KHACHIAN'S ELLIPSOIDE-METHODE VOOR LINEAIRE PROGRAMMERING

A. SCHRIJVER

0. INLEIDING

In februari 1979 verscheen in de Proceedings van de Sovjet Akademie van Wetenschappen een artikel van L.G. Khachian [23], waarin hij liet zien dat met de zgn. "ellipsoide-methode" lineaire programmeringsproblemen kunnen worden opgelost binnen polynomiaal begrensde tijd. Deze ellipsoide-methode was eerder ontwikkeld door D.B. Judin en A.S. Nemirovskii [20,21] en N.Z. Shor [37] om willekeurige convexe programmeringsproblemen te benaderen.

Het is lange tijd een onbeantwoorde vraag geweest of LP-problemen in polynomiale tijd zijn op te lossen. Voor het bestaan van een polynomiale algorithme bestonden een aantal aanwijzingen. Zo blijkt de bekende simplex-methode voor lineaire programmering in de praktijk een zeer efficiënte algorithme (de looptijd blijkt praktisch lineair in de afmetingen van het LP-probleem), hoewel er LP-problemen zijn geconstrueerd waarvoor de simplex-methode exponentieel lange tijd vergt. Ook was het LP-probleem een van de weinige bekende problemen in de klasse $\mathcal{NP}\cap\text{co}\mathcal{NP}$ waarvoor nog geen polynomiale methode was gevonden (zie Garey en Johnson [13] en Van Leeuwen [28]).

Vooralsnog is de doorbraak van Khachian echter voornamelijk van theoretisch belang. De simplex-methode is efficiënt in de praktijk maar niet in theorie, terwijl Khachian's methode efficiënt is in theorie (d.w.z. polynomiaal), maar (nog) niet in de praktijk. Het polynoom dat de looptijd van Khachian's methode begrenst heeft namelijk een zeer hoge graad, en daarnaast vereisen de berekeningen een dermate hoge precisie dat de methode praktisch numeriek instabiel is. Verder onderzoek en meer ervaring met de methode zullen de praktische relevantie moeten uitwijzen, waarvoor wel een aantal essentiële verbeteringen onmisbaar lijken.

Khachian's artikel heeft een stortvloed aan publiciteit en aan onderzoek teweeg gebracht. Aanvankelijk bleef het artikel enige tijd onopgemerkt, maar na rapporten van E.L. Lawler en van P. Gács en L. Lovász [11] werd het resultaat besproken in het wetenschappelijk tijdschrift Science [25], waarna het de voorpagina's van enige kranten haalde. Vaak werden hier zeer voorbarige conclusies getrokken, zoals dat nu ieder computerprogramma versneld kan worden, dat het handelsreizigersprobleem eenvoudig op te lossen is, dat weersvoorspellingen op langere termijn mogelijk zijn, en dat geheime codes nu sneller breekbaar zijn. De computerfabrikanten reageerden minder enthousiast; zij benadrukten terecht dat de nieuwe methode geen alternatief vormt voor de simplex-methode, en dat de bestaande software-pakketten hun waarde dus behouden, hoewel hierbij soms de valse vergelijking werd gemaakt tussen het "average case"-gedrag van de simplex-methode en het "worst case"-gedrag van de ellipsoïde-methode (vgl. McGall [31]).

Uiteraard richt een groot deel van het door Khachian's artikel aangezette onderzoek zich op de vraag hoe de nieuwe methode praktisch toepasbaar gemaakt kan worden. De totnogtoe gepubliceerde verbeteringen lijken de snelheid van de methode en haar numerieke stabiliteit echter nauwelijks te verbeteren. Daarnaast werden en worden verdere toepassingen van de methode onderzocht. Zo bleek de methode ook de polynomiale oplosbaarheid van convexe kwadratische programmering te geven (Kozlov, Tarasov en Khachian [26]), en verder tot een aantal nieuwe inzichten in de combinatorische optimalisering te leiden (Grötschel, Lovász en Schrijver [18], Padberg en Rao [35], Karp en Papadimitriou [22]). Voor een overzicht van het onderzoek naar de nieuwe methode verwijzen we naar de bibliografie van Wolfe [38] en naar het uitgebreide artikel van Bland, Goldfarb en Todd [2].

In dit artikel geven we een bespreking van de ellipsoide-methode, ingedeeld in de volgende hoofdstukken:

1. Lineaire programmering,
2. De simplex-methode,
3. Vooraf aan de ellipsoide-methode,
4. De ellipsoide-methode,
5. De vooronderstellingen,
6. De kleinste ellipsoide,
7. E_N is klein genoeg,
8. Precisie en praktische toepasbaarheid,
9. Optimaliserings- en scheidingsalgorithmen,
10. Kwadratische programmering,
11. Toepassingen in de combinatorische optimalisering,
12. Perfecte grafen en submodulaire functies,
13. Geheeltallige lineaire programmering.

1. LINEAIRE PROGRAMMERING

Een van de vele mogelijke verschijningsvormen van een lineair programmeringsprobleem (LP-probleem) is: bepaal

(1) $$M = \max \{cx \mid Dx \leq b\}.$$

Hierin is D een m×n-matrix, b een m-vector en zijn c en x n-vectoren, en is cx het inwending product van c en x. Zonder beperking van de algemeenheid mogen we aannemen dat D, c en b geheeltallig zijn. Gevraagd wordt een algorithme om (1) te bepalen waarvan de looptijd begrensd wordt door een polynoom in de afmetingen van het LP-probleem, d.w.z. in

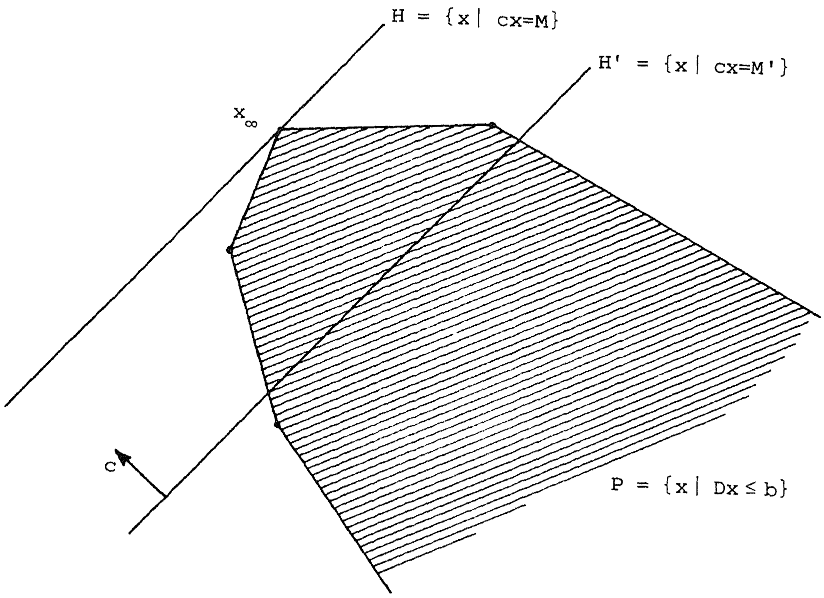
(2) $$(m+1)(n+1)\log T,$$

waarbij T gelijk is aan het grootste getal (in absolute waarde) +1, dat voorkomt in D, c en b.

"$Dx \leq b$" representeert een stelsel lineaire ongelijkheden, en de oplossingsverzameling

(3) $$P := \{x \mid Dx \leq b\}$$

is een polyeder in $\mathbb{Q}^n$. Meetkundig kan het LP-probleem (1) dan worden voorge-



Figuur 1.

steld als het parallel opschuiven van het hypervlak loodrecht op de vector c, net zolang als dit hypervlak nog punten van P bevat. Als x_∞ een punt van P is dat zich in het "laatste" hypervlak bevindt, dan is cx_∞ de oplossing voor (1). (Dit punt x_∞ hoeft niet uniek te zijn, noch hoeft zo'n laatste hypervlak altijd te bestaan.)

Meetkundig kan men zich eenvoudig voorstellen dat het laatste hypervlak H een niet-negative lineaire combinatie is van de facetten van P die in x_∞ samenkomen. Dit impliceert dat een vector y bestaat zo dat:

(4) $\quad y \geq 0,\ yD = c,\ yb = M.$

Dus

(5) $\quad M \geq \min\ \{yb \mid y \geq 0,\ yD = c\}.$

Omdat het maximum (1) niet echt groter kan zijn dan dit minimum (omdat $cx = yDx \leq yb$), volgt de Dualiteitsstelling van de lineaire programmering (Von Neumann [32], Gale Kuhn en Tucker [12]):

(6) $\quad \max\ \{cx \mid Dx \leq b\} = \min\ \{yb \mid y \geq 0,\ yD = c\}.$

Deze Dualiteitsstelling impliceert dat LP $\in$ *NP*$\cap$*coNP*, d.w.z. dat in polynomiaal begrensde tijd *bewezen* kan worden dat het maximum (1) een zekere waarde M heeft: hiervoor hoeft men slechts een x_∞ en een y_∞ te specificeren zo dat

(7) $\quad Dx_\infty \leq b,\ y_\infty \geq 0,\ y_\infty D = c,\ cx_\infty = M = y_\infty b,$

welke berekeningen in polynomiaal-begrensde tijd kunnen worden uitgevoerd. Hoe men dit bewijs, d.w.z. zo'n x_∞ en y_∞ kan *vinden* is een moeilijker probleem; Khachian's methode toont aan dat ook dit probleem in polynomiale tijd kan worden opgelost.

