Epäyhtälöistä, osa 2

Markku Halmetoja
Mäntän lukio

Välillä I määriteltyä funktiota sanotaan konveksiksi, jos sen kuvaaja on alaspäin kupera, eli jos kuvaajan mitkä tahansa kaksi pistettä yhdistävä jana ei alita kuvaajaa. Esimerkiksi funktiolla $f(x) = x^2$ on tämä ominaisuus. Konveksit funktiot toteuttavat Jensenin[1] epäyhtälön. Se on eräänlainen yleisepäyhtälö, jonka avulla voidaan todistaa lukuisa joukko muita epäyhtälöitä. Jensenin epäyhtälö on tämän kirjoituksen pääasia, mutta siihen pääsemiseksi on ensin käsiteltävä konveksien funktioiden ominaisuuksia. Aivan ensimmäiseksi on yllä annettu konveksisuuden geometrinen määritelmä muunnettava analyyttiseksi. Kirjoitukseen sisältyy muutama harjoitus, joiden pohtiminen syventää asian ymmärtämistä. Niiden ratkaisut julkaistaan myöhemmin Solmun nettisivulla.

Konveksisuuden määritelmä

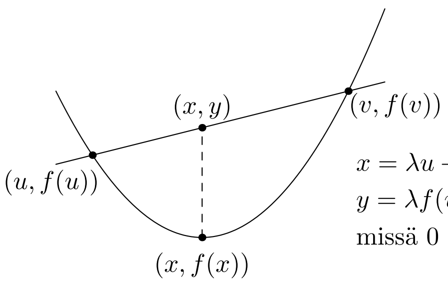
Olkoon f välillä I määritelty funktio ja $[u,v] \subseteq I$. Sen kuvaajan pisteet $(u,f(u))$ ja $(v,f(v))$ yhdistävän janan yhtälö on

$$y = y(\lambda) = \lambda f(u) + (1-\lambda)f(v),$$

missä $0 \le \lambda \le 1$. Näillä λ:n arvoilla myös

$$x = \lambda u + (1-\lambda)v \in [u,v].$$

[1] Johan Jensen (1859–1925), tanskalainen matemaatikko.



$x = \lambda u + (1-\lambda)v$ ja
$y = \lambda f(u) + (1-\lambda)f(v)$,
missä $0 \le \lambda \le 1$.

Jos kuvaaja on alaspäin kupera, niin piste $(x,f(x))$ ei ylitä mainittua janaa millään $x \in [u,v]$. Konveksisuuden analyyttinen määritelmä saadaan kirjoittamalla tämä geometrinen ehto epäyhtälöksi.

Määritelmä. Olkoon I reaalilukuväli ja f siinä määritelty funktio. Se on *konveksi*, jos epäyhtälö

$$f\big(\lambda u + (1-\lambda)v\big) \le \lambda f(u) + (1-\lambda)f(v) \qquad (1)$$

toteutuu kaikilla $u,v \in I$ ja $\lambda \in\,]0,1[$. Se on *aidosti konveksi*, jos

$$f\big(\lambda u + (1-\lambda)v\big) < \lambda f(u) + (1-\lambda)f(v) \qquad (2)$$

kaikilla $\lambda \in\,]0,1[$ ja kaikilla keskenään erisuurilla $u,v \in I$.

Huomautus. Parametri λ rajoitetaan määritelmässä avoimelle välille $]0,1[$, sillä tämän välin päätepisteissä epäyhtälö (1) on joka tapauksessa voimassa.

Jos erisuuruus epäyhtälöissä (1) tai (2) on vastakkaiseen suuntaan, niin funktio f on *konkaavi* tai *aidosti konkaavi* välillä I. Selvästi funktio f on (aidosti) konkaavi, jos ja vain jos $-f$ on (aidosti) konveksi, joten jatkossa voidaan rajoittua pelkästään konveksisuuteen.

Seuraavissa harjoituksissa ja esimerkissä sovelletaan konveksisuuden määritelmää.

Esim. Funktio $f(x) = |x|$ on konveksi, sillä kolmioepäyhtälöstä seuraa

$$|\lambda u + (1-\lambda)v| \le \lambda|u| + (1-\lambda)|v|$$

kaikilla $u,v \in \mathbb{R}$ ja $\lambda \in]0,1[$.

