘拟合思想，满足正规方程组

$$\begin{cases} 5a + b\sum\limits_{i=1}^{5} x_i^2 = \sum\limits_{i=1}^{5} f(x_i),\\ a\sum\limits_{i=1}^{5} x_i^2 + b\sum\limits_{i=1}^{5} x_i^4 = \sum\limits_{i=1}^{5} x_i^2 f(x_i).\end{cases}$$

代入数据表并计算，得到

$$\begin{cases} a + 30b = 50.65,\\ 30a + 354b = 644.45.\end{cases}$$

从而解得

$$a = -1.6132, b = 1.9572.$$

即满足观测数据表的线性拟合多项式为

$$y = -1.6132 + 1.9572x^2.$$

7. 若记母函数$\psi(t)$的傅里叶变换为$\hat{\psi}(\omega)$，计算小波函数$\psi_{a,b}(t)$的傅里叶变换。

证明：由$\psi_{a,b}(t)=|a|^{-\frac{1}{2}}\psi\left(\frac{t-b}{a}\right)$，以及傅里叶变换的定义，有

$$\begin{aligned}\hat{\psi}_{a,b}(\omega)&=\int_{-\infty}^{\infty}\psi(t)e^{-i\omega t}dt=\int_{-\infty}^{\infty}|a|^{-\frac{1}{2}}\psi\left(\frac{t-b}{a}\right)e^{-i\omega t}dt\\&=|a|^{-\frac{1}{2}}\int_{-\infty}^{\infty}\psi(u)e^{-i\omega(au+b)}adu=|a|^{\frac{1}{2}}e^{-i\omega b}\int_{-\infty}^{\infty}\psi(u)e^{-i(a\omega)u}du\\&=|a|^{\frac{1}{2}}e^{-i\omega b}\hat{\psi}(a\omega)\end{aligned}$$

第 七 章　习题解答与提示

1. 设有一根长为 的均匀柔软细弦，当它作微小横振动时,除受内部张力外，还受到周围介质所产生的阻尼力的作用，阻尼力与速度的平方成正比(比例系数为b)，试写出带有阻尼力的弦振动方程。

   解：设均匀弦的密度为ρ，张力为T（ρ和T均为常数），则阻尼外力的密度函数为$-b(\frac{\partial u}{\partial t})^2$，根据带强迫外力的弦振动方程可得本题方程为

   $$\frac{\partial^2 u}{\partial t^2} - a^2\frac{\partial^2 u}{\partial x^2} = f(x,t)$$

   其中$a^2 = \frac{T}{\rho}$, $f(x,t) = -\frac{b}{\rho}\left(\frac{\partial u}{\partial t}\right)^2$

2. 选已知函数 $w(x,t)$，并作函数代换 $v = u - w$ 将以下边界条件齐次化

   (1) $u(0,t) = t,u(2,t) = \sin t;$

   (2) $u(0,t) = 1,\ u_x(l,t) = 1 + t^2;$

   (3) $u_x(0,t) = \varphi(t),u_x(3,t) = \psi(t);$

   (4) $u_x(0,t) = t^2,u_x(2,t) + u(2,t) = t.$

   解：(1) 由$v(0,t) = 0,\ v(2,t) = 0$，可得 $w(0,t) = t, w(2,t) = \sin t$，从而可取$w$为关于$x$的线性函数，故$w(x,t) = t + \frac{x}{2}(\sin t - t)$；

   (2) 由$v(0,t) = 0,\ v_x(l,t) = 0$，可得$w(0,t) = 1, w_x(l,t) = 1 + t^2$，取$w(x,t) = a(t)x + b(t)$，将边界条件带入后可得$b(t) = 1, a(t) = 1 + t^2$，从而得到$w(x,t) = (1 + t^2)x + 1$

   (3) 由$v_x(0,t) = 0,\ v_x(3,t) = 0$，可得$w_x(0,t) = \varphi(t), w_x(3,t) = \psi(t)$，可取$w$为关于$x$的线性函数，故$w_x(x,t) = \varphi(t) + \frac{x}{3}(\psi(t) - \varphi(t))$，从而$w(x,t) = \varphi(t)x + \frac{x^2}{6}(\psi(t) - \varphi(t))$

