

红外图像处理技术 之红外光谱信号处理(上)

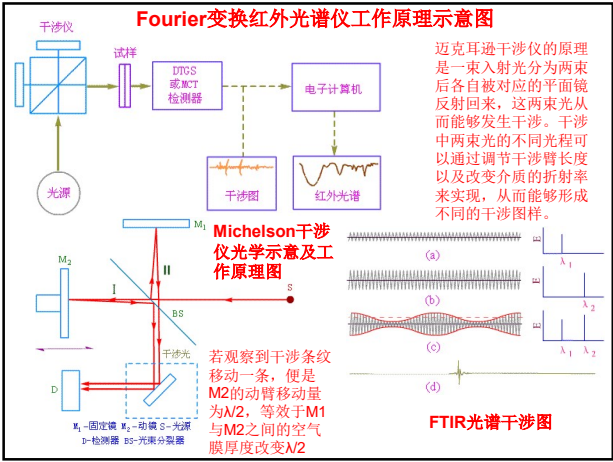
航空航天学院
陆哲明/郑阳明

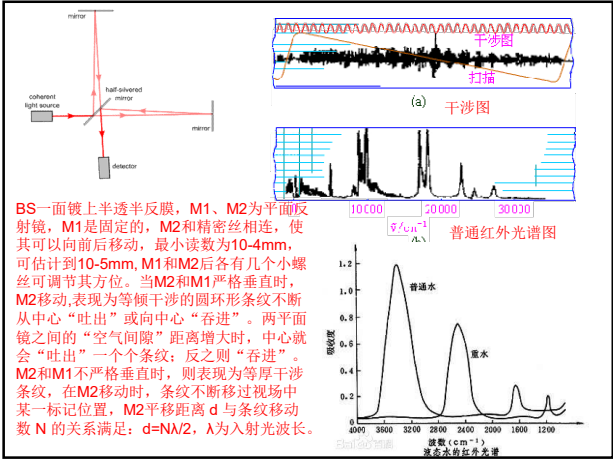
引言

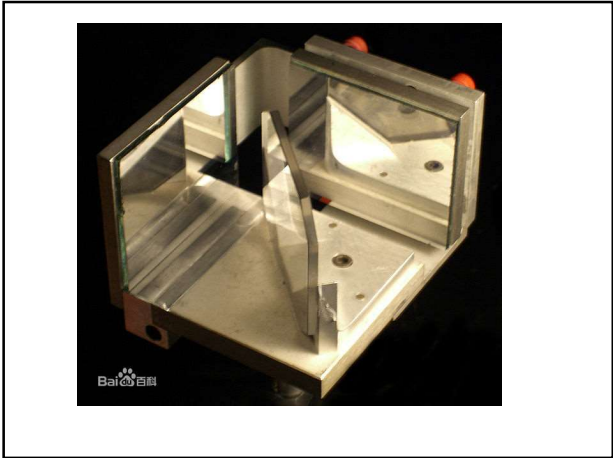
- 红外光谱信号处理包括**信号采集**、**信号转换**（即模数转换）、**信号的再生与复原**。
- 信号采集由**光电检测器**完成，通过检测器获得的电信号是一个时间函数，绘出的是**干涉图**。
- 对于这种包含有光谱信息的**时域干涉图**，很难直接进行光谱分析，因而需要通过模数转换进行转换后以数字信号的形式输入计算机，再经计算机和快速计算完成信号的再生与复原，最终获得以波数(或波长)为横坐标的**频域光谱图**，即人们分析时所用的一般光谱图。

• 红外光谱定义：

- 当一束具有连续波长的红外光通过物质，物质分子中某个基团的振动频率或转动频率和红外光的频率一样时，分子就吸收能量由原来的基态振(转)动能级跃迁到能量较高的振(转)动能级，分子吸收红外辐射后发生振动和转动能级的跃迁，该处波长的光就被物质吸收。
- 红外光谱法实质上是一种根据分子内部原子间的相对振动和分子转动等信息来确定物质分子结构和鉴别化合物的分析方法。
- 将分子吸收红外光的情况用仪器记录下来，就得到红外光谱图。红外光谱图通常用波长(λ)或**波数**(σ)为横坐标，表示吸收峰的位置，用**透光率**($T\%$)或者**吸光度**(A)为纵坐标，表示吸收强度。







内容

- 8.1信号采集
- 8.2数字信号处理预备知识
- 8.3模拟信号和数字信号的傅立叶变换
- 8.4信号采样与复原
- 8.5常用光谱数据处理
- 8.6回归分析



傅里叶 红外光谱仪

8.1 信号采集

- 红外光谱分析的信号特征不仅决定于**红外辐射的物理特征**，也决定了近红外光谱仪器应用的**光源、检测器与光学材料**。
- 常用的测光材料是半导体材料，如Si、PbS(硫化铅)、InSb(碲化铟)、InAs(砷化铟)等。近红外光谱区的检测效率很高，但温度引起的热噪声极大，因此此类检测器必须在**液氮**下工作。



采集到的红外模拟信号

- 红外光谱检测器的特性参数

✓ 响应度

□ 响应度（或称为灵敏度）是光电检测器输出信号与输入辐射功率之间关系的度量。

□ 它描述的是光电检测器件的光-电转换效能，定义为光电检测器输出电压 V_0 或输出电流 I_0 与**入射光**功率 P （或通量 ϕ ）之比，即：

$$\begin{cases} S_V = \frac{V_0}{P_i} \\ S_I = \frac{I_0}{P_i} \end{cases}$$

□ 式中， S_V 为电压响应度； S_I 为电流响应度。

□ 由于光电检测器的响应度随入射光的波长而变化，因此又有**光谱响应度**和**积分响应度**之分。

✓ 光谱响应度

- 光谱响应度 $S(\lambda)$ 是光电检测器的输出电压或输出电流与入射到检测器上的单色辐通量（光通量）之比。

$$\begin{cases} S(\lambda) = \frac{V_0}{\phi(\lambda)} \\ S(\lambda) = \frac{I_0}{\phi(\lambda)} \end{cases}$$

- 式中，如果 $\phi(\lambda)$ 为光通量，则 $S(\lambda)$ 的单位为V/lm； $\phi(\lambda)$ 为入射单色辐通量时， $S(\lambda)$ 的单位为V/W。
- 光谱响应度是表述入射的单色通量或光通量所产生的检测器和输出电压（或电流），它的值越大意味着检测器越灵敏。

✓ 积分响应度

- 积分响应度表示检测器对连续辐射通量的反应程度。
- 对包含有各种波长的辐射光源，总光通量为：

$$\phi = \int_0^{\infty} \phi_{\lambda} d\lambda$$

- 光电检测器输出的电流或电压与入射总光通量之比称为积分响应度。由于光电检测器输出的光电流是由不同的波长的光辐射引起的，所以输出光电流应为：

$$I_0 = \int_{\lambda_1}^{\lambda_0} d\lambda = \int_{\lambda_1}^{\lambda_0} S_{\lambda} \phi_{\lambda} d\lambda$$

- 可得积分响应度为
- $$S = \frac{\int_{\lambda_1}^{\lambda_0} S_{\lambda} \phi_{\lambda} d\lambda}{\int_0^{\infty} \phi_{\lambda} d\lambda}$$

- 式中， λ_0 为光电检测器的长波限； λ_1 为光电检测器的短波限。

✓ 响应时间

- 当入射辐射到光电检测器后或入射辐射遮断后，光电检测器的输出上升到稳定值或下降到照射前的值所需时间为响应时间。
- 把从10%上升到90%峰值处所需的时间称为检测器的上升时间，而把从90%下降到10%处所需的时间称为下降时间。

✓ 频率响应

- 由于光电检测器信号的产生和消失存在着一个滞后过程。所以入射光辐射的频率对光电检测器的响应将会有较大的影响。光电检测器的响应随入射辐射的调制频率而变化的特性称为频率响应，利用时间常数可得到光电检测器响应度与入射调制频率关系：

$$s(f) = \frac{S_0}{[1 + (2\pi f \tau)^2]^{1/2}}$$

- 式中， $s(f)$ 为频率，是 f 时的响应度； S_0 为频率，是零时的响应度； τ 为时间常数(等于 RC)。当 $s(f)/S_0 = 0.707$ 时，可得放大器的上限截止频率。显然，时间常数决定了光电检测器频率响应的带宽。