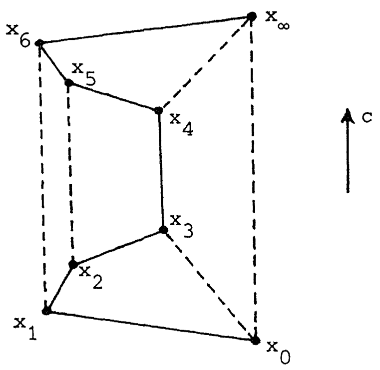
2. DE SIMPLEX-METHODE

Een algorithme om het maximum (1) expliciet te bepalen is de zgn. simplex-methode, ontwikkeld door Dantzig [4]. Meetkundig gezien maakt de simplex-methode een "reis" over het polyeder P door in een hoekpunt van P te beginnen, en dan via ribben van hoekpunt naar hoekpunt te reizen, totdat een optimaal hoekpunt is bereikt. Men kiest de te bereizen ribben zodanig dat de doelfunctie cx steeds toeneemt, althans niet afneemt.

Een meer preciese, algebraïsche uitwerking van deze ruw geschetste methode is vervat in de bekende schema's van "simplex-tableau's" en "pivot-regels", die in een eindig aantal stappen het maximum (1) opleveren. (Men moet een subroutine inbouwen om het "begin-hoekpunt" te vinden, alsmede een regel om het zgn. "cycling" te voorkomen.)

In de praktijk blijkt de simplex-methode een zeer efficiënte algorithme. Dantzig [5] rapporteert dat het aantal pivot-stappen meestal ongeveer $\frac{3}{2}m$ bedraagt, en vrijwel nooit meer dan 3m. Een recent resultaat van Dantzig [6] zegt dat onder zekere voorwaarden het gemiddelde aantal pivot-stappen (over alle LP-problemen) $m.\log m$ is. Dus gemiddeld is de simplex-methode zeker polynomiaal.

De simplex-methode blijkt echter een van de weinige bekende algorithmen te zijn waarvoor het "worst case"-gedrag veel slechter is dan het "average case"-gedrag. Klee en Minty [24] vonden een klasse van LP-problemen, met $m = 2n$, waarvoor bij een slechte keuze van de pivot-elementen, 2^n-1 pivot-stappen moeten worden gezet. Het bijbehorende polyeder P is een kubus met schuine zijvlakken (vgl. Figuur 2 voor n = 3).



Figuur 2

Als de vector c evenwijdig loopt aan de ribbe x_0x_∞, dan kan de keuze van de pivot-stappen leiden tot een reis van x_0 naar x_∞ via de omweg $x_1,x_2,x_3,x_4,$ x_5,x_6. Nu kan men natuurlijk ook rechtstreeks van x_0 naar x_∞ reizen, en er zijn verschillende pivot-regels opgesteld om te dwingen tot de overeenkomstige pivot-keuze (zoals "best improvement", "steepest edge"), maar ook voor deze pivot-regels zijn LP-problemen opgesteld die exponentieel veel pivot-stappen vergen (zie Jeroslow [19], Goldfarb en Sit [15]), hoofdzakelijk door de randjes van de kubus van Klee en Minty "af te schaven" zodat afsnijpaadjes over het hoofd worden gezien.

Het is nog een open vraag of er enige pivot-regel kan bestaan die ieder LP-probleem in polynomiale tijd oplost. Hieraan verwant is het nog onopgeloste probleem of er een polynoom p(n,m) bestaat zo dat op ieder polytoop in $\mathbb{R}^n$ tussen ieder tweetal hoekpunten een pad via ribben bestaat met ten hoogste p(m,n) ribben, waarbij m het aantal facetten van P is (zie Barnette [1]).

3. VOORAF AAN DE ELLIPSOIDE-METHODE

Alvorens een beschrijving te geven van de ellipsoide-methode, enige voorbereidende opmerkingen over ellipsoïden, vooronderstellingen en hoekpunten van P.

Ellipsoïden. Een *ellipsoïde* is een verzameling vectoren van de vorm

$$E = \{x \mid (x-x_0)^T A^{-1}(x-x_0) \leq 1\}, \tag{8}$$

waarbij x_0 een vaste vector is (het *middelpunt* van E), en A een positief-definiete matrix (d.w.z. A is symmetrisch en heeft alleen positive eigenwaarden, hetgeen equivalent is met: $A = B^T B$ voor zekere nonsinguliere matrix B). De ellipsoïden zijn precies de affiene transformaties van de eenheidsbol.

Vooronderstellingen. Om de beschrijving van de ellipsoïde-methode te vereenvoudigen maken we de volgende vooronderstellingen:

(i) P is volledig-dimensionaal (d.w.z. P is niet bevat in een hypervlak);

(ii) P is begrensd;

(iii) $\max\{cx \mid x \in P\}$ wordt aangenomen in precies één hoekpunt van P.

In §5 zullen we laten zien dat deze vooronderstellingen zonder beperking der algemeenheid mogen worden gemaakt. Ieder LP-probleem kan in polynomiaal begrensde tijd worden omgezet in een LP-probleem met de eigenschappen (i), (ii) en (iii).

De hoekpunten van P. Het is eenvoudig in te zien dat ieder hoekpunt x_0 van P voldoet aan een stelsel lineaire vergelijkingen van de vorm:

(9) $$D_0 x_0 = b_0,$$

waarbij D_0 een niet-singuliere deelmatrix van D is, en b_0 de overeenkomstige deelvector van b. Dit volgt uit het feit dat x_0 in de doorsnede van de in x_0 samenkomende facetten van P ligt. Het stelsel vergelijkingen (9) impliceert dat de coördinaten van x_0 te schrijven zijn als quotiënten van deeldeterminanten van de matrix $[D_0\ b_0]$. Aangezien deze determinanten in absolute waarde ten hoogste

(10) $$Q := n^n T^n$$

zijn, weten we dat de absolute waarden van de tellers en noemers die in x_0 voorkomen eveneens ten hoogste Q zijn. Bovendien volgt

(11) $$\|x_0\|^2 \le R := n^{3n} T^{2n}$$

(dit is een ruwe schatting). Merk op dat de afmetingen van Q en R (d.w.z. log Q en log R) polynomiaal begrensd worden door de afmetingen van het LP-probleem.

4. DE ELLIPSOIDE-METHODE

De ellipsoïde-methode construeert een rij ellipsoïden

(12) $$E_0, E_1, E_2, \ldots, E_N,$$

d.w.z. rijen positief-definiete matrixen en middelpunten (vgl. (8))

(13) $$A_0, A_1, A_2, \ldots, A_N,$$
$$x_0, x_1, x_2, \ldots, x_N,$$

waarbij

(14) $$N := 20\lceil n^2.\log 12 + 13n^5.\log n + 10n^5.\log T\rceil,$$

zo dat $x_\infty \in E_k$ voor iedere k, als volgt.

E_0 is een bol die P geheel omvat, bijvoorbeeld

(15) $$E_0 = \{x \mid x^T x \le R\}$$

vanwege (11). Dus dan is $A_0 = R.I$ en $x_0 = \underline{0}$. Als $E_0, \ldots, E_k$ gedefinieerd zijn, dan wordt E_{k+1} als volgt gevonden.

I. Als $x_k \notin P$, dan is er een ongelijkheid

(16) $\quad d_i x \leq b_i$

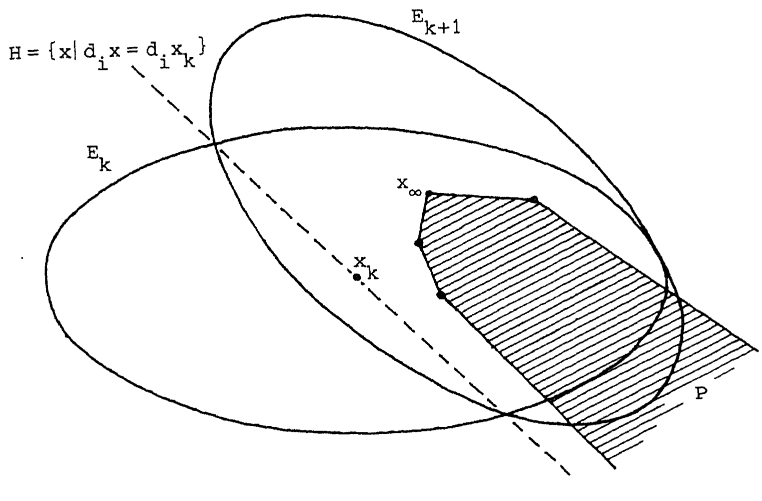
in het stelsel $Dx \leq b$ waaraan x_k niet voldoet, d.w.z.

