Ensimmäisen kertaluvun optimaalisuusehto epädifferentioituvaan optimointiin

Otso Helminen

LuK-tutkielma
Huhtikuu 2024

Ohjaaja:
Dos. Yury Nikulin

MATEMATIIKAN JA TILASTOTIETEEN LAITOS

TURUN YLIOPISTO
Matematiikan ja tilastotieteen laitos

Otso Helminen: Ensimmäisen kertaluvun optimaalisuusehto epädifferentioituvaan optimointiin
LuK-tutkielma, 13 s.
Matematiikka
Huhtikuu 2024

Tutkielmassa tutkitaan, mitä ehtoja optimointitehtävän sallitujen pisteiden pitää toteuttaa, jotta ne voisivat olla optimaalisia ratkaisuja. Tutkielma rajoittuu epädifferentioituvaan optimointiin ja käsittelemään ensimmäisen kertaluvun ehtoja. Esitetään Fritz John -tyyppiset optimaalisuusehdot sekä yhden tavoitteen optimoinnin että monitavoiteoptimoinnin tehtävälle. Esitettävät optimaalisuusehdot ovat välttämättömät ratkaisupisteen optimaalisuudelle. Lukijan odotetaan ymmärtävän jatkuvuuden, differentioituvuuden ja konveksisuuden määritelmät.

Asiasanat: Lipschitz-jatkuvuus, alidifferentiaali, epädifferentioituva optimointi, optimaalisuusehto.

Sisällys

1 Johdanto

Matemaattisessa optimointitehtävässä rajoitefunktiot muodostavat ratkaisuavaruuden pisteistä sallituksi alueeksi kutsutun osajoukon. Optimointitehtävän mahdolliset ratkaisut löytyvät tästä sallitusta alueesta. Kun halutaan tutkia, onko sallitun alueen piste lisäksi optimaalinen, voidaan tämä selvittää optimaalisuusehdoilla. Jotta ratkaisupiste on optimaalinen, sen täytyy toteuttaa tietyt ehdot, jotka saadaan muodostettua kohde- ja rajoitefunktioiden avulla. Optimaalisuusehtoja on sekä välttämättömiä että riittäviä. Ideaalitilanteessa optimaalisuusehto olisi sekä välttämätön että riittävä, mutta vain harvassa käytännön ongelmassa tämä on mahdollista.[4] Tässä tutkielmassa esitellään välttämättömiä ehtoja. Lisäksi optimaalisuusehdot voidaan jakaa ensimmäisen ja toisen kertaluvun ehtoihin. Ensimmäisen kertaluvun ehdoilla voidaan rajata sallitun alueen pisteitä optimien löytämiseksi. Toisen kertaluvun ehdoissa oletetaan funktioiden olevan kahdesti jatkuvasti differentioituvia, jolloin pystytään ensimmäisen kertaluvun ehtojen muodostamaa optimiratkaisujoukkoa supistamaan uusilla ehdoilla.[1] Tässä tutkielmassa keskitytään ensimmäisen kertaluvun ehtoihin, joista käsitellään vain Fritz John -tyyppisiä ehtoja.

Optimaalisuusehdot on määritelty sekä differentioituvaan että epädifferentioituvaan optimointiin. Tämä tutkielma keskittyy pelkästään epädifferentioituvaan optimointiin eli tilanteeseen, jossa kohdefunktioiden ja rajoitteiden ei tarvitse olla jatkuvasti differentioituvia. Koska oletusta jatkuvasta differentioituvuudesta ei ole, oletetaan sen sijaan funktioiden olevan lokaalisti Lipschitz-jatkuvia.[1]

Luvussa 2 käsitellään aluksi Lipschitz-jatkuvuutta, jonka jälkeen siirrytään tutkimaan konveksisuutta. Lisäksi käydään läpi alidifferentiaalien määritelmät sekä muutama tulos näille. Luku 3 esittää alidifferentiaaleille kolme Lipschitz-jatkuvuudesta saatavaa suurempaa tulosta, joita tarvitaan myöhempiin todistuksiin. Luvussa 4 esitetään ensimmäisen kertaluvun Fritz John -tyyppinen optimaalisuusehto kahdelle epädifferentioituvan optimoinnin tehtävälle. Ensimmäinen tehtävä käsittelee yhden tavoitteen optimointia. Toinen on puolestaan monitavoiteoptimointitehtävä. Lisäksi jälkimmäisen ehto käsittelee Pareto-optimaalisuutta, jonka määritelmiä esitellään luvussa 4.

Kaikki tulokset ja määritelmät on esitetty n-ulotteisen euklidisen avaruuden funktioille $f : \mathbb{R}^n \to \mathbb{R}$. Yksinkertaisuuden ja havainnollistamisen vuoksi kaikki esimerkit kuitenkin käsittelevät funktioita, jotka ovat tyyppiä $f : \mathbb{R} \to \mathbb{R}$.

2 Peruskäsitteitä

2.1 Lipschitz-jatkuvuus

Käydään kertauksena läpi Lipschitz-jatkuvuuden määritelmä yleistettynä n-ulotteisen euklidisen avaruuden funktioille $f : \mathbb{R}^n \to \mathbb{R}$.

Määritelmä 2.1. Palloympäristöllä $B(\boldsymbol{x}, r)$ tarkoitetaan pisteen $\boldsymbol{x} \in \mathbb{R}^n$ *ympäröivää* joukkoa $\{\boldsymbol{y} \in \mathbb{R}^n \mid \|\boldsymbol{x} - \boldsymbol{y}\| < r\}$.

Määritelmä 2.2. Funktio $f : \mathbb{R}^n \to \mathbb{R}$ on lokaalisti Lipschitz-jatkuva pisteessä $\boldsymbol{x}^* \in \mathbb{R}^n$, jos on olemassa luvut $m > 0, \delta > 0$ siten, että

$$|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leq m\|\boldsymbol{x} - \boldsymbol{y}\|, \text{ kaikilla } \boldsymbol{x}, \boldsymbol{y} \in B(\boldsymbol{x}^*, \delta).$$

Tarkemmin Lipschitz-jatkuvuus tarkoittaa sitä, että funktion kasvua tai vähenevyyttä on rajoitettu jollain skalaarilla m, jota kutsutaan Lipschitz-vakioksi. Lipschitz-jatkuvuus on tärkeä ominaisuus, kun halutaan tutkia epädifferentioituvia funktioita. Epädifferentioituvan optimoinnin teoria pohjautuukin Lipschitz-jatkuvuuteen

Epädifferentioituvuudella tarkoitetaan, että funktiot eivät ole kaikkialla jatkuvasti differentioituvia. Oletuksena kuitenkin on, että tällaisissa epädifferentioituvuuskohdissa funktio on lokaalisti Lipschitz-jatkuva, jolloin tarvittavat ominaisuudet säilyvät.[1] Havainnollistetaan lokaalia Lipschitz-jatkuvuutta yksinkertaisella esimerkillä.

