MS-A0202 Differentiaali- ja integraalilaskenta 2 (SCI)
Luento 6: Ääriarvojen luokittelu. Lagrangen kertojat.

Antti Rasila

Matematiikan ja systeemianalyysin laitos
Aalto-yliopisto

Syksy 2016

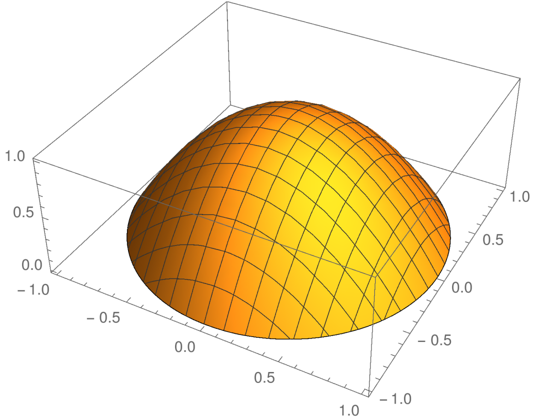
Kertausta: ääriarvot yhden muuttujan tapauksessa

- Funktiolla $f\colon I \subset \mathbb{R} \to \mathbb{R}$ on lokaali (paikallinen) maksimi pisteessä $a \in I$, jos $f(x) \leq f(a)$ kaikilla x:n arvoilla jossakin a:n ympäristössä (eli riittävän lähellä pistettä a).
- Vastaavasti lokaali minimi tarkoittaa sitä, että $f(x) \geq f(a)$ jossakin a:n ympäristössä.
- Maksimi tai minimi on globaali, jos kyseinen epäyhtälö on voimassa kaikilla $x \in I$.
- Ääriarvoja voi esiintyä: (i) Funktion f kriittisissä pisteissä $f'(x) = 0$, (ii) pisteissä joissa f:n derivaatta ei ole määritelty, ja (iii) määrittelyjoukon I reunalla.
- Seuraavaksi yleistetään näitä ehtoja funktion $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$ tapaukseen.

Ääriarvot ja usean muuttujan funktiot

- Funktiolla $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$ on pisteessä $\mathbf{x}_0 \in D$ lokaali maksimi, jos $f(\mathbf{x}) \leq f(\mathbf{x}_0)$ jossakin pisteen $\mathbf{x}_0$ ympäristössä.
- Vastaavasti $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$ on pisteessä $\mathbf{x}_0 \in D$ lokaali minimi, jos $f(\mathbf{x}) \geq f(\mathbf{x}_0)$ pisteen $\mathbf{x}_0$ ympäristössä.
- Ääriarvo on globaali eli absoluuttinen, jos kyseinen epäyhtälö on voimassa kaikilla $\mathbf{x} \in D$.
- Ääriarvoja voi esiintyä:
  1. Funktion f kriittissä pisteissä eli gradientin nollakohdissa $\nabla f(\mathbf{x}) = 0$,
  2. pisteissä joissa ∇f ei ole määritelty, sekä
  3. määrittelyjoukon D reunalla.
- Joukon D kriittistä pistettä $\mathbf{x}_0$, joka ei ole maksimi tai minimi, kutsutaan funktion $f\colon D \to \mathbb{R}$ satulapisteeksi.

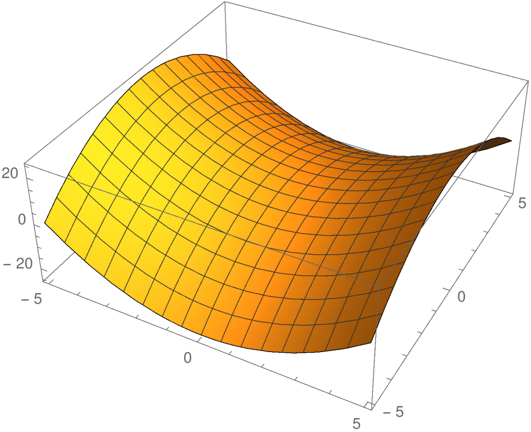
Esimerkki 1



Funktiolla $f(x, y) = 1 - x^2 - y^2$ on lokaali maksimi $f(0, 0) = 1$ pisteessä $(0, 0)$. Tämä piste on funktion f kriittinen piste, koska

