Interpoleerimine ja regressioon

Enne käesoleva peatüki põhitemaatika juurde minekut vaatleme ühte käsu plot lisavõimalust. Tuletame meelde, et käsk plot(x,y) joonestab tasandile koordinaatidega x ja y antud punktid ja ühendab need järjekorras sirglõikudega. Kui me ei soovi sirglõikude joonestamist punktide vahele, siis võib seda käsku kasutada järgmisel kujul:

plot(x,y,'*sümbol*')

kus *sümbol* on punktide tähis (see võib olla *, o, x või ˆ). Näiteks olgu antud vektorid $x = (1, 2, 3, 4, 5, 6)$ ja $y = (4, 5, 8, 6, 7, 5)$. Joonestame neile vastavad punktid $P(x, y)$ tasandile kasutades sümbolit $*$. Selleks sisestame käsud

```
x=[1,2,3,4,5,6];
y=[4,5,8,6,7,5];
plot(x,y,'*')
```

ja käivitame. Tulemuseks saame järgmise joonise.

Nii nagu ikka, saab ühe käsuga plot joonestada korraga mitu graafikut. Näiteks käsk

```
plot(t,u,'*',z,v)
```

joonestab tasandile:
1) punktid koordinaatidega t ja u ning tähistab need sümboliga $*$;
2) punktid koordinaatidega z ja v ning ühendab need sirglõikudega.

Vaatleme interpoleerimist splainidega. Splainiga interpoleerimiseks võib Matlab-Octaves kasutada käsku interp1, mille kuju on järgmine:

interp1(x,y,xi,'*meetod*')

kus x on sõlmede vektor, y on vastavate funktsiooni väärtuste vektor, xi on argumendi väärtuste vektor, mille korral soovitakse splaini välja arvutada ning *meetod* on splaini tüübi nimetus. Näiteks kui soovitakse lineaarsplaini klassist $S^{1,0}(x)$, siis tuleb *meetod*-i kohale kirjutada linear, kui aga soovitakse kuupsplaini klassist $S^{3,2}(x)$, siis tuleb *meetod*-i kohale kirjutada spline . Käsk interp1 annab tulemuseks argumendi väärtuste vektorile xi vastava splaini väärtuste vektori.

Näiteülesanne 1. Antud on järgmine funktsiooni $y = f(x)$ väärtuste tabel:

x	1	2	3	4	5	6	7
y	0	7	1	3	2	4	1

Interpoleerida seda funktsiooni kuupsplainiga $S^{3,2}(x)$. Joonestada interpolatsioonipunktid ja splain samas teljestikus lõigul $[1, 7]$.

Lahendus. Kirjutame skripti järgmised read:

```
x=[1,2,3,4,5,6,7];
y=[0,7,1,3,2,4,1];
xi=1:1e-3:7;
yi=interp1(x,y,xi,'spline');
plot(x,y,'*',xi,yi)
xlabel('x')
ylabel('y')
```

ja käivitame skripti. Tulemuseks saame järgmise joonise.

Näiteülesanne 2. Antud on järgmine funktsiooni $v = f(u)$ väärtuste tabel:

u	0	0.1	0.2	0.3	0.4	0.5	0.6
v	0	0.6	2.1	3.1	2.5	2	2.6

Interpoleerida seda funktsiooni lineaarsplainiga $S^{1,0}(u)$ ja kuupsplainiga $S^{3,2}(u)$.

Joonestada interpolatsioonipunktid ja splainid samas teljestikus lõigul $[0, 0.6]$. Arvutada $w_0 = S^{1,0}(0.35)$ ja $w_1 = S^{3,2}(0.35)$.

Lahendus. Kirjutame skripti järgmised read:

```
u=[0,0.1,0.2,0.3,0.4,0.5,0.6];
v=[0,0.6,2.1,3.1,2.5,2,2.6];
ui=0:1e-4:0.6;
vi1=interp1(u,v,ui,'linear');
vi2=interp1(u,v,ui,'spline');
plot(u,v,'*',ui,vi1,ui,vi2)
xlabel('u')
ylabel('v')
w0=interp1(u,v,0.35,'linear')
w1=interp1(u,v,0.35,'spline')
```

ja käivitame skripti. Tulemuseks saame järgmise joonise, millel roheline joon vastab lineaarspalinile ja punane joon kuupsplainile. Arvulised vastused on $w_0 = 2.8$ ja $w_1 = 2.9217$.

