

Excel を用いた検量線の作成手順と利用

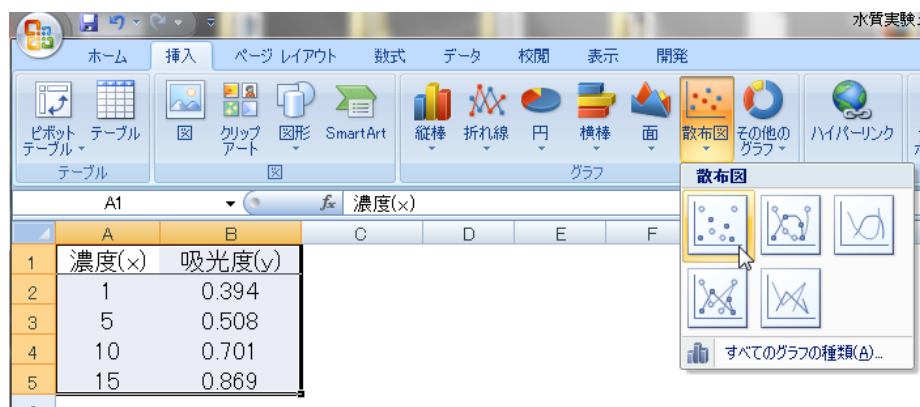
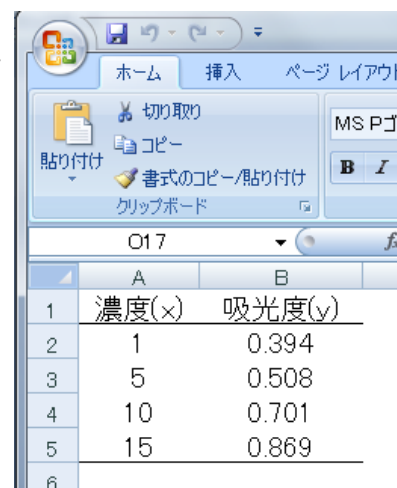
清水顕史

分光光度計で測定した検量線の定量結果が、例えば以下の場合を考えます。

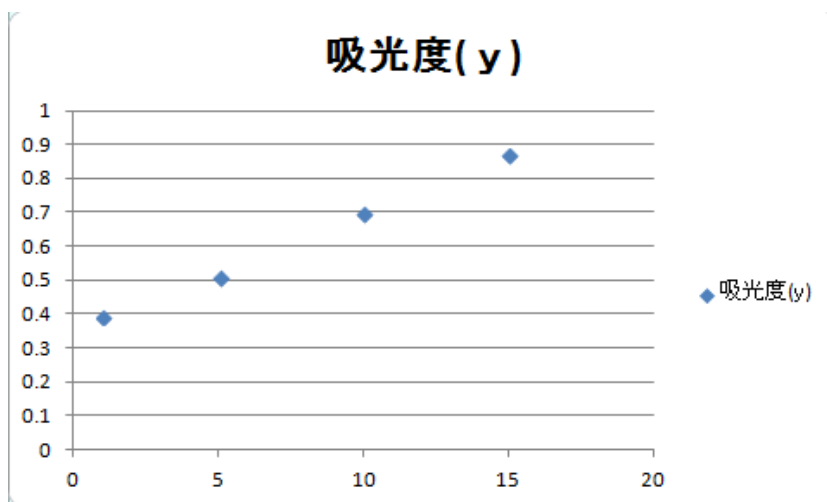
濃度(x)	吸光度(y)
1	0.394
5	0.508
10	0.701
15	0.869

濃度範囲 1~15 の間で、濃度と吸光度に直線関係が成り立つとき、濃度(x)に対する吸光度(y)の一次回帰式 $y = a + b x$ を推定します。Excel で回帰式を計算する場合には、散布図とその近似曲線の描出機能を使うと、数値の回帰式への当てはまりの良さを視覚的に確認できて便利です。