$$\nabla f(0, 0) = -2x\mathbf{i} - 2y\mathbf{j}\Big|_{(0,0)} = \mathbf{0}.$$

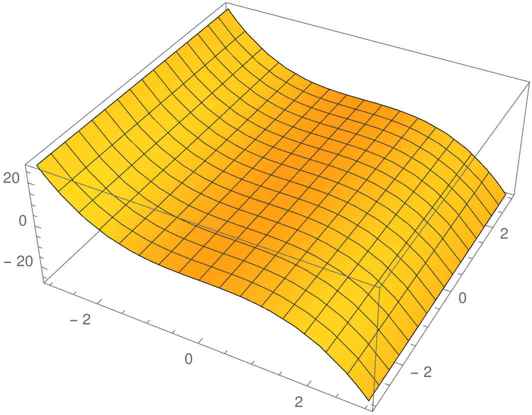
Esimerkki 2



Funktiolla $f(x, y) = y^2 - x^2$ on satulapiste $(0, 0)$. Tämä piste on funktion f kriittinen piste, koska

$$\nabla f(0,0) = -2x\mathbf{i} + 2y\mathbf{j}\Big|_{(0,0)} = \mathbf{0}.$$

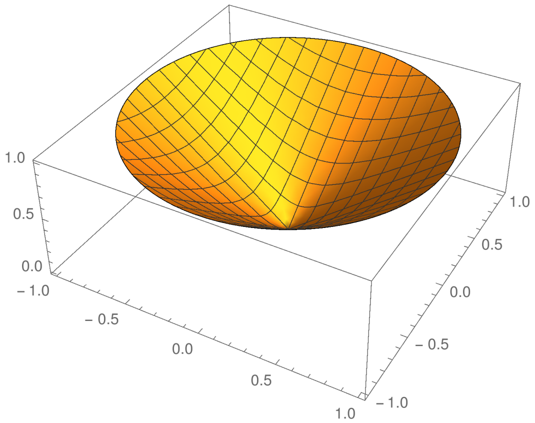
Esimerkki 3



Kaikki pisteet suoralla $x = 0$ ovat funktion $f(x, y) = -x^3$ satulapisteitä. Huomaa, että

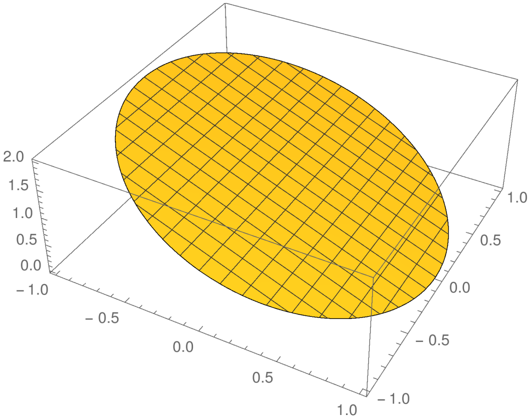
$$\nabla f(0, y) = -3x^2\mathbf{i}\Big|_{(0,y)} = \mathbf{0} \text{ kaikilla } y \in \mathbb{R}.$$

Esimerkki 4



Funktiolla $f(x, y) = \sqrt{x^2 + y^2}$ on lokaali minimi $f(0, 0) = 0$ pisteessä $(0, 0)$. Funktio f on jatkuva, mutta sen gradientti ∇f ei ole määritelty tässä pisteessä.

Esimerkki 5



Funktiolla $f(x, y) = 1 - x$ ei ole paikallisia ääriarvoja, jos sen määrittelyjoukko on koko taso $D = \mathbb{R}^2$. Jos määrittelyjoukoksi kuitenkin ajatellaan esimerkiksi kiekko $D = \{(x, y) : x^2 + y^2 < 1\}$, niin sen reunalla saadaan maksimi $f(-1, 0) = 2$ ja minimi $f(1, 0) = 0$.

