転置行列 14

デルタ関数的擾乱 114

ナイキスト周波数 61
並べ換え行列 13, 18–20

二次元二階偏微分方程式 142
二分法 4
二変数二分法 6

熱伝導方程式 107, 142

はさみうち法 4
波動方程式 79
反復しないスキーム 100, 107
反復するスキーム 101, 107

パワースペクトル密度 53

非線形移流方程式 128, 129
非線形不安定性 129
非中心差分 96
標本 38
標本分散 38
標本平均 38

ピボット 11, 13, 18, 21, 23

フーリエ級数法 83, 85
von Neumann 法 87

部分的ピボッティング 11–14

ヘッセンベルグ行列 25, 32, 34, 35
ヘッセンベルグ形 29, 32, 33, 35

冪乗法 30–32, 36
ベッセル関数 114

放射フラックス 15
方程式の解法 3
放物型方程式 142

母集団 38
母分散 38
母平均 38

摩擦項 109
摩擦方程式 107
松野 scheme 97

無作為抽出 38

有界性 83
有限差分法 77

陽解法 96, 142

ラグウインドウ 57
ランダム過程 49

連立一次方程式 9–18, 24
連立常微分方程式 93

惑星境界層 106

打ち切り誤差 82, 111

エイリアジング 61
エネルギー法 83, 85
エルゴード性 50

折り畳み周波数 61

解の反復改良 12–14, 24
拡散方程式 142
拡大行列 10–12
確率過程 49
火星 8
完全ピボッティング 11
緩和法 140

ガウスの消去法 9–14

気象業界 75
基本周波数 48
吸収係数 16
境界層 107
鏡像変換 33, 34
強定常確率過程 49
極冠 8

空間中心差分スキーム 119, 125

グラム ・シュミットの直交化 27, 29, 31–33
群速度 113

係数行列 9, 10, 12–14, 18, 22

原点移動 36

光学的厚さ 15, 16
高速フーリエ変換 64
後退差分近似式 78
後退代入 10–14, 21, 23
固有値問題 25, 26, 32, 36
コレスキー (Cholesky) 分解定理 14, 22
コレスキー分解法 9, 14, 23

最大エントロピー法 57
差分近似 77
差分スキーム 81

下三角行列 13, 14, 18–20

周期的境界条件 86
修正 Euler 法 91
収束性 81, 82
主小行列 22
振動項 109
振動方程式 100

時間発展問題 75, 95
自己相関関数 51
弱定常確率過程 49
条件が良い 24
条件が悪い 12, 24
条件数 24
条件が良い 24
条件が悪い 24
常微分方程式 90
上流差分スキーム 116

数値計算の歴史と現状 75
数値分散 113
スケーリング 11–14, 21, 23
スペクトルウインドウ 58

正確度 81, 111
正値対称行列 9, 14, 22
正方行列 18

前進差分近似式 77
前進消去 10, 11, 13, 14, 18, 21, 23
前進代入 13, 14, 22

増幅係数 87

大気の鉛直温度構造 9, 15
大気放射学 17

ダブル QR 法 36

中心差分 96
中心差分近似 114
中心差分近似式 78
調和解 113
直接法 56, 83, 85, 137

定常確率過程 49
適合性 81, 111
天気予報 75

索 引

2 次元移流方程式 127

Adames-Bashforth scheme 97
Arakawa スキーム 131
augmented matrix 10

back substitution 10
Blackman-Tukey 法 56

Cardano 法 3
CFL condition 112
complete pivoting 11
computational mode 97, 103
Courant-Friedfichs-Leny condition 112
Crank-Nicolson 法 143

DFT 64

ENIAC 75
enstrophy 129
Euler scheme 96
Euler 法 91

Ferrari 法 3
FFT 64
FFT のアルゴリズム 65
forward elimination 10
forward substitution 13

hamming 58
hanning 58
Heun scheme 97

ill-conditioned 24

Leapfrog scheme 97
leapfrog scheme 103
Liebmann 法 141
LU 分解 13, 14
LU 分解定理 13
LU 分解定理 18
LU 分解法 9, 13, 14, 21

Malne-Simpson scheme 97
MEM 57

Newton-Raphson 法 3

partial pivoting 11
physical mode 103
pivot 11

QR 分解 27–29, 32, 33, 35
QR 法 26, 27, 30, 32, 36

Regula-Falsi 法 4
Runge-Kutta 法 92

semi-Implicit 法 99
Split 法 110
split 法 110

three level scheme 97, 104, 108
two level scheme 96

von Neumann 法 87

well-conditioned 24
Wiener-Khinchine の公式 53

安定性 81, 83, 85

依存領域 82
一次元線形移流方程式 111
移流方程式 111, 113
陰解法 95

上三角行列 13, 18
上向きフラックス 15, 17

ベンダット J. S., ピアソル A. G., 1976: ランダムデータの統計的処理, 培風館.

森正武, 1973: 数値解析, 共立数学講座, 第 12 巻, 共立出版.

森正武, 1984: 数値解析法, 朝倉現代物理学講座, 第 7 巻, 朝倉書店.

森正武, 1988: FORTRAN77 数値計算プログラミング (増補版), 岩波書店.

山岡耕春, 1990: 火山性地震 ・火山性微動とマグマ活動, 火山, **34**, S227–S234.

Priestley, M. B., 1981: *Spectral Analysis and Time Series*, Academic Press.

Richardson, L. F., 1922: *Weather Prediction by Numerical Process*, Cambridge University Press.

Richtmyer, R. D., 1963: A survey of difference methods for nonsteady fluid dynamics, Technical Report 63–2, National Center for Atmospheric Research, Boulder, Colo.

Richtmyer, R. D. and Morton, K. W., 1967: *Difference Methods for Initial Value Problems*, Interscience.

Roberts, P. H., 1972: Kinematic Dynamo Models, *Phil. Trans. R. Soc. Lond.*, **A272**, 663–698.

Rossing, T. D., 1992: Acoustics of Drums, *Physics Today*.

Shuman, F. G., 1974: Analysis and experiment in nonlinear computational stability, in *Difference and Spectral Methods for Atmosphere and Ocean Dynamics Problems, Proc. Symp., Novosibirsk, Part I*, 51–81.

Thompson, P. D., 1961: *Numerical Weather Analysis and Prediction*, Macmillan.

Turner, J. S., 1973: *Buoyancy Effects in Fluids*, Cambridge.

Wurtele, M. G., 1961: On the problem of truncation error, *Tellus*, **13**, 379–391.

Young, J. A., 1968: Comparative properties of some time differencing schemes for linear and non-linear oscillations, *Mon. Wea. Rev.*, **96**, 357–364.

会田勝, 1982: 大気と放射過程, 気象学のプロムナード, 第 8 巻, 岩波書店.

小国力, 1991: 行列計算ソフトウェア – WS, スーパーコン, 並列計算機 –, 丸善.

川嶋 弘尚 ・酒井英昭, 1989: 現代スペクトル解析, 森北出版.

キッテル, 1986: 固体物理学入門, 丸善, 第 6 版.

ザイマン, 1976: 固体物性論の基礎, 丸善, 第 2 版.

杉江日出澄 ・他, 1986: FORTRAN77 による数値計算法, 培風館.

竹内均, 1972: 地球科学における諸問題, 裳華房.

竹内啓, 1989: 統計学辞典, 東洋経済新報社.

津田孝夫, 1988: 数値処理プログラミング, 岩波書店.

戸川隼人, 1971: マトリクスの数値計算, オーム社.

永田豊, 1976: 海洋物理学 II, 東京大学出版会, 第 12 章.

早野龍五, 高橋忠幸, 1989: 計算機実験 (平成元年度版), 東京大学理学部物理学教室.

日野幹雄, 1977: スペクトル解析, 朝倉書店.

平田邦男, 1989: Turbo C による科学計算入門, 共立出版.

ブラント S., 1976: データ解析の方法, みすず書房.

Goody, R. M. and Yung, Y. L., 1989: *Atmospheric Radiation Theoretical Basis*, Oxford University Press, second edition.

Jespersen, D. C., 1974: Arakawa's method is a finite-element method, *J. Comput. Phys.*, **16**, 383–390.

Kasahara, A., 1969: Simulation of the earth's atmosphere, Technical report, National Center for Atmospheric Research, Boulder, Colo.

Kreiss, H. and Oliger, J., 1973: *Methods for the approximate solution of time dependent problems*, Vol. 10 of *GARP Publications Series*, WMO/ICSU Joint Organizing Committee.

Krishnamurti, T. N., Kanamitsu, M., Ceselski, B. and Mathur, M. B., 1973: Florida State University's Tropical Prediction Model, *Tellus*, **25**, 523–535.

Kurihara, Y., 1965: On the use of implicit and iterative methods for the time integration of the wave equation., *Mon. Wea. Rev.*, **93**, 33–46.

Leith, C., 1965: Lagrangian advection in an atmospheric model, in *WMO-IUGG Symposium on Research and Development Aspects of Long Range Forecasting, Boulder, Colo., 1965, WMO Tech. Note, 66*, 168–176.

Lighthill, M. J., 1958: *An Introduction to Fourier Analysis and Generalised Functions*, Cambridge University Press.

Lilly, D. K., 1965: On the computational stability of numerical solutions of time-dependent non-linear geophysical fluid dynamics problems, *Mon. Wea. Rev.*, **93**, 11–26.

Marple, J. S. L., 1987: *Digital Spectral Analysis*, Prentice Hall.

Matsuno, T., 1966a: A finite difference scheme for time integrations of oscillatory equations with second order accuracy and sharp cut-off for high frequencies, *J. Meteor. Soc. Japan, Ser. 2*, **44**, 85–88.

Matsuno, T., 1966b: Numerical integrations of the primitive equations by a simulated backward difference method., *J. Meteor. Soc. Japan, Ser. 2*, **44**, 76–84.

Orszag, S. A., 1971: On the elimination of aliasing in finite-difference schemes by filtering high-wavenumber components, *J. Atmos. Sci.*, **28**, 1074.

Papoulis, A., 1962: *The Fourier Integral and Its Applications*, McGraw-Hill.

Papoulis, A., 1984: *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 2nd edition.

Phillips, N. A., 1956: The general circulation of the atmosphere : a numerical experiment, *Quart. J. Roy. Meteor. Soc.*, **82**, 123–164.

Platzman, G. W., 1964: An exact integral of complete spectral equations for unsteady one-dimensional flow, Technical Report 16, Dynamical Prediction Group, Dept. of the Geophysical Sciences, Univ. of Chicago.

Press, W. H., Flannary, B. P., Teukolsky, S. A. and Vetterling, W. T., 1986: *Numerical Recipes : the art of scientific computing*, Cambridge.

References

Ames, W. F., 1969: *Numerical Methods for Partial Differential Equations*, Nelson.

Anderson, D. and Fattahi, B., 1974: A comparison of numerical solutions of the advective equation, *J. Atmos. Sci.*, **31**, 1500–1506.

Arakawa, A., 1966: Computational design for long-term numerical integration of the equations of fluid motion : Two dimensional incompressible flow. Part I, *J. Comput. Phys.*, **1**, 119–143.

Blackman, R. B., 1959: *The Measurement of Power Spectra*, Dover Publications.

Brigham, E. O., 1974: *Fast Fourier Transform*, Prentice Hall.

Chandrasekhar, , 1961: *Hydrodynamic and Hydromagnetic Stability*, Oxford Univ. Press.

Charney, J. G., 1966: Some remaining problems in numerical weather prediction, in *Advances in Numerical Weather Prediction*, Hartford, Conn.,Travelers Research Center, Inc., 61–70.

Charney, J. G., Fjørtoft, R. and von Neuman, J., 1950: Numerical integration of the barotropic vorticity equation, *Tellus*, **2**, 237–254.

Courant, R., Friedrichs, K. and Lewy, H., 1928: Uber die partiellen Differenzengleichungen der mathematischen Physik, *Math. Annalen*, **100**, 32–74.

Deardorff, J. W., 1974: Three-dimensional numerical study of the height and mean structure of a heated planetary boundary layer, *Boundary-Layer Meteor.*, **7**, 81–106.

Doob, J. L., 1953: *Stochastic Processes*, John Wiley and Sons, Inc.

Egger, J., 1971: Mindestgrösse von Gebirgen und Konvektionsgebieten, die in den Modellen der numerischen Vorhersage berücksichtigt werden können, *Beitr. Phys. Atmos.*, **44**, 245–271.

Feller, W., 1966: *An Introduction to Probability Theory and Its Applications*, Vol. 2, John Wiley and Sons, Inc., 1st edition.

Fjørtoft, R., 1953: On the changes in the spectral distribution of kinetic energy for two-dimensional, nondivergent flow, *Tellus*, **5**, 225–230.

Fletcher, N. H. and Rossing, T. D., 1991: *The Physics of Musical Instruments*, Springer Verlag.

Gierasch, P. J. and Toon, O. B., 1973: Atmospheric pressure variation and the climate of Mars, *J. Atomos. Sci.*, **31**, 1502–1508.

第 9 章

著作権

Copyright © 1992, 地球惑星物理学科地球惑星物理学演習テキスト編集グループ

本テキストに関する著作権は、地球惑星物理学科地球惑星物理学演習テキスト編集グループ (以下 「グループ」と略す) に属します.
本テキストは 「グループ」に無断で複写・複製することを禁止します. また, 本テキストのソースプログラム (LaTeX とポストスクリプトファイル) も無断で再配布・変更してはいけません. 本テキストの一部もしくは全部を引用する場合は, 当 「グループ」による 「地球惑星物理学演習テキスト」からであることを明記して下さい.

なお、著作者 「地球惑星物理学科地球惑星物理演習テキスト編集グループ」の連絡先は

ta_comp@geoph.s.u-tokyo.ac.jp

です. 連絡はすべて電子メールでお願いします.

第 8 章

謝辞

3 章

例題のヒントを提供して下さった東京大学気候システム研究センターの中島映至助教授に感謝致します.

7 章

この章の内容のはほとんどは, 室井ちあし氏 (現気象庁) が編集した地球惑星物理学科地球流体グループの大気計算のための 数値算法ノート」によっています.

その他

テキスト作成に先立ち, さまざまな参考資料を提供していただいた地球惑星物理学科専攻課程の教官方に感謝致します. また, 気候システム研究センターの中島健介, 山中康裕両氏, ならびに, 東京都立大学の増田耕一氏には, テキスト作成を通じて多大なる助言をいただきました.

	Crank-Nicolson 法	解析解
$u_{1,j}$	0.081	0.076
$u_{2,j}$	0.114	0.108

7.21　その他の参考文献

本章の内容のうち, 常微分方程式の解法に関する記述は早野, 高橋, 1989; 杉江, 1986 から引用した. また, 偏微分方程式の差分解法と放物型方程式に関する記述は平田, 1989; 杉江, 1986 から引用した.

例題 1 の解法の指針

$h=\frac{1}{10}, r=1$ つまり $k=100$ とする. 系は $x=\frac{1}{2}$ に対して対象なので, $u_{i,j}(i=1,2,\cdots,5, j=1,2,\cdots)$ を求めれば良いことになる. 境界条件から, また対称性からであることを考慮して, 式 (7.132) を行列で表すと,

$$\begin{pmatrix} r_1 & & & & 0 \\ -1 & r_1 & -1 & & \\ & -1 & r_1 & -1 & \\ & & -1 & r_1 & -1 \\ 0 & & & -2 & r_1 \end{pmatrix} \begin{pmatrix} u_{1,j+1} \\ u_{2,j+1} \\ u_{3,j+1} \\ u_{4,j+1} \\ u_{5,j+1} \end{pmatrix} = \begin{pmatrix} -r_2 & 1 & & & 0 \\ 1 & -r_2 & 1 & & \\ & 1 & -r_2 & 1 & \\ & & 1 & -r_2 & 1 \\ 0 & & & 2 & -r_2 \end{pmatrix} \begin{pmatrix} u_{1,j} \\ u_{2,j} \\ u_{3,j} \\ u_{4,j} \\ u_{5,j} \end{pmatrix}$$

$$= \begin{pmatrix} -r_2 u_{1,j} + u_{2,j} \\ u_{1,j} - r_2 u_{2,j} + u_{3,j} \\ u_{2,j} - r_2 u_{3,j} + u_{4,j} \\ u_{3,j} - r_2 u_{4,j} + u_{5,j} \\ 2u_{4,j} - r_2 u_{5,j} \end{pmatrix}$$

となる. ここで

$$r_1 = \frac{2(r+1)}{r}, \quad r_2 = \frac{2(r-1)}{r}$$

である. 初期条件 $t=0$ すなわち $j=0$ における $u_{i,0}(i=1,2,\cdots,5)$ を式 (7.135) から求め, 式 (7.138) を三項対角連立一次方程式の解法 (Thomas 法) を用いて解けば, $u_{i,1}(i=1,2,\cdots,5)$ が求められる. 以下, 式 (7.135) を解いて $u_{i,j}(i=1,2,3,\cdots)$ を順次求めていく.

参考までに $t=0.1$ のときの計算結果のみを以下に示す.

j	t	$u_{1,j}$	$u_{2,j}$	$u_{3,j}$	$u_{4,j}$	$u_{5,j}$
10	0.100	0.0934	0.1776	0.2444	0.2873	0.3021

例題 2

初期条件

$$t=0, \quad 0<x<1 \quad \text{で} \quad u=1$$

境界条件

$$x=0 \text{ および } x=1, \quad t \geq 0 \text{ で} \quad u=0$$

のもとで, 放物型偏微分方程式

$$\frac{\partial u}{\partial x} = \frac{\partial^2 u}{\partial x^2}$$

を満たす u の数値解を区間 $0<t\leq 1$ で, $h=\frac{1}{4}, r=\frac{1}{2}$ (すなわち $k=\frac{1}{32}$) として, 陽解法および Crank-Nicolson 法で求め, 解析解と比較せよ. 解析解は次式で与えられる.

$$u = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \exp\{-(2n+1)^2 \pi^2 t\} \sin(2n+1)\pi x$$

ちなみに解の一部は, $t_j=0.25$ のとき

である.

例題 1

熱伝導方程式の計算を考える. 図 7.8に示すように, 棒の両端の温度を 0 に保ち, 棒の中心点を加熱すると, 棒の温度は中心点を境にして対象な直線状の温度分布となる.

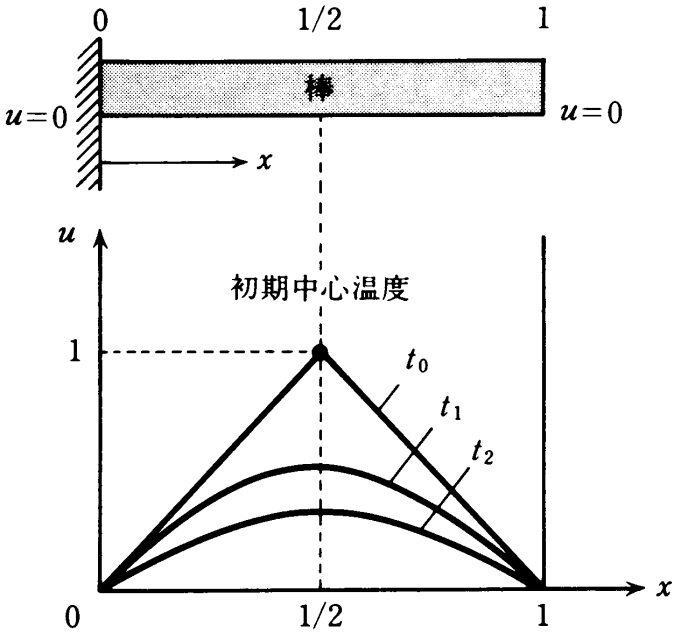


図 7.8: 例題 1

今, 中心点における加熱を中止すると, 棒の温度は時間の経過と共に図のように下がっていくであろう $(t_0 \to t_1 \to t_2 \to \cdots)$. これを数式で表現すると次のような熱伝導方程式になる.

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{7.133}$$

x は長さ, u は温度, t は時間の無次元量である. $x = 0$ および $x = 1$ における境界条件は

$$u = 0 \quad (t \geq 0) \tag{7.134}$$

であり, $t = 0$ における初期条件は

$$u = 2x \qquad (0 \leq x \leq \frac{1}{2}) \tag{7.135}$$

$$u = 2(1 - x) \quad (\frac{1}{2} \leq x \leq 1) \tag{7.136}$$

である. このとき無次元温度 u の時間変化を, 式 (7.133) から $x = 0.1, 0.2, 0.3, 0.4, 0.5$ の各点で Crank-Nicolson 法で求めてみよう. なお, 条件式 (7.134)~(7.136) を満足する式 (7.133) の解析解は次式で与えられる.

$$u = \frac{8}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \sin \frac{n\pi}{2} \sin n\pi x \exp(-n^2 \pi^2 t) \tag{7.137}$$

この解析解を Crank-Nicolson 法による近似解と比較してみよう.

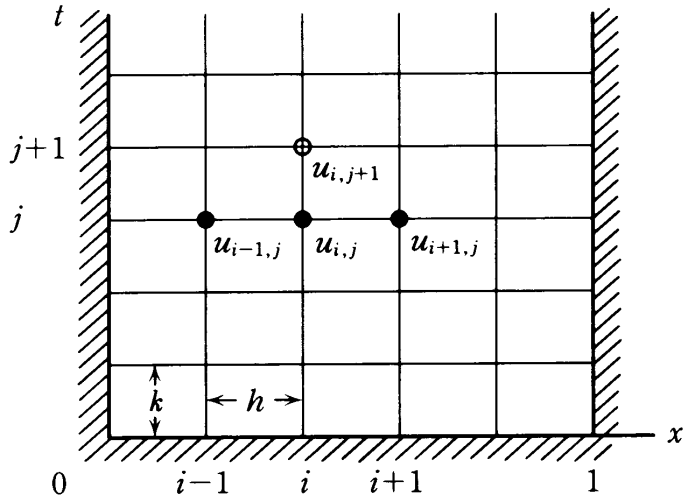


図 7.7: 陽解法

陽解法は計算方法が極めて簡単であるが, その代わりに式 (7.130) の rの値の選定に関する問題, つまり解法の収束安定性という問題がつきまとう. h と kとを共に小さくすればこの差分近似解が真の解に近付くであろうと単純に考えがちであるが, 必ずしもそうではない. この解法の安定性を考えて rの値を決める必要があるのである. 陽解法の安定性の条件は

$$0 < r \leq \frac{1}{2}$$

で与えられている.

Crank-Nicolson 法

陽解法は $0 < r \leq \frac{1}{2}$ つまり $k \leq \frac{1}{2}h^2$ が成立しないと不安定になり, 実用的な解が得られない. よって, 十分な精度を得るには h を小さくし, 時間刻 kも小さくしなければならない. しかしたとえば h を$\frac{1}{2}$にすると kは$\frac{1}{4}$にしなければならず, 格子点の数は 8 倍になり, 計算量も約 8 倍になる. このような計算量の爆発的増大の困難を克服するために考案されたのが Crank-Nicolson 法である. この方法が陽解法ともっとも異なる点は, $\dfrac{\partial^2 u}{\partial x^2}$ を, 第 j 時刻での u の二階差分だけではなく次の第 $j+1$ 時刻での u の二階差分との平均で置き代えたことにある. 具体的には, 式 (7.7) を

$$\frac{u_{i,j+1} - u_{i,j}}{k} = \frac{1}{2}\left(\frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{h^2} + \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2}\right) \tag{7.131}$$

と近似している. この式から

$$-u_{i-1,j+1} + 2\left(\frac{1+r}{r}\right)u_{i,j+1} - u_{i+1,j+1} = u_{i-1,j} - 2\left(\frac{1+r}{r}\right)u_{i,j} + u_{i+1,j} \tag{7.132}$$

が得られる. 式 (7.132) で $u_{i-1,j}, u_{i,j}, u_{i+1,j}$の値が既知であるとして, 未知数 $u_{i-1,j+1}, u_{i,j+1}, u_{i+1,j+1}$を求めるという方法である. 式 (7.132) は第 j時刻の $i = 1, 2, \cdots, M-1$ の $M-1$ 個の格子点上で成立しなければならないので, $M-1$ 元連立一次方程式を与えていることになる. よって Crank-Nicolson 法では未知数の決定に連立一次方程式の解を必要とするので, rのいかなる値に対しても安定 (無条件安定) であることが確かめられている.

7.20　補足 ・放物型方程式

一般に, x, yの関数 $u(x, y)$ の二次元二階偏微分方程式は

$$a\frac{\partial^2 u}{\partial x^2} + b\frac{\partial^2 u}{\partial x \partial y} + c\frac{\partial^2 u}{\partial y^2} + d\frac{\partial u}{\partial x} + e\frac{\partial u}{\partial y} + fu + g = 0$$

の形に書き表される. ここで, a, b, c, d, e, f, gは定数または独立変数 x, yや従属変数 u の関数である. 上の式は a, b, c の値によって, 式の形が以下のように分類される.

1. $b^2 - 4ac > 0$ のとき, 双曲型の式
2. $b^2 - 4ac = 0$ のとき, 方物型の式
3. $b^2 - 4ac < 0$ のとき, 楕円型の式

ここでは, 補足として放物型偏微分方程式の解法について簡単に述べる.

7.20.1　放物型方程式

時間 t を独立変数に含むような問題は放物型か双曲型になるのが普通である. 拡散方程式または熱伝導方程式として良く知られているもっとも簡単な方物型方程式は

$$\frac{\partial u}{\partial \theta} = c\frac{\partial^2 u}{\partial \xi^2} \quad (0 \le \xi \le l)$$

という形をしている. ただしは θ 時間, ξ は距離に関する変数で, c は定数である. いま

$$t = \frac{c}{l}\theta, \quad x = \frac{\xi}{l}$$

という置換えを行なって無次元化すれば

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad (0 \le x \le 1) \tag{7.128}$$

という方程式が得られる. 以下の説明は式 (7.128) について行なう.

陽解法

式 (7.128) に対する差分近似式は, 前節の解説により

$$\frac{u_{i,j+1} - u_{i,j}}{k} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \tag{7.129}$$

となる. $r \equiv k/h^2$ とおけば, 式 (7.129) は

$$u_{i,j+1} = r(u_{i+1,j} + u_{i-1,j}) + (1 - 2r)u_{i,j} \tag{7.130}$$

と書くことができる. 式 (7.130) と図 7.7からわかるように, この方法は第 j時刻の u の値 (既知の値, 図中の黒丸点) から第 $j + 1$ 時刻の u の値 (未知の値, 図中の白丸点) を順次繰り返して求める方法である. たとえば $j = 0$ の線上のすべての格子点の値すなわち境界値が与えられていれば (Dirichlet の問題) これを出発点にして式 (7.130) を用いて t 軸方向に一段上の x 軸に平行な線上のすべての格子点の値を求めることができる. このように, 一段ずつ各格子点の値を求めながら t 軸方向に前進する差分公式を陽解法 (explicit method) という.

とおいた. von Neumann 法を用いてこのスキームの安定性を調べると, 安定であるためには

$$\Delta t \leq \frac{d^2}{4}$$

であれば十分であることが分かる (各自確かめてみよ.).

後は, このスキームによって定常になるまで繰り返し計算すれば良い. 定常かどうかは残差

$$\begin{aligned} R^n_{i,j} &\equiv \nabla^2\psi_{i,j} - \rho_{i,j} \\ &= \frac{1}{d^2}(\psi^n_{i+1,j} + \psi^n_{i-1,j} + \psi^n_{i,j+1} + \psi^n_{i,j-1} - 4\psi^n_{ij}) - \rho_{ij} \end{aligned}$$

があらかじめ決めておいたある正数 ε より小さくなったかどうかで判定する.

7.19.4 緩和法の収束の速さ

定常に達する速さ (収束の速さ) を考えてみる.

安定の条件を満たすように, $\Delta t = \dfrac{d^2}{4}$ とすると, (7.125) より

$$\psi^{n+1}_{i,j} = \frac{1}{4}(\psi^n_{i+1,j} + \psi^n_{i-1,j} + \psi^n_{i,j+1} + \psi^n_{i,j-1} - d^2\rho_{i,j}) \tag{7.126}$$

である. よって, 完全に定常に達したときの解, つまり

$$\nabla^2\psi_{i,j} - \rho_{i,j} = 0$$

が成り立つような解を $\tilde{\psi}$ とし, それからの ずれを e とすると,

$$\begin{aligned} e^{n+1}_{i,j} &= \psi^{n+1}_{i,j} - \tilde{\psi}_{i,j} \\ &= \frac{1}{4}(e^n_{i+1,j} + e^n_{i-1,j} + e^n_{i,j+1} + e^n_{i,j-1}) \end{aligned} \tag{7.127}$$

となる. ここで, e に i, j 方向の (離散) Fourier 展開を施しその展開係数を $\hat{e}$ と書くと

$$\hat{e}^{n+1}_{k,m} = \frac{1}{2}(\cos k\Delta x + \cos m\Delta z)\hat{e}^n_{k,m}$$

となる. ただし, k, m は, それぞれ, x, z 方向の波数である. 従って, この波数の成分の減衰の速さは

$$\frac{1}{2}(\cos k\Delta x + \cos m\Delta z)$$

である. これは, スキーム (7.127) に von Neumann 法を適用した時の増幅 (または減衰) 係数である.

『Liebmann 法』では, 式 (7.125) の$\psi^n_{i-1,j}$, $\psi^n_{i,j-1}$ の代わりに, それぞれ, $\psi^{n+1}_{i-1,j}$, $\psi^{n+1}_{i,j-1}$ を用いる[24]. こうすると, 減衰係数は

$$\frac{1}{2}(\cos k\Delta x + \cos m\Delta z)^2$$

となり, 収束が速くなる (各自確かめてみよ).

[24] FORTRAN の DO ループなどの中で, スキャンしながら $\psi_{i,j}$ の値を更新していくと自然にこうなる.

7.19　緩和法

7.19.1　Introduction

ここでは, 緩和法といわれる数値計算法について説明する.

第 7.19 節と同様に

$$\mathcal{L}\psi = \rho \tag{7.121}$$

の近似解を求めることを考える (上の式は (7.112) と同じ.). これを差分化すると

$$L\boldsymbol{\psi} = \boldsymbol{\rho}$$

という形の連立一次方程式になるので (式 (7.113) と同じ) L の逆行列を計算すれば解が求められる. しかし, 行列 L が大きくなり過ぎて数値解を求めることが困難である場合も多い.

そこで, 式 (7.121) を解く代わりに, 適当な関数を ψ の初期値として

$$\frac{\partial\psi}{\partial t} = \mathcal{L}\psi - \rho \tag{7.122}$$

の定常解 ($\partial/\partial t = 0$ の解) を求めるという方法がある. これを, 緩和法という. もちろん, 上の方程式が定常解を持たないようなタイプのものである場合には, この方法は使えない.

以下の節では, 2 次元の Poisson 方程式

$$\nabla^2\psi(x,z) = \rho(x,z) \tag{7.123}$$

を例としてとり扱う.

7.19.2　熱伝導方程式の定常解としての Poisson 方程式の解

Poisson 方程式の場合, 式 (7.122) に相当するのは

$$\frac{\partial\psi}{\partial t} = \nabla^2\psi(x,z) - \rho(x,z) \tag{7.124}$$

である. これは, (熱源 $-\rho$ のある) 熱伝導方程式であるので定常解が存在する. そこで, 適当な初期値から出発して定常になるまで積分すれば, その時の ψ が Poisson 方程式 (7.123) の解である.

7.19.3　解の求め方

(7.124) を解くために差分化する. 例えば,

$$\frac{\psi_{i,j}^{n+1} - \psi_{i,j}^{n}}{\Delta t} = \nabla^2\psi_{i,j}^{n} - \rho_{i,j}$$

すなわち

$$\begin{aligned}\psi_{i,j}^{n+1} &= \psi_{i,j}^{n} + \Delta t(\nabla^2\psi_{i,j}^{n} - \rho_{i,j}) \\ &= \psi_{i,j}^{n} + \frac{\Delta t}{d^2}(\psi_{i+1,j}^{n} + \psi_{i-1,j}^{n} + \psi_{i,j+1}^{n} + \psi_{i,j-1}^{n} - 4\psi_{i,j}^{n} - d^2\rho_{i,j})\end{aligned} \tag{7.125}$$

とする. ただし

$$\Delta x = \Delta z \equiv d$$

7.18.4 方程式の解

Poisson 方程式を差分化した式を Fourier 変換した式 (7.120) を行列の形であらわそう. ψ が z 方向に周期 N_z を持つとする.

$$A \equiv \begin{pmatrix} -\frac{4\sin^2\frac{\pi k}{N_x}}{(\Delta x)^2} - \frac{2}{(\Delta z)^2} & \frac{1}{(\Delta z)^2} & 0 & \cdots & 0 & \frac{1}{(\Delta z)^2} \\ \frac{1}{(\Delta z)^2} & -\frac{4\sin^2\frac{\pi k}{N_x}}{(\Delta x)^2} - \frac{2}{(\Delta z)^2} & \frac{1}{(\Delta z)^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{(\Delta z)^2} & & & & \vdots \\ \vdots & & & \ddots & & 0 \\ 0 & & & & \ddots & \\ \frac{1}{(\Delta z)^2} & 0 & \cdots & 0 & & \ddots \end{pmatrix}$$

$$\tilde{\boldsymbol{\psi}} \equiv \begin{pmatrix} \tilde{\psi}_{k,0} \\ \vdots \\ \tilde{\psi}_{k,N_z-1} \end{pmatrix}, \qquad \tilde{\boldsymbol{\rho}} \equiv \begin{pmatrix} \tilde{\rho}_{k,0} \\ \vdots \\ \tilde{\rho}_{k,N_z-1} \end{pmatrix}$$

とおくと, 式 (7.120) は

$$A\tilde{\boldsymbol{\psi}} = \tilde{\boldsymbol{\rho}}$$

と書ける. よって, $\tilde{\psi}$ は

$$\tilde{\boldsymbol{\psi}} = A^{-1}\tilde{\boldsymbol{\rho}}$$

によって求めることができる. $\tilde{\psi}$ が求まれば, これを逆 Fourier 変換 (7.116) することによって, ψ を求めることができる.

この式に適切な境界条件を入れると[23]行列の形で書くことができて,

$$L \begin{pmatrix} \psi_{1,1} \\ \vdots \\ \psi_{1,N_z} \\ \psi_{2,1} \\ \vdots \\ \psi_{N_x,1} \\ \vdots \\ \psi_{N_x,N_z} \end{pmatrix} = \begin{pmatrix} \rho_{1,1} \\ \vdots \\ \rho_{1,N_z} \\ \rho_{2,1} \\ \vdots \\ \rho_{N_x,1} \\ \vdots \\ \rho_{N_x,N_z} \end{pmatrix}$$

となり, 原理的には解を求めることができるが, 行列が大きくなり過ぎて数値解を求めることが現実には困難である場合も多い. そこで, x 方向に (離散) Fourier 展開を行い, 一回に計算する次元を減らすことを考える.

7.18.3　Fourier 変換

x 方向の波数を k $(k = 0, 1, \ldots, N_x - 1)$ とすると, ψ と, その Fourier 変換 $\tilde{\psi}$は, 次のような関係にある.

$$\psi_{i,j} = \frac{1}{\sqrt{N_x}} \sum_{k=0}^{N_x-1} \tilde{\psi}_{k,j} \exp\left(2\pi \mathbf{i} \frac{ik}{N_x}\right) \tag{7.116}$$

$$\tilde{\psi}_{k,j} = \frac{1}{\sqrt{N_x}} \sum_{i=0}^{N_x-1} \psi_{i,j} \exp\left(-2\pi \mathbf{i} \frac{ik}{N_x}\right) \tag{7.117}$$

ただし, 虚数単位を i と表すことにする. ρ と, その Fourier 変換 $\tilde{\rho}$ についても同様に書ける.

$$\rho_{i,j} = \frac{1}{\sqrt{N_x}} \sum_{k=0}^{N_x-1} \tilde{\rho}_{k,j} \exp\left(2\pi \mathbf{i} \frac{ik}{N_x}\right) \tag{7.118}$$

$$\tilde{\rho}_{k,j} = \frac{1}{\sqrt{N_x}} \sum_{i=0}^{N_x-1} \rho_{i,j} \exp\left(-2\pi \mathbf{i} \frac{ik}{N_x}\right) \tag{7.119}$$

この (7.116) と (7.118) を, Poisson 方程式 (7.114) を差分化した式 (7.115) に代入して整理すると,

$$\begin{aligned} & \frac{1}{\sqrt{N_x}} \sum_{k=0}^{N_x-1} \exp\left(2\pi \mathbf{i} \frac{ik}{N_x}\right) \left\{ -\frac{\tilde{\psi}_{k,j}}{(\Delta x)^2} \cdot 4 \sin^2 \frac{\pi k}{N_x} + \frac{\tilde{\psi}_{k,j+1} - 2\tilde{\psi}_{k,j} + \tilde{\psi}_{k,j-1}}{(\Delta z)^2} \right\} \\ = \quad & \frac{1}{\sqrt{N_x}} \sum_{k=0}^{N_x-1} \tilde{\rho}_{k,j} \exp\left(2\pi \mathbf{i} \frac{ik}{N_x}\right) \end{aligned}$$

となる. よって, 各波数 k について

$$\frac{1}{(\Delta z)^2} \tilde{\psi}_{k,j-1} - \left\{ \frac{4}{(\Delta x)^2} \sin^2 \frac{\pi k}{N_x} + \frac{2}{(\Delta z)^2} \right\} \tilde{\psi}_{k,j} + \frac{1}{(\Delta z)^2} \tilde{\psi}_{k,j+1} = \tilde{\rho}_{k,j} \tag{7.120}$$

が成り立つ.

[23] 式 (7.115) の添字には $i \pm 1$, $j \pm 1$ が含まれているから, 何らかの境界条件がないと $i = 1, \ldots, N_x$, $j = 1, \ldots, N_z$ の全ての点について差分式 (7.115) を成立させることはできない.

7.18 直接法

7.18.1 Introduction

ここでは, 直接法といわれる数値計算法について説明する.

一般に, $\rho = \rho(x, y, z)$ が与えられているとき,

$$\mathcal{L}\psi = \rho, \quad (\psi = \psi(x, y, z)) \tag{7.112}$$

という形の偏微分方程式の近似解を求めることを考える. ここに, $\mathcal{L}$ は線形の微分演算子である. xyz 空間を格子に切って上の式を差分化すると

$$L\boldsymbol{\psi} = \boldsymbol{\rho} \tag{7.113}$$

という形の連立方程式になる. ただし, L は行列であり, $\boldsymbol{\psi}$ は格子点上での $\psi(x, y, z)$ の値 ψ_{ijk} を全部並べた縦ベクトルである. $\boldsymbol{\rho}$ についても同様とする. よって, L の逆行列を求めて (7.113) を解いてやれば, 偏微分方程式 (7.112) の近似解が求まることになる. このような数値計算法を直接法という.

以下の節では, 2 次元の Poisson 方程式

$$\nabla^2\psi(x, z) = \rho(x, z) \tag{7.114}$$

を例としてとり扱う.

7.18.2 格子の設定と方程式の差分化

計算範囲を $0 \leq x \leq \lambda,\ 0 \leq z \leq 1$ とする. この計算範囲を x 軸に垂直に N_x 等分, z 軸に垂直に N_z 等分し, 右図のような長方形状の格子をつくる. その格子の間隔を, x 方向に $\Delta x(= \lambda/N_x)$, z 方向に $\Delta z(= 1/N_z)$ とする. つまり, 格子点の座標は, $(x_i, z_j),\ i = 0, 1, \ldots N_x;\ j = 0, 1, \ldots N_z$ となる.

そして座標 (x_i, z_j) における ψ を

$$\psi_{i,j} = \psi(x_i, z_j)$$

のように書くことにする. ρ についても同様に書く.

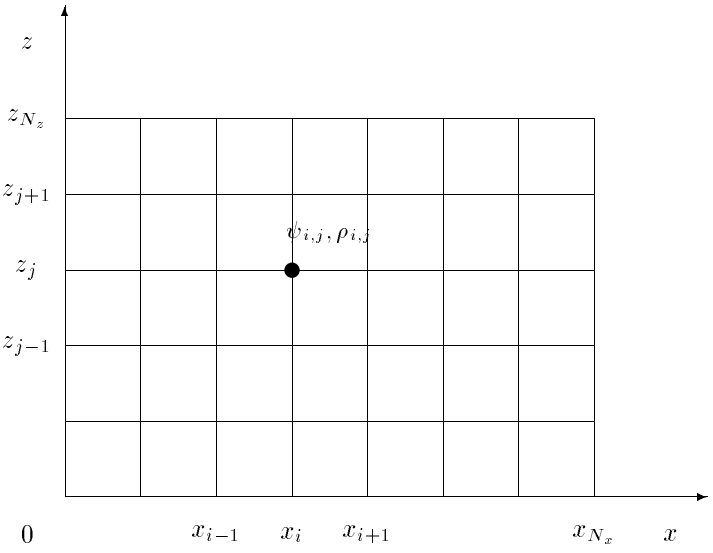


図 1: 格子

このように格子を設定することによって, Poisson 方程式 (7.114) は次のように差分化することができる.

$$\frac{\psi_{i+1,j} - 2\psi_{i,j} + \psi_{i-1,j}}{(\Delta x)^2} + \frac{\psi_{i,j+1} - 2\psi_{i,j} + \psi_{i,j-1}}{(\Delta z)^2} = \rho_{i,j} \tag{7.115}$$

と呼ばれる. Arakawa は, これら 3 つの性質を保存する 4 次の accuracy を持つスキームの作り方も示した. また, プリミティブ方程式中の移流項に対する近似も, 流速場が 2 次元非発散· 発散両方の場合に対して作られている. 最近, Arakawa Jacobian は有限要素法の特別な場合であることが示されている. (Jespersen, 1974)

$$= \frac{\partial}{\partial y}\left(q\frac{\partial p}{\partial x}\right) - \frac{\partial}{\partial x}\left(q\frac{\partial p}{\partial y}\right) \tag{7.108}$$
$$= \frac{\partial}{\partial x}\left(p\frac{\partial q}{\partial y}\right) - \frac{\partial}{\partial y}\left(p\frac{\partial q}{\partial x}\right)$$

それぞれに対する有限差分近似をうまく線形結合させてやることで (7.107) を満たすようにしよう. ここでは accuracy が 2 次のもののみを考えることにする. 空間微分を近似して (7.108) の値を求めるにあたり, 最も単純な中心差分を使うことにすると, 9 個の近接した格子点における p, q の値が必要となる. 格子間隔を d とし, 格子点 k における p, q の値を p_k,q_k とすると, 式 (7.108) の 3 つの式の近似として

$$\begin{aligned} J^{++}(p,q) &= \frac{1}{4d^2}\left[(p_1-p_3)(q_2-q_4)-(p_2-p_4)(q_1-q_3)\right] \\ J^{\times+}(p,q) &= \frac{1}{4d^2}\left[q_2(p_5-p_6)-q_4(p_8-p_7)-q_1(p_5-p_8)+q_3(p_6-p_7)\right] \\ J^{+\times}(p,q) &= \frac{1}{4d^2}\left[p_1(q_5-q_8)-p_3(q_6-q_7)-p_2(q_5-q_6)+p_4(q_8-q_7)\right] \end{aligned} \tag{7.109}$$

が得られる. 添字 $+$, $times$ は, どの格子点での値 p, q を使って近似を行ったか, ということを表す. 3 つの近似 (7.109) は, 全て consistent で, 2 次の accuracy をもつ. ここで, 線形結合

$$J(p,q) = \alpha J^{++} + \beta J^{\times+} + \gamma J^{+\times} \quad (\alpha+\beta+\gamma=1) \tag{7.110}$$

を考える. これも, consistent で 2 次の accuracy を持つヤコビアンの近似である. (7.110) を使って (7.107) の和を求めようとすると, 計算領域内の各点について 24 個の項がでてくる. (J^{++}から 8 個, $J^{\times+}$から 8 個, $J^{+\times}$から 8 個) これらの項は全て const. $\cdot\zeta_k\zeta_l\Psi_m$ という形をしている. ここで, 定数 α, β, γ を適当に選べば, 境界を除いて, 和をとったときこれらの項が全て打ち消し合う様に出来る. このとき, (7.107) は近似的に満たされる. 例えば, $k=0$ の点は, (7.107) に対して

$$\begin{aligned} \zeta_0 J_0(\zeta,\Psi) &= \frac{1}{4d^2}\zeta_0\{\alpha J^{++}(\zeta,\Psi)+\cdots\} \\ &= \frac{1}{4d^2}\{\alpha\zeta_0((\zeta_1-\zeta_3)(\Psi_2-\Psi_4)-\cdots)+\cdots\} \\ &= \frac{1}{4d^2}(\alpha\zeta_0\zeta_1\Psi_2 + 23 \text{ more terms}) \end{aligned}$$

という形で効いている. ここで, $\zeta_0\zeta_1\Psi_2$ という項は, $\zeta_1 J_1(\zeta,\Psi)$ にもでてくる. ヤコビアンの近似 (7.109) の形からみると, この項は $J^{+\times}$ 中の項 p_3q_6 からきている. これら 2 つの項は $\alpha=\beta$ ならば打ち消しあう. 同様に考えれば, 結局,

$$\alpha=\beta=\gamma=\frac{1}{3}$$

であれば, 式 (7.107) の第 2 式は, 項同士が打ち消し合って 0 となる. すなわち, 式 (7.107) が成立する. のみならず,

$$\zeta_1 J_1(\zeta,\Psi) = \frac{1}{4d^2}(\cdots - \beta\zeta_1\zeta_0\Psi_2+\cdots)$$

ならば, 式 (7.107) が成立するだけでなく, 平均の運動エネルギー, 平均の渦度保存の式においても, 項同士が全て打ち消し合って 0 となる, ということを示される (Arakawa, 1966; Lilly, 1965). よって, (7.96) に対して近似

$$J_A \equiv \frac{1}{3}(J^{++}+J^{\times+}+J^{+\times}) \tag{7.111}$$

を使えば, 平均の渦度, enstrophy, 運動エネルギーが保存される. これは, 非線形不安定性を予防するという観点からすれば, 十分過ぎるものである. この近似式 (7.111) は通常

Arakawa Jacobian

という式が成り立つ. そこで, p, q のどちらかが領域 A の境界上で定数ならば, Stokes の定理より

$$\overline{J(p,q)} = 0 \tag{7.105}$$

となる. 同じ境界条件で

$$\overline{pJ(p,q)} = 0, \quad \overline{qJ(p,q)} = 0 \tag{7.106}$$

も成立する. よって, Ψ が境界上で定数ならば, (7.103), (7.104) より

$$\frac{1}{2}\overline{\zeta^2} = \text{const.} \quad \text{かつ} \quad \overline{K} = \text{const.}$$

となる. ゆえに, 式 (7.102) より, 2 次元非発散では平均の波数は時間的に変化しないことが分かる. 当然, 系統的により高波数へとエネルギーが流れるようなことはおこらない.

その上, 高波数へ大きなエネルギーを移すのは難しい. というのは, エネルギーを移す時にも平均の enstrophy が保存するように移さなくてはならないが, 式 (7.101) より, enstrophy の高波数の成分を作るためには K_n と波数の 2 乗を掛け合わせねばならないから, 高波数へは少ししかエネルギーを移せないからである. (Fjørtoft, 1953) このことを, Charney, 1966 に従って力学のアナロジーを用いて説明してみよう. 式 (7.102) を

$$\overline{K}\lambda^2 = \sum_n K_n \lambda_n^2 = \text{const.}$$

と変形してみる. 左辺の 2 つの要素は定数である. そこでこの状況は, 重さの無い半無限の棒の左側には支点から距離λ^2 の所に重さ $\overline{K}$ のおもりが, 右側には距離$\lambda_1^2,\ \lambda_2^2,\ \lambda_3^2, ldots$ の所に重さ $K_1,\ K_2,\ K_3, \ldots$ のおもりがぶら下げられて, バランスを取っているのと同じ, と考えられる. 左側のおもりは動かせない. 右側のおもりは動かしてもいいが, バランスを崩してはいけない. おもりを付けてよい位置 $\lambda_1^2,\ \lambda_2^2,\ \lambda_3^2, \ldots$ は変えられないから, エネルギー (おもり) を動かすときには少なくとも 3 つ以上の要素をいっぺんに動かさねばならない. 特に, 高い周波数へは少ししかエネルギーを動かしてはいけないことが分かる.

(7.96) の数値解と, その非線形不安定にの話題に戻ろう. これまでの議論から, enstrophy と運動エネルギーの平均値が保存されるなら平均の波数は変化しないことがわかった. そこで, そのような性質を持つ有限差分スキームを作ってやれば, 高い周波数へ系統だってエネルギーが流れるようなことは起きない. Arakawa は, (7.106) を満たすようなヤコビアンの有限差分近似が作れることを示した. このようなスキームを使えば, 平均の enstrophy と運動エネルギーは保存し, 従って, 平均の波数も存するので, 非線形不安定性は予防される. その上, この近似では (7.105) も保存されるので, 平均の渦度も保存される. よって, 渦度の周波数分布が大局的にみて保存される. 差分近似をしない真の非発散渦度方程式では, 個々の流体粒子の渦度, 及びその面積が保存するから, 渦度の周波数分布の全モーメントは保存する. その有限差分近似に置いては, (7.105), (7.106) が成り立つので, 1 次と 2 次のモーメントは保存する.

7.17.3　Arakawa の方法

Arakawa の方法とは, この場合, (7.106) を満たすように, すなわち

$$\overline{\zeta_{ij} J_{ij}(\zeta,\ \Psi)} = \frac{1}{N}\sum_{i,j} \zeta_{ij} J_{ij}(\zeta,\ \Psi) = 0 \tag{7.107}$$

を満たすように近似をすることである. 但し, J はヤコビアンの有限差分近似を表し, N は格子点の総数を表す.

ヤコビアンに対する有限差分近似は沢山あるが, 3 つの等価な式

$$J(p,q) \quad = \quad \frac{\partial p}{\partial x}\frac{\partial q}{\partial y} - \frac{\partial p}{\partial y}\frac{\partial q}{\partial x}$$

エンストロフィー

Enstrophy の平均は

$$\frac{1}{2}\overline{\zeta^2} = \frac{1}{2}\overline{(\nabla^2\Psi)^2}$$

であるが, これも同様に分解できる. (7.98) 代入し, (7.97) を使って直交関数系 $\{\Psi_n\}_n$ で展開すれば

$$\frac{1}{2}\overline{\zeta^2} = \frac{1}{2}\sum_n \lambda_n^4 \overline{\Psi_n^2}$$

となる. ここで, 式 (7.100) を用いると

$$\frac{1}{2}\overline{\zeta^2} = \sum_n \lambda_n^2 K_n \tag{7.101}$$

であることがわかる.

平均の波数

これを (7.99) と比較すれば, 波数の平均値は平均の enstrophy · 平均の運動エネルギーと関係があることが分かる. 波数の平均値 λ をエネルギーの重みが付いた平均

$$\lambda \equiv \sqrt{\frac{\sum_n \lambda_n^2 K_n}{\sum_n K_n}}$$

によって定義する. この式に (7.101), (7.99) を代入すれば

$$\lambda = \sqrt{\frac{\frac{1}{2}\overline{\zeta^2}}{\overline{K}}} \tag{7.102}$$

が得られる. よって

速度が 2 次元非発散の時, 平均の波数は平均の enstrophy と
運動エネルギーとの比で決まる！

我々が知りたいのは, そもそも, 渦度方程式 (7.96) に支配される系では, エネルギースペクトルは時間的にどのように変化するか, ということであった. このために, 平均の波数 (7.102) の時間依存性を調べてみよう.

