Taivaanmekaniikka
Lineaarinen pienimmän neliösumman sovitus

Anni Järvenpää

22. lokakuuta 2015

1 Lineaarinen pienimmän neliösumman menetelmä

Lineaarisen pienimmän neliösumman menetelmän tavoitteena on sovittaa n muotoa (x_i, y_i) olevasta pisteestä koostuvaan havaintoaineistoon suora, joka edustaa pisteitä mahdollisimman hyvin. Tyypillisesti sovituksen hyvyyttä mitataan pisteiden vertikaalisena etäisyytenä $|e|$ sovitetusta suorasta (merkitty sinisellä kuvassa 1). Pienimmän neliösumman menetelmässä näiden vertikaalisten poikkeamien neliöiden summa pyritään minimoimaan, siis etsimään funktion $f(x) = \beta_1 + \beta_2 x + \beta_3 x^2 + ... + \beta_{m+1} x^m$ kertoimet $\beta_1 ... \beta_{m+1}$ siten, että virhe S on mahdollisimman pieni, kun S on määritelty yhtälön 1 mukaisesti. [3]

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

Virheen minimiarvot löytyvät derivaatan nollakohdista: [1]

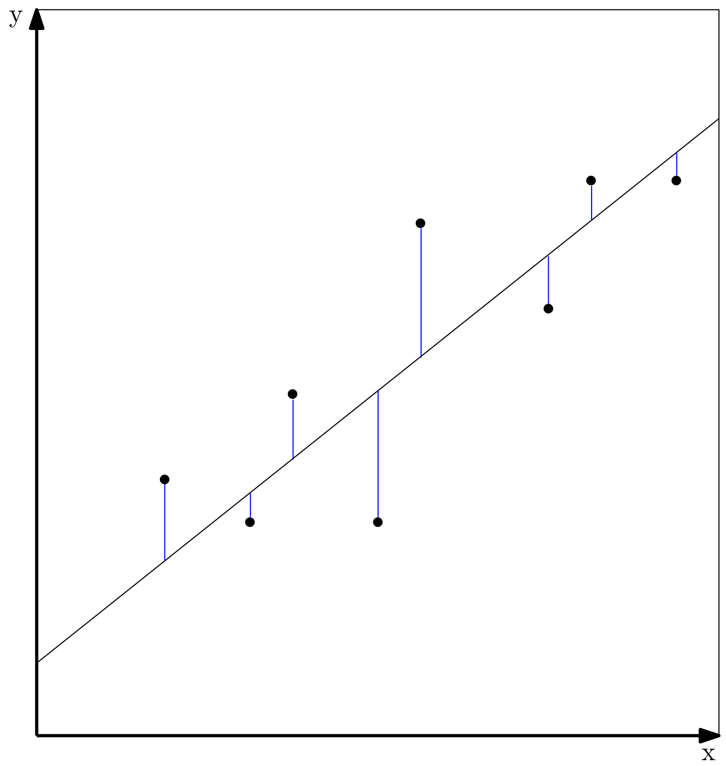
$$\begin{aligned}
\frac{\partial R^2}{\partial \beta_1} &= -2 \sum_{i=1}^{n} [y - (\beta_1 + \beta_2 x + ... + \beta_{m+1} x^m)] = 0 \\
\frac{\partial R^2}{\partial \beta_2} &= -2 \sum_{i=1}^{n} [y - (\beta_1 + \beta_2 x + ... + \beta_{m+1} x^m)] = 0 \\
&\vdots \\
\frac{\partial R^2}{\partial \beta_{m+1}} &= -2 \sum_{i=1}^{n} [y - (\beta_1 + \beta_2 x + ... + \beta_{m+1} x^m)] = 0
\end{aligned}$$

Näistä saadaan edelleen

$$\begin{aligned}
\beta_1 n + \beta_2 \sum_{i=1}^{n} x_i + ... + \beta_{m+1} \sum_{i=1}^{n} x_i^m &= \sum_{i=1}^{n} y_i \\
\beta_1 \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2 + ... + \beta_{m+1} \sum_{i=1}^{n} x_i^{m+1} &= \sum_{i=1}^{n} x_i y_i \\
\beta_1 \sum_{i=1}^{n} x_i^m + \beta_1 \sum_{i=1}^{n} x_i^{m+1} + ... + \beta_{m+1} \sum_{i=1}^{n} x_i^{2m} &= \sum_{i=1}^{n} x_i^m y_i
\end{aligned}$$

eli matriisimuodossa

$$\begin{bmatrix}
n & \sum_{i=1}^{n} x_i & ... & \sum_{i=1}^{n} x_i^m \\
\sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & ... & \sum_{i=1}^{n} x_i^{m+1} \\
\vdots & \vdots & \ddots & \vdots \\
\sum_{i=1}^{n} x_i^m & \sum_{i=1}^{n} x_i^{m+1} & ... & \sum_{i=1}^{n} x_i^{2m}
\end{bmatrix}
\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \\ \beta_{m+1} \end{bmatrix}
=
\begin{bmatrix} \sum_{i=1}^{n} y_1 \\ \sum_{i=1}^{n} y_2 \\ \vdots \\ \sum_{i=1}^{n} y_m \end{bmatrix}. \tag{2}$$



Kuva 1: Pistejoukko, johon sovitettu suora mustalla ja pisteiden vertikaaliset etäisyydet suorasta pisteinä.

