

課題：awk/perl によるデータ処理と gnuplot によるグラフ描画

学籍番号 035764C : 若津 大悟

平成 16 年 4 月 26 日

1 まえがき

実験によって得られたデータを Unix 上で加工し, 統計量やグラフを PostScript 形式で作成し, 結果を \LaTeX に取り込むための方法を習得する.

2 awk スクリプトで求める回帰直線

2.1 与えられたデータに対する回帰直線を求める awk スクリプト

<<<<<< awk プログラムソース >>>>>>

```
#!/usr/bin/awk -f

/^[0-9]/{
    Tx += $1;      #身長の総和を求める
    Ty += $2;      #体重の総和を求める
    Txy+= $1*$2;   #身長と体重を掛けた値の総和
    xi += $1^2;    #身長を 2 乗した値の総和を xi とする
    count++;       #1 づつカウントしていく
}
END{
    S_x2 = 1 / count * xi - 1 / (count^2) * (Tx^2);
    S_xy = 1 / count * Txy - 1 / (count^2) * Tx * Ty;
    a = S_xy / S_x2;
    b = Ty / count - a * (Tx/count);
    print"a:"a",b:"b;  #a,bを表示する
}
```

<<<<<< awk プログラムによる実行結果 >>>>>>

```
% awk -f j03064.awk example1.dat
a:0.763337,b:-63.8985
```

2.2 gnuplot を用いて実データと回帰直線をプロットしたグラフ

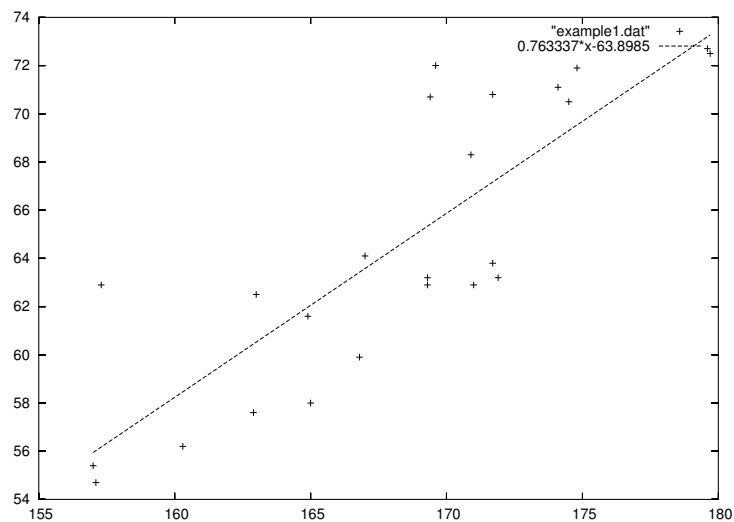


図 1: 回帰直線

3 perl または C 言語を用いて今回の実験と同様の回帰直線を求めるプログラムを作成せよ。

3.1 perl によるプログラム

```
<<<<<< perl プログラムソース >>>>>>
#!/usr/local/bin/perl

open(IN,"example1.dat");
while($x = <IN>){
    @DATA=split(" ", $x, 2);
```

```

    $Tx += $DATA[0];
    $Ty += $DATA[1];
    $Txy += $DATA[0]*$DATA[1];
    $xi += $DATA[0]*$DATA[0];
    $n++;
}

$S_x2 = $xi/$n - $Tx*$Tx/($n*$n);
$S_xy = $Txy/$n - $Tx*$Ty/($n*$n);

$a = $S_xy / $S_x2;
$b = $Ty / $n - $a * ($Tx / $n);

print "n=$n\n";
print "a=$a\n";
print "b=$b\n";

close(IN);

```

<<<<<< perl による実行結果 >>>>>>

```

% perl j03064.pl
n=24
a=0.763336518732883
b=-63.8984804940987

```

3.2 C 言語によるプログラム

<<<<<< C 言語プログラムソース >>>>>>

```

#include<stdio.h>

struct date{
    float height;
    float weight;
};

float Tx,Ty,Txy,S_x2,S_xy,xi,a,b;
int n=0;

```

```

main(){
    struct date sb[23];
    FILE *fp;
    char buf[256];

    fp=fopen("example1.dat","r");
    while(fgets(buf, 256, fp) != NULL){
        sscanf(buf, "%f%f",
            &sb[n].height,&sb[n].weight);

        Tx += sb[n].height;
        Ty += sb[n].weight;
        Txy += sb[n].height*sb[n].weight;
        xi += sb[n].height*sb[n].height;
        n++;
    }
    fclose(fp);

    S_x2 = xi / n - Tx*Tx / (n*n);
    S_xy = Txy / n - Tx*Ty / (n*n);

    a = S_xy / S_x2;
    b = Ty/n - a*(Tx/n);

    printf("a=%f\n",a);
    printf("b=%f\n",b);
}

```

<<<<<< C 言語による実行結果 >>>>>>

```

% a.out
a=0.763402
b=-63.909569

```