まず, 平均運動エネルギーと平均 enstrophy の時間変化を計算する. (7.96) より

$$\frac{\partial}{\partial t}\frac{1}{2}\overline{\zeta^2} = \overline{\zeta\frac{\partial}{\partial t}\zeta} = \overline{\zeta J(\zeta,\ \Psi)} \tag{7.103}$$

である. 一方, 再び境界条件 $\overline{\nabla\cdot(\Psi\nabla\Psi)} = 0$ を仮定すると

$$\begin{aligned}\frac{\partial}{\partial t}\overline{K} &= \frac{\partial}{\partial t}\frac{1}{2}\overline{(\nabla\Psi)^2} = \overline{\nabla\Psi\cdot\frac{\partial}{\partial t}\nabla\Psi} \\ &= -\overline{\Psi\frac{\partial}{\partial t}\nabla^2\Psi} = -\overline{\Psi J(\zeta,\ \Psi)}\end{aligned} \tag{7.104}$$

となる. ところが, 任意の 2 つのスカラー量 p, q の間には

$$J(p,q) = \boldsymbol{k}\cdot\nabla\times(p\nabla q) = -\boldsymbol{k}\cdot\nabla\times(q\nabla p)$$

となる. よって

$$\lambda^2_{n_1,n_2} = \left(\frac{2\pi n_1}{L_x}\right)^2 + \left(\frac{\pi n_2}{L_y}\right)^2$$

である. 領域の形や境界条件が異なれば, 当然 (7.97) の解の形は変わる.

再び一般の領域 · 一般の境界条件の場合に戻る. ある量 α の平均を

$$\overline{\alpha} \equiv \frac{1}{A}\int_A \alpha\, dA$$

と定義する. ここで知りたいのは, 単位質量当りの運動エネルギーの平均

$$\overline{K} = \frac{1}{2}\overline{(u^2+v^2)} = \frac{1}{2}\overline{\nabla\Psi\cdot\nabla\Psi}$$

である. そこで, 上の式に (7.98) を代入し, 項別微分, 項別積分が可能であると仮定すると

$$\overline{K} = \frac{1}{2}\overline{\nabla\sum_n \Psi_n\cdot\nabla\sum_m \Psi_m} = \frac{1}{2}\overline{\sum_n \nabla\Psi_n\cdot\sum_m \nabla\Psi_m} = \frac{1}{2}\sum_m\sum_n \overline{\nabla\Psi_m\cdot\nabla\Psi_n}$$

となる. さて,

$$\nabla\Psi_m\cdot\nabla\Psi_n = \nabla\cdot(\Psi_m\nabla\Psi_n) - \Psi_m\nabla^2\Psi_n$$

である. また,

$$\overline{\nabla\cdot(\Psi_m\nabla\Psi_n)} = \frac{1}{A}\oint \Psi_m v_{t,n}\, dl = 0$$

とを満たすような境界条件を仮定する[22] ここに, $\oint dl$ は境界を一周する線積分, v_t は境界に沿う方向の速度成分である. よって, 境界から質量の出入りがなく境界上で $\Psi = 0$ とおける場合や, 周期境界条件の場合などには, 上の境界条件は満たされている. このとき (7.97) を使うと

$$\overline{K} = -\frac{1}{2}\sum_m\sum_n \overline{\Psi_m\nabla^2\Psi_n} = \frac{1}{2}\sum_m\sum_n \lambda_n^2\overline{\Psi_m\Psi_n}$$

を得る. さらに, $\{\Psi_n\}_n$ が直交系を成すことを用いれば

$$\overline{K} = \frac{1}{2}\sum_n \lambda_n^2\overline{\Psi_n^2}$$

となる. 結局, 領域 A における平均の運動エネルギーは, 波数毎のエネルギーの和として

$$\overline{K} = \sum_n K_n \tag{7.99}$$

と書けることが分かる. ここに,

$$K_n \equiv \frac{1}{2}\lambda_n^2\overline{\Psi_n^2} \tag{7.100}$$

である. K_n は, 波数 λ_n の成分の平均の運動エネルギーへの寄与, すなわち, 運動エネルギースペクトルである. (7.100) より, 当然, 各 K_n は負にはならない.

もし流線関数 Ψ が与えられたなら, これを $\{\Psi_n\}_n$ の和に展開することができる. そして, よく使われる形の領域に対しては波数λ_n が計算されているから, (7.100) を使えば運動エネルギースペクトルが計算できることになる. このような方法は, ほかのスペクトルの計算と共に, 観測や数値計算によって得られた様々な変数の値の振舞いや構造を研究するのに, 気象学では非常によく使われている.

[22] このような境界条件の場合, $\{\Psi_n\}_n$ は直交系を成すことが示せる.

7.17 Arakawa スキーム

7.17.1 渦度方程式

Arakawa の方法を導入するにあたって, 渦度方程式

$$\frac{\partial \zeta}{\partial t} + \boldsymbol{v} \cdot \nabla \zeta = 0; \quad \zeta = \nabla^2 \Psi \tag{7.94}$$

$$\boldsymbol{v} = (u, v); \quad \nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)$$

を考える. ただし, 速度 $\boldsymbol{v}$ は非発散 $(\nabla \cdot \boldsymbol{v} = 0)$ であり,

$$\boldsymbol{v} = \left(-\frac{\partial \Psi}{\partial y}, \frac{\partial \Psi}{\partial x} \right) \tag{7.95}$$

と書かれている. これを (7.94) に代入すると

$$\frac{\partial}{\partial t} \nabla^2 \Psi = J(\nabla^2 \Psi, \ \Psi) \tag{7.96}$$

となる. ただし, J はヤコビアンである. この方程式は, 2 次元非発散の速度による移流の結果としての局所的な渦度の変化を表す非線形移流方程式である. この方程式に小さな現象を現す付加的な項を補ったものが, 10 年以上にわたって天気予報における基礎的な方程式として使われてきた.

7.17.2 渦度方程式の波数空間における性質

渦度方程式 (7.96) に対する Arakawa の方法を説明するためには, 解の波数空間における性質を少し知らねばならない. この方程式のもとでは, 異なった周波数間でどのようにエネルギーがやり取りされるのかを調べてみよう.

運動エネルギー

最初に, 運動エネルギースペクトルを考える. 領域 A 内での Helmholtz の方程式

$$\nabla^2 \Psi_n + \lambda_n^2 \Psi_n = 0 \tag{7.97}$$

の固有解の集合である完全直交関数系 $\{\Psi_n\}_n$ を用いて[21] 流線関数を

$$\Psi = \sum_n \Psi_n \tag{7.98}$$

のように展開する.

この完全直交関数系が具体的にはどのようなものであるか見るために, 例として, 領域 A を一辺が L_x , L_y の長方形であるとしてみよう. x 方向の境界条件は, 周期 L_x の周期的境界条件, また, 上下の境界では 0 とする. この時, Helmholtz の方程式 (7.97) の解は

$$\Psi_{n_1,n_2} = \left(a_{n_1,n_2} \cos \frac{2\pi n_1}{L_x} x + b_{n_1,n_2} \sin \frac{2\pi n_1}{L_x} x \right) \sin \frac{\pi n_2}{L_y} y$$

となる. これを微分すれば

$$\nabla^2 \Psi_{n_1,n_2} = -\left[\left(\frac{2\pi n_1}{L_x} \right)^2 + \left(\frac{\pi n_2}{L_y} \right)^2 \right] \Psi_{n_1,n_2}$$

[21] パラメーター λ_n は, 成分 Ψ_n の一般化された波数という. 各 λ_n^2 は負にはならないことが示せる.

式に対してしか用いることが出来ない.) Arakawa の保存スキームを使うときは, 移流項を人工的に減衰させる必要はない. このため, 移流項の統計的特性 — これは大気大循環の研究に役立つ特性であるが — が保存できるようになった.

短波長の波の位相の誤差は大きいから, これが非線形相互作用によって長波長の波の位相に影響を与える前に除去すべきである, という議論がときどき行われる. この議論は, いくつかの点を見落としている. たとえ短波長の波の位相速度が狂っていたとしても, 波の強度もまた狂っていない限り状況が悪くなるとは限らない. というのは, 地球物理的なシミュレーションにおいては, 興味があるのは, 個々の波がどれだけ正確に再現されているかどうかではなくてそれらが寄り集まった振舞い— 例えばエネルギーが全体としてどのように輸送されるか — といった問題である場合も多いからである. このような場合には, 統計をとったときに意味のある関数が出てくれば十分なので, 位相は少々狂っても統計的には使える関数を生み出すことはある. また, 興味がないからといって短波長の波をカットすると, 興味の対象である長波長の波のエネルギーまで一緒に減衰させてしまうこともある. 物理的意味もないのに計算の都合だけでエネルギーを減衰させることが好ましい訳はないのである.

7.16　非線形移流方程式の不安定性

7.16.1　非線形不安定性

Aliasing error が数値積分に及ぼす影響を考えてみる.

エネルギーのパワースペクトルを考えると, これは, 異なったスケールの現象の, 相対的な重要度を表す. 有限差分法においては, aliasing error によって, 物理的には存在しない過程によってもスペクトルの形が変化してしまうことが問題である. 第 7.15 節で述べたように, 最初は全ての波が k_{max} 以下であっても, 非線形項によって, k_{max} 以上の波がどんどん作られ, これが aliasing によって k_{max} 以下の別の波数を持つ波として取り扱われてしまう. これは, k_{max} 以下の波において偽のエネルギー流入が起きていることに相当する.

我々が扱うような問題では大きな波数 ($k \sim 0$) の成分が卓越するのが普通である. そのような成分 k_1, k_2 の相互作用で生じる波の波数 $k_1 + k_2$ は k_{max} より少し大きいだけであるから, aliasing error によって起こる偽のエネルギーの流入は k_{max} より少し小さい波数, すなわち高波数の部分で起こる (式 (7.93) で $n = 1$ として考えよ.). 積分を開始してからある程度時間が経過すると, 高波数のエネルギーがどんどん増えパワースペクトルは物理的に到底受け入れがたいような値となってしまう. さらに放置しておくと (経験によれば), ついには発散して計算不能となる. これは非線形項によって引き起こされる現象なので,

非線形不安定性 (nonlinear instability)

という. この現象は, Phillips, 1956 によって発見された.

7.16.2　非線形不安定性の抑止と予防

短時間の積分ならば非線形不安定を抑制しなくて済む場合もあるが, 長時間にわたって積分するならば, 非線形不安定性を抑制, ないしは予防する手段を考えねばならない.

格子点で表現できる $0 \leq |k| \leq k_{max}$ の波の内, 大きい波数の部分をカットすることによって aliasing error を防ごうとするなら, $\frac{2}{3}k_{max} < |k| \leq k_{max}$ の部分を落とせば十分である. なぜなら, $\frac{2}{3}k_{max} < |k| \leq k_{max}$ ならば, 発生する aliasing error もまた $\frac{2}{3}k_{max} < |k^*| \leq k_{max}$ となり (式 (7.93) で $n = 1$ として考えよ.), 次のステップで落とされてしまうからである. (Orszag, 1971)

この方法のように大きな波数の部分をを落としてしまって構わないのなら, いちいちフィルターをかけて高い周波数をカットするより, 何もせずとも大きな波数を減衰させてしまうスキーム (例えば Lax-Wendroff scheme) を使った方が簡単である (Richtmyer, 1963). こういったスキームを使う場合, 必ずしも最初から最後までこれ一つで通さなければならないわけではなく, 他のスキームの中に, ときたまこうしたスキームを入れてやれば十分である (Kasahara, 1969). しかし, このようなスキームでは減衰量を調節できないから, 本来減衰しないスキームに減衰項を付け加えた方がいい, という意見もある (Kreiss and Oliger, 1973).

非線形不安定を避けるもう一つのやり方は, 移流項の積分に Lagrange 的な考え方を使うことである. 時間ステップ Δt で, 1 群の格子点が移動するとそれに伴って様々な値が変化する . その値を計算する. また , その場所に対応する従属変数の値は, 空間的に内挿することによっても求められる. 移流による値の変化は, 内挿による値と格子点における値との差に等しいと置く. こうして得られるスキームは, Euler 的なスキームと同じ場合もあるが, 違ったスキームが得られるときもある. このようなやり方は最初 Leith, 1965 によって用いられた. 新しい例では, Krishnamurti et al., 1973 がある.

非線形不安定を取り扱うもう一つの考え方として, Arakawa の保存スキームを使う方法がある. これは, 渦度方程式においては enstrophy[20] が保存するという性質を利用するものである. (当然, これは渦度方程

[20] 渦度の 2 乗の 1/2, すなわち $\frac{1}{2}(\nabla \times v)^2 = \frac{1}{2}(\nabla^2\Psi)^2$ を enstrophy という.

7.15　非線形移流方程式のおばけ

非線形移流方程式

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0, \qquad u = u(x,t) \tag{7.92}$$

を考える. これは, shock equation とも呼ばれている (Shuman, 1974).

この一般解は,

$$u = f(x - ut)$$

である (Platzman, 1964 他). f は任意の関数である.

ここでは, 方程式 (7.92) の非線形項 $u\dfrac{\partial u}{\partial x}$ が引き起こす問題について述べる.

方程式 (7.92) を有限差分法で解こうとすると, aliasing error (「おばけ」) という誤差が発生する. これは, 格子点法では波長が $2\Delta x$ 以下, すなわち波数にして $k_{max} \equiv \dfrac{\pi}{\Delta x}$ 以上の波は表現できなところから生ずる誤差である. 例として波長 $l = \dfrac{4}{3}\Delta x$ の正弦波を考える. 格子点での値しか分かっていないとすると, この波は波長 $4\Delta x$ の正弦波と区別がつかない. これが aliasing error である.

次に 非線形項がどのようにして aliasing error を発生させるか見てみよう. u を

$$u = \sum_n u_n e^{ik_n x} \quad (k_{-n} = -k_n, \;\; u_{-n} = \overline{u_n})$$

と, 調和解の和で展開する. このとき, $u\dfrac{\partial u}{\partial x}$ には $e^{i(k_1+k_2)x}$ や $e^{i(k_1-k_2)x}$ に比例する項が含まれるから, 最初, 全ての波が $k \le k_{max}$ であっても, すぐに $k > k_{max}$ なる波が発生してしまうことが分かる.

さて, どの様な波が aliasing error の影響を受けるかを考えてみる. 非線形効果の結果, 波数 $k > k_{max}$ なる波が生じたとしよう. 任意の整数 n に対して

$$\exp[ikx] = \exp[i\{2nk_{max} - (2nk_{max} - k)\}x]$$

であるが, x が $x = j\Delta x$ (j は整数) と離散化されていることと $k_{max} = \dfrac{\pi}{\Delta x}$ を考えると

$$\begin{aligned} \exp[ikx] &= \exp[i2n\frac{\pi}{\Delta x}j\Delta x - i(2nk_{max} - k)x] \\ &= \exp[-i(2nk_{max} - k)x] \end{aligned}$$

よって, もし次のような波数 k^*

$$k^* \equiv 2nk_{max} - k \tag{7.93}$$

が $|k^*| \le k_{max}$ を満たしていれば (n の値を調整すれば常に可能である.), 波数 k であった波は, 有限差分の世界では波数 k^* なる波として認識 (サンプリング) されることになる. これが, 非線形効果による aliasing error である.

7.14　2 次元移流方程式の安定性

2 次元線形移流方程式

$$\frac{\partial u}{\partial t} + c_x \frac{\partial u}{\partial x} + c_y \frac{\partial u}{\partial y} = 0, \quad u = u(x, y, t), \quad c_x, c_y = \text{const.} \tag{7.90}$$

を考える. 移流の速さ c は

$$c = \sqrt{c_x^2 + c_y^2}$$

である. 1 次元移流方程式の場合と同様にして (7.90) の安定性を調べる. まず, 空間微分を 2 次の差分で近似して

$$\frac{\partial}{\partial t} u_{i,j} = -c_x \frac{u_{i+1,j} - u_{i-1,j}}{2\Delta x} - c_y \frac{u_{i,j+1} - u_{i,j-1}}{2\Delta y}$$

とする. ただし, 近似を行うとき $x = i\Delta x, \quad y = j\Delta y$ とし, $u(i\Delta x,\ j\Delta y)$ を $u_{i,j}$ と書いた. 解として, 仮に

$$u_{i,\ j} = Re[U(t)e^{i(kx+ly)}]$$

を代入すれば, 振動方程式

$$\frac{dU}{dt} = i\left(-\frac{c_x}{\Delta x}\sin k\Delta x - \frac{c_y}{\Delta y}\sin l\Delta y\right)\ U$$

が得られる. 時間微分の近似として, 例えば leapfrog スキームを取るならば, 安定であるための条件は, 許される任意の k, l に対して

$$\left|\left(\frac{c_x}{\Delta x}\sin k\Delta x + \frac{c_y}{\Delta y}\sin l\Delta y\right)\ \Delta t\right| \le 1 \tag{7.91}$$

が成立することである.

話を簡単にするため, $\Delta x = \Delta y \equiv d^*$, $c_x = c_y$ $\left(= \dfrac{c}{\sqrt{2}}\right)$ の場合のみを考える[19]. kd^*, ld^* を座標に取った波数平面を考えると, (7.91) の左辺が最大となるのは, $kd^* = ld^* = \dfrac{\pi}{2}$ のときである (このとき, 波長の x 成分, y 成分は共に $4d^*$ となる.). よって, 安定であるための条件は

$$\sqrt{2}c\frac{\Delta t}{\Delta x} \le 1$$

となる. ゆえに, $\dfrac{\Delta t}{\Delta x}$ は, 1 次元の時の値の $\dfrac{1}{\sqrt{2}}$ 倍以下でなければならない.

最も不安定となりやすいのは, 1 次元の時と同じく, 座標軸方向の波長が最短波長の 2 倍, $4d^*$ の時である. 波数ベクトルの大きさは $\sqrt{k^2 + l^2}$ であるので, 波の実際の波長は $\dfrac{4d^*}{\sqrt{2}}$ である.

[19] $c = \sqrt{c_x^2 + c_y^2} = \text{const.}$ という条件の下では, 不等式 (7.91) の左辺は, $c_x = c_y = \dfrac{c}{\sqrt{2}}$ の時に最大となる.

7.13.3　補足

正確度を上げるため, たくさんの格子点を用いるスキーム (すなわち, 高次のスキーム) を使うと

- 境界条件の扱いがややこしくなるという問題
- computational mode の問題

が生じる.

前者の問題のため高次のスキームに否定的な見解を持つのは特に数学屋に多い. 3 次以上のスキームでは得体の知れない境界条件を使わなくてはならない, というのがその根拠である. 実際 , 下手な境界条件が問題を引き起こすことは良くある. また, 移流方程式に対する 4 次のスキームが実際の予報に有効であると考える人々もいる. 中緯度の天気予報では, 高気圧, 低気圧 (すなわち, 波長の長い波) の位相の進行を正しく予報することが重要とされるからである. 一方, 高次のスキームは大気大循環モデルに対しては, このスキームは余り効果が無いと考えられている. この場合, スキームの正確度を上げるよりは, 格子間隔を小さくすることに努力した方がいいとされているのである (Thompson, 1961, p.157).

格子間隔と同じスケールの現象を扱うための本当に満足の行く有限差分法は, 結局の所存在しない. 格子間隔を小さくするのが確実な方法である. この場合問題になるのは, 計算時間が長くなることである. 格子間隔を 1/2 にすると, 2 次元問題では, 計算量は空間について 4 倍, さらに, CFL condition を満たすためにタイムステップも 1/2 にしなくてはならないので計算量は結局 8 倍なる. 1970 年代前半では, 移流項に対する空間的な打ち切り誤差は, 短期の天気予報に対する最大の誤差原因と見積られている. これについては, Anderson and Fattahi, 1974 参照. リファレンスもこれに豊富に載っている.

7.13　4 次の正確度を持つ空間中心差分スキーム

7.13.1　Introduction

空間差分に伴う様々な問題, 特に位相の誤差と計算上の発散 · 減衰という現象を避けるための 1 つの方法は, 近似の正確度を上げることである. ここでは, 4 次の正確度を持つスキームを取り上げる.

近似 u_j を Taylor 級数に展開し, 有限差分商に代入すると,

$$\frac{u_{j+1}-u_{j-1}}{2\Delta x}=\frac{\partial u}{\partial x}+\frac{1}{3!}\frac{\partial^3 u}{\partial x^3}(\Delta x)^2+O[(\Delta x)^4] \tag{7.87}$$

となる. 同様に

$$\frac{u_{j+2}-u_{j-2}}{4\Delta x}=\frac{\partial u}{\partial x}+\frac{4}{3!}\frac{\partial^3 u}{\partial x^3}(\Delta x)^2+O[(\Delta x)^4] \tag{7.88}$$

この 2 つの式から, $(\Delta x)^2$ の項を消去すれば,

$$\frac{4}{3}\frac{u_{j+1}-u_{j-1}}{2\Delta x}-\frac{1}{3}\frac{u_{j+2}-u_{j-2}}{4\Delta x}=\frac{\partial u}{\partial x}+O[(\Delta x)^4] \tag{7.89}$$

となり, このスキームが, 4 次の正確度を持つことがわかる.

近似 (7.89) は, 中心からの距離 d が, Δx よりも小さい点 (この場合, $d=\frac{2\Delta x}{3}$) における近似を, (7.87), (7.88) から外挿したものとも解釈できる. もちろん, d を小さくすれば近似の精度が常に上がるわけではない.

7.13.2　位相速度に対する効果

近似 (7.89) の, 移流方程式の位相速度に対する効果を見てみよう. (7.89) を

$$\frac{\partial u}{\partial t}+c\frac{\partial u}{\partial x}=0$$

に代入すれば

$$\frac{\partial u_j}{\partial t}+c\left(\frac{4}{3}\frac{u_{j+1}-u_{j-1}}{2\Delta x}-\frac{1}{3}\frac{u_{j+2}-u_{j-2}}{4\Delta x}\right)=0$$

となる. ここへ, 調和解

$$u_j(t)=Re[U(t)e^{ikj\Delta x}]\quad(k=p\frac{\pi}{J\Delta x};\ \ p=-J,-J+1,\ldots,-1,0,1,\ldots,J-1,J)$$

を代入すると, 数値解の位相速度 c^{**} は

$$c^{**}=c\left(\frac{4}{3}\frac{\sin k\Delta x}{k\Delta x}-\frac{1}{3}\frac{\sin 2k\Delta x}{2k\Delta x}\right)$$

となる. 2 次の空間微分の時は,

$$c^{*}=c\frac{\sin k\Delta x}{k\Delta x}$$

であった (式 (7.66) 参照). それぞれを, $k\Delta x=0$ の周りで展開すると,

$$\begin{aligned}c^{**}&=c\left(1-\frac{4}{5!}(k\Delta x)^4+\cdots\right)\\ c^{*}&=c\left(1-\frac{1}{3!}(k\Delta x)^2+\cdots\right)\end{aligned}$$

となる. よって, 位相の遅れはやはり存在するが, 小さな k に対しては, 4 次の差分の遅れは 2 次のものに比べ, ずっと小さくなっている.

をとる. よって, $|\lambda|$ の最小値は, 波長が長くなるにつれ, 単調に増大する.

このスキームの欠点は, 減衰が空間差分の格子間隔 Δx のみならず, タイムステップ Δt, 移流項の速度 c にもよるということである[18].

長所としては, 2 次の正確度を持ち, explicit で, 適当な格子をとれば安定であり, two level スキームであるから computational mode は存在しないといった事が挙げられる. これだけの長所を備えたスキームは他にないので, 気象学ではかなり広く使われている. スキームによる減衰は, 物理的なエネルギー散逸に比べて小さいときは問題ない. むしろ, 最短波長を落とせるので役に立つ. ただし, このどちらでも無いときは, 中立なスキームを使うべきである.

7.12.5　補足

スキームの選択においては, いい特徴を組み合わせればいいのであって, 積分の途中でスキームを変えることもよく行われる. 例えば National Center for Atmosphric Research, Boulder, Colo. の大気モデルでは, 減衰量を小さくし, かつ短波長を落とすため, 100 回に 1 回の割合で Lax-Wendroff スキーム を使い, 他は中立な leapfrog スキーム を使っている.

[18] 安定性, 正確性を追求すると, 減衰量を最適にするように格子をとることが必ずしもできない

と表される.

Lax-Wendroff スキーム は, 空間, 時間共に 2 次の正確度をもつ.

(7.86) は, 右辺第 2 項の添え字等を除けば Matsuno スキーム (7.83) と一致する. (7.86) の右辺第 2 項は, Δt を固定して$\Delta x \to 0$ とすると $(1/2)c^2\Delta t(\partial^2 u/\partial x^2)$ となり, これもやはり式 (7.83) の場合と同じく, この項は解を減衰させる. ただし, 今度の場合は, 波長が $2\Delta x$ の場合が最も減衰が激しい. 数値計算の場合, 面倒なことはたいてい波長が $2\Delta x$ 付近の波で起きるため, これは多くの場合望ましい性質として受け入れられている. 波長が $2\Delta x$ 付近の波で起きるトラブルは, 多くの場合, エネルギー散逸スキームを用いれば解決する. これは, 波長 $2\Delta x$ 付近の波を選択的におとすスキームである.

Lax-Wendroff スキームの安定性を調べる.

$$u_j^n = Re[U^{(n)}e^{ikj\Delta x}] \quad (k = p\frac{\pi}{J\Delta x};\ \ p = -J, -J+1, \ldots, -1, 0, 1, \ldots, J-1, J)$$

を (7.86) に代入すれば

$$U^{(n+1)} = [1 + \mu^2(\cos k\Delta x - 1) - i\mu \sin k\Delta x]U^{(n)}$$

となる (ただし $\mu \equiv c\dfrac{\Delta t}{\Delta x}$). よって, 増幅係数 λ は

$$\lambda = 1 + \mu^2(\cos k\Delta x - 1) - i\mu \sin k\Delta x$$

である. ここで,

$$\cos k\Delta x - 1 = -2\sin^2 \frac{k\Delta x}{2}$$

$$\sin k\Delta x = 2\sin \frac{k\Delta x}{2}\cos \frac{k\Delta x}{2}$$

であるから, 結局

$$|\lambda| = \left[1 - 4\mu^2(1-\mu^2)\sin^4 \frac{k\Delta x}{2}\right]^{\frac{1}{2}}$$

が得られる. $4\mu^2 \sin^4 \dfrac{k\Delta x}{2} \geq 0$ であるから, $|\lambda| \leq 1$ であるためには, $1 - \mu^2 \geq 0$ が必要十分である. よって, 安定の条件は

$$|c|\frac{\Delta t}{\Delta x} \leq 1$$

これは CFL condition である.

次に, 波長別の λ の振舞いを解析する. 最短波長 $2\Delta x$ においては, $k\Delta x = \pi$ であるから,

$$|\lambda| = (1 - 4\mu^2 + 4\mu^4)^{\frac{1}{2}} = |1 - 2\mu^2|$$

である. その 2 倍の波長 $4\Delta x$ においては, $k\Delta x = \dfrac{\pi}{2}$ であるから,

$$|\lambda| = (1 - \mu^2 + \mu^4)^{\frac{1}{2}}$$

である. また

$$\frac{d|\lambda|}{d\mu} = -\frac{4\mu(1-2\mu^2)\sin^4 \frac{k\Delta x}{2}}{\left[1 - 4\mu^2(1-\mu^2)\sin^4 \frac{k\Delta x}{2}\right]^{\frac{1}{2}}}$$

である. これより, λ は任意の $k\Delta x$ に対して$\mu = \dfrac{1}{\sqrt{2}}$ で最小値

$$\left(1 - \sin^4 \frac{k\Delta x}{2}\right)^{\frac{1}{2}}$$

(7.84) の両辺に u_j^* を乗じ, j について和をとれば,

$$\sum_{j=-J}^{J-1} u_j^*(u_j^{n+1} - u_j^n) = -\frac{1}{2}\mu \sum_{j=-J}^{J-1} u_j^*(u_{j+1}^* - u_{j-1}^*)$$

となる. ここで, 周期的境界条件により, 右辺は 0 となるので, 結局

$$\sum_{j=-J}^{J-1} u_j^*(u_j^{n+1} - u_j^n) = 0$$

となる. さて,

$$\sum_{j=-J}^{J-1} \frac{1}{2}\left[(u_j^{n+1})^2 - (u_j^n)^2\right] = \sum_{j=-J}^{J-1} \frac{1}{2}(u_j^{n+1} + u_j^n)(u_j^{n+1} - u_j^n)$$

であるが, ここへ (7.85) を代入し, $\beta = 1 - \alpha$ という関係によって β を消去すれば,

$$\sum_{j=-J}^{J-1} \frac{1}{2}\left[(u_j^{n+1})^2 - (u_j^n)^2\right] = \left(\alpha - \frac{1}{2}\right) \sum_{j=-J}^{J-1} (u_j^{n+1} - u_j^n)^2$$

となる. よって

$\alpha > \frac{1}{2}$:　スキームは不安定. 全 “エネルギー” は時間と共に単調に増加.

$\alpha = \frac{1}{2}$:　スキームは安定, かつ中立.

$\alpha < \frac{1}{2}$:　スキームは安定, かつ減衰. 全 “エネルギー” は時間と共に単調に減少.

となる.

7.12.4　two-step Lax-Wendroff スキーム

Lax-Wendroff スキームでは時間微分と空間微分の両方を差分化する.
まず, $u_{j+\frac{1}{2}}^{n+\frac{1}{2}}$, $u_{j-\frac{1}{2}}^{n+\frac{1}{2}}$ を

$$\begin{aligned}
\frac{u_{j+\frac{1}{2}}^{n+\frac{1}{2}} - \frac{1}{2}(u_{j+1}^n + u_j^n)}{\frac{1}{2}\Delta t} &= -c\frac{u_{j+1}^n - u_j^n}{\Delta x}, \\
\frac{u_{j-\frac{1}{2}}^{n+\frac{1}{2}} - \frac{1}{2}(u_j^n + u_{j-1}^n)}{\frac{1}{2}\Delta t} &= -c\frac{u_j^n - u_{j-1}^n}{\Delta x}
\end{aligned}$$

として定義する. この $u_{j+\frac{1}{2}}^{n+\frac{1}{2}}$ と $u_{j-\frac{1}{2}}^{n+\frac{1}{2}}$ を用いて

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = -c\frac{u_{j+\frac{1}{2}}^{n+\frac{1}{2}} - u_{j-\frac{1}{2}}^{n+\frac{1}{2}}}{\Delta x}$$

とする. $u_{j+\frac{1}{2}}^{n+\frac{1}{2}}$, $u_{j-\frac{1}{2}}^{n+\frac{1}{2}}$ を消去すれば, このスキームは

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = -c\frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} + \frac{1}{2}c^2\Delta t\frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2} \tag{7.86}$$

7.12.3　Matsuno スキーム, および一般の two level スキーム

(7.77) の時間微分項に対して Matsuno スキームを適用した場合を考える. Matsuno スキームでは, まず $u_j^{(n+1)^*}$ を forward スキーム を使って

$$\frac{u_j^{(n+1)^*} - u_j^n}{\Delta t} = -c\frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x}$$

と定義し, 次に backward スキームを使って

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = -c\frac{u_{j+1}^{(n+1)^*} - u_{j-1}^{(n+1)^*}}{2\Delta x} \tag{7.82}$$

とする. (7.82) から $u^{(n+1)^*}$ を消去すると,

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = -c\frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} + c^2\Delta t\frac{u_{j+2}^n - 2u_j^n + u_{j-2}^n}{(2\Delta x)^2} \tag{7.83}$$

となる. 右辺第 2 項を除けば, これは forward スキームによって得られた有限差分方程式と一致する. Δt を固定して, $\Delta x \to 0$ とすれば, 右辺第 2 項は $c^2\Delta t\dfrac{\partial^2 u}{\partial x^2}$ となる. この項は解を減衰させる働きがある. 減衰率は波長によって異なる. 波長 $\lambda = 4\Delta x$ の波が最も強く減衰し, 一方, 波長の一番短い波 ($\lambda = 2\Delta x$) に対する減衰率は 0 である[17]. 不安定を最も起こしやすいのは波長の一番短い波であるから, 不安定を抑える, という観点からすると, Matsuno スキームを移流方程式に適用するのは不適当である.

次に, (7.77) の時間差分として一般の two level スキームを適用した時の安定性を, エネルギー法を用いて調べる.

(7.77) を解くスキームは, 大抵のものが

$$u_j^{n+1} - u_j^n = -\frac{1}{2}\mu(u_{j+1}^* - u_{j-1}^*) \tag{7.84}$$

という形で書ける. 但し,

$$\mu \equiv c\frac{\Delta t}{\Delta x}$$

とし, u_j^* は $\{u_j^n\}$ を適当に線形結合したものとする.

例えば, 反復しない two level スキームならば, u_j^* は

$$u_j^* = \alpha u_j^n + \beta u_j^{n+1} \tag{7.85}$$

となり, 反復する two level スキームならば, u_j^* は

$$u_j^* = u_j^n - \frac{\beta}{2}\mu(u_{j+1}^n - u_{j-1}^n)$$

となり, Adams-Bashforth スキームならば, u_j^* は

$$u_j^* = \frac{3}{2}u_j^n - \frac{1}{2}u_j^{n-1}$$

となる.

ここでは, 反復しない two level スキームの安定性を, 周期的境界条件 $u_{-J} = u_J$ の元で調べてみる.

[17] 解として $u_j^n = Re(U^{(n)}\exp ikj\Delta x)$ を右辺第 2 項に代入すると, この項は $c^2U^{(n)}\Delta t\cos kj\Delta x\sin^2 k\Delta x/(\Delta x)^2$ となる. すなわちこの項は, $k\Delta x = \pi/2$ ($\lambda = 4\Delta x$) の時最大となり, $k\Delta x = \pi$ ($\lambda = 2\Delta x$) の時 0 となる.

となる. ここで,

$$p \equiv \omega\Delta t = -c\frac{\Delta t}{\Delta x}\sin k\Delta x \tag{7.81}$$

と書けば

$$U^{(n+1)} = U^{(n-1)} + 2ipU^{(n)}$$

となる. これは, (7.77) に leapfrog スキームを適用して (7.78) となった式に (7.79) を代入したものと一致する. よって, (7.78) の特性は leapfrog スキームを振動方程式に適用した場合の結果から導くことができる.

その結果から、安定であるためには $|p| \leq 1$ であることが必要かつ十分であることが分かっているから, 今の場合

$$\left| c\frac{\Delta t}{\Delta x}\sin k\Delta x \right| \leq 1$$

が $0 \leq |k| \leq \dfrac{\pi}{\Delta x}$ の任意の k に対して成立すること, すなわち,

$$|c|\frac{\Delta t}{\Delta x} \leq 1$$

が, 安定であるための必要十分条件である. これを

CFL condition, または Courant-Friedrichs-Lewy condition

という. 最も不安定となりやすい波は $|p|$ が最大となる波長を与える波, すなわち $k\Delta x = \dfrac{\pi}{2}$ なる波である. これは, 波長が $4\Delta x$, すなわち, 表現し得る最短波長の波の 2 倍のものである.

振動方程式に対する解析の結果から, physical mode と computational mode の振舞いを考えると, table 7.12 のようになる.

Table 7.12: leapfrog スキームを適用したときの各 mode の振舞い

	physical mode	computational mode
増幅係数	λ_1 $U_1^{(n)} = \lambda_1^n U_1^{(0)}$	λ_2 $U_2^{(n)} = \lambda_2^n U_2^{(0)}$
安定な場合, p≷0 に対して	$\lambda_1 = e^{i\theta}$ 但し $\theta = \arctan\left(\frac{p}{\sqrt{1-p^2}}\right)$	$\lambda_2 = e^{i(\pm\pi-\theta)} = -e^{-i\theta}$ 但し $\theta = \arctan\left(\frac{p}{\sqrt{1-p^2}}\right)$
(7.79) を用いた時の解	$u_j^n = Re\left[U_1^{(0)} e^{ik(j\Delta x + \frac{\theta}{k\Delta t}n\Delta t)}\right]$	$u_j^n = Re\left[(-1)^n U_2^{(0)} e^{ik(j\Delta x - \frac{\theta}{k\Delta t}n\Delta t)}\right]$ 時間 1 ステップ毎に符号が変わる
位相速度 その極限 $(\Delta x, \Delta t \to 0)$[16]	$c_1 = -\frac{\theta}{k\Delta t}$ $c_1 \to c$	$c_1 = \frac{\theta}{k\Delta t}$ $c_2 \to -c$

[16] $\Delta t \to 0$ の時$\theta \to p$, (7.81) より $\Delta x \to 0$ の時 $p \to -ck\Delta t$ となるため.

7.12　2 次の正確度を持つ空間中心差分スキーム

7.12.1　Introduction

空間差分に伴う様々な問題, 特に位相の誤差と計算上の発散 · 減衰という現象を避けるための 1 つの方法として, 近似の正確度を上げるという方法がある.

ここでは, 1 次元移流方程式

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0; \qquad c = \text{const.}, \qquad u = u(x,t) \tag{7.76}$$

に, 2 次の正確度を持つスキームを適用したときの特性を調べる.

この空間微分を, 隣接する 2 つの格子点における値を使って有限差分商で近似すると, 時間微分は

$$\frac{\partial u_j}{\partial t} = -c\frac{u_{j+1} - u_{j-1}}{2\Delta x} \tag{7.77}$$

と表すことができる. ここで添字は, 領域の中央に原点を取り, $j = -J, -J+1, \ldots, -1, 0, 1, \ldots, J-1, J$ とする. 格子点数は $(2J+1)$ である.

取り上げるスキームは

- (7.77) の時間微分に leapfrog スキームを適用したもの
- (7.77) の時間微分に Matsuno スキーム, および一般の two level スキームを適用したもの
- (7.76) に, Lax-Wendroff スキームを適用したもの

である.

7.12.2　leapfrog スキーム

(7.77) の時間微分を計算するのに, leapfrog スキームを使えば,

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t} = -c\frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} \tag{7.78}$$

となる.

このスキームの特性は, 以下のようにして, 振動方程式に対する leapfrog スキームの特性 (第 7.5.4 節参照.) に帰着することができる.

(7.77) に

$$u_j(t) = Re[U(t)e^{ikj\Delta x}] \qquad (k = p\frac{\pi}{J\Delta x};\ \ p = -J, -J+1, \ldots, -1, 0, 1, \ldots, J-1, J) \tag{7.79}$$

という形の解を代入すると,

$$\frac{dU}{dt} = i\left(-\frac{c}{\Delta x}\sin k\Delta x\right)U \tag{7.80}$$

となる. これは, 振動方程式で

$$\omega \equiv -\frac{c}{\Delta x}\sin k\Delta x$$

とした場合に相当する.

(7.80) を適当なスキームを使って近似したものは, (7.77) に同じスキームを適用して, それから解 (7.79) を代入したものに等しい. 例えば, (7.80) の近似として leapfrog スキームを使えば

$$U^{(n+1)} = U^{(n-1)} + 2i\left(-c\frac{\Delta t}{\Delta x}\sin k\Delta x\right)U^{(n)}$$

と書ける. ただし, a_p は任意の定数である. $\tau = 0$ を代入すれば,

$$u_j(0) = a_j$$

となる. a_j は, 任意の初期条件を満たすように取れるから, これが一般解である.

全ての初期値は, デルタ関数 (この場合, もどき) の和で表せること, 1 つ 1 つの解 (7.74) の振舞いを考えると, j における値 $u_j(\tau)$ は, その場所, 及びそれより上流の場所における初期値の影響の重ね合わせと解釈出来る.

得る最も短い波長, $L = 2\Delta x$ に対しては, Δt が非常に小さくない限り, 誤差は大変大きくなり得る. $\mu = \frac{1}{2}$ の時は, 1 ステップ時間を進めただけでこの波が完全に消えてしまう.

なぜ, このような誤差の波長依存性が現われるのだろうか. それは, 様々な波長の成分がどのようにして有限差分格子で表わされているかを考えればわかる. 表現し得る最も短い波では, 1 波長の中に格子点は 2 点しかない. 差分格子ではこのような波は十分表現されていないのであるから, 誤差は大きくなる. 波長が長くなればなるほど 1 波長の中の格子点の数は増加し波が十分表現されるようになるので, 誤差は減少する.

上流差分スキームは, 次のようにしても得られる. 移流方程式の特性曲線は $x - ct = (\text{const.})$ である. 点 $((j-1)\Delta x, n\Delta t)$ を通る特性曲線上での u の値はステップ n では u^n_{j-1} であるからステップ $(n+1)$ でもその値である。従って, その特性曲線と直線 $t = (n+1)\Delta t$ の交わる点 A $((j-1)\Delta x + c\Delta t, (n+1)\Delta t)$ における u の値は u^n_{j-1} である. 同様に, $(j\Delta x, n\Delta t)$ を通る特性曲線上を考えると, 点 B $(j\Delta x + c\Delta t, (n+1)\Delta t)$ における u の値は u^n_j である. 安定の条件より $\dfrac{\Delta t}{\Delta x} \le \dfrac{1}{c}$ なので, 値を求めたい格子点 $(j\Delta x, (n+1)\Delta t)$ は線分 AB (長さは Δx である.) 上にあり, 点 A からの距離は $\Delta x - c\Delta t$ である. そこで, 点 A, B の値を用いて u^{n+1}_j を 1 次の内挿によって求めると[14],

$$u^{n+1}_j = u^n_{j-1} + \frac{u^n_j - u^n_{j-1}}{\Delta x}(\Delta x - c\Delta t)$$

が得られる. これは, スキーム (7.72) と等しい[15].

差分方程式 (7.70) の性質を解析するため, その解析解を考えることにする (Wurtele, 1961 参照.). 無次元の時間 $\tau = \dfrac{ct}{\Delta x}$ を考える. このとき, 式 (7.70) は

$$\frac{d}{d\tau}u_j(\tau) + u_j(\tau) - u_{j-1}(\tau) = 0 \tag{7.73}$$

となる. Poisson の周波数関数

$$u_j(\tau) = \begin{cases} \dfrac{e^{-\tau}\tau^{j-p}}{(j-p)!} & \text{for } j \ge p \\ 0 & \text{for } j < p \end{cases} \tag{7.74}$$

は (7.73) の解である.

この関数は,

$$\sum_{j-p=0}^{\infty} \frac{e^{-\tau}\tau^{j-p}}{(j-p)!} = 1 \tag{7.75}$$

という性質を持つ. $\tau = 0$ の時は, 高さ $\dfrac{1}{\Delta x}$ の長方形となる. すなわち, これは, 時刻 $\tau = 0$ において, デルタ関数的な擾乱を与えた事に相当する. これが, 時間と共にどの様に変化するかを追ってみる.

この平均値は

$$\sum_{j-p=0}^{\infty} (j-p)\frac{e^{-\tau}\tau^{j-p}}{(j-p)!} = \tau$$

であるから, この擾乱の平均値は, 実際の波と同じ速度で移動する. だが, 最大値は実際の波よりも遅れる. また, (7.75) より, 擾乱の量は常に 1 で変化しない. だが, 最大値は速やかに減少する.

Poisson の周波数関数の性質が分かったところで, 元の微分差分方程式の一般解を求める. 前節と同様, p は任意で, 方程式は線形だから, 解は

$$u_j(\tau) = \sum_{p=-\infty}^{j} a_p \frac{e^{-\tau}\tau^{j-p}}{(j-p)!}$$

[14] 内挿ならば影響領域に格子点が入るが, 外挿ではそうはならない.
[15] 2 次で内挿すれば, Lax–Wendroff scheme が得られる.

7.11　上流差分スキーム

7.11.1　上流差分と下流差分

移流方程式の空間微分は, もちろん, 非中心差分で置き換えてもいい. 物理的に考えると, まず 1 点は求めたい点にとり, もう 1 点はその上流に取るのが自然である. そこで, 1 次元移流方程式

$$\frac{\partial}{\partial u}t + c\frac{\partial}{\partial u}x = 0$$

は

$$\frac{\partial}{\partial u_j}t + c\frac{u_j - u_{j-1}}{\Delta x} = 0, \qquad \text{for } c > 0 \tag{7.70}$$

$$\frac{\partial}{\partial u_j}t + c\frac{u_{j+1} - u_j}{\Delta x} = 0, \qquad \text{for } c < 0 \tag{7.71}$$

と近似できる. これもまた, 差分微分方程式である. ただし, (7.70) は後退差分, (7.71) は前進差分である. 2 つとも, 差分を上流に向けて計算している. そこで, この 2 つの差分は,

上流差分 (upstream differences)

と呼ばれる. 逆のものを

下流差分 (downstream differences)

という.

7.11.2　上流差分スキームの特徴

式 (7.70), (7.71) の時間微分を適合性のある方法で近似すれば, 移流方程式を解くスキームができる. できたスキームは空間に関して 1 次の accuracy しかもたないが, 物理的に波が進む方向にしか波が進まないので, デルタ関数的な擾乱を与えても寄生波が発生しないという利点がある.

例えば, 時間微分を前進差分で置き換えれば, $c > 0$ に対して,

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + c\frac{u_j^n - u_{j-1}^n}{\Delta x} = 0 \tag{7.72}$$

というスキームができる. このスキームではどのような波が減衰するのか調べてみよう.

式 (7.65) のように 解を x 方向に Fourier 展開し, von Neumann 法に従って増幅係数 λ を計算すると,

$$\begin{aligned} |\lambda|^2 &= 2(1-\cos k\Delta x)\left(\mu - \frac{1}{2}\right)^2 - \frac{1}{2}(1-\cos k\Delta x) + 1 \\ &= 1 - 2\mu(1-\mu)(1-\cos k\Delta x) \end{aligned}$$

となる. ただし, $\mu \equiv \dfrac{c\Delta t}{\Delta x}$ である. この式に従って, $k\Delta x$ の値を様々に固定しておいて, μ の変化に対する $|\lambda|$ の変化を調べてみる. まず $k \neq 0$ のときを考えると, $|\lambda|^2$ を縦軸に, μ を横軸にとった時の曲線は, 下に凸の放物線で, $\mu = \dfrac{1}{2}$ で最小値をとる. 次に, $k \to 0$ (波長大) のときを考えると, 増幅係数は μ の全ての値に対して 1 に近付くことが分かる.

また $(0 <)\mu < 1$ なる全ての μ に対して $|\lambda|^2 < 1$ (減衰) である. そして, 減衰率 $|\lambda|$ は, 波長が短くなるにつれて増加する. $(0 \leq k \leq \dfrac{\pi}{\Delta x}$ に注意.) 真の解ならば増幅係数 $\lambda = \text{const.} = 1$ だから, この減衰は, 有限差分による誤差を表わす. こうして, この誤差は波長が短くなるにつれて増加することが分かった. 求め

7.10 スキームのまとめ

ここでは移流方程式を解くために用いられるスキームについて空間差分の取り方と時間差分の取り方に注目してまとめる.

空間差分には中心差分及び上流差分 ・下流差分の 3 種類がある. それぞれのスキームに関する特徴をまとめると次のようになる.

- 上流差分
  上流差分スキームにはデルタ関数的な擾乱を与えても寄生波が発生しないというメリットがある. また上流差分スキームには, その依存領域が特性曲線を含んでおり, しかも上流差分スキームに比べて余分な領域が狭いという特徴もある.

- 下流差分
  下流差分スキームの依存領域は特性曲線を含んでいない. よって, 依存領域に関しては下流差分は推奨できない.

- 中心差分上流差分スキームの依存領域は特性曲線を含んではいるが, 不用な領域も含んでいる.

空間差分については, 問題に応じて必要な正確度を持つ中心差分スキームや上流差分を用いればよい.

また, 時間差分スキームについては常微分方程式を解く際に用いるスキームを使用することができて, 次のような種類がある.

- two level scheme
  two level scheme には Euler scheme や backward scheme, Matsuno scheme, Heun scheme などの種類がある. two level scheme を用いる場合, その安定性については特に注意を払う必要がある.

- three level scheme
  three level scheme にはleapfrog scheme や Adams-Bashforth scheme などの種類がある. three level scheme には最初のステップで使えないという欠点がある.

移流方程式を解くためには, 上で述べた空間差分スキームと時間差分スキームを適当に組み合せて方程式を差分化すればよい.

デルタ関数的擾乱

式 (7.64) の初期条件として, 原点にデルタ関数的なもの

$$u_j(0) = \begin{cases} 1 & \text{for } j = 0 \\ 0 & \text{for } j \neq 0 \end{cases} \tag{7.67}$$

を与えた場合の解の振る舞いを調べる.

まず, 式 (7.64) の解析解を求める. 無次元の時間 τ を,

$$\tau \equiv \frac{ct}{\Delta x}$$

と定義する.t を τ で書き換えると, 式 (7.64) は

$$2\frac{d}{d\tau}u_j(\tau) = u_{j-1}(\tau) - u_{j+1}(\tau) \tag{7.68}$$

となる. これは, ベッセル関数の漸化式である. よって, ベッセル関数 $J_l(\tau)$ の任意の重ね合わせ

$$u_j(\tau) = \sum_{p=-\infty}^{\infty} a_p J_{j-p}(\tau) \tag{7.69}$$

は (7.68) を満たす. ただし, a_p は任意の定数である. ここで, $J_j(0) \neq 0$ であるのは $j = 0$ のときのみであることを考えれば,

$$u_j(0) = a_j$$

となる.a_p は任意だから, (7.69) は任意の初期条件を満たす. よって, (7.69) は (7.64), (7.68) の一般解である.

初期条件 (7.67) を与えると, 解は

$$u_j(\tau) = J_j(\tau)$$

である.

長い波長の位相速度は, 真の解と比べても余り遅くはないから, 主たる擾乱は $j = \tau$ から少しだけ遅れたところにある. 短波長の波の群速度は負であるために, $j < 0$ においても, $u_j \neq 0$ となっている. すなわち, 実際の波とは逆に進むという論外の波が出来てしまう. これを寄生波 (parasitic wave) という.

このように, この数値解はひどいものになってしまい, 差分近似がこの様なデルタ関数的擾乱は苦手であることがわかる.

中心差分近似を時間微分に適用した, より一般的な解析解については, Egger, 1971 参照.

7.9.4　数値分散

位相速度と群速度の分散性

まず, 移流方程式 (7.58) の真の解の位相速度を求める.

移流方程式 (7.58) へ, 調和解

$$u(x,\ t) = Re[U(t)e^{ikx}]$$

を代入すれば

$$\frac{dU}{dt} + ikcU = 0$$

となる. よって真の解の位相速度は

$$c = \frac{\nu}{k} \qquad (\nu\text{は周波数})$$

で一定となり, 分散性はない.

次に, 数値解の位相速度を求める.

移流方程式 (7.58) を, 空間微分のみを差分で表すと

$$\frac{\partial u}{\partial t} + c\frac{u_{j+1} - u_{j-1}}{2\Delta x} = 0 \tag{7.64}$$

となる. この方程式の誤差は空間差分のみによる. 上と同様に調和解

$$u_j(t) = Re[U(t)e^{ikj\Delta x}] \quad (k = p\frac{\pi}{J\Delta x};\ \ p = -J, -J+1, \ldots, -1, 0, 1, \ldots, J-1, J) \tag{7.65}$$

を代入すれば,

$$\frac{dU}{dt} + ik\left(c\frac{\sin k\Delta x}{k\Delta x}\right)U = 0$$

となる. よって, 数値解の位相速度 c^* は

$$c^* = c\frac{\sin k\Delta x}{k\Delta x} \tag{7.66}$$

となり, これは k による. すなわち, 分散性がある.

