幾何学 II

目次

1 距離空間の完備性

定義 1.1 (X,d) を距離空間とし, $(x_k)_{k\in\boldsymbol{N}}$ を X の点列とする. すべての項が自然数である狭義単調増加数列 $k_1 < k_2 < \cdots < k_i < k_{i+1} < \cdots$ に対し, $(x_{k_i})_{i\in\boldsymbol{N}}$ という形の点列を $(x_k)_{k\in\boldsymbol{N}}$ の部分列という.

(X,d) を距離空間とするとき, $p\in X$, $r>0$ に対し, $B_d(p\,;r) = \{x\in X \mid d(x,p) < r\}$ とおき, これを中心 p, 半径 r の開球という.

補題 1.2 (X,d) を距離空間とし, $(x_k)_{k\in\boldsymbol{N}}$ を X の点列とする. X の点 p が条件「任意の正の実数 ε に対して $\{k\in\boldsymbol{N} \mid d(x_k,p) < \varepsilon\}$ は有限集合ではない.」を満たすとき, $(x_k)_{k\in\boldsymbol{N}}$ の部分列で p に収束するものが存在する.

証明 すべての項が自然数からなる狭義単調増加数列 $k_1 < k_2 < \cdots < k_i < k_{i+1} < \cdots$ を以下のように帰納的に定める. 仮定から $\{k\in\boldsymbol{N} \mid d(x_k,p) < 1\}$ は有限集合ではないため, この集合から自然数 k_1 を一つ選べば $d(x_{k_1},p) < 1$ が成り立つ. 帰納的に, すべての項が自然数からなる狭義単調増加数列 $k_1 < k_2 < \cdots < k_i$ が定まり, $j = 1,2,\ldots,i$ に対して $d\bigl(x_{k_j},p\bigr) < \dfrac{1}{j}$ が成り立つと仮定する. 仮定から $\left\{k\in\boldsymbol{N} \,\middle|\, d(x_k,p) < \dfrac{1}{i+1}\right\}$ は有限集合ではないため, この集合は k_i より大きな自然数を含む. その一つを k_{i+1} とおけば $d\bigl(x_{k_{i+1}},p\bigr) < \dfrac{1}{i+1}$ が成り立つ. このとき X の点列 $(x_{k_i})_{i\in\boldsymbol{N}}$ は任意の自然数 i に対して $d(x_{k_i},p) < \dfrac{1}{i}$ を満たすため, p に収束する. □

補題 1.3 (X,d) を距離空間とし, $(x_k)_{k\in\boldsymbol{N}}$ を X の点列とする. X の任意の点 p に対して正の実数 ε_p で $\{k\in\boldsymbol{N} \mid d(x_k,p) < \varepsilon_p\}$ が有限集合になるものが存在すれば, 距離関数 d から定まる X の位相に関して, X はコンパクトではない.

証明 $p\in X$ に対して $B_d(p\,;\varepsilon_p)$ は p を含む x の開集合だから $X = \bigcup_{p\in X} B_d(p\,;\varepsilon_p)$ である. もし X がコンパクトならば $p_1,p_2,\ldots,p_N\in X$ で $X = B_d(p_1\,;\varepsilon_{p_1})\cup B_d(p_2\,;\varepsilon_{p_2})\cup\cdots\cup B_d(p_N\,;\varepsilon_{p_N})$ を満たすものが存在する. 仮定から, 各 $i=1,2,\ldots,N$ に対して $x_k\in B_d(p_i\,;\varepsilon_{p_i})$ となる自然数 k は有限個しかないため, すべての $i=1,2,\ldots,N$ に対して $x_n\notin B_d(p_i\,;\varepsilon_{p_i})$ となる自然数 n が存在するが, このことは $X = B_d(p_1\,;\varepsilon_{p_1})\cup B_d(p_2\,;\varepsilon_{p_2})\cup\cdots\cup B_d(p_N\,;\varepsilon_{p_N})$ であることと矛盾する. □

定理 1.4 (X,d) を距離空間とするとき, 距離関数 d から定まる X の位相に関して X がコンパクトならば X の任意の点列は収束する部分列をもつ.

証明 収束する部分列をもたない X の点列 $(x_k)_{k\in\boldsymbol{N}}$ が存在すると仮定すれば, 補題 1.2 の対偶から X の任意の点 p に対して正の実数 ε_p で, $\{k\in\boldsymbol{N} \mid d(x_k,p) < \varepsilon_p\}$ が有限集合になるものが存在する. このとき補題 1.3 から X はコンパクトではない. 従って X がコンパクトならば X の任意の点列は収束する部分列をもつ. □

n 次元ユークリッド空間 $\boldsymbol{R}^n$ の距離関数 d_n を $d_n(\boldsymbol{x},\boldsymbol{y}) = \|\boldsymbol{x}-\boldsymbol{y}\|$ で定義し, 距離空間 $(\boldsymbol{R}^n,d_n)$ を考える. $\boldsymbol{R}^n$ に距離関数 d_n から定まる位相を与えれば $\boldsymbol{R}^n$ の有界閉集合はコンパクトだから, 定理 1.4 から次の結果が得られる.

系 1.5 X が $\boldsymbol{R}^n$ の有界閉集合ならば, X の任意の点列は収束する部分列をもつ.

定義 1.6 (1) 距離空間 (X,d) の点列 $(x_k)_{k\in\boldsymbol{N}}$ が次の条件を満たすとき $(x_k)_{k\in\boldsymbol{N}}$ を (X,d) の Cauchy 列という.

任意の $\varepsilon>0$ に対し, 自然数 N で条件「$m,n\geqq N$ ならば $d(x_m,x_n) < \varepsilon$」を満たすものが存在する.

(2) 距離空間 (X,d) の任意の Cauchy 列が収束するとき, (X,d) は完備であるという.

(3) 距離空間 (X,d) の点列 $(x_k)_{k\in\boldsymbol{N}}$ に対し, $\{x_1,x_2,\ldots,x_k,\ldots\}$ が (X,d) の有界な部分集合であるとき, $(x_k)_{k\in\boldsymbol{N}}$ は有界であるという.

命題 1.7　収束する点列は Cauchy 列である.

証明　$(x_k)_{k\in\boldsymbol{N}}$ を p に収束する距離空間 (X,d) の点列とすれば, 任意の正の実数 ε に対して自然数 N で, 条件「$n \geqq N$ ならば $d(x_n,p) < \frac{\varepsilon}{2}$」を満たすものがある. $m,n \geqq N$ ならば, 三角不等式から $d(x_m,x_n) \leqq d(x_m,p)+d(p,x_n) = d(x_m,p)+d(x_n,p) < \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon$ より $(x_k)_{k\in\boldsymbol{N}}$ は Cauchy 列である. □

補題 1.8　Cauchy 列は有界である.

証明　$(x_k)_{k\in\boldsymbol{N}}$ を距離空間 (X,d) の Cauchy 列とする. Cauchy 列の定義により, 自然数 N で,「$m,n \geqq N$ ならば $d(x_m,x_n) < 1$」を満たすものがある. そこで, $d(x_2,x_1), d(x_3,x_1), \dots, d(x_N,x_1)$ のうちで最大のものを R とすれば, すべての $k \in \boldsymbol{N}$ に対して $x_k \in B_d(x_1\,;R+1)$ が成り立つ. 実際 $k \leqq N$ ならば $d(x_k,x_1) \leqq R < R+1$ であり, $k > N$ ならば, 三角不等式より $d(x_k,x_1) \leqq d(x_k,x_N)+d(x_N,x_1) < 1+R$ である. 故に $(x_k)_{k\in\boldsymbol{N}}$ は有界である.□

補題 1.9　Cauchy 列が収束する部分列を含めば, その Cauchy 列は収束する.

証明　$(x_k)_{k\in\boldsymbol{N}}$ を距離空間 (X,d) の Cauchy 列とし, $(x_{k_i})_{i\in\boldsymbol{N}}$ を $(x_k)_{k\in\boldsymbol{N}}$ の収束する部分列とする. $\lim_{i\to\infty} x_{k_i} = p$ とおくと, 任意の $\varepsilon > 0$ に対して自然数 N_1 で,「$i \geqq N_1$ ならば $d(x_{k_i},p) < \frac{\varepsilon}{2}$」を満たすものがある. また, $(x_k)_{k\in\boldsymbol{N}}$ が Cauchy 列であることから, 自然数 N_2 で条件「$m,n \geqq N_2$ ならば $d(x_m,x_n) < \frac{\varepsilon}{2}$」を満たすものがある. k_{N_1} と N_2 の大きい方を N とする. $k \geqq N$ ならば $k_N \geqq N \geqq N_2$ だから, 三角不等式により $d(x_k,p) \leqq d(x_k,x_{k_N})+d(x_{k_N},p) < \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon$ が得られる. 従って $(x_k)_{k\in\boldsymbol{N}}$ は収束する. □

系 1.5 の応用として, 次の定理を示す. この定理の主張は $\boldsymbol{R}^n$ の完備性と呼ばれる.

定理 1.10　$(\boldsymbol{R}^n, d_n)$ は完備距離空間である. すなわち, $\boldsymbol{R}^n$ の任意の Cauchy 列は収束する.

証明　$(\boldsymbol{x}_k)_{k\in\boldsymbol{N}}$ を $\boldsymbol{R}^n$ の Cauchy 列とする. 補題 1.8 により, $(\boldsymbol{x}_k)_{k\in\boldsymbol{N}}$ は有界であるため, 正の実数 r で, すべての $k \in \boldsymbol{N}$ に対して $\boldsymbol{x}_k \in B_n(\boldsymbol{0}\,;r)$ となるものがある. $\overline{B}_n(\boldsymbol{0}\,;r) = \{\boldsymbol{x} \in \boldsymbol{R}^n \,|\, \|\boldsymbol{x}\| \leqq r\}$ とおけば $B_n(\boldsymbol{0}\,;r) \subset \overline{B}_n(\boldsymbol{0}\,;r)$ であり, $\overline{B}_n(\boldsymbol{0}\,;r)$ は有界閉集合だから, 系 1.5 により $(\boldsymbol{x}_k)_{k\in\boldsymbol{N}}$ は収束する部分列を含む. 故に補題 1.9 により, $(\boldsymbol{x}_k)_{k\in\boldsymbol{N}}$ は収束する. □

定義 1.11　X を集合, (Y,d) を距離空間, $f: X \to Y$ を写像とする. 各 $n \in \boldsymbol{N}$ に対して写像 $f_n : X \to Y$ が与えられていて, 任意の $\varepsilon > 0$ に対して自然数 N で, 条件「$n \geqq N$ ならば, すべての $x \in X$ に対して $d(f_n(x), f(x)) < \varepsilon$」を満たすものがあるとき, 写像の列 $(f_n)_{n\in\boldsymbol{N}}$ は f に一様収束するという.

命題 1.12　$(X,\mathcal{O})$ を位相空間, (Y,d) を距離空間とする. X から Y への連続写像の列 $(f_n)_{n\in\boldsymbol{N}}$ が X から Y への写像 f に一様収束すれば, f は連続写像である.

証明　任意の $p \in X$ と $\varepsilon > 0$ に対し, 仮定から自然数 N で, 条件「$n \geqq N$ ならば, すべての $x \in X$ に対して $d(f_n(x), f(x)) < \dfrac{\varepsilon}{3}$」を満たすものがある. また, f_N の p における連続性から, p を含む開集合 U で条件「$x \in U$ ならば $d(f_N(x), f_N(p)) < \dfrac{\varepsilon}{3}$」を満たすものがある. 従って三角不等式から, $x \in U$ ならば

$$d(f(x), f(p)) \leqq d(f(x), f_N(x)) + d(f_N(x), f_N(p)) + d(f_N(p), f(p)) < \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3} = \varepsilon$$

が成り立つため, f は p において連続である. □

定義 1.13　(X,d) を距離空間とする. 集合 S から X への写像 f の像が有界であるとき, f は有界であるという.

命題 1.14　$(X,\mathcal{O})$ を位相空間, (Y,d) を距離空間とし, X から Y への有界な連続写像全体からなる集合を $\boldsymbol{B}(X,Y)$ で表す. 関数 $\hat{d}: \boldsymbol{B}(X,Y) \times \boldsymbol{B}(X,Y) \to \boldsymbol{R}$ を $\hat{d}(f,g) = \sup\{d(f(x),g(x)) \,|\, x \in X\}$ で定義すれば, $\hat{d}$ は $\boldsymbol{B}(X,Y)$ の距離関数であり, (Y,d) が完備距離空間ならば $(\boldsymbol{B}(X,Y), \hat{d})$ も完備距離空間である.

証明 $f, g \in \boldsymbol{B}(X, Y)$ ならば $p, q \in Y$ と $K, L > 0$ で, すべての $x \in X$ に対して $d(f(x), p) \leqq K$, $d(g(x), q) \leqq L$ を満たすものがある．三角不等式から $d(f(x), g(x)) \leqq d(f(x), p) + d(p, q) + d(q, g(x)) \leqq K + L + d(p, q)$ が成り立つため, $K + L + d(p, q)$ は集合 $\{d(f(x), g(x)) \,|\, x \in X\}$ の上界である．故に幾何学 I 定理 10.12 より $\sup\{d(f(x), g(x)) \,|\, x \in X\}$ は存在する．$\hat{d}(f, g) = \hat{d}(g, f)$ と $\hat{d}(f, g) \geqq 0$ が成り立つことは明らかである．$\hat{d}(f, g) = 0$ ならばすべての $x \in X$ に対して $d(f(x), g(x)) \leqq \hat{d}(f, g) = 0$ だから $f = g$ である．$f, g, h \in \boldsymbol{B}(X, Y)$ ならば, 任意の $x \in X$ に対して三角不等式から $d(f(x), h(x)) \leqq d(f(x), g(x)) + d(g(x), h(x)) \leqq \hat{d}(f, g) + \hat{d}(g, h)$ が成り立つため, $\hat{d}(f, h) \leqq \hat{d}(f, g) + \hat{d}(g, h)$ が得られる．従って $\hat{d}$ は $\boldsymbol{B}(X, Y)$ の距離関数である．

$(f_n)_{n \in \boldsymbol{N}}$ を $(\boldsymbol{B}(X, Y), \hat{d})$ の Cauchy 列とする．任意の $\varepsilon > 0$ に対し, 自然数 N で条件「$m, n \geqq N$ ならば $\hat{d}(f_m, f_n) < \dfrac{\varepsilon}{2}$」を満たすものがある．このとき, 任意の $x \in X$ に対して $d(f_m(x), f_n(x)) \leqq \hat{d}(f_m, f_n) < \dfrac{\varepsilon}{2} < \varepsilon$ だから $(f_n(x))_{n \in \boldsymbol{N}}$ は Y の Cauchy 列である．従って仮定から $(f_n(x))_{n \in \boldsymbol{N}}$ は収束するため, 写像 $f : X \to Y$ が $f(x) = \lim\limits_{n \to \infty} f_n(x)$ で定義できる．$m, n \geqq N$ ならばすべての $x \in X$ に対して成り立つ不等式 $d(f_m(x), f_n(x)) < \dfrac{\varepsilon}{2}$ において $n \to \infty$ とすれば, $d(f_m(x), f(x)) \leqq \dfrac{\varepsilon}{2} < \varepsilon$ がすべての $x \in X$ に対して成り立つため, $(f_n)_{n \in \boldsymbol{N}}$ は f に一様収束する．よって命題 1.12 から f は連続写像である．

上でとくに $m = N$ の場合を考えれば, $d(f_N(x), f(x)) < \varepsilon$ がすべての $x \in X$ に対して成り立つ．また f_N が有界であることから, $q \in Y$ と $K > 0$ で, すべての $x \in X$ に対して $d(f_N(x), q) < K$ を満たすものが存在するため, $d(f(x), q) \leqq d(f(x), f_N(x)) + d(f_N(x), q) < \varepsilon + K$ となり, f も有界であることがわかる．さらに, $m \geqq N$ ならばすべての $x \in X$ に対して $d(f_m(x), f(x)) \leqq \dfrac{\varepsilon}{2}$ が成り立つため, $\hat{d}(f_m, f) \leqq \dfrac{\varepsilon}{2} < \varepsilon$ である．これは, $(\boldsymbol{B}(X, Y), \hat{d})$ の Cauchy 列 $(f_n)_{n \in \boldsymbol{N}}$ が $\boldsymbol{B}(X, Y)$ の点 f に収束することを意味する． □

命題 1.15 $(X, \mathcal{O})$ を位相空間, (Y, d) を完備距離空間とし, V を中心 $y_0 \in Y$ 半径 $\beta > 0$ の Y における開球 $B_d(y_0\,; \beta)$ とする．連続写像 $v : X \times V \to Y$ に対し, $0 \leqq k < 1$ である k が存在して, すべての $x \in X$, $y_1, y_2 \in V$ に対し, $d(v(x, y_1), v(x, y_2)) \leqq k d(y_1, y_2)$ および $d(v(x, y_0), y_0) < \beta(1 - k)$ が成り立つとする．このとき, 写像 $f : X \to V$ ですべての $x \in X$ に対し, $f(x) = v(x, f(x))$ を満たすものがただ一つ存在する．さらに, この写像は連続であり, ある $x_0 \in X$ に対し $v(x_0, y_0) = y_0$ ならば $f(x_0) = y_0$ である．

証明 各 $x \in X$ に対し, Y の点列 $(f_n(x))_{n \in \boldsymbol{N}}$ を帰納的に $f_0(x) = y_0$, $f_n(x) = v(x, f_{n-1}(x))$ $(n \geqq 1)$ により定める．このとき $f_0(x) = y_0 \in V$ であり, 帰納的に $f_0(x), \ldots, f_{n-1}(x) \in V$ であると仮定すれば,

$$d(f_m(x), f_{m-1}(x)) = d(v(x, f_{m-1}(x)), v(x, f_{m-2}(x))) \leqq k d(f_{m-1}(x), f_{m-2}(x))$$

より $d(f_m(x), f_{m-1}(x)) \leqq k^{m-1} d(f_1(x), y_0)$ を得る．従って, 三角不等式と仮定から $1 \leqq l \leqq n - m$ に対して

$$d(f_{m+l}(x), f_m(x)) \leqq \sum_{i=1}^{l} d(f_{m+i}(x), f_{m+i-1}(x)) \leqq \left(\sum_{i=1}^{l} k^{m+i-1}\right) d(f_1(x), y_0) = \frac{k^m(1 - k^l) d(v(x, y_0), y_0)}{1 - k}$$

より, 次の不等式 $(*)$ が成り立つ．

$$d(f_{m+l}(x), f_m(x)) \leqq \frac{k^m(1 - k^l) d(v(x, y_0), y_0)}{1 - k} < \beta k^m (1 - k^l) \cdots (*)$$

$m = 0$ とすれば $(*)$ から $d(f_l(x), y_0) < \beta(1 - k^l) \leqq \beta$ だから $f_l(x) \in V$ である．従って x を $f_n(x)$ に対応させる写像 f_n は有界であり, v の連続性から帰納的に f_n が連続であることがわかる．$0 \leqq k < 1$ で, すべての $x \in X$ に対して $d(f_{m+n}(x), f_m(x)) < \beta k^m$ が成り立つため, 写像の列 $(f_n)_{n \in \boldsymbol{N}}$ は距離空間 $(\boldsymbol{B}(X, Y), \hat{d})$ の Cauchy 列である．故に命題 1.14 から $(f_n)_{n \in \boldsymbol{N}}$ は連続写像に収束して, その極限を f とする．不等式 $(*)$ で $m = 0$, $l \to \infty$ とすれば $d(f(x), y_0) \leqq \dfrac{d(v(x, y_0), y_0)}{1 - k} < \beta$ が得られるため, $f(x) \in V$ である．従って f を X から V への写像とみなすことができる．

$f_n(x) = v(x, f_{n-1}(x))$ から $n \to \infty$ として $f(x) = v(x, f(x))$ を得る．$v(x_0, y_0) = y_0$ であれば, $f_0(x_0) = y_0$ より, n による帰納法ですべての n に対して $f_n(x_0) = y_0$ となることは容易に示されるため $f(x_0) = y_0$ である．

$y_1, y_2 \in V$ が $y_1 = v(x, y_1)$, $y_2 = v(x, y_2)$ を満たせば, $d(y_1, y_2) = d(v(x, y_1), v(x, y_2)) \leqq kd(y_1, y_2)$ より, $(1-k)d(y_1, y_2) \leqq 0$. $k < 1$ だから $d(y_1, y_2) = 0$ となるため $y_1 = y_2$ である. 従って, 各 $x \in X$ に対し, $f(x) = v(x, f(x))$ を満たす $f(x)$ は一意的に定まる. □

とくに, 上において X が 1 点からなる位相空間の場合には次の結果を得る.

命題 1.16 (不動点定理) (Y, d) を完備距離空間, $V = B_d(y_0\,; \beta)$ とする. 写像 $v : V \to Y$ に対して $0 \leqq k < 1$ を満たす k が存在して, 任意の $y_1, y_2 \in V$ に対し $d(v(y_1), v(y_2)) \leqq kd(y_1, y_2)$ かつ $d(v(y_0), y_0) < \beta(1-k)$ が成り立つとする. このとき, 点 $x \in V$ で $x = v(x)$ を満たすものがただ一つ存在する.

2 写像の微分

$m \times n$ 行列 $A = (a_{ij})$ に対して $\|A\| = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2}$ とおく.

命題 2.1 $m \times n$ 行列 A と $n \times k$ 行列 B に対して $\|AB\| \leqq \|A\|\|B\|$ が成り立つ.

証明 $A = (a_{ij})$, $B = (b_{ij})$ とし, A の第 i 行の成分を縦に並べて得られるベクトルを $\boldsymbol{a}_i$, B の第 j 列を $\boldsymbol{b}_j$ とすると, AB の (i, j) 成分が $(\boldsymbol{a}_i, \boldsymbol{b}_j)$ であることと Schwarz の不等式から $|(\boldsymbol{a}_i, \boldsymbol{b}_j)|^2 \leqq \|\boldsymbol{a}_i\|^2 \|\boldsymbol{b}_j\|^2$ だから

$$\begin{aligned} \|AB\|^2 &= \sum_{i=1}^{m} \sum_{j=1}^{k} |(\boldsymbol{a}_i, \boldsymbol{b}_j)|^2 \leqq \sum_{i=1}^{m} \sum_{j=1}^{k} \|\boldsymbol{a}_i\|^2 \|\boldsymbol{b}_j\|^2 = \left(\sum_{i=1}^{m} \|\boldsymbol{a}_i\|^2\right)\left(\sum_{j=1}^{k} \|\boldsymbol{b}_j\|^2\right) \\ &= \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2\right)\left(\sum_{j=1}^{k} \sum_{l=1}^{n} |b_{lj}|^2\right) = \|A\|^2 \|B\|^2 \end{aligned}$$

が得られる. $\|AB\|^2 \leqq \|A\|^2 \|B\|^2$ の辺の正の平方根をとれば $\|AB\| \leqq \|A\|\|B\|$ が得られる. □

$\boldsymbol{R}^n$ の中心 $\boldsymbol{p}$, 半径 r の開球 $\{\boldsymbol{x} \in \boldsymbol{R}^n \,|\, d_n(\boldsymbol{x}, \boldsymbol{y}) < r\}$ を以後 $B_n(\boldsymbol{p}\,; r)$ で表す. この節ではとくに断らない限り X, Y はそれぞれ $\boldsymbol{R}^n$, $\boldsymbol{R}^m$ の部分集合で, f は X から Y への写像である.

命題 2.2 $\boldsymbol{p} \in X$ に対し, 正の実数 r, L で条件「$\boldsymbol{x} \in B_n(\boldsymbol{p}\,; r) \cap X$ ならば $\|f(\boldsymbol{x}) - f(\boldsymbol{p})\| \leqq L\|\boldsymbol{x} - \boldsymbol{p}\|$」を満たすものがあれば, f は $\boldsymbol{p}$ で連続である.

証明 任意の $\varepsilon > 0$ に対し $\delta = \min\{r, \frac{\varepsilon}{L}\}$ とおけば, $\boldsymbol{x} \in B_n(\boldsymbol{p}\,; \delta) \cap X$ ならば $\|f(\boldsymbol{x}) - f(\boldsymbol{p})\| \leqq L\|\boldsymbol{x} - \boldsymbol{p}\| < L\delta \leqq \varepsilon$ だから f は $\boldsymbol{p}$ で連続である. □

定義 2.3 $\boldsymbol{p}$ を X の内点とする. $m \times n$ 行列 A で, 次の等式 (D) を満たすものがあるとき f は $\boldsymbol{p}$ で微分可能であるという.

$$\lim_{\boldsymbol{x} \to \boldsymbol{p}} \frac{f(\boldsymbol{x}) - f(\boldsymbol{p}) - A(\boldsymbol{x} - \boldsymbol{p})}{\|\boldsymbol{x} - \boldsymbol{p}\|} = \boldsymbol{0} \quad \cdots \ (D)$$

以後, 「f は $\boldsymbol{p}$ で微分可能である.」というときは $\boldsymbol{p}$ は f の定義域 X の内点であることを仮定する.

写像 f, $m \times n$ 行列 A, $\boldsymbol{p} \in X$ に対して, 写像 $\varepsilon = \varepsilon_{f,A,\boldsymbol{p}} : X \to \boldsymbol{R}^m$ を

$$\varepsilon_{f,A,\boldsymbol{p}}(\boldsymbol{x}) = \begin{cases} \dfrac{f(\boldsymbol{x}) - f(\boldsymbol{p}) - A(\boldsymbol{x} - \boldsymbol{p})}{\|\boldsymbol{x} - \boldsymbol{p}\|} & \boldsymbol{x} \neq \boldsymbol{p} \\ \boldsymbol{0} & \boldsymbol{x} = \boldsymbol{p} \end{cases}$$

で定義すれば, この定義と定義 2.3 から次のことがわかる.

命題 2.4　任意の $\boldsymbol{x} \in X$ に対して, 次の等式が成り立つ.

$$f(\boldsymbol{x}) = f(\boldsymbol{p}) + A(\boldsymbol{x}-\boldsymbol{p}) + \|\boldsymbol{x}-\boldsymbol{p}\|\varepsilon_{f,A,\boldsymbol{p}}(\boldsymbol{x})$$

f が $\boldsymbol{p}$ で微分可能であるためには, $\varepsilon_{f,A,\boldsymbol{p}}$ が $\boldsymbol{p}$ において連続, すなわち $\lim\limits_{\boldsymbol{x}\to\boldsymbol{p}} \varepsilon_{f,A,\boldsymbol{p}}(\boldsymbol{x}) = \boldsymbol{0}$ となるような, $m \times n$ 行列 A が存在することが必要十分である.

命題 2.5　f が $\boldsymbol{p} \in X$ で微分可能ならば $r, L > 0$ で $B_n(\boldsymbol{p}\,;r) \subset X$ かつ「$\boldsymbol{x} \in B_n(\boldsymbol{p}\,;r)$ ならば $\|f(\boldsymbol{x}) - f(\boldsymbol{p})\| \leqq L\|\boldsymbol{x}-\boldsymbol{p}\|$」を満たすものがある. 従って命題 2.2 から f は $\boldsymbol{p}$ で連続である.

証明　$m \times n$ 行列 A は定義 2.3 の (D) を満たすとする. 命題 2.4, 三角不等式および命題 2.1 から

$$\|f(\boldsymbol{x})-f(\boldsymbol{p})\| = \|A(\boldsymbol{x}-\boldsymbol{p})+\|\boldsymbol{x}-\boldsymbol{p}\|\varepsilon_{f,A,\boldsymbol{p}}(\boldsymbol{x})\| \leqq \|A(\boldsymbol{x}-\boldsymbol{p})\|+\|\boldsymbol{x}-\boldsymbol{p}\|\|\varepsilon_{f,A,\boldsymbol{p}}(\boldsymbol{x})\| \leqq (\|A\| + \|\varepsilon_{f,A,\boldsymbol{p}}(\boldsymbol{x})\|)\|\boldsymbol{x}-\boldsymbol{p}\|$$

である. 仮定と命題 2.4 から $\lim\limits_{\boldsymbol{x}\to\boldsymbol{p}} \|\varepsilon_{f,A,\boldsymbol{p}}(\boldsymbol{x})\| = 0$ だから, $r > 0$ で $B_n(\boldsymbol{p}\,;r) \subset X$ かつ「$\boldsymbol{x} \in B_n(\boldsymbol{p}\,;r)$, $\boldsymbol{x} \neq \boldsymbol{p}$ ならば $\|\varepsilon_{f,A,\boldsymbol{p}}(\boldsymbol{x})\| < 1$」を満たすものがとれる. 故に上式から $\boldsymbol{x} \in B_n(\boldsymbol{p}\,;r)$ ならば $\|f(\boldsymbol{x}) - f(\boldsymbol{p})\| < (\|A\|+1)\|\boldsymbol{x}-\boldsymbol{p}\|$ だから $L = \|A\| + 1$ とすればよい. □

定義 2.6　$\boldsymbol{p} \in X$, $\boldsymbol{v} \in \boldsymbol{R}^n$ とし, 十分小さな $r > 0$ に対して $|t| < r$ ならば $\boldsymbol{p} + t\boldsymbol{v} \in X$ であるとする. 極限

$$\lim_{t\to 0} \frac{f(\boldsymbol{p}+t\boldsymbol{v}) - f(\boldsymbol{p})}{t}$$

が存在するとき, f は $\boldsymbol{p}$ において $\boldsymbol{v}$ 方向に微分可能であるといい, この極限のベクトルを f の $\boldsymbol{p}$ における $\boldsymbol{v}$ 方向の微分という. とくに $Y = \boldsymbol{R}$, $\boldsymbol{v} = \boldsymbol{e}_j$ = ($\boldsymbol{R}^n$ の j 番目の基本ベクトル) の場合, 上の極限値を $\dfrac{\partial f}{\partial x_j}(\boldsymbol{p})$ で表す. これを f の $\boldsymbol{p}$ における j 番目の変数に関する偏微分係数といい, このとき f は j 番目の変数に関して $\boldsymbol{p}$ において偏微分可能であるという. さらに f が X の各点で j 番目の変数に関して偏微分可能なとき, $\boldsymbol{p} \in X$ を $\dfrac{\partial f}{\partial x_j}(\boldsymbol{p})$ に対応させる関数を j 番目の変数に関する偏導関数と呼んで $\dfrac{\partial f}{\partial x_j} : X \to \boldsymbol{R}$ で表す.

命題 2.7　f が $\boldsymbol{p}$ で微分可能なとき, 任意の $\boldsymbol{v} \in \boldsymbol{R}^n$ に対して f は $\boldsymbol{p}$ において $\boldsymbol{v}$ 方向に微分可能である. このとき定義 2.3 の等式 (D) における $m \times n$ 行列 A は

$$\lim_{t\to 0} \frac{f(\boldsymbol{p}+t\boldsymbol{v}) - f(\boldsymbol{p})}{t} = A\boldsymbol{v}$$

を満たす. とくに A の第 j 列は $\lim\limits_{t\to 0} \dfrac{f(\boldsymbol{p}+t\boldsymbol{e}_j) - f(\boldsymbol{p})}{t}$ で与えられるため定義 2.3 の等式 (D) を満たす行列 A は存在すればただ 1 つだけである.

証明　命題 2.4 の等式に $\boldsymbol{x} = \boldsymbol{p} + t\boldsymbol{v}$ を代入すれば,

$$f(\boldsymbol{p}+t\boldsymbol{v}) = f(\boldsymbol{p}) + tA\boldsymbol{v} + |t|\|\boldsymbol{v}\|\varepsilon_{f,A,\boldsymbol{p}}(\boldsymbol{p}+t\boldsymbol{v})$$

となるため, 次の等式が得られる.

$$\lim_{t\to 0} \frac{f(\boldsymbol{p}+t\boldsymbol{v}) - f(\boldsymbol{p})}{t} = A\boldsymbol{v} + \lim_{t\to 0} \frac{|t|}{t}\|\boldsymbol{v}\|\varepsilon_{f,A,\boldsymbol{p}}(\boldsymbol{p}+t\boldsymbol{v}) \cdots \ (*)$$

$\left\|\dfrac{|t|}{t}\|\boldsymbol{v}\|\varepsilon_{f,A,\boldsymbol{p}}(\boldsymbol{p}+t\boldsymbol{v})\right\| = \|\boldsymbol{v}\|\,\|\varepsilon_{f,A,\boldsymbol{p}}(\boldsymbol{p}+t\boldsymbol{v})\|$ であり, 仮定と命題 2.4 から $\lim\limits_{t\to 0}\|\varepsilon_{f,A,\boldsymbol{p}}(\boldsymbol{p}+t\boldsymbol{v})\| = 0$ だから $(*)$ から $\lim\limits_{t\to 0} \dfrac{f(\boldsymbol{p}+t\boldsymbol{v}) - f(\boldsymbol{p})}{t} = A\boldsymbol{v}$ が得られる. □

定義 2.8　命題 2.7 から定義 2.3 の等式 (D) を満たす行列 A は f と $\boldsymbol{p}$ を与えればただ 1 つに定まるため, これを $f'(\boldsymbol{p})$ で表して, f の $\boldsymbol{p}$ における微分という.

