

数値計算備忘録

motchy

2022 年 8 月 23 日 ~ 2022 年 9 月 6 日
ver 0.1.0-wip

目次

第 I 部	線形代数	2
第 1 章	Gauss-Seidel 法	3
1.1	定義	3
1.2	係数行列が狭義優対角ならば厳密解に収束すること	3
第 2 章	Cholesky 分解	5
2.1	rank-one update	5
第 3 章	LDL 分解	7
3.1	rank-one update	7

第 I 部

線形代数

第 1 章

Gauss-Seidel 法

1.1 定義

以下に述べる定義は Wikipedia の英語記事 [Gauss Seidel method](#) からの引用である。

$n \in \mathbb{N}$, $A \in \mathbb{C}^{n \times n}$, $\mathbf{b} \in \mathbb{C}^n$ とする。 A は正定値対称、または狭義優対角であるとする。Gauss-Seidel 法とは、線型方程式 $A\mathbf{x} = \mathbf{b}$ の解を求める反復法である。 $\mathbf{x}_1 \in \mathbb{C}^n$ を任意の初期解とし、次の漸化式で解候補を更新してゆく。

$$L_* \mathbf{x}_{k+1} = -U \mathbf{x}_k \quad (k = 1, 2, \dots)$$

ここに L_* は A の対角成分およびその下側の要素からなる下三角行列であり、 U は A の対角成分の上側の要素からなる上三角行列である。

1.2 係数行列が狭義優対角ならば厳密解に収束すること

Proof. A の次数を n とする。 $\hat{\mathbf{x}}$ を厳密解とすると $L_* \hat{\mathbf{x}} = \mathbf{b} - U \hat{\mathbf{x}}$ である。これを解の更新式から減じると次式を得る。

$$L_*(\mathbf{x}_{k+1} - \hat{\mathbf{x}}) = -U(\mathbf{x}_k - \hat{\mathbf{x}}) \quad (1)$$

$\mathbf{v}_k := \mathbf{x}_k - \hat{\mathbf{x}}$ とおくと、式 (1) より次式が成り立つ。

$$L_* \mathbf{v}_{k+1} = -U \mathbf{v}_k \quad (2)$$

$M_k := \max_{i=1, \dots, n} |v_{k,i}|$ ($v_{k,i}$ は \mathbf{v}_k の第 i 要素) とする。次の 2 つが同時に成り立つことが、 \mathbf{v}_k が $\mathbf{0}_n$ に収束するための十分条件である。

1. ある $k \in \mathbb{N}$ に対して $M_k = 0$ ならば $M_l = 0$ ($l = k+1, k+2, \dots$)
2. 適当な $0 < \alpha < 1$ が存在して $M_k > 0 \Rightarrow M_{k+1} < \alpha M_k$

L_* が正則であることと式 (2) より直ちに 1. が成り立つ。次に 2. を数学的帰納法で示す。 $\tilde{\alpha}$ を次式で定義する。

$$\tilde{\alpha} := \min_{i=1, 2, \dots, n} \frac{1}{|a_{i,i}|} \sum_{j=1, \dots, n \wedge j \neq i} |a_{i,j}|$$

A は優対角だから $0 < \tilde{\alpha} < 1$ である。

$$\begin{aligned} a_{1,1}v_{k+1,1} &= - \sum_{j=2}^n a_{1,j}v_{k,j} \\ |a_{1,1}||v_{k+1,1}| &= \left| \sum_{j=2}^n a_{1,j}v_{k,j} \right| \leq \sum_{j=2}^n |a_{1,j}||v_{k,j}| \leq M_k \sum_{j=2}^n |a_{1,j}| \\ |v_{k+1,1}| &\leq \frac{M_k}{|a_{1,1}|} \sum_{j=2}^n |a_{1,j}| \leq \tilde{\alpha} M_k \end{aligned}$$

