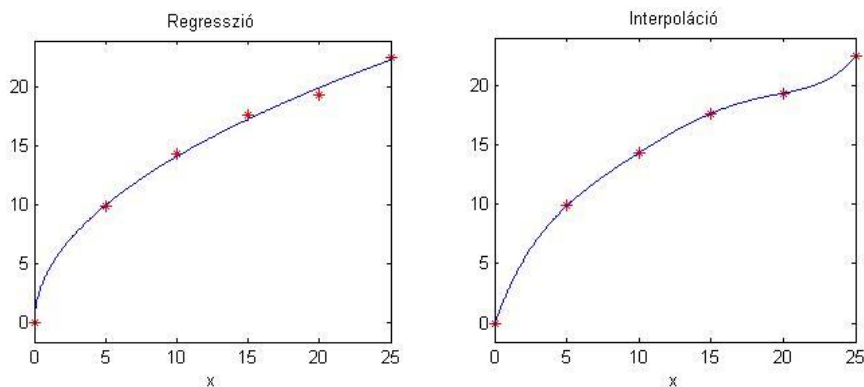


8. REGRESSZIÓ

A mérnöki gyakorlatban sokféle fizikai mennyiséget mérnek műszerekkel, amiket manapság többnyire digitalizálnak és a mérési eredményeket diszkrét pontonként tárolják. Ilyen például szilárdságtanból a szakítóvizsgálat, ahol az anyagok húzó igénybevétellel szembeni ellenállását mérik vagy geodéziából a teljes hullámalakos lézerszkennerek mérései. A méréseket különböző módokon használják fel, van amikor tudják, hogy az adott fizikai mennyiség milyen alakú összefüggéssel írható le és függvényillesztéssel (regresszió) keresik ennek a függvénynek a paramétereit, máskor a mért pontok között lenne szükség a becsült értékekre (interpoláció) vagy meg kellene becsülni, hogy a mért tartományon kívül hogyan alakulnának az adatok (extrapoláció).

Regresszió (függvényillesztés) esetén meghatározzuk a pontokra legjobban illeszkedő függvény paramétereit, ilyenkor a meghatározott függvény többnyire nem megy át a mért pontokon, csak közel halad hozzájuk. Interpoláció esetén az ismert pontok közötti értékeket szeretnénk megbecsülni, úgy, hogy a görbe (általában valamilyen polinom) minden mért ponton áthaladjon.



REGRESSZIÓ MINŐSÍTÉSE

Regresszió esetében a 'legjobban' illeszkedő függvényt keressük. Hogyan tudjuk mérni, mi a legjobban illeszkedő? Ehhez meg kell határozni az eltéréseket a mért pontok y_i koordinátái és az illesztett függvény adott pontbeli $f(x_i)$ függvényértéke között. Ezeket maradék eltéréseknek is szokták nevezni (r_i):

$$r_i = y_i - f(x_i)$$

A maradék eltéréseket kellene valamilyen módon minimalizálni, egyetlen mérőszámot alkalmazva az összes pontra. Lehetne egyszerűen összeadni a hibákat, de ebben az esetben előfordulhatna, hogy nagyon nagy pozitív és nagyon nagy negatív hibák vannak, amelyek kiejtik egymást és hiába lesz nulla az összes eltérés, az illeszkedés rossz lesz. Többféle megoldás lehetne erre, pl. minimalizálhatnánk a hibák abszolút értékeinek az összegét, akkor nem ejtenék ki egymást a hibák, vagy lehetne az abszolút értékben maximális eltérést minimalizálni, ezeket azonban nem olyan könnyű megvalósítani. A legáltalánosabban alkalmazott megoldás a hibák négyzetösszegének minimalizálása, a legkisebb négyzetek módszere, ami Gauss nevéhez köthető. Ezzel jól lehet mérni az illeszkedés minőségét és egyértelmű

megoldást ad a függvény paramétereire. Minimalizálandó az összes hiba négyzetösszege (S):

$$S = \sum_{i=1}^n r_i^2 = \sum_{i=1}^n (y_i - f(x_i))^2$$

A regresszió lokális minősítése a maradék eltérések (r_i - rezídiuumok) alapján történhet, a globális minősítés pedig ezeknek a korrigált tapasztalati szórásával:

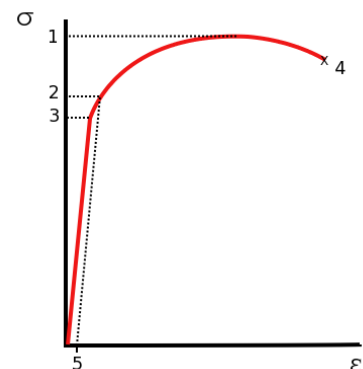
$$\sigma = \sqrt{\frac{\sum_{i=1}^n (y_i - f(x_i))^2}{n - np}} = \sqrt{\frac{\sum_{i=1}^n r_i^2}{f}} = \sqrt{\frac{S}{f}}$$

ahol n a rendelkezésre álló mérések, np a becsült paraméterek, f pedig a fölös mérések száma ($f=n-np$).

