浙 江 林 学 院 学 报 2000, **17**(4): 430～435
Journal of Zhejiang Forestry College

文章编号: 1000-5692(2000)04-0430-06

条件概率矩阵和信道理论及 Markov 链

管 宇

(浙江林学院 信息工程与基础科学系, 浙江 临安 311300)

摘要: 给出离散型随机变量的概率分布律所成的单位空间的定义, 将条件概率矩阵看作单位空间之间的映射, 研究其中一些性质。并应用于信息论, 得出信道容量的几何意义, 特别给出信道容量的一种较好算法。应用于 Markov 链, 得到状态有限的转移概率矩阵必有不动点。参 3

关键词: 条件概率矩阵; 信息率; Markov 链
中图分类号: O211.6　　文献标识码: A

1 单位空间

定义　设由 n 个实数构成的向量空间 $\Omega_n = \{(\omega_1, \omega_2, \cdots, \omega_n)\}$ 满足:

$$①\ \omega_i \geqslant 0,\ i = 1, 2, \cdots, n; \quad ②\sum_{i=1}^{n} \omega_i = 1。$$

则称 Ω_n 为 n 维的单位空间。此处 n 可为自然数或无穷大。实际上 n 维的单位空间就是取 n 个值的离散型随机变量的概率分布律全体所成空间。

在 Ω_n 中定义距离

$$d(u, v) = \sum_{i=1}^{n} | u_i - v_i |; u = (u_1, u_2, \cdots, u_n) \in \Omega_n, v = (v_1, v_2, \cdots, v_n) \in \Omega_n。$$

显然 Ω_n 按距离 $d(u, v)$ 为一完备的度量空间。为讨论方便, 若 $n < m$ 时, 将 Ω_n 当作 Ω_m 的子空间。

在 Ω_n 中定义加法: 设 $u = (u_1, u_2, \cdots, u_n) \in \Omega_n, v = (v_1, v_2, \cdots, v_n) \in \Omega_n$, 非负数 $a, b, a + b = 1$, $au + bv \overset{\text{Def}}{=} (au_1 + bv_1, au_2 + bv_2, \cdots, au_n + bv_n) \in \Omega_n$。对于 Ω_n 中向量组 $p_1, p_2, \cdots, p_s, p$, 如果存在一组非负数 $k_1, k_2, \cdots, k_s, \sum_{i=1}^{s} k_i = 1$, 使得 $p = \sum_{i=1}^{s} k_i p_i$, 称向量 p 为向量组 $p_1, p_2, \cdots, p_s$ 单位线性表出或单位线性组合。若 $q_1, q_2, \cdots, q_t (\in \Omega_n)$ 中任一向量都可经向量组 $p_1, p_2, \cdots, p_s$ 单位线性表出, 则称前者可经后者单位线性表出。显然单位线性表出具有传递性, 但不具对称性。如向量组 $p_1, p_2, \cdots, p_s$ 中有一向量可经其他向量单位线性表出, 则称此向量组单位线性相关; 如果其中任一向量都不能经其他向量单位线性表出, 则称为单位线性无关组。

单位线性相关线性无关与一般向量空间的向量组的线性相关线性无关的关系如下: 单位线性相关必线性相关, 反之不然; 线性无关必单位线性无关, 反之不然。

收稿日期: 1999-12-06; 修回日期: 2000-04-24
作者简介: 管 宇(1964—), 男, 浙江台州人, 讲师, 从事概率统计理论及应用研究。

一向量组的部分组称为一个极大单位线性无关组，如果这部分向量组单位线性无关，则向量组中其他向量都可以这部分组单位线性表出。关于极大单位线性无关组的求法：先取若干单位线性无关向量，再在剩余的向量中任取一个，若能由前面几个单位线性表出，则将其去除，不然加入前面的向量组，如此重复下去即可得。

设 $G = \left|\sum_{i=1}^{s} k_i p_i \middle| k_i \geqslant 0;\ i = 1, 2, \cdots, s;\ \left\{\sum_{i=1}^{s} k_i = 1; p_1, p_2, \cdots, p_s \in \Omega_n\right\} \in \Omega_n\right.$。

