数値計算

大阪大学基礎工学部

永原正章

2012年6月28日(5限)

連立線形方程式

- N 変数の連立線形方程式

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1N}x_N &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2N}x_N &= b_2 \\
&\vdots \\
a_{N1}x_1 + a_{N2}x_2 + \cdots + a_{NN}x_N &= b_N
\end{aligned}
$$

- ベクトル表現

$$
Ax = b, \quad A \in \mathbb{R}^{N\times N}, \quad x \in \mathbb{R}^N, \quad b \in \mathbb{R}^N.
$$

連立線形方程式

- $Ax = b$
- 行列 A は正則と仮定する.
- 厳密解

$$x^* = A^{-1}b$$

- 直接解法:Gauss の消去法(厳密解,計算量:大)
- 近似解法:Jacobi法,Gauss-Seidel法(近似解,計算量:小)

線形方程式と最小化問題

- 線形方程式 $Ax = b$
- 行列 A は正則かつ $A > 0$(正定値)とする.
- 次の2つは等価である.
  1. x^* は方程式 $Ax = b$ の解.
  2. x^* は次の関数を最小化する

$$J(x) = \frac{1}{2}x^\top Ax - x^\top b.$$

証明

- [2]→[1]:
  J(x) は$\mathbb{R}^N$上で微分可能であるので, J(x)を最小化するベクトルx*は次式を満たさなければならない.

$$\nabla J(x^*) = \mathbf{0}.$$

ここで,∇J(x) は関数 J(x) の**勾配**であり,

$$\nabla J(x) = \begin{bmatrix} \frac{\partial J(x)}{\partial x_1} & \cdots & \frac{\partial J(x)}{\partial x_N} \end{bmatrix}^\top$$

で定義される. 具体的に ∇J(x) を計算すると

$$\nabla J(x) = Ax - b$$

が得られるので,$Ax^* = b$.すなわち線形方程式の解であることがわかる.

証明

- [1]→[2]:
  ベクトルx^* が$Ax^* = b$ を満たすとする. このとき,任意のベクトル$x \in \mathbb{R}^N$ に対して

$$J(x) = J(x^* + x - x^*) = J(x^*) + \frac{1}{2}(x - x^*)^\top A(x - x^*)$$

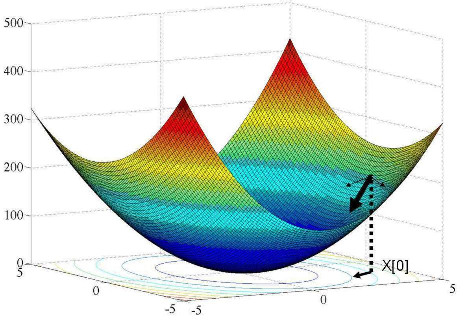
  が成り立つ. $A > 0$だから,上式より $x \neq x^*$ ならば $J(x) > J(x^*)$ となる. したがって, x^* は$J(x)$を最小化することがわかる.

2次形式の勾配

$$\nabla\left(\frac{1}{2}\mathsf{x}^\top \mathsf{A}\mathsf{x} - \mathsf{x}^\top \mathsf{b}\right) = \mathsf{A}\mathsf{x} - \mathsf{b}.$$

線形方程式と最小化問題

- 方程式 $Ax = b$ を解く ⇔ 関数 $J(x)$ を最小化する



- 初期ベクトルx[0]を決める.
- x[1] は $J(x[1]) < J(x[0])$ となるように決める.
- 最も急な傾斜の方向(最急降下方向)に向かって更新する.

線形方程式と最小化問題

- $x \in \mathbb{R}^N$ での$J(x)$の最急降下方向は負の勾配,すなわち

$$-\nabla J(x) = b - Ax$$

- 次の反復法により,x^*の近似解を得る.

$$x[n+1] = x[n] + \alpha[n] r[n], \quad n = 0, 1, 2, \ldots$$

- $\alpha[n] \in \mathbb{R}$: ステップサイズ(未定)
- $r[n] = b - Ax[n]$: 残差ベクトル

最適ステップサイズ

- 反復法

$$x[n+1] = x[n] + \alpha[n]r[n], \quad n = 0, 1, 2, \ldots$$

- $J(x + \alpha r)$ を最小化する$\alpha \in \mathbb{R}$を求める.