   (4) 由$w_x(0,t) = t^2, w_x(2,t) + w(2,t) = t$，取$w(x,t) = a(t)x + b(t)$，将边界条件带入后可得$a(t) = t^2, b(t) = t - 3t^2$，从而得到$w(x,t) = t^2x + t - 3t^2$。

3. 考虑如下有界弦振动方程定解问题

   $$\begin{cases} u_{tt} = a^2u_{xx}, 0 < x < l, t>0 \\ u(0,t) = 0, u(l,t) = 0, t \geq 0 \\ u(x,0) = 0, u_t(x,0) = \psi(x), 0 \leq x \leq l. \end{cases}$$

(1)将 $\psi(x)$ 在 $[0,l]$ 按正交函数系 $\{\sin\frac{n\pi}{l}x\}_{n\geq 1}$ 展成Fourier 级数,并求Fouier系数；

(2) 对于任意的整数$n\geq 1$，验证$u_n(x,t)=\frac{l}{n\pi a}\psi_n\sin\frac{n\pi a}{l}t\sin\frac{n\pi}{l}x$是如下问题的解

$$\begin{cases} u_{tt}=a^2u_{xx}, 0<x<l, t>0\\ u(0,t)=0, u(l,t)=0, t\geq 0\\ u(x,0)=0, u_t(x,0)=\psi_n\sin\frac{n\pi}{l}x, 0\leq x\leq l. \end{cases}$$

(3) 利用(2)中的结果试写出问题(1)中的定解问题的解。

解：(1) 设$\psi(x)=\sum\limits_{n=0}^{\infty}\psi_n\sin\frac{n\pi}{l}x$，则系数

$$\psi_n=\frac{\int_0^l\psi(x)\sin\frac{n\pi}{l}xdx}{\int_0^l\left(\sin\frac{n\pi}{l}x\right)^2dx}=\frac{2}{l}\int_0^l\psi(x)\sin\frac{n\pi}{l}xdx$$

(2) 直接代入进行验证即可；

(3) 根据线性叠加原理可得解为

$$u(x,t)=\sum_{n=1}^{\infty}u_n(x,t)=\sum_{n=1}^{\infty}\frac{l}{n\pi a}\psi_n\sin\left(\frac{n\pi a}{l}t\right)\sin\left(\frac{n\pi}{l}x\right)$$

4. 用古典显格式在$h=0.2, r=\dfrac{1}{6}$时, 计算

$$\begin{cases} \dfrac{\partial u}{\partial t}=\dfrac{\partial^2 u}{\partial t^2}, & 0<x<1,\quad t>0\\ u(x,0)=4x(1-x), & 0\leq x\leq 1\\ u(0,t)=u(1,t)=0, & t\geq 0 \end{cases}$$

前两层的差分解。

解：显格式为:

$$\begin{cases} u_i^{k+1}=\frac{1}{6}(u_{i+1}^k-u_{i-1}^k)+\frac{2}{3}u_i^k\\ u_i^0=4x_i(1-x_i)\\ u_0^k=u_5^k=0 \end{cases}$$

利用边界条件$u_0^k=u_5^k=0$, 记$\mathbf{u}^k=[u_1^k,u_2^k,u_3^k,u_4^k]^T$, 则显格式的矩阵形式为:

$$\mathbf{u}^{k+1}=\mathbf{A}\mathbf{u}^k=\begin{bmatrix} \frac{2}{3} & \frac{1}{6} & & \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & \\ & \frac{1}{6} & \frac{2}{3} & \frac{1}{6}\\ & & \frac{1}{6} & \frac{2}{3} \end{bmatrix}\mathbf{u}^k$$

其中 $\mathbf{u}^0=[u_1^0,u_2^0,u_3^0,u_4^0]^T=[0.64,0.96,0.96,0.64]^T$, 从而得到

$$\mathbf{u}^1=[0.5867,0.8533,0.7467,0.7467]^T,\ \mathbf{u}^2=[0.5333,0.7911,0.7644,0.6222]^T$$