Ääriarvojen luokittelu: johdanto

- Ääriarvojen luokittelu perustuu suureen $\Delta f = f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x})$ tarkasteluun kriittisessä pisteessä $\mathbf{x} \in D$.
- Jos Δf saa vain positiivisia arvoja (kun $\|\mathbf{h}\|$ on pieni), on piste $\mathbf{x}$ minimi ja negatiivisessa tapauksessa maksimi. Jos Δf vaihtaa merkkiä, niin piste $\mathbf{x}$ ei ole minimi eikä maksimi.
- Tämä johtaa funktion f toisen derivaatan tarkasteluun kriittisessä pisteessä. Yhden muuttujan tapauksessa:
  1. Jos $f''(x) < 0$, funktiolla f lokaali maksimi pisteessä x.
  2. Jos $f''(x) > 0$, funktiolla f lokaali minimi pisteessä x.
  3. Jos $f''(x) = 0$, testi ei anna vastausta, ja kysymys täytyy ratkaista muulla tavoin.
- Seuraavaksi yritetään yleistää tätä ajatusta monen muuttujan funktiolle.

Hessen matriisi

- Olkoon $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$ funktio, jolla on jatkuvat toisen kertaluvun osittaisderivaatat.
- Funktion f luonnollinen derivaattakäsite on gradientti, joka itsessään on vektoriarvoinen funktio $\nabla f\colon \mathbb{R}^n \to \mathbb{R}^n$.
- Siten funktion f toinen derivaatta on matriisi, jota nimitetään Hessen matriisiksi

$$H_f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2}{\partial x_1^2} f(\mathbf{x}) & \frac{\partial^2}{\partial x_2 \partial x_1} f(\mathbf{x}) & \dots & \frac{\partial^2}{\partial x_n \partial x_1} f(\mathbf{x}) \\ \frac{\partial^2}{\partial x_1 \partial x_2} f(\mathbf{x}) & \frac{\partial^2}{\partial x_2^2} f(\mathbf{x}) & \dots & \frac{\partial^2}{\partial x_n \partial x_2} f(\mathbf{x}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2}{\partial x_1 \partial x_n} f(\mathbf{x}) & \frac{\partial^2}{\partial x_2 \partial x_n} f(\mathbf{x}) & \cdots & \frac{\partial^2}{\partial x_n^2} f(\mathbf{x}) \end{bmatrix}.$$

- Koska f on kaksi kertaa jatkuvasti derivoituva, derivoinnin järjestystä voidaan vaihtaa, ja kyseinen matriisi on symmetrinen. Mutta mitä tarkoittaa matriisin positiivisuus?

Matriisin definiittisyys

- Symmetristä $n \times n$-matriisia A sanotaan positiividefiniitiksi, jos $\mathbf{x}^T A\mathbf{x} > 0$ kaikilla nollasta poikkeavilla vektoreilla $\mathbf{x} \in \mathbb{R}^n$. Tällaisilla matriiseilla on monia samoja ominaisuuksia kuin positiivisilla reaaliluvuilla.
- Yhtäpitävästi voidaan määritellä, että matriisi on positiividefiniitti, jos sen kaikki ominaisarvot ovat positiivisia.
- Toinen yhtäpitävä ehto on Sylvesterin kriteerio: Kaikilla $1 \leq m \leq n$ matriisin A vasemmasta yläkulmasta muodostettujen $m \times m$-matriisien determinantti on positiivinen (nolla ei käy).
- Vastaavalla tavalla voidaan määritellä negatiividefiniitti matriisi.
- Matriisin sanotaan olevan indefiniitti, jos sen kaikki ominaisarvot ovat nollasta poikkeavia ja sillä on vähintään yksi positiivinen sekä yksi negatiivinen ominaisarvo.
- Matriisi on negatiividefiniitti, jos Sylvesterin kriteeriossa esiintyvistä determinateista parittomat ovat negatiivisia ja parilliset positiivisia. Muussa tapauksessa matriisi on indefiniitti.

Toisen derivaatan testi monen muuttajan tapauksessa

Lause

Olkoon $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$ funktio, jolla on jatkuvat toiset osittaisderivaatat kriittisen pisteen $\mathbf{x} \in D$ ympäristössä. Tällöin:

a) Jos $H_f(\mathbf{x})$ on positiividefiniitti, niin f:llä on lokaali minimi pisteessä $\mathbf{x}$.

b) Jos $H_f(\mathbf{x})$ on negatiividefiniitti, niin f:llä on lokaali maksimi pisteessä $\mathbf{x}$.

c) Jos $H_f(\mathbf{x})$ on indefiniitti, niin f:llä on satulapiste pisteessä $\mathbf{x}$.

d) Muussa tapauksessa testi ei anna tietoa funktiosta f.

