三次元座標計測（第4回）
2005年度大学院講義
2005年11月15日

高増 潔
東京大学工学系研究科
精密機械工学専攻
E-mail: takamasu@pe.u-tokyo.ac.jp
HP: http://www.nano.pe.u-tokyo.ac.jp/

追加

RSS
アボガドロ数
球の体積，水の密度
三角関数の近似

二乗和の平方

- RSS: root sum square 二乗和の平方根（二乗和平方根）
- RSS: root-sum-square value 二乗和平方根値
- RMS: root mean square 二乗平均平方根（二乗平均根）
- rms: root mean square value 根二乗平均実効値
- ppm: parts per million 100万分の1，10^{-6}
- ppb: parts per billion 10億分の1，10^{-9}
- average, mean 平均
- variance 分散
- SD: standard deviation 標準偏差

$$\mathrm{RSS} = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

$$\mathrm{RMS} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} = \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

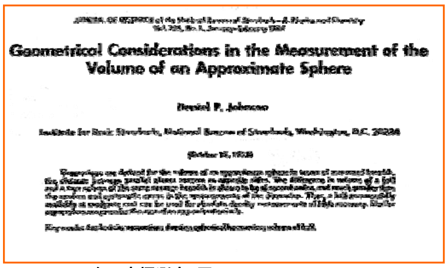
$$\sigma_x^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

基礎物理定数の不確かさ

- アボガドロ定数
  - 1モル（1グラム分子）の物質中に存在する分子の数．NAまたはLの記号であらわす．
  - 12 g の炭素12（C_{12}）の中に存在する原子数である．
  - 同体積の気体は同温，同圧のもとでは同数の分子をふくむという仮説をたてた(→ アボガドロの法則).
  - Amedeo Avogadro 1776～1856
  - 1986年 6.0221367 ± 0.0000035 x 10^{23}
    6.0221332 ～ 6.0221402 x 10^{23}
  - 1973年 6.022045 ± 0.00003 x 10^{23}
    6.022015 ～ 6.022075 x 10^{23}
- 磁束量子，ファラデー定数，素電荷，プランク定数，電子の静止質量なども同じく，1973年と1986年では不確かさを考慮しても整合しない．

アボガドロ数の精密測定

- 完全結晶シリコン球体の密度
  - 球体の直径を光波干渉測定する.
  - 体積を精密に決定する.
  - X線干渉技術でシリコンの格子定数を測定する.
  - シリコンの同位体組成比を求める.
  - アボガドロ数を長さ及び質量の一次標準（原器）に照らし合わせて求めることができる.


Geometrical Considerations in the Measurement of the Volume of an Approximate Sphere

光波干渉計による球の直径測定

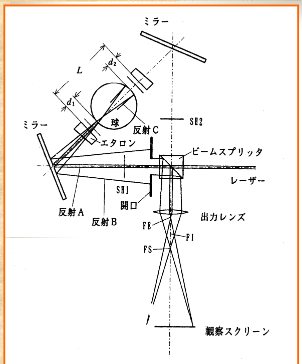
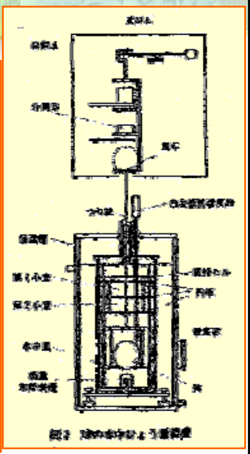


図１ 光波干渉計による球の直径測定の原理図

水の密度測定

表1 世界各地の降水の同位体組成(年間平均)と密度。第4欄は降水の密度(ρ)と標準平均海水(SMOW)の密度(ρ_{SMOW})との差を示す。SMOWの密度は標準海水の密度にだいたい等しい

