

# 収束

黒木玄

2018-04-18

- Copyright 2018 Gen Kuroki
- License: MIT <https://opensource.org/licenses/MIT> (<https://opensource.org/licenses/MIT>)

このファイルは次の場所できれいに閲覧できる:

- <http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/01%20convergence.ipynb>  
<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/01%20convergence.ipynb>

このファイルは [Julia Box](https://juliabox.com) (<https://juliabox.com>) で利用できる。

自分のパソコンに [Julia言語](https://julialang.org/) (<https://julialang.org/>) をインストールしたい場合には

- [WindowsへのJulia言語のインストール](http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f) (<http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f>)

を参照せよ。

論理的に完璧な説明をするつもりはない。細部のいい加減な部分は自分で訂正・修正せよ。

## 目次

### 1 収束について

#### 1.1 収束の定義

- 1.1.1 数列の収束の定義
- 1.1.2 級数の収束の定義
- 1.1.3 関数の値の収束の定義

#### 1.2 収束の判定法の基本

- 1.2.1 はさみうち
- 1.2.2 二項定理の復習
- 1.2.3 多項式関数より指数関数の方が速く増加すること
- 1.2.4 指数関数より階乗の方が速く増加すること

#### 1.3 まとめ

### 2 上限と下限と上極限と下極限

- 2.1 上限と下限
- 2.2 上極限と下極限

```
In [1]: 1  using Plots
2  gr(); ENV["PLOTS_TEST"] = "true"
3  #clibrary(:colorset)
4  clibrary(:misc)
5
6  ▼ function pngplot(P...; kwargs...)
7    sleep(0.1)
8    pngfile = tempname() * ".png"
9    savefig(plot(P...; kwargs...), pngfile)
10   showimg("image/png", pngfile)
11 end
12 pngplot(; kwargs...) = pngplot(plot!(; kwargs...))
13
14 ▼ showing(mime, fn) = open(fn) do f
15    base64 = base64encode(f)
16    display("text/html", """""")
17 end
18
19 using SymPy
20 #sympy[:init_printing](order="lex") # default
21 #sympy[:init_printing](order="rev-lex")
22
23 using SpecialFunctions
24 using QuadGK
```

## 1 収束について

### 1.1 収束の定義

#### 1.1.1 数列の収束の定義

数列  $a_n$  が  $n \rightarrow \infty$  で  $\alpha$  に収束するとは、任意の  $\varepsilon > 0$  に対して、ある番号  $N$  が存在して、 $N$  以降のすべての番号  $n$  について  $|a_n - \alpha| < \varepsilon$  が成立することである。

$a_n$  が  $\alpha$  に収束するとき、 $a_n$  の極限の値を  $\lim_{n \rightarrow \infty} a_n = \alpha$  と定義する。収束しないときには極限の値は定義されない。

### 1.1.2 級数の収束の定義

(無限)級数  $\sum_{n=1}^{\infty} a_n$  の収束は数列  $s_n = \sum_{k=1}^n a_k$  の収束で定義される。収束するとき  $\sum_{n=1}^{\infty} a_n = \lim_{n \rightarrow \infty} \sum_{k=1}^n a_k$  と書く。

### 1.1.3 関数の値の収束の定義

関数の  $x$  における値  $f(x)$  が  $x \rightarrow a$  で  $\alpha$  に収束するとは、任意の  $\varepsilon > 0$  に対して、ある  $\delta > 0$  が存在して、 $0 < |x - a| < \delta$  を満たすすべての  $x$  について  $|f(x) - \alpha| < \varepsilon$  が成立することである。

$f(x)$  が  $x \rightarrow a$  で  $\alpha$  に収束するとき、その極限の値を  $\lim_{x \rightarrow a} f(x) = \alpha$  と定義する。収束しないときには極限の値は定義されない。

**注意:** 上の定義における  $|f(x) - \alpha| < \varepsilon$  を  $|f(x) - \alpha| \leq \varepsilon$  に置き換えるても同値な条件になる。□

**問題:** これはなぜか? □

## 1.2 収束の判定法の基本

### 1.2.1 はさみうち

**定理:**  $A_n \leq a_n \leq B_n$  かつ  $A_n$  と  $B_n$  が  $\alpha$  に収束するならば  $a_n$  も  $\alpha$  に収束する。□

**定理:**  $A_n \leq a_n$  かつ  $A_n \rightarrow \infty$  ならば  $a_n \rightarrow \infty$  となる。□

### 1.2.2 二項定理の復習

二項係数を

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}$$

と定義する。例えば

$$\binom{n}{0} = 1, \quad \binom{n}{1} = n, \quad \binom{n}{2} = \frac{n(n-1)}{2}, \quad \binom{n}{3} = \frac{n(n-1)(n-2)}{6}.$$

$n$  も  $k$  も非負の整数のとき  $\binom{n}{k}$  は  $n$  個から  $k$  個選び出すときの場合の数に一致している。高校数学では  ${}_n C_k$  のように書いているようだが、一般には  $\binom{n}{k}$  と書くことの方が多いように感じられるので、 $\binom{n}{k}$  の方を使用する。(上の定義に従えば、 $k$  が非負の整数でありさえすれば、 $n$  が整数でなくとも  $\binom{n}{k}$  が定義されていることに注意せよ。この事実は二項展開を考えるときに必要になる。)

以上の記号法のもとで、二項定理は

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k \tag{*}$$

と書ける。高校数学で証明を習っているはずだが、忘れている人は二項定理を証明しておくこと。

**考え方:** 解析学では不等式を扱いたい。しかし、二項定理のような良い等式は良い不等式を生み出すために非常に役に立つ。解析学では非自明な等式は非自明な不等式を作る材料になる。次の節の証明を見よ！

**二項定理の証明1:**  $(x+y)^n$  を展開した結果の  $x^{n-k} y^k$  の係数は  $n - k$  個の  $x$  と  $k$  個の  $y$  の並べ方の個数に等しい。その個数は全部で  $n$  個並べるうち  $y$  を置く  $k$  個の場所の選び方の個数に等しい。それは

$$\frac{n!}{k!(n-k)!} = \binom{n}{k}$$

に等しい。□

**二項定理の証明2:**  $n$  に関する帰納法を使う。 $n = 0$  の場合に(\*)は成立している。(\*)が  $n$  について成立していると仮定する。そのとき、 $k$  が負の整数のとき  $\binom{n}{k} = 0$  と約束しておくと、

$$\begin{aligned}
\binom{n}{k-1} + \binom{n}{k} &= \frac{n(n-1)\cdots(n-k+2)}{(k-1)!} + \frac{n(n-1)\cdots(n-k+2)(n-k+1)}{k!} \\
&= \frac{n(n-1)\cdots(n-k+2)}{k!} \times (k + (n - k + 1)) \\
&= \frac{n(n-1)\cdots(n-k+2)}{k!} \times (n + 1) \\
&= \binom{n+1}{k}.
\end{aligned}$$

ゆえに  $n$  に関する(\*)を使うと,

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)(x+y)^n = (x+y) \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k \\
&= \sum_{k=0}^n \binom{n}{k} x^{n-k+1} y^k + \sum_{k=0}^n \binom{n}{k} x^{n-k} y^{k+1} \\
&= \sum_{k=0}^{n+1} \binom{n}{k} x^{n-k+1} y^k + \sum_{k=0}^{n+1} \binom{n}{k-1} x^{n-k+1} y^k \\
&= \sum_{k=0}^{n+1} \left( \binom{n}{k} + \binom{n}{k-1} \right) x^{n-k+1} y^k \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} x^{n-k+1} y^k
\end{aligned}$$

と  $n$  を  $n+1$  で置き換えた場合の(\*)も成立することがわかる. これで示すべきことが示された.  $\square$

### 1.2.3 多項式函数より指數函数の方が速く増加すること

**定理:**  $a > 1$  ならば  $n \rightarrow \infty$  のとき  $\frac{n^l}{a^n} \rightarrow 0$ .  $\square$

**証明:** 二項定理を使う.

$L$  は  $l$  より大きな正の整数であるとし,  $n$  は  $L$  以上であると仮定する.

そのとき特に  $n \rightarrow \infty$  のとき  $n^l/n^L \rightarrow 0$  が成立している.

$a > 1$  なので  $a = 1 + \alpha$ ,  $\alpha > 0$  と書ける. 二項定理より

$$\begin{aligned}
a^n &= (1 + \alpha)^n = \sum_{k=0}^n \binom{n}{k} \alpha^k = \sum_{k=0}^n \frac{n(n-1)\cdots(n-k+1)}{k!} \alpha^k \\
&\geq \frac{n(n-1)\cdots(n-L+1)}{L!} \alpha^L = n^L \frac{(1-1/n)\cdots(1-(L-1)/n)\alpha^L}{L!}.
\end{aligned}$$

ゆえに,  $n \rightarrow \infty$  のとき

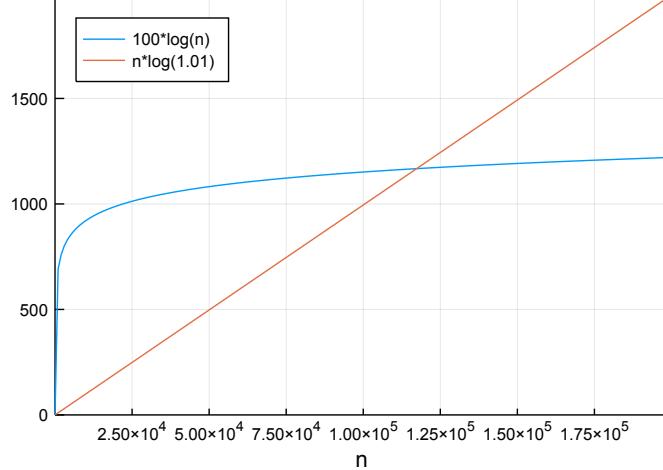
$$0 \leq \frac{n^l}{a^n} \leq \frac{n^l}{n^L} \times \frac{L!}{(1-1/n)\cdots(1-(L-1)/n)\alpha^L} \rightarrow 0 \times \frac{L!}{\alpha^L} = 0. \quad \square$$

**問題:**  $a = 1.01$ ,  $l = 100$  のとき,  $n^l$  と  $a^n$  の大きさを比較するためのグラフを描いてみよ.

**解答例:** 直接  $n^l$  と  $a^n$  を扱うと数値が巨大になり過ぎるので, それらの対数を比較することにする.

```
In [2]: 1 a = 1.01
2 l = 100
3 n = 1:10^3:2*10^5
4 f(n) = l*log(n)
5 g(n) = n*log(a)
6 plot(size=(500, 350), legend=:topleft, xlabel="n")
7 plot!(n, f.(n), label="$l*\log(n)")
8 plot!(n, g.(n), label="n*\log($a)")
```

Out[2]:



**問題:**  $a > 1$  ならば,  $x > 0$  が連続的に幾らでも大きくなるとき  $\frac{x^l}{a^x} \rightarrow 0$  となることを示せ.

**証明:** 正の整数  $n$  について  $n \leq x \leq n+1$  のとき,  $x^l$  は  $n^l$  と  $(n+1)^l$  のあいだにあり,  $a^x$  は  $a^n$  以上になる. ゆえに,  $x \rightarrow \infty$  のとき,  $n \rightarrow \infty$  となり,

$$0 \leq \frac{x^l}{a^x} \leq \frac{\max\{n^l, (n+1)^l\}}{a^n} \leq \max \left\{ \frac{n^l}{a^n}, a \frac{(n+1)^l}{a^{n+1}} \right\} \rightarrow 0.$$

最後に上方の定理を使った. これで示すべきことが示された.  $\square$

**問題:**  $x > 0$  について,  $x \rightarrow 0$  のとき  $x \log x \rightarrow 0$  となることを示せ.

**証明:**  $x > 0$  なので  $x = e^{-t}$ ,  $t \in \mathbb{R}$  と書ける.  $x \rightarrow 0$  のとき  $t \rightarrow \infty$  となるので,

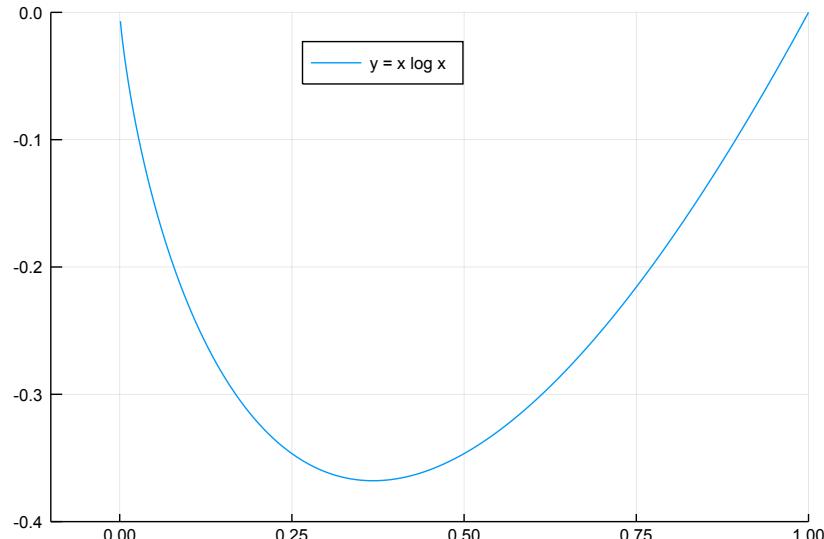
$$x \log x = e^{-t} \log e^{-t} = -\frac{t}{e^t} \rightarrow 0.$$

最後の上の問題の結果を使った.  $\square$

**注意:** 統計学に出て来る計算では  $0 \log 0 = 0$  と約束しておくことが適切な場合が多い.  $\square$

```
In [3]: 1 f(x) = x*log(x)
2 x = 0.001:0.001:1
3 plot(x, f.(x), xlim=(-0.1,1), ylim=(-0.4,0), label="y = x log x", legend=:top)
```

Out[3]:



#### 1.2.4 指数函数より階乗の方が速く増加すること

**定理:**  $a > 0$  と仮定する. そのとき  $n \rightarrow \infty$  ならば  $\frac{a^n}{n!} \rightarrow 0$  が成立する.

**証明:**  $N$  を十分大きくすると  $\frac{a}{N+1} < 1$  となる. そして,  $n > N$  とすると,

$$\frac{a^n}{n!} = \frac{a^N}{N!} \frac{a}{N+1} \frac{a}{N+2} \cdots \frac{a}{n} \leq \frac{a^N}{N!} \left( \frac{a}{N+1} \right)^{n-N} \rightarrow 0 \quad (n \rightarrow \infty). \quad \square$$

**問題:**  $a > 1$  と仮定する. そのとき  $n \rightarrow \infty$  ならば  $\frac{n!}{a^n} \rightarrow 0$  が成立することを示せ.

**解答例:** 指数函数は多項式函数より速く増加するので  $\frac{n!}{a^n} \rightarrow 0$ . ゆえにある番号  $N$  が存在して  $n > N$  ならば  $\frac{n!}{a^n} \geq \frac{1}{2}$  となる. したがって,  $n > N$  のとき,

$$\frac{n!}{a^{n^2}} = \frac{n!}{(a^n)n} = \frac{N!}{(a^n)^N} \frac{N+1}{a^n} \frac{N+2}{a^n} \cdots \frac{n}{a^n} \leq \frac{N!}{(a^N)^n} \left( \frac{1}{2} \right)^{n-N} \rightarrow 0 \quad (n \rightarrow \infty). \quad \square$$

### 1.3 まとめ

$n \rightarrow \infty$  のとき  $\frac{a_n}{b_n} \rightarrow 0$  となることを  $a_n < b_n$  と書くことにする.  $a > 1$  のとき

$$1 < \cdots < \log \log n < \log n < n < n^2 < \cdots < a^n < n! < a^{n^2} < \cdots$$

右に行くほど  $n \rightarrow \infty$  で速く増加する.

**例:** Stirlingの近似公式より,

$$n! = n^n \frac{1}{e^n} \sqrt{n} \sqrt{2\pi} (1 + \varepsilon_n) \quad (\varepsilon_n \rightarrow 0)$$

が成立することが知られている. 右辺に登場する  $n^n$ ,  $e^n$ ,  $\sqrt{n}$ ,  $\sqrt{2\pi}$  は左のものほど  $n \rightarrow \infty$  で速く増加する.

$\delta_n = \log(1 + \varepsilon_n)$  とおき, Stirlingの近似公式の両辺の対数を取ると,

$$\log n! = n \log n - n + \frac{1}{2} \log n + \frac{1}{2} \log 2\pi + \delta_n \quad (\delta_n \rightarrow 0).$$

右辺に登場する  $n \log n$ ,  $n$ ,  $\frac{1}{2} \log n$ ,  $\frac{1}{2} \log \sqrt{2\pi}$ ,  $\delta_n$  は左のものほど  $n \rightarrow \infty$  で速く増加する.  $\square$

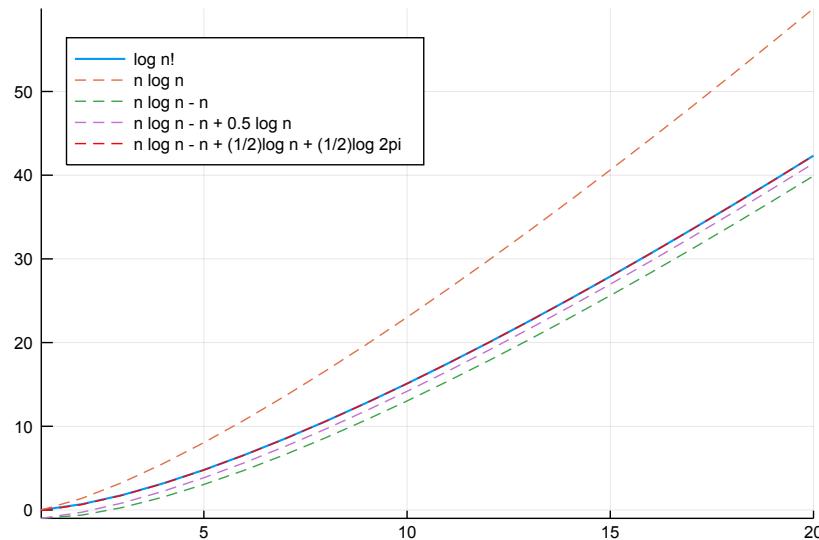
次のセルで  $\log n!$  がどのように近似されるかをプロットしよう.

```
In [4]: 1 lfact(n) = lgamma(n+1) # = log n!
2 f1(n) = n*log(n)
3 f2(n) = n*log(n) - n
4 f3(n) = n*log(n) - n + 0.5*log(n)
5 f4(n) = n*log(n) - n + 0.5*log(n) + 0.5*log(2pi)
6
7 ▼ function plot_logfactorial(n)
8     plot(legend=:topleft)
9     plot!(n, lfact.(n), label="log n!", lw=1.5)
10    plot!(n, f1.(n), label="n log n", ls=:dash)
11    plot!(n, f2.(n), label="n log n - n", ls=:dash)
12    plot!(n, f3.(n), label="n log n - n + 0.5 log n", ls=:dash)
13    plot!(n, f4.(n), label="n log n - n + (1/2)log 2pi", color=:red, ls=:dash)
14 end
```

Out[4]: plot\_logfactorial (generic function with 1 method)

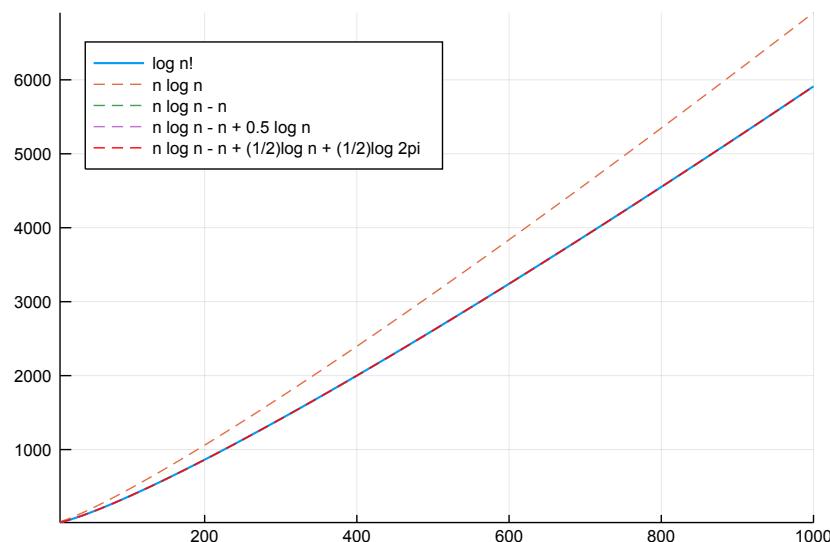
```
In [5]: 1 plot_logfactorial(1:20)
```

Out[5]:



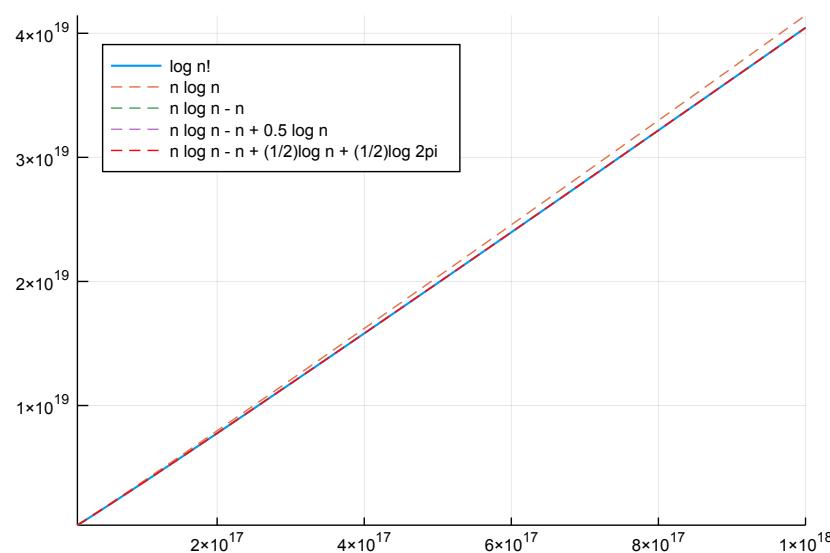
```
In [6]: 1 plot_logfactorial(10:10:1000)
```

Out[6]:



```
In [7]: 1 plot_logfactorial(10^16:10^16:10^18)
```

Out[7]:



すぐ上のプロットを見れば、n がこのように極めて大きな値であれば  $\log n!$  の  $\log n^n = n \log n$  による近似もそう悪くないことがわかる。

## 2 上限と下限と上極限と下極限

### 2.1 上限と下限

$X$  は空でない実数の集合であると仮定する。

実数  $M$  で任意の  $x \in X$  について  $x \leq M$  を満たすものを実数の集合  $X$  の**上界**(upper bound)と呼ぶ。実数  $m$  で任意の  $x \in X$  について  $m \leq x$  を満たすものを  $X$  の**下界**(lower bound)と呼ぶ。

$X$  の上界が1つ以上存在するとき、 $X$  は**上に有界**(bounded from above)であると言い、 $X$  の下界が1つ以上存在するとき、 $X$  は**下に有界**(bounded from below)であると言う。 $X$  が上に有界かつ下に有界であるとき、 $X$  は**有界**(bounded)であると言う。

$X$  の上に有界ば  $X$  の上界全体の集合の中に最小値が存在する。その最小値を  $X$  の**最小上界**(least upper bound)もしくは**上限**(supremum)と呼び。同様に、 $X$  が下に有界ならば  $X$  の下界全体の集合の中に最大値が存在する。その最大値を  $X$  の**最大下界**(greatest lower bound)または**下限**(infimum)と呼ぶ。 $X$  の上限と下限をそれぞれ

$$\sup X, \quad \inf X$$

と書く。実数値函数  $f(x)$  による集合  $X$  の像  $f(X) = \{f(x) \mid x \in X\}$  の上限と下限をそれぞれ

$$\sup f(X) = \sup_{x \in X} f(x), \quad \inf f(X) = \inf_{x \in X} f(x)$$

と書く。

$X$  が上に有界でないときには  $X$  は(有限な)上限を持たないが、 $\sup X = \infty$  と書くことがある。同様に  $X$  が下に有界でないときには  $X$  は(有限な)下限を持たないが、 $\inf X = -\infty$  と書くことがある。

$X$  に最大値  $\max X$  が存在するとき、 $\sup X = \max X$  となる。同様に  $X$  に最小値  $\min X$  が存在するとき、 $\inf X = \min X$  となる。

**上限と下限の便利なところ:**  $X$  に最大値や最小値が存在しなくとも、上限と下限は値として  $\pm\infty$  を許せば常に存在する。上限と下限は、最大値や最小値が存在しないときに、それらに代わるものとして用いることができる場合が結構ある。□

**例:** 開区間  $X = (0, 1)$  には最小値も最大値も存在しない。 $\inf X = 0, \sup X = 1$  となる。□

**例:**  $X = \{1, 1/2, 1/3, \dots\}$  には最大値 1 は存在するが、最小値は存在しない。このとき、

$$\sup X = \sup_{n=1,2,3,\dots} \frac{1}{n} = \max_{n=1,2,3,\dots} \frac{1}{n} = 1, \quad \inf X = \inf_{n=1,2,3,\dots} \frac{1}{n} = 0. \quad \square$$

**例:**  $X = \{1, 2, 3, \dots\}$  は上に有界でないので有限の上限を持たない。下限は最小値の 1 になる：

$$\sup \{1, 2, 3, \dots\} = \infty, \quad \inf \{1, 2, 3, \dots\} = \min \{1, 2, 3, \dots\} = 1.$$

### 2.2 上極限と下極限

実数の集合  $X$  とその部分集合  $Y$  を考える。 $X$  の上界は  $Y$  の上界にもなるので、 $X$  の上界の最小値は  $Y$  の上界の最小値以上になる。すなわち、 $\sup X \geq \sup Y$  となる(これは図を描いて考えれば明らかだろう)。部分集合の上限はもとの集合の上限以下になる。同様に部分集合の下限はもとの集合の下限以上になる。

したがって、実数列  $a_n$  に対して、数列  $\alpha_n = \inf_{k \geq n} a_k$  は単調増加し、数列  $\beta_n = \sup_{k \geq n} a_k$  は単調減少する。ゆえに  $\pm\infty$  も収束先に含めれば、それらの数列は常に収束することになる。それぞれの収束先を次のように書く。

$$\liminf_{n \rightarrow \infty} a_n = \liminf_{n \rightarrow \infty} \inf_{k \geq n} a_k, \quad \limsup_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} \sup_{k \geq n} a_k.$$

これらをそれぞれ**上極限**(limit superior)、**下極限**(limit inferior)と呼ぶ。

それを  $\liminf a_n, \limsup a_n$  と略して書くこともある。

自明に  $\inf_{k \geq n} a_k \leq \sup_{k \geq n} a_k$  が成立しているので、

$$\liminf_{n \rightarrow \infty} a_n \leq \limsup_{n \rightarrow \infty} a_n$$

が成立している。

**定理:** 実数列  $a_n$  について

$$-\infty < \liminf_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} a_n < \infty$$

が成立しているならば、実数列  $a_n$  は収束して、

$$\lim_{n \rightarrow \infty} a_n = \liminf_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} a_n.$$

**証明:**  $\alpha = \liminf_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} a_n$  とおき、任意に  $\varepsilon > 0$  と取って固定する。 $\alpha = \liminf_{n \rightarrow \infty} a_n$  より、 $\inf_{j \geq n} a_j$  が  $n$  について単調増加することに注意すれば、ある番号  $N'$  が存在して、

$$\alpha - \varepsilon \leq \inf_{j \geq n} a_j \leq \alpha \quad (n \geq N').$$

同様に,  $\sup_{j \geq n} a_j$  が  $n$  について単調減少することに注意すれば, ある番号  $N''$  が存在して,

$$\alpha \leq \sup_{j \geq n} a_j \leq \alpha + \varepsilon \quad (n \geq N'').$$

ゆえに  $n = \max\{N', N''\}$  のとき,  $k \geq n$  ならば

$$\alpha - \varepsilon \leq \inf_{j \geq n} a_k \leq a_k \leq \sup_{j \geq n} a_j \leq \alpha + \varepsilon.$$

これは  $k \geq n$  のとき,  $|a_k - \alpha| \leq \varepsilon$  が成立することを意味する. したがって, 数列の収束の定義より, 数列  $a_n$  は  $\alpha$  に収束する. これで示すべきことが示された.  $\square$

**上極限と下極限の便利な点:**  $\lim a_n$  と違って,  $\liminf a_n$  と  $\limsup a_n$  は収束先として  $\pm\infty$  を含めれば常に収束しているので収束性を確認せずに気軽に使うことができる. そして上の定理より,  $\liminf a_n$  と  $\limsup a_n$  が一致していれば,  $a_n$  は収束していて収束先はそれらと同じになる.

だから, 以下のような使われ方をすることが多い. 実数列  $a_n$  に対して, 任意の  $\varepsilon > 0$  に対して, それに依存して決まる別の数列  $A_n^{(\varepsilon)}$  と  $B_n^{(\varepsilon)}$  で同じ値  $\alpha$  に収束するものが存在して, さらに

$$A_n^{(\varepsilon)} - \varepsilon \leq a_n \leq B_n^{(\varepsilon)} + \varepsilon$$

が成立していると仮定する. このとき,  $n \rightarrow \infty$  とすると,

$$\alpha - \varepsilon \leq \liminf_{n \rightarrow \infty} a_n \leq \limsup_{n \rightarrow \infty} a_n \leq \alpha + \varepsilon.$$

$\varepsilon > 0$  は幾らでも小さくできるので,

$$\liminf_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} a_n = \alpha$$

となる. ゆえに数列  $a_n$  は  $\alpha$  に収束する:

$$\lim_{n \rightarrow \infty} a_n = \alpha.$$

「 $n \rightarrow \infty$  とすると」の段階では数列  $a_n$  が収束するかどうかわかっていないので,  $\lim$  ではなく,  $\liminf$  と  $\limsup$  を使わなければいけない.  $\square$

**問題:** 以下の実数列の上極限を求めよ.

$$(1) a_n = 10/n + (-1)^n.$$

$$(2) b_n = 10/n + \sin(4n).$$

$\pi$  が無理数であることを認めて使ってよい.

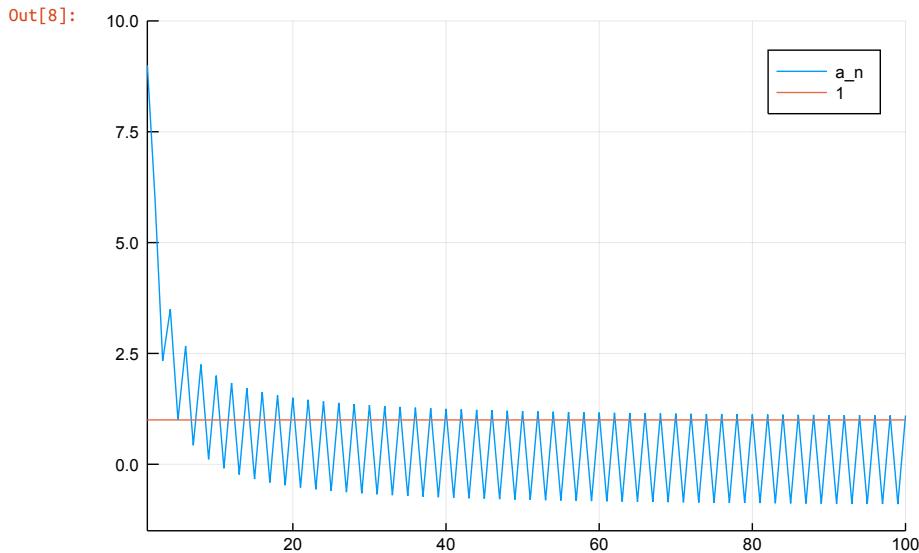
**解答例:** (1)  $\limsup a_n = \limsup(10/n + (-1)^n) = 1$ .

(2)  $\pi$  は無理数なので  $a = 4/(2\pi)$  も無理数である. そのとき,  $\sin(4n) = \sin(2\pi na)$ . 次の問題より, 任意の  $k > 0$  に対して, ある正の整数  $n_k$  で  $n_k a$  の小数点以下の部分と  $1/4$  の距離が  $1/k$  以下になるものが存在する. そのとき,  $n_k a$  の小数点以下の部分は  $1/4$  に収束するので,  $k \rightarrow \infty$  で  $\sin(2\pi n_k a) \rightarrow \sin(2\pi(1/4)) = 1$  となる. これより,  $\limsup b_n = 1$  となることがわかる.  $\square$

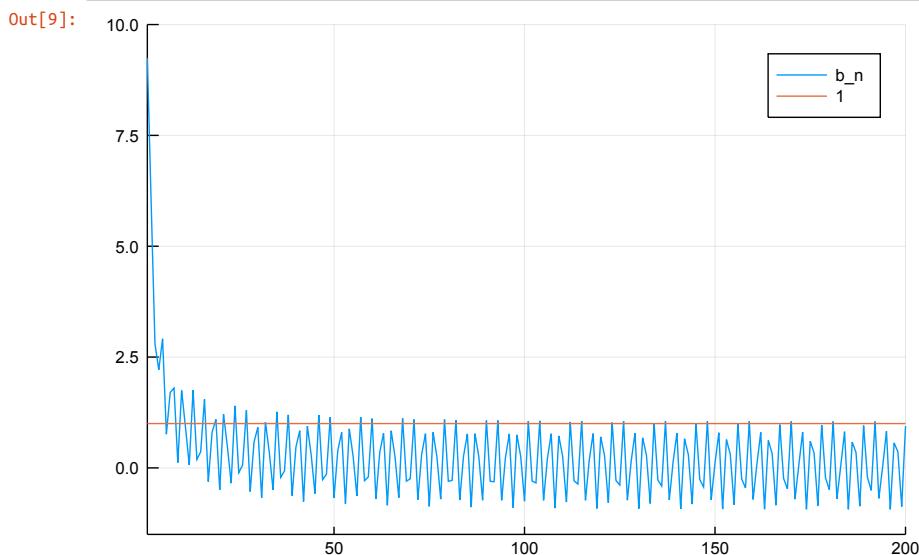
**問題:** 実数  $x$  に対して  $x$  以下の最大の整数を  $\lfloor x \rfloor$  と書き,  $f(x) = x - \lfloor x \rfloor$  とおく.  $f(x)$  は  $x$  の小数点以下の部分になっている.  $a$  は無理数であるとする. そのとき, 任意の  $\alpha \in [0, 1)$  と  $\varepsilon > 0$  に対して, ある正の整数  $n$  で  $f(na)$  と  $\alpha$  の距離が  $\varepsilon$  以下になるものが存在することを示せ.

**解答例:**  $a$  は無理数なので  $a, 2a, 3a, \dots$  は決して整数にはならない. そして,  $f(a), f(2a), f(3a), \dots$  の中の任意の2つは決して一致しない. なぜならば, もしも  $f(ma) = f(na)$ ,  $m < n$  ならば  $na - ma = (n - m)a$  が整数になって矛盾するからである.) だから,  $\{f(a), f(2a), f(3a), \dots\}$  は有限半開区間  $[0, 1)$  の無限部分集合になる. ゆえにある正の整数の組  $m < n$  で  $f(ma)$  と  $f(na)$  の距離が  $\varepsilon$  以下になるものが存在する. そのとき,  $(n - m)a$  とそれに最も近い整数のあいだの距離は  $\varepsilon$  以下になる. このとき,  $f(k(n - m))$ ,  $k = 1, 2, 3, \dots$  は  $[0, 1)$  の中に刻み幅が  $\varepsilon$  以下の等差数列を作るので, 正の整数  $k$  をうまく取れば  $f(k(n - m)a)$  と  $\alpha$  の距離を  $\varepsilon$  以下にできる.  $\square$

```
In [8]: 1 nmax = 100
2 n = 1:nmax
3 plot(n, 10./n .+ (-1).^n, label="a_n", ylims=(-1.5,10))
4 plot!(n, fill(1,nmax), label="1")
```



```
In [9]: 1 nmax = 200
2 n = 1:nmax
3 plot(n, 10./n + sin.(4.*n), label="b_n", ylims=(-1.5, 10))
4 plot!(n, fill(1,nmax), label="1")
```



```
In [ ]:
```

```
1
```

# 級数

黒木玄

2018-04-18

- Copyright 2018 Gen Kuroki
- License: MIT <https://opensource.org/licenses/MIT> (<https://opensource.org/licenses/MIT>)

このファイルは次の場所できれいに閲覧できる:

- <http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/02%20series.ipynb>  
<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/02%20series.ipynb>

このファイルは [Julia Box](https://juliabox.com) (<https://juliabox.com>) で利用できる。

自分のパソコンにJulia言語 (<https://julialang.org/>) をインストールしたい場合には

- [WindowsへのJulia言語のインストール](http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f) (<http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f>)

を参照せよ。

論理的に完璧な説明をするつもりはない。細部のいい加減な部分は自分で訂正・修正せよ。

## 目次

### 1 級数

#### 1.1 絶対収束と条件収束

- 1.1.1 絶対収束と条件収束の定義
- 1.1.2 絶対収束級数の収束先は和を見る順番によらないこと

#### 1.2 絶対収束級数の例

- 1.2.1 絶対収束級数の例1
- 1.2.2 絶対収束級数の例2

#### 1.3 条件収束級数の例

- 1.3.1 条件収束級数の例1
- 1.3.2 Euler変換による交代級数の数値計算
- 1.3.3 無限和と積分の交換に関する注意
- 1.3.4 条件収束級数の例2

### 2 複素数の級数

- 2.1 複素数の絶対収束級数
- 2.2 条件収束する複素級数の例

In [1]:

```
1  using Plots
2  gr(); ENV["PLOTS_TEST"] = "true"
3  #clibrary(:colorcet)
4  clibrary(:misc)
5
6  ▼ function pngplot(P...; kwargs...)
7      sleep(0.1)
8      pngfile = tempname() * ".png"
9      savefig(plot(P...; kwargs...), pngfile)
10     showimg("image/png", pngfile)
11  end
12  pngplot(; kwargs...) = pngplot(plot!(; kwargs...))
13
14  ▼ showimg(mime, fn) = open(fn) do f
15      base64 = base64encode(f)
16      display("text/html", """""")
17  end
18
19  using SymPy
20  #sympy[:init_printing](order="lex") # default
21  #sympy[:init_printing](order="rev-lex")
22
23  using SpecialFunctions
24  using QuadGK
```

## 1 級数

### 1.1 絶対収束と条件収束

### 1.1.1 絶対収束と条件収束の定義

**定理:**  $a_n \geq 0$  であると仮定する. そのとき, ある正の定数  $M$  が存在して, すべての番号  $n$  について  $\sum_{k=1}^n a_k \leq M$  が成立しているならば, 無限級数  $\sum_{n=1}^{\infty} a_n$  は有限の値に収束する.  $\square$

**定理:**  $\sum_{n=1}^{\infty} |a_n| < \infty$  ならば級数  $\sum_{n=1}^{\infty} a_n$  は収束する. そのとき, 級数  $\sum_{n=1}^{\infty} a_n$  は絶対収束するという.  $\square$

逆は成立しないことに注意に注意せよ. 級数  $\sum_{n=1}^{\infty} a_n$  が収束していても  $\sum_{n=1}^{\infty} |a_n| = \infty$  となる場合がある. そのとき, 級数  $\sum_{n=1}^{\infty} a_n$  は条件収束するという.

**注意:** 上の定理は  $a_n$  が実数列でも複素数列でも成立している. さらに  $a_n$  が  $\mathbb{R}^n$  や  $\mathbb{C}^n$  の点列でかつ絶対値をノルムで置き換えた場合にも成立している. 直観的には「長さが有限の紐の先端に対応する点が収束先として存在する」というような発想で考えればよい.

より一般にはノルム空間においては上の定理が成立することと完備性の条件は同値になっている. 実数列の場合には**実数の連続性**と同値になっている.

**実数の連続性**について理解していなくても, 絶対収束の概念を理解していれば足りることは多い.  $\square$

### 1.1.2 絶対収束級数の収束先は和を取る順番によらないこと

**定理:** 絶対収束する無限級数の収束先は足し上げる順序をどのように変えても同じになる.  $\square$

**系:** 特に  $a_n \geq 0$  のとき  $a_1, a_2, \dots$  の無限和は足し上げる順序をどのように変えても同じになる.  $\square$

**注意:**  $a_n \geq 0$  のとき  $a_1, a_2, \dots$  の無限和の値として  $\infty$  も許せば, その無限和は常に値を持っているとみなせる. 非負の実数の無限和を扱うときには収束先として  $\infty$  も許したくなる場合が多いので, その場合には収束先が有限の値になるか否かを明瞭に述べておいた方が誤解が少なくなる.  $\square$

**注意:** 条件収束級数の収束先は足し上げる順序によって変わるので扱いには注意が必要になる.  $\square$

## 1.2 絶対収束級数の例

### 1.2.1 絶対収束級数の例1

$s > 1$  と仮定する. このとき  $\sum_{n=1}^{\infty} \frac{1}{n^s} < \infty$  である. なぜならば  $n \rightarrow \infty$  のとき

$$\sum_{k=1}^n \frac{1}{k^s} \leq 1 + \int_1^n \frac{dx}{x^s} = 1 + \frac{1 - n^{1-s}}{s-1} \rightarrow 1 + \frac{1}{s-1}.$$

これより  $s > 1$  のとき

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \leq 1 + \frac{1}{s-1}$$

となることがわかった.  $s > 1$  の函数  $\zeta(s)$  を

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

と定める. これを**Riemannのゼータ函数**と呼ぶ. Riemannのゼータ函数の定義域は  $s \neq 1$  の複素数全体に自然に拡張される(解析接続される). Riemannのゼータ函数は整数論における最も基本的で重要な函数である.

$s > 1$  のとき,  $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^s}$  は絶対収束級数である. この級数は後で示すように  $s = 1$  でも収束している.(さらに複素平面全体に解析接続されることも示せる.)

**問題:**  $s > 1$  と仮定する. この仮定のもとで次の公式を証明せよ:

$$(1 - 2^{1-s})\zeta(s) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^s}.$$

**証明:** 絶対収束級数を扱っているので, 以下のように形式的に計算した結果がそのまま証明になる.

$$\begin{aligned}
(1 - 2^{1-s})\zeta(s) &= \left(1 - \frac{2}{2^s}\right) \sum_{n=1}^{\infty} \frac{1}{n^s} = \sum_{n=1}^{\infty} \frac{1}{n^s} - \sum_{n=1}^{\infty} \frac{2}{(2n)^s} \\
&= \left(\frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \dots\right) - \left(\frac{2}{2^s} + \frac{2}{4^s} + \dots\right) \\
&= \frac{1}{1^s} - \frac{1}{2^s} + \frac{1}{3^s} - \frac{1}{4^s} + \dots = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^s}. \quad \square
\end{aligned}$$

**問題:**  $\zeta(2)$  を数値計算して,  $\pi^2/6$  に近いことを確認せよ.

```
In [2]: 1 N = 10^8
2 SimpleSum = sum(n->1/n^2, 1:N)
3 TrueValue = pi^2/6
4 @show SimpleSum
5 @show TrueValue
6 @show SimpleSum - TrueValue;
```

```
SimpleSum = 1.6449340568482278
TrueValue = 1.6449340668482264
SimpleSum - TrueValue = -9.99999860695766e-9
```

**問題:**  $\zeta(2) = \pi^2/6$  の証明について調べよ.

**ヒント:** この問題は**Basel問題**などと呼ばれているようである. インターネットで検索してみよ.  $\square$

正の偶数  $2m$  について  $\zeta(2m)$  はすべて計算されている. しかし, 3以上の奇数  $2m+1$  に対する  $\zeta(3m+1)$  の数学的な素性はまだよくわかっていない.

### 1.2.2 絶対収束級数の例2

**問題:**  $\sum_{n=2}^{\infty} (\zeta(n) - 1) = 1$  を示せ.

**証明:**

$k \geq 2$  のとき

$$\sum_{n=2}^{\infty} \frac{1}{k^n} = \frac{1/k^2}{1 - 1/k} = \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}.$$

そして

$$\zeta(n) - 1 = \frac{1}{2^n} + \frac{1}{3^n} + \frac{1}{4^n} + \dots = \sum_{k=2}^{\infty} \frac{1}{k^n}$$

なので,

$$\begin{aligned}
\sum_{n=2}^{\infty} (\zeta(n) - 1) &= \sum_{n=2}^{\infty} \sum_{k=2}^{\infty} \frac{1}{k^n} = \sum_{k=2}^{\infty} \sum_{n=2}^{\infty} \frac{1}{k^n} = \sum_{k=2}^{\infty} \left( \frac{1}{k-1} - \frac{1}{k} \right) \\
&= \left( \frac{1}{1} - \frac{1}{2} \right) + \left( \frac{1}{2} - \frac{1}{3} \right) + \left( \frac{1}{3} - \frac{1}{4} \right) + \dots = 1.
\end{aligned}$$

2つ目の等号で  $1/k^n \geq 0$  なので足し上げる順序を変えても総和は等しいことを用了.  $\square$

以下はコンピューターによる計算.

```
In [3]: 1 N = 50
2 sum(n->zeta(n)-1, 2:N)
```

```
Out[3]: 1.0000000000000044
```

```
In [4]: 1 l, n = symbols("l n", integer=true, positive=true) # l = k-1
2 s = simplify(doit(Sum(1/(l+1)^n, (n,2,oo))))
3 S = doit(Sum(s, (l,1,oo)))
```

```
Out[4]: 1
```

**問題:** 上の問題の結果より,  $\sum_{n=2}^{\infty} (-1)^n (\zeta(n) - 1)$  は絶対収束する. その収束先を求めよ.

**証明:**

$k \geq 2$  のとき

$$\sum_{n=2}^{\infty} \frac{(-1)^n}{k^n} = \frac{1/k^2}{1+1/k} = \frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}.$$

そして

$$(-1)^n(\zeta(n) - 1) = \frac{(-1)^n}{2^n} + \frac{(-1)^n}{3^n} + \frac{(-1)^n}{4^n} + \dots = \sum_{k=2}^{\infty} \frac{(-1)^n}{k^n}$$

なので,

$$\begin{aligned} \sum_{n=2}^{\infty} (\zeta(n) - 1) &= \sum_{n=2}^{\infty} \sum_{k=2}^{\infty} \frac{(-1)^n}{k^n} = \sum_{k=2}^{\infty} \sum_{n=2}^{\infty} \frac{(-1)^n}{k^n} = \sum_{k=2}^{\infty} \left( \frac{1}{k} - \frac{1}{k+1} \right) \\ &= \left( \frac{1}{2} - \frac{1}{3} \right) + \left( \frac{1}{3} - \frac{1}{4} \right) + \left( \frac{1}{4} - \frac{1}{5} \right) + \dots = \frac{1}{2}. \end{aligned}$$

2つ目の等号で絶対収束級数なので足し上げる順序を変えても総和は等しいことを使った.  $\square$

以下はコンピューターによる計算.

```
In [5]: 1 N = 50
         2 sum(n->(-1)^n*(zeta(n)-1), 2:N)
```

```
Out[5]: 0.5000000000000018
```

```
In [6]: 1 l, n = symbols("l n", integer=true, positive=true) # l = k-1
         2 s = simplify(doit(Sum((-1)^n/(l+1)^n, (n,2,oo))))
         3 S = doit(Sum(s, (l,1,oo)))
```

```
Out[6]: 1/2
```

問題: 以上の例以外の絶対収束級数の例について調べ、ノートにまとめよ.

### 1.3 条件収束級数の例

定理: 一般に  $a_1 \geq a_2 \geq a_3 \geq \dots$  かつ  $a_n \geq 0$  かつ  $a_n \rightarrow 0$  のとき、交代級数

$$\sum_{n=1}^{\infty} (-1)^{n-1} a_n = a_1 - a_2 + a_3 - a_4 + \dots$$

は収束する.  $\square$

この定理中の  $a_n$  について  $\sum_{n=1}^{\infty} a_n = \infty$  が成立しているとき、交代級数は条件収束級数になる.

#### 1.3.1 条件収束級数の例1

$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$  は収束するが、 $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$  となる.

なぜならば、 $\sum_{k=1}^n \frac{1}{k} \geq \int_1^{n+1} \frac{dx}{x} = \log(n+1) \rightarrow \infty$  ( $n \rightarrow \infty$ ) が成立するからである.

問題:  $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$  を求めよ.

解答:  $\log 2$ .

証明:

一般に

$$\int_1^x \frac{dt}{t} = \log x.$$

ゆえに、

$$\int_0^1 \frac{dx}{1+x} = \int_1^2 \frac{dy}{y} = \log 2.$$

一方,  $|x| < 1$  のとき

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots$$

なので

$$\int_0^1 \frac{dx}{1+x} = \int_0^1 (1 - x + x^2 - x^3 + \dots) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}.$$

したがって,

$$\int_0^1 \frac{dx}{1+x} = \log 2 = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}. \quad \square$$

以下はコンピューターによる計算.

```
In [7]: 1 n = symbols("n", integer=true)
2 S = doit(Sum((-1)^(n-1)/n, (n, 1, oo)))
```

Out[7]:  $\log(2)$

```
In [8]: 1 N = 1000
2 SimpleSum = sum(n>->(-1)^(n-1)/n, 1:N)
3 TrueValue = log(2)
4 @show SimpleSum
5 @show TrueValue
6 @show SimpleSum - TrueValue;
```

```
SimpleSum = 0.6926474305598223
TrueValue = 0.6931471805599453
SimpleSum - TrueValue = -0.0004997500001230337
```

### 1.3.2 Euler変換による交代級数の数値計算

$\sum_{n=1}^N \frac{(-1)^{n-1}}{n}$  で  $N = 1000$  としても収束先の値を小数点以下第2桁までしか再現してくれない. 一般に条件収束する交代級数を定義通り足し上げると収束は遅い.

収束を加速するための様々な方法が知られているが, 交代級数の場合にはEuler変換が有名である. 詳しくは

- 交代ゼータ函数のオイラー変換 (<http://nbviewer.jupyter.org/gist/genkuroki/4a8fea5e2ed1b4e3b737c99acf237042>)

を参照せよ. 交代級数  $\sum_{k=0}^{\infty} (-1)^k a_k$  のEuler変換の結果は

$$\lim_{L \rightarrow \infty} \sum_{k=0}^{L-1} w_k^{(L)} (-1)^k a_k, \quad w_k^{(L)} = \sum_{n=k}^{L-1} 2^{-n-1} \binom{n}{k}.$$

すなわち交代級数のEuler変換は交代級数の各項にウェイト  $w_k^{(L)}$  をかけてから  $L \rightarrow \infty$  の極限を取る形になる.

以下のようにEuler変換を使えば数十項で精密な計算が可能になる.

```
In [9]: 1 binom(n,k) = exp(lgamma(n+1) - lgamma(k+1) - lgamma(n-k+1))
2
3 struct EulerWeight{T} W::T end
4 ▼ function EulerWeight(; Lmax=2^6)
5   W = zeros(Lmax,Lmax)
6   ▼ for L in 1:Lmax
7     ▼ for k in 0:L-1
8       W[L,k+1] = sum(n->2.0^(-n-1)*binom(n,k), k:L-1)
9     end
10   end
11   EulerWeight(W)
12 end
13 (eulerweight::EulerWeight)(L,k) = eulerweight.W[L,k+1]
14 eulerweight = EulerWeight()
```

Out[9]:  $EulerWeight\{Array\{Float64,2\}\}([0.5 \ 0.0 \ \dots \ 0.0 \ 0.0; 0.75 \ 0.25 \ \dots \ 0.0 \ 0.0; \dots; 1.0 \ 1.0 \ \dots \ 1.0842e-19 \ 0.0; 1.0 \ 1.0 \ \dots \ 3.52366e-18 \ 5.42101e-20])$

```
In [10]: 1 L = 2^6
2 EulerTran = sum(k->eulerweight(L,k)*(-1)^k/(k+1), 0:L-1)
3 TrueValue = log(2)
4 @show EulerTran
5 @show TrueValue
6 @show EulerTran - TrueValue
```

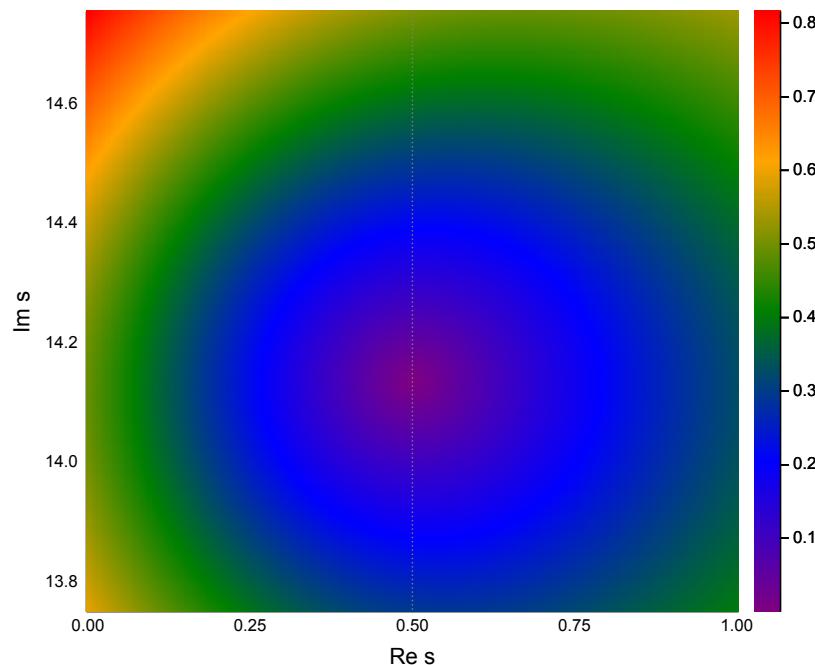
```
EulerTran = 0.6931471805599451
TrueValue = 0.6931471805599453
EulerTran - TrueValue = -2.220446049250313e-16
```

Out[10]: -2.220446049250313e-16

Euler変換を使うと、次のセルのように  $\zeta(s)$  の非自明な零点をプロットすることができる。

```
In [11]: 1 f(s) = sum(k -> eulerweight(L,k)*(-1)^k/(1+k)^s, 0:L-1)/(1-2^(1-s))
2 x = 0.0:0.005:1.0
3 y = 13.75:0.005:14.75
4 z = x' .+ im.*y
5 w = f.(z)
6 plot(size=(600, 500), title="first non-trivial zero of zeta(s)")
7 heatmap!(x, y, abs.(w), color=:rainbow, xlabel="Re s", ylabel="Im s")
8 vline!([0.5], color="gray", ls=:dot, label="")
```

Out[11]: first non-trivial zero of zeta(s)



複素平面における実部が 0 から 1 の縦に無限に長い帯状の領域を  $\zeta(s)$  の critical strip と呼ぶ。 $\zeta(s)$  の critical strip における零点を  $\zeta(s)$  の非自明な零点と呼ぶ。

**Riemann予想:**  $\zeta(s)$  の非自明な零点の実部はすべて  $1/2$  である. □

この予想は非常に有名であり、もしも証明されれば素数の分布に関する精密な評価が得られる。

**問題:**  $\zeta(s)$  のcritical stripにおける様子をプロットせよ. □

### 1.3.3 無限和と積分の交換に関する注意

上では  $\int_0^1 x^n dx = 1/(n+1)$  となることと、無限和  $1 - x + x^2 - x^3 + \dots$  と積分  $\int_0^1 dx$  の順序を交換できることを用いた。

無限和と積分の順序交換が可能なことは以下のようにして確認できる。まず  $|x| < 1$  で

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots + (-1)^{N-1} x^{N-1} + (-1)^N \frac{x^N}{1+x}$$

となることを確認せよ。この公式を使うと、

$$\int_0^1 \frac{dx}{1+x} = 1 - \frac{1}{2} + \frac{1}{3} + (-1)^{N-1} \frac{1}{N} + (-1)^N \int_0^1 \frac{x^N}{1+x} dx$$

さらに  $0 \leq x \leq 1$  のとき、 $1 \leq 1+x$  なので

$$0 \leq \frac{x^N}{1+x} \leq x^N$$

となるから,

$$0 \leq \int_0^1 \frac{x^N}{1+x} dx \leq \int_0^1 x^N dx = \frac{1}{N+1}$$

なので,  $N \rightarrow \infty$  のとき

$$\int_0^1 \frac{x^N}{1+x} dx \rightarrow 0.$$

以上合わせると,

$$\int_0^1 \frac{dx}{1+x} = \lim_{N \rightarrow \infty} \sum_{n=1}^N \frac{(-1)^{n-1}}{n} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$$

となることがわかる.

#### 1.3.4 条件収束級数の例2

問題:  $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n-1}$  を求めよ.

答え:  $\pi/4$ .

証明:

$a = \tan \theta$  の  $-\pi/2 < \theta < \pi/2$  における逆函数を  $\theta = \arctan a$  と書くと,

$$\int_0^a \frac{dt}{1+t^2} = \arctan a.$$

この公式は左辺で  $t = \tan \theta$  と置換すれば確認できる.

ゆえに,

$$\int_0^1 \frac{dx}{1+x^2} = \arctan 1 = \frac{\pi}{4}$$

一方,  $|x| < 1$  のとき

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \dots$$

なので

$$\int_0^1 \frac{dx}{1+x^2} = \int_0^1 (1 - x^2 + x^4 - x^6 + \dots) dx = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n-1}.$$

したがって,

$$\int_0^1 \frac{dx}{1+x^2} = \arctan 1 = \frac{\pi}{4} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n-1}. \quad \square$$

以下はコンピューターによる計算.

```
In [12]: 1 n = symbols("n", integer=True)
2 S = doit(Sum((-1)^(n-1)/(2n-1), (n, 1, oo)))
```

```
Out[12]: π
4
```

```
In [13]: 1 N = 10^8
2 SimpleSum = sum(n>-(-1)^(n-1)/(2n-1), 1:N)
3 TrueValue = π/4
4 @show SimpleSum
5 @show TrueValue
6 @show SimpleSum - TrueValue;
```

```
SimpleSum = 0.7853981608974486
TrueValue = 0.7853981633974483
SimpleSum - TrueValue = -2.499999651739415e-9
```

```
In [14]: 1 L = 2^6
2 EulerTran = sum(k->eulerweight(L,k)*(-1)^k/(2k+1), 0:L-1)
3 TrueValue = π/4
4 @show EulerTran
5 @show TrueValue
6 @show EulerTran - TrueValue
```

EulerTran = 0.7853981633974486  
TrueValue = 0.7853981633974483  
EulerTran - TrueValue = 3.3306690738754696e-16

Out[14]: 3.3306690738754696e-16

問題: 以上の計算における無限和と積分の正当化を自分で行ってみよ. □

問題: 以上の例以外の条件収束級数の例について調べ、ノートにまとめよ. □

以下はコンピューターによる  $\int_0^1 \frac{dx}{1+x^3}$  や  $\int_0^1 \frac{dx}{1+x^4}$  などの計算.

```
In [15]: 1 x = symbols("x", real=true, positive=true)
2 integrate(1/(1+x^3), (x,0,1))
```

Out[15]:  $\frac{1}{3} \log(2) + \frac{\sqrt{3}\pi}{9}$

```
In [16]: 1 n = symbols("n", integer=true, positive=true)
2 S = doit(Sum((-1)^(n-1)/(3n-2), (n,1,oo)))
```

Out[16]:  $\frac{1}{3} \log(2) - \frac{1}{3} e^{-\frac{5\pi i}{3}} \log\left(1 + e^{\frac{2\pi i}{3}}\right) - \frac{1}{3} e^{-\frac{i\pi}{3}} \log\left(1 + e^{-\frac{2\pi i}{3}}\right)$

```
In [17]: 1 x = symbols("x", real=true, positive=true)
2 integrate(x/(1+x^3), (x,0,1))
```

Out[17]:  $-\frac{1}{3} \log(2) + \frac{\sqrt{3}\pi}{9}$

```
In [18]: 1 n = symbols("n", integer=true, positive=true)
2 S = doit(Sum((-1)^(n-1)/(3n-1), (n,1,oo)))
```

Out[18]:  $-\frac{1}{3} \log(2) - \frac{1}{3} e^{-\frac{2\pi i}{3}} \log\left(1 + e^{-\frac{2\pi i}{3}}\right) - \frac{1}{3} e^{-\frac{10\pi i}{3}} \log\left(1 + e^{\frac{2\pi i}{3}}\right)$

```
In [19]: 1 x = symbols("x", real=true, positive=true)
2 integrate(1/(1+x^4), (x,0,1))
```

Out[19]:  $-\frac{\sqrt{2}}{8} \log(-\sqrt{2} + 2) + \frac{\sqrt{2}}{8} \log(\sqrt{2} + 2) + \frac{\sqrt{2}\pi}{8}$

```
In [20]: 1 n = symbols("n", integer=true, positive=true)
2 S = doit(Sum((-1)^(n-1)/(4n-3), (n,1,oo)))
```

Out[20]:  $-\frac{i}{4} e^{-\frac{i\pi}{4}} \log\left(1 + e^{\frac{3\pi i}{4}}\right) + \frac{1}{4} e^{-\frac{i\pi}{4}} \log\left(1 + e^{\frac{i\pi}{4}}\right) + \frac{i}{4} e^{-\frac{i\pi}{4}} \log\left(1 + e^{-\frac{i\pi}{4}}\right) - \frac{1}{4} e^{-\frac{i\pi}{4}} \log\left(1 + e^{-\frac{3\pi i}{4}}\right)$

```
In [21]: 1 x = symbols("x", real=true, positive=true)
2 s = symbols("s", real=true, positive=true)
3 J = simplify(integrate(1/(1+x^(1/s)), (x,0,1)))
```

Out[21]:  $s\Phi(e^{i\pi}, 1, s)$

## 2 複素数の級数

この節では複素数  $a_n$  の級数  $\sum_{n=1}^{\infty} a_n$  を扱う.