(17) $\quad d_i x_k > b_i$.

Dan is E_{k+1} de ellipsoïde met kleinste volume zo dat

(18) $\quad E_{k+1} \supset E_k \cap \{x \mid d_i x \leq d_i x_k\}$

(vgl. Figuur 3).



Figuur 3.

Het blijkt dat een dergelijke kleinste ellipsoïde uniek is, en dat de bijbehorende A_{k+1} en x_{k+1} eenvoudig uit A_k, x_k en d_i berekend kunnen worden - zie §6. Uit (16), (17) en (18) volgt direct dat

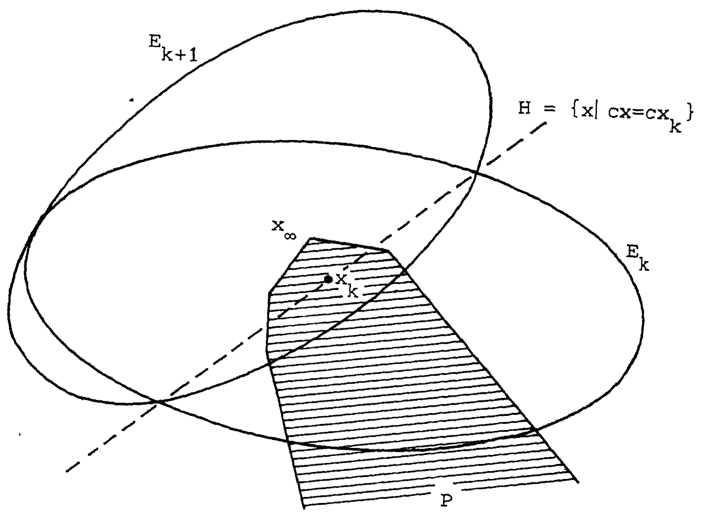
(19) $\quad P \cap E_k \subset P \cap E_{k+1}$.

Dus als $x_\infty \in E_k$ dan ook $x_\infty \in E_{k+1}$.

II. Als $x_k \in P$, dan wordt in principe dezelfde constructie uitgevoerd, maar nu met de doelvector c in plaats van d_i. Nu is E_{k+1} de kleinste ellipsoïde zo dat

(20) $$E_{k+1} \supset E_k \cap \{x \mid cx \geq cx_k\}$$

(vgl. Figuur 4), welke ellipsoïde weer uniek is.



Figuur 4

Omdat $cx_\infty \geq cx_k$ en $x_\infty \in E_k$ volgt dat $x_\infty \in E_{k+1}$.

Zowel in geval I als in geval II hebben we de helft van E_k benaderd met een nieuwe ellipsoide E_{k+1}. Hoewel het volume van de nieuwe ellipsoïde groter is dan de helft van het volume van de oude ellipsoïde, kan worden aangetoond dat het volumen met een vaste factor < 1 afneemt:

(21) $$\frac{\text{vol } E_{k+1}}{\text{vol } E_k} \leq \left(\frac{1}{2}\right)^{1/20n}$$

(zie §6). Omdat vol $E_0 \leq (2R)^n$, weten we dat

(22) $\quad \text{vol } E_N \leq \left(\frac{1}{2}\right)^{N/20n} . (2R)^n.$

Laat nu x_K het laatste element in de rij $x_0, \ldots, x_N$ zijn dat in P zit, d.w.z. zo dat geval II werd toegepast. Uit (19) volgt dat $x_K \in E_N$. Verder weten we dat $x_\infty \in E_N$. Het principe van de ellipsoïde-methode is nu dat vanwege (22) de ellipsoide E_N dermate klein is dat x_K een goede benadering voor x_∞ is. Bewezen kan worden dat

(23) $\quad \|x_K - x_\infty\| < \frac{1}{2Q^2}$

(zie §7). Omdat de noemers van de in x_∞ voorkomende coördinaten ten hoogste Q zijn, en omdat geen twee rationale getallen met noemers ten hoogste Q dichter dan $1/Q^2$ bij elkaar liggen, is x_∞ de enige vector, met noemers $\leq Q$, die aan (23) voldoet.

De coördinaten van x_∞ kunnen expliciet worden gevonden door de coördinaten van x_K stuk voor stuk te benaderen door een kettingbreukontwikkeling. De laatste benadering waarvan de noemer niet groter dan Q is, is de gezochte overeenkomstige coördinaat van x_∞.

Na deze ruwe schets van de ellipsoïde-methode moeten we het volgende nog preciseren:

- dat de vooronderstellingen zonder beperking van de algemeenheid mogen worden gemaakt - zie §5;
- hoe de ellipsoïde E_{k+1} uit de ellipsoïde E_k wordt bepaald - zie §6;
- dat de ellipsoïde E_N klein genoeg is om zeker te zijn van (23) - zie §7.

5. DE VOORONDERSTELLINGEN

We laten nu zien dat de in §3 gemaakte vooronderstellingen zonder beperking van de algemeenheid gemaakt mogen worden. De hieronder gegeven technieken om het oorspronkelijke LP-probleem te transformeren zijn van min of meer theoretisch belang. In de praktijk weet men vaak a priori dat aan de vooronderstellingen is voldaan, of kunnen deze op een eenvoudiger manier worden verkregen.

Ten eerste mogen we aannemen dat het polyeder P geen affiene deelruimten van dimensie ten minste 1, omvat. Als we hier niet zeker van zijn kunnen we (1) vervangen door

(24) $\quad \max \{cx - cx' \mid x, x' \geq 0,\ Dx - Dx' \leq b\}.$

Weliswaar verdubbelt de probleemgrootte hierdoor, maar deze blijft polynomiaal begrensd door de oorspronkelijke probleemgrootte.

Als P geen deelruimten omvat, dan wordt het maximum (1) bereikt in een hoekpunt van P (en misschien ook in andere punten van P), of is oneindig. Dus dan weten we dat, als M eindig is,

(25) $$|M| \leq nTQ \leq n^{2n}.T^{2n}.$$

Als we nu niet weten of P volledig-dimensionaal is, kunnen we (1) vervangen door

(26) $$\max \{cx - (3n^{3n}T^{3n})\lambda \mid \lambda \geq 0,\ -\lambda + Dx \leq b\},$$

waarbij λ een nieuwe variable is. Dan omvat het nieuwe polytoop

(27) $$P' = \{(x,\lambda) \mid \lambda \geq 0,\ -\lambda + Dx \leq b\}$$

weer geen deelruimten, dus (26) wordt bereikt in een hoekpunt (x_0,λ_0) van P'. Weer kan eenvoudig worden ingezien dat de tellers en noemers van de hoekpunten van P' ten hoogste Q in absolute waarde zijn. Dus als $\lambda_0 > 0$, dan is $\lambda_0 \geq 1/Q$, en is (26) ten hoogste

(28) $$nTQ - \frac{3n^{3n}T^{3n}}{Q} \leq -2n^{2n}T^{2n}.$$

Dus als (26) groter is dan $-2n^{2n}T^{2n}$, dan wordt dit maximum bereikt in een hoekpunt (x_0,λ_0) met $\lambda_0 = 0$, en is x_0 een oplossing voor (1). Als (26) kleiner is dan $-2n^{2n}T^{2n}$, dan is, volgens (25), $\lambda > 0$, en dan bestaat er kennelijk geen x met $Dx \leq b$. Bovendien kan eenvoudig worden ingezien dat (26) eindig is als en alleen als (1) eindig is.

Merk op dat het polyeder P' volledig-dimensionaal is, en dat de afmeting van het LP-probleem (26) polynomiaal begrensd wordt door de afmeting van het oorspronkelijke LP-probleem.

We mogen dus aannemen dat P volledig-dimensionaal is en geen affiene deelruimten omvat. Ook mogen we aannemen dat P begrensd is. Als we dit niet weten, kunnen we probleem (1) vervangen door de problemen: bepaal

(29) $$M_1 = \max \{cx \mid Dx \leq b,\ -2Q \leq x^i \leq +2Q\ (i = 1,\dots,n)\} \text{ en}$$
$$M_2 = \max \{cx \mid Dx \leq b,\ -3Q \leq x^i \leq +3Q\ (i = 1,\dots,n)\}.$$

Eenvoudig kan worden ingezien dat als $M_1 = M_2$ dan is $M = M_1$, en als $M_2 > M_1$ dan is $M = \infty$.