$$\begin{aligned}
J(x + \alpha r) &= \frac{1}{2}(x + \alpha r)^\top A(x + \alpha r) - (x + \alpha r)^\top b \\
&= \frac{1}{2}x^\top Ax - x^\top b - \alpha r^\top (b - Ax) + \frac{1}{2}\alpha^2 r^\top Ar \\
&= J(x) - \alpha r^\top r + \frac{1}{2}\alpha^2 r^\top Ar \\
&= \left(\frac{1}{2}r^\top Ar\right)\alpha^2 - (r^\top r)\alpha + J(x)
\end{aligned}$$

最適ステップサイズ

$$J(x+\alpha r) = \left(\frac{1}{2} r^\top A r\right)\alpha^2 - (r^\top r)\alpha + J(x)$$

を最小化する$\alpha \in \mathbb{R}$は

$$\alpha^* = \frac{r^\top r}{r^\top A r}$$

- なお,$A > 0$ であったので,任意の $r \neq 0$ に対して$r^\top A r > 0$ である.
- これを第 n ステップ目のステップサイズとする.

$$\alpha[n] = \frac{r[n]^\top r[n]}{r[n]^\top A r[n]}$$

最急降下法

- 方程式 $\mathsf{Ax} = \mathsf{b}$ の近似解を求める反復法

$$\mathsf{r}[\mathsf{n}] = \mathsf{b} - \mathsf{Ax}[\mathsf{n}],$$
$$\alpha[\mathsf{n}] = \frac{\mathsf{r}[\mathsf{n}]^\top \mathsf{r}[\mathsf{n}]}{\mathsf{r}[\mathsf{n}]^\top \mathsf{Ar}[\mathsf{n}]},$$
$$\mathsf{x}[\mathsf{n}+1] = \mathsf{x}[\mathsf{n}] + \alpha[\mathsf{n}]\mathsf{r}[\mathsf{n}], \quad \mathsf{n} = 0, 1, 2, \ldots$$

- この方法を最急降下法と呼ぶ.

最急降下法

残差ベクトルについて

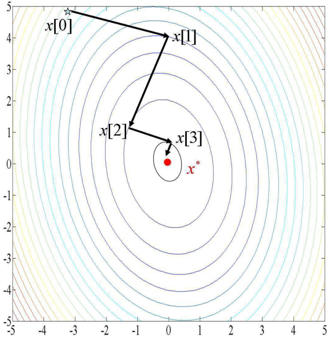
$$\begin{aligned} r[n+1] &= b - Ax[n+1] \\ &= b - Ax[n] - A\alpha[n]r[n] \\ &= r[n] - \alpha[n]Ar[n] \end{aligned}$$

より,$r[n+1]$ と $r[n]$ との内積を計算すると

$$\begin{aligned} r[n]^\top r[n+1] &= r[n]^\top (r[n] - \alpha[n]Ar[n]) \\ &= r[n]^\top r[n] - \alpha[n]r[n]^\top Ar[n] \\ &= r[n]^\top r[n] - \frac{r[n]^\top r[n]}{r[n]^\top Ar[n]} \cdot r[n]^\top Ar[n] \\ &= r[n]^\top r[n] - r[n]^\top r[n] \\ &= 0 \end{aligned}$$

すなわち,残差ベクトル $r[n+1]$ と $r[n]$ は直交する.

最急降下法



- x[0] から最急降下方向にまっすぐ進む
- その方向で最小になるところ(それ以上,その方向に進むとJの値が大きくなるところ)で止まる.その位置がx[1].
- 下に向く方向に90度回転する.(この方向が最急降下方向)
- 以下同様

最急降下法の収束性

- $A > 0$ ならば,最急降下法は任意の初期ベクトルに対して,唯一つの厳密解に収束する.