### 2.1 複素数の絶対収束級数

**定理:** 複素数  $a_1, a_2, \dots$  について, 級数  $\sum_{n=1}^{\infty} |a_n|$  が有限の値に収束するならば, 級数  $\sum_{n=1}^{\infty} a_n$  も収束する. このとき, 級数  $\sum_{n=1}^{\infty} a_n$  は絶対収束するという.  $\square$

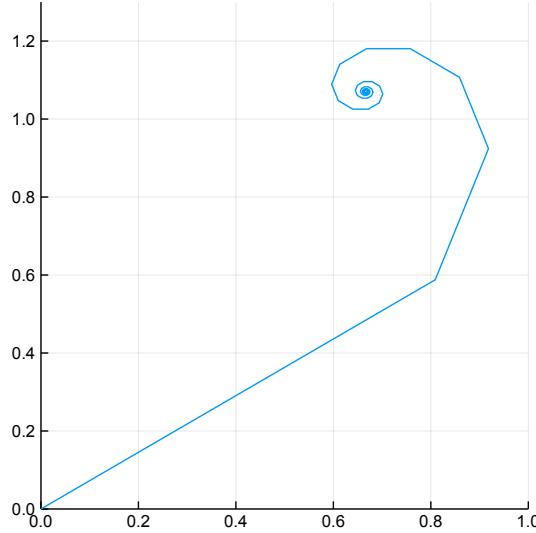
絶対収束していないが収束する級数は**条件収束**するという.

上の定理の証明は略す. しかし, 下の方で, 級数  $\sum_{n=1}^{\infty} |a_n|$  が有限の値に収束するならば, 級数  $\sum_{n=1}^{\infty} a_n$  も収束する直観的な理由を説明する.

**例:**  $\theta$  は実数であるとする. このとき, 複素級数  $\sum_{n=1}^{\infty} \frac{e^{in\theta}}{n^{3/2}}$  は絶対収束する. 以下のセルにおける計算とプロットを参照せよ.

```
In [22]:
1 θ = 2π/10
2 a(n) = exp(im*n*θ)/n^1.5
3 sum_a(n) = iszero(n) ? Complex(0.0) : sum(k->a(k), 1:n)
4 N = 70
5 n = 0:N
6 z = sum_a_(n)
7 plot(size=(400, 400), xlims=(0,1.0), ylims=(0,1.3))
8 plot!(real.(z), imag.(z), legend=false#, line=:arrow)
```

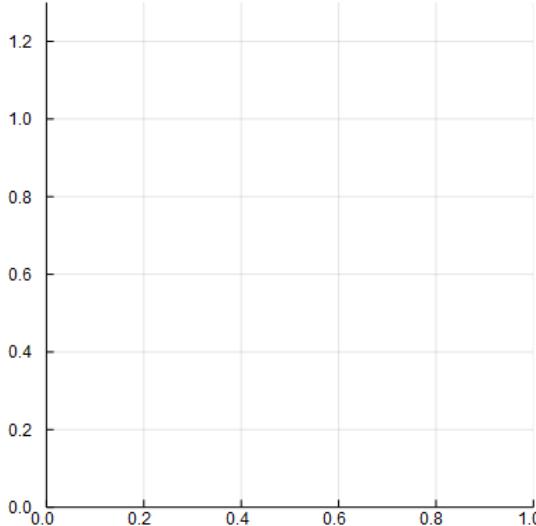
Out[22]:



```
In [23]: 1 θ = 2π/10
2 a(n) = exp(im*n*θ)/n^1.5
3 sum_a(n) = iszero(n) ? Complex(0.0) : sum(k->a(k), 1:n)
4 ▼ @time anim = @animate for N in 0:70
5     n = 0:N
6     z = sum_a.(n)
7     plot(size=(400, 400), xlims=(0,1.0), ylim=(0,1.3))
8     plot!(real.(z), imag.(z), legend=false, line=:arrow)
9 end
10
11 gifname = "sum_a.gif"
12 @time gif(anim, gifname; fps=10)
13 sleep(0.1)
14 showing("image/gif", gifname)
```

2.757190 seconds (19.47 M allocations: 502.184 MiB, 3.32% gc time)  
0.890623 seconds (89.10 k allocations: 4.771 MiB)

INFO: Saved animation to C:\Users\genkuroki\OneDrive\Calculus\sum\_a.gif



#### 絶対収束の概念の直観的な説明:

各複素数  $a_n$  は平面上の点と同一視できる(複素平面の考え方).

平面上の点と原点からその点への矢線(ベクトル)は同一視できる(ベクトルの考え方).

複素数の和はベクトルの和に対応している.

ベクトルの和は矢線としてのベクトルを連結することを意味する(ベクトルの和の幾何学的意味).

ゆえに,  $s_N = \sum_{n=1}^N a_n$  は各  $a_1, a_2, \dots, a_N$  の各々に対応する矢線を連結して得られる折れ線の先端の点に対応する複素数に等しい.

$L = \sum_{n=1}^{\infty} |a_n|$  は各  $a_n$  に対応する矢線の長さの総和を意味する.

ゆえに,  $L$  が有限の値になることは, 各  $a_n$  に対応する矢線をすべて連結して得られる折れ線の長さが有限になることを意味する.

そのとき,  $a_n$  に対応する矢線をすべて連結して得られる折れ線の「先端」の点に対応する複素数が級数  $\sum_{n=1}^{\infty} a_n$  の収束先になる.

1つ前のセルのアニメーションはそのような様子の一例になっている.

注意: 上の説明を理解できれば, 級数  $\sum_{n=1}^{\infty} |a_n|$  が有限の値に収束するならば, 級数  $\sum_{n=1}^{\infty} a_n$  も収束することも十分に納得できると思う. その直観は數

学的には完備性 (<https://www.google.co.jp/search?q=%E5%AE%8C%E5%82%99%E6%80%A7+E7%B5%B6%E5%AF%BE%E5%8F%8E%E6%9D%9F>) の概念によって正当化される. □

## 2.2 条件収束する複素級数の例

例:  $\theta$  は実数であり,  $e^{i\theta} \neq 1$  であると仮定する. このとき  $\sum_{n=1}^{\infty} \frac{e^{in\theta}}{n}$  は条件収束する. □

**例の解説のため的一般論:** 一般に  $S_n = \sum_{k=1}^n a_k b_k$  のとき,  $A_k = \sum_{j=1}^k a_j$ ,  $A_0 = 0$  とおくと,  $a_k = A_k - A_{k-1}$  なので

$$\begin{aligned} S_n &= \sum_{k=1}^n (A_k - A_{k-1}) b_k = \sum_{k=1}^n A_k b_k - \sum_{k=0}^{n-1} A_k b_{k+1} \\ &= A_n b_n - A_0 b_1 + \sum_{k=1}^{n-1} A_k (b_k - b_{k+1}). \end{aligned}$$

この公式は部分積分の公式の和に関するバージョンであることに注意せよ。(「部分積分」のアイデアは普遍的によく現われる。)

仮定の追加1: 数列  $b_k$  は単調減少実数列でかつ 0 に収束すると仮定する。

このとき,  $b_k - b_{k+1} \geq 0$  でかつ

$$\sum_{k=1}^{n-1} (b_k - b_{k+1}) = b_1 - b_n \rightarrow b_1 \quad (n \rightarrow \infty).$$

仮定の追加2:  $A_k$  は有界数列であると仮定する。すなわちある正の実数で  $|A_k| \leq M$  ( $k = 1, 2, \dots$ ) を満たすものが存在すると仮定する。

このとき,

$$\sum_{k=1}^{\infty} |A_k| |b_k - b_{k+1}| \leq M \sum_{k=1}^{\infty} (b_k - b_{k+1}) = Mb_1 < \infty$$

なので 級数  $\sum_{k=1}^{\infty} A_k (b_{k+1} - b_k)$  は絶対収束する。さらに,  $n \rightarrow \infty$  で  $b_n \rightarrow 0$  なので  $A_n b_n \rightarrow 0$  となる。ゆえに級数  $\sum_{k=1}^{\infty} a_k b_k$  は収束する。(積分の場合にも、絶対収束しない積分を部分積分によって絶対収束する積分に変形できる場合がある。各自自分で研究してみよ。) □

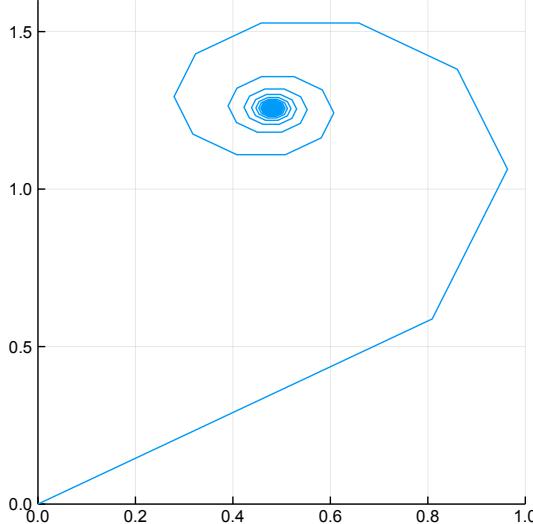
**例の解説:** 以上的一般論を上の例に適用しよう。すなわち  $a_n = e^{in\theta}$ ,  $b_n = 1/n$  の場合を考える。このとき

$$A_k = \sum_{j=1}^k a_j = \frac{e^{i\theta} - e^{i(n+1)\theta}}{1 - e^{i\theta}}$$

であるから,  $|A_k| \leq 2/|1 - e^{i\theta}|$  なので,  $A_k$  は有界である。実数列  $b_k$  は単調減少しながら 0 に収束する。したがって上の一般論より, 級数  $\sum_{n=1}^{\infty} \frac{e^{in\theta}}{n}$  は条件収束する。□

```
In [24]: 1 θ = 2π/10
2 a(n) = exp(im*n*θ)/n
3 sum_a(n) = iszero(n) ? Complex(0.0) : sum(k->a(k), 1:n)
4 N = 400
5 n = 0:N
6 z = sum_a.(n)
7 plot(size=(400, 400), xlims=(0,1.0), ylims=(0,1.6))
8 plot!(real.(z), imag.(z), legend=false) #, line=:arrow
```

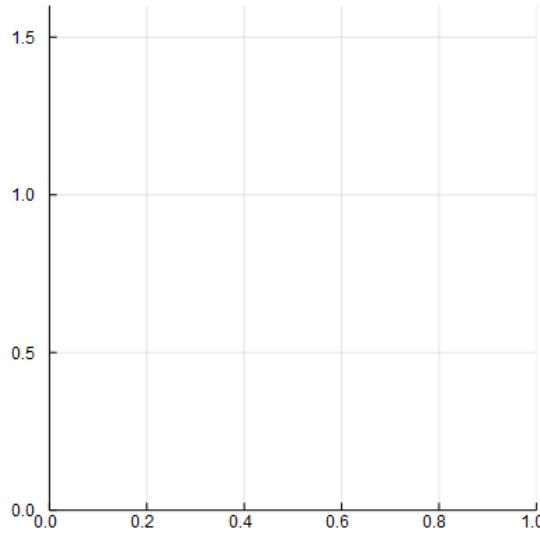
Out[24]:



```
In [25]: 1 θ = 2π/10
2 a(n) = exp(im*n*θ)/n
3 sum_a(n) = iszero(n) ? Complex(0.0) : sum(k->a(k), 1:n)
4 ▼ @time anim = @animate for N in 0:4:400
5     n = 0:N
6     z = sum_a.(n)
7     plot(size=(400, 400), xlims=(0,1.0), ylim=(0,1.6))
8     plot!(real.(z), imag.(z), legend=false, line=:arrow)
9 end
10
11 gifname = "sum_a.gif"
12 @time gif(anim, gifname; fps=10)
13 sleep(0.1)
14 showing("image/gif", gifname)
```

4.223874 seconds (40.28 M allocations: 1.076 GiB, 5.13% gc time)  
0.635724 seconds (1.05 k allocations: 69.656 KiB)

INFO: Saved animation to C:\Users\genkuroki\OneDrive\Calculus\sum\_a.gif



上のセルのアニメーションの折れ線の長さは無限大である。無限に長くなる折れ線の先端が無限に回転しながら収束している。

In [ ]:

1

# $\pi$ と e と Euler定数 γ ¶

黒木玄

2018-04-18

- Copyright 2018 Gen Kuroki
- License: MIT <https://opensource.org/licenses/MIT> (<https://opensource.org/licenses/MIT>)

このファイルは次の場所できれいに閲覧できる:

- <http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/03%20pi%2C%20e%2C%20and%20Euler's%20y.ipynb>  
<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/03%20pi%2C%20e%2C%20and%20Euler's%20y.ipynb>

このファイルは [Julia Box \(https://juliabox.com\)](https://juliabox.com) で利用できる。

自分のパソコンに[Julia言語 \(https://julialang.org/\)](https://julialang.org/)をインストールしたい場合には

- [WindowsへのJulia言語のインストール \(http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f\)](http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f)

を参照せよ。

論理的に完璧な説明をするつもりはない。細部のいい加減な部分は自分で訂正・修正せよ。

## 目次

### 1 円周率, Napier数, Euler定数

#### 1.1 準備: 曲線の長さ

#### 1.2 円周率

##### 1.2.1 円周率の定義1

##### 1.2.2 正弦函数の微積分

##### 1.2.3 円周率の定義2

##### 1.2.4 正接函数の微積分

#### 1.3 Napier数

##### 1.3.1 対数函数と指數函数

##### 1.3.2 Napier数の定義

#### 1.4 Euler定数

##### 1.4.1 Euler定数の定義

##### 1.4.2 条件収束交代級数への応用

##### 1.4.3 Euler定数がガンマ函数の無限積表示に出て来ること

##### 1.4.4 Euler定数が高次元単体上の一様分布に関連して自然に出て来ること

##### 1.4.5 $\zeta(s) - 1/(s-1)$ の $s \rightarrow 0$ での極限がEuler定数に等しいこと

```
In [1]: 1 using Plots; gr(); ENV["PLOTS_TEST"] = "true"
2 #clibrary(:colorcet)
3 clibrary(:misc)
4
5 using SymPy
6 using SpecialFunctions
```

## 1 円周率, Napier数, Euler定数

### 1.1 準備: 曲線の長さ

高校数学IIIの復習: 曲線の長さは速さの時間積分で定義される。すなわち, 曲線  $(x(t), y(t))$ ,  $a < t < b$  の長さ  $\ell$  は定積分

$$\ell = \int_a^b \sqrt{x'(t)^2 + y'(t)^2} dt$$

で定義される。被積分函数は点の運動の速さを意味する。

### 1.2 円周率

#### 1.2.1 円周率の定義1

円周率  $\pi$  の定義は单位半円周の長さである。单位半円周  $C$  は

$$C = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1, x > 0\}$$

と定義される。これは

$$x(t) = \sqrt{1 - t^2}, \quad y(t) = t$$

によって

$$C = \{(x(t), y(t)) \mid -1 < t < 1\}$$

と表わされ、ちょっとした計算で

$$\sqrt{x'(t)^2 + y'(t)^2} = \frac{1}{\sqrt{1 - t^2}}$$

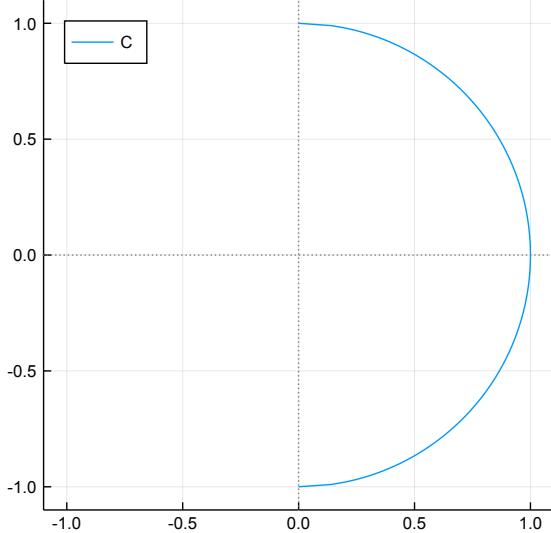
となることがわかる(自分で計算してみよ)。ゆえに、円周率は

$$\pi = \int_{-1}^1 \frac{dt}{\sqrt{1 - t^2}}$$

と表わされる。これを円周率の定義だと思ってよい。

```
In [2]: 1 xx(t) = sqrt(1-t^2)
2 yy(t) = t
3 t = -1:0.01:1
4 plot(size=(500, 400), aspect_ratio=1, xlim=(-1.1,1.1), ylim=(-1.1,1.1))
5 plot!(xx.(t), yy.(t), label="C", legend=:topleft)
6 hline!([0], color="gray", ls=:dot, label="")
7 vline!([0], color="gray", ls=:dot, label="")
```

Out[2]:



```
In [3]: 1 t = symbols("t", real=true)
2 x = sqrt(1-t^2)
3 y = t
4 simplify(diff(x,t)^2+diff(y,t)^2)
```

Out[3]:

$$-\frac{1}{t^2 - 1}$$

```
In [4]: 1 integrate(1/sqrt(1-t^2), (t,-1,1))
```

Out[4]:

$$\pi$$

### 1.2.2 正弦函数の微積分

前節の計算より、 $-1 \leq y \leq 1$  に対して、

$$F(y) = \int_0^y \frac{dt}{\sqrt{1 - t^2}}$$

は  $C$  上の点  $(1, 0)$  から  $(\sqrt{1 - y^2}, y)$  までの弧の長さの  $\pm 1$  倍になる。符号は  $y > 0$  なら  $+$  になり、 $y < 0$  なら  $-$  になる。すなわち、 $\theta(y)$  は弧度法の意味での角度である。単位円上の  $y$  座標を角度  $\theta$  の函数とみなしたもので  $y = \sin \theta$  と定義するのであった。これは、 $y = \sin \theta$  は  $\theta = F(y)$  の逆函数であることを意味する。 $F(y)$  は  $\arcsin y$  と書かれることが多い。

$\frac{d\theta}{dy} = F'(y) = \frac{1}{\sqrt{1-y^2}}$  ので、その逆函数の導函数は

$$\frac{d\sin\theta}{d\theta} = \frac{dy}{d\theta} = \sqrt{1-y^2} = \sqrt{1-\sin^2\theta}.$$

これで  $\sin\theta$  の導函数の計算の仕方もわかった。

以上の筋道に従えば、高校数学IIIまでの範囲内で、三角函数の微分積分学が適切に展開可能なことがわかる。

### 1.2.3 円周率の定義2

傾き  $t$  の直線  $y = tx$  と単位半円周  $C$  の交点を  $(X, Y) = (X(t), Y(t))$  と書くと、

$$X = X(t) = \frac{1}{\sqrt{1+t^2}}, \quad Y = Y(t) = \frac{t}{\sqrt{1+t^2}}.$$

このとき、

$$\sqrt{X'(t)^2 + Y'(t)^2} = \frac{1}{1+t^2}$$

となる(自分で工夫して楽な計算の仕方を見付けるように努力してみよ)。ゆえに円周率は

$$\pi = \int_{-\infty}^{\infty} \frac{dt}{1+t^2}$$

と表わされる。これを円周率の定義だと思ってよい。

**問題:** 積分  $\int_{-\infty}^{\infty} \frac{dt}{1+t^2}$  で  $t = \frac{s}{\sqrt{1-s^2}}$  と置換すると、

$$\int_{-\infty}^{\infty} \frac{dt}{1+t^2} = \int_{-1}^1 \frac{ds}{\sqrt{1-s^2}}$$

が得られることを確認せよ。□

**注意:**  $X = 1/\sqrt{1+t^2}$ ,  $Y = t/\sqrt{1+t^2}$  のとき、 $X^2 + Y^2 = 1$  より  $X = \sqrt{1-Y^2}$  となるので、 $t = Y/X = Y/\sqrt{1-Y^2}$ 。さらに  $Y = s$  とおけば上の問題の置換の式が得られる。□。

```
In [5]: 1 t = symbols("t", real=True)
2 X = 1/sqrt(1+t**2)
3 Y = t*X
4 simplify(sqrt(diff(X,t)**2+diff(Y,t)**2))
```

Out[5]:

$$\frac{1}{t^2 + 1}$$

```
In [6]: 1 s = symbols("s")
2 t = s/sqrt(1-s**2)
3 simplify(diff(t,s)/(1+t**2))
```

Out[6]:

$$\frac{1}{\sqrt{-s^2 + 1}}$$

### 1.2.4 正接函数の微積分

前節の計算より、

$$\theta = G(a) = \int_0^a \frac{dt}{1+t^2}$$

は単位半円周上の点  $(1, 0)$  から傾き  $a$  の直線  $y = ax$  と単位半円周  $C$  の交点までの弧の長さの  $\pm 1$  倍になり、符号  $a$  の符号に等しくなる。すなわち、弧度法の意味での角度に等しい。

正接函数の定義は傾きを角度の函数とみたものであった。これは  $a = \tan\theta$  が  $\theta = G(a)$  の逆函数になっていることを意味している。 $G(a)$  は  $\arctan a$  と書かれることが多い。

$\frac{d\theta}{da} = G'(a) = \frac{1}{1+a^2}$  ので、 $\frac{d\tan\theta}{d\theta} = \frac{da}{d\theta} = 1 + a^2 = 1 + \tan^2\theta$ .

以下の公式もよく使われる：

$$\int_0^1 \frac{dt}{1+t^2} = \frac{\pi}{4}, \quad \int_0^\infty \frac{dt}{1+t^2} = \frac{\pi}{2}.$$

```
In [7]: 1 t = symbols("t", real=True)
2 integrate(1/(1+t^2), (t,0,1)), integrate(1/(1+t^2), (t,0,oo))

Out[7]: (pi/4, pi/2)
```

### 1.3 Napier数

#### 1.3.1 対数函数と指數函数

このノートでは  $x > 0$  の(自然)対数函数  $y = \log x$  を

$$y = \log x = \int_1^x \frac{dt}{t}$$

で定義し、この逆函数で指數函数  $x = \exp y$  を定義しておく。

このとき、 $\log 1 = 0$ ,  $\exp 0 = 1$  となる。

さらに、 $\frac{dy}{dx} = \frac{d\log x}{dx} = \frac{1}{x}$  が成立し、その逆函数については  $\frac{dx}{dy} = \frac{d\exp y}{dy} = x = \exp y$ .

これらの条件によって  $\log x$  と  $\exp y$  は一意的に定まる事を示せるが、ここではその証明には触れない。

**問題:** 上の対数函数の定義だけを用いて  $\log(x_1 x_2) = \log x_1 + \log x_2$  ( $x_1, x_2 > 0$ ) となることを示せ。

**証明:**

$$\log(x_1 x_2) = \int_1^{x_1 x_2} \frac{dt}{t} = \int_1^{x_1} \frac{dt}{t} + \int_{x_1}^{x_1 x_2} \frac{dt}{t}.$$

2つ目の項の積分変数を  $t = x_1 s$  で置換すると、

$$\int_{x_1}^{x_1 x_2} \frac{dt}{t} = \int_1^{x_2} \frac{x_1 ds}{x_1 s} = \int_1^{x_2} \frac{ds}{s}.$$

$s$  を改めて  $t$  とおけば欲しい結果が得られる。□

**問題:** 上の問題の結果を用いて  $\exp(y_1 + y_2) = \exp(y_1)\exp(y_2)$  となることを示せ。

**証明:**  $x_i = \exp y_i$  とおくと、 $y_i = \log x_i$  であり、上の問題の結果より、 $y_1 + y_2 = \log(x_1 x_2)$  なので

$$\exp(y_1 + y_2) = \exp(\log(x_1 x_2)) = x_1 x_2 = \exp(y_1)\exp(y_2). \quad \square$$

$a > 0$  に対して、 $a^y$  を

$$a^y = \exp(y \log a)$$

によって定義する。このとき、整数  $n$  について

$$a^1 = \exp(\log a) = a, \quad a^n = \exp(n \log a) = (n \text{ 個の } \exp(\log a) \text{ の積}) = (n \text{ 個の } a \text{ の積}).$$

後者で上の問題の結果を使った。これより、整数の  $n$  について、上の定義に基いた  $a^n$  と  $n$  個の  $a$  の積が一致することがわかった。上のように定義しても混乱は生じない。

#### 1.3.2 Napier数の定義

このノートでは Napier数  $e$  を

$$e = \exp(1)$$

によって定義する。すなわち  $e$  は

$$\int_1^e \frac{dt}{t} = 1$$

という条件で定義されているものとする。

**定理:**

$$\lim_{n \rightarrow \infty} \left(1 + \frac{y}{n}\right)^n = \exp(y).$$

証明:

$$\left(1 + \frac{y}{n}\right)^n = \exp\left[n \log\left(1 + \frac{y}{n}\right)\right]$$

である.  $d(\log x)/dx = 1/x$  の  $x = 1$  の場合より,  $h \rightarrow 0$  のとき,

$$\frac{\log(1+h)}{h} = \frac{\log(1+h) - \log(1)}{h} \rightarrow 1.$$

これを  $h = y/n$ ,  $n \rightarrow \infty$  の場合に適用すると,

$$\frac{\log(1+y/n)}{1/n} = y \frac{\log(1+y/n)}{y/n} \rightarrow y.$$

ゆえに,  $\exp$  の連続性を用いると,  $n \rightarrow \infty$  のとき,

$$\left(1 + \frac{y}{n}\right)^n = \exp\left[n \log\left(1 + \frac{y}{n}\right)\right] \rightarrow \exp(y). \quad \square$$

上の定理の  $y = 1$  の特別な場合として次が得られる.

系:

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e. \quad \square$$

以上によって, 対数函数と指数函数とNapier数  $e$  の理論が微分積分学を用いてすべて再構成された.

## 1.4 Euler定数

以上によって, 数学における基本的な定数  $\pi$  と  $e$  の定義が明瞭になった.

しかし, それだけで終わってしまうと, 高校の数学で習うことを単に再構成したに過ぎなくなってしまう.

そこでこのノートではその次に知っておくべきだと思われるEuler定数  $\gamma$  について説明しよう.

Euler定数とその歴史については次の文献が詳しい:

- <https://arxiv.org/abs/1303.1856> (<https://arxiv.org/abs/1303.1856>)

### 1.4.1 Euler定数の定義

数列  $a_n, b_n$  を次のように定める:

$$a_n = \frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n} - \log(n+1)$$
$$b_n = \frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n} - \log n.$$

$y = 1/x$  は狭義単調減少函数なので

$$\log(n+1) - \log n = \int_n^{n+1} \frac{dx}{x} < \frac{1}{n} < \int_{n-1}^n \frac{dx}{x} = \log n - \log(n-1),$$
$$\log(n+1) = \int_1^{n+1} \frac{dx}{x} < \frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n}$$

となる(図を描いて確認せよ). 後者より,

$$0 < a_n < b_n$$

が成立していることがわかり, 前者より,

$$a_n - a_{n-1} = \frac{1}{n} + \log n - \log(n+1) > 0,$$
$$b_n - b_{n-1} = \frac{1}{n} + \log(n-1) - \log n < 0.$$

となり,  $a_n$  は単調増加し,  $b_n$  は単調減少することがわかる.(図を描けば,  $a_n$  が単調増加し,  $b_n$  が単調減少することは見た目的に明らかになる.)

したがって,  $a_n$  も  $b_n$  も収束することがわかる.

さらに,  $0 < \log(n+1) - \log n < 1/n$  より  $a_n$  と  $b_n$  の収束先が同じであることもわかる.

それらの収束先を  $\gamma$  と書き, Euler定数と呼ぶ.

$$\gamma = \lim_{n \rightarrow \infty} \left( \frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n} - \log(n+1) \right) = \lim_{n \rightarrow \infty} \left( \frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n} - \log n \right).$$

この定義は次のように書き換えられる:

$$\frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n} = \log n + \gamma + o(1).$$

ここで  $o(1)$  は  $n \rightarrow \infty$  で 0 に収束する項を表す.

**考え方:** 解析学では

$$(\text{調べたい量}) = (\text{相対的にわかりやすい量}) + (\text{誤差})$$

の形式で考える. Euler定数の定義は、調べたいもの

$$\frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n}$$

を相対的にわかりやすい量

$$\log n + \gamma$$

プラス誤差の形式で認識することであるとみなせる.  $n$  が大きなとき、第1項の  $\log n$  と比較して第2項の定数項  $\gamma$  は小さいとみなせ、第2項  $\gamma$  と比較して誤差項  $o(1)$  は小さいとみなせる.  $1/1 + 1/2 + \dots + 1/n$  を  $\log n$  で大雑把に近似するより、 $\log n + \gamma$  で近似する方がより精密な近似になる.

このように新たに項を付け加えることによって近似の精度を逐次的に高めて行くことは、解析学における典型的なパターンであり、あらゆる場所で出会うことになる.  $\square$

**注意:** digamma函数  $\psi(s)$  ( $s > 0$ ) を

$$\psi(s) = - \lim_{n \rightarrow \infty} \left[ \frac{1}{s} + \frac{1}{s+1} + \dots + \frac{1}{s+n} - \log n \right]$$

と定めることができる. これはガンマ函数が

$$\Gamma(s) = \lim_{n \rightarrow \infty} \frac{s^s n!}{s(s+1)\cdots(s+n)}$$

と書くことを使えば  $\psi(s) = d\log\Gamma(s)/ds$  とも表わされる. ガンマ函数に関するこの公式については

- [ガンマ分布の中心極限定理とStirlingの公式](https://genkuroki.github.io/documents/20160501StirlingFormula.pdf) (<https://genkuroki.github.io/documents/20160501StirlingFormula.pdf>)

の第8.2節を参照せよ. Euler定数の定義より  $\psi(1) = -\gamma$  であることがわかり、digamma函数の定義より  $\psi(s+1) = 1/s + \psi(s)$  であることがわかるので、帰納的に非負の整数  $n$  について

$$\psi(n+1) = \frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n} - \gamma$$

であることがわかる. この公式はdigamma函数  $\psi(s+1)$  は調和級数からEuler定数を引いたものの連続的な補間になっていることを意味している. したがって、 $\psi(s+1)$  の  $s \rightarrow \infty$  での漸近挙動がよくわかれば、調和級数の漸近挙動もよくわかることになる.

ガンマ函数については次のStirlingの近似公式がよく知られている:  $s \rightarrow \infty$  において

$$\log\Gamma(s+1) = s \log s - s + \frac{1}{2} \log s + \log \sqrt{2\pi} + \frac{1}{12s} + O\left(\frac{1}{s^2}\right).$$

これを  $s$  で微分することによって、

$$\psi(s+1) = \log s + \frac{1}{2s} - \frac{1}{12s^2} + O\left(\frac{1}{s^3}\right) \quad (s \rightarrow \infty).$$

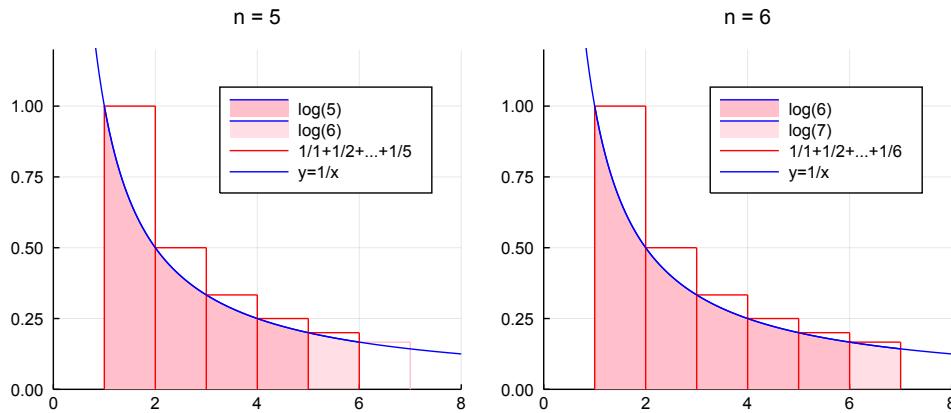
以上をまとめると次が得られる:

$$\frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n} = \log n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + O\left(\frac{1}{n^3}\right) \quad (s \rightarrow \infty).$$

このように調和級数  $\frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n}$  の  $n \rightarrow \infty$  における漸近挙動の詳細はガンマ函数に関するStirlingの公式の対数微分によって得られる.  $\square$

```
In [8]: 1 ▼ function plot_eulergamma(N, n)
2     f(x) = 1/x
3     p = plot(xlim=(0,N+2), ylim=(0, 1.2))
4     x = 1.0:0.05:n
5     plot!(p, x, f.(x), color=:blue, label="log($n)", fill=(0, 1.0, :pink))
6     x = n:0.05:n+1
7     plot!(p, x, f.(x), color=:blue, label="log($(n+1))", fill=(0, 0.5, :pink))
8     plot!(p, [1,1], [0,1], color=:red, label="1/1+1/2+...+1/$n")
9     for k in 1:N
10        x = k:0.05:k+1
11        g(x) = 1/k
12        c = ifelse(k > n, :pink, :red)
13        plot!(p, x, g.(x), color=c, label="")
14        plot!(p, [k+1,k+1], [0,1/k], color=c, label="")
15    end
16    x = 0.1:0.05:N+2
17    plot!(p, x, f.(x), color=:blue, label="y=1/x")
18    plot!(title="n = $n", titlefontsize=10)
19    p
20 end
21
22 N = 6
23 p1 = plot_eulergamma(N, N-1)
24 p2 = plot_eulergamma(N, N)
25 plot(p1, p2, size=(700, 300))
```

Out[8]:



```
In [9]: 1 a(n) = sum(k->1/k, 1:n) - log(n+1)
2 b(n) = sum(k->1/k, 1:n) - log(n)
3 [(n, a(n), b(n)) for n in [1:9; 10.^{1:7}]]
```

```
Out[9]: 16-element Array{Tuple{Int64,Float64,Float64},1}:
 (1, 0.306853, 1.0)
 (2, 0.401388, 0.806853)
 (3, 0.447039, 0.734721)
 (4, 0.473895, 0.697039)
 (5, 0.491574, 0.673895)
 (6, 0.50409, 0.658241)
 (7, 0.513416, 0.646947)
 (8, 0.520633, 0.638416)
 (9, 0.526383, 0.631744)
 (10, 0.531073, 0.626383)
 (100, 0.572257, 0.582207)
 (1000, 0.576716, 0.577716)
 (10000, 0.577166, 0.577266)
 (100000, 0.577211, 0.577221)
 (1000000, 0.577215, 0.577216)
 (10000000, 0.577216, 0.577216)
```

```
In [10]: 1 eulergamma
```

Out[10]:  $\gamma = 0.5772156649015\dots$

#### 1.4.2 条件収束交代級数への応用

次の2つの数列を考える:

$$A_n = \frac{1}{1} + \frac{1}{3} + \dots + \frac{1}{2n-1}, \quad B_n = \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2n}.$$

このとき

$$A_n + B_n = \sum_{k=1}^{2n} \frac{1}{k} = \log(2n) + \gamma + o(1), \quad 2B_n = \sum_{k=1}^n \frac{1}{k} = \log n + \gamma + o(1).$$

ここで  $o(1)$  は  $n \rightarrow \infty$  で 0 に収束する項を意味する. ゆえに,

$$A_n = \log 2 + \frac{1}{2} \log n + \frac{\gamma}{2} + o(1), \quad B_n = \frac{1}{2} \log n + \frac{\gamma}{2} + o(1).$$

これより,  $n \rightarrow \infty$  で

$$A_n - B_n \rightarrow \log 2.$$

すなわち,

$$\frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \log 2.$$

より一般に正の整数  $p, q$  について

$$A_{pn} = \log 2 + \frac{1}{2}(\log p + \log n) + \frac{\gamma}{2} + o(1), \quad B_n = \frac{1}{2}(\log q + \log n) + \frac{\gamma}{2} + o(1)$$

なので,  $n \rightarrow \infty$  のとき

$$A_{pn} - B_{qn} \rightarrow \log 2 + \frac{1}{2} \log \frac{p}{q}.$$

例:  $p = 2, q = 1$  のとき

$$\frac{1}{1} + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \cdots = \frac{3}{2} \log 2.$$

$p = 1, q = 2$  のとき

$$\frac{1}{1} - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \cdots = \frac{1}{2} \log 2. \quad \square$$

考え方: 以上のような計算が可能になったのは,

$$\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n}$$

が大雑把に  $\log n$  で近似できることだけではなく, より精密に

$$\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n} = \log n + \gamma + o(1)$$

の形で漸近挙動がわかっているからである.

```
In [11]: 1 A(n) = sum(k->1/(2k-1), 1:n)
2 B(n) = sum(k->1/(2k), 1:n)
3
4 n = 10^7
5 @show S = A(2n) - B(n)
6 @show T = 3/2*log(2)
7 @show S - T;
```

```
S = A(2n) - B(n) = 1.039720745839908
T = (3 / 2) * log(2) = 1.0397207708399179
S - T = -2.5000009840070447e-8
```

```
In [12]: 1 sum(k->1/(4k-3)+1/(4k-1)-1/(2k), 1:10^7) - T
```

```
Out[12]: -2.5000000736241645e-8
```

#### 1.4.3 Euler定数がガンマ函数の無限積表示に出て来ること

$s > 0$  のとき, ガンマ函数  $\Gamma(s)$  が

$$\Gamma(s) = \int_0^\infty e^{-x} x^{s-1} dx$$

によって定義され, digamma函数  $\psi(s)$  が

$$\psi(s) = \frac{d}{ds} \log \Gamma(s) = \frac{\Gamma'(s)}{\Gamma(s)} = \frac{1}{\Gamma(s)} \int_0^\infty e^{-x} x^{s-1} \log x dx$$

によって定義される.

ガンマ函数について次が成立することが知られている(Weierstrassの公式):

$$\frac{1}{\Gamma(s)} = e^{\gamma s} s \prod_{n=1}^{\infty} \left[ \left( 1 + \frac{s}{n} \right) e^{-s/n} \right].$$

この公式の証明については

- ガンマ分布の中心極限定理とStirlingの公式 (<https://genkuroki.github.io/documents/20160501StirlingFormula.pdf>)

の第8.2節を参照せよ.

Weierstrassの公式より

$$-\psi(s) = \gamma + \frac{1}{s} - \sum_{n=1}^{\infty} \left[ \frac{1}{n} - \frac{1}{n+s} \right].$$

特に

$$-\psi(1) = -\int_0^\infty e^{-x} \log x dx = \gamma = 0.5772156649015\dots$$

Euler定数  $\gamma$  は次の積分表示を持つことも示せる:

$$\gamma = \int_0^\infty \left( \frac{e^{-t}}{1-e^{-t}} - \frac{e^{-t}}{t} \right) dt.$$

この公式の証明については上の文献の第8.7節を参照.

In [13]: 1 -digamma(big(1))

Out[13]: 5.772156649015328606065120900824024310421593359399235988057672348848677267776685e-01

#### 1.4.4 Euler定数が高次元単体上の一様分布に関連して自然に出て来ること

$r > 0$  に対して,  $n-1$  次元単体  $\Delta^{n-1}(r)$  が

$$\Delta^{n-1}(r) = \{(x_1, \dots, x_n) \in \mathbb{R}_{\geq 0}^n \mid x_1 + \dots + x_n = r\}$$

と定義される. ここで  $\mathbb{R}_{\geq 0}$  は非負の実数全体の集合を表す. 例えば,  $n=3$  のとき

$$\Delta^2(r) = \{(x, y, z) \in \mathbb{R}^3 \mid x, y, z \geq 0, x+y+z=r\}.$$

は点  $(r, 0, 0), (0, r, 0), (0, 0, r)$  を頂点とする正三角形になる.

$\Delta^{n-1}(n)$  上の一様分布(例えば  $n=3$  の場合には正三角形上に一様に分布する確率分布)を考える.

そのとき,  $\Delta^{n-1}(n)$  の点の第  $i$  座標  $x_i$  の分布(例えば  $x_1$  の分布)は  $n \rightarrow \infty$  のとき確率密度函数が  $\exp(-x)$  の指数分布に近付くことを証明できる.

ゆえに,  $n \rightarrow \infty$  で  $\log x_i$  の期待値は  $\int_0^\infty e^{-x} \log x dx = \psi(1) = -\gamma = -0.5772\dots$  に近付く.

このような形でも自然にEuler定数  $\gamma$  が自然に出て来る.

このような現象は「統計力学」ではよく知られている. 詳しくは次のノートを参照せよ:

- Kullback-Leibler情報量とSanovの定理 (<https://genkuroki.github.io/documents/20160616KullbackLeibler.pdf>)

**問題:** 次のセルの `rand_Uniform_on_Simplex` フィルターがどうして  $\Delta^{n-1}(1)$  上の一様分布に従う擬似乱数を発生させる函数になっているかを説明せよ.

**ヒント:**  $(x_1, \dots, x_n)$  を  $(x_1, x_1+x_2, \dots, x_1+\dots+x_{n-1})$  に対応させる写像を考えよ. その写像を通して, 単体  $\Delta^{n-1}(1)$  上の一様分布と  $n-1$  次元立方体  $[0, 1]^{n-1}$  上の一様分布をどのように関係付けられるかについて考えよ.  $n-1$  次元立方体はその成分の並び方の順序で分類することによって  $(n-1)!$  個の部分集合に分割される(境界の重なりは無視してよい). □

```

In [14]: 1 ▼ # r = 1 の場合の n-1 次元単体上の一様分布に従う擬似乱数を発生させる函数
2 ▼ function rand_Uniform_on_Simplex(n)
3     local Z = sort(rand(n-1))
4     return vcat(Z, 1.0) - vcat(0.0, Z)
5 end
6
7 ▼ function rand_Uniform_on_Simplex(n, N)
8     local X = zeros(n, N)
9     for j in 1:N
10        X[:,j] = rand_Uniform_on_Simplex(n)
11    end
12    return X
13 end
14
15 # x_1 の分布と log x_i 達の平均の分布をプロットする函数
16 ▼ function plot_Uniform_on_Simplex(n; N=10^4)
17     X = n*rand_Uniform_on_Simplex(n, N)
18     ElogX = vec(mean(log.(X), 1))
19
20     p1 = plot(xlabel="x", xlim=(0, 7.5), title="n = $n")
21     histogram!(p1, X[1,:], normed=true, alpha=0.5, label="x_1")
22     plot!(p1, x->exp(-x), 0, 7.5, lw=2, label="exp(-x)")
23
24     p2 = plot(xlabel="x", legend=:topleft, legendfontsize=7)
25     xlims!(-euler_gamma-0.3, -euler_gamma+0.3)
26     histogram!(p2, ElogX, normed=true, alpha=0.5, linealpha=0.1, label="mean of log x_i")
27     vline!([-euler_gamma], lw=2, label="x = $(round(Float64(-euler_gamma), 5)))"
28
29     plot(p1, p2, size=(700, 300), layout=grid(1,2, widths=[0.35, 0.65]))
30 end

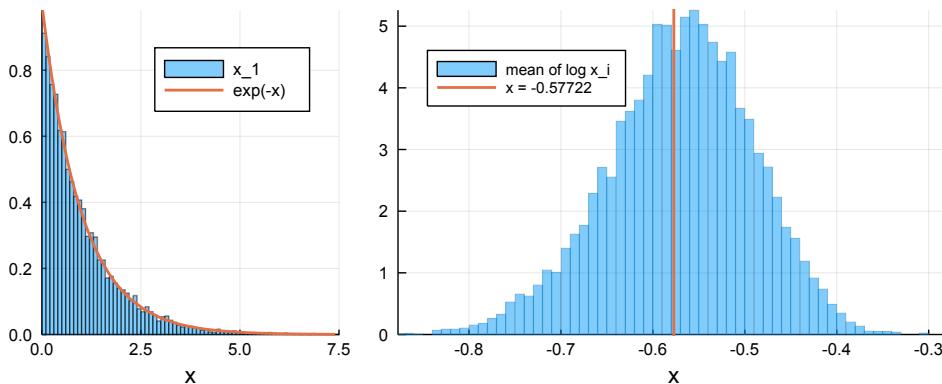
```

Out[14]: `plot_Uniform_on_Simplex` (generic function with 1 method)

In [15]: 1 `plot_Uniform_on_Simplex(100)`

Out[15]:

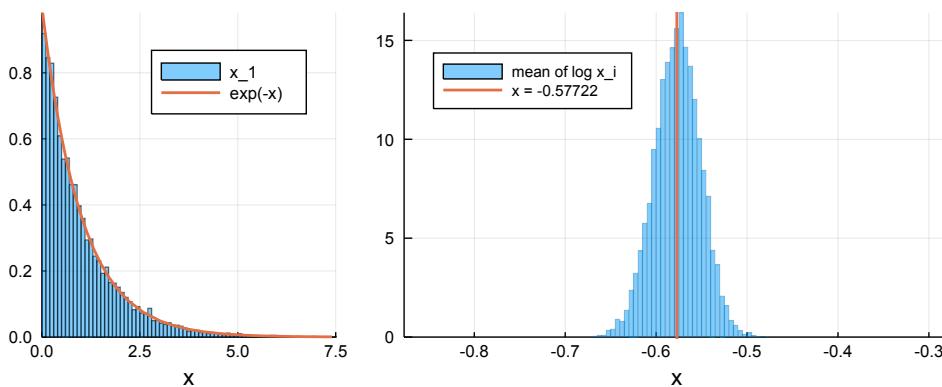
$n = 100$



In [16]: 1 `plot_Uniform_on_Simplex(1000)`

Out[16]:

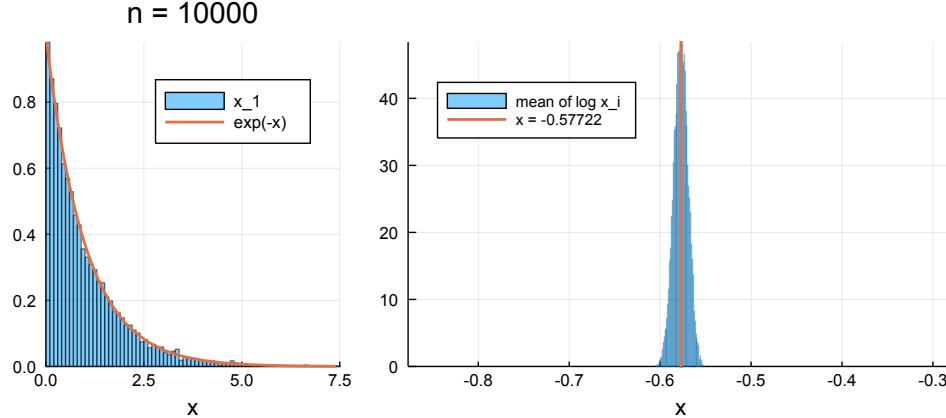
$n = 1000$



In [17]: 1 @time plot\_Uniform\_on\_Simplex(10000)

9.876582 seconds (1.04 M allocations: 5.992 GiB, 10.66% gc time)

Out[17]:



**問題:**  $n-1$  次元単体  $\Delta^{n-1}(n)$  上の一様分布に関する  $x_1$  の分布が  $n \rightarrow \infty$  で確率密度函数が  $\exp(-x)$  の指数分布に近付くことを示せ。直観的に明瞭な説明をしていれば必ずしも論理的に厳密な議論でなくてもよい。

**解答例:** 簡単のため,  $\Delta^{n-1}(r)$  を  $\mathbb{R}^{n-1} = \{(x_1, \dots, x_{n-1})\}$  への射影の像を  $D^{n-1}(r)$  と書くことにする:

$$D^{n-1}(r) = \{(x_1, \dots, x_{n-1}) \in \mathbb{R}_{\geq 0}^{n-1} \mid x_1 + \dots + x_{n-1} \leq r\}.$$

$\Delta^{n-1}(r)$  上の一様分布を  $D^{n-1}(r)$  に射影して得られる分布は  $D^{n-1}(r)$  上の一様分布になる。

$D^{n-1}(r)$  の体積は  $r^{n-1}/(n-1)!$  に等しい。そして,  $0 \leq a \leq r$  のとき,  $D^{n-1}(r)$  を  $x_1 \geq a$  に制限して得られる部分集合は  $D^{n-1}(r-a)$  に合同なのでその体積は  $(r-a)^{n-1}/(n-1)!$  に等しい。ゆえに,  $x_1 \geq a$  となる確率は  $P(x_1 \geq a)$  は

$$P(x_1 \geq a) = \frac{(r-a)^{n-1}/(n-1)!}{r^{n-1}/(n-1)!} = \left(1 - \frac{a}{r}\right)^{n-1}$$

となるので,  $r = n$  のとき,  $n \rightarrow \infty$  で

$$P(x_1 \geq a) = \left(1 - \frac{a}{n}\right)^{n-1} \rightarrow \exp(-a).$$

一方, 確率密度函数  $\exp(-x)$  において  $x \geq a$  となる確率は

$$\int_a^\infty \exp(-x) dx = [-\exp(-x)]_a^\infty = \exp(-a).$$

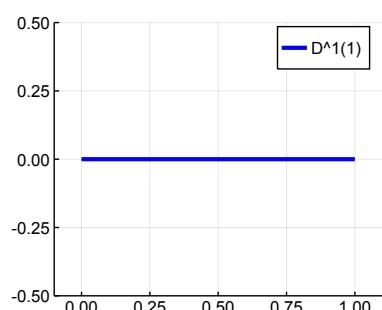
以上を比較することによって,  $D^{n-1}(n)$  上の一様分布に関する  $x_1$  の分布は  $n \rightarrow \infty$  で確率密度函数  $\exp(-x)$  の指数分布に近付くことがわかる。□

**注意:** 以上の議論において  $D^{n-1}(r)$  に射影したことは本質的ではなく,  $\Delta^{n-1}(r)$  の体積(面積)が  $r^{n-1}$  に比例することだけを使っても同様の議論を遂行できる。 $D^{n-1}(r)$  に射影するメリットは比例定数が  $1/(n-1)!$  になることが易しく求まることに過ぎない。その比例定数は確率を求めるときに分子分母でキャンセルして消えるので議論の本質には何も影響しない。□

**問題:**  $D^1(r), D^2(r), D^3(r)$  を図示し, それぞれの長さ, 面積, 体積がそれぞれ  $r, r^2/2, r^3/6$  であることを確認せよ。□

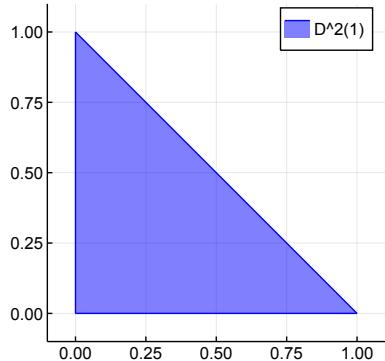
In [18]: 1 plot(size=(300,300), aspect\_ratio=1)  
2 plot!([0,1], [0,0], lw=3, color=:blue, label="D^1(1)")  
3 plot!(xlims=(-0.1,1.1), ylims=(-0.5,0.5))

Out[18]:

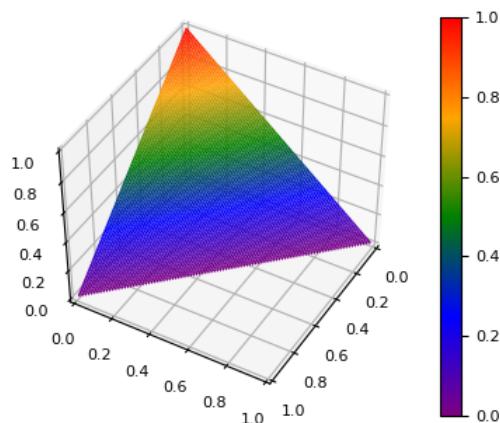


```
In [19]: 1 plot(size=(300,300), aspect_ratio=1)
2 plot!([0,1], [1,0], color=:blue, fill=(0, 0.5), label="D^2(1)")
3 plot!([0,0], [0,1], color=:blue, label="")
4 plot!([0,1], [0,0], color=:blue, label="")
5 plot!(xlims=(-0.1,1.1), ylims=(-0.1,1.1))
```

Out[19]:



```
In [20]: 1 x = 0:0.01:1
2 y = 0:0.01:1
3 f(x,y) = ifelse(x+y>1, NaN, 1-x-y)
4
5 pyplot()
6 p = plot(size=(350, 300), zlims=(0,1))
7 surface!(x, y, f, color=:rainbow, camera=(120, 45))
8 display(p)
9 gr();
```



#### 1.4.5 $\zeta(s) - 1/(s-1)$ の $s \rightarrow 0$ での極限が Euler 定数に等しいこと

以下ではこの先に進まないと知らないはずのことを勝手に用いる。

**問題:**  $h > 0$  のとき

$$\sum_{n=1}^{\infty} \left( \frac{1}{n^h} - \frac{1}{(n+1)^h} \right) = 1.$$

となることを示せ.  $\square$

解答略.

**問題:**  $h > 0$  ならば,  $n \rightarrow \infty$  で

$$\frac{1}{n^{1+h}} - \frac{1}{h} \left( \frac{1}{n^h} + \frac{1}{(n+1)^h} \right) = O\left(\frac{1}{n^{2+h}}\right)$$

となることを示せ.

**解答例:** 二項展開より

$$\left(1 + \frac{1}{n}\right)^{-h} = 1 - \frac{h}{n} + \frac{h(h+1)}{2n^2} + \dots = 1 - \frac{h}{n} + O\left(\frac{1}{n^2}\right)$$

なので

$$(左辺) = \frac{1}{hn^h} \left( \frac{h}{n} - 1 + \left( 1 + \frac{1}{n} \right)^{-h} \right)$$

$$= \frac{1}{hn^h} \left( \frac{h}{n} - 1 + \frac{h}{n} + O\left(\frac{1}{n^2}\right) \right) = O\left(\frac{1}{n^{2+h}}\right). \quad \square$$

**定義:**  $s > 1$  に対して,  $\zeta(s)$  を

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

と定める。この級数は絶対収束している。

**問題:**  $h > 0$  ならば

$$\zeta(1+h) - \frac{1}{h} = \sum_{k=0}^{\infty} \frac{(-h)^k}{k!} \gamma_k, \quad \gamma_k = \sum_{n=1}^{\infty} \left( \frac{(\log n)^k}{n} + \frac{(\log n)^{k+1} - (\log(n+1))^{k+1}}{k+1} \right)$$

となることを示せ.  $\square$

解答略.

**ヒント:** この問題の左辺は上の問題の左辺の  $n$  に関する総和に等しい.  $1/n^h = \exp(-h\log n)$  と  $1/(n+1)^h = \exp(-h\log(n+1))$  に  $\exp(x) = \sum_{k=1}^{\infty} \frac{x^k}{k!}$  を適用し,  $-h$  に関する同次の項をまとめて整理せよ. そのとき  $1/n^{1+h}$  は  $\exp(-h\log n)/n$  だとみなせ. □

**問題:** 上の問題の  $\gamma_0$  が Euler の定数  $\gamma$  に等しいことを示せ。

**解答例：**

$$\gamma_0 = \sum_{n=1}^{\infty} \left( \frac{1}{n} + \log n - \log(n+1) \right)$$

$$= \left( \frac{1}{1} - \log 2 \right) + \left( \frac{1}{2} + \log 2 - \log 3 \right) + \left( \frac{1}{3} + \log 3 - \log 4 \right) + \dots$$

であり、この級数の  $n = N$  までの和は

$$\frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{N} - \log(N+1)$$

に等しいので、その級数はEuler定数  $\gamma$  に収束する。□

以上によって

$$\zeta(1+h) = \frac{1}{h} + \gamma + O(h) \quad (h \rightarrow 0)$$

であることが示された。

**問題:**  $\zeta(1+h) = \frac{1}{h} + \gamma + O(h)$  ( $h \rightarrow 0$ ) を以上とは別の方法を使って示してみよ。□

例えばEuler-Maclaurinの和公式を使っても証明可能である。

**問題:**  $h \searrow 0$  のとき,  $\zeta(1+h) - \frac{1}{h} \rightarrow \gamma$  となることを数値的に確認せよ.  $\square$

In [ ]: 1

# 連続函数 ¶

黒木玄

2018-04-20

- Copyright 2018 Gen Kuroki
- License: MIT <https://opensource.org/licenses/MIT> (<https://opensource.org/licenses/MIT>)

このファイルは次の場所できれいに閲覧できる:

- <http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/04%20continuous%20functions.ipynb>  
<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/04%20continuous%20functions.ipynb>

このファイルは [Julia Box](https://juliabox.com) (<https://juliabox.com>) で利用できる。

自分のパソコンにJulia言語 (<https://julialang.org/>) をインストールしたい場合には

- [WindowsへのJulia言語のインストール](http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f) (<http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f>)

を参照せよ。

論理的に完璧な説明をするつもりはない。細部のいい加減な部分は自分で訂正・修正せよ。

## 目次

### [1 連続函数](#)

#### [1.1 連続函数の定義](#)

#### [1.2 中間値の定理](#)

#### [1.3 不動点定理](#)

#### [1.4 閉区間に上の実数値連続函数は最大最小を持つ](#)

#### [1.5 例](#)

##### [1.5.1 例1: コンパクト台を持つ滑らかな連続函数](#)

##### [1.5.2 例2: 無限に振動する不連続函数](#)

##### [1.5.3 例3: 無限に振動する連続函数](#)

##### [1.5.4 例4: Weierstrass函数](#)

##### [1.5.5 例5: 高木函数](#)

### [2 一様連續性](#)

#### [2.1 閉区間に上の連続函数が一様連續であること](#)

#### [2.2 連続だが一様連續ではない函数の例](#)

### [3 一様収束](#)

#### [3.1 函数列の各点収束の定義](#)

#### [3.2 函数列の一様収束の定義](#)

#### [3.3 連続函数列の一様収束先も連続になること](#)

#### [3.4 Peano曲線](#)

In [1]:

```
1  using Plots
2  ▼ gr(); ENV["PLOTS_TEST"] = "true"
3  #clibrary(:colorcet)
4  clibrary(:misc)
5
6  function pngplot(P...; kwargs...)
7      sleep(0.1)
8      pngfile = tempname() * ".png"
9      savefig(plot(P...; kwargs...), pngfile)
10     showimg("image/png", pngfile)
11 end
12 pngplot(; kwargs...) = pngplot(plot!(; kwargs...))
13
14 showing(mime, fn) = open(fn) do f
15     base64 = base64encode(f)
16     display("text/html", """""")
17 end
18
19 using SymPy
20 using SpecialFunctions
21 using QuadGK
```

## 連続函数

### 連続函数の定義

函数  $f(x)$  が  $x = \alpha$  で連続であるとは,  $x \rightarrow \alpha$  のとき  $f(x) \rightarrow f(\alpha)$  が成立することであると定める.

**定理:** 以下の2つの条件は互いに同値である:

(1)  $f(x)$  は  $x = \alpha$  で連続である。(すなわち  $x \rightarrow \alpha$  のとき  $f(x) \rightarrow f(\alpha)$  となる.)

(2)  $\alpha$  に収束する任意の数列  $a_n$  について  $n \rightarrow \infty$  のとき  $f(a_n) \rightarrow f(\alpha)$ .

**証明:** (1)ならば(2)が成立することは自明なので, (2)ならば(1)を示せば十分である. (2)ならば(1)の対偶を示すために, (1)の否定を仮定する. すなわち,  $x \rightarrow \alpha$  のとき  $f(x) \rightarrow \alpha$  とならないと仮定する.

そのとき, ある  $\varepsilon > 0$  が存在して, 任意の正の整数  $n$  についてある  $a_n$  で  $|a_n - \alpha| < 1/n$  と  $|f(a_n) - f(\alpha)| \geq \varepsilon$  を満たすものが存在する. このとき,  $a_n$  は  $\alpha$  に収束しているが,  $f(a_n)$  は  $f(\alpha)$  に収束していない.

以上によって(1)の否定から(2)の否定が導かれることがわかった. これで示すべきことがすべて示された.  $\square$

## 中間値の定理

**中間値の定理:**  $f$  は閉区間  $[a, b]$  上の実数値連続函数であるとする. このとき  $f(a)$  と  $f(b)$  のあいだにある任意の実数  $\alpha$  に対して, ある  $\xi \in [a, b]$  で  $f(\xi) = \alpha$  を満たすものが存在する.  $\square$

**補足:**  $x$  が  $a$  と  $b$  のあいだにあるとは  $a \leq x \leq b$  または  $a \geq x \geq b$  が成立していることだと定める.  $\square$

**注意:** 中間値の定理において,  $f(a) < \alpha < f(b)$  または  $f(a) > \alpha > f(b)$  ならば  $f(\xi) = \alpha$  をみたす  $\xi \in [a, b]$  は  $a < \xi < b$  となるように取れる.  $\square$

**中間値の定理の証明:**  $f(x)$  の代わりに  $f(x) - \alpha$  を考えることによって,  $\alpha = 0$  の場合に証明すれば十分である. 必要なら  $f(x)$  の代わりに  $-f(x)$  を考えることによって,  $f(a) \leq 0, f(b) \geq 0$  の場合に証明すれば十分である. 以下, そのように仮定する.

$(a_1, b_1) = (a, b)$  とおく.  $n$  について帰納的に  $f(a_n) \leq 0 \leq f(b_n)$  を満たす  $(a_n, b_n)$  を次のように定めることができる.

- $f((a_n + b_n)/2) \geq 0$  ならば  $(a_{n+1}, b_{n+1}) = (a_{n+1}, (a_n + b_n)/2)$ .
- $f((a_n + b_n)/2) < 0$  ならば,  $(a_{n+1}, b_{n+1}) = ((a_n + b_n)/2, b_n)$ .