Tenslotte mogen we aannemen dat het maximum (1) in precies één hoekpunt van P wordt aangenomen. Vervang anders c door

$$(30) \qquad c' = S^n.c + (1,S,S^2,\dots,S^{n-1}),$$

waarbij $S = 4.Q^{3n}$. Dan wordt $\max\{c'x \mid x \in P\}$ in precies één hoekpunt van P aangenomen, zeg in x_∞, en bovendien geldt dat $M = cx_\infty$.

Aangezien elk van deze transformaties in polynomiaal-begrensde tijd kan worden uitgevoerd, kan het oorspronkelijke LP-probleem in polynomiale tijd worden omgevormd tot een probleem dat aan de genoemde vooronderstellingen voldoet.

6. DE KLEINSTE ELLIPSOIDE

We laten nu zien hoe de parameters A_{k+1} en x_{k+1} voor de ellipsoïde E_{k+1} berekend kunnen worden uit die voor E_k, en dat het volume van E_{k+1} klein genoeg is ten opzichte van dat van E_k.

De ellipsoïde E_k werd gegeven door

$$(31) \qquad E_k = \{x \mid (x-x_k)^T A_k^{-1}(x-x_k) \le 1\}.$$

Verder werd een halfruimte H gegeven door, zeg,

$$(32) \qquad H = \{x \mid ax \ge ax_k\},$$

voor zekere vector a ($a = -d_i$ en $a = c$ in geval I, resp. II). Nodig is de kleinste ellipsoïde E_{k+1} te bepalen zo dat

$$(33) \qquad E_{k+1} \supset E_k \cap H.$$

Nu bestaat er, zoals eerder werd opgemerkt, een affiene transformatie die de ellipsoïde E_k overbrengt naar de eenheidsbol $\{x \mid x^T x \le 1\}$, en de halfruimte H naar de halfruimte $\{x \mid x^1 \ge 0\}$ (waarbij $x = (x^1,x^2,\dots,x^n)$). Het is een eenvoudige meetkundige opgave de unieke kleinste ellipsoïde E te bepalen zo dat

$$(34) \qquad E \supset \{x \mid x^T x \le 1,\ x^1 \ge 0\}.$$

Omdat de volumens van meetkundige figuren onder een affiene transformatie met een constante factor worden vermenigvuldigd (nl. met de absolute waarde van de determinant van de bijbehorende matrix), krijgen we door de inverse affiene transformatie op E toe te passen de kleinste ellipsoïde E_{k+1} die $E_k \cap H$ omvat.

Uitwerking geeft de volgende formules voor A_{k+1} en x_{k+1}:

$$(35) \qquad A_{k+1} = \frac{n^2}{n^2-1}\left(A_k - \frac{2}{n+1} b_k b_k^T\right),$$

$$(36) \qquad x_{k+1} = x_k + \frac{1}{n+1} b_k,$$

waarbij

$$(37) \qquad b_k = A_k a/\sqrt{a^T A_k a}.$$

Nu weten we dat

$$(38) \qquad \frac{\text{vol } E_{k+1}}{\text{vol } E_k} = \sqrt{\frac{\det A_{k+1}}{\det A_k}} = \left(\frac{n^2}{n^2-1}\right)^{\frac{1}{2}n}\left(\frac{n-1}{n+1}\right)^{\frac{1}{2}} \leq (\tfrac{1}{2})^{1/20n}$$

omdat we zonder beperking der algemeenheid mogen aannemen dat $A_k = I$ en $a = (1,0,\ldots,0)^T$. Uit (38) volgt

$$(39) \qquad \text{vol } E_N = \leq (\tfrac{1}{2})^{N/20n} \cdot (2R)^n$$

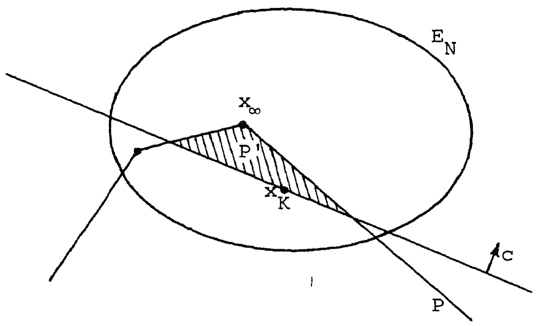
Er schuilt een adder onder het gras: om b_k te berekenen moeten we de wortel trekken. Dit impliceert dat we bij toepassing van de methode op een eindige precisie-computer moeten afronden, en dat we daardoor de ellipsoïde E_{k+1} iets ruimer moeten nemen om afrondingsfouten te neutraliseren. Het zal blijken dat we kunnen volstaan met een afronding tot op een polynomiaal aantal bits, en dat de iets ruimer gekozen ellipsoïden ook de gewenste polynomiale algorithme voor lineaire programmering opleveren - zie §8.

7. E_N IS KLEIN GENOEG

We laten nu zien dat E_N inderdaad klein genoeg is om te weten dat de afstand tussen de vectoren x_K en x_∞ kleiner dan $1/2Q^2$ is.

We weten dat $x_K \in E_N$, $x_\infty \in E_N$ en dat

$$(40) \qquad P' := P \cap \{x \mid cx \geq cx_K\} \subset E_N.$$



Figuur 5

We bewijzen dat alle punten van P' dichter dan $1/2Q^2$ bij x_∞ liggen. Want stel dat een punt $x'_K \neq x_\infty$ in P' bestaat zo dat

(41) $\qquad \|x'_K - x_\infty\| \geq 1/2Q^2.$

Zonder beperking der algemeenheid mogen we aannemen dat x'_K een hoekpunt van P' is. Als P' nog andere hoekpunten van P bevat behalve x_∞, dan kunnen we voor x'_K een hoekpunt van P ongelijk x_∞ nemen zo dat cx'_K zo groot mogelijk is (hoekpunten van P liggen nooit dichter dan $1/2Q^2$ bij elkaar).

Laat $D'x \leq b'$ het stelsel van die ongelijkheden uit $Dx \leq b$ zijn die voor $x = x_\infty$ overgaan in gelijkheid. Bekijk het polytoop

(42) $\qquad P'' = \{x \mid D'x \leq b', \ cx \geq cx_\infty - 1\}.$

P" is begrensd omdat $\max\{cx \mid x \in P\}$ in precies één hoekpunt van P wordt aangenomen. P" is volledig-dimensionaal omdat P volledig-dimensionaal is.

Uit een redenering analoog aan die gemaakt in §3 volgt dat de tellers en de noemers van de coördinaten van de hoekpunten van P" in absolute waarde niet groter dan $n^{2n^2}T^{2n^2}$ zijn. Hieruit volgt ten eerste dat

(43) $\qquad \text{vol } P'' \geq (n^{5n^4}.T^{4n^4})^{-1},$

omdat P" n+1 affien onafhankelijke hoekpunten heeft, waarvan het convex omhulsel een volumen heeft gelijk aan 1/n! maal de determinant van een matrix met noemers ten hoogste $n^{4n^2}T^{4n^2}$, dus dit volumen is ten minste het rechterlid van (43). Ten tweede volgt dat ieder tweetal hoekpunten van P", en dus ook ieder tweetal willekeurige punten in P", een afstand ten hoogste

$$(44)\qquad 2n^{3n^2}.T^{2n^2}$$

hebben. Omdat

$$(45)\qquad P' \supset \{x \mid D'x \le b',\ cx \ge cx_\infty - (cx_\infty - cx_K')\},$$

volgt dat

$$(46)\qquad \text{vol } P' = (cx_\infty - cx_K')^n.\text{vol } P'' \ge (cx_\infty - cx_K')^n (n^{5n^4}.T^{4n^4})^{-1}.$$

Evenzo bevat het rechterlid van (45) geen twee punten op afstand groter dan

$$(47)\qquad (cx_\infty - cx_K')\,2n^{3n^2}T^{2n^2}.$$

In het bijzonder volgt de tegenspraak

$$(48)\qquad \frac{1}{2Q^2} \le \|x_\infty - x_K'\| \le (cx_\infty - cx_K')\,2n^{3n^2}T^{2n^2} \le 2n^{3n^2}T^{2n^2}n^{5n^3}T^{4n^3}(\text{vol } P')^{1/n} \le$$
$$\le 2n^{8n^3}T^{6n^3}(\text{vol } E_N)^{1/n} \le 2n^{8n^3}T^{6n^3}(1/2)^{N/20n^2}\,2R \le 4n^{11n^3}T^{8n^3}(1/2)^{N/20n}$$
$$< \frac{1}{2Q^2},$$

gebruik makend van resp. (41), (47), (46), (40), (39), (11) en (10).

