Taru Tiensuu

LAGRANGEN MENETELMÄ JA KUHN-TUCKERIN EHDOT

Optimointi taloustieteissä

Kandidaatintyö
Tekniikan ja luonnontieteiden tiedekunta
Tarkastaja: Merja Laaksonen
Huhtikuu 2023

TIIVISTELMÄ

Taru Tiensuu: Lagrangen menetelmä ja Kuhn-Tuckerin ehdot
Kandidaatintyö
Tampereen yliopisto
Tekniikka ja luonnontieteet, TkK
Huhtikuu 2023

Matemaattisessa optimoinnissa etsitään parasta ratkaisua kaikkien mahdollisten ratkaisujen joukosta. Optimoinnin tavoitteena voi olla esimerkiksi yrityksen taloudellisen tuloksen maksimointi, hävikin minimointi tai sijoitusstrategian optimointi. Usein tällaisissa tilanteissa tulee ottaa huomioon mahdollisiin ratkaisuihin vaikuttavat resurssit, kuten budjetti. Tässä kandidaatintyössä esitellään kaksi menetelmää yhtälörajoitetun optimointitehtävän ratkaisemiseksi: Lagrangen menetelmä sekä sen laajennus epäyhtälöehtoihin Kuhn-Tuckerin ehtojen avulla. Molemmissa menetelmissä optimointitehtävä esitetään Lagrangen funktiona, jonka avulla muodostetaan tarvittavat yhtälöt sekä selvitetään optimointitehtävän ratkaisu. Tässä työssä esiteltävien menetelmien merkittävimmät hyödyt saavutetaan optimointitehtävien ratkaisemisessa niin, ettei eksplisiittinen parametrisointi ehtofunktioiden suhteen ole tarpeellista.

Tässä kandidaatintyössä määritellään aluksi usean muuttujan peruskäsitteitä, kuten osittaisderivaatta ja gradientti. Lisäksi tarkastellaan, mitä vaaditaan usean muuttujan funktion ääriarvokohdalta ja mitä puolestaan satulapisteeltä sekä esitellään ylöspäin kupera usean muuttujan funktio. Matemaattisia taustatietoja soveltaen esitetään Lagrangen menetelmän ja Kuhn-Tuckerin ehtojen teoriat. Lagrangen menetelmän teoria esitellään aluksi kahden muuttujan funktion ääriarvojen löytämiseksi yhdellä ehtoyhtälöllä, jonka jälkeen esitetään yleinen menetelmä. Työssä havainnollistetaan sekä Lagrangen menetelmän että Kuhn-Tuckerin ehtojen tulkintoja myös geometrisesti. Lisäksi esitetään Kuhn-Tuckerin ehtojen yhteys Lagrangen funktion satulapisteeseen.

Matemaattisen teorian lisäksi tässä kandidaatintyössä esitellään muutamia menetelmien sovelluskohteita taloustieteiden näkökulmasta. Lagrangen menetelmää sovelletaan kuvitteellisen yrityksen tulosfunktion maksimointiin budjetti huomioon ottaen. Kuhn-Tuckerin ehtoja puolestaan sovelletaan niin ikään tulosfunktion maksimointiin kahdessa esimerkkitilanteessa, joista toisessa yrityksen hinnoittelu perustuu kysyntään ja toisessa monopoliyrityksen toimintaa säännellään.

Avainsanat: Lagrangen menetelmä, Kuhn-Tuckerin ehdot, optimointi taloustieteissä

ALKUSANAT

Haluan kiittää ohjaajaani Merja Laaksosta hyvin sujuneesta yhteistyöstä sekä kehittävistä ajatuksista kirjoitustyön aikana.

Kiitos erityisesti opiskeluystävilleni, joiden kanssa olen saanut jakaa kuluneet opiskeluvuodet sekä yhteisen kiinnostuksen luonnontieteitä kohtaan. Osoittamanne apu, kannustus ja ymmärtäväisyys ovat olleet korvaamattoman tärkeitä.

Kaunis kiitos myös muille ystävilleni sekä rakkaille läheisilleni. Kiitos, että tarjoatte tukea ja tasapainoa huomaavaisuudella, arvostuksella ja välittämisellä.

Tampereella, 20. huhtikuuta 2023

Taru Tiensuu

SISÄLLYSLUETTELO

1. JOHDANTO

Tässä kandidaatintyössä käsitellään Lagrangen menetelmää sekä sen laajennusta Kuhn-Tuckerin ehtoihin. Lagrangen menetelmän kehitti italialais-ranskalainen matemaatikko Joseph-Louis Lagrange määritelläkseen yleiset tasapainoehdot, joita statistiikassa tarvittiin. Lagrangen kertoimet mahdollistivat differentiaalilaskennan minimi- ja maksimikohtien käsittelyn samalla tavalla kuin statistiikassa silloin, kun pisteen tai pisteiden täytyi olla tasapainotilassa. [2]

Yhdysvaltalainen Harold W. Kuhn sekä kanadalainen Albert W. Tucker puolestaan esittelivät teoriansa epälineaarisesta ohjelmoinnista vuonna 1950 tuoden termin uutena matemaattiseen kirjallisuuteen. Työssään he esittelivät epälineaariseen ohjelmointiin liittyvän tehtävän sekä osoittivat teorian tämän ratkaisemiseen. Kyseinen teoria, joka opittiin tuntemaan nimellä Kuhn-Tuckerin teoria, esittää vaadittavat ehdot epälineaarisen ohjelmointitehtävän optimaaliselle ratkaisulle. Kuhn ja Tucker eivät kuitenkaan olleet teorian ensimmäiset todistajat. Myös amerikkalainen matemaatikko William Karush oli osoittanut teorian jo vuonna 1939, joten teoria tunnetaan myös nimellä Karush-Kuhn-Tuckerin ehdot. [5]

Funktion optimointi asetetuilla ehdoilla on hyödyllistä useilla tieteenaloilla. Tässä kandidaatintyössä käsitellään funktion optimointia Lagrangen menetelmää sekä Kuhn-Tuckerin ehtoja soveltaen. Luvussa 2 esitellään menetelmiin liittyvä matemaattinen tausta sekä usean muuttujan funktioiden peruskäsitteitä. Aluksi määritellään usean muuttujan funktio, funktion osittaiderivaatta sekä gradientti. Lisäksi tarkastellaan, mitä vaaditaan usean muuttujan funktion ääriarvokohdalta ja mitä puolestaan satulapisteeltä sekä esitellään ylöspäin kupera usean muuttujan funktio.

Luvussa 3 esitellään aluksi Lagrangen menetelmän teoria kahden muuttujan funktion ääriarvojen löytämiseksi yhdellä ehtoyhtälöllä. Tämän jälkeen esitetään yleinen menetelmä, jossa muuttujia ja ehtoyhtälöitä voi olla useampia. Teorian todistusta havainnollistetaan myös graafisesti, minkä lisäksi esitellään Lagrangen menetelmän geometrinen tulkinta xy-tasossa sekä yksinkertainen laskuesimerkki. Luvussa 4 puolestaan esitellään Kuhn-Tuckerin ehtojen teoria. Myös tätä todistusta havainnollistetaan graafisesti. Lopuksi esitetään ehtojen yhteys Lagrangen funktion satulapisteeseen.

Luvussa 5 käsitellään muutamia työssä esiteltävien menetelmien sovelluskohteita taloustieteiden näkökulmasta. Ensimmäisessä esimerkissä maksimoidaan kuvitteellisen yrityksen tulosfunktio Lagrangen menetelmää soveltaen. Toisessa ja kolmannessa esimerkissä esitetään teoriat niin ikään tulosfunktioiden maksimointiin Kuhn-Tuckerin ehtoja soveltaen. Ensimmäisessä esimerkissä yrityksen hinnoittelu perustuu kysyntään ja jälkimmäisessä esimerkissä monopoliyrityksen toimintaa säännellään.

2. MATEMAATTINEN TAUSTA

Tässä luvussa esitellään Lagrangen menetelmässä ja Kuhn-Tuckerin ehdoissa tarvittavia usean muuttujan peruskäsitteitä. Luvussa on käytetty lähdettä [1], ellei toisin mainita. Määritellään aluksi reaalifunktio f, jossa on n reaalista muuttujaa.

Määritelmä 2.1. Funktio $f : \mathbb{R}^n \to \mathbb{R}$ on sääntö, joka määrää yksikäsitteisen reaalisen arvon $f(x_1, x_2, \dots, x_n)$ jokaista määrittelyjoukon $D \in \mathbb{R}$ alkiota kohti. Määrittelyjoukon D pisteiden asettamaa reaaliluvuista $f(x_1, x_2, \dots, x_n)$ koostuvaa joukkoa kutsutaan funktion f arvojoukoksi.

Derivoituvalla funktiolla, jossa on n muuttujaa, on myös n kappaletta ensimmäisen kertaluvun osittaisderivaattoja, mikäli nämä ovat olemassa. Jokaista funktion muuttujaa kohti on siis oma ensimmäisen kertaluvun osittaisderivaatta.

Määritelmä 2.2. Funktion f osittaisderivaatta muuttujan x_j suhteen on funktio $f_{x_j}(x_1, x_2, \dots, x_n)$, joka määritellään erotusosamäärän raja-arvona

$$f_{x_j}(x_1, x_2, \dots, x_n) = \lim_{h \to 0} \frac{f(x_1, x_2, \dots, x_j + h, \dots, x_n) - f(x_1, x_2, \dots, x_n)}{h},$$

mikäli raja-arvo on olemassa. Raja-arvo voidaan kirjoittaa myös standardivektorin $\mathbf{e}_j$ avulla muodossa

$$\lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{e}_j) - f(\mathbf{x})}{h},$$

missä standardivektorin muut termit ovat nollia ja indeksiä j vastaava termi on 1 [6].

Huomataan, että $f_{x_j}(x_1, x_2, \dots, x_n)$ on funktion f ensimmäinen derivaatta tarkasteltuna ainoastaan muuttujan x_j funktiona. Tällöin funktion muita muuttujia pidetään vakioina derivoimisen suhteen. Osittaisderivaatta on olemassa, kun tämä raja-arvo on äärellinen. Merkitään ensimmäisen kertaluvun osittaisderivaattoja

$$\frac{\partial}{\partial x_j} f(x_1, x_2, \dots, x_n) = f_{x_j}(x_1, x_2, \dots, x_n) = D_{x_j} f(x_1, x_2, \dots, x_n).$$

Määritellään usean muuttujan funktion gradientti, joka kuvaa funktion osittaisderivaatat vektoriarvoiseksi funktioksi sovitussa järjestyksessä. Gradientti siis ilmaisee usean muuttujan funktion muutosnopeuden.

