MS-A0202 Differentiaali- ja integraalilaskenta 2 (SCI) Luento 7: Pienimmän neliösumman menetelmä ja Newtonin menetelmä.

Antti Rasila

Matematiikan ja systeemianalyysin laitos
Aalto-yliopisto

Syksy 2016

Esimerkki 1

- Yritetään etsiä Lagrangen kertojien menetelmällä funktion $f(x, y) = y$ minimi ehdolla $g(x, y) = y^3 - x^2 = 0$.
- Selvästi nähdään, että minimi $f(x, y) = 0$ saavutetaan pisteessä $(0, 0)$.
- Muodostetaan Lagrangen funktio

$$L(x, y, \lambda) = y + \lambda(y^3 - x^2).$$

- Saadaan yhtälöt

$$-2\lambda x = 0, \quad 1 + 3\lambda y^2 = 0, \text{ ja } y^3 - x^2 = 0.$$

- Nämä yhtälöt ovat keskenään ristiriidassa, joten ratkaisua ei ole.
- Huomaa, että $\nabla g(0, 0) = \mathbf{0}$.
- Tästä nähdään, että ehto $\nabla g(0, 0) \neq \mathbf{0}$ on välttämätön menetelmän toimimisen kannalta.

Esimerkki 2 1/3

- Etsitään ääriarvot funktiolle $f(x, y, z) = xy + 2z$ ehdoilla $x + y + z = 0$ ja $x^2 + y^2 + z^2 = 24$.
- Koska f on jatkuva ja annettujen leikkausjoukkojen leikkaus on ympyräviiva (eli rajoitettu), niin ääriarvot ovat olemassa.
- Muodostetaan Lagrangen funktio

$$L(x, y, z, \lambda, \mu) = xy + 2z + \lambda(x + y + z) + \mu(x^2 + y^2 + z^2 - 24).$$

- Lagrangen funktion osittaisderivaatoista saadaan yhtälöt

$$y + \lambda + 2\mu x = 0, \quad x + \lambda + 2\mu y = 0, \quad 2 + \lambda + 2\mu z = 0,$$

$$z + y + z = 0, \text{ ja } x^2 + y^2 + z^2 - 24 = 0.$$

- Kahden ensimmäisen yhtälön erotus johtaa yhtälöön $(x - y)(1 - 2\mu) = 0$, joten joko $\mu = 1/2$ tai $x = y$. Tutkitaan molemmat tapaukset.

Esimerkki 2 2/3

- **Tapaus I** ($\mu = 1/2$): Toisen ja kolmannen yhtälön perusteella

$$x + \lambda + y = 0 \text{ ja } 2 + \lambda + z = 0, \text{ siis } x + y = 2 + z.$$

- Neljännestä yhtälöstä saadaan $z = -1$ ja $x + y = 1$. Viimeisen yhtälön perusteella $x^2 + y^2 = 24 - z^2 = 23$.
- Koska $x^2 + y^2 + 2xy = (x + y)^2 = 1$, saadaan $2xy = 1 - 23 = -22$ ja $xy = -11$.
- Nyt $(x - y)^2 = x^2 + y^2 - 2xy = 23 + 22 = 45$, joten $x - y = \pm 3\sqrt{5}$.
- Yhdessä yhtälön $x + y = 1$ tästä saadaan kaksi kriittistä pistettä

$$P_1 = \big((1+3\sqrt{5})/2, (1-3\sqrt{5})/2, -1\big) \text{ ja } P_2 = \big((1-3\sqrt{5})/2, (1+3\sqrt{5})/2, -1\big).$$

- Kummassakin pisteessä $f(x, y, z) = -11 - 2 = -13$.

Esimerkki 2 3/3

- **Tapaus II** ($x = y$): Neljännestä yhtälöstä nähdään, että $z = -2x$, ja viimeisen yhtälön perusteella $6x^2 = 24$ eli $x = \pm 2$.
- Näin ollen, kriittiset pisteet ovat

$$P_3 = (2, 2, -4) \text{ ja } P_4 = (-2, -2, 4).$$

- Saadaan

$$f(2, 2, -4) = 4 - 8 = -4 \text{ ja } f(-2, -2, 4) = 4 + 8 = 12.$$

- Siten funktion f maksimi on 12 ja minimi -13.

Regressio-ongelma

- Regressioanalyysissa pyritään valitsemaan regressiomallin parametrin β arvo siten, että kaikista virhetermeistä ε_j tulee samanaikaisesti mahdollisimman pieniä.
- Pyritään siis valitsemaan parametri β siten, että käyrä

$$y = f(x; \beta)$$

kulkisi mahdollisimman läheltä jokaista havaintopistettä

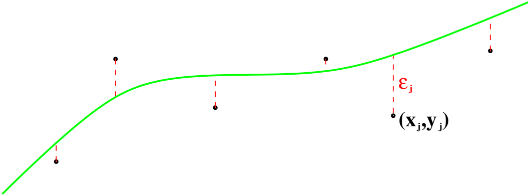
$$(x_j, y_j) \in \mathbb{R}^2, \; j = 1, 2, \ldots, n.$$

- Erään ratkaisun tähän käyränsovitusongelmaan tarjoaa pienimmän neliösumman menetelmä.

