経済学のための数学

京都大学経済研究所　原千秋・梶井厚志

平成 20 年 6 月 4 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 $x_1, x_2, x_3, \dots$ の並びを数列と言い，これをまとめて数列 $\big(x_n\big)$ $(n = 1, 2, 3, \dots)$ などと表す．

1.1 定義 (**数列の収束**)**.** $\boldsymbol{R}$ 上の数列 $\big(x_n\big)$ 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 $\big|x_n - x\big| < \epsilon$ が常に成り立つ時，数列 $\big(x_n\big)$ は x に収束すると言う． またこれを，

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

と書く．

数列 $\big(x_n\big)$ 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 $\big(x_n\big)$ 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 $\big(x_n\big)$ 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

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

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

1.2 例. 数列 $\big(x_n\big)$ 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 $\big(x_n\big)$ 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

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

であったから，その否定は

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

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， $\big|x_n - x\big| \geq \epsilon$ 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

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

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 $x_1, x_2, x_3, \dots$ の並び，すなわち (x_n) $(n = 1, 2, 3, \dots)$ 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 $x = (x^1, x^2, x^3, \dots, x^L)$ に対して，

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

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 (x_n) と点 $x \in \boldsymbol{R}^L$ を考える．

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

となる時，点列 (x_n) 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 (x_n) の極限であると言う．

点列 (x_n) が x に収束するとは， 点列 (x_n) 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 $x_n = (x_n^1, x_n^2, x_n^3, \dots, x_n^L)$ について，1.4 定義は以下の条件と同値である．

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

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 $\|x_n\| \to 0$ $(n \to \infty)$ 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.

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

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，$\partial^2 F / \partial y \partial x$ と $\partial^2 F / \partial x \partial y$ 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. $h,k \in \boldsymbol{R} \setminus \{0\}$ 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

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

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

$$\begin{aligned}
\frac{\partial^2 F}{\partial x \partial y} &: \boldsymbol{R}^2 \to \boldsymbol{R} \\
\frac{\partial^2 F}{\partial y \partial x} &: \boldsymbol{R}^2 \to \boldsymbol{R}
\end{aligned}$$

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，

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

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，

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

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，

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

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，$F(x,y) = c$ 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

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

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，

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

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

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

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

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

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

$$f\left(x+h\right) = f\left(x\right) + f'\left(x+\theta h\right)h$$

つまり

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

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

$$Ax = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n \end{pmatrix} = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n \in \boldsymbol{R}^m$$

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

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

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 $(j = 1, 2, 3, \ldots, n)$ 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 $e_1, e_2, e_3, \ldots, e_n$ の線形結合で表すことができる．$e_1, e_2, e_3, \ldots, e_n$ 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

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

と置けば,

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

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．

$a_1, a_2, \ldots, a_n$ 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

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

であるから

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

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 $a_1, a_2, \ldots, a_n$ が 1 次独立であれば

$$\begin{gathered} \forall j: \quad a_j \neq 0 \\ \text{かつ} \\ \forall j, k: \quad j \neq k \quad \to a_j \neq a_k \end{gathered}$$

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

$$\begin{gathered} \exists j: \quad a_j = 0 \\ \text{または} \\ \exists j, k,\ j \neq k: \quad a_j = a_k \end{gathered}$$

であれば，$a_1, a_2, \ldots, a_n$ 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，$a_1, a_2, \ldots, a_n$ が 1 次独立なベクトルの組であれば，その部分集合 $a_{i_1}, a_{i_2}, \ldots, a_{i_k} (k \leq n)$ 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 $a_1, a_2, \ldots, a_n$ 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 $a_1, a_2, \ldots, a_n, a_{n+1}, \ldots, a_{n+k}$ 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

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

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 $x_1 = x_2 = \cdots = x_n$ 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 $x_1, \ldots, x_n$ と m の方程式より成る連立 1 次方程式

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

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）$(x_1, \ldots, x_n)$ 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 $v_{l+1},\dots,v_n$ 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

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

となる．これから $Av_{l+1},\dots,Av_n$ 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 $a_1,a_2,\dots,a_n$ 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 $v_1,v_2,\dots,v_l,v_{l+1},v_{l+2},\dots,v_n$ 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，$v_1,v_2,\dots,v_l$ 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 $Av_{l+1},Av_{l+2},\dots,Av_n$ の線形結合で表される．これにより

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

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 $y\in\left\langle Av_{l+1},Av_{l+2},\dots,Av_n\right\rangle$ 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

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

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

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

次に

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

を仮定すると

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

が成り立つから

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

よって

$$\exists z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_l \end{pmatrix} \in \boldsymbol{R}^l : \quad z_{l+1}v_{l+1} + z_{l+2}v_{l+2} + \cdots + z_n v_n = z_1 v_1 + z_2 v_2 + \cdots + z_l v_l$$

すなわち

$$z_1 v_1 + z_2 v_2 + \cdots + z_l v_l - z_{l+1}v_{l+1} - z_{l+2}v_{l+2} - \cdots - z_n v_n = 0$$

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

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

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

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

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

(解の自由度) = (未知数の数) − (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, $\langle a_1,\ldots,a_n\rangle = \boldsymbol{R}^n$ 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, $a_1,\ldots,a_n$ 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, $a_1,\ldots,a_n$ 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 $a_1,\ldots,a_n$ の線形結合として表される．これは,
$\langle a_1,\ldots,a_n\rangle = \boldsymbol{R}^n$ 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, $a_1,\ldots,a_n$ 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 sgn(π) 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

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

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, 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, det$A^\top$ に他ならない. □

2.20 定義 (小行列式). n 次正方行列

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

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 $(l \leq n)$ 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 $v \notin \operatorname{Ker}(A-\lambda I)$) 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

$$\begin{aligned} Av_1 &= \lambda v_1 \\ Av_2 &= v_1 + \lambda v_2 \end{aligned}$$

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

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

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

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

と置けば，(2.26) を

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

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 $P'^{-1}A'_1P'$ 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,

$$\begin{aligned} P^{-1}AP &= \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix}^{-1} \begin{pmatrix} \alpha_1 & 0 \\ 0 & A_1' \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix} \\ &= \begin{pmatrix} \alpha_1 & 0 \\ 0 & P'^{-1}A_1'P' \end{pmatrix} \end{aligned}$$

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,

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

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

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

を満たす整数の列とする．ここで行列 m 次正方行列 B を

$$B_M = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1m} \\ b_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ b_{m1} & \ldots\ldots & & b_{mm} \end{pmatrix} = \begin{pmatrix} a_{i_1i_1} & a_{i_1i_2} & \ldots & a_{i_1i_m} \\ a_{i_2i_1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{i_mi_1} & \ldots\ldots\ldots & & a_{i_mi_m} \end{pmatrix}$$

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:
<u>例</u>

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

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:
<u>例</u>

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

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

$$C^c = \left\{ x \in \boldsymbol{R}^n \;\middle|\; x \notin C \right\}$$

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

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

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

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

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

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

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

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

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

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

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. $\forall z \in \boldsymbol{R} : \quad \big\{x \in C \;\big|\; F(x) \leq z\big\}$ 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. $\forall z \in \boldsymbol{R} : \big\{x \in C \;\big|\; F(x) \geq z\big\}$ 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.

$\square$

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

$$\begin{aligned}
\widehat{b}^{\top} &= \sum_{j=1}^{J-1} w_j \widehat{a}_j \\
&= \sum_{j=1}^{J-1} w_j \left(a_j - \frac{x' \cdot a_j}{x' \cdot a_J} a_J \right) \\
&= \sum_{j=1}^{J-1} w_j a_j - \left(\sum_{j=1}^{J-1} w_j \frac{x' \cdot a_j}{x' \cdot a_J} \right) a_J \qquad (3.9)
\end{aligned}$$

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

$$\begin{aligned}
b^{\top} &= \sum_{j=1}^{J-1} w_j a_j + \left(\frac{x' \cdot b}{x' \cdot a_J} - \sum_{j=1}^{J-1} w_j \frac{x' \cdot a_j}{x' \cdot a_J} \right) a_J \\
&= w^{\top} A' + \left(\frac{x' \cdot b}{x' \cdot a_J} - \sum_{j=1}^{J-1} w_j \frac{x' \cdot a_j}{x' \cdot a_J} \right) a_J \qquad (3.10)
\end{aligned}$$

$a_J \cdot x' < 0$ 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, $j = 1, 2, \dots, J$ 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

$$z^\top \widetilde{A} = (\underbrace{0, 0, \dots, 0}_{N}, 1)$$

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

$$\widetilde{A}\widetilde{x} \in \boldsymbol{R}^J_+ \tag{3.24}$$

$$(0, 0, \dots, 0, 1)\widetilde{x} < 0 \tag{3.25}$$

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，

$$\begin{aligned} \forall n: \quad & \nabla f_n(x^*) \cdot v > 0 \\ \forall m \leq K: \quad & \nabla g_m(x^*) \cdot v > 0 \end{aligned}$$

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 $(\mu_1, \ldots, \mu_N, \lambda_1, \ldots \lambda_K, \lambda_{K+1}, \ldots, \lambda_M) \in \boldsymbol{R}_+^{N+K}$ 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，$(\mu_1, \ldots, \mu_N, \lambda_1, \ldots, \lambda_M)$ 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 $(t\mu_1, \ldots, t\mu_N, t\lambda_1, \ldots, t\lambda_M)$ も同様に条件を満たすことを意味する．よって，$(\mu_1, \ldots, \mu_N, \lambda_1, \ldots, \lambda_M)$ 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 $(\mu_1, \ldots, \mu_N, \lambda_1, \ldots, \lambda_M) \in \boldsymbol{R}_+^{N+M}$ が存在し

1. $\forall m: \quad g_m(x^*) \geq 0;$
2. $(\mu_1, \ldots, \mu_N, \lambda_1, \ldots, \lambda_M) \in \boldsymbol{R}_{++}^{N};$
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 3.3.3 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 $\boldsymbol{X} = \left\{x \in \boldsymbol{R}^2 \mid x_1 > -1,\ x_2 > -1\right\}$ 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

$$\begin{aligned} \max_{x \in \boldsymbol{R}^2_+} \quad & u(x_1, x_2) = \sqrt{x_1} + \sqrt{x_2} \\ \text{subject to} \quad & w - p_1 x_1 - p_2 x_2 \geq 0. \end{aligned} \tag{3.30}$$

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

$$\begin{aligned} \frac{\partial u}{\partial x_1}(x_1, x_2) \to \infty \quad (x_1 \to 0) \\ \frac{\partial u}{\partial x_2}(x_1, x_2) \to \infty \quad (x_2 \to 0) \end{aligned}$$

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

$$\begin{aligned} \max_{x \in \boldsymbol{R}^2_{++}} \quad & u(x_1, x_2) = \sqrt{x_1} + \sqrt{x_2} \\ \text{subject to} \quad & w - p_1 x_1 - p_2 x_2 \geq 0. \end{aligned} \tag{3.31}$$

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

$$\begin{aligned} \max_{x \in \boldsymbol{R}^2_+} \quad & u(x) \\ \text{subject to} \quad & w - x_1 - x_2 \geq 0. \end{aligned} \tag{3.32}$$

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 $\left\{ x \in \boldsymbol{R}^2 \mid x_1 > 0,\ x_2 \geq 0 \right\}$ 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

$$\boldsymbol{X} = \left\{ x \in \boldsymbol{R}^2 \mid x_1 > 0 \right\}$$

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

$$\begin{aligned} \max_{x \in \boldsymbol{X}} \quad & f(x, p) \\ \text{subject to} \quad & g_1(x, p) \geq 0, \\ & g_2(x, p) \geq 0, \\ & \quad \vdots \\ & g_M(x, p) \geq 0. \end{aligned} \tag{3.35}$$

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． □

3.3.3　クーン＝タッカー条件の応用

3.3.1 項では，クーン＝タッカー条件を一般的な形式で導入した．この項では非線形計画法の問題

$$\begin{aligned} \max_{x\in V} \quad & f(x) \\ \text{subject to} \quad & g_j(x) \geq 0, \qquad j=1,2,\ldots,m \end{aligned} \tag{3.41}$$

を例にしながら，後半の議論のために使いやすい形で結果をまとめておく．V は $\boldsymbol{R}^n$ の開部分集合，関数 $f : V \to \boldsymbol{R}$, $g_j : V \to \boldsymbol{R}$ $(j = 1, 2, \ldots m)$ はいずれも微分可能であるとする．

3.19 定理 (**フリッツ・ジョン**)**.** ベクトル x^* が問題 (3.41) の解であるとする．このとき $m+1$ 個の非負の数 $\lambda_0, \lambda_1, \ldots, \lambda_m$ が存在し，

$$\lambda_0 \nabla f(x^*) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x^*) = 0 \tag{3.42}$$

$$\lambda_j g_j(x^*) = 0, \quad j = 1, 2, \ldots, m \tag{3.43}$$

が成立する．ただし，$\lambda_0, \lambda_1, \ldots, \lambda_m$ のうち少なくとも一つは厳密に正の値をとる．

3.19 定理における (3.42) を 1 階の条件（First Order Condition: FOC），(3.43) を相補性条件（Complementary Slackness）と呼ぶ．もし $\lambda_0 > 0$ であれば，(3.42) と (3.43) を λ_0 で割り，$\lambda_1, \lambda_2, \ldots, \lambda_m$ を置き直すことによって

$$\nabla f(x^*) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x^*) = 0 \tag{3.44}$$

$$g_j(x^*) = 0, \quad j = 1, 2, \ldots, m \tag{3.45}$$

のように条件を書き直すことができる．(3.44) と (3.45) を合わせたものは，クーン＝タッカー条件と呼ばれる．応用上はこちらの方が有益であるが，3.19 定理では $\lambda_0 = 0$ のケースを排除していない．そこで必要となるのが「制約想定（Constraint Qualification）」という追加的な条件であり，次の定理はこの点を取り込んだものである．

3.20 定理 (**クーン＝タッカー条件の必要性：一般型**)**.** ベクトル x^* が問題 (3.41) の解であるとする．さらに x^* において何らかの制約想定条件が満たされているとする．このとき m 個の非負の数 $\lambda_1, \lambda_2, \ldots, \lambda_m$ が存在し，

$$\nabla f(x^*) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x^*) = 0$$

$$g_j(x^*) = 0, \quad j = 1, 2, \ldots, m$$

が成立する．ただし，$\lambda_1, \lambda_2, \ldots, \lambda_m$ のうち少なくとも一つは厳密に正の値をとる．

3.20 定理で言う制約想定条件としては，様々なものが知られている[11]．経済学への応用という点でいうと，おそらく次の条件を制約想定として用いるのが便利であろう．

1. 関数 g_j はいずれも準凹関数である．
2. あるベクトル $x_0 \in \boldsymbol{R}^n$ が存在し，全ての $j = 1, 2, \ldots, m$ で $g_j(x_0) > 0$ となる．

[11] 3.19 定理における $\lambda_0 \neq 0$ を保障するためには，クーン＝タッカー条件で正の値をとる乗数 $\lambda_j > 0$ について，それらに対応する $\nabla g_j(x^*)$ たちが（正の）1 次従属になるケースを排除してやればよい.「制約想定」条件として様々なものがあるというのは，こういったケースを排除するためのテクニックが色々と考えられているということである．

この二つの条件が満たされていれば，3.20 定理の制約想定条件は満足される．ちなみに，上の条件 2 はスレーター条件（Slater Condition）と呼ばれるものである．このスレーター条件によって制約想定条件を具体化すれば，3.20 定理は以下のように書き直すことができる．

3.21 定理 (クーン＝タッカー条件の必要性：スレーター条件使用). ベクトル x^* が問題 (3.41) の解であるとする．さらに（1）関数 g_j はいずれも準凹関数で, (2) スレーター条件が満たされているとする．このとき m 個の非負の数 $\lambda_1, \lambda_2, \ldots, \lambda_m$ が存在し，

$$\nabla f(x^*) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x^*) = 0$$

$$g_j(x^*) = 0, \quad j = 1, 2, \ldots, m$$

が成立する．ただし，$\lambda_1, \lambda_2, \ldots, \lambda_m$ のうち少なくとも一つは厳密に正の値をとる．

解の十分条件についても，次の定理でまとめておこう．

3.22 定理 (クーン＝タッカー条件の十分性). ベクトル x^* が全ての制約式を満たすとする．また x^* と m 個の非負の数 $\lambda_1, \lambda_2, \ldots, \lambda_m$ がクーン＝タッカー条件 (3.44) と (3.45) を満たしているとする．このとき全ての関数 g_j が準凹関数で，なおかつ関数 f が凹関数の単調変換[12]であれば，x^* は問題 (3.41) の解である．

実際の応用上は，次の定理[13]とクーン＝タッカー条件の必要性（3.20 定理もしくは 3.21 定理）を組み合わせるテクニックも役に立つ．

3.23 定理 (ワイヤシュトラスの最大値定理). $\boldsymbol{R}^n$ の有界な閉集合（つまりコンパクト集合）で定義された連続関数 f については，それを最大化する元が存在する．

問題 (3.41) の制約式を満たすベクトルの集合が有界な閉集合であると示すことができれば，3.23 定理により問題に解が存在することが分かる．そしてその解は 3.20 定理（3.21 定理）によりクーン＝タッカー条件を必ず満たすので，(3.44) と (3.45) を分析すれば最適化の性質を知ることができるのである．

最後にクーン＝タッカー条件とラグランジュ関数との関係について簡単にまとめておく[14]．問題 (3.41) に対して

$$L(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x)$$

で定義される関数 $L : V \times \boldsymbol{R}_+^m \to \boldsymbol{R}$ を，ラグランジュ関数という．また，任意の $x \in V$ と任意の $\lambda \in \boldsymbol{R}_+^m$ に対して

$$L(x, \lambda^*) \leq L(x^*, \lambda^*) \leq L(x^*, \lambda)$$

が成り立つとき，$(x^*, \lambda^*) \in V \times \boldsymbol{R}_+^m$ はラグランジュ関数 $L(x, \lambda)$ の鞍点（saddle point）であるという．

3.24 定理 (鞍点の必要条件). 問題 (3.41) において，$(x^*, \lambda^*) \in V \times \boldsymbol{R}_+^m$ がラグランジュ関数 $L(x, \lambda)$ の鞍点ならば，次の四つの条件が成立する．

1. $\dfrac{\partial}{\partial x_i} L(x^*, \lambda^*) \leq 0, \qquad i = 1, 2, \ldots, n$

[12] ある凹関数 h と，$\phi' > 0$ を満たす関数 $\phi : \boldsymbol{R} \to \boldsymbol{R}$ を用いて，$f(x) \equiv \phi\big(h(x)\big)$ と書けるということ．
[13] 詳しくは岡田 4 章を参照．ただし興味のある人のみでかまわない．
[14] 証明などの詳しい説明は岡田 6 章の 214-223 頁を参照．

2. $x_i^* \cdot \dfrac{\partial}{\partial x_i} L(x^*, \lambda^*) = 0, \qquad i = 1, 2, \ldots, n$

3. $\dfrac{\partial}{\partial \lambda_j} L(x^*, \lambda^*) \geq 0, \qquad j = 1, 2, \ldots, m$

4. $\lambda_j^* \cdot \dfrac{\partial}{\partial \lambda_j} L(x^*, \lambda^*) = 0, \qquad j = 1, 2, \ldots, m$

3.24 定理の条件は，それぞれ

$$\begin{aligned}
\frac{\partial}{\partial x_i} L(x^*, \lambda^*) &= \frac{\partial}{\partial x_i} f(x^*) + \sum_{j=1}^{m} \lambda_j^* \frac{\partial}{\partial x_i} g_j(x^*) \leq 0, \quad i = 1, 2, \ldots, n && (3.46)\\
x_i^* \cdot \frac{\partial}{\partial x_i} L(x^*, \lambda^*) &= x_i^* \left(\frac{\partial}{\partial x_i} f(x^*) + \sum_{j=1}^{m} \lambda_j^* \frac{\partial}{\partial x_i} g_j(x^*) \right) = 0, \quad i = 1, 2, \ldots, n && (3.47)\\
\frac{\partial}{\partial \lambda_j} L(x^*, \lambda^*) &= g_j(x^*) \geq 0, \quad j = 1, 2, \ldots, m && (3.48)\\
\lambda_j^* \cdot \frac{\partial}{\partial \lambda_j} L(x^*, \lambda^*) &= \lambda_j^* g_j(x^*) = 0, \quad j = 1, 2, \ldots, m && (3.49)
\end{aligned}$$

と書き直すことができる．(3.48) は問題の制約条件であり，(3.49) は相補性条件である．また全ての i について $x_i^* > 0$ であるならば，(3.46) と (3.47) は

$$\frac{\partial}{\partial x_i} f(x^*) + \sum_{j=1}^{m} \lambda_j^* \frac{\partial}{\partial x_i} g_j(x^*) = 0, \qquad i = 1, 2, \ldots, n$$

すなわち

$$\nabla f(x^*) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x^*) = 0$$

に等しく，これは 1 階の条件に他ならない．したがって問題 (3.41) の解 x^* は

$$\nabla f(x^*) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x^*) = 0 \tag{3.50}$$

$$g_j(x^*) = 0, \quad j = 1, 2, \ldots, m \tag{3.51}$$

$$g_j(x^*) \geq 0, \quad j = 1, 2, \ldots, m \tag{3.52}$$

を満たすと言え，これはクーン＝タッカー条件に制約式を合わせたものである[15]．つまり，クーン＝タッカー条件はラグランジュ関数の鞍点のための必要条件となっているのである．

3.25 定理 (鞍点と最適解の必要性). f と g_j $(j = 1, 2, \ldots, m)$ が凹関数であり，スレーター条件が満たされているとする．このとき x^* が問題 (3.41) の解であるならば，$\lambda^* \in \boldsymbol{R}_+^m$ が存在し，$(x^*, \lambda^*) \in V \times \boldsymbol{R}_+^m$ はラグランジュ関数 $L(x, \lambda)$ の鞍点となる．

3.25 定理によれば，f と g_j $(j = 1, 2, \ldots, m)$ が凹関数であり，なおかつスレーター条件が満たされている場合，ラグランジュ関数の鞍点は問題 (3.41) の最適解の必要条件となる．したがって 3.24 定理を合わせて考えれば，その場合にクーン＝タッカー条件は問題 (3.41) の解の必要条件であると言える．

また次の定理が示すように，一定の条件の下で，クーン＝タッカー条件はラグランジュ関数の鞍点の十分条件にもなる．

[15] 鞍点のための必要条件という理解から，(3.50) と (3.51) に加えて (3.52) を合わせたものをクーン＝タッカー条件と呼ぶことも多い．梶井は (3.50) と (3.51) をクーン＝タッカー条件と呼ぶことにしているが，岡田本をはじめ多くの教科書が制約式を含めたものをクーン＝タッカー条件と呼んでいる．

3.26 定理 (**鞍点の十分条件**). f と $g_j(j=1,2,\ldots,m)$ が凹関数であるとする．このとき $(x^*,\lambda^*)\in V\times\boldsymbol{R}_+^m$ が 3.24 定理の四つの条件を満たすならば，(x^*,λ^*) はラグランジュ関数 $L(x,\lambda)$ の鞍点である．

さらに，次の定理により 鞍点であれば問題 (3.41) の解となることが言えるから，クーン＝タッカー条件は問題 (3.41) の解の十分条件にもなるのである．

3.27 定理 (**鞍点と最適解の十分性**). $(x^*,\lambda^*)\in V\times\boldsymbol{R}_+^m$ がラグランジュ関数 $L(x,\lambda)$ の鞍点であるならば，x^* は問題 (3.41) の解である．

以上のように，クーン＝タッカー条件はラグランジュ関数と密接な関わりを持っている．クーン＝タッカー条件の覚え方としても，ラグランジュ関数 $L(x,\lambda)$ の x に関する偏微分が全てゼロになることと覚えておけばよい．

練習問題 3.15. 関数 $f:\boldsymbol{R}\to\boldsymbol{R}$ は 2 回微分可能で，$f>0$，$f'>0$，$f''\geq 0$ を満たしているとする．この時，次のような最適化問題を考える．

$$\min_{x\in\boldsymbol{R}}\quad g(x):=\frac{1}{x}f(x)$$

まず，$x>0$ の領域で g は凸関数であるか調べよ．また，$g'(x)=0$ を満たす x が問題の解となるか調べよ.（ヒント：$g'(x)=0$ を満たす x が存在すれば，それは唯一であることを示す.）

4 無限次元の最適化問題：動的計画法

4.1 無限期間の最適化問題

この章[16]では無限期間の最適化問題を解く方法として，動的計画法を説明する．その準備として，まずは簡単な 2 期間の最適化問題を考えよう．全体として $\bar{x}$ 単位の消費可能な財があるとして，それを 2 期間かけて消費するような場合に，各期にどれだけの消費を割り当てるべきかを考えることにする．このとき問題は,

$$\begin{aligned} \max_{(c_0,c_1)} \quad & u(c_0)+\delta u(c_1) \\ \text{subject to} \quad & c_0+c_1 \leq \bar{x} \\ & c_0 \geq 0 \\ & c_1 \geq 0 \end{aligned} \tag{4.1}$$

のように書ける．ここで u は効用関数を，c_0 および c_1 はそれぞれ 0 期と 1 期の消費量を表す．δ は割引因子である．

問題 (4.1) の解 (c_0^*, c_1^*) が内点解であると仮定すれば，クーン＝タッカー必要条件により,

$$\begin{aligned} u'(c_0^*) &= \lambda \\ \delta u'(c_1^*) &= \lambda \end{aligned}$$

が成立しているはずである．よってラグランジュ乗数 λ を消去して，予算制約と合わせれば

$$\begin{aligned} -u'(c_0^*)+\delta u'(c_1^*) &= 0 \\ c_0^*+c_1^* &= \bar{x} \end{aligned}$$

のように整理することができ，この連立法的式を解けば解を得られる．

考慮する期間を長くしても，この問題の構造は変わらない．例えばより一般的に，$\bar{x}$ 単位の消費可能な財を T 期間かけて消費するような場合を考えよう．すると先ほどの問題 (4.1) は,

$$\begin{aligned} \max_{(c_0,\dots,c_{T-1})} \quad & u(c_0)+\delta u(c_1)+\delta^2 u(c_2)+\cdots+\delta^{T-1}u(c_{T-1}) \\ \text{subject to} \quad & c_0+c_1+\cdots+c_{T-1} \leq \bar{x} \\ & c_0 \geq 0 \\ & c_1 \geq 0 \\ & \vdots \\ & c_{T-1} \geq 0 \end{aligned} \tag{4.2}$$

と書き直すことができる．2 期間の場合と同様に，問題 (4.2) の解 $(c_0^*, \dots, c_{T-1}^*)$ が内点解であると仮定すれば，クーン＝タッカー必要条件により

$$\begin{aligned} u'(c_0^*) &= \lambda \\ \delta u'(c_1^*) &= \lambda \\ &\vdots \\ \delta^{T-1}u'(c_{T-1}^*) &= \lambda \end{aligned}$$

[16] この章の内容をより詳しく知るためには，Stokey-Lucas, chapeter4 を見よ．

が導かれる．よって予算制約と合わせれば

$$-\delta^{t-1}u'(c^*_{t-1})+\delta^t u'(c^*_t)=0,\quad t=1,2,\dots,T{-}1$$
$$c^*_0+c^*_1+\cdots+c^*_{T-1}=\bar{x}$$

のように整理することができ，この連立法的式を解けば解を得られる．

以上から，期間 T が有限である限りで，この問題を解くことは可能であると分かる．では，期間 T が無限となるようなケースについてはどうであろうか．$\bar{x}$ 単位の財を無限期間かけて消費するような問題は，

$$\begin{aligned}\max_{\{c_t:t=0,1,\dots\}}\quad & \sum_{t=0}^{\infty}\delta^t u(c_t) \\ \text{subject to}\quad & \sum_{t=0}^{\infty}c_t\le\bar{x} \\ & c_t\ge 0,\quad t=0,1,\dots\end{aligned}\tag{4.3}$$

のように書くことができる．しかしこの無限期間の最適化問題に対しては，上記の方法で解を見つけることができない．同様の方法を適用しても無限個の変数と無限本の方程式が導かれてしまい，その連立方程式体系を解くことができないからである．このような問題に答えるために必要とされるのが，本章で説明する動的計画法（Dynamic Programming: DP）である．

4.2 動的計画法

これ以降の議論では $\boldsymbol{X}\subseteq\boldsymbol{R}^n_+$ とし，$x_t\in\boldsymbol{X}$ と約束する．また δ は $0<\delta<1$ を満たす定数であり，関数 $f:\boldsymbol{X}\times\boldsymbol{X}\to\boldsymbol{R}$ は与えられているものとする．

4.2.1 動的計画法の定式化

動的計画法の問題 DP($\bar{x}_0$) は，一般に次のような最適化問題として定式化される．

$$\begin{aligned}\max_{\{x_t:t=1,2,\dots\}}\quad & \sum_{t=0}^{\infty}\delta^t f(x_t,x_{t+1}) \\ \text{subject to}\quad & \Gamma(x_t,x_{t+1})\ge 0,\quad t=0,1,\dots \\ & x_0=\bar{x_0}\end{aligned}$$

ただし，任意の $x\in\boldsymbol{X}$ について，$\Gamma(x,y)\ge 0$ を満たす $y\in\boldsymbol{X}$ が存在するものと仮定する．

4.1 例 (ケーキ食べ問題). 自由に消費できる $\bar{x}$ 単位のケーキが用意されており，その最適な消費スケジュールを決める問題を考えよう．この個人がケーキを z 単位だけ消費することで得られる効用は $u(z)$，割引因子は δ であるとする．

この問題を，各期に消費するケーキの量 $\{c_t\}$ を決定するものであると考えれば，これはちょうど (4.3) のように定式化できる．しかし一方で，これを次の期に残しておくケーキの量を決定する問題であると捉えることもできる．t 期のはじめの時点で残っているケーキの量を $\{x_t\}$ と置けば，

$$c_t=x_t-x_{t+1}$$
$$\bar{x}=x_0$$

の関係が成り立つから，問題 (4.3) は

$$
\begin{aligned}
\max_{\{x_t:t=0,1,\dots\}} \quad & \sum_{t=0}^{\infty} \delta^t u\left(x_t - x_{t+1}\right) \\
\text{subject to} \quad & x_0 = \bar{x} \\
& x_t - x_{t+1} \geq 0, \quad t = 0, 1, \dots
\end{aligned}
$$

のように書ける．ここで

$$
\begin{aligned}
f\left(x_t, x_{t+1}\right) &:= u\left(x_t - x_{t+1}\right) \\
\Gamma\left(x_t, x_{t+1}\right) &:= x_t - x_{t+1}
\end{aligned}
$$

のように $f\left(x_t, x_{t+1}\right)$ および $\Gamma(x_t, x_{t+1})$ を定義すれば，

$$
\begin{aligned}
\max_{\{x_t:t=0,1,\dots\}} \quad & \sum_{t=0}^{\infty} \delta^t f\left(x_t, x_{t+1}\right) \\
\text{subject to} \quad & x_0 = \bar{x} \\
& \Gamma\left(x_t, x_{t+1}\right) \geq 0, \quad t = 0, 1, \dots
\end{aligned}
$$

