Käyrän sovitusta PNS - menetelmällä

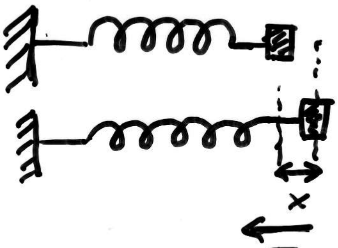
PNS : Pienimmän neliösumman ...
LSQ : Least squares

Lähteitä : Lay : 6.6 : Application to Linear Models
KRE : 20.5 : Least Squares Method
[Esitystapa poikkeaa täysin luentotyylistä ja on kovasti rajoittunut]
Leon : Linear Algebra with applications, Pearson, 2006, Ch 5. [Hyvä oheislukemistokirja]

Johdantoesim : Jousivakio

Hooken laki : Jousivoima on verrannollinen venymään.

$F = kx$



Jousi lepotilassa

venytetty jousi

x

F : jousivoima

Teekkarit T_1 ja T_2 fysiikan labratöissä. Jousivakio?

Mittaukset :

F/N	x/cm
3	4
5	7
8	11

$$\begin{cases} 4k = 3 & \Rightarrow k = \frac{3}{4} \\ 7k = 5 & \Rightarrow k = \frac{5}{7} \\ 11k = 8 & \Rightarrow k = \frac{8}{11} \end{cases}$$

Ristiriitaiset yhtälöt.

Mikä neuvoksi ? Teekkari T_i ($i = 1$ tai 2)
Mikäs se olikaan se PNS ?

$$\begin{bmatrix} 4 \\ 7 \\ 11 \end{bmatrix} k = \begin{bmatrix} 3 \\ 5 \\ 8 \end{bmatrix}$$

(Harvinaisen laiha kerroinmatriisi!)

$$\begin{bmatrix} 4 & 7 & 11 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \\ 11 \end{bmatrix} k = \begin{bmatrix} 4 & 7 & 11 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ 8 \end{bmatrix}$$

$$\Leftrightarrow \quad 186k = 135 \quad \Rightarrow \quad \boxed{k \simeq 0{,}726}$$

Teekkari T_{2-i+1}: "Entä, jos oltaisiin laskettu keskiarvot" Tällöin $\boxed{k \simeq 0.731}$

T_i : "Ainakin edellisessä oli tyyliä!"

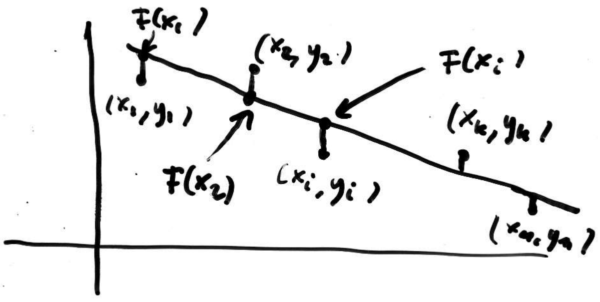
Jos lukua k pidetään vakiofunktiona, tämä voidaan ajatella 0-asteisena käyrän (polynomin) sovituksena

PNS - suora

Olkoon annettu mittauspisteet $(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)$.

Teoreettisesta (fysikaalisesta, ..., psykologisesta, silmämääräisestä) syystä voidaan olettaa, että taustalla olevaa ilmiötä voidaan mallintaa suoralla:

$$F(x) = c_0 + c_1 x$$



Neliösumma:

$$S(c_0, c_1) = \sum_{i=1}^{m} (y_i - F(x_i))^2$$

Tehtävä: Määrätään kertoimet c_0 ja c_1

s.e. $S(c_0, c_1) = \min$.