Harjoituksia

1. Osoita, että funktio $f(x) = x^2$ on aidosti konveksi.

2. Osoita, että funktio $g(x) = x^{-1}$, $x > 0$, on aidosti konveksi.

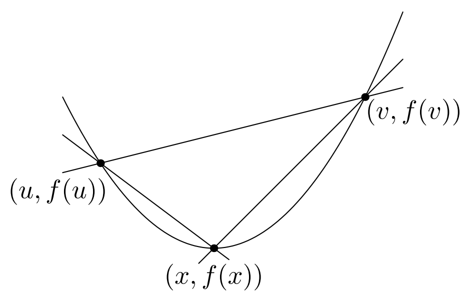
3. Keksi esimerkki konveksista funktiosta $f : [0,1] \to \mathbb{R}$, joka on **a)** jatkuva, **b)** epäjatkuva.

4. Osoita, että jos f ja g ovat välillä I konvekseja funktioita ja jos a sekä b ovat ei-negatiivisia, niin lineaarikombinaatio $af + bg$ on konveksi.

Konveksisuuden luonnehdintaa

Funktion konveksisuus on alkeellisia tapauksia lukuunottamatta varsin hankalaa todistaa suoraan määritelmän perusteella. Seuraavassa johdetaan eräitä yhtäpitäviä ja riittäviä ehtoja konveksisuudelle.

Jos f on välillä I määritelty konveksi funktio ja $x \in]u,v[\subset I$, niin kuvioon



piirrettyjen sekanttien kulmakertoimet toteuttavat kaksoisepäyhtälön

$$\frac{f(x)-f(u)}{x-u} \le \frac{f(v)-f(u)}{v-u} \le \frac{f(v)-f(x)}{v-x}, \qquad (3)$$

mikä johtaa tärkeään konveksisuuden luonnehdintaan.

Lause 1. Välillä I määritelty funktio f on konveksi, jos ja vain jos kaikille ehdon $u < x < v$ toteuttaville välin I luvuille on voimassa mikä tahansa kaksoisepäyhtälöstä (3) saatava erotusosamääriä koskeva epäyhtälö.

Todistus. Käsitellään tapaus

$$\frac{f(x)-f(u)}{x-u} \le \frac{f(v)-f(x)}{v-x},$$

ja muut jätetään harjoitustehtäviksi.

Olkoot $u,x,v \in I$ ja $u < x < v$. Tällöin on olemassa $\lambda \in]0,1[$ siten, että $x = \lambda u + (1-\lambda)v$. Koska $x - u > 0$, $v - x > 0$, $v - u > 0$,

$$\lambda = \frac{v-x}{v-u} \quad \text{ja} \quad 1-\lambda = \frac{x-u}{v-u},$$

ovat epäyhtälöt

$$f(x) \le \lambda f(u) + (1-\lambda)f(v)$$
$$\lambda f(x) + (1-\lambda)f(x) \le \lambda f(u) + (1-\lambda)f(v)$$
$$\lambda\big(f(x)-f(u)\big) \le (1-\lambda)\big(f(v)-f(x)\big)$$
$$\frac{v-x}{v-u}\big(f(x)-f(u)\big) \le \frac{x-u}{v-u}\big(f(v)-f(x)\big)$$
$$\frac{f(x)-f(u)}{x-u} \le \frac{f(v)-f(x)}{v-x}$$

keskenään yhtäpitäviä, joten lause on todistettu.

Seuraus. Funktio $f : I \to \mathbb{R}$ on aidosti konveksi, jos ja vain jos kaikille ehdon $u < t < v$ toteuttaville välin I luvuille on voimassa

$$\frac{f(x)-f(u)}{x-u} < \frac{f(v)-f(x)}{v-x}.$$

Derivoituvan funktion konveksisuudelle saadaan helpohko kriteeri.

Lause 2. Olkoon $f : I \to \mathbb{R}$ derivoituva. Jos ja vain jos derivaatta on (aidosti) kasvava, niin f on (aidosti) konveksi.

