2015.7.15

情報処理 I：行列の収束計算

ヤコビ法，ガウス・ザイデル法

松谷茂樹

ガウス・ザイデル法、ヤコビ法がなぜ線型計算できるのかということに関して、代数的な観点から書かれた教科書がないようなので以下にまとめる。

行列の分解

与えられた $n\times n$ 行列 A と n 次元ベクトル b,

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1\,n-1} & a_{1n}\\ a_{21} & a_{22} & a_{23} & \cdots & a_{2\,n-1} & a_{2n}\\ a_{31} & a_{32} & a_{33} & \cdots & a_{3\,n-1} & a_{3n}\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ a_{n-1\,1} & a_{n-1\,2} & a_{n-1\,3} & \cdots & a_{n-1\,n-1} & a_{n-1\,n}\\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{n\,n-1} & a_{nn}\end{pmatrix}$$ を以下

のように、厳密上三角行列 U と対角成分 D と厳密下三角行列 L に分ける:

$$A=D+L+U$$

つまり、

$$U=\begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n-1} & a_{1n}\\ 0 & 0 & a_{23} & \cdots & a_{2n-1} & a_{2n}\\ 0 & 0 & 0 & \cdots & a_{3n-1} & a_{3n}\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & a_{n-1n}\\ 0 & 0 & 0 & \cdots & 0 & 0\end{pmatrix},$$

$$L=\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0\\ a_{21} & 0 & 0 & \cdots & 0 & 0\\ a_{31} & a_{32} & 0 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ a_{n-1\,1} & a_{n-1\,2} & a_{n-1\,3} & \cdots & 0 & 0\\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{n\,n-1} & 0\end{pmatrix}$$

$$D=\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 & 0\\ 0 & a_{22} & 0 & \cdots & 0 & 0\\ 0 & 0 & a_{33} & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & a_{n-1n-1} & 0\\ 0 & 0 & 0 & \cdots & 0 & a_{nn}\end{pmatrix}$$

また、

$$H=L+D$$

とする。

$$A=D+H$$

$$\overrightarrow{b}=\begin{pmatrix} b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}$$ に対し $$\overrightarrow{x}=\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}$$

に関する方程式

$$A\overrightarrow{x}=\overrightarrow{b}$$

を解くことを考える。

ヤコビ法

$A\overrightarrow{x}=\overrightarrow{b}$ が、$(D+H)\overrightarrow{x}=\overrightarrow{b}$ に注意する。
初期値としてあるベクトル $\overrightarrow{x}^{(0)}$ に対して、

$$\overrightarrow{x}^{(n+1)}=D^{-1}(\overrightarrow{b}-H\overrightarrow{x}^{(n)})$$

を考え，$n=0,1,2,\dots$ の逐次計算を行い，収束した先を $A\overrightarrow{x}=\overrightarrow{b}$ の解 $\overrightarrow{x}^*$ であると考えるのがヤコビ法である。
$A\overrightarrow{x}=\overrightarrow{b}$ の解 $\overrightarrow{x}^*$ に対しては,

$$\overrightarrow{x}^*=D^{-1}(\overrightarrow{b}-H\overrightarrow{x}^*)$$

を満たすので，$|\overrightarrow{x}^{(n+1)}-\overrightarrow{x}^{(n)}|$ は，よい初期値 $\overrightarrow{x}^{(0)}$ に対してはより早く小さくなる。（初期値依存性が存在する。）

