Helsingin yliopisto
Matemaattis-luonnontieteellinen tiedekunta
Matematiikan ja tilastotieteen osasto

Kandidaatintutkielma

Konveksit funktiot

Sini Järveläinen

Matematiikan, fysiikan ja kemian opettajan kandiohjelma
Ohjaaja: Mika Koskenoja 13. toukokuuta 2024

Tiivistelmä

Tämä kandidaatintutkielma käsittelee konvekseja funktioita ja niiden ominaisuuksia, ja se liittyy siten analyysiin alaan. Konveksit funktiot ovat rakenteeltaan erityisiä ja siksi merkittäviä funktioita esimerkiksi optimoinnissa.

Tutkielman aluksi luvussa 2 määritellään konveksisuus yhden muuttujan reaalifunktiolle. Algebrallisesta määritelmästä johdetaan erilaisia lauseita, jotka kuvaavat konveksin funktion geometrista muotoilua. Konveksi funktio on funktio, jonka kuvaajalta voidaan valita mitkä tahansa kaksi pistettä ja näiden pisteiden välinen jana on aina funktion kuvaajan yläpuolella (tai päällä).

Luvussa 3 tutkitaan konveksien funktioiden ominaisuuksia. Konveksi funktio $f : (a, b) \to \mathbb{R}$ osoitetaan jatkuvaksi. Konveksilla funktiolla on myös vasemman- ja oikeanpuoleiset derivaatat kaikkialla määrittelyjoukossaan. Konveksi funktio on suhteellisen sileän rakenteensa takia melkein kaikkialla derivoituva, mutta sen ei ole kuitenkaan pakko olla jokaisessa pisteessä derivoituva.

Konveksi funktio on joko monotoninen tai ensin vähenevä ja sitten kasvava. Tästä seuraa myös merkittävä ominaisuus liittyen funktion minimiin: jos konveksilla funktiolla on lokaali minimi, niin kyseessä on myös globaali minimi. Aidosti konveksilla funktiolla globaali minimi on yksikäsitteinen. Tämä ominaisuus tekee konveksista funktiosta merkittävän työkalun optimoinnissa.

Konveksien funktioiden ominaisuuksia käsittelevä luku päättyy derivoituviin konvekseihin funktioihin. Tutkielmassa todistetaan lause, jonka mukaan derivoituva funktio on konveksi, jos ja vain jos sen toiselle derivaatalle pätee $f'' \geq 0$ kaikkialla funktion määrittelyjoukossa. Tämän lauseen avulla derivoituva funktio on helppo näyttää konveksiksi.

Tutkielman lopuksi luvussa 4 esitellään Jensenin epäyhtälö. Jensenin epäyhtälö johdetaan konveksin funktion määritelmästä. Kyseessä on yksi matematiikan tärkeimmistä epäyhtälöistä, ja sen avulla todistetaankin usein muita epäyhtälöitä.

Sisällys

1 Johdanto 1

2 Konveksisuus 2
2.1 Konveksi funktio 2
2.2 Geometrinen tulkinta 2

3 Konveksien funktioiden ominaisuuksia 6
3.1 Jatkuvuus 6
3.2 Derivoituvuus 6
3.3 Monotonisuus 8
3.4 Suurin arvo 9
3.5 Globaali minimi 9
3.6 Derivoituvat konveksit funktiot 10

4 Jensenin epäyhtälö 14

5 Lopuksi 16

1 Johdanto

Reaalifunktion kuvaaja voi olla alaspäin kupera, ylöspäin kupera, suora tai jotain muuta. Tällöin funktio on konveksi, konkaavi, molempia tai ei kumpaakaan. Konveksin funktion kuvaaja on siis alaspäin kupera. Tässä työssä tutkitaan tällaisiä funktioita. Konveksi funktio liittyy analyysin alaan.

Kuva 1: Funktioiden kuvaajia

Kyseessä on siis rakenteeltaan erityinen funktio. Konvekseilla funktioilla onkin ominaisuuksia, jotka tekevät niistä merkittäviä funktioita useilla matematiikan ja sen sovellusten alueilla. Konvekseilla funktioilla on tärkeä asema niin matemaattisessa optimoinnissa kuin myös taloustieteessä, insinööritieteissä ja koneoppimisessa. Tässä tutkielmassa tutustutaan näihin konveksien funktioiden ominaisuuksiin.

Aluksi määritellään konveksi funktio algebrallisesti ja tulkitaan sitten määritelmää geometrisesti. Luvun 2 tarkoituksena on muodostaa lukijalle käsitys konveksista funktiosta ja sen rakenteesta. Tämän jälkeen funktioiden ominaisuuksien tutkiminen ja ymmärtäminen on helpompaa.

Luvussa 3 esitellään konveksien funktioiden ominaisuuksia. Luvussa annetaan konveksien funktioiden analyysiin liittyviä lauseita ja niille melko yksityiskohtaisetkin todistukset. Yksi tärkeimmistä ominaisuuksista on luvussa 3.5 esitelty globaali minimi, joka liittyy vahvasti optimointiin.

Lopuksi luvussa 4 esitellään vielä konveksin funktion tärkeä sovellus, Jensenin epäyhtälö. Jensenin epäyhtälö on yksi matematiikan tärkeimmistä epäyhtälöistä, ja sen avulla todistetaan useita muita epäyhtälöitä.

Tutkielman tavoitteena on tarjota lukijalle kattava ja ymmärrettävä johdatus konveksien funktioiden teoriaan. Tutkielma esittelee konveksien funktioiden keskeisiä matemaattisia ominaisuuksia pyrkien näyttämään konveksien funktioiden merkittävän aseman eri aloilla.

2 Konveksisuus

Tarkastellaan yhden muuttujan reaalifunktioita. Määritellään aluksi konveksi funktio ja tarkastellaan sitten sen rakennetta tulkiten määritelmää geometrisesti. Luvussa 2.2 esiintyvät lauseet löytyvät Andrew Browderin kirjasta [1]. Lauseiden todistukset ovat tässä yksityiskohtaisempia kuin lähteessä.

2.1 Konveksi funktio

Määritelmä 2.1. Olkoon $I \subseteq \mathbb{R}$ väli. Funktio $f : I \to \mathbb{R}$ on konveksi, jos

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) \tag{2.2}$$

kaikilla $x, y \in I$ ja $\lambda \in [0, 1]$. Jos määritelmässä vaaditaan aito epäyhtälö $f((1-\lambda)x + \lambda y) < (1-\lambda)f(x) + \lambda f(y)$ kaikilla $x, y \in I$ ja $\lambda \in [0, 1]$, niin funktio $f : I \to \mathbb{R}$ on aidosti konveksi. Jos funktio f on (aidosti) konveksi, niin funktio $-f$ on (aidosti) konkaavi.

