第Ⅴ部

数学上級編

第11章　線形代数：上級編

第3章（初級編）では，連立方程式を解くとき逆行列の計算を経由するのは実際的でなく，避けなくてはいけないと述べました．
計算の精度や効率を考えるのは，純粋の数学とは離れた議論のように感じるかも知れませんが，決してそうではありません．特に計算を主題とするならなおさらです．計算は有限の資源（時間や手間，計算機の規模）に依存しており，計算アルゴリズムの適切な選択によって，不可能な計算が実行可能な計算に変わることはよくあります．

11.1　連立方程式の解法

11.1.1　解法の復習

逆行列の計算あるいは連立方程式の計算，固有値・固有ベクトルの計算は，勿論多くの共通性があります．ここでは連立方程式

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

を解くことを考えてみましょう．A は $n \times n$ 行列，$\boldsymbol{x}, \boldsymbol{b}$ は n 個の成分からなる列ベクトルです．

MATLAB ではこの演算は，A を $n \times n$ 行列, b を n 個の成分からなる列ベクトルと定義した後,

$$x = A \backslash b$$

と書きます.（MATLAB を含む）実用的なプログラムでは，逆行列を計算して b にかけるというアルゴリズムはとりません．MATLAB では、行列 A の特徴により，以下で述べる掃き出し法，LU 分解あるいは他の方法を自動的に選択します．
http://matlab.izmiran.ru/help/techdoc/ref/mldivide.html

（MATLAB の線形代数関数および行列演算は，Fortran サブルーチン・ライブラリとして評価の高い LAPACK を基に構築されています.）

11.1.2　行列の分解

11.1.2.1　ガウス - ジョルダンの掃き出し法

通常，教科書などで教えるのはガウス - ジョルダンの「掃き出し法」と呼ばれる次の手順です：

- **[Step1]**　$n \times n$ 行列 A および列ベクトル $\boldsymbol{b}$ を
$$[A \ \boldsymbol{b}]$$
と並べて n 行 $n+1$ 列（$n \times (n+1)$）の行列を作る.

- **[Step2]** 「基本行演算（行の定数倍，加減のみからなる基本変形演算）」を繰り返し，$[A \ \boldsymbol{b}]$ を
$$[E \ \boldsymbol{x}]$$
となるようにする．E は $n \times n$ の単位行列である．$\boldsymbol{b}$ があったところに答えである列ベクトル $\boldsymbol{x}$ が現れる．ただし基本行演算により成分の順番が入れ替わっていることがある.

- [**計算量**] 上の計算に必要な演算回数は乗除算が n^3 のオーダーとなる．掃き出し法では，連立方程式の大きさ n が例えば 10 倍になれば，演算回数は 1000 倍になり，すぐに現実的な時間内で結果を得ることができなくなる.

図 11.1に掃き出し法で逆行列を計算した結果を示しておきましょう．行列 A と単位行列を並べて，基本行操作を行ったものです．コマンド rref を用います．結果では，単位行列の所に逆行列が現れます．その確認のための計算が右に示してあります.

```
B=[2 1 0 1 0 0;1 2 1 0 1 0;0 1 2 0 0 1]
B =
    2    1    0    1    0    0
    1    2    1    0    1    0
    0    1    2    0    0    1
>> R=rref(B)
R =
   1.0000     0        0      0.7500  -0.5000   0.2500
        0   1.0000     0     -0.5000   1.0000  -0.5000
        0        0   1.0000   0.2500  -0.5000   0.7500
```

```
>>  A=[2 1 0;1 2 1;0 1 2]
A =
    2    1    0
    1    2    1
    0    1    2
>> inv(A)
ans =
   0.7500  -0.5000   0.2500
  -0.5000   1.0000  -0.5000
   0.2500  -0.5000   0.7500
```

図 11.1: ガウス - ジョルダンの掃き出し法

11.1.2.2 LU 分解

A を，下三角行列 L と上三角行列 U の積

$$A = LU = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ l_{n1} & \cdots & l_{nn-1} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix}$$

に書き換える LU 分解という方法があります．この方法は実際によく使われます.

演算の数は，A が密行列（要素の大部分が 0 ではない行列）である場合には，ガウス-ジョルダンの掃き出し法と比べて大きく異なることはありません．A が疎（要素の大部分が 0 である行列）である大行列の場合にも，LU 分解したときの L および U はやはり疎であるという性質を保ちます．そのため A が疎である場合には計算量に大きな差 (演算回数が n^2 程度の量になることもある）が生じます.