EGYENES ILLESZTÉS

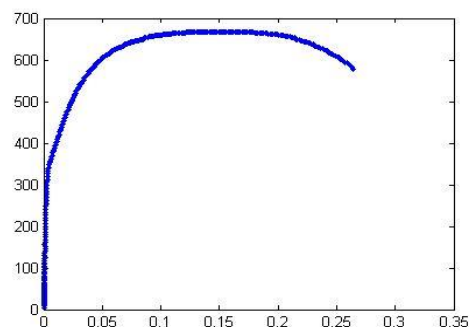
Nézzünk egy konkrét példát szilárdságtanból! Egy jól megmunkálható, ötvözött acél szakítóvizsgálatából származó méréseket kell feldolgozni. A teszt során a terhelés folyamatosan növekszik egy maximum értékig, utána csökken, majd tönkremegy (eltörik) az anyag. A vizsgálatból határozzuk meg

- 1) tönkrementelt, a szakadás helyét (ábrán: 4)
- 2) a szakítószilárdságot (ábrán: 1)
- 3) az arányossági határt (ábrán: 3)
- 4) a rugalmassági modulust (Young-modulus) (E)
- 5) 0.2% maradó alakváltozáshoz (ábrán: 5) tartozó egyezményes folyáshatárt (ábrán: 2)



Ehhez először töltsük be a mérési adatainkat a **szakitovizsgalat.txt** fájlból! Ebben a mért fajlagos alakváltozásokhoz (ϵ - %) tartozó feszültség (σ - Mpa) értékek találhatóak meg.

```
> % szakítóvizsgálat
> clc; clear all; close all;
> data =
  load('szakitovizsgalat.txt');
> x = data(:,1); % fajlagos
  alakváltozás, epszilon
> y = data(:,2); % feszültség, szigma
> figure(1)
> plot(x,y, '.')
```



Azt, hogy hol megy tönkre az anyag, könnyű meghatározni, az utolsó mérési eredményt kell venni, ez a maximális ϵ érték is egyben. A szakítószilárdság meghatározásához a maximális feszültség (σ) értéket kell megkeresnünk:

```
> disp('Tönkrementelhez tartozó fajlagos alakváltozás:')
> x(end) % ugyanaz, mint a max(x) => 0.2644 %
```


```
> disp('Szakítószilárdság:')
> max(y) % 668.3606 Mpa
```

Tehát 26.4 % fajlagos alakváltozásnál szakadt el az anyag, és 668 MPa volt a maximális feszültség, a szakítószilárdság. A rugalmassági modulus meghatározásához azt a szakaszt kell megkeressük, ahol lineáris az összefüggés a fajlagos alakváltozás és a feszültség között (hook-törvény), és erre kell egy egyenest illeszteni. Az egyenes meredeksége lesz a rugalmassági modulus értéke. Ehhez nagyítsunk bele egy kicsit az ábrába, 0.6 % alakváltozásig és 400 MPa feszültségig!

```
> axis([0 0.006 0 400])
```

A fenti ábrán látszik, hogy a mérés elején, az origó közelében nem tekinthető lineárisnak a mérés a műszer bizonytalansága miatt, így az egyenes illesztéshez le kell vágni az adatok elejét, most vágjuk le, ami kisebb, mint 0.02 %. Meg kell keresni a rugalmassági határ (arányossági határ) felső végét is, vegyük ezt most 0.15 %-nak.

Megjegyzés: a lineáris szakasz végét az ábra alapján állapítottuk meg. Segítségül a 'data

cursor' gombra  kattintva lekérdezhethetjük az egyes pontok adatait. Válogassuk le logikai indexelést használva a lineáris szakasz pontjait, ahol az x koordináta nagyobb, mint 0.0002 és kisebb, mint 0.0015!

```
> feltetel=and(x>0.0002,x<0.0015);
> x1 = x(feltetel);
> y1 = y(feltetel);
> hold on;
> plot(x1,y1,'r*');
```

Vizsgáljuk meg, hogy a leválogatott pontok között tényleg lineáris kapcsolat áll-e fent. Ehhez számoljuk ki a lineáris korrelációs együtthatót:

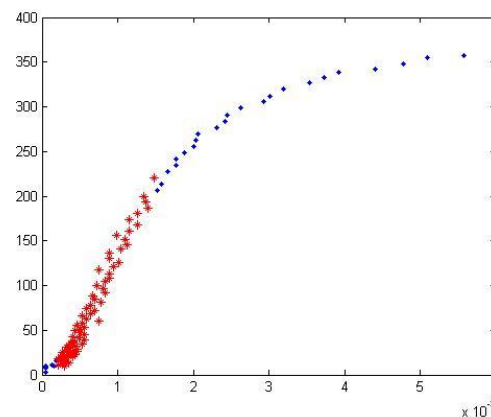
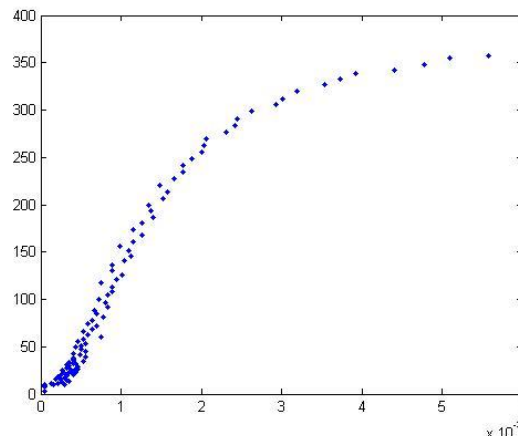
$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

```
> xs = x1-mean(x1)
> ys = y1-mean(y1)
> r = sum(xs.*ys)/sqrt(sum(xs.^2)*sum(ys.^2)) % 0.9805
```

