TAMPEREEN YLIOPISTO

Kandidaattitutkielma

Kim Sarén

Taylorin sarja ja polynomi usean muuttujan tapauksessa

Luonnontieteiden tiedekunta

Matematiikka

Tammikuu 2018

Sisältö

1 Johdanto

Tämän tutkielman tarkoituksena on perehdyttää lukija Taylorin lauseeseen useamman muuttujan funktioiden tapauksessa. Erityisesti sovellamme lausetta ääriarvotestin todistukseen. Aiheen laajempi käsittely tapahtuu luvussa 3, jossa todistamme tutkielman päätuloksen. Aloitamme luvun selostamalla aihealuetta yleisesti läpi erityisesti jatkuvan derivoituvuuden sekä todistuksen kannalta relevantin yhdysjanan olemassaolon osalta. Esitämme varsinaisen todistuksen Taylorin lauseelle usean muuttujan tapauksessa pykälässä 3.2.

Luvussa 4 sovellamme aiemmin johdettua Taylorin useamman muuttujan funktioiden lausetta ääriarvotestin todistamiseen. Esitämme muutaman aihetta havainnollistavan esimerkin.

Käymme luvussa 2 luettelonomaisesti läpi muutamia aihealueeseen liittyviä määritelmiä ja lauseita. Todistamme matemaattisella induktiolla pykälän 3.2 kannalta relevantin aputuloksen johtamalla symbolisen esityksen todistuksessa 3.1 esiintyvälle derivaatalle.

Esitietoina oletamme lukijan hallitsevan potenssisarjaesityksen käsitteen erityisesti Taylorin lauseen kannalta, usean muuttujan funktioiden differentiaalilaskentaa sekä jatkuvuuden käsitteen yhden ja useamman muuttujan funktioille. Lineaarialgebrasta olisi hyvä ymmärtää joitakin alkeistuloksia, varsinkin Hessen matriisin käsite. Käytännössä vaatimukset vastaavat yliopiston ensimmäisiä analyysin ja lineaarialgebran kursseja sekä usean muuttujan differentiaalilaskennan kurssia.

2 Alustavia tarkasteluja

2.1 Taylorin lause yhden muuttujan funktioille

Luku 2 sisältää aihealueen läpikäymiseksi vaadittavat määritelmät, lauseet sekä aputulokset. Esitellään ensin pykälässä 2.1 Taylorin lauseen käsite yhden muuttujan funktioiden tapauksessa.

Lause 2.1 (Taylorin lause)**.** *Olkoot f funktio ja sen derivaatat $f', f'', \ldots, f^{(n)}$ jatkuvia jollakin välillä $[a, b]$ ja derivaatta $f^{(n+1)}$ olemassa välillä $]a, b[$. Tällöin on olemassa luku $\xi \in]a, b[$ siten, että*

$$f(b) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(b-a)^{n+1}. \tag{2.1}$$

Todistus. Sivuutetaan. Ks. [1, s. 132-133]. □

Määritelmä 2.1 (Taylorin polynomi ja jäännöstermi)**.** Taylorin lauseessa 2.1 esiintyvää summaa

$$P_n(b) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k \tag{2.2}$$

kutsutaan *Taylorin polynomiksi* pisteessä b. Erityistapauksesta $a = 0$ käytetään termiä *Maclaurinin polynomi*. Lauseen jälkimmäistä osaa

$$R_n(b) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(b-a)^{n+1} \tag{2.3}$$

kutsutaan sarjan *jäännöstermiksi*.

Huomautus. Koko lause rakentuu polynomin ja jäännöstermin summasta. Siis

$$f(x) = P_n(x) + R_n(x).$$

2.2 Aiheen kannalta relevantteja määritelmiä ja lauseita

Tässä pykälässä esitetään neljä määritelmää: usean muuttujan funktioiden n-kertainen jatkuvan derivoituvuus, matriisin definiittisyys sekä konveksin ja tähden muotoisen joukon käsitteet. Esitetään myös matriisien definiittisyyteen liittyvä seurauslause, joka havainnollistaa matriisin ominaisarvojen ja definiittisyyden välistä yhteyttä.

Määritelmä 2.2 (Useamman muuttujan funktioiden n-kertainen jatkuva derivoituvuus)**.** Olkoon $f\colon U \to \mathbb{R}$, missä $U \subset \mathbb{R}^m$ alue. Sanotaan, että funktio on *jatkuvasti derivoituva alueessa* U, mikäli osittaisderivaattafunktiot

$$D_i f\colon U \to \mathbb{R},\ \forall\ 1 \leq i \leq m$$

ovat jatkuvia alueessa $U \subset \mathbb{R}^m$. Merkitään tällöin $f \in \mathbf{C}^1(U)$. Edelleen, jos kaksinkertaiset osittaisderivaattafunktiot

$$D_{ij} f\colon U \to \mathbb{R},\ \forall\ 1 \leq i \leq m,\ 1 \leq j \leq m$$

ovat jatkuvia alueessa U, on funktio *kahdesti jatkuvasti derivoituva* ja merkitään $f \in \mathbf{C}^2(U)$. Mikäli em. ehdot pätevät myös n-kertaisille osittaisderivaatoille, on funktio *n-kertaa jatkuvasti derivoituva* alueessa U ja käytetään merkintää $f \in \mathbf{C}^{(n)}(U)$.