LU 分解の例を図 11.2示します．LU 分解を行うために，コマンド lu が用意されています．[L,U,P] = lu(A) は、L*U = P*A となる上三角行列 U，単位対角をもつ下三角行列 L，および置換行列 P を返します.

```
>>  A=[2 1 0;1 2 1;0 1 2]
A =
    2    1    0
    1    2    1
    0    1    2
```

```
>> [L,U,P]=lu(A)
L =
   1.0000         0         0
   0.5000    1.0000         0
        0    0.6667    1.0000
U =
   2.0000    1.0000         0
        0    1.5000    1.0000
        0         0    1.3333
P =
    1    0    0
    0    1    0
    0    0    1
```

図 11.2: LU 分解

11.1.2.3　コレスキー（Cholesky）分解

行列 A が正定値エルミートであるとき，行列 A を上三角行列 U と U のエルミート共役 U^* との積に分解することができます．これをコレスキー分解と呼び，やはりよく用いられます．A のエルミート性を利用した LU 分解の特別な場合です．これのためにも U=chol(A) または U=chol(A,'upper') というコマンドが用意されています．

$$\mathrm{A} = \mathrm{U}' * \mathrm{U}$$

(U' は U の転置．　数学の記号なら t または T を右または左の上付きとします.）また下三角行列で求めるときは L=chol(A,'lower') とします．：

$$\mathrm{A} = \mathrm{L} * \mathrm{L}'$$

11.1.2.4　シュール（Schur）分解

任意の n 次正方行列 A に対しては次の性質が成り立ちます．

$$A = UTU^* \quad .$$

```
>> A=[1 2 3;4 5 1;2 3 4]
A =
    1    2    3
    4    5    1
    2    3    4
>> rank(A)
ans =
    3
>> eig(A)
ans =
   8.5233
  -0.3255
   1.8022
```

```
>> A=[1 2 3;4 5 1;2 3 4]
A =
    1    2    3
    4    5    1
    2    3    4
>> [U,T]=schur(A)
U =
   0.4233    0.8063   -0.4131
   0.6575   -0.5871   -0.4722
   0.6233    0.0717    0.7787
T =
   8.5233    0.2591   -1.1689
        0   -0.3255    2.7507
        0         0    1.8022
```

図 11.3: シュール分解とその性質.

U はユニタリ行列，T は上三角行列で，その対角成分は A の固有値 $(\lambda_1, \lambda_2, \cdots ,)$ が並びます．U^* は U のエルミート共役です．シュール分解のコマンドは schur です（図 11.3）.

11.1.2.5　正規行列のシュール（Schur）分解

A が正規行列（$A^*A = AA^*$) である場合には，シュール分解は以下のようになります（図 11.4）.

$$A = UDU^* \quad .$$

U は直交行列，D は対角に A の固有値が並んだ対角行列です．U の第 j 列は，A の固有値 λ_j に対応した固有ベクトルです.

11.2　特異値分解

実際の連立方程式の場合には，常に解が得られるとは限りません．行列 A が正方行列でない場合，あるいは正方行列であっても正則でない場合

```
>> A=[1 3;3 1]
A =
    1    3
    3    1
>> [U,D]=schur(A)
U =
  -0.7071   0.7071
   0.7071   0.7071
```

```
D =
   -2    0
    0    4
>> U*D*U'
ans =
    1.0000    3.0000
    3.0000    1.0000
```

```
>> eig(A)
ans =
   -2
    4
>> U'*U
ans =
    1.0000    0.0000
    0.0000    1.0000
```

図 11.4: 正規行列のシュール分解

などです．より厄介な問題は，A のいくつかの固有値が他と比べて非常に小さくなっていて，ほとんど正則でない状態にある場合です．このようなときは，固有値問題あるいはシュール分解と深い関係にある以下の性質が大変有用です.

11.2.1　特異値分解の概要

11.2.1.1　特異値分解定理

A を階数 r の $m \times n$ 複素行列とします．このとき

$$A = U\Sigma V^*$$

という A の分解が存在します．ここで U は $m \times m$ のユニタリ行列，V^* は $n \times n$ のユニタリ行列 V のエルミート共役です．Σ は $m \times n$ 行列で，下のような形になります：

$$\Sigma = \begin{pmatrix} \mathrm{diag}(\sigma_1, \cdots \sigma_r) & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix}$$

これを複素行列 A の特異値分解と呼び，Σ の (i,i) 要素 σ_i を A の**特異値**と呼びます．ここで，$\mathrm{diag}(\sigma_1, \cdots \sigma_r)$ は $\sigma_1, \cdots \sigma_r$ を対角成分とする $r \times r$ の対角行列であり，$O_{l,k}$ は $l \times k$ のゼロ行列です.

```
>> A=[1 2 3 4;5 6 7 8]
A =
    1    2    3    4
    5    6    7    8
>> [U S V]=svd(A)
U =
  -0.3762   0.9266
  -0.9266  -0.3762
S =
  14.2274    0        0        0
       0   1.2573    0        0
```

```
V =
   -0.3521  -0.7590  -0.4001  -0.3741
   -0.4436  -0.3212   0.2546   0.7970
   -0.5352   0.1165   0.6910  -0.4717
   -0.6268   0.5542  -0.5455   0.0488
```

```
>> U*S*V'
ans =
    1.0000    2.0000    3.0000    4.0000
    5.0000    6.0000    7.0000    8.0000
```

図 11.5: 特異値分解の例.

