「数値解析」講義ノート

2016年度·後期

Karel Svadlenka · 京都大学理学研究科数学教室

概要

偏微分方程式は自然現象のモデルとして現れるが、実際の問題では方程式が複雑だったり、非線形になったり、方程式が成り立つ領域が複雑な形状をしたりして、解析的に解くことができないことがよくある。この講義では、このような複雑な偏微分方程式に対する(初期値)境界値問題を差分法を用いて近似的に解く数値解法とその収束性や安定性などの数学的な性質を紹介する。また、差分法を適用することにより得られる大規模な連立線形方程式を効率よく解く方法についても触れる。

1 楕円型方程式

ポアソン方程式に対する境界値問題という簡単な例題を用いて,数値計算と数値解析の主な手法を具体的に見る. その後,この考察から現れる課題を解決するのに必要な理論を展開する.

1.1 モチベーションとなる例題

 $\Omega \subset \mathbb{R}^2$ を十分滑らかな境界をもつ領域とする。 Ω の境界 $\partial\Omega$ を二つの部分 $\overline{\Gamma}_D$ (Dirichlet 境界) と $\overline{\Gamma}_N$ (Neumann 境界)に分ける: $\partial\Omega = \overline{\Gamma}_D \cup \overline{\Gamma}_N$. ただし, Γ_D と Γ_N がそれぞれ端点を含まない有限個の弧からなり, $\Gamma_D \cap \Gamma_N = \emptyset$ が成り立つとする。

記号.

• 関数 $v \in C^2(\Omega)$ に対し、 ラプラス作用素 Δ を次の式で定義する:

$$\Delta v = \sum_{i=1}^{2} \frac{\partial^2 v}{\partial x_i^2}.$$

- 点 $x \in \partial \Omega$ に対し, $\mathbf{n}(x)$ を点 x における $\partial \Omega$ への外向き単位法線ベクトルとする.
- 関数 $v \in C^1(\overline{\Omega})$ と点 $x \in \partial \Omega$ に対し、点 x における v の 法線微分 を次で定義する:

$$\frac{\partial v}{\partial n}(x) = \nabla v(x) \cdot \boldsymbol{n}(x).$$

ただし、 $\nabla v = (\frac{\partial v}{\partial x_1}, \frac{\partial v}{\partial x_2})$ は v の勾配ベクトルである.

「記号」の終わり _____

ポアソン方程式に対する境界値問題を考える.

解く問題

次の条件を満たす関数 $u \in C^2(\overline{\Omega})$ を求める:

$$-\Delta u = f \qquad \text{in } \Omega \tag{1.1}$$

$$u = u_D \quad \text{on } \Gamma_D$$
 (1.2)

$$\frac{\partial u}{\partial n} + ku = \varphi_N \quad \text{on } \Gamma_N \tag{1.3}$$

このような問題は電磁気学,流体力学,弾性体理論などに応用がある.

この問題の入力データは Ω , Γ_D , Γ_N , f, u_D , φ_N とk で,未知関数はu である.データに対して次の仮定をおく.

- (A1) $f \in C(\overline{\Omega}); u_D \in C(\overline{\Gamma}_D); \varphi_N, k \in C(\overline{\Gamma}_N)$
- (A2) $k \geq 0$ on $\overline{\Gamma}_N$
- (A3) 次の条件のうち、一つ以上の条件が成り立つ:
 - Γ_D は空でない弧を含む
 - $k(\tilde{x}) > 0$ を満たす $\tilde{x} \in \overline{\Gamma}_N$ が存在する

1.2 数値スキームを作る

問題(1.1)-(1.3)の差分法による解き方を二つのステップに分ける.

1. 格子を構成する

基準点 $(x_0, y_0) \in \mathbb{R}^2$ と差分幅 h > 0 を決めて、 \mathbb{R}^2 での格子 S_h を

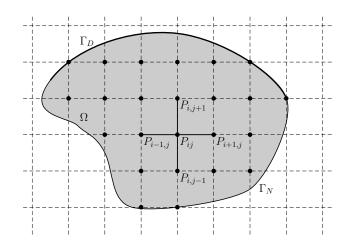
$$S_h = \{P_{ij} = (x_i, y_j) = (x_0 + ih, y_0 + jh), i, j = 0, \pm 1, \pm 2, \dots\}$$

で定義する. P_{ij} を 節点, $x = x_i, y = y_j$ を格子線の方程式とよぶ.

 Ω での格子を

$$\Omega_h = \mathcal{S}_h \cap \Omega$$

とする.



点 $P_{i-1,j}, P_{i+1,j}$ と $P_{i,j-1}, P_{i,j+1}$ を端点にもつ開線分がすべて Ω に含まれるような点 P_{ij} を正則な節点 と言う.

2. 格子点において問題を離散化する

以下では、 $u \in C^4(\overline{\Omega})$ であると仮定する.

(a) $P_{ij} \in \Omega_h$ が正則な節点の場合

微分方程式を P_{ij} で近似するために、テーラー展開を用いる。

$$u(P_{i\pm 1,j}) = u(P_{ij}) \pm h \frac{\partial u}{\partial x}(P_{ij}) + \frac{h^2}{2} \frac{\partial^2 u}{\partial x^2}(P_{ij}) \pm \frac{h^3}{3!} \frac{\partial^3 u}{\partial x^3}(P_{ij}) + \frac{h^4}{4!} \frac{\partial^4 u}{\partial x^4}(x_i \pm \theta_i^{\pm} h, y_j)$$

ただし、 $\theta_i^{\pm} \in [0,1]$. この式を足して、整理すると、

$$\frac{\partial^{2} u}{\partial x^{2}}(P_{ij}) = \frac{u(P_{i-1,j}) - 2u(P_{ij}) + u(P_{i+1,j})}{h^{2}} - \frac{h^{2}}{4!} \left\{ \frac{\partial^{4} u}{\partial x^{4}}(x_{i} + \theta_{i}^{+}h, y_{j}) + \frac{\partial^{4} u}{\partial x^{4}}(x_{i} - \theta_{i}^{-}h, y_{j}) \right\}$$

を得る. 最後の行の誤差項を $\varepsilon_h^x(P_{ij})$ と書くと,

$$|\varepsilon_h^x(P_{ij})| \le \frac{2h^2}{4!} \max_{(x,y)\in\overline{\Omega}} \left| \frac{\partial^4 u}{\partial x^4}(x,y) \right| = K_x h^2$$

のように評価できる. 同様にして、yについての微分を計算しておくと、

$$\frac{\partial^2 u}{\partial y^2}(P_{ij}) = \frac{u(P_{i,j-1}) - 2u(P_{ij}) + u(P_{i,j+1})}{h^2} + \varepsilon_h^y(P_{ij}), \qquad |\varepsilon_h^y(P_{ij})| \le K_y h^2.$$

上の2式を節点 P_{ij} での偏微分方程式(1.1)に代入すると,

$$\frac{1}{h^2} \left\{ -u(P_{i-1,j}) - u(P_{i+1,j}) - u(P_{i,j-1}) - u(P_{i,j+1}) + 4u(P_{ij}) \right\} = f(P_{ij}) + \varepsilon_h(P_{ij}). \tag{1.4}$$

ここで、 P_{ij} と h に依らない K があり、誤差 ε_h は

$$\varepsilon_h(P_{ij}) = \varepsilon_h^x(P_{ij}) + \varepsilon_h^y(P_{ij}), \qquad |\varepsilon_h(P_{ij})| \le Kh^2, \quad K = K_x + K_y$$

を満たす.よって,節点 P_{ij} における離散化による誤差は $\varepsilon_h(P_{ij}) = O(h^2)$ である. この誤差を無視すると,真の解の値 $u(P_{ij})$ は誤差項を落とした式 (1.4) を正確に満たされなくなる.そこで,誤差項を落とした式 (1.4) を正確に満たす近似値を u_h として, $U_{ij} := u_h(P_{ij})$ という記号を導入すれば,

$$\frac{1}{h^2} \left\{ -U_{i-1,j} - U_{i+1,j} - U_{i,j-1} - U_{i,j+1} + 4U_{ij} \right\} = f_{ij}. \tag{1.5}$$

もし $X \in \{P_{i\pm 1,j}, P_{i,j\pm 1}\}$ が $\overline{\Gamma}_D$ 上にある節点なら, $u_h(X) := u_D(X)$ とおいて,近似式 (1.5) にその値を代入し,右辺に移項する.

例 1.1 ポアソン方程式 $-\Delta u = f$ を正方形領域 $\Omega = (0,1) \times (0,1)$ で考える. $\Gamma_D = \partial \Omega$, $\Gamma_N = \emptyset$ として,境界条件を $u|_{\partial \Omega} = u_D$ とする.

基準点を $(x_0, y_0) = (0, 0)$ にして,領域の両辺を N 等分すると, $h = \frac{1}{N}$, $P_{ij} = (x_i, y_j) = (ih, jh)$ を得る.この場合,

$$\Omega_h = \{ P_{ij}; i, j = 1, \dots, N-1 \}$$

となり、すべての節点 $P_{ij}\in\Omega_h$ が正則な節点である。 近似方程式は

$$\frac{1}{h^2} \left\{ -U_{i-1,j} - U_{i+1,j} - U_{i,j-1} - U_{i,j+1} + 4U_{ij} \right\} = f_{ij}, \quad i, j = 1, \dots, N-1$$

と書けるが、Uの添え字のどちらかが0またはNのときの値は境界条件より決まる:

$$U_{0j} = u_D(P_{0j}), \ U_{Nj} = u_D(P_{Nj}), \ U_{i0} = u_D(P_{i0}), \ U_{iN} = u_D(P_{iN}), \ i, j = 1, \dots N-1.$$

したがって、近似解 u_h の $n=(N-1)^2$ 個の節点 $P_{ij}\in\Omega_h$ における未知の値に対する $n=(N-1)^2$ 本の方程式が得られた.

この連立方程式を行列を用いて書くが、式を書きやすくするために未知数と近似方程式に番号をふっておく。節点を $\Omega_h = \{P_1, \dots, P_n\}$ と表したとき,i-番目の未知数を節点 P_i における未知数(つまり, $u_h(P_i)$)とし,i-番目の方程式を節点 P_i において構成された方程式とする。未知数のベクトルを $\mathbf{u}_h = (u_h(P_1), \dots, u_h(P_n))^T$ と定義すれば,上記の方程式を

$$A_h \boldsymbol{u}_h = \boldsymbol{F}_h$$

と書ける.

行列 A_h について次のことが言える:

- 対角線上の要素は ⁴/₁₂ であるから,正である.
- 対角線以外の要素は0または $-\frac{1}{62}$ であるから、非正である.
- 行列は優対角である(定義2.5を参照).

- $P_{i\pm 1,j}, P_{i,j\pm 1}$ のどれかが Γ_D に位置するような節点 P_{ij} において構成された方程式 に対応する行列 A_h の行において、狭義優対角性が成り立つ。
- A_h の有向グラフが強連結であるから, A_h は既約行列である(定義2.8と定理2.10を参照).

したがって、 A_h は既約優対角で、対称で、正定値行列である。対称性は、j-番目の方程式における未知数 $u_h(P_j)$ の係数が、j-番目の方程式における未知数 $u_h(P_i)$ の係数に等しいことから従う。

$$\frac{u(P_{i-1,j}) - 2u(P_{ij}) + u(P_{i+1,j})}{h^2} = \frac{\partial^2 u}{\partial x^2}(P_{ij}) + \frac{2h^2}{4!} \frac{\partial^4 u}{\partial x^4}(P_{ij}) + O(h^4)$$

がわかる。4階微分を2階微分の差分化として書いて

$$\frac{\partial^4 u}{\partial x^4}(P_{ij}) = \frac{\frac{\partial^2 u}{\partial x^2}(P_{i-1,j}) - 2\frac{\partial^2 u}{\partial x^2}(P_{ij}) + \frac{\partial^2 u}{\partial x^2}(P_{i+1,j})}{h^2} + O(h^2),$$

上の式に代入して, 差分化オペラータを

$$\delta_x^2 u(P_{ij}) = \frac{u(P_{i-1,j}) - 2u(P_{ij}) + u(P_{i+1,j})}{h^2}$$

で定義すると,

$$\delta_x^2 u(P_{ij}) = \left(1 + \frac{h^2}{12} \delta_x^2\right) \frac{\partial^2 u}{\partial x^2} (P_{ij}) + O(h^4)$$

を得る。y-方向に対して同じような計算を行って、まとめると、ポアソン方程式 (1.1) のより高精度の離散化

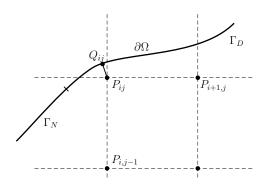
$$\left(1 + \frac{h^2}{12}\delta_x^2\right)^{-1}\delta_x^2 u(P_{ij}) + \left(1 + \frac{h^2}{12}\delta_y^2\right)^{-1}\delta_y^2 u(P_{ij}) = f(P_{ij}) + O(h^4)$$

にたどりつく.整理して、 $O(h^4)$ の誤差項を無視すると、近似方程式

$$\frac{1}{6} \left(U_{i+1,j+1} + U_{i+1,j-1} + U_{i-1,j+1} + U_{i-1,j-1} \right) + \frac{2}{3} \left(U_{i+1,j} + U_{i-1,j} + U_{i,j+1} + U_{i,j-1} \right)
- \frac{10}{3} U_{ij} = \frac{h^2}{12} (f_{i+1,j} + f_{i-1,j} + f_{i,j+1} + f_{i,j-1} + 8f_{ij})$$
(1.6)

を得る。この式では節点 P_{ij} とその近隣の 8 点での近似値を利用しているので、9 点スキームとよばれる。

(b) $P_{ij} \in \Omega_h$ が正則でない節点の場合



- (α) P_{ij} が $\overline{\Gamma}_D$ の近くにある場合、離散化の方法を三通り紹介する.
 - (i) $\overline{\Gamma}_D$ 上の境界条件を直接 P_{ij} に移転させる方法 節点 P_{ij} に最も近い $\overline{\Gamma}_D$ の点 Q_{ij} をとって,

$$U_{ij} = u_h(P_{ij}) := u_D(Q_{ij})$$

とし、近似方程式においてこの項を右辺に移項する。そうすると、対応する係数行列の行は狭義優対角になる。

離散化による誤差,すなわち, $u(P_{ij})$ の値を $u(Q_{ij}) = u_D(Q_{ij})$ で置き換えることによって生じる誤差,を調べる.まず, $\alpha, \beta \in [-\sqrt{2}, \sqrt{2}]$ が存在し,

$$Q_{ij} = (x_i + \alpha h, y_j + \beta h)$$

と書けることに注意する. 点 P_{ij} と点 Q_{ij} を端点にもつ線分が Ω に含まれると仮定して、テーラー展開を用いる:

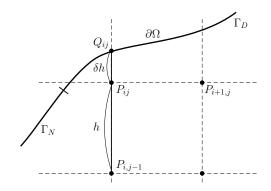
$$u(Q_{ij}) = u(P_{ij}) + h\left(\alpha \frac{\partial u}{\partial x}(P_{ij}) + \beta \frac{\partial u}{\partial y}(P_{ij})\right) + \frac{h^2}{2} \left[\alpha \frac{\partial}{\partial x} + \beta \frac{\partial}{\partial y}\right]^2 u(\tilde{P}_{ij}).$$

ただし、 \tilde{P}_{ij} は $P_{ij}Q_{ij}$ 上の点である。近似の誤差 $\tilde{\varepsilon}_h(P_{ij})$ を $u(Q_{ij})=u(P_{ij})+\tilde{\varepsilon}_h(P_{ij})$ で定義すると、

$$\begin{split} |\tilde{\varepsilon}(P_{ij})| & \leq h\left(\left|\alpha\right|\left|\frac{\partial u}{\partial x}(P_{ij})\right| + \left|\beta\right|\left|\frac{\partial u}{\partial y}(P_{ij})\right|\right) \\ & + \frac{h^2}{2}2\left(\left|\frac{\partial^2 u}{\partial x^2}(\tilde{P}_{ij})\right| + 2\left|\frac{\partial^2 u}{\partial x \partial y}(\tilde{P}_{ij})\right| + \left|\frac{\partial^2 u}{\partial y^2}(\tilde{P}_{ij})\right|\right) \\ & \leq h\sqrt{2}\max_{\overline{\Omega}}\left(\left|\frac{\partial u}{\partial x}\right| + \left|\frac{\partial u}{\partial y}\right|\right) + h^2\max_{\overline{\Omega}}\left(\left|\frac{\partial^2 u}{\partial x^2}\right| + 2\left|\frac{\partial^2 u}{\partial x \partial y}\right| + \left|\frac{\partial^2 u}{\partial y^2}\right|\right) \\ & \leq K_1h + K_2h^2. \end{split}$$

よって, $|\tilde{\epsilon}_h(P_{ij})| = O(h)$ となり,正則な節点と比べて,近似オーダーが一つ落ちる.

(ii) Collatz による線形補間



節点 $P_{i,j-1}$ と節点 P_{ij} を結ぶ直線の $\overline{\Gamma}_D$ との交点を Q_{ij} として, $P_{ij}Q_{ij}$ の長さ を δh とする. $P_{i,j-1}$ と Q_{ij} の間の線形補間をした関数の P_{ij} での値を用いて $u(P_{ij})$ を近似すると,

$$u(P_{ij}) = \frac{u(Q_{ij}) + \delta u(P_{i,j-1})}{1 + \delta} + \varepsilon_h(P_{ij})$$

の式になるが、テーラー展開により $|\varepsilon_h(P_{ij})| = O(h^2)$ という誤差オーダーが得られる.

 $u(Q_{ij})$ は境界条件より定まるので、それを右辺に移項し、

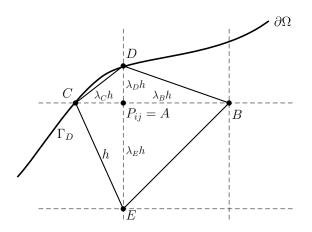
$$-\delta u(P_{i,j-1}) + (1+\delta)u(P_{ij}) = u_D(Q_{ij}) + (1+\delta)\varepsilon_h(P_{ij})$$

より, 近似方程式

$$-\delta U_{i,j-1} + (1+\delta)U_{ij} = u_D(Q_{ij})$$

を得る。この式を連立方程式に加えると、係数行列 A_h の対称性が崩れる。

(iii) 一般 5 点スキームを用いた近似



点 A, B, C, D, E とパラメータ $\lambda_B, \lambda_C, \lambda_D, \lambda_E \in (0, 1]$ を図のように決める. 四 角形 R = BECD における微分方程式を用いて、近似方程式を導く.

関数 $w \in C(\overline{R})$ は次の条件を満たす関数とする:

 $w|_{\partial R}=0$, w(A)=1, w は三角形 ABD, ABE, ACD, ACE 上で線形

関数wを微分方程式(1.1)にかけて,Rで積分し,グリーンの定理を適用する.

$$\int_{R} fw \, dx \, dy = -\int_{R} \Delta u \cdot w \, dx \, dy = -\int_{R} \left(\frac{\partial^{2} u}{\partial x^{2}} + \frac{\partial^{2} u}{\partial y^{2}} \right) w \, dx \, dy$$

$$= -\int_{\partial R} \left(\frac{\partial u}{\partial x} w n_{x} + \frac{\partial u}{\partial y} w n_{y} \right) \, dS + \int_{R} \left(\frac{\partial u}{\partial x} \frac{\partial w}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial w}{\partial y} \right) \, dx \, dy$$

w が R の境界で消えるので, $u \in C^2(\overline{\Omega})$ であれば, 結局,

$$\int_{B} fw \, dx \, dy = \int_{B} \left(\frac{\partial u}{\partial x} \frac{\partial w}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial w}{\partial y} \right) \, dx \, dy \tag{1.7}$$

この式での積分領域を $R=R_1\cup R_2$, $R_1=BDE$, $R_2=DCE$ と分けて,積分を数値積分により近似し,微分を差分により近似する:

$$\begin{split} \frac{\partial u}{\partial x}\Big|_{R_1} &\approx \frac{u(B) - u(A)}{\lambda_B h}, \quad \frac{\partial w}{\partial x}\Big|_{R_1} \approx \frac{w(B) - w(A)}{\lambda_B h} = -\frac{1}{\lambda_B h} \\ \frac{\partial u}{\partial x}\Big|_{R_2} &\approx \frac{u(A) - u(C)}{\lambda_C h}, \quad \frac{\partial w}{\partial x}\Big|_{R_2} \approx \frac{w(A) - w(C)}{\lambda_C h} = \frac{1}{\lambda_C h} \end{split}$$

式(1.7)の右辺の第一項を計算すると、

$$\int_{R_1} \frac{\partial u}{\partial x} \frac{\partial w}{\partial x} dx dy \approx \frac{u(B) - u(A)}{\lambda_B h} \cdot \frac{-1}{\lambda_B h} |R_1|$$

$$= \frac{u(A) - u(B)}{(\lambda_B h)^2} \frac{1}{2} \lambda_B h^2 (\lambda_D + \lambda_E)$$

$$\approx \frac{u(A) - u(B)}{\lambda_B}$$

$$\int_{R_2} \frac{\partial u}{\partial x} \frac{\partial w}{\partial x} dx dy \approx \frac{u(A) - u(C)}{\lambda_C}$$

$$\int_{R} \frac{\partial u}{\partial x} \frac{\partial w}{\partial x} dx dy \approx -\frac{u(B)}{\lambda_B} - \frac{u(C)}{\lambda_C} + \left(\frac{1}{\lambda_B} + \frac{1}{\lambda_C}\right) u(A)$$

同様にして,

$$\int_{B} \frac{\partial u}{\partial y} \frac{\partial w}{\partial y} \, dx \, dy \approx -\frac{u(D)}{\lambda_{D}} - \frac{u(E)}{\lambda_{E}} + \left(\frac{1}{\lambda_{D}} + \frac{1}{\lambda_{E}}\right) u(A).$$

これを合わせると,

$$-\int_{R} \Delta u \cdot w \, dx \, dy \approx -\frac{u(B)}{\lambda_{B}} - \frac{u(C)}{\lambda_{C}} - \frac{u(D)}{\lambda_{D}} - \frac{u(E)}{\lambda_{E}} + \left(\frac{1}{\lambda_{B}} + \frac{1}{\lambda_{C}} + \frac{1}{\lambda_{D}} + \frac{1}{\lambda_{E}}\right) u(A).$$

(1.7) の左辺の積分を

$$\int_{R} fw \, dx \, dy \approx f(A) \int_{R} w \, dx \, dy$$

と近似するので、 $\int_{\mathbb{R}} w$ の近似値も必要である:

$$\int_{R} w \, dx \, dy \approx w(\frac{A+B}{2})|R_{1}| + w(\frac{A+C}{2})|R_{2}|$$

$$= \frac{1}{2}w(A)\left[\frac{1}{2}\lambda_{B}h^{2}(\lambda_{E} + \lambda_{D}) + \frac{1}{2}\lambda_{C}h^{2}(\lambda_{E} + \lambda_{D})\right]$$

$$= \frac{1}{4}(\lambda_{B} + \lambda_{C})(\lambda_{D} + \lambda_{E})h^{2}.$$

したがって、式(1.7)の近似を最終的に書けば、

$$\begin{split} \frac{1}{h^2} \left[-\frac{u(B)}{\lambda_B} - \frac{u(C)}{\lambda_C} - \frac{u(D)}{\lambda_D} - \frac{u(E)}{\lambda_E} + \left(\frac{1}{\lambda_B} + \frac{1}{\lambda_C} + \frac{1}{\lambda_D} + \frac{1}{\lambda_E} \right) u(A) \right] \\ &= \frac{1}{4} (\lambda_B + \lambda_C) (\lambda_D + \lambda_E) f(A) + \varepsilon_h(A) \end{split}$$

となるので、誤差項を無視すると、近似方程式

$$\frac{1}{h^2} \left[-\frac{U_B}{\lambda_B} - \frac{U_C}{\lambda_C} - \frac{U_D}{\lambda_D} - \frac{U_E}{\lambda_E} + \left(\frac{1}{\lambda_B} + \frac{1}{\lambda_C} + \frac{1}{\lambda_D} + \frac{1}{\lambda_E} \right) u(A) \right]
= \frac{1}{4} (\lambda_B + \lambda_C) (\lambda_D + \lambda_E) f(A)$$
(1.8)

を得る。A が正則な節点ならば, $\lambda_{B,C,D,E}=1$ となるので,このとき上の式が正則な節点の場合に導いたしきと一致することに注意しよう。

真の解が $u\in C^3(\overline{\Omega})$ を満たすと仮定して,この一般 5 点スキームによる離散 化の誤差 $\varepsilon_h(A)$ を評価する.

$$u(B) = u(A) + \lambda_B h \frac{\partial u}{\partial x}(A) + \frac{1}{2} \lambda_B^2 h^2 \frac{\partial^2 u}{\partial x^2}(A) + O(h^3)$$

$$u(C) = u(A) - \lambda_C h \frac{\partial u}{\partial x}(A) + \frac{1}{2} \lambda_C^2 h^2 \frac{\partial^2 u}{\partial x^2}(A) + O(h^3)$$

$$u(D) = u(A) + \lambda_D h \frac{\partial u}{\partial y}(A) + \frac{1}{2} \lambda_D^2 h^2 \frac{\partial^2 u}{\partial y^2}(A) + O(h^3)$$

$$u(E) = u(A) - \lambda_E h \frac{\partial u}{\partial y}(A) + \frac{1}{2} \lambda_E^2 h^2 \frac{\partial^2 u}{\partial y^2}(A) + O(h^3)$$

よって, 誤差は

$$\varepsilon_{h}(A) = -\frac{1}{h^{2}} \left[\frac{1}{\lambda_{B}} (u(B) - u(A)) + \frac{1}{\lambda_{C}} (u(C) - u(A)) + \cdots \right] \\
-\frac{1}{4} (\lambda_{B} + \lambda_{C}) (\lambda_{D} + \lambda_{E}) f(A) \\
= -\frac{1}{2} \left[\lambda_{B} \frac{\partial^{2} u}{\partial x^{2}} (A) + \lambda_{C} \frac{\partial^{2} u}{\partial x^{2}} (A) + \lambda_{D} \frac{\partial^{2} u}{\partial y^{2}} (A) + \lambda_{E} \frac{\partial^{2} u}{\partial y^{2}} (A) \right] \\
-\frac{1}{4} (\lambda_{B} + \lambda_{C}) (\lambda_{D} + \lambda_{E}) f(A) + O(h) \\
= -\frac{\partial^{2} u}{\partial x^{2}} (A) - \frac{\partial^{2} u}{\partial y^{2}} (A) \\
-\frac{1}{2} \left[(\lambda_{B} + \lambda_{C} - 2) \frac{\partial^{2} u}{\partial x^{2}} (A) + (\lambda_{D} + \lambda_{E} - 2) \frac{\partial^{2} u}{\partial y^{2}} (A) \right] \\
-f(A) - \left[\frac{1}{4} (\lambda_{B} + \lambda_{C}) (\lambda_{D} + \lambda_{E}) - 1 \right] f(A) + O(h) \\
= O(1)$$

となり、 $h \to 0$ のとき誤差は0 に収束しない。これは正しい離散化では許されないことである("discretization crime"ともいう)が、後に分くるように、数値スキーム全体の収束には影響しない。

例 1.2 $\Gamma_D = \partial \Omega$ として、次の問題を考える:

$$-\Delta u = f$$
 in Ω , $u|_{\partial\Omega} = u_D$.