このとき,  $a_1 \leq a_2 \leq \dots \leq b_2 \leq b_1$  かつ  $|a_n - b_n| = |a - b|/2^{n-1}$  であることから,  $a_n$  と  $b_n$  は同一の値  $\xi$  に収束する. そのとき  $f(a_n)$  と  $f(b_n)$  も同一の値  $f(\xi)$  に収束し,  $f(a_n) \leq 0 \leq f(b_n)$  なのでその収束先は 0 でなければいけない. ゆえに  $f(\xi) = 0$ .  $\square$

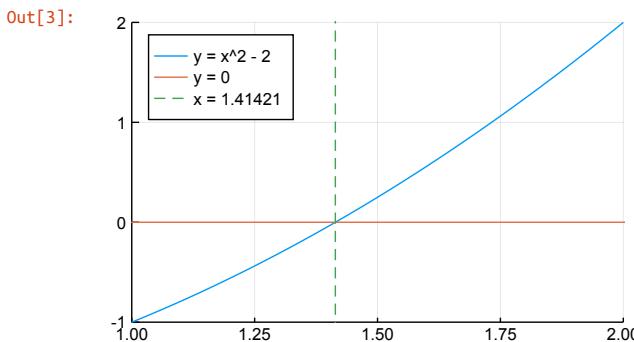
**注意:** 上の証明法の**二分法 (bisection method)**と呼ぶことがある. 実際の数値計算でも利用可能な方法である.  $\square$

```
In [2]: 1 function bisection(f, a, b; atol=eps(), maxiter=10^3)
2     f(a)*f(b) > 0 && return NaN
3     a_n, b_n = a, b
4     k = 0
5     while abs(a_n - b_n) > atol
6         k > maxiter && return NaN
7         c = (a_n+b_n)/2
8         if f(b_n)*f(c) ≥ 0
9             a_n, b_n = a_n, c
10        else
11            a_n, b_n = c, b_n
12        end
13        k += 1
14    end
15    return (a_n+b_n)/2
16 end
```

Out[2]: bisection (generic function with 1 method)

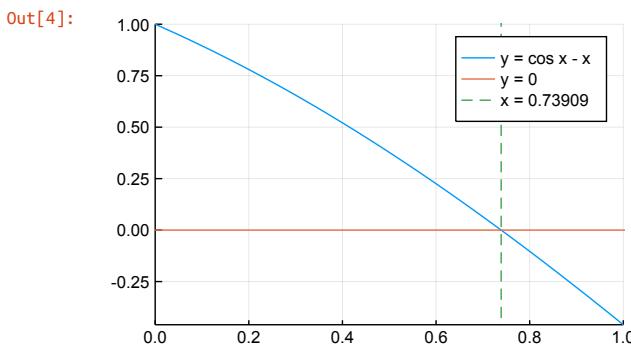
```
In [3]: 1 f(x) = x^2 - 2
2 @show ξ = bisection(f, 1, 2)
3 @show √2
4 @show round(f(ξ), 15)
5
6 x = 1:0.005:2
7 plot(size=(400,250), legend=:topleft)
8 plot!(x, f.(x), label="y = x^2 - 2")
9 ▼ hline!([0], label="y = 0")
10 ▼ vline!([ξ], label="x = $(round(ξ, 5))", ls=:dash)
```

$\xi = \text{bisection}(f, 1, 2) = 1.414213562373095$   
 $\sqrt{2} = 1.4142135623730951$   
 $\text{round}(f(\xi), 15) = -0.0$



```
In [4]: 1 f(x) = cos(x) - x
2 @show ξ = bisection(f, 0, 1)
3 @show round(f(ξ), 15)
4
5 x = 0:0.005:1
6 plot(size=(400,250))
7 plot!(x, f.(x), label="y = cos x - x")
8 ▼ hline!([0], label="y = 0")
9 ▼ vline!([ξ], label="x = $(round(ξ, 5))", ls=:dash)
```

$\xi = \text{bisection}(f, 0, 1) = 0.7390851332151606$   
 $\text{round}(f(\xi), 15) = 0.0$



## 不動点定理

**不動点定理:** 閉区間  $[a, b]$  からそれ自身への連続写像  $f$  に対して, ある  $\xi \in [a, b]$  で  $f(\xi) = \xi$  を満たすものが存在する. (そのような  $\xi$  を  $f$  の不動点と呼ぶ.)

**証明:**  $g(x) = f(x) - x$  とおく. このとき  $g$  は閉区間  $[a, b]$  上の実数値連続函数であり,  $g(a) = f(a) - a \geq 0$  かつ  $g(b) = f(b) - b \leq 0$  となり, 0 は  $g(a)$  と  $g(b)$  のあいだにある実数になる. ゆえに中間値の定理より, ある  $\xi \in [a, b]$  で  $g(\xi) = 0$  すなわち  $f(\xi) = \xi$  を満たすものが存在する.  $\square$

**補足:** 以下の形の  $\mathbb{R}$  の部分集合を区間と呼ぶ:

$$\begin{aligned} [a, b] &= \{ x \in \mathbb{R} \mid a \leq x \leq b \}, \\ [a, b) &= \{ x \in \mathbb{R} \mid a \leq x < b \}, \\ (a, b] &= \{ x \in \mathbb{R} \mid a < x \leq b \}, \\ (a, b) &= \{ x \in \mathbb{R} \mid a < x < b \}. \end{aligned}$$

$[a, b]$  は閉区間と呼ばれ,  $(a, b)$  は開区間と呼ばれる.  $(-\infty, b]$  のように無限に長い区間を扱うこともある.  $\square$

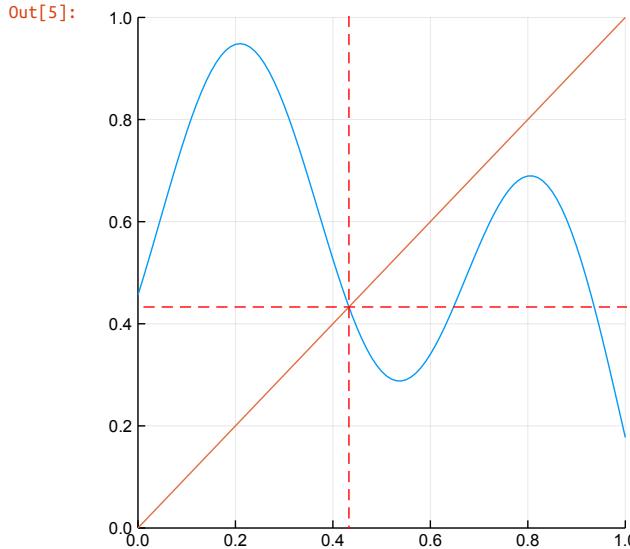
**補足:** 閉区間  $[a, b]$  からそれ自身への写像とは  $x \in [a, b]$  に  $f(x) \in [a, b]$  を対応させる函数のことである. そのグラフ  $G = \{ (x, f(x)) \mid a \leq x \leq b \}$  は正方形  $[a, b] \times [a, b] = \{ (x, y) \in \mathbb{R}^2 \mid a \leq x, y \leq b \}$  に含まれる.  $\square$

**注意:** 上の不動点定理は以下のように説明することもできる。簡単のため  $[a, b] = [-1, 1]$  の場合を考える。もしも  $f$  の不動点が存在しなかつたならば、すべての  $x \in [-1, 1]$  について  $f(x) \neq x$  となる。ゆえに  $x \in [-1, 1]$  の連続函数  $g(x)$  を

$$g(x) = \frac{x - f(x)}{|x - f(x)|}$$

と定めることができる。すべての  $x \in [-1, 1]$  について  $g(x)$  は  $\pm 1$  のどちらかになり、 $-1 < f(-1)$  なので  $g(-1) = -1$  となり、 $1 > f(1)$  なので  $g(1) = 1$  となる。しかし、そのような  $[-1, 1]$  上の連続函数は存在するはずがないので、不動点が存在するはずである。□

```
In [5]: 1 f(x) = 0.6 + 0.4x - 0.8x^2 + 0.3sin(10x-0.5)
2 ξ = bisection(x->f(x)-x, 0, 1)
3 x = 0:0.01:1
4 plot(size=(400,400), aspect_ratio=1, legend=false)
5 plot!(xlim=(0,1), ylim=(0,1))
6 plot!(x, f.(x))
7 ▼ plot!([ξ], [0,1])
8 ▼ hline!([ξ], ls=:dash, color=:red)
9 ▼ vline!([ξ], ls=:dash, color=:red)
```



**注意:** 上の易しい不動点定理はBrouwerの不動点定理の次元1の特別な場合である。□

**Brouwerの不動点定理:**  $n$  次元閉円盤

$$D = \{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1^2 + \dots + x_n^2 \leq 1 \}$$

からそれ自身への連続写像  $f$  は不動点を持つ。すなわち、ある  $\xi \in D$  で  $f(\xi) = \xi$  を満たすものが存在する。□

証明略。

**研究課題:**  $n = 2$  の場合にBrouwerの不動点定理が直観的には当然成立すべき結果であることがわかるような説明を見付けよ。□

**ヒント:** 上にある注意の方法をこの場合に適用してみよ。□

## 閉区間上の実数値連続函数は最大最小を持つ

**定理:**  $f$  は閉区間  $[a, b]$  上の実数値連続函数であるとする。このとき、実数の集合  $f([a, b]) = \{ f(x) \mid a \leq x \leq b \}$  には最大値  $M$  と最小値  $m$  が存在する。□

証明略。

**注意:** 上の定理の状況である  $\xi, \eta \in [a, b]$  で  $f(\xi) = M, f(\eta) = m$  を満たすものが存在する。□

## 例

### 例1: コンパクト台を持つ滑らかな連続函数

函数

$$f(x) = \begin{cases} 0 & (x \leq 0) \\ e^{-1/x} & (x > 0) \end{cases}$$

は  $x \searrow 0$  のとき  $e^{-1/x} \rightarrow 0$  なので連続函数である。ここで  $x \searrow 0$  は  $x$  が上から(正の値のまま)0に近付くことを意味している( $x \rightarrow +0$  と書くこともある)。 $f(x)$  は任意有限回微分可能である(そのとき  $C^\infty$  函数であるという)。

**問題:**  $n$  が非負の整数のとき,  $x \searrow 0$  のとき  $\frac{e^{-1/x}}{x^n} \rightarrow 0$  となることを示せ.  $\square$

**証明:**  $t = 1/x$  とおくと, 示すべきことは,  $t \rightarrow \infty$  のとき  $\frac{t^n}{e^t} \rightarrow 0$  となることと同値である. しかし, それは指数関数が多項式関数より速く増加することを意味しているので, 成立することがすでにわかっている.  $\square$

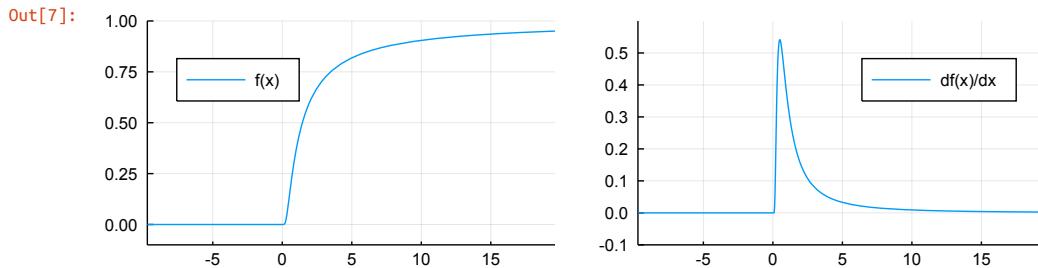
**注意:**  $e^{-1/x}$  の  $n$  階の導函数は  $\frac{e^{-1/x}}{x^n}$  に  $x$  の多項式をかけた形になる. そのことより,  $x \searrow 0$  のとき  $e^{-1/x}$  の  $n$  階の導函数も 0 に収束することがわかる.  $\square$

```
In [6]: 1 x = symbols("x")
2 diffexpneginvx(n) = diff(exp(-1/x), x, n)
3 [diffexpneginvx(n) for n in 0:3]
```

Out[6]:

$$\begin{bmatrix} e^{-\frac{1}{x}} \\ \frac{e^{-\frac{1}{x}}}{x^2} \\ \frac{e^{-\frac{1}{x}}}{x^3} \left( -2 + \frac{1}{x} \right) \\ \frac{e^{-\frac{1}{x}}}{x^4} \left( 6 - \frac{6}{x} + \frac{1}{x^2} \right) \end{bmatrix}$$

```
In [7]: 1 f(x) = x > 0 ? exp(-1/x) : zero(x)
2 g(x) = x > 0 ? exp(-1/x)/x^2 : zero(x)
3
4 p1 = plot(ylims=(-0.1, 1.0), legend=:topleft)
5 plot!(f, -10, 20, label="f(x)")
6
7 p2 = plot(ylims=(-0.1, 0.6), legend=:topright)
8 plot!(g, -10, 20, label="df(x)/dx")
9
10 plot(p1, p2, size=(700,200))
```



**問題:**  $x \leq 0$  のとき最小値 0 になり,  $x \geq 1$  のとき最大値 1 になる函数で任意有限回微分可能なものが存在することを示せ.

**解答例:** 関数

$$f(x) = \begin{cases} 0 & (x \leq 0 \text{ or } 1 \geq x) \\ \exp\left(-\frac{1}{x} + \frac{1}{x-1}\right) & (0 < x < 1) \end{cases}$$

は任意有限回微分可能である.  $Z = \int_0^1 f(x) dx$  とおき,  $g(x) = f(x)/Z$  とおくと,  $\int_0^1 g(x) dx = 1$  となる. このとき, 関数

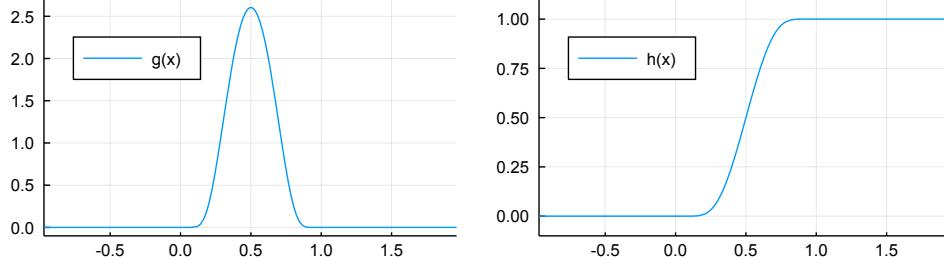
$$h(x) = \int_{-\infty}^x g(t) dt$$

は問題の条件を満たしている.  $\square$

```
In [8]: 1 f(x) = 0 < x < 1 ? exp(-1/x+1/(x-1)) : zero(x)
2 @show Z, err = quadgk(f, 0, 1) # = ∫_0^1 f(x) dx, error
3 g(x) = f(x)/Z
4 ▼ h(x) = quadgk(g, -1, x)[1] # = ∫_{-1}^x g(t) dt
5
6 p1 = plot(legend=:topleft, ylims=(-0.1, 2.7))
7 plot!(g, -1, 2, label="g(x)")
8
9 p2 = plot(legend=:topleft, ylims=(-0.1, 1.1))
10 plot!(h, -1, 2, label="h(x)")
11
12 plot(p1, p2, size=(700, 200))
```

(Z, err) = quadgk(f, 0, 1) = (0.007029858406609658, 3.23578265950559e-11)

Out[8]:



問題:  $x \leq 0$  または  $4 \leq x$  のとき最小値 0 になり,  $1 \leq x \leq 3$  のとき最大値 1 になる函数で任意有限回微分可能なものが存在することを示せ.

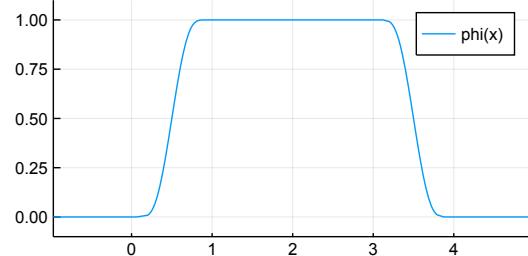
解答例: 上の問題の函数  $h(x)$  を用いて, 函数  $\phi(x)$  を

$$\phi(x) = h(x)h(4-x)$$

と定めると条件を満たしている.  $\square$

```
In [9]: 1 φ(x) = h(x)*h(4-x)
2
3 plot(size=(400, 200), ylims=(-0.1, 1.1))
4 plot!(φ, -1, 5, label="phi(x)")
```

Out[9]:



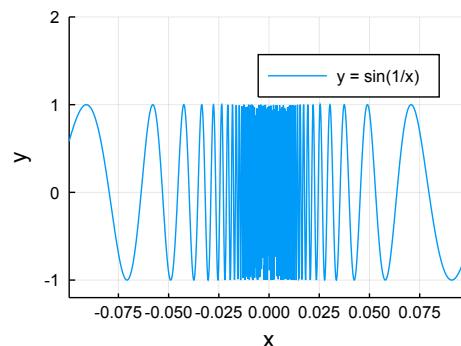
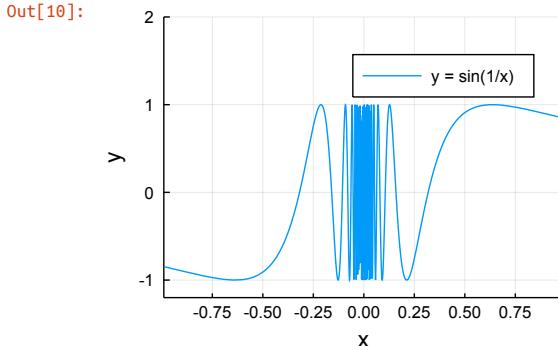
## 例2: 無限に振動する不連続函数

函数

$$f(x) = \begin{cases} \sin(1/x) & (x \neq 0) \\ 0 & (x = 0) \end{cases}$$

は  $x = 0$  で不連続である. なぜならば  $x \rightarrow 0$  のとき  $f(x)$  は振幅を保ったまま無限に振動するからである.

```
In [10]: 1 f(x) = x == 0 ? 0 : sin(1/x)
2
3 p1 = plot(size=(400, 250), xlabel="x", ylabel="y")
4 plot!(f, -1.0, 1.0, ylims=(-1.2, 2.0), label="y = sin(1/x)")
5
6 p2 = plot(size=(400, 250), xlabel="x", ylabel="y")
7 plot!(f, -0.1, 0.1, ylims=(-1.2, 2.0), label="y = sin(1/x)")
8
9 plot(p1, p2, size=(700, 250))
```



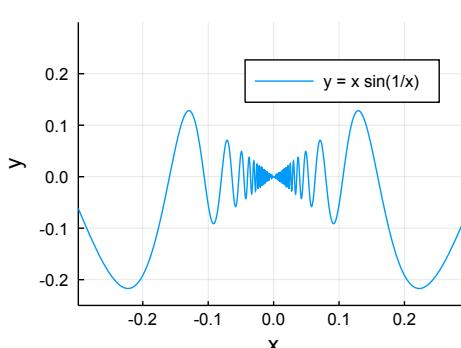
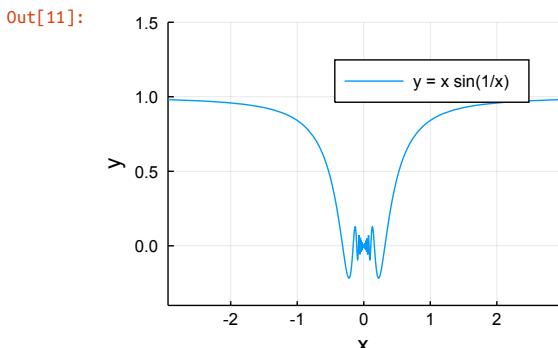
### 例3: 無限に振動する連続函数

函数

$$g(x) = \begin{cases} x \sin(1/x) & (x \neq 0) \\ 0 & (x = 0) \end{cases}$$

は  $x \rightarrow 0$  のとき無限に振動しながら 0 に収束するので連続である。

```
In [11]: 1 f(x) = x == 0 ? 0 : x*sin(1/x)
2
3 p1 = plot(size=(400, 250), xlabel="x", ylabel="y")
4 plot!(f, -3, 3, ylims=(-0.4, 1.5), label="y = x sin(1/x)")
5
6 p2 = plot(size=(400, 250), xlabel="x", ylabel="y")
7 plot!(f, -0.3, 0.3, ylims=(-0.25, 0.3), label="y = x sin(1/x)")
8
9 plot(p1, p2, size=(700, 250))
```



### 例4: Weierstrass函数

$0 < a < 1$  について, 函数

$$F(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x)$$

は連続函数であり, Weierstrass 函数と呼ばれる.  $b$  が正の奇数であり,  $ab > 1 + 3\pi/2$  のとき至るところ微分不可能なことが知られている.

- [ワイエルシュトラス関数 - Wikipedia](#)

(<https://ja.wikipedia.org/wiki/%E3%83%AF%E3%82%A4%E3%82%A8%E3%83%AB%E3%82%B7%E3%83%A5%E3%83%88%E3%83%A9%E>)

注意: 一般に, 区間  $I$  上の連続函数列  $f_n(x)$  について,  $x$  によらない非負実数列  $M_n$  で

$$|f_n(x)| \leqq M_n \quad (x \in I), \quad \sum_{n=1}^{\infty} M_n < \infty$$

を満たしているならば,

$$F(x) = \sum_{n=1}^{\infty} f_n(x) \quad (x \in I)$$

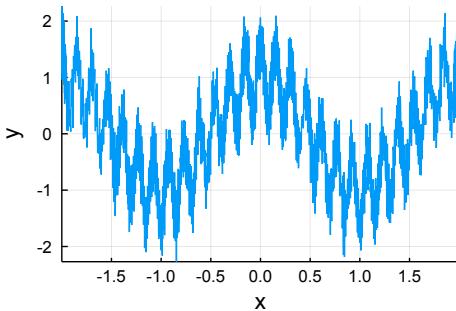
で定義される函数  $F(x)$  は  $I$  上の連續函数になる. 証明略.  $\square$

```
In [12]: 1 ▼ struct WeierstrassFunction{R<:Real, S<:Integer, T<:Integer} <: Function
2     a::R; b::S; N::T
3 end
4 (F::WeierstrassFunction)(x) = sum(n->F.a^n*cos(F.b^n*pi*x), 0:F.N)
```

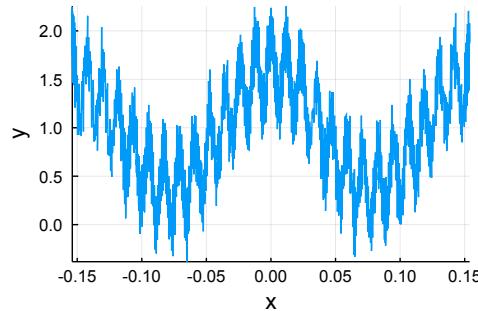
```
In [13]: 1 F = WeierstrassFunction(0.6, 13, 100)
2 @show F.a*F.b, 1+3pi/2
3
4 p1 = plot(legend=false, titlefontsize=11)
5 plot!(title="Weierstrass function: a = $(F.a), b = $(F.b)")
6 plot!(size=(600, 420), xlabel="x", ylabel="y")
7 plot!(F, -2, 2)
8
9 p2 = plot(legend=false, titlefontsize=11)
10 plot!(title="Weierstrass function: a = $(F.a), b = $(F.b)")
11 plot!(size=(600, 420), xlabel="x", ylabel="y")
12 plot!(F, -2/F.b, 2/F.b)
13
14 plot(p1, p2, size=(700, 250))
```

(F.a \* F.b, 1 + (3π) / 2) = (7.8, 5.71238898038469)

Out[13]: Weierstrass function: a = 0.6, b = 13



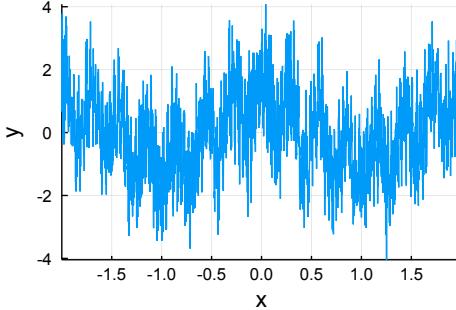
Weierstrass function: a = 0.6, b = 13



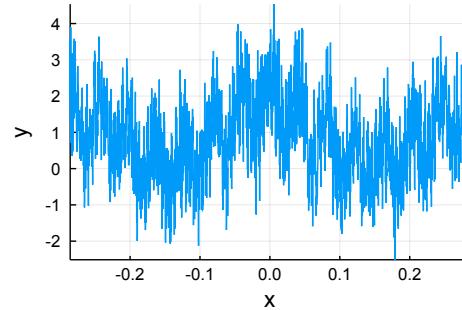
```
In [14]: 1 F = WeierstrassFunction(0.85, 7, 100)
2 @show F.a*F.b, 1+3pi/2
3
4 p1 = plot(legend=false, titlefontsize=11)
5 plot!(title="Weierstrass function: a = $(F.a), b = $(F.b)")
6 plot!(size=(600, 420), xlabel="x", ylabel="y")
7 plot!(F, -2, 2)
8
9 p2 = plot(legend=false, titlefontsize=11)
10 plot!(title="Weierstrass function: a = $(F.a), b = $(F.b)")
11 plot!(size=(600, 420), xlabel="x", ylabel="y")
12 plot!(F, -2/F.b, 2/F.b)
13
14 plot(p1, p2, size=(700, 250))
```

(F.a \* F.b, 1 + (3π) / 2) = (5.95, 5.71238898038469)

Out[14]: Weierstrass function: a = 0.85, b = 7



Weierstrass function: a = 0.85, b = 7



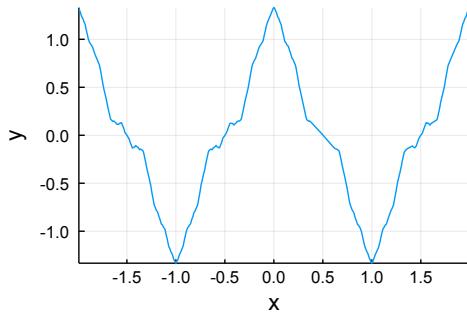
```

In [15]: 1 F = WeierstrassFunction(1/4, 3, 100)
          2 @show F.a*F.b, 1+3π/2
          3
          4 p1 = plot(legend=false, titlefontsize=11)
          5 plot!(title="Weierstrass function: a = $(F.a), b = $(F.b)")
          6 plot!(size=(600, 420), xlabel="x", ylabel="y")
          7 plot!(F, -2, 2)
          8
          9 p2 = plot(legend=false, titlefontsize=11)
          10 plot!(title="Weierstrass function: a = $(F.a), b = $(F.b)")
          11 plot!(size=(600, 420), xlabel="x", ylabel="y")
          12 plot!(F, -2/F.b, 2/F.b)
          13
          14 plot(p1, p2, size=(700, 250))

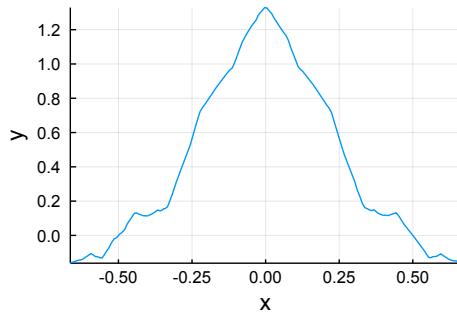
```

$$(F.a * F.b, 1 + (3\pi) / 2) = (0.75, 5.71238898038469)$$

Out[15]: Weierstrass function: a = 0.25, b = 3



Weierstrass function: a = 0.25, b = 3



#### 例5: 高木函数

$0 < w < 1$  について

$$F(x) = \sum_{n=0}^{\infty} w^n s(2^n x), \quad s(x) = \min_{k \in \mathbb{Z}} |x - k|$$

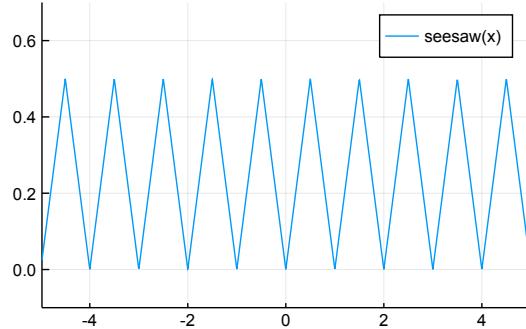
で定義される函数  $F(x)$  は連続函数である.  $w = 1/2$  のとき  $F(x)$  は高木函数と呼ばれる.

```

In [16]: 1 seesaw(x) = abs(x - round(x))
          2 plot(seesaw, size=(400, 250), ylims=(-0.1, 0.7), label="seesaw(x)")

```

Out[16]:



```

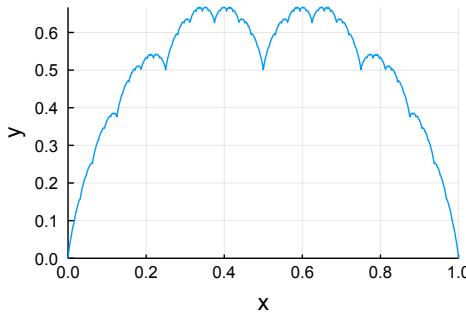
In [17]: 1 ▼ struct TakagiFunction{R<:Real, T<:Integer} <: Function
          2     w::R; N::T
          3 end
          4 (F::TakagiFunction)(x) = sum(n->F.w^n*seesaw(2^n*x), 0:F.N)

```

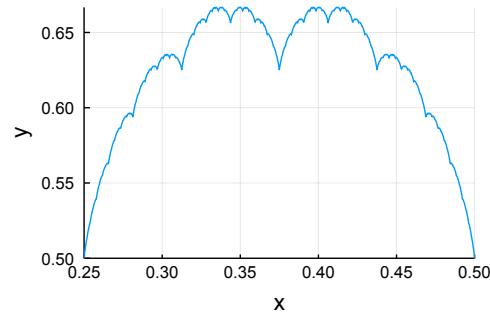
```
In [18]: 1 F = TakagiFunction(0.5, 100)
2
3 p1 = plot(legend=false, titlefontsize=11)
4 plot!(title="Takagi function: w = $(F.w)")
5 plot!(size=(600, 420), xlabel="x", ylabel="y")
6 x = 0:0.001:1
7 plot!(x, F.(x))
8
9 p2 = plot(legend=false, titlefontsize=11)
10 plot!(title="Takagi function: w = $(F.w)")
11 plot!(size=(600, 420), xlabel="x", ylabel="y")
12 x = 0.25:0.00025:0.50
13 plot!(x, F.(x))
14
15 plot(p1, p2, size=(700, 250))
```

Out[18]:

Takagi function:  $w = 0.5$



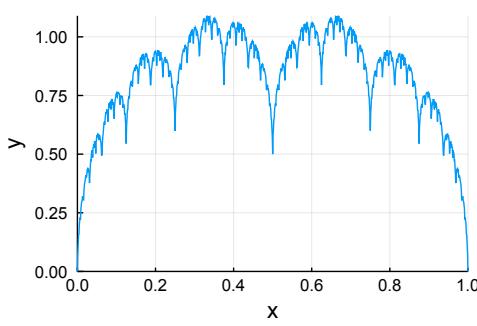
Takagi function:  $w = 0.5$



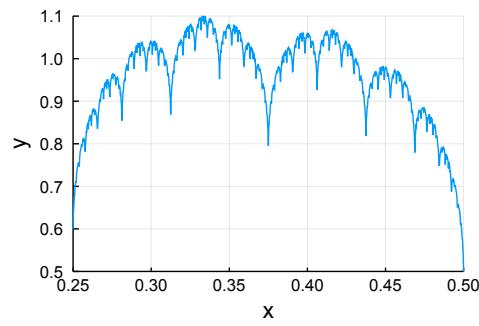
```
In [19]: 1 F = TakagiFunction(0.7, 100)
2
3 p1 = plot(legend=false, titlefontsize=11)
4 plot!(title="Takagi function: w = $(F.w)")
5 plot!(size=(600, 420), xlabel="x", ylabel="y")
6 x = 0:0.001:1
7 plot!(x, F.(x))
8
9 p2 = plot(legend=false, titlefontsize=11)
10 plot!(title="Takagi function: w = $(F.w)")
11 plot!(size=(600, 420), xlabel="x", ylabel="y")
12 x = 0.250:0.00025:0.50
13 plot!(x, F.(x))
14
15 plot(p1, p2, size=(700, 250))
```

Out[19]:

Takagi function:  $w = 0.7$

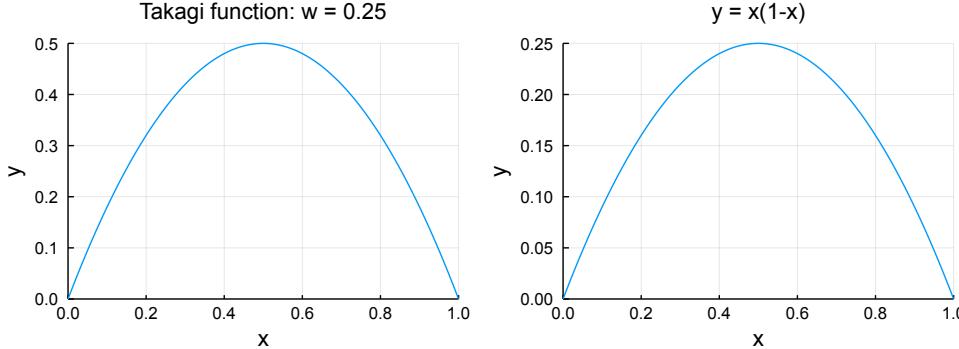


Takagi function:  $w = 0.7$



```
In [20]: 1 F = TakagiFunction(1/4, 100)
2
3 p1 = plot(legend=false, titlefontsize=11)
4 plot!(title="Takagi function: w = $(F.w)")
5 plot!(size=(600, 420), xlabel="x", ylabel="y")
6 x = 0:0.001:1
7 plot!(x, F.(x))
8
9 p2 = plot(legend=false, titlefontsize=11)
10 plot!(title="y = x(1-x)")
11 plot!(size=(600, 420), xlabel="x", ylabel="y")
12 x = 0:0.001:1
13 plot!(x, x.*(1.-x))
14
15 plot(p1, p2, size=(700, 250))
```

Out[20]:



## 一様連続性

$f(x)$  が  $x = a$  で連続であるとは、 $x \rightarrow a$  のとき  $f(x) \rightarrow f(a)$  となることであった。区間  $I$  上で定義された函数  $f(x)$  が連続であるとはすべての  $x = a \in I$  で  $f(x)$  が連続なことであると定める。

しかし、一般に、連続な  $f(x)$  について、各  $a \in I$  ごとに  $x \rightarrow a$  のときの  $f(x) \rightarrow f(a)$  の「収束の速さ」は一般に異なる。 $x \rightarrow a$  のときの  $f(x) \rightarrow f(a)$  の「収束の速さ」がすべての  $a \in I$  について同時に上からおさえられるとき、 $f(x)$  は  $I$  上で一様連続であるという。

正確には一様連続性は次のように定義される。次の条件が成立しているとき、 $f(x)$  は  $I$  上で一様連続であるという：任意に  $\epsilon > 0$  を与えたとき、ある  $\delta > 0$  が存在して、任意の  $x, a \in I$  について  $|x - a| < \delta$  ならば  $|f(x) - f(a)| < \epsilon$  が成立する。

$x \rightarrow a$  のとき  $f(x) \rightarrow f(a)$  が成立するということは、 $x$  を  $a$  に近付けることによって  $f(a)$  を  $f(x)$  でいくらでも近似できるということである。 $x \rightarrow a$  のとき  $f(x) \rightarrow f(a)$  であるということは、許される誤差  $\epsilon > 0$  を与えたとき、ある  $\delta > 0$  が存在して、 $f(x)$  と  $f(a)$  の違いを  $\epsilon$  未満におさえるためには  $x$  と  $a$  の違いを  $\delta$  未満におさえればよいということである。

単なる連続性の定義では  $\delta$  の取り方は各  $a$  ごとに違っていてもよい。 $\delta$  を小さく取らなければいけない  $a$  においては  $x \rightarrow a$  のときの  $f(x) \rightarrow f(a)$  の「収束の速さ」は「遅い」ということである。

一様連続の定義は、すべての  $a$  について共通の  $\delta > 0$  を選べることを意味している(すなわち  $\delta > 0$  は  $\epsilon > 0$  に依存するが、 $a \in I$  には依存しないようにできることを意味している)。

## 閉区間上の連続函数が一様連続であること

**定理:** 閉区間上の連続函数は一様連続である。□

証明略。

閉区間上の函数については、連続函数と一様連続函数を区別しなくてもよい。

## 連続だが一様連続ではない函数の例

**問題:**  $x > 0$  の連続函数  $f(x) = 1/x$  は区間  $(0, 1]$  上では一様連続ではないことを示せ。

**解答例:** 図を描けば、 $\epsilon > 0$  と  $a > 0$  について、

$$x > 0, |x - a| < \delta \implies |f(x) - f(a)| < \epsilon \quad (*)$$

すなわち

$$x > 0, a - \delta < x < a + \delta \implies \frac{1}{a} - \epsilon < \frac{1}{x} < \frac{1}{a} + \epsilon$$

かつ  $\delta > 0$  となるためには、

$$\delta \leq a - \frac{1}{1/a + \epsilon} = \frac{a^2 \epsilon}{1 + a\epsilon} < a^2 \epsilon$$

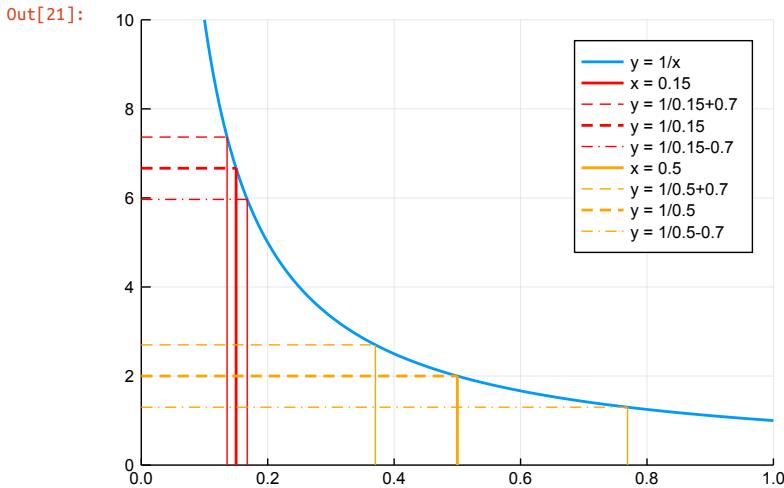
でなければいけないことがわかる。したがって  $a > 0$  を小さくすれば  $\delta > 0$  はいくらでも小さくなる。ゆえに  $a$  と無関係に(\*)を満たす  $\delta > 0$  を取ることは不可能である。つまり  $f(x)$  は  $(0, 1]$  上で一様連続ではない。□

**問題:** 任意に与えられた  $r > 0$  について,  $f(x) = 1/x$  は区間  $[r, \infty)$  上で一様連続であることを示せ。

**解答例:**  $e > 0$  あるとする。 $a^2 e / (1 + ae)$  は  $a$  について単調増加函数なので,  $a$  が  $[r, \infty)$  内を動くとき, その最小値は  $r^2 e / (1 + re)$  になる。それを  $\delta$  とおけばすべての  $a \in [r, \infty)$  について同時に (\*) が成立する。これで  $f(x) = 1/x$  が  $[r, \infty)$  上一様連続であることがわかった。□

**注意:** 微分可能函数については, 区間  $I$  上でのグラフの傾きが幾らでも大きくなる函数は  $I$  上で一様連続ではなくなるが, 区間  $I$  上でのグラフの傾きに上限がある場合には一様連続になる。□

```
In [21]: 1 f(x) = 1/x
2 x = 0.1:0.001:1
3
4 ε = 0.7
5
6 plot(xlims=(0,1), ylims=(0,10), size=(500, 350))
7 plot!(x, f.(x), label="y = 1/x", lw=2)
8
9 a = 0.15
10 δ₀ = a - 1/(1/a+ε)
11 δ₁ = 1/(1/a-ε) - a
12 c = :red
13 ▼ plot!([a-δ₀,a-δ₀], [0,f(a-δ₀)], color=c, label="")
14 ▼ plot!([a,a], [0,f(a)], color=c, lw=2, label="x = $a")
15 ▼ plot!([a+δ₁,a+δ₁], [0,f(a+δ₁)], color=c, label="")
16 ▼ plot!([0,a-δ₀], [f(a-δ₀),f(a-δ₀)], ls=:dash, color=c, label="y = 1/$a+$ε")
17 ▼ plot!([0,a], [f(a),f(a)], color=c, ls=:dash, lw=2, label="y = 1/$a")
18 ▼ plot!([0,a+δ₁], [f(a+δ₁),f(a+δ₁)], ls=:dashdot, color=c, label="y = 1/$a-$ε")
19
20 a = 0.5
21 δ₀ = a - 1/(1/a+ε)
22 δ₁ = 1/(1/a-ε) - a
23 c = :orange
24 ▼ plot!([a-δ₀,a-δ₀], [0,f(a-δ₀)], color=c, label="")
25 ▼ plot!([a,a], [0,f(a)], color=c, lw=2, label="x = $a")
26 ▼ plot!([a+δ₁,a+δ₁], [0,f(a+δ₁)], color=c, label="")
27 ▼ plot!([0,a-δ₀], [f(a-δ₀),f(a-δ₀)], ls=:dash, color=c, label="y = 1/$a+$ε")
28 ▼ plot!([0,a], [f(a),f(a)], color=c, ls=:dash, lw=2, label="y = 1/$a")
29 ▼ plot!([0,a+δ₁], [f(a+δ₁),f(a+δ₁)], ls=:dashdot, color=c, label="y = 1/$a-$ε")
```



## 一様収束

### 函数列の各点収束の定義

**各点収束の定義:** 区間  $I$  上の函数の列  $f_1(x), f_2(x), f_3(x), \dots$  が函数  $\varphi(x)$  に**各点収束**するとは, 各  $x \in I$  ごとに, 任意の  $\varepsilon > 0$  に対して, ある番号  $N$  が存在して,

$$|f_n(x) - f(x)| < \varepsilon \quad (n \geq N)$$

が成立することだと定める。

$$|f_n(x) - \varphi(x)| < \varepsilon \quad (n \geq N, x \in I)$$

が成立することだと定める。すなわち, 各  $x \in I$  ごとに,

$$f_n(x) \rightarrow \varphi(x) \quad (n \rightarrow \infty)$$

が成立するとき, 函数列  $f_n(x)$  は  $I$  上で  $\varphi(x)$  に各点収束するという. 各点収束においては, 各点  $x \in I$  ごとに  $f_n(x)$  の  $\varphi(x)$  への収束の速さは全く違っていてもよい.  $\square$

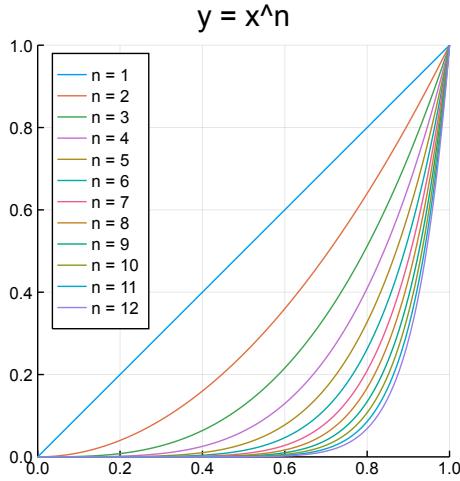
**問題:** 閉区間  $I = [0, 1]$  上で函数列  $f_n(x) = x^n$  は函数

$$\varphi(x) = \begin{cases} 0 & (0 \leq x < 1) \\ 1 & (x = 1) \end{cases}$$

に各点収束することを示せ.  $0 \leq x < 1$  が 1 に近ければ近いほど,  $f_n(x) = x^n$  の 0 に収束する速さは遅くなることに注意せよ.  $\square$

```
In [22]: 1 f(n,x) = x^n
2
3 x = 0:0.005:1
4 p1 = plot(xlims=(0,1), legend=:topleft)
5 for n in 1:12
6   plot!(x, f.(n,x), label="n = $n")
7 end
8 plot(p1, size=(350,350), aspect_ratio=1, title="y = x^n")
```

Out[22]:



## 函数列の一様収束の定義

**一様収束の定理:** 区間  $I$  上の函数の列  $f_1(x), f_2(x), f_3(x), \dots$  が函数  $\varphi(x)$  に**一様収束**するとは, 任意の  $\varepsilon > 0$  に対して, ある番号  $N$  が存在して, すべての  $x \in I$  について一斉に,

$$|f_n(x) - \varphi(x)| < \varepsilon \quad (n \geq N)$$

が成立することだと定める.  $\square$

この定義は次と同値である(証明略).

**同値な定義:** 区間  $I$  上の函数列  $f_n(x)$  が函数  $\varphi(x)$  に**一様収束**するとは,  $\forall x \in I$  に無関係な非負実数列  $c_n$  で

$$|f_n(x) - \varphi(x)| \leq c_n \rightarrow 0 \quad (n \rightarrow \infty)$$

を満たすものが存在することだと定める.  $\square$

すなわち, すべての  $x \in I$  について一斉に  $f_n(x)$  と  $\varphi(x)$  の違いの大きさを  $n \rightarrow \infty$  で一様に小さくできるときに(すなわち 0 に収束する非負実数列  $c_n$  で違いの大きさが上からおさえられるとき),  $f_n(x)$  は  $\varphi(x)$  に一様収束するという.

**問題:** 閉区間  $I = [0, 1]$  上で函数列  $f_n(x) = x^n$  は函数

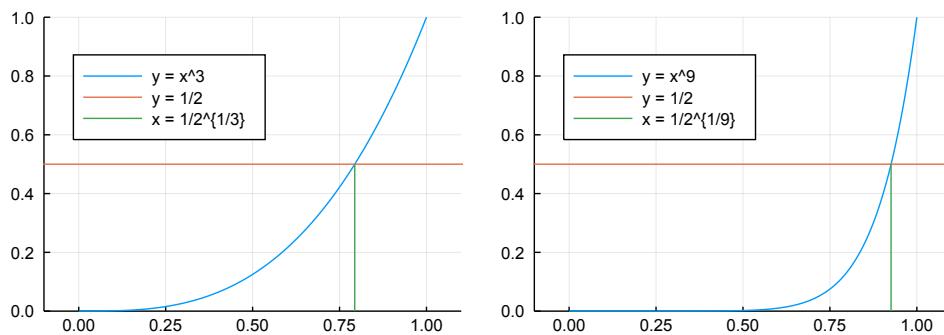
$$\varphi(x) = \begin{cases} 0 & (0 \leq x < 1) \\ 1 & (x = 1) \end{cases}$$

に一様収束しないことを示せ.

**解答例:**  $0 < x < 1$  のとき  $|x| \geq 1/2^{1/n}$  とすると,  $f_n(x) \geq 1/2$  となる. これで,  $f_n(x)$  と  $\varphi(x)$  の違いの大きさをすべての  $x \in I$  について一斉に  $1/2$  より小さくできないことがわかった.  $\square$

```
In [23]: 1 f(n,x) = x^n
2
3 x = 0:0.005:1
4
5 n = 3
6 p1 = plot(xlims=(-0.1,1.1), legend=:topleft)
7 plot!(x, f.(n,x), label="y = x^n")
8 ▼ hline!([1/2], label="y = 1/2")
9 ▼ plot!(([1/2^(1/n), 1/2^(1/n)], [0, 1/2], label="x = 1/2^{1/n}")
10
11 n = 9
12 p2 = plot(xlims=(-0.1,1.1), legend=:topleft)
13 plot!(x, f.(n,x), label="y = x^n")
14 ▼ hline!([1/2], label="y = 1/2")
15 ▼ plot!(([1/2^(1/n), 1/2^(1/n)], [0, 1/2], label="x = 1/2^{1/n}")
16
17 plot(p1, p2, size=(700, 250))
18
```

Out[23]:



### 連続函数列の一様収束先も連続になること

**定理:** 区間  $I$  上の連続な函数の列  $f_n(x)$  が函数  $\varphi(x)$  に一様収束するとき, 収束先の  $\varphi(x)$  も  $I$  上の連続函数になる.  $\square$

証明略.

**定理:** 区間  $I$  上の函数の列  $f_n(x)$  について,

$$F_N(x) = \sum_{n=1}^N |f_n(x)|$$

が  $I$  上一様収束するとき, 函数項級数  $\sum_{n=1}^{\infty} f_n(x)$  も一様収束する. このとき, その函数項級数は一様絶対収束するという.  $\square$

証明略.

**注意:** 上の定理では収束先の函数が不明であっても収束することがわかる.  $\square$

**定理:** すべての  $f_n(x)$  が区間  $I$  上の連続函数であり,  $\sum_{n=1}^{\infty} f_n(x)$  が  $I$  上一様絶対収束するならば, その収束先も連続函数になる.  $\square$

証明略.

**定理:** 区間  $I$  上の函数の列  $f_n(x)$  に対して, 非負実数列  $c_n$  で

$$|f_n(x)| \leq c_n \quad (x \in I), \quad \sum_{n=1}^{\infty} c_n < \infty$$

が成立していれば, 函数列

$$s_N(x) = \sum_{n=1}^N f_n(x)$$

は  $I$  上一様絶対収束する.  $\square$

証明略.

**注意:** 函数の数値計算で区間  $I$  の点を有限個の点  $x_i \in I, x_1 < \dots < x_L$  で代表し, 函数  $f_n(x)$  の代わりに配列  $[f_n(x_1), \dots, f_n(x_L)]$  を扱うことがよくある. そのとき収束するかどうかの判定を  $f_n(x_i)$  と  $f_{n+1}(x_i)$  の違いの大きさがすべての  $i = 1, \dots, L$  について一斉に最初に決めておいた  $\epsilon > 0$  より小さくなつたかどうかで判定することは, 一様収束に近い考え方で収束を判定していることになる.  $\square$

**問題:**  $s > 1$  の函数  $f_n(s)$  を

$$f_n(s) = \frac{1}{n^s}$$

と定める. 1より真に大きい実数  $a$  を任意に取る. このとき  $s \geq a$  において, 関数項級数

$$\sum_{n=1}^{\infty} f_n(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

が一様絶対収束することを示せ.

**解答例:** 各  $n \geq 1$  ごとに  $f_n(s) = 1/n^s$  は  $s$  に関する単調減少函数である. ゆえに

$$|f_n(s)| = \frac{1}{n^s} \leq \frac{1}{n^a} \quad (s \geq a).$$

そして,  $a > 1$  より  $\sum_{n=1}^{\infty} \frac{1}{n^a}$  は有限の値に収束する. 以上によって,  $\sum_{n=1}^{\infty} f_n(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$  は  $s \geq a$  で絶対収束する.  $\square$

**注意:** 上の問題の  $a > 1$  はいくらでも 1 に近付けることができる.

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \quad (s > 1)$$

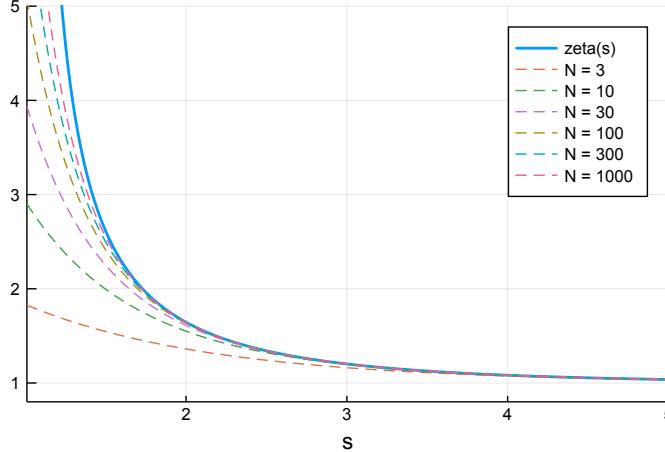
は  $s > 1$  における連続函数になることがわかる.  $\square$

**注意:** 実際にはもっとよいことが言える.  $\zeta(s)$  は  $s > 1$  でいくらでも項別微分可能であり,  $s = 1$  を除いた複素平面全体に解析接続される.  $\square$

```
In [24]: 1 g(N, s) = sum(n->1/n^s, 1:N)
2 s = 1.01:0.01:5.00
3
4 p1 = plot(title="sum of 1/n^s for n=1,...,N")
5 plot!(ylims=(0.8, 5), xlabel="s")
6 plot!(s, zeta.(s), label="zeta(s)", lw=2)
7 ▼ for N in [3,10,30,100,300,1000]
8 plot!(s, g.(N,s), label="N = $N", ls=:dash)
9 end
10 plot(p1, size=(500, 350))
```

Out[24]:

sum of  $1/n^s$  for  $n=1,\dots,N$



### Peano曲線

以下のプログラムのように  $n$  について帰納的に定義された  $\leq t \leq 1$  の連続函数  $x_n(t) = \text{Peano}_x(n, t)$  と  $y_n(t) = \text{Peano}_y(n, t)$  は  $n \rightarrow \infty$  で一様収束し, その収束先  $x(t), y(t)$  は平面上の連続曲線  $(x(t), y(t))$  ( $0 \leq t \leq 1$ ) を定める. その曲線は**Peano曲線**と呼ばれており, 正方形  $[0, 1] \times [0, 1]$  の内側を埋め尽くす.

**問題:**  $x_n(t), y_n(t)$  がともに閉区間  $[0, 1]$  上で一様収束することを示せ.

**解答例:**  $f_n(t)$  は  $x_n(t)$  または  $y_n(t)$  であるとする. そのとき

$$|f_n(t) - f_{n-1}(t)| \leq \frac{1}{2^n} \quad (0 \leq t \leq 1)$$

となっているので(下の方のプロットを参照せよ),  $0 \leq t \leq 1$  のとき

$$\sum_{n=1}^{\infty} |f_n(t) - f_{n-1}(t)| \leq \sum_{n=1}^{\infty} \frac{1}{2^n} = 1 < \infty$$

となる. ゆえに函数項級数

$$f_0(t) + (f_1(t) - f_0(t)) + (f_2(t) - f_1(t)) + (f_3(t) - f_2(t)) + \cdots$$

は閉区間  $[0, 1]$  上で一様絶対収束する。これは  $f_n(t)$  が一様収束することを意味する。□

**注意:**  $0 \leq t \leq 1$  のとき

$$f(t) = f_n(t) + \sum_{k=n}^{\infty} (f_{k+1}(t) - f_k(t)), \quad \sum_{k=n}^{\infty} |f_{k+1}(t) - f_k(t)| \leq \sum_{k=n}^{\infty} \frac{1}{2^{k+1}} = \frac{1}{2^n}$$

なので、

$$|f_n(t) - f(t)| \leq \frac{1}{2^n} \quad (0 \leq t \leq 1)$$

であることもわかる。□

**問題:**  $x_n(t), y_n(t)$  の極限をそれぞれ  $x(t), y(t)$  と書く。そのとき、任意の正方形上の点  $(a, b) \in [0, 1] \times [0, 1]$  に対して、ある実数  $t \in [0, 1]$  で  $(x(t), y(t)) = (a, b)$  を満たすものが存在することを示せ。

**解答例:**  $x_n(t), y_n(t)$  の作り方より(下の方のプロットを参照せよ)、任意の正の整数  $n$  に対して、ある  $t_n \in [0, 1]$  で

$$|x_n(t_n) - a| \leq \frac{1}{2^n}, \quad |y_n(t_n) - b| \leq \frac{1}{2^n}$$

を満たすものが存在する。数列  $t_n$  は閉区間  $[0, 1]$  上の数列なので、そのある部分列  $t_{n_k}$  ( $1 \leq n_1 < n_2 < \dots$ ) で収束するものが存在する。その収束先を  $t \in [0, 1]$  と書く。任意に  $\varepsilon > 0$  を取る。 $x(t), y(t)$  の連続性より、十分に  $k$  を大きくすると

$$|x(t) - x(t_{n_k})| \leq \varepsilon, \quad |y(t) - y(t_{n_k})| \leq \varepsilon.$$

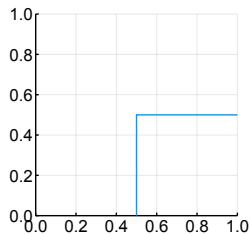
ゆえに、以上と上の注意の内容を合わせると、 $k \rightarrow \infty$  で、

$$\begin{aligned} |x(t) - a| &\leq |x(t) - x(t_{n_k})| + |x(t_{n_k}) - x_{n_k}(t_{n_k})| + |x_{n_k}(t_{n_k}) - a| \leq \varepsilon + \frac{1}{2^{n_k}} + \frac{1}{2^{n_k}} \rightarrow \varepsilon, \\ |y(t) - b| &\leq |y(t) - y(t_{n_k})| + |y(t_{n_k}) - y_{n_k}(t_{n_k})| + |y_{n_k}(t_{n_k}) - b| \leq \varepsilon + \frac{1}{2^{n_k}} + \frac{1}{2^{n_k}} \rightarrow \varepsilon. \end{aligned}$$

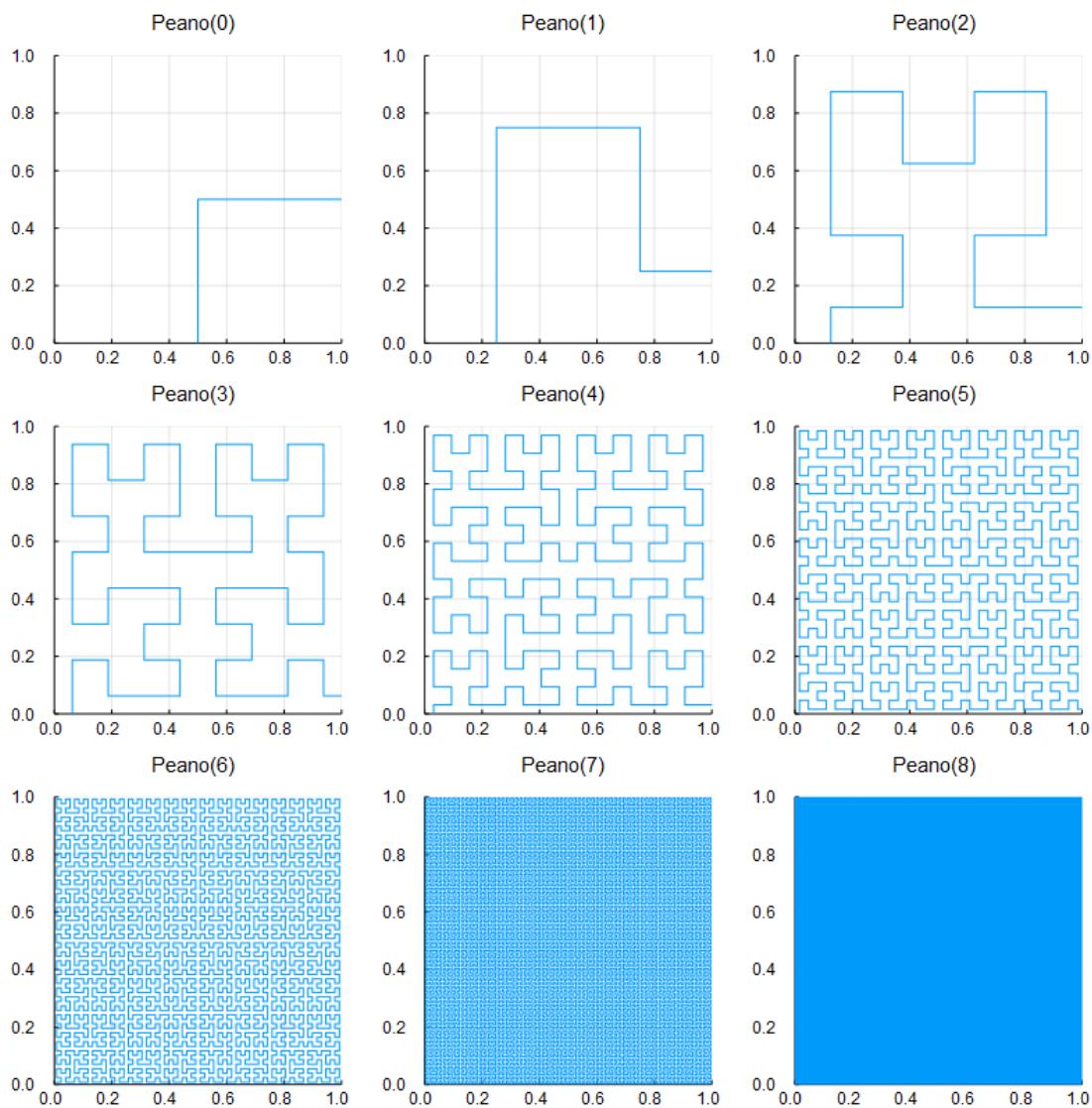
$\varepsilon > 0$  はいくらでも小さくできるので、 $(x(t), y(t)) = (a, b)$ 。□

```
In [25]: 1 peano_x₀(t) = t < 1/2 ? 1/2 : t
2 peano_y₀(t) = t < 1/2 ? t : 1/2
3
4 n = 0
5 t = 0:1/(2*4^n):1
6 x = peano_x₀.(t)
7 y = peano_y₀.(t)
8 plot(size=(200,200))
9 plot!(title="Peano($n)", titlefontsize=10)
10 plot!(xlims=(0,1), ylims=(0,1))
11 plot!(aspect_ratio=1, legend=:false)
12 plot!(x, y)
```

Out[25]: Peano(0)

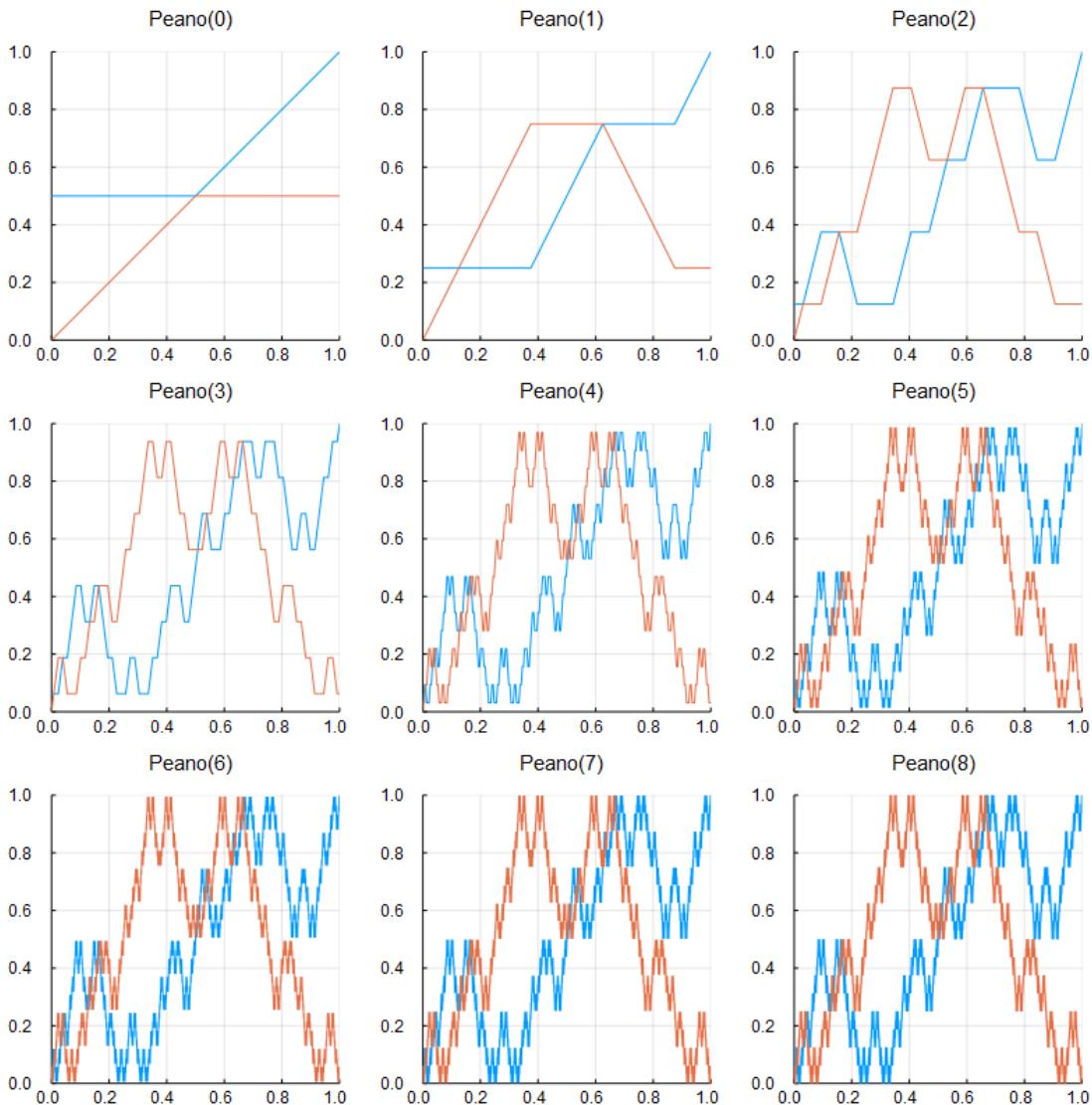


```
In [26]: 1 function peano_x(n, t)
2     if n == 0
3         peano_x_0(t)
4     else
5         if t < 1/(2*4^n)
6             1/(2*2^n)
7         elseif t < 1/4
8             peano_y(n-1, 4t)/2
9         elseif t < 2/4
10            peano_x(n-1, 4t-1)/2
11        elseif t < 3/4
12            (1-peano_x(n-1, 3-4t))/2 + 1/2
13        else
14            (1-peano_y(n-1, 4-4t))/2 + 1/2
15        end
16    end
17
18
19 function peano_y(n, t)
20     if n == 0
21         peano_y_0(t)
22     else
23         if t < 1/(2*4^n)
24             2^n*t
25         elseif t < 1/4
26             peano_x(n-1, 4t)/2
27         elseif t < 2/4
28             peano_y(n-1, 4t-1)/2 + 1/2
29         elseif t < 3/4
30             peano_y(n-1, 3-4t)/2 + 1/2
31         else
32             peano_x(n-1, 4-4t)/2
33         end
34     end
35
36
37 P = Plots.Plot[]
38 for n in 0:8
39     t = 0:1/(2*4^n):1
40     x = peano_x.(n,t)
41     y = peano_y.(n,t)
42     p = plot()
43     plot!(title="Peano($n)", titlefontsize=10)
44     plot!(xlims=(0,1), ylims=(0,1))
45     plot!(aspect_ratio=1, legend=:false)
46     plot!(x, y)
47     push!(P, p)
48 end
49 plot(P..., size=(800,800), layout=@layout(grid(3,3)))
50 pngplot()
```