5. 用显格式解 取$r=1,h=0.2$, 求$k=1,2$层上的差分解.

$$\begin{cases}\dfrac{\partial^2 u}{\partial t^2}-\dfrac{\partial^2 u}{\partial x^2}=0, & 0<x<1,\quad t>0\\ u(x,0)=\sin\pi x, & \dfrac{\partial u(x,0)}{\partial t}=x(1-x),\quad 0\le x\le 1\\ u(0,t)=u(1,t)=0,\quad t\ge 0\end{cases}$$

解：显格式为:

$$\begin{cases}u_i^{k+1}=u_{i+1}^k+u_{i-1}^k-u_i^{k-1}\\ u_i^0=\sin\pi x_i,\ \frac{u_i^1-u_i^{-1}}{2\tau}=x_i(1-x_i)\\ u_0^k=u_5^k=0\end{cases}$$

由$r=(\tau/h)^2,h=0.2$可得$\tau=0.2$，从而得到

$$u_i^1=\frac{1}{2}(u_{i+1}^0+u_{i-1}^0)+0.2x_i(1-x_i),\ u_i^2=u_{i+1}^1+u_{i-1}^1-u_i^0$$

因此有

$$\begin{aligned}\mathbf{u}^0&=[u_0^0,u_1^0,u_2^0,u_3^0,u_4^0,u_5^0]^T=[0,0.5878,0.9511,0.9511,0.5878,0]^T\\ \mathbf{u}^1&=[u_0^1,u_1^1,u_2^1,u_3^1,u_4^1,u_5^1]^T=[0,0.5076,0.8175,0.8175,0.5076,0]^T\\ \mathbf{u}^2&=[u_0^1,u_1^1,u_2^1,u_3^1,u_4^1,u_5^1]^T=[0,0.2297,0.3740,0.3740,0.2297,0]^T\end{aligned}$$

6. 用有限元法解下列边值问题

$$\begin{cases}-\Delta u+12u=18xy+2,\quad (x,y)\in\Omega\\ \dfrac{\partial u}{\partial n}+3u\Big|_{\partial\Omega}=0\end{cases}$$

其中Ω是图7.1表示的正方形, $\partial\Omega$是其边界, 设Ω剖分成两个三角形.

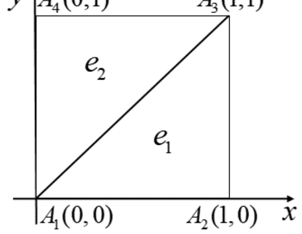

图 7.1: 三角剖分示意图

解：（1）写出微分方程对应的变分形式： $\min_u J(u) = \frac{1}{2}a(u,u) - (f,u)$, 其中

$$a(u,v) = \iint\limits_{\Omega} \left[\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} \right) + 12uv \right] dxdy + \int\limits_{\partial\Omega} 3uvds,$$

$$(f,u) = \iint\limits_{\Omega} (18xy+2)vdxdy.$$

（2）可选用三角形剖分对求解区域进行分割，确定节点编号及单元编号，求出节点坐标（略）；

(3) 计算刚度矩阵及单元载荷向量，结合边界条件形成有限元方程组（略）

(4) 计算求解节点处的近似值（略）

第 八 章　习题解答与提示

1. 为了调查某广告对销售收入的影响，某商店记录了7个月的销售收入y(万元)与广告费用x(万元)，数据如下表，

x	1	2	3	4	5	6	7
y	10	10	16	20	23	40	43

求参数β_0, β_1的最小二乘估计。

解：因为 $n=7$，$\bar{x}=4$，$\bar{y}=23.143$，$\sum\limits_{i=1}^{n} x_i^2=140$

$$l_{xx}=\sum_{i=1}^{n} x_i^2-n\left(\bar{x}\right)^2=28,$$

$$l_{xy}=\sum_{i=1}^{n}(x_i-\bar{x})y_i=166,$$

参数β_0, β_1的最小二乘估计

$$\hat{\beta}_1=l_{xy}/l_{xx}=5.929,$$

$$\hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}=-0.571$$

2. 10组观测数据由下表给出:

x	0.5	0.6	0.7	0.8	0.9	1	1.1	1.2	1.3	1.4
y	1.1	1.2	1.3	1.5	1.6	1.7	1.9	2	2.1	2.2

应用正态线性回归模型$y=\beta_0+\beta_1x+\varepsilon$， 其中误差项$\varepsilon$服从正态分布$N(0,\sigma^2)$，