G 就是由向量组 $p_1, p_2, \cdots, p_s$ 的所有单位线性组合全体所成的空间，称之为由 $p_1, p_2, \cdots, p_s$ 生成的空间。

定理 1　设向量组 $q_1, q_2, \cdots, q_t$ 经向量组 $p_1, p_2, \cdots, p_s$ 单位线性表出，则由向量组 $q_1, q_2, \cdots, q_t$ 生成的空间包含于由向量组 $p_1, p_2, \cdots, p_s$ 生成的空间。从而知任一向量组生成的空间完全等同于其极大单位线性无关组所生成的空间。

2　条件概率矩阵

定义　称 $P = (p_{ij})_{r\times s}$ 为条件概率矩阵，如果：

$$(1)\ p_{ij} \geqslant 0,\ i = 1, 2, \cdots, r;\ j = 1, 2, \cdots, s;\quad (2)\ \sum_{j=1}^{s} p_{ij} = 1,\ i = 1, 2, \cdots, r。$$

其中 r 和 s 为自然数或无穷大。

如果将条件概率矩阵的每一行看成 s 维单位空间 Ω_s 中的一个向量，那么 $r\times s$ 的条件概率矩阵可认为是由 Ω_s 中的 r 个向量组成的，而向量组的生成空间亦称为（相应的）条件概率矩阵的生成空间。显然，按普通矩阵的乘法，两个条件概率矩阵的乘积（如果可乘的话），仍是条件概率矩阵。

定理 2　向量组的极大单位线性无关组是唯一的。从而定义该极大单位线性无关组的个数为向量组的单位秩。条件概率矩阵 P 的单位秩定义为其行向量组的单位秩，记为 $\tau(P)$。

证明　设 Ω_s 向量组 $p_1, p_2, \cdots, p_r$ 有 2 个极大单位线性无关组 $q_1, q_2, \cdots, q_n$ 和 $q_1^*, q_2^*, \cdots, q_m^*$；由这 2 个极大单位线性无关组向量为行向量分别组成 $n\times s$、$m\times s$ 条件概率矩阵 Q 和 Q^*，则 Q 与 Q^* 中的行向量分别可经 Q^* 与 Q 的行向量单位线性表出，而表出的系数可分别作为 Ω_m 与 Ω_n 中的向量。即存在 $n\times m$ 与 $m\times n$ 的条件概率矩阵 M 与 N，使得 $Q = M\times Q^*$，$Q^* = N\times Q$；因此有 $Q = M\times N\times Q$。

由 Q 是行单位线性无关知，$M\times N$ 必为单位矩阵。再由 M 与 N 的元素的非负性及每行元素之和为 1 知，必 $m = n$ 且 M 与 N 的元素只有 0 和 1。故向量组的极大单位线性无关组是唯一的。

定理 3　设有 $r\times n$ 与 $n\times s$ 的条件概率矩阵 L 与 Q，$L\times Q = P$，则：

①P 的行向量可经 Q 行向量单位线性表出；②P 的生成空间包含于 Q 的生成空间。反之，若 P 的生成空间包含于 Q 的生成空间，即 P 的行向量可经 Q 的行向量单位线性表出，则必存在条件概率矩阵 L，使得 $P = L\times Q$。

证明　略。

定理 4　设 P 为 $r\times s$ 的条件概率矩阵，τ 与 κ 分别表示 P 的单位秩与在通常意义下的矩阵秩，则两者关系如下：

① 当 $\kappa = 1, 2$ 时，$\kappa = \tau$；② 当 $\kappa \geqslant 3$ 时，$\kappa \leqslant \tau$。

证明　① 当 $\kappa = 1$ 时，因 P 是条件概率矩阵，故 P 所有行必相同，即 $\tau = 1$。

当 $\kappa = 2$ 时，P 的所有行向量对应的点位于一直线上，只要行数 r 有限，必有两个最外侧的点，其他点（内部点）对应的向量可经这两端点对应的向量单位线性表出，即 $\tau = 2$。