例 2.9　実数を成分とする $m \times n$ 行列 A と $\boldsymbol{b} \in \boldsymbol{R}^m$ に対して写像 $f : \boldsymbol{R}^n \to \boldsymbol{R}^m$ を $f(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$ で定めると, $\displaystyle\lim_{\boldsymbol{x}\to\boldsymbol{p}} \frac{f(\boldsymbol{x}) - f(\boldsymbol{p}) - A(\boldsymbol{x}-\boldsymbol{p})}{\|\boldsymbol{x}-\boldsymbol{p}\|} = \lim_{\boldsymbol{x}\to\boldsymbol{p}} \frac{(A\boldsymbol{x}+\boldsymbol{b}) - (A\boldsymbol{p}+\boldsymbol{b}) - A(\boldsymbol{x}-\boldsymbol{p})}{\|\boldsymbol{x}-\boldsymbol{p}\|} = \boldsymbol{0}$ だから微分の定義によって f の $\boldsymbol{p} \in \boldsymbol{R}^n$ における微分 $f'(\boldsymbol{p})$ は $f'(\boldsymbol{p}) = A$ で与えられる.

命題 2.10　写像 $f : X \to Y$ に対し, $\boldsymbol{x} \in X$ を $f(\boldsymbol{x}) \in Y$ の第 i 成分に対応させる X で定義された実数値関数を $f_i : X \to \boldsymbol{R}$ とする.

(1) f が $\boldsymbol{p}$ で微分可能ならば, 各 f_i は $\boldsymbol{p}$ で微分可能で, $f'(\boldsymbol{p})$ の (i,j) 成分は $\dfrac{\partial f_i}{\partial x_j}(\boldsymbol{p})$ である.

(2) 逆に各 f_i が $\boldsymbol{p}$ で微分可能ならば f は $\boldsymbol{p}$ で微分可能である.

証明　(1) $f'(\boldsymbol{p})$ の第 i 行を A_i とすると $\dfrac{f(\boldsymbol{x}) - f(\boldsymbol{p}) - f'(\boldsymbol{p})(\boldsymbol{x}-\boldsymbol{p})}{\|\boldsymbol{x}-\boldsymbol{p}\|}$ の第 i 成分は $\dfrac{f_i(\boldsymbol{x}) - f_i(\boldsymbol{p}) - A_i(\boldsymbol{x}-\boldsymbol{p})}{\|\boldsymbol{x}-\boldsymbol{p}\|}$ だから

$$\left|\frac{f_i(\boldsymbol{x}) - f_i(\boldsymbol{p}) - A_i(\boldsymbol{x}-\boldsymbol{p})}{\|\boldsymbol{x}-\boldsymbol{p}\|}\right| \leqq \left\|\frac{f(\boldsymbol{x}) - f(\boldsymbol{p}) - f'(\boldsymbol{p})(\boldsymbol{x}-\boldsymbol{p})}{\|\boldsymbol{x}-\boldsymbol{p}\|}\right\|$$

が成り立つ. $\boldsymbol{x} \to \boldsymbol{p}$ のとき, 上式の右辺は 0 に近づくため $\displaystyle\lim_{\boldsymbol{x}\to\boldsymbol{p}} \frac{f_i(\boldsymbol{x}) - f_i(\boldsymbol{p}) - A_i(\boldsymbol{x}-\boldsymbol{p})}{\|\boldsymbol{x}-\boldsymbol{p}\|} = 0$ となり, f_i は $\boldsymbol{p}$ で微分可能で $f_i'(\boldsymbol{p}) = A_i$ である. A の (i,j) 成分は A_i の第 j 列だから命題 2.7 と偏微分の定義から $A_i\boldsymbol{e}_j = \dfrac{\partial f_i}{\partial x_j}(\boldsymbol{p})$ に等しくなる.

(2) A を $f_i'(\boldsymbol{p})$ を第 i 行とする $m \times n$ 行列とすれば, m 次元ベクトル $\dfrac{f(\boldsymbol{x}) - f(\boldsymbol{p}) - A(\boldsymbol{x}-\boldsymbol{p})}{\|\boldsymbol{x}-\boldsymbol{p}\|}$ の第 i 成分は $\dfrac{f_i(\boldsymbol{x}) - f_i(\boldsymbol{p}) - f_i'(\boldsymbol{p})(\boldsymbol{x}-\boldsymbol{p})}{\|\boldsymbol{x}-\boldsymbol{p}\|}$ だから 仮定と幾何学 I 命題 2.26 から $\displaystyle\lim_{\boldsymbol{x}\to\boldsymbol{p}} \frac{f(\boldsymbol{x}) - f(\boldsymbol{p}) - A(\boldsymbol{x}-\boldsymbol{p})}{\|\boldsymbol{x}-\boldsymbol{p}\|} = \boldsymbol{0}$ である. □

定理 2.11　(合成写像の微分法) X, Y, Z をそれぞれ $\boldsymbol{R}^n, \boldsymbol{R}^m, \boldsymbol{R}^l$ の部分集合とする. $f : X \to Y$ が $\boldsymbol{p}$ で微分可能であり, $g : Y \to Z$ が $f(\boldsymbol{p})$ で微分可能ならば, 合成写像 $g \circ f : X \to Z$ も $\boldsymbol{p}$ で微分可能で, 次の等式が成り立つ.

$$(g \circ f)'(\boldsymbol{p}) = g'(f(\boldsymbol{p}))f'(\boldsymbol{p})$$

証明　写像 $\varepsilon_{f,f'(\boldsymbol{p}),\boldsymbol{p}}$, $\varepsilon_{g,g'(f(\boldsymbol{p})),f(\boldsymbol{p})}$ をそれぞれ $\varepsilon_{f,\boldsymbol{p}}$, $\varepsilon_{g,f(\boldsymbol{p})}$ で表すことにすれば, 命題 2.4 から

$$\begin{aligned} f(\boldsymbol{x}) &= f(\boldsymbol{p}) + f'(\boldsymbol{p})(\boldsymbol{x}-\boldsymbol{p}) + \|\boldsymbol{x}-\boldsymbol{p}\|\varepsilon_{f,\boldsymbol{p}}(\boldsymbol{x}) \ \cdots\ (i) \\ g(\boldsymbol{y}) &= g(f(\boldsymbol{p})) + g'(f(\boldsymbol{p}))(\boldsymbol{y}-f(\boldsymbol{p})) + \|\boldsymbol{y}-f(\boldsymbol{p})\|\varepsilon_{g,f(\boldsymbol{p})}(\boldsymbol{y}) \ \cdots\ (ii) \end{aligned}$$

が成り立つ. (ii) の等式の $\boldsymbol{y}$ に $f(\boldsymbol{x})$ を代入すれば次の等式が得られる.

$$g(f(\boldsymbol{x})) = g(f(\boldsymbol{p})) + g'(f(\boldsymbol{p}))(f(\boldsymbol{x}) - f(\boldsymbol{p})) + \|f(\boldsymbol{x}) - f(\boldsymbol{p})\|\varepsilon_{g,f(\boldsymbol{p})}(f(\boldsymbol{x}))$$

$\boldsymbol{x} \neq \boldsymbol{p}$ として, この等式の右辺の第 2 項の $f(\boldsymbol{x})$ に (i) の右辺を代入して, 両辺を $\|\boldsymbol{x}-\boldsymbol{p}\|$ で割って整理すれば

$$\frac{(g\circ f)(\boldsymbol{x}) - (g\circ f)(\boldsymbol{p}) - g'(f(\boldsymbol{p}))f'(\boldsymbol{p})(\boldsymbol{x}-\boldsymbol{p})}{\|\boldsymbol{x}-\boldsymbol{p}\|} = g'(f(\boldsymbol{p}))\varepsilon_{f,\boldsymbol{p}}(\boldsymbol{x}) + \frac{\|f(\boldsymbol{x}) - f(\boldsymbol{p})\|}{\|\boldsymbol{x}-\boldsymbol{p}\|}\varepsilon_{g,f(\boldsymbol{p})}(f(\boldsymbol{x})) \ \cdots\ (iii)$$

が得られる. 従って $\boldsymbol{x} \to \boldsymbol{p}$ としたときに, 上式の右辺が $\boldsymbol{0}$ に近づくことが示されれば, $g \circ f : X \to Z$ は $\boldsymbol{p}$ で微分可能で, 定義 2.8 により, $(g\circ f)'(\boldsymbol{p}) = g'(f(\boldsymbol{p}))f'(\boldsymbol{p})$ であることがわかる.

三角不等式から (iii) の右辺の長さについて, 次の不等式が成り立つ.

$$\left\|g'(f(\boldsymbol{p}))\varepsilon_{f,\boldsymbol{p}}(\boldsymbol{x}) + \frac{\|f(\boldsymbol{x}) - f(\boldsymbol{p})\|}{\|\boldsymbol{x}-\boldsymbol{p}\|}\varepsilon_{g,f(\boldsymbol{p})}(f(\boldsymbol{x}))\right\| \leqq \|g'(f(\boldsymbol{p}))\varepsilon_{f,\boldsymbol{p}}(\boldsymbol{x})\| + \frac{\|f(\boldsymbol{x}) - f(\boldsymbol{p})\|}{\|\boldsymbol{x}-\boldsymbol{p}\|}\|\varepsilon_{g,f(\boldsymbol{p})}(f(\boldsymbol{x}))\| \ \cdots\ (iv)$$

$A = g'(f(\boldsymbol{p}))$ として命題 2.1 を用いると $\|g'(f(\boldsymbol{p}))\varepsilon_{f,\boldsymbol{p}}(\boldsymbol{x})\| \leqq \|A\|\|\varepsilon_{f,\boldsymbol{p}}(\boldsymbol{x})\|$ であり, $\displaystyle\lim_{\boldsymbol{x}\to\boldsymbol{p}} \varepsilon_{f,\boldsymbol{p}}(\boldsymbol{x}) = \boldsymbol{0}$ と命題 2.4 から $\boldsymbol{x}$ が $\boldsymbol{p}$ に近づくとき, この不等式の右辺は 0 に近づくため

$$\lim_{\boldsymbol{x}\to\boldsymbol{p}} \|g'(f(\boldsymbol{p}))\varepsilon_{f,\boldsymbol{p}}(\boldsymbol{x})\| = 0 \ \cdots\ (v)$$

である．命題 2.5 から, $r, L > 0$ で $B_n(\boldsymbol{p}\,; r) \subset X$ かつ $\|\boldsymbol{x}-\boldsymbol{p}\| < r$ ならば $\|f(\boldsymbol{x})-f(\boldsymbol{p})\| \leqq L\|\boldsymbol{x}-\boldsymbol{p}\|$ を満たすものがあるため, $0 < \|\boldsymbol{x}-\boldsymbol{p}\| < r$ ならば次の不等式が成り立つ.

$$\frac{\|f(\boldsymbol{x})-f(\boldsymbol{p})\|}{\|\boldsymbol{x}-\boldsymbol{p}\|}\left\|\varepsilon_{g,f(\boldsymbol{p})}(f(\boldsymbol{x}))\right\| \leqq L\left\|\varepsilon_{g,f(\boldsymbol{p})}(f(\boldsymbol{x}))\right\| \cdots (vi)$$

命題 2.5 から, f は $\boldsymbol{p}$ で連続だから, $\lim_{\boldsymbol{x}\to\boldsymbol{p}} f(\boldsymbol{x}) = f(\boldsymbol{p})$ が成り立つため, $\lim_{\boldsymbol{y}\to f(\boldsymbol{p})} \varepsilon_{g,f(\boldsymbol{p})}(\boldsymbol{y}) = \varepsilon_{g,f(\boldsymbol{p})}(f(\boldsymbol{p})) = \boldsymbol{0}$ と幾何学 I 命題 4.15 から, $\lim_{\boldsymbol{x}\to\boldsymbol{p}} \varepsilon_{g,f(\boldsymbol{p})}(f(\boldsymbol{x})) = \boldsymbol{0}$ である．従って $\boldsymbol{x}\to\boldsymbol{p}$ のとき, (vi) の右辺は 0 に近づくため,

$$\lim_{\boldsymbol{x}\to\boldsymbol{p}}\frac{\|f(\boldsymbol{x})-f(\boldsymbol{p})\|}{\|\boldsymbol{x}-\boldsymbol{p}\|}\left\|\varepsilon_{g,f(\boldsymbol{p})}(f(\boldsymbol{x}))\right\| = 0 \cdots (vii)$$

が成り立つ. (v), (vii) と (iv) から, $\boldsymbol{x}\to\boldsymbol{p}$ のとき, (iii) の右辺は $\boldsymbol{0}$ に近づくため, 主張が示された. □

定理 2.11 で $Z = \boldsymbol{R}$ の場合を考える．$\boldsymbol{x} \in X$ を $f(\boldsymbol{x}) \in Y$ の第 i 成分に対応させる X で定義された実数値関数 $f_i : X \to \boldsymbol{R}$ を考えれば, 命題 2.10 の (1) から $m \times n$ 行列 $f'(\boldsymbol{p})$ の (i,j) 成分は $\dfrac{\partial f_i}{\partial x_j}(\boldsymbol{p})$ であり, $1 \times m$ 行列 $g'(f(\boldsymbol{p}))$ の $(1,j)$ 成分は $\dfrac{\partial g}{\partial y_j}(f(\boldsymbol{p}))$ である．一方 $(g\circ f)'(\boldsymbol{p})$ の $(1,j)$ 成分は $\dfrac{\partial(g\circ f)}{\partial x_j}(\boldsymbol{p})$ だから, 定理 2.11 で示した等式 $(g\circ f)'(\boldsymbol{p}) = g'(f(\boldsymbol{p}))f'(\boldsymbol{p})$ の両辺の $(1,j)$ 成分を比較すれば次の結果が得られる.

系 2.12 $f : X \to Y$ が $\boldsymbol{p}$ で微分可能であり, $g : Y \to \boldsymbol{R}$ が $f(\boldsymbol{p})$ で微分可能ならば, 次の等式が成り立つ.

$$\frac{\partial(g\circ f)}{\partial x_j}(\boldsymbol{p}) = \sum_{k=1}^{m}\frac{\partial g}{\partial y_k}(f(\boldsymbol{p}))\frac{\partial f_k}{\partial x_j}(\boldsymbol{p})$$

3 偏微分

補題 3.1 $x_1, x_2, \ldots, x_n \in \boldsymbol{R}$ に対し, 不等式 $|x_1| + |x_2| + \cdots + |x_n| \leqq \sqrt{n}\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ が成り立つ.

証明 $\left(\sqrt{n}\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}\right)^2 - (|x_1| + |x_2| + \cdots + |x_n|)^2 = \sum_{1\leqq i<j\leqq n}(|x_i| - |x_j|)^2 \geqq 0$ より結果が得られる. □

この節ではとくに断らない限り X, Y はそれぞれ $\boldsymbol{R}^n$, $\boldsymbol{R}^m$ の開集合であるとする.

定理 3.2 写像 $f : X \to Y$ に対し, $\boldsymbol{x} \in X$ を $f(\boldsymbol{x}) \in Y$ の第 i 成分に対応させる X で定義された実数値関数を $f_i : X \to \boldsymbol{R}$ とする．$i = 1, 2, \ldots, m$ に対し, f_i は X の各点ですべての変数に関して偏微分可能で, 偏導関数 $\dfrac{\partial f_i}{\partial x_j} : X \to \boldsymbol{R}$ は連続であるとする．このとき, f は X の各点で微分可能である.

証明 命題 2.10 の (2) から $Y = \boldsymbol{R}$ の場合に対して主張を示せばよい．$\boldsymbol{x}, \boldsymbol{p} \in X$ の第 j 成分をそれぞれ x_j, p_j として, $g_i : [0,1] \to \boldsymbol{R}$ を

$$g_i(t) = f\left(\sum_{j=1}^{i-1} p_j\boldsymbol{e}_j + (p_i + t(x_i - p_i))\boldsymbol{e}_i + \sum_{j=i+1}^{n} x_j\boldsymbol{e}_j\right)$$

で定めると, g_i は $(0,1)$ の各点で微分可能であり,

$$g_i'(t) = (x_i - p_i)\frac{\partial f}{\partial x_i}\left(\sum_{j=1}^{i-1} p_j\boldsymbol{e}_j + (p_i + t(x_i - p_i))\boldsymbol{e}_i + \sum_{j=i+1}^{n} x_j\boldsymbol{e}_j\right) \cdots (i)$$

が成り立つ．さらに $i = 2, 3, \ldots, n$ に対し, $g_i(1) = g_{i-1}(0)$ であり, $g_1(1) = f(\boldsymbol{x})$, $g_n(0) = f(\boldsymbol{p})$ だから $f(\boldsymbol{x}) - f(\boldsymbol{p}) = \sum_{i=1}^{n}(g_i(1) - g_i(0))$ が成り立つ．平均値の定理から $g_i(1) - g_i(0) = g_i'(\theta_i)$ を満たす $0 < \theta_i < 1$ がある

ため, $\boldsymbol{c}_i = \sum_{j=1}^{i-1} p_j\boldsymbol{e}_j + (p_i + \theta_i(x_i - p_i))\boldsymbol{e}_i + \sum_{j=i+1}^{n} x_j\boldsymbol{e}_j$ とおけば (i) より $g_i'(\theta_i) = (x_i - p_i)\dfrac{\partial f}{\partial x_i}(\boldsymbol{c}_i)$ だから

$$\begin{aligned} f(\boldsymbol{x}) - f(\boldsymbol{p}) - \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(\boldsymbol{p})(x_i - p_i) &= \sum_{i=1}^{n}\left(g_i(1) - g_i(0) - \frac{\partial f}{\partial x_i}(\boldsymbol{p})(x_i - p_i)\right) = \sum_{i=1}^{n}\left(g_i'(\theta_i) - \frac{\partial f}{\partial x_i}(\boldsymbol{p})(x_i - p_i)\right) \\ &= \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}(\boldsymbol{c}_i) - \frac{\partial f}{\partial x_i}(\boldsymbol{p})\right)(x_i - p_i) \ \cdots\ (ii) \end{aligned}$$

である. $\dfrac{\partial f}{\partial x_i}$ の連続性から, 任意の $\varepsilon > 0$ に対し, $\delta > 0$ で $\|\boldsymbol{v} - \boldsymbol{p}\| < \delta$ ならば $i = 1, 2, \ldots, n$ に対して $\left|\dfrac{\partial f}{\partial x_i}(\boldsymbol{v}) - \dfrac{\partial f}{\partial x_i}(\boldsymbol{p})\right| < \dfrac{\varepsilon}{\sqrt{n}}$ を満たすものがある. 一方 $0 < \theta_i < 1$ より

$$\|\boldsymbol{c}_i - \boldsymbol{p}\| = \sqrt{\theta_i^2(x_i - p_i)^2 + \sum_{j=i+1}^{n}(x_j - p_j)^2} \leqq \|\boldsymbol{x} - \boldsymbol{p}\|$$

だから $\|\boldsymbol{x} - \boldsymbol{p}\| < \delta$ ならば $\left|\dfrac{\partial f}{\partial x_i}(\boldsymbol{c}_i) - \dfrac{\partial f}{\partial x_i}(\boldsymbol{p})\right| < \dfrac{\varepsilon}{\sqrt{n}}$ である. 故に $0 < \|\boldsymbol{x} - \boldsymbol{p}\| < \delta$ ならば (ii) と補題 3.1 から次の不等式が得られる.

$$\left|f(\boldsymbol{x}) - f(\boldsymbol{p}) - \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(\boldsymbol{p})(x_i - p_i)\right| \leqq \sum_{i=1}^{n}\left|\frac{\partial f}{\partial x_i}(\boldsymbol{c}_i) - \frac{\partial f}{\partial x_i}(\boldsymbol{p})\right||x_i - p_i| < \sum_{i=1}^{n}\frac{\varepsilon}{\sqrt{n}}|x_i - p_i| \leqq \varepsilon\|\boldsymbol{x} - \boldsymbol{p}\|$$

$A = \begin{pmatrix}\dfrac{\partial f}{\partial x_1}(\boldsymbol{p}) & \dfrac{\partial f}{\partial x_2}(\boldsymbol{p}) & \cdots & \dfrac{\partial f}{\partial x_n}(\boldsymbol{p})\end{pmatrix}$ とおけば上の不等式の左端の辺は $|f(\boldsymbol{x}) - f(\boldsymbol{p}) - A(\boldsymbol{x} - \boldsymbol{p})|$ となるため, $0 < \|\boldsymbol{x} - \boldsymbol{p}\| < \delta$ ならば $\dfrac{|f(\boldsymbol{x}) - f(\boldsymbol{p}) - A(\boldsymbol{x} - \boldsymbol{p})|}{\|\boldsymbol{x} - \boldsymbol{p}\|} < \varepsilon$ が成り立つ. これは $\displaystyle\lim_{\boldsymbol{x} \to \boldsymbol{p}}\frac{f(\boldsymbol{x}) - f(\boldsymbol{p}) - A(\boldsymbol{x} - \boldsymbol{p})}{\|\boldsymbol{x} - \boldsymbol{p}\|} = 0$ を意味するため, f は $\boldsymbol{p}$ で微分可能である. □

定義 3.3 関数 $f : X \to \boldsymbol{R}$ と $1 \leqq i_s \leqq n$ $(s = 1, 2, \ldots, r)$ を満たす整数の列 $i_1, i_2, \ldots, i_r$ に対し, f の r 次偏導関数

$$\frac{\partial^r f}{\partial x_{i_r}\partial x_{i_{r-1}}\ldots\partial x_{i_1}} : X \longrightarrow \boldsymbol{R}$$

を次のように帰納的に定義する. f が X の各点で i_1 番目の変数に関して偏微分可能であるとき, $\dfrac{\partial^1 f}{\partial x_{i_1}}$ は $\boldsymbol{x} \in X$ を $\dfrac{\partial f}{\partial x_{i_1}}(\boldsymbol{x})$ に対応させる関数とする. $\dfrac{\partial^{r-1} f}{\partial x_{i_{r-1}}\ldots\partial x_{i_1}} : X \to \boldsymbol{R}$ $(r \geqq 2)$ が定まり, X の各点で i_r 番目の変数に関して偏微分可能であるとき, $\dfrac{\partial^r f}{\partial x_{i_r}\partial x_{i_{r-1}}\ldots\partial x_{i_1}}$ を

$$\frac{\partial^r f}{\partial x_{i_r}\partial x_{i_{r-1}}\ldots\partial x_{i_1}}(\boldsymbol{x}) = \frac{\partial}{\partial x_{i_r}}\left(\frac{\partial^{r-1} f}{\partial x_{i_{r-1}}\ldots\partial x_{i_1}}\right)(\boldsymbol{x})$$

によって定義する.

$1 \leqq i_s \leqq n$ $(s = 1, 2, \ldots, r)$ を満たす任意の整数の列 $i_1, i_2, \ldots, i_r$ に対し, f の r 次偏導関数が定義できるとき, 「f は X で r 回偏微分可能である」という.

定理 3.4 $\boldsymbol{p} \in X$ とする. 関数 $f : X \to \boldsymbol{R}$ は X の各点で i 番目と j 番目の変数に関して偏微分可能であるとし, $\dfrac{\partial f}{\partial x_i}, \dfrac{\partial f}{\partial x_j} : X \to \boldsymbol{R}$ がともに連続であるとする. さらに, $\dfrac{\partial^2 f}{\partial x_j\partial x_i} : X \to \boldsymbol{R}$ が存在して, $\boldsymbol{p} \in X$ において連続ならば $\dfrac{\partial f}{\partial x_j}$ は $\boldsymbol{p}$ で i 番目の変数に関して偏微分可能であり, $\dfrac{\partial}{\partial x_i}\left(\dfrac{\partial f}{\partial x_j}\right)(\boldsymbol{p}) = \dfrac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{p})$ が成り立つ.

証明 仮定から $B_n(\boldsymbol{p}\,;r)\subset X$ を満たす $r>0$ がある. $J\subset \boldsymbol{R}^2$ を $J=\left\{\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)\in\boldsymbol{R}^2\,\middle|\,|x|,|y|<\frac{r}{\sqrt{2}}\right\}$ で定め, $F:J\to\boldsymbol{R}$ を

$$F\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)=f(\boldsymbol{p}+x\boldsymbol{e}_i+y\boldsymbol{e}_j)-f(\boldsymbol{p}+x\boldsymbol{e}_i)-f(\boldsymbol{p}+y\boldsymbol{e}_j)+f(\boldsymbol{p})$$

で定義する. $t\in\left(-\frac{r}{\sqrt{2}},\frac{r}{\sqrt{2}}\right)$ を固定して $\varphi_t:\left(-\frac{r}{\sqrt{2}},\frac{r}{\sqrt{2}}\right)\to\boldsymbol{R}$ を $\varphi_t(x)=f(\boldsymbol{p}+x\boldsymbol{e}_i+t\boldsymbol{e}_j)-f(\boldsymbol{p}+x\boldsymbol{e}_i)$ で定めれば, $\varphi_t'(x)=\dfrac{\partial f}{\partial x_i}(\boldsymbol{p}+x\boldsymbol{e}_i+t\boldsymbol{e}_j)-\dfrac{\partial f}{\partial x_i}(\boldsymbol{p}+x\boldsymbol{e}_i)$ および $\varphi_t(x)-\varphi_t(0)=F\left(\begin{smallmatrix}x\\t\end{smallmatrix}\right)$ が成り立つ. $s\in\left(-\frac{r}{\sqrt{2}},\frac{r}{\sqrt{2}}\right)$ に対し, 0 と s を両端とする区間において平均値の定理を用いると $0<\theta_1<1$ で $\varphi_t(s)-\varphi_t(0)=s\varphi_t'(\theta_1 s)$ を満たすものがあるため,

$$F\left(\begin{smallmatrix}s\\t\end{smallmatrix}\right)=s\left(\frac{\partial f}{\partial x_i}(\boldsymbol{p}+\theta_1 s\boldsymbol{e}_i+t\boldsymbol{e}_j)-\frac{\partial f}{\partial x_i}(\boldsymbol{p}+\theta_1 s\boldsymbol{e}_i)\right)\ \cdots\ (i)$$

が成り立つ. ここで θ_1 は s と t の両方に依存することに注意する. さらに, s, t を固定して $\psi:\left(-\frac{r}{\sqrt{2}},\frac{r}{\sqrt{2}}\right)\to\boldsymbol{R}$ を $\psi(u)=\dfrac{\partial f}{\partial x_i}(\boldsymbol{p}+\theta_1 s\boldsymbol{e}_i+u\boldsymbol{e}_j)$ で定義すれば, 仮定から ψ は微分可能で, 0 と t で挟まれた区間で平均値の定理を用いると, $0<\theta_2<1$ で $\psi(t)-\psi(0)=t\psi'(\theta_2 t)$ を満たすものがある. $\psi'(u)=\dfrac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{p}+\theta_1 s\boldsymbol{e}_i+u\boldsymbol{e}_j)$ だから

$$\frac{\partial f}{\partial x_i}(\boldsymbol{p}+\theta_1 s\boldsymbol{e}_i+t\boldsymbol{e}_j)-\frac{\partial f}{\partial x_i}(\boldsymbol{p}+\theta_1 s\boldsymbol{e}_i)=t\frac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{p}+\theta_1 s\boldsymbol{e}_i+\theta_2 t\boldsymbol{e}_j)\ \cdots\ (ii)$$

が得られる. (i), (ii) から

$$F\left(\begin{smallmatrix}s\\t\end{smallmatrix}\right)=st\frac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{p}+\theta_1 s\boldsymbol{e}_i+\theta_2 t\boldsymbol{e}_j)\ \cdots\ (iii)$$

を満たす $0<\theta_1,\theta_2<1$ の存在が示せたことになる. $s,t\in\left(-\frac{r}{\sqrt{2}},\frac{r}{\sqrt{2}}\right)$ を固定して $\lambda_s:\left(-\frac{r}{\sqrt{2}},\frac{r}{\sqrt{2}}\right)\to\boldsymbol{R}$ を $\lambda_s(y)=f(\boldsymbol{p}+s\boldsymbol{e}_i+y\boldsymbol{e}_j)-f(\boldsymbol{p}+y\boldsymbol{e}_j)$ で定めて, 0 と t で挟まれた区間で平均値の定理を用いれば, $\lambda_s(t)-\lambda_s(0)=t\lambda_s'(\theta_3 t)$ を満たす $0<\theta_3<1$ がある. $\lambda_s(t)-\lambda_s(0)=F\left(\begin{smallmatrix}s\\t\end{smallmatrix}\right)$, $\lambda_s'(y)=\dfrac{\partial f}{\partial x_j}(\boldsymbol{p}+s\boldsymbol{e}_i+y\boldsymbol{e}_j)-\dfrac{\partial f}{\partial x_j}(\boldsymbol{p}+y\boldsymbol{e}_j)$ より

$$F\left(\begin{smallmatrix}s\\t\end{smallmatrix}\right)=t\left(\frac{\partial f}{\partial x_j}(\boldsymbol{p}+s\boldsymbol{e}_i+\theta_3 t\boldsymbol{e}_j)-\frac{\partial f}{\partial x_j}(\boldsymbol{p}+\theta_3 t\boldsymbol{e}_j)\right)\ \cdots\ (iv)$$

である. 任意の $\varepsilon>0$ に対し, $\dfrac{\partial^2 f}{\partial x_j\partial x_i}$ の $\boldsymbol{p}$ における連続性から $0<\delta<1$ で, $x^2+y^2<\delta^2$ ならば

$$\left|\frac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{p}+x\boldsymbol{e}_i+y\boldsymbol{e}_j)-\frac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{p})\right|<\varepsilon\ \cdots\ (v)$$

を満たすものがある. 従って, $s^2+t^2<\delta^2$ ならば (iii), (iv), (v) から

$$\left|\frac{1}{s}\left(\frac{\partial f}{\partial x_j}(\boldsymbol{p}+s\boldsymbol{e}_i+\theta_3 t\boldsymbol{e}_j)-\frac{\partial f}{\partial x_j}(\boldsymbol{p}+\theta_3 t\boldsymbol{e}_j)\right)-\frac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{p})\right|=\left|\frac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{p}+\theta_1 s\boldsymbol{e}_i+\theta_2 t\boldsymbol{e}_j)-\frac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{p})\right|<\varepsilon$$

が成り立つ. $\dfrac{\partial f}{\partial x_j}$ の連続性から, 上式で $t\to 0$ とすれば $|s|<\delta$ ならば

$$\left|\frac{1}{s}\left(\frac{\partial f}{\partial x_j}(\boldsymbol{p}+s\boldsymbol{e}_i)-\frac{\partial f}{\partial x_j}(\boldsymbol{p})\right)-\frac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{p})\right|\leqq\varepsilon$$

が成り立つことが分かる. $\varepsilon>0$ は任意だから上式は $\displaystyle\lim_{s\to 0}\frac{1}{s}\left(\frac{\partial f}{\partial x_j}(\boldsymbol{p}+s\boldsymbol{e}_i)-\frac{\partial f}{\partial x_j}(\boldsymbol{p})\right)=\frac{\partial^2 f}{\partial x_j\partial x_i}(\boldsymbol{p})$ であることを示しているので定理の主張が成り立つ. □

定義 3.5　r を 0 以上の整数とする．$\boldsymbol{R}^n$ の開集合 X で定義された実数値関数 f が X の各点で r 回偏微分可能であり，r 次までのすべての偏導関数が連続であるとき，f を C^r 級関数という．0 以上のすべての整数に対して，C^r 級関数である関数を C^∞ 級関数という．$\boldsymbol{R}^n$ の開集合 X から $\boldsymbol{R}^m$ の部分集合への写像 f に対し，$\boldsymbol{x} \in X$ を $f(\boldsymbol{x})$ の第 i 成分に対応させる実数値関数を f_i で表すとき，すべての $i = 1, 2, \ldots, m$ に対して f_i が C^r 級関数ならば，f を C^r 級写像という．

定理 3.4 から，f が C^r 級関数ならば，f の r 次偏導関数は偏微分する変数の順序には依存しないことが分かる．

命題 3.6　X を $\boldsymbol{R}^n$ の開集合とする．C^r 級関数 $f, g : X \to \boldsymbol{R}$ の和と積は C^r 級関数である．