• 红外光谱检测器的工作方式

✓ 按照工作方式可分为**单通道**和**多通道**两种类型。

□ 单通道检测器只有一个检测单元，一次只能接收一个光信号，得到全部光谱需要经过光谱扫描；

□ 多通道检测器可同时接收检测面上不同波长的光信号，不需扫描，速度很快。

□ 多通道检测器类型主要有二极管阵列(PDA)和电荷耦合器件，其材料决定工作波长范围，检测面上的阵列数目决定仪器的波长分辨率，常见的阵列有 256×1 、 512×1 、 1024×1 、 2048×1 ，阵列数目越大分辨率越高，但价格也越贵，因此选用多通道检测器阵列的数目要与光栅的分光性能相匹配。

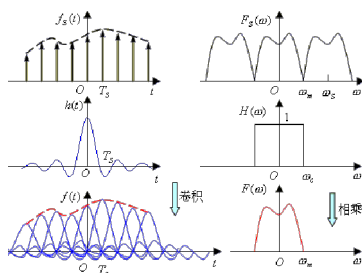
• 光电检测信号的模数转换

✓ 通过光电检测器获得的红外光与样品作用后的**模拟量电信号**，是一**时间函数**。必须变换为适于数字处理的形式（二进制数码），送入计算机、数字存储设备、数据传输设备处理和存储，才能进一步进行计算、分析和处理。完成这一模拟量电信号到离散数字信号转换的器件叫模数转换器，简称ADC或A/D转换器。

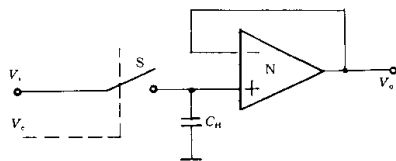
✓ 在A/D转换器中，因为输入的模拟信号在时间上是连续量，而输出的数字信号代码是离散量，所以进行转换时必须在一系列选定的瞬间（亦即时间坐标轴上的一些规定点上）对输入的模拟信号取样，然后再把这些取样值转换为输出的数字量。因此，一般的A/D转换过程是通过**取样**、**保持**、**量化**和**编码**这四个步骤完成的。在实际电路中，这些过程是并行进行的，例如，取样和保持、量化和编码往往都是在转换过程中同时实现的。

✓ 取样和保持

□ 取样随时间连续变化的模拟量。通过分析可看到，取样信号 $S(t)$ 的频率越高，所取得信号经低通滤波器后越能真实的复现输入信号。但带来的问题是数据量增大，为保证有合适的取样频率，它必须满足**取样定理**。



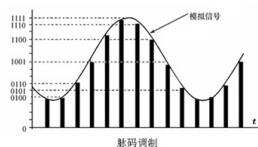
- 将取样电路每次取样的模拟信号转换为数字信号都需要一定时间，为了给后续的量化编码过程提供一个稳定值，每次取得的模拟信号必须通过保持电路**保持**一段时间。



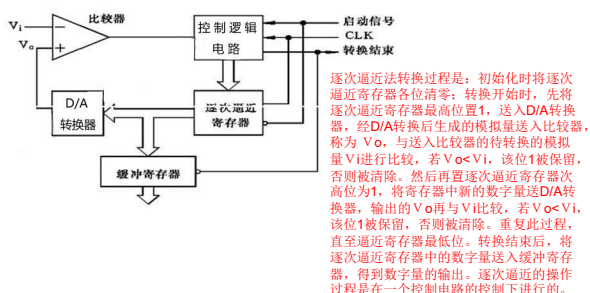
在输入逻辑电平控制下出于“采样”或“保持”两种工作状态。“采样”状态下电路的输出跟踪输入模拟信号，在“保持”状态下电路的输出保持前次采样结束时刻的瞬时输入模拟信号，直至进入下一次采样状态为止

✓ 量化和编码

- 数字信号不仅在时间上是离散的，而且在幅值上也是不连续的。任何一个数字量的大小只能是某个规定的最小数量单位的整数倍。
- 为了将模拟信号转换为数字量，在A/D转换过程中，还必须将取样-保持电路的输出电压，按某种近似的方式规划到相应的离散电平上，这一转化过程称为数值量化，简称**量化**。
- 量化是将模拟信号量程分成许多离散量级，并确定输入信号所属的量级。然后对每一量级分配唯一的数字码，并确定与输入信号相对应的代码。最普遍的码是二进制，它由 2^n （ n 为位数）个量级，可依次逐个**编号**。

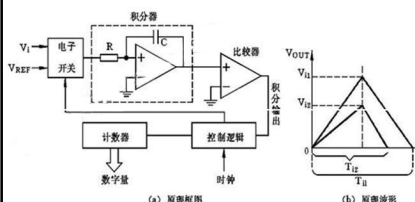


- A/D转换器的种类很多，按其工作原理不同分为直接A/D转换器和间接A/D转换器两类。
- 直接A/D转换器可将模拟信号直接转换为数字信号，这类A/D转换器具有较快的转换速度，其典型电路有并行比较型A/D转换器、逐次比较型A/D转换器。



□ 而间接A/D转换器则是先将模拟信号转换成某一中间电量（时间或频率），然后再将中间电量转换为数字量输出。此类A/D转换器的速度较慢，典型电路是双积分型A/D转换器、电压频率转换型A/D转换器。

□ 这样，通过光电检测器获得的红外光与样品作用后的**模拟量电信号**，经过数模转换就变成了计算机能够识别、计算和分析处理的**数字信号**。



双积分法A/D转换的过程是：先将开关接通待转换的模拟量 V_i ， V_i 采样输入到积分器，积分器从零开始进行固定时间 T_1 的正向积分。时间 T_1 到后，开关再接通与 V_i 极性相反的基准电压 V_{REF} ，将 V_{REF} 输入到积分器，进行反向积分，直到输出为0V时停止积分。 V_i 越大，积分器输出电压越大，反向积分时间也越长。计数器在反向积分时间内所计的数值，就是输入模拟电压 V_i 所对应的数字量，实现了A/D转换。

8.2 数字信号处理预备知识

• 数字信号

- ✓ 物理信号：温度、压力、流量、光等
- ✓ 电流信号&电压信号
- ✓ 时域信号&空域信号
- ✓ 实信号&复信号
- ✓ 连续时间信号：模拟信号
- ✓ 离散时间信号：离散时间序列
- ✓ 通过光电检测器获得的红外光与样品作用后的连续模拟量电信号，经过模数转换后就变成了计算机能够识别、计算和分析处理的数字信号。

• 数字信号处理

- ✓ 信号处理的方法有：**变换、平滑、卷积和积分**。
- ✓ 目前一般将连续模拟信号转换为数字信号存储在计算机内，这些信息的提取则随后通过数字法来实现。如卡尔曼滤波、傅里叶变换、小波变换、主成份分析等方法可以把信号和噪声背景分离，达到提高分析信号质量的目的。
- ✓ 傅里叶分析包含了连续信号和离散信号的傅里叶变换和傅里叶级数，内容丰富。

8.3 模拟信号和数字信号的傅立叶变换

• 傅立叶变换光谱学的基本方程

$$F(x) = \int_{-\infty}^{\infty} A(\bar{v}) e^{j2\pi f\bar{x}} d\bar{v}$$

$$A(\bar{v}) = \int_{-\infty}^{\infty} F(x) e^{-j2\pi f\bar{x}} dx$$

✓ $F(x)$ 和 $A(\bar{v})$ 互为傅里叶变换和逆变换。适用条件是：

- 函数 $F(x)$ 处处有界，并且仅存在有限个第一类不连续点；
- 除了有限个点，函数 $F(x)$ 和 $A(\bar{v})$ 的微分处处存在；
- $\int_{-\infty}^{\infty} |F(x)| dx < \infty$ ，即 $F(x)$ 是绝对可积函数；
- 不难看出，物理上所研究的光谱图和实验测量获得的干涉图都满足这些条件，它们的傅里叶积分变换是可行的。

