

Part III-B: Analysis Chemistry

Lecture by 王敏

Note by THF

2024 年 10 月 15 日

目录

1	概论	2
2	误差和分析数据处理	3
2.1	准确度和精密度	3
2.1.1	准确度和误差	3
2.1.2	精密度和偏差	5
2.1.3	误差	6
2.2	如何提高分析结果的准确度	7
2.3	有效数字	8
2.3.1	修约规则	8
2.3.2	运算规则	8
2.4	有限量分析数据的统计处理	9
2.4.1	偶然误差的正态分布	9
2.4.2	t 分布	10
2.4.3	平均值的精密度和置信区间	11
2.4.4	可疑数据的取舍和显著性检验	12
2.4.5	相关和回归	13

3 酸碱滴定法	14
3.1 概述	14
3.2 基本原理	14
3.2.1 酸碱定义	14
3.2.2 酸碱的强度	15
3.2.3 分布系数	15
3.2.4 缓冲溶液的 pH	17

Lecture 1

1 概论

20 世纪 20-30 年代: 分析化学出现四大反应平衡理论的建立

20 世纪 40-50 年代: 光电色谱仪器设备出现

Notation. Bloch F and Purcell E M 建立核磁共振测定

Martin A J P and Synge R L M 建立气相色谱

Heyrovsky J 建立极谱分析法

20 世纪 70 年代以来: 计算机参与自动化

Notation. 分析化学分析方法: 3S+2A

3S: Sensitivity, Selectivity, Speediness

2A: Accuracy, Automatics

分析化学主要发展趋势:

{
在线分析
原位分析
实时分析
活体分析
...
}

Notation. 分析过程的步骤:

1. 分析方法选择
2. 取样 (Sampling, 具代表性的样本)
3. 制备试样 (适合与选定的分析方法, 消除可能的干扰)
4. 分析测定 (优化条件, 仪器校正, 方法验证)
5. 结果处理和表达 (统计学分析, 测量结果的可靠性分析, 书面报告)

Notation. 制备试样首先需要进行样品前处理

方法验证: 线性性, 灵敏性等

Lecture 2

2 误差和分析数据处理

Notation. “挑数据”: 做标准曲线

标品浓度 (0,1,2,3,4,5,6,7)

吸光度 (0,0.1,0.2,0.3,0.4,0.001,0.6,0.7)

对于可疑数据, 需要通过其他方法进行确认 (至少 3 次试验)

2.1 准确度和精密度

2.1.1 准确度和误差

Definition. 准确度 (Accuracy): 测量值和真值的接近程度, 准确度的高低用误差大小衡量

Definition. 误差 (error): 测量结果和真值的差值

误差具有客观性和普遍性

实验结果都有误差, 测量值只能尽可能接近真实值

Definition. 约定真值: 由国际计量大会定义的单位及我国的法定计量单位

Example. 国际单位制基本单位 “米”、“克” 等

Notation. 约定真值是有一个量的真值的近似值, 误差可以忽略不计

Definition. 标准值: 采用可靠的分析方法、在不同实验室、由不同的分析人员、对同一个试样反复多次测定后将大量数据用数理统计求得的测量值

误差的表示方法:

$$\delta = x - \mu.$$

Definition. 绝对误差: δ

测量值: x

真值: μ

相对误差: RE%

绝对误差 (Absolute Error) 可正可负, 单位为测量值的单位

绝对误差的绝对值越小, 准确度越高

相对误差 (Relative Error, RE):

$$\text{RE}\% = \frac{\delta}{\mu} \times 100\%.$$

或:

$$\text{RE}\% = \frac{\delta}{x} \times 100\%.$$

相对误差无单位, 可正可负

Example. 有真实值为 0.0020g 和 0.5000g 的两个样品, 称量结果分别为 0.0021g 和 0.5001g, 计算相对误差和绝对误差

$$\delta_1 = 0.0021 - 0.0020 = 0.0001, \quad \delta_2 = 0.5001 - 0.5000 = 0.0001 = \delta_1$$

$$\text{RE}_1\% = \frac{0.0001}{0.0020} \times 100\% = 5\%$$

$$\text{RE}_2\% = \frac{0.0001}{0.5000} \times 100\% = 0.02\%.$$

Notation. RE 要求: 测高含量组分, RE 可小; 测低含量组分, RE 可大
高含量组分对应化学分析法; 低含量组分对应仪器分析法

2.1.2 精密度和偏差

Definition. 精密度: 在规定的测定条件下, 多次平行测定结果相互吻合的程度, 精密度高低用偏差衡量

偏差: 单个测量值和测量平均值的差距

Definition. 绝对偏差: d

相对偏差: $d\%$

平均偏差: \bar{d}

标准偏差: SD

相对标准偏差: RSD

绝对偏差:

$$d = x_i - \bar{x}.$$

相对偏差:

$$d\% = \frac{d}{\bar{x}}.$$

平均偏差:

$$\bar{d} = \frac{\sum |x_i - \bar{x}|}{n}.$$

相对平均偏差:

$$\frac{\bar{d}}{\bar{x}} \times 100\% = \frac{\sum |x_i - \bar{x}|}{n \cdot \bar{x}} \times 100\%.$$