として，ケーキ食べ問題を上記の動的計画法 *DP* の形に書くことができる．

練習問題 4.1 (**ケーキ食べ問題**)**.** 上記のケーキ食べ問題の想定を少し変え，各期に残されたケーキの量が次の期の始めに β 倍 ($\beta > 0$) になるとする．すなわち，t 期の消費量と $t+1$ 期に残されるケーキの量との間には，

$$x_{t+1} = \beta(x_t - c_t)$$

の関係が成り立つものとする．この想定の下でのケーキ食べ問題を，DP の形で書きなさい．

練習問題 4.2 (**貯蓄と消費**)**.** t 期における資産保有を x_t，消費を c_t，貯蓄を s_t と置き，貯蓄に対する金利は i で一定とする．資産と消費，貯蓄の間には

$$c_t + s_t = x_t$$

の関係が成立し，t 期の貯蓄と $t+1$ 期の資産は

$$x_t = (1+i)s_t$$

で関係付けられている．この制約の下で，消費から得られる生涯効用

$$\sum_{t=0}^{\infty} \delta^t u\left(c_t\right)$$

を最大化する問題を，上記のような DP の形に書きなさい．

4.1 定義 (**実行可能プラン**)**.** ベクトル列 $\{x_t : t = 0, 1, \dots\}$ が問題 DP$(\bar{x}_0)$ の制約を全て満たすとき，$\{x_t : t = 0, 1, \dots\}$ を（$\bar{x}_0$ を始点とする）実行可能プラン（あるいは実行可能経路）と呼ぶ．また特に，$\Gamma\left(x_t, x_{t+1}\right) > 0$ がすべての t について成立するとき，その実行可能プランは内点であるという．

4.2.2　オイラー方程式

以上のようにして定式化される動的計画法について，その解の求め方を見ていこう．この節では $\boldsymbol{X}$ を開集合とし，$f : \boldsymbol{X} \times \boldsymbol{X} \to \boldsymbol{R}$ を連続微分可能であると仮定する．有限の非線形計画法におい

てクーン＝タッカー必要条件を用いて解を求めたように，無限の最適化問題でも必要条件から解を特定してゆくことになる．動的計画法において必要条件に対応するものは，以下のオイラー方程式（Euler equation）と呼ばれるものである．

$$\frac{\partial}{\partial x_t} f\left(x_{t-1}, x_t\right) + \delta \frac{\partial}{\partial x_t} f\left(x_t, x_{t+1}\right) = 0, \quad t = 1, 2, \ldots \tag{4.4}$$

このオイラー方程式が問題 DP$(\bar{x}_0)$ の内点解が満たすべき必要条件になっていることを，以下の定理を証明しながら確認する．

4.1 定理 (解の必要条件). ベクトル列 $\{x_t : t = 0, 1, 2, \ldots\}$ が問題 DP$(\bar{x}_0)$ の内点解であれば，オイラー方程式 (4.4) が成立する．

4.1 定理の証明. ベクトル列 $\{x_t\}$ が問題 DP$(\bar{x}_0)$ の内点解であるとする．この時，目的関数の値は

$$\sum_{t=0}^{\infty} \delta^t f\left(x_t, x_{t+1}\right) = f\left(x_0, x_1\right) + \delta f\left(x_1, x_2\right) + \cdots + \delta^{t-1} f\left(x_{t-1}, x_t\right) + \delta^t f\left(x_t, x_{t+1}\right) + \ldots \tag{4.5}$$

である．ここで時点 t を任意に選び，変数 x_t のみを操作することによってこの目的関数の値を変化させることを考えよう．すなわち，

$$V\left(z\right) = f\left(x_0, x_1\right) + \delta f\left(x_1, x_2\right) + \cdots + \delta^{t-1} f\left(x_{t-1}, z\right) + \delta^t f\left(z, x_{t+1}\right) + \ldots \tag{4.6}$$

なる関数 $V(z)$ について，制約条件を加えた最適化問題

$$\begin{array}{rl} \max\limits_{z} & V\left(z\right) \\ \text{subject to} & \Gamma\left(x_{t-1}, z\right) \geq 0 \\ & \Gamma\left(z, x_{t+1}\right) \geq 0 \end{array} \tag{4.7}$$

を考える．ここでもし $z = x_t$ が解でなかったとすれば，制約式を満たす x_t' が存在し

$$V\left(x_t'\right) > V\left(x_t\right)$$

すなわち

$$f\left(x_0, x_1\right) + \cdots + \delta^{t-1} f\left(x_{t-1}, x_t'\right) + \delta^t f\left(x_t', x_{t+1}\right) + \cdots > \sum_{t=0}^{\infty} \delta^t f\left(x_t, x_{t+1}\right)$$

が成立する．経路 $\{x_0, x_1, \ldots x_t', \ldots\}$ は実行可能プランであるから，これは $\{x_0, x_1, \ldots x_t, \ldots\}$ が問題 DP$(\bar{x}_0)$ の解であることに反する．したがって (4.8) の解は $z = x_t$ でなければならず，$z = x_t$ は (4.8) の 1 階条件

$$\left.\frac{d}{dz} V(z)\right|_{z=x_t} = 0 \quad \Leftrightarrow \quad \delta^{t-1} \frac{\partial}{\partial x_t} f\left(x_{t-1}, x_t\right) + \delta^t \frac{\partial}{\partial x_t} f\left(x_t, x_{t+1}\right) = 0$$

を満たさなければならない．さらに両辺を δ^{t-1} で割れば，

$$\frac{\partial}{\partial x_t} f\left(x_{t-1}, x_t\right) + \delta \frac{\partial}{\partial x_t} f\left(x_t, x_{t+1}\right) = 0 \tag{4.8}$$

となる．時点 t は任意に選んであったから，結局この必要条件はオイラー方程式 (4.4) に他ならない． □

4.2 例 (ケーキ食べ問題). ケーキ食べ問題についてオイラー方程式を求めてみよう．問題は，

$$\begin{aligned}
\max_{\{x_t:t=0,1,\dots\}} \quad & \sum_{t=0}^{\infty} \delta^t u\left(x_t - x_{t+1}\right) \\
\text{subject to} \quad & x_t - x_{t+1} \geq 0, \quad t = 0, 1, \dots \\
& x_0 = \bar{x}
\end{aligned}$$

であったから，オイラー方程式は

$$-u'\left(x_{t-1} - x_t\right) + \delta u'\left(x_t - x_{t+1}\right) = 0$$

である．ちなみに，これを各期の消費量 c_t で表してみると，

$$-u'\left(c_{t-1}\right) + \delta u'\left(c_t\right) = 0 \quad \Leftrightarrow \quad \frac{\delta u'\left(c_t\right)}{u'\left(c_{t-1}\right)} = 1 \tag{4.9}$$

となる．$t-1$ 期の消費の限界効用は $\delta^{t-1}u'\left(c_{t-1}\right)$，$t$ 期の消費の限界効用は $\delta^t u'\left(c_t\right)$ であるから，(4.9) の左辺は $t-1$ 期の限界効用と t 期の限界効用の比に相当する．一方の右辺は，$t-1$ 期のケーキと t 期のケーキの価格比である[17]．よってこの問題のオイラー方程式 (4.9) が，限界効用の比が価格比に等しくなるという 1 階条件を表すものであることを確認できる．

4.2.3 横断面条件

オイラー方程式は，あくまで問題 DP($\bar{x}_0$) の解が満たすべき必要条件である．したがって，ある実行可能経路がオイラー方程式を満たしたからと言って，それが最適解であるとは限らない．有限の非線形計画法の場合は，目的関数や制約式が凹関数であれば 1 階条件が最適性の十分条件にもなったが，ここで問題としている無限の場合では十分性のための追加的な条件を考える必要がある．その十分条件の一部として用いられるのが，

$$\lim_{t\to\infty} \delta^t \frac{\partial}{\partial x_t} f\left(x_t, x_{t+1}\right) x_t = 0 \tag{4.10}$$

という条件であり，これを横断面条件（Transversality Condition: TC）と呼ぶ．

4.2 定理 (解の十分条件). 関数 $f\left(x_t, x_{t+1}\right)$ は $\boldsymbol{X}\times\boldsymbol{X} \subset \boldsymbol{R}\times\boldsymbol{R}$ 上の有界な凹関数で，$\dfrac{\partial}{\partial x_t} f\left(x_t, x_{t+1}\right) \gg 0$ が $\boldsymbol{X}\times\boldsymbol{X}$ 上で成立しているとする．この時，内点の実行可能経路 $\{x_t : t = 0, 1, \dots\}$ が

1. オイラー方程式 (4.4)
2. 横断面条件 (4.10)

を満たすならば，$\{x_t : t = 0, 1, \dots\}$ は問題 DP($\bar{x}_0$) の解である．

4.2 定理の証明. 内点の実行可能経路 $\{x_t : t = 0, 1, \dots\}$ がオイラー方程式 (4.4) と横断面条件 (4.10) を満たしているとする．この時，任意の実行可能経路 $\{x'_t : t = 0, 1, \dots\}$ に対して，

$$\sum_{t=0}^{\infty} \delta^t f\left(x_t, x_{t+1}\right) \geq \sum_{t=0}^{\infty} \delta^t f\left(x'_t, x'_{t+1}\right)$$

が成立することを示せばよい．

[17] この問題ではどの期のケーキの価格も等しい．

まず極限操作の順序交換に注意すれば

$$
\begin{aligned}
\sum_{t=0}^{\infty} \delta^t f(x_t, x_{t+1}) - \sum_{t=0}^{\infty} \delta^t f\left(x'_t, x'_{t+1}\right) &= \lim_{T\to\infty} \left(\sum_{t=0}^{T} \delta^t f(x_t, x_{t+1}) - \sum_{t=0}^{T} \delta^t f\left(x'_t, x'_{t+1}\right) \right) \\
&= \lim_{T\to\infty} \left(\sum_{t=0}^{T} \delta^t \Big(f(x_t, x_{t+1}) - f\left(x'_t, x'_{t+1}\right) \Big) \right) \quad (4.11)
\end{aligned}
$$

を得る．また $f(x_t, x_{t+1})$ が凹関数であることにより

$$
\begin{aligned}
f\left(x'_t, x'_{t+1}\right) - f(x_t, x_{t+1}) &\leq Df(x_t, x_{t+1}) \cdot \left(x'_t - x_t, x'_{t+1} - x_{t+1}\right) \\
&= f_1(x_t, x_{t+1})(x'_t - x_t) + f_2(x_t, x_{t+1})\left(x'_{t+1} - x_{t+1}\right)
\end{aligned}
$$

すなわち

$$
f(x_t, x_{t+1}) - f\left(x'_t, x'_{t+1}\right) \geq f_1(x_t, x_{t+1})(x_t - x'_t) + f_2(x_t, x_{t+1})\left(x_{t+1} - x'_{t+1}\right)
$$

が成り立つから

$$
\begin{aligned}
&\lim_{T\to\infty} \left(\sum_{t=0}^{T} \delta^t \Big(f(x_t, x_{t+1}) - f\left(x'_t, x'_{t+1}\right) \Big) \right) \\
&\qquad \geq \lim_{T\to\infty} \left(\sum_{t=0}^{T} \delta^t \Big(f_1(x_t, x_{t+1})(x_t - x'_t) + f_2(x_t, x_{t+1})\left(x_{t+1} - x'_{t+1}\right) \Big) \right) \quad (4.12)
\end{aligned}
$$

である．さらに $x_0 = x'_0$ であることに注意すれば

$$
\begin{aligned}
&\sum_{t=0}^{T} \delta^t \Big(f_1(x_t, x_{t+1})(x_t - x'_t) + f_2(x_t, x_{t+1})\left(x_{t+1} - x'_{t+1}\right) \Big) \\
&= f_1(x_0, x_1)(x_0 - x'_0) \ + \\
&\quad f_2(x_0, x_1)(x_1 - x'_1) \ + \ \delta f_1(x_1, x_2)(x_1 - x'_1) \ + \\
&\quad \delta f_2(x_1, x_2)(x_2 - x'_2) \ + \ \delta^2 f_1(x_2, x_3)(x_2 - x'_2) \ + \\
&\quad \delta^2 f_2(x_2, x_3)(x_3 - x'_3) \ + \ \delta^3 f_1(x_3, x_4)(x_3 - x'_3) \ + \\
&\qquad \vdots \\
&\quad \delta^{T-1} f_2(x_{T-1}, x_T)(x_T - x'_T) \ + \ \delta^T f_1(x_T, x_{T+1})(x_T - x'_T) \ + \\
&\quad \delta^T f_2(x_T, x_{T+1})\left(x_{T+1} - x'_{T+1}\right) \\
&= \sum_{t=0}^{T-1} \delta^t \Big(f_2(x_t, x_{t+1}) + \delta f_1(x_{t+1}, x_{t+2}) \Big) \left(x_{t+1} - x'_{t+1}\right) \ + \ \delta^T f_2(x_T, x_{T+1})\left(x_{T+1} - x'_{T+1}\right)
\end{aligned}
$$

と整理することができるので，(4.12) の右辺は

$$
\begin{aligned}
&\lim_{T\to\infty} \left(\sum_{t=0}^{T} \delta^t \Big(f_1(x_t, x_{t+1})(x_t - x'_t) + f_2(x_t, x_{t+1})\left(x_{t+1} - x'_{t+1}\right) \Big) \right) \\
&\qquad = \lim_{T\to\infty} \left(\sum_{t=0}^{T-1} \delta^t \Big(f_2(x_t, x_{t+1}) + \delta f_1(x_{t+1}, x_{t+2}) \Big) \left(x_{t+1} - x'_{t+1}\right) \right) \\
&\qquad\qquad + \lim_{T\to\infty} \delta^T f_2(x_T, x_{T+1})\left(x_{T+1} - x'_{T+1}\right) \quad (4.13)
\end{aligned}
$$

である．ここでオイラー方程式 (4.4) により

$$
f_2(x_t, x_{t+1}) + \delta f_1(x_{t+1}, x_{t+2}), \qquad t = 0, 1, 2, \ldots
$$

が成立するから，(4.13) の右辺の第 1 項は

$$\lim_{T\to\infty}\left(\sum_{t=0}^{T-1}\delta^t\Big(f_2\left(x_t,x_{t+1}\right)+\delta f_1\left(x_{t+1},x_{t+2}\right)\Big)\left(x_{t+1}-x'_{t+1}\right)\right)=0 \tag{4.14}$$

であり，第 2 項についても

$$\lim_{T\to\infty}\delta^T f_2\left(x_T,x_{T+1}\right)\left(x_{T+1}-x'_{T+1}\right)=-\lim_{T\to\infty}\delta^{T+1}f_1\left(x_{T+1},x_{T+2}\right)\left(x_{T+1}-x'_{T+1}\right) \tag{4.15}$$

と書き直すことができる．(4.11)，(4.12)，(4.13)，(4.14)，(4.15) を合わせれば，

$$\sum_{t=0}^{\infty}\delta^t f\left(x_t,x_{t+1}\right)-\sum_{t=0}^{\infty}\delta^t f\left(x'_t,x'_{t+1}\right)\geq-\lim_{T\to\infty}\delta^{T+1}f_1\left(x_{T+1},x_{T+2}\right)\left(x_{T+1}-x'_{T+1}\right) \tag{4.16}$$

となる．この (4.16) の右辺は，仮定により $\dfrac{\partial}{\partial x_t}f\left(x_t,x_{t+1}\right)\gg 0$ であり，なおかつ x'_{T+1} は非負ベクトルであるから，

$$\begin{aligned}-\lim_{T\to\infty}\delta^{T+1}f_1\left(x_{T+1},x_{T+2}\right)\left(x_{T+1}-x'_{T+1}\right) &= -\lim_{T\to\infty}\delta^{T+1}f_1\left(x_{T+1},x_{T+2}\right)x_{T+1}\\ &\quad+\lim_{T\to\infty}\delta^{T+1}f_1\left(x_{T+1},x_{T+2}\right)x'_{T+1}\\ &\geq -\lim_{T\to\infty}\delta^{T+1}f_1\left(x_{T+1},x_{T+2}\right)x_{T+1}\end{aligned} \tag{4.17}$$

を満たす．また TC により

$$\lim_{T\to\infty}\delta^T f_1\left(x_T,x_{T+1}\right)x_T=0$$

が成立するから，(4.17) の右辺は 0 であることが言える．よって (4.16) と (4.17) を合わせれば，

$$\sum_{t=0}^{\infty}\delta^t f\left(x_t,x_{t+1}\right)-\sum_{t=0}^{\infty}\delta^t f\left(x'_t,x'_{t+1}\right)\geq 0 \tag{4.18}$$

が成り立つ．最後に，関数 $f\left(x_t,x_{t+1}\right)$ が有界であることから

$$\sum_{t=0}^{\infty}\delta^t f\left(x'_t,x'_{t+1}\right)<\infty$$

が言えるので，(4.18) は

$$\sum_{t=0}^{\infty}\delta^t f\left(x_t,x_{t+1}\right)\geq\sum_{t=0}^{\infty}\delta^t f\left(x'_t,x'_{t+1}\right)$$

したがって，$\{x_t:t=0,1,\dots\}$ は問題 DP$(\bar{x}_0)$ の最適解である． □

4.1 注意. 4.2 定理の中にある「関数 f が有界である」と言う条件は,「任意の実行可能経路 $\{x'_t:t=0,1,\dots\}$ について $\sum_{t=0}^{\infty}\delta^t f\left(x'_t,x'_{t+1}\right)$ が絶対収束する」という条件に等しい.

関数 f が有界であるという条件は，そのままだと使いにくい．一例を挙げると，f が

$$f(x_t,x_{t+1})=u\left(g(x_t)-x_{t-1}\right)$$

のような形状をとるようなケースには応用上頻繁に遭遇するが，効用関数が $u=\ln(z)$ で定義されている場合などは f が $\boldsymbol{R}$ 上で有界でなくなってしまう．したがって，定理をそのまま当てはめようとするとすぐに困ることになる．このような時は，定義域 $\boldsymbol{X}$ を工夫するのが定石である．例えばケーキ食べ問題などでは，実行可能経路においては x が必ず減少してゆくので定義域 $\boldsymbol{X}$ を $\boldsymbol{R}$ 全体にする必要はない．最初から $\boldsymbol{X}$ を $[0,M]$ のような有界な形にとっておいても，最適経路の選択に問題は生じないのである．

4.3 例 (ケーキ食べ問題). ケーキ食べ問題の効用関数が

$$u(z) = \ln(z)$$

で与えられる場合について，具体的に問題を解いてみよう．このとき問題は

$$\begin{aligned} \max_{\{x_t : t=0,1,\dots\}} \quad & \sum_{t=0}^{\infty} \delta^t \ln(x_t - x_{t+1}) \\ \text{subject to} \quad & x_t - x_{t+1} \geq 0, \quad t = 0, 1, \dots \\ & x_0 = \bar{x} \end{aligned}$$

である．まず必要条件から考えると，オイラー方程式は

$$-\frac{1}{x_t - x_{t+1}} + \delta \frac{1}{x_{t+1} - x_{t+2}} = 0, \quad t = 0, 1, \dots$$

である．$c_t = x_t - x_{t+1}$ であったから，これは

$$-\frac{1}{c_t} + \delta \frac{1}{c_{t+1}} = 0 \quad \Leftrightarrow \quad c_{t+1} = \delta c_t \tag{4.19}$$

と書ける．(4.19) が $t = 0, 1, 2, \dots$ で成立するから，

$$c_t = \delta^t c_0, \quad t = 0, 1, \dots \tag{4.20}$$

の関係が成立し，最適な消費経路はこれに従うはずである．

しかしながら，(4.20) を満たす消費経路は無数に存在する．(4.20) が表す条件は c_0 に依存しており，その c_0 の選び方については何も言っていないからである．従って，オイラー方程式だけで解を特定することはできず，解を一意的に決定するためには外生的に与えられた値，すなわち x_0 を用いた条件が必要となる．そこで (4.20) に加えて,「最終的にケーキを食べつくさなければならない」という条件を設けることによって，解を絞り込むことを考える．(4.20) を満たす消費経路上のケーキの総消費量を考えると

$$\begin{aligned} \sum_{t=0}^{\infty} c_t &= c_0 \left(1 + \delta + \delta^2 + \dots\right) \\ &= c_0 \frac{1}{1-\delta} \end{aligned}$$

であるから，最終的にケーキを食べつくすためには

$$c_0 \frac{1}{1-\delta} = x_0 \quad \Leftrightarrow \quad x_0 - c_0 = \delta x_0$$

が満たされなければならない．ここで $c_0 = x_0 - x_1$ の関係を用いれば,

$$\begin{aligned} x_1 &= x_0 - c_0 \\ &= \delta x_0 \end{aligned}$$

さらに $c_1 = x_1 - x_2$ の関係を用いれば,

$$\begin{aligned} x_2 &= x_1 - c_1 \\ &= \delta (x_0 - c_0) \\ &= \delta^2 x_0 \end{aligned}$$

同様の作業を繰り返せば，

$$x_t = \delta^t x_0, \quad t = 0, 1, \ldots \tag{4.21}$$

の関係を得る．ここで x_0 は外生的に与えられているため，この条件を満たす経路は一意に決まる．

以上から，(4.21) を満たす実行可能経路 $\{x_t : t = 0, 1, \ldots\}$ が問題の解であると予想できる．次に十分性を確認することによって，この経路が最適解であることを確かめよう．いま $f(z) = \ln(z)$ と定義していたから，4.2 定理の条件の中で，関数 f が凹関数であることと，$\dfrac{\partial}{\partial x_t} f\left(x_t, x_{t+1}\right) \gg 0$ が満たされることは容易に確認できる．また上で述べたように，あらかじめ定義域を有界にとっておけば $f\left(x_t, x_{t+1}\right)$ も実質的に有界であると考えてよい．よって，あとは TC が満たされていることを確認できさえすれば十分である．そこで TC をチェックすると

$$\begin{aligned}
\tfrac{\partial}{\partial x_t} f\left(x_t, x_{t+1}\right) &= \frac{1}{x_t - x_{t+1}} \\
&= \frac{1}{c_t} \\
&= \frac{1}{\delta^t c_0}
\end{aligned}$$

であるから，

$$\begin{aligned}
\lim_{t \to \infty} \delta^t \frac{\partial}{\partial x_t} f\left(x_t, x_{t+1}\right) x_t &= \lim_{t \to \infty} \delta^t \frac{1}{\delta^t c_0} \delta^t x_0 \\
&= \lim_{t \to \infty} \delta^t \frac{x_0}{c_0} \\
&= \lim_{t \to \infty} \delta^t \frac{1}{1 - \delta} \\
&= 0
\end{aligned}$$

が成立し，TC が満たされていることを確認できる．以上から，(4.21) に従って生成される経路が問題の解である．

4.2.4 ベルマン方程式

オイラー方程式を用いた解法は便利であるが，目的関数や制約式が微分可能でない場合や凸性の条件を満たさないケースには当てはめることができない．従って，関数の微分可能性などが満たされないような場合には別の解法を用いる必要があり，それがこの節で扱うベルマン方程式（Bellman equation）と呼ばれるものである[18]．

4.2 定義 (ベルマン方程式)． 関数 $v : \boldsymbol{X} \to \boldsymbol{R}$ について

$$\forall x \in \boldsymbol{X}, \qquad v(x) = \max_{y:\, \Gamma(x,y) \geq 0} \left\{ f(x, y) + \delta v(y) \right\} \tag{4.22}$$

で表される関数方程式を，ベルマン方程式という．

ベルマン方程式は，関数 v を変数とする関数方程式である．したがってこの方程式の解くことは，(4.22) を満たす関数 v を見つけ出すことに相当する．

4.3 定義 (Policy Function)． ある関数 $v(x)$ が，ベルマン方程式を満たしているとしよう．この時，

$$\phi(x) \in \max_{y:\, \Gamma(x,y) \geq 0} \left\{ f(x, y) + \delta v(y) \right\}$$

で定義される関数 $\phi(x)$ を Policy Function（最適政策関数，最適戦略関数）と呼ぶ．

[18] よって，この節の議論では微分可能性を前提としない．

4.3 定理 (最適性原理：必要性). 問題 DP($\bar{x}_0$) が，全ての $\bar{x}_0 \in \boldsymbol{X}$ に関して解 $\{x_t^* : t = 0, 1, \dots\}$ を持つとする．このとき

$$v^*(\bar{x}_0) = \sum_{t=0}^{\infty} \delta^t f\left(x_t^*, x_{t+1}^*\right) \tag{4.23}$$

によって定義される関数 v^* は，ベルマン方程式 (4.22) を満たす．

(4.23) によって定義される関数 v^* は Value Function（価値関数）と呼ばれるもので，$x_0 = \bar{x}_0$ を所与として目的関数が実現し得る最大の「価値」を表している．この 4.3 定理は，実行可能プラン $\{x_t^* : t = 0, 1, \dots\}$ が問題 DP($\bar{x}_0$) の解であるならば，それに対応する Value Function はベルマン方程式を満たしていなければならないという，最適性の必要条件を示している．

4.3 定理の証明. 仮定により全ての $\bar{x}_0 \in \boldsymbol{X}$ に対して問題 DP($\bar{x}_0$) の解 $\{\bar{x}_0, x_1^*, x_2^*, x_3^*, \dots\}$ が存在するので，任意の $\bar{x}_0 \in \boldsymbol{X}$ について Value Function v^* は

$$\begin{aligned} v^*(\bar{x}_0) &= \sum_{t=0}^{\infty} \delta^t f\left(x_t^*, x_{t+1}^*\right) \\ &= f\left(\bar{x}_0, x_1^*\right) + \delta f\left(x_1^*, x_2^*\right) + \delta^2 f\left(x_2^*, x_3^*\right) + \dots \end{aligned} \tag{4.24}$$

のように書ける．$v^*(\bar{x}_0)$ について，背理法の仮定として

$$v^*(\bar{x}_0) < f\left(\bar{x}_0, \hat{x}_1\right) + \delta v^*(\hat{x}_1) \tag{4.25}$$

を満たす $\hat{x}_1 \in \{x \in \boldsymbol{X} | \, \Gamma(\bar{x}_0, x) \geq 0\}$ が存在したとしよう．するとこの $\hat{x}_1$ を初期値とした問題

$$\begin{aligned} \max_{\{x_t : t = 2, 3, \dots\}} \quad & \sum_{t=0}^{\infty} \delta^t f\left(x_{t+1}, x_{t+2}\right) \\ \text{subject to} \quad & \Gamma\left(x_t, x_{t+1}\right) \geq 0, \quad t = 1, 2, \dots \\ & x_1 = \hat{x}_1 \end{aligned}$$

を考えることができ，その解を $\{\hat{x}_1, \hat{x}_2, \hat{x}_3, \dots\}$ と置けば，Value Function の定義により，

$$v^*(\hat{x}_1) = f\left(\hat{x}_1, \hat{x}_2\right) + \delta f\left(\hat{x}_2, \hat{x}_3\right) + \delta^2 f\left(\hat{x}_3, \hat{x}_4\right) + \dots \tag{4.26}$$

と書ける．この時 (4.25) は，(4.24) および (4.26) により，

$$\begin{aligned} f\left(\bar{x}_0, x_1^*\right) + \delta f\left(x_1^*, x_2^*\right) + \dots &< f\left(\bar{x}_0, \hat{x}_1\right) + \delta v^*(\hat{x}_1) \\ &= f\left(\bar{x}_0, \hat{x}_1\right) + \delta f\left(\hat{x}_1, \hat{x}_2\right) + \delta^2 f\left(\hat{x}_2, \hat{x}_3\right) + \dots \end{aligned}$$

なる実行可能プラン $\{\bar{x}_0, \hat{x}_1, \hat{x}_2, \dots\}$ が存在することを意味する．しかしこれは，$\{\bar{x}_0, x_1^*, x_2^*, \dots\}$ が問題 DP($\bar{x}_0$) の解であることに矛盾する．従って仮定 (4.25) は誤りで

$$\forall x_1 \in \{\Gamma(\bar{x}_0, x_1) \geq 0\}, \quad v^*(\bar{x}_0) \geq f\left(\bar{x}_0, x_1\right) + \delta v^*(x_1) \tag{4.27}$$

でなければならない．$\bar{x}_0 \in \boldsymbol{X}$ が任意であったことに注意すれば，これは

$$\forall \bar{x}_0 \in \boldsymbol{X}, \quad v^*(\bar{x}_0) = \max_{x_1 :\, \Gamma(\bar{x}_0, x_1) \geq 0} \Big\{ f(\bar{x}_0, x_1) + \delta v(x_1) \Big\}$$

に他ならず，v^* がベルマン方程式 (4.22) を満たすことが確認できる．□

4.4 例 (ケーキ食べ問題). ケーキ食べ問題を例にして，4.3 定理を確認しよう．効用関数が

$$u(z) = \ln(z)$$

で定義されるケースでは，最適解は

$$x_t = \delta^t x_0 \quad \Leftrightarrow \quad c_t = (1-\delta)\delta^t x_0$$

であったから，Value Function は

$$v^*(x) = \sum_{t=0}^{\infty} \delta^t \ln\big((1-\delta)\delta^t x\big)$$

となる．したがってベルマン方程式は

$$\forall x \in \boldsymbol{X}, \qquad v^*(x) \;=\; \max_{y:\,\Gamma(x,y)\geq 0} \Big\{ \ln(x-y) + \delta \sum_{t=0}^{\infty} \delta^t \ln\big((1-\delta)\delta^t y\big) \Big\} \tag{4.28}$$

である．ここで

$$\begin{aligned} \max_{y} \quad & \ln(x-y) + \delta \sum_{t=0}^{\infty} \delta^t \ln\big((1-\delta)\delta^t y\big) \\ \text{subject to} \quad & \Gamma(x,y) \geq 0 \end{aligned} \tag{4.29}$$

は通常の凹関数の最大化問題であるから，y に関する 1 階の条件を確認すればよい．

$$\ln\big((1-\delta)\delta^t y\big) = \ln y + \ln(1-\delta)\delta^t$$

に注意すれば，1 階の条件は

$$-\frac{1}{x-y} + \delta \sum_{t=0}^{\infty} \delta^t \frac{1}{y} = 0$$

である．これを y について解くことにより

$$y = \delta x$$

が最大化問題の解であることが分かる．従って，これを (4.29) の目的関数に代入することにより

$$\begin{aligned} \max_{y:\,\Gamma(x,y)\geq 0} \Big\{ \ln(x-y) + \delta \sum_{t=0}^{\infty} \delta^t \ln\big((1-\delta)\delta^t y\big) \Big\} &= \ln(x-\delta x) + \delta \sum_{t=0}^{\infty} \delta^t \ln\big((1-\delta)\delta^t \delta x\big) \\ &= \ln\big((1-\delta)x\big) + \sum_{t=0}^{\infty} \delta^{t+1} \ln\big((1-\delta)\delta^{t+1} x\big) \\ &= \sum_{t=0}^{\infty} \delta^t \ln\big((1-\delta)\delta^t x\big) \\ &= v^*(x) \end{aligned}$$

となるから，ベルマン方程式 (4.28) が満たされていることが確認できる．

練習問題 4.3. 効用関数が

$$u(z) = \ln(z)$$

で定義されるとき，練習問題 4.1 および練習問題 4.2 について，Value Function を求めよ．

4.3 定理により，実行可能プラン $\{x^*_t : t = 0, 1, \dots\}$ が問題 $\mathrm{DP}(\bar{x}_0)$ の解であるならば，対応する Value Function はベルマン方程式を満たすことが分かる．実は逆に，ベルマン方程式を満たす関数 v が，ある種の境界条件が満たされる限りで，問題 $\mathrm{DP}(\bar{x}_0)$ の解に対応する Value Function であることも示せる．つまり，最適化問題の解はベルマン方程式を解くことによっても求められるのである．

4.4 定理 (**最適性原理：十分性**). ある関数 v について，

1. v はベルマン方程式 (4.22) の解である．
2. 任意の実行可能プラン $\{x_t : t = 0, 1, \dots\}$ について $\lim_{t\to\infty} \delta^t v(x_t) = 0$ が成り立つ[19].

