Epäyhtälöistä, osa 1

Markku Halmetoja
Mäntän lukio

Tässä kirjoituksessa tarkastellaan eräitä koulumatematiikan keinoin todistuvia epäyhtälöitä. Erityisesti keskitytään niihin, joiden oikeaksi todistaminen perustuu neliön ei-negatiivisuuteen. Kirjoitukseen sisältyy muutama lukijan aktivoimiseksi tarkoitettu harjoitustehtävä. Myöhemmin ehkä ilmestyvässä toisessa osassa tarkastellaan yhden muuttujan konvekseja funktioita, jolloin saadaan menetelmiä hieman hankalampien epäyhtälöiden käsittelyyn.

Neliön ei-negatiivisuus

Viimeistään lukio-opiskelun alussa käy selväksi, että $x^2 \geq 0$ kaikilla $x \in \mathbb{R}$, ja yhtäsuuruus on voimassa, jos ja vain jos $x = 0$. Tämän perusepäyhtälön avulla voidaan esimerkiksi todistaa, että positiivisen luvun ja sen käänteisluvun summa on vähintään 2, mikä ei ole aivan ilmeinen asia, jos tarkasteltava luku on lähellä ykköstä. Seuraavassa esimerkissä todistetaan tämä väite.

Esim. Osoita, että jos u ja v ovat positiivisia lukuja, niin

$$\frac{u}{v} + \frac{v}{u} \geq 2,$$

missä yhtäsuuruus on voimassa, jos ja vain jos $u = v$.

Ratk. Epäyhtälöt

$$\frac{u}{v} + \frac{v}{u} \geq 2 \ \mid \cdot uv \quad (uv > 0, \text{kertominen sallittu})$$

$$u^2 + v^2 \geq 2uv$$

$$u^2 - 2uv + v^2 \geq 0$$

$$(u - v)^2 \geq 0$$

ovat keskenään yhtäpitäviä, ja koska viimeinen niistä on tosi, ovat kaikki tosia. Selvästi yhtäsuuruus on voimassa, jos ja vain jos $u = v$.

Epäyhtälön todistaminen tapahtuu usein niin, että johdetaan todistettavan epäyhtälön kanssa yhtäpitäviä epäyhtälöitä, kunnes päästään sellaiseen, mikä nähdään todeksi esimerkiksi neliön ei-negatiivisuuden perusteella. Näin meneteltiin edellisessä esimerkissä. Joskus voidaan epäyhtälön toista puolta sieventämällä päästä lopulta sellaiseen epäyhtälöön, mikä nähdään todeksi jonkin tunnetun asian perusteella. Seuraavissa harjoituksissa sovelletaan näitä periaatteita.

Harjoituksia

1. Kahden positiivisen luvun harmoninen, geometrinen, aritmeettinen ja kontraharmoninen keskiarvo määritellään yhtälöillä

$$\mathcal{H} = \frac{2}{\frac{1}{u}+\frac{1}{v}},\ \mathcal{G} = \sqrt{uv},\ \mathcal{A} = \frac{u+v}{2} \text{ ja } \mathcal{C} = \frac{u^2+v^2}{u+v}.$$

Todista keskiarvojen suuruusjärjestys $\mathcal{H} \leq \mathcal{G} \leq \mathcal{A} \leq \mathcal{C}$. Milloin yhtäsuuruus on voimassa?

2. Osoita, että jos u_1 ja u_2 ovat positiivisia lukuja, joille $u_1 + u_2 = 1$, niin

$$\frac{1}{u_1} + \frac{1}{u_2} \geq 4.$$

Milloin vallitsee yhtäsuuruus?

3. Osoita, että jos u_1, u_2 ja u_3 ovat positiivisia lukuja, joille $u_1 + u_2 + u_3 = 1$, niin

$$\frac{1}{u_1} + \frac{1}{u_2} + \frac{1}{u_3} \geq 9.$$

Milloin vallitsee yhtäsuuruus?

4. Yleistä tehtävien 2. ja 3. epäyhtälöt mielivaltaiselle määrälle positiivisia lukuja, ja todista näin saamasi epäyhtälö yhtäsuuruusehtoineen.