标准偏差:

$$S_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (n \leq 20).$$

$n - 1$ 称为自由度

相对标准偏差 (RSD , 变异系数):

$$RSD = \frac{S_x}{\bar{x}} \times 100\%.$$

RSD 越小, 数据越集中, 精密度越高

Notation. 方法的精密度考察:

1. 重复性 (repeatability): 同一实验室, 较短时间间隔, 同一分析人员对同一试样测定所得结果的接近程度
2. 中间精密度 (intermediate precision): 改变某些试验条件对同一试样测定结果的接近程度
3. 重现性 (reproducibility): 不同实验室, 不同人员对同一试样测试结果的接近程度

Notation. 精密度和准确度的关系:

1. 精密度高是准确度高的前提
2. 精密度高, 准确度不一定高
3. 只有精密度和准确度都高的数据才可取

Notation. 准确度体现结果的正确性, 精密度体现结果的重现性

2.1.3 误差

Notation. 误差分类: 系统误差、偶然误差 (随机误差)

Definition. 系统误差 (可定误差): 分析中由某些确定原因造成的误差

特点:

1. 重现性
2. 单向性
3. 可以校正消除
4. 影响准确度

Notation. 系统误差分类:

$$\left\{ \begin{array}{l} \text{方法误差: 方法不完善} \\ \text{仪器误差: 仪器本身缺陷} \\ \text{试剂误差: 试剂有杂质} \\ \text{操作误差、主观误差} \end{array} \right.$$

Definition. 偶然误差 (随机误差): 由偶然因素影响

特点:

1. 无单向性 (方向大小不确定)
2. 符合统计学规律 (大误差出现的概率小, 随机误差的正态分布)
3. 可以通过增加平行测定次数减小
4. 影响精密度

Definition. 过失: 由分析工作者粗心大意造成, 可以避免

Lecture 3

2.2 如何提高分析结果的准确度

1. 选择合适的分析方法:

Notation. 对于化学分析法: 适用于常量组分 ($>1\%$) 的分析

对于仪器分析法: 适用于微量组分 ($0.01 \sim 1\%$) 或痕量组分 ($<0.01\%$) 的分析

2. 减少相对误差: 增大取样量
3. 减少偶然误差: 增加平行测试次数 (≥ 3 , 活体测试 ≥ 6)
4. 消除系统误差:

Notation. 常用方法:

1. 与经典方法比较: (用其他方法测试该方法)
2. 对照试验 (control test): (用其他样本对比)
3. 回收试验/标准加入法: 测量原样本 x_1 , 加入某标准量 x_2 , 测量加入后的样本 x_3 , 计算回收率:

$$\text{Recovery}\% = \frac{x_3 - x_1}{x_2} \times 100\% \approx 95\% \sim 105\%.$$

4. 空白试验: 不加试样测试, 得到并扣去空白值, 用于检验由试剂、容器等引入杂质导致的系统误差
5. 仪器校正

2.3 有效数字

Definition. 有效数字: 分析工作中实际上能测量到的数字

原则上:

1. 在记录测量数据时, 只允许保留一位可疑数字 (欠准数字)
2. 误差是末尾数的 ± 1 个单位

Example. 在分析天平上称出 $m = 21.5370\text{g}$, 则真值为 $21.5370 \pm 0.0001\text{g}$

Notation. 有效数字的末尾 0 不可省略: 反映相对误差

Question. 如何判断有效数字的位数?

