04104-106通信原理II第三讲

March 13, 2007

1 量化

连续随机变量X的熵是无穷大，量化器把它映射到M进制的离散随机变量Y，传输Y最多只需要$\log_2 M$bits/symbol，如果Y的这M个量化电平不是等概出现的话，还能更少。

对于通常的量化设计，如果已知X，那么Y就是确定的，但已知Y并不能确定X，即$I(X;Y) = H(Y) < H(X)$。因此Y和X之间的失真D必不为零。给定D的情况下，理论上最好的量化器所需要的传输速率是$R(D)$（率失真函数）。

对于一些具体的量化器，它的速率也是失真的函数。比如，若Y均匀分布，那么均匀量化器的量化信噪比是$\frac{S}{N_q} = M^2$。不妨假设信号功率$S = E[X^2] = 1$(即X在$[-\sqrt{3}, +\sqrt{3}]$内均匀分布)。量化噪声$N_q \triangleq E[(Y-X)^2]$就是均方失真$D$。而传输速率$R$是每样值（即每符号）$\log_2 M$bit，此情形下速率与失真的函数关系是$R(D) = -\log_2 \sqrt{D}$，其曲线与图7.7.1类似。若要求$D=0$，则$R$就得无穷大（因为$M$无限大）；如果允许$D=1$的话，$R=0$，说明不用传就可以。这种方案是，无论$Y$是多少，收端始终输出$X=0$。

注意，信息论讲的是我们最好能做到什么程度（理论上的可能性）。7.9–7.11节讲的是一些具体做法（真实存在的技术），前面的Huffman编码也是一种具体技术。假如我们通过适当的数学模型能算出7.8节给出的$R(D)$，那么任何实际的量化系统都不可能做到比它更好。

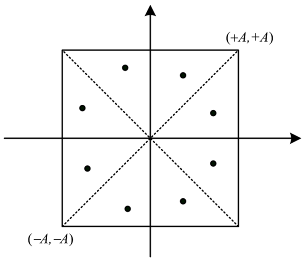
7.9.3讲的是标量量化，均匀量化是这种量化的一种最简单形式。就一般而言，量化的设计包括两点：

1. 把X的取值范围$(-\infty, +\infty)$或者$(-A, +A)$分割成M个不重叠的子区间（量化区间）；
2. 对每个区间设计一个代表此区间的点：量化电平。

发送端传给接收端的是子区间的编号。收端只要知道X在那个子区间就够了。接收端知道每个区间的量化电平是多少，因为这是事先设计的。

均匀量化只是一种简单的量化方式。给定M时，能使D最小的量化方法见课本326页。它的思路是，把失真D表达成多元函数，其自变量是分界点和代表点（量化电平）。再用数学中求多元函数极值的方法得到最佳的设计。其主要结论是

1. 边界必须是代表点的等距分界点；
2. 代表点必须是概率质心，或者说是条件数学期望。



上述量化是对标量X进行的，也可以扩展到矢量，即把$(X_1, X_2, \cdots, X_N)$当作一个整体，将它进行空间划分，再把矢量所在的区域编号传给收端。收端得到这个编号后，用预先为该区域所设计的代表点$(Y_1, Y_2, \cdots, Y_N)$来近似信源端的矢量$(X_1, X_2, \cdots, X_N)$。这些代表点的集合叫做码本(codebook) 。发端储存着划分空间的方法，收端储存着码本。

比如说，(X_1, X_2)的取值范围是平面上的一个正方形，这个正方形的左下角的坐标是$(-A, -A)$，右上角的坐标是(A, A) 。采用笛卡儿积的记号方法，(X_1, X_2)的取值范围就是$(-A, A) \times (-A, A)$。$M = 8$的矢量量化就是把这个正方形分割成8个区域，并为每个区域设计一个代表点。如下图所示

如果X_1, X_2落在某个区域中，就把这个编号告诉收端。这样做的传输速率是$R = \frac{\log_2 8}{2} = 1.5$bits/symbol，相应的失真度取决于区域的划分方法和代表点的选择。对于最佳的设计，边界线是代表点之间的等距线，代表点是该区域的概率质心（条件数学期望）。

2 相关信源

如果信源序列之间有相关性，表明它们之间的互信息不为0，每一个符号单独新增的信息并不一定很大。预测编码和变换编码的原理都是一样的，就是先把信源改造成不相关序列。改变相关性的常用做法是进行线性变换。