• 模拟信号的傅里叶变换

✓ 设 $x(t)$ 为一连续时间信号，若 $x(t)$ 属于 L_1 空间，即满足：

$$\int_{-\infty}^{\infty} |x(t)| dt < \infty$$

✓ 那么 $x(t)$ 的傅里叶变换存在，并定义为

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

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

✓ 式中， $\omega = 2\pi f$ 为角频率，单位为 rad/s； $X(j\omega)$ 是 ω 的连续函数，称为信号 $x(t)$ 的频谱密度函数，或简称频谱； $x(t)$ 称为时(空)域函数； $X(j\omega)$ 称为频率域函数，即时(空)域函数与频率域函数可通过一个变换来统一。

✓ 不具备傅里叶变换条件的周期信号，在引入了冲激信号后也可以作傅里叶变换。当然，变换结果应从冲激信号角度理解。

✓ 时域连续周期信号的傅里叶变换在频率上是离散非周期的。时域连续非周期信号的傅里叶变换在频域上是连续非周期的。

✓ 一般地 $X(j\omega)$ 即是常见的图谱， $x(t)$ 为波形图(干涉图)。在傅里叶型仪器中，直接采集到的是 $x(t)$ 信号，然后通过快速傅里叶变换，而得到图谱 $X(j\omega)$ 。傅里叶变换在数据处理中应用的主要目的是加快提取信息的过程，通过压缩数据使信息提取方便有效，同时可去掉干扰信号和噪声。

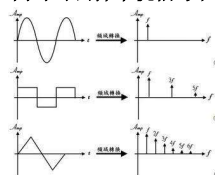


图 8-1 (a) 正弦波周期信号 (b) 三角波周期信号 (c) 三角波非周期信号

• 数字信号的傅立叶变换

✓ 通常光谱分析中通过信号采集获得的 $x(t)$ 是有限长度，等间隔数据点构成的序列，即离散时间信号，它必须通过离散傅里叶变换才能进行频谱分析。在讨论离散傅里叶变换之前，先讨论离散时间信号的傅里叶变换和离散时间周期信号的傅里叶级数。

✓ 离散时间傅里叶变换

□ 离散时间傅里叶变换即是离散时间信号从时域变换到频域和从频域变换到时域的一对线性变换。设 $h(n)$ 为一线性平移不变(LSI)系统的单位抽样响应，可以求出该系统的频率响应：

$$H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h(n)e^{-j\omega n}$$

□ 此式即为离散时间序列的傅里叶变换，即DTFT。

$$H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h(n)e^{-j\omega n}$$

□ $H(e^{j\omega})$ 是 ω 的连续周期函数，周期为 2π 。上式也可以看作是周期信号 $H(e^{j\omega})$ 在频域内展成的傅里叶级数，其傅里叶系数是时域信号 $h(n)$ 。

✓ 离散时间周期信号的傅里叶级数

□ 设 $\tilde{x}(nT_s)$ 是周期信号 $\tilde{x}(t)$ 的抽样， $\tilde{x}(t)$ 的周期为 T ，每个周期内抽 N 个点，即 $T = NT_s$ 。这样， $\tilde{x}(nT_s)$ 也是周期的，周期为 NT_s 或 N 。将 $\tilde{x}(t)$ 展开成傅里叶级数，得

$$\tilde{x}(t) = \sum_{k=-\infty}^{\infty} \tilde{X}(k\Omega_0)e^{jk\Omega_0 t} \quad k = 0, \pm 1, \dots, \pm \infty$$

□ $\tilde{X}(k\Omega_0)$ 是 $\tilde{x}(t)$ 的傅里叶系数，所以它是离散的，

$$\Omega_0 = 2\pi / T = 2\pi / (NT_s)$$

□ 现对上式抽样得：

$$\begin{aligned} \tilde{x}(nT_s) &= \tilde{x}(t)|_{t=nT_s} = \sum_{k=-\infty}^{\infty} \tilde{X}(k\Omega_0)e^{jk\Omega_0 nT_s} \\ &= \sum_{k=-\infty}^{\infty} \tilde{X}(k\Omega_0)e^{jk\frac{2\pi}{NT_s}nT_s} = \sum_{k=-\infty}^{\infty} \tilde{X}(k\Omega_0)e^{j\frac{2\pi}{N}nk} \end{aligned}$$

□ 由离散时间的傅立叶变换过程可知离散信号的频谱是周期的，周期为 Ω_s ，所以 $\tilde{X}(k\Omega_0)$ 应是周期的。由于

$$\Omega_s = 2\pi / T_s = 2\pi N / NT_s = N\Omega_0$$

□ Ω_s 是 $\tilde{x}(t)$ 的基波频率。也就是说，在 $\tilde{X}(k\Omega_0)$ 的一个周期内，包含了基波、二次谐波、直至 N 次谐波。所以 $\tilde{X}(k\Omega_0)$ 应是 N 点的周期序列。取一个周期，并简记为 $X(k)$ ，又由于

$$e^{j\frac{2\pi}{N}nk} = e^{j\frac{2\pi}{N}n(k+IN)}$$

□ I 为任意整数，所以上式的求和只在一个周期内进行，即

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

□ 在该式中, 当 $n=0,1,\dots,N-1$ 和 $n=N,\dots,2N-1$ 时, 所求出的结果是一样的, 也即该式只能计算出 N 个 $\tilde{x}(NT_s)$ 的值。这 N 个值即是 $\tilde{x}(NT_s)$ 的一个周期, 记为 $x(n), n=0,1,\dots,N-1$ 。这样上式左边可以换成 $x(NT_s)$

□ 对上式两边作如下运算:

$$\sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}ln} = \sum_{n=0}^{N-1} \left[\sum_{k=0}^{N-1} X(k) e^{j\frac{2\pi}{N}nk} \right] e^{-j\frac{2\pi}{N}nl} = \sum_{n=0}^{N-1} X(k) \sum_{k=0}^{N-1} e^{j\frac{2\pi}{N}(k-l)n}$$

□ 由于 $\sum_{k=0}^{N-1} e^{j\frac{2\pi}{N}(k-l)n} = \begin{cases} N & k-l=0, N, 2N, \dots \\ 0 & \text{others} \end{cases}$

□ 故上式的右边等于 $NX(k)$, 于是有:

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

□ 对离散周期信号可得到如下变换公式:

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

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

□ 上式称为离散周期序列的傅里叶级数 (DFS), 尽管式中标注的 n, k 都是从 $-\infty$ 至 $+\infty$, 实际上只能算出 N 个独立的值。DFS在时域、频域都是周期的, 且是离散的。

✓ 离散傅里叶变换

□ 正如前面提到的, 通常光谱分析中通过信号采集获得的 $x(t)$ 是有限长度的, 等间隔数据点 $x(n)$ 构成的序列, 即数字信号, 也称离散信号, 它是时域和频域上都离散信号。要想在计算机上实现信号的频谱分析及其他方面的处理工作, 那么对信号的要求也是要在时域和频域都应是离散的, 且都应是有限长。

