

## 一般化線形モデルにおける交互作用：交互作用項があるときの主効果の回帰係数の意味について

2010. 3. 11

粕谷英一

### 交互作用とは

回帰や分散分析における交互作用 (interaction) とは、ある説明変数が目的変数に与える効果が、他の説明変数がどんな値をとるかによって異なることである。一般化線形モデル (GLM) では、交互作用の項を説明変数に含めるのは普通の使われ方であるが、その際の解釈を少し整理してみる。

2 説明変数の線形重回帰を考えてみる。回帰式は

$$y=b1\cdot x1+b2\cdot x2+a$$

と書ける。ここで、 $y$  は目的変数 (応答変数)、 $x1$  と  $x2$  はそれぞれ第 1 と第 2 の説明変数である。 $a$  は定数項である。 $x2$  を一定にしておいて、 $x1$  が 1 増えると  $y$  は  $b1$  増加する。このことは、 $x2$  の値がどんなものであっても、変わらない。言い換えると、説明変数  $x1$  の値が目的変数  $y$  に与える効果は、他の説明変数である  $x2$  がどんな値をとってしようと一定であり変わらない。つまり交互作用はないことになる。同様に  $x1$  を一定にしておいて、 $x2$  が 1 増えると  $y$  は  $b2$  増加し、やはり  $x1$  の値にはよらない。つまり、線形重回帰の式は、交互作用がない場合を表現していることがわかる。

### 一般化線形モデルでの交互作用の入れ方

では、回帰の拡張とも見ることが可能な、一般化線形モデル (GLM) では、どのようにして交互作用を扱っているのだろうか。2つの説明変数の間の交互作用は、説明変数の積として扱われている。たとえば、identity リンクで 2 説明変数の場合、それぞれの説明変数そのものと交互作用を説明変数にしているなら、考えている回帰式は、

$$y=b1\cdot x1+b2\cdot x2+c\cdot(x1\cdot x2)+a$$

となる。交互作用の項の回帰係数 (上記の例なら  $c$ ) に対して、説明変数そのものの回帰係数 ( $b1$  と  $b2$ ) は主効果 (main effect) と呼ばれることが多い。ここでは、主効果の回帰係数と呼ぶ。

以下、2つの説明変数、identity リンクで等分散の正規分布を例として使っていく。なお、以下、誤差に関する内容はほとんど出て来ず、また、主な内容

はリンクには関係がない。つまり他のリンクや誤差分布でも同様のことが成り立つと考えられる。なお、“交互作用の項が入っているときに主効果の回帰係数の意味するもの”がこの文章の主題なので、

$$y=c \cdot x1 \cdot x2+a$$

のような交互作用項のみのモデルはここでは扱わない。

### GLMにおいて交互作用があるときの主効果の意味は何か

2つの説明変数で、主効果と交互作用の項が入る場合なので、扱う回帰式は、

$$y=b1 \cdot x1+b2 \cdot x2+c \cdot (x1 \cdot x2)+a$$

である。交互作用の項がない場合には、両説明変数と目的変数の3次元空間では、回帰式は平面を表している。では、3次元空間にこの主効果と交互作用の項が入っているモデル（回帰式）を図示したらどんなものになるのだろうか。片方の説明変数の値を固定すると、 $y$ はもう片方の説明変数の一次式であらわされることから、直線を動かしていったときにできる線織面（ruled surface）であることがわかる。次に2次元局面の分類の理論にしたがって、上の式の係数を検討してみると、線織面の中でも双曲放物面になることがわかる。つまり、交互作用を含む上記の回帰式は、2つの説明変数と目的変数の、3次元の空間では双曲放物面を表し、この回帰式をデータにあてはめるとは、もっともよくあう（対数尤度最大などの基準で）双曲放物面を求めていることである。

この式で、他の説明変数は一定で  $x1$  や  $x2$  が1増加したときに  $y$  に与える効果を見ると

$$x1 \text{ については } b1 + c \cdot (x2)$$

$$x2 \text{ については } b2 + c \cdot (x1)$$

となり、主効果のみの場合、つまり  $y=b1 \cdot x1+b2 \cdot x2+ a$  における

$$x1 \text{ については } b1$$

$$x2 \text{ については } b2$$

のようなシンプルさは無くなっている。交互作用がある場合なので、他の説明変数の値がどうなっているかを決めてやらないと、問題としている説明変数の変化が目的変数に与える効果が決まらないのである。逆にいえば、他の説明変数の値がどんなものであるときに、問題としている説明変数の効果を見るのかを決めればよいことになる。