DATA		MALLIN ENNUSTE
x_1	y_1	$F(x_1) = c_0 + c_1 x_1$
x_2	y_2	$F(x_2) = c_0 + c_1 x_2$
$\vdots$	$\vdots$	
x_n	y_n	$F(x_n) = c_0 + c_1 x_n$

$$\begin{cases} c_0 + c_1 x_1 = y_1 \\ c_0 + c_1 x_2 = y_2 \\ \quad\vdots \\ c_0 + c_1 x_n = y_n \end{cases} \qquad \underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} c_0 \\ c_1 \end{bmatrix}}_{\vec{c}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{\vec{y}}$$

---

yhtälöryhmän $A\vec{c} = \vec{y}$ PNS-ratkaisu merkitsee sitä, että etsitään col A:n vektori $c_0[1, 1, \dots, 1]^T + c_1[x_1, x_2, \dots, x_n]^T$, jonka etäisyys vektorista $\vec{y} = [y_1, \dots, y_n]^T$ on mahdollisimman pieni.

Tämä etäisyys on juuri

$$(c_0 + c_1 x_1 - y_1)^2 + (c_0 + c_1 x_2 - y_2)^2 + \dots + (c_0 + c_1 x_n - y_n)^2$$

$$= S(c_0, c_1).$$

Ratkaisu normaaliyhtälöillä:

$$A^T A \vec{c} = A^T \vec{y}$$

[PNS-suoran tapauksessa normaaliyhtälöratkaisu on numeerisestikin ihan ok., sillä $A^T A$ on 2×2.]

Esim. Määritä PNS - suora, kun

xdata : 2 5 7 8
ydata : 1 2 3 3

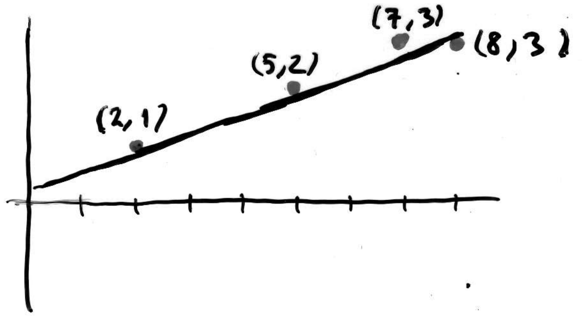
$$\underline{X} = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix}, \quad \vec{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

(Merk. vaihtoehs. $\underline{X}$)

NORMAALIYHT.: $\underline{X}^T \underline{X} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \underline{X}^T \vec{y} \iff \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$

$$\left(\begin{matrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \\ [\text{— '' —}] [1\ 2\ 3\ 3]^T = [9\ \ 57]^T \end{matrix} \right)$$

$$\Rightarrow \vec{c} = \left[\frac{2}{7}, \frac{5}{14}\right]^T; \quad y = \frac{2}{7} + \frac{5}{14}x$$



Matlab:

```
>> xd = [2 5 7 8]; yd=[1 2 3 3];
>> m = length(xd) % yleensä datan paljous
>> X = [ones(m,1) xd]
>> XX = X'*X; XY = X'*yd';
>> c = XX\XY
>> x = [0 9]; y = c(1)+c(2)*x;
>> plot(x,y,'r'); hold on; plot(xd,yd,'o'); grid
```

Yleinen lineaarinen malli

Edellinen on helppo yleistää polynomimallille. Jos haluttaisiin sovittaa dataan 2. asteen polynomi, niin edellä oleva funktiomalli saisi muodon $F(x) = c_0 + c_1 x + c_2 x^2$. Tämä johtaisi PNS-tehtävään:

$$\underbrace{\begin{bmatrix} 1 & x_0 & x_0^2 \\ 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix}}_{\underline{X}} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \underbrace{\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}}_{\vec{y}}$$

(Vaihdettiin data-indeksointi alkamaan 0:sta → n+1 datapistettä.)

