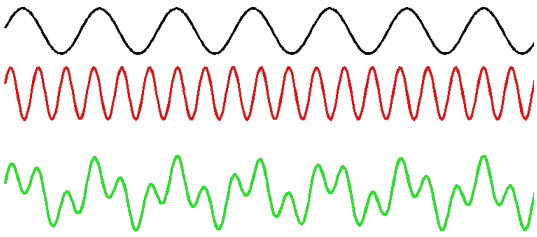
THE ABEL PRIZE 2021

Lovász-Lenstra-Lenstra reduksjonsalgoritme for gittere

For rundt 200 år siden publiserte den franske matematikeren Jean Baptiste Joseph Fourier et bevis for at enhver kontinuerlig funksjon kan skrives som en uendelig sum av sinus- og cosinus-bølger. Resultatet har stor betydning ved opptak og lagring av lyd. Rene sinusbølger kan konverteres til lyd gjennom en høyttaler og lyden fra et instrument kan betraktes som en sum av sinusbølger med heltallsmultipler av den grunnleggende frekvensen. Hver frekvens bidrar til lyden med sin individuelle amplitude. Utfordringen ved opptak av lyd er å separere de ulike frekvensene.



I en matematisk setting betrakter vi hver frekvens som en separat koordinat i et stort vektorrom. Ved å introdusere et passende indre-produkt på rommet danner frekvens-koordinatsystemet en ortogonal basis for sinusbølger med ulike frekvenser. Ortogonaliteten av denne basisen forenkler prosessen med å dekomponere lyden. Dette er derfor et godt eksempel på hvorfor det er relevant å lete etter ortogonale basiser for vektorrom.

En effektiv algoritme for å produsere ortogonale basiser er Gram-Schmidt prosessen.

Lemma. La $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\} \subset \mathbb{R}^n$ være en basis for et vektorrom V, og la

$$\mu_{ij} = \frac{\langle \mathbf{b}_i, \mathbf{b}_j^* \rangle}{\|\mathbf{b}_j^*\|^2}, \quad 1 \leq j < i \leq n$$

Disse størrelsene kalles Gram-Schmidt koeffisientene til basisen $\mathcal{B}$. De koder informasjon om den ortogonale projeksjonen av en basisvektor langs en annen. La

$$\mathbf{b}_k^* = \mathbf{b}_k - \sum_{i=1}^{k-1} \mu_{ki} \mathbf{b}_i^*$$

Da er $\mathcal{B}^* = \{\mathbf{b}_1^*, \mathbf{b}_2^*, \ldots, \mathbf{b}_n^*\}$ en ortogonal basis for $\mathbb{R}^n$.

En ortogonal basis for et vektorrom gir opphav til ortonormal basis, dvs. ortogonal og slikt at hver basisvektor har lengde 1.

Computer science er i sin natur mye sterkere knyttet til diskret matematikk enn til kontinuumsmatematikk. Vi er derfor mer eller mindre tvunget til å jobbe med gitre $\mathbb{Z}^n$, framfor vektorrom $\mathbb{R}^n$. Men det er fortsatt ønskelig å kunne ha ortogonale basiser tilgjengelig, og også ha tilgang til "korte" basisvektorer. Definisjonen av en "kort" vektor er nokså vag, det refererer til den vanlige euklidske lengden av vektoren, og en vektor i et gitter sies å være kort dersom lengden er i nærheten av den minimale lengden av en hvilken som helst annen vektor i gitteret. Som et

eksempel kan man betrakte gitteret generert av de to "lange" vektorene $(6386, 51)$ og $(71999, 575)$. En annen basis for det samme gitteret er gitt ved de "korte" vektorene $(1, 0)$ og $(0, 1)$.

Det er to vesentlige forskjeller mellom et gitter og et vektorrom:

i) Det er ikke åpenbart at det eksisterer en ortogonal basis for gitteret,

Ii) Det kan være tidkrevende å finne "korte" vektorer i gitteret.

Hvis du ikke kan få det beste, må du nøye deg med det nest beste. Det nest beste er i denne situasjonen å fravike ønsket om å finne en ortonormal basis, å heller lete etter en Lovász-Lenstra-Lenstra-redusert basis (eller enklere sagt en LLL-redusert basis):

Definisjon. En basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ for et gitter $\mathcal{L}$ er LLL-redusert dersom det finnes en parameter $\delta \in (0.25, 1]$ slik at følgende er oppfyllt:

(i) For $1 \leq j < i \leq n$: $|\mu_{ij}| \leq 0.5$ (lengde-reduksjon)

(ii) For $k = 2, 3, \ldots, n$: $\delta \|\mathbf{b}_{k-1}^*\|^2 \leq \|\mathbf{b}_k^* + \mu_{k,k-1}\mathbf{b}_{k-1}^*\|^2$ (Lovász-betingelsen)

hvor $\mathcal{B}^*$ er den tilhørende Gram-Schmidt-basisen.