では、式  $y=b1 \cdot x1+b2 \cdot x2+c \cdot (x1 \cdot x2)+a$  の  $b1$  や  $b2$  は、他の説明変数がどんな値の時に、問題の説明変数の効果を見ることに対応しているのだろうか。上記のように説明変数が1増加したときに  $y$  に与える効果を見ると

$x_1$  については  $b_1 + c \cdot (x_2)$

$x_2$  については  $b_2 + c \cdot (x_1)$

である。ここから、他の説明変数が 0 のときに問題の説明変数の効果を見ることが主効果の回帰係数に対応していることがわかる。つまり、 $b_1$  については、 $x_2=0$  のときに  $x_1$  が 1 増えた時に  $y$  に与える効果  
 $b_2$  については、 $x_1=0$  のときに  $x_2$  が 1 増えた時に  $y$  に与える効果  
 である（ここでは、とくに断らない限り、他の説明変数  $x_2$  は一定で  $x_1$  が 1 増加したときに  $y$  に与える効果のことを、 $x_1$  の効果ということにする）。

以下、説明変数が量的なものかそれとも名義変数かにより、4つの場合に分けて、説明変数  $x_1$  の（主効果の）回帰係数  $b_1$  の意味するものを見ていく。名義変数の場合は 2 値的であるとする。2 値的変数は、1 つのダミー変数（値が 1 か 0）として扱われることが多く、ここでもそのように扱っていくが、他の扱い方（たとえば線形対比など）があることを忘れてはいけない。なお、計算例は、すべて R2. 10. 0 による。

### 【1】 $x_1$ が量的な変数で、 $x_2$ が名義変数の場合

$x_2$  は 1 か 0 のどちらかの値である。 $x_2$  がある特定の値をとるデータをここでは、サブグループと呼ぶことにする。いまのケースではデータ全体は、2 つのサブグループからなっており、片方は  $x_2=1$  でもう 1 つは  $x_2=0$  である。

式  $y=b_1 \cdot x_1 + b_2 \cdot x_2 + c \cdot (x_1 \cdot x_2) + a$  は、

$x_2=0$  のとき、

$$y=b_1 \cdot x_1 + a$$

$x_2=1$  のとき、

$$y=b_1 \cdot x_1 + b_2 + c \cdot x_1 + a = (b_1 + c) \cdot x_1 + (b_2 + a)$$

となる。それぞれのサブグループについての式を見ると、 $b_1$  と  $a$  をサブグループ  $x_2=0$  のデータにもっともよく合うように決めてしまっても、 $y=b_1 \cdot x_1 + a$  に含まれていない  $c$  や  $b_2$  の値を変えることで、サブグループ  $x_2=0$  のデータにもっともよく合うように  $y=b_1 \cdot x_1 + b_2 + c \cdot x_1 + a$  を決めてやることができることがわかる。そこで、 $b_1$  とは、サブグループ  $x_2=0$  のデータのみについて、説明変数が  $x_1$  のみで分析したときの  $x_1$  の回帰係数に他ならないだろうと想像される。ここでは省略するが、正規方程式を立てて連立一次方程式を解くと（identity リンクで等分散の正規分布なので、最小二乗法によるこの求め方で最尤法と同じ解になる）、その想像が一般的に正しいことが確かめられる。

以下、数値例を見てみよう。まず、データである。x2 がデータフレーム mydata1 の fx20 にあたる。目的変数 y が mydata1 の y1 にあたる。

```
> mydata1
  x1 y1 fx20
1  1.1 2.0  a
2  1.5 1.5  b
3  2.0 1.9  a
4  2.1 3.0  b
5  2.4 2.1  a
6  3.0 4.0  b
7  3.5 2.1  a
8  4.1 5.1  b
9  5.0 1.9  a
10 5.4 6.0  b
```

fx20 が a のデータと b のデータ(サブグループ)に分けて分析してみる。

```
> glm(y1~x1,data=mydata1,family=gaussian,subset=(fx20=="a"))
```

```
Call:  glm(formula = y1 ~ x1, family = gaussian, data = mydata1, subset = (fx20 == "a"))
```