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

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

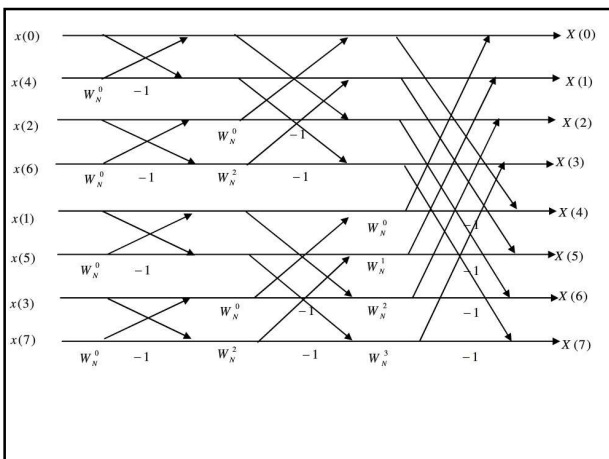
- DFT对应的是在时域、频域都是有限长，且又都是离散的。
- 显然DFT并不是一个新的傅里叶变换形式，它实际上来自于DFS，只不过在时域、频域各取一个周期而已。这样，对任意有限长序列 $x(n)$ ，都可按上式方便地求频谱。但需要注意：只要使用上式，不管 $x(n)$ 本身是否来自周期序列，都应把它看作某一周期序列的一个周期。不过，求出一段 $X(k)$ ，需要 N 次复数乘法，求出 N 点 $X(k)$ ，需要 N^2 次复数乘法。当 N 很大时，其计算量是相当大的。

✓ 快速离散傅里叶变换

- 1965年Cooley和Turky发表了快速傅立叶变换算法，使 N 点DFT的计算量下降为

$$\frac{N}{2} \log_2 N$$

- DFT和卷积是信号处理中两个最基本也是常用的运算，它们涉及信号与系统的分析与综合这一广泛的信号处理领域。卷积可化为DFT来实现，实际上其他许多算法，如相关、滤波、谱估计等也都可化为DFT来实现。



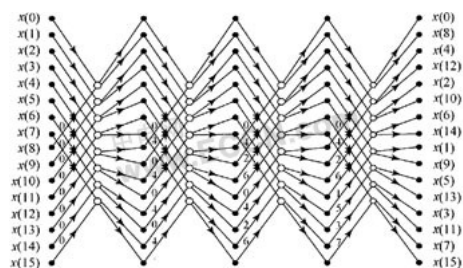
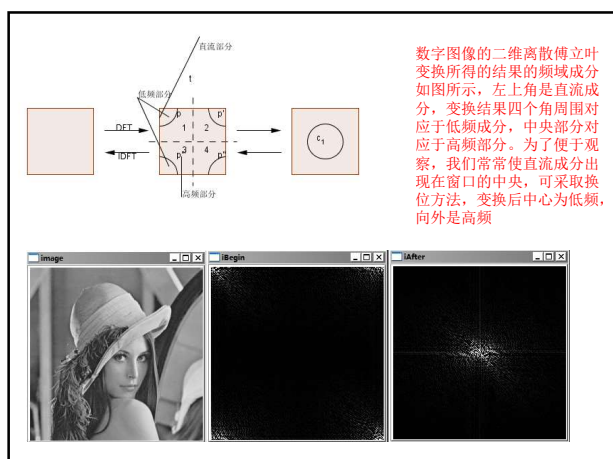


图1 固定几何结构 FFT 运算流程图

• 傅里叶变换在光谱处理中的应用

- ✓ 傅立叶变换在光谱信号处理中的作用主要有两方面：压缩数据/特征提取和滤波处理。通常 $x(n)$ 信号的数量很大，如近红外光谱，一般有成千上万个数据点，这个数据处理增加了很大的计算量，但是如果把 $x(t_n)$ 变换为 $X(k)$ 来处理，数据点可以减少很多。因此光谱的主要信息集中在 $X(k)$ 的前若干个值上，这些数据点代表了原光谱的主要信息特征，因此采用傅立叶变换来进行光谱分析计算将显著降低计算量。
- ✓ 在傅里叶变换滤波方面的应用很多，可以通过选用不同宽度的脉冲函数来卷积，得到的光谱分辨率是不同的，而高频随机噪声可通过选用较窄的脉冲函数来卷积而被滤掉，当然脉冲函数宽度的选取要依靠实际经验而定。

- ✓ 实际上，时间域的脉冲函数与频率域sinc函数是对应的，在时间域用方波函数来卷积也就相当于频率域乘以sinc函数，而且时间域脉冲函数越宽频率域的sinc函数越窄，反之亦然。sinc函数是一个窗口平滑函数，它的宽窄影响着滤波的效果。通过高频滤波可以去掉高频噪声，而低频滤波器则可以扣除漂移。在实际算法中，一般采用离散快速傅里叶变换算法。



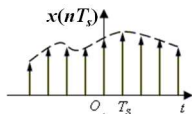
8.4 信号采样与复原

• 信号采样

- ✓ 将连续信号变成数字信号是在计算机上实现信号数字处理的必要步骤。在实际工作中，信号的采样是通过A/D芯片来实现的。通过A/D，将连续信号 $x(t)$ 变成数字信号 $x(nT_s)$ ， $x(t)$ 的傅里叶变换 $X(j\Omega)$ 变成 $X(e^{j\omega})$

✓ 问题：

- $x(nT_s)$ 是否包含有 $x(t)$ 的全部信息？
- $X(e^{j\omega})$ 与 $X(j\Omega)$ 是什么关系？



$$X(e^{j\omega}) = X_s(j\Omega) \big|_{\Omega=\omega/T_s} = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} X_a(j\Omega - jk\Omega_s)$$

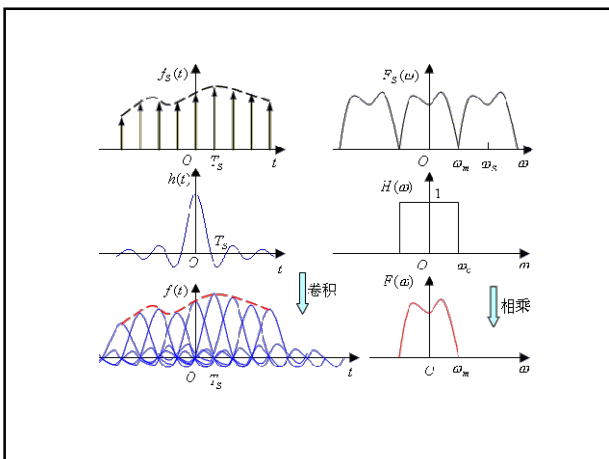
- 如何由 $x(nT_s)$ 恢复出 $x(t)$ ？

- ✓ 将连续信号 $x_a(t)$ 经抽样变成 $x(nT_s)$ 后， $x(nT_s)$ 的频谱变成周期的。相对频率 Ω_s ，周期为 $\Omega_s = 2\pi / T_s = 2\pi f_s$ ，相对圆频率 ω ，周期为 2π 。变为周期的方法是将 $X_a(j\Omega)$ 在频率轴上以 Ω_s 为周期作移位后再叠加，并除以 T_s 。这种现象又称为频谱的周期延拓。
- ✓ 若在 $|\Omega| \geq \Omega_s / 2$ 时 $|X_a(j\Omega)| \equiv 0$ ，也即 $X_a(j\Omega)$ 是有限带宽的，那么作周期延拓后， $X_s(j\Omega)$ 的每一个周期等于 $X_a(j\Omega)$ 。（差一定标因子 $1/T_s$ ）。反之，若 T_s 过大，或者 $X_a(j\Omega)$ 本身就不是有限带宽的，那么做周期延拓后将发生频域的“混叠”现象，以致一个周期中的 $X_s(j\Omega)$ 不等于 $X_a(j\Omega)$ 。由此产生的结果是无法由 $x(nT_s)$ 来恢复出 $x_a(t)$ 。
- ✓ 抽样定理：若连续信号 $x(t)$ 是有限带宽的，其频谱的最高频率为 f_c ，对 $x(t)$ 抽样是，若保证抽样频率：
$$f_s \geq 2f_c$$
- ✓ 那么可由 $x(nT_s)$ 恢复出 $x(t)$ ，即 $x(nT_s)$ 保留了 $x(t)$ 的全部信息。

• 信号复原

$$x_a(t) = \sum_{n=-\infty}^{\infty} x(nT_s) \frac{\sin[\pi(t-nT_s)/T_s]}{\pi(t-nT_s)/T_s}$$