Vagy ugyanez egyszerűbben a matlab beépített **corr2** parancsával:

```
> r = corr2(x1,y1) % 0.9805
```

Minél jobban közelít a korrelációs tényező abszolút értéke az 1-hez, annál inkább lineáris a kapcsolat a két változó között. Mivel most 0.98 lett ez az érték, a kapcsolatot lineárisnak tekinthetjük és illeszthetünk rá egy egyenest. Ehhez nézzük meg az egyenes egyenletét: $y = m \cdot x + b$. Ebben két ismeretlen található m, az egyenes



meredeksége és b eltolás paraméter, ahol az egyenes metszi az y tengelyt. A leválogatott mérések alapján 75 összetartozó x,y értékünk van, ezek alapján 75 egyenletet tudunk felírni, amelyek lineárisak az ismeretleneket tekintve (m,b):

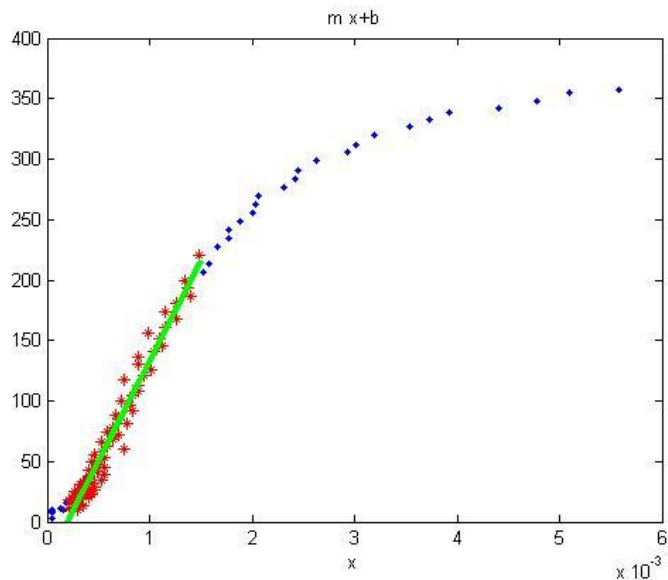
$$\begin{aligned}
 m \cdot x_1 + b &= y_1 \\
 m \cdot x_2 + b &= y_2 \\
 &\vdots \\
 m \cdot x_{75} + b &= y_{75}
 \end{aligned}
 \quad \text{Mátrixos alakban} \quad (A \cdot x = B):
 \quad A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_{75} & 1 \end{pmatrix}; B = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{75} \end{pmatrix}$$

A fenti egyenletrendszerben két ismeretlen van (m,b) és 75 egyenlet, vagyis van 73 fölös mérésünk. Ez egy túlhatározott egyenletrendszer, ahol a maradék eltérések négyzetösszegének minimalizálásával szeretnénk megkapni a legkisebb hibájú megoldást. A korábbi órák alapján ehhez használhatjuk például a túlhatározott esetben QR felbontást alkalmazó $x=A \setminus B$ alakú parancsot, vagy az SVD felbontást használó $x=\text{pinv}(A)*B$ -t is.

Először elő kell állítanunk az A alakmátrixot az ismeretlenek együtthatóival. Az első oszlopban m együtthatói lesznek, vagyis x_i értékei (x_i^1), a másodikban pedig b együtthatója, ami mindig 1, ezt írhatjuk az egyszerűség kedvéért x_i^0 alakba is.

```

> A = [x1.^1 x1.^0]
> B = y1;
> mb = A \ B
> % egyenes egyenlete
> m = mb(1)
> b = mb(2)
> f = @(x) m*x+b
> hold on;
> fplot(f,[0 0.0015], 'g', 'Linewidth',3);
> axis([0 0.006 0 400])
    
```



Az illesztett egyenes meredeksége lesz a rugalmassági modulus értéke:

```

> format long
> E = m % E = 1.656261744954783e+05
    
```

Vagyis a mérés alapján meghatározott E rugalmassági modulus értéke: 165 626 N/mm² (MPa).