11.2.1.2　特異値分解の例

[U,S,V] = svd(A) は，$A = USV^*$ となるように，行列 A の特異値分解を行います（図 11.5）.

11.2.2　疑似逆行列

これまでの多くの場合では，行列 A の逆行列が存在していると仮定し，それをいかに効率よく求められるか，ということを考えました．また逆行列が存在しない場合にも行列の分割という手段で対応しました．ここでは，逆行列が存在しない場合にも，逆行列に代わる概念である疑似逆行列という考え方を紹介しましょう（ムーア-ペンローズの擬似逆行列）.
A の疑似逆行列 A^+ は一意に決まり，以下の性質を満たします.
(1) $AA^+A = A$
(2) $A^+AA^+ = A^+$
(3) $(AA^+)^* = AA^+$，すなわち AA^+ はエルミート行列.
(4) $(A^+A)^* = A^+A$，すなわち A^+A はエルミート行列.
この性質から A が正則な場合には，$A^+ = A^{-1}$ であることが分かります.
A^+ を A の疑似逆行列と呼びます.

11.2.2.1　疑似逆行列の具体的な形

- $n > m = r$ の場合
$$A = U\Sigma V^*$$
としてもう少し踏み込んでみましょう.
$$\Sigma = (\mathrm{diag}(\sigma 1, \cdots, \sigma_m) \;\; 0_{m,n-m})$$
です. また
$$\Sigma^+ = \begin{pmatrix} \mathrm{diag}(1/\sigma_1, \cdots, 1/\sigma_m^{-1}) \\ 0_{n-m,m} \end{pmatrix}$$
とすると A^+ は
$$A^+ = V\Sigma^+ U^*$$
です. 実際これが疑似逆行列の持つべき性質を満たすことは簡単に確かめることができます.

- $m > n = r$ の場合にも同じように書かれます. ただしこの場合には
$$\Sigma = \begin{pmatrix} \mathrm{diag}(\sigma_1, \cdots \sigma_n) \\ 0_{n-m,m} \end{pmatrix}$$
とします.

ユニタリ行列 U および V を, 正規直交（列）ベクトル $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$
$$\boldsymbol{u}_i^* \cdot \boldsymbol{u}_j = \delta_{ij} \;,\; i, j = 1 \sim m$$
および正規直交（列）ベクトル $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$
$$\boldsymbol{v}_i^* \cdot \boldsymbol{v}_j = \delta_{ij} \;,\; i, j = 1 \sim n$$
を用いて書き換えてみましょう：
$$U = (\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m)$$
および
$$V = (\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n) \;,\; V^* = \begin{pmatrix} \boldsymbol{v}_1^* \\ \vdots \\ \boldsymbol{v}_n^* \end{pmatrix}$$

```
C=[1 2;3 4]
C =
    1    2
    3    4
>> rank(C)
ans =
    2
>> D=pinv(C)
D=
   -2.0000    1.0000
    1.5000   -0.5000
>> inv(C)
ans =
   -2.0000    1.0000
    1.5000   -0.5000
```

```
>> A=[2 2;1 1]
A =
    2    2
    1    1
>> rank(A)
ans =
    1
>> B=pinv(A)
B =
    0.2000    0.1000
    0.2000    0.1000
>> A*B*A
ans =
    2.0000    2.0000
    1.0000    1.0000
>> B*A*B
ans =
    0.2000    0.1000
    0.2000    0.1000
```

```
>> (A*B)'
ans =
    0.8000    0.4000
    0.4000    0.2000
>> A*B
ans =
    0.8000    0.4000
    0.4000    0.2000

>> (B*A)'
ans =
    0.5000    0.5000
    0.5000    0.5000
>> B*A
ans =
    0.5000    0.5000
    0.5000    0.5000
```

図 11.6: 疑似逆行列の例.