証明　r による帰納法で主張を示す．関数の和と積の微分法から次の等式が成り立つ．

$$\frac{\partial(f+g)}{\partial x_j} = \frac{\partial f}{\partial x_j} + \frac{\partial g}{\partial x_j} \ \cdots (i) \qquad \frac{\partial(fg)}{\partial x_j} = \frac{\partial f}{\partial x_j} g + f \frac{\partial g}{\partial x_j} \ \cdots (ii)$$

$r = 1$ の場合，連続関数の和と積は連続関数だからの右辺は連続関数である．故に $r = 1$ のときは主張が成り立つ．$r - 1$ のときに主張が成り立つと仮定して，f, g を C^r 級関数とすれば，$\dfrac{\partial f}{\partial x_j}$, $\dfrac{\partial g}{\partial x_j}$ は C^{r-1} 級関数であり，f, g は C^{r-1} 級関数でもあるので帰納法の仮定から (i), (ii) の右辺は C^{r-1} 級関数だから $f + g$, fg は C^r 級関数である．□

命題 3.7　X, Y, Z をそれぞれ $\boldsymbol{R}^n$, $\boldsymbol{R}^m$, $\boldsymbol{R}^l$ の開集合とする．C^r 級写像 $f : X \to Y$ と $g : Y \to Z$ の合成写像 $g \circ f : X \to Z$ は C^r 級写像である．

証明　$\boldsymbol{x} \in X$, $\boldsymbol{y} \in Y$ に対して $f(\boldsymbol{x})$, $(g \circ f)(\boldsymbol{x})$, $g(\boldsymbol{y})$ の第 i 成分を対応させる関数をそれぞれ f_i,: $X \to \boldsymbol{R}$, $(g \circ f)_i$,: $X \to \boldsymbol{R}$, $g_i : Y \to \boldsymbol{R}$ とすると $(g \circ f)_i(\boldsymbol{x}) = g_i(f(\boldsymbol{x}))$ だから，合成写像の微分法から次の等式を得る．

$$\frac{\partial(g\circ f)_i}{\partial x_j}(\boldsymbol{x}) = \frac{\partial(g_i \circ f)}{\partial x_j}(\boldsymbol{x}) = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}(f(\boldsymbol{x}))\frac{\partial f_k}{\partial x_j}(\boldsymbol{x}) = \sum_{k=1}^{m}\Big(\frac{\partial g_i}{\partial y_k} \circ f\Big)(\boldsymbol{x})\frac{\partial f_k}{\partial x_j}(\boldsymbol{x}) \ \cdots (*)$$

r による帰納法で主張を示す．$r = 1$ の場合，連続写像の合成は連続写像であり，連続関数の積，和は連続関数だから (i) の右辺は $\boldsymbol{x}$ の連続写像である．故に $(g \circ f)_i$ は C^1 級関数だから命題 2.10 より $r = 1$ のときは主張が成り立つ．$r - 1$ のときに主張が成り立つと仮定して，f, g を C^r 級写像とすれば $\dfrac{\partial f_k}{\partial x_j}$, $\dfrac{\partial g_i}{\partial y_k}$ は C^{r-1} 級関数で，f は C^{r-1} 級関数でもあるので帰納法の仮定から $\dfrac{\partial g_i}{\partial y_k} \circ f$ は C^{r-1} 級関数である．命題 3.6 より，C^{r-1} 級関数の和と積は C^{r-1} 級関数だから $(*)$ の右辺は $\boldsymbol{x}$ の C^{r-1} 級関数である．故にすべての $i = 1, 2, \ldots, l$, $j = 1, 2, \ldots, n$ に対して $\dfrac{\partial(g \circ f)_i}{\partial x_j}$ C^{r-1} 級関数だから $g \circ f$ は C^r 級写像となり，帰納法によって主張が示される．□

命題 3.8　X を $\boldsymbol{R}^n$ の開集合，Y を $\boldsymbol{R}^m$ の部分集合，$f : X \to Y$ を C^1 級写像とする．$\boldsymbol{x}, \boldsymbol{y} \in X$ と任意の $0 \leqq t \leqq 1$ に対して $\boldsymbol{y} + t(\boldsymbol{x} - \boldsymbol{y}) \in X$ であるとし，$t \in [0, 1]$ を $\|f'(\boldsymbol{y} + t(\boldsymbol{x} - \boldsymbol{y}))\|$ に対応させる関数の最大値を M とすれば次の不等式が成り立つ．

$$\|f(\boldsymbol{x}) - f(\boldsymbol{y})\| \leqq \sqrt{m}M\|\boldsymbol{x} - \boldsymbol{y}\|$$

証明　$i = 1, 2, \ldots, m$ に対して $g_i : [0, 1] \to \boldsymbol{R}$ を $g_i(t) = f_i(\boldsymbol{y} + t(\boldsymbol{x} - \boldsymbol{y}))$ で定めると，g_i は $(0, 1)$ の各点で微分可能であり，合成写像の微分法から $g_i'(t) = f_i'(\boldsymbol{y} + t(\boldsymbol{x} - \boldsymbol{y}))(\boldsymbol{x} - \boldsymbol{y})$ が成り立つ．平均値の定理から $g_i(1) - g_i(0) = g_i'(\theta_i)$ を満たす $0 < \theta_i < 1$ があるため，上の等式と命題 2.1 から次の式が得られる．

$$|f_i(\boldsymbol{x}) - f_i(\boldsymbol{y})| = |g_i(1) - g_i(0)| = |g_i'(\theta_i)| = |f_i'(\boldsymbol{y} + \theta_i(\boldsymbol{x} - \boldsymbol{y}))(\boldsymbol{x} - \boldsymbol{y})| \leqq \|f_i'(\boldsymbol{y} + \theta_i(\boldsymbol{x} - \boldsymbol{y}))\|\|\boldsymbol{x} - \boldsymbol{y}\|$$

命題 2.10 から $f_i'(\boldsymbol{y}+\theta_i(\boldsymbol{x}-\boldsymbol{y}))$ は $f'(\boldsymbol{y}+\theta_i(\boldsymbol{x}-\boldsymbol{y}))$ の第 i 行だから $\|f_i'(\boldsymbol{y}+\theta_i(\boldsymbol{x}-\boldsymbol{y}))\|^2 \leqq \|f'(\boldsymbol{y}+\theta_i(\boldsymbol{x}-\boldsymbol{y}))\|^2 \leqq M^2$ である．故に上の不等式から

$$\|f(\boldsymbol{x}) - f(\boldsymbol{y})\|^2 = \sum_{i=1}^{m} |f_i(\boldsymbol{x}) - f_i(\boldsymbol{y})|^2 \leqq \sum_{i=1}^{m} \|f_i'(\boldsymbol{y} + \theta_i(\boldsymbol{x} - \boldsymbol{y}))\|^2\|\boldsymbol{x} - \boldsymbol{y}\|^2 \leqq mM^2\|\boldsymbol{x} - \boldsymbol{y}\|^2$$

が得られて, この不等式の両端の辺の正の平方根をとれば $\|f(\boldsymbol{x})-f(\boldsymbol{y})\| \leqq \sqrt{m}M\|\boldsymbol{x}-\boldsymbol{y}\|$ が得られる. □

定義 3.9 V を $\boldsymbol{R}$ 上のベクトル空間とする. V の部分集合 X が次の条件を満たすとき X は凸であるという.

「$\boldsymbol{x}, \boldsymbol{y} \in X$ かつ $0 \leqq t \leqq 1$ ならば $\boldsymbol{y}+t(\boldsymbol{x}-\boldsymbol{y}) \in X$」

系 3.10 X を $\boldsymbol{R}^n$ の凸である開集合, Y を $\boldsymbol{R}^m$ の部分集合, $f: X \to Y$ を C^1 級写像とする. 実数 K がすべての $\boldsymbol{x} \in X$ と $i=1,2,\dots,m$, $j=1,2,\dots,n$ に対して $\left|\dfrac{\partial f_i}{\partial x_j}(\boldsymbol{x})\right| \leqq K$ を満たすとき, 任意の $\boldsymbol{x}, \boldsymbol{y} \in X$ に対して次の不等式が成り立つ.

$$\|f(\boldsymbol{x})-f(\boldsymbol{y})\| \leqq Km\sqrt{n}\|\boldsymbol{x}-\boldsymbol{y}\|$$

証明 命題 2.10 から $f'(\boldsymbol{x})$ の (i,j) 成分は $\dfrac{\partial f_i}{\partial x_j}(\boldsymbol{x})$ だから, 仮定からすべての $\boldsymbol{x} \in X$ に対して $\|f'(\boldsymbol{x})\| \leqq \sqrt{mn}K$ となるため, 任意の $\boldsymbol{x}, \boldsymbol{y} \in X$ に対し $t \in [0,1]$ を $\|f'(\boldsymbol{y}+t(\boldsymbol{x}-\boldsymbol{y}))\|$ に対応させる関数の最大値は $\sqrt{mn}K$ 以下である. 従って命題 3.8 から $\|f(\boldsymbol{x})-f(\boldsymbol{y})\| \leqq Km\sqrt{n}\|\boldsymbol{x}-\boldsymbol{y}\|$ である. □

系 3.11 $X, Y, f: X \to Y$ を命題 3.10 と同様とする. $\boldsymbol{p} \in X$ とし, 実数 K がすべての $\boldsymbol{x} \in X$ と $i=1,2,\dots,m$, $j=1,2,\dots,n$ に対して $\left|\dfrac{\partial f_i}{\partial x_j}(\boldsymbol{x})-\dfrac{\partial f_i}{\partial x_j}(\boldsymbol{p})\right| \leqq K$ を満たすとき, 任意の $\boldsymbol{x}, \boldsymbol{y} \in X$ に対して次の不等式が成り立つ.

$$\|f(\boldsymbol{x})-f(\boldsymbol{y})-f'(\boldsymbol{p})(\boldsymbol{x}-\boldsymbol{y})\| \leqq Km\sqrt{n}\|\boldsymbol{x}-\boldsymbol{y}\|$$

証明 $g(\boldsymbol{x})=f(\boldsymbol{x})-f'(\boldsymbol{p})(\boldsymbol{x}-\boldsymbol{p})$ で $g: X \to \boldsymbol{R}^m$ を定めれば, 例 2.9 から $g'(\boldsymbol{x})=f'(\boldsymbol{x})-f'(\boldsymbol{p})$ であることに注意して, g に対して系 3.10 の結果を用いればよい. □

4 常微分方程式の解の存在と一意性

定義 4.1 U を $\boldsymbol{R}^n$ の開集合, I を $\boldsymbol{R}$ の開集合とする. 写像 $F: I \times U \to \boldsymbol{R}^n$ と I に含まれる開区間 J に対して, 微分可能な写像 $f: J \to U$ が微分方程式 $\boldsymbol{x}'=F(t,\boldsymbol{x})$ の解であるとは, 任意の $t \in J$ に対して, $f'(t)=F(t,f(t))$ が成り立つことである. 従って F が連続ならば f は J において C^1 級写像である.

I を $\boldsymbol{R}$ の部分集合, $[a,b] \subset I$ とする. 写像 $f: I \to \boldsymbol{R}^n$ に対し, 各 $t \in I$ を $f(t)$ の第 i 成分に対応させる関数を f_i で表す. $i=1,2,\dots,n$ に対して $\displaystyle\int_a^b f_i(t)\,dt$ が存在するとき, この値を第 i 成分とする n 次元数ベクトルを $\displaystyle\int_a^b f(t)\,dt$ で表す.

命題 4.2 不等式 $\left\|\displaystyle\int_a^b f(t)\,dt\right\| \leqq \displaystyle\int_a^b \|f(t)\|\,dt$ が成り立つ.

証明 三角不等式より, 任意の自然数 N に対して次の不等式が成り立つ.

$$\left\|\sum_{k=1}^{N}\frac{b-a}{N}f\left(a+\frac{k(b-a)}{N}\right)\right\| \leqq \sum_{k=1}^{N}\frac{b-a}{N}\left\|f\left(a+\frac{k(b-a)}{N}\right)\right\|$$

$\displaystyle\sum_{k=1}^{N}\frac{b-a}{N}f\left(a+\frac{k(b-a)}{N}\right)$ の第 i 成分は, $\displaystyle\sum_{k=1}^{N}\frac{b-a}{N}f_i\left(a+\frac{k(b-a)}{N}\right)$ で, $N \to \infty$ とすれば, $\displaystyle\int_a^b f_i(t)\,dt$ に近づくため, $N \to \infty$ のとき, 上式の左辺は $\left\|\displaystyle\int_a^b f(t)\,dt\right\|$ に近づく. また, 上式の右辺は $N \to \infty$ のとき, $\displaystyle\int_a^b \|f(t)\|\,dt$ に近づくため, $\left\|\displaystyle\int_a^b f(t)\,dt\right\| \leqq \displaystyle\int_a^b \|f(t)\|\,dt$ が得られる. □

補題 4.3　定義 4.1 における F は連続であるとする．$t_0 \in J$ と $\boldsymbol{x}_0 \in U$ に対し，$f : J \to U$ が $f(t_0) = \boldsymbol{x}_0$ を満たす $\boldsymbol{x}' = F(t, \boldsymbol{x})$ の解であるための必要十分条件は，f は連続かつ $f(t) = \boldsymbol{x}_0 + \displaystyle\int_{t_0}^{t} F(s, f(s))\,ds$ がすべての $t \in J$ に対して成り立つことである．

証明　f が $\boldsymbol{x}' = F(t, \boldsymbol{x})$ の解ならば f' は連続だから $f(t) - f(t_0) = \displaystyle\int_{t_0}^{t} f'(s)\,ds = \int_{t_0}^{t} F(s, f(s))\,ds$ が得られる．従って $f(t_0) = \boldsymbol{x}_0$ ならば $f(t) = \boldsymbol{x}_0 + \displaystyle\int_{t_0}^{t} F(s, u(s))\,ds$ である．逆に $f(t) = \boldsymbol{x}_0 + \displaystyle\int_{t_0}^{t} F(s, f(s))\,ds$ がすべての $t \in J$ に対して成り立てば，この両辺の導関数を考えると，微積分学の基本定理より $f'(t) = F(t, f(t))$ である．□

定理 4.4　(常微分方程式の解の存在定理) I, U を定義 4.1 と同様とする．連続写像 $F : I \times U \to \boldsymbol{R}^n$ が下記の条件 $(*)$ を満たせば，任意の $t_0 \in I$ と $\boldsymbol{x}_0 \in U$ に対し，t_0 を含み I に含まれる開区間 J が存在して，微分方程式 $\boldsymbol{x}' = F(t, \boldsymbol{x})$ の J における解 $f : J \to U$ で $f(t_0) = \boldsymbol{x}_0$ を満たすものが存在する．K が t_0 を含み I に含まれる開区間で，$g : K \to U$ も $g(t_0) = \boldsymbol{x}_0$ を満たす $\boldsymbol{x}' = F(t, \boldsymbol{x})$ の解ならば，$t \in J \cap K$ に対して $g(t) = f(t)$ である．

$(*)$ 任意の $(t, \boldsymbol{x}) \in I \times U$ に対して，$t \in J \subset I$ である開区間 J と，$\boldsymbol{x}$ を中心として U に含まれる開球 B および $k \geqq 0$ で，$s \in J$, $\boldsymbol{y}_1, \boldsymbol{y}_2 \in B$ ならば $\|F(s, \boldsymbol{y}_1) - F(s, \boldsymbol{y}_2)\| \leqq k\|\boldsymbol{y}_1 - \boldsymbol{y}_2\|$ が成り立つようなものが存在する．

証明　まず，任意の $(t_0, \boldsymbol{x}_0) \in I \times U$ に対し，条件 $(*)$ を満たす開区間を J_0，開球を B，B の半径を β とする．$[t_0 - a, t_0 + a] \subset J_0$ となる $a > 0$ をとれば，$[t_0 - a, t_0 + a] \times \{\boldsymbol{x}_0\}$ はコンパクトだから F の連続性から $F([t_0 - a, t_0 + a] \times \{\boldsymbol{x}_0\})$ もコンパクトである．従って $F([t_0 - a, t_0 + a] \times \{\boldsymbol{x}_0\})$ は有界だから，$M > 0$ で条件「$t \in [t_0 - a, t_0 + a]$ ならば $\|F(t, \boldsymbol{x}_0)\| \leqq M$」を満たすものがある．$0 < r \leqq a$ に対し，$F_r = \boldsymbol{B}([t_0 - r, t_0 + r], \boldsymbol{R}^n)$ とおくと，$(F_r, \hat{d})$ は命題 1.14 により完備距離空間である．B_r を $[t_0 - r, t_0 + r]$ から $\boldsymbol{x}_0$ への定値写像 $c_{\boldsymbol{x}_0}$ を中心とし，半径 β の F_r における開球 $B_{\hat{d}}(c_{\boldsymbol{x}_0}; \beta)$ とすると，$s \in [t_0 - r, t_0 + r]$, $g \in B_r$ ならば $g(s) \in B$ である．写像 $v : B_r \to F_r$ を $v(g)(t) = \boldsymbol{x}_0 + \displaystyle\int_{t_0}^{t} F(s, g(s))ds$ で定める．

$g_1, g_2 \in B_r$ と $s \in [t_0 - r, t_0 + r]$ に対し，仮定より $\|F(s, g_1(s)) - F(s, g_2(s))\| \leqq k\|g_1(s) - g_2(s)\| \leqq k\hat{d}(g_1, g_2)$ である．従って命題 4.2 から，任意の $t \in [t_0 - r, t_0 + r]$ に対して，次の不等式が成り立つ．

$$\|v(g_1)(t) - v(g_2)(t)\| = \left\| \int_{t_0}^{t} (F(s, g_1(s)) - F(s, g_2(s)))ds \right\| \leqq kr\hat{d}(g_1, g_2)$$

故に $\hat{d}(v(g_1), v(g_2)) \leqq kr\hat{d}(g_1, g_2)$ である．一方，任意の $s \in [t_0 - r, t_0 + r]$ に対して $\|F(s, \boldsymbol{x}_0)\| \leqq M$ だから

$$\|v(c_{\boldsymbol{x}_0})(t) - c_{\boldsymbol{x}_0}(t)\| = \left\| \int_{t_0}^{t} F(s, \boldsymbol{x}_0)ds \right\| \leqq Mr$$

が任意の $t \in [t_0 - r, t_0 + r]$ に対して成り立つため $\hat{d}(v(c_{\boldsymbol{x}_0}), c_{\boldsymbol{x}_0}) \leqq Mr$ となる．$r < \dfrac{\beta}{M + k\beta}$ を満たすように $r > 0$ を選んでおけば $kr < 1$ かつ $Mr < \beta(1 - kr)$ が成り立つため，命題 1.16 が v に適用できて $v(f) = f$ となる $f \in B_r$ がある．$J = (t_0 - r, t_0 + r)$ とすれば補題 4.3 から $f : J \to B$ は $f(t_0) = \boldsymbol{x}_0$ を満たす $\boldsymbol{x}' = F(t, \boldsymbol{x})$ の解である．

K が t_0 を含み I に含まれる開区間で，$\bar{f} : K \to U$ も $\bar{f}(t_0) = \boldsymbol{x}_0$ を満たす $\boldsymbol{x}' = F(t, \boldsymbol{x})$ の解であるとして，$A = \{t \in J \cap K \,|\, f(t) = \bar{f}(t)\}$ とおく．A は連続写像 $\bar{f} - f : J \cap K \to \boldsymbol{R}^n$ による閉集合 $\{\boldsymbol{0}\}$ の逆像で，$t_0 \in A$ だから A は $J \cap K$ の空でない閉部分集合である．$t_1 \in A$ に対し，$\bar{f}(t_1) = f(t_1) = \boldsymbol{x}_1$ とおいて，上の議論において t_0 を t_1, $\boldsymbol{x}_0$ を $\boldsymbol{x}_1$ に置き換えて同じ議論を行えば，$r' > 0$ で，$t \in [t_1 - r', t_1 + r']$ に対して $v(g)(t) = \boldsymbol{x}_1 + \displaystyle\int_{t_1}^{t} F(s, g(s))ds$ によって定められた写像 $v : B_{r'} = B_{\hat{d}}(c_{\boldsymbol{x}_1}; \beta) \to F_{r'} = \boldsymbol{B}([t_1 - r', t_1 + r'], \boldsymbol{R}^n)$ が命題 1.16 の条件を満たすようなものがとれる．このとき，f, $\bar{f}$ の定義域を $[t_1 - r', t_1 + r']$ に制限したものをそれぞれ f_1, $\bar{f}_1$ で表せば，$v(f_1) = f_1$ と $v(\bar{f}_1) = \bar{f}_1$ が成り立つため，命題 1.16 から $f_1 = \bar{f}_1$ である．従って A の各点は $J \cap K$ の内点だから，A は $J \cap K$ の開部分集合でもある．$J \cap K$ は開区間で連結だから $A = J \cap K$ が成り立つ．□

補題 4.5　$(X, \mathcal{O}_X)$ をコンパクトな位相空間，$(Y, \mathcal{O}_Y)$, $(Z, \mathcal{O}_Z)$ を位相空間とする．A を Y の部分集合，U を Z の開

集合 として, U が連続写像 $g: X\times Y\to Z$ による $X\times A$ の像を含めば, A を含む Y の開集合 W で $g(X\times W)\subset U$ となるものが存在する.

証明 $g^{-1}(U)$ は $X\times A$ を含む $X\times Y$ の開集合だから, 任意の $(x,y)\in X\times A$ に対して x, y の開近傍 $V_{x,y}$, $W_{x,y}$ で $V_{x,y}\times W_{x,y}\subset g^{-1}(U)$ を満たすものがある. 各 $y\in A$ に対し, X のコンパクト性から $x_1,x_2,\dots,x_{n_y}\in X$ で $X=\bigcup_{i=1}^{n_y}V_{x_i,y}$ となるものがあり, $W_y=\bigcap_{i=1}^{n_y}W_{x_i,y}$ とおく. このとき $X\times W_y=\bigcup_{i=1}^{n_y}(V_{x_i,y}\times W_y)$ で, 各 $V_{x_i,y}\times W_y$ は $g^{-1}(U)$ に含まれるから $X\times W_y\subset g^{-1}(U)$ である. W_y は y の開近傍だから $W=\bigcup_{y\in A}W_y$ とおけばよい. □

実数を成分とする $m\times n$ 行列全体からなる $\boldsymbol{R}$ 上のベクトル空間を $M_{m,n}(\boldsymbol{R})$ で表す. $A,B\in M_{m,n}(\boldsymbol{R})$ に対して $(A,B)=\mathrm{tr}({}^tAB)$ とおけば $A=(a_{ij})$, $B=(b_{ij})$ のとき $(A,B)=\sum_{1\leqq i\leqq m,1\leqq j\leqq n}a_{ij}b_{ij}$ だから (A,B) は $M_{m,n}(\boldsymbol{R})$ の内積であることが確かめられる. この内積によって $M_{m,n}(\boldsymbol{R})$ を $\boldsymbol{R}$ 上の計量ベクトル空間とみなす. このとき $A\in M_{m,n}(\boldsymbol{R})$ の長さ $\sqrt{(A,A)}$ は A のすべての成分の 2 乗の和の正の平方根だから, 命題 2.1 の前で定義した $\|A\|$ に一致する. $A,B\in M_{m,n}(\boldsymbol{R})$ に対して $d_{m,n}(A,B)=\|A-B\|$ によって関数 $d_{m,n}: M_{m,n}(\boldsymbol{R})\times M_{m,n}(\boldsymbol{R})\to\boldsymbol{R}$ を定義すれば, $d_{m,n}$ は $M_{m,n}(\boldsymbol{R})$ の距離関数だから, この距離関数によって $M_{m,n}(\boldsymbol{R})$ を距離空間とみなす.

命題 4.6 U を $\boldsymbol{R}^n$ の開集合とする. 写像 $F: I\times U\to\boldsymbol{R}^m$ の各成分の関数 F_i $(i=1,2,\dots,m)$ が 2 番目の変数 x_1 から $n+1$ 番目の変数 x_n に関して偏微分可能で, 偏導関数 $\dfrac{\partial F_i}{\partial x_j}: I\times U\to\boldsymbol{R}$ が連続ならば, F は定理 4.4 の条件 $(*)$ を満たす.

証明 $(t,\boldsymbol{x})\in I\times U$ に対して, $\dfrac{\partial F_i}{\partial x_j}(t,\boldsymbol{x})$ を (i,j) 成分とする $m\times n$ 行列を対応させる写像を $D_2F: I\times U\to M_{m,n}(\boldsymbol{R})$ で表せば, 仮定から D_2F は連続である. $(t,\boldsymbol{x})\in I\times U$ に対して, $[t-a,t+a]\subset I$ となる $a>0$ をとると, $[t-a,t+a]\times\{\boldsymbol{x}\}$ はコンパクトだから, DF_2 の連続性から $D_2F([t-a,t+a]\times\{\boldsymbol{x}\})$ はコンパクトであり, 従って有界である. そこで, $k>0$ を, $M_{m,n}(\boldsymbol{R})$ において零行列を中心とし, 半径 $\dfrac{k}{m\sqrt{n}}$ の開球を B_1 としたとき, B_1 が $D_2F([t-a,t+a]\times\{\boldsymbol{x}\})$ を含むように定める. 補題 4.5 により, $\boldsymbol{x}$ を中心とした開球 B で, $D_2F([t-a,t+a]\times B)\subset B_1$ を満たすものがある. このとき, $(s,\boldsymbol{y})\in[t-a,t+a]\times B$ ならば $D_2F(s,\boldsymbol{y})$ の各成分の絶対値は $\dfrac{k}{m\sqrt{n}}$ より小さいため, 系 3.10 から $s\in(t-a,t+a)$, $\boldsymbol{y}_1,\boldsymbol{y}_2\in B$ ならば $\|F(s,\boldsymbol{y}_1)-F(s,\boldsymbol{y}_2)\|\leqq k\|\boldsymbol{y}_1-\boldsymbol{y}_2\|$ が得られる. □

5 陰関数定理・逆写像定理

実数を成分とする n 次正方行列全体 $M_{n,n}(\boldsymbol{R})$ を $M_n(\boldsymbol{R})$ で表す.

補題 5.1 X を $\boldsymbol{R}^m$ の開集合とし, 連続写像 $f: X\to M_n(\boldsymbol{R})$ が与えられているとき, $f(\boldsymbol{x})$ が正則行列であるようなベクトル全体からなる X の部分集合は $\boldsymbol{R}^m$ の開集合である.

証明 $\boldsymbol{R}-\{0\}$ は $\boldsymbol{R}$ の開集合であり, 連続写像の合成写像 $\det\circ f: X\to\boldsymbol{R}$ は連続だから, $(f\circ\det)^{-1}(\boldsymbol{R}-\{0\})$ は X の開集合である. この集合は $f(\boldsymbol{x})$ が正則行列であるような $\boldsymbol{x}$ 全体からなる X の部分集合である. □

$\boldsymbol{x}\in\boldsymbol{R}^n$, $\boldsymbol{y}\in\boldsymbol{R}^m$ に対し, $\boldsymbol{x}$, $\boldsymbol{y}$ の第 i 成分をそれぞれ x_i, y_i とするとき, $\left(\begin{smallmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{smallmatrix}\right)$ によって, 第 i 成分が $1\leqq i\leqq n$ ならば x_i, $n+1\leqq i\leqq m+n$ ならば y_{i-n} である $\boldsymbol{R}^{m+n}$ のベクトルを表す. $\boldsymbol{R}^n$ と $\boldsymbol{R}^m$ の直積空間 $\boldsymbol{R}^n\times\boldsymbol{R}^m$ 要素 $(\boldsymbol{x},\boldsymbol{y})$ を $\boldsymbol{R}^{m+n}$ のベクトル $\left(\begin{smallmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{smallmatrix}\right)$ に対応させることによって, $\boldsymbol{R}^n\times\boldsymbol{R}^m$ と $\boldsymbol{R}^{m+n}$ を同一視する.