Koska konveksin (ja konkaavin) funktion määritelmässä sallitaan yhtäsuuruus, funktio voi olla sekä konveksi että konkaavi. Lineaarinen funktio on sekä konveksi että konkaavi.

Esimerkki 2.3. Olkoon $f : \mathbb{R} \to \mathbb{R}, f(x) = |x|$. Sijoitetaan epäyhtälöön 2.2 f:n paikalle itseisarvofunktio, jolloin saadaan

$$|(1-\lambda)x + \lambda y| \leq (1-\lambda)|x| + \lambda|y|.$$

Kolmioepäyhtälön mukaan kaikille reaaliluvuilla x ja y on voimassa $|x + y| \leq |x| + |y|$. Lisäksi luvut λ ja $1 - \lambda$ on määritelty ei-negatiivisiksi. Siten

$$|(1-\lambda)x + \lambda y| \leq |(1-\lambda)x| + |\lambda y| = (1-\lambda)|x| + \lambda|y|,$$

mikä on haluttu väite. Siis f on konveksi.

Esimerkki 2.4. Reaalifunktiot

a) $f(x) = ax + b$

b) $g(x) = ax^2 + bx + c$, kun $a > 0$ ja

c) $h(x) = e^{ax}$, kun $a \neq 0$

ovat konvekseja. Nämä todistetaan konvekseiksi luvussa 3.6.

2.2 Geometrinen tulkinta

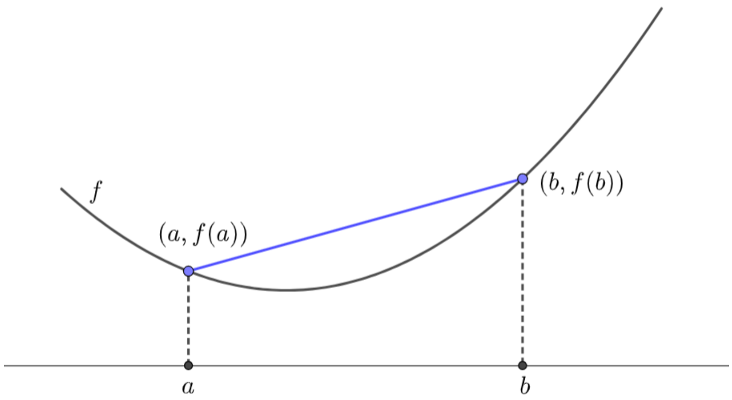
Geometrisesti konveksisuus tarkoittaa sitä, että kun funktion f kuvaajalta valitaan mitkä tahansa kaksi pistettä $(a, f(a))$ ja $(b, f(b))$, niiden välinen jana on funktion f kuvaajan yläpuolella (tai sen päällä). Funktion kuvaaja on siis alaspäin kupera. Konveksisuuden määritelmä 2.1 onkin yhtäpitävä tässä luvussa esitettyjen lauseiden kanssa.

Lause 2.5. *Olkoot $I \subseteq \mathbb{R}$ väli ja $f : I \to \mathbb{R}$. Funktio f on konveksi, jos ja vain jos*

$$f(x) \leq f(a) + \frac{f(b) - f(a)}{b - a}(x - a)$$

kaikilla $a, b \in I$ ja $a < x < b$.

Huomaa, että tässä pisteiden $(a, f(a))$ ja $(b, f(b))$ kautta kulkevan suoran yhtälö on $y = f(a) + \frac{f(b)-f(a)}{b-a}(x - a)$.



Kuva 2: Konveksi funktio

Lauseen 2.5 todistus. Oletetaan ensin, että f on konveksi. Valitaan $\lambda = \dfrac{x-a}{b-a}$. Nyt kun $\lambda = 0$, niin $x = a$ ja kun $\lambda = 1$, niin $x = b$. Siis kun $\lambda \in (0, 1)$, niin $a < x < b$. Nyt

$$\lambda = \frac{x - a}{b - a} \Leftrightarrow x - a = \lambda b - \lambda a \Leftrightarrow x = a - \lambda a + \lambda b = (1 - \lambda)a + \lambda b.$$

Nyt konveksisuuden määritelmästä 2.1 seuraa, että

$$\begin{aligned} f(x) \leq (1-\lambda)f(a) + \lambda f(b) &= f(a) - \frac{x-a}{b-a}f(a) + \frac{x-a}{b-a}f(b) \\ &= f(a) + \frac{(x-a)f(b)}{b-a} - \frac{(x-a)f(a)}{b-a} = f(a) + \frac{f(b)-f(a)}{b-a}(x-a). \end{aligned}$$

Oletetaan sitten, että lauseen ehto on voimassa. Olkoon $\lambda \in (0, 1)$. Tällöin on olemassa $x \in (a, b)$ siten, että $\lambda = \dfrac{x-a}{b-a}$, jolloin $x = (1 - \lambda)a + \lambda b$. Lauseen ehdosta seuraa, että

$$\begin{aligned} f((1-\lambda)a + \lambda b) &\leq f(a) + \frac{f(b)-f(a)}{b-a}((1-\lambda)a + \lambda b - a) \\ &= f(a) + \frac{f(b)-f(a)}{b-a}\lambda(b-a) = f(a) + \lambda f(b) - \lambda f(a) \\ &= (1-\lambda)f(a) + \lambda f(b). \end{aligned}$$

Siis f on konveksi. $\square$

Lause 2.6. *[1, 4.10 Proposition] Olkoot $I \subseteq \mathbb{R}$ väli ja $f : I \to \mathbb{R}$. Funktio f on konveksi, jos ja vain jos*

$$f(x) \leq \frac{x-a}{b-a} f(b) + \frac{b-x}{b-a} f(a)$$

kaikilla $a, b \in I$ ja $a < x < b$.

Huomaa, että tässä pisteiden $(a, f(a))$ ja $(b, f(b))$ kautta kulkevan suoran yhtälö on $\dfrac{x-a}{b-a} f(b) + \dfrac{b-x}{b-a} f(a)$. Yhtälö voidaan löytää lineaarisen interpolaation kaavalla [2].

