Analyysi II

Jari Taskinen

21. elokuuta 2002

Luku 6

Sisältö

6. Ääriarvojen teoriaa

Ääriarvojen teoriaa

Tutkitaan aluksi *neliömuotoja* kahden muuttujan x, y tapauksessa:

$$P(x,y) = ax^2 + 2bxy + cx^2, \tag{1}$$

missä $a, b, c \in \mathbf{R}$ annettuja kertoimia.

Esimerkki Määritellään

$$P(x,y) = 2x^2 - 10xy + 3y^2.$$

Nähdään että $P(0,0) = 0$

1. Jos $P(x,y) \neq 0$ aina, kun $(x,y) \neq (0,0)$, niin P on *definiitti*.
   a) Jos $P(x,y) > 0$ aina, kun $(x,y) = (0,0)$, niin P on *positiivisesti definiitti*.
   b) Jos $P(x,y) < 0$ aina, kun $(x,y) = (0,0)$, niin P on *negatiivisesti definiitti*.
2. Jos $P(x,y) \geq 0 \; \forall \; (x,y) \in \mathbf{R}^2$ tai jos $P(x,y) \leq 0 \; \forall \; (x,y) \in \mathbf{R}^2$, niin P on *semidefiniitti*. (Huomautus! Jos P on definiitti, on se myös semidefiniitti.)
3. Jos P saa positiivisia ja negatiivisia arvoja, se on *indefiniitti*.

Yleisemmin, neliömuoto $\mathbf{R}^n$:ssä on muotoa

$$P(\overline{x}) = \overline{x}^T A \overline{x} = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} 2 a_{ij} x_i x_j,$$

Sisältö:

missä A on $n \times n$ matriisi $(a_{ij})_{i,j=1}^{n}$.

Lause 6.1 Neliömuoto (1) on

Definiitti, jos $ac - b^2 > 0$,

semidefiniitti, jos $ac - b^2 = 0$,

indefiniitti, jos $ac - b^2 < 0$.

Merkitään $D := ac - b^2$.

Todistus. Oletetaan, että $D > 0$. Silloin $a \neq 0$ ja $P(x, y)$ voidaan kirjoittaa muodossa

$$P(x, y) = \frac{1}{a}[(ax + by)^2 + Dy^2].$$

Jos $P(x, y) = 0$, niin

$$0 = aP(x, y) = (ax + by)^2 + Dy^2.$$

Tämä on mahdollista vain jos $Dy^2 = 0$ ja $(ax + by)^2 = 0$ eli $y = 0$ ja $(ax + 0)^2 = 0$ eli $y = 0$ ja $x = 0$ (koska $a \neq 0$).

Oletetaan nyt $D < 0$ ja $a \neq 0$. Tällöin

$$P(x, y) = \frac{1}{a}\Big((ax+by)^2 + Dy^2\Big) = \frac{1}{a}\Big((ax+by) - \sqrt{|D|}y\Big)\Big((ax+by) + \sqrt{|D|}y\Big).$$

Nähdään, että P häviää xy-tason suorilla

$$ax + (b - \sqrt{|D|})y = 0 \text{ ja } ax + (b + \sqrt{|D|})y = 0.$$

Siis P on indefiniitti.

Tapaus $a = 0$, $D < 0$. Tällöin

$$P(x, y) = y(2bx + cy).$$

Koska $D = ac - b^2 < 0$ ja $a = 0$, niin täytyy olla $b \neq 0$. Otetaan $y = 1$:

$$P(x, 1) = 2bx + c,$$

joka saa negatiivisia ja positiivisia arvoja kun $x \in \mathbf{R}$. P on siis indefiniitti.

Tapaus $D = 0$.