$\boldsymbol{R}^n\times\boldsymbol{R}^m$ の開集合 U から $\boldsymbol{R}^k$ の部分集合 Y への写像 f が $\boldsymbol{p}\in U$ で微分可能であるとき, $k\times n$ 行列 $D_1f(\boldsymbol{p})$ と $k\times m$ 行列 $D_2f(\boldsymbol{p})$ を以下のように定めれば, $f'(\boldsymbol{p})=\begin{pmatrix}D_1f(\boldsymbol{p}) & D_2f(\boldsymbol{p})\end{pmatrix}$ が成り立つ.

$$D_1f(\boldsymbol{p})=\begin{pmatrix}f'(\boldsymbol{p})\boldsymbol{e}_1 & f'(\boldsymbol{p})\boldsymbol{e}_2 & \cdots & f'(\boldsymbol{p})\boldsymbol{e}_n\end{pmatrix}\qquad D_2f(\boldsymbol{p})=\begin{pmatrix}f'(\boldsymbol{p})\boldsymbol{e}_{n+1} & f'(\boldsymbol{p})\boldsymbol{e}_{n+2} & \cdots & f'(\boldsymbol{p})\boldsymbol{e}_{n+m}\end{pmatrix}$$

定理 5.2　(陰関数定理) F は $\boldsymbol{R}^n\times\boldsymbol{R}^m$ の開集合 U から $\boldsymbol{R}^m$ への C^1 級写像で, U の点 $\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}$ に対し, $F\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}=\boldsymbol{0}$ かつ $D_2F\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}$ は正則行列であるとする．このとき $\boldsymbol{R}^n$ における $\boldsymbol{x}_0$ の開近傍 U_0 が存在して, U_0 に含まれる $\boldsymbol{x}_0$ の任意の連結な開近傍 U_1 に対し, C^1 級写像 $f:U_1\to\boldsymbol{R}^m$ で $f(\boldsymbol{x}_0)=\boldsymbol{y}_0$ かつ, すべての $\boldsymbol{x}\in U_1$ に対して $\binom{\boldsymbol{x}}{f(\boldsymbol{x})}\in U$, $F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}=\boldsymbol{0}$ を満たすものがただ一つ存在する．さらに f の微分は $f'(\boldsymbol{x})=-D_2F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}^{-1}D_1F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}$ により与えられる．

証明　(第一段階) $A=D_2F\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}$ とおいて, 写像 $v:U\to\boldsymbol{R}^m$ を $v\binom{\boldsymbol{x}}{\boldsymbol{y}}=\boldsymbol{y}-A^{-1}F\binom{\boldsymbol{x}}{\boldsymbol{y}}$ で定義するとき v を $\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}$ の十分小さな近傍に制限すれば命題 1.15 が適用できることを示す．$\binom{\boldsymbol{x}}{\boldsymbol{y}}\in U$ を $F\binom{\boldsymbol{x}}{\boldsymbol{y}}$ の第 i 成分に対応させる関数を F_i で表す．F は C^1 級写像だから中心 $\boldsymbol{x}_0$, $\boldsymbol{y}_0$, 半径 α, β の $\boldsymbol{R}^n$, $\boldsymbol{R}^m$ における開球 $U_0=B_n(\boldsymbol{x}_0;\alpha)$, $V_0=B_m(\boldsymbol{y}_0;\beta)$ で条件「$U_0\times V_0\subset U$ かつ $\boldsymbol{x}\in U_0$, $\boldsymbol{y}\in V_0$ ならばすべての $i=1,2,\dots,m$, $j=1,2,\dots,n+m$ に対して $\left|\dfrac{\partial F_i}{\partial x_j}\binom{\boldsymbol{x}}{\boldsymbol{y}}-\dfrac{\partial F_i}{\partial x_j}\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}\right|\leqq\dfrac{1}{2m\sqrt{n}\|A^{-1}\|}$」を満たすものがある．$F'\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}=\left(D_1F\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}\ \ D_2F\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}\right)$ だから $F'\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}\left(\binom{\boldsymbol{x}}{\boldsymbol{y}_1}-\binom{\boldsymbol{x}}{\boldsymbol{y}_2}\right)=D_2F\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}(\boldsymbol{y}_1-\boldsymbol{y}_2)$ が成り立つことに注意すれば, 系 3.11 から $\boldsymbol{x}\in U_0$, $\boldsymbol{y}_1,\boldsymbol{y}_2\in V_0$ ならば $\left\|F\binom{\boldsymbol{x}}{\boldsymbol{y}_1}-F\binom{\boldsymbol{x}}{\boldsymbol{y}_2}-D_2F\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}(\boldsymbol{y}_1-\boldsymbol{y}_2)\right\|\leqq\dfrac{1}{2\|A^{-1}\|}\|\boldsymbol{y}_1-\boldsymbol{y}_2\|$ が成り立つため, 命題 2.1 から次が成り立つ．

$$\begin{aligned}\left\|v\binom{\boldsymbol{x}}{\boldsymbol{y}_1}-v\binom{\boldsymbol{x}}{\boldsymbol{y}_2}\right\|&=\left\|\boldsymbol{y}_1-\boldsymbol{y}_2-A^{-1}\left(F\binom{\boldsymbol{x}}{\boldsymbol{y}_1}-F\binom{\boldsymbol{x}}{\boldsymbol{y}_2}\right)\right\|=\left\|A^{-1}\left(A(\boldsymbol{y}_1-\boldsymbol{y}_2)-\left(F\binom{\boldsymbol{x}}{\boldsymbol{y}_1}-F\binom{\boldsymbol{x}}{\boldsymbol{y}_2}\right)\right)\right\|\\&\leqq\|A^{-1}\|\left\|D_2F\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}(\boldsymbol{y}_1-\boldsymbol{y}_2)-\left(F\binom{\boldsymbol{x}}{\boldsymbol{y}_1}-F\binom{\boldsymbol{x}}{\boldsymbol{y}_2}\right)\right\|\leqq\frac{1}{2}\|\boldsymbol{y}_1-\boldsymbol{y}_2\|\end{aligned}$$

さらに F の $\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}$ における連続性より, 必要なら α を小さくとりなおして $\boldsymbol{x}\in U_0$ ならば $\left\|F\binom{\boldsymbol{x}}{\boldsymbol{y}_0}\right\|<\dfrac{\beta}{2\|A^{-1}\|}$ が成り立つようにできる．このとき, $\left\|v\binom{\boldsymbol{x}}{\boldsymbol{y}_0}-\boldsymbol{y}_0\right\|=\left\|-A^{-1}F\binom{\boldsymbol{x}}{\boldsymbol{y}_0}\right\|<\dfrac{\beta}{2}$ が成り立つ．以上より命題 1.15 が写像 $v:U_0\times V_0\to\boldsymbol{R}^m$ に適用できて, 写像 $f:U_0\to V_0$ で, すべての $\boldsymbol{x}\in U_0$ に対し, $v\binom{\boldsymbol{x}}{f(\boldsymbol{x})}=f(\boldsymbol{x})$ を満たすものがただ一つ存在して f は連続である．$v\binom{\boldsymbol{x}}{f(\boldsymbol{x})}=f(\boldsymbol{x})$ は $F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}=\boldsymbol{0}$ と同値で, $F\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}=\boldsymbol{0}$ だから $v\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}=\boldsymbol{y}_0$ となるため $f(\boldsymbol{x}_0)=\boldsymbol{y}_0$ である．

(第二段階) 次に U_1 を U_0 に含まれる $\boldsymbol{x}_0$ の連結な開近傍とするとき, 連続写像 $f:U_1\to\boldsymbol{R}^m$ で, $f(\boldsymbol{x}_0)=\boldsymbol{y}_0$ かつ, すべての $x\in U_1$ に対し, $\binom{\boldsymbol{x}}{f(\boldsymbol{x})}\in U$, $F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}=0$ を満たすものはただ一つしかないことを示す．F は C^1 級写像だから $\boldsymbol{x}\in U_0$ を $D_2F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}$ に対応させる写像 $U_0\to M_m(\boldsymbol{R})$ は連続であり, $D_2F\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}$ は正則行列だから, 補題 5.1 より, $D_2F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}$ が正則行列であるような $\boldsymbol{x}\in U_0$ 全体からなる集合は $\boldsymbol{x}_0$ の開近傍である．従って必要なら再び α を小さくとりなおすことにより, すべての $\boldsymbol{x}\in U_0$ に対して $D_2F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}$ が正則行列であると仮定してよい．$g:U_1\to\boldsymbol{R}^m$ も f と同じ条件を満たすものとする．$V=\{\boldsymbol{x}\in U_1|\,f(\boldsymbol{x})=g(\boldsymbol{x})\}$ とおくと V は $\boldsymbol{x}\in U_1$ を $f(\boldsymbol{x})-g(\boldsymbol{x})$ に対応させる写像 $f-g$ による閉集合 $\{\boldsymbol{0}\}$ の逆像である．f, g は連続だから, $f-g$ も連続であるため, V は U_1 の閉集合であり, $f(\boldsymbol{x}_0)=g(\boldsymbol{x}_0)=\boldsymbol{y}_0$ より V は $\boldsymbol{x}_0$ を含む．$\boldsymbol{a}\in V$ に対し, $\boldsymbol{b}=f(\boldsymbol{a})$ とおくと $\binom{\boldsymbol{a}}{\boldsymbol{b}}\in U$ で $D_2F\binom{\boldsymbol{a}}{\boldsymbol{b}}$ は正則行列だから $\binom{\boldsymbol{x}_0}{\boldsymbol{y}_0}$ のかわりに $\binom{\boldsymbol{a}}{\boldsymbol{b}}$ に対して第一段階の結果を用いれば $\boldsymbol{a}$, $\boldsymbol{b}$ の開近傍 $U_{\boldsymbol{a}}$, $V_{\boldsymbol{b}}$ が存在して, $h(\boldsymbol{a})=\boldsymbol{b}$ かつ「すべての $\boldsymbol{x}\in U_{\boldsymbol{a}}$ に対して $F\binom{\boldsymbol{x}}{h(\boldsymbol{x})}=\boldsymbol{0}$」を満たす写像 $h:U_{\boldsymbol{a}}\to V_{\boldsymbol{b}}$ はただ一つしかない．$f(\boldsymbol{a})=g(\boldsymbol{a})=\boldsymbol{b}=h(\boldsymbol{a})$ より $U_2=U_{\boldsymbol{a}}\cap f^{-1}(V_{\boldsymbol{b}})\cap g^{-1}(V_{\boldsymbol{b}})$ とおくと U_2 は $\boldsymbol{a}$ の開近傍で, U_2 の各点 $\boldsymbol{x}$ で $F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}=F\binom{\boldsymbol{x}}{g(\boldsymbol{x})}=F\binom{\boldsymbol{x}}{h(\boldsymbol{x})}=\boldsymbol{0}$ が成り立つため $h(\boldsymbol{x})$ の一意性から $f(\boldsymbol{x})=g(\boldsymbol{x})=h(\boldsymbol{x})$ である．従って $U_2\subset V$ となるため V は開集合でもある．U_1 の連結性から $U_1=V$, すなわち f の一意性が得られる．

(第三段階) f は U_0 で C^1 級であることを示す．$\boldsymbol{x}\in U_0$ に対し, $S_{\boldsymbol{x}}=D_1F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}$, $T_{\boldsymbol{x}}=D_2F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}$ とおくと, F が $\binom{\boldsymbol{x}}{f(\boldsymbol{x})}$ で微分可能であることと, $F'\binom{\boldsymbol{x}}{f(\boldsymbol{x})}\binom{\boldsymbol{u}}{\boldsymbol{v}}=\left(D_1F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}\ D_2F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}\right)\binom{\boldsymbol{u}}{\boldsymbol{v}}=S_{\boldsymbol{x}}\boldsymbol{u}+T_{\boldsymbol{x}}\boldsymbol{v}$ から,

$$\lim_{\binom{\boldsymbol{u}}{\boldsymbol{v}}\to\binom{\boldsymbol{0}}{\boldsymbol{0}}}\frac{F\binom{\boldsymbol{x}+\boldsymbol{u}}{f(\boldsymbol{x})+\boldsymbol{v}}-F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}-S_{\boldsymbol{x}}\boldsymbol{u}-T_{\boldsymbol{x}}\boldsymbol{v}}{\left\|\binom{\boldsymbol{u}}{\boldsymbol{v}}\right\|}=\boldsymbol{0}$$

だから, 任意の $\varepsilon>0$ に対して $\delta>0$ で, $\|\boldsymbol{u}\|,\|\boldsymbol{v}\|<\delta$ ならば次の不等式を満たすものがある．

$$\left\|F\binom{\boldsymbol{x}+\boldsymbol{u}}{f(\boldsymbol{x})+\boldsymbol{v}}-F\binom{\boldsymbol{x}}{f(\boldsymbol{x})}-S_{\boldsymbol{x}}\boldsymbol{u}-T_{\boldsymbol{x}}\boldsymbol{v}\right\|\leqq\varepsilon\left\|\binom{\boldsymbol{u}}{\boldsymbol{v}}\right\|=\varepsilon\sqrt{\|\boldsymbol{v}\|^2+\|\boldsymbol{v}\|^2}\leqq\varepsilon(\|\boldsymbol{u}\|+\|\boldsymbol{v}\|)\ \cdots(i)$$

$\boldsymbol{x}, \boldsymbol{x}+\boldsymbol{u} \in U_0$ に対し, $\boldsymbol{v} = f(\boldsymbol{x}+\boldsymbol{u}) - f(\boldsymbol{x})$ とおくと, $F\begin{pmatrix} \boldsymbol{x}+\boldsymbol{u} \\ f(\boldsymbol{x})+\boldsymbol{v} \end{pmatrix} = F\begin{pmatrix} \boldsymbol{x}+\boldsymbol{u} \\ f(\boldsymbol{x}+\boldsymbol{u}) \end{pmatrix} = \boldsymbol{0}$ であり, f の連続性から $0 < \gamma \leqq \delta$ で, $\|\boldsymbol{u}\| < \gamma$ ならば $\|\boldsymbol{v}\| < \delta$ となるものがある. (i) より $\|S_{\boldsymbol{x}}\boldsymbol{u} + T_{\boldsymbol{x}}\boldsymbol{v}\| \leqq \varepsilon(\|\boldsymbol{u}\| + \|\boldsymbol{v}\|)$ であり, $T_{\boldsymbol{x}}$ は正則行列だから, 命題 2.1 と上の不等式から次の不等式が得られる.

$$\|\boldsymbol{v} + T_{\boldsymbol{x}}^{-1}S_{\boldsymbol{x}}\boldsymbol{u}\| = \|T_{\boldsymbol{x}}^{-1}(S_{\boldsymbol{x}}\boldsymbol{u} + T_{\boldsymbol{x}}\boldsymbol{v})\| \leqq \|T_{\boldsymbol{x}}^{-1}\|\|S_{\boldsymbol{x}}\boldsymbol{u} + T_{\boldsymbol{x}}\boldsymbol{v}\| \leqq \varepsilon\|T_{\boldsymbol{x}}^{-1}\|(\|\boldsymbol{u}\| + \|\boldsymbol{v}\|) \cdots (ii)$$

$c = 2\|T_{\boldsymbol{x}}^{-1}S_{\boldsymbol{x}}\| + 1$ とおくと, $\varepsilon \leqq \dfrac{1}{2\|T_{\boldsymbol{x}}^{-1}\|}$ ならば (ii) から

$$\|\boldsymbol{v}\| - \frac{c-1}{2}\|\boldsymbol{u}\| = \|\boldsymbol{v}\| - \|T_{\boldsymbol{x}}^{-1}S_{\boldsymbol{x}}\|\|\boldsymbol{u}\| \leqq \|\boldsymbol{v}\| - \|T_{\boldsymbol{x}}^{-1}S_{\boldsymbol{x}}\boldsymbol{u}\| \leqq \|\boldsymbol{v} + T_{\boldsymbol{x}}^{-1}S_{\boldsymbol{x}}\boldsymbol{u}\| \leqq \frac{1}{2}(\|\boldsymbol{u}\| + \|\boldsymbol{v}\|)$$

だから $\|\boldsymbol{v}\| \leqq c\|\boldsymbol{u}\|$ が得られる. 故に $\|\boldsymbol{u}\| < \gamma$ ならば

$$\left\|f(\boldsymbol{x}+\boldsymbol{u}) - f(\boldsymbol{x}) + D_2F\left(\begin{smallmatrix} \boldsymbol{x} \\ f(\boldsymbol{x}) \end{smallmatrix}\right)^{-1} D_1F\left(\begin{smallmatrix} \boldsymbol{x} \\ f(\boldsymbol{x}) \end{smallmatrix}\right)\boldsymbol{u}\right\| = \|\boldsymbol{v} + T_{\boldsymbol{x}}^{-1}S_{\boldsymbol{x}}\boldsymbol{u}\| \leqq \varepsilon(c+1)\|T_{\boldsymbol{x}}^{-1}\|\|\boldsymbol{u}\|$$

が成り立つため, これは f が $\boldsymbol{x}$ で微分可能で $f'(\boldsymbol{x})$ が $-D_2F\left(\begin{smallmatrix} \boldsymbol{x} \\ f(\boldsymbol{x}) \end{smallmatrix}\right)^{-1} D_1F\left(\begin{smallmatrix} \boldsymbol{x} \\ f(\boldsymbol{x}) \end{smallmatrix}\right)$ であることを意味する. さらに $\boldsymbol{x} \in U_0$ を $-D_2F\left(\begin{smallmatrix} \boldsymbol{x} \\ f(\boldsymbol{x}) \end{smallmatrix}\right)^{-1} D_1F\left(\begin{smallmatrix} \boldsymbol{x} \\ f(\boldsymbol{x}) \end{smallmatrix}\right) \in M_{m,n}(\boldsymbol{R})$ に写す写像 $U_0 \to M_{m,n}(\boldsymbol{R})$ の各成分の関数は, $\dfrac{\partial F_i}{\partial x_j}$ と f の連続性によって連続だから f は U_0 で C^1 級写像である. □

定理 5.2 においてとくに $m = n = 1$ の場合を考えれば, 次の主張が成り立つ.

系 5.3 $\boldsymbol{R} \times \boldsymbol{R}$ の開集合 U から $\boldsymbol{R}$ への C^1 級写像 F が $\left(\begin{smallmatrix} x_0 \\ y_0 \end{smallmatrix}\right) \in U$ に対して $F\left(\begin{smallmatrix} x_0 \\ y_0 \end{smallmatrix}\right) = 0$ かつ $\dfrac{\partial F}{\partial y}\left(\begin{smallmatrix} x_0 \\ y_0 \end{smallmatrix}\right) \neq 0$ を満たすとき, x_0 を含む開区間 U_1 と C^1 級関数 $f : U_1 \to \boldsymbol{R}$ で $f(x_0) = y_0$ かつ, すべての $x \in U_1$ に対して $\left(\begin{smallmatrix} x \\ f(x) \end{smallmatrix}\right) \in U$, $F\left(\begin{smallmatrix} x \\ f(x) \end{smallmatrix}\right) = 0$ を満たすものがただ 1 つ存在する. さらに f の微分は $f'(x) = -\dfrac{\frac{\partial F}{\partial x}\left(\begin{smallmatrix} x \\ f(x) \end{smallmatrix}\right)}{\frac{\partial F}{\partial y}\left(\begin{smallmatrix} x \\ f(x) \end{smallmatrix}\right)}$ で与えられる.

命題 5.4 定理 5.2 の仮定のもとで F が $\left(\begin{smallmatrix} \boldsymbol{x}_0 \\ \boldsymbol{y}_0 \end{smallmatrix}\right)$ の近傍で C^r 級写像ならば f は $\boldsymbol{x}_0$ の近傍で C^r 級写像である.

証明 定理 5.2 より f は C^1 級写像である. $G\left(\begin{smallmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{smallmatrix}\right) = -D_2F\left(\begin{smallmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{smallmatrix}\right)^{-1}D_1F\left(\begin{smallmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{smallmatrix}\right)$ によって写像 G を定めれば, G の各成分の関数は $\dfrac{\partial F_i}{\partial x_j}$ の連続性により $\left(\begin{smallmatrix} \boldsymbol{x}_0 \\ \boldsymbol{y}_0 \end{smallmatrix}\right)$ の近傍で G は C^{r-1} 級写像である. 帰納的に f が $\boldsymbol{x}_0$ の近傍で C^k 級写像 $(1 \leqq k \leqq r-1)$ であると仮定すると, $\boldsymbol{x}$ を $\left(\begin{smallmatrix} \boldsymbol{x} \\ f(\boldsymbol{x}) \end{smallmatrix}\right)$ に対応させる写像は C^k 級写像である. 定理 5.2 より $f'(\boldsymbol{x}) = G\left(\begin{smallmatrix} \boldsymbol{x} \\ f(\boldsymbol{x}) \end{smallmatrix}\right)$ だから f' は $\boldsymbol{x}$ を $\left(\begin{smallmatrix} \boldsymbol{x} \\ f(\boldsymbol{x}) \end{smallmatrix}\right)$ に対応させる C^k 級写像と G との合成写像であるため, 命題 3.7 により C^k 級写像である. 従って命題 2.10 の (1) により f は C^{k+1} 級写像だから, 帰納法によって C^r 写像である. □

定理 5.5 (逆写像定理) f を $\boldsymbol{x}_0 \in \boldsymbol{R}^n$ の開近傍 U から $\boldsymbol{R}^n$ への C^r 級写像とする. $f'(\boldsymbol{x}_0)$ が正則行列であれば U に含まれる $\boldsymbol{x}_0$ の開近傍 U_0 で f の U_0 への制限が $\boldsymbol{R}^n$ における $\boldsymbol{y}_0 = f(\boldsymbol{x}_0)$ のある開近傍の上への同相写像になるようなものが存在する. さらに, この同相写像の逆写像 $g : f(U_0) \to U_0$ も C^r 級写像である.

証明 $F : U \times \boldsymbol{R}^n \to \boldsymbol{R}^n$ を $F\left(\begin{smallmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{smallmatrix}\right) = f(\boldsymbol{x}) - \boldsymbol{y}$ で定義すれば, $F\left(\begin{smallmatrix} \boldsymbol{x}_0 \\ \boldsymbol{y}_0 \end{smallmatrix}\right) = \boldsymbol{0}$ であり, $D_1F\left(\begin{smallmatrix} \boldsymbol{x}_0 \\ \boldsymbol{y}_0 \end{smallmatrix}\right) = f'(\boldsymbol{x}_0)$ は正則行列だから, $\boldsymbol{x}$ と $\boldsymbol{y}$ を入れかえた形の定理 5.2 が F に適用できる. 命題 5.4 より $\boldsymbol{y}_0$ の開近傍 U_1 と C^r 級写像 $g : U_1 \to U$ で, $g(\boldsymbol{y}_0) = \boldsymbol{x}_0$ かつ任意の $\boldsymbol{y} \in U_1$ に対して $F\left(\begin{smallmatrix} g(\boldsymbol{y}) \\ \boldsymbol{y} \end{smallmatrix}\right) = \boldsymbol{0}$ すなわち $f(g(\boldsymbol{y})) = \boldsymbol{y}$ を満たすものがあり, $g'(\boldsymbol{y}_0) = -D_1F\left(\begin{smallmatrix} \boldsymbol{x}_0 \\ \boldsymbol{y}_0 \end{smallmatrix}\right)^{-1} D_2F\left(\begin{smallmatrix} \boldsymbol{x}_0 \\ \boldsymbol{y}_0 \end{smallmatrix}\right) = f'(\boldsymbol{x}_0)^{-1}$ だから $g'(\boldsymbol{y}_0)$ は正則行列である. f のかわりに g に対して上の議論を行なうと, U に含まれる $\boldsymbol{x}_0$ の開近傍 U_0 と C^r 級写像 $h : U_0 \to U_1$ で, $h(\boldsymbol{x}_0) = \boldsymbol{y}_0$ かつ任意の $\boldsymbol{x} \in U_0$ に対して $g(h(\boldsymbol{x})) = \boldsymbol{x}$ を満たすものが存在する. $\boldsymbol{x} \in U_0$ ならば $h(\boldsymbol{x}) \in U_1$ だから $f(\boldsymbol{x}) = f(g(h(\boldsymbol{x}))) = h(\boldsymbol{x})$ が成り立つため, $g(f(\boldsymbol{x})) = \boldsymbol{x}$ である. 従って $g(f(U_0)) = U_0$ だから $f(U_0) \subset g^{-1}(U_0)$ であり, $\boldsymbol{y} \in g^{-1}(U_0)$ ならば $g(\boldsymbol{y}) \in U_0$ だから $\boldsymbol{y} = f(g(\boldsymbol{y})) \in f(U_0)$ が成り立つため, $g^{-1}(U_0) \subset f(U_0)$ である. 故に $f(U_0) = g^{-1}(U_0)$ が成り立ち, g の連続性から $f(U_0)$ は $\boldsymbol{R}^n$ の開集合である. 以上から f を U_0 から $f(U_0)$ への写像とみなし, g を $f(U_0) = g^{-1}(U_0)$ から U_0 への写像とみなせば, g は f の逆写像になるため f は U_0 から $f(U_0)$ の上への同相写像である. □

6 Lagrange 乗数

$\boldsymbol{R}^{n+m}$ の開集合 X から $\boldsymbol{R}^m$ への写像 F に対し, 条件 $F(\boldsymbol{x})=\boldsymbol{0}$ を満たす点全体からなる X の部分集合を M とする. X で定義された実数値関数 φ が与えられたとき, φ の定義域を M に制限して得られる関数 $\varphi|_M$ が M の点 $\boldsymbol{p}$ で極値をとるための必要条件は次の定理で与えられる.

定理 6.1 X を $\boldsymbol{R}^{n+m}$ の開集合とし, $F:X\to\boldsymbol{R}^m$ は C^1 級写像で, $M=\{\boldsymbol{x}\in X|\,F(\boldsymbol{x})=\boldsymbol{0}\}$ の任意の点 $\boldsymbol{x}$ において $F'(\boldsymbol{x})$ の階数は m であるとする. C^1 級関数 $\varphi:X\to\boldsymbol{R}$ の定義域を M に制限した関数 $\varphi|_M:M\to\boldsymbol{R}$ が $\boldsymbol{p}\in M$ において極値をとるならば, $\lambda_1,\lambda_2,\dots,\lambda_m\in\boldsymbol{R}$ で $\varphi'(\boldsymbol{p})=\big(\lambda_1\ \lambda_2\cdots\lambda_m\big)F'(\boldsymbol{p})$ を満たすものが存在する.

証明 $\boldsymbol{x}\in X$ に対し, 定理 5.2 の証明で用いた $m\times n$ 行列 $D_1F(\boldsymbol{x})$, m 次正方行列 $D_2F(\boldsymbol{x})$ を考える. $F'(\boldsymbol{p})$ の階数は m だから $\boldsymbol{R}^{n+m}$ の座標の順序を入れ替えることにより, $D_2F(\boldsymbol{p})$ は正則であると仮定する. $\boldsymbol{p}=\left(\begin{smallmatrix}\boldsymbol{p}_1\\ \boldsymbol{p}_2\end{smallmatrix}\right)$ ($\boldsymbol{p}_1\in\boldsymbol{R}^n$, $\boldsymbol{p}_2\in\boldsymbol{R}^m$) とおくと, 定理 5.2 から $\boldsymbol{p}_1$ の開近傍 U と C^1 級写像 $f:U\to\boldsymbol{R}^m$ で, $f(\boldsymbol{p}_1)=\boldsymbol{p}_2$ であり, 各 $\boldsymbol{x}\in U$ に対して $\left(\begin{smallmatrix}\boldsymbol{x}\\ f(\boldsymbol{x})\end{smallmatrix}\right)\in X$ かつ $F\left(\begin{smallmatrix}\boldsymbol{x}\\ f(\boldsymbol{x})\end{smallmatrix}\right)=\boldsymbol{0}$ を満たすものがある. 関数 $\psi:U\to\boldsymbol{R}$ を $\psi(\boldsymbol{x})=\varphi\left(\begin{smallmatrix}\boldsymbol{x}\\ f(\boldsymbol{x})\end{smallmatrix}\right)$ で定めると, ψ は $\boldsymbol{p}_1$ で極値をとるため, $\psi'(\boldsymbol{p}_1)=O$ である. $\boldsymbol{x}\in X$ に対して

$$D_1\varphi(\boldsymbol{x})=\left(\frac{\partial\varphi}{\partial x_1}(\boldsymbol{x})\quad\frac{\partial\varphi}{\partial x_2}(\boldsymbol{x})\quad\cdots\quad\frac{\partial\varphi}{\partial x_n}(\boldsymbol{x})\right),\quad D_2\varphi(\boldsymbol{x})=\left(\frac{\partial\varphi}{\partial x_{n+1}}(\boldsymbol{x})\quad\frac{\partial\varphi}{\partial x_{n+2}}(\boldsymbol{x})\quad\cdots\quad\frac{\partial\varphi}{\partial x_{n+m}}(\boldsymbol{x})\right)$$

とおくと, 合成写像の微分法から $\boldsymbol{x}\in U$ に対し, $\psi'(\boldsymbol{x})=D_1\varphi\left(\begin{smallmatrix}\boldsymbol{x}\\ f(\boldsymbol{x})\end{smallmatrix}\right)+D_2\varphi\left(\begin{smallmatrix}\boldsymbol{x}\\ f(\boldsymbol{x})\end{smallmatrix}\right)f'(\boldsymbol{x})$ である. さらに, 定理 5.2 により $f'(\boldsymbol{x})=-D_2F\left(\begin{smallmatrix}\boldsymbol{x}\\ f(\boldsymbol{x})\end{smallmatrix}\right)^{-1}D_1F\left(\begin{smallmatrix}\boldsymbol{x}\\ f(\boldsymbol{x})\end{smallmatrix}\right)$ が成り立つため,

$$\psi'(\boldsymbol{x})=D_1\varphi\left(\begin{smallmatrix}\boldsymbol{x}\\ f(\boldsymbol{x})\end{smallmatrix}\right)-D_2\varphi\left(\begin{smallmatrix}\boldsymbol{x}\\ f(\boldsymbol{x})\end{smallmatrix}\right)D_2F\left(\begin{smallmatrix}\boldsymbol{x}\\ f(\boldsymbol{x})\end{smallmatrix}\right)^{-1}D_1F\left(\begin{smallmatrix}\boldsymbol{x}\\ f(\boldsymbol{x})\end{smallmatrix}\right)$$

が得られる. $\psi'(\boldsymbol{p}_1)=O$, $f(\boldsymbol{p}_1)=\boldsymbol{p}_2$, $\boldsymbol{p}=\left(\begin{smallmatrix}\boldsymbol{p}_1\\ \boldsymbol{p}_2\end{smallmatrix}\right)$ だから上式より, $D_1\varphi(\boldsymbol{p})=D_2\varphi(\boldsymbol{p})D_2F(\boldsymbol{p})^{-1}D_1F(\boldsymbol{p})$ を得る. そこで $D_2\varphi(\boldsymbol{p})D_2F(\boldsymbol{p})^{-1}=\big(\lambda_1\ \lambda_2\cdots\lambda_m\big)$ とおくと, 次の等式が成り立つ.

$$D_1\varphi(\boldsymbol{p})=\big(\lambda_1\ \lambda_2\cdots\lambda_m\big)D_1F(\boldsymbol{p}),\qquad D_2\varphi(\boldsymbol{p})=\big(\lambda_1\ \lambda_2\cdots\lambda_m\big)D_2F(\boldsymbol{p})$$

一方, $D_1\varphi(\boldsymbol{p})$, $D_2\varphi(\boldsymbol{p})$, $D_1F(\boldsymbol{p})$, $D_2F(\boldsymbol{p})$ の定義から $\varphi'(\boldsymbol{p})=\Big(D_1\varphi(\boldsymbol{p})\ D_2\varphi(\boldsymbol{p})\Big)$, $F'(\boldsymbol{p})=\Big(D_1F(\boldsymbol{p})\ D_2F(\boldsymbol{p})\Big)$ だから, 上式から $\varphi'(\boldsymbol{p})=\big(\lambda_1\ \lambda_2\cdots\lambda_m\big)F'(\boldsymbol{p})$ を得る. □

上の定理における $1\times m$ 行列 $\big(\lambda_1\ \lambda_2\cdots\lambda_m\big)$ を Lagrange 乗数という. とくに $m=1$ の場合に上の定理は, 次のようになる.

系 6.2 X を $\boldsymbol{R}^{n+1}$ の開集合とし, $F:X\to\boldsymbol{R}$ は C^1 級関数で, $M=\{\boldsymbol{x}\in X|\,F(\boldsymbol{x})=0\}$ の任意の点 $\boldsymbol{x}$ に対し, $\dfrac{\partial F}{\partial x_j}(\boldsymbol{x})\neq 0$ となる $j=1,2,\dots,n+1$ があるとする. C^1 級関数 $\varphi:X\to\boldsymbol{R}$ の定義域を M に制限した関数 $\varphi|_M:M\to\boldsymbol{R}$ が $\boldsymbol{p}\in M$ において極値をとるならば, $\lambda\in\boldsymbol{R}$ で, すべての $j=1,2,\dots,n+1$ に対して $\dfrac{\partial\varphi}{\partial x_j}(\boldsymbol{p})=\lambda\dfrac{\partial F}{\partial x_j}(\boldsymbol{p})$ を満たすものがある.

さらに $n=1$ の場合は次のようになる.

系 6.3 X を $\boldsymbol{R}^2$ の開集合とし, $F:X\to\boldsymbol{R}$ は C^1 級関数で, $M=\{\boldsymbol{x}\in X|\,F(\boldsymbol{x})=0\}$ の任意の点 $\boldsymbol{x}$ 対し, $\dfrac{\partial F}{\partial x}(\boldsymbol{x})$ または $\dfrac{\partial F}{\partial y}(\boldsymbol{x})$ は 0 でないとする. C^1 級関数 $\varphi:X\to\boldsymbol{R}$ の定義域を M に制限した関数 $\varphi|_M:M\to\boldsymbol{R}$ が $\boldsymbol{p}\in M$ において極値をとるならば, $\lambda\in\boldsymbol{R}$ で, $\dfrac{\partial\varphi}{\partial x}(\boldsymbol{p})=\lambda\dfrac{\partial F}{\partial x}(\boldsymbol{p})$ かつ $\dfrac{\partial\varphi}{\partial y}(\boldsymbol{p})=\lambda\dfrac{\partial F}{\partial y}(\boldsymbol{p})$ を満たすものがある.

7　線形代数学の復習

$\boldsymbol{K}$ によって実数全体の集合 $\boldsymbol{R}$ または複素数全体の集合 $\boldsymbol{C}$ を表す．また $\boldsymbol{K}$ の要素を成分とする $m\times n$ 行列全体のなす $\boldsymbol{K}$ 上のベクトル空間を $M_{m,n}(\boldsymbol{K})$ で表し，$M_{n,n}(\boldsymbol{K})=M_n(\boldsymbol{K})$ とおく．

定義 7.1　$A=(a_{jk})\in M_{m,n}(\boldsymbol{C})$ に対し，$\bar{a}_{jk}$ を (j,k) 成分とする $m\times n$ 行列 を A の共役行列と呼んで $\overline{A}$ で表し，a_{kj} を (j,k) 成分とする $n\times m$ 行列を A の転置行列と呼んで ${}^tA$ で表す．このとき，${}^t(\overline{A})=\overline{{}^tA}$ が成り立つが，${}^t(\overline{A})$ を A の随伴行列と呼んで A^* で表す．

定義 7.2　$A=(a_{jk})\in M_n(\boldsymbol{K})$ に対し，対角成分すべての和 $\sum\limits_{j=1}^{n} a_{jj}$ を A のトレースといい，$\operatorname{tr} A$ で表す．

次の命題 7.3, 命題 7.4 は容易に確かめられる．

命題 7.3　(1) $A,B\in M_{m,n}(\boldsymbol{K})$, $c\in\boldsymbol{K}$ に対して ${}^t(A+B)={}^tA+{}^tB$, $\overline{A+B}=\overline{A}+\overline{B}$, $(A+B)^*=A^*+B^*$, ${}^t(cA)=c\,{}^tA$, $\overline{cA}=\bar{c}\overline{A}$, $(cA)^*=\bar{c}A^*$ が成り立つ．また $m=n$ ならば $|\overline{A}|=|A^*|=\overline{|A|}$ である．

(2) $A\in M_{m,n}(\boldsymbol{K})$, $B\in M_{n,k}(\boldsymbol{K})$ に対して ${}^t(AB)={}^tB\,{}^tA$, $\overline{AB}=\overline{A}\,\overline{B}$, $(AB)^*=B^*A^*$ が成り立つ．

命題 7.4　(1) $A,B\in M_n(\boldsymbol{K})$, $c\in\boldsymbol{K}$ に対して以下の等式が成り立つ．

$$\operatorname{tr}(A+B)=\operatorname{tr}A+\operatorname{tr}B,\quad \operatorname{tr}(cA)=c\operatorname{tr}A,\quad \operatorname{tr}{}^tA=\operatorname{tr}A,\quad \operatorname{tr}\overline{A}=\operatorname{tr}A^*=\overline{\operatorname{tr}A}$$

(2) $A\in M_{m,n}(\boldsymbol{K})$, $B\in M_{n,m}(\boldsymbol{K})$ に対して $\operatorname{tr}(AB)=\operatorname{tr}(BA)$ が成り立つ．とくに $m=n$ の場合，n 次正則行列 P に対して $\operatorname{tr}(P^{-1}AP)=\operatorname{tr}A$ が成り立つ．

$A,B\in M_{m,n}(\boldsymbol{K})$ に対して A と B の内積 (A,B) を $(A,B)=\operatorname{tr}(AB^*)$ で定める．このとき，命題 7.4 から $(A,B)=\operatorname{tr}(B^*A)=\operatorname{tr}({}^t(B^*A))=\operatorname{tr}({}^tA\overline{B})$ であることがわかる．

命題 7.5　$A,B,C\in M_{m,n}(\boldsymbol{K})$, $c\in\boldsymbol{K}$ とするとき，次のことが成り立つ．

(1) $(A+B,C)=(A,C)+(B,C)$, $(A,B+C)=(A,B)+(A,C)$.