Lauseen 2.6 todistus. Tehdään todistus samaan tapaan kuin lauseen 2.5 todistus. Oletetaan ensin, että f on konveksi ja valitaan $\lambda = \dfrac{x-a}{b-a}$, jolloin $x = (1-\lambda)a + \lambda b$. Konveksisuuden määritelmästä seuraa, että

$$f(x) \leq (1-\lambda) f(a) + \lambda f(b) = \left(1 - \frac{x-a}{b-a}\right) f(a) + \frac{x-a}{b-a} f(b) = \frac{x-a}{b-a} f(b) + \frac{b-x}{b-a} f(a).$$

Oletetaan sitten, että lauseen väite pätee. Olkoon $\lambda \in (0,1)$. Tällöin on olemassa $x \in (a,b)$ siten, että $\lambda = \dfrac{x-a}{b-a}$ ja $x = (1-\lambda)a + \lambda b$. Lauseen ehdosta seuraa, että

$$\begin{aligned} f((1-\lambda)a + \lambda b) &\leq \lambda f(b) + \frac{b - (1-\lambda)a - \lambda b}{b-a} f(a) = \lambda f(b) + \frac{(1-\lambda)b - (1-\lambda)a}{b-a} f(a) \\ &= \lambda f(b) + \frac{(1-\lambda)(b-a)}{b-a} f(a) = (1-\lambda) f(a) + \lambda f(b). \end{aligned}$$

Siis f on konveksi. $\square$

Lauseista 2.5 ja 2.6 seuraa suoraan se, että välin (a,b) ulkopuolella pisteiden $(a, f(a))$ ja $(b, f(b))$ kautta kulkeva suora on funktion f kuvaajan alapuolella.

Lause 2.7. *[1, 4.11 Proposition] Olkoot $I \subseteq \mathbb{R}$ väli, $f : I \to \mathbb{R}$ ja $a < b < c \in I$. Jos funktio f on konveksi, niin*

$$f(x) \geq \frac{x-a}{b-a} f(b) + \frac{b-x}{b-a} f(a) \text{ kaikilla } b < x < c$$

ja

$$f(x) \geq \frac{c-x}{c-b} f(b) + \frac{x-b}{c-b} f(c) \text{ kaikilla } a < x < b.$$

Todistus. Oletetaan, että f on konveksi. Tällöin siihen voidaan soveltaa lausetta 2.6. Jos $b < x < c$, niin on voimassa $a < b < x$. Lauseesta 2.6 seuraa, että

$$f(b) \leq \frac{b-a}{x-a} f(x) + \frac{x-b}{x-a} f(a).$$

Kun ratkaistaan yhtälöstä $f(x)$, saadaan

$$f(x) \geq \frac{x-a}{b-a}f(b) - \frac{x-b}{b-a}f(a) = \frac{x-a}{b-a}f(b) + \frac{b-x}{b-a}f(a).$$

Jos taas $a < x < b$, niin on voimassa $x < b < c$. Nyt lauseesta 2.6 seuraa, että

$$f(b) \leq \frac{b-x}{c-x}f(c) + \frac{c-b}{c-x}f(x).$$

Kun ratkaistaan yhtälöstä $f(x)$, saadaan

$$f(x) \geq \frac{c-x}{c-b}f(b) - \frac{b-x}{x-a}f(a) = \frac{c-x}{c-b}f(b) + \frac{x-b}{x-a}f(a).$$

□

Lause 2.8. *[1, 4.13 Propostion] Jos $f : I \to \mathbb{R}$ on konveksi, niin*

$$\frac{f(x) - f(a)}{x-a} \leq \frac{f(b) - f(a)}{b-a} \leq \frac{f(b) - f(x)}{b-x} \tag{2.9}$$

kaikilla $a, b \in I$ ja $a < x < b$.

Todistus. Oletetaan, että f on konveksi. Tällöin lauseen 2.6 mukaan

$$f(x) \leq \frac{x-a}{b-a}f(b) + \frac{b-x}{b-a}f(a).$$

Vähennetään epäyhtälöstä puolittain $f(a)$ ja jaetaan luvulla $x - a$, jolle pätee $x - a > 0$, koska $a < x$. Saadaan

$$\begin{aligned}
f(x) - f(a) &\leq \frac{x-a}{b-a}f(b) + \frac{b-x}{b-a}f(a) - f(a) \\
f(x) - f(a) &\leq \frac{x-a}{b-a}f(b) + f(a)\left(\frac{b-x}{b-a} - 1\right) = \frac{x-a}{b-a}f(b) - \frac{x-a}{b-a}f(a) \\
\frac{f(x) - f(a)}{x-a} &\leq \frac{f(b) - f(a)}{b-a}.
\end{aligned}$$

Tehdään sitten sama siten, että suoritetaankin vähennys luvulla $f(b)$. Saadaan

$$\begin{aligned}
f(x) - f(b) &\leq \frac{x-a}{b-a} - f(b) + \frac{b-x}{b-a}f(a) \\
-\frac{f(b) - f(x)}{b-x} &\leq -\frac{f(b) - f(a)}{b-a} \\
\frac{f(b) - f(x)}{b-x} &\geq \frac{f(b) - f(a)}{b-a}.
\end{aligned}$$

□

Epäyhtälö 2.9 voidaan tulkita geometrisesti siten, että pisteiden $(a, f(a))$ ja $(b, f(b))$ kautta kulkevan suoran kulmakerroin kasvaa, kun b kasvaa tai kun a kasvaa.

Tähän saakka olemme määritelleet konveksin funktion algebrallisesti sekä geometrisesti. Konveksisuudelle voidaan antaa myös derivaattamuotoilu, jos funktio on derivoituva. Tästä lisää luvussa 3.6.

3 Konveksien funktioiden ominaisuuksia

3.1 Jatkuvuus

Lause on Niels Lauritzenin kirjan [3, Theorem 7.7.] mukainen. Todistus on Browderin teoksen [1, 4.12 Corollary] mukainen ja se hyödyntää edellä todistettuja lauseita. Toisenlaisia todistuksia samalle lauseelle löytyy teoksista [3] ja [4, 1.1.2 Theorem].