の条件が満たされる時，対応する政策関数 ϕ に従って生成される実行可能プラン $\{\bar{x}_0, \phi(\bar{x}_0), \phi^2(\bar{x}_0), \dots\}$ は問題 $\mathrm{DP}(\bar{x}_0)$ の解である．

4.4 定理の証明. 仮定により v はベルマン方程式を満たすから，

$$\forall x \in \boldsymbol{X}, \qquad v(x) \;=\; \max_{y:\ \Gamma(x,y)\ge 0} \Big\{ f(x,y) + \delta v(y) \Big\} \tag{4.30}$$

が成立する．右辺の問題

$$\begin{array}{cl} \displaystyle\max_{y} & f(x,y) + \delta v(y) \\ \text{subject to} & \Gamma(x,y) \ge 0 \end{array}$$

の解が政策関数 $\phi(x)$ であるから，(4.30) は ϕ を用いて

$$\forall x \in \boldsymbol{X}, \qquad v(x) = f\big(x, \phi(x)\big) + \delta v\big(\phi(x)\big) \tag{4.31}$$

と表せる．さらに $\phi(x) \in \boldsymbol{X}$ であるから (4.31) は $x = \phi(x)$ としても成立し

$$\forall x \in \boldsymbol{X}, \qquad v\big(\phi(x)\big) = f\big(\phi(x), \phi\circ\phi(x)\big) + \delta v\big(\phi\circ\phi(x)\big)$$

である．この操作を繰り返し行えば，

$$\begin{aligned} \forall x \in \boldsymbol{X}, \qquad v(x) &= f\big(x, \phi(x)\big) + \delta v\big(\phi(x)\big) \\ &= f\big(x, \phi(x)\big) + \delta f\big(\phi(x), \phi^2(x)\big) + \delta^2 v\big(\phi^2(x)\big) \\ &\ \ \vdots \\ &= \sum_{t=0}^{\infty} \delta^t f\big(\phi^t(x), \phi^{t+1}(x)\big) \end{aligned}$$

と書ける．ただし

$$\begin{aligned} \phi^0(x) &= x \\ \phi^t(x) &= \underbrace{\phi\circ\phi\circ\cdots\circ\phi}_{t}(x), \qquad t \ge 1 \end{aligned}$$

である．x は任意であったから，$x = \bar{x}_0$ とすれば，

$$v(\bar{x}_0) = \sum_{t=0}^{\infty} \delta^t f\big(\phi^t(\bar{x}_0), \phi^{t+1}(\bar{x}_0)\big) \tag{4.32}$$

であり，(4.32) における $\{\bar{x}_0, \phi(\bar{x}_0), \phi^2(\bar{x}_0), \dots\}$ はその構成方法からして明らかに実行可能プランである．

[19]詳しくは解説しないが，この条件は横断性条件と基本的に同じであることを示すことができる．

一方，v がベルマン方程式を満たすことから，$\Gamma(\bar{x}_0, x_1) \geq 0$ を満たす任意の x_1 について，

$$v(\bar{x}_0) \geq f(\bar{x}_0, x_1) + \delta v(x_1) \tag{4.33}$$

が成り立つ．さらに，ここで任意に選んだ x_1 を所与として，$\Gamma(x_1, x_2) \geq 0$ を満たす任意の x_2 について

$$v(x_1) \geq f(x_1, x_2) + \delta v(x_2)$$

が成り立つから，(4.33) は

$$\begin{aligned} v(\bar{x}_0) &\geq f(\bar{x}_0, x_1) + \delta v(x_1) \\ &\geq f(\bar{x}_0, x_1) + \delta f(x_1, x_2) + \delta^2 v(x_2) \end{aligned}$$

と書ける．この操作を繰り返し行うことによって，任意の実行可能プラン $\{x_t : t = 0, 1, \dots\}$ に対して

$$\begin{aligned} v(\bar{x}_0) &\geq f(\bar{x}_0, x_1) + \delta v(x_1) \\ &\geq f(\bar{x}_0, x_1) + \delta f(x_1, x_2) + \delta^2 v(x_2) \\ &\geq f(\bar{x}_0, x_1) + \delta f(x_1, x_2) + \delta^2 f(x_2, x_3) + \delta^3 v(x_3) \\ &\vdots \\ &\geq \sum_{t=0}^{T} \delta^t f(x_t, x_{t+1}) + \delta^T v(x_T) \\ &\vdots \\ &\geq \sum_{t=0}^{\infty} \delta^t f(x_t, x_{t+1}) \end{aligned} \tag{4.34}$$

を得る．ただし最後の不等式は，仮定により

$$\begin{aligned} \lim_{T\to\infty} \left(\sum_{t=0}^{T} \delta^t f(x_t, x_{t+1}) + \delta^T v(x_T) \right) &= \lim_{T\to\infty} \sum_{t=0}^{T} \delta^t f(x_t, x_{t+1}) + \lim_{T\to\infty} \delta^T v(x_T) \\ &= \sum_{t=0}^{\infty} \delta^t f(x_t, x_{t+1}) \end{aligned}$$

が成り立つことによる．

(4.32) および (4.34) により，任意の実行可能プラン $\{x_t : t = 0, 1, \dots\}$ に対して

$$\sum_{t=0}^{\infty} \delta^t f\big(\phi^t(\bar{x}_0), \phi^{t+1}(\bar{x}_0)\big) \geq \sum_{t=0}^{\infty} \delta^t f(x_t, x_{t+1})$$

が成立する．これは，政策関数 ϕ に従って生成される経路 $\{\bar{x}_0, \phi(\bar{x}_0), \phi^2(\bar{x}_0), \dots\}$ が問題 $\mathrm{DP}(\bar{x}_0)$ の解であることに他ならない． □

ベルマン方程式は微分可能性や凸性に関する仮定を必要としないため，幅広い問題に適用できる．ただしベルマン方程式は関数方程式であるから，一般には式の変形を施すだけでは解くことができない．そのため問題の解や Policy Function がある程度予想できる場合には，それに基づいて Value Function v を計算し，v がベルマン方程式を満たすかどうかを確認することによって問題を解くことになる．

しかし当然ながら，中にはもっともらしい予想が困難なケースもある．そのような場合には，次のようなシステマティックな方法でベルマン方程式を解くことができる．まず関数 $v : \boldsymbol{X} \to \boldsymbol{R}$ を適当に予想して，

$$Tv(x) \;:=\; \max_{y:\Gamma(x,y)\geq 0} \left\{ f(x,y) + \delta v(y) \right\} \tag{4.35}$$

のように $Tv(x)$ を置く．ここでもし，右辺の問題

$$\begin{array}{ll} \displaystyle\max_{y} & f(x,y) + \delta v(y) \\ \text{subject to} & \Gamma(x,y) \geq 0 \end{array}$$

を解いた結果として

$$\forall x \in \boldsymbol{X}, \qquad Tv(x) = v(x) \tag{4.36}$$

が成立すれば，これはすなわち

$$\forall x \in \boldsymbol{X}, \qquad v(x) \;=\; \max_{y:\Gamma(x,y)\geq 0} \left\{ f(x,y) + \delta v(y) \right\}$$

が成立することに他ならないから，偶然にも最初に予想した v がベルマン方程式の解であったことになる．しかし v は適当に予想したものであるため，一般に (4.36) は成り立たない．そこで予想を修正することになるのであるが，修正された新たな予想として，(4.35) で定義した Tv を用いるのである．そして最初の v と同様にして，

$$T^2v(x) \;:=\; \max_{y:\Gamma(x,y)\geq 0} \left\{ f(x,y) + \delta Tv(y) \right\} \tag{4.37}$$

のように $T^2v(x)$ を置き，右辺の問題を解く．その結果として

$$\forall x \in \boldsymbol{X}, \qquad T^2v(x) = Tv(x) \tag{4.38}$$

が成立すれば，これはすなわち

$$\forall x \in \boldsymbol{X}, \qquad Tv(x) \;=\; \max_{y:\Gamma(x,y)\geq 0} \left\{ f(x,y) + \delta Tv(y) \right\}$$

が成立することに他ならないから，予想を修正した Tv がベルマン方程式の解である．もし (4.38) が成り立たない場合には，今度はさらに修正した予想として，(4.37) で定義した T^2v を用いる．同様の手続きを繰り返すことによって，$T^n v$ を得ることができる．そしてこの予想は最終的にある関数 v^* へと収束し，その v^* についてベルマン方程式が満たされるのである．最初の v を任意に選んだことを考えると，これはずいぶんと乱暴な議論にも見えるが，次の定理はこのような素朴な方法でベルマン方程式が「解ける」ことを示している．

4.5 定理. 問題 DP$(\bar{x}_0)$ における目的関数 f が連続で，定義域 $\boldsymbol{X}$ が有界な閉集合であるとする．この時，全ての $x \in \boldsymbol{X}$ について $T^n v(x)$ の極限が存在し，その極限を

$$v^*(x) := \lim_{n\to\infty} T^n v(x)$$

と置くと，

$$\forall x \in \boldsymbol{X}, \qquad v^*(x) \;=\; \max_{y:\Gamma(x,y)\geq 0} \left\{ f(x,y) + \delta v^*(y) \right\}$$

が成立する．なお，v^* は v の選び方に依存しない．

4.5 定理は，最初の予想 v の選び方によらず，$T^n v$ がベルマン方程式の解 v^* に収束することを示している．つまりどのような関数 v から始めても，結局は v^* に行き着くのである．

4.3 オイラー方程式とベルマン方程式の関係

ベルマン方程式とオイラー方程式は，いずれも無限次元の最適化問題に対する解法である．同一の問題に対する解の条件を与えるものである以上，それらの間には何らかの関係性があると予想されよう．実際，二つの方程式は密接に関連しており，次のような手順でベルマン方程式からオイラー方程式を導出することができる．

いまベルマン方程式の解として，関数 $v:\boldsymbol{X}\to\boldsymbol{R}$ が得られたとしよう．つまりこの v について，

$$\forall x\in\boldsymbol{X},\qquad v(x)\ =\max_{y:\,\Gamma(x,y)\geq 0}\Big\{f(x,y)+\delta v(y)\Big\}$$

が成立したとする．ここで v の微分可能性と，右辺の問題

$$\begin{array}{ll}\displaystyle\max_{y} & f(x,y)+\delta v(y)\\ \text{subject to} & \Gamma(x,y)\geq 0\end{array}$$

の内点解を仮定する．すると，Policy Function を $\phi(x)$ と書けば，1 階の条件から

$$\forall x\in\boldsymbol{X},\quad f_2(x,\phi(x))+\delta v'(\phi(x))=0 \tag{4.39}$$

が成り立つはずである．一方，ベルマン方程式でパラメタを x と見れば，包絡線定理により

$$\frac{d}{dx}v(x)=\frac{\partial}{\partial x}\Big\{f(x,y)+\delta v(y)\Big\}\bigg|_{y=\phi(x)}$$

すなわち

$$v'(x)=f_1(x,\phi(x))$$

が成立するから，これを (4.39) に代入することにより

$$\forall x\in\boldsymbol{X},\quad f_2(x,\phi(x))+\delta f_1(\phi(x),\phi^2(x))=0 \tag{4.40}$$

を得る．$\{x_0,x_1,x_2,\dots\}$ が最適なプランであるとすれば，Policy Function の定義により

$$\phi(x_{t-1})=x_t,\quad \phi^2(x_{t-1})=x_{t+1},\quad t=1,2,\dots$$

と表せるから，(4.40) において $x=x_{t-1}\in\boldsymbol{X}$ と置くことで，オイラー方程式

$$f_2\left(x_{t-1},x_t\right)+\delta f_1\left(x_t,x_{t+1}\right)=0,\quad t=1,2,\dots$$

が導かれる．

4.4 応用：簡単な仕入れ問題

ベルマン方程式を活用する例として，次のような「仕入れ問題」を考えよう．いま x を非負の整数として，x 単位の財を在庫として持っている店があるとする．この店では在庫さえあれば，1 期間あたり 1 単位の財が価格 p で自動的に売れてゆく．在庫を増やすためには，仕入れのための固定費用 $a>0$ に加えて単位あたりの限界費用 $c>0$ が必要で，例えば y 単位の財を仕入れる費用は $a+cy$ となる．また，仕入れた財が在庫に追加されるのは次の期の期首である．割引因子を δ で表す．

この問題を DP の形に書いてみる．在庫の量を変数と考え，t 期の期首における在庫を $x_t\in\boldsymbol{X}$ で表す．ただし $\boldsymbol{X}$ は非負整数の集合，すなわち $\boldsymbol{X}=\{0,1,2,\dots\}$ である．$x_t\geq 1$ であれば t 期に 1 単位売れるので，t 期末の在庫は x_t-1 となる．$t+1$ 期の期首の在庫を x_{t+1} で表すと，t 期の仕入れ

量は $x_{t+1}-(x_t-1)$ で表すことができるから，t 期に発生する仕入れの費用は $a+c(x_{t+1}-x_t+1)$ である．

目的関数 f は 1 期間の利益，すなわち

$$f(x_t,x_{t+1})=\begin{cases}\pi(x_t)-\bigl(a+c(x_{t+1}-x_t+1)\bigr) & (\ x_{t+1}-x_t+1>0\text{ のとき }),\\ \pi(x_t) & (\ x_{t+1}-x_t+1\leq 0\text{ のとき }).\end{cases}$$

とおく．ただしここで

$$\pi(x_t)=\begin{cases}p & (\ x_t>0\text{ のとき }),\\ 0 & (\ x_t\leq 0\text{ のとき }).\end{cases}$$

である．また制約条件は

$$\Gamma(x_t,x_{t+1})=x_{t+1}-x_t+1\geq 0$$

と書ける．

練習問題 4.4. 上で定義した f, π, Γ からなる DP が，仕入れ問題に対応していることを確認せよ．

仕入れる度に固定費用 a が発生するので，仕入れの回数はなるべく少ないほうがよいことは容易に分かる．しかし一度に大量に仕入れた場合，仕入れに伴うコストはすぐに発生する一方で売り上げは 1 期間に 1 単位分しか生じないため，利益は割引率の分だけ目減りすることになる．つまり，大量に仕入れればよいという訳でもないのである．したがって最も自然な仕入れ方は,（1）在庫が 2 単位以上あるうちには仕入れをせず,（2）在庫が 1 単位，あるいは 0 単位になった時点で「適度な」量を仕入れる，というものになろう．そこで以下の議論では，今述べた仕入れ方が最適な方法であると予想して Value Function の候補を計算し，それがベルマン方程式を満たすことを確認するという作戦をとることにする．

このような手順で議論を進めるためには，まず「適度な」量を確定しておく必要がある．n を正の整数として n 単位だけ仕入れ，その後は仕入れを行わずにいた場合の割引利益は

$$\begin{aligned}r_n &:= p(\delta+\delta^2+\cdots+\delta^n)-(a+cn)\\ &= p\frac{1-\delta^n}{1-\delta}-(a+cn)\end{aligned}$$

となる．仕入れた在庫が尽きる度に（すなわち n 期ごとに）これを繰り返すと，割引総利益は

$$\begin{aligned}r_n\Bigl(1+\delta^n+\bigl(\delta^n\bigr)^2+\dots\Bigr) &= r_n\frac{1}{1-\delta^n}\\ &= p\frac{1}{1-\delta}-\frac{a+cn}{1-\delta^n}\end{aligned}\tag{4.41}$$

である．

1 単位だけ仕入れることで得る利益は $p\delta$ で，そのために必要なコストは $a+c$ である．したがって $p\delta>a+c$ であれば，仕入れをしない（したがって将来に渡って販売を行わない）よりも $n=1$ とした方が利益 (4.41) は大きくなるから，仕入れを全く行わないというやり方が最適となることはない．一方で，(4.41) の値はある数 n を境に減少してゆくことが分かるので，割引総利益を最大にするような仕入れ数 n^* が存在すると言える．

練習問題 4.5. 上記の議論を確認せよ.（ヒント：n を実数と見れば，ある n を超えると (4.41) が減少に転じることが確認できるはず.）

これ以降の議論では $p\delta>a+c$ を仮定し，(4.41) の最大値を Π^* と置くことにする．つまり

$$\Pi^*:=p\frac{1}{1-\delta}-\frac{a+cn^*}{1-\delta^{n^*}}$$

と書く．ここで n^* は，在庫が尽きるたびに仕入れるという方針をとった場合の利益 (4.41) を最大にする仕入れ量であるが，これが上で言う「適度な」量，すなわち DP の解に対応する仕入れ量であるとすれば，Value Function v は

$$
\begin{aligned}
v(1) &= p + \Pi^* \\
v(2) &= p + \delta v(1) = (p + \delta p) + \delta \Pi^* \\
v(3) &= p + \delta v(2) = (p + \delta p + \delta^2 p) + \delta^2 \Pi^* \\
&\vdots \\
v(x) &= p + \delta v(x-1) = (p + \delta p + \ldots \delta^{x-1} p) + \delta^{x-1} \Pi^* \qquad (4.42)
\end{aligned}
$$

すなわち $x \geq 1$ に対して

$$
v(x) = p\frac{1-\delta^x}{1-\delta} + \delta^{x-1}\Pi^* = \frac{1}{1-\delta}p + \left(p - \frac{a+cn^*}{1-\delta^{n^*}}\right)\delta^{x-1} \qquad (4.43)
$$

のように書けるはずである．

練習問題 4.6. $v(0)$ がどのようになるか書きなさい．

任意の実行可能解について，$\lim_{t\to\infty} \delta^t v(x_t) = 0$ が成り立つことに注意しよう．なぜなら，$x \to \infty$ のとき $\left(p - \frac{a+cn^*}{1-\delta^{n^*}}\right)\delta^{x-1}$ はゼロに収束するため，任意の実行可能解に対して $v(x_t)$ が発散することはないからである．

練習問題 4.7. 上の主張をより丁寧に示せ．

したがって 4.4 定理により，後は (4.43) で予想した v がベルマン方程式を満たすことが示されれば，上で記述した自然な戦略が最適な仕入れ方法であることが言える．$x \geq 1$ の在庫を抱えた状態を考えよう．もしも $y \geq 1$ 単位の仕入れを行えば今期の利益は $p - (a + cy)$，次期以降の総割引価値は $\delta v(x - 1 + y)$ となる．一方，仕入れを行わなければ今期の利益は p，次期以降の総割引価値は $\delta v(x-1)$ である．よって $x \geq 1$ の時のベルマン方程式は

$$
v(x) = \max\left\{p + \delta v(x-1),\ \max_{y\geq 1}\Big(p - (a+cy) + \delta v(x-1+y)\Big)\right\} \qquad (4.44)
$$

と書ける．

練習問題 4.8. $y \geq 1$ において，$p - (a+cy) + \delta v(x-1+y)$ が y に関する減少関数となる（したがって，(4.44) における右辺の第 2 項は $y = 1$ の時に最大化される）ことを示せ．また，$x \geq 1$ ならば $p + \delta v(x-1) \geq p - (a+c) + \delta v(x)$ であることを示せ.（ヒント：(4.43) を直接代入する．また，Π^* の決め方から $\frac{a+cn^*}{1-\delta^{n^*}} \leq \frac{a+cn}{1-\delta^n}$ であり，特に $n = 1$ の時にもこれが成立するすることに注意すること.）

以上から，(4.44) の成立を示すためには $x \geq 1$ の時に $v(x) = p + \delta v(x-1)$ となることを示せば十分であると言えるが，これは (4.42) から直接確認することができる．

練習問題 4.9. $x = 0$ の時もベルマン方程式が成立していることを示せ．

5 差分方程式と動学系の安定性

Policy Function は，初期値を所与として，目的関数を最大化する変数が辿る動学的な経路を描写している．したがって，最適な経路上で状態変数がある一定の水準に収束するのか，それとも発散するのかといったことは，Policy Function の形状と初期値の水準に依存している．この章[20]では Policy Function のような動学を規定する方程式について，どのような条件が満たされる時に状態変数が安定的に推移するのか，あるいはしないのかを考える．

まずは，動学系に対する一般的な表現方法を導入しよう．n 次元ベクトルを実数に対応させる n 個の関数 $h_k : \boldsymbol{R}^n \to \boldsymbol{R}$ $(k = 1, 2, \ldots, n)$ を考え，いずれも連続微分可能であるとする．これら $h_1, h_2, \ldots, h_n$ を用いて，関数 $h : \boldsymbol{R}^n \to \boldsymbol{R}^n$ を

$$\forall x \in \boldsymbol{R}^n, \quad h(x) = \big(h_1(x),\ h_2(x),\ \ldots,\ h_n(x)\big)$$

と定義する．するとこの h を用いた差分方程式

$$x_{t+1} = h(x_t), \quad t = 0, 1, 2, \ldots \tag{5.1}$$

は，n 次元ベクトル x_0 を初期値としたベクトル列を定めることになる．つまり，方程式 (5.1) と初期値 x_0 が与えられれば，

$$\big\{\ x_0, h(x_0), h^2(x_0), h^3(x_0), \ldots\ \big\}$$

という一つの動学経路が定まるのである．以下では，このようにして (5.1) によって定まるベクトル列 $\{x_t : t = 0, 1, 2, \ldots\}$ の動きを考えてゆくが，その前にいくつか基本的な概念を確認しておこう．

5.1 定義 (定常点). (5.1) で定義される動学において，ある $\bar{x} \in \boldsymbol{R}^n$ が

$$\bar{x} = h(\bar{x})$$

を満たす時，$\bar{x}$ をこの動学の定常点（stationary point）という．

5.2 定義 (定常点の安定性). $\bar{x} \in \boldsymbol{R}^n$ を，(5.1) で定義される動学の定常点の一つとする．定義域上の任意の点 x_0 を初期値として，x_0 と (5.1) とで定まるベクトル列 $\{x_t : t = 0, 1, 2, \ldots\}$ が

$$\lim_{t \to \infty} x_t = \bar{x}$$

を満たすとき，点 $\bar{x}$ は大域において安定的（globally stable）であるという．また，$\bar{x}$ のある近傍が存在し，その近傍に含まれる任意の点 x_0 を初期値として，x_0 と (5.1) とで定まるベクトル列 $\{x_t : t = 0, 1, 2, \ldots\}$ が

$$\lim_{t \to \infty} x_t = \bar{x}$$

を満たすとき，点 $\bar{x}$ は局所において安定的（locally stable）であるという．

定常点とは，一端そこに至るとそれ以降は動学がその水準で推移し続けるような点である．この定常点に注目することによって経済の状態がどのように推移してゆくかを予想できるため，動学的な分析においては定常点の分析が重要な役割を果たす．

[20] この章の内容をより詳しく知るためには，Stokey-Lucas, chapter 6 を見よ．

5.1 線形静学

5.1.1 1 変数の線形動学

まずは，関数 h が線形関数である場合を考えよう．さしあたって変数の数を 1 とし，最も簡単な

$$h(x) = \beta x,$$

というケースで定常点を求めてみる．ただし，β は $\beta \neq 1, 0$ を満たす定数とする．定常点を見つけるには

$$x = h(x)$$

を解けばよいから，

$$\begin{aligned} x = \beta x \quad &\Leftrightarrow \quad (1-\beta)x = 0 \\ &\Leftrightarrow \quad x = 0 \qquad (\beta \neq 1 \text{ による}) \end{aligned}$$

となり，$\bar{x} = 0$ が定常点であると分かる．次に，この定常点 $\bar{x}$ が安定的であるための条件を考えよう．数列 $\{x_t : t = 0, 1, 2, \ldots\}$ の動きを定める差分方程式は

$$x_{t+1} = \beta x_t$$

であるから，初期値を x_0 として

$$x_t = \beta^t x_0, \quad t = 0, 1, 2, \ldots$$

が成立する．従って任意の初期値 x_0 に対して

$$\lim_{t\to\infty} x_t = \begin{cases} 0 & (\ |\beta| < 1 \text{ のとき }), \\ \text{発散} & (\ |\beta| > 1 \text{ のとき }). \end{cases}$$

である．これは，β の絶対値が 1 より小さい時，そしてその時のみ $\bar{x}$ が安定的であることを意味する．

関数の形をより一般化して，定数項 $c \neq 0$ を加えた

$$h(x) = \beta x + c$$

についてはどうだろうか．先ほどと同様にして $h(x) = x$ と置けば，

$$\bar{x} = \frac{c}{1-\beta}$$

という定常点が求められる．ここで

$$z_t = x_t - \bar{x}, \qquad t = 0, 1, 2, \ldots$$

と $\{z_t\}$ を定めると

$$\begin{aligned} z_{t+1} &= x_{t+1} - \bar{x} \\ &= \beta x_t + c - \frac{c}{1-\beta} \\ &= \beta x_t - \beta \frac{c}{1-\beta} \\ &= \beta(x_t - \bar{x}) \\ &= \beta z_t \end{aligned}$$

の関係が成立するので，

$$\lim_{t\to\infty} z_t = \begin{cases} 0 & (\ |\beta| < 1 \text{ のとき }), \\ \text{発散} & (\ |\beta| > 1 \text{ のとき }). \end{cases}$$

である．

$$\begin{aligned} \lim_{t\to\infty} z_t = 0 \quad &\Leftrightarrow \quad \lim_{t\to\infty}(x_t - \bar{x}) = 0 \\ &\Leftrightarrow \quad \lim_{t\to\infty} x_t = \bar{x} \end{aligned}$$

であるから，この場合にも β の絶対値が 1 より小さい時，そしてその時のみ定常点 $\bar{x}$ は安定的である．以上から，1 変数の線形動学においては，定常点の安定性は係数 β の絶対値と 1 との比較によって判断できることが分かる．

5.1.2 多変数の線形動学

次に多変数のケース，すなわち $n \times n$ 行列 A が存在し，$h(x)$ が

$$h(x) = Ax$$

によって表される場合を考えよう．ただし，$I - A$ は正則であると仮定する[21]．この時ベクトル列 $\{x_t : t = 0, 1, 2, \ldots\}$ の動きは

$$x_{t+1} = Ax_t, \quad t = 0, 1, 2, \ldots$$

によって定められる．定常点は方程式

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

の解であるが，$I - A$ が正則であることにより逆行列 $(I - A)^{-1}$ が存在するので，

$$\begin{aligned} (I - A)x = 0 \quad &\Leftrightarrow \quad (I - A)^{-1}(I - A)x = 0 \\ &\Leftrightarrow \quad x = 0 \end{aligned}$$

となり，$\bar{x} = 0$ と一意に決まる．したがって定常点 $\bar{x}$ の安定性を調べるには，$\lim_{t\to\infty} x_t = 0$，すなわち

$$\lim_{t\to\infty} A^t x_0 = 0$$

が成立するかどうかを考えればよいことになる．このような定常点での安定性を考える前に，関数の形をより一般化した

$$h(x) = Ax + a_0$$

のケースで定常点を確認しておこう．ベクトル列 $\{x_t : t = 0, 1, 2, \ldots\}$ の動きは

$$x_{t+1} = Ax_t + a_0, \quad t = 0, 1, 2, \ldots$$

によって定められ，定常点は方程式

$$x = Ax + a_0$$

の解である．$I - A$ が正則であることにより，定常点は

$$\bar{x} = (I - A)^{-1} a_0$$

[21] これは，1 変数の場合に $(1 - \beta) \neq 0$ と仮定したことに対応している．

のように導かれる．1 変数の場合と同様にして，ここで

$$z_t = x_t - \bar{x}, \qquad t = 0, 1, 2, \ldots \tag{5.2}$$

のように $\{z_t\}$ を定めると

$$\begin{aligned}
z_{t+1} &= x_{t+1} - \bar{x} \\
&= Ax_t + a_0 - (I - A)^{-1} a_0 \\
&= Ax_t + (I - A)(I - A)^{-1} a_0 - (I - A)^{-1} a_0 \\
&= Ax_t - A(I - A)^{-1} a_0 \\
&= A(x_t - \bar{x}) \\
&= Az_t
\end{aligned}$$

の関係が成立し，

$$\begin{aligned}
\lim_{t\to\infty} z_t = 0 \quad &\Leftrightarrow \quad \lim_{t\to\infty} (x_t - \bar{x}) = 0 \\
&\Leftrightarrow \quad \lim_{t\to\infty} x_t = \bar{x}
\end{aligned}$$

であるから，この場合も定常点 $\bar{x}$ の安定性を調べるには，$\lim_{t\to\infty} z_t = 0$，すなわち

$$\lim_{t\to\infty} A^t z_0 = 0$$

が成立するかどうかを考えれば十分である．

5.1 定理 (線形動学経路の安定性)． 差分方程式

$$x_{t+1} = Ax_t + a_0, \quad t = 0, 1, 2, \ldots \tag{5.3}$$

で記述される動学の定常点を $\bar{x}$ と書く．この時，以下の二つの条件は同値である．

1. (5.3) を満たす全てのベクトル列 $\{x_t : t = 0, 1, 2, \ldots\}$ について，$\lim_{t\to\infty} x_t = \bar{x}$ が成立する．
2. 特性方程式 $\det(A - \lambda I) = 0$ の解の絶対値[22]は，いずれも 1 より小さい．

5.1 定理の証明（アイディア）． 行列 A が実対称行列の場合についてのみ，証明を与える[23]．A が実対称行列であれば，その固有ベクトルより成る直行行列 H が存在し，

$$H^\top AH = \Lambda$$

が成立する．ただしここで，λ_i $(i = 1, 2, \ldots, n)$ を A の固有値として

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

[22]特性方程式の解は A が実対称行列ならば全て実数で，A の固有値に等しい．したがってその場合は，行列 A 固有値の絶対値を調べればよい．ただし一般には複素数の範囲で解を持つため，たとえば複素数 $a + bi$ の絶対値 $\sqrt{a^2 + b^2}$ を見ることになる．

[23]一般の行列に関しては，固有値を用いてジョルダン標準形に持ち込み，同様の議論をする

である．すでに確認した通り，(5.2) のようにして $\{z_t\}$ を定めると

$$z_{t+1} = Az_t \tag{5.4}$$

と書ける．H が直行行列であり，したがって $H^\top = H^{-1}$ であることに注意すれば，(5.4) の両辺に左から $H^\top$ を掛けることにより

$$\begin{aligned} H^\top z_{t+1} &= H^\top A z_t \\ &= H^\top AHH^\top z_t \\ &= \Lambda H^\top z_t \end{aligned} \tag{5.5}$$