地名	$\delta_O \times 10^3$	$\delta_D \times 10^3$	$\rho-\rho_{SMOW}$ (ppm)
ボンベイ	−1.5	−7	−0.5
ハワイ	−2.5	−8	−0.7
リオ・デ・ジャネイロ	−4	−26	−1.3
ニューデリー	−4	−27	−1.4
マニラ	−5	−26	−1.6
メルボルン	−5	−29	−1.6
ジャカルタ	−6	−33	−1.9
カラチ	−6	−35	−2.0
東京	−6	−33.5	−2.8
ウィーン	−8.5	−54.5	−3.1
香港	−6	−37	−2.1
シュツットガルト	−10	−76	−3.6
シカゴ	−10.5	−77	−3.7
エドモントン	−19.5	−141	−6.7
ノア	−21	−158	−7.5
ホワイトホース	−22	−174	−8.0



2005/11/15 三次元座標測定4回 7

三角関数の近似

x(rad)	x(度)	sin(x)	cos(x)	tan(x)
0.0001	0.006	0.000100000000	0.999999995000	0.000100000000
0.001	0.057	0.000999999833	0.999999500000	0.001000000333
0.01	0.573	0.009999833334	0.999950000417	0.010000333347
0.1	5.730	0.099833416647	0.995004165278	0.100334672085

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \approx x$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \approx 1 - \frac{x^2}{2} \approx 1$$

$$\tan x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \cdots \approx x$$

$x \approx 0, y \approx 0$

$$\sqrt{1+x} \approx 1 + \frac{x}{2}$$

$$\frac{1}{1+x} = \frac{1-x}{1-x^2} \approx 1-x$$

$$\frac{y}{1+x} = \frac{y(1-x)}{1-x^2} \approx y(1-x) \approx y$$

2005/11/15 三次元座標測定4回 8

最小二乗法入門

パラメータの設定
線形モデルの最適パラメータの決定

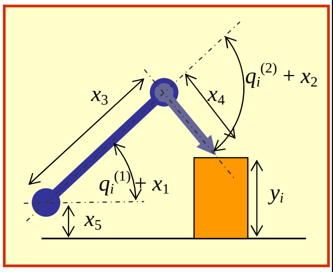
パラメータの設定(1)

- ベクトルや行列は太字で書く.
  量(変化する)を表す記号はイタリックで書く(ベクトルや行列の場合は, イタリックにしないこともある), 真の値は右肩に0をつける
  数字(1, 2. . .)はイタリックにしない. 太字にもしない.
- 測定値:さし当たって測定値はスカラーとする.
  *n*個の測定値が得られた. それぞれの測定値は, それぞれの測定条件*q*に よって区別される.
  $\boldsymbol{y} = \{y_1, y_2, \ldots y_n\}$
- 測定条件を指定する量(横座標):さし当たって誤差がないとする
  $\boldsymbol{q}_i = \{q_i^{(1)}, q_i^{(2)}, \ldots q_i^{(l)}\} = \boldsymbol{q}_i^0$
- 測定条件と測定値の間には, モデル $f(\ )$ がある. モデルそのものは正しいとしておく(真のモデル $f_0(\ )$). モデルが持つパラメータは分かっていない.
- 真のパラメータの組
  $\boldsymbol{x}_o = \{x_1^0, x_2^0, \ldots x_m^0\}$
- 真のモデル, 真のパラメータが分かり, 真の測定条件を与えれば, その測定条件における真の測定値が得られる.
  $y_i^0 = f^0(q_i^{0(1)}, \ldots q_i^{0(l)}; x_1^0, \ldots x_m^0) = f^0(\boldsymbol{q}_i^0; \boldsymbol{x}^0) = f_i^0(\boldsymbol{x}^0)$

2005/11/15 三次元座標測定4回 10

パラメータの設定(2)

- 平面2関節機構による高さ測定機
- *i* 番目の測定に対応した
- 測定値:高さ y_i
- 横座標(測定量):エンコーダの読み1と2 $q_i^{(1)} = q_i^{0(1)}, q_i^{(2)} = q_i^{0(2)}$
- パラメータ:運動学パラメータ
  エンコーダのオフセット x_1, x_2
  腕の長さ1と2 x_3, x_4
  原点の高さのオフセット x_5
- モデル:真のモデル
  y_i^0
  $= f^0(q_i^{0(1)}, q_i^{0(2)}; x_1^0, x_2^0, x_3^0, x_4^0, x_5^0)$
  $= x_5^0$
  $+ x_3 \sin(q_i^{0(1)} + x_1^0)$
  $+ x_4^0 \sin(q_i^{0(1)} + x_1^0 + q_i^{0(2)} + x_2^0)$