と書くと（$\boldsymbol{v}_i^*$ は行ベクトル）

$$
\begin{aligned}
A &= U\Sigma V^* = \sum_{i=1} \sigma_i \boldsymbol{u}_i\ \boldsymbol{v}_i^* \\
A^+ &= V\Sigma^+ U^* = \sum_{i=1} \frac{1}{\sigma_i} \boldsymbol{v}_i\ \boldsymbol{u}_i^*
\end{aligned}
$$

となります $\boldsymbol{u}_i\ \boldsymbol{v}_i^*$ は $m \times n$ 行列です．これらの式から

$$
\begin{aligned}
A\boldsymbol{v}_i &= \sigma_i \boldsymbol{u}_i \\
A^+ \boldsymbol{u}_i &= \frac{1}{\sigma_i} \boldsymbol{v}_i
\end{aligned}
$$

も得られます．このように固有値問題が成り立たない場合にも，こので示したような一般化固有値問題が成り立ち，大変に有用です．

疑似逆行列の計算を図 11.6に示しましょう．B = pinv(A) は行列 A の Moore-Penrose 疑似逆行列を返します。

11.3　特異値分解の応用

ここでは簡単のために，A は $m \times n$ の実行列であり，その階数は r であるとします．このとき，特異値分解

$$A = U\Sigma V^T \quad .$$

が存在します．ただし U および V はそれぞれ，$m \times m$ および $n \times n$ の直交行列で，Σ は $m \times n$ 行列です．特異値 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$ は Σ の対角に並びます.

11.3.1　最小二乗法

上の A に対して実ベクトル $\boldsymbol{b}$ が与えられたとします．このとき

$$||A\boldsymbol{x} - \boldsymbol{b}||^2 \equiv (A\boldsymbol{x} - \boldsymbol{b})^T(A\boldsymbol{x} - \boldsymbol{b})$$

を最小にするように $\boldsymbol{x}$ を決める問題を考えます．これを最小 2 乗問題といい，よく知られた最小 2 乗法と同等の問題です．またこの解を**最小 2 乗解**といいます.
今の場合，最小 2 乗問題の解 $\boldsymbol{x}'$ は

$$A^T A\boldsymbol{x}' = A^T\boldsymbol{b}$$

の解です.
連立方程式系 $A\boldsymbol{x} = \boldsymbol{b}$ が与えられたとき，この方程式系の解がない場合には，A の疑似逆行列 A^+ を用いて，$\boldsymbol{x}' = A^+\boldsymbol{b}$ が最適解になります．これは上の問題の最小 2 乗解でもあります.

11.3.1.1　簡単な例

$$A\boldsymbol{x} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \boldsymbol{b} \quad .$$

この方程式は解を持ちません.

これを $A^TA\boldsymbol{x} = A^T\boldsymbol{b}$ と変形すれば

$$A^TA\boldsymbol{x}' = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = A^T\boldsymbol{b} \quad .$$

これを解いて

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

を得ます．これが最小 2 乗解です．
この計算を疑似逆行列を用いて解いた結果を，図 11.7に示しましょう．

```
>> A=[1 1;1 1]
A =
    1    1
    1    1
>> b=[3;1]
b =
    3
    1
```

```
>> rank(A)
ans =
    1
>> P = pinv(A)*b
P =
    1.0000
    1.0000
```

図 11.7: 連立方程式の最小 2 乗解

11.3.2 主成分分析 (Principal Component Analysis)

多くのデータが与えられ，それらが互いに相関のある多数の変数からなっているとき，互いに相関の無い少数の変数（**主成分**）で全体のデータを記述することが望まれます．そのような成分を探す方法の 1 つが**主成分分析**です．

11.3.2.1　データ行列の特異値分解

11.3.2.1.1　データの標準化

N 個のサンプルについて其々P 個の変数（データ）を持つデータ群が与えられていて，サンプル $i(i = 1 \sim N)$ についての生のデータを $\{t_{i\alpha}\}(\alpha = 1 \sim P)$ とします．このデータが，N 行 P 列の**データ行列** T にまとめられているとします：

$$T = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1P} \\ t_{21} & t_{22} & \cdots & t_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ t_{N1} & t_{N2} & \cdots & t_{NP} \end{pmatrix} .$$

一般にこれらの P 個の変数はそれぞれ単位も異なり，値やそのバラツキの大きさも異なります．そのため一般にデータの**標準化**という手法がとられます．すなわちデータの中心を 0 と定め，値をそれぞれの標準偏差で規格化します：

$$\bar{t}_s = \frac{1}{N}\sum_{m=1}^{N} t_{ms} \ , \quad \sigma_{t_s}^2 = \frac{1}{N-1}\sum_{m=1}^{N}(t_{ms} - \bar{t}_s)^2 \ ,$$
$$x_{ns} = \frac{t_{ns} - \bar{t}_s}{\sigma_{t_s}} \ , \quad \sum_{m=1}^{N} x_{ms} = 0 \ , \quad \sum_{m=1}^{N} x_{ms}^2 = N - 1 \ .$$