Huomautus. Edellisessä määritelmässä n-kertainen jatkuva derivoituvuus tarkoittaa tietysti, että kaikkien n-kertaisten tai sitä pienempien osittaisderivaattojen kombinaatioiden on oltava olemassa ja jatkuvia. Yleinen esitysmuoto on seuraava:

$$\frac{\partial^k f}{\partial x_1^{k_1} \partial x_2^{k_2} \dots \partial x_m^{k_m}},\ k = |k_1 + \cdots + k_m| = n.$$

Määritelmä 2.3 (Matriisin definiittisyys)**.** Olkoon $\boldsymbol{A}$ symmetrinen $n \times n$ matriisi. Tällöin matriisia kutsutaan

1. positiivisesti definiitiksi, jos neliömuoto $\boldsymbol{x}^{\boldsymbol{T}} \boldsymbol{A} \boldsymbol{x} > 0$ kaikilla $\boldsymbol{x} \in \mathbb{R}^n,\ \boldsymbol{x} \neq \boldsymbol{0}$,
2. negatiivisesti definiitiksi, jos neliömuoto $\boldsymbol{x}^{\boldsymbol{T}} \boldsymbol{A} \boldsymbol{x} < 0$ kaikilla $\boldsymbol{x} \in \mathbb{R}^n,\ \boldsymbol{x} \neq \boldsymbol{0}$,
3. indefiniitiksi, jos on olemassa sellaiset vektorit $\boldsymbol{x} \in \mathbb{R}^n,\ \boldsymbol{x} \neq \boldsymbol{0}$ ja $\boldsymbol{y} \in \mathbb{R}^n,\ \boldsymbol{y} \neq \boldsymbol{0}$, että $\boldsymbol{x}^{\boldsymbol{T}} \boldsymbol{A} \boldsymbol{x} > 0$ ja $\boldsymbol{y}^{\boldsymbol{T}} \boldsymbol{A} \boldsymbol{y} < 0$.

Tästä määritelmästä voidaan johtaa seurauslauseena tulos, joka yhdistää symmetrisen matriisin ominaisarvojen laadun sen definiittisyyteen. Useamman kuin kahden muuttujan funktioissa neliömuodon operointi voi olla teknisesti raskasta, joten tämä seurauslause antaa mielekkään vaihtoehdon matriisin definiittisyyden määrittämiseen. Ominaisarvojen laadun tunnistamisella on lisäksi teoreettista arvoa, jota hyödynnetään myöhemmin myös tässä tutkielmassa ääriarvotestiä todistettaessa.

Lause 2.2 (Matriisin definiittisyys ja ominaisarvot)**.** *Olkoon* $\boldsymbol{A}$ *symmetrinen* $n \times n$ *matriisi. Tällöin matriisin ominaisarvot ovat*

1. *positiivisia, jos ja vain jos matriisi on positiivisesti definiitti,*

2. *negatiivisia, jos ja vain jos matriisi on negatiivisesti definiitti,*

3. *positiivisia ja negatiivisia, jos ja vain jos matriisi on indefiniitti.*

Todistus. Sivuutetaan. Ks. [4, s. 856-858]. □

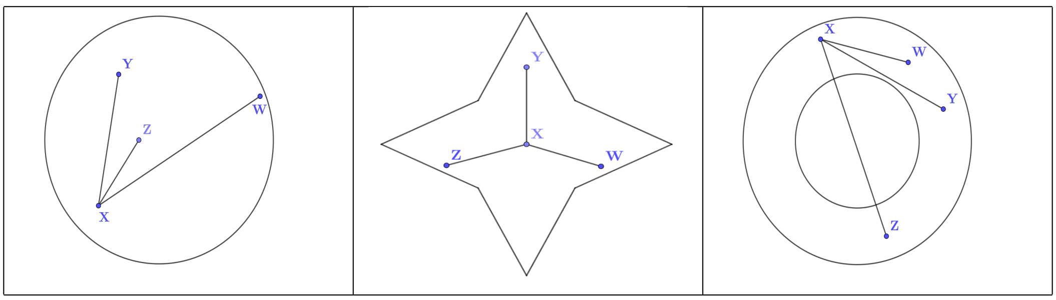
Määritelmä 2.4 (Konveksi joukko)**.** Joukkoa $S \subset \mathbb{R}^m$ sanotaan *konveksiksi*, mikäli kaikille sen alkioille $\boldsymbol{x}, \boldsymbol{y}$ pätee

$$\boldsymbol{x}, \boldsymbol{y} \in S \implies t\boldsymbol{x} + (1-t)\boldsymbol{y} \in S, \ t \in [0, 1].$$

Euklidisessa avaruudessa joukko on konveksi, jos kaikki sen alkioiden väliset janat sisältyvät samaan joukkoon. Esimerkiksi ellipsi on konveksi, mutta renkaan muotoinen alue ei.

Määritelmä 2.5 (Tähden muotoinen joukko)**.** Joukkoa $S \subseteq \mathbb{R}^m$ sanotaan *tähden* muotoiseksi pisteen $\boldsymbol{x} \in S$ suhteen, jos sen ja mielivaltaisen joukon pisteen $\boldsymbol{y}$ väliset yhdysjanat kuuluvat kokonaisuudessaan joukkoon S. Tällöin pisteen $\boldsymbol{x}$ suhteen pätee siis

$$\boldsymbol{x}, \boldsymbol{y} \in S \implies t\boldsymbol{x} + (1-t)\boldsymbol{y} \in S, \ t \in [0, 1].$$



Kuva 2.1: Vasemmalta oikealle: konveksi joukko, tähden muotoinen joukko, ei konveksi eikä tähden muotoinen joukko.

2.3 Aputuloksia

Lause 2.3 (Derivaatan symbolinen esitys lauseessa 3.1)**.** *Olkoon f ja g funktioita siten, että $f\colon \mathbb{R}^n \to \mathbb{R}$ ja $g\colon \mathbb{R} \to \mathbb{R}^n$. Olkoon $\boldsymbol{a} = (a_1, \dots, a_n)$ ja $\boldsymbol{v} = (x_1, \dots, x_n)$ pisteet, joiden yhdysjana on funktion $g(t) = t\boldsymbol{v} - (1-t)\boldsymbol{a}$ kuva. Tällöin, jos yhdistetty funktio $F(t) = f(g(t))$ on m kertaa ($m \geq 1$) jatkuvasti derivoituva, saamme sen derivaatoille yleisen esityksen*

$$F^{(j)}(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j}{\partial x_{i_1} \cdots \partial x_{i_j}} f(g(t))(x_{i_1} - a_{i_1}) \cdots (x_{i_j} - a_{i_j}),$$

kun $1 \leq j \leq m$.

Todistus. Todistetaan väite käyttämällä matemaattista induktiota. Käydään aloitusaskeleena kaksi ensimmäistä derivaattaa, tehdään sitten induktio-oletus ja todistetaan lopuksi, että väite pätee myös oletusta seuraavalle derivatalle.