を得る．ここで改めて

$$y_t = H^\top z_t, \qquad t = 0, 1, 2, \ldots \tag{5.6}$$

のように $\{y_t\}$ を定めれば，(5.5) は $y_{t+1} = \Lambda y_t$，すなわち

$$\begin{pmatrix} y_{t+1}^1 \\ y_{t+1}^2 \\ \vdots \\ y_{t+1}^n \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix} \begin{pmatrix} y_t^1 \\ y_t^2 \\ \vdots \\ y_t^n \end{pmatrix} = \begin{pmatrix} \lambda_1 y_t^1 \\ \lambda_2 y_t^2 \\ \vdots \\ \lambda_n y_t^n \end{pmatrix}$$

と書くことができる．これは結局，各変数 y_t^i $(i = 1, 2, \ldots, n)$ の動きが

$$y_t^i = (\lambda_i)^t y_0^i, \quad t = 0, 1, 2, \ldots$$

という単純な線形関係で特徴づけられることに他ならなず，1 変数の線形動学で見たように，λ_i の絶対値が 1 より小さい時，そしてその時のみ $\lim_{t\to\infty} y_t^i = 0$ が成立する．したがってベクトル列 $\{y_t : t = 0, 1, 2, \ldots\}$ について，

$$\begin{aligned} |\lambda_i| < 1, \quad i = 1, 2, \ldots, n \quad &\Leftrightarrow \quad \lim_{t\to\infty} y_t^i = 0, \quad i = 1, 2, \ldots, n && (5.7) \\ &\Leftrightarrow \quad \lim_{t\to\infty} y_t = 0 && (5.8) \end{aligned}$$

が言える．また H が直行行列であることに注意すれば，y_t の定め方により

$$\begin{aligned} \lim_{t\to\infty} y_t = 0 \quad &\Leftrightarrow \quad \lim_{t\to\infty} H^\top z_t = 0 \\ &\Leftrightarrow \quad H^\top \lim_{t\to\infty} z_t = 0 \\ &\Leftrightarrow \quad HH^\top \lim_{t\to\infty} z_t = 0 \\ &\Leftrightarrow \quad \lim_{t\to\infty} z_t = 0 \end{aligned} \tag{5.9}$$

さらに z_t の定義から

$$\begin{aligned} \lim_{t\to\infty} z_t = 0 \quad &\Leftrightarrow \quad \lim_{t\to\infty} (x_t - \bar{x}) = 0 \\ &\Leftrightarrow \quad \lim_{t\to\infty} x_t = \bar{x} \end{aligned} \tag{5.10}$$

以上，(5.8)(5.9)(5.10) を合わせれば

$$|\lambda_i| < 1, \quad i = 1, 2, \ldots, n \quad \Leftrightarrow \quad \lim_{t\to\infty} x_t = \bar{x}$$

が言える． □

5.1 定理の条件 1 は，(5.3) で定義される動学の大域的安定性に他ならない．つまり，線形動学の定常点が大域において安定的であるかどうかは，その動学を定義する行列の固有値の大きさを調べることで判断できるのである．

練習問題 5.1. $n=1$ として，$x_{t+1}=ax_t$ で定まる動学の安定性を調べよ．

練習問題 5.2. $n=2$ として

$$\begin{pmatrix} x_{t+1} \\ y_{t+2} \end{pmatrix} = \begin{pmatrix} ax_t \\ by_t \end{pmatrix}$$

で定まる動学系の安定性を調べよ．

5.1.3 安定多様体

絶対値が 1 を超える固有値が存在する場合についても，収束の条件を考えておこう．いま行列 A の n 個の固有値のうちで，m（$m<n$）個の絶対値は 1 より小さく，残りの $n-m$ 個については絶対値が 1 を超えていたとしよう．5.1 定理の証明と同様に $\{y_t\}$ を定義すれば，

$$\begin{pmatrix} y_{t+1}^1 \\ y_{t+1}^2 \\ \vdots \\ y_{t+1}^n \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \begin{pmatrix} y_t^1 \\ y_t^2 \\ \vdots \\ y_t^n \end{pmatrix}$$

と書ける．方程式を適当に並べ替えることで λ_i の順序は変更できるから，$\lambda_1, \lambda_2, \dots, \lambda_m$ を絶対値が 1 より小さいものとして，また $\lambda_{m+1}, \lambda_{m+2}, \dots, \lambda_n$ を絶対値が 1 より大きいものと考えて構わない．したがって最初の m 個の変数 $y_t^1, y_t^2, \dots, y_t^m$ については，任意の $y_0^1, y_0^2, \dots, y_0^m$ に対して

$$\lim_{t\to\infty} y_t^i = 0, \quad i=1,2,\dots,m$$

である．一方，残りの $n-m$ 個の変数 $y_t^{m+1}, y_t^{m+2}, \dots, y_t^n$ については，対応する初期値 $y_0^{m+1}, y_0^{m+2}, \dots, y_0^n$ が全てゼロであれば

$$\lim_{t\to\infty} y_t^i = 0, \quad i=m+1, m+2, \dots, n$$

が成立する．よって

$$y_0 = \begin{pmatrix} y_0^1 \\ \vdots \\ y_0^m \\ y_0^{m+1} \\ \vdots \\ y_0^n \end{pmatrix}$$

の $y_0^{m+1}, y_0^{m+2}, \dots, y_0^n$ が全てゼロとなるように初期値 x_0 をとれば，

$$\lim_{t\to\infty} y_t = 0 \quad \text{したがって} \quad \lim_{t\to\infty} x_t = \bar{x}$$

が成立し，動学経路は定常点に収束することが分かる．最初の m 個の項目については任意に選ぶことができるから，定常点に収束するような初期値の集合は m 次元空間になる．初期値がそのような m 次元空間に含まれる場合には定常点に収束し，それ以外の点から出発する場合については定常点への収束は保証されない．

上記の条件を満たす初期値 x_0 の集合を安定多様体（stable manifold）と言うが，この安定多様体にもう少しフォーマルな定義を与えておこう．いま（線形）関数 $\psi : \boldsymbol{R}^n \to \boldsymbol{R}^{n-m}$ を

$$\psi(x) := \begin{pmatrix} 0 & 0 \\ 0 & I_{n-m} \end{pmatrix} H^\top (x - \bar{x})$$

のように定義する．ここで $y_t = H^\top(x_t - \bar{x})$ であったことに注意すれば，

$$\begin{aligned} \psi(x_t) &= \begin{pmatrix} 0 & 0 \\ 0 & I_{n-m} \end{pmatrix} H^\top (x_t - \bar{x}) \\ &= \begin{pmatrix} 0 & 0 \\ 0 & I_{n-m} \end{pmatrix} y_t \\ &= \begin{pmatrix} y_t^{m+1} \\ \vdots \\ y_t^n \end{pmatrix} \end{aligned}$$

と書けるから，

$$\psi(x_0) = 0 \quad \Leftrightarrow \quad \begin{pmatrix} y_0^{m+1} \\ \vdots \\ y_0^n \end{pmatrix} = 0$$

が成立する．つまり初期値 x_0 が $\psi(x_0) = 0$ を満せば，その動学経路は定常点に収束するのである．したがって，いま問題としている差分方程式で定まる任意のベクトル列 $\{x_t : t = 0, 1, 2, \ldots\}$ について

$$x_0 \in \left\{ x \in \boldsymbol{R}^n \;\middle|\; \psi(x) = 0 \right\} \quad \Rightarrow \quad \lim_{t \to \infty} x_t = \bar{x}$$

が成立し，この集合 $\{x \in \boldsymbol{R}^n \mid \psi(x) = 0\}$ が安定多様体である．

練習問題 5.3. $n = 2$ として

$$\begin{pmatrix} x_{t+1} \\ y_{t+2} \end{pmatrix} = \begin{pmatrix} 2x_t - 1 \\ \frac{1}{2} y_t \end{pmatrix}$$

で定まる動学系の定常点を調べ，それに対応する安定多様体を求めよ．

5.2 非線形動学

次に非線形動学の話に入ろう．動学を定める関数 h が線形でないときには，定常点の周辺で h を線形近似し，線形の場合と同様の議論を当てはめることによってその安定性を判断することになる．したがって非線形動学での議論は，基本的に線形動学で確認した結果の応用である．

まずは，線形近似を行うためのヤコビ行列を定義しておく．動学を表す関数 $h : \boldsymbol{R}^n \to \boldsymbol{R}^n$ は

$$\forall x \in \boldsymbol{R}^n, \quad h(x) = \big(h_1(x),\ h_2(x),\ \ldots,\ h_n(x)\big)$$

で定義されていたが，この関数 h の点 $x \in \boldsymbol{R}^n$ におけるヤコビ行列を，$Dh(x)$ で表すことにする．すなわち，

$$Dh(x) = \begin{pmatrix} \frac{\partial h_1(x)}{\partial x_1} & \frac{\partial h_1(x)}{\partial x_2} & \cdots & \frac{\partial h_1(x)}{\partial x_n} \\ \frac{\partial h_2(x)}{\partial x_1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \frac{\partial h_n(x)}{\partial x_1} & \cdots & & \frac{\partial h_n(x)}{\partial x_n} \end{pmatrix}$$

である．

5.2 定理 (非線形動学の局所安定性). $\bar{x}$ を定常点として，$I - Dh(x)$ は正則行列であるとする．この時，以下の 3 点が成立する．

1. 特性方程式 $\det(Dh(\bar{x}) - \lambda I) = 0$ の解の絶対値がいずれも 1 より厳密に小さいならば，$\bar{x}$ は局所安定的である．
2. 特性方程式の解の中に絶対値が 1 よりも厳密に大きいものがあれば，$\bar{x}$ は局所安定的でない．
3. 一般に，特性方程式の解で絶対値が 1 よりも厳密に小さいものの数を m とすれば，$\bar{x}$ の近傍で定義される連続微分可能な関数 $\psi : \boldsymbol{R}^n \to \boldsymbol{R}^{n-m}$ が存在し，(5.1) で定まるベクトル列のうちで初期値が $\psi(x_0) = 0$ を満たすものについて，$\lim_{t\to\infty} x_t = \bar{x}$ が成立する．

5.2 定理における関数 ψ が，線形動学で定義した ψ と同趣旨のものであることは明らかであろう．したがって，$\psi(x) = 0$ で定まる $\boldsymbol{R}^n$ の「m 次元」部分集合 $\{x \in \boldsymbol{R}^n \mid \psi(x) = 0\}$ は安定多様体と呼ばれる．安定多様体は，線形の場合にはある線形部分空間を平行移動したものであるが，一般には滑らかな境界面を持つ図形になる．

5.1 注意. 5.2 定理において，絶対値が 1 に等しい固有値が存在する場合については何も言及されていない．というのも，そのような場合には局所安定的であることもあればそうでないこともあり，定常点の安定性をにわかには判定できないからである．例えば $n = 1$ として，$h(x) = x$ と $h(x) = -x^3 + x$ という二つのケースを考えてみよう．いずれの動学においても $x = 0$ が定常点であることは明らかで，さらに $h'(0) = 1$ であるから特性方程式 $h'(0) - \lambda = 0$ の解は 1 に等しい．しかしながら，前者のケースでは定常点は局所安定的でないが，後者では局所安定的となる．

練習問題 5.4. $n = 1$ のケースを考える．関数 h が $h(x) = x(x-1)(x-2) + x$ で与えられる時，$x_{t+1} = h(x_t)$ で定まる動学系の定常点を求め，その安定性を調べよ．

6 積分法

6.1 不定積分

6.1.1 原始関数と不定積分

6.1 定義 (**原始関数**). I を $\boldsymbol{R}$ に含まれる区間であるとする．関数 $f : I \to \boldsymbol{R}$ に対して I 上で微分可能な関数 F が存在し，

$$\forall x \in I, \quad F'(x) = f(x)$$

が成立する時，F は区間 I における f の原始関数であるという．

もっとも，全ての関数 f に対して原始関数が存在するとは限らない．

6.1 定理. F は $f : I \to \boldsymbol{R}$ の原始関数の一つであるとする．この時，以下の 2 点が成立する．

1. 任意の定数 C について，
$$\hat{F}(x) = F(x) + C$$
で定義される関数 $\hat{F}$ は，f の原始関数である．

2. 関数 G が f の原始関数であれば，必ずある定数 C が存在して
$$G(x) = F(x) + C$$
と書ける．

6.1 定理の証明. 関数 F は f の原始関数の一つであるから，定義により

$$\forall x \in I, \quad F'(x) = f(x)$$

である．したがって

$$\forall x \in I, \quad \hat{F}'(x) = F'(x) = f(x)$$

が成立し，$\hat{F}$ が f の原始関数となるから 1 点目の結果を得る．

次に 2 点目について

$$\forall x \in I, \quad G'(x) = f(x)$$

であるから，

$$h(x) := G(x) - F(x)$$

と h を定義すれば，

$$\begin{aligned} \forall x \in I, \quad h'(x) &= G'(x) - F'(x) \\ &= f(x) - f(x) \\ &= 0 \end{aligned} \tag{6.1}$$

が成立する．一方，平均値の定理により

$$\forall a, b \in I \ : \quad \frac{h(b) - h(a)}{b - a} = h'(c)$$

を満たすような c が，a と b で区切られた区間に存在する．(6.1) により右辺はゼロであるから，結局

$$\forall a, b \in I, \quad h(a) = h(b)$$

となり，これは関数 h が区間 I で一定の値をとることに他ならない．したがってその値を C と置けば

$$\begin{aligned} \forall x \in I, \quad C &= h(x) \\ &= G(x) - F(x) \end{aligned}$$

すなわち

$$\forall x \in I, \quad G(x) = F(x) + C$$

を得る． □

6.1 定理の一点目によれば，F に任意の定数を加えたものは必ず原始関数となる．したがって関数 f に対して原始関数が存在する場合には，それは一意に定まらない．一方の二点目は，任意の原始関数は F に適当な定数を加えた形で書けること，逆に言えば $F(x) + C$ によって表現できるもの以外に原始関数は存在しないことをを示している．つまり関数 f の原始関数は無数に存在するが，それは定数部分を除いて一意的に定まるのである．

6.2 定義 (不定積分). 区間 $I \subset \boldsymbol{R}$ で定義された関数 f について，その原始関数の総称を不定積分といい $\int f(x)dx$ で表す[24]．すなわち，F を f の原始関数の一つとすれば，任意の定数 C に対して

$$\int f(x)dx = F(x) + C \tag{6.2}$$

である．また (6.2) のような形で f の不定積分を求めることを f を積分するといい，その時 f は被積分関数，C は積分定数と呼ばれる．

6.1.2 不定積分の性質

不定積分の基本的な性質について，次の定理を確認しておく．

6.2 定理 (不定積分の線形性). f と g はともに I で定義される関数で，いずれも原始関数を持つものとする．この時 α, β を任意の定数として

$$\int \big\{\alpha f(x) + \beta g(x)\big\}dx = \alpha \int f(x)dx + \beta \int g(x)dx$$

が成立する．

6.2 定理の証明. F と G をそれぞれ f と g の原始関数の一つとすれば，不定積分の定義により c_1, c_2 を任意の定数として

$$\int f(x)dx = F(x) + c_1$$

$$\int g(x)dx = G(x) + c_2$$

と書けるから，

$$\int f(x)dx + \int g(x)dx = F(x) + G(x) + c_1 + c_2 \tag{6.3}$$

である．一方，

$$\begin{aligned} \frac{d}{dx}\big\{F(x) + G(x)\big\} &= F'(x) + G'(x) \\ &= f(x) + g(x) \end{aligned}$$

[24] 不定積分に対しては定積分を用いた定義が与えられることもある．

であるから，$F(x)+G(x)$ は $f(x)+g(x)$ の原始関数の一つであり，c_3 を任意の定数として

$$\int \bigl\{f(x)+g(x)\bigr\}dx = F(x)+G(x)+c_3 \tag{6.4}$$

と書ける．c_1，c_2 及び c_3 は任意の定数であるから，$c_1+c_2=c_3$ と置いて構わない．すると (6.3) と (6.4) により

$$\int \bigl\{f(x)+g(x)\bigr\}dx = \int f(x)dx + \int g(x)dx \tag{6.5}$$

が言える．一方，任意の定数 α（$\alpha \neq 0$）に対して，

$$\begin{aligned}\frac{d}{dx}\bigl\{\alpha F(x)\bigr\} &= \alpha F'(x)\\ &= \alpha f(x)\end{aligned}$$

が成立するから，$\alpha F(x)$ は $\alpha f(x)$ の原始関数の一つである．したがって c_4 を定数として

$$\begin{aligned}\int \alpha f(x)dx &= \alpha F(x)+c_4\\ &= \alpha\Bigl\{F(x)+\frac{c_4}{\alpha}\Bigr\}\\ &= \alpha\int f(x)dx\end{aligned}$$

と書ける．同様の議論から

$$\int \beta g(x)dx = \beta\int g(x)dx$$

も言えるから，(6.5) と合わせれば

$$\begin{aligned}\int \bigl\{\alpha f(x)+\beta g(x)\bigr\}dx &= \int \alpha f(x)dx + \int \beta g(x)dx\\ &= \alpha\int f(x)dx + \beta\int g(x)dx\end{aligned}$$

が成立する． □

6.3 定理 (部分積分). f と g はともに I で定義される関数で，いずれも原始関数を持つものとする．この時 f と g が微分可能であれば

$$\int f'(x)g(x)dx = f(x)g(x) - \int f(x)g'(x)dx$$

が成立する．

6.3 定理の証明. 関数の積の微分の公式により

$$\frac{d}{dx}\bigl\{f(x)g(x)\bigr\} = f'(x)g(x)+f(x)g'(x)$$

であるから，$f(x)g(x)$ は $f'(x)g(x)+f(x)g'(x)$ の原始関数である．したがって c を任意の定数として

$$\int \bigl\{f'(x)g(x)+f(x)g'(x)\bigr\}dx = f(x)g(x)+c$$

と書け，6.2 定理を用いれば

$$\int f'(x)g(x)dx + \int f(x)g'(x)dx = f(x)g(x)+c$$

である．c は任意なので 0 と置けば定理の結果を得る． □

6.4 定理 (**置換積分**). f は I で定義される関数で原始関数を持つものとし，変数 x と変数 t とが

$$x = \varphi(t)$$

によって関係付けられているとする．この時，関数 φ が微分可能であれば

$$\int f'(x)dx = \int f(\varphi(t))\varphi'(t)dt$$

が成立する．

6.4 定理の証明. 関数 f の原始関数を F とすれば，c_1 を任意の定数として

$$\begin{aligned} \int f(x)dx &= F(x) + c_1 \\ &= F\bigl(\varphi(t)\bigr) + c_1 \end{aligned} \tag{6.6}$$

と書ける．一方，合成関数の微分の公式により

$$\begin{aligned} \frac{d}{dt}F\bigl(\varphi(t)\bigr) &= F'\bigl(\varphi(t)\bigr)\varphi'(t) \\ &= f\bigl(\varphi(t)\bigr)\varphi'(t) \end{aligned}$$

であるから，$F\bigl(\varphi(t)\bigr)$ は $f\bigl(\varphi(t)\bigr)\varphi'(t)$ の原始関数．したがって c_2 を任意の定数として

$$\int f\bigl(\varphi(t)\bigr)\varphi'(t)dt = F\bigl(\varphi(t)\bigr) + c_2 \tag{6.7}$$

と表せる．任意の定数 c_1 と c_2 とを等しく定めれば，(6.6) と (6.7) により定理の結果を得る． □

6.2 定積分（リーマン積分）

6.2.1 定積分

$\boldsymbol{R}$ 上の区間 $[a.b]$ を考える．$[a,b]$ に含まれる $n+1$ 個の点 $x_0, x_1, x_2, \ldots, x_n$ が，

$$a = x_0 < x_1 < \cdots < x_n = b$$

を満たす時，集合 $\{x_0, x_1, \ldots, x_n\}$ を区間 $[a,b]$ の分割といい

$$\Delta = \{x_0, x_1, \ldots, x_n\}$$

で表す．また，区間 $[a,b]$ の全ての分割の集合を $\mathcal{D}$ と表す．すなわち

$$\mathcal{D} := \Bigl\{ \{x_0, x_1, \ldots, x_n\} \Bigm| a = x_0 < x_1 < \cdots < x_n = b,\ n \in \boldsymbol{N} \Bigr\}$$

である．分割 $\Delta \in \mathcal{D}$ が一つ与えられると，区間 $[a,b]$ を n 個の部分区間 $[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n]$ に分けることができ，それぞれの部分区間から適当に $\xi_i \in [x_{i-1}, x_i]$ を選べば

$$S(\Delta, \xi) = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1})$$

のような和を考えられる．これはリーマン和と呼ばれるもので，$f(x)$ の下方の面積を近似したものと見なせる．

リーマン和の値は分割 Δ と各 $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ の取り方に応じて定まり，分割の取り方を細かくすればするほど，$f(x)$ 下の面積の値に近づいてゆくことが予想されよう．分割の取り方を細かくする作業は，区間 $[a,b]$ の分割 Δ に対して n 個の部分区間の長さの最大値を

$$\delta(\Delta) := \max_{1 \le i \le n} (x_i - x_{i-1})$$

と置き，この $\delta(\Delta)$ の値をゼロに近づけてゆくことに相当する．したがって，$\delta(\Delta) \to 0$ の極限をとった時に $S(\Delta, \xi)$ がある値に収束するのであれば，その収束先を $f(x)$ の下方の面積と見なすことができる．$f(x)$ の区間 $[a,b]$ における定積分（リーマン積分）とは，このような区間分割と極限操作によって面積を求めることに他ならない．

6.3 定義 (定積分). f を区間 $[a,b]$ で定義される関数とする．また Δ を $[a,b]$ の分割とし，Δ が定める部分区間の最大幅を $\delta(\Delta)$ と置く．この時ある $s \in \boldsymbol{R}$ が存在して

$$\forall \epsilon > 0,\ \exists \delta(\epsilon): \qquad \delta(\Delta) < \delta(\epsilon) \quad \Rightarrow \quad |\, S(\Delta, \xi) - s \,| < \epsilon$$

が成立するならば，f は区間 $[a,b]$ で定積分可能であるといい

$$s = \lim_{\delta(\Delta) \to 0} S(\Delta, \xi)$$

と書く．また，極限値 s を区間 $[a,b]$ における $f(x)$ の定積分といい，記号

$$s = \int_a^b f(x)dx$$

で表す．また，積分の上限 a と下限 b とを逆転させた記号を

$$\int_b^a f(x)dx := -\int_a^b f(x)dx$$

で定義する．

積分可能性に関する重要なポイントとしては,「連続関数であれば定積分可能」というものがある．ただし一般にその逆は成立せず，不連続点をもっていても定積分は可能でありうる．連続関数の積分可能性については，以下の定理によって保証される．

6.5 定理. 関数 f が閉区間 $[a,b]$ で連続であれば，f は $[a,b]$ でリーマン積分可能である．

6.5 定理に証明を与えるためには準備が必要である．まずは，連続関数の基本的な性質を確認しておこう．

6.6 定理 (最大・最小値の定理). $\boldsymbol{R}$ 上の閉区間 $[a,b]$ で連続な関数 f は，$[a,b]$ 上で最大値と最小値を持つ．

6.5 定理の f は $[a,b]$ で連続であるから，分割 Δ に対応する n 個の部分区間でも f は連続である．したがって 6.6 定理により，f は各部分区間 $[x_{i-1}, x_i]$（$i = 1, 2, \ldots, n$）で最大値 M_i と最小値 m_i を持つ．部分区間 $[x_{i-1}, x_i]$ において $f(x)$ の最大値を与える x の値を $\bar{x}_i$，最小値を与える x の値を $\underline{x}_i$ と置けば

$$f(\bar{x}_i) = M_i, \quad f(\underline{x}_i) = m_i$$

である．また，M_i および m_i を用いて

$$S(\Delta) := \sum_{i=1}^n M_i (x_i - x_{i-1})$$

$$s(\Delta) := \sum_{i=1}^{n} m_i(x_i - x_{i-1})$$

のような 2 種類の和を分割ごとに定義することができる．このようにして定義される $S(\Delta)$ と $s(\Delta)$ に関して，次の定理が成り立つ．

6.7 定理. ある分割 Δ に対して，新たな分点をいくつか加えたものを Δ' と置く．すなわち，$\Delta \subset \Delta'$ である．この時

$$s(\Delta) \leq s(\Delta') \quad \text{かつ} \quad S(\Delta') \leq S(\Delta)$$

が成立する．

この定理を踏まえれば，以下の定理が導かれる．

6.8 定理. 任意の分割 $\Delta, \Delta' \in \mathcal{D}$ について

$$s(\Delta) \leq S(\Delta')$$

が成立する．

この 6.8 定理において分割 Δ は任意であるから，$s(\Delta)$ の全体の集合 $\{s(\Delta)|\Delta \in \mathcal{D}\}$ は上に有界であることが分かる．したがって $\{s(\Delta)|\Delta \in \mathcal{D}\}$ の上限 s が存在し

$$\forall \Delta \in \mathcal{D}, \quad s(\Delta) \leq s$$

である．また 6.8 定理により，$S(\Delta)$ が $\{s(\Delta)|\Delta \in \mathcal{D}\}$ の上界に属することも分かるから

$$\forall \Delta \in \mathcal{D}, \quad s(\Delta) \leq s \leq S(\Delta)$$

が言える．

最後に，一様連続の概念を導入しておく．

6.4 定義 (一様連続). $\boldsymbol{R}$ の区間 I で定義される関数 f について

$$\forall \epsilon,\ \exists \delta(\epsilon),\ \forall x, y \in I: \quad |x - y| < \delta(\epsilon) \quad \Rightarrow \quad |f(x) - f(y)| < \epsilon$$

が成立する時，f は区間 I で一様連続であるという．

連続性の定義と比較すれば明らかなように，連続よりも一様連続の方が f に対する要請が強い．したがって，f が区間 I で一様連続ならば f は区間 I で連続であるが，その逆は一般に成立しない．ただし，一定の条件の下で両概念は一致することが知られており，次の定理はその条件を与えるものである．

6.9 定理. 連続関数 f の定義域 I がコンパクトであれば，f は I で一様連続である．

閉区間 $[a, b]$ はコンパクトであるから，この定理により，6.5 定理における関数 f は $[a, b]$ で一様連続である．

それでは，以上で確認した記号や諸定理を用いて，6.5 定理に証明を与えよう．

6.5 定理の証明. M_i と m_i の定め方により，

$$m_i \leq f(\xi_i) \leq M_i, \qquad i = 1, 2, \ldots, n$$

したがって

$$m_i(x_i - x_{i-1}) \leq f(\xi_i)(x_i - x_{i-1}) \leq M_i(x_i - x_{i-1}), \qquad i = 1, 2, \ldots, n$$

が言えるから，

$$\sum_{i=1}^{n} m_i(x_i - x_{i-1}) \leq \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) \leq \sum_{i=1}^{n} M_i(x_i - x_{i-1}) \quad i = 1, 2, \ldots, n$$

である．よって任意の Δ と ξ に対して

$$s(\Delta) \leq S(\Delta, \xi) \leq S(\Delta) \tag{6.8}$$

が成立する．また 6.8 定理により，s を $\{s(\Delta) | \Delta \in \mathcal{D}\}$ の上限として

$$\forall \Delta \in \mathcal{D}, \quad s(\Delta) \leq s \leq S(\Delta)$$

が成立するから，(6.8) と合わせれば任意の Δ と ξ に対して

$$\forall \Delta \in \mathcal{D}, \quad S(\Delta) - s(\Delta) > |S(\Delta, \xi) - s| \tag{6.9}$$

が言える．

一方 f は $[a, b]$ で一様連続であるから，任意の正の数 ϵ を所与として

$$\forall x, x' \in [a, b] : \qquad |x - x'| < \delta(\epsilon) \quad \Rightarrow \quad |f(x) - f(x')| < \epsilon \tag{6.10}$$

となるような正の数 $\delta(\epsilon)$ が一つ定まる．ここで $\delta(\Delta) < \delta(\epsilon)$，すなわち

$$|x_i - x_{i-1}| < \delta(\epsilon), \quad i = 1, 2, \ldots, n$$

となるように分割 Δ を選べば，$\bar{x}_i$ と $\underline{x}_i$ は区間 $[x_{i-1}, x_i]$ に含まれるから

$$|\bar{x}_i - \underline{x}_i| < \delta(\epsilon), \quad i = 1, 2, \ldots, n$$

であり，(6.10) と合わせれば

$$|f(\bar{x}_i) - f(\underline{x}_i)| < \epsilon \qquad \text{すなわち} \qquad M_i - m_i < \epsilon$$

よって

$$\begin{aligned} s(\Delta) - S(\Delta) &= \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) \\ &< \epsilon \sum_{i=1}^{n} (x_i - x_{i-1}) \\ &= \epsilon(b - a) \end{aligned}$$

が成り立つ．(6.9) と合わせれば，任意の $\epsilon > 0$ に対して $\delta(\epsilon) > 0$ が存在し，$\delta(\Delta) < \delta(\epsilon)$ を満たす任意の分割 Δ について

$$\begin{aligned} |S(\Delta, \xi) - s\,| &< S(\Delta) - s(\Delta) \\ &< \epsilon(b - a) \end{aligned} \tag{6.11}$$

が成立する．$(b - a) > 0$ でありかつ ϵ は任意に選べるから，(6.11) の右辺は任意の正の数である．したがって定義により，f は区間 $[a, b]$ において定積分可能である． □

6.2.2 定積分の性質

定積分の基本的な性質について，次の定理を確認しておく．

6.10 定理 (定積分の線形性). f と g は $[a,b]$ で連続な関数であるとする．この時，α, β を任意の定数として

$$\int_a^b \big\{\alpha f(x) + \beta g(x)\big\} dx = \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx$$

が成立する．

6.10 定理の証明. $[a,b]$ の任意の分割 $\Delta \in \mathcal{D}$ に対して，関数 $\alpha f(x) + \beta g(x)$ のリーマン和を考えると，

$$\begin{aligned} S(\Delta,\xi) &= \sum_{i=1}^n \big\{\alpha f(\xi_i) + \beta g(\xi_i)\big\}(x_i - x_{i-1}) \\ &= \alpha \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) + \beta \sum_{i=1}^n g(\xi_i)(x_i - x_{i-1}) \end{aligned}$$

と書ける．いま f と g は $[a,b]$ において定積分可能であるから

$$\lim_{\delta(\Delta)\to 0} \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) = \int_a^b f(x)dx$$

$$\lim_{\delta(\Delta)\to 0} \sum_{i=1}^n g(\xi_i)(x_i - x_{i-1}) = \int_a^b g(x)dx$$

したがって，

$$\begin{aligned} \lim_{\delta(\Delta)\to 0} S(\Delta,\xi) &= \alpha \lim_{\delta(\Delta)\to 0} \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) + \beta \lim_{\delta(\Delta)\to 0} \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \\ &= \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx \end{aligned}$$

が成立し，定理の結果を得る． □

6.11 定理 (定積分の加法性). f は $[a,b]$ で連続な関数であるとする．この時，$a < c < b$ を満たす任意の c について

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

が成立する[25]．

6.11 定理の証明. $[a,b]$ の分割 $\Delta = \{x_0, x_1, \dots, x_n\}$ のうちで

$$a = x_0 < \cdots < x_k = c < x_{k+1} < \cdots < x_n = b,$$

のように c を一つの分点とするものを考えれば，このような分割のリーマン和は

$$\sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) = \sum_{i=1}^k f(\xi_i)(x_i - x_{i-1}) + \sum_{i=k+1}^n f(\xi_i)(x_i - x_{i-1})$$

