数字信号处理中的几个关键概念

主讲人 宋 辉

清华大学电子工程系 博士
滴滴AI Labs 语音技术部

2.1 数字信号及其基本运算

2.2 采样定理

2.3 时频分析与傅里叶变换

2.4 作业

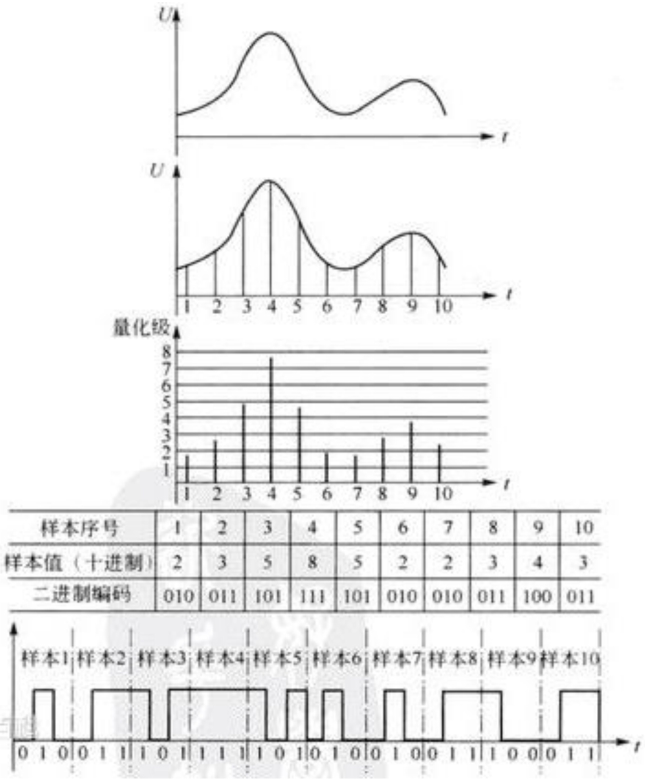
2.1 数字信号及其基本运算

信号是信息的物理载体，信息是信号的具体内容。

- 连续时间信号：在连续时间范围内定义的信号，信号的幅度可以是连续的（模拟信号），也可以是离散的。
- 离散时间信号：时间为离散变量的信号，即独立变量时间被量化了，而幅度仍是连续变化的。
- 数字信号：时间离散而幅度量化的信号。

由模拟信号到数字信号

模拟数据 → 采样 → 量化 → 编码 → 数字信号



样本序号	1	2	3	4	5	6	7	8	9	10
样本值（十进制）	2	3	5	8	5	2	2	3	4	3
二进制编码	010	011	101	111	101	010	010	011	100	011