証明

- 行列 A の固有値を $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N > 0$ とおく.
- このとき,任意のベクトル $x \in \mathbb{R}^N$,ただし$x \neq 0$ に対して

$$0 < x^\top A x \leq \lambda_1 x^\top x, \qquad \therefore \frac{1}{x^\top A x} \geq \frac{1}{\lambda_1 x^\top x}$$

- また $A > 0$ に対して,A^{-1} の固有値は $0 < 1/\lambda_1 \leq 1/\lambda_2 \leq \cdots \leq 1/\lambda_N$ となり,さらに

$$x^\top A^{-1} x \leq \frac{1}{\lambda_N} x^\top x, \qquad \therefore x^\top x \geq \lambda_N x^\top A^{-1} x$$

- 以上から,任意の$x \in \mathbb{R}^N$,ただし$x \neq 0$ に対して

$$\frac{x^\top x}{x^\top A x} \geq \frac{\lambda_N x^\top A^{-1} x}{x^\top A x} \geq \frac{\lambda_N x^\top A^{-1} x}{\lambda_1 x^\top x}$$

証明

- 残差ベクトル $\mathrm{r}[n] = \mathrm{b} - \mathrm{A}\mathrm{x}[n]$ に関して

$$\mathrm{r}[n+1] = \mathrm{r}[n] - \alpha[n]\mathrm{A}\mathrm{r}[n]$$

であったので,

$$\begin{aligned}
&\mathrm{r}[n+1]^\top \mathrm{A}^{-1}\mathrm{r}[n+1] \\
&\quad = (\mathrm{r}[n] - \alpha[n]\mathrm{A}\mathrm{r}[n])^\top \mathrm{A}^{-1} (\mathrm{r}[n] - \alpha[n]\mathrm{A}\mathrm{r}[n]) \\
&\quad = \mathrm{r}[n]^\top \mathrm{A}^{-1}\mathrm{r}[n] - \alpha[n]\mathrm{r}[n]^\top \mathrm{r}[n]
\end{aligned}$$

また,

$$\alpha[n] = \frac{\mathrm{r}[n]^\top \mathrm{r}[n]}{\mathrm{r}[n]^\top \mathrm{A}\mathrm{r}[n]} \geq \frac{\lambda_N \mathrm{r}[n]^\top \mathrm{A}^{-1}\mathrm{r}[n]}{\lambda_1 \mathrm{r}[n]^\top \mathrm{r}[n]}$$

証明

- したがって

$$
\begin{aligned}
\mathsf{r}[\mathsf{n}+1]^\top \mathsf{A}^{-1}\mathsf{r}[\mathsf{n}+1] &\le \mathsf{r}[\mathsf{n}]^\top \mathsf{A}^{-1}\mathsf{r}[\mathsf{n}] - \frac{\lambda_{\mathsf{N}}}{\lambda_1}\mathsf{r}[\mathsf{n}]^\top \mathsf{A}^{-1}\mathsf{r}[\mathsf{n}] \\
&= \left(1 - \frac{\lambda_{\mathsf{N}}}{\lambda_1}\right)\mathsf{r}[\mathsf{n}]^\top \mathsf{A}^{-1}\mathsf{r}[\mathsf{n}] \\
&\le \left(1 - \frac{\lambda_{\mathsf{N}}}{\lambda_1}\right)^2 \mathsf{r}[\mathsf{n}-1]^\top \mathsf{A}^{-1}\mathsf{r}[\mathsf{n}-1] \\
&\le \cdots \\
&\le \left(1 - \frac{\lambda_{\mathsf{N}}}{\lambda_1}\right)^{\mathsf{n}+1} \mathsf{r}[0]^\top \mathsf{A}^{-1}\mathsf{r}[0] \\
&\to 0 \quad (\mathsf{n} \to \infty)
\end{aligned}
$$

証明

- ゆえに $\mathsf{r}[\mathsf{n}]^\top \mathsf{A}^{-1}\mathsf{r}[\mathsf{n}] = \|\mathsf{r}[\mathsf{n}]\|^2_{\mathsf{A}^{-1}} \to \mathbf{0}, \mathsf{n} \to \infty$
- したがって,$\lim_{\mathsf{n}\to\infty} \mathsf{r}[\mathsf{n}] = \lim_{\mathsf{n}\to\infty}(\mathsf{b} - \mathsf{A}\mathsf{x}[\mathsf{n}]) = \mathbf{0}$
- すなわち,x[n] は方程式 $\mathsf{A}\mathsf{x} = \mathsf{b}$ の厳密解に収束する.
- 解の一意性は,A が正則であることより明らか.

練習問題

正則行列 $\mathsf{A} \in \mathbb{R}^{\mathsf{N}\times\mathsf{N}}$ の固有値を $\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_\mathsf{N}$ としたとき, 逆行列A^{-1}の固有値は$1/\boldsymbol{\lambda}_1, \ldots, 1/\boldsymbol{\lambda}_\mathsf{N}$となることを示せ.