Märkus. Graafiku legendis ja telgede märgendites esinevate keerukamate sümbolite loomiseks saab kasutada nn Latexi koodi. Näiteks sümboli $^\circ$ saab käsuga `^\circ`. Latexi koodi saab kasutada ka kreeka tähetede kirjutamisel. Näiteks α saab käsuga `\alpha`. Täpset infot erinevatele sümbolitele vastavate käskude kohta on võimalik leida internetist otsides märksõnaga "latex symbols".

Järgnevalt vaatleme vähimruutude meetodit ehk regressiooni tabelina antud funktsioonide lähendamisel. Alustame lineaarse ja mittelineaarse regressiooniga (st polünomiaalse lähendamisega) juhul, kui ülesandes esinevad kaalud on võrdsed (sel juhul võib nad võtta võrdseks ühega). Taolise ülesande lahendamiseks on Matlab-Octaves olemas spetsiaalne käsk `polyfit`, mille kuju on järgmine:

```
polyfit(x,y,k)
```

kus `x` ja `y` on tabelis antud argumendi ja funktsiooni väärtused ning `k` on polünoomi aste. Kui `k = 1`, siis on tegemist lineaarse regressiooniga, kui `k = 2`, on tegemist ruutregressiooniga jne. Käsk väljastab regressioonipolünoomi kordajate vektori.

Näiteülesanne 3. Antud on järgmine funktsiooni $y = f(t)$ väärtuste tabel:

t	1	2	5	8	11	13
y	3	6	13	24	31	39

Leida lineaarne regressioon võrdsete kaaludega.

Lahendus. Koostame järgmise skripti:

```
t=[1,2,5,8,11,13];
y=[3,6,13,24,31,39];
c=polyfit(t,y,1)
```

ja käivitame. Matlab-Octave annab vastuse

```
c =
    2.9545   -0.3636
```

See tähendab, et otsitava lineaarfunktsiooni valem on $\Phi(t) = 2.9545t - 0.3636$.

Ettantud kordajatega polünoomi väärtuste arvutamiseks sobib kõige paremini käsk `polyval` , mille kuju on järgmine:

```
polyval(c,xi)
```

kus `c` on polünoomi kordajate vektor ja `xi` on argumendi väärtuste vektor. Käsk annab `xi` -le vastava polünoomi väärtuste vektori.

Näiteülesanne 4. Arvutada näiteülesandes 3 saadud lineaarse lähendi väärtused $t = 3$ ja $t = 10$ korral.

Lahendus. Oletame, et peale näiteülesande 3 lahendamist ei ole Matlab-Octavest väljutud. Siis on vektor c veel mälus olemas (vastasel juhul tuleb näiteülesandes 3 olevad käsud uuesti täita). Koostame järgmise skripti:

```
Vastus1=polyval(c,3)
Vastus2=polyval(c,10)
```

ja käivitame. Tulemus on järgmine:

```
Vastus1 =
      8.5000
Vastus2 =
      29.1818
```

Näiteülesanne 5. Joonestada näiteülesandes 3 saadud lineaarne lähend ja tabeli punktid samas teljestikus lõigul $[1, 13]$. Lisada telgede märgendid.

Lahendus. Eeldame jällegi, et Matlab-Octavest ei ole vahepeal väljutud, st vektorid t, y ja c on mälus olemas. Kirjutame skripti järgmised read:

```
ti=1:1e-2:13;
yi=polyval(c,ti);
plot(ti,yi,t,y,'*')
xlabel('t')
ylabel('y')
```

ja käivitame skripti. Kuvatakse joonis.

Näiteülesanne 6. Antud on järgmine funktsiooni $u = f(x)$ väärtuste tabel:

x	1	2	3	4	5	8	9	11	12
u	10	13	19	28	30	24	16	11	6

Leida kuupregressioon (so kuuplähend vähimruutude mõttes) võrdsete kaaludega. Joonestada tabeli punktid ja lähend samas teljestikus lõigul $[1, 12]$. Lisada telgede märgendid.