と書ける．f は区間 $[a,b]$ で連続であるから部分区間 $[a,c]$ と $[c,b]$ でも連続で，いずれの区間でも定積分可能となる．したがって，c を分点に含みながら極限 $\delta(\Delta) \to 0$ をとれば，左辺は

$$\lim_{\delta(\Delta)\to 0} \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) = \int_a^b f(x)dx$$

[25] これは，f が $[a,c]$ および $[c,b]$ で積分可能であれば $[a,b]$ で積分可能であることを意味する．

となり，右辺については

$$\lim_{\delta(\Delta)\to 0}\sum_{i=1}^{k} f(\xi_i)(x_i - x_{i-1}) + \lim_{\delta(\Delta)\to 0}\sum_{i=k+1}^{n} f(\xi_i)(x_i - x_{i-1}) = \int_a^c f(x)dx + \int_c^b f(x)dx$$

が成り立つから，定理の結果を得る． □

6.12 定理 (定積分の単調性). f と g は $[a,b]$ で連続な関数であるとする．区間 $[a,b]$ において

$$\forall x \in [a,b], \quad f(x) \geq g(x)$$

であるならば,

$$\int_a^b f(x)dx \geq \int_a^b g(x)dx$$

が成立する．またこの時,

$$\int_a^b f(x)dx = \int_a^b g(x)dx \quad \Leftrightarrow \quad \forall x \in [a,b],\ f(x) = g(x)$$

の関係が成り立つ．

6.12 定理の証明. 区間 $[a,b]$ で常に $f(x) \geq g(x)$ が成り立つから，$[a,b]$ の分割 $\Delta = \{x_0, x_1, \ldots, x_n\}$ と $\xi_i \in [x_{i-1}, x_i]$ をどのように選んでも

$$f(\xi_i)(x_i - x_{i-1}) \geq g(\xi_i)(x_i - x_{i-1}), \quad i = 1, 2, \ldots, n$$

である．この時，任意の分割 Δ について

$$\sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) \geq \sum_{i=1}^{n} g(\xi_i)(x_i - x_{i-1})$$

であるから，極限 $\delta(\Delta) \to 0$ をとっても不等号は維持され，結局

$$\int_a^b f(x)dx \geq \int_a^b g(x)dx$$

が言える．さらに，この時

$$\int_a^b f(x)dx = \int_a^b g(x)dx \tag{6.12}$$

であるならば

$$\forall x \in [a,b],\ f(x) = g(x)$$

でなければならないことを示す．背理法の仮定として，$f(c) > g(c)$ なる $c \in [a,b]$ が存在したとしよう．すると，f と g の連続性により c の近傍では常に $f(x) > g(x)$ となるから

$$\forall x \in [\alpha, \beta], \quad f(x) - \frac{1}{2}f(c) > g(x) - \frac{1}{2}g(c)$$

が成り立つように，c を含む十分小さな部分区間 $[\alpha, \beta] \subset [a,b]$ をとることができる．この部分区間 $[\alpha, \beta]$ の任意の分割 $\Delta = \{x_0, x_1, \ldots, x_n\}$ と，任意の $\xi_i \in [x_{i-1}, x_i]$ について,

$$f(\xi_i)(x_i - x_{i-1}) - \frac{1}{2}f(c)(x_i - x_{i-1}) > g(\xi_i)(x_i - x_{i-1}) - \frac{1}{2}g(c)(x_i - x_{i-1}), \quad i = 1, 2, \ldots, n$$

となるから，両辺の和をとれば

$$\sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) - \frac{1}{2}f(c)\sum_{i=1}^{n}(x_i - x_{i-1}) > \sum_{i=1}^{n} g(\xi_i)(x_i - x_{i-1}) - \frac{1}{2}g(c)\sum_{i=1}^{n}(x_i - x_{i-1})$$

すなわち

$$\sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) - \frac{1}{2}f(c)(\beta - \alpha) > \sum_{i=1}^{n} g(\xi_i)(x_i - x_{i-1}) - \frac{1}{2}g(c)(\beta - \alpha)$$

である．ここで両辺について $\delta(\Delta) \to 0$ の極限をとると

$$\int_{\alpha}^{\beta} f(x)dx - \frac{1}{2}f(c)(\beta - \alpha) \geq \int_{\alpha}^{\beta} g(x)dx - \frac{1}{2}g(c)(\beta - \alpha)$$

したがって

$$\begin{aligned} \int_{\alpha}^{\beta} f(x)dx - \int_{\alpha}^{\beta} g(x)dx &\geq \frac{1}{2}f(c)(\beta - \alpha) - \frac{1}{2}g(c)(\beta - \alpha) \\ &> 0 \end{aligned} \tag{6.13}$$

となる．ただし，最後の不等式は $f(c) > g(c)$ による．一方，積分の加法性により

$$\int_{a}^{b} f(x)dx = \int_{a}^{\alpha} f(x)dx + \int_{\alpha}^{\beta} f(x)dx + \int_{\beta}^{b} f(x)dx$$

$$\int_{a}^{b} g(x)dx = \int_{a}^{\alpha} g(x)dx + \int_{\alpha}^{\beta} g(x)dx + \int_{\beta}^{b} g(x)dx$$

であり，また $[a, \alpha]$ および $[\beta, b]$ において $f(x) \geq g(x)$ であることにより

$$\int_{a}^{\alpha} f(x)dx \geq \int_{a}^{\alpha} g(x)dx, \quad \int_{\beta}^{b} f(x)dx \geq \int_{\beta}^{b} g(x)dx$$

であるから，

$$\int_{a}^{b} f(x)dx - \int_{a}^{b} g(x)dx \geq \int_{\alpha}^{\beta} f(x)dx - \int_{\alpha}^{\beta} g(x)dx$$

が成り立つ．しかし，これを (6.13) と合わせれば

$$\int_{a}^{b} f(x)dx - \int_{a}^{b} g(x)dx > 0$$

となり，(6.12) に矛盾する．よって

$$\forall x \in [a, b], \ f(x) \geq g(x), \ \text{かつ} \ \int_{a}^{b} f(x)dx = \int_{a}^{b} g(x)dx, \quad \Rightarrow \quad \forall x \in [a, b], \ f(x) = g(x)$$

でなければならない．この逆が成立することは明らかであるから，以上で定理の結果を得る． □

練習問題 6.1. f は区間 $[a, b]$ で連続であるとする．上記の性質を用いて，

$$\int_{a}^{b} f(x)dx \leq \int_{a}^{b} |f(x)|dx$$

が成り立つこと，及び

$$\forall x \in [a, b], \ f(x) \geq 0, \ \text{かつ} \ \int_{a}^{b} f(x)dx = 0, \quad \Rightarrow \quad \forall x \in [a, b], \ f(x) = 0$$

であることを示せ．

6.2.3 微積分学の基本定理

以上の議論を踏まえて，この節では微積分学の基本定理について述べる．まずはその準備として，平均値の定理を確認しておこう．

6.13 定理 (積分の平均値の定理). 関数 f が区間 $[a,b]$ で連続であるならば

$$\frac{1}{b-a}\int_a^b f(x)dx = f(c)$$

となるような $c \in (a,b)$ が存在する．

6.13 定理の証明. f は区間 $[a,b]$ で連続であるから，6.6 定理により $[a,b]$ において最大値 M と最小値 m を持つ．f の最大値を与えるものを $\bar{x} \in [a,b]$，最小値を与えるものを $\underline{x} \in [a,b]$ と置けば，

$$f(\bar{x}) = M, \quad f(\underline{x}) = m$$

である．もし $M = m$ ならば，f は $[a,b]$ 上で一定のはずであるからその値を k と置くと

$$\begin{aligned} S(\Delta,\xi) &= \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \\ &= k\sum_{i=1}^n (x_i - x_{i-1}) \\ &= k(b-a) \end{aligned}$$

が $[a,b]$ の任意の分割 Δ で成立する．ここで，両辺について $\delta(\Delta) \to 0$ の極限をとれば

$$\int_a^b f(x)dx = k(b-a)$$

である．したがって，任意の $c \in [a,b]$ について

$$\begin{aligned} \frac{1}{b-a}\int_a^b f(x)dx &= \frac{1}{b-a}k(b-a) \\ &= k \\ &= f(c) \end{aligned}$$

が成立する．一方，$M > m$ ならば

$$\forall x \in [a,b], \quad m \leq f(x) \leq M$$

に加えて

$$\exists x \in [a,b], \quad m < f(x) < M$$

である．したがって積分の単調性により

$$\int_a^b m dx < \int_a^b f(x)dx < \int_a^b M dx$$

すなわち

$$m(b-a) < \int_a^b f(x)dx < M(b-a)$$

が成立し，これは

$$f(\underline{x}) = m < \frac{1}{b-a}\int_a^b f(x)dx < M = f(\bar{x})$$

と書き直せる．ここで一般性を失うことなく $\underline{x} < \bar{x}$ と仮定すれば，f が区間 $[\underline{x}, \bar{x}] \subset [a, b]$ で連続であることから，中間値の定理により

$$\frac{1}{b-a}\int_a^b f(x)dx = f(c)$$

を満たす $c \in (\underline{x}, \bar{x})$ が存在し，この c が定理の条件を満たす． □

6.14 定理 (微積分学の基本定理). f を $[a, b]$ で連続な関数とする．任意の $x \in [a, b]$ に対して部分区間 $[a, x]$ を考えると，f は $[a, x]$ で積分可能であり，その積分の値は x の取り方に応じて定まる．そこで任意の $x \in [a, b]$ に対して

$$F(x) := \int_a^x f(t)dt$$

と F を定義すれば，F は $[a, b]$ で微分可能であり，

$$\forall x \in [a, b], \quad F'(x) = f(x)$$

が成立する．

6.14 定理の証明. 任意の $x \in [a, b]$ に対して，$x + h$ が $[a, b]$ に含まれるように $h \neq 0$ をとる．$h > 0$ の時，f は部分区間 $[x, x + h] \subset [a, b]$ で連続であるから，平均値の定理により

$$\frac{1}{h}\int_x^{x+h} f(t)dt = f(c) \tag{6.14}$$

なる $c \in (x, x + h)$ が存在する．一方，積分の加法性により

$$\int_a^{x+h} f(t)dt = \int_a^x f(t)dt + \int_x^{x+h} f(t)dt$$

であるから，F の定義により

$$F(x + h) - F(x) = \int_x^{x+h} f(t)dt \tag{6.15}$$

と書ける．よって (6.14) と (6.15) を合わせれば

$$\frac{F(x + h) - F(x)}{h} = f(c), \qquad x < c < x + h$$

となる．ここで両辺について $h \to +0$ の極限をとると，f の連続性により

$$\begin{aligned} \lim_{h\to+0}\frac{F(x + h) - F(x)}{h} &= \lim_{h\to+0} f(c) \\ &= f(\lim_{h\to+0} c) \\ &= f(x) \end{aligned} \tag{6.16}$$

を得る．$h < 0$ の時も，f は部分区間 $[x + h, x] \subset [a, b]$ で連続であるから，平均値の定理により

$$-\frac{1}{h}\int_{x+h}^x f(t)dt = f(c) \tag{6.17}$$

なる $c \in (x + h, x)$ が存在する．先ほどと同様に積分の加法性を用いれば，

$$\int_a^x f(t)dt = \int_a^{x+h} f(t)dt + \int_{x+h}^x f(t)dt$$

であるから，F の定義により

$$F(x+h) - F(x) = -\int_{x+h}^{x} f(t)dt \tag{6.18}$$

と書ける．よって (6.17) と (6.18) を合わせれば

$$\frac{F(x+h) - F(x)}{h} = f(c), \qquad x+h < c < x$$

となる．両辺について $h \to -0$ の極限をとると，f の連続性により

$$\begin{aligned} \lim_{h \to -0} \frac{F(x+h) - F(x)}{h} &= \lim_{h \to -0} f(c) \\ &= f(\lim_{h \to -0} c) \\ &= f(x) \end{aligned} \tag{6.19}$$

を得る．(6.16) と (6.19) により，

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

であるから，F は任意の $x \in [a,b]$ において微分可能であり，なおかつその値は $f(x)$ に等しいと言える． □

6.14 定理は，$\int_a^x f(t)dt$ が f の原始関数の一つとなることを示している．したがって，$[a,b]$ 上で定義される連続関数 f の任意の原始関数 F は，C を積分定数として

$$F(x) = \int_a^x f(t)dt + C$$

のように表すことができる．

6.15 定理. f を $[a,b]$ で連続な関数とすると，f の任意の原始関数 F について

$$\int_a^b f(x)dx = F(b) - F(a) \tag{6.20}$$

が成立する．

6.15 定理の証明. 微積分学の基本定理により，C を積分定数として

$$F(x) = \int_a^x f(t)dt + C$$

と書ける．したがって

$$\begin{aligned} F(b) - F(a) &= \int_a^b f(t)dt + C - \int_a^a f(a)dt - C \\ &= \int_a^b f(t)dt \end{aligned}$$

が成立する． □

なお，記号 $\big[F(x)\big]_a^b$ を

$$\big[F(x)\big]_a^b := F(b) - F(a)$$

と約束すると，6.15 定理における (6.20) は，

$$\int_a^b f(x)dx = \big[F(x)\big]_a^b$$

と表すことができる．

6.1 例. 微積分学の基本定理を用いて，具体的な関数の定積分を求めてみよう．例えば関数 f と区間 $[a,b]$ を

$$f(x)=\frac{1}{1+x},\quad [a,b]=[0,1]$$

と定め，f の定積分を求めることを考える．区間 $[0,1]$ の分割として $\Delta=\{0,\frac{1}{n},\frac{2}{n},\dots,\frac{n}{n}\}$ を用い，$\xi_i=\frac{i}{n}\in[\frac{i-1}{n},\frac{i}{n}]$（$i=1,2,\dots,n$）とすれば，リーマン和は

$$\begin{aligned}S(\Delta,\xi) &= \sum_{i=1}^{n}f\Big(\frac{i}{n}\Big)\Big(\frac{i}{n}-\frac{i-1}{n}\Big)\\ &= \frac{1}{n}\left(\frac{1}{1+\frac{1}{n}}+\frac{1}{1+\frac{2}{n}}+\cdots+\frac{1}{1+\frac{n}{n}}\right)\\ &= \left(\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{n+n}\right)\end{aligned}$$

のように書ける．ここで $n\to\infty$ の極限をとれば $\delta(\Delta)\to 0$ となるから，定積分は

$$\int_0^1 f(x)dx=\lim_{n\to\infty}\left(\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{n+n}\right)$$

であるが，右辺の値がはっきりしない．そこで基本定理を用いれば，$F(x)=\log(1+x)$ は明らかに $f(x)$ の原始関数であるから

$$\begin{aligned}\int_0^1 f(x)dx &= F(1)-F(0)\\ &= \log(1+1)-\log(1+0)\\ &= \log(2)\end{aligned}$$

が成立する．つまり，

$$\lim_{n\to\infty}\left(\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{n+n}\right)=\log(2)$$

であることが分かる．

6.2 例. 2 変数の連続関数についても，基本定理を適用してみよう．f を定義域上で連続な 2 変数関数とする．まずは $f(x,y)$ の変数 y を固定して

$$g(x,y)=\int_b^x f(t,y)dt$$

と置けば，微積分学の基本定理により

$$\frac{\partial}{\partial x}g(x,y)=\frac{\partial}{\partial x}\int_b^x f(t,y)dt=f(x,y)$$

である．さらに $g(x,y)$ の変数 x を固定して

$$F(x,y)=\int_a^y g(x,u)du$$

と置けば，微積分学の基本定理により

$$\frac{\partial}{\partial y}F(x,y)=\frac{\partial}{\partial y}\int_a^y g(x,u)du=g(x,y)$$

を得る．したがって，

$$F(x,y)=\int_a^y\int_b^x f(t,u)dtdu$$

で定義される F について

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

が成り立つことが分かる．

6.3 例． 関数の凹凸は，普通は微分によって特徴づけられる．しかし微分と積分とを結びつける基本定理を用いれば，関数の凹凸を積分によって特徴づけることもできる．例えば f を区間 $[a,b]$ で連続微分可能な凹関数であるとすれば，f' は区間 $[a,b]$ で連続な非増加関数となる．これを逆から考えれば，g が区間 $[a,b]$ において連続な非増加関数である時，微積分学の基本定理により

$$G(x)=\int_a^x g(t)dt$$

で定義される関数 G は $[a,b]$ で連続微分可能な凹関数であると言える．つまり連続微分可能な凹（凸）関数は，積分を介して連続な非増加（非減少）関数と一対一に対応付けられるのである．

6.4 例． 凹関数と非増加関数が対応付けられるように，増加関数と非負関数も一対一の関係にある．例えばある関数 f が $[a,b]$ において非負の値をとるならば，微積分学の基本定理により，

$$F(x)=\int_a^x f(t)dt$$

で定義される関数 F は $[a,b]$ において増加関数である．そして逆に連続微分可能な関数 G が $[a,b]$ で増加関数である時，$[a,b]$ で非負の値をとるある関数 g が存在し，

$$G(x)=\int_a^x g(t)dt$$

の関係が成立するのである．

6.1 注意．

$$\lim_{n\to\infty}\int_a^\infty f_n(x)\,dx=\int_a^\infty\left\{\lim_{n\to\infty}f_n(x)\right\}dx$$

は必ずしも成立しないが，

$$\lim_{n\to\infty}\int_a^b f_n(x)\,dx=\int_a^b\left\{\lim_{n\to\infty}f_n(x)\right\}dx$$

は，$f_n(x)$ や $\left\{\lim_{n\to\infty}f_n(x)\right\}$ が連続関数ならば成立する．また，

$$\frac{\partial}{\partial y}\int_a^b f(x,y)\,dx=\int_a^b\left\{\frac{\partial}{\partial y}f(x,y)\right\}dx$$

は必ずしも成立しないが，$\frac{\partial}{\partial y}f(x,y)$ が連続ならば成立する．(6.3 節を参照)

以上の微積分学の基本定理を用いて，前の節で不定積分の性質として確認した部分積分や置換積分の公式が定積分にも当てはまることを見ておこう．

6.16 定理 (部分積分)． f と g は $[a,b]$ で連続な関数であるとする．この時 f と g が微分可能であれば

$$\int_a^b f'(x)g(x)dx=\Big[f(x)g(x)\Big]_a^b-\int_a^b f(x)g'(x)dx$$

が成立する．

6.16 定理の証明. 関数の積の微分の公式により

$$\forall x \in [a,b], \quad \frac{d}{dx}\big\{f(x)g(x)\big\} = f'(x)g(x) + f(x)g'(x)$$

であるから，$f(x)g(x)$ は $f'(x)g(x)+f(x)g'(x)$ の原始関数の一つである．したがって 6.15 定理を用いれば

$$\int_a^b \big\{f'(x)g(x) + f(x)g'(x)\big\}dx = \Big[f(x)g(x)\Big]_a^b$$

と書け，定積分の線形性により

$$\int_a^b f'(x)g(x)dx + \int_a^b f(x)g'(x)dx = \Big[f(x)g(x)\Big]_a^b$$

が言える． □

6.17 定理 (置換積分). f は $[a,b]$ で連続な関数であり，変数 x と変数 t とが

$$x = \varphi(t)$$

によって関係付けられているとする．また φ は $[\alpha,\beta]$ で微分可能で，

$$\varphi(\alpha) = a, \qquad \varphi(\beta) = b$$

とする．この時 φ の $[\alpha,\beta]$ における値域が $[a,b]$ に含まれるならば，すなわち任意の $t \in [\alpha,\beta]$ について $\varphi(t) \in [a,b]$ であるならば，

$$\int_a^b f'(x)dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)dt \tag{6.21}$$

が成立する．

6.17 定理の証明. 関数 F を

$$F(x) = \int_a^x f(t)dt$$

と定めれば，微積分学の基本定理により F は $[a,b]$ で微分可能で

$$\forall x \in [a,b], \quad F'(x) = f(x)$$

が成立する．したがって F は f の原始関数の一つであるから，6.15 定理により

$$\int_a^b f(x)dx = F(b) - F(a) \tag{6.22}$$

と書ける．仮定により $\{\varphi(t)|\alpha \le t \le \beta\} \subset [a,b]$ であるから，F は任意の $x \in \{\varphi(t)|\alpha \le t \le \beta\}$ で微分可能で，なおかつ φ も任意の $t \in [\alpha,\beta]$ で微分可能であるから，合成関数の微分の公式により

$$\begin{aligned}\forall t \in [\alpha,\beta], \quad \frac{d}{dt}F\big(\varphi(t)\big) &= F'\big(\varphi(t)\big)\varphi'(t) \\ &= f\big(\varphi(t)\big)\varphi'(t)\end{aligned}$$

が成り立つ．よって $F\big(\varphi(t)\big)$ は，$[\alpha,\beta]$ における $f\big(\varphi(t)\big)\varphi'(t)$ の原始関数の一つである．したがって 6.15 定理により

$$\begin{aligned}\int_\alpha^\beta f(\varphi(t))\varphi'(t)dt &= F\big(\varphi(\beta)\big) - F\big(\varphi(\alpha)\big) \\ &= F(b) - F(a)\end{aligned}$$

が言え，(6.22) と合わせれば定理の結果を得る． □

6.2 注意. 6.17 定理に出てくる $\forall t \in [\alpha, \beta]$, $\varphi(t) \in [a, b]$ という条件は，(6.21) の右辺の定積分が存在することを保証するためのものである．定理の中では f については区間 $[a, b]$ における連続性が仮定されているだけなので，$[a, b]$ 以外の点における積分可能性は保証されていない．したがって，$\varphi(t) \notin [a, b]$ なる $t \in [\alpha, \beta]$ が存在したとすると，f は点 $\varphi(t)$ において連続でないかもしれず，その場合には (6.21) の右辺の定積分は存在しない可能性が出てくるのである．もっともこの説明から明らかなように，$\forall t \in [\alpha, \beta]$, $\varphi(t) \in [a, b]$ という条件は (6.21) の成立に必ずしも必要ではない．たとえ φ の値域が $[a, b]$ をはみ出すものであっても，そのはみ出した部分を含む φ の値域上で f が連続であれば置換積分の公式は成立する．

6.5 例. 部分積分や置換積分の公式を用いて，実際に定積分を計算してみよう．ここでは例として，$f(x) = x^4$ の $[0, 1]$ における定積分を求めてみる．まずは微積分学の基本定理を用いて解くと，$f(x) = x^4$ の原始関数の一つとして $F(x) = \frac{1}{5}x^5$ を考えることができるから，

$$\int_0^1 x^4 dx = F(1) - F(0) = \frac{1}{5}$$

を得る．次に，部分積分を用いた解放を考える．

$$f(x) = x \cdot x^3 = \left\{\frac{1}{2}x^2\right\}' \cdot x^3$$

と見れば，

$$\int_0^1 x^4 dx = \left[\frac{1}{2}x^2 \cdot x^3\right]_0^1 - \int_0^1 \frac{1}{2}x^2 \cdot 3x^2 dx = \frac{1}{2} - \frac{3}{2}\int_0^1 x^4$$

と書ける．よって，これを $\displaystyle\int_0^1 x^4 dx$ について解けば

$$\int_0^1 x^4 dx = \frac{1}{5}$$

が求まる．最後に，置換積分を用いて解いてみよう．例えば $\varphi(t) = 2\sqrt{t}$ と置けば，φ は微分可能で $\varphi'(t) = t^{-\frac{1}{2}}$ である．また $\varphi(0) = 0$, $\varphi(\frac{1}{4}) = 1$ であり，$[0, \frac{1}{4}]$ で φ が単調増加であることに注意すれば，$\forall t \in [0, \frac{1}{4}]$, $\varphi(t) \in [0, 1]$ が満たされることも確認できる．したがって置換積分の公式を当てはめることができ，

$$\begin{aligned}
\int_0^1 x^4 dx &= \int_0^{\frac{1}{4}} f\bigl(\varphi(t)\bigr)\varphi'(t)dt \\
&= \int_0^{\frac{1}{4}} \bigl(2\sqrt{t}\bigr)^4 t^{-\frac{1}{2}} dt \\
&= 16\int_0^{\frac{1}{4}} t^{\frac{3}{2}} dt \\
&= 16\left[\frac{2}{5}t^{\frac{5}{2}}\right]_0^{\frac{1}{4}} \\
&= \frac{1}{5}
\end{aligned}$$

を得る．

6.6 例 (岡田, 241 頁, 例題 7.6). $f(x) = (x^2+1)^{-\frac{1}{2}}$ で定義される関数 f について，区間 $[0, 1]$ における定積分を求める．置換積分の公式で用いる関数 φ として，$(x^2+1)^{\frac{1}{2}} = t - x$ を x について解いたものを考える．すなわち，

$$\varphi(t) = \frac{t^2 - 1}{2t} \tag{6.23}$$

と置く．すると当然ながら

$$\begin{aligned}\left(\big(\varphi(t)\big)^2+1\right)^{\frac{1}{2}} &= t-\varphi(t)\\ &= \frac{t^2+1}{2t}\end{aligned}$$

が成り立つ．またこの φ について，$\varphi(1)=0,\ \varphi(1+\sqrt{2})=1$ であり

$$\forall t\in[1,1+\sqrt{2}\,],\quad \varphi'(t)=\frac{t^2+1}{2t^2}$$

が成立する．したがって，置換積分の公式を用いれば

$$\begin{aligned}\int_0^1(x^2+1)^{-\frac{1}{2}}dx &= \int_1^{1+\sqrt{2}} f\big(\varphi(t)\big)\varphi'(t)dt\\ &= \int_1^{1+\sqrt{2}}\left(\big(\varphi(t)\big)^2+1\right)^{-\frac{1}{2}}\frac{t^2+1}{2t^2}dt\\ &= \int_1^{1+\sqrt{2}}\frac{2t}{t^2+1}\frac{t^2+1}{2t^2}dt\\ &= \int_1^{1+\sqrt{2}}\frac{1}{t}dt\\ &= \big[\log t\big]_1^{1+\sqrt{2}}\\ &= \log(1+\sqrt{2})\end{aligned}$$

を得る．

6.6 例において，$\varphi(t)=0$ の解は $t=1$ 及び $t=-1$ である．したがって (6.23) で定義される φ に対して，$\alpha=-1,\ \beta=1+\sqrt{2}$ と置いても $\varphi(\alpha)=a,\ \varphi(\beta)=b$ は成立し，置換積分の条件を満たしているかに見える．しかしながら，φ は区間 $[-1,1+\sqrt{2}]$ において微分可能でなく，$t=0$ においては定義すらされていない．したがって，$\alpha=-1,\ \beta=1+\sqrt{2}$ では置換積分ができないのである．

練習問題 6.2. 6.6 例において，$\varphi(t)=1$ の解は $t=1+\sqrt{2}$ と $t=1-\sqrt{2}$ であるから，α,β の置き方としては $\alpha=-1,\ \beta=1-\sqrt{2}$ という組み合わせも考えられる．このようにして定義した α,β の下で，置換積分が可能かどうか検討せよ．

6.2.4 広義積分

以上の議論では，関数 f が閉区間 $[a,b]$ で連続な場合に限って定積分を定義した．本節ではこの定義を拡張し，f が積分区間で有限個の不連続点を持つ場合，あるいは積分区間が $[a,\infty)$ のような無限区間となる場合の定積分を考える．

6.5 定義 (広義積分). f を半開区間 $(a,b]$ で連続な関数であるとして，$(a,b]$ における積分を考える．$0<\epsilon<b-a$ を満たす任意の ϵ について，f は $[a+\epsilon,b]$ で連続であるから積分可能で

$$I(\epsilon)=\int_{a+\epsilon}^b f(x)dx$$

が存在する．この $I(\epsilon)$ について $\epsilon\to+0$ の極限が存在するならば，f は $(a,b]$ において（広義）積分可能であるといい

$$\int_a^b f(x)dx=\lim_{\epsilon\to+0}\int_{a+\epsilon}^b f(x)dx$$

で表す．関数 f が半開区間 $[a, b)$ で連続である場合も，$0 < \epsilon < b - a$ を満たす任意の ϵ について

$$I(\epsilon) = \int_a^{b-\epsilon} f(x)dx$$

で定義される $I(\epsilon)$ が $\epsilon \to +0$ で収束するならば，f は $[a, b)$ において（広義）積分可能であるといい

$$\int_a^b f(x)dx = \lim_{\epsilon \to +0} \int_a^{b-\epsilon} f(x)dx$$

と書く．さらに関数 f が開区間 (a, b) で連続である場合については

$$I(\epsilon_1, \epsilon_2) = \int_{a+\epsilon_1}^{b-\epsilon_2} f(x)dx$$

と置き，$I(\epsilon_1, \epsilon_2)$ が $\epsilon_1 \to +0, \epsilon_2 \to +0$ で極限を持つならば，f は (a, b) において（広義）積分可能であるといい

$$\int_a^b f(x)dx = \lim_{\substack{\epsilon_1 \to +0 \\ \epsilon_2 \to +0}} \int_{a+\epsilon_1}^{b-\epsilon_2} f(x)dx$$

で定義する．

具体的な広義積分の値を計算する際には，定積分と同様に微積分学の基本定理を用いる．例えば f の $[a, b)$ における積分を計算する場合，f の原始関数の一つを F として

$$\begin{aligned} I(\epsilon) &= \int_a^{b-\epsilon} f(x)dx \\ &= F(b - \epsilon) - F(a) \end{aligned}$$

と書けるから，広義積分は結局

$$\int_a^b f(x)dx = \lim_{\epsilon \to +0} F(b - \epsilon) - F(a)$$

を計算することになる．

なお，f が当該区間で定積分可能である場合には，定積分の値と広義積分の値とは一致する．つまり f が閉区間 $[a, b]$ で定義され，なおかつ $[a, b]$ で積分可能であるとき，f は $(a, b], (a, b], (a, b)$ のそれぞれで広義積分可能で，いずれの広義積分の値も定積分の値と等しい．f が $[a, b]$ で定積分可能であるとき，ϵ を $0 < \epsilon < b - a$ を満たす任意の数とすれば，$[a, b - \epsilon], [b - \epsilon, b]$ のそれぞれの区間で f は定積分可能である．したがって加法性により，f の $[a, b]$ における定積分は

$$\int_a^b f(x)dx = \int_a^{b-\epsilon} f(x)dx + \int_{b-\epsilon}^b f(x)dx$$

と書ける．ここで $\epsilon \to +0$ の極限を考えると，右辺の第 2 項がゼロに収束することに注意すれば

$$\int_a^b f(x)dx = \lim_{\epsilon \to +0} \int_a^{b-\epsilon} f(x)dx$$

が成立する．これは，f の $[a, b)$ における広義積分が収束し（右辺），その値が $[a, b]$ における定積分（左辺）と一致することに他ならない．$(a, b]$ や (a, b) の広義積分についても，ほぼ同様の議論が適用できる．