Todistus. Olkoot x,t ja y ehdon $x < t < y$ toteuttavia I:n lukuja. Väliarvolauseen mukaan on olemassa $\xi_1 \in]x,t[$ ja $\xi_2 \in]t,y[$ siten, että

$$\frac{f(t)-f(x)}{t-x} = f'(\xi_1) \quad \text{ja} \quad \frac{f(y)-f(t)}{y-t} = f'(\xi_2).$$

Koska $\xi_1 < \xi_2$ ja f' on kasvava, on $f'(\xi_1) \le f'(\xi_2)$, joten

$$\frac{f(t)-f(x)}{t-x} = f'(\xi_1) \le f'(\xi_2) = \frac{f(y)-f(t)}{y-t},$$

eli väite on tosi lauseen 1 perusteella. Jos f' on aidosti kasvava, niin epäyhtälöstä $\xi_1 < \xi_2$ seuraa $f'(\xi_1) < f'(\xi_2)$, joten

$$\frac{f(t)-f(x)}{t-x} = f'(\xi_1) < f'(\xi_2) = \frac{f(y)-f(t)}{y-t},$$

eli väite on tosi lauseen 1 seurauksen perusteella.

Seuraus. Jos f on kahdesti derivoituva ja jos f'' on ei-negatiivinen, niin f on konveksi. Jos f'' on ei-negatiivinen ja sillä on enintään erillisiä nollakohtia, niin f on aidosti konveksi.

Esim. Funktiot $f(x) = \mathrm{e}^x$ ja $g(x) = -\ln x$ ovat aidosti konvekseja, sillä $f''(x) = \mathrm{e}^x > 0$ kaikilla $x \in \mathbb{R}$ ja $g''(x) = x^{-2} > 0$ kaikilla $x \in \mathbb{R}_+$.

Esim. Funktiot $f_n(x) = x^{2n}$, $n \in \mathbb{Z}_+$, ovat aidosti konvekseja. Funktiot $g_n(x) = x^{2n+1}$, $n \in \mathbb{Z}_+$, ovat aidosti konvekseja välillä $[0,\infty[$.

Harjoituksia

5. Osoita, että välillä I määritelty funktio f on konveksi, jos ja vain jos

$$\frac{f(v)-f(u)}{v-u} \le \frac{f(v)-f(x)}{v-x}$$

kaikille ehdon $u < x < v$ toteuttaville välin I luvuille.

6. Osoita, että välillä I määritelty funktio f on konveksi, jos ja vain jos

$$\frac{f(x)-f(u)}{x-u} \le \frac{f(v)-f(u)}{v-u}$$

kaikille ehdon $u < x < v$ toteuttaville välin I luvuille.

7. Todista, että *avoimella välillä* määritelty konveksi funktio on jatkuva. Ohje: Jos I on avoin väli ja $x_0 \in I$, niin on olemassa luvut $a,b,u,v \in I$ siten, että $a < u < x_0 < v < b$. Valitse $x \in\,]u,x_0[$, ja osoita kaksoisepäyhtälöä (3) soveltaen, että $f(x) \to f(x_0)$, kun $x \to x_0$ vasemmalta. Osoita sitten samalla tavalla, että $f(x) \to f(x_0)$, kun $x \to x_0$ oikealta.

8. Osoita, että funktio $f : \mathbb{R}_+ \to \mathbb{R}$, $f(x) = x \ln x$, on aidosti konveksi.

Jensenin epäyhtälö

Konveksisuuden määritelmä voidaan kirjoittaa seuraavasti:

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

$$f(\lambda_1 x_1 + \lambda_2 x_2) \le \lambda_1 f(x_1) + \lambda_2 f(x_2)$$

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

Jensenin epäyhtälö yleistää määritelmän epäyhtälön useammalle λ:n ja x:n arvolle.