Ved å bruke det vanlige euklidske indreproduktet har vi

$$\langle \mathbf{b}_i, \mathbf{b}_j^* \rangle = \|\mathbf{b}_i\| \, \|\mathbf{b}_j^*\| \, \cos(\theta)$$

hvor θ er vinkelen mellom $\mathbf{b}_i$ og $\mathbf{b}_j^*$. Den første betingelsen i definisjonen er da gitt ved

$$\|\mathbf{b}_j^*\| \, |\cos(\theta)| \leq \frac{\|\mathbf{b}_i\|}{2}$$

Ulikheten er oppfylt enten dersom $\mathbf{b}_i$ og $\mathbf{b}_j$ er nær ved å være ortogonale, dvs. $\cos(\theta)$ er nær ved 0, eller dersom basisvektorene i store trekk er ordnet etter lengde. Dersom vi har en ortogonal basis sier Lovász-betingelsen at

$$\delta^{n-1}\|\mathbf{b}_1^*\|^2 \leq \delta^{n-2}\|\mathbf{b}_2^*\|^2 \leq \cdots \leq \delta\|\mathbf{b}_{n-1}^*\|^2 \leq \|\mathbf{b}_n^*\|^2$$

Det betyr at i begge tilfellene vil basisvektorene langt på vei være ordnet etter lengde.

I en artikkel fra 1972 introduserte László Lovász i et samarbeid med Lenstra-brødrene, Arjen og Hendrik, den såkalte LLL algoritmen. LLL-algoritmen er laget med det formålet å produsere en LLL-redusert basis med en vilkårlig basis for et gitter som input. Algoritmen har to hovedkomponenter; lengdereduksjon og basisvektor-ombytting.

Lengdereduksjon tas hånd om ved en Gram-Schmidt-type prosess, og ombytting av basisvektorer er nødvendig for å opprettholde lengde-ordningen av basisvektorene.

Vi kan se på et eksempel som illustrerer resultatet av algoritmen: La $\mathcal{L} \simeq \mathbb{Z}^3$ være gitteret utspent av de tre vektorene

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \mathbf{b}_2 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \quad \mathbf{b}_3 = \begin{pmatrix} 3 \\ 5 \\ 6 \end{pmatrix}$$

Den euklidske lengden av vektorene er $\sqrt{3}$, $\sqrt{5}$ og $\sqrt{70}$. Etter at vi har gjennomført tilstrekkelig antall iterasjoner av LLL-algoritmen ender vi opp med basisen

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$$

hvor $\|\mathbf{v}_1\| = 1$, $\|\mathbf{v}_2\| = \sqrt{2}$ and $\|\mathbf{v}_3\| = \sqrt{5}$. Gram-Schmidt-koeffisientene til resultat-basisen er $\mu_{21} = \mu_{31} = 0$, $\mu_{32} = \frac{1}{2}$, som alle er mindre enn eller lik $\frac{1}{2}$.

LLL-algoritmen har hatt stor betydningt innen kryptografi.

En annen interessant anvendelse gjelder Mertens-formodningen. Mertens-formodningen ble framsatt mot slutten av det nittende århundret av den polske matematikeren Franz Mertens. Formodningen dreier seg i bunn og grunn om primtallenes fordeling, og kanskje den mest spektakulære konsekvensen av formodningen er Riemann-hypotesen. Dessverre har det vist seg at Mertens-formodningen ikke stemmer, med den konsekvens at et angrep mot Riemann-hypotesen langs denne linja ikke ville bære noen frukter.

Hovedingrediensen i Mertens-formodningen er Möbiusfunksjonen. Möbiusfunksjonen $\mu = \mu)n =$ for et positivt heltall n, ble introdusert av den tyske matematikeren August Ferdinand Möbius i 1832. Den er definert som følger:

- Hvis n er delelig med et kvadrattall, så er $\mu(n) = 0$.
- Hvis n kvadrat-fri med r primfaktorer, så er $\mu(n) = (-1)^r$.

For noen av de minste tallene har vi $\mu(1) = 1$, $\mu(2) = \mu(3) = \mu(5) = -1$, $\mu(4) = 0$ og $\mu(6) = 1$.

Enhver tallfølge kan samles sammen i en Dirichelet-rekke:

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = 1 - \frac{1}{2^s} - \frac{1}{3^s} - \frac{1}{5^s} + \frac{1}{6^s} - \frac{1}{7^s} \cdots$$
$$= (1 - \frac{1}{2^s})(1 - \frac{1}{3^s})(1 - \frac{1}{5^s}) \ldots$$

hvor s er en kompleks variabel. Det er en nær relasjon mellom Dirichelet-rekken til en Möbiusfunksjon og Riemanns zeta-funksjon.

