経済学のための数学

京都大学経済研究所 原千秋

平成 20 年 9 月 19 日

目 次

1 微分法

1.1 1 変数関数の連続性

1.1.1 数列の収束

$\boldsymbol{R}$ の要素 $x_1, x_2, x_3, \ldots$ の並びを数列と言い，これをまとめて数列 $\left(x_n\right)$ $(n = 1, 2, 3, \ldots)$ などと表す．

1.1 定義 (**数列の収束**)**.** $\boldsymbol{R}$ 上の数列 $\left(x_n\right)$ と点 $x \in \boldsymbol{R}$ を考える． 任意の正の数 ϵ に対してある自然数 N が存在し， その N より大きい n について $\left|x_n - x\right| < \epsilon$ が常に成り立つ時，数列 $\left(x_n\right)$ は x に収束すると言う． またこれを，

$$\forall\epsilon > 0, \quad \exists N \in \boldsymbol{N}, \quad \forall n > N: \quad \left|x_n - x\right| < \epsilon$$

と書く．

数列 $\left(x_n\right)$ が x に収束するとは，直感的には， n の値が大きくなるにつれて x_n の値が限りなく x に近づいていくことを意味しており，

$$\lim_{n\to\infty} x_n = x \quad \text{あるいは} \quad x_n \to x \ (n \to \infty)$$

などと表す．またこの時，x は数列 $\left(x_n\right)$ の極限であると言う．

1.1 例. $x_n = 1/n$ で定義される数列を考えれば，この数列 $\left(x_n\right)$ は 0 に収束する． 実際，任意の正の実数 ϵ に対して $N \geq 1/\epsilon$ を満たすように自然数 $N \in \boldsymbol{N}$ を取れば， 任意の $n > N$ について

$$0 < x_n = \frac{1}{n} < \frac{1}{N} \leq \epsilon$$

が常に成立する．よって

$$\forall n > N: \quad \left|x_n - 0\right| < \epsilon$$

これは数列 $\left(x_n\right)$ が 0 に収束することの定義に他ならない．

1.2 例. 数列 $\left(x_n\right)$ を

$$x_n = \begin{cases} \dfrac{1}{n} & (\ n \text{ が奇数のとき}\ ) \\ 1 & (\ n \text{ が偶数のとき}) \end{cases}$$

によって定義すれば，この数列は $\boldsymbol{R}$ 上のいずれの点にも収束しない． これを確認するために，まずは $\left(x_n\right)$ が 0 に収束しないことを示そう． 数列がある値に収束しないことを示すためには， 定義 1.1 の否定が満たされることを言えばよい． 定義 1.1 の条件は

$$\forall\epsilon > 0,\ \exists N \in \boldsymbol{N},\ \forall n > N: \quad \left|x_n - x\right| < \epsilon$$

であったから，その否定は

$$\exists\epsilon > 0,\ \forall N \in \boldsymbol{N},\ \exists n > N: \quad \left|x_n - x\right| \geq \epsilon \tag{1.1}$$

である．つまりある正の実数 ϵ を取った時， 任意の自然数 N に対してその N より大きな数 n が存在し， $\left|x_n - x\right| \geq \epsilon$ が成り立つことを示せばよい．

この例では $\epsilon = 1/2$ とすれば，任意の偶数 n に対して

$$\left|x_n - 0\right| = \left|1 - 0\right| = 1 > \epsilon \tag{1.2}$$

が成立する．よって ϵ が $1/2$ という値を取る時，どんなに大きな自然数 N を取っても，それより大きな偶数 n について必ずこの不等式が成り立つ．これは条件 (1.1) を満たすので，$\left(x_n\right)$ が 0 に収束することはないと言える．ほぼ同様の議論から，$\left(x_n\right)$ が 1 に収束しないことも示せる．(任意の奇数 n を考えよ.)

練習問題 1.1. 例 1.2 で定義した数列 $\left(x_n\right)$ が，$\boldsymbol{R}$ 上のいかなる点にも収束することはないことを示せ．

練習問題 1.2. 数列の極限が存在するとすれば，それは一意であることを示せ．

1.1.2 1 変数関数の連続性

関数 $f:\boldsymbol{R}\to\boldsymbol{R}$ は数列 $\left(x_n\right)$ に対応して，その像の列 $f(x_1), f(x_2), f(x_3), \ldots$ を定める．よって数列の写像も数列となり，数列自体の極限を考えるのと同様に，数列の像の極限を考えることもできる．

1.3 例. 関数 $f:\boldsymbol{R}\to\boldsymbol{R}$ を

$$f(x) = x^3 + 1 \tag{1.3}$$

によって定義しよう．さらに数列 $\left(x_n\right)$ を，1.1 例と同様に $x_n = 1/n$ で定義する．この時

$$f(x_n) = x_n^3 + 1 = \frac{1}{n^3} + 1$$

であるから

$$f(x_n) \to 1 \ \ (n \to \infty)$$

となり，$f(x_n)$ は 1 に収束することが分かる．

また別の数列として $x_n = (-1)^n/n$ を定義し，その像の収束を考えてみる．この時

$$f(x_n) = x_n^3 + 1 = \frac{(-1)^n}{n^3} + 1$$

であるから

$$f(x_n) \to 1 \ \ (n \to \infty)$$

となり，これも先ほどの数列と同様に，1 に収束することが分かる．

例 1.3 に挙げた二つの数列 $\left(x_n\right)$ は，いずれも $\lim_{n\to\infty} x_n = 0$，つまり 0 に収束する数列であった．実は (1.3) で定義される関数 f については，0 に収束する任意の数列 $\left(x_n\right)$ に対して，

$$f(x_n) \to 1 \ \ (n \to \infty)$$

が成立する．これは関数の連続性と呼ばれる性質である．

1.2 定義 (関数の連続性). x を $\boldsymbol{R}$ の要素，f を $\boldsymbol{R}$ 上で定義される関数とする．$x_n \to x \ (n\to\infty)$ を満たす任意の数列 $\left(x_n\right)$ に対して，

$$f(x_n) \to f(x) \ \ (n \to \infty) \quad \text{あるいは} \quad \lim_{n\to\infty} f(x_n) = f(x)$$

が成り立つ時，関数 f は x において連続であると言う．このとき

$$f(y) \to f(x) \ \ (y \to x) \quad \text{あるいは} \quad \lim_{y\to x} f(y) = f(x)$$

などとも書く．

関数の連続性とは，極限を取る操作と像を取る操作の順序をとりかえてもその値が変化しないことを意味している． つまり, 数列の極限 $(x_n \to x)$ を取ってから その極限における関数の像 $(f(x))$ を取った場合と， 数列の像 $(f(x_n))$ を取ってからその極限 $\left(\lim_{n\to\infty} f(x_n)\right)$ を取った場合とで， 収束先の値が異なることがないということである． （つまり，$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right)$ が成り立つ.）例 1.3 で定義された関数 f はこの性質を満たしているため， 0 において連続であると言える．

また関数の連続性の定義 1.2 は，以下のようにも言い換えられる．

$$\forall\epsilon > 0,\ \exists\delta > 0,\ \forall y \in \boldsymbol{R}: \quad \left|x - y\right| < \delta \quad \Rightarrow \quad \left|f(x) - f(y)\right| < \epsilon$$

練習問題 1.3. 上記の関数の連続性の定義が定義 1.2 と同値であることを証明せよ。

1.4 例. $\boldsymbol{R}$ を定義域とする関数 f を，

$$f(x) = \begin{cases} 1 & （\ x \geq 0 \text{ のとき }） \\ 0 & （\ x < 0 \text{ のとき }） \end{cases}$$

と定義しよう．するとこの関数は，$x = 0$ において不連続である．

ある点において関数が不連続であることを確認するには， その問題となる点に収束する数列をいくつか考え，それらの中に関数の連続性の定義を満たさないものがあることを示せばよい． ここでは 0 に収束する二つの数列， 例えば $x_n = 1/n$ で定義される数列と， $x_n = -1/n$ で定義される数列を考えることにする．

まず前者の数列については

$$\forall n, \quad x_n = \frac{1}{n} > 0$$

であるから，関数 f の定義により

$$\forall n, \quad f(x_n) = 1$$

となり，その極限は

$$f(x_n) \to 1 \ \ (n \to \infty)$$

である． 関数 f の定義により $f(0) = 1$ であったから，この時

$$f(x_n) \to f(0) \ \ (n \to \infty)$$

が成立し，この数列に関しては点 0 における連続性の要件が満たされる [1]．

一方，後者の数列については

$$\forall n, \quad x_n = -\frac{1}{n} < 0$$

であるから，関数 f の定義により

$$\forall n, \quad f(x_n) = 0$$

となり，その極限は

$$f(x_n) \to 0 \ \ (n \to \infty)$$

として 0 に収束してしまう． よってこの時，数列の極限の像と数列の像の極限とが一致しない． すなわち

$$f(x_n) \nrightarrow f(0) \ \ (n \to \infty)$$

である．つまりこちらの数列に関しては，点 0 における連続性の要件が満たされていない [2]．

関数 f が点 0 で連続であると言えるためには， 0 に収束する任意 [3] の数列で要件が満たされなければならなかったから， この関数は点 0 で連続ではないこということになる．

[1] ある点に対して，実数直線上を右側から近づいていく数列の取り方によらずにある極限が存在するとき，この値をこの点における右極限と言う．またこの例のように，右極限が，関数がその点において取る値と一致している時， 関数はその点において右連続であると言う．

[2] 右極限（右連続）と同様にして，左極限（左連続）が定義される．この関数は左連続でない．

[3] これはつまり，右極限と左極限をチェックする必要があるということ．

1.2 多変数関数の連続性

1.2.1 点列の収束

L 次元空間 $\boldsymbol{R}^L$ の要素 $x_1, x_2, x_3, \ldots$ の並び，すなわち $\left(x_n\right)$ $(n = 1, 2, 3, \ldots)$ を点列という．点列の収束は数列の収束とほぼ同様にして定義できるが，まずはノルム（Euclidean Norm）の概念を導入しておく必要がある．

1.3 定義 (**ノルム**)**.** $\boldsymbol{R}^L$ の要素 $x = \left(x^1, x^2, x^3, \ldots, x^L\right)$ に対して，

$$\|x\| = \sqrt{\left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2 + \cdots + \left(x^L\right)^2} \tag{1.4}$$

で定義される $\|x\|$ の値を，$x \in \boldsymbol{R}^L$ のノルムと言う．

$L = 2,\ 3$ の場合で確認できるように，ノルムはベクトルの長さに他ならない．

1.4 定義. $\boldsymbol{R}^L$ 上の点列 $\left(x_n\right)$ と点 $x \in \boldsymbol{R}^L$ を考える．

$$\lim_{n\to\infty}\left\|x_n - x\right\| = 0 \quad \text{あるいは} \quad \left\|x_n - x\right\| \to 0\ (n \to \infty)$$

となる時，点列 $\left(x_n\right)$ は x に収束すると言う． またこの時

$$\lim_{n\to\infty} x_n = x \quad \text{あるいは} \quad x_n \to x\ (n \to \infty) \tag{1.5}$$

などと表し，x は点列 $\left(x_n\right)$ の極限であると言う．

点列 $\left(x_n\right)$ が x に収束するとは， 点列 $\left(x_n\right)$ のそれぞれの成分が x の対応する成分に収束することに他ならない． つまり点列 $x_n = \left(x_n^1, x_n^2, x_n^3, \ldots, x_n^L\right)$ について，1.4 定義は以下の条件と同値である．

$$\forall l = 1, 2, 3, \ldots, L:\ x_n^l \to x^l\ (n \to \infty) \tag{1.6}$$

練習問題 1.4. 定義 1.4 と (1.6) とが同値であることを証明せよ．

1.5 例. $\boldsymbol{R}^2$ 上の点 $0 = (0, 0)$ に収束する点列を考えよう． 例えば

$$x_n = \left(\frac{1}{n}, 0\right) \tag{1.7}$$

$$x_n = \left(0, \frac{1}{n}\right) \tag{1.8}$$

$$x_n = \left(\frac{1}{n}, \frac{1}{n}\right) \tag{1.9}$$

$$x_n = \left(\frac{1}{n}\cos n, \frac{1}{n}\sin n\right) \tag{1.10}$$

のように点列を定義すれば， これらはいずれも $\left\|x_n\right\| \to 0$ $(n \to \infty)$ を満たすから，点 $0 \in \boldsymbol{R}^2$ に収束する点列である． しかし，それぞれの点列が極限である点 0 に近付く方向は全く異なっている．

1.2.2 多変数関数の連続性

以上のように点列の収束を定義すれば， L 次元空間を定義域とする多変数関数 $F:\boldsymbol{R}^L\to\boldsymbol{R}^M$ についても，1 変数関数の場合と同様に連続性を考えることができる．

$\boldsymbol{R}^L$ を定義域，$\boldsymbol{R}^M$ を値域とする関数 $F:\boldsymbol{R}^L\to\boldsymbol{R}^M$ を考える． このような多変数関数は，

$$F(x)=F(x^1,x^2,\ldots,x^L)$$

のように定義域のベクトル $x\in\boldsymbol{R}^L$ が成分で表示されることもあり， またさらに

$$F(x)=\begin{pmatrix}F_1(x^1,x^2,\ldots,x^L)\\F_2(x^1,x^2,\ldots,x^L)\\\vdots\\F_M(x^1,x^2,\ldots,x^L)\end{pmatrix}$$

のように値域の次元が明示されることもある．

1.5 定義. x を $\boldsymbol{R}^L$ の要素，F を $\boldsymbol{R}^L$ 上で定義される多変数関数とする． $x_n\to x\ (n\to\infty)$ を満たす任意の点列 (x_n) に対して

$$F(x_n)\to F(x)\ (n\to\infty)$$

が成り立つ時，すなわち $\|x_n-x\|\to 0\ (n\to\infty)$ を満たす任意の点列 (x_n) に対して

$$\big\|F(x_n)-F(x)\big\|\to 0\ (n\to\infty)$$

が成り立つ時，関数 F は x において連続であると言う． これを

$$\lim_{y\to x}F(y)=F(x)\quad \text{あるいは}\quad F(y)\to F(x)\ (y\to x)$$

などとも書く．

関数 F が点 0 で連続であるとは，様々な近付き方をする点列のいずれについても， その点列の像が $F(0)$ に収束することを意味する．

また 1 変数関数の連続性の定義と同様に， 多変数関数の連続性も ϵ-δ 論法を用いて言い換えることができる．つまり定義 1.5 は

$$\forall\epsilon>0,\ \exists\delta>0,\ \forall x\in\boldsymbol{R^L}:\quad \big\|x-u\big\|<\delta\quad\Rightarrow\quad\big\|f(x)-f(u)\big\|<\epsilon$$

と同値である．ただし，ここでノルムが定義されている空間の次元は異なることに注意せよ．

多変数関数の連続性が 1 変数の連続性と大きく異なる点は， 「ある点に収束する任意の点列」と言った場合に，より多様な点列を想定しなければならないことにある．1 変数の場合であれば，数列がある点に収束する時， その近付く方向は水平方向に限られていた． しかし多次元空間における点列の収束となると，同じ 1 点に収束するものであっても，近付く方向としては様々なものが考えられるからである．

1.3 多変数関数の微分

ベクトル $(x^1,x^2,\ldots,x^L)$ の変化に対して この多変数関数 F の値がどのように変化するかを見るための方法が， 多変数関数の微分である．

1.3.1 偏微分

理解し易さの便宜から， $L=2$，$M=1$ とし，$F(x^1,x^2)=F(x,y)$ と置いて話をすすめる．

1.6 定義 (**1 変数関数の微分可能性**)**.** 1 変数関数 $f:\boldsymbol{R}\to\boldsymbol{R}$ の定義域上の点 $x\in\boldsymbol{R}$ を所与として，極限

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$$

が存在する時，f は点 x において微分可能であると言い，その極限の値を

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=\frac{df}{dx}(x,y)=f'(x)$$

などと表す． これは f の x における微分係数と呼ばれる． また f が定義域上の任意の点 $x\in\boldsymbol{R}$ において微分可能である時，f は $\boldsymbol{R}$ において微分可能であると言う．f が $\boldsymbol{R}$ において微分可能であるとき，$\boldsymbol{R}$ の各点にその点における微分係数を対応付ける関数 $x\mapsto f'(x)$ を f の導関数といい，$f':\boldsymbol{R}\to\boldsymbol{R}$ と表す． f' が x において連続なとき，f は x において連続微分可能という． もし各点 x において連続微分可能なら，f は（$\boldsymbol{R}$ において）連続微分可能という．

練習問題 1.5. n をゼロまたは正の整数とし，関数 $f:\boldsymbol{R}\to\boldsymbol{R}$ を

$$f(x)=\begin{cases} x^n\sin\left(\dfrac{1}{x}\right) & (x\neq 0 \text{ に対して}) \\ 0 & (x=0 \text{ に対して}) \end{cases}$$

と定義する．

1. $n=0$ ならば（つまり，$x\neq 0$ に対しては $f(x)=\sin(1/x)$ ならば），f は $x=0$ において連続ではないことを証明せよ．
2. $n=1$ ならば，f は $x=0$ において連続だが微分可能ではないことを証明せよ．
3. $n=2$ ならば，f は $x=0$ において微分可能だが連続微分可能ではないことを証明せよ．

1.7 定義 (**偏微分可能性**)**.** $F:\boldsymbol{R}^2\to\boldsymbol{R}$ の定義域上の点 $(x,y)\in\boldsymbol{R}^2$ を所与として，

$$g(h)=F(x+h,y)$$

によって関数 $g:\boldsymbol{R}\to\boldsymbol{R}$ を定義する． g が $h=0$ において微分可能である時，すなわち極限

$$\lim_{h\to 0}\frac{g(h)-r(0)}{h}=\lim_{h\to 0}\frac{F(x+h,y)-F(x,y)}{h}$$

が存在する時，F は点 (x,y) において変数 x について偏微分可能であると言い，その極限の値を

$$\frac{\partial F}{\partial x}(x,y)$$

と表す． また，F が変数 x について定義域上の任意の点 $(x,y)\in\boldsymbol{R}^2$ において偏微分可能である時，F は変数 x について偏微分可能であると言う．

変数 y についても同様に，極限

$$\lim_{h\to 0}\frac{F(x,y+h)-F(x,y)}{h}$$

が存在する時，F は点 (x,y) において変数 y について偏微分可能であると言い，その収束先の値を，

$$\frac{\partial F}{\partial y}(x,y)$$

と表す． また F が変数 y について定義域上の任意の点 $(x,y)\in\boldsymbol{R}^2$ において偏微分可能である時，F は変数 y について偏微分可能であると言う．

定義 1.7 からも分かるように，偏微分は当該変数以外の変数を定数と見なした時の微分に相当し，1 変数の微小な変化に対する関数 F の変化率を表す．

1.6 例. 資本と労働を投入して財を生産する経済を考え，投入ベクトルが (K, L) から $(K+dk, L+dL)$ へと変化した時の生産量の変化を見てみよう．生産関数を

$$Q = F(K, L)$$

で定義し，F はいずれの変数についても (K, L) において偏微分可能であるとする．この時，資本投入量が K から $K + dK$ へと変化することによる Q の変化分は $\frac{\partial F}{\partial K}(K, L)dK$，労働投入量が L から $L + dL$ へと変化することによる Q の変化分は $\frac{\partial F}{\partial L}(K, L)dL$ であるから，全体としての Q の変化量 dQ は

$$dQ = \frac{\partial F}{\partial K}(K, L)dK + \frac{\partial F}{\partial L}(K, L)dL$$

と書くことがよく見受けられる [4]．

例 1.6 のような偏微分による分析は非常に便利である．しかし偏微分による 1 次近似は，変数の変化に対する F の反応を把握する上で適切な方法であるとは限らない．特に複数の変数が同時に変化する場合，F の変化量を偏微分を用いて近似することには問題が伴う可能性がある．

1.7 例. 2 変数関数 $F(x, y)$ を

$$F(x, y) = \begin{cases} 0 & (\ xy\ =\ 0\ \text{のとき}\ ) \\ \dfrac{x+y}{2} & (\ xy\ \neq\ 0\ \text{のとき}) \end{cases}$$

によって定義し，任意のゼロでない数 h について ベクトル (x, y) が $(0, 0)$ から (h, h) へと変化した時に，F の値がどのように変化するかを考えよう．$F(x, y)$ は各変数について点 $(0, 0)$ において偏微分可能であり，各変数についての偏微分は

$$\frac{\partial F}{\partial x}(0, 0) = \lim_{h \to 0} \frac{F(0 + h, 0) - F(0, 0)}{h} = 0$$

$$\frac{\partial F}{\partial y}(0, 0) = \lim_{h \to 0} \frac{F(0, 0 + h) - F(0, 0)}{h} = 0$$

と計算できる．ここで F の変化量 dF を偏微分によって近似すれば

$$dF = \frac{\partial F}{\partial x}(0, 0)h + \frac{\partial F}{\partial y}(0, 0)h = 0 \tag{1.11}$$

となり，この偏微分による変化の近似を見る限り ベクトルの微小な変化は F の値に影響を与えないように見える．しかし実は，(1.11) による近似は適切ではない．正確な変化量は

$$dF = F(0 + h, 0 + h) - F(0, 0) = h - 0 = h$$

ゆえに

$$\lim_{h \to 0} \frac{F(0 + h, 0 + h) - F(0, 0)}{h} = 1$$

によって求められるから，実際にはこの微小な変化が F の値に長さに比例的な変化をもたらすのである．

[4] 正確な Q の変化量は $dQ = F(K + dk, L + dL) - F(K, L)$ であるため，ここでは偏微係数を用いることで近似的に変化量を求めていることになる．

1.3.2 方向微分

例 1.7 が示すように， ある関数が偏微分可能であっても， 1 変数の変化に対する反応と複数変数の同時変化に対する反応とが大きく異なる場合， 偏微分による分析ではよい近似を与えることができない． 偏微分による近似が持つこのような制約のために， 複数の変数が同時に変化した場合の F の反応を把握するためには，偏微分とはやや異なる概念が必要となる．

1.8 定義 (方向微分). 関数 $F : \boldsymbol{R}^2 \to \boldsymbol{R}$ の定義域上の点 $(x, y) \in \boldsymbol{R}^2$ を所与とし， 変化のベクトル $(u, v) \in \boldsymbol{R}^2$ を用いて関数 $g : \boldsymbol{R} \to \boldsymbol{R}$ を

$$g(\epsilon) = F(x + \epsilon u, y + \epsilon v)$$

によって定義する． g が $\epsilon = 0$ において微分可能である時，すなわち

$$\lim_{\epsilon\to 0} \frac{g(\epsilon) - g(0)}{\epsilon} = \lim_{\epsilon\to 0} \frac{F(x + \epsilon u, y + \epsilon v) - F(x, y)}{\epsilon} \tag{1.12}$$

が存在する時，F は点 (x, y) において (u, v) の方向に微分可能であると言う． またこの時，極限 (1.12) の値を F の点 (x, y) にける方向微分と呼ぶ．

方向微分とは，2 変数がある一定の方向 (u, v) へと微小に変化した時の 関数 F の変化率を表し，偏微分の概念を一般化したものである．実際，定義 1.8 における変化の方向ベクトル $(u, v) \in \boldsymbol{R}^2$ を $(u, v) = (1, 0)$ あるいは $(u, v) = (0, 1)$ と置けば， それぞれ x についての偏微分と y についての偏微分の定義に対応する．