② 当 $\kappa \geqslant 3$ 时，只要考虑 κ 顶点的多面体内部可放置任意多个顶点的多面体，就可知 $\tau \geqslant \kappa \geqslant 3$。

定理 5　设单位空间 Ω_r 到单位空间 Ω_s 的映射 ρ：

$$\forall x \in \Omega_r,\ \rho\circ x = x\times P \in \Omega_s,\ P = \left(P_{ij}\right)_{r\times s},$$

则　$\forall x, y \in \Omega_r,\ x = \left(x_1, x_2, \cdots, x_r\right),\ y = \left(y_1, y_2, \cdots, y_r\right),\ \rho\circ x = x\times P,\ \rho\circ y = y\times P,$

成立不等式：

$$d(\rho\circ x, \rho\circ y) \leqslant \left[1 - \inf_{ij}\left\{p_{ij}\right\}\right] d\left(x, y\right)。$$

证明
$$
\begin{aligned}
d(\rho\circ x,\rho\circ y) &= d(x\times P, y\times P)=\sum_j\left|\sum_i (x_i-y_i)p_{ij}\right| \\
&=\sum_j\left|\sum_{i_1}(x_{i_1}-y_{i_1})p_{i_1j}-\sum_{i_2}|x_{i_2}-y_{i_2}|p_{i_2j}\right| \\
&=\sum_{j_1}\left[\sum_{i_1}(x_{i_1}-y_{i_1})p_{i_1j_1}-\sum_{i_2}|x_{i_2}-y_{i_2}|p_{i_2j_1}\right]+ \\
&\quad\sum_{j_2}\left[\sum_{i_2}|x_{i_2}-y_{i_2}|p_{i_2j_2}-\sum_{i_1}(x_{i_1}-y_{i_1})p_{i_1j_2}\right] \\
&=\sum_{i_1}\left[|x_{i_1}-y_{i_1}|\left(\sum_{j_1}p_{i_1j_1}-\sum_{j_2}p_{i_1j_2}\right)\right]+ \\
&\quad\sum_{i_2}\left[|x_{i_2}-y_{i_2}|\left(\sum_{j_2}p_{i_2j_2}-\sum_{j_1}p_{i_2j_1}\right)\right] \\
&\leqslant\sum_{i_1}\left[|x_{i_1}-y_{i_1}|\sum_{j_1}p_{i_1j_1}\right]+\sum_{i_2}\left[|x_{i_2}-y_{i_2}|\sum_{j_2}p_{i_2j_2}\right] \\
&\leqslant\sum_{i_1}\left[|x_{i_1}-y_{i_1}|\right]\left[1-\inf_{ij}\{p_{ij}\}\right]+ \\
&\quad\sum_{i_2}\left[|x_{i_2}-y_{i_2}|\right]\left[1-\inf_{ij}\{p_{ij}\}\right] \\
&=\left[1-\inf_{ij}\{p_{ij}\}\right]d(x,y)。
\end{aligned}
$$

其中 i_1, i_2, j_1, j_2 分别使得：$(x_{i_1}-y_{i_1})\geqslant 0$, $(x_{i_2}-y_{i_2})<0$, 和

$$\sum_{i1}(x_{i_1}-y_{i_1})p_{i_1j_1}\geqslant\sum_{i_2}|x_{i_2}-y_{i_2}|p_{i_2j_2},\ \sum_{i_1}(x_{i_1}-y_{i_1})p_{i_1j_2}<\sum_{i_2}|x_{i_2}-y_{i_2}|p_{i_2j_2}。$$

定理 5 中等号成立的充要条件是：$\forall i_1, i_2, \sum_{j_2}p_{i_1j_2}=0, \sum_{j_1}p_{i_2j_1}=0$, 即 $\forall i_1, i_2, j_1, j_2$ $p_{i_1j_2}=0, p_{i_2j_1}=0$; 或 $\forall i_1, i_2, j_1, j_2, p_{i_1j_2}=p_{i_2j_1}=\inf_{i,j}\{p_{ij}\}$。