(2) $(cA,B)=c(A,B)$, $(A,cB)=\bar{c}(A,B)$.

(3) $(B,A)=\overline{(A,B)}$.

(4) $(A,A)\in\boldsymbol{R}$ かつ $(A,A)\geqq 0$ であり，$A\neq O$ ならば $(A,A)>0$ である．

証明　(1), (2), (3) は命題 7.3 と命題 7.4 から明らかである．$A=(a_{jk})$, $B=(b_{jk})$ とすれば (A,B) の定義により，$(A,B)=\sum\limits_{j=1}^{m}\sum\limits_{k=1}^{n}a_{jk}\overline{b_{jk}}$ だから $(A,A)=\sum\limits_{j=1}^{m}\sum\limits_{k=1}^{n}|a_{jk}|^2$ となるため，(4) の主張が成り立つことがわかる．　□

行列 A の長さ $\|A\|$ を $\|A\|=\sqrt{(A,A)}$ により定める．

注意 7.6　(1) 1×1 行列 $\begin{pmatrix} z \end{pmatrix}$ $(z=x+yi,\ x,y\in\boldsymbol{R})$ の長さは，複素数 z の絶対値 $|z|=\sqrt{x^2+y^2}$ に他ならない．
(2) 以後，$M_{m,1}(\boldsymbol{K})$ を $\boldsymbol{K}$ 上の m 次元数ベクトル空間 $\boldsymbol{K}^m$ と同一視することにより，$\boldsymbol{K}^m$ を $M_{m,n}(\boldsymbol{K})$ の特別な場合とみなして $\boldsymbol{K}^m$ における内積を導入する．とくに，複素数全体の集合 $\boldsymbol{C}$ を $M_m(\boldsymbol{C})$ の特別な場合と考える．

定理 7.7　$A,B\in M_{m,n}(\boldsymbol{K})$ に対し，以下の不等式が成り立つ．

(1) $|(A,B)|\leqq\|A\|\,\|B\|$　(Schwarz の不等式)　　(2) $\|A+B\|\leqq\|A\|+\|B\|$　(三角不等式)

証明　(1) $A=O$ ならば両辺はともに 0 になって不等式は成り立つため，$A\neq O$ と仮定する．命題 7.5 の (1), (2), (3) を用いれば，任意の $t\in\boldsymbol{K}$ に対して次の等式が成り立つ．

$$(tA+B,tA+B)=(tA,tA+B)+(tA,B+B)=(tA,tA)+(tA,B)+(B,tA)+(B,B)=|t|^2\|A\|^2+t(A,B)+\bar{t}\overline{(A,B)}+\|B\|^2$$

命題 7.5 の (4) から $\|A\| \neq 0$ だからとくに $t = -\dfrac{\overline{(A,B)}}{\|A\|^2}$ の場合, $(tA+B, tA+B) = \dfrac{1}{\|A\|^2}(\|A\|^2\|B\|^2 - |(A,B)|^2)$ が成り立つ. 一方, 任意の $t \in \boldsymbol{K}$ に対して命題 7.5 の (4) から $(tA+B, tA+B) \geqq 0$ が成り立つため, 上式から $\|A\|^2\|B\|^2 - |(A,B)|^2 \geqq 0$ が得られる.

(2) (1) で得た等式で $t = 1$ とすれば, 次の等式が得られる.

$$\|A+B\|^2 = (A+B, A+B) = \|A\|^2 + (A,B) + \overline{(A,B)} + \|B\|^2 = \|A\|^2 + 2\mathrm{Re}(A,B) + \|B\|^2$$

(1) より $\mathrm{Re}(A,B) \leqq |(A,B)| \leqq \|A\|\,\|B\|$ だから $\|A+B\|^2 \leqq \|A\|^2 + 2\|A\|\,\|B\| + \|B\|^2 = (\|A\| + \|B\|)^2$ である.□

命題 7.8　$A = (a_{ij}) \in M_{m,n}(\boldsymbol{K})$ とする.

(1) $\lambda \in \boldsymbol{K}$ に対し $\|\lambda A\| = |\lambda|\|A\|$.

(2) $\|A\| \geqq 0$ であり, $\|A\| = 0$ となるのは $A = O$ の場合に限る.

(3) $\|A^*\| = \|{}^tA\| = \|\overline{A}\| = \|A\|$.

(4) 任意の $1 \leqq i \leqq m$, $1 \leqq j \leqq n$ に対し, $|a_{ij}| \leqq \|A\| \leqq \sum_{j=1}^{m}\sum_{k=1}^{n}|a_{jk}|$.

証明　(1), (2) は命題 7.5 からただちにわかる. (3), (4) は $\|A\| = \sqrt{\sum_{j=1}^{m}\sum_{k=1}^{n}|a_{jk}|^2}$ から明らかである. □

命題 2.1 は複素数を成分とする行列に対しても成り立つ.

命題 7.9　$A \in M_{m,n}(\boldsymbol{K})$, $B \in M_{n,k}(\boldsymbol{K})$ に対して不等式 $\|AB\| \leqq \|A\| \cdot \|B\|$ が成り立つ.

証明　$A = (a_{ij})$, $B = (b_{ij})$ とし, A の第 i 行の成分を縦に並べて得られるベクトルを $\boldsymbol{a}_i$, B の第 j 列を $\boldsymbol{b}_j$ とすると, AB の (i,j) 成分が $(\boldsymbol{a}_i, \bar{\boldsymbol{b}}_j)$ であることと Schwarz の不等式から $|(\boldsymbol{a}_i, \bar{\boldsymbol{b}}_j)|^2 \leqq \|\boldsymbol{a}_i\|^2\|\bar{\boldsymbol{b}}_j\|^2 = \|\boldsymbol{a}_i\|^2\|\boldsymbol{b}_j\|^2$ だから

$$\begin{aligned}\|AB\|^2 &= \sum_{i=1}^{m}\sum_{j=1}^{k}|(\boldsymbol{a}_i, \bar{\boldsymbol{b}}_j)|^2 \leqq \sum_{i=1}^{m}\sum_{j=1}^{k}\|\boldsymbol{a}_i\|^2\|\boldsymbol{b}_j\|^2 = \left(\sum_{i=1}^{m}\|\boldsymbol{a}_i\|^2\right)\left(\sum_{j=1}^{k}\|\boldsymbol{b}_j\|^2\right) = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)\left(\sum_{j=1}^{k}\sum_{l=1}^{n}|b_{lj}|^2\right)\\ &= \|A\|^2\|B\|^2\end{aligned}$$

が成り立つ. $\|AB\|^2 \leqq \|A\|^2\|B\|^2$ の両辺の正の平方根をとれば $\|AB\| \leqq \|A\|\|B\|$ が得られる. □

注意 7.6 の (2) で述べたように $\boldsymbol{K}^n = M_{n,1}(\boldsymbol{K})$ とみなして, $\boldsymbol{K}^n$ の内積を考える.

定義 7.10　$\boldsymbol{K}^n$ の零ベクトルでない 2 つのベクトル $\boldsymbol{x}$, $\boldsymbol{y}$ の内積が 0 であるとき, これらは直交するという. $\boldsymbol{K}^n$ の零でないベクトル $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ の相異なる 2 つのベクトルが直交するとき, これらのベクトルを直交系という. さらに, 長さがすべて 1 である直交系を正規直交系といい, 正規直交系である基底を正規直交基底という.

命題 7.11　(1) $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ が $\boldsymbol{K}^n$ の直交系ならば $\dfrac{1}{\|\boldsymbol{v}_1\|}\boldsymbol{v}_1, \dfrac{1}{\|\boldsymbol{v}_2\|}\boldsymbol{v}_2, \ldots, \dfrac{1}{\|\boldsymbol{v}_k\|}\boldsymbol{v}_k$ は $\boldsymbol{K}^n$ の正規直交系である.

(2) $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ が $\boldsymbol{K}^n$ の直交系で, $\boldsymbol{x} = \sum_{j=1}^{k}\lambda_j\boldsymbol{v}_j$ ならば $\lambda_j = \dfrac{(\boldsymbol{x}, \boldsymbol{v}_j)}{\|\boldsymbol{v}_j\|^2}$ である. 故に直交系は 1 次独立である.

(3) $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ が $\boldsymbol{K}^n$ の正規直交基底ならば任意の $\boldsymbol{x} \in \boldsymbol{K}^n$ は $\boldsymbol{x} = \sum_{j=1}^{n}(\boldsymbol{x}, \boldsymbol{v}_j)\boldsymbol{v}_j$ と表せる.

証明　(1) 命題 7.8 の (1) により $j = 1, 2, \ldots, n$ に対して $\left\|\dfrac{1}{\|\boldsymbol{v}_j\|}\boldsymbol{v}_j\right\| = 1$ である.

(2) $\boldsymbol{y} = \sum_{j=1}^{k}\lambda_j\boldsymbol{v}_j$ ならば $(\boldsymbol{y}, \boldsymbol{v}_l) = \left(\sum_{j=1}^{k}\lambda_j\boldsymbol{v}_j, \boldsymbol{v}_l\right) = \sum_{j=1}^{k}\lambda_j(\boldsymbol{v}_j, \boldsymbol{v}_l) = \lambda_l\|\boldsymbol{v}_l\|^2$ より結果が得られる.

(3) は (2) からただちにわかる. □

補題 7.12　$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ を $\boldsymbol{K}^n$ の基底とする．$\boldsymbol{K}^n$ のベクトル $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ が $\boldsymbol{w}_j = a_{1j}\boldsymbol{v}_1 + a_{2j}\boldsymbol{v}_2 + \cdots + a_{jj}\boldsymbol{v}_j$ $(a_{ij} \in \boldsymbol{K},\ i, j = 1, 2, \ldots, n)$ の形に表されていて，$a_{11}a_{22}\cdots a_{nn} \neq 0$ ならば $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ は $\boldsymbol{K}^n$ の基底である．

証明　各 $j = 1, 2, \ldots, n$ に対し $\boldsymbol{v}_j = b_{1j}\boldsymbol{w}_1 + b_{2j}\boldsymbol{w}_2 + \cdots + b_{jj}\boldsymbol{w}_j$ を満たす $b_{1j}, b_{2j}, \ldots, b_{jj} \in \boldsymbol{K}$ が存在することを j による帰納法で示す．$j = 1$ の場合は $\boldsymbol{w}_1 = a_{11}\boldsymbol{v}_1$ と $a_{11} \neq 0$ より，$b_{11} = \dfrac{1}{a_{11}}$ とおけば $\boldsymbol{v}_1 = b_{11}\boldsymbol{w}_1$ である．$k = 1, 2, \ldots, j-1$ に対して $\boldsymbol{v}_k = b_{1k}\boldsymbol{w}_1 + b_{2k}\boldsymbol{w}_2 + \cdots + b_{kk}\boldsymbol{w}_k$ を満たす $b_{1k}, b_{2k}, \ldots, b_{kk} \in \boldsymbol{K}$ が存在すると仮定する．$\boldsymbol{w}_j = a_{1j}\boldsymbol{v}_1 + a_{2j}\boldsymbol{v}_2 + \cdots + a_{jj}\boldsymbol{v}_j$ と $a_{jj} \neq 0$ より，

$$\boldsymbol{v}_j = -\frac{1}{a_{jj}}\left(\sum_{k=1}^{j-1} a_{kj}\boldsymbol{v}_k + \boldsymbol{w}_j\right) = -\frac{1}{a_{jj}}\left(\sum_{k=1}^{j-1} a_{kj}\left(\sum_{i=1}^{k} b_{ik}\boldsymbol{w}_i\right) + \boldsymbol{w}_j\right) = -\frac{1}{a_{jj}}\left(\sum_{i=1}^{j-1}\left(\sum_{k=i}^{j-1} a_{kj}b_{ik}\right)\boldsymbol{w}_i + \boldsymbol{w}_j\right)$$

となるため，$b_{ij} = -\dfrac{1}{a_{jj}}\sum\limits_{k=i}^{j-1} a_{kj}b_{ik}$ $(i = 1, 2, \ldots, j-1)$，$b_{jj} = -\dfrac{1}{a_{jj}}$ によって $b_{1j}, b_{2j}, \ldots, b_{jj}$ を定めれば $\boldsymbol{v}_j = b_{1j}\boldsymbol{w}_1 + b_{2j}\boldsymbol{w}_2 + \cdots + b_{jj}\boldsymbol{w}_j$ が成り立つ．

$\boldsymbol{K}^n$ の任意のベクトル $\boldsymbol{x}$ は $\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n$ と表され，各 $\boldsymbol{v}_j$ は $\boldsymbol{v}_j = b_{1j}\boldsymbol{w}_1 + b_{2j}\boldsymbol{w}_2 + \cdots + b_{jj}\boldsymbol{w}_j$ と表されるため，$\boldsymbol{x}$ は $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ の 1 次結合である．従って $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ は $\boldsymbol{K}^n$ を生成する．

$c_1\boldsymbol{w}_1 + c_2\boldsymbol{w}_2 + \cdots + c_n\boldsymbol{w}_n = \boldsymbol{0}$ が成り立つとき，$\boldsymbol{w}_j = a_{1j}\boldsymbol{v}_1 + a_{2j}\boldsymbol{v}_2 + \cdots + a_{jj}\boldsymbol{v}_j$ $(j = 1, 2, \ldots, n)$ を代入して

$$\sum_{k=1}^{n}\left(\sum_{j=k}^{n} a_{kj}c_j\right)\boldsymbol{v}_k = \sum_{j=1}^{n}\sum_{k=1}^{j} a_{kj}c_j\boldsymbol{v}_k = \boldsymbol{0}$$

を得る．$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ は 1 次独立だから $j = 1, 2, \ldots, n$ に対して $\sum\limits_{j=k}^{n} a_{kj}c_j = 0$ である．とくに $k = n$ の場合 $a_{nn}c_n = 0$ で $a_{nn} \neq 0$ だから $c_n = 0$ である．$c_n = c_{n-1} = \cdots = c_{i+1} = 0$ $(1 \leqq i \leqq n-1)$ が成り立つと仮定すれば $\sum\limits_{j=i}^{n} a_{ij}c_j = 0$ より $a_{ii}c_i = 0$ を得る．$a_{ii} \neq 0$ だから $c_i = 0$ となり，帰納的に $c_n = c_{n-1} = \cdots = c_1 = 0$ が示されるため $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ は 1 次独立である．□

定理 7.13　$\boldsymbol{K}^n$ の基底 $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ に対し，$\boldsymbol{K}^n$ の正規直交基底 $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ で次の条件を満たすものがある．

$(*)$ 各 $j = 1, 2, \ldots, n$ に対し，$a_{1j}, a_{2j}, \ldots, a_{j\,j-1} \in \boldsymbol{K}$ と正の実数 a_{jj} で $\boldsymbol{w}_j = a_{1j}\boldsymbol{v}_1 + a_{2j}\boldsymbol{v}_2 + \cdots + a_{jj}\boldsymbol{v}_j$ を満たすものがある．

証明　$\boldsymbol{u}_1 = \boldsymbol{v}_1$ とおき，$j = 2, 3, \ldots, n$ に対して $\boldsymbol{u}_j$ を帰納的に次のように定める．

$$\boldsymbol{u}_j = \boldsymbol{v}_j - \sum_{i=1}^{j-1}\frac{(\boldsymbol{v}_j, \boldsymbol{u}_i)}{\|\boldsymbol{u}_i\|^2}\boldsymbol{u}_i = \boldsymbol{v}_j - \frac{(\boldsymbol{v}_j, \boldsymbol{u}_1)}{\|\boldsymbol{u}_1\|^2}\boldsymbol{u}_1 - \frac{(\boldsymbol{v}_j, \boldsymbol{u}_2)}{\|\boldsymbol{u}_2\|^2}\boldsymbol{u}_2 - \cdots - \frac{(\boldsymbol{v}_j, \boldsymbol{u}_{j-1})}{\|\boldsymbol{u}_{j-1}\|^2}\boldsymbol{u}_{j-1}$$

このとき，各 j に対して $\boldsymbol{u}_j = b_{1j}\boldsymbol{v}_1 + b_{2j}\boldsymbol{v}_2 + \cdots + b_{j-1\,j}\boldsymbol{v}_{j-1} + \boldsymbol{v}_j$ を満たす $b_{1j}, b_{2j}, \cdots, b_{j-1\,j} \in \boldsymbol{K}$ が存在する．実際，$j = 1$ のときは明らかで，帰納的に $j = 1, 2, \ldots, k-1$ に対してこの主張が正しいと仮定し，$b_{ii} = 1$ とおくと

$$\boldsymbol{u}_k = \boldsymbol{v}_k - \sum_{i=1}^{k-1}\frac{(\boldsymbol{v}_k, \boldsymbol{u}_i)}{\|\boldsymbol{u}_i\|^2}\boldsymbol{u}_i = \boldsymbol{v}_k - \sum_{i=1}^{k-1}\frac{(\boldsymbol{v}_k, \boldsymbol{u}_i)}{\|\boldsymbol{u}_i\|^2}\sum_{l=1}^{i} b_{li}\boldsymbol{v}_l = \boldsymbol{v}_k - \sum_{i=1}^{k-1}\sum_{l=1}^{i}\frac{(\boldsymbol{v}_k, \boldsymbol{u}_i)}{\|\boldsymbol{u}_i\|^2}b_{li}\boldsymbol{v}_l = \boldsymbol{v}_k - \sum_{l=1}^{k-1}\left(\sum_{i=l}^{k-1}\frac{(\boldsymbol{v}_k, \boldsymbol{u}_i)}{\|\boldsymbol{u}_i\|^2}b_{li}\right)\boldsymbol{v}_l$$

だから，$b_{lk} = -\sum\limits_{i=l}^{k-1}\dfrac{(\boldsymbol{v}_k, \boldsymbol{u}_i)}{\|\boldsymbol{u}_i\|^2}b_{li}$ とおけば $j = k$ に対しても主張が成り立つ．従って，もし $\boldsymbol{u}_j = \boldsymbol{0}$ となる j があれば，$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_j$ が 1 次独立であることに矛盾するため，すべての $j = 1, 2, \ldots, n$ に対して $\boldsymbol{u}_j \neq \boldsymbol{0}$ である．

j に関して帰納的に $j > i, k$ かつ $i \neq k$ ならば $(\boldsymbol{u}_i, \boldsymbol{u}_k) = 0$ であると仮定する．$\boldsymbol{u}_j$ の定義から $j > k$ に対し，

$$(\boldsymbol{u}_j, \boldsymbol{u}_k) = (\boldsymbol{v}_j, \boldsymbol{u}_k) - \sum_{i=1}^{j-1}\frac{(\boldsymbol{v}_j, \boldsymbol{u}_i)}{\|\boldsymbol{u}_i\|^2}(\boldsymbol{u}_i, \boldsymbol{u}_k) = (\boldsymbol{v}_j, \boldsymbol{u}_k) - \frac{(\boldsymbol{v}_j, \boldsymbol{u}_k)}{\|\boldsymbol{u}_k\|^2}(\boldsymbol{u}_k, \boldsymbol{u}_k) = 0$$

が成り立つため $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ は直交系である．$\boldsymbol{w}_j = \dfrac{1}{\|\boldsymbol{u}_j\|}\boldsymbol{u}_j$ で $\boldsymbol{w}_j$ を定めれば，$a_{ij} = \dfrac{b_{ij}}{\|\boldsymbol{u}_j\|}$ であり，補題 7.12 から $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ は与えられた条件を満たす $\boldsymbol{K}^n$ の正規直交基底である．□

上の定理の証明のように，与えられた $\boldsymbol{K}^n$ の基底 $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ から正規直交基底 $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ を構成することを，「$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ を直交化する.」といい，この構成法を Schmidt の正規直交化法という．

定義 7.14　$A \in M_n(\boldsymbol{K})$ とする.

(1) $A\boldsymbol{v} = \lambda\boldsymbol{v}$ を満たす $\lambda \in \boldsymbol{K}$ と零でない $\boldsymbol{v} \in \boldsymbol{K}^n$ が存在するとき, λ を A の固有値, $\boldsymbol{v}$ を λ に対する固有ベクトルという.

(2) A の固有値 λ に対し, $V_\lambda = \{\boldsymbol{x} \in \boldsymbol{K}^n \mid A\boldsymbol{x} = \lambda\boldsymbol{x}\}$ とおき, V_λ を λ に対する A の固有空間という.

(3) x を変数とする n 次多項式 $|xE_n - A|$ を A の固有多項式と呼んで $F_A(x)$ で表す. また, n 次方程式 $F_A(x) = 0$ を A の固有方程式という.

命題 7.15　n 次正方行列 A, n 次正則行列 P に対して $F_{P^{-1}AP}(x) = F_A(x)$ が成り立つ.

証明　行列式の性質 $|AB| = |A||B|$ と $|E_n| = 1$ から $F_{P^{-1}AP}(x) = |xE_n - P^{-1}AP| = |P^{-1}(xE_n - A)P| = |P^{-1}||xE_n - A||P| = |P^{-1}||P||xE_n - A| = |P^{-1}P||xE_n - A| = |E_n||xE_n - A| = |xE_n - A| = F_A(x)$ である. □

定理 7.16　$\lambda_1, \lambda_2, \ldots, \lambda_k \in \boldsymbol{K}$ を $A \in M_n(\boldsymbol{K})$ の相異なる固有値として, V_{λ_j} を λ_j に対する A の固有空間とする. $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \boldsymbol{K}^n$ が $j = 1, 2, \ldots, k$ に対して $\boldsymbol{x}_j \in V_{\lambda_j}$ を満たし, さらに $\boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_k = \boldsymbol{0}$ が成り立てば $\boldsymbol{x}_1 = \boldsymbol{x}_2 = \cdots = \boldsymbol{x}_k = \boldsymbol{0}$ である.

証明　k による帰納法で主張を示す. まず $k = 1$ の場合は主張は明らかに成立する. $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_i \in \boldsymbol{K}^n$ が $j = 1, 2, \ldots, i$ に対して $\boldsymbol{x}_j \in V_{\lambda_j}$ を満たし, さらに $\boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_i = \boldsymbol{0}$ が成り立つとき, この等式の両辺に左から A をかけた式から, 両辺を λ_i 倍した式を辺々引くことにより $(\lambda_1 - \lambda_i)\boldsymbol{x}_1 + (\lambda_2 - \lambda_i)\boldsymbol{x}_2 + \cdots + (\lambda_{i-1} - \lambda_i)\boldsymbol{x}_{i-1} = \boldsymbol{0}$ が得られる. そこで, $j = 1, 2, \ldots, i-1$ に対して $\boldsymbol{x}'_j = (\lambda_j - \lambda_i)\boldsymbol{x}_j$ とおくと $\boldsymbol{x}'_j \in V_{\lambda_j}$ かつ $\boldsymbol{x}'_1 + \boldsymbol{x}'_2 + \cdots + \boldsymbol{x}'_{i-1} = \boldsymbol{0}$ が成り立つため, $k = i - 1$ の場合の帰納法の仮定により $\boldsymbol{x}'_1 = \boldsymbol{x}'_2 = \cdots = \boldsymbol{x}'_{i-1} = \boldsymbol{0}$ である. $j < i$ ならば $\lambda_j \neq \lambda_i$ だから $\boldsymbol{x}_1 = \boldsymbol{x}_2 = \cdots = \boldsymbol{x}_{i-1} = \boldsymbol{0}$ が得られ, さらに $\boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_i = \boldsymbol{0}$ より $\boldsymbol{x}_i = \boldsymbol{0}$ である. 故に $k = i$ の場合も主張が成り立つ. □

命題 7.17　$\lambda \in \boldsymbol{K}$ が $A \in M_n(\boldsymbol{K})$ の固有値であるためには $|\lambda E_n - A| = 0$ が成り立つことが必要十分である.

証明　λ を A の固有値とする. λ に対する A の固有ベクトル $\boldsymbol{v}$ をとると, $A\boldsymbol{v} = \lambda\boldsymbol{v}$ より $(\lambda E_n - A)\boldsymbol{v} = \boldsymbol{0}$ である. もし $\lambda E_n - A$ が正則ならば $(\lambda E_n - A)\boldsymbol{v} = \boldsymbol{0}$ の両辺に左から $\lambda E_n - A$ の逆行列をかければ $\boldsymbol{v} = \boldsymbol{0}$ が得られ, $\boldsymbol{v}$ が固有ベクトルであるという仮定に反する. 従って $\lambda E_n - A$ は正則ではないため $|\lambda E_n - A| = 0$ である.

逆に $|\lambda E_n - A| = 0$ ならば $\lambda E_n - A$ は正則ではないため, $\lambda E_n - A$ を係数行列とする斉次連立 1 次方程式は零ベクトルではない解 $\boldsymbol{v}$ をもつ. このとき $(\lambda E_n - A)\boldsymbol{v} = \boldsymbol{0}$ より $A\boldsymbol{v} = \lambda\boldsymbol{v}$ となるため λ は A の固有値である. □

n 次方程式の解の存在については「代数学の基本定理」と呼ばれる次の定理が知られている.

定理 7.18　複素数を係数にもつ方程式 $a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0$ $(a_n \neq 0)$ は $\boldsymbol{C}$ の中に解をもつ. 従って, 複素数を係数にもつ 1 変数の多項式は 1 次式の積に因数分解される.

n 次正方行列 $A = (a_{ij})$ が条件「$i > j$ ならば $a_{ij} = 0$」をみたすとき, A を上半三角行列という.

命題 7.19　A を n 次正方行列, P を n 次正則行列とする. $P^{-1}AP$ が λ_j を (j, j) 成分とする上半三角行列ならば $F_A(x) = (x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n)$ である.

証明　命題 7.15 により

$$F_A(x) = F_{P^{-1}AP}(x) = \begin{vmatrix} x-\lambda_1 & * & \cdots & * \\ & x-\lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ \text{\huge 0} & & & x-\lambda_n \end{vmatrix} = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n).$$

□

命題 7.20　n 次正方行列 A の固有多項式 $F_A(x)$ が $(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$ と因数分解されるとき, 等式 $|A| = \lambda_1\lambda_2\cdots\lambda_n$, $\operatorname{tr} A = \lambda_1+\lambda_2+\cdots+\lambda_n$ が成り立つ.

証明　まず $F_A(x) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$ の両辺に $x=0$ を代入して $|-A| = (-\lambda_1)(-\lambda_2)\cdots(-\lambda_n)$ を得る. この左辺は $(-1)^n|A|$ に等しく, 右辺は $(-1)^n\lambda_1\lambda_2\cdots\lambda_n$ に等しいため, $|A| = \lambda_1\lambda_2\cdots\lambda_n$ が得られる.

また $A = (a_{ij})$ とすれば, $F_A(x) = |xE_n - A| = (x-a_{11})(x-a_{22})\cdots(x-a_{nn}) + (x$ の $n-2$ 次以下の多項式$) = x^n - (\operatorname{tr} A)x^{n-1} + (x$ の $n-2$ 次以下の多項式) となるため $|xE_n - A|$ の x^{n-1} の係数は $-\operatorname{tr} A$ である. 一方 $(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$ の x^{n-1} の係数は $-(\lambda_1+\lambda_2+\cdots+\lambda_n)$ だから $\operatorname{tr} A = \lambda_1+\lambda_2+\cdots+\lambda_n$ を得る.□

定義 7.21　$A \in M_n(\boldsymbol{K})$ に対し, 正則行列 $P \in M_n(\boldsymbol{K})$ で $P^{-1}AP$ が対角行列になるものが存在するとき, A は $\boldsymbol{K}$ 上対角化可能であるという. このとき P は A を対角化するという.

第 j 成分が 1 で, それ以外の成分はすべて 0 である $\boldsymbol{K}^n$ のベクトルを $\boldsymbol{e}_j$ で表す. また, $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n \in \boldsymbol{K}^m$ に対して, $\boldsymbol{a}_j$ を第 j 列とする $m\times n$ 行列を $\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n\end{pmatrix}$ で表す. このとき, $k\times m$ 行列 B に対して等式 $B\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n\end{pmatrix} = \begin{pmatrix}B\boldsymbol{a}_1 & B\boldsymbol{a}_2 & \cdots & B\boldsymbol{a}_n\end{pmatrix}$ が成り立つことに注意する.

定理 7.22　$A \in M_n(\boldsymbol{K})$ が $\boldsymbol{K}$ 上対角化可能ならば A の固有ベクトルからなる $\boldsymbol{K}^n$ の基底が存在する. 逆に $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ を A の固有ベクトルからなる $\boldsymbol{K}^n$ の基底とし, $j = 1, 2, \ldots, n$ に対して $A\boldsymbol{v}_j = \lambda_j\boldsymbol{v}_j$ が成り立つとき, $P = \begin{pmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n\end{pmatrix}$ とおけば, P は正則で

$$P^{-1}AP = \begin{pmatrix}\lambda_1 & & & \text{\huge 0} \\ & \lambda_2 & & \\ & & \ddots & \\ \text{\huge 0} & & & \lambda_n\end{pmatrix}$$

が成り立つ. 従って A が $\boldsymbol{K}$ 上対角化可能であるためには A の固有ベクトルからなる $\boldsymbol{K}^n$ の基底が存在することが必要十分である.

証明　$A \in M_n(\boldsymbol{K})$ が $\boldsymbol{K}$ 上対角化可能であるとする. 正則行列 P で $P^{-1}AP = \begin{pmatrix}\lambda_1\boldsymbol{e}_1 & \lambda_2\boldsymbol{e}_2 & \cdots & \lambda_n\boldsymbol{e}_n\end{pmatrix}$ となるものを考えれば, $AP = P\begin{pmatrix}\lambda_1\boldsymbol{e}_1 & \lambda_2\boldsymbol{e}_2 & \cdots & \lambda_n\boldsymbol{e}_n\end{pmatrix} = \begin{pmatrix}\lambda_1P\boldsymbol{e}_1 & \lambda_2P\boldsymbol{e}_2 & \cdots & \lambda_nP\boldsymbol{e}_n\end{pmatrix}$ より AP の第 j 列 $AP\boldsymbol{e}_j$ は $\lambda_jP\boldsymbol{e}_j$ となる. すなわち λ_j は A の固有値で, P の第 j 列は λ_j に対する A の固有ベクトルである. また, P は正則行列であるため, P の列ベクトル $P\boldsymbol{e}_1, P\boldsymbol{e}_2, \ldots, P\boldsymbol{e}_n$ は $\boldsymbol{K}^n$ の基底である.

$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ を A の固有ベクトルからなる $\boldsymbol{K}^n$ の基底とし, 各 j に対して $A\boldsymbol{v}_j = \lambda_j\boldsymbol{v}_j$ が成り立つとき, $P = \begin{pmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n\end{pmatrix}$ とおけば, $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ は $\boldsymbol{K}^n$ の基底であるため, P は正則である. $P\boldsymbol{e}_j = \boldsymbol{v}_j$ より

$$\begin{aligned} AP &= A\begin{pmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n\end{pmatrix} = \begin{pmatrix}A\boldsymbol{v}_1 & A\boldsymbol{v}_2 & \cdots & A\boldsymbol{v}_n\end{pmatrix} = \begin{pmatrix}\lambda_1\boldsymbol{v}_1 & \lambda_2\boldsymbol{v}_2 & \cdots & \lambda_n\boldsymbol{v}_n\end{pmatrix} \\ &= \begin{pmatrix}\lambda_1P\boldsymbol{e}_1 & \lambda_2P\boldsymbol{e}_2 & \cdots & \lambda_nP\boldsymbol{e}_n\end{pmatrix} = P\begin{pmatrix}\lambda_1\boldsymbol{e}_1 & \lambda_2\boldsymbol{e}_2 & \cdots & \lambda_n\boldsymbol{e}_n\end{pmatrix} \end{aligned}$$

となるため $P^{-1}AP$ は対角行列 $\begin{pmatrix}\lambda_1\boldsymbol{e}_1 & \lambda_2\boldsymbol{e}_2 & \cdots & \lambda_n\boldsymbol{e}_n\end{pmatrix}$ である. □

注意 7.23　$A \in M_n(\boldsymbol{K})$ が $\boldsymbol{K}$ 上対角化可能ならば命題 7.19 により A の固有多項式は $\boldsymbol{K}$ の範囲で 1 次式の積に因数分解される.

