

総和記号演算

前川眞一

070419,20,25,26,27,080211,13,
20191220, 20200422, 20230404

目次

1	添字付き変数と総和記号	1
1.1	添え字付き変数とは	1
1.2	総和記号	2
1.3	<i>spec</i> の利用	3
2	総和記号の演算公式	4
3	複数の変数を添え字付き変数で表す場合	6
3.1	複数の変数	6
3.2	関数の和の微積分に関する一般公式	6
3.3	多変量関数の添字付き変数による表現	6
3.4	総和記号を含む関数の添字付き変数に関する微分	6
3.5	多変数合成関数の微分	9
4	行列演算との関係	12
5	総積記号	14

1 添字付き変数と総和記号

1.1 添え字付き変数とは

添字付き変数とは、多くの値を共通の名前 *root* と添字 *sub* を用いて表すための記法であり、

$$root_{sub} \quad \text{or} \quad root_{sub1,sub2} \quad \text{or} \quad root_{sub1,sub2,sub3,\dots} \quad (1)$$

の様な形をしている。添字が二つ以上の場合、誤解を招かないようであれば添字の間のコンマは省略しても良い。
たとえば、5 個の値

$$a, b, c, d, e, f, g, h, i \quad (2)$$

を取り扱う場合、それらを格納する長さが 5 のベクトル \boldsymbol{x} を考え、*root* を x として、その要素を

$$x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_5 \quad (3)$$

とする。すなわち

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \\ e \end{pmatrix} \quad \text{or} \quad \boldsymbol{x} = (x_1, x_2, x_3, x_4, x_5) \\ = (a, b, c, d, e) \quad (4)$$

である。また、このベクトルの各要素を、一般的に *sub* を i として、また $n = 5$ として

$$x_i, i = 1, 2, \dots, n \quad (5)$$

として参照する。たとえば x_2 の値は b であり、 x_4 の値は d である。なお、添字として用いる文字 *sub* は自由であり、同じ 5 つの要素を

$$x_j, j = 1, 2, \dots, 5 \quad (6)$$

と書いてもよい。ただし、一度添字付き変数が定義された後は、 x_i と x_j は $i \neq j$ の場合、異なる要素を表す。

また、*root* として用いる文字も自由であるが、一般的に x_i と y_i は異なる。例えば、 y_i には別の値 $\alpha, \beta, \gamma, \dots$ が格納されているかもしれない。

1.2 総和記号

添字付き変数を用いると、たとえば、先ほどの例では、5つの値のうち、最初の3つの和は、 $a + b + c$ であるが、それは $x_1 + x_2 + x_3$ と書ける。また、これは、総和記号を用いて、

$$\sum_{i=1}^3 x_i \quad (7)$$

と書くことが出来る。

総和記号とは、一般的に

$$\sum_{sub=from}^{to} body \quad (8)$$

の形をしており、これは、任意の数式 $body$ に含まれる添字 sub を $from$ から to まで順に変えることにより得られる数式を、全て足し合わせることを示している。上記の例では

$$sub = i, from = 1, to = 3, body = x_i \quad (9)$$

である。また、総和記号は、文中では $\sum_{sub=from}^{to} body$ の様に表記されることもある。

たとえば、 $x_i, i = 1, 2, \dots, n$ の総和は

$$\sum_{i=1}^n x_i \quad (10)$$

である。

総和記号に用いられる添字は何でもよく、

$$\sum_{i=1}^n x_i = \sum_{j=1}^n x_j \quad (11)$$

である。

添字がふたつ以上付いてる変数もある。たとえば、2元配置の分散分析で要因 A の第 i 水準、要因 B の第 j 水準の組み合わせ（セル）でそれぞれ一つだけデータが観測されているばあい、そのデータを y_{ij} と記す。ただし、要因の数をそれぞれ A, B とし、 $i = 1, 2, \dots, A, j = 1, 2, \dots, B$ である。

一般的に、ふたつの添字が付いている変数は、 $A \times B$ の行列の第 i 行第 j 列要素（ ij 要素）を表していると考えられ、最初の添字が行を二番目の添字が列をあらわす。

この場合、要因 A の第 i 水準の元に観測された全ての観測値の和、ならびに、要因 B の第 j 水準の元に観測された全ての観測値の和は

$$r_i = \sum_{j=1}^B y_{ij} \quad \text{and} \quad c_j = \sum_{i=1}^A y_{ij} \quad (12)$$

と書ける。和を取るのに用いられた以外の添字は、結果として得られる変数にそのまま付与されていることに注意。より一般的に言えば、和を取られた添字は、結果から消えている。

和を取る際に使う添字の記号は自由であるから、 r_i を

$$r_i = \sum_{k=1}^B y_{ik} \quad (13)$$

と書いても良い。また、要素を $y_{ij}, i = 1, 2, \dots, A, j = 1, 2, \dots, B$ で定義した行列に関して、**第 j 行の和** を

$$r_j = \sum_{k=1}^B y_{jk}, j = 1, 2, \dots, A \quad (14)$$

と定義しても間違いではないが、なるべく、要素を定義した時の添字（この場合は第 1 番目が i で行を表し、2 番目が j で列を表す）を使うべきである。

全ての観測値（AB 個）の総和は、行ごとの和 r_i の和であり、しかも各列の和であるから

$$t = \sum_{i=1}^A r_i = \sum_{i=1}^A \sum_{j=1}^B y_{ij} \quad (15)$$

$$= \sum_{j=1}^B c_j = \sum_{j=1}^B \sum_{i=1}^A y_{ij} \quad (16)$$

である。これは、総和記号の *scope, from, to* が他の添字に依存しない場合、総和記号の順序を入れ替えても良いことを示している。すなわち、加算は順序に依存しないので、どのように加えても良い。しかし、総和記号の解釈としては、右側（内側）に書かれた総和記号から順に和を取っていきと解釈するべきである。すなわち、上式の第 1 番目は、まず j に関する和を取り、その後、その結果を i に関して加えると解釈するべきである。