データの標準化により，**（標準化された）データ行列**は

$$X = \{x_{ij}\} \ , \quad i = 1 \sim N, \quad j = 1 \sim P$$

です．データの各要素 x_{mi} については

$$\text{平均値} \ : \ \frac{1}{N}\sum_{m=1}^{N} x_{ms} = 0$$
$$\text{標準偏差} \ : \ \frac{1}{N-1}\sum_{m=1}^{N} x_{ms}^2 = 1$$

となります．

11.3.2.1.2　データ行列と共分散行列

$$Q = \frac{1}{N-1} X^T X = \begin{pmatrix} \sigma_{x_1x_1} & \sigma_{x_1x_2} & \cdots & \sigma_{x_1x_P} \\ \sigma_{x_2x_1} & \sigma_{x_2x_2} & \cdots & \sigma_{x_2x_P} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{x_Px_1} & \sigma_{x_Px_2} & \cdots & \sigma_{x_Px_P} \end{pmatrix}$$

と定義される $P \times P$ 行列を**共分散行列**と呼びます．ただし

$$\sigma_{x_{s_1}x_{s_2}} = \frac{1}{N-1}\sum_{m=1}^{N} x_{ms_1}x_{ms_2} \ , \quad \sigma_{x_{s_1}x_{s_1}} = \sigma^2_{x_{s_1}} = 1 \ .$$

11.3.2.1.3　共分散行列の特異値分解

正規直交行列

$$V = (\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_P) = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1P} \\ v_{21} & v_{22} & \cdots & v_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ v_{P1} & v_{P2} & \cdots & v_{PP} \end{pmatrix}$$

を用いて

$$Z = XV \quad , \quad (Z)_{m\alpha} = (XV)_{m\alpha} = \sum_{\beta=1}^{P} x_{m,\beta} v_{\beta,\alpha}$$

という変換を行いましょう．これは各データを特徴付ける $1 \sim P$ の変数の変数の線形結合で新しい変数を定義したことになります．
ここで V を

$$V^T Q V = V^T \ \frac{X^T X}{N-1} \ V = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_P^2 \end{pmatrix} = \Sigma$$

を満たすように決めるとすれば

$$QV = V\Sigma \quad , \quad Q\boldsymbol{v}_\alpha = \sigma_\alpha^2 \boldsymbol{v}_\alpha \ , \quad \sum_{\alpha=1}^{P} \sigma_\alpha^2 = P$$

となり，Q の固有値問題であることが分かります．つまり主成分分析ではまず共分散行列の固有空間への分解が行われます．
これまでのことを少し整理しましょう．

1. Q の固有値 σ_α^2 $(\alpha = 1 \sim P)$ を対角に並べた行列 $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_P)$ は共分散行列 Q を対角化（固有空間に分解）したものです．

2. 行列 $X^T X$ の固有値問題；

$$X^T X \boldsymbol{v}_\alpha = \lambda_\alpha^2 \boldsymbol{v}_\alpha \ ,$$

固有値 $\{\lambda_\alpha^2\}$ が大きな順に

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_P > 0$$

と並べられているとします．$(\lambda_\alpha, \boldsymbol{v}_\alpha)$ を**第 α 主成分**といいます．

3. 上の諸量をまとめて，

$$\begin{aligned} \Lambda &= \mathrm{diag}(\lambda_1, \cdots, \lambda_P) \\ \boldsymbol{u}_\alpha &= \frac{1}{\lambda_\alpha} X \boldsymbol{v}_\alpha \\ U &= (\boldsymbol{u}_1, \cdots, \boldsymbol{u}_P) \end{aligned}$$

を定義します（$\lambda_\alpha^2 = (N-1)\sigma_\alpha^2$）．

$$\begin{aligned} X \boldsymbol{v}_\alpha &= \lambda_\alpha \boldsymbol{u}_\alpha \\ X^T \boldsymbol{u}_\alpha &= \lambda_\alpha \boldsymbol{v}_\alpha \end{aligned}$$

などは定義から直ぐに証明できます．

4. $\boldsymbol{v}_\alpha$ は正規直交ベクトルですから，$\boldsymbol{u}_\alpha$ も正規直交ベクトルです ($U^T U = E$：$P \times P$ 単位ベクトル）．

5. 上から以下の式が成り立つことが分かります：

$$\begin{aligned} XV &= (X\boldsymbol{v}_1, \cdots, X\boldsymbol{v}_P) = (\lambda_1 \boldsymbol{u}_1, \cdots, \lambda_P \boldsymbol{u}_P) \\ &= (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots)\Lambda = U\Lambda \end{aligned}$$

これは

$$X = U \Lambda V^T$$

と書き直すことができ，X の特異値分解であることが分かります．

6. また行列 X^TX、XX^T のスペクトル分解；

$$\begin{aligned} X^TX &= \sum_{\alpha=1}^{P} \lambda_\alpha^2 \boldsymbol{v}\boldsymbol{v}^T \ , \\ XX^T &= \sum_{\alpha=1}^{P} \lambda_\alpha^2 \boldsymbol{u}\boldsymbol{u}^T \end{aligned}$$

が成り立ちます.