In [27]:

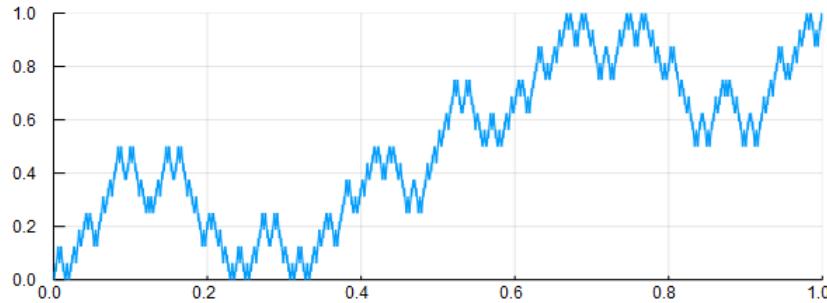
```
1 P = Plots.Plot[]
2 for n in 0:8
3     t = 0:1/(2*4^n):1
4     x = peano_x.(n,t)
5     y = peano_y.(n,t)
6     p = plot()
7     plot!(title="Peano($n)", titlefontsize=10)
8     plot!(xlims=(0,1), ylims=(0,1))
9     plot!(aspect_ratio=1, legend=:false)
10    plot!(t, x)
11    plot!(t, y)
12    push!(P, p)
13 end
14 plot(P..., size=(800,800), layout=@layout(grid(3,3)))
15 pngplot()
```



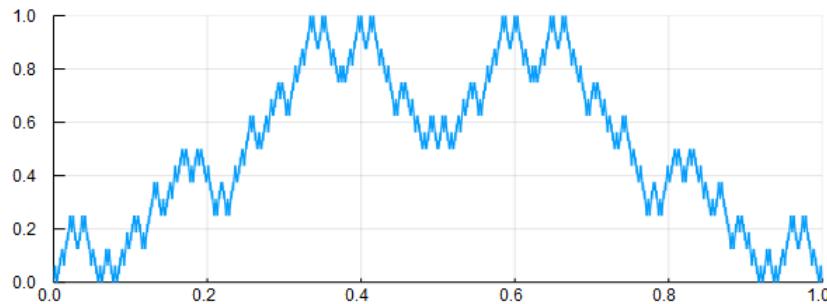
In [28]:

```
1 n = 8
2 px = plot()
3 py = plot()
4 t = 0:1/(2*4^n):1
5 x = peano_x.(n,t)
6 y = peano_y.(n,t)
7 plot!(px, title="Peano_x($n)", titlefontsize=10)
8 plot!(py, title="Peano_y($n)", titlefontsize=10)
9 plot!(px, xlims=(0,1), ylims=(0,1))
10 plot!(py, xlims=(0,1), ylims=(0,1))
11 plot!(px, legend=:topleft)
12 plot!(py, legend=:topleft)
13 plot!(px, t, x, label="$n", lw=0.3)
14 plot!(py, t, y, label="$n", lw=0.3)
15 plot(px, py, size=(600,500), layout=@layout(grid(2,1)), legend=false)
16 pngplot()
```

Peano\_x(8)

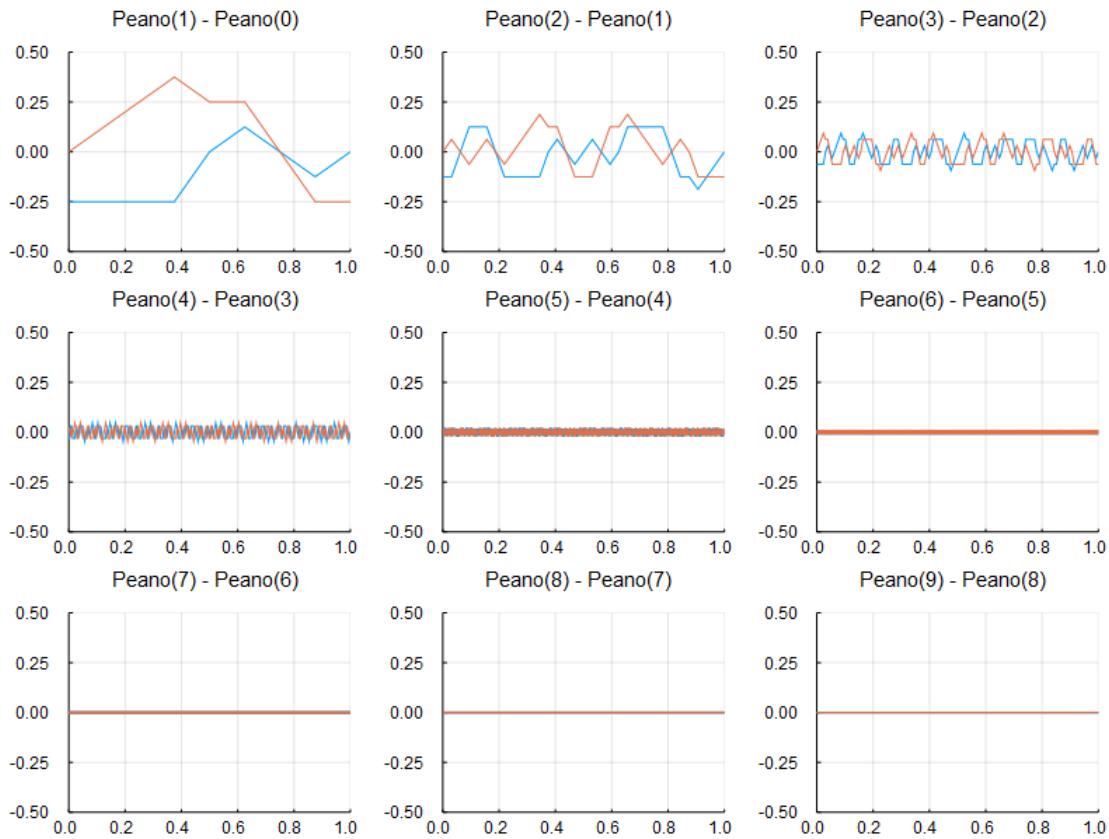


Peano\_y(8)



In [29]:

```
1 P = Plots.Plot[]
2 for n in 1:9
3     t = 0:1/(2*4^n):1
4     dx = peano_x.(n,t) - peano_x.(n-1,t)
5     dy = peano_y.(n,t) - peano_y.(n-1,t)
6     p = plot()
7     plot!(title="Peano($n) - Peano($(n-1))", titlefontsize=10)
8     plot!(xlims=(0,1), ylims=(-1/2,1/2))
9     plot!(!legend=:false)
10    plot!(t, dx, lw=0.7)
11    plot!(t, dy, lw=0.7)
12    push!(P, p)
13 end
14 plot(P..., size=(800,600))
15 pngplot()
```



In [ ]:

```
1
```

# 微分可能函数

黒木玄

2018-04-20

- Copyright 2018 Gen Kuroki
- License: MIT <https://opensource.org/licenses/MIT> (<https://opensource.org/licenses/MIT>)

このファイルは次の場所できれいに閲覧できる:

- <http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/05%20differentiable%20functions.ipynb>  
<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/05%20differentiable%20functions.ipynb>

このファイルは [Julia Box](https://juliabox.com) (<https://juliabox.com>) で利用できる。

自分のパソコンに [Julia言語](https://julialang.org/) (<https://julialang.org/>) をインストールしたい場合には

- [WindowsへのJulia言語のインストール](http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f) (<http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f>)

を参照せよ。

論理的に完璧な説明をするつもりはない。細部のいい加減な部分は自分で訂正・修正せよ。

## 目次

<a href="#">1 Landau記号</a>
<a href="#">1.1 数列版</a>
<a href="#">1.2 連続版</a>
<a href="#">2 Landau記号を使った微分可能性の定義</a>
<a href="#">2.1 連続函数</a>
<a href="#">2.2 微分可能函数</a>
<a href="#">2.3 <math>C^n</math> 級函数</a>
<a href="#">3 微分の基本性質</a>
<a href="#">3.1 Leibnitz則</a>
<a href="#">3.2 合成函数の微分則</a>
<a href="#">4 平均値の定理</a>
<a href="#">4.1 微分可能函数の平均値の定理</a>
<a href="#">4.2 <math>C^1</math> 級函数の積分版平均値の定理</a>
<a href="#">4.3 Cauchyの平均値の定理</a>

```
In [1]: 1  using Plots
2  gr(); ENV["PLOTS_TEST"] = "true"
3  #clibrary(:colorcet)
4  clibrary(:misc)
5
6  ▼ function pngplot(P...; kwargs...)
7    sleep(0.1)
8    pngfile = tempname() * ".png"
9    savefig(plot(P...; kwargs...), pngfile)
10   showimg("image/png", pngfile)
11 end
12 pngplot(; kwargs...) = pngplot(plot!(; kwargs...))
13
14 ▼ showing(mime, fn) = open(fn) do f
15   base64 = base64encode(f)
16   display("text/html", """""")
17 end
18
19 using SymPy
20 #sympy[:init_printing](order="lex") # default
21 #sympy[:init_printing](order="rev-lex")
22
23 using SpecialFunctions
24 using QuadGK
```

## 1 Landau記号

### 1.1 数列版

数列  $a_n, b_n$  について

$$\frac{a_n}{b_n} \rightarrow 0 \quad (n \rightarrow \infty)$$

が成立するとき, 量  $a_n$  を  $o(b_n)$  と書くことがある:

$$a_n = o(b_n).$$

例えば

- $n = o(n^2)$ .
- $n^{100} = o(e^n)$ .
- $\log n = o(n)$ .

ある非負の定数  $M$  が存在して, 十分大きな  $n$  について

$$|a_n| \leq M|b_n|$$

が成立するとき, 量  $a_n$  を  $O(b_n)$  と書くことがある:

$$a_n = O(b_n).$$

例えば

- $3n^2 + 100n + 26 = O(n^2)$ .
- $n \log n - n + \frac{1}{2}\log(2\pi n) = O(n \log n)$ .

以上で定義した  $o(b_n)$  と  $O(b_n)$  を**Landau記号**と呼ぶ.

違うものを同じLandau記号の  $o(b_n)$  や  $O(b_n)$  で表すことがあるので注意しなければいけない. 例えば  $O(n) - O(n) = 0$  とはならず,  $O(n) - O(n) = O(n)$  となる.

$o(b_n)$  の量は  $O(b_n)$  の量にもなる.

**問題:**  $n \rightarrow \infty$  のとき  $a_n/b_n$  が収束するならば  $a_n = O(b_n)$  であることを示せ.

**解答例:**  $a_n/b_n$  が  $C$  に収束しているなら,  $|a_n/b_n|$  は  $|C|$  に収束するので,  $M = |C| + 1$  とおくと, 十分大きな  $n$  について  $|a_n/b_n| \leq |C| + 1 = M$  すなわち  $|a_n| \leq M|b_n|$  が成立する.  $\square$

## 1.2 連続版

$x \rightarrow \infty$  におけるLandau記号  $o(g(x))$ ,  $O(g(x))$  は上の数列の場合と同様に定義される.

有限の  $a$  に関する  $x \rightarrow a$  におけるLandau記号  $o(g(x))$ ,  $O(g(x))$  も同様に定義される.

$$\frac{f(x)}{g(x)} \rightarrow 0 \quad (x \rightarrow a)$$

が成立するとき,  $f(x) = o(g(x))$  と書き, ある非負の定数  $M$  とある  $\delta > 0$  が存在して

$$|f(x)| \leq M|g(x)| \quad (|x - a| \leq \delta)$$

が成立するとき,  $f(x) = O(g(x))$  と書く.

例えば,

- $x \rightarrow 0$  において  $x \log x \rightarrow 0$  なので

$$x = o(1/\log x), \quad \log x = o(1/x) \quad (x \rightarrow 0).$$

- $x \rightarrow a$  において,

$$(x - a)^2 + (x - a)^3 + (x - a)^4 + \dots = O((x - a)^2) = o(x - a) \quad (x \rightarrow a).$$

- どんなに大きな  $n$  についても,  $x \searrow 0$  において  $e^{-1/x}/x^n \rightarrow 0$  なので

$$\exp\left(-\frac{1}{x}\right) = o(x^n) \quad (x \searrow 0).$$

**問題:** 以上に出て来た例を確認(証明)せよ. さらに以上に出て来ていない例で面白そうなものを自分で作れ.  $\square$

## 2 Landau記号を使った微分可能性の定義

### 2.1 連続函数

函数  $f(x)$  が  $x$  において連続であることの定義は, Landau記号を使うと

$$f(x + h) = f(x) + o(1) \quad (h \rightarrow 0)$$

と書ける.  $o(1)$  は  $h \rightarrow 0$  で 0 に収束する項を意味する.

## 2.2 微分可能函数

函数  $f(x)$  が  $x$  において微分可能であることの定義は、ある  $f'(x)$  が存在して、

$$\frac{f(x+h)-f(x)}{h} \rightarrow f'(x) \quad (h \rightarrow 0)$$

が成立することである。この条件は Landau 記号を使うと、

$$f(x+h) = f(x) + f'(x)h + o(h)$$

と書き直される。ここで  $o(h)$  は  $h \rightarrow 0$  のとき  $o(h)/h \rightarrow 0$  となるような量を意味する。 $o(h)$  を  $h$  より高次の微小量と呼ぶことがある。

以上をまとめると、

**連続性:**  $f(x+h) = f(x) + o(1) (h \rightarrow 0)$ .

**微分可能性:**  $f(x+h) = f(x) + f'(x)h + o(h) (h \rightarrow 0)$ .

このように書けばこれらは

$$f(x+h) = a_0 + a_1 h + a_2 h^2 + \cdots + a_k h^k + o(h^k)$$

の形に一般化可能であることが明瞭になる。後で実際にこの一般化を実行する(Taylor の定理)。

## 2.3 $C^n$ 級函数

連続函数を  $C^0$  級函数と呼ぶ。

函数  $f$  が微分可能でかつ導函数が連続なとき  $f$  は  $C^1$  級であるという。そのとき  $f$  は連続微分可能であると言うこともある。

より一般に、 $f$  が  $n$  回微分可能でかつその  $n$  階の導函数  $f^{(n)}$  が連続なとき、 $f$  は  $C^n$  級であるという。

任意有限回微分可能な函数を  $C^\infty$  級函数と呼ぶ。

$C^n$  級函数(class- $C^n$  function)を  $C^n$  函数 ( $C^n$ -function) と呼ぶことが多い。

**問題:** 微分可能だが  $C^1$  級でない函数の例を挙げよ。

**解答例:** 函数  $f$  を

$$f(x) = \begin{cases} x^2 \sin(1/x) & (x \neq 0) \\ 0 & (x = 0) \end{cases}$$

と定める。このとき、 $x \neq 0$  ならば  $f'(x) = 2x \sin(1/x) - \cos(1/x)$  である。そして、 $x \rightarrow 0$  のとき

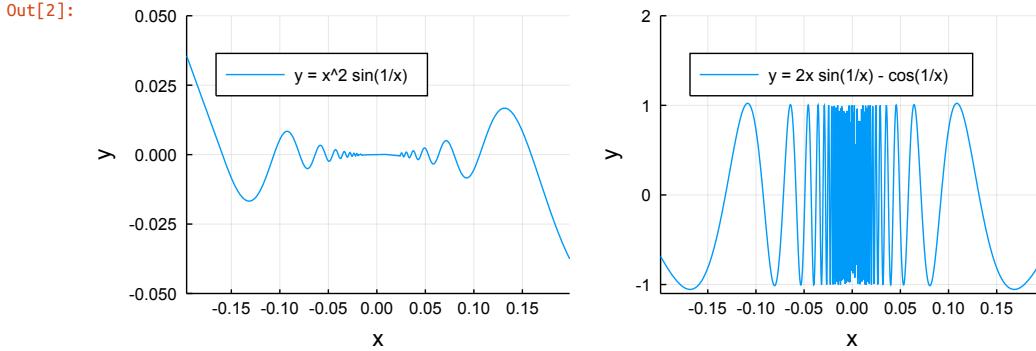
$$\frac{f(x)-f(0)}{x-0} = \frac{x^2 \sin(1/x)}{x} = x \sin(1/x) \rightarrow 0$$

なので、 $f$  は  $x = 0$  でも微分可能であり、 $f'(0) = 0$  である。これで  $f$  が微分可能であることがわかった。しかし、 $x \rightarrow 0$  で  $f'(x)$  の中の  $\cos(1/x)$  が無限に振動して収束しないので、 $f'$  は  $x = 0$  で不連続である。これで  $f$  は  $C^1$  級ではないことがわかった。□

**問題:** 上の問題の  $x^2 \sin(1/x)$  を非負の整数  $n$  に対する  $x^n \sin(1/x)$  で置き換えた場合について考えてみよ。その場合に得らえる函数は何回微分可能であるか？最大回数の微分の結果得られる導函数は連続であるか？まず、 $n = 0, 1, 3, 4$  の場合について考えてみよ。□

解答例略。

```
In [2]: 1 f(x) = x == 0 ? 0 : x^2*sin(1/x)
2 g(x) = x == 0 ? 0 : 2x*sin(1/x) - cos(1/x)
3
4 p1 = plot(size=(400, 250), xlabel="x", ylabel="y", legend=:topleft)
5 plot!(f, -0.2, 0.2, ylims=(-0.05, 0.05), label="y = x^2 sin(1/x)")
6
7 p2 = plot(size=(400, 250), xlabel="x", ylabel="y", legend=:topleft)
8 plot!(g, -0.2, 0.2, ylims=(-1.1, 2.0), label="y = 2x sin(1/x) - cos(1/x)")
9
10 plot(p1, p2, size=(700, 250))
```



### 3 微分の基本性質

#### 3.1 Leibnitz則

微分可能な函数  $f(x), g(x)$  の積  $h(x) = f(x)g(x)$  も微分可能であり,

$$h'(x) = f'(x)g(x) + f(x)g'(x).$$

**証明:**  $f, g$  は微分可能なので,

$$\begin{aligned} f(x+h) &= f(x) + f'(x)h + \varepsilon(h)h, & \varepsilon(h) \rightarrow 0 & (h \rightarrow 0), \\ g(x+h) &= g(x) + g'(x)h + \delta(h)h, & \delta(h) \rightarrow 0 & (h \rightarrow 0) \end{aligned}$$

と書ける. ゆえに

$$\begin{aligned} h(x+h) &= f(x+h)g(x+h) = f(x)g(x) + (f'(x)g(x) + f(x)g'(x))h + \eta(h)h, \\ \eta(h) &= f'(x)g'(x)h + (f(x) + f'(x)h)\delta(h) + (g(x) + g'(x)h)\varepsilon(h) + \varepsilon(h)\delta(h)h \end{aligned}$$

となることを確認できる. そして,  $h \rightarrow 0$  のとき  $\eta(h) \rightarrow 0$  なので,  $h$  も微分可能で  $h'(x) = f'(x)g(x) + f(x)g'(x)$  となることがわかる.  $\square$

**問題:** 部分可能な函数  $f, g$  について

$$\left(\frac{d}{dx}\right)^n (f(x)g(x)) = \sum_{k=0}^n \binom{n}{k} f^{(k)}(x)g^{(n-k)}(x)$$

となることを示せ.

**解法1:**  $n$  に関する帰納法.  $\square$

**解法2:** 二項定理より,

$$\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)^n (f(x)g(y)) = \sum_{k=0}^n \binom{n}{k} \left(\frac{\partial}{\partial x}\right)^k \left(\frac{\partial}{\partial y}\right)^{n-k}$$

より, これを  $f(x)g(y)$  に作用させると,

$$\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)^n (f(x)g(y)) = \sum_{k=0}^n \binom{n}{k} f^{(k)}(x)g^{(n-k)}(y).$$

以上において  $y = x$  と置けば欲しい結果が得られる.  $\square$

#### 3.2 合成函数の微分則

微分可能な函数  $f(y), g(x)$  の合成  $h(x) = f(g(x))$  も微分可能であり,

$$h'(x) = f'(g(x))g'(x).$$

**証明:**  $f, g$  は微分可能なので,

$$\begin{aligned} f(y+k) &= f(y) + f'(y)k + \varepsilon(k)k, & \varepsilon(k) \rightarrow 0 & (k \rightarrow 0), \\ g(x+h) &= g(x) + g'(x)h + \delta(h)h, & \delta(h) \rightarrow 0 & (h \rightarrow 0) \end{aligned}$$

と書ける。ゆえに  $y = g(x)$ ,  $k = g'(x)h + \delta(h)h$  とおくと,  $h \rightarrow 0$  のとき  $k \rightarrow 0$  であり,

$$\begin{aligned} h(x+h) &= f(g(x+h)) = f(y+k) = f(g(x)) + f'(g(x))g'(x)h + \eta(h)h, \\ \eta(h) &= f'(g(x))\delta(h) + e(k)(g'(x) + \delta(h)) \end{aligned}$$

であり,  $h \rightarrow 0$  のとき  $\eta(h) \rightarrow 0$  となるので,  $h$  も微分可能かつ  $h'(x) = f'(g(x))g'(x)$  となることがわかる。□

## 4 平均値の定理

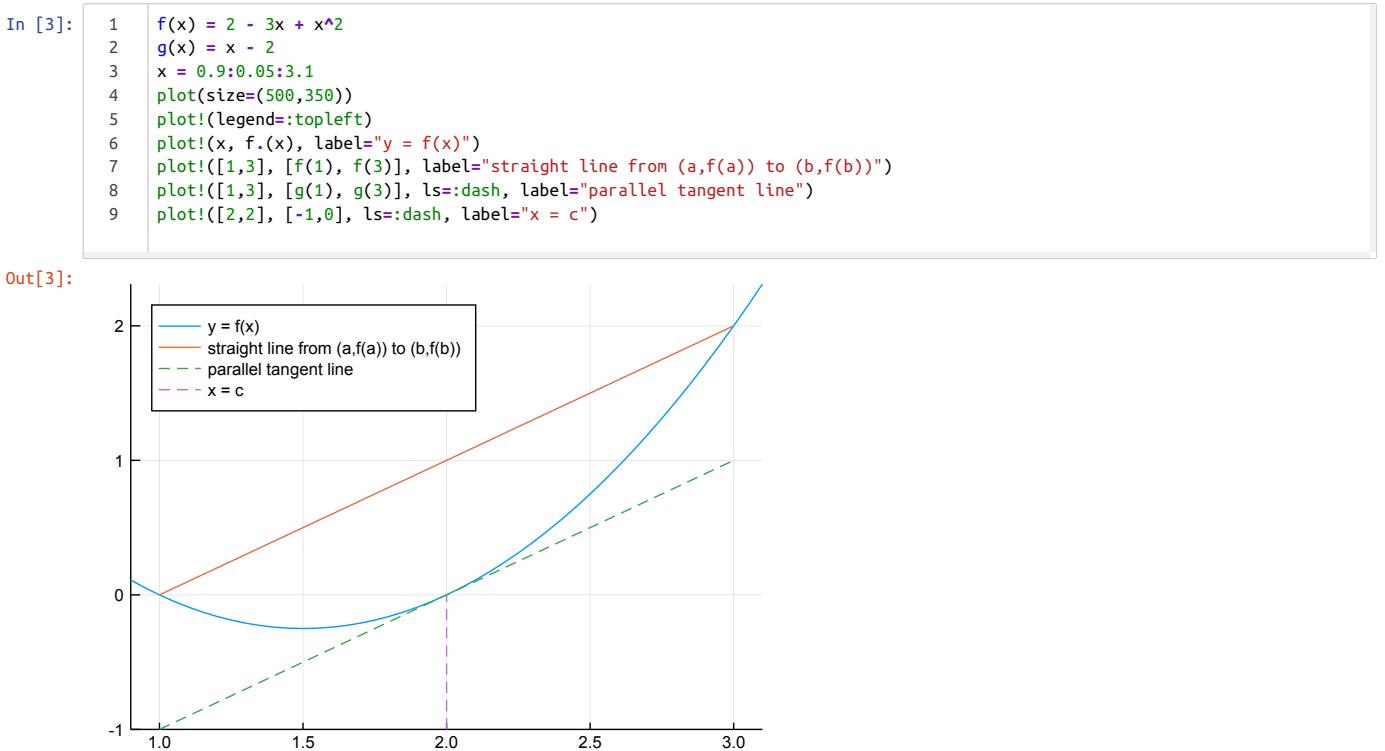
### 4.1 微分可能函数の平均値の定理

**平均値の定理:**  $f$  が微分可能函数ならば,  $a < b$  のとき, ある実数  $c$  で

$$f(b) - f(a) = f'(c)(b - a), \quad a < c < b$$

を満たすものが存在する。□

証明略。次の図を見よ。



**平均値の定理の不等式版:**  $f$  は微分可能函数であるとし,  $a < b$  であるとする。 $a < t < b$  で  $|f'(t)| \leq M$  となっているならば  $|f(b) - f(a)| \leq M|b - a|$  となる。

**証明:** 平均値の定理より, ある実数  $c$  で  $f(b) - f(a) = f'(c)(b - a)$ ,  $a < c < b$  を満たすものが存在する。そのとき

$$|f(b) - f(a)| = |f'(c)||b - a| \leq M|b - a|. \quad \square$$

平均値の定理はこの不等式版の形式で使われることが多い。

**問題:**  $f$  は微分可能函数であるとし,  $a < b$  であるとする。 $a < t < b$  で  $f'(t) = 0$  ならば  $f(b) = f(a)$  であることを示せ。

**解答例1:** 平均値の定理より, ある実数  $c$  で  $a < c < b$ ,  $f(b) - f(a) \leq f'(c)(b - a)$  を満たすものが存在する。 $a < t < b$  で  $f'(t) = 0$  なので  $f'(c) = 0$  となるので,  $f(b) - f(a) = f'(c)(b - a) = 0$  である。ゆえに  $f(b) = f(a)$  となる。□

**解答例2:**  $a < t < b$  で  $|f'(t)| = 0$  なので, 平均値の定理(の不等式版)より,  $|f(b) - f(a)| \leq 0|b - a| = 0$ 。ゆえに  $f(b) = f(a)$  となる。□

上の問題の結果は, 導函数が恒等的に 0 になる微分可能函数は定数函数になることを意味している。

**問題:**  $f, g$  は区間  $I$  上の微分可能函数であるとする。区間  $I$  上で恒等的に  $f' = g'$  ならば  $f - g$  は区間  $I$  上の定数函数になる。

**解答例:**  $h = f - g$  とおくと, 区間  $I$  上で恒等的に  $f' = g'$  ならば区間  $I$  上で恒等的に  $h' = 0$  となるので, 上の問題の結果より,  $h$  は区間  $I$  上の定数函数になる。□

### 4.2 $C^1$ 級函数の積分版平均値の定理

**平均値の定理の積分版:**  $a < b$  であるとする。閉区間  $[a, b]$  上の  $C^1$  級函数  $f$  について

$$f(b) - f(a) = \int_a^b f'(t) dt. \quad \square$$

証明略.

多くの議論を微分可能性の仮定だけで実行可能なのだが、導函数が連続函数になるとは限らないので、議論が技巧的になってしまふ傾向がある。そうなることを避けたければ  $C^1$  級まで仮定を強めた方がよい。平均値の定理の積分版の式は積分について知っている人ならばみんな知っている公式に過ぎない。

解析学ユーザー的には  $C^1$  級まで条件を強めても困ることはない。

**平均値の定理の不等式版:**  $f$  は  $C^1$  級函数であるとし、 $a < b$  であるとする。 $a < t < b$  で  $|f'(t)| \leq M$  となっているならば  $|f(b) - f(a)| \leq M|b - a|$  となる。

証明:

$$|f(b) - f(a)| = \left| \int_a^b f'(t) dt \right| \leq \int_a^b |f'(t)| dt \leq \int_a^b M dt = M|b - a|. \quad \square$$

**中間値の定理から導かれる補題:**  $a < b$  であるとする。閉区間  $[a, b]$  上の実数値連続函数  $g(t)$  に対して、ある実数  $c \in [a, b]$  で

$$\int_a^b g(t) dt = g(c)(b - a)$$

を満たすものが存在する。(数値例が下の方で計算されている。)

証明:  $[a, b]$  上での  $g(t)$  の最小値と最大値をそれぞれ  $g(\alpha), g(\beta)$  ( $\alpha, \beta \in [a, b]$ ) と書くと、

$$\int_a^b g(t) dt \geq \int_a^b g(\alpha) dt = g(\alpha)(b - a), \quad \int_a^b g(t) dt \leq \int_a^b g(\beta) dt = g(\beta)(b - a)$$

より

$$g(\alpha) \leq \frac{1}{b - a} \int_a^b g(t) dt \leq g(\beta).$$

ゆえに、中間値の定理より、ある実数  $c \in [a, b]$  で

$$g(c) = \frac{1}{b - a} \int_a^b g(t) dt, \quad \text{すなわち} \quad \int_a^b g(t) dt = g(c)(b - a)$$

を満たすものが存在する。□

上の中間値の定理から導かれる補題を平均値の定理の積分版に適用すると次が得られる。

**平均値の定理の積分版2:**  $a < b$  であるとする。 $f$  が  $[a, b]$  上の実数値  $C^1$  級函数ならば、ある実数  $c \in [a, b]$  が存在して、

$$f(b) - f(a) = \int_a^b f'(t) dt = f(c)(b - a). \quad \square$$

```

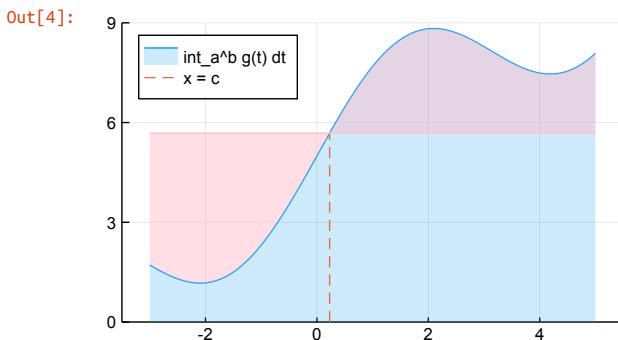
In [4]: 1 ▼ # 中間値の定理の証明に使われる二分法の函数としての実装
2 ▼ function bisection(f, a, b; atol=eps(), maxiter=10^3)
3     f(a)*f(b) > 0 && return NaN
4     a_n, b_n = a, b
5     k = 0
6 ▼     while abs(a_n - b_n) > atol
7         k > maxiter && return NaN
8         c = (a_n+b_n)/2
9 ▼         if f(b_n)*f(c) ≥ 0
10            a_n, b_n = a_n, c
11        else
12            a_n, b_n = c, b_n
13        end
14        k += 1
15    end
16    return (a_n+b_n)/2
17 end
18
19 # 中間値の定理から導かれる補題を次の g(x) と a,b に適用してみる。
20 g(x) = 5 + x + 2sin(x)
21 a, b = -3, 5
22 x = a:0.05:b
23
24 t = symbols("t", real=true)
25 @show g(t)
26 @show S = integrate(g(t), (t,a,b))
27 @show N(S)
28
29 h(x) = g(x) - S/(b-a)
30 @show h(t)
31 @show c = bisection(h, a, b)
32
33 plot(size=(400,250), legend=:topleft)
34 plot!(xLims=(a-0.5, b+0.5), lims=(0,9))
35 plot!(x, g.(x), label="int_a^b g(t) dt", fill=(0, 0.2))
36 plot!([c,c], [0,g(c)], ls=:dash, label="x = c")
37 plot!(x, fill(g(c), length(x)), fill=(g, 0.5), color=:pink, label="")

```

```

g(t) = t + 2*sin(t) + 5
S = integrate(g(t), (t, a, b)) = 2*cos(3) - 2*cos(5) + 48
S = N(S) = 45.452690635872656
h(t) = t + 2*sin(t) - 0.681586329484082
c = bisection(h, a, b) = 0.22851790816319195

```



### 4.3 Cauchyの平均値の定理

**Cauchyの平均値の定理:**  $f(t)$ ,  $g(t)$  は微分可能函数であるとし,  $a < b$  であるとし,  $f(b) \neq f(a)$  であると仮定する. このとき, ある実数  $c$  で

$$\frac{g(b) - g(a)}{f(b) - f(a)} = \frac{g'(c)}{f'(c)}, \quad a < c < b \quad (*)$$

を満たすものが存在する.  $\square$

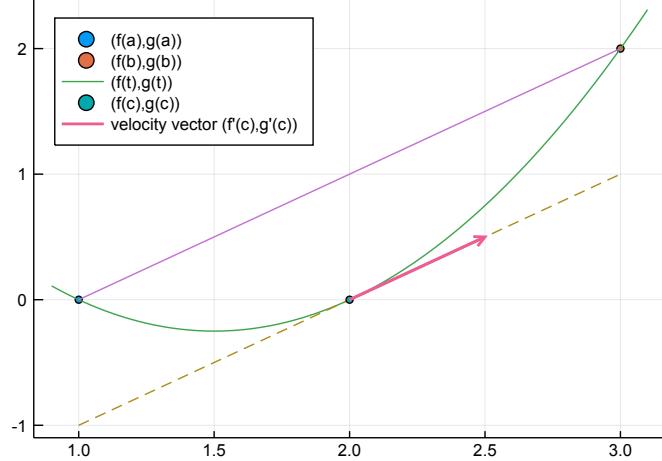
証明略.

$(f(t), g(t))$  の描く軌跡は平面上の曲線になる.  $(*)$  の最初の式の左辺は, 曲線上の点  $(f(a), g(a))$ ,  $(f(b), g(b))$  を通る直線の傾きになる. その直線を平行移動して曲線に接するようにできる. 接点  $(f(c), g(c))$  における測度ベクトル  $(f'(c), g'(c))$  の傾きが  $(*)$  の最初の式の右辺になる.

**問題:** 図を描いてみよ.  $\square$

```
In [5]: 1 f(t) = t/2
2 g(t) = 2 - 3t/2 + t^2/4
3 h(x) = x - 2
4 t = 1.8:0.05:6.2
5 plot(size=(500,350))
6 plot!(legend=:topLeft)
7 scatter!([f(2)], [g(2)], markersize=3, label="(f(a),g(a))")
8 scatter!([f(6)], [g(6)], markersize=3, label="(f(b),g(b))")
9 plot!(f.(t), g.(t), label="(f(t),g(t))")
10 plot!([f(2),f(6)], [g(2), g(6)], label="")
11 plot!([1,3], [h(1), h(3)], ls=:dash, label="")
12 scatter!([f(4)], [g(4)], markersize=3, label="(f(c),g(c))")
13 plot!([2,2.5], [h(2), h(2.5)], line=:arrow, lw=2, label="velocity vector (f'(c),g'(c))")
```

Out[5]:



In [ ]:

1

# Taylorの定理

黒木玄

2018-04-20

- Copyright 2018 Gen Kuroki
- License: MIT <https://opensource.org/licenses/MIT> (<https://opensource.org/licenses/MIT>)

このファイルは次の場所できれいに閲覧できる:

- <http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/06%20Taylor%27s%20theorems.ipynb>  
<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/06%20Taylor%27s%20theorems.ipynb>

このファイルは [Julia Box](https://juliabox.com) (<https://juliabox.com>) で利用できる。

自分のパソコンに [Julia言語](https://julialang.org/) (<https://julialang.org/>) をインストールしたい場合には

- [WindowsへのJulia言語のインストール](http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f) (<http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f>)

を参照せよ。

論理的に完璧な説明をするつもりはない。細部のいい加減な部分は自分で訂正・修正せよ。

## 目次

### 1 Taylorの定理

#### 1.1 積分型剩余項型Taylorの定理

- 1.1.1 積分型剩余項の絶対値の上からの評価
- 1.1.2 積分型剩余項の書き直し(1)
- 1.1.3 積分型剩余項の書き直し(2)
- 1.1.4 Landau記号を用いたTaylorの定理

#### 1.2 $n$ 回の微分可能性だけを仮定したTaylorの定理について

### 2 Taylor展開

- 2.1 基本的なTaylor展開の例
- 2.2 コンピューターを用いた計算例

### 3 べき級数の収束

- 3.1 べき級数の定義
- 3.2 べき級数の収束

### 4 超幾何級数

```
In [1]: 1  using Plots
2  gr(); ENV["PLOTS_TEST"] = "true"
3  #clibrary(:colorcet)
4  clibrary(:misc)
5
6  ▼ function pngplot(P...; kwargs...)
7    sleep(0.1)
8    pngfile = tempname() * ".png"
9    savefig(plot(P...; kwargs...), pngfile)
10   showimg("image/png", pngfile)
11 end
12 pngplot(; kwargs...) = pngplot(plot!(; kwargs...))
13
14 ▼ showimg(mime, fn) = open(fn) do f
15    base64 = base64encode(f)
16    display("text/html", """""")
17 end
18
19 using SymPy
20 #sympy[:init_printing](order="lex") # default
21 #sympy[:init_printing](order="rev-lex")
22
23 using SpecialFunctions
24 using QuadGK
```

## 1 Taylorの定理

$$f(x) = f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2 + \cdots + \frac{1}{k!}f^{(k)}(a)(x - a)^k + \cdots$$

の最後の  $+ \cdots$  の部分を  $f$  の  $k+1$  階の導函数の式で書いた結果を **Taylorの定理** と呼ぶことにする。その表示の仕方には様々な形がある。

## 1.1 積分型剩余項型Taylorの定理

$f$  は  $C^n$  級函数であると仮定する.

まず,  $n = 4$  の場合を例として扱う. そのとき

$$f'''(x) = f'''(a) + \int_a^x f^{(4)}(x_4) dx_4.$$

ゆえに

$$\begin{aligned} f''(x) &= f''(a) + \int_0^x f'''(x_3) dx_3 \\ &= f''(a) + f'''(a)(x - a) + \int_a^x \left[ \int_a^{x_3} f^{(4)}(x_4) dx_4 \right] dx_3. \end{aligned}$$

括弧が増えると書くのが面倒になるので

$$\int_a^x dx_3 \int_a^{x_3} dx_4 f^{(4)}(x_4) := \int_a^x \left[ \int_a^{x_3} f^{(4)}(x_4) dx_4 \right] dx_3$$

と書くことにする. 積分が3重以上でも同様に書くことにする. そのように書くと,

$$f''(x) = f''(a) + f'''(a)(x - a) + \int_a^x dx_3 \int_a^{x_3} dx_4 f^{(4)}(x_4).$$

ゆえに

$$\begin{aligned} f'(x) &= f'(a) + \int_a^x dx_2 f''(x_2) \\ &= f'(a) + f''(a)(x - a) + \frac{1}{2} f'''(a)(x - a)^2 + \int_a^x dx_2 \int_a^{x_2} dx_3 \int_a^{x_3} dx_4 f^{(4)}(x_4). \end{aligned}$$

さらに,

$$\begin{aligned} f(x) &= f(a) + \int_a^x dx_1 f'(x_1) \\ &= f(a) + f'(a)(x - a) + \frac{1}{2} f''(a)(x - a)^2 + \frac{1}{3!} f'''(a)(x - a)^3 + \int_a^x dx_1 \int_a^{x_1} dx_2 \int_a^{x_2} dx_3 \int_a^{x_3} dx_4 f^{(4)}(x_4). \end{aligned}$$

以上をまとめると,

$$\begin{aligned} f(x) &= f(a) + f'(a)(x - a) + \frac{1}{2} f''(a)(x - a)^2 + \frac{1}{3!} f'''(a)(x - a)^3 + R_4, \\ R_4 &= \int_a^x dx_1 \int_a^{x_1} dx_2 \int_a^{x_2} dx_3 \int_a^{x_3} dx_4 f^{(4)}(x_4). \end{aligned}$$

上と同様にして次が成立することを示せる.

**Taylorの定理:**  $f$  は  $C^n$  級函数であると仮定する. そのとき,

$$\begin{aligned} f(x) &= f(a) + f'(a)(x - a) + \frac{1}{2} f''(a)(x - a)^2 + \cdots + \frac{1}{(n-1)!} f^{(n-1)}(a)(x - a)^{n-1} + R_n, \\ R_n &= \int_a^x dx_1 \int_a^{x_1} dx_2 \cdots \int_a^{x_{n-1}} dx_n f^{(n)}(x_n). \end{aligned}$$

$R_n$  を剩余項と呼ぶ.  $\square$

この形のTaylorの定理は  $f^{(n)}(x)$  を何度も積分するだけで得られ、技巧的な証明の工夫をする必要が一切ない.

**問題:** 上の形のTaylorの定理の証明を完成せよ.  $\square$

### 1.1.1 積分型剩余項の絶対値の上からの評価

以下,  $a$  と  $x$  のあいだの  $t$  について  $|f^{(n)}(t)| \leq M$  が成立しているとき,

$$|R_n| \leq \left| \int_a^x dx_1 \int_a^{x_1} dx_2 \cdots \int_a^{x_{n-1}} dx_n M \right| = \frac{1}{n!} M|x - a|^n.$$

ここで,

$$\int_a^x dx_1 \int_a^{x_1} dx_2 \cdots \int_a^{x_{n-1}} dx_n 1 = \frac{1}{n!}(x - a)^n$$

となることをを使った. 1 を  $a$  から  $x$  まで積分する操作を  $n$  回繰り返すとそうなることは容易に確かめられる.

$$|R_n| \leq \frac{M|x - a|^n}{n!}$$

この形の不等式はよく使われる.

### 1.1.2 積分型剩余項の書き直し (1)

**定理:**

$$R_n = \int_a^x f^{(n)}(x_n) \frac{(x - x_n)^{n-1}}{(n-1)!} dx_n.$$

**証明:** 簡単のため  $a < x$  と仮定する( $a > x$  の場合も同様である). そのとき, 積分

$$R_n = \int_a^x dx_1 \int_a^{x_1} dx_2 \cdots \int_a^{x_{n-1}} dx_n f^{(n)}(x_n)$$

の中の  $x_i$  たちは  $a \leqq x_n \leqq x_{n-1} \leqq \cdots \leqq x_2 \leqq x_1 \leqq x$  を動く. ゆえに, この積分は

$$R_n = \int_a^x \left[ f^{(n)}(x_n) \int_{x_n}^x dx_1 \int_{x_n}^{x_1} dx_2 \cdots \int_{x_n}^{x_{n-2}} dx_{n-1} 1 \right] dx_n$$

と書き直される. そして,

$$\int_{x_n}^x dx_1 \int_{x_n}^{x_1} dx_2 \cdots \int_{x_n}^{x_{n-2}} dx_{n-1} 1 = \frac{(x - x_n)^{n-1}}{(n-1)!}$$

なので上の定理が得られる.  $\square$

**注意:**  $f$  が  $C^{n+1}$  級函数ならば, 上の定理の  $n$  の場合の結果から, 部分積分によって

$$R_n = f^{(n)}(a) \frac{(x - a)^n}{n!} + R_{n+1}$$

が得られる(問題: 示してみよ). ゆえに, 部分積分を用いた  $n$  に関する帰納法によっても上の定理を示せる. 部分積分を用いる計算法はよく使われる.  $\square$

### 1.1.3 積分型剩余項の書き直し (2)

簡単のため  $a < x$  であると仮定する.

$f$  は  $C^n$  級函数であるとする. そのとき  $f^{(n)}$  は連続函数になる.  $f^{(n)}(t)$  の  $a \leqq t \leqq x$  における最小値と最大値はそれぞれ  $f^{(n)}(\alpha), f^{(n)}(\beta)$  であるとする. そのとき,

$$f^{(n)}(\alpha) \frac{(x - a)^n}{n!} \leqq R_n \leqq f^{(n)}(\beta) \frac{(x - a)^n}{n!}$$

が成立する. なぜならば  $R_n$  の中の  $f^{(n)}(x_n)$  を定数  $C$  で置き換えると  $R_n$  の値は  $C(x - a)^n/n!$  に変わるからである. 中間値の定理より, ある実数  $\xi$  で

$$R_n = f^{(n)}(\xi) \frac{(x - a)^n}{n!}, \quad a < \xi < x$$

を満たすものが存在する.  $a < x$  の仮定を  $a > x$  で置き換えると,  $a < \xi < x$  は  $a > \xi > x$  に変わる.

剩余項の

$$R_n = f^{(n)}(\xi) \frac{(x - a)^n}{n!}, \quad a < \xi < x$$

の形もよく使われる.

### Landau記号を用いたTaylorの定理

$C^n$  級函数  $f$  について  $a$  と  $x$  のあいだのある実数  $\xi$  で

$$R_n = f^{(n)}(\xi) \frac{(x - a)^n}{n!}$$

を満たすものが存在することを上で示した.  $f^{(n)}$  は連続なので  $x \rightarrow a$  のとき  $f^{(n)}(\xi) \rightarrow f^{(n)}(a)$  となる. ゆえに

$$R_n - f^{(n)}(a) \frac{(x - a)^n}{n!} = (f^{(n)}(\xi) - f^{(n)}(a)) \frac{(x - a)^n}{n!} = o((x - a)^n) \quad (x \rightarrow a).$$

これで次が成立することがわかった:

$$f(x) = f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2 + \cdots + \frac{1}{n!}f^{(n)}(a)(x - a)^n + o((x - a)^n) \quad (x \rightarrow a).$$

Taylorの定理はこの形でもよく使われる.

さらに  $f$  が  $C^{n+1}$  級ならば,  $n+1$  次の剩余項は

$$R_{n+1} = \frac{1}{(n+1)!} f^{(n+1)}(\xi)(x - a)^{n+1} = O((x - a)^{n+1}), \quad (\xi \text{ は } a \text{ と } x \text{ のあいだにある})$$

の形をしているので,

$$f(x) = f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2 + \cdots + \frac{1}{n!}f^{(n)}(a)(x - a)^n + O((x - a)^{n+1}) \quad (x \rightarrow a).$$

も成立していることがわかる.

### **$n$ 回の微分可能性だけを仮定した Taylor の定理について**

以上では  $C^n$  級の仮定のもとで Taylor の定理を証明したが、 $n$  回微分可能の仮定のもとでも Taylor の定理を導くことができる。例えば、

- 高木貞治『解析概論』改定第三版

の第25節のpp.61-63に詳しい説明がある。このノートでは技巧的になり過ぎるので扱わないことにする。

## **Taylor 展開**

$C^\infty$  函数  $f$  について、ある  $r > 0$  が存在して、

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(a)(x - a)^k$$

が  $|x - a| < r$  で収束するとき  $f$  は  $x = a$  で Taylor 展開可能であるという。そのような  $r$  の最大値を Taylor 展開の収束半径と呼ぶ。Taylor 展開はべき級数展開と呼ばれることが多い。

$x = 0$  における Taylor 展開は MacLaurin 展開と呼ばれることがある。

### **基本的な Taylor 展開の例**

以下の Taylor 展開はよく使われる。

$$\begin{aligned} (1+x)^a &= \sum_{k=0}^{\infty} \binom{a}{k} x^k & (|x| < 1), & \binom{a}{k} = \frac{a(a-1)\cdots(a-k+1)}{k!}. \\ (1-x)^{-a} &= \sum_{k=0}^{\infty} \frac{(a)_k}{k!} x^k & (|x| < 1), & (a)_k = a(a+1)\cdots(a+k-1). \\ e^x &= \sum_{k=0}^{\infty} \frac{x^k}{k!} & (|x| < \infty). \\ \cos x &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} & (|x| < \infty). \\ \sin x &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} & (|x| < \infty). \\ \log(1+x) &= \sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^k}{k}, & (|x| < 1). \\ -\log(1-x) &= \sum_{k=1}^{\infty} \frac{x^k}{k} & (|x| < 1). \\ \arcsin x &= \sum_{k=0}^{\infty} (-1)^k \binom{-1/2}{k} \frac{x^{2k+1}}{2k+1} = \sum_{k=0}^{\infty} \frac{1}{2^{2k}} \frac{(2k)!}{(k!)^2} \frac{x^{2k+1}}{2k+1} & (|x| < 1). \\ \arctan x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} & (|x| < 1). \end{aligned}$$

**問題:** 以上の公式をすべて示せ。□

解答は略す。まずは自分で  $f(0), f'(0), f''(0), f'''(0)$  を計算してみよう。そして、個人的には最初の段階では Taylor 展開が収束するか否かにはあまり神経を使わなくてもよいと思う。

他人が書いた解説に単に従うだけだと数学は決して理解できない。自分で何を計算するべきであるかを考え、自分が持っている力でその計算を実行してみて、その計算結果を見て次に何を計算するべきであるかを考え、また計算を繰り返す。試行錯誤抜きに数学をまともに理解できるようになることはありえない。

しかし、以下の説明も読んでおくこと。

**復習:**  $\binom{a}{k}$  は二項係数を表すのであった：

$$\binom{a}{k} = \frac{a(a-1)\cdots(a-k+1)}{k!}.$$

例えば

$$\binom{a}{0} = 1, \quad \binom{a}{1} = a, \quad \binom{a}{2} = \frac{a(a-1)}{2}, \quad \binom{a}{3} = \frac{a(a-1)(a-2)}{6}.$$

$(a)_k = a(a+1) \cdots (a+k-1)$  もよく使われる記号法である。例えば

$$(a)_0 = 1, \quad (a)_1 = a, \quad (a)_2 = a(a+1), \quad (a)_3 = a(a+1)(a+2).$$

**参考:** 収束半径については複素解析によって以下が知られている。 $x = a$  で Taylor 展開可能な函数  $f(x)$  を  $|z - a| < r$  を満たす複素数  $z$  の正則函数に拡張できるとき、その Taylor 展開の収束半径は  $r$  以上になる。だから、 $f(x)$  を可能な限り解析接続したものを  $f(z)$  と書くとき、 $f$  の  $x = a$  での Taylor 展開の収束半径は  $a$  から  $f(z)$  の正則ではない点までの最短距離に等しい。

例えば、 $a$  が非負の整数以外であるとき、 $f(x) = (1+x)^a = \exp(a \log(1+x))$  を複素平面に解析接続によって拡張すると、 $x = 0$  に最も近い  $f$  の正則ではない点は  $x = -1$  なので収束半径は 1 になる。例えば  $a = -1$  のとき

$$(1+x)^{-1} = \frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots$$

と公比  $-x$  の等比級数が出て来るが、これが成立するのは  $|x| < 1$  においてである。 $x = 0$  から最も近い函数  $f(x) = 1/(1+x)$  の特異点は  $x = -1$  である。

同様の理由で  $f(x) = \log(1+x)$  や  $f(x) = -\log(1-x)$  の場合も  $x = 0$  での Taylor 展開の収束半径は 1 になる。

例えば、 $\exp x, \sin x, \cos x$  は複素平面全体上の正則函数に拡張できることと、それらの Taylor 展開の収束半径が  $\infty$  になることは同値になり、実際にどちらも成立している。

**逆三角函数に関するヒント:**  $\arcsin x$  と  $\arctan x$  について

$$\arcsin x = \int_0^x \frac{dt}{\sqrt{1-t^2}}, \quad \arctan x = \int_0^x \frac{dt}{1+t^2}$$

を用いてよい。さらに  $|t| < 1$  で

$$\frac{1}{\sqrt{1-t^2}} = (1-t^2)^{-1/2} = \sum_{k=0}^{\infty} (-1)^k \binom{-1/2}{k} t^{2k}, \quad \frac{1}{1+t^2} = \sum_{k=0}^{\infty} (-1)^k t^{2k}.$$

これを項別に積分すれば欲しい結果が得られる。

二項係数  $\binom{-1/2}{k}$  の部分は以下のように処理できる：

$$\begin{aligned} (-1)^k \binom{-1/2}{k} &= (-1)^k \frac{-\frac{1}{2}(-\frac{1}{2}-1) \cdots (-\frac{1}{2}-k+1)}{k!} \\ &= \frac{\frac{1}{2}(\frac{1}{2}+1) \cdots (\frac{1}{2}+k-1)}{k!} = \frac{1 \cdot 3 \cdots (2k-1)}{2^k k!} \\ &= \frac{1 \cdot 3 \cdots (2k-1)}{2^k k!} \frac{2 \cdot 4 \cdots (2k)}{2^k k!} = \frac{1}{2^{2k}} \frac{(2k)!}{(k!)^2}. \quad \square \end{aligned}$$

**問題:**  $e^x$  と  $\cos x$  と  $\sin x$  の  $x = 0$  における Taylor 展開が  $x$  が複素数である場合にも成立していると仮定して、

$$e^{ix} = \cos x + i \sin x$$

が成立していることを示せ。□

**解答例:**

$$\begin{aligned} e^{ix} &= \sum_{n=0}^{\infty} \frac{i^n x^n}{n!} = \sum_{k=0}^{\infty} \frac{i^{2k} x^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{i^{2k+1} x^{2k+1}}{(2k+1)!} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} = \cos x + i \sin x. \end{aligned}$$

1つ目の等号で  $e^z$  の  $z = 0$  における Taylor 展開を  $z = ix$  に適用した。2つ目の等号では和を  $n$  が偶数の場合と奇数の場合に分けた。3つ目の等号では  $i^{2k} = (-1)^k$  を使った。4つ目の等号では  $\cos z, \sin z$  の  $z = 0$  における Taylor 展開を  $z = x$  に適用した。□

**問題:**  $\sqrt{10}$  を

$$(1+x)^{1/2} = \sum_{k=0}^{\infty} \binom{1/2}{k} x^k = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} + O(x^4) \quad (|x| < 1)$$

を用いて小数点以下第2桁まで求めよ。

**解答例:**

$$\sqrt{10} = (9+1)^{1/2} = 3(1+1/9)^{1/2} = 3 + 3 \frac{1}{2} \frac{1}{9} - 3 \frac{1}{8} \frac{1}{9^2} + 3 \frac{1}{16} \frac{1}{9^3} + \dots.$$

そして、

$$3\frac{1}{2}\frac{1}{9} = \frac{1}{6} = 0.166666 \dots, \quad -3\frac{1}{8}\frac{1}{9^2} = -\frac{1}{216} = -0.004629 \dots, \quad 3\frac{1}{16}\frac{1}{9^3} = 0.000257 \dots,$$

より

$$3 + 3\frac{1}{2}\frac{1}{9} - 3\frac{1}{8}\frac{1}{9^2} = 3.162 \dots$$

$\sqrt{10} = 3.16 \dots$  である.  $\square$

**注意:** 上では誤差項をきちんと評価せずに計算したがきちんと評価するには次のようにすればよい. 剰余項付きのTaylorの定理より,  $0 < x < 1$  のとき ある実数  $\xi$  が存在して,

$$(1+x)^{1/2} = 1 + \frac{x}{2} - \frac{x}{8} + \frac{x^3}{16(1+\xi)^{5/2}}, \quad 0 < \xi < x,$$

$$0 < \frac{x^3}{16(1+\xi)^{5/2}} < \frac{x^3}{16}.$$

ゆえに,  $x = 1/9$  のとき,  $\sqrt{10}$  と  $3 + 1/6 - 1/216 = 3.162 \dots$  の差は

$$3\frac{1}{16}\frac{1}{9^3} = 0.000257 \dots$$

より小さい. このことに注意すれば本当に小数点以下第2桁まで正確に計算できているという数学的な証明が得られる.  $\square$

```
In [2]: 1 x = symbols("x")
          2 series((1+x)^(Sym(1)/2), n=5)
```

Out[2]:  $1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \frac{5x^4}{128} + \mathcal{O}(x^5)$

```
In [3]: 1 3/2/9, -3/8/9^2, 3/16/9^3
```

Out[3]: (0.1666666666666666, -0.004629629629629629, 0.000257201646090535)

```
In [4]: 1 3 + 3/2/9 -3/8/9^2, 3/16/9^3
```

Out[4]: (3.1620370370368, 0.000257201646090535)

$\sqrt{ }$  は  $\backslash\sqrt$  TAB で入力できる.

<https://docs.julialang.org/en/stable/manual/unicode-input/> (<https://docs.julialang.org/en/stable/manual/unicode-input/>)

```
In [5]: 1 √(10)
```

Out[5]: 3.1622776601683795

```
In [6]: 1 factorial(10)
```

Out[6]: 3628800

**問題:**  $\tan x$  の  $x = 0$  でのTaylor展開 5 次の項まで求めよ.

**解答例:**  $f(x) = \tan x$  とおくと  $f' = 1 + \tan^2 x = 1 + f^2$  となるので,

$$\begin{aligned} f'' &= 2ff' = 2f + 2f^3, \\ f''' &= 2f' + 6f^2f' = 2 + 8f^2 + 6f^4, \\ f^{(4)} &= 16ff' + 24f^3f' = 16f + 40f^3 + 24f^5, \\ f^{(5)} &= 16f' + 120f^2f' + 120f^4f' = 16 + 136f^2 + 240f^4 + 120f^6. \end{aligned}$$

ゆえに

$$f(0) = f''(0) = f^{(4)}(0) = 0, \quad f'(0) = 1, \quad f'''(0) = 2, \quad f^{(5)}(0) = 16.$$

したがって,

$$\tan x = f(x) = x + \frac{2}{3!}x^3 + \frac{16}{5!}x^5 + \dots = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \dots. \quad \square$$

**別の解答例:**  $\sin x$  と  $\cos x$  のMaclaurin展開から  $\tan x$  のMaclaurin展開を求めよう:

$$\begin{aligned}
\tan x &= \frac{\sin x}{\cos x} = \frac{\sin x}{1 - (1 - \cos x)} \\
&= \sin x (1 + (1 - \cos x) + (1 - \cos x)^2 + \dots) \\
&= \left( x - \frac{x^3}{6} + \frac{x^5}{120} + O(x^7) \right) \left( 1 + \left( \frac{x^2}{2} - \frac{x^4}{24} \right) + \left( \frac{x^2}{2} - \frac{x^4}{24} \right)^2 + O(x^6) \right) \\
&= \left( x - \frac{x^3}{6} + \frac{x^5}{120} + O(x^7) \right) \left( 1 + \frac{x^2}{2} + \left( -\frac{1}{24} + \frac{1}{4} \right) x^4 + O(x^6) \right) \\
&= \left( x - \frac{x^3}{6} + \frac{x^5}{120} + O(x^7) \right) \left( 1 + \frac{x^2}{2} + \frac{5}{24} x^4 + O(x^6) \right) \\
&= x + \left( \frac{1}{2} - \frac{1}{6} \right) x^3 + \left( \frac{5}{24} - \frac{1}{12} + \frac{1}{120} \right) x^5 + O(x^7) \\
&= x + \frac{1}{3} x^3 + \frac{2}{15} x^5 + O(x^7). \quad \square
\end{aligned}$$

```
In [7]: 1 x = symbols("x")
2 sympy[:init_printing](order="rev-lex")
3 ▼ for k in 0:5
4     expand(diff(tan(x), x, k)) ▷ display
5 end
6 sympy[:init_printing](order="lex") # default
7
8 x = symbols("x")
9 series(tan(x), n=10)
```

$$\begin{aligned}
\tan(x) & \\
1 + \tan^2(x) & \\
2 \tan(x) + 2 \tan^3(x) & \\
2 + 8 \tan^2(x) + 6 \tan^4(x) & \\
16 \tan(x) + 40 \tan^3(x) + 24 \tan^5(x) & \\
16 + 136 \tan^2(x) + 240 \tan^4(x) + 120 \tan^6(x) &
\end{aligned}$$

Out[7]:  $x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \frac{62x^9}{2835} + \mathcal{O}(x^{10})$

```
In [8]: 1 s = series(sin(x), n=7)
2 cinv = series(1/cos(x), n=6)
3 display(s)
4 display(cinv)
5 expand(s * cinv)
```

$$\begin{aligned}
x - \frac{x^3}{6} + \frac{x^5}{120} + \mathcal{O}(x^7) & \\
1 + \frac{x^2}{2} + \frac{5x^4}{24} + \mathcal{O}(x^6) &
\end{aligned}$$

Out[8]:  $x + \frac{x^3}{3} + \frac{2x^5}{15} + \mathcal{O}(x^7)$

**問題:**  $e = e^1$  の値を  $e^x$  のMaclaurin展開を用いて誤差100万の1以下で正確に計算するためには  $\sum_{k=0}^9 \frac{1}{k!}$  を計算すれば十分であることを示せ.  $e < 3$  と  $10! = 3628800$  を認めて使ってよい.