Lause 3. Olkoon $f : I \to \mathbb{R}$ konveksi funktio ja $\lambda_1, \lambda_2, \ldots, \lambda_n$ positiivisia lukuja, joiden summa on 1. Tällöin kaikilla $x_1, x_2, \ldots, x_n \in I$

$$f(\lambda_1 x_1 + \ldots + \lambda_n x_n) \le \lambda_1 f(x_1) + \ldots + \lambda_n f(x_n). \quad (4)$$

Todistus. Väite on siis tosi, kun $n = 2$. Jos se on tosi $(n-1)$:lle luvulle, niin, koska

$$\begin{aligned}&\lambda_1 x_1 + \ldots + \lambda_{n-1} x_{n-1} + \lambda_n x_n \\ &= (1-\lambda_n)\left(\frac{\lambda_1}{1-\lambda_n} x_1 + \ldots + \frac{\lambda_{n-1}}{1-\lambda_n} x_{n-1}\right) + \lambda_n x_n,\end{aligned}$$

on

$$\begin{aligned}&f(\lambda_1 x_1 + \ldots + \lambda_{n-1} x_{n-1} + \lambda_n x_n) \\ &\le (1-\lambda_n) f\left(\frac{\lambda_1}{1-\lambda_n} x_1 + \ldots + \frac{\lambda_{n-1}}{1-\lambda_n} x_{n-1}\right) \\ &\qquad + \lambda_n f(x_n) \\ &\le (1-\lambda_n)\left(\frac{\lambda_1}{1-\lambda_n} f(x_1) + \ldots + \frac{\lambda_{n-1}}{1-\lambda_n} f(x_{n-1})\right) \\ &\qquad + \lambda_n f(x_n) \\ &= \lambda_1 f(x_1) + \ldots + \lambda_{n-1} f(x_{n-1}) + \lambda_n f(x_n),\end{aligned}$$

joten väite on induktioperiaatteen nojalla tosi kaikilla $n \in \mathbb{Z}$, $n \ge 2$.

Valitsemalla kaikki λ:t yhtä suuriksi saadaan monessa yhteydessä käyttökelpoinen seurauslause.

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

$$f\left(\frac{x_1 + x_2 + \ldots + x_n}{n}\right) \le \frac{f(x_1) + f(x_2) + \ldots + f(x_n)}{n}. \quad (5)$$

Milloin epäyhtälöissä (4) ja (5) vallitsee yhtäsuuruus? Jos esimerkiksi f on ensimmäisen asteen polynomifunktio (se on konveksi), niin yhtäsuuruus on voimassa kaikilla $x_1, \ldots, x_n \in \mathbb{R}$. Kysymys yhtäsuuruuden voimassaolosta on siis kiinnostava vain, jos funktio on aidosti konveksi.

Lause 4. Olkoon $f : I \to \mathbb{R}$ aidosti konveksi, $x_1, \ldots, x_n \in I$ ja $\lambda_1 \ldots, \lambda_n$ positiivisia lukuja, joiden summa on 1. Tällöin

$$f(\lambda_1 x_1 + \ldots + \lambda_n x_n) = \lambda_1 f(x_1) + \ldots + \lambda_n f(x_n) \quad (6)$$

jos ja vain jos $x_1 = x_2 = \ldots = x_n$.

Todistus. Selvästi (6) on voimassa, jos $x_1 = x_2 = \ldots = x_n$. Oletetaan nyt, että (6) on voimassa, ja tehdään vastaoletus, että luvut $x_1, \ldots, x_n$ eivät ole kaikki yhtä suuria. Voidaan rajoituksetta olettaa, että ne ovat pienimmästä alkaen suuruusjärjestyksessä. Tällöin x_1 on niistä pienin. Olkoon x_k ensimmäinen siitä poikkeava luku. Siis

$$x_1 = \ldots = x_{k-1} < x_k \le \ldots \le x_n.$$

Merkitään vielä $\mu = \lambda_1 + \ldots + \lambda_{k-1}$, jolloin $\lambda_k + \ldots + \lambda_n = 1 - \mu$. Koska f on aidosti konveksi ja

$$x_1 < \frac{\lambda_k}{1-\mu} x_k + \ldots + \frac{\lambda_n}{1-\mu} x_n,$$

on

$$\begin{aligned} & f\big(\lambda_1 x_1 + \ldots + \lambda_n x_n\big) \\ &= f\left(\mu x_1 + (1-\mu)\left(\frac{\lambda_k}{1-\mu}x_k + \ldots + \frac{\lambda_n}{1-\mu}x_n\right)\right) \\ &< \mu f(x_1) + (1-\mu) f\left(\frac{\lambda_k}{1-\mu}x_k + \ldots + \frac{\lambda_n}{1-\mu}x_n\right) \\ &\le (\lambda_1 + \ldots + \lambda_{k-1}) f(x_1) + \lambda_k f(x_k) + \ldots + \lambda_n f(x_n) \\ &= \lambda_1 f(x_1) + \lambda_2 f(x_2) + \ldots + \lambda_n f(x_n), \end{aligned}$$

