数値解析　第６回

<u>最小２乗法</u>

観測値（実験）		推算式（理論）
(x_i, y_i) , $(i = 1,2, \dots , N)$	⇔	$y = f(x)$
N回の観測		$f(x)$は実験で定めるべき n 個の未知量 a_j , $(j = 1,2, \dots , n)$で表されるとする。

x_i に対応する推算値を f_i と書くことにする。$f_i = f(x_i)$ であり，f_iは a_j , $(j = 1,2, \dots , n)$ の関数でもあるから，$f_i = f_i(a_1, a_2, \cdot\cdot\cdot, a_n)$ と書くことができる。

ここで，強引に $y_i = f_i$ とおくことにすると，$N < n$ の場合は，未知数の数に対して方程式の数が不足するので a_j , $(j = 1,2, \dots , n)$ の値は決定できないが，$N = n$ であれば，連立方程式を解くことで a_jの値を決定できる。$N > n$ の場合は，N 組の観測結果から n 組のデータを選んで連立方程式を作れば a_j の値が計算できる。ただし，選んだデータによって a_j の計算結果が異なるので，n 組の抽出データによる a_j の計算を複数回行って平均をとることで，a_j の値を最確値に近づけることができる。a_j の凡その値を知りたいのであればこの方法で十分であるが，与えられた観測結果における a_j の最確値を知りたい場合には，全てのデータの組み合わせについて計算しなければならない。この計算は N が小さければ実行可能だが，N が大きい場合には計算量が膨大になり現実的でない。

$N > n$ の場合，以下に述べる最小２乗法が有効である。

観測値と推算値の差を残差 ρ とすると各観測における残差は，

$$\rho_i = y_i - f_i \ , (i = 1,2, \dots , N)$$

残差の２乗和 S は，

$$S = \sum_{i=1}^{N} \rho_i^2$$

で計算される。この S を最小にする a_j , $(j = 1,2, \dots , n)$ が見つかれば，その a_j が与えられた観測結果における最確値である。

一般には，S が最小になる条件が各 a_j について $\partial S/\partial a_j = 0$ となることから，

$$\sum_{i=1}^{N} \rho_i \frac{\partial \rho_i}{\partial a_j} = 0 \quad , (j = 1,2, \dots , n)$$

となることを利用して解析的に解くことが可能である。特に，f_i が a_j の線形関数である場合，a_j の連立１次方程式が得られ，容易に a_j の最確値が求められる。（教科書 p102～109，最小二乗近似）

今回は，これとは別に a_j の値を少しずつ変えながら S の値を計算し S を最小にする a_j を探すことで最確値を求める。実用的には a_j の最確値は無限に細かな桁まで求める必要はなく，計算は単純な計算手続きの繰り返しなので，コンピュータを用ることで，最確値を求められる。

【練習問題】

次のデータに対して最小２乗法を利用して１次関数 $y = a\,x$ のあてはめを行い，a の値を小数点以下３桁で求める。ヒント： $1.45 \le a \le 1.55$ の範囲に解は存在する。

（データ）

i	x	y
1	1.01	1.53
2	2.02	3.03
3	3.00	4.51
4	3.99	5.99
5	5.01	7.52

残差の２乗和は,

$$S = \sum_{i=1}^{5} \left(y_i - a\,x_i \right)^2$$

で計算する。

(1) 初期値の設定

$S_{\min} = 10.0$ （大きな数なら何でもよい）と置く。

$a_{\mathrm{f}} = 1.45$（予想される a の範囲の最小値）と置く。

$a = a_{\mathrm{f}}$ と置く。

$\delta a = 0.001$（必要な精度に対応する a の刻み幅）と置く。

(2) 残差の２乗和 S の値を計算する。

(3) $S_{\min}$，a_{f} の更新

$S_{\min} > S$ の場合： $S_{\min}$ を S で，a_{f} を a で 置き換える。

$S_{\min} \le S$ の場合： $S_{\min}$，a_{f} はそのまま。

(4) a の更新

$a < 1.55$ （予想される a の範囲の最大値）の場合：

a を $a + \delta a$ で 置き換えて，(2) に戻って計算を繰り返す。

$a \ge 1.55$ の場合： a_{f} （必要なら $S_{\min}$ も）を表示して終了する。

【未知量が２つの場合】

上記練習問題のデータに対して最小２乗法を利用して１次関数 $y = a\,x + b$ のあてはめを行い，a，b の値を小数点以下３桁で求める。ヒント： $1.45 \le a \le 1.55$， $-0.5 \le b \le 0.5$ の範囲に解は存在する。

残差の２乗和は,

$$S = \sum_{i=1}^{5} \left\{ y_i - f\left(x_i\right) \right\}^2$$

ただし， $f\left(x\right) = a\,x + b$

で計算する。

(1) 初期値の設定

$S_{\min} = 10.0$ （大きな数なら何でもよい）と置く。

$a_{\min} = 1.45$（予想される a の最小値），

$a_{\max} = 1.55$（予想される a の最大値）と置く。

$a_{\mathrm{f}} = a_{\min}$ と置く。

$a = a_{\mathrm{f}}$ と置く。

$\delta a = 0.001$（必要な精度に対応する a の刻み幅）と置く。

$b_{\min} = -0.5$（予想される b の最小値），

$b_{\max} = 0.5$（予想される b の最大値）と置く。

$b_{\mathrm{f}} = b_{\min}$ と置く。

$b = b_{\mathrm{f}}$ と置く。

$\delta b = 0.001$（必要な精度に対応する b の刻み幅）と置く。

(2) 残差の２乗和 S の値を計算する。

(3) $S_{\min}$，a_{f}，b_{f} の更新

$S_{\min} > S$ の場合： $S_{\min}$ を S で，a_{f} を a で，b_{f} を b で 置き換える。

$S_{\min} \le S$ の場合： $S_{\min}$，a_{f}，b_{f} はそのまま。

(4) b の更新

$b < b_{\max}$ の場合：b を $b + \delta b$ で 置き換えて，(2) に戻って計算を繰り返す。

$b \ge b_{\max}$ の場合：b を $b_{\min}$ で 置き換えて， (5) へ進む。

(5) a の更新

$a < a_{\max}$ の場合：a を $a + \delta a$ で 置き換えて，(2) に戻って計算を繰り返す。

$a \ge a_{\max}$ の場合： a_{f}，b_{f} （必要なら $S_{\min}$ も）を表示して終了する。