このように, 真の解の位相速度には分散性がないのに, 差分化したために数値解には分散性が生じてしまった. これを, 数値分散という.

(7.66) より

$$\frac{c^*}{c} = \frac{\sin k\Delta x}{k\Delta x} <\ 1$$

である. よって, 数値解の位相速度は, 常に真の解の位相速度よりも遅い. また, 式 (7.65) より $0 \le k\Delta x \le \pi$ だから, 波長が短いほど, 真の位相速度より遅くなることも分かる. 特に, 最小の波長 $2\Delta x$ の波は, $k\Delta x = \pi$ であるから $c^* = 0$, すなわち定在波となる. [13]

このため, 様々な波長の波が重なって出来ている波は, 計算して行くと崩れてくる. 従って, 様々な波長の波が重なって出来ているものの数値計算においては, 数値分散は重大な問題である.

次に, 群速度について考える. 真の解の群速度 c_g は

$$\tau_c = \frac{d(kc)}{dk} =\ c$$

である. 一方, 数値解の群速度は, (7.66) より

$$c_g^* = \frac{d(kc^*)}{dk} = c\ \cos k\Delta x$$

となる. よって, $k\Delta x$ が増加すると, c_g^* は単調に減少し, 最短波長 $2\Delta x$ のときは, $c_g^* = -c_g$ となる.

[13] 式 (7.65) において, 最も短い波長 $(k = \pi/\Delta x)$ の波は, $u_j(t) = Re[U(t)e^{ij\pi}]$ と表される. この波は $j = 0, 1, 2, 3\ldots$ に対して $u_j = U, -U, U, -U\ \ldots$ と格子点ごとに符号を変えるので, 式 (7.64) より $\frac{\partial u}{\partial t} = 0$ となり, 定在波となることが良く分かる.

という. これを

$$\varepsilon = O(\Delta x),\ \varepsilon = O(\Delta t)$$

または

$$\varepsilon = O(\Delta x,\ \Delta t)$$

と書く.

スキームによっては, Δx, Δt それぞれに対しての正確度の次数が違うことがあるので, この 2 つをわけて考えることは有用である.

2 番目の例として, 時間差分は leapfrog, 空間差分は中心差分とすると

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t} + c\frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} = 0 \tag{7.62}$$

となる.

このスキームにも適合性がある. すなわち, 差分間隔が 0 に近付く時, 元の微分方程式 (7.58) に近付く.

打ち切り誤差を Taylor 展開すると,

$$\varepsilon = \frac{1}{6}\frac{\partial^3 u}{\partial t^3}(\Delta t)^2 + \cdots + c\left(\frac{1}{6}\frac{\partial^3 u}{\partial x^3}(\Delta x)^2 + \cdots\right) \tag{7.63}$$

となる. 最低次の項は, Δx について 2 次, Δt について 2 次である. 従って

このスキームは Δx について 2 次, Δt について 2 次の正確度 (accuracy) を持つ!

つまり,

$$\varepsilon = O(\Delta x^2),\ \varepsilon = O(\Delta t^2)$$

または

$$\varepsilon = O(\Delta x^2,\ \Delta t^2)$$

ということになる.

このように, 同じ方程式でも, 適用するスキ− ムによって正確度が違ってくる.

7.9.3　CFL condition

一般に, スキームが安定であるための必要十分条件を

CFL condition (Courant-Friedfichs-Leny condion)

という. 次のような形をしていることが多い.

$$c\Delta t \leq \Delta x$$

7.9　1 次元移流方程式の数値計算法

7.9.1　Introduction

ここでは, 一次元線形移流方程式

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0 \tag{7.58}$$

を数値計算する際に生ずる問題点に付いて述べる.

この方程式の解析解は

$$u = f(x - ct)$$

である. ただし, f は任意の関数であり, 初期条件によって決まる.

以下の節では次の問題について論じる.

- 正確度
- 収束性と安定性
- CFL condition
- 数値分散　(computational dispersion または numerical dispersion)

7.9.2　正確度

方程式 (7.58) の数値解を求めるため, x-t 平面に格子を設定し, 点 $(j\Delta x, n\Delta t)$ における u の値 u_j^n を求めることを考える.

最初の例として, 時間微分は前進差分, 空間微分は後退差分で置き換えてみると,

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + c\frac{u_j^n - u_{j-1}^n}{\Delta x} = 0 \tag{7.59}$$

となる.

このスキームは適合性がある. すなわち, 差分間隔が 0 に近付く時, 元の微分方程式 (7.58) に近付く.

スキームの誤差は, 次のようにして定義される. 真の解 $u(j\Delta x, n\Delta t)$ を (7.59) に代入した時, 本来ならば右辺は 0 となるはずだが, 誤差のため 0 とはならない. これを ε と書き, 打ち切り誤差 (truncation error) という.

すなわち,

$$\frac{u(j\Delta x, (n+1)\Delta t) - u(j\Delta x, n\Delta t)}{\Delta t} + c\frac{u(j\Delta x, n\Delta t) - u((j-1)\Delta x, n\Delta t)}{\Delta x} = \varepsilon \tag{7.60}$$

である.

左辺を Taylor 展開すると, この場合

$$\varepsilon = \frac{1}{2}\frac{\partial^2 u}{\partial t^2}\Delta t + \frac{1}{6}\frac{\partial^3 u}{\partial t^3}(\Delta t)^2 + \cdots - c\left(\frac{1}{2}\frac{\partial^2 u}{\partial x^2}\Delta x - \frac{1}{6}\frac{\partial^3 u}{\partial x^3}(\Delta x)^2 + \cdots\right) \tag{7.61}$$

と書ける. 最低次の項は, Δx について 1 次, Δt についても 1 次である. そこで

このスキームは Δx について 1 次, Δt について 1 次の正確度 (accuracy) を持つ!

7.8 Split 法

微分方程式の表す現象を遅い現象と速い現象に分けることができるならば, 各現象ごとに異なる時間ステップをとり計算量を節約することが可能な場合がある.

例として, two level scheme を考える.two level scheme は, 一般に

$$\mathbf{F}^{(n+1)} = (\mathbf{E} + \mathbf{G}\Delta t)\mathbf{F}^{(n)}$$

という形に書ける. ここで, $\mathbf{E}$ は単位行列を表し, $\mathbf{G}$ はある作用行列を表す.

さて, 作用行列 $\mathbf{G}$ が

$$\mathbf{G} = \mathbf{G}_1 + \mathbf{G}_2 + \mathbf{G}_3$$

という形に書けるとする. このとき, 時間差分 Δt が十分小さければ

$$\begin{aligned}
\mathbf{F}^{(n+1)} &= (\mathbf{E} + (\mathbf{G}_1 + \mathbf{G}_2 + \mathbf{G}_3)\Delta t)\mathbf{F}^{(n)} \\
&\cong (\mathbf{E} + \mathbf{G}_1\Delta t)(\mathbf{E} + \mathbf{G}_2\Delta t)(\mathbf{E} + \mathbf{G}_3\Delta t)\mathbf{F}^{(n)} \\
&\cong (\mathbf{E} + \mathbf{G}_1\Delta t)(\mathbf{E} + \mathbf{G}_2\frac{\Delta t}{n})^n(\mathbf{E} + \mathbf{G}_3\frac{\Delta t}{mn})^{mn}\mathbf{F}^{(n)} \\
&= (\mathbf{E} + \mathbf{G}_1\Delta t_1)(\mathbf{E} + \mathbf{G}_2\Delta t_2)^n(\mathbf{E} + \mathbf{G}_3\Delta t_3)^{mn}\mathbf{F}^{(n)}
\end{aligned}$$

と変形することができる. ただし

$$\Delta t_1 \equiv \Delta t, \qquad \Delta t_2 \equiv \frac{\Delta t}{n}, \qquad \Delta t_3 \equiv \frac{\Delta t}{mn}$$

この時, $\mathbf{G}_3$ に速い現象, $\mathbf{G}_2$ に少し遅い現象, $\mathbf{G}_1$ に遅い現象を選べば, 速い現象は時間間隔を小さく取り, 遅い現象は時間差分を大きく取って計算することができることもある. つまり, Semi-Implicite 法 [1] を用いて, 少し遅い現象は速い現象を m 回計算する間に 1 回計算すればよく, 遅い現象は速い現象を mn 回計算する間に 1 回計算すればよい. これを

split 法！

という.

Table 7.11: three level schemes

	leapfrog sheme	Adams-Bashforth scheme
式	$U^{(n+1)} = U^{(n-1)} - 2KU^{(n)}$ $\lambda^2 + 2K\lambda - 1 = 0$	$U^{(n+1)} = U^{(n)} - K(\frac{3}{2}U^{(n)} - \frac{1}{2}U^{(n-1)})$ $\lambda^2 - (1 - \frac{3}{2}K)\lambda - \frac{1}{2}K = 0$
physical mode $K \to 0$	$\lambda_1 = -K + \sqrt{1 + K^2}$ $\lambda_1 \to 1$	$\lambda_1 = \frac{1}{2}\left(1 - \frac{3}{2}K + \sqrt{1 - K + \frac{9}{4}K^2}\right)$ $\lambda_1 \to 1$
computational mode $K \to 0$	$\lambda_2 = -K - \sqrt{1 + K^2}$ $\lambda_2 \to -1$	$\lambda_2 = \frac{1}{2}\left(1 - \frac{3}{2}K - \sqrt{1 - K + \frac{9}{4}K^2}\right)$ $\lambda_2 \to 0$
安定性	$K > 0$ で $\lambda_2 < -1$ computational mode は常に不安定. 摩擦方程式には適さない.	Kが十分小さければ安定 (computational mode は落ちる)

7.7　方程式が複数の項を含む場合のスキームの選択

方程式が

$$\frac{dU}{dt} = i\omega U - \kappa U \tag{7.56}$$

のように振動項と, 摩擦項の両方を含んでいるときは, 例えば leapfrog scheme は摩擦項があるから使えない. こういったときには, 振動項と摩擦項に対して違うスキームを使えばよい. 例えば, 振動項には leapfrog scheme, 摩擦項には Euler scheme を用いて,

$$\frac{U^{(n+1)} - U^{(n-1)}}{2\Delta t} = i\omega U^{(n)} - \kappa U^{(n-1)} \tag{7.57}$$

のように差分化するのは一つの有効な方法である. もちろん, 他の組合せも可能である.

Table 7.9: 反復しない two level schemes

種類	$(\alpha,\ \beta)$	性質
Euler scheme	(1,0)	$0 < K \leq 2$ で安定 解は $K < 1$ で振動しない
backward scheme	(0,1)	$K > 0$ で安定 解は振動しない
台形 scheme	$\left(\frac{1}{2}, \frac{1}{2}\right)$	$K > 0$ で安定 解は $K < 2$ で振動しない

$$U^{(n+1)} = (1 - K + \beta K^2)U^{(n)} \tag{7.55}$$

となる. K の定義は前と同じ (式 (7.53)) である. $\beta = 1$ と $\beta = \dfrac{1}{2}$ のスキームについて, table 7.10 にまとめる.

Table 7.10: 反復する two level schemes

種類	β	性質
Matsuno scheme	1	$0 < K \leq 1$ で安定 解は振動しない
Heun scheme	$\frac{1}{2}$	$0 < K \leq 2$ で安定 解は振動しない

Matsuno, Heun 両スキームとも, K が十分小さければ安定である.

7.6.3　three level scheme の特性

次に, three level scheme について以下の表にまとめる.

7.6　摩擦方程式に適用したときのスキームの特性

7.6.1　Introduction

適用する方程式によって, スキームの特性は変わる. ここでは, 摩擦方程式に適用したときの, スキームの特性を考える. ただし, 摩擦方程式とは,

$$\frac{dU}{dt} = -\kappa U, \quad U = U(t), \quad \kappa > 0 \tag{7.50}$$

という方程式である. 応用としては, 速度に比例した摩擦力が働く場合 (大気の地面付近の境界層など), 熱伝導方程式をこれへ帰着させる場合[12] などがある.

この一般解は,

$$U(t) = U(0)e^{-\kappa t} \tag{7.51}$$

である. すなわち, この解は, 実数部も虚数部も時間と共に指数的に減少する.

以下に, 摩擦方程式に様々な差分スキームを適用したときの性質を, von Neumann 法で解析する.

7.6.2　two level scheme の特性

反復しないスキーム

反復しないスキーム

$$\begin{aligned} U^{(n+1)} &= U^{(n)} + \Delta t\ (\alpha f^{(n)} + \beta f^{(n+1)}) \\ f^{(n)} &\equiv f(U^{(n)}, n\Delta t) \\ \alpha &+ \beta = 1 \end{aligned}$$

を摩擦方程式に代入すれば,

$$U^{(n+1)} = U^{(n)} - \kappa\Delta t\ (\alpha U^{(n)} + \beta U^{(n+1)}) \tag{7.52}$$

となる. ここで,

$$K \equiv \kappa\Delta t \tag{7.53}$$

とかけば,

$$U^{(n+1)} = \frac{1-\alpha K}{1+\beta K} U^{(n)} \tag{7.54}$$

となる. 従って, $(\alpha,\beta) = (1,0)$, $(\alpha,\beta) = (0,1)$, $(\alpha,\beta) = (\frac{1}{2},\frac{1}{2})$ の各スキームの性質は table 7.9 のようになる.

反復するスキーム

反復するスキーム

$$\begin{aligned} U^{(n+1)*} &= U^{(n)} + \Delta t \cdot f^{(n)} \\ U^{(n+1)} &= U^{(n)} + \Delta t\ (\alpha f^{(n)} + \beta f^{(n+1)*}) \\ f^{(n+1)*} &\equiv f(U^{(n+1)*}, (n+1)\Delta t) \\ \alpha &+ \beta = 1 \end{aligned}$$

を考えると,

[12] $u \propto e^{ik\cdot x}$ とすると, 熱伝導方程式 $\frac{\partial u}{\partial t} = \chi\triangle u$ は $\frac{\partial u}{\partial t} = -\chi k^2 u$ に帰着する. Fourier 展開しない数値解法については 7.20.1 節を参照して欲しい.

Adams-Bashforth scheme

Adams-Bashforth scheme を振動方程式に適用すると,

$$U^{(n+1)} = U^{(n)} + i\omega\Delta t\left(\frac{3}{2}U^{(n)} - \frac{1}{2}U^{(n-1)}\right) \tag{7.48}$$

となる.

式 (7.39) を代入して

$$\lambda^2 - \left(1 + i\frac{3}{2}p\right)\lambda + i\frac{1}{2}p = 0$$

となる. これより増幅係数は

$$\begin{aligned}\lambda_1 &= \tfrac{1}{2}\left(1 + i\tfrac{3}{2}p + \sqrt{1 - \tfrac{9}{4}p^2 + ip}\right)\\ \lambda_2 &= \tfrac{1}{2}\left(1 + i\tfrac{3}{2}p - \sqrt{1 - \tfrac{9}{4}p^2 + ip}\right)\end{aligned} \tag{7.49}$$

と求めることができる. よって, モード毎に整理すると次の表のようになる.

Table 7.8: Adams-Bashforth scheme の mode

	physical mode	computational mode[10]
増幅係数[11]	$\lambda_1 = \frac{1}{2}\left(1 + i\frac{3}{2}p + \sqrt{1 - \frac{9}{4}p^2 + ip}\right)$ $= 1 + ip - \frac{1}{2}p^2 + i\frac{1}{4}p^3 - \frac{1}{8}p^4 + \cdots$ $= \left(1 - \frac{1}{2}p^2 - \frac{1}{8}p^4 - \cdots\right)$ $+i\left(p + \frac{1}{4}p^3 + \cdots\right)$	$\lambda_2 = \frac{1}{2}\left(1 + i\frac{3}{2}p - \sqrt{1 - \frac{9}{4}p^2 + ip}\right)$ $= i\frac{1}{2}p + \frac{1}{2}p^2 - i\frac{1}{4}p^3 + \frac{1}{8}p^4 + \cdots$ $= \left(\frac{1}{2}p^2 + \frac{1}{8}p^4 + \cdots\right)$ $+i\left(\frac{1}{2}p - \frac{1}{4}p^3 - \cdots\right)$
$p \to 0$	$\lambda_1 \to 1$	$\lambda_2 \to 0$
$\lvert\lambda\rvert$	$\lvert\lambda_1\rvert = \left(1 + \frac{1}{2}p^4 + \cdots\right)^{\frac{1}{2}}$ $= 1 + \frac{1}{4}p^4 + \cdots$ > 1	$\lvert\lambda_2\rvert = \left(\frac{1}{4}p^2 + \cdots\right)^{\frac{1}{2}}$ $= \frac{1}{2}p + \cdots$

これより, physical mode の増幅係数が 1 より大きいので, このスキームは不安定であることがわかる. しかし, Heun scheme と同様 (4 次の係数は Heun scheme の 2 倍であるが), その不安定性は非常に弱いので, 少し危険と感じたら Δt を小さくすれば, 実用には差し支えない. 実際, このスキームは気象学ではしばしば使われている. (例えば Deardorff の惑星境界層のシミュレーション (Deardorff, 1974))

この他の方法については, Lilly, 1965; Kurihara, 1965; Young, 1968 などの論文を見ること.

実際のスキームの選択では, ここに記した特性の他に, 計算機の容量なども考えねばならない.

[10] computational mode の増幅係数を十分小さくできるので, Adams-Bashforth scheme では, computational mode は問題とならない.

[11] p は小さくとる, ということを考えて級数展開する.

Table 7.7: leapfrog scheme の mode

		physical mode	computational mode
増幅係数		$\lambda_1 = \sqrt{1-p^2} + ip$	$\lambda_2 = -\sqrt{1-p^2} + ip$
$\Delta t \to 0$	$(p \to 0)$	$\lambda_1 \to 1$ 真の値に近づく	$\lambda_2 \to -1$ 真の値に近づかない
$\|p\| < 1$	性質	$\|\lambda_1\| = 1$ 安定かつ中立	$\|\lambda_2\| = 1$ 安定かつ中立
	位相	$\theta_1 = \arctan\left(\frac{p}{\sqrt{1-p^2}}\right)$ $= p + \frac{1}{6}p^3 + \cdots$	$\theta_2 = \arctan\left(-\frac{p}{\sqrt{1-p^2}}\right)$
	$p > 0$[6]	$0 < \theta_1 < \frac{\pi}{2}$	$\frac{\pi}{2} < \theta_2 < \pi$
	$p < 0$[7]	$-\frac{\pi}{2} < \theta_1 < 0$	$-\pi < \theta_2 < -\frac{\pi}{2}$
	$p \to 0$	$\theta_1 \to p$	$\theta_2 \to \pm\pi - p$
$\|p\| = 1$	性質	$\lambda_1 = ip$ $\|\lambda_1\| = 1$ 安定かつ中立	$\lambda_2 = ip$ $\|\lambda_2\| = 1$ 安定かつ中立
	位相[8] $p = \pm 1$	$\theta_1 = \pm\frac{\pi}{2}$	$\theta_2 = \pm\frac{\pi}{2}$
$\|p\| > 1$	性質	$\lambda_1 = i(p + \sqrt{p^2-1})$ $\|\lambda_1\| = \|p + \sqrt{p^2-1}\|$	$\lambda_2 = i(p - \sqrt{p^2-1})$ $\|\lambda_2\| = \|p - \sqrt{p^2-1}\|$
	$p > 1$	$\|\lambda_1\| > 1$ 不安定[9]	
	$p < -1$		$\|\lambda_2\| > 1$ 不安定
	位相	$\theta_1 = \pm\frac{\pi}{2}$	$\theta_2 = \pm\frac{\pi}{2}$

補足説明する.

$|p| < 1$ のとき, leapflog scheme では位相は進む. ただし, その速度は, Matsuno scheme の $1/4$ である.

θ_1 を微分すると

$$\frac{d\theta_1}{dp} = \frac{1}{\sqrt{1-p^2}}$$

となる. よって, $p \to 1$ の時, 位相の誤差は急速に増加して, $\frac{\theta_1}{p} \to \frac{\pi}{2}$ となる.

$|p| > 1$ のとき, 表の増幅係数と位相から, 2 つのモードは, $p \gtrless 1$ のとき,

$$\begin{aligned} U_1^{(n)} &= |p + \sqrt{p^2-1}|^n U_1^{(0)} e^{\pm in\frac{\pi}{2}} \\ U_2^{(n)} &= |p - \sqrt{p^2-1}|^n U_2^{(0)} e^{\pm in\frac{\pi}{2}} \end{aligned} \tag{7.47}$$

となる.

複素平面でみると, 2 つのモードは 1 ステップ毎に$\pm\frac{\pi}{2}$ で回転している. ただし, この場合, 大きさは単調に増加していく.

1 つの方程式を解くのではなく方程式系を解く場合には, physical mode のみ計算する方法がある. (ここでは触れない.)

Table 7.6: $U^{(0)}$ と $U^{(1)}$ の関係

	たまたま $U^{(1)} = U^{(0)}$	たまたま $U^{(1)} = -U^{(0)}$
すべての n に対して	$U^{(n+1)} = U^{(n)}$	$U^{(n+1)} = -U^{(n)}$
$p = 0$ の時の	$U^{(n+1)} = \lambda_1 U^{(n)}$ に対応	$U^{(n+1)} = \lambda_2 U^{(n)}$ に対応
解はすべて	physical mode となる	computational mode となる

となる. 初期条件として与えられているのは $U^{(0)}$ であるから, 式 (7.43) より, $U^{(2)}, U^{(4)}, U^{(6)}, \cdots$ は決定される. 問題は $U^{(1)}, U^{(3)}, U^{(5)}, \cdots$ である. 2 つの極端な場合を表にまとめる.

実際には, この 2 つの中間となる. ここで, 式 (7.41) は線形であるから, 解の重ね合わせがきく. よって,

$$U^{(n)} = a\lambda_1^n U_1^{(0)} + b\lambda_2^n U_2^{(0)} \tag{7.44}$$

とかける. ただし, a, b は定数である.

初期条件は,

$$\begin{aligned} U^{(0)} &= aU_1^{(0)} + bU_2^{(0)}, \\ U^{(1)} &= a\lambda_1 U_1^{(0)} + b\lambda_2 U_2^{(0)} \end{aligned} \tag{7.45}$$

となる.

a, b を消去すれば

$$U^{(n)} = \frac{1}{\lambda_1 - \lambda_2}\left[\lambda_1^n (U^{(1)} - \lambda_2 U^{(0)}) - \lambda_2^n (U^{(1)} - \lambda_1 U^{(0)})\right] \tag{7.46}$$

となる.

よって, physical mode が強く出るか, computational mode が強く出るかは, この 2 つの項のうち, どちらが大きいかによる. λ_1, λ_2, $U^{(0)}$ は定数であるから, これは $U^{(1)}$ のみによる. もし, うまく $U^{(1)} = \lambda_1 U^{(0)}$ となれば, 解はすべて physical mode となる. 不幸にして, $U^{(1)} = \lambda_2 U^{(0)}$ としてしまえば, 解はすべて computational mode となる.

このように, $U^{(1)}$ の選択は重要であるが, 実際には, 解が分からないから計算しているのであって, 仕方がないから two level scheme, 例えば Euler scheme や, Heun scheme を使って計算することになる. このとき, computational mode がどの程度強く出ているかは分からない. 理論的には, computational mode という観点からは, Heun scheme の方が優れている. また実際には丸め誤差があるから, たとえ初期値を完璧に選んでも, computational mode を完全になくすことはできない. ただ, 実際にやってみると, 大気モデル, 更に一般に偏微分方程式を解くとき, 通常, 丸め誤差による computational mode の出現は余り重要ではないことが分かる.

three level scheme 全般の議論に戻る. スキームは安定でなければならないこと, computational mode は完全には除去できないことを考えると, λ_1, λ_2 は共に 1 より小さくなければならない. ここで 3 つの場合 ($|p| < 1$, $|p| = 1$, $|p| > 1$) を考えるのが便利である.

これらを以下の表にまとめてみる.

---

[6] $\theta_2 = \pi - \theta_1$

[7] $\theta_2 = -\pi - \theta_1$

[8] 2 つのモードは

$$U^{(n)} = U^{(0)} e^{\pm in\frac{\pi}{2}}$$

とかける. よって, 複素平面でみると, 位相は 1 ステップ毎に $\pm\frac{\pi}{2}$ 変化する. 真の解では, ± 1 しか変化しない. この様に, $p = 1$ の時は, 位相の誤差は大きい.

[9] 不安定性は, $|p|$ が 1 を越えると急速に増加する. ($\frac{d|\lambda_1|}{dp}$ は, $p > 1$ において, 単調に増加する.)

7.5.4　Three level schemes の特性

Three level schemes, あるいはもっと高い次元のスキームの問題点は, 物理的に考えれば初期条件は 1 つですむはずなのに, 2 つ以上の初期条件が必要で, これをより低い次元のスキームを使って計算しなければならないということである.

また, 第 2 の問題として, 物理的な解の他に, computational mode と呼ばれる, 数値計算をした事によって発生したモードが現れる. ここでは, これらを扱う.

leapfrog scheme

three level scheme の一例として, leapfrog scheme を考える. 振動方程式に適用すると, このスキームは,

$$U^{(n+1)} = U^{(n-1)} + i2\omega\Delta t\ U^{(n)} \tag{7.38}$$

となる. 増幅係数 λ は

$$U^{(n)} = \lambda U^{(n-1)}, \qquad U^{(n+1)} = \lambda^2 U^{(n-1)} \tag{7.39}$$

である. これを式 (7.38) に代入すれば

$$\lambda^2 - i2p\lambda - 1 = 0$$

これを解いて, 2 つの増幅係数

$$\begin{aligned} \lambda_1 &= \sqrt{1-p^2} + ip, \\ \lambda_2 &= -\sqrt{1-p^2} + ip \end{aligned} \tag{7.40}$$

を得る. このように, three level scheme では, λ の解は必ず 2 つ求まってしまう. 一般に, n level scheme では, $n-1$ 個の λ が求まる. 1 つ 1 つの λ の値に対応する解を, モードという.

λ_1 は, $\Delta t \to 0$ (すなわち, $p \to 0$) のとき, $\lambda_1 \to 1$ となる. すなわち, 真の解に近づく. 一般に, この様な λ に対応する解を

physical mode !

という.

一方 λ_2 は, $\Delta t \to 0$ のとき, $\lambda_2 \to -1$ となる. すなわち, 真の解には近づかない. この様な λ に対応する解を

computational mode !

という. この computational mode は, 数値計算のみに生ずる解であり, 物理的には全く無意味である.

computational mode について

この 2 つのモードの関係をはっきりさせるため, 簡単な例を考える. 振動方程式において, $\omega = 0$ とすると,

$$\frac{dU}{dt} = 0 \tag{7.41}$$

となる. この解は

$$U = \text{const.} \tag{7.42}$$

である. さて, leapfrog scheme を式 (7.41) に適用すると,

$$U^{(n+1)} = U^{(n-1)} \tag{7.43}$$

を代入すれば,

$$\begin{aligned} U^{(n+1)*} &= U^{(n)} + i\omega\Delta t\ U^{(n)} \\ U^{(n+1)} &= U^{(n)} + i\omega\Delta t\ (\alpha U^{(n)} + \beta U^{(n+1)*}) \end{aligned} \tag{7.36}$$

$U^{(n+1)*}$ を消去して, $p \equiv \omega\Delta t$ とすれば,

$$U^{(n+1)} = (1 - \beta p^2 + ip)U^{(n)}$$

よって,

$$\lambda = 1 - \beta p^2 + ip \tag{7.37}$$

である.

$\beta = 1, \beta = \dfrac{1}{2}$ の場合についてスキームの特性を table 7.5 にまとめる.

Table 7.5: 反復する two level schemes

種類	β	$\lambda,\ \lvert\lambda\rvert$	安定性
松野 scheme	1	$\lambda = 1 - p^2 + ip$ $\lvert\lambda\rvert = (1 - p^2 + p^4)^{\frac{1}{2}}$ $\dfrac{\theta}{p} = \frac{1}{p}\arctan\frac{p}{1-p^2}$ $= 1 + \dfrac{2}{3}p^2 + \cdots$	$\lvert p\rvert \le 1 (i.e. \Delta t \le \dfrac{1}{\lvert\omega\rvert})$ の時, 安定 (条件付き安定[7]). 位相は進む. 例えば, $p = 1$ の場合, $\dfrac{\theta}{p} = \dfrac{\pi}{2}$である.
Heun scheme	$\frac{1}{2}$	$\lambda = 1 - \frac{1}{2}p^2 + ip$ $\lvert\lambda\rvert = (1 + \dfrac{1}{4}p^4)^{\frac{1}{2}}$ $= 1 + \dfrac{1}{8}p^4 + \cdots$ $\dfrac{\theta}{p} = \dfrac{1}{p}\arctan\dfrac{p}{1 - \frac{1}{2}p^2}$	常に不安定. しかし, $\lvert\lambda\rvert = 1 + O[(\Delta t)^4]$ より, 不安定性が弱く, Δt が小さければ十分使える.

特に, 松野 scheme の $|\lambda|$ を微分すれば,

$$\frac{d|\lambda|}{dp} = \frac{p}{(1 - p^2 + p^4)^{\frac{1}{2}}}(1 - 2p^2)$$

となる. よって, Matsuno scheme は, $p = \frac{1}{\sqrt{2}}$の時, $|\lambda|$ が最少となる. 最も高い周波数を 1 番強く減衰させるため, Matsuno, 1966b に従って, 必要な周波数が $0 < p < \frac{1}{\sqrt{2}}$となるように時間間隔をとり, backward implicit scheme で高い周波数を落とすということが行われる. このようにして高い周波数を落とすテクニックは, 現在の大気モデルでは, 観測されたデータを近似することによって発生する, 高い周波数のノイズを落とすため, 広く使われている. Matsuno, 1966a によれば, このテクニックは, より高い accuracy を持った scheme でも使える.

振動方程式の差分化による位相の狂いは, 余り重要ではない. まず第一に, 安定性に影響しないからであり, 第二には, これらを実際に運動方程式に適用してみると, 通常, 空間差分による位相のずれの方がより大きいからである.

[5] 松野 scheme では, $\Delta t \le \frac{1}{|\omega|}$ から明らかなように, 周波数が高い程 Δt についての条件は厳しくなる.

となる.

以後, 見やすくするため,

$$p \equiv \omega\Delta t \tag{7.33}$$

と書き, $U^{(n+1)}$ について解けば,

$$U^{(n+1)} = \frac{1+i\alpha p}{1-i\beta p}\, U^{(n)} \tag{7.34}$$

となる.

よって,

$$\lambda = \frac{1+i\alpha p}{1-i\beta p} = \frac{1}{1+\beta^2 p^2}(1-\alpha\beta p^2 + ip) \tag{7.35}$$

である.

各スキームの特性を表にまとめる.

Table 7.4: 反復しない two level schemes

種類	$(\alpha,\ \beta)$	$\lambda,\ \|\lambda\|,\ theta$	安定性
Euler scheme	(1,0)	$\lambda = 1+ip$ $\|\lambda\| = (1+p^2)^{\frac{1}{2}}$ $= 1+\frac{1}{2}p^2+\cdots$ $\frac{\theta}{p} = \frac{1}{p}\arctan p < 1$	常に不安定. $\|\lambda\|$ を展開した式は一見安定[3]のようだが, むやみに適用しても余りうまく行かない. 位相は遅れる. $p=1$ の時は, $\frac{\theta}{p} = \frac{\pi}{4}$ である.
backward scheme	(0,1)	$\lambda = = \frac{1}{1+p^2}(1+ip)$ $\|\lambda\| = = (1+p^2)^{-\frac{1}{2}}$ $\frac{\theta}{p} = = \frac{1}{p}\arctan p < 1$	Δt の値によらず安定. 特に高い周波数[4]ω で減衰し, 高周波カットのフィルターとなる. 位相は遅れる. $p=1$ の時は, $\frac{\theta}{p} = \frac{\pi}{4}$ である.
台形 scheme	$(\frac{1}{2},\ \frac{1}{2})$	$\lambda = \frac{1}{1+\frac{1}{4}p^2}(1-\frac{1}{4}p^2+ip)$ $\|\lambda\| = 1$ $\frac{\theta}{p} = \frac{1}{p}\arctan\frac{p}{1-\frac{1}{4}p^2}$	常に安定. 増幅係数については真の解同様に振舞う.

反復しないスキームのうち, 後退スキームと, 台形スキームは, Δt がどんなに大きくとも, 安定であることがわかる.

反復するスキーム

振動方程式に

$$\begin{aligned} U^{(n+1)*} &= U^{(n)} + \Delta t \cdot f^{(n)} \\ U^{(n+1)} &= U^{(n)} + \Delta t\ (\alpha f^{(n)} + \beta f^{(n+1)*}) \\ f^{(n+1)*} &\equiv f(U^{(n+1)*}, (n+1)\Delta t) \\ \alpha &+ \beta = 1 \end{aligned}$$

[3] $|\lambda| = 1 + O[(\Delta t)^2]$ だから, 辛うじて von Neumann の条件を満たすかに見える.

[4] 高い周波数は, 初期条件に左右されやすく, 異常に大きな値が出やすいため, 実際の計算では, 高い周波数をわざと減衰させたいことが多い.

7.5　振動方程式に適用したときのスキームの特性

7.5.1　Introduction

数値解の振舞いは, どの様な方程式に対して適用したかによって異なる. ここでは, 振動方程式に対して適用した場合を調べてみる.

振動方程式とは, $f(U,t) = i\omega U$, すなわち

$$\frac{dU}{dt} = i\omega U, \qquad U = U(t)$$

という形をした方程式のことである. ただし, ω は実数とする.

この解は,

$$U(t) = U(0)e^{i\omega t}$$

である. 離散的な値　$t = n\Delta t$　に対しては

$$U(n\Delta t) = U(0)e^{in\omega\Delta t}$$

となる.

7.5.2　特性の調べ方

この方程式の数値解の性質は, von Neumann 法を用いて調べるのが便利である.

増幅係数 λ (第 7.2.4 節参照) を見れば

$$\begin{array}{ll} \text{不安定} & > \\ \text{中立} & |\lambda| = 1 \\ \text{減衰} & < \end{array}$$

といえる.

また, 1 タイムステップ毎の位相の変化θを見れば, 位相は

$$\begin{array}{ll} \text{進む} & > \\ \text{中立} & \dfrac{\theta}{\omega\Delta t} = 1 \\ \text{遅れる} & < \end{array}$$

といえる.

以下, 実際のスキームに適用する.

7.5.3　two level schemes の特性

反復しないスキーム

反復しないスキーム

$$\begin{aligned} U^{(n+1)} &= U^{(n)} + \Delta t\ (\alpha f^{(n)} + \beta f^{(n+1)}) \\ f^{(n)} &\equiv f(U^{(n)}, n\Delta t) \end{aligned}$$

を振動方程式に適用すれば,

$$U^{(n+1)} = U^{(n)} + i\omega\Delta t\ (\alpha U^{(n)} + \beta U^{(n+1)}) \tag{7.32}$$

7.4 semi-Implicit 法

7.4.1 Introduction

気象業界においては, 速やかに変動する現象 (雲の発達など) と, 緩やかに変動する現象 (低気圧の発達など) とが入り交じっている. しかも, この 2 つの現象は相互に影響を及ぼし合っている. さて, 我々に興味があるのは, 主として緩やかに変動する現象である. 緩やかに変動する現象を観察するためには, 長時間積分しなければならない. しかし, 方程式を積分するためには, 速やかに変動する現象も計算しなければならない. ところが, 安定性を保つためには, 時間差分 Δt を空間差分に比べ必要以上に小さく取らねばならない. こうすると, 時間も金もかかる. なんとか楽をして解を求めよう, 多少の誤差には目をつぶろう, どうせ色々な所で誤差が入ってくるのだ, 大きく狂わなければいいではないか, というのが以下の話である.

以下, 様々なごまかしをする. これが許容誤差範囲内で正しいかどうかは, 実際の天気図と照らし合わせてみても分からない. 自然現象をモデル化するときに誤差が入ったのか, 積分するときにごまかしをしたが効いているのか, 判断できないからである. そこで, 通常, なにかごまかしをする時には, 時間差分を小さく取った真面目な計算を 1 回行い, これを control とする

7.4.2 semi-Implicit 法の原理

気象学において, 速やかに変動する現象には余り興味がない. しかし, ないと困る. そこで, 速い現象はいい加減に計算し, 緩やかに変動する現象は真面目に計算しよう, という考えが浮かぶ. ここで, 振動方程式

$$\frac{\partial F}{\partial t} = i(\alpha + \beta)F$$

を考える. ここで, α は速い現象に対応し, β は遅い現象に対応する. このとき

$$\alpha \gg \beta$$

である. これに対して

$$\frac{F^{(n+1)} - F^{(n-1)}}{2\Delta t} = i\alpha\frac{F^{(n+1)} + F^{(n-1)}}{2} + i\beta F^{(n)}$$

というスキームをくむ. これが大した面倒もなく解けるならば, 時間差分 Δt は通常, 全部をまじめに計算する場合の $3 \sim 4$ 倍に取ることができる (これは経験則) ので, 有効である. これを semi-Implicit 法という. このスキームの増幅係数を λ とした時, λ^2 を計算すると

$$\text{Physical mode} \quad : \quad 1 + \frac{4\alpha\beta\Delta t^2\sqrt{1+\alpha^2\Delta t^2}}{1+\alpha^2\Delta t^{2\,2}}$$

$$\text{Computational mode}^{\,1} \quad : \quad 1 - \frac{4\alpha\beta\Delta t^2\sqrt{1+\alpha^2\Delta t^2}}{1+\alpha^2\Delta t^{2\,2}}$$

となる. ここで, $\alpha \gg \beta$ という条件からβ の 2 次以上の項を落とした. これより, Physical mode は発散気味, Computational mode は抑制気味であることがわかる.

る. 一方, 偏微分方程式では, 初期条件を離散的にしか与えなかったことによる誤差が無視出来ない. スキームの時間に関する accuracy だけを上げても無意味である. ところが, 空間に関する accuracy を上げようにも, 実際問題としてはそれを定めるだけの初期値のデータ密度がないため無駄, とされたのである.

- 安定な解を得るためには, 時間間隔を十分に小さく取らねばならない. この時, 誤差は主として空間差分など他の原因で起きており, 時間微分項だけがんばって計算しても無意味である.

もちろん, だからスキームはどうでもいいのではない. 正確度以外の特性にも十分注意を払わねばならない, ということである.

反復するスキーム

反復するスキームとは前進 scheme を使って $n+1$ ステップでの仮の値を計算し (これを $*$ で表す), 本来 implicit scheme である A2, A3 のスキームに仮の値を代入するものである. こうすることにより, A2, A3 のスキームは explicit となる. 反復すると, 安定性が良くなる.

反復する two level scheme は, まとめて

$$\begin{aligned} U^{(n+1)*} &= U^{(n)} + \Delta t \cdot f^{(n)} \\ U^{(n+1)} &= U^{(n)} + \Delta t\ (\alpha f^{(n)} + \beta f^{(n+1)*}) \\ f^{(n+1)*} &\equiv f(U^{(n+1)*}, (n+1)\Delta t) \\ \alpha &+ \beta = 1 \end{aligned} \tag{7.31}$$

とかける. 表 2 に主な反復する two level scheme を示す.

Table 7.2: 反復する two level schemes

	種類	(α, β)	代入されるスキーム	陽 ・陰	正確度
A4	松野 scheme[5]	(0,1)	後退スキーム (A2)	explicit	1 次
A5	Heun scheme	$(\frac{1}{2}, \frac{1}{2})$	台形スキーム (A3)	explicit	2 次

7.3.10　three level scheme の定義

$U^{(n+1)}$ を計算するのに, 何等かの形で $U^{(n-1)}$ を使うスキームを three level scheme と呼ぶ. もちろん, 最初のステップでは使えない. また, 自由度が 1 つ増えるため, computational mode の問題が生じる.

主なスキームを表で示す.

Table 7.3: 主な three level schemes

	種類	式	陽 ・陰	正確度
B1	Leapfrog scheme [2]	$U^{(n+1)} = U^{(n-1)} + 2\Delta t \cdot f^{(n)}$	explicit	2 次
B2	Adames-Bashforth scheme	$U^{(n+1)} = U^{(n)} + \Delta t(\frac{3}{2} f^{(n)} - \frac{1}{2} f^{(n-1)})$	explicit	2 次 [3]

7.3.11　補足

要は 式 (7.29) の積分をいかに正しくやるかであるから, 上表の他, Simpson 則を用いたスキーム (これらをまとめて Malne-Simpson scheme という) など, 他にもたくさんある. これらは, Young, 1968 を参照.

通常気象学では, 1 次か 2 次といった低次の accuracy を持つスキームが使われる. この理由は,

- 偏微分方程式では, 常微分方程式とは異なり, 高次の accuracy を持つスキームを使っても (少なくともこれまでは) 余り意味がないとされてきた. 常微分方程式では, 初期条件として 1 点での関数及び必要な導関数の値を与えれば十分であり, それで答えが出なかったら, それはスキームが悪いことにな

[5] Euler-backward scheme ともいう.
[2] 気象のモデル作りで広く使われる.
[3] 本来の Adams-Bashforth scheme は, 4 次の正確度を持つ.

- explicit scheme (陽解法)

  f を計算するのに, n ステップ目までの値のみを使うもの. すなわち, 既に求まっているもののみを用いるスキーム.

の 2 つに分けられる.

偏微分方程式系の implicit scheme では, $U^{(n+1)}$ を求めるために連立方程式を解くことになる為, 計算量が膨大になることが多い. しかし, 常微分方程式ならば, その操作は比較的容易である.

差分は, 格子点と, 格子点を使って求めた導関数との関係によって,

- 中心差分 (centered difference)

  使用した格子点の中心点における導関数を求めるもの. たとえば

$$\Delta u_{j+\frac{1}{2}} = u_{j+1} - u_j$$

- 非中心差分 (uncentered difference)

  格子点の中心ではない点における導関数を求めるもの. たとえば

$$\Delta u_j = u_{j+1} - u_j$$

という 2 つに分けられる. 中心差分の方が通常良く使われる.

7.3.9　two level scheme の定義

最初のステップでは, two level scheme しか使えない.

反復しないスキーム (仮の値を使わないスキーム)

反復しない two level scheme は, まとめて

$$\begin{aligned} U^{(n+1)} &= U^{(n)} + \Delta t \left(\alpha f^{(n)} + \beta f^{(n+1)}\right) \\ f^{(n)} &\equiv f(U^{(n)}, n\Delta t) \end{aligned} \tag{7.30}$$

とかける. ただし, スキームには適合性がなければならないから

$$\alpha + \beta = 1$$

である. 表 1 に主な反復しない two level scheme を示す.

Table 7.1: 反復しない two level schemes

	種類	(α, β)	中心・非中心	陽・陰	正確度
A1	前進 scheme[2]	(1,0)	uncentered	explicit	1 次[3]
A2	後退 scheme [4]	(0,1)	uncentered	implicit	1 次 [3]
A3	台形 scheme	$(\frac{1}{2}, \frac{1}{2})$	centered	implicit	2 次

[2] Euler scheme ともいう.
[3] 一般に two level の uncentered scheme は, 1 次の accuracy を持つ.
[4] backward scheme ともいう.

例題 3

二階常微分方程式

$$\frac{d^2y}{dx^2} = x\frac{dy}{dx} + y$$

を, 初期条件

$$(x_0, y_0) = (0, 1), \quad \left(\frac{dy}{dx}\right)_{x=x_0, y=y_0} = 1$$

のもとで Runge-Kutta 法で解け. もともとの方程式を

$$\begin{cases} \dfrac{dy_1}{dx} = y_2 \\ \dfrac{dy_2}{dx} = xy_2 + y_1 \end{cases}$$

という連立常微分方程式に変換して求めれば良い.

7.3.7　時間発展問題

さて次は具体的な時間発展的な問題の考察に入る。

$$\frac{dU}{dt} = f(U, t), \qquad U = U(t) \tag{7.28}$$

ただし, ここでの独立変数 t は時間とする. 時間の grid interval をΔt とし, 時間 $n\Delta t$ における U の近似値を $U^{(n)}$ と書く.

$$f^{(n)} \equiv f(U^{(n)}, n\Delta t)$$

ここでの問題は, $U^{(n)}$, $U^{(n-1)}$, $\cdots$ が分かっていた時, どうやって $U^{(n+1)}$ を求めるか, ということである. その答えは解析的には,

$$U^{(n+1)} = U^{(n)} + \int_{n\Delta t}^{(n+1)\Delta t} f(U, t)dt \tag{7.29}$$

となる. 数値計算では (7.29) を近似して計算することになる. その近似方法をスキームと呼ぶ. スキームにはたくさんの種類がある. 以下では, まずスキームの分類を行い, そのあと時間微分項を積分するために用いられるいくつかのスキームを定義しよう.

7.3.8　スキームの分類

スキームは, n や $n+1$ といった各時間 level における依存変数を関係付ける. スキームは時間 level を幾つ用いるかで次のように分類される.

- two level scheme : n と $n+1$ の 2 つを用いたもの.
- three level scheme : 更に $n-1$ も加え 3 つを用いたもの.

また, そのレベルの使い方からスキームは

- implicit scheme (陰解法)

  $U^{(n+1)}$ の値を決めるために用いる f を計算する際に, $n+1$ ステップ目での値 $U^{(n+1)}$, すなわち, これから求められるべきものを用いるスキーム.

すなわち

$$\begin{bmatrix} k_{11} \\ k_{12} \\ \vdots \\ k_{1n} \end{bmatrix} = \begin{bmatrix} hf_1(x_0, y_{10}, y_{20}, \cdots, y_{n0}) \\ hf_2(x_0, y_{10}, y_{20}, \cdots, y_{n0}) \\ \vdots \\ hf_n(x_0, y_{10}, y_{20}, \cdots, y_{n0}) \end{bmatrix}$$

のように表される. 以下同様に k_2, k_3, k_4 についてもその値が求められ,

$$\begin{aligned} \boldsymbol{k} &= \frac{1}{6}(\boldsymbol{k}_1 + 2\boldsymbol{k}_2 + 2\boldsymbol{k}_3 + \boldsymbol{k}_4) \\ &= \frac{1}{6}\left\{ \begin{bmatrix} k_{11} \\ k_{12} \\ \vdots \\ k_{1n} \end{bmatrix} + 2 \begin{bmatrix} k_{21} \\ k_{22} \\ \vdots \\ k_{2n} \end{bmatrix} + 2 \begin{bmatrix} k_{31} \\ k_{32} \\ \vdots \\ k_{3n} \end{bmatrix} + \begin{bmatrix} k_{41} \\ k_{42} \\ \vdots \\ k_{4n} \end{bmatrix} \right\} \end{aligned}$$

が求められる. これより,

$$\boldsymbol{y}_1 = \boldsymbol{y}_0 + \boldsymbol{k}$$

として次のステップの値を求めることができる.

例題 1

微分方程式

$$\frac{dx}{dt} = 1 - x$$

初期条件を

$$x(0) = 0$$

のもとで, 上に示した三つの解法について h の刻幅を変えて計算し, 解析解との差 (誤差) を調べよ. ちなみに, 解析解は

$$x(t) = 1 - e^{-t}$$

である.

例題 2

次の常微分方程式を Runge-Kutta 法で解け. 但し計算区間は $[0, 15]$ とし, 計算間隔 h は 0.5 と 0.25 の両者について計算を行ない, 真の値と比較検討せよ.

$$\frac{dy}{dx} = -xy, \quad y(0) = 15$$

この方程式の一般解は

$$y = c \exp\left(-\frac{x^2}{2}\right) \qquad (c \text{ は定数})$$

である.

$$
\begin{aligned}
k_3 &= hf(t_n + \frac{h}{2}, x_n + \frac{k_2}{2}) \\
k_4 &= hf(t_n + h, x_n + k_3)
\end{aligned}
$$

である. k_1, k_2, k_3, k_4を求めた後, $x_{n+1} - x_n$の打ち切り誤差が h^4 の項まで消えるように重み $k_1 : k_2 : k_3 : k_4 = 1 : 2 : 2 : 1$ をとして平均した量を求め, それを kとして $t_{n+1} = t_n + k$により x_nを定める (図 7.6 参照) .

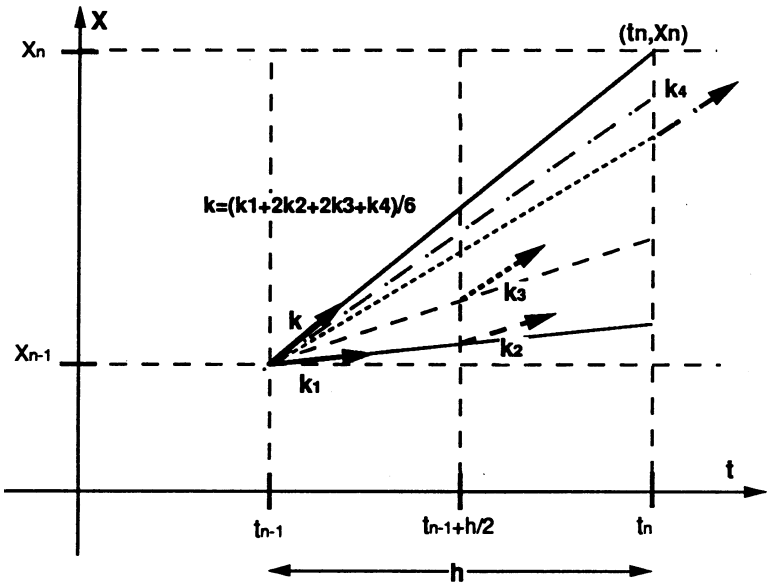


図 7.6: Runge-Kutta 法

7.3.6　連立常微分方程式

前述の連立常微分方程式を Runge-Kutta 法を用いて解くことを考えてみる. ここでは, 連立常微分方程式をベクトル形式で

$$
\frac{d\boldsymbol{y}}{dx} = \boldsymbol{f}(x, \boldsymbol{y})
$$

と表すことにする. ここで

$$
\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \boldsymbol{f} = \begin{bmatrix} f_1(x, y_1, y_2, \cdots, y_n) \\ f_2(x, y_1, y_2, \cdots, y_n) \\ \vdots \\ f_n(x, y_1, y_2, \cdots, y_n) \end{bmatrix}
$$

とする. このようなベクトル表示を用いれば, Runge-Kutta 法において k_1を求める式は

$$
\boldsymbol{k}_1 = h\boldsymbol{f}(x_0, \boldsymbol{y}_0)
$$

$$x_{n+1} - x_n = \frac{1}{2}\left(f(t_n, x_n) + f(t_{n+1}, x_{n+1})\right)$$

が得られる (数値積分で言えば台形則) . これを修正 Euler 法と呼ぶ. 具体的には, 図 7.5に示すように, まず点 (t_n, x_n) での傾き k_1 を計算し, 次に 1 ステップ Euler 法で前進したところで再び傾き k_2 を計算する. これらの平均の傾き $\frac{1}{2}(k_1 + k_2)$ を (t_n, x_n) での傾きと定める.