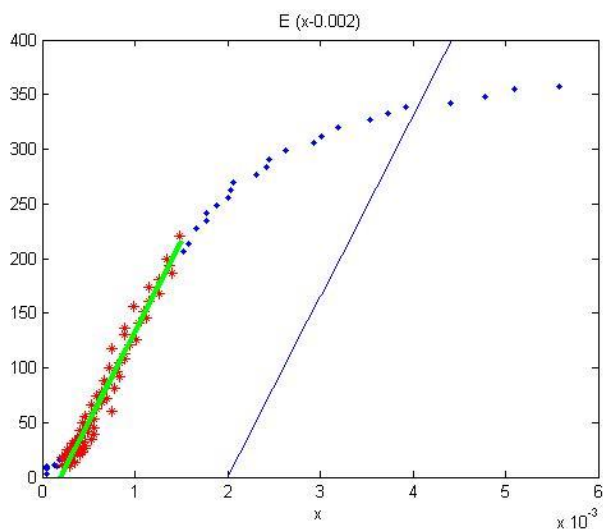
PARABOLA ILLESZTÉS

A folyáshatár megállapítása ebben az esetben nem egyértelmű az ábrából, ennek az anyagnak nincs jól látható folyáshatára. Ilyenkor az egyezményes folyáshatárt szokás használni, ami a 0.2% maradó alakváltozáshoz tartozó feszültség érték. Ezt a szakítódiagramból úgy lehet meghatározni, hogy 0.2% fajlagos nyúlás értékétől párhuzamos egyenest húznak a lineáris szakasszal, vagyis E meredekséggel kell berajzolni egy egyenest ebből a pontból, és ahol ez metszi a szakítógörbét, ott kell leolvasni a feszültséget. Definiáljuk ezt az egyenest és rajzoljuk be az ábrába!

```
> % 0.2%-os folyáshatár
> fhatar = @(x) E*(x-0.002)
> fplot(fhatar,[0 0.006])
> axis([0 0.006 0 400])
```

Hogyan határozzuk meg ennek az egyenesnek és a szakítógörbének a metszéspontját? Jó lenne illeszteni egy függvényt arra a szakaszra is, ahol a metszéspont is található, és ennek a metszéspontját megkeresni az egyenessel. Ehhez válogassuk le az előzőekhez hasonlóan a $x > 0.0015$ és $x < 0.006$ pontokat:

```
> % 0.0015 és 0.006 közötti szakasz
> feltetel2=and(x>0.0015,x<0.006);
> xp = x(feltetel2);
> yp = y(feltetel2);
> plot(xp,yp,'m*')
```

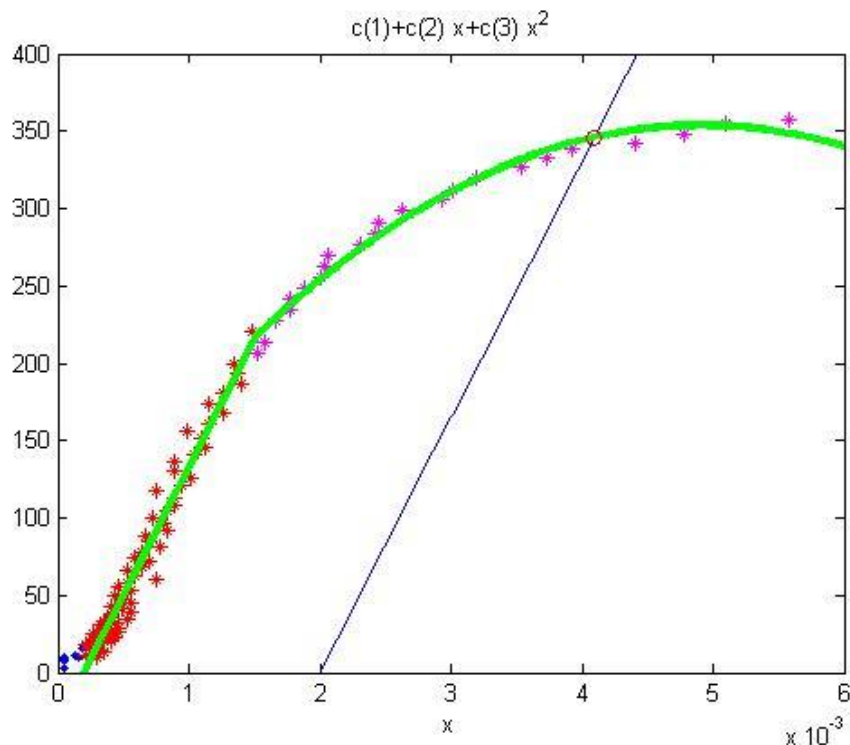


A fenti feltételnek összesen 23 pont felel meg. Erre most illesszünk egy másodfokú polinomot, egy parabolát $y = c_0 + c_1 \cdot x + c_2 \cdot x^2$ alakban!

$$\begin{aligned}
 y_1 &= c_0 + c_1 x_1 + c_2 x_1^2 \\
 y_2 &= c_0 + c_1 x_2 + c_2 x_2^2 \\
 &\vdots \\
 y_n &= c_0 + c_1 x_n + c_2 x_n^2
 \end{aligned}
 \quad \text{Mátrixos alakban:}
 \quad A = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix}; b = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

A parabola (másodfokú polinom) illesztéséhez elő kell állítanunk ismét a megfelelő alakmátrixot (az ismeretlen c_0 , c_1 és c_2 együtthatóit), és megoldanunk egy túlhatározott lineáris egyenlet rendszert! 3 ismeretlenünk van és 23 egyenletet tudunk felírni, tehát van összesen 20 fölös mérésünk.

```
> A = [xp.^0 xp.^1 xp.^2]
> b = yp;
> % túlhatározott lin. egy. rsz. megoldása
> c = A\b
> % parabola egyenlete
> f2 = @(x) c(1) + c(2).*x + c(3)*x.^2
> fplot(f2,[0.0015 0.006], 'g', 'Linewidth', 3)
> axis([0 0.006 0 400])
```



Most már csak a metszéspontot kell megkeressük. Ezt a korábbi tanulmányaink alapján szintén könnyen megtehetjük. Egy $f(x)$ és egy $g(x)$ függvény metszéspontjában $f(x)=g(x)$. Ezt nullára rendezve a $h(x) = f(x)-g(x) = 0$ nemlineáris egyenlet gyökeit kell megkeresni, amit megtehetünk például a beépített **fzero** függvényt alkalmazva!