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \dots$$

Vi multipliserer $\zeta(s)$ med $\frac{1}{2^s}$ og trekker fra, og får

$$\begin{aligned}(1 - \frac{1}{2^s})\zeta(s) &= (1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \dots) \\ &\quad - (\frac{1}{2^s} + \frac{1}{4^s} + \frac{1}{6^s} + \dots) \\ &= 1 + \frac{1}{3^s} + \frac{1}{5^s} + \frac{1}{7^s} \dots\end{aligned}$$

Ved å gjøre tilsvarende for alle primtall får vi

$$(1 - \frac{1}{2^s})(1 - \frac{1}{3^s})(1 - \frac{1}{5^s}) \cdots = \frac{1}{\zeta(s)}$$

Det betyr at Möbiusfunksjonen svarer til den multiplikative inversen til Riemanns zetafunksjon.

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}$$

Mertens-formodningen ble lansert av Thomas Joannes Stieltjes, i 1885 i et brev til Charles Hermite og i trykt versjon av Franz Mertens i 1897. Mertens-formodningen gir en øvre grense for den akkumulerte Möbiusfunksjonen;

Formodning. Vi har

$$M(x) = \sum_{n<x} \mu(n) \leq \sqrt{x}$$

for alle $x \geq 1$.

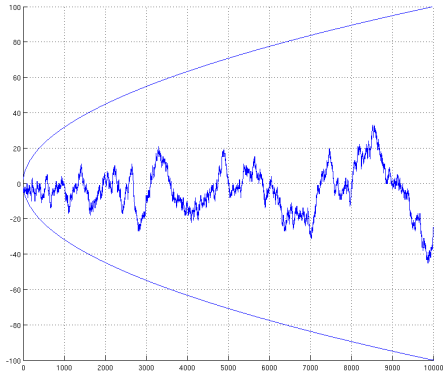
Merk at siden $M(x)$ er konstant på hvert intervall $[n, n+1)$ så har vi

$$\begin{aligned}\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} &= \sum_{n=1}^{\infty} \frac{M(n) - M(n-1)}{n^s} \\ &= \sum_{n=1}^{\infty} M(n)\frac{1}{n^s} - \sum_{n=0}^{\infty} M(n)\frac{1}{(n+1)^s} \\ &= \sum_{n=1}^{\infty} M(n)\left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right) \\ &= \sum_{n=1}^{\infty} M(n) \cdot \int_n^{n+1} \frac{s\,dx}{x^{s+1}} \\ &= s\int_0^{\infty} \frac{M(x)\,dx}{x^{s+1}}\end{aligned}$$

Anta nå at Mertens-formodningen er sann, dvs. at $M(x) \leq \sqrt{x}$. Da har vi

$$\frac{1}{\zeta(s)} \leq s\int_0^{\infty} \frac{\sqrt{x}\,dx}{x^{s+1}} = s\int_0^{\infty} \frac{dx}{x^{s+\frac{1}{2}}}$$

Det siste integralet definerer en analytisk funksjon for $\Re(s) > \frac{1}{2}$, og dermed en analytisk fortsettelse av $\frac{1}{\zeta(s)}$ til området hvor $\Re(s) > \frac{1}{2}$. Spesielt gir dette at $\zeta(s)$ ikke har noen nullpunkter for $\Re(s) > \frac{1}{2}$, Dette er nøyaktig innholdet i Riemann-hypotesen. Det betyr at hvis Mertens-formodningen hadde vært sann, så ville det samme være tilfelle for Riemann-hypotesen. Dessverre er ikke Mertens-formodningen sann. Den ble motbevist av Andrew Odlyzko og Herman te Riele i 1985. Likefullt er Mertens-formodningen et slående eksempel på en matematisk antagelse som er motbevist på tross av en stor mengde indikasjoner på det motsatte. Dette er illustrert i figuren under. Den ytre kurven er funksjonen $f(x) = \pm\sqrt{x}$ og den indre sikksakk-kurven er Mertens-funksjonen $M(x)$. Antagelsen sier at sikksakk-kurven holder seg innenfor den ytre kurven for alle $x \in \mathbb{R}_+$.



Et helt sentralt poeng i beviset til Odlyzko og te Riele var en anvendelse av gitter-reduksjons-algoritmen til Lenstra, Lenstra and Lovász. Ved hjelp av denne algoritmen var de i stand til å vise at

$$\limsup_{x\to\infty} \frac{M(x)}{\sqrt{x}} > 1.06$$

og

$$\liminf_{x\to\infty} \frac{M(x)}{\sqrt{x}} < -1.009$$

noe som motsier ulikheten i formodningen. I argumentet til Lenstra, Lenstra and Lovász ligger det ikke noen eksakt verdi for hvor formodningen svikter, men på et senere tidspunkt er det vist at moteksempelet ligger i intervallet fra 10^{16} til 10^{40}.