由定理5知条件概率矩阵可对应于单位空间之间的映射，除一部分点之间是等距映射外，其他点之间是压缩映射，因此亦是连续映射。

3 信道理论

在信道理论中条件概率矩阵常称为信道矩阵，一般记作：

$$P(Y|X)=[p(b_j/a_i)]_{r\times s}$$

信源（输入）空间 $[X;P]$：$\begin{cases} X & a_1 & a_2 & \cdots & a_r \\ P(X) & p(a_1) & p(a_2) & \cdots & p(a_r), \end{cases}$

信缩（输出）空间 $[Y;P]$：$\begin{cases} Y & b_1 & b_2 & \cdots & b_s \\ P(Y) & p(b_1) & p(b_2) & \cdots & p(b_s)。 \end{cases}$

其中 $\sum_{i=1}^{r}p(a_i)=1, \sum_{j=1}^{s}p(b_j)=1, p(a_i)\geqslant 0, p(b_j)\geqslant 0, i=1,2,\cdots,r, j=1,2,\cdots,s$。输入概率 $P(X)$ 取值于整个 r 维单位空间，因 $P(Y)=P(X)P(Y/X)$，故输出概率 $P(Y)$ 依赖于信道矩阵，取值于 s 维单位空间的某子空间。输入与输出的平均交互信息量（信息传输率）定义为

$$I(X,Y)=\sum_{i=1}^{r}\sum_{j=1}^{s}p(a_i)p(b_j/a_i)\log\frac{p(b_i/a_i)}{p(b_j)},$$

其中log常以2或e为底当信道矩阵给定后，$I(X,Y)$ 是输入概率的凸函数[1]，故有极大值。此最大信息传输率就定义为信道容量：$C=\max_{P(X)}\{I(X,Y)\}$。

定理 6 设 P, L 与 Q 分别是 $r\times s, r\times n$ 与 $n\times s$ 的信道矩阵，且 $P=L\times Q$，则：

① 信道矩阵 P 对应的输出空间 $G_P\subseteq$ 信道矩阵 Q 对应的输出空间 G_Q；② P 的信道容量 $C_P\leqslant Q$ 的信道容量 C_Q。

证明 ① 因 $P(Y)=P(X)P(Y/X)$，故输出概率就是信道矩阵的行向量的单位线性组合，输出空

间就是由信道矩阵的极大单位线性无关组为顶点所成的整个多面体。由定理 3 得 $G_P \subseteq G_Q$。② 设 $P=[p(b_j/a_i)]_{r\times s}$, $L=[(p(c_k/a_i)]_{r\times n}$, $Q=[P(b_j/c_k)]_{n\times s}$，输入概率：$P(X)=[p(a_1), p(a_2), \cdots, p(a_r)]$，经过信道 L 输出 $P(Z)=[p(c_1), p(c_2), \cdots, p(a_n)]$，再经过信道 Q 输出：$P(Y)=P(Z)\times Q=P(X)\times L\times Q=[p(b_1), p(b_2), \cdots, p(b_s)]$。分别考虑 X 与 Y, Z 与 Y 的平均交互信息量：

$$I(X, Y)=\sum_{i=1}^{r}\sum_{j=1}^{s}p(a_i)p(b_j/a_i)\log\frac{p(b_j/a_i)}{p(b_j)}$$

$$=\sum_{i=1}^{r}\sum_{k=1}^{n}\sum_{j=1}^{s}p(a_i)p(c_k/a_i)p(b_j/c_k)\log\frac{p(b_j/a_i)}{p(b_j)},$$

$$I(Z, Y)=\sum_{k=1}^{n}\sum_{j=1}^{s}p(c_k)p(b_j/c_k)\log\frac{p(b_j/C_k)}{p(b_j)}$$

