TÉMAVÁZLAT 8-11. ALKALOM

Kémiai Számítástechnika Gyakorlat, Kémia BSc I. évf. 2018/2019 I. félév

(összeállította: Tóth Gergely)

TÉMAVÁZLAT

Octave (ingyenes MatLab klón, internetről telepíthető)

Könyvtár-naplózás beállítása

```
> pwd
> cd oda\odaam (Linuxnál) cd c:\a\oktatas\szamkem (Windows-nál)
> diary amitgepelek.dat
> ls
> quit
```

Kezdő lépések

```
> x=2
> y=3;
> z=x+y**3
> x=cos(z)
> disp(" muveleti sorrendek")
> function y=f(x,z) y=x+z*3-sqrt(x);endfunction;
> f(3,4)
> x=2
> z=6
> f(x,z)
> y=exp(x)*log(x)-log10(x)+sqrt(x)-ceil(x)+floor(x)-fmod(x,2)+tan(x)
> round(2.4)
> round(2.6)
> sign(3.2)
> sign(-3.2)
> factor(12312)
> lookfor factor
> factorial(12)
> help primes
> primes(100)
```

> Sor- és oszlopvektorok

```
> v=[2;34;2]
> v=[2,34,2]
> u=rand(1,3)
> size(u)
> v=resize(v,1,4)
> u=resize(u,1,4)
> v*u'
> dot(v,u)
> sum(v)
> prod(v)
> sumsq(v)
> sumsq(v.*u)        # . = elemenkent szoroz
> v
> sort(v) # sorbarendezés
```

Komplex számok

```
> z=0.1+2i
> real(z)
> imag(z)
> abs(z)
```

Koordináta transzformációk

```
> [theta,r]=cart2pol(1,0)
theta = 0
r =  1
> cart2pol(1,0)
ans = 0
> [x,y]=pol2cart(0,1)
x =  1
y = 0
> [x,y,z]=sph2cart(0,0,1)
x =  1
y = 0
z = 0
```

> Grafika (csak grafikus környezetben, startx után)

```
> plot(u,v,"4+") #4 egyik szín, + jelölő
```

```
> print -djpg proba.jpg #elmenti az utolsó képet jpg formátumban
```

Adatok tárolása – mátrix

n sor egy-egy mintát (objektumot) jelent

m oszlop: egy-egy tulajdonságot jelent

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix}$$ *n* sor, *m* oszlop

Adat centrálása: értékekből oszlopátlagok kivonása

$$c_{ij} = d_{ij} - \frac{\sum_{i=1}^{n} d_{ij}}{n} = d_{ij} - \overline{d}_j$$

Adatok skálázása: értékek osztása az adott oszlop szórásával

$$c_{ij} = \frac{d_{ij}}{s_j} = \frac{d_{ij}}{\sqrt{\frac{\sum_{i=1}^{n}\left(d_{ij} - \overline{d}_j\right)^2}{n-1}}}$$

Standardizálás (néha studentizálás, z-score, normál eloszlásnál szinte minden adat -3;+3 közé kerül)

$$c_{ij} = \frac{d_{ij} - \overline{d}_j}{s_j}$$

```
> D=rand(10,5)
> D(:,2)=cos(D(:,1))
> D(:,3)=D(:,1)*2+rand()
> D(:,4)=exp(D(:,1))
> D(:,5)=D(:,2)+D(:,4)+rand()*0.1
> mean(D)
> median(D)
> meansq(D)
> std(D)
> var(D)
> sortrows(D,2)
> statistics(D)
> help statistics
> center(D)
> A=studentize(D)
> mean(A) #standardizált átlaga 0
> std(A)  #standardizált szórása 1
```