Esimerkki 2.1. Olkoon $f : \mathbb{R} \to \mathbb{R},\ f(x) = |x^3 + x|$. Funktio f on jatkuva muttei jatkuvasti differentioituva, sillä epädifferentioituvuuskohdassa $x = 0$ toispuoleiset derivaatat ovat erisuuret. Valitaan esimerkiksi $m = 2$ ja $\delta = \frac{1}{2}$, jolloin tarkastellaan palloympäristöä $\left(-\frac{1}{2}, \frac{1}{2}\right) \subset \mathbb{R}$. Huomautuksena $|x^3 - y^3| \leq |x - y|$, kun $x, y \in \left(-\frac{1}{2}, \frac{1}{2}\right)$. Nyt kolmioepäyhtälön avulla

$$\begin{aligned} |f(x) - f(y)| &= ||x^3 + x| - |y^3 + y|| \\ &\leq |x^3 + x - y^3 - y| \\ &= |(x^3 - y^3) + (x - y)| \\ &\leq |x^3 - y^3| + |x - y| \\ &\leq |x - y| + |x - y| \\ &= 2|x - y|, \text{ kaikilla } x, y \in \left(-\frac{1}{2}, \frac{1}{2}\right), \end{aligned}$$

mistä seuraa funktion f lokaali Lipschitz-jatkuvuus kohdassa $x = 0$.

2.2 Konveksisuudesta

Optimaalisuusehtojen ja epädifferentioituvan optimoinnin käsittelemiseksi tarvitaan ymmärrystä konvekseista funktioista sekä näiden ominaisuuksista. Esitetään tässä alaluvussa konveksisuuteen ja epädifferentioituvuuteen liittyviä määritelmiä ja lauseita havainnollistavine esimerkkeineen ja kuvineen. Käydään kuitenkin ensiksi läpi suunnattujen derivaattojen määritelmät.

Määritelmä 2.3. [3] Funktion $f : \mathbb{R}^n \to \mathbb{R}$ suunnattu derivaatta pisteessä $\boldsymbol{x}$ vektorin $\boldsymbol{v}$ suuntaisesti on

$$f'(\boldsymbol{x};\boldsymbol{v}) = \lim_{t\to 0^+} \frac{f(\boldsymbol{x}+t\boldsymbol{v}) - f(\boldsymbol{x})}{t}.$$

Määritellään seuraavaksi yleistetty suunnattu derivaatta. Lipschitz-jatkuvilla funktioilla ei välttämättä tarvitse olla olemassa määritelmän 2.3 mukaista suunnattua derivaattaa, minkä takia tarvitsee määritellä aina olemassa oleva vastine. Lipschitz-jatkuvilla funktioilla on aina yleistetty suunnattu derivaatta.[2]

Määritelmä 2.4. [5] Olkoon $f : \mathbb{R}^n \to \mathbb{R}$ lokaalisti Lipschitz-jatkuva pisteessä $\boldsymbol{x} \in \mathbb{R}^n$. Funktion f yleistetty suunnattu derivaatta pisteessä $\boldsymbol{x}$ vektorin $\boldsymbol{v} \in \mathbb{R}^n$ suuntaisesti määritellään seuraavasti:

$$f^\circ(\boldsymbol{x};\boldsymbol{v}) = \limsup_{\substack{\boldsymbol{y}\to\boldsymbol{x}\\ t\to 0^+}} \frac{f(\boldsymbol{y}+t\boldsymbol{v}) - f(\boldsymbol{y})}{t} = \inf_{\delta>0} \sup_{\|\boldsymbol{y}-\boldsymbol{x}\|<\delta,\, 0<t<\delta} \frac{f(\boldsymbol{y}+t\boldsymbol{v}) - f(\boldsymbol{y})}{t},$$

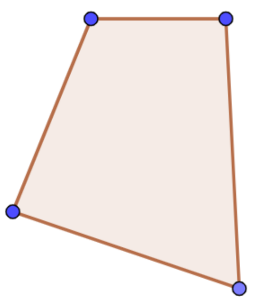
missä $\boldsymbol{y} \to \boldsymbol{x}$ tarkoittaa $\| \boldsymbol{y} - \boldsymbol{x} \| \to 0$.

Esitellään nyt konveksisuuteen liittyviä määritelmiä ja lauseita. Ensimmäisenä on tulos, joka yhdistää konveksisuuden ja lokaalin Lipschitz-jatkuvuuden.

Lause 2.5. *Olkoon funktio $f : \mathbb{R}^n \to \mathbb{R}$ konveksi. Tällöin se on myös lokaalisti Lipschitz-jatkuva kaikilla $\boldsymbol{x} \in \mathbb{R}^n$.*

Todistus. Ks. [6]. □

Joukon A peite tarkoittaa toista joukkoa tai joukkojen unionia $B = \bigcup B_i$, johon A kokonaisuudessaan sisältyy. Konveksipeite eroaa yleisestä peitteestä siten, että sen täytyy tietenkin olla konveksi ja lisäksi minimaalinen.



Kuva 1: Neljän pisteen joukon konveksipeite.

Määritelmä 2.6. [3] Joukon $E \subset \mathbb{R}^n$ konveksipeitteellä tarkoitetaan pienintä konveksia joukkoa, joka sisältää joukon E pisteet, ja siitä käytetään merkintää conv$E$. Joukon E konveksipeite on

$$\text{conv}E = \left\{ \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x_i}, \sum_{i=1}^{k} \lambda_i = 1,\ \ \lambda_i \geq 0,\ \ \boldsymbol{x_i} \in E,\ \ k = 1, 2, ... \right\},$$

missä $\sum_{i=1}^{k} \lambda_i \boldsymbol{x_i}$ on *konveksi kombinaatio* lukumäärästä k joukon E pisteitä.

Funktioille on mahdollista määritellä gradientin kaltainen ominaisuus riippumatta siitä, ovatko ne kaikkialla jatkuvasti differentioituvia. Epädifferentioituvuuskohdissa funktioilla ei ole yksikäsitteistä gradienttivektoria, joten määritellään sen sijaan aligradientin käsite. Aligradientti vastaa tavanomaista gradienttia, tai derivaattaa tapauksessa $f : \mathbb{R} \to \mathbb{R}$, kaikkialla, missä funktio on jatkuvasti differentioituva. Epädifferentioituvuuskohdissa aligradientteja voi kuitenkin olla ylinumeroituvasti ääretön määrä.

Määritelmä 2.7. [3] Olkoon E epätyhjä konveksi joukko ja $f : E \to \mathbb{R}$ konveksi funktio. Vektori $\boldsymbol{\xi} \in \mathbb{R}^n$ on funktion f aligradientti pisteessä $\boldsymbol{x}^*$, jos kaikille joukon E pisteille $\boldsymbol{x}$ pätee

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*) + \boldsymbol{\xi}^T(\boldsymbol{x} - \boldsymbol{x}^*).$$

Lisäksi konveksin funktion f alidifferentiaalilla tarkoitetaan pisteessä $\boldsymbol{x}^*$ määriteltyjen aligradienttien joukkoa. Tätä alidifferentiaalia merkitään

$$\partial_c f(\boldsymbol{x}^*) = \{\boldsymbol{\xi} \mid \boldsymbol{\xi} \text{ on funktion } f \text{ aligradientti pisteessä } \boldsymbol{x}^*\}.$$

Määritellään myöhemmin yleistetty funktion alidifferentiaali, missä funktion f ei oleteta olevan konveksi. Kun viitataan nimenomaan konveksin funktion alidifferentiaaliin, käytetään merkintää ∂_c. Yleistetyssä tapauksessa alaindeksi jätetään pois.