**解答例:** Taylorの定理より, 0 と 1 のあいだにある実数  $\xi$  で次を満たすものが存在する.

$$e^1 = \sum_{k=0}^9 \frac{1}{k!} + \frac{e^\xi}{10!}, \quad \frac{e^\xi}{10!} < \frac{e}{10!} < \frac{3}{3000000} = \frac{1}{10^6}.$$

不等式は  $\xi < 1$  と  $e < 3$ ,  $10! > 3000000$  より得られる. これで示すべきことが示された.  $\square$

Taylorの定理を使えば誤差項の大きさを上から評価できる.

```
In [9]: 1 sum(1/factorial(k) for k in 0:9), exp(1)
```

Out[9]: (2.7182815255731922, 2.718281828459045)

**問題:**  $\arcsin x$ ,  $\arctan x$  の  $x = 0$  でのTaylor展開の  $x \rightarrow 1$  の極限を取ることによって

$$\frac{\pi}{2} = \sum_{k=0}^{\infty} \frac{1}{2^{2k}} \frac{(2k)!}{(k!)^2} \frac{1}{2k+1}, \quad \frac{\pi}{4} = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}$$

が得られることを確認せよ.  $\square$

解答略. 次の節におけるコンピューターによる計算例も参照せよ.

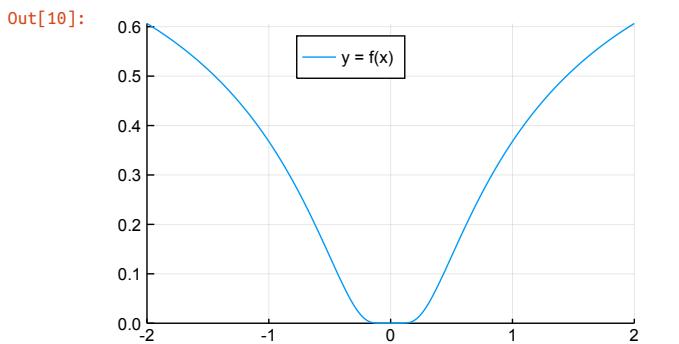
**問題:**  $\mathbb{R}$  上の  $C^\infty$  関数  $f(x)$  で, 任意の実数  $x$  について  $\sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0)x^n$  が収束しているのに, その収束先が  $x \neq 0$  のとき  $f(x)$  と一致しないもののが存在する. そのような  $f(x)$  の例を1つ挙げよ.

**解答例:**  $f(x)$  を

$$f(x) = \begin{cases} \exp(-1/|x|) & (x \neq 0) \\ 0 & (x = 0) \end{cases}$$

と定めると,  $f$  は  $C^\infty$  関数でかつ  $f^{(n)}(0) = 0$  ( $n = 0, 1, 2, \dots$ ) となることを示せる(自分で示してみよ). ゆえにそのとき,  $\sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0)x^n$  は常に 0 に収束する. しかし,  $x \neq 0$  のとき  $f(x) > 0$  なので収束先の 0 は  $f(x)$  に一致しない.  $\square$

```
In [10]: 1 f(x) = x == 0 ? zero(x) : exp(-1/abs(x))
2 x = -2:0.001:2
3 plot(x, f.(x), label="y = f(x)", size=(400,250), legend=:top)
```



## コンピューターを用いた計算例

私が個人的に知る限りの範囲内では, 手計算もきちんとできる人の方がコンピューターを上手に使っている場合が多い. おそらく, その人自身がコンピューターを頼らなくてもある程度の計算ができるおかげで, コンピューターを使った人間には不可能な計算の価値をよく理解できるからなのだろう.

```
In [11]: 1 x = symbols("x")
2 series(e^x, n=10)
```

Out[11]:  $1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \frac{x^6}{720} + \frac{x^7}{5040} + \frac{x^8}{40320} + \frac{x^9}{362880} + \mathcal{O}(x^{10})$

```
In [12]: 1 series(cos(x), n=10)
```

Out[12]:  $1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720} + \frac{x^8}{40320} + \mathcal{O}(x^{10})$

```
In [13]: 1 series(sin(x), n=10)
```

Out[13]:  $x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040} + \frac{x^9}{362880} + \mathcal{O}(x^{10})$

```
In [14]: 1 series(log(1+x), n=10)
```

Out[14]:  $x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \frac{x^6}{6} + \frac{x^7}{7} - \frac{x^8}{8} + \frac{x^9}{9} + \mathcal{O}(x^{10})$

```
In [15]: 1 series(-log(1-x), n=10)
```

Out[15]:  $x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \frac{x^5}{5} + \frac{x^6}{6} + \frac{x^7}{7} + \frac{x^8}{8} + \frac{x^9}{9} + \mathcal{O}(x^{10})$

```
In [16]: 1 series(asin(x), n=10)

Out[16]:  $x + \frac{x^3}{6} + \frac{3x^5}{40} + \frac{5x^7}{112} + \frac{35x^9}{1152} + \mathcal{O}(x^{10})$ 
```

```
In [17]: 1 k = symbols("k", integer=true)
2 doit(Sum(1/2^(2k)*factorial(2k)/(factorial(k)^2)*x^(2k+1), (k,0,4)))

Out[17]:  $\frac{35x^9}{1152} + \frac{5x^7}{112} + \frac{3x^5}{40} + \frac{x^3}{6} + x$ 
```

```
In [18]: 1 series(atan(x), n=10)

Out[18]:  $x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} + \mathcal{O}(x^{10})$ 
```

```
In [19]: 1 Sum((-1)^k/(2k+1), (k,0,oo))

Out[19]:  $\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}$ 
```

```
In [20]: 1 doit(Sum((-1)^k/(2k+1), (k,0,oo)))

Out[20]:  $\frac{\pi}{4}$ 
```

次のセルの  $f(n,k)$  は

$$f(n, k) = \frac{1}{2^n} \binom{n}{k} = \frac{1}{2^n} \frac{n!}{k!(n-k)!}$$

を意味している。これは  $p = 1/2$  に対する二項分布における確率である。

`lgamma(n+1) = log(nの階乗)`

であることに注意せよ。

次のセルで一度全体の対数を取った式を計算してから `exponential` を取っている理由は階乗を単体で計算すると簡単にオーバーフローしてしまうからである。階乗の対数ならばオーバーフローしない。さらに、 $\binom{n}{k}$  と  $1/2^n$  を別々に計算した後で掛け合わせる計算の仕方もよくない。巨大な数を巨大な数で割る数値計算はオーバーフローやアンダーフローを起こしやすい。対数を取って適度な大きさの数値にして差を取ってから、まとめて `exponential` するのがよい。

しかし、そのように注意深く書いたコードであっても、もとの級数の収束は極めて遅いので、1億項を足し合わせても小数点以下第4桁までしか正確に計算できていない。

```
In [21]: 1 f(n,k) = exp(lgamma(n+1)-lgamma(k+1)-lgamma(n-k+1)-n*log(2))
2 @time sum(k->f(2k,k)/(2k+1), 0:10^8-1), pi/2
```

11.108824 seconds (26.78 k allocations: 1.371 MiB, 0.08% gc time)

```
Out[21]: (1.5707399078387505, 1.5707963267948966)
```

## べき級数の収束

### べき級数の定義

$\sum_{k=0}^{\infty} a_k x^k$  の形式の級数を**べき級数**と呼ぶ。

$\sum_{k=0}^{\infty} |a_k x^k|$  が収束するとき、べき級数  $\sum_{k=0}^{\infty} a_k x^k$  は絶対収束するという。

### べき級数の収束

**準備:** 実数列  $\alpha_n$  が  $n \rightarrow \infty$  のとき  $\alpha_n \rightarrow \infty$  となることを、任意の実数  $M$  に対して(どんなに大きな実数  $M$  に対して), ある番号  $N$  で

$$\alpha_n \geq M \quad (n \geq N)$$

を満たすものが存在すること(すなわちある番号以降の  $n$  について  $\alpha_n \geq M$  成立すること)であると定める。□

**定理:**  $k \rightarrow \infty$  のとき  $\frac{|a_k|}{|a_{k+1}|} \rightarrow \rho > 0$  が成立していると仮定する(上の準備のもとで  $\rho = \infty$  であってもよいとする). このとき, べき級数  $\sum_{k=0}^{\infty} a_k x^k$  は  $|x| < \rho$  のとき絶対収束する.

**証明:**  $0 < r < \rho$  と仮定する.  $r$  は幾らでも  $\rho$  に近付き得るので  $|x| < r$  のときべき級数が絶対収束することを示せば十分である.  $|x| < r$  と仮定する. 仮定より, ある番号  $N$  が存在して,  $k \geq N$  のとき,

$$\frac{|a_k|}{|a_{k+1}|} \geq r, \quad \text{i.e.} \quad |a_{k+1}| \leq \frac{1}{r} |a_k|.$$

ゆえに

$$|a_{N+n}| \leq \frac{1}{r^n} |a_N|.$$

したがって,

$$|a_{N+n} x^{N+n}| \leq \frac{1}{r^n} |a_N| |x|^{N+n} = |a_n| |x|^N \left( \frac{|x|}{r} \right)^n.$$

$|x| < r$  という仮定より,  $|x|/r < 1$  なので, 右辺の  $n = 0, 1, 2, \dots$  に関する和は有限の値に収束する. ゆえに左辺の同様の和も有限の値に収束する. これより, べき級数

$$\sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{N-1} a_k x^k + \sum_{n=0}^{\infty} a_{N+n} x^{N+n}$$

が絶対収束することがわかる.  $\square$

**注意:** 上の定理は **d'Alembert の判定法** (ダランベールの判定法) と呼ばれている. それ以外にもべき級数の判定法が存在する. 興味がある人は自分で調べてみて欲しい.  $\square$

- [Googleで「収束半径の求め方」を検索\(\)](#)

**例:** べき級数  $\sum_{k=0}^{\infty} \frac{1}{k!} x^k$  について,  $k \rightarrow \infty$  のとき,

$$\frac{1/k!}{1/(k+1)!} = k+1 \rightarrow \infty$$

なので, そのべき級数は  $|x| < \infty$  で絶対収束する.  $\square$

**例:** べき級数  $\sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!}$  を  $x^2$  のべき級数とみなす.

$$\frac{1/(2k)!}{1/(2(k+1))!} = (2k+1)(2k+2) \rightarrow \infty \quad (k \rightarrow \infty)$$

なので, そのべき級数は  $|x| < \infty$  で絶対収束する.

**問題:** べき級数  $\sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}$  が  $|x| < \infty$  で絶対収束することを示せ.

**ヒント:** そのべき級数と  $x^2$  のべき級数と  $x$  の積とみなしてから, d'Alembertの判定法を適用せよ.  $\square$

## 超幾何級数

次のべき級数を **Gaussの超幾何級数** (Gauss's hypergeometric series) と呼ぶ:

$${}_2F_1(a, b; c; x) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k k!} x^k.$$

ここで

$$(\alpha)_k = \alpha(\alpha+1) \cdots (\alpha+k-1).$$

より一般に

$${}_rF_s(a_1, \dots, a_r; b_1, \dots, b_s; x) = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_r)_k}{(b_1)_k \cdots (b_s)_k k!} x^k$$

と定める.

**問題:** Gaussの超幾何級数  ${}_2F_1$  が  $|x| < 1$  で絶対収束することを示せ.

**解答例:** d'Alembertの判定法を使おう.

$$\frac{(a)_k(b)_k/((c)_kk!)}{(a)_{k+1}(b)_{k+1}/((c)_{k+1}(k+1)!)} = \frac{(c+k)(k+1)}{(a+k)(b+k)} \rightarrow 1 \quad (k \rightarrow \infty)$$

より, Gaussの超幾何級数は  $|x| < 1$  で絶対収束する.  $\square$

**問題:** 以下を示せ:

$$(1)_k = k!, \quad (2)_k = (k+1)!, \quad (1/2)_k k! = \frac{(2k)!}{4^k}, \quad (2k+1)(1/2)_k = (3/2)_k.$$

**ヒント:** 3つ目の公式を除けば易しい. 3つ目の公式は以下のようにして示される:

$$(1/2)_k k! = \frac{1}{2} \frac{3}{2} \cdots \frac{2k-1}{2} k! = \frac{1 \cdot 3 \cdots (2k-1)}{2^k} \frac{2 \cdot 4 \cdots (2k)}{2^k} = \frac{(2k)!}{4^k}. \quad \square$$

**例:** 初等函数を以下のように表すことができる:

$${}_0F_0(; ; x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k = e^x.$$

$${}_0F_1(; 1/2; -x^2/4) = \sum_{k=0}^{\infty} \frac{(-x^2/4)^k}{(1/2)_k k!} = \sum_{k=0}^{\infty} \frac{(-x^2)^k}{(2k)!} = \cos x.$$

$${}_1F_0(a; ; x) = \sum_{k=0}^{\infty} \frac{(a)_k}{k!} x^k = (1-x)^{-a}.$$

$${}_2F_1(1, 1; 2; x) = \sum_{k=0}^{\infty} \frac{k! k!}{(k+1)! k!} x^k = \sum_{k=0}^{\infty} \frac{x^k}{k+1} = -\frac{\log(1-x)}{x}.$$

$$x {}_2F_1(1/2, 1; 3/2; -x^2) = x \sum_{k=0}^{\infty} \frac{(1/2)_k k!}{(3/2)_k k!} (-x^2)^k = x \sum_{k=0}^{\infty} \frac{(-x^2)^k}{2k+1} = \arctan x.$$

**問題:** 上の例の結果を確認せよ.  $\square$

**問題:**  $\sin x$  を  ${}_0F_1$  を使って表せ.  $\square$

**問題:**  $\arcsin x$  を  ${}_2F_1$  を使って表せ.  $\square$

**ヒント:** 「超幾何 初等函数」を検索 (<https://www.google.co.jp/search?q=%E8%B6%85%E5%B9%BE%E4%BD%95+%E5%88%9D%E7%AD%89%E5%87%BD%E6%95%BD>).

**問題:** ガンマ函数とベータ函数を

$$\Gamma(s) = \int_0^\infty e^{-x} x^{s-1} dx, \quad B(p, q) = \int_0^1 t^{p-1} (1-t)^{q-1} dt \quad (s, p, q > 0)$$

と定めると,

$$(x)_k = \frac{\Gamma(x+k)}{\Gamma(x)}, \quad B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

が成立することを認めて,  $c > b > 0$  のとき

$${}_2F_1(a, b; c; x) = \frac{1}{B(b, c-b)} \int_0^1 t^{b-1} (1-t)^{c-b-1} (1-xt)^{-a} dt$$

が成立することを示せ. これを **Gaussの超幾何函数の積分表示** と呼ぶ.

**解答例:**

$$\begin{aligned} \frac{(a)_k(b)_k}{(c)_k k!} x^k &= \frac{\Gamma(b+k)\Gamma(c)}{\Gamma(b)\Gamma(c+k)} \frac{(a)_k}{k!} x^k = \frac{B(b+k, c-b)}{B(b, c-b)} \frac{(a)_k}{k!} x^k \\ &= \frac{1}{B(b, c-b)} \int_0^1 t^{b-1} (1-t)^{c-b-1} \frac{(a)_k}{k!} (xt)^k dt. \end{aligned}$$

ゆえに, これを  $k = 0, 1, 2, \dots$  について足し上げて目的の公式を得る.  $\square$

**問題:** 以下の公式を示せ:  $c > b > 0$  のとき,

$${}_1F_1(b; c; x) = \frac{1}{B(b, c-b)} \int_0^1 t^{b-1} (1-t)^{c-b-1} e^{xt} dt$$

となることを示せ. これを **Kummerの合流型超幾何函数の積分表示** と呼ぶ.

$$\begin{aligned}\frac{(b)_k}{(c)_k k!} x^k &= \frac{\Gamma(b+k)\Gamma(c)}{\Gamma(b)\Gamma(x+k)} \frac{x^k}{k!} = \frac{B(b+k, c-b)}{B(b, c-b)} \frac{x^k}{k!} \\ &= \frac{1}{B(b, c-b)} \int_0^1 t^{b-1} (1-t)^{c-b-1} \frac{(xt)^k}{k!} dt.\end{aligned}$$

ゆえに、これを  $k = 0, 1, 2, \dots$  について足し上げて目的の公式を得る.  $\square$

In [ ]:

1

# 漸近展開の有名な例

黒木玄

2018-05-20

- Copyright 2018 Gen Kuroki
- License: MIT <https://opensource.org/licenses/MIT> (<https://opensource.org/licenses/MIT>)

このファイルは次の場所できれいに閲覧できる:

- <http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/07%20example%20of%20asymptotic%20expansion.ipynb> (<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/07%20example%20of%20asymptotic%20expansion.ipynb>)

このファイルは [Julia Box \(https://juliabox.com\)](https://juliabox.com) で利用できる。

自分のパソコンに[Julia言語 \(https://julialang.org/\)](https://julialang.org/)をインストールしたい場合には

- [WindowsへのJulia言語のインストール \(http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f\)](http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f)

を参照せよ。

論理的に完璧な説明をするつもりはない。細部のいい加減な部分は自分で訂正・修正せよ。

## 目次

### 1 漸近展開の有名な例

- [1.1  \$F\_n\(x\)\$  の定義](#)
- [1.2  \$F\_n\(x\) = O\(x^{n+1}\)\$](#)
- [1.3  \$F\_0\(x\)\$  の漸近展開](#)
- [1.4  \$F\_0\(x\)\$  の漸近展開の別の導出の仕方](#)
- [1.5 以上のまとめ](#)
- [1.6  \$F\_0\(1/10\)\$  の漸近展開を用いた数値計算](#)
- [1.7 漸近展開の入門書](#)
- [1.8 指数積分函数](#)
- [1.9 連分数展開による数値計算](#)

```
In [1]: 1  using Plots
2  gr(); ENV["PLOTS_TEST"] = "true"
3  #clibrary(:colorcet)
4  clibrary(:misc)
5
6  ▼ function pngplot(P...; kwargs...)
7    sleep(0.1)
8    pngfile = tempname() * ".png"
9    savefig(plot(P...; kwargs...), pngfile)
10   showimg("image/png", pngfile)
11 end
12 pngplot(; kwargs...) = pngplot(plot!(; kwargs...))
13
14 ▼ showimg(mime, fn) = open(fn) do f
15   base64 = base64encode(f)
16   display("text/html", """""")
17 end
18
19 using SymPy
20 #sympy[:init_printing](order="lex") # default
21 #sympy[:init_printing](order="rev-lex")
22
23 using SpecialFunctions
24 using QuadGK
```

## 1 漸近展開の有名な例

### 1.1 $F_n(x)$ の定義

$n$  は非負の整数であるとする。 $x > 0$  の函数  $F_n(x)$  を

$$F_n(x) = n! e^{1/x} \int_{1/x}^{\infty} e^{-t} t^{-n-1} dt$$

と定める。 $F_n(x)$  の  $x > 0$  が 0 に近いときの様子を調べたい。

## 1.2 $F_n(x) = O(x^{n+1})$

$x > 0$  のとき  $t \geq 1/x$  ならば  $t^{-n-1} \leq x^{n+1}$  なので

$$0 < F_n(x) \leqq n! e^{1/x} \int_{1/x}^{\infty} e^{-t} t^{n+1} dt = n! x^{n+1}.$$

ゆえに,

$$F_n(x) = O(x^{n+1}) \quad (x \searrow 0).$$

## 1.3 $F_0(x)$ の漸近展開

$(-e^{-t})' = e^{-t}$  を用いた部分積分によって

$$\begin{aligned} F_n(x) &= n! e^{1/x} \left( [-e^{-t} t^{-n-1}]_{1/x}^{\infty} + \int_{1/x}^{\infty} e^{-t}(-(n+1)t^{-n-2}) dt \right) \\ &= n! x^{n+1} - F_{n+1}(x). \end{aligned}$$

ゆえに  $x > 0$  において,

$$\begin{aligned} F_0(x) &= 0! x - F_1(x) = 0! x - 1! x^2 + F_2(x) = \dots \\ &= 0! x - 1! x^2 + \dots + (-1)^n n! x^{n+1} + (-1)^{n+1} F_{n+1}(x) \\ &= \sum_{k=0}^n (-1)^k k! x^{k+1} + (-1)^{n+1} F_{n+1}(x). \end{aligned}$$

## 1.4 $F_0(x)$ の漸近展開の別の導出の仕方

$x > 0$  であるとする.  $F_0(x)$  は  $t = 1/x + u$  という置換によって次のように書き直される:

$$F_0(x) = e^{1/x} \int_{1/x}^{\infty} \frac{e^{-t}}{t} dt = x \int_0^{\infty} \frac{e^{-u}}{1+xu} du.$$

ゆえに

$$\frac{1}{1+z} = \sum_{k=0}^{n-1} (-1)^k z^k + (-1)^n \frac{z^n}{1+z}$$

を  $z = xu$  に適用した結果を使うと,

$$\begin{aligned} F_0(x) &= \sum_{k=0}^{n-1} (-1)^k x^{k+1} \int_0^{\infty} e^{-u} u^k du + (-1)^n x^{n+1} \int_0^{\infty} \frac{e^{-u} u^n}{1+xu} du \\ &= \sum_{k=0}^{n-1} (-1)^k k! x^{k+1} + (-1)^n x^{n+1} \int_0^{\infty} \frac{e^{-u} u^n}{1+xu} du. \end{aligned}$$

さらに

$$0 < x^{n+1} \int_0^{\infty} \frac{e^{-u} u^n}{1+xu} du \leqq x^{n+1} \int_0^{\infty} e^{-u} u^n du = n! x^{n+1}$$

なので

$$x^{n+1} \int_0^{\infty} \frac{e^{-u} u^n}{1+xu} du = O(x^{n+1}) \quad (x \searrow 0).$$

## 1.5 以上のまとめ

以上をまとめると

$$F_0(x) = \sum_{k=0}^{n-1} (-1)^k k! x^{k+1} + (-1)^n F_n(x) = \sum_{k=0}^{n-1} (-1)^k k! x^{k+1} + O(x^{n+1}) \quad (x \searrow 0).$$

ここで

$$F_n(x) = n! e^{1/x} \int_{1/x}^{\infty} e^{-t} t^{-n-1} dt = x^{n+1} \int_0^{\infty} \frac{e^{-u} u^n}{1+xu} du = O(x^{n+1}) \quad (x \searrow 0).$$

しかし,  $x > 0$  で

$$\sum_{k=0}^{\infty} (-1)^k k! x^{k+1} = 0! x - 1! x^2 + 2! x^3 - 3! x^4 + \dots$$

は決して収束しない.

このように全てを足し上げると発散する場合であっても、有限項と剰余項の和の形式で表わせば、すべてがwell-definedな量だけを使って議論を進めることができる場合がある。上の場合には剰余項は $(-1)^n F_n(x)$ である。

## 1.6 $F_0(1/10)$ の漸近展開を用いた数値計算

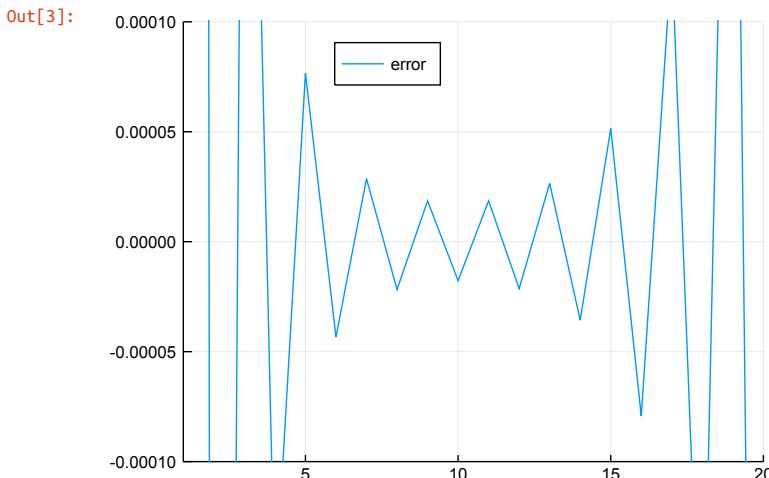
$x > 0$  のとき  $0 < F_n(x) \leq n! x^{n+1}$  となるのであった。 $x = 1/10$  のとき  $n! x^{n+1}$  が最小になる  $n$  は  $n = 9, 10$  であることを確認できる。そのとき  $n! x^{n+1} = 3.6288 \times 10^{-5}$  である。したがって、 $n = 9, 10$  のときの  $\sum_{k=0}^{n-1} (-1)^k k! x^{k+1}$  による  $F_0(x)$  の近似は最良になると予想される。

In [2]: 1 [(n, factorial(n)\*(1/10)^(n+1)) for n in 7:13]

Out[2]: 7-element Array{Tuple{Int64,Float64},1}:  
 (7, 5.04e-5)  
 (8, 4.032e-5)  
 (9, 3.6288e-5)  
 (10, 3.6288e-5)  
 (11, 3.99168e-5)  
 (12, 4.79002e-5)  
 (13, 6.22702e-5)

In [3]: 1  $F_0(x) = \exp(1/x) * \text{quadgk}(u \rightarrow \exp(-1/(x*u))/u, 0, 1)[1]$   
 2  $F_0\_ae(x, n=10) = \text{sum}(k \rightarrow (-1)^k k! \text{factorial}(k) * x^k, 0:n-1)$   
 3  
 4 @show Y = F\_0(1/10)  
 5 @show Y\_ae9 = F\_0\_ae(1/10, 9)  
 6 @show Y\_ae10 = F\_0\_ae(1/10, 10)  
 7  
 8 n = 1:20  
 9 plot(size=(500,350))  
 10 plot!(ylims=(-0.0001,0.0001))  
 11 plot!(:legend=:top)  
 12 plot!(n, F\_0\_ae.(1/10, n)-Y, label="error")

Y = F\_0(1 / 10) = 0.09156333393978826  
 Y\_ae9 = F\_0\_ae(1 / 10, 9) = 0.09158192  
 Y\_ae10 = F\_0\_ae(1 / 10, 10) = 0.091545632



## 1.7 漸近展開の入門書

このように  $x > 0$  で決して収束しないべき級数であっても、適切な解釈のもとで数学的に十分な意味を持つことがある。このような理由で数学では発散級数に関するたくさんの深い研究がある。

その方面的入門書としては次の文献がある。非常に面白い本なのでおすすめできる。

- 大久保謙二郎・河野實彦共著『漸近展開』、新しい応用の数学12、教育出版、1976、1996 (<https://www.amazon.co.jp/dp/4316376306>)。

## 1.8 指数積分函数

$F_0(x)$  は実用的に重要な函数である。それと本質的に同じ

$$E_1(z) = e^{-z} F_0(1/z) = \int_z^\infty \frac{e^{-t}}{t} dt = \int_0^1 \frac{e^{-zt}}{u} du$$

は指数積分函数と呼ばれる特殊函数の1つである( $t = z/u$ )。この函数の数値計算については

- <http://nbviewer.jupyter.org/github/stevengj/18S096/blob/iap2017/pset3/pset3-solutions.ipynb>  
[\(http://nbviewer.jupyter.org/github/stevengj/18S096/blob/iap2017/pset3/pset3-solutions.ipynb\)](http://nbviewer.jupyter.org/github/stevengj/18S096/blob/iap2017/pset3/pset3-solutions.ipynb)

が非常に面白い解説になっている。特殊函数の数値計算の最適化に興味がある人は是非とも閲覧して欲しい。

## 1.9 連分数展開による数値計算

以下は  $x = 1/z$  のときの

$$G(z) = F_0(1/z) = e^z E_1(z) = e^z \int_0^1 \frac{e^{-yu}}{u} du$$

の連分数展開による数値計算である。この函数の連分数展開については

- 一松信著『特殊関数入門』数学選書、森北出版、1999 (<https://www.amazon.co.jp/dp/4627038216>)

の第3章の最後のp.74にある例3.10および第6章のpp.132-133にある例6.4を参照せよ。この本も非常に面白い本なのでおすすめできる。

```
In [4]: 1 ▼ function G_cf(z; n::Int=2)
2      cf = 1 + (n+1)/z
3 ▼      for i = n:-1:1
4          cf = z + (1+i)/cf
5          cf = 1 + i/cf
6      end
7      return 1 / (z + 1/cf)
8  end
9
10 G(z) = exp(z)*quadgk(u->exp(-z/u)/u, 0, 1)[1]
```

Out[4]: G (generic function with 1 method)

```
In [5]: 1 sympy[:init_printing](order="lex")
2 z = symbols("z")
3 ▼ for n in 1:3
4     cf = G_cf(z, n=n)
5     display(cf)
6 end
```

$$\begin{aligned} & \frac{1}{z + \frac{1}{1 + \frac{1}{z + \frac{2}{1 + \frac{2}{z}}}}} \\ & \frac{1}{z + \frac{1}{1 + \frac{1}{z + \frac{2}{1 + \frac{2}{z + \frac{3}{1 + \frac{3}{z}}}}}}} \\ & \frac{1}{z + \frac{1}{1 + \frac{1}{z + \frac{2}{1 + \frac{2}{z + \frac{3}{1 + \frac{3}{z + \frac{4}{1 + \frac{4}{z}}}}}}}}} \end{aligned}$$

```
In [6]: 1 sympy[:init_printing](order="rev-lex")
2 display(series(sum((-1)^n*factorial(n)*z^(n+1) for n in 0:9), n=11))
3 ▼ for n in 1:3
4     display(series(G_cf(1/z, n=n), n=11))
5 end
```

$$\begin{aligned} & z - z^2 + 2z^3 - 6z^4 + 24z^5 - 120z^6 + 720z^7 - 5040z^8 + 40320z^9 - 362880z^{10} \\ & z - z^2 + 2z^3 - 6z^4 + 24z^5 - 108z^6 + 504z^7 - 2376z^8 + 11232z^9 - 53136z^{10} + \mathcal{O}(z^{11}) \\ & z - z^2 + 2z^3 - 6z^4 + 24z^5 - 120z^6 + 720z^7 - 4896z^8 + 35712z^9 - 269568z^{10} + \mathcal{O}(z^{11}) \\ & z - z^2 + 2z^3 - 6z^4 + 24z^5 - 120z^6 + 720z^7 - 5040z^8 + 40320z^9 - 360000z^{10} + \mathcal{O}(z^{11}) \end{aligned}$$

$n = 1, 2, 3$  の連分数近似であっても、大きめの  $z$  における近似が相當にうまく行っていることを以下で確認する。

```
In [7]: 1 G_cf(4), G(4)
```

Out[7]: (0.20636792452830188, 0.20634564990105575)

```
In [8]: 1 x = 0.05:0.05:4
2 @time y = G.(x)
3 @time y_cf1 = G_cf.(x, n=1)
4 @time y_cf2 = G_cf.(x, n=2)
5 @time y_cf3 = G_cf.(x, n=3)
6 plot(size=(500,350))
7 plot!(title="y = G(x) = e^x E_1(x)", titlefontsize=11)
8 plot!(xlims=(0,maximum(x)), ylims=(0,2.3))
9 plot!(x, y, label="numerical integration")
10 plot!(x, y_cf1, label="continued fractional approximation n=1", ls=:dash)
11 plot!(x, y_cf2, label="continued fractional approximation n=2", ls=:dash)
12 plot!(x, y_cf3, label="continued fractional approximation n=3", ls=:dash)
```

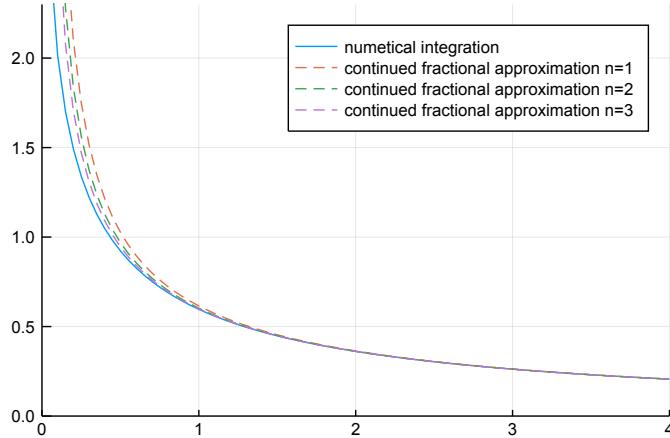
2.022534 seconds (266.78 k allocations: 13.742 MiB)

0.047182 seconds (16.07 k allocations: 832.076 KiB)

0.016745 seconds (5.59 k allocations: 284.821 KiB)

0.016846 seconds (5.59 k allocations: 285.634 KiB)

Out[8]:  $y = G(x) = e^x E_1(x)$



In [ ]:

1

# 函数の凸性と不等式への応用

黒木玄

2018-06-11

- Copyright 2018 Gen Kuroki
- License: MIT <https://opensource.org/licenses/MIT> (<https://opensource.org/licenses/MIT>)

このファイルは次の場所できれいに閲覧できる:

- <http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/08%20convexity.ipynb>  
(<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/08%20convexity.ipynb>)

このファイルは [Julia Box](https://juliabox.com) (<https://juliabox.com>) で利用できる。

自分のパソコンにJulia言語 (<https://julialang.org/>) をインストールしたい場合には

- [WindowsへのJulia言語のインストール](http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f) (<http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f>)

を参照せよ。

論理的に完璧な説明をするつもりはない。細部のいい加減な部分は自分で訂正・修正せよ。

## Table of Contents

<a href="#">1 凸函数の定義</a>
<a href="#">2 2階の導函数の符号と函数の凸性</a>
<a href="#">3 Jensenの不等式</a>
<a href="#">3.1 期待値函数</a>
<a href="#">3.2 Jensenの不等式とその証明</a>
<a href="#">3.3 相加相乗平均の不等式</a>
<a href="#">3.4 Youngの不等式</a>
<a href="#">3.5 Hölderの不等式</a>
<a href="#">3.5.1 <math>L^p</math> ノルム</a>
<a href="#">3.5.2 一般化されたヘルダーの不等式</a>
<a href="#">3.5.3 分配函数の対数凸性</a>
<a href="#">3.6 Minkowskiの不等式</a>

```
In [1]: 1 using Plots
2 gr(); ENV["PLOTS_TEST"] = "true"
3 #clibrary(:colorcet)
4 clibrary(:misc)
5
6 ▼ function pngplot(P...; kwargs...)
7   sleep(0.1)
8   pngfile = tempname() * ".png"
9   savefig(plot(P...; kwargs...), pngfile)
10  showimg("image/png", pngfile)
11 end
12 pngplot(; kwargs...) = pngplot(plot!(; kwargs...))
13
14 ▼ showimg(mime, fn) = open(fn) do f
15   base64 = base64encode(f)
16   display("text/html", """""")
17 end
18
19 using SymPy
20 #sympy[:init_printing](order="lex") # default
21 #sympy[:init_printing](order="rev-lex")
22
23 using SpecialFunctions
24 using QuadGK
```

## 1 凸函数の定義

**定義:** 区間  $I$  上の実数値函数  $f$  が下に凸な函数であるとは、任意の  $a, b \in I$  と実数  $t$  について、

$$0 \leq t \leq 1 \implies f((1-t)a + tb) \leq (1-t)f(a) + tf(b)$$

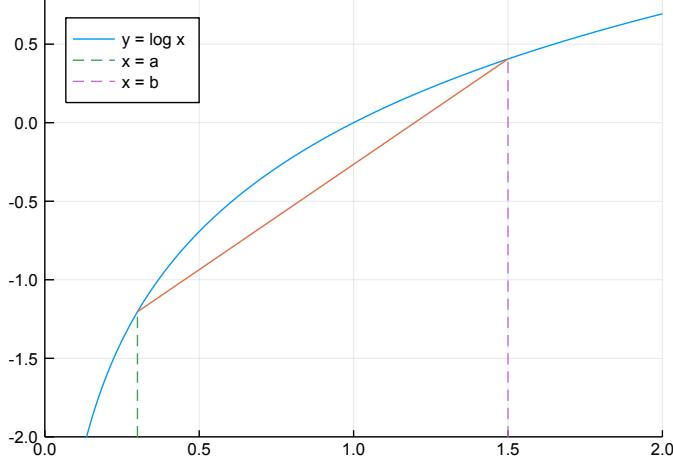
を満たしていることである。逆向きの不等式で上に凸な函数を定義する。すなわち、区間  $I$  上の実数値函数  $f$  が上に凸な函数であるとは、任意の  $a, b \in I$  と実数  $t$  について、

$$0 \leq t \leq 1 \implies f((1-t)a + tb) \geq (1-t)f(a) + tf(b)$$

を満たしていることである。

```
In [2]: 1 ▼ # log x は上に凸な函数
2
3 x = 0:0.01:2.0
4 a, b = 0.3, 1.5
5 f(x) = log(x)
6 t = 0:0.01:1.0
7 g(a,b,t) = (1-t)*f(a) + t*f(b)
8 X(a,b,t) = (1-t)*a + t*b
9 plot(size=(500,350), legend=:topleft, xlims=(0,2.0), ylims=(-2.0, 0.8))
10 plot!(x, f.(x), label="y = log x")
11 plot!(X.(a,b,t), g.(a,b,t), label="")
12 plot!([a,a], [-10.0, f(a)], label="x = a", ls=:dash)
13 plot!([b,b], [-10.0, f(b)], label="x = b", ls=:dash)
```

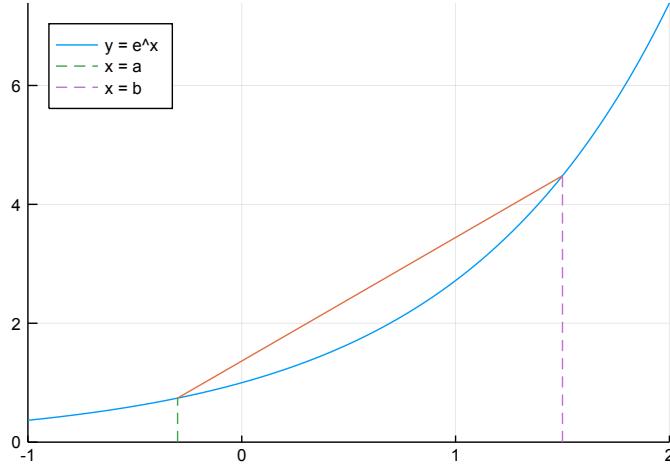
Out[2]:



In [3]:

```
1 ▼ # e^x は下に凸な函数
2
3 x = -1:0.01:2
4 a, b = -0.3, 1.5
5 f(x) = e^x
6 t = 0:0.01:1.0
7 g(a,b,t) = (1-t)*f(a) + t*f(b)
8 X(a,b,t) = (1-t)*a + t*b
9 plot(size=(500,350), legend=:topleft, xlims=(-1,2))
10 plot!(x, f.(x), label="y = e^x")
11 plot!(X.(a,b,t), g.(a,b,t), label="")
12 plot!([a,a], [-0.0, f(a)], label="x = a", ls=:dash)
13 plot!([b,b], [-0.0, f(b)], label="x = b", ls=:dash)
```

Out[3]:



## 2 階の導函数の符号と函数の凸性

**定理:** 区間  $I$  上の実数値函数  $f$  が  $C^2$  級かつ  $I$  上で  $f'' \geq 0$  ならば  $f$  は下に凸である。同様に区間  $I$  上の実数値函数  $f$  が  $C^2$  級かつ  $I$  上で  $f'' \leq 0$  ならば  $f$  は上に凸である。

**証明:** 前者のみを証明する。区間  $I$  上の実数値函数  $f$  が  $C^2$  級かつ  $I$  上で  $f'' \geq 0$  であると仮定する。 $a, b \in I$  と仮定し,  $c = (1-t)a + tb$  とおく。 $0 \leq t \leq 1$  であると仮定する。 $(1-t)f(a) + tf(b) \geq f(c)$  を示せばよい。 $f$  は  $C^2$  級なので Taylor の公式より,

$$\begin{aligned}f(a) &= f(c) + f'(c)(a - c) + R(a), \\f(b) &= f(c) + f'(c)(b - c) + R(b), \\R(x) &= \int_c^x dx_1 \int_c^{x_1} dx_2 f''(x_2).\end{aligned}$$

$f'' \geq 0$  という仮定より,  $R(x) \geq 0$  となることがわかる. その理由は以下の通り.

$\$F(x)=\text{int\_c}^x;$

(1)  $x \leq c$  のとき. 積分  $R(x)$  の中の  $x_1$  は  $x \leq x_1 \leq c$  の範囲を動き, そのとき  $f'' \geq 0$  より  $\int_c^{x_1} dx_2 f''(x_2) \leq 0$  となるので,

ゆえに

$$\begin{aligned}f(a) &\geq f(c) + f'(c)(a - c), \\f(b) &\geq f(c) + f'(c)(b - c).\end{aligned}$$

1つ目の不等式の両辺に  $(1 - t)$  をかけ, 2つめの不等式の両辺に  $t$  をかけると

$$\begin{aligned}(1 - t)f(a) &\geq (1 - t)f(c) + f'(c)((1 - t)a - (1 - t)c), \\tf(b) &\geq tf(c) + f'(c)(tb - tc).\end{aligned}$$

これらをたすと,  $c = (1 - t)a + tb$  とおいたことより,

$$(1 - t)f(a) + tf(b) \geq f(c) + f'(c)((1 - t)a + tb - c) = f(c).$$

これで  $f$  が下に凸であることが示された.  $\square$

**注意:** 上の証明の本質はTaylorの公式の剩余項  $R, S$  が0以上になることである.  $f$  が  $C^2$  級であるという仮定を2回微分可能性の仮定にゆるめても,  $c$  と  $a$  のあいだのある実数  $\xi$  が存在して,

$$f(a) = f(c) + f'(c)(a - c) + R, \quad R = \frac{1}{2}f''(\xi)(c - a)$$

となるという形でTaylorの定理が成立している.  $f'' \geq 0$  ならば  $R \geq 0$  となる. このことを使えば上の定理で  $C^2$  級の仮定を2回微分可能であるという仮定に弱めることができる.  $\square$

**問題:** 上と同様にして  $f'' \leq 0$  ならば  $f$  が上に凸であることの証明を書き下せ. さらに上の定理における  $C^2$  級の仮定を2回微分可能性に弱めた場合の証明も書き下せ.  $\square$

**例:**  $f(x) = e^{ax}$  とおくと  $f'(x) = ae^{ax}, f''(x) = a^2e^{ax} \geq 0$  を満たしているので,  $f$  は下に凸な函数である.  $\square$

**例:**  $x > 0$  に対して  $f(x) = \log x$  とおくと  $f'(x) = 1/x, f''(x) = -1/x^2 < 0$  なので,  $f$  は上に凸な函数である.  $\square$

**例題:**  $a$  は実数であるとし,  $x > 0$  に対して  $f(x) = x^a$  とおく.  $f(x)$  は凸性について調べよ.

**解答例:**  $x > 0$  であるとする.  $f'(x) = ax^{a-1}, f''(x) = a(a-1)x^{a-2}$  であり,  $x^{a-2} > 0$  である. ゆえに,  $a(a-1) \geq 0$  のとき, すなわち  $a \leq 0$  または  $1 \leq a$  のとき,  $f$  は下に凸になり,  $a(a-1) \leq 0$  のとき, すなわち  $0 \leq a \leq 1$  のとき,  $f$  は上に凸になる.  $\square$

## Jensenの不等式

### 期待値汎函数

函数を数に対応させる函数を汎函数 (functional)と呼ぶことがある.

**函数の例:**  $x$  を  $x$  に対応させる函数を  $x$  と書く.

すべての  $x$  を一定の値  $\alpha$  に対応させる函数も同じ記号の  $\alpha$  で書き, 定数函数と呼ぶ.  $f(x) = \alpha$  とおいて記号  $f$  を利用すると煩雑になるので, 単に  $\alpha$  と書いて代用するということである. 数としての  $\alpha$  と定数函数としての  $\alpha$  を同じ記号で書くので混乱しないように注意して欲しい.

例えば, すべての  $x$  を 1 に対応させる函数も単に 1 と書く.  $\square$

**定義:** 函数  $f$  を数  $E[f]$  に対応させる函数  $E[\cdot]$  (汎函数)が以下の条件を満たしているとき,  $E[\cdot]$  は期待値汎函数 (expectation functional)であると言ふことにする:

(1) 函数  $f, g$  と数  $\alpha, \beta$  に対して,  $E[\alpha f + \beta g] = \alpha E[f] + \beta E[g]$  (線形性).

(2) 函数  $f, g$  のあいだで常に  $f \leq g$  が成立しているならば  $E[f] \leq E[g]$  (単調性).

(3) 定数函数  $\alpha$  について,  $E[\alpha] = \alpha$  (規格化条件).  $\square$

**例:** 区間  $I$  上の函数を数  $E[f]$  に対応させる函数  $E[\cdot]$  を以下のように定めると,  $E[\cdot]$  は期待値汎函数になる. まず  $a_1, \dots, a_n \in I$  を任意に取る.  $w_1, \dots, w_n$  は0以上の実数でかつ  $w_1 + \dots + w_n = 1$  を満たしていると仮定する. そして区間  $I$  上の函数  $f$  に対して

$$E[f] = \sum_{i=1}^n w_i f(a_i) = w_1 f(a_1) + \dots + w_n f(a_n)$$

と定める。このようにして定められた  $E[\cdot]$  が実際に期待値汎函数の性質を満たしていることは以下のようにして確認できる。

(1) 区間  $I$  上の函数  $f, g$  と数  $\alpha, \beta$  について

$$\begin{aligned} E[\alpha f + \beta g] &= \sum_{i=1}^n w_i (\alpha f(a_i) + \beta g(a_i)) = \sum_{i=1}^n (\alpha w_i f(a_i) + \beta w_i g(a_i)) \\ &= \alpha \sum_{i=1}^n w_i f(a_i) + \beta \sum_{i=1}^n w_i g(a_i) = \alpha E[f] + \beta E[g]. \end{aligned}$$

(2) 区間  $I$  上で  $f \leq g$  が成立していると仮定すると,  $w_i \geq 0$  より  $w_i f(a_i) \leq w_i g(a_i)$  も成立するので

$$E[f] = \sum_{i=1}^n w_i f(a_i) \leq \sum_{i=1}^n w_i g(a_i) = E[g].$$

(3) 定数函数  $f(x) = \alpha$  について  $f(a_i) = \alpha$  が成立しているので,  $w_1 + \dots + w_n = 1$  という仮定より,

$$E[\alpha] = E[f] = \sum_{i=1}^n w_i f(a_i) = \sum_{i=1}^n w_i \alpha = \alpha.$$

これで示すべきことがすべて示された。

特に  $w_i = 1/n$  であるとき,  $E[f]$  は

$$E[f] = \frac{f(a_1) + \dots + f(a_n)}{n}$$

と  $f(a_i)$  たちの加法平均になる.  $\square$

抽象的な期待値汎函数の概念が難しいと感じる人は, 一般的な期待値汎函数の定義を忘れてこの例の  $E[f]$  またはさらにその特別な場合である加法平均の場合のみを考えれば十分である。

**問題:**  $a < b$  であるとする. 閉区間  $I = [a, b]$  上の連続函数  $f$  を

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

に対応させる函数  $E[\cdot]$  は期待値汎函数であることを示せ.  $\square$

**問題:**  $\mathbb{R}$  上の函数  $p(x)$  は  $p(x) \geq 0$  ( $x \in \mathbb{R}$ ) と  $\int_{-\infty}^{\infty} p(x) dx = 1$  を満たしていると仮定する。(このような  $p(x)$  は確率密度函数と呼ばれる。) このとき,  $\mathbb{R}$  上の(適当によい条件を仮定した)函数  $f$  を

$$E[f] = \int_{-\infty}^{\infty} f(x)p(x) dx$$

に対応させる函数  $E[\cdot]$  が期待値汎函数であることを示せ。 $(E[f]$  は確率密度函数  $p(x)$  が定める確率分布に関する確率変数  $f$  の期待値と呼ばれる。)

**注意:** 上の問題はこの問題で

$$p(x) = \begin{cases} \frac{1}{b-a} & (a \leq x \leq b) \\ 0 & (\text{それ以外}) \end{cases}$$

とした場合になっている.  $\square$

**解答例:** 以下では  $\int_{-\infty}^{\infty}$  を  $\int$  と略記する。

(1) 数  $\alpha, \beta$  と函数  $f, g$  について

$$E[\alpha f + \beta g] = \int (\alpha f(x) + \beta g(x))p(x) dx = \alpha \int f(x)p(x) dx + \beta \int g(x)p(x) dx = \alpha E[f] + \beta E[g].$$

(2)  $\mathbb{R}$  上で  $f \leq g$  が成立していると仮定すると,  $p(x) \geq 0$  より  $f(x)p(x) \leq g(x)p(x)$  も成立するので

$$E[f] = \int f(x)p(x) dx \leq \int g(x)p(x) dx = E[g].$$

(3) 定数函数  $f(x) = \alpha$  について,  $\int p(x) dx = 1$  という仮定より,

$$E[\alpha] = \int \alpha p(x) dx = \alpha.$$

これで示すべきことがすべて示された.  $\square$

### Jensenの不等式とその証明

**Jensenの不等式:**  $E[\cdot]$  は区間  $I$  上の函数の期待値汎函数であるとし,  $f$  は区間  $I$  上の上に凸(もしくは下に凸)な函数であるとする。このとき

$$E[f(x)] \leq f(E[x]) \quad (\text{もしくは } E[f(x)] \geq f(E[x])).$$

例えば

$$E[f(x)] = \sum_{i=1}^n w_i f(a_i), \quad w_i \geq 0, \quad \sum_{i=1}^n w_i = 1$$

のとき,

$$w_1 f(a_1) + \cdots + w_n f(a_n) \leq f(w_1 a_1 + \cdots + w_n a_n)$$

(もしくは  $w_1 f(a_1) + \cdots + w_n f(a_n) \geq f(w_1 a_1 + \cdots + w_n a_n)$ ).

特にこの  $n = 2$  の場合には  $f$  が上に凸(もしくは下に凸)であることの定義とこれらは同値であることに注意せよ。すなわち、函数  $f$  が Jensen の不等式を満たすことと上(もしくは下)に凸なことは同値になる。この意味で Jensen の不等式は函数の凸性を言い換えた結果になっている。

**証明:** 簡単のため  $f$  は  $C^1$  級であると仮定する。 $(C^1$  級でない場合にも同様の方法で証明できるが、「接線」の存在を別に証明する必要が生じる。)

$f$  は上に凸であると仮定する。

$\mu = E[x]$  とおく。 $E[f(x)] \leq f(\mu)$  を示せばよい。 $x = \mu$  における  $y = f(x)$  の接線を  $y = a(x - \mu) + f(\mu)$  と書く。 $f$  は上に凸であると仮定したので

$$f(x) \leq a(x - \mu) + f(\mu)$$

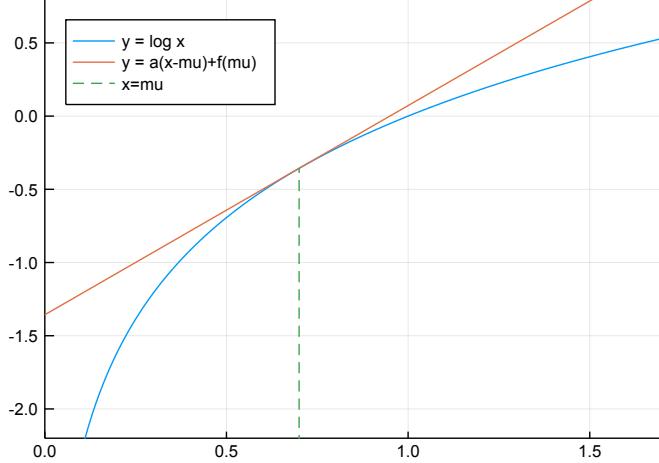
が成立している(図を描いてみよ)。ゆえに  $E[\cdot]$  の性質より

$$E[f(x)] \leq E[a(x - \mu) + f(\mu)] = a(E[x] - E[\mu]) + E[f(\mu)] = f(\mu).$$

最初の不等号は期待値汎函数の単調性より、1つ目の等号は線形性より、2つ目の等号は規格化条件より。これで示すべきことがすべて示された。□

```
In [4]: 1 x = 0:0.01:2.0
2 μ = 0.7
3 f(x) = log(x)
4 g(μ,x) = (1/μ)*(x-μ) + f(μ)
5 plot(size=(500,350), legend=:topleft, xlims=(0,1.7), ylims=(-2.2, 0.8))
6 plot!(x, f.(x), label="y = log x")
7 plot!(x, g.(μ,x), label="y = a(x-μ)+f(μ)")
8 plot!([μ, μ], [-3.0, f(μ)], label="x=μ", ls=:dash)
```

Out[4]:



**問題:** 上に凸な函数  $f(x)$  と

$$E[g(x)] = w_1 g(a_1) + \cdots + w_n g(a_n), \quad w_i \geq 0, \quad w_1 + \cdots + w_n = 1$$

に関する Jensen の不等式

$$w_1 f(a_1) + \cdots + w_n f(a_n) \leq f(w_1 a_1 + \cdots + w_n a_n) \quad (*)$$

を  $n$  に関する数学的帰納法を用いて直接証明せよ。

**解答例:**  $n = 1$  のとき、 $(*)$  は  $f(a_1) \leq f(a_1)$  を意味するので、自明に正しい。

$n = N \geq 1$  に関する  $(*)$  が常に成立していると仮定し(帰納法の仮定),

$$w_i \geq 0, \quad w_1 + \cdots + w_N = 1$$

と仮定する。もしも  $w_{N+1} = 1$  ならば  $n = N + 1$  の場合の  $(*)$  は自明に成立している。以下では  $w_{N+1} < 1$  と仮定し,

$$t_i = \frac{w_i}{1 - w_{N+1}}$$

とおく。このとき、 $w_i \geq 0, w_1 + \cdots + w_N = 1 - w_{N+1}$  より,

$$t_i \geq 0, \quad t_1 + \cdots + t_N = \frac{w_1 + \cdots + w_N}{1 - w_{N+1}} = 1.$$

したがって,  $n = N$  の場合の(\*)より,

$$t_1 f(a_1) + \cdots + t_N f(a_N) \leq f(t_1 a_1 + \cdots + t_N a_N).$$

$f$  は上に凸と仮定してあつたので,  $w_i = (1 - w_{N+1})t_i$  を使うと,

$$\begin{aligned} w_1 f(a_1) + \cdots + w_N f(a_N) + w_{N+1} f(a_{N+1}) &= (1 - w_{N+1})(t_1 f(a_1) + \cdots + t_N f(a_N)) + w_{N+1} f(a_{N+1}) \\ &\leq (1 - w_{N+1})f(t_1 a_1 + \cdots + t_N a_N) + w_{N+1} f(a_{N+1}) \\ &\leq f((1 - w_{N+1})(t_1 a_1 + \cdots + t_N a_N) + w_{N+1} a_{N+1}) \\ &= f(w_1 a_1 + \cdots + w_N a_N + w_{N+1} a_{N+1}). \end{aligned}$$

ここで, 2行目で  $n = N$  の場合の(\*)を使い, 3行目で  $f$  が上に凸であることを使つた.

ゆえに数学的帰納法によってすべての  $n$  について (\*) が成立する.  $\square$

## 相加相乗平均の不等式

例:  $a_i > 0$  であるとし,

$$E[f(x)] = \frac{f(a_1) + \cdots + f(a_n)}{n}$$

とおく.  $x > 0$  に対して  $f(x) = \log x$  とおく. このとき,  $f(x) = \log x$  は下に凸なので, Jensenの不等式より  $E[f(x)] \leq f(E[x])$ . そして,

$$E[f(x)] = \frac{\log a_1 + \cdots + \log a_n}{n} = \log(a_1 \cdots a_n)^{1/n}, \quad f(E[x]) = \log \frac{a_1 + \cdots + a_n}{n}$$

なので,  $\log$  の単調増加性より,

$$(a_1 \cdots a_n)^{1/n} \leq \frac{a_1 + \cdots + a_n}{n}.$$

これでJensenの不等式が相加相乗平均の不等式を含んでいることがわかつた.  $\square$

例題: 下に凸な函数  $f(x) = e^x$  に関するJensenの不等式を用いて相加相乗平均の不等式を証明し直せ.

解答例:  $a_i > 0$  であるとし,  $A_i = \log a_i$ ,  $E[f(x)] = \frac{f(A_1) + \cdots + f(A_n)}{n}$  とおくと,  $f(x) = e^x$  のとき

$$\begin{aligned} E[f(x)] &= E[e^x] = \frac{e^{A_1} + \cdots + e^{A_n}}{n} = \frac{a_1 + \cdots + a_n}{n}, \\ f(E[x]) &= e^{(A_1 + \cdots + A_n)/n} = (e^{A_1} \cdots e^{A_n})^{1/n} = (a_1 \cdots a_n)^{1/n}. \end{aligned}$$

下に凸な函数  $f(x) = e^x$  に関するJensenの不等式  $E[f(x)] \geq f(E[x])$  より,

$$\frac{a_1 + \cdots + a_n}{n} \geq (a_1 \cdots a_n)^{1/n}.$$

これで示すべきことが示された.  $\square$

## Youngの不等式

Youngの不等式の一般化:  $p_i > 0$ ,  $\frac{1}{p_1} + \cdots + \frac{1}{p_n} = 1$ ,  $a_i > 0$  のとき

$$a_1 \cdots a_n \leq \frac{d_1^{p_1}}{p_1} + \cdots + \frac{d_n^{p_n}}{p_n}$$

$n = 2$  の場合をYoungの不等式と呼ぶ. すなわち  $p > 0$ ,  $q > 0$ ,  $\frac{1}{p} + \frac{1}{q} = 1$ ,  $a > 0$ ,  $b > 0$  のときの

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

という不等式をYoungの不等式と呼ぶ.

証明: 上に凸な函数  $f(x) = \log x$  に関するJensenの不等式より

$$\log(a_1 \cdots a_n) = \frac{\log a_1^{p_1}}{p_1} + \cdots + \frac{\log a_n^{p_n}}{p_n} \leq \log \left( \frac{d_1^{p_1}}{p_1} + \cdots + \frac{d_n^{p_n}}{p_n} \right)$$

ゆえに

$$a_1 \cdots a_n \leq \frac{d_1^{p_1}}{p_1} + \cdots + \frac{d_n^{p_n}}{p_n} \quad \square$$

**問題:**  $a, b > 0, p, q > 0, \frac{1}{p} + \frac{1}{q} = 1$  のとき

$$ab \leqq \frac{a^p}{p} + \frac{b^q}{q}$$

となること(つまりYoungの不等式)を直接的に証明してみよ.