- ✓ 此式为由抽样后的离散信号重建原信号的公式。不难发现，这是一个插值公式，插值函数为sinc函数，插值间距为 T_s ，权重为 $x(nT_s)$ 。只要满足抽样定理，那么由无穷多加权sinc函数移位后的和即可重建出原信号。
- ✓ 在实际中，将离散信号变成模拟信号可以通过模/数转换D/A来实现。D/A把二进制数字量电信号转换为与其数值成正比的模拟量电信号。在D/A参数中一个组重要的参数就是分辨率，它是指输入数字量发生单位数码变化时，所对应输出模拟量(电压或电流)的变化量，从而实现实时在线检测和控制等。



8.5 常用光谱数据处理

• 光谱数据采集参数的选择与作用

✓ 采样间隔

□ 光谱仪测定干涉图时不可能对**光程**进行连续采样，它只能进行**时间域**的连续采样，干涉图理论上是时间域的连续信号，但这就需用采样间隔参数来确定采集数据点的间距，干涉图就用这些数字化点表示。

□ **采样间隔**是指连续记录的两个光谱信号间的波长差。很明显，间隔越小，采集的数据点越多，样品信息越丰富，所得到的干涉图与真实情况吻合越好，但光谱存储空间也越大；间隔过大则可能丢失一部分真实信息而使图像失真，比较合适的数据采样间隔设计应当小于仪器的分辨率。

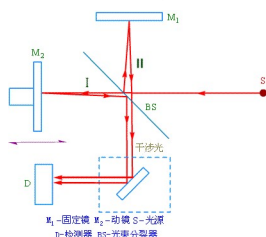
- Nyquist准则：任何简谐函数波形必须至少以其基频宽的二倍频率进行采样，所采集的数组才能真实地代表原来图形。
- 对于高频波的红外图谱测试，采样间隔应小一些；对于低频波的情况，因光谱范围的波长长，采样间隔可大一些。
- 测定中红外区的光谱，采样间隔可取隔一个过零点；在近红外测定中，采样间隔要小些；在远红外测定中，采样间隔要加大。
- ✓ 采集干涉图的数据点数
- 采集数据点数的多少与所得图谱的分辨率有关，在采样间隔一定的情况下，采集的数据点越多，也就是动镜移动的距离越长，所得图谱分辨率越高。光谱分辨率与动镜移动距离成倒数关系。

✓ 傅里叶变换数据点数

- 假定采集干涉图的数据点为 N ，在对这个干涉图进行快速傅立叶变换FFT，其光谱数组中只有 $N/2$ 个数据点是实数点，如直接用它进行运算，所得光谱由于各点直接相连形成“台阶”而得不到平滑的光谱。
- 可以采用充零方法改善图谱的精度，即在采集到干涉图后，在进行FFT以前，将 N 个零充填到已采集的 N 个数据点后面，然后再进行FFT运算。这时傅里叶变换的数据点数为 $2N$ ，进行FFT的时间要加倍，所得的光谱中在原来数据点之间加一个数据点，因而能得到表现光滑的光谱。
- 如果在 N 个数据点之后填充入 $3N$ 个零，则FFT的数据点数为 $4N$ 个，FFT时间为原来的4倍，在所得光谱中数据点间距为原来的 $1/4$ ，因而所得光谱更加平滑。
- 但由于真正决定光谱的数据点数并没有增加，因而光谱的分辨率并未改变，充零只是改善由于干涉图采样不连续而破坏的图形。

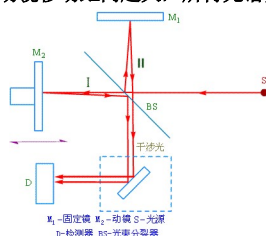
✓ 干涉图峰的位置

- 在零光程差处的峰就是干涉图峰。
- 在实际测量干涉图时，在零光程差左右二边采样，二边采样的数据点可以相同，也可以不同。
- 但峰前至少需采集一定数量的数据点，因此需要一个参数来确定以采集数据点为基准的干涉图峰的位置。



✓ 动镜移动速度和距离

- 动镜移动速度与所用检测器种类有关。用于光声光谱、衰减全反射或其他低光通量附件测量时的动镜移动速度要慢一些；MCT检测器的响应速度大于TGS检测器，因此用MCT检测器时，动镜移动速度要比采用TGS的快一些。
- 一些中高档的FTNIR光谱仪可以改变动镜移动距离，在一定的采样间隔下，可以改变采集的数据点数，从而获得不同分辨率的谱图。动镜移动距离越大，所得光谱分辨率越高。



✓ 扫描次数

- 噪声来源很多，完全消除是不可能的，只能采用各种方法降低噪声以提高信噪比。数据点信号累加平均法与平滑技术可提高信噪比。
- 吸收信号是样品对红外光吸收所产生的，它在每次采样中均为正值，噪声是随机的，有正有负。如果对一样品进行多次扫描，将每次测得的信号用计算机累加起来取平均值，由于样品的信号都是正值，平均后其数值不变，而噪声信号有正有负，累加平均能使部分噪声相互抵消。
- 因此，增加扫描次数取数据点信号累加平均值的总效果可使图谱的信噪比得到改善，这是一种提高信噪比的方法。

✓ 扫描延迟

- 在红外光谱仪操作中，为了使机械系统的瞬间变化不致影响测定，需等待仪器系统平衡后再进行测试，一般设置扫描延迟参数使扫描延迟一定时间开始，本参数以时间(秒)为单位。
- 在改变光阑孔径，改变光路等参数时，都需要延迟一定时间以保证系统的稳定性。

✓ 光阑孔径

- 由于红外光谱仪是通过干涉图得到红外光谱的，因此理论上要求由光源发出的光必须平行投向干涉仪，如果光束发散就会产生光程差而引起光的干涉，会干扰测试结果。
- 实际上，有光源发射出的光总是发散的，为了减少干扰，许多仪器在光源与干涉仪中间加光阑，其孔径发散大；如果要得到高分辨率的光谱或使用高灵敏度检测器时，可选用小孔径光阑。

✓ 高通噪声滤波、低通噪声滤波

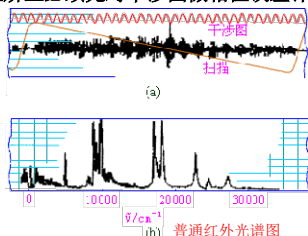
- 为了使采集的数据能正确代表真实的干涉图，除需要正确选定采样间隔外，还需要去频率过高或过低的辐射，使它们不能被检测器所检测，以免干扰所得峰形。
- 高频噪声滤波与低频噪声滤波就是选用光学滤波器去干扰辐射的参数，不同仪器选择不同的滤波频率。

✓ 放大器增益

- 采集到的信号如果太小，可以增加放大器增益使信号放大。加大增益使所有信号都加强，噪声也同时放大。
- 增益过大会使进入模数转换器的信号溢出以致导出错误信息，同时可能使电子系统饱和而使干涉图变形。
- 可以由显示的干涉图的峰-峰高度来粗略估算所用增益是否合适，50%~80%为合适，否则，应调整增益，增益取1，2，4，8...等数值。

✓ 相位校正

- 理论上的干涉图应该具有左右对称的峰形。实际测得的干涉图却具有不对称的峰形，这是由各方面原因引起的相位误差导致的图形畸变。
- 需要从实测所得的失真的光谱中消除由于相位误差引起的与真实光谱的差异，以得到原来的无畸变的真实谱，相位校正就是用来矫正这一误差。
- 进行相位校正必须先对干涉图做相位误差计算然后进行修正。



✓ 数据文件的尺寸

- 在计算机采集数据时，每个文件中除存放所采集的数据点外，还要存储一些与数据采集有关的资料如采集时间、标题、参数、简要的说明等。
- 因此每个数据文件大小必须要大于所采集的数据点数，至少要等于或大于数据点数的1.4倍。