- > %% 0.2 %-hoz tartozó folyáshatár megállapítása
- > % fhatar(x) = f2(x), vagyis $h(x) = fhatar(x) - f2(x) = 0$
- > h = @(x) fhatar(x) - f2(x)
- > metszes = fzero(h,0.004)
- > folyashatar = f2(metszes) % 345.818 MPa
- > plot(metszes, folyashatar, 'ro')

Az egyezményes, 0.2% maradék alakváltozáshoz tartozó folyáshatár a mérésünknel 346 MPa-ra adódott.

Hasonlóan az eddigiekhez harmad, negyed stb. fokú polinomokat is illeszthetünk az adatainkra. A magasabb fokú polinomoknál azonban vigyáznunk kell, mert az alakmátrixunk rosszul kondicionált lesz és bizonytalan lesz a megoldásunk, a mérési pontjainkra lehet, hogy tökéletesen fog illeszkedni a polinom, de közöttük oszcilláció léphet fel. Erre fogunk majd példát látni az interpolációnál.

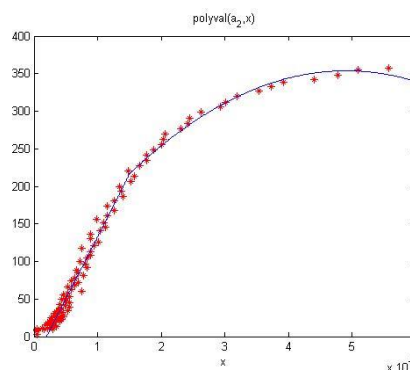
POLINOM ILLESZTÉS MATLAB BEÉPÍTETT FÜGGVÉNYEIVEL (POLYFIT, POLYVAL)

A rugalmassági modulus meghatározásához a feladat első részében egy egyenes illesztésre volt szükség, ami megfelel egy elsőfokú polinomnak, a feladat második részében pedig másodfokú polinomot illesztettünk. Matlab-ban van egy parancs (**polyfit**), amivel összetartozó pontpárokhoz határozhatjuk meg tetszőleges fokszámú polinom együtthatóit. Az eredménye ennek egy vektor lesz, ami a legkisebb négyzetek módszerével illesztett polinom együtthatóit tartalmazza, a legnagyobb fokú tagtól kezdve visszafelé a konstans tagig. Ennek a parancsnak van egy párja is, a **polyval** parancs, ami kiszámolja egy tetszőleges pontban a polinom értékét, ha megadtuk azt a vektort, ami az együtthatókat tartalmazza. Ez utóbbit meghívhatjuk egy konkrét x értékre, vagy definiálhatjuk függvényként x független változóval. Nézzük meg, hogyan oldhattuk volna meg az előző feladat függvény illesztéseit ezekkel a parancsokkal!

```
> % egyenes illesztése (elsőfokú polinom)
> a1 = polyfit(x1,y1,1)
> p1 = @(x) polyval(a1,x)
> % parabola illesztése (másodfokú polinom)
> a2 = polyfit(xp,yp,2)
> p2 = @(x) polyval(a2,x)
>
> figure(2); plot(x,y,'r*'); hold on;
> fplot(p1,[0 0.0015]);
> fplot(p2,[0.0015 0.006]);
> axis([0 0.006 0 400])
```

Vizsgáljuk meg az egyenes és a parabola illesztés minőségét, határozzuk meg az empirikus tapasztalati szórásokat!

```
> r1 = y1-p1(x1);
> S1 = sum(r1.^2);
> n1 = length(x1);
> np1 = 2; f1 = n1-np1 % 73
> szoras1 = sqrt(S1/f1) % 11.063
>
> rp = yp-p2(xp);
> Sp = sum(rp.^2);
> np = length(xp);
> npp = 3; fp = np-npp % 20
> szorasp = sqrt(Sp/fp) % 6.519
```



ISMERT ALAKÚ FÜGGVÉNYEK LINEÁRIS PARAMÉTEREINEK MEGHATÁROZÁSA

Természetesen nem csak polinomillesztések oldhatóak meg lineáris regresszióval, sok más probléma is erre az esetre vezet vissza. Ilyen például, ha szinusz-koszinusz függvénnyel leírható kapcsolat amplitúdóját keressük (pl. harmonikus rezgőmozgás), vagy bármilyen más függvény lineáris paramétereit kell meghatároznunk. Nézzünk erre két példát, ahol a meghatározandó paraméterek (a, b) lineáris alakban szerepelnek:

$$y = a \cdot \sin(\pi \cdot x) + b$$

$$y = a \cdot |x - 0.5|^{3/2} + b$$

Természetesen ezek is lineáris regresszióval megoldhatóak. Nézzük meg a gyakorlatban az első példát, néhány megadott pont esetében!