**解答例1:**  $f(x) = \log x$  が上に凸であることより,  $A, B > 0$  について

$$\frac{1}{p} \log A + \frac{1}{q} \log B \leqq \log \left( \frac{A}{p} + \frac{B}{q} \right).$$

ゆえに  $A = a^p, B = b^q$  とおくと,

$$\log(ab) = \frac{1}{p} \log a^p + \frac{1}{q} \log b^q \leqq \log \left( \frac{a^p}{p} + \frac{b^q}{q} \right).$$

したがって,

$$ab \leqq \frac{a^p}{p} + \frac{b^q}{q}. \quad \square$$

**解答例2:**  $f(x) = e^x$  が下に凸であることより, 実数  $A, B$  について

$$\frac{1}{p} e^A + \frac{1}{q} e^B \geqq e^{A/p+B/q} = e^{A/p} e^{B/q}.$$

ゆえに  $A = \log a^p, B = \log b^q$  とおくと,

$$\frac{a^p}{p} + \frac{b^q}{q} \geqq ab. \quad \square$$

**解答例3:**  $f(x)$  は  $x \geqq 0$  の狭義単調増加連続函数で  $f(0) = 0$  を満たすものであるとし,

$$X = \{(x, y) \mid 0 \leqq x \leqq a, 0 \leqq y \leqq f(x)\}, \quad Y = \{(x, y) \mid 0 \leqq x \leqq f^{-1}(y), 0 \leqq y \leqq b\}$$

とおくと,  $[0, a] \times [0, b] \subset X \cup Y$  となるので(図を描いてみよ), 面積を比較することによって

$$ab \leqq \int_0^a f(x) dx + \int_0^b f^{-1}(y) dy. \quad (\star)$$

$f(x) = x^{p-1}$  とおくと,

$$\frac{1}{p} + \frac{1}{q} = 1 \iff (p-1)(q-1) = 1$$

より,  $f^{-1}(y) = y^{1/(p-1)} = y^{q-1}$  となることがわかる. これに上の不等式を適用すると

$$ab \leqq \frac{a^p}{p} + \frac{b^q}{q}$$

が得られる.  $\square$

**注意:**  $a^{p-1} = b$  ならば  $(\star)$  で等号が成立する.

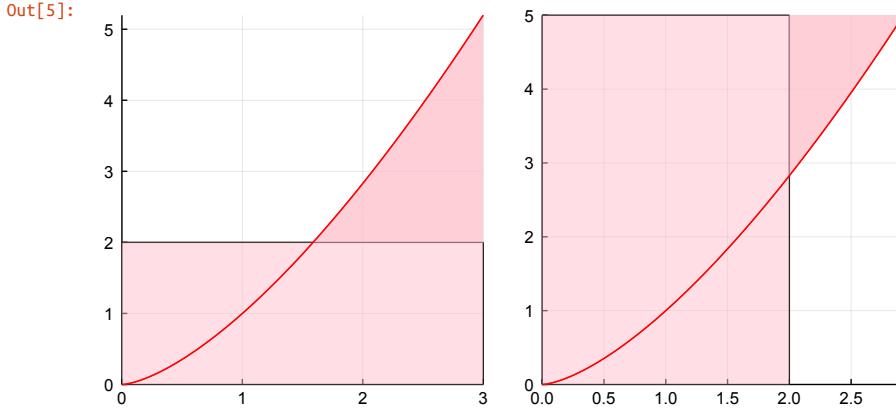
$$\frac{1}{p} + \frac{1}{q} = 1 \iff (p-1)(q-1) = 1 \iff (p-1)q = p$$

なので  $a^{p-1} = b$  と  $a^p = b^q$  は同値である(前者の両辺を  $q$  乗すると後者になる).  $\square$

```

In [5]: 1 ▼ function plot_Young(a, b; p=2.5, kwarg...)
2     q = p/(p-1)
3     f(x) = x^(p-1)
4     g(y) = y^(q-1)
5     A = max(a, g(b))
6     x = 0:a/200:a
7     y = 0:b/200:b
8     x1 = 0:A/200:A
9     P = plot(legend=false; kwarg...)
10    plot!([a,a], [0,b], color=:black)
11    plot!([0,a], [b,b], color=:black)
12    plot!(x, f.(x), color=:red, fill=(0, 0.5, :pink))
13    plot!(x1, f.(x1), color=:red, fill=(b, 0.5, :pink))
14    if A > a
15        x2 = a:(A-a)/200:A
16        plot!(x2, f.(x2), color=:red, fill=(b, 0.5, :pink))
17    end
18    P
19 end
20
21 P1 = plot_Young(3,2)
22 P2 = plot_Young(2,5)
23 plot(P1, P2, size=(600, 300))

```



## Hölderの不等式

### $L^p$ ノルム

区間  $[a, b]$  乗の連続函数  $f$  に対して  $\|f\|_p$  を

$$\|f\|_p = \left( \int_a^b |f(x)|^p dx \right)^{1/p}$$

と定める。これを函数  $f$  の  $L^p$  ノルムと呼ぶ。

**問題:** 定数  $\alpha$  について  $\|\alpha f\|_p = |\alpha| \|f\|_p$  が成立していることを示せ。

**解答例:**

$$\|\alpha f\|_p = \left( \int_a^b |\alpha f(x)|^p dx \right)^{1/p} = \left( |\alpha|^p \int_a^b |f(x)|^p dx \right)^{1/p} = |\alpha| \left( \int_a^b |f(x)|^p dx \right)^{1/p} = |\alpha| \|f\|_p. \quad \square$$

**注意:** 上の問題の結果より、特に  $g = \frac{f}{\|f\|_p}$  とおくと  $\|g\|_p = 1$  となることがわかる。□

### 一般化されたヘルダーの不等式

$p > 0, p_i > 0$  であるとし、

$$\frac{1}{p_1} + \dots + \frac{1}{p_n} = \frac{1}{p}$$

と仮定する。このとき、区間  $[a, b]$  上の連続函数  $f_i$  達について

$$\|f_1 \cdots f_n\|_p \leq \|f_1\|_{p_1} \cdots \|f_n\|_{p_n} \quad (*)$$

が成立する。これを**一般化されたヘルダーの不等式**と呼ぶことにする。 $n = 2, p = 1$  の場合はヘルダーの不等式と呼ばれる。すなわち、 $p > 0, q > 0$ かつ

$$\frac{1}{p} + \frac{1}{q} = 1$$

のときの

$$\|fg\|_1 \leq \|f\|_p \|g\|_q$$

という不等式はヘルダーの不等式と呼ばれる。さらにこれの  $p = q = 2$  の場合はCauchy-Schwarzの不等式の特別な場合である。

**一般化されたヘルダーの不等式の証明の概略:**  $g_i(x) = \frac{f_i(x)}{\|f_i\|_{p_i}}$  とおく。そのとき  $\|g_i\|_{p_i} = 1$  となる。

$\frac{1}{p_1} + \dots + \frac{1}{p_n} = \frac{1}{p}$  は  $\frac{p}{p_1} + \dots + \frac{p}{p_n} = 1$  と同値なので,  $q_i = p_i/p$  とおくと,

$$q_i > 0, \quad pq_i = p_i, \quad \frac{1}{q_1} + \dots + \frac{1}{q_n} = 1.$$

ゆえに  $a_i = |g_i(x)|^p$  にYoungの不等式の一般化を適用することによって,

$$|g_1(x) \cdots g_n(x)|^p = |g_1(x)|^p \cdots |g_n(x)|^p \leq \frac{|g_1(x)|^{pq_1}}{q_1} + \cdots + \frac{|g_n(x)|^{pq_n}}{q_n} = \frac{|g_1(x)|^{p_1}}{q_1} + \cdots + \frac{|g_n(x)|^{p_n}}{q_n}.$$

これを  $x$  で積分すると,

$$(\|g_1 \cdots g_n\|_p)^p \leq \frac{(\|g_1(x)\|_{p_1})^{p_1}}{q_1} + \cdots + \frac{(\|g_n(x)\|_{p_n})^{p_n}}{q_n} = \frac{1}{q_1} + \cdots + \frac{1}{q_n} = 1.$$

ゆえに,

$$\frac{\|f_1 \cdots f_n\|_p}{\|f_1\|_{p_1} \cdots \|f_n\|_{p_n}} = \|g_1 \cdots g_n\|_p \leq 1.$$

したがって,

$$\|f_1 \cdots f_n\|_p \leq \|f_1\|_{p_1} \cdots \|f_n\|_{p_n}. \quad \square$$

**問題:** (一般化されていない)Youngの不等式から(一般化されていない)ヘルダーの不等式を導け。

**解答例の概略:**  $p, q > 0, \frac{1}{p} + \frac{1}{q} = 1$  と仮定する。

$$\|fg\|_1 \leq \|f\|_p \|g\|_q$$

を証明したい。左辺を右辺で割ったものが 1 以下になることを示せばよい。

$$a(x) = \frac{|f(x)|}{\|f\|_p}, \quad b(x) = \frac{|g(x)|}{\|g\|_q}$$

とおくと,  $\|a\|_p = \|b\|_q = 1$  となる。 $a(x), b(x)$  にYoungの不等式を適用すると,

$$a(x)b(x) \leq \frac{a(x)^p}{p} + \frac{b(x)^q}{q}.$$

両辺を  $x$  について積分すると,

$$\begin{aligned} \int_a^b a(x)b(x) dx &= \|ab\|_1 = \frac{\|fg\|_1}{\|f\|_p \|g\|_q}, \\ \int_a^b \left( \frac{a(x)^p}{p} + \frac{b(x)^q}{q} \right) dx &= \frac{(\|a\|_p)^p}{p} + \frac{(\|b\|_q)^q}{q} = \frac{1}{p} + \frac{1}{q} = 1. \end{aligned}$$

ゆえに

$$\frac{\|fg\|_1}{\|f\|_p \|g\|_q} \leq 1.$$

これより, ヘルダーの不等式が成立することがわかる。□

### 分配函数の対数凸性

**問題:**  $a < b$  と仮定する。 $Z(\beta) = \int_a^b e^{-\beta f(x)+g(x)} dx$  の形の函数  $Z(\beta)$  について,  $\log Z(\beta)$  が  $\beta$  の下に凸な函数になることをヘルダーの不等式を使って示せ。

**解答例:**  $0 < t < 1$  であるとし,  $p = 1/(1-t)$ ,  $q = 1/t$  とおく。そのとき  $\frac{1}{p} + \frac{1}{q} = 1$  となる。ゆえにヘルダーの不等式より,

$$\begin{aligned}
Z((1-t)\alpha + t\beta) &= Z\left(\frac{\alpha}{p} + \frac{\beta}{q}\right) = \int_a^b e^{-(\alpha/p + \beta/q)f(x) + (1/p + 1/q)g(x)} dx = \int_a^b (e^{-\alpha f(x) + g(x)})^{1/p} (e^{-\beta f(x) + g(x)})^{1/q} dx \\
&\leq \left( \int_a^b e^{-\alpha f(x) + g(x)} dx \right)^{1/p} \left( \int_a^b e^{-\beta f(x) + g(x)} dx \right)^{1/q} \\
&= Z(\alpha)^{1/p} Z(\beta)^{1/q} = Z(\alpha)^{1-t} Z(\beta)^t.
\end{aligned}$$

ゆえに

$$\log Z((1-t)\alpha + t\beta) \leq (1-t) \log Z(\alpha) + t \log Z(\beta).$$

これは  $\log Z(\beta)$  が  $\beta$  に関する下に凸な函数であることを意味している.  $\square$

**注意:** 上の問題の  $Z(\beta)$  の形の函数を**分配函数**と呼ぶことがある(統計力学). 上の問題より, ヘルダーの不等式は分配函数の対数が下に凸な函数になることを意味していることがわかる.  $\square$

**例(ガンマ函数の対数凸性):**  $s > 0$  に対して  $\Gamma(s) = \int_0^\infty e^{-x} x^{s-1} dx = \int_0^\infty e^{(s-1)\log x - x} dx$  とおくと, その対数  $\log \Gamma(s)$  は  $s > 0$  の下に凸な函数になる.  $\square$

**注意:** 上の問題の記号のもとで, 函数  $h(x)$  に対する  $\langle h(x) \rangle = \langle h(y) \rangle$  を次のように定める:

$$\langle h(x) \rangle = \frac{1}{Z(\beta)} \int_a^b h(x) e^{-\beta f(x) + g(x)} dx, \quad \langle h(y) \rangle = \frac{1}{Z(\beta)} \int_a^b h(y) e^{-\beta f(y) + g(y)} dy.$$

さらに函数  $h(x, y)$  に対する  $\langle h(x, y) \rangle$  を次のように定める:

$$\langle h(x, y) \rangle = \frac{1}{Z(\beta)^2} \int_a^b \int_a^b h(x, y) e^{-\beta f(x) + g(x)} e^{-\beta f(y) + g(y)} dx dy.$$

このとき  $\langle f(x) \rangle^2 = \langle f(x)f(y) \rangle$  などが成立している.

もしも上の問題の  $Z(\beta)$  が  $\beta$  について2回微分可能でかつ微分と積分の順序交換を自由にできるならば

$$\begin{aligned}
-\frac{d}{d\beta} \log Z(\beta) &= \frac{\int_a^b e^{-\beta f(x) + g(x)} f(x) dx}{Z(\beta)} = \langle f(x) \rangle, \\
\frac{d^2}{d\beta^2} \log Z(\beta) &= \langle f(x)^2 \rangle - \langle f(x) \rangle^2 = \langle f(x)f(x) \rangle - \langle f(x)f(y) \rangle \\
&= \frac{1}{2} (\langle f(x)f(x) \rangle - 2\langle f(x)f(y) \rangle + \langle f(y)f(y) \rangle) = \frac{1}{2} \langle (f(x) + f(y))^2 \rangle \geq 0.
\end{aligned}$$

これからも,  $\log Z(\beta)$  が下に凸な函数であることがわかる.  $\square$

## Minkowskiの不等式

$L^1$  ノルムについては, 三角不等式  $|f(x) + g(x)| \leq |f(x)| + |g(x)|$  より,

$$\|f + g\|_1 = \int_a^b |f(x) + g(x)| dx \leq \int_a^b (|f(x)| + |g(x)|) dx = \|f\|_1 + \|g\|_1$$

が成立している. この  $L^1$  ノルムに関する三角不等式はそのまま  $L^p$  ノルム( $p > 1$ )に一般化される.

$p > 1$  であると仮定する.  $q = \frac{p}{p-1}$  とおくと,  $\frac{1}{p} + \frac{1}{q} = 1$ . このとき,

$$\|\varphi^{p-1}\|_q = \|\varphi^{p-1}\|_{p/(p-1)} = \left( \int_a^b |\varphi(x)|^p dx \right)^{(p-1)/p} = (\|\varphi\|_p)^{p-1}.$$

これを  $\varphi = f + g$  に適用することによって,

$$\begin{aligned}
(\|f + g\|_p)^p &= \int_a^b |f(x) + g(x)|^p dx = \int_a^b |f(x) + g(x)| |f(x) + g(x)|^{p-1} dx \\
&\leq \int_a^b |f(x)| |f(x) + g(x)|^{p-1} dx + \int_a^b |g(x)| |f(x) + g(x)|^{p-1} dx \\
&\leq \|f\|_p \|\varphi^{p-1}\|_q + \|g\|_p \|\varphi^{p-1}\|_q \\
&= \|f\|_p (\|f(x) + g(x)\|_p)^{p-1} + \|g\|_p (\|f(x) + g(x)\|_p)^{p-1} \\
&= (\|f\|_p + \|g\|_p) (\|f(x) + g(x)\|_p)^{p-1}
\end{aligned}$$

ここで, 2行目の不等式で三角不等式  $|f(x) + g(x)| \leq |f(x)| + |g(x)|$  を使い, 3行目の不等式でヘルダーの不等式を使った. したがって,

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

これを**Minkowskiの不等式**と呼ぶ.

In [ ]: 1

# 積分 ¶

黒木玄

2018-06-21

- Copyright 2018 Gen Kuroki
- License: MIT <https://opensource.org/licenses/MIT> (<https://opensource.org/licenses/MIT>)

このファイルは次の場所できれいに閲覧できる:

- <http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/09%20integration.ipynb>  
<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/09%20integration.ipynb>.

このファイルは [Julia Box](https://juliabox.com) (<https://juliabox.com>) で利用できる。

自分のパソコンにJulia言語 (<https://julialang.org/>) をインストールしたい場合には

- [WindowsへのJulia言語のインストール](http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f) (<http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f>)

を参照せよ。

論理的に完璧な説明をするつもりはない。細部のいい加減な部分は自分で訂正・修正せよ。

## Table of Contents

<a href="#">1 積分の基本性質</a>
<a href="#">  1.1 基本性質(A)線形性..(B)積分区間の連結..(C)定数の積分..(D)単調性</a>
<a href="#">  1.2 性質(E) 積分の絶対値の上からの評価</a>
<a href="#">  1.3 性質(F) 微分積分学の基本定理1</a>
<a href="#">  1.4 性質(G) 微分積分学の基本定理2</a>
<a href="#">2 部分積分と置換積分</a>
<a href="#">  2.1 置換積分</a>
<a href="#">  2.2 部分積分</a>
<a href="#">3 積分の近似</a>
<a href="#">  3.1 Riemann和</a>
<a href="#">  3.2 台形公式</a>
<a href="#">  3.3 Simpsonの公式</a>
<a href="#">4 広義積分</a>
<a href="#">  4.1 絶対収束する広義積分の例</a>
<a href="#">  4.2 条件収束する広義積分の例</a>
<a href="#">5 積分と極限の交換</a>
<a href="#">  5.1 積分と極限が交換できない場合</a>
<a href="#">  5.2 Dirichlet積分とその一般化</a>

```
In [1]: 1  using Plots
2 ▼ gr(); ENV["PLOTS_TEST"] = "true"
3 #clibrary(:colorcet)
4 clibrary(:misc)
5
6 function pngplot(P...; kwargs...)
7     sleep(0.1)
8     pngfile = tempname() * ".png"
9     savefig(plot(P...; kwargs...), pngfile)
10    showimg("image/png", pngfile)
11 end
12 pngplot(; kwargs...) = pngplot(plot!(; kwargs...))
13
14 showing(mime, fn) = open(fn) do f
15     base64 = base64encode(f)
16     display("text/html", """""")
17 end
18
19 using SymPy
20 ▼ #sympy[:init_printing](order="lex") # default
21 ▼ #sympy[:init_printing](order="rev-lex")
22
23 using SpecialFunctions
24 using QuadGK
```

## 積分の基本性質

1変数函数  $f(x)$  の積分  $\int_a^b f(x) dx$  の構成の仕方は無数に存在する.

Riemann積分とLebesgue積分は有名であり, それ以外にも様々な方法が存在する. 例えは,

- S. ラング, 『ラント 現代の解析学』, 共立出版, 1981 (原書: Serge Lang, Real Analysis, 1969)

の第5章「微分法」の第1節では,

$$f(t) = v_i \quad (a_{i-1} < t < a_i, a = a_0 \leqq a_1 \leqq \dots \leqq a_n = b)$$

を満たす函数(階段函数と呼ぶことにする)の積分を

$$\int_a^b f(t) dt = \sum_{i=1}^n v_i(a_i - a_{i-1})$$

と定め, これを階段函数の一様収束先に拡張することによって, 積分を導入している. 閉区間上の連続函数は階段函数の一様収束先になっているので, 連続函数の積分もこれによって定義されることになる.

積分の構成の仕方は無数にあるが, 積分は構成の仕方によらない基本性質を持つ. だから, その基本性質のみを使って示された結果はどの積分論においても利用可能になる.

このノートでは積分の構成の仕方には深く踏み込まずに, 積分の基本性質のみを使って積分論を展開する.

ただし, 積分の基本性質の厳密な説明を避けて, 大らかなスタイルで議論を進めることにする.

### 基本性質(A)線形性, (B)積分区間の連結, (C)定数の積分, (D)単調性

**基本性質(A) 線形性:** 定数  $\alpha, \beta$  について

$$\int_a^b (\alpha f(x) + \beta g(x)) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx.$$

**基本性質(B) 積分区間の連結:**

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

**注意:**

(1) (B)において  $a = b = c$  と仮定すると  $\int_a^a f(x) dx = 0$  が導かれる.

(2) (B)において  $c = a$  と仮定すると  $\int_b^a f(x) dx = -\int_a^b f(x) dx$  が導かれる.

(3) ゆえに  $a \leqq b$  のときに  $\int_a^b f(x) dx$  が定義されていれば,  $a > b$  の場合には  $\int_a^b f(x) dx$  を  $\int_a^b f(x) dx = -\int_b^a f(x) dx$  で定義することによって,  $a, b$  の大小に条件を付ける必要がなくなる.

(4) (B)は  $\int_a^c f(x) dx - \int_b^a f(x) dx = \int_b^c f(x) dx$  と同値である. 基本性質(B)はこの形でもよく使われる.

**基本性質(C) 定数の積分:** 定数  $\alpha$  について,

$$\int_a^b \alpha dx = \alpha(b - a).$$

**注意:**

(1)  $\int_a^b 1 dx = b - a$  と(A)から(C)が導かれる.

(2) 函数  $f(x)$  が  $a < x < b$  で一定の値  $\alpha$  の値を取るならば(すなわち  $f(x) = \alpha$  ( $a < x < b$ )) ならば  $\int_a^b f(x) dx = \alpha(b - a)$  となるという(C)の一般化を仮定すると, (B)と合わせて, 階段函数の積分が確定する.

**基本性質(D) 単調性:**  $a \leqq b$  と仮定する. 1変数実数值函数  $f(x), g(x)$  について

$$f(x) \leqq g(x) \quad (a < x < b) \quad \Rightarrow \quad \int_a^b f(x) dx \leqq \int_a^b g(x) dx.$$

### 性質(E) 積分の絶対値の上からの評価

**(E)**  $a \leqq b$  と仮定する. 1変数実数值函数  $f(x)$  について

$$\left| \int_a^b f(x) dx \right| \leqq \int_a^b |f(x)| dx. \tag{1}$$

さらに  $a \leqq x \leqq b$  で  $|f(x)| \leqq M$  ならば

$$\left| \int_a^b f(x) dx \right| \leqq M|b - a|. \tag{2}$$

**注意:** (2)は  $a > b$  であっても成立している. そのために絶対値を付けた.

**証明:**  $-|f(x)| \leq f(x) \leq |f(x)|$  なので積分の単調性(D)より

$$-\int_a^b |f(x)| dx \leq \int_a^b f(x) dx \leq \int_a^b |f(x)| dx.$$

これは(1)が成立することを意味する.  $a \leq x \leq b$  で  $|f(x)| \leq M$  ならば積分の単調性(D)と(C)より,

$$\int_a^b |f(x)| dx \leq \int_a^b M dx = M(b-a) = M|b-a|.$$

これで(2)も成立することが示された.  $\square$

**解説:** 多くの場面で「小さいこと」を示したい誤差項は積分で表わされる. (Taylorの定理における積分型剰余項はその典型例である.) 成立(E)は積分の絶対値が「小さいこと」を示すためによく使われることになる.  $\square$

**注意:**  $f(x)$  が複素数値函数であっても性質(E)は成立している. さらに絶対値を任意のノルムで置き換えればベクトル値函数であっても性質(E)は成立している. しかし、そのことを示すためには性質(A)~(D)を仮定するだけでは足りないので、複素数値函数やベクトル値函数まで拡張された性質(E)も必要ならば、最初から拡張された性質(E)も仮定しておく方がよいかもしれません.

拡張された性質(E)が主要などの積分論でも成立する理由は、どの積分論においても積分が有限和  $v_1 + \dots + v_n$  の極限で構成されているからである. ここで  $v_i$  は複素数またはベクトルである. 絶対値およびより一般にノルムは三角不等式

$$\|v_1 + \dots + v_n\| \leq \|v_1\| + \dots + \|v_n\|$$

を満たしている. これの極限として複素数値もしくはベクトル値函数  $f(x)$  についても

$$\left\| \int_a^b f(x) dx \right\| \leq \int_a^b \|f(x)\| dx$$

が成立することが示される. この不等式を点の運動  $x(t)$  の速度ベクトル  $v(t) = x'(t)$  に適用した結果は

$$\|x(b) - x(a)\| = \left\| \int_a^b v(t) dt \right\| \leq \int_a^b \|v(t)\| dt$$

となる. 速さ  $\|v(t)\| = \|x'(t)\|$  の時刻  $a$  から  $b$  までの積分の値は点の運動の軌跡(曲線)の長さなので、この不等式は  $x(a)$  と  $x(b)$  の距離は  $x(a)$  と  $x(b)$  を結ぶ曲線の長さ以下になることを意味する.

これより、ベクトル値函数に拡張された性質(E)は、直観的には当然そうなるべき結果を意味していることがわかる.

## 性質(F) 微分積分学の基本定理1

(F) 微分積分学の基本定理1: 連続函数  $f(x)$  について

$$\frac{d}{dx} \int_a^x f(\xi) d\xi = f(x).$$

**証明:**  $F(x) = \int_a^x f(\xi) d\xi$  とおく. このとき(B)より

$$F(x+h) - F(x) = \int_a^{x+h} f(\xi) d\xi - \int_a^x f(\xi) d\xi = \int_x^{x+h} f(\xi) d\xi.$$

さらに  $f(x)$  は積分変数  $\xi$  について定数なので、(C)より

$$f(x)h = f(x)((x+h)-x) = \int_x^{x+h} f(x) d\xi.$$

ゆえに

$$F(x+h) - F(x) - f(x)h = \int_x^{x+h} f(\xi) d\xi - \int_x^{x+h} f(x) d\xi = \int_x^{x+h} (f(\xi) - f(x)) d\xi.$$

任意の  $\epsilon > 0$  を取る.  $f(\xi)$  は  $\xi$  の連続函数なので  $|h|$  を十分に小さくすると

$$|f(\xi) - f(x)| \leq \epsilon$$

となる(図を描いてみよ). ゆえに(E)より

$$|F(x+h) - F(x) - f(x)h| = \left| \int_x^{x+h} (f(\xi) - f(x)) d\xi \right| \leq \epsilon|x+h-x| = \epsilon|h|.$$

両辺を  $|h|$  で割ることによって、 $\epsilon > 0$  ごとに  $|h| > 0$  を十分に小さくすることによって、

$$\left| \frac{F(x+h) - F(x) - f(x)h}{h} \right| \leq \epsilon$$

となることわかる. これは  $h \rightarrow 0$  のとき

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

が成立することを意味する。これで  $F'(x) = f(x)$  となることを示せた。□

**注意:** 上の証明では  $\varepsilon$ - $\delta$  論法を使った。 $\varepsilon$ - $\delta$  論法を未習もしくはまだ理解していない人であっても、 $f(x)$  のグラフを適当に描いて、積分が「面積」を表すことを思い出しながら、証明の各ステップに現われる積分がグラフのどの部分の「面積」になっているかを確認すれば、直観的には明らかな議論しか行っていないことがわかるだろう。むしろ  $\varepsilon$ - $\delta$  論法を使った証明を使いこなせるようになるためには、 $\varepsilon$ - $\delta$  論法を理解していない段階でのそのような直観的な議論を十分にしておいた方がよい。慣れてしまえば、 $\varepsilon$ - $\delta$  論法自体が直観的に明らかな議論に見えて来ることになるだろう。□

## 性質(G) 微分積分学の基本定理2

**準備:** 微分可能函数  $f(x)$  の導函数  $f'(x)$  が恒等的に 0 ならば  $f(x)$  は定数函数になることを(G)の証明では認めて使うことにする。その結果は平均値の定理

$$f(x) - f(a) = f'(\xi)(x - a) \quad (\xi = a + t(x - a), 0 < t < 1)$$

を使えば容易に導かれる( $\xi$  は  $a, x$  のあいだのある実数である)。 $f'$  が恒等的に 0 ならば特に  $f'(\xi) = 0$  となるので  $f(x) = f(a)$  となり、 $f(x)$  の値は  $x$  の値によらず  $f(a)$  に等しくなる。□

**(G) 微分積分学の基本定理2:**  $F(x)$  は  $C^1$  級函数(微分可能かつ導函数が連続な函数)であるとすると、

$$\int_a^b F'(x) dx = F(b) - F(a). \quad (1)$$

**注意:** (1)の右辺を  $[F(x)]_a^b$  と書くことが多い:

$$[F(x)]_a^b = F(b) - F(a).$$

**証明:**  $G(x) = \int_a^x F'(\xi) d\xi - (F(x) - F(a))$  とおく。このとき、(B)より

$$G(a) = \int_a^a F'(\xi) d\xi - (F(a) - F(a)) = 0$$

となり、(F)より

$$G'(x) = F'(x) - F'(x) = 0.$$

ゆえに  $G(x)$  は定数函数になり、恒等的に 0 になる。ゆえに

$$0 = G(b) = \int_a^b F'(x) dx - (F(b) - F(a)).$$

これは(1)が成立することを意味する。□

以上によって、実1変数実数値函数の微分積分学の基本定理が積分の基本性質(A)～(D)だけから導かれることがわかった。

**注意:** 実1変数の複素数値函数もしくはベクトル値函数に関する微分積分学の基本定理は積分の基本性質(A)～(D)と拡張された性質(E)のみを使って導き出せる。性質(A)～(E)は高校数学レベルの積分に関する直観があれば「明らか」だと感じられるシンプルな性質の集まりである。そのような「明らか」でかつシンプルな性質のみを使って微分積分学の基本定理は導出可能なのである。□

## 部分積分と置換積分

微分積分学の基本定理さえ証明できてしまえば、部分積分と置換積分の導出は微分学の結果をそのまま積分学に応用する話に過ぎなくなる。

### 置換積分

$f(x), g(t)$  は  $C^1$  級函数であると仮定し、

$$F(x) = \int_a^x f(\xi) d\xi$$

とおく。微分の性質

$$\frac{d}{dt} F(g(t)) = F'(g(t))g'(t) = f(g(t))g'(t)$$

に微分積分学の基本定理(G)を適用すると、 $a = g(\alpha), b = g(\beta)$  のとき、

$$\int_a^b f(x) dx = [F(x)]_a^b = [F(g(t))]_{\alpha}^{\beta} = \int_{\alpha}^{\beta} \frac{d}{dt} F(g(t)) dt = \int_{\alpha}^{\beta} f(g(t))g'(t) dt.$$

つまり、次が成立している。

**(H) 置換積分:**  $a = g(\alpha), b = g(\beta)$  のとき、

$$\int_a^b f(x) dx = \int_{\alpha}^{\beta} f(g(t))g'(t) dt.$$

左辺の積分を右辺で計算することを**置換積分**と呼ぶ。 $x(t) = g(t)$  とおいて、

$$\int_a^b f(x) dx = \int_a^b f(x(t)) \frac{dx(t)}{dt} dt$$

と書き直すと、「分子分母の  $dt$  が約分で消える公式」のように見えて見易い。

**問題:** 任意の微積分の教科書を参照して、置換積分を使って計算する積分計算の面白い問題を見付けて、計算練習せよ。□

## 部分積分

$f(x), g(x)$  は  $C^1$  級函数であると仮定する。

微分の性質

$$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$$

に微分積分学の基本定理(G)と積分の線形性(A)を適用すると

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

これは以下と同値である。

(I) **部分積分:**  $f(x), g(x)$  は  $C^1$  級函数であるとき、

$$\begin{aligned} \int_a^b f'(x)g(x) dx &= [f(x)g(x)]_a^b - \int_a^b f(x)g'(x) dx, \\ \int_a^b f(x)g'(x) dx &= [f(x)g(x)]_a^b - \int_a^b f'(x)g(x) dx. \end{aligned}$$

左辺の積分を右辺で計算することを**部分積分**と呼ぶ。

**問題:** 任意の微積分の教科書を参照して、部分積分を使って計算する積分計算の面白い問題を見付けて、計算練習せよ。□

## 積分の近似

$a < b$  であると仮定し、 $f$  は閉区間  $[a, b]$  上の連続函数であると仮定する。

### Riemann和

閉区間  $[a, b]$  の分割  $a = x_0 \leq x_1 \leq \dots \leq x_N = b$  と  $x_i^* \in [x_{i-1}, x_i]$  と  $\Delta x_i = x_i - x_{i-1}$  に対して、和

$$\sum_{i=1}^N f(x_i^*) \Delta x_i = \sum_{i=1}^N f(x_i^*)(x_i - x_{i-1})$$

を**Riemann和**と呼ぶ。

$\Delta x_i$  の最大値を小さくすることによって(分割を細かくすることによって)、Riemann和で積分  $\int_a^b f(x) dx$  を幾らでも近似できることを積分の基本性質のみを使って証明しよう。

**証明:** 任意の  $\varepsilon > 0$  を取る。 $f(x)$  は閉区間  $[a, b]$  上の連続函数であると仮定してあったので、 $f(x)$  は閉区間  $[a, b]$  上の一様連続函数になる。 $\Delta x_i$  達の最大値を十分小さくすれば

$$|f(x) - f(x_i^*)| \leq \frac{\varepsilon}{b-a} \quad (x_{i-1} \leq x \leq x_i)$$

となる。したがって、

$$\begin{aligned} \left| \sum_{i=1}^N f(x_i^*) \Delta x_i - \int_a^b f(x) dx \right| &= \left| \sum_{i=1}^N \int_{x_{i-1}}^{x_i} f(x_i^*) dx - \sum_{i=1}^N \int_{x_{i-1}}^{x_i} f(x) dx \right| \\ &= \left| \sum_{i=1}^N \int_{x_{i-1}}^{x_i} (f(x_i^*) - f(x)) dx \right| \\ &\leq \sum_{i=1}^N \left| \int_{x_{i-1}}^{x_i} (f(x_i^*) - f(x)) dx \right| \\ &\leq \sum_{i=1}^N \int_{x_{i-1}}^{x_i} |f(x_i^*) - f(x)| dx \\ &\leq \sum_{i=1}^N \int_{x_{i-1}}^{x_i} \frac{\varepsilon}{b-a} dx = \sum_{i=1}^N \frac{\varepsilon}{b-a} \Delta x_i \\ &= \frac{\varepsilon}{b-a} (b-a) = \varepsilon \end{aligned}$$

以上によって、閉区間  $[a, b]$  の分割を細かくすることによって、Riemann和で積分を誤差  $\varepsilon$  以下で近似できることを示せた。□

## 台形公式

閉区間  $[a, b]$  の分割  $a = x_0 \leqq x_1 \leqq \dots \leqq x_N = b$  に対する和

$$\sum_{i=1}^N \frac{f(x_{i-1}) + f(x_i)}{2} \Delta x_i$$

で積分  $\int_a^b f(x) dx$  を近似することもできる。これを**台形公式**(trapezoidal rule)と呼ぶ。

## Simpsonの公式

閉区間  $[a, b]$  の分割  $a = x_0 \leqq x_1 \leqq \dots \leqq x_N = b$  に対する和

$$\sum_{i=1}^N \frac{f(x_{i-1}) + 4f((x_{i-1} + x_i)/2) + f(x_i)}{6} \Delta x_i$$

で積分  $\int_a^b f(x) dx$  を近似することもできる。これを**Simpsonの公式**(Simpson's rule)と呼ぶ。

**Simpsonの公式の由来:**  $a < b$  であるとし、 $c = (a + b)/2$  とおく。このとき、 $f(a) = \alpha, f(c) = \gamma, f(b) = \beta$  を満たす2次以下の多項式  $f(x)$  が一意的に決まり、

$$f(x) = \alpha \frac{(x - b)(x - c)}{(a - b)(a - c)} + \beta \frac{(x - a)(x - c)}{(b - a)(b - c)} + \gamma \frac{(x - a)(x - b)}{(c - a)(c - b)}$$

と一意に表わされる。このとき、

$$\begin{aligned} \int_a^b \frac{(x - b)(x - c)}{(a - b)(a - c)} dx &= \frac{1}{6}(b - a), \\ \int_a^b \frac{(x - b)(x - c)}{(a - b)(a - c)} dx &= \frac{1}{6}(b - a), \\ \int_a^b \frac{(x - b)(x - c)}{(a - b)(a - c)} dx &= \frac{4}{6}(b - a) \end{aligned}$$

となることを確認できるので

$$\int_a^b f(x) dx = \frac{f(a) + 4f((a+b)/2) + f(b)}{6} (b - a)$$

となることがわかる。

つまり、 $f(x)$  が2次函数のとき閉区間  $[a, b]$  の自明な分割  $a = a_0 < a_1 = b$  に関するSimpsonの公式は誤差無しで積分の値を計算する公式になっている。□

Simpsonの公式の他の解釈については

- 黒木玄、[Euler-Maclaurinの和公式の一般化](https://genkuroki.github.io/documents/20170724EulerMaclaurin.pdf) (<https://genkuroki.github.io/documents/20170724EulerMaclaurin.pdf>)

の第3.5と第3.6節を参照せよ。

**問題:** 上のSimpsonの公式の由来の解説において結果だけを書いた積分計算を自分で実行せよ。□

解答略。

```
In [2]: 1 a, b, x = symbols("a b x", real=True)
2 c = (a+b)/2
3 [
4     factor(simplify(integrate((x-b)*(x-c), (x, a, b)))/((a-b)*(a-c))),
5     factor(simplify(integrate((x-a)*(x-c), (x, a, b)))/((b-a)*(b-c))),
6     factor(simplify(integrate((x-a)*(x-b), (x, a, b)))/((c-a)*(c-b))),
7 ]
```

```
Out[2]: ⎡ -1/6(a - b) ⎤
⎢ -1/6(a - b) ⎥
⎢ -2/3(a - b) ⎥
```

**問題:** Simpsonの公式の由来の説明に基いて、以下の公式を導け。

$$\int_a^b (x - a)(x - b) dx = -\frac{(b - a)^3}{6}, \quad \int_a^b (x - a) \left( x - \frac{a+b}{2} \right) dx = \frac{(b - a)^3}{12}. \quad \square$$

解答略。

```
In [3]: 1 a, b, x = symbols("a b x", real=true)
2 c = (a+b)/2
3
4 sol = [
5     integrate((x-a)*(x-b), (x,a,b)),
6     integrate((x-a)*(x-c), (x,a,b)),
7 ]
8 factor(simplify(sol))
```

Out[3]: 
$$\left[ \frac{1}{6}(a-b)^3, -\frac{1}{12}(a-b)^3 \right]$$

問題: 円周率を表わす2種類の積分

$$4 \int_0^1 \sqrt{1-x^2} dx = \pi, \quad 4 \int_0^1 \frac{dx}{1+x^2} = \pi$$

の近似計算を、閉区間  $[0, 1]$  の分割  $x_i = i/N$  ( $i = 0, 1, \dots, N$ ) と  $x_i^* = (x_{i-1} + x_i)/2$  に関するRiemann和と台形公式とSimpsonの公式を用いて実行してみよ。 $N = 10$  の場合にどうなるか? □

```
In [4]: 1 function RiemannSum(f, a, b, N)
2     Δx = (b-a)/N
3     sum(i->f(a+(i-1/2)*Δx)*Δx, 1:N)
4 end
5
6 function TrapezoidalRule(f, a, b, N)
7     Δx = (b-a)/N
8     sum(i->(f(a+(i-1)*Δx)+f(a+i*Δx))/2*Δx, 1:N)
9 end
10
11 function SimpsonRule(f, a, b, N)
12     Δx = (b-a)/N
13     sum(i->(f(a+(i-1)*Δx)+4f(a+(i-1/2)*Δx)+f(a+i*Δx))/6*Δx, 1:N)
14 end
15
16 function SymPyIntegrate(f, a, b)
17     x = symbols("x", real=true)
18     integrate(f(x), (x, a, b))
19 end
```

Out[4]: SymPyIntegrate (generic function with 1 method)

```
In [5]: 1 f(x) = 4x + 3
2 a, b = 0.0, 1.0
3 N = 10
4
5 @show RiemannSum(f, a, b, N)
6 @show TrapezoidalRule(f, a, b, N)
7 @show SimpsonRule(f, a, b, N)
8 @show s = SymPyIntegrate(f, Sym(0), Sym(1))
9 float(s)
```

RiemannSum(f, a, b, N) = 5.0  
TrapezoidalRule(f, a, b, N) = 5.0  
SimpsonRule(f, a, b, N) = 5.0  
s = SymPyIntegrate(f, Sym(0), Sym(1)) = 5

Out[5]: 5.0

```
In [6]: 1 f(x) = 2x^2 - x + 3
2 a, b = 0.0, 1.0
3 N = 10
4
5 @show RiemannSum(f, a, b, N)
6 @show TrapezoidalRule(f, a, b, N)
7 @show SimpsonRule(f, a, b, N)
8 @show s = SymPyIntegrate(f, Sym(0), Sym(1))
9 float(s)
```

RiemannSum(f, a, b, N) = 3.1650000000000005  
TrapezoidalRule(f, a, b, N) = 3.1700000000000001  
SimpsonRule(f, a, b, N) = 3.1666666666666667  
s = SymPyIntegrate(f, Sym(0), Sym(1)) = 19/6

Out[6]: 3.1666666666666665

```
In [7]: 1 f(x) = 4*sqrt(1-x^2)
2 a, b = 0.0, 1.0
3 N = 10
4
5 @show RiemannSum(f, a, b, N)
6 @show TrapezoidalRule(f, a, b, N)
7 @show SimpsonRule(f, a, b, N)
8 @show s = SymPyIntegrate(f, Sym(0), Sym(1))
9 float(s)
```

```
RiemannSum(f, a, b, N) = 3.1524114332616446
TrapezoidalRule(f, a, b, N) = 3.1045183262483182
SimpsonRule(f, a, b, N) = 3.136447064257202
s = SymPyIntegrate(f, Sym(0), Sym(1)) = pi
```

Out[7]: 3.141592653589793

```
In [8]: 1 f(x) = 4/(1+x^2)
2 a, b = 0.0, 1.0
3 N = 10
4
5 @show RiemannSum(f, a, b, N)
6 @show TrapezoidalRule(f, a, b, N)
7 @show SimpsonRule(f, a, b, N)
8 @show s = SymPyIntegrate(f, Sym(0), Sym(1))
9 float(s)
```

```
RiemannSum(f, a, b, N) = 3.142425985001098
TrapezoidalRule(f, a, b, N) = 3.1399259889071587
SimpsonRule(f, a, b, N) = 3.1415926529697855
s = SymPyIntegrate(f, Sym(0), Sym(1)) = pi
```

Out[8]: 3.141592653589793

```
In [9]: 1 SimpsonRule(f, a, b, N) - pi
```

Out[9]: -6.200076008155975e-10

$4 \int_0^1 \frac{dx}{1+x^2} = \pi$  の左辺を10分割のSimpsonの公式で計算すると、小数点以下8桁まで円周率を正確に求めることができている！たったの10分割しかしていないのに！

**注意:** 数値積分は非常に深い分野である。多くの功名な方法が開発されており、多くの場合に以上で説明した方法よりもずっと効率的である。□

## 広義積分

無限区間もしくは区間  $(a, b)$  の端で発散しているかもしれない函数  $f(x)$  について極限

$$\lim_{\alpha \searrow a, \beta \nearrow b} \int_{\alpha}^{\beta} f(x) dx$$

を**広義積分**と呼ぶ。極限が存在しなければ広義積分は存在しない。広義積分は

$$\lim_{\alpha \searrow a, \beta \nearrow b} \int_{\alpha}^{\beta} |f(x)| dx \quad (*)$$

が有限の値に収束するとき収束し、そのときその広義積分は**絶対収束**するという。もとの広義積分は収束するが、(\*)が有限の値に収束しないとき、その広義積分は**条件収束**するという。

応用上重要な積分のかなりの割合が無限区間  $(-\infty, \infty)$ ,  $(0, \infty)$  上の積分になり、有限区間  $(a, b)$  上の積分であっても広義積分になっていることが多い。

例えば、**Gauss積分**や**Fourier変換**は無限区間  $(-\infty, \infty)$  上の積分で定義され、**ガンマ函数**は半無限区間  $(0, \infty)$  上の積分で定義され、**ベータ函数**は有限開区間  $(0, 1)$  上の広義積分で定義される。

**注意:** 広義積分を表すときに毎回  $\lim$  記号を書くのは面倒なので、通常の積分と同じ記号法で

$$\int_a^b f(x) dx = \lim_{\alpha \searrow a, \beta \nearrow b} \int_{\alpha}^{\beta} f(x) dx$$

と書くことが多い。さらに  $f(x) = F'(x)$  のとき、

$$\lim_{\alpha \searrow a, \beta \nearrow b} \int_{\alpha}^{\beta} f(x) dx = \lim_{\alpha \searrow a, \beta \nearrow b} [F(x)]_{\alpha}^{\beta} = \lim_{\beta \nearrow b} F(\beta) - \lim_{\alpha \searrow a} F(\alpha).$$

であるが、この右辺を記号の簡単のため

$$[F(x)]_a^b = \lim_{\beta \nearrow b} F(\beta) - \lim_{\alpha \searrow a} F(\alpha)$$

と書くことにする.  $\square$

**問題:** 積分は積分の基本性質を満たしていると仮定する.  $a < \alpha \leq \beta < b$  のとき積分  $\int_a^\beta f(x) dx$  が定義されており,

$$\lim_{\alpha \searrow a, \beta \nearrow b} \int_a^\beta |f(x)| dx$$

が有限の値に収束するならば,

$$\lim_{\alpha \searrow a, \beta \nearrow b} \int_a^\beta f(x) dx$$

も収束することを示せ.

**解答例:**  $a$  に収束する単調減少数列  $a_n$  と  $b$  に収束する単調増加数列  $b_n$  で  $a_n \leq b_n$  を満たすものを任意に取る. このとき仮定より

$$\lim_{n \rightarrow \infty} \int_{a_n}^{b_n} |f(x)| dx$$

は有限の値に収束する. この条件のもとで

$$\lim_{n \rightarrow \infty} \int_{a_n}^{b_n} f(x) dx$$

が収束することを示せばよい.  $A_n, B_n$  を

$$\begin{aligned} A_1 &= \int_{a_1}^{b_1} |f(x)| dx, & A_n &= \int_{a_n}^{a_{n-1}} |f(x)| dx + \int_{b_{n-1}}^{b_n} |f(x)| dx & (n \geq 2) \\ B_1 &= \int_{a_1}^{b_1} f(x) dx, & B_n &= \int_{a_n}^{a_{n-1}} f(x) dx + \int_{b_{n-1}}^{b_n} f(x) dx & (n \geq 2) \end{aligned}$$

と定める. このとき,

$$|B_k| \leq A_k, \quad \int_{a_n}^{b_n} |f(x)| dx = \sum_{k=1}^n A_k, \quad \int_{a_n}^{b_n} f(x) dx = \sum_{k=1}^n B_k.$$

そして,

$$\sum_{k=1}^{\infty} |B_k| \leq \sum_{k=1}^{\infty} A_k = \lim_{n \rightarrow \infty} \int_{a_n}^{b_n} |f(x)| dx < \infty$$

なので、「級数の各項の絶対値を足し上げた結果が有限の値になればその級数は収束する」(所謂絶対収束)という結果より, 級数  $\sum_{k=1}^{\infty} B_k$  は収束する. それは  $\int_{a_n}^{b_n} f(x) dx$  が  $n \rightarrow \infty$  で収束することを意味する.  $\square$

**注意:** 部分積分や置換積分は極限操作に注意を払えばそのまま広義積分にも適用可能である.  $\square$

**注意:** 部分積分

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

は, 特に  $[f(x)g(x)]_a^b = 0$  すなわち

$$\lim_{\alpha \searrow a} f(\alpha)g(\alpha) = \lim_{\beta \nearrow b} f(\beta)g(\beta)$$

のときに便利である. そのとき, 部分積分の公式は

$$\int_a^b f'(x)g(x) dx = - \int_a^b f(x)g'(x) dx$$

となる. これは  $f$  の微分を  $g$  に移動すると  $-1$  倍になることを意味している. 部分積分はこの形で使われることが多い.

特に,  $[f(x)g(x)]_a^b = 0$  の十分条件として,

$$\lim_{\alpha \searrow a} f(\alpha)g(\alpha) = 0, \quad \lim_{\beta \nearrow b} f(\beta)g(\beta) = 0$$

という条件は非常によく出て来る.  $\square$

**問題(Fourier変換):** 函数  $f(x)$  は  $C^1$  級であり,  $x \rightarrow \pm\infty$  で  $|f(x)|$  と  $|f'(x)|$  は十分な速さで 0 に近付くと仮定し,

$$\hat{f}(p) = \int_{-\infty}^{\infty} e^{-ipx} f(x) dx$$

とおく.  $\hat{f}(p)$  は  $f(x)$  のFourier変換と呼ばれる. このとき

$$(xf)^{\wedge}(p) = -i \frac{\partial}{\partial p} \hat{f}(p), \quad \left( -i \frac{\partial}{\partial x} f \right)^{\wedge}(p) = p \hat{f}(p)$$

が成立することを示せ.

**解答例:** 前者は次のように示される:

$$-i \frac{\partial}{\partial p} \hat{f}(p) = -i \int_{-\infty}^{\infty} (-ix) e^{-ipx} f(x) dx = \int_{-\infty}^{\infty} e^{-ipx} xf(x) dx = (xf)^{\wedge}(p).$$

後者を示そう. この場合にはすぐ上の注意のタイプの部分積分を利用できるので,

$$\left( -i \frac{\partial}{\partial x} f \right)^{\wedge}(p) = -i \int_{-\infty}^{\infty} e^{-ipx} \frac{\partial}{\partial x} f(x) dx = i \int_{-\infty}^{\infty} (-ip) e^{-ipx} f(x) dx = p \hat{f}(p).$$

2つ目の等号で部分積分を使った.  $\square$

このようにFourier変換には  $x$  をかける操作と微分する操作を交換する性質を持っている.

### 絶対収束する広義積分の例

以下は絶対収束する広義積分の例である.

(1)  $\arcsin y = \int_0^y \frac{dt}{\sqrt{1-t^2}}$  なので  $\int_{-1}^1 \frac{dt}{\sqrt{1-t^2}} = \pi$ . 積分区間  $(-1, 1)$  の両端で被積分函数  $\frac{1}{\sqrt{1-t^2}}$  は発散している.

(2)  $\arctan a = \int_0^a \frac{dt}{1+t^2}$  なので  $\int_{-\infty}^{\infty} \frac{dt}{1+t^2} = \pi$ . これは無限区間の積分になっている.

以上の2つの  $\pi$  は単位半円周の長さを意味する.

(3)  $s > 0$  に対するガンマ函数の定義  $\Gamma(s) = \int_0^{\infty} e^{-x} x^{s-1} dx$  は無限区間の積分になっており,  $0 < s < 1$  のとき  $x \searrow 0$  で被積分函数は発散している.

$x \rightarrow \infty$  で  $e^{-x} x^{s-1}$  は急速に 0 に近付くので  $\int_1^{\infty} e^{-x} x^{s-1} dx$  はどんな  $s$  の値に対しても収束している.

$s \neq 0$  のとき  $\int x^{s-1} dx = \frac{x^s}{s}$  なので  $s > 0$  ならば  $s < 1$  であっても  $\int_0^1 x^{s-1} ds = \frac{1}{s}$  (有限の値)となる. これより  $s > 0$  のとき  $\int_0^1 e^{-x} x^{s-1} dx$  も収束することがわかる. ( $x > 0$  で  $e^{-x} < 1$  となることに注意せよ.)

**問題:** 任意の実数  $s$  に対して, 積分  $\int_1^{\infty} e^{-x} x^{s-1} dx$  が有限の値に収束することを示せ.

**解答例:** 指数函数は多項式函数よりも速く増大するので,  $R > 1$  を十分大きくすると,  $x \geq R$  で  $e^{-x/2} x^{s-1} \leq 1$  となる. ゆえに

$$e^{-x} x^{s-1} \leq e^{-x/2} \quad (x \geq R).$$

ゆえに

$$\begin{aligned} \int_1^{\infty} e^{-x} x^{s-1} dx &= \int_1^R e^{-x} x^{s-1} dx + \int_R^{\infty} e^{-x} x^{s-1} dx \\ &\leq \int_1^R e^{-x} x^{s-1} dx + \int_R^{\infty} e^{-x/2} dx \\ &= \int_1^R e^{-x} x^{s-1} dx + 2e^{-R/2} < \infty. \end{aligned} \quad \square$$

**注意:** 指数函数部分  $e^{-x}$  を  $e^{-x/2}$  と  $e^{-x/2}$  に分けてその片方をもう1つの因子と比較するというような方法はよく使われる.  $\square$

(4)  $p > 0, q > 0$  に対するベータ函数の定義  $B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx$  の被積分函数は,  $0 < p < 1$  ならば積分区間の左端で,  $0 < q < 1$  ならば積分区間の右端で発散している.

しかし,  $p > 0$  ならば  $p < 1$  であっても  $\int_0^1 x^{p-1} ds = \frac{1}{p}$  (有限の値)となることから左端でも積分は収束する. 同様にして,  $q > 0$  ならば右端でも積分は収束する.

(5)  $a > 0$  と仮定する.  $f(x) = e^{-x^2/a}$  のFourier変換

$$\hat{f}(p) = \int_{-\infty}^{\infty} e^{-ipx} f(x) dx = \int_{-\infty}^{\infty} e^{-x^2/a} \cos(px) dx$$

の積分区間は無限区間である。 $x \rightarrow \pm\infty$  で  $e^{-x^2/a}$  が急速に 0 に近付くのでこの積分は絶対収束する。

**問題(三角函数のLaplace変換):**  $s > 0$  と仮定する. 次の公式を示せ:

$$\int_0^\infty e^{-st} \cos(pt) dt = \frac{s}{s^2 + p^2}, \quad \int_0^\infty e^{-st} \sin(pt) dt = \frac{p}{s^2 + p^2}.$$

**解答例1:**  $f(s) = \int_0^\infty e^{-st} \cos(pt) dt$  とおく.  $\frac{\partial}{\partial t} \left( \frac{e^{-st}}{-s} \right) = e^{-st}$  を用了部分積分を2回行う.

$$\begin{aligned}
f(s) &= \int_0^\infty e^{-st} \cos(pt) dt = \left[ \frac{e^{-st}}{-s} \cos(pt) \right]_0^\infty - \int_0^\infty \frac{e^{-st}}{-s} (-p \sin(pt)) dt \\
&= \frac{1}{s} - \frac{p}{s} \int_0^\infty e^{-st} \sin(pt) dt = \frac{1}{s} - \frac{p}{s} \left( \left[ \frac{e^{-st}}{-s} \sin(pt) \right]_0^\infty - \int_0^\infty \frac{e^{-st}}{-s} p \cos(pt) dt \right) \\
&= \frac{1}{s} - \frac{p^2}{s^2} f(s).
\end{aligned}$$

これを  $f(s)$  について解けば前者の公式が得られる.

$$g(s) = \int_0^\infty e^{-st} \sin(pt) dt \text{ とおく. 上と同様にして,}$$

$$\begin{aligned} g(s) &= \int_0^\infty e^{-st} \sin(pt) dt = \left[ \frac{e^{-st}}{-s} \sin(pt) \right]_0^\infty - \int_0^\infty \frac{e^{-st}}{-s} p \cos(pt) dt \\ &= \frac{p}{s} \int_0^\infty e^{-st} \cos(pt) dt = \frac{p}{s} \left( \left[ \frac{e^{-st}}{-s} \cos(pt) \right]_0^\infty - \int_0^\infty \frac{e^{-st}}{-s} (-p \sin(pt)) dt \right) \\ &= \frac{p}{s} \left( \frac{1}{s} - \frac{p}{s} g(s) \right). \end{aligned}$$

これを  $g(s)$  について解けば後者の公式が得られる.  $\square$

**解答例2:**  $e^{ipx} = \cos(px) + i \sin(px)$  を用いて一挙に計算する.

$$\int_0^\infty e^{-st} e^{ipx} dx = \int_0^\infty e^{-(s-ip)t} dx = \left[ \frac{e^{-(s-ip)t}}{-(s-ip)} \right]_0^\infty = \frac{1}{s-ip} = \frac{s+ip}{s^2+p^2}.$$

この結果の実部と虚部のそれぞれから示したい公式の前者と後者が得られる.  $\square$

部分積分を2回繰り返すよりも計算が圧倒的に平易になっている。

```
In [10]: 1 t, s, p = symbols("t s p", positive=true)
2 [
3     simplify(integrate(exp(-s*t)*cos(p*t), (t,0,oo))),
4     simplify(integrate(exp(-s*t)*sin(p*t), (t,0,oo)))
5 ]
```

Out[10]:  $\begin{bmatrix} \frac{s}{p^2+s^2} \\ \frac{p}{p^2+s^2} \end{bmatrix}$

### 条件収束する広義積分の例

以下は条件収束する広義積分に関する有名な公式である。

(1) Dirichlet積分(ディリクレ積分):  $\int_0^\infty \frac{\sin x}{x} dx = \frac{\pi}{2}$ .

## (2) Dirichlet積分の一般化:

$$\int_0^\infty \frac{e^{-ax} \sin x}{x} dx = \int_a^\infty \frac{dt}{1+t^2} = \frac{\pi}{2} - \arctan a \quad (a \geq 0).$$

これはDirichlet積分になる  $q = 0$  の場合にのみ条件収束する。

### (3) Fresnel積分の特殊値(フレネル積分の特殊値):

$$\int_{-\infty}^{\infty} \cos(t^2) dt = \int_{-\infty}^{\infty} \sin(t^2) dt = \sqrt{\frac{\pi}{2}}.$$

(4) 上の続き. (3)における被積分函数は  $t$  の偶函数なので  $(0, \infty)$  上の積分に書き直して,  $t = \sqrt{x}$  と置換すると

$$\int_0^\infty \frac{\cos x}{\sqrt{x}} dx = \int_0^\infty \frac{\sin x}{\sqrt{x}} dx = \sqrt{\frac{\pi}{2}}.$$

これらの公式の証明については

- 黒木玄, ガンマ分布の中心極限定理とStirlingの公式 (<https://genkuroki.github.io/documents/20160501StirlingFormula.pdf>)

を参照せよ. (1), (2)の証明は第5.5節とを参照せよ. (1), (3), (4)の証明は第8.6節を参照せよ.

**Fresnel積分:** 次の不定積分をFresnel積分と呼ぶ:

$$C(t) = \int_0^t \cos(s^2) ds, \quad S(t) = \int_0^t \sin(s^2) ds.$$

上で紹介したように、条件収束する広義積分の例として、

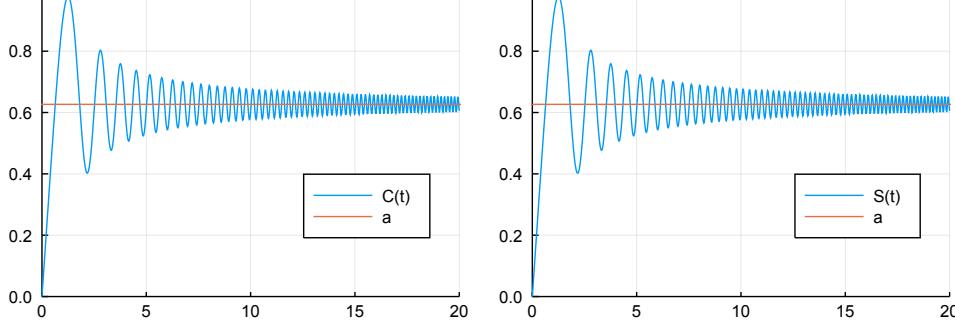
$$\lim_{t \rightarrow \infty} C(t) = \lim_{t \rightarrow \infty} S(t) = \frac{1}{2} \sqrt{\frac{\pi}{2}}$$

は有名である.  $C'(t)^2 + S'(t)^2 = 1$  なので点  $(C(t), S(t))$  の運動の速さは常に 1 である. だから時刻 0 から  $t$  までの軌跡の長さは  $t$  に等しい. だから軌跡の長さは  $t \rightarrow \infty$  で無限大に発散する. 無限に長い軌跡の最終到着地点が  $a = (1/2)\sqrt{\pi/2}$  に対する点  $(a, a)$  になる.

Fresnel積分  $C(t), S(t)$  のそれぞれをプロットすると以下のようにになる. それらが  $t \rightarrow \infty$  で  $a = (1/2)\sqrt{\pi/2}$  に収束する様子も確認できる.

```
In [11]: 1 ▼ C(t) = quadgk(s->cos(s^2), 0, t)[1]
2 ▼ S(t) = quadgk(s->sin(s^2), 0, t)[1]
3 a = sqrt(pi/2)/2
4
5 PP = []
6
7 tmin, tmax = 0, 20
8 t = tmin:(tmax-tmin)/1000:tmax
9 ▼ PC = plot(t, C.(t), label="C(t)"); hline!([a], label="a")
10 ▼ PS = plot(t, S.(t), label="S(t)"); hline!([a], label="a")
11 push!(PP, PC, PS)
12
13 plot(PP..., size=(700, 250), legend=:bottomright)
```

Out[11]:



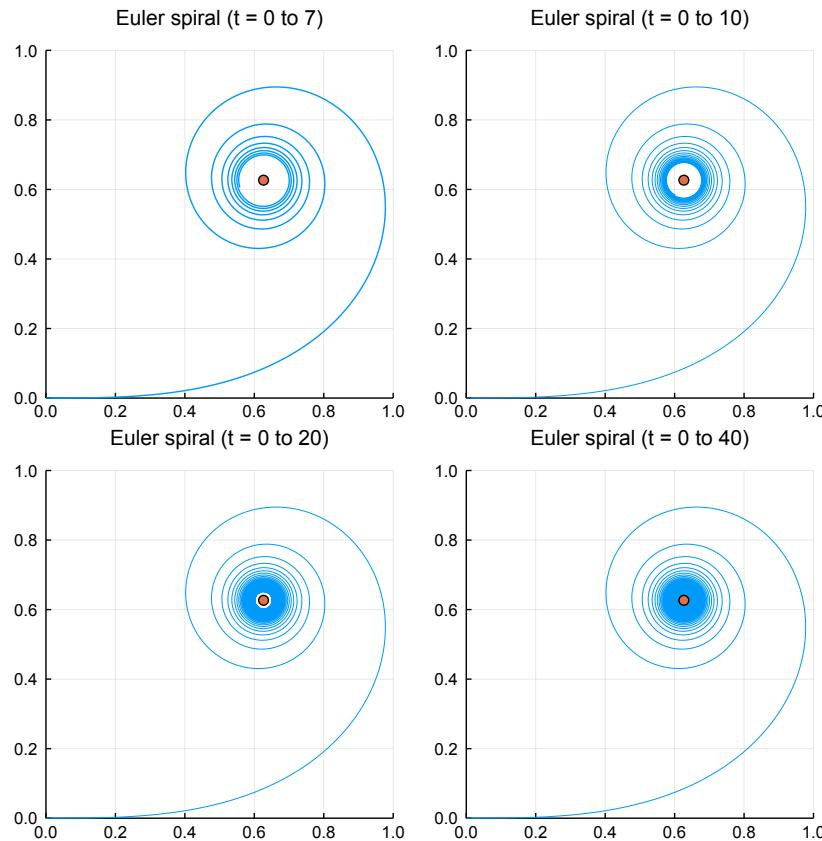
次のセルで  $(C(t), S(t))$  の軌跡をプロットしてみよう.

```

In [12]: 1 ▼ C(t) = quadgk(s->cos(s^2), 0, t)[1]
2 ▼ S(t) = quadgk(s->sin(s^2), 0, t)[1]
3 a = sqrt(pi/2)/2
4
5 P = []
6 ▼ for (lw, tmax) in [(1,7), (0.6,10), (0.4,20), (0.25,40)]
7   t = 0:0.01:tmax
8   p = plot(xlims=(0,1), ylims=(0,1), legend=false)
9   plot!(title="Euler spiral (t = 0 to $tmax)", titlefontsize=10)
10  plot!(C.(t), S.(t), lw=lw)
11  scatter!([a],[a])
12  push!(P, p)
13 end
14 plot(P..., size=(600, 600))
15 #pngplot(P..., size=(600, 600))

```

Out[12]:



```

In [13]: 1 x = symbols("x", positive=true)
2 integrate(x^(-1)*sin(x), (x,0,oo))

```

Out[13]:  $\frac{\pi}{2}$

```

In [14]: 1 a, x = symbols("a, x", positive=true)
2 integrate(exp(-a*x)*x^(-1)*sin(x), (x,0,oo))

```

Out[14]:  $-\operatorname{atan}(a) + \frac{\pi}{2}$

```

In [15]: 1 x = symbols("x", positive=true)
2 [
3   integrate(cos(x^2), (x,-oo,oo)),
4   integrate(sin(x^2), (x,-oo,oo))
5 ]

```

Out[15]:  $\begin{bmatrix} \frac{\sqrt{2}\sqrt{\pi}}{2} \\ \frac{\sqrt{2}\sqrt{\pi}}{2} \end{bmatrix}$

```
In [16]: 1 x = symbols("x", positive=true)
2 [
3     integrate(x^(-Sym(1)/2)*cos(x), (x,0,oo)),
4     integrate(x^(-Sym(1)/2)*sin(x), (x,0,oo))
5 ]
```

Out[16]:  $\left[ \frac{\sqrt{2}\sqrt{\pi}}{2}, \frac{\sqrt{2}\sqrt{\pi}}{2} \right]$

**問題:**  $a > 0$  のとき, 広義積分

$$\int_1^\infty \frac{\sin x}{x^a} dx = \lim_{R \rightarrow \infty} \int_1^R \frac{\sin x}{x^a} dx$$

が収束することを示せ.

**解答例:** 部分積分によって

$$\begin{aligned} \int_1^R \frac{\sin x}{x^a} dx &= \int_1^R x^{-a} (-\cos x)' dx \\ &= [-x^{-a} \cos x]_1^R - \int_1^R (x^{-a})' (-\cos x) dx \\ &= -R^{-a} \cos R + \cos 1 - a \int_1^R x^{-a-1} \cos x dx. \end{aligned}$$

第1項の  $-R^{-a} \cos R$  は  $R \rightarrow \infty$  で 0 に収束し, 第3項の積分は

$$\int_1^R |x^{-a-1} \cos x| dx \leq \int_1^R x^{-a-1} dx = \left[ \frac{x^{-a}}{-a} \right]_1^R = \frac{1 - R^{-a}}{a}$$

なので  $R \rightarrow \infty$  で絶対収束する. これで示すべきことがすべて示された.  $\square$

**解説:**

$$\int_1^R x^{-a} dx = \begin{cases} \frac{1 - R^{1-a}}{1 - a} & (a \neq 1) \\ \log R & (a = 1) \end{cases}$$

なので,  $a \leq 1$  のとき, この積分は  $R \rightarrow \infty$  で無限大に発散する. しかし,  $x^{-a}$  を微分すると,  $-ax^{-(a+1)}$  になり,  $a > 0$  ならば  $a+1 > 1$  となり,  $x^{-(a+1)}$  の 1 から  $\infty$  までの積分が収束することを使える. 上の問題の解答例では, 部分積分によってそれを実現している. 絶対収束しない積分を上手に部分積分すると絶対収束する形に直せることがよくある. 部分積分はこの意味での非常に有用である.  $\square$

**問題:**  $a(x), b(x)$  は  $x \geq 1$  の連続函数であるとし,  $A(x) = \int_1^x a(\xi) d\xi$  とおく.  $A(x)$  は  $x \geq 1$  で有界であると仮定する:  $A(x) \leq M$  ( $x \geq 1$ ). さらに  $b(x)$  は単調減少な  $C^1$  函数でかつ  $x \rightarrow \infty$  で  $b(x) \rightarrow 0$  となっていると仮定する(特に  $b(x) \geq 0$  ( $x \geq 0$ ) である). このとき, 広義積分

$$\int_1^\infty a(x)b(x) dx = \lim_{R \rightarrow \infty} \int_1^R a(x)b(x) dx \quad (1)$$

が収束することを示せ.