$|v_{k+1,j}| \leq \tilde{\alpha} M_k$ ($j = 1, 2, \dots, l$) ($l \in \{1, 2, \dots, n-1\}$) ならば $|v_{k+1,l+1}| \leq \tilde{\alpha} M_k$ であることを示す。式 (2) の $l+1$ 行目を展開すると次式を得る。

$$\begin{aligned} \sum_{j=1}^{l+1} a_{l+1,j}v_{k+1,j} &= - \sum_{j=l+2}^n a_{l+1,j}v_{k,j} \\ a_{l+1,l+1}v_{k+1,l+1} &= - \sum_{j=1}^l a_{l+1,j}v_{k+1,j} - \sum_{j=l+2}^n a_{l+1,j}v_{k,j} \\ |a_{l+1,l+1}||v_{k+1,l+1}| &= \left| - \sum_{j=1}^l a_{l+1,j}v_{k+1,j} - \sum_{j=l+2}^n a_{l+1,j}v_{k,j} \right| \leq \sum_{j=1}^l |a_{l+1,j}||v_{k+1,j}| + \sum_{j=l+2}^n |a_{l+1,j}||v_{k,j}| \\ &\leq M_k \sum_{j=1, \dots, n \wedge j \neq l+1} |a_{l+1,j}| \\ |v_{k+1,l+1}| &\leq \frac{M_k}{|a_{l+1,l+1}|} \sum_{j=1, \dots, n \wedge j \neq l+1} |a_{l+1,j}| \leq \tilde{\alpha} M_k \end{aligned}$$

以上より帰納的に $|v_{k+1,j}| \leq \tilde{\alpha} M_k$ ($j = 1, 2, \dots, n$) が成り立つ。すなわち $M_{k+1} \leq \tilde{\alpha} M_k$ が成り立つ。
 $\tilde{\alpha} < \alpha < 1$ となるように α を定めることで 2. が示される。 \square

第 2 章

Cholesky 分解

2.1 rank-one update

主張

$n \in \mathbb{N}$, $A \in \mathbb{C}^{n \times n}$, $A \succeq O$, $\mathbf{x} \in \mathbb{C}^n$ とし、 A は Hermite 行列であるとする。 $A + \mathbf{x}\mathbf{x}^*$ に対して Cholesky 分解のアルゴリズムを適用すると $O(n^3)$ の計算量を要する。しかし、 A の Cholesky 分解 LL^* が既に得られているとき、 $A + \mathbf{x}\mathbf{x}^*$ の Cholesky 分解を $O(n^2)$ で得ることができる。 $\mathbf{x}\mathbf{x}^*$ の階数が 1 以下である (特に 0 となるのは $\mathbf{x} = \mathbf{0}$ の時かつその時に限る) ことから、この方法は “rank-one update” と呼ばれている。

導出. 方針としては、 $n \times n$ 行列の rank-one update を $(n-1) \times (n-1)$ 行列の問題に帰着させ、以降同様に逐次的に行列の次数を縮小しながら解を構築する。このアルゴリズムの総計算量が $O(n^2)$ となるのは明らかであろう。

$A + \mathbf{x}\mathbf{x}^*$ の Cholesky 分解を FF^* とする。 L の第 i 列ベクトルを $\mathbf{l}_i = [0, \dots, 0, l_{i,i}, \dots, l_{n,i}]^\top \in \mathbb{C}^{n \times n}$ とし、同様に F の第 i 列ベクトルを $\mathbf{f}_i = [0, \dots, 0, f_{i,i}, \dots, f_{n,i}]^\top \in \mathbb{C}^{n \times n}$ とすると次式が成り立つ。

$$\begin{aligned} \sum_{i=1}^n \mathbf{f}_i \mathbf{f}_i^* &= \mathbf{x}\mathbf{x}^* + \sum_{i=1}^n \mathbf{l}_i \mathbf{l}_i^* \\ \mathbf{f}_1 \mathbf{f}_1^* + \sum_{i=2}^n \mathbf{f}_i \mathbf{f}_i^* &= \mathbf{x}\mathbf{x}^* + \mathbf{l}_1 \mathbf{l}_1^* + \sum_{i=2}^n \mathbf{l}_i \mathbf{l}_i^* \end{aligned} \quad (1)$$