1. 在数据中 1 ~ 9 均为有效数字 (0 待定)
2. 算式中的倍数、分数及某些常数 (π, e 等) 可看为无数位有效数字
3. 变化单位时有效数字的位数必须保持不变, 如 $0.0015\text{g} = 1.5\text{mg}$
4. pH 和 pK_a 等对数值, 有效数字仅取决于小数部分, 例: $\text{pH} = 4.23$, 有效数字 2 位

Notation. 0 的位置和有效数字:

1. 小数前的 0 起定位作用, 后面的 0 为有效数字
2. 整数后的 0 不一定是有效数字

Example. $36000 \Rightarrow 3.60 \times 10^4$: 3 位有效数字

2.3.1 修约规则

1. 四舍六入五留双, 五后有数需进位
2. 修约标准偏差: 只进不舍, 降低精密度, 提高可信度
 - 2.1. 可多保留一位有效数字进行计算
 - 2.2. 与标准限度值比较不应修约

2.3.2 运算规则

1. 加减法: 结果的小数位数以小数点后位数最少的为标准

表 1: 修约为两位

原值	修约值	原因
3.249	3.2	四舍
8.361	8.4	六入
6.550	6.6	五留双
6.250	6.2	五留双
6.252	6.3	五后有数

Example.

$$0.0121 + 25.64 + 1.057 = 26.7091 \approx 26.71.$$

2. 乘除法：取相对误差最大的为标准

Example.

$$\frac{0.0325 \times 5.10 \times 60.1}{139.8} \approx 0.0712.$$

Notation. 对于高含量组分 ($w > 10\%$), 分析结果一般保留 4 位有效数字

对于中等含量组分 ($1\% < w < 10\%$), 保留 3 位

对于微量组分 ($w < 1\%$), 保留 2 位

2.4 有限量分析数据的统计处理

2.4.1 偶然误差的正态分布

正态分布的概率密度函数：

$$y = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}.$$

Notation. x : 测量值

μ : 无限次测量的总体平均值

σ : 总体偏差

特点：

1. 当 $x = \mu$ 时, y 最大：大部分测量值集中在算术平均值附近

2. 函数图像以 $x = \mu$ 的直线对称: 正负误差出现的概率相同
3. $x \rightarrow -\infty$ 或 $x \rightarrow +\infty$ 时: 无限趋近 x 轴
4. μ 越大, 数据越分散, 函数图像矮小、坡度较缓
- 5.

$$\int_{-\infty}^{+\infty} f(x) dx = 1.$$

Notation. 若令

$$u = \frac{x - \mu}{\sigma}.$$

则:

$$y = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-u^2/2}.$$

称为标准正态分布曲线

Notation. 3σ 标准: 测量结果需要有 99.7% 以上的数据在真值范围内, 即:

$$\int_{\mu-3\sigma}^{\mu+3\sigma} f(x) dx \approx 99.7\%.$$

2.4.2 t 分布

Definition. t 分布为有限量数据 n 平均值的概率误差分布

$$t = \frac{|\bar{x} - \mu|}{S} \sqrt{n}.$$

\bar{x} : 样本平均值

S : 样品标准偏差

t 值随自由度 $f (f = n - 1)$ 而变

当 $f \rightarrow +\infty$ 时为正态分布

Notation. t 分布和正态分布的异同:

相同: 形状相似, 积分面积表示概率

不同: t 不同时概率不同

Notation. t 分布相关概念:

1. 自由度 f

2. 置信区间: 以测定区间为中心, 包括总体平均值在内的可信范围

$$\bar{x} \pm \frac{tS}{\sqrt{n}}.$$

3. 置信水平 (置信度 P): 样本平均值落在置信区间的概率

$$P = \mu \pm \frac{tS}{\sqrt{n}}.$$

4. 显著性水平 $\alpha = 1 - P$

t 值的表达: 一定 P 下, $t \rightarrow t_{\alpha, f}$

2.4.3 平均值的精密度和置信区间

Example. 有一个样品, m 个人各测量 n 次, 计算出每个人测得的平均值, 平均值的分布符合正态分布

平均值的标准偏差/平均值的精密度:

$$S_{\bar{x}} = \frac{S_x}{\sqrt{n}}.$$

即: 增加平行测定次数 n 可以减小平均值的标准偏差 $S_{\bar{x}}$

Notation. 增加测定次数, 平均值的标准偏差呈反比变化, 一般 3 到 4 次已经可以达到目标, 继续增加效果不显著

Lecture 4

10.08

Notation. 常用估计法: 点估计

求 x 和 \bar{x} 进行比较

平均值置信区间:

1. 单次

$$\mu = x \pm u\sigma.$$

其中 $u = (x - \mu)/\sigma$ (见标准正态分布曲线)

2. 多次: 使用 $\sigma_{\bar{x}}$ 代替 σ

$$\mu = \bar{x} \pm u\sigma_{\bar{x}} = \bar{x} \pm u \cdot \frac{\sigma}{\sqrt{n}}.$$

