確率と統計（O）「Octaveを用いた計算機実習」*

東京工業大学 計算工学専攻
杉山 将
sugi@cs.titech.ac.jp　*http://sugiyama-www.cs.titech.ac.jp/~sugi/*

1　Octaveの起動

左下の Finder アイコンをクリックし，開いたウインドウのアプリケーションのところにある octave.app をクリックすれば，octave が起動する．まずは Octave を電卓のように使ってみよう．

```
> (1+2)*(3-4)*5^2
ans = -75
> exp(10)*log(5)*sin(0.5*pi)
ans =  3.5450e+04
> exit
```

exit は Octave を終了するコマンドである．

2　ベクトル，行列の演算

Octave の特徴は，ベクトルや行列を直接扱うことができるところである．ベクトルや行列は四角括弧で定義する．横方向の要素はスペースで区切り，縦方向の要素はセミコロンで区切る．

```
> A=[1 2;3 4]
A =
  1  2
  3  4

> b=[3;4]
b =
  3
  4
```

ベクトルや行列の演算はそのままスカラーの場合と同じように行う．

```
> A*b
ans =
  11
  25
```

行末にセミコロンをつけると値を表示しなくなる．

```
> C=A*b;
```

単に変数名を入力すれば，変数の値が表示される．

```
> C
C =
  11
  25
```

ベクトルや行列の転置はアポストロフィをつける．

```
> C'
ans =
  11  25
```

行列の要素を取り出すときは，丸括弧で要素を指定する．

```
> A(2,1)
ans = 3
```

また，行列の要素に直接値を代入することもできる．

```
> A(2,1)=5
A =
  1  2
  5  4
```

一列（一行）の要素を全て取り出したいときはコロンを用いる．

```
> A(:,1)
ans =
  1
  5
```

ベクトルや行列のスカラ倍は，そのまま記述する．

```
> D=3*A
D =
   3   6
  15  12
```

行列の要素同士の掛け算や割り算は以下のように記述する．

```
> A.*D
ans =
   3  12
  75  48

> A./D
ans =
  0.33333  0.33333
  0.33333  0.33333
```

数列のベクトルを作りたいときは以下のようにする．

*この資料を作成するに当たり，東京農工大学の田中聡久先生に見せて頂いた資料を参考にした．Octave のより詳しい説明は，例えば *http://www.gnu.org/software/octave/* や *http://www.obihiro.ac.jp/~suzukim/masuda/octave/* を参照せよ．

```
> e=10:5:30
e =
  10   15   20   25   30
```

2 番目の引数を省略すれば，公差は 1 になる．

```
> f=1:4
f =
   1   2   3   4
```

3 数値演算関数

Octave には沢山の数値演算のための関数が組み込まれている．例えば，

```
> cos(2/3*pi)
ans = -0.50000
```

引数にベクトルや行列を指定すれば，要素ごとにその関数の値を計算する．

```
> sin(f)
ans =
   0.84147   0.90930   0.14112  -0.75680
```

他には例えば，sin, cos, tan, acos, atan, tanh, exp, log, sqrt などがある．各関数の説明は help で参照できる．

```
> help exp
```

また，詳細は help -i で info ファイルを参照できる．

```
> help -i exp
```

ベクトル要素の最大値，最小値，和，積は，max, min, sum, prod により得ることができる．

```
> max([1 3 5 2 4])
ans = 5
```

逆行列は inv で計算できる．

```
> A=[1 2;3 4]; inv(A)
ans =
  -2.00000   1.00000
   1.50000  -0.50000
```

行列の固有ベクトル，固有値は eig で求められる．

```
> [eigvec eigval]=eig(A)
eigvec =
  -0.82456  -0.41597
   0.56577  -0.90938
eigval =
  -0.37228   0.00000
   0.00000   5.37228
```

ベクトルの平均，標準偏差，分散は mean, std, var でそれぞれ計算できる．

```
> var([1 2 3 4])
ans = 1.6667
```

0 から 1 までの間の一様乱数は rand で，平均 0 分散 1 の標準正規分布に従う乱数は randn で生成できる．引数を指定すれば，その大きさの乱数行列を生成できる．

```
> randn(1,4)
ans =
   0.22797  -1.63055   0.38559  -1.66595
```

ソートは sort 関数で行なう．

```
>[sorted index]=sort([7 3 6 1 2])
sorted =
  1  2  3  6  7
index =
  4  5  2  3  1
```

4 描画関数

Octave には沢山のグラフ作成のための関数が組み込まれている．標準的な折れ線グラフは plot 関数で描画する．

```
> x=[-2:0.1:2]; y=sin(x); plot(x,y)
```