ヤコビ法の原理

1. $n=1$ の場合：
   $\overrightarrow{x}^{(1)}=D^{-1}(\overrightarrow{b}-H\overrightarrow{x}^{(0)})$
2. $n=2$ の場合：
   $\overrightarrow{x}^{(2)}=D^{-1}(\overrightarrow{b}-H\overrightarrow{x}^{(1)})$
   を考える。これは
   $$\begin{aligned}\overrightarrow{x}^{(2)}&=D^{-1}(\overrightarrow{b}-H\overrightarrow{x}^{(1)})\\ &=D^{-1}(\overrightarrow{b}-HD^{-1}(\overrightarrow{b}-H\overrightarrow{x}^{(0)}))\\ &=D^{-1}(I-HD^{-1})\overrightarrow{b}-(D^{-1}H)^2\overrightarrow{x}^{(0)}\end{aligned}$$
3. $n=3$ の場合
   $$\begin{aligned}\overrightarrow{x}^{(3)}&=D^{-1}(\overrightarrow{b}-H\overrightarrow{x}^{(2)})\\ &=D^{-1}(\overrightarrow{b}-HD^{-1}(\overrightarrow{b}-H\overrightarrow{x}^{(1)})\\ &=D^{-1}(\overrightarrow{b}-HD^{-1}(\overrightarrow{b}-HD^{-1}(\overrightarrow{b}-H\overrightarrow{x}^{(0)})))\\ &=D^{-1}(I-HD^{-1}+(HD^{-1})^2)\overrightarrow{b}-(D^{-1}H)^3\overrightarrow{x}^{(0)}\end{aligned}$$
4. 一般の n の場合
   $$\overrightarrow{x}^{(n)}=D^{-1}\left(\sum_{\ell=0}^{n-1}(-HD^{-1})^\ell\right)\overrightarrow{b}-(D^{-1}H)^n\overrightarrow{x}^{(0)},\quad\text{(J-1)}$$
   となる。
5. 一般の n が十分大きい場合:
   $$-D^{-1}\sum_{\ell=0}^{n}(-1)^\ell(HD^{-1})^\ell-A^{-1}=(-1)^n(D^{-1}H)^{n+1}$$
   となることが知られている。
   ある条件の下（対角優位： $\min a_{ii}>\sum_{i\neq j}|a_{i,j}|$) において $n\to\infty$ に対して，
   $$D^{-1}\left(\sum_{\ell=0}^{n-1}(-HD^{-1})^\ell\right)\longrightarrow A^{-1},\quad (D^{-1}H)^n\longrightarrow 0$$
   となる。(レゾルベントと現代数学ではいう）
   よって，(J-1) 式の第一項は $A^{-1}b$ に漸近する。他方、第二項
   $$(HD^{-1})^\ell\overrightarrow{x}^{(0)}\to\overrightarrow{0}$$
   に近づく。従ってヤコビ法により，n を十分大きくすることで
   $$\overrightarrow{x}^{(n)}\longrightarrow A^{-1}\overrightarrow{b}$$
   となり、解に漸近する。

ガウス・ザイデル法

$A\overrightarrow{x}=\overrightarrow{b}$ が、$(D+L+U)\overrightarrow{x}=\overrightarrow{b}$ に注意するとヤコビ法より，収束の早い方法が考えられる。
初期値としてあるベクトル $\overrightarrow{x}^{(0)}$ に対して、

$$\overrightarrow{x}^{(n+1)}=D^{-1}(\overrightarrow{b}-L\overrightarrow{x}^{(n+1)}-U\overrightarrow{x}^{(n)})$$

を考え，$n=0,1,2,\dots$ の逐次計算を行い，収束した先を $A\overrightarrow{x}=\overrightarrow{b}$ の解 $\overrightarrow{x}^*$ であると考えるのがガウス・ザイデル法である。
$A\overrightarrow{x}=\overrightarrow{b}$ の解 $\overrightarrow{x}^*$ に対しては,

$$\overrightarrow{x}^*=D^{-1}(\overrightarrow{b}-L\overrightarrow{x}^*-U\overrightarrow{x}^*)$$

を満たすので，$|\overrightarrow{x}^{(n+1)}-\overrightarrow{x}^{(n)}|$ は，よい初期値 $\overrightarrow{x}^{(0)}$ に対してはより早く小さくなる。（初期値依存性が存在する。）

ガウス・ザイデル法の収束の原理

1. 基本的な関係式：

   $A\overrightarrow{x} = \overrightarrow{b}$ が、$(D+H)\overrightarrow{x} = \overrightarrow{b}$ に注意するとヤコビ法より，収束の早い方法が考えられる。