8. PRECISIE EN PRAKTISCHE TOEPASBAARHEID

In §6 merkten we op dat een expliciete, exacte uitvoering van de ellipsoïde-methode zoals hierboven beschreven op een eindige precisie-computer niet mogelijk is omdat bij de overgang van E_k op E_{k+1} wortel moet worden getrokken. Dit kan worden opgelost door alle berekeningen uit te voeren tot op

$$(49)\qquad p := 5N\left\lceil \log \frac{12\sqrt{n}}{R^2}\right\rceil$$

binaire bits nauwkeurig, door de ellipsoïden iets ruimer te kiezen, nl. door A_{k+1} in plaats van als in (35) als volgt te geven

$$(50) \qquad A_{k+1} = \frac{2n^2+3}{2n^2}\left(A_k - \frac{2}{n+1} b_k b_k^T\right),$$

en door N te verhogen tot

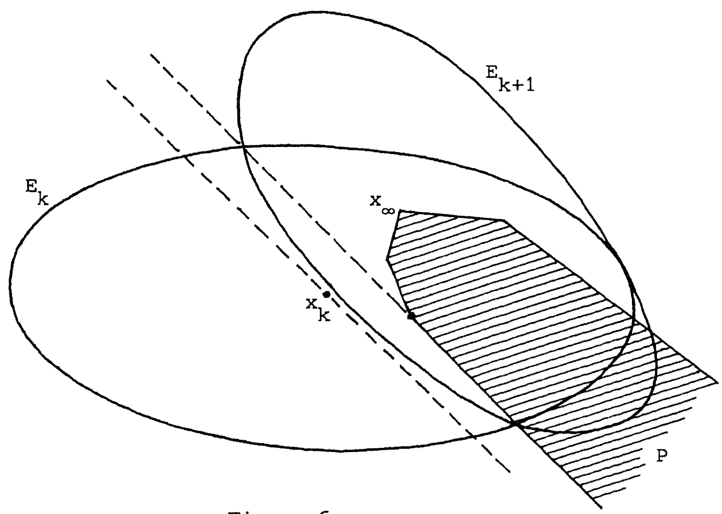
$$(51) \qquad 20\lceil n^5.\log 12 + 13n^8.\log n + 10n^8.\log T\rceil$$

Dit wordt bewezen in [18]. Merk op dat p en N begrensd zijn door een polynoom in de afmeting van het LP-probleem, zodat de hele methode uitvoerbaar is binnen polynomiaal begrensde tijd.

Toch zijn zowel het aantal stappen N als de vereiste nauwkeurigheid p veel te groot voor de computers van dit moment om de ellipsoïde-methode in deze vorm ook maar enige praktische toepasbaarheid toe te schrijven. Weliswaar zijn hierboven, ter bevordering van de eenvoud, de berekeningen aan de ruime kant uitgevoerd, maar ook scherpere afschattingen van N en p lijken de methode niet essentieel te verbeteren.

Er zijn verschillende wijzigingen voorgesteld om de efficiëntie en de numerieke stabiliteit van de methode te verhogen, zoals:

(i) In plaats van de ellipsoïde E_k te "halveren" d.m.v. een hypervlak door het middelpunt x_k, en voor E_{k+1} de kleinste ellipsoïde te nemen die de "goede" helft van E_k omvat, kunnen we het hypervlak opschuiven tot tegen P.



Figuur 6

D.w.z., als geval I zich voordoet (cf. §4), en $d_i x \le b_i$ is een vergelijking uit het stelsel $Dx \le b$ zo dat $d_i x_k > b_i$, dan is E_{k+1} de kleinste ellipsoïde met

(52) $$E_{k+1} \supset E_k \cap \{x \mid d_i x \le b_i\}.$$

De bij deze "diepe snede-methode" behorende formules voor A_{k+1} en x_{k+1} zijn bepaald door Padberg en Rao [33] en verschillende anderen (zie Wolfe [38]). Men kan deze methode nog verscherpen door niet één ongelijkheid uit het stelsel $Dx \le b$ te kiezen, maar een niet-negatieve lineaire combinatie $d'x \le b'$ van ongelijkheden uit $Dx \le b$, zo dat $d'x_k > b'$, en zo dat $E_{k+1} \supset$ $\supset E_k \cap \{x \mid d'x \le b'\}$ zo klein mogelijk gekozen kan worden. Dit is de methode van de "surrogaat sneden" - zie Goldfarb en Todd [16]. Bland, Goldfarb en Todd [2] lieten zien dat ook deze wijzigingen niet leiden tot een essentiële verbetering van de ellipsoïde-methode.

(ii) Als te grof wordt afgerond, d.w.z. als p te klein wordt gekozen, dan hoeven de matrixen A_k niet positief-definiet te blijven, zodat de deling door $\sqrt{a^T A_k a}$ als in (37) niet steeds uitvoerbaar is. Dit kan verholpen worden door A_k te schrijven als $A_k = J_k J_k^T$, waarbij J_k een niet-singuliere matrix is, en door de rij $A_0, A_1, \ldots$ te vervangen door $J_0, J_1, \ldots$. Zie Bland, Goldfarb en Todd [2] voor de expliciete formules.

Daarnaast zijn allerlei verbeteringen mogelijk door het onderhavige LP-probleem nader te beschouwen. Vaak zal bijvoorbeeld de begin-ellipsoïde E_0 zuiniger gekozen kunnen worden dan hierboven, doordat het polytoop P a priori beter begrensd kan worden.

Maar basisidee van de ellipsoïde-methode is te eindigen met een zeer kleine ellipsoïde E_N, hetgeen als basisproblemen een grote precisie en een groot aantal halveringen onvermijdelijk lijken te maken.

Een probleem is ook dat men in het algemeen niet eerder dan na N stappen zeker is dat men dichtbij het optimum x_∞ is. Een eenvoudige omvorming van de methode geeft ons echter ook een polynomiale methode om te beslissen of het open polyeder

(53) $$P^o = \{x \mid Dx < b\}$$

leeg is of niet, en zo niet, een punt in P^o te vinden. Eveneens kan ieder LP-probleem worden teruggebracht tot het probleem een oplossing voor een stelsel strikte lineaire ongelijkheden $Dx < b$ te vinden. Het voordeel van deze alternatieve formulering is dat vaak niet alle N stappen hoeven worden uitgevoerd, maar dat gestopt kan worden zodra een x_k gevonden is met $Dx_k < b$. In feite is dit de oorspronkelijke methode van Khachian [23] - zie Gács en Lovász [11]. Vergelijkt men de ellipsoïde-methode met de simplex-methode dan lijkt het daarom niet helemaal eerlijk de zeere grote N te vergelijken met de $\frac{3}{2}m$ pivot-stappen van de simplex-methode, d.w.z. een theoretische bovengrens voor het "worst case"-gedrag van de ellipsoïde-methode met het praktische "average case"-gedrag van de simplex-methode.

Het moet overigens niet uitgesloten worden geacht dat een combinatie van simplex- en ellipsoïde-methode nog eens zal leiden tot een zowel in praktisch als in theoretisch opzicht efficiënte algorithme voor lineaire programmering. De simplex-methode heeft het voordeel dat de bezochte punten "mooi" blijven, nl. hoekpunten van het polyeder, maar het nadeel dat een verkeerde keuze van de reisrichting (d.w.z. van de pivot-elementen) tot lange reizen kan leiden. Omgekeerd heeft de ellipsoïde-methode het voordeel dat de gevolgde reisrichting in polynomiale tijd tot het optimum leidt, maar het nadeel dat de punten $x_0, x_1, \ldots$ grote noemers kunnen hebben. De vraag is dus: kunnen de beide voordelen gecombineerd worden?

9. OPTIMALISERINGS- EN SCHEIDINGSALGORITHMEN

In de inleiding werd al opgemerkt dat het toepassingsgebied van de ellipsoïde-methode ruimer is dan lineaire programmering. In de rest van dit artikel geven we hiervan een ruwe schets.

Stel dat een begrensde, gesloten, volledig-dimensionale, convexe verzameling P in $\mathbb{Q}^n$ is gegeven, en dat we het volgende probleem willen oplossen.

Optimaliseringsprobleem. Gegeven $c \in \mathbb{Q}^n$, bepaal $\max\{cx \mid x \in P\}$.

Nadere beschouwing van de hierboven gegeven ellipsoïde-methode leert dat dit optimaliseringsprobleem kan worden benaderd in polynomiaal begrensde tijd als het volgende probleem kan worden opgelost in polynomiaal begrensde tijd.

Scheidingsprobleem. Gegeven $y \in \mathbb{Q}^n$, bepaal of $y \in P$, en zo niet, vind een $a \in \mathbb{Q}^n$ zo dat $ay > \max\{ax \mid x \in P\}$.