 Ω_h のすべての節点 P_{ij} において,方程式を (1.8) のスキームに従って離散化すると,近似方程式は $A_h u_h = F_h$ と書ける。 $A,B \in \Omega_h$ ならば,節点 A で構成した近似方程式における U_B の係数は $-\frac{1}{\lambda_B h^2}$ で,これは節点 B で構成した近似方程式における U_A の係数に等しいので,係数行列 A_h は対称行列である。また,行列 A_h は優対角であるが,(どの h に対しても)境界 $\overline{\Gamma}_D$ 上に位置する節点が存在すれば,狭義優対角にもなる.

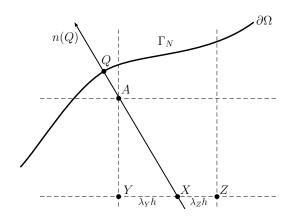
_____「例」の終わり _____

(eta) $P_{ij}=A\in\Omega_h$ が $\overline{\Gamma}_N$ の近くにある場合

ここでロバン条件 (1.3)

$$\frac{\partial u}{\partial n} + ku = \varphi_N \qquad \text{on } \Gamma_N$$

を近似することになる.



境界 Γ_N に直交で節点 A を通る直線を考えて, Γ_N と交わる点を Q,格子線と交わる A に最も近い点を X とする。すると,Q での単位法線ベクトル $\mathbf{n}(Q)=(n_x,n_y)$ はこの直線と同じ向きをもつ。

X はこの格子線にある長さ h の線分 YZ の点であるような節点 Y,Z が存在する. また,Q が境界 $\overline{\Gamma}_N$ に最も近い点であるため, $AQ \subset \overline{\Omega}$ が成り立つ.X と Y の距離を $\lambda_Y h$,そして X と Z の距離を $\lambda_Z h$ とすると, $0 \leq \lambda_Y, \lambda_Z \leq 1, \lambda_Y + \lambda_Z = 1$ となる.

 $u \in C^2(\overline{\Omega})$ を仮定し、 $\frac{\partial u}{\partial n}(A) := \nabla u(A) \cdot \boldsymbol{n}(Q)$ と定義したとき、次の近似を行う:

$$\frac{\partial u}{\partial n}(Q) = \nabla u(Q) \cdot \boldsymbol{n}(Q) \approx \frac{\partial u}{\partial n}(A).$$

 $\frac{\partial u}{\partial n}(A)$ &

$$\frac{\partial u}{\partial n}(A) = \frac{\partial u}{\partial x}(A)n_x + \frac{\partial u}{\partial y}(A)n_y$$
$$= \left(\frac{\partial u}{\partial x}(A)|A - X|n_x + \frac{\partial u}{\partial y}(A)|A - Y|n_y\right)\frac{1}{|A - X|}$$

のように書き直し, $X-A=-(|A-X|n_x,|A-X|n_y)$ に注意しながらテーラー展開を用いると,

$$u(X) = u(A) - \frac{\partial u}{\partial x}(A)|A - X|n_x - \frac{\partial u}{\partial y}(A)|A - X|n_y + O(|A - X|^2)$$

となるので、さらに $h \le |A-X| \le \sqrt{2}h$ に注意すると、

$$\frac{\partial u}{\partial n}(A) = \frac{u(A) - u(X)}{|A - X|} + O(h)$$

を得る。さらに、 $u \in C^2(\overline{\Omega})$ と仮定しているため $\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \in C^1(\overline{\Omega})$ で、 $\frac{\partial u}{\partial x}(Q) = \frac{\partial u}{\partial x}(A) + O(h), \frac{\partial u}{\partial y}(Q) = \frac{\partial u}{\partial y}(A) + O(h)$ が言えるので、

$$\frac{\partial u}{\partial n}(Q) = \frac{\partial u}{\partial n}(A) + O(h), \qquad u(Q) = u(A) + O(h)$$

もわかる。

線分 YZ で線形補間

$$u(X) = \lambda_Z u(Y) + \lambda_Y u(Z) + O(h^2)$$

を適用し、以上で導いた式を全て点 Q におけるロバン条件 (1.3) に代入すれば、

$$\frac{u(A) - \lambda_Z u(Y) - \lambda_Y u(Z) + O(h^2)}{|A - X|} + k(Q)(u(A) + O(h)) + O(h) = \varphi_N(Q),$$

すなわち,

$$-\frac{\lambda_Z}{|A-X|}u(Y) - \frac{\lambda_Y}{|A-X|}u(Z) + u(A)\frac{1+|A-X|k(Q)}{|A-X|} = \varphi_N(Q) + O(h)$$
 (1.9)

を得る。この式の誤差項を無視すれば、Aにおける近似方程式が導出されたことになる。

注. 節点 A が $\overline{\Gamma}_N$ 上にあることもあり得る. この場合も、方程式 (1.9) を基にして同様に近似を行う.

以上の離散化で導いた近似方程式を整理しよう。 $\overline{\Gamma}_D$ の近傍にある節点について,一般 5 点スキームを用いることにする.

問題の離散化

正則な節点 P_{ij}

$$\frac{1}{h^2} \left\{ -u(P_{i-1,j}) - u(P_{i+1,j}) - u(P_{i,j-1}) - u(P_{i,j+1}) + 4u(P_{ij}) \right\} = f(P_{ij}) + \varepsilon_h(P_{ij})$$
(1.10)

誤差:
$$|\varepsilon_h(P_{ij})| \leq Kh^2$$
, $h \in (0, h_0)$

 $\overline{\Gamma}_D$ の近くにある正則でない節点 A

$$\frac{1}{h^2} \left[-\frac{u(B)}{\lambda_B} - \frac{u(C)}{\lambda_C} - \frac{u(D)}{\lambda_D} - \frac{u(E)}{\lambda_E} + \left(\frac{1}{\lambda_B} + \frac{1}{\lambda_C} + \frac{1}{\lambda_D} + \frac{1}{\lambda_E} \right) u(A) \right] \\
= \frac{1}{4} (\lambda_B + \lambda_C) (\lambda_D + \lambda_E) f(A) + \varepsilon_h(A) \tag{1.11}$$

誤差: $|\varepsilon_h(A)| \leq K$, $h \in (0, h_0)$

 $\overline{\Gamma}_N$ 上または近くにある正則でない節点 A

$$-\frac{\lambda_{Z}}{|A-X|}u(Y) - \frac{\lambda_{Y}}{|A-X|}u(Z) + \frac{1 + |A-X|k(Q)}{|A-X|}u(A) = \varphi_{N}(Q) + \varepsilon_{h}(A) \quad (1.12)$$

誤差: $|\varepsilon_h(P_{ij})| \leq Kh$, $h \in (0, h_0)$

近似方程式は $P_{ij} \in \tilde{\Omega}_h := \Omega_h \cup [\mathcal{S}_h \cap \overline{\Gamma}_N]$ の各点において構成した。また、未知数は $\tilde{\Omega}_h$ の各点における値であるので、

近似方程式の数 = 未知数の数 = card $\tilde{\Omega}_h = n_h$.

 $\tilde{\Omega}_h$ の節点に $\tilde{\Omega}_h = \{P_1, \dots, P_{n_h}\}$ のように任意に番号をふって,i-番目の未知数を P_i における近似値とし,i-番目の近似方程式を節点 P_i で構成された近似方程式としておく.これにより, $n_h \times n_h$ の行列 A_h を係数行列にもつ連立一次方程式

$$A_h \boldsymbol{U}_h = \boldsymbol{F}_h$$

が得られる.ここで, U_h は節点 P_i における近似値 $u_h(P_i)$ を並べたベクトルを意味する.

係数行列 A_h の性質を調べる.

- 対角要素は正で、それ以外の要素はゼロ以下である。
- 優対角行列である、狭義の優対角性は次の場合に満たされる:
- (1.10) 正則な節点:となりの節点が $\overline{\Gamma}_D$ 上にあるような節点 P_{ij} が存在する.
- (1.11) $\overline{\Gamma}_D$ の近くの正則でない節点:節点 B, C, D, E のうちのどれかが $\overline{\Gamma}_D$ 上にある (必ず満たされる).
- (1.12) $\overline{\Gamma}_N$ の近くの正則でない節点:k(Q) > 0 である.

狭義優対角になる近似方程式が一つ以上あるかが重要となる。そこで、最初においた仮定(A3)を思い出す:「次の条件のうち、一つ以上の条件が成り立つ

- $-\Gamma_D$ は空でない弧を含む
- $-k(\tilde{x}) > 0$ を満たす $\tilde{x} \in \overline{\Gamma}_N$ が存在する」

最初の「 Γ_D は空でない弧を含む」の場合,格子が十分細ければ, $\bar{\Gamma}_D$ の近くにある Ω_h の節点が必ず存在する.そのとき,(1.10) または (1.11) を使って離散化するので,この節点に対応する近似方程式は狭義優対角になる.後者の「 $k(\tilde{x})>0$ を満たす $\tilde{x}\in\bar{\Gamma}_N$ が存在する」の場合,k が $\bar{\Gamma}_N$ で連続であるから,近傍 $O(\tilde{x})$ が存在しすべての $\xi\in O(\tilde{x})\cap\bar{\Gamma}_N$ に対し $k(\xi)>0$ が成り立つ.よって,格子が十分細ければ,k(Q)>0 を満たす近似方程式 (1.12) が存在する.結論として, $h_1\in(0,h_0)$ が存在し,すべての $h\in(0,h_1)$ に対し狭義優対角になる近似方程式が一つ以上存在するということが言える.

• 係数行列の既約性について考えると、格子が粗いとき、 A_h のグラフが連結でない可能性がある。しかし、格子が十分細ければ、グラフは連結になるということがわかる(証明は境界の滑らかさを利用するが、ここでは割愛する)。

系 1.3 すべての $h \in (0, h_1)$ に対し、連立近似方程式 $A_h U_h = F_h$ がただ一つの解をもつような $h_1 > 0$ が存在する.

Proof. 上で(証明なしで)述べたように、h が十分小さければ、係数行列 A_h が狭義優対角または既約優対角になる。よって(定理2.28)、 A_h は単調行列で、その解は一意に決まる(定理2.27)。

注.

- 係数行列は大きい $(n_h \gg 1)$ が,疎行列であるためメモリーへの負担が大きくない.また,連立方程式を解くとき,疎行列であることを利用して反復法を適用することが多い.
- $\overline{\Gamma}_D = \partial \Omega$ のとき、正則でない節点で (1.11) の近似を用いれば、 A_h は対称行列で、正定値となる.

1.3 スキームの誤差評価

次に近似解 U_h と真の解 u の差について調べる。近似解 U_h は連立方程式 $A_hU_h=F_h$ の (唯一の) 解であるが,真の解と比べるために,真の解の $\tilde{\Omega}_h$ の節点における値を並べたベクトル $u_h:=(u(P_1),\dots,u(P_{n_h}))^T$ を導入する。このベクトルは上記の連立方程式を離散化による誤差 ε_h を除いて満たす: $A_hu_h=F_h+\varepsilon_h$. 数値スキームの 誤差 を $\eta_h=u_h-U_h$ で定義すれば,この誤差は連立方程式

 $A_h \boldsymbol{\eta}_h = \boldsymbol{\varepsilon}_h$

を満たす。ただし、以降では (1.10)–(1.12) のスキームを用いることにして、 $h \in (0, h_1)$ に対して A_h が単調行列であるような $h_1 > 0$ が存在することを認める.

定理 1.4 仮定 (A1)–(A3) が満たされ、境界値問題 (1.1)–(1.3) の解 $u \in C^4(\overline{\Omega})$ が存在するとする。そのとき、定数 $h_2 \in (0,h_1)$ と M>0 が存在し、

$$\|\boldsymbol{\eta}_h\|_{\infty} \le Mh \qquad \forall h \in (0, h_2)$$

が成り立つ.

ただし、 $\overline{\Gamma}_D = \partial \Omega$ の場合、この主張を次のように改良できる:

$$\|\boldsymbol{\eta}_h\|_{\infty} \le Mh^2 \qquad \forall h \in (0, h_2).$$

Proof. $z \in C^4(\overline{\Omega})$ を次の問題の解とする:

$$-\Delta z=1 \quad \text{in } \Omega, \qquad z|_{\Gamma_D}=1, \qquad rac{\partial z}{\partial n}+kz=1 \quad \text{on } \Gamma_N.$$

証明の基本方針は Gershgorin の方法を用いることである(定理2.29)。z を求める問題の離散化による係数行列はu を求める問題の離散化による係数行列 A_h と一致する。対応する近似連立方程式を $A_h \mathbf{Z}_h = \boldsymbol{\vartheta}_h$ と書き,真の解のベクトルを $\mathbf{z}_h := (z(P_1), \dots, z(P_{n_h}))^T$ とすると,

$$A_h \boldsymbol{z}_h = \boldsymbol{\vartheta}_h + \boldsymbol{\tau}_h, \qquad \boldsymbol{\tau}_h = (\tau_h(P_1), \dots, \tau_h(P_{n_h}))^T$$

の式を得る.

上で計算した離散化による誤差を思い出すと、それぞれの節点の種類に対して、 τ_h を次のように評価できる:

$$(1.10) |\tau_h(P_{ij})| \le \tilde{K}h^2, h \in (0, h_0)$$

(1.11)
$$|\tau_h(A)| \le \tilde{K}, \qquad h \in (0, h_0)$$

(1.12)
$$|\tau_h(A)| \le \tilde{K}h, \quad h \in (0, h_0).$$

右辺ベクトル ϑ_h の成分を具体的に見ていくと,

(1.10)
$$\vartheta_h(P_{ij}) = 1$$
 または $1 + \frac{1}{h^2}$ 、いずれにしても ≥ 1

(1.11)
$$\vartheta_h(A) = \frac{1}{4}(\lambda_B + \lambda_C)(\lambda_D + \lambda_E) + \frac{1}{\lambda_X h^2} \ge \frac{1}{h^2}$$

$$(1.12) \vartheta_h(A) = 1$$

(1.11) では、A のとなりの節点のどれか(ここで X とした)が $\overline{\Gamma}_D$ 上にあるからそこでの値を右辺に移項することを考慮した。以上のことを用いて、 ε_h を $\vartheta_h + \tau_h$ でおさえる。

(1.10) このとき $|\tau_h(P_{ij})| \leq \tilde{K}h^2$ であるから, $h_2 \in (0, h_1)$ が存在し,

$$\vartheta_h(P_{ij}) + \tau_h(P_{ij}) \ge \vartheta_h(P_{ij}) - |\tau_h(P_{ij})| \ge 1 - \tilde{K}h^2 \ge \frac{1}{2} \qquad \forall h \in (0, h_2).$$

したがって,

$$|\varepsilon_h(P_{ij})| \le Kh^2 = \frac{1}{2}2Kh^2 \le 2Kh^2(\vartheta_h(P_{ij}) + \tau_h(P_{ij})) \qquad \forall h \in (0, h_2).$$

(1.11) 同様に考えて,

$$\vartheta_h(A) + \tau_h(A) \ge \frac{1}{h^2} - \tilde{K} = \frac{1}{h^2} (1 - \tilde{K}h^2) \ge \frac{1}{2h^2} \quad \forall h \in (0, h_2).$$

したがって,

$$|\varepsilon_h(A)| \leq K = 2Kh^2 \frac{1}{2h^2} \leq 2Kh^2 (\vartheta_h(A) + \tau_h(A)) \qquad \forall h \in (0, h_2).$$

(1.12) 前項と同様なので、詳細を省略して、

$$|\varepsilon_h(A)| \le 2Kh(\vartheta_h(A) + \tau_h(A)) \quad \forall h \in (0, h_2).$$

以上とまとめると、 $h_2 > 0$ が存在し、

$$|\varepsilon_h(P_{ij})| \le K^* h(\vartheta_h(P_{ij}) + \tau_h(P_{ij})) \qquad \forall P_{ij} \in \tilde{\Omega}_h \ \forall h \in (0, h_2).$$

一方で,

$$A_h(K^*h\boldsymbol{z}_h) = K^*h(\boldsymbol{\vartheta}_h + \boldsymbol{\tau}_h), \qquad A_h\boldsymbol{\eta}_h = \boldsymbol{\varepsilon}_h$$

が成り立つので、 A_h の単調性より、

$$|\eta_h(P_{ij})| \le K^* h|z_h(P_{ij})| \le K^* \max_{\overline{\Omega}} |z_h|h.$$

よって,

$$\|\boldsymbol{\eta}_h\|_{\infty} = \max_{P_{ij} \in \tilde{\Omega}_h} |\eta_h(P_{ij})| \le Mh \qquad \forall h \in (0, h_2)$$

が示された.

 $\overline{\Gamma}_D = \partial\Omega$ のとき、(1.12) の式が連立方程式に現れない。上の証明を振り返ると、Mh が Mh^2 に改善されることがわかる。

定義 1.5 定数 $M, h_2 > 0$ が存在し,

$$\|\boldsymbol{\eta}_h\|_{\infty} \leq Mh^p \qquad \forall h \in (0, h_2)$$

が成り立つとき、数値スキームが(最大値ノルムで) オーダーp であると言う.

系 1.6 (1.10)-(1.12) のスキームはオーダー1で、最大値ノルムで収束する:

$$\lim_{h\to 0+} \|\boldsymbol{\eta}_h\|_{\infty} = 0.$$

ただし、 $\overline{\Gamma}_D = \partial \Omega$ のときはオーダー 2 となる.

注.

• (1.11) で "離散化の犯罪" を犯したが、Dirichlet 条件に救われた:

$$(A_h \boldsymbol{z}_h)|_{A=P_{ij}} = \frac{1}{4}(\lambda_B + \lambda_C)(\lambda_D + \lambda_E) + \frac{1}{\lambda_X h^2} + \tau_h(P_{ij}).$$

 $\frac{1}{\lambda_X h^2}$ の項のおかげで、 $\tau_h(P_{ij})=O(1)$ でも、結局 $\vartheta_h(A)+\tau_h(A)\geq \frac{1}{2h^2}$ を得ることができた。

• この解析は $u,z\in C^4(\overline{\Omega})$ を仮定している。 $u,z\in C^3(\overline{\Omega})$ のみのとき,正則な節点ではテーラー展開より $\varepsilon_h(P_{ij})=O(h)$ 得て, $\|\boldsymbol{\eta}_h\|_{\infty}=O(h)$ となり,スキームはオーダー 1 となる。さらに, $u,z\in C^2(\overline{\Omega})$ のとき,スキームはオーダー 0 となるが,収束することが証明できる.

「注」の終わり _

次に考案したスキームが最大値原理の離散版を満たすことを確かめる。簡単のために、ラプラス作用素に対する Dirichlet 問題を考える。

定理 1.7 $\Gamma_D = \partial \Omega$ とした問題を (1.11) に従って離散化する。領域 Ω において $f \leq 0$ ならば, Ω のすべての節点 P_{ij} での近似解の値は境界値 u_D の最大値より大きくならない。同様に, Ω において $f \geq 0$ ならば, Ω のすべての節点 P_{ij} での近似解の値は境界値 u_D の最小値より小さくならない。

Proof. Ω において f ≤ 0 のとき,

$$\left(\frac{1}{\lambda_B} + \frac{1}{\lambda_C} + \frac{1}{\lambda_D} + \frac{1}{\lambda_E}\right) U_A \le \frac{U_B}{\lambda_B} + \frac{U_C}{\lambda_C} + \frac{U_D}{\lambda_D} + \frac{U_E}{\lambda_E}.$$
(1.13)

すべての節点における近似値の最大値が節点 A で達成され、境界上の最大値 $\max u_D$ より大きいと仮定する。節点 B,C,D,E がすべての Ω 内部の節点のとき(すなわち、A が正則な節点のとき)、上の式は

$$U_A \le \frac{1}{4} (U_B + U_C + U_D + U_E)$$

となり、 U_A が極値なら $U_B = U_C = U_D = U_E = U_A$ を得る。よって、A と連結しているすべての節点 P_{ij} での値が U_A と一致し、 U_A の値が達成される正則でない節点(または境界上の節点) A' が存在する。この節点が境界上にあれば、矛盾が導かれた。一方では、A' が Ω 内部の節点のとき、となりにある 4 点 B' 、C' 、D' 、E' における近似値は U_A と一致する(となりの点が領域内部の節点の場合)か、 U_A より小さい(境界上の節点の場合)。境界上の点が一つ以上あるから、矛盾が得られた。 $f \geq 0$ のときの証明は同様である。

一般の2階楕円型方程式について構成したスキームが最大値原理を満たすとは限らない。例えば、方程式が混合微分の項 $2b\partial^2 u/\partial x\partial y$ を含むとき、この項を中心差分で離散化すると最大値原理の離散版は成り立たない。一方で、bが正でa,cと比べて十分小さいとき、前進差分と後退差分の平均を用いて離散化すれば、最大値原理の離散版は成り立つ。

注. 長方形でない領域における差分法について

方程式が成り立つ領域が長方形でないとき、上で説明した方法で差分スキームを構成できるが、 座標変換を行うことで長方形領域に変換することが可能な問題もある。そうしたときに微分方程 式の形が変わるが、それを差分化することによって得られる係数行列が対称な正定値行列でなく なる可能性がある。しかし、差分化のしかたで工夫をこらせば、対称性と正定値性を保つことが できるときがある。

例えば,極座標で書いたポアソン方程式

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = f(r,\theta)$$

を次のように離散化できる:

$$\frac{1}{r_i} \left(r_{i+1/2} \frac{u_{i+1,j} - u_{ij}}{\Delta r} - r_{i-1/2} \frac{u_{ij} - u_{i-1,j}}{\Delta r} \right) \frac{1}{\Delta r} + \frac{1}{r_i^2} \frac{u_{i,j+1} - 2u_{ij} + u_{i,j-1}}{\Delta \theta^2} = f_{ij}.$$

このとき、原点における条件を与える必要があるが、半径 $\Delta r/2$ の円盤で微分方程式を積分することにより得られる式を離散化することによりそのような条件を得ることができる。

問題

問題 1.1 格子のx-方向の離散幅がy-方向の離散幅が異なるときも最大値原理の離散版が成り立っことを示せ.

問題 1.2 9 点スキームに対して、離散版の最大値原理を述べて、証明せよ.

問題 1.3 ポアソン方程式を領域 $\Omega = \{(x,y);\ 0 < x < 1,\ 0 < y < H(x)\}$ で考える。新しい座標 $\xi = x, \eta = y/H(x)$ で微分方程式を書き直し、正定値で対称な行列をもつ差分スキームを構成せよ。

問題 1.4 極座標で書かれたポアソン方程式を領域 $\Omega=\{(r,\theta);\ 0< r< s(\theta),\ 0< \theta< 2\pi\}$ で考える。新しい座標 $\rho=r/s(\theta),\phi=\theta$ で微分方程式を書き直し,正定値で対称な行列をもつ差分スキームを構成せよ。

2 付録:行列解析の基礎

正則行列 $A \in \mathbb{R}^{n \times n}$, 右辺ベクトル $b \in \mathbb{R}^n$ と未知ベクトル $x \in \mathbb{R}^n$ に対する連立線形方程式

$$Ax = b (2.1)$$

を解く問題を考える.

解法は基本的に2種類ある:

- 直接法
- 反復法

消去法に代表される直接法では、メモリーや誤差の関係で扱える行列の大きさに限界がある。そのため、反復法が優先されることが多い。反復法の基本的な考え方は、(2.1)を

$$x = Hx + c$$

のように書き換えることである。ただし、 反復行列 とよばれる H は $n \times n$ 行列, $\mathbf{c} \in \mathbb{R}^n$ である。 反復スキームは任意の $\mathbf{x}^0 \in \mathbb{R}^n$ から始め,次の計算を繰り返す:

$$\boldsymbol{x}^{k+1} = H\boldsymbol{x}^k + \boldsymbol{c}, \qquad k \ge 0.$$

これによりベクトルの列 $\{x^k\}_{k=0}^{\infty}$ が得られるが, $k \to \infty$ のとき x^k が (2.1) の解 x に収束することを要求する.この収束が成り立つための必要十分条件は,H のスペクトル半径 $\rho(H)$ が 1 より小さいことである.そのため,まずは与えられた行列のスペクトル半径の評価のしかたを紹介する.

例 2.1 最もシンプルな反復法の一つはヤコビ法である。行列 A を対角行列 D と対角がゼロの行列 R にわけて

$$A = D + R = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} + \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ a_{21} & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & 0 \end{pmatrix},$$

連立方程式を $x = D^{-1}(b - Rx)$ と書き直す、そこで、反復行列 H を

$$H = -D^{-1}R$$

のように定義すれば、 $x^{k+1} = Hx^k + D^{-1}b$ という反復法が得られる.