この方向微分の考え方を用いて，先ほどの例 1.7 における関数 F が 2 変数の同時変化に対してどのように反応するかを見てみよう． $(0, 0)$ から (ϵ, ϵ) への変化に対する関数 F の変化率は，$(u, v) = (1, 1)$ 方向への方向微分を考えることに他ならないから

$$\frac{F(0 + \epsilon, 0 + \epsilon) - F(0, 0)}{\epsilon} = \frac{\epsilon - 0}{\epsilon} = 1$$

となる．つまり $(0, 0)$ から $(u, v) = (1, 1)$ 方向へとベクトルが変化する時，関数 F の値は 1 だけ変化するのである．

1.3.3 全微分

方向微分は偏微分を一般化したものであるが， それでも一定の方向に限って変化を考える概念であった． つまり，(1.10) の点列のようにあたかも方向 (u, v) が様々に変化する時の微分可能性には触れていないのである． この部分にまで踏み込んで微分可能性を考えるのが， 次に述べる全微分という概念である． 全微分可能性を定義するにあたって，まずは 1 変数の微分可能性の定義を確認しておこう．

定義（再掲：1 変数関数の微分可能性） 1 変数関数 $f : \boldsymbol{R} \to \boldsymbol{R}$ の定義域上の点 $x \in \boldsymbol{R}$ を所与として，極限

$$\lim_{h\to 0} \frac{f(x + h) - f(x)}{h}$$

が存在する時，f は点 x において微分可能であると言い，その極限の値を

$$\lim_{h\to 0} \frac{f(x + h) - f(x)}{h} = \frac{df}{dx}(x, y) = f'(x)$$

などと表す． また f が定義域上の任意の点 $x \in \boldsymbol{R}$ において微分可能である時，f は微分可能であると言う．

この 1 変数関数の微分可能性の定義 1.6 は，次のように言い換えることができたことを思い出そう． すなわち

1. f は点 x において微分可能である．
2. ある実数 $c \in \boldsymbol{R}$ が存在し，
$$\lim_{h\to 0}\frac{f(x+h)-\bigl(f(x)+ch\bigr)}{|h|}=0 \tag{1.13}$$
が成立する．

は同値であり，この時 (1.13) における c は
$$c=f'(x)$$
と一意に決まるのであった． 言い換えれば，関数 f の微分可能性は (1.13) によっても定義することができるのである． 多変数関数の全微分可能性の定義は， 1 変数関数の場合における (1.13) と同様の形式で与えられる．

1.9 定義 (**全微分可能性**)**.** $F:\boldsymbol{R}^2\to\boldsymbol{R}$ の定義域上の点 $(x,y)\in\boldsymbol{R}^2$ を所与として， あるベクトル $(c,d)\in\boldsymbol{R}^2$ が存在し，
$$\lim_{(h,k)\to(0,0)}\frac{F(x+h,y+k)-\bigl(F(x,y)+ch+dk\bigr)}{\|(h,k)\|}=0 \tag{1.14}$$
が成立する時，F は点 (x,y) において全微分可能であると言う．またこの時 F は (x,y) においていずれの変数についても偏微分可能であり，ベクトル (c,d) は，
$$(c,d)=\left(\frac{\partial F}{\partial x}(x,y),\frac{\partial F}{\partial y}(x,y)\right)$$
と一意に定まる．

定義から明らかなように，F が (x,y) において全微分可能であれば， F は (x,y) において全ての変数について偏微分可能である． しかし，F が全ての変数について偏微分可能であるからといって，F が全微分可能であるとは限らない．

全微分の定義で注意が必要な点は，(1.14) が
$$\lim_{n\to\infty}(h_n,k_n)=(0,0)$$
を満たす任意の点列 $\{(h_n,k_n)\}$ で成立する必要があるということである． 偏微分可能性の条件として想定されていた点列の収束は，例 1.5 において (1.7) や (1.8) に対応するグラフが示すような 水平方向もしくは垂直方向からの収束のみである． 方向微分可能性の条件を満たすためには， 加えて (1.9) に対応するような点列の収束も考慮しなければならなかった． 全微分可能性の条件は，さらに (1.10) のような収束も含めたあらゆる方向からの収束に対して，(1.14) が成立することを求めているのである．

練習問題 1.6. 設問および解答の内容に応じて証明・反例・値等を与えること．

2 変数関数 $F:\boldsymbol{R}^2\to\boldsymbol{R}$ を
$$F(x,y)=\sqrt{|x||y|}.$$
と定義する．

1. $(x,y)=(0,0)$ において F は偏微分可能か？

2. $(x,y)=(0,0)$ において F はどの方向に微分可能（方向微分が存在する）か？

3. $(x,y)=(0,0)$ において F は全微分可能か？

1.3.4 勾配ベクトル

定義 1.9 におけるベクトル (c, d) は (x, y) における勾配ベクトル (gradient vector) と呼ばれ,

$$\nabla F(x,y) = \left(\frac{\partial F}{\partial x}(x,y), \frac{\partial F}{\partial y}(x,y)\right)$$

と表す.

ここで少し，この勾配ベクトルの意味について確認しておこう． 勾配ベクトルが何を表しているかを見るために，次のような最大化問題を考えてみる．

$$\max_{(h,k)} F(x+h, y+k) - F(x,y)$$
$$\text{subject to} \qquad \|(h,k)\| \leq \epsilon$$

これは，(x, y) から $(x + h, y + k)$ へとベクトルを変化させることによって生じる 関数 F の変化量を最大化する問題であるが，ベクトル (x, y) が変化できる領域を点 (x, y) を中心とした半径 ϵ の円内に限定して考えたものである． F は全微分可能であるとしてこの問題を解くと，後で与えられるクーン＝タッカー条件により，$\nabla F \neq 0$ ならば,

$$(h_\epsilon, k_\epsilon) = \frac{\epsilon}{\|\nabla F(x+h_\epsilon, y+k_\epsilon)\|}\nabla F(x+h_\epsilon, y+k_\epsilon) \tag{1.15}$$

を得る．ここで

$$\lim_{\epsilon\to 0} h_\epsilon = 0, \qquad \lim_{\epsilon\to 0} k_\epsilon = 0$$

であるから， (1.15) の両辺を ϵ で割って $\epsilon \to 0$ の極限を取り， なおかつ ∇F が連続であると仮定すると

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}(h_\epsilon, k_\epsilon) = \frac{1}{\|\nabla F(x,y)\|}\nabla F(x,y) \tag{1.16}$$

が導かれる．

(1.16) の左辺は， 各変数を点 (x, y) からごく微小な範囲で変化させることを考えた時に， 関数 F の変化量を最大化する単位ベクトルを表し， 右辺は $\nabla F(x, y)$ 方向への単位ベクトルである． よって (1.16) から， 勾配ベクトルの方向は 関数 F の値を最も急激に増加（あるいは減少）させる変化の方向を示すものであることが分かる．

次に，勾配ベクトルの長さ $\|\nabla F(x, y)\|$ の意味するところを確認しよう． 今 F が (x, y) において全微分可能であると仮定すれば,

$$\lim_{(\epsilon h,\epsilon k)\to(0,0)} \frac{F(x+\epsilon h, y+\epsilon k) - \left(F(x,y) + \frac{\partial F}{\partial x}(x,y)\epsilon h + \frac{\partial F}{\partial y}(x,y)\epsilon k\right)}{\|(\epsilon h, \epsilon k)\|} = 0$$

すなわち,

$$\lim_{\epsilon\to 0}\frac{F(x+\epsilon h, y+\epsilon k) - F(x,y)}{\epsilon\|(h,k)\|} = \frac{\frac{\partial F}{\partial x}(x,y)h + \frac{\partial F}{\partial y}(x,y)k}{\|(h,k)\|}$$

が，任意のベクトル (h, k) で成立する．ここで,

$$(h,k) = \frac{1}{\|\nabla F(x,y)\|}\nabla F(x,y) = \frac{1}{\|\nabla F(x,y)\|}\left(\frac{\partial F}{\partial x}(x,y), \frac{\partial F}{\partial y}(x,y)\right) \tag{1.17}$$

と置けば，$\|(h, k)\| = 1$ であるから,

$$\lim_{\epsilon\to 0}\frac{F(x+\epsilon h, y+\epsilon k) - F(x,y)}{\epsilon} = \|\nabla F(x,y)\| \tag{1.18}$$

となる．(1.18) は (h, k) 方向への方向微分が勾配ベクトルの長さに等しいことを意味している．(1.17) により (h, k) は勾配ベクトルを長さ 1 に正規化したものに選んであったから， 結局 (1.18) からは,勾配ベクトル方向へと各変数が微小に変化した時の関数 F の（最大の）変化率が， 勾配ベクトルの長さだと分かる．

1.4 微分に関する諸定理

1.4.1 偏微分可能性と全微分可能性

ある点において関数が全微分可能であるためには， その点に収束するあらゆる点列を考慮する必要があるため， 全微分可能性をチェックすることは容易ではない．しかしある点において偏微分可能な関数が一定の条件を満たせば， それは同時に全微分可能でもあることが知られている．

1.10 定義. 関数 $F:\boldsymbol{R}^L\to\boldsymbol{R}$ が，$U\subset\boldsymbol{R}^L$ 上で k 階までの全ての偏導関数が存在して連続である時，F は U 上で k 階連続微分可能である，または C^k 級であるという。

1.1 定理. 関数 $F:\boldsymbol{R}^2\to\boldsymbol{R}$ は，$\boldsymbol{R}^2$ で偏微分可能であるとする． この時，F が (x,y) において 1 階連続微分可能であれば， F は (x,y) で全微分可能である．

証明は高木貞治「解析概論」（岩波書店）にある． これは平均値の定理の良い応用と言える。

1.4.2 高階偏導関数

関数 $F:\boldsymbol{R}^2\to\boldsymbol{R}$ は，定義域上の任意の点 (x,y) で 変数 x について偏微分可能であるとする． この時，各点 (x,y) に (x,y) における偏微分

$$\frac{\partial F}{\partial x}(x,t)$$

を対応付ける関数を偏導関数という．

$$\frac{\partial F}{\partial x}:\boldsymbol{R}^2\to\boldsymbol{R} \tag{1.19}$$

がさらに点 (x,y) において偏微分可能なとき， その偏微分を F の 2 階の偏微分といい，それぞれ

$$\frac{\partial^2 F}{\partial^2 x}(x,y)$$
$$\frac{\partial^2 F}{\partial y\partial x}(x,y)$$

などと書く．変数 y に関しても同様で，偏導関数

$$\frac{\partial F}{\partial y}:\boldsymbol{R}^2\to\boldsymbol{R} \tag{1.20}$$

がさらに点 (x,y) において偏微分可能なとき， (1.20) の偏微分をそれぞれ

$$\frac{\partial^2 F}{\partial x\partial y}(x,y)$$
$$\frac{\partial^2 F}{\partial^2 y}(x,y)$$

などと書く．ここで，$\partial^2F/\partial y\partial x$ と $\partial^2F/\partial x\partial y$ の定義の違いに注意してほしい． 前者を得るには，まず x で偏微分し，次に y で偏微分する． 後者を得るには，まず y で偏微分し，次に x で偏微分する． しかしながら，多くの場合両者の値が一致する． その十分条件を与えよう．

1.2 定理 (ヤングの定理). 関数 $F:\boldsymbol{R}^2\to\boldsymbol{R}$ は，定義域上の任意の点 (x,y) で， 変数 x 及び変数 y について偏微分可能であるとする．この時，偏導関数

$$\frac{\partial F}{\partial x}:\boldsymbol{R}^2\to\boldsymbol{R}$$
$$\frac{\partial F}{\partial y}:\boldsymbol{R}^2\to\boldsymbol{R}$$

が, 点 (a,b) においていずれも全微分可能であれば

$$\frac{\partial^2 F}{\partial x \partial y}(a,b) = \frac{\partial^2 F}{\partial y \partial x}(a,b)$$

が成立する.

定理 1.2 の証明. $h,k \in \boldsymbol{R} \diagdown \{0\}$ について,

$$\Delta = F(a+h, b+k) - F(a+h, b) - F(a, b+k) + F(a,b)$$

とおく.

$$\phi(x) = F(x, b+k) - F(x, b)$$

とすると,

$$\Delta = \phi(a+h) - \phi(a)$$

が成立する. また, F は (a,b) で変数 x について偏微分可能なので,

$$\phi'(x) = \frac{\partial F}{\partial x}(x, b+k) - \frac{\partial F}{\partial x}(x, b)$$

そこで, 区間 $(a, a+h)$ に関して, 平均値の定理を $\phi(x)$ に適用すると, ある $\theta \in (0,1)$ が存在して,

$$\frac{\phi(a+h) - \phi(a)}{h} = \phi'(a+\theta h)$$

が成り立つ. つまり,

$$\frac{\Delta}{h} = \frac{\partial F}{\partial x}(a+\theta h, b+k) - \frac{\partial F}{\partial x}(a+\theta h, b)$$

$h=k$ とする. $\partial F / \partial x$ は全微分可能なので

$$\frac{\partial F}{\partial x}(a+\theta h, b+k) = \frac{\partial F}{\partial x}(a,b) + \theta h \frac{\partial^2 F}{\partial x^2}(a,b) + h \frac{\partial^2 F}{\partial y \partial x}(a,b) + o(h)$$
$$\frac{\partial F}{\partial x}(a+\theta h, b) = \frac{\partial F}{\partial x}(a,b) + \theta h \frac{\partial^2 F}{\partial x^2}(a,b) + o(h)$$

である. ただし, $o(\cdot)$ は高次の無限小 $\lim_{h\to 0} o(h)/h = 0$ である. 上式より

$$\frac{\Delta}{h^2} = \frac{\partial^2 F}{\partial y \partial x}(a,b) + \frac{o(h)}{h}$$

つまり

$$\lim_{h\to 0} \frac{\Delta}{h^2} = \frac{\partial^2 F}{\partial y \partial x}(a,b)$$

が成り立つ. 以上の議論を x と y を入れ替えて行うと

$$\lim_{h\to 0} \frac{\Delta}{h^2} = \frac{\partial^2 F}{\partial x \partial y}(a,b)$$

となり題意は示された. □

1.3 定理 (**シュワルツの定理を弱めたもの**). 関数 $F: \boldsymbol{R}^2 \to \boldsymbol{R}$ は, 定義域上の任意の点 (x,y) で, 変数 x 及び変数 y について偏微分可能であるとする. この時, 2 階の偏導関数

$$\frac{\partial^2 F}{\partial x \partial y}: \boldsymbol{R}^2 \to \boldsymbol{R}$$
$$\frac{\partial^2 F}{\partial y \partial x}: \boldsymbol{R}^2 \to \boldsymbol{R}$$

が任意の点 (x,y) で存在し，点 (a,b) においていずれも連続であれば

$$\frac{\partial^2 F}{\partial x \partial y}(a,b) = \frac{\partial^2 F}{\partial y \partial x}(a,b)$$

が成立する．

定理 1.3 の証明. ヤングの定理の証明の前半とまったく同じ議論により

$$\frac{\Delta}{h} = \frac{\partial F}{\partial x}(a+\theta h, b+k) - \frac{\partial F}{\partial x}(a+\theta h, b)$$

である． $\partial^2 F/\partial x \partial y$ が存在するので，左辺を y の 1 変数関数と見て平均値の定理を適用すると，ある $\theta' \in (0,1)$ が存在して，

$$\frac{\Delta}{hk} = \frac{\partial^2 F}{\partial y \partial x}(a+\theta h, b+\theta' k)$$

が成り立つ． $\partial^2 F/\partial x \partial y$ は (a,b) で連続なので

$$\lim_{(h,k)\to(0,0)} \frac{\Delta}{hk} = \frac{\partial^2 F}{\partial y \partial x}(a,b)$$

が成り立つ． 以上の議論を x と y を入れ替えて行うと

$$\lim_{(h,k)\to(0,0)} \frac{\Delta}{hk} = \frac{\partial^2 F}{\partial x \partial y}(a,b)$$

となり題意は示された． □

ヤングの定理とシュワルツの定理はいずれも

$$\frac{\partial^2 F}{\partial x \partial y}(a,b) = \frac{\partial^2 F}{\partial y \partial x}(a,b)$$

が成立するための条件を述べたものであるが， 各偏導関数に対する要請の方向性が異なっている．問題となる点 (a,b) における微分可能性の「深さ」に関して言えば，ヤングの定理がいずれの偏導関数の全微分可能性を要請する一方で， シュワルツの定理はそれを要請しないため， この点ではヤングの定理の方が要請が強い． 一方微分可能領域の「広さ」については， シュワルツの定理が異なる変数同士の 2 階の偏導関数の存在を任意の点 (x,y) において要請するのに対して， ヤングの定理は点 (a,b) においてのみ要請しているため， こちらに関してはシュワルツの定理の方が強い条件を課している． よって，どちらの定理がより一般的であるということはなく，状況に応じて使い分けることになる．

1.4.3 合成関数の微分

まずは 1 変数関数の場合の復習から始めよう．

1.4 定理 (**連鎖律 1**)**.** 関数 $f:\boldsymbol{R}\to\boldsymbol{R}$ と関数 $g:\boldsymbol{R}\to\boldsymbol{R}$ との合成関数，$f\circ g:\boldsymbol{R}\to\boldsymbol{R}$ を考える．F の第 1・第 2 変数に関する偏微分を $\partial F/\partial t_1$，$\partial F/\partial t_2$ と書く．この時， f 及び g がともに微分可能であれば，$f\circ g$ も微分可能となり，

$$\frac{df\bigl(g(x)\bigr)}{dx} = \frac{df(y)}{dy}\frac{dg(x)}{dx}$$

が成立する．

1.5 定理 (連鎖律 2). 関数 $F:\boldsymbol{R^2}\to\boldsymbol{R}$ と関数 $G:\boldsymbol{R}\to\boldsymbol{R^2}$ との合成関数, $F\circ G:\boldsymbol{R}\to\boldsymbol{R}$ を考える．この時, F が全微分可能で G が微分可能であれば, $F\circ G$ も微分可能となり,

$$\frac{dF\big(G_1(x),G_2(x)\big)}{dx}=\frac{\partial F(G_1\left(x\right),G_2\left(x\right))}{\partial t_1}\frac{dG_1(x)}{dx}+\frac{\partial F(G_1\left(x\right),G_2\left(x\right))}{\partial t_2}\frac{dG_2(x)}{dx}$$

が成立する．

1.6 定理 (連鎖律 3). 関数 $F:\boldsymbol{R^2}\to\boldsymbol{R}$ と関数 $G:\boldsymbol{R^2}\to\boldsymbol{R^2}$ との合成関数, $F\circ G:\boldsymbol{R^2}\to\boldsymbol{R}$ を考える．この時, F が全微分可能で G が両変数について全微分可能であれば, $F\circ G$ も全微分可能となり,

$$\frac{\partial F\big(G_1(x_1,x_2),G_2(x_1,x_2)\big)}{\partial x_1}=\frac{\partial F(G_1(x_1,x_2))}{\partial t_1}\frac{\partial G_1(x_1,x_2),}{\partial x_1}+\frac{\partial F(G_1(x_1,x_2))}{\partial t_2}\frac{\partial G_2(x_1,x_2)}{\partial x_1}$$
$$\frac{\partial F\big(G_1(x_1,x_2),G_2(x_1,x_2)\big)}{\partial x_2}=\frac{\partial F(G_1(x_1,x_2))}{\partial t_1}\frac{\partial G_1(x_1,x_2),}{\partial x_2}+\frac{\partial F(G_1(x_1,x_2))}{\partial t_2}\frac{\partial G_2(x_1,x_2)}{\partial x_2}$$

が成立する．

1.4.4 陰関数定理

1.7 定理 (陰関数定理). 関数 $F:\boldsymbol{R}^2\to\boldsymbol{R}$ は連続微分可能であるとし，$c\in\boldsymbol{R}$ 及び $(a,b)\in\boldsymbol{R^2}$ について

$$F(a,b)=c$$
$$\frac{\partial F}{\partial y}(a,b)\neq 0$$

を仮定する．この時，a を含む十分狭い開区間 $I\subset\boldsymbol{R}$ 及び，b を含む十分狭い開区間 $J\subset\boldsymbol{R}$ を取れば,

$$\forall x\in I\quad\forall y\in J:\qquad F(x,y)=c\quad\Leftrightarrow\quad y=g(x)$$

を満たす微分可能な関数 $g:I\to J$ が存在し

$$\forall x\in I\quad\forall y\in J:\qquad \frac{dg}{dx}(x)=-\frac{\frac{\partial F}{\partial x}\big(x,g(x)\big)}{\frac{\partial F}{\partial y}\big(x,g(x)\big)}$$

が成立する． つまり，$F\left(x,y\right)=c$ の解 (x,y) は (a,b) の周囲では，x の関数として表せるということである．

1.8 例. 生産関数 $F(K,L)$ について， ある一定量 c を生産する資本 K と労働 L は

$$F(K,L)=Q \tag{1.21}$$

という関係にある．ここである投入水準 (K,L) から資本投入量を限界的に減少させた時， 生産量 Q を維持するためにはどれだけの追加的労働が必要とされるかを考えよう．これは等量曲線上の点 (K,L) におけるグラフの傾きを求めることに対応する．

一つの方法としては，(1.21) を

$$F(K,L)=Q\quad\Leftrightarrow\quad L=g(K) \tag{1.22}$$

のようにして L について解き，関数 g の微分を取ることが考えられる． しかし (1.22) のような形に書き直せるとは限らないため，この方法は一般性を持たない． このような場合，陰関数定理を用い

ることにより，たとえ g の具体的な関数形が分からなくともその微分 g' の値を求めることができる．この例では

$$\frac{dg}{dK}(K) = -\frac{\frac{\partial F}{\partial K}\big(K, g(K)\big)}{\frac{\partial F}{\partial L}\big(K, g(K)\big)}$$

によって求まる値が，点 (K, L) における等量曲線の傾きである．

練習問題 1.7. 資本と労働から（*1* 種類の）消費財を産出する生産技術を表す生産関数 F を

$$Q = F(K, L) = K^2 + K^5 + L^7 + L^9$$

と定義する．但しここで，K は資本投入量，L は労働投入量，Q は産出量を表す．

1. $\nabla F(1, 1)$ を求めよ．

2. *1* 変数関数 $g(K)$ を $F(K, g(K)) = F(1, 1)$ で定義するとき，$g'(1)$ を求めよ．

3. *1* 変数関数 $h(L)$ を $F(h(L), L) = 38$ で定義するとき，$h'(1)$ を求めよ．

1.4.5 オイラーの定理

1.11 定義 (同次関数). 関数 $F : \boldsymbol{R}^L_{++} \to \boldsymbol{R}$ について，

$$\forall x \in \boldsymbol{R}^L_{++},\ \forall t > 0 : \quad F(tx) = t^m F(x)$$

が成立する時，関数 F は次数 m の同次関数であると言う [5]．

1.9 例. コブ＝ダグラス型の効用関数

$$U(x, y) = x^{1/2} y^{1/2}$$

を考えよう．この関数は任意の $t > 0$ に対して

$$\begin{aligned} U(tx, ty) &= t^{1/2} x^{1/2} t^{1/2} y^{1/2} \\ &= t x^{1/2} y^{1/2} \\ &= tU(x, y) \end{aligned}$$

が成り立つ．よってこの効用関数は 1 次同次関数である．

効用関数の同次性について 1 点だけ補足しておく．ミクロ経済学の文脈において，効用関数に単調変換を加えても選好関係は保存されるが，同次性は失われることがある．

1.10 例. 1.9 例の効用関数 U を対数変換し，新たな効用関数 V を作る．つまり

$$V(x, y) = \log U(x, y) = \frac{1}{2}\log x + \frac{1}{2}\log y$$

とおく．V は U を単調変換したものであるから，選好関係は保存される．しかし

$$\begin{aligned} V(tx, ty) &= \frac{1}{2}\log tx + \frac{1}{2}\log ty \\ &= \frac{1}{2}\log t + \frac{1}{2}\log x + \frac{1}{2}\log t + \frac{1}{2}\log y \\ &= \log t + V(x, y) \end{aligned}$$

であり，これはいかなる m に対しても $t^m V(x, y)$ と一般には異なるので（例えば $x = y = 1$ とおけ），もはや V は 1 次同次関数ではないことがわかる．

[5] ここで x と t が任意であることに注意．ある $x \in \boldsymbol{R}^n_{++}$ とある $t > 0$ について $F(tx) = t^m F(x)$ が成り立ったとしても，F が m 次同次関数であるとは限らない．

練習問題 1.8. 本問では α と β を正定数とする．また，変数 x と y はともに正の値のみとるものとする．以下で定義される 2 変数関数 F が同次関数であるか否かを判定せよ．また，同次関数なら，その次数を求めよ．

1. $F(x,y)=x^{\alpha}y^{\beta}$.
2. $F(x,y)=x^{\alpha}+y^{\beta}$.
3. 関数 $G(x,y)$ は同次関数であるとし，$F(x,y)=(G(x,y))^{\alpha}$.