Dit scheidingsprobleem vraagt dus om een hypervlak dat y van P scheidt. We kunnen een polynomiale algorithme voor het scheidingsprobleem invoeren als subroutine in de ellipsoïde-methode, nl. om te beslissen tussen de gevallen I en II uit §4, en om in geval I een halveringshypervlak te vinden. Omdat de ellipsoïde-methode niet meer vraagt over P dan dit, geeft dit ons een polynomiale algorithme om het optimaliseringsprobleem te benaderen. (In feite zou een convexe verzameling P in de computer gedeclareerd kunnen worden als een algorithme voor het scheidingsprobleem.) Dit wordt precieser geformuleerd en bewezen in [18]. Noodzakelijk is dan een punt a_0 en getallen r en R te weten zo dat

(54) $\quad S(a_0,r) \subset P \subset S(a_0,R),$

waarbij $S(a_0,\rho)$ de bol om a_0 met straal ρ voorstelt. "Polynomiaal" betekent: de looptijd wordt begrensd door een polynoom in $|\log r|$, $|\log R|$, $\log T$ en $|\log \varepsilon|$, waarbij T het grootste getal is dat voorkomt in tellers en noemers van a_0, c of y (in absolute waarde), en waarbij ε de nauwkeurigheid is die het antwoord moet hebben. D.w.z., het optimaliseringsprobleem vraagt om een x_∞ zo dat $d(x_\infty,P) < \varepsilon$ en $cx_\infty + \varepsilon > \max\{cx \mid x \in P\}$, en het scheidingsprobleem vraagt te bepalen of $d(y,P) < \varepsilon$, of een a in $\mathbb{Q}^n$ te vinden zo dat $ay + \varepsilon > \max\{ax \mid x \in P\}$.

De ellipsoïde-methode geeft ons dus de implicatie:

(55) $\quad \exists$ polynomiale scheidingsalgorithme $\Rightarrow \exists$ polynomiale optimaliseringsalgorithme.

De implicatie geldt echter ook in de omgekeerde richting. We kunnen zonder beperking der algemeenheid aannemen dat het hierboven omschreven punt a_0 gelijk is aan de oorsprong $\underline{0}$. Als we nu P vervangen door de duale convexe verzameling

(56) $\quad P^* = \{y \mid xy \le 1 \text{ voor iedere } x \text{ in } P\},$

dan blijken scheidings- en optimaliseringsprobleem in elkaar over te gaan. Dus dan:

(57) $\quad \exists$ polynomiale optimaliseringsalgorithme voor P $\Rightarrow \exists$ polynomiale scheidingsalgorithme voor $P^* \Rightarrow \exists$ polynomiale optimaliseringsalgorithme voor $P^* \Rightarrow \exists$ polynomiale scheidingsalgorithme voor P,

en daarom:

(58) $\exists$ polynomiale scheidingsalgorithme voor P $\Leftrightarrow$ $\exists$ polynomiale optimaliseringsalgorithme voor P.

In de volgende paragrafen geven we een paar toepassingen van dit principe.

10. KWADRATISCHE PROGRAMMERING

Een van de verschijningsvormen van een kwadratisch programmeringsprobleem (QP-probleem) is: bepaal

(59) $$\min \{x^T Bx + cx \mid Dx \le b\}.$$

Hierin is $Dx \le b$ weer een stelsel lineaire ongelijkheden, en cx het inwendig product van c en x, en B is een symmetrische positief semi-definiete $n \times n$-matrix. Weer mogen we zonder de algemeenheid te beperken aannemen dat B, c, D en b geheeltallig zijn. We kunnen (59) benaderen met de ellipsoïde-methode als volgt. Het QP-probleem is equivalent met: bepaal

(60) $$\min \{\lambda \mid Dx \le b,\ \lambda \ge x^T Bx + cx\}.$$

Als we definieren

(61) $$P = \{(x,\lambda) \mid Dx \le b,\ \lambda \ge x^T Bx + cx\},$$

dan is P een gesloten convexe verzameling, en is (60) een speciaal geval van het optimaliseringsprobleem voor P. De verzameling P kan worden opgevat als de grafiek van de functie $x^T Bx+cx$ gedefinieerd op $\{x \mid Dx \le b\}$, aangevuld met alle punten die boven de grafiek liggen. We kunnen P begrensd maken door P te vervangen door $P \cap \{(x,\lambda) \mid \lambda \le U\}$, waarbij U groot genoeg wordt gekozen (in het algemeen zal het niet moeilijk zijn een dergelijke U te vinden).

Om (60) in polynomiale tijd te vinden is het volgens de voorgaande paragraaf voldoende te laten zien dat het scheidingsprobleem voor P in polynomiale tijd kan worden opgelost. Dit laatste is niet moeilijk. Als we een (x',λ') kiezen dan kan door substitutie eenvoudig worden nagegaan of $(x',\lambda')\in P$. Als blijkt dat (x',λ') niet in P zit, dan is of aan $Dx' \leq b$ niet voldaan, of aan $\lambda' \geq x'^TBx'+cx'$. In het eerste geval levert de ongelijkheid waaraan niet is voldaan een scheidend hypervlak. In het tweede geval, d.w.z. als $\lambda' < x'^TBx'+cx'$, scheidt het hypervlak

(62) $$\{(x,\lambda) \mid (x'^TB+c)x - \lambda = (x'^TBx'+cx'-\lambda')\}$$

het punt (x',λ') van P, en kunnen we als oplossing voor het scheidingsprobleem de vector $a = (x'^TB+c,-1)$ nemen.

Het scheidingsprobleem voor P kan dus in polynomiale tijd worden opgelost, en dus ook het optimaliseringsprobleem, en daarom bestaat een polynomiale algorithme voor (convexe) kwadratische programmering. Voor een preciesere beschrijving verwijzen we naar Kozlov, Tarasov en Khachian [26].

11. TOEPASSINGEN IN DE COMBINATORISCHE OPTIMALISERING

In [18], [22] en [35] wordt aangetoond hoe de ellipsoïde-methode ook in de combinatorische optimalisering tot een aantal nieuwe inzichten leidt. Een voorbeeld hiervan is het volgende.

Stel we hebben een gerichte graaf $G = (V,E)$, waarin een vast punt r is gekozen. Een *r-vertakking* is een verzameling E' van kanten van G zo dat ieder punt van G bereikbaar is vanuit r via een gericht pad van kanten in E'. (Dus de minimale r-vertakkingen zijn precies de in r gewortelde gerichte bomen.) Nu is het een interessante vraag om, gegeven een "lengte" functie $\ell:E\longrightarrow\mathbb{Z}_+$, een r-vertakking E' te vinden met minimale totale lengte, d.w.z. met

(63) $$\sum_{e\in E'} \ell(e)$$

zo klein mogelijk. Er bestaat een polynomiale algorithme van Fulkerson [10] voor dit probleem, maar met de ellipsoïde-methode kan de polynomiale oplosbaarheid van dit probleem worden herleid tot die van een eenvoudiger probleem.

Definieer

(64) P := het convex omhulsel van de karakteristieke functies van de r-vertakkingen.

P is dus een convex polytoop in $\mathbb{Q}^E$, en het bovenbeschreven minimaliseringsprobleem is equivalent met

(65) $\min\{\ell x \mid x \in P\}$.

Volgens het principe beschreven in §9 kan dit minimum in polynomiaal-begrensde tijd worden bepaald, als er een polynomiale scheidingsalgorithme voor P bestaat. Nu heeft Fulkerson ook de volgende karakterisering van P gegeven:

(66) $$P = \{x \in \mathbb{Q}^E \mid 0 \le x(e) \le 1 \quad (e \in E), \quad \sum_{e \in E'} x(e) \ge 1 \quad (E' \text{ r-snede})\},$$

waarbij een verzameling E' van kanten van G een *r-snede* heet als er een niet-lege verzameling V' van punten van G bestaat zo dat $r \notin V'$ en E' is de verzameling kanten van $V \setminus V'$ naar V'.

Het is duidelijk dat P bevat moet zijn in het rechterlid van (66), omdat de karakteristieke vector van iedere r-vertakking voldoet aan de lineaire ongelijkheden. Fulkerson's stelling geeft de omgekeerde inclusie.

Dus om het scheidingsprobleem voor P op te lossen moeten we, gegeven een x in $\mathbb{Q}^E$, nagaan of x aan de ongelijkheden in (66) voldoet. Eenvoudig is in polynomiale tijd na te gaan of $0 \le x(e) \le 1$ voor iedere kant e. Als er een kant e bestaat die hieraan niet voldoet dan geeft deze kant ons een scheidend hypervlak. Om het tweede stelsel ongelijkheden na te gaan kunnen we niet alle r-sneden E' aflopen en $\sum_{e \in E'} x(e) \ge 1$ controleren, omdat er exponentieel veel (ongeveer $2^{|V|-1}$) r-sneden bestaan. Toch kan dit stelsel in polynomiale tijd worden nagegaan, en wel als volgt. Vat de functie x op als een capaciteitsfunctie op de kanten van G. We kunnen nu zoeken naar een r-snede met minimale capaciteit. Als deze minimale capaciteit ten minste 1 bedraagt besluiten we tot: $x \in P$. Zo niet, dan levert de r-snede met capaciteit kleiner dan 1 ons een scheidend hypervlak.