次に、各セルごとに n 個の観測値があるとする。この場合、データは $y_{ijk}, i = 1, 2, \dots, A, j = 1, 2, \dots, B, k = 1, 2, \dots, n$ である。したがって、各セルの和や、水準ごとの和、ならびの総和は

$$y_{ij+} = \sum_{k=1}^n y_{ijk}, \quad r_i = y_{i++} = \sum_{j=1}^B \sum_{k=1}^n y_{ijk}, \quad c_j = y_{+j+} = \sum_{i=1}^A \sum_{k=1}^n y_{ijk}, \quad t = y_{+++} = \sum_{j=1}^B \sum_{i=1}^A \sum_{k=1}^n y_{ijk} \quad (17)$$

と書けるが、総和記号の順序は自由である。

セルごとにそれぞれ観測値数が異なる場合、それを n_{ij} とする。ただし、 $i = 1, 2, \dots, A, j = 1, 2, \dots, B$ である。また、各セルの第 k 番目の観測値を y_{ijk} と書けば、 $i = 1, 2, \dots, A, j = 1, 2, \dots, B, k = 1, 2, \dots, n_{ij}$ である。

この場合、各セルの平均や要因 A の第 i 水準の元に観測された全ての観測値の和は

$$\bar{y}_{ij} = y_{ij.} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk} \quad (18)$$

$$c_i = y_{i++} = \sum_{j=1}^B \sum_{k=1}^{n_{ij}} y_{ijk} = \sum_{j=1}^B (n_{ij} \bar{y}_{ij}) \quad (19)$$

と書ける。しかし、2 番目の総和記号の添字の上限が 1 番目の総和記号の添字 j を含むため

$$c_i = \sum_{k=1}^{n_{ij}} \sum_{j=1}^B y_{ijk} \quad \text{間違い} \quad (20)$$

とは書けない。加算の性質から、どの順に加えようと結果は同じであるが、総和記号の記法上、上記のように表現することは出来ない。なぜなら、総和記号は右（内側）から実行されると考えられるため、外側の総和記号に出てくる j が定義されていないからである。

観測値の総和は

$$t = \sum_{i=1}^A \sum_{j=1}^B \sum_{k=1}^{n_{ij}} y_{ijk} = \sum_{j=1}^B \sum_{i=1}^A \sum_{k=1}^{n_{ij}} y_{ijk} \quad (21)$$

である。

1.3 spec の利用

総和記号の変形として、

$$\sum_{scope} body \quad (22)$$

の様な記法もある。spec には、どの添字に関する和を取るかを明記し、その後に、条件を付ける形が多く用いられる。たとえば、添字が偶数のものの和は

$$\sum_{j, j=\text{偶数}} x_j \quad (23)$$

の様に書ける。

総和記号の spec として以下のものがよく利用される。

- 範囲の省略

$$\sum x_i = \sum_i x_i = \sum_{i=1}^n x_i \quad \text{and} \quad \sum x_{ij} = \sum_{i=1}^n x_{ij} \quad (24)$$

- 行列の対角要素

$$\sum_{i,j,i=j} x_{ij} = \sum_i x_{ii} \quad (25)$$

- 行列の非対角要素

$$\sum_{i,j,i \neq j} x_{ij} = \sum_{i=1}^n \sum_{j=1}^n x_{ij} - \sum_{i,j,i=j} x_{ij} \quad (26)$$

- 行列の対角を含む、もしくは含まない下三角部分

$$\sum_{i,j,i \geq j} x_{ij} \quad \text{and} \quad \sum_{i,j,i > j} x_{ij} \quad (27)$$

- 行列の対角を含む、もしくは含まない上三角部分

$$\sum_{i,j,i \leq j} x_{ij} \quad \text{and} \quad \sum_{i,j,i < j} x_{ij} \quad (28)$$

- 一つの総和記号の分解

$$\sum_{i=1}^n body = \sum_{i=1}^{k-1} body + body_k + \sum_{i=k+1}^n body \quad (29)$$

$$= \sum_{i,i \neq k} body + body_k \quad (30)$$

ただし、 $body$ は添字 i を含み、 $body_k$ は、 $body$ に含まれる添字 i を k と於いたものであり、

$$\sum_{i,i \neq k} body = \sum_{i=1}^{k-1} body + \sum_{i=k+1}^n body = \sum_{i=1}^n body - body_k \quad (31)$$

である。この分解の結果 $body_k$ は $\sum_{i,i \neq k} body$ の部分から除外される。

2 総和記号の演算公式

- 定数

$$\sum_{scope} body = scope \text{ に含まれる要素数} \times body \quad (32)$$

ただし、 $body$ は $scope$ に記された添字を含まない。

- 分配の法則 1

$$\sum_{scope} (body1 + body2) = \sum_{scope} body1 + \sum_{scope} body2 \quad (33)$$

- 分配の法則 2

$$\sum_{scope} (body1 \times body2) = body1 \sum_{scope} body2 \quad (34)$$

ただし、 $body1$ は $scope$ に記された添字を含まない。

これらを利用すると以下のようなことが解る。

- 定数

$$\sum_i^n c = n \times c \quad (35)$$

- 総和記号の分配

$$\sum_i (x_i + y_i) = \sum_i x_i + \sum_i y_i \quad (36)$$

- 定数倍

$$\sum_i^n cy_i = c \sum_i y_i \quad (37)$$

- 総和記号の移動 1

$$\sum_i \sum_j x_i y_j = \sum_i x_i \sum_j y_j = \left(\sum_i x_i \right) \times \left(\sum_j y_j \right) \quad (38)$$

- 総和記号の移動 2

$$\sum_{sub1} \cdots \sum_{subK} \cdots \sum_{subS} (body1 \times body2) = \sum_{sub1} \cdots \sum_{subK-1} body1 \sum_{subK} \cdots \sum_{subS} body2 \quad (39)$$

ただし $body1$ は添字 $subK$ を含まず $body2$ は添字 $subK$ を含む場合。

- 二重総和記号の分解

$$\sum_i \sum_j x_{ij} = \sum_{i,i \neq p} \left(\sum_j x_{ij} \right) + \sum_j x_{pj} \quad (40)$$

$$= \sum_{i,i \neq p} \left(\sum_{j,j \neq q} x_{ij} + x_{iq} \right) + \sum_{j,j \neq q} x_{pj} + x_{pq} \quad (41)$$

$$= \sum_{i,i \neq p} \sum_{j,j \neq q} x_{ij} + \sum_{i,i \neq p} x_{iq} + \sum_{j,j \neq q} x_{pj} + x_{pq} \quad (42)$$