2005/11/15 三次元座標測定4回 11

パラメータの設定(3)

- 測定値: *n* 個(測定回数などで決まる), 添え字 *i*
  $\boldsymbol{y} = \{y_1, y_2, \ldots y_n\}$
- 横座標: *l* 個(モデルで決まる), 添え字 *j*
  $\boldsymbol{q}_i = \{q_i^{(1)}, q_i^{(2)}, \ldots q_i^{(l)}\} = \boldsymbol{q}_i^0$
- 真のパラメータの組: *m* 個(モデルで決まる)
  $\boldsymbol{x}_o = \{x_1^0, x_2^0, \ldots x_m^0\}$
- 真のモデル:
  $y_i^0 = f^0(q_i^{0(1)}, \ldots q_i^{0(l)}; x_1^0, \ldots x_m^0) = f^0(\boldsymbol{q}_i^0; \boldsymbol{x}^0) = f_i^0(\boldsymbol{x}^0)$

2005/11/15 三次元座標測定4回 12

最適値の条件

- 実際の測定値の組 $\{y_i\}$ とその誤差分布 $\{s_i\}$ から，できるだけ真のモデル式と真のパラメータの推定値 $\{x^*{}_j\}$ を決定したい．
- 最小二乗条件
  $\Sigma\,[y_i - f_i(\boldsymbol{x})]^2 / s_i^2 = \min$
- 残差
  $v_i = y_i - y_i^* = y_i - f_i(x^*)$
- 誤差分布の考え方
  - すべての測定の誤差は等しい⇒誤差分布を考えなくてよい
  - 誤差分布が分かっている⇒誤差を単位付きで指定する
  - 誤差の比のみが分かっている⇒比で考える

最小二乗法の前提と原理

1. 測定値の誤差には偏りがない．
2. 測定値の誤差の分散は既知である
3. 各測定は互いに独立であり，共分散はゼロとする．
4. 誤差の分布形は正規分布（Gauss分布）である．
5. パラメータ（m個）を含むモデルfが知られている．

- 以上の前提のもとで，パラメータを推定する最尤推定法が最小二乗法である．
- いくつかの条件は拡張できる
  - 2で分散の相対比が既知ならよい．
  - 3で独立でなくても，相関が分かればよい．
  - 4で正規分布でないときは，最尤推定法ではなくなるが，線形不偏推定法のうちでは最良となる．

線形モデルでの最適パラメータの決定

- モデルがパラメータの一次結合で表される場合
  - 横座標は測定するときに，実際の値が求まっているので，線形でもなんでも計算ができるので，線形の必要はない．
  - $f\,(q_i^{(1)}, \ldots\, q_i^{(l)}; x_1, x_2, \ldots, x_m)$
    $= f_i\,(x_1, x_2, \ldots, x_m)$
    $= A_{i1}\,x_1 + A_{i2}\,x_2 + \ldots + A_{im}\,x_m$
    $= \Sigma\,A_{ij}\,x_j\,(i = 1 \sim n)$
  - A_{ij} は横座標 $q_i^{(1)} \sim q_i^{(l)}$ で表される既知の定数である．計算はどんなに複雑でも計算できればよい．パラメータには依存しない．
- 最小二乗条件を適応
  - $\Sigma\,[y_i - f_i(\boldsymbol{x})]^2 / s_i^2 = \min$
    $\Sigma\,[y_i - \Sigma\,A_{ij}\,x_j]^2 / s_i^2 = \min$