$$=\sum_{i=1}^{r}\sum_{k=1}^{n}\sum_{j=1}^{s}p(a_i)p(c_k/a_i)p(b_j/c_k)\log\frac{p(b_j/c_k)}{p(b_j)},$$

$$I(X,Y)-I(Z,Y)=\sum_{i=1}^{r}\sum_{k=1}^{n}\sum_{j=1}^{s}p(a_i)p(c_k/a_i)p(b_j/c_k)\log\frac{p(b_j/a_i)}{p(b_j/c_k)}$$

$$\leqslant \log\sum_{i=1}^{r}\sum_{k=1}^{n}\sum_{j=1}^{s}p(a_i)p(c_k/a_i)p(b_j/c_k)\frac{p(b_j/a_i)}{p(b_j/c_k)}$$

$$=\log 1=0$$

即 $I(X, Y)\leqslant I(Z, Y)$（此处用到 log 的凸性）。故信道容量 $C_P\leqslant C_Q$。

① 与 ② 等号成立的充要条件均为 P 与 Q 有相同的极大单位线性无关组。

推论　设 $r\times s$ 的信道矩阵 P 的单位秩为 τ, 则其信道容量：

$$0\leqslant C_P\leqslant \min\{\tau, s\}。$$

定理 7　设信道矩阵 $P(Y/X)$ 中没有 1, $\varepsilon=\inf_{i,j}\{p(b_j/a_i)\neq 0\}$, 则, 信道的噪声熵：

$$H(Y/X)\geqslant -[\varepsilon\log\varepsilon+(1-\varepsilon)\log(1-\varepsilon)],$$

信道容量：

$$C\leqslant H(Y)+[\varepsilon\log\varepsilon+(1-\varepsilon)\log(1-\varepsilon)]。$$

引理　设 n 个正数 $k_1, k_2, \cdots, k_n$, $\sum_{i=1}^{n}k_i=1$, 必有：

$$\sum_{i=1}^{n}k_i\log k_i\leqslant k_1\log k_1+\left[\sum_{i=2}^{n}k_i\right]\log\sum_{i=2}^{n}k_i。$$

证明　$\sum_{i=1}^{n}k_i\log k_i=k_1\log k_1+\sum_{i=2}^{n}k_i\log k_i\leqslant k_1\log k_1+\sum_{i=2}^{n}k_i\log\sum_{j=2}^{n}k_j$

$$=k\log k+\left[\sum_{i=2}^{n}k_i\right]\log\sum_{i=2}^{n}k_i。$$

定理 7 证明　$H(Y/X)=-\sum_{i=1}^{r}\sum_{j=1}^{s}p(a_i)p(b_j/a_i)\log p(b_j/a_i)$

$$=-\sum_{i=1}^{r}p(a_i)\sum_{j=1}^{s}p(b_j/a_i)\log p(b_j/a_i)$$

$$\geqslant -\sum_{i=1}^{r}p(a_i)[p_i\log p_i+(1-p_i)\log(1-p_i)]$$

$$\geqslant -\sum_{i=1}^{r}p(a_i)[\varepsilon\log\varepsilon+(1-\varepsilon)\log(1-\varepsilon)]$$

$$=-[\varepsilon\log\varepsilon+(1-\varepsilon)\log(1-\varepsilon)]。$$

其中 $p_i=\inf_j\{p(b_j/a_i)\neq 0\}$, $i=1, 2, \cdots, r$, 并利用 $[\varepsilon\log\varepsilon+(1-\varepsilon)\log(1-\varepsilon)]$ 在 $[0, 0.5]$ 上单调增加的性质。

当 ε=0.01，0.1，0.2 时，相应的噪声熵下界分别为 log 0.945 5，log 0.722 5，log 0.606 2，因此当

$\varepsilon \geqslant 0.1$时，噪声熵将明显影响信道容量，信道矩阵列数不大于其行数时尤其显著。综上所述，从某种意义上讲信道容量大小反映输出空间大小。只有在疑义度 $H(X/Y)=0$ 时，即信道容量 $C=\log r$（r 为信道矩阵的行数）时，输出空间可认为与输入空间一样大小。否则输出空间将变小，换言之从输入到输出，部分信息可能要被丢失。

