# 数値解析
# 第拾弐回　最小二乗法

舟木　剛

平成26年12月17日２限

# シラバス

- 授業の目的
  - 工学分野でよく用いられる数値計算の算法ならびにそれらの数値的な特性について理解させる.
- 授業計画
  - 数値計算と誤差（1回）
    - 数値計算の必要性ならびに誤差の種類について説明するとともに，行列式とその性質，行及び列の交換などの基本演算，逆行列の定義と求め方について説明する.
  - 代数方程式（2回）
    - ２分法，Newton-Raphson 法，Bailey 法について解説する．また，多変数に対するNewton-Raphson 法と収束に関する留意点について述べる.
  - 連立方程式（３回）
    - Gauss の消去法，Gauss-Jordan の掃き出し法，LU 分解法の手順ならびに計算複雑度について解説する．また，Jacobi 法，Gauss-Seidel 法，SOR 法などの反復法について手順並びに収束条件について解説する.
  - 行列の固有値（３回）
    - 固有値の性質，べき乗法，Householder 法，QR 法について述べる．QR 法の収束定理について解説する.
  - 補間法（２回）
    - 線形補間，Lagrange 補間，Newton 補間，スプライン補間について述べる．多項式補間については算出方法をスプライン補間については３次のスプライン関数の導出方法を説明する．また，自然なスプライン関数とその特徴について解説する.
  - 関数近似（１回）
    - 最小２乗法による関数の近似について解説する.
  - 数値積分（１回）
    - 台形公式，シンプソンの公式による数値積分について解説する.
  - 常微分方程式（１回）
    - 単区分法である Euler 法，修正Euler 法，Runge-Kutta 法ならびに複区分法である Adams-Moulton の予測子・修正子法について解説する．また，数値不安定性についても解説する.
- 教科書 佐藤・中村共著「よくわかる数値計算」日刊工業新聞社

# 最小二乗法

- $n$組のデータ$(x_i, y_i)\ i = 1,2,\cdots,n$に対する近似式$f(x)$を求める
  - $m$次多項式による近似
    $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$
    - $n = m + 1$の場合→全てのデータを補間点とする補間多項式（与えられたデータに対する誤差は0）
    - $n > m + 1$の場合→多項式の次数よりデータ点数の方が多い
      データと近似式の誤差が最小となる関数を求める
      - 誤差の最大値を最小にする→ミニマックス近似（最良近似）
      - 誤差の二乗和（残差平方和）を最小にする→最小二乗近似

# 最小二乗法

- $m$次近似多項式$f(x)$の誤差の二乗和（残差平方和）$S$を最小にする係数$a_0, a_1, \cdots, a_m$を求める
  - 近似多項式
    - $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$
  - 誤差の二乗和
    - $S = \sum_{i=1}^{n} \{f(x_i) - y_i\}^2$
      - 下に凸な関数
      - $S$を最小化する多項式の係数$a_j$ $(j = 0,1,2,\cdots,m)$の条件
        - » $\frac{\partial S}{\partial a_j} = 0 \qquad (j = 0,1,2,\cdots,m)$

# 1次関数の最小二乗近似

- 1次関数$f(x) = a_0 + a_1 x$の誤差の二乗和が最小となる係数$a_0, a_1$を求める
  - $S = \sum_{i=1}^{n}\{f(x_i) - y_i\}^2 = \sum_{i=1}^{n}\{a_0 + a_1 x_i - y_i\}^2$
    - $a_0$の条件
      - $\frac{\partial S}{\partial a_0} = \frac{\partial}{\partial a_0}\sum_{i=1}^{n}\{a_0 + a_1 x_i - y_i\}^2 = \sum_{i=1}^{n} 2\{a_0 + a_1 x_i - y_i\}$
        $= 2\left\{a_0 \sum_{i=1}^{n} 1 + a_1 \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i\right\} = 0$
      - $a_0 n + a_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$
    - $a_1$の条件
      - $\frac{\partial S}{\partial a_1} = \frac{\partial}{\partial a_1}\sum_{i=1}^{n}\{a_0 + a_1 x_i - y_i\}^2 = \sum_{i=1}^{n} 2x_i\{a_0 + a_1 x_i - y_i\}$
        $= 2\left\{a_0 \sum_{i=1}^{n} x_i + a_1 \sum_{i=1}^{n} {x_i}^2 - \sum_{i=1}^{n} x_i y_i\right\} = 0$
      - $a_0 \sum_{i=1}^{n} x_i + a_1 \sum_{i=1}^{n} {x_i}^2 = \sum_{i=1}^{n} x_i y_i$