1° Induktioaskel. Ensinnäkin,

(2.4)
$$F'(t) = \nabla(f(g(t))) \cdot g'(t) = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix} \begin{bmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{bmatrix} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(g(t))(x_j - a_j).$$

Lasketaan ensin asiaa selkeyttävä aputulos, jonka avulla saadaan ratkaistua derivaatta $F''(t)$. Derivoimalla tuloksessa 2.4 esiintyvä termi tulokseksena on

$$\frac{d}{dt}\left(\frac{\partial f}{\partial x_j}(g(t))\right) = \left(\frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}\right)(g(t))g'(t) = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(g(t))(x_i - a_i),$$

jonka avulla voidaan esittää toisen kertaluvun derivaatta

$$F''(t) = \frac{d}{dt}\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(g(t))(x_j - a_j) = \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(g(t))(x_i - a_i)(x_j - a_j). \tag{2.5}$$

2° Induktio-oletus. Oletetaan, että pätee

$$F^{(k)}(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \frac{\partial^k}{\partial x_{i_1} \cdots \partial x_{i_k}} f(g(t))(x_{i_1} - a_{i_1}) \cdots (x_{i_k} - a_{i_k}), \tag{2.6}$$

kun $1 \leq k < m$.

3° Todistus. Kuten aiemmin, johdetaan ensin luettavuutta selkeyttävä aputulos derivoimalla induktio-oletuksessa esiintyvä termi. Siis

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial^k}{\partial x_{i_1}\cdots\partial x_{i_k}}f(g(t))\right) &= \left(\frac{\partial}{\partial x_{i_{k+1}}}\frac{\partial^k f}{\partial x_{i_1}\cdots\partial x_{i_k}}\right)(g(t))g'(t)\\ &= \sum_{i_{k+1}=1}^{n}\frac{\partial^{k+1} f}{\partial x_{i_1}\cdots\partial x_{i_{k+1}}}g(t)(x_{i_{k+1}}-a_{i_{k+1}}).\end{aligned}\tag{2.7}$$

Todistetaan sitten induktioväite seuraavasti:

$$\begin{aligned}F^{(k+1)}(t) = \frac{d}{dt}F^{(k)}(t) &\overset{\text{IO}}{=} \frac{d}{dt}\sum_{i_1=1}^{n}\cdots\sum_{i_k=1}^{n}\frac{\partial^k}{\partial x_{i_1}\cdots\partial x_{i_k}}f(g(t))(x_{i_1}-a_{i_1})\cdots(x_{i_k}-a_{i_k})\\ &= \sum_{i_1=1}^{n}\cdots\sum_{i_k=1}^{n}\frac{d}{dt}\left(\frac{\partial^k}{\partial x_{i_1}\cdots\partial x_{i_k}}f(g(t))\right)(x_{i_1}-a_{i_1})\cdots(x_{i_k}-a_{i_k})\\ &\overset{2.7}{=} \sum_{i_1=1}^{n}\cdots\sum_{i_{k+1}=1}^{n}\frac{\partial^{k+1}}{\partial x_{i_1}\cdots\partial x_{i_{k+1}}}f(g(t))(x_{i_1}-a_{i_1})\cdots(x_{i_{k+1}}-a_{i_{k+1}}).\end{aligned}$$

Väite pätee induktion nojalla, kun $1 < k+1 \leq m$. □

3 Taylorin sarja ja polynomi

3.1 Esioletuksia ja aiheen havainnollistamista

Luvussa 3 käsitellään Taylorin lausetta useamman muuttujan funktioiden tapauksessa. Tässä pykälässä käydään aihealuetta vähän yleisemmin läpi ennen varsinaisia todistuksia sekä tehdään muutama todistusten kannalta aiheellinen huomio.

Jotta Taylorin lausetta voitaisiin soveltaa useamman muuttujan funktioiden tapauksessa, asetetaan muutama ehto seuraavasti:

1. Funktion on oltava $m + 1$ kertaa jatkuvasti derivoituva tutkittavassa joukossa U.

2. Pisteiden $\boldsymbol{v}$ ja $\boldsymbol{a}$ ($\boldsymbol{v} \neq \boldsymbol{a}$) välisen yhdysjanan täytyy sisältyä kokonaisuudessaan tutkittavaan alueeseen U. Asetetaan oletus, että alue U on tähden muotoinen kehitettävän pisteen $\boldsymbol{a}$ suhteen. Tällöin pisteiden välinen yhdysjana sisältyy joukkoon U.

Huomautus. Ehto 2. toteutuu erityisesti konveksin joukon tapauksessa, jolloin tietysti kaikkien joukon pisteiden väliset yhdysjanat sisältyvät myös joukkoon.

3.2 Taylorin lause usean muuttujan funktioille

Muodostetaan tässä pykälässä Taylorin lauseen esitys useamman muuttujan funktioiden tapaukseen. Aloitetaan käymällä läpi lauseissa ja todistuksessa käytettäviä notaatioita.

Olkoon $n \in \mathbb{N}$ positiivinen kokonaisluku ja $J = (j_1, j_2, \ldots, j_n)$ epänegatiivista termeistä koostuva jono. Käytetään jonon normista merkintää $|J| = j_1 + j_2 + \cdots + j_n$. $J!$ tarkoittaa luonnollisesti kertomaa $J! = j_1!j_2 \cdots j_n!$. Jos $|J| = k$, voidaan lauseen osittaisderivaatoista käyttää merkintää

$$D_J = \frac{\partial^k}{\partial x_1^{j_1} \partial x_2^{j_2} \cdots \partial x_n^{j_n}}.$$

Siis, jos f on funktio, jolla on n kappaletta muuttujia, niin

$$D_J f = \frac{\partial^k f}{\partial x_1^{j_1} \partial x_2^{j_2} \cdots \partial x_n^{j_n}}.$$