Kovariancia – változók együttes mozgására és annak nagyságára utal

$$s_{xy}^2 = \frac{\sum_{i=1}^{n}(x_i - E_x)(y_i - E_y)}{n-1}$$

Korrelációs együttható, értéke [-1,1] intervallumba esik, lineáris kapcsolatra utal: $r_{xy} = \dfrac{\sum_{i=1}^{n}(x_i - E_x)(y_i - E_y)}{\sqrt{\sum_{i=1}^{n}(x_i - E_x)^2 \sum_{i=1}^{n}(y_i - E_y)^2}} = \dfrac{s_{xy}^2}{s_x s_y}$

```
> cov(D) #mátrixba rendezve…
> cor(D)
> cov(A) #standardizált kovarianciája=korrelációja az eredetinek
```

<u>Hipotézis vizsgálatok, varianciaanalízis</u>

F-próba

```
> p=var_test(D(:,1),D(:,2))
```

közvetlenül H_0 „valószínűségét” számolja:

ha $\alpha \leq$ p, H_0-t elfogadjuk

ha p < α, H_0-t elutasítjuk (H_a-t fogadjuk el).

Egymintás t-próba

```
> p=t_test(D(:,1),0.5)
> p=t_test(D(:,1),1)
> p=t_test(D(:,1),2,”<>”)
```

Megoldás kétoldalira:

ha $\alpha \leq$ p H_0-t elfogadjuk

ha p < α, H_0-t elutasítjuk (H_a-t fogadjuk el)

30 < N, lehet normál eloszlással dolgozni t-eloszlás helyett:

Megoldás egyoldalira:

```
> p=t_test(D(:,1),0.78,”<”)
> p=t_test(D(:,1),0.9,”>”)
```

Kétmintás t-próba

```
> p=t_test_2(D(:,1),D(:,2),”<>”)
```

Varianciaanalízis

```
> anova(D)
```

Mátrixműveletek

```
> A=[1,2;3,4]
> B=randn(2,2)
> A*B
> C=randn(2,3)
> C*A
error: operator *: nonconformant arguments (op1 is 2x3, op2 is 2x2)
> A*C
> C'*A
> det(A)
> inv(A)
> A*inv(A)
> eig(A)
> [vA,eA]=eig(A)
> A==B
> v=vec(A)
> u=v'
> B=eye(2)                  #Diagonal Matrix
> trace(A)
```

Nemlineáris egyenlet megoldása, maximum és minimum keresése

```
> function u=f(r) u=4*1*(3**12/r**12-3**6/r**6); endfunction; #
Lennard-Jones potenciál
> for i=1:10 v(i)=i*0.5+2.3;u(i)=f(v(i));endfor;
> plot(v,u)
> fsolve(@f,2.4)
> fsolve(@f,3.2)
> fsolve(@f,4)
> fzero(@f,[2,4])
> fminunc(@f,3)
> fminbnd(@f,2,4)
```

Inhomogén lineáris egyenletrendszer egyértelmű megoldása

```
> A=resize(A,3,3) #Ha volt korábban A
> A=[2,4,6;2,3,1;-1,0,5]
> b=[8;7;-2]
> inv(A)*b
```

```
> B=resize(B,3,3) ##Ha volt korábban B, Cramer-szabaly gyakorlasahoz
> B=A
> B(:,1)=b
> det(B)/det(A)
> B=A
> B(:,2)=b
> det(B)/det(A)
```

Túlhatározott lineáris egyenletrendszer megoldása

```
> A=resize(A,4,3)
> A(4,:)=[2,3,2]
> b=resize(b,4,1)
> b(4)=4
> inv(A'*A)*A'*b
```

> Konstans tag szerepeltetése az e.h. mátrixban

```
> A=resize(A,5,4)
> A(:,4)=1
> A(5,:)=[-1,5,-1,1]
> b=resize(b,5,1)
> b(5)=0
> inv(A'*A)*A'*b
```

Valószínűségi változó függvényének eloszlása

$y=y(x)$ és $x=x$(y) kölcsönösen egyértelmű függvények x és y valószínűségi változók között. Mi y eloszlása, ha x eloszlását ismerjük?

$f_y(Y)|dy| = f_x(x(Y))|dx|$, amiből átrendezéssel: $f_y(Y) = f_x(x(Y))\left|\frac{dx}{dy}\right|$

Tehát ismert kapcsolat esetén f_y sűrűségfüggvény megkapható minden olyan Y értékre, ahol a dx/dy derivált létezik és y folytonos.