Coefficients:

```
(Intercept)      x1
  2.03415      -0.01220
```

```
> glm(y1~x1,data=mydata1,family=gaussian,subset=(fx20=="b"))
```

```
Call:  glm(formula = y1 ~ x1, family = gaussian, data = mydata1, subset = (fx20 == "b"))
```

Coefficients:

```
(Intercept)      x1
  0.3842         1.0981
```

データ全体を使い、交互作用と両説明変数の主効果を入れて分析すると、

```
> glm(y1~x1*fx20,data=mydata1,family=gaussian)
```

```
Call:  glm(formula = y1 ~ x1 * fx20, family = gaussian, data = mydata1)
```

Coefficients:

(Intercept)	x1	fx20b	x1:fx20b
2.03415	-0.01220	-1.64996	1.11027

で x1 の回帰係数は、fx20 が a の場合のみのデータを使った場合と同じである。

まとめると、x1 の主効果の回帰係数は、x2=0 であるデータについてのみ、説明変数が x1 だけの分析を行ったときの x1 の回帰係数と同じであった。サンプルのうち一部のデータだけを使った時の x1 の効果を、サンプル全体についての x1 の主効果として知りたいというケースはおそらくまれだろう。そういうときには、サンプルの一部のデータ（つまりこの例では x2=0 であるデータ）を使って分析すればすむことであって、しかもその方が自然であり、おそらく多くの研究者はそうするからである。すなわち、ごく大部分の場合、交互作用を含むモデルの主効果の回帰係数として GLM が返してくるものは、研究者が期待するものとはかなりちがうであろう。

さて、主効果の回帰係数として得られる結果は、片方のサブグループのデータのみを分析した（当然、他の説明変数は含まず、説明変数は x1 で）ときの回帰係数と同じである。では、どちらのサブグループを分析したものになるのかと言え、x2=0 の方であるから、どちらのサブグループに x2=0 を割り振るのかにより、まったくちがうデータを分析することになり、当然、結果は異なるだろう。

実際に分析をすると、名義変数の名前の付け方などにより、2つのサブグループのどちらを分析した結果になるかが異なることになる。たとえば R では文字列をソートした時の順番で 0, 1 と割り振られる。aka と midori にすれば、aka に 0 が midori に 1 が割り振られるが、red と green なら green に 0 が red に 1 が割り振られる（なお、同じ文字列を使ってもロケールにより異なる可能性がある）。どちらに 0 が割り振られているのかを確認するのは分析者の責任である。

## 【2】x1 も x2 も量的な変数の場合

この【2】の場合が他の3つをも包含する一般的なケースではあるが、他の説明変数（ここでは x2）が名義変数か連続的な変数かにより、異なる点も多い。

この【2】の場合を理解するには、図 1 のように、縦軸に x1 が 1 増加した時に y に与える効果、横軸に x2 をとったプロットを考えてみるとよい。回帰式、

$y=b_1 \cdot x_1+b_2 \cdot x_2+c \cdot x_1 \cdot x_2+a$  を変形すると

$y=(b_1+c \cdot x_2)x_1+b_2 \cdot x_2+a$  であるから、交互作用があるときのモデルは図1のプロットでは直線であらわされ、この直線の切片の値つまり  $x_2=0$  のときに  $x_1$  が1増加することが  $y$  に与える効果が  $b_1$  であることがわかる。別の言い方をすると、交互作用と両変数の主効果を入れたモデルでは、図1のプロットが直線になることを仮定している。

既に述べた、【1】のような、 $x_2$  が名義変数の場合、ダミー変数を使うと、 $x_2=0$  となるデータ点は、データ全体から見ると一部ではもちろんあるが、必ずデータの中に存在する。だが、量的な変数の場合には  $x_2=0$  となるデータ点は存在しないこともある。長さ、質量、体積などでは、負のデータは普通存在しないし、値が0のデータは登場しないことも多いだろう。