Adott 5 pont x,y koordinátákkal, ezekre a pontokra illesztünk egy függvényt.

x	0	0.25	0.5	0.75	1
y	0	0.5	1	0.5	0

A hipotézisfüggvényünk alakja a következő:

$$y = a \cdot \sin(\pi \cdot x) + b$$

Határozzuk meg az a, b paramétert lineáris regresszióval! 5 pont alapján 5 (a paraméterekre nézve) lineáris egyenletet tudunk felírni.

$$\begin{aligned}
 y_1 &= a \cdot \sin(\pi \cdot x_1) + b \\
 y_2 &= a \cdot \sin(\pi \cdot x_2) + b \\
 y_3 &= a \cdot \sin(\pi \cdot x_3) + b \\
 y_4 &= a \cdot \sin(\pi \cdot x_4) + b \\
 y_5 &= a \cdot \sin(\pi \cdot x_5) + b
 \end{aligned}$$

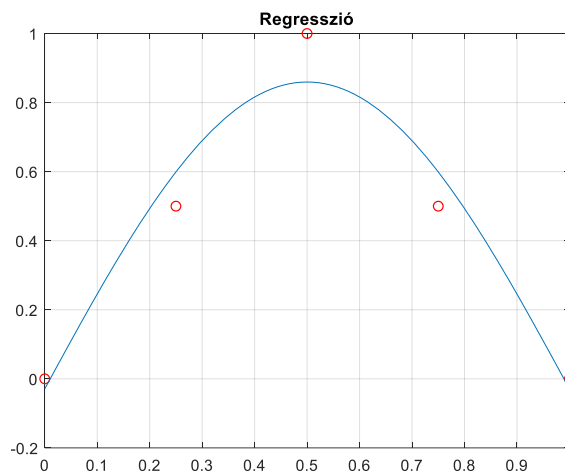
Mátrixos alakban:

$$A = \begin{pmatrix} \sin(\pi \cdot x_1) & 1 \\ \sin(\pi \cdot x_2) & 1 \\ \sin(\pi \cdot x_3) & 1 \\ \sin(\pi \cdot x_4) & 1 \\ \sin(\pi \cdot x_5) & 1 \end{pmatrix}; b = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix}$$

Megoldás Matlab-ban:

```

> xy = [0, 0; 0.25, 0.5; 0.5, 1; 0.75, 0.5; 1, 0]
> x = xy(:,1); y = xy(:,2);
> figure(1); plot(x,y,'ro')
> grid on; title('Regresszió')
> % Együtthatómátrix A
> A = [sin(pi*x), ones(size(x))]
> % paraméterek legkisebb négyzetek módszerével
> c = A\y
> a = c(1) % 0.8887
> b = c(2) % -0.0291
> % függvény megjelenítése
> f = @(x) a*sin(pi*x)+b
> hold on; fplot(f,xlim)
    
```



NEMLINEÁRIS REGRESSZIÓ LINEÁRIS ALAKBA ÍRÁSSAL

A valóságban nagyon sok olyan fizikai jelenség van, ahol a mennyiségek közötti kapcsolat nem lineáris. Például a légsűrűséget (ρ) a magasság (h) függvényében exponenciális függvénnyel lehet modellezni: $\rho = k \cdot e^{m h}$, egy elejtett tárgy v sebessége a megtett x út függvényében az alábbi függvénnyel írható le: $v^2 = 2 g x$.

Nagyon sok nemlineáris függvény van, de most csak azokkal fogunk foglalkozni, amelyeket át lehet úgy alakítani, hogy egy lineáris egyenletrendszer megoldásával megtaláljuk a paramétereket legkisebb négyzetek módszerét alkalmazva. Ilyen többek között a

- hatványfüggvény: $y = k x^m$
- exponenciális függvény: $y = k e^{m x}$ vagy $y = k 10^{m x}$
- reciprokl függvény: $y = \frac{1}{m x + c}$

Az algebrai polinomok is ilyenek, ezeknek az illesztését már láttuk az előző példában.

Az a kérdés, hogyan tudjuk a fenti függvényeket lineáris alakba írni?

Általában új változók bevezetésével tudjuk átalakítani a kétváltozós nemlineáris egyenletet, hogy az új változók (amelyek az eredeti változókból levezethetőek) már lineáris kapcsolatban álljanak a keresett paraméterekkel. Nézzünk erre egy példát, hozzuk lineáris alakra a **hatványfüggvényt**, vegyük mindkét oldal természetes alapú logaritmusát!

$$\ln(y) = \ln(k x^m) = m \ln(x) + \ln(k)$$

Vezessünk be új változókat, hogy $Y = c_1 X + c_2$ lineáris alakra hozzuk az egyenletet. Most legyen $Y = \ln(y)$, $X = \ln(x)$, $c_1 = m$, $c_2 = \ln(k)$:

$$\begin{aligned} \ln(y) &= m \ln(x) + \ln(k) \\ Y &= c_1 X + c_2 \end{aligned}$$

A fenti formában már alkalmazhatjuk a lineáris regressziót, és amint megkaptuk c_1, c_2 értékét az eredeti összefüggés keresett paraméterei könnyen meghatározhatóak:

$$m = c_1, \quad k = e^{c_2}$$

Sok más nemlineáris egyenlet is lineáris alakba hozható hasonlóan. Nézzük meg erre a bevezetőben említett egyik példát!