Feladat: Határozza meg $f_y(y)$-t, ha $y=\sqrt{x}$ és x λ=1 paraméterű exponenciális eloszlással írható le ($x\geq 0$). Octave-val: Tabulálja 0.1-es x illetve y felosztással a sűrűségfüggvényeket és ábrázolja azokat.

```
for (i=1:60) v(i)=0.1*i; endfor;
for (i=1:60) fx(i)=exp(-v(i)); endfor;
for (i=1:60) fy(i)=2*v(i)*exp(-v(i)*v(i)); endfor;
```

```
plot(v,fx)
plot(v,fy)
intx=sum(fx)*0.1 #görbe alatti terület ellenőrzése
inty=sum(fy)*0.1 #görbe alatti terület ellenőrzése
```

Feladat: Határozza meg $f_y(y)$-t, ha $y=x^2$ és x standard normális eloszlású. Octave-val: Tabulálja 0.2-es x illetve y felosztással a sűrűségfüggvényeket és ábrázolja azokat. Teljesül-e, hogy a görbe alatti terület 1? Ha nem, miért és hogyan tudja rendbe tenni? Mi x és y értelmezési tartománya?

Gauss-féle hibaterjedési szabály

Legyen $y=f(x_1,x_2,\ldots,x_n)=f(\underline{x})$ egy többváltozós függvény. Az $x_1\ldots x_n$ változókat azonban nem pontosan ismerjük, csak becsüljük $\xi_1\ldots\ \xi_1$ valószínűségi változókkal. Ezen változók mindegyikéhez számolható becsült várható érték $E_1\ldots E_n$, becsült szórás ($s_1\ldots s_n$), annak négyzete a variancia (s_i^2), és számolható a kovarianciájuk (s_{ij}^2). Mekkora y varianciája?

Különböző közelítéseket alkalmazva (Taylor-sorfejtes, lineáris tag felettiek elhanyagolása...) levezethető, hogy

$$s_y^2=\sum_{i=1}^{n}\left(\frac{\partial y}{\partial x_i}\right)^2 s_i^2+2\sum_{i<j}\sum_{j=1}^{n}\left(\frac{\partial y}{\partial x_i}\right)\left(\frac{\partial y}{\partial x_j}\right)s_{ij}^2=\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\partial y}{\partial x_i}\right)\left(\frac{\partial y}{\partial x_j}\right)s_{ij}^2$$

, ahol a parciális deriváltak számolásakor $x_1\ldots x_n$ változók $E_1\ldots E_n$ értékkel szerepelnek.

Feladat: vezessük le $y=a_0+\sum_{i=1}^{n}a_1x_1$ *függvény esetére* s_y^2 *képletét! Lineáris kombináció esetén* $E_y=a_0+\sum_{i=1}^{n}a_iE_i$ *módon számolható.*

Feladat: Hogyan egyszerűsödik a képlet, ha $\xi_1\ldots\ \xi_1$ páronként független valószínűségi változók?

Feladat: Generáljon egy 10x3-as X mátrixot, ahol az oszlopok közül legalább az egyik valamilyen hibával terhelt kapcsolatban áll a többivel (n=3 oszlop (=változó), m= 10 sor (=mérés)). Becsülje meg a fenti képlet segítségével y varianciáját. Számolja ki az y_i értékeket is az X mátrix egy-egy sora alapján is, és az ebből számolható y varianciát hasonlítsa össze a becsülttel! A következő függvényeket használja:

a) $y=x_1+2*x_2-3*x_3$

```
D=randn(10,3)
```

```
D(:,3)=D(:,1)+2*D(:,2)+0.1*randn()
CD=cov(D)
y=D(:,1)+2*D(:,2)-3*D(:,3)
PD=[1,2,-3;2,4,-6;-3,-6,9]
var(y)
sum(sum(CD.*PD))
```

b) $y=x_1*x_1+\cos(x_2)*x_3$

c) $y=x_1*x_1+\cos(x_2)+x_3$

A lineáris esetre ellenőrizze, hogy a várható érték pontosan számolható-e a két feladattal korábban szereplő képlettel!