Pulse Code Modulation

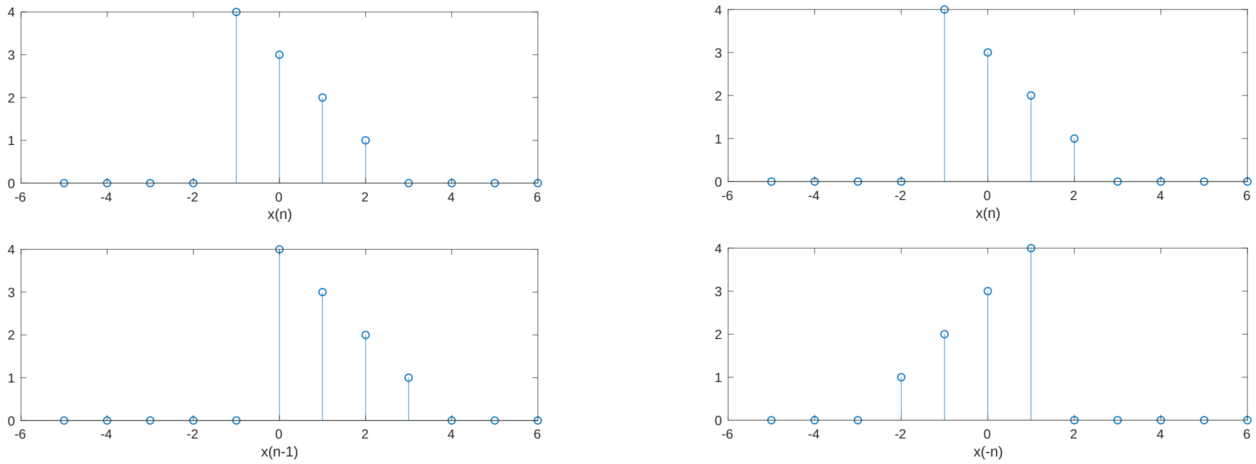
数字信号的基本运算

移位

设某一序列 x(n)，当 m>0 时，x(n-m) 表示序列 x(n) 逐项依次延时（右移）m 位。

翻褶

设某一序列 x(n)，则 x(-n) 是以 n=0 的纵轴为对称轴将 x(n) 加以翻褶。



数字信号的基本运算

和

$$z(n) = x(n) + y(n)$$

积

$$z(n) = x(n) \cdot y(n)$$

累加

$$y(n) = \sum_{k=-\infty}^{n} x(k)$$

差分

$$y(n) = x(n) - x(n-1)$$

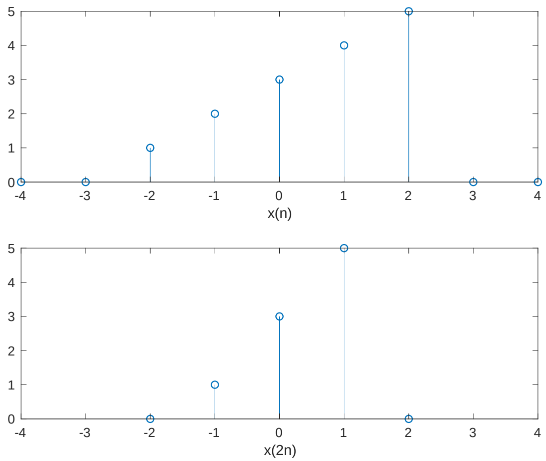
Q：一阶差分运算具有高通滤波的效果，为什么？

数字信号的基本运算

尺度变换

对于序列 $x(n)$，形如 $x(mn)$ 或者 $x(\frac{n}{m})$（m为正整数）的序列为 $x(n)$ 的尺度变换序列。

以 m=2 为例，$x(2n)$ 是以低一倍的抽样频率从 $x(n)$ 中每隔两点取一点，这种运算称为抽取，通常表示为 $\boxed{\downarrow M}$。类似的，$x(\frac{n}{2})$ 称为插值。