# 1次関数の最小二乗近似

- 行列表現
  - $\begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} {x_i}^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{bmatrix}$
  - $\begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} {x_i}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{bmatrix}$
    $= \dfrac{1}{n \sum_{i=1}^{n} {x_i}^2 - \left\{ \sum_{i=1}^{n} x_i \right\}^2} \begin{bmatrix} \displaystyle\sum_{i=1}^{n} {x_i}^2 & -\displaystyle\sum_{i=1}^{n} x_i \\ -\displaystyle\sum_{i=1}^{n} x_i & n \end{bmatrix} \begin{bmatrix} \displaystyle\sum_{i=1}^{n} y_i \\ \displaystyle\sum_{i=1}^{n} x_i y_i \end{bmatrix}$
  - $y = a_0 + a_1 x$ → 回帰直線
    - $a_1$ → 回帰係数

# m次関数の最小二乗近似

- $m$次多項式での一般化
- 誤差の二乗和
  - $S = \sum_{i=1}^{n} \{a_0 + a_1 x_i + a_2 {x_i}^2 + \cdots + a_m {x_i}^m - y_i\}^2$
  - $S$を最小化する係数$a_j$ $(j = 1,2,\cdots,m)$の条件
    - $\frac{\partial S}{\partial a_j} = \sum_{i=1}^{n} 2{x_i}^j \{a_0 + a_1 x_i + \cdots + a_j {x_i}^j + \cdots + a_m {x_i}^m - y_i\} = 0$
    - $a_0 \sum_{i=1}^{n} {x_i}^j + a_1 \sum_{i=1}^{n} {x_i}^{j+1} + \cdots + a_j \sum_{i=1}^{n} {x_i}^{j+j} + \cdots + a_m \sum_{i=1}^{n} {x_i}^{j+m} = \sum_{i=1}^{n} {x_i}^j y_i$
    - 正規方程式（残差二乗和を最小にする推定値を与える方程式）

  - $$\begin{bmatrix} n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} {x_i}^2 & \cdots & \sum_{i=1}^{n} {x_i}^m \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} {x_i}^2 & \sum_{i=1}^{n} {x_i}^3 & \cdots & \sum_{i=1}^{n} {x_i}^{m+1} \\ \vdots & \vdots & & \ddots & \vdots \\ \sum_{i=1}^{n} {x_i}^m & \sum_{i=1}^{n} {x_i}^{m+1} & \sum_{i=1}^{n} {x_i}^{m+2} & \cdots & \sum_{i=1}^{n} {x_i}^{2m} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} {x_i}^2 y_i \\ \vdots \\ \sum_{i=1}^{n} {x_i}^m y_i \end{bmatrix}$$
  - $y = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$ → 回帰曲線

# 指数関数の最小二乗近似

- 近似式（指数関数）$f(x) = ae^{bx}$
- 誤差の二乗和$S$
  - $S = \sum_{i=1}^{n}\{f(x_i) - y_i\}^2 = \sum_{i=1}^{n}\{ae^{bx_i} - y_i\}^2$
    - 多項式ではないので，係数の微分では正規方程式が得られない
    - 近似関数の変換（対数）
      - $g(x) = \log f(x) = \log ae^{bx} = \log a + bx = a' + bx$
        - » $a' = \log a$
    - データの変換（対数）
      - $y'_i = \log f(y_i) \qquad i = 1,2,\cdots,n$

