Osittaissummauksella ikävien summien kimppuun

Anne-Maria Ernvall-Hytönen
Matematiikan ja tilastotieteen laitos, Helsingin yliopisto

Kuvitellaan, että pihalla leikkii lapsia, joilla jokaisella on yhtä monta palloa kuin ikää vuosina. Naapuriin muuttaa lapsia, joilla ei ole lainkaan palloja. Solidaarisuuden nimissä lapset päättävät antaa naapureille osan palloistaan. Koska olisi kohtuutonta, jos kaikki luopuisivat yhtä monesta pallosta, kun niitä ei alun perinkään yhtä monta ollut, päätetään, että jokainen lapsi jättää itselleen (kokonaislukuun pyöristettynä) sen määrän palloja, mikä jää, kun oma alkuperäinen pallomäärä jaetaan oman iän luonnollisella logaritmilla (paitsi yksi- ja kaksivuotiaiden palloihin ei kosketa). Siispä, jos lapsen ikä on vaikka 5, alun perin hänellä on viisi palloa, ja lahjoituksen jälkeen $\frac{5}{\ln 5} \approx 3{,}1 \approx 3$ palloa. Kuinka monta palloa lapsille karkeasti ottaen jää yhteensä?

Oletetaan yksinkertaisuuden vuoksi, että pihalla on vain vähintään kolmevuotiaita lapsia. On helppo laskea, että jos alun perin pallomäärät (eli iät) ovat nousevassa järjestyksessä $a_1, a_2, \ldots, a_k$, ja niitä on yhteensä N kappaletta, niin uusi pallomäärä on

$$\begin{aligned}
&\frac{a_1}{\ln a_1} + \frac{a_2}{\ln a_2} + \cdots + \frac{a_k}{\ln a_k} \\
&\qquad \geq \frac{a_1}{\ln a_k} + \frac{a_2}{\ln a_k} + \cdots + \frac{a_k}{\ln a_k} \\
&\qquad = \frac{1}{\ln a_k}\left(a_1 + a_2 + \cdots + a_k\right) = \frac{N}{\ln a_k},
\end{aligned}$$

mikäli ei välitetä pyöristyksistä. Samoin saadaan, että pallomäärä on korkeintaan $\frac{N}{\ln a_1}$ (jälleen mikäli ei välitetä pyöristyksistä). Mikäli luvut a_1 ja a_k (eli nuorimman ja vanhimman iät) ovat lähellä toisiaan, on näin saatu varsin hyvät arviot summalle. Jos lukujen a_1 ja a_k ero on huomattava, on tuloksessakin heittoa. Esimerkiksi, jos lasten iät ovat $3, 4, \ldots, 17$, niin alun perin palloja on $3 + 4 + \cdots + 17 = 150$, ja $\frac{150}{\ln 3} \approx 137$ ja $\frac{150}{\ln 17} \approx 53$. Näiden tulosten välillä on valtava ero. Kumpi näistä on lähempänä totuutta? Ovatko molemmat päin honkia?

Koska luvut ovat varsin pieniä, voidaan helposti laskea pallojen todellinen määrä. Kolmevuotiaalla on $\frac{3}{\ln 3} \approx 2{,}7 \approx 3$ palloa, nelivuotiaalla $\frac{4}{\ln 4} \approx 3$ palloa, ja niin edelleen. Yhteensä 66 palloa. Pienempi raja oli siis selvästi parempi, ja se ei itse asiassa edes ollut täysin päin honkia, vaikkakaan ei ihan oikeinkaan.

Tällaisten summien käsittelyyn *osittaussummaus* on oiva työkalu. Esitellään seuraavissa luvuissa osittaissummaus formaalisti ja todistetaan se, käsitellään yhteys osittaisintegrointiin, ja lopuksi käydään läpi muutama esimerkki, mukaan lukien lukuteoreettinen sovellus. Mikäli osittaisintegrointi on täysin vierasta, voi jättää lukematta sen osan tekstiä, joka liittää osittaissummauksen osittaisintegrointiin. Tekstin muun osan ymmärtäminen ei vaadi tämän osan lukemista.