1. Jos $a \neq 0$, niin $P(x, y) = \frac{1}{a}(ax + by)^2$,
2. Jos $a = 0$, niin $P(x, y) = cy^2 \Rightarrow b = 0$.

Selvästikin P on semidefiniitti. $\square$

Esimerkki 1 Määritellään

$$P(x, y) = x^2 - 6xy + 2y^2.$$

Nyt $a = 1$, $b = -3$, $c = 2$ ja $D = 2 - 9 = -7$, eli P on indefiniitti. $P(x, y)$ häviää suorilla

$$y = \left(\frac{1}{3 \pm \sqrt{7}}\right) x.$$

Esimerkiksi

$$\begin{aligned} P(1,1) &= 1-6+2=-3<0 \\ P(1,-1) &= 1+6+2=9>0. \end{aligned}$$

Esimerkki 2 Määritellään

$$P(x,y)=4xy-4x^2-y^2.$$

Nyt $a=-4$, $b=2$, $c=-1$ ja $D=4-4=0$ (semidefiniitti). Itseasiassa $P(x,y)=-(2x-y)^2$, mistä nähdään

(i) $P(x,y)\leq 0\ \forall\ (x,y)\in \mathbf{R}^2$

(ii) $P(50,100)=-(100-100)^2=0.$

Esimerkki 3 Määritellään

$$P(x,y)=2xy-3x^2-y^2.$$

Nyt $a=-3$, $b=1$, $c=-1$ ja $D=2>0$ (definiitti). Esimerkiksi $P(1,1)=-2<0$. Siis $P(x,y)<0$, kun $(x,y)\neq\bar{0}$.

6.1. Taylorin kaava

Taylorin kaava

Olkoon $A \subset\circ \mathbf{R}^2$, $f : A \to \mathbf{R}$ funktio, jolla on kaikkien kertalukujen jatkuvat osittaisderivaatat. Olkoon $\overline{x} \in A$ tarkastelupiste ja $\overline{h} \in \mathbf{R}^2$ siten, että jana

$$J := \{\overline{x} + t\overline{h} \mid t \in [0,1]\} \subset A.$$

J on jana, jonka päätepisteet ovat $\overline{x}$ ja $\overline{x} + \overline{h}$. Määritellään

$$\begin{aligned} g &: [0,1] \to A \\ g(t) &= \overline{x} + t\overline{h} \\ g(0) &= \overline{x} \\ g(1) &= \overline{x} + \overline{h}. \end{aligned}$$

Tarkastellaan funktiota

$$f \circ g : [0,1] \to \mathbf{R},$$

joka on mielivaltaisen monta kertaa derivoituva yhden muuttujan funktio.

Ketjusäännön mukaan

$$(f \circ g)' = (D_1 f) \circ g \cdot g_1' + (D_2 f) \circ g_2' = h_1 (D_1 f) \circ g + h_2 (D_2 h) \circ g.$$

Huomautus

$$g(t) = \big(g_1(t), g_2(t)\big) = (x_1 + th_1, x_2 + th_2),$$

missä $\overline{x} = (x_1, x_2)$, $\overline{h} = (h_1, h_2)$ ja $g_1' = h_1$, $g_2' = h_2$. Edelleen

$$\begin{aligned} (f \circ g)'' &= h_1 \tfrac{d}{dt}(D_1 f \circ g) + h_2 \tfrac{d}{dt}(D_2 f \circ g) \\ &= h_1 \Big(h_1 (D_{11} f) \circ g + h_2 (D_{12} f) \circ g\Big) + h_2 \Big(h_1 (D_{21} f) \circ g + h_2 (D_{22} f) \circ g\Big) \\ &= h_1^2 (D_{11} f) \circ g + 2h_1 h_2 (D_{12} f) \circ g + h_2^2 (D_{22} f) \circ g \\ &= \Big((h_1 D_1 + h_2 D_2)^2 f\Big) \circ g, \end{aligned}$$

missä $D_1^2 = D_{11}$, $D_1D_2 = D_{12} = D_2D_1$ ja $D_2^2 = D_{22}$. Induktiolla voidaan todistaa kaava