Määritelmä 2.3. Funktion $f(x_1, x_2, \ldots, x_n)$ gradientti

$$\nabla f(x_1, x_2, \ldots, x_n) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right) = (f_{x_1}, f_{x_2}, \ldots, f_{x_n}).$$

Funktion suurimpien ja pienimpien arvojen löytäminen on tärkeää useissa laskennallisissa sovelluksissa eri tieteenaloilla. Seuraava lause esittää kolme vaihtoehtoa pisteille, joissa usean muuttujan funktion ääriarvot voivat esiintyä.

Lause 2.4. *Funktion $f(x_1, x_2, \ldots, x_n)$ globaali tai paikallinen ääriarvokohta voi olla pisteessä $A = (a_1, a_2, \ldots, a_n)$, jos piste A on*

i. *funktion f kriittinen piste eli piste A toteuttaa yhtälön $\nabla f(a_1, a_2, \ldots, a_n) = \mathbf{0}$,*

ii. *funktion f singulaaripiste eli piste, jossa $\nabla f(a_1, a_2, \ldots, a_n)$ ei ole määritelty tai*

iii. *funktion f määrittelyalueen reunapiste.*

Todistus. Oletetaan, että piste $A = (a_1, a_2, \ldots, a_n)$ kuuluu funktion f määrittelyjoukkoon. Jos piste A ei ole funktion f määrittelyalueen reunapiste, niin sen täytyy olla funktion f määrittelyalueen sisäpiste. Jos piste A ei myöskään ole funktion f singulaaripiste, täytyy funktion f gradientin olla määritelty pisteessä A. Lisäksi mikäli piste A ei ole funktion f kriittinen piste, poikkeaa funktion f gradientti nollasta pisteessä A. Näin ollen funktiolla f on positiivinen suunnattu derivaatta funktion f gradientin suuntaan sekä negatiivinen suunnattu derivaatta vastakkaiseen suuntaan pisteessä A. Tämä tarkoittaa, että funktio f kasvaa toiseen suuntaan liikuttaessa poispäin pisteestä A ja pienenee liikuttaessa vastakkaiseen suuntaan. Tästä seuraa, ettei funktiolla f voi olla ääriarvokohtaa pisteessä A. Näin ollen pisteen A täytyy olla funktion f kriittinen piste, singulaaripiste tai määrittelyalueen reunapiste, jotta kyseisessä pisteessä voi olla ääriarvokohta. □

Ääriarvokohtien lisäksi tässä työssä esiteltävissä menetelmissä esiintyy termi satulapiste. Tällöin täytyy tutkia funktion toisia osittaisderivaattoja Hessen matriisin avulla. Tarkastellaan seuraavassa määritelmässä erityisesti termiä indefiniitti, jota tarvitaan myöhemmin Hessen matriisin ja funktion satulapisteen välisen yhteyden tulkinnassa.

Määritelmä 2.5. Merkitään nyt kokoa $n \times n$ olevan neliömatriisin B ominaisarvoja kirjaimella μ. Matriisin B ominaisarvo on μ, jos on olemassa sellainen nollasta poikkeava pystyvektori $\mathbf{x}$, että $B\mathbf{x} = \mu\mathbf{x}$ eli

$$(B - \mu I)\mathbf{x} = \mathbf{0},$$

missä I on identiteettimatriisi, jonka koko on $n \times n$. Seuraavat kohdat toteutuvat, jos matriisi B on symmetrinen reaalinen matriisi.

i. Matriisi B on positiivisesti definiitti, jos kaikki sen ominaisarvot ovat positiivisia.

ii. Matriisi B on negatiivisesti definiitti, jos kaikki sen ominaisarvot ovat negatiivisia.

iii. Matriisi B on indefiniitti, jos sillä on ainakin yksi positiivinen ominaisarvo ja ainakin yksi negatiivinen ominaisarvo.

Näiden termien avulla voidaan tulkita Lauseessa (2.6) esitettävää Hessen matriisia. Todistus löytyy lähteestä [1, s. 787], mutta nyt se sivuutetaan. Tässä työssä tärkeää on erityisesti vaatimus Hessen matriisin indefiniittisyydestä funktion satulapisteessä.

Lause 2.6. *Oletetaan, että piste* $A = (a_1, a_2, \ldots, a_n)$ *on funktion* $f(\boldsymbol{x}) = f(x_1, x_2, \ldots, x_n)$ *kriittinen piste ja funktion* f *määrittelyalueen sisäpiste. Oletetaan myös, että kaikki funktion* f *toisen kertaluvun osittaisderivaatat ovat jatkuvia pisteen* A *ympäristössä. Tällöin myös Hessen matriisi*

$$\mathcal{H}(\boldsymbol{x}) = \begin{bmatrix} f_{x_1x_1}(\boldsymbol{x}) & f_{x_1x_2}(\boldsymbol{x}) & \ldots & f_{x_1x_n}(\boldsymbol{x}) \\ f_{x_2x_1}(\boldsymbol{x}) & f_{x_2x_2}(\boldsymbol{x}) & \ldots & f_{x_2x_n}(\boldsymbol{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\boldsymbol{x}) & f_{x_nx_2}(\boldsymbol{x}) & \ldots & f_{x_nx_n}(\boldsymbol{x}) \end{bmatrix}$$

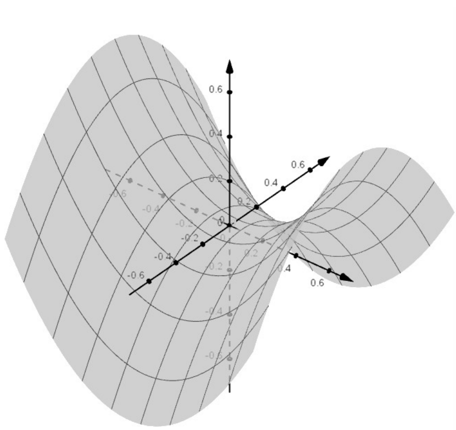
on jatkuva pisteen A *ympäristössä ja seuraavat kohdat toteutuvat.*

- **i.** *Jos* $\mathcal{H}(A)$ *on positiivisesti definiitti, niin pisteessä* A *on funktion* f *paikallinen minimikohta.*
- **ii.** *Jos* $\mathcal{H}(A)$ *on negatiivisesti definiitti, niin pisteessä* A *on funktion* f *paikallinen maksimikohta.*
- **iii.** *Jos* $\mathcal{H}(A)$ *on indefiniitti, niin pisteessä* A *on funktion* f *satulapiste.*
- **iv.** *Jos* $\mathcal{H}(A)$ *ei ole mikään aiemmista, niin testi ei anna informaatiota pisteestä* A*.*

Havainnollistetaan seuraavaksi kuvassa 2.1 satulapistettä graafisesti funktion $f(x, y) = x^2 - y^2$ avulla. Nyt

$$\mathcal{H}(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix},$$

jolloin ominaisarvot μ ovat -2 ja 2 ja siten $\mathcal{H}(x, y)$ on indefiniitti. Lisäksi $\nabla f(0, 0) = \mathbf{0}$, joten origossa on funktion f satulapiste.



Kuva 2.1. *Funktio* $f(x, y) = x^2 - y^2$*, jonka satulapiste on origossa.*

Esitellään vielä ylöspäin kupera usean muuttujan funktio, jota tarvitaan myöhemmin tässä työssä Kuhn-Tuckerin menetelmän sovelluskohteissa. Lauseessa (2.7) käytetään lähdettä [10].

Lause 2.7. *Olkoon funktio* $f(x_1,\ldots,x_n)$ *derivoituva ja olkoon sen määrittelyalue* D_f*. Funktio* $f(x_1,\ldots,x_n)$ *on ylöspäin kupera, jos*

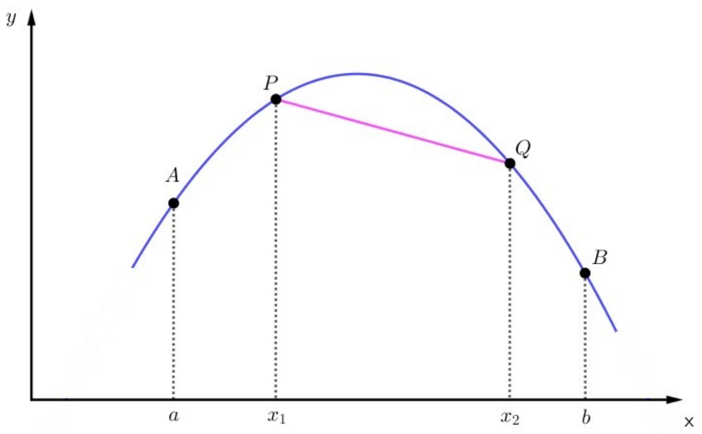
$$f(x_1\ldots,x_n) \leq f(a_1,\ldots,a_n) + \frac{\partial f}{\partial x_1}(a_1,\ldots,a_n)\cdot(x_1-a_1)+\cdots+\frac{\partial f}{\partial x_n}(a_1,\ldots,a_n)\cdot(x_n-a_n)$$

aina, kun $(a_1,\ldots,a_n)\in D_f$.

Alaspäin kuperan usean muuttujan funktion ehto saadaan muuttamalla Lauseen (2.7) epäyhtälömerkki vastakkaiseksi eli funktio $f(x_1,\ldots,x_n)$ on alaspäin kupera, jos

$$f(x_1\ldots,x_n) \geq f(a_1,\ldots,a_n) + \frac{\partial f}{\partial x_1}(a_1,\ldots,a_n)\cdot(x_1-a_1)+\cdots+\frac{\partial f}{\partial x_n}(a_1,\ldots,a_n)\cdot(x_n-a_n)$$

aina, kun $(a_1,\ldots,a_n)\in D_f$. Kuvassa 2.2 on havainnollistettu kahden muuttujan funktiota, joka on ylöspäin kupera. Tällöin siis funktion pisteiden välille piirretyn janan pisteiden arvojen tulee olla pienempiä tai yhtä suuria kuin funktion arvot ovat.



Kuva 2.2. *Ylöspäin kupera kahden muuttujan funktio.*

3. LAGRANGEN MENETELMÄ

Esitellään menetelmä aluksi kahden muuttujan funktion $f(x, y)$ ääriarvojen löytämiseksi yhdellä ehtoyhtälöllä $g(x, y) = 0$. Myöhemmin tässä luvussa esitellään yleistetty Lagrangen menetelmä, jossa muuttujia ja ehtoyhtälöitä voi olla useampia. Lauseessa (3.1) on käytetty lähdettä [1].