Osittaissummauksen pääperiaate

Esitän osittaissummauksen kuten se Inghamin kirjassa [1] on sivulla 18 esitetty vain hieman notaatiota muuttaen:

Lause. Olkoon $\lambda_1, \lambda_2, \ldots$ kasvava jono reaalilukuja, joka lähestyy ääretöntä. Olkoon

$$A(x) = \sum_{\lambda_n \le x} a_n,$$

missä luvut a_n voivat olla reaalisia tai kompleksisia. Jos $X \ge \lambda_1$ ja funktiolla $f(x)$ on jatkuva derivaatta. niin

$$\sum_{\lambda_n \le X} a_n f(\lambda_n) = A(X)f(X) - \int_{\lambda_1}^{X} A(x)f'(x)dx.$$

Todistus. Todistus on varsin yksinkertainen, ja se perustuu alkuperäisen summan ja termin $A(X)f(X)$ erotukseen:

$$\begin{aligned} A(X)f(X) &- \sum_{\lambda_n \le X} a_n f(\lambda_n) \\ &= \sum_{\lambda_n \le X} a_n (f(X) - f(\lambda_n)). \end{aligned}$$

Koska funktio f on jatkuvasti derivoituva, voidaan kirjoittaa

$$f(X) - f(\lambda_n) = \int_{\lambda_n}^{X} f'(x)dx,$$

jolloin

$$\begin{aligned} A(X)f(X) &- \sum_{\lambda_n \le X} a_n f(\lambda_n) \\ &= \sum_{\lambda_n \le X} a_n \int_{\lambda_n}^{X} f'(x)dx, \end{aligned}$$

jolloin integrointi- ja summausjärjestystä vaihtaen (Harjoitustehtävä: piirrä kuva muuttujien suuruuksista tai muutoin vakuuta itsesi, että näin todella on):

$$\begin{aligned} \sum_{\lambda_n \le X} a_n \int_{\lambda_n}^{X} f'(x)dx &= \int_{\lambda_1}^{X} \sum_{\lambda_n \le x} a_n f'(x)dx \\ &= \int_{\lambda_1}^{X} A(x)f'(x)dx, \end{aligned}$$

ja väite on todistettu.

Käytännössä jono λ_n on usein esimerkiksi $\lambda_n = n$ tai alkuluvuista koostuva jono, mutta lauseen määrittely antaa mahdollisuuden määritellä jonon paljon mielikuvituksellisemmillakin tavoilla.

Yhteys osittaisintegrointiin

Aluksi varoituksen sananen: hetken päästä puhun sujuvasti epäjatkuvien funktioiden integraaleista. Tällaisten funktioiden integraalia ei välttämättä pysty määrittelemään millään nätillä kaavalla. Silti muutoin riittävän siististi käyttäytyvät epäjatkuvat funktiot ovat usein integroituvia (ehtoja integroituvuudella käsitellään yliopiston kursseilla).

Osittaissummaus on osittaisintegroinnin diskreetti vastine: Siinä missä positiivisen funktion g integraali mittaa funktion ja x-akselin väliin jäävän alueen pinta-alaa, summa funktion arvoista mittaa sellaista pinta-alaa, joka saadaan muodostamalla funktion arvojen korkuisia ja ykkösen levyisiä palkkeja. Voi myös ajatella niin, että jos funktio g on määritelty kokonaislukupisteissä, niin täydennetään siitä kaikilla reaaliluvuilla määritelty funktio siten, että määrätään $g(x) = g(\lfloor x \rfloor)$. Näin summauksen ja integroinnin yhteys on ilmeinen.

Osittaisintegrointihan toimii seuraavasti:

$$\int_a^b g(x)f(x)dx = \Big/_{a}^{b} G(x)f(x) - \int_a^b G(x)f'(x)dx,$$

missä $G'(x) = g(x)$. Osittaissummauksen tilanteessa funktio f ei muutu mitenkään. Sen sijaan funktio g pitää ajatella kokonaislukupisteissä määriteltynä funktiona (tai siitä reaaliluvuille kaavalla $g(x) = g(\lfloor x \rfloor)$ yleistettynä) ja $G(x)$ puolestaan summafunktiona tai yleistetyn funktion integraalina. Osittaissummauksessa alasijoitukset jäävät pois, sillä summa on siellä tyhjä (eli osittaissummauksessa $G(a) = 0$).

Kuinka suuri on kertoma?

Luvun n kertoma, jota merkitään $n!$, on lukujen $1, 2, \ldots, n$ tulo, eli

$$n! = 1 \cdot 2 \cdot 3 \cdots n = \prod_{1 \le m \le n} m.$$

Ensimmäinen ajatus tätä lukua tuijottaessa ainakin minulle itselleni alun perin oli, että sen on pakko olla tuhottomasti pienempi kuin n^n. Alarajan määrittäminen tuntui paljon hankalammalta. Esimerkiksi arvio $n! \ge 1^n = 1$ tarjoaisi kyllä alarajan, mutta yleisesti ottaen tämä alaraja on aivan valtavan huono. Tarkastellaan nyt tätä lukua tarkemmin. Otetaan aluksi tulosta logaritmi:

$$\ln n! = \sum_{1 \le m \le n} \ln m.$$

Tämän summan voi joko arvioida edellisen Solmun numeron opeilla integraalien avulla, tai sitten voidaan hyödyntää osittaissummausta.