- Een r-snede met minimale capaciteit kan worden gevonden door voor ieder punt $s \neq r$ een snede E_s van minimale capaciteit te bepalen die r van s scheidt. Deze kan worden gevonden met de minimum-snede algorithme van Ford en Fulkerson [9]. Kiezen we die snede uit $\{E_s \mid s \neq r\}$ met de laagste capaciteit dan hebben we een r-snede met minimale capaciteit.

Zo wordt de polynomiale oplosbaarheid van het "minimum lengte van een r-vertakking"-probleem afgeleid uit die van het "minimum capaciteit van een r-snede"-probleem. Nu kan op analoge wijze omgekeerd worden laten zien dat ook de polynomiale oplosbaarheid van het laatste probleem afgeleid kan worden uit die van het eerste. Op deze manier vinden we een zekere dualiteit tussen combinatorische problemen. Zo blijkt de polynomiale oplosbaarheid van ieder van de onderstaande linkerproblemen equivalent met de polynomiale oplosbaarheid van de rechterproblemen.

(i) min.lengte van een r-vertakking;	(ii) min.capaciteit van een r-snede;
(iii) min.lengte van een pad van r naar s;	(iv) min.capaciteit van een r-s-snede;
(v) min.gewicht van een perfecte matching;	(vi) min.capaciteit van een oneven snede.

(De laatste twee problemen hebben betrekking op een ongerichte graaf G = (V,E) met V even. Een *perfecte matching* is een verzameling disjuncte kanten van G die V overdekken. Een *oneven snede* is de collectie kanten van V' naar V\V', waarbij V' een oneven verzameling punten van G is.)

Voor probleem (iii) is een eenvoudige polynomiale algorithme bekend ontworpen door Dijkstra [7], terwijl probleem (iv) kan worden opgelost met de iets moeilijkere algorithme van Ford en Fulkerson [9]. Voor probleem (v) is een vrij gecompliceerde algorithme opgesteld door Edmonds [8], terwijl probleem (vi) kan worden opgelost door een eenvoudige aanpassing van de "minimum snede"-algorithme van Ford en Fulkerson (zie Padberg en Rao [34]).

Voor meer toepassingen van dit dualiteitsprincipe verwijzen we naar [18]. Weliswaar kunnen de meeste van deze toepassingen, net als het bovenstaande r-vertakkingsprobleem, worden geformuleerd als een LP-probleem (en lijken dus op het eerste gezicht direct oplosbaar met Khachian's methode), maar de formulering van deze LP-problemen vergt in het algemeen al exponentieel veel tijd vanwege het grote aantal lineaire ongelijkheden (bijvoorbeeld r-sneden).

De toepassingen berusten sterk op de polyhedrale karakteriseringen van zekere polytopen, zoals de karakterisering (66) van het convex omhulsel P van de r-vertakkingen. Opgemerkt moet worden dat dergelijke polyhedrale karakteriseringen vaak verkregen werden als bijproduct van de constructie van een polynomiale algorithme. Nu met de ellipsoïde-methode is het juist omgekeerd: een polyhedrale karakterisering kan aanleiding geven tot een polynomiale algorithme, hetgeen verder onderzoek naar dergelijke karakteriseringen motiveert.

Overigens zijn de aldus met de ellipsoïde-methode gevonden algorithmen weliswaar polynomiaal, maar verre van efficiënt, vanwege de hoge graad van het polynoom dat de looptijd begrenst. Zij vormen dan ook geen alternatief voor de eerder ontwikkelde speciale algorithmen. In de volgende paragraaf zullen we echter laten zien dat met de ellipsoïde-methode ook de polynomiale oplosbaarheid van een aantal combinatorische problemen kan worden afgeleid die nog niet eerder in deze zin waren opgelost.

12. PERFECTE GRAFEN EN SUBMODULAIRE FUNCTIES

We zullen nu ingaan op twee combinatorische problemen waarvan de polynomiale oplosbaarheid alleen werd aangetoond met de ellipsoïde-methode. Uitwerking levert weer algorithmen die, hoewel polynomiaal, een zeer lage graad van efficiëntie hebben. Het bewijs van de polynomiale oplosbaarheid kan veeleer worden beschouwd als een motivering om, nu we eenmaal weten dat polynomiale algorithmen bestaan, te zoeken naar werkelijk efficiënte algorithmen. Hierbij is het onwaarschijnlijk dat een verscherping van de ellipsoïde-methode tot essentiële verbeteringen zal leiden. Voor de details verwijzen we naar [18].

Perfecte grafen. Een ongerichte graaf $G = (V,E)$ heet *perfect* als voor iedere geïnduceerde deelgraaf G' van G geldt:

(67) $\qquad \omega(G') = \chi(G').$

Hierbij is $\omega(G')$ het *kliekgetal* van G' (d.w.z. de afmeting van de grootste kliek in G'), en $\chi(G')$ het *kleurgetal* van G' (d.w.z. het minimale aantal kleuren dat nodig is om de punten van G' zo te kleuren dat ieder tweetal verbonden punten verschillend gekleurd zijn). Eenvoudig is in te zien dat $\omega(G) \leq \chi(G)$ voor iedere graaf G.

Het probleem om van een willekeurige graaf het kliekgetal (of het kleurgetal) te bepalen is NP-volledig (zie Garey en Johnson [13] en Van Leeuwen [28]). Voor verschillende deelklassen van de klasse der perfecte grafen bestaan echter polynomiale algorithmen om het kliekgetal te bepalen (zoals voor bipartite grafen, lijngrafen van bipartite grafen, getrianguleerde grafen, transitief-orientieerbare grafen, en hun complementen). Het was een open vraag of voor de klasse van *alle* perfecte grafen zo'n algorithme bestaat. Deze vraag kan bevestigend beantwoord worden met behulp van de ellipsoïde-methode. Omdat, zoals Lovász [29] bewees, het complement van een perfecte graaf weer perfect is, geeft dit tegelijk een polynomiale algorithme om het onafhankelijkheidsgetal $\alpha(G)$ van een perfecte graaf G te vinden ($\alpha(G)$ = het maximale aantal paarsgewijs niet-verbonden punten in G).

Lovász [30] liet zien dat als G = (V,E) een perfecte graaf is, dan is

(68) $\omega(G) = \max\{ \sum_{i,j=1}^{n} a_{ij} \mid A = (a_{ij})_{i,j=1}^{n}$ is een positief semi-definiete matrix zo dat $\mathrm{Tr}\,A = 1$ en $a_{ij} = 0$ als i en j verschillende niet-verbonden punten zijn$\}$,

waarbij is aangenomen dat $V = \{1,\ldots,n\}$. Het bepalen van dit maximum is een speciaal geval van het optimaliseringsprobleem voor de convexe verzameling P van alle matrixen A die voldoen aan de voorwaarden vermeld in (68). Volgens het principe uit §9 is dit probleem in polynomiale tijd oplosbaar als een polynomiale scheidingsalgorithme voor P bestaat. Nu kan, gegeven een matrix $A = (a_{ij})_{i,j=1}^{n}$, eenvoudig worden nagegaan of $\mathrm{Tr}\,A = 1$ en of $a_{ij} = 0$ als i en j verschillend en niet verbonden zijn. Als aan een van deze voorwaarden niet voldaan is kan op eenvoudige wijze een scheidend hypervlak worden bepaald. Ook of A positief semi-definiet is kan worden gecontroleerd binnen polynomiale tijd, en wel als volgt. Vind een hoofddeelmatrix A' van A (d.w.z. een deelmatrix symmetrisch om de hoofddiagonaal van A) met rang A' = rang A. Een dergelijke matrix kan worden gevonden met de gebruikelijke Gauss-eliminatie. Dan is A positief semi-definiet als en alleen als A' positief definiet is. Zonder beperking der algemeenheid is $A' = (a_{ij})_{i,j=1}^{k}$. Nu is A' positief definiet als en alleen als

(69) $$\det((a_{ij})_{i,j=1}^{k'}) > 0$$

voor $k' = 1,\ldots,k$. Deze determinanten kunnen in polynomiale tijd worden uitgerekend. Bovendien, als een dezer determinanten niet positief is, dan kan hieruit een hypervlak worden gevonden dat A van P scheidt.

Het scheidings- en het optimaliseringsprobleem voor P zijn dus polynomiaal oplosbaar, en dus kan $\omega(G)$ in polynomiale tijd worden bepaald. In [18] wordt aangetoond hoe hieruit ook polynomiale algorithmen kunnen worden afgeleid om expliciet een kliek ter grootte $\omega(G)$, en een goede kleuring met $\chi(G)$ kleuren te vinden.