評価が最小となるパラメータ

$$S(x) = \sum_{i=1}^{n} \frac{\left(y_i - \sum_{j=1}^{m} A_{ij} x_j\right)^2}{s_i^2} = \min$$

$$0 = \frac{\partial S(x)}{\partial x_j} = -2\sum_{i=1}^{n} v_i A_{ij} / s_i^2 = -2\sum_{i=1}^{n} \frac{y_i - \sum_{j'=1}^{m} A_{ij'} x_{j'}}{A_{ij} / s_i^2}$$

$$\sum_{j'=1}^{m}\sum_{i=1}^{n} \frac{A_{ij'} A_{ij}}{s_i^2} x_{j'} = \sum_{i=1}^{n} \frac{A_{ij}}{s_i^2} y_i$$

連立一次方程式

$$\begin{cases} B_{11}x_1 + B_{12}x_2 + \cdots + B_{1m}x_m + = b_1 \\ B_{21}x_1 + B_{22}x_2 + \cdots + B_{2m}x_m + = b_1 \\ \vdots \\ B_{m1}x_1 + B_{m2}x_2 + \cdots + B_{mm}x_m + = b_1 \end{cases}$$

$$B_{jj'} = \sum_{i=1}^{n} \frac{A_{ij'} A_{ij}}{s_i^2}$$

$$b_j = \sum_{i=1}^{n} \frac{A_{ij}}{s_i^2} y_i$$

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ \hat{x}_m \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1m} \\ B_{21} & B_{22} & \cdots & B_{2m} \\ \vdots & \vdots & & \vdots \\ B_{m1} & B_{m2} & \cdots & B_{mm} \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

行列での表現

- 観測方程式y，誤差行列S

$$\mathbf{y} = \mathbf{f}(\mathbf{q};\mathbf{x}) = \mathbf{Ax}$$

$$\mathbf{S} = \begin{pmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_n^2 \end{pmatrix}, \qquad \mathbf{W} = \mathbf{S}^{-1}$$

- ヤコビ行列A

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & A_{1m} \\ A_{21} & A_{22} & A_{2m} \\ & & \\ A_{n1} & A_{n2} & A_{nm} \end{pmatrix}, \quad A_{ij} = \frac{\partial f_i(\mathbf{x})}{\partial x_j} \quad (i = 1 - n, j = 1 - m)$$

- 最小二乗解（パラメータの値）
  - 誤差の重みを考慮する場合

$$\mathbf{B} = \mathbf{A}^t\mathbf{WA}$$
$$\mathbf{b} = \mathbf{A}^t\mathbf{Wy}$$
$$\hat{\mathbf{x}} = (\mathbf{A}^t\mathbf{WA})^{-1}(\mathbf{A}^t\mathbf{Wy}) = \mathbf{B}^{-1}\mathbf{b}$$

誤差に重みがない場合

$$\mathbf{B} = \mathbf{A}^t\mathbf{A}$$
$$\mathbf{b} = \mathbf{A}^t\mathbf{y}$$
$$\hat{\mathbf{x}} = (\mathbf{A}^t\mathbf{A})^{-1}(\mathbf{A}^t\mathbf{y}) = \mathbf{B}^{-1}\mathbf{b}$$

最小二乗法の計算プログラム

線形最小二乗法
非線形最小二乗法
円の求め方

直線の線形最小二乗法

- 観測方程式：パラメータ：傾き(x1)と切片(x2)

$$d = x_1 q + x_2 - y$$

- 偏微分：パラメータで偏微分する

$$\frac{\partial d}{\partial x_1} = q, \quad \frac{\partial d}{\partial x_2} = 1 \text{ なので,}$$

- ヤコビ行列A

$$\mathbf{A} = \begin{pmatrix} q_1 & 1 \\ \vdots & \vdots \\ q_n & 1 \end{pmatrix}, \quad \mathbf{d} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \text{となる.}$$

- 最小二乗解(パラメータの値)

$$\hat{\mathbf{x}} = (\mathbf{A}^t\mathbf{A})^{-1}(\mathbf{A}^t\mathbf{d})$$

線形最小二乗法(1次関数)

```
function testlsm1()
% 線形最小二乗法の例1
% 一次関数 y = x1 * q + x2
% 横座標に誤差はない, 測定値の誤差はすべて同じ
qi1 = [10 20 30 40 50]';
yi = [12 23 28 38 55]';
n = length(yi);
% A:ヤコビ行列, B:正規行列
A = [qi1 ones(n, 1)]
B = A' * A;
b = A' * yi;
x = inv(B) * b;
y = x(1) .* qi1 + x(2);
plot(qi1, yi, 'o', qi1, y, '-');
```