Voidaan myös huomata, että yhtälöön 2 voidaan päästä hyödyntämällä Vandermonden matriisia [2]

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_1^2 & \dots & x_1^n \\ 1 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_m^2 & \dots & x_m^n \end{bmatrix} \tag{3}$$

ja kirjoittamalla tämän avulla

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ 1 & x_3 & x_3^2 & \dots & x_3^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_m^2 & \dots & x_m^n \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{m+1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}. \tag{4}$$

Nyt voidaan kertoa edellinen yhtälö puolittain vasemmalta Vandermonden matriisin transpoosilla $\boldsymbol{X}^T$, jolloin yhtälö saadaan seuraavaan muotoon:

$$\begin{bmatrix} 1 & 1 & 1 & ... & 1 \\ x_1 & x_2 & x_3 & ... & x_m \\ x_1^2 & x_2^2 & x_3^3 & ... & x_m^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^n & x_2^n & x_3^n & ... & x_m^n \end{bmatrix} \begin{bmatrix} 1 & x_1 & x_1^2 & ... & x_1^n \\ 1 & x_2 & x_2^2 & ... & x_2^n \\ 1 & x_3 & x_3^2 & ... & x_3^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_m^2 & ... & x_m^n \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{m+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & ... & 1 \\ x_1^2 & x_2^2 & ... & x_m^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^n & x_2^n & ... & x_m^n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Suoritetaan yhtälön matriisien kertolaskut, jolloin saadaan

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_i & ... & \sum_{i=1}^{n} x_i^m \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & ... & \sum_{i=1}^{n} x_i^{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} x_i^m & \sum_{i=1}^{n} x_i^{m+1} & ... & \sum_{i=1}^{n} x_i^{2m} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{m+1} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_1 \\ \sum_{i=1}^{n} y_2 \\ \vdots \\ \sum_{i=1}^{n} y_m \end{bmatrix}. \tag{5}$$

Nyt voidaan huomata yhtälöiden 5 ja 2 olevan identtiset. Koska Vandermonden matriisi on kääntyvä[1], on kaikille yhtälöstä 4 yhtälöön 5 pääsemiseksi tehdyille operaatioille käänteisoperaatio. Täten yhtälöt 4 ja 2 ovat yhtäpitävät, eli pienimmän neliösumman antavat kertoimet β_i voidaan ratkaista yhtälöstä 4.

Otetaan kerroinvektorin ratkaisemisen helpottamiseksi käyttöön seuraavat merkinnät:

$$\boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \text{ja} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix},$$

jolloin yhtälö 2 voidaan kirjoittaa lyhyesti

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} \tag{6}$$

ja tästä ratkaista $\boldsymbol{\beta}$ seuraavasti:

$$\begin{aligned} \boldsymbol{Y} &= \boldsymbol{X}\boldsymbol{\beta} && ||\boldsymbol{X}^T\cdot \\ \boldsymbol{X}^T\boldsymbol{Y} &= \boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta} && ||(\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta})^{-1} \\ \boldsymbol{\beta} &= (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y} \end{aligned} \tag{7}$$

[1]Mikä voidaan helposti osoittaa, ks. esim `https://proofwiki.org/wiki/Inverse_of_Vandermonde's_Matrix`

Samankaltaisella päättelyketjulla voidaan ratkaista yleinen lauseke myös tilanteelle, jossa kaikkien pisteiden mittaustarkkuus ei ole sama, jolloin pisteille halutaan käyttää erilaisia painokertoimia sovitusta tehtäessä. Tähän voidaan käyttää varianssimatriisia $\boldsymbol{V}$, jonka avulla voidaan määritellä painokerroinmatriisi $\boldsymbol{W} = \boldsymbol{V}^{-1}\sigma^2$. Tässä σ on tutkittavan aineiston varianssi ja $\boldsymbol{V}$ määritelty seuraavasti: [3]

$$\boldsymbol{V} = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} & \dots & \sigma_{1,n} \\ \sigma_{2,1} & \sigma_2^2 & \dots & \sigma_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n,1} & \sigma_{n,2} & \dots & \sigma_n^2 \end{bmatrix}$$

Näitä merkintöjä käyttäen kerroinvektori $\boldsymbol{\beta}$ voidaan ratkaista yhtälöstä 8 [3].

$$\boldsymbol{\beta} = (\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{Y} \tag{8}$$