Lause 3.1. *Konveksi funktio $f : [t, u] \to \mathbb{R}$ on jatkuva avoimella välillä (t, u).*

Todistus. Olkoon $b \in (t, u)$. Koska kyseessä on avoin väli, niin on olemassa jotkin $a, c \in (t, u)$ siten, että $a < b < c$. Koska f on konveksi, niin lauseiden 2.3 ja 2.4 mukaan

$$\frac{c-x}{c-b}f(b) + \frac{x-b}{c-b}f(c) \leq f(x) \leq \frac{x-a}{b-a}f(b) + \frac{b-x}{b-a}f(a)$$

kaikilla $a < x < b$. Nyt koska

$$\lim_{x \to b} \frac{c-x}{c-b}f(b) + \frac{x-b}{c-b}f(c) = f(b) = \lim_{x \to b} \frac{x-a}{b-a}f(b) + \frac{b-x}{b-a}f(a),$$

niin $\lim_{x \to b^-} f(x) = f(b)$. Samoin lauseiden 2.3 ja 2.4 mukaan

$$\frac{x-a}{b-a}f(b) + \frac{b-x}{b-a}f(a) \leq f(x) \leq \frac{x-b}{c-b}f(c) + \frac{c-x}{c-b}f(b)$$

kaikilla $b < x < c$. Nyt koska

$$\lim_{x \to b} \frac{x-a}{b-a}f(b) + \frac{b-x}{b-a}f(a) = f(b) = \lim_{x \to b} \frac{x-b}{c-b}f(c) + \frac{c-x}{c-b}f(b),$$

niin $\lim_{x \to b^+} f(x) = f(b)$. Jatkuvuuden määritelmän [5, s. 69] mukaan f on vasemmalta jatkuva pisteessä b ja f on oikealta jatkuva pisteessä b. Siten f on jatkuva pisteessä b. $\square$

Funktion jatkuvuus on välttämätön mutta ei riittävä ehto funktion konveksisuudelle. Huomaa, että suljetulla välillä konveksin funktion ei tarvitse olla jatkuva välin päätepisteissä.

3.2 Derivoituvuus

Tutkitaan seuraavaksi konveksin funktion derivoituvuutta.

Lause 3.2. *([1], 4.14 Theorem) Olkoon f konveksi funktio avoimella välillä I. Tällöin funktiolla f on vasemman- ja oikeanpuoleiset derivaatat f'_- ja f'_+ jokaisessa pisteessä $b \in I$ ja kaikille $a < b < c$ pätee*

$$f'_+(a) \leq f'_-(b) \leq f'_+(b) \leq f'_-(c).$$

Todistus. Olkoon $b \in I$. Koska f on konveksi, niin lauseen 2.8 mukaan tällöin kaikille $t < b < u$ pätee

$$\frac{f(b)-f(t)}{b-t} \leq \frac{f(u)-f(b)}{u-b},$$

jolloin voidaan merkitä ylä- ja alarajat

$$\sup_{t<c} \frac{f(b)-f(t)}{b-t} \leq \inf_{u>c} \frac{f(u)-f(b)}{u-b} \tag{3.3}$$

ja epäyhtälön molemmat puolet ovat äärellisiä. Olkoot $s, v \in I$ siten, että $s < t < b < u < v$. Tällöin lauseen 2.8 vasemmanpuoleisesta epäyhtälöstä saadaan

$$\frac{f(u)-f(b)}{u-b} \leq \frac{f(v)-f(b)}{v-b},$$

joten funktio $u \mapsto \dfrac{f(u)-f(b)}{u-b}$ on kasvava joukossa $I \cap (b, \infty)$, ja siten oikeanpuoleinen derivaatta

$$f'_+(b) = \lim_{u \to b^+} \frac{f(u)-f(b)}{u-b} = \inf_{u>c} \frac{f(u)-f(b)}{u-b}$$

on olemassa. Vastaavasti lauseen 2.8 oikeanpuoleisesta epäyhtälön saadaan

$$\frac{f(b)-f(s)}{b-s} \leq \frac{f(b)-f(t)}{b-t},$$

joten funktio $t \mapsto \dfrac{f(b)-f(t)}{b-t}$ on kasvava joukossa $I \cap (-\infty, b)$, ja siten vasemmanpuoleinen derivaatta

$$f'_-(b) = \lim_{t \to b^-} \frac{f(b)-f(t)}{b-t} = \sup_{t<c} \frac{f(b)-f(t)}{b-t}$$

on olemassa. Nyt epäyhtälöstä 3.3 nähdään, että $f'_-(b) \leq f'_+(b)$, ja molemmat ovat reaalilukuja, joten ne todella ovat funktion f vasemman- ja oikeanpuoleiset derivaatat pisteessä b. Olkoot nyt $a < b < c$. Valitsemalla $t \in (a, b)$ saadaan

$$f'_+(a) \leq \frac{f(t)-f(a)}{t-a} \leq \frac{f(b)-f(t)}{b-t} \leq f'_-(b)$$

ja valitsemalla $u \in (b, c)$ saadaan

$$f'_+(b) \leq \frac{f(u)-f(b)}{u-b} \leq \frac{f(c)-f(u)}{c-u} \leq f'_-(c).$$

Siis $f'_+(a) \leq f'_-(b) \leq f'_+(b) \leq f'_-(c)$. □

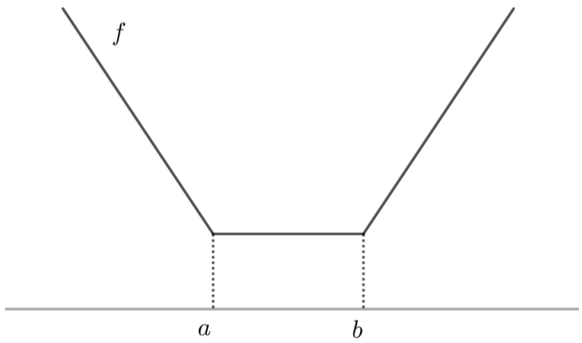
Konveksilla funktiolla on siis jokaisessa avoimen välin pisteessä vasemman- ja oikeanpuoleiset derivaatat. Funktion ei kuitenkaan tarvitse olla derivoituva jokaisessa pisteessä.

Esimerkki 3.4. Aiemmin osoitimme (esimerkki 2.3) funktion $f : \mathbb{R} \to \mathbb{R}, f(x) = |x|$ konveksiksi. Funktio ei ole derivoituva pisteessä $x = 0$. Sillä on kuitenkin vasemman- ja oikeanpuoleiset derivaatat $f'_-(0) = -1$ ja $f'_+(0) = 1$, joille pätee $f'_-(0) \leq f'_+(0)$.