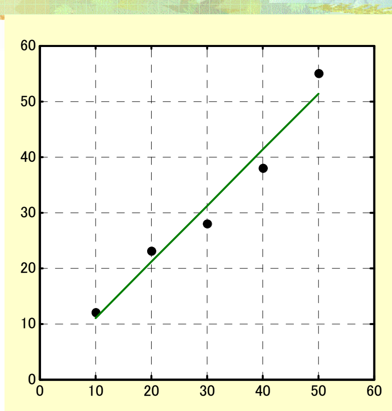
$$\mathbf{A} = \begin{bmatrix} 10 & 1 \\ 20 & 1 \\ 30 & 1 \\ 40 & 1 \\ 50 & 1 \end{bmatrix}$$

$$\mathbf{B} = \mathbf{A}^t\mathbf{A} = \begin{bmatrix} 5500 & 150 \\ 150 & 5 \end{bmatrix}$$

$$\mathbf{B}^{-1} = \begin{bmatrix} 0.001 & -0.030 \\ -0.030 & 1.100 \end{bmatrix}$$

$$\mathbf{b} = \mathbf{A}^t\mathbf{y} = \begin{bmatrix} 5690 \\ 156 \end{bmatrix}$$

$$\hat{\mathbf{x}} = (\mathbf{A}^t\mathbf{A})^{-1}(\mathbf{A}^t\mathbf{y}) = \mathbf{B}^{-1}\mathbf{b} = \begin{bmatrix} 1.010 \\ 0.900 \end{bmatrix}$$



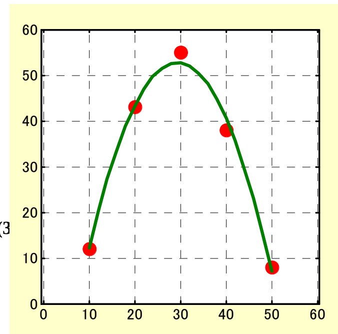
Matlabの関数

- ベクトルの代入
  qi1 = [10 20 30 40 50]';
- 転置
  yi = [12 23 28 38 55]';
- ベクトルの大きさ
  n = length(yi);
- 行列の大きさ
  [n, m]=size(xx);
- 行列の代入, 要素が1の行列
  A = [qi1 ones(n, 1)]
- 行列の積
  B = A' * A;
- 逆行列
  x = inv(B) * b;
- グラフ
  plot(qi1, yi, 'o', qi1, y, '-');
- 連続値
  q = [10:2:50]';
- 対角要素
  S = diag([2 2 5 5 10])

線形最小二乗法(2次関数)

```
function testlsm2()
qi1 = [10 20 30 40 50]';
yi = [12 43 55 38 8]';
n = length(yi);
A = [qi1.^2 qi1 ones(n, 1)];
B = A' * A;
b = A' * yi;
x = inv(B) * b;

q = [10:2:50]';
y = x(1) .* q .^ 2 + x(2) .* q + x(3
plot(qi1, yi, 'o', q, y, '-');
axis equal;
```

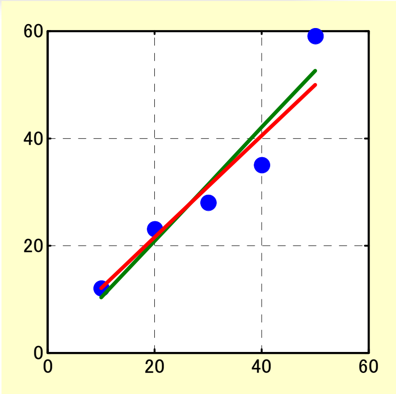


誤差を考慮した場合

```
qi1 = [10 20 30 40 50]';
yi = [12 23 28 35 59]';
S = diag([2 2 5 5 10])
W = inv(S);
n = length(yi);

A = [qi1 ones(n, 1)];
B1 = A' * A;
b1 = A' * yi;
x1 = inv(B1) * b1;

B2 = A' * W * A;
b2 = A' * W * yi;
x2 = inv(B2) * b2;
```