3. 少量: t 分布

$$\mu = \bar{x} \pm tS_{\bar{x}} = \bar{x} \pm \frac{tS}{\sqrt{n}}.$$

可以用 X_U 表示上限, X_L 表示下限

Notation. 置信水平越高, 置信区间越宽

通常置信水平取 $P = 0.95$ 或 $\alpha = 0.05$

置信区间反映估计的精密度

置信水平说明估计的把握程度

Notation. 1. μ 为一个定值, 无随机性

2. 单侧检验大于或小于总体均值, 双侧检验同时大于和小于总体均值

2.4.4 可疑数据的取舍和显著性检验

主要使用 G 检验 (可疑数据取舍) $\rightarrow F$ 检验 (精密度检验) $\rightarrow t$ 检验 (准确度检验)

Notation. F 检验和 t 检验合称显著性检验

Notation. G 检验

在一组平行测量数据中有过高或过低的数据, 称为可疑数据/异常值/逸出值

1. 确定可疑值 x_q , 求出包括可疑值在内的平均值 \bar{x}

2. 求出可疑值与平均值之差的绝对值 $|x_q - \bar{x}|$

3. 计算包括可疑值在内的标准偏差 S

$$G = \frac{|x_q - \bar{x}|}{S}.$$

4. 根据置信度 P 得到 α , 查表得 G 的临界值 $G_{\alpha,n}$, 若 $G > G_{\alpha,n}$, 则数据应当舍弃

Notation. F 检验

判断两组数据间存在偶然误差是否有显著不同

$$F = \frac{S_1^2}{S_2^2}, (S_1 > S_2).$$

f_1, f_2 为 S_1, S_2 的自由度

P 一定时, 查表得 F_{α, f_1, f_2} , 比较 $F < F_{\alpha, f_1, f_2}$ 则无显著性差异

Notation. t 检验

判断某一分析方法或操作过程中是否存在较大的系统误差

1. 使用 \bar{x} 和 μ 的比较: 已知真值的 t 检验

$$t = \frac{|\bar{x} - \mu|}{s} \sqrt{n}.$$

2. 两组样本平均值比较: 未知真值的 t 检验

两组数据:

$$\begin{cases} n_1, s_1, \bar{x}_1 \\ n_2, s_2, \bar{x}_2 \end{cases}.$$

当 $s_1 \approx s_2$, 令合并标准偏差:

$$s_R = \sqrt{\frac{\text{偏差平方和}}{\text{总自由度}}} = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 - 1 + n_2 - 1}}.$$

$$\Rightarrow t = \frac{|\bar{x}_1 - \bar{x}_2|}{s_R} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}.$$

查表得 $t_{\alpha, f}$ (若无真值使用总自由度 $f = f_1 + f_2$), 比较 $t, t_{\alpha, f}$

Notation. 更常用的数据组比较准确度方法: ANOVA 分析

通过比较 p 值判断, 用 * 号个数表示差异的大小

2.4.5 相关和回归

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \times \sum_{i=1}^n (y_i - \bar{y})^2}}.$$

作业: P25, 8、9 题

3 酸碱滴定法

3.1 概述

滴定分析法 (titrimetry) 是一种定量分析方法

Notation. 将已知准确浓度的试剂溶液低加到待测物质的溶液中, 通过滴加试剂的浓度和体积, 定量计算待测物质的含量

滴定分析法又称容量分析法

特点:

1. 准确度高、适于常量组分分析 (组分含量 $> 1\%$, 取样量 $> 0.1\text{g}$)
2. 易于操作
3. 快速

Notation. 基本术语:

滴定剂: 浓度准确已知的试样溶液

滴定: 将滴定剂通过滴定管逐滴滴加到被测物质溶液中的过程

化学计量点: 滴定剂与待测溶液按化学计量关系完全反应的点, 用 sp 表示

指示剂: 发生颜色改变指示终点的物质

滴定终点: 指示剂发生颜色改变的点, 用 ep 表示

滴定误差: 滴定终点和化学计量点的差距, 用 TE 表示

3.2 基本原理

3.2.1 酸碱定义

1. 电离理论

$$\begin{cases} \text{酸: 能电离出H}^+\text{的物质} \\ \text{碱: 能电离出OH}^-\text{的物质} \end{cases}$$

2. 质子理论

$$\begin{cases} \text{Brønster 酸: 能给出质子的物质} \\ \text{Brønster 碱: 能接受质子的物质} \end{cases}$$

Notation. 共轭酸碱对: 酸 \leftrightarrow 碱 (质子)