1.8 定理 (オイラーの定理). 関数 $F:\boldsymbol{R}^n_{++}\to\boldsymbol{R}$ が全微分可能で m 次同次関数であれば，

$$\forall x\in\boldsymbol{R}^n_{++}:\quad x\cdot\nabla F(x)=mF(x)$$

すなわち，

$$\forall x=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\in\boldsymbol{R}^n_{++}:\quad x_1\frac{\partial F}{\partial x_1}(x)+x_2\frac{\partial F}{\partial x_2}(x)+\cdots+x_n\frac{\partial F}{\partial x_n}(x)=mF(x_1,x_2,\dots,x_n)$$

が成立する [6].

定理 1.8 の証明. F が m 次同次関数なので，$t>0$ に対して

$$F(tx)=t^mF(x)$$

が成り立つ． 両辺を t で微分すると

$$x_1\frac{\partial F}{\partial x_1}(tx)+x_2\frac{\partial F}{\partial x_2}(tx)+\cdots+x_n\frac{\partial F}{\partial x_n}(tx)=mt^{m-1}F(x)$$

である． $t=1$ のときは

$$x_1\frac{\partial F}{\partial x_1}(x)+x_2\frac{\partial F}{\partial x_2}(x)+\cdots+x_n\frac{\partial F}{\partial x_n}(x)=mF(x_1,x_2,\dots,x_n)$$

となる． □

1.11 例. 生産関数 $Q=F(K,L)$ によって生産を行う企業を考えよう． F が 1 次同次関数であるとすれば，オイラーの定理により，

$$K\frac{\partial F}{\partial K}(K,L)+L\frac{\partial F}{\partial L}(K,L)=Q \tag{1.23}$$

が任意の (K,L) で成立する．ここで企業が選ぶ資本と労働の投入量が (K^*,L^*)，その結果得られる生産量が (Q^*) であったとしよう． この時 (1.23) は

$$K^*\frac{\partial F}{\partial K}(K^*,L^*)+L^*\frac{\partial F}{\partial L}(K^*,L^*)=Q^*$$

と書けるから，両辺に生産財価格 p を掛ければ

$$K^*p\frac{\partial F}{\partial K^*}(K^*,L^*)+L^*p\frac{\partial F}{\partial L}(K^*,L^*)=pQ^* \tag{1.24}$$

[6] 実はオイラーの定理の逆も言うことができる． よって

$$F\text{ は }m\text{ 次同次関数}\quad\Leftrightarrow\quad\forall x\in\boldsymbol{R}^n_{++}:\quad x\cdot\nabla F(x)=mF(x)$$

が成立する．

が成立することが分かる． もし企業が資本レンタル料・賃金率・生産財が一定であると想定する（価格受容的行動）ならば，利潤最大化の1階の条件により， $p\frac{\partial F}{\partial K}(K^*, L^*)$ は資本のレンタル料に，$p\frac{\partial F}{\partial L}(K^*, L^*)$ は賃金率に等しくなるので， 左辺は均衡における企業の総支出を表す． 一方の右辺は均衡における総収入であるから， (1.24) は 完全競争市場においては均衡における企業の利潤がゼロとなることを意味している．

1.5 テイラー展開

1.5.1 1変数関数のテイラー展開

1変数関数 $f : \boldsymbol{R} \to \boldsymbol{R}$ について $f(x)$ の値は分かっているとし，変数 x を $x+h$ へと変化させた時の関数 f の値 $f(x+h)$ を求めることを考える．

1.9 定理 ((テイラーの定理))．もし $f : \boldsymbol{R} \to \boldsymbol{R}$ が x において f が $k+1$ 回連続微分可能であれば，ある $\theta \in [0,1]$ が存在して

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2!}f''^2 + \cdots + \frac{1}{k!}f^{(k)}(x)h^k + R_{k+1}(h;x) \tag{1.25}$$
$$R_k(h;x) = \frac{1}{(k+1)!}f^{(k+1)}(x+\theta h)h^{k+1}$$

が成立する． これを f の x における k 次のテイラー展開という．ここで，$f^{(k)}(x)$ は f の k 階導関数を点 x で評価したものである． また右辺の最終項は剰余項と呼ばれるもので，

$$\lim_{h \to 0} \frac{R_{k+1}(h;x)}{h^k} = 0$$

が成り立つ．

よって h が十分に小さい時（したがって $x+h$ が x に十分に近い時）には

$$f(x+h) \approx f(x) + f'(x)h + f''^2 + \cdots + f^{(k)}(x)h^k$$

のように，$f(x+h)$ の値を，点 x における f の導関数の値を係数とした h の多項式によって近似できる．例えば $f(x+h)$ の1次近似（$k=1$ の場合）は

$$f(x+h) \approx f(x) + f'(x)h$$

であり，2次近似（$k=2$ の場合）は

$$f(x+h) \approx f(x) + f'(x)h + \frac{1}{2}f''^2$$

である． なお，$k=0$ の場合は

$$f(x+h) = f(x) + f'(x+\theta h)\,h$$

つまり

$$\frac{f(x+h) - f(x)}{h} = f'(x+\theta h)$$

となるので，テイラーの定理は平均値の定理の拡張であると言える．

1.5.2 多変数関数のテイラー展開

テイラー展開を多変数関数へと拡張しよう． n 変数関数 $F:\boldsymbol{R}^n\to\boldsymbol{R}$ について $F(x)$ の値は分かっているとし， ベクトル x を $x+h$ へと変化させた時の関数 F の値 $F(x+h)$ を求めることを考える．この時 x において F が $k+1$ 回連続微分可能であれば

$$\begin{aligned}F(x+h)=F(x)&+\sum_{i=1}^{n}\frac{\partial F(x)}{\partial x_i}h_i+\frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2F(x)}{\partial x_i\partial x_j}h_ih_j+\dots\\&+\frac{1}{k!}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_k=1}^{n}\frac{\partial^kF(x)}{\partial x_{i_1}\partial x_{i_2}\dots\partial x_{i_k}}h_{i_1}h_{i_2}\dots h_{i_k}+R_{k+1}(h;x)\end{aligned}$$

1 変数の場合と同様に右辺の最終項は剰余項と呼ばれるもので， ある $\theta\in[0,1]$ を用いて

$$R_k(h;x)=\frac{1}{(k+1)!}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_{k+1}=1}^{n}\frac{\partial^{k+1}F(x+\theta h)}{\partial x_{i_1}\partial x_{i_2}\dots\partial x_{i_{k+1}}}h_{i_1}h_{i_2}\dots h_{i_{k+1}}$$

と書け,

$$\lim_{h\to0}\frac{R_{k+1}(h;x)}{\|h\|^k}=0$$

が成り立つ． また多変数関数についても近似を考えることができ，例えば $F(x+h)$ の 1 次近似は

$$F(x+h)\approx F(x)+\nabla F(x)^\top h$$

2 次近似は

$$F(x+h)\approx F(x)+\nabla F(x)^\top h+\frac{1}{2}h^\top\nabla^2F(x)h$$

である． 以下，n 次項（$n\geq3$）は「多重線形関数」を用いて定義される．

練習問題 1.9. *2* 変数関数 F の *2* 階偏微分係数

$$\left|\frac{\partial^2F(x,y)}{\partial x^2}\right|,\left|\frac{\partial^2F(x,y)}{\partial x\partial y}\right|,\left|\frac{\partial^2F(x,y)}{\partial y\partial x}\right|,\left|\frac{\partial^2F(x,y)}{\partial y^2}\right|$$

は，どの $(x,y)\in\boldsymbol{R}^2$ においても，ある正定数 c を上回らないものとする. また，$|h|\leq1/3$ および $|k|\leq1/3$ とする. このとき，F の (x,y) における *1* 階の微分までのテイラー展開による $F(x+h,y+k)$ の近似の剰余項の値の上限を c を使って表せ.

2 線形代数

2.1 行列

2.1.1 線形写像と行列

m 行 n 列の行列 $A \in \boldsymbol{R}^{m\times n}$ は

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{m1} & \cdots\cdots & & a_{mn} \end{pmatrix}$$

によって表されるが，これを列ごとにまとめて

$$A = (a_1 \ a_2 \ \ldots \ a_n), \qquad a_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{m1} \end{pmatrix} \in \boldsymbol{R}^m, \quad j = 1, 2, 3, \ldots, n$$

と書くこともある．

この行列 A に n 次元ベクトル

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \boldsymbol{R}^n,$$

を右からかければ

$$Ax = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n \end{pmatrix} = x_1a_1 + x_2a_2 + \cdots + x_na_n \in \boldsymbol{R}^m$$

となる．これは $\boldsymbol{R}^n$ 上の点 x と $\boldsymbol{R}^m$ 上の点 Ax とを対応付ける写像

$$x \ \mapsto \ Ax$$

と見なすことができる．行列を一つ定めることは，一つの線形写像を定めることに他ならない．

2.1 定義 (**線形写像**)**.** 写像 $F : \boldsymbol{R}^n \to \boldsymbol{R}^m$ について

$$\forall x \in \boldsymbol{R}^n, \quad \forall y \in \boldsymbol{R}^n : \qquad F(x + y) = F(x) + F(y)$$
$$\forall x \in \boldsymbol{R}^n, \quad \forall t \in \boldsymbol{R} : \qquad F(tx) = tF(x)$$

が成り立つ時，F は線形写像であると言う．

2.1 定理. 任意の行列 $A \in \boldsymbol{R}^{m\times n}$ に対して，写像 $F : \boldsymbol{R}^n \to \boldsymbol{R}^m$ を，

$$F(x) = Ax$$

と定めれば，F は線形写像である．

逆に任意の線形写像 $F:\boldsymbol{R}^n \to \boldsymbol{R}^m$ に対して

$$\forall x \in \boldsymbol{R}^n, \quad F(x) = Ax$$

を満たす行列 $A \in \boldsymbol{R}^{m\times n}$ がただ一つ存在する．

2.1 **定理の証明.** 第 j 成分のみが 1 で，他はすべて 0 である $\boldsymbol{R}^n$ のベクトルを，e_j $(j=1,2,3,\dots,n)$ で表そう．すると F の定義域に含まれる任意の点 $x \in \boldsymbol{R}^n$ は

$$\begin{aligned} x &= \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \\ &= x_1e_1 + x_2e_2 + \cdots + x_ne_n \end{aligned}$$

のように，ベクトル $e_1, e_2, e_3, \dots, e_n$ の線形結合で表すことができる．$e_1, e_2, e_3, \dots, e_n$ は F の定義域に含まれるので，F の線形性により

$$\begin{aligned} F(x) &= F(x_1e_1 + x_2e_2 + \cdots + x_ne_n) \\ &= x_1F(e_1) + x_2F(e_2) + \cdots + x_nF(e_n) \end{aligned}$$

が成立する．ここで

$$F(e_j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{m1} \end{pmatrix} = a_j \quad j = 1,2,3,\dots,n, \qquad A = (a_1a_2\dots a_n)$$

と置けば,

$$\begin{aligned} F(x) &= x_1a_1 + x_2a_2 + \cdots + x_na_n \\ &= (a_1a_2\dots a_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= Ax \end{aligned}$$

と書くことができる．以上から，任意の線形写像 F に対して行列 $A \in \boldsymbol{R}^{m\times n}$ が存在し，$F(x) = Ax$ と表せることが示された． □

練習問題 2.1. 上の証明で存在を確認した行列 A が，関数 F に対して一意であることを示せ．もし $F(x) = Bx$ を満たすような $B \in \boldsymbol{R}^{m\times n}$ が存在するなら，それは必ず A と等しい（$A = B$ である）ことを示せばよい.

定理 2.1 および定理 2.1 により，行列を一つ定めることと一つの線形写像を定めることは同値であるとが言える．よって行列の性質を考えることは，それに対応する線形写像の性質を考えることに他ならない.

2.1.2 連立 1 次方程式と行列

連立 1 次方程式は，行列を用いて表現することができる．例えば

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n = b_m \end{cases}$$

は，行列

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{m1} & \dots\dots & & a_{mn} \end{pmatrix}$$

を用いることで

$$Ax = b$$

と簡潔に表すことができる．ただしここで

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \boldsymbol{R}^n, \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \boldsymbol{R}^m$$

である．また, $x_1a_1 + x_2a_2 + \cdots + x_na_n = b$ と書くことができる．

またこの時，b が行列 A の列ベクトルの 1 次結合で表せることと，対応する連立 1 次方程式に解が存在することは等しい．つまり，

1. 連立 1 次方程式 $Ax = b$ を満たす x が存在する．
2. b は行列 A の列ベクトルの 1 次結合で表すことができる．

は同値である．

2.1.3 ベクトルの 1 次独立性

2.2 定義 (**1 次独立性**). n 個の m 次元ベクトル $a_1, a_2, \dots, a_n$ について，

$$\forall x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \boldsymbol{R}^n : \quad x_1a_1 + x_2a_2 + \cdots + x_na_n = 0 \quad \Rightarrow \quad x = 0$$

が成り立つ時，$a_1, a_2, \dots, a_n$ は 1 次独立（線形独立）であると言う．また, 1 次独立でないとき, つまり,

$$\exists x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \boldsymbol{R}^n \setminus \{0\} : \quad x_1a_1 + x_2a_2 + \cdots + x_na_n = 0$$

が成り立つ時，$a_1, a_2, \dots, a_n$ は 1 次従属（線形従属）であると言う．

$a_1, a_2, \dots, a_n$ が 1 次従属である時，いずれかのベクトルが他のベクトルの 1 次結合で表される．定義 2.2 により

$$\exists x \in \boldsymbol{R}^n,\ \exists k,\ x_k \neq 0: \quad x_1 a_1 + x_2 a_2 + \dots + x_k a_k + \dots x_n a_n = 0$$

であるから

$$a_k = -\frac{x_1}{x_k} a_1 - \frac{x_2}{x_k} a_2 - \dots - \frac{x_{k-1}}{x_k} a_{k-1} - \frac{x_{k+1}}{x_k} a_{k+1} \dots - \frac{x_n}{x_k} a_n$$

と書ける．

2.1 命題. 1 次独立なベクトルは，すべて互いに異なる非ゼロベクトルである．つまり $a_1, a_2, \dots, a_n$ が 1 次独立であれば

$$\forall j: \quad a_j \neq 0$$

かつ

$$\forall j, k: \quad j \neq k \quad \to a_j \neq a_k$$

が成り立つ．逆に同じベクトルやゼロベクトルが含まれていれば，それらのベクトルは 1 次従属である．つまり

$$\exists j: \quad a_j = 0$$

または

$$\exists j, k,\ j \neq k: \quad a_j = a_k$$

であれば，$a_1, a_2, \dots, a_n$ は 1 次従属である．

2.2 定理. 1 次独立なベクトルの集合の部分集合は 1 次独立である．すなわち，$a_1, a_2, \dots, a_n$ が 1 次独立なベクトルの組であれば，その部分集合 $a_{i_1}, a_{i_2}, \dots, a_{i_k} (k \leq n)$ も 1 次独立である．

また対偶であるが，1 次従属なベクトルの集合に新たなベクトルを追加した集合は 1 次従属である．つまり $a_1, a_2, \dots, a_n$ が 1 次従属なベクトルの組であれば，これに k 個のベクトルを加えた集合 $a_1, a_2, \dots, a_n, a_{n+1}, \dots, a_{n+k}$ も 1 次従属である．

練習問題 2.2. n を正の整数，α をゼロではない実数とし，関数 $F: \boldsymbol{R}^n_{++} \to \boldsymbol{R}$ を

$$F(x) = (\min\{x_1, x_2, \dots, x_n\})^\alpha$$

と定義する．

1. F は同次関数であることを証明せよ．その際，同次性の次数を求めよ．
2. 任意の $x \in \boldsymbol{R}^n_{++}$ に対し，x において x の方向に関して（方向）微分可能であることを証明せよ．
3. 任意の $x \in \boldsymbol{R}^n_{++}$ に対し，もし $x_1 = x_2 = \dots = x_n$ ならば，いずれの変数に関しても偏微分可能ではないことを証明せよ．

練習問題 2.3. n と m を正の整数とし，$n > m$ を仮定するとき，n の未知数 $x_1, \dots, x_n$ と m の方程式より成る連立 1 次方程式

$$\begin{cases} a_{11} x_1 + \dots + a_{1n} x_n = 0, \\ \qquad \vdots \\ a_{m1} x_1 + \dots + a_{mn} x_n = 0 \end{cases}$$

は自明でない解（つまり，$x_j \neq 0$ となる j が少なくともひとつ存在するような解）$(x_1, \dots, x_n)$ が存在することを，m に関する帰納法を使って証明せよ．

2.3 定理. $a_1, a_2, \ldots, a_n$ が $\boldsymbol{R}^m$ の 1 次独立なベクトルの組であれば，$n \leq m$ である．

2.3 **定理の証明.** $n > m$ とする．$a_1, \ldots, a_n$ は 1 次独立なベクトルの組であるので，

$$\alpha_1 a_1 + \cdots + \alpha_n a_n = 0$$

が成り立つならば, $\alpha_1 = \cdots = \alpha_n = 0$ である．
　ところが,

$$\begin{cases} a_{11}\alpha_1 + \cdots + a_{1n}\alpha_n = 0, \\ \qquad\qquad \vdots \\ a_{m1}\alpha_1 + \cdots + a_{mn}\alpha_n = 0 \end{cases}$$

は少なくとも一つ自明でない解を持つ．つまり, $a_1, \ldots, a_n$ の n 個のベクトルのうち少なくとも一つのベクトルが他のベクトルの線形結合として表される．これは, $a_1, \ldots, a_n$ が 1 次独立なベクトルの組であることに矛盾する．従って, $n \leq m$. □

2.1.4　線形部分空間の基底

2.3 定義 (線形部分空間). $\boldsymbol{R}^m$ の部分集合 V について，

$$\forall v \in V,\ \forall w \in V: \quad v + w \in V$$
$$\forall v \in V,\ \forall t \in \boldsymbol{R}: \quad tv \in V$$

が成り立つ時，V は $\boldsymbol{R}^m$ の線形部分空間であると言う．

　例:

- $\boldsymbol{R}^2$ の原点を通る直線
- $\boldsymbol{R}^3$ の原点を通る平面

2.4 定理. n 個の m 次元ベクトル $a_1, a_2, \ldots, a_n$ の 1 次結合によって表されるベクトルの集合を，

$$\langle a_1, a_2, \ldots, a_n \rangle \subset \boldsymbol{R}^m$$

で表すことにしよう．この時，$\langle a_1, a_2, \ldots, a_n \rangle$ は $\boldsymbol{R}^m$ の線形部分空間である．

2.4 定義 (線形部分空間の基底). $\boldsymbol{R}^m$ の線形部分空間 V と，V の n 個のベクトル $a_1, a_2, \ldots, a_n$ について，

1. $a_1, a_2, \ldots, a_n$ は V を張る．すなわち，$\langle a_1, a_2, \ldots, a_n \rangle = V$
2. $a_1, a_2, \ldots, a_n$ は 1 次独立.

が同時に成り立つ時，$a_1, a_2, \ldots, a_n$ は V の基底であるという．

　基底とは，その部分空間に含まれる任意の要素をその線形結合で表すための過不足ないベクトルの集合である．任意の線形部分空間には少なくとも 1 つの基底が存在するが，基底は一意ではない．

2.1 例. $V = \boldsymbol{R}^2$ とした時の V の基底を考えてみよう．まず

$$a_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

と置けば， a_1, a_2 が 1 次独立なベクトルの組であることは明らかであり，また

$$\forall x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in V: \quad x = x_1 a_1 + x_2 a_2$$

と表せるから，$\langle a_1, a_2 \rangle = V$．よって a_1, a_2 は V の基底である．

一方，

$$a_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

と置いても a_1, a_2 は V の基底となる．というのも，a_1, a_2 が 1 次独立なベクトルの組であり，また

$$\forall x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in V: \quad x = (x_1 - x_2) a_1 + x_2 a_2$$

と表せ，$\langle a_1, a_2 \rangle = V$ を満たすからである．

2.5 定理. 線形部分空間の基底に含まれるベクトルの個数は，基底の取り方によらず一定である．つまり，$a_1, \ldots, a_k$ 及び $b_1, \ldots, b_l$ がともに線形部分空間 V の基底である時，$k = l$ が成立する．

2.5 **定理の証明.** $k > l$ とする．$a_1, \ldots, a_k$ 及び $b_1, \ldots, b_l$ はともに線形部分空間 V の基底であることから, 基底 $a_1, \ldots, a_k$ の各ベクトルは $b_1, \ldots, b_l$ の線形結合によって表される．すなわち,

$$\begin{cases} a_1 = \beta_1^1 b_1 + \cdots + \beta_l^1 b_l, \\ \qquad \vdots \\ a_k = \beta_1^k b_1 + \cdots + \beta_l^k b_l. \end{cases} \tag{$*$}$$

$a_1, \ldots, a_k$ は 1 次独立なベクトルの組であるので,

$$\alpha_1 a_1 + \cdots + \alpha_k a_k = 0 \tag{$**$}$$

が成り立つならば, $\alpha_1 = \cdots = \alpha_k = 0$ である．ここで $(*)$ に $(**)$ を代入し, $b_1, \ldots, b_l$ もまた 1 次独立なベクトルの組であることを用いて整理すると,

$$\begin{cases} \beta_1^1 \alpha_1 + \cdots + \beta_1^k \alpha_k = 0, \\ \qquad \vdots \\ \beta_l^1 \alpha_1 + \cdots + \beta_l^k \alpha_k = 0. \end{cases} \tag{$***$}$$

$k > l$ であるので, $(***)$ は少なくとも一つ自明でない解を持つ．これは, $a_1, \ldots, a_k$ が 1 次独立なベクトルの組であることに矛盾する．従って, $k \leq l$. 同様にして, $k \geq l$ であることも示すことができる．従って, $k = l$. □

2.5 定義 (**線形部分空間の次元**). 線形部分空間 V に対して，その基底に含まれるベクトルの個数を V の次元といい，$\dim V$ で表す.