Osittaissummauksen notaatiota käyttäen $a_m = 1$ ja $f(x) = \ln x$. Nyt

$$\begin{aligned} \sum_{1 \le m \le n} \ln m &= \sum_{1 \le m \le n} a_m f(m) \\ &= f(n) \sum_{1 \le m \le n} a_m - \int_1^n \left(\sum_{1 \le m \le x} a_m \right) f'(x)dx \\ &= n \ln n - \int_1^n \frac{\lfloor x \rfloor}{x} dx. \end{aligned}$$

Nyt saadaan helpolla arviolla

$$\int_1^n \frac{\lfloor x \rfloor}{x} dx \leq \int_1^n 1 dx = n - 1 < n.$$

Siispä $\ln n! \geq n \ln n - n$, eli $n! \geq \left(\frac{n}{e}\right)^n$, joten oikeastaan n^n ei ole yhtään hullumpi arvio luvulle $n!$, vaikka se reilusti yläkanttiin onkin.

Lapset jälleen pihamaalla

Palataan nyt alun esimerkkiin ja solidaarisiin lapsiin leikkimässä pihamaalla ja jakamassa pallojaan. Oletetaan, että lasten iät ovat $3, 4, 5, \ldots, n$, ja tarkastellaan miten pallojen määrä käyttäytyy, kun n kasvaa. (On ehkä syytä olettaa, että eletään jollain mystisellä planeetalla, jolla täysikäisyyden raja on ääretön, eli kaikki ovat lapsia iästä riippumatta.)

Lapsen, jonka ikä on k vuotta, pallomäärä on $\frac{k}{\ln k}$ pyöristettynä kokonaislukuun. Lapsia on alle n kappaletta, joten jos unohdetaan pyöristykset ja eletään ikään kuin pallomäärä olisi $\frac{k}{\ln k}$, niin virhe on korkeintaan luvun n kokoinen kaikilta lapsilta yhteensä.

Muistetaan vielä, että niiden lasten, joiden ikä on korkeintaan m, ikien summa on $\frac{m(m+1)}{2} - 3$. Nyt

$$\begin{aligned}\sum_{3 \leq k \leq n} \frac{k}{\ln k} &= \frac{n(n+1)}{2 \ln n} - \frac{3}{\ln n} \\ &\quad - \int_3^n \left(\sum_{3 \leq k \leq x} k \right) \cdot \frac{-1}{x (\ln x)^2} dx.\end{aligned}$$

Huomataan ensinnäkin, että

$$\frac{n(n+1)}{2 \ln n} - \frac{3}{\ln n}$$

on hyvin tarkasti $\frac{n(n+1)}{2 \ln n}$, eli pallojen alkuperäinen määrä jaettuna logaritmilla. Keskitytään nyt integraalin arviointiin. Integraalin tarkka määrittäminen olisi viheliäinen tehtävä, joten pelataan likiarvoilla.

Nyt huomataan, että

$$\sum_{3 \leq k \leq x} k \leq \frac{x(x+1)}{2} - 3,$$

kun x on mikä tahansa reaaliluku. Voidaan siis korvata integraali lausekkeella

$$\begin{aligned}&- \int_3^n \left(\sum_{3 \leq k \leq x} k \right) \cdot \frac{-1}{x (\ln x)^2} dx \\ &\qquad \leq \int_3^n \left(\frac{x(x+1)}{2} - 3 \right) \cdot \frac{1}{x (\ln x)^2} dx.\end{aligned}$$

Termin

$$\int_3^n \frac{3}{x (ln x)^2} dx \leq 3 \int_3^n \frac{1}{x} dx \leq 3 \ln n$$

vaikutus on hyvin pieni. Voidaan siis keskittyä integraalin

$$\int_3^n \frac{x(x+1)}{2} \cdot \frac{1}{x (\ln x)^2} dx = \int_3^n \frac{x+1}{2 (\ln x)^2} dx$$

arviointiin. Koska olemme kiinnostuneita pallojen määrän käyttäytymisestä, kun n lähestyy ääretöntä, voidaan olettaa $\sqrt{n} \geq 3$. Jaetaan integraali kahtia pisteestä $\sqrt{n}$, jolloin saadaan

$$\int_3^n \frac{x+1}{2 (\ln x)^2} dx = \int_3^{\sqrt{n}} \frac{x+1}{2 (\ln x)^2} dx + \int_{\sqrt{n}}^n \frac{x+1}{2 (\ln x)^2} dx.$$