Olkoot $\boldsymbol{v} = (x_1, x_2, \ldots, x_n)$ ja $\boldsymbol{a} = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ kaksi mielivaltaisen monta termiä sisältävää vektoria. Määritellään merkintä $(\boldsymbol{v} - \boldsymbol{a})^J$ seuraavasti:

$$(\boldsymbol{v} - \boldsymbol{a})^J = (x_1 - a_1)^{j_1} (x_2 - a_2)^{j_2} \cdots (x_n - a_n)^{j_n}.$$

Esimerkki 3.1. Olkoon $n = 3$ ja $J = (1, 0, 4)$. Tällöin em. ehtojen mukaisesti

$$|J| = 5, \qquad J! = 1! \cdot 0! \cdot 4! = 24$$
$$D_J = \frac{\partial^4 f}{\partial x_1^1 \partial x_2^0 \partial x_2^4} = \frac{\partial^4 f}{\partial x_1^1 \partial x_2^4}$$

ja

$$(\boldsymbol{v} - \boldsymbol{a})^J = (x_1 - a_1)(x_3 - a_3)^4.$$

Seuraavassa lauseessa käytetään summamerkintää $\sum_{|J| \leq m}$, joka tarkoittaa kaikkien jonon J mahdollisten kombinaatioiden summaa siten, että $|J| = j_1 + j_2 + \cdots + j_n \leq m$.

Esimerkki 3.2. Olkoon $m = 1$ ja $n = 2$. Tällöin summanotaatiossa yhteenlaskettavat kombinaatiot olisivat

$$J = (0, 0) \qquad J = (0, 1) \qquad J = (1, 0).$$

Esitellään sitten tämän tutkielman keskeisin lause.

Lause 3.1 (Taylorin lause)**.** *Olkoon $U \subset \mathbb{R}^n$ tähden muotoinen joukko, funktio $f : U \to \mathbb{R}$ ja funktion osittaisderivaatat jatkuvia kertalukuun $m + 1$ asti joukossa U. Olkoon lisäksi $\boldsymbol{v}$ tutkittava joukon piste ja $\boldsymbol{a}$ piste, jonka suhteen sarjaa kehitetään. Tällöin voimme esittää funktion f sarjakehitelmänä*

$$f(\boldsymbol{v}) = \sum_{|J| \leq m} \frac{1}{J!} (D_J f)(\boldsymbol{a})(\boldsymbol{v} - \boldsymbol{a})^J + R_m(\boldsymbol{v}), \tag{3.1}$$

jossa jäännöstermi R_p esitetään muodossa

$$R_m(\boldsymbol{v}) = \sum_{|J| \leq m+1} \frac{1}{J!} (D_J f)(\boldsymbol{c})(\boldsymbol{v} - \boldsymbol{a})^J, \tag{3.2}$$

kun arvo $\boldsymbol{c}$ sijaitsee pisteiden $\boldsymbol{a}$ ja $\boldsymbol{v}$ välisellä yhdysjanalla.

Todistus. (vrt. [2, s. 155–157]) Koska joukko U on tähden muotoinen, pisteiden $\boldsymbol{v}$ ja $\boldsymbol{a}$ välinen yhdysjana kuuluu kokonaisuudessaan joukkoon. Määritellään ensin käyrä, jonka kuva on funktio $g : [0,1] \to \mathbb{R}^n$ seuraavasti:

$$g(t) = t\boldsymbol{v} + (1-t)\boldsymbol{a}.$$

Funktion g kuvajoukko $Im(g)$ esittää pisteiden $\boldsymbol{v}$ ja $\boldsymbol{a}$ välistä yhdysjanaa; siis, $Im(g) \subset U$. Koska joukko U on avoin, pätee tietysti myös

$$]g(-\epsilon), g(1+\epsilon)[\subset U$$

jollain arvolla $\epsilon > 0$.

Funktiolla $g(t)$ on ensimmäisen kertaluvun derivaatta $D_t g = \boldsymbol{v} - \boldsymbol{a}$ ja sitä suuremman kertaluvun derivaatat katoavat. Yhdisteellä $F(t) = f(g(t))$ on siis derivaattoja kertalukuun $m+1$ asti kun $t \in]-\epsilon, 1+\epsilon[$. Näillä merkinnöillä 2.1 on muotoa

$$F(t) = \sum_{j=0}^{m} \frac{F^{(j)}(0)}{j!} t^j + \frac{F^{(m+1)}(\boldsymbol{b})}{(m+1)!} t^{m+1}, \tag{3.3}$$

kun $\boldsymbol{b}$ kuuluu pisteiden $\mathbf{a}$ ja $\boldsymbol{v}$ välisellä yhdysjanalle. Erityisesti kun $t = 1$ esitykselle saadaan muoto

$$F(1) = \sum_{j=0}^{m} \frac{F^{(j)}(0)}{j!} + \frac{F^{(m+1)}(\boldsymbol{b})}{(m+1)!} = f(\mathbf{v}), \tag{3.4}$$

joka on funktion f arvo pisteessä $\boldsymbol{v}$.

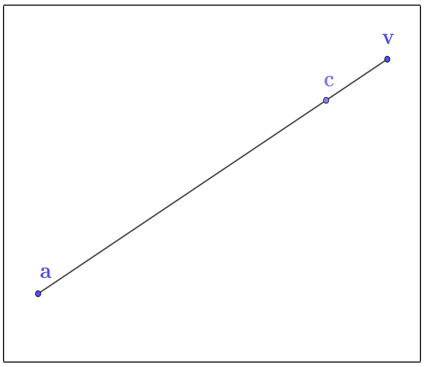
Tavoitteena on määrittää yleinen esitys derivaatalle $F^{(j)}(t)$. Sijoittamalla derivaatta edelliseen yhtälöön saadaan Taylorin lauseen esityksen useamman muuttujan funktoille. Periaatteessa ongelma ratkeaa yksinkertaisesti ketjusäännön toistuvalla soveltamisella $m+1$ kertaa.

