Minsta-kvadratmetoden

1 Inledning

Ett ofta förekommande problem inom teknik och vetenskap är att koppla samman mätdata med en formel eller kurva som man vill verifiera eller bygga upp.

2 Minsta-kvadratmetoden

Ett klassiskt problem är att anpassa en rät linje $y = a+b\cdot t$ till givna mätdata $(t_i, y_i),\ i = 1, \cdots, n$. Hur skall vi välja a och b? Vi kan ju inte få den räta linjen att gå igenom mer än högst två punkter. Problemet vi skall lösa är följande överbestämda ekvationssystem

$$\begin{cases} a + b \cdot t_1 = y_1 \\ a + b \cdot t_2 = y_2 \\ \qquad \vdots \\ a + b \cdot t_n = y_n \end{cases}$$

(det är a och b som är de obekanta!) På matrisform,

$$\underbrace{\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} a \\ b \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{\mathbf{y}}$$

Grundidén i minsta-kvadratmetoden är att projicera vektorn $\mathbf{y}$ ortogonalt på kolonnrummet för matrisen $\mathbf{A}$ ($Col(\mathbf{A})$) och sedan lösa ekvationen $\mathbf{Ax} = \mathbf{p}$ där $\mathbf{p}$ är projektionen. På så vis får vi en lösning $\hat{\mathbf{x}}$ där avståndet $\|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}\|$ är det minsta möjliga och väljer vi nu $a = \hat{x_1},\ b = \hat{x_2}$ så har vi minimerat summan av kvadraterna på avvikelserna:

$$\sum_{i=1}^{n} (a + b \cdot t_i - y_i)^2$$

Lösningen $\hat{\mathbf{x}}$ till problemet $\mathbf{Ax} = \mathbf{p}$ ovan säges vara *minsta-kvadratlösningen* till det ursprungliga ekvationssystemet $\mathbf{Ax} = \mathbf{y}$. Vi finner den genom att lösa den så kallade *normalekvationen*,

$$\mathbf{A}^\mathsf{T}\mathbf{Ax} = \mathbf{A}^\mathsf{T}\mathbf{y} \tag{1}$$

(Notera att vi inte behöver bestämma projektionen $\mathbf{p}$ för att bestämma $\hat{\mathbf{x}}$.)

Det *kvadratiska medelfelet*, den genomsnittliga avvikelsen, ges av

$$\varepsilon = \|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}\|/\sqrt{n}$$

där n är antalet mätdata.

Låt oss nu bestämma den räta linje som i minsta-kvadratmening är bäst anpassad till följande data:

t	-1	0	1	2	3	4	6	7
y	-0.75	0.3	3	4	5.6	7	6.4	8.4

Då är

$$\mathbf{A}^{\mathsf{T}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 & 3 & 4 & 6 & 7 \end{bmatrix}, \ \mathbf{A}^{\mathsf{T}}\mathbf{A} = \begin{bmatrix} 8 & 22 \\ 22 & 116 \end{bmatrix}, \ \mathbf{A}^{\mathsf{T}}\mathbf{y} = \begin{bmatrix} 33.95 \\ 153.75 \end{bmatrix}$$

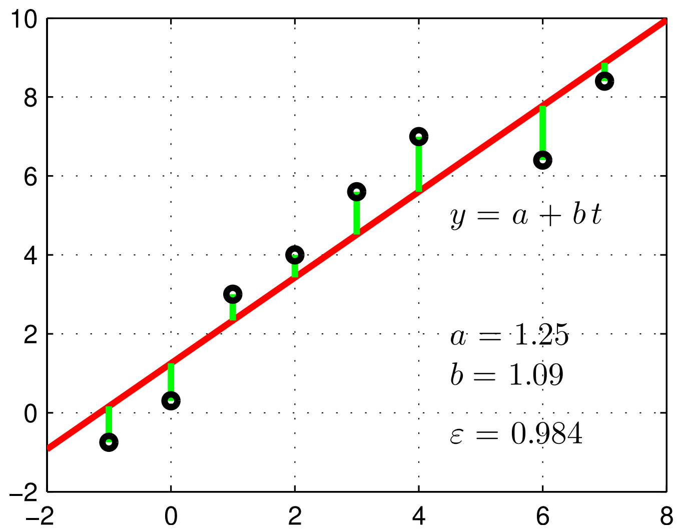
och normalekvationen är

$$\begin{bmatrix} 8 & 22 \\ 22 & 116 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 33.95 \\ 153.75 \end{bmatrix}$$