Lause 3.1. *Oletetaan, että funktioiden $f : f(x, y)$ ja $g : g(x, y)$ ensimmäisen kertaluvun osittaisderivaatat ovat jatkuvia käyrän $C : g(x, y) = 0$ pisteen $P_0 = (x_0, y_0)$ ympäristössä. Oletetaan lisäksi, että*

i. *rajoituttuessa käyrälle C funktiolla $f(x, y)$ on paikallinen ääriarvokohta pisteessä P_0,*

ii. *P_0 ei ole käyrän C päätepiste ja*

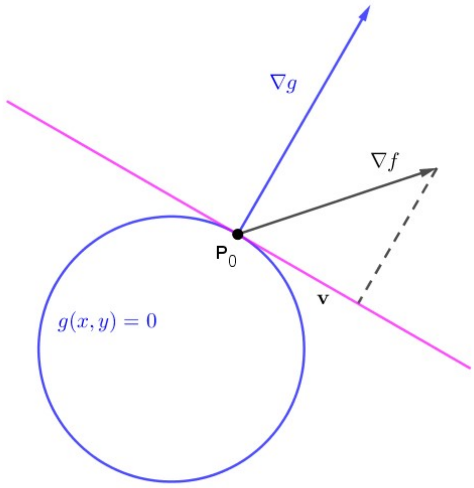
iii. $\nabla g(P_0) \neq \mathbf{0}$.

Tällöin piste (x_0, y_0, λ_0) on Lagrangen funktion

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

kriittinen piste.

Todistus. Oletusten ii. ja iii. perusteella funktio $g(x, y)$ on derivoituva pisteessä P_0 ja funktion g gradientti pisteessä P_0 on käyrän C tangentin normaali. Todistusta havainnollistaa kuva 3.1, jossa on ehtoyhtälön $g(x, y) = 0$ kuvaaja sekä tangenttisuora ja funktioiden f ja g gradientit pisteessä P_0, joka sijaitsee käyrällä $g(x, y)$.



Kuva 3.1. *Funktioiden f ja g gradientit pisteessä P_0.*

Jos funktioiden f ja g gradientit eivät ole yhdensuuntaiset pisteessä P_0, kuten kuvassa 3.1, on $\text{proj}_\mathbf{v}\nabla f \neq \mathbf{0}$ pisteessä P_0. Tällöin funktiolla f on positiivinen suunnattu derivaatta pisteessä P_0 vektoriprojektion **v** suuntaisesti ja negatiivinen suunnattu derivaatta vastakkaiseen suuntaan. Siis funktio f kasvaa tai pienenee liikuttaessa käyrää C pitkin poispäin pisteestä P_0 vektoriprojektion **v** suuntaisesti. Näin ollen funktiolla f ei voi olla paikallista maksimia tai minimiä pisteessä P_0. Oletuksen mukaan funktiolla f on paikallinen ääriarvokohta pisteessä P_0, joten

$$\text{proj}_\mathbf{v}\nabla f = \text{proj}_\mathbf{v}\nabla g = \mathbf{0}$$

pisteessä P_0. Oletuksen iii. mukaan funktion g gradientti pisteessä P_0 poikkeaa nollasta, joten täytyy olla olemassa sellainen vakio λ_0, että

$$\nabla f(P_0) = -\lambda_0 \nabla g(P_0)$$

eli

$$\nabla f(P_0) + \lambda_0 \nabla g(P_0) = \mathbf{0}.$$

Tämän vektoriyhtälön perusteella täytyy Lagrangen funktion L osittaisderivaattojen muuttujien x ja y suhteen olla nollia pisteessä (x_0, y_0, λ_0). Lisäksi Lagrangen funktion L kriittisessä pisteessä täytyy toteutua yhtälö

$$\frac{\partial L}{\partial \lambda} = g(x, y) = 0.$$

Tämä toteutuu pisteessä (x_0, y_0, λ_0), koska piste P_0 sijaitsee käyrällä C. Näin ollen piste (x_0, y_0, λ_0) on Lagrangen funktion $L(x, y, \lambda)$ kriittinen piste.

$\square$

Lauseen 3.1 perusteella yhtälöryhmän

$$\begin{cases} \dfrac{\partial L}{\partial x} = f_x(x,y) + \lambda g_x(x,y) = 0 \\ \dfrac{\partial L}{\partial y} = f_y(x,y) + \lambda g_y(x,y) = 0 \\ \dfrac{\partial L}{\partial \lambda} = g(x,y) = 0 \end{cases}$$

tulee toteutua Lagrangen funktion

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

kriittisessä pisteessä. Yhtälöryhmän kaksi ensimmäistä yhtälöä kuvaavat funktioiden f ja g gradienttien vaadittua yhdensuuntaisuutta. Kolmas yhtälö toteutuu, kun piste (x,y) sijaitsee käyrällä C.

Lagrangen funktiossa L esiintyvää kerrointa λ kutsutaan Lagrangen kertoimeksi. Lisäksi huomioidaan, että Lauseessa 3.1 oletetaan, että ratkaisu on olemassa. Lause 3.1 ei kuitenkaan takaa, että ratkaisu olisi olemassa.

Esitellään seuraavaksi yleinen Lagrangen menetelmä, kun muuttujia on n kappaletta. Esitetään muuttujat vektorina $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Nyt ehtoyhtälöitä on m kappaletta, missä $m \leq n - 1$. Lauseessa (3.2) on käytetty lähdettä [1].

Lause 3.2. *Oletetaan, että ratkaisu on olemassa pisteessä P_0. Oletetaan lisäksi, että funktion f ja kaikkien ehtoyhtälöiden g_j ensimmäiset osittaisderivaatat ovat olemassa ja ne ovat jatkuvia pisteen P_0 ympäristössä. Ehtoyhtälöiden leikkauksen täytyy myös olla ehtoyhtälöiden ratkaisujoukko lähellä pistettä P_0.*

Tällöin piste P_0 on $n + m$ muuttujaa sisältävän Lagrangen funktion

$$L(\boldsymbol{x}, \lambda_1, \ldots, \lambda_m) = f(\boldsymbol{x}) + \sum_{j=1}^{m} \lambda_j g_j(\boldsymbol{x})$$

kriittisten pisteiden joukossa. Jokaisen kriittisen pisteen tulee toteuttaa $n + m$ yhtälöä

$$\frac{\partial L}{\partial x_i} = 0, \qquad 1 \leq i \leq n,$$

$$\frac{\partial L}{\partial \lambda_j} = g_j(\boldsymbol{x}) = 0, \qquad 1 \leq j \leq m.$$

Tarkastellaan menetelmää vielä esimerkin kautta. Ratkaistaan kahden ehtofunktion esimerkki hyödyntäen Lausetta 3.2.

Esimerkki 3.3. Etsitään funktion

$$f(x, y, z) = xyz$$

suurin arvo ehdoilla

$$s(x, y, z) = x + y + z - 28 = 0$$

ja

$$t(x, y, z) = x - y + z - 8 = 0.$$

Oletetaan lisäksi, että $x, y, z \geq 0$. Käytännössä etsitään siis funktion f kriittisiä pisteitä tasojen s ja t leikkaussuoralta. Lagrangen funktio on nyt muotoa

$$L = (x, y, z, \lambda_1, \lambda_2) = f(x, y, x) + \lambda_1 s(x, y, z) + \lambda_2 t(x, y, z).$$

Lasketaan aluksi funktioiden f, s ja t gradientit.

$$\nabla f(x, y, z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = (yz, xz, xy),$$

$$\nabla s(x, y, z) = \left(\frac{\partial s}{\partial x}, \frac{\partial s}{\partial y}, \frac{\partial s}{\partial z}\right) = (1, 1, 1)$$

ja

$$\nabla t(x, y, z) = \left(\frac{\partial t}{\partial x}, \frac{\partial t}{\partial y}, \frac{\partial t}{\partial z}\right) = (1, -1, 1).$$

Lauseen 3.2 mukaisesti muodostetaan yhtälöryhmä

$$\begin{cases} \frac{\partial L}{\partial x} = f_x(x, y, z) + \lambda_1 s_x(x, y, z) + \lambda_2 t_x(x, y, z) = yz + \lambda_1 + \lambda_2 = 0 \\ \frac{\partial L}{\partial y} = f_y(x, y, z) + \lambda_1 s_y(x, y, z) + \lambda_2 t_y(x, y, z) = xz + \lambda_1 - \lambda_2 = 0 \\ \frac{\partial L}{\partial z} = f_z(x, y, z) + \lambda_1 s_z(x, y, z) + \lambda_2 t_z(x, y, z) = xy + \lambda_1 + \lambda_2 = 0 \\ \frac{\partial L}{\partial \lambda_1} = s(x, y, z) = x + y + z - 28 = 0 \\ \frac{\partial L}{\partial \lambda_2} = t(x, y, z) = x - y + z - 8 = 0. \end{cases}$$

Yhtälöryhmä saadaan muotoon

$$\begin{cases} \lambda_1 + \lambda_2 = -yz \\ \lambda_2 = \lambda_2 + 81 \\ \lambda_1 + \lambda_2 = -xy \\ x + z = 18 \\ y = 10 \end{cases}$$

eli

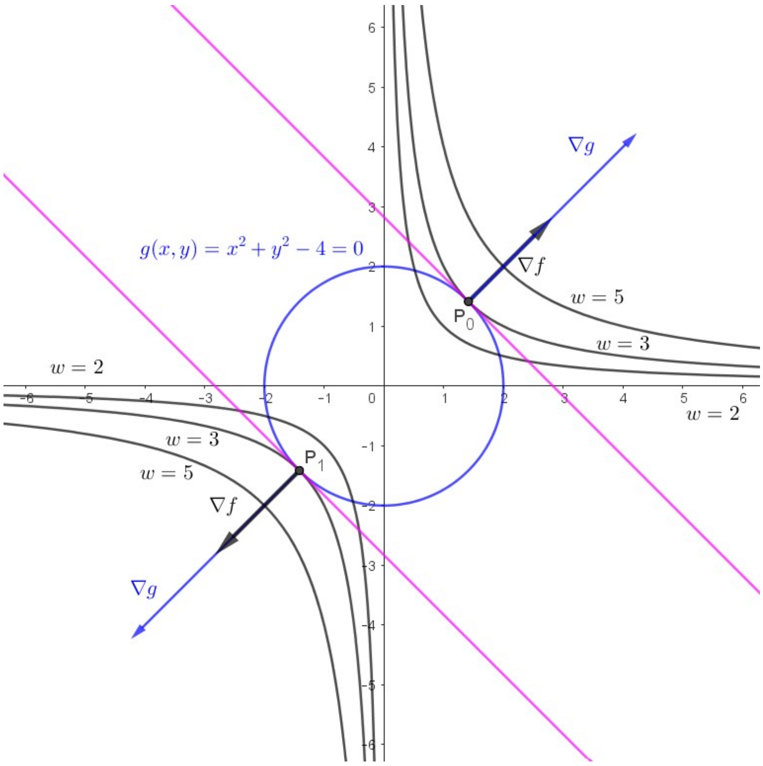
$$\begin{cases} x = 9 \\ y = 10 \\ z = 9 \\ \lambda_1 = -\frac{171}{2} \\ \lambda_2 = -\frac{9}{2}. \end{cases}$$

Nyt tiedetään, että funktion f kriittinen piste on $(9, 10, 9)$. Funktion suurin arvo on siis

$$f(9, 10, 9) = 9 \cdot 10 \cdot 9 = 810.$$

Etsitään seuraavaksi kuvissa 3.2 ja 3.3 esimerkin avulla geometrisesti kriittisiä pisteitä käyrälle $w = f(x, y) = xy + 1$ yhdellä ehtoyhtälöllä $g(x, y) = x^2 + y^2 - 4 = 0$. Tarkastellaan siis funktion f kriittisiä pisteitä origokeskisellä ympyrällä, jonka säde on 2.