非線形最小二乗法の条件

- ガウス・ニュートン法
- 非線形性が小さい
- 横座標に対する測定値が正しく求まる
- 偏微分係数が大体求まる：数値的でもよい
- 初期値がかなり正しく求まる：数値的や別の近似でもよい

- 機械などの分野では，かなり上記のことが成り立つ．

円の非線形最小二乗法

- 観測方程式：パラメータ：中心（x, y）と半径rの3つ

$$d=\sqrt{(x-x_0)^2+(y-y_0)^2}-r$$

- 偏微分：パラメータで偏微分する

$$\frac{\partial d}{\partial x_0}=\frac{-(x-x_0)}{\sqrt{(x-x_0)^2+(y-y_0)^2}},\quad \frac{\partial d}{\partial y_0}=\frac{-(y-y_0)}{\sqrt{(x-x_0)^2+(y-y_0)^2}},\quad \frac{\partial d}{\partial r}=-1\ \text{なので,}$$

- ヤコビ行列A

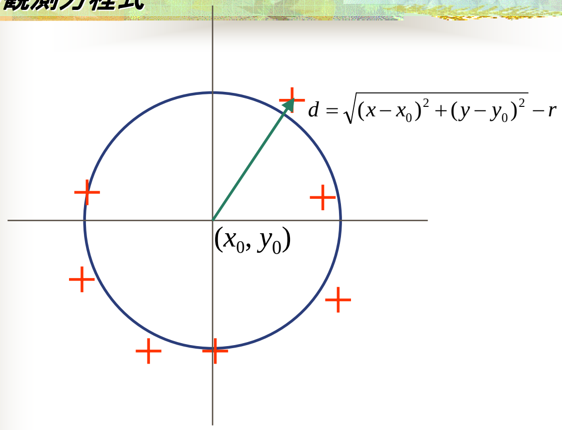
$$\mathbf{A}=\begin{pmatrix}-(x_1-x_0)/r_1 & -(y_1-y_0)/r_1 & -1\\ \vdots & \vdots & \vdots\\ -(x_n-x_0)/r_n & -(y_n-y_0)/r_n & -1\end{pmatrix},\quad \mathbf{d}=\begin{pmatrix}r_1-r\\ \vdots\\ r_n-r\end{pmatrix}\text{となる.}$$

$$\text{ここで,}\ r_i=\sqrt{(x_i-x_0)^2+(y_i-y_0)^2}$$

- 線形近似

$$(x_i-x_0)^2+(y_i-y_0)^2-r^2=-2x_ix_0-2y_iy_0+(x_0^2+y_0^2-r^2)+(x_i^2+y_i^2)$$
$$=-2x_ix_0-2y_iy_0+p+(x_i^2+y_i^2)$$

円の観測方程式



二次曲面で円の初期値を求める

```
function [xy0, r] = lsmCircle2dinit(xx)
[n, m] = size(xx);

% 2次曲面近似のヤコビ行列a，測定値ベクトルbを作る
a = [2*xx(:, 1), 2*xx(:, 2), -1*ones(n, 1)];
b = xx(:, 1).^2 + xx(:, 2).^2;

% 最小二乗法で計算
p = inv(a'*a) * (a'*b);
xy0 = p(1:2);
r = sqrt(p(1)^2 + p(2)^2 - p(3));
```


円の非線形最小二乗法のメイン

```
function [xy0, r]=lsmCircle2d(xx)
LSMMAXLOOP = 20;LSMCONVERGED = 0.00000000001;
[n, m]=size(xx);
% 初期値をlsmCircle2dinitで決定する
[xy0, r] = lsmCircle2dinit(xx);
% 更新量がLSMCONVERGED以下になるまで繰り返す
for ii = 1:LSMMAXLOOP;
  [ll, xy0,r] = lsmCircle2dRe(xx, xy0, r);
  se = lsmCircle2dSe(xx, xy0, r);
if ll < LSMCONVERGED;
    break
  end
end
```