   $\overrightarrow{x}^{(n+1)} = D^{-1}(\overrightarrow{b} - L\overrightarrow{x}^{(n+1)} - U\overrightarrow{x}^{(n)})$ は

   $(E + D^{-1}L)\overrightarrow{x}^{(n+1)} = D^{-1}(\overrightarrow{b} - U\overrightarrow{x}^{(n)})$ と書けることより，基本関係式

   $$\overrightarrow{x}^{(n+1)} = (D+L)^{-1}(\overrightarrow{b} - U\overrightarrow{x}^{(n)})$$

   を計算している事と一致する．

   但し、$(AB)^{-1} = B^{-1}A^{-1}$ より $(E+D^{-1}L)^{-1}D^{-1} = (D(E+D^{-1}L))^{-1} = (D+L)^{-1}$ を利用した。これより，

2. 一般の n の場合

   $$\overrightarrow{x}^{(n)} = (D+L)^{-1}\left(\sum_{\ell=0}^{n-1}(-(D+L)^{-1}U)^{\ell}\right)\overrightarrow{b}$$
   $$-(U(D+L)^{-1})^{n}\overrightarrow{x}^{(0)}, \quad \text{(G-1)}$$

   となり，ある条件の下（対角優位： $\min a_{ii} > \sum_{i\neq j}|a_{i,j}|$）において

   $n \to \infty$ に対して，

   $$(D+L)^{-1}\left(\sum_{\ell=0}^{n-1}(-(D+L)^{-1}U)^{\ell}\right) \longrightarrow A^{-1},$$

   $$((D+L)^{-1}U)^{n} \longrightarrow 0$$

   となる。(レゾルベントと現代数学ではいう）

   よって，(G-1) 式の第一項は $A^{-1}b$ に漸近する。他方、第二項

   $$(U(D+L)^{-1})^{\ell}\overrightarrow{x}^{(0)} \to \overrightarrow{0}$$

   に近づく。従ってガウス・ザイデル法により，n を十分大きくすることで

   $$\overrightarrow{x}^{(n)} \longrightarrow A^{-1}\overrightarrow{b}$$

   となり、解に漸近する。

ヤコビ法，ガウス・ザイデル法がうまく行くわけ

以下、単純化して 1 次元にする：

1. テイラー展開：

   テイラー展開より，次の等式が成り立つ：

   $$\frac{1}{1+x} = 1-x+x^2-x^3+x^4-\cdots = \sum_{\ell=0}^{\infty}(-x)^{\ell}, \quad ,(|x|<1), \quad \text{(T-1)}$$

   $x \ll 1$ に対しては $\sum_{\ell=0}^{n}(-x)^{\ell}$ は有限の n でもよい近似を示す。

2. 数学的事実：

   d,h をある実数とする (T-1) を使って $d > h$ のとき

   $$\frac{1}{d+h} = \frac{1}{d(1+h/d)} = \frac{1}{d}\frac{1}{1+h/d}$$
   $$= \frac{1}{d}(1 - \frac{h}{d} + \left(\frac{h}{d}\right)^2 - \left(\frac{h}{d}\right)^3 + \cdots)$$
   $$= \frac{1}{d}\sum_{\ell=0}^{\infty}\left(-\frac{h}{d}\right)^{\ell}$$

3. ヤコビ法のイメージ

   $(d+h)x = b$ となる x を計算するのにおいて，ある初期値 $x^{(0)}$ を用意して、

   $$x^{(n+1)} = \frac{1}{d}(b - hx^{(n)})$$

   を逐次的に計算する。

   $$x^{(n+1)} = \textstyle\sum_{\ell=0}^{n}\left(-\frac{h}{d}\right)^{\ell} b - \left(-\frac{h}{d}\right)^{n} x^{(0)}$$

4. ガウス・ザイデルのイメージ

   上記 h を $h = \ell + u$ に分解して，$(d+\ell+u)x = b$ となる x を計算するのにおいて，ある初期値 $x^{(0)}$ を用意して、

   $$x^{(n+1)} = \frac{1}{d+\ell}(b - ux^{(n)})$$

   を逐次的に計算する。

   $$x^{(n+1)} = \textstyle\sum_{m=0}^{n}\left(-\frac{h}{d+\ell}\right)^{m} b - \left(-\frac{u}{d+\ell}\right)^{n} x^{(0)}$$