$\mathbf{f}_i \mathbf{f}_i^*$, $\mathbf{l}_i \mathbf{l}_i^*$ ($i = 2, 3, \dots, n$) の第 1 行および第 1 列は 0 であるから、 $\mathbf{f}_1 \mathbf{f}_1^*$ と $\mathbf{x}\mathbf{x}^* + \mathbf{l}_1 \mathbf{l}_1^*$ の第 1 行および第 1 列が一致する。これより次式が成り立つ。

$$f_{1,1} = \sqrt{l_{1,1}^2 + |x_1|^2} =: r, \quad f_{k,1} = \frac{1}{r} (l_{1,1} l_{k,1} + \overline{x_1} x_k) \quad (k = 2, 3, \dots, n) \quad (2)$$

ただし L の対角成分が非負の実数であることを前提としている。以上より、 $\tilde{\mathbf{l}}_1 := [0, l_{2,1}, l_{3,1}, \dots, l_{n,1}]^\top$, $\tilde{\mathbf{x}} := [0, x_2, x_3, \dots, x_n]^\top$ とすると次式が成り立つ。

$$\mathbf{f}_1 = r \mathbf{e}_1 + \frac{l_{1,1}}{r} \tilde{\mathbf{l}}_1 + \frac{\overline{x_1}}{r} \tilde{\mathbf{x}}$$

ここに \mathbf{e}_1 は第 1 要素が 1 で他は 0 であるベクトルである。 $\mathbf{f}_1 \mathbf{f}_1^*$ の右下 $(n-1) \times (n-1)$ 行列を評価すると

次式を得る。

$$\begin{aligned}
& \frac{1}{r^2} \left(l_{1,1} \tilde{l}_1 + \overline{x_1} \tilde{x} \right) = \frac{1}{r^2} \left(l_{1,1}^2 \tilde{l}_1 \tilde{l}_1^* + l_{1,1} x_1 \tilde{l}_1 \tilde{x}^* + |x_1|^2 \tilde{x} \tilde{x}^* + l_{1,1} \overline{x_1} \tilde{x} \tilde{l}_1^* \right) \\
& = \left(1 - \frac{|x_1|^2}{r^2} \right) \tilde{l}_1 \tilde{l}_1^* + \frac{l_{1,1} x_1}{r^2} \tilde{l}_1 \tilde{x}^* + \left(1 - \frac{l_{1,1}^2}{r^2} \right) \tilde{x} \tilde{x}^* + \frac{\overline{x_1} l_{1,1}}{r^2} \tilde{x} \tilde{l}_1^* \\
& = \tilde{l}_1 \tilde{l}_1^* + \tilde{x} \tilde{x}^* - \frac{1}{r^2} \left(|x_1|^2 \tilde{l}_1 \tilde{l}_1^* + l_{1,1}^2 \tilde{x} \tilde{x}^* - x_1 l_{1,1} \tilde{l}_1 \tilde{x}^* - \overline{x_1} l_{1,1} \tilde{x} \tilde{l}_1^* \right) \\
& = \tilde{l}_1 \tilde{l}_1^* + \tilde{x} \tilde{x}^* - \mathbf{y} \mathbf{y}^* \quad \text{where} \quad \mathbf{y} = \frac{1}{r} \left(l_{1,1} \tilde{x} - x_1 \tilde{l}_1 \right)
\end{aligned}$$

上式の $\tilde{l}_1 \tilde{l}_1^* + \tilde{x} \tilde{x}^*$ は $\mathbf{x} \mathbf{x}^* + l_1 l_1^*$ の右下 $(n-1) \times (n-1)$ 行列である。以上より次式が成り立つ。

$$\mathbf{f}_1 \mathbf{f}_1^* = \mathbf{x} \mathbf{x}^* + l_1 l_1^* - \mathbf{y} \mathbf{y}^*$$