Esimerkki 2.2. Olkoon $f : \mathbb{R} \to \mathbb{R},\ f(x) = x^2$ ja $x^* = 2$. Nyt funktion f aligradientit pisteessä x^* toteuttavat epäyhtälön:

$$\begin{aligned} f(x) &\geq f(2) + \xi(x-2) \\ x^2 &\geq 4 + \xi x - 2\xi, \qquad \forall x \in \mathbb{R}. \end{aligned}$$

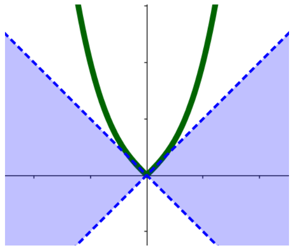
Luku $\xi = 4$ on ainoa luku, joka toteuttaa epäyhtälön. Funktion f ainoa aligradientti pisteessä $x^* = 2$ on täten 4. Samoin alidifferentiaali $\partial_c f(2) = \{4\}$. Huomatuksena funktio $f : \mathbb{R} \to \mathbb{R},\ f(x) = x^2$ on kaikkialla jatkuvasti differentioituva, joten aligradientti vastaa tavanomaista derivaattaa kaikilla $x \in \mathbb{R}$.

Seuraava lause yhdistää suunnatun derivaatan ja alidifferentiaalin. Käytännössä konveksin funktion f alidifferentiaali $\partial_c f$ voidaan määritellä suunnatun derivaatan avulla.

Lause 2.8. *Olkoon $f : \mathbb{R}^n \to \mathbb{R}$ konveksi. Tällöin jokaisessa pisteessä $\boldsymbol{x} \in \mathbb{R}^n$*

$$\partial_c f(\boldsymbol{x}) = \left\{\boldsymbol{\xi} \in \mathbb{R}^n \mid f'(\boldsymbol{x}; \boldsymbol{v}) \geq \boldsymbol{\xi}^T \boldsymbol{v},\ \text{kaikilla } \boldsymbol{v} \in \mathbb{R}^n\right\}.$$

Todistus. Ks. [2] s.14–15. □



Kuva 2: Tummanvihreällä näkyy esimerkin 2.1 funktio pisteen $x = 0$ ympäristössä. Tummansinisellä on piirretty vasemman- ja oikeanpuolisten derivaattojen tangentit pisteessä $x = 0$. Tangenttisuorien väliin jäävä sininen alue havainnollistaa alidifferentiaalia $\partial_c f(0)$.

Esimerkki 2.3. Tutkitaan esimerkin 2.1 tilannetta. Funktio ei ole derivoituva pisteessä $x = 0$, mutta suunnatut derivaatat voidaan kuitenkin laskea määritelmän 2.3 mukaisesti. Tässä tapauksessa yksiulotteinen vektori $v = \pm 1$. Suunnatut derivaatat ovat:

$$f'(0;1) = \lim_{t\to 0^+} \frac{f(0+t)-f(0)}{t} = \lim_{t\to 0^+} \frac{t^3+t}{t} = \lim_{t\to 0^+} (t^2+1) = 1,$$
$$f'(0;-1) = \lim_{t\to 0^+} \frac{f(0-t)-f(0)}{t} = \lim_{t\to 0^+} \frac{-(-t)^3-(-t)}{t} = \lim_{t\to 0^+} (t^2+1) = 1,$$

Suuntaan $v = -1$ tangentin kulmakerroin on 1. Tämä vastaa suuntaan $v = 1$ tangenttia, jonka kulmakerroin on -1. Nyt edellisen lauseen 2.8 mukaan

$$\partial_c f(0) = \{\xi \in \mathbb{R} \mid f'(0;v) \geq \xi v, \ \text{kaikilla } v \in \mathbb{R}\} = [-1,1]\,.$$

Kuten kuvassa 2 on havainnollistettu piirtämällä kaikki mahdolliset tangenttisuorat, alidifferentiaali kohdassa $x = 0$ sisältää kaikki luvut suljetulla välillä $[-1, 1]$. Tästä seuraa, että epädifferentioituvuuskohdissa aligradientteja voi olla useita.

Määritellään nyt yleistetyn alidifferentiaalin käsite samalla tavalla kuin lauseessa 2.8. Koska määritelmä tehdään yleistetyn suunnatun derivaatan avulla, tarvitaan lokaalia Lipschitz-jatkuvuutta.

Määritelmä 2.9. [2] Olkoon $f : \mathbb{R}^n \to \mathbb{R}$ lokaalisti Lipschitz-jatkuva pisteessä $\boldsymbol{x} \in \mathbb{R}^n$. Funktion f alidifferentiaali pisteessä $\boldsymbol{x}$ on tällöin

$$\partial f(\boldsymbol{x}) = \left\{\boldsymbol{\xi} \in \mathbb{R}^n \mid f^\circ(\boldsymbol{x};\boldsymbol{v}) \geq \boldsymbol{\xi}^T\boldsymbol{v}, \ \textit{kaikilla } \boldsymbol{v} \in \mathbb{R}^n\right\}.$$

Esitetään vielä alidifferentiaaleille muutama pienempi tulos, jotka auttavat myöhemmin todistamaan Lipschitz-jatkuvuudesta seuraavia ominaisuuksia.

Lause 2.10. *Konveksille funktiolle $f : \mathbb{R}^n \to \mathbb{R}$ pätee*

$$\partial_c f(\boldsymbol{x}) = \partial f(\boldsymbol{x}).$$

Todistus. Ks. [2] s.35–36. $\square$

Lause 2.11. *Olkoon f lokaalisti Lipschitz-jatkuva pisteessä $\boldsymbol{x} \in \mathbb{R}^n$. Tällöin kaikille $\lambda \in \mathbb{R}$*

$$\partial(\lambda f)(\boldsymbol{x}) = \lambda \partial f(\boldsymbol{x}).$$

Todistus. Ks. [2] s.38. □

Lause 2.12. *Olkoot $h : \mathbb{R}^n \to \mathbb{R}^m$, missä jokainen $h_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, ..., m$, on lokaalisti Lipschitz-jatkuva pisteessä $\boldsymbol{x}$ ja $g : \mathbb{R}^m \to \mathbb{R}$ lokaalisti Lipschitz-jatkuva pisteessä $h(\boldsymbol{x})$. Tällöin yhdistetty kuvaus $f : \mathbb{R}^n \to \mathbb{R}, f = g \circ h$, on lokaalisti Lipschitz-jatkuva pisteessä $\boldsymbol{x}$. Lisäksi,*

$$\partial f(\boldsymbol{x}) \subset \operatorname{conv}\left\{\sum_{i=1}^{m} a_i \boldsymbol{\xi}_i \mid \boldsymbol{\xi}_i \in \partial h_i(\boldsymbol{x}) \text{ ja } \boldsymbol{a} \in \partial g(h(\boldsymbol{x}))\right\}.$$

Todistus. Ks. [2] s.41–42. □

3 Alidifferentiaalien ominaisuuksia

Esitellään seuraavaksi muutama funktioiden lokaalista Lipschitz-jatkuvuudesta seuraava ominaisuus alidifferentiaaleille. Näitä tuloksia tullaan tarvitsemaan optimaalisuusehtojen todistuksissa. Ensimmäinen lause esittelee alidifferentiaalien subadditiivisuutta.