ヒストグラムは hist 関数で描画する．

```
> a=randn(1000,1); hist(a)
```

3 次元のグラフは surf, mesh で，等高線は contour で描画する．

```
> x=[-3:0.1:3]; y=stdnormal_pdf(x); surf(y'*y);
```

グラフの保存は print コマンドで行う．グラフを LaTeX などで取り込む場合は eps 形式で保存するとよい．

```
> print -deps graph.eps
```

5 Octave によるプログラミング

これまでは，コマンドラインに直接コマンドを打ち込むことにより計算を行ってきた．より複雑な処理をさせるときには，スクリプトや関数を用いる．

スクリプト: スクリプトは，これまでコマンドラインに打ち込んでいたものをあらかじめテキストファイルに記述したものである．例えば，適当なテキストエディタで次のようなスクリプトを作成し，myscript.m というファイル名で保存する．

```
A=[1 2; 3 4];
b=[5;6];
C=A*b;
```

但し，スクリプトのファイルは現在 octave が実行されているディレクトリに作成すること．現在のディレクトリは，Octave のコマンドライン上で

```
> pwd
```

と入力することにより確認できる．スクリプト myscript.m を作成し所定のディレクトリに置いたら，Octave のコマンドライン上で

```
> myscript
```

と入力することにより，myscript.m に記述されたコマンドが実行される．

繰り返しと条件分岐: C 言語などと同じように，繰り返しには for 文を使う．例えば，1 から 100 までの整数の和を求めるスクリプトは次のようになる．

```
n=100;
a=0;
for i=1:n
  a=a+i;
end
```

Octave でプログラミングするときに，非常に重要なことは**できる限り for 文を使わない**ということである．例えば，1 から 100 万までの整数の和を求めるとき，上記のプログラムで n=1000000 とすると非常に時間がかかるが，

```
a=sum(1:1000000);
```

とすれば，一瞬で計算が終わる．
　条件分岐には if 文を使う．例えば，

```
x=-10:0.1:10;
for xx=1:length(x)
  if x(xx)>0
    y(xx)=x(xx);
  else
    y(xx)=-x(xx);
  end
end
plot(x,y)
```

このプログラムも，次のようにすれば非常に高速になる．

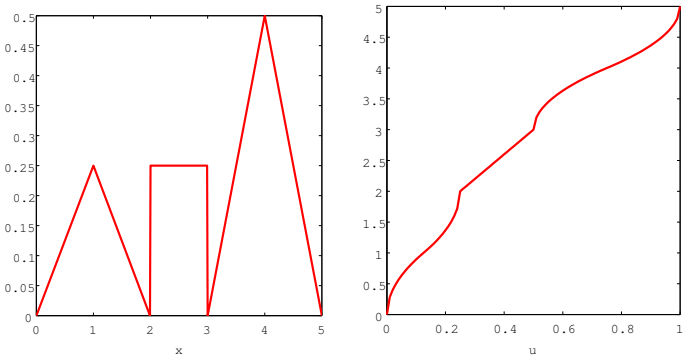
```
x=-10:0.1:10;
y=zeros(size(x));
y(x>0)=x(x>0);
y(x<=0)=-x(x<=0);
plot(x,y)
```

そのほか while 文や switch 文もある．

関数: スクリプトの先頭行で function と宣言することによって，関数を定義できる．ここでは，正規分布の確率密度関数の値を計算する関数を定義しよう．次のファイルを g_pdf.m という名前で保存する．

```
function z=g_pdf(x,mu,sigma)

z=1/sqrt(2*pi*sigma^2)*exp(-(x-mu).^2/(2*sigma^2));
```



(a) $f(x)$　　(c) $F^{-1}(u)$

図 1: 分布の例

そして，以下のスクリプトを実行すれば，正規分布の確率密度関数が表示・保存される．

```
clear all;
mu=0;
sigma=1;
x=[-3:0.1:3];
y=g_pdf(x,mu,sigma);

figure(1);
clf
plot(x,y);
print -deps gauss_pdf.eps
```

6　演習：棄却法および逆関数法

　本節では，次の確率密度関数に従う確率変数を棄却法および逆関数法で生成する（図 1(a) 参照）．

$$f(x) = \begin{cases} \frac{1}{4}x & 0 \le x < 1 \\ \frac{1}{2} - \frac{1}{4}x & 1 \le x < 2 \\ \frac{1}{4} & 2 \le x < 3 \\ -\frac{3}{2} + \frac{1}{2}x & 3 \le x < 4 \\ \frac{5}{2} - \frac{1}{2}x & 4 \le x \le 5 \end{cases} \tag{1}$$