$$(f \circ g)^{(k)} = \Big((n_1D_1 + h_2D_2)^k f\Big) \circ g.$$

Taylorin kaavasta yhden muuttujan funktiolle $F : B \to \mathbf{R}$, missä $[0,1] \subset B \subset\circ\, \mathbf{R}$, saadaan

$$F(1) = \sum_{k=0}^{n-1} \frac{1}{k!} F^{(k)}(0) + \frac{1}{n!} F^{(n)}(\theta), \tag{2}$$

missä $\theta \in \,]0,1[$ (jos $n = 1$, tämä on väliarvolause).

Jos F toteuttaa tietyt (ankarat) vaatimukset, sille pätee kehitelmä

$$F(x) = \sum_{k=0}^{\infty} \frac{1}{k!} F^{(k)}(0) x^k,$$

missä $x \in B(0,r)$.

Taylorin kehitelmä (2) pätee funktiolle F, jos se on n kertaa jatkuvasti derivoituva. Sovelletaan tätä, kun $F := f \circ g$:

$$(f \circ g)(1) = \sum_{k=0}^{n-1} \frac{(f \circ g)^{(k)}(0)}{k!} + \frac{(f \circ g)^{(n)}(\theta)}{n!}.$$

Sijoittamalla k:nen derivaatan lauseke ja ottamalla huomioon

$$g(0) = \overline{x}, \quad g(1) = \overline{x} + \overline{h}, \quad g(\theta) = \overline{x} + \theta\overline{h},$$

saadaan

$$\begin{aligned} f(\overline{x}+\overline{h}) &= f(g(1)) = f\circ g(1) \\ &= \sum_{k=0}^{n-1}\tfrac{1}{k!}(h_1D_1+h_2D_2)^k f(\overline{x}) + \tfrac{1}{n!}(h_1D_1+h_2D_2)^n f(\overline{x}+\theta\overline{h}), \end{aligned}$$

missä $\theta \in\,]0,1[$. Tämä kehitelmä pätee, kun f on n kertaa jatkuvasti derivoituva. Arvolla $n=2$ kehitelmästä seuraa

Lause 6.2 Olkoon $f : A \to \mathbf{R}$, $A \subset\circ\, \mathbf{R}^2$, kaksi kertaa jatkuvasti derivoituva. Olkoon $\overline{x}\in A$, $\overline{h}\in\mathbf{R}^2$ ja $\overline{h}\in B(\overline{0},r)$, missä r on niin pieni, että $\overline{x}+B(\overline{0},r)\subset A$. Tällöin

$$\begin{aligned} f(\overline{x}+\overline{h}) - f(\overline{x}) &= h_1D_1f(\overline{x}) + h_2D_2f(\overline{x}) \\ &\quad+\tfrac{1}{2}\Big(h_1^2D_{11}f(\overline{x}) + 2h_1h_2D_{12}f(\overline{x}) + h_2^2D_{22}f(\overline{x})\Big) \\ &\quad+\left|\,\overline{h}\,\right|^2\varepsilon(h), \end{aligned}$$

missä $\varepsilon(\overline{h})\to 0$, kun $\overline{h}\to\overline{0}$.

6.2. Ääriarvoista

Ääriarvoista

Olkoon $f : A \to \mathbf{R}$, $A \subset\circ\, \mathbf{R}^m$, $m\in\mathbf{N}$. Tällöin

- funktiolla f on (lokaali)maksimi pisteessä $a\in A$, jos on olemassa sellainen $r>0$, että $f(\overline{x})\le f(\overline{a})$, kun $\overline{x}\in B(\overline{a},r)$

- funktiolla f on (lokaali)minimi pisteessä $a \in A$, jos on olemassa sellainen $r > 0$, että $f(\overline{x}) \geq f(\overline{a})$, kun $\overline{x} \in B(\overline{a}, r)$
- ääriarvo on minimi tai maksimi
- ääriarvopiste on lähtöjoukon piste, jossa ääriarvo saavutetaan
- ääriarvo on oleellinen, jos $f(\overline{x}) \neq f(\overline{a})$, kun $\overline{x} \in B(\overline{a}, r) \setminus \{\overline{a}\}$ (tässä r kuten maksimin tai minimin määritrelmässä).