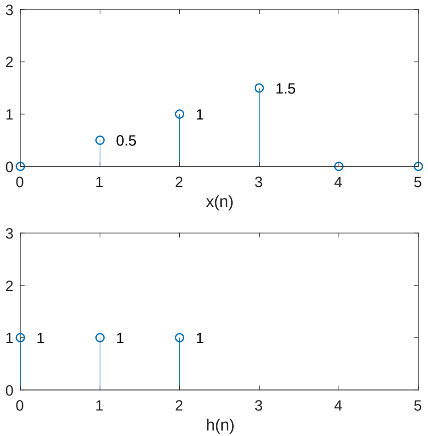


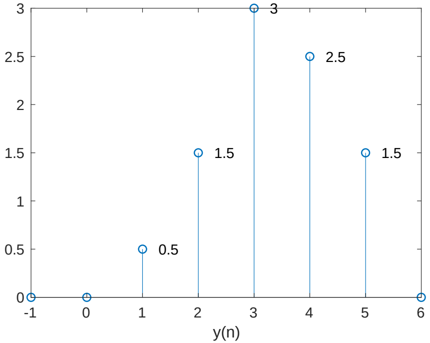
数字信号的基本运算

线性卷积（linear convolution）

$$y(n)=\sum_{m=-\infty}^{\infty}x(m)h(n-m)=x(n)*h(n)$$

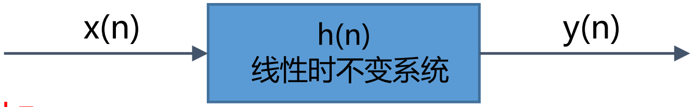
由卷积的定义可知，卷积在图形表示上可分为四步：翻褶、移位、相乘、相加。



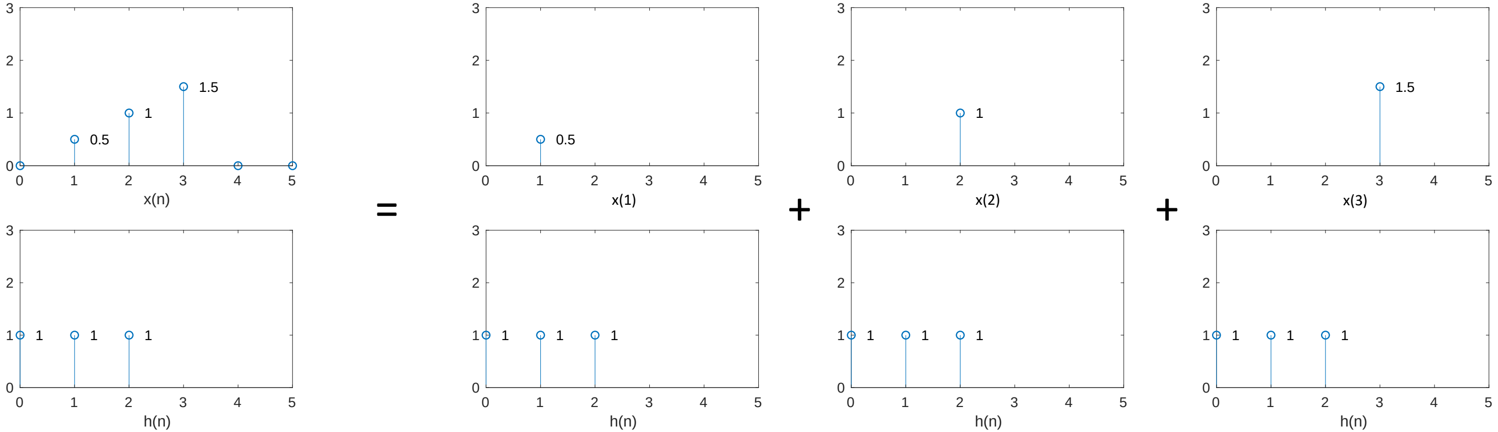


数字信号的基本运算

线性卷积的另一种理解



线性卷积实质上是一个信号在另一个信号上的**加权叠加**。

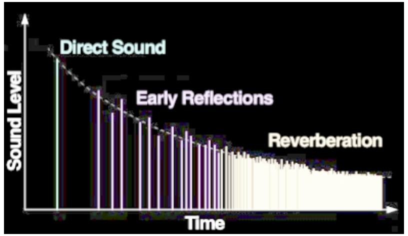


对于线性时不变系统，如果已知系统的单位冲激响应，那么将单位冲激响应与输入信号求卷积，就得到了输出信号。

数字信号的基本运算

线性卷积的应用：模拟远场数据

*



=

near-field speech

room impulse response

simulated farfield speech

数字信号的基本运算

圆周移位（circular shift）

$$x_m(n) = x((n+m))_N R_N(n)$$

其中，$x((n+m))_N$ 表示 $x(n)$ 经过周期（N）延拓后的序列，再移位 m。

$$R_N(n) = \begin{cases} 1 & 0 \leqslant n \leqslant N-1 \\ 0 & other\ n \end{cases}$$