Kuvaan 3.2 on piirretty optimoitavan funktion f satunnaisia käyriä $w = 2, 3, 5$. Kuten huomataan, käyrällä $w = 5$ ei ole pisteitä ympyrällä. Sen sijaan käyrä $w = 3$ sivuaa ympyrää pisteessä P_0. Pisteessä P_0 on siis funktion $f(x, y)$ paikallinen maksimi, sillä liikuttaessa ympyrän kehää pitkin päädytään molemmissa suunnissa pienempiin käyrän w arvoihin. Samoin tapahtuu liikuttaessa ympyrän kehää pitkin pisteestä P_1. Molemmissa pisteissä P_0 ja P_1 sijaitsee siis optimoitavan funktion f paikalliset maksimikohdat.



Kuva 3.2. *Funktion f paikalliset maksimikohdat pisteissä P_0 ja P_1.*

Kuvassa 3.2 on myös piirretty tangenttisuorat ympyrän ja käyrän $w = 3$ leikkauspisteisiin P_0 ja P_1. Näiden lisäksi leikkauspisteisiin on asetettu funktioiden f ja g gradientit

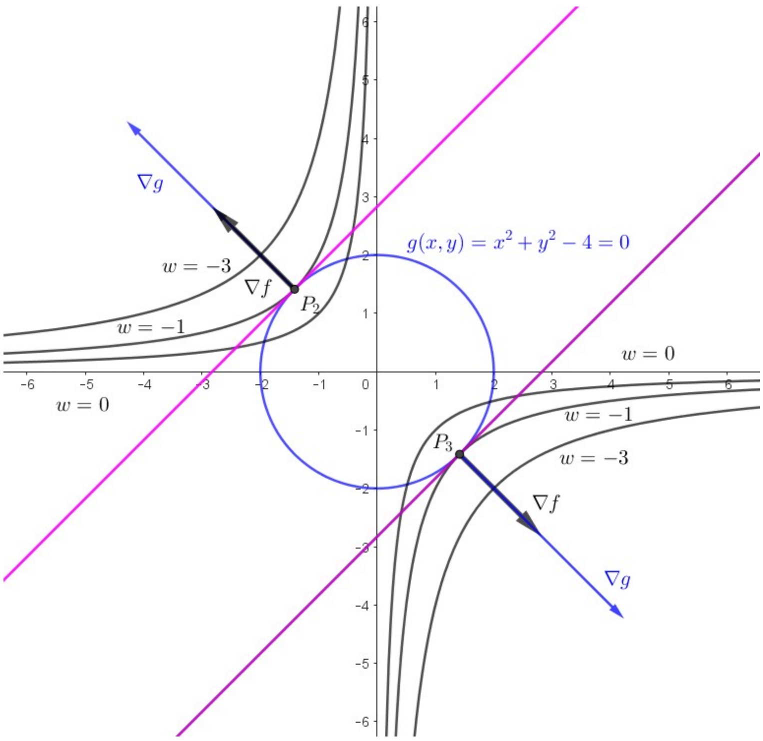
$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (y, x)$$

ja

$$\nabla g = \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right) = (2x, 2y)$$

Kuvasta nähdään, että gradientit ovat Lagrangen menetelmän vaatimusten mukaisesti kohtisuorassa tangentteja vastaan pisteissä P_0 ja P_1 sekä keskenään yhdensuuntaiset.

Vastaavasti tarkastellaan optimoitavan funktion f paikallisia minimikohtia. Kuvassa 3.3 on esitetty satunnaisia käyriä $w = f(x, y) = xy + 1$, $w = 0, -1, -3$. Huomataan, että optimoitavan funktion paikalliset minimikohdat löydetään käyrältä $w = -1$ pisteistä P_2 ja P_3. Nyt liikuttaessa ympyrän kehää pitkin päädytään molemmissa suunnissa suurempiin käyrän w arvoihin. Kuvassa 3.3 pisteisiin P_2 ja P_3 on asetettu tangenttisuorat molempiin ympyrän ja käyrän $w = -1$ leikkauskohtiin sekä funktioiden f ja g gradientit. Kuvasta nähdään jälleen, että gradientit ovat Lagrangen menetelmän vaatimusten mukaisesti kohtisuorassa tangentteja vastaan pisteissä P_2 ja P_3 sekä keskenään yhdensuuntaiset.



Kuva 3.3. *Funktion f paikalliset minimikohdat pisteissä P_2 ja P_3.*

4. KUHN-TUCKERIN EHDOT

Matemaattisessa optimoinnissa Kuhn-Tuckerin ehdot, jotka tunnetaan myös nimellä Karush-Kuhn-Tuckerin ehdot, hyödyntävät Lagrangen menetelmän tavoin ensimmäisen kertaluvun osittaisderivaattoja. Näitä sovelletaan erityisesti epälineaarisessa ohjelmoinnissa. Lagrangen menetelmän tavoin optimointitehtävä esitetään Lagrangen funktiona, mutta nyt ehdot voivat olla epäyhtälöitä.

Seuraavat Lauseessa (4.1) [7] esitettävät Kuhn-Tuckerin ehdot (myöhemmin KT-ehdot) ovat välttämättömiä yleisessä yhtälörajoitetussa optimointitehtävässä. Käydään KT-ehdot läpi seuraavassa minimointitehtävässä, jossa optimoitava funktio on $f(\mathbf{x})$ sekä ehtoepäyhtälöt

$$g_i(\mathbf{x}) \leq 0, \qquad i = 1, 2, \ldots, m.$$

Hyödyntämällä apumuuttujia $y_i,\ i = 1, 2, \ldots, m$ voidaan ehtoyhtälö kirjoittaa muodossa

$$g_i(\mathbf{x}) + y_i^2 = 0, \qquad i = 1, 2, \ldots, m.$$

Lause 4.1. *Olkoon optimoitava funktio $f(\boldsymbol{x})$ sekä rajoitteena ehtojoukko*

$$K = \{\boldsymbol{x} | g_i(\boldsymbol{x}) \leq 0, \qquad i = 1, 2, \ldots, m\},$$

missä $\boldsymbol{x} \in \mathbb{R}^n$. Oletetaan, että kaikki osittaisderivaatat f_{x_j} ja $\frac{\partial g_i(x)}{\partial x_j}$, $j = 1, 2, \ldots, n$, $i = 1, 2, \ldots, m$ ovat olemassa. Mikäli optimoitava funktio $f(\boldsymbol{x})$ saavuttaa paikallisen minimin pisteessä $\boldsymbol{x}^*$, niin on olemassa sellainen Lagrangen kertoimista koostuva vektori $\boldsymbol{\lambda}^*$, että

$$\frac{\partial f(\boldsymbol{x}^*)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i^* \frac{\partial g_i(\boldsymbol{x}^*)}{\partial x_j} = 0, \qquad j = 1, 2, \ldots, n, \tag{4.1}$$

$$g_i(\boldsymbol{x}^*) \leq 0, \qquad i = 1, 2, \ldots, m, \tag{4.2}$$

$$\lambda_i^* g_i(\boldsymbol{x}^*) = 0, \qquad i = 1, 2, \ldots, m \tag{4.3}$$

ja

$$\lambda_i^* \geq 0, \qquad i = 1, 2, \ldots, m. \tag{4.4}$$

KT-ehdot kohdissa (4.1)–(4.4) ovat vaadittavia ehtoja paikalliselle minimikohdalle. Etsittäessä paikallista maksimikohtaa korvataan vaatimus $\lambda_i^* \geq 0$ yhtälöstä (4.4) ehdolla $\lambda_i^* \leq 0$. Lagrangen kertoimet eivät siis saa olla negatiivisia minimointitehtävissä, kun taas maksimointitehtävissä ne eivät saa olla positiivisia.

Todistus. Lagrangen funktio voidaan nyt esittää muuttujien $\mathbf{x}$, apumuuttujien $\mathbf{y}$ ja Lagrangen kertoimien $\boldsymbol{\lambda}$ avulla muodossa

$$L(\mathbf{x}, \mathbf{y}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i (g_i(\mathbf{x}) + y_i^2).$$

Optimoitavan funktion f paikallisen minimikohdan $\mathbf{x}^*$ tulee toteuttaa ehdot

$$\frac{\partial L}{\partial x_j} = \frac{\partial f(\mathbf{x}^*)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i^* \frac{\partial (g_i(\mathbf{x}^*) + (y_i^*)^2)}{\partial x_j} = 0, \qquad j = 1, 2, \ldots, n, \tag{4.5}$$

$$\frac{\partial L}{\partial y_i} = 2\lambda_i^* y_i^* = 0, \qquad i = 1, 2, \ldots, m \tag{4.6}$$

ja

$$\frac{\partial L}{\partial \lambda_i} = g_i(\mathbf{x}^*) + (y_i^*)^2 = 0, \qquad i = 1, 2, \ldots, m. \tag{4.7}$$

Osoitetaan, että ehdot yhtälössä (4.6) vastaavat KT-ehtoja yhtälössä (4.3). Oletetaan, että $\lambda_i^* = 0$, jolloin $\lambda_i^* y_i^* = \lambda_i^* g_i(\mathbf{x}^*) = 0$. Tällöin molemmat kyseisistä ehdoista toteutuvat.

Jos $\lambda_i^* \neq 0$, niin yhtälön (4.6) perusteella $y_i^* = 0$, jolloin $(y_i^*)^2 = -g_i(\mathbf{x}^*) = 0$. Myös tällöin yhtälö (4.3) toteutuu. Toisaalta yhtälön (4.3) mukaisesti $g_i(\mathbf{x}^*) = 0$, jolloin $y_i^* = 0$ ja jälleen yhtälö (4.6) toteutuu. Apumuuttujat y_i, $i = 1, 2, \ldots, m$ voidaan eliminoida ehdoista (4.5) ja (4.7), jolloin jäljelle jää KT-ehtoja vastaavat yhtälöt.