(1) 求回归方程;

(2) σ^2的无偏估计值$\hat{\sigma}^2$;

(3) 求β_1的置信水平为95%的置信区间;

(4) 在显著性水平$\alpha=0.05$下相关系数的显著性检验;

(5) 观察值y的置信水平为95%的预测区间。

解：因为 $n=10$，$\bar{x}=0.95$，$\bar{y}=1.66$，$\sum\limits_{i=1}^{n} x_i^2=9.85$

$$l_{xx}=\sum_{i=1}^{n} x_i^2-n\left(\bar{x}\right)^2=0.825,$$

$$l_{xy} = \sum_{i=1}^{n} (x_i - \bar{x}) y_i = 1.05,$$

$$\hat{\beta}_1 = l_{xy}/l_{xx} = 1.273,$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 0.451$$

于是， 所求的回归方程为

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = 0.451 + 1.273x$$

σ^2的无偏估计值为

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{n} (y_i - \hat{\beta}_0 + \hat{\beta}_1 x_i)^2 = 0.001$$

由于

$$r = \frac{l_{xy}}{\sqrt{l_{xx} l_{yy}}} = 0.997 > r_\alpha(n-2) = r_{0.05}(8) = 0.632$$

因此， 在显著性水平为0.05条件下，认为一元线性回归效果显著。

参数β_1的置信水平为95%的置信区间为

$$\left(\hat{\beta}_1 - t_{\alpha/2}\,(n-2)\,\frac{\hat{\sigma}}{\sqrt{l_{xx}}}\;,\; \hat{\beta}_1 + t_{\alpha/2}\,(n-2)\,\frac{\hat{\sigma}}{\sqrt{l_{xx}}} \right) = (1.194,\ 1.351)$$

给定$x_0 = x_i$，$i = 1, 2, \cdots, 10$， 根据

$$\left(\hat{y}_0 \pm t_{\alpha/2}(n-2) \hat{\sigma} \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{l_{xx}}} \right)$$

对应的y_0的95%的预测区间为

序号	y_0	y_0的下限	y_0的上限	序号	y_0	y_0的下限	y_0的上限
1	1.1	1.005	1.17	6	1.7	1.649	1.798
2	1.2	1.135	1.294	7	1.9	1.775	1.927
3	1.3	1.265	1.419	8	2	1.901	2.055
4	1.5	1.393	1.545	9	2.1	2.026	2.185
5	1.6	1.522	1.671	10	2.2	2.15	2.315

3. 对于一元线性回归的样本模型

$$\begin{cases} y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (i = 1, 2, \cdots, n) \\ \varepsilon_i \text{相互独立},\ E(\varepsilon_i) = 0, \mathrm{var}(\varepsilon_i) = \sigma^2 \end{cases}$$

请给出参数β_0, β_1的最小二乘估计值$\hat{\beta}_0, \hat{\beta}_1$ 的推导。

解：略

4. 某课题研究四种衣料内棉花吸附十硼氢量。 每种衣料各做五次测量，所得数据如下表，试检验各种衣料棉花吸附十硼氢量有没有差异。 ($\alpha = 0.05$)

衣料1	衣料2	衣料3	衣料4
2.33	2.48	3.06	4.00
2.00	2.34	3.06	5.13
2.93	2.68	3.00	4.61
2.73	2.34	2.66	2.80
2.33	2.22	3.06	3.60

解：方差分析表

来源	平方和	自由度	均方	F比
因素	8.434	3	2.811	11.164
误差	4.029	16	0.252	
总和	12.463	19		

因为$F_{0.05}(3, 16) = 3.24$，故在水平0.05的条件下拒绝H_0，认为各种衣料棉花吸附十硼氢量有差异。

5. 设有三种机器A，B，C制造一种产品，对每种机器各观测5天，其日产量如下表所示，问机器与机器之间是否存在差异?($\alpha = 0.05$)