I MATLAB löser vi (1) med kommandot \ enligt

```
>> td=[-1 0 1 2 3 4 6 7]';                 % t-data
>> yd=[-0.75 0.3 3 4 5.6 7 6.4 8.4]';      % y-data
>> A=[ones(size(td)) td];                  % Designmatrisen
>> x=A\yd; a=x(1); b=x(2);                 % Minsta-kvadratlösningen
>> n=length(td);                           % Antalet mätdata
>> e=norm(A*x-yd)/sqrt(n);                 % Kvadratiska medelfelet
```

Vi kan nu rita upp följande figur



Vi minimerar summan av kvadraterna på de lodräta sträckorna.

Uppgift 1. Antag att variablerna t och y uppfyller sambandet $y = a + b \cdot t$. För att betämma koefficienterna a och b utför vi mätningar av t och y:

t	5	6	7	8	9	10
y	19.5888	23.4043	25.5754	29.1231	31.9575	35.8116

Tabell 1: Mätvärden av y för vissa t.

(a). Lös normalekvationen (1) med minsta kvadratmetoden i MATLAB. Bestäm också kvadratiska medelfelet.

(b). Rita upp datapunkterna (t_i, y_i) och den anpassade funktionen $y = a + b \cdot t$ i samma figur.

3 Tillämpning, Arrhenius ekvation

Uppgift 2. Arrhenius ekvation lyder $k = k_0 e^{-E/(RT)}$.

(a). Logaritmera för hand på papper båda sidorna av Arrhenius ekvation.

(b). Bestäm konstanten k_0 och kvoten E/R från informationen i Tabell 2. Detta gör vi genom att den logaritmerade ekvationen ger en linjär relation mellan $y = \ln(k)$ och $t = 1/T$ på formen $y = a + b \cdot t$. Använd Tabell 2 till att generera datapunkter (t_i, y_i), där $t_i = 1/T_i$ och $y_i = \ln(k_i)$. Bilda ett linjärt ekvationssystem och lös det med minsta kvadratmetoden. Bestäm också kvadratiska medelfelet. I MATLAB ges $\ln(k)$ av `log(k)`.

(c). Rita ut datapunkterna (t_i, y_i) och den anpassade funktionen $y = a + b \cdot t$ i samma figur.

T[K]	343	353	363	373	383	393	403
k[s^{-1}]	2.8 10^{-5}	5.6 10^{-5}	11.2 10^{-5}	22.4 10^{-5}	44.8 10^{-5}	89.6 10^{-5}	179.2 10^{-5}

Tabell 2: Data till Arrhenius ekvation.

4 Allmän formulering

Härnäst skall vi bestämma den kurva

$$y = c_1 \cdot f_1(t) + c_2 \cdot f_2(t) + \cdots + c_k \cdot f_k(t)$$

där $f_1,\ f_2,\ \cdots,\ f_k$ är kända funktioner, som i minsta kvadratmening är bäst anpassad till givna mätdata

t	t_1	t_2	$\cdots$	t_n
y	y_1	y_2	$\cdots$	y_n

Matrisformuleringen av det överbestämda ekvationssystem vi intresserar oss för ges av