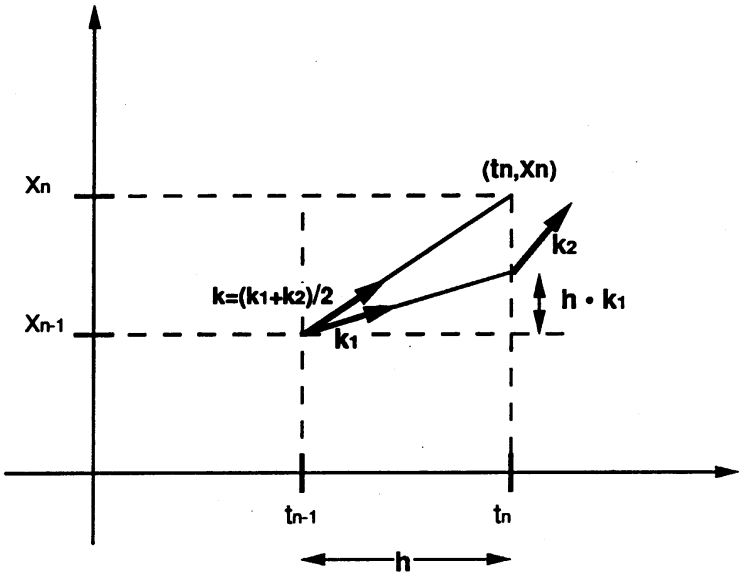


図 7.5: 修正 Euler 法

(t_{n+1}, x_{n+1}) を求める手続きは,

$$\begin{aligned} k_1 &= f(t_n, x_n) \\ k_2 &= f(t_n + h, x_n + hk_1) \\ x_{n+1} &= x_n + \frac{h}{2}(k_1 + k_2) \\ t_{n+1} &= t_n + h \end{aligned}$$

と表せる. この方法によって精度は Euler 法に比べて良くなり, $O(h^2)$ で評価される.

7.3.5　Runge-Kutta 法

修正 Euler 法では (t_n, x_n) から次のステップに進む時の傾きとして二つの傾きの平均値を用いたが, Runge-Kutta 法では四つの量を用いる. t_{n+1} での x_{n+1} の値は次の諸式で与えられる.

$$x_{n+1} = x_n + \frac{k_1 + 2k_2 + 2k_3 + k_4}{6}$$

ここで

$$\begin{aligned} k_1 &= hf(t_n, x_n) \\ k_2 &= hf(t_n + \frac{h}{2}, x_n + \frac{k_1}{2}) \end{aligned}$$

7.3.3　Euler 法

微分方程式 (7.26) を初期値 x_0, x_1のもとで

$$\begin{aligned} x_1 &= x_0 + hf(t_0, x_0) \\ x_2 &= x_1 + hf(t_1, x_1) \\ &\cdots \\ x_m &= x_0 + hf(t_{m-1}, x_{m-1}) \end{aligned}$$

のように前の点に於ける傾きを辿って行って解く方法を Euler 法と言う (図 7.4参照).

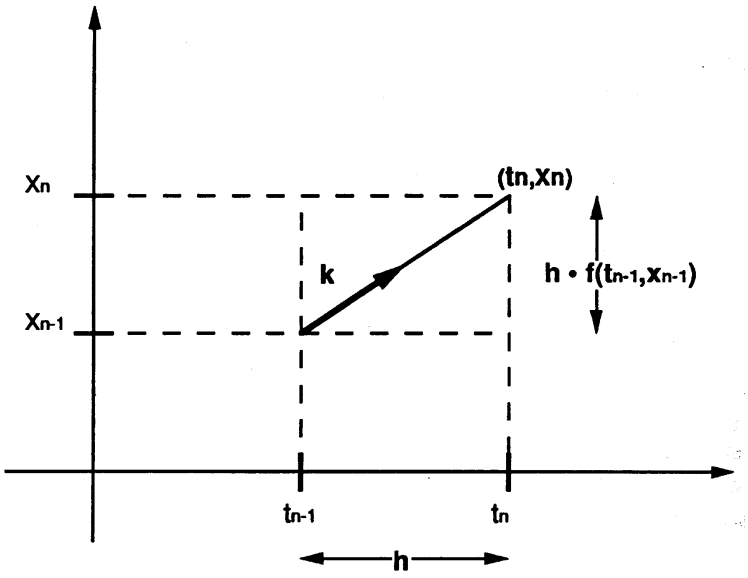


図 7.4: Euler 法

これは微分方程式 (7.26) を t_nから t_{n+1}まで積分した時, 右辺の $f(t,x)$ がその区間で一定であると近似して得られる方法であり, 数値積分法では区分求積法に対応している. ここで,

$$\begin{aligned} x(t+h) &= x(t) + hx'(t) + \frac{1}{2}h^2 x''(\xi) \qquad (t < \xi < t+h) \\ &= x(t) + hf(t,x) + \frac{1}{2}h^2 \frac{df(\xi, x(\xi))}{dt} \end{aligned}$$

なので, 式 (7.27) で表されるこの解法には

$$\frac{1}{2}h^2 \frac{df(\xi, x(\xi))}{dt}$$

に相当する誤差が含まれていることになる. これは後述する二つの方法に比べて非常に悪い.

7.3.4　修正 Euler 法

微分方程式 (7.26) を t_nから t_{n+1}まで積分するとき, $f(t,x)$ がその区間で $f(t_n, x_n)$ と $f(t_{n+1}, x_{n+1})$ を通る直線であると近似すると,

7.3　常微分方程式を解く各種スキーム

7.3.1　Introduction

ここでは, 常微分方程式を解く各種 scheme の分類と定義について説明する. 偏微分方程式を解く scheme を考える場合でも, 常微分方程式を解く scheme についての知識を得ておくことは必要である. なぜなら

- 偏微分方程式の時間微分項の近似をするとき, そのまま使える
- スペクトル法では, 偏微分方程式は, 常微分方程式の組に帰着される
- Lagrange 的な記述を用いると, 常微分方程式の組が出てくる

といった理由があげられるからである.

7.3.2　常微分方程式の一般的な解法

まずは非常に一般的に常微分方程式を解くことを考える. ここでは,

- Euler 法
- 修正 Euler 法
- Runge-Kutta 法

の三方法について簡単に解説する. このほかにも Runge-Kutta-Gill 法, Milne 法などの解法があるが, ここでは触れない.

独立変数を t, t の未知関数を $x(t)$ として, $x(t)$ の満たすべき微分方程式と初期条件

$$\frac{d}{dx}t = f(t,x) \quad (a \le t \le b) \tag{7.26}$$
$$x(a) = x_0$$

を与えて関数 $x(t)(a \le t \le b)$ を求めるという問題を考えてみる. ここで, 区間 $[a,b]$, 右辺 $f(t,x)$ の関数, 初期値 x_0を与えることにより常微分方程式の初期値問題を定めることができる. この問題を離散的に扱うためには変数 t の区間 $[a,b]$ を $h=(b-a)/m$ によって分解し, $t_n = a+nh(n=0,1,2,\cdots,m)$ における解を数値的に求めればよい. もっとも簡単な方法としては, 式 (7.26) の左辺を

$$\frac{dx}{dt} \simeq \frac{x_{n+1}-x_n}{t_{n+1}-t_n} = \frac{x_{n+1}-x_n}{h} \tag{7.27}$$

で置き代えて整理し,

$$x_{n+1} = x_n + hf_n(t_n, x_n)$$

としたもので, Euler 法と呼ばれる. 連立微分方程式の場合も未知関数が $x^1(t), x^2(t), \cdots, x^l(t)$ と l個あること以外には形式的に全く同じで, 微分方程式

$$\frac{dx^i}{dt} = f^i(t, x^1, \cdots, x^l) \quad (i = 1, \cdots, l \ ; \ a \le t \le b)$$

と, 初期条件

$$x^i(a) = x^i{}_0 \quad (i, \cdots, l)$$

を与えて $x^i(t)(a \le t \le b)$ を求めることが問題になる. 更に高階の微分方程式の場合も,

$$x^1 = x, \quad x^2 = \frac{dx}{dt}, \quad x^3 = \frac{d^2x}{dt^2}, \cdots$$

などとおけば連立微分方程式の場合に帰着する.

$$1 - \mu \geq 0$$

と求められた. これは, 直接法やエネルギー法の結果と同じものである.

一般に, 真の解では振幅は変わらない. 即ち $|\lambda| = 1$ である. 数値解の $|\lambda|$ の値により

$$\begin{array}{cc} \text{不安定} & |\lambda| > 1 \\ \text{中立} & |\lambda| = 1 \\ \text{減衰} & |\lambda| < 1 \end{array}$$

と分類できる. もちろん $|\lambda|$ はできるだけ 1 に近い方が望ましい.

次に, λ の位相について考える. λ を, 実数部と虚数部とにわけ,

$$\lambda \equiv \lambda_{re} + i\lambda_{im}$$

と書く. この時, 式 (7.20) より数値解の $|\lambda|$ の位相は,

$$\theta = \arctan \frac{\lambda_{im}}{\lambda_{re}}$$

あるいは, 数値解の $|\lambda|$ の位相と $\omega\Delta t$　(ただし, $\omega = ck$) との比をとると

$$\frac{\theta}{\omega\Delta t} = \frac{1}{\omega\Delta t} \arctan \frac{\lambda_{im}}{\lambda_{re}}$$

となる. $\frac{\theta}{\omega\Delta t}$の値によって

数値解の位相は調和解の位相より

$$\begin{array}{cc} \text{進む} & \frac{\theta}{\omega\Delta t} > 1 \\ \text{中立} & \frac{\theta}{\omega\Delta t} = 1 \\ \text{遅れる} & \frac{\theta}{\omega\Delta t} < 1 \end{array}$$

と分類できる. もちろん, $\frac{\theta}{\omega\Delta t}$ もできるだけ 1 に近いことが望ましい. また, computational mode [2] の $|\lambda|$ は, 出来るだけ小さいことが望ましい.

[2] computational mode とは, 物理的な解とは関係のない, スキームによって発生する, 数値計算特有のものである.

と定義する. また

$$\lambda \equiv |\lambda| e^{i\theta} \tag{7.20}$$

とする. このとき, (7.19) と (7.20) から

$$U^{(n)} = |\lambda|^n U^{(0)} e^{in\theta} \tag{7.21}$$

となる.θ は, 1 タイムステップ毎の位相の変化を表す. (7.19) の絶対値をとると

$$|U^{(n+1)}| = |\lambda||U^{(n)}|$$

となる. 式 (7.17) の各々の調和解が安定であるためには, ある正数 B が存在して,

$$|U^{(n)}| = |\lambda|^n |U^{(0)}| < B$$

であることが必要十分である. この条件式の対数をとると

$$n \ln |\lambda| < \ln(B/|U^{(0)}|) \equiv B'$$

となる.B' は新たな定数である. さらに $n = \frac{t}{\Delta t}$ を使うと, 安定であるための条件は

$$\ln |\lambda| < \frac{B'}{t} \Delta t \tag{7.22}$$

となる.

一方, 条件 (7.22) は,

$$\ln |\lambda| \leq O(\Delta t)$$

と書ける. ここで δ を

$$|\lambda| \equiv 1 + \delta$$

と定義する. 安定性の条件を $ln(1+\delta)$ のべき展開の形で書くと

$$\delta \leq O(\Delta t)$$

となる.$|\lambda| = 1 + \delta$ を再び使うと

$$|\lambda| \leq 1 + O(\Delta t) \tag{7.23}$$

となる. これが von Neumann による安定性の必要条件である.

von Neumann の条件では, 解は, あまり速く成長しなければ, 指数的に成長することが許される. しかし, 例 (7.16) のように, 真の解が成長しないと分かっている時は, 式 (7.23) を十分条件

$$|\lambda| \leq 1 \tag{7.24}$$

で置き換えてしまう. この条件は, 元の安定性の条件と比べて, それほど緩いものではない. 例に戻ると, 式 (7.19) を式 (7.18) に代入すれば,

$$\lambda = 1 - \mu + \mu e^{-ik\Delta x}$$

となる. これより,

$$|\lambda|^2 = 1 - 2\mu(1-\mu)(1-\cos k\Delta x) \tag{7.25}$$

が得られる. よって安定性の条件 (7.24) は

である. 式 (7.13) と (7.14) を使うと, もし $1-\mu \geq 0$ ならば, 式 (7.12) から

$$\sum_{j=0}^{J-1}(u_j^{n+1})^2 \leq \left[(1-\mu)^2+2\mu(1-\mu)+\mu^2\right]\sum_{j=0}^{J-1}(u_j^n)^2$$

すなわち

$$\sum_j (u_j^{n+1})^2 \leq \sum_j (u_j^n)^2$$

という不等式が得られることが分かる. よって,

$$1-\mu \geq 0$$

は, 周期的境界条件下では式 (7.8) の安定性の十分条件となることが証明された.

von Neumann 法 ここでは, von Neumann 法, 別名フーリエ級数法について述べる. この方法は, スキームの安定性を調べるのにもっともよく使われる方法である.

それは, 次のような手順で行われる. まず, 解をフーリエ級数に分解する. そして 1 つ 1 つの成分の安定性を調べる. もしこれで適当な成分全てが安定ならば, スキームの安定性が示されたことになる.

通常 von Neumann 法は非線形方程式の安定性を調べることはできない. そこで, 非線形方程式の安定性を調べる代わりに, その線形化した方程式の解析をする.

この方法を具体的に説明するにあたって, 最初に元の移流方程式 (7.6) の解析解を調和解

$$u(x,t) = Re\left[U(t)e^{ikx}\right] \tag{7.15}$$

の形で求めておく. ここで, $U(t)$ は波の強度, k は波数である. 式 (7.15) を前の方程式に代入して,

$$\frac{dU}{dt} + ikcU = 0$$

を得る. このようにして, 偏微分方程式を解く問題が常微分方程式を解くことに帰着したのである. この解は

$$U(t) = U(0)e^{-ikct}$$

ただし, $U(0)$ は波の強度の初期値である. よって, 求める解析解は

$$u(x,t) = Re\left[U(0)e^{ik(x-ct)}\right] \tag{7.16}$$

である. 各々の波の要素は, このように, 定速度 c で強度を変化させずに x 軸方向へと進んでいく.

さて, (7.6) の解析解が求まったので, 次にこれに対応する有限差分方程式 (7.8) の解の振舞いを考える. (7.7) の差分形として

$$u_j^n = Re\left[U^{(n)}e^{ikj\Delta x}\right] \tag{7.17}$$

という形を考える. ただし, $U^{(n)}$ は時間ステップ n における強度である.

これを (7.8) に代入すれば,

$$U^{(n+1)} = (1-\mu)U^{(n)} + \mu U^{(n)}e^{-ik\Delta x} \tag{7.18}$$

となる. この方程式を使えば, n が増加する時 $U^{(n)}$ がどのようにふるまうか解析出来る. まず, 増幅係数 λ を

$$U^{(n+1)} \equiv \lambda U^{(n)} \tag{7.19}$$

一方, $\mathrm{Max}_{(j)}|u_j^n|$ は定義により

$$|u_l^n| \le \mathrm{Max}_{(j)}|u_j^n|, \quad |u_{l-1}^n| \le \mathrm{Max}_{(j)}|u_j^n|$$

である. よって

$$(1-\mu)|u_l^n| + \mu|u_{l-1}^n| \le \mathrm{Max}_{(j)}|u_j^n| \tag{7.11}$$

が成り立つ. 式 (7.10),(7.11) をあわせれば, 結局, 任意の n に対して

$$\mathrm{Max}_{(j)}|u_j^{n+1}| \le \mathrm{Max}_{(j)}|u_j^n|$$

となる. これは数値解の有界性を示す. よって,

$$1-\mu \ge 0$$

は, スキーム (7.8) の安定性の十分条件であることがわかる.

エネルギー法　ここでは, 真の解が有界であることが分かっている時, u_j^n のノルム $\|u^n\|$ もまた有界かどうかを調べる. もしノルムが有界なら, 誤差 $u_j^n - u(j\Delta x, n\Delta t)$ は有界であり, スキームは安定である. 特に, ノルムとして $\|u^n\| = \sqrt{\sum_j (u_j^n)^2}$ をとると, $\sum_j (u_j^n)^2$ の有界性を調べることによってノルム $\|u^n\|$ の有界性を示すことができる. このようにしてスキームの安定性を調べる方法をエネルギー法という. 物理学では, u^2 は多くの場合なんらかの形態のエネルギーに比例しているのでこう呼ばれる. この方法は, 大変広い適用範囲を持っており (実はあまり正しくないが), 非線形方程式に対してすら適用できる. これより先,

$$u_0 \equiv u_J$$

という周期的境界条件の下で考える.

式 (7.8) の両辺を 2 乗し, j について和をとれば

$$\sum_{j=0}^{J-1} (u_j^{n+1})^2 = \sum_{j=0}^{J-1} \left[(1-\mu)^2 (u_j^n)^2 + 2\mu(1-\mu) u_j^n u_{j-1}^n + \mu^2 (u_{j-1}^n)^2\right] \tag{7.12}$$

となる. 周期的境界条件により,

$$\sum_{j=0}^{J-1} (u_{j-1}^n)^2 = \sum_{j=0}^{J-1} (u_j^n)^2 \tag{7.13}$$

となる. さて, Schwartz の不等式

$$\sum_{j=0}^{J-1} a_j b_j \le \sqrt{\sum_{j=0}^{J-1} a_j^2} \sqrt{\sum_{j=0}^{J-1} b_j^2}$$

と式 (7.13) から,

$$\begin{aligned} \sum_{j=0}^{J-1} u_j^n u_{j-1}^n &\le \sqrt{\sum_{j=0}^{J-1} (u_j^n)^2 \sum_{j=0}^{J-1} (u_{j-1}^n} \\ &= \sum_{j=0}^{J-1} (u_j^n)^2 \end{aligned} \tag{7.14}$$

7.2.4 安定性の判定法

Introduction 数値解の安定性を調べる方法はいろいろある.

- 直接法

  誤差, あるいは数値解が有界であることを直接調べる.

- エネルギー法

  数値解のノルムが有界であるかどうかを調べる.

- von Neumann 法 (フーリエ級数法)

  解をフーリエ級数に分解し, 一つ一つの成分の安定性を調べる.

ここでは, 1 次元線形移流方程式

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0; \quad u = u(x,t), \quad c = \text{正の定数} \tag{7.6}$$

を差分化した式

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + c\frac{u_j^n - u_{j-1}^n}{\Delta x} = 0 \tag{7.7}$$

にこれらの方法を適用し, 数値解が安定であるための条件を求める.

直接法 誤差 $u_j^n - u(j\Delta x, n\Delta t)$ あるいは数値解 u_j^n が有界であることを直接調べる方法を直接法という. この方法は簡明であるが, かなり限られた数のスキームに対してしか適用できない. ここでは, スキーム (7.7) から得られる数値解の有界性を示す.

スキーム (7.7) は

$$u_j^{n+1} = (1-\mu)u_j^n + \mu u_{j-1}^n \tag{7.8}$$

と書ける. ただし,

$$\mu \equiv c\frac{\Delta t}{\Delta x}$$

である. もし $1-\mu \geq 0$ ならば, $c \geq 0$ に限っているので$\mu \geq 0$ であるから,

$$|u_j^{n+1}| \leq (1-\mu)|u_j^n| + \mu|u_{j-1}^n| \tag{7.9}$$

となる. これはたまたま適合性の必要条件でもある.

さて, 左辺を考えると, ある n に注目して j をいろいろ変化させたときの $|u_j^{n+1}|$ の最大値を $\mathrm{Max}_{(j)}|u_j^{n+1}|$ と書く. その最大値を与える j を l とすると

$$\mathrm{Max}_{(j)}|u_j^{n+1}| = |u_l^{n+1}|$$

である. このときも式 (7.9) は成立するから,

$$\mathrm{Max}_{(j)}|u_j^{n+1}| = |u_l^{n+1}| \leq (1-\mu)|u_l^n| + \mu|u_{l-1}^n| \tag{7.10}$$

となる. これが任意の n に対して成立する.

定性と収束性との関係については, Richtmyer and Morton, 1967; Kreiss and Oliger, 1973; Ames, 1969 といった本を参照すればよい[1].

[1] しかし, 大気モデルに数値計算法を用いるためならば, ここに書かれている以上に安定性や収束性の概念を精密にする前にすることがある. すなわち, 位相速度の誤差, 非線形不安定性, 時空間格子が数値解に及ぼす影響等を議論する方がより重要なのである.

線上では, u の値は全て等しいから, 図 7.3における原点の u の値と, 点 A における u の値は等しい. よって本来ならば, 原点での u の値を計算することのみによって, A における u の値が決められなければならない. ところが実際には, 点 A の数値解は, スキーム (7.3) を見ると, 白丸をつけた格子点を使って計算されている. つまり, この白丸をつけた領域が依存領域となる. 原点は依存領域の外となっている. 従って, 点 A における数値解は, 原点での値によっては決められない. よって, 例えば, 依存領域の初期値として適当な値をとれば, 点 A での値は原点での値とは関係のない任意の値をとり得る. 逆にいうと, 誤差はいくらでも大きくなり得る. Δx と Δt との比率が等しいままなら, $\Delta x,\ \Delta t \to 0$ としても状況はなんら改善されない. 原点はやはり依存領域外にある. 収束性はないままである. 収束するための必要十分条件は, 数値解を求めようとしている点の特性曲線が, その点の依存領域内部にあることである. この場合は,

$$c\Delta t \leq \Delta x$$

である. これがスキーム (7.3) が収束するための必要条件である. これは CFL condition と呼ばれる.

安定性 (**stability**)　差分間隔 Δx , Δt が一定のとき, ステップ数 n が増加しても, 即ちどんなに時間がたっても, 数値解の絶対値が発散しないならば, その数値解は安定であるという. また, 任意の初期条件に対し, 安定な数値解を与える時, そのスキームは安定であるという. 安定性は, 実用上大切な概念である. 適合性があって, 高次の正確度があり, その上, 解が収束する差分スキームでも, $n \to \infty$ にともなって, 真の解から速やかに発散するスキームだと計算不可能となってしまう.

数値解の安定性は, 厳密には関数解析の概念を用いて定義され, 解の有界性に依存する. 真の解が有界であると分かっている時には, 誤差 $u_j^n - u(j\Delta x, n\Delta t)$ の有界性を用いて安定性を定義出来る. すなわち, 真の解が有界とわかっていて, 差分間隔Δx , Δt 一定の条件のもとで n を増加させた時, 誤差 $u_j^n - u(j\Delta x, n\Delta t)$ が有界ならば, その数値解は安定である. しかし, 必ずしも真の解の有界性が分からないから一般には議論できない.

安定性を判定する方法はいろいろある. 例えば

- 直接法

  誤差, あるいは数値解が有界であることを直接証明する.

- エネルギー法

  数値解のノルムが有界であることを証明する.

- von Neumann 法 (フーリエ級数法)

  解をフーリエ級数に分解し, 一つ一つの成分の安定性を調べる.

などがある. これらを移流方程式とそのスキーム (7.3) に適用してみると, このスキームが安定であるための十分条件は

$$c\Delta t \leq \Delta x$$

であることがわかる.

補足　境界値問題を解くときには, 境界において常に正しい答えが与えられているから, 少々無茶をやっても大丈夫なことが多い. しかし, 初期値問題では, 正しい答えは全く分からない. そこで, 収束性や安定性が大事となってくる. 収束性や安定性があるスキームを使えば, 正しい答が得られると信じることにするからである. 安定性がないと, 解が発散してしまい, 計算できなくなってしまう. そこで実際上は安定性が必須条件となる. より正確な安定性, 収束性の定義や, 安定性の判定条件はたくさんある. これらのことや, 安

スキームの誤差は, 次のようにして定義される. 真の解 $u(j\Delta x, n\Delta t)$ を (7.3) に代入した時, 本来ならば右辺は 0 となるはずだが, 誤差のため 0 とはならない. これを ε と書き, 打ち切り誤差 (truncation error) という. すなわち,

$$\frac{u(j\Delta x,(n+1)\Delta t)-u(j\Delta x,n\Delta t)}{\Delta t}+c\frac{u(j\Delta x,n\Delta t)-u((j-1)\Delta x,n\Delta t)}{\Delta x}=\varepsilon \tag{7.4}$$

である. (7.4) の左辺を Taylor 展開すると

$$\varepsilon=\frac{1}{2}\frac{\partial^2 u}{\partial t^2}\Delta t+\frac{1}{6}\frac{\partial^3 u}{\partial t^3}(\Delta t)^2+\cdots-c\left(\frac{1}{2}\frac{\partial^2 u}{\partial x^2}\Delta x-\frac{1}{6}\frac{\partial^3 u}{\partial x^3}(\Delta x)^2+\cdots\right) \tag{7.5}$$

と書ける. 最低次の項は, Δx で 1 次, Δt で 1 次である. そこで, このスキームは Δx について 1 次, Δt について 1 次の正確度 (accuracy) を持つという. これを

$$\varepsilon=O(\Delta x),\ \varepsilon=O(\Delta t)$$

または

$$\varepsilon=O(\Delta x,\ \Delta t)$$

と書く. スキームによっては, Δx, Δt それぞれに対しての accuracy の次数が違うので, この 2 つをわけて考えることは有用である. 正確度が n 次のスキームを, 単に n 次のスキームと呼ぶ.

7.2.3　収束性 (convergence)

適合性のあるスキームであっても, 差分間隔$\Delta x,\ \Delta t\to 0$ のとき, 数値解の誤差 $\to 0$ とは必ずしもならない. ここで, 収束性 (convergence) という概念が登場する.

全時間 $n\Delta t$ を固定して, 差分間隔 $\Delta x,\ \Delta t\to 0$ とした時, 数値解の誤差 $u_j^n-u(j\Delta x,n\Delta t)\to 0$ となるならば, この数値解は収束するという. 又, あるスキームが任意の初期値に対して収束する解を与えるならば, そのスキームは収束するという. 差分スキームが適合性があるからといって, 収束するとは限らない. それは, 依存領域についても考慮しなければならないからである. 依存領域 (domain of dependence) というのは, 注目する点における関数の値を計算するために用いる格子点の集合のことである.

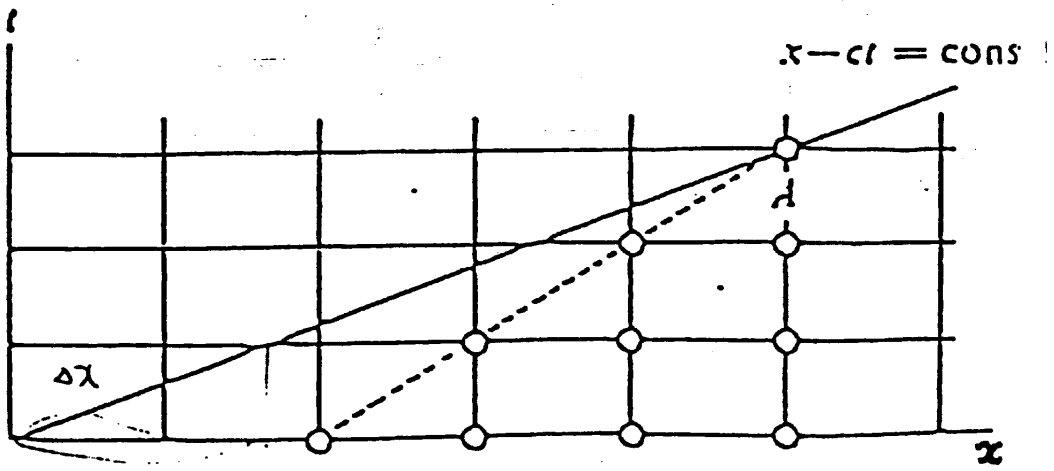


図 7.3: 特性曲線と依存領域の位置関係の例 (F.Mesinger & A.Arakawa, 1976).

例として, 前で取り上げた移流方程式 (7.1) を考える. ここで我々は, 図 7.3において, 黒丸で表わされている点 A における数値解を求めたい. 特性曲線 $x-ct=\text{const.}$ が, 斜めの実線で表わされている. 特性曲

7.2.2 有限差分スキームにおける基礎概念

Introduction 微分方程式中の導関数を適当な有限差分近似で置き換えて得られる代数方程式を, 有限差分近似式または有限差分スキームという. ここでは, 有限差分スキームを考える場合に, なくてはならない 4 つの基礎概念

- 適合性 (consistency)$\cdots$ 差分間隔を小さくすると, 近似された導関数が真の導関数に収束する性質
- 正確度 (accuracy) $\cdots$ どの程度適合性があるのかを表す度合い
- 収束性 (convergence) $\cdots$ 差分間隔を小さくすると, 数値解が真の解に収束する性質
- 安定性 (stability) $\cdots$ 差分間隔一定で計算を続けていっても (タイムステップ数を増やしていっても), 数値解が発散しない性質

について説明する.

適合性 (consistency) 導関数の近似は, 差分間隔 $\Delta x, \Delta t \to 0$ の時, 導関数に収束しなければならない. この内容は, たいてい近似は適合性 (consistency) を持たねばならない, と表現する. 適合性を持つためには, 近似の誤差が Δx , Δt について 1 次以上であることが必要かつ十分である. すなわち, 誤差を ε とすれば,

$$\varepsilon = O(\Delta x), \qquad \varepsilon = O(\Delta t)$$

であることが必要十分である. 同様に, 差分間隔が 0 に収束するとき, 差分方程式が微分方程式に収束する場合, このスキームには適合性があるという. 初期値問題においては, スキームに適合性をもたせることが特に重要である. 境界値問題ならば, 境界において真の値が与えられているため, 途中で真の値からずれても, 正しい値に戻ることが可能である. しかし, 初期値問題では, 一旦ずれたらもはや正しい値に戻ることはできない. よって, 初期値問題では, スキームの適合性の有無はきちんと調べておかねばならない.

正確度 (accuracy) 適合性のあるスキームならば, 差分間隔が 0 に収束するとき, スキームは元の式に収束する. だがそのとき, 格子間隔の何乗に比例して収束するのか ?これを表現するのが正確度 (accuracy) という概念である. 以下, 例として 1 次元線形移流方程式

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0; \quad u = u(x,t), \quad c = \text{正の定数} \tag{7.1}$$

を用いて説明する. この解析解は

$$u = f(x - ct) \tag{7.2}$$

である. 但し, f は任意の関数であり, 初期条件によって決まる.

$x - t$ 平面上の直線

$$x - ct = \text{const.}$$

上では, u は一定の値を取る. このような直線を特性曲線 (characteristics) という.

数値解を求めるため, $x - t$ 平面に格子を設定する. 点 $(j\Delta x, n\Delta t)$ における u の値 u_j^n を求めることを考える. 真の解を見れば, 時間微分と空間微分は次のように置き換えるのがいいことが分かる.

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + c\frac{u_j^n - u_{j-1}^n}{\Delta x} = 0 \tag{7.3}$$

となる. このスキームは適合性がある. すなわち, 差分間隔が 0 に近付く時, 元の微分方程式 (7.1) に近付く.

ここで, $L = 500$ (単位は任意) とし, または $F(x)$ 以下のように与えられるとする. $F(x)$ の形が $0 \leq i \leq 6$ と $7 \leq i \leq 50$ で異なるのは, 弦の一部に正弦波の一個分を作り出したことを意味する. n は差分のステップ数を表す.

$$\begin{cases} x & = & i \cdot \Delta x & (i = 0, 1, 2, \cdots, n) \\ F(i \cdot \Delta x) & = & 5 \sin(\dfrac{\pi \Delta x}{6} i) & (i = 0, 1, 2, \cdots, n) \\ F(i \cdot \Delta x) & = & 0 & (i = 7, 8, 9, \cdots, n) \end{cases}$$

x は格子点の x 座標, Δx は x 方向の刻みである. 計算にあたっては t の刻みΔt を 1 にする. また, L, v, n らは定数であって値は何でも良いのだが, 便宜的にここでは $L = 500, v = 0.1, n = 50$ とする. 言わずもがなのことだが, 微分方程式を解くとは, 任意の地点 x, 任意の時刻 t における $y(x, t)$ の値を求めるということである. 適当なグラフィックルーチンを用いて $(x, y(x, t))$ をプロットすれば, パルス波の伝搬の様子がアニメーション的に理解されることになる.

$$
\begin{aligned}
\left(\frac{\partial^2 u}{\partial x^2}\right)_p &= \left(\frac{\partial^2 u}{\partial x^2}\right)_{i,j} \\
&= \left[\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right)\right]_{i,j} \\
&\simeq \left[\frac{\partial}{\partial x}\left(\frac{u(x+\Delta x,y)-u(x,y)}{\Delta x}\right)\right] \\
&= \left[\frac{1}{\Delta x}\left(\frac{\partial u(x+\Delta x,y)}{\partial x}-\frac{\partial u(x,y)}{\partial x}\right)\right]_{i,j} \\
&\simeq \left[\frac{u(x+2\Delta x,y)-2u(x+\Delta x,y)+u(x,y)}{(\Delta x)^2}\right]_{i,j} \\
&= \frac{u_{i+2,j}-2u_{i+1,j}+u_{i,j}}{(\Delta x)^2}
\end{aligned}
$$

2. 後退差分近似式

$$
\begin{aligned}
\left(\frac{\partial u}{\partial x}\right)_p &= \left(\frac{\partial u}{\partial x}\right)_{i,j} \\
&\simeq \frac{u_{i,j}-u_{i-1,j}}{\Delta x} \\
\left(\frac{\partial^2 u}{\partial x^2}\right)_p &= \left(\frac{\partial^2 u}{\partial x^2}\right)_{i,j} \\
&\simeq \frac{u_{i,j}-2u_{i-1,j}+u_{i-2,j}}{(\Delta x)^2}
\end{aligned}
$$

3. 中心差分近似式

$$
\begin{aligned}
\left(\frac{\partial u}{\partial x}\right)_p &= \left(\frac{\partial u}{\partial x}\right)_{i,j} \\
&\simeq \frac{u_{i+1,j}-u_{i-1,j}}{2\Delta x} \\
\left(\frac{\partial^2 u}{\partial x^2}\right)_p &= \left(\frac{\partial^2 u}{\partial x^2}\right)_{i,j} \\
&\simeq \frac{u_{i+1,j}-2u_{i,j}+u_{i-1,j}}{(\Delta x)^2}
\end{aligned}
$$

例題

両端を固定した長さ L の弦の上を往復するパルス波の伝搬を解くプログラムを作ってみよう. 但し, 波動の伝搬速度を v とし, 一次元の波動方程式を以下の境界条件のもとで解くものとする.

- 境界条件 : $y(0,t)=0, \quad y(L,t)=0$

これは, 長さ L の弦の両端を固定する境界条件である.

- 初期条件 : $y(x,0)=F(x)$

と表すことができる. これを後退差分近似式と言う. 前進と後退のふたつの近似式の差を取ると

$$y(x+\Delta x)-y(x-\Delta x)=2\Delta x\cdot y'(x)+\frac{2}{3!}(\Delta x)^3\cdot y'''(x)+\cdots$$

となるが, $(\Delta x)^3$以上の高次項を無視すると

$$y'(x)\simeq\frac{y(x+\Delta x)-y(x-\Delta x)}{2\Delta x}$$

となる. これを導関数の中心差分近似式と言い, 図 7.1の線分 AB の勾配に相当している. これは前の二つの近似式に比べて精度が高い.

ふたつの独立変数 x,yの関数を $u(x,y)$ とすると, 上述の方法と全く同様にして導関数の差分近似式を導くことができる. その結果だけを示しておこう. 但し, ここでは次の省略記号を用いる. すなわち, 図 7.2に示すように, 座標平面を間隔$\Delta x,\Delta y$の格子に分割し, ある点 (x,y) の座標を

$$\begin{cases} x_i=x_0+i\cdot\Delta x & (i=0,1,2,\cdots)\\ y_i=y_0+j\cdot\Delta y & (j=0,1,2,\cdots)\end{cases}$$

と表すことにする. 但し, x_0,y_0は基準点での座標である. このようにすると関数 $u(x,y)$ のある格子点での値は次のように表すことができる. 但し, $x_0=0,y_0=0$ とすると,

$$u(x_i,y_i)=u(0+i\cdot\Delta x,0+j\cdot\Delta y)\equiv u_{i,j}$$

以上の省略記号を用いると, 差分近似式は次のようになる.

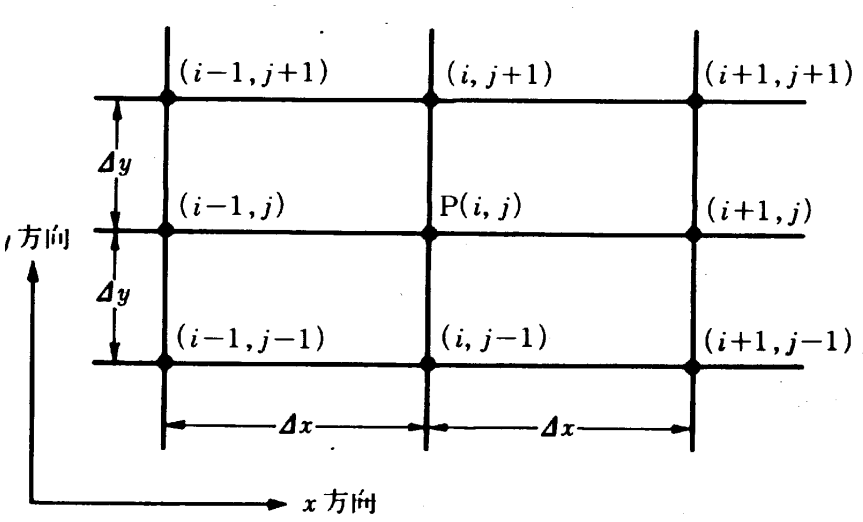


図 7.2: x,y平面上の格子上分割

1. 前進差分近似式

$$\left(\frac{\partial u}{\partial x}\right)_p=\left(\frac{\partial u}{\partial x}\right)_{i,j}\simeq\frac{u_{i+1,j}-u_{i,j}}{\Delta x}$$

7.2 有限差分法の原理

7.2.1 導関数の差分近似

関数を yを x の関数とし, $y = y(x)$ とおく. $y(x)$ が x の狭い範囲で微分可能であるとして, $y(x)$ をテーラー展開すると

$$y(x+\Delta x) = y(x) + \Delta x \cdot y'(x) + \frac{(\Delta x)^2}{2!} \cdot y''(x) + \frac{(\Delta x)^3}{3!} \cdot y'''(x) + \cdots$$

Δx を微小量と考えて展開式の第二項まで取ると,

$$y(x+\Delta x) \simeq y(x) + \Delta x \cdot y'(x)$$

となり, これより

$$y'(x) \simeq \frac{y(x+\Delta x) - y(x)}{\Delta x}$$

を得る. この式を前進差分近似式と言う. この式は図 7.1の点 P での接線の勾配を線分 PB の傾きで近似することを意味している. この場合の計算誤差はΔx 程度で, かなり大きい.

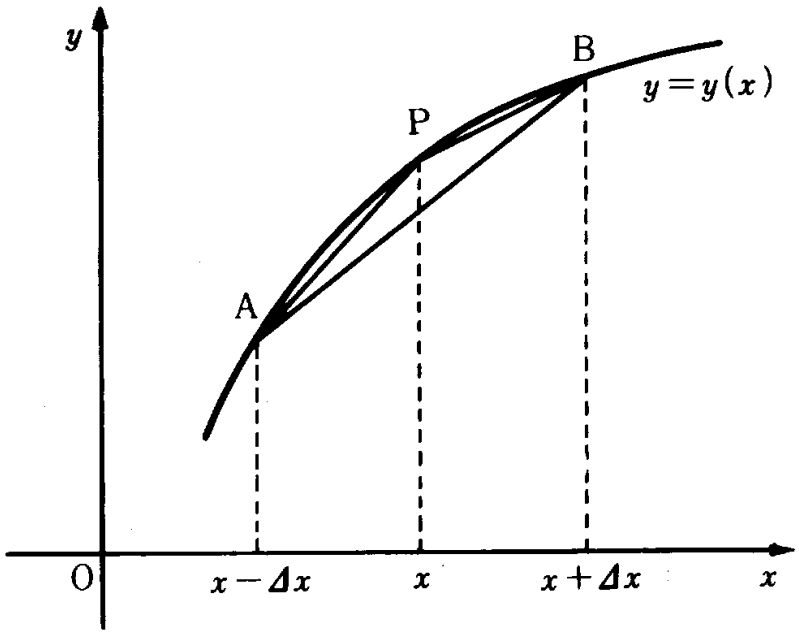


図 7.1: グラフの接線の勾配と導関数

$y(x)$ のテーラー級数展開として

$$y(x+\Delta x) = y(x) - \Delta x \cdot y'(x) + \frac{(\Delta x)^2}{2!} \cdot y''(x) - \frac{(\Delta x)^3}{3!} \cdot y'''(x) + \cdots$$

を用いると, 点 P での接線の勾配を線分 AP の傾きで近似することも可能で, この場合は

$$y'(x) \simeq \frac{y(x) - y(x-\Delta x)}{\Delta x}$$

- 数値シミュレーション

  何か適当な条件を設定して, どのようになるかやってみるもの

の様に自由なものまで, いろいろなことが行われるようになってきている.

第 7 章

時間発展問題の解法

この章では, 時間発展問題を数値的に解く方法について解説する. 時間発展問題のほとんどのものは微分方程式の解法へ帰着される. ここでは, 解説を常微分方程式と偏微分方程式の解法および解の安定性に絞り, 具体例を挙げながら解説を進める.

なお, 本章の内容は地球惑星物理学科地球流体グループに蓄積されたノートをもとにしており, その意味で指向が地球流体力学に偏っているが, そのあたりは深く考えないで読み進んで欲しい.

7.1　気象業界における数値計算の歴史と現状

7.1.1　歴史

気象業界における数値計算の歴史について簡単な年表を以下に示す.

1904 年, Wilhelm Bjerknes :　運動方程式を積分することによる天気予報の可能性を指摘した. Bjerknes が実際に数値積分を実行しようとしていたかどうかはわからない.

1922 年, Richardson :　実際に手計算で積分を行なう. 非常に長い時間を要して得た結果は間違っていたが, めげずにこれを本にする (Richardson, 1922). 当時は, 実地に応用出来るかは疑問視されていた.

1930 年頃,　Courant et al., 1928,　Rossby (1930 年代): 運動方程式の解の形を研究することにより, 簡略化された方程式, 例えば渦度方程式などを解けばいいということを示す.

1945 年, ENIAC の発明 :　Charney et al., 1950 により, ENIAC というコンピュータを用いて初めてちゃんとした数値予報がなされる. これ以降, 計算機の進歩が新しい計算方法を生み出し, 計算機と計算方法は, 相補的に進歩していく.

1970 年代に入って, 計算機, 計算方法共に進歩したため, 数値計算は実用に耐え得るものとなってきた. 具体的にいうと, 渦度方程式や発散方程式などの簡略化された方程式だけではなく, 基礎方程式を直接解くことができるようになってきた. これに伴い, 数値計算の応用範囲も広がってきている.

気象学では,

- 数値予報

  流体力学の方程式を, 現在の観測値を初期値として積分し, 未来の天気を予測する

という目的のはっきりしたものから,

$\delta(t)$ (Dirac のデルタ函数) $\Longleftarrow CTFT \Longrightarrow$ $I(f) = 1$ (恒等函数)

$I(t) = 1$ (恒等函数) $\Longleftarrow CTFT \Longrightarrow$ $\delta(f)$ (Dirac のデルタ函数)

$$\mathrm{sgn}(t)\left\{\begin{array}{ll} -1, & t<0 \\ 0, & t=0 \\ 1, & t>0 \end{array}\right\} \text{(符号函数)} = \Longleftarrow CTFT \Longrightarrow \frac{2}{j2\pi f}$$

$$\frac{1}{t} \Longleftarrow CTFT \Longrightarrow -j\pi\ \mathrm{sgn}(f)$$

$$H(t) = \left\{\begin{array}{ll} 0, & t<0 \\ 1/2, & t=0 \\ 1, & t>0 \end{array}\right\} \Longleftarrow CTFT \Longrightarrow \frac{1}{2}\left\{\delta(f) + \frac{2}{j2\pi f}\right\}$$

(Heaviside の単位階段函数)

$$e^{\pm j2\pi f_0 t} \Longleftarrow CTFT \Longrightarrow \delta(f \mp f_0)$$

$$\cos(2\pi f_0 t) \Longleftarrow CTFT \Longrightarrow \frac{1}{2}\{\delta(f-f_0) + \delta(f+f_0)\}$$

$$\sin(2\pi f_0 t) \Longleftarrow CTFT \Longrightarrow \frac{1}{2j}\{\delta(f-f_0) - \delta(f+f_0)\}$$

$$T\sum_{n=-\infty}^{\infty}\delta(t-nT) = \sum_{n=-\infty}^{\infty} e^{jn2\pi t/T} \Longleftarrow CTFT \Longrightarrow \sum_{n=-\infty}^{\infty}\delta(f-n/T)$$

定理 8 . の測度 Rを　$\{X(t)\}$のスペクトル測度[3](spectral measure) といい , $R(\lambda)$ の増加点を　$\{X(t)\}$のスペクトル (spectrum) という .

付記２　;Priestley(1981) の第 4 章 **Wiener-Khintchine** の定理の説明ではこのメモと同じ結論が導いてある . (残念である ！)

付記３　;自己相関函数のスペクトル表現のみならず, 弱定常過程のスペクトル表現も, 実は Hilbert 空間中のユニタリー作用素のスペクトル分解定理から導ける事は, 函数解析の分野では 1940 年代に確立されたというが , それについての説明は後日を期したい .

本メモに関する参考文献は以下の通りである.

Brigham, 1974 ; Papoulis, 1962 ; Lighthill, 1958 ; Blackman, 1959 ; Feller, 1966 ; 川嶋 弘尚 ・酒井, 1989 ; Marple, 1987 ; Priestley, 1981 ; 竹内, 1989 ; Doob, 1953 ; Papoulis, 1984 ;

6.9　付録 2:フーリエ変換の性質と例

$$g(t)=\int_{-\infty}^{\infty}G(f)e^{j2\pi ft}df \quad \Longleftarrow CTFT \Longrightarrow \quad G(f)=\sum_{-\infty}^{\infty}g(t)e^{-j2\pi ft}dt$$

$$g(at) \quad \Longleftarrow CTFT \Longrightarrow \quad (1/|a|)G(f/a)$$

$$(1/|a|)g(t/a) \quad \Longleftarrow CTFT \Longrightarrow \quad G(af)$$

$$\bar{g}(at) \quad \Longleftarrow CTFT \Longrightarrow \quad (1/|a|)\bar{G}(-f/a)$$

$$G(at) \quad \Longleftarrow CTFT \Longrightarrow \quad (1/|a|)g(-f/a)$$

$$\bar{G}(at) \quad \Longleftarrow CTFT \Longrightarrow \quad (1/|a|)\bar{g}(f/a)$$

$$g(t\pm t_0) \quad \Longleftarrow CTFT \Longrightarrow \quad G(f)e^{\pm j2\pi ft_0}$$

$$e^{\pm j2\pi f_0 t}g(t) \quad \Longleftarrow CTFT \Longrightarrow \quad G(f\mp f_0)$$

$$a_1g_1(t)+a_2g_2(t) \quad \Longleftarrow CTFT \Longrightarrow \quad a_1G_1(f)+a_2G_2(f)$$

$$g_1(t)*g_2(t)\ \equiv\ \sum_{-\infty}^{\infty}g_1(\tau)g_2(t-\tau)d\tau \quad \Longleftarrow CTFT \Longrightarrow \quad G_1(f)G_2(f)$$

$$g_1(t)g_2(t) \quad \Longleftarrow CTFT \Longrightarrow \quad G_1(f)*G_2(f)\ \equiv\ \int_{-\infty}^{\infty}G_1(z)g_2(f-z)dz$$

[3]通信工学ではパワースペクトル分布函数 (power spectrum distribution function) と呼ぶ .

さて Khintchine(1934) の自己相関函数 $R(t-u)$ は, アンサンブル平均 $E(\mathbf{x}_t)=0$, 分散 $E(\mathbf{x}_t^2)=1$ という弱定常実過程に対して, $E(\mathbf{x}_t\mathbf{x}_u)$ で定義された函数であり $R(0)=1$ を満足している (規格化されている). Khintchine(1934) が証明した定理は次の通りである.

"$R(t)$ が連続な実定常確率過程の自己相関函数である為の必要充分条件は

$$R(t) \;=\; \int_{-\infty}^{\infty} \cos(tx)dF(x) \qquad \text{Khintchine(1934)}\ \ (2) \tag{6.115}$$

と表現出来ることである": ここに $F(x)$ は 1 つの確率分布函数である. また連続な定常確率過程とは, その自己相関函数 $R(t)$ が原点で連続であり ($R(+0)=R(0)=1$), 従ってあらゆる点で連続であるような過程[1] である.

上の定理の証明に用いたのは次の Bochner(1932) の定理である.

"$R(0)=1$ を満足する $R(t)$ が連続で正定値であれば, $F(x)$ を 1 つの確率分布函数であるとして

$$R(t) \;=\; \int_{-\infty}^{\infty} \exp(itx)dF(x) \tag{6.116}$$

と表現出来る":

以上から分かるように, Wiener(1930) の式 (5.40)(6.15) は確定関数についてであるが, この確定関数を, エルゴード性を持つ定常確率過程の 1 つの実現函数と解釈する事も可能である. それに比べると Khintchine(1934) の式 (2) は平均連続な弱定常確率過程ならば成立するのであり, エルゴード性は仮定する必要は無い.

時系列のスペクトル解析の分野で Wiener-Khintchine の公式と呼ばれているのは, 上述の Wiener(1930)-(5.40), Khintchine(1934)-(2) という Fourier-Stieltjes 積分であるというのがこのメモの結論である. 単調な有界変動函数は微分が存在する絶対連続函数と階段函数の和で表わされるとして整理したのが別紙のまとめである. (Wiener(1930) には, 連続ではあるが, ほとんど全ての点で微分係数が 0 であるという函数 (連続特異函数) の例もあげてある. Lebesgue 積分論では **Lebesgue** の分解定理により, 連続特異函数を考慮するのが常識であるらしい. しかし通常の物理的な応用では考える必要は無いと判断して, ここでは連続特異函数を無視している .)

付記 1　;以上の結論を得た後で, 思いついて遅ればせながら岩波数学辞典 (第 3 版)(1985) や統計学辞典 (1989) を調べたら **Wiener-Khintchine** の定理という名前の定理は見つからなかった.　数学者の間では, 上述の **Bochner** の定理の方が基本として紹介されているようだ.

参考: 統計学辞典 (1988) p.1016.

定理 5. ボホナーの定理 Bochner's theorem : 実変数 t の連続関数 ϕ がある (確率) 分布の特性関数であるための必要十分条件は, ϕ が正定符号かつ $\phi(0)=1$ を満たすことである.

同上　p.1017.

定理 8.[2] : 確率過程 $\{X(t);-\infty<t<\infty\}$ において, $E\left(X(t+s)\overline{X(s)}\right)$ が s に無関係な t の連続関数 $\rho(t)$ ならば, 単調非減少で右連続な関数 $R(\lambda)$ が存在して

$$\rho(t) \;=\; \int_{-\infty}^{\infty} \exp(i\lambda t)dR(\lambda) \tag{6.117}$$

と表される. 特に $X(t)$ が実数値を取るならば, 次のようになる.

$$\rho(t) \;=\; \int_{-\infty}^{\infty} \cos(\lambda t)dR(\lambda) \tag{6.118}$$

[1] 現在では平均連続 (continuous in the mean) と呼ぶ.