この分解の結果、 x_{pq} は最後の項にしか含まれない。これを総和記号の分解公式と呼ぶ。以下の図を参照のこと。なお、少し冗長ではあるが、

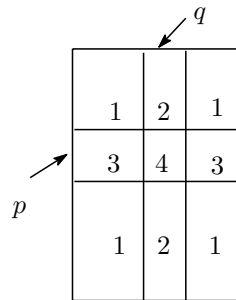


Figure 1: Divided Double Summation

$$\sum_i \sum_j x_{ij} = \sum_{i,i \neq p} \sum_{j,j \neq q} x_{ij} + \sum_i x_{iq} + \sum_j x_{pj} - x_{pq} \quad (43)$$

と書くことも出来るが、この場合、第2項と第3項にも x_{pq} が含まれることになる。

3 複数の変数を添え字付き変数で表す場合

3.1 複数の変数

これまでは x_i 等のベクトルの要素を特定の値として説明をしてきたが、複数の変数を並べたものをベクトルとして扱うことも出来る。例えば地図上の点はその緯度と経度を指定することにより表すことが出来るが \boldsymbol{x} として x_1 を緯度、 x_2 を経度とするベクトルを考えることが出来る。また、緯度と経度のを与えたときにその土地の高度 y_1 を返す関数を f_1 、また、その土地の地価 y_2 を表す関数を f_2 とすれば

$$y_i = f_i(x_1, x_2) = f_i(\boldsymbol{x}), \quad i = 1, 2 \quad (44)$$

と書くことが出来る。ここに、 $\boldsymbol{y} = (y_1, y_2)$ は高度と地価という二つの変数を持つベクトルである。

以下に、例えば、地価と緯度経度の関係を明らかにしたいような場合に、地価という変数を緯度や経度という変数で微分する様な場合に必要となる添え字付き変数による関数の微分に関して説明する。

(実際に N 個の個体に関するそれらの変数の値がデータとしてある場合には、それらは行列に格納されることになる。)

3.2 関数の和の微積分に関する一般公式

n 個の関数を $f_i(x), i = 1, 2, \dots, n$ で表すと、微積分の性質から、通常の間数では、

$$\frac{d \sum_{i=1}^n f_i(x)}{dx} = \sum_{i=1}^n \frac{d f_i(x)}{dx} \quad (45)$$

$$\int_X \sum_{i=1}^n f_i(x) dx = \sum_{i=1}^n \int_X f_i(x) dx \quad (46)$$

である。すなわち、和を取ってから微積分を行っても、微積分を行ってから和をとっても同じ結果が得られる。

3.3 多変量数関数の添字付き変数による表現

ある地点の緯度 a と経度 b を与えた時にその地点の高度 g を与えるような関数 $g = f(a, b)$ を多変量数関数と呼ぶ。関数への入力が多変量数であり、出力はスカラーである。一般的にそのような関数を

$$y = f(x_1, x_2, \dots, x_n) \quad (47)$$

と書く。 x_1, x_2, \dots は a, b, \dots 等の変数に順に名前を付けた物であり、 $i \neq j$ であれば x_i と x_j は別の変数を表す。 f をある特定の x_k で偏微分したものを

$$\frac{\partial y}{\partial x_k} = \frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_k} \quad (48)$$

と書く。なお、 x_k に関する偏微分とは $x_i, i \neq k$ の変数を定数として取り扱った微分の事である。

3.4 総和記号を含む関数の添字付き変数に関する微分

添字付き変数とは、多数の異なる変数を、共通の名前 *root* と添字を用いて統一的に表したものであった。したがって、 $i \neq j$ の場合、 x_i と x_j は異なる変数を表している。また、添字が二つ以上付いた場合でも、 $i = I \ \& \ j = J \ \& \ \dots$ でなければ $x_{ij\dots}$ と $IJ\dots$ は別物である。

したがって、ある添字付き変数を、別の添字付き変数で微分すると

$$\frac{\partial x_i}{\partial x_k} = \delta_{ik} = \begin{cases} 0 & \text{if } i \neq k \\ 1 & \text{if } i = k \end{cases} \quad (49)$$

$$\frac{\partial x_{ij\dots}}{\partial x_{k\ell\dots}} = \delta_{ii\dots, k\ell\dots} = \begin{cases} 0 & \text{if } \text{not}(i = k \ \& \ j = \ell \ \& \ \dots) \\ 1 & \text{if } i = k \ \& \ j = \ell \ \& \ \dots \end{cases} \quad (50)$$

となる。(ここから多変量数の関数の話が出てくるので、偏微分の記号を用いる。)

たとえば、変数 a, b, \dots を順に $x_i, i = 1, 2, \dots, n$ で表し、それらの変数の関数を

$$g(a, b, c, d, \dots) = a^2 + b^2 + \dots = \sum_{i=1}^n x_i^2 \quad (51)$$

とする。これを、たとえば第 2 番目の変数 b で微分すると

$$\frac{\partial g(a, b, \dots)}{\partial b} = \frac{\partial (a^2 + b^2 + \dots)}{\partial b} = 2b \quad (52)$$

であるから $x_2 = b$ として $g(x_1, x_2, \dots) = \sum_i x_i^2$ の x_2 に関する偏微分は

$$\frac{\partial \sum_i x_i^2}{\partial x_2} = \frac{\partial (x_1^2 + x_2^2 + \dots)}{\partial x_2} = 2x_2 \quad (53)$$

となる。したがって、一般的に、 $g(x_1, x_2, \dots) = \sum_i x_i^2$ を添字付き変数 $x_i, i = 1, 2, \dots$ のある特定の変数 x_k に関して微分すると

$$\frac{\partial \sum_i x_i^2}{\partial x_k} = \frac{\partial (x_1^2 + x_2^2 + \dots)}{\partial x_k} = 2x_k \quad (54)$$

と書ける。

より一般的には、添字 i に依存する関数を $f_i()$ と書き

$$g((x_1, x_2, \dots) = \sum_i f_i(x_i) \quad (55)$$

とすると、 x_i と x_j は $i \neq j$ であれば無関係である時には、

$$\frac{\partial g(x_1, x_2, \dots)}{\partial x_k} = \frac{\partial \sum_{i=1}^n f_i(x_i)}{\partial x_k} \quad (56)$$

$$= \sum_{i=1}^n \frac{\partial f_i(x_i)}{\partial x_k} = \sum_{i=1}^n \frac{\partial f_i(x_i)}{\partial x_i} \frac{\partial x_i}{\partial x_k} \quad (57)$$

$$= \sum_{i=1}^n f'_i(x_i) \frac{\partial x_i}{\partial x_k} = \sum_{i=1}^n f'_i(x_i) \delta_{ik} \quad (58)$$

$$= f'_k(x_k) \quad (59)$$

となる。(先ほどの例では $f_i(x) = x^2$ として全て同じ関数を用いたが、 $f_2(x) = \log x$ の様な場合もある。)