Lauseen 2.3 mukaan, käyttämällä ketjusääntöä toistuvasti funktioon $F(t)$, derivaatta saa yleisen muodon

$$F^{(j)}(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j}{\partial x_{i_1} \cdots \partial x_{i_j}} f(g(t))(x_{i_1} - a_{i_1}) \cdots (x_{i_j} - a_{i_j}),$$

kun $1 \leq j \leq m+1$. Käyttämällä aiemmin määriteltyjä notaatioita lause supistuu muotoon

$$F^j(t) = \sum_{|J|=j} \frac{|J|!}{J!} (D_J f)(g(t))(\boldsymbol{v} - \boldsymbol{a})^J. \tag{3.5}$$



Kuva 3.1: Pisteiden $\boldsymbol{a}$ ja $\boldsymbol{v}$ välinen yhdysjana sekä janalle kuuluva piste $\boldsymbol{c}$. Funktio $F(t)$ kuvastaa tätä janaa arvoilla $t \in [0, 1]$.

Edelleen, kun $t = 0$, $g(0) = \boldsymbol{a}$ voidaan kirjoittaa 3.5 esitys

$$F^{j}(0) = \sum_{|J|=j} \frac{|J|!}{J!}(D_J f)(\mathbf{a})(\boldsymbol{v} - \boldsymbol{a})^J, \tag{3.6}$$

joka on tarvittu tulos Taylorin lauseen yleistyksen tekemiseksi. Sijoittamalla 3.6 alussa määriteltyyn Taylorin lauseen esitykseen 3.4, saadaan ratkaistua *Taylorin polynomi*

$$P_m(\boldsymbol{v}) = \sum_{j=0}^{m} \frac{F^{(j)}(0)}{j!} = \sum_{|J|\le m} \frac{1}{J!}(D_J f)(\boldsymbol{a})(\boldsymbol{v} - \boldsymbol{a})^J. \tag{3.7}$$

Jäännöstermi saadaan tekemällä erotus

$$\begin{aligned} R_m(\boldsymbol{v}) = f(\boldsymbol{v}) - P_m(\boldsymbol{v}) &= F(t) - \sum_{j=0}^{m} \frac{F^{(j)}(0)}{j!} \\ &= \frac{F^{(m+1)}(b)}{(m+1)!} \qquad \text{jollakin arvolla } b \in]0, 1[\\ &= \sum_{|J|=m+1} \frac{1}{J!}(D_J f)(\boldsymbol{c})(\boldsymbol{v} - \boldsymbol{a})^J, \end{aligned} \tag{3.8}$$

kun $\boldsymbol{c} = g(b)$, $\boldsymbol{c}$ kuuluu pisteiden $\boldsymbol{a}$ ja $\boldsymbol{v}$ väliselle yhdysjanalle.

Summaamalla Taylorin polynomin 3.7 ja jäännöstermin 3.8 olemme todistaneet lauseen 3.1

$$\begin{aligned} f(\boldsymbol{v}) &= P_m(\boldsymbol{v}) + R_m(\boldsymbol{v}) \\ &= \sum_{|J|\le m} \frac{1}{J!}(D_J f)(\boldsymbol{a})(\boldsymbol{v} - \boldsymbol{a})^J + \sum_{|J|=m+1} \frac{1}{J!}(D_J f)(\boldsymbol{c})(\boldsymbol{v} - \boldsymbol{a})^J(\boldsymbol{v}) \\ &= \sum_{|J|\le m} \frac{1}{J!}(D_J f)(\boldsymbol{a})(\boldsymbol{v} - \boldsymbol{a})^J + R_m(\boldsymbol{v}). \end{aligned}$$

$\square$

Esimerkki 3.3. Muodostetaan funktiolle $f(x, y) = x^3 + xy^2$ Taylorin toisen asteen sarjakehitelmä pisteen $(2, 1)$ suhteen. Aloitetaan laskemalla osittaisderivaatat

$$
\begin{aligned}
&f_x = 3x^2 + y^2, \quad f_y = 2xy, \quad f_{xx} = 6x, \quad f_{xy} = f_{yx} = 2y,\\
&f_{yy} = 2x, \quad f_{xxx} = 6, \quad f_{xxy} = 0, \quad f_{xyx} = f_{yxx} = 0\\
&f_{yyx} = 2, \quad f_{yyy} = 0.
\end{aligned}
$$

Soveltamalla lausetta 3.1 kahden muuttujan funktioiden tapaukseen saamme esityksen, jossa

$$
\begin{aligned}
f(x, y) = f(a, b) + f_x(a, b) + f_y(a, b) + \frac{1}{2}\left[\frac{\partial^2 f(a, b)}{\partial x^2}(x - a)^2\right.\\
\left. + 2\frac{\partial^2 f(a, b)}{\partial x \partial y}(x - a)(y - b) + (y - b)^2\right] + R_2,
\end{aligned} \tag{3.9}
$$

jossa jäännöstermi R_2 esitetään kuten

$$
\begin{aligned}
R_2 = \frac{1}{6}\left[\frac{\partial^3 f(\xi, \eta)}{\partial x^3}(x - a)^3 + 3\frac{\partial^3 f(\xi, \eta)}{\partial x^2 \partial y}(x - a)^2(y - b)\right.\\
\left. + 3\frac{\partial^3 f(\xi, \eta)}{\partial x \partial y^2}(x - a)(y - b)^2 + \frac{\partial^3 f(\xi, \eta)}{\partial y^3}(y - b)^3\right],
\end{aligned} \tag{3.10}
$$

jossa (ξ, η) kuuluu pisteiden (x, y) ja $(2, 1)$ väliselle yhdysjanalle.