[2] **Khintchine** の定理とか **Bochner-Khintchine** の定理とか呼ぶ本もある.

6.8　付録 1:Wiener-Khintchine の定理について

“平均パワー有限の函数に対して自己相関函数とパワースペクトル密度とが互いに Fourier 変換の関係にある” というのが一般に言われている Wiener-Khintchine の定理 (または公式) である．しかしパワースペクトル密度は自己相関函数の Fourier 変換であると定義している書物 (例えば, Blackman and Tukey(1959),p.6, p.85. Papoulis(1984),p.265.) を見たりすると, 一体 Wiener や Khintchine は何を証明したのかという疑問が湧いてきた．Blackman and Tukey(1959) には Wiener や Khintchine の名前は登場しない．Papoulis(1984) は “パワースペクトル密度が正 (非負) である” という定理を Wiener-Khintchine の定理と呼んでいる (p.276) が，その証明には (実現不能 ?の) 理想的フィルターを使用しているから説得力があるとは言えない．手元のスペクトル解説本も二，三 (例えば, 日野幹雄 「スペクトル解析」や大崎順彦 「地震動のスペクトル解析入門」等) 当たって見たが, Wiener-Khintchine の定理の説明がどうもはっきりしないので，仕方が無いから原論文 Wiener(1930) および Khintchine(1934) に当たって見た．以下の記号は原論文のものを使用する．(積分は両者共に Lebesgue 積分を用いている．)

まず Wiener(1930) の方には，確定函数 $f(t)$ について以下の事項の証明がある．

“自己相関函数 $\phi(x) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} f(x+t)\bar{f}(t)dt$　(式 (3.01)) が全ての x で存在すると仮定すると,

integrated periodogram　$S(u) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi(x)\frac{e^{iux}-1}{ix}dx$　(式 (3.27)) が定義出来て，自己相関函数$\phi(x)$ は次のように表現出来る．(この時 $\phi(x)$ が連続という条件が必要である .)

$$\phi(\lambda) = \int_{-\infty}^{\infty} e^{-i\lambda u}dS(u) \qquad \text{Wiener(1930)}\ \ (5.40) \tag{6.111}$$

ここに $S(u)$ は単調非減少で有界変動であり，上式の積分は Fourier-Stieltjes 積分である．
なお $S(u)$ の性質として $\phi(0) = S(+\infty) - S(-\infty)$ が成り立つ :

これで当面の課題については充分であるが，更に次式も導かれていて Wiener(1933) や Wiener(1949) にも引用されているのでここにも載せて置く．

$$s(u) = \frac{1}{2\pi}\int_{-1}^{1} f(x)\frac{e^{iux}-1}{ix}dx + \frac{1}{2\pi}\underset{\mathbf{A}\to\infty}{\text{l.i.m.}}\left[\int_{1}^{A}+\int_{-A}^{-1}\right]\frac{f(x)e^{iux}}{ix}dx \qquad \text{Wiener(1930)}\ \ (5.51) \tag{6.112}$$

を定義すると，

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|f(x)|^2dx = \lim_{\mu\to 0}\frac{1}{2\mu}\int_{-\infty}^{\infty}|s(u+\mu)-s(u-\mu)|^2du \qquad \text{Wiener(1930)}\ \ (5.53)\ \ (6.113)$$

この式についての Wiener(1930) の説明は次の通りである．

“The meaning of (5.53) is that if $f(x)$ is quadratically summable over every finite range, and $f(x)/x$ is quadratically summable over any infinite range excluding the origin, then if either side of (5.53) exists, the other side exists and assumes the same value.”

更に :

$$\phi(t) = \lim_{\epsilon\to 0}\frac{1}{2\epsilon}\int_{-\infty}^{\infty} e^{-itu}|s(u+\epsilon)-s(u-\epsilon)|^2du \qquad \text{Wiener(1930)}\ \ (6.15) \tag{6.114}$$

“This formula holds in the same sense as (5.53) for each t independently.”

(式 (6.15) で積分と極限の順序を交換して, $P(u) = \lim_{\epsilon\to 0}\frac{1}{2\epsilon}|s(u+\epsilon)-s(u-\epsilon)|^2$ と置けるなら, $\phi(t)$ と $P(u)$ は通常の Fourier 変換対をなしていて, $P(u)\geq 0$ も成立し, 話は簡単である．普通のテキストでは, この場合のみについて言及しているだけである．実はこの式の右辺の極限は, 現代風に言えばデルタ函数を含んでいる場合もあるので Wiener は式 (5.40)(5.53)(6.15) のままにして置いたのであろう .)

6.7　参考文献

データ解析および時系列解析に関するこのマニュアルの執筆には永田, 1976 ; ブラント, 1976 ; ベンダット, ピアソル, 1976 ; Press et al., 1986 ; 日野, 1977 ; を参考にした.

同一条件のもとで得られる数回の測定結果の平均値をとる.

(b) 分割平均

全データを l個の部分に分割し, その各々の区間の FFT スペクトル $P(f)$ の平均をとる.

$$\hat{P}(f) = \frac{1}{l}[P_1(f) + P_2(f) + \cdots + P_l(f)]$$

(c) スペクトルウインドー

スペクトルウインドーによる平滑化を行なう. 例えば, FFT スペクトルの移動平均をとる. あるいは, 次のような $f = k/T$を中心とする三角ウインドーも用いられる.

$$\hat{P}(f) = \frac{1}{\acute{l}^2} \sum_{j=-\acute{l}+1}^{\acute{l}-1} (\acute{l} - |j|) P\left(\frac{k-j}{T}\right)$$

6.6 演習問題

スペクトル解析の練習問題の素材としては特に地球物理の分野では個々の分野でそれぞれ多くの観測データ等があると思われるが, ここでは生データを用いることなくスペクトル解析をやってみよう.

問題

本章でも触れたように, データは本来無限長のものであると考えれば, 得られた生のデータにはもともと次の箱型 (box car) ウインドーがかかっているものと考えられる.

$$w_j = \begin{cases} 0 & (j < 0) \\ 1 & (0 \le j < N) \\ 0 & (N \le j) \end{cases}$$

ただし, FFT にかけるデータは $\{x_j : j = 0, ..., N-1\}$ の N個とする. このままスペクトルを推定すると, このウインドーのために (有限データによる誤差のために) 高周波領域で大きくばらついてスペクトルを推定しにくい. そこで, この box car ウインドーの影響を押えるために, 本章で紹介した他にもいろいろな種類のデータウインドーが提唱されている. その例を次にいくつか挙げる.

$$\begin{aligned} w_j &= 1 - \left| \frac{j - \frac{1}{2}(N-1)}{\frac{1}{2}(N+1)} \right| \\ &\equiv \text{“Parzen window”} \end{aligned}$$

$$\begin{aligned} w_j &= \frac{1}{2}\left[1 - \cos\left(\frac{2\pi j}{N-1}\right)\right] \\ &\equiv \text{“Hanning window”} \end{aligned}$$

$$\begin{aligned} w_j &= 1 - \left(\frac{j - \frac{1}{2}(N-1)}{\frac{1}{2}(N+1)}\right)^2 \\ &\equiv \text{“Welch window”} \end{aligned}$$

box car ウインドーとこれらのウインドーのパワースペクトルを FFT によって求め, 周波数領域でデータのスペクトルに対してどのような影響を与えるかを考えてみよ. FFT のプログラムは自分で組んでも良いし, 計算機のライブラリを用いても良い.

た方が良い. 例えば, データの始めと終り $1/10$ ずつに cosine 型の滑らかなウインドーを掛ける.

$$W(t) = \begin{cases} 0 & (t < 0) \\ \frac{1}{2}\{1 - \cos\frac{10\pi t}{T}\} & (0 \le t < \frac{1}{10}T) \\ 1 & (\frac{1}{10}T \le t < \frac{9}{10}T) \\ \frac{1}{2}\{1 - \cos\frac{10\pi(T-t)}{T}\} & (\frac{9}{10}T \le t \le T) \\ 0 & (T < t) \end{cases}$$

この結果歪みはほぼ消えるが, パワースペクトルの強さを

$$\mu = \left[\frac{1}{T}\int_0^T W^2(t)\mathrm{dt}\right]^{-1} = \frac{1}{0.875}$$

倍しなければならない.

(3) FFT の実行と生のスペクトル　計算する式は

$$X_k = \frac{T}{N}\sum_{j=0}^{N-1} x_j \exp\left(-i\frac{2\pi}{N}kj\right) \qquad k = \text{integer}, 0 \le k \le N/2$$

であるが, 普通 FFT のプログラムでは T/N を省略した共役な複素数

$$A_k + iB_k = \sum_{j=0}^{N-1} x_j \left[\cos\frac{2\pi jk}{N} + i\sin\frac{2\pi jk}{N}\right]$$

を計算する. 従って, $X_k = X_{\mathrm{r}k} + iX_{\mathrm{i}k}$の実数部と虚数部はそれぞれ

$$\begin{aligned} X_{\mathrm{r}k} &= \frac{T}{N}A_k \\ X_{\mathrm{i}k} &= -\frac{T}{N}B_k \end{aligned}$$

となる. ただし, $k = 0, 1, ..., N/2$ である.

パワースペクトル (パワースペクトル密度関数) $P(f)$ $(f = k/T)$ は

$$P(f) = \frac{1}{T}E[X_k {X_k}^*] = \frac{1}{T}E[|X_k|^2]$$

あるいは

$$P(f) = \frac{T}{N^2}E[{A_k}^2 + {B_k}^2]$$

もし, データウインドーを用いたなら,

$$\mu P(f) \Longrightarrow P(f)$$

のように修正する.

(4) 生のスペクトルの平滑化　こうして得られた生の FFT スペクトルは B-T 法の場合に比べて推定誤差が大きく, 特に高周波数領域で激しい振動を示す. そこで, 次の平滑化操作のいずれかを 1 つまたは 2 対上組み合わせて行ない, 分散の少ないスペクトル推定値を求める.

(a) 統計平均

とすると, 第 2 段階の演算は

$$C_{pQ+q} = \sum_{\tilde{q}=0}^{Q-1} D_q^p e^{i2\pi\left(\frac{\tilde{q}q}{Q}+\frac{\tilde{p}q}{PQ}\right)}$$

であり, もしも P, Q がさらに分解できればこの演算を繰り返すことができることを示している. 例えば, $N = P^l$ と分解できればこのような和計算を l 回繰り返せばよい. この場合の演算回数は NlP である. 前の $N = 8$ の例では, 行列の列を並べかえて第 1 段階をさらに分解し次のように計算できる.

$$\begin{bmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{bmatrix} = \begin{bmatrix} 1 & & & & 1 & & & \\ & 1 & & & & W & & \\ & & 1 & & & & i & \\ & & & 1 & & & & iW \\ 1 & & & & -1 & & & \\ & 1 & & & & -W & & \\ & & 1 & & & & -i & \\ & & & 1 & & & & -iW \end{bmatrix}$$

$$\times \left[\begin{array}{cccc|cccc} 1 & & 1 & & & & & \\ & 1 & & i & & & 0 & \\ 1 & & -1 & & & & & \\ & 1 & & -i & & & & \\ \hline & & & & 1 & & 1 & \\ & & 0 & & & 1 & & i \\ & & & & 1 & & -1 & \\ & & & & & 1 & & -i \end{array}\right] \left[\begin{array}{cccc|cccc} 1 & 1 & & & 0 & & & \\ 1 & -1 & & & & & & \\ & & 1 & 1 & & & 0 & \\ & 0 & 1 & -1 & & & & \\ \hline & & & & 1 & 1 & & \\ & & 0 & & 1 & -1 & 0 & \\ & & & & & & 1 & 1 \\ & & & & 0 & & 1 & -1 \end{array}\right] \left[\begin{array}{c} x_0 \\ x_4 \\ \hline x_2 \\ x_6 \\ \hline x_1 \\ x_5 \\ \hline x_3 \\ x_7 \end{array}\right]$$

これは結局, $N = 2^3$ と分解した例に他ならない. FFT のプログラムとして良く使われるのは, $N = 2^l$ の FFT であるが, このように任意の自然数に対して高速化できることがわかる.

スペクトル推定の具体的手順

(1) データ数の決定　前記のように FFT は任意のデータ数に対して演算可能であるが, 普通は 2^p のデータ数に対するサブルーチンが用意されていることが多い. データ数が一致しない時は, データを削るか, あるいは, 0 データを加える. 0 データを加えても, 求められるスペクトルの周波数間隔が細かくなるだけである.

(2) データウインドー　$[0, T]$ の間の有限の長さのデータのフーリエ変換を行なうことは, 無限長データ $x(t)$ に

$$W(t) = \begin{cases} 0 & (t < 0) \\ 1 & (0 \le t \le T) \\ 0 & (T < t) \end{cases}$$

の箱型 (box car) ウインドーを掛けてから変換することになる. このウインドーはスペクトル領域で大きな歪みをもつので, データをそのままフーリエ変換するのではなく, 何らかの別のデータウインドーを掛け

この行列の列を入れかえると,

$$
\begin{bmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{bmatrix}
=
\begin{bmatrix}
1 & & & & 1 & & & \\
 & 1 & & & & W & & \\
 & & 1 & & & & i & \\
 & & & 1 & & & & iW \\
1 & & & & -1 & & & \\
 & 1 & & & & -W & & \\
 & & 1 & & & & -i & \\
 & & & 1 & & & & -iW
\end{bmatrix}
\times
\left[\begin{array}{cccc|cccc}
1 & 1 & 1 & 1 & & & & \\
1 & i & -1 & -i & & & & \\
1 & -1 & 1 & -1 & & 0 & & \\
1 & -i & -1 & i & & & & \\
\hline
 & & & & 1 & 1 & 1 & 1 \\
 & & & & 1 & i & -1 & -i \\
 & 0 & & & 1 & -1 & 1 & -1 \\
 & & & & 1 & -i & -1 & i
\end{array}\right]
\begin{bmatrix} x_0 \\ x_2 \\ x_4 \\ x_6 \\ \hline x_1 \\ x_3 \\ x_5 \\ x_7 \end{bmatrix}
$$

となる. ここで, 最初に演算する行列計算は虚数単位を掛けるだけだから, 実際の計算上はデータの実部 (虚部) と虚部 (実部) の加算と減算だけであり, 元のフーリエ変換の式よりも容易な演算になる. この例では, 第 1 段階の行列演算に要する演算回数は $4 \times 4 \times 2 = 32$, 第 2 段階の行列演算に要する演算回数は $4 \times 2 \times 2 = 16$ で計 48 であるが, 元のフーリエ変換を直接計算すると, $8 \times 8 = 64$ なので, このように分解することで計算の能率も上がる.

一般的には N はどんな自然数でも構わないので, 以上の計算方法を一般化すると,

$$N = P \cdot Q \qquad\qquad P, Q = \text{integer}$$

と分解できるとし, 添え字の自然数 j, kを

$$j = \tilde{q}P + \tilde{p}\ ,\ k = pQ + q \qquad p, \tilde{p} = 0, 1, ..., P-1\ ;\ q, \tilde{q} = 0, 1, ..., Q-1$$

と書く. 離散化フーリエ変換は次のように変形される.

$$
\begin{aligned}
C_{pQ+q} &= \sum_{\tilde{q}=0}^{Q-1} \sum_{\tilde{p}=0}^{P-1} x_{\tilde{q}P+\tilde{p}} W^{(\tilde{q}P+\tilde{p})(pQ+q)} \\
&= \sum_{\tilde{q}=0}^{Q-1} e^{i2\pi\left(\frac{\tilde{q}q}{Q} + \frac{\tilde{p}q}{PQ}\right)} \sum_{\tilde{p}=0}^{P-1} x_{\tilde{q}P+\tilde{p}} e^{i2\pi\frac{\tilde{p}p}{P}}
\end{aligned}
$$

前述の $N = 8$ の例では, $P = 4$, $Q = 2$, 第 1 段階の行列計算が$\sum_{\tilde{p}=0}^{P-1}$, 第 2 段階の行列演算が$\sum_{\tilde{q}=0}^{Q-1}$ に当たる. 第 1 段階に要する演算回数が $P \cdot Q \cdot P = NP$, 第 2 段階に要する演算回数が $P \cdot Q \cdot Q = NQ$ で, 計 $N(P+Q)$ となり, Nが大きければ, 直接計算した場合の N^2よりも演算回数が少なくなる. 第 1 段階で求められる結果を

$$D_{\tilde{q}}^{p} \equiv \sum_{\tilde{p}=0}^{P-1} x_{\tilde{q}P+\tilde{p}} e^{i2\pi\frac{\tilde{p}p}{P}}$$

を離散化して, 次の有限離散化フーリエ逆変換が得られる.

$$x_j = \sum_{k=-N/2}^{N/2-1} X_k \exp(i2\pi \frac{k}{T}\frac{jT}{N})\frac{1}{T} \tag{6.109}$$

$$= \sum_{k=0}^{N-1} X_k \exp(i2\pi \frac{k}{T}\frac{jT}{N})\frac{1}{T} \tag{6.110}$$

(2) 高速フーリエ変換 (FFT) の基本的なアルゴリズム　FFT は前記の DFT を高速に求める手段である. これを如何に速く計算するかに FFT のアルゴリズムの真髄がある.

簡単のため, 用いるデータは全時間 T, データ数 Nの有限実数データ $\{x_j\}$ とする. Wを

$$W = \exp \frac{i2\pi}{N}$$

とすると, 前記の DFT の式は,

$$X_k = \frac{T}{N}\sum_{j=0}^{N-1} x_j W^{-jk}$$

となるが, さらに複素共役を考えて,

$$X_k = \frac{T}{N}\sum_{j=0}^{N-1} \left({x_j}^* W^{-jk*}\right)^*$$

$$= \frac{T}{N}\sum_{j=0}^{N-1} \left(x_j W^{jk}\right)^*$$

従って, $C_k \equiv \sum_{j=0}^{N-1} x_j W^{jk}$ を計算すれば良いことになる. 複素データの場合はこれからわかるように, その共役な複素数を考えれば良い.

$N = 8$ の場合について考察してみる. この時, $W = \exp \frac{i\pi}{4}$ である. $W^2 = i$, $W^3 = iW$, $W^4 = -1$, $W^5 = -W$, $W^6 = -i$, $W^7 = -iW$ に注意すれば,

$$\begin{bmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \\ C_7 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & W & W^2 & W^3 & W^4 & W^5 & W^6 & W^7 \\ 1 & W^2 & W^4 & W^6 & W^8 & W^{10} & W^{12} & W^{14} \\ 1 & W^3 & W^6 & W^9 & W^{12} & W^{15} & W^{18} & W^{21} \\ 1 & W^4 & W^8 & W^{12} & W^{16} & W^{20} & W^{24} & W^{28} \\ 1 & W^5 & W^{10} & W^{15} & W^{20} & W^{25} & W^{30} & W^{35} \\ 1 & W^6 & W^{12} & W^{18} & W^{24} & W^{30} & W^{36} & W^{42} \\ 1 & W^7 & W^{14} & W^{21} & W^{28} & W^{35} & W^{42} & W^{49} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & W & i & iW & -1 & -W & -i & -iW \\ 1 & i & -1 & -i & 1 & i & -1 & -i \\ 1 & iW & -i & W & -1 & -iW & i & -W \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -W & i & -iW & -1 & W & -i & iW \\ 1 & -i & -1 & i & 1 & -i & -1 & i \\ 1 & -iW & -i & -W & -1 & iW & i & W \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix}$$

6.5.2　直接法 (記録の直接フーリエ変換による方法)

Blackman-Tukey 法による計算時間は主として自己相関関数の計算の部分に費やされ, 技術の進歩により良質のデータが大量に得ることができるようになると, より計算時間の短いスペクトル計算法が求められるようになった. その様な頃, 1965 年に Cooley and Tukey によってフーリエ変換を効率よく計算する方法 (FFT) が発表され, 大きな反響を巻きおこした. これは自己相関関数を求めることなくパワースペクトルを推定する方法であり, この方法が現在では一般的である. ここでは, まず FFT の原理を述べ, その後スペクトル推定の具体的手順について述べる.

FFT (高速フーリエ変換) の原理

(1) 有限離散化フーリエ変換 (DFT)　何らかの手段で得られるデータから計算機でスペクトルを計算しようとすると有限の長さで且つ離散化されたデータを扱わなければならない. ここではこのような有限離散化されたデータのフーリエ変換 (有限離散化フーリエ変換 :Discrete Fourier Transform [DFT]) について簡単に述べる.

連続時系列データ $x(t)$ のフーリエ変換 $x(t) \rightarrow X(f)$ を複素形式で表して

$$X(f) = \int_{-\infty}^{\infty} x(t) \exp(-i2\pi ft) dt$$

とする. ここに, N 個の有限複素離散化データ $\{x_j\}$ $(j = 0, 1, ..., N-1)$ が与えられた時, この DFT が X_k $(k = 0, 1, ..., N-1)$ と求められるとすると, データの記録時間を T として

(連続)		(離散)
dt	$\Longrightarrow$	$\frac{T}{N}$
t	$\Longrightarrow$	$\frac{jT}{N}$
f	$\Longrightarrow$	$\frac{k}{T}$

のように対応するから, DFT は (直観的に) 次のように書ける.

$$\begin{aligned} X_k &= \sum_{j=0}^{N-1} x_j \exp(-i2\pi \frac{k}{T}\frac{jT}{N})\frac{T}{N} \\ &= \frac{T}{N}\sum_{j=0}^{N-1} x_j \exp(-i\frac{2\pi}{N}kj) \qquad\qquad k = -\frac{N}{2}, \cdots, \frac{N}{2}-1 \end{aligned} \tag{6.107}$$

ここで, 離散化データのサンプリング周波数は $\frac{N}{T}$ であるが, エイリアシングを起こさない最大の周波数 f_N (ナイキスト [Nyquist] 周波数) は, その半分の $\frac{N}{2T}$ である. したがって, 意味のある DFT は, $|f| \leq f_N$ でしか得られない. つまり, $\{X_k : -N/2 \leq k \leq N/2\}$ が意味のある DFT の結果である. ただし, 式 (6.107) からわかる通り X_k は周期 N を持つため, $X_{-N/2}$ と $X_{N/2}$ は独立ではない (互いに相等しい). また, k の範囲を $-N/2 \leq k \leq N/2-1$ の代わりに $0 \leq k \leq N-1$ にとり x_j の j の範囲と一致させることもある. ここで, $X_{-k} = X_{N-k}$ である.

また, 全く同様に, 連続時系列データのフーリエ逆変換

$$x(t) = \int_{-\infty}^{\infty} X(f) \exp(-i2\pi ft) df \tag{6.108}$$

が得られたら,

$$m = \frac{1}{N}\sum_i \eta_i' \tag{6.98}$$

を計算し,

$$\eta_i = \eta_i' - m \tag{6.99}$$

として平均値 0 の時系列 $\{\eta_i\}$ に置き換える.

(4)　プリホワイトニング　解析しようとする現象に強い周期性があるとか, 対象とする周波数範囲のエネルギーレベルの変化が大きい場合には, プリホワイトニングの操作を施して, あらかじめスペクトルを平坦化するようにする.

(5)　自己相関関数の計算　自己相関関数は

$$C(l) = C(-l) = \frac{1}{N-1}\sum_{i=1}^{N-1} \eta_i \cdot \eta_{i+l} \tag{6.100}$$

において, $l = 0$ から m まで計算する. m の取り方は周波数分解能Δfとスペクトルの推定値の信頼度のかねあいで決まる. 即ち,

$$\Delta f = \frac{1}{2m\Delta t} \tag{6.101}$$

の様に, m が大きいほど周波数分解能は高くなるが, 推定値の信頼度が落ちる.

(6) ラグウインドウによる平滑化及びフーリエ変換　ラグウインドウ (例として Hanning をとる) をかけてからフーリエ変換すると

$$S(f)\Delta f = \frac{1}{2m}\sum_{l=0}^{m} \delta_l \left\{ C(l)\left(1+\cos\frac{\pi l}{m}\right)\cos\frac{\pi r l}{m} \right\} \tag{6.102}$$

ただし,

$$\delta_l = \begin{cases} 1 & (l \neq 0, m) \\ \frac{1}{2} & (l = 0, m) \end{cases} \tag{6.103}$$

$$\Delta f = \frac{1}{2m\Delta t} \tag{6.104}$$

この式は $l \geq 0$ の範囲の総和を取っているので, $2D_2(\tau)$ に対応する $(1+\cos\frac{\pi l}{m})$ を掛けてある. また,

$$f = r\Delta f \tag{6.105}$$

であり, rについては 0 から m まで計算する. 即ち, スペクトルは周波数

$$f = r\Delta f \qquad\qquad (r = 0, 1, \ldots, m) \tag{6.106}$$

の各点で求まる.

(7) リカラリング (recolouring)　プリホワイトニングを行った場合, その効果を修正する必要がある.

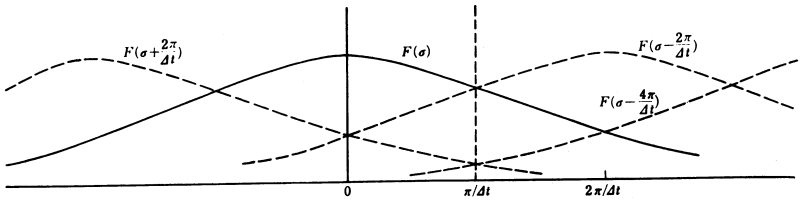


図 6.4: 真のスペクトルの値が $|f| > 1/2\Delta t$ にも存在する場合のエイリアジング. ただし, 横軸は$\sigma = 2\pi f$.

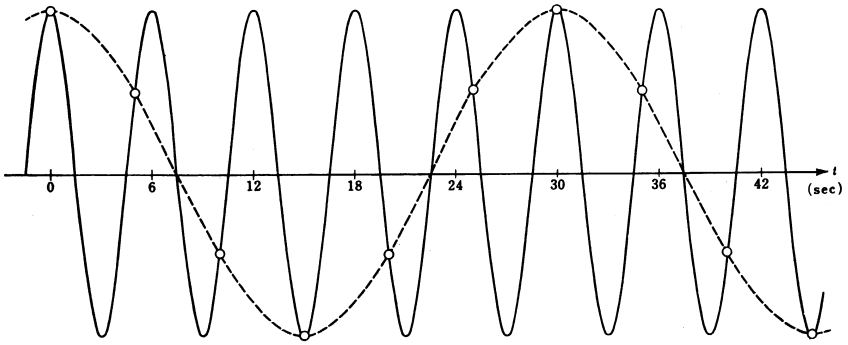


図 6.5: エイリアジングの原理

も短い周期の半分以下にするか, あるいは, 予め記録に適当な低域通過フィルター (low-pass filter) をかけておいて Nyquist 周波数よりも高周波数成分を取り除いておく等の操作により, 誤差を除去する事が可能である.

6.5　パワースペクトルを求める実際の計算手順

6.5.1　Blackman-Tukey 法

解析すべきデータは等間隔に取られた有限の記録とする.

(1)　トレンドの除去など　解析する前に生の記録$\eta'(t)$ をながめて, トレンドやドリフト (記録の長さより長い低周波成分のこと) が無いことを確かめる. もし, 存在するときは平均勾配を最小自乗法などで推定することによりさっ引いておく. またデータを確率過程とみなすのが適当であるかどうかなども検討しておくべきである.

(2)　読み取り間隔, ラグの決定　読み取り間隔は, 記録に現れている最も短い変動の周期の半分以下に定めるべきである. 即ち, Nyquist 周波数 $f_N = \frac{1}{2\Delta t}$ が記録に含まれる最大周波数 f_{max} より大きくなるようにΔt を定める. 計算能力等の理由によりΔt を大きくとるときはあらかじめ記録に低域通過フィルターをかけて f_Nより高周波成分を除いておくべきである.

(3)　座標変換　等間隔Δt で N個のデータ,

$$\eta'_1, \eta'_2, \ldots, \eta'_N \tag{6.97}$$

を用いて

$$S_a(f) = A\left(f;\frac{1}{\Delta t}\right) * S(f) \tag{6.93}$$

$$= \sum_{n=-\infty}^{\infty} S\left(f-\frac{1}{\Delta t}n\right) \tag{6.94}$$

$$= S(f)+S\left(f-\frac{1}{\Delta t}\right)+S\left(f-2\frac{1}{\Delta t}\right)+\cdots+S\left(f+\frac{1}{\Delta t}\right)+S\left(f+2\frac{1}{\Delta t}\right)+\cdots \tag{6.95}$$

となる.

いま, 真のスペクトルが $|f| < 1/2\Delta t$ でのみ値を持つとすると, $S_a(f)$ は $S(f)$ が $1/\Delta t$ 間隔で繰り返して現われるだけであるから, $S_a(f)$ の $|f| > 1/2\Delta t$ の部分をゼロにしてしまえば, そのまま真のスペクトル $S(f)$ が得られることになる. この様子を図 (6.3) に示す.

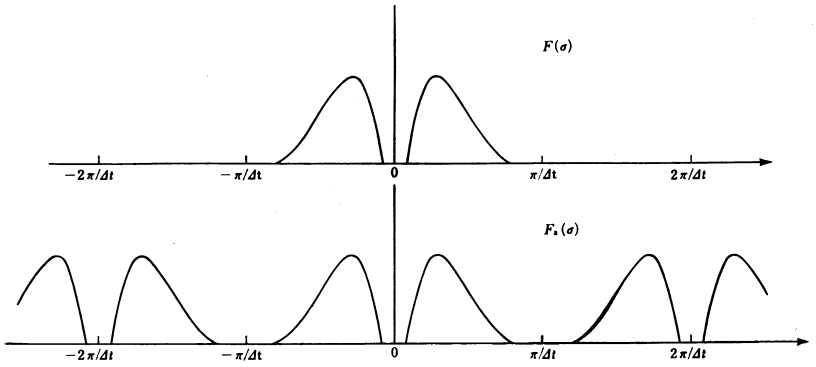


図 6.3: 真のスペクトルとエイリアジングを受けたスペクトル (真のスペクトルの値が $|f| < 1/2\Delta t$ に限られている場合) . ただし, 横軸は$\sigma = 2\pi f$.

一方, 真のスペクトルが $|f| > 1/2\Delta t$ でも有限の値を持つとすると, 推定されるスペクトル $S_a(f)$ は $|f| < 1/2\Delta t$ の範囲内でも, 重なりあいがおこってくる. いま, スペクトルは偶関数なので $0 < f < 1/2\Delta t$ の範囲のみ考えるとすると $f > 0$ の部分のスペクトルを

$$f = 1/2\Delta t, 2/2\Delta t, \ldots \tag{6.96}$$

のところで折り目をつけて畳み込んだ形になっている. それが $1/\Delta t$ の間隔で繰り返し現われることになっている. 従って, 真のスペクトル $S_a(f)$ の推定値としての $S(f)$ は $0 < f < 1/2\Delta t$ の範囲のみを考えればよい. ただし, この場合 $f > 1/2\Delta t$ に本来あったエネルギーが $0 < f < 1/2\Delta t$ の範囲に折り畳まれるように加算されている. このような, 高周波数成分が低周波数成分に重なってくる現象をエイリアジング (aliasing) と呼ぶ. また, 周波数 $f_N = 1/2\Delta t$ を折り畳み周波数 (folding frequency) またはナイキスト周波数 (Nyquist frequency) と呼ぶ. この折り畳みの様子を図 (6.4) に示す.

エイリアジングの例を図 (6.5) に示す. これは, 周期 6sec の波を 5sec 間隔で読みとった場合である. このように, 周期が読み取り間隔の 2 倍以下の波はより長周期の波 (ここでは周期 30sec) と区別することが出来ない.

記録の不連続性からくる誤差の推定

上記の通り, 記録の不連続性は Nyquist 周波数より高周波数成分の重ね合わせに効いてくるのみであり, スペクトル推定値のランダムな変動には寄与しない. このため, Δt を十分小さくとって, 記録に含まれる最

従うことが解っている. 即ち, スペクトルウインドウの等価幅 $W_e = 1/T_m$ のなかに周波数の最小スケール $1/T_n$ がいくら入っているかを数えればよい. すると

$$k = 2\frac{1/T_m}{1/T_n} = \frac{2T_n}{T_m} \tag{6.85}$$

ファクター 2 は一つの波は振幅と位相の二つの自由度を持つことからくると解釈できる (もちろん, ある種の近似のもとで上記の結果は数学的に導出可能である.) .

誤差の分散成分が既知の分布に従うと分かってしまえば, その評価は容易である. いま, χ^2-分布は, その分散と平均値との間に

$$\text{分散} = \frac{2(\text{平均値})^2}{k} \tag{6.86}$$

の関係があるため, 正規化された標準偏差 ϵ は

$$\epsilon = \frac{\sqrt{\text{分散}}}{\text{平均値}} = \sqrt{\frac{2}{k}} \tag{6.87}$$

となる. また, スペクトルピークの信頼限界もχ^2-分布の表から求められる.

即ち, T_n に較べて T_m が小さい程, パワースペクトル推定値の分散は減少し安定性は増大することが分かった. しかし, あまり T_m を小さくとると, 周波数分解能が粗くなってしまうためスペクトルの形が歪んだり, ピークが埋もれてしまったりするので, 実際には T_m の大きさは, 安定性と分解能とのトレードオフで決まる. 通常の目的には, T_m は T_n の $5 \sim 10\%$ くらいが適当とされている.

6.4.2　記録の不連続性の効果

実際の記録は, 厳密な意味で連続ではなく, とびとびの値しか得ることはできない. ここでは記録が等間隔Δt で与えられている場合について考え, 記録の不連続性が推定されるパワースペクトルに与える影響について議論する.

いま,

$$t = 0, \Delta t, 2\Delta t, \ldots, n\Delta t, \ldots \tag{6.88}$$

の各時刻における値のみ知られているとすると, 自己相関関数も

$$\tau = 0, \pm \quad \Delta t, \pm \quad 2\Delta t, \ldots, \pm \quad n\Delta t, \ldots \tag{6.89}$$

に対する値のみ計算可能である. このような, 離散的な自己相関関数をフーリエ変換することによって得られたパワースペクトルが, 真のパワースペクトルからどの様にずれてくるかを考える.

とびとびの値を持つ離散的な自己相関関数は, ディラックのデルタ関数を用いて

$$C_a(\tau) = \nabla(\tau; \Delta t) \cdot C(\tau) \tag{6.90}$$

ここで, $C(\tau)$ は真の自己相関関数であり, また

$$\nabla(\tau; \Delta t) = \Delta t \sum_{n=-\infty}^{\infty} \delta(\tau - n\Delta t) \tag{6.91}$$

である.

式 (6.90) は式 (6.71) と比べて, $D_0(\tau)$ が $\nabla(\tau; \Delta t)$ に変わっただけであるから, $C_a(\tau)$ のフーリエ変換 $S_a(f)$ は, $\nabla(\tau; \Delta t)$ のフーリエ変換

$$A\left(f; \frac{1}{\Delta t}\right) = \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{1}{\Delta t}n\right) \tag{6.92}$$

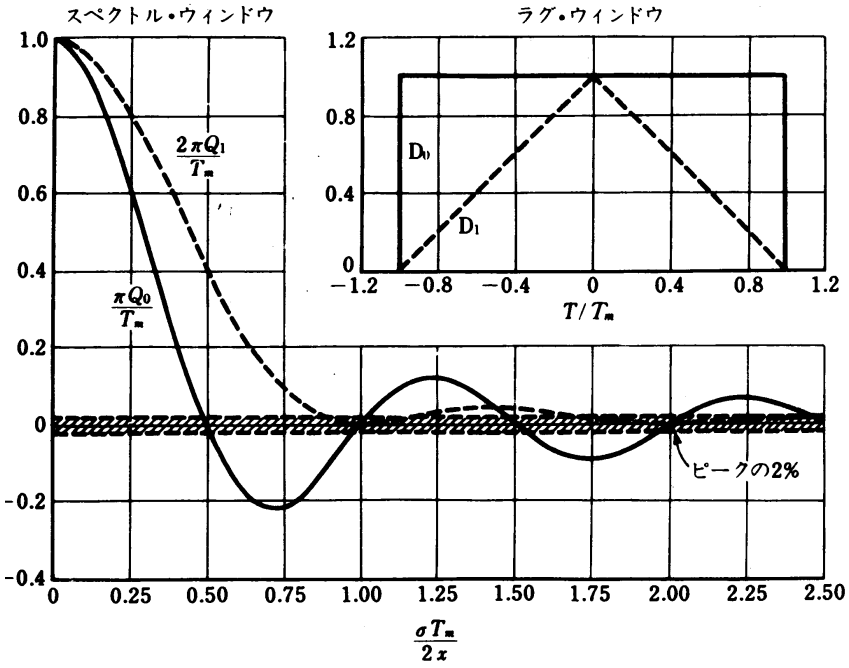


図 6.1: ラグウインドウとスペクトルウインドウの例 (1). ただし, $\sigma = 2\pi f$ である.

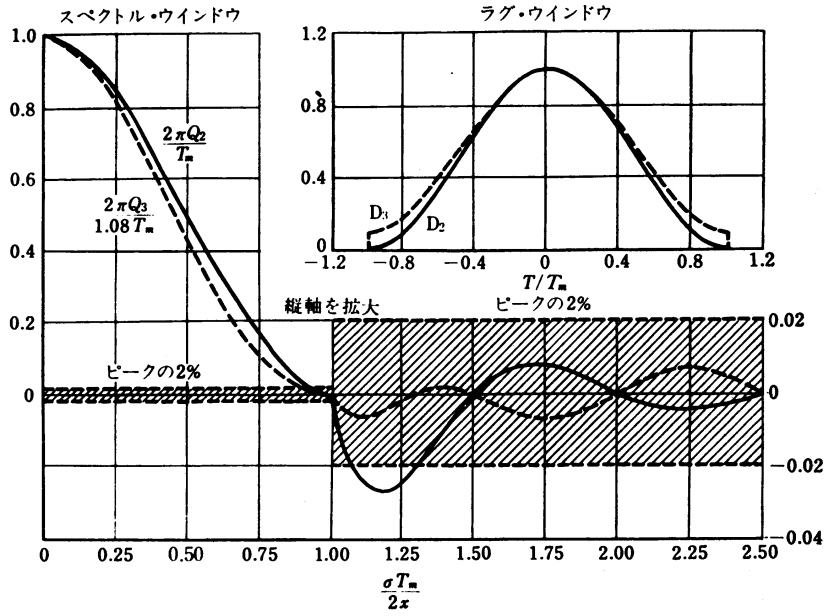


図 6.2: ラグウインドウとスペクトルウインドウの例 (2). ただし, $\sigma = 2\pi f$ である.

最小スケール$\Delta f = 1/T_n$, 及びウインドウによって平滑化が行なわれる等価周波数幅, 即ち分解できる最小の周波数幅 $W_e \cong 1/T_m$ に対応している.

まず, 誤差のうちのバイアス成分だが, これは真のスペクトルの形が最小の周波数分解能 $1/T_m$ の間で大きく変化しないくらい T_m を大きくとれば, ゼロになる.

一方誤差の分散成分は, スペクトル推定に用いた独立なデータの数を等価自由度 k とする χ^2-分布にほぼ

であるから, フーリエ変換の合成積定理より

$$E[S_0(f)] = Q_0(f) * S(f) \tag{6.77}$$
$$\equiv \int_{-\infty}^{\infty} Q_0(f-\xi)S(\xi)d\xi \tag{6.78}$$

即ち, 我々に計算可能な量 $S_0(f)$ は, 真のスペクトル $S(f)$ と既知の関数 $Q_0(f)$ の convolution の一つの推定値である. $Q_0(f)$ の具体的な形は図 (6.1) に示す.

ラグウインドウとスペクトルウインドウ

$D_0(\tau)$ の様なものをラグウインドウと呼ぶ一方, $Q_0(f)$ の様なものはスペクトルウインドウと呼ばれる. convolution の計算はスペクトルウインドウ $Q_0(f)$ を重み関数とした移動平均である. そのため, スペクトルウインドウとしては, なるべく $f=0$ の近くのピークに比べて, 両側の副次的な振動が小さいほうが望ましい.

ラグウインドウに対する条件としては, $|\tau| \geq T_m$ に対しては 0 であり, また全パワー $C(0)$ を変化させないために $\tau = 0$ で 1 となることのみである. これらを満たしさえすれば $D_0(\tau)$ 以外にもいろいろ考えられる. 例えば, $|\tau| \geq T_m$ に対して

$$D_1(\tau) = 1 - \frac{|\tau|}{T_m} \tag{6.79}$$
$$D_2(\tau) = \frac{1}{2}\left(1 + \cos\frac{\pi\tau}{T_m}\right) \tag{6.80}$$
$$D_3(\tau) = 0.54 + 0.46\cos\frac{\pi\tau}{T_m} \tag{6.81}$$

等が考えられる. これらに対するスペクトルウインドウは, それぞれ

$$Q_1(f) = T_m\left(\frac{\sin\pi f T_m}{\pi f T_m}\right)^2 \tag{6.82}$$
$$Q_2(f) = \frac{1}{2}Q_0(f) + \frac{1}{4}\left\{Q_0(f)\left(2\pi f + \frac{\pi}{T_m}\right) + Q_0(f)\left(2\pi f - \frac{\pi}{T_m}\right)\right\} \tag{6.83}$$
$$Q_3(f) = 0.54Q_0(f) + 0.23\left\{Q_0(f)\left(2\pi f + \frac{\pi}{T_m}\right) + Q_0(f)\left(2\pi f - \frac{\pi}{T_m}\right)\right\} \tag{6.84}$$

となる. これらの実際の形は図 (6.1), 図 (6.2) に示す. 図からも $Q_2(f)$, $Q_3(f)$ の特性は $Q_0(f)$ よりも格段によいことが分かる. $Q_3(f)$ の副次的な山の高さは $Q_2(f)$ に比べて $1/3$ であるが, $|f|$ が大きくなった時の減衰の速さは $Q_2(f)$ のほうが優れており, 両者の優劣は簡単には論じられない. この両者は 実際にもよく用いられ, $Q_2(f)$ を用いることを **hanning**, $Q_3(f)$ を用いることを **hamming** と呼ぶ. また $Q_1(f)$ の特徴は, 正値のみをとるということであるが, 実際に用いられることは少ない.

記録の有限性からくる誤差の推定

有限時間 T_n の長さの記録から, 有限の長さのラグ $|\tau| \leq T_m$ に対して定義された自己相関関数 $C(\tau)$ を計算し, そのフーリエ変換から得られたパワースペクトル $S_0(f)$ が持つ誤差について考える. 一般的な議論には複雑な計算が必要だが, ここでは結果及びその直観的な説明に話を限り, 詳細は他書にゆずることにする.

推定値の誤差には二通りある. 即ち, 系統的なずれの成分と, ランダムな変動成分とである. 普通, 前者をバイアス, 後者を分散と呼ぶ. 誤差の原因としては, 全記録長 T_n が有限であること, 及び, 用いた自己相関関数の最大ラグ T_m が有限であることの二つである. それぞれ周波数空間に焼き直すと, 周波数の原理的な

Burg による **MEM**（最大エントロピー法）などがある．今では，実際には FFT や MEM が一般には用いられているが，ここではまず，古い Blackman-Tukey 法について述べる．

現実の生データは無限の長さを持つことは有り得ず，必ず有限の長さであり，また厳密に連続的なデータというものも有り得ない．これらのことから生じる推定値の誤差に関する議論も行う．

6.4.1　記録の有限性の効果

エルゴード性を持つ確率過程であれば，無限に長い記録から，確定した値を持つ自己相関関数とパワースペクトルを求め得る．しかし，有限の長さの記録から推定したパワースペクトルは，無限の長さの記録から求めたものとは一般には一致しない．無限の長さの記録からは，無限個の有限長さの記録が作れる．従って，ある有限長さの記録から推定されたパワースペクトルは一種の確率変数と考えるのが自然である．言い換えると，同じ長さの記録の集合である母集団を考え，測られた記録をその母集団から任意に選び出された一つの標本という考えにたつわけである．このように考えることによって，推定されたスペクトルの信頼性を統計的に論じることが可能となってくる．このように，誤差理論に対する理解がスペクトル解析の結果の解釈等に関しても重要である．

記録の長さが有限である場合，自己相関関数を求めるのに，時間に関して $(-\infty, \infty)$ の範囲で積分を実行することは不可能である．そこで，見かけの（apparent）自己相関関数

$$C_{00}(\tau) = \frac{1}{T_n - |\tau|} \int_{-(T_n-|\tau|))/2}^{(T_n-|\tau|)/2} \eta\left(t - \frac{\tau}{2}\right) \eta\left(t + \frac{\tau}{2}\right) dt = C_{00}(-\tau) \tag{6.69}$$

の形を用いる．ここで，T_n は記録の長さである．当然 $|\tau|$ が T_n を越えると，計算が不可能になる．$|\tau|$ の最大値 T_m があまり T_n に近いと大きな $|\tau|$ に対する $C_{00}(\tau)$ は用いた資料があまりに少ないため信頼度が非常に悪くなる．従って，実際の計算では T_m は T_n の 5〜10% 程度が適当とされている．この際，$|\tau| \geq T_m$ に対する $C_{00}(\tau)$ の値の推定値を計算することはできない．ところが，一般のランダム変動に対して $C(\tau) \to 0 \quad (|\tau| \to \infty)$ であることが知られているので，

$$D_0(\tau) = \begin{cases} 1 & (|\tau| \leq T_m) \\ 0 & (|\tau| > T_m) \end{cases} \tag{6.70}$$

で定義されるラグウインドウ $D_0(\tau)$ を $C_{00}(\tau)$ にかけて，修正した見かけの（modified apparent）自己相関関数

$$C_0(\tau) = D_0(\tau) C_{00}(\tau) \tag{6.71}$$

を得る．当然

$$C_0(\tau) = 0 \text{for} |\tau| > T_m \tag{6.72}$$

である．$C_0(\tau)$ のフーリエ変換を $S_0(f)$ とすると，

$$S_0(f) = \int_{-\infty}^{\infty} C_0(\tau) e^{-i2\pi f\tau} d\tau \tag{6.73}$$

である．さらに，$D_0(\tau)$ のフーリエ変換を $Q_0(f)$ とすると，

$$Q_0(f) = \int_{-\infty}^{\infty} D_0(\tau) e^{-i2\pi f\tau} d\tau \tag{6.74}$$

$$= \frac{T_m}{\pi} \frac{\sin(2\pi f T_m)}{f T_m} \tag{6.75}$$

である．式（6.71）の期待値を取ると

$$E[C_0(\tau)] = D_0(\tau) C(\tau) \qquad \text{for all } \tau \tag{6.76}$$

(自己相関関数)　(パワースペクトル密度) $= \sum_{n=-\infty}^{\infty} \Phi[nf_0]\delta(f-nf_0)$

$$\text{平均パワー } R(0) = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2}|g(t)|^2dt = \int_{-\infty}^{\infty}S(f)df \tag{6.61}$$

(ii) 非周期函数 - 離散スペクトル,　概周期函数　(直流成分も含む)

(i) の場合現れる振動数は基本振動数の整数倍であったが , ここでは一般に二つの振動数の比は無理数である .

$$g(t) = \sum_{n=-\infty}^{\infty} G[f_n]e^{j2\pi f_n t} \Leftarrow CTFT \Rightarrow G(f) = \sum_{n=-\infty}^{\infty} G[f_n]\delta(f-f_n) \tag{6.62}$$

$$R(\tau) = \sum_{n=-\infty}^{\infty} |G[f_n]|^2 e^{j2\pi f_n \tau} \Leftarrow CTFT \Rightarrow S(f) = \sum_{n=-\infty}^{\infty} |G(f_n)|^2\delta(f-f_n) \tag{6.63}$$

$$\text{平均パワー } R(0) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|g(t)|^2dt = \int_{-\infty}^{\infty}S(f)df \tag{6.64}$$

(iii) 非周期函数 - 連続スペクトル,　持続函数 (上記の (i) や (ii) は取り除いてあるとする)

$$R(\tau) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\bar{g}(t)g(t+\tau)dt \Longleftarrow CTFT \Longrightarrow S(f) = \int_{-\infty}^{\infty}R(\tau)e^{-j2\pi f\tau}d\tau \tag{6.65}$$

(自己相関函数)　(パワースペクトル密度)

$$\text{平均パワー } R(0) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|g(t)|^2dt = \int_{-\infty}^{\infty}S(f)df \tag{6.66}$$

一方で

$$S(f) = \lim_{T\to\infty}\frac{|G_T(f)|^2}{2T} \tag{6.67}$$

が示される . ここに

$$g_T(t) = \left\{\begin{array}{ll} g(t), & |t|<T \\ 0, & |t|>T \end{array}\right\} \Longleftarrow CTFT \Longrightarrow G_T(f) \tag{6.68}$$

注意 . $\lim_{T\to\infty} G_T(f)$ 自身は存在しない .

平均パワーが有限の函数に対して自己相関函数 $R(\tau)$ とパワースペクトル密度 $S(f)$ とがフーリエ変換対をなすことを示す (23),(26),(28) 式は Wiener(-Khintchine) の定理と呼ばれている .

6.4　パワースペクトルの推定とその誤差

前節では, 自己相関関数, パワースペクトル密度, エネルギースペクトル密度などの概念, およびそれらの具体的表式について述べてきた.