命題 7.24　$\lambda \in \boldsymbol{K}$ を $A \in M_n(\boldsymbol{K})$ の固有値, V_λ を λ に対する A の固有空間とする. $\dim V_\lambda = m$ とおけば, A の固有多項式 $F_A(x)$ は $(x-\lambda)^m$ を因数にもつ.

証明　$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m$ を V_λ の基底とし, $\boldsymbol{v}_{m+1}, \boldsymbol{v}_{m+2}, \ldots, \boldsymbol{v}_n \in \boldsymbol{K}^n$ を選んで $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ が $\boldsymbol{K}^n$ の基底になるよ

うにする．$P = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{pmatrix}$ とおくと，P は正則で，$j = 1, 2, \ldots, m$ に対して $A\boldsymbol{v}_j = \lambda \boldsymbol{v}_j = \lambda P\boldsymbol{e}_j$ だから

$$AP = \begin{pmatrix} A\boldsymbol{v}_1 & A\boldsymbol{v}_2 & \cdots & A\boldsymbol{v}_n \end{pmatrix} = \begin{pmatrix} \lambda P\boldsymbol{e}_1 & \lambda P\boldsymbol{e}_2 & \cdots & \lambda P\boldsymbol{e}_m & A\boldsymbol{v}_{m+1} & A\boldsymbol{v}_{m+2} & \cdots & A\boldsymbol{v}_n \end{pmatrix}$$

である．従って $\begin{pmatrix} P^{-1}A\boldsymbol{v}_{m+1} & \cdots & P^{-1}A\boldsymbol{v}_n \end{pmatrix}$ の第 $m+1$ 行以下の行からなる $n-m$ 次正方行列を B とすると

$$P^{-1}AP = \begin{pmatrix} \lambda\boldsymbol{e}_1 & \lambda\boldsymbol{e}_2 & \cdots & \lambda\boldsymbol{e}_m & P^{-1}A\boldsymbol{v}_{m+1} & P^{-1}A\boldsymbol{v}_{m+2} & \cdots & P^{-1}A\boldsymbol{v}_n \end{pmatrix} = \begin{pmatrix} \lambda E_m & * \\ O & B \end{pmatrix}$$

より，命題 7.15 を用いれば

$$F_A(x) = F_{P^{-1}AP}(x) = \begin{vmatrix} (x-\lambda)E_m & * \\ O & xE_{n-m} - B \end{vmatrix} = |(x-\lambda)E_m||xE_{n-m} - B| = (x-\lambda)^m F_B(x)$$

である．故に $F_A(x)$ は $(x-\lambda)^m$ を因数にもつ． □

補題 7.25 $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m \in \boldsymbol{K}^n$ とし，$\boldsymbol{K}^n$ の部分空間 $V_1, V_2, \ldots, V_k$ が次の条件 $(*)$ を満たすとする．

$(*)$ $\boldsymbol{K}^n$ のベクトル $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ が，$i = 1, 2, \ldots, k$ に対して $\boldsymbol{x}_i \in V_i$ を満たし $\boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_k = \boldsymbol{0}$ を満たすならば $\boldsymbol{x}_1 = \boldsymbol{x}_2 = \cdots = \boldsymbol{x}_k = \boldsymbol{0}$ である．

各 $i = 1, 2, \ldots, k$ に対して $\boldsymbol{v}_{s_{i-1}+1}, \boldsymbol{v}_{s_{i-1}+2}, \ldots, \boldsymbol{v}_{s_i}$ $(0 = s_0 < s_1 < \cdots < s_k = m)$ が V_i の 1 次独立なベクトルならば $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m$ は 1 次独立である．

証明 $x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_m\boldsymbol{v}_m = \boldsymbol{0}$ が成り立つとして $\boldsymbol{x}_i = \sum\limits_{j=s_{i-1}+1}^{s_i} x_j\boldsymbol{v}_j$ とおくと $\boldsymbol{x}_i \in V_i$ $(i = 1, 2, \ldots, k)$ であり $\boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_k = \boldsymbol{0}$ が成り立つため仮定により $\boldsymbol{x}_1 = \boldsymbol{x}_2 = \cdots = \boldsymbol{x}_k = \boldsymbol{0}$ である．従って，各 $i = 1, 2, \ldots, k$ に対して $\sum\limits_{j=s_{i-1}+1}^{s_i} x_j\boldsymbol{v}_j = \boldsymbol{0}$ となるため，$\boldsymbol{v}_{s_{i-1}+1}, \boldsymbol{v}_{s_{i-1}+2}, \ldots, \boldsymbol{v}_{s_i}$ の 1 次独立性により $x_{s_{i-1}+1} = x_{s_{i-1}+2} = \cdots = x_{s_i} = 0$ である．故に $x_1 = x_2 = \cdots = x_m = 0$ となるため $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m$ は 1 次独立である． □

固有空間の概念を用いれば，対角化可能であるための条件は次のように言い換えられる．

定理 7.26 $\lambda_1, \lambda_2, \ldots, \lambda_k \in \boldsymbol{K}$ を $A \in M_n(\boldsymbol{K})$ の相異なる固有値の全体として V_{λ_i} を λ_i に対する A の固有空間とする．このとき次の (1), (2), (3) は同値である．

(1) A は $\boldsymbol{K}$ 上対角化可能である．

(2) 任意の $\boldsymbol{x} \in \boldsymbol{K}^n$ に対し $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_k$ となる $\boldsymbol{x}_i \in V_{\lambda_i}$ $(i = 1, 2, \ldots, k)$ がある．

(3) $\dim V_{\lambda_i} = m_i$ とおくと A の固有多項式は $(x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_k)^{m_k}$ と因数分解される．

証明 (1)⇒(2), (3); 定理 7.22 により A の固有ベクトルからなる $\boldsymbol{K}^n$ の基底 $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ が存在するため，各 $j = 1, 2, \ldots, n$ に対して $A\boldsymbol{v}_j = \lambda_{l(j)}\boldsymbol{v}_j$ を満たす $1 \leqq l(j) \leqq k$ がある．$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ の番号を付け直すことにより $1 = l(1) \leqq l(2) \leqq \cdots \leqq l(n) = k$ が成り立つようにして，各 $i = 1, 2, \ldots, k$ に対し，$l(j) = i$ を満たす最大の j を s_i とおく．任意の $\boldsymbol{x} \in \boldsymbol{K}^n$ は $\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n$ と表され，$\boldsymbol{x}_i = \sum\limits_{j=s_{i-1}+1}^{s_i} x_j\boldsymbol{v}_j$ とおくと，$\boldsymbol{x}_i \in V_{\lambda_i}$ $(i = 1, 2, \ldots, k)$ であり，$\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_k$ が成り立つため (2) が成り立つ．

とくに $\boldsymbol{x} \in V_{\lambda_i}$ の場合，$\boldsymbol{x}_1 + \cdots + \boldsymbol{x}_{i-1} + (\boldsymbol{x}_i - \boldsymbol{x}) + \boldsymbol{x}_{i+1} + \cdots + \boldsymbol{x}_k = \boldsymbol{0}$ となり，$\boldsymbol{x}_l \in V_{\lambda_l}$，$\boldsymbol{x}_i - \boldsymbol{x} \in V_{\lambda_i}$ に注意すれば定理 7.16 により $\boldsymbol{x}_i - \boldsymbol{x} = \boldsymbol{0}$ すなわち $\boldsymbol{x} = \boldsymbol{x}_i = \sum\limits_{j=s_{i-1}+1}^{s_i} x_j\boldsymbol{v}_j$ である．故に 1 次独立なベクトル $\boldsymbol{v}_{s_{i-1}+1}, \boldsymbol{v}_{s_{i-1}+2}, \ldots, \boldsymbol{v}_{s_i}$ は V_{λ_i} を生成するため，これらは V_{λ_i} の基底である．従って $i = 1, 2, \ldots, k$ に対して $m_i = \dim V_{\lambda_i} = s_i - s_{i-1}$ である．$P = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{pmatrix}$ とおけば，定理 7.22 から

$$P^{-1}AP = \begin{pmatrix} \lambda_1 E_{m_1} & & & \text{\huge 0} \\ & \lambda_2 E_{m_2} & & \\ & & \ddots & \\ \text{\huge 0} & & & \lambda_k E_{m_k} \end{pmatrix}$$

となるため, 命題 7.15 により $F_A(x) = F_{P^{-1}AP}(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_k)^{m_k}$ と因数分解される.

(3)⇒(1); A の固有多項式の次数は n であるため, 仮定から $m_1+m_2+\cdots+m_k = n$ である. $s_0 = 0$, $s_i = m_1+m_2+\cdots+m_i$ $(i=1,2,\dots,k)$ とおいて, $\boldsymbol{v}_{s_{i-1}+1}, \boldsymbol{v}_{s_{i-1}+2}, \dots, \boldsymbol{v}_{s_i}$ を V_{λ_i} の基底とする. このとき, $s_k = n$ であり, 定理 7.16 と補題 7.25 により $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ は $\boldsymbol{K}^n$ の n 個の 1 次独立なベクトルだから, これらは $\boldsymbol{K}^n$ の基底である. $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ はすべて A の固有ベクトルだから定理 7.22 により A は $\boldsymbol{K}$ 上対角化可能である.

(2)⇒(1); 各 $i = 1,2,\dots,k$ に対して $\boldsymbol{v}_{s_{i-1}+1}, \boldsymbol{v}_{s_{i-1}+2}, \dots, \boldsymbol{v}_{s_i}$ $(0 = s_0 < s_1 < \cdots < s_k)$ が V_{λ_i} の基底になるように $\boldsymbol{K}^n$ のベクトル $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_{s_k}$ を選ぶ. 任意の $\boldsymbol{x} \in \boldsymbol{K}^n$ に対し $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_k$ となる $\boldsymbol{x}_i \in V_{\lambda_i}$ $(i=1,2,\dots,k)$ が存在し, さらに各 i に対して $\boldsymbol{x}_i$ は $\boldsymbol{v}_{s_{i-1}+1}, \boldsymbol{v}_{s_{i-1}+2}, \dots, \boldsymbol{v}_{s_i}$ の 1 次結合になるため, $\boldsymbol{x}$ は $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_{s_k}$ の 1 次結合である. 故に $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_{s_k}$ は $\boldsymbol{K}^n$ を生成する. また, 定理 7.16 と補題 7.25 より $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_{s_k}$ は 1 次独立でもあるから $\boldsymbol{K}^n$ の基底である. 従って $s_k = n$ である. さらに $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ はすべて A の固有ベクトルだから定理 7.22 により A は $\boldsymbol{K}$ 上対角化可能である. □

8 正規行列の対角化

定義 8.1 A を正方行列とする. $AA^* = A^*A$ が成り立つとき A を正規行列, $A^* = A$ が成り立つとき A をエルミート行列といい, $A^* = -A$ が成り立つとき A を歪エルミート行列, $A^*A = E_n$ が成り立つとき A をユニタリー行列という.

注意 8.2 (1) 対角行列, エルミート行列, 歪エルミート行列, ユニタリー行列はいずれも正規行列である.

(2) $A^*A = E_n$ ならば両辺の行列式を考えれば命題 7.3 から $\overline{|A|}|A| = 1$ が得られるため, ユニタリー行列の行列式の値は絶対値が 1 の複素数である. とくに A は正則行列で,$A^* = A^{-1}$ である. 逆に $A^* = A^{-1}$ ならば $A^*A = E_n$ だから A はユニタリー行列である. A が実直交行列ならば A はユニタリー行列で, $|A|$ は実数だから $|A| = \pm 1$ である.

補題 8.3 $A \in M_n(\boldsymbol{K})$, $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{K}^n$ に対して $(A\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, A^*\boldsymbol{y})$ が成り立つ.

証明 内積の定義から $(A\boldsymbol{x}, \boldsymbol{y}) = {}^t(A\boldsymbol{x})\bar{\boldsymbol{y}} = {}^t\boldsymbol{x}{}^tA\bar{\boldsymbol{y}} = {}^t\boldsymbol{x}\overline{\overline{{}^tA}}\bar{\boldsymbol{y}} = {}^t\boldsymbol{x}\overline{A^*\boldsymbol{y}} = (\boldsymbol{x}, A^*\boldsymbol{y})$. □

補題 8.4 λ が正規行列 $A \in M_n(\boldsymbol{K})$ の固有値で, $\boldsymbol{x}$ が λ に対する A の固有ベクトルならば $\bar{\lambda}$ は A^* の固有値で, $\boldsymbol{x}$ は $\bar{\lambda}$ に対する A^* の固有ベクトルである.

証明 補題 8.3 と命題 7.5 より次の等式が得られ, $(A^*\boldsymbol{x} - \bar{\lambda}\boldsymbol{x}, A^*\boldsymbol{x} - \bar{\lambda}\boldsymbol{x}) = 0$ が導かれるため $A^*\boldsymbol{x} = \bar{\lambda}\boldsymbol{x}$ である.

$$\begin{aligned}(A^*\boldsymbol{x}, A^*\boldsymbol{x} - \bar{\lambda}\boldsymbol{x}) &= (A^*\boldsymbol{x}, A^*\boldsymbol{x}) - (A^*\boldsymbol{x}, \bar{\lambda}\boldsymbol{x}) = (AA^*\boldsymbol{x}, \boldsymbol{x}) - \lambda(A^*\boldsymbol{x}, \boldsymbol{x}) = (A^*A\boldsymbol{x}, \boldsymbol{x}) - \lambda(\boldsymbol{x}, A\boldsymbol{x}) \\ &= (A\boldsymbol{x}, A\boldsymbol{x}) - (\lambda\boldsymbol{x}, A\boldsymbol{x}) = (A\boldsymbol{x} - \lambda\boldsymbol{x}, A\boldsymbol{x}) = (\boldsymbol{0}, A\boldsymbol{x}) = 0 \\ (\bar{\lambda}\boldsymbol{x}, A^*\boldsymbol{x} - \bar{\lambda}\boldsymbol{x}) &= (\bar{\lambda}\boldsymbol{x}, A^*\boldsymbol{x}) - (\bar{\lambda}\boldsymbol{x}, \bar{\lambda}\boldsymbol{x}) = \bar{\lambda}(\boldsymbol{x}, A^*\boldsymbol{x}) - \bar{\lambda}\lambda(\boldsymbol{x}, \boldsymbol{x}) = (A\boldsymbol{x}, \lambda\boldsymbol{x}) - (\lambda\boldsymbol{x}, \lambda\boldsymbol{x}) \\ &= (A\boldsymbol{x} - \lambda\boldsymbol{x}, \lambda\boldsymbol{x}) = (\boldsymbol{0}, \lambda\boldsymbol{x}) = 0\end{aligned}$$

□

定義 8.5 $\boldsymbol{K}^n$ の部分空間 V, W が条件「$\boldsymbol{x} \in V$, $\boldsymbol{y} \in W$ ならば $(\boldsymbol{x}, \boldsymbol{y}) = 0$」を満たすとき, V と W は直交するという.

命題 8.6 λ, μ を正規行列 $A \in M_n(\boldsymbol{K})$ の相異なる固有値とし, V_λ, V_μ をそれぞれ λ, μ に対する A の固有空間とすれば V_λ と V_μ は直交する.

証明 $\boldsymbol{x} \in V_\lambda$, $\boldsymbol{y} \in V_\mu$ とすれば $A\boldsymbol{x} = \lambda\boldsymbol{x}$, $A\boldsymbol{y} = \mu\boldsymbol{y}$ だから補題 8.4 により $A^*\boldsymbol{y} = \bar{\mu}\boldsymbol{y}$ が成り立つ. 従って内積の性質と補題 8.3 から $\lambda(\boldsymbol{x}, \boldsymbol{y}) = (\lambda\boldsymbol{x}, \boldsymbol{y}) = (A\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, A^*\boldsymbol{y}) = (\boldsymbol{x}, \bar{\mu}\boldsymbol{y}) = \mu(\boldsymbol{x}, \boldsymbol{y})$ となるため $(\lambda - \mu)(\boldsymbol{x}, \boldsymbol{y}) = 0$ が得られる. 仮定から $\lambda \neq \mu$ であるから $(\boldsymbol{x}, \boldsymbol{y}) = 0$ である. □

命題 8.7 $A \in M_n(\boldsymbol{K})$ がユニタリー行列であるためには, A の列ベクトル $A\boldsymbol{e}_1, A\boldsymbol{e}_2, \dots, A\boldsymbol{e}_n$ が $\boldsymbol{K}^n$ の正規直交系であることが必要十分である.

証明 $A=(a_{ij})$ とすれば A^*A の (j,i) 成分は $\sum_{k=1}^{n} a_{ki}\bar{a}_{kj}=(A\boldsymbol{e}_i, A\boldsymbol{e}_j)$ であることから, $A^*A=E_n$ であるためには $A\boldsymbol{e}_1, A\boldsymbol{e}_2, \dots, A\boldsymbol{e}_n$ が $\boldsymbol{K}^n$ の正規直交系であることが必要十分である. □

補題 8.8 (1) A, B が n 次ユニタリー行列ならば AB, A^{-1} も n 次ユニタリー行列である.

(2) A が m 次ユニタリー行列, B が n 次ユニタリー行列ならば $\begin{pmatrix} A & O \\ O & B \end{pmatrix}$ は $m+n$ 次ユニタリー行列である.

(3) U を n 次ユニタリー行列, A を n 次正規行列とすれば, $U^{-1}AU$ は正規行列である.

(4) A を正方行列とする. ユニタリー行列 U で $U^{-1}AU$ が正規行列になるものが存在すれば A は正規行列である.

証明 (1) $A^*A=B^*B=E_n$ より $(AB)^*(AB)=B^*A^*AB=B^*B=E_n$ であり, $A^*=A^{-1}$ より $(A^{-1})^*=(A^*)^*=A=(A^{-1})^{-1}$ となるため AB, A^{-1} もユニタリー行列である.

(2) $A^*A=E_m$, $B^*B^{-1}=E_n$ より $\begin{pmatrix} A & O \\ O & B \end{pmatrix}^*\begin{pmatrix} A & O \\ O & B \end{pmatrix}=\begin{pmatrix} A^* & O \\ O & B^* \end{pmatrix}\begin{pmatrix} A & O \\ O & B \end{pmatrix}=\begin{pmatrix} A^*A & O \\ O & B^*B \end{pmatrix}=E_{m+n}$ となるため $\begin{pmatrix} A & O \\ O & B \end{pmatrix}$ はユニタリー行列である.

(3) $U^{-1}=U^*$ より次の等式が成り立つため $U^{-1}AU$ は正規行列である.

$$\begin{aligned}(U^{-1}AU)^*(U^{-1}AU)&=(U^*AU)^*(U^{-1}AU)=U^*A^*(U^*)^*U^{-1}AU=U^*A^*UU^{-1}AU=U^*A^*AU=U^*AA^*U\\&=U^{-1}AUU^{-1}A^*U=U^{-1}AUU^*A^*(U^*)^*=(U^{-1}AU)(U^*AU)^*=(U^{-1}AU)(U^{-1}AU)^*\end{aligned}$$

(4) $U^{-1}AU=B$ とおくと $A=(U^{-1})^{-1}BU^{-1}$ であり, (1) により U^{-1} はユニタリー行列だから (3) によって A は正規行列である. □

定理 8.9 $A\in M_n(\boldsymbol{K})$ に対し, A の固有方程式の解がすべて $\boldsymbol{K}$ に属するとき, $\boldsymbol{K}$ の要素を成分にもつユニタリー行列 U で, $U^{-1}AU$ が上半三角行列になるものが存在する.

証明 n による帰納法で主張を示す. まず $n=1$ の場合は, 主張は明らかである. 固有方程式の解がすべて $\boldsymbol{K}$ に属する任意の $n-1$ 次正方行列 B に対して $\boldsymbol{K}$ の要素を成分にもつユニタリー行列 V で, $V^{-1}BV$ が上半三角行列になるものが存在すると仮定する.

λ を A の固有値, $\boldsymbol{v}_1$ を λ に対する A の固有ベクトルとする. $\boldsymbol{K}^n$ のベクトル $\boldsymbol{v}_2, \boldsymbol{v}_3, \dots, \boldsymbol{v}_n$ で, $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ が $\boldsymbol{K}^n$ の基底になるものを選ぶ. $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ を直交化して得られる $\boldsymbol{K}^n$ の正規直交基底を $\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n$ とすれば, $\boldsymbol{w}_1=\frac{1}{\|\boldsymbol{v}_1\|}\boldsymbol{v}_1$ だから $\boldsymbol{w}_1$ も λ に対する A の固有ベクトルである. $\boldsymbol{w}_j$ を第 j 列にもつ n 次正方行列を U_1 とすれば命題 8.7 により U_1 はユニタリー行列である. (AU_1 の第 1 列) $=A\boldsymbol{w}_1=\lambda\boldsymbol{w}_1=\lambda U_1\boldsymbol{e}_1=U_1(\lambda\boldsymbol{e}_1)$ だから $U_1^{-1}AU_1$ の第 1 列は $\lambda\boldsymbol{e}_1$ に等しい. そこで $U_1^{-1}AU_1$ の $(i+1,j+1)$-成分を (i,j)-成分とする $n-1$ 次正方行列を B とすれば $U_1^{-1}AU_1=\begin{pmatrix} \lambda & * \\ \boldsymbol{0} & B \end{pmatrix}$ である. 命題 7.15 から

$$F_A(x)=F_{U_1^{-1}AU_1}(x)=\begin{vmatrix} x-\lambda & * \\ \boldsymbol{0} & xE_{n-1}-B \end{vmatrix}=(x-\lambda)|xE_{n-1}-B|=(x-\lambda)F_B(x)$$

となるため, B の固有方程式の解は A の固有方程式の解である. 従って B の固有方程式の解はすべて $\boldsymbol{K}$ に属するため, 帰納法の仮定によって $\boldsymbol{K}$ の要素を成分にもつユニタリー行列 V で, $V^{-1}BV$ が上半三角行列になるものが存在する. $U_2=\begin{pmatrix} 1 & {}^t\boldsymbol{0} \\ \boldsymbol{0} & V \end{pmatrix}$ とおくと補題 8.8 の (2) により U_2 はユニタリー行列であり, $U=U_1U_2$ とおけば補題 8.8 の (1) により U もユニタリー行列である.

$$U^{-1}AU=(U_1U_2)^{-1}A(U_1U_2)=U_2^{-1}U_1^{-1}AU_1U_2=\begin{pmatrix} 1 & {}^t\boldsymbol{0} \\ \boldsymbol{0} & V^{-1} \end{pmatrix}\begin{pmatrix} \lambda & * \\ \boldsymbol{0} & B \end{pmatrix}\begin{pmatrix} 1 & {}^t\boldsymbol{0} \\ \boldsymbol{0} & V \end{pmatrix}=\begin{pmatrix} \lambda & * \\ \boldsymbol{0} & V^{-1}BV \end{pmatrix}$$

より $U^{-1}AU$ は上半三角行列である. □

補題 8.10　上半三角行列が正規行列ならば対角行列である．

証明　$A = (a_{ij})$ を n 次上半三角行列である正規行列とすれば，$i > j$ ならば $a_{ij} = 0$ である．故に A^*A の (j,j) 成分は $\sum_{k=1}^{n} |a_{kj}|^2 = \sum_{k=1}^{j} |a_{kj}|^2$，$AA^*$ の (j,j) 成分は $\sum_{k=1}^{n} |a_{jk}|^2 = \sum_{k=j}^{n} |a_{jk}|^2$ だから $\sum_{k=1}^{j-1} |a_{kj}|^2 = \sum_{k=j+1}^{n} |a_{jk}|^2$ が $j = 1, 2, \ldots, n$ に対して成り立つ．とくに $j = n$ の場合，$\sum_{k=1}^{n-1} |a_{kn}|^2 = 0$ だから $a_{1n} = a_{2n} = \cdots = a_{n-1\,n} = 0$ である．j に関して帰納的に「$k = j+1, j+2, \ldots, n$ に対して $a_{1k} = a_{2k} = \cdots = a_{k-1\,k} = 0$」が成り立つと仮定すれば，とくに $a_{j\,j+1} = a_{j\,j+2} = \cdots = a_{j\,n} = 0$ だから $\sum_{k=1}^{j-1} |a_{kj}|^2 = \sum_{k=j+1}^{n} |a_{jk}|^2 = 0$ より $a_{1j} = a_{2j} = \cdots = a_{j-1\,j} = 0$ となるため $i < j$ ならば $a_{ij} = 0$ である．従って A は対角行列である．□

定理 8.11　$A \in M_n(\boldsymbol{K})$ が正規行列で，A の固有方程式の解がすべて $\boldsymbol{K}$ に属するならば，$\boldsymbol{K}$ の要素を成分にもつユニタリー行列 U で $U^{-1}AU$ が対角行列になるものが存在する．

証明　定理 8.9 により，$\boldsymbol{K}$ の要素を成分にもつユニタリー行列 U で，$U^{-1}AU$ が上半三角行列になるものが存在する．補題 8.8 の (3) から $U^{-1}AU$ は正規行列だから補題 8.10 によって $U^{-1}AU$ は対角行列である．□

命題 8.12　A を正方行列とする．

(1) A がエルミート行列ならば A の固有方程式の解はすべて実数である．

(2) A が歪エルミート行列ならば A の固有方程式の解はすべて純虚数である．

(3) A がユニタリー行列ならば A の固有方程式の解はすべて絶対値 1 の複素数である．

証明　まず定理 7.18 により，複素数の範囲で A の固有方程式の解が存在することに注意する．$\lambda \in \boldsymbol{C}$ を A の固有方程式の解として，A を複素を成分とする行列とみなせば，命題 7.17 により λ は A の固有値である．そこで λ に対する A の固有ベクトル $\boldsymbol{v} \in \boldsymbol{C}^n$ を 1 つ選ぶ．

(1) 仮定から $A^* = A$ で $A\boldsymbol{v} = \lambda\boldsymbol{v}$ だから $\lambda(\boldsymbol{v},\boldsymbol{v}) = (\lambda\boldsymbol{v},\boldsymbol{v}) = (A\boldsymbol{v},\boldsymbol{v}) = (\boldsymbol{v},A^*\boldsymbol{v}) = (\boldsymbol{v},A\boldsymbol{v}) = (\boldsymbol{v},\lambda\boldsymbol{v}) = \bar{\lambda}(\boldsymbol{v},\boldsymbol{v})$ となるため $(\lambda - \bar{\lambda})(\boldsymbol{v},\boldsymbol{v}) = 0$. $\boldsymbol{v} \neq \boldsymbol{0}$ より $\lambda - \bar{\lambda} = 0$, 従って λ は実数である．

(2) 仮定から $A^* = -A$ で $A\boldsymbol{v} = \lambda\boldsymbol{v}$ だから $\lambda(\boldsymbol{v},\boldsymbol{v}) = (\lambda\boldsymbol{v},\boldsymbol{v}) = (A\boldsymbol{v},\boldsymbol{v}) = (\boldsymbol{v},A^*\boldsymbol{v}) = (\boldsymbol{v},-A\boldsymbol{v}) = (\boldsymbol{v},-\lambda\boldsymbol{v}) = -\bar{\lambda}(\boldsymbol{v},\boldsymbol{v})$ となるため $(\lambda + \bar{\lambda})(\boldsymbol{v},\boldsymbol{v}) = 0$. $\boldsymbol{v} \neq \boldsymbol{0}$ より $\lambda + \bar{\lambda} = 0$, 従って λ は純虚数である．

(3) 仮定から $A^*A = E_n$ で $A\boldsymbol{v} = \lambda\boldsymbol{v}$ だから $|\lambda|^2(\boldsymbol{v},\boldsymbol{v}) = (\lambda\boldsymbol{v},\lambda\boldsymbol{v}) = (A\boldsymbol{v},A\boldsymbol{v}) = (\boldsymbol{v},A^*A\boldsymbol{v}) = (\boldsymbol{v},\boldsymbol{v})$ となるため $(|\lambda|^2 - 1)(\boldsymbol{v},\boldsymbol{v}) = 0$. $\boldsymbol{v} \neq \boldsymbol{0}$ より $|\lambda|^2 = 1$, 従って λ は絶対値 1 の複素数である．□

命題 8.13　A を正規行列とする．

(1) A の固有方程式の解がすべて実数ならば A はエルミート行列である．

(2) A の固有方程式の解がすべて純虚数ならば A は歪エルミート行列である．

(3) A の固有方程式の解がすべて絶対値 1 の複素数ならば A はユニタリー行列である．

証明　$A \in M_n(\boldsymbol{C})$ とみなせば，定理 7.18 と定理 8.11 より，複素数を成分とするユニタリー行列 U で $U^{-1}AU$ が対角行列になるものがある．そこで $D = U^{-1}AU$ とおくと D は A の固有方程式の解を対角成分にもつ対角行列で，$A = UDU^{-1} = UDU^*$ だから $A^* = (UDU^*)^* = (U^*)^*D^*U^* = UD^*U^{-1}$ である．

(1) A の固有方程式の解がすべて実数ならば D は実数を成分とする対角行列だから $D^* = \overline{D} = D$ となるため $A^* = UD^*U^{-1} = UDU^{-1} = A$ である．

(2) A の固有方程式の解がすべて純虚数ならば D は純虚数を成分とする対角行列だから $D^* = \overline{D} = -D$ となるため $A^* = UD^*U^{-1} = U(-D)U^{-1} = -A$ である．

(3) A の固有方程式の解がすべて絶対値 1 の複素数ならば $D^*D = \overline{D}D = E_n$ だから $A^*A = UD^*U^{-1}UDU^{-1} = UD^*DU^{-1} = UU^{-1} = E_n$ である．□

9 行列の級数

$d_{m,n}: M_{m,n}(\boldsymbol{K}) \times M_{m,n}(\boldsymbol{K}) \to \boldsymbol{R}$ を $d_{m,n}(A,B) = \|A-B\|$ で定義すれば, 命題 7.8 の (1), (2) と定理 7.7 の (2) から, $d_{m,n}$ は $M_{m,n}(\boldsymbol{K})$ の距離関数である. 従って, 距離空間 $(M_{m,n}(\boldsymbol{K}), d_{m,n})$ を考えることにより, 行列の列の収束や Cauchy 列の概念が定義できる.

命題 9.1 $A_l = \left(a_{ij}^{(l)}\right), B = (b_{ij}) \in M_{m,n}(\boldsymbol{K})$ とするとき, $\lim_{l\to\infty} A_l = B$ であるためには, すべての $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$ について $\lim_{l\to\infty} a_{ij}^{(l)} = b_{ij}$ が成り立つことが必要十分である.

証明 $\lim_{l\to\infty} A_l = B$ とすれば, 命題 7.8 の (4) を用いると, 任意の i, j と任意の $\varepsilon > 0$ に対して, 「$l \geqq N$ ならば $|a_{ij}^{(l)} - b_{ij}| \leqq \|A_l - B\| < \varepsilon$」が成り立つような自然数 N があるため $\lim_{l\to\infty} a_{ij}^{(l)} = b_{ij}$ である.

逆にすべての $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$ について $\lim_{l\to\infty} a_{ij}^{(l)} = b_{ij}$ とすると, 任意の $\varepsilon > 0$ に対して自然数 N_{ij} で「$l \geqq N_{ij}$ ならば $|a_{ij}^{(l)} - b_{ij}| < \dfrac{\varepsilon}{mn}$」を満たすものがあるため N を N_{ij} $(1 \leqq i \leqq m, 1 \leqq j \leqq n)$ のうちで最大のものとすると命題 7.8 の (4) から $l \geqq N$ ならば $\|A_l - B\| < \varepsilon$ が成り立つ. 従って $\lim_{l\to\infty} A_l = B$ である. □

注意 9.2 $z = x + yi \in \boldsymbol{C}$ $(x, y \in \boldsymbol{R})$ ならば $|x|, |y| \leqq |z| \leqq |x| + |y|$ だから, $z_l = x_l + y_l i$, $w = u + vi$ $(x_l, y_l, u, v \in \boldsymbol{R})$ のとき $\lim_{l\to\infty} z_l = w$ であるためには $\lim_{l\to\infty} x_l = u$ かつ $\lim_{l\to\infty} y_l = v$ であることが必要十分である.

定理 9.3 距離空間 $(M_{m,n}(\boldsymbol{K}), d_{m,n})$ は完備である.