円の1回分の非線形最小二乗法

```
function [ll, xy0, r] = lsmCircle2dRe(xx, xy0, r)
[n, m] = size(xx);
% ヤコビ行列j，測定値行列dを計算
x1 = xx(:, 1) - xy0(1);
x2 = xx(:, 2) - xy0(2);
ri = sqrt(x1.^2 + x2.^2);
d = r*ones(n, 1) - ri;
j = [-x1./ri, -x2./ri, -1*ones(n, 1)];
% 最小二乗解pを計算して更新
p = inv(j'*j) * (j'*d);
ll = sqrt(sum(p.*p));
xy0 =xy0 + p(1:2);
r = r + p(3);
```


測定値の重みつき平均

測定値の重みつき平均

- 棒の長さをマイクロメータとノギスで測定したとき，それぞれの測定値 y_1 と y_2 から，どのようにして測定結果を求めることがより精度の高い求め方になるのかを考える．
- それぞれの測定機の誤差にはかたよりがなく，誤差は正規分布となり，分散 s_1 および s_2 が分かっているとする．
- このような場合の，測定値の誤差は次のように計算できる．

$$y = k_1 y_1 + k_2 y_2 \quad \bar{y} = k_1 \bar{y}_1 + k_2 x\bar{y}_2$$
$$s^2 = \overline{(y-\bar{y})^2} = k_1^2 s_1^2 + k_2^2 s_2^2 + 2k_1 k_2 s_{12}$$
$$= k_1^2 s_1^2 + k_2^2 s_2^2$$

共分散が0の場合

- 普通は，マイクロメータの測定とノギスの測定とは独立なので，s_{12} は0となり，y, y_1 および y_2 の平均は等しくなくてはならないので，$k_2 = (1 - k_1)$ である．
- このとき k_1 としては，s^2 が最小になるような値を選ぶべきであり，

$$s^2 = k_1^2 s_1^2 + (1-k_1)^2 s_2^2$$
$$\frac{\partial s^2}{\partial k_1} = 2k_1 s_1^2 - 2(1-k_1)s_2^2 = 0$$
$$k_1 = \frac{s_2^2}{s_1^2 + s_2^2} \qquad s^2 = \frac{s_1^2 s_2^2}{s_1^2 + s_2^2}$$

例と一般の形

- 具体的に，マイクロメータの分散が (3 μm)^2，ノギスの分散が (15 μm)^2 とすると，$k_1 = 0.9615$，$k_2 = 0.0385$ となり，s_2 は 8.6538 = (2.9417 μm)^2 となり，マイクロメータだけの場合より少し精度があがっている．
- 一般的に，n 個の測定値の重みつき平均（線形和）で測定値を表す場合

$$x = k_1 y_1 + k_2 y_2 + k_3 y_3 + \cdots + k_n y_n$$
$$k_1 + k_2 + k_3 + \cdots + k_n = 1$$
$$k_i = \frac{\frac{1}{s_i^2}}{\frac{1}{s_1^2} + \frac{1}{s_2^2} + \frac{1}{s_3^2} + \cdots + \frac{1}{s_n^2}} \qquad s^2 = \frac{1}{\frac{1}{s_1^2} + \frac{1}{s_2^2} + \frac{1}{s_3^2} + \cdots + \frac{1}{s_n^2}}$$

単純平均の場合

- 単にn 個のデータの単純平均を求めると，標準偏差は和の n 分の一となる．
- さらに，$s_1 = s_2 = ... = s_n$ の場合は，同じ分散を持つ測定結果を n 個の和をとると分散は n 分の一，標準偏差は，ルート n 分の一となる．

$$x = kx_1 + kx_2 + kx_3 + \cdots + kx_n$$
$$k = \frac{1}{n}$$
$$s^2 = \frac{s_1^2 + s_2^2 + s_3^2 + \cdots + s_n^2}{n^2}$$