11.3.2.2 主成分分析の目的

与えられた N 個の各データが P 個の変数で特徴付けられています．X^TX の特異値を順次大きい順番で考えると，大きい固有値が行列 X^TX の特徴を決めることが分かります（スペクトル分解）．小さい固有値を無視すれば，少しの変数でデータの本質を捉えられると期待できます.

11.3.2.2.1 寄与率

以下に定義される**寄与率** は，得られた主成分がデータをどの程度説明しているかを示す尺度です.
第 α 固有値 (特異値) λ_α は各空間がどの程度 X^TX の性格を決めているかが分かりますから，寄与率として通常用いられるのはこれを規格化した

$$\frac{\lambda_\alpha}{\sum_{\beta=1}^{P} \lambda_\beta}$$

です．また，寄与率を第 1 主成分から順に累積していったもの（例えば α 主成分まで足したもの）を累積寄与率といいます：

$$\frac{\sum_{\beta=1}^{\alpha} \lambda_i}{\sum_{\gamma=1}^{P} \lambda_\gamma} \ .$$

11.3.2.2.2 因子負荷量

分散共分散行列の固有ベクトルから，例えば第 α 主成分の主軸方向の長さを計算できます．この長さを**主成分スコア**といいます．また固有ベクトルが因子負荷量として使われます.

11.3.2.3　主成分分析の例

MATLAB が提供しているデータ hald（セメントの発熱と混合成分）を例に主成分分析の演習をしましょう．13 種類のセメント組成に対して，ingredients 部分（4 つの列）には 4 種類のセメント原料の組成率が与えられています．この部分をデータ行列 T とします．

$$T = \begin{array}{rrrr} 7 & 26 & 6 & 60 \\ 1 & 29 & 15 & 52 \\ 11 & 56 & 8 & 20 \\ 11 & 31 & 8 & 47 \\ 7 & 52 & 6 & 33 \\ 11 & 55 & 9 & 22 \\ 3 & 71 & 17 & 6 \\ 1 & 31 & 22 & 44 \\ 2 & 54 & 18 & 22 \\ 21 & 47 & 4 & 26 \\ 1 & 40 & 23 & 34 \\ 11 & 66 & 9 & 12 \\ 10 & 68 & 8 & 12 \end{array}$$

- 各列について要素の平均値 ts および標準偏差 sigt を計算します．平均値は mean, 分散は var が与えます．

$$\begin{aligned} ts = \mathrm{mean}(T) &= \quad 7.4615 \quad 48.1538 \quad 11.7692 \quad 30.0000 \\ sigts = \mathrm{sqrt}(\mathrm{var}(T)) &= \quad 5.8824 \quad 15.5609 \quad 6.4051 \quad 16.7382 \end{aligned}$$

- 用意されているコマンド zscore を使い，平均値，標準偏差，標準化されたデータ行列を計算します．

$$[X, ts, sigts] = \mathrm{zscore}(T)$$

  $ts, sigts, X$ はそれぞれ，もとのデータ行列の各列ベクトル成分に関する平均値 $\bar{t}_s$，標準偏差 σ_{t_s}，および標準化されたデータ行列 X です．mean, var を用いた計算と比較して確認しましょう．

- X の共分散行列 Q は（ここでは $N = 13$ ですから）MATLAB の言葉で書けば

$$Q = X' * X/12$$

  です．共分散行列の固有値 σ_α^2 を対角に並べた行列 $SIGMA2$, 固有ベクトル（主成分 $\alpha = 1 \sim P$) を各列成分（もとの ingredients=1 $\sim P$ を行成分に）並べた行列 V を計算します：

$$[V, SIGMA2] = \mathrm{eig}(Q)\ .$$

$$V = \begin{array}{rrrr} 0.2411 & 0.6755 & 0.5090 & -0.4760 \\ 0.6418 & -0.3144 & -0.4139 & -0.5639 \\ 0.2685 & 0.6377 & -0.6050 & 0.3941 \\ 0.6767 & -0.1954 & 0.4512 & 0.5479 \end{array}$$

$$SIGMA2 = \begin{array}{rrrr} 0.0016 & 0 & 0 & 0 \\ 0 & 0.1866 & 0 & 0 \\ 0 & 0 & 1.5761 & 0 \\ 0 & 0 & 0 & 2.2357 \end{array}$$