Lahendus. Kirjutame skripti järgmised read:

```
x=[1 2 3 4 5 8 9 11 12];
u=[10 13 19 28 30 24 16 11 6];
c=polyfit(x,u,3)
xi=1:1e-3:12;
ui=polyval(c,xi);
plot(x,u,'*',xi,ui)
xlabel('x')
ylabel('u')
```

ja käivitame skripti. Antakse järgmine vastus:

```
c =
     0.050102  -1.639367   13.299637   -4.320540
```

ja kuvatakse joonis. Järelikult on otsitava kuuplähendi valem järgmine:

$$\Phi(x) = 0.050102x^3 - 1.639367x^2 + 13.299637x - 4.320540.$$

Mittevõrdsete kaalude korral puudub Matlab-Octaves käsk, millega saab regressioonipolünoomi otseselt leida. Sel juhul tuleb regressiooniülesandele vastav lineaarne süsteem ise koostada ja lahendada. Selle süsteemi maatriksis esinevad summad mitmesugustest vektoritest. Vektori x elementide summa saab Matlab-Octaves lihtsalt käsuga sum(x) . Vaatlemegi kõigepealt taoliste summade arvutamist.

Näiteülesanne 7. Sisestada vektorid $t = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10)$, $x = (4, 5, 4, 6, 4, 5, 4, 6, 4, 5)$ ja leida vektor s, mille komponendid on

$$s_1 = \sum_{i=1}^{10} t_i, \quad s_2 = \sum_{i=1}^{10} x_i, \quad s_3 = \sum_{i=1}^{10} t_i x_i, \quad s_4 = \sum_{i=1}^{10} t_i x_i^2.$$

Lahendus. Skripti koostamisel tuleb arvestada sellega, et summa all toimub korrutamine ja astendamine komponentide kaupa, seega tuleb kasutada käske .* ja .^ Skript on järgmine:

```
%Andmete sisestamine
t=[1 2 3 4 5 6 7 8 9 10];
x=[4 5 4 6 4 5 4 6 4 5];
%Summade all olevate vektorite leidmine
s3vektor=t.*x;
s4vektor=t.* x.^2;
%Vektori s komponentide leidmine
s(1)=sum(t);
s(2)=sum(x);
s(3)=sum(s3vektor);
s(4)=sum(s4vektor);
%Vektori s kuvamine
s
```

Samaväärne, kuid lühem skript oleks

```
%Andmete sisestamine
t=[1 2 3 4 5 6 7 8 9 10];
x=[4 5 4 6 4 5 4 6 4 5];
%Vektori s komponentide leidmine
s(1)=sum(t);
s(2)=sum(x);
s(3)=sum(t.*x);
s(4)=sum(t.* x.^2);
%Vektori s kuvamine
s
```

Peale skripti käivitamist annab Matlab-Octave vektori s komponendid. Vastus on $s = (55, 47, 262, 1282)$.

Näiteülesanne 8. Argumendi x, funktsiooni y ja kaalude κ väärtused on antud tabelis

x	-4	-3	-2	-1	0	1	2	3	4	5
y	2.3	3.0	5.2	6.8	7.0	9.6	11	13.6	14	16.6
κ	1	1	2	2	3	3	2	2	1	1

Leida lineaarne regressioon etteantud kaaludega. Joonestada lähend ja tabeli punktid samas teljestikus.

Lahendus. Lähendi $\Phi(x) = c_1 x + c_2$ kordajate leidmiseks tuleb lahendada süsteem $Ac = b$, mille maatriks ja parema poole vektor avalduvad kujul

$$A = \begin{pmatrix} \sum\limits_{i=1}^{10} \kappa_i x_i^2 & \sum\limits_{i=1}^{10} \kappa_i x_i \\ \sum\limits_{i=1}^{10} \kappa_i x_i & \sum\limits_{i=1}^{10} \kappa_i \end{pmatrix}, \qquad b = \begin{pmatrix} \sum\limits_{i=1}^{10} \kappa_i x_i y_i \\ \sum\limits_{i=1}^{10} \kappa_i y_i \end{pmatrix}.$$