✓ 变迹函数

- 测定干涉图时动镜的移动距离是有限的，因此需要通过数学处理来完成干涉图，用变迹函数乘以干涉图就能起这个作用，它还能抑制旁瓣负峰，改善谱线形状。常用的变迹函数有三角波函数、矩形波函数、Happ-Genzel函数，这三种变迹函数对光谱的影响略有差别。
- 采用三角波函数处理后的谱图旁波瓣显著降低，但分辨率比较低；
- 采用矩形波函数处理后的谱图旁波瓣高，分辨率比用三角波函数的好；
- 采用Happ-Genzel函数处理所得的谱图其旁波瓣很小，但分辨率最低。因此，压制旁波瓣是以损失一些分辨率为代价的，用户可根据需要选用变迹函数。
- 除了上述三种常用的变迹函数外，还可以应用梯形函数、高斯函数、指数函数等函数作为变迹函数。

• 光谱坐标变换

✓ 横坐标变换

□ 波数 σ 和波长 λ 是红外光谱横坐标常用的两种坐标单位，它们之间有如下关系：

$$\sigma(\text{cm}^{-1}) = \frac{1}{\lambda(\mu\text{m})} \times 10^4$$

□ 式中， σ 为以波数为单位的红外辐射频率（ cm^{-1} ）； λ 为以微米为单位的红外辐射波长（ μm ）。

✓ 纵坐标变换

□ 红外图谱纵坐标有透射率 T 和吸光度 A 两种常用的表示法，透射率 T 是红外光透过样品的光强与入射光强的比值。

$$T = I / I_0$$

□ 式中， I 为红外光的入射强度； I_0 为红外光的投射强度。在实际工作中常采用百分透射率 $T\%$ ，

$$T\% = (I / I_0)\%$$

□ 吸光度 A 是百分吸收率 I_0/I 的对数： $A = \lg(I_0 / I)$ ，透射率与吸光度之间的关系为： $A = \lg(1 / T)$

□ 以透射率表示的红外谱图是标准谱图普遍采用的标准格式，能直观地看出样品对红外光的吸收情况。

□ 以吸光度表示的谱图的优点是吸光度值 A 在一定范围内与样品厚度及样品浓度呈正比关系，根据朗伯-比尔定律

$$A = KCL$$

□ 式中 A 为样品在特定波数下的吸光度； C 为样品浓度(摩尔/升)； L 为样品厚度(厘米或毫米)； K 为比例常数(升/厘米·摩尔)。

□ 采用以吸光度为纵坐标标度的红外谱图适用于定量分析以应用定量原理进行运算的应用软件。采用公式 $A = \lg(1 / T)$ 编程的软件可以方便地使透射谱与吸收谱之间互相变换，这实际上是对数或反对数变换。

□ 在吸收度 A 过大时，谱图上相应区域的吸收峰呈平头峰或呈现空白区。

✓ 差谱

□ 差谱即差示光谱，是计算机对已存储的数字化光谱所做的一种重要的数据处理功能。

□ 它有两种处理方法，一种是透射率相除法；另一种是吸光度相减法，一般说的差谱软件是吸光度相减法。

✓ 加谱

□ 在混合物光谱中，某一波数处的总吸收光度是该体系中各组分在该处产生的吸收光度值的总和。

□ 因此对于混合组分的样品，当红外光通过样品时，混合物中每个组分吸收相应的红外光，在波数 σ 处总的吸光度是各组分吸光度之和，即吸光度有加和性。
$$A(\sigma) = \lg \frac{I_0(\sigma)}{I(\sigma)} = \sum_{i=1}^n K_i(\sigma) C_i L$$

□ 利用吸光度加和性就可以将2个或2个以上的光谱加和起来成为一张新光谱。相加的2个谱图如果是同一中样品的光谱，则加谱可以提高原谱的信噪比，提高谱图质量。相加的2个谱图如果是2种不同样品的光谱，加和后得到无分子作用力的混合物光谱，这种方法可用于人工制造定量标准曲线。

✓ 乘谱

□ 按照朗伯-比尔定律，任何一种红外光谱乘以 m 倍，所得乘谱的吸光度值与浓度间仍呈原谱同样的正比关系。

$$A = KCL$$

$$mA = mKCL = K(mC)L$$

□ 乘谱可使光谱纵坐标强度发生变化，计算机根据上述原理进行乘谱运算，谱图不会失真，可用于定量分析。但它不能提高信噪比。

✓ 导数光谱（微分光谱）

□ 红外光谱仪的倒数光谱是用计算机对光谱进行微分处理而得到的光谱，就是沿光谱曲线计算出每个数据点处的斜率，连成曲线就是导数光谱。

□ $\frac{dA}{d\lambda} - \lambda$ 曲线是一阶导数谱； $\frac{d^2A}{d\lambda^2} - \lambda$ 是二阶倒数谱，依次类推，可得到 n 阶倒数谱。

✓ 平滑技术

- 它是降低谱图噪声的一种方法，通过对数据作一定的数字均匀化计算，使误差重新分配，从而达到降低噪声的目的。
- 对于已经采集好数据存储于计算机中的光谱图，可用平滑技术降低谱图的噪声，尤其是降低高频噪声。
- 可用于平滑技术的数学方法很多，它们都采用各种算法以周围的数据点去平均每个数据点值，最常用的是萨维特斯基-戈莱法。
- 平滑级数对基线与吸收峰的影响不同，平滑级数越高，吸收峰失真越多。宽峰与窄峰所受到的影响也不一样，不同分辨率的谱图受相同级数平滑的影响也不同，扫描是采集的数据点数的多少会影响平滑效果。一般来讲，对于窄峰最好用较高分辨率采样，使谱图有较多的数据点。平滑级数不要过高，这样经平滑处理后的谱图失真较小。

✓ 峰值及峰谷位置的求取

- 峰值即吸收峰峰尖处的横坐标值，峰尖与峰谷的位置可通过多种数学方法计算出来。峰尖与峰谷处斜率变化为0，它们的位置可由一阶导数值来确定，计算机由左到右对谱图逐步求导就可得到吸收峰的峰尖与峰谷的数值。
- 用一阶导数求峰值的方法对于连续函数是精确的，但对于由离散数据点组成的谱图就会有误差。数据点越少，误差越大，即对于分辨率低的谱图，求出的峰值离真实值偏差大。计算机通过二次插值法能够更正确地找出峰值。

✓ 吸收峰积分强度的计算

- 红外光谱中吸收峰积分强度的测定是定量分析的一种常用数据处理方法，所用光谱必须为吸收谱。在 $\sigma_1 - \sigma_2$ 区间吸收峰的面积可用积分吸光度表示如下：

$$area\ of\ peak = \int_{\sigma_1}^{\sigma_2} A_{\sigma} d\sigma$$

✓ 基线校正

- 通过实验测到的红外谱图的基线有时会发生漂移、倾斜或弯曲，尤其在固体制样的谱图中常会遇到这种情况。
- 一个谱图可以看成是由许多间距很小的直线连接而成，其间距由进行FFT时的数据点决定。基线可通过一些数学方法进行矫正，例如根据下述方程可以矫正每个波数处的基线：
$$Y'(\sigma) = Y(\sigma) - (T + C \times \sigma)$$
- 式中， $Y(\sigma)$ 为原图中波数 σ 处的纵坐标（吸光度）值； $Y'(\sigma)$ 为相应的矫正后的纵坐标值； T 为与基线漂移有关的数值； C 为与该处基线倾斜有关的数值。
- 如果整个光谱只需做水平矫正，则 $C=0$ 、 T 为常数；
- 如果只做斜度矫正，则 $T=0$ 、 C 为常数；
- 如果同时做水平与斜度矫正，则 C 、 T 均不为零。
- 如果这部分光谱呈弯曲状，则可通过逐步改变 T 、 C 数值而达到矫正的目的。

✓ 计算机分峰

- 红外光谱中有许多由各种吸收峰叠加起来的谱带，它们可能是同一化合物在不同吸收谱带的叠加。导数光谱可以分辨出肩峰，分峰程序一般有二种：最小二乘拟合法和解卷积法。
- 用洛伦兹函数与高斯函数的组合峰表征各个子峰，将若干个子峰加和起来拟合成一个谱带，将其与真实谱带相比较，计算出真实谱带与计算谱带间位置、强度与峰形，使这个方差之和趋向最小，就可以把重叠峰分解成若干个子峰。这种方法就是最小二乘拟合法。