Cauchyn-Bunjakovskin-Schwarzin epäyhtälö

Matemaattiset teoreemat on tapana nimetä löytäjänsä mukaan. Jos useat henkilöt päätyvät toisistaan riippumatta samoihin tuloksiin, on oikeudenmukaista nimetä tulos kaikkien keksijöiden mukaan. Cauchy[1] löysi ensimmäisenä nyt tarkasteltavan epäyhtälön, ja Schwarz[2] yleisti sen integraaleja koskevaksi. Siitä puhutaan yleisesti Cauchyn ja Schwarzin nimillä. Myöhemmin on osoittautunut, että Bunjakovski[3] oli todistanut epäyhtälön integraalimuodon 25 vuotta ennen Schwarzia, ks. [3]. Siksi on oikein käyttää otsikossa olevaa nimihirviötä, mikä jatkossa lyhennetään CBS-epäyhtälöksi. Tämä tärkeä epäyhtälö voidaan todistaa lukion pitkän matematiikan kakkoskurssin tiedoilla.

Lause. *CBS-epäyhtälö.* Reaaliluvut $u_1, u_2, \ldots, u_n$ ja $v_1, v_2, \ldots, v_n$ toteuttavat epäyhtälön

$$|u_1v_1 + \ldots + u_nv_n| \leq \sqrt{u_1^2 + \ldots + u_n^2}\sqrt{v_1^2 + \ldots + v_n^2}, \quad (1)$$

missä yhtäsuuruus on voimassa, jos ja vain jos on olemassa t_0, jolle $u_\imath t_0 = v_\imath$ kaikilla $\imath \in \{1, 2, \ldots, n\}$, tai kaikki u-luvut ovat nollia tai kaikki v-luvut ovat nollia.

Todistus. On selvää, että (1):ssä vallitsee yhtäsuuruus, jos kaikki u-luvut tai kaikki v-luvut ovat nollia. Olkoot $u_1, \ldots, u_n$ ja $v_1, \ldots, v_n$ reaalilukuja, joista kaikki eivät ole nollia. Voidaan rajoituksetta olettaa, että vähintään yksi u-luku on nollasta eroava. Tällöin funktio

$$g(t) = (u_1t - v_1)^2 + \ldots + (u_nt - v_n)^2 \quad (2)$$

voidaan kehittää toisen asteen polynomiksi

$$g(t) = t^2\sum_{\imath=1}^{n} u_\imath^2 - 2t\sum_{\imath=1}^{n} u_\imath v_\imath + \sum_{\imath=1}^{n} v_\imath^2.$$

Sillä on enintään yksi nollakohta, sillä neliöt $(u_\imath t - v_\imath)^2$ ovat ei-negatiivisia. Diskriminanttitarkastelulla saadaan

$$\left(\sum_{\imath=1}^{n} u_\imath v_\imath\right)^2 \leq \sum_{\imath=1}^{n} u_\imath^2 \sum_{\imath=1}^{n} v_\imath^2,$$

mistä (1) seuraa. Selvästi yhtäsuuruus vallitsee, jos ja vain jos kaikki neliöt (2):ssa häviävät samalla t:n arvolla, eli jos ja vain jos on olemassa t_0, jolle $u_\imath t_0 = v_\imath$ kaikilla $\imath \in \{1, 2, \ldots, n\}$. Muussa tapauksessa yhtäsuuruus on voimassa ainoastaan, jos kaikki u-luvut tai kaikki v-luvut ovat nollia.

Harjoitus

5. Jos kaikki u-luvut ovat nollasta eroavia, niin CBS-epäyhtälön yhtäsuuruusehto voidaan kirjoittaa muotoon

$$\frac{v_1}{u_1} = \frac{v_2}{u_2} = \ldots = \frac{v_n}{u_n}.$$

Osoita, että jos u-lukujen summa ei ole nolla, niin

$$\frac{v_\imath}{u_\imath} = \frac{v_1 + v_2 + \ldots + v_n}{u_1 + u_2 + \ldots + u_n}$$

kaikilla $\imath \in \{1, 2, \ldots, n\}$.