Tätä voidaan jatkaa n-asteiseen polynomiin (n+1 kerrointa) saakka.
Tällöin ollaankin neliömatriisissa ja interpolaatiotehtävässä, jolla on aina 1-käs. ratkaisu, kun x_i-pisteet erillisiä.
Voidaan osoittaa esim. Lagrangen menetelmällä.
(Ks. vaikka Luento1-36.html:n "Solmu"-linkki.)
Mallia ei ole yleensä järkevää pakottaa kulkemaan tarkasti datapisteiden kautta. Korkeaasteinen polynomi heilahtelee voimakkaasti. Ääriesimerkkinä voisi olla suoraa mallintava data, jossa on hiukan virhettä. Tällöin esim. 10 datapisteen kautta kulkeva 9 asteinen polynomi johtaisi totaalisesti harhaan.
Ks. L/PNSsuorajaINTpol.html

<u>Yleisesti</u>: Dataa kuvaava lineaarinen funktiomalli voidaan muodostaa annettujen "kantafunktioiden" lineaarikombinaationa:

$$y = F(x) = c_0 \phi_0(x) + c_1 \phi_1(x) + \dots + c_k \phi_k(x)$$

Polynomimallissa siis $\phi_j(x) = x^j$.

Annettu data: $x_0, \dots, x_n$
$y_0, \dots, y_n$

$$F(x_i) = c_0 \phi_0(x_i) + c_1 \phi_1(x_i) + \dots + c_k \phi_k(x_i)$$

Aivan samoin kuin edellä, PNS-tehtävä

$$\min_{c_0, \dots, c_k} \sum (y_i - F(x_i))^2$$

johtaa ylimääräiseen yhtälöryhmään $X\vec{c} = \vec{y}$,

missä

$$X = \begin{bmatrix} \phi_0(x_0) & \phi_1(x_0) & \cdots & \phi_k(x_0) \\ \phi_0(x_1) & \phi_1(x_1) & \cdots & \phi_k(x_1) \\ \vdots & \vdots & & \vdots \\ \phi_0(x_n) & \phi_1(x_n) & \cdots & \phi_k(x_n) \end{bmatrix}$$

<u>Esim.</u> Jos data on jaksollista, voidaan ottaa kantafunktioiksi esim.

$\phi_0 \equiv 1$, $\phi_1(x) = \sin x$, $\phi_2(x) = \cos(x)$,
$\phi_3(x) = \sin 2x$, $\phi_4(x) = \cos 2x$
...

<u>HUOM!</u> "Lineaarinen" tarkoittaa lineaarisuutta parametrien $c_0, c_1, \dots$ suhteen. Mallintava funktio $F(x)$ voi olla vaikka kuinka "kiharainen".

Muita keinoja PNS - laskentaan

Normaaliyhtälöt : $A^T A \vec{x} = A^T \vec{b}$.
Numeerisesti raskas, jos mallimatriisissa A on paljon sarakkeita (ja jos dataa paljon).

Mikä neuvoksi? 1) QR - hajotelma

Lause 12 (Lay 6.4 s. 405)

A (m×n), LRT sarakkeet. Tällöin A voidaan esittää muodossa $A = QR$, missä
Q (m×n), ON sarakkeet
R (n×n), yläkolmiomatr., diag. alkiot $\neq 0$.

Tod Luonn.: Gram-Schmidt-ortonormeeraus
$\Rightarrow$ Q:n sarakkeet.
Menetelmä antaa kolmiomaisen matriisin kertoimia r_{kj}, mistä rakentuu R $\square$

Tätä voidaan pitää lähinnä olemassaolo-todistuksena.
Oikeasti hajotelma muodostetaan ns. Householderin muunnoksilla vähän samaan tyyliin kuin LU-hajotelma rakennettiin alkeismatriiseilla kertomalla.

Kts. KRE[9] 20.9 : Tridiagonalization and QR - Fact.

KRE - kirjassa QR - hajotelma esitetään menetelmänä ominaisarvojen ja -vektorien laskentaan.

PNS - ratk. $A^T A \vec{x} = A^T \vec{b}$, $A = QR$, $Q^T Q = I$
R kääntyvä yläkolmiom.

$\underbrace{(QR)^T}_{R^T Q^T} Q R \vec{x} = (QR)^T \vec{b}$, $Q^T Q = I$

$\Leftrightarrow R^T R \vec{x} = R^T Q^T \vec{b}$ | kerr. vas $(R^T)^{-1}$:llä

$\Leftrightarrow R \vec{x} = Q^T \vec{b}$

RATKAISU TAKAISINSIJOITUKSELLA