Ensimmäinen palikka on hyvin helppo arvioida:

$$\int_3^{\sqrt{n}} \frac{x+1}{2 (\ln x)^2} dx \leq n,$$

koska integroitava on pienempi kuin $\sqrt{n}$, kuten myös integraalin pituus. Toinen termi yksinkertaistetaan logaritmin avulla. Tässä ei edes menetetä paljon tarkkuutta, sillä logaritmi kasvaa hyvin hitaasti. Siispä

$$\begin{aligned}\int_{\sqrt{n}}^n \frac{x+1}{2 (\ln x)^2} dx &\leq \int_{\sqrt{n}}^n \frac{x+1}{2 (\ln \sqrt{n})^2} dx \\ &= \frac{2}{(\ln n)^2} \int_{\sqrt{n}}^n (x+1) dx \\ &\leq \frac{2}{(\ln n)^2} \cdot \frac{(n+2)n}{2} = \frac{(n+2)n}{(\ln n)^2}.\end{aligned}$$

Luvun n kasvaessa tämä viimeksi saatu termi on suurin virhetermeistä. Kuitenkin sekin on selvästi pienempi kuin $\frac{n(n+1)}{2 \ln n}$, joten pallojen määrä jaettuna suurimmalla logaritmilla osoittautuu varsin hyväksi arvioksi, kun lasten iät ovat peräkkäisiä kokonaislukuja, ja vanhin lapsista on hyvin vanha.

Pienimmän yhteisen jaettavan suuruusluokka

Tarkastellaan miten suuri on sellainen luku, joka on n ensimmäisen positiivisen kokonaisluvun pienin yhteinen jaettava, eli joka on pienin positiivinen kokonaisluku, joka on jaollinen kaikilla luvuista $1, 2, \ldots, n$. Huomataan aluksi (väitteen varmistaminen jätetään harjoitustehtäväksi), että tämän luvun pitää olla muotoa

$$\prod_{p \leq n,\ p \text{ alkuluku}} p^{\lfloor \log_p n \rfloor}.$$

Koska $\lfloor \log_p n \rfloor = 1$, kun $\sqrt{n} < p \leq n$, voidaan tulo kirjoittaa tulon

$$\prod_{p \leq \sqrt{n},\ p \text{ alkuluku}} p^{\lfloor \log_p n \rfloor}$$

ja tulon

$$\prod_{\sqrt{n}<p\leq n,\ p \text{ alkuluku}} p$$

tulona. Aloitetaan tarkastelu jälkimmäisestä termistä. Tämän käsittely on lähes samanlainen kuin kertomankin tapauksessa. Yksinkertaisuuden vuoksi otetaan tulosta logaritmi, jolloin saadaan se muunnettua summaksi:

$$\ln \prod_{\sqrt{n}<p\leq n,\ p \text{ alkuluku}} p = \sum_{\sqrt{n}<p\leq n,\ p \text{ alkuluku}} \ln p.$$

Alkulukulause on hyvin syvällinen tulos (sitä ei todisteta tässä, mutta tulos löytyy useista analyyttisen lukuteorian kirjoista), joka kertoo, kuinka paljon annettua lukua pienempiä alkulukuja on. Sen nojalla lukua x pienempiä alkulukuja on

$$\frac{x}{\ln x} + o\left(\frac{x}{\ln x}\right),$$

missä $o\left(\frac{x}{\ln x}\right)$ tarkoittaa, että sitä vastaavan virhetermin ja luvun $\frac{x}{\ln x}$ osamäärä lähestyy nollaa, kun x lähestyy ääretöntä. Lukujen $\sqrt{n}$ ja n välissä on siis

$$\frac{n}{\ln n} + o\left(\frac{n}{\ln n}\right) - \frac{\sqrt{n}}{\ln\sqrt{n}} - o\left(\frac{\sqrt{n}}{\ln\sqrt{n}}\right) = \frac{n}{\ln n} + o\left(\frac{n}{\ln n}\right)$$

alkulukua. Osittaissummauksen avulla

$$\begin{aligned}\sum_{\sqrt{n}<p\leq n} \ln p &= \ln n \cdot \left(\frac{n}{\ln n} + o\left(\frac{n}{\ln n}\right)\right) \\ &\quad - \int_{\sqrt{n}}^{n} \left(\frac{x}{\ln x} + o\left(\frac{x}{\ln x}\right)\right) x^{-1} dx \\ &= n + o(n) - \int_{\sqrt{n}}^{n} \left(\frac{1}{\ln x} + o\left(\frac{1}{\ln x}\right)\right) dx.\end{aligned}$$