天数机器	1	2	3	4	5
A	41	48	41	49	57
B	65	57	54	72	64
C	45	51	56	48	48

解：方差分析表

来源	平方和	自由度	均方	F比
因素	667.733	2	333.867	8.959
误差	447.2	12	37.267	
总和	1114.933	14		

因为$F_{0.05}(2, 12) = 3.89$，故在水平0.05的条件下拒绝H_0，认为机器与机器之间存在差异。

6. 以淀粉为原料生产葡萄糖过程中，残留的许多糖蜜可用于酱色生产。 生产酱色之前应尽可能彻底除杂，以保证酱色质量。 今选用4 种除杂方法，每种方法做4 次试验，试验结果如下表，试分析不同除杂方法的除杂效果有无差异? $(\alpha = 0.05)$

除杂方法	除杂质量			
方法一	25.6	24.7	21.6	25.3
方法二	26.4	28.9	19.7	24.6
方法三	22.1	23.1	24.2	30
方法四	25.1	24.6	22.6	25.7

解：方差分析表

来源	平方和	自由度	均方	F比
因素	41.723	3	13.908	2.889
误差	57.775	12	4.815	
总和	99.498	15		

因为$F_{0.05}(3,12) = 3.49$，故在水平0.05的条件下拒绝H_0，认为同除杂方法的除杂效果无差异。

7. 高新技术公司的工程师为了确认不同厂家的塑胶材料对成型品拉拔力的影响，分别对该公司正在使用的5家供应商提供的塑胶料成型品进行拉拔试验，取得数据如下：

供应商A	供应商B	供应商C	供应商D	供应商E
5.86	4.95	5.60	5.88	4.97
5.95	5.72	6.21	6.42	4.86
5.74	5.36	5.81	6.51	5.03
5.62	5.89	5.30	5.98	6.05
5.08	4.99	4.97	5.74	4.10

试分析不同供应商所提供的塑胶材料是否有显著不同？$(\alpha = 0.05)$

解：方差分析表

因为$F_{0.05}(4,20) = 2.87$，故在水平0.05的条件下拒绝H_0，认为不同供应商所提供的塑胶材料有显著不同。

来源	平方和	自由度	均方	F比
因素	3.241	4	0.81	3.619
误差	4.478	20	0.224	
总和	7.72	24		

8. 设$\boldsymbol{X} = (X_1, X_2, \cdots X_n)'$是来自二项分布总体$B(m,p)$的$i.i.d$样本，其中$m$已知，$0 < p < 1$未知。

1) 验证$\pi(p) \propto p^{-1}(1-p)^{-1}$是广义先验分布;

2) 在上述先验分布下，求p的后验分布。 并找出p的充分统计量;

3) 若p的先验分布取为$\beta(a,b)$, 验证$\beta(a,b)$是否为二项分布总体$B(m,p)$共轭先验分布。

解：（1）由

$$\pi(p)p(x|p) \propto p^{-1}(1-p)^{-1}\prod_{i=1}^{n} C_m^{x_i} p^{x_i}(1-p)^{m-x_i} \propto p^{\sum\limits_{i=1}^{n} x_i - 1}(1-p)^{nm-\sum\limits_{i=1}^{n} x_i - 1}$$

可知

$$\int_0^1 p^{x-1}(1-p)^{m-x-1}dp = \frac{\Gamma(\sum\limits_{i=1}^{n} x_i)\Gamma(nm-\sum\limits_{i=1}^{n} x_i)}{\Gamma(nm)} < +\infty$$

所以是广义先验分布;

（2）由（1）可知，

$$h(\theta|x) \propto p^{\sum\limits_{i=1}^{n} x_i - 1}(1-p)^{nm-\sum\limits_{i=1}^{n} x_i - 1}$$

表明后验分布是参数为的贝塔分布，由此可知$\sum\limits_{i=1}^{n} X_i$即为充分统计量。

（3）由（2）可知

$$h(\theta|x) \propto p^{\sum\limits_{i=1}^{n} x_i - 1}(1-p)^{nm-\sum\limits_{i=1}^{n} x_i - 1}$$

即后验分布是参数为$\sum\limits_{i=1}^{n} x_i, nm-\sum\limits_{i=1}^{n} x_i$的贝塔分布，所以$\beta(a,b)$是否为二项分布总体$B(m,p)$共轭先验分布.