この例は、 δ 記法を用いて、先に微分を取ってから加える方法であるが、逆に、以下のように、和を総和記号の分解公式を用いて分解し、最後に微分をとる方法もある。

$$\frac{\partial \sum_{i=1}^n f_i(x_i)}{\partial x_k} = \frac{\partial \left(\sum_{i, i \neq k} f_i(x_i) + f_k(x_k) \right)}{\partial x_k} \quad (60)$$

$$= \frac{\partial \sum_{i, i \neq k} f_i(x_i)}{\partial x_k} + \frac{\partial f_k(x_k)}{\partial x_k} \quad (61)$$

$$= 0 + \frac{\partial f_k(x_k)}{\partial x_k} \quad (62)$$

$$= \frac{\partial f_k(x_k)}{\partial x_k} = f'_k(x_k) \quad (63)$$

どちらの方法を用いても同じ結果となる。

例

- 一次式の微分

$$\frac{\partial \sum_j a_j x_j}{\partial x_k} \quad (64)$$

これは δ 記法を用いると、 $f_i(x_i) = a_i x_i$ として、

$$\frac{\partial \sum_j a_j x_j}{\partial x_k} = \sum_j \frac{\partial a_j x_j}{\partial x_k} \quad (65)$$

$$= \sum_j a_j \frac{\partial x_j}{\partial x_k} = \sum_j a_j \delta_{jk} \quad (66)$$

$$= a_k \quad (67)$$

である。同様に、分解公式により

$$\sum_j a_j x_j = \sum_{j,j \neq k} a_j x_j + a_k x_k \quad (68)$$

したがって、

$$\frac{\partial \sum_j a_j x_j}{\partial x_k} = a_k \quad (69)$$

である。

これまでの例では、結果として得られる微分からは δ 記法が消えていたが、それは元の関数が、総和記号を含んでいたからであった。しかし、常にそうなるとは限らない。たとえば、もう少し複雑な例として、ユークリッド距離の2乗

$$d_{ij}^2 = \sum_{a=1}^r (x_{ia} - x_{ja})^2 \quad (70)$$

の x_{kb} に関する微分を考えると以下ようになる。

$$\frac{\partial \sum_{a=1}^r (x_{ia} - x_{ja})^2}{\partial x_{kb}} = \sum_{a=1}^r \frac{\partial (x_{ia} - x_{ja})^2}{\partial x_{kb}} \quad (71)$$

$$= 2 \sum_{a=1}^r (x_{ia} - x_{ja}) \frac{\partial (x_{ia} - x_{ja})}{\partial x_{kb}} \quad (72)$$

$$= 2 \sum_{a=1}^r (x_{ia} - x_{ja}) \left(\frac{\partial x_{ia}}{\partial x_{kb}} - \frac{\partial x_{ja}}{\partial x_{kb}} \right) \quad (73)$$

$$= 2 \sum_{a=1}^r (x_{ia} - x_{ja}) (\delta_{ia,kb} - \delta_{ja,kb}) \quad (74)$$

$$= 2(x_{ib} - x_{jb}) (\delta_{ib,kb} - \delta_{jb,kb}) \quad (75)$$

$$= 2(x_{ib} - x_{jb}) (\delta_{ik} - \delta_{jk}) \quad (76)$$

この場合も、全ての距離の和の微分を考えると以下のように δ 記法は消えてしまう。

$$\frac{\partial \sum_i \sum_j \sum_{a=1}^r (x_{ia} - x_{ja})^2}{\partial x_{kb}} = \sum_i \sum_j \frac{\partial \sum_{a=1}^r (x_{ia} - x_{ja})^2}{\partial x_{kb}} \quad (77)$$

$$= 2 \sum_i \sum_j (x_{ib} - x_{jb}) (\delta_{ik} - \delta_{jk}) \quad (78)$$

$$= 2 \sum_i \sum_j ((x_{ib} - x_{jb}) \delta_{ik} - (x_{ib} - x_{jb}) \delta_{jk}) \quad (79)$$

$$= 2 \left(\sum_i \sum_j (x_{ib} - x_{jb}) \delta_{ik} - \sum_i \sum_j (x_{ib} - x_{jb}) \delta_{jk} \right) \quad (80)$$

$$= 2 \left(\sum_j (x_{kb} - x_{jb}) - \sum_i (x_{ib} - x_{kb}) \right) \quad (81)$$

$$= 4 \sum_j (x_{kb} - x_{jb}) \quad (82)$$

3.5 多変数合成関数の微分

関数 $\phi(x)$ が変数 x の関数 $g(x)$ の関数として

$$\phi(x) = F(g(x)) \quad (83)$$

と書き表されるとき、この $\phi(x)$ を x で微分する際に一変数の場合の合成関数の微分公式（推移率）

$$\frac{d\phi(x)}{dx} = \frac{dF(g(x))}{dg} = \frac{dF(g)}{dg} \frac{dg(x)}{dx} \quad (84)$$

が使える。ここでは、それを多変数の場合に拡張する。

ふたつの多変数関数 $g_1 = f_1(x_1, x_2)$ と $g_2 = f_2(x_1, x_2)$ を用いて合成関数

$$\phi(x_1, x_2) = F(g_1, g_2) = F(f_1(x_1, x_2), f_2(x_1, x_2)) \quad (85)$$

を定義する。これを x_1 や x_2 で偏微分すると

$$\frac{\partial \phi(x_1, x_2)}{\partial x_1} = \frac{\partial F(f_1(x_1, x_2), f_2(x_1, x_2))}{\partial x_1} = \frac{\partial F(g_1, g_2)}{\partial x_1} \quad (86)$$

$$= \frac{\partial F(g_1, g_2)}{\partial g_1} \frac{\partial g_1(x_1, x_2)}{\partial x_1} + \frac{\partial F(g_1, g_2)}{\partial g_2} \frac{\partial g_2(x_1, x_2)}{\partial x_1} \quad (87)$$

$$\frac{\partial \phi(x_1, x_2)}{\partial x_2} = \frac{\partial F(f_1(x_1, x_2), f_2(x_1, x_2))}{\partial x_2} = \frac{\partial F(g_1, g_2)}{\partial x_2} \quad (88)$$

$$= \frac{\partial F(g_1, g_2)}{\partial g_1} \frac{\partial g_1(x_1, x_2)}{\partial x_2} + \frac{\partial F(g_1, g_2)}{\partial g_2} \frac{\partial g_2(x_1, x_2)}{\partial x_2} \quad (89)$$