4 信道容量的计算

为求信道容量，作辅助函数：$F=I(X,Y)-\lambda\sum_{i=1}^{r}p(a_i)$，其中 λ 为待定常数。辅助函数 F 关于 $p(a_i)(i=1,2,\cdots,r)$，取偏导数并使其为零[2]，运算整理最后得矩阵方程：

$$\log[e^{1-\lambda}P(Y)]=\begin{bmatrix}1-\lambda+\log p(b_1)\\ \vdots\\ 1-\lambda+\log p(b_s)\end{bmatrix},\ -h(Y/X)=\begin{bmatrix}\sum_{j=1}^{s}p(b_j/a_i)\log(b_j/a_1)\\ \vdots\\ \sum_{j=1}^{s}p(b_j/a_r)\log(b_j/a_r)\end{bmatrix},$$

$$P(Y/X)\log[e^{1-\lambda}P^T(Y)]=-h(Y/X)。$$

定理 8 对信道矩阵的某行两元素作微小变动，信道容量亦至多微小变化。

证明 不失一般性，对信道矩阵 $P(Y/X)$ 的第 1 行的第 1，2 两元素作微小变动，得：

$$P'(Y/X)=[p'(b_j/a_i)],\ p'(b_1/a_1)=p(b_1/a_1)+\varepsilon,\ p'(b_2/a_1)=p(b_2/a_1)-\varepsilon>0$$

其他元素都相同。则

$$\begin{aligned}I'(X,Y)-I(X,Y)=&\sum_{i=1}^{r}\sum_{j=1}^{s}p(a_i)p'(b_j/a_i)\log\frac{p'(b_j/a_i)}{p'(b_j)}-\\&\sum_{i=1}^{r}\sum_{j=1}^{s}p(a_i)p(b_j/a_i)\log\frac{p(b_j/a_i)}{p(b_j)}\\=&\sum_{j=1}^{2}p(a_i)p'(b_j/a_i)\log\frac{p'(b_j/a_1)}{p'(b_j)}-\sum_{j=1}^{2}p(a_1)p(b_j/a_1)\log\frac{p(b_j/a_1)}{p(b_j)}+\\&\sum_{i=2}^{r}\sum_{j=1}^{2}p(a_i)p(b_j/a_i)\left[\log\frac{p(b_j/a_i)}{p'(b_j)}-\log\frac{p(b_j/a_i)}{p(b_j)}\right]\\=&\sum_{j=1}^{2}p(a_1)p(b_j/a_1)\log\frac{p'(b_j/a_1)p(b_j)}{p(b_j/a_1)p'(b_j)}+\varepsilon p(a_1)\log\frac{p'(b_1/a_1)p(b_1)}{p(b_1/a_1)p'(b_1)}-\\&\varepsilon p(a_1)\log\frac{p'(b_2/a_1)p(b_2)}{p(b_2/a_1)p'(b_2)}-\sum_{i=2}^{r}\sum_{j=1}^{2}p(a_i)p(b_j/a_i)\log\frac{p'(b_j)}{p(b_j)}\\=&\sum_{j=1}^{2}p(a_1)p'(b_j/a_1)\left[\log\frac{p'(b_j/a_1)}{p(b_j/a_1)}-\log\frac{p'(b_j)}{p(b_j)}\right]+O(\varepsilon)\\=&O(\varepsilon)。\end{aligned}$$

如果说某个 $p(b_j)=\sum_{i=1}^{r}p(a_i)p(b_j/a_i)=0$，则所有的 $p(a_i)p(b_j/a_i)=0,\ i=1,2,\cdots,r$，上述式子有意义。

利用定理 8 和前面的结论及（a）式易推出对一般信道容量的计算：

①当信道矩阵 P 是列线性无关时，对（a）式两边作初等变换，将 P 变为前 s 行组成单位矩阵，而后 $r-s$ 行全为零。如此就能由（a）式解得唯一的 λ 值，与相应的输出概率 $P(Y)$。