预测编码中，我们不是直接传送第L个信源符号X_L（对语音信号，就是抽样值），而是传送它和预测值$\hat{X}_L$的差：$Y_L = X_L - \hat{X}_L$。整体看，等于是把序列$\{X_i\}$变换成了序列$\{Y_i\}$。理想的情况下，$\{Y_i\}$应该是个独立序列，表明预测值完全包含了过去符号的贡献，使得扣除预测值后的预测误差只包含新增的信息。注意，理想情况不是$\hat{X}_L = X_L$，除非$H_\infty(X) = 0$，否则这是不可能的。

预测器用过去的结果计算出预测值，可记为$\hat{X}_L = f(X_1, X_2, \cdots, X_{L-1})$。这个$f$一般用线性函数，即$\hat{X}_L$是$X_1, X_2, \cdots, X_{L-1}$的线性组合。这样，$Y_L = X_L - \hat{X}_L$自然也是$X_1, X_2, \cdots, X_{L-1}$的线性组合。因此，把信源$\{X_i\}$映射为$\{Y_i\}$的过程就是经过了一个因果的滤波器（线性系统）。这个例子中的规则是：得到Y_L不得利用未来的符号$\{X_{L+1}, X_{L+2}, \cdots\}$（因果性）。

在另外一些应用中，信源$\{X_1, \cdots, X_L\}$不一定表示按时间顺序先后到达的符号（也许是空间次序，比如图片）。或者虽然是时间顺序，但我们可以批量处理，一次处理L个，下次再处理L个。它和刚才的差别是，得到Y_k时，我们可以用它后面的X_{k+1}。对于固定长度为L的序列，能去除相关性的线性变换就是一个矩阵变换：$[Y_1, Y_2, \cdots, X_L]^T = \mathbf{A}[X_1, X_2, \cdots, X_L]^T$。

总而言之，相关信源中的相关性是一种冗余，去除相关性就能提高传输效率。预测编码和变换编码都是借助线性变换来改变相关性，只不过一个是因果

滤波器，一个是矩阵变换。它们的实质是相同的。实际的技术或许出于复杂性或其他原因不能完全消除相关性，但至少能使它显著变弱。

3　信道

信道是一个抽象的概念，指所研究问题中从发送端到接收端之间的一切环节。不同的情景下，“发送端”和“接收端”所指不同，因而信道的含义也不同。例如：

1. 在无线通信中，发送天线到接收天线之间的电波传播空间是一种信道，这个信道也叫“狭义信道”或者“传输媒质”；
2. 从调制器的输出端到解调器的输入端的一切环节是一种信道，它的内部包含了上下变频器、放大器、收发滤波器、收发天线、传输媒质等；
3. 从A律十三折线编码器的输出端到接收机中A律十三折线译码器的输入端之间的一切环节也是一种信道，它的内部包括了调制解调器以及前面所提到的信道。

再比如，下图是一个BPSK系统的框图。图中的“编码器”和“译码器”指的是第九章要讲的信道编码，目的是为了纠正BPSK解调后仍然存在的误码。对于不同的研究，信道可能是$A \to B$（无线传输媒质）、$C \to D$（BPSK调制的带通信道）、$I \to J$（BPSK调制的等效基带信道）、$G \to H$（离散时间的BPSK等效信道）、$E \to F$（二元数字信道）等等。 无论是哪一种情况，总

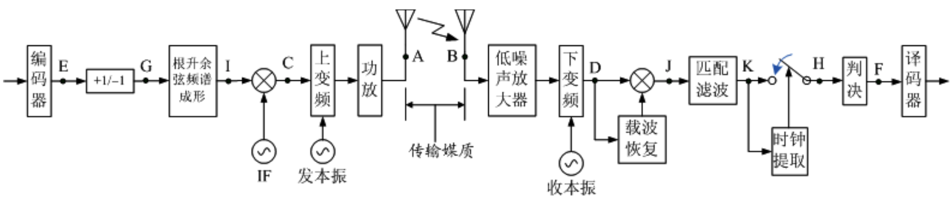


Figure 1: 某BPSK系统

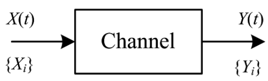


Figure 2: 信道模型

可以把问题抽象为Figure. 2：信道的输入或者是连续时间的随机过程$X(t)$，或者是离散时间的随机序列$\{X_i\}$。信道的输出也可能是随机过程$Y(t)$或者随机序列$\{Y_i\}$。信道所起的作用就是把输入的随机过程或序列映射为另外一个随机过程或者序列。这种映射其本身往往具有随机性。也就是说，观察到输出Y，并不能确定发送的X，也即$I(X;Y) < H(X)$。

