it-ex05-ch7

1. 解释闭凸集。

答：设 K 是 n 维欧式空间的一个闭的点集，若任意两点 $X^{(1)}\in K, X^{(2)}\in K$ 的连线上的所有点 $\alpha X^{(1)}+(1-\alpha)X^{(2)}\in K,(0\le\alpha\le 1)$; 则 K 为闭凸集；

2. 证明: 在给定条件分布的前提下，互信息关于边际分布是凹函数。(证明要点提示:

(1) 线性函数没有凹凸性的考虑。因此互信息的条件熵部分对于凹凸性没有影响。

(2) 互信息的边际熵部分可以看成复合函数:首先是线性组合函数的作用，其次是凹函数的作用。)

证明：$I(X;Y)=H(Y)-H(Y|X)=H(Y)-\sum_x p(x)H(Y|X=x)$

不妨设固定 p(y|x),则由全概率公式，p(y)是关于 p(x)的线性函数。因而，关于 p(y)的凹函数 H(Y)也是 p(x)的凹函数。上式中第二项是关于 p(x)的线性函数。因此，它们的差仍然是 p(x)的凹函数.

3. 叙述并证明 KT 条件。

$f(\alpha)$定义在 R^n 上的凹（concave）函数，$\alpha=(\alpha_1,\cdots\alpha_n)$ 是概率矢量。

假定$\dfrac{\partial f(\alpha)}{\partial\alpha_i}$ （i=1,2,…,n）均存在，且在R^n 上连续，则 $f(\alpha)$ 在R^n 上取极大值的充要条件是

$$\begin{cases}\dfrac{\partial f(\alpha)}{\partial\alpha_i}=\lambda,\forall\alpha_i>0\\ \dfrac{\partial f(\alpha)}{\partial\alpha_i}\le\lambda,\forall\alpha_i=0\end{cases}$$

证明：

充分性证明：

设 f 在点α 处满足$\begin{cases}\dfrac{\partial f(\alpha)}{\partial\alpha_i}=\lambda,\forall\alpha_i>0\\ \dfrac{\partial f(\alpha)}{\partial\alpha_i}\le\lambda,\forall\alpha_i=0\end{cases}$，下面证明$f(\alpha)$为极大值

即对任意的概率矢量$\beta\in R^n$ ，成立 $f(\beta)-f(\alpha)\le 0$ 。

$f(\alpha)$ 是定义在R^n 上的凹函数，所以对任意的$\theta\in(0,1)$ ，有

$$\theta f(\beta)+(1-\theta)f(\alpha)\le f(\theta\beta+(1-\theta)\alpha)$$

$$f(\beta)-f(\alpha)\le\frac{f(\theta\beta+(1-\theta)\alpha)-f(\alpha)}{\theta}$$

其中$\theta\beta+(1-\theta)\alpha$ 也是概率矢量（注解：全体概率矢量形成闭凸集）。

由于$\theta\in(0,1)$ 的任意性，可令$\theta\to 0^+$，得到

$$f(\beta)-f(\alpha)\le\left.\frac{df(\theta\beta+(1-\theta)\alpha)}{d\theta}\right|_{\theta=0}$$

$$=\left.\frac{df(\alpha+\theta(\beta-\alpha))}{d\theta}\right|_{\theta=0}$$

$$=\sum_{i=1}^{n}\frac{\partial f(\alpha)}{\partial\alpha_i}\quad(\beta_i-\alpha_i)$$

$$\le\sum_{i=1}^{n}\lambda(\beta_i-\alpha_i)$$

$$=\lambda(\sum_{i=1}^{n}\beta_i-\sum_{i=1}^{n}\alpha_i)$$

$$=\lambda(1-1)$$

$$=0$$

充分性获证。

充分性证明中的一个注解：

方向导数的定义$\frac{\partial f(\alpha)}{\partial d}=\lim_{\theta\to0^+}\frac{f(\alpha+\theta d)-f(\alpha)}{\theta\|d\|},d=\beta-\alpha$

$$\frac{\partial f(\alpha)}{\partial d}=<\nabla f(\alpha),\frac{d}{\|d\|}>$$

$$\lim_{\theta\to0^+}\frac{f(\alpha+\theta d)-f(\alpha)}{\theta}=\frac{\partial f(\alpha)}{\partial d}\|d\|=<\nabla f(\alpha),\frac{d}{\|d\|}>\|d\|=<\nabla f(\alpha),d>$$

必要性证明：

设 f 在点α 处取得极大值，且$\frac{\partial f(\alpha)}{\partial\alpha_i}$ （i=1,2,…,n）在点α 处连续，