Haladó feladat: Az előző feladat valamelyik esetére adjunk meg konfidencia intervallumot E_y-ra. Ennek megadásához azonban ismerni kell, hogy mekkora szabadsági fokú t-eloszlással dolgozzunk. Ennek értéke a következő képlettel becsülhető:

$$\nu_y = \frac{s_y^4}{\sum_{i=1}^{n}\left(\frac{\partial y}{\partial x_i}\right)^4 s_i^4 \frac{1}{m-1}}$$

, ahol m a mérések (X sorainak) száma.

Vessük össze ezt a konfidencia intervallumot azzal, mintha közvetlenül y_i értékekből dolgoznánk.

<u>Integrálás, Polinomok, Interpoláció</u>

> Integrálás

```
> u=rand(1,20);
> for i=1:20 v(i)=i*0.1; endfor;
> u
> v
> trapz(v,u)
ans =  0.87450
> quad('cos',0,pi()/2)
ans =  1.0000
```

> Polinomok

```
> c=[2,3,4,5,1];        #koefficiensek (utolsó a konstans tag)
> a=[1,2,3,4];
> conv(a,c)             #szorzat
> polyder(a)            #deriválása
> q=polyder(a,c)        #szorzat deriváltja
```

```
> [q,r]=polyder(a,c)    #racionális törtfüggvény
> roots(a)              #gyökök
> for i=1:20 u(i)=v(i)+cos(v(i)*0.3); endfor;
> p=polyfit(v,u,5)      #polinom illesztése
> pint=polyint(p)       #polinom integrálása,határozatlan integrálás
> polyval(pint,pi()/2)-polyval(pint,0)   #helyettesítési érték
kiszámolása, Newton-Leibniz formula alapján
```

> Interpoláció

```
> interp1(v,u,1.05)
> interp1(v,u,1.05,'linear')
> interp1(v,u,1.05,'cubic')
```

Feladat

Polinom illesztése mért adatokra, integrálás

Egy mérés során az idő függvényében rögzítik az adatokat, majd ennek idő szerinti integrálját próbálják meghatározni. Technikai okból a folyamat középső szakasza nem mérhető, de a mért mennyiség jól közelíthető 3-ad fokú polinommal.

Illesszen harmadfokú polinomot a fenti adatokra, majd számolja ki a polinom integrálját a *t*€[0;10] intervallumon. Hasonlítsa össze a kapott értéket a mért görbe közvetlenül trapézformulával meghatározott integráljával. Ez utóbbinál lineárisan extrapoláljon a *t*=0 esetre.

Mért adatok:

t	*y(t)*
0,5	-4,5
1	-6,8
1,5	-8,9
2	-12,5
2,5	-15,9
7,5	17,0
8	35,6
8,5	57,6
9	84,9
9,5	117,7
10	155,9

Fourier-transzformáció

Jean-Baptiste Joseph Fourier (1768—1830) francia matematikus és fizikus

Ugyanaz az információ, de más változó terében…

Fourier-transzformáció matematikai háttere:

Fourier-sorfejtés (Bevezető matematika 2-ből ismert, ott T=2π)

$$f(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos\frac{2\pi nt}{T} + \sum_{n=1}^{\infty} b_n \sin\frac{2\pi nt}{T}$$

$$\frac{a_n}{2} = \frac{1}{T}\int_{t_0}^{t_0+T} f(t)\cos\left(\frac{2\pi nt}{T}\right)dt \qquad \frac{b_n}{2} = \frac{1}{T}\int_{t_0}^{t_0+T} f(t)\sin\left(\frac{2\pi nt}{T}\right)dt$$

$$a_0 = \frac{1}{T}\int_{t_0}^{t_0+T} f(t)dt$$

Periodikus függvény Fourier-sorfejtése exponenciális formában:

$$e^{it} = \cos(t) + i\sin(t) \qquad f(t) = \sum_{n=-\infty}^{\infty} c_n e^{i\omega t}$$

$$\omega = \frac{2\pi n}{T} \qquad \nu = \frac{n}{T} \qquad c_n = \bar{c}_{-n} = \frac{a_n - ib_n}{2},\ n\neq 0$$