数字信号的基本运算

圆周卷积（circular convolution）

如果 $x_1(n)$ 和 $x_2(n)$ 都是长度为N的有限长序列（$0 \leqslant n \leqslant N-1$），并且

$$DFT[x_1(n)] = X_1(k) \qquad DFT[x_2(n)] = X_2(k)$$

$$Y(k) = X_1(k)X_2(k)$$

则

$$y(n) = IDFT[Y(k)] = \left[\sum_{m=0}^{N-1} x_1(m)x_2((n-m))_N\right] R_N(n)$$

$$= \left[\sum_{m=0}^{N-1} x_2(m)x_1((n-m))_N\right] R_N(n)$$

定义为 $x_1(n)$ 和 $x_2(n)$ 的圆周卷积。

注意：与线性卷积相比，圆周卷积多了周期延拓和取主值序列两个步骤。因此必须指定圆周卷积的点数N。

圆周卷积和线性卷积的关系

圆周卷积

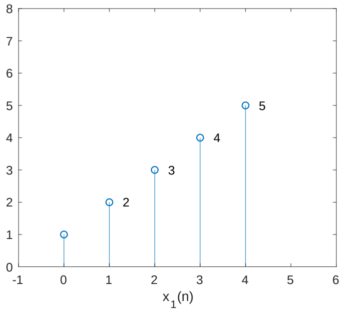
$$y(n) = IDFT[Y(k)] = \left[\sum_{m=0}^{N-1} x_1(m)x_2((n-m))_N\right] R_N(n)$$

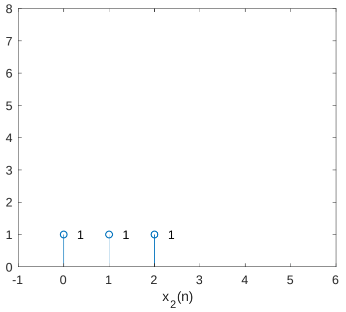
$$= \left[\sum_{m=0}^{N-1} x_2(m)x_1((n-m))_N\right] R_N(n)$$

线性卷积

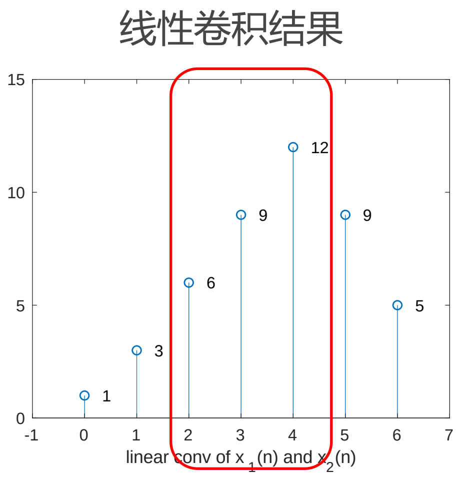
$$y(n) = \sum_{m=-\infty}^{\infty} x(m)h(n-m) = x(n) * h(n)$$

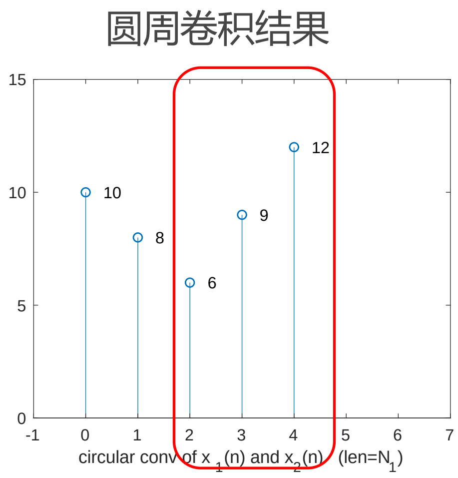
给定两个有限长序列 $x_1(n)$ 和 $x_2(n)$，他们的长度分别为：$N_1 = 5$ ，$N_2 = 3$ 。相应的取值如下图，我们重点研究 $0 \leqslant n \leqslant N_1 - 1$ 这个区间内，线性卷积和圆周卷积的关系。





圆周卷积和线性卷积的关系





观察：圆周卷积结果的后一部分与线性卷积的结果

一般的，如果两个有限长序列的长度为 N_1 和 N_2 ，且满足 $N_1 \geqslant N_2$ ，则圆周卷积的后 N_1-N_2+1 个点，与线性卷积的结果一致。

数字信号的基本运算

线性相关（linear correlation)

$$r_{xy}(m)=\sum_{n=-\infty}^{\infty}x(n)y^*(n-m)$$

请大家思考: 线性相关定义的形式与线性卷积的相似与不同?