两性物质: 能给出和接受质子的物质 (例: 水)

3.2.2 酸碱的强度

1. 酸 HA 的解离常数:

$$K_a = \frac{[\text{H}_3\text{O}^+][\text{A}^-]}{[\text{HA}]}$$

2. 碱 BOH 的解离常数:

$$K_b = \frac{[\text{OH}^-][\text{B}^+]}{[\text{BOH}]}$$

Lecture 5

10.15

3.2.3 分布系数

Definition. 分布系数: δ

用 $\delta_0, \delta_1, \delta_2 \dots$ 来表示电荷数为 0, 1, 2... 的组分的分布系数

一元弱酸的浓度 (分析浓度) 为 c , 解离平衡后:

$$\begin{aligned}\delta_{\text{HA}} &= \frac{[\text{HA}]}{c} = \frac{[\text{HA}]}{[\text{HA}] + [\text{A}^-]} \\ &= \frac{1}{1 + K_a/[\text{H}^+]} = \frac{[\text{H}^+]}{[\text{H}^+] + K_a} \\ \delta_{\text{A}^-} &= \frac{K_a}{[\text{H}^+] + K_a}\end{aligned}$$

易得 $\delta_{\text{HA}} + \delta_{\text{A}^-} = 1$

Example. 计算 pH=5 是 0.10mol/L HAc 溶液中各组分 (HAc, Ac⁻, H⁺) 的分布系数 δ 和平衡浓度 c

解: $[\text{H}^+] = 10^{-5}\text{mol/L}, K_a = 1.7 \times 10^{-5}$

计算 δ :

$$\delta_0 = \delta_{\text{HAc}} = \frac{[\text{H}^+]}{[\text{H}^+] + K_a} = 0.37.$$

$$\delta_1 = \delta_{\text{Ac}^-} = 1 - \delta_{\text{HAc}} = 1 - 0.37 = 0.63.$$

....

Notation. 一元强酸 H^+ 的浓度 (精确式)

$$[\text{H}^+] = \frac{c_a + \sqrt{c_a^2 + 4K_w}}{2}.$$

忽略水的解离 ($c_a \geq 20[\text{OH}^-]$), 可以近似为下式 (最简式):

$$\begin{cases} [\text{H}^+] = [\text{A}^-] = c_a \\ \text{pH} = -\lg[\text{H}^+] = -\lg c_a \end{cases}.$$

Notation. 一元弱酸的 pH 计算:

精确式:

$$[\text{H}^+]^2 = K_a (c_a - [\text{H}^+]).$$

$$[\text{H}^+] = \frac{-K_a + \sqrt{K_a^2 + 4c_a \cdot K_a}}{2}.$$

近似式 ($c_a K_a < 20K_w, c_a/K_w \geq 500$)

$$[\text{H}^+] = \sqrt{c_a K_a + K_w}.$$

最简式 ($c_a K_a \geq 20K_w, c_a/K_w \geq 500$)

$$[\text{H}^+] = \sqrt{K_a \cdot c_a}.$$

Notation. 两性物质 pH 计算

精确式:

$$[\text{H}^+] = \sqrt{\frac{K_{a1} (K_{a2} [\text{HB}] + K_w)}{K_{a1} + [\text{HB}]}.$$

近似式:

$$[\text{H}^+] = \sqrt{\frac{K_{a1} (K_{a2} c + K_w)}{K_{a1} + c}}.$$

最简式:

$$[\text{H}^+] = \sqrt{K_{a1} K_{a2}}.$$

$$\text{pH} = \frac{1}{2} (\text{p}K_{a1} + \text{p}K_{a2}).$$

3.2.4 缓冲溶液的 pH

Notation. 缓冲溶液的种类:

1. 共轭酸碱对
2. 两性物质, 如 $\text{H}_2\text{PO}_4 \sim \text{HPO}_4^{2+}$ (PBS, 磷酸缓冲盐)
3. 高浓度的强酸/强碱

Notation. 缓冲溶液有缓冲能力/缓冲容量 β , 与缓冲溶液的总浓度与组分有关

1. 总浓度 $\propto \beta$
2. 缓冲组分的浓度比 $\rightarrow 1 \propto \beta$
- 3.

缓冲溶液的 Henderson 缓冲公式:

$$[\text{H}^+] = K_a \frac{c_a}{c_b}.$$

$$\text{pH} = \text{p}K_a + \lg \frac{c_b}{c_a}.$$

Notation. 生物上常用的缓冲溶液:

1. 三(羟甲基)氨基甲烷 +HCl (tris-HCL)
2. HEPES