となる。

たとえば、

$$g_1 = x_1 + x_2 \quad \text{and} \quad g_2 = x_1 x_2, \quad (90)$$

$$\phi(x_1, x_2) = F(g_1, g_2) = g_1^2 - g_2 = (x_1 + x_2)^2 - x_1 x_2 \quad (91)$$

とすると、 x_1 と x_2 は別の変数であるから、

$$\frac{\partial \phi(x_1, x_2)}{\partial x_1} = \frac{\partial (x_1 + x_2)^2 - x_1 x_2}{\partial x_1} = 2x_1 + x_2 \quad (92)$$

$$\frac{\partial \phi(x_1, x_2)}{\partial x_2} = \frac{\partial (x_1 + x_2)^2 - x_1 x_2}{\partial x_2} = x_1 + 2x_2 \quad (93)$$

$$(94)$$

であるが、上記の公式を使うと

$$\frac{\partial \phi(x_1, x_2)}{\partial x_1} = \frac{\partial g_1^2 - g_2}{\partial g_1} \frac{\partial x_1 + x_2}{\partial x_1} + \frac{\partial g_1^2 - g_2}{\partial g_2} \frac{\partial x_1 x_2}{\partial x_1} \quad (95)$$

$$= 2g_1 \times 1 + (-1) \times x_2 = 2(x_1 + x_2) - x_2 \quad (96)$$

$$\frac{\partial \phi(x_1, x_2)}{\partial x_1} = \frac{\partial g_1^2 - g_2}{\partial g_1} \frac{\partial x_1 + x_2}{\partial x_2} + \frac{\partial g_1^2 - g_2}{\partial g_2} \frac{\partial x_1 x_2}{\partial x_2} \quad (97)$$

$$= 2g_1 \times 1 + (-1) \times x_1 = 2(x_1 + x_2) - x_1 \quad (98)$$

となり、同様の結果を得る。

一般的に x_1, x_2, \dots, x_n の関数 $g_\ell = f_\ell(x_1, x_2, \dots, x_n)$, $\ell = 1, 2, \dots, m$ を考え、それらの合成関数

$$\phi(x_1, x_2, \dots, x_n) = F(g_1, g_2, \dots, g_m) \quad (99)$$

$$= F(f_1(x_1, x_2, \dots, x_n), f_2(x_1, x_2, \dots, x_n), \dots, f_m(x_1, x_2, \dots, x_n)) \quad (100)$$

を定義する。これをある特定の変数 x_k で微分したものは

$$\frac{\partial \phi(x_1, x_2, \dots, x_n)}{\partial x_k} = \sum_{\ell=1}^m \frac{\partial F(g_1, g_2, \dots, g_m)}{\partial g_\ell} \frac{\partial g_\ell(x_1, x_2, \dots, x_n)}{\partial x_k} \quad (101)$$

となる。

上記の公式の特殊例として

$$g_\ell(x_1, x_2, \dots, x_n) = x_\ell, \ell = 1, 2, \dots, m = n \quad (102)$$

の場合を考える。この場合、

$$\phi(x_1, x_2, \dots, x_n) = F(x_1, x_2, \dots, x_n) \quad (103)$$

であるから

$$\frac{\partial \phi(x_1, x_2, \dots, x_n)}{\partial x_k} = \frac{\partial F(x_1, x_2, \dots, x_n)}{\partial x_k} \quad (104)$$

$$= \sum_{\ell=1}^n \frac{\partial F(x_1, x_2, \dots, x_n)}{\partial x_\ell} \frac{\partial x_\ell}{\partial x_k} \quad (105)$$

となるが、 $\ell \neq k$ であれば x_ℓ と x_k は別の変数であるから、

$$\frac{\partial x_\ell}{\partial x_k} = \delta_{\ell k} = \begin{cases} 0 & \text{if } \ell \neq k \\ 1 & \text{if } \ell = k \end{cases} \quad (106)$$

と書ける。したがって、添字付き変数の多変量関数 F のある特定の変数 x_k に関する偏微分は

$$\frac{\partial F(x_1, x_2, \dots, x_n)}{\partial x_k} = \sum_{\ell=1}^n \frac{\partial F(x_1, x_2, \dots, x_n)}{\partial x_\ell} \frac{\partial x_\ell}{\partial x_k} = \sum_{\ell=1}^n \frac{\partial F(x_1, x_2, \dots, x_n)}{\partial x_\ell} \delta_{\ell k} \quad (107)$$

となる。すなわち、関数 F の引数である添字付き変数のそれぞれに関して連鎖律による偏微分を行ったものを加えていく。

たとえば、 $n = 2$, $F(x_1, x_2) = x_1 x_2$ と於いた例を考えると

$$\frac{\partial F(x_1, x_2)}{\partial x_k} = \frac{\partial x_1 x_2}{\partial x_k} \quad (108)$$

$$= \frac{\partial x_1 x_2}{\partial x_1} \delta_{1k} + \frac{\partial x_1 x_2}{\partial x_2} \delta_{2k} \quad (109)$$

$$= x_2 \delta_{1k} + x_1 \delta_{2k} \quad (110)$$

となる。(これは先ほどの2乗距離の微分同様、 δ 記法が残っているが、 $k=1, \text{ or } 2$ と於くことにより、正しい答えが得られている。

次に、以下の例を考えてみる。

$$\phi(x_1, x_2, \dots, x_n) = \sum_i \sum_j F(x_i, x_j) \quad (111)$$

の x_k に関する微分である。この場合、まず、偏微分記号を総和記号の中に入れ、その後、上記の公式を用いると

$$\frac{\partial \sum_i \sum_j F(x_i, x_j)}{\partial x_k} = \sum_i \sum_j \frac{\partial F(x_i, x_j)}{\partial x_k} \quad (112)$$

$$= \sum_i \sum_j \sum_{\ell=1}^2 \frac{\partial F(x_i, x_j)}{\partial x_\ell} \frac{\partial x_\ell}{\partial x_k} \quad (113)$$

$$= \sum_i \sum_j \left(\frac{\partial F(x_i, x_j)}{\partial x_i} \frac{\partial x_i}{\partial x_k} + \frac{\partial F(x_i, x_j)}{\partial x_j} \frac{\partial x_j}{\partial x_k} \right) \quad (114)$$

$$= \sum_i \sum_j \left(\frac{\partial F(x_i, x_j)}{\partial x_i} \delta_{ik} + \frac{\partial F(x_i, x_j)}{\partial x_j} \delta_{jk} \right) \quad (115)$$