Sovellus

Ensinäkemältä (1) vaikuttaa sekavalta, mutta jos oikealla puolella olevat neliösummat ovat positiivisia, niin se voidaan kirjoittaa kaksoisepäyhtälöksi

$$-1 \leq \frac{u_1v_1 + \ldots + u_nv_n}{\sqrt{u_1^2 + \ldots + u_n^2}\sqrt{v_1^2 + \ldots + v_n^2}} \leq 1, \quad (3)$$

missä näkyy tuttuja piirteitä. Jos $n = 2$ tai $n = 3$, niin keskellä oleva lauseke esittää kahden vektorin välisen kulman γ kosinia, ja kaksoisepäyhtälö kertoo sen, että $-1 \leq \cos\gamma \leq 1$. Voiko tällaista tulkintaa tehdä, jos $n > 3$? Kaksi- ja kolmiulotteisen avaruuden vektorit voidaan samastaa lukupareihin ja lukukolmikoihin,

$$\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} = (u_1, u_2)$$

[1] Augustin Louis Cauchy (1789 – 1857), ranskalainen matemaatikko.

[2] Hermann Amandus Schwarz (1843 – 1921), saksalainen matemaatikko.

[3] Viktor Jakovlevitš Bunjakovski (1804 – 1889), venäläinen matemaatikko.

ja

$$\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k} = (v_1, v_2, v_3).$$

Mikään ei estä määrittelemästä yleisesti n-ulotteisen avaruuden vektoria pistämällä yksinkertaisesti n kappaletta reaalilukuja jonoon. Siis esimerkiksi

$$\mathbf{u} = (u_1, u_2, \ldots, u_n) \quad \text{ja} \quad \mathbf{v} = (v_1, v_2, \ldots, v_n).$$

Kaikkien tällaisten vektorien joukko on tulojoukko

$$\underbrace{\mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}}_{n\text{kpl}} = \mathbb{R}^n.$$

Nollavektorissa on n kappaletta nollia, ja kaksi vektoria ovat samat, jos niillä on samat koordinaatit. Yhteenlasku ja reaaliluvulla kertominen määritellään samalla tavalla kuin 2- ja 3-ulotteisissa tapauksissa. Kantavektorit $\mathbf{i} = (1{,}0)$ jne. yleistyvät vektoreiksi

$$\begin{aligned}
\mathbf{e}_1 &= (1{,}0{,}0, \ldots ,0),\\
\mathbf{e}_2 &= (0{,}1{,}0, \ldots ,0),\\
\ldots &\quad \ldots \ \ldots\\
\mathbf{e}_n &= (0{,}0, \ldots ,0{,}1),
\end{aligned}$$

ja jokaisella $\mathbb{R}^n$:n vektorilla on yksikäsitteinen esitys tässä kannassa. Esimerkiksi

$$\mathbf{u} = (u_1, u_2, \ldots, u_n) = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + \ldots + u_n\mathbf{e}_n.$$

Skalaaritulo on

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \ldots + u_nv_n,$$

ja se noudattaa samoja laskusääntöjä kuin $\mathbb{R}^2$:n ja $\mathbb{R}^3$:n vektorien skalaaritulo. Vektorin pituus on $|\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$. Vektorien $\mathbf{u}$ ja $\mathbf{v}$ välinen kulma γ määritellään asettamalla

$$\cos\gamma = \frac{u_1v_1 + \ldots + u_nv_n}{\sqrt{u_1^2 + \ldots + u_n^2}\sqrt{v_1^2 + \ldots + v_n^2}}.$$

CBS-epäyhtälön (3) perusteella tämä määritelmä on mielekäs. Kosini määrää vektorien välisen kulman γ yksikäsitteisesti, sillä $\gamma \in [0°, 180°]$. Itse CBS-epäyhtälö on vektorimuodossa hyvin yksinkertainen:

$$|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}||\mathbf{v}|.$$

Sen avulla voidaan todistaa mm. kolmioepäyhtälö, ks. teht. 7.

Harjoituksia

6. a) Sijoita yksikkökuutio $\mathbb{R}^3$:n positiivisten akselien rajaamaan soppeen siten, että yksi kärki tulee origoon ja kolme siitä lähtevää särmää yhtyy koordinaattiakseleihin. Määritä kuution kärkipisteiden koordinaatit. Laske samasta kärjestä lähtevän särmän ja avaruuslävistäjän välinen kulma.

b) Sijoita yksikkökuutio $\mathbb{R}^4$:n positiivisten akselin osien raajaamaan osaan siten, että yksi kärki tulee origoon ja neljä siitä lähtevää särmää yhtyy koordinaattiakseleihin. Määritä kuution kärkipisteiden koordinaatit. Laske samasta kärjestä lähtevän särmän ja avaruuslävistäjän välinen kulma.