練習問題 6.3. 関数 f が $[a, b)$ で（広義）積分可能であるとする．このとき $a \leq c \leq b$ を満たす任意の c について，$[a, c)$ および $[c, b)$ における f の（広義）積分はともに存在し，

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

が成り立つことを示せ．

有限個の不連続点が存在する場合には，積分区間は不連続点を端点としたいくつかの（半）開区間に分けることができる．そしてそれぞれの（半）開区間について広義積分可能であれば，区間全体の広義積分は各部分区間における広義積分の和として定義される．例えば，f が区間 (a,b) におて有限個の点 $c_1 < c_2 < \cdots < c_n$ を除いて連続であるとしよう．このとき (a,b) は $(a,c_1),(c_1,c_2),\dots(c_n,b)$ に分割することができ，それぞれの部分区間で f は連続である．そして，各区間における f の広義積分が収束するならば

$$\int_a^b f(x)dx = \int_a^{c_1} f(x)dx + \int_{c_1}^{c_2} f(x)dx + \cdots + \int_{c_n}^b f(x)dx$$

として，f の (a,b) における広義積分を定義するのである．

6.7 例. 関数 f を

$$f(x) = \begin{cases} 0 & (\ x \leq 0 \text{ のとき }), \\ x^{-\frac{1}{2}} & (\ x > 0 \text{ のとき }). \end{cases}$$

で定義する．$x \to +0$ をとると $f(x)$ は発散してしまうので，f は $[-1,1]$ において定積分可能でない．しかしながら，f は $[-1,0)$ と $(0,1]$ のぞれぞれで広義積分が収束するから，区間 $[-1,1]$ における広義積分は可能で

$$\int_{-1}^1 f(x)dx = \int_{-1}^0 f(x)dx + \int_0^1 f(x)dx$$

である．

練習問題 6.4. $a < c_1 < c_2 < b$ として，f は (a,c_1)，(a,c_2)，(c_1,b)，(c_2,b) のそれぞれで広義積分可能であるとする．このとき

$$\int_a^{c_1} f(x)dx + \int_{c_1}^b f(x)dx = \int_a^{c_2} f(x)dx + \int_{c_2}^b f(x)dx$$

が成り立つことを示せ．

広義積分として拡張された定義の下では，上限や下限が無限であるような場合にも積分を考えることができる．例えば有限な閉区間 $[a,t]$ における定積分

$$I(t) = \int_a^t f(x)dx$$

について，$t \to +\infty$ で $I(t)$ の極限が存在するならば，$[a,\infty)$ における積分を

$$\int_a^\infty f(x)dx = \lim_{t\to+\infty} \int_a^t f(x)dx$$

と書く．$(-\infty,b]$ における積分についても同様に

$$I(t) = \int_t^b f(x)dx$$

で定義される $I(t)$ が，$t \to -\infty$ で収束するならば

$$\int_{-\infty}^b f(x)dx = \lim_{t\to-\infty} \int_t^b f(x)dx$$

と定義する．また積分区間が $(-\infty,\infty)$ のときは

$$I(s,t) = \int_s^t f(x)dx$$

と置いて，$I(s,t)$ について $s \to -\infty, t \to \infty$ の極限を考えればよい．そして極限が存在するならば

$$\int_{-\infty}^\infty f(x)dx = \lim_{\substack{s\to-\infty \\ t\to\infty}} \int_s^t f(x)dx$$

と書くのである．

6.3 微積分の順序交換

以下の議論は教科書では触れられていないが，経済学では頻繁に出てくる話題なので簡単にまとめておく[26]．経済学で積分を扱う場合，積分の結果はしばしばパラメタの関数となる．つまり，例えば 2 変数関数 $f(x,y)$ について

$$F(y) = \int_a^b f(x,y)dx$$

で定義されるような，パラメタ y の関数 F が分析の対象となる．いま，y を微小に変化させた時の積分値の変化を見るために，$F'(y)$ の値を求める必要があるとしよう．f がいくつかの条件を満たせば

$$\begin{aligned} F'(y) &= \frac{d}{dy}\int_a^b f(x,y)dx \\ &= \int_a^b \frac{d}{dy} f(x,y)dx \end{aligned}$$

が成り立つことが知られている．つまり $F'(y)$ を求める際に，f を x について積分してから y で微分するのではなく，先に y で微分してから x について定積分をとっても構わない．次の定理は，その条件を与えるものである．

6.18 定理 (微分と積分の順序交換)． 2 変数関数 $f(x,y)$ が，区間 $[a,b]\times[c,d]$ において

1. x に関して連続
2. y に関して偏微分可能
3. 偏導関数 $f_y(x,y)$ が両変数について連続

の各条件を満たすとする．このとき，x に関する積分と y に関する微分は順序を交換することができる．すなわち

$$\frac{d}{dy}\int_a^b f(x,y)dx = \int_a^b \frac{d}{dy} f(x,y)dx$$

が成立する．

6.18 定理の証明は与えないが，そのもっともらしさは積分が基本的に足し算であることを考えれば理解できる．定積分の定義に立ち返れば

$$\begin{aligned} F(y) &= \int_a^b f(x,y)dx \\ &= \lim_{\delta(\Delta)\to 0}\sum_{i=1}^{n} f(\xi_i,y)(x_i - x_{i-1}) \end{aligned}$$

と書けるから，自然な推論として

$$\begin{aligned} F'(y) &= \frac{d}{dy}\int_a^b f(x,y)dx \\ &= \frac{d}{dy}\left\{\lim_{\delta(\Delta)\to 0}\sum_{i=1}^{n} f(\xi_i,y)(x_i - x_{i-1})\right\} \\ &= \lim_{\delta(\Delta)\to 0}\sum_{i=1}^{n}\left\{\frac{d}{dy} f(\xi_i,y)\right\}(x_i - x_{i-1}) \\ &= \int_a^b \frac{d}{dy} f(x,y)dx \end{aligned}$$

[26]応用例としては 7.4 節を見よ．

が導かれる．この推論における唯一の問題点は 2 行目から 3 行目への移行であり，そこで上記の条件が必要になるのである．

ただし，広義積分の場合には追加的な条件が必要である．なぜなら 6.1 注意で述べたように，

$$\lim_{n\to\infty}\int_a^\infty f_n(x)\,dx=\int_a^\infty\left\{\lim_{n\to\infty}f_n(x)\right\}dx$$

という極限操作の順序交換は必ずしも可能ではないからである．広義積分の場合の条件を，次の定理で与えておく．

6.19 定理 (微分と広義積分の順序交換)． 2 変数関数 $f(x,y)$ が，区間 $[a,b)\times[c,d]$ において[27]

1. x に関して連続
2. y に関して偏微分可能
3. 偏導関数 $f_y(x,y)$ が両変数について連続

の各条件を満たすとする．このとき広義積分

$$\int_a^b f(x,y)dx=\lim_{\epsilon\to+0}\int_a^{b-\epsilon}f(x,y)dx$$

が y に関して一様収束すれば，x に関する積分と y に関する微分は順序を交換することができる．すなわち

$$\frac{d}{dy}\int_a^b f(x,y)dx=\int_a^b\frac{d}{dy}f(x,y)dx$$

が成立する．

6.19 定理における広義積分の一様収束とは，

$$\forall\epsilon>0,\ \exists\bar{h}:\quad \bar{h}<h<b\ \ \Rightarrow\ \ \left|\int_a^h f(x,y)dx-\int_a^b f(x,y)dx\right|<\epsilon$$

が任意の $y\in[c,d]$ で成立することに相当する．

6.4 簡単な変分法の例

次のような問題を考えよう．これは「変分法」と呼ばれる問題の単純な例である．

$$\max_{f\in\mathcal{F}}\int_a^b G\big(f(t)\big)dt\qquad\text{subject to}\quad\int_a^b\Gamma\big(f(t)\big)dt\geq 0\tag{6.24}$$

ただしここで，G と Γ は連続微分可能な凹関数で，$\mathcal{F}$ は区間 $[a,b]$ において連続な関数全体の集合とする．

積分が入っているので厄介に見えるが，「積分は足し算」であると思えば，実はこれはよく見た非線形計画法の問題として以下のように理解できる．記号を簡略化するために $[a,b]=[0,1]$ の場合を考え，区間 $[0,1]$ を n 等分して $[0,\frac{1}{n})$，$[\frac{1}{n},\frac{2}{n})$，... $[\frac{n-1}{n},1]$ のように分割する．記号 $(x_1,\ldots,x_n)$ によって，k 番目の区間上で常に x_k をとる関数を表すことにする．$(x_1,\ldots,x_n)\in\boldsymbol{R}^n$ と見なせるから，$\mathcal{F}_n:=\boldsymbol{R}^n$ と置けば，$\mathcal{F}_n$ は上記の各区間上では一定の値をとる関数全体と見なすことができる．その上で問題

$$\max_{f\in\mathcal{F}_n}\int_a^b G\big(f(t)\big)dt\qquad\text{subject to}\quad\int_a^b\Gamma\big(f(t)\big)dt\geq 0\tag{6.25}$$

[27] $b=+\infty$ の場合を含む．

を考えると,（広義）積分の定義により

$$\begin{aligned}\int_0^1 G\big(f(t)\big)dt &= \int_0^{\frac{1}{n}} G(x_1)dt + \int_{\frac{1}{n}}^{\frac{2}{n}} G(x_2)dt + \cdots + \int_{\frac{n-1}{n}}^{1} G(x_n)dt \\ &= G(x_1)\left(\frac{1}{n}-0\right) + G(x_2)\left(\frac{2}{n}-\frac{1}{n}\right) + \cdots + G(x_n)\left(1-\frac{n-1}{n}\right) \\ &= \frac{1}{n}\sum_{k=1}^{n} G(x_k)\end{aligned}$$

と書けるから，問題 (6.25) は

$$\max_{(x_1,x_2,\ldots,x_n)\in \boldsymbol{R}^n} \sum_{k=1}^{n} G(x_k) \qquad \text{subject to} \quad \sum_{k=1}^{n} \Gamma(x_k) \geq 0 \tag{6.26}$$

と同値である．仮定により G, Γ は凹関数であるから，クーン=タッカー必要条件を満たすことが (6.26) の解として必要十分である．つまり，ある $\lambda \geq 0$ が存在して

$$\begin{aligned}G'(x_k) + \lambda\Gamma'(x_k) &= 0, \quad k = 1, 2, \ldots, n \\ \sum_{k=1}^{n} \Gamma(x_k) &\geq 0 \\ \lambda \sum_{k=1}^{n} \Gamma(x_k) &= 0\end{aligned} \tag{6.27}$$

が成立することが必要十分条件となる．

問題 (6.25) と問題 (6.26) の同値関係を考えれば，(6.24) をあたかも有限のものと見て，おのおのの $f(t)$ を変数と見なしてクーン=タッカー必要条件のようなものを捻出することができる．つまり，問題 (6.24) の解の条件として

$$\begin{aligned}G'\big(f(t)\big) + \lambda\Gamma'\big(f(t)\big) &= 0, \quad t \in [0,1] \\ \int_a^b \Gamma\big(f(t)\big)dt &\geq 0 \\ \lambda \int_a^b \Gamma\big(f(t)\big)dt &= 0\end{aligned} \tag{6.28}$$

を考えるのである．少々野蛮にも見えるが，実は連続関数 f がこの条件を満たすことと，f が問題 (6.24) の解であることは同値になる．詳細を証明するにはこの授業の範囲を超えた準備が必要となるので，以下にアイディアのみを示すことにする．

いま，f^* が問題 (6.24) の解であるとする．仮に

$$\frac{G'\big(f^*(t_1)\big)}{-\Gamma'\big(f^*(t_1)\big)} > \frac{G'\big(f^*(t_2)\big)}{-\Gamma'\big(f^*(t_2)\big)} \tag{6.29}$$

となる 2 点 t_1, t_2 が存在するならば，これは「費用対効果」$\frac{G'}{-\Gamma'}$ が点 t_1 において点 t_2 よりも勝ることを意味するから，f^* を t_1 の付近で若干大きくし，t_2 の付近で制約式を満たす範囲で若干小さくすれば，目的関数の値を上昇させることができる．しかしこれは，f^* が問題 (6.24) の解であることに反する．したがって (6.29) を満たすような t_1, t_2 は存在せず，任意の $t \in [0,1]$ に対して費用対効果は一定の値をとる．そこでその値を λ と置けば，

$$G'\big(f^*(t)\big) + \lambda\Gamma'\big(f^*(t)\big) = 0, \qquad t \in [0,1]$$

となるような $\lambda \geq 0$ が存在すると言え，このことから条件 (6.28) が満たされていることを示せる．

今度は逆に，f^* が条件 (6.28) を満たしているとする．このとき制約式を満たす任意の関数 f に対して，

$$\int_a^b G\big(f^*(t)\big)dt \geq \int_a^b G\big(f(t)\big)dt \tag{6.30}$$

が成立することを示せばよい．まず点 $t_n^{*k} \in [\frac{k}{n}, \frac{k+1}{n})$ をそれぞれの k について適当に選び

$$\forall t \in \left[\frac{k}{n}, \frac{k+1}{n}\right): \quad f_n^*(t) = f^*\left(t_n^{*k}\right)$$

と置けば，f_n^* は k 番目の区間上で常に一定の値 $f^*\left(t_n^{*k}\right)$ をとる関数となるから $f_n^* \in \mathcal{F}_n$ である．その上で，各 $t_n^{*k} \in [\frac{k}{n}, \frac{k+1}{n})$ を

$$\int_a^b \Gamma\big(f_n^*(t)\big)dt = \frac{1}{n}\sum_{k=1}^n \Gamma\left(f^*\left(t_n^{*k}\right)\right) = 0$$

を満たすように定める．仮定により f^* は条件 (6.28) を満たしているので，

$$G'\left(f^*\left(t_n^{*k}\right)\right) + \lambda\Gamma'\left(f^*\left(t_n^{*k}\right)\right) = 0, \quad k = 1, 2, \dots n$$

も成り立つ．この時 $f^*\left(t_n^{*k}\right)$, $k = 1, 2, \dots, n$ はクーン＝タッカー条件 (6.27) を満たし，(6.26) の解となる．一方，制約条件を満たす任意の f に対しても点 $t_n^k \in [\frac{k}{n}, \frac{k+1}{n})$ を適当に選び

$$\forall t \in \left[\frac{k}{n}, \frac{k+1}{n}\right): \quad f_n(t) = f\left(t_n^k\right)$$

と置けば，f_n は k 番目の区間上で常に一定の値 $f\left(t_n^k\right)$ をとる関数となるから $f_n \in \mathcal{F}_n$ である．その上で，各 $t_n^k \in [\frac{k}{n}, \frac{k+1}{n})$ を

$$\int_a^b \Gamma\big(f_n(t)\big)dt = \frac{1}{n}\sum_{k=1}^n \Gamma\left(f\left(t_n^k\right)\right) = 0$$

を満たすように定める．すると $f\left(t_n^k\right)$, $k = 1, 2, \dots, n$ と $f^*\left(t_n^{*k}\right)$, $k = 1, 2, \dots, n$ について，後者が (6.26) の最適解なのだから

$$\sum_{k=1}^n G\left(f^*\left(t_n^{*k}\right)\right) \geq \sum_{k=1}^n G\left(f\left(t_n^k\right)\right)$$

が言える．したがって，(6.26) と (6.25) が同値であることに注意すれば，この時 f_n と f_n^* について後者が (6.25) の解となり

$$\int_a^b G\big(f_n^*(t)\big)dt \geq \int_a^b G\big(f_n(t)\big)dt$$

が成立するはずである．最後に，$f^* = \lim_{n\to\infty} f_n^*$ かつ $f = \lim_{n\to\infty} f_n$ に注意すれば，片々について $n \to \infty$ の極限をとることで (6.30) を得る．

練習問題 6.5. c を正の定数として，次の問題を解け．

$$\max_{f\in\mathcal{F}} \int_0^1 f(t)dt \qquad \text{subject to} \quad \int_0^1 \big(f(t)\big)^2 dt \leq c$$

7 確率と期待値

7.1 確率論の基礎

起こり得る全ての事柄の集合を標本空間 Ω と書き，Ω の部分集合からなる集合族を $\mathcal{F}$ で表す．$\mathcal{F}$ の各要素を $\boldsymbol{R}$ 上の一点と対応づける集合関数を $P:\mathcal{F}\to\boldsymbol{R}$ として，P が次のような確率の公理を満たすとする．

1. $\forall A\in\mathcal{F}: \quad 0\le P(A)\le 1$
2. $P(\Omega)=1$
3. $A_k\in\mathcal{F},\ \forall k\neq l: A_k\bigcap A_l=\emptyset, \ \Rightarrow \ P\left(\bigcup_{k=1}^{\infty}A_k\right)=\sum_{k=1}^{\infty}P\left(A_k\right)$

この時 $(\Omega,\mathcal{F},P)$ の組を確率空間と呼び，$A\in\mathcal{F}$ に対して $P(A)$ を事象 A の確率と言う．確率 $P(A)$ は，標本空間の部分集合 A に対してその起こりやすさの度合いを表すものである．

7.1.1 確率変数と累積分布関数

ある確率空間 $(\Omega,\mathcal{F},P)$ から実数への写像を確率変数 (Random Variable: RV) といい，通常大文字のアルファベットを用いて $X:\Omega\to\boldsymbol{R}$ のように書く．このような確率変数を用いれば，確率を扱う際に標本空間 Ω を捨象して考えることができるようになる．言い換えれば，X を $\boldsymbol{R}$ 上で確率的に変化する数と考えることによって，背後にある確率空間を記述せずに確率を扱えるようになるのである．

例えばある事象 A が生じる確率 $P(A)$ を考えた時，$E=\{X(\omega)|\omega\in A\}$ と置けば，$P(A)=P\big(\{\omega\in\Omega|X(\omega)\in E\}\big)$ と書ける．つまり事象 A が生じる確率は，確率変数 X が対応する集合 E 内に値をとる確率と読み替えることができるのである．ここでは簡単化のために，$P\big(\{\omega\in\Omega|X(\omega)\in E\}\big)$ を $\Pr(X\in E)$ で表す．また，該当する集合 E が区間 $(-\infty,x)$ である場合には，$\Pr(X\le x)$ などと書く．

7.1 定義 (累積分布関数). 確率変数 $X:\Omega\to\boldsymbol{R}$ に対して

$$F_X(x):=\Pr\left(X\le x\right)$$

で定義される関数 $F_X:\boldsymbol{R}\to[0,1]$ を，累積分布関数（Cumulative Distribution Function: CDF）という．

累積分布関数 $F_X(x)$ とは，確率変数 X が x 以下の値をとる確率を記述したものである．まずは累積分布関数の性質を，次の定理で確認しておく．

7.1 定理. 累積分布関数 $F_X:\boldsymbol{R}\to[0,1]$ は，次の性質を満たす．

1. $x<y \ \Rightarrow \ F_X(x)\le F_X(y)$（非減少）
2. $\forall x: \ \lim_{h\to+0}F_X(x+h)=F_X(x)$（右連続）
3. $\lim_{x\to-\infty}F_X(x)=0,$ かつ $\lim_{x\to\infty}F_X(x)=1$

また逆に，7.1 定理の三つの条件を満たす関数 F が与えられれば，それに対応する確率変数が一意に定まることも知られている．したがって，上記の条件を満たすような関数に対しては，累積分布を考えるのが「基本」である．

7.1.2 連続型確率変数と密度関数

累積分布関数 $F_X(x)$ が連続であるとき，X は連続型確率変数であるという．直感的には，連続な区間に値をとる確率変数のことである．X が連続型であるとき，任意の区間 $[a,b] \subset \boldsymbol{R}$ に対して

$$F_X(b) - F_X(a) = \int_a^b f_X(x)dx \tag{7.1}$$

となるような関数 $f_X : \boldsymbol{R} \to \boldsymbol{R}_+$ が存在し，このような $f_X(x)$ を確率変数 X の密度関数（density function）と呼ぶ．また $f_X(x)$ を点 x における確率密度というが，$F_X(x)$ が微分可能である場合には (7.1) を b で微分することで

$$\frac{d}{dx}F_X(x) = f_X(x)$$

によって密度を求めることができる．ここで確率密度 $f_X(x)$ が意味するものは

$$\begin{aligned} f_X(x) &= \lim_{h\to 0} \frac{F_X(x+h) - F_X(x)}{h} \\ &= \lim_{h\to 0} \frac{\Pr(x \le X \le x+h)}{h} \end{aligned}$$

すなわち X が $[x, x+h]$ に値をとる確率と h の比について，$h \to 0$ で極限をとったものである．

また，確率変数が区間 $[a, a+h]$ に値をとる確率は

$$\begin{aligned} \Pr(a \le x \le a+h) &= \Pr(x \le a+h) - \Pr(x \le a) \\ &= F_X(a+h) - F_X(a) \end{aligned}$$

であるから，$h \to +0$ の極限をとれば

$$\begin{aligned} \Pr(x = a) &= F_X(a) - F_X(a) \\ &= 0 \end{aligned}$$

したがって，連続型の確率変数が特定の値をとる確率（一点集合の確率）はゼロである．なお，確率変数が区間 $[a,b]$ に値をとる確率を

$$\begin{aligned} \Pr(a \le x \le b) &= \Pr(x \le b) - \Pr(x \le a) \\ &= F_X(b) - F_X(a) \\ &= \int_a^b f_X(z)dz \end{aligned}$$

と表せることにも注意しておく．

7.1 注意. 連続型の確率変数に対して，密度関数は必ずしも一意に定まらない．例えば g を積分可能な関数とすれば，積分区間のある 1 点において g の値を変更したとしても依然として積分可能であり，その積分の値も変わらない．したがって $f_X(x)$ が X の密度関数である時，$f_X(x)$ を 1 点においてのみ変更した関数もまた X の密度関数となるのである．これを逆に言えば，離散的な有限個の点を除いて密度関数が定まる限りで，その有限個の点における密度関数の値はどのようなものでもかまわないことになる．

7.1 例. 累積分布関数 $F_X(x)$ が，次のような形で与えられている確率変数 X を考える．（グラフを描

いてみること）

$$F_X(x)=\begin{cases} 0 & （\ x<0\ のとき\ ） \\ \frac{1}{4}+x & （\ 0\le x<\frac{1}{4}\ のとき\ ） \\ \frac{1}{2} & （\ \frac{1}{4}\le x<\frac{1}{2}\ のとき\ ） \\ \frac{1}{4}+\frac{1}{2}x & （\ \frac{1}{2}\le x<\frac{3}{2}\ のとき\ ） \\ 1 & （\ \frac{3}{2}\le x\ のとき\ ） \end{cases}$$

関数 $F_X(x)$ が累積分布関数の性質を全て満たしていることは容易に確認できる．また $F_X(x)$ は点 $x=0$ を除いて連続であるから，区間 $(-\infty,0)$ と区間 $(0,\infty)$ に限ってみれば，それぞれの区間においては密度関数を持つ連続型の確率変数と見なせる．$x=\frac{1}{2}$ 及び $x=\frac{3}{2}$ の 2 点を除けばいずれの区間においても $F_X(x)$ は微分可能であるから，密度関数は $F_X(x)$ の微係数として与えられるはずである．そこで微分可能な区間において

$$f_X(x)=\frac{d}{dx}F_X(x)$$

によって $f_X(x)$ を定義すると，

$$f_X(x)=\begin{cases} 0 & （\ x<0\ のとき\ ） \\ 1 & （\ 0\le x<\frac{1}{4}\ のとき\ ） \\ 0 & （\ \frac{1}{4}\le x<\frac{1}{2}\ のとき\ ） \\ \frac{1}{2} & （\ \frac{1}{2}\le x<\frac{3}{2}\ のとき\ ） \\ 0 & （\ \frac{3}{2}\le x\ のとき\ ） \end{cases}$$

を得る．そして実際にこの $f_X(x)$ について

$$F_X(b)-F_X(a)=\int_a^b f_X(x)dx \tag{7.2}$$

が任意の $a,b>0$ で成立し，$f_X(x)$ は X の密度関数であることが確認できる．

7.1.3 離散型確率変数と確率関数

分布関数 $F_X(x)$ の値が不連続点（離散点）以外で増加しない時，X は離散型確率変数であると言う．離散型確率変数 X に対して，関数 $p_X(x)$ を

$$p_X(x)=F_X(x)-\lim_{h\to+0}F_X(x-h)$$

で定義し，$p_X(x)$ を X の確率関数（probability function）と呼ぶ．実際，X が区間 $[a-h,a]$ に値をとる確率は

$$\begin{aligned}\Pr(a-h\le x\le a) &= \Pr(x\le a)-\Pr(x\le a-h)\\ &= F_X(a)-F_X(a-h)\end{aligned}$$

であり，$h\to+0$ の極限をとれば

$$\Pr(x=a)=F_X(a)-\lim_{h\to+0}F_X(a-h)$$

となるから，$p_X(a)$ は確率変数 X が a の値をとる「確率」を表すことが分かる．

練習問題 7.1. x_1, x_2, x_3 を，それぞれ確率 $p_X(x_1), p_X(x_2), p_X(x_3)$ でとるような確率変数 X を考える．すなわち

$$p_X(x_i) \geq 0, \quad i = 1, 2, 3, \quad \text{かつ} \quad \sum_{i=1}^{3} p_X(x_i) = 1$$

とする．この時，X に対応する累積分布関数を求めよ．

7.2 期待値と分散

7.2.1 期待値

7.2 定義 (離散型確率変数の期待値). 離散型の確率変数 X に対して，$F_X(x)$ を $[a, b]$ 上の累積分布関数とする[28]．また $F_X(x)$ の不連続点を，x_i（$i = 1, 2, \ldots$）と置く．この時，X の実現値を定義域とする関数 h に対して

$$E[h(X)] = \sum_{i=1}^{\infty} h(x_i) p_X(x_i)$$

で定義される $E[h(X)]$ を，$h(X)$ の期待値（exptected value）あるいは平均値（mean）という．ただしここで，関数 $p_X(x)$ は

$$p_X(x) = F_X(x) - \lim_{h \to +0} F_X(x - h)$$

で定義されるものとする．

7.3 定義 (連続型確率変数の期待値). 連続型の確率変数 X に対して，$F_X(x)$ を $[a, b]$ 上の累積分布関数とする．この時，X の実現値を定義域とする関数 h に対して

$$E[h(X)] = \int_a^b h(z) f_X(z) dz$$

で定義される $E[h(X)]$ を，$h(X)$ の期待値（平均値）という．ただしここで，関数 $f_X(x)$ は

$$F_X(b) - F_X(a) = \int_a^b f_X(x) dx \tag{7.3}$$

を満たす密度関数である．

7.2 定義と 7.3 定義において $h(x) = x$ としたものを，確率変数 X の期待値（平均値）という．たとえば X が連続型の確率変数である場合，$F_X(x)$ が微分可能えあればその微係数が密度関数となるので

$$E[X] = \int_a^b x \left(\frac{d}{dx} F_X(x) \right) dx$$

が X の期待値である．

確率変数が離散型の場合と連続型の場合とで期待値の定義を分けて書いているが，いずれも同じ基準を用いて定義が与えられていることに注意しよう．X が連続型であれ離散型であれ，$h(X)$ の期待値は

$$E[h(X)] = F_X(b)h(b) - F_X(a)h(a) - \int_a^b F_X(z) h'(z) dz \tag{7.4}$$

のように統一的に表現でき，特に $F_X(a) = 0,\ F_X(b) = 1$ となるような a と b が存在すれば

$$E[h(X)] = h(b) - \int_a^b F_X(z) h'(z) dz$$

と書けるのである．このような表現を見れば，たとえば X が離散型でも連続型でもない場合でも，同じ基準に基づいて期待値の定義が与えられていることが分かる．

[28] すなわち，$F_X(a) = 0,$ かつ $F_X(b) = 1$ である

練習問題 7.2. x_1, x_2, x_3 を，それぞれ確率 $p_X(x_1), p_X(x_2), p_X(x_3)$ でとるような確率変数 X を考える．すなわち

$$p_X(x_i) \geq 0, \quad i = 1, 2, 3, \quad \text{かつ} \quad \sum_{i=1}^{3} p_X(x_i) = 1$$

である．この時，確率変数 X の期待値が

$$E[X] = b - \int_a^b F_X(z)dz$$

の形で書けることを確認せよ．

7.2 注意. 区間 $[a, \infty]$ で議論する時には，積分の定義により

$$\begin{aligned} \int_a^\infty h(x)\left(\frac{d}{dx}F_X(z)\right)dz &= \lim_{b\to\infty}\int_a^b h(x)\left(\frac{d}{dx}F_X(z)\right)dz \\ &= \lim_{b\to\infty}\left\{h(b) - \int_a^b F_X(z)h'(z)dz\right\} \end{aligned}$$

となる．しかしながら，右辺の $h(b)$ と $\int_a^b F_X(z)h'(z)dz$ について，それぞれ個別に $b \to \infty$ の極限が存在するとは限らないので

$$\int_a^\infty h(x)\left(\frac{d}{dx}F_X(z)\right)dz = \lim_{b\to\infty} h(b) - \lim_{b\to\infty}\int_a^b F_X(z)h'(z)dz$$

あるいは

$$\int_a^\infty h(x)\left(\frac{d}{dx}F_X(z)\right)dz = h(\infty) - \int_a^\infty F_X(z)h'(z)dz$$

のように書くことはできない．区間を $(-\infty, \infty)$ として考える場合も同様である．

7.2.2 累積分布関数による期待値の表現

期待値に対する (7.4) のような表現は，累積分布関数を用いて期待値を計算できることを示している．そこで期待値を表す際に，累積分布関数 $F_X(x)$ を記号として直接用いることを考える．すなわち，

$$\int_a^b h(x)dF_X = E[h(X)] \qquad \text{あるいは} \qquad \int_a^b h(x)F_X(dx) = E[h(X)]$$

と約束するのである[29]．(7.4) と合わせれば

$$\int_a^b h(z)dF_X = F_X(b)h(b) - F_X(a)h(a) - \int_a^b F_X(z)h'(z)dz \tag{7.5}$$

であり，$F_X(a) = 0,\ F_X(b) = 1$ である場合には

$$\int_a^b h(x)dF_X = h(b) - \int_a^b F_X(z)h'(z)dz$$

と書ける．

また (7.4) の式は，累積分布関数 $F_X(x)$ をいくつかの関数に分解して関数ごとに「期待値」を計算し，その和をとることで期待値が求められることを示している．例えば，任意の $z \in [a, b]$ に対して

$$F_X(z) = F_1(z) + F_2(z) \tag{7.6}$$

[29] これをスティルチェス型積分（Stieltjes Integral）という．詳しくは定義が必要であるが，実際上は，離散型と連続型で別の表現を用いるのが面倒なのでこの形で一括して表記するのだ，と記憶しておいて問題は生じないだろう．

が成り立つように，$F_X(x)$ を $F_1(x)$ と $F_2(x)$ に分解できたとしよう．ただし，$F_k(x) \geq 0$（$k=1,2$）とする．このとき期待値 $E[h(X)]$ は，(7.4) により

$$\begin{aligned} E[h(X)] &= \int_a^b h(x)dF_X \\ &= h(b) - \int_a^b F_X(z)h'(z)dz \\ &= h(b) - \int_a^b \Big\{F_1(z)+F_2(z)\Big\}h'(z)dz \\ &= h(b) - \int_a^b F_1(z)h'(z)dz - \int_a^b F_2(z)h'(z)dz \end{aligned} \tag{7.7}$$