となるが、これは δ_{ik} が $i = k$ の時にだけ 1 となることに注意すれば、

$$= \sum_i \sum_j \left(\frac{\partial F(x_i, x_j)}{\partial x_i} \delta_{ik} \right) + \sum_i \sum_j \left(\frac{\partial F(x_i, x_j)}{\partial x_j} \delta_{jk} \right) \quad (116)$$

$$= \sum_j \sum_i \left(\frac{\partial F(x_i, x_j)}{\partial x_i} \delta_{ik} \right) + \sum_i \sum_j \left(\frac{\partial F(x_i, x_j)}{\partial x_j} \delta_{jk} \right) \quad (117)$$

$$= \sum_j \frac{\partial F(x_k, x_j)}{\partial x_k} + \sum_i \frac{\partial F(x_i, x_k)}{\partial x_k} \quad (118)$$

となる。

より一般的には

$$\frac{\partial \sum_i \sum_j \dots F_{ij\dots}(x_i, x_j, \dots)}{\partial x_k} = \sum_q A_{qk} \quad (119)$$

ただし、 A_{qk} は $F_{ij\dots}(x_i, x_j, \dots)$ の第 q 番目の添字を k で置き換えたものを x_k で偏微分したものを k を除く全ての i, j, \dots に関して加えたものである。

例

- ベクトル x の 2 次形式 $x'Ax$ の x の要素に関するの微分

$$\frac{\partial \sum_{i=1}^p \sum_{j=1}^p a_{ij} x_i x_j}{\partial x_k} = \sum_{i=1}^p \sum_{j=1}^p \frac{\partial a_{ij} x_i x_j}{\partial x_k} \quad (120)$$

$$= \sum_{i=1}^p \sum_{j=1}^p \left(a_{ij} x_i \frac{\partial x_j}{\partial x_k} + a_{ij} x_j \frac{\partial x_i}{\partial x_k} \right) \quad (121)$$

$$= \sum_{i=1}^p \sum_{j=1}^p a_{ij} x_i \delta_{kj} + \sum_{i=1}^p \sum_{j=1}^p a_{ij} x_j \delta_{ik} \quad (122)$$

$$= \sum_{i=1}^p a_{ik} x_i + \sum_{j=1}^p a_{kj} x_j \quad (123)$$

$$= \sum_{i=1}^p a_{ik} x_i + \sum_{i=1}^p a_{ki} x_i \quad (124)$$

$$= \sum_{i=1}^p (a_{ik} + a_{ki}) x_i \quad (125)$$

同じ例を分解公式を用いれば以下のようになる。すなわち、

$$\sum_{i,i \neq k}^p \sum_{j=1}^p a_{ij} x_i x_j = \sum_{i,i \neq k}^p \sum_{j=1}^p a_{ij} x_i x_j + \sum_{j=1}^p a_{kj} x_k x_j \quad (126)$$

$$= \sum_{i,i \neq k} \left(\sum_{j,j \neq k} a_{ij} x_i x_j + a_{ik} x_i x_k \right) + \sum_{j,j \neq k} a_{kj} x_k x_j + a_{kk} x_k x_k \quad (127)$$

$$= \sum_{i,i \neq k} \sum_{j,j \neq k} a_{ij} x_i x_j + \sum_{i,i \neq k} a_{ik} x_i x_k + \sum_{j,j \neq k} a_{kj} x_k x_j + a_{kk} x_k x_k \quad (128)$$

となるため、

$$\frac{\partial \sum_{i,i \neq k}^p \sum_{j=1}^p a_{ij} x_i x_j}{\partial x_k} = 0 + \sum_{i,i \neq k} a_{ik} x_i + \sum_{j,j \neq k} a_{kj} x_j + 2a_{kk} x_k \quad (129)$$

$$= \sum_{i,i \neq k} a_{ik}x_i + a_{kk}x_k + \sum_{j,j \neq k} a_{kj}x_j + a_{kk}x_k \quad (130)$$

$$= \sum_{i=1}^p a_{ik}x_i + \sum_{j=1}^p a_{kj}x_j \quad (131)$$

$$= 2 \sum_{i=1}^p a_{ik}x_i \quad \text{if } a_{ij} = a_{ji} \quad (132)$$

である。

4 行列演算との関係

\mathbf{a} と \mathbf{b} を $r \times 1$ の縦ベクトルとする。この二つのベクトルの和や差を \mathbf{c} とすると、

$$\mathbf{c} = \mathbf{a} \pm \mathbf{b}, \quad c_i = a_i \pm b_i \quad (133)$$

となる。二つのベクトルの内積を c とすると

$$c = \mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a}, \quad c = \sum_{i=1}^r a_i b_i \quad (134)$$

である。

\mathbf{A} と \mathbf{B} を $r \times c$ の行列とする。この二つの行列の和や差を \mathbf{C} とすると、

$$\mathbf{C} = \mathbf{A} \pm \mathbf{B}, \quad c_{ij} = a_{ij} \pm b_{ij} \quad (135)$$

となる。 \mathbf{C} 行列は元の行列と同じサイズである。 \mathbf{A} を $ra \times ca$ \mathbf{B} を $rb \times cb$ の行列とする。いま、 $ca = rb$ の場合、積 \mathbf{AB} が定義でき、

$$\mathbf{C} = \mathbf{AB}, \quad c_{ij} = \sum_{k=1}^{ca} a_{ik} b_{kj}, \quad i = 1, 2, \dots, ra, \quad j = 1, 2, \dots, cb \quad (136)$$