Lause 3.1. *Olkoot funktiot $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, ..., k$ lokaalisti Lipschitz-jatkuvia pisteessä $\boldsymbol{x} \in \mathbb{R}^n$. Tällöin painoille $w_i \in \mathbb{R}$ on voimassa:*

$$\partial\left(\sum_{i=1}^{k} w_i f_i\right)(\boldsymbol{x}) \subset \sum_{i=1}^{k} w_i \partial f_i(\boldsymbol{x}). \tag{1}$$

Joukot ovat samat, jos vähintään $k-1$ kappaletta funktioista f_i on jatkuvasti differentioituvia tai jos funktiot ovat konvekseja ja painot w_i ovat positiivisia.

Todistus. (Kuten [2] s.39) Riittää osoittaa, että kaava (1) pätee, kun $k = 2$. Yleinen tapaus voidaan osoittaa induktiolla.

$$\begin{aligned}
(f_1 + f_2)^\circ(\boldsymbol{x}; \boldsymbol{v}) &= \limsup_{\substack{\boldsymbol{y} \to \boldsymbol{x} \\ t \to 0^+}} \frac{(f_1 + f_2)(\boldsymbol{y} + t\boldsymbol{v}) - (f_1 + f_2)(\boldsymbol{y})}{t} \\
&= \limsup_{\substack{\boldsymbol{y} \to \boldsymbol{x} \\ t \to 0^+}} \frac{f_1(\boldsymbol{y} + t\boldsymbol{v}) + f_2(\boldsymbol{y} + t\boldsymbol{v}) - f_1(\boldsymbol{y}) - f_2(\boldsymbol{y})}{t} \\
&\leq \limsup_{\substack{\boldsymbol{y} \to \boldsymbol{x} \\ t \to 0^+}} \frac{f_1(\boldsymbol{y} + t\boldsymbol{v}) - f_1(\boldsymbol{y})}{t} + \limsup_{\substack{\boldsymbol{y} \to \boldsymbol{x} \\ t \to 0^+}} \frac{f_2(\boldsymbol{y} + t\boldsymbol{v}) - f_2(\boldsymbol{y})}{t} \\
&= f_1^\circ(\boldsymbol{x}; \boldsymbol{v}) + f_2^\circ(\boldsymbol{x}; \boldsymbol{v}).
\end{aligned}$$

Tästä seuraa

$$\partial(f_1 + f_2)(\boldsymbol{x}) \subset \partial f_1(\boldsymbol{x}) + \partial f_2(\boldsymbol{x}).$$

Nyt lauseen 2.11 nojalla

$$\partial(w_1 f_1 + w_2 f_2)(\boldsymbol{x}) \subset \partial(w_1 f_1)(\boldsymbol{x}) + \partial(w_2 f_2)(\boldsymbol{x}) = w_1 \partial f_1(\boldsymbol{x}) + w_2 \partial f_2(\boldsymbol{x}).$$

Todistuksen loppuosa seuraa funktioiden f_i regulaarisuudesta. Ehdot funktion regulaarisuudelle on määritelty [2] s.37. □

Lause 3.2. *Olkoot funktiot $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, ..., k$ lokaalisti Lipschitz-jatkuvia pisteessä $\boldsymbol{x} \in \mathbb{R}^n$. Tällöin funktio $f : \mathbb{R}^n \to \mathbb{R}$,*

$$f(\boldsymbol{x}) = \max_{i=1,...,k} f_i(\boldsymbol{x}),$$

on myös lokaalisti Lipschitz-jatkuva pisteessä $\boldsymbol{x}$. Lisäksi

$$\partial f(\boldsymbol{x}) \subset \operatorname{conv}\{\partial f_i(\boldsymbol{x}) \mid i \in I(\boldsymbol{x})\},$$

missä $I(\boldsymbol{x}) \subset \{1, ..., k\}$ tarkoittaa indeksijoukkoa, jolle $f(\boldsymbol{x}) = f_i(\boldsymbol{x})$.

Todistus. (Kuten [2] s.47–49) Koska jokainen funktio f_i on lokaalisti Lipschitz-jatkuva pisteessä $\boldsymbol{x}$, on selvästi myös f lokaalisti Lipschitz-jatkuva pisteessä $\boldsymbol{x}$. Määritellään lauseen toisen osion todistusta varten kaksi apufunktiota:

$$\begin{aligned} &g : \mathbb{R}^k \to \mathbb{R}, \quad g(a_1, a_2, ..., a_k) = \max_{i=1,...,k}\{a_i\} \text{ ja} \\ &h : \mathbb{R}^n \to \mathbb{R}^k, \;\; h(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), ..., f_k(\boldsymbol{x})). \end{aligned}$$

Selvästi $f = g \circ h$. Kaikille $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^k$ ja $\lambda \in [0, 1]$ on voimassa

$$\begin{aligned} g(\lambda \boldsymbol{u} + (1-\lambda)\boldsymbol{v}) &= \max_{i=1,...,k}\{\lambda u_i + (1-\lambda) v_i\} \\ &\le \lambda \max_{i=1,...,k}\{u_i\} + (1-\lambda)\max_{1,...,k}\{v_i\} \\ &= \lambda g(\boldsymbol{u}) + (1-\lambda) g(\boldsymbol{v}). \end{aligned}$$

Funktio g on täten konveksi ja lauseen 2.5 mukaan lokaalisti Lipschitz-jatkuva pisteessä $h(\boldsymbol{x})$. Olkoon joukko $J(\boldsymbol{u}) = \{i \in \{1, ..., k\} \mid u_i = g(\boldsymbol{u})\}$. Nyt funktion g suunnattu derivaatta on muotoa:

$$\begin{aligned} g'(\boldsymbol{u}; \boldsymbol{v}) &= \lim_{t\to 0^+} \frac{g(\boldsymbol{u} + t\boldsymbol{v}) - g(\boldsymbol{u})}{t} = \lim_{t\to 0^+} \max_{i=1,...,k} \frac{\{u_i + tv_i\} - g(\boldsymbol{u})}{t} \\ &= \lim_{t\to 0^+} \max_{i\in J(\boldsymbol{u})} \frac{\{u_i + tv_i\} - g(\boldsymbol{u})}{t} = \lim_{t\to 0^+} \max_{i\in J(\boldsymbol{u})} \frac{\{u_i + tv_i - u_i\}}{t}, \end{aligned}$$

eli

$$g'(\boldsymbol{u}; \boldsymbol{v}) = \max_{i\in J(\boldsymbol{u})} v_i.$$

Lauseiden 2.8 ja 2.10 nojalla

$$\partial g(\boldsymbol{u}) = \{\boldsymbol{a} \in \mathbb{R}^k \mid \max_{i\in J(\boldsymbol{u})} v_i \ge \boldsymbol{a}^T \boldsymbol{v}, \;\; \forall \boldsymbol{v} \in \mathbb{R}^k\},$$

mistä voidaan todeta, että

$$\boldsymbol{a} \in \partial g(\boldsymbol{u}) \iff \begin{cases} a_i \geq 0, & i = 1, ..., k, \\ \sum_{i=1}^{k} a_i = 1, & \\ a_i = 0, & \text{kun } i \notin J(\boldsymbol{u}). \end{cases}$$