これを式 (1) に適用して次式を得る。

$$\sum_{i=2}^n \mathbf{f}_i \mathbf{f}_i^* = \mathbf{y} \mathbf{y}^* + \sum_{i=2}^n l_i l_i^*$$

これは $(n-1) \times (n-1)$ 行列の rank-one update である。このようにして行列の次数を逐次的に縮小し、最後はスカラーの計算に帰着する。次数 k の問題に対し式 (2) の計算量は $O(k)$ であるから、このアルゴリズムの総計算量は $n(n+1)/2$ に比例する。 \square

このアルゴリズムの Julia 1.8.0 による実装例を `Cholesky-decomposition_rank-one_update.ipynb` に記した。本文書の Git リポジトリ内でファイル検索すれば見つかる。

第 3 章

LDL 分解

3.1 rank-one update

主張

$n \in \mathbb{N}$, $A \in \mathbb{C}^{n \times n}$, $A \succeq O$, $\mathbf{x} \in \mathbb{C}^n$ とし、 A は Hermite 行列であるとする。 $A + \mathbf{x}\mathbf{x}^*$ に対して LDL 分解のアルゴリズムを適用すると $O(n^3)$ の計算量を要する。しかし、 A の LDL 分解 LDL^* が既に得られているとき、 $A + \mathbf{x}\mathbf{x}^*$ の LDL 分解を $O(n^2)$ で得ることができる。 $\mathbf{x}\mathbf{x}^*$ の階数が 1 以下である (特に 0 となるのは $\mathbf{x} = \mathbf{0}$ の時かつその時に限る) ことから、この方法は “rank-one update” と呼ばれている。

導出. 方針は Cholesky 分解の rank-one update と同様である。 $A + \mathbf{x}\mathbf{x}^*$ の LDL 分解を FGF^* とする。 D, G の第 i 対角成分をそれぞれ d_i, g_i とする。但し $d_i \geq 0$ を前提とする。 L の第 i 列ベクトルを $\mathbf{l}_i = [0, \dots, 0, 1, l_{i+1,i}, \dots, l_{n,i}]^\top \in \mathbb{C}^{n \times n}$ とし、同様に F の第 i 列ベクトルを $\mathbf{f}_i = [0, \dots, 0, 1, f_{i+1,i}, \dots, f_{n,i}]^\top \in \mathbb{C}^{n \times n}$ とすると次式が成り立つ。

$$\begin{aligned} \sum_{i=1}^n \mathbf{f}_i g_i \mathbf{f}_i^* &= \mathbf{x}\mathbf{x}^* + \sum_{i=1}^n \mathbf{l}_i d_i \mathbf{l}_i^* \\ \mathbf{f}_1 g_1 \mathbf{f}_1^* + \sum_{i=2}^n \mathbf{f}_i g_i \mathbf{f}_i^* &= \mathbf{x}\mathbf{x}^* + \mathbf{l}_1 d_1 \mathbf{l}_1^* + \sum_{i=2}^n \mathbf{l}_i d_i \mathbf{l}_i^* \end{aligned} \quad (1)$$

$\mathbf{f}_i g_i \mathbf{f}_i^*$, $\mathbf{l}_i d_i \mathbf{l}_i^*$ ($i = 2, 3, \dots, n$) の第 1 行および第 1 列は 0 であるから、 $\mathbf{f}_1 g_1 \mathbf{f}_1^*$ と $\mathbf{x}\mathbf{x}^* + \mathbf{l}_1 d_1 \mathbf{l}_1^*$ の第 1 行および第 1 列が一致する。これより次式が成り立つ。

$$g_1 = d_1 + |x_1|^2 =: g, \quad f_{k,1} = \frac{1}{g} (d_1 l_{k,1} + \overline{x_1} x_k) \quad (k = 2, 3, \dots, n) \quad (2)$$