圆周相关（circular correlation)

如果: $R_{xy}(k)=X(k)Y^*(k)$

则 $r_{xy}(m)=IDFT[R_{xy}(k)]=\sum_{n=0}^{N-1}y^*(n)x((n+m))_N R_N(m)=\sum_{n=0}^{N-1}x(n)y^*((n-m))_N R_N(m)$

定义为 $x(n)$ 和 $y(n)$ 的圆周相关。

圆周相关和线性相关的关系

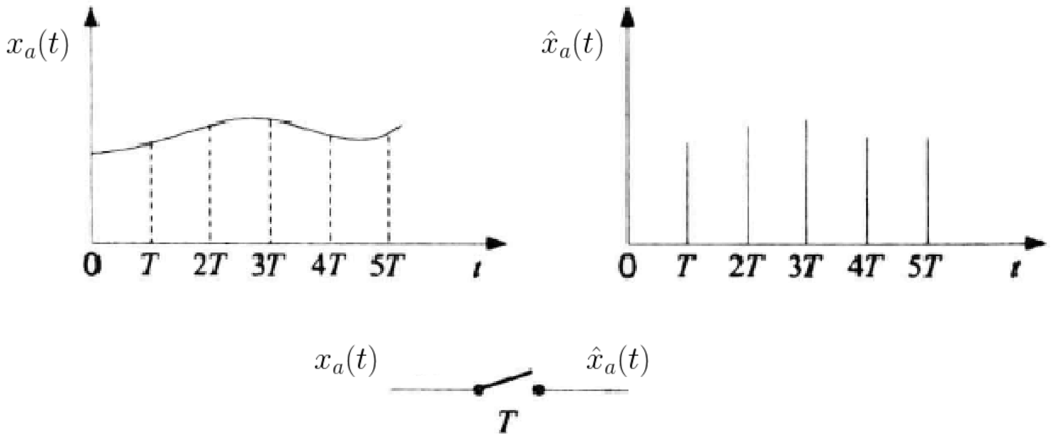
线性相关和圆周相关的关系，大家也可以用上面的方法尝试验证。

这里直接给出结论：

一般的，如果两个有限长序列的长度为 N_1 和 N_2 ，且满足 $N_1 \geqslant N_2$ ，则有圆周相关的前 $N_1 - N_2 + 1$ 个点，与线性相关的结果一致。

2.2 采样定理（Nyquist-Shannon sampling theorem)

模拟信号的采样：



2.2 采样定理（Nyquist-Shannon sampling theorem)

采样：利用周期性冲激函数序列，从连续信号 $x_a(t)$ 中抽取一系列的离散值，得到采样信号，即离散时间信号 $\hat{x}_a(t)$ 。

冲激函数序列:
$$\delta_T(t) = \sum_{m=-\infty}^{\infty} \delta(t - mT) \tag{2.1}$$

则，采样信号:
$$\hat{x}_a(t) = x_a(t) \cdot \delta_T(t) \tag{2.2}$$

将(2.1)代入(2.2)，得:
$$\hat{x}_a(t) = \sum_{m=-\infty}^{\infty} x_a(t)\delta(t - mT) \tag{2.3}$$

由于 $\delta(t - mT)$ 只在 $t = mT$ 处不为零，因此:
$$\hat{x}_a(t) = \sum_{m=-\infty}^{\infty} x_a(mT)\delta(t - mT) \tag{2.4}$$

2.2 采样定理

采样后信号频谱的变化

$$\hat{x}_a(t) = x_a(t) \cdot \delta_T(t)$$

$$\Big\downarrow \text{等式两端取DTFT}$$

$$\hat{X}_a(j\Omega) = \frac{1}{2\pi}[\Delta_T(j\Omega) * X_a(j\Omega)] = \frac{1}{T}\sum_{k=-\infty}^{\infty} X_a(j(\Omega - k\Omega_s)) \qquad (2.5)$$