**解答例:** 部分積分によって,

$$\int_1^R a(x)b(x) dx = A(R)b(R) - A(1)b(1) - \int_1^R A(x)b'(x) dx. \quad (2)$$

そして,  $x \geq 1$  のとき,  $|A(x)| \leq M$  かつ  $b(x)$  の単調減少性より  $b'(x) \leq 0$  なので,

$$|A(x)b'(x)| \leq -Mb'(x) \quad (x \geq 1).$$

ゆえに,  $b(x) \geq 0$  であることも使うと,

$$\int_1^R |A(x)b'(x)| dx \leq - \int_1^R Mb'(x) dx = M(b(1) - b(R)) \leq Mb(1) < \infty.$$

これで, 上の(2)の右辺の積分が絶対収束することがわかった. (2)の右辺の  $A(R)b(R)$  は  $R \rightarrow \infty$  のとき  $b(R) \rightarrow 0$  より 0 に収束する. これで, (1)の広義積分が収束することがわかった.  $\square$

**注意:** 上の解答例と

- 黒木玄, [02 級数 \(<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/02%20series.ipynb>\) の 2.2 条件収束する複素級数の例 \(<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/02%20series.ipynb#%E6%9D%A1%E4%BB%B6%E5%8F%8E%E6%9D%9F>\)](http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/02%20series.ipynb)

にある「例の解説のため的一般論」を比較してみよ. 積分と和の違いに過ぎず, 本質的には同じ内容であることがわかる.  $\square$

## 積分と極限の交換

このノートを含む一連のノート群では安全な極限の交換しか扱わない。

読者は積分と極限の交換、積分と微分の交換、積分順序の交換などを形式的に自由にできることを認めて読んで欲しい。

もしくは以下の結果が証明抜きで自由に使用されると考えてもよい。

(J) 積分の順序交換:  $\int_a^b \left( \int_c^d |f(x, y)| dy \right) dx < \infty$  ならば

$$\int_a^b \left( \int_c^d f(x, y) dy \right) dx = \int_c^d \left( \int_a^b f(x, y) dx \right) dy.$$

(K) 積分記号化での微分: 各パラメーター  $\alpha$  について  $\int_a^b |f(\alpha, x)| dx < \infty$  であり、各  $x \in (a, b)$  について偏導函数(偏微分)

$$\frac{\partial f(\alpha, x)}{\partial \alpha} = \lim_{h \rightarrow 0} \frac{f(\alpha + h, x) - f(\alpha, x)}{h}$$

が存在し、ある非負値函数  $\varphi(x) \geqq 0$  で \$\$かつ

$$\int_a^b \varphi(x) dx < \infty, \quad \left| \frac{\partial f(\alpha, x)}{\partial \alpha} \right| \leqq \varphi(x) \quad (a < x < b)$$

を満たすものが存在するならば、

$$\frac{d}{d\alpha} \int_a^b f(\alpha, x) dx = \int_a^b \frac{\partial f(\alpha, x)}{\partial \alpha} dx.$$

(L) Lebesgue(ルベーグ)の収束定理: 函数列  $f_n(x)$  は各点収束している(すなわち各  $x \in (a, b)$  ごとに数列  $f_n(x)$  が収束している)とし、ある非負値函数  $\varphi(x) \geqq 0$  で

$$\int_a^b \varphi(x) dx < \infty$$

を満たすものが存在すると仮定する。このとき、

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \lim_{n \rightarrow \infty} f_n(x) dx.$$

以上は少々荒い説明の仕方になっている。正確でかつより一般的な結果およびその証明については

- 猪狩惺、『実解析入門 (<https://www.amazon.co.jp/dp/4000054449>)』、岩波書店、1996

など多くの教科書に書いてある。この本は品切れで中古本にプレミア価格がついている。大学の図書館などで借りて読めば安上がりである。その本において、(L)はp.78の定理3.23に、(K)はp.79の例3.3に、(J)のより正確な表現であるFubiniの定理はpp.120-123に書いてある。

## 積分と極限が交換できない場合

このノートを含む一連のノート群では安全な極限の交換しか行わないが、積分と極限が交換できない例は非常に簡単に作れる。以下を見よ。

例(Lebesgueの収束定理の結論が成立しない場合1): 函数列  $f_n(x)$  を

$$f_n(x) = \begin{cases} 1/n & (0 < x \leqq n) \\ 0 & (\text{otherwise}) \end{cases}$$

と定める。さらに函数  $g(x)$  を

$$g(x) = \max \{f_n(x)\}_{n=1}^{\infty} = \begin{cases} 1/n & (n-1 < x \leqq n) \\ 0 & (\text{otherwise}) \end{cases}$$

と定める。このとき、任意の実数  $x$  について  $\lim_{n \rightarrow \infty} f_n(x) = 0$  だが、 $\int_0^{\infty} f_n(x) dx = 1$  なので、

$$\lim_{n \rightarrow \infty} \int_0^{\infty} f_n(x) dx = 1 \neq 0 = \int_0^{\infty} \lim_{n \rightarrow \infty} f_n(x) dx$$

となる。そして、

$$\int_0^{\infty} g(x) dx = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

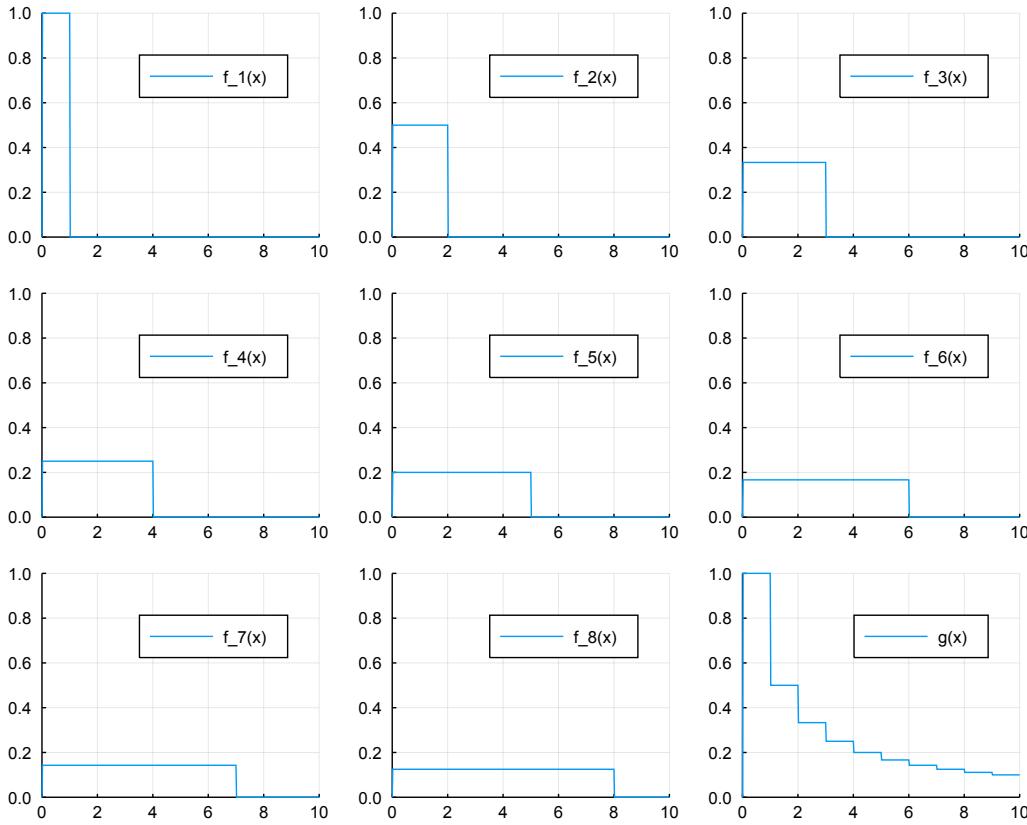
であることにも注意せよ。□

問題: 上の問題の  $f_n(x), g(x)$  のグラフを描け。

解答例: 次のセルをみよ。□

```
In [17]: 1 f(n,x) = 0 < x ≤ n ? 1/n : zero(x)
2 g(x) = x > 0 ? 1/ceil(x) : zero(x)
3 x = 0.0:0.02:10.0
4
5 P = []
6 for n in 1:8
7     p = plot(x, f.(n,x), label="f_$(n)(x)")
8     push!(P, p)
9 end
10 p = plot(x, g.(x), label="g(x)")
11 push!(P, p)
12 plot(P..., size=(750, 600), ylims=(0,1))
13 #pngplot(P..., size=(750, 600), ylims=(0,1))
```

Out[17]:



## Dirichlet積分とその一般化

問題: Dirichlet積分の一般化に関する公式

$$\int_0^\infty \frac{e^{-ax} \sin x}{x} dx = \int_a^\infty \frac{dt}{1+t^2} = \frac{\pi}{2} - \arctan a \quad (a \geq 0)$$

を  $a > 0$  の場合について示せ.  $a = 0$  の場合の公式

$$\int_0^\infty \frac{\sin x}{x} dx = \frac{\pi}{2}$$

の左辺を**Dirichlet積分**と呼ぶことがある. Dirichlet積分は条件収束する広義積分の典型例である.

**解答例1:**  $a > 0$  と仮定し,  $I(a) = \int_0^\infty \frac{e^{-ax} \sin x}{x} dx$  とおく.

$a \rightarrow \infty$  で  $I(a) \rightarrow 0$  である( $I(\infty) = 0$  とおく). 積分記号化での微分(K)と上の方にある三角函数のLaplace変換の問題の結果より

$$I'(a) = \int_0^\infty \frac{\partial}{\partial a} \frac{e^{-ax} \sin x}{x} dx = - \int_0^\infty e^{-ax} \sin x dx = \frac{-1}{1+a^2}.$$

ゆえに

$$I(a) = I(\infty) + \int_\infty^a \frac{-1}{1+t^2} dt = \int_a^\infty \frac{dt}{1+t^2}.$$

さらに

$$\int_a^\infty \frac{dt}{1+t^2} = \int_0^\infty \frac{dt}{1+t^2} - \int_0^a \frac{dt}{1+t^2} = \frac{\pi}{2} - \arctan a.$$

これで示すべきことがすべて示された.  $\square$

**解答例2:**  $a > 0$  であると仮定する.  $x > 0$  のとき,  $\frac{e^{-ax}}{x} = \int_a^\infty e^{-tx} dt$  なので

$$\int_0^\infty \frac{e^{-ax} \sin x}{x} dx = \int_0^\infty \left( \int_a^\infty e^{-tx} \sin x dt \right) dx.$$

積分の順序交換(J)によって,

$$\int_0^\infty \left( \int_a^\infty e^{-tx} \sin x dt \right) dx = \int_a^\infty \left( \int_0^\infty e^{-tx} \sin x dx \right) dt.$$

上方の三角函数のLaplace変換の問題の結果より,

$$\int_a^\infty \left( \int_0^\infty e^{-tx} \sin x dx \right) dt = \int_a^\infty \frac{1}{1+t^2} dt.$$

これで  $a > 0$  の場合の公式(2)が示された.  $\square$

**注意:**  $a > 0$  の場合には絶対収束する積分しか出て来ないので積分順序を自由に交換できて簡単である.  $a = 0$  の場合にも積分順序の交換が可能なことを示せば  $a = 0$  の場合の(1)も証明される.  $\square$

**問題(次の問題の解答例のための準備):** 次の公式を示せ:  $t > 0$  のとき,

$$\int e^{-tx} \sin x dx = \frac{-e^{-tx}}{1+t^2}(t \cos x - \sin x), \quad \int e^{-tx} \sin x dx = \frac{-e^{-tx}}{1+t^2}(t \sin x + \cos x).$$

**解答例:**

$$\begin{aligned} - \int e^{-tx} e^{ix} dx &= - \int_0^\infty e^{-(t-i)x} dx = \frac{e^{-(t-i)x}}{t-i} = \frac{t+i}{t^2+1} e^{-tx} (\cos x + i \sin x) \\ &= \frac{e^{-tx}}{1+t^2} ((t \cos x - i \sin x) + i(t \sin x + \cos x)). \end{aligned}$$

この等式の両辺の実部と虚部が示すべき公式にちょうどなっている.  $\square$

この問題の結果より,  $t > 0$  のとき,

$$\begin{aligned} \int_0^R e^{-tx} \cos x dx &= \frac{-e^{-tR}}{1+t^2}(t \cos R - \sin R) + \frac{t}{1+t^2}, \\ \int_0^R e^{-tx} \sin x dx &= \frac{-e^{-tR}}{1+t^2}(t \sin R + \cos R) + \frac{1}{1+t^2}. \end{aligned}$$

```
In [18]: 1 t, x = symbols("t x", positive=true)
2 ▼ [
3     factor(integrate(exp(-t*x)*cos(x), x)),
4     factor(integrate(exp(-t*x)*sin(x), x))
5 ]
```

```
Out[18]: ⎡ −e−tx ⎤ ⎡ −e−tx ⎤
      ⎣ ⎢ t cos (x) − sin (x) ⎥ ⎢ t sin (x) + cos (x) ⎥ ⎦
```

**問題:** Dirichlet積分の公式

$$\int_0^\infty \frac{\sin x}{x} dx = \lim_{R \rightarrow \infty} \int_0^R \frac{\sin x}{x} dx = \frac{\pi}{2}$$

を示せ. (1つ目の等号は広義積分の定義であり, 示す必要はない.)

**解答例:**  $R > 0$  であるとする.  $x > 0$  のとき,  $\frac{1}{x} = \int_0^\infty e^{-tx} dt$  なので

$$\int_0^R \frac{\sin x}{x} dx = \int_0^R \left( \int_0^\infty e^{-tx} \sin x dt \right) dx.$$

積分の順序交換(J)によって

$$\int_0^R \frac{\sin x}{x} dx = \int_0^\infty \left( \int_0^R e^{-tx} \sin x dx \right) dt.$$

すぐ上の問題の結果より,

$$\int_0^R \frac{\sin x}{x} dx = \int_0^\infty \frac{-e^{-tR}}{1+t^2}(t \sin R + \cos R) dt + \int_0^\infty \frac{dt}{1+t^2}.$$

右辺の前者の積分は  $R \rightarrow \infty$  で 0 に収束し, 後者の積分は  $\pi/2$  に等しいので,

$$\int_0^\infty \frac{\sin x}{x} dx = \lim_{R \rightarrow \infty} \int_0^R \frac{\sin x}{x} dx = \frac{\pi}{2}.$$

これで公式(1)が示された.  $\square$

**問題(次の問題の準備):** 次を示せ.

$$\int_0^\infty \frac{dt}{1+t^4} = \int_0^\infty \frac{t^2 dt}{1+t^4} = \frac{\sqrt{2}\pi}{4}$$

```
In [19]: 1 t = symbols("t", real=True)
2 [
3     integrate(t^2/(t^4+1), (t,0,oo)),
4     integrate(1/(t^4+1), (t,0,oo)),
5 ]
```

$$\begin{bmatrix} \frac{\sqrt{2}\pi}{4} \\ \frac{\sqrt{2}\pi}{4} \end{bmatrix}$$

**解答例1:**

$t = 1/u$  とおくと,

$$\int_0^\infty \frac{dt}{1+t^4} = \int_\infty^0 \frac{-u^{-2} du}{1+u^{-4}} = \int_0^\infty \frac{u^2 du}{u^4+1} = \int_0^\infty \frac{t^2 dt}{1+t^4}.$$

天下りになってしまふが,  $F(t)$  を次のように定める:

$$F(t) = \frac{\sqrt{2}}{8} \log \frac{t^2 + \sqrt{2}t + 1}{t^2 - \sqrt{2}t + 1} + \frac{\sqrt{2}}{4} (\arctan(\sqrt{2}t + 1) + \arctan(\sqrt{2}t - 1)).$$

この導函数を計算すると  $F'(t) = 1/(1+t^4)$  となることを確認できる(かなり面倒な計算が必要になる). したがって,

$$\int_0^\infty \frac{dt}{1+t^4} = [F(t)]_0^\infty = F(\infty) = \frac{\sqrt{2}}{4}\pi. \quad \square$$

$1/(1+t^4)$  の原始函数の求め方については

- E. ハイラー, G. ワナー, 『解析教程 上』(<https://www.amazon.co.jp/dp/4621062034>)』

の第2章(5.16)式の周辺(pp.172-174)が詳しい.  $1/(1+t^4)$  の部分分数展開を使う. 下の方の解答例2の直後の問題とその解答例も見よ.

```
In [20]: 1 t = symbols("t")
2 r = sqrt(Sym(2))
3 F = r/8*log((t^2+r*t+1)/(t^2-r*t+1)) + r/4*(atan(r*t+1)+atan(r*t-1))
4 display(F)
5 simplify(diff(F, t))
```

$$\frac{\sqrt{2}}{4}(\arctan(\sqrt{2}t - 1) + \arctan(\sqrt{2}t + 1)) + \frac{\sqrt{2}}{8} \log \left( \frac{t^2 + \sqrt{2}t + 1}{t^2 - \sqrt{2}t + 1} \right)$$

$$\frac{1}{t^4 + 1}$$

**解答例2:** 複素解析を使う方法.  $R > 0$  とし,  $C_1$  は 0 から  $R$  への線分  $t, 0 \leq t \leq R$  であるとし,  $C_2$  は  $R$  から  $iR$  への曲線  $Re^{i\theta}, 0 \leq \theta \leq \pi/2$ ,  $C_3$  は  $Ri$  から 0 への線分であるとする. このとき,

$$\int_{C_3} \frac{dz}{1+z^4} = - \int_0^R \frac{d(it)}{1+(it)^4} = -i \int_0^R \frac{dt}{1+t^4} = -i \int_{C_1} \frac{dt}{1+t^4}.$$

$1/(1+z^4)$  の  $z = e^{i\pi/4}$  における留数は

$$\frac{1}{(e^{i\pi/4} - e^{i3\pi/4})(e^{i\pi/4} - e^{i5\pi/4})(e^{i\pi/4} - e^{i7\pi/4})} = \frac{1}{\sqrt{2} \cdot \sqrt{2}(1+i) \cdot \sqrt{2}i} = \frac{1}{2\sqrt{2}(1+i)i}$$

なので、留数定理より,

$$\frac{2\pi i}{2\sqrt{2}(1+i)i} = (1-i) \int_{C_1} \frac{dz}{1+z^4} + \int_{C_2} \frac{dz}{1+z^4}.$$

$R \rightarrow \infty$  で右辺の  $C_1$  上での積分は  $\int_0^\infty dt/(1+t^4)$  に収束し,  $C_2$  上での積分は 0 に収束するので,

$$\frac{2\pi i}{2\sqrt{2}(1+i)i} = (1-i) \int_0^\infty \frac{dt}{1+t^4}.$$

ゆえに,

$$\int_0^\infty \frac{dt}{1+t^4} = \frac{2\pi}{2\sqrt{2}\cdot 2} = \frac{\sqrt{2}\pi}{4}. \quad \square$$

問題:  $1/(t^4 + 1)$  の原始函数を求めよ.

解答例: 以下, 記号の簡単のため  $r = \sqrt{2}$  とおく( $r^2 = 2$ ). 因数分解

$$t^4 + 1 = (t^2 + rt + 1)(t^2 - rt + 1)$$

より,

$$\frac{2r}{t^4 + 1} = \frac{t+r}{t^2 + rt + 1} - \frac{t-r}{t^2 - rt + 1}.$$

一般に  $a > 0$  のとき

$$\int \frac{s \, ds}{s^2 + a^2} = \frac{1}{2} \log(s^2 + a^2), \quad \int \frac{ds}{s^2 + a^2} = \frac{1}{a} \arctan \frac{s}{a}.$$

$t = s - r/2$  とおくと

$$\begin{aligned} \int \frac{t+r}{t^2 + rt + 1} dt &= \int \frac{s+r/2}{s^2 + 1/2} ds = \frac{1}{2} \log(2s^2 + 1) + \arctan(rs) \\ &= \frac{1}{2} \log 2 + \frac{1}{2} \log(t^2 + rt + 1) + \arctan(rt + 1). \end{aligned}$$

$t = s + r/2$  とおくと

$$\begin{aligned} - \int \frac{t-r}{t^2 - rt + 1} dt &= - \int \frac{s-r/2}{s^2 + 1/2} ds = -\frac{1}{2} \log(2s^2 + 1) + \arctan(rs) \\ &= -\frac{1}{2} \log 2 - \frac{1}{2} \log(t^2 - rt + 1) + \arctan(rt - 1). \end{aligned}$$

ゆえに

$$2r \int \frac{dt}{t^4 + 1} = \frac{1}{2} \log \frac{t^2 + rt + 1}{t^2 - rt + 1} + \arctan(rt + 1) + \arctan(rt - 1).$$

したがって,

$$\int \frac{dt}{t^4 + 1} = \frac{1}{4r} \log \frac{t^2 + rt + 1}{t^2 - rt + 1} + \frac{1}{2r} (\arctan(rt + 1) + \arctan(rt - 1)). \quad \square$$

```
In [21]: 1 s = symbols("s", real=true)
2 a = symbols("a", positive=true)
3 display(integrate(s/(s^2+a^2), s))
4 display(integrate(1/(s^2+a^2), s))
```

$$\frac{1}{2} \log(a^2 + s^2)$$

$$\frac{1}{a} \operatorname{atan}\left(\frac{s}{a}\right)$$

```
In [22]: 1 r, s, t = symbols("r s t", real=true)
2 v = sqrt(Sym(2))
3
4 A = (t + r)/(t^2 + r*t + 1)
5 B = -(t - r)/(t^2 - r*t + 1)
6 F = expand(1/((t^2 + r*t + 1)*(t^2 - r*t + 1)))
7
8 display(A)
9 display(B)
10 display(F(r=>v))
11 display(simplify(A+B-2r*F)) # Verify 2r*F == A+B
12
13 # integrate A
14 As = simplify(simplify(A(t=>s-r/2))(r^2=>2))
15 intAs = integrate(s/(s^2+Sym(1)/2), s) + integrate(r/2/(s^2+Sym(1)/2), s)
16 intA = simplify(expand(intAs(s=>t+r/2)))
17 display(As)
18 display(simplify(intAs(r=>v)))
19 display(simplify(intA(r=>v)))
20
21 # integrate B
22 Bs = simplify(simplify(B(t=>s+r/2))(r^2=>2))
23 intBs = -integrate(s/(s^2+Sym(1)/2), s) + integrate(r/2/(s^2+Sym(1)/2), s)
24 intB = simplify(expand(intBs(s=>t-r/2)))
25 display(Bs)
26 display(simplify(intBs(r=>v)))
27 display(simplify(intB(r=>v)))
28
29 G = simplify((1/(2r)*(intA + intB))(r=>v))
30 display(G)
31 simplify(diff(G, t)) # Verify dF/dt = 1/(t^4+1)
```

$$\begin{aligned} & \frac{r+t}{rt+t^2+1} \\ & \frac{r-t}{-rt+t^2+1} \\ & \frac{1}{t^4+1} \\ & 0 \\ & \frac{r+2s}{2s^2+1} \\ & \frac{1}{2}\log(2s^2+1) + \operatorname{atan}(\sqrt{2}s) \\ & \frac{1}{2}\log(t^2+\sqrt{2}t+1) + \operatorname{atan}(\sqrt{2}t+1) + \frac{1}{2}\log(2) \\ & \frac{r-2s}{2s^2+1} \\ & -\frac{1}{2}\log(2s^2+1) + \operatorname{atan}(\sqrt{2}s) \\ & -\frac{1}{2}\log(t^2-\sqrt{2}t+1) + \operatorname{atan}(\sqrt{2}t-1) - \frac{1}{2}\log(2) \\ & \frac{\sqrt{2}}{8}(-\log(t^2-\sqrt{2}t+1) + \log(t^2+\sqrt{2}t+1) + 2\operatorname{atan}(\sqrt{2}t-1) + 2\operatorname{atan}(\sqrt{2}t+1)) \end{aligned}$$

Out[22]:  $\frac{1}{t^4+1}$

問題: Fresnel積分の特殊値の公式

$$\int_0^\infty \frac{\cos x}{\sqrt{x}} dx = \int_0^\infty \frac{\sin x}{\sqrt{x}} dx = \sqrt{\frac{\pi}{2}}$$

を示せ。

解答例: Gauss積分の公式より,  $x > 0$  のとき

$$2 \int_0^\infty e^{-xt^2} dt = \int_{-\infty}^\infty e^{-xt^2} dt = \sqrt{\frac{\pi}{x}}$$

ゆえに

$$\int_0^\infty \frac{e^{ix}}{\sqrt{x}} dx = \frac{2}{\sqrt{\pi}} \int_0^\infty \left( \int_0^\infty e^{-xt^2} e^{ix} dt \right) dx$$

積分順序を交換すると(本当はこの部分を上の問題と同様の方法で正当化する必要がある),

$$\frac{2}{\sqrt{\pi}} \int_0^\infty \left( \int_0^\infty e^{-xt^2} e^{ix} dt \right) dx = \frac{2}{\sqrt{\pi}} \int_0^\infty \left( \int_0^\infty e^{-xt^2} e^{ix} dx \right) dt.$$

三角函数のLaplace変換の問題の解答例2より,

$$\frac{2}{\sqrt{\pi}} \int_0^\infty \left( \int_0^\infty e^{-xt^2} e^{ix} dx \right) dt = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{t^2 + i}{t^4 + 1} dt$$

上の問題の結果より,

$$\frac{2}{\sqrt{\pi}} \int_0^\infty \frac{t^2 - i}{t^4 + 1} dt = \frac{2}{\sqrt{\pi}} (1+i) \frac{\sqrt{2}\pi}{4} = (1+i) \sqrt{\frac{\pi}{2}}.$$

したがって,

$$\int_0^\infty \frac{e^{ix}}{\sqrt{x}} dx = (1+i) \sqrt{\frac{\pi}{2}}.$$

両辺の実部と虚部を見れば示したい結果が得られる.  $\square$

上の解答例と本質的に同じ解答の詳しい解説が

- E. ハイラー, G. ワナー, 『[解析教程 下 \(https://www.amazon.co.jp/dp/4621061909\)](https://www.amazon.co.jp/dp/4621061909)』

の第4章の演習問題5.14(p.254, この問題はこの本の最後に載っている問題)の解答解説(pp.313-314)にある.

**注意:** 複素解析を使えばFresnel積分の特殊値の公式

$$\int_0^\infty \cos t^2 dt = \int_0^\infty \sin t^2 dt = \frac{1}{2} \sqrt{\frac{\pi}{2}}$$

を以下のように比較的容易に示すことができる.

$R > 0$  であるとし,  $C_1$  は 0 から  $R$  への線分,  $C_2$  は  $R$  から  $Re^{\pi i/4}$  への円弧,  $C_3$  は 0 から  $Re^{\pi i/4}$  への線分とすると, Cauchyの積分定理より,

$$\int_{C_3} e^{-z^2} dz = \int_{C_1} e^{-z^2} dz + \int_{C_2} e^{-z^2} dz.$$

そして,

$$\begin{aligned} \int_{C_3} e^{-z^2} dz &= \int_0^R \exp(-(e^{\pi i/4} t)^2) d(e^{\pi i/4} t) dt = e^{\pi i/4} \int_0^R e^{-it^2} dt, \\ \int_{C_1} e^{-z^2} dz &= \int_0^R e^{-t^2} dt \rightarrow \frac{\sqrt{\pi}}{2} \quad (R \rightarrow \infty), \\ \int_{C_2} e^{-z^2} dz &\rightarrow 0 \quad (R \rightarrow \infty). \end{aligned}$$

したがって,

$$e^{\pi i/4} \int_0^\infty e^{-it^2} dt = \frac{\sqrt{\pi}}{2}.$$

両辺に  $e^{-\pi i/4} = \frac{1-i}{\sqrt{2}}$  をかけると,

$$\int_0^\infty e^{-it^2} dt = (1-i) \frac{1}{2} \sqrt{\frac{\pi}{2}}.$$

$e^{-it^2} = \cos t^2 - i \sin t^2$  に注意して, 両辺の実部と虚部を比較すれば欲しい結果が得られる.  $\square$

上の複素解析を用いたFresnel積分の特殊値の導出のより詳しい解説が

- 高木貞治, 『[解析概論 \(https://www.amazon.co.jp/dp/4000051717\)](https://www.amazon.co.jp/dp/4000051717)』

の第62節の例2(pp.224-225)にある.

ガンマ函数との関連については

- 黒木玄, [ガンマ分布の中心極限定理とStirlingの公式 \(https://genkuroki.github.io/documents/20160501StirlingFormula.pdf\)](https://genkuroki.github.io/documents/20160501StirlingFormula.pdf)

の第8.6節を参照せよ. 一般に, ちょっとした複素解析の利用によって,  $0 \leq s < 1$  のとき

$$\int_0^\infty x^{s-1} \sin x dx = \Gamma(s) \sin \frac{\pi s}{2} = \Gamma(s+1) \frac{\sin(\pi s/2)}{s}$$

となることを示せる. ( $\sin(\pi s/2)/s$  の  $s = 0$  での値は極限を取ることによって自然に  $\pi/2$  だと定義される.) これの  $s = 0$  の場合が  
 $\int_0^\infty \frac{\sin x}{x} dx = \frac{\pi}{2}$  であり,  $s = 1/2$  の場合が  $\int_0^\infty \frac{\sin x}{\sqrt{x}} dx = \sqrt{\frac{\pi}{2}}$  である. このように, Dirichlet積分とFresnel積分(の特殊値)はガンマ函数によって統一的に取り扱うことができる. 別々に公式を証明する必要はない.  $\square$

In [ ]:

1

# Gauss積分, ガンマ函数, ベータ函数

黒木玄

2018-06-21

- Copyright 2018 Gen Kuroki
- License: MIT <https://opensource.org/licenses/MIT> (<https://opensource.org/licenses/MIT>)

このファイルは次の場所できれいに閲覧できる:

- <http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/10%20Gauss%2C%20Gamma%2C%20Beta.ipynb>  
<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/10%20Gauss%2C%20Gamma%2C%20Beta.ipynb>

このファイルは [Julia Box \(https://juliabox.com\)](https://juliabox.com) で利用できる。

自分のパソコンに[Julia言語 \(https://julialang.org/\)](https://julialang.org/)をインストールしたい場合には

- [WindowsへのJulia言語のインストール \(http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f\)](http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f)

を参照せよ。

論理的に完璧な説明をするつもりはない。細部のいい加減な部分は自分で訂正・修正せよ。

## Table of Contents

### 1 Gauss積分

- [1.1 Gauss積分の公式](#)
- [1.2 Gauss積分を使う簡単な計算例](#)
- [1.3 Gauss分布のFourier変換](#)
- [1.4 Gauss積分の公式の導出](#)
  - [1.4.1 方法1: 高さ  \$\pi\$  で輪切りにする方法](#)
  - [1.4.2 方法2: 極座標を使う方法](#)
  - [1.4.3 方法3:  \$y = x \tan \theta\$  と変数変換する方法](#)

### 2 ガンマ函数とベータ函数

- [2.1 ガンマ函数とベータ函数の定義](#)
- [2.2 ガンマ函数の函数の特殊値と函数等式](#)
- [2.3 ベータ函数とガンマ函数の関係](#)
  - [2.3.1 方法1: 極座標変換を使う方法](#)
  - [2.3.2 方法2: 置換積分と積分の順序交換のみを使う方法](#)
  - [2.3.3 ベータ函数とガンマ函数の関係の応用](#)

### 3 Stirlingの公式とLaplaceの方法

- [3.1 Stirlingの公式](#)
- [3.2 Laplaceの方法](#)
- [3.3 Laplaceの方法の弱形](#)

In [1]:

```
1  using Plots
2  gr(); ENV["PLOTS_TEST"] = "true"
3  #clibrary(:colorset)
4  clibrary(:misc)
5
6  ▼ function pngplot(P...; kwargs...)
7      sleep(0.1)
8      pngfile = tempname() * ".png"
9      savefig(plot(P...; kwargs...), pngfile)
10     showimg("image/png", pngfile)
11  end
12  pngplot(; kwargs...) = pngplot(plot!(; kwargs...))
13
14  ▼ showing(mime, fn) = open(fn) do f
15      base64 = base64encode(f)
16      display("text/html", """""")
17  end
18
19  using SymPy
20  #sympy[:init_printing](order="lex") # default
21  #sympy[:init_printing](order="rev-lex")
22
23  using SpecialFunctions
24  using QuadGK
```

## 1 Gauss積分

## 1.1 Gauss積分の公式

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

をGauss積分の公式と呼ぶことにする。証明は後で行う。

このノートの筆者は大学新入生が習う積分の公式の中でこれが最も重要であると考えている。ガウス積分が重要だと考える理由は以下の通り。

(1) この公式自体が非常に面白い形をしている。左辺を見てもどこにも円周率は見えないが、右辺には円周率が出て来る。しかも円周率がそのまま出て来るのではなく、その平方根が出て来る。

(2) 様々な方法を使ってGauss積分の公式を証明できる。

(3) Gauss積分の公式は確率論や統計学で正規分布を扱うときには必須である。正規分布は中心極限定理によって特別に重要な役目を果たす確率分布である。

(4) Gauss積分はガンマ函数に一般化される。

(5) Gauss積分はLaplaceの方法の基礎である。Laplaceの方法はある種の積分の漸近挙動を調べるために最も基本的な方法であり、解析学の応用において基本的かつ重要である。

(6) 特にGauss積分で階乗に等しい積分を近似することによって、Stirlingの公式が得られる。(Stirlingの公式  $n! \sim n^n e^{-n} \sqrt{2\pi n}$  の平行根の因子はGauss積分を経由して得られる。)

以上のようにGauss積分は純粹数学的にも応用数学的にも基本的かつ重要である。

## 1.2 Gauss積分を使う簡単な計算例

問題: 上の公式を使って、 $a > 0$  のとき、

$$\int_{-\infty}^{\infty} e^{-y^2/a} dy = \sqrt{a\pi}$$

となることを示せ。

注意:  $a$  を  $1/a$  で置き換えれば

$$\int_{-\infty}^{\infty} e^{-ay^2} dy = \sqrt{\frac{\pi}{a}}$$

も得られる。

解答例: Gauss積分の公式で  $x = \frac{y}{\sqrt{a}}$  と置換積分すると

$$\sqrt{\pi} = \int_{-\infty}^{\infty} e^{-x^2} dx = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} e^{-y^2/a} dy$$

なので、両辺に  $\sqrt{a}$  をかけば示したい公式が得られる。□

問題: 分散  $\sigma^2 > 0$ 、平均  $\mu$  の正規分布の確率密度函数  $p(x)$  が

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)}$$

で定義される。このとき

$$\int_{-\infty}^{\infty} p(x) dx = 1$$

となることを示せ。(この問題より、確率統計学においてGauss積分の公式は必須であることがわかる。)

解答例:  $x = y + \mu$  と置換し、上の問題の結果を使うと、

$$\begin{aligned} \int_{-\infty}^{\infty} p(x) dx &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-(y+\mu-\mu)^2/(2\sigma^2)} dy = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-y^2/(2\sigma^2)} dy \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \sqrt{2\sigma^2\pi} = 1. \quad \square \end{aligned}$$

問題(Lebesgueの収束定理の結論が成立しない場合2): 函数列  $f_n(x)$  を

$$f_n(x) = \frac{1}{\sqrt{n\pi}} e^{-x^2/n}$$

と定める。以下を示せ。

$$(1) \int_{-\infty}^{\infty} f_n(x) dx = 1.$$

$$(2) \text{各 } x \in \mathbb{R} \text{ ごとに } \lim_{n \rightarrow \infty} f_n(x) = 0.$$

$$(3) \text{したがって } \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x) dx \neq \int_{-\infty}^{\infty} \lim_{n \rightarrow \infty} f_n(x) dx.$$

**解答例:** (1)はGauss積分の公式から得られる(詳細は自分で計算して確認せよ). (3)は(1)と(2)からただちに得られるので, あとは(2)のみを示せば十分である.  $x \in \mathbb{R}$  を任意に取って固定する. このとき  $n \rightarrow \infty$  で  $\frac{x^2}{n} \rightarrow 0$ ,  $\frac{1}{\sqrt{n\pi}} \rightarrow 0$  となるので,  $f_n(x) \rightarrow 0$  となることもわかる.  $\square$

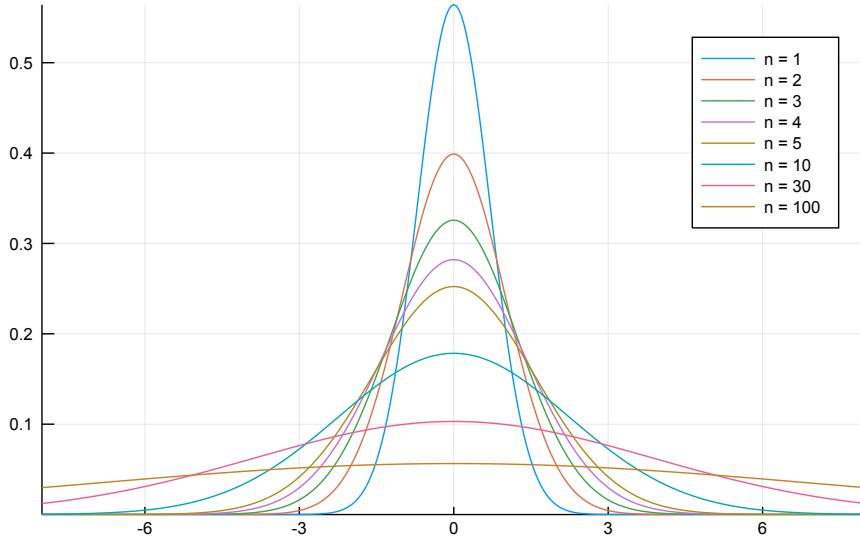
**問題:** すぐ上の問題の函数  $f_n(x)$  のグラフを描いてみよ.

**解答例:** 以下のセルのようになる.

$n$  が大きくなると,  $f_n(x)$  の「分布」は広く拡がる.  $\square$

```
In [2]: 1 f(n,x) = exp(-x^2/n)/sqrt(n*pi)
2 x = -8.0:0.05:8.0
3 p = plot(size=(640,400))
4 for n in [1,2,3,4,5,10, 30, 100]
5 plot!(x, f.(n,x), label="n = $n")
6 end
7 p
```

Out[2]:



**問題:** 次を示せ:  $a > 0$  と  $k = 0, 1, 2, \dots$  について

$$\int_{-\infty}^{\infty} e^{-ax^2} x^{2k} dx = \sqrt{\pi} \frac{1 \cdot 3 \cdots (2k-1)}{2^k} a^{-(2k+1)/2} = \sqrt{\pi} \frac{(2k)!}{2^{2k} k!} a^{-(2k+1)/2}. \quad (1)$$

**注意:**  $a$  を  $1/a$  で置き換えれば次も得られる:

$$\int_{-\infty}^{\infty} e^{-x^2/a} x^{2k} dx = \frac{1 \cdot 3 \cdots (2k-1)}{2^k} \sqrt{a^{2k+1} \pi} = \frac{(2k)!}{2^{2k} k!} \sqrt{a^{2k+1} \pi}. \quad (2)$$

**解答例:** Gauss積分の公式から得られる公式

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\pi} a^{-1/2}$$

の両辺を  $a$  で微分して  $-1$  倍する操作を繰り返すと((K)を使う),

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-ax^2} x^2 dx &= \sqrt{\pi} \frac{1}{2} a^{-3/2}, \\ \int_{-\infty}^{\infty} e^{-ax^2} x^4 dx &= \sqrt{\pi} \frac{1}{2} \frac{3}{2} a^{-5/2}, \\ \int_{-\infty}^{\infty} e^{-ax^2} x^6 dx &= \sqrt{\pi} \frac{1}{2} \frac{3}{2} \frac{5}{2} a^{-7/2}. \end{aligned}$$

$k$  回その操作を繰り返すと,

$$\int_{-\infty}^{\infty} e^{-ax^2} x^{2k} dx = \sqrt{\pi} \frac{1}{2} \frac{3}{2} \cdots \frac{2k-1}{2} a^{-(2k+1)/2}.$$

これより, (1)の前半が成立することがわかる. 後半の成立は

$$\frac{1 \cdot 3 \cdots (2k-1)}{2^k} = \frac{1 \cdot 3 \cdots (2k-1)}{2^k} \frac{2 \cdot 4 \cdots (2k)}{2^k k!} = \frac{(2k)!}{2^{2k} k!} \quad (3)$$

によって確認できる.  $\square$

**注意:** 奇数の積  $1 \cdot 3 \cdots (2k-1)$  について(3)の計算法はよく使われる:

$$1 \cdot 3 \cdots (2k-1) = 1 \cdot 3 \cdots (2k-1) \frac{2 \cdot 4 \cdots (2k)}{2^k k!} = \frac{(2k)!}{2^k k!}$$

例えば二項係数に関する

$$\begin{aligned} (-1)^k \binom{-1/2}{k} &= (-1)^k \frac{(-1/2)(-3/2) \cdots(-(2k-1)/2)}{k!} \\ &= \frac{1 \cdot 3 \cdots (2k-1)}{2^k k!} = \frac{(2k)!}{2^{2k} k! k!} = \frac{1}{2^{2k}} \binom{2k}{k} \end{aligned}$$

もよく出て来る.  $\square$

### 1.3 Gauss分布のFourier変換

$a > 0$  であるとする.  $e^{-x^2/a}$  型の函数を**Gauss分布函数**と呼ぶことがある.

一般に函数  $f(x)$  に対して,

$$\hat{f}(p) = \int_{-\infty}^{\infty} e^{-ipx} f(x) dx$$

を  $f$  の**Fourier変換**(フーリエ変換)と呼ぶ. もしも実数値函数  $f(x)$  が偶函数であれば,

$$e^{-ipx} f(x) = f(x) \cos(px) - if(x) \sin(px)$$

の虚部は奇函数になり, その積分は消えるので

$$\hat{f}(p) = \int_{-\infty}^{\infty} f(x) \cos(px) dx$$

となる.

**問題:**  $a > 0$  とする.  $f(x) = e^{-x^2/a}$  のFourier変換を求めよ.

**解答例:**  $\cos(px) = \sum_{k=0}^{\infty} \frac{(-p^2)^k x^{2k}}{(2k)!}$  より,

$$\begin{aligned} \hat{f}(p) &= \int_{-\infty}^{\infty} e^{-x^2/a} \cos(px) dx = \sum_{k=0}^{\infty} \frac{(-p^2)^k}{(2k)!} \int_{-\infty}^{\infty} e^{-x^2/a} x^{2k} dx \\ &= \sum_{k=0}^{\infty} \frac{(-p^2)^k}{(2k)!} \frac{(2k)!}{2^{2k} k!} \sqrt{a^{2k+1} \pi} = \sqrt{a\pi} \sum_{k=0}^{\infty} \frac{(-ap^2/4)^k}{k!} = \sqrt{a\pi} e^{-ap^2/4}. \end{aligned}$$

3つ目の等号で上の問題の結果を用いた.  $\square$

上の問題の別解については

- 黒木玄, [ガンマ分布の中心極限定理とStirlingの公式](https://genkuroki.github.io/documents/20160501StirlingFormula.pdf) (<https://genkuroki.github.io/documents/20160501StirlingFormula.pdf>).

の第6節を参照せよ.

### 1.4 Gauss積分の公式の導出

$I = \int_{-\infty}^{\infty} e^{-x^2} dx$  とおく.  $I = \sqrt{\pi}$  であることを示したい. そのためには

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} e^{-x^2} dx \cdot \int_{-\infty}^{\infty} e^{-y^2} dy \\ &= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} e^{-x^2} dx \right) e^{-y^2} dy = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx \right) dy \end{aligned}$$

が  $\pi$  に等しいことを証明すればよい. 上の計算の2つ目と3つ目の等号で積分の線形性(A)を用いた.

#### 1.4.1 方法1: 高さ $z$ で輪切りにする方法

$I^2 = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx \right) dy$  は2変数函数  $z = e^{-(x^2+y^2)}$  の  $xyz$  空間内のグラフと  $xy$  平面  $z = 0$  のあいだに挟まれた山型の領域の体積を意味する。

なぜならば、 $S(y) = \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx$  はその領域の  $y$  を固定したときの切断面の面積に等しく、 $\int_{-\infty}^{\infty} S(y) dy$  はその切断面の面積の積分なので領域全体の体積に等しいからである。一般に、長さを積分すれば面積になり、面積を積分すれば体積になる。

その山型の領域の体積は高さ  $z$  での切断面の面積の  $z = 0$  から  $z = 1$  までの積分に等しい。高さ  $z$  での切断面は半径  $\sqrt{x^2 + y^2} = \sqrt{-\log z}$  の円盤になり、その面積は  $-\pi \log z$  になる。ゆえに

$$I^2 = \int_0^1 (-\pi \log z) dz = -\pi [\log z - z]_0^1 = \pi.$$

これより  $I = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$  であることがわかる。

#### 1.4.2 方法2: 極座標を使う方法

以下の方法は2重積分の積分変数の変換の仕方(Jacobianが出て来る)を知っておかなければ使えない。

$x = r \cos \theta, y = r \sin \theta$  とおくと、

$$I^2 = \int_0^{2\pi} d\theta \int_0^{\infty} r e^{-r^2} dr = 2\pi \left[ \frac{e^{-r^2}}{-2} \right]_0^{\infty} = 2\pi \frac{1}{2} = \pi.$$

ゆえに  $I = \sqrt{\pi}$ 。

#### 1.4.3 方法3: $y = x \tan \theta$ と変数変換する方法

$I^2$  は次のようにも表せる：

$$I^2 = 2 \int_0^{\infty} \left( \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dy \right) dx.$$

この積分内で  $x$  は  $x > 0$  を動くと考える。

内側の積分で積分変数を  $-\infty < y < \infty$  から  $-\pi/2 < \theta < \pi/2$  に  $y = x \tan \theta$  によって変換すると、

$$\cos \theta > 0, \quad dy = \frac{x}{\cos^2 \theta} d\theta, \quad x^2 + y^2 = x^2(1 + \tan^2 \theta) = \frac{x^2}{\cos^2 \theta}$$

なので

$$I^2 = 2 \int_0^{\infty} \left( \int_{-\pi/2}^{\pi/2} \exp\left(-\frac{x^2}{\cos^2 \theta}\right) \frac{x}{\cos^2 \theta} d\theta \right) dx.$$

ゆえに積分の順序を交換すると((J)を使う)、

$$\begin{aligned} I^2 &= 2 \int_{-\pi/2}^{\pi/2} \left( \int_0^{\infty} \exp\left(-\frac{x^2}{\cos^2 \theta}\right) \frac{x}{\cos^2 \theta} dx \right) d\theta \\ &= 2 \int_{-\pi/2}^{\pi/2} \left[ \frac{1}{-2} \exp\left(-\frac{x^2}{\cos^2 \theta}\right) \right]_{x=0}^{x=\infty} d\theta = 2 \int_{-\pi/2}^{\pi/2} \frac{1}{2} d\theta = 2 \frac{\pi}{2} = \pi. \end{aligned}$$

したがって  $I = \sqrt{\pi}$ 。

**注意:** 極座標変換  $(x, y) = (r \cos \theta, r \sin \theta)$  が有効な場面では、 $y = x \tan \theta$  という変数変換も有効なことが多い。 $\tan \theta$  の幾何的な意味は「原点を通る直線の傾き」であった。その意味でも  $y = x \tan \theta$  は自然な変数変換だと言える。□

Gauss積分の別の計算の仕方については

- 黒木玄, ガンマ分布の中心極限定理とStirlingの公式 (<https://genkuroki.github.io/documents/20160501StirlingFormula.pdf>)

の第7節を参照せよ。

## 2 ガンマ函数とベータ函数

### 2.1 ガンマ函数とベータ函数の定義

$s > 0, p > 0, q > 0$  と仮定する。 $\Gamma(s)$  と  $B(p, q)$  を次の積分で定義する：

$$\Gamma(s) = \int_0^{\infty} e^{-x} x^{s-1} dx, \quad B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx.$$

$\Gamma(s)$  をガンマ函数と、 $B(p, q)$  をベータ函数と呼ぶ。

問題(ガンマ函数のGauss積分型の表示): 次を示せ:

$$\Gamma(s) = 2 \int_0^\infty e^{-y^2} y^{2s-1} dy.$$

解答例: ガンマ函数の積分による定義式において  $x = y^2$  と置換すると,

$$\Gamma(s) = \int_0^\infty e^{-y^2} y^{2s-2} 2y dy = 2 \int_0^\infty e^{-y^2} y^{2s-1} dy. \quad \square$$

注意: この公式より, ガンマ函数は本質的にGauss積分の一般化になっていることがわかる.  $\square$

問題(ベータ函数の別の表示): 次を示せ:

$$B(p, q) = 2 \int_0^{\pi/2} (\cos \theta)^{2p-1} (\sin \theta)^{2q-1} d\theta = \int_0^\infty \frac{t^{p-1}}{(1+t)^{p+q}} dt.$$

ベータ函数のこれらの表示もよく使われる.

解答例:  $B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx$  で  $x = \cos^2 \theta$  と置換すると,

$$dx = -2 \cos \theta \sin \theta d\theta$$

より,

$$B(p, q) = 2 \int_0^{\pi/2} (\cos \theta)^{2p-1} (\sin \theta)^{2q-1} d\theta.$$

$B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx$  で  $x = \frac{t}{1+t} = 1 - \frac{1}{1+t}$  と置換すると,

$$dx = \frac{dt}{1+t}$$

より,

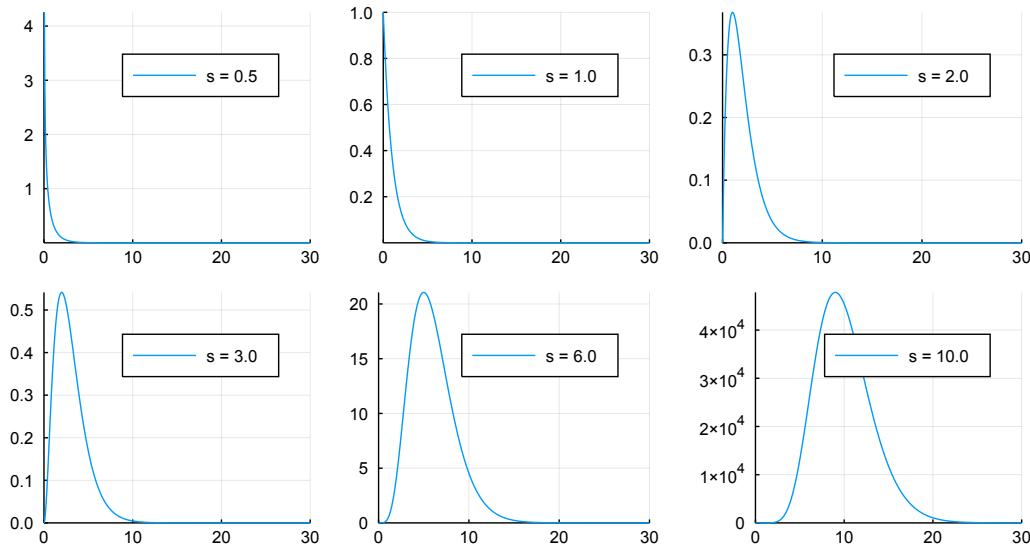
$$B(p, q) = \int_0^\infty \left(\frac{t}{1+t}\right)^{p-1} \left(\frac{1}{1+t}\right)^{q-1} \frac{dt}{(1+t)^2} = \int_0^\infty \frac{t^{p-1}}{(1+t)^{p+q}} dt \quad \square$$

問題: ガンマ函数を定義する積分の被積分函数のグラフを色々な  $s > 0$  について描いてみよ.

解答例: 次のセルを見よ.  $\square$

```
In [3]: 1 f(s,x) = e^(-x)*x^(s-1)
2 x = 0.00:0.05:30.0
3 P = []
4 for s in [1/2, 1, 2, 3, 6, 10]
5     p = plot(x, f.(s,x), label="s = $s", xlims=(0,30))
6     push!(P, p)
7 end
8 plot(size=(750, 400), P...)
```

Out[3]:

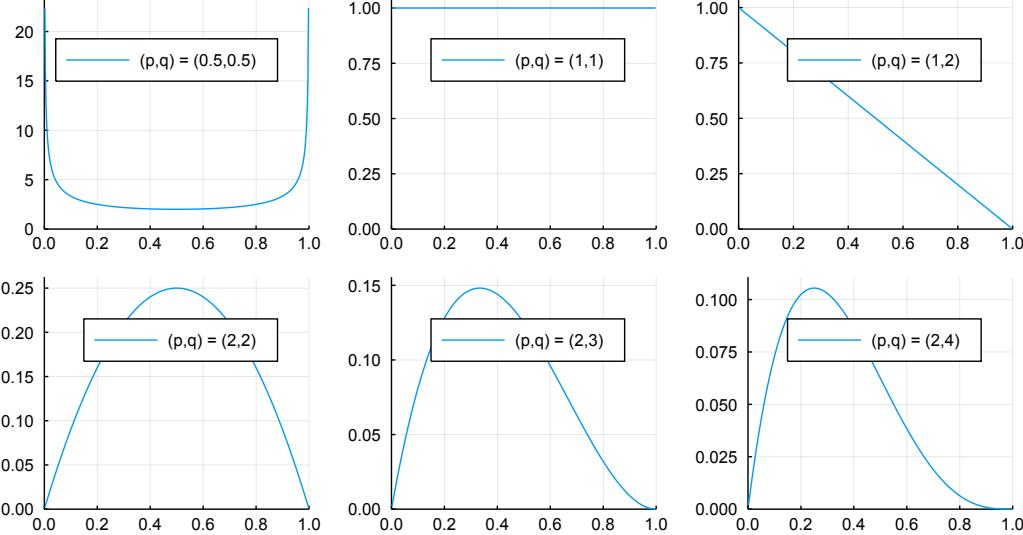


問題: ベータ函数を定義する積分の被積分函数のグラフを色々な  $p, q > 0$  について描いてみよ.

解答例: 次のセルを見よ.  $\square$

```
In [4]: 1 f(p,q,x) = x^(p-1)*(1-x)^(q-1)
2 x = 0.002:0.002:0.998
3 P = []
4 ▼ for (p,q) in [(1/2,1/2), (1,1), (1,2), (2,2), (2,3), (2,4)]
5     y = f.(p,q,x)
6     p = plot(x, y, label="(p,q) = ($p,$q)", xlims=(0,1), ylims=(0,1.05*maximum(y)))
7     push!(P, p)
8 end
9 plot(size=(750, 400), P...)
```

Out[4]:



## 2.2 ガンマ函数の函数の特殊値と函数等式

問題(ガンマ函数の最も簡単な特殊値):  $\Gamma(1) = 1$  と  $\Gamma(1/2) = \sqrt{\pi}$  を示せ.

解答例: 前者は

$$\Gamma(1) = \int_0^\infty e^{-x} dx = [-e^{-x}]_0^\infty = 1.$$

と容易に示される. 後者を示すためには  $\Gamma(1/2)$  がGauss積分  $\int_{-\infty}^\infty e^{-y^2} dy = \sqrt{\pi}$  に等しいことを示せばよい.  $x = y^2$  で置換積分すると,

$$\begin{aligned} \Gamma(1/2) &= \int_0^\infty e^{-x} x^{1/2-1} dx = \int_0^\infty e^{-y^2} \frac{1}{y} 2y dy \\ &= 2 \int_0^\infty e^{-y^2} dy = \int_{-\infty}^\infty e^{-y^2} dy = \sqrt{\pi}. \quad \square \end{aligned}$$

注意: 上の問題の解答より,  $\Gamma(1/2)$  は本質的にGauss積分に等しい. その意味でガンマ函数はGauss積分の一般化になっていると言える.  $\square$

問題(ガンマ函数の函数等式):  $s > 0$  のとき  $\Gamma(s+1) = s\Gamma(s)$  となることを示せ.

解答例: 部分積分を使う.  $s > 0$  と仮定する. このとき

$$\begin{aligned} \Gamma(s+1) &= \int_0^\infty e^{-x} x^s dx = \int_0^\infty (-e^{-x})' x^s dx \\ &= \int_0^\infty e^{-x} (x^s)' dx = \int_0^\infty e^{-x} s x^{s-1} dx = s\Gamma(s). \end{aligned}$$

3つ目の等号で部分積分を行った. そのとき,  $x \searrow 0$  でも  $x \rightarrow \infty$  でも  $e^{-x} x^s \rightarrow 0$  となることを使った( $s > 0$  と仮定したことにも注意せよ). (積分以外の項が消える.)  $\square$

注意: 上の問題の結果を使えば,  $s < 0, s \neq 0, -1, -2, \dots$  のとき  $s+n > 0$  となる整数  $n$  を取れば,

$$\Gamma(s) = \frac{\Gamma(s+n)}{s(s+1)\cdots(s+n-1)}$$

の右辺はwell-definedになるので, この公式によってガンマ函数を  $s < 0, s \neq 0, -1, -2, \dots$  の場合に自然に拡張できる.  $\square$

注意(ガンマ函数は階乗の一般化): 以上の問題の結果より, 非負の整数  $n$  について

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \cdots = n(n-1)\cdots 1 \Gamma(1) = n!.$$

すなわち,  $\Gamma(s+1)$  は階乗  $n!$  の連続変数  $s$  への拡張になっていることがわかる.  $\square$

問題(ガンマ函数の正の半整数での値): 次を示せ: 非負の整数  $k$  に対して

$$\Gamma((2k+1)/2) = \frac{1 \cdot 3 \cdots (2k-1)}{2^k} \sqrt{\pi} = \frac{(2k)!}{2^{2k} k!} \sqrt{\pi}$$

**解答例1:** ガンマ函数の函数等式と  $\Gamma(1/2) = \sqrt{\pi}$  より

$$\begin{aligned}\Gamma\left(\frac{2k+1}{2}\right) &= \frac{2k-1}{2} \Gamma\left(\frac{2k-1}{2}\right) = \frac{2k-1}{2} \frac{2k-3}{2} \Gamma\left(\frac{2k-3}{2}\right) = \dots \\ &= \frac{2k-1}{2} \frac{2k-3}{2} \cdots \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = \frac{1 \cdot 3 \cdots (2k-1)}{2^k} \sqrt{\pi}.\end{aligned}$$

これで示したい公式の1つ目の等号は示せた。2つ目の等号は上方のGauss積分の応用問題で使った方法を使えば同様に示される。□

**解答例2:** ガンマ函数の  $\Gamma(s) = 2 \int_0^\infty e^{-y^2} y^{2s-1} dy$  という表示を使うと、

$$\Gamma((2k+1)/2) = 2 \int_0^\infty e^{-y^2} y^{2k} dy = \int_{-\infty}^\infty e^{-y^2} y^{2k} dy$$

なので、上方のGauss積分の応用問題に関する結果から欲しい公式が得られる。□

### 2.3 ベータ函数とガンマ函数の関係

ベータ函数はガンマ函数によって

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \quad (*)$$

と表わされる。これを証明したい。そのためには

$$\Gamma(p)\Gamma(q) = \int_0^\infty \left( \int_0^\infty e^{-(x+y)} x^{p-1} y^{q-1} dy \right) dx$$

が

$$\Gamma(p+q)B(p, q) = \int_0^\infty e^{-z} z^{p+q-1} dz \int_0^1 t^{p-1} (1-t)^{q-1} dt$$

に等しいことを示せばよい。ガンマ函数とベータ函数の別の表示を使えば右辺も別の形になることに注意せよ。

#### 2.3.1 方法1: 極座標変換を使う方法

この方法は2重積分に関する知識が必要になる。2重積分について知らない人は次の節の別の方法を参照せよ。

$x = X^2, y = Y^2$  と変数変換すると、

$$\Gamma(p)\Gamma(q) = 4 \int_0^\infty \int_0^\infty e^{-(X^2+Y^2)} X^{2p-1} Y^{2q-1} dX dY.$$

さらに  $X = r \cos \theta, Y = r \sin \theta$  と変数変換すると、

$$\begin{aligned}\Gamma(p)\Gamma(q) &= 4 \int_0^{\pi/2} d\theta \int_0^\infty e^{-r^2} (r \cos \theta)^{2p-1} (r \sin \theta)^{2q-1} r dr \\ &= 4 \int_0^{\pi/2} (\cos \theta)^{2p-1} (\sin \theta)^{2q-1} d\theta \int_0^\infty e^{-r^2} r^{2(p+q)-1} dr = B(p, q)\Gamma(p+q).\end{aligned}$$

最後の等号でベータ函数の三角函数を用いた表示とガンマ函数のGauss積分に似た表示を用いた。□

#### 2.3.2 方法2: 置換積分と積分の順序交換のみを使う方法

ガンマ函数とベータ函数のあいだの関係式は1変数の置換積分と積分の順序交換のみを使って証明可能である。条件  $A$  に対して、 $x, y$  が条件  $A$  をみたすとき値が 1 になり、それ以外のときに値が 0 になる  $x, y$  の函数を  $1_A(x, y)$  と書くことにすると、

$$\begin{aligned}\Gamma(p)\Gamma(q) &= \int_0^\infty \left( \int_0^\infty e^{-(x+y)} x^{p-1} y^{q-1} dy \right) dx \\ &= \int_0^\infty \left( \int_x^\infty e^{-z} x^{p-1} (z-x)^{q-1} dz \right) dx \\ &= \int_0^\infty \left( \int_0^\infty 1_{x<z}(x, z) e^{-z} x^{p-1} (z-x)^{q-1} dz \right) dx \\ &= \int_0^\infty \left( \int_0^\infty 1_{x<z}(x, z) e^{-z} x^{p-1} (z-x)^{q-1} dx \right) dz \\ &= \int_0^\infty \left( \int_0^z e^{-z} x^{p-1} (z-x)^{q-1} dx \right) dz \\ &= \int_0^\infty \left( \int_0^1 e^{-z} (zt)^{p-1} (z-zt)^{q-1} z dt \right) dz \\ &= \int_0^\infty e^{-z} z^{p+q-1} dz \int_0^1 t^{p-1} (1-t)^{q-1} dt = \Gamma(p+q)B(p, q).\end{aligned}$$

2つ目の等号で  $y = z - x$  と置換積分し, 4つ目の等号で積分の順序を交換し, 6つ目の等号で  $x = zt$  と置換積分した.