Funktion g alidifferentiaali kohdassa $h(\boldsymbol{x}) \in \mathbb{R}^k$ on tällöin

$$\partial g(h(\boldsymbol{x})) = \{\boldsymbol{a} \in \mathbb{R}^k \mid a_i \geq 0,\ \sum_{i=1}^{k} a_i = 1,\ a_i = 0,\ \text{jos } i \notin J(\boldsymbol{u})\}.$$

Lauseen 2.12 nojalla

$$\begin{aligned} \partial(g \circ h)(\boldsymbol{x}) = \partial f(\boldsymbol{x}) \subset \operatorname{conv} & \left\{ \sum_{i=1}^{k} a_i \boldsymbol{\xi}_i \mid \boldsymbol{\xi}_i \in \partial h_i \quad \text{ja} \quad \boldsymbol{a} \in \partial g(h(\boldsymbol{x})) \right\} \\ = \operatorname{conv} & \left\{ \sum_{i \in I(\boldsymbol{x})} a_i \partial f_i(\boldsymbol{x}) \mid a_i \geq 0 \quad \text{ja} \quad \sum_{i \in I(\boldsymbol{x})} a_i = 1 \right\} \\ = \operatorname{conv} & \left\{ \partial f_i(\boldsymbol{x}) \mid i \in I(\boldsymbol{x}) \right\}. \end{aligned}$$

□

Seuraava tulos näyttää nollavektorin sisältyvän lokaalisti Lipschitz-jatkuvan funktion minimikohdan alidifferentiaaliin. Tulos toimii ehtona sille, onko piste yhden tavoitteen optimointitehtävän minimi, jos kohdefunktio on konveksi.

Lause 3.3. *Olkoon $f : \mathbb{R}^n \to \mathbb{R}$ lokaalisti Lipschitz-jatkuva pisteessä $\boldsymbol{x} \in \mathbb{R}^n$, ja se saavuttaa lokaalin miniminsä pisteessä $\boldsymbol{x}$. Silloin*

$$\boldsymbol{0} \in \partial f(\boldsymbol{x}). \tag{2}$$

Jos f on konveksi, on ehto (2) riittävä minimikohdan $\boldsymbol{x}$ globaalisuuteen.

Todistus. (Kuten [2] s.38–39) Koska $\boldsymbol{x}$ on funktion f lokaali minimi, on olemassa $\epsilon > 0$ siten, että $f(\boldsymbol{x} + t\boldsymbol{v}) \geq f(\boldsymbol{x})$, kaikilla $t \in (0, \epsilon)$ ja $\boldsymbol{v} \in \mathbb{R}^n$. Nyt

$$f^\circ(\boldsymbol{x}; \boldsymbol{v}) = \limsup_{\substack{\boldsymbol{y} \to \boldsymbol{x} \\ t \to 0^+}} \frac{f(\boldsymbol{y} + t\boldsymbol{v}) - f(\boldsymbol{y})}{t} \geq \limsup_{t \to 0^+} \frac{f(\boldsymbol{x} + t\boldsymbol{v}) - f(\boldsymbol{x})}{t} \geq 0,$$

eli kaikille $\boldsymbol{v} \in \mathbb{R}^n$

$$f^\circ(\boldsymbol{x}; \boldsymbol{v}) \geq 0 = \boldsymbol{0}^T \boldsymbol{v}.$$

Alidifferentiaalin määritelmän nojalla $\boldsymbol{0} \in \partial f(\boldsymbol{x})$. □

Esimerkissä 2.3 kulmakerroin 0 sisältyi käsiteltävään alidifferentiaaliin ja kohta $x = 0$ oli selvästi lokaali minimikohta. Lisäksi esimerkin funktio $f(x) = |x^3 + x|$ on konveksi, joten minimi oli globaali.

4 Optimaalisuusehtoja

4.1 Yhden tavoitteen optimointia

Tarvittavat tulokset itse ensimmäisen kertaluvun optimaalisuusehtojen käsittelemiseen on nyt esitetty. Ensimmäistä kertalukua olevia optimaalisuusehtoja on useammantyyppisiä, mutta keskitytään tässä tutkielmassa vain ensimmäisen kertaluvun Fritz John -tyyppisiin ehtoihin. Fritz Johnin kehittämät optimaalisuusehdot ovat välttämättömiä ratkaisupisteen optimaalisuudelle mutta eivät kuitenkaan aina riittäviä.[4] Tarkastellaan ensin yhden tavoitteen optimoinnin ongelmaa ja tämän ratkaisupisteitä, minkä jälkeen siirrytään monitavoiteoptimointiin. Seuraavana esitellään ensimmäinen välttämätön ehto ratkaisupisteen optimaalisuudelle.

Lause 4.1. (Fritz Johnin välttämätön optimaalisuusehto) *Välttämätön ehto sille, että piste $\boldsymbol{x}^* \in \mathbb{R}^n$ olisi lokaali minimi optimointitehtävälle*

$$\begin{aligned} &\text{min.} \quad f(\boldsymbol{x}) \\ &\text{s.t.} \quad\; g(\boldsymbol{x}) = (g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), ..., g_m(\boldsymbol{x}))^T \leq \mathbf{0}, \end{aligned}$$

missä kohdefunktio $f : \mathbb{R}^n \to \mathbb{R}$ sekä rajoitteet $g_j : \mathbb{R}^n \to \mathbb{R}$, $j = 1, ..., m$, ovat lokaalisti Lipschitz-jatkuvia pisteessä $\boldsymbol{x}^*$, on, että on olemassa luku $\lambda \in \mathbb{R}$ ja vektori $\boldsymbol{\mu} \in \mathbb{R}^m$, joille $\lambda \geq 0, \boldsymbol{\mu} \geq \mathbf{0}$ ja $(\lambda, \boldsymbol{\mu}) \neq (0, \mathbf{0})$, ja seuraavat ovat voimassa:

$$\begin{aligned} &\text{(i)} \quad \mathbf{0} \in \lambda \partial f(\boldsymbol{x}^*) + \sum_{j=1}^{m} \mu_j \partial g_j(\boldsymbol{x}^*) \\ &\text{(ii)} \quad \mu_j g_j(\boldsymbol{x}^*) = 0, \quad \textit{kaikilla } j = 1, ..., m. \end{aligned}$$

Todistus. Ks. [4] □

4.2 Pareto-optimaalisuus

Todistetaan tässä alaluvussa lauseen 4.1 kaltainen ehto Pareto-optimaalisille pisteille. Pareto-optimaalisuus tarkoittaa tilannetta, jossa optimointitehtävän ratkaisua ei voida parantaa jonkin ominaisuuden suhteen huonontamatta samalla toista. Ilmiötä esiintyy varsinkin monitavoiteoptimoinnissa, kun jokin kohdefunktio voisi saada parempia arvoja mutta samalla yhden tai useamman muun kohdefunktion arvot huononevat. Ratkaisupisteitä voi siten olla useita, ja niiden vertailu on mahdollisesti haastavaa. Esitetään vielä formaali määritelmä Pareto-optimaalisuudelle.