则对任意概率矢量$\beta\in R^n$ ， $\theta\in(0,1)$ 成立

$$f(\theta\beta+(1-\theta)\alpha)-f(\alpha)\le0$$

$$\frac{f(\theta\beta+(1-\theta)\alpha)-f(\alpha)}{\theta}\le0$$

其中$\theta\beta+(1-\theta)\alpha$ 也是概率矢量。

由于$\theta\in(0,1)$ 的任意性，可令$\theta\to0^+$ ，得到

$$
\begin{aligned}
&\left.\frac{df(\theta\beta+(1-\theta)\alpha)}{d\theta}\right|_{\theta=0} \\
&=\left.\frac{df(\alpha+\theta(\beta-\alpha))}{d\theta}\right|_{\theta=0} \\
&=\sum_{i=1}^{n}\frac{\partial f(\alpha)}{\partial\alpha_i}(\beta_i-\alpha_i)\leq 0
\end{aligned}
$$

由于α 是概率矢量，所以至少有一个分量是严格正的（不妨设为$\alpha_1 > 0$，且令$\lambda = \frac{\partial f(\alpha)}{\partial \alpha_1}$ ），

又由β 的任意性，不妨设

$$\beta_i = \begin{cases} \alpha_1 - \varepsilon (i=1) \\ \alpha_k + \varepsilon (i=k \neq 1) \\ \alpha_i (i \neq k, i \neq 1) \end{cases}$$

$0 < \varepsilon < \alpha_1$

$$\varepsilon \frac{\partial f(\alpha)}{\partial \alpha_k} - \varepsilon \frac{\partial f(\alpha)}{\partial \alpha_1} \le 0 \Rightarrow \frac{\partial f(\alpha)}{\partial \alpha_k} \le \lambda$$

若$\alpha_k > 0$ ，则还可取$\varepsilon < 0$

$$\varepsilon \frac{\partial f(\alpha)}{\partial \alpha_k} - \varepsilon \frac{\partial f(\alpha)}{\partial \alpha_1} \le 0 \Rightarrow \frac{\partial f(\alpha)}{\partial \alpha_k} \le \lambda$$

从而$\frac{\partial f(\alpha)}{\partial \alpha_k} = \lambda$

若$\alpha_k = 0$ ，则只可取$\varepsilon > 0$ ，$\frac{\partial f(\alpha)}{\partial \alpha_k} \le \lambda$ 必要性获证。

4. 叙述并解释迭代算法。
答：
Algorithm for channel capacity

Input： $P = (p_{ij})_{n \times m}$

Initialization: $S^{(0)} = (\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n})$

Iteration:

$$\varphi_{ji}^{(k)} = \frac{s_i^{(k)} p_{ij}}{\sum_{i=1}^{n} s_i^{(k)} p_{ij}}$$

$$s_i^{(k+1)} = \frac{\exp(\sum_{j=1}^{m} p_{ij} \log \varphi_{ji}^{(k)})}{\sum_{i=1}^{n} \exp(\sum_{j=1}^{m} p_{ij} \log \varphi_{ji}^{(k)})}$$

$$C^{(k+1)} = \log(\sum_{i=1}^{n} \exp(\sum_{j=1}^{m} p_{ij} \log \varphi_{ji}^{(k)}))$$

5. 叙述并解释信道编码定理。

答：信道编码定理：对于离散无记忆信道，小于信道容量 C 的所有码率都是可达的。具体来说，对任意码率 R<C，存在一个 $(2^{nR}, n)$ 码序列，它的最大误差概率为 $\lambda^{(n)} \to 0$.

反之，任何满足 $\lambda^{(n)} \to 0$ 的 $(2^{nR}, n)$ 码序列必定有 R<=C.

这个定理是信道编码的理论依据，可以看出：信道容量是一个明确的分界点，当取分界点以下的信息传输率时,以指数趋进于 0; 当取分界点以下的信息传输率时,以指数趋进于 1；因此在任何信道中，信道容量都是可达的、最大的可靠信息传输率.

这个定理是一个存在定理，它没有给出一个具体可构造的编码方法，在它的证明过程中，码序列是随机的选取的，它有助于指导各种通信系统的设计，有助于评价各种系统及编码的效率。

6. 解释信道编码。

答：由于移动通信存在干扰和衰落，在信号传输过程中将出现差错，故对数字信号必须采用纠、检错技术，即纠、检错编码技术，以增强数据在信道中传输时抵御各种干扰的能力，提高系统的可靠性。对要在信道中传送的数字信号进行的纠、检错编码就是信道编码。

7. 解释 ARQ.

答：反馈重发（automatic repeat request, ARQ）,发送端发送检错码，如果循环冗余校验码(CRC),接收端通过检测接收码是否符合编码规律来判断该码是否存在差错。若判定码组有错，则通过反向信道通知发端重发该码，如此反复，直到收端认为正确接收为止。

8. 解释 FEC.

答：前向纠错(forward error correction, FEC), 发送端信息经纠错码后实行传送，而接收端通过纠错码自动纠正传递过程中的差错。所谓“前向”，指纠错过程在收端独立进行，不存在差错信息的反馈。