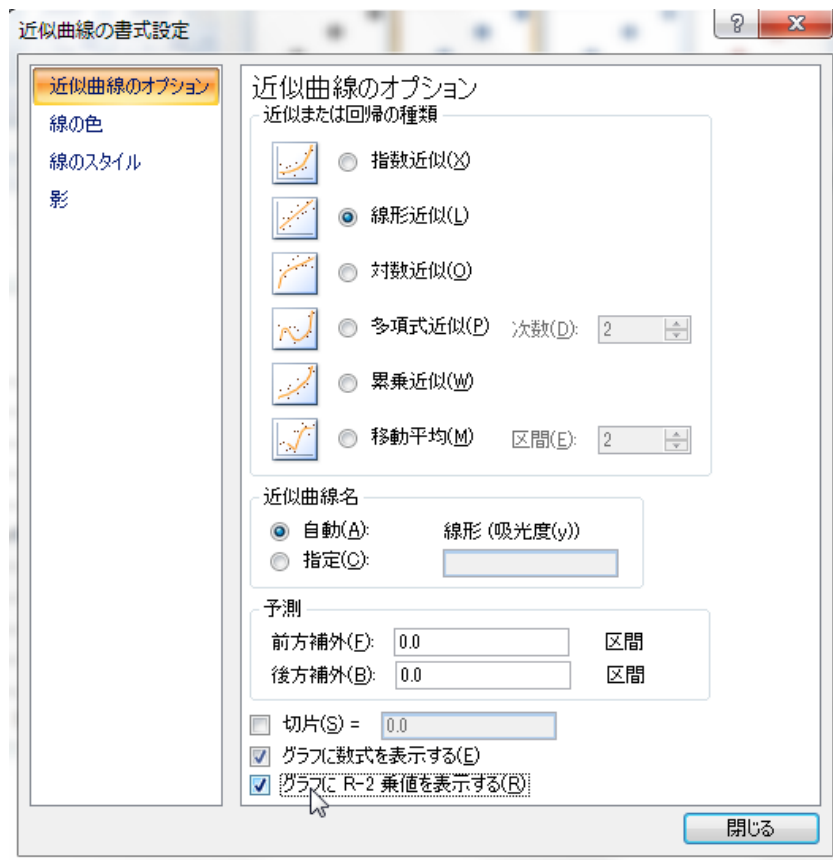
1. まず Excel で、右図のように A 列の 2~5 行に濃度を、そして対応する吸光度を B 列の 2~5 行に入力しておきます。A 列の 1 行目と B 列の 1 行目には変数のタイトルを入力しておきます。
2. 入力した 2 列のデータ(各列が各変数に対応)を、タイトルも含めてドラッグし (左ボタンを押さえたままでマウスを移動)、ファイルメニューの[挿入]を選んで[散布図]を選びます。



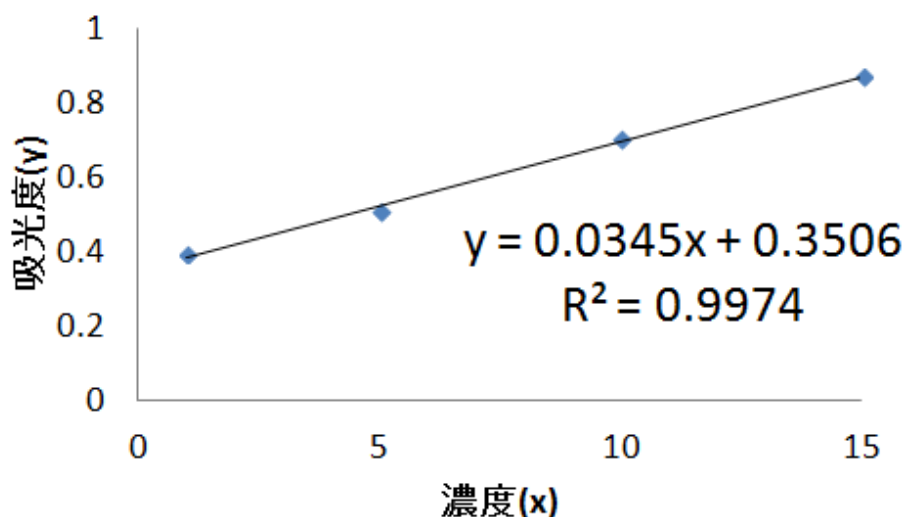
3. 濃度(x)を横軸、吸光度(y)を縦軸にした散布図が自動作成されます。



4. 図中のプロット(デフォルトでは菱形)を右クリックしてメニューを開き、[近似曲線の追加]を選ぶと開く書式設定ウィンドウで、線形近似(デフォルト)のまま[グラフに数式を表示する]、[グラフに R-2 乗値を表示する]をチェックします。