Määritelmä 4.2. [1] Olkoon $E \subset \mathbb{R}^n$ optimointitehtävän sallittujen ratkaisupisteiden joukko ja minimoitavat kohdefunktiot $f_i : E \to \mathbb{R}$. Ratkaisupiste $\boldsymbol{x}^* \in E$ on Pareto-optimaalinen, jos ei ole olemassa toista pistettä $\boldsymbol{x} \in E$, jolle $f_i(\boldsymbol{x}) \leq f_i(\boldsymbol{x}^*)$, kaikilla $i = 1, ..., k$, ja $f_j(\boldsymbol{x}) < f_j(\boldsymbol{x}^*)$, jollain $j \in \{1, ..., k\}$.

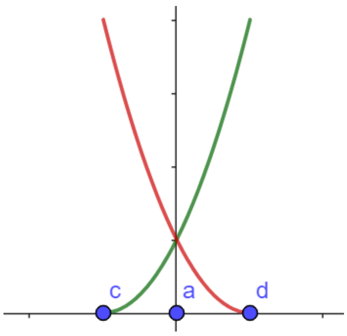
Lisäksi voidaan määritellä Pareto-optimaalisia pisteitä laajemmat ja suppeammat joukot: heikosti ja vahvasti Pareto-optimaaliset pisteet. Heikoksi Pareto-optimaaliksi kutsutaan pistettä, jossa ei voida parantaa kaikkia ominaisuuksia. Vahva Pareto-optimaali on taas piste, jossa on olemassa vähintään yksi pari ominaisuuksia, joista toista parannettaessa toinen heikentyy. Tämä vaihtokauppa ominaisuuksien välillä on kuitenkin rajoitettu, mikä tarkoittaa, että äärelliset parannukset toisessa ominaisuudessa aiheuttavat kohtuullista heikentymistä toisessa. Vahvasti Pareto-optimaaliset pisteet sisältyvät Pareto-optimaalisten pisteiden joukkoon ja Pareto-optimaaliset pisteet heikosti Pareto-optimaalisten pisteiden joukkoon.[1]

Määritelmä 4.3. [1] Olkoon $E \subset \mathbb{R}^n$ optimointitehtävän sallittujen ratkaisupisteiden joukko ja minimoitavat kohdefunktiot $f_i : E \to \mathbb{R}$. Piste $\boldsymbol{x}^* \in E$ on heikosti Pareto-optimaalinen, jos ei ole olemassa toista sallittua ratkaisupistettä $\boldsymbol{x} \in E$, jolle $f_i(\boldsymbol{x}) < f_i(\boldsymbol{x}^*)$, kaikilla $i = 1, ..., k$.

Määritelmä 4.4. [1] (Geoffrion, 1968) Olkoon $E \subset \mathbb{R}^n$ optimointitehtävän sallittujen ratkaisupisteiden joukko ja minimoitavat kohdefunktiot $f_i : E \to \mathbb{R}$. Piste $\boldsymbol{x}^* \in E$ on vahvasti Pareto-optimaalinen, jos se on Pareto-optimaalinen ja on olemassa reaaliluku $M > 0$ siten, että jokaiselle f_i ja $\boldsymbol{x} \in E$, joille $f_i(\boldsymbol{x}) < f_i(\boldsymbol{x}^*)$, on olemassa ainakin yksi f_j, jolle $f_j(\boldsymbol{x}^*) < f_j(\boldsymbol{x})$, ja

$$\frac{f_i(\boldsymbol{x}^*) - f_i(\boldsymbol{x})}{f_j(\boldsymbol{x}) - f_j(\boldsymbol{x}^*)} \leq M.$$

Havainnollistetaan vahvaa Pareto-optimaalisuutta esimerkillä.



Kuva 3: Kuvassa on vihreällä esimerkin 4.1 funktio f_1 ja punaisella funktio f_2

Esimerkki 4.1. Tarkastellaan minimointitehtävää. Kuvassa 3 on esitetty kohdefunktion f_1 kuvaaja vihreällä ja kohdefunktion f_2 kuvaaja punaisella rajoitettuna välille $[c, d]$. Piste a on vahvasti Pareto-optimaalinen, sillä jos pisteestä a siirrytään funktion f_1 suhteen parempaan suuntaan, funktio f_2 saa suurempia arvoja eli kaikille pisteille $x \in [c, d]$, joille $f_1(x) < f_1(a)$, on voimassa $f_2(a) < f_2(x)$. Samoin kaikille pisteille $x' \in [c, d]$, joille $f_2(x') < f_2(a)$, on voimassa $f_1(a) < f_1(x')$. Lisäksi symmetrian nojalla $\frac{f_1(a)-f_1(x)}{f_2(x)-f_2(a)} \leq M \geq \frac{f_2(a)-f_2(x')}{f_1(x')-f_1(a)}$, missä $M = f_1(c) + f_2(d)$.

Tarkastellaan yleistä tapausta epädifferentioituville kohde- ja rajoitusfunktioille. Seuraava lause määrää monitavoiteoptimointitehtävän ratkaisupisteelle ehdon Pareto-optimaalisuudelle.

Lause 4.5. (Fritz Johnin välttämätön Pareto-optimaalisuusehto) *Välttämätön ehto sille, että piste $\boldsymbol{x}^* \in S$ olisi Pareto-optimaalinen optimointitehtävälle*

$$\begin{aligned} &\text{min.} \quad \{f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), ..., f_k(\boldsymbol{x})\} \\ &\text{s.t.} \quad\ \boldsymbol{x} \in S = \{\boldsymbol{x} \in \mathbb{R}^n \mid g(\boldsymbol{x}) = (g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), ..., g_m(\boldsymbol{x}))^T \leq \mathbf{0}\}, \end{aligned}$$

missä kohdefunktiot $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, ..., k$, ja rajoitteet $g_j : \mathbb{R}^n \to \mathbb{R}$, $j = 1, ..., m$, ovat lokaalisti Lipschitz-jatkuvia pisteessä $\boldsymbol{x}^*$, on, että on olemassa vektorit $\boldsymbol{\lambda} \in \mathbb{R}^k$ ja $\boldsymbol{\mu} \in \mathbb{R}^m$, joille $\boldsymbol{\lambda} \geq \mathbf{0}, \boldsymbol{\mu} \geq \mathbf{0}$ ja $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \neq (\mathbf{0}, \mathbf{0})$, ja seuraavat ehdot ovat voimassa:

$$\begin{aligned} &\text{(A)} \quad \mathbf{0} \in \sum_{i=1}^{k} \lambda_i \partial f_i(\boldsymbol{x}^*) + \sum_{j=1}^{m} \mu_j \partial g_j(\boldsymbol{x}^*), \\ &\text{(B)} \quad \mu_j g_j(\boldsymbol{x}^*) = 0, \quad \textit{kaikilla } j = 1, ..., m. \end{aligned}$$