つまり、【1】ではまだしも、 $x_2$  が実際に観察されている値の時に  $x_1$  の効果を見ているのだが、この【2】ではそうとは限らない。 $x_2$  が正の値だけという場合には、 $x_2=0$  という場合は実際にはない。図1のようなプロットを使ってみると、 $x_1$  の主効果は【1】ではデータのうち一部だけみたときの効果だったが、この【2】で  $x_2$  が正の値だけという場合には、実際には  $x_2$  の値が観察されていないところにまで直線を伸ばして行って  $x_2=0$  の時の縦軸の値を読んだものであることがわかる。これは一種の外挿(補外)と見ることができる。 $x_2=0$  のときの効果ということに研究上明確な意味がある場合を除けば(そのような例が多いとは思えないが)、 $x_1$  の主効果として分析する側が期待しているものと主効果の回帰係数の意味するところが食い違う度合いは、【1】をさらに上回るように思われる。

外挿(補外)のようになるのを避ける対策として、(標準化のように)平均値を0にするように予めしておいて、 $x_2=0$  の両側にデータ点があるようにしておくことも考えられる。しかしこの操作ではうまくいかない場合があることは分散分析の方法をめぐる議論ではすでに指摘されている。このことは次の【3】で述べる。

以下、数値例を見ていく。まず、データであるが、目的変数  $y$  が  $y_1$  にあたる。なお、 $x_2$  は、2,3,4 のいずれかの値をとっている。

```
> testdata1
  x1 x2  y1
1  1  2  5.97
2  2  2  6.74
3  3  2  7.51
4  4  2  8.28
5  5  2  9.05
6  6  2  9.82
```

```

7 7 2 10.59
8 8 2 11.36
9 1 3 7.23
10 2 3 8.16
11 3 3 9.09
12 4 3 10.02
13 5 3 10.95
14 6 3 11.88
15 7 3 12.81
16 8 3 13.74
17 1 4 8.49
18 2 4 9.58
19 3 4 10.67
20 4 4 11.76
21 5 4 12.85
22 6 4 13.94
23 7 4 15.03
24 8 4 16.12

```

まず、 $x_2$  の値によってデータを分けて、分析してみる

```
> glm(y1~x1,family=gaussian,data=testdata1,subset=(x2==2))
```

```
Call: glm(formula = y1 ~ x1, family = gaussian, data = testdata1, subset = (x2 == 2))
```

Coefficients:

```
(Intercept)      x1
          5.20      0.77
```

```
> glm(y1~x1,family=gaussian,data=testdata1,subset=(x2==3))
```

```
Call: glm(formula = y1 ~ x1, family = gaussian, data = testdata1, subset = (x2 ==3))
```

Coefficients:

```
(Intercept)      x1
          6.30      0.93
```

```
> glm(y1~x1,family=gaussian,data=testdata1,subset=(x2==4))
```

```
Call: glm(formula = y1 ~ x1, family = gaussian, data = testdata1, subset = (x2 ==4))
```

Coefficients:

```
(Intercept)      x1
```

7.40      1.09

次に、データ全体を使い、交互作用も入れた場合である、

```
> glm(y1~x1*x2,family=gaussian,data=testdata1)
```

```
Call: glm(formula = y1 ~ x1 * x2, family = gaussian, data = testdata1)
```

Coefficients:

(Intercept)	x1	x2	x1:x2
3.00	0.45	1.10	0.16

x1 の主効果の係数は 0.45 で、x2 がどの値の場合の場合よりも小さくなっている。この数値例では、x1 の回帰係数は、x2 が 4, 3, 2 と減るにしたがい、1.09, 0.93, 0.77 と減っており、データ全体を使い交互作用も入れた場合の 0.45 は、このペースで直線的に減少して x2=0 のときとほぼ同じである。上で述べたように、外挿した場合にあたっていることがうかがえる。

次の例は、いまの例とそっくりのデータで、ただ x2 がみな 2 小さくなっていることだけがちがっている。つまり、すぐ上の例の x2 から 2 を引いたものが次の例の x2 である。

```
> testdata2
```

	x1	x2	y1
1	1	0	5.97
2	2	0	6.74
3	3	0	7.51
4	4	0	8.28
5	5	0	9.05
6	6	0	9.82
7	7	0	10.59
8	8	0	11.36
9	1	1	7.23
10	2	1	8.16
11	3	1	9.09
12	4	1	10.02
13	5	1	10.95
14	6	1	11.88
15	7	1	12.81
16	8	1	13.74



```

17 1 2 8.49
18 2 2 9.58
19 3 2 10.67
20 4 2 11.76
21 5 2 12.85
22 6 2 13.94
23 7 2 15.03
24 8 2 16.12