Koostame skripti

```
%Andmete sisestamine
```

```
x=[-4 -3 -2 -1 0 1 2 3 4 5];
y=[2.3 3.0 5.2 6.8 7.0 9.6 11 13.6 14 16.6];
kappa=[1 1 2 2 3 3 2 2 1 1];
%Maatriksi A ja vektori b komponentide leidmine
A(1,1)=sum(kappa.*x.^2);
A(1,2)=sum(kappa.*x);
A(2,1)=A(1,2);
A(2,2)=sum(kappa);
b(1)=sum(kappa.*x.*y);
b(2)=sum(kappa.*y);
%Süsteemi Ac=b lahendamine (NB! võib kasutada ka käsku c=A\b')
c=inv(A)*b'
%Graafiku joonestamine
xi=-4:1e-3:5;
yi=polyval(c,xi);
plot(xi,yi,x,y,'*')
```

Peale käivitamist saame järgmise joonise. Kordajate väärtused on $c_1 = 1.6055$, $c_2 = 8.0250$.

NÄITEÜLESANNE 9. Argumendi t, funktsiooni z ja kaalude κ väärtused on antud tabelis

t	1	3	5	7	9
z	10	15	19	21	22
κ	1	2	3	2	1

Leida ruutregressioon etteantud kaaludega. Joonestada lähend ja tabeli punktid samas teljestikus.

Lahendus. Lähendi $\Phi(x) = c_1x^2+c_2x+c_3$ kordajate leidmiseks tuleb lahendada süsteem $Ac = b$, mille maatriks ja parema poole vektor avalduvad kujul

$$A = \begin{pmatrix} \sum\limits_{i=1}^{5} \kappa_i t_i^4 & \sum\limits_{i=1}^{5} \kappa_i t_i^3 & \sum\limits_{i=1}^{5} \kappa_i t_i^2 \\ \sum\limits_{i=1}^{5} \kappa_i t_i^3 & \sum\limits_{i=1}^{5} \kappa_i t_i^2 & \sum\limits_{i=1}^{5} \kappa_i t_i \\ \sum\limits_{i=1}^{5} \kappa_i t_i^2 & \sum\limits_{i=1}^{5} \kappa_i t_i & \sum\limits_{i=1}^{5} \kappa_i \end{pmatrix}, \qquad b = \begin{pmatrix} \sum\limits_{i=1}^{5} \kappa_i t_i^2 z_i \\ \sum\limits_{i=1}^{5} \kappa_i t_i z_i \\ \sum\limits_{i=1}^{5} \kappa_i z_i \end{pmatrix}.$$

Koostame skripti

```
%Andmete sisestamine
t=[1 3 5 7 9];
z=[10 15 19 21 22];
kappa=[1 2 3 2 1];
%Maatriksi A ja vektori b komponentide leidmine
A(1,1)=sum(kappa.*t.^4);
A(1,2)=sum(kappa.*t.^3);
A(1,3)=sum(kappa.*t.^2);
A(2,1)=A(1,2);
A(2,2)=A(1,3);
A(2,3)=sum(kappa.*t);
A(3,1)=A(1,3);
A(3,2)=A(2,3);
A(3,3)=sum(kappa);
b(1)=sum(kappa.*t.^2.*z);
b(2)=sum(kappa.*t.*z);
b(3)=sum(kappa.*z);
%Süsteemi Ac=b lahendamine (NB! võib kasutada ka käsku c=A\b')
```

```
c=inv(A)*b'
%Graafiku joonestamine
ti=1:1e-3:9;
zi=polyval(c,ti);
plot(ti,zi,t,z,'o')
```

Peale käivitamist saame järgmise joonise. Kordajate väärtused on $c_1 = -0.1833$, $c_2 = 3.3333$, $c_3 = 6.7833$.

Lõpuks vaatleme põgusalt ka eksponentsiaalset regressiooni.

Näiteülesanne 10. Leida järgmise tabeliga antud funktsiooni eksponentsiaalne regressioon võrdsete kaaludega.

x	0	5	10	15	20
y	100	50	1	0.2	0.015

Joonestada lähend ja tabeli punktid samas teljestikus. Lähendit kasutades arvutada y väärtus, kui $x = 13$.