Osoitetaan vielä, että minimointitehtävän epäyhtälöehtoihin liittyvät Lagrangen kertoimet eivät saa olla negatiivisia. Ehtojoukon

$$K = \{\mathbf{x} | g_i(\mathbf{x}) \leq 0, \qquad i = 1, 2, \ldots, m\}$$

epäyhtälöehdon $g_i(\mathbf{x}) \leq 0$ oikean puolen kasvattaminen nollasta voi ainoastaan pienentää minimoitavan funktion f arvoa. Näin ollen epäyhtälöehtoihin liittyvät Lagrangen kertoimet eivät voi olla negatiivisia minimointitehtävän tapauksessa. [9] Vastaavasti Lagrangen kertoimet eivät voi olla positiivisia maksimointitehtävän tapauksessa.

□

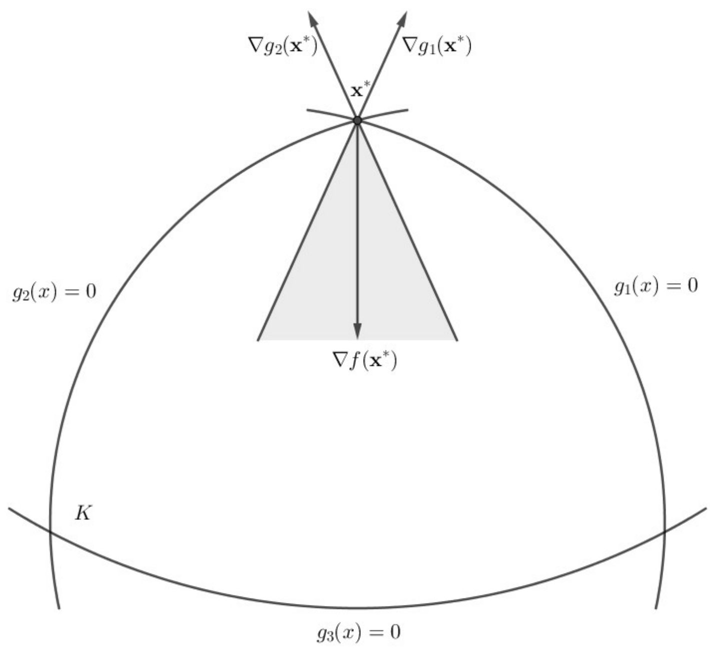
KT-menetelmä siis yleistää Lagrangen menetelmän hyväksymällä myös epäyhtälöehdot. Esitellään seuraavaksi ehtojen geometrinen tulkinta. Kirjoitetaan nyt KT-ehto yhtälöstä (4.1) muodossa

$$\frac{\partial f(\mathbf{x}^*)}{\partial x_j} = -\sum_{i=1}^{m} \lambda_i^* \frac{\partial g_i(\mathbf{x}^*)}{\partial x_j}, \qquad j = 1, 2, \ldots, n$$

eli

$$\nabla f(\mathbf{x}^*) = -\sum_{i=1}^{m} \lambda_i^* \nabla g_i(\mathbf{x}^*).$$

Nyt $\nabla f(\mathbf{x})$ on optimoitavan funktion f gradienttivektori ja $\nabla g_i(\mathbf{x})$, $i = 1, 2, \ldots, m$ ovat ehtofunktioiden gradienttivektoreita. Siis optimoitavan funktion f gradienttivektorin tulee olla ei-positiivinen summa ehtofunktioiden g_i gradienttivektoreista ratkaisupisteessä $\mathbf{x}^*$. Tällöin optimoitavan funktion f gradienttivektorin tulee sijaita alueella, jonka rajaavat sisäänpäin suuntautuvat normaalit ratkaisupisteessä $\mathbf{x}^*$. [7] Kuvassa 4.1 geometrinen tulkinta on esitetty esimerkin avulla optimointitehtävälle, jossa $n = 2,\ m = 3$.



Kuva 4.1. *KT-ehtojen geometrinen tulkinta esimerkkitilanteessa.*

Soveltamalla Lagrangen funktiota ilman apumuuttujia y_i voidaan Lauseessa (4.1) esitelty optimointitehtävä kirjoittaa muodossa

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}), \tag{4.8}$$

jolloin KT-ehdot ovat

$$\frac{\partial L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial x_j} = 0, \qquad j = 1, 2, \ldots, n, \tag{4.9}$$

$$\frac{\partial L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial \lambda_i} \leq 0, \qquad i = 1, 2, \ldots, m, \tag{4.10}$$

$$\lambda_i^* \frac{\partial L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial \lambda_i} = 0, \qquad i = 1, 2, \ldots, m \tag{4.11}$$

ja

$$\lambda_i^* \geq 0, \qquad i = 1, 2, \ldots, m. \tag{4.12}$$

Useissa sovelluksissa erityisesti taloustieteissä vaaditaan, etteivät Lagrangen kertoimet saa olla negatiivisia. Nämä voitaisiin sisällyttää ehtojoukkoon $g_i(\mathbf{x}) \leq 0,\ i = 1, 2, \ldots, m$. Seuraavaksi kuitenkin esitetään, kuinka ei-negatiivisuutta vastaavat Lagrangen kertoimet voidaan eliminoida ja ehto kannattaa siksi käsitellä erikseen.

Minimoidaan funktio $f(\mathbf{x})$ ehdoilla $g_i(\mathbf{x}) \leq 0$, $i = 1, 2, \ldots, m$ ja $-x_j \leq 0$, $j = 1, 2, \ldots, n$. Kirjoitetaan aluksi Lagrangen funktio

$$L(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{w}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^{n} w_j(-x_j). \tag{4.13}$$

Tällöin Lauseen (4.1) perusteella KT-ehdot ovat

$$\frac{\partial L}{\partial x_j} = \frac{\partial f(\mathbf{x}^*)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i^* \frac{\partial g_i(\mathbf{x}^*)}{\partial x_j} - w_j^* = 0, \qquad j = 1, 2, \ldots, n \tag{4.14}$$

eli

$$\frac{\partial f(\mathbf{x}^*)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i^* \frac{\partial g_i(\mathbf{x}^*)}{\partial x_j} = w_j^*, \qquad j = 1, 2, \ldots, n. \tag{4.15}$$

Lisäksi

$$\frac{\partial L}{\partial \lambda_i} = g_i(\mathbf{x}^*) \leq 0, \qquad i = 1, 2, \ldots, m, \tag{4.16}$$

$$\lambda_i^* \frac{\partial L}{\partial \lambda_i} = \lambda_i^* g_i(\mathbf{x}^*) = 0, \qquad i = 1, 2, \ldots, m, \tag{4.17}$$

$$\frac{\partial L}{\partial w_j} = -x_j^* \leq 0, \qquad j = 1, 2, \ldots, n \tag{4.18}$$

ja

$$w_j^* \frac{\partial L}{\partial w_j} = w_j^*(-x_j^*) = 0, \qquad j = 1, 2, \ldots, n. \tag{4.19}$$

Yhtälön (4.15) mukaisesti voidaan kirjoittaa

$$x_j^* \left(\frac{\partial f(\mathbf{x}^*)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i^* \frac{\partial g_i(\mathbf{x}^*)}{\partial x_j} \right) = 0, \qquad j = 1, 2, \ldots, n, \tag{4.20}$$

$$\lambda_i^* \geq 0, \qquad i = 1, 2, \ldots, m, \tag{4.21}$$

$$w_j^* \geq 0, \qquad j = 1, 2, \ldots, n. \tag{4.22}$$

Lagrangen funktion (4.8) avulla KT-ehdot yhtälöistä (4.15)-(4.22) voidaan kirjoittaa muuttujien $\mathbf{x}$ ja $\boldsymbol{\lambda}$ avulla

$$\frac{\partial L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial x_j} \geq 0, \qquad j = 1, 2, \ldots, n, \tag{4.23}$$

$$x_j^* \frac{\partial L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial x_j} = 0, \qquad j = 1, 2, \ldots, n, \tag{4.24}$$

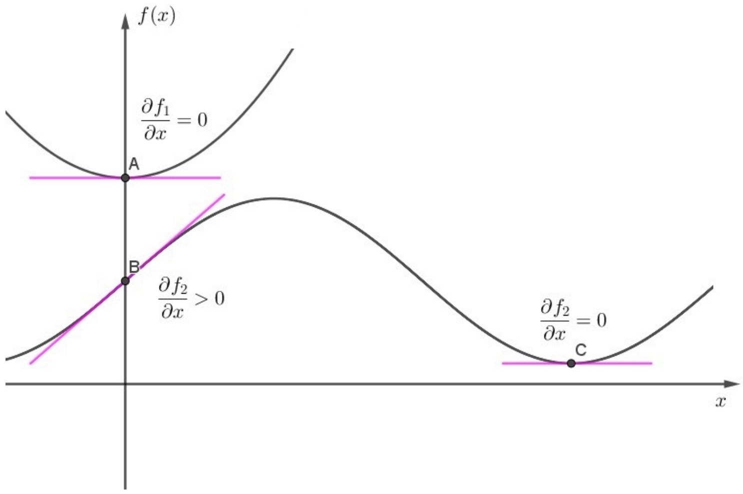
$$x_j^* \geq 0, \qquad j = 1, 2, \ldots, n, \tag{4.25}$$

$$\frac{\partial L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial \lambda_i} \leq 0, \qquad i = 1, 2, \ldots, m, \tag{4.26}$$

$$\lambda_i^* \frac{\partial L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial \lambda_i} = 0, \qquad i = 1, 2, \ldots, m, \tag{4.27}$$

$$\lambda_i^* \geq 0, \qquad i = 1, 2, \ldots, m. \tag{4.28}$$

Nyt siis ehto ei-negatiivisuudesta johtaa siihen, että KT-ehto (4.1) on korvattu kahdella ehdolla (4.23) ja (4.24). Tarkastellaan vielä KT-ehtojen perusteita graafisesti tämän havainnollistamiseksi. Kuten jo aiemmin mainittiin, KT-ehdot laajentavat Lagrangen menetelmän hyväksymällä myös epäyhtälörajoitteet. Tällöin otetaan huomioon se, että ääriarvokohta voi olla funktion määrittelyalueen reunapisteen lisäksi myös funktion määrittelyalueen sisäpiste. Laskennalliset vaatimukset ovat yleisesti sopivia pisteille, joissa ääriarvokohdan kaikki muuttujat, apumuuttujat y_i mukaan lukien, poikkeavat nollasta. [7] Tarkastellaan funktion $f(x)$ minimointia ehdolla $x \geq 0$ kuvassa 4.2, jossa on esitetty kaksi funktiota $f_1(x)$ ja $f_2(x)$.