```

まず、 $x_2$  の値によってデータを分けて、分析してみる。

```
> glm(y1~x1,family=gaussian,data=testdata2,subset=(x2==0))
```

```
Call: glm(formula = y1 ~ x1, family = gaussian, data = testdata2, subset = (x2==0))
```

Coefficients:

```
(Intercept)      x1
          5.20      0.77
```

```
> glm(y1~x1,family=gaussian,data=testdata2,subset=(x2==1))
```

```
Call: glm(formula = y1 ~ x1, family = gaussian, data = testdata2, subset = (x2==1))
```

Coefficients:

```
(Intercept)      x1
          6.30      0.93
```

```
> glm(y1~x1,family=gaussian,data=testdata2,subset=(x2==2))
```

```
Call: glm(formula = y1 ~ x1, family = gaussian, data = testdata2, subset = (x2==2))
```

Coefficients:

```
(Intercept)      x1
          7.40      1.09
```

当然のことだが、1つ前の例の、 $x_2=2, 3, 4$  の例の  $x_1$  の回帰係数が、この例の  $x_2=0, 1, 2$  の場合とそれぞれ同じである。

次に、データ全体を使い、交互作用も入れた場合である、

```
> glm(y1~x1*x2,family=gaussian,data=testdata2)
```

```
Call: glm(formula = y1 ~ x1 * x2, family = gaussian, data = testdata2)
```

Coefficients:

```
(Intercept)      x1      x2      x1:x2
```

5.20      0.77      1.10      0.16

$x_1$  の主効果の回帰係数は、 $x_2=0$  の場合（つまり 1 つ前の例では  $x_2=2$  の場合に  
あたる）と同じである。主効果として返されるのが、他の説明変数が 0 である  
ときに見た効果であるということからすれば当然の結果ではある。だが、この  
2 つの例の比較から、他の説明変数に定数を加えたり引いたりするような操作  
をすると、ある説明変数の主効果として返されるものが変わってしまうことが  
わかる（定数の増減をされる変数と主効果に影響が出る変数は別）。距離、温  
度（あるいは時間でも）など、自明な 0 を持たない間隔尺度の変数では定数だ  
けの変化はよく起こりそうなことである。直線回帰や線形の重回帰では、変数  
の値への定数の増減は回帰係数には影響しなかった。交互作用の項を含む場合  
には、そのようなシンプルさは失われる。

### 【3】 $x_1$ も $x_2$ も名義変数の場合

このケースは、【1】で  $x_1$  を 1 か 0 しかとらない 2 値的な変数とした特殊な  
場合にあたり、【1】で述べたことはそのまま当てはまる。

また、この【3】のケースは、分散分析で過去、多くの研究があり、論争が  
あったことでも知られる（分散分析における論争については、いくつもの異なる  
意見があり、決着を見ていない）。論争の焦点は、ここでの言葉で言えば、  
“ $x_2$  の値がどんなときにおける  $x_1$  の効果を、 $x_1$  の主効果とするか” である。  
データを得る前から決めているある特定の  $x_2$  の値（つまり 0 か 1）ということ  
も 1 つの答えとして考えられるであろうし、 $x_2$  の平均ということも考えられる  
だろう。だが、どちらにも難点はある。特定の値とすると、他の値の時は  $x_1$  の  
主効果の大きさはちがうのにそれは無視することになる。また、 $x_2$  の平均とす  
ると、サブグループごとの得られたデータの数により、見ている内容つまり仮  
説の中身が異なる可能性がある（ $x_2=0$  のデータと  $x_2=1$  のデータの相対的な数に  
より見ている効果の内容は異なる一量的変数で平均を使おうとする場合にも同  
じことはあてはまる）。そもそも、交互作用があるということは  $x_2$  の値により  
 $x_1$  の効果の大きさは異なるのだから、そういうときには  $x_1$  の効果を見ること  
をやめるという考え方もある（Cochran など）。

分散分析での過去の議論を振り返ると、“交互作用があるときの主効果”  
とは分散分析においてすでに困難さが認識されている問題を、（分散分析では

名義変数だったが、GLM や一般線形モデルでは量的な変数もあるから) 少し拡張した問題ととらえることができる。

以下、数値例である。まず、データであるが、目的変数 $y$ が  $y_1$  にあたる。説明変数  $x_1$  が  $f1a$ 、 $x_2$  は  $f1p$  で、それぞれ2水準の名義変数である。

```
> mydata0
  y1 f1a f1p
1 0.0  a  p
2 1.1  a  p
3 2.2  a  p
4 0.1  a  q
5 1.0  a  q
6 0.9  a  q
7 1.2  b  p
8 2.2  b  p
9 3.0  b  p
10 3.1 b  q
11 4.0 b  q
12 5.1 b  q
```