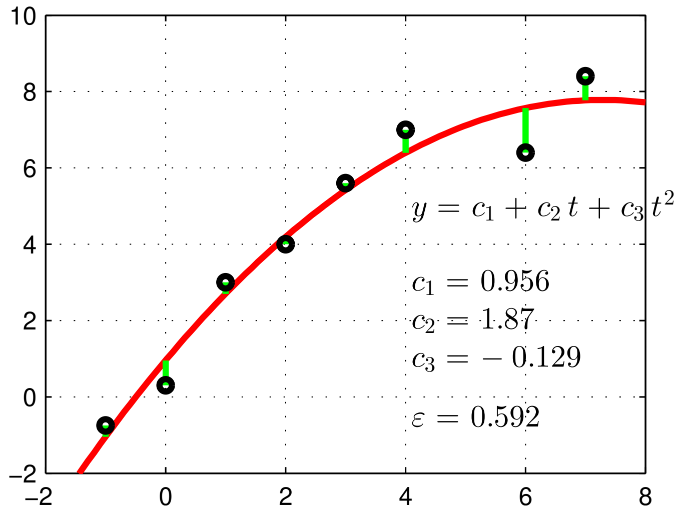
$$\underbrace{\begin{bmatrix} f_1(\begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}) & f_2(\begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}) & \cdots & f_k(\begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}) \end{bmatrix}}_{\text{A (designmatrisen)}} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{\mathbf{y}}$$

Precis som tidigare finner vi minsta-kvadratlösningen genom att lösa normalekvationen (1). Låt oss återgå till vårt första exempel ovan och istället anpassa ett andragradspolynom, $y = c_1 + c_2 \cdot t + c_3 \cdot t^2$, till givna mätpunkter. Här är $f_1(t) = 1,\ f_2(t) = t,\ f_3(t) = t^2$ och designmatrisen $\mathbf{A}$ och minsta-kvadratlösningen skapas i MATLAB på liknande sätt som tidigare

```
>> td=[-1 0 1 2 3 4 6 7]';              % t-data
>> yd=[-0.75 0.3 3 4 5.6 7 6.4 8.4]';   % y-data
>> A=[ones(size(td)) td td.^2];         % Designmatrisen
>> c=A\yd;                              % Minsta-kvadratlösningen
```

```
>> n=length(td);                          % Antalet mätdata
>> e=norm(A*c-yd)/sqrt(n);                % Kvadratiska medelfelet
>> y=@(t)c(1)+c(2)*t+c(3)*t.^2;
>> t=linspace(-2,8);
>> plot(td,yd,'ko',t,y(t),'r','linewidth',2)
```

Vi kan nu rita figuren



Vi minimerar summan av kvadraterna på de lodräta sträckorna.

Uppgift 3. Tabellen nedan visar medeltemperaturen under tidsperioden 1961-1990 för årets tolv månader i Göteborg.

Månad	1	2	3	4	5	6	7	8	9	10	11	12
Temp °C	-0.9	-0.9	2.0	6.0	11.6	15.5	16.6	16.2	12.8	9.1	4.4	1.0

Eftersom detta är ett periodiskt förlopp skall vi anpassa följande modell till mätdata:

$$y(t) = c_1 + c_2 \sin(\omega t) + c_3 \cos(\omega t)$$

Här är

$$f_1(t) = 1,\;\; f_2(t) = \sin(\omega t),\;\; f_3(t) = \cos(\omega t)$$

Välj ett lämpligt värde på ω (tänk på periodlängder) och bestäm c_1, c_2 och c_3 med minsta-kvadratmetoden. Bestäm också kvadratiska medelfelet. Rita grafen av funktionen $y(t)$ och rita ut mätdata i samma bild så att man kan se om modellen ansluter väl till mätdata.

Uppgift 4. Tabellen nedan visar (en del av) utfallet y av ett försök då man varierar två faktorer s och t.

s	41.9	43.4	43.9	44.5	47.3	47.5	47.9	50.2	⋯
t	29.1	29.3	29.5	29.7	29.9	30.3	30.5	30.7	⋯
y	251.3	251.3	248.3	267.5	273.0	276.5	270.3	274.9	⋯

(a). Datafilen `labdata.mat` på studiohemsidan innehåller hela uppsättningen data lagrad i tre variabler `sd`, `td` och `yd`. Hämta filen och ladda in i MATLAB med `load` och rita upp data med `plot3` enligt

```
>> load('labdata')
>> smin=min(sd); smax=max(sd); tmin=min(td); tmax=max(td);
>> plot3(sd,td,yd,'ro','linewidth',2), hold on
>> axis([smin smax tmin tmax]), axis vis3d, box on
```

Vänd och vrid för att se om data ligger nära ett plan.