2.1 行列のスペクトル半径の評価

定義 2.2

• \mathbb{C} の要素をもつ $n \times n$ 行列 A の <mark>固有値</mark> は次の条件を満たす $\lambda \in \mathbb{C}$ である: $\mathbf{x} \in \mathbb{C}^n$ が存在し,

$$Ax = \lambda x$$

が成り立つ. ベクトル なを 固有ベクトル とよぶ.

- Aの スペクトル Sp(A) は Aのすべての固有値を集めた集合である。
- *A* の スペクトル半径 は次の式で定義される:

$$\rho(A) = \max\{|\lambda|; \ \lambda \in Sp(A)\}.$$

• 行列 A の ノルム は A が定義する線形写像 $\mathbb{R}^n \to \mathbb{R}^n$ のノルムである。つまり、 $\|\cdot\|$ が \mathbb{R}^n におけるノルムであれば、このノルムが 誘導 する A のノルムは

$$||A|| = \sup_{\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x} \neq \boldsymbol{0}} \frac{||A\boldsymbol{x}||}{||\boldsymbol{x}||}.$$

「定義」の終わり

||A|| が \mathbb{R}^n のノルム $||\cdot||$ に誘導されるノルムであれば,上の定義より

$$||Ax|| \le ||A|| \cdot ||x|| \qquad \forall x \in \mathbb{R}^n \tag{2.2}$$

が成り立つ.

例 2.3

- (1) ユークリッドノルム $\|x\|_E := (\sum_{i=1}^n |x_i|^2)^{1/2}$ が誘導する行列ノルムを $\|A\|_E$ とする.この ノルムを与える公式が存在しない.
- (2) ノルム $\|x\|_1 := \sum_{i=1}^n |x_i|$ が誘導する行列ノルムは

$$||A||_1 = \max_{j=1,\dots,n} \sum_{i=1}^n |a_{ij}|.$$

(3) 一方で、ノルム $\|\boldsymbol{x}\|_{\infty}\coloneqq\max_{i=1,\dots,n}|x_i|$ が誘導する行列ノルムは

$$||A||_{\infty} = \max_{i=1,\dots,n} \sum_{j=1}^{n} |a_{ij}|.$$

(4) 行列 A を \mathbb{R}^{n^2} のベクトルとして見なし,

$$||A|| := \left(\sum_{i,j=1,\dots,n} |a_{ij}|^2\right)^{1/2}$$

というノルムを考えることができる. しかし, 誘導ノルムではないので, (2.2) のような不 等式は成立しない.

「例」の終わり _

定理 **2.4** $\|\cdot\|$ が \mathbb{R}^n のノルム $\|\cdot\|$ に誘導された行列ノルムであるならば,

$$\rho(A) \le ||A|| \qquad \forall A \in \mathbb{R}^{n \times n}.$$

Proof. 固有値 $\lambda \in Sp(A)$ を任意にとり、対応する固有ベクトルを x とすると、

$$|\lambda| \|x\| = \|Ax\| \le \|A\| \|x\|.$$

したがって、すべての A の固有値の絶対値が $\|A\|$ 以下であり、定理が証明された.

定義 2.5 $n \times n$ 行列 A が 優対角 であるとは,

$$\Lambda_i := \sum_{j \neq i} |a_{ij}| \le |a_{ii}| \qquad \forall i = 1, \dots, n$$

が成り立つということである。上の式がくで成り立つ場合、行列は 狭義優対角 であるという.

系 2.6 行列 A が狭義優対角ならば、ヤコビ法は収束する.

定理 **2.7** (Gershgorin) 集合 Sp(A) は次の n 個の円盤の和集合の部分集合である:

$$Sp(A) \subset D_A := \bigcup_{i=1}^n \{ \xi \in \mathbb{C}; |\xi - a_{ii}| \le \Lambda_i \}.$$

Proof. $\lambda \in Sp(A)$ を任意に選ぶ. 対応する固有ベクトル $x \neq 0$ が

$$|x_i| \le |x_r| = 1, \qquad j = 1, \dots, n$$

を満たすようにとることができる。 $Am{x}=\lambdam{x}$ なので、 $\sum_{j=1}^n a_{rj}x_j=\lambda x_r$ より

$$\sum_{j \neq r} a_{rj} x_j = (\lambda - a_{rr}) x_r.$$

両辺の絶対値をとり、評価すれば、

$$|\lambda - a_{rr}| \cdot |x_r| \le \sum_{j \ne r} |a_{rj}| \cdot |x_j| \le \sum_{j \ne r} |a_{rj}| = \Lambda_r.$$

定義 2.8 $n \times n$ 行列 A が 可約 であるとは,

$$a_{\rho_{\nu}\sigma_{\mu}} = 0$$
 $\forall \nu = 1, \dots, r \quad \forall \mu = 1, \dots, n - r$

が成立するような置換

$$\begin{pmatrix} 1 & \dots & r & r+1 & \dots & n \\ \rho_1 & \dots & \rho_r & \sigma_1 & \dots & \sigma_{n-r} \end{pmatrix}$$
 (2.3)

が存在するという意味である.

Aが 既約 であるとは、可約ではないということである。

「定義」の終わり _____

上記の置換をAの行と列に適用すると,

$$\tilde{A} := \begin{pmatrix} \tilde{A}_{11} & O \\ \tilde{A}_{21} & \tilde{A}_{22} \end{pmatrix}$$

という形の行列が得られる。ここで,O は $r \times (n-r)$ の零行列, \tilde{A}_{11} は $r \times r$ 行列, \tilde{A}_{21} は $(n-r) \times r$ 行列, \tilde{A}_{22} は $(n-r) \times (n-r)$ 行列を意味する。よって, $A\mathbf{x} = \mathbf{b}$ という連立方程式を二つ以上の独立な連立方程式 $\tilde{A}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$ として書き換えることができる。

定義 **2.9** 頂点の集合 $M = \{P_1, \dots, P_n\}$ に対しグラフ $G \subset M \times M$ を考える. G の元 (P_i, P_j) を G の辺とよぶ. グラフ上の任意の 2 頂点間に道が存在するグラフのことを 連結 なグラフと言う. ここで、道は、各頂点とその次の頂点との間に辺が存在するような頂点の列である.

行列 A が与えられたとき、 $M = \{P_1, \dots, P_n\}$ 上のグラフ G_A を次で定義する:

$$(P_i, P_j) \in G_A \quad \Leftrightarrow \quad a_{ij} \neq 0.$$

注. 一般の行列 A では、有向グラフを考える必要がある。すなわち、 $a_{ij} \neq 0$ のとき P_i から P_j に向かう辺があると考える。一方で、行列 A が対称行列のとき、 $a_{ij} \neq 0$ と $a_{ji} \neq 0$ は同値であるから、 P_i と P_j を結ぶ辺は必ず両方向に存在し、向きを考える必要がない。以下の議論は、辺の向きについて特記しないが、行列の対称性に関係なく通用する。

「定義」の終わり _____

定理 **2.10** $n \times n$ 行列 A $(n \ge 2)$ が既約であることとグラフ G_A が連結であることとが同値である.

Proof. A が可約であるとすれば,(2.3) のような置換が存在する.そのとき,頂点 $P_{\rho_1},\ldots,P_{\rho_r}$ の どの頂点をとっても,頂点 $P_{\sigma_1},\ldots,P_{\sigma_{n-r}}$ のどれもとつなげる道が存在しない.もしそのような 道が存在すれば, $a_{\rho_{\nu}\sigma_{\mu}}\neq 0$ となるような ν,μ があり,矛盾が得られる.よって,グラフは連結で はない.

一方で、 G_A が連結でないと仮定すると、道が存在しないような頂点 P_i, P_j が存在する。グラフ G_A の頂点 P_1, \ldots, P_n の番号をふり直すことで、次の頂点の集合を定義できる:

$$B := \{P_{\alpha} \in M; P_i \geq P_{\alpha} \text{ の間に道が存在する}\} \cup \{P_i\} = \{P_{\rho_1}, \dots, P_{\rho_r}\}$$
 $C := M \setminus B = \{P_{\sigma_1}, \dots, P_{\sigma_{n-r}}\}.$

頂点 P_i が集合 C に属するので、C は空集合ではない。これにより置換 (2.3) が決まるが、

$$a_{\rho_{\nu}\sigma_{\mu}} = 0$$
 $\forall \nu = 1, \dots, r, \quad \forall \mu = 1, \dots, n-r$

であることを確かめる。上の性質が成立しない μ,ν が存在するならば, $(P_{\rho_{\nu}},P_{\sigma_{\mu}})\in G_A$ となるので, $P_{\sigma_{\mu}}\in B$ となり,矛盾が得られる.

次に Gershgorin 定理の精密化を紹介する.

定理 2.11 A が既約で,その固有値 $\lambda \in Sp(A)$ が D_A の境界上にあれば($\lambda \in \partial D_A$),次が成り立つ:

$$|\lambda - a_{ii}| = \Lambda_i \quad \forall i = 1, \dots, n.$$

Proof. $\lambda \in Sp(A)$ なら、 $x \in \mathbb{C}^n$ $(x \neq 0)$ が存在し、 $Ax = \lambda x$ が成り立つ.

$$|x_i| \le 1 = |x_r| \qquad \forall i = 1, \dots, n \tag{2.4}$$

となるようにできる.

$$(\lambda - a_{rr})x_r = \sum_{j \neq r} a_{rj}x_j$$

より

$$|\lambda - a_{rr}| \le \sum_{j \ne r} |a_{rj}| |x_j| \le \sum_{j \ne r} |a_{rj}| = \Lambda_r.$$

 $\lambda \in \partial D_A$ より、 $|\lambda - a_{rr}| = \Lambda_r$ 、したがって、

$$|\lambda - a_{rr}| = \sum_{j \neq r} |a_{rj}| |x_j| = \sum_{j \neq r} |a_{rj}|.$$

もし $a_{rj} \neq 0$ であれば, $|x_j| = 1$ でなければならない.

任意の $i \in \{1,\ldots,n\}$ をとり、 $|\lambda - a_{ii}| = \Lambda_i$ が成り立つことを示す。 G_A が連結なので、 P_i と P_r の間に道がある。よって、 $r = \alpha_0, \alpha_1, \ldots, \alpha_s = i$ が存在し、

$$(P_{\alpha_j}, P_{\alpha_{j+1}}) \in G_A \qquad \forall j = 0, \dots, s-1.$$

つまり、 $\nu=0,\ldots,s-1$ について $a_{\alpha_{\nu}\alpha_{\nu+1}}\neq 0$ が成り立つ。 $a_{\alpha_{0}\alpha_{1}}=a_{r\alpha_{1}}\neq 0$ なので、 $|x_{\alpha_{1}}|=1$ と結論できる.

いま,(2.4) において r を α_1 に変えても (2.4) は成り立つ。 $|\lambda - a_{\alpha_1\alpha_1}| = \Lambda_{\alpha_1}$ なので,上と同じような議論により, $|x_{\alpha_2}| = 1$ を得る.これを繰り返すことで,

$$|\lambda - a_{\alpha_{\nu}\alpha_{\nu}}| = \Lambda_{\alpha_{\nu}} \qquad \nu = 0, \dots, s$$

が従う。とくに、 $\nu = s$ とおいて、i の任意性を考慮すれば、定理の主張が示された。

定義 **2.12** 行列 A は 既約優対角行列 (IDDM) であるとは、A が既約で次の 2 条件が成り立つという意味である:

- $\sum_{j \neq i} |a_{ij}| \le |a_{ii}| \ \forall i = 1, \dots, n$
- $\sum_{j\neq i_0}|a_{i_0j}|<|a_{i_0i_0}|$ となる $i_0\in\{1,\ldots,n\}$ が存在する

定理 2.13 行列 A は狭義優対角または既約優対角行列であれば、次の性質をもつ:

- (1) A は正則行列である
- (2) $a_{ii} > 0 \ \forall i = 1, \dots, n \$ \$\text{\$\text{\$\text{\$t\$}\$}\$, } \\ \text{Re}(\lambda) > 0 \ \forall \lambda \in Sp(A)

Proof. (1) A は正則行列であるとは、 $0 \not\in Sp(A)$ ということであるので、 $0 \in Sp(A)$ と仮定して矛盾を導く。 $|x_i| \leq |x_r| = 1$ かつ $A \boldsymbol{x} = \boldsymbol{0}$ を満たす固有ベクトル \boldsymbol{x} をとる。 $-a_{rr}x_r = \sum_{j \neq r} a_{rj}x_j$ より、

$$|a_{rr}| \le \sum_{j \ne r} |a_{rj}| = \Lambda_r$$

が成り立つ. A が狭義優対角ならば、これは矛盾である。既約優対角行列の場合、 $r=i_0$ ならば、矛盾である(i_0 は上記定義と同じ意味をもつ)。 $r\neq i_0$ ならば、A の既約性より $a_{\alpha_{\nu}\alpha_{\nu+1}}\neq 0 \ \forall \nu=0,\ldots,s-1$ が成立するような $r=\alpha_0\neq\alpha_1\neq\cdots\neq\alpha_s=i_0$ が存在する。r 行目については狭義優対角性が満たされないから、 $\Lambda_r=|a_{rr}|$.

$$\sum_{j \neq r} |a_{rj}| |x_j| = \sum_{j \neq r} |a_{rj}|$$

より $|x_{\alpha_1}|=1$ を得る。 α_1 行目で狭義優対角性が成り立つならば,矛盾が得られる。そうでなければ,最初に戻って,r を α_1 に置き換えて,議論を繰り返す。途中で狭義優対角性が成り立つ行に出会わなければ, i_0 にたどりついて矛盾を得る。

(2) $a_{ii} > 0 \ \forall i = 1, ..., n$ ならば、 $|\lambda - a_{rr}| \le \Lambda_r \le |a_{rr}|$ を満たすような $r \in \{1, ..., n\}$ が存在する.

定理の主張を背理法で示す。 $\operatorname{Re}(\lambda)=0$ ならば,A の正則性より $\operatorname{Im}(\lambda)\neq 0$ で,

$$|a_{rr}| \ge |\lambda - a_{rr}| = \sqrt{a_{rr}^2 + (\operatorname{Im} \lambda)^2} > |a_{rr}|$$

より矛盾が従う、 $Re(\lambda) < 0$ ならば、

$$|a_{rr}| \ge |\lambda - a_{rr}| = \sqrt{(\operatorname{Re} \lambda - a_{rr})^2 + (\operatorname{Im} \lambda)^2} > |a_{rr}|$$

より矛盾が従う.

定義 2.14 A を $n \times n$ 行列とする.

- $A^T=A$ が成り立つとき、A は 対称行列 であるという(ここで、 $A=(a_{ij})_{i,j=1}^n$ のとき、 $A^T:=(a_{ji})_{i,i=1}^n$)
- $A^*=A$ が成り立つとき、A は エルミート行列 であるという(ここで、 $A=(a_{ij})_{i,j=1}^n$ のとき、 $A^*:=(\overline{a}_{ji})_{j,i=1}^n$)

定義 2.15 エルミート行列 Aが

$$\mathbf{x}^* A \mathbf{x} > 0 \qquad \forall \mathbf{x} \in \mathbb{C}^n, \ \mathbf{x} \neq \mathbf{0}$$

を満たすとき、A は正定値行列であるという。上の式で \geq のみ言えるときはA は半正定値であるという。ただし、 $\mathbf{x}=(x_1,\ldots,x_n)^T, \mathbf{x}^*=(\overline{x}_1,\ldots,\overline{x}_n)$.

_ 問題

問題 2.1 次を示せ:

- (1) A はエルミート行列 \Rightarrow $Sp(A) \subset \mathbb{R}$
- (2) A は正定値行列 \Leftrightarrow $Sp(A) \subset (0, +\infty)$
- (3) A は半正定値行列 \Rightarrow $Sp(A) \subset [0, +\infty)$

定理 **2.16** A がエルミート行列ならば、 \mathbb{C}^n の直交基底を成すような A の固有ベクトルが存在する.

定理2.13より

系 2.17 エルミート行列 A が狭義優対角または既約優対角で, $a_{ii}>0 \ \forall i=1,\ldots,n$ ならば,A は正定値行列である.

定理 2.18 ユークリッドノルムに誘導される行列ノルムは次を満たす:

$$||A||_E = \sqrt{\rho(A^*A)}.$$

Proof. 行列 *A*A* の性質をまとめる:

- (1) エルミート行列である
- (2) 半正定値行列である. よって、その固有値は $0 \le \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n = \rho(A^*A)$ を満たす.
- (3) $A^*A\alpha_i = \lambda_i\alpha_i$ を満たすような \mathbb{C}^n の直交基底 $\alpha_1, \ldots, \alpha_n \in \mathbb{C}^n$ が存在する.

誘導ノルムの定義より

$$||A||_E = \sup_{\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n} \frac{||A\mathbf{x}||_E}{||\mathbf{x}||_E}$$

であるが, ここで,

$$\|x\|_E = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2} = (x^*x)^{1/2}$$

 $\|Ax\|_E^2 = (Ax)^*(Ax) = x^*A^*Ax.$

さらに、 $m{x} = \sum_{i=1}^n c_i m{lpha}_i$ と書けるような c_i が存在し、そのとき $m{x}^*m{x} = \sum_{i=1}^n |c_i|^2$ が成り立つ.

 $\mathbf{0} \neq \mathbf{x} \in \mathbb{C}^n$ に対し次のように計算できる:

$$\begin{split} \frac{\|A\boldsymbol{x}\|_{E}^{2}}{\|\boldsymbol{x}\|_{E}^{2}} &= \frac{\boldsymbol{x}^{*}A^{*}A\boldsymbol{x}}{\boldsymbol{x}^{*}\boldsymbol{x}} \\ &= \frac{\left(\sum_{j=1}^{n}c_{j}\boldsymbol{\alpha}_{j}\right)^{*}(A^{*}A\sum_{i=1}^{n}c_{i}\boldsymbol{\alpha}_{i})}{\sum_{j=1}^{n}|c_{j}|^{2}} \\ &= \frac{\left(\sum_{j=1}^{n}\bar{c}_{j}\boldsymbol{\alpha}_{j}^{*}\right)\left(\sum_{i=1}^{n}c_{i}(A^{*}A\boldsymbol{\alpha}_{i})\right)}{\sum_{j=1}^{n}|c_{j}|^{2}} \\ &= \frac{\left(\sum_{j=1}^{n}\bar{c}_{j}\boldsymbol{\alpha}_{j}^{*}\right)\left(\sum_{i=1}^{n}\lambda_{i}c_{i}\boldsymbol{\alpha}_{i}\right)}{\sum_{j=1}^{n}|c_{j}|^{2}} \\ &= \frac{\sum_{i=1}^{n}\lambda_{i}|c_{i}|^{2}}{\sum_{j=1}^{n}|c_{j}|^{2}} \\ &\leq \lambda_{n} = \rho(A^{*}A) \end{split}$$

両辺の $x \in \mathbb{R}^n$ についての sup をとることで、 $||A||_E \leq \sqrt{\rho(A^*A)}$ を得る.

逆の不等式は $x = \alpha_n$ とすることによって示される。実際,

$$||A||_E \ge \frac{||A\alpha_n||_E^2}{||\alpha_n||_E^2} = \lambda_n = \rho(A^*A)$$

が成り立つから、 $||A||_E$ が $\rho(A^*A)$ 以上であることがわかる.

 $\rho(A^2) = (\rho(A))^2$ に気づけば、次の結果を得る:

系 2.19 エルミート行列 A については $||A||_E = \rho(A)$ が成り立つ.

2.2 連立方程式を解くための反復法

正則行列 $A \in \mathbb{R}^{n \times n}$, 右辺ベクトル $\boldsymbol{b} \in \mathbb{R}^n$ と未知ベクトル $\boldsymbol{x} \in \mathbb{R}^n$ に対する連立線形方程式

$$Ax = b (2.5)$$

を解くのに利用できる基本的な反復法を紹介する。反復法の一般的な形は

$$\boldsymbol{x}^{k+1} = H\boldsymbol{x}^k + \boldsymbol{c} \tag{2.6}$$

と書ける。このとき、Hを反復行列とよぶ。

最初に扱う反復法は行列 A の分解に基づいているので、その記号を導入する。A を対角行列

D, 下三角行列 -E と上三角行列 -F に分ける。つまり、

$$A = D - E - F = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} - \begin{pmatrix} 0 & 0 & \dots & 0 \\ a_{21} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & 0 \end{pmatrix} - \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ 0 & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

2.2.1 Jacobi 法

ヤコビ法は

$$D\boldsymbol{x}^{k+1} = \boldsymbol{b} + (E+F)\boldsymbol{x}^k$$

という反復法である. よって、(2.6) において、 $H = D^{-1}(E+F)$ 、 $c = D^{-1}b$ となる. 成分ごとで書くと、

$$a_{ii}x_i^{k+1} = b_i - \sum_{j < i} a_{ij}x_j^k - \sum_{j > i} a_{ij}x_j^k.$$
(2.7)

定理 2.20 行列 A が狭義優対角行列または既約優対角行列ならば、Jacobi 法は収束する.

Proof. 定理2.4と優対角性より

$$\rho(H) \le ||H||_{\infty} = ||D^{-1}(E+F)||_{\infty} = \max_{i=1,\dots,n} \frac{\sum_{j \ne i} |a_{ij}|}{|a_{ii}|} = \max_{i=1,\dots,n} \frac{\Lambda_i}{|a_{ii}|} \le 1.$$

A は狭義優対角行列のとき, $\Lambda_i < |a_{ii}|$ がすべての $i=1,\ldots,n$ について成り立つので, $\rho(H) < 1$ が示せた.

一方、A は既約優対角行列のとき、 $|\lambda|=1$ となる $\lambda\in Sp(H)$ があれば、 λ が Gershgorin の円盤の和集合の境界に位置することになるので、定理2.11より λ はすべての円盤($i=1,\ldots,n$)の境界に位置することになる。これは $\Lambda_i=|a_{ii}|,\ i=1,\ldots,n$ という意味になり、既約優対角性との矛盾である.

2.2.2 Gauss-Seidel 法

ガウスーザイデル法はヤコビ法 (2.7) の右辺の最初の和において $x_j, j < i,$ の (k+1)-近似がすでに求まっていることを利用して、ヤコビ法を次のように改良したものである:

$$a_{ii}x_i^{k+1} = b_i - \sum_{j \le i} a_{ij}x_j^{k+1} + \sum_{j \ge i} a_{ij}x_j^k.$$
(2.8)

これを行列を用いて書くと, $Dx^{k+1}=b+Ex^{k+1}+Fx^k$ となるので,この場合の反復行列は $H=(D-E)^{-1}F$,そして $\mathbf{c}=(D-E)^{-1}\mathbf{b}$ である.よって,Gauss-Seidel 法は

$$\mathbf{x}^{k+1} = (D-E)^{-1}F\mathbf{x}^k + (D-E)^{-1}\mathbf{b}.$$

定理 2.21 行列 A が狭義優対角行列または既約優対角行列ならば, Gauss-Seidel 法は収束する.

Proof. $H = (D - E)^{-1}F$ の固有値 λ に対し, $1 = |x_r| \ge |x_i|, i = 1, ..., n$, を満たす固有ベクトル x をとる.このとき,

$$-\sum_{j=i+1}^{n} a_{ij} x_j = \lambda \sum_{j=1}^{i} a_{ij} x_j, \qquad 1 = |x_r| \ge |x_i|, \qquad i = 1, \dots, n$$

が成り立つので、これをi=rについて用いると、次のように評価できる:

$$|\lambda| \left(|a_{rr}| - \sum_{j=1}^{r-1} |a_{rj}| \right) \leq |\lambda| \left(|a_{rr}x_r| - \sum_{j=1}^{r-1} |a_{rj}x_j| \right)$$

$$\leq |\lambda| \left(|a_{rr}x_r| - \left| \sum_{j=1}^{r-1} a_{rj}x_j \right| \right)$$

$$\leq |\lambda| \left| \sum_{j=1}^{r} a_{rj}x_j \right|$$

$$= \left| \sum_{j=r+1}^{n} a_{rj}x_j \right|$$

$$\leq \sum_{j=r+1}^{n} |a_{rj}|$$

$$\leq \sum_{j=r+1}^{r-1} |a_{rj}|$$

$$= \Lambda_r - \sum_{j=1}^{r-1} |a_{rj}|.$$

さて、A が狭義優対角のとき、 $|a_{rr}| > \Lambda_r$ が成立するので、

$$|\lambda| \le \frac{\Lambda_r - \sum_{j=1}^{r-1} |a_{rj}|}{|a_{rr}| - \sum_{j=1}^{r-1} |a_{rj}|} < 1$$

が従う.

一方,A が既約優対角のとき, $|a_{i_0i_0}|>\Lambda_{i_0}$ を満たす i_0 が存在する. $r=i_0$ ならば,証明は終わるが,そうでないとき, $|a_{rr}|=\Lambda_r$ であるので, $\lambda\geq 1$ と仮定すれば,

$$|a_{rr}| - \sum_{j=1}^{r-1} |a_{rj}x_j| \le \left| \sum_{j=1}^r a_{rj}x_j \right| \le \left| \sum_{j=r+1}^n a_{rj}x_j \right| \le \sum_{j=r+1}^n |a_{rj}| |x_j|.$$

よって,

$$|a_{rr}| \le \sum_{j \ne r} |a_{rj}| |x_j| \le \sum_{j \ne r} |a_{rj}| \le \Lambda_r \le |a_{rr}|$$

より $a_{rj} \neq 0$ ならば $|x_j| = 1$ が言える.この先の議論は定理2.11の証明と同様であるので,省略する.

2.2.3 **SOR** 法

Gauss-Seidel 法による近似を \tilde{x}_i^{k+1} とおくと, すなわち,

$$a_{ii}\tilde{x}_{i}^{k+1} = b_{i} - \sum_{j < i} a_{ij}x_{j}^{k+1} + \sum_{j > i} a_{ij}x_{j}^{k}$$

とすると、SOR 法はパラメータ $\omega > 0$ をとり、

$$x_i^{k+1} = \omega \tilde{x}_i^{k+1} + (1 - \omega) x_i^k$$

とする反復法である (Gauss-Seidel 法の解 \tilde{x}^{k+1} と前ステップの解 x^k の????).