$x_2$  の値によってデータを分けて、分析してみる。

```
> glm(y1~f1a,data=mydata0,family=gaussian,subset=(f1p=="p"))
```

```
Call:  glm(formula = y1 ~ f1a, family = gaussian, data = mydata0, subset = (f1p == "p"))
```

Coefficients:

(Intercept)	f1ab
1.100	1.033

```
> glm(y1~f1a,data=mydata0,family=gaussian,subset=(f1p=="q"))
```

```
Call:  glm(formula = y1 ~ f1a, family = gaussian, data = mydata0, subset = (f1p == "q"))
```

Coefficients:

(Intercept)	f1ab
0.6667	3.4000

次にデータ全体を使って分析し、交互作用の項も入れた場合である。

```
> glm(y1~f1a*f1p,data=mydata0,family=gaussian)
```

Call: `glm(formula = y1 ~ f1a * f1p, family = gaussian, data = mydata0)`

Coefficients:

(Intercept)	f1ab	f1pq	f1ab:f1pq
1.1000	1.0333	-0.4333	2.3667

f1a (つまり x1) の主効果の回帰係数は、【1】と同じく、他の説明変数 (ここでは f1p) が 0 であるデータのみ、つまり f1p が p であるデータのみを使って、説明変数は f1a だけで分析したときと同じである。

#### 【4】 x1 が名義変数で、x2 が量的な変数の場合

【1】とちょうど対称に、x1 の値により、データを 2 つのサブグループに分けてみると、回帰式は、

x1=1 の方のサブグループについては、

$$(b2+c) \cdot x2 + b1 + c = y$$

x1=0 の方のサブグループについては、

$$(b2) \cdot x2 + c = y$$

となる。x2 と y のグラフを描くとそれぞれ直線になる (図 2)。【1】の場合と同様に、それぞれのサブグループにもっともよく合う直線をそれぞれ決めることができる。x1 が 1 増加した時の y に与える効果は、図 2 の 2 本の直線の、縦 (y) 方向の差である。交互作用があるとは、2 本の直線が平行ではないということである。交互作用の項も入れたモデルでの x1 の主効果の回帰係数は、x2=0 のところで 2 本の直線の高さの差である。

やはり x2 が正の値ばかりのときには、存在しない値のところでは x1 の効果を見ることになり、いわば外挿のようになるという点については、【2】と同様である。そこで、図 3 のように、x2 の最小値と 0 の間で 2 本の直線が交差する場合には、実際に見られている x2 の範囲全体で、x1 が 1 である方が x1 が 0 であるよりも y は大きくても、x1 の主効果の回帰係数はそれと反対のことを示すことも起こりうる。

この【4】は、【3】の両方とも量的な変数の場合という一般的なケースの、特殊なものと考えてもよい。

余談: 交互作用の項で表現できるもの

交互作用は他の説明変数の値により問題としている説明変数が応答変数(目的変数)に与える効果が異なることである。だが、交互作用を説明変数の積で扱うことでは、そのうち、一部しか表現できない。以下は、両方の説明変数が量的である場合の数値例である。この例でも、x2の値は、2, 3, 4のいずれかである。

```
> testdata21
  x1 x2  y1
1  1  2  5.5
2  2  2  5.8
3  3  2  6.1
4  4  2  6.4
5  5  2  6.7
6  6  2  7.0
7  7  2  7.3
8  8  2  7.6
9  1  3  6.8
10 2  3  7.3
11 3  3  7.8
12 4  3  8.3
13 5  3  8.8
14 6  3  9.3
15 7  3  9.8
16 8  3 10.3
17 1  4  7.7
18 2  4  8.0
19 3  4  8.3
20 4  4  8.6
21 5  4  8.9
22 6  4  9.2
23 7  4  9.5
24 8  4  9.8
```