以上の計算は

- 黒木玄, ガンマ分布の中心極限定理とStirlingの公式 (<https://genkuroki.github.io/documents/20160501StirlingFormula.pdf>).

の第7.4節からの引き写しだある.

### 2.3.3 ベータ函数とガンマ函数の関係の応用

問題: ベータ函数とガンマ函数の関係を用いて  $\Gamma(1/2) = \sqrt{\pi}$  を証明せよ.

解答例:

$$\Gamma(1/2)^2 = \frac{\Gamma(1/2)\Gamma(1/2)}{\Gamma(1)} = B(1/2, 1/2) = 2 \int_0^{\pi/2} (\cos \theta)^{2 \cdot 1/2 - 1} (\sin \theta)^{2 \cdot 1/2 - 1} d\theta = 2 \int_0^{\pi/2} d\theta = \pi.$$

1つ目の等号で  $\Gamma(1) = 1$  を使い, 2つ目の等号でベータ函数とガンマ函数の関係を用い, 3つ目の等号でベータ函数の三角函数を用いた表示を使つた. ゆえに  $\Gamma(1/2) = \sqrt{\pi}$ .  $\square$

注意: この問題の解答例はGauss積分の公式の別証明  $\int_{-\infty}^{\infty} e^{-x^2} dx = \Gamma(1/2) = \sqrt{\pi}$  を与える.  $\square$

問題: 次の積分を計算せよ:

$$A = \int_0^1 x^5 (1 - x^2)^{3/2} dx.$$

解答例:  $x = t^{1/2}$  と置換すると  $dx = \frac{1}{2}t^{-1/2} dt$  なので,

$$A = \int_0^1 t^{5/2} (1 - t)^{3/2} \frac{1}{2}t^{-1/2} dt = \frac{1}{2} \int_0^1 t^2 (1 - t)^{3/2} dt = \frac{1}{2} B(3, 5/2) = \frac{\Gamma(3)\Gamma(5/2)}{2\Gamma(3 + 5/2)}.$$

3つ目の等号で  $2 = 3 - 1, 3/2 = 5/2 - 1$  とみなしてからベータ函数の表示を得ていることに注意せよ. このステップでよく間違う.

一般に非負の整数  $n$  について

$$\Gamma(n+1) = n!, \quad \frac{\Gamma(s)}{\Gamma(s+n)} = \frac{1}{s(s+1)\cdots(s+n-1)}$$

なので,

$$\Gamma(3) = 2! = 2, \quad \frac{\Gamma(5/2)}{\Gamma(3 + 5/2)} = \frac{1}{(5/2)(7/2)(9/2)} = \frac{2^3}{5 \cdot 7 \cdot 9}.$$

したがって

$$A = \frac{2}{2} \frac{2^3}{5 \cdot 7 \cdot 9} = \frac{8}{315}. \quad \square$$

```
In [5]: 1 x = symbols("x", real=True)
2 integrate(x^5*(1-x^2)^(Sym(3)/2), (x, 0, 1))
```

```
Out[5]: 8/315
```

問題(Gaussの公式): ベータ函数とガンマ函数の関係を用いて, 次の公式を示せ.

$$\Gamma(s) = \lim_{n \rightarrow \infty} \frac{n^s n!}{s(s+1)\cdots(s+n)}.$$

解答例: 右辺をベータ函数と表示することを考える. 以下では  $n$  は正の整数であるとし,  $s > 0$  と仮定する. ベータ函数とガンマ函数の函数等式および  $\Gamma(n+1) = n!$  より,

$$B(s, n+1) = \frac{\Gamma(s)\Gamma(n+1)}{\Gamma(s+n+1)} = \frac{n!}{s(s+1)\cdots(s+n)}.$$

ゆえに

$$n^s B(s, n+1) = \frac{n^s n!}{s(s+1)\cdots(s+n)}.$$

左辺を  $n \rightarrow \infty$  での極限を取り易い形に変形しよう.  $x = t/n$  と置換することによって,  $n \rightarrow \infty$  のとき

$$\begin{aligned} n^s B(s, n+1) &= n^s \int_0^1 x^{s-1} (1-x)^n dt \\ &= \int_0^n t^{s-1} \left(1 - \frac{t}{n}\right)^n dt \rightarrow \int_0^\infty t^{s-1} e^{-t} dt = \Gamma(s). \end{aligned}$$

以上をまとめると示したい結果が得られる.  $\square$

この問題の続きについては

- 黒木玄, ガンマ分布の中心極限定理とStirlingの公式 (<https://genkuroki.github.io/documents/20160501StirlingFormula.pdf>)

の第8.2節以降を参照せよ.

### 3 Stirlingの公式とLaplaceの方法

一般に数列  $a_n, b_n$  について

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1$$

が成立するとき,

$$a_n \sim b_n$$

と書くこととする.

#### 3.1 Stirlingの公式

Stirlingの(近似)公式:  $n \rightarrow \infty$  のとき,

$$n! \sim n^n e^{-n} \sqrt{2\pi n}.$$

さらに, 両辺の対数を取ることによって,  $n \rightarrow \infty$  のとき,

$$\log n! = n \log n - n + \frac{1}{2} \log n + \log \sqrt{2\pi} + o(1).$$

Stirlingの公式の「物理学的」もしくは「情報理論的」な応用については

- 黒木玄, Kullback-Leibler情報量とSanovの定理 (<https://genkuroki.github.io/documents/20160616KullbackLeibler.pdf>)

の第1節を参照せよ.

**Stirlingの公式の証明:**

$$n! = \Gamma(n+1) = \int_0^\infty e^{-x} x^n dx$$

で  $x = n + \sqrt{n} y$ ,  $y = n(1 + y/\sqrt{n})$  と置換すると,

$$n! = n^n e^{-n} \sqrt{n} \int_{-\sqrt{n}}^\infty e^{-\sqrt{n}y} \left(1 + \frac{y}{\sqrt{n}}\right)^n dy = n^n e^{-n} \sqrt{n} \int_{-\sqrt{n}}^\infty f_n(y) dy.$$

ここで, 被積分函数を  $f_n(y)$  と書いた. そのとき  $n \rightarrow \infty$  で

$$\begin{aligned} \log f_n(y) &= -\sqrt{n}y + n \log \left(1 + \frac{y}{\sqrt{n}}\right) = -\sqrt{n}y + n \left(\frac{y}{\sqrt{n}} - \frac{y^2}{2n} + O\left(\frac{1}{n\sqrt{n}}\right)\right) \\ &= -\frac{y^2}{2} + O\left(\frac{1}{\sqrt{n}}\right) \rightarrow -\frac{y^2}{2}. \end{aligned}$$

すなわち  $f_n(y) \rightarrow e^{-y^2/2}$  となる. ゆえに

$$\frac{n!}{n^n e^{-n} \sqrt{2\pi n}} = \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{n}}^\infty f_n(y) dy \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-y^2/2} dy = 1.$$

最後の等号でGauss積分の公式  $\int_{-\infty}^\infty e^{-y^2/2} dy = \sqrt{\pi}$  を用いた.  $\square$

**Stirlingの公式の証明の解説:** 上の証明のポイントは  $x = n + \sqrt{n}y$  という積分変数変換である. この変数変換の「正体」は  $\Gamma(n+1) = \int_0^\infty e^{-x} x^n dx$  の被積分函数  $\varphi(x) = e^{-x} x^n$  のグラフを描いてみれば見当がつく.

$\psi(x) = \log \varphi(x) = n \log x - x$  の導函数は  $\psi'(x) = n/x - 1$  は  $x$  について単調減少であり,  $x = n$  で 0 になる. ゆえに  $\psi(x) = \log \varphi(x)$  は  $x = n$  で最大になる. そこで  $x = n$  における  $\psi(x) = \log \varphi(x)$  のTaylor展開を求めてみよう.  $\psi''(x) = -n/x^2$ ,  $\psi'''(x) = 2n/x^3$  なので,  $\psi(n) = n \log n - n$ ,  $\psi'(n) = 0$ ,  $\psi''(n) = -1/n$ ,  $\psi'''(n) = 2/n^2$  なので,

$$\psi(x) = n \log n - n - \frac{(x-n)^2}{2n} + \frac{(x-n)^3}{3n^2} + \dots$$

これの2次の項が  $-y^2/2$  になるような変数変換がちょうど  $x = n + \sqrt{n}y$  になっている. これが上の証明で用いた変数変換の「正体」である.  $\square$

問題:  $n = 1, 2, \dots, 10$  について Stirling の公式の相対誤差

$$\frac{n^2 e^{-n} \sqrt{2\pi n}}{n!} - 1$$

を求めよ.

解答例: 以下のセルを参照せよ.  $n = 5$  で相対誤差は2%を切っている. □

```
In [6]: 1 f(n) = factorial(n)
2 g(n) = n^n * exp(-n) * sqrt(2*pi*n)
3 [(n, f(n), g(n), g(n)/f(n)-1) for n in 1:10]
```

```
Out[6]: 10-element Array{Tuple{Int64,Int64,Float64,Float64},1}:
(1, 1, 0.922137, -0.077863)
(2, 2, 1.919, -0.0404978)
(3, 6, 5.83621, -0.0272984)
(4, 24, 23.5062, -0.020576)
(5, 120, 118.019, -0.0165069)
(6, 720, 710.078, -0.0137803)
(7, 5040, 4980.4, -0.0118262)
(8, 40320, 39902.4, -0.0103573)
(9, 362880, 3.59537e5, -0.00921276)
(10, 3628800, 3.5987e6, -0.00829596)
```

参考: 上の計算を見れば,  $n^n e^{-n} \sqrt{2\pi n}$  は  $n!$  よりも微小に小さいことがわかる. その分を補正したより精密な近似式

$$n! = n^n e^{-n} \sqrt{2\pi n} \left( 1 + \frac{1}{12n} + O\left(\frac{1}{n^2}\right) \right)$$

が成立している. (実際には  $O(1/n^2)$  の部分についてもっと詳しいことがわかる.)

1/(12n) で補正した近似式の相対誤差は  $n = 1$  ですでに0.1%程度と非常に小さくなる. 次のセルを見よ. □

```
In [7]: 1 f(n) = factorial(n)
2 g1(n) = n^n * exp(-n) * sqrt(2*pi*n) * (1+1/(12n))
3 [(n, f(n), g1(n), g1(n)/f(n)-1) for n in 1:10]
```

```
Out[7]: 10-element Array{Tuple{Int64,Int64,Float64,Float64},1}:
(1, 1, 0.998982, -0.00101824)
(2, 2, 1.99896, -0.000518567)
(3, 6, 5.99833, -0.000278913)
(4, 24, 23.9959, -0.00017137)
(5, 120, 119.986, -0.000115383)
(6, 720, 719.94, -8.28033e-5)
(7, 5040, 5039.69, -6.22504e-5)
(8, 40320, 40318.0, -4.84771e-5)
(9, 362880, 3.62866e5, -3.88063e-5)
(10, 3628800, 3.62868e6, -3.17601e-5)
```

問題(Wallisの公式): Stirlingの公式を用いて次を示せ:

$$\frac{1}{2^{2n}} \binom{2n}{n} \sim \frac{1}{\sqrt{\pi n}}.$$

解答例:

$$\frac{1}{2^{2n}} \binom{2n}{n} = \frac{(2n)!}{2^{2n}(n!)^2} \sim \frac{(2n)^{2n} e^{-2n} \sqrt{4\pi n}}{2^{2n} n^{2n} e^{-2n} 2\pi n} = \frac{1}{\sqrt{\pi n}}. \quad \square$$

注意: この形のWallisの公式は1次元の単純ランダムウォーク(コイントス)の逆正弦法則に関係している.

- 黒木玄, [単純ランダムウォークの逆正弦法則](https://genkuroki.github.io/documents/#2016-11-02) ([手書きのノートのPDF](https://genkuroki.github.io/documents/20161102ArcSineLawForSimpleRandomWalks.pdf))

を参照せよ. 特に手書きのノートのPDFファイルの12頁以降にまとまつた解説がある. 1次元の単純ランダムウォークの場合には高校数学レベルの組み合わせ論的な議論とWallisの公式から逆正弦法則を導くことができる. 1次元の一般ランダムウォークの場合にはTauber型定理を使ってWallisの公式に対応する漸近挙動を証明することになる. □

注意: 次の公式もWallisの公式と呼ばれている:

$$\prod_{n=1}^{\infty} \frac{2n \cdot 2n}{(2n-1)(2n+1)} = \frac{\pi}{2}.$$

こちらの公式の証明については

- 黒木玄, [ガンマ分布の中心極限定理とStirlingの公式](https://genkuroki.github.io/documents/20160501StirlingFormula.pdf) (<https://genkuroki.github.io/documents/20160501StirlingFormula.pdf>)

の第8.4節以降を参照せよ. □

Stirlingの公式の重要な応用については

- 黒木玄, 11 Kullback-Leibler情報量 (<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/11%20Kullback-Leibler%20information.ipynb>)

も参照せよ. 「Stirlingの公式」とその応用としての「KL情報量に関するSanovの定理」についてはできるだけ早く理解しておいた方がよい. □

### 3.2 Laplaceの方法

**Laplaceの方法:** Stirlingの公式の証明の解説のようにして見付かる変数変換はより一般の場合に非常に有用である. 以下では  $\int_{-\infty}^{\infty}$  や  $\int_0^{\infty}$  を単に  $\int$  と書くことにし,

$$Z_n = \int e^{-nf(x)} g(x) dx$$

とおく. ただし,  $f(x)$  は実数値函数で唯一つの最小値を持ち, そこで,

$$f(x) = f(x_0) + \frac{a}{2}(x - x_0)^2 + O((x - x_0)^3), \quad a = f''(x_0) > 0$$

とTaylor展開されていると仮定するし, さらに, 0 以上の値を持つ実数値函数  $g(x)$  は積分  $Z_n$  がうまく定義されるような適当な条件を満たしていると仮定し,  $x_0$  の近傍で  $g(x) > 0$  を満たしていると仮定する. (ここで,  $x_0$  の近傍で  $g(x) > 0$  が成立しているとは, ある  $\delta > 0$  が存在して,  $|x - x_0| < \delta$  ならば  $g(x) > 0$  となることである.) このとき,

$$Z_n = e^{-nf(x_0)} \int \exp\left(-n\left(\frac{a}{2}(x - x_0)^2 + O((x - x_0)^3)\right)\right) g(x) dx.$$

$x = x_0 + y/\sqrt{n}$  と変数変換すると

$$Z_n = \frac{e^{-nf(x_0)}}{\sqrt{n}} \int \exp\left(-\frac{a}{2}y^2 + O\left(\frac{1}{\sqrt{n}}\right)\right) g\left(x_0 + \frac{y}{\sqrt{n}}\right) dy.$$

そして,  $n \rightarrow \infty$  で

$$\int \exp\left(-\frac{a}{2}y^2 + O\left(\frac{1}{\sqrt{n}}\right)\right) g\left(x_0 + \frac{y}{\sqrt{n}}\right) dy \rightarrow \int \exp\left(-\frac{a}{2}y^2\right) g(x_0) dy = \sqrt{\frac{2\pi}{a}} g(x_0).$$

$a = f''(x_0)$  とおいたことを思い出しながら, 以上をまとめると,  $n \rightarrow \infty$  で

$$Z_n \sim \frac{e^{-nf(x_0)}}{\sqrt{n}} \sqrt{\frac{2\pi}{f''(x_0)}} g(x_0).$$

すなわち,

$$-\log Z_n = nf(x_0) + \frac{1}{2}\log n - \log\left(\sqrt{\frac{2\pi}{f''(x_0)}} g(x_0)\right) + o(1).$$

$Z_n$  の  $n \rightarrow \infty$  における漸近挙動を調べるために以上の方法を**Laplaceの方法**(Laplace's method)と呼ぶ. □

**問題(Stirlingの公式):**  $n! = \int_0^{\infty} e^{-t} t^n dt$  にLaplaceの方法を適用して, Stirlingの公式を導出せよ.

**解答例:** 積分変数を  $t = nx$  で置換すると,

$$n! = \int_0^{\infty} e^{-t+n\log t} dt = n^{n+1} \int_0^{\infty} e^{-n(x-\log x)} dx.$$

$f(x) = x - \log x$ ,  $g(x) = 1$  とおく.  $f'(x) = 1 - 1/x$ ,  $f''(x) = 1/x^2$  なので  $f(x)$  は  $x_0 = 1$  で最小になり,  $f(1) = f''(1) = 1$  となる. ゆえに, それらにLaplaceの方法を適用すると,

$$n! \sim n^{n+1} \frac{e^{-n}}{\sqrt{n}} \sqrt{2\pi} = n^n e^{-n} \sqrt{2\pi n}. \quad \square$$

Laplaceの方法は本質的にGauss積分の応用である. Gauss積分をガンマ函数に置き換えることによって得られる一般化されたLaplaceの方法の素描については

- 黒木玄, 一般化されたLaplaceの方法 (<https://genkuroki.github.io/documents/20161014GeneralizedLaplace.pdf>).

を参照せよ. 一般化されたLaplaceの方法は

- 渡辺澄夫, ベイズ統計の理論と方法 (<https://www.amazon.co.jp/dp/4339024627>), 2012

の第4章の主結果であるベイズ統計における自由エネルギーの

$$F_n = -\log Z_n = nS + \lambda \log n - (m-1) \log \log n + O(1)$$

の形の漸近挙動を導く議論を初等化するために役に立つ. 特異点解消は本質的に不可避だが, この形の漸近挙動だけが欲しいのであればゼータ函数を用いた精密な議論は必要ない.

### 3.3 Laplaceの方法の弱形

**Laplaceの方法の弱形:** Laplaceの方法が使える状況では,

$$Z_n = \int e^{-nf(x)} g(x) dx$$

について、特に、 $n \rightarrow \infty$  のとき、

$$-\frac{1}{n} \log Z_n \rightarrow f(x_0) = \min_{a \leq x \leq b} f(x), \quad \text{i.e.} \quad Z_n = \int e^{-nf(x)} g(x) dx = \exp(-n \min_{a \leq x \leq b} f(x) + o(n))$$

が成立している。この結論を**Laplaceの方法の弱形**と呼ぶことにする。Laplaceの方法のような精密な形でなくても、こちらの弱形だけで用が足りることは結構多い。□

**問題(Laplaceの方法の弱形が明瞭に成立する場合):** 閉区間  $[a, b]$  上の実数値連続函数  $f(x)$  と 0 以上の値を持つ実数値函数  $g(x)$  は、  
 $f(x_0) = \min_{a \leq x \leq b} f(x)$  を満たすある  $x_0 \in [a, b]$  の近傍で  $g(x) > 0$  を満たしていると仮定する。このとき、 $n \rightarrow \infty$  において、

$$\int_a^b e^{-nf(x)} g(x) dx = \exp\left(-n \min_{a \leq x \leq b} f(x) + o(n)\right)$$

が成立していることを示せ。すなわち、 $n \rightarrow \infty$  のとき、

$$-\frac{1}{n} \log \int_a^b e^{-nf(x)} g(x) dx \rightarrow \min_{a \leq x \leq b} f(x)$$

が成立していることを示せ。

**解答例:**  $f_0(x) = f(x) - \min_{a \leq x \leq b} f(x)$  とおくと、 $f_0(x)$  の最小値は 0 になり、

$$-\frac{1}{n} \log \int_a^b e^{-nf(x)} g(x) dx = \min_{a \leq x \leq b} f(x) - \frac{1}{n} \log \int_a^b e^{-nf_0(x)} g(x) dx$$

なので、 $n \rightarrow \infty$  のとき

$$-\frac{1}{n} \log \int_a^b e^{-nf_0(x)} g(x) dx \rightarrow 0$$

となることを示せばよい。

$\varepsilon > 0$  を任意に取って固定し、 $A = \{x \in [a, b] \mid f_0(x) \leq \varepsilon\}$  とおき、その  $[a, b]$  での補集合を  $A^c$  と書き、

$$Z_n = \int_a^b e^{-nf_0(x)} g(x) dx = I_n + J_n, \quad I_n = \int_A e^{-nf_0(x)} g(x) dx, \quad J_n = \int_{A^c} e^{-nf_0(x)} g(x) dx.$$

とおく。 $n \rightarrow \infty$  のとき  $-\frac{1}{n} \log Z_n \rightarrow 0$  となることを示したい。

$x \in A$  について  $\varepsilon \geq f_0(x) \geq 0$  なので、 $e^{-ne} \leq e^{-nf_0(x)} \leq 1$  となるので、

$$e^{-ne} \int_A g(x) dx \leq I_n \leq \int_A g(x) dx.$$

$f(x_0) = \min_{a \leq x \leq b} f(x)$  を満たすある  $x_0 \in [a, b]$  の近傍で  $g(x) > 0$  となっていると仮定したことより、 $\int_A g(x) dx > 0$  となることにも注意せよ。

$x \in A^c$  について  $f_0(x) > \varepsilon$  なので、 $0 < e^{-nf_0(x)} < e^{-ne}$  となるので、

$$0 \leq J_n \leq e^{-ne} \int_{A^c} g(x) dx.$$

以上をまとめると、

$$e^{-ne} \int_A g(x) dx \leq Z_n \leq \int_A g(x) dx + e^{-ne} \int_{A^c} g(x) dx.$$

この全体の対数を取って  $-1/n$  倍すると、

$$\varepsilon - \frac{1}{n} \log \int_A g(x) dx \geq -\frac{1}{n} \log Z_n \geq -\frac{1}{n} \log \left( \int_A g(x) dx + e^{-ne} \int_{A^c} g(x) dx \right).$$

したがって、 $n \rightarrow 0$  とすると、

$$\varepsilon \geq \liminf_{n \rightarrow \infty} \left( -\frac{1}{n} \log Z_n \right) \geq \limsup_{n \rightarrow \infty} \left( -\frac{1}{n} \log Z_n \right) \geq 0.$$

$\varepsilon > 0$  は幾らでも小さくできるので、下極限と上極限が等しくなることがわかり、

$$\lim_{n \rightarrow \infty} \left( -\frac{1}{n} \log Z_n \right) = 0$$

が得られる.  $\square$

In [ ]:

1

# Kullback-Leibler情報量 ¶

黒木玄

2018-06-24

- Copyright 2018 Gen Kuroki
- License: MIT <https://opensource.org/licenses/MIT> (<https://opensource.org/licenses/MIT>)

このファイルは次の場所できれいに閲覧できる:

- <http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/11%20Kullback-Leibler%20information.ipynb> (<http://nbviewer.jupyter.org/github/genkuroki/Calculus/blob/master/11%20Kullback-Leibler%20information.ipynb>)

このファイルは [Julia Box \(https://juliabox.com\)](https://juliabox.com) で利用できる。

自分のパソコンにJulia言語 (<https://julialang.org/>) をインストールしたい場合には

- [WindowsへのJulia言語のインストール \(http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f\)](http://nbviewer.jupyter.org/gist/genkuroki/81de23edcae631a995e19a2ecf946a4f)

を参照せよ。

論理的に完璧な説明をするつもりはない。細部のいい加減な部分は自分で訂正・修正せよ。

## Table of Contents

### [1 Kullback-Leibler情報量の基本性質](#)

[1.1 KL情報量の定義](#)

[1.2 KL情報量の正値性](#)

[1.3 KL情報量の凸性](#)

[1.4 KL情報量が  \$L^1\$  ノルムを上から押さえること](#)

[1.5 KL情報量の使われ方](#)

### [2 多項分布のSanovの定理](#)

[2.1 Sanovの定理とその証明](#)

[2.2 Sanovの定理の使い方1](#)

[2.3 二項分布の場合](#)

[2.4 Sanovの定理の使い方2](#)

[2.5 正規分布の場合](#)

[2.6 多項分布の中心極限定理](#)

In [1]:

```
1 using Plots
2 gr(); ENV["PLOTS_TEST"] = "true"
3 #clibrary(:colorcet)
4 clibrary(:misc)
5
6 ▶ function pngplot(P...; kwargs...)
7     sleep(0.1)
8     pngfile = tempname() * ".png"
9     savefig(plot(P...; kwargs...), pngfile)
10    showimg("image/png", pngfile)
11 end
12 pngplot(; kwargs...) = pngplot(plot!(; kwargs...))
13
14 ▶ showimg(mime, fn) = open(fn) do f
15     base64 = base64encode(f)
16     display("text/html", """""")
17 end
18
19 using SymPy
20 #sympy[:init_printing](order="lex") # default
21 #sympy[:init_printing](order="rev-lex")
22
23 using SpecialFunctions
24 using QuadGK
```

## 1 Kullback-Leibler情報量の基本性質

### 1.1 KL情報量の定義

一般の確率分布についても定義可能だが、このノートでは簡単のため、有限集合

$$\Omega = \{1, 2, \dots, r\}$$

の上の確率分布のみを扱う.  $p = (p_1, \dots, p_r)$  が  $\Omega = \{1, 2, \dots, r\}$  上の確率分布であるとは

$$p_i \geq 0 \quad (i = 1, 2, \dots, r) \quad \sum_{i=1}^r p_i = 1$$

が成立していることであると定める.

$\Omega = \{1, 2, \dots, r\}$  上の確率分布  $q, p$  に対して,  $(q, p)$  の Kullback-Leibler 情報量  $D(q||p)$  を

$$D(q||p) = \sum_{i=1}^r q_i \log \frac{q_i}{p_i} = \left( - \sum_{i=1}^r q_i \log p_i \right) - \left( - \sum_{i=1}^r q_i \log q_i \right)$$

と定める. ただし,  $x \log y$  は  $x = 0$  のとき 0 になり,  $\log 0 = -\infty$  であると約束しておく. Kullback-Leibler 情報量の  $-1$  倍を相対エントロピーと呼ぶ.

記述の簡単のため以下では Kullback-Leibler を KL と略記することにする. さらに,  $q, p$  は  $\Omega = \{1, 2, \dots, r\}$  上の確率分布であると仮定する.

## 1.2 KL情報量の正値性

KL情報量の正値性:  $D(q||p) \geq 0$ .

証明:  $x \geq 0$  に対して  $f(x) = x \log x$  とおく. ただし,  $f(0) = \lim_{x \searrow 0} f(x) = 0$  と約束しておく.  $f'(x) = \log x + 1, f''(x) = 1/x > 0$  なので  $f(x)$  は下に凸な函数である. 期待値汎函数  $g(x) \mapsto E[g(x)]$  を

$$E[g(x)] = \sum_{i=1}^r p_i g(q_i/p_i)$$

と定めることができる. ゆえに Jensen の不等式より,

$$E[f(x)] \geq f(E[x]).$$

そして,

$$E[f(x)] = \sum_{i=1}^r p_i \frac{q_i}{p_i} \log \frac{q_i}{p_i} = D(q||p), \quad E[x] = \sum_{i=1}^r p_i \frac{q_i}{p_i} = \sum_{i=1}^r q_i = 1, \quad f(E[x]) = f(1) = 0$$

なので  $D(q||p) \geq 0$  を得る.  $\square$

## 1.3 KL情報量の凸性

KL情報量の凸性:  $D(q||p)$  は  $q$  の函数として下に凸である.

証明:  $\Omega$  上の確率分布  $q, q'$  と  $0 < t < 1$  が任意に与えられているとする.  $f(x) = x \log x$  とおくと,  $f(x)$  は下に凸な函数である. ゆえに

$$\begin{aligned} D((1-t)q + tq' || p) &= \sum_{i=1}^r p_i f\left(\frac{(1-t)q_i + tq'_i}{p_i}\right) \\ &\leq \sum_{i=1}^r p_i \left( (1-t)f\left(\frac{q_i}{p_i}\right) + t f\left(\frac{q'_i}{p_i}\right) \right) \\ &= (1-t)D(q||p) + tD(q'||p). \end{aligned} \quad \square$$

注意:  $f(x)$  は強い意味で下に凸なので  $D(q||p)$  も  $q$  の函数として強い意味で下に凸になる. このことより,  $q$  の函数としての  $D(q||p)$  の最小値を与える  $q = p$  とは異なる  $q \neq p$  において  $D(q||p) > 0$  となることがわかる. より精密な結果については次の節を見よ.  $\square$

## 1.4 KL情報量が $L^1$ ノルムを上から押さえること

$q = p$  ならば  $q_i \log(q_i/p_i) = 0$  より,  $D(q||p) = 0$  となる.

- 黒木玄, [Kullback-Leibler 情報量と Sanov の定理 \(<https://genkuroki.github.io/documents/20160616KullbackLeibler.pdf>\)](https://genkuroki.github.io/documents/20160616KullbackLeibler.pdf)

の第5.3節で次の不等式が証明されている:

$$D(q||p) \geq \frac{1}{2} \left( \sum_{i=1}^r |q_i - p_i| \right)^2.$$

これより,  $q \neq p$  ならば  $D(q||p) > 0$  となることがわかる. 特に  $D(q||p) = 0$  と  $p = q$  が同値であることがわかる.

さらに, 確率分布の列  $q^{(1)}, q^{(2)}, \dots$  について,  $k \rightarrow \infty$  で  $D(q^{(k)}||p) \rightarrow 0$  ならば  $q^{(k)} \rightarrow p$  となり, 確率分布の列  $p^{(1)}, p^{(2)}, \dots$  について,  $k \rightarrow \infty$  で  $D(q||p^{(k)}) \rightarrow 0$  ならば  $p^{(k)} \rightarrow q$  となることもわかる.  $\square$

## 1.5 KL情報量の使われ方

KL情報量  $D(q||p)$  は確率分布  $q$  と確率分布  $p$  の違いの大きさを表す指標としてよく使われる。一般に  $D(q||p) \neq D(p||q)$  であるが、そうなることには必然性がある。「KL情報量の意味」は下の方で説明するSanovの定理を見ればわかる。Sanovの定理はStirlingの公式を知つていれば容易に証明可能である。Stirlingの公式を理解しておけば「情報理論」「統計力学」「統計学」などなどで基本的なSanovの定理をすぐに理解できるようになる。

- 階乗に関するStirlingの近似公式 → KL情報量に関するSanovの定理

という経路をできるだけ早く踏破しておいた方が数学の応用について楽に考えることができるようになると思う。

統計学へのSanovの定理の応用については赤池弘次氏による以下の2つの論説を参照せよ：

- 赤池弘次、エントロピーとモデルの尤度(〈講座〉物理学周辺の確率統計)  
([https://www.jstage.jst.go.jp/article/butsuri1946/35/7/35\\_7\\_608/\\_article/-char/ja/](https://www.jstage.jst.go.jp/article/butsuri1946/35/7/35_7_608/_article/-char/ja/))、日本物理学会誌、35巻(1980)7号、608-614
- 赤池弘次、統計的推論のパラダイムの変遷について([https://ismrepo.ism.ac.jp/?action=pages\\_view\\_main&active\\_action=repository\\_view\\_main\\_item\\_detail&item\\_id=32568&item\\_no=1&page\\_id=13&block\\_id=21](https://ismrepo.ism.ac.jp/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=32568&item_no=1&page_id=13&block_id=21))、統計数理研究所彙報、27巻、1号、5-12、1980-03

さらに

- 黒木玄、Kullback-Leibler情報量とSanovの定理(<https://genkuroki.github.io/documents/20160616KullbackLeibler.pdf>)

も参照して欲しい。「Kullback-Leibler情報量とSanovの定理(<https://genkuroki.github.io/documents/20160616KullbackLeibler.pdf>)」の第10節には平均符号長とKL情報量の関係も解説してある。「情報量」という呼び方の由来はそちらを見た方が分かり易いかもしれない。□

## 2 多項分布のSanovの定理

1回の試行で  $\Omega = \{1, 2, \dots, r\}$  から  $i$  を取り出す確率は  $p_i$  であるとし、 $k_i$  は非負の整数であり、 $k_1 + k_2 + \dots + k_r = n$  を満たしているとする。このとき  $n$  回の独立試行で  $1, 2, \dots, r$  のそれぞれが  $k_1, k_2, \dots, k_r$  回取り出される確率は

$$\frac{n!}{k_1!k_2!\dots k_r!} p_1^{k_1} p_2^{k_2} \dots p_r^{k_r} \quad (*)$$

になる。この式における階乗の因子はそれぞれ  $k_i$  個の  $i$  達を全部で  $n$  個順番に並べる方法の総数に等しいので、多項係数と呼ばれる。 $(k_1, \dots, k_r)$  全体の集合上の確率分布 (\*) は多項分布と呼ばれる。

### 2.1 Sanovの定理とその証明

**Sanovの定理:**  $n \rightarrow \infty$  で  $k_i$  達が  $k_i/n \sim q_i$  ( $i = 1, \dots, r$ ) のように振る舞うならば、 $n \rightarrow \infty$  で

$$\frac{n!}{k_1!k_2!\dots k_r!} p_1^{k_1} p_2^{k_2} \dots p_r^{k_r} = \exp(-nD(q||p) + o(n)).$$

すなわち

$$-\log\left(\frac{n!}{k_1!k_2!\dots k_r!} p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}\right) = nD(q||p) + o(n).$$

**証明:**  $k_i/n \sim q_i$  という仮定より、 $k_i = nq_i + o(n)$  である。 $q_i$  は固定されているので、 $n$  が大きくなると  $k_i$  も  $O(n)$  のオーダーで大きくなる。だから、Stirlingの公式

$$n! \sim n^n e^{-n} \sqrt{2\pi n}, \quad k_i! \sim k_i^{k_i} e^{-k_i} \sqrt{2\pi k_i}$$

を使用できる。ゆえに、 $k_1 + \dots + k_r = n$  を使うと、

$$\frac{n!}{k_1!k_2!\dots k_r!} \sim \frac{n^n e^{-n} \sqrt{2\pi n}}{k_1^{k_1} e^{-k_1} \sqrt{2\pi k_1} \dots k_r^{k_r} e^{-k_r} \sqrt{2\pi k_r}} = \left( \prod_{i=1}^r \left( \frac{k_i}{n} \right)^{k_i} \sqrt{(2\pi n)^{r-1} \prod_{i=1}^r \frac{k_i}{n}} \right)^{-1}.$$

したがって、

$$\begin{aligned} \frac{n!}{k_1!k_2!\dots k_r!} p_1^{k_1} p_2^{k_2} \dots p_r^{k_r} &\sim \left( \prod_{i=1}^r \left( \frac{k_i/n}{p_i} \right)^{k_i} \sqrt{(2\pi n)^{r-1} \prod_{i=1}^r \frac{k_i}{n}} \right)^{-1} \\ &= \exp\left( -\sum_{i=1}^r k_i \log \frac{k_i/n}{p_i} - (r-1) \log n - \log \sqrt{(2\pi)^{r-1} \prod_{i=1}^r \frac{k_i}{n}} \right). \end{aligned}$$

これに  $k_i/n = q_i + o(1)$ ,  $k_i = nq_i + o(n)$  を代入すれば示したい結果が得られる。 $\log n$  や定数の項は  $n$  で割って  $n \rightarrow \infty$  とすると 0 に収束するので、それらは  $o(n)$  の項にまとめられることに注意せよ。さらにKL情報量の定義が

$$D(q||p) = \sum_{i=1}^r q_i \log \frac{q_i}{p_i}$$

であったことも思い出して計算せよ。□

要するに多項分布にStirlingの近似公式と  $k_i = nq_i + o(n)$  を単純に代入するだけで、Sanovの定理が得られる。

### 2.2 Sanovの定理の使い方1

有限集合  $\Omega = \{1, 2, \dots, r\}$  から確率  $p_i$  で  $i$  を選ぶ独立試行を  $n$  回行った結果,  $i$  が  $k_i$  回取り出されたとき,  $\Omega = \{1, 2, \dots, r\}$  上の確率分布  $(k_1/n, \dots, k_r/n)$  を経験分布と呼ぶ。Sanovの定理における  $k_i/n \sim q_i$  ( $n \rightarrow \infty$ ) という条件は,  $n$  を大きくしたとき経験分布が確率分布  $q$  に近いことを意味している。だから, Sanovの定理は  $n$  が大きなとき

- (確率分布  $p$  のもとでの独立試行の経験分布が確率分布  $q$  に近い確率)  $\sim \exp(-nD(q||p) + o(n))$ .

が成立することを意味している。

$o(n)$  は  $n$  で割ると 0 に収束する量なので,  $n$  が十分に大きなとき  $-nD(q||p)$  と比較すると無視できるほど小さな量になる。ゆえに,  $n$  を十分に大きくすると, 異なる  $q$  ごとの「経験分布が  $q$  に近くなる確率」の大小関係をKL情報量  $D(q||p)$  の大小関係で測って良い状況になる。 $D(q||p)$  が小さな  $q$  ほど相対的に経験分布が近くなる確率は高くなる。

経験分布が確率分布  $q^*$  に近くなる確率で経験分布が確率分布  $q$  に近くなる確率を割った結果は

$$\exp(-n(D(q||p) - D(q^*||p)) + o(n))$$

のように振る舞う。 $D(q||p) > D(q^*||p)$  ならばこの確率の商は  $n \rightarrow \infty$  で 0 に収束する。そのとき, 経験分布が  $q$  に近くなる確率は経験分布が  $q^*$  に近くなる確率と比較すると相対的に無視できるくらい小さくなる。

確率分布の集合  $A$  が与えられているとする。ただし,  $A$  の内部の閉包が  $A$  閉包に一致するような「良い集合」のみを考える。例えば  $r = 2$  (二項分布のとき)で  $0 \leq a < b \leq 1$  のとき  $A = \{(q_1, 1-q_1) \mid a \leq q_1 \leq b\}$  はそのような「良い集合」の例になっている。(「内部」「閉包」の意味を知らない人はこの例を思い浮かべておけばよい。)

$q$  が  $A$  上を動くとき,  $D(q||p)$  を最小にする  $q \in A$  を  $q^*$  と書く。

このとき, 経験分布が  $A$  に入るという条件のもとでの条件付き確率分布を考えると,  $q^*$  以外の  $q \in A$  に経験分布が近くなる確率は  $n \rightarrow \infty$  で 0 に近付く。これは、経験分布が  $A$  に入るという条件のもとでの条件付き確率分布が  $n \rightarrow \infty$  で  $q^*$  に台を持つデルタ分布に近付くことを意味している。

## 2.3 二項分布の場合

$r = 2$  の二項分布の場合を考える。このとき,  $q_1 = q$ ,  $q_2 = 1 - q$ ,  $p_1 = p$ ,  $p_2 = 1 - p$  と書くと,

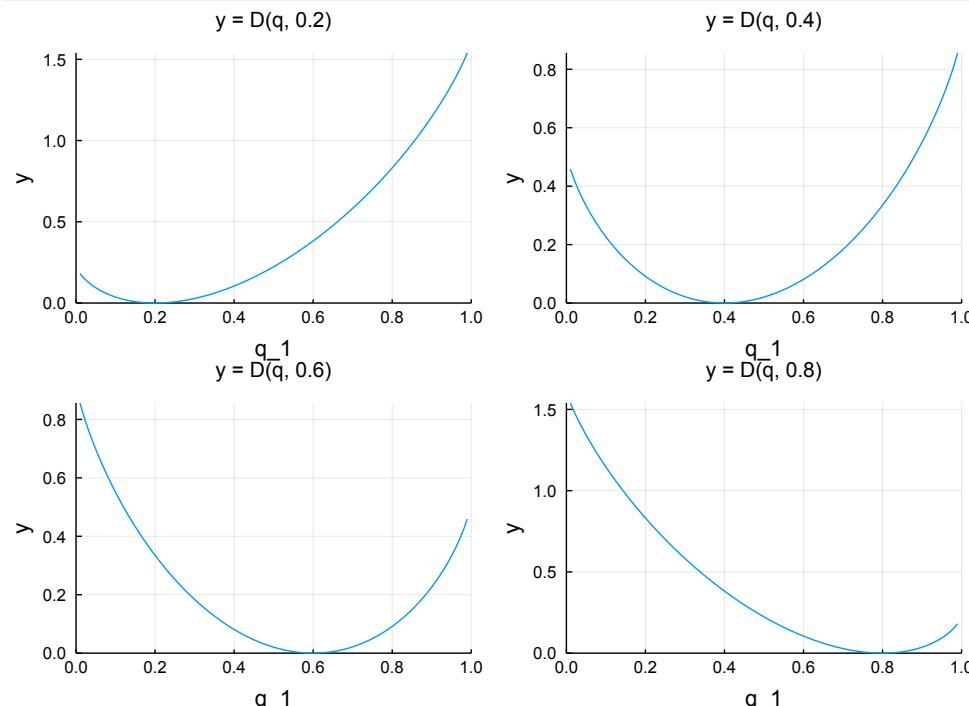
$$D(q||p) = q \log \frac{q}{p} + (1-q) \log \frac{1-q}{1-p}.$$

問題: この函数は  $q = p$  で最小値 0 になることを自分で確認してみよ。□

次のセルのプロットを見ればわかるように  $D(q||p)$  は  $q$  の函数として  $q = p$  で最小になり, 下に凸な函数になる。

```
In [2]:
1  D(q,p) = q*log(q/p) + (1-q)*log((1-q)/(1-p))
2  q = 0:0.01:1
3  PP = []
4  for p in [0.2, 0.4, 0.6, 0.8]
5      P = plot(Legend=false)
6      plot!(title="y = D(q, $(round(p,3)))", titlefontsize=10)
7      plot!(xlabel="q_1", ylabel="y")
8      plot!(q, D.(q,p))
9      push!(PP, P)
10 end
11 plot(PP..., size=(700, 500))
```

Out[2]:



**問題:** 実数  $E_1, \dots, E_N$  について,  $n \rightarrow \infty$  のとき,

$$-\frac{1}{n} \log \sum_{i=1}^N e^{-nE_i} \rightarrow \min\{E_1, \dots, E_N\} \quad (*)$$

となることを示せ.

**解答例:**  $E_1 = \dots = E_k < E_{k+1} \leq \dots \leq E_N$  と仮定してよい. このとき,

$$-\frac{1}{n} \log \sum_{i=1}^N e^{-nE_i} = -\frac{1}{n} \log \left( e^{-nE_1} \left( k + \sum_{i=k+1}^n e^{-n(E_i-E_1)} \right) \right) = E_1 - \frac{1}{n} \log \left( k + \sum_{i=k+1}^n e^{-n(E_i-E_1)} \right).$$

これより, (\*)が成立することがわかる.  $\square$

**注意:** コンピューターの計算では[logsumexp](https://www.google.co.jp/search?q=logsumexp) (<https://www.google.co.jp/search?q=logsumexp>) という函数を適切に定義して使うことがよくある.  $\square$

確率  $p$  で当たるくじを引く独立試行を  $n$  回繰り返したとき, 当たりが  $k$  回出る確率は

$$\binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

になる.  $0 \leq a < b \leq 1$  であるとする. このとき当たりが出た回数の割合  $k/n$  が  $a$  以上  $b$  以下になる確率は

$$\sum_{a \leq k/n \leq b} \binom{n}{k} p^k (1-p)^{n-k}$$

になる. この確率の対数の  $-1/n$  倍

$$-\frac{1}{n} \log \sum_{a \leq k \leq b} \binom{n}{k} p^k (1-p)^{n-k}$$

が  $n \rightarrow \infty$  でどうなるかについて考えよう.

Sanovの定理より,  $k/n \sim q$  のとき,

$$\binom{n}{k} p^k (1-p)^{n-k} = \exp(-nD(q||p) + o(n)).$$

したがって, 上の問題の結果の一般化によって,  $n \rightarrow \infty$  で

$$-\frac{1}{n} \log \sum_{a \leq k \leq b} \binom{n}{k} p^k (1-p)^{n-k} \rightarrow \max_{a \leq q \leq b} D(q||p) = \begin{cases} 0 & (a \leq p \leq b) \\ D(b||p) & (b < p) \\ D(a||p) & (p < a) \end{cases}$$

となると考えられる. 要するに  $a$  以上  $b$  以下の  $k/n$  で  $p$  に最も近いものが生じる確率の寄与だけが, 効いて来るということである. (前節の説明も再読せよ.)

以上の議論は少々厳密さに欠けた荒い議論だが, 結論は正しい. 正確な議論の仕方については

- 黒木玄, [Kullback-Leibler情報量とSanovの定理](https://genkuroki.github.io/documents/20160616KullbackLeibler.pdf) (<https://genkuroki.github.io/documents/20160616KullbackLeibler.pdf>)

の第3節を参照せよ.

## 2.4 Sanovの定理の使い方2

KL情報量  $D(q||p)$  は, Sanovの定理によって, 確率分布  $p$  の試行を独立に  $n$  回繰り返したとき, 経験分布として  $q$  に近い分布が出て来る確率が  $\exp(-nD(q||p) + o(n))$  になるという性質を持っているのであった.

これより, 確率分布  $p$  の独立試行によって確率分布  $q$  をシミュレートしたときの, 予測誤差の大きさが  $D(q||p)$  で測られることを意味していると考えることができる.

$q = p$  と  $D(q||p) = 0$  が同値であることも, この解釈をサポートしている.

さらに, 最も簡単な次の場合について考えてみよう. 確率  $q$  で雨が降るときに確率 1 で雨が降ると予報を出すことを  $n$  回繰り返すときに, 予報の出し方が間違っていたことがばれずに済む確率は  $q^n = \exp(-n(-\log q))$  である. そして, この場合のKL情報量は

$$D(q||1) = 1 \log \frac{1}{q} + (1-1) \log \frac{1-1}{1-q} = -\log q$$

となるので(二項分布の場合の記号法をそのまま使った),  $q^n = \exp(-nD(q||1))$  となる.  $D(q||1)$  は予報の出し方が間違っていたことがばれる速さを表していると考えられる. この極端で単純な場合も, KL情報量  $D(q||p)$  を確率分布  $p$  による確率分布  $q$  のシミュレーションの予測誤差を表しているという解釈をサポートしている.

より一般的の場合には, Sanovの定理は

- (確率分布  $p$  の独立試行の経験分布が  $q$  に近い確率) =  $\exp(-nD(q||p) + o(n))$

と要約でき, 確率分布  $p$  の独立試行の経験分布が  $q$  に近いことは, 確率分布  $p$  の独立試行により確率分布  $q$  の予測が間違っていることがばれずに済んでいることを意味していると解釈できるので, やはり  $D(q||p)$  は予測が間違っていたことがばれるまでの速さになっていると解釈できる.

KL情報量は以上の意味で予測誤差になっていると考えられる:

- $D(q||p) = (\text{確率分布 } p \text{ の独立試行による確率分布 } q \text{ のシミュレーションの予測誤差}).$

## 2.5 正規分布の場合

連続的な確率分布の場合であっても、確率密度函数  $q(x), p(x)$  に対するKL情報量が

$$D(q||p) = \int q(x) \log \frac{q(x)}{p(x)} dx$$

と定義され(  $\int$  は定積分を意味する), 適当な条件と適切な定式化のもとでSanovの定理が成立していることが知られている:

- $D(q||p) = (\text{確率分布 } p \text{ の独立試行による確率分布 } q \text{ のシミュレーションの予測誤差}).$

正確な内容については「[大偏差原理 Sanovの定理 \(https://www.google.co.jp/search?q=%E5%A4%A7%E5%81%8F%E5%B7%AE%E5%8E%9F%E7%90%86+Sanov%E3%81%AE%E5%AE%9A%E7%90%86\)](https://www.google.co.jp/search?q=%E5%A4%A7%E5%81%8F%E5%B7%AE%E5%8E%9F%E7%90%86+Sanov%E3%81%AE%E5%AE%9A%E7%90%86)」について検索して欲しい。(連続的な  $x$  軸を  $r$  個に区切れれば,  $x$  軸上の確率分布は有限集合  $\Omega = \{1, \dots, r\}$  上の確率分布で近似できることに注意せよ.)

確率分布  $p$  による独立試行で確率分布  $q$  をシミュレートする状況を考えよう.  $\int_{-\infty}^{\infty}$  を単に  $\int$  と書くことにすると, その場合の予測誤差は

$$D(q||p) = \int q(x) \log \frac{q(x)}{p(x)} dx = G(q||p) - S(q).$$

ここで

$$G(q||p) = - \int q(x) \log p(x) dx, S(q) = - \int q(x) \log q(x) dx$$

といおいた.  $S(q)$  はShannon情報量と呼ばれ,  $G(q||p)$  は汎化誤差と呼ばれる.

例として, 正規分布

$$p(x) = p_N(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)}$$

による独立試行で確率分布  $q(x)$  をシミュレートする場合について考えよう. 予測誤差を小さくするためにには, パラメーター  $\mu, \sigma$  を調節して汎化誤差  $G(q||p)$  を小さくすればよい.  $p$  が正規分布の場合には

$$\begin{aligned} G(q||p) &= \int q(x) \left( \frac{(x-\mu)^2}{2\sigma^2} + \frac{1}{2} \log \sigma^2 + \log \sqrt{2\pi} \right) dx \\ &= \frac{1}{\sigma^2} \int q(x)(x-\mu)^2 dx + \frac{1}{2} \log \sigma^2 + \log \sqrt{2\pi}. \end{aligned}$$

$q(x)$  の平均と分散を

$$\mu_q = \int q(x)x dx, \quad \sigma_q^2 = \int q(x)(x - \mu_q)^2 dx$$

と書くと,

$$\int q(x)(x - \mu)^2 dx = \int q(x)((x - \mu_q) - (\mu - \mu_q))^2 dx = \sigma_q^2 + (\mu - \mu_q)^2$$

なので

$$G(q||p) = \frac{(\mu - \mu_q)^2}{2\sigma^2} + \frac{\sigma_q^2}{2\sigma^2} + \frac{1}{2} \log \sigma^2 + \log \sqrt{2\pi}.$$

これより,  $G(q||p)$  を最小にする  $\mu, \sigma$  は  $\mu = \mu_q, \sigma = \sigma_q$  となることがわかる.

すなわち, 正規分布による確率分布  $q(x)$  のシミュレーションの予測誤差を最小にするためには, 正規分布の平均と分散を確率分布  $q(x)$  の平均と分散に一致させねばよいことがわかった.

この結果は未知の確率分布  $q(x)$  のサンプルに基いてその平均と分散の両方を推定することは, 未知の確率分布  $q(x)$  を最も小さな予測誤差でシミュレートできる正規分布を推定することと同じことであることがわかる.

平均と分散に関する統計学は, KL情報量のSanovの定理の立場から見ると, 本質的に正規分布の統計学であるということがわかった.

以上のように, KL情報量のSanovの定理は, 「パラメーターの推定」という考え方をより普遍的な「確率分布の推定」という考え方へ高めるため役に立つ.

問題: 上の計算で省略した部分を埋めよ.

解答例:  $\sigma$  を固定したときの  $G(q||p)$  の最小値は  $\mu = \mu_q$  で与えられることは明らか.

$$f(s) = \frac{a}{s} + \log s$$

とおくと,

$$f'(s) = -\frac{a}{s^2} + \frac{1}{s} = \frac{s-a}{s^2}$$

なので  $f(s)$  は  $s = a$  で最小になることがわかる。これより,  $G(q|p)$  は  $\mu = \mu_q$ ,  $\sigma = \sigma_q$  のときに最小になることがわかる。□

**問題:** 次の確率密度函数で定義される分布をロジスティック分布と呼ぶ:

$$q(x) = \frac{1}{4s} \operatorname{sech}^2\left(\frac{x-\mu}{2s}\right).$$

ここで

$$\operatorname{sech} x = \frac{1}{\cosh x}, \quad \cosh x = \frac{e^x + e^{-x}}{2}.$$

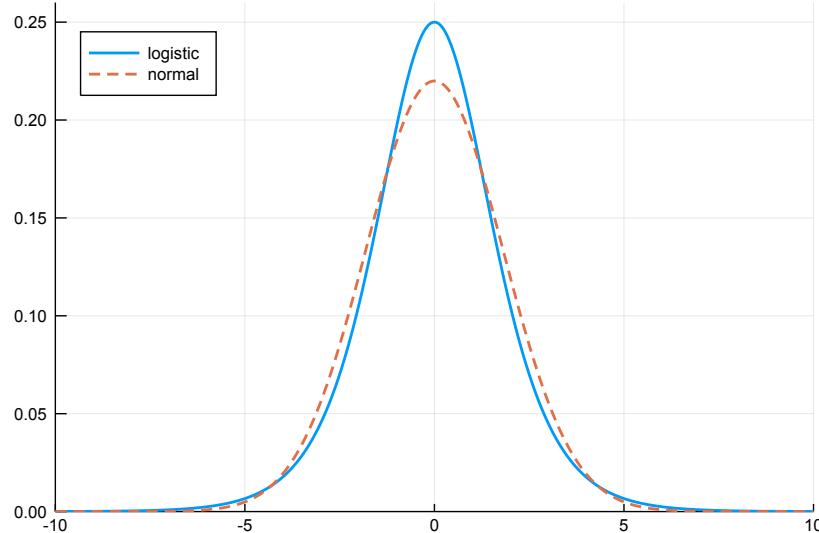
ロジスティック分布の平均と分散はそれぞれ

$$\mu_q = \mu, \quad \sigma_q^2 = \frac{\pi^2 s^2}{3}.$$

ロジスティック分布を最もよくシミュレートする正規分布の確率密度函数とともにロジスティック分布の確率密度函数を同時にプロットしてみよ。□

```
In [3]: 1 pdflogistic(mu,s,x) = 1/(4s)*sech((x-mu)/(2s))^2
2 pdfnormal(mu,s,x) = exp(-(x-mu)^2/(2*s^2))/sqrt(2*pi*s^2)
3
4 mu = 0.0
5 s = 1.0
6 sigma = pi*s/sqrt(3)
7 x = -10:0.02:10
8
9 P = plot(legend=:topleft, ylims=(0, 0.26))
10 plot!(x, pdflogistic.(mu,s,x), label="logistic", lw=2)
11 plot!(x, pdfnormal.(mu,s,x), label="normal", lw=2, ls=:dash)
```

Out[3]:



ロジスティック分布の正規分布による近似に関しては

- 黒木玄, [ロジスティック分布の2通りの正規分布近似の比較](http://nbviewer.jupyter.org/gist/genkuroki/96b0508a5773035e3a5247beff1d4f99) (<http://nbviewer.jupyter.org/gist/genkuroki/96b0508a5773035e3a5247beff1d4f99>)
- も参照せよ。

## 2.6 多項分布の中心極限定理

Sanovの定理の証明によって,  $k_i/n \sim q_i$  ( $n \rightarrow \infty$ ) のとき,  $n \rightarrow \infty$  で,

$$\frac{n!}{k_1! k_2! \cdots k_r!} p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r} \sim \left( \prod_{i=1}^r \left( \frac{k_i/n}{p_i} \right)^{k_i} \sqrt{(2\pi n)^{r-1} \prod_{i=1}^r \frac{k_i}{n}} \right)^{-1}.$$

これを

$$\frac{k_i}{n} = p_i + \frac{1}{\sqrt{n}} \varepsilon_i, \quad \varepsilon_i = o\left(\frac{1}{\sqrt{n}}\right)$$

の場合に適用すると,

$$\begin{aligned}
\sum_{i=1}^r \left( \frac{x_i}{\sqrt{n}} + \varepsilon_i \right) &= \sum_{i=1}^r \frac{k_i}{n} - \sum_{i=1}^r p_i = 0, \\
\log \left( \frac{k_i/n}{p_i} \right)^{k_i} &= np_i \left( 1 + \frac{x_i}{\sqrt{n} p_i} + \frac{\varepsilon_i}{p_i} \right) \log \left( 1 + \frac{x_i}{\sqrt{n} p_i} + \frac{\varepsilon_i}{p_i} \right) = n \left( \frac{x_i}{\sqrt{n}} + \varepsilon_i \right) + \frac{1}{2} \frac{x_i^2}{p_i} + o(1), \\
\therefore \log \prod_{i=1}^r \left( \frac{k_i/n}{p_i} \right)^{k_i} &= \sum_{i=1}^r \log \left( \frac{k_i/n}{p_i} \right)^{k_i} = \frac{1}{2} \sum_{i=1}^r \frac{x_i^2}{p_i} + o(1), \\
\prod_{i=1}^r \frac{k_i}{n} &= p_1 \cdots p_r (1 + o(1))
\end{aligned}$$

なので

$$\frac{n!}{k_1! k_2! \cdots k_r!} p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r} \sim \frac{\exp \left( -\frac{1}{2} \sum_{i=1}^r \frac{x_i^2}{p_i} \right)}{\sqrt{(2\pi n)^{r-1} p_1 \cdots p_r}}.$$

両辺に  $dk_1 \cdots dk_{r-1} = \sqrt{n} dx_1 \cdots \sqrt{n} dx_{r-1}$  をかけると,

$$\frac{n!}{k_1! k_2! \cdots k_r!} p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r} dk_1 \cdots dk_{r-1} \sim \frac{\exp \left( -\frac{1}{2} \sum_{i=1}^r \frac{x_i^2}{p_i} \right)}{\sqrt{(2\pi)^{r-1} p_1 \cdots p_r}} dx_1 \cdots dx_{r-1}.$$

$k_1 + \cdots + k_r = n$  より  $x_1 + \cdots + x_r = o(1)$  であることに注意せよ. 右辺は  $(x_1, \dots, x_{r-1})$  の多変量正規分布の確率密度函数の形になっている.

以上の結果は多項分布が  $n$  が大きなときに多変量正規分布で近似できることを意味している. この結果を**多項分布の中心極限定理**と呼ぶ.

**二項分布の中心極限定理:**  $r = 2$  のとき,  $0 \leq p \leq 1$ ,  $k/n = p + x/\sqrt{n} + o(1/\sqrt{n})$  のとき,

$$\binom{n}{k} p^k (1-p)^{n-k} dk \sim \frac{\exp \left( -\frac{1}{2} \left( \frac{x^2}{p} + \frac{(-x)^2}{1-p} \right) \right)}{\sqrt{2\pi p(1-p)}} dx = \frac{\exp \left( -\frac{x^2}{2p(1-p)} \right)}{\sqrt{2\pi p(1-p)}} dx.$$

この結果を**二項分布の中心極限定理**と呼ぶ. □

より一般的の場合の中心極限定理の証明については

- 黒木玄, [確率論入門](https://genkuroki.github.io/documents/IntroProbability.pdf) (<https://genkuroki.github.io/documents/IntroProbability.pdf>)

の第2.4節を参照せよ. 本質的にTaylorの定理のみを使って中心極限定理は証明される. (積率母函数や特性函数の理論(本質的にFourier-Laplace解析の理論)を使わなくても証明できる.)

**問題:** 二項分布の中心極限定理を実際にグラフをプロットすることによって確認せよ. □

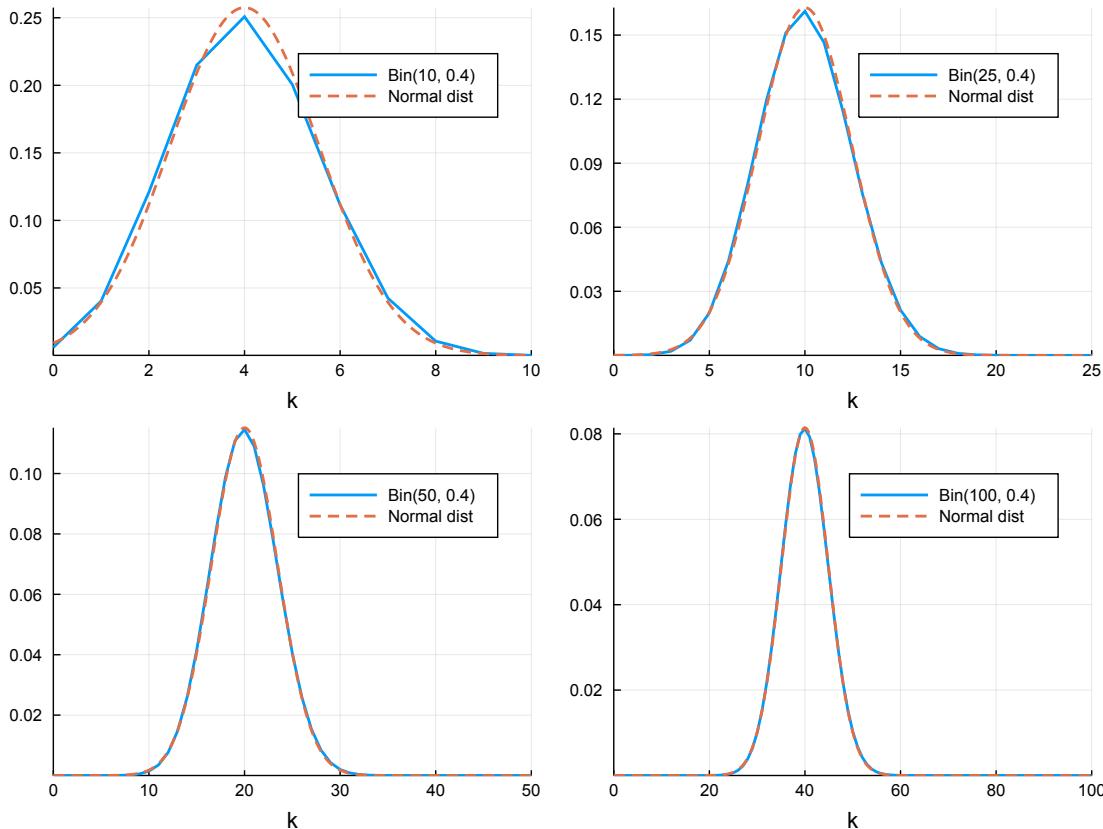
次のセルを見よ.

```

In [4]: 1 binomcoeff(n,k) = exp(lgamma(n+1)-lgamma(k+1)-lgamma(n-k+1))
2 pdfbinom(n,p,k) = 0 ≤ k ≤ n ? binomcoeff(n,k)*p^k*(1-p)^(n-k) : 0.0
3 pdfnormal(n,p,x) = exp(-x^2/(2p*(1-p)))/sqrt(2pi*n*p*(1-p))
4
5 PP = []
6 ▶ for (n,p) in [(10,0.4), (25,0.4), (50,0.4), (100,0.4)]
7   k = 0:n
8   kk = 0:n/400:n
9   x = (kk .- n*p)./sqrt(n)
10  P = plot(legend=:topright, xlabel="k")
11  plot!(k, pdfbinom.(n,p,k), label="Bin($n, $p)", lw=2)
12  plot!(kk, pdfnormal.(n,p,x), label="Normal dist", lw=2, ls=:dash)
13  push!(PP, P)
14 end
15 plot(PP..., size=(800, 600))
16 #pngplot(PP..., size=(800, 600))

```

Out[4]:



In [ ]:

1