Laskemalla funktion ja osittaisderivaattojen arvot pisteen $(2, 1)$ suhteen sekä tekemällä sijoitus kohtiin 3.9 ja 3.10, saadaan Taylorin sarjalle esitys

$$f(x, y) = x^3 + xy^2 = 16 + 14(x - 2) + 4(y - 1) + \frac{1}{2!}\left[12(x - 2)^2 + 4(y - 1)^2\right] + R_2,$$

jossa jäännöstermi on muotoa

$$R_2 = \frac{1}{3!}\left[6(\xi - 2)^3 + 2(\xi - 2)(\eta - 1)^2\right].$$

Esimerkki 3.4. Lasketaan funktion $F(x, y) = x \ln(1 + xy)$ Taylorin sarja kehitettynä pisteessä $(a, b) = (0, 0)$. Muistetaan logaritmifunktion alkeisfunktio yhden muuttujan Taylorin lauseen tapauksessa, joka on muotoa

$$\ln(1 + t) = \sum_{n=0}^{\infty} \frac{(-1)^{n-1} t^n}{n}.$$

Pystymme muodostamaan pyydetyn Taylorin sarjan soveltamalla kyseistä alkeisfunktiota (vrt. [3, s. 120]) tekemällä sijoitus $t = xy$ ja kertomalla sarja muuttujalla x, jolloin

$$F(x, y) = x \ln(1 + xy) = x \sum_{n=0}^{\infty} \frac{(-1)^{n-1}(xy)^n}{n} = x \sum_{n=0}^{\infty} \frac{(-1)^{n-1} x^n y^n}{n},$$

joka on haluttu Taylorin sarja.

4 Ääriarvot

4.1 Ääriarvotesti

Johdetaan tässä pykälässä esitys ääriarvotestille useamman muuttujan tapauksessa. Käsitellään todistuksessa tapausta $n = 2$, mutta todistuksen voi helposti yleistää myös useamman muuttujan funktioille tekemällä sopivat oletukset.

Lause 4.1 (Ääriarvotesti useamman muuttujan funktioiden tapauksessa)**.** *Olkoon $U \subseteq \mathbb{R}^2$ avoin ja funktio $f\colon U \to \mathbb{R}$ kahdesti jatkuvasti derivoituva funktio. Olkoon lisäksi $\boldsymbol{v_0}$ funktion kriittinen piste joukossa U ja $\boldsymbol{H_f}(\boldsymbol{v_0})$ funktion Hessen matriisi samassa pisteessä. Tällöin pätee seuraavat väitteet:*

1. *Jos $\boldsymbol{H_f}(\boldsymbol{v_0})$ on positiivisesti definiitti, niin $\boldsymbol{v_0}$ on paikallinen minimi,*
2. *Jos $\boldsymbol{H_f}(\boldsymbol{v_0})$ on negatiivisesti definiitti, niin $\boldsymbol{v_0}$ on paikallinen maksimi,*
3. *Jos $\boldsymbol{H_f}(\boldsymbol{v_0})$ on indefiniitti, niin $\boldsymbol{v_0}$ ei ole paikallinen minimi eikä maksimi.*

Todistus. (vrt. [2, s. 194–196]) Todistetaan lause tapauksessa $n = 2$. Olkoon $f\colon \mathbb{R}^2 \to \mathbb{R}$ funktio. Kehitetään pisteen $\boldsymbol{v_0} = (x_0, y_0)$ suhteen Taylorin sarja, kun $\boldsymbol{v} = (x, y)$ on mielivaltainen piste joukossa siten, että niiden yhdysjana sisältyy joukkoon. Käytetään erotuksesta $\boldsymbol{v} - \boldsymbol{v_0}$ merkintää $\boldsymbol{u}$. Tällöin saamme sarjakehitelmän

$$f(x, y) = f(\boldsymbol{v_0}) + \nabla f(\boldsymbol{v_0}) \cdot \boldsymbol{u} + R_1,$$

jossa jäännöstermi R_1 voidaan kirjoittaa muotoon

$$R_1 = \frac{1}{2}\left(\boldsymbol{u}^T \cdot \boldsymbol{H_f}(\xi, \eta)\boldsymbol{u}\right).$$

Kaavassa esiintyvä $\boldsymbol{H_f}$ on luontaisesti funktion Hessen matriisi yhdysjanalle kuuluvassa pisteessä $(\xi, \eta) \in]\boldsymbol{v_0}, \boldsymbol{v}[$. Edellä mainittujen kohtien perusteella saamme kirjoitettua erotuksen

(4.1)

$$\begin{aligned} f(\boldsymbol{v}) - f(\boldsymbol{v_0}) &= \frac{1}{2}\left(\boldsymbol{u}^T \cdot \boldsymbol{H_f}(\xi, \eta)\boldsymbol{u}\right) \\ &= \frac{1}{2}\left(f_{xx}(\xi, \eta)(x - x_0)^2 + 2f_{xy}(\xi, \eta)(x - x_0)(y - y_0) + f_{yy}(\xi, \eta)(y - y_0)^2\right), \end{aligned}$$

jossa ensimmäiset derivaatat ovat kadonneet, sillä määritelmän nojalla $\boldsymbol{v_0}$ on joukon kriittinen piste. Jatkuvuusehdot pätevät tietenkin lausekkeessa 4.1, sillä sen neliömuodon osat muodostuvat jatkuvista osittaisderivaatoista sekä niiden kertoimista.

Tutkitaan alustavasti pistettä $\boldsymbol{v_0}$ ja tehdään sen kautta johtopäätökset myös pisteen (ξ, η) suhteen. Ensinnäkin, koska Hessen matriisi on *symmetrinen*, seuraa siitä, että sille on olemassa reaaliarvoiset ominaisarvot. Voimme kirjoittaa pisteen $\boldsymbol{v_0}$ Hessen matriisin, kuten

$$\boldsymbol{H_f}(\boldsymbol{v_0}) = \begin{bmatrix} f_{xx}(\boldsymbol{v_0}) & f_{xy}(\boldsymbol{v_0}) \\ f_{yx}(\boldsymbol{v_0}) & f_{yy}(\boldsymbol{v_0}) \end{bmatrix}. \tag{4.2}$$

.

Oletetaan, että Hessen matriisi $\boldsymbol{H_f}(\boldsymbol{v_0})$ positiivisesti definiitti. Tällöin määritelmän 2.3 mukaan sen määräämä neliömuoto on positiivinen kaikilla vektoreilla $\boldsymbol{x} \in \mathbb{R}^2$. Tällöin

$$\boldsymbol{u}^T \cdot \boldsymbol{H_f}(\boldsymbol{v_0})\boldsymbol{u} > 0,$$

josta seuraa, että on olemassa sellainen luvun $\epsilon > 0$ määräämä ympäristö $B(\boldsymbol{v_0}, \epsilon)$, että pätee myös

$$f(\boldsymbol{v}) - f(\boldsymbol{v_0}) = \frac{1}{2}\left(\boldsymbol{u}^T \cdot \boldsymbol{H_f}(\xi, \eta)\boldsymbol{u}\right) > 0.$$

Toisin sanoen, tällöin piste $\boldsymbol{v_0}$ on *paikallinen minimi* ja kohta 1. on todistettu.