$$s^2 = \frac{ns_1^2}{n^2} = \frac{s_1^2}{n}$$
$$s = \frac{s_1}{\sqrt{n}}$$

2つの測定値の分散を考慮した平均

- 表に，s_1（1に固定）と s_2（1から100のいくつかの場合）の組み合わせに対して，分散を考えた場合と単純平均を取った場合
- 単純平均の場合，k_1 は常に 0.5 である．
- 単純平均をすると，誤差の大きい測定値に引っ張られ，誤差が大きくなるので，誤差の大きい測定値は捨てたほうがよい．
- 分散を考慮すれば，誤差が大きい測定値でも少しは役に立つ．

s_1	s_2	k_1	分散を使ったs	単純平均s
1	1	0.500	0.707	0.707
1	2	0.800	0.894	1.118
1	4	0.941	0.970	2.062
1	10	0.990	0.995	5.025
1	100	0.9999	0.99995	50.002

数値計算の注意

標準偏差の計算(1)

- C言語によるアルゴリズム事典　奥村晴彦　技術評論社
- Numerical Recipes in C（日本語版）　技術評論社
- 数値計算の常識　伊理正夫，藤野和建　共立出版
- 計算手法
  - 普通に計算する．データを2度スキャンする必要がある．（平均を求めるときと標準偏差を求めるとき）
  - 変形した式を使うと1度のスキャンで平均と標準偏差を求められるが，大きい数（データの二乗和）と大きい数（平均の二乗のn倍）の差を取ることで桁落ちする．
  - 仮の平均を使うと1度のスキャンで求められ，誤差も少ない

$$\sigma_x^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n}(x_i^2-2\bar{x}x_i+\bar{x}^2)}{n-1} = \frac{\sum_{i=1}^{n}x_i^2-n\bar{x}^2}{n-1}$$

標準偏差の計算(2)

- 1回のスキャンで行う方法

```
s1 = s2 = n = 0;
while (scanf("%f", &x) == 1) {
    n++;
    s1 += x;
    s2 += x * x;
}
s2 = (s2 - n * s1 * s1) / (n - 1);
if (s2 > 0)
    s2 = sqrt(s2);
else
    s2 = 0;
```

- 仮の平均を使う方法

```
s1 = s2 = n = 0;
while (scanf("%f", &x) == 1) {
    n++;
    x -= s1;
    /* 仮の平均との差を取る */

    s1 += x / n;
    /* 仮の平均を更新 */

    s2 += (n - 1) * x * x / n;
    /* 平方和を計算 */
}
s2 = sqrt(s2 / (n - 1));
```


桁落ち(1)

- 小さいxに対する3つの計算方法
  - 結果の有効精度が変化する
- 絶対値がごく近い2つの数を足したり引いたりして結果の絶対値が小さくなるような計算をしてはいけない
- 多数の和を取るとき：非常に大きな和に小さい数を足すと積み残しが起こる．

$$x = 0.0031834$$

$$1-\frac{1}{\sqrt{1+x}} = 1-\frac{1}{\sqrt{1.0032}} = 1-\frac{1}{1.0016} = 1-0.99840 = 0.00160$$

$$\frac{\sqrt{1+x}-1}{\sqrt{1+x}} = \frac{\sqrt{1.0032}-1}{\sqrt{1.0032}} = \frac{1.0016-1}{1.0016} = \frac{0.0016}{1.0016} = 0.0015974$$

$$\frac{x}{(1+x)+\sqrt{1+x}} = \frac{0.0031834}{1.0032+1.0016} = \frac{0.0031834}{2.0048} = 0.0015879$$

桁落ち(2)

$$x \approx 0$$

$$\sqrt{1+x}-1 = \frac{x}{\sqrt{1+x}+1} \quad \text{（分子の有理化）}$$

$$x-\sin x = \frac{x^3}{3!}-\frac{x^5}{5!}+\frac{x^7}{7!}-\cdots \quad \text{（級数展開）}$$

$$\cos^2 x-\sin^2 x = \cos 2x$$

$$1-\cos x = 2\sin^2\frac{x}{2}$$

$$x = \frac{-b\pm\sqrt{b^2-4ac}}{2a} = \frac{2c}{-b\mp\sqrt{b^2-4ac}} \quad \text{（重根に近いとき）}$$