Kuva 4.2. *KT-ehtojen graafiset perusteet funktioiden $f_1(x)$ ja $f_2(x)$ avulla.*

Tarkastellaan aluksi pistettä C, jossa muuttujan x arvoa voidaan suurentaa tai pienentää. Tiedetään, että pisteessä C täytyy olla $\frac{\partial f_2}{\partial x} = 0$, sillä muuttujan x arvon muuttaminen kasvattaisi funktion $f_2(x)$ arvoa, eikä pisteessä C voisi olla paikallista minimikohtaa.

Tarkastellaan seuraavaksi kahta vaihtoehtoa minimikohdalle, kun $x = 0$. Tiedetään, että pisteessä A voi olla paikallinen minimikohta, koska $\frac{\partial f_1}{\partial x} = 0$. Pisteessä B puolestaan voi olla minimikohta, koska $\frac{\partial f_2(x)}{\partial x} > 0$ eikä muuttujan x arvoa voida pienentää.

Yleistetään tämä vielä derivoituvalle funktiolle $f(x_1, x_2, \ldots, x_n)$. Funktion määrittelyalueen sisäpisteessä sijaitsevan ääriarvokohdan tulee toteuttaa yhtälö $\frac{\partial f}{\partial x_j} = 0,\ j = 1, 2, \ldots, n$. Sen sijaan funktion määrittelyalueen reunapisteessä sijaitsevan minimikohdan tulee toteuttaa epäyhtälö $\frac{\partial f}{\partial x_j} \geq 0,\ j = 1, 2, \ldots, n$. Minimikohdan graafista tarkastelua vastaavasti saadaan määrittelyalueen reunapisteessä sijaitsevalle maksimikohdalle ehtoepäyhtälö $\frac{\partial f}{\partial x_j} \leq 0,\ j = 1, 2, \ldots, n$. [7]

Esitellään vielä KT-ehtojen yhteys Lagrangen funktion satulapisteeseen. Käytetään apuna minimointitehtävää, jossa optimoitava funktio on $f(\mathbf{x})$ ja ehtoepäyhtälöt ovat $g_i(\mathbf{x}) \leq 0, i = 1, 2, \ldots, m$ ja $-x_j \leq 0, j = 1, 2, \ldots, n$. Tarkastellaan Lagrangen funktion $L(\mathbf{x}, \boldsymbol{\lambda})$ paikallista minimiä koskevia ehtoja aluksi muuttujien $\mathbf{x}$ suhteen. Kun $x_j \geq 0, j = 1, 2, \ldots, n$, saadaan ehdot

$$\frac{\partial L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial x_j} \geq 0, \qquad j = 1, 2, \ldots, n,$$

$$x_j^* \frac{\partial L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial x_j} = 0, \qquad j = 1, 2, \ldots, n$$

ja

$$\lambda_i^* \geq 0, \qquad i = 1, 2, \ldots, m.$$

jotka ovat KT-ehdot. Toisaalta Lagrangen funktion L paikalliselle maksimille muuttujien $\boldsymbol{\lambda}$ suhteen KT-ehdot ovat

$$\frac{\partial L(\mathbf{x}^*, \boldsymbol{\lambda}^*)}{\partial \lambda_i} \leq 0, \qquad i = 1, 2, \ldots, m,$$

$$\lambda_i^* \frac{\partial L(\mathbf{x}^*, \boldsymbol{\lambda})}{\partial \lambda_i} = 0, \qquad i = 1, 2, \ldots, m$$

ja

$$\lambda_i^* \geq 0, \qquad i = 1, 2, \ldots, m.$$

Näistä yhdessä seuraa Lause (4.2) Lagrangen funktion L satulapisteelle. [7]

Lause 4.2. *Jos Lagrangen funktion $L(\boldsymbol{x}, \boldsymbol{\lambda})$ satulapiste on $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$, niin piste $\mathbf{x}^*$ on ratkaisu optimointitehtävään, jossa minimoitava funktio on $f(\boldsymbol{x})$ ja ehtoepäyhtälöt*

$$g_i(\boldsymbol{x}) \leq 0, \qquad i = 1, 2, \ldots, m,$$

$$-x_j \leq 0, \qquad j = 1, 2, \ldots, n.$$

Lauseen (4.2) mukaisesti siis alkuperäisen optimointitehtävän ratkaisupiste $\mathbf{x}^*$ ja sitä vastaavat Lagrangen kertoimet $\boldsymbol{\lambda}^*$ muodostavat Lagrangen funktion satulapisteen $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$.

5. SOVELLUSKOHTEITA TALOUSTIETEISSÄ

Tässä luvussa esitellään muutamia Lagrangen menetelmän ja Kuhn-Tuckerin ehtojen sovelluskohteita taloustieteiden näkökulmasta, missä optimointi on tärkeää. Taloustieteissä käsitellään usein esimerkiksi optimaalista resurssien kohdistamista. Tällöin kohdetta, kuten myytyjen kappaleiden määrää, kuvaava funktio tulee optimoida resurssien asettamilla ehdoilla.

5.1 Tulosfunktion maksimointi

Lagrangen menetelmän soveltaminen taloustieteissä perustuu Lagrangen kertoimen λ tulkintoihin. Kerroin kuvaa, kuinka paljon funktion f optimaalinen arvo muuttuu pienillä ehtofunktion muutoksilla [4]. Taloustieteissä ehtofunktio g kuvaa usein rajoitettua resurssia, kuten budjettia tai työtä. Tällöin kerroin λ ilmaisee kyseisen resurssin varjohinnan. Varjohinta on kustannus, joka aiheutuu päämäärän saavuttamisesta ja jota verrataan saavutettavaan hyötyyn. [10]

Lagrangen menetelmä on sovellettavissa monipuolisesti taloustieteissä. Esimerkiksi G. C. Chow esittelee teoksessaan *Dynamic Economics: Optimization by the Lagrange Method* [3] menetelmän soveltamista mm. talouden kasvun –, tasapainon – ja kiertokulun arviointiin sekä sijoitusmalleihin. Tarkastellaan aluksi yksinkertaisen esimerkin kautta kuvitteellisen yrityksen liikevoiton maksimointia Lagrangen menetelmää soveltaen, kun käytettävissä oleva budjetti otetaan huomioon.

Esimerkki 5.1. Erotuomaripillejä valmistava yritys REF-Whistle on kehittänyt tulosfunktion

$$f(x, y) = 4x^2 + 3xy + 6y^2,$$

missä muuttuja x kuvaa tuhansien myytyjen pillien määrää ja muuttuja y kuvaa mainostukseen käytettyjä tunteja kuukaudessa. Pillien myyntiin käytettävää budjettia rajoittaa yhtälö

$$g(x, y) = x + y - 56 = 0.$$

Etsitään sellaiset arvot muuttujille x ja y, että yrityksen liikevoitto saadaan maksimoitua ottamalla budjetti huomioon.

Lasketaan aluksi funktioiden f ja g gradientit. Funktion f gradientti

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (8x + 3y, 3x + 12y)$$

ja funktion g gradientti

$$\nabla g(x, y) = \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y} \right) = (1, 1).$$

Lauseen 3.1 mukaisesti muodostetaan yhtälöryhmä

$$\begin{cases} \frac{\partial L}{\partial x} = f_x(x, y) + \lambda g_x(x, y) = 8x + 3y + \lambda = 0 \\ \frac{\partial L}{\partial y} = f_y(x, y) + \lambda g_y(x, y) = 3x + 12y + \lambda = 0 \\ \frac{\partial L}{\partial \lambda} = g(x, y) = x + y - 56 = 0. \end{cases}$$

Ratkaisemalla kyseinen yhtälöryhmä saadaan muuttujien x, y ja λ arvoiksi

$$\begin{cases} x = 36 \\ y = 20 \\ \lambda = -348. \end{cases}$$

Tulosfunktion f maksimiarvo saavutetaan, kun pillejä myydään 36 000 kpl ja mainostamiseen käytetään 20 tuntia kuukaudessa.

5.2 Kysyntäperusteinen hinnoittelu

Esitellään seuraavaksi Kuhn-Tuckerin menetelmän soveltamista hinnoittelun sovittamiseen kysynnän mukaiseksi. Nyt yrityksen tarjoaman tuotteen kysyntä siis riippuu kellonajasta niin, että tuottohuippujen aikana yrityksen kapasiteetti voidaan hyödyntää täysin, kun taas ajoittain kapasiteetista voidaan hyödyntää vain pieni osa (tuottovaje). Tällainen hinnoittelu on kuluttajalle tuttua mm. sähköyhtiöiden osalta.

Lauseessa (5.2) ja sen todistuksessa käytetään lähdettä [7]. Tässä vähimmäisvalmistuskustannuksilla tarkoitetaan kustannuksia, jotka tulevat tuotteen tekemisestä, kuten materiaalikustannuksista. Pääoman tuotolla tarkoitetaan yrityksen suhteellisesta kannattavuutta. Rajakustannus kuvaa puolestaan hyötyä, joka saadaan. Nettovoitolla tarkoitetaan myyntihinnan ja kustannusten erotusta, kun esimerkiksi veroja ei oteta huomioon.