其中 $\Delta_T(j\Omega) = DTFT[\delta_T(t)]$

结论：频谱产生了周期延拓，周期为 Ω_s 。

因此，只要各延拓分量与原频谱分量不发生频率交叠，则可以恢复原信号。

2.2 采样定理

公式推导

$$\hat{X}_a(j\Omega) = \frac{1}{2\pi}[\Delta_T(j\Omega) * X_a(j\Omega)] = \frac{1}{T}\sum_{k=-\infty}^{\infty} X_a(j(\Omega - k\Omega_s))$$

其中 $\Delta_T(j\Omega) = DTFT[\delta_T(t)]$ 由于 $\delta_T(t)$ 是周期信号（周期为T），则可以表示成傅里叶级数

$$\delta_T(t) = \sum_{k=-\infty}^{\infty} A_k e^{jk\Omega_s t}$$

其中 $\Omega_s = \dfrac{2\pi}{T}$ 为采样频率，

$$A_k = \frac{1}{T}\int_T \delta_T(t)e^{-jk\Omega_s t}dt = \frac{1}{T}\int_T \sum_{m=-\infty}^{\infty} \delta(t - mT)e^{-jk\Omega_s t}dt$$

$$= \frac{1}{T}\int_T \delta(t)e^{-jk\Omega_s t}dt = \frac{1}{T}$$

因为在一个积分区间T [-T/2, T/2] 内，只有一个冲激函数。

2.2 采样定理

公式推导

$$\hat{X}_a(j\Omega) = \frac{1}{2\pi}\left[\Delta_T(j\Omega) * X_a(j\Omega)\right] = \frac{1}{T}\sum_{k=-\infty}^{\infty} X_a(j(\Omega - k\Omega_s))$$

2.2 采样定理

奈奎斯特采样定理：

要想采样后能够无失真的还原出原信号，则采样频率必须大于两倍信号谱的最高频率。

$$f_s > 2f_h$$

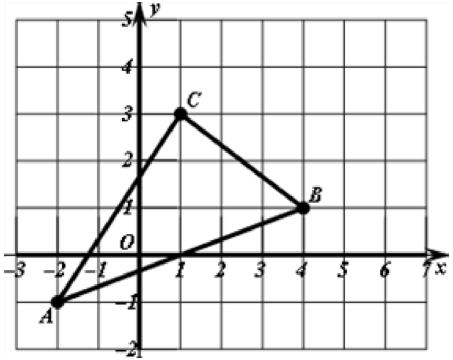
16kHz语音如何降采样到8kHz?

空间“采样定理”

	频域	空域
采样	冲激函数序列	麦克风
采样率	f_s	d
信号的最高频率	f_h	λ_{min}
防混叠条件	$f_s > 2f_h$	$d < \lambda_{min}/2$

2.3 时频分析与傅里叶变换

变换是一种常用的数学工具



A(-2, -1) $-2\mathbf{e}_x - \mathbf{e}_y$

B(4, 1) $4\mathbf{e}_x + \mathbf{e}_y$

C(1, 3) $\mathbf{e}_x + 3\mathbf{e}_y$

其中 $\mathbf{e}_x$ 和 $\mathbf{e}_y$ 构成标准正交基，满足如下条件：

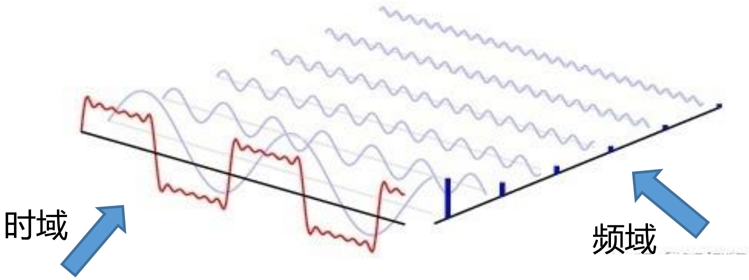
$$\begin{cases} \|\mathbf{e}_x\| = \|\mathbf{e}_y\| = 1 \\ \langle \mathbf{e}_x, \mathbf{e}_y \rangle = 0 \end{cases}$$