Korollaari 3.5. *[1, 4.15 Corollary] Olkoon* $f : (a, b) \to \mathbb{R}$*. Tällöin on olemassa numeroituva joukko* $S \subset (a, b)$ *siten, että kaikille* $x \in (a, b) \setminus S$ *on olemassa* $f'(x)$*.*

Siis joukko S, joka sisältää ne pisteet, joissa funktio f ei ole derivoituva, on aina löydettävissä ja sen alkiot voidaan luetella. Tämän joukon ulkopuolella määrittelyjoukossaan f on derivoituva. Esimerkiksi mainitulle itseisarvofunktiolle merkitään $S = \{0\}$ ja f on derivoituva joukossa $\mathbb{R} \setminus S$. Korollaari viittaa siihen, että vaikka derivaatta voi puuttua joissakin pisteissä, konveksin funktion yleinen sileys ei ”häiriinny” liikaa ja derivaatta onkin olemassa melkein kaikkialla. Korollaari vahvistaa konveksien funktioiden säännöllisyyttä ja ennustettavuutta.

Esimerkki 3.6. Kuvassa 3 funktio f on derivoituva kaikkialla muualla paitsi pisteissä a ja b. Siten $S = \{a, b\}$. Huomaa, että derivaatta noudattaa nousevaa järjestystä. Kun $x < a$, niin $f'(x) = -1$. Kun $a < x < b$, niin $f'(x) = 0$. Kun $x > b$, niin $f'(x) = 1$. Lisäksi $f'_-(a) = -1, f'_+(a) = 0, f'_-(b) = 0$ ja $f'_+(b) = 1$.



Kuva 3: Ei-derivoituva konveksi funktio

3.3 Monotonisuus

Funktio on monotoninen, kun se on kasvava tai vähenevä. Tutkitaan seuraavaksi konveksin funktion monotonisuutta.

Lause 3.7. *[4, 1.1.4 Lemma] Jos* $f : (a, b) \to \mathbb{R}$ *on konveksi, niin* f *on joko monotoninen tai on olemassa sellanen piste* $\varepsilon \in I$*, että* f *on vähenevä välillä* (a, ε) *ja kasvava välillä* (ε, b)*.*

Todistus. Oletetaan, että f ei ole monotoninen. Valitaan mielivaltaiset $t, u \in (a, b)$ siten, että $t < u$. Merkitään $m = \inf\{f(x) : x \in [t, u]\}$. Koska f on konveksi, niin se on jatkuva välillä $[t, u]$ (lause 3.1). Siten m saavutetaan jossakin pisteessä $\varepsilon \in [t, u]$ ja tällöin merkitään $m = f(\varepsilon)$.

Jos $t \leq x < y < \varepsilon$, niin lauseen 2.6 mukaan

$$f(y) \leq \frac{y-x}{\varepsilon - x} f(\varepsilon) + \frac{\varepsilon - y}{\varepsilon - x} f(x).$$

Koska $f(\varepsilon) = m$, niin $f(x) \geq f(\varepsilon)$. Lisäksi, koska funktioiden $f(\varepsilon)$ ja $f(x)$ kertoimien summa on 1, niin voidaan asettaa yläraja $f(y) \leq f(x)$. Siten f on vähenevä välillä $[t, \varepsilon]$.

Jos $\varepsilon < x < y \leq u$, niin samaan tapaan voidaan asettaa

$$f(x) \leq \frac{x-\varepsilon}{y-\varepsilon} f(y) + \frac{y-x}{y-\varepsilon} f(\varepsilon) \leq f(y),$$

jolloin f on kasvava välillä $[\varepsilon, u]$. □

Korollaari 3.8. *[4, 1.1.5 Corollary] a) Jos konveksi funktio $f : [a, b] \to \mathbb{R}$ ei ole monotoninen, niin sillä on globaali minimi jossain välin (a, b) pisteessä.*
b) Jos konveksi funktio $f : \mathbb{R} \to \mathbb{R}$ on rajoitettu ylhäältä, niin se on vakiofunktio.

Siis konveksi funktio on aina joko monotoninen koko määrittelyjoukossaan tai sitten se on ensin vähenevä ja sitten kasvava siten, että tämä käännekohta on funktion globaali minimi. Puhumme lisää konveksien funktioiden globaaleista minimeistä luvussa 3.5.

3.4 Suurin arvo

Konveksien funktioiden erityinen geometrinen ominaisuus (ks. luku 2.2) antaa tärkeän tuloksen, kun tutkitaan konvekseja funktioita suljetuilla väleillä.

Lause 3.9. *Konveksi funktio $f : [a, b] \to \mathbb{R}$ saavuttaa suurimman arvonsa välin $[a, b]$ jommassa kummassa päätepisteessä.*

Huomaa, että lauseessa funktio on määritelty suljetulla välillä. Konveksin funktion ei aina tarvitse saavuttaa suurinta arvoaan. Esimerkiksi $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$ ei saavuta suurinta arvoaan, mutta $f : [-1, 1] \to \mathbb{R}, f(x) = x^2$ saavuttaa suurimman arvonsa 1 välinsä päätepisteissä.

3.5 Globaali minimi

Määritellään ensin funktion globaali minimi.

Määritelmä 3.10. [3, Definition 7.5] Funktion $f : I \to \mathbb{R}$ globaali minimipiste on $x_0 \in I$, jos ja vain jos

$$f(x_0) \leq f(x)$$

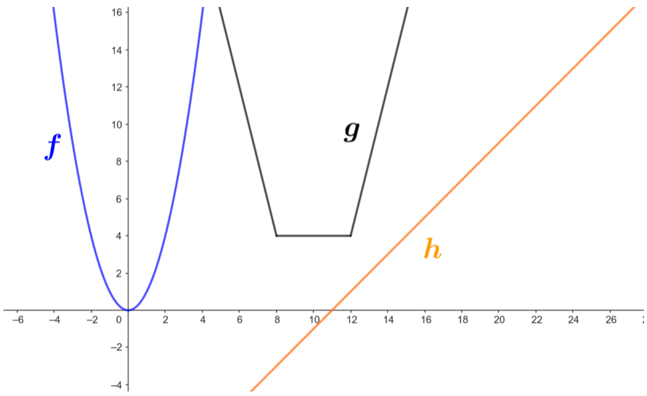
kaikilla $x \in I$. Piste x_0 on lokaali minimi, jos on olemassa $\varepsilon > 0$ siten, että $f(x_0) \leq f(x)$ kaikilla $x \in I \cap (x_0 - \varepsilon, x_0 + \varepsilon)$.

Globaali minimi on selvästi lokaali minimi, mutta sama ei päde toiseen suuntaan. Tarkastellaan sitten konveksien funktioiden globaalia minimiä. Lauseesta 3.7 seuraa seuraava merkittävä tulos.