$SIGMA2$ は Q の固有値行列ですから，主成分分散 σ_α^2 を対角に並べたものです．列は σ_α の昇順です．

さらに固有ベクトルをサンプルの成分を行成分（$j = 1 \sim N$) で書き表した行列 $Z = XV$ を計算してみてください．

- これまでの計算を特異値分解で計算します：

$$[U1, Lambda1, V1] = \mathrm{svd}(X) \ .$$

$$Lambda1 = \begin{array}{rrrr} 5.1796 & 0 & 0 & 0 \\ 0 & 4.3489 & 0 & 0 \\ 0 & 0 & 1.4964 & 0 \\ 0 & 0 & 0 & 0.1396 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

$$V1 = \begin{array}{rrrr} 0.4760 & -0.5090 & 0.6755 & 0.2411 \\ 0.5639 & 0.4139 & -0.3144 & 0.6418 \\ -0.3941 & 0.6050 & 0.6377 & 0.2685 \\ -0.5479 & -0.4512 & -0.1954 & 0.6767 \end{array}$$

ここの V は前項の V と同じもの（列の並べ方は降順），$Lambda1$ は X の特異値 λ_α を並べた行列です．$\lambda_\alpha^2 = (N-1)\sigma_\alpha^2$ より，各成分に対して $Lambda1 = \sqrt{12 \times SIGMA2}$ であることが確かめられます．

- これまでの計算（主成分分析）を一度に行うコマンドが pca です：

$$[COEFF, SCORE, LATENT] = \mathrm{pca}(X)\ .$$

これらの返り値の列は主成分分散 σ_α^2 の大きい方からの順（降順）に並びます．$COEFF$ は $N \times p$ (標準化された）データ行列 X の主成分係数 (負荷量）$V1$ を，$SCORE$ は主成分スコア $Z = X * V$ を，$LATENT$ は主成分分散 σ_α^2 （$SIGMA2$）を返します：

$$LATENT = \begin{matrix} 2.2357 \\ 1.5761 \\ 0.1866 \\ 0.0016 \end{matrix}$$

第12章　非線形微分方程式

12.1　相空間と安定性

捕食系の振る舞いをよく捉えたモデルとして，ボルテラ系（あるいはロトカ・ボルテラ方程式）があります:

$$\begin{aligned}\frac{dx}{dt} &= x - xy \\ \frac{dy}{dt} &= xy - y\end{aligned}$$

x, y はそれぞれ被捕食者、捕食者の個体数です.
これから t を消去すると

$$\frac{dx}{dy} = \frac{dx/dt}{dy/dt} = \frac{x - xy}{xy - y} = \frac{-1 + \frac{1}{y}}{1 - \frac{1}{x}}$$

となり

$$(x - \log x) + (y - \log y) = C$$

が得られます．解は，x-y 空間で

$$xy\ \mathrm{e}^{-x-y} = f = 一定$$

の曲線上を動くことが分かります.
このような解析は重要であり，ここで x-y 空間を相空間と呼びます．相空間での振る舞いを知ることが，解の性質，特に解の安定性（充分時間がたったとき，解が有限の領域に留まらないとき，その解を ‘不安定’ といいます.）を知るために大変重要です．相空間内において，上で得られた解の運動を調べたものが図 12.1です．$f =$ 一定の等高線図の書き方が分かるでしょう.

```
[x,y]=meshgrid(0:0.01:3);f=x.*y.*exp(-x-y);
[C,h]=contour(x,y,f,5)
clabel(C,h,'manual')
hold on
%
[x,y]=meshgrid(-3:0.1:3);f=x.*y.*exp(-x-y);
[C,h]=contour(x,y,f,25)
clabel(C,h,'manual')
[C,h]=contour(x,y,f,500)
clabel(C,h,'manual')
hold on
```

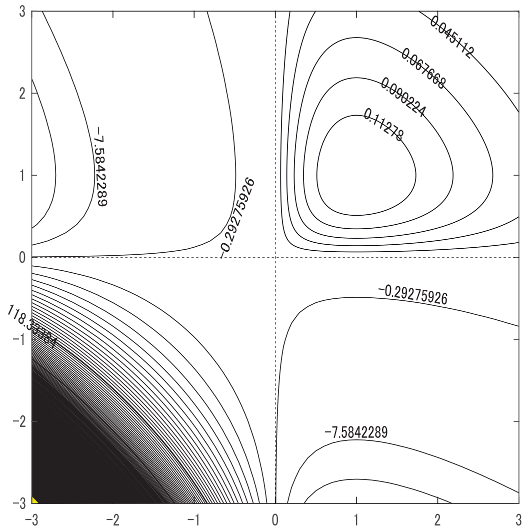

図 12.1: 相空間と解の安定性

- 図 12.1では，$x, y > 0$, $x, y < 0$，$xy < 0$ において，等高線の間隔（スケール）をそれぞれ変えてある．
- $[\mathrm{x}, \mathrm{y}] = \mathrm{meshgrid}(\cdots)$ では，2 次元空間 xx-y のメッシュを定義 ([始点：間隔：終点])．
- contour で等高線を表示する．最後のパラメータで等高線数を選択．
- clabel で等高線の値をラベルに与える．パラメータ 'manual' はラベル付けする等高線を手で選択することを指示．