であり、 \mathbf{C} 行列のサイズは $ra \times cb$ である。また、 $cb = ra$ の場合、積 \mathbf{BA} が定義でき、

$$\mathbf{C} = \mathbf{BA}, \quad c_{ij} = \sum_{k=1}^{ca} b_{ik} a_{kj}, \quad i = 1, 2, \dots, rb, \quad j = 1, 2, \dots, ca \quad (137)$$

であり、 \mathbf{C} 行列のサイズは $rb \times ca$ である。この場合、スカラー同士の掛け算は順序に依存しないため

$$\mathbf{C} = \mathbf{BA}, \quad c_{ij} = \sum_{k=1}^{ca} a_{kj} b_{ik}, \quad i = 1, 2, \dots, rb, \quad j = 1, 2, \dots, ca \quad (138)$$

と書いても良い。また、 $ra = rb$ の場合、積 $\mathbf{A}'\mathbf{B}$ が定義でき、

$$\mathbf{C} = \mathbf{A}'\mathbf{B}, \quad c_{ij} = \sum_{k=1}^{ca} a_{ki} b_{kj}, \quad i = 1, 2, \dots, ca, \quad j = 1, 2, \dots, cb \quad (139)$$

であり、 \mathbf{C} 行列のサイズは $ca \times cb$ である。すなわち、Eq.(136) のパターンに於いて、転置となっている行列 \mathbf{A} の添字を交換すればよい。また、 $rb = ra$ の場合、積 $\mathbf{B}'\mathbf{A}$ が定義でき、

$$\mathbf{C} = \mathbf{B}'\mathbf{A}, \quad c_{ij} = \sum_{k=1}^{ca} b_{ki} a_{kj} = \sum_{k=1}^{ca} a_{kj} b_{ki}, \quad i = 1, 2, \dots, cb, \quad j = 1, 2, \dots, ca \quad (140)$$

であり、 \mathbf{C} 行列のサイズは $cb \times ca$ である。これも、Eq.(137) の \mathbf{B} の添字を交換してもものとなっている。自分自身の転置との積は常に定義可能で、

$$\mathbf{C} = \mathbf{A}'\mathbf{A}, \quad c_{ij} = \sum_{k=1}^{ra} a_{ki} a_{kj}, \quad i = 1, 2, \dots, ca, \quad j = 1, 2, \dots, ca \quad (141)$$

$$\mathbf{C} = \mathbf{AA}', \quad c_{ij} = \sum_{k=1}^{ca} a_{ik} a_{jk}, \quad i = 1, 2, \dots, ra, \quad j = 1, 2, \dots, ra \quad (142)$$

となるが、 $A'A$ のサイズは $ca \times ca$ 、 AA' のサイズは $ra \times ra$ である。

また、対角行列が入っている場合は、対角行列の掛け算に関する性質を用いて、対角行列の個数だけ総和記号を減らすことが出来る。例えば、 A 行列が対角行列の場合は、 AB は B の各行を A の対角要素倍したものであり、同様に、 B 行列が対角の場合は、 AB は A の各列 B の対角要素倍したものとなる。したがっていずれの場合にも総和記号は必要でなくなり、Eq.(136) は以下ようになる。なお、対角行列を強調するために $\text{diag}()$ の記号を用いている。

$$C = \text{diag}(A)B, \quad c_{ij} = a_{ii}b_{ij} \quad (143)$$

$$C = A\text{diag}(B), \quad c_{ij} = a_{ij}b_{jj} \quad (144)$$

$$C = \text{diag}(A)\text{diag}(B), \quad c_{ij} = \begin{cases} a_{ii}b_{ii} & i = j \\ 0 & i \neq j \end{cases} \quad (145)$$

なお、最後の式は、対角行列同士の積がまた対角行列になることを示している。

行列とベクトルとの積は、上記の公式で、ベクトルに相当する行列が、列ベクトルの場合には行の添字が常に 1、すなわち $ca = 1$ または $cb = 1$ 、行ベクトルの場合には列の添字が常に 1、すなわち $rca = 1$ または $rb = 1$ 、であると考えればよい。例えば、Eq.(136) に於いて B が縦ベクトルの場合、 $cb = 1$ であるから、

$$c = Ab, \quad c_i = c_{i1} = \sum_{k=1}^{ca} a_{ik}b_{k1} = \sum_{k=1}^{ca} a_{ik}b_k \quad (146)$$

となり、 c は $ra \times 1$ の縦ベクトルとなる。Eq.(139) に於いて A が列ベクトルの場合は $ca = 1$ であるから

$$c' = a'B, \quad c_i = c_{1i} = \sum_{k=1}^{ca} a_{k1}b_{kj} = \sum_{k=1}^{ca} a_k b_{kj} \quad (147)$$

となり c は $1 \times cb$ の行ベクトルとなる。

この特殊形として、ベクトルが単位ベクトル $\mathbf{1}$ の場合を考えてみると、

$$c = A\mathbf{1}, \quad c_i = c_{i1} = \sum_{k=1}^{ca} a_{ik}1 = \sum_{k=1}^{ca} a_{ik} \quad (148)$$

$$c' = \mathbf{1}'B, \quad c_i = c_{1i} = \sum_{k=1}^{ca} 1b_{kj} = \sum_{k=1}^{ca} b_{kj} \quad (149)$$

となる。すなわち、行列の行や列の和がえられる。

3つの行列の積、例えば、 $D = ABC$ は、これを $D = A(BC)$ もしくは $D = (AB)C$ と考え、上記の公式を2度適用すればよい。前者の場合、 BC の結果として得られる行列の pq 要素を $[BC]_{pq}$ と書けば、

$$D = A(BC), \quad d_{ij} = \sum_{k=1}^{ca} a_{ik}[BC]_{kj} = \sum_{k=1}^{ca} a_{ik} \left(\sum_{\ell=1}^{cb} b_{k\ell}c_{\ell j} \right) \quad (150)$$

$$= \sum_{k=1}^{ca} a_{ik} \sum_{\ell=1}^{cb} b_{k\ell}c_{\ell j} = \sum_{k=1}^{ca} \sum_{\ell=1}^{cb} a_{ik}b_{k\ell}c_{\ell j} \quad (151)$$