mikä on ristiriidassa oletuksen (6) kanssa. Siis ei ole ensimmäistäkään lukua x_k, joka poikkeaisi pienimmästä luvusta x_1, joten kaikki x:t ovat samoja.

Muodostamalla sopivia aidosti konvekseja funktiota voit Jensenin epäyhtälön avulla keksiä aivan uusia, ennennäkemättömiä epäyhtälöitä! Seuraava lienee kuitenkin yleisesti tunnettu.

Lause 5. Olkoot $u_1, \ldots, u_n$ positiivisia lukuja ja $\mathcal{A}$ niiden aritmeettinen keskiarvo. Tällöin

$$\mathcal{A}^{\mathcal{A}} \le \sqrt[n]{u_1^{u_1} u_2^{u_2} \ldots u_n^{u_n}},$$

missä yhtäsuuruus on voimassa, jos ja vain jos $u_1 = \ldots = u_n$.

Todistus. Positiivisille luvuille määritelty funktio $f(x) = x \ln x$ on aidosti konveksi, joten Jensenin epäyhtälön (5) mukaan

$$\left(\frac{u_1 + \ldots + u_n}{n}\right) \ln \left(\frac{u_1 + \ldots + u_n}{n}\right) \le \frac{1}{n} u_1 \ln u_1 + \ldots + \frac{1}{n} u_n \ln u_n,$$

mistä väite seuraa.

Harjoituksia

9. Osoita, että jos positiivisten lukujen $u_1, \ldots, u_n$ aritmeettinen keskiarvo on 1, niin

$$u_1 u_2 \ldots u_n \le 1 \le u_1^{u_1} u_2^{u_2} \ldots u_n^{u_n},$$

ja että yhtäsuuruudet ovat voimassa, jos ja vain jos $u_k = 1$ kaikilla $k \in \{1, \ldots, n\}$.

10. Osoita **a)** edellisen tehtävän erikoistapauksena, **b)** pelkästään lukion oppimäärään tukeutuen, että

$$(1-x)^{1-x}(1+x)^{1+x} > 1$$

kaikilla $x \in]0,1[$.

11. Osoita, että jos a,b,c ovat positiivisia ja $a+b+c = 1$, niin

$$a^2 + b^2 + c^2 \ge \tfrac{1}{3}.$$

Milloin yhtäsuuruus on voimassa?

12. Osoita, että jos luvut u_k ovat positiivisia ja jos $u_1 + u_2 + \ldots + u_n = 1$, niin

$$\big(u_1^2 + u_2^2 + \ldots + u_n^2\big)^2 \le u_1^3 + u_2^3 + \ldots + u_n^3.$$

Milloin yhtäsuuruus on voimassa?

Keskiarvoepäyhtälöitä

Kirjoituksen ykkösosassa (Solmu 2/2010) tutkittiin keskiarvojen $\mathcal{H}$, $\mathcal{G}$, $\mathcal{A}$ ja $\mathcal{C}$ suuruusjärjestystä. Jensenin epäyhtälön avulla voidaan todistaa sama järjestys myös vastaaville painotetuille keskiarvoille. Ne määritellään yhtälöillä

$$\begin{aligned} \mathcal{H}_\lambda &= \frac{1}{\frac{\lambda_1}{u_1} + \ldots + \frac{\lambda_n}{u_n}} \\ \mathcal{G}_\lambda &= u_1^{\lambda_1} \ldots u_n^{\lambda_n} \\ \mathcal{A}_\lambda &= \lambda_1 u_1 + \ldots + \lambda_n u_n \\ \mathcal{C}_\lambda &= \frac{\lambda_1 u_1^2 + \ldots + \lambda_n u_n^2}{\lambda_1 u_1 + \ldots + \lambda_n u_n} \end{aligned}$$

missä u- ja λ-luvut ovat positiivisia ja λ-lukujen summa on 1. Järjestys todistuu kolmessa vaiheessa.