2.6 定理. 線形写像 $F: \boldsymbol{R}^n \to \boldsymbol{R}^m$ を行列 $A = (a_1 \ a_2 \ \ldots \ a_n) \in \boldsymbol{R}^{m \times n}$ を用いて $F(x) = Ax$ で定義する．この時

1. F が全射 $\quad \Leftrightarrow \quad \langle a_1, a_2, \ldots, a_n \rangle = \boldsymbol{R}^n$

2. F が単射 $\quad \Leftrightarrow \quad a_1, a_2, \ldots, a_n$ は 1 次独立

3. F が全単射 $\quad \Leftrightarrow \quad a_1, a_2, \ldots, a_n$ は $\boldsymbol{R}^n$ の基底

が成立する．

2.6 定義 (列空間). $m \times n$ 行列

$$A = (a_1 \ a_2 \ \dots \ a_n), \qquad a_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{m1} \end{pmatrix} \in \boldsymbol{R}^m, \quad j = 1, 2, 3, \dots, n$$

について，A の列ベクトルが張る $\boldsymbol{R^m}$ の線形部分空間を A の列空間といい $\mathrm{Col}A$ と表す．つまり，

$$\mathrm{Col}A = \langle a_1, a_2, \dots, a_n \rangle$$

2.7 定義 (行空間). $m \times n$ 行列

$$A = \begin{pmatrix} a^1 \\ a^2 \\ \vdots \\ a^m \end{pmatrix}, \qquad a^i = (a_{i1} \ a_{i2} \ \dots \ a_{in}) \in \boldsymbol{R}^n, \quad i = 1, 2, 3, \dots, m$$

について，A の行ベクトルが張る空間を A の行空間といい，$\mathrm{Row}A$ と表す．つまり，

$$\mathrm{Row}A = \langle a^1, a^2, \dots, a^m \rangle$$

2.8 定義 (核). m 本からなる n 変数の連立 1 次方程式 $Ax = 0$ を満たす解 $x \in \boldsymbol{R}^n$ の集合を行列 A の核といい，$\mathrm{Ker}A$ で表す．つまり

$$\mathrm{Ker}A = \left\{ x \in \boldsymbol{R}^n \,\middle|\, Ax = 0 \right\}$$

KerA は $\boldsymbol{R}^n$ の線形部分空間である．

2.7 定理. 任意の $m \times n$ 行列 $A \in \boldsymbol{R}^{m \times n}$ について

$$\dim(\mathrm{Col}A) + \dim(\mathrm{Ker}A) = n$$

が成り立つ．

2.7 **定理の証明.** $v_1, v_2, \dots, v_l$ が $\mathrm{Ker}A$ の基底の組であるとすると，次元の定義により

$$\dim(\mathrm{Ker}A) = l$$

である．

まず, $\mathrm{Ker}A$ と $\boldsymbol{R}^n$ が等しい場合，すなわち $\langle v_1, v_2, \dots, v_l \rangle = \boldsymbol{R}^n$ の場合を考えよう．この時 $l = n$ で，任意の $x \in \boldsymbol{R}^n$ で $Ax = 0$ が成立するので $A = 0$ でなければならず

$$\dim(\mathrm{Col}A) = 0$$

よって，

$$\dim(\mathrm{Col}A) + \dim(\mathrm{Ker}A) = 0 + n = n$$

が成立することを確認できる．次に，$\langle v_1, v_2, \dots, v_l \rangle \neq \boldsymbol{R}^n$ の場合を考えよう．この時，$v_1, v_2, \dots, v_l$ の線形結合として表せないベクトルが存在する．これを v_{l+1} とすると, $v_1, \dots, v_l, v_{l+1}$ もまた 1 次独

立となる．同様の操作を高々$n-l$回繰り返して，適当な$(n-l)$個のベクトル$v_{l+1},\ldots,v_n$を加えることで$\boldsymbol{R}^n$の基底を作ることができる．つまり

$$\left\langle v_1,v_2,\ldots,v_l,v_{l+1},v_{l+2},\ldots,v_n\right\rangle=\boldsymbol{R}^n$$

となる．これから$Av_{l+1},\ldots,Av_n$がCol$A$の基底であることを示す．
ここで$y\in$Col$A$とすれば，yはAの列ベクトル$a_1,a_2,\ldots,a_n$の線形結合で表せるから

$$\exists x\in\boldsymbol{R}^n:\quad y=Ax$$

と書ける．また$v_1,v_2,\ldots,v_l,v_{l+1},v_{l+2},\ldots,v_n$は$\boldsymbol{R}^n$の基底であるから，この$x$は

$$\exists z=\begin{pmatrix}z_1\\z_2\\\vdots\\z_n\end{pmatrix}\in\boldsymbol{R}^n:\quad x=z_1v_1+z_2v_2+\cdots+z_lv_l+z_{l+1}v_{l+1}+\cdots+z_nv_n$$

と書ける．よって

$$\begin{aligned}y&=A\left(z_1v_1+z_2v_2+\cdots+z_lv_l+z_{l+1}v_{l+1}+\cdots+z_nv_n\right)\\&=z_1Av_1+z_2Av_2+\cdots+z_lAv_l+z_{l+1}Av_{l+1}+\cdots+z_nAv_n\end{aligned}$$

ここで，$v_1,v_2,\ldots,v_l$はKer$A$の基底であったから

$$Av_1=Av_2=\cdots=Av_l=0$$

よって

$$y=z_{l+1}Av_{l+1}+\cdots+z_nAv_n$$

のように，yは$(n-l)$個のベクトル$Av_{l+1},Av_{l+2},\ldots,Av_n$の線形結合で表される．これにより

$$y\in\mathrm{Col}A\quad\Rightarrow\quad y\in\left\langle Av_{l+1},Av_{l+2},\ldots,Av_n\right\rangle$$

すなわち

$$\mathrm{Col}A\subseteq\left\langle Av_{l+1},Av_{l+2},\cdots,Av_n\right\rangle$$

が言える．逆に$y\in\left\langle Av_{l+1},Av_{l+2},\ldots,Av_n\right\rangle$とすれば，

$$\exists z=\begin{pmatrix}c_{l+1}\\c_{l+2}\\\vdots\\c_n\end{pmatrix}\in\boldsymbol{R}^{n-l}:\quad y=c_{l+1}Av_{l+1}+c_{l+2}Av_{l+2}+\cdots+c_nAv_n$$

$$\therefore y=A(c_{l+1}v_{l+1}+c_{l+2}v_{l+2}+\cdots+c_nv_n).$$

よって

$$y\in\left\langle Av_{l+1},Av_{l+2},\ldots,Av_n\right\rangle\quad\Rightarrow\quad y\in\mathrm{Col}A$$

つまり，

$$\left\langle Av_{l+1},Av_{l+2},\cdots,Av_n\right\rangle\subseteq\mathrm{Col}A.$$

である．以上から

$$\left\langle Av_{l+1},Av_{l+2},\ldots,Av_n\right\rangle=\mathrm{Col}A$$

次に

$$z_{l+1}Av_{l+1} + z_{l+2}Av_{l+2} + \cdots + z_nAv_n = 0$$

を仮定すると

$$A\left(z_{l+1}v_{l+1} + z_{l+2}v_{l+2} + \cdots + z_nv_n\right) = 0$$

が成り立つから

$$z_{l+1}v_{l+1} + z_{l+2}v_{l+2} + \cdots + z_nv_n \in \mathrm{Ker}A$$

よって

$$\exists z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_l \end{pmatrix} \in \boldsymbol{R}^l : \quad z_{l+1}v_{l+1} + z_{l+2}v_{l+2} + \cdots + z_nv_n = z_1v_1 + z_2v_2 + \cdots + z_lv_l$$

すなわち

$$z_1v_1 + z_2v_2 + \cdots + z_lv_l - z_{l+1}v_{l+1} - z_{l+2}v_{l+2} - \cdots - z_nv_n = 0$$

と書ける．$v_1, v_2, \ldots, v_l, v_{l+1}, v_{l+2}, \ldots, v_n$ は $\boldsymbol{R}^n$ の基底であるから 1 次独立で，この時

$$z_1 = z_2 = \cdots = z_l = -z_{l+1} = -z_{l+2} = \cdots = -z_n = 0$$

が成立する．よって

$$z_{l+1}Av_{l+1} + z_{l+2}Av_{l+2} + \cdots + z_nAv_n = 0 \quad \Rightarrow \quad z_{l+1} = z_{l+2} = \cdots = z_n = 0$$

が言えるから，$Av_{l+1}, Av_{l+2}, \ldots, Av_n$ は 1 次独立なベクトルの組であることが分かる．以上から，$(n-l)$ 個のベクトル $Av_{l+1}, Av_{l+2}, \ldots, Av_n$ が $\mathrm{Col}A$ の基底なので

$$\dim\left(\mathrm{Col}A\right) = n - l$$

また

$$\dim\left(\mathrm{Ker}A\right) = l$$

であったから

$$\dim\left(\mathrm{Col}A\right) + \dim\left(\mathrm{Ker}A\right) = n - l + l = n$$

が確認できる． □

2.9 定義 (**直交補空間**)．$\boldsymbol{R}^n$ の線形部分空間 V について，V の全てのベクトルと直交するような n 次元ベクトルの集合

$$V^{\perp} = \left\{x \in \boldsymbol{R}^n \;\middle|\; x \cdot z = 0,\ z \in V\right\}$$

を V の直交補空間という．これは $\boldsymbol{R}^n$ の線形部分空間である．

2.8 定理． $\boldsymbol{R}^n$ の線形部分空間 V について

$$\dim V + \dim V^{\perp} = n$$

が成り立つ．

2.8 **定理の証明.** まず, V の基底 $v_1, \ldots, v_r$ から, すべての i と j について, $e_i \cdot e_i = 1, e_i \cdot e_j = 0\ (i \neq j)$ を満たすような基底 $e_1, \ldots, e_r$ を作る．(シュミットの直交化法.[7])

$\boldsymbol{R}^n$ の任意のベクトル x に対して,

$$x_1 = \sum_{i=1}^{r} (x \cdot e_i) e_i$$

と置けば, $x_1 \in V$ である．$x_2 = x - x_1$ と置くと, x_2 は $e_1, \ldots, e_r$ と直交するので, $x_2 \in V^\perp$ である．よって, $\boldsymbol{R}^n$ は V と $V^\perp$ との和である．また, $V \cap V^\perp = \{0\}$ であるので, $\boldsymbol{R}^n$ は V と $V^\perp$ との直和である．従って, $\dim \boldsymbol{R}^n = \dim V + \dim V^\perp$. □

2.9 定理. 任意の $m \times n$ 行列 $A \in \boldsymbol{R}^{m \times n}$ について

$$\dim(\mathrm{Col}A) = \dim(\mathrm{Row}A)$$

が成り立つ．これは $\dim\ (\mathrm{Col}A) = \dim\ (\mathrm{Col}A^\top)$, $\dim\ (\mathrm{Row}A) = \dim\ (\mathrm{Row}A^\top)$ などと書いても同じことである。

2.9 **定理の証明.** x が $Ax = 0$ の解であることは

$$\begin{aligned} x \in \mathrm{Ker}A \quad &\Leftrightarrow \quad a^1 \cdot x = a^2 \cdot x = \cdots = a^m \cdot x = 0 \\ &\Leftrightarrow \quad \forall z \in \mathrm{Row}A:\ z \cdot x = 0 \\ &\Leftrightarrow \quad x \in (\mathrm{Row}A)^\perp \end{aligned}$$

のように言い換えることができるから

$$\mathrm{Ker}A = (\mathrm{Row}A)^\perp$$

が成立する．よって

$$\dim(\mathrm{Ker}A) = \dim\left((\mathrm{Row}A)^\perp\right)$$

ここで，定理 2.7 と定理 2.8 により

$$\begin{aligned} \dim(\mathrm{Ker}A) &= n - \dim(\mathrm{Col}A) \\ \dim\left((\mathrm{Row}A)^\perp\right) &= n - \dim(\mathrm{Row}A) \end{aligned}$$

であるから結局

$$\dim(\mathrm{Col}A) = \dim(\mathrm{Row}A)$$

が成立する． □

定理 2.7 を使えば, 逆に定理 2.8 が定理 2.9 から導出されることも示される．

定理 2.8 および定理 2.9 から，任意の $m \times n$ 行列 $A \in \boldsymbol{R}^{m \times n}$ について

$$\dim(\mathrm{Ker}A) = n - \dim(\mathrm{Row}A) \tag{2.1}$$

が成り立つことが分かる．連立 1 次方程式 $Ax = 0$ において $\dim(\mathrm{Ker}A)$ は解の自由度，n は未知数の数，$\dim(\mathrm{Row}A)$ は実質的な方程式の本数を表すから，[8] (2.1) は

$$(\text{解の自由度}) = (\text{未知数の数}) - (\text{実質的な方程式の本数})$$

の関係が成り立つことを示している．

2.10 定義 (rank A). 任意の行列 A について

$$\dim(\mathrm{Col}A) = \dim(\mathrm{Row}A) = \mathrm{rank}A$$

が成立する．

[7]詳しくは, 例えば, 齋藤正彦, 線形代数入門, 東京大学出版会, p.121 を参照のこと．

[8]例えば, $a^m = c_1 a^1 + \cdots + c_{m-1} a^{m-1}$ かつ $a^1 \cdot x = \cdots = a^{m-1} \cdot x = 0$ ならば, $a^m \cdot x = 0$ も成り立つ．

2.1.5 正則行列

行の数と列の数が等しい行列は, 様々な重要な性質を持つ.

2.11 定義 (正方行列). 行列 A の行の数と列の数が等しい時, A は正方行列であるという.

2.12 定義 (逆行列). 正方行列 $A \in \boldsymbol{R}^{n\times n}$ に対して, 正方行列 $B \in \boldsymbol{R}^{n\times n}$ が存在し

$$AB = I \quad \text{かつ} \quad BA = I$$

が成り立つ時, 行列 B を行列 A の逆行列という.

2.10 定理. 任意の正方行列 $A \in \boldsymbol{R}^{n\times n}$ に対しては高々1 つの逆行列が存在する.

A の逆行列が存在する時, それを A^{-1} と表す.

2.11 定理. 二つの行列 A 及び B がともに n 次の正方行列なら, $AB = I$ と $BA = I$ は同値である.

よって, B が A の逆行列であることを証明するためには, 逆行列の定義における $AB = I$ と $BA = I$ のどちらか一方を示せば十分である.

練習問題 2.4. 3×2 行列 A を

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

とする.

1.
$$BA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
を満たす 2×3 行列 B は存在するか？もし存在するなら, そのような B を全て挙げよ.

2.
$$AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
を満たす 2×3 行列 B は存在するか？もし存在するなら, そのような B を全て挙げよ.

練習問題 2.5. a を定数として, 2×3 行列 A を

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & a \end{bmatrix}$$

とする.

1.
$$BA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
を満たす 3×2 行列 B は存在するか？もし存在するなら, そのような B を全て挙げよ.

2.
$$AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
を満たす 3×2 行列 B は存在するか？もし存在するなら, そのような B を全て挙げよ.

2.12 定理. 正方行列 $A \in \boldsymbol{R}^{n\times n}$ に対応する線形写像を $F : \boldsymbol{R}^n \to \boldsymbol{R}^n$ とする．この時，以下の 3 条件は同値である．

1. F が全射
2. F が単射
3. F が全単射

2.12 **定理の証明.** 1 $\Rightarrow$ 2: F が全射であるので, $\langle a_1, \dots, a_n\rangle = \boldsymbol{R}^n$ が成り立つ．また, $\dim \boldsymbol{R}^n = \dim(\mathrm{ColA}) = n$ であるので, $\dim(\mathrm{KerA}) = 0$ でなければならない．すなわち,

$$0 = \alpha_1 a_1 + \cdots + \alpha_n a_n$$

が成り立つならば, $\alpha_1 = \cdots = \alpha_n = 0$ である．これは, $a_1, \dots, a_n$ が 1 次独立であることを意味する．よって, F は単射である．

2 $\Rightarrow$ 1: $\dim \boldsymbol{R}^n = n$ である．また, F は単射であるので, $a_1, \dots, a_n$ は n 個の 1 次独立なベクトルである．従って, $\boldsymbol{R}^n$ の任意のベクトル x は $a_1, \dots, a_n$ の線形結合として表される．これは, $\langle a_1, \dots, a_n\rangle = \boldsymbol{R}^n$ が成り立つことを意味する．よって, F は全射である．

以上のことから, $a_1, \dots, a_n$ が $\boldsymbol{R}^n$ の基底をなすことは明らかであるので, 3 条件は同値である． □

2.13 定理. 写像 $F : \boldsymbol{R}^n \to \boldsymbol{R}^n$ が逆写像 $F^{-1} : \boldsymbol{R}^n \to \boldsymbol{R}^n$ を持つとする．この時，F が線形写像なら F^{-1} も線形写像である．

2.13 定義 (**正則行列**)**.** 正方行列 $A \in \boldsymbol{R}^{n\times n}$ について，A が正則であるとは，$F : \boldsymbol{R}^n \to \boldsymbol{R}^n$ を $F(x) = Ax$ で定義すると, F は全単射で F^{-1} が存在することを言う．

2.14 定理. 正方行列 $A \in \boldsymbol{R}^{n\times n}$ に対応する線形写像 $F : \boldsymbol{R}^n \to \boldsymbol{R}^n$ が全単射である時，A は正則である．

2.14 **定理の証明.** まず正方行列 $A \in \boldsymbol{R}^{n\times n}$ が正則であるとしよう．すると正則行列の定義により，線形写像 $F : \boldsymbol{R}^n \to \boldsymbol{R}^n$ を

$$F(x) = Ax$$

と定めれば，F は全単射となる．この時, 逆写像 $F^{-1} : \boldsymbol{R}^n \to \boldsymbol{R}^n$ が存在し, 定理 2.13 により, F^{-1} も線形写像となる．行列と線形写像との対応関係により，F^{-1} に対応する n 次の正方行列行列 B が存在し

$$F^{-1}(x) = Bx$$

となる．ここで写像 F と写像 F^{-1} との合成写像 $F \circ F^{-1} : \boldsymbol{R}^n \to \boldsymbol{R}^n$ を考えると

$$(F \circ F^{-1})(x) = F(Bx) = ABx \tag{2.2}$$

一方，F と F^{-1} が全単射であることに注意すれば，$F \circ F^{-1} : \boldsymbol{R}^n \to \boldsymbol{R}^n$ は恒等写像となることが分かるので

$$(F \circ F^{-1})(x) = x \tag{2.3}$$

(2.2) と (2.3) から

$$AB = I$$

ここで A と B はともに n 次の正方行列であるから，定理 2.11 により

$$BA = I$$

も成立する．よって B は A の逆行列であり，$B = A^{-1}$． □

2.2 行列式

m 行 n 列の行列は $m \times n$ 個の成分によって構成されており，行列が持つ情報を完全に把握するためには，それら全ての成分を見る必要がある．しかし，それぞれの行列が持つ情報の全てを把握することが常に必要とされるわけではない．ある行列が与えられた時，その行列（したがってそれに対応する線形写像）の主要な性質をまとめて表す指標があれば便利である．特に逆行列の存在を考える際などには，各正方行列に対して定まる一つの実数によって判断することができれば分析が簡便になる．そこで，正方行列 $A \in \boldsymbol{R}^{n\times n}$ に対して

$$\det : \boldsymbol{R}^{n\times n} \to \boldsymbol{R}$$

なる写像 $\det A$ を定義したいという動機が生じてくる．このようにして考えられるようになったのが行列式であり，$\det A$ あるいは $|A|$ などと表される．

2.2.1 行列式の直感的理解

行列式の成り立ちを直感的に理解するために，まずは 2 次正方行列について考えよう．

2.14 定義 (2 次行列式). 2 次の正方行列

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \boldsymbol{R}^{2\times 2}$$

に対して一つの実数を対応させる写像 $\det : \boldsymbol{R}^{2\times 2} \to \boldsymbol{R}$ を

$$\det A = a_{11}a_{22} - a_{12}a_{21}$$

によって定義し，これを A の行列式と呼ぶ．

定義 2.14 によって定義される行列式 $\det A$ は，A が持つ情報のうちのどのような部分を抽出しているのであろうか．A は 2 次の正方行列であるから，対応する線形写像は 2 次元平面上の各点を同じ 2 次元平面上に移し変える写像 (変換) である．そこで 2 次元平面上の各単位ベクトル

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad 及び \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

を A に対応する線形写像によって移し変えることを考えよう．行列 A によって

$$e_1 \mapsto Ae_1, \quad 及び \quad e_2 \mapsto Ae_2$$

のような，各単位ベクトルに対応した点 Ae_1 と点 Ae_2 とを定めることができる．行列式の値は，この各単位ベクトル e_1, e_2 と その A による像 Ae_1, Ae_2 とに関係している．なお，点 Ae_1 と点 Ae_2 とを成分で表せば

$$Ae_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad 及び \quad Ae_2 = \begin{pmatrix} a_{21} \\ a_{22} \end{pmatrix}$$

であるから，それぞれが A の列ベクトル a_1，a_2 に対応しており，したがって行列式の値はその行列の列ベクトルが持つ情報と関係していることになる．

まず，行列式の符号が何によって定められているかを考えよう．行列式の符号を定めているのは，各ベクトルの相対的な位置関係である．つまり二つのベクトルがなす角を基準として，列ベクトル a_1 と a_2 との相対的な位置関係が単位ベクトル e_1 と e_2 との相対的な位置関係と等しい時，行列式の

符号は正となる．ここで e_1 と e_2 との相対的な位置関係とは，二つのベクトルがなす角（この場合は直角）を基準として，e_1 から e_2 が左回りとなる関係にあることを意味する．よって行列式が正の値となるのは，二つの列ベクトルがなす角（180 度以下の角）を基準として，a_1 から a_2 が左回りとなる関係にある場合である．

逆に，二つのベクトルがなす角を基準として，列ベクトル a_1 と a_2 との相対的な位置関係と単位ベクトル e_1 と e_2 との相対的な位置関係とが異なる時，すなわち a_1 から a_2 が右回りとなる関係にある場合には，行列式は負の値をとる．

次に，行列式の大きさが何を表しているのかを考えよう．行列式の絶対値は，単位ベクトルによって張られる正方形の面積と，列ベクトルによって張られる平行四辺形の面積との比率を表している．単位ベクトルによって張られる正方形の面積は 1 であるから，行列式の絶対値は結局列ベクトルによって張られる平行四辺形の面積と等しくなる．

2.2.2 行列式の性質

2.15 定義 (**多重線形性**)**.** 行列式が多重線形性を持つとは，

1.
$$\det\begin{pmatrix} a_{11} & a_{12}+a'_{12} \\ a_{21} & a_{22}+a'_{22}\end{pmatrix} = \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix} + \det\begin{pmatrix} a_{11} & a'_{12} \\ a_{21} & a'_{22}\end{pmatrix}$$

2. 行列式が 任意の $t \in \boldsymbol{R}$ に対して
$$\det\begin{pmatrix} a_{11} & ta_{12} \\ a_{21} & ta_{22}\end{pmatrix} = t\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix} \quad \text{かつ} \quad \det\begin{pmatrix} ta_{11} & a_{12} \\ ta_{21} & a_{22}\end{pmatrix} = t\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & {}_{22}\end{pmatrix}$$

を満たすことを言う．

2.16 定義 (**交代性**)**.** 行列式が交代性を持つとは
$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix} = \det\begin{pmatrix} a_{12} & a_{11} \\ a_{22} & a_{21}\end{pmatrix}$$
を満たすことを言う．

2.15 定理. 行列式 $\det : \boldsymbol{R}^{2\times 2} \to \boldsymbol{R}$ は

1. 多重線形性
2. 交代性
3. 単位行列の写像が 1

という三つの性質を満たす．

定義 2.14 によって定義される行列式がこれらの性質を持つことは容易に確認できる．重要なことは，逆にこれらの三つの性質を満たすような写像は一つしか存在しないという点である．

2.16 定理.