棄却法: この確率密度関数の値は，次の関数 f.m で求められる．

```
function y=f(x)
  y=zeros(size(x));

  flag=(0<=x & x<1);
  y(flag)=0.25*x(flag);

  flag=(1<=x & x<2);
  y(flag)=-0.25*x(flag)+0.5;

  flag=(2<=x & x<3);
  y(flag)=0.25*ones(size(x(flag)));
```

```
  flag=(3<=x & x<4);
  y(flag)=0.5*x(flag)-1.5;

  flag=(4<=x & x<=5);
  y(flag)=-0.5*x(flag)+2.5;
```

このf.mを用いれば，棄却法によって次のようにして $f(x)$ に従う確率変数を生成することができる．

```
clear all;
n=10000;
u=5*rand(n,1);
v=0.6*rand(n,1);
flag=(v<=f(u));

figure(1);
clf
hist(u(flag==1),50);

figure(2);
clf
hold on
plot(u(flag==1),v(flag==1),'bo');
plot(u(flag~=1),v(flag~=1),'gx');
x=[0:0.1:5];
y=f(x);
plot(x,y,'r-');
```

逆関数法: 上記の確率密度関数 $f(x)$ の累積分布関数 $F(x)$ の逆関数 $F^{-1}(u)$ は次式で与えられる．

$$F^{-1}(u) = \begin{cases} \sqrt{8u} & 0 \le u < \frac{1}{8} \\ 2 - \sqrt{2 - 8u} & \frac{1}{8} \le u < \frac{1}{4} \\ 1 + 4u & \frac{1}{4} \le u < \frac{1}{2} \\ 3 + \sqrt{-2 + 4u} & \frac{1}{2} \le u < \frac{3}{4} \\ 5 - \sqrt{4 - 4u} & \frac{3}{4} \le u \le 1 \end{cases} \tag{2}$$

この逆関数の値は，次の関数 Finv.m で求められる．

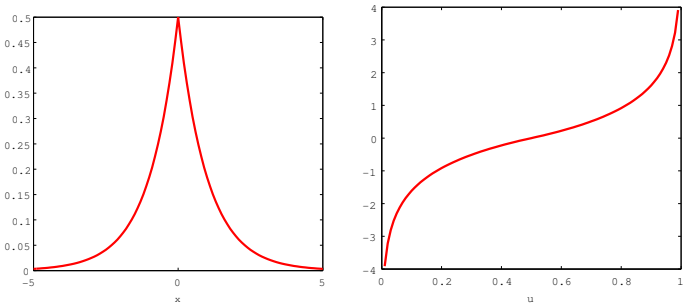
```
function x=Finv(u)
  x=zeros(size(u));

  flag=(0<=u & u<1/8);
  x(flag)=sqrt(8*u(flag));

  flag=(1/8<=u & u<1/4);
  x(flag)=2-sqrt(2-8*u(flag));

  flag=(1/4<=u & u<1/2);
  x(flag)=1+4*u(flag);

  flag=(1/2<=u & u<3/4);
  x(flag)=3+sqrt(4*u(flag)-2);
```



(a) $f(x)$ (c) $F^{-1}(u)$

図 2: ラプラス分布

```
  flag=(3/4<=u & u<=1);
  x(flag)=5-sqrt(4-4*u(flag));
```

このFinv.mを用いれば，逆関数法によって次のようにして $f(x)$ に従う確率変数を生成することができる．

```
clear all;
n=10000;
u=rand(1,n);
x=Finv(u);

figure(1);
clf
hist(x,50);
```

7 宿題

問1: 次の確率密度関数を持つ分布をラプラス分布 (Laplacian distribution) と呼ぶ．

$$f(x) = \frac{1}{2}\exp(-|x|) \tag{3}$$

(a) ラプラス分布の累積分布関数 $F(x)$，および，その逆関数 $F^{-1}(u)$ の式を求めよ．またその概形を octave などを使って図示せよ．

(b) 逆関数法および棄却法を用いて，ラプラス分布に従う確率変数を生成せよ．但し，棄却法では $-5 \le x \le 5$ の範囲だけを考えることにする．

問2: 自分で好きな確率分布を定義し，その分布に従う確率変数を逆関数法と棄却法で生成せよ．但し，確率密度関数は

$$f(x) \ge 0, \int f(x)dx = 1 \tag{4}$$

を満たさなければならないことに注意すること．

問3: 実際に逆関数法と棄却法で乱数を発生させてみた経験を元に，それぞれの手法の長所と短所を自分の言葉で述べよ．