12.2　非線形微分方程式の例

ボルテラ系のような非線形微分方程式を，MATLAB を用いて数値的に議論してみましょう．

12.2.1　ボルテラ系

前の図 12.1を数値的に取り扱ってみましょう（図 12.2）．

```
ode=@(t,y)dydt0(t,y); % Define the Ord. Diff.Eq.
[t,y]=ode45(ode,[0,20],[1.0,1.8]); % Solver ode45
subplot(2,1,1);plot(t,y(:,1));
hold on
plot(t,y(:,2));
xlabel('t');ylabel('solution x and y');
txt2=sprintf('%.6f',1.0);
txt3=sprintf('%.6f',1.8)
title(['Lotka-Volterra equation: Starting Pt y0=(',txt2,' , ',txt3,')'])
legend('x(t)','y(t)')
hold off
subplot(2,1,2);plot(y(:,1),y(:,2));
xlim([0 2.0]);ylim([0 2.0])
xlabel('x(t)');ylabel('y(t)');
title(['Lotka-Volterra equation : Starting Pt y0=(',txt2,' , ',txt3,')'])
hold on
```

```
function dydt0 = dydt0(t,y)
%Lotka?Volterra equation
% dy1/dt=y1-y1*y2
% dy2/dt=y1*y2-y2
dydt0 = [y(1)-y(1)*y(2);y(1)*y(2)-y(2)]
```

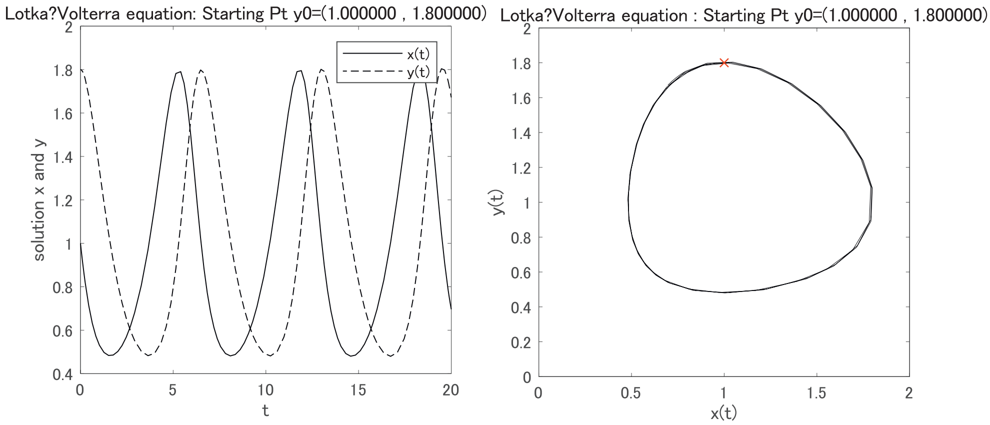

図 12.2: ボルテラ系（あるいはロトカ・ボルテラ方程式）

- 連立微分方程式（上段右に function dydt0 を定義）として関数を定義したのが 1 行目.
- 2 行目は微分方程式を解くスクリプトとして ODE45 を用いる．時間 t は (0,20) の範囲で変化させ，初期値は $x(0) = 1.0, y(0) = 1.8$ とする.
- １つの図に 2 つの図を収めることを，subplot コマンドで定義する．subplot$(2, 1, n)$ は，図を全体として 2 行 1 列に並べ，そのうちの n 番目の図として配する，という意味.

- txt2=sprintf($\cdots$), txt3=sprintf($\cdots$) で，図の説明に書き込む変数（初期値）を定義. sprintf はデータから書式を指定した文字列データに変換する.
- 第 1 の図は $x(t), y(t)$ を時間 t の関数として，第 2 の図は $(x(t), y(t))$ の軌道を表す．ここでは $x, y > 0$ の領域で，周期軌道（閉軌道）を表す．注意をして見れば，数値的誤差のために，軌道が完全には前の軌道をなぞっていないことが分かる.
- x（被捕食者=餌）の数が減少（増加）すると y（捕食者）の数が減少（増加）し出す．しばらくすると y の数が減った（増えた）ため，x（餌）の数が増加（減少）し出す．このようなサイクルが繰り返される.