   こちらの方が収束が早いと思われる。

5. $d = 16, \ell = 1, u = 2, h = \ell + u = 3$ の場合, $b = 1, x^{(0)} = 0$ として

\	1	2	3	4	5
J 型	0.05078	0.05298	0.05257	0.05264	0.05263
GS 型	0.05190	0.05272	0.05262	0.05263	0.05263

   となり，$1/19 = 0.0526316$ への収束は当然であるが後者の方がよい事が見える。

付録：行列の公式

1. $A=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$ と $B=\begin{pmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{pmatrix}$ との掛け算

$$\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{pmatrix}$$
$$=\begin{pmatrix}a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22}\\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22}\end{pmatrix}$$

2. $A=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$ と $\overrightarrow{b}=\begin{pmatrix}b_1\\ b_2\end{pmatrix}$ との掛け算

$$\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}b_1\\ b_2\end{pmatrix}=\begin{pmatrix}a_{11}b_1+a_{12}b_2\\ a_{21}b_1+a_{22}b_2\end{pmatrix}$$

3. $A=\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}$ と $\overrightarrow{b}=\begin{pmatrix}b_1\\ b_2\\ b_3\end{pmatrix}$ との掛け算

$$\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}\begin{pmatrix}b_1\\ b_2\\ b_3\end{pmatrix}=\begin{pmatrix}a_{11}b_1+a_{12}b_2+a_{13}b_3\\ a_{21}b_1+a_{22}b_2+a_{23}b_3\\ a_{31}b_1+a_{32}b_2+a_{33}b_3\end{pmatrix}$$

4. $A=\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}$ と $B=\begin{pmatrix}b_{11} & b_{12} & b_{13}\\ b_{21} & b_{22} & b_{23}\\ b_{31} & b_{32} & b_{33}\end{pmatrix}$ との掛け算

$$\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}\begin{pmatrix}b_{11} & b_{12} & b_{13}\\ b_{21} & b_{22} & b_{23}\\ b_{31} & b_{32} & b_{33}\end{pmatrix}$$
$$=\begin{pmatrix}a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31} & a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32} & a_{11}+b_{13}+a_{12}b_{23}+a_{13}b_{33}\\ a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31} & a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{32} & a_{21}+b_{13}+a_{22}b_{23}+a_{23}b_{33}\\ a_{31}b_{11}+a_{32}b_{21}+a_{33}b_{31} & a_{31}b_{12}+a_{32}b_{22}+a_{33}b_{32} & a_{31}+b_{13}+a_{32}b_{23}+a_{33}b_{33}\end{pmatrix}$$

5. $A=\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}$ と $B=\begin{pmatrix}b_{11} & b_{12} & b_{13} & d_1\\ b_{21} & b_{22} & b_{23} & d_2\\ b_{31} & b_{32} & b_{33} & d_3\end{pmatrix}$ との掛け算

$$\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}\begin{pmatrix}b_{11} & b_{12} & b_{13} & d_1\\ b_{21} & b_{22} & b_{23} & d_2\\ b_{31} & b_{32} & b_{33} & d_3\end{pmatrix}$$
$$=\begin{pmatrix}a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31} & a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32} & a_{11}b_{13}+a_{12}b_{23}+a_{13}b_{33} & a_{11}d_1+a_{12}d_2+a_{13}d3\\ a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31} & a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{32} & a_{21}b_{13}+a_{22}b_{23}+a_{23}b_{33} & a_{21}d_1+a_{22}d_2+a_{23}d_3\\ a_{31}b_{11}+a_{32}b_{21}+a_{33}b_{31} & a_{31}b_{12}+a_{32}b_{22}+a_{33}b_{32} & a_{31}b_{13}+a_{32}b_{23}+a_{33}b_{33} & a_{31}d_1+a_{32}d_2+a_{33}d_3\end{pmatrix}$$