行列 $L = D^{-1}E$ と $U = D^{-1}F$ を定義すると、SOR 法は次のように表現される:

$$\mathbf{x}^{k+1} = (I - \omega L)^{-1} [(1 - \omega)I + \omega U] \mathbf{x}^k + \omega (I - \omega L)^{-1} D^{-1} \mathbf{b}.$$

ただし、 I は単位行列である.

定理 2.22 A が正定値のエルミート行列で, $a_{ii} > 0 \, \forall i, \, \omega \in (0,2)$ ならば,SOR 法は収束する.

Proof.

$$\mathcal{L}_{\omega} := (I - \omega L)^{-1} \left[(1 - \omega)I + \omega U \right]$$

とおいて、 $\lambda \in Sp(\mathcal{L}_{\omega})$ とする. すると、固有ベクトル $x \in \mathbb{R}^n$ が存在し、

$$\lambda(D - \omega E)\boldsymbol{x} = [(1 - \omega)D + \omega F]\,\boldsymbol{x}$$

が成り立つ. これを整理すれば,

$$\lambda \left[(2-\omega)D + \omega A + \omega (F-E) \right] \boldsymbol{x} = \left[(2-\omega)D - \omega A + \omega (F-E) \right] \boldsymbol{x}.$$

上式に左から x^* をかける。そこで現れる量は次を満たす:

$$d := \boldsymbol{x}^* D \boldsymbol{x} > 0$$
, $a := \boldsymbol{x}^* A \boldsymbol{x} > 0$, $c := \boldsymbol{x}^* (F - E) \boldsymbol{x}$, $\operatorname{Re} c = 0$.

そこで,

$$\lambda(2-\omega)d + \lambda\omega a + \lambda\omega c = (2-\omega)d - \omega a + \omega c.$$

よって,

$$|\lambda| = \left| \frac{(2-\omega)d - \omega a + \omega c}{(2-\omega)d + \omega a + \omega c} \right| = \frac{\sqrt{[(2-\omega)d - \omega a]^2 + \omega^2 |c|^2}}{\sqrt{[(2-\omega)d + \omega a]^2 + \omega^2 |c|^2}} < 1,$$

au au au au au au au

行列 A がある条件を満たせば, \mathcal{L}_{ω} のスペクトル半径が最小になるような ω_{opt} を求めることができる:

$$\omega_{\rm opt} = \frac{2}{1 + \sqrt{1 - \rho^2(B)}}.$$

ただし, $B = D^{-1}(E + F)$ はヤコビ法の反復行列である.

行列 A が上記のような条件を満たしていれば,

$$\rho(\mathcal{L}_1) = \rho^2(B), \qquad \rho(\mathcal{L}_{\omega \text{ opt}}) < \rho(\mathcal{L}_{\omega}) \ \forall \omega \neq \omega_{\text{opt}}$$

が成り立つので、ヤコビ法(反復行列 B)、Gauss-Seidel 法(反復行列 \mathcal{L}_1)、最適な SOR 法(反復行列 $\mathcal{L}_{\omega \text{opt}}$)の順番に収束のスピードがあがることがわかる。以上の反復法に関する詳しいことは、[2,4,6,7] などを参照されたい。

注.

- 最適な ω を正確に決めることが難しいが、どちらかといえば、大きくとった方がよい。これは $\omega_{\rm opt}$ の計算過程でわかることである。
- 実際の計算では $\rho(B)$ の計算がコストが高いので, $\rho(B)$ を近似的に求める次のようなアルゴリズムを用いることが多い.このアルゴリズムは $\rho^2(B)=\rho(\mathcal{L}_1)$ という事実と,power method という固有値を近似的に求める方法に基づいている(つまり, \mathbf{g}_0 を適当に決めて, $\mathbf{g}_{k+1}=\mathcal{L}_1\mathbf{g}_k$ で $\{\mathbf{g}_k\}$ の列を計算したとき, $\rho(\mathcal{L}_1)=\lim_{n\to\infty}\|\mathbf{g}_{k+1}\|/\|\mathbf{g}_k\|$ が成り立つことを用いて,最大固有値の近似を計算する。)
 - (1) $\omega = 1$ で SOR 法の計算を始める(すなわち,反復行列 \mathcal{L}_1 を使用).
 - (2) x^k, x^{k+1} に対し、 $g_k = x^{k+1} x^k$ と定義したとき、 $g_{k+1} = \mathcal{L}_1 g_k$ となる。 $\rho(\mathcal{L}_1) = \lim_{k \to \infty} \|g_{k+1}\|/\|g_k\|$ なので、 $\|g_{k+1}\|/\|g_k\|$ が変化しなくなるまでこれを繰り返す。
 - (3) (2) で求めた $\rho(\mathcal{L}_1)$ の近似を利用し、 $\rho(B) = \sqrt{\rho(\mathcal{L}_1)}$ といて、対応する ω_{opt} を計算し、SOR 法の計算をこの ω_{opt} のパラメータを用いて続ける.

適切な ω の使用により、計算を本質的に高速化できる.

• 上に言及した A に対する条件がすべて満たされなくても、SOR 法がよい結果を出すことがある。

2.2.4 最小化法

正定値行列 A に対し,

$$J(\boldsymbol{x}) := \frac{1}{2} \boldsymbol{x}^T A \boldsymbol{x} - \boldsymbol{x}^T \boldsymbol{b}$$

を定義する.

定理 2.23 Ax = b が成り立つのは、 $J(x) = \min_{y \in \mathbb{R}^n} J(y)$ のときとそのときのみである.

Proof. $J(x) = \min_{y \in \mathbb{R}^n} J(y)$ ならば,

$$0 = \frac{\partial J}{\partial x_i}(\boldsymbol{x}) = \frac{\partial}{\partial x_i} \left(\frac{1}{2} \sum_{k,l=1}^n x_k a_{kl} x_l - \sum_{k=1}^n x_k b_k \right) = \sum_{k=1}^n a_{ik} x_k - b_i$$

がすべての $i=1,\ldots,n$ について成立する. したがって、Ax=bが示せた.

一方, Ax = b ならば,

$$J(\boldsymbol{x} + \boldsymbol{v}) - J(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{x} + \boldsymbol{v})^T A(\boldsymbol{x} + \boldsymbol{v}) - (\boldsymbol{x} + \boldsymbol{v})^T \boldsymbol{b} - \frac{1}{2} \boldsymbol{x}^T A \boldsymbol{x} + \boldsymbol{x}^T \boldsymbol{b}$$

$$= \frac{1}{2} \boldsymbol{x}^T A \boldsymbol{v} + \frac{1}{2} \boldsymbol{v}^T A \boldsymbol{x} + \frac{1}{2} \boldsymbol{v}^T A \boldsymbol{v} - \boldsymbol{v}^T \boldsymbol{b}$$

$$= \boldsymbol{v}^T (A \boldsymbol{x} - \boldsymbol{x}) + \frac{1}{2} \boldsymbol{v}^T A \boldsymbol{v}$$

$$= \frac{1}{2} \boldsymbol{v}^T A \boldsymbol{v} > 0$$

が任意のベクトル $v \neq 0$ について正しいので、 $J(x) = \min_{y \in \mathbb{R}^n} J(y)$ が示せた.

最急降下法は関数 J が最も急に下がる方向(つまり、勾配のマイナス方向)を探索方向とする 反復法である。

- 1. 解の初期近似 x^0 を適当に定める.
- 2. x^k まで求まったとして、 x^{k+1} を式

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + t_k \boldsymbol{v}^k$$

で計算する。ただし、 $v_k = -\nabla J(x^k)$ (x^k における J のマイナス勾配)として、 t_k を v の方向に J を制限した 2 次関数の最小値を与える量とする。すなわち、

$$t_k = \operatorname{argmin}_{t>0} J(\boldsymbol{x}^k + t\boldsymbol{v}^k).$$

 $\frac{d}{dt}J(\boldsymbol{x}^k+t\boldsymbol{v}^k)|_{t=t_k}=0$ という条件より t_k を計算すると,

$$t_k = \frac{(\boldsymbol{v}_k)^T (\boldsymbol{b} - A\boldsymbol{x}^k)}{(\boldsymbol{v}^k)^T A \boldsymbol{v}^k}$$

を得る.

3. $J(x^k)$ の値が指定した範囲で変わらなくなるまでステップ 2 を繰り返す.

最急降下法を改良したのは共役勾配法(Conjugate Gradient Method, CG 法)という反復法である。そのアルゴリズムのみ述べる。

- 1. 解の初期近似 x^0 を定め、 $r^0 := b Ax^0$ 、 $v^0 := r^0$ 、k := 0 とする.
- 2. $r^k \cdot r^k \le \varepsilon$ ならば終了する(ここで、 ε は予め決めた tolerance である).
- 3. 次の量を計算する。ただし、 $\mathbf{x} \cdot \mathbf{y} \coloneqq \mathbf{x}^T \mathbf{y}$ という意味で使う。

$$egin{array}{lll} t_k &:=& rac{oldsymbol{r}^k \cdot oldsymbol{r}^k}{oldsymbol{v}^k \cdot A oldsymbol{v}^k}, \ oldsymbol{x}^{k+1} &:=& oldsymbol{x}^k + t_k oldsymbol{v}^k, \ oldsymbol{r}^{k+1} &:=& oldsymbol{r}^{k-1} \cdot oldsymbol{r}^{k+1}, \ oldsymbol{v}^{k+1} &:=& oldsymbol{r}^{k+1} + s_k oldsymbol{v}^k, \end{array}$$

4. k := k+1とおいて,ステップ2に戻る.

コンピュータのメモリーに保存する配列は A, x, v, Av と r である.次の補題でまとめたように,ベクトル v^k は A-直交である($v^k \cdot Av^l = 0, k \neq l$).このことを共役なベクトルであるとも言う.

補題 2.24 CG 法で定義されるベクトルは以下の性質をもつ.

- $(1) \ \mathbf{r}^k = \mathbf{b} A\mathbf{x}^k$
- (2) $\mathbf{x}^j \cdot \mathbf{x}^l = 0, \quad j \neq l$
- (3) $\mathbf{v}^j \cdot A \mathbf{v}^l = 0, \quad j \neq l$

したがって、CG 法は遅くてもn ステップで終了する.

CG 法は最急降下法に比べはやく収束するが、収束をさらにはやくするには前処理(preconditioning)という操作を行う。前処理のアイデアは、行列 A に"近い"行列 Q ($Q^{-1}A$ が単位行列に近い)で、Q を係数行列とした連立方程式が高速で解けるような行列 Q を構成して、CG 法を $Q^{-1}Ax = Q^{-1}b$ という連立方程式に適用することである。 $Q^{-1}A$ が単位行列に近ければ近いほど、CG 法の収束がはやい。前処理つきの CG 法のアルゴリズムの例を挙げる。

- 1. 解の初期近似 x^0 を定め、 $r^0 := b Ax^0$ 、 $z^0 := Q^{-1}r^0$ 、 $v^0 := z^0$ 、k := 0 とする.
- 2. $\mathbf{r}^k \cdot \mathbf{r}^k \leq \varepsilon$ ならば終了する(ここで、 ε は予め決めた tolerance である).
- 3. 次の量を計算する.

$$egin{array}{lll} lpha_k &\coloneqq& rac{oldsymbol{r}^k \cdot oldsymbol{z}^k}{oldsymbol{v}^k \cdot A oldsymbol{v}^k}, \ oldsymbol{x}^{k+1} &\coloneqq& oldsymbol{x}^k + lpha_k oldsymbol{v}^k, \ oldsymbol{z}^{k+1} &\coloneqq& oldsymbol{Q}^{-1} oldsymbol{r}^{k+1}, \ eta_k &\coloneqq& rac{oldsymbol{r}^{k+1} \cdot oldsymbol{z}^{k+1}}{oldsymbol{r}^k \cdot oldsymbol{z}^k}, \ oldsymbol{v}^{k+1} &\coloneqq& oldsymbol{z}^{k+1} + eta_k oldsymbol{v}^k. \end{array}$$

4. k := k + 1とおいて、ステップ2に戻る。

CG 法とその一般化に関する詳細は、[2,3,4,6,7] または [5] などを参照されたい.

2.3 単調行列について

定義 2.25 $n \times r$ 行列 A, B が $A \ge B$ を満たすとは,

$$a_{ij} \ge b_{ij}$$
 $\forall i = 1, \dots, n, \ \forall j = 1, \dots, r$

が成り立つということである。同様に、 $(x_1,\ldots,x_n)^T \geq (y_1,\ldots,y_n)^T$ とは $x_i \geq y_i \ \forall i=1,\ldots,n$ ということである。

定義 2.26 $n \times n$ 行列 A が 単調行列 であるとは

$$oldsymbol{z} \in \mathbb{C}^n, \ Aoldsymbol{z} \geq oldsymbol{0} \quad \Rightarrow \quad oldsymbol{z} \geq oldsymbol{0}$$

が成り立つことである.

定理 **2.27** A が単調行列であるのは, $A^{-1} \ge O$ が成り立つときとそのときのみである.

Proof. $A^{-1} \geq O$ ならば、 $z \in \mathbb{C}^n$ に対して

$$Az \geq \mathbf{0} \quad \Rightarrow \quad z = A^{-1}(Az) \geq \mathbf{0}.$$

よって、Aは単調行列である.

一方で、Aが単調ならば、正則行列である。実際、Az = 0となる $z \in \mathbb{C}^n$ は $Az \leq 0$ と $Az \geq 0$ の両方を満たすので、単調性より $z \leq 0$ と $z \geq 0$ とがわかり、z = 0 が従う。

したがって、 A^{-1} は意味をもち、その列を $\boldsymbol{b}_1,\dots,\boldsymbol{b}_n$ とすれば、 $\boldsymbol{b}_r \geq \boldsymbol{0} \ \forall r$ を示せばよい。しかし、 $A\boldsymbol{b}_r = \boldsymbol{e}_r \geq 0$ と A の単調性より、 $\boldsymbol{b}_r \geq \boldsymbol{0}$ がすぐに従う.

最後に挙げた定理は大きい行列 A が単調であるかどうかを判定するのに向いていないので、別の判定法を紹介する。

定理 2.28 $n \times n$ 行列 A が狭義優対角または既約優対角であるとする。また、A の要素は $a_{ii} > 0 \, \forall i, \, a_{ij} \leq 0 \, \forall i \neq j$ を満たすとする。そのとき、A は単調行列である。

Proof. 任意の $z \in \mathbb{C}^n$ に対して $Az \geq 0$ であれば, $z \geq 0$ が成り立つことを示す.

まず、z が実数ベクトルであることを示す。実際、A が実数行列であるから、 $Az = A(\operatorname{Re} z + i\operatorname{Im} z) \geq \mathbf{0}$ より $A\operatorname{Im} z = \mathbf{0}$ がわかる。定理2.13を思い出すと、 $\operatorname{Im} z = \mathbf{0}$ が言える。

次にz > 0を背理法で示す.

$$z_r = \min_{j=1,\dots,n} z_j < 0 (2.9)$$

と仮定する。 $Az \ge \mathbf{0}$ の r-番目の成分を書くと, $\sum_{j=1}^n a_{rj}z_j \ge 0$ となる。また,定理の仮定より, $j \ne r$ に対し, $a_{rj}z_r \ge a_{rj}z_j$ が正しいから,

$$\sum_{j=1}^{n} a_{rj} z_j \le z_r \sum_{j=1}^{n} a_{rj} < 0$$

が言え、矛盾となる.

最後の<0の不等式は次のように証明できる:

- A が狭義優対角の場合: $\sum_{j\neq r}|a_{rj}|<|a_{rr}|$, すなわち, $-\sum_{j\neq r}a_{rj}< a_{rr}$, 整理すれば $\sum_{j=1}^n a_{rj}>0$
- Aが既約優対角の場合:次のことがわかっている:

$$0 \le \sum_{j=1}^{n} a_{rj} z_j \le z_r \sum_{j=1}^{n} a_{rj},$$

$$\sum_{j=1}^{n} a_{ij} \ge 0 \quad \forall i = 1, \dots, n, \qquad \exists i_0 : \sum_{j=1}^{n} a_{i_0 j} > 0.$$

r-番目の行に対して狭義優対角性が成り立てば、すぐに矛盾が得られる。そうでない場合、 $\sum_{i=1}^n a_{ri}=0$ が成り立つ。よって、

$$0 \le \sum_{j=1}^{n} a_{rj} z_j \le z_r \sum_{j=1}^{n} a_{rj} = 0,$$

すなわち

$$\sum_{j \neq r} a_{rj}(z_j - z_r) = 0.$$

 $\mathcal{L}_{i} < 0$ $\mathcal{L}_{i} < 0$ $\mathcal{L}_{i} = \mathcal{L}_{i}$ $\mathcal{L}_{i} = \mathcal$

Aが既約だから, $a_{\alpha_i\alpha_{i+1}}<0$, $i=0,\ldots,s-1$ を満たす $r=a_0\neq\alpha_1\neq\cdots\neq\alpha_s=i_0$ が存在する.上で示したことより, $z_r=z_{\alpha_0}=z_{\alpha_1}$.今,(2.9) に戻り,r の代わりに α_1 と書く. α_1 -番目の行で狭義優対角性が成り立てば,矛盾を得る.そうでないとき, $z_{\alpha_1}=z_{\alpha_2}=\min_{j=1,\ldots,n}z_j$ が従う.このように続けると,必ず矛盾が得られる α_i が現れる(遅くても $\alpha_s=i_0$).

定理 2.29 $n \times n$ 行列 A が単調行列であるとする.二つの連立方程式 Ax = b, Ay = c $(x, y, b, c \in \mathbb{C}^n)$ の右辺ベクトルが $|b_i| \leq c_i \ \forall i$ を満たすなら,

$$|x_i| \le y_i \quad \forall i = 1, \dots, n.$$

Proof. $|x|:=(|x_1|,\ldots,|x_n|)^T$ という記号を用いる。A が単調なので, $A^{-1}\geq O$ が成り立つ。 $x=A^{-1}b,\ y=A^{-1}c$ より

$$|x| \le A^{-1}|b| \le A^{-1}c = y.$$

3 付録:楕円型方程式

楕円型方程式に関する基本的な事項を述べる.

3.1 問題の定式化

楕円型方程式の代表例はポアソン方程式に対する Dirichlet 問題

$$-\Delta u = f \text{ in } \Omega, \qquad u = u_D \text{ on } \partial\Omega \tag{3.1}$$

が挙げられる。この方程式は様々な現象のモデルとして現れる。例えば、熱源fの下での物体の定常な温度分布を表す。Dirichlet境界条件は境界上の温度が一定の温度に保たれることを意味する。

他に,水平な薄板の垂直方向の変位を表す重調和方程式

$$\Delta(\Delta u) = f$$

や非圧縮粘性流体の定常な流れを表すストークス方程式

$$\Delta u - p_x = f_1, \quad \Delta v - p_y = f_2, \quad u_x + v_y = 0$$

やコーシー・リーマン方程式

$$u_x - v_y = 0, \qquad u_y + v_x = 0$$

なども楕円型方程式である.

一般の2階楕円型方程式を定義する.

定義 3.1 $\Omega \subset \mathbb{R}^2$ を領域とする. 準線形方程式

$$a(x,y)u_{xx} + 2b(x,y)u_{xy} + c(x,y)u_{yy} + d(x,y,u,u_x,u_y) = f(x,y) \qquad (x,y) \in \Omega$$
(3.2)

は $a, c > 0, ac - b^2 > 0$ を満たすとき、 楕円型方程式 であるという.

方程式につける境界条件によって、様々な問題が得られる.

(1) Dirichlet 問題

$$u = u_D \qquad \text{on } \partial\Omega \tag{3.3}$$

(2) Neumann 問題

$$\frac{\partial u}{\partial n} = \varphi_N \qquad \text{on } \partial\Omega \tag{3.4}$$

(3) Robin 問題

$$\frac{\partial u}{\partial n} + ku = \varphi_N \qquad \text{on } \partial\Omega \tag{3.5}$$

(4) 混合問題: Ω の境界 $\partial\Omega$ を二つの部分 $\overline{\Gamma}_D$ (Dirichlet 境界)と $\overline{\Gamma}_N$ (Neumann 境界)に分けて $(\partial\Omega = \overline{\Gamma}_D \cup \overline{\Gamma}_N, \ \Gamma_D \cap \Gamma_N = \emptyset)$, Γ_D 上で Dirichlet 条件, Γ_N 上で Neumann 条件(または Robin 条件)を指定する.

3.2 解の存在と一意性

詳しいことは偏微分方程式を扱う本を参考にしていただくとして,ここでは二つの基本的なアプローチがあることだけ記しておく.

- フレドホルムの交代定理と最大値原理に基づいた直接法.
- 対応する変分問題を満たす弱解の存在と一意性を Lax-Milgram の定理を用いて示し、弱解が古典解と同じ正則性をもつことをあとで証明する方法。

変分問題は、方程式に試験関数vをかけて、領域 Ω で積分して、グリーンの定理を適用す

ることによって得られる。例えば、上の(3.1)の問題は次のように変形される:

$$u - \tilde{u}_D \in H_0^1(\Omega), \qquad \int_{\Omega} \nabla u \cdot \nabla v \, dx \, dy = \int_{\Omega} f v \, dx \, dy \quad \forall v \in H_0^1(\Omega).$$

ただし、 H^1 はソボレフ空間で、 \tilde{u}_D は u_D を Ω 全体に拡張した H^1 -関数である.

注.

$$-\Delta u = f \quad \text{in } \Omega, \qquad \frac{\partial u}{\partial n} = 0 \quad \text{on } \partial \Omega$$

のような Neumann 問題では、解が微分した状態のみで出ているため、解は定数を除いて決まる。 さらに、方程式を積分してみると、 $\int_{\Omega} f \, dx = 0$ が成り立つことが解の存在の必要事件であることがわかる。

3.3 解の正則性

微分方程式の解の滑らかさは数値スキームの収束のはやさに影響するため,正則性について考えることが重要である.

楕円型方程式の特徴は、その解がデータより滑らかであることである。滑らかさは方程式の最高微分の階数だけ上がる。例えば、2 階方程式 (3.2) の場合、解u はデータ(f など)よりさらに 2 回微分可能である。

注. 波動方程式 $u_{xx} - u_{yy} = 0$ のような双曲型方程式の解がデータより滑らかになるとは限らない. 解が不連続になることさえあり得る.

楕円型方程式 (3.2) の係数が定数で $\Omega = \mathbb{R}^2$ の場合, つまり,

$$au_{xx} + 2bu_{xy} + cu_{yy} + d_1u_x + d_2u_y + eu = f$$
 in \mathbb{R}^2

のとき, フーリエ変換を用いて, 次の解の評価が導かれる:

$$||u||_{H^{s+2}}^2 \le C_s(||f||_{H^s}^2 + ||u||_{L^2}^2).$$

ノルム $\|\cdot\|_{H^s}$ は s-階微分までのすべての微分の L^2 -ノルムを含むノルムである。よって、データ f の s-階までの微分が L^2 に属するならば、階の (s+2)-階までの微分は $L^2(\mathbb{R}^2)$ -関数になる。

有界領域 Ω の場合, 同様にして内部評価

$$||u||_{H^{s+2}(K)}^2 \le C_s(\Omega, K)(||f||_{H^s(\Omega)}^2 + ||u||_{L^2(\Omega)}^2)$$

が導かれる.ここで,K は $\overline{K} \subset \Omega$ を満たす集合である.また,ノルム $\|\cdot\|_{H^s(\Omega)}$ は $\|\cdot\|_{H^s}$ と同様に定義されるが,積分領域は Ω に制限される.

注. 係数がx,yの関数のときも同様な評価が得られるが、証明が複雑になる.

正則性に関する具体的な結果をいくつか紹介する。ここでは、方程式を

$$-(a_1u_x + a_2u_y)_x + (a_3u_x + a_4u_y)_y + (b_1u)_x + (b_2u)_y + cu = f$$

という形で考え, 境界条件(3.3)-(3.5)をそれぞれ考える.

(1) Dirichlet 境界条件 (3.3) の場合: $k \geq 0$ が存在し、 $\partial \Omega$ が C^{k+2} -多様体、 $a_i, b_i \in C^k(\overline{\Omega})$ 、 $f \in H^k(\Omega)$ 、 $u_D \in H^{3/2}(\partial \Omega)$ であれば、解は $u \in H^{k+2}(\Omega)$ を満たす.

境界が十分滑らかであれば、解の境界上の法線微分は境界値の境界に沿った微分だけの正則性をもつ.しかし、境界値が不連続であれば、解の法線微分が不連続な点で発散する.

- (2) 上と同じ Dirichlet 条件で領域が多角形領域の場合を考える。このとき、境界はリプシッツ にしかならず、(1) の結果が成立しない。簡単のために、ポアソン方程式 $-\Delta u = f$ に限って話を進める。
 - $f \in L^2(\Omega)$ で領域 Ω が凸であれば、上と同様に $u \in H^2(\Omega)$ が証明できる.
 - 領域 Ω が凸でないとき,解は H^2 -関数 + $r^{\pi/\omega}$ の特異性をもつ関数と書ける.ここで, $\omega > \pi$ は凹角の大きさである.よって, $u \in H^{3/2}(\Omega)$ が常に成り立つ.
- (3) Neumann と Robin 条件 (3.5) のとき,Dirichlet 条件と同じ仮定に加え, $k \in C^{k+1}(\partial\Omega)$ と $\varphi_N \in H^{k+1/2}(\partial\Omega)$ を仮定すれば, $u \in H^{k+2}(\Omega)$ が言える.(多角形領域における斉次 Neumann 境界条件つきのポアソン方程式の場合, $f \in L^2(\Omega)$ であれば, $u \in H^2(\Omega)$ (凸領域)または $u \in H^{3/2}(\Omega)$ (一般領域)が示せる.ただし, $\int_{\Omega} f = 0$ という条件が解の存在のために必要であることに注意する.)