証明 複素数列 $(z_l)_{l\geqq 0}$ が Cauchy 列ならば $z_l = x_l + y_l i$ $(x_l, y_l \in \boldsymbol{R})$ とおくと, $|x_l - x_k|, |y_l - y_k| \leqq |z_l - z_k|$ だから実数列 $(x_l)_{l\geqq 0}$, $(y_l)_{l\geqq 0}$ はともに Cauchy 列になるため, 注意 9.2 と定理 1.10 から $(z_l)_{l\geqq 0}$ は収束する. 故に $\boldsymbol{C}$ は完備である. $(A_l)_{l\geqq 0}$ が $(M_{m,n}(\boldsymbol{K}), d_{m,n})$ の Cauchy 列ならば命題 7.8 の (4) から, A_l の各成分からなる数列は Cauchy 列であり, $\boldsymbol{R}$ と $\boldsymbol{C}$ は完備だからこれらは収束する. 従って, 命題 9.1 から $(A_l)_{l\geqq 0}$ は収束する. □

命題 9.4 (1) $(A_l)_{l\geqq 0}, (B_l)_{l\geqq 0}$ がともに収束する $m \times n$ 行列の列で, $\lambda, \mu \in \boldsymbol{K}$ とすると, 次の等式が成り立つ.

$$\lim_{l\to\infty}(\lambda A_l + \mu B_l) = \lambda \lim_{l\to\infty} A_l + \mu \lim_{l\to\infty} B_l, \quad \lim_{l\to\infty} {}^t\!A_l = {}^t(\lim_{l\to\infty} A_l), \quad \lim_{l\to\infty} \overline{A_l} = \overline{\lim_{l\to\infty} A_l}, \quad \lim_{l\to\infty} A_l^* = (\lim_{l\to\infty} A_l)^*$$

(2) $m \times n$ 行列の列 $(A_l)_{l\geqq 0}$ と $n \times k$ 行列の列 $(B_l)_{l\geqq 0}$ がともに収束するならば $\lim_{l\to\infty} A_l B_l = \Big(\lim_{l\to\infty} A_l\Big)\Big(\lim_{l\to\infty} B_l\Big)$.

証明 $\lim_{l\to\infty} A_l = A$, $\lim_{l\to\infty} B_l = B$ とおき, ε を任意の正の実数とする.

(1) 自然数 N_1, N_2 で,「$n > N_1$ ならば $\|A_l - A\| < \dfrac{\varepsilon}{1+2|\lambda|}$」,「$n > N_2$ ならば $\|B_l - B\| < \dfrac{\varepsilon}{1+2|\mu|}$」を満たすものがある. $l \geqq \max\{N_1, N_2\}$ ならば $\|A_l - A\| < \dfrac{\varepsilon}{1+2|\lambda|}$ かつ $\|B_l - B\| < \dfrac{\varepsilon}{1+2|\mu|}$ だから定理 7.7 の (2) より $\|(\lambda A_l + \mu B_l) - (\lambda A + \mu B)\| = \|\lambda(A_l - A) + \mu(B_l - B)\| \leqq |\lambda|\|A_l - A\| + |\mu|\|B_l - B\| < \dfrac{\varepsilon|\lambda|}{1+2|\lambda|} + \dfrac{\varepsilon|\mu|}{1+2|\mu|} < \varepsilon$ が成り立つため, $\lim_{l\to\infty}(\lambda A_l + \mu B_l) = \lambda A + \mu B$ である. $\|{}^t\!A_l - {}^t\!A\| = \|\overline{A_l} - \overline{A}\| = \|A_l^* - A^*\| = \|A_l - A\|$ だから, 残りの 3 つの等式は明らかである.

(2) 自然数 N_3, N_4 で,「$l \geqq N_3$ ならば $\|A_l - A\| < \dfrac{\varepsilon}{2(1+\|B\|)+\varepsilon}$」,「$l \geqq N_4$ ならば $\|B_l - B\| < \dfrac{\varepsilon}{2(1+\|A\|)}$」を満たすものがある. $\dfrac{\varepsilon}{2(1+\|B\|)+\varepsilon} < \min\left\{1, \dfrac{\varepsilon}{2(1+\|B\|)}\right\}$ であることに注意すると $l \geqq \max\{N_3, N_4\}$ ならば

$$\begin{aligned}
\|A_l\| &= \|A_l - A + A\| \leqq \|A_l - A\| + \|A\| < 1 + \|A\| \\
\|A_l B_l - AB\| &= \|A_l(B_l - B) + (A_l - A)B\| \leqq \|A_l\|\|B_l - B\| + \|A_l - A\|\|B\| \\
&\leqq (1 + \|A\|)\frac{\varepsilon}{2(1+\|A\|)} + \frac{\varepsilon}{2(1+\|B\|)+\varepsilon}\|B\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2(1+\|B\|)} < \varepsilon
\end{aligned}$$

が成り立つため, $\lim_{l\to\infty} A_l B_l = AB$ である. □

定義 9.5　$m \times n$ 行列の列 $(A_l)_{l \geqq 0}$ に対し，$S_l = \sum_{k=0}^{l} A_k$ とおく．$\lim_{l \to \infty} S_l = S$ が存在すれば，$m \times n$ 行列の級数 $\sum_{l=0}^{\infty} A_l = A_0 + A_1 + \cdots + A_l + \cdots$ は S に収束するといい，S も $\sum_{l=0}^{\infty} A_l$ で表す．

命題 9.6　$m \times n$ 行列の列 $(A_l)_{l \geqq 0}$ に対し，$s_l = \sum_{k=0}^{l} \|A_k\|$ とおく．数列 $(s_l)_{l \geqq 0}$ が上に有界ならば $\sum_{l=0}^{\infty} A_l$ は収束する．

証明　$(s_l)_{l \geqq 0}$ は単調増加数列で上に有界だから収束するため，命題 1.7 により，この数列は Cauchy 列である．$S_l = \sum_{k=0}^{l} A_k$ とおくと，$l \leqq r$ ならば定理 7.7 の (2) から $\|S_r - S_l\| = \left\| \sum_{k=l+1}^{r} A_k \right\| \leqq \sum_{k=l+1}^{r} \|A_k\| = s_r - s_l$ となるため $(S_l)_{l \geqq 0}$ は $m \times n$ 行列の Cauchy 列である．従って命題 9.3 により $(S_l)_{l \geqq 0}$ は収束する．□

定義 9.7　$m \times n$ 行列の列 $(A_l)_{l \geqq 0}$ が命題 9.6 の仮定を満たすとき，行列の級数 $\sum_{l=0}^{\infty} A_l$ は絶対収束するという．

命題 9.6 において $(s_l)_{l \geqq 0}$ は単調増加数列だから，命題 1.7, 命題 1.8, 定理 1.10 および上に有界な単調増加数列が収束することから数列 $(s_l)_{l \geqq 0}$ が「上に有界であること」,「Cauchy 列であること」,「収束すること」はすべて同値であることに注意する．

命題 9.8　$(A_l)_{l \geqq 0}$, $(B_l)_{l \geqq 0}$ を $m \times n$ 行列の列とし，行列の級数 $\sum_{l=0}^{\infty} A_l$, $\sum_{l=0}^{\infty} B_l$ はともに収束すると仮定する．このとき $\lambda, \mu \in \boldsymbol{K}$ に対し，行列の級数 $\sum_{l=0}^{\infty} (\lambda A_l + \mu B_l)$ は $\lambda \sum_{l=0}^{\infty} A_l + \mu \sum_{l=0}^{\infty} B_l$ に収束する．また，$\sum_{l=0}^{\infty} A_l$, $\sum_{l=0}^{\infty} B_l$ が絶対収束すれば $\sum_{l=0}^{\infty} (\lambda A_l + \mu B_l)$ は絶対収束する．

証明　$S_i = \sum_{l=0}^{i} A_l$, $T_i = \sum_{l=0}^{i} B_l$ とおくと $\lambda S_i + \mu T_i = \sum_{l=0}^{i} (\lambda A_l + \mu B_l)$ であり，命題 9.4 により次の等式が成り立つ．

$$\sum_{l=0}^{\infty} (\lambda A_l + \mu B_l) = \lim_{i \to \infty} (\lambda S_i + \mu T_i) = \lambda \lim_{i \to \infty} S_i + \mu \lim_{i \to \infty} T_i = \lambda \sum_{l=0}^{\infty} A_l + \mu \sum_{l=0}^{\infty} B_l$$

$\sum_{l=0}^{\infty} A_l$, $\sum_{l=0}^{\infty} B_l$ が絶対収束すれば，定数 α, β で，すべての i に対して $\sum_{l=0}^{i} \|A_l\| \leqq \alpha$, $\sum_{l=0}^{i} \|B_l\| \leqq \beta$ を満たすものがある．故に任意の i に対して $\sum_{l=0}^{i} \|\lambda A_l + \mu B_l\| \leqq \sum_{l=0}^{i} |\lambda| \|A_l\| + \sum_{l=0}^{i} |\mu| \|B_l\| = |\lambda| \sum_{l=0}^{i} \|A_l\| + |\mu| \sum_{l=0}^{i} \|B_l\| \leqq |\lambda| \alpha + |\mu| \beta$ となるため，後半の主張も成り立つ．□

命題 9.9　$(A_l)_{l \geqq 0}$, $(B_l)_{l \geqq 0}$ をそれぞれ $m \times n$, $n \times p$ 行列の列とし，$C_l = \sum_{k=0}^{l} A_k B_{l-k}$ とおく．行列の級数 $\sum_{l=0}^{\infty} A_l$, $\sum_{l=0}^{\infty} B_l$ がともに絶対収束するならば $\sum_{l=0}^{\infty} C_l$ は $\left(\sum_{l=0}^{\infty} A_l \right) \left(\sum_{l=0}^{\infty} B_l \right)$ に絶対収束する．とくに $A_0 = A$, $A_l = O$ $(l \geqq 1)$ と $B_0 = B$, $B_l = O$ $(l \geqq 1)$ の場合を考えれば $A \left(\sum_{l=0}^{\infty} B_l \right) = \sum_{l=0}^{\infty} A B_l$, $\left(\sum_{l=0}^{\infty} A_l \right) B = \sum_{l=0}^{\infty} A_l B$ が成り立つ．

証明　$a_i = \sum_{l=0}^{i} \|A_l\|$, $b_i = \sum_{l=0}^{i} \|B_l\|$, $c_i = \sum_{l=0}^{i} \|C_l\|$ とおくと，仮定から定数 α, β で，すべての i に対して $a_i \leqq \alpha$, $b_i \leqq \beta$ を満たすものがとれる．このとき定理 7.7 の (2) と命題 7.9 から

$$c_i \leqq \sum_{l=0}^{i} \sum_{k=0}^{l} \|A_k B_{l-k}\| \leqq \sum_{l=0}^{i} \sum_{k=0}^{l} \|A_k\| \cdot \|B_{l-k}\| \leqq \sum_{l=0}^{i} \sum_{k=0}^{i} \|A_k\| \cdot \|B_l\| = a_i b_i \leqq \alpha\beta$$

となるため $(c_l)_{l \geqq 0}$ は上に有界である．故に，命題 9.6 から行列の級数 $\sum_{l=0}^{\infty} C_l$ は絶対収束する．

とくに $d_l = \sum_{k=0}^{l} \|A_k\| \cdot \|B_{l-k}\|$, $e_i = \sum_{l=0}^{i} d_l$ とおいて上の議論の $m = n = p = 1$ の場合を適用すれば，単調増加数

列 $(e_l)_{l\geqq 0}$ は上に有界であるため Cauchy 列であることに注意する．

$S=\sum\limits_{l=0}^{\infty}A_l$, $T=\sum\limits_{l=0}^{\infty}B_l$, $S_i=\sum\limits_{l=0}^{i}A_l$, $T_i=\sum\limits_{l=0}^{i}B_l$, $R_i=\sum\limits_{l=0}^{i}C_l$ とおくと，定理 7.7 の (2) と命題 7.9 から不等式

$$\|S_iT_i-R_i\|=\left\|\sum_{u,v\leqq i,u+v>i}A_uB_v\right\|\leqq\sum_{u,v\leqq i,u+v>i}\|A_uB_v\|\leqq\sum_{u,v\leqq i,u+v>i}\|A_u\|\cdot\|B_v\|\leqq\sum_{l=i+1}^{2i}d_l=e_{2i}-e_i$$

を得る．この不等式と $ST-R_i=(S-S_i)T+S_i(T-T_i)+S_iT_i-R_i$ に注意すれば，定理 7.7 の (2) と命題 7.9 から

$$\|ST-R_i\|\leqq\|(S-S_i)T\|+\|S_i(T-T_i)\|+\|S_iT_i-R_i\|\leqq\|S-S_i\|\cdot\|T\|+\|S_i\|\cdot\|T-T_i\|+e_{2i}-e_i$$

が得られる． $i\to\infty$ のとき，$\|S-S_i\|$, $\|T-T_i\|\to 0$, $\|S_i\|\to\|S\|$ であり，$(e_l)_{l\geqq 0}$ は Cauchy 列だから $\lim\limits_{i\to\infty}(e_{2i}-e_i)=0$ である．従って上の不等式から $\lim\limits_{i\to\infty}R_i=ST$ がわかる． □

命題 9.10 x の整級数 $f(x)=\sum\limits_{l=0}^{\infty}a_lx^l=a_0+a_1x+\cdots+a_lx^l+\cdots$ の収束半径が R ならば $\|A\|<R$ を満たす $A\in M_n(\boldsymbol{C})$ に対し，行列の級数 $f(A)=\sum\limits_{l=0}^{\infty}a_lA^l=a_0E_n+a_1A+\cdots+a_lA^l+\cdots$ は絶対収束する．

証明 $|x|<R$ ならば $f(x)$ は絶対収束するため，$\bar f(x)=\sum\limits_{l=0}^{\infty}|a_l|x^l$ で整級数 $\bar f(x)$ を定めれば $\bar f(x)$ も $|x|<R$ の範囲で絶対収束する．$s_l=\sum\limits_{k=0}^{l}\left\|a_kA^k\right\|$ とおけば

$$s_l\leqq|a_0|\|E_n\|+\sum_{k=1}^{l}|a_k|\|A\|^k=\sqrt{n}|a_0|-|a_0|+\sum_{k=0}^{l}|a_k|\|A\|^k\leqq(\sqrt{n}-1)|a_0|+\bar f(\|A\|)$$

だから $(s_l)_{l\geqq 0}$ は上に有界である．故に命題 9.6 から $f(A)=\sum\limits_{l=0}^{\infty}a_lA^l$ は絶対収束する． □

命題 9.11 $\boldsymbol{F}$, $\boldsymbol{K}$ は $\boldsymbol{R}$, $\boldsymbol{C}$ のいずれかで，$\boldsymbol{K}=\boldsymbol{R}$ の場合は $\boldsymbol{F}=\boldsymbol{R}$ であるとする．x の整級数 $f(x)=\sum\limits_{l=0}^{\infty}a_lx^l$ の収束半径が R で，すべての $l=0,1,2,\ldots$ に対して $a_l\in\boldsymbol{F}$ とする．また，写像 $\Phi:M_n(\boldsymbol{K})\to M_n(\boldsymbol{K})$ は連続で，任意の $r,s\in\boldsymbol{F}$ と $A,B\in M_n(\boldsymbol{K})$ に対して $\Phi(rA+sB)=r\Phi(A)+s\Phi(B)$ を満たし，$\Phi(A^l)=\Phi(A)^l$ が任意の $A\in M_n(\boldsymbol{K})$ と 0 以上の整数 k に対して成り立つとする．このとき $\|A\|<R$ を満たす $A\in M_n(\boldsymbol{K})$ に対して $f(\Phi(A))=\Phi(f(A))$ が成り立つ．

証明 $S_k=\sum\limits_{l=0}^{k}a_lA^l$ とおけば，仮定から $\Phi(S_k)=\sum\limits_{l=0}^{k}\Phi(a_lA^l)=\sum\limits_{l=0}^{k}a_l\Phi(A^l)=\sum\limits_{l=0}^{k}a_l\Phi(A)^l$ だから，Φ の連続性より $f(\Phi(A))=\sum\limits_{l=0}^{\infty}a_l\Phi(A)^l=\lim\limits_{k\to\infty}\sum\limits_{l=0}^{k}a_l\Phi(A)^l=\lim\limits_{k\to\infty}\Phi(S_k)=\Phi\Big(\lim\limits_{k\to\infty}S_k\Big)=\Phi\Big(\sum\limits_{l=0}^{\infty}a_lx^l\Big)=\Phi(f(A))$ が得られる． □

注意 9.12 P を n 次正則行列とし，$\Phi_P,\Phi_t,\Phi_c,\Phi_a:M_n(\boldsymbol{C})\to M_n(\boldsymbol{C})$ をそれぞれ $\Phi_P(A)=P^{-1}AP$, $\Phi_t(A)={}^tA$, $\Phi_c(A)=\overline{A}$, $\Phi_a(A)=A^*$ で定める．行列の積の線形性と結合法則および命題 7.3, 命題 9.4 の (1) から Φ_P と Φ_t は $\boldsymbol{F}=\boldsymbol{K}=\boldsymbol{C}$ に対して上の命題の仮定を満たし，Φ_c と Φ_a は $\boldsymbol{F}=\boldsymbol{R}$, $\boldsymbol{K}=\boldsymbol{C}$ に対して上の命題の仮定を満たすため，収束半径が R の x の整級数 $f(x)$ に対して $\|A\|<R$ ならば $f(P^{-1}AP)=P^{-1}f(A)P$, $f({}^tA)={}^tf(A)$ が成り立ち，さらに $f(x)$ の係数がすべて実数ならば $f(\overline{A})=\overline{f(A)}$, $f(A^*)=f(A)^*$ が成り立つ．

命題 9.13 $A\in M_n(\boldsymbol{C})$ の固有多項式が $F_A(x)=(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$ と因数分解し，整級数 $f(x)=\sum\limits_{l=0}^{\infty}a_lx^l$ の収束半径を R とする．$\|A\|<R$ ならば $F_{f(A)}(x)=(x-f(\lambda_1))(x-f(\lambda_2))\cdots(x-f(\lambda_n))$ である．

証明 定理 8.9 からユニタリー行列 P で，$P^{-1}AP$ が上半三角行列になるものがある．$P^{-1}AP$ の (i,i) 成分を λ_i とすれば $(P^{-1}AP)^k$ の (i,i) 成分は λ_i^k だから注意 9.12 から $P^{-1}f(A)P=f(P^{-1}AP)$ は (i,i) 成分が $f(\lambda_i)$ であるような上半三角行列である．故に命題 7.19 から結果が得られる． □

10 行列の指数写像

整級数 $f(x)$ が $x=0$ における e^x の Taylor 展開 $\sum_{k=0}^{\infty}\frac{x^k}{k!}=1+x+\frac{x^2}{2!}+\cdots+\frac{x^k}{k!}+\cdots$ の場合に命題 9.10 を用いれば, 行列の指数写像が次のように定義できる.

定義 10.1 写像 $\exp: M_n(\boldsymbol{C})\to M_n(\boldsymbol{C})$ を $\exp A=\sum_{k=0}^{\infty}\frac{1}{k!}A^k=E_n+A+\frac{A^2}{2!}+\cdots+\frac{A^k}{k!}+\cdots$ で定義する. この写像を行列の指数写像という. とくに $n=1$ の場合, $z\in\boldsymbol{C}$ に対して級数 $\sum_{k=0}^{\infty}\frac{z^k}{k!}$ は絶対収束するため, $e^z=\sum_{k=0}^{\infty}\frac{z^k}{k!}$ とおき, $z\in\boldsymbol{C}$ を e^z に対応させる $\boldsymbol{C}$ から $\boldsymbol{C}$ への関数を指数関数という.

注意 10.2 (1) 写像 $\exp: M_n(\boldsymbol{C})\to M_n(\boldsymbol{C})$ の定義から $\exp O=E_n$ である.

(2) 実数 x に対し, 数列 $(a_k(x))_{k\geqq 0}$, $(b_k(x))_{k\geqq 0}$ を $a_k(x)=\frac{(xi)^k}{k!}$ (k は偶数), $a_k(x)=0$ (k は奇数), $b_k(x)=\frac{(xi)^k}{k!}$ (k は奇数), $b_k(x)=0$ (k は偶数) で定めると, $\cos x=\sum_{l=0}^{\infty}(-1)^l\frac{x^{2l}}{(2l)!}=\sum_{k=0}^{\infty}a_k(x)$, $\sin x=\sum_{l=0}^{\infty}(-1)^l\frac{x^{2l+1}}{(2l+1)!}=\frac{1}{i}\sum_{k=0}^{\infty}b_k(x)$ であり, すべての $k=0,1,\ldots$ に対して $a_k(x)+b_k(x)=\frac{(xi)^k}{k!}$ が成り立つことに注意すれば命題 9.8 から $e^{xi}=\sum_{k=0}^{\infty}\frac{(xi)^k}{k!}=\sum_{k=0}^{\infty}(a_k(x)+b_k(x))=\sum_{k=0}^{\infty}a_k(x)+\sum_{k=0}^{\infty}b_k(x)=\cos x+i\sin x$ が得られる. このことから, e^{xi} は絶対値が 1 の複素数であり, 逆に絶対値が 1 の複素数は e^{xi} ($x\in\boldsymbol{R}$) の形に表せることがわかる.

命題 10.3 指数写像 $\exp: M_n(\boldsymbol{C})\to M_n(\boldsymbol{C})$ は次の性質をもつ.

(1) $AB=BA$ ならば $\exp(A+B)=(\exp A)(\exp B)$.

(2) $\exp(-A)=(\exp A)^{-1}$. 従って, exp は常に正則行列を値にとる.

(3) $\det\exp A=e^{\operatorname{tr}A}$

証明 (1) $A_l=\frac{A^l}{l!}$, $B_l=\frac{B^l}{l!}$ とおけば $\exp A=\sum_{l=0}^{\infty}A_l$, $\exp B=\sum_{l=0}^{\infty}B_l$ である. $C_l=\sum_{k=0}^{l}A_kB_{l-k}$ とおけば, $AB=BA$ より二項定理から $C_l=\frac{1}{l!}\sum_{k=0}^{l}\frac{l!}{k!(l-k)!}A^kB^{l-k}=\frac{1}{l!}(A+B)^l$ だから $\exp(A+B)=\sum_{l=0}^{\infty}C_l$ である. 故に命題 9.9 から $\exp(A+B)=(\exp A)(\exp B)$ が得られる.

(2) $B=-A$ を (1) の式に代入すれば $(\exp A)(\exp(-A))=\exp O=E_n$ だから $\exp(-A)=(\exp A)^{-1}$ である.

(3) $F_A(x)=(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$ ならば命題 9.13 から $F_{\exp A}(x)=(x-e^{\lambda_1})(x-e^{\lambda_2})\cdots(x-e^{\lambda_n})$ である. とくに $n=1$ の場合, (1) により任意の $z,w\in\boldsymbol{C}$ に対して「指数法則」$e^{z+w}=e^ze^w$ が成り立つことから, 命題 7.20 より $\det\exp A=e^{\lambda_1}e^{\lambda_2}\cdots e^{\lambda_n}=e^{\lambda_1+\lambda_2+\cdots+\lambda_n}=e^{\operatorname{tr}A}$ が得られる. □

例 10.4 $x\in\boldsymbol{R}$ に対して指数写像の定義と e^x のテイラー展開から $\exp\begin{pmatrix}x&0\\0&x\end{pmatrix}=\begin{pmatrix}e^x&0\\0&e^x\end{pmatrix}$ が成り立つ. $y\in\boldsymbol{R}$ に対し, $\begin{pmatrix}0&-y\\y&0\end{pmatrix}^2=-y^2E_2$ だから $\begin{pmatrix}0&-y\\y&0\end{pmatrix}^{2k}=(-1)^ky^{2k}E_2$ および $\begin{pmatrix}0&-y\\y&0\end{pmatrix}^{2k+1}=(-1)^ky^{2k+1}\begin{pmatrix}0&-1\\1&0\end{pmatrix}$ が成り立つことと, $\cos y$, $\sin y$ のテイラー展開から, $\exp\begin{pmatrix}0&-y\\y&0\end{pmatrix}=\begin{pmatrix}\cos y&-\sin y\\\sin y&\cos y\end{pmatrix}$ が成り立つ. $\begin{pmatrix}x&0\\0&x\end{pmatrix}$ と $\begin{pmatrix}0&-y\\y&0\end{pmatrix}$ の積は交換可能だから, 命題 10.3 の (1) から次の等式が得られる.

$$\exp\begin{pmatrix}x&-y\\y&x\end{pmatrix}=\exp\left(\begin{pmatrix}x&0\\0&x\end{pmatrix}+\begin{pmatrix}0&-y\\y&0\end{pmatrix}\right)=\exp\begin{pmatrix}x&0\\0&x\end{pmatrix}\exp\begin{pmatrix}0&-y\\y&0\end{pmatrix}=\begin{pmatrix}e^x\cos y&-e^x\sin y\\e^x\sin y&e^x\cos y\end{pmatrix}$$

注意 10.5 $\boldsymbol{K}$ の要素を成分をもつ n 次正則行列全体からなる集合を $GL_n(\boldsymbol{K})$ で表す．命題 10.3 の (2) から，任意の $A \in M_n(\boldsymbol{C})$ に対して $\exp A \in GL_n(\boldsymbol{C})$ だから，以後指数写像を $M_n(\boldsymbol{C})$ から $GL_n(\boldsymbol{C})$ への写像とみなす．また，$A \in M_n(\boldsymbol{R})$ ならば，$\exp A$ の各成分は実数だから $\exp A \in GL_n(\boldsymbol{R})$ である．

命題 10.6 次の等式が成り立つ．従って，exp は O で微分可能であり，$\exp' O$ は $M_n(\boldsymbol{C})$ の恒等写像である．

$$\lim_{X\to O} \frac{\exp X - (\exp O + id_{M_n(\boldsymbol{C})}(X-O))}{\|X-O\|} = O$$

証明 指数写像の定義から $\exp X - (\exp O + id_{M_n(\boldsymbol{C})}(X-O)) = \exp X - E_n - X = \sum_{k=2}^{\infty} \frac{1}{k!}X^k$ が成り立つ．$\|X\| \leqq 1$ ならば，2 以上の自然数 N に対し，定理 7.7 の (2) と命題 7.9 から次の不等式が得られる．

$$\left\|\sum_{k=2}^{N} \frac{1}{k!}X^k\right\| \leqq \|X\|^2 \sum_{k=2}^{N} \frac{1}{k!}\|X\|^{k-2} \leqq \|X\|^2 \sum_{k=2}^{N} \frac{1}{k!} = \|X\|^2 \left(\sum_{k=0}^{\infty} \frac{1}{k!} - 2\right) = \|X\|^2(e-2)$$

故に $\left\|\sum_{k=2}^{\infty} \frac{1}{k!}X^k\right\| \leqq \|X\|^2(e-2)$ だから，最初の等式から $\left\|\frac{\exp X - (\exp O + id_{M_n(\boldsymbol{C})}(X-O))}{\|X-O\|}\right\| \leqq \|X\|(e-2)$ が得られる．従って $X \to O$ のとき，$\frac{\exp X - (\exp O + id_{M_n(\boldsymbol{C})}(X-O))}{\|X-O\|}$ は O に近づく． □

$A \in M_n(\boldsymbol{R})$, $\boldsymbol{v} \in \boldsymbol{R}^n$ に対し，写像 $f : \boldsymbol{R} \to \boldsymbol{R}^n$ を $f(t) = \exp(tA)\boldsymbol{v}$ で定める．

系 10.7 各 $t \in \boldsymbol{R}$ に対して f は t で微分可能で，$f'(t) = A\exp(tA)\boldsymbol{v} = Af(t)$ が成り立つ．

証明 $A = O$ の場合は主張は明らかだから，$A \neq O$ と仮定する．$h \to 0$ のとき，$hA \to O$ だから，命題 10.6 より $\lim_{h\to 0}\left\|\frac{\exp(hA) - (E_n + hA)}{\|hA\|}\right\| = 0$ が成り立つ．この左辺は $\frac{1}{\|A\|}\lim_{h\to 0}\left\|\frac{1}{h}(\exp(tA) - E_n) - A\right\|$ に等しいため，$\lim_{h\to 0}\left(\frac{1}{h}(\exp(hA) - E_n) - A\right) = O$，すなわち $\lim_{h\to 0}\frac{1}{h}(\exp(hA) - E_n) = A$ である．従って，f の定義と定理 10.3 の (1) から次の等式が成り立つため主張が示される．

$$\begin{aligned}\lim_{h\to 0}\frac{1}{h}(f(t+h) - f(t)) &= \lim_{h\to 0}\frac{1}{h}(\exp((t+h)A)\boldsymbol{v} - \exp(tA)\boldsymbol{v}) = \lim_{h\to 0}\frac{1}{h}(\exp(hA+tA)\boldsymbol{v} - \exp(tA)\boldsymbol{v}) \\ &= \lim_{h\to 0}\frac{1}{h}(\exp(hA)\exp(tA)\boldsymbol{v} - \exp(tA)\boldsymbol{v}) = \lim_{h\to 0}\frac{1}{h}(\exp(hA) - E_n)\exp(tA)\boldsymbol{v} \\ &= A\exp(tA)\boldsymbol{v}\end{aligned}$$

□

注意 10.8 上の結果は $f(t) = \exp(tA)\boldsymbol{v}$ で定義される写像 $f : \boldsymbol{R} \to \boldsymbol{R}^n$ は A を係数行列とする定数係数斉次線形連立微分方程式 $\boldsymbol{x}' = A\boldsymbol{x}$ の $f(0) = \boldsymbol{v}$ を満たす解であることを示している．

記号 10.9 また n 次歪エルミート行列全体のなす $M_n(\boldsymbol{C})$ の部分空間を $SH(n)$ で表し，n 次ユニタリー行列全体のなす $M_n(\boldsymbol{C})$ の部分空間を $U(n)$ で表す．

補題 10.10 $U(n)$ はコンパクトである．

証明 写像 $f : M_n(\boldsymbol{C}) \to M_n(\boldsymbol{C})$ を $f(A) = A^*A$ で定めれば，f は連続写像であり，$U(n)$ は f による閉集合 $\{E_n\}$ の逆像だから，$U(n)$ は $M_n(\boldsymbol{C})$ の閉集合である．また $U \in U(n)$ ならば命題 8.7 より U の各列ベクトルの長さは 1 だから $\|U\| = \sqrt{n}$ である．故に $U(n)$ は $M_n(\boldsymbol{C})$ の有界な閉集合である．$M_n(\boldsymbol{C})$ は距離空間として $\boldsymbol{C}^{n^2}$ と同一視され，さらに $\boldsymbol{C}^{n^2}$ は距離空間として $\boldsymbol{R}^{2n^2}$ と同一視されるため，$U(n)$ は $\boldsymbol{R}^{2n^2}$ の有界な閉集合とみなせる． □

$SH(1)$ は純虚数全体の集合，$U(1)$ は絶対値 1 の複素数全体の集合とみなせるため，$SH(n)$, $U(n)$ はそれぞれこれらの一般化と考えれば，次の命題 10.11 は複素数を変数とする指数関数が純虚数全体の集合を絶対値が

1 の複素数全体の集合の上に写すという事実を一般化している．さらに $SH_0(n) = \{A \in SH(n) | \operatorname{tr} A = 0\}$, $SU(n) = \{A \in U(n) | \det A = 1\}$ とおく．

命題 10.11 exp による $SH(n)$ の像は $U(n)$ であり, $SH_0(n)$ の像は $SU(n)$ である．

証明 $A \in SH(n)$ とすると, $A^* = -A$ だから注意 9.12 と命題 10.3 の (2) から $(\exp A)^* = \exp A^* = \exp(-A) = (\exp A)^{-1}$ となるため $\exp A \in U(n)$ である．さらに $\operatorname{tr} A = 0$ ならば命題 10.3 の (3) から $\det \exp A = e^0 = 1$ だから $\exp A \in SU(n)$ である．従って exp は $SH(n)$ を $U(n)$ に写し, $SH_0(n)$ を $SU(n)$ に写す．任意の $U \in U(n)$ に対して, $P^{-1}UP$ が対角行列になるようなユニタリー行列 P をとると補題 8.8 の (1) から $P^{-1}UP$ もユニタリー行列だから各対角成分は絶対値が 1 である複素数である．故に注意 10.2 から $P^{-1}UP$ の (j, j) 成分を $e^{x_j i}$ $(x_j \in \boldsymbol{R})$ とおくことができる．B を (j, j) 成分が $x_j i$ であるような対角行列とすれば, $B \in SH(n)$ であり, $\exp B = P^{-1}UP$ となる．従って, 注意 9.12 から $\exp(PBP^{-1}) = U$ が得られる．$P^{-1} = P^*$ と $B^* = -B$ から $PBP^{-1} \in SH(n)$ がわかるため, exp による $SH(n)$ の像は $U(n)$ である．また, $U \in SU(n)$ ならば対角行列 $P^{-1}UP$ の行列式は 1 だから, 対角成分の積は 1 である．従って, 上記の実数 $x_1, x_2, \ldots, x_n$ は $x_n = -x_1 - x_2 - \cdots - x_{n-1}$ が成り立つように選ぶことができる．このとき, $\operatorname{tr} PBP^{-1} = \operatorname{tr} B = i(x_1 + x_2 + \cdots + x_n) = 0$ だから $PBP^{-1} \in SH_0(n)$ である． □

11 行列の対数写像

$\boldsymbol{C}^n$ の標準的な内積を考える．n 次正方行列 A と $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{C}^n$ に対して $y^* A\boldsymbol{x} = {}^t(\boldsymbol{y}^* A\boldsymbol{x}) = {}^t(A\boldsymbol{x}){}^t(\boldsymbol{y}^*) = {}^t(A\boldsymbol{x})\bar{\boldsymbol{y}} = (A\boldsymbol{x}, \boldsymbol{y})$ が成り立つため, A がエルミート行列ならば $\overline{\boldsymbol{x}^* A\boldsymbol{x}} = \overline{(A\boldsymbol{x}, \boldsymbol{x})} = (\boldsymbol{x}, A\boldsymbol{x}) = (\boldsymbol{x}, A^*\boldsymbol{x}) = (A\boldsymbol{x}, \boldsymbol{x}) = \boldsymbol{x}^* A\boldsymbol{x}$ だから $\boldsymbol{x}^* A\boldsymbol{x}$ は実数である．

定義 11.1 n 次エルミート行列 A が条件「$\boldsymbol{x} \in \boldsymbol{C}^n$ かつ $\boldsymbol{x} \neq \boldsymbol{0}$ ならば $\boldsymbol{x}^* A\boldsymbol{x} > 0$」を満たすとき, A を正値エルミート行列という．また, n 次実対称行列 A が条件「$\boldsymbol{x} \in \boldsymbol{R}^n$ かつ $\boldsymbol{x} \neq \boldsymbol{0}$ ならば ${}^t\boldsymbol{x} A\boldsymbol{x} > 0$」を満たすとき, A を正値対称行列という．

命題 11.2 エルミート行列 A が正値エルミート行列であるためには, A の固有値がすべて正の実数であることが必要十分である．とくに正値エルミート行列のトレースと行列式の値は正の実数である．

証明 n 次エルミート行列 A を対角化するユニタリー行列 U を 1 つ選び, $U^{-1}AU$ の (j, j) 成分を λ_j とおく．$\boldsymbol{x} \in \boldsymbol{C}^n$ に対し, $\boldsymbol{y} = U^{-1}\boldsymbol{x}$ とおき, $\boldsymbol{y}$ の第 j 成分を y_j とすれば, $\boldsymbol{x} = U\boldsymbol{y}$ より次の等式が成り立つ．

$$\boldsymbol{x}^* A\boldsymbol{x} = (U\boldsymbol{y})^* AU\boldsymbol{y} = \boldsymbol{y}^* U^* AU\boldsymbol{y} = \boldsymbol{y}^* U^{-1} AU\boldsymbol{y} = \lambda_1 |y_1|^2 + \lambda_2 |y_2|^2 + \cdots + \lambda_n |y_n|^2$$

A の固有値がすべて正の実数で, $\boldsymbol{x} \neq \boldsymbol{0}$ ならば $\boldsymbol{y} \neq \boldsymbol{0}$ だから上式の左辺は正になるため, A は正値エルミート行列である．逆に A が正値エルミート行列ならば, $\boldsymbol{x} = U\boldsymbol{e}_j$ の場合, U は正則行列だから $\boldsymbol{x} \neq \boldsymbol{0}$ であり, 上式より $\lambda_j = \boldsymbol{x}^* A\boldsymbol{x} > 0$ である．このとき命題 7.20 から $\operatorname{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n > 0$, $|A| = \lambda_1 \lambda_2 \cdots \lambda_n > 0$ である． □

記号 11.3 $H(n)$, $S(n)$ でそれぞれ n 次エルミート行列全体, n 次実対称行列全体を表し, $H^+(n)$, $S^+(n)$ でそれぞれ n 次正値エルミート行列全体, n 次正値対称行列全体を表す．

$H(1) = S(1)$ は実数全体の集合, $H^+(1) = S^+(1)$ は正の実数全体の集合とみなせるため, $H(n)$, $S(n)$ は実数全体の集合, $H^+(n)$, $S^+(n)$ は正の実数全体の集合を一般化したものであると考えられる．このような見方をすれば, 次の定理 11.4 は実数を変数とする指数関数が実数全体の集合から正の実数全体の集合への全単射を与えているという事実を一般化するものであると解釈される．

定理 11.4 exp による $H(n)$ の像は $H^+(n)$ であり, exp の定義域を $H(n)$ に制限した写像は単射である．同様に

exp による $S(n)$ の像は $S^+(n)$ であり, exp の定義域を $S(n)$ に制限した写像は単射である.