A légsűrűséget (ρ) a magasság (h) függvényében **exponenciális függvénnyel** lehet modellezni: $\rho = k \cdot e^{m h}$. A következő táblázatban különböző magasságokban mért légsűrűség értékek találhatóak:

h [m]	1	4000	8000	12000	16000	20000
ρ [kg/m ³]	1.225	0.820	0.525	0.309	0.168	0.092

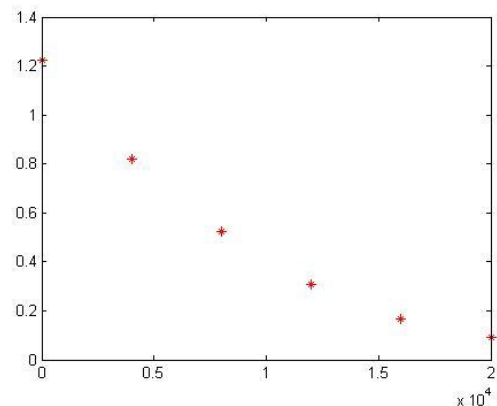
Két kérdésre szeretnénk választ kapni:

- 1) Mekkora lesz a 8850 m magas Csomolungmán a légsűrűség?
- 2) Milyen magasan lesz 1 kg/m³ a légsűrűség?

Lineáris regressziót használva határozzuk meg a legjobban illeszkedő függvényhez a k és m együtthatókat! Vizsgáljuk meg lokálisan és globálisan a görbeillesztés hibáit is!

A megoldáshoz töltsük be a **legsuruseg.txt** állományt, amiben a fenti adatok vannak.

```
> % légsűrűség a magasság függvényében
> clear all; close all; clc;
> data = load('legsuruseg.txt')
> % p=k*e^(m*h) - exponential function,
  m,k = ?
> h = data(:,1) % magasság
> p = data(:,2) % légsűrűség
> figure(1)
> plot(h,p,'r*')
```



Hozzuk lineáris alakra az exponenciális függvényt! $\rho = k \cdot e^{m h}$

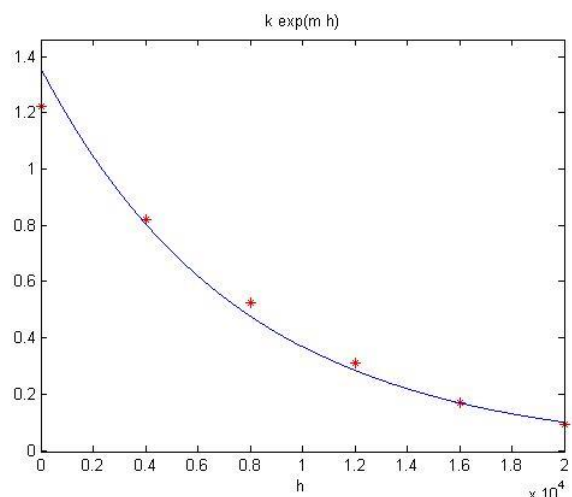
$$\ln(\rho) = m h + \ln(k)$$

$$Y = c_1 X + c_2$$

```
> Y = log(p)
> X = h
> % A*x=b alakban felírva
> A = [X.^1 X.^0]
> b = Y
> % megoldás:
> c = A\b % c1 = m, c2 = ln(k)
```

A keresett paraméterek: $m = c_1$, $k = e^{c_2}$

```
> % a keresett paraméterek
> m = c(1)
> k = exp(c(2))
> % illesztett függvény
> f = @(h) k*exp(m*h)
> % felrajzolva
> hold on;
> fplot(f,[0 20000])
```



Válaszoljunk a kérdésekre. Az egyenletet használva mekkora lesz a 8850 m magas Csomolungmán a légsűrűség? Milyen magasan lesz 1 kg/m³ a légsűrűség? Az első kérdést egy egyszerű behelyettesítéssel megválaszolhatjuk, a másodikhoz az $f(x) = 1$ egyenletet át kell alakítani $g(x) = f(x) - 1 = 0$ alakra és megkeresni ennek a nemlineáris egyenletnek a gyökeit. Ehhez szükséges egy kezdőértéket is megadni, amit vehetünk az ábrából körülbelül 2000-nek (0.2×10^4).

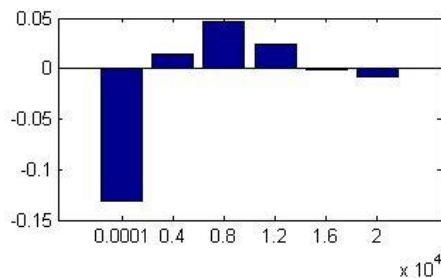
```
> % legnyomas 8850 m magasban
> p8850 = f(8850) % 0.4285
>
> % Milyen magasan lesz 1 kg/m^3 a légsűrűség?
> g = @(h) f(h)-1
> h06 = fzero(g,2000) % 2.3415e+03
```

Tehát a Csomolungmán 0.4285 kg/m³ a levegő sűrűsége, és 2342 m-en lesz pont 1 kg/m³ a sűrűség.