- 从实验测得的红外光谱是真实光谱与仪器线性函数卷积的结果，这个卷积降低了光谱的分辨率而使谱宽加大。去除这个线性函数的影响称为解卷积或去卷积，这可以提高光谱的表现分辨率，从而可将重叠的谱带分解成各个子谱带，达到分峰的目的。解卷积可以保留吸收峰的积分面积，因而能提供定量的信息。
- 在傅立叶空间进行谱带解卷积很方便。解卷积技术的参数少，容易完成优化过程，在分离重叠谱带方面应用较广。解卷积可以提高光谱的分辨率，但噪声也会增加，因此进行解卷积的谱图的信噪比一定要好，最好所选光谱区间的二端是基线这样就可以比较出噪声的影响程度。进行光谱解卷积的最大吸光度不要超过1，否则易发生谱图失真。

✓ 使谱图某一区域变为或生成直线

- 要使谱图的某一区域变为空白，计算机就删除这一区域的所有数据点，就是在空白区的数据都丢失。
- 由于目前的红外光谱仪一般都是单光束光谱仪，背景光谱与样品光谱不能同时测定，环境中的CO₂等往往对谱图有干扰，在2400~2240cm⁻¹区间一般没有其他吸收峰而只有CO₂峰，因此使这二点间生成直线就可消除CO₂峰的干扰，这在谱图处理中经常使用。
- 二点间生成直线并不删除这一区间的数据点，只是以直线上的数值代替了原来的数据点。

- 数据处理系统的发展
- ✓ 操作系统灵活性
- ✓ 谱图显示多样性
- ✓ 高速数据采集及高密度存储数据
- ✓ 工作智能化

8.6回归分析

- 多元线性回归
- ✓ 概述
- 线性回归分为：一元线性回归和多元线性回归。
- 两个或两个以上自变量对一个因变量的数量变化关系，称为多元线性回归分析，表现这一数量关系的数学公式，称为多元线性回归模型。
- 多元回归研究的是随机变量 y 与多个普通变量 x_1, x_2, \dots, x_n 的相互关系。多元线性回归（Multiple linear regression, MLR）是一种使用非常广泛的矫正方法，如均匀实验设计等方法中均用到它的原理。

- 传统的分析化学测量，一般都是对单一组分的某个性质进行测量，即根据响应量与被测物浓度的关系来进行矫正分析。这就是基于一元线性回归的单变量矫正方法。
- 但是由于实际的化学体系往往是复杂的多变量系统，在单一组分测定时经常有各种各样的干扰因素。在实际的化学测量中，很多情况下都要对多组分体系进行研究分析。
- 当样品中多个组分共存时，由于各组分的响应信号之间可能产生重叠，而给各组分的定量带来误差，这种情况下常常难以找到对单一组分进行选择性的测量的条件。

✓ 多元线性回归的原理

- 对于一个多组分共存的复杂体系，某一测量信号与多个因素有关，假设响应为 y ，各个因素分别是： x_1, x_2, \dots, x_n 。如果响应 y 与各因素之间为线性关系，则有：

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$$

- 通过对 y 和各自变量 x_1, x_2, \dots, x_n 的一系列观察值，用最小二乘法确定模型的系数 $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ ，从而建立起 y 对 x_i 的线性回归方程。

- 由于响应量的测量往往含有一定的误差，建立在上式基础上对 m 个自变量组成各不相同的观察对象进行测量时，所得测量值可建立以下方程：

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_n x_{1n} + \varepsilon_1$$

$$y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_n x_{2n} + \varepsilon_2$$

...

$$y_m = \beta_0 + \beta_1 x_{m1} + \beta_2 x_{m2} + \dots + \beta_n x_{mn} + \varepsilon_m$$

- 式中， ε 为服从正态分布的随机误差，上式含有 $n+1$ 个未知数，要求 $m \geq n+1$ 。用最小二乘法求解，最小二乘法即使上式中各误差平方和有极小值的一组解，即：

$$\sum_{i=1}^m \varepsilon_i^2 \rightarrow \min$$

□ 设 $Q = \sum_{i=1}^m [y_i - (\beta_0 + \beta_1 x_{i1} + \dots + \beta_n x_{in})]^2 = \sum_{i=1}^m \varepsilon_i^2$

- 欲使 $\sum_{i=1}^m \varepsilon_i^2$ 有极小值，则有：

$$\frac{\partial Q}{\partial \beta_0} = -2 \sum_{i=1}^m [y_i - (\beta_0 + \beta_1 x_{i1} + \dots + \beta_n x_{in})] = 0$$

$$\frac{\partial Q}{\partial \beta_1} = -2 x_{11} \sum_{i=1}^m [y_i - (\beta_0 + \beta_1 x_{i1} + \dots + \beta_n x_{in})] = 0 \quad \dots$$

$$\frac{\partial Q}{\partial \beta_n} = -2 x_{n1} \sum_{i=1}^m [y_i - (\beta_0 + \beta_1 x_{i1} + \dots + \beta_n x_{in})] = 0$$

□ 又 $x_{i1} \neq 0$ ，方程组可变形为：

$$\begin{aligned}\sum y_i &= \beta_0 m + \beta_1 \sum x_{i1} + \dots + \beta_r \sum x_{im} \\ \sum x_{i1} y_i &= \beta_0 \sum x_{i1} + \beta_1 \sum x_{i1}^2 + \beta_2 \sum x_{i1} x_{i2} + \dots + \beta_r \sum x_{i1} x_{im} \\ &\dots \\ \sum x_{im} y_i &= \beta_0 \sum x_{im} + \beta_1 \sum x_{i1} x_{im} + \beta_2 \sum x_{i2} x_{im} + \dots + \beta_r \sum x_{im}^2\end{aligned}$$

□ 上式称为正规方程组，其方程数据与未知数目相等。

□ 上式右边的系数矩阵为：

$$\begin{bmatrix} m & \sum x_{i1} & \dots & \sum x_{im} \\ \sum x_{i1} & \sum x_{i1}^2 & \dots & \sum x_{i1} x_{im} \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_{i1} & \sum x_{i1} x_{im} & \dots & \sum x_{im}^2 \end{bmatrix} = X^T X$$

□ 而其左边为：

$$\begin{bmatrix} \sum y_i \\ \sum x_{i1} y_i \\ \vdots \\ \sum x_{im} y_i \end{bmatrix} = X^T y$$

□ 故上式的矩阵形式 $X^T y = X^T X \beta$

□ 当 $X^T X$ 的逆矩阵 $(X^T X)^{-1}$ 存在时，系数向量 β 的估计值为：

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

□ 假如 X 的各行或各列之间存在相关性，则称 $X^T X$ 为病态（亦称线性相关或奇异），则其逆将不可能正确。

□ 应尽可能使 X 的各行（或各列）线性无关，一般来说按正交设计原则或其他实验设计原则安排 X 矩阵中各元素，对多元线性回归求解是有利的。

□ 这时有：

$$\hat{y} = X \hat{\beta} = X (X^T X)^{-1} X^T y$$

□ 上式表示了实验值 y_i 与其拟合值之间的关系，它们两者之间有可能相符，也有可能并不相符，甚至相差很大，即拟合结果不理想。

□ 因而对拟合结果进行的优劣应进行统计检验。通常采用复相关系数 r 来评价拟合值 \hat{y}_i 和实验数据 y_i 之间的关系，并进行检验。

□ 根据方差分析的思想，将 y 的数据波动，即 y 的总方差和 $SS_T(\text{total})$ 分解成两部分：一部分是由于各自变量的变化而引起的 y 的波动，即 **回归差方** 和 $SS_{\text{Reg}}(\text{regression})$ ；一部分是由于随机误差或其他未知因素所引起的 y 的波动，即 **残余差方** 和 $SS_{\text{Res}}(\text{residual})$ 。