9. 设$\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$是来自Poisson分布总体$P(\lambda)$的$i.i.d$。 样本, λ的先验分布取为Γ 分布 $\Gamma(\alpha,\mu)$, 则 Γ 分布族 $\{\Gamma(\alpha,\mu); \alpha > 0, \mu > 0\}$ 是 Poisson 分布族 $\{P(\lambda), \lambda > 0\}$ 的共轭先验分布族。

证明：利用贝叶斯公式，可得后验分布

$$h(\theta|x) \propto \pi(\lambda)p(x|\lambda) \propto \lambda^{\alpha-1}e^{-\lambda\mu}\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda} \propto \lambda^{\alpha+\sum\limits_{i=1}^{n} x_i - 1}e^{-\lambda(\mu+n)}$$

即后验分布是参数为$\alpha+\sum\limits_{i=1}^{n} x_i, \mu+n$的 Γ 分布，所以 Γ 分布族 $\{\Gamma(\alpha,\mu);\alpha>0,\mu>0\}$ 是Poisson 分布族 $\{P(\lambda),\lambda>0\}$ 的共轭先验分布族。

10. 设$\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$是来自0-1分布总体$B(1,p)$ 的$i.i.d$样本，设p的先验分布为

$$\pi(p)=\begin{cases}\frac{1}{p},\ p_0<p<1\\ 0,\ \text{其它}\end{cases}$$

其中p_0是$(0,1)$中的固定常数。求证p的条件期望估计为

$$\hat{p}=\frac{\sum\limits_{i=1}^{n} X_i}{n+2}\cdot\frac{1-I_{p_0}\left(\sum\limits_{i=1}^{n} X_i+2,n-\sum\limits_{i=1}^{n} X_i+1\right)}{1-I_{p_0}\left(\sum\limits_{i=1}^{n} X_i+1,n-\sum\limits_{i=1}^{n} X_i+1\right)}$$

其中

$$I_{p_0}(\alpha,\beta)=\frac{1}{B(\alpha,\beta)}\int_0^{p_0} p^{\alpha-1}(1-p)^{\beta-1}\mathrm{d}p$$

证明： $\pi(p)p(x\,|p)=\frac{1}{p}\prod\limits_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}=p^{\sum\limits_{i=1}^{n} x_i-1}(1-p)^{n-\sum\limits_{i=1}^{n} x_i}$

$$\begin{aligned}h(p|x)&=\frac{\pi(p)p(x\,|p)}{\int_{p_0}^1\pi(p)p(x\,|p)dp}\\&=\frac{p^{\sum\limits_{i=1}^{n} x_i-1}(1-p)^{n-\sum\limits_{i=1}^{n} x_i}}{\int_0^1\pi(p)p(x\,|p)dp-\int_0^{p_0}\pi(p)p(x\,|p)dp}\\&=\frac{p^{\sum\limits_{i=1}^{n} x_i-1}(1-p)^{n-\sum\limits_{i=1}^{n} x_i}}{B(\sum\limits_{i=1}^{n} x_i,n-\sum\limits_{i=1}^{n} x_i+1)-I_{p_0}\cdot B(\sum\limits_{i=1}^{n} x_i,n-\sum\limits_{i=1}^{n} x_i+1)}\\&=\frac{p^{\sum\limits_{i=1}^{n} x_i-1}(1-p)^{n-\sum\limits_{i=1}^{n} x_i}}{B(\sum\limits_{i=1}^{n} x_i,n-\sum\limits_{i=1}^{n} x_i+1)\,[1-I_{p_0}]}\end{aligned}$$