2 Pienimmän neliösumman ratkaisu kaksiulotteiselle havaintoaineistolle

Kun pienimmän neliösumman sovitusta sovelletaan asteroidin radanmääritykseen, on käytössä kaksiulotteinen havaintoaineisto, joka koostuu asteroidin paikoista (x_i, y_i) ajanhetkillä t_i. Asteroidin paikkavektori $\boldsymbol{r}(t)$ mielivaltaisella ajanhetkellä t voidaan ratkaista asteroidin rataelementtien avulla. Rataelementeiksi voidaan valita asteroidin paikka $\boldsymbol{R}_0 = (X_0, Y_0)$ ja nopeus $\boldsymbol{V}_0 = (\dot{X}_0, \dot{Y}_0)$ ajanhetkellä $t = 0$, jolloin asteroidin paikka saadaan seuraavasti:

$$\boldsymbol{r}(t) = \boldsymbol{R}_0 + \boldsymbol{V}_0 t. \tag{9}$$

Havaintoaineiston ollessa kaksiulotteinen, täytyy $\boldsymbol{Y}$-vektorin tallentaa sekä x- että y-koordinaatit, jolloin

$$\boldsymbol{Y} = \begin{bmatrix} x_1 \\ y_2 \\ x_2 \\ y_2 \\ \vdots \\ x_n \\ y_n \end{bmatrix}.$$

Koska tahdotaan ratkaista aiemmin määritellyt rataelementit, on myös ratkaistava kerroinvektori erilainen:

$$\boldsymbol{\beta} = \begin{bmatrix} X_0 \\ Y_0 \\ \dot{X}_0 \\ \dot{Y}_0 \end{bmatrix}.$$

Myös havainnot sisältävän matriisin $\boldsymbol{X}$ täytyy mukautua kaksiulotteiseen havaintoaineistoon, jotta kertolasku $\boldsymbol{Y}$-vektorin kanssa on mahdollinen ja fysikaalisesti mielekäs. Tähän sopii alla esitetty matriisi, jolla $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$ tuottaa kooltaan $4 \times 2n$ matriisin, joka voidaan edelleen kertoa edellä määritellyllä vektorilla $\boldsymbol{Y}$, jolloin saadaan kooltaan matriisia $\boldsymbol{\beta}$ vastaava matriisi. Ajanhetki t_i on pisteen (x_i, y_i) havaintoaika.

$$\boldsymbol{X} = \begin{bmatrix} 1 & 0 & t_1 & 0 \\ 0 & 1 & 0 & t_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & t_n & 0 \\ 0 & 1 & 0 & t_n \end{bmatrix}$$

Näillä uudelleenmääritellyillä on mahdollista käyttää suoraan aiemmin ratkaistuja yhtälöitä 7 (kaikkien havaintojen painokerroin sama) tai 8 (painokertoimet). Mikäli käytetään painokertoimia, on myös painokerroinmatriisin koko luonnollisesti $2n \times 2n$. Ratkaisuna saatavasta vektorista $\boldsymbol{\beta}$ saadaan suoraan rataelementit $\boldsymbol{R_0}$ ja $\boldsymbol{V_0}$.

Virhearvio

3 Ennusteen virhearvio lineaarisessa mallissa

Yhtälön 8 mukaisesti ratkaistujen parametrien $\boldsymbol{\beta}$ virhe voidaan laskea kun tiedetään, että lineaarikombinaatiossa $\boldsymbol{P}^T\boldsymbol{Y}$, missä $\boldsymbol{P}$ on vakiovektori, saadaan varianssimatriisi seuraavasti [3]:

$$\mathrm{var}(\boldsymbol{P}^T\boldsymbol{Y}) = \boldsymbol{P}^T\mathrm{var}(\boldsymbol{Y})\boldsymbol{P}.$$

Tällöin saadaan

$$\mathrm{var}(\boldsymbol{\beta}) = (\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{W}\mathrm{var}(\boldsymbol{\beta})\boldsymbol{W}(\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}. \tag{10}$$

Hyödyntämällä aiemmin määriteltyä $\boldsymbol{W} = \boldsymbol{V}^{-1}\sigma^2$ ja tietoa $\mathrm{var}(\boldsymbol{\beta}) = \boldsymbol{V}$ eli $\mathrm{var}(\beta) = \sigma^2\boldsymbol{W}^{-1}$ saadaan yhtälö 10 muotoon

$$\mathrm{var}(\boldsymbol{\beta}) = \sigma^2(\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1} \tag{11}$$

Kun parametrien β_i varianssit σ_i tunnetaan, saadaan niistä helposti keskihajonnat s, kun tiedetään, että $\sigma = s^2$ [4].

Viitteet

[1] Least squares fitting–polynomial. http://mathworld.wolfram.com/LeastSquaresFittingPolynomial.html. Luettu: 15.10.2015.

[2] Vandermonde matrix. http://mathworld.wolfram.com/VandermondeMatrix.html. Luettu: 16.10.2015.

[3] Peter J. Bickel. *Mathematical statistics: basic ideas and selected topics.* Holden-Day, Oakland, CA, 1977.

[4] Raimo Seppänen. *MAOL-taulukot : matematiikka, fysiikka, kemia.* Otava, Helsingissä, 2005. Lisäpainokset: 2.-3:. p. 2006. - 4.-5. p. 2007. - 2.-9. p. 2011.

A Liittyvä liite.

Liian laaja leipätekstiin.