ここでは, これらの理論に基づき具体的に不規則現象を解析する手順および推定値の信頼度について述べる. パワースペクトルの推定法には, FFT を用いた直接法と, いったん自己相関関数を求めてそのフーリエ変換によりパワースペクトルを求める方法 (**Blackman-Tukey 法**), さらに全く異なる原理に基づく,

6.3.2　平均パワー有限の関数

平均パワー $\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|g(t)|^2dt = \text{finite} > 0$. この時全エネルギー $\int_{-\infty}^{\infty}|g(t)|^2dt$ は当然無限大となる．

(i) 周期関数 (直流成分も含む)

有限な周期を T_0 とすると $g(t+T_0)=g(t)$ である．

平均パワーは1 周期 T_0 についての平均パワーに等しい．

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|g(t)|^2dt \;=\; \frac{1}{T_0}\int_{-T_0/2}^{T_0/2}|g(t)|^2dt \tag{6.51}$$

$g(t)$ をフーリエ級数展開すると

$$G[nf_0] \;=\; \frac{1}{T_0}\int_{-T_0/2}^{T_0/2}g(t)e^{-jn2\pi f_0 t}dt \;(n=0,\pm1,\pm2,\ldots) \tag{6.52}$$

$$g(t) \;=\; \sum_{n=-\infty}^{\infty}G[nf_0]e^{jn2\pi f_0 t} \;(-\infty<t<\infty) \tag{6.53}$$

ただし $f_0=1/T_0$ である. この時 $g(t) \Longleftarrow CTFS \Longrightarrow G[nf_0]$ と表現する．

$g(t)$ が実数値函数の場合は $G[-nf_0]=\bar{G}[nf_0]$ である．

相互相関定理　$g_1(t)\Longleftarrow CTFS\Longrightarrow G_1(nf_0)$,　$g_2(t)\Longleftarrow CTFS\Longrightarrow G_2(nf_0)$ とすると

$$\phi_{12}(\tau) \;\equiv\; \frac{1}{T_0}\int_{-T_0/2}^{T_0/2}\bar{g}_1(t)g_2(t+\tau)dt \quad\Longleftarrow CTFS\Longrightarrow\quad \Phi_{12}[nf_0]=\bar{G}_1[nf_0]\;G_2[nf_0] \tag{6.54}$$

(相互相関函数)　　(相互パワースペクトル)

$\tau=0$ として次の **Parseval** の定理を得る．

$$\frac{1}{T_0}\int_{-T_0/2}^{T_0/2}\bar{g}_1(t)g_2(t)dt \;=\; \sum_{n=-\infty}^{\infty}\bar{G}_1[nf_0]G_2[nf_0] \tag{6.55}$$

自己相関定理

$$R(\tau) \;\equiv\; \frac{1}{T_0}\int_{-T_0/2}^{T_0/2}\bar{g}(t)\;g(t+\tau)dt \;\Longleftarrow CTFS\Longrightarrow\; \Phi[nf_0]=|G[nf_0]|^2\geq 0 \tag{6.56}$$

(自己相関函数)　　(パワースペクトル)

$\tau=0$ として次のパワーの定理を得る．

$$\text{平均パワー}\;\; R(0)=\frac{1}{T_0}\int_{-T_0/2}^{T_0/2}|g(t)|^2dt \;=\; \sum_{n=-\infty}^{\infty}\Phi[nf_0] \tag{6.57}$$

合成積定理

$$g_1(t)*g_2(t) \;\equiv\; \frac{1}{T_0}\int_{-T_0/2}^{T_0/2}g_1(\tau)g_2(t-\tau)d\tau \;\Longleftarrow CTFS\Longrightarrow\; G_1[nf_0]G_2[nf_0] \tag{6.58}$$

以上が通常のフーリエ級数理論である．これらの結果は超函数のフーリエ変換理論を用いて，次のようにも表わせる．

$$g(t) \;=\; \sum_{n=-\infty}^{\infty}G[nf_0]e^{jn2\pi f_0 t} \;\Longleftarrow CTFT\Longrightarrow\; G(f)=\sum_{n=-\infty}^{\infty}G[nf_0]\delta(f-nf_0) \tag{6.59}$$

$$R(\tau) \;=\; \sum_{n=-\infty}^{\infty}|G[nf_0]|^2e^{jn2\pi f_0\tau} \;\Longleftarrow CTFT\Longrightarrow\; S(f)=\sum_{n=-\infty}^{\infty}|G[nf_0]|^2\delta(f-nf_0) \tag{6.60}$$

6.3.1　過渡関数の場合

$g(t) \to 0 (t \to \pm\infty)$ を満たす過渡関数の重要な性質は全エネルギー$\int_{-\infty}^{\infty} |g(t)|^2 dt$ が有限正定値であることである. 従ってこの時, 平均パワー $\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} |g(t)|^2 dt$ は零になる.

ここで, $g(t)$ とフーリエ変換対となる関数を $G(f)$ を書くと

$$G(f) = \int_{-\infty}^{\infty} g(t) e^{-j2\pi ft} dt \quad (-\infty < f < \infty) \quad \text{フーリエ変換 (CTFT)} \tag{6.41}$$

$$g(t) = \int_{-\infty}^{\infty} G(f) e^{j2\pi ft} df \quad (-\infty < t < \infty) \quad \text{フーリエ逆変換 (ICTFT)} \tag{6.42}$$

である. この時, $g(t) \Longleftarrow CTFT \Longrightarrow G(f)$ と表現することにする.

このとき成り立つ定理をいくつか述べる.

$g(t) \Longleftarrow CTFT \Longrightarrow G(f)$ であれば , $g(-t) \Longleftarrow CTFT \Longrightarrow G(-f)$, $\bar{g}(\pm t) \Longleftarrow CTFT \Longrightarrow \bar{G}(\mp f)$, $G(\pm t) \Longleftarrow CTFT \Longrightarrow g(\mp f)$, $\bar{G}(\pm t) \Longleftarrow CTFT \Longrightarrow \bar{g}(\pm f)$　(以上複号同順) 等が成り立つ .

相互相関定理　$g_1(t) \Longleftarrow CTFT \Longrightarrow G_1(f)$, $g_2(t) \Longleftarrow CTFT \Longrightarrow G_2(f)$ とすると

$$\phi_{12}(\tau) \equiv \int_{-\infty}^{\infty} \bar{g}_1(t) g_2(t+\tau) dt \Longleftarrow CTFT \Longrightarrow \Phi_{12}(f) = \bar{G}_1(f) G_2(f) \tag{6.43}$$

(相互相関函数)　(相互エネルギースペクトル密度)

$\tau = 0$ の場合を考えれば , 次の **Parseval** の定理を得る .

$$\int_{-\infty}^{\infty} \bar{g}_1(t) g_2(t) dt = \int_{-\infty}^{\infty} \bar{G}_1(f) G_2(f) df \tag{6.44}$$

自己相関定理

$$\phi(\tau) = \int_{-\infty}^{\infty} \bar{g}(t) g(t+\tau) dt \Longleftarrow CTFT \Longrightarrow \Phi(f) = |G(f)|^2 \geq 0 \tag{6.45}$$

(自己相関函数)　(エネルギースペクトル密度)

一般に $\phi(0) > |\phi(\tau)|$ $(\tau \neq 0)$ が言える .

$\tau = 0$ の時を考えると , 次のエネルギーの定理を得る .

$$\text{全エネルギー } \phi(0) = \int_{-\infty}^{\infty} |g(t)|^2 dt = \int_{-\infty}^{\infty} \Phi(f) df \tag{6.46}$$

なお $g(t)$ が実数値函数の場合は ϕ も Φ も実偶函数となりフーリエ変換は次の形になる .

$$\phi(\tau) = \int_{-\infty}^{\infty} \Phi(f) \cos(2\pi f\tau) df \tag{6.47}$$

$$\Phi(f) = \int_{-\infty}^{\infty} \phi(\tau) \cos(2\pi f\tau) d\tau \tag{6.48}$$

合成積定理

$$g_1(t) * g_2(t) \equiv \int_{-\infty}^{\infty} g_1(\tau) g_2(t-\tau) d\tau \Longleftarrow CTFT \Longrightarrow G_1(f) G_2(f) \tag{6.49}$$

$$g_1(t) g_2(t) \Longleftarrow CTFT \Longrightarrow G_1(f) * G_2(f) \equiv \int_{-\infty}^{\infty} G_1(z) G_2(f-z) dz \tag{6.50}$$

$\tau = 0$ とすると

$$C(0) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \{\eta(t)_T\}^2 dt = \int_{-\infty}^{\infty} \left[\lim_{T\to\infty} \frac{X_T(f) X_T^*(f)}{2T} \right] df = \int_{-\infty}^{\infty} S(f) df \tag{6.35}$$

ただし,

$$S(f) = \lim_{T\to\infty} \frac{X(f)_T X_T^*(f)}{2T} \tag{6.36}$$

式 (6.35) で, 左辺は時系列の標本の平均パワーを表す. また, $S(f)$ は平均パワーへの各周波数成分の寄与率を表し, パワースペクトル密度 (**power spectral density**) と呼ばれる. 式 (6.36) より, パワースペクトル密度は生の時系列の標本$\eta_T(t)$ のフーリエ変換 $X_T(f)$ (振幅スペクトル) を用いて表せることが分かる. さらに式 (6.34) より,

$$C(\tau) = \int_{-\infty}^{\infty} S(f) e^{i2\pi f\tau} df \tag{6.37}$$

$$S(f) = \int_{-\infty}^{\infty} C(\tau) e^{-i2\pi f\tau} d\tau \tag{6.38}$$

即ち, $C(\tau)$ と $S(f)$ は互いにフーリエ変換の関係にある. この関係は **Wiener-Khinchine** の公式と呼ばれる. Wiener-Khinchine の定理 (公式) に関するより詳しい議論は章末の付録 1 に載せる.

ここでは $S(f)$ は $(-\infty, \infty)$ の範囲で定義されている. いま $S(f)$ も $C(f)$ 同様偶関数である. 本来負の周波数というのは考えにくいが周波数領域を $(0, \infty)$ とした時の周波数 f成分からの寄与 $G(f)$ を

$$G(f) = S(f) + S(-f) \tag{6.39}$$

$$= 2S(f) \tag{6.40}$$

と考えればよい. ここで $S(f)$ は two-sided spectrum, $G(f)$ は one-sided spectrum と呼ばれる.

この様にして, もし$-\infty < t < \infty$ に対して一つの時系列$\eta(t)$ が観測されたならば, 式 (6.29) を用いて自己相関関数 $C(\tau)$ を計算し, さらに式 (6.38) からパワースペクトル $S(f)$ が計算される (**Blackman-Tukey** 法). ここで, $\eta(t)$ が確率過程であっても, $C(\tau)$ や $S(f)$ は決定論的な関数であり, これらより時系列データの統計的性質を議論することができるはずである.

また, 1985 年に Cooley and Tukey によって FFT のアルゴリズムが発表されて以来, 生のデータ$\eta(t)$ を直接フーリエ変換して, パワースペクトルを求める方法が広く用いられるようになっている (直接法).

ところが, どちらにせよ実際上$-\infty < t < \infty$ の観測は不可能であるため, 次に問題になるのは有限の時間内の観測資料から, どの様にして真の $S(f)$ を推定するか, また推定値の信頼度はどの程度であるかということである. これについては 6.4節で議論する.

その前に, 様々なデータ (確定データおよびランダム過程の標本データ) に対するエネルギースペクトルまたはパワースペクトルの具体的表式について次に議論する.

6.3　スペクトル密度関数の具体的表式

時間に関して連続な確定関数 $g(t), -\infty < t < \infty$ が与えられた時そのエネルギースペクトルまたはパワースペクトルの具体的表式を以下に列挙する. $g(t)$ は本当の確定関数であっても良いし, 確率過程の標本関数 (実現関数) であっても良い.

1. が成立する時は，平均パワーは 0 であり，
2. が成立する時は，全エネルギーは無限大である．

次に $g(t)$ が確定関数または確率過程の実現関数（標本関数）であれば，（一般）フーリエ解析により

1. の場合は，$E = \int_{-\infty}^{\infty} \Phi(f)df$ と表わせて$\Phi(f)(\geq 0)$ をエネルギースペクトル密度と呼ぶ．
2. の場合は，$P = \int_{-\infty}^{\infty} S(f)df$ と表わせて $S(f)(\geq 0)$ をパワースペクトル密度と呼ぶ．

但し，2 の場合の $S(f)$ としては，通常の函数以外に超函数のデルタ函数の和$\sum_n \sigma_n \delta(f - f_n)$ で表わされる時もある．これが嫌な時には $P = \sum_n \sigma_n(f_n)$ と表わして$\sigma_n(f_n)$ をパワースペクトルと呼ぶ（のが普通？）．(1. の場合には$\Phi(f)$ は通常の函数であり，デルタ函数を含む場合は無い．)

6.2.4　Wiener-Khintchine の公式

様々な場合のエネルギースペクトル密度やパワースペクトル密度の具体的表式を示す前に, 平均パワーが有限である関数について成り立つ自己相関関数とパワースペクトル密度との間の関係式（Wiener-Khintchine の公式）を示す.

時系列の標本を$\eta(t), -\infty < t < \infty$ とする. また, $\eta_T(t)$ を

$$\eta_T(t) \equiv \begin{cases} \eta_(t) & \text{for} \quad |t| < T \\ 0 & \text{for} \quad |t| > T \end{cases} \tag{6.26}$$

で定義する. $\eta_T(t)$ のフーリエ変換を $X_T(f)$ とすると,

$$\eta_T(t) = \int_{-\infty}^{\infty} X_T(f) e^{i2\pi ft} df \tag{6.27}$$

逆に,

$$X_T(f) = \int_{-\infty}^{\infty} \eta_T(t) e^{-i2\pi ft} dt \tag{6.28}$$

一方, 自己相関関数 $C(\tau)$ は

$$\begin{aligned} C(\tau) &= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \eta_T(t)\eta_T(t+\tau)dt & (6.29) \\ &= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \eta_T(t) \left[\int_{-\infty}^{\infty} X_T(f) e^{i2\pi f(t+\tau)} df \right] dt \\ &= \lim_{T\to\infty} \frac{1}{2T} \int_{-\infty}^{\infty} X_T(f) e^{i2\pi f\tau} \left[\int_{-T}^{T} \eta_T(t) e^{i2\pi ft} dt \right] df & (6.30) \\ & & (6.31) \end{aligned}$$

ここで, $\eta_T(t)$　が $(-T, T)$ 以外で 0 であるから, 積分範囲を $(-\infty, \infty)$ に置き換えられる. いま,

$$\begin{aligned} \int_{-\infty}^{\infty} \eta_T(t) e^{i2\pi ft} dt &= X_T(-f) & (6.32) \\ &= X_T^*(f) & (6.33) \end{aligned}$$

であるから

$$C(\tau) = \int_{-\infty}^{\infty} \left[\lim_{T\to\infty} \frac{X(f)_T X_T^*(f)}{2T} \right] e^{i2\pi f\tau} df \tag{6.34}$$

で完全に表し得る. ここで, 過程が (弱) 定常であると仮定して共分散の表式を変形してみる. (弱) 定常過程においては, $\overline{\eta(t_i)}$ は定数となるから, 簡単のため$\overline{\eta(t_i)} \equiv 0$ となる様に座標を選ぶことにする. そのとき,

$$c_{ij} =< {}^k\eta(t_i)^k\eta(t_j) >_k \tag{6.21}$$

さらに, (弱) 定常過程では t_iの選び方にはよらないので, c_{ij}は時間の差$\tau = t_j - t_i$のみの関数となる. c_{ij}をτの関数として表したもの, 即ち

$$C(\tau) =< {}^k\eta(t_i)^k\eta(t_j + \tau) >_k \tag{6.22}$$

を自己相関関数 (autocorrelation function) という. あるいは, エルゴード性を仮定すると

$$C(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \eta(t)\eta(t+\tau)dt \tag{6.23}$$

と定義できる. つまり, (弱) 定常ガウス過程は平均値と自己相関関数の二つで完全に記述されることになる.

重要なことは, $\eta(t)$ が確率過程であっても, 過程が (弱) 定常過程であれば $C(\tau)$ は決定論的な関数である, ということである.

6.2.2　自己相関関数の性質

過程が (弱) 定常であれば

$$\begin{aligned} C(\tau) &= < {}^k\eta(t)^k\eta(t+\tau) >_k \\ &= < {}^k\eta(t-\tau)^k\eta(t) >_k \\ &= C(-\tau) \end{aligned} \tag{6.24}$$

であり, $C(\tau)$ は偶関数である.

また, ランダムな過程では

$$C(\tau) \to 0 \qquad (\tau \to \infty) \tag{6.25}$$

であることが知られている. これは, ある程度時間がたつと過去のことは忘れてしまうことに対応している. また, 両極端の場合として, 周期関数の自己相関関数はまた周期関数, 白色雑音の自己相関関数はデルタ関数であることが知られている.

6.2.3　スペクトル密度関数

まず $g(t)$ $(-\infty < t < \infty)$ という連続時間関数対して, 全エネルギーまたは平均パワーを以下のように定義する.

1. $\displaystyle\int_{-\infty}^{\infty} g(t)^2 dt$ が有限な正値 Eをとる場合，E を $g(t)$ の全エネルギーという．
2. $\displaystyle\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} g(t)^2 dt$ が有限な正値 P を取る場合，P を $g(t)$ の平均パワーという．

上の定義は 1Ω の抵抗を流れる電流が $g(t)$ A である時に，消費される全エネルギーまたは平均パワーを考えれば納得できるものである．

次には $g(t)$ が如何なる物理量を表現する函数であっても，上の定義の仮定に合致する場合それをエネルギーとか平均パワーと呼ぶようになった．(この場合には，エネルギーや平均パワーの物理的な次元を持たない事が多い．) また, 一般に次のことがいえる.

エルゴード性

自然現象を対象とする限り, 同一条件のもとでの観測を, 数多く繰り返して行うことはまず不可能であり, 普通得られる確率過程の標本は一つきりである. そこで, 集合平均のかわりに時間平均,

$$< F({}^k\eta(t)) >_t = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} F({}^k\eta(t))dt \tag{6.15}$$

を考える. 一般には, この時間平均は先に述べた集合平均とは一致しないが, これらが一致するような確率過程をエルゴード性を持つという. この仮定は, 十分長い時間を考えれば, 一つの記録の値は事実上その取り得る値の全てを経過すると予想するものである. 実際の時系列解析においては常にエルゴード性は仮定される. エルゴード性の仮定は, 実際には実行できない集合平均を時間平均で置き換えることを可能とするものであり, 時系列解析上の統計操作は必ず時間またはそのフーリエ変換の変数である周波数に対しておこなわれることになる.

ガウス過程

ガウス過程の形式的な定義は, ある確率過程${}^k\eta(t)$ について, 任意の n 個の時刻, $t_1, t_2, \ldots, t_n$における値, ${}^k\eta(t_1), {}^k\eta(t_2), \ldots, {}^k\eta(t_n)$ の同時確率密度関数が, n 次のガウス分布, 即ち,

$$p_r({}^k\eta(t_1), {}^k\eta(t_2), \ldots, {}^k\eta(t_n)) =$$
$$\frac{1}{(2\pi)^{n/2}|C|^{1/2}} \exp\left\{ -\frac{1}{2|C|} \sum_{i,j=1}^{n} |c_{ij}| \{ {}^k\eta(t_i) - \overline{\eta(t_i)} \} \{ {}^k\eta(t_j) - \overline{\eta(t_j)} \} \right\} \tag{6.16}$$

の関係を有する場合をいう. ここで,

$$\overline{\eta(t_i)} = < {}^k\eta(t_i) >_k \tag{6.17}$$
$$c_{ij} = < \{ {}^k\eta(t_i) - \overline{\eta(t_i)} \} \{ {}^k\eta(t_j) - \overline{\eta(t_j)} \} >_k \tag{6.18}$$

であり, また $|C|$ は c_{ij}によって作られる行列の行列式, $|c_{ij}|$ はこの行列式の c_{ij}の余因数である.

式 (6.16) からわかる通り, ガウス過程においては, その全ての統計的性質は平均値と共分散によって記述される. パワースペクトルの推定値の信頼限界を議論する際にも, 通常ガウス性が仮定される. 過程のガウス性は, 一般の物理現象に於いても良く成り立つと考えられ, また数学的にも中心極限定理により正当化される.

最後に, 不規則データの解析においてはここで述べた定常性 ・エルゴード性 ・ガウス性は通常, 全て仮定される.

6.2　自己相関関数およびスペクトル密度関数

6.2.1　自己相関関数

前に述べたようにガウス過程においては, その統計的性質は平均値

$$\overline{\eta(t_i)} = < {}^k\eta(t_i) >_k \tag{6.19}$$

および共分散

$$c_{ij} = < \{ {}^k\eta(t_i) - \overline{\eta(t_i)} \} \{ {}^k\eta(t_j) - \overline{\eta(t_j)} \} >_k \tag{6.20}$$

過渡的データの重要な性質は, 周波数スペクトルが離散的ではなく連続的であるということである.

6.1.2　不規則データについて

不規則データの分類は, その数学的モデルである確率過程の持つ種々の性質に基づいてなされる. そこで, ここでは確率過程の定義から始めて, 確率過程に関する諸々の概念について述べていく.

確率過程とは

過程 (process) とは, 時間的に変化していくもの一般を指すものである. 確率過程またはランダム過程 (stochastic process or random process) とは, ある時刻におけるその過程の値をはっきり決めることができないような不確定の現象ではあるが, その値の現れ方がある確率法則に従うような過程である. 例えば, 風波を例にとってみると, 風速や吹送距離, 吹送時間といった巨視的な条件が全く同一であっても, 個々の峰が観測点に到達する時刻やその高さをいちいち予測するのはまず不可能である. しかし, ある種の平均的な波高と風の強さとの間には, かなりはっきりした関係があることが知られている. このように不確定の現象であっても, 統計量については決定論的な議論が出来ると考えられる場合がよくある.

集合平均 (ensemble mean)

数学的な確率過程のモデルを考える際には, 統計量の取り方をはっきり定義づけておく必要がある. いま, $\{{}^k\eta(t)\}$ を同じ巨視的条件のもとに取られた波の記録のような, 同一の母集団に属する確率過程の集合であると考える. ここで, 左肩の数は標本の番号を表し, 個々の標本はそれぞれ時間 t の関数である. 確率論的にいって, 本来の統計量はこの集合についての平均であるべきである. 即ち, $\eta(t)$ の任意の関数　$F(\eta(t)$ の集合平均が以下のように任意の時間 $t = t_1$ に関して得られる.

$$< F({}^k\eta(t_1)) >_k = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} F({}^k\eta(t_1)) \tag{6.11}$$

この平均操作によって, 任意の時間における平均値や分散が得られる.

定常確率過程

集合平均をとった全ての統計量が時間的に不変な場合この過程を定常確率過程, または, 特に強定常確率過程と呼ぶ. 即ち,

$$< F({}^k\eta(t_1)) >_k =< F({}^k\eta(t_2)) >_k \tag{6.12}$$

な場合である.

また, 特に平均値と自己相関関数のみが時間によらない場合, 即ち

$$< {}^k\eta(t_1) >_k \quad = \quad < {}^k\eta(t_2) >_k \tag{6.13}$$

$$< {}^k\eta(t_1)^k\eta(\tau + t_1) >_k \quad = \quad < {}^k\eta(t_2)^k\eta(\tau + t_2) >_k \tag{6.14}$$

のとき, この過程を弱定常確率過程と呼ぶ. 自然現象は厳密な意味においては, 全て一度きりの現象であり, 定常性の仮定はある種の理想化である. しかし, 定常性の仮定をはずすと, モデルとしての数学的簡潔さが著しく損なわれてしまう. 実際には, 観測時間を限って定常性を仮定することになる.

正弦波的周期データ

正弦波的データは, 次のような式で定義される時間変化関数である.

$$x(t) = X\sin(2\pi f_0 t + \theta) \tag{6.1}$$

当然, 振動数スペクトルでみると一つの特定の振動数 f_0 成分のみからなる.

複合周期データ

複合周期データとは,

$$x(t) = x(t + nT_p) \qquad n = 1, 2, 3, \cdots \tag{6.2}$$

を満たすようなデータ, 即ち一般的な周期的データである. もちろん, 正弦波的データも複合周期データの一つである. フーリエ級数定理より, 複合周期データは次のように展開できる.

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos 2\pi n f_1 t + b_n \sin 2\pi n f_1 t) \tag{6.3}$$

ここで

$$f_1 = \frac{1}{T_p} \tag{6.4}$$

$$a_n = \frac{2}{T_p}\int_0^{T_p} x(t)\cos 2\pi n f_1 t dt \qquad n = 0, 1, 2, \cdots \tag{6.5}$$

$$b_n = \frac{2}{T_p}\int_0^{T_p} x(t)\sin 2\pi n f_1 t dt \qquad n = 1, 2, 3, \cdots \tag{6.6}$$

つまり, 複合周期データは基本周波数 T_p およびその整数倍の高調波の重ね合わせで表せる. また振動数スペクトルでみると, 離散的でありかつ各振動数成分は等間隔にならんでいる.

概周期データ

複合周期データのところで述べたことから分かるように, 二つの正弦波の和はそれらの周波数の比が有理数である時, また周期的でありそれらを高調波とするような基本周波数が存在する. ここで概周期データとは, 非周期的であるデータのうちでも, 正弦波の重ね合わせで表されるが, それらの成分波の周波数の比が無理数となる場合をいう. 数学的に書き表すと (確定的データだから当然あらわせる)

$$x(t) = \sum_{n=1}^{\infty} X_n \sin(2\pi f_n t + \theta_n) \tag{6.7}$$

ただし, f_m/f_n がすべて有理数ではない. つまり, 振動数スペクトルでみると, 離散的ではあるが振動数成分の比が無理数である場合になる.

過渡的非周期データ

過渡的データとは, 概周期データを除く全ての非周期データといって良い. たとえば

$$x(t) = \begin{cases} Ae^{-at} & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{6.8}$$

$$x(t) = \begin{cases} Ae^{-at}\cos bt & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{6.9}$$

$$x(t) = \begin{cases} A & 0 \leq t \leq c \\ 0 & t < 0, c < t \end{cases} \tag{6.10}$$

第 6 章

時系列解析の基礎

時系列解析とは時間の関数として与えられた観測値が時間のどのような関数であるかの予想が全くつかない場合に, どのように時間に関して依存しているかを統計的変動をなるべく押えながらさぐるということである. 前章でやった時間依存性が線形などの簡単な場合と異なり, かなり数学的にもやっかいな部分もあるが, ここでは数学 ・統計学には深入りせずに基本的な概念および具体的手続きの初歩についてのみ述べることにする.

時間的 (または空間的) に与えられた観測値の系列を取り扱う場合に, 決定論的な立場と確率論的な立場がある. 決定論的な立場とは, 外力がはっきりしており, 現象が, 数学的定式化可能な因果関係によって記述されるような場合に対応する取り扱いである. この場合, 何時何分にどんな値を取るかが予測可能であると考えられる. 即ち, 振幅だけでなく, 位相の情報も同時に取り扱われる. 一方, 確率論的立場においては, 現象はランダムであり何時何分にどんな値を取るかは原理的に予測不可能であり, 統計的な平均値のみ意味を持つと考える. つまり, 現象をある確率過程のひとつの標本と考え, その確率過程としての平均的性質のみに注目する.

もちろん, ほとんど全ての地球物理現象において, 完全な因果関係が知られている場合も, また完全にランダムな現象と見なせる場合もあり得ないので, この二つの立場はそれぞれ現象を異なった方向に理想化したものと考えられる. いずれにせよ, 心に留めておくべきことは時系列解析の前段階として現象のモデル化が必要不可欠であるということである.

6.1 物理的データの基礎的記述

一般に, 物理的データは確定的データと不規則データに大別できる.

確定的データとは, 明確な数学的関係式で記述できるもの, 例えばバネにつるしたおもりの運動などである. このような場合は, 予測することも可能であり, 実際, 実験によって誤差の範囲内で観測値が予想値と一致することも確かめられる. このようなデータの解析は, 個々の現象に応じて適当なモデル化をすることも多いが, (形式的には) スペクトル解析の対象ともなり得るのでまず 6.1.1小節で整理しておく.

不規則データとは, 十分な精度で予測することが不可能なデータ, 例えば荒れた海の波の高さなどである. この不規則データこそ, スペクトル解析の主たる対象である. 6.1.2小節では, 不規則データの数学的モデルとしての確率過程に関する種々の概念について述べる.

6.1.1 確定的データについて

確定的な現象を表すデータは, 周期的および非周期的なものに分類できる. 非周期的データは, さらに概周期的なものと過渡的なものとに分類できる. ここでは, これらについて簡単に述べる.

測定量 y_j は 変数 t_j と線形の関係があるとして

$$\eta_j = x_1 + x_2 t_j \tag{5.82}$$

の係数 x_1, x_2 を表 (5.3) の 4 個の測定データから決めることを考えよ. ここで $\eta_j \equiv y_j - \epsilon_j$ は真の (誤差がない時の) 測定量である. また推定値 $\tilde{x_1}, \tilde{x_2}$ の誤差についても論ぜよ.

(注) 手で解けます ！

Table 5.1: 練習問題のデータ

j	1	2	3	4
t_j	0.0	1.0	2.0	3.0
y_j	1.4	1.5	3.7	4.1
σ_j	0.5	0.5	0.5	0.5

この式は $r \leq N$ の時 x について解くことができる. その解, $\tilde{x}$ は

$$\tilde{\boldsymbol{x}} = -(A^{\mathrm{T}} C_y^{-1} A)^{-1} A^{\mathrm{T}} C_y^{-1} (\boldsymbol{y} + \boldsymbol{a}_0) \tag{5.68}$$

となる.

次に測定誤差が最適推定値 $\tilde{x}$におよぼす影響について考える. データが互いに独立な場合の誤差の伝播は式 (5.45) ですでにあたえられているが, ここでまず, もう少し一般的な場合について行列の形式で考えてみる. すなわち, 関数 $\boldsymbol{y} = (y_1, y_2, \ldots, y_r)$ が, N 個の変数 $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$ に関して線形である場合に, $\boldsymbol{y}$ の共分散行列を$\boldsymbol{x}$ の共分散行列で表すことを考える. $\boldsymbol{y}$ と $\boldsymbol{x}$ の関係を

$$\boldsymbol{y} = T\boldsymbol{x} + \boldsymbol{a} \tag{5.69}$$

とおく. T, $\boldsymbol{a}$ はそれぞれ定行列および定ベクトルである. この時

$$\begin{aligned}
C_y &= E[(\boldsymbol{y} - \bar{\boldsymbol{y}})(\boldsymbol{y} - \bar{\boldsymbol{y}})^{\mathrm{T}}] && (5.70)\\
&= E[(T\boldsymbol{x} + \boldsymbol{a} - T\bar{\boldsymbol{x}} - \boldsymbol{a})(T\boldsymbol{x} + \boldsymbol{a} - T\bar{\boldsymbol{x}} - \boldsymbol{a})^{\mathrm{T}}] && (5.71)\\
&= E[T(\boldsymbol{x} - \bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}} T^{\mathrm{T}}] && (5.72)\\
&= TE[(\boldsymbol{x} - \bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}}] T^{\mathrm{T}} && (5.73)\\
&= TC_x T^{\mathrm{T}} && (5.74)
\end{aligned}$$

となる. ここで,

$$\begin{aligned}
\bar{\boldsymbol{x}} &= E[\boldsymbol{x}] && (5.75)\\
\bar{\boldsymbol{y}} &= E[\boldsymbol{y}] && (5.76)
\end{aligned}$$

である. この結果を用いると, 推定値 $\tilde{x}$ の共分散行列 $C_{\tilde{x}}$ は, 式 (5.68) より

$$\begin{aligned}
C_{\tilde{x}} &= [(A^{\mathrm{T}} C_y^{-1} A)^{-1} A^{\mathrm{T}} C_y^{-1}] C_y [(A^{\mathrm{T}} C_y^{-1} A)^{-1} A^{\mathrm{T}} C_y^{-1}]^{\mathrm{T}} && (5.77)\\
&= (A^{\mathrm{T}} C_y^{-1} A)^{-1} A^{\mathrm{T}} C_y^{-1} C_y C_y^{-1} A (A^{\mathrm{T}} C_y^{-1} A)^{-1} && (5.78)\\
&= (A^{\mathrm{T}} C_y^{-1} A)^{-1} (A^{\mathrm{T}} C_y^{-1} A)(A^{\mathrm{T}} C_y^{-1} A)^{-1} && (5.79)\\
&= (A^{\mathrm{T}} C_y^{-1} A)^{-1} && (5.80)
\end{aligned}$$

となる. ただしここで, 行列 $C_y, C_y^{-1}, (A^{\mathrm{T}} C_y^{-1} A)$ は対称行列だから, 転置行列と等しいこと, および公式

$$(ABC)^{\mathrm{T}} = C^{\mathrm{T}} B^{\mathrm{T}} A^{\mathrm{T}} \tag{5.81}$$

を使った.

5.3 練習問題

形式的な議論のみだと分かりにくいかも知れないので, 具体的な線形回帰および誤差伝播に関する問題を解いて, 理解を確実なものにしよう.

問題

いま, 変数 t_j の関数である測定量 y_j を考える. ただし, t_j の値は十分精度よく決められ, 誤差は測定値 y_j のみを考える. 測定値誤差の標準偏差 σ_j も知られているとする.

さらに

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_N \end{pmatrix} \quad \boldsymbol{a}_0 = \begin{pmatrix} a_{10} \\ a_{20} \\ \vdots \\ a_{N0} \end{pmatrix} \quad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & & & \\ a_{N1} & a_{N2} & \cdots & a_{Nr} \end{pmatrix} \tag{5.54}$$

と定義すると, 式 (5.53) は

$$\boldsymbol{\eta} + \boldsymbol{a}_0 + A\boldsymbol{x} = 0 \tag{5.55}$$

と書ける. また, 実際の測定値を y_i としその誤差 ϵ_i は正規分布をするとすると

$$y_i = \eta_i + \epsilon_i \tag{5.56}$$

$$E[\epsilon_i] = 0 \tag{5.57}$$

$$E[\epsilon^2] = \sigma_j^2 \tag{5.58}$$

となる. N 個の測定値 y_iはたがいに独立であるから, 共分散行列 C_y は対角行列になる.

$$C_y = \begin{pmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_N^2 \end{pmatrix} \tag{5.59}$$

先ほどと同様に 観測値と誤差を表す N元ベクトル $\boldsymbol{y}$, $\boldsymbol{\epsilon}$ を定義すると

$$\boldsymbol{y} = \boldsymbol{\eta} + \boldsymbol{\epsilon} \tag{5.60}$$

となる. これを式 (5.55) に代入すると

$$\boldsymbol{y} - \boldsymbol{\epsilon} + \boldsymbol{a}_0 + A\boldsymbol{x} = 0 \tag{5.61}$$

となる. ここで, 最適な x を最尤法によって求めることをつぎに考える. すると, 式 (5.31) で $y_i - y(x_i)$ をϵ_i で置き換えると全く同様である. すなわち,

$$\chi^2 = \sum_{i=1}^{N} \frac{\epsilon_i^2}{\sigma_i^2} \tag{5.62}$$

$$= \sum_{i=1}^{N} \frac{(y_i + a_{i0} + \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x})^2}{\sigma_i^2} \tag{5.63}$$

を最小にすればよいことがわかる. 式 (5.63) を行列を用いて書くと

$$\chi^2 = \boldsymbol{\epsilon}^{\mathrm{T}} C_y^{-1} \boldsymbol{\epsilon} \tag{5.64}$$

$$= (\boldsymbol{y} + \boldsymbol{a}_0 + A\boldsymbol{x})^{\mathrm{T}} C_y^{-1} (\boldsymbol{y} + \boldsymbol{a}_0 + A\boldsymbol{x}) \tag{5.65}$$

となる. χ^2 は x_i に関する偏微分がゼロのとき最小になる. つまり

$$\frac{\partial \chi^2}{\partial x_i} = 0 \qquad i = 1, 2, \ldots, r \tag{5.66}$$

これを行列で書くと

$$2A^{\mathrm{T}} C_y^{-1} (\boldsymbol{y} + \boldsymbol{a}_0 + A\boldsymbol{x}) = 0 \tag{5.67}$$

と書き換えられる. さらに

$$\Delta \equiv SS_{xx} - (S_x)^2 \tag{5.42}$$

と書くと,

$$a = \frac{S_{xx}S_y - S_xS_{xy}}{\Delta} \tag{5.43}$$

$$b = \frac{SS_{xy} - S_xS_y}{\Delta} \tag{5.44}$$

と, a, b の値が得られる. 最後に, 推定値 a, b の誤差 (不定性) について考える. それぞれのデータ y_iの不定性は互いに独立であるとすると, 誤差の伝播は

$$\sigma_f^2 = \sum_{i=1}^{N} \sigma_i^2 \left(\frac{\partial f}{\partial y_i}\right)^2 \tag{5.45}$$

となる. ここで, σ_fは, 任意の関数 fの標準偏差である. y_iによる a, b の偏微分はそれぞれ

$$\frac{\partial a}{\partial y_i} = \frac{S_{xx} - S_x x_i}{\sigma_i^2 \Delta} \tag{5.46}$$

$$\frac{\partial b}{\partial y_i} = \frac{Sx_i - S_x}{\sigma_i^2 \Delta} \tag{5.47}$$

となる. これを, 式 (5.45) に代入して計算すると

$$\sigma_a^2 = \frac{S_{xx}}{\Delta} \tag{5.48}$$

$$\sigma_b^2 = \frac{S}{\Delta} \tag{5.49}$$

これで, 推定値の分散が得られたことになる.

一般の線形問題

次に, 数個の未知量 $x_1, x_2, \ldots, x_r$ がある一般の場合について考える. この型の問題は第 3章の練習問題にも出てくる通り, 測定データの個数 N が未知量の数より多い場合の最適な未知量の推定の際に良く用いられる.

ここでは, 未知量 $x_1, x_2, \ldots, x_r$と, 測定可能な量 $\eta_1, \eta_2, \ldots, \eta_N$ の間には, 線形の関係

$$\eta_i = p_{i0} + p_{i1}x_1 + p_{i2}x_2 + \ldots + p_{ir}x_r \qquad (i = 1, 2, \ldots, N) \tag{5.50}$$

少し書き換えて,

$$\eta_i + a_{i0} + a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{ir}x_r = 0 \qquad (i = 1, 2, \ldots, N) \tag{5.51}$$

ここで, 列ベクトル

$$\boldsymbol{a}_i = \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{ir} \end{pmatrix} \qquad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} \tag{5.52}$$

を定義すると, 式 (5.51) は次のように書き直せる.

$$\eta_i + a_{i0} + \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x} = 0 \qquad (i = 1, 2, \ldots, N) \tag{5.53}$$

の最小化に帰着する. また, 簡単な場合として σ_i が一定値 σ のときを考えると, 単に, 測定誤差の自乗の和

$$Q = \sum_{i=1}^{N} (y_i - y(x_i))^2 \tag{5.32}$$

の最小化に帰着される.

5.2.2　最小自乗法–線形回帰

最小自乗法とは, 測定誤差の自乗の和の最小化 (least-squares fitting) を maximum likelihood estimator として用いるモデルフィットの手法である. この方法の暗黙の仮定は上述のとおりである. より一般的な χ^2-fitting なども考えられる. また, 確率変数間の関係式 (モデル) も一般の関数形でよいが, ここでは最も簡単な線形の関係の場合について述べる.

直線のあてはめ

第一の例題として数個の測定点が与えられた時に, それを最も良く説明する直線を決めることを考える. 即ち, 二つの確率変数 x, yの間に

$$y = a + bx \tag{5.33}$$

という線形の関係が予想される場合について考える. ただし, 独立変数 x は十分精度良く決めることができ従属変数 y のみ測定誤差を含むとする. 実際の実験では制御パラメータが独立変数 x に対応する. いま, N個のデータ,

$$(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N) \tag{5.34}$$

と従属変数 y の標準偏差

$$\sigma_1, \sigma_2, \ldots, \sigma_N \tag{5.35}$$

が与えられた時, パラメータ a, b の最ももっともらしい推定値を求めることを考える. いま, 最小化するべきχ^2は a, b の関数として

$$\chi^2(a,b) = \sum_{i=1}^{N} \left(\frac{y_i - a - bx_i}{\sigma_i} \right)^2 \tag{5.36}$$

と書ける. これを最小にする a, b の値を求める. $\chi^2(a,b)$ の a, b による微分は最小値をとる点でゼロになることより,

$$\frac{\partial \chi^2}{\partial a} = -2 \sum_{i=1}^{N} \frac{y_i - a - bx_i}{\sigma_i^2} = 0 \tag{5.37}$$

$$\frac{\partial \chi^2}{\partial b} = -2 \sum_{i=1}^{N} \frac{x_i(y_i - a - bx_i)}{\sigma_i^2} = 0 \tag{5.38}$$

となる. ここで, 簡単のため下記の諸量を定義する.

$$\begin{array}{lll} S \equiv \sum_{i=1}^{N} \frac{1}{\sigma_i^2} & S_x \equiv \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} & S_y \equiv \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2} \\ S_{xx} \equiv \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} & S_{xy} \equiv \sum_{i=1}^{N} \frac{x_i y_i}{\sigma_i^2} & \end{array} \tag{5.39}$$

このとき, 式 (5.37) , (5.38) は

$$aS + bS_x = S_y \tag{5.40}$$

$$aS_x + bS_{xx} = S_{xy} \tag{5.41}$$

となる. ここで$\Gamma(n/2)$ はガンマ関数である. χ^2_nの平均値及び分散はそれぞれ

$$\mu_{\chi^2} = n \tag{5.25}$$
$$\sigma^2_{\chi^2} = 2n \tag{5.26}$$

である. χ^2-分布は, 自由度が大きくなるにつれ正規分布に近づく.

5.2　最小自乗法–データのモデルフィットの簡単な例

適当なモデルを考え, 最もデータにフィットするようにモデルパラメータを決定するという手法は, データ解析における重要な考え方である.

ここでは, 一般的な最尤法について簡単に説明したあと, 最も簡単な例である, 最小自乗当てはめ (least-squares fit) による線形回帰について述べる.

5.2.1　最尤法 (maximum likelihood method)

モデルフィットのためのパラメータ推定

N個のデータ, $(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)$ が与えられたとき, もっともらしいモデルパラメータを決定することを考える. 統計学では, ある推定値が正しいということを証明することは, 原理的に不可能である. しかし, ある推定値がほぼあり得ないということは言える. 即ち, あるパラメータ値を仮定した時, 逆に観測データ, $(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)$ が実現される確率が十分小さい時, もとのパラメータ値が間違いであったと言える. このような, " 事後確率 " のことを尤度と呼ぶ. 従って, 尤度が最大になる様に, パラメータを決定すればよい.

いま, yは x の関数であり, yの測定誤差 (観測値とモデルによる推定値との差のことをこう呼ぶ) は各データで独立かつガウス的 (正規的) であるとすると, 観測データ,

$$(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N) \tag{5.27}$$

が実現される確率 Pは, 各データ (x_i, y_i) が実現される確率

$$\frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(y_i - y(x_i))^2}{2\sigma_i^2}\right) \tag{5.28}$$

の積である. 即ち,

$$P = (2\pi)^{-N/2} \prod_{i=1}^{N} \left\{ \frac{1}{\sigma_i} \exp\left(-\frac{(y_i - y(x_i))^2}{2\sigma_i^2}\right) \right\} \tag{5.29}$$

である. ここで σ_i は y_i の標準偏差である. この Pを最大にするように関数形 $y(x)$ のモデルパラメータを決定すればよい. または, $L = -\ln P$を最小にしてもよい. 即ち,

$$L = \frac{N}{2} + \sum_{i=1}^{N} \ln \sigma_i + \sum_{i=1}^{N} \frac{(y_i - y(x_i))^2}{2\sigma_i^2} \tag{5.30}$$

を最小化する. ここで, N, σ_iは定数であるから, 結局 L の最小化は, 測定誤差の自乗に分散の逆数の重みを掛けたものの和

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_i - y(x_i))^2}{\sigma_i^2} \tag{5.31}$$

$$E[(\eta_j-\mu)^2] = \sigma^2 \tag{5.17}$$

$$E[(\overline{\eta}-\mu)^2] = E\left[\left(\frac{1}{N}\sum_{j=1}^{N}\eta_j-\mu\right)^2\right] \tag{5.18}$$

$$= \frac{1}{N^2}E\left[\left(\sum_{j=1}^{N}(\eta_j-\mu)\right)^2\right] \tag{5.19}$$

ここで, η_jは互いに独立だから

$$E[(\overline{\eta}-\mu)^2] = \frac{1}{N^2}E\left[\left(\sum_{j=1}^{N}(\eta_j-\mu)\right)^2\right]$$

$$= \frac{1}{N^2}(N\sigma^2) = \frac{\sigma^2}{N} \tag{5.20}$$

以上, 式 (5.12), (5.16), (5.17), (5.20) より

$$E[\hat{\sigma^2}] = \frac{1}{N-1}(N\sigma^2-\sigma^2) = \sigma^2 \tag{5.21}$$

即ち, $\frac{1}{N}\sum_{j=1}^{N}(\eta_j-\overline{\eta})^2$ではなく, $\frac{1}{N-1}\sum_{j=1}^{N}(\eta_j-\overline{\eta})^2$が偏りのない分散の推定値である.

さらに式 (5.20) より平均の不偏推定量$\bar{\eta}$の分散はσ^2/Nとなることが分かる. これはNの値が大きくなるほど推定値のばらつきが減少することを示している. 誤差の大きさ (標準偏差) でいうと$1/\sqrt{N}$で減少している. このように誤差伝播における有名な事実が具体的な確率分布の形を仮定することなく示された.

5.1.2　重要な確率分布

観測から得られたデータの平均値や分散自体が, ある確率分布に従う確率変数と見なせる場合がよくある. この考え方が, 誤差の推定の基本となる. 従って, ここで, データ解析の分野でよく出てくるいくつかの重要な分布について説明しておく.

正規分布 (ガウス分布)

確率密度関数 $p(x)$ が

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma_x}\exp\left[-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right] \tag{5.22}$$

で表されるとき, 確率変数 x は正規分布に従うという. ここでμ_x, σ_x^2 はそれぞれ, x の平均値及び分散である. 物理的問題において正規分布が重要であるひとつの理由は, 中心極限定理である. これは, 互いに独立な確率変数の和は正規分布に近づいていくというものである.

χ^2-分布

$z_1, z_2, \ldots, z_n$はそれぞれ, 平均値ゼロ, 分散 1 の正規分布をする互いに独立な確率変数とする. 新たな確率変数として

$$\chi_n^2 = z_1^2 + z_2^2 + \cdots + z_n^2 \tag{5.23}$$

を考える. この確率変数χ_n^2が従う分布を自由度 n のχ^2-分布という. 確率密度関数は

$$p(\chi_n^2) = \frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)}(\chi^2)^{\frac{n}{2}-1}e^{-\frac{\chi^2}{2}} \quad (\chi^2 \geq 0) \tag{5.24}$$

今の場合, 母平均μおよび母分散σ^2の推定値としては, 標本平均$\bar{\eta}$および標本分散$\bar{\sigma^2}$がすぐ思いつく. しかし, これらが最も適切な推定量であるかどうかは, 必ずしも明らかではない.

推定値が持つべき性質のひとつとして不偏ということがある. これはある母集団の持つ統計量 (母数) α の推定値$\hat{\alpha}$が

$$E[\hat{\alpha}] = \alpha \tag{5.5}$$

を満たす時これを不偏推定量 (unbiased estimator) と呼ぶ, として定義されるものである.

実際, 母平均μおよび母分散σ^2に対する不偏推定量はそれぞれ

$$\hat{\mu} = \frac{1}{N}\sum_{j=1}^{N}\eta_j = \bar{\eta} \tag{5.6}$$

$$\hat{\sigma^2} = \frac{1}{N-1}\sum_{j=1}^{N}(\eta_j - \bar{\eta})^2 \neq \bar{\sigma^2} \tag{5.7}$$

と書けることが知られている. 即ち, 標本平均は母平均の不偏推定量であるが, 標本分散は母分散の不偏推定量ではない.

ここで, 上記の結果の証明を示す. まず母平均の推定量については,

$$E[\bar{\eta}] = E[\frac{1}{N}\sum_{j=1}^{N}\eta_j] \tag{5.8}$$

$$= \frac{1}{N}\sum_{j=1}^{N}E[\eta_j] \tag{5.9}$$

$$= \mu \tag{5.10}$$

となり, 確かに標本平均が不偏推定量である. 次に分散の推定量$\hat{\sigma^2}$であるが, これは$\frac{1}{N}\sum_{j=1}^{N}(\eta_j - \overline{\eta})^2$ではなく, $\frac{1}{N-1}\sum_{j=1}^{N}(\eta_j - \overline{\eta})^2$であることを簡単に述べる. 分散の推定値$\hat{\sigma^2}$の期待値は,

$$E[\hat{\sigma^2}] = E\left[\frac{1}{N-1}\sum_{j=1}^{N}(\eta_j - \overline{\eta})^2\right] \tag{5.11}$$

$$= \frac{1}{N-1}E\left[\sum_{j=1}^{N}(\eta_j - \overline{\eta})^2\right] \tag{5.12}$$

である. ここで

$$\sum_{j=1}^{N}(\eta_j - \overline{\eta})^2 = \sum_{j=1}^{N}(\eta_j - \mu + \mu - \overline{\eta})^2 \tag{5.13}$$

$$= \sum_{j=1}^{N}(\eta_j - \mu)^2 - 2(\overline{\eta} - \mu)\sum_{j=1}^{N}(\eta_j - \mu) + \sum_{j=1}^{N}(\overline{\eta} - \mu)^2 \tag{5.14}$$

$$= \sum_{j=1}^{N}(\eta_j - \mu)^2 - 2(\overline{\eta} - \mu)N(\overline{\eta} - \mu) + N(\overline{\eta} - \mu)^2 \tag{5.15}$$

$$= \sum_{j=1}^{N}(\eta_j - \mu)^2 - N(\overline{\eta} - \mu)^2 \tag{5.16}$$

いま,

第 5 章

データ解析の基礎

いわゆるデータ解析を行なう際には, 統計学的な概念の理解が必要不可欠である. 本章の目的はデータ解析の基礎となる統計学上の用語と手法についての最小限の解説を与えることである.

5.1　データの統計学的記述

5.1.1　平均と分散

得られたデータの特徴を一般的に記述することは, 統計学の重要な役割の一つである. 一見して不規則なデータの中から普遍的な"統計的"性質を引き出すことがデータ解析の目的でもある. その中でも, モデルにはよらずに決めることのできる量として, 平均値 (mean) μ, 分散 (variance) σ^2 などがある. 一般的に, 平均値とは, データ値ηの期待値 $E[\eta]$, 分散とは, データ値ηの平均値$\mu \equiv E[\eta]$ からのずれの自乗の期待値 $E[(\eta-\mu)^2]$ である. 分散の平方根を標準偏差 (standard deviation) という.

次に, サンプリングによるこれら統計量の推定について述べる. ひとつひとつの実験 (観測) は, ある複数の要素を持った集合の要素と考えられる. この集合は, あらゆる考えられる限りの事象から成り立っているものでありこの集合を母集団 (**population**) と呼ぶ. 各要素が持つ変数ηの母集団中での実現確率が確率密度関数 $f(\eta)$ で表現されているとすると, 変数ηの母平均μおよび母分散σ^2はそれぞれ

$$\mu = \int_{-\infty}^{\infty} \eta f(\eta) d\eta \tag{5.1}$$

$$\sigma^2 = \int_{-\infty}^{\infty} (\eta-\mu)^2 f(\eta) d\eta \tag{5.2}$$

と表される. ところが実際には母集団の分布は知られていないのが普通なので, 直接この式から母平均や母分散を計算することはできない. 実際にはいくつかの標本 (sample) (データ) からこれを推定することになる. N個の標本 (データ) , $\eta_1, \eta_2, \ldots, \eta_N$ が与えられたとき, 以下のように標本平均$\bar{\eta}$ および標本分散$\bar{\sigma^2}$ が定義される.

$$\bar{\eta} = \frac{1}{N}\sum_{j=1}^{N} \eta_j \tag{5.3}$$

$$\bar{\sigma^2} = \frac{1}{N}\sum_{j=1}^{N} (\eta_j - \bar{\eta})^2 \tag{5.4}$$

この標本抽出作業がランダムである時, 無作為抽出 (random sampling) と呼び, ひとつひとつの標本が従う確率密度分布は母集団の持つ確率密度分布と一致する. 従ってこの場合, 標本の持つ統計的な性質から母集団の持つ統計的性質を推定することが可能となるはずである.

3. バネの強さやオモリの重さがランダムな場合.
4. さらにダッシュポットを連ねるとどうなるか.

問題 2 好きな楽器の振動をモデル化して固有振動を解け. (参考文献 : Rossing, 1992 ; Fletcher & Rossing, 1991)

問題 3 マグマ溜りの振動をモデル化して固有振動を解け. 単色の火山性微動との関わりを考察せよ. (参考文献 : 山岡, 1990)

問題 4 弾性球の自由振動の分散関係を求めよ. (参考書 : 竹内, 1972)

問題 5 非圧縮流体の重力波の分散関係を求めよ. (参考書 : Turner, 1973)

1. 非粘性の場合.
2. ダンピングのある場合.

問題 6 Bénard 対流の線型安定論の分散関係を求めよ. (参考書 : Chandrasekhar, 1961)

問題 7 $\alpha - \omega$ ダイナモの分散関係を求めよ. (参考文献 : Roberts, 1972)

4.6 参考文献

固有値問題の数値解法に関しては, 戸川, 1971 ; 森, 1973 ; 森, 1984 ; 小国, 1991 ; Press et al., 1986 がある. このうち最後のものが一番実用的である.