となる。また、後者の場合、 AB の結果として得られる行列の pq 要素を $[AB]_{pq}$ と書けば、

$$D = (AB)C, \quad d_{ij} = \sum_{k=1}^{cb} [AB]_{ik}c_{kj} = \sum_{k=1}^{cb} \left(\sum_{\ell=1}^{ca} a_{i\ell}b_{\ell k} \right) c_{kj} \quad (152)$$

$$= \sum_{k=1}^{cb} \sum_{\ell=1}^{ca} a_{i\ell}b_{\ell k}c_{kj} = \sum_{\ell=1}^{ca} a_{i\ell} \sum_{k=1}^{cb} b_{\ell k}c_{kj} = \sum_{k=1}^{ca} a_{ik} \sum_{\ell=1}^{cb} b_{k\ell}c_{\ell j} \quad (153)$$

となるが、最後の式の展開は単なる添字の付け替えである。

より一般的には、複数の行列の掛け算は、以下の5つの行列の掛け算

$$F = ABCDE \quad (154)$$

$$f_{ij} = \sum_{i1=1}^{ca} \sum_{i2=1}^{cb} \sum_{i3=1}^{cc} \sum_{i4=1}^{cd} a_{i,i1} b_{i1,i2} c_{i2,i3} d_{i3,i4} e_{i4,j} \quad (155)$$

$$= \sum_{i1=1}^{ca} a_{i,i1} \sum_{i2=1}^{cb} b_{i1,i2} \sum_{i3=1}^{cc} c_{i2,i3} \sum_{i4=1}^{cd} d_{i3,i4} e_{i4,j} \quad (156)$$

に示される様なパターンを行列の数により拡大縮小した形となっている。また、いくつかの行列が、転置の形で入っている場合、それらの行列の添字を入れ替えればよい。例えば、

$$\mathbf{F} = \mathbf{A}\mathbf{B}'\mathbf{C}\mathbf{D}'\mathbf{E} \quad (157)$$

$$f_{ij} = \sum_{i1=1}^{ca} \sum_{i2=1}^{rb} \sum_{i3=1}^{cc} \sum_{i4=1}^{rd} a_{i,i1} b_{i2,i1} c_{i2,i3} d_{i4,i3} e_{i4,j} \quad (158)$$

$$= \sum_{i1=1}^{ca} a_{i,i1} \sum_{i2=1}^{rb} b_{i2,i1} \sum_{i3=1}^{cc} c_{i2,i3} \sum_{i4=1}^{rd} d_{i4,i3} e_{i4,j} \quad (159)$$

また、対角行列が入っている場合は、先に述べた対角行列の性質を使って、対角行列の個数だけ総和記号を減らすことが出来る。例えば、 \mathbf{A} 行列と \mathbf{D} 行列が対角行列の場合は、まず、 $\text{diag}(b\mathbf{A})\mathbf{B}$ をまとめることにより、総和記号が一つ減る。次に、 \mathbf{D} 行列であるが、これを \mathbf{E} 行列の左からかかっていると見るか、 \mathbf{C} 行列の右からかかっていると見るかで以下のような二つの同値な結果を得る。

$$\mathbf{F} = \text{diag}(\mathbf{A})\mathbf{B}\mathbf{C}\text{diag}(\mathbf{D})\mathbf{E} = (\text{diag}(\mathbf{A})\mathbf{B})\mathbf{C}(\text{diag}(\mathbf{D})\mathbf{E}) \quad (160)$$

$$f_{ij} = \sum_{i1=1}^{ca} \sum_{i2=1}^{cc} (a_{i,i1} b_{i,i1}) c_{i1,i2} (d_{i2,i2} e_{i2,j}) \quad (161)$$

$$= \sum_{i1=1}^{ca} a_{i,i1} b_{i,i1} \sum_{i2=1}^{cc} c_{i1,i2} d_{i2,i2} e_{i2,j} \quad (162)$$

$$\mathbf{F} = \text{diag}(\mathbf{A})\mathbf{B}\mathbf{C}\text{diag}(\mathbf{D})\mathbf{E} = (\text{diag}(\mathbf{A})\mathbf{B})(\mathbf{C}\text{diag}(\mathbf{D}))\mathbf{E} \quad (163)$$

$$f_{ij} = \sum_{i1=1}^{ca} \sum_{i2=1}^{cc} (a_{i,i1} b_{i,i1}) (c_{i1,i2} d_{i2,i2}) e_{i2,j} \quad (164)$$

$$= \sum_{i1=1}^{ca} a_{i,i1} b_{i,i1} \sum_{i2=1}^{cc} c_{i1,i2} d_{i2,i2} e_{i2,j} \quad (165)$$

5 総積記号

総積記号とは、一般的に

$$\prod_{sub=from}^{to} body \quad (166)$$

の形をしており、これは、任意の数式 $body$ に含まれる添字 sub を $from$ から to まで順に変えることにより得られる数式を、全て掛け合わせることを示している。したがって、対数を取ると $body$ に関する総和記号の演算

$$\log \left(\prod_{sub=from}^{to} body \right) = \sum_{sub=from}^{to} \log(body) \quad (167)$$

となる。したがって、総和記号に関する全ての公式を使い上式の右辺を計算し、それをを用いて

$$\prod_{sub=from}^{to} body = \exp \left(\sum_{sub=from}^{to} \log(body) \right) \quad (168)$$

とすればよい。

以下に総積記号に対応する公式を挙げる。

- 定数

$$\prod_{scope} body = body^{scope \text{ に含まれる要素数}} \quad (169)$$

ただし、 $body$ は $scope$ に記された添字を含まない。

- 分配の法則

$$\prod_{scope} (body1 \times body2) = \left(\prod_{scope} body1 \right) \left(\times \prod_{scope} body2 \right) \quad (170)$$

これらを利用すると以下のようなことが解る。

- 定数

$$\prod_i^n c = c^n \quad (171)$$

- 総積記号の分配

$$\prod_i (x_i \times y_i) = \left(\prod_i x_i \right) \times \left(\prod_i y_i \right) \quad (172)$$

- 定数倍

$$\prod_i^n cy_i = c^n \prod_i y_i \quad (173)$$

- 総積記号の移動

$$\prod_{i=1}^A \prod_{j=1}^B x_i y_j = \left(\prod_i x_i \right)^B \left(\prod_j y_j \right)^A \quad (174)$$