境界が滑らかであれば、解の境界上の微分は φ_N だけの正則性をもつ。

(4) 混合問題ではデータが滑らかでも解が滑らかになるとは限らない。データと境界が滑らかで 解が $H^s(\Omega), s < \frac{3}{2}$ にしか属さないような問題の例が存在する。

注. 境界值問題

$$-\varepsilon \Delta u + \boldsymbol{b} \cdot \nabla u = f$$
 in Ω , $u = 0$ on $\partial \Omega$

について注意する. ただし, $\varepsilon > 0$ で, **b**は div **b** = 0 を満たす滑らかな関数であるとする.

方程式に解 u をかけて、積分して、グリーンの定理を用いると、

$$\varepsilon \int_{\Omega} |\nabla u|^2 dx = \int_{\Omega} f u dx$$

を得る。これより、ポアンカレ不等式を適用すれば、

$$\|\nabla u\|_{L^2(\Omega)} \le \frac{C_{\Omega}}{\varepsilon} \|f\|_{L^2(\Omega)}$$

というアプリオリ評価が従う.

そこで、 ε が小さいとき、解の勾配が十分にコントロールできない可能性がある。さらに、通

常の数値解法で得られた近似解の誤差評価も $\frac{1}{\varepsilon}$ の係数を含み、性能が悪くなる。真の解が境界層をもつとき、数値解は境界の近傍で不自然な振動を起こす傾向にある。これに対処するには、数値スキームの「安定化」という特別な配慮が必要である。

以上をまとめると、Dirichlet・Neumann・Robin 境界条件のつく楕円型問題は領域の内部と滑らかな境界の近傍でおとなしい(データより正則な)解をもつ。しかし、混合問題で境界条件の種類が変わる点や滑らかでない境界点において解の微分に特異性が現れる。この事実を楕円型方程式の数値解法に反映させる必要がある。

問題

問題 3.1 次の境界値問題の解の正則性 (特に境界上の正則性) を調べよ.

(1)
$$\Delta u = 0$$
 in $\Omega := \{(x,y); y > 0\}, \quad u(x,0) = \begin{cases} 0 & (x > 0) \\ 1 & (x < 0) \end{cases}$ (解は $u(x,y) = \frac{1}{\pi} \arctan \frac{y}{x}$)

(2)
$$\Delta u = 0$$
 in $\Omega := \{(x,y); y > 0\}$, $\frac{\partial u}{\partial y}(x,0) = \begin{cases} 0 & (x > 0) \\ \sqrt{|x|} & (x < 0) \end{cases}$ (解は極座標で $u(r,\theta) = -\frac{2}{3}r^{3/2}\cos\frac{3}{2}\theta$)

(3)
$$\Delta u = 0$$
 in $\Omega := \{(x,y); y > 0\}, \quad u(x,0) = 0 \ (x > 0), \quad u_y(x,0) = 0 \ (x < 0)$ (解は極座標で $u(r,\theta) = r^{1/2} \sin \frac{1}{2}\theta$)

(4)
$$\Delta u = 0$$
 in $\Omega := \{(r,\theta); \ 0 < r < r_0, 0 < \theta < \frac{3}{2}\pi\}, \quad u = 0$ on $\theta = 0, \frac{3}{2}\pi$ (解は $u(r,\theta) = r^{2/3}\sin\frac{2}{3}\theta$)

3.4 最大值原理

最大値原理は2階偏微分方程式を扱うための重要な道具となる。例として、2つの定理を挙げる。

定理 3.2 有界領域 Ω において関数 a,b,c が a,c>0, $b^2<ac$ を満たすとする. 関数 u が

$$au_{xx} + 2bu_{xy} + cu_{yy} \ge 0$$
 in Ω

を満たすならば、u の $\overline{\Omega}$ 上の最大値は境界 $\partial\Omega$ 上で達成される.

定理 3.3 関数 u が領域 Ω において楕円型方程式

$$au_{xx} + 2bu_{xy} + cu_{yy} + d_1u_x + d_2u_y + eu = 0$$

を満たし、係数が a,c>0、そして $e\leq 0$ を満たす関数であれば、u は Ω の内部に正の極大値または負の極小値を持たない.

4 放物型方程式

熱方程式に対する初期値境界値問題を例題として,時間発展を表す放物型方程式の差分法による解き方について考える.

4.1 例題

 $\Omega\subset\mathbb{R}^d$ を十分滑らかな境界をもつ領域とし, $Q_T:=\Omega\times(0,T)$ とする。この章では,楕円型オペレータ L が現れる方程式

$$\frac{\partial u}{\partial t} = Lu + f$$

に Dirichlet 境界条件

$$u\big|_{\partial\Omega\times(0,T)} = u_D$$

と初期条件

$$u(x,0) = u^0(x), \qquad x \in \Omega$$

をつけた問題を考える.

楕円型方程式についての付録でも触れたが、楕円型オペレータは例えば、次のものがある:

- $\Omega:=(a,b)\subset\mathbb{R}$ のとき、 $L=rac{\partial^2}{\partial x^2}$ とする.
- $\Omega \subset \mathbb{R}^d$ (一般次元) のとき, $L = \Delta$ (ラプラス作用素) とする.
- $\Omega \subset \mathbb{R}^d$ (一般次元) のとき,

$$L = \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial}{\partial x_j} \right)$$

とする. ただし, L が楕円型になるには $a_{ij} \in C^1(\Omega)$ が

$$\lambda |\boldsymbol{\xi}|^2 \le \sum_{i,j=1}^d a_{ij} \xi_i \xi_j \le \Lambda |\boldsymbol{\xi}|^2, \qquad \lambda, \Lambda > 0$$

を満たさなければならない.

以降では、次の問題の数値解法について考える。

解く問題

次の条件を満たす関数 $u: \overline{Q}_T \to \mathbb{R}$, $\frac{\partial u}{\partial t}$, $\frac{\partial^2 u}{\partial x^2} \in C(\overline{Q}_T)$ を求める:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f$$
 in $Q_T = (a, b) \times (0, T)$, (4.1)

$$u(a,t) = A(t), \quad u(b,t) = B(t) \quad \text{for } t \in (0,T),$$
 (4.2)

$$u(x,0) = u^{0}(x)$$
 for $x \in (a,b)$. (4.3)

物理的な解釈として、(4.1) がまっすぐな針金 (a,b) のなかの熱伝導を近似的に表していると言える。そのとき、u は温度で f は熱源の密度の意味をもつ。

4.2 差分法による近似

まず, \overline{Q}_T において格子を構成する.ここでは,空間領域(この場合, $x\in[a,b]$)の格子と時間区間(この場合, $t\in[0,T]$)の格子の積として \overline{Q}_T の格子を定義する.

• 空間の差分幅を

$$h := \frac{b - a}{N}, \quad N \in \mathbb{N}$$

とし, [a,b] の節点を $x_i = a + ih$, i = 0,...,N とする.

• 時間の差分幅を

$$\tau := \frac{T}{r}, \quad r \in \mathbb{N}$$

とし, [0,T] の節点を $t_k = k\tau$, k = 0, ..., r とする.

ullet \overline{Q}_T の節点は次のように設定する:

$$\{(x_i, t_k); i = 0, \dots, N, k = 0, \dots, r\}.$$

次に、節点で関数の微分を近似する。そのため、u が \overline{Q}_T において連続な微分 $\frac{\partial^2 u}{\partial t^2}$, $\frac{\partial^4 u}{\partial x^4}$ があると仮定する。節点 $(x_i,t_k)\in \overline{Q}_T$ をとって、 $x_i\in (a,b)$ であれば、

$$\frac{\partial^2 u}{\partial x^2}(x_i, t_k) = \frac{u(x_{i-1}, t_k) - 2u(x_i, t_k) + u(x_{i+1}, t_k)}{h^2} + O(h^2).$$

式を見やすくするために,

$$u_i^k := u(x_i, t_k)$$

という記号を導入すれば $(f_i^k$ なども同様に定義する),

$$\frac{\partial^2 u}{\partial x^2}(x_i, t_k) = \frac{1}{h^2} \left(u_{i-1}^k - 2u_i^k + u_{i+1}^k \right) + O(h^2).$$

時間微分については二通りの近似のしかたがある:

(a)
$$t_k \in [0,T)$$
 のとき,

$$\frac{\partial u}{\partial t}(x_i, t_k) = \frac{1}{\tau} \left(u_i^{k+1} - u_i^k \right) + O(\tau).$$

(b) t_k ∈ (0, T] \emptyset とき,

$$\frac{\partial u}{\partial t}(x_i, t_k) = \frac{1}{\tau} \left(u_i^k - u_i^{k-1} \right) + O(\tau).$$

このようにして、二つのスキームを得る:

(a)
$$(x_i, t_k)$$
, $i = 1, ..., N-1$, $k = 0, ..., r-1$ に対し,

$$\frac{u_i^{k+1} - u_i^k}{\tau} = \frac{u_{i-1}^k - 2u_i^k + u_{i+1}^k}{h^2} + f_i^k + O(\tau + h^2).$$

(b) (x_i, t_k) , i = 1, ..., N-1, k = 0, ..., r-1 に対し,

$$\frac{u_i^{k+1} - u_i^k}{\tau} = \frac{u_{i-1}^{k+1} - 2u_i^{k+1} + u_{i+1}^{k+1}}{h^2} + f_i^{k+1} + O(\tau + h^2).$$

誤差項 $O(\tau + h^2)$ を無視して、節点 (x_i, t_k) における近似解の値を U_i^k と書くことにすると、

(a)
$$\frac{U_i^{k+1} - U_i^k}{\tau} = \frac{U_{i-1}^k - 2U_i^k + U_{i+1}^k}{h^2} + f_i^k, \quad i = 1, \dots, N-1, \ k = 0, \dots, r-1,$$

(b)
$$\frac{U_i^{k+1} - U_i^k}{\tau} = \frac{U_{i-1}^{k+1} - 2U_i^{k+1} + U_{i+1}^{k+1}}{h^2} + f_i^{k+1}, \quad i = 1, \dots, N-1, \ k = 0, \dots, r-1$$

という数値スキームを導いた。さらに、境界条件と初期条件より以下の関係式を得る:

境界条件より
$$U_0^k = A(t_k), \ U_N^k = B(t_k), \ k=1,\ldots,r$$
 初期条件より $U_i^0 = u^0(x_i), \ i=0,\ldots,N.$

(a) のスキームでは,k=0 からはじめ, $U_i^k, i=0,\ldots,N$ の既知の値から $U_i^{k+1}, i=1,\ldots,N-1$ を計算する.このとき, U_i^{k+1} が陽に求まるから,この解き方を <mark>陽解法</mark> と言う.一方では,(b) のスキームで時刻 t_k での既知の値 $U_i^k, i=0,\ldots,N$ より次の時刻 t_{k+1} での値 $U_i^{k+1}, i=1,\ldots,N-1$ を計算するには未知数 $U_1^{k+1},\ldots,U_{N-1}^{k+1}$ についての 1 次連立方程式を解かなければならないので,(b) のアプローチを 陰解法 と言う.

(b) の式に $\gamma \in [0,1]$ をかけて, (a) に $(1-\gamma)$ をかけた式と足し合わせると,

$$\frac{U_i^{k+1} - U_i^k}{\tau} = \gamma \frac{U_{i-1}^{k+1} - 2U_i^{k+1} + U_{i+1}^{k+1}}{h^2} + (1 - \gamma) \frac{U_{i-1}^k - 2U_i^k + U_{i+1}^k}{h^2} + \gamma f_i^{k+1} + (1 - \gamma) f_i^k$$

というより一般的なスキームを得る。ただし、 $\gamma=0$ のとき陽解法となり、 $\gamma=1$ のとき陰解法となる。 $\gamma\in(0,1]$ のときも連立方程式を解くことになるので、陰解法である。 $\gamma=\frac{1}{2}$ のケースは特別な性質をもつから、Crank-Nicolson 法という名前がついている。

次に離散化の誤差を調べる.

定理 **4.1** 問題 (4.1)-(4.3) の解 u が $\frac{\partial^2 u}{\partial t^2}$, $\frac{\partial^4 u}{\partial x^4} \in C(\overline{Q}_T)$ を満たすならば,

$$\frac{u_i^{k+1} - u_i^k}{\tau} = \gamma \frac{u_{i-1}^{k+1} - 2u_i^{k+1} + u_{i+1}^{k+1}}{h^2} + (1 - \gamma) \frac{u_{i-1}^k - 2u_i^k + u_{i+1}^k}{h^2} + \gamma f_i^{k+1} + (1 - \gamma) f_i^k + \varepsilon_i^k$$

が $i=1,\ldots,N-1,\ k=0,\ldots,r-1,\ \gamma\in[0,1]$ について成り立つ。ただし、 $\varepsilon_i^k=\varepsilon_i^k(h,\tau)=O(\tau+h^2)$ 、すなわち、定数 $M,h_0,\tau_0>0$ が存在し、次が成り立つ:

$$|\varepsilon_i^k| \le M(\tau + h^2) \quad \forall \tau \in (0, \tau_0) \ \forall h \in (0, h_0).$$

さらに、 $\frac{\partial^3 u}{\partial t^3}$, $\frac{\partial^4 u}{\partial x^4}$ $\in C(\overline{Q}_T)$ で、 $\gamma = \frac{1}{2}$ ならば、

$$\varepsilon_i^k = O(\tau^2 + h^2), \quad i = 1, \dots, N - 1, \quad k = 0, \dots, r - 1.$$

Proof. 前半は楕円型方程式と同様なテーラー展開により示されるので省略して、Crank-Nicolson $(\gamma = \frac{1}{2})$ のスキーム

$$\frac{1}{\tau} \left(u_i^{k+1} - u_i^k \right) = \frac{1}{2h^2} \left(u_{i-1}^{k+1} - 2u_i^{k+1} + u_{i+1}^{k+1} + u_{i-1}^k - 2u_i^k + u_{i+1}^k \right) + \frac{1}{2} \left(f_i^{k+1} + f_i^k \right) + \varepsilon_i^k$$

のみ調べる.

$$\frac{1}{2} \left(\frac{\partial u(x_i, t_k)}{\partial t} + \frac{\partial u(x_i, t_{k+1})}{\partial t} \right) = \frac{u_i^{k+1} - u_i^k}{\tau} + O(\tau^2)$$
(4.4)

を示せばよいということに注意する.

テーラー展開より,

$$u_{i}^{k+1} = u(x_{i}, t_{k+1}) = u(x_{i}, t_{k} + \frac{\tau}{2}) + \frac{\tau}{2} \frac{\partial u}{\partial t} (x_{i}, t_{k} + \frac{\tau}{2}) + \frac{1}{2} (\frac{\tau}{2})^{2} \frac{\partial^{2} u}{\partial t^{2}} (x_{i}, t_{k} + \frac{\tau}{2}) + O(\tau^{3}),$$

$$u_{i}^{k} = u(x_{i}, t_{k}) = u(x_{i}, t_{k} + \frac{\tau}{2}) - \frac{\tau}{2} \frac{\partial u}{\partial t} (x_{i}, t_{k} + \frac{\tau}{2}) + \frac{1}{2} (\frac{\tau}{2})^{2} \frac{\partial^{2} u}{\partial t^{2}} (x_{i}, t_{k} + \frac{\tau}{2}) + O(\tau^{3}).$$

これらの式の差をとると,

$$\frac{u_i^{k+1} - u_i^k}{\tau} = \frac{\partial u}{\partial t}(x_i, t_k + \frac{\tau}{2}) + O(\tau^2). \tag{4.5}$$

一方で, テーラー展開を適用して,

$$\frac{\partial u}{\partial t}(x_i, t_{k+1}) = \frac{\partial u}{\partial t}(x_i, t_k + \frac{\tau}{2}) + \frac{\tau}{2} \frac{\partial^2 u}{\partial t^2}(x_i, t_k + \frac{\tau}{2}) + O(\tau^2),$$

$$\frac{\partial u}{\partial t}(x_i, t_k) = \frac{\partial u}{\partial t}(x_i, t_k + \frac{\tau}{2}) - \frac{\tau}{2} \frac{\partial^2 u}{\partial t^2}(x_i, t_k + \frac{\tau}{2}) + O(\tau^2).$$

を得るので,これらの式の和をとると,

$$\frac{1}{2} \left[\frac{\partial u}{\partial t}(x_i, t_{k+1}) + \frac{\partial u}{\partial t}(x_i, t_k) \right] = \frac{\partial u}{\partial t}(x_i, t_k + \frac{\tau}{2}) + O(\tau^2).$$

上で導いた関係式 (4.5) と合わせると, (4.4) が示せた.

あとは,

$$\frac{1}{h^2} \left(u_{i-1}^k - 2u_i^k + u_{i+1}^k \right) = \frac{\partial^2 u}{\partial x^2} (x_i, t_k) + O(h^2),$$

そして k を k+1 で置き換えた式が成立することを用いると、証明の冒頭で定義した誤差 ε_i^k に対して次の表現を得る:

$$\begin{split} \varepsilon_{i}^{k} &= \frac{1}{2} \left[\frac{\partial u}{\partial t}(x_{i}, t_{k+1}) + \frac{\partial u}{\partial t}(x_{i}, t_{k}) \right] - \frac{1}{2} \frac{\partial^{2} u}{\partial x^{2}}(x_{i}, t_{k+1}) - \frac{1}{2} \frac{\partial^{2} u}{\partial x^{2}}(x_{i}, t_{k}) - \frac{1}{2} f_{i}^{k+1} - \frac{1}{2} f_{i}^{k} \\ &+ O(\tau^{2}) + O(h^{2}) \\ &= O(\tau^{2} + h^{2}). \end{split}$$

最後の等号が正しいのは、u が $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f$ の解であるから上式の 1 行目の右辺が消えるためである。なお、f の正則性が必要でなかったことに注意しよう。

陰解法で時間レベル k+1 での未知の近似解を求めるには $n_h:=(N-1)$ 本の連立方程式を解く。未知ベクトルを $U_h^k=(U_1^k,\ldots,U_{n_h}^k)^T\in\mathbb{R}^{n_h}$ とする。境界条件より決まっている $U_{i-1}^k,U_{i+1}^k,U_{i-1}^{k+1},U_{i+1}^{k+1}$ の値は f の値を含む右辺に移項する。したがって,

$$\frac{1}{\tau} \left(\boldsymbol{U}_{h}^{k+1} - \boldsymbol{U}_{h}^{k} \right) = -\gamma A_{h} \boldsymbol{U}_{h}^{k+1} - (1 - \gamma) A_{h} \boldsymbol{U}_{h}^{k} + \boldsymbol{\varphi}_{h}^{k}$$

$$\tag{4.6}$$

の連立方程式となる。 ただし, A_h は

$$A_h := \frac{1}{h^2} \begin{pmatrix} 2 & -1 & 0 & 0 & \dots \\ -1 & 2 & -1 & 0 & \dots \\ 0 & -1 & 2 & -1 & \dots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix}$$

で定義される行列である。 A_h は既約優対角行列で正定値である。また、右辺ベクトルは

$$\varphi_h^k = \left(\gamma f_1^{k+1} + (1-\gamma)f_1^k + \gamma A(t_{k+1}) + (1-\gamma)A(t_k), \gamma f_2^{k+1} + (1-\gamma)f_2^k, \dots, \gamma f_{N-2}^{k+1} + (1-\gamma)f_{N-2}^k, \gamma f_{N-1}^{k+1} + (1-\gamma)f_{N-1}^k + \gamma B(t_{k+1}) + (1-\gamma)B(t_k)\right)^T$$

と書け、初期ベクトル U_h^0 は初期条件より $U_h^0 = (u^0(x_1), \dots, u^0(x_{N-1}))^T$ のように与えられる.

問題

問題 4.1 次の初期値境界値問題の差分スキームを書いて、(4.6)と同じ形の連立方程式が得られ

ることを確かめよ、

$$\begin{split} &\frac{\partial u}{\partial t} = \Delta u + f & \text{in } Q_T = (0,1)^2 \times (0,T), \\ &u(x,y,t) = u_D(x,y,t), & (x,y) \in \partial \Omega, \ t \in (0,T), \\ &u(x,y,0) = u^0(x,y), & (x,y) \in \Omega. \end{split}$$

4.3 差分スキームと近似解の性質

幅広い問題に対応するために、以降では任意のベクトル $U_h^k, \varphi_h^k \in \mathbb{R}^{n_h}$ と $n_h \times n_h$ 行列 A_h に対するスキーム (4.6) を考える。ただし、 A_h が正定値の対称行列であると仮定する。

(4.6)を書き換えると,

$$(I + \gamma \tau A_h) \boldsymbol{U}_h^{k+1} = (I - (1 - \gamma)\tau A_h) \boldsymbol{U}_h^k + \tau \boldsymbol{\varphi}_h^k$$

となるので,

$$A_H = I + \gamma \tau A_h, \qquad A_D = I - (1 - \gamma)\tau A_h$$

とおけば,

$$A_H \boldsymbol{U}_h^{k+1} = A_D \boldsymbol{U}_h^k + \tau \boldsymbol{\varphi}_h^k \tag{4.7}$$

という形の連立方程式を得る。

補題 4.2 行列 A_H は対称で正定値行列である。よって,すべての $\tau > 0, \gamma \in [0,1]$ について正則行列で,スキーム (4.7) はただ一つの解をもつ。

Proof. A_H は二つの対称な正定値行列 (I と $\gamma \tau A_h$) の和で書かれるので、主張は明らかである. \square

次にスキームの収束性を調べる。まず、記号を整理する。 $U_h^k \in \mathbb{R}^{n_h}$ は近似解ベクトルである。u は真の解であるが、それを(k-番目の時間レベルにおいて近似解を求める)節点に制限することにより得られるベクトルを u_h^k とする。すると、

$$oldsymbol{\eta}_h^k \coloneqq oldsymbol{u}_h^k - oldsymbol{U}_h^k \ \in \mathbb{R}^{n_h}$$

は k-番目の時間レベルにおけるスキームの誤差となる。この誤差を \mathbb{R}^{n_h} のノルム $\|\cdot\|$ を用いて 測る.

定義 **4.3** (0,0) が集合 $\mathcal{M} \subset \{(h,\tau); h > 0, \tau > 0\}$ の集積点であるとする.

$$\lim_{\substack{(h,\tau)\to(0,0)\\(h,\tau)\in\mathcal{M}}}\max_{t_k\in[0,T]}\|\boldsymbol{\eta}_h^k\|=0$$

が成り立つとき,スキーム (4.7) が集合 M においてノルム $\|\cdot\|$ で収束すると言う.ただし, \max では離散値 t_0, t_1, \ldots, t_r のみで最大値をとる.

誤差 η_h^k が満たす方程式を導く。離散化の誤差を ε_h^k とすれば、真の解は次の式を満たす:

$$A_H \boldsymbol{u}_h^{k+1} = A_D \boldsymbol{u}_h^k + \tau \boldsymbol{\varphi}_h^k + \tau \boldsymbol{\varepsilon}_h^k.$$

この式から (4.7) をひくと, $A_H \eta_h^{k+1} = A_D \eta_h^k + \tau \varepsilon_h^k$ を得る.また,近似解の初期条件と境界条件は真の解のものと一致するため,スキームの誤差はゼロの初期条件とゼロの境界条件に従うことがわかる.よって,列 $\{\eta_h^k\}_{k=0}^r$ について,下記の方程式が成り立つ.

$$\eta_h^0 = \mathbf{0},$$

$$A_H \eta_h^{k+1} = A_D \eta_h^k + \tau \varepsilon_h^k, \quad k = 0, \dots, r-1.$$

$$(4.8)$$

定義 4.4

• 定数 $K, h_0, \tau_0 > 0$ が存在し,

$$\max_{t_k \in [0,T]} \|\boldsymbol{\varepsilon}_h^k\| \le K \left(h^p + \tau^q \right) \qquad \forall h \in (0,h_0) \ \forall \tau \in (0,\tau_0)$$

が成り立つとき,スキームは空間オーダーp,時間オーダーqであると言う.

- 空間オーダー p と時間オーダー q が p,q>0 を満たすとき,スキームは 適合性 をもつと 言う
- 定数 C > 0 が存在し, (4.8) の解 η_h^k が

$$\max_{t_k \in [0,T]} \|\boldsymbol{\eta}_h^k\| \le C \max_{t_j \in [0,T]} \|\boldsymbol{\varepsilon}_h^j\| \qquad \forall (h,\tau) \in \mathcal{M}$$

を満たすとき、スキームはMにおいてノルム $\|\cdot\|$ で右辺に対して $\overline{\text{安定}}$ であると言う.

なお, $\mathcal{M} = (0, h_1) \times (0, \tau_1)$ とできるような $h_1, \tau_1 > 0$ が存在すれば,スキームが 無条件安定 であると言う.「無条件」の部分を省略することが多い.一方で, \mathcal{M} をこのように表すことができないときは,条件つき安定性 となる.

上の定義よりすぐに次の定理を得る.

定理 4.5 スキームは M においてノルム $\|\cdot\|$ で右辺に対して安定で、空間オーダー p、時間 オーダー q ならば、M>0 が存在し、次が成り立つ:

$$\max_{t_k \in [0,T]} \|\boldsymbol{\eta}_h^k\| \le M(h^p + \tau^q) \qquad \forall h, \tau \in \mathcal{M} \cap (0,h_0) \times (0,\tau_0).$$

この結果より、次の主張も明らかである.

定理 4.6 スキームは M においてノルム $\|\cdot\|$ で右辺に対して安定で、適合性をもつならば、M 上でノルム $\|\cdot\|$ で収束する。すなわち、

$$\max_{t_k \in [0,T]} \|\boldsymbol{\eta}_h^k\| \to 0 \quad \text{as } h, \tau \to 0, \quad (h,\tau) \in \mathcal{M}.$$

標語的に言えば,

安定性 + 適合性 ⇒ 収束性.

以上の理論を我々の場合に適用する。定理4.1より、真の解uが十分なめらかであれば、

$$\|\varepsilon_h^k\|_{\infty} \le K(h^2 + \tau) \qquad h \in (0, h_0), \ \tau \in (0, \tau_0)$$
 (4.9)

が成立するので、スキームは適合性をもつ $(\gamma = \frac{1}{2}$ の場合は、 $K(h^2 + \tau^2)$ とできる).空間次元が 2 以上の場合も同様な結果が導かれる.