Lause 3.11. *[3, Lemma 7.6] Olkoon $f : I \to \mathbb{R}$ konveksi funktio. Jos $x_0 \in I$ on lokaali minimi, niin x_0 on globaali minimi. Jos f on aidosti konveksi, niin globaali minimi on yksikäsitteinen.*

Konveksilla funktiolla ei siis ole lokaalia minimiä, joka ei olisi globaali minimi. Tämän ominaisuuden takia konveksi funktio on erittäin hyödyllinen funktio optimoinnissa. Se, että lokaalin minimin löytäminen tarkoittaa myös globaalin minimin löytymistä, nopeuttaa optimointiprosessia. Tällöin ei tarvitse jumittua paikallisiin minimikohtiin. Lisäksi konveksin funktion minimi on helppo löytää: liikutaan pisteestä aina siihen suuntaan, jossa funktion arvot pienenevät.

Jos funktio ei ole aidosti konveksi, lokaaleja minimejä voi olla useita, jolloin nämä minimit muodostavat yhtenäisen välin.



Kuva 4: Erilaisia konvekseja funktioita

Kuvassa 4 on kolme erilaista konveksia funktiota, jotka on määritelty koko reaalilukujen joukossa $\mathbb{R}$. Funktio f on aidosti konveksi, ja sillä on yksikäsitteinen globaali minimi origossa. Funktio g on konveksi, mutta ei aidosti konveksi, ja sen minimit muodostavat yhtenäisen välin. Lineaarisella funktiolla $h : \mathbb{R} \to \mathbb{R}$ ei ole minimejä.

3.6 Derivoituvat konveksit funktiot

Kuten aiemmin mainittu, konveksin funktion ei tarvitse olla derivoituva. Se kuitenkin voi olla derivoituva, jolloin konveksisuuden määritelmälle voidaan antaa derivaattamuotoilu. Sen avulla derivoituva funktio on helppo tunnistaa konveksiksi.

Lause 3.12. *[3, Theorem 7.19] Olkoon $f : (a, b) \to \mathbb{R}$ derivoituva funktio. Tällöin f on konveksi, jos ja vain jos f' on kasvava. Jos f' on aidosti kasvava, niin f on aidosti konveksi.*

Todistus. Oletetaan aluksi, että f derivoituva konveksi funktio. Olkoon $a < x \leq y < b$.

Tällöin valitsemalla riittävän pieni $h > 0$ lauseesta 2.8 seuraa

$$\frac{f(x) - f(x-h)}{h} \leq \frac{f(y+h) - f(x-h)}{y - x + 2h}$$

ja

$$\frac{f(y+h) - f(x-h)}{y - x + 2h} \leq \frac{f(y+h) - f(y)}{h}.$$

Yhdistämällä nämä saadaan

$$\frac{f(x) - f(x-h)}{h} \leq \frac{f(y+h) - f(y)}{h}.$$

Kun $h \to 0$, niin derivaatan määritelmän mukaisesti epäyhtälön vasemmalla puolella on derivaatan $f'(x)$ vasemmanpuoleinen erotusosamäärä ja epäyhtälön oikealla puolella on derivaatan $f'(y)$ oikeanpuoleinen erotusosamäärä. Siis $f'(x) \leq f'(y)$, kun $x \leq y$. Siten f' on kasvava.

Oletetaan sitten, että f' on kasvava. Halutaan osoittaa, että tällöin $f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y)$ kaikilla $x, y \in (a, b)$ ja $\lambda \in [0, 1]$. Tämä on selvää, kun $\lambda = 0$ tai $\lambda = 1$. Oletetaan sitten, että $0 < \lambda < 1$ ja $x < y$. Merkitään $z = (1-\lambda)x + \lambda y$. Nyt väliarvolauseesta (engl. mean value theorem) [5, s. 114] seuraa, että on olemassa $\varepsilon_1 \in (x, z)$ ja $\varepsilon_2 \in (z, y)$ siten, että

$$f'(\varepsilon_1) = \frac{f(z) - f(x)}{z - x} \text{ ja } f'(\varepsilon_2) = \frac{f(y) - f(z)}{y - z}.$$

Koska f' on kasvava, niin $f'(\varepsilon_1) \leq f'(\varepsilon_2)$. Muokataan tätä lauseketta ja saadaan

$$\begin{aligned}
\frac{f((1-\lambda)x + \lambda y) - f(x)}{(1-\lambda)x + \lambda y - x} &\leq \frac{f(y) - f((1-\lambda)x + \lambda y)}{y - (1-\lambda)x - \lambda y} \\
\frac{f((1-\lambda)x + \lambda y) - f(x)}{\lambda(y - x)} &\leq \frac{f(y) - f((1-\lambda)x + \lambda y)}{(1-\lambda)(y - x)} \\
(1-\lambda)(f((1-\lambda)x + \lambda y) - f(x)) &\leq \lambda(f(y) - f((1-\lambda)x + \lambda y)) \\
f((1-\lambda)x + \lambda y) - f(x) + \lambda f(x) &\leq \lambda f(y) \\
f((1-\lambda)x + \lambda y) &\leq f(x) - \lambda f(x) + \lambda f(y) \\
f((1-\lambda)x + \lambda y) &\leq (1-\lambda)f(x) + \lambda f(y),
\end{aligned}$$

jolloin päädyttiin haluttuun tulokseen. Siis f on konveksi. □

Derivaatan ominaisuuksien mukaan se, että $f''(x) \geq 0$ kaikkialla, tarkoittaa, että f' on kasvava. Siten derivaatan ominaisuuksista ja yllä olevasta lauseesta seuraa seuraava ominaisuus.

Korollaari 3.13. *[3, Corollary 7.20] Olkoon $f : (a, b) \to \mathbb{R}$ kahdesti derivoituva funktio. Tällöin f on konveksi, jos ja vain jos $f''(x) \geq 0$ kaikilla $x \in (a, b)$. Jos $f''(x) > 0$, kaikille $x \in (a, b)$, niin f on aidosti konveksi.*

Tämän lauseen avulla derivoituva funktio on helppo tunnistaa konveksiksi.

Esimerkki 3.14. Olkoon $f(x) = ax + b$. Nyt $f'(x) = a$ ja $f''(x) = 0$. Funktio on siis konveksi.

Esimerkki 3.15. Olkoon $g(x) = ax^2 + bx + c$, missä $a > 0$. Nyt $f'(x) = 2ax + b$ ja $f''(x) = 2a$. Koska $a > 0$, niin $f''(x) > 0$. Funktio on siis aidosti konveksi.