Nézzük meg a maradék eltérések alakulását! Rajzoljuk fel őket egy oszlopdiagramra, és számítsuk ki az eltérések négyzetösszegét és a maradék eltérések korrigált tapasztalati szórását!

```

> % regresszió maradék hibái
> r = p - f(h)
> % maradék eltérések felrajzolása
> figure(2)
> bar(h,r)
> % hibák négyzetösszege
> S = sum(r.^2) % 0.0203
> % korrigált tapasztalati szórás
> n = length(h) % mérések száma
> np = 2 % becsült paraméterek száma: k, m
> szoras = sqrt(S/(n-np)) % 0.0712
    
```



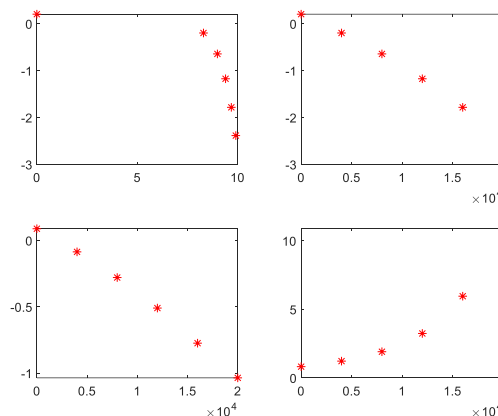
Foglaljunk össze egy táblázatban néhány nemlineáris egyenletet, amit hasonlóképpen megoldhatnánk lineáris regresszióval!

Nemlineáris egyenlet	Lineáris alak	$Y = c_1 X + c_2$ alakban	Keresett paraméterek	Értékek a lineáris regresszióhoz (képe egyenes)
$y = k x^m$	$\ln(y) = m \ln(x) + \ln(k)$	$Y = \ln(y), X = \ln(x),$ $c_1 = m, c_2 = \ln(k)$	$m = c_1$ $k = e^{c_2}$	$\ln(x), \ln(y)$
$y = k e^{m x}$	$\ln(y) = m x + \ln(k)$	$Y = \ln(y), X = x,$ $c_1 = m, c_2 = \ln(k)$	$m = c_1$ $k = e^{c_2}$	$x, \ln(y)$
$y = k 10^{m x}$	$\lg(y) = m x + \lg(k)$	$Y = \lg(y), X = x,$ $c_1 = m, c_2 = \lg(k)$	$m = c_1$ $k = 10^{c_2}$	$x, \lg(y)$
$y = \frac{1}{m x + k}$	$\frac{1}{y} = m x + k$	$Y = 1/y, X = x,$ $c_1 = m, c_2 = k$	$m = c_1$ $k = c_2$	$x, \frac{1}{y}$
$y = \frac{m x}{x + k}$	$\frac{1}{y} = \frac{k}{m} \frac{1}{x} + \frac{1}{m}$	$Y = 1/y, X = 1/x,$ $c_1 = k/m, c_2 = 1/m$	$m = 1/c_2$ $k = c_1/c_2$	$\frac{1}{x}, \frac{1}{y}$

NEMLINEÁRIS EGYENLET TÍPUSÁNAK KIVÁLASZTÁSA¹

Az előző feladatban rendelkezésünkre állt egy modell, hogy milyen alakú összefüggés áll fent a mennyiségek között. Előfordulhat azonban olyan eset is, amikor nem ismerjük a kapcsolatot leíró függvény alakját. Hogyan választhatjuk ki a megfelelő illesztendő nemlineáris egyenletet ilyen esetben? Célszerű felrajzolni a pontokat és a fenti táblázat utolsó oszlopában lévő értékeket. Ha van olyan, amelyik esetében a pontok nagyjából egy egyenes mentén helyezkednek el, akkor azt a függvény típust választjuk a regresszióhoz! Néhányat ábrázoljunk az előző példához ezek közül! Matlab-ban a természetes alapú logaritmus a **log** függvény, 10-es alapú logaritmus a **log10** függvény, exponenciális függvény pedig az **exp** függvény. Az ábrázoláshoz használjuk a **subplot** függvényt, amivel egy ábrára több dolgot is fel tudunk rajzolni. A parancsot a **subplot(n,m,i)** formában hívhatjuk meg, ahol n a sorok, m az oszlopok száma, i pedig az adott rajz sorszáma balról jobbra és fentről le számolva.

```
> figure(3)
> x=h; y = p;
> subplot(2,2,1)
> plot(log(x), log(y), 'r*')
> subplot(2,2,2)
> plot(x,log(y), 'r*')
> % egyenes lett!
> subplot(2,2,3)
> plot(x,log10(y), 'r*')
> % egyenes lett!
> subplot(2,2,4)
> plot(x,1./y, 'r*')
```



A fenti rajzon a második és a harmadik képen lett a pontok képe közelítőleg egyenes, amikor x és $\ln(y)$ illetve $\log(y)$ lett ábrázolva. A táblázatra ránézve látszódik, hogy abban az esetben, ha nem ismernénk a függvénykapcsolatot, akkor is exponenciális függvény illesztésével lenne érdemes próbálkozni.

A FEJEZETBEN HASZNÁLT ÚJ FÜGGVÉNYEK

- | | |
|---------|---|
| axis | - Tengelyek minimális, maximális értékeinek megadása |
| mean | - Vektor elemeinek számtani közepe, átlaga |
| sum | - Vektor elemeinek összege |
| corr2 | - Lineáris korrelációs együttható |
| polyfit | - Megadott fokszámú polinom illesztése az adatokra |
| polyval | - Együttható vektorral megadott polinom értékének kiszámítása |
| bar | - Ábrázolás oszlopdigrammon |
| subplot | - Egy grafikus ablakon belül több ábra |

¹ Kiegészítő anyag