Esimerkki 6 1/2

- Etsitään ja luokitellaan funktion

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

  kriittiset pisteet.
- Yhtälöt kriittisille pisteille ovat

$$\begin{aligned} 0 &= f_1(x,y,z) = 2xy - 2, \\ 0 &= f_2(x,y,z) = x^2 + 2yz, \\ 0 &= f_3(x,y,z) = y^2 + 2z. \end{aligned}$$

- Nämä yhtälöt ratkaisemalla nähdään, että funktion f ainoa kriittinen piste on $P = (1, 1, -1/2)$.

Esimerkki 6 2/2

- Lasketaan Hessen matriisi:

$$H_f(1,1,-1/2) = \begin{bmatrix} 2 & 2 & 0 \\ 2 & -1 & 2 \\ 0 & 2 & 2 \end{bmatrix}.$$

- Käytetään Sylvesterin kriteeriota:

$$2 > 0, \quad \begin{vmatrix} 2 & 2 \\ 2 & -1 \end{vmatrix} = -6 < 0, \quad \begin{vmatrix} 2 & 2 & 0 \\ 2 & -1 & 2 \\ 0 & 2 & 2 \end{vmatrix} = -20 < 0.$$

- Siten funktiolla f on satulapiste pisteessä $P = (1,1,-1/2)$.

Sovellus: Lagrangen kertojien menetelmä 1/2

- Usein optimointitehtävissä halutaan asettaa rajoitteita optimoitaville muuttujille.
- Tyypillinen esimerkki tällaisesta tehtävästä on peltipurkin muodon optimointi: Halutaan minimoida purkin pinta-ala (eli käytetty materiaali) $f(h,r) = 2\pi rh + 2\pi r^2$ niin, että tilavuus $V(r,h) = \pi r^2 h$ on vakio.
- Asetetaan tehtävä: Minimoi $f(x,y)$ ehdolla $g(x,y) = 0$.
- Havaitaan, että mikäli ongelmalla on ratkaisu, niin ratkaisupisteessä (a,b) vektorien ∇f ja ∇g on oltava samansuuntaisia.
- Miksi? Koska muussa tapauksessa funktiolla f olisi nollasta poikkeva suunnattu derivaatta käyrän $g(x,y) = 0$ tangentin suuntaan pisteessä (a,b), ja siis minimi ei voi olla pisteessä (a,b).

Sovellus: Lagrangen kertojien menetelmä 2/2

- Tästä voidaan päätellä, että mikäli optimipiste on olemassa, se on Lagrangen funktion

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$$

kriittinen piste (eli gradientin nollakohta).

- Menetelmä yleistyy myös useammalle muuttujalle. Esimerkiksi kolmen muuttujan ja kahden rajoitusehdon tapauksessa Lagrangen funktio on

$$L(x, y, z, \lambda, \mu) = f(x, y, z) + \lambda g(x, y, z) + \mu h(x, y, z),$$

missä f on minimoitava funktio ja rajoite-ehdot ovat $g(x, y, z) = 0$ sekä $h(x, y, z) = 0$.

Esimerkki 7 1/2

- Minimoidaan funktio $f(x, y) = x^2 + y^2$ ehdolla $g(x, y) = x^2y - 16 = 0$.
- Muodostetaan Lagrangen funktio

$$L(x, y, \lambda) = x^2 + y^2 + \lambda(x^2y - 16).$$

- Yhtälöt kriittisille pisteille ovat

$$\begin{aligned} 0 &= \frac{\partial L}{\partial x} = 2x(1 + \lambda y), \\ 0 &= \frac{\partial L}{\partial y} = 2y + \lambda x^2, \\ 0 &= \frac{\partial L}{\partial \lambda} = x^2y - 16. \end{aligned}$$

Esimerkki 7 2/2

- Ensimmäisestä yhtälöstä saadaan $x = 0$ tai $\lambda y = -1$, mutta $x = 0$ on ristiriidassa kolmannen yhtälön kanssa.
- Siten
$$0 = 2y^2 + \lambda y x^2 = 2y^2 - x^2.$$
- Tästä saadaan edelleen $x = \pm\sqrt{2}y$, ja $2y^3 = 16$ eli $y = 2$.
- Minimejä on siis kaksi: $(\pm 2\sqrt{2}, 2)$.