前面的系数表示平面中的点在这个基向量方向上有多少个单位长度。

2.3 时频分析与傅里叶变换

利用正弦波模拟方波

何为“频域”？



2.3 时频分析与傅里叶变换

傅里叶级数 （Fourier Series）

如果 $x(t)$ 是一个周期为 T_0 的周期性连续函数，则 $x(t)$ 可展开成傅里叶级数：

$$x(t) = \sum_{k=-\infty}^{\infty} X(jk\Omega_0) e^{jk\Omega_0 t}$$

$$X(jk\Omega_0) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) e^{-jk\Omega_0 t} dt$$

解读： $\Omega_0 = 2\pi F = \dfrac{2\pi}{T_0}$ 傅里叶级数系数的计算，实质上是通过内积的方式， “抽取”对应频率分量的系数。

2.3 时频分析与傅里叶变换

连续傅里叶变换 （Fourier Transform）

连续非周期信号 $x(t)$ 的傅里叶变换可以表示为：

$$X(j\Omega) = \int_{-\infty}^{\infty} x(t) e^{-j\Omega t} dt$$

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(j\Omega) e^{j\Omega t} d\Omega$$

解读：这应该是大家在“信号与系统”里学到的第一个傅里叶变换公式。它仍然是通过内积的方式，“抽取”对应频率分量的系数。

与傅里叶级数不同的是，由于时域信号非周期，因此频域中是连续谱。

离散时间傅里叶变换 （Discrete Time Fourier Transform）

离散非周期信号 x(n) 的 DTFT 可以表示为：

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n}$$

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} d\omega$$

解读：DTFT与傅里叶级数互为正反变换。

离散傅里叶变换 （Discrete Fourier Transform）

离散周期信号 x(n) 的 DFT 可以表示为：

$$X(k)=\sum_{n=0}^{N-1}x(n)e^{-j\frac{2\pi}{N}nk}=\sum_{n=0}^{N-1}x(n)W_N^{nk}$$

$$x(n)=\frac{1}{N}\sum_{k=0}^{N-1}X(k)e^{j\frac{2\pi}{N}nk}=\frac{1}{N}\sum_{k=0}^{N-1}X(k)W_n^{-nk}$$

$$W_N=e^{-j\frac{2\pi}{N}}$$

解读：DFT只针对有限长序列或周期序列。

DFT相当于对DTFT中的正变换加以采样，造成时域信号的周期性，因此时域信号应限制在一个周期内。凡是用到离散傅里叶变换的时候，有限长序列都是作为周期序列的一个周期来表示的，都隐含有周期性意义。

2.3 时频分析与傅里叶变换

离散傅里叶变换的矩阵形式

定义Fourier矩阵：

$$\mathbf{F}=\begin{bmatrix}1 & 1 & 1 & \dots & 1\\ 1 & W_N & W_N^2 & \dots & W_N^{N-1}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \dots & W_N^{(N-1)(N-1)}\end{bmatrix}$$

$$W_N=e^{-j\frac{2\pi}{N}}$$

Fourier矩阵的性质： $\mathbf{F}^H\mathbf{F}=\mathbf{F}\mathbf{F}^H=N\mathbf{I} \Rightarrow \mathbf{F}^{-1}=\frac{1}{N}\mathbf{F}^H$

$$\mathbf{X}(k)=\mathbf{F}\mathbf{x}(n)$$

$$\mathbf{x}(n)=\mathbf{F}^{-1}\mathbf{X}(k)$$

傅里叶变换的四种形式

	时间函数	正变换	反变换	频率函数
Fourier Transform	连续 非周期	$X(j\Omega)=\int_{-\infty}^{\infty}x(t)e^{-j\Omega t}dt$	$x(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}X(j\Omega)e^{j\Omega t}d\Omega$	非周期 连续
Fourier Series	连续 周期	$X(jk\Omega_0)=\frac{1}{T_0}\int_{-T_0/2}^{T_0/2}x(t)e^{-jk\Omega_0 t}dt$	$x(t)=\sum_{k=-\infty}^{\infty}X(jk\Omega_0)e^{jk\Omega_0 t}$	非周期 离散
DTFT	离散 非周期	$X(e^{j\omega})=\sum_{n=-\infty}^{\infty}x(n)e^{-j\omega n}$	$x(n)=\frac{1}{2\pi}\int_{-\pi}^{\pi}X(e^{j\omega})e^{j\omega n}d\omega$	周期 连续
DFT	离散 周期	$X(k)=\sum_{n=0}^{N-1}x(n)W_N^{nk}$	$x(n)=\frac{1}{N}\sum_{k=0}^{N-1}X(k)W_n^{-nk}$	周期 离散