まず、x2の値によってデータを分けて、分析してみる。

```
> glm(y1~x1,family=gaussian,data=testdata21,subset=(x2==2))
Call:  glm(formula = y1 ~ x1, family = gaussian, data = testdata21, subset = (x2 == 2))
Coefficients:
```

```
(Intercept)      x1
              5.2    0.3
```

```
> glm(y1~x1,family=gaussian,data=testdata21,subset=(x2==3))
```

```
Call:  glm(formula = y1 ~ x1, family = gaussian, data = testdata21, subset = (x2==3))
```

```
Coefficients:
```

```
(Intercept)      x1
              6.3    0.5
```

```
> glm(y1~x1,family=gaussian,data=testdata21,subset=(x2==4))
```

```
Call:  glm(formula = y1 ~ x1, family = gaussian, data = testdata21, subset = (x2== 4))
```

```
Coefficients:
```

```
(Intercept)      x1
              7.4    0.3
```

データ全体を使い、交互作用と両説明変数の主効果を入れて分析すると、

```
> glm(y1~x1*x2,family=gaussian,data=testdata21)
```

```
Call:  glm(formula = y1 ~ x1 * x2, family = gaussian, data = testdata21)
```

```
Coefficients:
```

```
(Intercept)      x1      x2      x1:x2
 3.000e+00  3.667e-01  1.100e+00 -6.835e-18
```

となる。

この例では、 $x_2$  の値の減少（あるいは増加）に伴い、 $x_1$  が  $y$  に与える効果が傾向的に変化していない。 $x_2=2$  では 0.3、 $x_2=3$  では 0.5 と大きくなっているが  $x_2=4$  ではまた 0.3 と小さくなっている。図 1 のようなプロットを引き、そこによくあった直線を（無理やりだが）引けば、水平になる。その直線が  $x_2=0$  のときの  $x_1$  の効果の値は、水平なので、交互作用も入れて全体のデータで分析した時の  $x_1$  の主効果の係数は、3つの場合 ( $x_2=2, 3, 4$ ) の間にくるだろう。実際に、0.3667 となっている。一見データ全体での平均的な効果を反映しているように見えるかもしれないが、このプロットでは、3つの点は直線からかなり離れている。このデータでは、問題としている説明変数 ( $x_1$ ) が応答変数 (目的変数、 $y$ ) に与える効果は、他の説明変数 ( $x_2$ ) の値により異なるのだが、交互作用項を入れたモデルが表す関係とは異なっているためにうまく表現できていない。交互作用項を入れたモデルは、説明変数の効果が他の説明変数の値に依存するという意味での交互作用の特定のパターンを表すことができるにすぎないのである。

## まとめ

以上のことから、交互作用項を含んでいるときの主効果の回帰係数は、その説明変数自体の何か一般的な効果とか交互作用の効果を除いた説明変数自体の純粋な効果といったものではないことがわかる。交互作用項を含んでいるときの主効果の回帰係数は、他の説明変数の値が0のときの問題の説明変数の効果の予測値とでも言うべきものであり、他の説明変数が名義変数のときには一部のデータのみを用いて問題の説明変数だけを使って分析した結果と同じであり、他の説明変数が正の値をとる量的な変数のときには、実際の値の範囲を超えて他の説明変数が0のところだったらどうなるかをというものである。いずれも、研究者が説明変数の主効果として期待するものと一致することは少ないと思われる。

おそらく、“説明変数の主効果”という名のもとにデータを分析する人が見たい内容はいくつかの異なったものを含んでいる。データを分析する人が“説明変数の主効果”として見たいものが、本人にとってはっきりしている場合には、主効果という名前にとらわれず、分析方法を探すのが早道である。実は既存のよく知られた方法が想定している状況に該当することもある。その際、“説明変数の主効果”として人が期待するものは、実際には驚くほど様々であるかもしれないということを頭に置き、（交互作用と主効果が説明変数に入っているなど）形式的に似た他の例の分析や意味付けを鵜呑みにしないことが大切だろう。大きな誤りを避けるうえで、（定石ではあるが）まず、グラフを何通りか描いてみるのが大切だ。

自分で（あるいは他人でも）交互作用の項も入れて“説明変数の主効果”を使って分析したくなったら、質問（自分なら自問）してみるといいと思う：「あなたがその説明変数の主効果として見たいものは何か説明していただけますか」あるいは名義変数が登場していれば「一部のデータだけで分析した時と同じことですがそれでいいのですか」また量的な変数なら「他の説明変数が0のときに効果を見る理由は何ですか」と。