5. 図中に表示される数式が、求めたい回帰式です(軸にタイトルを付け、図中の余分な線を除いています。図中の縦・横の補助線は、理系の報告書では描出しないことの方が多)。ここで注目すべきなのは、 R^2 の値で、この値はプロットの回帰式への当てはまりの良さを示し、1 に近づくほど検量線の精度が良いといえます。定量実験で R^2 は 0.98 以上が望ましく 0.95 以下であれば検量線を希釈しなおすべきでしょう。



6. 推定された回帰式 $y = a + b x$ で、 a を回帰定数、 b を回帰係数といいます。これらの係数は最小 2 乗法による推定量で、正式には \hat{a} と \hat{b} と記しますが($\hat{}$ はハットと呼ぶ)、ここでは簡単のため a と b のままで表記します。未知試料の濃度は、測定した吸光度からこの回帰式を用いて推定します。まず a と b は Excel 関数 intercept と slope を用いて計算できます。つまり、それぞれ y の数列と x の数列を以下のように入力します(下図)。

	A	B
1	濃度(x)	吸光度(y)
2	1	0.394
3	5	0.508
4	10	0.701
5	15	0.869
6		
7	回帰係数	=SLOPE(B2:B5,A2:A5)
8	回帰定数	=INTERCEPT(B2:B5,A2:A5)
9		

7. 得られた回帰定数 (a) と回帰係数 (b) を用いて、試料の濃度を推定するには、(吸光度 $-a$)/ b を計算します。Excel のオートフィル機能と絶対参照を使うと複数試料の濃度

を簡単に算出できます。右図のように B 列の 7 行目 (Excel では入力単位をセルと呼び、セル位置の列と行の組合せで例えば B7 と表記する) に回帰係数を、B 列の 8 行目 (B8 セル) に回帰定数を計算した場合を考えます。1 つめの試料の吸光度を A12 セルに入力しておき、濃度の推定値を入力する B12 セルには、

$$= (A11-B\$8)/B\$7 \quad \text{式 1}$$

という式を入力します。このまま Enter キーを押すと計算結果が B12 に表示されます。式中の A12 や B\$9 はそのセル位置に入力されている値の代入を意味します (参照という)。

	A	B	C
1	濃度(x)	吸光度(y)	
2	1	0.394	
3	5	0.508	
4	10	0.701	
5	15	0.869	
6			
7	回帰係数	0.034501129	
8	回帰定数	0.350616253	
9			
10	試料の吸光度	濃度の推定値	
11	0.504	$= (A11-B\$8)/B\7	
12	0.453		
13	0.395		

【相対参照と絶対参照について】 Excel ではセルを列と行で指定し、セルに記入された値を参照します。つまり式 1 の A11 は A 列の 11 行目の値 (=0.504) を意味します。式 1 が上図のように B11 セルに入力されているとき、B11 セルをコピーして B12 または B13 にペーストすると、

$$= (A12-B\$8)/B\$7 \quad \text{(B12 セル)}$$

$$= (A13-B\$8)/B\$7 \quad \text{(B13 セル)}$$

とそれぞれ入力され、式 1 の A11 が A12 および A13 に自動的に変わります (相対参照)。例えば、A1 を相対参照する或るセルを+a 列+b 行のセルにペーストすると参照内容はセルの移動に合わせて A+a 列目の 1+b 行目のセルに変わります。

	A	B	C	D	E
1	A1セルの値				
2					
3			=A1		
4					
5					
6					

→

	A	B	C	D	E	F
1	A1セルの値					
2						
3				A1セルの値		
4						
5						=C3
6						

相対参照例。D3セルがA1セルを相対参照するとき、D3セルをF5セルにコピー&ペーストすると、参照先はC3になる (ペースト先が右に2つ下に2つずれると、参照先も右に2つ下に2つずれる)。

式 1 の B\$8 や B\$7 のセルが移動しないのは、列や行の前に\$を付けて絶対参照しているためです (下図)。

	A	B	C	D	E
1	A1セルの値				
2					
3			=\$A\$1		
4					
5					
6					

→

	A	B	C	D	E	F
1	A1セルの値					
2						
3				A1セルの値		
4						
5						=\$A\$1
6						

絶対参照の例。D3セルのA1セルへの参照時、行や列の前に\$を付けると絶対参照になる(\$A\$1)。絶対参照の場合、セルをコピーしてペーストしても、ペースト先(例えばF5)でも参照セルは\$A\$1のみである。

絶対参照と相対参照を駆使すると、推定した検量線を用いた複数サンプルの定量が可能になります。その際、オートフィル機能が便利です。式1を入力済みの B12 セルを選びマウスカursorをセルの右下角に置くと+マークに代わるので、そのまま B13、B14、、、へドラッグすると、A 列の同一行の吸光度に対応した濃度が B 列に計算されます。