次にスキームの安定性について確かめる。 $S := A_H^{-1} A_D$ とおけば、(4.8)の式は

$$\boldsymbol{\eta}_h^{k+1} = S\boldsymbol{\eta}_h^k + \tau A_H^{-1}\boldsymbol{\varepsilon}_h^k$$

と書き直される。以下では、行列ノルム $\|\cdot\|$ として、 \mathbb{R}^{n_h} のベクトルノルム $\|\cdot\|$ に誘導されるものだとして議論を進める。

定理 4.7 定数 c > 0 が存在し、

$$||A_H^{-1}|| \le c, \quad ||S^{\ell}|| \le c \qquad \forall \ell = 0, 1, \dots, \ \forall (h, \tau) \in \mathcal{M}$$

が成り立つならば、スキーム (4.6) は M においてノルム ||·||で右辺に対して安定である.

Proof. $\eta_h^k=\mathbf{0},$ $\eta_h^{k+1}=S\eta_h^k+ au A_H^{-1}arepsilon_h^k$ という漸化式より,

$$\eta_h^1 = \tau A_H^{-1} \varepsilon_h^0, \qquad \eta_h^2 = \tau S A_H^{-1} \varepsilon_h^0 + \tau A_H^{-1} \varepsilon_h^1, \quad \dots$$

のようになるので,

$$\eta_h^k = \tau \sum_{\nu=1}^k S^{k-\nu} A_H^{-1} \varepsilon_h^{\nu-1}$$

と推測する (証明は数学的帰納法を用いる). したがって, $(h,\tau) \in M$ であれば,

$$\begin{split} \| \boldsymbol{\eta}_{h}^{k} \| & \leq & \tau \sum_{\nu=1}^{k} \| S^{k-\nu} \| \| A_{H}^{-1} \| \| \boldsymbol{\varepsilon}_{h}^{\nu-1} \| \\ & \leq & \tau c^{2} \sum_{\nu=1}^{k} \| \boldsymbol{\varepsilon}_{h}^{\nu-1} \| \\ & \leq & \tau k c^{2} \max_{t_{j} \in [0,T]} \| \boldsymbol{\varepsilon}_{h}^{j} \| \\ & \leq & c^{2} T \max_{t_{j} \in [0,T]} \| \boldsymbol{\varepsilon}_{h}^{j} \|. \end{split}$$

よって、 $C = c^2T$ とおけば、

$$\max_{t_k \in [0,T]} \|\boldsymbol{\eta}_h^k\| \le C \max_{t_j \in [0,T]} \|\boldsymbol{\varepsilon}_h^j\| \qquad \forall (h,\tau) \in \mathcal{M}$$

を示した.

上の補題の条件を確かめるには, \mathbb{R}^{n_h} における具体的なノルムを決める必要がある. $(n_h=N-1,h=\frac{b-a}{N}$ だったことを思い出そう.)最大値ノルム $\|\cdot\|_{\infty}$ をとることができるが,評価するときに問題が起き,結果が制限される.もう一つの選択肢は離散 L^2 -ノルムである. $\eta_h=(\eta_1,\dots,\eta_{n_h})^T\in\mathbb{R}^{n_h}$ に対して,

$$\|\boldsymbol{\eta}_h\|_2 := \left(h\sum_{i=1}^{n_h} \eta_i^2\right)^{1/2} = \sqrt{h}\|\boldsymbol{\eta}_h\|_E$$
 (4.10)

と定義される。(4.10) の第 2 式より, $\|S\|_2 = \|S\|_E$ が従う。また,S が対称行列なら, $\|S\|_2 = \rho(S)$ も正しい。

注. L^2 -ノルムの定義式 (4.10) において h が係数としてつく理由について考える。 $u \in C([a,b])$ に対し, $\mathbf{u}_h := (u(x_1), \dots, u(x_{n_h}))^T$ とする(ここで, $x_i, i = 0, \dots, n_h + 1$ はスキームの導出で導入した節点である)。そのとき,

$$\|\boldsymbol{u}_h\|_2 = \left(h\sum_{i=1}^{n_h} u^2(x_i)\right)^{1/2}$$

であるが, $h \rightarrow 0+$ とすることにより,

$$\lim_{h \to 0+} \|\mathbf{u}_h\|_2 = \left(\int_a^b u^2(x) \, dx\right)^{1/2} = \|u\|_{L^2(a,b)}$$

を得るので、係数れの役割ははっきりする。

補題 4.8 行列 A_h が対称な正定値行列ならば,

(1)
$$Sp(A_H^{-1}) = \left\{ \mu = \frac{1}{1 + \tau \gamma \nu}; \ \nu \in Sp(A_h) \right\}$$

(2)
$$Sp(S) = \left\{ \mu = \frac{1 - \tau(1 - \gamma)\nu}{1 + \tau\gamma\nu}; \ \nu \in Sp(A_h) \right\}$$

Proof. 証明の方法が同様のため、(2) のみ示す. $\mu \in Sp(S)$ は $\det(S - \mu I) = 0$ と同値である. そこで、 $\mu = \frac{1 - \tau(1 - \gamma)\nu}{1 + \tau \gamma \nu}$ を満たすように ν を選ぶと(つまり、 $\nu := \frac{1 - \mu}{\tau((1 - \gamma) + \gamma \mu)}$ とおくと)、

$$S - \mu I = (I + \tau \gamma A_h)^{-1} (I - \tau (1 - \gamma) A_h) - \frac{1 - \tau (1 - \gamma) \nu}{1 + \tau \gamma \nu} I$$
$$= \frac{\tau}{1 + \tau \gamma \nu} (I + \tau \gamma A_h)^{-1} (\nu I - A_h)$$

を得る. $\frac{\tau}{1+\tau\gamma\nu}\det{(I+\tau\gamma A_h)^{-1}}\neq0$ が常に成り立つので,

$$\det(S - \mu I) = 0 \quad \Leftrightarrow \quad \det(\nu I - A_h) = 0$$

がわかった. これは主張(2)と同じことである.

補題 $\mathbf{4.9}$ 行列 A_h が対称な正定値行列ならば, A_H^{-1} と S は対称行列である.

Proof.

- $A_H = I + \tau \gamma A_h$ はすべての $\tau > 0$, $\gamma \in [0,1]$ について対称で正定値だから, A_H^{-1} も対称で正定値である.
- $S^T = S$ とは, $(A_H^{-1}A_D)^T = A_H^{-1}A_D$ ということで, A_H^{-1} と A_D が対称であるから,これは $A_DA_H^{-1} = A_H^{-1}A_D$ と同じである.つまり, A_D と A_H が可換であることを示せばよい.これは A_D , A_H の定義式を書いて確かめればよい.

系 4.10

$$||A_H^{-1}||_2 = \rho(A_H^{-1}), \qquad ||S||_2 = \rho(S).$$

定理 **4.11** (離散 L^2 -ノルムでの安定性)

- (1) $||A_H^{-1}||_2 \le 1 \quad \forall (h, \tau) \in (0, +\infty)^2$
- (2) $\gamma \in [0, \frac{1}{2}) \ \mathcal{O} \ \xi \ \xi$,

$$0 < \tau \le \frac{2}{\rho(A_h)} \frac{1}{1 - 2\gamma}, \qquad 0 < h$$

を満たす (h,τ) に対し $||S||_2 \le 1$ が成り立つ.

(3) $\gamma \in [\frac{1}{2},1]$ のとき、すべての $(h,\tau) \in (0,+\infty)^2$ に対し $\|S\|_2 \leq 1$ が成り立つ。

注. $\gamma \in [0, \frac{1}{2})$ のとき,スキームは $\|\cdot\|_2$ -ノルムで条件つき安定である.安定性が成り立つ集合は

$$\mathcal{M} = \left\{ (h, \tau); \ h > 0, \ 0 < \tau \le \frac{2}{\rho(A_h)} \frac{1}{1 - 2\gamma} \right\}$$

である. 一方, $\gamma \in [\frac{1}{2}, 1]$ のとき, 無条件安定である.

Proof.

- (1) $\|A_H^{-1}\|_2 = \rho(A_H^{-1})$ なので, $\rho(A_H^{-1}) \le 1$ を示せばよい.任意の固有値 $\vartheta \in Sp(A_H^{-1})$ に対し, $\vartheta = \frac{1}{1+\tau\gamma\nu}$ となる $\nu \in Sp(A_h)$ が存在する.ここで, $\nu > 0$ (A_h は正定値), $\tau > 0$, $\gamma \in [0,1]$ であるため, $0 < \vartheta \le 1$ を得る. $\vartheta \in Sp(A_H^{-1})$ が任意だったので, $\rho(A_H^{-1}) \le 1$.
- (2,3) $\|S\|_2 = \rho(S)$ であるから, $\rho(S) \le 1$ を示せばよい.(1) と同様に, $\mu \in Sp(S)$ を任意に選ぶ. すると, $\nu \in Sp(A_h)$ が存在し, $\mu = \frac{1-\tau(1-\gamma)\nu}{1+\tau\gamma\nu}$ が成り立つ.このとき, $|\mu| \le 1$ という条件は,次の条件と同値である:

$$[1 - \tau(1 - \gamma)\nu]^2 \leq [1 + \tau\gamma\nu]^2$$
$$-2\tau\nu + \tau^2(1 - 2\gamma)\nu^2 \leq 0$$
$$\tau\nu(1 - 2\gamma) \leq 2.$$

今, $\gamma \in [0, \frac{1}{2})$ なら, 最後の条件は

$$\tau \le \frac{2}{\nu} \frac{1}{1 - 2\gamma} \qquad \forall \nu \in Sp(A_h)$$

と同じである.

一方, $\gamma \in [\frac{1}{2}, 1]$ のとき,この条件は常に満たされる.

応用では, $\|\cdot\|_2$ の行列ノルムは計算しづらいため, ∞ -ノルムなど計算しやすいノルムで置き換えることが望ましい.そこで, \mathbb{R}^{n_h} のノルムに誘導される行列ノルム $\|\cdot\|$ なら, $\rho(A_h) \leq \|A_h\|$

が常に成り立つので、 $\frac{1}{\rho(A_h)} \ge \frac{1}{\|A_h\|_{\infty}}$ より、

$$0 < \tau \le \frac{2}{\|A_h\|_{\infty}} \frac{1}{1 - 2\gamma}$$

が $\gamma \in [0, \frac{1}{2})$ のときの安定性のための十分条件となることがわかる.

例題として考えた問題に対する (4.6) の具体的な行列を見ると, $\|A_h\|_{\infty}=rac{4}{h^2}$ であるから,

$$0<\tau\leq \frac{h^2}{2}\,\frac{1}{1-2\gamma}\qquad \text{for }\,\gamma\in[0,\tfrac{1}{2})$$

という安定性の条件となる。空間の格子を細かくしたときに、これは厳しい制限である。さらに、時間の刻み幅を小さくすることによって、計算しなければならない時間ステップの数が増え、丸め誤差の影響も無視できなくなってしまう。その意味では、連立方程式を解く必要があっても、 $\gamma \in [\frac{1}{5},1]$ に対する陰解法を選択した方がよいことがよくある。

例 4.12 空間 2次元の領域 $(0,1)^2 \times (0,T)$ において問題 $\frac{\partial u}{\partial t} = \Delta u + f$ を考えたとき, $\|A_h\|_{\infty} = \frac{8}{h^2}$ より,安定性条件は

$$0 < \tau \le \frac{h^2}{4} \frac{1}{1 - 2\gamma}, \qquad \gamma \in [0, \frac{1}{2})$$

となる.ただし,離散 L^2 -ノルムは $\|m{\eta}_h\|_2 = (h^2 \sum_{P_{ij} \in \Omega_h} \eta_{ij}^2)^{1/2}$ のようになる.

最後に例題の収束のはやさについて調べる。安定性(定理4.11)より解の誤差 η_h を離散化の誤差 ε_h でおさえることができ, ε_h の評価は (4.9) で計算されている。しかし,安定性は L^2 -ノルムについて示したのに対し,評価は ∞ -ノルムで行っているから,この二つのノルムの関係を確認しておく必要がある:

$$\|\boldsymbol{\varepsilon}_h^k\|_2 = \left(h\sum_{i=1}^{n_h} (\boldsymbol{\varepsilon}_i^k)^2\right)^{1/2} \leq \sqrt{h} \left(n_h \|\boldsymbol{\varepsilon}_h\|_{\infty}^2\right)^{1/2} = \sqrt{hn_h} \|\boldsymbol{\varepsilon}_h\|_{\infty} \leq \sqrt{b-a} \|\boldsymbol{\varepsilon}_h\|_{\infty}.$$

以上を合わせて,次の定理を得る.

定理 **4.13** 問題 (4.1)-(4.3) の解 u が $\frac{\partial^2 u}{\partial t^2}$, $\frac{\partial^4 u}{\partial x^4} \in C(\overline{Q}_T)$ を満たすならば,

(1) $\gamma \in [0, \frac{1}{2}) \mathcal{O} \mathcal{E}$,

$$\max_{t_j \in [0,T]} \|\boldsymbol{\eta}_h^j\|_2 \le TK\sqrt{b-a}(h^2+\tau) \qquad \forall h \in (0,h_0), \ \forall \tau \in (0,\min\{\tau_0,\frac{h^2}{2},\frac{1}{1-2\gamma}\}).$$

(2) $\gamma \in [\frac{1}{2},1]$ のとき、上の評価式は $\forall h \in (0,h_0), \ \forall \tau \in (0,\tau_0)$ について成り立つ。

(3)
$$\gamma=rac{1}{2}$$
 で、さらに $rac{\partial^3 u}{\partial t^3}\in C(\overline{Q}_T)$ ならば、

$$\max_{t_j \in [0,T]} \| \boldsymbol{\eta}_h^j \|_2 \le TK\sqrt{b-a}(h^2 + \tau^2) \qquad \forall h \in (0,h_0), \ \forall \tau \in (0,\tau_0).$$

5 双曲型方程式

波動方程式に対する初期値境界値問題を例題として,時間発展を表す双曲型方程式の差分法に よる解き方について考える.

5.1 例題

両端の高さを指定した弦の振動を表す問題を例として用いる.

解く問題

次の条件を満たす関数 $u:\overline{Q}_T:=\Omega imes (0,T) o \mathbb{R}$, $\frac{\partial^2 u}{\partial t^2}, \frac{\partial^2 u}{\partial x^2} \in C(\overline{Q}_T)$ を求める:

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + f \qquad \text{in } Q_T = (a, b) \times (0, T), \tag{5.1}$$

$$u(a,t) = A(t), \quad u(b,t) = B(t) \quad \text{for } t \in (0,T),$$
 (5.2)

$$u(x,0) = g^{0}(x)$$
 for $x \in (a,b)$, (5.3)

$$u_t(x,0) = g^1(x)$$
 for $x \in (a,b)$. (5.4)

条件 (5.3) は弦の初期形状,条件 (5.4) は弦の初速度を表現している.

5.2 差分法による近似

まず、 \overline{Q}_T において格子を構成する。放物型問題と同様である。

• 空間の差分幅を

$$h := \frac{b-a}{N}, \quad N \in \mathbb{N}$$

とし, [a,b] の節点を $x_i = a + ih$, i = 0,...,N とする.

• 時間の差分幅を

$$\tau := \frac{T}{r}, \quad r \in \mathbb{N}$$

とし、[0,T] の節点を $t_k = k\tau$, k = 0, ..., r とする.

 \bullet \overline{Q}_T の節点は次のように設定する:

$$\{(x_i, t_k); i = 0, \dots, N, k = 0, \dots, r\}.$$

次に、節点で関数の微分を近似する. $u \in C^4(\overline{Q}_T)$ ならば、

$$\frac{\partial^2 u}{\partial x^2}(x_i, t_k) = \frac{u(x_{i-1}, t_k) - 2u(x_i, t_k) + u(x_{i+1}, t_k)}{h^2} + \varepsilon_h^x(x_i, t_k).$$

ただし、h,i,kに依らない K が存在し、

$$|\varepsilon_h^x(x_i, t_k)| \le Kh^2 \quad \forall h \in (0, h_0).$$

同じように,

$$\frac{\partial^2 u}{\partial t^2}(x_i, t_k) = \frac{u(x_i, t_{k-1}) - 2u(x_i, t_k) + u(x_i, t_{k+1})}{\tau^2} + \varepsilon_h^t(x_i, t_k).$$

ただし、 τ , i, k に依らない K が存在し、

$$|\varepsilon_h^t(x_i, t_k)| \le K\tau^2 \quad \forall \tau \in (0, \tau_0).$$

放物型方程式のときと同様に、 $u_i^k := u(x_i, t_k)$ 、 $f_i^k := f(x_i, t_k)$ の記号を導入すると、波動方程式は次のように近似される:

$$\frac{1}{\tau^2} \left(u_i^{k+1} - 2u_i^k + u_i^{k-1} \right) = \frac{1}{h^2} \left(u_{i-1}^k - 2u_i^k + u_{i+1}^k \right) + f_i^k + \varepsilon_{h\tau}(x_i, t_k), \qquad |\varepsilon_{h\tau}(x_i, t_k)| \le K(h^2 + \tau^2).$$

誤差 $\varepsilon_{h\tau}$ を無視して、近似解を U_i^k の記号で表すと、

$$\frac{1}{\tau^2} \left(U_i^{k+1} - 2U_i^k + U_i^{k-1} \right) = \frac{1}{h^2} \left(U_{i-1}^k - 2U_i^k + U_{i+1}^k \right) + f_i^k, \qquad i = 1, \dots, N-1, \quad k = 1, \dots, r-1$$

という数値スキームが導かれる。さらに、境界条件と初期条件より以下の関係式を得る:

境界条件より
$$U_0^k = A(t_k), \ U_N^k = B(t_k), \ k = 1, ..., r$$
 初期条件より $U_i^0 = g^0(x_i), \ i = 0, ..., N.$

ここで,(k+1)-番目の時間レベルにおける近似解 U_i^{k+1} を求めるには,k-番目と (k-1)-番目のレベルの近似解の情報を用いるので,スキームをスタートさせるには U_i^0 , $i=1,\ldots,N-1$ に加えて, U_i^1 , $i=1,\ldots,N-1$ の値を計算しておく必要がある.

 U_i^1 の近似値は u_t に対する初期条件より求める.

$$u(x_i, t_1) = u(x_i, t_0) + \tau \frac{\partial}{\partial t} u(x_i, t_0) + \frac{\tau^2}{2} \frac{\partial^2}{\partial t^2} u(x_i, \tilde{t}_0), \qquad \tilde{t}_0 \in [0, \tau].$$
 (5.5)

初期条件を適用すれば,

$$u_i^1 = g^0(x_i) + \tau g^1(x_i) + \varepsilon_{\tau}^t(x_i, t_0), \qquad i = 1, \dots, N - 1, \ |\varepsilon_{\tau}^t(x_i, t_0)| \le K\tau^2, \ \tau \in (0, \tau_0)$$

となるので,

$$U_i^1 = g^0(x_i) + \tau g^1(x_i), \qquad i = 1, \dots, N-1$$

と定義できる。このようにして、陽解法のスキームを得る.

しかし, $t=t_1$ における初期条件の近似が十分に正確であるかという疑問点がある。実際, (5.5) より

$$\frac{\partial u}{\partial t}(x_i, 0) = \frac{u_i^1 - u_i^0}{\tau} - \frac{\tau}{2} \frac{\partial^2 u}{\partial t^2}(x_i, \tilde{t}_0)$$

と変形でき、最後の項は $O(\tau)$ である。これは $O(\tau^2+h^2)$ という方程式の離散化の誤差と比べて精度がおちる。そのため、 U_i^1 のより正確な近似を導く。

$$u_i^1 = u_i^0 + \tau \frac{\partial}{\partial t} u(x_i, t_0) + \frac{\tau^2}{2} \frac{\partial^2}{\partial t^2} u(x_i, t_0), +O(\tau^3)$$

のテーラー展開における 2 階微分の項は未知であるが、微分方程式を用いれば求めることができる。 $u \in C^4(\overline{Q}_T)$ の仮定を利用して (5.1) の方程式を \overline{Q}_T に拡張すれば、

$$\frac{\partial^2}{\partial t^2}u(x_i, t_0) = \frac{\partial^2}{\partial x^2}u(x_i, t_0) + f_i^0 = \frac{u_{i-1}^0 - 2u_i^0 + u_{i+1}^0}{h^2} + f_i^0 + O(h^2)$$

がわかる. これを上のテーラー展開に代入し、 $u_i^0 = g^0(x_i)$ を用いると、

$$u_i^1 = g^0(x_i) + \tau g^1(x_i) + \frac{\tau^2}{2} \left(\frac{g^0(x_{i-1}) - 2g^0(x_i) + g^0(x_{i+1})}{h^2} + f_i^0 \right) + O(h^2 \tau^2 + \tau^3)$$

というより正確な近似を得る.

次に、陰解法のスキームを導出する。節点 $(x_i,t_k)\in Q_T$ での微分方程式の近似を以下の対称な方法で行う:

$$\frac{\partial^2}{\partial x^2}u(x_i,t_k) = \frac{1}{2}\left(\frac{\partial^2}{\partial x^2}u(x_i,t_{k+1}) + \frac{\partial^2}{\partial x^2}u(x_i,t_{k-1})\right) + O(\tau^2).$$

ただし、右辺の微分はいつもの近似を施す:

$$\frac{\partial^2}{\partial x^2} u(x_i, t_{k\pm 1}) = \frac{u_{i-1}^{k\pm 1} - 2u_i^{k\pm 1} + u_{i+1}^{k\pm 1}}{h^2} + O(h^2).$$

すると,

$$\frac{1}{\tau^2} \left(U_i^{k+1} - 2U_i^k + U_i^{k-1} \right) = \frac{1}{2h^2} \left(U_{i-1}^{k+1} - 2U_i^{k+1} + U_{i+1}^{k+1} + U_{i-1}^{k-1} - 2U_i^{k-1} + U_{i+1}^{k-1} \right) + f_i^k,
i = 1, \dots, N-1$$
(5.6)

に初期条件と境界条件の情報をつけた連立方程式を解くことになる.

問題

問題 5.1 膜の振動方程式

$$\frac{\partial^2 u}{\partial t^2} = \Delta u + f$$
 in $(0,1)^2 \times (0,T)$

対称な正定値行列

$$A_h := \frac{1}{h^2} \begin{pmatrix} 2 & -1 & 0 & 0 & \dots \\ -1 & 2 & -1 & 0 & \dots \\ 0 & -1 & 2 & -1 & \dots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix}$$

を用いて、差分法の式を書き直す。 $U_h^k:=(U_1^k,\dots,U_{N-1}^k)^T\in\mathbb{R}^{n_h},\ n_h=N-1$ を k-番目の時間 レベルでの近似解の節点における値を並べたベクトルとすると、スキームは

陽解法
$$\frac{1}{\tau^2} \left(\boldsymbol{U}_h^{k+1} - 2\boldsymbol{U}_h^k + \boldsymbol{U}_h^{k-1} \right) + A_h \boldsymbol{U}_h^k = \boldsymbol{F}_h^k, \quad k = 1, \dots, r-1$$
 陰解法
$$\frac{1}{\tau^2} \left(\boldsymbol{U}_h^{k+1} - 2\boldsymbol{U}_h^k + \boldsymbol{U}_h^{k-1} \right) + \frac{1}{2} A_h \left(\boldsymbol{U}_h^{k+1} + \boldsymbol{U}_h^{k-1} \right) = \boldsymbol{F}_h^k, \quad k = 1, \dots, r-1$$

のように書ける。ただし, U_h^0, U_h^1 は初期条件より既知で, $F_h^k \in \mathbb{R}^{n_h}$ は f_i^k の値と境界条件による値で定まるベクトルである.

5.3 差分スキームの性質

このあとの解析では, A_h を一般の対称な正定値行列と考える.陰解法の係数行列は $\frac{1}{\tau^2}I+\frac{1}{2}A_h$ となり,これは対称で正定値,よって正則な行列なので,連立方程式はただ一つの解をもつ.

放物型問題のように、双曲型問題でも安定性が重要な役割を担う。安定性は一般的に、ある時間レベルで発生した誤差の大きさが先の時間レベルに進んでも有界にとどまるというスキームの性質のことを指す。

陽解法の場合、正確に計算した近似解は

$$\frac{1}{\tau^2} \left(\boldsymbol{U}_h^{k+1} - 2\boldsymbol{U}_h^k + \boldsymbol{U}_h^{k-1} \right) + A_h \boldsymbol{U}_h^k = \boldsymbol{F}_h^k, \quad k = 1, \dots, r-1$$

と初期条件 $m{U}_h^0 = m{g}_h^0, m{U}_h^1 = m{g}_h^1$ を満たす.一方では,不正確な(摂動を加えた)初期値

$$oldsymbol{ ilde{U}}_h^0 = oldsymbol{g}_h^0 + oldsymbol{artheta}_h^0, \quad oldsymbol{ ilde{U}}_h^1 = oldsymbol{g}_h^1 + oldsymbol{artheta}_h^1$$

からはじめた近似解を $ilde{m U}_h^k$ とする。この二つの近似解の差を $m Z_h^k:= ilde{m U}_h^k-m U_h^k$ とおくと, $m Z_h^k$ は次を満たす:

$$\frac{1}{\tau^2} \left(\mathbf{Z}_h^{k+1} - 2\mathbf{Z}_h^k + \mathbf{Z}_h^{k-1} \right) + A_h \mathbf{Z}_h^k = \mathbf{0}, \quad k = 1, 2, \dots, \qquad \mathbf{Z}_h^0 = \boldsymbol{\vartheta}_h^0, \quad \mathbf{Z}_h^1 = \boldsymbol{\vartheta}_h^1. \tag{5.7}$$

スキームが安定であるとは、 $\{oldsymbol{Z}_h^k\}_{k=1,2,\dots}$ が有界であるということを意味する。そこで、k が増え

たときの \mathbf{Z}_h^k の振る舞いを調べればよい.