Fourier-transzformáció – folytonos, nem periodikus függvényre

Fourier-transzformáció:

$$F(\nu) = \int_{-\infty}^{\infty} f(t)e^{-i2\pi\nu t}dt$$

inverz Fourier-transzformáció:

$$f(t) = \int_{-\infty}^{\infty} F(\nu)e^{i2\pi\nu t}d\nu$$

(több kicsit eltérő képlet van, a lényeg, hogy megfelelő oda-vissza párt használjunk)

Diszkrét Fourier-transzformáció

N mintavétel Δt gyakorisággal

jelölések: $\Delta\nu = \frac{1}{N\Delta t}$ $\qquad 2\pi\nu t = \frac{2\pi nk}{N}$

$$F(\nu) = F(k\Delta\nu) = F_k \qquad f(t) = f(n\Delta t) = f_n$$

Fourier-transzformáció:

$$F_k \equiv \frac{1}{N}\sum_{n=0}^{N-1} f_n e^{-\frac{i2\pi nk}{N}}$$

inverz Fourier-transzformáció:

$$f_n \equiv \sum_{k=0}^{N-1} F_k e^{\frac{i2\pi nk}{N}}$$

Gyors Fourier-transzformáció: speciális számítási módszer, ha $N=2^k$ adatunk van. Ha kevesebb, érdemes ennyire kiegészíteni 0-kal, ez nem befolyásolja az eredményt.

Feladatok

1 sin transzformációja :

```
> n=256
> dt=1/n
> for i=1:n t(i)=(i-1)*dt; endfor;
> f=sin(t*15*2*pi);
> plot(t,f) # valami hiba miatt a plot parancsot kétszer kell kiadni
> dnu=1/n/dt
> for i=1:n nu(i)=(i-1)*dnu; endfor;
> F=fft(f);
> Fa=abs(F);
> Fr=real(F);
> plot(nu,Fa)
> plot(nu,Fr)
> Fa(15)
> Fa(16)
> Fa(17)
> finv=ifft(F);
> plot(t,finv)
```

3 sin keveréke, egyik frekvencia kivágása (frekvenciaszűrés):

```
> f=sin(t*15*2*pi)+sin(t*5*2*pi)+sin(t*30*2*pi);
> plot(t,f)
> F=fft(f);
> Fa=abs(F);
> plot(nu,Fa)
> Fa(16)
> Fa(242)
> F(16)=0+0i;
> F(242)=0+0i;
> finv=ifft(F);
> plot(t,finv)
```

```
> plot(t,f)
```

3 sin keveréke + zaj, zaj szűrése (amplitúdószűrés):

```
> f=sin(t*15*2*pi)+sin(t*5*2*pi)+sin(t*30*2*pi);
> fz=f;
> fz=f+2*randn(1,256);
> plot(t,fz)
> F=fft(fz);
> Fa=abs(F);
> plot(nu,Fa)
> for i=1:n if(Fa(i)<80) F(i)=0+0i; endif; endfor;
> finv=ifft(F);
> plot(t,finv)
> plot(t,f)
```

Gauss fv. transzformálása

```
> dx=0.1;
> for i=1:n x(i)=(i-1)*dx; endfor;
> s=1 #utána innen többször s=0.5-tel es s=3-mal is!
> for i=1:n g(i)=1/sqrt(2*pi()*s*s)*exp(-x(i)*x(i)/2/s/s); endfor;
> plot(x,g)
> dnu=1/dx/n;
> for i=1:n nu(i)=(i-1)*dnu; endfor;
> G=fft(g);
> Ga=abs(G);
> plot(nu,G);
```