応用問題に関しては, キッテル, 1986 ; ザイマン, 1976 ; Rossing, 1992 ; Fletcher and Rossing, 1991 ; 山岡, 1990 ; 竹内, 1972 ; Turner, 1973 ; Chandrasekhar, 1961 ; Roberts, 1972.

ただし, I_j は $j \times j$ の単位行列, $U(\theta)$ は二次元の回転行列で,

$$U(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{4.67}$$

の形をしている. ここで, θ を

$$\sin\theta = \frac{a_{i+1,i}^{(i-1)}}{\sqrt{|a_{i,i}^{(i-1)}|^2 + |a_{i+1,i}^{(i-1)}|^2}} \tag{4.68}$$

$$\cos\theta = \frac{a_{i,i}^{(i-1)}}{\sqrt{|a_{i,i}^{(i-1)}|^2 + |a_{i+1,i}^{(i-1)}|^2}} \tag{4.69}$$

と選んでおけば良い.

行列 $A^{(i)}$ と $A^{(i-1)}$ との違いは第 i 行と第 $(i+1)$ 行だけだから, 行列 $A^{(i)}$ を求める時は次の計算をするだけで良い.

$$a_{i,j}^{(i)} = a_{i,j}^{(i-1)} \cos\theta + a_{i+1,j}^{(i-1)} \sin\theta \tag{4.70}$$

$$a_{i+1,j}^{(i)} = -a_{i,j}^{(i-1)} \sin\theta + a_{i+1,j}^{(i-1)} \cos\theta \tag{4.71}$$

4.4　二, 三の補足

ここまでで, 一応固有値問題が解けるようになっているはずだ. が, 多少の注意をしておこう.

1. QR 法のような冪乗法的考え方では, 固有値の絶対値の違いを利用して固有ベクトルへの収束を図っている[9]. だから, 絶対値の同じ固有値があると (例えば, 固有値の縮退があったり, 複素共役な固有値が出てくるような場合) 右上三角行列には収束しない. この場合はブロック右上三角行列に収束する (対角線上に絶対値が等しい固有値の組に対応するブロックが並ぶ). その小ブロックを取り出して改めて固有値問題を解く必要がある.

2. 特に非対称実行列では, 複素共役な固有値のペアが出てくるが, その場合は 「ダブル QR 法」という方法を用いるのが普通である. 詳しくは第 4.6 節に挙げたような参考書を見られたい.

3. 絶対値の等しい固有値がなくても, 絶対値の近い固有値があると収束が非常に遅くなる. この困難をできるだけ避けるために 「原点移動」というテクニックを普通用いる. これも詳しくは第 4.6 節に挙げたような参考書を見られたい.

4.5　応用問題

QR 法がわかったら, いくつか応用問題を解いてみよう.

問題 1　バネとオモリを交互に N 個連ねた系の固有振動の分散関係を求めよ. (参考書 : キッテル, 1986 ; ザイマン, 1976)

1. バネの強さがすべて k, オモリの重さがすべて m の場合.
2. 一つだけ重さが他と違うオモリが入っている場合.

[9] 第 4.2.6 節参照

ただし

$$\boldsymbol{v} = \frac{\boldsymbol{b} - \begin{pmatrix} s \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}}{\left\| \boldsymbol{b} - \begin{pmatrix} s \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \right\|} \tag{4.59}$$

とすれば良い.

4.3.4　ヘッセンベルグ行列の QR 分解

さて, わざわざ行列をヘッセンベルグ形に直したご利益は QR 分解が安定に速くできるということであった. これからその方法を考えよう.

QR 分解

$$A = QR \tag{4.60}$$

を考える. これを

$$Q^{-1}A = R \tag{4.61}$$

と読み換えよう. すなわち, 問題は, ヘッセンベルグ行列 A に左からかけると右上三角行列を生み出すようなうまいユニタリ行列 Q^{-1} を求めよ, ととらえることもできる.

ここでまた例によって, はじめに第一列目の第二行目以下を 0 にし, 次に第二列目の第三行目以下を 0 にし, ... ということを順繰りに繰り返して右上三角行列に持っていきたい.

$$A^{(0)} = A \tag{4.62}$$

$$A^{(i)} = P^{(i)}A^{(i-1)} \qquad (i = 1, 2, \cdots, n-1) \tag{4.63}$$

ただし, $A^{(i)}$ は第 i 列目まで右上三角になったヘッセンベルグ行列, $P^{(i)}$ はユニタリ行列である. もしこうなっていれば,

$$Q^{-1} = P^{(n-1)}P^{(n-2)}\cdots P^{(2)}P^{(1)} \tag{4.64}$$

$$Q = P^{(1)\dagger}P^{(2)\dagger}\cdots P^{(n-2)\dagger}P^{(n-1)\dagger} \tag{4.65}$$

で Q が計算できる.

行列 $P^{(i)}$ としては次のような回転行列を使えば良いことが簡単に確かめられる.

$$P^{(i)} = \begin{pmatrix} I_{i-1} & O & O \\ O & U(\theta) & O \\ O & O & I_{n-i-1} \end{pmatrix} \tag{4.66}$$

変換で A は次の形にまでなっているはずだ.

$$A = \left(\begin{array}{ccccc|ccccc} & & & & & & & & & \\ & & & & & & & & & \\ & & C & & & & & D & & \\ & & & & & & & & & \\ & & & & & & & {b'}^T & & \\ \hline - & - & - & - & - & - & - & - & - & - \\ 0 & 0 & \cdots & 0 & & & & & & \\ 0 & 0 & \cdots & 0 & & & & & & \\ \vdots & \vdots & & \vdots & \boldsymbol{b} & & & B & & \\ 0 & 0 & \cdots & 0 & & & & & & \\ 0 & 0 & \cdots & 0 & & & & & & \end{array}\right) \tag{4.53}$$

ここで C は $j \times j$ のヘッセンベルグ行列, B は $(n-j) \times (n-j)$ の正方行列, D は $(j-1) \times (n-j)$ の行列, $\boldsymbol{b}$, $\boldsymbol{b}'$ は $(n-j) \times 1$ の列ベクトルである. この時,

$$P^{-1}AP = PAP \tag{4.54}$$

$$= \left(\begin{array}{ccccc|ccccc} & & & & & & & & & \\ & & & & & & & & & \\ & & C & & & & & Q_{n-j}D & & \\ & & & & & & & & & \\ & & & & & & & {b'}^T Q_{n-j} & & \\ \hline - & - & - & - & - & - & - & - & - & - \\ 0 & 0 & \cdots & 0 & & & & & & \\ 0 & 0 & \cdots & 0 & & & & & & \\ \vdots & \vdots & & \vdots & Q_{n-j}\boldsymbol{b} & & & Q_{n-j}BQ_{n-j} & & \\ 0 & 0 & \cdots & 0 & & & & & & \\ 0 & 0 & \cdots & 0 & & & & & & \end{array}\right) \tag{4.55}$$

となるので,

$$Q_{n-j}\boldsymbol{b} = \begin{pmatrix} s \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \tag{4.56}$$

の形となる Q_{n-j} が見つかれば目的は達せられる. 前に述べた鏡像変換の性質を使うとこれは容易にできる.

$$s = \|\boldsymbol{b}\| \text{ もしくは } -\|\boldsymbol{b}\|^{8} \tag{4.57}$$

として,

$$Q_{n-j} = I_{n-j} - 2\boldsymbol{v}\boldsymbol{v}^T \tag{4.58}$$

[8] 符号は後の $\boldsymbol{v}$ の計算で桁落ちを起こさないように $\boldsymbol{b}$ の第一成分と逆符号にとる.

さらに, 元の行列 $A(=A_1)$ をヘッセンベルグ形に直しておけば, $A_2, A_3, \ldots$ は自動的にヘッセンベルグ形になっている (このことの図形的な意味を各自考えてみよ). したがって, 行列をヘッセンベルグ形に直すのは始めに一回やっておけば良い.

ヘッセンベルグ形に直した後では, QR 分解は第 4.3.4 節の方法で行なうと良い. 第 4.2.3 節に書いたような手順でまともにグラム・シュミットの直交化をしても良いように思われるかも知れないけれが, それでは桁落ちの可能性があるし, 遅い. 第 4.3.4 節に示したような消去法的に右上三角行列を作る方法の方が数値計算上安定だし, 速い. これもヘッセンベルグ形を用いることのご利益の一つである.

4.3.2　鏡像変換

任意の行列をヘッセンベルグ形に直すことができるようになるための準備として, 鏡像変換というものを導入する. 鏡像変換とは空間の任意の一点 x を, 原点を含む平面 U に関して対称な点 x' に移すような変換である. この変換は次のように表される (各自確かめよ).

$$\boldsymbol{x}' = P\boldsymbol{x} \tag{4.47}$$

$$P = I - 2\boldsymbol{u}\boldsymbol{u}^T \tag{4.48}$$

ただし, I は単位行列, $\boldsymbol{u}$ は平面 U の単位法線ベクトルである. 行列 P は対称かつ直交行列なので, 逆行列が自分自身に等しくなり, 計算に大変便利である.

次の性質は非常に重要である. 長さの等しい任意の二つのベクトル x, y が与えられた時, x を y に移すような鏡像変換が必ず存在するということである. 幾何学的に考えればこれはほとんど自明である. 二点 x, y を線分で結び, その中点を通る平面で, その線分に垂直なものを持ってくれば良い.

数学的に書くと次のようになる. $\|x\| = \|y\|$ となるような任意の二つのベクトル x, y に対し,

$$P\boldsymbol{x} = \boldsymbol{y} \tag{4.49}$$

$$P = I - 2\boldsymbol{u}\boldsymbol{u}^T \tag{4.50}$$

となるような $\boldsymbol{u}$ が存在する. それは

$$\boldsymbol{u} = \frac{\boldsymbol{x} - \boldsymbol{y}}{\|\boldsymbol{x} - \boldsymbol{y}\|} \tag{4.51}$$

で与えられる.

4.3.3　任意の行列をヘッセンベルグ形に直す法

さて, 任意の $n \times n$ 行列 A を相似変換でヘッセンベルグ形に直すことを考えよう.

方針としては, はじめに第一列目の第三行目以下を 0 にし, 次に第二列目の第四行目以下を 0 にし, ... ということを順繰りに行なっていくことにしたい. そのためには, j 回目の相似変換が $i(i \leq j-1)$ 列目に影響を与えないような変換でなければ駄目だ. そこで, (ここは論理的脈絡はないが) j 回目の相似変換には次の形の鏡像変換を試してみてはどうかと考える.

$$P = \begin{pmatrix} I_j & O \\ O & Q_{n-j} \end{pmatrix} \tag{4.52}$$

ここで, I_j は $j \times j$ の単位行列, Q_{n-j} は $(n-j) \times (n-j)$ の鏡像変換行列である. $(j-1)$ 回目までの相似

2. 第 1 近似
   P_0 を A で一次変換して, それにグラム ・シュミットの直交化をして作った回転行列を, 求める回転行列の第 1 近似 P_1 とする.

$$AP_0 = P_1 R_1 \tag{4.41}$$

   求める右上三角行列の第 1 近似は

$$A_2 = P_1^\dagger A P_1 \tag{4.42}$$

3. 以下同様
   第 k 近似
   P_{k-1} を A で一次変換して, それにグラム ・シュミットの直交化をして作った回転行列を, 求める回転行列の第 k 近似 P_k とする.

$$AP_{k-1} = P_k R_k \tag{4.43}$$

   求める右上三角行列の第 k 近似は

$$A_{k+1} = P_k^\dagger A P_k \tag{4.44}$$

この操作を続けると A_k は右上三角行列になる. 固有値は絶対値の大きい順番に左上から右下に並ぶということも想像されるだろう. また, P_k は固有ベクトルにグラムシュミットの直交化を施したものになる.

これが QR 法の考え方である. すなわち QR 法とは

反復の度ごとに直交化を繰り返す冪乗法

である. 以上の操作を簡略化すると第 4.2.2 節で与えた処方箋になる. 対応関係は

$$P_k = Q_1 Q_2 \cdots Q_k \tag{4.45}$$

となっている (各自確かめてみよ).

厳密な証明を知りたい方は, 第 4.6 節に挙げたような参考書を見られたい.

4.3　ヘッセンベルグ行列の利用

4.3.1　ヘッセンベルグ行列を用いるご利益

以上のやり方で原理的には固有値問題を QR 法で解けるようになったはずである. しかしながら, 第 4.1 節で述べたように一旦行列をヘッセンベルグ形という簡単な形に直してから QR 分解をやった方が速い.

ヘッセンベルグ行列とは

$$\begin{pmatrix} * & * & * & \cdots & * & * & * \\ * & * & * & \cdots & * & * & * \\ 0 & * & * & \cdots & * & * & * \\ 0 & 0 & * & \cdots & * & * & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & * & * & * \\ 0 & 0 & 0 & \cdots & 0 & * & * \end{pmatrix} \tag{4.46}$$

の形をした行列である (* は零でない成分を示す) . すなわち, $i \geq j+2$ の範囲の成分がすべて零であるような行列である. 任意の行列は 4.3.3 節の方法でヘッセンベルグ形に直すことができる.

なっているはずである. 何故かというと, 固有ベクトルは一次変換 A によって方向を変えないベクトルだからである.

もうちょっときちんと考えてみよう. A の固有値を $\lambda_1, \lambda_2, \cdots, \lambda_n$ (ただし, $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$ [5]) とし, それに対応する固有ベクトルを$\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots, \boldsymbol{t}_n$ とする. ただし, $\boldsymbol{t}_i (i = 1, \cdot, n)$ は規格化されているものとする. いま,

$$\boldsymbol{x} = \sum_{j=1}^{n} f_j \boldsymbol{t}_j \tag{4.32}$$

と書けたとする. これを A で k 回変換したものを $\boldsymbol{x}^{(k)}$ と書く.

$$\begin{aligned} \boldsymbol{x}^{(k)} &= A^k \boldsymbol{x} && (4.33)\\ &= \sum_{j=1}^{n} f_j A^k \boldsymbol{t}_j && (4.34)\\ &= \sum_{j=1}^{n} f_j \lambda_j^k \boldsymbol{t}_j && (4.35) \end{aligned}$$

となり, $f_1 \neq 0$ ならば $k \to \infty$ のとき,

$$\lim_{k\to\infty} \frac{\boldsymbol{x}^{(k)}}{\|\boldsymbol{x}^{(k)}\|} = \pm \boldsymbol{t}_1 \tag{4.36}$$

になることが見てとれる[6] . すなわち, $\boldsymbol{x}$ ($\boldsymbol{t}_1$ と直交しない任意のベクトル) を何度も一次変換すると絶対値が最大の固有値の固有ベクトルに近付く. 「冪乗法」とはこのようにして絶対値最大の固有値と固有ベクトルを求める方法である.

他の固有ベクトルは, $\boldsymbol{t}_1$ と直交する空間を調べれば出てくる. $f_2 \neq 0$ のとき

$$\lim_{k\to\infty} \frac{\boldsymbol{x}^{(k)} - (\boldsymbol{x}^{(k)}, \boldsymbol{t}_1)\boldsymbol{t}_1}{\|\boldsymbol{x}^{(k)} - (\boldsymbol{x}^{(k)}, \boldsymbol{t}_1)\boldsymbol{t}_1\|} = \pm \frac{\boldsymbol{t}_2 - (\boldsymbol{t}_2, \boldsymbol{t}_1)\boldsymbol{t}_1}{\|\boldsymbol{t}_2 - (\boldsymbol{t}_2, \boldsymbol{t}_1)\boldsymbol{t}_1\|} \tag{4.37}$$

となり, 次に大きい固有値の固有ベクトルが見えてくる. このようにしていけば, 固有ベクトルにグラム・シュミットの直交化をしたもの, すなわち求めるべき回転行列, が一列ずつ求まってくるはずだ.

しかし良く考えてみると, 式 4.37 はそのままでは使えない. $\boldsymbol{t}_1$ は式 4.36 から求まるのだから, 式 4.37 を本当に使おうとすると極限をどう取っていいかわからないし, 正しく$\boldsymbol{t}_1$ がわかっていたところで桁落ちしそうな式であるからだ. 正しいやり方は, 常に反復の最中で $\boldsymbol{x}^{(k)}$ と直交する空間の中で「冪乗法」を適用していくことだ. 以上をまとめると, 次のような方策になる.

1. 第 0 近似
   出発点では求める回転行列の第 0 近似 P_0 を単位行列 I にとる.

$$P_0 = I \tag{4.38}$$

   求める右上三角行列の第 0 近似は A 自身になる[7].

$$\begin{aligned} A_1 &= P_0^{\dagger} A P_0 && (4.39)\\ &= A && (4.40) \end{aligned}$$

[5] 絶対値が等しい固有値がある場合については第 4.4 節を見よ
[6] ここで $\pm$ の符号をつけたのは固有値が負の時, k が一つ変わるたびにプラスマイナスが交互に出てくるということを表したつもりである. 固有ベクトルは方向だけが問題となる量なので符号のことは気にしなくて良い.
[7] 都合上, 添字は近似の次数に 1 加えておく

4.2.4　右上三角行列

ここで, QR 法の原理を考える準備として右上三角行列の図形的な意味を考察しておこう.

右上三角行列 R を考える.

$$R = (\boldsymbol{r}_1 \boldsymbol{r}_2 \cdots \boldsymbol{r}_n), \tag{4.28}$$

$$\boldsymbol{r}_1 = \begin{pmatrix} r_{11} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \boldsymbol{r}_2 = \begin{pmatrix} r_{12} \\ r_{22} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots \boldsymbol{r}_n = \begin{pmatrix} r_{1n} \\ r_{2n} \\ r_{3n} \\ \vdots \\ r_{nn} \end{pmatrix}. \tag{4.29}$$

また, 単位ベクトルを $\boldsymbol{e}_1, \boldsymbol{e}_2 \cdots \boldsymbol{e}_n$ とおく.

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \cdots \boldsymbol{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \tag{4.30}$$

R を変換行列とする一次変換を考える. 単位ベクトル $\boldsymbol{e}_i$ の R による像は $\boldsymbol{r}_i$ となる. すなわち $\boldsymbol{e}_i$ の像は $\{\boldsymbol{e}_j (j \leq i)\}$ の張る空間内に入る. これが右上三角行列の図形的な意味である.

4.2.5　右上三角行列を与える回転変換

それでは任意の行列 (ここでは簡単のため実行列としておく) をどのように回転したら右上三角になるであろうか.

A を $n \times n$ 行列とする. A の固有ベクトルを (仮にわかっているとして) $\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots, \boldsymbol{t}_n$ とする.

ここで右上三角行列の図形的意味を考えると, 次のようにして回転行列を作れば右上三角になることがわかるだろう.

まず, $\boldsymbol{t}_1$ 方向の単位ベクトルを $\boldsymbol{p}_1$ とする. 次に, $\boldsymbol{t}_1$ と $\boldsymbol{t}_2$ の張る空間内で $\boldsymbol{t}_1$ に直交する単位ベクトルを $\boldsymbol{p}_2$ とする. さらに, $\boldsymbol{t}_1$, $\boldsymbol{t}_2$, $\boldsymbol{t}_3$ の張る空間内で $\boldsymbol{t}_1$ と $\boldsymbol{t}_2$ の張る空間に直交する単位ベクトルを $\boldsymbol{p}_3$ とする. このようにして単位ベクトル $\boldsymbol{p}_i (i = 1, \cdots, n)$ を作る. これはとりも直さず, A の固有ベクトルにグラム ・シュミットの直交化を施したものである. 求める回転行列は

$$P = (\boldsymbol{p}_1 \boldsymbol{p}_2 \cdots \boldsymbol{p}_n) \tag{4.31}$$

で与えられる.

4.2.6　QR 法の考え方

実際には前もって固有ベクトルがわかっているわけではないから, すぐに右上三角にはできない. 固有ベクトルを求める手立てを考えなければならない.

ここで, いわゆる 「冪乗法」の考え方が役に立つ. 「冪乗法」とは何か. いま勝手にベクトル x を持ってくる. x_1 を一次変換 A で変換して Ax を作り, その Ax を変換して A^2x を作り, ... といったことを繰り返すときっとどこかに収束するんじゃないかと考える. 収束するとすれば, それは固有ベクトルの一つに

1. $\boldsymbol{a}_1$ を規格化して $\boldsymbol{q}_1$ を作る.

$$\boldsymbol{q}_1 = \frac{\boldsymbol{a}_1}{\|\boldsymbol{a}_1\|} \tag{4.15}$$

2. $\boldsymbol{a}_2$ から $\boldsymbol{q}_1$ に直交する部分だけを取り出す.

$$\boldsymbol{u}_2 = \boldsymbol{a}_2 - (\boldsymbol{a}_2, \boldsymbol{q}_1)\boldsymbol{q}_1 \tag{4.16}$$

それを規格化する.

$$\boldsymbol{q}_2 = \frac{\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|} \tag{4.17}$$

3. $\boldsymbol{a}_3$ から $\boldsymbol{q}_1$ と $\boldsymbol{q}_2$ とに直交する部分を取り出す.

$$\boldsymbol{u}_3 = \boldsymbol{a}_3 - (\boldsymbol{a}_3, \boldsymbol{q}_1)\boldsymbol{q}_1 - (\boldsymbol{a}_3, \boldsymbol{q}_2)\boldsymbol{q}_2 \tag{4.18}$$

それを規格化する.

$$\boldsymbol{q}_3 = \frac{\boldsymbol{u}_3}{\|\boldsymbol{u}_3\|} \tag{4.19}$$

4. 以下同様

とする. こうしてできた

$$\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_n \tag{4.20}$$

が正規直交系である.

以上のグラム・シュミットの直交化をきちんと行列で書くと,

$$A = QR \tag{4.21}$$

ただし

$$A = (\boldsymbol{a}_1 \boldsymbol{a}_2 \cdots \boldsymbol{a}_n) \tag{4.22}$$

$$Q = (\boldsymbol{q}_1 \boldsymbol{q}_2 \cdots \boldsymbol{q}_n) \tag{4.23}$$

$$R = (r_{jk}) \tag{4.24}$$

$$\boldsymbol{u}_k = \begin{cases} \boldsymbol{a}_1 & k = 1 \\ \boldsymbol{a}_k - \sum_{j=1}^{k-1} (\boldsymbol{a}_k, \boldsymbol{q}_j)\boldsymbol{q}_j & k = 2, 3, \cdots, n \end{cases} \tag{4.25}$$

$$\boldsymbol{q}_k = \frac{\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|} \tag{4.26}$$

$$r_{j,k} = \begin{cases} (\boldsymbol{a}_k, \boldsymbol{q}_j) & j < k \\ \|\boldsymbol{u}_k\| & j = k \\ 0 & j > k \end{cases} \tag{4.27}$$

となる. これが QR 分解である.

この分解の一意性も証明できるのだが, 興味があれば第 4.6 節に挙げてあるような適当な教科書を見られたい.

実際の数値計算では, ヘッセンベルグ形を経由する時は QR 分解に以上の式は使わない. 第 4.3.4 節の消去法的な方法の方が数値的に安定だし速い.

(a) A_k を QR 分解する[3].

$$A_k = Q_k R_k \tag{4.5}$$

(b) 積 $R_k Q_k$ を作る.

$$A_{k+1} = R_k Q_k \tag{4.6}$$

3. この操作を繰り返せば A_k は右上三角行列 (A がエルミート行列なら対角行列) に近付き[4], 対角成分に固有値が並んでいる.

さらに, 固有値は絶対値の大きい順番に左上から右下に並ぶ, ということも知られている.

固有ベクトルは次のようにして求まる. 以上の操作をずっと繰り返すと結局次のような変換が行なわれたことになっている.

$$A_\infty = P^\dagger A P \tag{4.7}$$

ただし,

$$P = Q_1 Q_2 \cdots \tag{4.8}$$

である. A_∞ は右上三角行列になっているはずである. いま, $A_\infty = (a_{ij}^{(\infty)})$ と書き, P を

$$P = (\boldsymbol{p}_1 \boldsymbol{p}_2 \cdots \boldsymbol{p}_n) \tag{4.9}$$

と書く.

- A がエルミートなら, A_∞ は対角行列であり, $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n$ が固有ベクトルである.
- A_∞ に非対角成分が残っている場合, 固有ベクトル $\boldsymbol{t}_i \quad (i = 1, 2, \cdots, n)$ は次のようにして求まる (各自確かめよ).

$$\boldsymbol{t}_1 = \boldsymbol{p}_1 \tag{4.10}$$

$$\boldsymbol{t}_i = \boldsymbol{p}_i + \sum_{j=1}^{i-1} b_j \boldsymbol{p}_j \qquad (i = 2, 3, \ldots, n) \tag{4.11}$$

ただし, b_j は次の表式によって $j = i-1$ から順繰りに $j = 1$ まで計算される.

$$b_{i-1} = \frac{a_{i-1,i}^{(\infty)}}{a_{i,i}^{(\infty)} - a_{i-1,i-1}^{(\infty)}} \tag{4.12}$$

$$b_j = \frac{a_{j,i}^{(\infty)} + \sum_{k=j+1}^{i-1} b_k a_{j,k}^{(\infty)}}{a_{i,i}^{(\infty)} - a_{j,j}^{(\infty)}} \qquad (j = i-2, i-3, \ldots, 1) \tag{4.13}$$

4.2.3　QR 分解とグラム・シュミットの直交化

さて, QR 分解とは何かをまとめておこう.

QR 分解とはグラム・シュミットの直交化を行列表現したものである. グラム・シュミットの直交化とは任意の一次独立なベクトルの組から正規直交系をなすあるベクトルの組を作る方法である. それは次のような方法で作られる. まず, n 次元ベクトル空間における一次独立なベクトルの組

$$\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n \tag{4.14}$$

を持ってくる. そのベクトルに対し,

[3] このやり方は, 第 4.2.3 節および 第 4.3.4 節を見よ
[4] 厳密にいうと, 右上三角行列に近付いても非対角成分は収束はしない. だがこのことはいまの問題に関しては重要ではない

2. QR 法の処方箋 : 早速やり方だけはわかる.

3. QR 分解とグラム ・シュミットの直交化: 準備その 1 .

4. 右上三角行列 : 準備その 2 .

5. 右上三角行列を与える回転変換 : 準備その 3.

6. QR 法の考え方: なぜ QR 法はうまくいくのか.

4.2.2 QR 法の処方箋

早速 QR 法を使いたいという人もいるだろうし, 前 subsection の説明で Q とは何だ, R とは何だと思った人もいるだろうし, 第 4.1 節で分解とは何だと思った人もいるだろう. 本 subsection では以上のような要求に答えるべく言葉の説明と処方箋だけを与える. 理窟付けは後でゆっくりやる.

まず前節で与えた定義を数学的に書くと次のようになる. いま, 固有値を求めたい $n \times n$ 行列を A と書く. そして, 次の手順を踏め.

1. $A_1 = A$ とおく.

2. $k = 1, 2, \cdots$ に対して次の操作を繰り返す.

   (a) とあるユニタリ行列 Q_k があるものとする.

   $$Q_k^{-1} = Q_k^\dagger \tag{4.2}$$

   ただし, † は行列の転置複素共役を示す.

   (b) Q_k で A_k をユニタリ変換する.

   $$A_{k+1} = Q_k^\dagger A_k Q_k \tag{4.3}$$

3. この操作を繰り返せば A_k は右上三角行列に近付き, 対角成分に固有値が並んでいる.

QR 法の Q とはこの 「とあるユニタリ行列」のことである. では, この 「とあるユニタリ行列」はどのように求めたら良いのだろうか.

任意の正則な正方行列 A は次のように一意的に分解できることが知られている [2]. これを QR 分解と呼ぶ.

$$A = QR \tag{4.4}$$

ただし, Q はユニタリ行列, R は右上三角行列である.

「とあるユニタリ行列」としては A_k を QR 分解した時のユニタリ行列をとれ, というのが QR 法である (どうしてそうすればいいのかは追い追い説明していく). したがって, QR 法の定義を文字の通り書いてしまえば,

QR 分解を利用した反復法によって, 固有値を求める方法

である.

改めて QR 法の手順を書く. 固有値を求めたい行列を A とする.

1. $A_1 = A$ とする.

2. $k = 1, 2, \cdots$ に対して次の操作を繰り返す.

[2] 詳しくは第 4.2.3 節を見よ.

このうち, ハウスホルダー法は汎用性があり, 決定版といわれているので, ここではこれを学習する.
第二ステップで使われる方法には次のようなものがある.

- 分解 (factorization) を用いる方法: LR 法, QR 法
- 特性多項式を用いる方法 : スツルム法, ...

このうち, 非対称行列にも使えて汎用性がありしかも比較的安定な方法であるといわれている QR 法をここでは学習する.
本章は次の節から構成されている.

1. 概観 : いま読んでいるところ.
2. QR 法: 上述の第二ステップである. これでともかく固有値が計算できるようになる.
3. ヘッセンベルグ行列の利用 : 上述の第一ステップである. これで計算が早くなるはずだ.
4. 二, 三の補足 : 少し補足を.
5. 応用問題 : せっかくだから, 物理的な応用問題を解いてみよう.
6. 参考文献 : 詳しく知りたい方は他の本も読んでみるべし.

4.2　QR 法

4.2.1　QR 法の基本方針

固有値問題を解くということは相似変換によって行列を対角化するということである. すなわち, A を固有値を求めたい行列とする時, うまい正則行列 P を持って来て,

$$\tilde{A} = P^{-1}AP \tag{4.1}$$

を対角行列にしたい. 一般には, 対角化は有限回の操作ではできないので相似変換を繰り返すことで対角行列に近付けるということになる.
ところで, 数値的にこの相似変換をする時には, 桁落ちを生じないような性質の良い相似変換をしていきたい. 特に逆行列が入ってくるところが危ない. 相似変換がユニタリ変換だけで済めばこの問題は回避される. なぜかというと, ユニタリ変換は, 回転のようなものだから, 座標系をゆがめないからである. 逆行列は転置複素共役をとるだけで良く, 割算が入ってこない.
とはいえ, ユニタリ変換で対角化できるのは, エルミート行列に限られる. それだけでも結構役に立つが, 一般的な行列には役に立たないかというとそんなことはない. 一般に, ユニタリ変換だけで行列を右上三角行列にできる. 右上三角行列の対角成分には固有値が並んでいるから, 固有値を求めるだけならそれで十分だし, 固有ベクトルを求めるのも容易である.
そういう考えで QR 法はできている. QR 法とは,

　　ユニタリ変換を反復することで行列を右上三角行列に近付けていく

という方法である.
本節は次の subsection から構成されている.

1. 基本方針 : いま読んでいるところ.

第 4 章

行列計算その 2 : 固有値固有ベクトル

4.1 固有値問題の解法の概観

固有値問題の数値解法には実にさまざまのものがある. 現在広く用いられているアルゴリズムは大抵次の二つのステップを踏んでいる.

1. 元の行列を,
   - 三重対角行列 (元が対称行列の場合)
   - ヘッセンベルグ行列[1] (元が非対称行列の場合)

   に相似変換する.
2. 三重対角行列もしくはヘッセンベルグ行列に対して固有値を計算する.

第一ステップでは, ある成分 (達) を 0 にするというような消去法を用いて行列の形を簡単にする. これは有限の計算回数で行なうことができる. 第二ステップでは, 何らかの反復法で固有値を求める.

計算ステップを二つに分ける理由は次の通りである. 固有値問題の計算は一般には有限回の計算では不可能であり, 何らかの反復法を使わざるを得ない. 消去法を反復して使って行列を対角形に近付けていく, という解法が概念的にはもっとも単純だ. が, 大きい行列ではこの方法では効率が悪いことが知られている. 第二ステップで用いられるような 「何らかの反復法」で最初から最後まで計算することも可能だ. けれども, 予め有限回の手続きで行列を単純な形にしておいて, それからその単純な形に適した収束の早い 「何らかの反復法」を適用することにすれば計算は格段に早くなるだろう. そういうわけで, ちょっと面倒だが 2 ステップの計算をするのが普通になっている.

第一ステップで使われる方法には次のようなものがある.

- ハウスホルダー法
- ギブンス法
- ランチョス法
- ガウス法
- ヘッセンベルグ法

[1] 第 4.3.1 節を参照せよ

3.7.5　付録 5: 条件数

解の反復改良の話をしたときに　条件が悪い行列」という言い方がでてきたが, ここではこの　条件が悪い」とか, あるいは反対に　条件が良い」とかいうのはどいうことを指すのかを説明する.

そのためにまず条件数というものを定義しよう.

連立一次方程式 $Ax = b$ を数値的に解いて得られた近似解を x_1 とする. 残差ベクトル r は

$$\boldsymbol{r} = \boldsymbol{b} - A\boldsymbol{x}_1$$

である. x_1 は, b が $b - r$ にずれた時の正確な解である.

ここで $\|\cdot\|$ を, ベクトルおよび行列に関する任意のノルムとする. このとき

$$\|\boldsymbol{b}\| = \|A\boldsymbol{x}\| \leq \|A\|\|\boldsymbol{x}\|$$

より

$$\frac{1}{\|\boldsymbol{x}\|} \leq \frac{\|A\|}{\|\boldsymbol{b}\|} \tag{3.49}$$

である. 一方, A が正則なら

$$\|\boldsymbol{x} - \boldsymbol{x}_1\| = \|A^{-1}\boldsymbol{r}\| \leq \|A^{-1}\|\|\boldsymbol{r}\| \tag{3.50}$$

であるから. 式 (3.49) と式 (3.50) を辺々かけ算すると, 相対誤差の上界が

$$\frac{\|\boldsymbol{x} - \boldsymbol{x}_1\|}{\|\boldsymbol{x}\|} \leq \|A\|\|A^{-1}\| \cdot \frac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|} \tag{3.51}$$

のように得られる. r は b の誤差としての意味を持つから, 式 (3.51) は解の誤差を b の誤差によって見積もっていることになる. ここで

$$cond(A) \equiv \|A\|\|A^{-1}\| \tag{3.52}$$

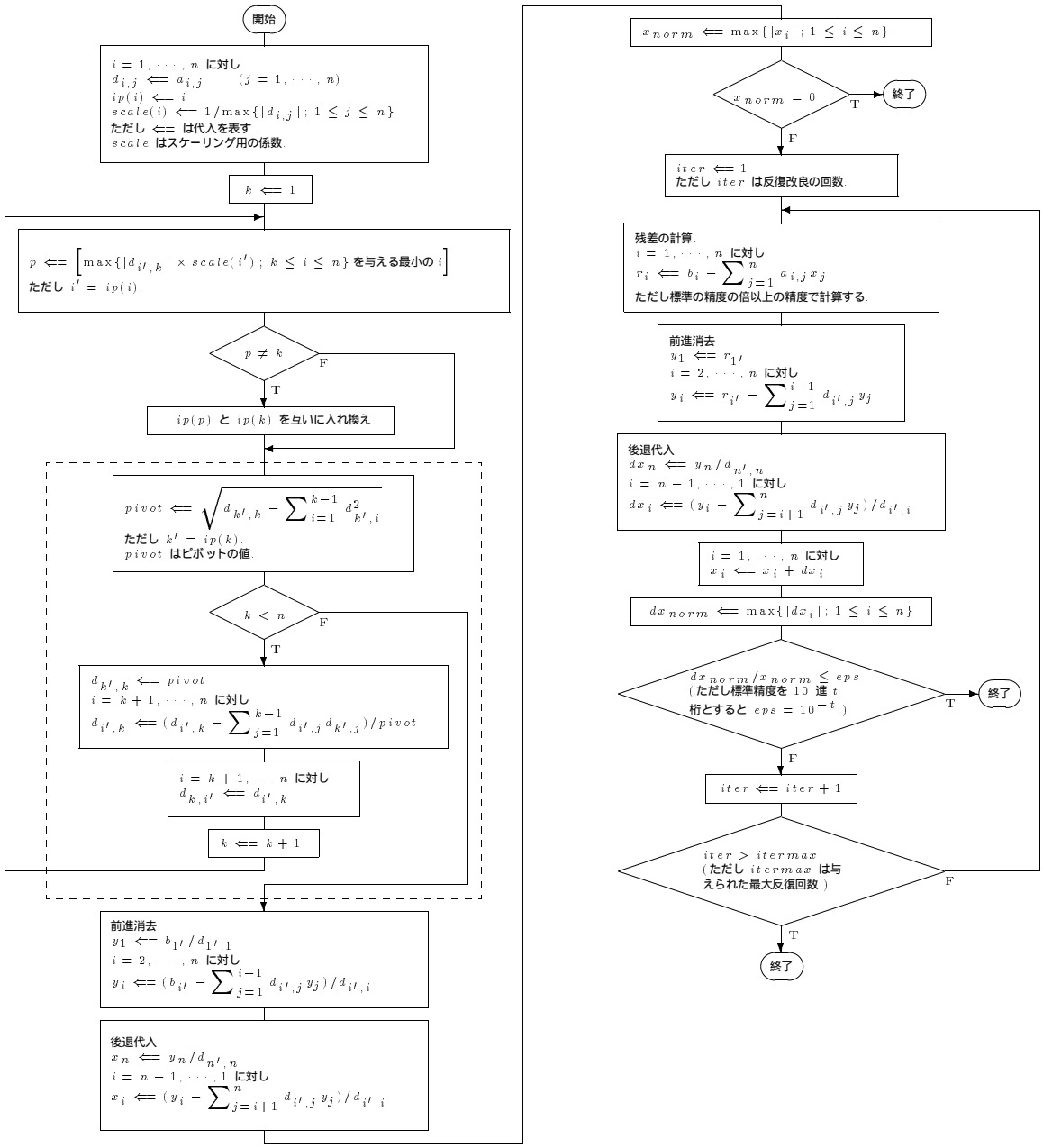
を条件数と言う.

$$1 = \|I\| = \|AA^{-1}\| \leq \|A\|\|A^{-1}\|$$

から条件数は 1 以上である.

式 (3.51) からわかるように, 条件数が 1 よりはるかに大きいと, b が少しずれただけで解の誤差は巨大なものになる可能性がある. このような場合に条件が悪い (ill-conditioned) という. 逆に条件数が 1 に近ければ, b の相対誤差と解の相対誤差は同程度の大きさである. このような場合を条件が良い (well-conditioned) という.

3.7.4　付録 4: コレスキー分解法のアルゴリズムのフローチャート



3.7.3　付録 3: コレスキー分解定理の証明

係数行列 A の次数 k に関する帰納法を用いる. すなわち A が $k \times k$ の正値対称行列である時にコレスキー分解定理が成り立つと仮定すれば $(k+1) \times (k+1)$ の正値対称行列 A' についてもコレスキー分解定理が成り立つことを示す.

$k = 1$ のときは, $L = 1$ として自明. 次数 $k+1$ の正値対称行列 A' の次数 k の主小行列を A とすると, A は正値対称行列である. A' は A とある k 次元列ベクトル $\boldsymbol{y}$, 実数 $a_{k+1,k+1}$ を用いて

$$A' = \begin{pmatrix} A & \boldsymbol{y} \\ \mathbf{y}^T & a_{k+1,k+1} \end{pmatrix} \tag{3.41}$$

と書ける. 帰納法の仮定より

$$A = LL^T$$

と分解されるが, このときに $(k+1) \times (k+1)$ の下三角行列 L'

$$L' = \begin{pmatrix} L & \mathbf{o} \\ \boldsymbol{l}'^T & l'_{k+1,k+1} \end{pmatrix} \tag{3.42}$$

が存在して

$$A' = L'L'^T \tag{3.43}$$

となることを示す. $\boldsymbol{l}'$, $l'_{k+1,k+1}$ はそれぞれ A' から算出されるベクトルと定数である. 式 (3.43) の両辺を比較すると

$$A = LL^T \tag{3.44}$$

$$\boldsymbol{y} = L\boldsymbol{l}' \tag{3.45}$$

$$\boldsymbol{y}^T = \boldsymbol{l}'^T L^T \tag{3.46}$$

$$a_{k+1,k+1} = l'^2_{k+1,k+1} + \boldsymbol{l}'^T \boldsymbol{l}' \tag{3.47}$$

を示せば良いことがわかるが, 式 (3.44) は帰納法の仮定であり, 式 (3.45) と (3.46) は同値な関係である. したがって式 (3.45) と式 (3.47) を示せばよい.

式 (3.45) については L と $\boldsymbol{y}$ が与えられているから前進代入によって $\boldsymbol{l}'$ を求められる.

式 (3.47) については,

$$l'_{k+1,k+1} = \sqrt{a_{k+1,k+1} - \boldsymbol{l}'^T \boldsymbol{l}'} \tag{3.48}$$

によって $l'_{k+1,k+1}$ が確定することはわかる. あとはこれが 0 でない実数であること, すなわち右辺の平方根内が正であることが示せれば良い. そのために

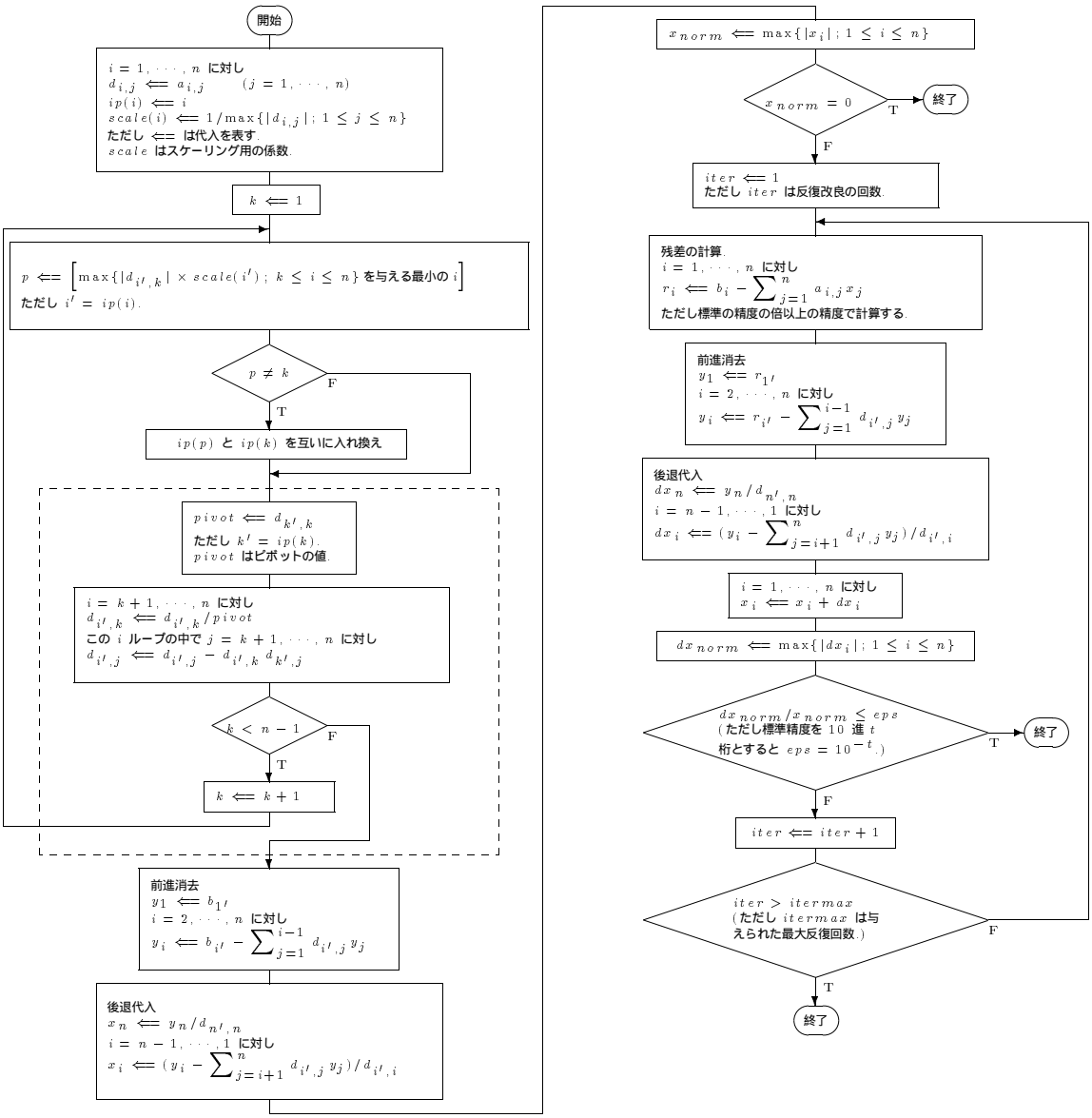
$$\boldsymbol{x} = \begin{pmatrix} A^{-1}\boldsymbol{y} \\ -1 \end{pmatrix}$$

とおいて $\boldsymbol{x}^T A' \boldsymbol{x}$ を計算すると, 式 (3.41) により

$$\begin{aligned} \boldsymbol{x}^T A' \boldsymbol{x} &= a_{k+1,k+1} - \boldsymbol{y}^T A^{-1} \boldsymbol{y} \\ &= a_{k+1,k+1} - \boldsymbol{y}^T (LL^T)^{-1} \boldsymbol{y} \\ &= a_{k+1,k+1} - (L^{-1}\boldsymbol{y})^T (L^{-1}\boldsymbol{y}) \\ &= a_{k+1,k+1} - \boldsymbol{l}'^T \boldsymbol{l}' \end{aligned}$$

となるが (最後の等式は式 (3.45) を使用), A' は正値であるから右辺も正である. よって $l'_{k+1,k+1}$ は 0 でない実数である. [証明終り]

3.7.2 付録 2: LU 分解法のアルゴリズムのフローチャート



となる. これは対角要素が 1 の下三角行列である.

以上により, 一つの並べ換え行列 P_k および対角要素がすべて 1 の下三角行列 M_k により

$$A_k = M_k P_k A \tag{3.40}$$

となることが帰納法の仮定から示された.

また, $k = 0$ の場合には, $A_0 = A$ であるから $M_0 = P_0 = I_n$ とすればよいから式 (3.40) の関係が成立する. これで補題が証明された. [補題の証明終り]

さて, 正則な下三角行列 M の逆行列 M^{-1} はやはり下三角行列であるから, $M^{-1} = L$ とおけば

$$PA = LU$$

となる. [定理の証明終り]

ただし I_{k-1} は $(k-1)\times(k-1)$ の単位行列である. 以上の 2 つの操作をまとめると,

$$A_k = M'_k P'_k A_{k-1} \tag{3.33}$$

となって A_k が得られる. P'_k が一対の行の交換を表すのに注意すると, $P'_k P'_k$ は単位行列になるので

$$P'_k M_{k-1} = P'_k M_{k-1} P'_k P'_k$$

となる. したがって式 (3.33) は

$$\begin{aligned} A_k &= M'_k P'_k A_{k-1} \\ &= M'_k P'_k (M_{k-1} P_{k-1} A) \qquad (\text{帰納法の仮定}) \\ &= M'_k (P'_k M_{k-1} P'_k)(P'_k P_{k-1}) A \end{aligned}$$

と変形できる. ここで

$$P_k \equiv P'_k P_{k-1}$$

とおくと, P_k は並べ換え行列の一種である.

次に

$$M_k \equiv M'_k (P'_k M_{k-1} P'_k) \tag{3.34}$$

と置き, 帰納法の仮定として

$$M_{k-1} = \begin{pmatrix} L_{k-1} & 0 \\ C & I_{n-k+1} \end{pmatrix} \tag{3.35}$$

であるとしたときに,

$$M_k = \begin{pmatrix} L_k & 0 \\ D & I_{n-k} \end{pmatrix} \tag{3.36}$$

が導かれることを示す. ただし, L_k は対角要素が 1 の $k\times k$ の下三角行列である.

さて, 式 (3.32) より

$$M'_k = \begin{pmatrix} I_k & 0 \\ B & I_{n-k} \end{pmatrix}$$

と書ける. 一方, P'_k が A_{k-1} の最初の $k-1$ 行に対しては影響しないので

$$P'_k = \begin{pmatrix} I_{k-1} & 0 \\ 0 & P''_k \end{pmatrix} \qquad (P''_k \text{は} (n-k+1)\times(n-k+1) \text{の並べ換え行列})$$

である. これと帰納法の仮定 (3.35) から

$$P'_k M_{k-1} P'_k = \begin{pmatrix} L_{k-1} & 0 \\ C' & I_{n-k+1} \end{pmatrix} \tag{3.37}$$

の形となる ($P''_k P''_k = I_{n-k+1}$ であることに注意!). 式 (3.37) は

$$P'_k M_{k-1} P'_k = \begin{pmatrix} L_k & 0 \\ C'' & I_{n-k} \end{pmatrix} \tag{3.38}$$

と書き換えられるので式 (3.34) は

$$M_k = \begin{pmatrix} I_k & 0 \\ B & I_{n-k} \end{pmatrix} \begin{pmatrix} L_k & 0 \\ C'' & I_{n-k} \end{pmatrix} = \begin{pmatrix} L_k & 0 \\ D & I_{n-k} \end{pmatrix} \tag{3.39}$$

3.7　付録

3.7.1　付録 1: LU 分解定理の証明

まず次の補題を証明する.

> 補題: $A = (a_{ij})\,(i, j = 1, \cdots, n)$ を $n \times n$ 正方行列とする. ピボットを 0 にしない任意の並べ換え行列 P に対し, ある上三角行列 U と, ある正則な下三角行列 M が存在して
>
> $$MPA = U$$
>
> とできる.

ただし並べ換え行列とは, 各行, 各列に 1 が 1 個しかなく, 他の要素がすべて 0 であるような $n \times n$ 正方行列である. 例えば

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

などである. この例の P は, 任意の行列 A の左から掛けると A の第 1 行と第 2 行を交換し, 右から掛けると第 1 列と第 2 列を交換する.

[補題の証明] 連立一次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ を考えよう. 前進消去で $x_1, \cdots, x_{k-1}$ までを消去し終った段階での係数行列を A_{k-1} と書くことにする (したがって $A_0 = A$ である). ある並べ換え行列を P_{k-1} とし, M_{k-1} を対角要素がすべて 1 のある下三角行列としたとき,

$$A_{k-1} = M_{k-1} P_{k-1} A$$

を仮定すれば

$$A_k = M_k P_k A$$

が導かれ, P_k がある並び換え行列, M_k が対角要素が 1 の下三角行列となることを示せば良い.