② 当信道矩阵 P 列线性相关时，若行数 r 小于列数 s，则在 P 中加上 $r-s$ 行可经前 r 行单位线性表出的向量。由定理 6 知这样做信道容量及相应的输出概率不变。若 $r\geqslant s$，则不变动。然后对新的信道矩阵的若干行作微调：将某些行的某 2 个元素分别 $+\varepsilon$ 与 $-\varepsilon$（其中不同行所选取的二元素的列数一般不同），使信道矩阵的列为线性无关。再重复 ① 过程，可解得 λ 值与相应的输出概率。由于信道容量 C 必存

在且唯一, 相应的输出概率亦唯一[1]; 如果让任意小正数 ε 趋于零, 就能得到 λ 的精确解与相应的输出概率。最后, 信道容量 $C = \lambda + \log e$, 再由 $P(Y) = P(X)P(Y/X)$ 求出相应的输入概率 $P(X)$, 当然 $P(X)$ 解可不唯一。如果是用计算机计算, 只要将 ε 取得足够小, 就可得较理想近似解。

不难看出, 上述计算信道容量的算法明显优于 [1], [2] 中所给迭代算法。

5　Markov 链

至于 Markov 链的转移概率矩阵, 自然可看作相同维数的单位空间之间的连续映射所对应的矩阵。我们利用定理 5 来说明这样一个事实。

定理 9　设有限 r 阶的转移概率矩阵 $P = (p_{ij})_{r\times r}$, 则存在一 $q(x) \in \Omega_r$, 使得 $q(x)P = q(x)$。

证明　当 $\delta = \inf_{i,j}\{p_{ij}\} > 0$ 时, 则 P 对应于一严格的压缩映射 ρ, 必有唯一一个不动点 $q(x)$, 使 $\rho(q(x)) = q(x)P = q(x)$。

当 $\delta = 0$ 时, 将 P 中所有零元素换成任意小的正数 ε, 再将所在行中某个非零元素同时相应减去同样多的 ε, 得新的转移概率矩阵 $P(\varepsilon)$。那么必有唯一一个不动点 $q(\varepsilon)$, 使得 $q(\varepsilon)P(\varepsilon) = q(\varepsilon)$。设一列 $\varepsilon_n \downarrow 0$, $P(\varepsilon_n) \xrightarrow{n\to\infty} P$, $q(\varepsilon_n)P(\varepsilon_n) = q(\varepsilon_n)$。因为 r 维单位空间是有界的完备的度量空间, 故不动点集 $\{q(\varepsilon_n)\}$ 有界。则有收敛的子向量序列: $\{q(\varepsilon_{n_k})\} \downarrow q(x) \in \Omega_r$, 使 $q(\varepsilon_{n_k})P(\varepsilon_{n_k}) = q(\varepsilon_{n_k})$; 两边取极限即可。

一般的 Markov 链理论[3] 多从到达状态的时间来进行研究, 本文从映射角度作一尝试。

参考文献:

[1] 常迥. 信息理论基础[M]. 北京: 清华大学出版社, 1993. 60—70.

[2] 姜丹, 钱玉美. 信息理论与编码[M]. 合肥: 中国科学技术大学出版社, 1992. 243—256.

[3] 钱敏平. 随机过程引论[M]. 北京: 北京大学出版社, 1990. 67—104.

On conditioned probability matrix and channel theory and Marcov chains

CUAN Yu

(Department of Information Engineering and Basic Science, Zhejiang Forestry College, Lin'an 311300, Zhejiang, China)

Abstract: Regarding all dispersed stochastic variable's distribution as "unit space" and treating conditioned probability matrixes as a mapping between unit spaces, the author applies these to information theory and obtains a better algorithm of channel capacity, and proves that there must be a steady point if the state of transition probability matrix in Marcov Chains is limited.

Key words: conditioned probability matrix; information rate; Marcov chains