Lause 5.2. *Olkoon yrityksen tuotto sellainen, että yrityksen hinnat tuottovajeen aikana kattavat ainoastaan vähimmäisvalmistuskustannukset, kun taas tuottohuippujen aikana yrityksen hinnat ylittävät ne. Vähimmäisvalmistuskustannukset ylittävän osuuden summa kaikilta tuottohuipuilta muodostaa pääoman tuoton eli summa vastaa kasvavia rajakustannuksia.*

Todistus. Olkoon kysyntä päivän 24 tunnin mukaisesti $x_1, x_2, \ldots, x_{24}$ ja kysyntää vastaava hinta $p_1, p_2, \ldots, p_{24}$. Oletetaan, että jokainen $x_i > 0,\ i = 1, 2, \ldots, 24$ eli kysyntää on päivän jokaisena tuntina. Olkoon tunnin tuottokapasiteetti y. Funktio $C(x_1, x_2, \ldots, x_{24})$ kuvaa päivän valmistuskuluja ja funktio $g(y)$ päivän pääoman tuottoa. Oletetaan, että vähimmäisvalmistuskustannukset $\frac{\partial C}{\partial x_i}$ ja rajakustannukset $\frac{dg}{dy}$ ovat positiivisia. Oletetaan lisäksi, ettei yrityksen tuotto vaikuta hintoihin eli $\frac{\partial p_i}{\partial x_i} = 0,\ i = 1, 2, \ldots, 24.$

Yrityksen tulee siis maksimoida päivän nettovoitto

$$E = \sum_{i=1}^{24} p_i x_i - C(x_1, x_2, \ldots, x_{24}) - g(y)$$

ehdoilla

$$x_i \leq y, \quad i = 1, 2, \ldots, 24,$$

$$x_i > 0, \quad i = 1, 2, \ldots, 24$$

ja

$$y \geq 0.$$

Tällöin Lagrangen funktio on muotoa

$$L(\mathbf{x}, y, \boldsymbol{\lambda}) = \sum_{i=1}^{24} p_i x_i - C(x_1, x_2, \ldots, x_{24}) - g(y) + \sum_{i=1}^{24} \lambda_i (y - x_i).$$

Nyt KT-ehdot ovat

$$\frac{\partial L}{\partial x_i} = p_i - \frac{\partial C}{\partial x_i} - \lambda_i \leq 0, \quad i = 1, 2, \ldots, 24, \tag{5.1}$$

$$x_i \frac{\partial L}{\partial x_i} = x_i \left(p_i - \frac{\partial C}{\partial x_i} - \lambda_i \right) = 0, \quad i = 1, 2, \ldots, 24, \tag{5.2}$$

$$\frac{\partial L}{\partial y} = -\frac{dg}{dy} + \sum_{i=1}^{24} \lambda_i \leq 0, \tag{5.3}$$

$$y \frac{\partial L}{\partial y} = y \left(-\frac{dg}{dy} + \sum_{i=1}^{24} \lambda_i \right) = 0, \tag{5.4}$$

$$\frac{\partial L}{\partial \lambda_i} = y - x_i \geq 0, \quad i = 1, 2, \ldots, 24, \tag{5.5}$$

$$\lambda_i \frac{\partial L}{\partial \lambda_i} = \lambda_i (y - x_i) = 0, \quad i = 1, 2, \ldots, 24, \tag{5.6}$$

$$\lambda_i \geq 0, \quad i = 1, 2, \ldots, 24. \tag{5.7}$$

Yhtälön (5.5) perusteella $y > 0$, koska oletettiin, että $x_i > 0$, $i = 1, 2, \ldots, 24$. Tämä tarkoittaa, ettei kapasiteetti voi olla nolla. Nyt tarkastellaan ainoastaan ratkaisuja, joissa kaikki x_i ja y ovat positiivisia, joten epäyhtälöt (5.2) ja (5.4) saadaan muotoon

$$p_i - \frac{\partial C}{\partial x_i} - \lambda_i = 0, \quad i = 1, 2, \ldots, 24 \tag{5.8}$$

ja

$$-\frac{\partial g}{\partial y} + \sum_{i=1}^{24} \lambda_i = 0. \tag{5.9}$$

Määritelmän mukaisesti jokaisella tuottovajeen ajanhetkellä t kapasiteetti on suurempi kuin kysyntä eli $y > x_t$. Tällöin yhtälön (5.6) perusteella tuottovajeen aikana $\lambda_t = 0$. Nyt yhtälö (5.8) on tuottovajeen aikana

$$p_t = \frac{\partial C}{\partial x_t}.$$

Tuottovajeen aikana tuotteen hinta kannattaa siis asettaa vastaamaan vähimmäisvalmistuskustannusta $\frac{\partial C}{\partial x_t}$. Koska kapasiteettia ei pystytä hyödyntämään täysin, kannattaa kysyntää edistää asettamalla hinta mahdollisimman alhaiseksi kuitenkin niin, ettei tästä aiheudu yritykselle tappiota.

Tuottohuippujen s aikana yrityksen kapasiteetti hyödynnetään täydellisesti, jolloin $x_s = y$. Oletuksen $\frac{dg}{dy} > 0$ perusteella yhtälö (5.9) saadaan muotoon

$$\frac{dg}{dy} = \sum_{s=1}^{24} \lambda_s > 0$$

eli ainakin osassa tuottohuipuista Lagrangen kertoimien λ tulee olla positiivisia. Nyt yhtälön (5.8) perusteella tuottohuippujen optimaaliseksi hinnoitteluksi saadaan

$$p_s = \frac{\partial C}{\partial x_s} + \lambda_s.$$

Tällöin hinta ylittää vähimmäisvalmistuskustannukset summalla, jota vastaa Lagrangen kerroin λ_s. Lisäksi yhtälön (5.9) perusteella tuottohuippujen valmistuskustannukset ylittävä summa vastaa täsmälleen rajakustannusta $\frac{\partial g}{\partial y}$. Koska tuottohuippu vastaa kapasiteettia, niin jokainen lisäys tuottohuippuun vaatii kapasiteetin lisäystä. Näin ollen sen täytyy kattaa pääoman tuotto $\frac{dg}{dy}$. □

5.3 Monopoliyrityksen sääntely

Tarkastellaan vielä monopoliasemassa olevan yrityksen sääntelyä käyttäen lähdettä [7]. Luonnollisia monopoleja syntyy, kun tavaroiden tuottaminen on halvempaa yhdellä yrityksellä kuin muilla. Tätä asemaa säädellään, jottei kyseisellä yrityksellä ole liikaa valtaa suhteessa asiakkaisiin. Lisäksi pyritään laadun ja luotettavuuden takaamiseen taloudellisesti tai poliittisesti halutuilla hinnoilla.

Monopolin tuoton maksimitasolla tuotto vastaa marginaalikustannuksia. Tämä tarkoittaa muutosta valmistuskustannuksissa, kun valmistetaan yksi tuote enemmän [8]. Tällä tuotolla hinnat ylittävät marginaalikustannukset. Monopoliyrityksen kannattavuus riippuu hinnan ja keskiarvokustannuksen välisestä suhteesta. Yksi lähestymistapa monopoliyrityksen sääntelyyn on sallia yritykselle hinnoittelu, joka ylittää keskiarvohinnan riittävästi mahdollistaen kohtuullisen tuoton.

Tarkastellaan yksinkertaista mallia, jossa monopoliyritys tuottaa yhtä tuotetta käyttäen pääomaa ja työtä. Merkitään näitä vastaavia määriä nyt q, x_1 ja x_2. Olkoon yrityksen valmistaman tuotteen yksikköhinta p. Oletetaan, että yritys voi ostaa pääomaa ja työtä rajoittamattomasti vakiohinnoilla c_1 ja c_2, jolloin yrityksen tuottofunktio

$$E = pq - c_1 x_1 - c_2 x_2. \tag{5.10}$$

Oletetaan, että $x_1, x_2 > 0$. Lisäksi tuloksen maksimointi vaatii, että

$$\frac{\partial E}{\partial x_1} = \frac{\partial pq}{\partial x_1} - c_1 = 0 \tag{5.11}$$

ja

$$\frac{\partial E}{\partial x_2} = \frac{\partial pq}{\partial x_2} - c_2 = 0, \tag{5.12}$$

jolloin

$$\frac{\frac{\partial pq}{\partial x_1}}{\frac{\partial pq}{\partial x_2}} = \frac{c_1}{c_2}. \tag{5.13}$$

Oletetaan nyt, että valtio rajoittaa yrityksen tuottoa kohtuulliselle tasolle. Tällöin jäljelle jäävän nettovoiton tulisi olla juuri riittävä kompensoimaan sijoitukset yrityksen välineistöön, kun yritys vähentää toiminnasta aiheutuvat kustannukset bruttotuloista. Jos tuoton astetta eli nettovoiton suhdetta välineistön arvoon pidetään liian suurena, niin yrityksen tulee laskea hintoja. Mikäli tämä suhde on liian matala, niin yrityksen on mahdollista nostaa hintojaan. Lauseessa (5.3) esiintyy termi rajasubstituutiosuhde. Se kuvaa suhdetta, jolla päätöksentekijä on valmis vaihtamaan hyödykkeitä keskenään [11].

Lause 5.3. *Nyt yritys ei suhteuta rajasubstituutiosuhdetta panoksen hintojen suhteeseen. Yrityksellä on sen arvon kasvattamiseen kannustin, jonka mukaan käytetyn pääoman määrä monopolin sääntelyehdolla ei saa olla vähemmän kuin ilman ehtoa olisi.*

Todistus. Olkoon yrityksen tuotantofunktio $q = f(x_1, x_2)$, missä $\frac{\partial f}{\partial x_1} > 0$, $\frac{\partial f}{\partial x_2} > 0$ ja

$$f(0, x_2) = f(x_1, 0) = 0.$$

Käänteinen kysyntäfunktio voidaan puolestaan kirjoittaa $p = p(q)$, missä $\frac{dp}{dq} < 0$. Tuottofunktio

$$E = pq - c_1 x_1 - c_2 x_2$$

esitettiin yhtälössä (5.10).

Olkoon x_1 välineistön lukumäärä perustasolla, b_1 sen hankintahinta perustasolla, β_1 välineistön arvonalenema ja B_1 jo kertynyt arvonalenema. Olkoon sääntelyehto

$$\frac{pq - c_2 x_2 - \beta_1}{b_1 x_1 - B_1} \leq s, \tag{5.14}$$

missä nettovoitto työkustannuksista ja pääoman arvonalenemasta määrää perustason prosenttiosuuden maksimin s. Oletetaan, että arvonalenemat β_1 ja B_1 ovat nollia ja hankintahinta b_1 on 1, jolloin perustason arvo vastaa fyysistä pääomaa. Hinta c_1 vastaa siis välineistön pitämisestä aiheutuvia kustannuksia. Tällöin sääntelyehto (5.14) voidaan kirjoittaa muodossa

$$\frac{pq - c_2 x_2}{x_1} \leq s \tag{5.15}$$

eli

$$pq - sx_1 - c_2x_2 \leq 0. \tag{5.16}$$

Nyt kohtuullista voittoa kuvaava s on enimmäismäärä, joka välineistöstä sallitaan monopoliyritykselle kompensoimaan sen pääomakustannuksia. Jos $s < c_1$, niin sallittu voitto on pienempi kuin todelliset pääomakustannukset ja yritys vetäytyisi markkinoilta. Siksi täytyy olettaa, että $s > c_1$ eli sallitun voiton tulee vähintään kattaa pääomakustannukset.