□ SS_T 、 SS_{Reg} 和 SS_{Res} 的自由度为 $m-1$ ， n 和 $m-n-1$ 。

$$SS_T = \sum_{i=1}^m (y_i - \bar{y})^2 = \sum_{i=1}^m (\bar{y}_i - \bar{y})^2 + \sum_{i=1}^m (y_i - \bar{y}_i)^2 = SS_{\text{Reg}} + SS_{\text{Res}}$$

□ 复相关系数 r 反映 y_i 与 x_i 之间线性相关程度的大小。

$$r = \sqrt{\frac{SS_{Reg}}{SS_T}} = \sqrt{1 - \frac{SS_{Res}}{SS_T}}$$

- 当残余方差和 $SS_{Res}=0$ 时, $r=1$, 说明 y 与 x_i 之间完全按线性关系变化, y 的方差全部是由 n 个自变量贡献的, 因而各自变量与 y 完全相关。
- 当 $SS_{Res}=0$, 即 $r=0$, 说明全部自变量对 y 的方差没有贡献, 因而 y 与各自变量无关。
- 通常情况下, 有 $0 \leq r \leq 1$, r 越接近于1, 表明 y 与各自变量相关关系越好。

• 多元非线性回归

- ✓ 在发射光谱分析中, 用感光板记录的发射光谱分析线对的黑度差 $\Delta s/r$ 和被测组分含量 c 之间的关系, 就不是线性关系
- ✓ 它们之间的关系: $\Delta s/r = \lg \alpha + b \lg c$
- ✓ 处理非线性问题的方法有如下几种: 一种是进行变量代换, 将非线性问题变成线性问题去处理; 另一种是非线性回归法 (Multiple nonlinear regression), 即最小二乘估计非线性模型中的参数, 进而建立非线性的回归模型。

✓ 变量代换

□ 例如上式的非线性模型, 可作如下代换:

$$y = \Delta s/r, a = \lg \alpha, b = \lg c, x = b$$

- 于是变成: $y=a+bx$
- 这就可以用一元线性回归方法来处理了。
- ✓ 非线性最小二乘法
- 处理各种非线性矫正问题的一般方法是用非线性最小二乘法(nonlinear least squares)对非线性模型中的各参数进行估计, 从而建立非线性矫正模型并用于预报。

- 假设响应量 y 与自变量 x 之间的关系可用以下非线性模型拟合（注意自变量 x 可能不止一个，但为了问题简单化，这里仅讨论常见的一个自变量的情况）：

$$y_i = f(x_i, b_1, b_2, \dots, b_n) \quad i = 1, 2, \dots, m$$

- 式中， b_j 为待估计参数； y 是 x 的非线性函数。
 □ 要从 m 次实验所得到的测量值 (y_i, x_i) ($i=1, 2, \dots, m$)对参数 b_j ($j=1, 2, \dots, n$)作出估计。由于非线性函数一般难以直接得其解析解，通常采用迭代（即逼近）的方法来估计参数 b_j 。
 □ 首先给参数 b_j ($j=1, 2, \dots, n$)赋予初值 $b_j^{(0)}$ 它与真值 b_j 之差为 Δb_j ，即有：

$$b_j = b_j^{(0)} + \Delta b_j$$

 □ 所以对参数 b_j 的估计就变成在逐步逼近过程中求 Δb_j 的问题

- $b_j^{(0)}$ 比较接近于 b_j ，则可将函数 $y = f(x, b_1, b_2, \dots, b_n)$ 在初值 $b_j^{(0)}$ 处作Taylor级数展开。为使问题简单化，略去高此项、只保留一次项，对于第 i 次实验观察值有：

$$y_i = f(x_i, b_1, b_2, \dots, b_n) = f_i^{(0)} + \frac{\partial f_i}{\partial b_1} \Delta b_1 + \frac{\partial f_i}{\partial b_2} \Delta b_2 + \dots + \frac{\partial f_i}{\partial b_n} \Delta b_n + \varepsilon_i$$

- 式中， $f_i^{(0)} = f(x_i, b_1^{(0)}, \dots, b_n^{(0)})$
 □ 它是各参数为初值 $b_j^{(0)}$ 时第 i 次实验时的函数计算值； $\frac{\partial f_i}{\partial b_j}$ 是参数为 $b_j^{(0)}$ 时函数的各偏导数值。所以当给定初值 $b_j^{(0)}$ 后，对于一组自变量 (x_1, x_2, \dots, x_m) ，可分别求得 $\frac{\partial f_i}{\partial b_j}$ 和 $f_i^{(0)}$ 的值：

- 显然通过Taylor级数展开，仅保留一次项处理后可得一组线性方程组。由最小二乘法即可解出方程组中 Δb_j 。

- 最小二乘法要求每次测量误差平方和最小，即

$$\sum_{i=0}^m e_i^2 \Rightarrow \min$$

- 矩阵形式：

$$\begin{bmatrix} y_1 - f_1^{(0)} \\ y_2 - f_2^{(0)} \\ \vdots \\ y_m - f_m^{(0)} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial b_1} & \frac{\partial f_1}{\partial b_2} & \dots & \frac{\partial f_1}{\partial b_n} \\ \frac{\partial f_2}{\partial b_1} & \frac{\partial f_2}{\partial b_2} & \dots & \frac{\partial f_2}{\partial b_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial b_1} & \frac{\partial f_m}{\partial b_2} & \dots & \frac{\partial f_m}{\partial b_n} \end{bmatrix} \begin{bmatrix} \Delta b_1 \\ \Delta b_2 \\ \vdots \\ \Delta b_n \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{bmatrix}$$

$$\Delta y = \begin{bmatrix} y_1 - f_1^{(0)} \\ y_2 - f_2^{(0)} \\ \vdots \\ y_m - f_m^{(0)} \end{bmatrix} \quad F = \begin{bmatrix} \frac{\partial f_1}{\partial b_1} & \frac{\partial f_1}{\partial b_2} & \dots & \frac{\partial f_1}{\partial b_n} \\ \frac{\partial f_2}{\partial b_1} & \frac{\partial f_2}{\partial b_2} & \dots & \frac{\partial f_2}{\partial b_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial b_1} & \frac{\partial f_m}{\partial b_2} & \dots & \frac{\partial f_m}{\partial b_n} \end{bmatrix} \quad \Delta b = \begin{bmatrix} \Delta b_1 \\ \Delta b_2 \\ \vdots \\ \Delta b_n \end{bmatrix}$$

- 故可写作矩阵简式: $\Delta y = F \Delta b$
- 其最小二乘解为: $\Delta b = (F^T F)^{-1} F^T \Delta y$
- 由此即得各系数的增量 $\Delta b_1, \Delta b_2, \dots, \Delta b_n$ 。将各 Δb_j 代入, 可得各参数的新的估计值

$$b_j^{(1)} = b_j^{(0)} + \Delta b_j \quad j = 1, 2, \dots, n$$
- 以 $b_j^{(1)}$ 为第二次迭代初值, 重复以上步骤进行计算, 直至收敛。

- 与线性回归的情况不同, 矩阵 F (常称之为Jacobian矩阵) 与参数 b_j 有关, 每一次新的迭代矩阵 F 都要重新计算 ($F^{(0)} \rightarrow F^{(1)} \rightarrow F^{(2)} \rightarrow \dots$)。而在线性回归中, 系数矩阵不随变量而变化, 因此它不需要因参数的变化而更新。

- 迭代计算的每一步误差的差方和由下式给出:

$$ss(b) = \sum_{i=1}^m [y_i - \hat{y}_i(b)]^2$$

- 在迭代过程中, 根据上式可以知道误差的差方和是否下降。同时可以由下式来判断收敛的情况:

$$|b_j^{(k)} - b_j^{(k+1)}| / |b_j^{(k)}| < \delta \quad j = 1, 2, \dots, n$$

- 式中, δ 为一个预先设定的较小的数 (例如 10^{-4})。

- 迭代结束后, 误差差方和 $ss(b) = \sum_{i=1}^m [e_i]^2$ 可以用来估计误差的方差:

$$s_e^2 = \sum e_i^2 / (m - n)$$

- 一般来说, 非线性回归的运算量较大, 而参数初值的选择迭代收敛的速度有较大的影响。