Lahendus. Teatavasti saab eksponetsiaalselt regressioonilt logaritmimist kasutades üle minna lineaarsele regresioonile. Selleks tuleb leida funktsiooni $z = \ln y$ lineaarne lähend vähimruutude mõttes. Olgu selleks $z = c_1x + c_2$. Siis avaldub otsitav eksponentsiaalne regressioon kujul $y = e^z = e^{c_1x+c_2}$ ehk $y = ae^{bx}$, kus $a = e^{c_2}$ ja $b = c_1$. Arvutuste teostamiseks koostame skripti

```
%Algandmete sisestamine ja logaritmimine
x=[0,5,10,15,20];
y=[100,50,1,0.2,0.015];
z=log(y);
%Regressiooni leidmine
c=polyfit(x,z,1);
%Graafiku joonestamine
xi=0:1e-2:20;
yi=exp(c(1)*xi+c(2));
plot(xi,yi,x,y,'*')
%y väärtuse arvutamine x=13 korral
Vastus=exp(c(1)*13+c(2))
```

Käivitamisel saame joonise. Arvuline vastus on $y = 0.4290$.

Harjutusülesanne 1. Töötava mootori mähise temperatuuri on mõõdetud 10-minutilise intervalliga ajalõigul 0 kuni 60min. Tulemused on toodud järgmises tabelis:

t(min)	0	10	20	30	40	50	60
T(°C)	30	51	68	79	85.5	89	90

Interpoleerida temperatuurifunktsiooni splainiga $S^{3,2}(t)$. Joonestada interpolatsioonipunktid ja splain samas teljestikus lõigul $[0, 60]$. Kanda joonisele telgede märgendid ja legend. Arvutada temperatuuri ligikaudne väärtus ajahetkel 54min splaini kasutades. Skript salvestada nime z40.m all.

Lahendus.

Harjutusülesanne 2. Tabelis on antud soojusvaheti kasutegur sõltuvalt seda läbivast õhu kogusest ajaühikus.

$Q(\frac{\mathrm{m}^3}{\mathrm{h}})$	100	200	300	400	500
$\eta(\%)$	70	71	70	66	61

Leida kasuteguri ruutregressioon võrdsete kaaludega. Joonestada ruutlähend ja tabeli punktid samas teljestikus lõigul $[100, 500]$. Lisada telgede märgendid ja sobiv legend. Skript salvestada nime z41.m all.

Lahendus.

Harjutusülesanne 3. Asünkroonmootori pöördemomenti on mõõdetud erinevate sageduste korral. Tulemused on toodud järgmises tabelis:

$n\left(\frac{\mathrm{p}}{\mathrm{min}}\right)$	0	500	1000	1500	2000	2500	3000	3500	4000
M (Nm)	55	60	66	73	81	93	94	80	60

Leida selle mehaanilise karakteristiku 4-nda astme lähend vähimruutude mõttes võrdsete kaaludega. Joonestada tabeli punktid ja lähend samas teljestikus lõigul $[0, 4000]$. Lisada telgede märgendid. Arvutada lähendit kasutades pöördemomendi ligikaudne väärtus sageduse $\mathrm{n} = 2400\frac{\mathrm{p}}{\mathrm{min}}$ korral. Skript salvestada nime z42.m all.

Lahendus.

Harjutusülesanne 4. Tööstusmasina ebatäpsus sõltub selle masina töö kestvusest. Järgmises tabelis on antud masina täpsuse ε väärtused sõltuvalt ajast t:

t(h)	30	33	34	35	39	44	45
ε(mm)	1.1	1.21	1.23	1.25	1.3	1.4	1.42

Leida täpsuse lineaarne regressioon järgmiste kaaludega:

$\varkappa$	1	1	1	1.5	1.5	2	2

Joonestada lähend ja tabeli punktid samas teljestikus. Lisada telgede märgendid ja sobiv legend. Arvutada täpsuse väärtus, kui t =50h. Skript salvestada faili z43.m.

Lahendus.

Harjutusülesanne 5. Tabelis on antud tulu, mis on saadud teatud arvutite müügist.

t(aasta)	2008	2009	2010	2011
$R(10^9\$)$	3	4.2	6	9.2

Leida eksponentsiaalne regressioon võrdsete kaaludega. Joonestada lähend ja tabeli punktid samas teljestikus. Lisada telgede märgendid. Prognoosida tulu 2012.a. Skript salvestada nime z44.m all.

Lahendus.