Yrityksen tulee siis maksoimoida tulosfunktio yhtälössä (5.10) ehdoilla (5.16) ja $x_1 > 0$, $x_2 > 0$. Nyt Lagrangen funktio on

$$L(x_1, x_2, \lambda) = p(q)q - c_1x_1 - c_2x_2 - \lambda\big(p(q)q - sx_1 - sx_2\big),$$

missä $q = f(x_1, x_2)$. KT-ehdot maksimille pisteessä $(x_1^*, x_2^*, \lambda^*)$ ovat

$$\frac{\partial L}{\partial x_1} = (1-\lambda)(p + p'q)f_{x_1} - c_1 + \lambda s \leq 0, \tag{5.17}$$

$$x_1\frac{\partial L}{\partial x_1} = x_1\big((1-\lambda)(p + p'q)f_{x_1} - c_1 + \lambda s\big) = 0, \tag{5.18}$$

$$\frac{\partial L}{\partial x_2} = (1-\lambda)(p + p'q)f_{x_2} - (1-\lambda)c_2 \leq 0, \tag{5.19}$$

$$x_2\frac{\partial L}{\partial x_2} = x_2\big((1-\lambda)(p + p'q)f_{x_2} - (1-\lambda)c_2\big) = 0, \tag{5.20}$$

$$\frac{\partial L}{\partial \lambda} = -\big(p(q)q - sx_1 - c_2x_2\big) \geq 0, \tag{5.21}$$

$$\lambda\frac{\partial L}{\partial \lambda} = \lambda\big(p(q)q - sx_1 - c_2x_2\big) = 0, \tag{5.22}$$

$$\lambda \geq 0. \tag{5.23}$$

Nyt kuitenkin $x_1^* > 0$ ja $x_2^* > 0$. Lisäksi oletetaan, että $\lambda^* > 0$, joten epäyhtälöt (5.17), (5.19) ja (5.21) voidaan kirjoittaa yhtälömuodossa

$$(1-\lambda)(p + p'q)f_{x_1} + \lambda s = c_1, \tag{5.24}$$

$$(p + p'q)f_{x_2} = c_2, \tag{5.25}$$

$$pq - sx_1 - c_2x_2 = 0. \tag{5.26}$$

Tässä $(p + p'q)$ kuvaa rajatuloja ja funktiot f_{x_1}, f_{x_2} kuvaavat vastaavasti pääoman ja työn tason muutosta. Huomataan, että yhtälöt (5.24)-(5.26) määrittävät ratkaisun x_1^*, x_2^* ja λ^*. Yhtälöstä (5.24) seuraa, että $\lambda^* > 0$ vääristää yhtäsuuruutta tuoton muutoksessa todellisen hinnan c_1 verran. Näin ollen yrityksen käyttämien pääoman ja työn suhteelliset osuudet muuttuvat ja suhde $\frac{f_{x_1}}{f_{x_2}}$ ei enää vastaa ostohintojen suhdetta. Näin ensimmäinen osuus Lauseesta (5.3) on osoitettu.

Yhtälöstä (5.11) nähdään, että oletuksen $\lambda > 0$ perusteella $c_1 = s$, kun $\lambda = 1$. Toisaalta jos $c_1 = s$, niin yhtälö (5.24) pelkistyy yhtälöksi (5.11), joka vastaa rajoittamattoman monopoliyrityksen toimintaa. Näin ollen tarkennetaan olettamusta $s \geq c_1$ olettamukseen $s > c_1$ ja saadaan $\lambda^* \neq 1$.

Nyt yläindeksi * ilmaisee optimointitehtävän ratkaisua rajoitetulle monopoliyritykselle, kun taas yläindeksi $^\circ$ rajoittamattomalle monopoliyritykselle. Lauseke $(p + p'q)f_{x_1}$ kuvaa pääoman marginaalituottoa ja lauseke $(p + p'q)f_{x_2}$ puolestaan työvoiman marginaalituottoa. Merkitään näitä nyt MR_1 ja MR_2 (eng. marginal revenue). Lisäämällä termi $c_1\lambda^*$ puolittain yhtälöön (5.24) ja järjestämällä termit saadaan rajoitetulle monopoliyritykselle

$$MR_1^* = c_1 - \frac{(s - c_1)}{1 - \lambda^*}\lambda^*. \tag{5.27}$$

Nyt oletuksien $s > c_1$ ja $\lambda < 1$ perusteella yhtälöstä (5.27) seuraa $MR_1^* < c_1$ eli pääoman marginaalituotto on rajoitetulle monopoliyritykselle pienempi kuin pääoman ostamisen vakiohinta. Oletetaan, että tuottofunktio

$$G = pf(x_1, x_2)$$

on ylöspäin kupera, jolloin MR_1 eli pääoman marginaalituotto ei kasva käytetyn pääoman x_1 funktiona. Tällöin käytetyn pääoman määrä sääntelyä noudattaen (x_1^*) ei ole vähemmän kuin ilman sääntelyä (x_1°). Koska nyt funktio G on ylöspäin kupera, niin

$$\frac{\partial MR_1}{\partial x_1} < 0,$$

jolloin $x_1^* > x_1^\circ$. Lisäksi yhtälöistä (5.24) ja (5.25) seuraa, että

$$\frac{MR_1}{MR_2} = \frac{c_1}{c_2} - \frac{(s - c_1)}{c_2}\frac{\lambda}{(1 - \lambda)} < \frac{c_1}{c_2}. \tag{5.28}$$

Nyt siis

$$\frac{MR_1}{MR_2} < \frac{c_1}{c_2}$$

eli pääoman ja työvoiman marginaalituottojen suhde on pienempi kuin niiden ostamisen vakiohintojen suhde. Tuotanto voidaan siis valmistaa suuremmalla pääomalla ja vähemmällä työllä verrattuna rajoittamattomaan optimiin. □

Tämä tunnetaan Averch-Johnsonin ilmiönä. Mikäli siis sääntelyn mahdollistama tulon taso on suurempi kuin pääomakustannus, mutta pienempi kuin tulon taso olisi ilman sääntelyä, niin yrityksen tulee korvata pääoma eri tuotannontekijällä. Tällöin yrityksen tuotannon tulee toimia siten, ettei kustannuksia ole minimoitu.

6. YHTEENVETO

Tässä kandidaatintyössä esiteltiin Lagrangen menetelmä, sen yleistys epäyhtälöehtoihin Kuhn-Tuckerin ehtojen avulla sekä näiden sovelluskohteita taloustieteiden näkökulmasta. Taloustieteiden lisäksi muita mahdollisia sovelluskohteita työssä esitetyille menetelmille ovat mm. epälineaarinen ohjelmointi, säätöteoria sekä sovellukset sähköjärjestelmissä, kuten virranjakelun rajoittaminen. Todettiin, että menetelmien merkittävimmät hyödyt saavutetaan optimointitehtävien ratkaisemisessa niin, ettei eksplisiittinen parametrisointi ehtofunktioiden suhteen ole tarpeellista.

Luvussa 2 esiteltiin työn menetelmissä tarvittavia usean muuttujan funktioiden matemaattisia taustatietoja. Lagrangen menetelmä esiteltiin aluksi yksinkertaisessa tapauksessa Lauseessa (3.1). Tällöin optimoitavassa funktiossa oli kaksi muuttujaa ja ehtoyhtälöitä oli yksi. Tämän jälkeen teoria laajennettiin Lauseen (3.2) yleiseen Lagrangen menetelmään, jolloin muuttujia oli n kappaletta. Tällöin ehtoyhtälöitä oli m kappaletta, missä $m \leq n - 1$. Tuloksena saatiin Lagrangen funktiota osittaiderivoimalla muodostettava yhtälöryhmä, jonka ratkaisu on myös kyseisen optimointitehtävän ratkaisu. Yhtälöryhmä perustui optimoitavan funktion ja ehtofunktioiden gradienttien yhdensuuntaisuuksiin ratkaisupisteessä sekä kyseisen ratkaisupisteen vaadittuun sijaintiin ehtofunktiolla. Lisäksi esitettiin Lagrangen menetelmän geometrinen tulkinta xy-tasossa.

Luvussa 4 esiteltiin Lagrangen menetelmän laajennus Kuhn-Tuckerin ehtoihin. Lagrangen menetelmän tavoin optimointitehtävä esitettiin Lagrangen funktiona, mutta ehtoina oli nyt epäyhtälöitä. Muodostetun Lagrangen funktion avulla asetettiin ratkaisupisteelle ehdot Lauseessa (4.1). Myös Kuhn-Tuckerin ehtojen geometrinen tulkinta esitettiin xy-tasossa. Lisäksi luvussa 4 esiteltiin optimoitavan funktion ratkaisupisteen välinen yhteys Lagrangen funktion satulapisteeseen.

Luvussa 5 esiteltiin muutamia Lagrangen menetelmän ja Kuhn-Tuckerin ehtojen sovelluskohteita taloustieteiden näkökulmasta. Lagrangen menetelmän osalta käytiin läpi yksinkertainen esimerkki kuvitteellisen yrityksen tulosfunktion optimoinnista budjetti huomioon ottaen. Kahdessa muussa esimerkissä käsiteltiin kysyntäperusteista hinnoittelua sekä monopoliyrityksen sääntelyä niin, että molemmissa optimoitiin jälleen yrityksen tulos. Näistä muodostettiin Kuhn-Tuckerin ehtoja soveltaen Lauseet (5.2) ja (5.3), jotka myös todistettiin.

LÄHTEET

[1] R. A. Adams. *Calculus: A Complete Course*. Pearson Education Canada, 2003.

[2] P. Bussotti. *On the Genesis of the Lagrange Multipliers*. Kesäkuu 2003. URL: https://doi.org/10.1023/A:1023952102705 (viitattu 05. 04. 2023).

[3] G. C. Chow. *Dynamic Economics: Optimization by the Lagrange Method*. New York: Oxford University Press, 1997.

[4] M. Herzing. *Using the Lagrangian Method to Solve Optimization Problems*. URL: http://www.ne.su.se/polopoly_fs/1.295573.1473167836!/menu/standard/file/Lagrangian%20method.pdf (viitattu 07. 02. 2023).

[5] T. H. Kjeldsen. *A Contextualized Historical Analysis of the Kuhn–Tucker Theorem in Nonlinear Programming: The Impact of World War II*. Marraskuu 2000. URL: https://www.sciencedirect.com/science/article/pii/S0315086000922894 (viitattu 05. 04. 2023).

[6] M. Laaksonen. *Usean muuttujan funktiot*. Opintomoniste, Tampereen yliopisto, 2021.

[7] M. Luptácik. *Mathematical Optimization and Economic Analysis*. Springer Science+Business Media, LLC, 2010.

[8] *Marginal Cost Meaning, Formula, and Examples*. 1. elokuuta 2022. URL: https://www.investopedia.com/terms/m/marginalcostofproduction.asp (viitattu 27. 03. 2023).

[9] *More on Lagrange multipliers*. 21. huhtikuuta 2015. URL: https://sboyles.github.io/teaching/ce377k/class23.pdf (viitattu 09. 04. 2023).

[10] K. Myllylä ja N. Korteslahti. *Matematiikan perusteet taloustieteilijöille 1b*. Opintomoniste, Oulun yliopisto, 2013.

[11] *Sanasto*. URL: https://www.core-econ.org/the-economy/book/fi/text/50-02-glossary.html (viitattu 27. 03. 2023).