以上より、 $\tilde{\mathbf{l}}_1 := [0, l_{2,1}, l_{3,1}, \dots, l_{n,1}]^\top$, $\tilde{\mathbf{x}} := [0, x_2, x_3, \dots, x_n]^\top$ とすると次式が成り立つ。

$$\mathbf{f}_1 = \mathbf{e}_1 + \frac{d_1}{g} \tilde{\mathbf{l}}_1 + \frac{\overline{x_1}}{g} \tilde{\mathbf{x}}$$

ここに \mathbf{e}_1 は第 1 要素が 1 で他は 0 であるベクトルである。 $\mathbf{f}_1 g_1 \mathbf{f}_1^*$ の右下 $(n-1) \times (n-1)$ 行列を評価する

と次式を得る。

$$\begin{aligned}
& \frac{1}{g} \left(d_1 \tilde{l}_1 + \tilde{x} \tilde{x}^* \right) \left(d_1 \tilde{l}_1 + \tilde{x} \tilde{x}^* \right)^* = \frac{d_1}{g} \tilde{l}_1 d_1 \tilde{l}_1^* + \frac{|x_1|^2}{g} \tilde{x} \tilde{x}^* + \frac{d_1}{g} \left(x_1 \tilde{l}_1 \tilde{x}^* + \overline{x_1} \tilde{x} \tilde{l}_1^* \right) \\
&= \frac{g - |x_1|^2}{g} \tilde{l}_1 d_1 \tilde{l}_1^* + \frac{g - d_1}{g} \tilde{x} \tilde{x}^* + \frac{d_1}{g} \left(x_1 \tilde{l}_1 \tilde{x}^* + \overline{x_1} \tilde{x} \tilde{l}_1^* \right) \\
&= \tilde{l}_1 d_1 \tilde{l}_1^* + \tilde{x} \tilde{x}^* - \frac{d_1}{g} \left[|x_1|^2 \tilde{l}_1 \tilde{l}_1^* + \tilde{x} \tilde{x}^* - x_1 \tilde{l}_1 \tilde{x}^* - \overline{x_1} \tilde{x} \tilde{l}_1^* \right] \\
&= \tilde{l}_1 d_1 \tilde{l}_1^* + \tilde{x} \tilde{x}^* - \mathbf{y} \frac{d_1}{g} \mathbf{y}^* \quad \text{where} \quad \mathbf{y} = x_1 \tilde{l}_1 - \tilde{x}
\end{aligned}$$

上式の $\tilde{l}_1 d_1 \tilde{l}_1^* + \tilde{x} \tilde{x}^*$ は $\mathbf{x} \mathbf{x}^* + \mathbf{l}_1 d_1 \mathbf{l}_1^*$ の右下 $(n-1) \times (n-1)$ 行列である。以上より次式が成り立つ。

$$\mathbf{f}_1 g_1 \mathbf{f}_1^* = \mathbf{x} \mathbf{x}^* + \mathbf{l}_1 d_1 \mathbf{l}_1^* - \mathbf{y} \frac{d_1}{g} \mathbf{y}^*$$

これを式 (1) に適用して次式を得る。

$$\sum_{i=2}^n \mathbf{f}_i g_i \mathbf{f}_i^* = \mathbf{y} \frac{d_1}{g} \mathbf{y}^* + \sum_{i=2}^n \mathbf{l}_i d_i \mathbf{l}_i^*$$

これは $(n-1) \times (n-1)$ 行列の rank-one update である。このようにして行列の次数を逐次的に縮小し、最後はスカラーの計算に帰着する。次数 k の問題に対し式 (2) の計算量は $O(k)$ であるから、このアルゴリズムの総計算量は $n(n+1)/2$ に比例する。□

このアルゴリズムの Julia 1.8.0 による実装例を `LDL-decomposition_rank-one_update.ipynb` に記した。本文書の Git リポジトリ内でファイル検索すれば見つかる。