1. 多重線形性
2. 交代性

3. 単位行列の写像が 1

の 3 性質を満たすなら，関数 $F: \boldsymbol{R}^{2\times 2} \to \boldsymbol{R}$ は行列式のみである．

定理 2.15 と定理 2.16 により，ある写像が行列式であることとその写像が三つの性質を満たすことは同値であると言える．よって行列式を「多重線形性と交代性を持ち単位行列の像が 1 であるような写像」として，定義し直すことができる．

2.2.3　3 次以上の行列式

行列式の概念を 3 次以上の正方行列に一般化することを考えよう．まずは 2.16 定理の内容を，n 次元空間へと一般化しておく．

2.17 定理.

1. 多重線形性
2. 交代性
3. 単位行列の写像が 1

の 3 性質を満たす関数 $F: \boldsymbol{R}^{n\times n} \to \boldsymbol{R}$ は, ただ一つ存在する．

この定理 2.17 を用いれば，n 次正方行列についても 2 次正方行列の場合と整合的に行列式を定義することができる．つまり，多重線形性と交代性を持ち単位行列の像が 1 であるような写像が行列式である．

2.2.4　置換による行列式の定義

行列式は, 置換を用いて与えられることも多い．特に行列式の値を計算するには, こちらの方がわかりやすい．

2.17 定義 (置換). n 個の数 $1, 2, \ldots, n$ からなる集合

$$M = \{1, 2, 3, \ldots, n\}$$

について，その上で定義される写像 $\pi: M \to M$ を考える．この写像が全単射である時，π を M の置換という．

2.18 定義 (互換). 2 文字のみを入れ替え，他の数字は動かさない置換のことを互換という．つまり，$k = 1, 2, 3, \ldots, n$ について

$$\sigma(k) = \begin{cases} i & (\ k = j \text{ のとき }) \\ j & (\ k = i \text{ のとき }) \\ k & (\ k \neq i, j \text{ のとき }) \end{cases}$$

となる置換 $\sigma: M \to M$ が互換である．

2.18 定理. 任意の置換は，適当な互換の合成写像によって表せる．

定理 2.18 について注意が必要な点は，任意の置換は，適当な互換の合成写像によって表せるが，その表し方は一意的ではないことである．つまり一つの置換に対して，それと同値関係となるような互換の合成写像はいくつも考えられる．ただし，合成する互換の数の偶奇については一意的に定まる．

2.19 定理. 任意の置換 π が

$$\pi = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_r$$
$$\pi = \pi_1 \circ \pi_2 \circ \cdots \circ \pi_s$$

のように互換の合成写像として 2 通りに表されたとする．このとき，$r-s$ は必ず偶数となる．つまり r と s の偶奇は一致する．

この 2.19 定理を用いて，置換の符号を定めることができる．

2.19 定義 (置換の符号). 偶数の互換の合成写像によって表される置換を偶置換，奇数の互換の合成写像によって表される置換を奇置換といい，その符号 $\mathrm{sgn}(\pi)$ を

$$\mathrm{sgn}(\pi) = \begin{cases} 1 & (\ \pi \text{ が偶置換のとき }) \\ -1 & (\ \pi \text{ が奇置換のとき }) \end{cases}$$

によって定める．

置換によって行列式の定義を与える準備ができたので，以上の概念を用いて行列式を定義し直し，関連する定理をいくつか紹介する．

2.20 定理. $M = \{1, 2, 3, \ldots, n\}$ 上の全ての置換の集合を Π で表すとする．この時，任意の n 次正方行列

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \ldots\ldots\ldots & & a_{nn} \end{pmatrix} \in \boldsymbol{R}^{n \times n}$$

に関して

$$\sum_{\pi \in \Pi} (\mathrm{sgn}\pi) a_{\pi(1)1} a_{\pi(2)2} \ldots a_{\pi(n)n} = \mathrm{det}A$$

が成立する．(これは多重線形性, 交代性, $\mathrm{det}I$=1 から従うので, 一意性を示していることになる)

2.21 定理. 任意の n 次正方行列 A について，その転置行列を $A^\top$ と置けば

$$\mathrm{det}A = \mathrm{det}A^\top$$

が成立する．

2.21 **定理の証明.**

$$\mathrm{detA} = \sum_{\pi \in \Pi} (\mathrm{sgn}\pi) a_{\pi(1)1} a_{\pi(2)2} \ldots a_{\pi(n)n}.$$

π が Π を動く時, π の逆置換 π^{-1} も Π を動くので,

$$\mathrm{detA} = \sum_{\pi \in \Pi} (\mathrm{sgn}\pi^{-1}) a_{\pi^{-1}(1)1} a_{\pi^{-1}(2)2} \ldots a_{\pi^{-1}(n)n}.$$

$\pi^{-1}(1), \pi^{-1}(2), \ldots, \pi^{-1}(n)$ を小さい順に並べ替えると, $\pi^{-1}(i) = k$ と置けば $i = \pi(k)$ であるので,

$$\mathrm{detA} = \sum_{\pi \in \Pi} (\mathrm{sgn}\pi) a_{1\pi(1)} a_{2\pi(2)} \ldots a_{n\pi(n)}.$$

これは, $\mathrm{det}A^\top$ に他ならない． □

2.20 定義 (小行列式). n 次正方行列

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \dots\dots\dots & & a_{nn} \end{pmatrix} \in \boldsymbol{R}^{n\times n}$$

について，第 i 行と第 j 列の全ての成分を取り除いた $n-1$ 次正方行列

$$A_{ij} = \begin{pmatrix} a_{11} & \dots & a_{1\,j-1} & a_{1\,j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1\,1} & \dots & a_{i-1\,j-1} & a_{i-1\,j+1} & \dots & a_{i-1\,n} \\ a_{i+1\,1} & \dots & a_{i+1\,j-1} & a_{i+1\,j+1} & \dots & a_{i+1\,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{n\,j-1} & a_{n\,j+1} & \dots & a_{nn} \end{pmatrix} \in \boldsymbol{R}^{(n-1)\times(n-1)}$$

を A_{ij} と書き，この行列式 $\det(A_{ij})$ を A の $n-1$ 次小行列式という．

2.22 定理. n 次正方行列 A について

$$\sum_{j=1}^{n} a_{ij}\,(-1)^{i+j}\,\mathrm{det}A_{ij} = \mathrm{det}A \quad i \text{ 行に関する展開}$$

$$\sum_{i=1}^{n} a_{ij}\,(-1)^{i+j}\,\mathrm{det}A_{ij} = \mathrm{det}A \quad j \text{ 列に関する展開}$$

が成立する．

2.23 定理 (クラメルの公式). A が n 次正則行列, $b \in R^n$ ならば, $Ax = b$ の解 x は

$$x_j = \frac{\det(a_1, \dots, a_{j-1}, b, a_{j+1} \dots, a_n)}{\mathrm{det}A}$$

で与えられる．

2.3 固有値と固有ベクトル

2.3.1 固有値の考え方

行列式は正方行列の性質を一つの実数に要約したものであり，行列が持つ主要な性質を把握しやすくする方法としては便利である．しかし，行列式を用いることで得られるこのような簡便さは，行列が持つ情報の多くが捨象されてしまうことと裏腹の関係にある．よって各行列の性質をより詳細に把握するためには，行列式では把握できない追加的な情報を表す指標が必要となる．固有値や固有ベクトルを考える必要性は，この点にある．

2.2 例. 以下のような 4 種類の 2 次正方行列を考えよう．

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} -4 & 0 \\ 0 & -4 \end{pmatrix}, \quad C = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}, \quad D = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$$

それぞれの行列について行列式の値を計算すれば

$$\mathrm{det}A = \mathrm{det}B = \mathrm{det}C = \mathrm{det}D = 16$$

となるから，行列式の値から得られる情報は四つの行列で変わらないことが分かる．よって，それぞれを有意味な方法で区別するためには，行列式とは異なった観点からの特徴づけがなされなければならない．

各行列に対応する線形写像は，2 次元平面上の各点を同じ 2 次元平面上に移し変える写像である．そこで，2 次元平面上の単位ベクトル

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{及び} \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

を，各行列に対応する線形写像によって移し変えることを考えよう．例えば行列 A については

$$e_1 \mapsto Ae_1, \quad \text{及び} \quad e_2 \mapsto Ae_2$$

のようにして，単位ベクトルに対応した点 Ae_1 と点 Ae_2 とを定めることができる．行列 A の固有値や固有ベクトルは，単位ベクトル e_1, e_2 とその A による像 Ae_1, Ae_2 との位置に関係している．

行列 A，B，C については単位ベクトルの傾きとその像との傾きとが一致しており，単位ベクトルを座標軸に沿って移動させる写像に対応している．A と B が各単位ベクトルを同じ比率で変換しているのに対して C は異なる変換比率を与えている点で異なるが，いずれにしても，各行列による単位ベクトルの像は，もとの単位ベクトルのスカラー倍で表すことができる．つまり e_1 と e_2 に対して，A は 4 と 4，B は -4 と -4，C は 8 と 2，という変換倍率を与えるのである．

$$Ae_1 = \begin{pmatrix} 4 \\ 0 \end{pmatrix} = 4e_1, \quad \text{及び} \quad Ae_2 = \begin{pmatrix} 0 \\ 4 \end{pmatrix} = 4e_2$$

$$Be_1 = \begin{pmatrix} -4 \\ 0 \end{pmatrix} = -4e_1, \quad \text{及び} \quad Be_2 = \begin{pmatrix} 0 \\ -4 \end{pmatrix} = -4e_2$$

$$Ce_1 = \begin{pmatrix} 8 \\ 0 \end{pmatrix} = 8e_1, \quad \text{及び} \quad Ce_2 = \begin{pmatrix} 0 \\ 2 \end{pmatrix} = 2e_2$$

これらの行列は，行列式の値という点からは区別されなかったが，各単位ベクトルをどの程度の長さに変換するかという変換倍率を考えることで，異なった特徴づけが与えられることになる．

しかし行列 D については，このような単純な特徴づけを与えることはできない．単位ベクトルが写像によってその傾きを変えてしまうため，D による像をもとベクトルのスカラー倍として表すことができないからである．よって，単位ベクトルの変化倍率だけを考えていては各行列に共通の視点から特徴づけを与えることはできそうにない．そこで，単位ベクトル e_1，e_2 の代わりに

$$b_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \text{及び} \quad b_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

なる二つのベクトルを D で変換することを考えてみよう．すると各ベクトルの像 Db_1 と Db_2 とは

$$Db_1 = \begin{pmatrix} 8 \\ 8 \end{pmatrix} = 8b_1, \quad \text{及び} \quad Db_2 = \begin{pmatrix} -2 \\ 2 \end{pmatrix} = 2b_2$$

のように，もとのベクトルのスカラー倍で表せることがわかる．e_1 と e_2 に対して行列 C が 8 と 2 という変換倍率を与えたのと同じように，D は b_1 と b_2 に対して 8 と 2 という変換倍率を与えるものとして，行列 D を特徴づけることができるのである．

以上の例によって示唆されるように，各行列は，対応する写像によってスカラー倍に変換されるベクトルとその変換倍率とによって特徴付けられる．そしてそのベクトルと変化倍率とを，固有ベクトルと固有値と呼ぶのである．

2.21 定義 (固有値と固有ベクトル). n 次正方行列 A について

$$Ax = \lambda x, \quad x \in \boldsymbol{R}^n \setminus \{0\}, \quad \lambda \in \boldsymbol{R}$$

をみたす実数 λ を A の固有値，非ゼロベクトル x を λ に対応する固有ベクトルという．

固有値と固有ベクトルという概念を用いて例 2.2 の内容を言い直せば，行列 A の固有値は 4 のみで，e_1 と e_2 という二つのベクトルが対応する固有ベクトルであり，行列 B の固有値も -4 のみで，同じく e_1 と e_2 が対応する固有ベクトルである．行列 C は 8 と 2 という二つの固有値を持ち，それぞれの固有値に対応する固有ベクトルは e_1 と e_2 である．最後の行列 D については 8 と 2 が固有値であり，対応する固有ベクトルは b_1 と b_2 ということになる．

もっとも，全ての行列に対してこのような特徴づけができるというわけではない．固有ベクトルを持たない行列も存在するからである．

2.3 例. 次のような 2 次正方行列を考えよう．

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

この行列に対して

$$Ax = \lambda x, \quad x \in \boldsymbol{R}^2 \setminus \{0\}, \quad \lambda \in \boldsymbol{R} \tag{2.4}$$

を満たすような実数 λ と非ゼロベクトル x は存在しない．

これは背理法によって示すことができる．まずは背理法の仮定として，(2.4) を満足する実数 λ と非ゼロベクトル x が存在するとしよう．するとこれは

$$(A - \lambda I)\, x = 0 \tag{2.5}$$

が x について非自明解（$x = 0$ 以外の解）を持つことを意味する．ここでもし $(A - \lambda)$ が正則であるとすれば，その逆行列 $(A - \lambda)^{-1}$ が存在することになる．これを (2.5) の両辺に左からかければ

$$(A - \lambda I)^{-1} (A - \lambda I)\, x = 0 \quad \Leftrightarrow \quad x = 0$$

となるから，(2.5) は自明解しか持たないことになり，仮定に矛盾する．よってまずは，(2.5) が非自明解を持つ時，$(A - \lambda)$ が正則であってはならないことが確認される．

ある行列が正則でなければ，その行列式の値はゼロに等しい．$(A - \lambda)$ が正則でないならば

$$\det (A - \lambda I) = 0$$

が成り立つ．つまり，A が固有値と固有ベクトルを持つとすれば

$$\det (A - \lambda I) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0 \tag{2.6}$$

が成り立っていなければならないことになる．しかしながら，この (2.6) を満足するような実数 λ は存在しない．よって，A が固有値と固有ベクトルを持つという仮定は誤りである．

2.4 例. 2 次の正方行列が固有値と固有ベクトルを持つ場合でも，その数は常に 2 であるとは限らない．次のような例を考えよう．

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

この行列に対して

$$Ax = \lambda x, \quad x \in \boldsymbol{R}^2 \setminus \{0\}, \quad \lambda \in \boldsymbol{R} \tag{2.7}$$

を満たすような非ゼロベクトル x は，そのスカラー倍を除いてただ一つしか存在しない．以下でこのことを示そう．

例 2.3 と同様にして考えれば，固有値 λ は

$$\det(A-\lambda I)=0$$

を満たすはずなので，これを計算すると

$$\begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (\lambda-1)^2=0 \quad \Leftrightarrow \quad \lambda=1$$

として固有値の値が求まる．また固有値に対応する固有ベクトルは (2.7) を満たすはずなので

$$Ax=x \quad \Leftrightarrow \quad (A-I)\,x=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=0$$

これを満足する非ゼロベクトルは

$$x=t\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad t\in \boldsymbol{R}\setminus\{0\} \tag{2.8}$$

に限られる．よって行列 A の固有値は 1 のみで，対応する固有ベクトルはスカラー倍を除いてただ一つである．

2.3.2 固有値・固有ベクトルの求め方

既に例の中で用いたものであるが，ここで固有値と固有ベクトルの求め方を一般化しておこう．n 次正方行列 A について

$$\begin{aligned} & Ax=\lambda x \text{ が非自明解を持つ} \\ \Leftrightarrow \quad & (A-\lambda I) \text{ が正則でない} \\ \Leftrightarrow \quad & \det(A-\lambda I)=0 \end{aligned} \tag{2.9}$$

と言い換えることができるので，λ が固有値ならば (2.9) 式を満たし，また (2.9) を満たす λ が固有値であるといえる．よって固有値を求めるためには，$\det(A-\lambda I)=0$ を λ について解けばよい固有値を求める際に用いられる $\det(A-\lambda I)=0$ は λ の n 次多項式であり，行列 A の固有多項式（あるいは特性多項式）と呼ばれる．代数学の基本定理によりその実数解は高々n 個であるから，n 次正方行列の固有値も高々n 個である．

また，求めた固有値 λ のそれぞれを $Ax=\lambda x$ に代入することで，対応する固有ベクトルが求められる．つまり

$$Ax=\lambda x \quad \Leftrightarrow \quad (A-\lambda I)\,x=0$$

の非自明解 x が，各 λ に対応する固有ベクトルである．

2.4 行列の対角化

2.4.1 対角化の必要十分条件

一つの線形写像は一つの行列と対応関係にあり，その行列の分析を通して線形写像の性質を考察することができる．よって行列表現をより見やすいものに変形することができれば，対応する線形写像の考察も容易になる．固有値や固有ベクトルを求める理由の一つは，行列表現をより見やすいものに変形するという作業に関係しており，その作業を行列の対角化という．

2.22 定義 (行列の対角化). n 次正方行列 A について，ある n 次正則行列 P 及び対角行列 Λ が存在し

$$A = P\Lambda P^{-1}$$

が成り立つ時，行列 A は対角化可能であると言う.

換言すれば, ある正則行列 P が存在して, $P^{-1}AP$ が対角行列となる. さらに上式を書き換えれば,

$$AP = P\Lambda.$$

$$P = (v_1 \cdots v_n) \quad \Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

とすれば

$$Av_i = \lambda_i v_i \ \forall i = 1, \ldots n.$$

対角化と固有値・固有ベクトルとがどのように関わりあっているのかは，次の定理を確認することで理解できる.

2.24 定理 (対角化の必要十分条件). 任意の n 次正方行列 A に対して，その固有ベクトルのみよりなる $\boldsymbol{R}^n$ の基底が存在するとき, 及びそのときに限り, A は対角化可能である.

2.24 **定理の証明.** 仮定により固有値と固有ベクトルが存在し

$$\forall i = 1, 2, \ldots, n: \quad Av_i = \lambda_i v_i \tag{2.10}$$

が成立する. また固有ベクトル $v_1, v_2, \ldots, v_n$ が $\boldsymbol{R}^n$ の基底を構成しているので

$$\forall x \in \boldsymbol{R}^n, \quad \exists c = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \in \boldsymbol{R}^n: \quad x = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n \tag{2.11}$$

と表すことができる. ここで行列 P を

$$P = (v_1 \ v_2 \ \ldots \ v_n)$$

と置けば，P は n 次正則行列であり (2.11) は

$$x = Pc \quad \Leftrightarrow \quad c = P^{-1}x \tag{2.12}$$

と書ける. 一方 A を (2.11) の両辺に左から掛ければ (2.10) により

$$\begin{aligned} Ax &= c_1 Av_1 + c_2 Av_2 + \cdots + c_n Av_n \\ &= c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_n \lambda_n v_n \end{aligned}$$

この式について

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

と置けば

$$Ax = P\begin{pmatrix} c_1\lambda_1 \\ c_2\lambda_2 \\ \vdots \\ c_n\lambda_n \end{pmatrix} = P\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} c \quad \Leftrightarrow \quad Ax = P\Lambda c \tag{2.13}$$

と表せる．よって (2.12) を (2.13) の c に代入すれば

$$Ax = P\Lambda P^{-1}x \quad \Leftrightarrow \quad A = P\Lambda P^{-1}$$

を得る． □

定理 2.24 では対角化の十分条件を示したが，実はこの十分条件はそのまま必要条件にもなっている．

練習問題 2.6. 定理 2.24 の必要性を証明せよ．

2.25 定理. $P^{-1}AP$ が対角行列であるとき，およびそのときに限り，P の列ベクトルが $\boldsymbol{R}^n$ の基底をなす．

2.25 **定理の証明.** 十分性: A の n 個の固有ベクトルを $v_1, \dots, v_n$ とする．行列 P を

$$P = (v_1, \dots, v_n)$$

と表すとすると，P は $\boldsymbol{R}^n$ の標準基底 $e_1, \dots, e_n$ を新しい基底 $v_1, \dots, v_n$ に変換する．従って，P の列ベクトルが $\boldsymbol{R}^n$ の基底をなす．

必要性: P の第 i 列ベクトルとして A の固有値 λ_i に対する固有ベクトル v_i をとる．$(i = 1, \dots, n)$．$P^{-1}AP$ の第 i 列ベクトルを b_i とすれば，

$$b_i = P^{-1}Av_i = P^{-1}\lambda_i v_i = \lambda_i e_i$$

となる．従って，$P^{-1}AP$ は対角行列となる．

□

2.5 例.

$$D = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix} \quad P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

定理 2.24 により，「A の固有ベクトルよりなる $\boldsymbol{R}^n$ の基底が存在すること」が「A が対角化可能であること」の必要十分条件であることが分かった．それでは，この必要十分条件が成立するための条件はどのようなものであろうか．

2.26 定理. n 次正方行列 A が相異なる l $(l \le n)$ 個の固有値を持つ時，対応する l 個の固有ベクトル $v_1, v_2, \dots, v_l$ は 1 次独立となる．よって，特に n 個の相異なる固有値が存在すれば対角化可能である．

2.26 **定理の証明.** $\lambda_1, \dots, \lambda_l$ を A の固有値，$v_1, \dots, v_l$ を対応する固有ベクトルとする．

$v_1, \dots, v_l$ が 1 次従属であるとすると，$v_1, \dots, v_{k-1}$ は 1 次独立である一方，$v_1, \dots, v_{k-1}, v_k$ は 1 次従属となるような k が存在する．つまり，

$$v_k = \alpha_1 v_1 + \dots + \alpha_{k-1}v_{k-1} \tag{$*$}$$

と表される．A を $(*)$ の両辺に左から掛ければ,

$$\begin{aligned} Av_k &= \alpha_1 Av_1 + \cdots + \alpha_{k-1} Av_{k-1} \\ \Leftrightarrow \lambda_k v_k &= \alpha_1 \lambda_1 v_1 + \cdots + \alpha_{k-1}\lambda_{k-1} v_{k-1}. \end{aligned}$$

一方, $(*)$ の両辺に λ_k を掛ければ,

$$\lambda_k v_k = \alpha_1 \lambda_k v_1 + \cdots + \alpha_{k-1}\lambda_k v_{k-1}. \tag{$**$}$$

$(*)$ と $(**)$ より,

$$\alpha_1(\lambda_1 - \lambda_k)v_1 + \cdots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = 0.$$

$v_1, \ldots, v_{k-1}$ は仮定により 1 次独立であるので,

$$\alpha_1(\lambda_1 - \lambda_k) = \cdots = \alpha_{k-1}(\lambda_{k-1} - \lambda_k) = 0.$$

$\lambda_1 - \lambda_k = \cdots = \lambda_{k-1} - \lambda_k \neq 0$ より, $v_k = 0$ でなければならない．ところが, これは v_k が A の固有ベクトルであることに矛盾する．従って, A の相異なる l 個の固有値に対応する固有ベクトルは 1 次独立となる． □

1 次独立な n 個のベクトルは，$\boldsymbol{R}^n$ の基底になる．よってこの定理 2.26 により，n 次正方行列 A の固有多項式が相異なる n 個の実数解を持つ時，固有ベクトルよりなる $\boldsymbol{R}^n$ の基底が存在し，したがって行列 A は対角化可能であることがわかる．

ただし，固有多項式が相異なる n 個の実数解を持つ必要は必ずしもない．1 つの固有値に対して複数の固有ベクトルが対応することがありうるからである．相異なる固有値の数は n 個に満たなくとも，固有ベクトルが n 個存在してそれらが 1 次独立であるならば，対角化は可能である．例えば, 単位行列 I が良い例である．

2.4.2 対角化できない行列

では逆に，対角化が不可能となるのはどのような場合であろうか．$n = 2$ のケースについて，簡単な例を挙げながら確認してみよう．

2.6 例 (**固有値が 1 つの場合**)．2 次正方行列 A が，唯一の固有値 $\lambda \in \boldsymbol{R}$ を持つとしよう．もし行列 A と固有値 λ が

$$A - \lambda I = 0$$

を満たしていれば，これは対角化可能である．というのも

$$\forall x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : \quad Ax = \lambda x \tag{2.14}$$

が成立するからである．この時 $\boldsymbol{R}^2$ の任意のベクトルが λ に対応する固有ベクトルとなり，固有ベクトルからなる $\boldsymbol{R}^2$ の基底が存在することになる．例 2.2 で挙げた行列 A は，このケースに対応する．

一方

$$A - \lambda I \neq 0$$

である場合には

$$\dim\{\mathrm{Ker}\,(A - \lambda I)\} = 1$$

のように固有空間の次元が 1 となってしまうため，固有ベクトルからなる $\boldsymbol{R}^2$ の基底は存在しない．よって対角化は不可能である．例 2.4 で挙げた行列はこのケースに対応する．

ただし，このように対角化が不可能である場合にも，対角行列に近い行列を用いて行列を表現し直すことはできる．ここで，固有ベクトルでない非ゼロベクトル v (つまり $v \notin \mathrm{Ker}\,(A-\lambda I)$) を用いて

$$v_1 = (A-\lambda I)v \tag{2.15}$$
$$v_2 = v \tag{2.16}$$

という二つのベクトル v_1 及び v_2 を考えてみよう．

練習問題 2.7.