Todistus. (Kuten [1] s.47–48) Oletuksen $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \neq (\mathbf{0}, \mathbf{0})$ nojalla voidaan asettaa $\sum_{i=1}^{k} \lambda_i + \sum_{j=1}^{m} \mu_j = 1$. Olkoon $\boldsymbol{x}^*$ Pareto-optimaalinen. Määritellään todistuksen avuksi funktio

$$F : \mathbb{R}^n \to \mathbb{R}, \ F(\boldsymbol{x}) = \max\{f_i(\boldsymbol{x}) - f_i(\boldsymbol{x}^*), g_j(\boldsymbol{x}) \mid i = 1, ..., k, \ j = 1, ..., m\}.$$

Osoitetaan, että kaikille $\boldsymbol{x} \in \mathbb{R}^n$ pätee

$$F(\boldsymbol{x}) \geq 0. \tag{3}$$

Tehdään vastaoletus, että jollekin $\boldsymbol{x}' \in \mathbb{R}^n$, $F(\boldsymbol{x}') < 0$. Nyt $g_j(\boldsymbol{x}') < 0$, kaikilla $j = 1, ..., m$ ja piste $\boldsymbol{x}'$ on siten sallittu. Toisaalta $f_i(\boldsymbol{x}') - f_i(\boldsymbol{x}^*) < 0 \iff f_i(\boldsymbol{x}') < f_i(\boldsymbol{x}^*)$, kaikilla $i = 1, ..., k$, mikä on ristiriidassa pisteen $\boldsymbol{x}^*$ Pareto-optimaalisuuden kanssa. (3) on siis välttämättä tosi.
Koska $\boldsymbol{x}^*$ on sallittu piste, on $g(\boldsymbol{x}^*) < \mathbf{0}$, mistä seuraa, että $F(\boldsymbol{x}^*) = 0$. Edellisen ja välituloksen (3) avulla voidaan päätellä, että F saavuttaa minimikohtansa pisteessä $\boldsymbol{x}^*$. Koska kaikki funktiot f_i ja g_j ovat lokaalisti Lipschitz-jatkuvia, on lauseen 3.2 perusteella myös F lokaalisti Lipschitz-jatkuva. Nyt lauseen 3.3 nojalla

$$\mathbf{0} \in \partial F(\boldsymbol{x}^*).$$

Lisäksi, koska $f_i(\boldsymbol{x}^*)$ on vakio,

$$\partial(f_i(\boldsymbol{x}) - f_i(\boldsymbol{x}^*)) = \partial(f_i(\boldsymbol{x})). \tag{4}$$

Olkoon joukko $J(\boldsymbol{x}^*) = \{j = 1, ..., m \mid F(\boldsymbol{x}^*) = g_j(\boldsymbol{x}^*)\}$. Lauseen 3.2 nojalla ja kaavan (4) avulla saadaan

$$\mathbf{0} \in \operatorname{conv}\{\partial f(\boldsymbol{x}^*), \partial g_j(\boldsymbol{x}^*) \mid i = 1, ..., k, \ j \in J(\boldsymbol{x}^*)\}.$$

Konveksipeitteen määritelmän mukaan on olemassa reaalialkioiset vektorit $\boldsymbol{\lambda}$ ja $\boldsymbol{\mu}$, joille $\lambda_i \geq 0$, kaikilla $i = 1, ..., k$, $\mu_j \geq 0$, kaikilla $j \in J(\boldsymbol{x}^*)$, ja

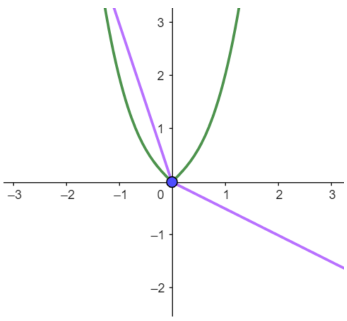
$$\sum_{i=1}^{k} \lambda_i + \sum_{j \in J(\boldsymbol{x}^*)} \mu_j = 1.$$

Vektoreille $\boldsymbol{\lambda}$ ja $\boldsymbol{\mu}$ on voimassa

$$\mathbf{0} \in \sum_{i=1}^{k} \lambda_i \partial f_i(\boldsymbol{x}^*) + \sum_{j \in J(\boldsymbol{x}^*)} \mu_j \partial g_j(\boldsymbol{x}^*).$$

Nyt lauseen ehto (A) saadaan, kun asetetaan $\mu_j = 0$, kun $j \in \{1, ..., m\} \setminus J(\boldsymbol{x}^*)$. Jos taas $g_j(\boldsymbol{x}^*) \neq 0 \Rightarrow g_j(\boldsymbol{x}^*) < 0$, jollekin j, niin $j \in \{1, ..., m\} \setminus J(\boldsymbol{x}^*)$. Tästä seuraa $\mu_j = 0$, jolloin saadaan ehto (B). $\square$

Käytetään lausetta 4.5 esimerkissä.



Kuva 4: Kuvaan on piirretty vihreällä esimerkin 4.2 funktio $f_1(x)$ ja violetilla funktio $f_2(x)$ kohdan $x = 0$ ympäristössä.

Esimerkki 4.2. Tarkastellaan tilannetta, jossa halutaan minimoida kahta funktiota

$$f_1 : \mathbb{R} \to \mathbb{R},\ f_1(x) = |x^3 + x|,$$
$$f_2 : \mathbb{R} \to \mathbb{R},\ f_2(x) = \begin{cases} -3x, & x \leq 0, \\ -\frac{1}{2}x, & x > 0. \end{cases}$$

Rajoitteina toimivat funktiot $g_1 : \mathbb{R} \to \mathbb{R}$, $g_1(x) = -x - 1$, $g_2 : \mathbb{R} \to \mathbb{R}$, $g_2(x) = x - 1$ ja $g_3 : \mathbb{R} \to \mathbb{R}$, $g_3(x) = -x^2$, jolloin sallituksi alueeksi muodostuu $S = \left\{x \in \mathbb{R} \mid g(x) = (g_1(x), g_2(x), g_3(x))^T \leq \mathbf{0}\right\} = [-1, 1]$.

Kuten kuvasta 4 nähdään, kohta $x = 0$ on selvästi funktion $f_1(x)$ minimikohta. Funktio $f_2(x)$ saisi kuitenkin vielä pienempiä arvoja, kun x kasvaa. Piste $x = 0$ on Pareto-optimaalinen, koska ominaisuutta f_2 voidaan parantaa, mutta samalla f_1 huonontuu ja ei ole toista pistettä $x' \in \mathbb{R}$, jolle $f_1(x') \leq f_1(x)$ ja $f_2(x') \leq f_2(x)$. Tarkistetaan Pareto-optimaalisuus vielä lauseen 4.5 avulla.