と書ける．一方，$k=1,2$ について

$$\int_a^b h(z)dF_k(z) = F_k(b)h(b) - F_k(a)h(a) - \int_a^b F_k(z)h'(z)dz$$

と表せば，$F_1(b)+F_2(b)=F_X(b)=1$ および $F_1(a)+F_2(a)=F_X(a)=0$ に注意して

$$\int_a^b h(z)dF_1(z) + \int_a^b h(z)dF_2(z) = h(b) - \int_a^b F_1(z)h'(z)dz - \int_a^b F_2(z)h'(z)dz \tag{7.8}$$

を得る．したがって，(7.7)，(7.8) により

$$E[h(X)] = \int_a^b h(z)dF_1(z) + \int_a^b h(z)dF_2(z)$$

の関係を確認することができる．

この関係を用いれば，分布関数を離散型の部分と連続型の部分とに分解して期待値を求めることが可能となる．つまり (7.6) において，$F_1(z)$ が離散型，$F_2(z)$ が連続型となるように $F_X(z)$ を分解すれば

$$\int_a^b h(z)dF_1(z) = \sum_{i=1}^{\infty} h(x_i)p(x_i)$$

$$\int_a^b h(z)dF_2(z) = \int_a^b h(z)\frac{d}{dz}F_2(z)dz$$

となるから，期待値は

$$E[h(X)] = \sum_{i=1}^{\infty} h(x_i)p(x_i) + \int_a^b h(z)\frac{d}{dz}F_2(z)dz$$

によって計算することができる．ただし，$F_1(z)$ や $F_2(z)$ 自体は累積分布関数ではないので注意すること．例えば

$$\sum_{i=1}^{\infty} p(x_i) + \int_a^b \frac{d}{dz}F_2(z)dz = 1$$

であるから，

$$\sum_{i=1}^{\infty} p(x_i) < 1 \quad \text{かつ} \quad \int_a^b \frac{d}{dz}F_2(z)dz < 1$$

となる．

7.2 例. 確率変数 X について，その分布関数 $F_X(x)$ が

$$F_X(x) = \begin{cases} 0 & (\ x < 0 \text{ のとき }), \\ \min\{\frac{1}{2}+\frac{1}{2}x, 1\} & (\ 0 \leq x \text{ のとき }). \end{cases}$$

で与えられているとする．$F_X(0)=0$ かつ $F_X(1)=1$ であるから，期待値をとる区間としては $[0,1]$ を考えればよい．ここで $F_1(x)$ と $F_2(x)$ を

$$F_1(x)=\begin{cases} 0 & (\ x<0\ のとき\ ), \\ \dfrac{1}{2} & (\ 0\le x\ のとき\ ). \end{cases}$$

$$F_2(x)=\min\left\{\frac{1}{2}+\frac{1}{2}x,1\right\}$$

と定義すれば，任意の $x\in[0,1]$ に対して

$$F_X(x)=F_1(x)+F_2(x)$$

が成り立つ．$F_1(x)$ は離散型で，不連続点は $x=0$ のみであるから

$$\begin{aligned}\int_0^1 xdF_1(x) &= 0\times p(0)\\ &= 0\end{aligned}$$

である．一方の $F_2(x)$ は連続型で，密度は任意の $x\in[0,1]$ において

$$\frac{d}{dx}\left(\frac{1}{2}+\frac{1}{2}x\right)=\frac{1}{2}$$

であるから

$$\begin{aligned}\int_0^1 xdF_2(x) &= \int_0^1 \frac{1}{2}xdx\\ &= \frac{1}{4}\end{aligned}$$

を得る．したがって X の期待値は

$$\begin{aligned}E[h(X)] &= \int_0^1 xdF_1(x)+\int_0^1 xdF_2(x)\\ &= 0+\frac{1}{4}\\ &= \frac{1}{4}\end{aligned}$$

のように計算できる．

練習問題 7.3. 7.2 例について (7.4) を用いて $E[X]$ を計算し，上で求めた結果と一致することを確認せよ．

7.2.3 期待値の性質

期待値には以下のような性質がある．

7.2 定理 (期待値の線形性). X を確率変数とすると，任意の実数 α,β に対して

$$E[\alpha X+\beta]=\alpha E[X]+\beta$$

が成立する．

7.2 定理の証明. 期待値の定義により

$$
\begin{aligned}
E[\alpha X+\beta] &= \int_a^b (\alpha x+\beta)dF_X(x)\\
&= \alpha b+\beta-\int_a^b F_X(x)\alpha dx\\
&= \alpha\left(b-\int_a^b F_X(x)dx\right)+\beta\\
&= \alpha\int_a^b x dF_X(x)+\beta\\
&= \alpha E[X]+\beta
\end{aligned}
$$

が直ちに導かれる. □

7.3 定理 (期待値の単調性). 区間 $[a,b]$ で連続な関数 f および g が

$$\forall x\in[a,b],\quad f(x)\le g(x)$$

を満たしているとする. このとき確率変数 X に対して

$$E[f(X)]\le E[g(X)]$$

が成立する.

7.3 定理の証明. まず, X が連続型の場合を考える. 密度関数 f_X は非負の値をとり, 仮定により $f(x)\le g(x)$ が任意の $x\in[a,b]$ で成り立つから

$$\forall x\in[a,b]:\quad (f(x)-g(x))\,f_X(x)\le 0$$

が言える. したがって積分の線形性と単調性により

$$
\begin{aligned}
E[f(X)]-E[g(X)] &= \int_a^b f(x)f_X(x)dx-\int_a^b g(x)f_X(x)dx\\
&= \int_a^b (f(x)-g(x))\,f_X(x)dx\\
&\le 0
\end{aligned}
$$

が成り立つ. X が離散型の場合も同様である. 確率関数 p_X は非負の値をとり, 仮定により $f(x_i)\le g(x_i)$ が $i=1,2,\ldots$ で成り立つから

$$(f(x_i)-g(x_i))\,p_X(x_i)\le 0,\qquad i=1,2,\ldots$$

が言える. よって

$$
\begin{aligned}
E[f(X)]-E[g(X)] &= \sum_{i=1}^{\infty} f(x_i)p_X(x_i)-\sum_{i=1}^{\infty} g(x_i)p_X(x_i)\\
&= \sum_{i=1}^{\infty}(f(x_i)-g(x_i))\,p_X(x_i)\\
&\le 0
\end{aligned}
$$

を得る. □

7.2.4 分散と標準偏差

分散と標準偏差についても簡単にまとめておく．

7.4 定義 (分散). 確率変数 X に対して

$$Var[X] = E\Big[(X - E[X])^2\Big]$$

で定義される $Var[X]$ を，X の分散（variance）と呼ぶ．$Var[X]$ は $\sigma^2[X]$ とも表す．

期待値の線形性を用いれば

$$\begin{aligned} E\Big[(X - E[X])^2\Big] &= E\Big[X^2 - 2E[X]X + \big(E[X]\big)^2\Big] \\ &= E\big[X^2\big] - 2E[X]E[X] + \big(E[X]\big)^2 \\ &= E\big[X^2\big] - \big(E[X]\big)^2 \end{aligned}$$

であるから，

$$Var[X] = E\big[X^2\big] - \big(E[X]\big)^2$$

と書ける．したがって，a, b を定数とした時の $aX + b$ の分散について

$$\begin{aligned} Var[aX+b] &= E\big[(aX+b)^2\big] - \big(E[aX+b]\big)^2 \\ &= E\big[a^2X^2 + 2abX + b^2\big] - \big(aE[X] + b\big)^2 \\ &= a^2E[X^2] + 2abE[X] + b^2 - a^2\big(E[X]\big)^2 - 2abE[X] - b^2 \\ &= a^2\Big(E[X^2] - \big(E[X]\big)^2\Big) \\ &= a^2Var[X] \end{aligned} \tag{7.9}$$

が成り立つことが分かる．

7.5 定義 (標準偏差). 確率変数 X に対して

$$\sigma[X] = \sqrt{Var[X]}$$

で定義される $\sigma[X]$ を，X の標準偏差（standard diviation）と呼ぶ．

(7.9) を用いれば，$aX + b$ の標準偏差は

$$\begin{aligned} \sigma[aX+b] &= \sqrt{Var[aX+b]} \\ &= \sqrt{a^2Var[X]} \\ &= |a|\sigma[X] \end{aligned}$$

として計算できる．

7.3 注意. 期待値や分散，標準偏差を持たない確率変数も存在する．連続関数の期待値が存在することは積分 $\int_a^b xf(x)dx$ が存在することにほかならないが，$xf(x)$ が x に関して積分可能であるとは限らないからである．分散や標準偏差についても場合も同様で，$x^2f(x)$ が x に関して積分可能でなければならない．密度関数 $f(x)$ 自体は x について積分可能であっても，それは $xf(x)$ や $x^2f(x)$ の積分可能性を保証するものではない．

7.3 期待効用と確率優位

本節では，確率分布の「大小関係」を考えることにする[30]．というのも，確率分布の大小関係を考えることは対応する確率変数を順序づけることに他ならず，異なる確率変数の間に何らかの意味で順序関係を定めることができれば，確率変数の実現値の増減に対する個人の反応を分析できるからである．

F と G という二つの累積分布関数について，これらを順序づける方法を考えてみよう．おそらく最も単純な方法は，

$$\forall x, \;\; F(x) \geq G(x)$$

が成り立つ時に，F と G の順序関係を定めるというものである．全ての累積分布関数の組に対してこのような大小関係が確定できるわけではないが，ひとまずこの順序づけの方法が意味するものについて考えてみる．

極限操作に伴う微妙な問題を回避するため[31]，これ以降では a, b が存在し

$$F(a) = G(a) = 0, \qquad F(b) = G(b) = 1 \tag{7.10}$$

が成り立つと仮定する．つまり F や G を累積分布関数とする確率変数は，いずれも区間 $[a, b]$ の範囲に実現値をとるものと考える．

7.3.1 1 次の確率優位

u を区間 $[a, b]$ で定義された微分可能な非減少関数とすれば，$u(X)$ の期待値は，X が F に従う場合と G に従う場合でそれぞれ

$$\int_a^b u(x)dF(x) = u(b) - \int_a^b F(z)u'(z)dz$$

$$\int_a^b u(x)dG(x) = u(b) - \int_a^b G(z)u'(z)dz$$

である．ここで

$$\begin{aligned}\int_a^b u(x)dF(x) - \int_a^b u(x)dG(x) &= u(b) - \int_a^b F(z)u'(z)dz - u(b) + \int_a^b G(z)u'(z)dz \\ &= \int_a^b G(z)u'(z)dz - \int_a^b F(z)u'(z)dz \\ &= \int_a^b \Big\{G(z) - F(z)\Big\}u'(z)dz\end{aligned}$$

と書けるから，$u'(x) \geq 0$ に注意すれば

$$\forall x \in [a, b], \;\; F(x) \geq G(x) \quad \Rightarrow \quad \int_a^b u(x)dG(x) \geq \int_a^b u(x)dF(x) \tag{7.11}$$

が成立することが分かる．

一般的に経済学では効用関数を増加関数と仮定するので，上記の u は効用関数とみなすことができる．u を効用関数とみなした場合，$u(X)$ の期待値は X の確率分布から得る期待効用を表すものと考えられる．したがって (7.11) は，単調非減少の効用関数を持つどのような個人にとっても，G の期待効用が F の期待効用を上回ることを意味している．つまり $F(x) \geq G(x)$ という順序関係に対し

[30] この節の内容を詳しく学びたければ，MWG の第 6 章を見よ．
[31] 7.2 注意を参照．

ては，一般的な効用関数を持つ個人にとって F よりも G の方が好ましいという経済学的な意味を与えることができるのである．そして実は，(7.11) における逆方向の命題も成立することが知られている．

7.4 定理. (7.10) を満たす二つの累積分布関数 F と G に関して，次の 2 条件は同値である．

1. 任意の $x \in [a,b]$ について，$F(x) \geq G(x)$
2. $[a,b]$ で定義された任意の微分可能な非減少関数 u について，$\displaystyle\int_a^b u(x)dG(x) \geq \int_a^b u(x)dF(x)$

7.4 定理の証明（アイディア）． 条件 1 が条件 2 から導かれることを背理法によって示す．ある $\hat{x} \in [a,b]$ が存在して，$F(\hat{x}) < G(\hat{x})$ が成立していると仮定しよう．すると F と G の右連続性により，$\hat{x}$ を含む十分小さな区間 $[\alpha,\beta]$ が存在し

$$\forall x \in [\alpha,\beta]: \quad F(x) - G(x) < 0 \tag{7.12}$$

が成立する[32]．ここで

$$\forall x \in [\alpha,\beta]: \;\; u'(x) > 0, \quad \text{かつ} \quad \forall x \in [a,b] \setminus [\alpha,\beta]: \;\; u'(x) = 0 \tag{7.13}$$

を満たす関数 u' を考えると，微積分学の基本定理により

$$u(x) = \int_a^x u'(t)dt$$

によって定義される u は，$[a,b]$ 上で微分可能な単調非減少関数となる．したがって条件 2 により，この u について

$$\int_a^b u(x)dG(x) - \int_a^b u(x)dF(x) \geq 0$$

が成立しなければならない．しかし一方で，(7.12) と (7.13) により

$$\begin{aligned}
\int_a^b u(x)dF(x) - \int_a^b u(x)dG(x) &= u(b) - \int_a^b F(z)u'(z)dz - u(b) + \int_a^b G(z)u'(z)dz \\
&= \int_a^b \Big\{G(z) - F(z)\Big\}u'(z)dz \\
&= \int_\alpha^\beta \Big\{G(z) - F(z)\Big\}u'(z)dz \\
&< 0
\end{aligned}$$

となるから矛盾が導かれる．よって仮定は誤りで，条件 1 が成立する． □

7.4 定理により，任意の $x \in [a,b]$ について $F(x) \geq G(x)$ が成り立つ時，そしてその時のみ，経済学に登場するどの個人も F よりも G を好むという意味で，G の方が F よりも優位であると解釈できる．このような順序関係に対しては，以下のような定義が与えられている．

7.6 定義 (**1 次の確率優位**)**.** 区間 $[a,b]$ に値をとる確率変数 X, Y について，対応する累積分布関数をそれぞれ，$F_X(x)$, $F_Y(x)$ と置く．任意の $x \in [a.b]$ に対して $F_X(x) \leq F_Y(x)$ が成立する時，X は Y よりも 1 次の確率優位にある（X first-order stochastically dominates Y）という．あるいは累積分布関数について，F_X は F_Y よりも 1 次の確率優位にあるという．

[32] この部分には別途証明が必要．

練習問題 7.4. 任意の $x \in [a,b]$ について $F(x) \geq G(x)$ が成立し，なおかつ $F(\hat{x}) > G(\hat{x})$ となる $\hat{x} \in [a,b]$ が存在したとする．このとき任意の $x \in [a,b]$ について $u'(x) > 0$ であるならば，G よりも F の方が期待効用は厳密に小さくなる，すなわち

$$\int_a^b u(x)dF(x) - \int_a^b u(x)dG(x) < 0$$

が成立することを示せ.（一般形が難しければ，連続型の場合のみ示せ.）

1 次の確率優位と期待値との関係についても簡単に触れておく．確率変数 X が確率変数 Y よりも 1 次の確率優位にある場合，任意の微分可能な単調非減少関数 u について

$$E[u(X)] \geq E[u(Y)] \tag{7.14}$$

が成立する．$u(z) = z$ は微分可能な単調非減少関数であるから，このように選んだ u についても (7.14) は成り立つはずで

$$E[X] \geq E[Y]$$

が言える．つまり X が Y よりも 1 次の確率優位にあるならば，X の期待値は Y の期待値よりも大きい．もっとも，その逆は一般に成立しない．

7.3.2　2 次の確率優位とリスク

経済学において「リスク」を扱おうとする時，それをどのように定義するかが問題となる.「リスク」の定義を考える際に参考になるのが，ジェンセンの不等式（Jensen's Inequality）と呼ばれる次の定理である．

7.5 定理 (ジェンセンの不等式). $[a,b]$ で定義された連続関数 u について，次の二つの条件は同値である．

1. u は凹関数である．
2. $[a,b]$ に値をとる任意の確率変数 X について，$E[u(X)] \leq u\big(E[X]\big)$ が成立する．

7.5 定理の証明. まずは，条件 1 から条件 2 が導かれることを示そう．区間 $[a,b]$ に値をとる任意の確率変数 X について，$E[X] = \mu$ と置く．u が凹関数であると仮定すれば，ある実数 s が存在して

$$\forall x \in [a,b]: \quad u(x) - u(\mu) \leq s(x - \mu) \tag{7.15}$$

が成立する[33]．期待値の単調性により，

$$E\big[u(X) - u(\mu)\big] \leq E\big[s(X - \mu)\big]$$

であるが，さらに期待値の線形性から

$$\begin{aligned} E\big[u(X)\big] - u(\mu) &\leq s(E[X] - \mu) \\ &= 0 \end{aligned}$$

が言えるから，条件 2 を得る[34]．

[33] u が微分可能であれば，$s = u'(\mu)$ が (7.15) を満たす．微分可能性を仮定しない場合を含めた一般のケースについては，分離超平面定理を用いることで s の存在を示すことができる．

[34] 確率変数 X が離散的で，X が正の確率でとる値が $x_1, x_2, \ldots x_n$ のように有限個である場合には，凹関数の定義から直接これを示すこともできる．u が凹関数なら $\sum p(x_i)u(x_i) \leq u\big(\sum p(x_i)x_i\big)$ で，これは条件 2 に他ならない．

次に，条件 2 を仮定して条件 1 が成立することを背理法によって示す．仮に u が凹関数でないとすると，$x_1,\ x_2 \in [a,b]$ と $p \in (0,1)$ が存在して

$$pu(x_1) + (1-p)u(x_2) > u\big(px_1 + (1-p)x_2\big)$$

が成立する．しかしこれは，確率 p で x_1 をとり確率 $(1-p)$ で x_2 をとるような確率変数 X に対して

$$E[u(X)] > u\big(E[X]\big)$$

が成り立つことに他ならず，条件 2 に矛盾する．したがって条件 2 の下では，条件 1 が成立しなければならない．□

7.5 定理は，凹関数を効用関数として持つ個人は確率変数 X よりその期待値 $E[X]$ を確率 1 で得ることを好み，逆に確率変数よりもその期待値の方を好む個人の効用関数は凹関数でなければならないということを示している．一般に，実現値が確定していない確率変数 X を得る方が，期待値 $E[X]$ を確実に受け取るよりもリスクが大きいと考えられよう．したがって 7.5 定理の意味は，効用関数が凹関数の個人はリスクが大きいものを好まず，リスクが大きいものを好まない個人の効用関数は凹関数である，と言い換えることもできる．

このような感覚と整合的に「リスク」を定義する方法としては,「リスク」を「凹関数を効用関数とする個人が好まないもの」とみなすことが考えられる．そしてこの考え方を自然に拡張すれば，凹関数の効用関数を持つ個人が X よりも Y を好む時，X の方が Y よりも「リスクが大きい」と見なすことができよう．

それでは「凹関数を効用関数として持つ個人が好まない」という関係は，数学的にはどのように表現できるだろうか．そこで，対応する累積分布関数を $F,\ G$ とする二つの確率変数 $X,\ Y$ について，凹関数を効用関数として持つ任意の個人が Y よりも X を好む時，すなわち

$$\int_a^b u(x)dF(x) - \int_a^b u(x)G(x) \geq 0$$

が任意の凹関数 u で成り立つ時，F と G との関係をどのように表現できるかを考えてみる．ここでは特に，X の期待値と Y の期待値とが等しいケースを扱うことにする．

7.6 定理． $F(a) = G(a) = 0$ 及び $F(b) = G(b) = 1$ を満たし，なおかつ平均値の等しい二つの累積分布関数 $F,\ G$ について，次の 2 条件は同値である．

1. 任意の $x \in [a,b]$ について，$\displaystyle\int_a^x F(t)dt \leq \int_a^x G(t)dt$
2. $[a,b]$ 上で定義される任意の微分可能な凹関数 u について，$\displaystyle\int_a^b u(x)dF(x) - \int_a^b u(x)G(x) \geq 0$

7.6 定理を証明する準備として，まずは X の期待値と Y の期待値とが等しい状況をどのように表現できるかについて確認しておこう．累積分布関数を用いて期待値を表現すれば

$$E[X] = b - \int_a^b F(x)dx$$

$$E[Y] = b - \int_a^b G(x)dx$$

と書ける．したがって，期待値が等しいという表現を

$$E[X] = E[Y] \quad \Leftrightarrow \quad \int_a^b \Big\{F(x) - G(x)\Big\}dx = 0 \tag{7.16}$$

のように言い換えることができる[35].

また便宜上，関数 $\hat{F}(x)$ を

$$\hat{F}(x)=\frac{1}{\left(\int_a^b F(x)dx\right)}\int_a^x F(t)dt \tag{7.17}$$

のように定義しておく．この $\hat{F}$ について，$\hat{F}(a)=0$, $\hat{F}(b)=1$ が成り立つことは明らかであろう．さらに F が累積分布関数であることにより，$\hat{F}(x)$ の分母部分は正の定数で，なおかつ分子は x に関して連続非減少となるため，$\hat{F}(x)$ は非減少の連続関数である．つまり，$\hat{F}$ は累積分布関数の性質を満たす．したがって $[a,b]$ に値をとるある確率変数が存在し，$\hat{F}(x)$ はその確率変数に対応する累積分布関数であると考えられる．

またこの $\hat{F}(x)$ について，微積分学の基本定理から

$$\begin{aligned}\frac{d}{dx}\hat{F}(x) &= \frac{1}{\left(\int_a^b F(x)dx\right)}\frac{d}{dx}\int_a^x F(t)dt\\ &= \frac{1}{\left(\int_a^b F(x)dx\right)}F(x)\end{aligned}$$

が成立する．ここで f を $[a,b]$ で定義された微分可能な関数とすると，部分積分により

$$\begin{aligned}\int_a^b f(x)dF(x) &= F(b)f(b)-F(a)f(a)-\int_a^b F(x)f'(x)dx\\ &= f(b)-\int_a^b f'(x)F(x)dx\end{aligned}$$

と書けることに注意すれば，

$$\begin{aligned}\int_a^b f'(x)d\hat{F}(x) &= \int_a^b f'(x)\frac{d}{dx}\hat{F}(x)dx\\ &= \frac{1}{\left(\int_a^b F(x)dx\right)}\int_a^b f'(x)F(x)dx\\ &= \frac{1}{\left(\int_a^b F(x)dx\right)}\left(f(b)-\int_a^b f(x)dF(x)\right)\end{aligned} \tag{7.18}$$

と書くことができる．以上の結果と 7.4 定理を用いて，7.6 定理を証明しよう．

7.6 定理の証明. まずは条件 1 を仮定する．X の期待値と Y の期待値は等しいことから，(7.17) と合わせて

$$\int_a^b F(x)dx=\int_a^b G(x)dx \tag{7.19}$$

が成立する．この式は正の値をとるので，(7.19) の各辺を用いて条件 1 を割ることにより

$$\frac{1}{\left(\int_a^b F(x)dx\right)}\int_a^x F(t)dt \ \le\ \frac{1}{\left(\int_a^b G(x)dx\right)}\int_a^x G(t)dt$$

を得る．(7.18) の要領で $\hat{F}(x)$, $\hat{G}(x)$ を定義すれば，これは $\hat{F}(a)=\hat{G}(a)=0$ 及び $\hat{F}(b)=\hat{G}(b)=1$ を満たす二つの累積分布関数について，

$$\forall x\in[a,b]:\quad \hat{F}(x)\le\hat{G}(x) \tag{7.20}$$

[35] b が有限の場合には $\int_a^b F(x)dx=\int_a^b G(x)dx$ と同じことであるが，b が無限の場合に曖昧さが生じるために，本文中の表記を選んだ．

が成り立つことに他ならない．したがって 7.6 定義から，$\hat{F}(x)$ は $\hat{G}(x)$ よりも 1 次の確率優位にある．ここで $[a,b]$ 上で定義された微分可能な凹関数 u を考えると，$-u'$ は $[a,b]$ で単調非減少の関数となるから，7.4 定理により

$$\int_a^b \big(-u'(x)\big)d\hat{F}(x) \geq \int_a^b \big(-u'(x)\big)d\hat{G}(x)$$

すなわち

$$\int_a^b u'(x)d\hat{F}(x) \leq \int_a^b u'(x)d\hat{G}(x)$$

が言える．これを (7.18) を用いて書き直せば，

$$\frac{1}{\left(\int_a^b F(x)dx\right)}\left(u(b)-\int_a^b u(x)dF(x)\right) \;\leq\; \frac{1}{\left(\int_a^b G(x)dx\right)}\left(u(b)-\int_a^b u(x)dG(x)\right)$$

さらに (7.19) に注意して整理することにより，

$$\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x)$$

したがって条件 2 を得る．

逆に，条件 2 から条件 1 が導かれることを示す．$[a,b]$ 上で定義される任意の微分可能な凹関数 u について，

$$\int_a^b u(x)dF(x) - \int_a^b u(x)G(x) \geq 0$$

が成り立つならば，上の議論を逆にたどることにより，

$$\int_a^b u'(x)d\hat{F}(x) \leq \int_a^b u'(x)d\hat{G}(x) \tag{7.21}$$

が言える．ここで，$[a,b]$ で定義された任意の非増加な連続関数 v' に対して

$$v(x) = \int_a^x v'(t)dt$$

と置く．すると v は $[a,b]$ で連続微分可能な凹関数となるから（一対一参照），$u = v$ と置いても (7.21) は成立するはずで

$$\int_a^b v'(x)d\hat{F}(x) \leq \int_a^b v'(x)d\hat{G}(x)$$

すなわち

$$\int_a^b \big(-v'(x)\big)d\hat{F}(x) \geq \int_a^b \big(-v'(x)\big)d\hat{G}(x)$$

を得る．v' は任意の非増加関数であったから，$-v'$ は任意の非減少関数である．したがって 7.4 定理により

$$\forall x \in [a,b]: \quad \hat{F}(x) \leq \hat{G}(x) \tag{7.22}$$

が成立し，$\hat{F}$, $\hat{G}$ の定義と (7.19) に注意すれば，条件 1 が導かれる． □

7.6 定理により，任意の $x \in [a,b]$ について，$\displaystyle\int_a^x F(t)dt \leq \int_a^x G(t)dt$ が成り立つ時，そしてその時のみ，経済学に登場するリスク回避的な個人は G よりも F を好むという意味で，F の方が G よりも優位であると解釈できる．このような順序関係に対しては，以下のような定義が与えられている．

7.7 定義 (2 **次の確率優位**). 区間 $[a,b]$ に値をとる確率変数 X, Y について，対応する累積分布関数をそれぞれ，$F_X(x)$, $F_Y(x)$ と置く．任意の $x \in [a.b]$ に対して $\int_a^x F_X(t)dt \leq \int_a^x F_Y(t)dt$ が成り立つ時，X は Y よりも 2 次の確率優位にある（X second-order stochastically dominates Y）という．あるいは累積分布関数について，F_X は F_Y よりも 2 次の確率優位にあるという．

練習問題 7.5. $F(a) = G(a) = 0$ 及び $F(b) = G(b) = 1$ を満たし，なおかつ平均値の等しい二つの累積分布関数 F, G を考える．またこれらの累積分布関数について，

$$\forall x \in [a,b]: \quad \int_a^x F(t)dt \leq \int_a^x G(t)dt$$
$$\exists \hat{x} \in [a,b]: \quad \int_a^{\hat{x}} F(t)dt < \int_a^{\hat{x}} G(t)dt$$

が成立しているとする．このとき任意の $x \in [a,b]$ について $u''(x) > 0$ であるならば，G よりも F の方が期待効用は厳密に大きくなる，すなわち

$$\int_a^b u(x)dF(x) - \int_a^b u(x)dG(x) > 0$$

が成り立つことを示せ.（一般形が難しければ，連続型の場合のみ示せ.）

以上で説明した 2 次の確率優位の考え方は,「凹関数を持つ個人が好まない」＝「リスクが大きい」という関係を前提として，確率変数の間にリスクに基づく順序を与えている．しかし次の結果が示すように，まずリスクに関する大小関係を直接的に定義して，その大小関係が 2 次の確率優位の順序と対応していることを示すことも可能である．

確率変数 X, Y は，それぞれ累積分布関数 F_X, F_Y に従うとする．ここで，X とは独立[36]で期待値がゼロの確率変数 Z が存在し，$Y = X + Z$ と書けたとしよう．すると X と Z の独立性により

$$\begin{aligned} E[X+Z] &= E[X] + E[Z] \\ &= E[X] \end{aligned}$$

が成り立つから，Y の期待値と X の期待値は等しくなる．このような X と Y について,「Y は X の mean-preserving spread である」と言うことがある．そして，Y が X の mean-preserving spread である時に「Y の方が X よりもリスクが大きい」と考えれば，このリスクに関する大小関係は 2 次の確率優位の順序と対応するのである．

7.7 定理. 区間 $[a,b]$ に値をとる確率変数 X, Y について，次の 2 条件は同値である．

1. X が Y よりも 2 次の確率優位にある．
2. X とは独立で期待値がゼロの確率変数 Z が存在し，$Y = X + Z$ と書ける．

7.7 定理の証明は多少入り組んでいるので，ここでは例を示して雰囲気を味わうにとどめる．

7.3 例. コインを投げて表が出たら 1 ドル，裏が出たら -1 ドルもらえるギャンブルを考え，一度目に裏が出た場合にはもう一度だけギャンブルを繰り返すものとする．一度目のギャンブルで得られる金額によって定まる確率変数を X とし，最終的に得られる金額で定まる確率変数として Y を考える．ここで二度目のギャンブルによって定まる金額を Z とすれば，$Y = X + Z$ の関係にあり，なおかつ $E[Z] = 0$ である．さらに X と Z は独立なので，7.7 定理の条件 2 が成立していることが分かる．

練習問題 7.6. 7.3 例において，7.7 定理の条件 1 が成立していることを確かめよ．

[36] $\Pr(X \leq x,\ Z \leq z)$ が $\Pr(X \leq x)\Pr(Z \leq z)$ と等しいということ．すなわち，X が x 以下となることと Z が z 以下となることが同時に生じる確率は，$F_X(x)F_Z(z)$ で与えられる．

7.4 応用：簡単なポートフォリオ選択問題

以上の議論を踏まえて，この節では次のような問題を考えることにする．

$$\max_{x\in[0,w]} E\big[u(w+xR)\big] \tag{7.23}$$

ここで $w>0$ は正の定数，R は $[\alpha,\beta]$ に値をとる確率変数である．ただし，α と β は $\alpha<0<\beta$ の関係を満たすものとする．また，u は $(0,\infty)$ で定義された 2 回連続微分可能な関数で，$u'>0$ かつ $u''<0$ と仮定する．