7. a) Olkoon $\mathbf{u} \in \mathbb{R}^n$, ja $\gamma_\imath$ sen sekä kantavektorin $\mathbf{e}_\imath$ välinen kulma kaikilla $\imath \in \{1, 2, \ldots, n\}$. Osoita, että

$$\cos^2\gamma_1 + \cos^2\gamma_2 + \ldots + \cos^2\gamma_n = 1.$$

b) Osoita, että $\mathbb{R}^n$:n vektorit toteuttavat kolmioepäyhtälön

$$|\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|.$$

Ohje: Sovella vektorimuotoista CBS-epäyhtälöä.

Sovelluksen sovellus

Mihin moniulotteisia avaruuksia ja niiden vektoreita tarvitaan? Useamman muuttujan funktioiden teoria perustuu oleellisesti niihin, ja näihin funktioihin puolestaan perustuvat monet matematiikan sovellukset. Tässä katsotaan lyhyesti yhtä tilastotieteen sovellusta, jossa tarvitaan vain vektoreita.

Oletetaan, että halutaan selvittää, onko ylipainoisilla henkilöillä keskimäärin muita korkeampi verenpaine (alapaine). Valitaan satunnaisesti n henkilöä käsittävä koeryhmä. Esittäkööt u_1, u_2, …, u_n heidän painoindeksejään ja v_1, v_2, …, v_n verenpaineitaan. Siis $u_\imath$ ja $v_\imath$ ovat samaa henkilöä koskevat mittaustulokset. Olkoot painoindeksien keskiarvo on $\bar{u}$ ja keskimääräinen verenpaine $\bar{v}$. Muodostetaan *havaintovektorit* asettamalla

$$\mathbf{u} = (u_1 - \bar{u}, \ldots, u_n - \bar{u})$$

ja

$$\mathbf{v} = (v_1 - \bar{v}, \ldots, v_n - \bar{v}).$$

Jos koehenkilöiden enemmistöllä molemmat arvot ovat keskiarvon samalla puolella, niin silloin pieni painoindeksi useimmiten yhdistyy matalaan verenpaineeseen ja suuri korkeaan. Tällöin tulot $(u_\imath - \bar{u})(v_\imath - \bar{v})$ ovat enimmäkseen positiivisia, havaintovektorien skalaaritulo on positiivinen, ja havaintovektorit ovat lähempänä saman- kuin vastakkaissuuntaisuutta. Jos taas usean henkilön arvot ovat keskiarvojen eri puolilla, niin skalaaritulon termit ovat negatiivisia ja skalaaritulokin on negatiivinen. Tällöin havaintovektorit ovat lähempänä vastakkais- kuin samansuuntaisuutta, ja suurta painoindeksiä vastaa matala verenpaine ja päin vastoin. Jos tulojen $(u_\imath - \bar{u})(v_\imath - \bar{v})$ merkit jakautuvat tasaisesti positiivisiin ja negatiivisiin, niin skalaaritulo on todennäköisesti lähellä nollaa, eikä tutkittavien ilmiöiden välillä ei ole havaittavissa selvää riippuvuutta. Havaintovektorit ovat tällöin lähellä keskinäistä kohtisuoruutta.

Jos ylipainon lisäksi halutaan tutkia jotakin toista selittävää tekijää korkealle verenpaineelle, muodostetaan uudesta tekijästä havaintovektori $\mathbf{w}$. Skalaaritulot $\mathbf{u}\cdot\mathbf{v}$ ja $\mathbf{w}\cdot\mathbf{v}$ eivät kuitenkaan välttämättä ole vertailukelpoisia, sillä mitattavien suureiden lukuarvot voivat olla aivan eri suuruusluokissa. Miten saadaan nämä skalaaritulot vertailukelpoisiksi? Muodostamalla havaintovektorien suuntaisten yksikkövektorien väliset skalaaritulot. Ne ovat alkuperäisten vektorien pituuksista riippumattomia, ja ne esittävät niiden välisten kulmien kosineita. Tilastotieteessä tätä kosinia kutsutaan *korrelaatiokertoimeksi*, ja se merkitään r-kirjaimella, ks. [2], s. 53. Siis

$$\cos\gamma = \frac{\mathbf{u}}{|\mathbf{u}|}\cdot\frac{\mathbf{v}}{|\mathbf{v}|} = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}||\mathbf{v}|} = r.$$

Jos r on lähellä ykköstä, niin ylipainon katsotaan olevan yhteydessä kohonneeseen verenpaineeseen. Tämä ei kuitenkaan *todista* näiden ilmiöiden välistä syyseuraus-suhdetta, vaan ainoastaan sen, että ilmiöt esiintyvät usein samalla yksilöllä.