証明 A を任意のエルミート行列とすれば, 注意 9.12 から $\exp A$ もエルミート行列であり, さらに A の固有値はすべて実数だから命題 9.13 により, $\exp A$ の固有値はすべて正の実数である. 従って $\exp A$ は正値エルミート行列になるため, exp による $H(n)$ の像は $H^+(n)$ に含まれる.

C を任意の正値エルミート行列とすれば, $P^{-1}CP$ が対角行列になるようなユニタリー行列 P をとると $P^{-1}CP$ も正値エルミート行列であることから各対角成分は正の実数である. 故に $P^{-1}CP$ の (j,j) 成分を e^{x_j} $(x_j \in \boldsymbol{R})$ とおくことができる. B を (j,j) 成分が x_j であるような対角行列とすれば, $B \in H(n)$ であり, $\exp B = P^{-1}CP$ となる. 従って, 注意 9.12 から $\exp(PBP^{-1}) = C$ が得られる. $P^{-1} = P^*$ と $B^* = B$ から $PBP^{-1} \in H(n)$ だから exp による $H(n)$ の像は $H^+(n)$ に一致する.

$A, B \in H(n)$ に対して, A, B を対角化するユニタリー行列 P, Q をとる. $P^{-1}AP$, $Q^{-1}BQ$ の (j,j) 成分をそれぞれ λ_j, μ_j とすれば, 命題 8.13 の (1) からこれらはすべて実数だから P, Q の列を入れ換えることにより, $\lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_n$, $\mu_1 \leqq \mu_2 \leqq \cdots \leqq \mu_n$ と仮定してよい. 命題 9.13 により, $\exp A$, $\exp B$ の固有値はそれぞれ $e^{\lambda_1} \leqq e^{\lambda_2} \leqq \cdots \leqq e^{\lambda_n}$, $e^{\mu_1} \leqq e^{\mu_2} \leqq \cdots \leqq e^{\mu_n}$ となるため, $\exp A = \exp B$ と仮定すれば, $e^{\lambda_j} = e^{\mu_j}$ が $j = 1, 2, \ldots, n$ に対して成り立つ. 故に $j = 1, 2, \ldots, n$ に対して $\lambda_j = \mu_j$ だから $P^{-1}AP = Q^{-1}BQ$ である. そこで $D = P^{-1}AP = Q^{-1}BQ$, $U = Q^{-1}P$ とおくと, $U(\exp D) = Q^{-1}(\exp A)P = Q^{-1}(\exp B)P = (\exp D)U$ となるため, U の (j,k) 成分を u_{jk} として両辺の成分を比較すれば, $e^{\lambda_k}u_{jk} = e^{\lambda_j}u_{jk}$ が得られる. この等式から $\lambda_k \neq \lambda_j$ ならば $u_{jk} = 0$ がわかるため, $\lambda_k u_{jk} = \lambda_j u_{jk}$ が成り立つが, これは $UD = DU$ であることを意味する. 故に $Q^{-1}BQ = D = UDU^{-1} = (Q^{-1}P)(P^{-1}AP)(P^{-1}Q) = Q^{-1}AQ$ となって $A = B$ が得られるため exp は単射である.

$H(n)$, $H^+(n)$ の要素で成分がすべて実数の行列の全体がそれぞれ $S(n)$, $S^+(n)$ であり, exp は実数を成分にもつ行列を実数を成分にもつ行列に写すため, 上で示したことから exp は $S(n)$ を $S^+(n)$ に写す単射である. C を任意の正値対称行列とすれば, $P^{-1}CP$ が対角行列になるような直交行列 P がとれるため, 上と同じ議論で $\exp A = C$ を満たす $A \in S(n)$ が存在することが示される. 従って, exp による $S(n)$ の像は $S^+(n)$ に一致する. □

$A \in H(n)$ を $\exp A \in H^+(n)$ に写す写像も以後 $\exp : H(n) \to H^+(n)$ で表す. 定理 11.4 からこの写像は全単射であり, 逆写像を $\log : H^+(n) \to H(n)$ で表す.

補題 11.5 $C \in H^+(n)$ に対して次が成り立つ.

(1) 自然数 k に対し, $\log C^k = k \log C$ が成り立つ.

(2) ユニタリー行列 U に対し, $\log(U^{-1}CU) = U^{-1}(\log C)U$ が成り立つ.

(3) C が正の実数 λ_j を (j,j) 成分とする対角行列ならば $\log C$ は $\log \lambda_j$ を (j,j) 成分とする対角行列である.

証明 (1) 命題 10.3 の (1) から $\exp(k \log C) = (\exp(\log C))^k = C^k = \exp(\log C^k)$ であり, $\exp : H(n) \to H^+(n)$ は単射だから結果が得られる.

(2) U はユニタリー行列だから $U^{-1}(\log C)U \in H(n)$, $U^{-1}CU \in H^+(n)$ であり, 注意 9.12 から $\exp(U^{-1}(\log C)U) = U^{-1}(\exp(\log C))U = U^{-1}CU$ となるため, 結果が得られる.

(3) $\log \lambda_j$ を (j,j) 成分とする対角行列を D とおけば $\exp D$ は $e^{\log \lambda_j} = \lambda_j$ を (j,j) 成分とする対角行列で, C に一致するため $\log C = D$ である. □

12 正則行列の極分解

複素数を成分とする 1 次正則行列全体の集合 $GL_1(\boldsymbol{C})$ は 0 でない複素数全体の集合とみなせるため, 複素数を成分とする n 次正則行列全体の集合 $GL_n(\boldsymbol{C})$ は 0 でない複素数全体の集合を高次元に一般化している. 次の定理 12.1 は 0 でない複素数が正の実数と絶対値が 1 の複素数の積の形に一通りに表されるという事実を高次元に一般化している

と考えられる.

定理 12.1　複素数を成分とする n 次正則行列 A に対し, $A = HU$ を満たす正値エルミート行列 H とユニタリー行列 U がただ 1 組存在する. また, 実数を成分とする n 次正則行列 A に対し, $A = HU$ を満たす正値対称行列 H と直交行列 U がただ 1 組存在する.

証明　$C = AA^*$ とおくと命題 7.3 の (2) から $C^* = C$ である. 任意の $\boldsymbol{x} \in \boldsymbol{C}^n$ に対して

$$\boldsymbol{x}^*C\boldsymbol{x} = \boldsymbol{x}^*AA^*\boldsymbol{x} = (A^*\boldsymbol{x})^*A^*\boldsymbol{x} = {}^t(\overline{A^*\boldsymbol{x}})\overline{\overline{A^*\boldsymbol{x}}} = (\overline{A^*\boldsymbol{x}}, \overline{A^*\boldsymbol{x}}) = \|A^*\boldsymbol{x}\|^2 \geqq 0$$

であり, A^* は正則行列だから $\boldsymbol{x} = \boldsymbol{0}$ の場合に限り $\boldsymbol{x}^*C\boldsymbol{x} = 0$ となるため, C は正値エルミート行列である. $H = \exp\left(\frac{1}{2}\log(AA^*)\right)$, $U = H^{-1}A$ とおくと, 命題 10.3 の (1) から $H^2 = \exp(\log(AA^*)) = AA^*$, $A = HU$ である. $(H^{-1})^* = H^{-1}$ であることに注意すれば, 次の等式から U はユニタリー行列である.

$$U^*U = A^*H^{-1}H^{-1}A = A^*(H^2)^{-1}A = A^*(AA^*)^{-1}A = A^*(A^*)^{-1}A^{-1}A = E_n$$

正値エルミート行列 G とユニタリー行列 P に対して $A = GP$ ならば $AA^* = GPP^*G^* = G^2$ が成り立つため, $H = \exp\left(\frac{1}{2}\log(AA^*)\right) = \exp\left(\frac{1}{2}\log G^2\right) = \exp(\log G) = G$ だから $A = HU = HP$ となり $P = U$ が得られる.

A が実数を成分とする n 次正則行列ならば定理 11.4 の後半の主張により, log は $S^+(n)$ を $S(n)$ に写すため, A に対して上と同様に定めた H は正値対称行列である. $U = H^{-1}A$ は実数を成分にもつ行列で, 上の結果から U はユニタリー行列だから直交行列である. □

正則行列 A に対し, $A = HU$ を満たす正値エルミート行列 H とユニタリー行列 U の対 (H, U) を求めることを, A の極分解を求めるという.

命題 12.2　$A \in GL_n(\boldsymbol{C})$ に対し, 正値エルミート行列 AA^* を対角化するユニタリー行列を P として

$$P^{-1}AA^*P = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}, \quad H = P\begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix}P^{-1}, \quad U = H^{-1}A$$

とおけば, (H, U) は A の極分解を与える.

証明　補題 11.5 の (2) と (3) より

$$P^{-1}\log(AA^*)P = \log(P^{-1}AA^*P) = \begin{pmatrix} \log\lambda_1 & 0 & \cdots & 0 \\ 0 & \log\lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \log\lambda_n \end{pmatrix}$$

となるため,

$$P^{-1}\exp\left(\frac{1}{2}\log(AA^*)\right)P = \exp\left(\frac{1}{2}P^{-1}\log(AA^*)P\right) = \exp\begin{pmatrix} \frac{1}{2}\log\lambda_1 & 0 & \cdots & 0 \\ 0 & \frac{1}{2}\log\lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{2}\log\lambda_n \end{pmatrix}$$

$$= \exp\begin{pmatrix} \log\sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \log\sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \log\sqrt{\lambda_n} \end{pmatrix} = \begin{pmatrix} e^{\log\sqrt{\lambda_1}} & 0 & \cdots & 0 \\ 0 & e^{\log\sqrt{\lambda_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\log\sqrt{\lambda_n}} \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix}$$

である．従って

$$H=\exp\left(\frac{1}{2}\log(AA^*)\right)=P\begin{pmatrix}\sqrt{\lambda_1}&0&\cdots&0\\0&\sqrt{\lambda_2}&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&\sqrt{\lambda_n}\end{pmatrix}P^{-1},\quad U=H^{-1}A=P\begin{pmatrix}\frac{1}{\sqrt{\lambda_1}}&0&\cdots&0\\0&\frac{1}{\sqrt{\lambda_2}}&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&\frac{1}{\sqrt{\lambda_n}}\end{pmatrix}P^{-1}A$$

とおけば, (H,U) が A の極分解である． □

13　弱位相と写像の列の広義一様収束

定義 13.1　$(X,\mathcal{O})$ を位相空間, $\mathcal{K}$ を X の部分空間からなる集合とする．次の条件 (WT) が成り立つとき, $(X,\mathcal{O})$ は $\mathcal{K}$ に関して弱位相 (weak topology) をもつという．

　(WT) X の部分集合 O が条件「$K\in\mathcal{K}$ ならば $O\cap K$ は K の開集合である.」を満たせば $O\in\mathcal{O}$ である．

とくに $\mathcal{K}$ が X のコンパクトな部分空間全体からなる集合の場合に $(X,\mathcal{O})$ が $\mathcal{K}$ に関して弱位相をもつとき, $(X,\mathcal{O})$ はコンパクト生成位相をもつといい, $(X,\mathcal{O})$ を CG 空間 (compactly generated space) と呼ぶ.

命題 13.2　X の部分空間からなる集合 $\mathcal{K}$ が $X=\displaystyle\bigcup_{K\in\mathcal{K}}K^i$ を満たせば, 位相空間 $(X,\mathcal{O})$ は $\mathcal{K}$ に関して弱位相をもつ.

証明　X の部分集合 O が条件「$K\in\mathcal{K}$ ならば $O\cap K$ は K の開集合である.」を満たすとする．仮定から任意の $p\in O$ に対して, $p\in K^i$ を満たす $K\in\mathcal{K}$ と $O\cap K=U\cap K$ を満たす $U\in\mathcal{O}$ が存在する．このとき, $p\in U\cap K^i\in\mathcal{O}$ であり, $U\cap K^i\subset U\cap K=O\cap K\subset O$ だから, p は O の内点であり, O は X の開集合である． □

　上の結果から直ちに次の結果が得られる．

系 13.3　X の各点がコンパクトな近傍をもつとき, $(X,\mathcal{O})$ は CG 空間である．

命題 13.4　$(X,\mathcal{O})$ を位相空間, $\mathcal{K}$ を X の部分空間からなる集合とする．$(X,\mathcal{O})$ が $\mathcal{K}$ に関して弱位相をもつためには次の条件 (WT') が成り立つことが必要十分である．

　(WT') X の部分集合 A が条件「$K\in\mathcal{K}$ ならば $A\cap K$ は K の閉集合である.」を満たせば A は X の閉集合である．

証明　$(X,\mathcal{O})$ が $\mathcal{K}$ に関して弱位相をもつと仮定し, X の部分集合 A が条件「$K\in\mathcal{K}$ ならば $A\cap K$ は K の閉集合である.」を満たすと仮定すれば, 任意の $K\in\mathcal{K}$ に対し, $A\cap K=(X-O)\cap K$ を満たす X の開集合 O が存在する．このとき, $(X-A)\cap K=K-(A\cap K)=K-((X-O)\cap K)=O\cap K$ だから, $X-A$ は K の開集合である．$(X,\mathcal{O})$ が $\mathcal{K}$ に関して弱位相をもつことから, $X-A$ は X の開集合である．従って A は X の閉集合であるため, 条件 (WT') が成り立つ．

　逆に条件 (WT') が成り立つと仮定して, X の部分集合 O が条件「$K\in\mathcal{K}$ ならば $O\cap K$ は K の開集合である.」を満たすと仮定すれば, 任意の $K\in\mathcal{K}$ に対し, $O\cap K=(X-A)\cap K$ を満たす X の閉集合 A が存在する．このとき, $(X-O)\cap K=K-(O\cap K)=K-((X-A)\cap K)=A\cap K$ だから, $X-O$ は K の閉集合である．条件 (WT') が成り立つことから, $X-O$ は X の閉集合である．従って O は X の開集合であるため, $(X,\mathcal{O})$ は $\mathcal{K}$ に関して弱位相をもつ． □

命題 13.5　$(X,\mathcal{O}_X)$ を位相空間とし, X の部分空間からなる集合 $\mathcal{K}$ に関して $(X,\mathcal{O}_X)$ は弱位相をもつとする．位相空間 $(Y,\mathcal{O}_Y)$ に対し, 写像 $f:X\to Y$ が連続であるためには, 任意の $K\in\mathcal{K}$ に対して, f の定義域を K に制限して得られる写像 $f|_K:K\to Y$ が連続であることが必要十分である．

証明　任意の $K\in\mathcal{K}$ に対して, $f|_K:K\to Y$ が連続であると仮定する．Y の 任意の開集合 O に対し, $f|_K$ の連続性から $f^{-1}(O)\cap K=(f|_K)^{-1}(O)$ は K の開集合だから, $f^{-1}(O)$ は X の開集合である．故に f は連続写像である．逆は明らかである． □

定義 13.6　$(X,\mathcal{O})$ を位相空間, (Y,d) を距離空間とし, $f, f_n : X \to Y$ $(n=1,2,\dots)$ を写像とする. 写像の列 $(f_n)_{n\in\boldsymbol{N}}$ が次の条件 $(*)$ を満たすとき, $(f_n)_{n\in\boldsymbol{N}}$ は f に広義一様収束するという.

$(*)$ X の任意のコンパクトな部分空間 K に対して, f_n の定義域を K に制限して得られる写像 $f_n|_K : K \to Y$ の列 $(f_n|_K)_{n\in\boldsymbol{N}}$ が f の定義域を K に制限して得られる写像 $f|_K : K \to Y$ に一様収束する.

命題 1.12 と命題 13.5 から次の結果が得られる.

命題 13.7　$(X,\mathcal{O})$ を CG 空間, (Y,d) を距離空間とし, $f, f_n : X \to Y$ $(n=1,2,\dots)$ を写像とする. すべての n に対して f_n が連続で, 写像の列 $(f_n)_{n\in\boldsymbol{N}}$ が f に広義一様収束するならば f は連続写像である.

命題 13.8　x の整級数 $\sum\limits_{k=0}^{\infty} a_k x^k$ の収束半径を R とし, $D_R=\{X\in M_n(\boldsymbol{C})\,|\,\|X\|<R\}$ とおく. 写像 $f:D_R\to M_n(\boldsymbol{C})$ を $f(X)=\sum\limits_{k=0}^{\infty} a_k X^k = a_0E_n + a_1X + \cdots + a_kX^k + \cdots$ で定義すれば, f は連続である.

証明　自然数 N に対して写像 $f_N : D_R \to M_n(\boldsymbol{C})$ を $f_N(X) = a_0E_n + a_1X + \cdots + a_kX^k + \cdots + a_NX^N$ で定義すれば, f_N は連続である. また, $\bar{f}(x) = \sum\limits_{k=0}^{\infty} |a_k|x^k$ で整級数 $\bar{f}(x)$ を定め, 自然数 N に対して x の多項式 $\bar{f}_N(x)$ を $\bar{f}_N(x) = \sum\limits_{k=0}^{N} |a_k|x^k$ で定める. 正の実数 r に対して $\bar{D}_r = \{X \in M_n(\boldsymbol{C})\,|\,\|X\| \leqq r\}$ とおく. 自然数 N と $0<r<R$ を満たす実数 r に対して, $X \in \bar{D}_r$ ならば

$$\|f_N(X) - f(X)\| = \left\|\sum_{k=N+1}^{\infty} a_kX^k\right\| \leqq \sum_{k=N+1}^{\infty} |a_k|\|X\|^k \leqq \sum_{k=N+1}^{\infty} |a_k|r^k = \bar{f}(r) - \bar{f}_N(r)$$

であり, $\lim\limits_{N\to\infty} \bar{f}_N(r) = \bar{f}(r)$ だから写像の列 $(f_N)_{N\in\boldsymbol{N}}$ は $\bar{D}_r$ において f に一様収束する. D_R の任意のコンパクトな部分空間 K に対して $r = \max\{\|X\|\,|\,X \in K\}$ とおけば $K \subset \bar{D}_r$ だから, $(f_N)_{N\in\boldsymbol{N}}$ は D_R において f に広義一様収束するため, 系 13.3 と命題 13.7 から, f は連続である. □

上の結果から, 行列の指数写像 $\exp : M_n(\boldsymbol{C}) \to GL_n(\boldsymbol{C})$ は連続である.

定理 13.9　$\exp : H(n) \to H^+(n)$ は同相写像である.

証明　定理 11.4 と命題 13.8 から $\exp : H(n) \to H^+(n)$ は連続な全単射だから, $H(n)$ の閉集合 F に対して $\exp(F)$ が $H^+(n)$ の閉集合になることを示せばよい. $(B_k)_{k\in\boldsymbol{N}}$ が $H^+(n)$ の点 B に収束する $\exp(F)$ の点列であると仮定する. 各自然数 k に対して $\exp(A_k) = B_k$ を満たす $A_k \in F$ を選び, さらに $U_k^{-1}A_kU_k$ が対角行列になるようなユニタリー行列 U_k を選んで, ユニタリー行列からなる点列 $(U_k)_{k\in\boldsymbol{N}}$ を考える. 補題 10.10 から $U(n)$ はコンパクトだから, 定理 1.4 により $(U_k)_{k\in\boldsymbol{N}}$ の部分列 $(U_{k_i})_{i\in\boldsymbol{N}}$ で $U(n)$ の点 U に収束するものが存在する. このとき, $(B_k)_{k\in\boldsymbol{N}}$ の部分列 $(B_{k_i})_{i\in\boldsymbol{N}}$ も B に収束するため, $(A_{k_i})_{i\in\boldsymbol{N}}$, $(B_{k_i})_{i\in\boldsymbol{N}}$, $(U_{k_i})_{i\in\boldsymbol{N}}$ をそれぞれ, $(A_k)_{k\in\boldsymbol{N}}$, $(B_k)_{k\in\boldsymbol{N}}$, $(U_k)_{k\in\boldsymbol{N}}$ で置き換えて, $(U_k)_{k\in\boldsymbol{N}}$ は U に収束すると仮定してよい. このとき $\lim\limits_{k\to\infty} \exp(U_k^{-1}A_kU_k) = \lim\limits_{k\to\infty} U_k^{-1}\exp(A_k)U_k = \lim\limits_{k\to\infty} U_k^{-1}B_kU_k = U^{-1}BU$ であり, すべての自然数 k に対して $U_k^{-1}A_kU_k$ は対角行列だから, $\exp(U_k^{-1}A_kU_k)$ も対角行列であるため, $U^{-1}BU$ も対角行列である. $U^{-1}BU$ の (i,i) 成分を λ_i とおけば, $B \in H^+(n)$ より, $i=1,2,\dots,n$ に対して $\lambda_i > 0$ である. さらに $U_k^{-1}A_kU_k$ の (i,i) 成分を $\lambda_{k,i}$ とすれば, $\lim\limits_{k\to\infty} \exp(U_k^{-1}A_kU_k) = U^{-1}BU$ より, $e^{\lim\limits_{k\to\infty}\lambda_{k,i}} = \lim\limits_{k\to\infty} e^{\lambda_{k,i}} = \lambda_i = e^{\log\lambda_i}$ だから $\lim\limits_{k\to\infty} \lambda_{k,i} = \log\lambda_i$ が得られる. 従って, $\log\lambda_i$ を (i,i) 成分とする対角行列を D で表せば, $\lim\limits_{k\to\infty} U_k^{-1}A_kU_k = D$ である. 故に $\lim\limits_{k\to\infty} A_k = \lim\limits_{k\to\infty} U_k(U_k^{-1}A_kU_k)U_k^{-1} = UDU^{-1}$ が成り立ち, $(A_k)_{k\in\boldsymbol{N}}$ は閉集合 F の点列だから, $UDU^{-1} \in F$ であることがわかる. 一方, 指数写像の連続性から $\exp(UDU^{-1}) = \exp\left(\lim\limits_{k\to\infty} A_k\right) = \lim\limits_{k\to\infty} \exp A_k = \lim\limits_{k\to\infty} B_k = B$ だから, $B \in \exp(F)$ が得られ, $\exp(F)$ が閉集合であることが示された. □

定理 13.9 から $\log : H^+(n) \to H(n)$ は連続な全単射で, 定理 11.4 から exp は $S(n)$ から $S^+(n)$ への連続な全単射を与えるため, log は $S^+(n)$ から $S(n)$ への連続な全単射を与える. 従って次の結果が得られる.

系 13.10 $A \in S(n)$ を $\exp A \in S^+(n)$ に対応させる $S(n)$ から $S^+(n)$ への写像は同相写像である.

$O(n)$ を n 次直交行列全体からなる $M_n(\boldsymbol{C})$ の部分空間とする. 定理 12.1 の証明と, $\exp : H(n) \to H^+(n)$, $\log : H^+(n) \to H(n)$ の連続性から次の結果が得られる.

系 13.11 (1) 写像 $\Phi_{\boldsymbol{C}} : GL_n(\boldsymbol{C}) \to H^+(n) \times U(n)$, $\Psi_{\boldsymbol{C}} : H^+(n) \times U(n) \to GL_n(\boldsymbol{C})$ を以下のように定めれば, これらは互いに逆写像で, 同相写像である.

$$\Phi_{\boldsymbol{C}}(A) = \left(\exp\left(\frac{1}{2}\log(AA^*)\right), \exp\left(-\frac{1}{2}\log(AA^*)\right)A\right), \qquad \Psi_{\boldsymbol{C}}(H,U) = HU$$

(2) 写像 $\Phi_{\boldsymbol{R}} : GL_n(\boldsymbol{R}) \to S^+(n) \times O(n)$, $\Psi_{\boldsymbol{R}} : S^+(n) \times O(n) \to GL_n(\boldsymbol{R})$ を以下のように定めれば, これらは互いに逆写像で, 同相写像である.

$$\Phi_{\boldsymbol{R}}(A) = \left(\exp\left(\frac{1}{2}\log(AA^*)\right), \exp\left(-\frac{1}{2}\log(AA^*)\right)A\right), \qquad \Psi_{\boldsymbol{R}}(S,T) = ST$$

$H(n)$, $S(n)$ は $\boldsymbol{R}$ 上のベクトル空間として, それぞれ $\boldsymbol{R}^{n^2}$, $\boldsymbol{R}^{\frac{n(n+1)}{2}}$ と同型で, この同型写像は同相写像でもある. 従って定理 13.9, 系 13.10, 系 13.11 から次の結果が得られる.

定理 13.12 $GL_n(\boldsymbol{C})$ は $\boldsymbol{R}^{n^2} \times U(n)$ と同相であり, $GL_n(\boldsymbol{R})$ は $\boldsymbol{R}^{\frac{n(n+1)}{2}} \times O(n)$ と同相である.

$H_0(n) = \{A \in H(n) | \operatorname{tr} A = 0\}$, $H_1^+(n) = \{A \in H^+(n) | \det A = 1\}$ とおく.

命題 13.13 $\exp : H(n) \to H^+(n)$ は $H_0(n)$ から $H_1^+(n)$ への同相写像を与える.

証明 命題 10.3 の (3) から $\exp : H(n) \to H^+(n)$ による $H_0(n)$ の像は $H_1^+(n)$ に含まれる. 定理 11.4 により任意の $B \in H_1^+(n)$ に対して $\exp A = B$ を満たす $A \in H(n)$ が存在する. エルミート行列の対角成分は実数だから $\operatorname{tr} A$ が実数であることに注意すれば命題 10.3 の (3) から $e^{\operatorname{tr} A} = \det B = 1$ だから $\operatorname{tr} A = 0$ である. 従って exp による $H_0(n)$ の像は $H_1^+(n)$ だから定理 13.9 から主張が成り立つ. □

$\boldsymbol{K}$ の要素を成分にもち, 行列式の値が 1 である n 次正方行列全体からなる $M_n(\boldsymbol{K})$ の部分空間を $SL_n(\boldsymbol{K})$ で表す. また行列式の値が 1 である n 次ユニタリー行列全体からなる $U(n)$ の部分空間を $SU(n)$ で表し, 行列式の値が 1 である n 次直交行列全体からなる $O(n)$ の部分空間を $SO(n)$ で表す.

命題 13.14 $A \in SL_n(\boldsymbol{C})$ の極分解を (H, U) とすれば $\det H = \det U = 1$ である.

証明 $\det A = 1$ ならば $\det A^* = \overline{\det A} = 1$ だから $\det(AA^*) = (\det A)(\det A^*) = 1$ である. 従って $AA^* \in H_1^+(n)$ だから命題 13.13 により $\log(AA^*) \in H_0(n)$ である. $\boldsymbol{R}$ 上のベクトル空間として $H_0(n)$ は $H(n)$ の部分空間だから $\frac{1}{2}\log(AA^*) \in H_0(n)$ となるため, 命題 13.13 により $H = \exp\left(\frac{1}{2}\log(AA^*)\right) \in H_1^+(n)$ である. さらに $U = H^{-1}A$ であり, $\det H = \det A = 1$ だから $\det U = 1$ である. □

$S_1^+(n) = S(n) \cap H_1^+(n)$ とおく. 命題 13.13, 命題 13.14, 系 13.11 と $\boldsymbol{R}$ 上のベクトル空間として $H_0(n)$ は $H(n)$ の次元より 1 次元低い $H(n)$ の部分空間であることから次の結果が得られる.

系 13.15 系 13.11 で定義した写像 $\Phi_{\boldsymbol{C}} : GL_n(\boldsymbol{C}) \to H^+(n) \times U(n)$ は $SL_n(\boldsymbol{C})$ から $H_1^+(n) \times SU(n)$ への同相写像を与え, $\Phi_{\boldsymbol{R}} : GL_n(\boldsymbol{R}) \to S^+(n) \times O(n)$ は $SL_n(\boldsymbol{R})$ から $S_1^+(n) \times SO(n)$ への同相写像を与える. 従って $SL_n(\boldsymbol{C})$ は $\boldsymbol{R}^{n^2-1} \times SU(n)$ と同相であり, $SL_n(\boldsymbol{R})$ は $\boldsymbol{R}^{\frac{n(n+1)}{2}-1} \times SO(n)$ と同相である.