# 指数関数の最小二乗近似

- 誤差の二乗和$S'$
  - $S' = \sum_{i=1}^{n}\{g(x_i) - y'_i\}^2 = \sum_{i=1}^{n}\{a' + bx_i - y'_i\}^2$
  - 一次関数の最小二乗近似として扱える
    - $\frac{\partial S'}{\partial a'} = 2\{a'n + b\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y'_i\} = 0$
    - $\frac{\partial S'}{\partial b} = 2\{a'\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} {x_i}^2 - \sum_{i=1}^{n} x_i y'_i\} = 0$
    - $\begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} {x_i}^2 \end{bmatrix}\begin{bmatrix} a' \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y'_i \\ \sum_{i=1}^{n} x_i y'_i \end{bmatrix}$
      - $a = e^{a'}$
    - ただし$ae^{bx} - y$ではなく，$\log a + bx - \log y$を評価しているので，値の小さい$y_i$の重みが大きい

# 一般的な関数の最小二乗近似

- 近似関数$f(x, c_0, c_1, \cdots, c_m)$
  - ただし$x$:変数，$c_j(j = 1, \cdots, m)$:係数（未知のパラメータ）
- $n$組のデータ$(x_i, y_i)\ i = 1,2, \cdots, n$に対する誤差の二乗和を最小化する係数の同定
  - $S = \sum_{i=1}^{n}\{f(x_i, c_0, c_1, \cdots, c_m) - y_i\}^2$
  - 繰り返し計算より$c_j(j = 1, \cdots, m)$を求めることを考える

# 一般的な関数の最小二乗近似

- 関数$f(x)$の$x = a$近傍でのテイラー展開
  - $f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots$
  - 2次以上の項を無視
    - $f(x) \cong f(a) + f'^{(a)}(x-a) = f(a) + f'(a)\Delta a$
      - $\Delta a = x - a$として
- 偏微分の定義
  - $x, y$の関数$f(x, y)$が，$x, y$に関して偏微分可能な場合
    - $df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$
- $f(x, c_0, c_1, \cdots, c_m)$ の$c_0 = {c_0}^0, c_1 = {c_1}^0, \cdots, c_m = {c_m}^0$近傍でのテイラー展開（2次以上の項を無視）
  - $f(x, c_0, c_1, \cdots, c_m) = f(x, {c_0}^0, {c_1}^0, \cdots, {c_m}^0) + \frac{\partial f}{\partial c_0}\Delta c_0 + \frac{\partial f}{\partial c_1}\Delta c_1 + \cdots + \frac{\partial f}{\partial c_m}\Delta c_m$
    - ただし繰り返し計算における係数$c_j$の初期値を${c_j}^0$，最適値$c_j$との関係は$\Delta c_j = c_j - {c_j}^0$

# 一般的な関数の最小二乗近似

- $S = \sum_{i=1}^{n} \{ f(x_i, c_0, c_1, \cdots, c_m) - y_i \}^2$
  $\cong \sum_{i=1}^{n} \{ f(x_i, {c_0}^0, {c_1}^0, \cdots, {c_m}^0)$
  $$+ \frac{\partial f}{\partial c_0} \Delta c_0 + \frac{\partial f}{\partial c_1} \Delta c_1 + \cdots + \frac{\partial f}{\partial c_m} \Delta c_m - y_i \Big\}^2$$
  - $\Delta c_j$に関する二次方程式となっている
  - $S$を最小とする$\Delta c_j$は$\frac{\partial S}{\partial \Delta c_j} = 0$を満たす