(1) $(A-\lambda I)^2$ はゼロ行列であることを示せ.

(2) v_1, v_2 は $\boldsymbol{R}^2$ の基底であることを示せ.

練習問題 2.7 の結果を利用すれば

$$(A-\lambda I)v_1 = (A-\lambda I)^2 v = 0$$

により，v_1 は λ に対応する固有ベクトルであり

$$Av_1 = \lambda v_1 \tag{2.17}$$

が成立することが分かる．(2.15) と (2.16)，(2.17) をあわせれば

$$Av_1 = \lambda v_1$$
$$Av_2 = v_1 + \lambda v_2$$

すなわち

$$A(v_1\ v_2) = (v_1\ v_2)\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix} \tag{2.18}$$

と表すことができる．最後に

$$P = (v_1\ v_2) \in \boldsymbol{R}^{2\times 2}$$

と置けば P は正則であり，この P を用いて (2.18) は

$$P^{-1}AP = \begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix}$$

のように書ける．

2.7 例 (**固有値が** 0 **個の場合**)**.** 固有値が 0 個の場合，つまり固有方程式が実数解を持たない場合を考えよう．固有多項式 $\det(A-\lambda I)=0$ は複素数解を持つから，それらを

$$\lambda_1 = \mu + i\nu \tag{2.19}$$
$$\lambda_2 = \mu - i\nu$$

と置くことにする．ただし i は虚数単位 $(i^2=1)$, μ 及び ν は実数で，$\nu \neq 0$ である．この時 $Ax=\lambda_1 x$ を満たす $x \in \boldsymbol{C}^2$ が存在するので，これを

$$x = u + iw \tag{2.20}$$

とする．また同様に $Ax = \lambda_2 x$ を満たす $x \in \boldsymbol{C}^2$ も存在するが，これは

$$x = u - iw$$

と表されることが容易に示される．ただし u 及び w は $\boldsymbol{R}^2$ の要素で，$w \neq 0$ である．

(2.19) 及び (2.20) により

$$\begin{aligned} Ax = \lambda_1 x &\Leftrightarrow A(u + iw) = (\mu + i\nu)(u + iw) \\ &\Leftrightarrow Au + iAw = \mu u - \nu w + i(\mu w + \nu u) \end{aligned}$$

つまり

$$Au = \mu u - \nu w \tag{2.21}$$

$$Aw = \mu w + \nu u \tag{2.22}$$

が成立する．$\lambda_2 = \mu - i\nu$, $x = u - iw$ に対して $Ax = \lambda_2 x$ を考えても，同じ式が得られる．

ここで

$$P = (u\ w) \in \boldsymbol{R}^{2\times 2}$$

と置けば P は正則であり，(2.21) と (2.22) は

$$P^{-1}AP = \begin{pmatrix} \mu & \nu \\ -\nu & \mu \end{pmatrix} \tag{2.23}$$

と書くことができる．

(2.23) は何を意味しているのだろうか．これを理解するために $\nu < 0$ と仮定し，λ_1 及び λ_2 の絶対値を

$$\rho = \sqrt{\mu^2 + \nu^2}$$

と置いてみよう．すると

$$\left(\frac{\mu}{\rho}\right)^2 + \left(\frac{\nu}{\rho}\right)^2 = 1$$

であるから，ある $\theta \in (0, \pi)$ が存在し

$$\mu = \rho \cos\theta \tag{2.24}$$

$$-\nu = \rho \sin\theta \tag{2.25}$$

と書けることが分かる．(2.24) と (2.25) を用いれば (2.23) は

$$\begin{aligned} \begin{pmatrix} \mu & \nu \\ -\nu & \mu \end{pmatrix} &= \begin{pmatrix} \rho\cos\theta & -\rho\sin\theta \\ \rho\sin\theta & \rho\cos\theta \end{pmatrix} \\ &= \rho \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \end{aligned}$$

と書けるから，これはベクトルを θ だけ回転させて ρ 倍する行列であることが分かる．

2.5 対称行列と 2 次形式

2.5.1 2 次形式

m 行 n 列の行列 A は $F(x) = Ax$ で定義される線形写像と対応関係にあることは確認したが，実は $F(y, x) = y^\top Ax$ で定義される写像とも対応関係にある．

2.27 定理. m 行 n 列の行列 $A \in \boldsymbol{R}^{m\times n}$ を用いて，写像 $F : \boldsymbol{R}^m \times \boldsymbol{R}^n \to \boldsymbol{R}$ を

$$F(y,x) = y^\top Ax, \qquad y \in \boldsymbol{R}^m, \quad x \in \boldsymbol{R}^n$$

と定義すれば，この写像 F は多重線形性 (多 1 次形式) を有する．また逆に，任意の写像 $F : \boldsymbol{R}^m \times \boldsymbol{R}^n \to \boldsymbol{R}$ が多重線形性を有する時

$$F(y,x) = y^\top Ax, \qquad y \in \boldsymbol{R}^m, \quad x \in \boldsymbol{R}^n$$

を満たす行列 $A \in \boldsymbol{R}^{m\times n}$ がただ一つ存在する．

このようにして定義される写像について，$m = n$，$y = x$ とした特殊なケースが 2 次形式と呼ばれるものに対応する．

2.23 定義 (**2 次形式**)**.** n 次元ベクトル $x \in \boldsymbol{R}^n$ について

$$Q(x) = \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}x_i x_j \tag{2.26}$$

として定義される写像 $Q : \boldsymbol{R}^n \to \boldsymbol{R}$ を 2 次形式と呼ぶ.

定義 2.23 における写像 $Q : \boldsymbol{R}^n \to \boldsymbol{R}$ は多重線形性を持つので，定理 2.27 により対応する n 次正方行列が存在する．これは

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{pmatrix} \in \boldsymbol{R}^{n\times n}$$

と置けば，(2.26) を

$$\sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}x_i x_j = x^\top Ax$$

と表せることからも明らかである．

2.24 定義 (**対称行列**)**.** 正方行列 A が, $A = A^\top$ を満たす時, A を対称行列と言う.

一般に Q から A は一意に定まらない. 例えば,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

に対し

$$Q(x) = x^\top Ax = x_1^2 + x_2^2$$

だが,

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

に対しても

$$Q(x) = x^\top Ax = x_1^2 + x_1x_2 - x_1x_2 + x_2^2 = x_1^2 + x_2^2$$

が成立する. 特に

$$x^\top Ax = x^\top\Big(\frac{1}{2}A + \frac{1}{2}A^\top\Big)x$$

なので, 任意の Q に対して

$$Q(x) = x^\top Ax$$

なる <u>対称</u> 行列 A が存在する. このような A は一意に定まる.

2.5.2 2 次形式と行列の定符号

2.25 定義 (n **次対称行列の定符号**). n 次対称行列 A について

$$\forall x \in \boldsymbol{R}^n \setminus \{0\}: \quad x^\top Ax > 0$$

が成立する時, A は正値定符号（postive definite) 行列であると言い,

$$\forall x \in \boldsymbol{R}^n: \quad x^\top Ax \geq 0$$

が成立する時, A は正値半定符号 (positive semidefinite) 行列であると言う.

また逆に

$$\forall x \in \boldsymbol{R}^n \setminus \{0\}: \quad x^\top Ax < 0$$

が成立する時，A は負値定符号 (negative definite) 行列であると言い,

$$\forall x \in \boldsymbol{R}^n: \quad x^\top Ax \leq 0$$

が成立する時，A は負値半定符号 (negative semidefinite) 行列であると言う.

2.8 例. 行列 I は正値定符号行列である．実際，定義 2.25 において行列 $A = I$ とおけば

$$\forall x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \boldsymbol{R}^n \setminus \{0\}: \quad x^\top Ax = x^\top Ix = x_1^2 + \quad x_2^2 + \cdots + x_n^2 > 0$$

となるから，これは正値定符号行列の定義を満たす.

ちなみに $A = -I \in \boldsymbol{R}^{n\times n}$ とすれば

$$\forall x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \boldsymbol{R}^n \setminus \{0\}: \quad x^\top Ax = -x^\top Ix = -x_1^2 - x_2^2 - \cdots - x_n^2 < 0$$

となるから，こちらは負値定符号行列である.

行列が対角行列の場合，その定符号の判定は容易に行うことができる.

2.28 定理 (**対角行列の定符号**). 対角行列

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \in \boldsymbol{R}^{n\times n}$$

について

1. 行列 A は正値定符号行列（正値半定符号行列）
2. $\forall i: \ \lambda_i > 0 \quad (\forall i: \ \lambda_i \geq 0)$

の 2 条件は同値である. 同様に, 行列 A が負値定符号行列（負値半定符号行列）であることとすべての i について $\lambda_i < 0$ ($\lambda_i \leq 0$) であることとは同値である.

定理 2.28 により，行列の対角化を行うことでその定符号の判定が極めて容易になることが分かる.

2.26 定義 (正規直交基底). $\boldsymbol{R}^n$ の基底 $v_1, v_2, \ldots, v_n$ について

$$v_i \cdot v_j = \begin{cases} 0 & (i \neq j \text{ の時}) \\ 1 & (i = j \text{ の時}) \end{cases} \tag{2.27}$$

が成立する時，$v_1, v_2, \ldots, v_n$ を $\boldsymbol{R}^n$ の正規直交基底と言う.

2.27 定義 (正規直交行列). 正規直交基底を $v_1, v_2, \ldots, v_n$ とした時，これらを列ベクトルとする行列

$$P = (v_1 v_2 \ldots v_n) \in \boldsymbol{R}^{n \times n}$$

を正規直交行列と言う.

任意の正規直交行列 P に対し,

$$P^\top P = PP^\top = I \quad \text{すなわち} \quad P^\top = P^{-1}$$

が成り立つ. 特に, P は正則行列である. また, P が正規直交行列であるならば, $P^\top = P^{-1}$ もまた正規直交行列である.

2.29 定理 (実対称行列の対角化). 任意の n 次正方行列 A に対して

1. A は実対称行列

2. A は正規直交行列 P を用いて対角化できる

の 2 条件は同値である.

注意: 単なる正則行列ではなく, 正規直交行列を使う. つまり, A の固有ベクトルよりなる正規直交基底が存在する.

2.29 **定理の証明.** $1 \Rightarrow 2$: 実対称行列 A の固有値は実数である. 従って, 固有ベクトルの各成分も実数とできる. このことを踏まえて, n 次実対称行列 A は正規直交行列 P を用いて対角化可能であること示す. すなわち,

$$P^{-1}AP = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$$

とできる. ただし, $\alpha_1, \ldots, \alpha_n$ は重複度を含めて A の固有値である.

$n = 1$ の時, 定理は自明であるので, n に関する帰納法を用いる. α_1 を A の一つの固有値とし, 対応する固有ベクトルを v_1 とする. v_1 の張る部分空間 W_1 の直交補空間を $W_1^\perp$ とすると, $W_1^\perp$ は A-不変である.

正規直交基底 $e_1', e_2', \ldots, e_n'$ を $\langle e_1' \rangle = \langle v_1 \rangle$, $\langle e_2', \ldots, e_n' \rangle = W_1^\perp$ となるように取り, この基底に関して A を表現する行列を A' とすれば

$$A' = \begin{pmatrix} \alpha_1 & 0 \\ 0 & A_1' \end{pmatrix}$$

となる. 基底の変換 $\{e_i\} \to \{e_i'\}$ の行列を P_1 とおけば, P_1 は正規直交行列であるから A' も実対称行列となる. 従って, A_1' も実対称行列であるので, 帰納法の仮定より, A_1' は対角化可能である. すなわち, ある $n-1$ 次正規直交行列 P' が存在して $P'^{-1}A_1'P'$ が対角行列になる.

ここで,

$$P = P_1 \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix}$$

とすると,

$$\begin{aligned} P^{-1}AP &= \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix}^{-1} \begin{pmatrix} \alpha_1 & 0 \\ 0 & A'_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix} \\ &= \begin{pmatrix} \alpha_1 & 0 \\ 0 & P'^{-1}A'_1P' \end{pmatrix} \end{aligned}$$

となる.

$2 \Rightarrow 1$: 練習問題 2.8

□

練習問題 2.8. 任意の n 次正方行列 A に対し, もし $P^{-1}AP$ が対角行列となるような n 次直交行列 P が存在するならば, A は対称であることを証明せよ.

練習問題 2.9. A を n 次対称行列, v を A の固有ベクトルとする. 任意の $x \in \boldsymbol{R}^n$ に対し, もし $v \cdot x = 0$ ならば, $v \cdot (Ax) = 0$ が成立することを証明せよ.

実対称行列の対角化において正規直交行列を使う理由は, 次の定理によって説明される.

2.30 定理. A を実対称行列, P を正規直交行列および Λ を対角行列とする. この時, $\Lambda = P^{-1}AP$ ならば

$$\forall x \in \boldsymbol{R}^n : (P^{-1}x)^\top \Lambda (P^{-1}x) = x^\top Ax$$

が成り立つ.

2.30 **定理の証明.** P は正規直交行列であるから

$$\begin{aligned} (P^{-1}x)^\top \Lambda (P^{-1}x) &= x^\top (P^{-1})^\top \Lambda (P^{-1}x) \\ &= x^\top P \Lambda P^{-1} x \\ &= x^\top Ax \end{aligned}$$

となる. □

P が単に正則行列であるだけなら $(P^{-1})^\top = P$ とは限らないので, 定理 2.30 が必ずしも成立するわけではない. また,

$$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

$P = (v_1, \ldots, v_n)$ とし, ある $z \in \boldsymbol{R}^n$ に対して

$$x = z_1 v_1 + \cdots + z_n v_n$$

であるならば,

$$x^\top Ax = \lambda_1 z^2 + \cdots + \lambda_n z^2$$

となり, 非常に簡単な表現となる.

この定理 2.30 から, 実対称行列が正 (負) 値定符号であることはその行列の固有値からなる対角行列が正 (負) 値定符号であることに等しいことが分かる. 定理 2.28 とあわせれば, 次の定理が導かれる.

2.31 定理 (実対称行列の定符号). 任意の n 次対称行列に対し，行列 A が正値定符号 (正値半定符号) であることと A のすべての固有値が正 (非負) であることとは同値である．同様に, 任意の n 次対称行列に対し，行列 A が負値定符号 (負値半定符号) であることと A のすべての固有値が負 (非正) であることとは同値である．

最後に，行列式を用いた定符号の判定方法も示しておこう．

2.28 定義 (主小行列式). A を n 次実対称行列とし，$i_1, i_2, \ldots, i_m$ を

$$1 \le i_1 < i_2 < \cdots < i_m \le n$$

を満たす整数の列とする．ここで行列 m 次正方行列 B を

$$B_M = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1m} \\ b_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ b_{m1} & \ldots\ldots\ldots & & b_{mm} \end{pmatrix} = \begin{pmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \cdots & a_{i_1 i_m} \\ a_{i_2 i_1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{i_m i_1} & \ldots\ldots\ldots & & a_{i_m i_m} \end{pmatrix}$$

と置けば，B_M も対称行列である．この時，行列式 $\det B_M$ を行列 A の m 次の主小行列式（principal minor）と言う．また特に，$i_1 = 1, i_2 = 2, \ldots, i_m = m$ のものを A の m 次の主座小行列式（leading principal minor）と言う．

m 次の主座行列式は一意に定まるが, 主小行列式は一意に定まらないことに注意が必要である.

主小行列式も主座行列式も, もとの行列から同じ数だけの行と列とを取り除いて得られた行列の行列式である．行列 A を n 次対称行列とすれば，主小行列式は $2^n - 1$ 種類，主座小行列式は n 種類ある．

2.9 例. 3 行 3 列の行列

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$$

を例にとろう．この行列 A の主小行列式は

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}, \quad \det \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad \det \begin{pmatrix} 1 & 3 \\ 3 & 6 \end{pmatrix}, \quad \det \begin{pmatrix} 4 & 5 \\ 5 & 6 \end{pmatrix}, \quad \det \begin{pmatrix} 1 \end{pmatrix}, \quad \det \begin{pmatrix} 4 \end{pmatrix}, \quad \det \begin{pmatrix} 6 \end{pmatrix}$$

の七つである．また主座小行列式はこの中で

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}, \quad \det \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad \det \begin{pmatrix} 1 \end{pmatrix}$$

の三つとなる．

2.32 定理. 任意の n 次対称行列とし，A の k 次の主小行列式を $\det B_k$，A の k 次の主座小行列式を $\det B_k^*$ とする．この時

1. A が正値定符号 $\Leftrightarrow$ 任意の主座小行列式が正
2. A が負値定符号 $\Leftrightarrow$ m が偶数 (奇数) なら m 次主座小行列式が正 (負)

3. A が正値半定符号 $\Leftrightarrow$ 任意の主小行列式が非負

4. A が負値半定符号 $\Leftrightarrow$ m が偶数 (奇数) なら, 任意の m 次主小行列式が非負 (非正)

が成立する.

注意: いずれの条件も必要性は容易に示すことができる.

注意: 1, 2 は A が対角行列であるならば容易に示すことができる.

注意: 3, 4 では主座小行列式をチェックするだけでは不十分である:
例

$$A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$$

この時, $\det B_1 = 0$, $\det B_2 = 0$ だが A は正値半定符号ではない. (事実,

$$x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

を考えれば, $x^\top Ax = -1 < 0$ となる.)

注意: 先の基準は非対称行列にはあてはまらない:
例

$$A = \begin{pmatrix} 1 & 1 \\ -5 & 1 \end{pmatrix}$$

に対し,

$$Q(x) = x^\top Ax = x_1^2 - 4x_1x_2 + x_2^2$$

だが,

$$B = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$$

に対しても

$$Q(x) = x^\top Ax = x_1^2 - 4x_1x_2 + x_2^2$$

である. この時, 非対称行列 A の主座行列式は $\det(1) > 0$, $\det A > 0$ となる一方で, 対称行列 B の主座行列式は $\det(1) > 0$, $\det B = 1 - 4 = -3 < 0$ となる.

3 凸解析

3.1 集合と関数の凸性

3.1.1 開集合・閉集合・凸集合

凸解析の説明に入る前に，必要となる基本的な概念を整理しておく．

3.1 定義 (開集合). $\boldsymbol{R}^n$ の部分集合 C について

$$\forall x \in C,\ \exists \epsilon > 0,\ \forall y \in \boldsymbol{R}^n: \quad \|x - y\| < \epsilon \quad \Rightarrow \quad y \in C$$

が成り立つ時，C は開集合であるという．

3.2 定義 (閉集合). $\boldsymbol{R}^n$ の部分集合 C について，C の補集合

$$C^c = \{x \in \boldsymbol{R}^n \mid x \notin C\}$$

が開集合であるとき，C は閉集合であるという．

3.1 定理 (開集合).

$$C \text{ が開集合} \Leftrightarrow \forall x \in C\ \forall x \text{ に収束する点列 } (x_n)\ \exists N\ \forall n > N : x_n \in C.$$

3.2 定理 (閉集合).

$$C \text{ が閉集合} \Leftrightarrow \forall x \in \boldsymbol{R}^n\ \forall x \text{ に収束する点列 } (x_n)\ \forall n : x_n \in C \text{ ならば } x \in C.$$

練習問題 3.1. 上の定義 3.1 と定理 3.1，及び定義 3.2 と定理 3.2 が同値であることを証明せよ．

以上の諸定理より，

閉集合 $\Leftrightarrow$ 境界を含む
開集合 $\Leftrightarrow$ 境界は全く外側で含まない

という図形的な形状には依らない性質が導かれる．

3.3 定義 (凸集合). 集合 C について，

$$\forall x, y \in C,\ \forall t \in [0,1]: \quad tx + (1-t)y \in C$$

が成り立つ時，C は凸集合であるという．言い換えれば, $\forall x, y \in C$ について, 線分 $[x, y]$ が C にすっぽり含まれることである．

3.1.2 関数の凹凸

3.4 定義 (凸関数). C を凸集合とする．C 上で定義される写像 $F: C \to \boldsymbol{R}$ について

$$\forall x, y \in C,\ \forall t \in [0,1]: \quad F(tx + (1-t)y) \le tF(x) + (1-t)F(y)$$

が成立するとき，F は凸関数であるという．

3.5 定義 (凹関数). C を凸集合とする．C 上で定義される写像 $F : C \to \boldsymbol{R}$ について，$-F$ が凸関数である時，つまり

$$\forall x,\, y \in C,\ \forall t \in [0,1] : \quad F(tx + (1-t)y) \geq tF(x) + (1-t)F(y)$$

が成立するとき，F は凹関数であるという．

関数の凹凸については，定義 3.4 や定義 3.5 から判断するだけでなく，ヘッセ行列の定符号を用いて判断する方法が一般的である．その方法を説明する前に，まずは微分可能な関数に関して凸関数の定義を書き直せることを確認しておこう．

3.3 定理. C を開凸集合とする．C 上で定義される全微分可能な関数 $F : C \to \boldsymbol{R}$ について F が凸関数である時，そしてその時のみ

$$\forall x \in C,\ \forall h \in \boldsymbol{R}^n :\ F(x+h) \geq F(x) + \nabla F(x) \cdot h \tag{3.1}$$

が成立する．

練習問題 3.2. 定理 3.3 を証明せよ．(ヒント: 方向微分可能性を示すだけで十分である.)

練習問題 3.3. 任意の全微分可能関数 $F : \boldsymbol{R}^L \to \boldsymbol{R}$ に対し，F が凸ならば，任意の $x \in \boldsymbol{R}^L$ および $y \in \boldsymbol{R}^L$ に対し，$F(y) \geq F(x) + \nabla F(x)(y-x)$ であることを示せ．

3.3 定理における (3.1) は

$$F(x+h) - \left\{F(x) + \nabla F(x) \cdot h\right\} \geq 0$$

と書き換えることができる．テーラー展開により

$$F(x+h) = F(x) + \nabla F(x) \cdot h + h^\top \nabla^2 F(x) h + R_2(h;x) \quad \text{ただし} \quad \lim_{h \to 0} \frac{R_2(h;x)}{\|h\|^2} = 0$$

であるから，結局 (3.1) は

$$h^\top \nabla^2 F(x) h + R_2(h;x) \geq 0 \tag{3.2}$$

のように，テーラー展開の 2 次以降の項の和によって表せることが分かる．よって

$$\left(\frac{1}{\|h\|}h\right)^\top \nabla^2 F(x) \left(\frac{1}{\|h\|}h\right) + \frac{R_2(h;x)}{\|h\|^2} \geq 0.$$

今, $h = tv, \|v\| = 1$ とし, $t \to 0$ とすれば, これは $v^\top \nabla^2 F(x) v \geq 0$ に収束する．ここで $h \to 0$ の極限を取った時，$h^\top \nabla^2 F(x) h$ の方が $R_2(h;x)$ よりも速く収束することに注意すれば，(3.2) はヘッセ行列を用いた 2 次形式によって

$$v^\top \nabla^2 F(x) v \geq 0$$

と書くことができる．

次の定理は，この逆も成立することを主張するものである．

3.4 定理. C を開凸集合とする．C 上で定義される 2 回連続微分可能な関数 $F : C \to \boldsymbol{R}$ について

1. F が凸関数
2. 任意の $x \in C$ について F のヘッセ行列 $\nabla^2 F(x)$ が正値半定符号

は同値である．

練習問題 3.4. 任意の全微分可能関数 $F : \boldsymbol{R}^L \to \boldsymbol{R}$ に対し, F が 2 回連続微分可能かつ凸ならば，任意の $x \in \boldsymbol{R}^L$ に対し，$\nabla^2 F(x)$ は正値半定符号であることを示せ.