行列 A_{k-1} の i, j 要素を簡単のために $a_{i,j}$ と書く. すべての $i \geq k$ に対して $a_{i,k} = 0$ ならば, 前進消去をそれ以上進めることが出来ない. 今は前進消去が最後まで実行できる場合のみを考えているので, $j \geq k$ を満たすある j について $a_{j,k} \neq 0$ として, 第 j 行と第 k 行を交換する. これを並べ換え行列 P'_k で表す. 次に第 $k+1$ 行から第 n 行迄について l, m 成分の更新

$$a_{l,m} \longleftarrow a_{l,m} - \frac{a_{i,k}}{a_{k,k}} a_{k,m} \tag{3.31}$$

を, $l = k+1, \cdots, n$ と $m = k, \cdots, n$ について行なう. この操作は A_{k-1} の第 k 行を $-a_{l,k}/a_{k,k}$ 倍して第 l 行に加えるという操作であって, A_{k-1} の左から次の下三角行列 M'_k を掛けるのと同じである.

$$M'_k = \begin{pmatrix} I_{k-1} & & & & & \\ & 1 & & & 0 & \\ & -a_{k+1,k}/a_{k,k} & 1 & & & \\ 0 & -a_{k+2,k}/a_{k,k} & 0 & 1 & & \\ & \cdots & & & & \\ & -a_{n,k}/a_{k,k} & 0 & \cdots & 0 & 1 \end{pmatrix} \tag{3.32}$$

- 大気上端での各波長の上向きフラックス

j	λ_j (μm)	F_j (10^{-7} W/m^2)
1	15.0000000000000	4.87939724940278
2	14.2565082250148	6.15280600216324
3	13.4024604461355	8.04975586004209
4	12.4214171674480	10.1765040618791
5	11.2944943670388	7.09856042168303

- 温度分布 (答え)

i	p_i (atm)	Δp_i (atm)	T_i (K)
1	0.2	0.2	220
2	0.4	0.2	240
3	0.6	0.2	260
4	0.8	0.2	280
5	1.0	0.2	300

この問題では波長やフラックスのデータが 10 進 15 桁の範囲まで与えてあるが, もし興味があればこれらのデータの桁数を落していくと得られる温度がどのようにずれていくかを調べてみると良いかも知れない. また, 解の反復改良を行なわないとどうなるかをやってみると面白いと思う.

最後に少し補足しておく. ここでは連立一次方程式の解法のみを使って解けるようにするために波長のサンプル数を高度のサンプル数に合わせたが, 波長のサンプル数を高度のサンプル数よりも多くとれば, データ解析の章で述べられる最小自乗法を使ってもっと信頼性の高い結果を得ることが出来る.

大気放射学に関する参考書として参考文献リストに 2 冊挙げておいた. 会田, 1982 は日本語で書かれた大気放射学の入門書としては今のところ唯一のものであろう. Goody and Yung, 1989 は世界的な標準的教科書であり, 大気放射学に関してかなり多くのことを網羅してあるが, 通読するには困難であり, むしろ辞書的な使い方をしたほうが良い本である.

となる. ただし p_0 は地表面気圧である. 更に, 光学的厚さが圧力 p に比例するとすれば

$$\tau_\lambda(p) = k_\lambda p \tag{3.27}$$

である. 係数 k_λ は波長の関数である[4]. また, ここでは Planck 関数の波長依存性が無視できるような狭い波長域でのみ考えることにし, 簡単のために Planck 関数が温度 T に比例する, すなわち

$$B_\lambda(p) = \beta T(p)$$

であるとする. ただし β は波長に依存しない定数, $T(p)$ は p で指定される高度の絶対温度である. このとき式 (3.26) は

$$F_{t,\lambda} = k_\lambda \beta \pi \int_o^{p_0} e^{-k_\lambda p} T(p)\, dp \tag{3.28}$$

となる.

式 (3.28) を, 圧力と波数に関して離散化すると

$$F_j = k_j \beta \pi \sum_{i=1}^{N} e^{-k_j p_i} T_i\, \Delta p_i \tag{3.29}$$

となる. ただし T_i は第 i 層 (圧力 p_i) での温度であり, 波長サンプルは鉛直方向のサンプルと同じ N 個 $(\lambda_1, \cdots, \lambda_N)$ とることにする. また, $F_j = F_{t,\lambda_j}$, $k_j = k_{\lambda_j}$ である. このようにすると,

$$x_i = T_i$$

$$b_j = F_j$$

$$a_{i,j} = k_j \beta \pi e^{-k_j p_i}\, \Delta p_i$$

とおき,

$$\boldsymbol{x} = (x_1, \cdots, x_N)^T$$

$$\boldsymbol{b} = (b_1, \cdots, b_N)^T$$

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \cdots & \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix}$$

とすれば,

$$A\boldsymbol{x} = \boldsymbol{b} \tag{3.30}$$

という連立一次方程式になる. これを解けば各高度の温度を求めることが出来る.

ここでは k_j としては

$$k_j = \frac{c}{(\lambda_j - \lambda_0)^2}$$

という関数形を仮定する. c, λ_0 は定数である. 以下に解くのに必要なデータを与えておく. この例では鉛直と波長のサンプル数は 5 である.

- パラメータ:

N:	5
c (μm^2/atm):	10
β (W/m^2·K):	6.0×10^{-9}
λ_0(mu$m$):	10

[4] このような係数を吸収係数という. 吸収係数は, 吸収物質をどのような単位で測るかによって単位が異なる. また普通は吸収係数の値は温度や圧力にも依存する.

3.6 例題

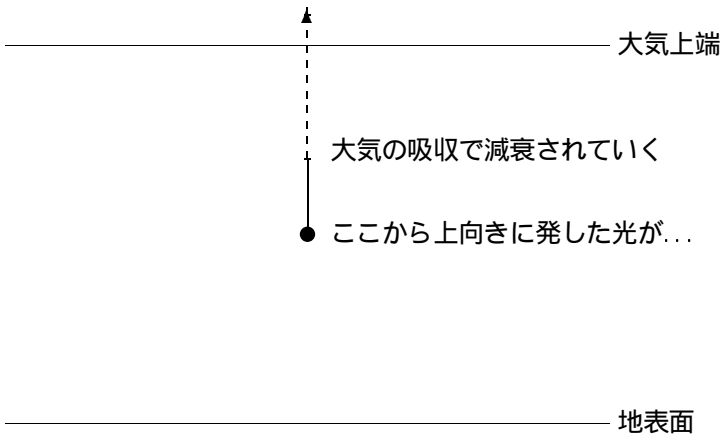
ここでは連立一次方程式の解法を使って, 大気上端の放射フラックスから大気の鉛直温度構造を計算する問題を考える.

問題

以下の解説に述べる設定のもとで大気の鉛直温度構造を求めよ.

解説

地球大気のうち約 40 km 以下の部分は局所的に熱力学平衡になっており, 黒体輻射に関する Planck の法則に従って温度に見合った光を放出している. 一方, 大気は外から来た光を吸収するので, 光が通過した大気の層が厚ければ厚いほど, 光は減衰する. この減衰の仕方は光の波長によって異なり, 同じ大気の層を通ってきても, ある波長の光はあまり減衰されないが, 別の波長の光はほとんど減衰されてしまうということが起こる.



大気が発した上向き赤外放射とその減衰の摸式図

光と大気のこのような関係を使うと, 大気上端からさまざまな波長の光を観測し, そのデータを使って大気の鉛直温度構造を推定することが出来る.

水平方向に一様な大気を考える. ここでは散乱は考えない. 波長 λ の光の上向きフラックス[2] を F_λ, 大気上端から測った光学的厚さ[3] を τ_λ, Planck 関数を B_λ とすると,

$$\frac{dF_\lambda}{d\tau_\lambda} = F_\lambda - \pi B_\lambda \tag{3.25}$$

となる. ここで地表面および大気圏外からの放射はないとして式 (3.25) を解いて大気上端での上向きフラックス $F_{t,\lambda}$ を求めると,

$$F_{t,\lambda} = -\pi \int_0^{p_0} B_\lambda(p) \frac{\partial e^{-\tau_\lambda}}{\partial p}\, dp \tag{3.26}$$

[2] 単位波長, 単位時間あたりに単位面積の水平面を通して上向きに流れる放射エネルギー. つまり上向きの放射フラックス.
[3] 放射フラックスが $1/e$ だけ減衰するような物質の量を単位として測った, 光路上の吸収物質の量. 波長の関数である.

3.4.1　LU 分解法のアルゴリズム

今までに述べたことをまとめると, ガウスの消去法で連立一次方程式が解ける場合には, 次の手順で解くことが出来るということになる.

1. スケーリング
2. 部分的ピボッティング
3. 係数行列 A の LU 分解
4. $U\boldsymbol{y} = \boldsymbol{b}$ から前進代入で $\boldsymbol{y}$ を求める. (前進消去)
5. $L\boldsymbol{x} = \boldsymbol{y}$ から後退代入で $\boldsymbol{x}$ を求める.
6. 解の反復改良

ただし前の脚注で述べたように, 実際の計算では部分的ピボッティングのための行交換を LU 分解の前にまとめることは困難であるから, 部分的ピボッティングのための行交換の手続き 2. と LU 分解の手続き 3. は同時進行の形となる.

これをもう少し具体的なフローチャートにすると 3.7.2 節のようになる. 本質的な部分はこのフローチャートの左半分である. 点線の四角の中が係数行列 A の LU 分解. L の対角成分は 1 になるように分解しているので, L と U は $d_{i,j}$ という一つの行列で表現している. $d_{i,j}$ のうち $i > j$ の部分が L の i, j 成分, $i \leq j$ の部分が U の i, j 成分である. 点線の四角より前の手続きがスケーリングと部分的ピボッティング, 点線の四角よりあとの部分が前進消去と後退代入である. フローチャートの右半分は解の反復改良の手続きである.

LU 分解法のプログラムについては, 津田, 1988; 森, 1988 あるいは小国, 1991 に FORTRAN77 のプログラムソースコードが載っているのでそちらを参考にして欲しい.

3.5　コレスキー分解法

係数行列 A が正値対称行列の時, すなわちゼロベクトルでない任意のベクトル $\boldsymbol{x}$ について

$$\boldsymbol{x}^T A\boldsymbol{x} > 0$$

のときには, 以下に示すコレスキー (**Cholesky**) 分解定理によって

$$A = LL^T \tag{3.24}$$

のように下三角行列, とその転置行列の積に分解でき, LU 分解法の親戚であるコレスキー分解法という方法で LU 分解法の約半分の演算時間で連立一次方程式が解ける. ただし $\boldsymbol{x}^T$, L^T は $\boldsymbol{x}$, L の転置行列である.

コレスキー分解法の根拠となるコレスキー分解定理とは以下のようなものである.

> コレスキー分解定理:$n \times n$ の正値対称行列 A を, 下三角行列 L によって式 (3.24) のように分解できる. ただし L は一意に定まり, 対角成分 $l_{i,i}$ は必ずしも 1 ではない.

この定理の証明は 3.7.3 節に示す.

コレスキー分解法のフローチャートを 3.7.4 節に示す. LU 分解法のフローチャートと異なるのは, 主に係数行列 A を LL^T に分解する手続きである点線の四角の部分である. あとは L の対角成分が 1 でないことに対応して前進消去が多少異なる.

を解く. もし x_2 が誤差を持たなければ$x_1 + x_2$ が正確な解である. しかし実際には x_1 と同様にx_2 も 10 進 3 桁の精度しか持たない. この時点で x_1, x_2 の関係は以下のようになっている.

$$\begin{array}{lcccccccccc} x_1 \text{ の } 1 \text{ 成分}: & 0. & \heartsuit & \heartsuit & \heartsuit & \spadesuit & \spadesuit & \spadesuit & \spadesuit & \spadesuit & \times 10^n \\ x_2 \text{ の同成分}: & 0. & 0 & 0 & 0 & \heartsuit & \heartsuit & \heartsuit & \spadesuit & \spadesuit & \times 10^n \end{array}$$

♡ はそこに正確な数値が入っていることを示し, ♠ はその数字がランダムな誤差によるものであることを示す. ただし x_2 の 「正確な数字」の部分は, x_1 の ♠ のうちの上 3 桁を修正する意味で正確な数字であることに注意. したがって $x_1 + x_2$ は 10 進 6 桁の精度を持っている.

この手順をもう一回繰り返せば 10 進 8 桁の精度の解を得ることが出来る. このような繰り返しの手続きを解の反復改良という.

3.4　LU 分解法

前節までに話したスケーリングおよび部分的ピボッティングつきガウス消去法によって, (場合によっては解の反復改良を施せば) 連立一次方程式を必要な精度で数値的に解くことが出来る.

ここではスケーリングと, 部分的ピボッティングの行交換を施した後のことを考えよう[1]. この時にはすぐにガウスの消去法に移れるわけだが, ガウスの消去法の 2 大手順である前進消去と後退代入を係数行列の分解で表現する次のような定理があるのでこれを利用する.

> LU 分解定理: ピボットを 0 にしない任意の並べ換え行列 P に対し,
>
> $$PA = LU \tag{3.21}$$
>
> と一意に分解できる. ただし L は対角成分が全て 1 の下三角行列, U は上三角行列である.

この定理の証明を 3.7.1 節に示す.　「並べ換え行列」 P とは部分的ピボッティングに必要な行交換の手続きを行列で表現したもの. 詳しくは 3.7.1 節を参照のこと. 行列をこのように分解することを LU 分解という. PA がこのように LU 分解できると次のようにして連立一次方程式を解くことが出来る. ただし部分的ピボッティングのための行交換はすべて終了しているとし, その並べ換え行列 P と係数行列 A の積 PA をあらためて A と書くことにする.

まず連立一次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ を A の LU 分解に即して

$$U\boldsymbol{x} = \boldsymbol{y} \tag{3.22}$$

$$L\boldsymbol{y} = \boldsymbol{b} \tag{3.23}$$

のように分解する. 式 (3.22) は前進消去が終了した状態に対応しており, $\boldsymbol{y}$ は同じ前進消去の手続きによって $\boldsymbol{b}$ から得られるものである. 次に式 (3.23) から前進代入 (**forward substitution**) によって $\boldsymbol{y} = (y_1, \cdots y_n)^T$ を y_1, y_2, $\cdots$, y_n の順に求める. 最後にその $\boldsymbol{y}$ と式 (3.22) から後退代入によって $\boldsymbol{x} = (x_1, \cdots x_n)^T$ を x_n, x_{n-1}, $\cdots$, x_1 の順に求めれば良い. このようにガウス消去法の 2 大手順の前進消去と後退代入は係数行列 A を下三角行列 L と上三角行列 U に分解することで表現できる.

[1] x_k を消去するために行なう部分的ピボッティングの行交換の手続きと, x_1 から x_{k-1} までの消去の手続きとは交換可能であるから, 変数の消去の手続きの前に行交換の手続きすべてを持って来ることが出来る. ただし実際の計算ではこのような手続きの交換を行なうことは大変難しい. なぜなら交換すべき行を具体的に特定しなければならないが, それは各消去の段階で $|a_{i,k}^{(k)}|\,(i = k, \cdots, n)$ の最大値を計算してからでないと特定できないからである.

$$x + y = 2 \tag{3.12}$$

その際に 1.23×10^4 のように 10 進法で仮数部 3 桁の浮動小数点の数体系で扱って, 部分的ピボッティングつきガウスの消去法を行なうことにする. そうすると行交換は起こらずに

$$2.00x + 2.00 \times 10^4 y = 20000$$

$$-1.00 \times 10^4 y = -1.00 \times 10^4 \tag{3.13}$$

となる. ここで式 (3.13) に注目して欲しい. 例えば y の係数は, $1.00 - 1.00 \times 10^4$ となるので, 10 進 3 桁の精度では -9999 ではなくて -1.00×10^4 となるのである. 同様にして右辺は -9998 ではなくて -1.00×10^4 となる. したがってこのまま解くと

$$x = 0.00, \qquad y = 1.00 \tag{3.14}$$

となって, 厳密解

$$x = \frac{10000}{9999}, \qquad y = \frac{9998}{9999} \tag{3.15}$$

とは似ても似つかない解になってしまう.

このようなことを (完全ではないけれども) 防ぐ手段として, 係数行列 A の各行と対応する右辺 $\boldsymbol{b}$ の行 (すなわち拡大行列 $\tilde{A}$ の各行) のうちで絶対値最大の要素で各要素を割ってから部分的ピボッティングつきガウス消去法を行なうという方法がある. このような処理をスケーリングという. 今の例でいけば, 式 (3.11) の両辺を 20000 で割って

$$1.0 \times 10^{-4} x + 1.00 y = 1.00 \tag{3.16}$$

$$1.00x + 1.00y = 2.00 \tag{3.17}$$

を部分的ピボッティングつきガウス消去法で解けば良いことになる. 部分的ピボッティングによって, 式 (3.17) を 1.0×10^{-4} 倍して式 (3.16) から引くことによって式 (3.16) の x を消去する. すると

$$1.00x + 1.00y = 2.00$$

$$1.00y = 1.00$$

これを後退代入で解けば

$$x = 1.00, \qquad y = 1.00 \tag{3.18}$$

となって, 式 (3.15) で示した厳密解に 10 進 3 桁の精度で一致することになる.

3.3.3　解の反復改良

いま標準の精度が 10 進 8 桁とし, 連立一次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ を解いた結果得られた解 $\boldsymbol{x}_1$ が 10 進 3 桁しか正しくなく, 残りの 5 桁はランダムな誤差による数値であるとする. このようなことは, 行列 A が 3.7.5 節に示すような 「条件が悪い行列」の場合に起こる. 「条件が悪い行列」とは, 簡単に言うと, 「解を高精度に求めるのに都合の悪い行列」ということである.

ここで

$$\boldsymbol{r}_1 \equiv \boldsymbol{b} - A\boldsymbol{x}_1 \tag{3.19}$$

を標準の 2 倍以上の精度で求め, $\boldsymbol{r}_1$ は誤差を持たないとする. $\boldsymbol{x}_1$ が誤差を持たない解であれば, 当然 $\boldsymbol{r}_1$ は 0 である. このとき標準の精度で

$$A\boldsymbol{x}_2 = \boldsymbol{r}_1 \tag{3.20}$$

と書き, 前進消去を始めて x_1 から x_k まで消去したときの拡大行列を $\tilde{A}^{(k)} = (a_{i,j}^{(k)})$ と書くと, x_{k+1} を消去する手順は, $\tilde{A}^{(k)}$ から $\tilde{A}^{(k+1)}$ を求める手順に相当する. このとき, 行の添字 i について $i = 1$ から $i = k$ までは

$$[\tilde{A}^{(k+1)}\text{の第 } i \text{ 行}] = [\tilde{A}^{(k)}\text{の第 } i \text{ 行}] \tag{3.9}$$

$i = k + 1$ から $i = n$ までは

$$[\tilde{A}^{(k+1)}\text{の第 } i \text{ 行}] = [\tilde{A}^{(k)}\text{の第 } i \text{ 行}] - \frac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}} \times [\tilde{A}^{(k)}\text{の第 } k\text{行}] \tag{3.10}$$

となる. この操作を $\tilde{A}^{(n)}$ を求めるところまで帰納的に続ければ, 前進消去が完了する. あとは後退代入によって x_n から順に x_1 までを計算すれば良い.

3.3　単純なガウスの消去法の問題点とその改善法

実際には単純なガウスの消去法ではまずい点がいくつかある. ここではこれらの問題点とその解決法について述べる.

3.3.1　ピボットの選択

式 (3.10) を見てもらいたい. この式に分母として現れる $a_{k,k}^{(k)}$ をピボット (**pivot**) というが, これが正確に 0 であれば前述の消去法はその時点で実行不可能となる. また 0 でなくても絶対値が非常に小さいと, 数値的に解を求めた時には解の誤差が非常に大きくなる. したがってピボットは出来る限り絶対値が大きなものを選択する必要がある.

そこで x_1 から x_k までを消去した段階の拡大行列$\tilde{A}^{(k)}$ の k 列目の下側 $(n-k+1)$ 個の要素$a_{k,k}^{(k)}, a_{k+1,k}^{(k)}, \cdots, a_{n,k}^{(k)}$ の絶対値を比較し, 絶対値最大のものが $a_{p,k}^{(k)}$ であれば第 p 行と第 k 行を入れ換えてからガウス消去法の手順を続行することを考える. このような方法を部分的ピボッティング (**partial pivoting**) という.

$$\tilde{A}^{(k)} = \begin{pmatrix} \cdots & \vdots & \cdots \\ & a_{k,k}^{(k)} & \\ \cdots & \vdots & \cdots \\ & a_{n,k}^{(k)} & \end{pmatrix}$$

図 3.3.1: $\tilde{A}^{(k)}$ の第 k 列目の下側 $(n-k+1)$ 個の要素

今のような縦方向の走査のみならず, 横方向も走査して全ての $a_{i,j}^{(k)}$ $(k \le i \le n, k \le j \le n)$ の中から絶対値最大のものを見つけ, 行のみならず列の入れ換えも行なう方法を完全ピボッティング (**complete pivoting**) というが, 完全ピボッティングは部分的ピボッティングよりも計算時間がはるかにかかるうえに, 実用上は部分的ピボッティングで十分であることがわかっているので普通は部分的ピボッティングを採用する.

3.3.2　スケーリング

次の連立一次方程式を解くことを考える.

$$2x + 20000y = 20000 \tag{3.11}$$

という式が得られる. この式 (3.5) から y を消せば良い. これも今までと同じ要領で

$$[\text{式 } (3.4)] \times (-3) + [\text{式 } (3.5)]$$

とすると

$$7z = 7 \tag{3.6}$$

として得られる.

したがって, 連立一次方程式 (3.1), (3.2), (3.3) と同値な方程式 (3.1), (3.4), (3.6)

$$x + y + z = 4$$

$$-6y - 2z = -8$$

$$7z = 7$$

を解けば良いことになる. 元の 3 本の方程式のうち上のものから順々に変数を消去してきたので, ここまでの手段を前進消去 (**forward elimination**) という. あとは式 (3.6) から $z = 1$, これを式 (3.4) に代入して $y = 1$, 最後に式 (3.1) より $x = 2$ として解を求めることが出来る. このように前進消去で得られた連立一次方程式の最後の式から順々に値を代入していって解を求める手順を後退代入 (**back substitution**) という. このようにして変数を消去しながら連立一次方程式を解く方法をガウスの消去法という.

3.2.2　一般化すると···

前述のガウスの消去法を一般化して書いてみよう. 連立一次方程式は $n \times n$ の係数行列 A と n 次元ベクトル $\boldsymbol{x}$, $\boldsymbol{b}$ を使って

$$A\boldsymbol{x} = \boldsymbol{b} \tag{3.7}$$

と書ける. 求めたいのは $\boldsymbol{x}$ である. 係数行列 A とベクトル $\boldsymbol{b}$ から作られる $n \times (n+1)$ の拡大行列 (**augmented matrix**)

$$\tilde{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & a_{1,n+1} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & a_{2,n+1} \\ & \cdots & \cdots & & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} & a_{n,n+1} \end{pmatrix} \tag{3.8}$$

を用いると, ガウスの消去法は拡大行列$\tilde{A}$ の行に対する操作で実現できる. ただし

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ & \cdots & \cdots & \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} a_{1,n+1} \\ a_{2,n+1} \\ \cdots \\ a_{n,n+1} \end{pmatrix}$$

である.

さて,

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

第 3 章

行列計算その 1 : 連立 1 次方程式

3.1 Introduction

連立一次方程式の解法として有名なものにガウスの消去法がある. ここではガウスの消去法を数値解法に適する形に改良した方法である LU 分解法について述べる. まずは本家本元のガウスの消去法について述べ, 次に数値解を求める際に問題となる点を挙げ, これらの問題点を改良した 「改良版ガウス消去法」として LU 分解法を紹介する. 更に, 連立一次方程式の係数行列が正値対称行列のときのみ適用できるコレスキー分解法についても述べる. コレスキー分解法は LU 分解法よりも約半分の演算時間で解を求めることが出来る. また, 例題として簡単なモデル大気の鉛直温度構造を決める問題を載せてある.

3.2 ガウスの消去法

3.2.1 まずは例から···

次のような連立一次方程式を考える.

$$x + y + z = 4 \tag{3.1}$$

$$2x - z = 3 \tag{3.2}$$

$$4x - 2y + 2z = 8 \tag{3.3}$$

まず式 (3.2) から x を消そう. そのために

$$[\text{式 } (3.1)] \times (-2) + [\text{式 } (3.2)]$$

とする. その結果

$$-2y - 3z = -5 \tag{3.4}$$

という x のない式が得られる.

次に式 (3.3) から x, y の二つを消す. そのためにまず式 (3.3) から x を消すことを考える. これは式 (3.2) から x を消した時と同じ要領で

$$[\text{式 } (3.1)] \times (-4) + [\text{式 } (3.3)]$$

とすれば

$$-6y - 2z = -8 \tag{3.5}$$

問題 5

連立方程式

$$\begin{cases} x^2 + 2\sqrt{3}xy - y^2 = 4\sin\dfrac{(\sqrt{3}-1)x + (\sqrt{3}+1)y}{2\sqrt{2}} \\ x^2 - 2\sqrt{3}xy + 3y^2 = 4\ln(3x^2 + 2\sqrt{3}xy - 8\sqrt{3} + y^2 - 8y + 40) - 4\ln 2 \end{cases}$$

の解を二変数二分法を用いて求めよ. 因みに正解は $(x, y) = (-0.35, -2.39)$ である.

2.6　参考になる文献

本章の記述は杉江, 1986 から多く引用した. 地球物理学への応用例の典型としては, Gierasch and Toon, 1973 に記述されている火星極冠のエネルギーバランス方程式を解く過程がちょうどこれにあたる.

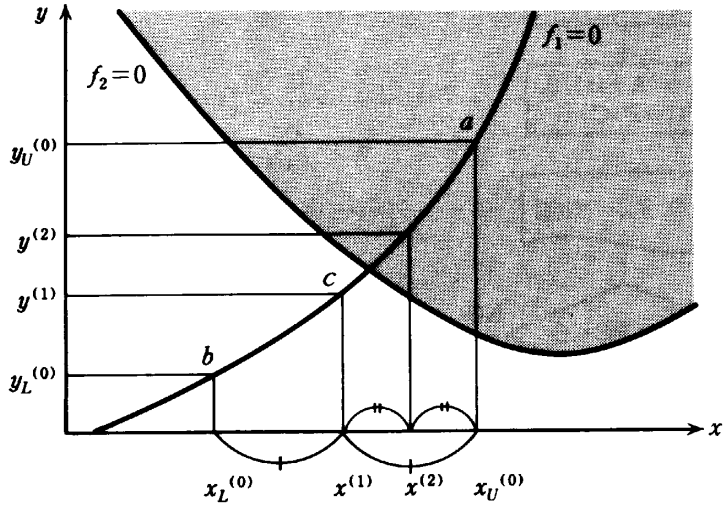


図 2.5: 二変数二分法

ごとに

$$x_n, \quad f(x_n), \quad f'(x_n), \quad \Delta x_n = -\frac{f(x_n)}{f'(x_n)}$$

の値の変化を調べよ.

問題 2

方程式

$$x \ln x - 1 = 0$$

の解を, Newton-Raphson 法 ・Regula-Falsi 法 ・二分法を用いて求め, 収束の速さと解の精度について検討を加えよ.

問題 3

問題 1 に於いて区間 $[0, 1]$ の解を Regula-Falsi 法と二分法で求めよ. 但し$\varepsilon = 10^{-5}$とする. また, 解の収束の様子を見るために, 反復ごとに $c, f(c)$ の値の変化を調べよ.

問題 4

連立方程式

$$\begin{cases} y = \sqrt{\ln(x^2 - 4x + 10) + \ln 2} \\ x^2 - y^2 = 2 \sin \dfrac{1}{\sqrt{2}}(x + y) \end{cases}$$

の解を x の区間 $[0, 3]$ で二変数二分法を用いて求めよ. この場合, 第一式を $f_1 = 0$, 第二式を $f_2 = 0$ として考えると良い.

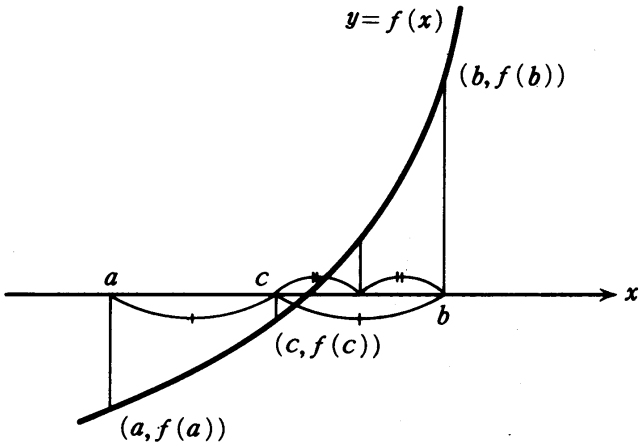


図 2.4: 二分法

2.5　二変数二分法

今度は, 複素根を持つ方程式や連立方程式を解く場合を考える. 関数 $f(x)=0$ の複素解を求めるとき, 根を$\alpha+i\beta$とすると

$$f(\alpha+i\beta)=f_1(\alpha,\beta)+if_2(\alpha,\beta)=0$$

と書き替えられるならば, $f(x)=0$ を解くことは連立方程式

$$\begin{cases} f_1(\alpha,\beta)=0 \\ f_2(\alpha,\beta)=0 \end{cases}$$

を解くことに帰着される. ここでは, 一般的な連立方程式

$$\begin{cases} f_1(x,y)=0 \\ f_2(x,y)=0 \end{cases} \tag{2.2}$$

の実数解を二分法を用いて求める方法について述べる.

式 (2.2) の二式のうちで, 方程式 $f(x,y)=0$ から x に値を与えて数値解法で y について解きやすい方の関数を f_1 と定義する. まず図 2.5に示したように, 解をはさんで x の上限 ${x_U}^{(0)}$ および下限 ${x_L}^{(0)}$ を与えて, 曲線 $f_1=0$ 上に点 $a=({x_U}^{(0)},{y_U}^{(0)})$ と点 $b=({x_L}^{(0)},{y_L}^{(0)})$ を求める. いま, 曲線 $f_2=0$ を境界線として xy平面を二つの領域に分けたとき, 点 a と点 b は互いに異なる領域に存在する. 関数 f_2の値は境界線上で 0, そして互いに他の領域では異なる符合を持つ. 従って

$$f_2({x_U}^{(0)},{y_U}^{(0)})f_2({x_L}^{(0)},{y_L}^{(0)})<0$$

となる. 次に x 軸上で ${x_U}^{(0)}$ と ${x_L}^{(0)}$の中点 $x^{(1)}$を求め, 曲線 $f_1=0$ 上に点 $c=(x^{(1)},y^{(1)})$ を求める. そして点 c における $f_2(x^{(1)},y^{(1)})$ の符合を調べて, 点 c がいずれの領域に存在するかを知り, x 軸上の解の存在範囲を $1/2$ に縮小する. 以下同様の操作を繰り返し, 所定の精度内に収束させ, 正しい解を得る.

問題 1

$f(x)=e^x-3x=0$ の解を Newton-Raphson 法で求めよ. ただし, $x_0=0$ とし, 収束判定条件としてふたつの近似解間の相対差を取り, $\varepsilon=10^{-5}$ とせよ. また, 解の収束の様子を見るために, 式 (2.1) による反復

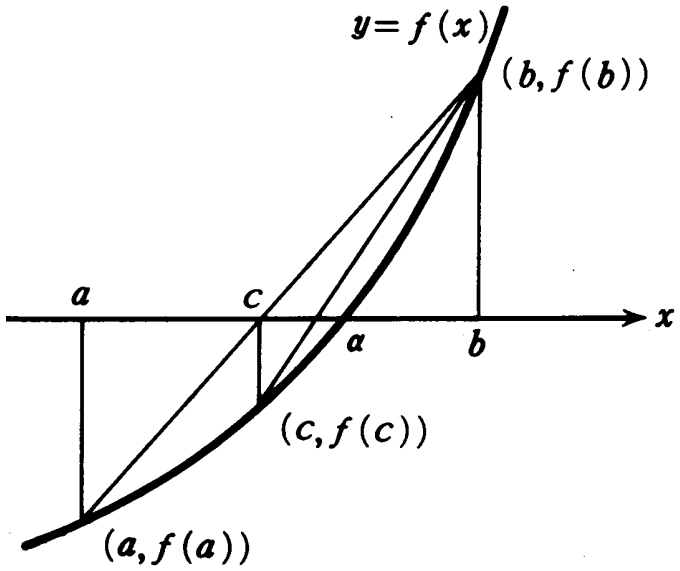


図 2.2: Regula-Falsi 法

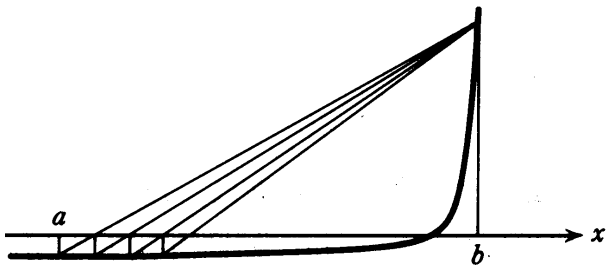


図 2.3: Regula-Falsi 法の悪い例

そして,

$$f(a)f(c) < 0 \quad \text{ならば} \quad b \leftarrow c,\ f(b) \leftarrow f(c)$$

$$f(b)f(c) < 0 \quad \text{ならば} \quad a \leftarrow c,\ f(a) \leftarrow f(c)$$

の変換を行ない, Regula-Falsi 法と同じ操作を繰り返して c を順次求め, 解を得る.

収束判定条件としては

1. $|b-a| < \varepsilon$
2. $|f(c)| < \varepsilon$

がある. この方法の特徴は, 例えば判定条件 1 の場合, 収束解を得るのに必要な反復計算回数 n が初期値 a, b を用いて

$$\frac{|b-a|}{2^n} < \varepsilon$$

という式から計算でき, あらかじめ n の値を知ることができる点である. したがって, n 回の反復計算を実行しても収束解が得られないのであればプログラムに誤りがあると判定でき, 計算機資源の無駄使いが避けられるのである.

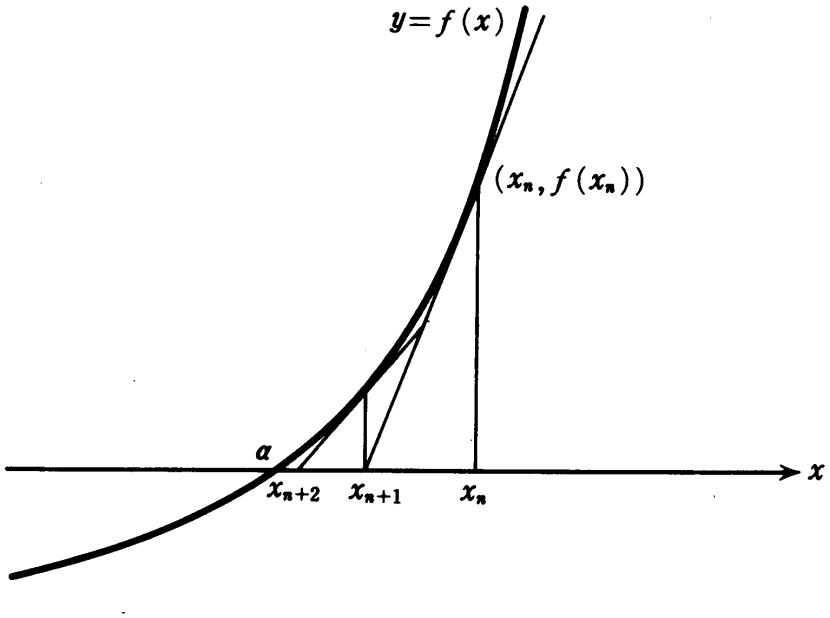


図 2.1: Newton-Raphson 法

2.3 Regula-Falsi 法

この方法ははさみうち法とも呼ばれる. 関数 $y = f(x)$ が区間 $[a, b]$ で連続で, 解がひとつ存在するとき, $f(a), f(b)$ は互いに符合が異なる, すなわち $f(a)f(b) < 0$ となる. 従って $(a, f(a))$ と $(b, f(b))$ の二点を直線で結び, この直線と x 軸との交点を c とすると, 次の式のように表せる.

$$c = \frac{af(b) - bf(a)}{f(b) - f(a)}$$

次に $f(c)$ の値を求め,

$$f(a)f(c) < 0 \quad \text{ならば} \quad b \leftarrow c,\ f(b) \leftarrow f(c)$$

$$f(b)f(c) < 0 \quad \text{ならば} \quad a \leftarrow c,\ f(a) \leftarrow f(c)$$

の変換を行ない, 解の存在区間を狭める. この操作を繰り返して c を順次求め, 解を得る. 収束の判定条件としては次のものがある.

1. $(c - a)(b - c) < \varepsilon$
2. $|f(c)| < \varepsilon$

この方法は区間 $[a, b]$ で関数 $f(x)$ が直線に近いものであればあるほど収束が速い. しかし, 図 2.3のような場合には計算回数を多く取る必要があるので, 次の二分法を用いた方が良い.

2.4 二分法

前節の Regula-Falsi 法では, 二点 $(a, f(a)), (b, f(b))$ を直線で結んで x 軸上に点 $(c, 0)$ を求めたが, 二分法では式 (2.3) の代わりに次の式を用いる.

$$c = \frac{a + b}{2}$$

第 2 章

高次代数方程式の数値解法

2.1 はじめに

3 次や 4 次の方程式までは Cardano 法や Ferrari 法などの直接解法がある. しかし, 5 次以上の方程式や非線形方程式では一般に直接的解法がないので, 数値解法を用いなければならない. 方程式の解を求める数値解法の代表的なものとしては,

1. Newton-Raphson 法
2. Regula-Falsi 法
3. 二分法

などがある. ここでは, これら三つの解法について簡単に述べる.

2.2 Newton-Raphson 法

この方法は関数 $f(x)$ が単調連続で変曲点がなく, かつ $f(x)$ の導関数 $f'(x)$ が求められる時に利用できる収束の速い方法である. 解の適当な初期値から始めて, 反復公式

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (n = 0, 1, \cdots) \tag{2.1}$$

を繰り返す. この方法の幾何学的意味は図 2.1に示す通りで, 曲線上の点 $(x_n, f(x_n))$ に於ける接線と x 軸との交点の座標を $(x_{n+1}, 0)$ とすると,

$$f'(x_n) = -\frac{f(x_n)}{x_{n+1} - x_n}$$

すなわち式 (2.1) を得る. 収束判定の条件としては, 与えられた ε に対し, 次の数種類のものがある.

1. ふたつの近似解間の誤差 : $|x_{n+1} - x_n| < \varepsilon$
2. ふたつの近似解間の相対差 : $\left|\dfrac{x_{n+1} - x_n}{x_n}\right| < \varepsilon$
3. ふたつの近似解関数値間の差 : $|f(x_{n+1}) - f(x_n)| < \varepsilon$

$f'(x)$ の値が大きいと, $|x_{n+1} - \alpha| > \varepsilon$にも関わらずの 1 と 2 の条件が満足されてしまう場合がある. しかし, 3 を収束の判定条件として用いると誤った解を出す危険性がない.

決まっているのであり, 修正すれば良い. しかし, 数値計算の結果をチェックする方法は自分自身の経験による他はない. 解析的な解がわからないから計算するのであり, したがって結果をあらかじめ知っていることはないのである. 計算精度の善し悪しの判断は, 計算結果に関する偏見をどれだけたくさん持っているかにかかっているが, また, 同時に計算がうまくいってない時には何が起こるかという経験にも支えられる. かくして経験を積むことにより, 4 年生になってから, あるいは進学・就職してから使うことになるであろう「できあいのソフトウェアパッケージ」がどんなによくできているか, 劣悪であるかが理解されるようになるものと思われる. 本テキストは以上のような基礎的経験を積むための技術教本であり, したがって, 地球惑星物理学的な問題の抽出, 定式化, 考察に関してはあまり語ってはいない. これらは各担当教官の「味付け」に任されている.

各題目は各々別の教官が担当する. したがって, 学生諸君に対する教官は計 5 人である. 各教官は, その担当の期間の適当な時期にそれぞれの得意技に応じて, 地球惑星物理学の諸分野での問題の抽出・定式化・処理の例を示し, あるいは, 計算・情報処理技術に関しての補足的解説を行なう. 解説の行なわれる日時, 集合場所等々は各教官から指示されるので注意されたし.

単位取得のために学生諸君に要請される義務は次の通りである.

- 上の 5 つの題目のうち, 最初の一つを除く計四つの題目に対して, それぞれ自分で適当に問題を設定し, 修練を積み, その内容と結果をレポートとして作成する. ただし, 題目 2 に対するレポートは純粋に理論的なもの (計算技術の修行を含まない地球惑星物理学的議論, 数理物理的考察, レビュー) であって良いものとし, 題目 3 ~ 5 (行列計算, 時系列処理, 時間発展方程式の積分) に対してはそれぞれの計算の理論または技術の実際の修練を何らかの形で含むような問題を作ることとする.
- レポートは A4 サイズの紙で図表を含め 15 枚以内とする. レポートのテーマには, 理論的背景の勉強の成果, 地球惑星物理学の定式化の説明, 計算技術の詳細, 自分の作ったサブルーチンリファレンス, 等々何を設定しても良い. ただし, 何について書かれているかが明解に第三者に理解されるように科学技術文章の基本作法をよく研究すること. 学生諸君個々人が選んだ問題とその考察は, もはや学生諸君個々人が専門家であり, 第三者にはまるで未知なるものである. それをどうやって第三者に伝えるかをよく考えてもらいたい.
- 計 4 冊のレポートは, 各題目担当教官の指定する期日までに当該教官のところへ提出し面接試問 (レポートの書き方, 問題の定式化と処理, 等々の作法の指導) を受ける. あらかじめ担当教官に連絡し訪問時刻等の予約をすること.

成績評価は機械的に行なわれる. 4 つ以上提出すれば優, 3 つならば良, 2 つならば可, それ以下の場合には不可である.

なお, 担当教官が研究集会, あるいは, 長期出張等で不在のこともあるので, レポートの作成はなるべく早めに行なうように. 〆切間際で集中しても面接試問のための時間がとれないことになり授業料を損するので気を付けられたい. 常識的な判断を期待する.

第 1 章

はじめに

地球惑星物理学演習の主たる目的は

> 地球惑星物理学の諸分野において, 問題を抽出し, あるいは, 処理する際に必要となる計算・情報処理の基礎技術を習得すること

にある. 本演習ではさらに

> 地球惑星物理学の諸分野における問題の抽出, 数理物理的定式化, その処理, 等々の例に触れること

をも目指している.

演習全体の構成はおおむね次のようになっている:

1. 情報環境入門 (第 1 ~ 2 週, テキスト 「計算情報環境入門」)
   UNIX / X window / Internet 環境 に習熟し, 電子メール, ニュースのやりとり, プログラムの作成と実行, ドキュメンテーション等々ができるようになる. 練習問題として方程式の解を求めることを自習する (テキスト第 2 分冊 第 2 章).

2. 数理物理的定式化と考察 (第 3 ~ 5 週, テキストにはない)
   地球惑星物理の諸問題の数理物理的定式化と考察の例題. いわゆる理論演習.

3. 行列を扱う (第 6 ~ 8 週, テキスト 「数値計算入門」第 3 ~ 4 章)
   連立 1 次方程式の解法, 固有値・固有ベクトル問題の解法の理論的背景を学ぶ.

4. データを処理する (第 9 ~ 11 週, テキスト 「数値計算入門」第 5 ~ 6 章)
   主に時系列データの処理を念頭において, 最小自乗法, フーリエ変換等々の時系列処理の基本技術の理論的背景を学ぶ.

5. 時間発展方程式を解く (第 12 ~ 14 週, テキスト 「数値計算入門」第 7 章)
   ルンゲクッタをはじめとする常微分方程式の時間積分スキーム, 流体運動等の移流方程式の時間積分で基本となる CFL 条件などを学び, 適当な系を設定して時間積分を行なう.

以上の題目において, 数値計算する際には, できあいの数値計算ソフトウェアパッケージに頼らずに実際にプログラムをつくって解くことが要請される. 計算・情報環境に習熟することを目指しているので, 特に, 計算プログラムの作成においてはなるべく苦労していただくようにしている. 図や文章を書くソフトウェアに関しては結果の善し悪しがその結果を見ることにより直ちに確認できる. 図がハミ出ておれば悪いに

7.12.1 Introduction . . . 119
7.12.2 leapfrog スキーム . . . 119
7.12.3 Matsuno スキーム, および一般の two level スキーム . . . 121
7.12.4 two-step Lax-Wendroff スキーム . . . 122
7.12.5 補足 . . . 124
7.13 4 次の正確度を持つ空間中心差分スキーム . . . 125
7.13.1 Introduction . . . 125
7.13.2 位相速度に対する効果 . . . 125
7.13.3 補足 . . . 126
7.14 2 次元移流方程式の安定性 . . . 127
7.15 非線形移流方程式のおばけ . . . 128
7.16 非線形移流方程式の不安定性 . . . 129
7.16.1 非線形不安定性 . . . 129
7.16.2 非線形不安定性の抑止と予防 . . . 129
7.17 Arakawa スキーム . . . 131
7.17.1 渦度方程式 . . . 131
7.17.2 渦度方程式の波数空間における性質 . . . 131
7.17.3 Arakawa の方法 . . . 134
7.18 直接法 . . . 137
7.18.1 Introduction . . . 137
7.18.2 格子の設定と方程式の差分化 . . . 137
7.18.3 Fourier 変換 . . . 138
7.18.4 方程式の解 . . . 139
7.19 緩和法 . . . 140
7.19.1 Introduction . . . 140
7.19.2 熱伝導方程式の定常解としての Poisson 方程式の解 . . . 140
7.19.3 解の求め方 . . . 140
7.19.4 緩和法の収束の速さ . . . 141
7.20 補足 ・放物型方程式 . . . 142
7.20.1 放物型方程式 . . . 142
7.21 その他の参考文献 . . . 146

8 謝辞 **147**

9 著作権 **148**

7 時間発展問題の解法 75
7.1 気象業界における数値計算の歴史と現状 . . . 75
7.1.1 歴史 . . . 75
7.2 有限差分法の原理 . . . 77
7.2.1 導関数の差分近似 . . . 77
7.2.2 有限差分スキームにおける基礎概念 . . . 81
7.2.3 収束性 (convergence) . . . 82
7.2.4 安定性の判定法 . . . 85
7.3 常微分方程式を解く各種スキーム . . . 90
7.3.1 Introduction . . . 90
7.3.2 常微分方程式の一般的な解法 . . . 90
7.3.3 Euler 法 . . . 91
7.3.4 修正 Euler 法 . . . 91
7.3.5 Runge-Kutta 法 . . . 92
7.3.6 連立常微分方程式 . . . 93
7.3.7 時間発展問題 . . . 95
7.3.8 スキームの分類 . . . 95
7.3.9 two level scheme の定義 . . . 96
7.3.10 three level scheme の定義 . . . 97
7.3.11 補足 . . . 97
7.4 semi-Implicit 法 . . . 99
7.4.1 Introduction . . . 99
7.4.2 semi-Implicit 法の原理 . . . 99
7.5 振動方程式に適用したときのスキームの特性 . . . 100
7.5.1 Introduction . . . 100
7.5.2 特性の調べ方 . . . 100
7.5.3 two level schemes の特性 . . . 100
7.5.4 Three level schemes の特性 . . . 103
7.6 摩擦方程式に適用したときのスキームの特性 . . . 107
7.6.1 Introduction . . . 107
7.6.2 two level scheme の特性 . . . 107
7.6.3 three level scheme の特性 . . . 108
7.7 方程式が複数の項を含む場合のスキームの選択 . . . 109
7.8 Split 法 . . . 110
7.9 1 次元移流方程式の数値計算法 . . . 111
7.9.1 Introduction . . . 111
7.9.2 正確度 . . . 111
7.9.3 CFL condition . . . 112
7.9.4 数値分散 . . . 113
7.10 スキームのまとめ . . . 115
7.11 上流差分スキーム . . . 116
7.11.1 上流差分と下流差分 . . . 116
7.11.2 上流差分スキームの特徴 . . . 116
7.12 2 次の正確度を持つ空間中心差分スキーム . . . 119

4.2.3 QR 分解とグラム ・シュミットの直交化 . . . 28
4.2.4 右上三角行列 . . . 30
4.2.5 右上三角行列を与える回転変換 . . . 30
4.2.6 QR 法の考え方 . . . 30
4.3 ヘッセンベルグ行列の利用 . . . 32
4.3.1 ヘッセンベルグ行列を用いるご利益 . . . 32
4.3.2 鏡像変換 . . . 33
4.3.3 任意の行列をヘッセンベルグ形に直す法 . . . 33
4.3.4 ヘッセンベルグ行列の QR 分解 . . . 35
4.4 二, 三の補足 . . . 36
4.5 応用問題 . . . 36
4.6 参考文献 . . . 37

5 データ解析の基礎 38
5.1 データの統計学的記述 . . . 38
5.1.1 平均と分散 . . . 38
5.1.2 重要な確率分布 . . . 40
5.2 最小自乗法–データのモデルフィットの簡単な例 . . . 41
5.2.1 最尤法 (maximum likelihood method) . . . 41
5.2.2 最小自乗法–線形回帰 . . . 42
5.3 練習問題 . . . 45

6 時系列解析の基礎 47
6.1 物理的データの基礎的記述 . . . 47
6.1.1 確定的データについて . . . 47
6.1.2 不規則データについて . . . 49
6.2 自己相関関数およびスペクトル密度関数 . . . 50
6.2.1 自己相関関数 . . . 50
6.2.2 自己相関関数の性質 . . . 51
6.2.3 スペクトル密度関数 . . . 51
6.2.4 Wiener-Khintchine の公式 . . . 52
6.3 スペクトル密度関数の具体的表式 . . . 53
6.3.1 過渡関数の場合 . . . 54
6.3.2 平均パワー有限の関数 . . . 55
6.4 パワースペクトルの推定とその誤差 . . . 56
6.4.1 記録の有限性の効果 . . . 57
6.4.2 記録の不連続性の効果 . . . 60
6.5 パワースペクトルを求める実際の計算手順 . . . 62
6.5.1 Blackman-Tukey 法 . . . 62
6.5.2 直接法 (記録の直接フーリエ変換による方法) . . . 64
6.6 演習問題 . . . 69
6.7 参考文献 . . . 70
6.8 付録 1:Wiener-Khintchine の定理について . . . 71
6.9 付録 2:フーリエ変換の性質と例 . . . 73

目 次

1 はじめに **1**

2 高次代数方程式の数値解法 **3**
2.1 はじめに . . . 3
2.2 Newton-Raphson 法 . . . 3
2.3 Regula-Falsi 法 . . . 4
2.4 二分法 . . . 4
2.5 二変数二分法 . . . 6
2.6 参考になる文献 . . . 8

3 行列計算その 1 : 連立 1 次方程式 **9**
3.1 Introduction . . . 9
3.2 ガウスの消去法 . . . 9
3.2.1 まずは例から . . . 9
3.2.2 一般化すると . . . 10
3.3 単純なガウスの消去法の問題点とその改善法 . . . 11
3.3.1 ピボットの選択 . . . 11
3.3.2 スケーリング . . . 11
3.3.3 解の反復改良 . . . 12
3.4 LU 分解法 . . . 13
3.4.1 LU 分解法のアルゴリズム . . . 14
3.5 コレスキー分解法 . . . 14
3.6 例題 . . . 15
3.7 付録 . . . 18
3.7.1 付録 1: LU 分解定理の証明 . . . 18
3.7.2 付録 2: LU 分解法のアルゴリズムのフローチャート . . . 21
3.7.3 付録 3: コレスキー分解定理の証明 . . . 22
3.7.4 付録 4: コレスキー分解法のアルゴリズムのフローチャート . . . 23
3.7.5 付録 5: 条件数 . . . 24

4 行列計算その 2 : 固有値固有ベクトル **25**
4.1 固有値問題の解法の概観 . . . 25
4.2 QR 法 . . . 26
4.2.1 QR 法の基本方針 . . . 26
4.2.2 QR 法の処方箋 . . . 27

地球惑星物理学演習

— 数値計算入門 —

東京大学理学部地球惑星物理学科
地球惑星物理学演習テキスト編集グループ

1994 年 1 月 25 日 (Ver.1.3)