$$\Omega_0=\frac{2\pi}{T_0}$$

离散傅里叶变换的几个问题

频谱泄漏

频谱泄漏是指由于信号截断造成的原始信号频谱扩散现象。

产生频谱泄漏的原因是对信号的截断。信号的截断相当于在原始信号 $x(n)$ 与一个窗函数 $w(n)$ 相乘，在频域中相当于各自频谱的卷积过程。卷积的结果造成原始信号频谱的“扩散”（或拖尾、变宽），这就是频谱泄漏。

思考题1：请比较两种窗函数：矩形窗和汉明窗。在对语音信号进行分帧加窗的过程中，哪一种窗函数更好？请说明原因。

$$w_{rec}(n) = R_N(n)$$

$$w_{hamming}(n) = \left[0.54 - 0.46cos\left(\frac{2\pi n}{N-1}\right)\right] R_N(n)$$

离散傅里叶变换的几个问题

栅栏效应

因为DFT计算频谱只限制在离散点上的频谱，也就是 F_0 的整数倍处的谱，而无法看到连续频谱函数，这就像通过一个“栅栏”观看景象一样，只能在离散点的地方看到真实景象。这种现象称为“栅栏效应”。

减小栅栏效应的方法就是要是频域抽样更密，即增加频域抽样点数，就好像距离“栅栏”的距离边远一些。在不改变时域信号的情况下，必然是在时域信号末端补零。补零后的时域数据，在频谱中的谱线更密，原来看不到的谱分量就有可能看到了。

离散傅里叶变换的几个问题

语音信号DFT的共轭对称性

时域中的语音信号，经过离散傅里叶变换DFT后的频谱是共轭对称的。

可通过DFT的公式证明，请大家思考其中的原因，并解答下面的思考题。

思考题2：对于一个长度为N的实数序列$x(n)$，离散傅里叶变换为$X(k)$。

(1) 请写出$X(0)$与$X\left(\frac{N}{2}\right)$的值的表达式。

(2) 请证明$X(k)$的共轭对称性，即：

$$X(n-i)=X^*(i)\quad 0<i<\frac{N}{2}$$

本章回顾

- 2.1 数字信号及其基本运算
- 2.2 采样定理
- 2.3 时频分析与傅里叶变换

2.4 作业

一、思考题

思考题1：请比较两种窗函数：矩形窗和汉明窗。在对语音信号进行分帧加窗的过程中，哪一种窗函数更好？请说明原因。

$$w_{rec}(n) = R_N(n)$$

$$w_{hamming}(n) = \left[0.54 - 0.46cos\left(\frac{2\pi n}{N-1}\right)\right] R_N(n)$$

思考题2：对于一个长度为N的实数序列 $x(n)$，离散傅里叶变换为 $X(k)$。

(1) 请写出 $X(0)$与 $X\left(\frac{N}{2}\right)$的值的表达式；

(2) 请证明$X(k)$的共轭对称性，即：

$$X(n-i) = X^*(i) \quad 0 < i < \frac{N}{2}$$

2.4 作业

二、计算题

给定两个有限长序列：

$$x_1(n)=\{1,\ 5,\ 7,\ 3,\ 2,\ 1,\ 6,\ 9\}\quad 0\leqslant n\leqslant 7\quad N_1=8$$

$$x_2(n)=\{2,\ 4,\ 6,\ 1,\ 3\}\quad 0\leqslant n\leqslant 4\quad N_2=5$$

分别计算 $x_1(n)$ 和 $x_2(n)$ 的线性卷积和圆周卷积（点数为8）结果，观察二者在哪些位置具有相同的值。

感谢聆听！

Thanks for Listening