Esimerkki 3.16. Olkoon $h(x) = e^{ax}$, missä $a \neq 0$. Nyt $f'(x) = ae^{ax}$ ja $f''(x) = a^2e^{ax}$. Koska $a \neq 0$, niin $a^2 > 0$. Lisäksi $e^{ax} > 0$. Siten $f''(x) > 0$. Funktio on siis aidosti konveksi.

Muistetaan kuitenkin, että konveksin funktion ei tarvitse olla derivoituva. Tällöin toista derivaattaa ei voi käyttää konveksisuuden näyttämiseen, vaan on käytettävä konveksin funktion määritelmää. Toinen tärkeä ominaisuus derivoituville konvekseille funktioille on seuraava lause.

Lause 3.17. *[3, Theorem 7.21] Olkoon $f : (a, b) \to \mathbb{R}$ derivoituva funktio. Tällöin f on konveksi, jos ja vain jos*

$$f(y) \geq f(x) + f'(x)(y - x)$$

kaikilla $x, y \in (a, b)$.

Todistus. Oletetaan ensin, että derivoituva funktio f on konveksi. Olkoon $x < x_1 < y$. Koska f on konveksi, lauseesta 2.8 seuraa, että

$$\frac{f(x_1) - f(x)}{x_1 - x} \leq \frac{f(y) - f(x)}{y - x}.$$

Koska kyseessä on derivoituva funktio, derivaatan määritelmästä

$$f'(x) = \lim_{x_1 \to x^+} \frac{f(x_1) - f(x)}{x_1 - x}$$

saadaan

$$f'(x) \leq \frac{f(y) - f(x)}{y - x}.$$

Samoin voidaan todistaa, jos $y < x$. Muokataan epäyhtälöä ja saadaan

$$\begin{aligned} f'(x)(y - x) &\leq f(y) - f(x) \\ f(y) &\geq f(x) + f'(x)(y - x). \end{aligned}$$

Tämä pätee kaikilla $x, y \in (a, b)$. Oletetaan sitten, että $f(y) \geq f(x) + f'(x)(y - x)$ kaikilla $x, y \in (a, b)$. Olkoon $x < y$. Oletuksesta seuraa myös

$$f(x) \geq f(y) + f'(y)(x - y) = f(y) - f'(y)(y - x).$$

Kun yhdistetään nämä ja muokataan epäyhtälöä, saadaan

$$\begin{aligned} f(x) + f'(y)(y - x) &\geq f(x) + f'(x)(y - x) \\ f'(x) &\leq f'(y) \end{aligned}$$

kaikilla $x < y$. Siten lauseen 3.12 mukaan f on konveksi. □

Lause 3.17 antaa toisen ajatuksen siitä, miksi derivoituvat konveksit funktiot ovat erityisiä. Se itseasiassa kertoo, että jos $f : (a, b) \to \mathbb{R}$ on derivoituva konveksi funktio ja pisteelle $x_0 \in (a, b)$ pätee $f'(x_0) = 0$, niin tällöin $f(x_0) \leq f(x)$ kaikille $x \in (a, b)$. Toisin sanoen x_0 on globaali minimi.

Esimerkki 3.18. Olkoon $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$. Tällöin $f'(x) = 2x$. Asetetaan $f'(x_0) = 0$, josta saadaan $f'(x_0) = 0 \Leftrightarrow 2x_0 = 0 \Leftrightarrow x_0 = 0$. Asetetaan sitten tämä x_0 lauseen 3.17 yhtälöön, jolloin saadaan $f(y) \geq f(0) + f'(0)(y - 0) = 0$. Koska $f(y) = y^2 \geq 0$ kaikilla y, niin yhtälö pätee ja tämä osoittaa, että f saavuttaa globaalin miniminsä, kun $x_0 = 0$.

4 Jensenin epäyhtälö

Tanskalainen matemaatikko Johan Ludwig William Valdemar Jensen löysi todella mielenkiintoisen epäyhtälön noin vuonna 1905. [3] Yhtälö ilmestyi arvostetun lehden Acta Mathematica tunnetussa julkaisussa vuonna 1906. Tässä luvussa lause 4.2 on yhdenmukainen Lauritzenin kirjan [3, Corollary 7.4.] kanssa ja sitä jatketaan matematiikkalehti Solmussa julkaistun Markku Halmetojan kirjoittaman artikkelin [6] mukaisesti.

Konveksisuuden määritelmä 2.1 voidaan kirjoittaa seuraavasti:

Funktio $f : I \to \mathbb{R}$ on konveksi, jos

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) \tag{4.1}$$

kaikilla $x_1, x_2 \in I$ ja $\lambda_1, \lambda_2 \in]0, 1[$, joille $\lambda_1 + \lambda_2 = 1$.

Jensenin epäyhtälö yleistää määritelmän epäyhtälön useammalle muuttujien λ ja x arvolle.

Lause 4.2. *Olkoot $f : I \to \mathbb{R}$ konveksi funktio ja $\lambda_1, \lambda_2, \ldots, \lambda_n$ positiivisia lukuja siten, että $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$. Tällöin*

$$f(\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \cdots + \lambda_n f(x_n)$$

kaikilla $x_1, x_2, \ldots, x_n \in I$.

Todistus. Todistetaan lause induktiolla luvun n suhteen. Väite on siis tosi, kun $n = 2$ (4.1). Tehdään sitten induktio-oletus, että

$$f(\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \cdots + \lambda_k f(x_k)$$

kaikilla $x_1, x_2, \ldots, x_k \in I$, kun $k > 2$. Todistetaan, että tällöin väite pätee myös seuraavalla luvulla $n = k + 1$. Koska

$$\lambda_1 x_1 + \cdots + \lambda_k x_k + \lambda_{k+1} x_{k+1} = (1 - \lambda_{k+1}) \left(\frac{\lambda_1}{1 - \lambda_{k+1}} x_1 + \cdots + \frac{\lambda_k}{1 - \lambda_{k+1}} x_k \right) + \lambda_{k+1} x_{k+1},$$

niin

$$\begin{aligned} & f(\lambda_1 x_1 + \cdots + \lambda_k x_k + \lambda_{k+1} x_{k+1}) \\ & = (1 - \lambda_{k+1}) f\left(\frac{\lambda_1}{1 - \lambda_{k+1}} x_1 + \cdots + \frac{\lambda_k}{1 - \lambda_{k+1}} x_k \right) + \lambda_{k+1} f(x_{k+1}) \\ & \overset{\text{IO}}{\leq} (1 - \lambda_{k+1}) \left(\frac{\lambda_1}{1 - \lambda_{k+1}} f(x_1) + \cdots + \frac{\lambda_k}{1 - \lambda_{k+1}} f(x_k) \right) + \lambda_{k+1} f(x_{k+1}) \\ & = \lambda_1 f(x_1) + \cdots + \lambda_k f(x_k) + \lambda_{k+1} f(x_{k+1}). \end{aligned}$$

Siis väite on tosi kaikilla kokonaisluvuilla $n \geq 2$. $\square$

Kun valitaan kaikki λ:t yhtä suuriksi, saadaan merkittävä seurauslause.