この問題 (7.23) は，個人が投資量と貯蓄量の割合を選択することによって，来期の受取額から得られる期待効用を最大化する問題と解釈することができる．例えば貯蓄による 1 期間の固定金利が i，投資による 1 期間の変動利回りが Y の下で，合計 z 円の資金を投資と貯蓄に振り分けようとしているとしよう．変動利回り Y は，$[\alpha+i,\beta+i]$ に値をとる確率変数である．投資額を x 円と決めれば，貯蓄額は $z-x$ 円であるから，来期に受け取る金額は

$$(z-x)(1+i)+x(1+Y)=(1+i)z+(Y-i)x$$

である．ここで $w:=(1+i)z,\ R:=(Y-i)$ と置けば，x 円の投資から得られる来期の期待効用は $w+xR$ 円で，その期待効用は $E\big[u(w+xR)\big]$ と書ける．そしてこの期待効用を最大化しようとする問題が，(7.23) であると解釈できるのである．なお，w は z 円を全て貯蓄に回した場合の来期の受取額に等しい．

7.4.1 最適な投資量の決定

確率変数 R の累積分布関数を F とすれば，来期の期待効用は F を用いて

$$\begin{aligned} E\big[u(w+xR)\big] &= \int_\alpha^\beta u(w+xr)F(dr) \\ &= u(w+x\beta)-\int_\alpha^\beta F(r)\frac{\partial}{\partial r}u(w+xr)dr \end{aligned}$$

と表すことができる．右辺の第 1 項については，仮定により x に関して微分可能で

$$\frac{\partial}{\partial x}u(w+x\beta)=u'(w-x\beta)\beta$$

である．また第 2 項についても，微分と積分の順序交換の定理により x に関して微分可能で

$$\frac{\partial}{\partial x}\int_\alpha^\beta F(r)\frac{\partial}{\partial r}u(w+xr)dr=\int_\alpha^\beta \frac{\partial}{\partial x}\Big(F(r)\frac{\partial}{\partial r}u(w+xr)\Big)dr$$

が成立する．したがって，u が 2 回連続微分可能であることに注意すれば，期待効用は x に関して偏微分可能で

$$\begin{aligned} \frac{\partial}{\partial x}E\big[u(w+xR)\big] &= \frac{\partial}{\partial x}u(w+x\beta)-\int_\alpha^\beta \frac{\partial}{\partial x}\left(F(r)\frac{\partial}{\partial r}u(w+xr)\right)dr \\ &= \frac{\partial}{\partial x}u(w+x\beta)-\int_\alpha^\beta F(r)\frac{\partial}{\partial r}\left(\frac{\partial}{\partial x}u(w+xr)\right)dr \\ &= \int_\alpha^\beta \frac{\partial}{\partial x}u(w+xr)F(dr) \end{aligned}$$

である．これは期待値記号を用いて

$$\frac{\partial}{\partial x}E\big[u(w+xR)\big]=E\left[\frac{\partial}{\partial x}u(w+xR)\right]=E\big[u'(w+xR)R\big]$$

と書ける．また同様にして，期待効用の 2 回微分に関しても

$$\frac{\partial^2}{\partial x^2}E\big[u(w+xR)\big]=E\left[\frac{\partial^2}{\partial x^2}u(w+xR)\right]=E\big[u''(w+xR)R^2\big]$$

が成り立つことが確認できる．

練習問題 7.7. $E\big[u(w+xR)\big]$ を x の関数と見たとき，その 2 回微分を計算することによって凹関数であることを示せ．

問題 (7.23) に戻ろう．問題の 1 階の条件を求めると，λ, μ をラグランジュ乗数として

$$\begin{aligned} E\big[u'(w+xR)R\big] &= \lambda+\mu \qquad (7.24)\\ \lambda x &= 0 \qquad (7.25)\\ \mu(w-x) &= 0 \qquad (7.26)\end{aligned}$$

を得る．練習問題 7.7 の結果から目的関数は凹関数であることが確認されているので，最適な x の値はこの 1 階の条件によって定まる．

1 階の条件において $\lambda=\mu=0$，すなわち最適解が内点となるための条件を与えておく．内点解の必要十分条件は，R の期待値がゼロよりも大きく，なおかつ正の確率で R が負の値をとるというものである．つまり，問題 (7.23) の解を x^* とすると，$E[R]>0$ かつ $\Pr(R<0)>0$ ならば $x^*\in(0,w)$ となる．逆に R がこの条件を満たさない場合は端点解となり，$E[R]\leq 0$ であれば $x^*=0$，$\Pr(R\geq 0)=1$ であれば $x^*=w$ が最適な投資額と言える．これは次のようにして示される．

まずは，$E[R]\leq 0$ であれば $x^*=0$ となることを示す．$E\big[u(w+xR)\big]$ を x の関数と見なして

$$f(x)=E\big[u(w+xR)\big] \qquad (7.27)$$

と置けば，その 1 階の導関数は

$$f'(x)=E\big[u'(w+xR)R\big]$$

である．$f'(x)$ をさらに x で微分すれば

$$f''(x)=E\big[u''(w+xR)R^2\big]<0 \qquad (7.28)$$

となるから，$f'(x)$ は x に関して減少関数である．ただし，最後の不等号は $u''<0$ であることによる．ここでもし $E[R]\leq 0$ であるとすると

$$\begin{aligned} f'(0) &= E\big[u'(w)R\big]\\ &= u'(w)E[R]\\ &\leq 0\end{aligned}$$

である．$f'(x)$ が x に関して減少関数であることを考えれば，これは任意の $x\in[0,w]$ に対して $f'(x)\leq 0$ となることを意味する．つまり $f(x)=E\big[u(w+xR)\big]$ が $[0,w]$ の範囲で減少関数となるから，このとき最適な投資額は $x=0$ になる．

次に，$\Pr(R\geq 0)=1$ であれば $x^*=w$ となることを示そう．仮定により $u'>0$ であることを考えれば，$R\geq 0$ が確率 1 で成り立つ時，正の確率で実現する任意の値 r について

$$u'(w+wr)r\geq 0$$

となる．このとき期待値の単調性により $E\big[u'(w+wR)R\big]\geq 0$ が成り立つから，先ほどと同じように (7.27) で $f(x)$ を定義すれば，

$$f'(w)=E\big[u'(w+wR)R\big]\geq 0$$

が成立する．(7.28) により $f'(x)$ が x に関して減少関数であったから，これは任意の $x \in [0, w]$ に対して $f'(x) \le 0$ が成り立つことを意味する．したがって $f(x) = E\big[u(w + xR)\big]$ は $[0, w]$ の範囲で増加関数となり，最適な投資額は明らかに $x = w$ である．

最後に，$E[R] > 0$ かつ $\Pr(R < 0) > 0$ ならば $x^* \in (0, w)$ となることを確認する．$E[R] > 0$ により

$$f'(0) = u'(w)E[R] > 0$$

であり，また $\Pr(R < 0) > 0$ により

$$f'(w) = E\big[u'(w + wR)R\big] < 0$$

が言える．したがって，(7.28) により $f'(x)$ が x に関して減少関数であったことに注意すれば，唯一の $x^* \in (0, w)$ が存在して

$$f'(x^*) = E\big[u'(w + x^*R)R\big] = 0$$

が成立する．(7.24)，(7.25)，(7.26) において，$x = x^*$，$\lambda = \mu = 0$ と置けば 1 階の条件は満たされるので，この x^* が問題の解となる．

これ以降の議論では，$E[R] > 0$ かつ $\Pr(R < 0) > 0$ を仮定する．したがって問題の最適解は

$$E\big[u'(w + xR)R\big] = 0$$

によって決定されることになる．また，この方程式の解は 効用関数 u の形状と w の値，および確率変数 R に依存して定まるので，最適値を $x^* = \phi(u, w, R)$ と書くことにする．

7.4.2 収益の 1 次変化および 2 次変化に対応する投資量の変化

u や w，R が変化した際に，最適な投資額 $\phi(u, w, R)$ がどのように反応するかという比較静学の問題を考える．はじめに，確率変数が R から $\hat{R}$ へと変化した場合の最適な投資額の反応を調べてみる．すなわち，

$$E\big[u'(w + xR)R\big] = 0 \tag{7.29}$$

$$E\big[u'(w + x\hat{R})\hat{R}\big] = 0 \tag{7.30}$$

の解をそれぞれ $\phi(u, w, R)$，$\phi(u, w, \hat{R})$ と置き，$\phi(u, w, R)$ と $\phi(u, w, \hat{R})$ との大小関係を検討する．

このような比較静学においては，最初に大まかな予想を立ててみることが大切である．例えば，$\hat{R}$ が R よりも 1 次の確率優位にある場合，あるいは $\hat{R}$ が R よりも 2 次の確率優位にある場合を考えてみる．$\hat{R}$ が 1 次の確率優位にあるということは，$\hat{R}$ の方が投資の利回りが高いと見なせるので，おそらく $\phi(u, w, \hat{R})$ が $\phi(u, w, R)$ よりも大きいだろうと予想できる．また $\hat{R}$ が 2 次の確率優位にあるようなケースでは，平均的な利回りは等しいものの，$\hat{R}$ の方がリスクが小さいと解釈できる．したがってこちらのケースでも，おそらく $\phi(u, w, \hat{R})$ の方が $\phi(u, w, R)$ よりも大きくなるだろう．もちろん厳密な議論には追加的な条件や証明が必要であるが，式を変形させてゆく前にある程度の見通しを持っておいて損はない．

これ以降の議論は，x と $E\big[u'(w + xR)R\big]$ との関係をグラフにして描いてみると分かりやすいだろう．次の定理も，$E\big[u'(w + xR)R\big]$ や $E\big[u'(w + xR)R\big]$ のグラフを描いてみれば明らかである．

7.8 定理. (7.29)，(7.30) の解をそれぞれ $\phi(u, w, R)$，$\phi(u, w, \hat{R})$ とするとき，次の 2 条件は同値である．

1. $\phi(u, w, R) \le \phi(u, w, \hat{R})$
2. $E\Big[u'\Big(w + \phi(u, w, R)\hat{R}\Big)\hat{R}\Big] \ge 0$

7.8 定理の証明. $x^* = \phi(u, w, R)$, $\hat{x} = \phi(u, w, \hat{R})$ と置く．$\hat{x}$ は (7.30) の解であるから

$$E\big[u'(w + \hat{x}\hat{R})\hat{R}\big] = 0 \tag{7.31}$$

である．$E\big[u'(w + x\hat{R})\hat{R}\big]$ が x に関して減少関数であることに注意すれば，(7.31) の下で

$$\begin{aligned} x^* \leq \hat{x} \quad &\Leftrightarrow \quad E\big[u'(w + x^*\hat{R})\hat{R}\big] \geq E\big[u'(w + \hat{x}\hat{R})\hat{R}\big] \\ &\Leftrightarrow \quad E\big[u'(w + x^*\hat{R})\hat{R}\big] \geq 0 \end{aligned}$$

が成り立つ． □

7.8 定理から，$x^* = \phi(u, w, R)$ と $\hat{x} = \phi(u, w, \hat{R})$ との大小関係を見るためには $E\big[u'(w + x^*\hat{R})\hat{R}\big]$ の符号を見ればよいことが分かる．つまり $E\big[u'(w + x^*R)R\big]$ を R の関数と見なし，R が $\hat{R}$ へと変化した時の反応を調べればよいのである．

7.9 定理. $x^* = \phi(u, w, R)$, $\hat{x} = \phi(u, w, \hat{R})$ と置き，関数 g を $g(r) = u'(w + x^*r)r$ で定義する．この時，以下の二つの命題が成立する．

1. g が微分可能な増加関数である時，$\hat{R}$ が R より 1 次の確率優位であれば，$\hat{x} \geq x^*$ となる．
2. g が微分可能な凹関数である時，$\hat{R}$ が R より 2 次の確率優位であれば，$\hat{x} \geq x^*$ となる．

7.9 定理の証明. 関数 g の定義により

$$E\big[u'(w + x^*R)R\big] = E[g(R)] = 0 \tag{7.32}$$

$$E\big[u'(w + x^*\hat{R})\hat{R}\big] = E[g(\hat{R})] \tag{7.33}$$

と書けることに注意しておく．まずは命題 1 を示そう．g は増加関数であるから，$\hat{R}$ が R より 1 次の確率優位であれば，7.4 定理により

$$E\big[g(\hat{R})\big] \geq E\big[g(R)\big]$$

が成立する．(7.32)，(7.33) を用いれば

$$E\big[u'(w + x^*\hat{R})\hat{R}\big] \geq 0 \tag{7.34}$$

と書き直せるが，7.8 定理により，これは $\hat{x} \geq x^*$ に等しい．

次に命題 2 を示す．g は凹関数であるから，$\hat{R}$ が R より 2 次の確率優位であれば，7.6 定理により

$$E\big[g(\hat{R})\big] \geq E\big[g(R)\big]$$

が成立する．(7.32)，(7.33) を用いれば

$$E\big[u'(w + x^*\hat{R})\hat{R}\big] \geq 0 \tag{7.35}$$

と書き直せるが，7.8 定理により，これは $\hat{x} \geq x^*$ に等しい． □

7.4 注意. 一般に，たとえ u が単調増加の凹関数であっても g が増加関数（あるいは凹関数）であるかは分からない．g が増加関数かどうかは u' と u'' との関係で定まり，g が凹関数であるかどうかは u'' と u''' との関係で決まるからである．これは，凹関数を効用関数とするリスク回避的な個人であっても，R の変化に対する反応が同じとは限らないことを意味している．リスク回避的な個人にとって，このポートフォリオ問題は一見すると同質の問題のように見えるかもしれないが，実は各個人の u''' の形状に依存して答えが違ってくるのである．

練習問題 7.8. $\kappa,\ \gamma,\ \theta > 0$ として，効用関数 u が

$$u(x) = \kappa x - \frac{1}{2\gamma}x^2, \qquad u(x) = -e^{-\theta x}$$

で与えられるそれぞれの場合について，$g(r) = u'(w + x^* r)r$ が増加関数であるか，また凹関数であるか検討せよ．

7.4.3 効用関数の変化に対応する投資量の変化

次に，効用関数 u が変化した際の最適な投資量の反応を考える．変化の方向に関する見通しとしては，u が「より凹に」なった時に，最適な投資量は減少することが予想されよう．効用関数が「より凹に」変化するとは，その個人がよりリスク回避的になったと考えられるからである．以下の議論でこのことを示してゆくが，その準備として次の定理を確認しておく．

7.10 定理. $\alpha < 0 < \beta$ として，区間 $[\alpha, \beta]$ で定義された二つの関数 h と g を考える．h は正の微分可能関数で，任意の $z \in [\alpha, \beta]$ に対して $h'(z) \leq 0$ であるとする．一方 g は連続関数で，任意の $z \in [\alpha, 0)$ に対して $g(z) \leq 0$，任意の $z \in [0, \beta]$ に対して $g(z) \geq 0$ が成立し，なおかつ $\displaystyle\int_\alpha^\beta g(x)dx = 0$ を満たすものとする．このとき

$$\int_\alpha^\beta h(x)g(x)dx \leq 0$$

が成立する．

7.10 定理の証明. 関数 G を

$$G(z) = \int_\alpha^z g(x)dx$$

で定義すれば $G(\alpha) = 0$ で，仮定により $G(\beta) = 0$ である．また微積分学の基本定理により

$$\forall z \in [\alpha, \beta] : \quad G'(z) = g(z)$$

が成り立つから，G は g の原始関数の一つである．よって部分積分を用いて

$$\begin{aligned}
\int_\alpha^\beta h(x)g(x)dx &= \Big[h(z)G(z)\Big]_\alpha^\beta - \int_\alpha^\beta h'(z)G(z)dz \\
&= h(\alpha)G(\alpha) - h(\beta)G(\beta) - \int_\alpha^\beta h'(z)G(z)dz \\
&= -\int_\alpha^\beta h'(z)G(z)dz \qquad (7.36)
\end{aligned}$$

と書ける．仮定により任意の $z \in [\alpha, 0)$ に対して $G'(z) \leq 0$，任意の $z \in [0, \beta]$ に対して $G'(z) \geq 0$ であるから，G は $[\alpha, 0)$ で単調非増加，$[0, \beta]$ では単調非減少となる．したがって，

$$\forall z \in [\alpha, 0) : \quad G(z) \leq G(\alpha) = 0$$

$$\forall z \in [0, \beta] : \quad G(z) \leq G(\beta) = 0$$

であり，任意の $z \in [\alpha, \beta]$ に対して $G(z) \leq 0$ が成立する．さらに任意の $z \in [\alpha, \beta]$ で $h'(z) \leq 0$ であることを考えれば，任意の $z \in [\alpha, \beta]$ に対して $h'(z)G(z) \geq 0$ となる．このとき

$$\int_\alpha^\beta h'(z)G(z)dz \geq 0$$

であり，これを (7.36) と合わせれば定理の結果を得る． □

7.11 定理. 効用関数 u と v は，それぞれ $(0,\infty)$ で定義された 2 回連続微分可能な関数で，$u'>0,\ v'>0$ かつ $u''<0,\ v''<0$ を満たすものとする．この時ある 2 解微分可能な凹関数 f が存在して

$$\forall z\in(0,\infty):\quad v(z)=f\big(u(z)\big) \tag{7.37}$$

と書けるならば，これまでの前提を満たす任意の $R,\ w$ に対して

$$\phi(u,w,R)\geq\phi(v,w,R)$$

が成立する．

7.11 定理の証明. 以下では R を連続と仮定して証明するが，一般の場合も同様に議論可能である．w を任意に固定し，R の密度関数を $f_R(z)$ で表すことにする．$\phi(v,w,R)$ は効用関数 が v の時の最適値であるから，1 階の条件により

$$E\big[v'(w+\phi(v,w,R)R)R\big]=0$$

が成り立つ．ここで $E\big[v'(w+\phi(u,w,R)R)R\big]\leq 0$ が言えれば，

$$E\big[v'(w+\phi(u,w,R)R)R\big]\leq E\big[v'(w+\phi(v,w,R)R)R\big]$$

となるから，$E\big[v'(w+xR)R\big]$ が x の減少関数であったことに注意すれば，$\phi(v,w,R)\geq\phi(u,w,R)$ が得られる．したがって定理の結果を得るには，$E\big[v'(w+\phi(u,w,R)R)R\big]\leq 0$ を示せば十分である．

ここからは表記の簡単化のために，$x^*=\phi(u,w,R)$ と置くことにしよう．仮定により，ある 2 解微分可能な凹関数 f が存在して $v(z)=f\big(u(z)\big)$ と書けるから

$$v(w+xR)=f\big(u(w+xR)\big)$$

両辺を x に関して微分し，$x=x^*$ で評価すれば

$$v'(w+x^*R)R=f'\big(u(w+x^*R)\big)u'(w+x^*R)R \tag{7.38}$$

となる．したがって

$$\begin{aligned} E\big[v'(w+x^*R)R\big] &= E\big[f'\big(u(w+x^*R)\big)u'(w+x^*R)R\big] \\ &= \int_\alpha^\beta f'\big(u(w+x^*z)\big)u'(w+x^*z)zf_R(z)dz \end{aligned} \tag{7.39}$$

である．ここで

$$h(z)=f'\big(u(w+x^*z)\big) \tag{7.40}$$

$$g(z)=u'(w+x^*z)zf_R(z) \tag{7.41}$$

と置く．まず h について，$u'>0,\ v'>0$ に注意して (7.38) を変形すれば

$$f'\big(u(w+x^*z)\big)=\frac{v'(w+x^*R)}{u'(w+x^*R)}>0$$

を得るので，$h(z)>0$ を満たすことが分かる．また f は凹関数であるから $f''\leq 0$ で，$u'>0,\ x^*\in(0,w)$ と合わせて

$$h'(z)=f''\big(u(w+x^*z)\big)u'(w+x^*z)x^*\leq 0$$

が言える．次に g については，$u'>0,\ f_R\geq 0$ であることにより，任意の $z\in[\alpha,0)$ に対して $g(z)\leq 0$，任意の $z\in[0,\beta]$ に対して $g(z)\geq 0$ が成立する．さらに，x^* は効用関数が u の時の最適値であるから 1 階の条件により

$$E\big[u'(w+x^*R)R\big]=\int_\alpha^\beta u'(w+x^*z)zf_R(z)dz=0$$

でなければならず，したがって

$$\int_\alpha^\beta g(z)dz = \int_\alpha^\beta u'(w + x^* z) z f_R(z) dx = 0$$

が満たされることが確認できる．つまり，(7.40) と (7.41) で定義される $h(z)$, $g(z)$ は 7.10 定理の条件を全て満たすから，

$$\int_\alpha^\beta h(z)g(z)dz = \int_\alpha^\beta f'\big(u(w + x^* z)\big) u'(w + x^* z) z f_R(z) dz \leq 0$$

が成り立つ．最後にこれを (7.39) と合わせれば，$E\big[v'(w + x^* R)R\big] \leq 0$ を得る． □

以上では効用関数 u と確率変数 R に関する比較静学を見てきたが，u と R を固定して，w が変化した時に最適な投資量 $\phi(u, w, R)$ がどのように反応するかを見ることもできる．この比較静学に関しては，

$$r_A(z) = -\frac{u''(z)}{u'(z)}$$

で定義される絶対的危険回避度 $r_A(z)$ が，z に関して減少するかどうかが判断基準になる[37]．詳しくは，例えば MWG の本の 6.C を見よ．

7.5 確率と最適化

一般論はこの授業の守備範囲を超えるので，職探し（Job Search）モデルを例にとり，確率動学的最適化問題の意味と解法について解説する．

7.5.1 職探しモデル

現在の給与水準 w を所与として，$t = 0, 1, 2, \ldots$ の各期で新しい仕事を探すか現在の仕事に留まるかを選択する問題を考えよう．モデルの設定は以下の通りである．

- t 期において仕事を「探す」を選択した場合，t 期の収入はゼロとなり，次の $t+1$ 期の賃金は $[0, \bar{w}]$ 上に値をとる確率変数によって定まるとする．確率変数は連続型で，密度関数は f で表す．
- t 期において現在の仕事に「留まる」を選択した場合，t 期の収入は w となる．ただし，次の期には確率 θ で解雇されるものとする．つまり $(1-\theta)$ の確率で賃金は w のままであるが，$t+1$ 期の賃金は確率 θ でゼロをとる．
- 給与 w から得る効用を $u(w) \geq 0$ で表し，$u(0) = 0$ と正規化しておく．なお，効用関数 u は $u' > 0$ 及び $u'' < 0$ を満たすものとする．また割引率を $\delta \in (0, 1)$ と置く．
- 個人は期待効用の割引和，$\displaystyle\sum_{t=0}^{\infty} \delta^t E[u(W_t)]$ を最大化するように行動する．

[37] r_A(z) が z に関して減少関数となる時，u は絶対的危険回避度減少（decreasing absolute risk aversion: DARA）であるという．

7.5.2 最適解

Policy Function の形式や，Value Function が満たすべき条件を考えておく．基本的には不確実性がない場合と同じであるが，選択変数 y が $\{$探す, 留まる$\}$ に対応する二つの値しかとらないケースであることには注意が必要である．

- 個人は期ごとに給与水準 w を所与として「探す」か「留まる」かを決める．したがって Policy Function は $\phi(w) \in \{$探す, 留まる$\}$ と書ける．
- 給与水準 w から始めて最適な政策を選んだ時に得られる効用和，すなわち Value Function を $v(w)$ で表す．すると，この $v(w)$ は次のベルマン方程式を満たす．

$$v(w) = \max\left\{u(w) + \delta\big\{(1-\theta)v(w) + \theta v(0)\big\},\ \delta\int_0^{\bar{w}} v(z)f(z)dz\right\} \tag{7.42}$$

このように毎期の選択肢が二つだけの時は，次のテクニックが有効である．まず v がすでに求まっているふりをして，定数 A を

$$A = \delta\int_0^{\bar{w}} v(z)f(z)dz \tag{7.43}$$

と置く．効用が負の値をとることはなく，したがって $v(w) \geq 0$ であることに注意すれば，(7.42) により

$$\begin{aligned} v(0) &= \max\left\{u(0) + \delta\big\{(1-\theta)v(0) + \theta v(0)\big\},\ \delta\int_0^{\bar{w}} v(z)f(z)dz\right\} \\ &= \max\left\{\delta v(0),\ A\right\} \\ &= A \end{aligned}$$

である．$w = 0$ は職探しをする期の給与水準に等しいから，この式により，A は最適な政策に従って「探す」を選んだ時に将来に渡って得る効用和の期待値であると解釈できる．

一方で「留まる」を選んだ場合に得られる効用は，その期に得られる賃金の効用 $u(w)$ に次の期以降に得られる効用 $(1-\theta)v(w) + \theta v(0)$ を（割り引いて）加えたものである．確率 θ で失業するので，その場合には次期以降の効用和は $w = 0$ の時の Value，すなわち $v(0)$ となることに注意すること．$u(w)$ も $(1-\theta)v(w) + \theta v(0)$ も今期の賃金 w に依存して決まり，直感的には w が高ければ高いほど大きな値をとるはずである．実際，$v(w)$ が w の単調増加関数であることは容易に確認でき，さらに v が w について連続であることも示せる．

つまり，w がとても低いときには探した方がよく，十分に w が高くなれば留まったほうがよいと予想できる．そこで，ちょうどその分岐点になる水準 w^* を求めてみることにしよう．w^* は方程式

$$v(w) = u(w) + \delta\{(1-\theta)v(w) + \theta v(0)\} = A \tag{7.44}$$

を w に関して解くことで求まる．上で述べたように中央の項は w に関して増加関数であるから，この方程式には一意な解 w^* が存在して

$$\begin{aligned} \forall w \in [0, w^*):\quad & u(w) + \delta\{(1-\theta)v(w) + \theta v(0)\} < A \\ \forall w \in [w^*, \bar{w}]:\quad & u(w) + \delta\{(1-\theta)v(w) + \theta v(0)\} \geq A \end{aligned}$$

が成立する．つまり，$w \geq w^*$ ならば 留まった方がよく，$w < w^*$ ならば探しにいって効用 A を得た方がよいということである．そこでこの w^* を用いて

$$v^*(w) = \begin{cases} A & (\ w < w^* \text{ のとき }), \\ \frac{1}{1-\delta(1-\theta)}\Big(u(w) + \delta\theta A\Big) & (\ w \geq w^* \text{ のとき }). \end{cases} \tag{7.45}$$

のように v^* を定義してみる．w^* が (7.44) の解であることから

$$v(w^*) = u(w^*) + \delta\{(1-\theta)v(w^*) + \theta v(0)\} = A$$

が成り立つので，$u(w^*) + \delta\{(1-\theta)v(w^*) + \theta v(0)\} = A$ を，$v(w^*) = A$ と $v(0) = A$ を用いて整理することにより

$$u(w^*) = (1-\delta)A \tag{7.46}$$

を得る．これを (7.45) と合わせれば

$$\begin{aligned} v^*(w^*) &= \frac{1}{1-\delta(1-\theta)}\Big(u(w^*) + \delta\theta A\Big) \\ &= \frac{1}{1-\delta(1-\theta)}\Big((1-\delta)A + \delta\theta A\Big) \\ &= A \end{aligned}$$

となるから，v^* が $w = w^*$ において連続であることが確認できる．つまり関数 v^* のグラフは，$u(w) + \delta\{(1-\theta)v(w) + \theta A\}$ のグラフと A の値で水平な直線とで得られる形状になる．

(7.45) で定義される関数 v^* がベルマン方程式 (7.42) を満たすことは容易に確認できる．また v^* は有界であるから，$\delta^t v(w_t)$ は任意の実行可能プランについて $t \to \infty$ で 0 に収束する．したがって，v^* は最適政策に対応する Value Function である．v^* の形状から明らかに，Policy Function ϕ は

$$\phi(w) = \begin{cases} \text{探す} & (w < w^* \text{ のとき}), \\ \text{留まる} & (w \geq w^* \text{ のとき}). \end{cases}$$

という形をとり，このようにして決まる w^* を留保価格（reservation price ）あるいは留保賃金率（reservation wage）と呼ぶ．

最後に，定数 A と w^* を求める方法を考えよう．留保価格定数 A の定義 (7.43) と (7.45) から

$$\begin{aligned} A &= \delta\int_0^{\bar{w}} v(z)f(z)\,dz \\ &= \delta\left[\int_0^{w^*} v(z)f(z)\,dz + \int_{w^*}^{\bar{w}} v(z)f(z)\,dz\right] \\ &= \delta\left[\int_0^{w^*} Af(z)\,dz + \int_{w^*}^{\bar{w}} \frac{u(z)+\delta\theta A}{1-\delta(1-\theta)}f(z)\,dz\right] \\ &= \delta\left[A\int_0^{w^*} f(z)\,dz + \int_{w^*}^{\bar{w}} \frac{u(z)}{1-\delta(1-\theta)}f(z)\,dz + \frac{\delta\theta A}{1-\delta(1-\theta)}\int_{w^*}^{\bar{w}} f(z)\,dz\right] \\ &= \delta\left[AF(w^*) + \frac{1}{1-\delta(1-\theta)}\int_{w^*}^{\bar{w}} u(z)f(z)\,dz + \frac{\delta\theta A}{1-\delta(1-\theta)}\big(1-F(w^*)\big)\right] \end{aligned} \tag{7.47}$$

を得る．ただしここで

$$F(w) = \int_0^w f(z)\,dz \qquad \left(\text{したがって} \quad \int_w^{\bar{w}} f(z)\,dz = 1 - F(w)\right)$$

と置いている．(7.47) は

$$\begin{aligned} \frac{\delta}{1-\delta(1-\theta)}\int_{w^*}^{\bar{w}} u(z)f(z)\,dz &= \left(1-\delta F(w^*) - \frac{\delta^2\theta}{1-\delta(1-\theta)}\big(1-F(w^*)\big)\right)A \\ &= \frac{1+\theta\delta-\delta F(w^*)}{1-\delta(1-\theta)}(1-\delta)A \end{aligned}$$

のように整理できるから，さらに (7.46) を利用して A を消去すれば

$$\begin{aligned}\frac{\delta}{1+\delta\theta-\delta F(w^*)}\int_{w^*}^{\bar{w}} u(z)\,f(z)\,dz &= (1-\delta)\,A \\ &= u(w^*) \qquad (7.48)\end{aligned}$$

を得る．関数 f, F, u はいずれも既知の関数であり，δ と θ は既知のパラメタであるから，これは w^* に関する方程式である．汗を流して工夫すれば，原理的には解けるはずである．

7.4 例. 上の問題で $\bar{w}=1$ とし，確率変数が $[0,1]$ 上の一様分布に従う場合を考える．このとき任意の $z\in[0,1]$ に対して

$$f(z)=1, \qquad F(z)=z$$

である．また，効用関数が $u(w)=w$ で与えられているとすれば

$$\begin{aligned}\int_{w^*}^{\bar{w}} u(z)\,f(z)\,dz &= \int_{w^*}^{1} z dz \\ &= \frac{1}{2}\left(1-(w^*)^2\right)\end{aligned}$$

と書ける．したがってこの場合の方程式 (7.48) は

$$\frac{\delta}{1+\delta\theta-\delta w^*}\frac{1}{2}\left(1-(w^*)^2\right)=w^*$$

となり，これを解くことで

$$w^*=\frac{1}{\delta}\left(\theta\delta+\sqrt{2\theta\delta-\delta^2+\theta^2\delta^2+1}+1\right)$$

を得る．