	A	B
6		
7	回帰係数	0.034501129
8	回帰定数	0.350616253
9		
10	試料の吸光度	濃度の推定値
11	0.504	4.445760272
12	0.453	2.967547762
13	0.395	1.286443339
14		

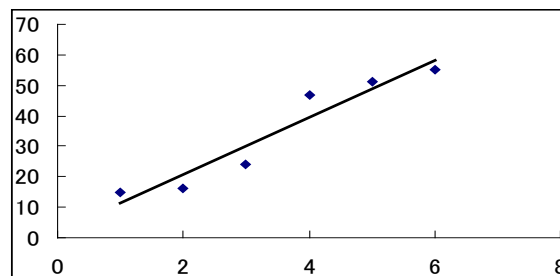
8. 注意点

検量線は 4~5 点で作成し、その際の R^2 は 0.98 以上が望ましいです (0.95 以下の場合、希釈しなおすこと。測定機器や試薬に問題があるかもしれません)。

最小 2 乗法による回帰直線の推定

例として、温度 x_i の変化にともなう或る化合物の生成量 y_i を考えてみよう。この場合、温度は自分で制御可能な変数である。 (x_i, y_i) という表記で実験から選ばれた 6 個のデータ(実測値)を表すと、(1, 15) (2, 16) (3, 24) (4, 47) (5, 51) (6, 55) となったとする。

各データのプロットは右図の通り。



更に図には回帰直線 $y = \alpha + \beta x$ が示されている。この直線は、どの点 (x_i, y_i) からの y 座標のズレもなるべく小さくするように選ばれている。 y 座標のズレは $y_i - \alpha - \beta x_i$ なので、

$$S = \sum_{i=1}^6 (y_i - \alpha - \beta x_i)^2 \quad \cdots \text{①}$$

としたときの S を最小にする α と β を求める。 S を α と β でそれぞれ偏微分し、

$$\frac{\partial S}{\partial \alpha} = 0 \quad \frac{\partial S}{\partial \beta} = 0 \quad \text{つまり、} \quad \frac{\partial S}{\partial \alpha} = -2 \sum_{i=1}^6 (y_i - \alpha - \beta x_i) = 0 \quad \frac{\partial S}{\partial \beta} = -2 \sum_{i=1}^6 x_i (y_i - \alpha - \beta x_i) = 0$$

この連立方程式の解 $\hat{\alpha}, \hat{\beta}$ は S を最小にする。これが最小 2 乗法である。

上の問題は、設定した温度条件下で生成される化合物の量は偶然誤差によってばらつくとするモデル $y_i = \alpha + \beta x_i + \varepsilon_i$ $E(\varepsilon_i) = 0, V(\varepsilon_i) = \sigma^2$ で、誤差を最小にする推定量を考える

ことと同じである。具体的には、

$$\begin{aligned} y_1 &= \alpha + \beta x_1 + \varepsilon_1 \\ y_2 &= \alpha + \beta x_2 + \varepsilon_2 \\ &\vdots \\ y_6 &= \alpha + \beta x_6 + \varepsilon_6 \end{aligned} \quad \text{で、これをベクトルと行列で表示すると} \quad \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_6 \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_6 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_6 \end{bmatrix}$$

となる。ここで、 $\mathbf{A} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_6 \end{bmatrix}$ 、 $\boldsymbol{\mu} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ とすると、の式は $\mathbf{Y} = \mathbf{A}\boldsymbol{\mu} + \boldsymbol{\varepsilon}$ …②と書ける。

先の①式は、 $S = (\mathbf{Y} - \mathbf{A}\boldsymbol{\mu})' (\mathbf{Y} - \mathbf{A}\boldsymbol{\mu})$ …③と書き直せる。 S を $\boldsymbol{\mu}$ の各成分で偏微分して 0 とおくと、③は、 $\mathbf{A}'\mathbf{A}\hat{\boldsymbol{\mu}} = \mathbf{A}'\mathbf{Y}$ …④となる。④は正規方程式と呼ばれる。②式の中身を、

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1s} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{ns} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad \text{という一般的な形にしても正規方程式④は同じ形のまま}$$

までである。 \mathbf{A} を計画行列、 $\boldsymbol{\mu}$ は未知母数で、 $\boldsymbol{\varepsilon}$ は誤差項である。