Lause 6.3 Olkoon $f : A \to \mathbf{R}$, $A \subset\circ \mathbf{R}^m$, kerran derivoituva. Jos $\overline{a} \in A$ on f:n ääriarvopiste, niin $D_i f(\overline{a}) = 0$, kun $i = 1, \ldots, m$.

Esimerkki 1 Määritellään

$$f(x, y) = x^2 + y^2.$$

Nyt

$$\begin{aligned} D_1 f(x, y) &= 2x \\ D_2 f(x, y) &= 2y \\ D_j f(0, 0) &= 0, \quad \text{kun } j = 1, 2 \end{aligned}$$

Tunnetusti f:llä on lokaali minimi pisteessä $\overline{0}$.

Esimerkki 2 Määritellään

$$f(x, y) = x^2 - y^2.$$

Nyt pätee $D_1 f(0, 0) = 0 = D_2 f(0, 0)$. Toisaalta $(0, 0)$ ei ole ääriarvopiste:

$$\begin{aligned} f(0, r) &= -r^2 < 0 \\ f(r, 0) &= r^2 > 0 \end{aligned}$$

Lauseen 6.3 (sivu 9) todistus:

Todistus. Tehdään antiteesi: Oletetaan että $\overline{a} \in A$ on A:n ääripiste ja $D_k f(\overline{a}) \neq 0$, jollekin $k \in \{1, \ldots, m\}$. Määritellään funktio

$$g(x) = f(a_1, \ldots, a_{k-1}, x, a_{k+1}, \ldots, a_m) \tag{3}$$

missä $x \in B(a_k, r) \subset \mathbf{R}$ jollakin $r > 0$.

Nyt

$$\frac{dg}{dx}(a_k) = D_k f(\overline{a}) \neq 0.$$

Näin ollen g saa a_k:n ympäristössä sekä suurempia että pienempiä arvoja kuin $g(a_k)$. Kohdan 3 nojalla sama pätee funktiolle f, joten $\overline{a}$ ei ole funktion f ääriarvopiste. □

Lause 6.4 Olkoon $A \subset\!\circ\, \mathbf{R}^2$, $f : A \to \mathbf{R}$. Oletetaan että funktiolla f on jatkuvat kertaluvun 1. ja 2. osittaisderivaatat. Jos pisteessä $\overline{a} \in A$ pätee

$$D_1 f(\overline{a}) = 0 = D_2 f(\overline{a})$$

ja

$$\mathcal{D} := D_{11} f(\overline{a}) D_{22} f(\overline{a}) - D_{12}(\overline{a})^2 > 0, \tag{4}$$

niin $\overline{a}$ on oleellinen ääriarvopiste;

- maksimi, jos $D_{11} < 0$
- minimi, jos $D_{11} > 0$.

Määritelmä Satulapiste. Olkoon $A, \overline{a}, f$ kuten lauseen 6.4 oletuksessa.

$$D_1 f(\overline{a}) = D_2 f(\overline{a}) = 0.$$

Jos f saa $\overline{a}$:n mielivaltaisessa ympäristössä sekä suurempia että pienempiä arvoja kuin $f(\overline{a})$, niin $\overline{a}$ on *satulapiste*.