Vaihe 1. $\mathcal{G}_\lambda \le \mathcal{A}_\lambda$. Olkoot $u_1, \ldots, u_n$ positiivisia lukuja ja $\lambda_1, \ldots, \lambda_n$ positiivisia lukuja, joiden summa on 1. On olemassa luvut $x_1, \ldots, x_n$ siten, että $u_k = e^{x_k}$ kaikilla $k \in \{1, 2, \ldots, n\}$. Koska funktio $f(x) = e^x$ on aidosti konveksi, on Jensenin epäyhtälön mukaan

$$e^{\lambda_1 x_1 + \ldots + \lambda_n x_n} \le \lambda_1 e^{x_1} + \ldots + \lambda_n e^{x_n}.$$

Yhtälöiden $u_k = e^{x_k}$ avulla tämä sievenee muotoon

$$\mathcal{G}_\lambda = u_1^{\lambda_1} u_2^{\lambda_2} \ldots u_n^{\lambda_n} \le \lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_n u_n = \mathcal{A}_\lambda.$$

Vaihe 2. $\mathcal{H}_\lambda \le \mathcal{G}_\lambda$. Olkoot u- ja λ-luvut kuten yllä. Soveltamalla edellistä epäyhtälöä u-lukujen käänteislukuihin saadaan

$$\frac{1}{u_1^{\lambda_1}} \cdots \frac{1}{u_n^{\lambda_n}} \le \frac{\lambda_1}{u_1} + \ldots + \frac{\lambda_n}{u_n},$$

mistä välittömästi seuraa

$$\mathcal{H}_\lambda = \frac{1}{\frac{\lambda_1}{u_1} + \ldots + \frac{\lambda_n}{u_n}} \le u_1^{\lambda_1} \ldots u_n^{\lambda_n} = \mathcal{G}_\lambda.$$

Vaihe 3. $\mathcal{A}_\lambda \le \mathcal{C}_\lambda$. Olkoot u- ja λ-luvut edelleen kuten yllä. Funktio $g(x) = x^2$ on aidosti konveksi, joten Jensenin epäyhtälöä soveltaen saadaan

$$(\lambda_1 u_1 + \ldots + \lambda_n u_n)^2 \le \lambda_1 u_1^2 + \ldots + \lambda_n u_n^2,$$

mistä seuraa

$$\mathcal{A}_\lambda = \lambda_1 u_1 + \ldots + \lambda_n u_n \leq \frac{\lambda_1 u_1^2 + \ldots + \lambda_n u_n^2}{\lambda_1 u_1 + \ldots + \lambda_n u_n} = \mathcal{C}_\lambda.$$

Siis

$$\mathcal{H}_\lambda \leq \mathcal{G}_\lambda \leq \mathcal{A}_\lambda \leq \mathcal{C}_\lambda.$$

Koska funktiot $f(x) = \mathrm{e}^x$ ja $g(x) = x^2$ ovat aidosti konvekseja, vallitsee tässä ketjussa yhtäsuuruus, jos ja vain jos luvut u_k ovat kaikki keskenään yhtäsuuria.

Lähdeluettelo

[1] J.Michael Steele, *The Cauchy-Schwarz Master Class, An Introduction to the Art of Mathematical Inequatilies*, Cambridge University Press, 2007.

[2] The MacTutor History of Mathematics archive, `http://www-groups.dcs.st-and.ac.uk/~history/`

Diplomitehtävien oheislukemistoa

Osoitteessa `http://solmu.math.helsinki.fi/diplomi.html` on diplomitehtäville oheislukemistoa, joka varmaan kiinnostaa muitakin kuin diplomien tekijöitä:

Desimaaliluvut, mitä ne oikeastaan ovat?

Murtolukujen laskutoimituksia

Hiukan osittelulaista

Lausekkeet, kaavat ja yhtälöt

Äärettömistä joukoista

Erkki Luoma-aho: Matematiikan peruskäsitteiden historia

Gaussin jalanjäljissä

K. Väisälä: Algebra