Optimointitehtävä on muotoa

$$\begin{aligned} \text{min.} \quad & \{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\} \\ \text{s.t.} \quad & \boldsymbol{x} \in S = [-1, 1]\,. \end{aligned}$$

Kohdefunktiot f_1 ja f_2 sekä rajoitteet g_1, g_2 ja g_3 ovat lokaalisti Lipschitz-jatkuvia pisteessä $x = 0$, jossa lisäksi kohdefunktioilla on kummallakin epädifferentioituvuuskohta. Nimetään tätä pistettä symbolilla x^*. Jos piste x^* on Pareto-optimaalinen, on lauseen 4.5 nojalla olemassa vektorit $\boldsymbol{\lambda} \in \mathbb{R}^2$ ja $\boldsymbol{\mu} \in \mathbb{R}^3$, joille $\boldsymbol{\lambda} \geq \mathbf{0}$, $\boldsymbol{\mu} \geq \mathbf{0}$ ja ehdot (A) ja (B) ovat voimassa. Koska $g_1(x^*) = g_2(x^*) = -1$ ja $g_3(x^*) = 0$, voidaan valita alkioiksi $\mu_1 = \mu_2 = 0$, jolloin

$$\mu_j g_j(x^*) = 0, \text{ kaikilla } j = 1, 2, 3,$$

ja ehto (B) toteutuu. Ehdon (A) tutkimiseksi pitää selvittää funktioiden f_1, f_2 ja g_3 alidifferentiaalit pisteessä x^*. Alidifferentiaalit $\partial g_1(x^*)$ ja $\partial g_2(x^*)$ voidaan jättää huomiotta, sillä $\mu_1 = \mu_2 = 0$. Esimerkissä 2.3 on jo selvitetty, että $\partial f_1(x^*) = [-1, 1]$. Samalla metodilla voidaan laskea alidifferentiaali $\partial f_2(x^*) = \left[-3, -\frac{1}{2}\right]$. Koska $g_3(x)$ on jatkuvasti differentioituva, on alidifferentiaali $\partial g_3(x^*) = 0$. Nyt voidaan valita alkioiksi $\lambda_1 = \frac{1}{2}$, $\lambda_2 = 0$ ja $\mu_3 = \frac{1}{2}$. Tästä seuraa, että

$$\begin{aligned} 0 &\in \sum_{i=1}^{2} \lambda_i \partial f_i(x^*) + \sum_{j=1}^{3} \mu_j \partial g_j(x^*) \\ &= \frac{1}{2}[-1, 1] + \frac{1}{2}\{0\} = \left[-\frac{1}{2}, \frac{1}{2}\right], \end{aligned}$$

joten ehto (A) toteutuu, sekä lisäksi $\sum_{i=1}^{2} \lambda_i + \sum_{j=1}^{3} \mu_j = 1$ ja $(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \left(\frac{1}{2}, 0, 0, 0, \frac{1}{2}\right) \neq (\mathbf{0}, \mathbf{0})$. Pisteen x^* Pareto-optimaalisuus on täten vahvistettu lauseen 4.5 avulla.

Koska heikosti Pareto-optimaalisten pisteiden joukko sisältää Pareto-optimaalisten pisteiden joukon, on seuraava tulos voimassa.

Seuraus 4.6. *Olkoon $E \subset \mathbb{R}^n$ optimointitehtävän sallittujen ratkaisupisteiden joukko. Lause 4.5 toimii välttämättömänä ehtona myös sille, että piste $\boldsymbol{x} \in E$ olisi heikosti Pareto-optimaalinen.*

Fritz John -tyyppistä optimaalisuusehtoa ei ole Geoffrionin mukaisille vahvasti Pareto-optimaalisille pisteille, kun kohde- ja rajoitefunktiot ovat epädifferentioituvia. Sen sijaan voidaan kehitellä itse asiassa riittävä optimaalisuusehto yhdistettyjen konveksien funktioiden avulla. Ehto säilyttää oletuksen lokaalista Lipschitz-jatkuvuudesta mutta vaatii, että funktiot ovat konvekseja. Tarkat määrittelyt voi löytää Jeyakumarin ja Yangin artikkelista vuodelta 1993.[7]

5 Yhteenveto

Tässä tutkielmassa on esitelty Lipschitz-jatkuvuutta ja konveksisuutta sekä määritelty suunnatut derivaatat ja alidifferentiaali. Päätuloksena on esitetty, miten optimointitehtävän sallitun alueen ratkaisupisteille voidaan määritellä ehdot niiden optimaalisuudelle. Optimaalisuusehdoilla rajoitetaan sallittua aluetta optimipisteiden löytämiseksi. Käsitellyt Fritz John -tyyppiset ehdot yhden tavoitteen optimoinnin ratkaisun optimaalisuudelle sekä monitavoiteoptimoinnin ratkaisun Pareto-optimaalisuudelle ovat välttämättömiä mutta eivät välttämättä kuitenkaan riittäviä. Varsinkin, jos kohdefunktiot ja rajoitteet eivät ole konvekseja, ehtojen riittävyys on kyseenalaista [4]. Voidaan kuitenkin olla varmoja, että piste ei ole optimaalinen tai Pareto-optimaalinen, jos se ei toteuta vastaavaa Fritz John -ehtoa.

Yhden tavoitteen tehtävässä vaaditaan olevan olemassa kohdefunktion ja rajoitteiden alidifferentiaaleille reaaliarvoiset skalaarit siten, että nollavektori sisältyy näiden skaalattujen alidifferentiaalien summaan. Lisäksi rajoitteen tai rajoitetta vastaavan skalaarin pitää olla arvoltaan 0 tarkasteltavassa ratkaisupisteessä. Vähintään yhden rajoitteen tulee kuitenkin saavuttaa nollakohtansa ratkaisupisteessä ja kohdefunktion alidifferentiaalin skalaari pitää olla erisuuri kuin 0. Monitavoiteoptimoinnin ratkaisun Pareto-optimaalisuutta määrittävä ehto on vastaavanlainen, mutta jokaisen kohdefunktion alidifferentiaalille on oma skalaarinsa ja vähintään yhden niistä täytyy olla erisuuri kuin 0.

Viitteet

[1] K. Miettinen: *Nonlinear Multiobjective Optimization.* 1999

[2] M. Mäkelä, P. Neittaanmäki: *Nonsmooth Optimization.* 1992

[3] R. T. Rockafellar: *Convex Analysis.* 1970

[4] O. L. Mangasarian, S. Fromovitz: *The Fritz John Necessary Optimality Conditions in the Presence of Equality and Inequality Constraints* lehdessä *Journal of Mathematical Analysis and Applications 17.* 1967, s.37–47

[5] F. H. Clarke: *Optimization and nonsmooth analysis - Society for Industrial and Applied Mathematics.* 1987

[6] A. W. Roberts, D. E. Varberg: *Another Proof that Convex Functions are Locally Lipschitz* lehdessä *The American Mathematical Monthly.* Taylor and Francis Group, 1974, s.1014–1016

[7] V. Jeyakumar, X. Q. Yang: *Convex composite multi-objective non-smooth programming* teoksessa *Mathematical Programming* 59. 1993, s.325–343