Integraalin voi käsitellä usealla tavalla riippuen siitä, mitä tarkkuutta tavoitellaan. Kuitenkin koska päätermi on n, ja ensimmäinen virhetermi $o(n)$, olisi turhaa kikkailua käyttää energiaa saadakseen integraalille tarkan arvon. Voidaan siis helposti arvioida:

$$\begin{aligned}&\int_{\sqrt{n}}^{n} \left(\frac{1}{\ln x} + o\left(\frac{1}{\ln x}\right)\right) dx \\ &\quad\leq \int_{\sqrt{n}}^{n} \left(\frac{1}{\ln\sqrt{n}} + o\left(\frac{1}{\ln\sqrt{n}}\right)\right) dx \\ &\quad\leq \frac{n}{\ln\sqrt{n}} + o\left(\frac{n}{\ln\sqrt{n}}\right) = o(n).\end{aligned}$$

Olemme nyt saaneet valmiiksi suurten alkulukujen osuuden määrittämisen. Nyt voidaan siirtyä pieniin alkulukuihin, eli tuloon

$$\prod_{p\leq\sqrt{n},\ p \text{ alkuluku}} p^{\lfloor \log_p n\rfloor}.$$

Otetaan ensimmäiseksi tästäkin logaritmi:

$$\ln \prod_{p\leq\sqrt{n},\ p \text{ alkuluku}} p^{\lfloor \log_p n\rfloor} = \sum_{p\leq\sqrt{n},\ p \text{ alkuluku}} \lfloor \log_p n\rfloor \ln p.$$

Ennen kuin jatketaan pidemmälle, halutaan päästä eroon alaspäinpyöristyksistä. Tämä voidaan tehdä yksinkertaisesti seuraavasti:

$$\begin{aligned}&\sum_{p\leq\sqrt{n},\ p \text{ alkuluku}} \lfloor \log_p n\rfloor \ln p \\ &\quad= \sum_{p\leq\sqrt{n},\ p \text{ alkuluku}} \log_p n \ln p \\ &\qquad- \sum_{p\leq\sqrt{n},\ p \text{ alkuluku}} (\log_p n - \lfloor \log_p n\rfloor) \ln p.\end{aligned}$$

Viimeisen rivin termistä tulee virhetermi:

$$\begin{aligned}&\sum_{p\leq\sqrt{n},\ p \text{ alkuluku}} (\log_p n - \lfloor \log_p n\rfloor) \ln p \\ &\qquad\leq \sum_{p\leq\sqrt{n},\ p \text{ alkuluku}} \ln p,\end{aligned}$$

ja samoin kuin suurten alkulukujen kontribuutiota määritettäessä saadaan

$$\sum_{p\leq\sqrt{n},\ p \text{ alkuluku}} \ln p = \sqrt{n} + o(\sqrt{n}) = o(n).$$

Voidaan siirtyä summan

$$\sum_{p\leq\sqrt{n},\ p \text{ alkuluku}} \log_p n \ln p$$

tarkastelemiseen. Huomataan ensin, että

$$\log_p n = \frac{\ln n}{\ln p},$$

jolloin summa sievenee helposti

$$\sum_{p\leq\sqrt{n},\ p \text{ alkuluku}} \log_p n \ln p = \ln n \sum_{p\leq\sqrt{n},\ p \text{ alkuluku}} 1.$$

Alkulukulauseen nojalla

$$\sum_{p\leq\sqrt{n},\ p \text{ alkuluku}} 1 = \frac{\sqrt{n}}{\ln\sqrt{n}} + o\left(\frac{\sqrt{n}}{\ln\sqrt{n}}\right),$$

joten

$$\ln n \sum_{p\leq\sqrt{n},\ p \text{ alkuluku}} 1 = o(n).$$

Täten pienimmän yhteisen jaettavan luonnollinen logaritmi on suuruusluokkaa $n+o(n)$, joten pienin yhteinen jaettava kasvaa hyvin karkeasti ottaen yhtä nopeasti kuin e^n.

Viitteet

[1] A. E. Ingham, *The distribution of prime numbers.* Cambridge University Press, 1990.