Moni tilastomatematiikan peruskurssin suorittanut pitää korrelaatiokerrointa käsittämättömän monimutkaisena ulkoa opeteltavana kaavarumiluksena. Vektoritulkinnan kautta se on kuitenkin täysin selvä ja luonnollinen käsite.

Useamman luvun keskiarvoista

Keskiarvot yleistyvät useammalle positiiviselle luvulle yhtälöillä

$$\mathcal{H} = \frac{n}{\frac{1}{u_1}+\ldots+\frac{1}{u_n}},$$

$$\mathcal{G} = \sqrt[n]{u_1\ldots u_n},$$

$$\mathcal{A} = \frac{u_1+\ldots+u_n}{n} \quad \text{ja}$$

$$\mathcal{C} = \frac{u_1^2+\ldots+u_n^2}{u_1+\ldots+u_n}.$$

Ne ovat lukujen $u_1, u_2, \ldots, u_n$ symmetrisiä funktioita, ja jokainen niistä sijaitsee pienimmän ja suurimman u-luvun välissä. Myös epäyhtälöketju

$$\mathcal{H} \le \mathcal{G} \le \mathcal{A} \le \mathcal{C}$$

on voimassa. Siinä vallitsee yhtäsuuruus jos ja vain jos kaikki u-luvut ovat keskenään yhtäsuuria. Tämän todistamisen jätetään harjoitustehtäviksi.

Harjoituksia

8. Osoita, että keskiarvot $\mathcal{H}$, $\mathcal{G}$, $\mathcal{A}$ ja $\mathcal{C}$ sijaitsevat pienimmän ja suurimman u-luvun välissä.

9. Todista epäyhtälö $\mathcal{G} \le \mathcal{A}$ yhtäsuuruusehtoineen. Ohje: Tälle epäyhtälölle löytyy useita erilaisia todistuksia, mutta teoksessa [1] esitetään seuraava erityisen nerokas ajatus. Voidaan rajoituksetta olettaa, että $u_1 \le u_2 \le \ldots \le u_n$. Geometrinen keskiarvo $\mathcal{G}$ sijaitsee pienimmän ja suurimman u-luvun välissä, joten on olemassa k, jolle $u_k \le \mathcal{G} \le u_{k+1}$. Koska

$$\sum_{\imath=1}^{k}\int_{u_\imath}^{\mathcal{G}}\left(\frac{1}{t}-\frac{1}{\mathcal{G}}\right)\mathrm{d}t + \sum_{\imath=k+1}^{n}\int_{\mathcal{G}}^{u_\imath}\left(\frac{1}{\mathcal{G}}-\frac{1}{t}\right)\mathrm{d}t \ge 0, \quad (1)$$

ja koska integroitavat ovat ei-negatiivisia, on yhtäsuuruus voimassa, jos ja vain jos $u_\imath = \mathcal{G}$ kaikilla $\imath \in \{1,2,\ldots,n\}$. Osoita, että (1) sievenee epäyhtälöksi $\mathcal{G} \le \mathcal{A}$.

10. a) Todista epäyhtälö $\mathcal{H} \le \mathcal{G}$ yhtäsuuruusehtoineen. Ohje: Sovella epäyhtälöä $\mathcal{G} \le \mathcal{A}$ sopivasti valittuihin lukuihin.

b) Todista epäyhtälö $\mathcal{A} \le \mathcal{C}$ yhtäsuuruusehtoineen. Ohje: Sovella CBS-epäyhtälöä sopivasti valittuihin lukuihin.

Kiitän dosentti Jorma Merikoskea kirjoitustani koskeneista kommenteista.

Lähdeluettelo

[1] Martin Aigner, Günter M. Ziegler, *Proofs from the Book*, Springer, 2004.

[2] Raimo Seppänen, Martti Kervinen, Irma Parkkila, Lea Karkela, Pekka Meriläinen, *Maol taulukot*, Otava, 2006.

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