Pienimmän neliösumman menetelmä

- Pienimmän neliösumman menetelmässä pyritään minimoimaan regressiomallin
$$y_j = f(x_j; \beta) + \varepsilon_j\,,\; j = 1, 2, \ldots, n$$
virhetermien ε_j neliöden summaa muuttamalla parametrivektorin $\beta = (\beta_0, \beta_1, \ldots, \beta_m)$ arvoa eli funktiota
$$F(\beta) = \sum_{j=1}^{n} \varepsilon_j^2 = \sum_{j=1}^{n} \big(y_j - f(x_j; \beta)\big)^2.$$
- Optimaalinen β:n arvo on parametrin β *pienimmän neliösumman estimaatti* eli *PNS-estimaatti*.

PNS-sovitus



PNS-estimaatin sovittaminen: Kuvassa vihreällä parametreista β_j riippuva sovitettava funktio $f(x;\beta)$ jollakin parametrin arvolla, datapisteet (x_j, y_j) ja vastaavat virhetermit ε_j.

Minimin löytäminen lineaarisessa tapauksessa 1/2

- Tarkastellaan funktiota

$$F(\beta_0, \beta_1) = \sum_i (y_i - \beta_0 - \beta_1 x_i)^2.$$

- Etsitään piste (β_0, β_1) siten, että

$$\nabla F(\beta_0, \beta_1) = 0.$$

- Lasketaan osittaisderivaatta

$$\frac{\partial}{\partial \beta_0} F(\beta_0, \beta_1) = 2\big(\beta_1 \sum x_i + n\beta_0 - \sum y_i\big).$$

- Ratkaistaan nollakohta

$$\beta_0 = \frac{1}{n} \sum y_i - \frac{\beta_1}{n} \sum x_i = \bar{y} - \beta_1 \bar{x}.$$

Tässä $\bar{x}$ on vektorin $\mathbf{x}$ komponenttien aritmeettinen keskiarvo

Minimin löytäminen lineaarisessa tapauksessa 2/2

- Lasketaan seuraavaksi osittaisderivaatta

$$\frac{\partial}{\partial \beta_1} F(\beta_0, \beta_1) = 2\big(\beta_0 \sum x_i + \beta_1 \sum x_i^2 - \sum x_i y_i\big).$$

- Sijoittamalla β_0:n lauseke, saadaan

$$n\bar{x}\bar{y} - n\beta_1 \bar{x}^2 + \beta_1 \sum x_i^2 - \sum x_i y_i = 0.$$

- Ratkaistaan nollakohta

$$\beta_1 = \frac{n\bar{x}\bar{y} - \sum x_i y_i}{n\bar{x}^2 - \sum x_i^2} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}.$$

□

Esimerkki 3

- Sovitetaan PNS-suora dataan ja estimoidaan y, kun $x = 5$:

x_i	0.0	1.0	2.0	3.0	4.0
y_i	2.10	1.92	1.84	1.71	1.64

- Saadaan $\bar{x} = 2.0$, $\bar{y} = 1.842$, ja

$$\beta_1 = \frac{-1.13}{10.0} = -0.113.$$

- Siten $\beta_0 = 1.842 + 0.113 \cdot 2.0 = 2.068$.
- Näin ollen, $y = -0.113x + 2.068$, ja haluttu estimaatti pisteessä $x = 5$ on $y = -0.113 \cdot 5 + 2.068 = 1.503$.

Esimerkki 4: Toisen asteen sovitus 1/2

- Tutkitaan lisäaineen määrän x vaikutusta kuivumisaikaan y. Eri lisäaineen määrillä x_i (grammaa) saatiin kuivumisajat y_i (tuntia), $i = 1, \ldots, 9$:

x_i	0.0	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0
y_i	11.0	9.4	9.1	7.0	6.2	7.1	6.6	7.5	8.2

- Huomataan, että kuivumisajan riippuvuus lisäaineen määrästä on epälineaarista. Minimikohdan estimoimiseksi sovitetaan havaintoihin paraabeli $y = \beta_0 + \beta_1 x + \beta_2 x^2$.
- Pienimmän neliösumman yhtälöryhmä mallille on

$$\frac{\partial}{\partial \beta_k} \sum (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2 = 0, \qquad k = 0, 1, 2.$$

Esimerkki 4: Toisen asteen sovitus 2/2

- Näistä saadaan yhtälöryhmä

$$\begin{cases} n\beta_0 + \beta_1 \sum x_i + \beta_2 \sum x_i^2 & = & \sum y_i, \\ \beta_0 \sum x_i + \beta_1 \sum x_i^2 + \beta_2 \sum x_i^3 & = & \sum x_i y_i \\ \beta_0 \sum x_i^2 + \beta_1 \sum x_i^3 + \beta_2 \sum x_i^4 & = & \sum x_i^2 y_i. \end{cases}$$