Vastaavasti kuin edellisessä kohdassa, oletetaan sitten, että Hessen matriisi on nyt negatiivisesti definiitti. Tällöin sen määräämä neliömuoto on negatiivinen kaikille vektoreille $\boldsymbol{x} \in \mathbb{R}^2$ ja edelleen

$$\boldsymbol{u}^T \cdot \boldsymbol{H_f}(\boldsymbol{v_0})\boldsymbol{u} < 0.$$

On jälleen olemassa sellainen luvun ϵ määräämä ympäristö $B(\boldsymbol{v_0}, \epsilon)$, että

$$f(\boldsymbol{v}) - f(\boldsymbol{v_0}) = \frac{1}{2}\left(\boldsymbol{u}^T \cdot \boldsymbol{H_f}(\xi, \eta)\boldsymbol{u}\right) < 0$$

kaikkialla kyseisessä ympäristössä. Siis tällöin $\boldsymbol{v_0}$ on *paikallinen maksimi* ja kohta 2. on todistettu.

Esitetään tämän kohdan todistus yksinkertaistettuna, koska pitäydymme tapauksessa $n = 2$. Oletetaan, että $\boldsymbol{H_f}(\boldsymbol{v_0})$ on indefiniitti. Tällöin matriisilla on määritelmän 2.2 mukaan niin positiivisia kuin negatiivisia ominaisarvoja. Kuten edellisissä kohdissa, on olemassa ympäristö $B(\boldsymbol{v_0}, \epsilon)$ siten, että matriisi on indefiniitti myös tässä ympäristössä. Olkoon (ξ, η) piste ympäristössä $B(\boldsymbol{v_0}, \epsilon)$, jolloin matriisi on tietenkin indefiniitti myös pisteessä (ξ, η). Algebran peruslauseen nojalla voidaan matriisin karakteristinen polynomi esittää pisteessä seuraavasti:

$$
\begin{aligned}
det(\boldsymbol{H_f}(\xi, \eta) - \lambda \boldsymbol{I}) &= \lambda^2 - \lambda(f_{xx}(\xi, \eta) + f_{yy}(\xi, \eta)) + f_{xx}(\xi, \eta) f_{yy}(\xi, \eta) - 2 f_{xy}(\xi, \eta) \\
&= \lambda^2 - tr(\boldsymbol{H_f}(\xi, \eta))\lambda + det(\boldsymbol{H_f}(\xi, \eta)) = (\lambda_1 - \lambda)(\lambda_2 - \lambda),
\end{aligned}
$$

josta johdetaan tulos

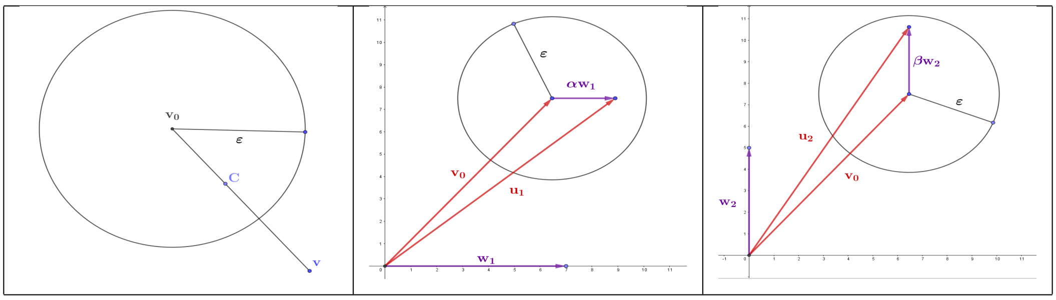
$$\lambda_1 \lambda_2 = det(\boldsymbol{H_f}(\xi, \eta)),$$

kun $\lambda = 0$. Koska matriisi on indefiniitti, ovat sen kaksi reaalista ominaisarvoa erimerkkiset 2.2 mukaisesti, jolloin tietenkin tulo

$$\lambda_1 \lambda_2 = det(\boldsymbol{H_f}(\xi, \eta)) = f_{xx}(\xi, \eta) f_{yy}(\xi, \eta) - f_{xy}^2(\xi, \eta) < 0. \tag{4.3}$$

Tällöin neliömuoto 4.1 saa niin positiivisia kuin negatiivisia arvoja kaikissa tapauksissa, jotka toteuttavat ehdon 4.3. Siis piste (ξ, η) ei ole minimi eikä maksimi ja kohta 3. on todistettu. □

Huomautus. Indefiniitin tapauksen voisi todistaa yleisemmin tutkimalla erillisiä vektoreita $\boldsymbol{w_1}$ ja $\boldsymbol{w_2}$, kun $\boldsymbol{w_1}$ vastaa negatiivisia ominaisarvoja, ja $\boldsymbol{w_2}$ vastaavasti positiivisia ominaisarvoja. Tällöin voidaan todistaa, että neliömuodon arvot ovat erimerkkiset lähestyttäessä erikseen vektoreiden $\boldsymbol{w_1}$ ja $\boldsymbol{w_2}$ suuntaan (ks. [2, s. 194–196]).



Kuva 4.1: Ensimmäinen kuva vasemmalta esittää positiivisesti/negatiivisesti definiitin matriisin tapausta, jossa piste $\boldsymbol{C} = (\xi, \eta)$ kuuluu tutkittavaan ympäristöön $B(\boldsymbol{v_0}, \epsilon)$. Keskimmäinen ja oikeanpuoleinen kuva esittävät lauseen 4.1 huomautuksen kuvaamaa tapausta indefiniitille matriisille, kun pisteeltä $\boldsymbol{v_0}$ siirrytään pisteiden $\boldsymbol{w_1}$ tai $\boldsymbol{w_2}$ suuntaan ympäristössä $B(\boldsymbol{v_0}, \epsilon)$.