Esimerkki 1 Määritellään

$$f(x, y) = x^2 - y^2.$$

Derivaatat ovat

$$\begin{array}{rclrcl} D_1 & = & 2x & D_2 & = & -2y \\ D_{11} & = & 2 & D_{22} & = & -2 \\ D_{12} & = & 0 & & & \end{array}$$

joten

$$\begin{array}{lcccc} D_1 f(x, y) = 0 & \iff & 2x = 0 & \iff & x = 0 \\ D_2 f(x, y) = 0 & \iff & -2y = 0 & \iff & y = 0. \end{array}$$

Pisteessä $\overline{a} = (0, 0)$

$$D_{11} f(\overline{a}) D_{22} f(\overline{a}) - D_{12} f(\overline{a})^2 < 0.$$

Esimerkki 2 Määritellään

$$f(x, y) = x^2 + xy + y^2 + x - y.$$

Tehtävänä on etsiä ääriarvo- ja satulapisteet.

Ratkaisu. Osittaisderivaatat

$$\begin{array}{rcl} D_1 f & = & 2x + y + 1 \\ D_2 f & = & x + 2y - 1. \end{array}$$

Jos (x, y) on kriittinen piste (eli derivaattojen nollakohta), niin

$$\begin{cases} 2x & + & y & + & 1 & = & 0 \\ x & + & 2y & - & 1 & = & 0 \end{cases} \Rightarrow \begin{cases} y & = & 1 \\ x & = & -1. \end{cases}$$

Pisteessä $(-1, 1)$ pätee

$$\left.\begin{array}{rcl} D_{11} & = & 2 \\ D_{22} & = & 2 \\ D_{12} & = & 1 \end{array}\right\} \Rightarrow \mathcal{D}(-1, 1) > 0,$$

eli kyseessä on ääriarvopiste (minimi).

Esimerkki 3 Määritellään

$$f(x, y) = x^3 - y^3 + 3xy.$$

Derivaatat ovat

$$\begin{array}{rcl} D_1 f & = & 3x^2 + 3y \\ D_2 f & = & -3y^2 + 3x. \end{array}$$

Kriittiset pisteet ovat

$$\begin{cases} 3x^2 & + & 3y & = & 0 \\ -3y^2 & + & 3x & = & 0 \end{cases} \iff \begin{cases} x & = & x^4 \\ y & = & -x^2 \end{cases}$$
$$\iff (x, y) = (0, 0) \text{ tai } (x, y) = (1, -1).$$

Edelleen

$$D_{11} f = 6x, \quad D_{22} f = -6y, \quad D_{12} f = 3,$$

joten

$$\begin{aligned}\mathcal{D}(0,0) &= -9<0 \quad \text{(satulapiste eli ei ääriarvopiste)}\\ \mathcal{D}(-1,1) &= 6\cdot 1\cdot(-6)\cdot(-1)-9 = 36-9>0 \quad \text{(minimi.)}\end{aligned}$$

Lauseen 6.4 (sivu 10) todistuksesta:

Olkoon $\overline{h}\in B(0,r)$, $r>0$. Lauseen 6.2 (sivu 8) mukaan pätee

$$\begin{aligned}f(\overline{a}+\overline{h})-f(\overline{a}) &= D_1 f(\overline{a})h_1 + D_2 f(\overline{a})h_2\\ &+\tfrac{1}{2}\Big(D_{11}f(\overline{a})h_1^2 + 2D_{12}f(\overline{a})h_1h_2\\ &+D_{22}f(\overline{a})h_2^2\Big) + \left|\,\overline{h}\,\right|\varepsilon(\overline{h}),\end{aligned}$$

missä $\varepsilon(\overline{h})\to 0$, kun $\overline{h}\to 0$. Koska $\overline{h}$ on pieni ja $D_1 f(\overline{a}) = D_2 f(\overline{a}) = 0$, määrää lauseke

$$D_{11}f(\overline{a})h_1^2 + 2D_{12}f(\overline{a})h_1h_2 + D_{22}f(\overline{a})h_2^2 \tag{5}$$

erotuksen $f(\overline{a}+\overline{h})-f(\overline{a})$ etumerkin. Lauseke (5) on neliömuoto $P(h_1,h_2)$, kertoimet $a=D_{11}f(\overline{a})$, $b=D_{12}f(\overline{a})$, $c=D_{22}f(\overline{a})$. Ylläolevista määritelmistä seuraa, että $\overline{a}$ on ääriarvopiste jos ja vain jos neliömuoto P on definiitti. Huomaa, että

$$D>0 \iff \mathcal{D}(\overline{a})>0.$$

Yllä olevat tarkastelut olivat lokaaleja.