- Laskemalla yhtälöryhmän kertoimet havainnoista saadaan

$$\begin{cases} 9\beta_0 + 36\beta_1 + 204\beta_2 & = & 72.1 \\ 36\beta_0 + 204\beta_1 + 1296\beta_2 & = & 266.6 \\ 204\beta_0 + 1296\beta_1 + 8772\beta_2 & = & 1515.4 \end{cases}$$

- Ratkaisuna ovat $\beta_0 = 11.15$, $\beta_1 = -1.806$ ja $\beta_2 = 0.1803$. Pienimmän neliösumman mielessä parhaiten havaintoihin liittyvä paraabeli on siis

$$y = 11.15 - 1.806x + 0.1803x^2.$$

Newtonin menetelmä

- Newtonin menetelmä on yksi käytetyimmistä numeerisen analyysin menetelmistä. Menetelmän avulla voidaan löytää funktion $f\colon \mathbb{R} \to \mathbb{R}$ nollakohta eli yhtälön $f(x) = 0$ ratkaisu.
- Silloin kun menetelmä toimii, se yleensä suppenee hyvin nopeasti.
- Menetelmän idea on seuraava: Lähdetään liikkeelle jostakin pisteestä x_0, joka on alkuarvaus yhtälön ratkaisulle.
- Arvioidaan funktiota f sen tangenttisuoralla pisteessä, eli funktiolla $l(x) = f(x_0) + f'(x_0)(x - x_0)$.
- Ratkaistaan yhtälö $l(x) = 0$. Tämän yhtälön ratkaisu on seuraava alkuarvaus x_1. Toistetaan edellinen käyttäen alkuarvauksena lukua x_1.
- Tämä menettely johtaa algoritmiin, jossa iteraatioaskeleet saadaan kaavasta
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \qquad n = 0, 1, 2, \ldots.$$
- Menetelmä on toimiessaan nopea, mutta suppeneminen riippuu alkuarvauksesta.

Esmerkki 5

- Etsitään likiarvo luvulle $\sqrt{5}$.
- Selvästi $2 = \sqrt{4} < \sqrt{5} < \sqrt{9} = 3$. Valitaan $x_0 = 2$.
- Tässä $f(x) = x^2 - 5$, joten $f'(x) = 2x$. Saadaan

$$x_1 = x_0 - \frac{f(2)}{f'(2)} = 2 - \frac{4-5}{2 \cdot 2} = \frac{9}{4}.$$

$$x_2 = x_1 - \frac{f(9/4)}{f'(9/4)} = \frac{9}{4} - \frac{27/16 - 5}{2 \cdot 9/4} = \frac{161}{72} \approx 2.2361.$$

- Huomaa, että $\sqrt{5} \approx 2.236068$, eli jo kahdella iteraatiolla saatiin varsin hyvä likiarvo.

Esimerkki 6

- Etsitään funktion $f(x) = x^3 - x + 1$ nollakohdat.
- Piirtämällä kuvaaja voidaan havaita, että funktiolla on vain yksi nollakohta jossain pisteiden -2 ja -1 välissä. Asetetaan $x_0 = -1$.
- Koska $f'(x) = 3x^2 - 1$ iteratioksi saadaan

$$x_{n+1} = x_n - \frac{x_n^3 - x_n + 1}{3x_n^2 - 1}.$$

- Saadaan

$$x_0 = -1, \quad x_1 = -1.5, \quad x_2 = -1.347826, \quad x_3 = -1.325200.$$

$$x_4 = -1.324718, \quad x_5 = -1.324717, \quad \ldots$$

Newtonin menetelmä monen muuttujan tapauksessa

- Newtonin menetelmää voidaan käyttää myös funktion $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^n$ tapauksessa.
- Tällöin iteraatioaskeleeksi saadaan

$$\mathbf{x}_{n+1} = \mathbf{x}_n - J_{\mathbf{f}}(\mathbf{x}_n)^{-1}\mathbf{f}(\mathbf{x}_n), \quad n = 0, 1, 2, \ldots .$$

- Tässä $J_{\mathbf{f}}(\mathbf{x}_n)^{-1}$ tarkoittaa funktion $\mathbf{f}$ Jacobin matriisin käänteismatriisia.

Esimerkki 7

- Etsitään $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, kun $\mathbf{x}_0 = (1, 0, 1)$ ja

  $$\mathbf{f}(x, y, z) = (x^2 + y^2 + z^2 - 3)\mathbf{i} + (x^2 + y^2 - z - 1)\mathbf{j} + (x + y + z - 3)\mathbf{k}.$$

- Saadaan

  $$J_{\mathbf{f}}(x, y, z) = \begin{bmatrix} 2x & 2y & 2z \\ 2x & 2y & -1 \\ 1 & 1 & 1 \end{bmatrix}.$$

- Voidaan laskea

  $$\mathbf{x}_1 = (3/2, 1/2, 1), \quad \mathbf{x}_2 = (5/4, 3/4, 1) \text{ ja } \mathbf{x}_2 = (9/8, 7/8, 1).$$

- Nähdään, että iteraatiot konvergoivat kohti pistettä $(1, 1, 1)$, joka on tehtävän tarkka ratkaisu.