(b). Anpassa en linjär modell

$$y = a + b \cdot s + c \cdot t$$

till data med minsta-kvadratmetoden. Rita ut, tillsammans med mätdata, det plan som beskriver modellen.

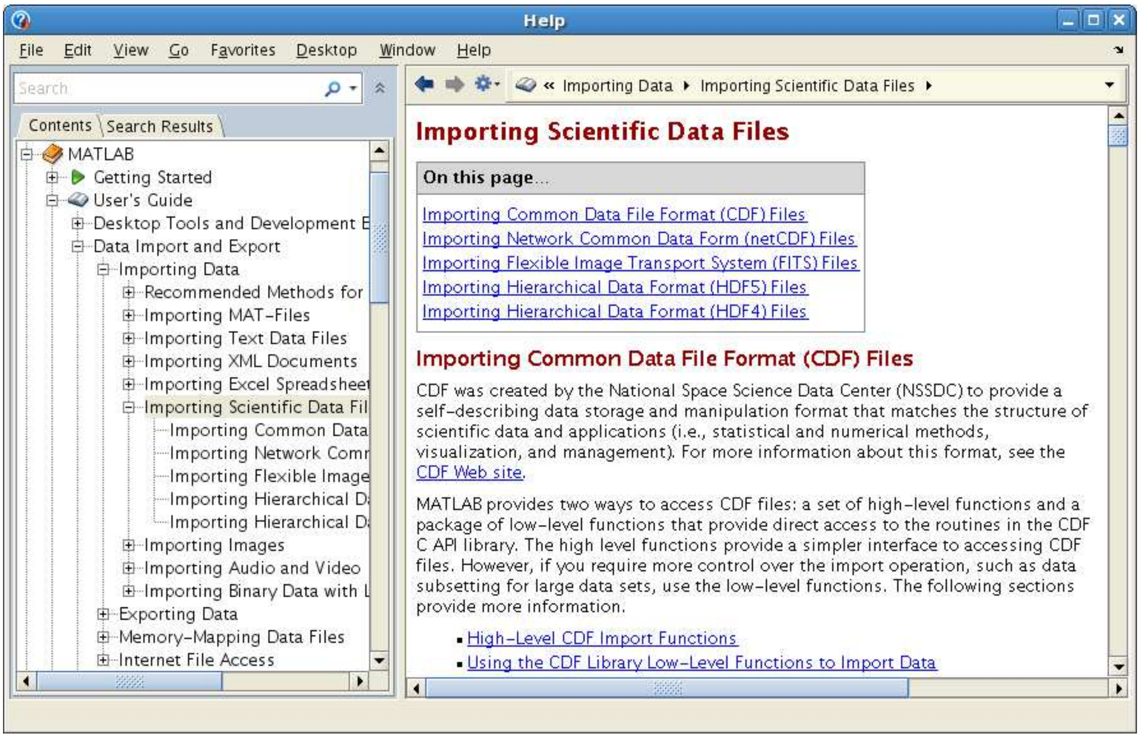
Om koefficienterna ligger samlade i en vektor `c` så ritar du upp modellen med

```
>> y=@(s,t)c(1)+c(2)*s+c(3)*t;
>> S=[smin smax smax smin]; T=[tmin tmin tmax tmax];
>> fill3(S,T,y(S,T),'b','facealpha',0.2)
>> xlabel('s'), ylabel('t'), zlabel('y(s,t)'), grid on
```

Vänd och vrid för att se om ert plan ligger nära data.

5 Import av data

För att hantera mätdata behöver man bl.a. kunna hantera olika fil-format. För import av data har MATLAB en Import Wizard som man kan läsa om i Helpdesk.



Vid behov läser man mer på detta i Helpdesk.