Esimerkki 4.1. Määritetään funktion $f(x, y) = x^2 - 4x + 2y^2 + 7$ kriittiset pisteet sekä niiden tyypit. Derivoimalla saadaan $f_x = 2x - 4$ ja $f_y = 4y$. Näistä voidaan päätellä, että

$$f_x = 2x - 4 = 0, \qquad f_y = 4y = 0 \implies \begin{cases} x = 2 \\ y = 0, \end{cases}$$

eli funktion kriittinen piste on $(2, 0)$. Lasketaan sitten funktion osittaisderivaatat

$$f_{xx}(x, y) = 2, \quad f_{xy} = 0, \qquad f_{yy} = 4,$$

ja edelleen muodostetaan niiden avulla funktion Hessen matriisi kuten

$$\boldsymbol{H}_f(2, 0) = \begin{bmatrix} f_{xx}(2, 0) & f_{xy}(2, 0) \\ f_{yx}(2, 0) & f_{yy}(2, 0) \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}.$$

Lasketaan ominaisarvot matriisin karakterisen polynomin determinantista

$$det(\boldsymbol{H}_f(2, 0) - \lambda \boldsymbol{I}) = (2 - \lambda)(4 - \lambda) - 0 = \lambda^2 - 6\lambda + 8 = 0 \implies \begin{cases} \lambda_1 = 2 \\ \lambda_2 = 4 \end{cases}.$$

Siis, ominaisarvot ovat positiivisia, eli Hessen matriisi on positiivisesti definiitti ja lauseen 4.1 kohdan 1. mukaan piste on funktion *lokaali minimi*.

Esimerkki 4.2. Määritetään funktion $f(x, y) = \frac{1}{2}x^2 + 2x - 2y - y^2$ kriittiset pisteet sekä niiden tyyppi. Kuten aiemmin, määritetään kriittiset pisteet ensimmäisten derivaattojen nollakohdista, kuten

$$f_x = x + 2 = 0, \quad f_y = -2y - 2 = 0 \implies \begin{cases} x = -2 \\ y = -1, \end{cases}$$

eli $(-2, -1)$ on funktion kriittinen piste. Lasketaan funktion kaksinkertaiset osittaisderivaatat kriittisessä pisteessä, jolloin tulokseksi saadaan

$$f_{xx}(x, y) = 1, \quad f_{xy} = 0, \qquad f_{yy} = -2,$$

ja edelleen funktion Hessen matriisi kriittisessä pisteessä on

$$\boldsymbol{H}_f(-2, -1) = \begin{bmatrix} f_{xx}(-2, -1) & f_{xy}(-2, -1) \\ f_{yx}(-2, -1) & f_{yy}(-2, -1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}.$$

Laskemalla karakteristinen polynomi, tulokseksi saadaan

$$det(\boldsymbol{H_f}(-2,-1) - \lambda \boldsymbol{I}) = (1-\lambda)(-2-\lambda) - 0 = \lambda^2 - \lambda - 2 = 0 \implies \begin{cases} \lambda_1 = -1 \\ \lambda_2 = 2 \end{cases}.$$

Siis, Hessen matriisi on indefiniitti, ja lauseen 4.1 kohdan 3. mukaan piste $(-2,-1)$ ei ole ääriarvo.

Esimerkki 4.3. Määritetään vielä funktion $f(x, y) = x^2 + 2xy + y^2$ kriittiset pisteet sekä ääriarvot. Lasketaan ensimmäisten derivaattojen nollakohdat

$$f_x = 2x + 2y = 0, \quad f_y = 2y + 2x = 0 \implies \begin{cases} x = -y \\ y = -x, \end{cases}$$

josta nähdään, että jokainen piste $(x, -x)$ ja $(-y, y)$ ovat funktion kriittisiä pisteitä. Tutkitaan esimerkiksi pistettä $(0, 0)$. Kuten aiemmissa esimerkeissä, muodostetaan funktion Hessen matriisi, joka on

$$\boldsymbol{H_f}(0,0) = \begin{bmatrix} f_{xx}(0,0) & f_{xy}(0,0) \\ f_{yx}(0,0) & f_{yy}(0,0) \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$

Edelleen, sen karakteristinen on

$$det(\boldsymbol{H_f}(0,0) - \lambda \boldsymbol{I}) = (2-\lambda)(2-\lambda) - 4 = \lambda^2 - 4\lambda = \lambda(\lambda - 4) = 0 \implies \begin{cases} \lambda_1 = 0 \\ \lambda_2 = 4 \end{cases},$$

joten matriisi on *positiivisesti semidefiniitti* pisteessä $(0, 0)$. Lause 4.1 ei kuitenkaan määrittele semidefiniittiä tapausta, joten ääriarvotesti epäonnistuu.

On kuitenkin helppo todistaa, että piste $(0, 0)$ on funktion paikallinen ääriarvo ja vieläpä minimi. Ensinnäkin, funktio voidaan esittää muodossa

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

jonka funktion kuvaaja $S = \{(x, y, f(x, y)) | (x, y) \in R^2\}$ on ylöspäin aukeava paraboloidi. Edelleen, koska $f(x, y) \geq 0$, $f(0, 0) = 0$ ja $f(x, y) > 0$ kaikilla $(0, 0) \neq (x, y) \in R^2$, niin kyseessä on funktion lokaali ja myös globaali minimi.

Kirjallisuutta

[1] Protter, Murray. Morrey, Charles. *Intermediate calculus*, 2nd ed. New York [u.a.]: Springer, 1996.

[2] Lawrence, Corwin. Robert, Szczarba. *Multivariable calculus*, New York: Marcel Dekker, 1982.

[3] Aatos, Lahtinen. Erkki, Pehkonen. *Matematiikkaa soveltajille 2*, Helsinki: Kirjayhtymä, 1988.

[4] Pitkäranta, Juhani. *Calculus Fennicus*, Helsinki: Avoimet oppimateriaalit ry, 2015 [Viitattu 20.01.2017]. URL https://mycourses.aalto.fi/pluginfile.php/123037/mod_resource/content/2/calculusfennicus.pdf.