于是p的条件期望估计为

$$\hat{p}=E(p|x)=\int_{p_0}^{1}p\cdot\frac{p^{\sum\limits_{i=1}^{n}x_i-1}(1-p)^{n-\sum\limits_{i=1}^{n}x_i}}{B(\sum\limits_{i=1}^{n}x_i,n-\sum\limits_{i=1}^{n}x_i+1)\,[1-I_{p_0}]}dp$$

$$=\frac{\int_{p_0}^{1}p^{\sum\limits_{i=1}^{n}x_i-1}(1-p)^{n-\sum\limits_{i=1}^{n}x_i}dp}{B(\sum\limits_{i=1}^{n}x_i,n-\sum\limits_{i=1}^{n}x_i+1)\,[1-I_{p_0}]}$$

$$=\frac{\sum\limits_{i=1}^{n}X_i}{n+2}\cdot\frac{1-I_{p_0}\left(\sum\limits_{i=1}^{n}X_i+2,n-\sum\limits_{i=1}^{n}X_i+1\right)}{1-I_{p_0}\left(\sum\limits_{i=1}^{n}X_i+1,n-\sum\limits_{i=1}^{n}X_i+1\right)}$$

11. 设$\boldsymbol{X}=(X_1,X_2,\cdots X_n)'$是来自$\Gamma$分布总体$\Gamma\left(\frac{1}{2},\frac{1}{2\theta}\right)$的$i.i.d$ 样本,

(1) 求证逆Γ分布$I\Gamma\left(\alpha,\lambda\right)$分布是$\Gamma$ 分布$\Gamma\left(\frac{1}{2},\frac{1}{2\theta}\right)$的共轭先验分布;

(2) 求θ的满足杰弗莱原则的先验分布;

(3) 在(1), (2)两种先验分布下，求θ的条件期望估计与最大后验估计。

证明：(1) 注意到

$$h(\theta|\,x)\propto\pi(\theta)p(x\,|\theta)\propto\theta^{-(\alpha+1)}e^{-\frac{\lambda}{\theta}}\left(\frac{1}{2\theta}\right)^{\frac{n}{2}}e^{-\frac{1}{2\theta}\sum\limits_{i=1}^{n}x_i}\propto\theta^{-(\alpha+1)-\frac{n}{2}}\cdot e^{-\frac{1}{\theta}\left(\lambda+\frac{1}{2}\sum\limits_{i=1}^{n}x_i\right)}$$

可知逆Γ分布$I\Gamma\left(\alpha,\lambda\right)$分布是$\Gamma$分布$\Gamma\left(\frac{1}{2},\frac{1}{2\theta}\right)$的共轭先验分布。

（2） 似然函数是

$$L(\theta)=\prod_{i=1}^{n}\frac{\left(\frac{1}{2\theta}\right)^{\frac{1}{2}}}{\Gamma(\frac{1}{2})}x_i^{-\frac{1}{2}}e^{-\frac{x_i}{2\theta}}=\frac{2^{-\frac{n}{2}}\theta^{-\frac{n}{2}}}{\Gamma^n(\frac{1}{2})}\cdot\prod_{i=1}^{n}x_i^{-\frac{1}{2}}\cdot e^{-\frac{1}{2\theta}\sum\limits_{i=1}^{n}x_i}$$

$$l(\theta)=\ln L(\theta)=-\frac{n}{2}\ln 2-\frac{n}{2}\ln\theta-n\ln\Gamma(\frac{1}{2})+\sum_{i=1}^{n}\ln x_i^{-\frac{1}{2}}-\frac{1}{2\theta}\sum_{i=1}^{n}x_i$$

$$l'(\theta)=-\frac{n}{2\theta}+\frac{1}{2\theta^2}\sum_{i=1}^{n}x_i$$

$$l''(\theta)=\frac{n}{2\theta^2}+(-2)\frac{1}{2\theta^3}\sum_{i=1}^{n}x_i=\frac{1}{2\theta^2}\left[n-\frac{2}{\theta}\sum_{i=1}^{n}x_i\right]$$

于是，Jeffrey原则的先验分布的核为

$$|l''(\theta)|^{\frac{1}{2}}=\frac{1}{\sqrt{2}\theta}\left[n-\frac{2}{\theta}\sum_{i=1}^{n}x_i\right]^{\frac{1}{2}}$$