4　无线多径信道

考虑图1中的信道$A \to B$，如果信道的输入是$X(t)$，输出$Y(t)$必然会包括噪声和干扰的影响。除此之外，无线传播环境也会带来许多复杂的影响。由于后者

是无线通信信道最具特征的东西，所以在无线信道建模时，我们关键是要搞清楚信道传播对信号带来的影响。故此，我们忽略噪声和干扰，先来看传播过程造成的影响。

设想手机中藏有一个观察者，那么在这个观察者看来

1. 接收信号的强度在随机变化。此现象叫做衰落(fading)

2. 接收信号的频率在变化。此现象叫多普勒频移（Doppler shift）。

他观察到这些的原因在于：(1)用户在移动；(2)到达手机的无线信号是经由多个路径到达的。

Doppler的原理在于，波的相位是时间和空间的函数。如果频率是f_0，那么在静止的地点，经过Δt时间后观察到相位变化是$\frac{\Delta t}{T_0} \times 2\pi = 2\pi f_0 \Delta t$($T_0$是周期)。如果用户同时还在以速度$v$移动，那么经过$\Delta t$时间后空间相位的变化是$\frac{\Delta t \times v}{\lambda} \times 2\pi$（$\lambda$是波长）。藏在手机中的观察者不知道这两种相位变化的区别，他看到的只是，经过Δt时间后，到达手机的信号的相位变化了$\frac{\Delta t \times v}{\lambda} \times 2\pi + 2\pi f_0 \Delta t = 2\pi(f_0 + v/\lambda)\Delta t$。于是，他以为发送信号的频率是$f_0 + v/\lambda$，因为他不知道手机在移动（想象一下，他只能看示波器之类的仪表，不许他看手机外面的风景）。或者说，在他看来这和发送频率就是$f_0 + v/\lambda$没有区别。如果考虑到移动方向θ，上述v应改为$v\cos\theta$。$v/\lambda\cos\theta$叫做方向θ上的Doppler频移，其最大值v/λ叫最大多普勒频移。

接收信号所呈现出的衰落还可分为平衰落和频率选择性衰落。假设发送一个复包络为$b(t)$ 的带通信号，这个信号将经过许多路径到达手机。每一径所经历的幅度衰减μ_i，相移φ_i、时延τ_i都可能不同，并且它们可能会随时间变化（因为手机在移动）。通常，手机移动造成的某一径上的时延变化不大。因为电磁波以光速传播，时延变化$1\mu s$意味着位置相隔300m。在短时间内，手机位置不可能发生这么大的变化。因此一般可以忽略τ_i随时间的变化。因而第i径在t时刻接收到的带通信号的复包络是$\mu_i(t)e^{j\varphi_i(t)}b(t-\tau_i)$。总接收信号是

$$r(t) = \mathrm{Re}\left\{\sum_i \mu_i(t)e^{j\varphi_i(t)}b(t-\tau_i)e^{j\omega_c t}\right\} \tag{1}$$

如果各路径之间的时延差相对于$b(t)$来说非常小，比如，各路径的时延的差别在μs数量级，而$b(t)$是速率为1kbps的双极性NRZ信号（BPSK的复包络），那么式(1)就可以近似成

$$r(t) \approx \mathrm{Re}\left\{\left[b(t-\bar{\tau})\sum_i \mu_i(t)e^{j\varphi_i(t)}\right]e^{j\omega_c t}\right\} \tag{2}$$

相当于让复包络乘上了一个时变的复数$h(t) = \sum_i \mu_i(t)e^{j\varphi_i(t)}$。这种情形叫平衰落。如果手机不动，则$\mu_i$和$\varphi_i$都不是时间的函数，因而$h$是个复常数。如果信号所经过的系统只是乘了复常数，这个系统的传递函数在所有频率上都一样，即其频域特性是平的。平衰落在任何一个瞬时都是一个频域特性为常数的线性系统。它不会对发送的信号造成波形失真（假设$h(t)$的变化比$b(t)$的变化慢得多），但不同瞬时的信噪比是随机变化的。

注意，以上讨论中，我们假设信道的时变性仅由于手机移动。实际上，发送端的移动，以及周边环境中物体的移动也会产生同样的效果。