Esimerkki Funktion suurin tai pienin arvo joukossa $B\subset\mathbf{R}^2$. Olkoon B kompakti (eli rajoitettu ja suljettu). Nyt pätee: Jos $f:B\to\mathbf{R}$ on jatkuva, niin f saa B:ssä suurimman ja pienimmän arvon. Esimerkiksi jos $C := [0, \infty[$, niin funktio $f(x) = \frac{1}{x}$ ei saa joukossa C pienintä arvoaan. Mutta C ei olekaan kompakti joukko.

Huomautus Se, että funktio $f : B \to \mathbf{R}$ saa pienimmän arvonsa pisteessä $\overline{x}_0 \in B$, tarkoittaa, että

$$f(\overline{y}) \geq f(\overline{x}_0) \; \forall \; y \in B.$$

Olkoon B kompakti, $f : B \to \mathbf{R}$ jatkuva ja kaksi kertaa jatkuvasti derivoituva. Ainoat pisteet, joissa f voi saada suurimman tai pienimmän arvonsa ovat

1. f:n lokaalit ääriarvopisteet
2. B:n reunapisteet.

Yllä esitettyä lokaalien ääriarvojen tarkastelua voidaan siten käyttää hyväksi myös globaalien ääriarvojen eli suurimman ja pienimmän arvon löytämiseksi. Lopuksi, ilman todistuksia esitetään n:n muuttujan funktioiden ääriarvojen teoriaa.

Lause 6.5 Olkoon $n \in \mathbf{N}$, $n \geq 2$, $A \subset\!\circ\, \mathbf{R}^n$, $f : A \to \mathbf{R}$ derivoituva. Jos funktiolla f on ääriarvo pisteessä $\overline{a} \in A$, niin

$$\nabla f(\overline{a}) = 0 \quad \text{eli} \quad D_1 f(\overline{a}) = D_2 f(\overline{a}) = \ldots = D_n f(\overline{a}) = 0.$$

Todistus. Sama kuin $n = 2$. $\square$

Lause 6.6 Olkoon $A \subset\!\circ\, \mathbf{R}^n$, f kaksi kertaa jatkuvasti derivoituva ja

$$\nabla f(\overline{a}) = 0 \text{ pisteessä } \overline{a} \in A.$$

Muodostetaan funktion f Hessen matriisi

$$H = \begin{pmatrix} D_{11}f(\overline{a}) & D_{12}f(\overline{a}) & \dots & D_{1n}f(\overline{a}) \\ D_{21}f(\overline{a}) & D_{22}f(\overline{a}) & \dots & D_{2n}f(\overline{a}) \\ \vdots & & \ddots & \\ D_{n1}f(\overline{a}) & D_{n2}f(\overline{a}) & \dots & D_{nn}f(\overline{a}) \end{pmatrix}$$

ja diagonalisoidaan se; saadaan muotoa

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & \lambda_n \end{pmatrix}$$

oleva matriisi. Seuraavat tulokset pätevät:

a) Jos kaikki ominaisarvot ovat suurempia kuin nolla, on $\overline{a}$ minimi.

b) Jos kaikki ominaisarvot ovat pienempiä kuin nolla, on $\overline{a}$ maksimi

c) Jos matriisilla on sekä positivisia että negatiivisia ominaisarvoja, niin $\overline{a}$ ei ole ääriarvopiste.