この安定性をフーリエの方法を用いて調べる。行列 A_h は対称,正定値なので, A_h の固有ベクトルからなら \mathbb{R}^{n_h} の正規直交基底 $\{z_1,\ldots,z_{n=n_h}\}$ が存在する。つまり, $\{\lambda_1,\ldots,\lambda_n\}=Sp(A_h)$ に対し,

$$A_h \boldsymbol{z}_j = \lambda_j \boldsymbol{z}_j, \qquad j = 1, \dots, n.$$

上で定義したベクトル Z_h^k は \mathbb{R}^n のベクトルであるから,

$$oldsymbol{Z}_h^k = \sum_{j=1}^n lpha_j^{(k)} oldsymbol{z}_j$$

と書ける。ただし、 $\alpha_j^{(k)}={m Z}_h^k\cdot{m z}_j$ (\mathbb{R}^n の内積)はフーリエ係数と呼ばれる。これを (5.7) に代入すれば、

$$\frac{1}{\tau^2} \sum_{j=1}^n \left(\alpha_j^{(k+1)} - 2\alpha_j^{(k)} + \alpha_j^{(k-1)} \right) \mathbf{z}_j + A_h \left(\sum_{j=1}^n \alpha_j^{(k)} \mathbf{z}_j \right) = \mathbf{0}$$

$$\sum_{j=1}^n \left(\frac{\alpha_j^{(k+1)} - 2\alpha_j^{(k)} + \alpha_j^{(k-1)}}{\tau^2} + \lambda_j \alpha_j^{(k)} \right) \mathbf{z}_j = \mathbf{0}$$

という関係を得るが、ベクトル z_i が独立なため、

$$\frac{\alpha_j^{(k+1)} - 2\alpha_j^{(k)} + \alpha_j^{(k-1)}}{\tau^2} + \lambda_j \alpha_j^{(k)} = 0 \qquad \forall j = 1, \dots, n$$

が従う. なお, $\alpha_j^{(0)}, \alpha_j^{(1)}$ は初期誤差 $\pmb{\vartheta}_h^0, \pmb{\vartheta}_h^1$ により与えられている.

これは定数係数の線形漸化式なので、具体的に解ける。j を固定し、 $\sigma_j := \frac{\lambda_j \tau^2}{2}$ とおくと、漸化式は

$$\alpha_j^{(k+1)} - 2(1 - \sigma_j)\alpha_j^{(k)} + \alpha_j^{(k-1)} = 0, \qquad k = 1, 2, \dots$$

となる. 特性方程式 $\eta_j^2 - 2(1 - \sigma_j)\eta_j + 1 = 0$ の根は

$$\eta_j^{\pm} = 1 - \sigma_j \pm \sqrt{\sigma_j^2 - 2\sigma_j}$$

であるので、三つの場合にわけて考える.

(a) $\sigma_j > 2$: このとき, $\eta_j^- < -1$, そして $-1 < \eta_j^+ < 0$ であるので,

$$\alpha_j^{(k)} = A_j(\eta_j^-)^k + B_j(\eta_j^+)^k, \qquad k = 0, 1, \dots$$

 $(A_i, B_i$ は初期条件より決まる.)

(b) $\sigma_j = 2$:このとき、 $\eta_j^{\pm} = -1$ なので、

$$\alpha_j^{(k)} = A_j(-1)^k + B_j k(-1)^k, \qquad k = 0, 1, \dots$$

(c) $\sigma_j \in (0,2)$: このとき $\eta_j^{\pm} = 1 - \sigma_j \pm i \sqrt{\sigma_j(2-\sigma_j)}$ となり, $|\eta_j^{\pm}| = 1$ と確認できる.

$$\alpha_i^{(k)} = A_j(\eta_i^-)^k + B_j(\eta_i^+)^k, \qquad k = 0, 1, \dots$$

目標は $\{\mathbf{Z}_h^k\}_k$ の有界性を示すことであるが、これは $\{\alpha_j^{(k)}\}_k, j=1,\dots,n$ の有界性と同値である.

定義 5.1

(a) $j=1,\ldots,n$ が存在し、 $\sigma_j>2$ となるとき、スキームは 不安定 であるという。(これは $\tau>2/\sqrt{\lambda_j}$ を満たす j があることと同値である。)

注:上の計算では $|\eta_j^-|>1$ であるから, $|\alpha_j^{(k)}|\to\infty$ as $k\to\infty$ となる.もし A_j がゼロであっても,数値計算の丸め誤差などによって η_j^- の影響が現れてしまう.

(b) $\sigma_i \in (0,2]$ のとき、スキームは 弱安定 であるという。

注: $\sigma_j=2$ のとき, $\eta_j^{\pm}=-1$ となり, $|\alpha_j^{(k)}|$ は k について線形的に増加する.弱安定性ではこのような緩い増加を許すという意味である.

(c) $\sigma_j \in (0,2)$ のとき(つまり, $\tau < 2/\sqrt{\lambda_j} \ \forall j=1,\ldots,n$,または $\tau < 2/\sqrt{\rho(A_h)}$ のとき),スキームは 安定 であるという.

注:このとき、 $|lpha_j^{(k)}| \leq |A_j| + |B_j| \, orall k = 0, 1, \ldots$

注. $\rho(A_h) \leq ||A_h||_{\infty}$ が成り立つので,

$$0 < \tau \le \frac{2}{\sqrt{\|A_h\|_{\infty}}}$$

が(弱)安定性のための十分条件である.

例題 (5.1) の場合, $\|A_h\|_{\infty} = \frac{4}{h^2}$ を得るので,

$$0 < \tau \le h$$

という CFL 条件 (Courant-Friedrichs-Lewy condition) になる.

陰解法のスキームについて簡単に述べる。 初期誤差 $m{Z}_h^0 = m{artheta}_h^0, m{Z}_h^1 = m{artheta}_h^1$ に対して、

$$\frac{1}{\tau^2} \left(\boldsymbol{Z}_h^{k+1} - 2 \boldsymbol{Z}_h^k + \boldsymbol{Z}_h^{k-1} \right) + \frac{1}{2} A_h \left(\boldsymbol{Z}_h^{k+1} + \boldsymbol{Z}_h^{k-1} \right) = \boldsymbol{0}, \quad k = 1, 2, \dots$$

を解く.フーリエ係数 $\alpha_{j}^{(k)}$ に対する方程式は

$$\frac{\alpha_j^{(k+1)} - 2\alpha_j^{(k)} + \alpha_j^{(k-1)}}{\tau^2} + \frac{1}{2}\lambda_j \left(\alpha_j^{(k+1)} + \alpha_j^{(k-1)}\right) = 0 \qquad k = 1, 2, \dots$$
$$(1 + \sigma_j)\alpha_j^{(k+1)} - 2\alpha_j^{(k)} + (1 + \sigma_j)\alpha_j^{(k-1)} = 0$$

となり、漸化式の特性方程式 $(1+\sigma_j)\eta_j^2-2\eta_j+(1+\sigma_j)=0$ の根は

$$\eta_j^{\pm} = \frac{1 \pm i \sqrt{\sigma_j(\sigma_j + 2)}}{1 + \sigma_j}$$

という形になる。 $|\eta_i^{\pm}|=1$ に注意すると,

$$|\alpha_j^{(k)}| = |A_j(\eta_j^-)^k + B_j(\eta_j^+)^k| \le |A_j| + |B_j|$$

のように有界である. よって、 $\{ m{Z}_h^k \}_{k=0,1,\dots}$ は au の値によらず有界にとどまり、スキームは 無条件安定 となる.

問題

問題 5.2 フーリエの方法を用いて、放物型問題に対する差分スキームの安定性を調べよ.

5.4 1階双曲型方程式

本節で考える問題の保存形を書くと、次のようになる:

$$\frac{\partial \boldsymbol{w}}{\partial t} + \sum_{j=1}^{N} \frac{\partial \boldsymbol{F}_{j}(\boldsymbol{w})}{\partial x_{j}} = 0 \quad \text{in } Q = \mathbb{R}^{N} \times (0, +\infty)$$
 (5.8)

$$\boldsymbol{w}(\boldsymbol{x},0) = g(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^N.$$
 (5.9)

ただし、 $\mathbf{w}: Q \to \mathbb{R}^m$ と考えるので、連立方程式である。また、 $\mathbf{F}_j: \mathbb{R}^m \to \mathbb{R}^m$ の成分を $\mathbf{F}_j = (F_{j1}, \dots, F_{jm}), j = 1, \dots, N$ と書くことにすると、 $F_{ji}: \mathbb{R}^m \to \mathbb{R}, i = 1, \dots, m, j = 1, \dots, N$ となる。

 $F_{ji} \in C^1(\mathbb{R}^m)$ のとき,

$$A_j(\boldsymbol{w}) = \frac{D\boldsymbol{F}_j(\boldsymbol{w})}{D\boldsymbol{w}}$$

で写像 F_i のヤコビ行列を定義する。そのとき, $\mathbf{w} \in [C^1(Q)]^m$ が上の微分方程式系の解ならば,

$$\frac{\partial \boldsymbol{w}}{\partial t} + \sum_{j=1}^{N} A_j(\boldsymbol{w}) \frac{\partial \boldsymbol{w}}{\partial x_j} = 0$$
 in Q

という 非保存形 の形に書くことができる.

定義 5.2 偏微分方程式系 (5.8) が 双曲型 であるとは、行列

$$P(\boldsymbol{w}, \boldsymbol{n}) := \sum_{j=1}^{N} n_j A_j(\boldsymbol{w}), \qquad \boldsymbol{n} \in \mathbb{R}^N, \ \boldsymbol{w} \in \mathbb{R}^m$$

が対角化可能ということである. すなわち,正則な行列 T(w) と対角行列 D(w,n) が存在し,

$$P(\boldsymbol{w},\boldsymbol{n}) = T(\boldsymbol{w})D(\boldsymbol{w},\boldsymbol{n})T^{-1}(\boldsymbol{w}), \qquad D(\boldsymbol{w},\boldsymbol{n}) = \begin{pmatrix} \lambda_1(\boldsymbol{w},\boldsymbol{n}) & & & \\ & \ddots & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & \lambda_m(\boldsymbol{w},\boldsymbol{n}) \end{pmatrix}$$

と書ける. ただし、 $\lambda_i(\boldsymbol{w},\boldsymbol{n}), i=1,\ldots,m$ は $P(\boldsymbol{w},\boldsymbol{n})$ の固有値で、すべて実数である.

例 5.3

(1) N = 1, m = 1 のとき,以下の方程式になる:

$$\frac{\partial w}{\partial t} + \frac{\partial F(w)}{\partial x} = 0 \qquad \text{in } Q = \mathbb{R} \times (0, +\infty).$$

(2) 線形な連立方程式の場合は、 $m \times m$ -行列 A_j に対して $F_j(\boldsymbol{w}) = A_j \boldsymbol{w}$ なので、

$$\frac{\partial \mathbf{w}}{\partial t} + \sum_{j=1}^{N} A_j \frac{\partial \mathbf{w}}{\partial x_j} = 0.$$

(3) 最も簡単な例は空間 1 次元における一本の線形な方程式である:

$$\frac{\partial w}{\partial t} + a \frac{w}{\partial x} = 0.$$

このような双曲型方程式系は流体が流れる現象のモデルとして現れることが多い。例えば、保存則であるオイラー方程式は非線形な双曲型方程式に分類される。また、水文学における多孔質内の流体の浸透を表すモデルも双曲型である。他に、交通分野では、高速道路の車の流れは非線形な双曲型方程式を用いて記述される。

このようなときに、 F_i は量wの x_i -方向の流束(flux)を意味する。すると、

$$\mathcal{P}(\boldsymbol{w},\boldsymbol{n}) := \sum_{j=1}^{N} F_j(\boldsymbol{w}) n_j$$

はベクトルnで示される方向に沿った流束を表す。上で導入した行列Pは

$$P(\boldsymbol{w}, \boldsymbol{n}) = \frac{D\mathcal{P}(\boldsymbol{w}, \boldsymbol{n})}{D\boldsymbol{w}}$$

を満たす.

5.4.1 空間1次元の1階双曲型方程式の数値解法

次の問題を考える:

$$\frac{\partial \mathbf{w}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{w})}{\partial x} = 0 \quad \text{in } Q = \mathbb{R} \times (0, +\infty)$$
 (5.10)

$$\boldsymbol{w}(x,0) = \boldsymbol{g}(x), \qquad x \in \mathbb{R}. \tag{5.11}$$

この問題のデータ(すなわち、関数g)が滑らかであっても、解の正則性は言うまでもなく解の存在すら保証されないことに注意しよう.

そこで、超関数の意味での弱解の概念を導入する。そのため、試験関数を関数空間 $[C_0^\infty(\mathbb{R}\times [0,\infty))]^m$ からとる。これは、無限回微分可能な関数 φ で台 $K_\varphi\subset\mathbb{R}\times [0,\infty)$ がコンパクト集合であるような関数の空間である。解 w が C^1 -関数であると仮定し、試験関数 $\varphi\in [C_0^\infty(\mathbb{R}\times [0,\infty))]^m$ を微分方程式にかけて、Q で積分し、部分積分を施すと

$$0 = \int_{Q} \left(\frac{\partial \boldsymbol{w}}{\partial t} + \frac{\partial \boldsymbol{F}(\boldsymbol{w})}{\partial x} \right) \cdot \boldsymbol{\varphi} \, dx \, dt = -\int_{Q} \left(\boldsymbol{w} \cdot \frac{\partial \boldsymbol{\varphi}}{\partial t} + \boldsymbol{F}(\boldsymbol{w}) \cdot \frac{\partial \boldsymbol{\varphi}}{\partial x} \right) \, dx \, dt - \int_{\mathbb{R}} \boldsymbol{w}(x,0) \cdot \boldsymbol{\varphi}(x,0) \, dx$$
 を得る.

定義 5.4 $g \in [L^{\infty}(\mathbb{R})]^m$ に対し、次の条件を満たす関数 $\mathbf{w} \in [L^{\infty}(Q)]^m$ を問題 (5.10)–(5.11) の弱解とよぶ:

$$\int_{Q} \left(\boldsymbol{w} \cdot \frac{\partial \boldsymbol{\varphi}}{\partial t} + \boldsymbol{F}(\boldsymbol{w}) \cdot \frac{\partial \boldsymbol{\varphi}}{\partial x} \right) dx dt + \int_{\mathbb{R}} \boldsymbol{g}(x) \cdot \boldsymbol{\varphi}(x,0) dx = 0 \qquad \forall \boldsymbol{\varphi} \in [C_0^{\infty}(\mathbb{R} \times [0,\infty))]^m.$$

問題 (5.10)-(5.11) を差分法で離散化する.時間の差分幅 $\tau>0$ と空間の差分幅 h>0 を決めて, \overline{Q} の格子点を

$$x_j = jh, \ j \in \mathbb{Z}, \qquad t_k = k\tau, \ k \in \mathbb{Z}_0^+$$

と定義する。スキームでは、二つの節点の中点も用いるので、その記号を $x_{j+\frac{1}{2}}:=x_j+\frac{h}{2}$ とする。

さて、 $\mathbf{w}\in[C^1(\overline{Q})]^m$ が古典解であるとして、微分方程式 (5.10) を集合 $(x_{j-\frac{1}{2}},x_{j+\frac{1}{2}})\times(t_k,t_{k+1})$ の上で積分すれば、

$$\begin{aligned} \mathbf{0} &= \int_{t_k}^{t_{k+1}} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \left(\frac{\partial \boldsymbol{w}}{\partial t} + \frac{\partial \boldsymbol{F}(\boldsymbol{w})}{\partial x} \right) \, dx \, dt &= \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \left(\boldsymbol{w}(x,t_{k+1}) - \boldsymbol{w}(x,t_k) \right) \, dx \\ &+ \int_{t_k}^{t_{k+1}} \left(\boldsymbol{F}(\boldsymbol{w}(x_{j+\frac{1}{2}},t)) - \boldsymbol{F}(\boldsymbol{w}(x_{j-\frac{1}{2}})) \right) \, dt \end{aligned}$$

を得る.節点 (x_j,t_k) における近似解を $oldsymbol{w}_j^k pprox oldsymbol{w}(x_j,t_k)$ とおく.右辺の第1項を

$$h\left(\boldsymbol{w}(x_j,t_{k+1})-\boldsymbol{w}(x_j,t_k)\right)$$

で近似し、右辺の第 2 項において、 $x_{j+\frac{1}{2}}$ が節点ではないので、数値流束 f という関数を用いて、次のように近似する:

$$F(\boldsymbol{w}(x_{j+\frac{1}{2}},t_k)) \approx f(\boldsymbol{w}(x_j,t_k),\boldsymbol{w}(x_{j+1},t_k)).$$

したがって,

$$h(\boldsymbol{w}_{i}^{k+1} - \boldsymbol{w}_{i}^{k}) + \tau(\boldsymbol{f}(\boldsymbol{w}_{i}^{k}, \boldsymbol{w}_{i+1}^{k}) - \boldsymbol{f}(\boldsymbol{w}_{i-1}^{k}, \boldsymbol{w}_{i}^{k})) = \boldsymbol{0}$$

というスキームとなる. これを書き換えると,

$$oldsymbol{w}_j^{k+1} = oldsymbol{w}_j^k - rac{ au}{h} \left(oldsymbol{f}(oldsymbol{w}_j^k, oldsymbol{w}_{j+1}^k) - oldsymbol{f}(oldsymbol{w}_{j-1}^k, oldsymbol{w}_j^k)
ight), \qquad j \in \mathbb{Z}, \; k \in \mathbb{Z}_0^+$$

となり、さらに $oldsymbol{f}_{j+rac{1}{2}}^k := oldsymbol{f}(oldsymbol{w}_j^k, oldsymbol{w}_{j+1}^k)$ という記号を導入すれば、

$$oldsymbol{w}_j^{k+1} = oldsymbol{w}_j^k - rac{ au}{h} \left(oldsymbol{f}_{j+rac{1}{2}}^k - oldsymbol{f}_{j-rac{1}{2}}^k
ight)$$

というコンパクトな形で表される。初期条件は $w_j^0=g(x_j)$ とすればよいが, $g\in [L^\infty(\mathbb{R})]^m$ としか仮定していないため,g が各点で定義されない可能性がある。そこで,

$$\mathbf{w}_{j}^{0} = \frac{1}{h} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \mathbf{g}(x) dx$$
 (5.12)

とおくことにする.

以上により、陽解法のスキームを導いた。fの上の添え字をkからk+1に変更すれば、陰解法スキームが得られるが、無限個の非線形な連立方程式を解くことになる。

注. 数値流束の関数 $f: \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ については、連続関数で適合であると仮定する。すなわち、

$$f(w, w) = F(w)$$
 $\forall w \in \mathbb{R}^m$.

例 5.5 以下で使う例題を一つ紹介する. N=m=1 として, F を線形関数とすると,

$$\frac{\partial w}{\partial t} + a \frac{\partial w}{\partial x} = 0 \quad in \ \mathbb{R} \times (0, \infty)$$
 (5.13)

$$w(x,0) = g(x) x \in \mathbb{R} (5.14)$$

という問題になる。 $g \in L^{\infty}(\mathbb{R})$ ならば、弱解がただ一つ存在することが証明できる。g が区分的に滑らかな関数のとき、この弱解を特性曲線法を用いて求めることができる。

数値スキームとして、次の3通りのスキームが考えられる.

(a)
$$\frac{w_j^{k+1} - w_j^k}{\tau} + a \frac{w_{j+1}^k - w_{j-1}^k}{2h} = 0$$

(b)
$$\frac{w_j^{k+1} - w_j^k}{\tau} + a \frac{w_{j+1}^k - w_j^k}{h} = 0$$

(c)
$$\frac{w_j^{k+1} - w_j^k}{\tau} + a \frac{w_j^k - w_{j-1}^k}{h} = 0$$

これらのスキームに対応する数値流束は

(a)
$$f(u,v) = a\frac{u+v}{2}$$
, (b) $f(u,v) = av$, (c) $f(u,v) = au$

と表現できる。上記のスキームがすべて適切なわけではない。実際、a が情報が伝播する速度を意味するので、 $\frac{\partial w}{\partial x}$ を近似するときに、a の方向をさかのぼったところにおける値を使用することが自然である。つまり、a>0 であれば、(c) のスキームのみ適切な近似である。

次に収束性など差分スキームの性質を調べる。まず、スキームが正しい近似解を与えること、すなわち、近似解が弱解に収束することを確かめる。そのため、すべての $(x,t) \in \mathbb{R} \times [0,\infty)$ に対して定義される近似解を定義する必要がある。差分幅 $h,\tau>0$ に対して,関数 $\boldsymbol{w}_{h\tau}(x,t)$ を次のように定める:

$$\mathbf{w}_{h\tau}(x,t) = \mathbf{w}_{j}^{k}$$
 for $x \in [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}), t \in [t_{k}, t_{k+1}), j \in \mathbb{Z}, k \in \mathbb{Z}_{0}^{+}$.

ただし、以下では初期値の近似に(5.12)の方法が適用されるとする.

定理 **5.6** (Lax-Wendroff の定理) 差分スキームが適合性をもつ(すなわち、数値流束 f が適合である)とする。また、定数 C>0 が存在し、

$$\|\boldsymbol{w}_{h\tau}\|_{L^{\infty}(Q)} \le C \qquad \forall h, \tau > 0$$

が成り立つと仮定する。 さらに、関数 w が存在し、

$$\boldsymbol{w}_{h\tau} \to \boldsymbol{w}$$
 a.e. in Q as $h, \tau \to 0$

が満たされるとする。そのとき、wは問題()の弱解である。

問題

問題 **5.3** 例5.5のスキーム (a)-(c) の空間上と時間上の近似オーダーを求めよ。つまり、w, f が滑らかな関数であると仮定した上で、定数 $M, h_0, \tau_0 > 0$ が存在し次の関係式が成り立つような最大の p, q を求めればよい:

$$|\varepsilon_j^k| \le M(h^p + \tau^q) \qquad \forall j \in \mathbb{Z} \ \forall k = 0, 1, 2, \dots \ \forall h \in (0, h_0) \ \forall \tau \in (0, \tau_0).$$

ただし、 ε_{j}^{k} は離散化による相対誤差で、次で定義される:

$$w(x_j, t_{k+1}) = w(x_j, t_k) - \frac{\tau}{h} \Big(f(w(x_j, t_k), w(x_{j+1}, t_k)) - f(w(x_{j-1}, t_k), w(x_j, t_k)) \Big) + \tau \varepsilon_j^k.$$

スキームの安定性

数値スキームの安定性は、「初期値 $\{w_j^0\}_{j\in\mathbb{Z}}$ が有界であれば、すべての時刻 $k=1,2,\ldots$ での近似解 $\{w_j^k\}_{j\in\mathbb{Z}}$ も有界である」という概念である。ここでは、基本的には線形な方程式にしか適用できないフォン・ノイマンの安定性判定法を紹介する。

定義 5.7 近似解をフーリエ級数に展開する:

$$\mathbf{w}_{j}^{k} = \sum_{n=-\infty}^{\infty} \mathbf{F}_{k}(n)e^{injh} = \sum_{n=-\infty}^{\infty} \mathbf{F}_{k}(n)e^{inx}, \quad j \in \mathbb{Z}, \quad k = 0, 1, 2, \dots$$

このとき、スキームがフォン・ノイマンによる安定性条件を満たすとは、定数C > 0が存在し、

$$\|F_k(n)\| \le C \|F_0(n)\| \quad \forall n \in \mathbb{Z} \ \forall k = 0, 1, 2, \dots$$

が成立するということである。

例 5.8 上に紹介したスキーム (c) のフォン・ノイマン安定性を調べるには、近似解のフーリエ展開を近似式に代入する:

$$\sum_{n=-\infty}^{\infty} F_{k+1}(n)e^{inx} = \sum_{n=-\infty}^{\infty} F_k(n)e^{inx} - \frac{\tau a}{h} \left(\sum_{n=-\infty}^{\infty} F_k(n)e^{inx} - \sum_{n=-\infty}^{\infty} F_k(n)e^{in(x-h)} \right)$$
$$= \sum_{n=-\infty}^{\infty} F_k(n)e^{inx} \left(1 - \frac{\tau a}{h} + \frac{\tau a}{h}e^{-inh} \right)$$

関数 $\{e^{inx}\}_{n=-\infty}^{\infty}$ が線形独立であるから,

$$F_{k+1}(n) = F_k(n) \left(1 - \frac{\tau a}{h} + \frac{\tau a}{h} e^{-inh} \right) \qquad n \in \mathbb{Z}$$

が得られる. そこで,

$$G = G(n, h, \tau) = 1 - \frac{\tau a}{h} + \frac{\tau a}{h} e^{-inh}$$

と置いて, G を 増幅定数 (amplification factor) とよぶ.

 $|F_{k+1}(n)| = |G| |F_k(n)|$ より、フォン・ノイマンの安定性条件が満たされるのは $|G| \le 1$ のと

きとそのときのみである。よって、|G|を計算すればよい:

$$|G|^{2} = \left|1 - \frac{\tau a}{h} (1 - \cos nh + i \sin nh)\right|^{2}$$

$$= \left(1 - \frac{\tau a}{h} (1 - \cos nh)\right)^{2} + \left(\frac{\tau a}{h} \sin nh\right)^{2}$$

$$= 1 - 2(1 - \cos nh)\frac{\tau a}{h} \left(1 - \frac{\tau a}{h}\right).$$

つまり、 $|G| \le 1$ は次の条件と同値である:

$$-2(1-\cos nh)\frac{\tau a}{h}\left(1-\frac{\tau a}{h}\right) \le 0.$$

係数 a の正負によって結論が異なる:

a > 0 のとき、上の条件は

$$\tau \leq \frac{h}{a}$$
 CFL 安定性条件

と同値である.