定理 3.3 や定理 3.4 は凸関数についてのものだが，凹関数についても同様の定理が成り立つ.

3.5 定理. C を開凸集合とする．C 上で定義される全微分可能な関数 $F : C \to \boldsymbol{R}$ について，F が凹関数である時，そしてその時のみ

$$\forall x \in C,\ \forall h \in \boldsymbol{R}^n :\ \ F(x+h) \le F(x) + \nabla F(x) \cdot h$$

が成立する.

3.6 定理. C を開凸集合とする．C 上で定義される 2 回連続微分可能な関数 $F : C \to \boldsymbol{R}$ について

1. F が凹関数
2. 任意の $x \in C$ について F のヘッセ行列 $\nabla^2 F(x)$ が負値半定符号

は同値である.

3.1.3 準凹関数・擬凹関数

以上の諸定理は凸関数や凹関数の判断に用いられるものであるが，経済学に用いられる概念としては，準凸（凹）関数や擬凸（凹）関数の方が重要である.

3.6 定義 (準凸関数). C を凸集合とする．C 上で定義される関数 $F : C \to \boldsymbol{R}$ について

$$\forall x, y \in C,\ \forall t \in [0,1] :\quad F\big(tx + (1-t)y\big) \le max\{F(x), F(y)\}$$

が成り立つ時，F は準凸関数 (quasi-convex function) であるという.

練習問題 3.5. F が凸関数ならば, 準凸関数であることを示せ.

3.7 定義 (準凹関数). C を凸集合とする．C 上で定義される関数 $F : C \to \boldsymbol{R}$ について, $-F$ が準凸関数である時, つまり

$$\forall x, y \in C,\ \forall t \in [0,1] :\quad F\big(tx + (1-t)y\big) \ge min\{F(x), F(y)\}$$

が成り立つ時，F は準凹関数 (quasi-concave function) であるという.

3.7 定理. C を凸集合とする．C 上で定義される関数 $F : C \to \boldsymbol{R}$ について

1. F が準凸関数
2. $\forall z \in \boldsymbol{R} :\quad \big\{x \in C \mid F(x) \le z\big\}$ が凸集合

は同値である.

3.8 定理. C を凸集合とする．C 上で定義される関数 $F : C \to \boldsymbol{R}$ について

1. F が準凹関数
2. $\forall z \in \boldsymbol{R} :\ \big\{x \in C \mid F(x) \ge z\big\}$ が凸集合

は同値である．

3.9 定理. C を凸集合とする．C 上で定義される準凸（準凹）関数 $F : C \to \boldsymbol{R}$, および $\boldsymbol{R}$ 上で定義される単調増加関数 $G : \boldsymbol{R} \to \boldsymbol{R}$ を考える．この時，合成関数 $G \circ F : C \to \boldsymbol{R}$ は準凸（準凹）関数となる．

練習問題 3.6. コブ＝ダグラス型関数

$$F(x_1, x_2) = x_1^{\alpha_1} x_2^{\alpha_2}, \qquad \alpha_1,\ \alpha_2 > 0$$

は必ず準凹関数であることを示せ．また

$$\alpha_1 + \alpha_2 \leq 1$$

である時，そしてその時のみ，凹関数となることを証明せよ．

以下では C は開集合であるとする．

3.10 定理. C を凸集合とする．C 上で定義される関数 $F : C \to \boldsymbol{R}$ について，F が準凸関数である時，そしてその時のみ

$$\forall y \in \boldsymbol{R}^n : \ \ \nabla F(x) \cdot y = 0 \quad \Rightarrow \quad y^\top \nabla^2 F(x) y \geq 0 \tag{3.3}$$

が成立する．

勾配ベクトルが張る空間 $\langle \nabla F(x) \rangle$ の直交補空間を $\langle \nabla F(x) \rangle^\perp$ と書けば，定理 3.10 における (3.3) の条件は，ヘッセ行列 $\nabla^2 F(x)$ が $\langle \nabla F(x) \rangle^\perp$ 上で正値半定符号であることを意味する.[9]

3.11 定理. C を凸集合とする．C 上で定義される関数 $F : C \to \boldsymbol{R}$ について

1. F が準凹関数
2. $\forall y \in \boldsymbol{R}^n : \nabla F(x) \cdot y = 0 \Rightarrow y^\top \nabla^2 F(x) y \leq 0$

は同値である．

今, C を凸集合とする．C 上で定義される 1 回連続微分可能関数 $F : C \to \boldsymbol{R}$ について

$$F \text{ が凸関数} \Rightarrow \forall x\ \forall y : F(y) - F(x) \geq \nabla F(x) \cdot (y - x)$$

が成り立つ．このことから, 以下の二つのことがわかる．$\forall x\ \forall y$:

(i) $\nabla F(x) \cdot (y - x) > 0 \Rightarrow F(y) > F(x)$. (I.e., $F(y) \leq F(x) \Rightarrow \nabla F(x) \cdot (y - x) \leq 0$.)

(ii) $\nabla F(x) \cdot (y - x) \geq 0 \Rightarrow F(y) \geq F(x)$. (I.e., $F(y) < F(x) \Rightarrow \nabla F(x) \cdot (y - x) < 0$.)

(i) は準凸性と同値である．また, (ii) を擬凸性と呼ぶ.

3.8 定義 (擬凸関数). C を凸集合とする．C 上で定義される関数 $F : C \to \boldsymbol{R}$ について

$$\forall x, y \in C : \nabla F(x) \cdot (y - x) \geq 0 \Rightarrow F(y) \geq F(x)$$

が成り立つ時，F は擬凸関数 (pseudo-convex function) であるという．

練習問題 3.7. (i) は準凸性と同値であることを示せ.

[9] F が凸関数であれば，ヘッセ行列 $\nabla^2 F(x)$ は $\boldsymbol{R}^n$ 上で正値半定符号となることを想起せよ．

3.1 命題.

1. (i) $\Rightarrow$ (ii). つまり, F が擬凸関数 $\Rightarrow$ F は準凸関数.
2. $\forall x\ \nabla F(x) \neq 0$ ならば, (i) $\Rightarrow$ (ii).

なお, (i) $\Rightarrow$ (ii) が成立するためには, $\nabla F(x) \neq 0$ であることは必要不可欠である.

例 $C = \boldsymbol{R}$, $F(x) = x^3$

3.1 命題の証明. (i) の証明: $x \in C,\ y \in C,\ \nabla F(x) \cdot (y - x) > 0$ とすると, 十分小さな $\varepsilon > 0$ に対して

$$z = (1-\varepsilon)x + \varepsilon y = x + \varepsilon(y - x)$$

と定義すれば

$$F(z) > F(x) \tag{3.4}$$
$$\nabla F(z) \cdot (y - z) > 0. \tag{3.5}$$

従って, (3.5) と (ii) により, $F(y) \geq F(z)$. よって, (3.4) により $F(y) > F(x)$. これで (i) が示された.

(ii) の証明: $x \in C,\ y \in C,\ \nabla F(x) \cdot (y - x) \geq 0$ とすると, 十分小さな $\varepsilon > 0$ に対して

$$z = y + \varepsilon \nabla F(x)$$

と定義すれば, C が開集合であることより $z \in C$ となり, なおかつ $\nabla F(x) \neq 0$ より

$$\nabla F(x) \cdot (z - x) > 0$$

となる. よって, (i) より $F(z) > F(x)$. $\varepsilon \to 0$ の時, $z \to y$ なので $F(y) \geq F(x)$. よって, (ii) が示された.

□

まとめると，連続微分可能な関数 $F : C \to \boldsymbol{R}$ について

(i) F が凸関数 $\Rightarrow$ F が擬凸関数 $\Rightarrow$ F が準凸関数

(ii) $\forall x \in C : \nabla F(x) \neq 0$ とすると, F が擬凸関数 $\Leftrightarrow$ F が準凸関数

となる.

3.2 ミンコフスキー＝ファルカスの補題と分離超平面定理

3.2.1 ミンコフスキー＝ファルカスの補題

3.12 定理 (ミンコフスキー＝ファルカスの補題). J 行 N 列の行列 $A \in \boldsymbol{R}^{J \times N}$，及び n 次元ベクトル $b \in \boldsymbol{R}^N$ について

1. $\exists z \in \boldsymbol{R}^J_+ : \quad b^\top = z^\top A$
2. $\exists x \in \boldsymbol{R}^N : \quad Ax \in \boldsymbol{R}^J_+ \quad$ かつ $\quad b^\top x < 0$

の 2 条件のいずれかが，排他的に成立する.

条件 1 の意味を考えてみよう．これは連立方程式

$$A^\top z = b$$

の解 z が存在し，なおかつ z の全ての要素が非負の値をとることを意味する．

ここで

$$z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_J \end{pmatrix} \in \boldsymbol{R}^J_+ \quad \text{及び} \quad A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_J \end{pmatrix} \in \boldsymbol{R}^{J\times N}, \quad a_j = (a_{j1}a_{j2}\dots a_{jN}),\ j = 1, 2, \dots, J$$

のようにベクトル z と行列 A を表すとすれば，$b^\top = z^\top A$ は

$$b^\top = z^\top A = z_1 a_1 + z_2 a_2 + \cdots + z_J a_J \qquad (z_j \geq 0,\ j = 1, 2, \dots, J)$$

と書けるから，条件 1 が満たされることは，$b^\top$ が行列 A の行ベクトルの非負結合で表せることに他ならない．幾何的には，これは行ベクトルがつくる錘 (cone) に b が含まれていることを意味する．

次に条件 2 の意味を考えてみよう．ここで $Ax \in \boldsymbol{R}^J_+$ が

$$Ax \in \boldsymbol{R}^J_+ \quad \Leftrightarrow \quad a_j \cdot x \geq 0,\ j = 1, 2, \dots, J$$

を意味することに注意すれば，条件 2 は

$$a_1 \cdot x \geq 0,\ \ a_2 \cdot x \geq 0,\ \ \dots\ \ , a_J \cdot x \geq 0 \quad \text{かつ} \quad b \cdot x < 0$$

を満たす $x \in \boldsymbol{R}^N$ が存在することに他ならない．つまり，行列 A のいずれの行ベクトルともなす角が 90 度以下で，なおかつ b とのなす角は 90 度よりも大きいようなベクトル x が存在することを意味する．これは幾何的には，行ベクトルがつくる錘と点 b とを分断する超平面の存在を保証するものである．

3.9 定義 (射影). $a \in \boldsymbol{R}^N$ 及び $b \in \boldsymbol{R}^N$ について，$a \cdot b \neq 0$ が成立しているとする．また

$$a^\perp = \left\{ x \in \boldsymbol{R}^N \;\middle|\; a \cdot x = 0 \right\}$$

と定義しよう．この時，任意の $x \in \boldsymbol{R}^N$ に対して

$$x = v + \lambda b$$

を満たす唯一の $v \in a^\perp$ と $\lambda \in \boldsymbol{R}$ が存在し，このような v を，x の $a^\perp$ への b に沿った射影という．

練習問題 3.8. x の $a^\perp$ への b に沿った射影 v について

$$v = x - \frac{a \cdot x}{a \cdot b} b$$

が成り立つことを示せ．

3.12 **定理の証明.** 最初に，二つの条件が同時に成立することはないことを背理法を用いて示そう．背理法の仮定として $z \in \boldsymbol{R}^J_+$ 及び $x \in \boldsymbol{R}^N$ が存在し，二つの条件が同じに満たされているとする．条件 1 により，この時 z について

$$b^\top = z^\top A$$

が成立するので，両辺に右から x を掛けることにより

$$b^\top x = z^\top A x \tag{3.6}$$

を得る．一方，z の要素が全て非負であることに注意すれば，条件 2 により

$$z^\top Ax = z_1 a_1 \cdot x + z_2 a_2 \cdot x + \cdots + z_J a_J \cdot x \geq 0 \quad \text{かつ} \quad b^\top x < 0$$

が成立する．しかしこれは，(3.6) と矛盾する．よって二つの条件が同時に満たされることはない．後は，条件 1 が満たされない時に条件 2 が必ず成立することを示せば十分である．これを J に関する帰納法を用いて示そう．$J = 1$ の場合についの証明のステップは練習問題とし，$J \geq 2$ について考える．帰納法の仮定として，$J-1$ について条件 1 が満たされない時に条件 2 が必ず満たされるとしよう．この仮定のもとで

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{J-1} \\ a_J \end{pmatrix} \in \boldsymbol{R}^{J \times N}, \quad A' = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{J-1} \end{pmatrix} \in \boldsymbol{R}^{(J-1) \times N}$$

のような A と A' を考え，行列 A について条件 1 が満たされていないとする．この時，行列 A' についても条件 1 は成立しないから

$$b^\top = z'^\top A' \tag{3.7}$$

を満たすような $z' \in \boldsymbol{R}_+^{J-1}$ は存在しない．よって帰納法の仮定により

$$A'x' \in \boldsymbol{R}_+^{J-1} \quad \text{かつ} \quad b^\top x' < 0$$

を満たす $x' \in \boldsymbol{R}^N$ が存在する．もし $a_J \cdot x' \geq 0$ であれば，この x' は

$$Ax' \in \boldsymbol{R}_+^J \quad \text{かつ} \quad b^\top x' < 0$$

を満たすことになるので条件 2 が成立し，証明は完了する．よって以下では，$a_J \cdot x' < 0$ の場合を考えることにする．ここで，$a_1, a_2, \ldots, a_{J-1}$ 及び $b^\top$ の $x'^\perp$ への a_J に沿った射影をそれぞれ，$\widehat{a}_1, \widehat{a}_2, \ldots, \widehat{a}_{J-1}$ 及び $\widehat{b^\top}$（いずれも列ベクトル）と書き，行列 $\widehat{A}$ を

$$\widehat{A} = \begin{pmatrix} \widehat{a}_1 \\ \widehat{a}_2 \\ \vdots \\ \widehat{a}_{J-1} \end{pmatrix} \in \boldsymbol{R}^{(J-1) \times N}$$

と定義する．このように定義された $\widehat{b^\top}$ 及び $\widehat{A}$ について

$$\widehat{b^\top} = w^\top \widehat{A} \tag{3.8}$$

を満たす $w \in \boldsymbol{R}_+^{J-1}$ が存在しないことを，背理法によって示そう．今，式を満たす $w \in \boldsymbol{R}_+^{J-1}$ が存在すると仮定して，その成分を

$$w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_{J-1} \end{pmatrix} \in \boldsymbol{R}_+^{J-1}$$

と表せば

$$\begin{aligned} \widehat{b^\top} &= w^\top \widehat{A} \\ &= \sum_{j=1}^{J-1} w_j \widehat{a}_j \end{aligned}$$

$\widehat{a}_j$ は a_j の $x'^{\perp}$ への a_J に沿った射影であるから，練習問題 3.8 の結果を用いることにより

$$\widehat{a}_j = a_j - \frac{x' \cdot a_j}{x' \cdot a_J} a_J, \quad j = 1, 2, \ldots, J-1$$

と書ける．よって

$$\begin{aligned} \widehat{b}^{\top} &= \sum_{j=1}^{J-1} w_j \widehat{a}_j \\ &= \sum_{j=1}^{J-1} w_j \left(a_j - \frac{x' \cdot a_j}{x' \cdot a_J} a_J \right) \\ &= \sum_{j=1}^{J-1} w_j a_j - \left(\sum_{j=1}^{J-1} w_j \frac{x' \cdot a_j}{x' \cdot a_J} \right) a_J \end{aligned} \tag{3.9}$$

さらに $\widehat{b}^{\top}$ も $b^{\top}$ の $x'^{\perp}$ への a_J に沿った射影であるから

$$\widehat{b}^{\top} = b^{\top} - \frac{x' \cdot b}{x' \cdot a_J} a_J$$

と書け，したがって (3.9) は

$$\begin{aligned} b^{\top} &= \sum_{j=1}^{J-1} w_j a_j + \left(\frac{x' \cdot b}{x' \cdot a_J} - \sum_{j=1}^{J-1} w_j \frac{x' \cdot a_j}{x' \cdot a_J} \right) a_J \\ &= w^{\top} A' + \left(\frac{x' \cdot b}{x' \cdot a_J} - \sum_{j=1}^{J-1} w_j \frac{x' \cdot a_j}{x' \cdot a_J} \right) a_J \end{aligned} \tag{3.10}$$

$a_J \cdot x' < 0$ かつ $w_j \geq 0, j = 1, 2, \ldots J-1$ であり，また帰納法の仮定により

$$A'x' \in \boldsymbol{R}_{+}^{J-1} \quad \text{かつ} \quad b^{\top} x' < 0$$

でったことに注意すれば

$$\frac{x' \cdot b}{x' \cdot a_J} - \sum_{j=1}^{J-1} w_j \frac{x' \cdot a_j}{x' \cdot a_J} \geq 0$$

である．よって (3.10) において

$$w_J = \frac{x' \cdot b}{x' \cdot a_J} - \sum_{j=1}^{J-1} w_j \frac{x' \cdot a_j}{x' \cdot a_J}$$

と置き

$$z = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_{J-1} \\ w_J \end{pmatrix}$$

と z を定義すれば，この z の成分はすべて非負で

$$\begin{aligned} b^{\top} &= w^{\top} A' + w_J a_J \\ &= z^{\top} A \end{aligned}$$

を満足する．これは，A について条件 1 が成立しないとした当初の仮定と矛盾する．よって，$\widehat{b}^\top$ 及び $\widehat{A}$ について

$$\widehat{b}^\top = w^\top \widehat{A}$$

を満たす $w \in \boldsymbol{R}_+^{J-1}$ は存在せず，行列 $\widehat{A}$ についても条件 1 は成立しない．したがって，帰納法の仮定により行列 $\widehat{A}$ は条件 2 を満たすので

$$\widehat{A}\widehat{x} \in \boldsymbol{R}_+^{J-1} \tag{3.11}$$

$$\widehat{b}^\top \widehat{x} < 0 \tag{3.12}$$

を満たす $\widehat{x} \in \boldsymbol{R}^N$ が存在することになる．

最後に，$\widehat{x}$ の $a_J^\perp$ への x' に沿った射影を x と表せば

$$Ax \in \boldsymbol{R}_+^J \quad \text{かつ} \quad b^\top x < 0$$

が成立し，A について条件 2 が満たされることを確認する．まず，射影の定義により $\widehat{a}_j$, $j = 1, 2, \ldots J-1$ 及び $\widehat{b}$ について

$$\exists \lambda \in \boldsymbol{R}: \quad \widehat{a}_j = a_j + \lambda a_J \tag{3.13}$$

$$\widehat{a}_j \cdot x' = 0 \tag{3.14}$$

$$\exists \lambda \in \boldsymbol{R}: \quad \widehat{b}^\top = b^\top + \lambda a_J \tag{3.15}$$

$$\widehat{b}^\top \cdot x' = 0 \tag{3.16}$$

が成立する．また，x について

$$\exists \lambda \in \boldsymbol{R}: \quad x = \widehat{x} + \lambda x' \tag{3.17}$$

$$a_J \cdot x = 0 \tag{3.18}$$

が成り立つ．ここで，(3.13) と (3.18) により

$$a_j \cdot x = \widehat{a}_j \cdot x \tag{3.19}$$

また (3.14) と (3.17)，(3.11) により

$$\widehat{a}_j \cdot x = \widehat{a}_j \cdot \widehat{x} \geq 0 \tag{3.20}$$

が成立するから，(3.19) と (3.20) をあわせれば

$$a_j \cdot x \geq 0 \quad j = 1, 2, \ldots J-1 \tag{3.21}$$

が成り立つ．よって (3.18) と (3.21) から

$$Ax \in \boldsymbol{R}_+^J$$

である．一方，(3.15) と (3.18) により

$$b^\top \cdot x = \widehat{b}^\top \cdot x \tag{3.22}$$

また (3.16) と (3.17)，(3.12) により

$$\widehat{b}^\top \cdot x = \widehat{b}^\top \cdot \widehat{x} < 0 \tag{3.23}$$

が成立するから (3.22) と (3.23) をあわせれば

$$b^\top \cdot x < 0$$

も確認できる． □

3.2.2 分離超平面定理

ミンコフスキー=ファルカスの補題は, ベクトル b が列ベクトル a_j, $j = 1, 2, \ldots, J$ が張る錘（錘は凸閉集合）に含まれないならば，b と錘とを分離する超平面が存在することを保証するものである．この補題における錘を凸閉集合一般の場合に拡張したものが，以下の分離超平面定理である．

3.13 定理 (**分離超平面定理**)**.** $\boldsymbol{R}^N$ の凸で閉な部分集合 C，及び $\boldsymbol{R}^N$ の C に含まれないベクトル $b \in \boldsymbol{R}^N \setminus C$ について

$$\forall c \in C : \quad x \cdot c \geq d > x \cdot b$$

を満たす $x \in \boldsymbol{R}^N$ 及び $d \in \boldsymbol{R}$ が存在する．

練習問題 3.9. 定理 3.13 を用いてミンコフスキー=ファルカスの補題を証明せよ．($d = 0$ と取れることを示すことが重要である.)