ヒント

- 行列 A の固有値をλ,対応する固有ベクトルをxとおくと,

$$\mathsf{Ax} = \lambda \mathsf{x}.$$

- A は正則だから,$\lambda \neq 0$.

解答例

- 行列 A の固有値をλ,対応する固有ベクトルをxとおくと

$$Ax = \lambda x$$

が成り立つ.

- 行列 A は正則だから,$\lambda \neq 0$ である.したがって,

$$\lambda^{-1} x = A^{-1} x$$

が成り立つ.

- すなわち,λ^{-1} は逆行列 A^{-1} の固有値であり, x が固有ベクトルとなることがわかる.
- 上の議論はAの任意の固有値$\lambda_1, \ldots, \lambda_N$について成り立つので, A^{-1}の固有値は $1/\lambda_1, \ldots, 1/\lambda_N$となることがわかる.

最急降下法の収束の速さ

- $A > 0$ とし,$x \in \mathbb{R}^N$ に対して,
$$\|x\|_A = \sqrt{x^\top A x}$$
とおく(これは$\mathbb{R}^N$のノルムとなる).
- 行列Aの条件数を
$$k(A) = \|A^{-1}\|_2 \|A\|_2$$
とおく.
- このとき, $k(A) = \lambda_{max}(A)/\lambda_{min}(A) \geq 1$ かつ
$$\|x[n] - x^*\|_A \leq \left(\frac{k(A) - 1}{k(A) + 1}\right)^n \|x[0] - x^*\|_A$$
が成り立つ.

最急降下法の収束の速さ

$k(A) = \lambda_{\max}(A)/\lambda_{\min}(A) \geq 1$ かつ

$$\|x[n] - x^*\|_A \leq \left(\frac{k(A) - 1}{k(A) + 1}\right)^n \|x[0] - x^*\|_A$$

より

- 収束の次数は1
- 条件数 $k(A)$ が1に近い,すなわち行列 A の最大固有値と最小固有値が近いほど,収束は速い.
- 条件数 $k(A)$ が大きいほど,すなわち行列 A の最大固有値と最小固有値の比が大きいほど, 収束は遅くなる可能性がある.

線形方程式の丸め誤差

- 方程式 $Ax = b$,A は正則,$b \neq 0$とする.
- 行列 A とベクトルb に丸め誤差

$$A \to A + \Delta A, \quad b \to b + \Delta b$$

- 丸められた方程式

$$(A + \Delta A)x = b + \Delta b$$

の解を $x^* + \Delta x^*$ とおく.

丸め誤差の影響

行列 A に対する丸め誤差 ΔA は十分小さく,

$$\|\Delta A\| < \frac{1}{\|A^{-1}\|}$$

が成立しているとする. このとき次の不等式が成り立つ.

$$\frac{\|\Delta x^*\|}{\|x^*\|} \leq \frac{k(A)}{1 - k(A)\|\Delta A\| \, \|A\|^{-1}} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|} \right).$$

ここで k(A) は行列 A の条件数 $k(A) := \|A\|\|A^{-1}\|$.

丸め誤差の影響

$$\frac{\|\Delta x^*\|}{\|x^*\|} \leq \frac{k(A)}{1 - k(A)\|\Delta A\| \, \|A\|^{-1}} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|} \right).$$

- 条件数 k(A) が小さければ,丸め誤差の方程式への影響は少ない.
- 逆に条件数 k(A) が大きければ,丸め誤差の方程式への影響は大きくなる可能性がある.

反復法のブロック線図表現

- Jacobi法

$$\begin{aligned} x[n+1] &= -D^{-1}(E+F)x[n] + D^{-1}b \\ &= x[n] - D^{-1}(D+E+F)x[n] + D^{-1}b \\ &= x[n] - D^{-1}(Ax[n] - b) \end{aligned}$$

- Gauss-Seidel法

$$x[n+1] = x[n] - (D+E)^{-1}(Ax[n] - b)$$

- 最急降下法

$$x[n+1] = x[n] - \alpha[n](Ax[n] - b)$$

反復法のブロック線図表現

$$x[n+1] = x[n] - P(Ax[n] - b)$$