• a < 0 のとき、上の条件はどんな $\tau > 0$ をとっても成立しない。このとき数値解法は 無条件不安定 であると言う。

この結果は前に行った物理的な考察と合致する。a の符号は情報が伝わる方向を示すため,a の値によって差分の仕方を変えるべきである。a>0 のときは情報が正の方向(左から右へ)に 伝播するので,後退差分(スキーム(c))を用いる。同様に,a<0 のときは前進差分(スキーム(b))を用いる。ここでわかったことをまとめると,

$$a>0$$
 のときスキーム (c) を用い, $0<\tau\leq \frac{h}{a}$ を満たす時間ステップを選ぶ $a<0$ のときスキーム (b) を用い, $0<\tau\leq \frac{h}{|a|}$ を満たす時間ステップを選ぶ

微分の近似のしかたを考える点での速度の向きによって調節する方法は 風上差分 と呼ばれる.

問題

問題 5.4 スキーム (a) のフォン・ノイマン安定性を調べ、すべての a に対し無条件不安定であることを示せ.

a < oときに使われるスキーム (b) の数値流束は f(u,v) = av であり,a > 0 のときに使われるスキーム (c) の数値流束は f(u,v) = au であるので,それらをまとめて一つの流束の定義式で書くと,

$$f(u,v) = a^+ u + a^- v$$

となる. ただし, $a^+ = \max\{0, a\}$, $a^- = \min\{a, 0\}$ とした. この式は任意の a に対し使用でき,

非線形方程式に対する数値流束の一般化の出発点になる.

その一般化を具体的に説明するためには、上の式を次のように解釈する:

$$f(u,v) = \tilde{F}^+(u) + \tilde{F}^-(v)$$
 ただし, $(\tilde{F}^+)' = a^+ \ge 0$ および $(\tilde{F}^-)' = a^- \le 0$.

これよりはm=1の場合(一つの微分方程式)についてのみ考える。非線形問題の数値流束を設計するとき、consistent でなければならない(つまり、f(w,w)=F(w))という条件を考慮しながら、流束を次のように正の部分と負の部分に分ける:

$$F(w) = \tilde{F}^+(w) + \tilde{F}^-(w)$$
 with $(\tilde{F}^+)' \ge 0$ and $(\tilde{F}^-)' \le 0$.

そこで,数値流束を次のように定義する:

$$f(u,v) = \tilde{F}^{+}(u) + \tilde{F}^{-}(v). \tag{5.15}$$

Engquist-Osher の方法 流束を (5.15) に従って分解する. ただし,

$$\tilde{F}^{+}(w) = \frac{1}{2}F(0) + \int_{0}^{w} (F')^{+}(q) dq$$

$$\tilde{F}^{-}(w) = \frac{1}{2}F(0) + \int_{0}^{w} (F')^{-}(q) dq.$$

この式を (5.15) に代入すると,

$$f(u,v) = \frac{1}{2} \left\{ F(u) + F(v) - \int_{u}^{v} |F'(q)| \, dq \right\},\tag{5.16}$$

が得られる. ここで, $|a| = a^+ - a^-$.

問題

問題 5.5 Engquist-Osher のスキームの数値流束が consistent であることを確かめ, $F(u) = \frac{1}{2}u^2$ のときの数値流束を具体的に求めよ

フォン・ノイマンの安定性解析は各節点 (x_i, t_k) において線形化した方程式

$$w_t + F'(w_j^k)w_x = 0$$

に適用される. 数値流束を

$$f_{EO}(u,v) = F'(w_j^k)^+ u + F'(w_j^k)^- v$$

と置くと、このスキームが

$$0 < \tau \le \frac{h}{|F'(w_j^k)|} \qquad \forall j \in \mathbb{Z} \ \forall k = 0, 1, 2, \dots$$

$$(5.17)$$

の条件のもとで安定であることがわかる。しかし、この条件は実用性がないため、実際には各時刻において異なる時間ステップ $\tau_k = t_{k+1} - t_k$ を選び、 τ_k がすべての $k = 0, 1, 2, \ldots$ に対し次の条件を満たすようにする:

$$0 < \tau_k \le \frac{h}{|F'(w_j^k)|} \qquad \forall j \in \mathbb{Z}$$

ここで、 w_i^0 は全ての j に対して初期条件から既知である.

Lax-Friedrichs の方法 中心差分を用いたスキームは

$$w_j^{k+1} = w_j^k - \frac{\tau}{2h} \left(F(w_{j+1}^k) - F(w_{j-1}^k) \right)$$

となるが、このスキームは無条件不安定である。一方では、上の式を

$$w_j^{k+1} = \frac{1}{2}(w_{j-1}^k + w_{j+1}^k) - \frac{\tau}{2h} \left(F(w_{j+1}^k) - F(w_{j-1}^k) \right)$$

と変更すれば、 τ が条件 (5.17) を満たすときに安定なスキームになる。このスキームは Lax-Friedrichs スキームとよばれ、Engquist-Osher スキームより拡散の効果が強い(すなわち、不連続性がよりなだらかになってしまう)。

5.4.2 オイラー方程式の数値計算

非粘性流体の体積力なしの断熱運動は未知変数 $\rho, v_1, v_2, v_3, p, E$ に対する オイラー方程式系 で記述される:

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) = 0 \tag{5.18}$$

$$\frac{\partial \rho v_i}{\partial t} + \operatorname{div}(\rho v_i \boldsymbol{v}) = -\frac{\partial p}{\partial x_i}, \qquad i = 1, 2, 3$$
(5.19)

$$\frac{\partial \rho E}{\partial t} + \operatorname{div}(\rho E \boldsymbol{v}) = -\operatorname{div}(p \boldsymbol{v}) \tag{5.20}$$

ここで, (5.18) は質量保存, (5.19) は運動量保存, (5.20) はエネルギー保存を表している. これに 状態方程式

$$p = (\kappa - 1)\rho(E - \frac{1}{2}|\boldsymbol{v}|^2)$$

が加えられる。ここで、 $v=(v_1,v_2,v_3)$ は速度、 ρ は密度、p は圧力、E はエネルギー密度、そして κ はポアソン定数である。

 $e = \rho E$ とおいて、上の連立方程式を次のように書き換えることができる:

$$\frac{\partial}{\partial t} \begin{pmatrix} \rho \\ \rho v_1 \\ \rho v_2 \\ \rho v_3 \\ e \end{pmatrix} + \frac{\partial}{\partial x_1} \begin{pmatrix} \rho v_1 \\ \rho v_1^2 + p \\ \rho v_1 v_2 \\ \rho v_1 v_3 \\ (e+p)v_1 \end{pmatrix} + \frac{\partial}{\partial x_2} \begin{pmatrix} \rho v_2 \\ \rho v_1 v_2 \\ \rho v_2^2 + p \\ \rho v_2 v_3 \\ (e+p)v_2 \end{pmatrix} + \frac{\partial}{\partial x_3} \begin{pmatrix} \rho v_3 \\ \rho v_1 v_3 \\ \rho v_2 v_3 \\ \rho v_3^2 + p \\ (e+p)v_3 \end{pmatrix} = 0.$$

また, 状態方程式は

$$p = (\kappa - 1)(e - \frac{1}{2}\rho|\boldsymbol{v}|^2).$$

となる。未知変数のベクトル $\mathbf{w}=(\rho,\rho v_1,\rho v_2,\rho v_3,e)^T$ を導入すると、さらに以下のように書ける:

$$\frac{\partial}{\partial t} \boldsymbol{w} + \frac{\partial}{\partial x_1} \boldsymbol{F}_1(\boldsymbol{w}) + \frac{\partial}{\partial x_2} \boldsymbol{F}_2(\boldsymbol{w}) + \frac{\partial}{\partial x_3} \boldsymbol{F}_3(\boldsymbol{w}) = 0.$$

ただし、非粘性流束(inviscid fluxes) とよばれる関数 $m{f}_i$ はベクトル $m{w}$ の成分を用いて表される。 例えば、

$$F_1(\mathbf{w}) = \left(w_2, \frac{w_2^2}{w_1} + (\kappa - 1)\left(w_5 - \frac{w_2^2 + w_3^2 + w_4^2}{2w_1}\right), \dots\right)^T.$$

これにより、オイラー方程式は

$$\frac{\partial \mathbf{w}}{\partial t} + \sum_{i=1}^{3} \frac{\partial \mathbf{F}_i}{\partial x_i}(\mathbf{w}) = \mathbf{0}$$
 (5.21)

というわかりやすい形で表される。これに適切な初期条件と境界条件がつく、

オイラー方程式の主な性質について簡単に述べる。

1. 系は 双曲型 である。 つまり、

$$R(\boldsymbol{w}, \boldsymbol{n}) = \sum_{s=1}^{3} n_s \boldsymbol{F}_s(\boldsymbol{w}), \qquad P(\boldsymbol{w}, \boldsymbol{n}) = \frac{DR(\boldsymbol{w}, \boldsymbol{n})}{D\boldsymbol{w}}$$

とすると、行列 P(w,n) は実数の固有値をもち、対角化可能である.

2. 系は 回転に対し不変 である. つまり、単位ベクトル n を角 $\alpha \in [-\pi, \pi], \beta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ を用いて $n = (\cos \alpha \cos \beta, \sin \alpha \cos \beta, \sin \beta)$ と表し回転行列を

$$Q(\mathbf{n}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \cos \alpha \cos \beta & \sin \alpha \sin \beta & \sin \beta & 0 \\ 0 & -\sin \alpha & \cos \alpha & 0 & 0 \\ 0 & -\cos \alpha \sin \beta & -\sin \alpha \sin \beta & \cos \beta & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

とすると、次が成り立つ:

$$R(\boldsymbol{w}, \boldsymbol{n}) = Q^{-1}(\boldsymbol{n}) \boldsymbol{F}_1(Q \boldsymbol{w}).$$

Qの真ん中の部分を取り出した行列

$$Q_0 = \begin{pmatrix} \cos \alpha \cos \beta & \sin \alpha \sin \beta & \sin \beta \\ -\sin \alpha & \cos \alpha & 0 \\ -\cos \alpha \sin \beta & -\sin \alpha \sin \beta & \cos \beta \end{pmatrix}$$

を用いて,変数変換を

$$\begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{pmatrix} = Q_0 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \tilde{\boldsymbol{\sigma}}, \qquad \tilde{\boldsymbol{\sigma}} \in \mathbb{R}^3,$$

とすれば、w が方程式 (5.21) を満たすことと関数 $q(\tilde{x},t):=Q(n)w(Q_0^{-1}(n)(\tilde{x}-\tilde{\sigma}),t)$ が方程式

$$\frac{\partial \mathbf{q}}{\partial t} + \sum_{i=1}^{3} \frac{\partial \mathbf{F}_i}{\partial x_i}(\mathbf{q}) = \mathbf{0}$$

を満たすことが同値であることが示される.

3. 流東 F_s と R は 1 次同次関数 である. よって,

$$F_s(w) = \frac{DF_s(w)}{Dw}w, \qquad R(w, n) = P(w, n)w.$$

実際, $F \in [C^1(Q)]^m$ ならば,

$$\frac{D\boldsymbol{F}_s(\boldsymbol{w})}{D\boldsymbol{w}}\boldsymbol{w} = \lim_{\alpha \to 0} \frac{\boldsymbol{F}_s(\boldsymbol{w} + \alpha \boldsymbol{w}) - \boldsymbol{F}_s(\boldsymbol{w})}{\alpha} = \lim_{\alpha \to 0} \frac{(1 + \alpha)\boldsymbol{F}_s(\boldsymbol{w}) - \boldsymbol{F}_s(\boldsymbol{w})}{\alpha} = \boldsymbol{F}_s(\boldsymbol{w}).$$

次に、有限体積法によるオイラー方程式の離散化について解説する。よく知られているように(双曲型方程式のノートを参照)、(5.21) で定義される古典解がある有限な時間区間でしか存在しないことがある。それ以降の時間で解に不連続性が生じ、衝撃波が出現する。そのため、一般化された弱解を定義する。リプシッツ境界をもつ部分領域 $\sigma \subset \mathbb{R}^3$ と $0 \le t' < t'' < \infty$ に対して方程式を $V = \sigma \times (t',t'')$ 上で積分し、部分積分を施すと、

$$\int_{\sigma} [\boldsymbol{w}(\boldsymbol{x}, t'') - \boldsymbol{w}(\boldsymbol{x}, t')] d\boldsymbol{x} + \int_{t'}^{t''} \int_{\partial \sigma} \sum_{s=1}^{3} \boldsymbol{F}_{s}(\boldsymbol{w}) n_{s} dS dt = \boldsymbol{0}$$
 (5.22)

を得る。ただし、n は σ への外向き単位法線ベクトルである。そこで、弱解を (5.22) を任意のリプシッツ領域 $\sigma \subset \mathbb{R}^3$ と $t',t'' \in [0,\infty), t' < t''$ について満たすような区分的に滑らかな関数 w と定義する。

1. まず、時間区間 (0,T) を部分区間 $0 < t_1 < t_2 < \ldots$ に分割し、空間領域 Ω を四面体 $T_i, i = 1, \ldots, N$ に分割する(有限要素法と同様)、 T_i の体積を $|T_i|$ で表し、 T_j の方向に指す T_i への外向き単位法線ベクトルを n_{ij} という記号を用いて書く、 $\partial T_i \cap \partial T_j$ が共通の面の

とき,その面を Γ_{ij} とし, T_i と T_j とが隣り合っているという.また,次のような記号を導入する:

$$S(i) = \{j: T_i \ \ \ \ \ \ \ \ \ \ \ \ \ \} \cup \bigcup_k \{-k (= 境界に位置する面の引数)\}.$$

こうすることによって、 T_i の境界を

$$\partial T_i = \bigcup_{j \in S(i)} \Gamma_{ij}$$

と書くことができる。なお、領域 Ω の境界を流入の部分 Γ_I 、流出部分 Γ_O 、壁境界 Γ_W と 周期部分 Γ_P に分ける: $\partial\Omega = \Gamma_I \cup \Gamma_O \cup \Gamma_W \cup \Gamma_P$.飛行機の翼のまわりの流れなど多くの 場合、時間が進むにつれ、境界の流入部分と流出部分が変化する.

2. 弱解の定義に従って、オイラー方程式系を $(t_k, t_{k+1}) \times T_i$ で積分し、グリーンの定理を適用する:

$$\begin{split} &\int_{t_k}^{t_{k+1}} \int_{T_i} \frac{\partial \boldsymbol{w}}{\partial t} \, dx \, dt + \int_{t_k}^{t_{k+1}} \int_{T_i} \sum_{s=1}^3 \frac{\partial \boldsymbol{F}_s}{\partial x_s}(\boldsymbol{w}) \, dx \, dt = \boldsymbol{0} \\ &\int_{T_i} \left[\boldsymbol{w}(x, t_{k+1}) - \boldsymbol{w}(x, t_k) \right] dx + \int_{t_k}^{t_{k+1}} \int_{\partial T_i} \sum_{s=1}^3 \boldsymbol{F}_s(\boldsymbol{w}(x, t)) n_s \, dS \, dt = \boldsymbol{0}. \end{split}$$

近似解がそれぞれの四面体 T_i 上で定数関数であると仮定し、2項目の時刻を固定すると、上の式は

$$|T_i|(\boldsymbol{w}_i^{k+1} - \boldsymbol{w}_i^k) + (t_{k+1} - t_k) \sum_{j \in S(i)} \int_{\Gamma_{ij}} \sum_{s=1}^3 (\boldsymbol{n}_{ij})_s \boldsymbol{F}_s(\boldsymbol{w}(x, t_k)) dS = \boldsymbol{0}$$

というふうに近似される. Γ_{ij} 上の積分記号下の式は、時刻 t_k における方向 n_{ij} の面 Γ_{ij} を 通しての単位時間当たりの w の流束として解釈できる. この式を数値流束 H を用いて、次 のように近似する:

$$\sum_{s=1}^{3} (\boldsymbol{n}_{ij})_s \boldsymbol{F}_s(\boldsymbol{w}(x,t_k)) dS \approx \boldsymbol{H}(\boldsymbol{w}_i^k, \boldsymbol{w}_j^k, \boldsymbol{n}_{ij}).$$

このようにして、以下のスキームにたどり着いた:

$$\mathbf{w}_{i}^{k+1} = \mathbf{w}_{i}^{k} - \frac{\tau_{k}}{|T_{i}|} \sum_{j \in S(i)} \mathbf{H}(\mathbf{w}_{i}^{k}, \mathbf{w}_{j}^{k}, \mathbf{n}_{ij}) \mu_{ij}, \qquad i = 1, \dots, N$$

ただし、 μ_{ij} は Γ_{ij} の面積で、 $\tau_k = t_{k+1} - t_k$ である。これは陽的スキームであるが、陰的スキームは同様に定義できる。陽的スキームは時間のステップ幅に対する CFL 条件 $\tau = O(h)$ が課される。一方では、陰的スキームはそのような制限がないが、各時間レベルで非線形連立方程式を解かなければならない。定常問題の場合(すなわち、 $\frac{\partial}{\partial t} = 0$)、経験上では陽的

スキームの方が計算が速い.

注. 関数 w のどの点での値を未知数とするかによって、cell-centered (四面体の重心)、cell-vertex (四面体の頂点)、cell-edge (辺の中点)など様々なスキームがある.



Cell-centred, cell-vertex と cell-edge タイプのスキーム

- 3. あとは、適切な数値流束を構成すればよい、数値流束に要求される条件は以下の2点である:
 - (1) consistency: $H(\boldsymbol{w}, \boldsymbol{w}, \boldsymbol{n}_{ij}) = \sum_{s=1}^{3} \boldsymbol{F}_{s}(\boldsymbol{w}) \boldsymbol{n}_{s}$
 - (2) conservativity: $H(\boldsymbol{w}_i, \boldsymbol{w}_j, \boldsymbol{n}_{ij}) = -H(\boldsymbol{w}_j, \boldsymbol{w}_i, -\boldsymbol{n}_{ij}) \ \forall \boldsymbol{w}_i, \boldsymbol{w}_j$

上で線形方程式の場合に勉強した風上差分の考え方は Godunov スキームにつながるが、それを以下で説明する.

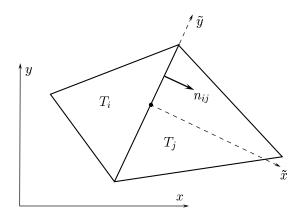
Godunov 型のスキーム

線形なスカラー問題 (5.13)-(5.14) の正しい数値流束は $H(u_i^k,u_{i+1}^k)=a^+u_i^k+a^-u_{i+1}^k$ により与えられるが、これはちょうどリーマン問題

$$u_t + au_x = 0 \quad \text{in } \mathbb{R} \times (0, \infty)$$
$$u(x, 0) = \begin{cases} u_i^k & x < 0 \\ u_{i+1}^k & x > 0 \end{cases}$$

の解の x=0 における値と一致する。つまり、リーマン問題の解は u_i^k と u_{i+1}^k という値の境目で解の流れがどのように変化していくかを教えてくれるのである。数値流束を風上差分の考え方に従って構成するとき、このリーマン問題の情報を利用する。

上の考察をオイラー方程式の解法に一般化する。簡単のため、図と変換を2次元の場合のみ示す。したがって、下図のような2つの隣り合う要素を考える。



方程式が双曲型であるため, 行列

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & n_1 & n_2 & 0 \\ 0 & n_2 & -n_1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

を用いて座標を回転させ,

$$R(\boldsymbol{w}, \boldsymbol{n}) = \sum_{s=1}^{2} \boldsymbol{F}_{s}(\boldsymbol{w}) n_{s} = Q^{-1} \boldsymbol{F}_{1}(Q\boldsymbol{w})$$

となるようにできる。 ただし、 $n_{ij}=(n_1,n_2)$ としている。

そして,数値流束を次のように定義する:

$$H(\boldsymbol{w}_{i}, \boldsymbol{w}_{j}, \boldsymbol{n}_{ij}) \approx \sum_{s=1}^{3} \boldsymbol{F}_{s}(\boldsymbol{w}) n_{s}|_{\Gamma_{ij}} = R(\boldsymbol{w}, \boldsymbol{n})|_{\Gamma_{ij}} = Q^{-1} \boldsymbol{F}_{1}(\boldsymbol{q})|_{\Gamma_{ij}} \approx Q^{-1} \boldsymbol{F}_{1}(\hat{\boldsymbol{q}}(\tilde{x}=0, t))$$

ここで、q = Qwで、 \hat{q} は以下のリーマン問題の解の $\tilde{x} = 0$ における値である:

$$\frac{\partial \mathbf{q}}{\partial t} + \frac{\partial \mathbf{F}_1(\mathbf{q})}{\partial \tilde{x}_1} = 0$$

$$\mathbf{q}(\tilde{x}_1, 0) = \begin{cases} \mathbf{q}_L = Q\mathbf{w}_i & \tilde{x}_1 < 0 \\ \mathbf{q}_R = Q\mathbf{w}_i & \tilde{x}_1 > 0 \end{cases}$$

以下の量,

$$\boldsymbol{f}_{R}(\boldsymbol{q}_{L},\boldsymbol{q}_{R}) = \boldsymbol{F}_{1}(\hat{\boldsymbol{q}}(\tilde{x}=0))$$

を リーマン・ソルバー とよぶ.

(5.16) でわかったように、スカラーの非線形問題の場合にはリーマン・ソルバーを

$$f_R(q_L, q_R) = \frac{1}{2} \left\{ F(q_L) + F(q_R) - \int_{q_L}^{q_R} |F'(q)| \, dq \right\}$$

と置くことができる。オイラー方程式は連立方程式なので、この考えをさらに拡張する必要がある。詳細を省いて、結論だけ述べると、リーマン・ソルバーを次のように定めることができる:

$$\boldsymbol{f}_{R}(\boldsymbol{q}_{L},\boldsymbol{q}_{R}) = \frac{1}{2} \left\{ \boldsymbol{F}(\boldsymbol{q}_{L}) + \boldsymbol{F}(\boldsymbol{q}_{R}) - \int_{\boldsymbol{q}_{l}}^{\boldsymbol{q}_{R}} |A(\boldsymbol{q})| d\boldsymbol{q} \right\}$$
(5.23)

ここで、Fのヤコビ行列Aを

$$A = TDT^{-1}, \qquad D = \operatorname{diag}(\lambda_1, \dots, \lambda_5),$$

と対角化し,

$$A^{\pm} = TD^{\pm}T^{-1}, \qquad D^{\pm} = \operatorname{diag}(\lambda_1^{\pm}, \dots, \lambda_5^{\pm})$$

とした. 同様に,

$$A = T|D|T^{-1},$$
 $|D| = \operatorname{diag}(|\lambda_1|, \dots, |\lambda_5|)$

と定義する。このリーマン・ソルバーの定義が意味をもつのは,(5.23) における積分 $\int_{q_L}^{q_R}$ が積分の経路に依らないときのみである。この仮定が実際には満たされないにもかかわらず,上記のリーマン・ソルバーはいわゆる flux vector splitting schemes においてオイラー方程式を解くのに使われる。このようなスキームは主に 2 種類ある:

- 1. (5.23) の積分は適切な経路に沿った積分として計算される (Osher-Solomon 法など).
- 2. (5.23) の積分は適切な数値積分法により近似される.
 - (a) Vijayasundaram スキーム は次の近似を用いる:

$$\begin{split} \int_{\boldsymbol{q}_L}^{\boldsymbol{q}_R} |A(\boldsymbol{q})| \, d\boldsymbol{q} &\approx \left| A \left(\frac{\boldsymbol{q}_L + \boldsymbol{q}_R}{2} \right) \right| (\boldsymbol{q}_R - \boldsymbol{q}_L) \\ \frac{1}{2} \left(\boldsymbol{F}(\boldsymbol{q}_L) + \boldsymbol{F}(\boldsymbol{q}_R) \right) &\approx \boldsymbol{F} \left(\frac{\boldsymbol{q}_L + \boldsymbol{q}_R}{2} \right) = A \left(\frac{\boldsymbol{q}_L + \boldsymbol{q}_R}{2} \right) \frac{\boldsymbol{q}_L + \boldsymbol{q}_R}{2}. \end{split}$$

これらを適用すると、リーマン・ソルバーは

$$oldsymbol{f}_V = A^+ \left(rac{oldsymbol{q}_L + oldsymbol{q}_R}{2}
ight) oldsymbol{q}_L + A^- \left(rac{oldsymbol{q}_L + oldsymbol{q}_R}{2}
ight) oldsymbol{q}_R$$

という形になる。

(b) $\underline{\text{Steger-Warming } Z + - \Delta}$ は $F^{\pm}(q) = A^{\pm}(q)q$ という近似を用いる。リーマン・ソルバーの形は

$$\boldsymbol{f}_{SW}(\boldsymbol{q}_L, \boldsymbol{q}_R) = A^+(\boldsymbol{q}_L)\boldsymbol{q}_L + A^-(\boldsymbol{q}_R)\boldsymbol{q}_R.$$

(c) Van Leer スキーム は次のリーマン・ソルバーを使う:

$$\boldsymbol{f}_{VL}(\boldsymbol{q}_L,\boldsymbol{q}_R) = \frac{1}{2} \left\{ \boldsymbol{F}(\boldsymbol{q}_L) + \boldsymbol{F}(\boldsymbol{q}_R) - \left| A \Big(\frac{\boldsymbol{q}_L + \boldsymbol{q}_R}{2} \Big) \right| (\boldsymbol{q}_R - \boldsymbol{q}_L) \right\}.$$

オイラー方程式の数値解法についてのよい参考書は[1]である.

References

- [1] M. Feistauer, J. Felcman, I. Straškraba, *Mathematical and Computational Methods for Compressible Flow*, Oxford science publications, Clarendon Press, 2003.
- [2] G.H. Golub, C.F. Van Loan, Matrix Computations, The Johns Hopkins University Press, 2013.
- [3] J. Málek, Z. Strakoš, *Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs*, SIAM, 2015.

- [4] A. Quarteroni, R. Sacco, F. Saleri, Numerical Mathematics, Springer, 2007.
- [5] J.R. Shewchuk, *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*, 1994. (http://www.cs.cmu.edu/~./quake-papers/painless-conjugate-gradient.pdf)
- [6] 菊地 文雄、齊藤 宣一、「数値解析の原理」、岩波書店、2016.
- [7] 齊藤 宣一,「数値解析入門」,東京大学出版会,2012.