# 一般的な関数の最小二乗近似

- $\frac{\partial S}{\partial \Delta c_j} = \frac{\partial S}{\partial \Delta c_j} \sum_{i=1}^{n} \left\{ f(x_i, {c_0}^0, {c_1}^0, \cdots, {c_m}^0) + \frac{\partial f}{\partial c_0} \Delta c_0 + \frac{\partial f}{\partial c_1} \Delta c_1 + \cdots + \frac{\partial f}{\partial c_j} \Delta c_j + \cdots + \frac{\partial f}{\partial c_m} \Delta c_m - y_i \right\}^2$
  $= \sum_{i=1}^{n} 2 \frac{\partial f}{\partial c_j} \left\{ f(x_i, {c_0}^0, {c_1}^0, \cdots, {c_m}^0) + \frac{\partial f}{\partial c_0} \Delta c_0 + \frac{\partial f}{\partial c_1} \Delta c_1 + \cdots + \frac{\partial f}{\partial c_m} \Delta c_m - y_i \right\} = 0$
  - $\sum_{i=1}^{n} \frac{\partial f}{\partial c_j} \left\{ f(x_i, {c_0}^0, {c_1}^0, \cdots, {c_m}^0) + \frac{\partial f}{\partial c_0} \Delta c_0 + \frac{\partial f}{\partial c_1} \Delta c_1 + \cdots + \frac{\partial f}{\partial c_m} \Delta c_m - y_i \right\} = 0$
  - $\Delta c_0 \sum_{i=1}^{n} \frac{\partial f}{\partial c_j} \frac{\partial f}{\partial c_0} + \Delta c_1 \sum_{i=1}^{n} \frac{\partial f}{\partial c_j} \frac{\partial f}{\partial c_1} + \cdots + \Delta c_m \sum_{i=1}^{n} \frac{\partial f}{\partial c_j} \frac{\partial f}{\partial c_m} = \sum_{i=1}^{n} \frac{\partial f}{\partial c_j} \{ y_i - f(x_i, {c_0}^0, {c_1}^0, \cdots, {c_m}^0) \}$

# 一般的な関数の最小二乗近似

- $$\begin{bmatrix} \sum_{i=1}^{n}\frac{\partial f}{\partial c_0}\frac{\partial f}{\partial c_0} & \sum_{i=1}^{n}\frac{\partial f}{\partial c_0}\frac{\partial f}{\partial c_1} & \sum_{i=1}^{n}\frac{\partial f}{\partial c_0}\frac{\partial f}{\partial c_2} & \cdots & \sum_{i=1}^{n}\frac{\partial f}{\partial c_0}\frac{\partial f}{\partial c_m} \\ \sum_{i=1}^{n}\frac{\partial f}{\partial c_1}\frac{\partial f}{\partial c_0} & \sum_{i=1}^{n}\frac{\partial f}{\partial c_1}\frac{\partial f}{\partial c_1} & \sum_{i=1}^{n}\frac{\partial f}{\partial c_1}\frac{\partial f}{\partial c_2} & \cdots & \sum_{i=1}^{n}\frac{\partial f}{\partial c_1}\frac{\partial f}{\partial c_m} \\ \vdots & \vdots & & \ddots & \vdots \\ \sum_{i=1}^{n}\frac{\partial f}{\partial c_m}\frac{\partial f}{\partial c_0} & \sum_{i=1}^{n}\frac{\partial f}{\partial c_m}\frac{\partial f}{\partial c_1} & \sum_{i=1}^{n}\frac{\partial f}{\partial c_m}\frac{\partial f}{\partial c_2} & \cdots & \sum_{i=1}^{n}\frac{\partial f}{\partial c_m}\frac{\partial f}{\partial c_m} \end{bmatrix} \begin{bmatrix} \Delta c_0 \\ \Delta c_1 \\ \Delta c_2 \\ \vdots \\ \Delta c_m \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n}\frac{\partial f}{\partial c_0}\{y_i - f(x_i, {c_0}^0, {c_1}^0, \cdots, {c_m}^0)\} \\ \sum_{i=1}^{n}\frac{\partial f}{\partial c_1}\{y_i - f(x_i, {c_0}^0, {c_1}^0, \cdots, {c_m}^0)\} \\ \sum_{i=1}^{n}\frac{\partial f}{\partial c_2}\{y_i - f(x_i, {c_0}^0, {c_1}^0, \cdots, {c_m}^0)\} \\ \vdots \\ \sum_{i=1}^{n}\frac{\partial f}{\partial c_m}\{y_i - f(x_i, {c_0}^0, {c_1}^0, \cdots, {c_m}^0)\} \end{bmatrix}$$
- 求まった$\Delta c_j$より，${c_j}^1 = {c_j}^0 + \Delta c_j$として，${c_j}^0$に換えて繰り返し計算を行う

# ミニマックス法

- 誤差の最大絶対値を最小化するように近似
- 区間$[a,b]$で連続な関数$f(x)$の近似関数$p(x)$
  - $\min\left\{\max|p(x)-f(x)|\right\}$
  - $p(x)$が$n$次の多項式の場合ミニマックス近似多項式（最良近似多項式）