Korollaari 4.3. *Jos $f : I \to \mathbb{R}$ on konveksi ja $x_1, x_2, \ldots, x_n \in I$, niin tällöin*

$$f\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right) \leq \frac{f(x_1) + f(x_2) + \cdots + f(x_n)}{n}.$$

Joskus Jensenin epäyhtälö esitetäänkin tässä muodossa.

Jensenin epäyhtälö on yksi matematiikan tärkeimmistä epäyhtälöistä. Sen avulla todistetaan lukuisia muita epäyhtälöjä kuten esimerkiksi aritmeettis-geometrinen epäyhtälö.

Tässä luvussa esiteltiin Jensenin epäyhtälön äärellinen muoto. Se voidaan kuitenkin kirjoittaa myös mittateorian tai todennäköisyysteorian avulla. Eri muotoilut soveltuvat eri konteksteihin ja laajentavat epäyhtälön käyttöaluetta. Epäyhtälöä sovelletaankin monilla tieteen ja tekniikan aloilla kuten optimoinnissa, taloustieteessä ja todennäköisyyslaskennassa.

5 Lopuksi

Tässä tutkielmassa on käsitelty konveksien funktioiden tavallisia ja joitain tärkeimpiä ominaisuuksia. Kyseessä on rakenteeltaan erityinen funktio, jonka ominaisuuksia voidaan hyödyntää eri aloilla. Tässä työssä määritelmät ja ominaisuudet on yksinkertaisuuden vuoksi esitetty yhden muuttujan reaalifunktiolla. Konveksisuus voidaan kuitenkin määritellä vastaavasti myös useamman muuttujan funktioille sekä joukoille. Esimerkiksi Lauritzenin kirjan [3] luku 9 käsittelee useamman muuttujan konvekseja funktioita helposti lähestyttävällä tavalla.

Tämän työn luvussa 3.5 esiteltiin globaali minimi. Kun tämä konveksien funktioiden erityinen ominaisuus laajennetaan käsittämään myös monimutkaisempia konvekseja malleja, se on hyödyksi monenlaisten optimointiongelmien ja minimointitehtävien ratkaisussa. Konveksianalyysin avulla optimointiongelmien ratkaisusta tulee tehokkaampaa, kun lokaalin minimin löytyminen tarkoittaa myös globaalin minimin löytymistä.

Konvekseja funktioita hyödynnetäänkin laajasti erilaisten teknisten, taloudellisten ja tieteellisten ongelmien ratkaisemiseen, kun tavoitteena on minimoida kustannukset, energia, aika tai jokin muu resurssi. Konvekseja malleja käytetään esimerkiksi logistiikassa, reittien suunnittelussa, varastonhallinnassa ja tuotantoprosessien optimoinnissa. Konveksioptimoinnista voit lukea lisää esimerkiksi Dimitri Bertsekasin kirjasta [7], joka antaa syvällisen johdatuksen asiaan.

Konvekseja funktioita voidaan hyödyntää myös koneoppimisessa, kun mallien kouluttamisessa pyritään minimoimaan virhefunktion arvo. Koneoppimisesta ja konveksioptimoinnista voit lukea Elad Hazadin kirjasta [8].

Molemmissa mainituissa teoksissa [7] ja [8] esitellään aligradientin (engl. subgradient) käsite, joka laajentaa gradientin käsitteen ei-derivoituville funktioille. Gradientti kertoo funktion muutosnopeuden ja muutoksen suunnan pisteessä, ja se edellyttää funktion derivoituvuutta. Aligradienttimenetelmällä voidaan löytää konveksien funktioiden minimit myös silloin, kun funktio ei ole derivoituva.

Kiinnostavaa on myös se, että useissa talous- ja rahoitusalan ilmiöissä esiintyy luonnostaan konveksisuutta [9]. Siksi monia rahoitusalan ilmiöitä ja optimointiongelmia mallinnetaan ja ratkotaan konveksianalyysillä. Näitä voivat olla esimerkiksi riskien minimointi, salkunhallinta ja optimaalisten sijoitusstrategioiden kehittäminen.

Viitteet

[1] Andrew Browder. *Mathematical Analysis: An Introduction.* Springer New York, 2001. ISBN: 978-1-4612-6879-6.

[2] Heikki Apiola. ”Polynomit, interpolaatio ja funktion approksimointi”. *Matematiikkalehti Solmu* 3 (2004). ISSN: 1458-8048. URL: https://matematiikkalehtisolmu.fi/2004/3/.

[3] Niels Lauritzen. *Undergraduate Convexity.* 1. painos. World Scientific Publishing Company, 2013. ISBN: 9814412511.

[4] Constantin P. Niculescu ja Lars-Erik Persson. *Convex Functions and Their Applications: A Contemporary Approach.* 2. painos. Springer International Publishing, 2018. ISBN: 3-319-78337-8.

[5] P. E. Kopp. *Analysis / P.E. Kopp.* eng. Modular mathematics series. Oxford: Elsevier, 1996. ISBN: 0080928722.

[6] Markku Halmetoja. ”Epäyhtälöistä, osa 2”. *Matematiikkalehti Solmu* 3 (2010). ISSN: 1458-8048. URL: https://matematiikkalehtisolmu.fi/2010/3/.

[7] Dimitri P. Bertsekas. *Convex Optimization Theory.* 1. painos. Athena Scientific, 2009. ISBN: 978-1-886529-31-1.

[8] Elad Hazad. ”Introduction to Online Convex Optimization”. *Foundations and Trends® in Optimization* 2.3-4 (2015), s. 157–325. DOI: 10.1561/2400000013.

[9] Teemu Pennanen. ”Introduction to convex optimization in financial markets”. *Mathematical programming* 134.1 (2012), s. 157–186. ISSN: 0025-5610. DOI: 10.1007/s10107-012-0573-4.