3.14 定理. 行列 $A \in \boldsymbol{R}^{J \times N}$ について

1. $\exists z \in \boldsymbol{R}^J_+ \setminus \{0\} : \quad z^\top A = 0$
2. $\exists x \in \boldsymbol{R}^N : \quad Ax \in \boldsymbol{R}^J_{++}$

の 2 条件のいずれかが，排他的に成立する．

練習問題 3.10. 定理 3.14 における条件 1 と条件 2 とが，同時に成立することはないことを示せ．

3.14 **定理の証明.** 練習問題 3.10 により，二つの条件は同じに成立しないことは示されているので，以下では条件 1 が成立しない時に条件 2 が必ず成立することを確認すれば十分である．まず $\boldsymbol{R}^J$ 上のベクトル e を

$$e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \boldsymbol{R}^J$$

と定義し，これを用いて行列 $\widetilde{A}$ を

$$\widetilde{A} = (A, e) \in \boldsymbol{R}^{J \times (N+1)}$$

と定義しよう．この時条件 1 は

$$z^\top \widetilde{A} = (\underbrace{0, 0, \ldots, 0}_{N}, 1)$$

を満たす $z \in \boldsymbol{R}^J_+$ が存在することと同値である．よってミンコフスキー=ファルカスの補題により，条件 1 が満たされていない時

$$\widetilde{A}\widetilde{x} \in \boldsymbol{R}^J_+ \tag{3.24}$$

$$(0, 0, \ldots, 0, 1)\widetilde{x} < 0 \tag{3.25}$$

を満たす $\widetilde{x} \in \boldsymbol{R}^{N+1}$ が存在する．ここで

$$\widetilde{x} = \begin{pmatrix} x \\ x_{N+1} \end{pmatrix} \in \boldsymbol{R}^n$$

と $x \in \boldsymbol{R}^N$ を置けば

$$\begin{aligned}\widetilde{A}\widetilde{x} &= (A, e)\begin{pmatrix} x \\ x_{N+1} \end{pmatrix} \\ &= Ax + x_{N+1}e \end{aligned} \tag{3.26}$$

と書ける．ここで (3.25) により

$$x_{N+1} = (0, 0, \ldots, 0, 1)\widetilde{x} < 0$$

であるから (3.26) から

$$\widetilde{A}\widetilde{x} < Ax$$

が導かれる．よって (3.24) と合わせれば $Ax \in \boldsymbol{R}^J_{++}$ が言え，条件 2 が満たされていることが分かる． □

3.14 定理の意味を考えよう．まず条件 1 について，これは連立方程式

$$A^\top z = 0$$

に自明でない解 z が存在し，なおかつ z の全ての要素が非負の値をとることを意味する．幾何的には，この条件は, A の行ベクトルがいずれもゼロでない限り, それらが張る錘に少なくとも一本の直線が含まれることに対応する．逆に言えば，条件 1 が満たされていないとは，錘が必ず原点でとがっていることに他ならない．

一方，条件 2 は

$$a_1 \cdot x > 0,\ \ a_2 \cdot x > 0,\ \ \ldots\ \ , a_J \cdot x > 0$$

なる $x \in \boldsymbol{R}^N$ が存在することに他ならないから，行列 A のいずれの行ベクトルともなす角が 90 度未満となるベクトル x が存在することを意味する．これは幾何的には，空間 $\boldsymbol{R}^N$ を，行ベクトルがつくる錘を含む部分空間と錘を含まない部分空間とに厳密に分断する，一本の直線を引くことができることを意味する．よって, この定理の主張は, A の行ベクトルが張る錘が原点でとがっている時, およびその時のみ, 上述のような直線を引くことができるということである．

3.3 制約付き最大化問題

L，M，N を正の整数とする．一般に制約付き最大化問題は，$\boldsymbol{R}^L$ の部分集合 X と N 個の目的関数 $f_n : X \to \boldsymbol{R}$ $(n = 1, 2, \ldots, N)$，及び M 個の制約式 $g_m : X \to \boldsymbol{R}$ $(m = 1, 2, \ldots, M)$ の組 $(X, f_1, f_2, \ldots, f_N, g_1, g_2, \ldots, g_M)$ によって定義され，

$$\begin{aligned} &\max_{x \in X} (f_1(x), f_2(x), \ldots, f_N(x)) \\ &\text{subject to} \quad g_1(x) \geq 0, \\ &\qquad\qquad\quad g_2(x) \geq 0, \\ &\qquad\qquad\qquad \vdots \\ &\qquad\qquad\quad g_M(x) \geq 0. \end{aligned} \tag{3.27}$$

のように表される．定義域上の点 $x^* \in X$ が全ての制約式を満たし，なおかつ

$$\begin{aligned} \forall m :&\quad g_m(x) \geq 0, \\ \forall n :&\quad f_n(x) \geq f_n(x^*), \\ \exists n :&\quad f_n(x) > f_n(x^*) \end{aligned}$$

を満たす $x \in X$ が存在しない時，x^* はこの制約付き最大化問題の解であると言う．以下では X を開集合とし，f_n 及び g_m を連続微分可能であると仮定して話を進める．

3.3.1 クーン＝タッカー条件

3.15 定理 (**クーン＝タッカー必要条件**)**.** $x^* \in \boldsymbol{X}$ が制約付き最大化問題 (3.27) の解であるならば，$N+M$ 個の非負実数からなるベクトル $(\mu_1, \mu_2, \ldots, \mu_N, \lambda_1, \lambda_2, \ldots, \lambda_M) \in \boldsymbol{R}_+^{N+M}$ が存在し

1. $\mu_1, \ldots, \mu_N, \lambda_1, \ldots, \lambda_M$ のうち，少なくとも 1 つは厳密に正;
2. $\forall m: \quad g_m(x^*) > 0 \quad \Rightarrow \quad \lambda_m = 0$;
3. $\displaystyle\sum_{n=1}^{N} \mu_n \nabla f_n(x^*) + \sum_{m=1}^{M} \lambda_m \nabla g_m(x^*) = 0.$

の 3 条件が同時に成立する．

3.15 **定理の証明 (スケッチ).** まずは M 本の制約式を，等号で満たされているもの ($g_m(x^*) = 0$) と，等号で満たされていないもの ($g_m(x^*) > 0$) とに分けて考える．制約式を適当に並べ替えて添え字を振り直すことによって，最初の $K\,(\leq M)$ 本の制約式が等号で成立しているものと一般性を失うことなく仮定できる．その上で，

$$\begin{aligned} \forall n: &\qquad \nabla f_n(x^*) \cdot v > 0 \\ \forall m \leq K: &\qquad \nabla g_m(x^*) \cdot v > 0 \end{aligned}$$

を満たすベクトル $v \in \boldsymbol{R}^L$ が存在しないことを，背理法を用いて示すことができる．

すると，これは

$$\begin{pmatrix} \nabla f_1(x^*) \\ \vdots \\ \nabla f_N(x^*) \\ \nabla g_1(x^*) \\ \vdots \\ \nabla g_K(x^* \end{pmatrix} v \in \boldsymbol{R}_{++}^{N+K}$$

なるベクトル $v \in \boldsymbol{R}^L$ が存在しないことを意味するから，3.14 定理により

$$(\mu_1, \ldots, \mu_N, \lambda_1, \ldots, \lambda_K) \begin{pmatrix} \nabla f_1(x^*) \\ \vdots \\ \nabla f_N(x^*) \\ \nabla g_1(x^*) \\ \vdots \\ \nabla g_K(x^* \end{pmatrix} = 0$$

すなわち

$$\sum_{n=1}^{N} \mu_n \nabla f_n(x^*) + \sum_{m=1}^{K} \lambda_m \nabla g_m(x^*) = 0$$

を満たすベクトル $(\mu_1, \mu_2, \ldots, \mu_N, \lambda_1, \lambda_2, \ldots, \lambda_K) \in \boldsymbol{R}_+^{N+K} \setminus \{0\}$ が存在する．

最後に，このベクトルに $M-K$ 個の実数

$$\lambda_m = 0, \quad m > K$$

を加えて $N+M$ 次元ベクトル $(\mu_1,\dots,\mu_N,\lambda_1,\dots\lambda_K,\lambda_{K+1},\dots,\lambda_M) \in \boldsymbol{R}_+^{N+K}$ を作れば，この $N+M$ 次元ベクトルについて 3.15 定理の条件が全て成立していることを確認できる． □

3.1 注意. クーン＝タッカー必要条件は，μ 及び λ に関する 0 次同次で成立する．

この注意は，$(\mu_1,\dots,\mu_N,\lambda_1,\dots,\lambda_M)$ が 3.15 定理における三つ条件を満たしているなら，任意の $t>0$ について $(t\mu_1,\dots,t\mu_N,t\lambda_1,\dots,t\lambda_M)$ も同様に条件を満たすことを意味する．よって，$(\mu_1,\dots,\mu_N,\lambda_1,\dots,\lambda_M)$ の中の 0 でない成分を一つだけ選び，それを 1 に正規化してもよいことになる．多くの教科書で $\mu_1=1$ と置いているのは，このような性質があるためである．ただし，$\mu_1 \neq 0$ であることは保証されていないため，常に $\mu_1=1$ と仮定できるわけではない[10].

ちなみに 3.15 定理において，未知数は x について L 個，μ について N 個，λ について K 個である．0 次同次性を考慮すれば実質的な未知数の数は一つ減るので，全体では $L+N+K-1$ 個となる．一方，方程式の本数は

$$\sum_{n=1}^{N} \mu_n \nabla f_n(x^*) + \sum_{m=1}^{K} \lambda_m \nabla g_m(x^*) = 0$$

について L 本,

$$\begin{aligned} g_1(x^*) &= 0, \\ g_2(x^*) &= 0, \\ &\vdots \\ g_K(x^*) &= 0. \end{aligned}$$

について K 本であるから，合計 $L+K$ 本である．よって解の自由度は $N-1$ であり，目的関数の数が一つであれば解は一意的に定められることが分かる．

3.16 定理 (クーン＝タッカー十分条件). 制約付き最大化問題 (3.27) において，$\boldsymbol{X}$ を凸集合，f_n を擬凹関数，g_m を準凹関数としよう．この時，$x^* \in \boldsymbol{X}$ 及び $(\mu_1,\dots,\mu_N,\lambda_1,\dots,\lambda_M) \in \boldsymbol{R}_+^{N+M}$ が存在し

1. $\forall m: \quad g_m(x^*) \geq 0;$
2. $(\mu_1,\dots,\mu_N,\lambda_1,\dots,\lambda_M) \in \boldsymbol{R}_{++}^{N};$
3. $\forall m: \quad g_m(x^*) > 0 \quad \Rightarrow \quad \lambda_m = 0;$
4. $\displaystyle\sum_{n=1}^{N} \mu_n \nabla f_n(x^*) + \sum_{m=1}^{M} \lambda_m \nabla g_m(x^*) = 0.$

が成立していれば，x^* は問題 (3.27) の解である．

3.16 定理の証明 (スケッチ). まず M 本の制約式を，等号で成立しているもの $(g_m(x^*)=0)$ と等号で成立していないもの $(g_m(x^*)>0)$ とに分けて考える．制約式を適当に並べ替えて添え字を振り直

[10] この点については **??** 節を参照.

すことによって，最初の $K\ (\leq M)$ 本の制約式が等号で成立しているものと一般性を失うことなく仮定できる．ここで背理法の仮定として，x^* が問題の解でないとすれば

$$\begin{aligned}\forall n: &\quad \nabla f_n(x^*)\cdot v \geq 0\\ \exists n: &\quad \nabla f_n(x^*)\cdot v > 0\\ \forall m\leq K: &\quad \nabla g_m(x^*)\cdot v \geq 0\end{aligned}$$

を満たすベクトル $v\in \boldsymbol{R}^L$ が存在する．この時，条件 1，条件 2 及び条件 3 により

$$\sum_{n=1}^{N}\mu_n\nabla f_n(x^*)\cdot v+\sum_{m=1}^{M}\lambda_m\nabla g_m(x^*)\cdot v>0$$

が成立する．しかしこの不等式の左辺は

$$\left(\sum_{n=1}^{N}\mu_n\nabla f_n(x^*)+\sum_{m=1}^{M}\lambda_m\nabla g_m(x^*)\right)\cdot v$$

であるから，条件 4 と矛盾する．よって x^* は問題の解である． □

練習問題 3.11. x^* が解であるための十分条件として，3.16 定理における条件 2 と条件 3 が不可欠であることを示す例を挙げよ．

3.16 定理の中で設けられている定義域が開集合でなければならないという条件は，具体的な問題を解くにあたって障害となる場合がある．しかしこの点は，効用関数の定義域を広げ，適当な制約を新たに加えることによって解決できる．

3.1 例. 一人の消費者が予算制約の下で 2 種類の財を購入し，効用を最大化する問題を考えてみよう．今，消費者の効用関数 $u:\boldsymbol{R}_+^2\to\boldsymbol{R}$ が

$$u(x_1,x_2)=(x_1+1)(x_2+1)$$

によって定義されているとする．それぞれの財価格を p_1,p_2，資産を w で表せば，効用最大化問題は

$$\begin{aligned}\max_{x\in\boldsymbol{R}_+^2} &\quad (x_1+1)(x_2+1)\\ \text{subject to} &\quad w-p_1x_1-p_2x_2\geq 0.\end{aligned}\tag{3.28}$$

と書ける．しかし効用関数の定義域 $\boldsymbol{R}_+^2$ は閉集合であるから，このままではクーン＝タッカー十分条件を適用することはできない．そこで，定義域を開集合 $\boldsymbol{X}=\{x\in\boldsymbol{R}^2\mid x_1>-1,\ x_2>-1\}$ へと拡張し，制約式に $x_1\geq 0$ と $x_2\geq 0$ を追加することを考える．つまり

$$\begin{aligned}\max_{x\in\boldsymbol{X}} &\quad (x_1+1)(x_2+1)\\ \text{subject to} &\quad w-p_1x_1-p_2x_2\geq 0,\\ &\quad x_1\geq 0,\\ &\quad x_2\geq 0.\end{aligned}\tag{3.29}$$

のように問題を書き直すのである．これにより，問題 (3.29) を解くことはもとの問題 (3.28) を解くことと等しく，なおかつクーン＝タッカー十分条件を適用できるようになる．

3.1 例では，効用関数の定義域を広げ，適当な制約を新たに加えることによって問題を解決したが，逆に定義域を限定することによってクーン＝タッカー十分条件を適用できるようになる場合もある．

3.2 例. 3.1 例と同様に，一人の消費者が予算制約の下で 2 種類の財を購入し，効用を最大化する問題を考えてみよう．ただし今度は，消費者の効用関数 $u : \boldsymbol{R}^2_+ \to \boldsymbol{R}$ が

$$u(x_1, x_2) = \sqrt{x_1} + \sqrt{x_2}$$

によって定義されているものとする．それぞれの財価格を p_1, p_2，資産を w で表せば，効用最大化問題は

$$\begin{aligned} &\max_{x \in \boldsymbol{R}^2_+} && u(x_1, x_2) = \sqrt{x_1} + \sqrt{x_2} \\ &\text{subject to} && w - p_1 x_1 - p_2 x_2 \geq 0. \end{aligned} \tag{3.30}$$

と書けるが，効用関数の定義域 $\boldsymbol{R}^2_+$ は閉集合であるからこのままではクーン＝タッカー十分条件を適用することはできない．しかし効用関数の形状により

$$\frac{\partial u}{\partial x_1}(x_1, x_2) \to \infty \quad (x_1 \to 0)$$
$$\frac{\partial u}{\partial x_2}(x_1, x_2) \to \infty \quad (x_2 \to 0)$$

であるから，$x^* = (x_1^*, x_1^*)$ が解であれば，必ず $x_1^* > 0$ かつ $x_1^* > 0$，すなわち $x^* \in \boldsymbol{R}^2_{++}$ でなければならない．これにより，定義域を開集合 $\boldsymbol{R}^2_{++}$ に限定した効用最大化問題

$$\begin{aligned} &\max_{x \in \boldsymbol{R}^2_{++}} && u(x_1, x_2) = \sqrt{x_1} + \sqrt{x_2} \\ &\text{subject to} && w - p_1 x_1 - p_2 x_2 \geq 0. \end{aligned} \tag{3.31}$$

の解は，もとの問題 (3.30) の解と等しいことが言える．よって，問題 (3.31) にクーン＝タッカー十分条件を適用することで，もとの問題の解を導くことができる．

3.1 例と 3.2 例で用いた二つの方法を組み合わせる必要がある場合もある．この点を，以下の例と練習問題によって確認しよう．

3.3 例. 一人の消費者が予算制約の下で 2 種類の財を購入し，効用を最大化する問題を考えてみよう．財価格はいずれの財についても 1，資産は $w(> 0)$，消費者の効用関数 $u : \boldsymbol{R}^2_+ \to \boldsymbol{R}$ を

$$u(x_1, x_2) = x_1^{\frac{1}{2}} e^{x_2}$$

によって定義する（効用関数の定義域が開集合ではないことに注意）．この時，消費者の効用最大化問題は

$$\begin{aligned} &\max_{x \in \boldsymbol{R}^2_+} && u(x) \\ &\text{subject to} && w - x_1 - x_2 \geq 0. \end{aligned} \tag{3.32}$$

と書ける．

練習問題 3.12. $x \in \boldsymbol{R}^2_+$ が効用最大化問題 (3.32) の解であれば，x_1 は厳密に正の値をとることを示せ．

練習問題 3.12 により，効用関数の定義域を $\left\{x \in \boldsymbol{R}^2 \;\middle|\; x_1 > 0,\ x_2 \geq 0\right\}$ に限定して考えても問題の解は変わらないことになる．

その上で，今度は効用関数 u について，その定義域を

$$\boldsymbol{X} = \left\{x \in \boldsymbol{R}^2 \;\middle|\; x_1 > 0\right\}$$

に広げることを考えよう．関数 e^{x_2} は任意の $x_2 \in \boldsymbol{R}$ について定義されるから，効用関数全体も $\boldsymbol{X}$ 上で定義し直して構わない．関数 $g_1 : \boldsymbol{X} \to \boldsymbol{R}$ を

$$g_1(x) = w - x_1 - x_2$$

で定義すれば，定義域を拡張した効用最大化問題は

$$\begin{aligned} \max_{x \in \boldsymbol{X}} \quad & u(x) \\ \text{subject to} \quad & g_1(x) \geq 0. \end{aligned} \tag{3.33}$$

と書ける．新たな定義域 $\boldsymbol{X}$ は開集合であるから，この問題 (3.33) にはクーン＝タッカー十分条件を適用することが可能である．ただし，ここでは無条件に定義域を広げているため，問題 (3.33) の解が当初の問題の解と一致する保証はない．そこで，新たな制約を付け加える必要が生じる．

練習問題 3.13. 効用最大化問題

$$\begin{aligned} \max_{x \in \boldsymbol{X}} \quad & u(x) \\ \text{subject to} \quad & g_1(x) \geq 0, \\ & g_2(x) \geq 0. \end{aligned} \tag{3.34}$$

が，当初の効用最大化問題 (3.32) と等しくなるように，関数 $g_2 : \boldsymbol{X} \to \boldsymbol{R}$ を定義せよ．

練習問題 3.14. 練習問題 3.13 を踏まえて，問題 (3.34) にクーン＝タッカー十分条件を適用することで，当初の効用最大化問題 (3.32) の解を求めよ．

3.3.2 包絡線定理

K，L，M を正の整数とし，$\boldsymbol{X}$ を $\boldsymbol{R}^L$ の部分集合，$\boldsymbol{P}$ を $\boldsymbol{R}^K$ の部分集合とする．また $f : \boldsymbol{X} \times \boldsymbol{P} \to \boldsymbol{R}$ 及び $g_m : \boldsymbol{X} \times \boldsymbol{P} \to \boldsymbol{R}\,(m = 1, 2, \ldots, M)$ を，$\boldsymbol{X} \times \boldsymbol{P}$ 上で定義される 2 回連続微分可能な実数値関数とする．$p \in \boldsymbol{P}$ を所与として，制約付き最大化問題

$$\begin{aligned} \max_{x \in \boldsymbol{X}} \quad & f(x, p) \\ \text{subject to} \quad & g_1(x, p) \geq 0, \\ & g_2(x, p) \geq 0, \\ & \vdots \\ & g_M(x, p) \geq 0. \end{aligned} \tag{3.35}$$

を考える．

問題 (3.35) についてパラメーター p の値を変化させれば，制約付き最大化問題の一つのクラスを考えることができ，この時，集合 $\boldsymbol{P}$ をそのクラスのパラメーター空間という．以下では，任意の $p \in \boldsymbol{P}$ に対して制約付き最大化問題 (3.35) の解が一意に定まると仮定して話を進める．

3.10 定義 (Policy Function). 任意の $p \in \boldsymbol{P}$ に対して，制約付き最大化問題 (3.35) の解が一意に定まると仮定すれば，その解は $\boldsymbol{P}$ 上で定義される関数 $a : \boldsymbol{P} \to \boldsymbol{X}$ となる．この時，$a(p)$ を Policy Function という．

3.11 定義 (Value Function). Policy Function を目的関数に代入した関数 $b : \boldsymbol{P} \to \boldsymbol{R}$ を

$$b(p) = f(a(p), p)$$

によって定義する．この時，$b(p)$ を Value Function という．

以下では勾配ベクトルを行ベクトルとして扱う．

3.17 定理. パラメーター $p^* \in \boldsymbol{P}$ を所与として，クーン＝タッカー乗数 $(1, \lambda_1, \ldots, \lambda_K) \in \boldsymbol{R}^{1+K}_{++}$ とベクトル $x^* \in \boldsymbol{X}$ が問題 (3.35) のクーン＝タッカー十分条件を満たしているとする．また，$(L{+}M)\times(L{+}M)$ 行列

$$\begin{pmatrix} \nabla_x^2 f(x^*, p^*) + \sum_{m=1}^{M} \lambda_m \nabla_x^2 g_m(x^*, p^*) & \nabla_x g_1(x^*, p^*)^\top & \ldots & \nabla_x g_M(x^*, p^*)^\top \\ \nabla_x g_1(x^*, p^*) & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \nabla_x g_M(x^*, p^*) & 0 & \ldots & 0 \end{pmatrix}$$

は正則行列であると仮定する．この時，p^* を含む開部分集合 $\boldsymbol{Q} \subset \boldsymbol{P}$ が存在し，Policy Function，$a(p)$，及び Value Function，$b(p)$ は，$\boldsymbol{Q}$ 上で連続微分可能である．

3.17 **定理の証明.** 陰関数定理の適用による． □

3.18 定理 (包絡線定理). パラメーター $p^* \in \boldsymbol{P}$ を所与として，クーン＝タッカー乗数 $(1, \lambda_1, \ldots, \lambda_K) \in \boldsymbol{R}^{1+K}_{++}$ とベクトル $x^* \in \boldsymbol{X}$ が問題 (3.35) のクーン＝タッカー十分条件を満たしているとする．また，Policy Function，$a(p)$，及び Value Function，$b(p)$ は連続微分可能であるとする．この時

$$\nabla b(p^*) = \nabla_p f(x^*, p^*) + \sum_{m=1}^{M} \lambda_m \nabla_p g_m(x^*, p^*)$$

が成立する．

3.18 **定理の証明（スケッチ）.** 仮定により，クーン＝タッカー乗数は全て厳密に正の値をとるから

$$\forall p \in \boldsymbol{P}: \quad g_m(a(p), p) = 0, \quad m = 1, 2, \ldots, M$$

が成立する．よってこれを p で微分して $p = p^*$ と置くことによって

$$\nabla_x g(x^*, p^*) \nabla a(p^*) + \nabla_p g_m(x^*, p^*) = 0, \quad m = 1, 2, \ldots, M \tag{3.36}$$

を得る．(3.36) の両辺に λ_m を掛け，m について和を取れば

$$\sum_{m=1}^{M} \lambda_m \nabla_x g(x^*, p^*) \nabla a(p^*) + \sum_{m=1}^{M} \lambda_m \nabla_p g_m(x^*, p^*) = 0 \tag{3.37}$$

一方，$(1, \lambda_1, \ldots, \lambda_K) \in \boldsymbol{R}^{1+K}_{++}$ 及び $x^* \in \boldsymbol{X}$ はクーン＝タッカー十分条件を満たすから

$$\nabla_x f(x^*, p^*) + \sum_{m=1}^{M} \lambda_m \nabla_x g_m(x^*, p^*) = 0 \tag{3.38}$$

ここで (3.38) を (3.37) に代入すれば

$$-\nabla_x f(x^*, p^*) \nabla a(p^*) + \sum_{m=1}^{M} \lambda_m \nabla_p g_m(x^*, p^*) = 0 \tag{3.39}$$

と書ける．Value Function の定義により

$$\forall p \in \boldsymbol{P}: \quad b(p) = f\left(a(p), p\right)$$

であるから，これを p で微分して $p = p^*$ と置くことによって

$$\nabla b(p^*) = \nabla_x f(x^*, p^*) \nabla a(x^*) + \nabla_p f(x^*, p^*) \tag{3.40}$$

この (3.40) に先の (3.39) を代入することにより，定理の結果を得る． □