Submodulaire functies. Een tweede toepassing van de ellipsoïde-methode is het vinden van de minimale waarde van een zgn. submodulaire functie. Een functie f gedefinieerd op de deelverzamelingen van een eindige verzameling X heet *submodulair* als

$$(70) \qquad f(X') + f(X'') \geq f(X' \cap X'') + f(X' \cup X'')$$

voor ieder tweetal deelverzamelingen X', X'' van X. Voorbeelden van submodulaire functies zijn: (i) X is de verzameling rijen van een matrix, en $f(X')$ is de rang van de rijen in X'; (ii) X is de verzameling punten van een gerichte graaf $G = (X,E)$, $c:E \longrightarrow \mathbb{Z}_+$ is een "capaciteits"-functie, en $f(X')$ is de totale capaciteit van de kanten die X' verlaten. Weliswaar wordt de minimum waarde van deze submodulaire functies trivialerwijs door $X' = \emptyset$ aangenomen, maar interessantere problemen kunnen worden afgeleid uit deze functies. Bijvoorbeeld, gegeven een gewichtsfunctie $w:X \longrightarrow \mathbb{Q}$, kan het minimum van $f(X') - \sum_{x \in X'} w(x)$ gevonden worden (dit definieert weer een submodulaire functie). Verder, als de submodulaire functie g gedefinieerd wordt door

$$(71) \qquad g(X') = f(X' \cup \{r\})$$

voor $X' \subset X \setminus \{r,s\}$, dan is, voor voorbeeld (ii), het vinden van de minimale waarde van g equivalent met het vinden van een r-s-snede van minimale capaciteit.

We moeten voorzichtig zijn met de manier waarop de functie f is gegeven. Als f gespecificeerd wordt door een lijst van alle deelverzamelingen van X met hun waarden onder f, dan kan de minimale waarde natuurlijk eenvoudig worden gevonden door de hele lijst af te gaan. Maar vaak vereist de declaratie van f minder ruimte, bijvoorbeeld als f is als in de bovenstaande voorbeelden (i) en (ii), in welk geval de vraag naar een polynomiale algorithme minder triviaal wordt. "Polynomiaal" betekent dan: polynomiaal in $|X|$ en $\log B$, waarbij B de grootste teller of noemer van $|f(X')|$ is. Het blijkt dat het probleem van de minimale waarde van een submodulaire functie gereduceerd kan worden tot het scheidingsprobleem voor polyeders van de vorm

$$(72) \qquad P = \{y \in \mathbb{Q}^X \mid \sum_{x \in X'} y(x) \leq f(X') \text{ voor iedere } X' \subset X\}.$$

Bovendien kan het optimaliseringsprobleem voor P eenvoudig binnen polynomiale tijd worden opgelost met de zgn. "greedy" algorithme. Het principe uit §9 geeft ons dan een polynomiale methode om het minimum van f te vinden. Voor een meer gedetailleerde uitwerking wordt verwezen naar [18].

13. GEHEELTALLIGE LINEAIRE PROGRAMMERING

Geheeltallige lineaire programmeringsproblemen vragen om een geheeltallige oplossing voor (1), d.w.z. naar

(73) $\max \{cx \mid Dx \leq b,\ x \text{ geheeltallig}\}.$

Dit "ILP-probleem" blijkt moeilijker dan het LP-probleem zonder geheeltalligheidsvoorwaarden. Zo is het ILP-probleem een NP-volledig probleem, en bestaat er (nog) geen Dualiteitsstelling (een min-max relatie) voor (73). Er is daarom vooralsnog weinig hoop op een polynomiale algorithme. Er bestaan verschillende methoden om (73) te bepalen, zoals "branch-and-bound"-methoden en de methode van de "Gomory-sneden", maar hun looptijd is niet polynomiaal-begrensd (zie Garfinkel en Nemhauser [14]).

De Gomory-sneden-methode kan als volgt worden geformaliseerd. Stel een polyeder P is gegeven. Dan is

(74) $\max \{cx \mid x \in P,\ x \text{ geheeltallig}\} = \max \{cx \mid x \in P_I\},$

waarbij P_I het convex omhulsel is van de geheeltallige vectoren in P. Omdat het ILP-probleem NP-volledig is, is het optimaliseringsprobleem voor P_I kennelijk ook NP-volledig, gegeven een scheidingsalgorithme voor P. Als we nu P' als volgt definieren:

(75) $P' = \cap H_I,$

waarbij de doorsnede loopt over alle halfruimten H zo dat $P \subset H$. Merk op dat als $H = \{x \mid wx \leq d\}$, waarbij w een geheeltallige vector is waarvan de coördinaten relatief priem zijn, dan is $H_I = \{x \mid wx \leq \lfloor d \rfloor\}$, waarbij $\lfloor d \rfloor$ het gehele deel van d is. H_I ontstaat uit H door H op te schuiven tot het begrenzende hypervlak van H geheeltallige punten bevat.

Eenvoudig is in te zien dat

(76) $P_I \subset P',$

omdat de inclusie $P \subset H$ impliceert dat $P_I \subset H_I$. Nu kan bewezen worden dat P' weer een polyeder is, en dat

(77) $P_I = P^{(t)}$

voor zekere t, waarbij $P^{(t)} = P^{(t-1)}$'. Dit is de essentie van de Gomory-sneden (zie Gomory [17], Chvátal [3], Schrijver [36]).

Nu zou een antwoord op de volgende vragen een aanzet kunnen zijn tot het toepassen van de ellipsoïde-methode op geheeltallige lineaire programmering. Gegeven een optimaliseringsalgorithme voor het polytoop P, bestaat er een polynomiale scheidingsalgorithme voor P'? (polynomiaal in de looptijd van de optimaliseringsalgorithme voor P). Of is het scheidingsprobleem voor P' NP-volledig? Merk op dat dit in ieder geval een probleem uit de klasse $\mathcal{NP}$ is: men kan, gegeven een $y \notin P'$, in polynomiale tijd *bewijzen* dat $y \notin P'$. Hiertoe moet men een halfruimte H geven zo dat $P \subset H$ en $y \notin H_I$. D.w.z., men moet een geheeltallige vector w specificeren zo dat

$$wy > \lfloor \max\{wx \mid x \in P\} \rfloor. \tag{78}$$

Deze ongelijkheid kan in polynomiale tijd worden bewezen met de optimaliseringsalgorithme voor P. Hoe men, gegeven een y, een dergelijke w in polynomiale tijd kan *vinden* is de kern van het scheidingsprobleem voor P'.

14. LITERATUUR

[1] D. BARNETTE, *Path problems and extremal problems for convex polytopes*, in: "Relations between combinatorics and other parts of mathematics", Proc. Symp. Pure Math. 34, Amer. Math. Soc., Providence, R.I., 1979, pp. 25-34.

[2] R.G. BLAND, D. GOLDFARB en M.J. TODD, *The ellipsoid method: a survey*, Tech. Report 476, School of Oper. Res. and Industrial Engineering, Cornell University, Ithaca, N.Y., 1980.

[3] V. CHVÁTAL, *Edmonds polytopes and a hierarchy of combinatorial problems*, Discrete Math. 4 (1973) 305-337.

[4] G.B. DANTZIG, *Maximization of a linear functional of variables subject to linear inequalities*, in: "Activity analysis of production and allocation" (T.C. Koopmans, ed.), J. Wiley, New York, 1951, pp. 339-347.

[5] G.B. DANTZIG, *Linear programming and extensions*, Princeton Univ. Press, Princeton, N.J., 1962.

[6] G.B. DANTZIG, *Expected number of steps of the simplex method for a linear program with convexity constraints*, Tech. Report SOL 80-3, Dept. of Oper. Res., Stanford Univ., Stanford, Ca., 1980.

[7] E.W. DIJKSTRA, *A note on two problems in connexion with graphs*, Numer. Math. 1 (1959) 269-271.

[8] J. EDMONDS, *Maximum matching and a polyhedron with 0,1-vertices*, J. Nat. Bur. Standards Sect. B 69 (1965) 125-130.

[9] L.R. FORD, Jr en D.R. FULKERSON, *Maximum flow through a network*, Canad. J. Math. 8 (1956) 399-404.

[10] D.R. FULKERSON, *Packing rooted directed cuts in a weighted directed graph*, Math. Programming 6 (1974) 1-13.

[11] P. GÁCS en L. LOVÁSZ, *Khachiyan's algorithm for linear programming*, Math. Programming study 14 (1981) 61-68.

[12] D. GALE, H.W. KUHN en A.W. TUCKER, *Linear programming and the theory of games*, in: "Activity analysis of production and allocation" (T.C. Koopmans, ed.), J. Wiley, New York, 1951, pp. 317-329.

[13] M.R. GAREY en D.S. JOHNSON, *Computers and intractability: a guide to the theory of NP-completeness*, Freeman, San Francisco, 1979.

[14] R.S. GARFINKEL en G.L. NEMHAUSER, *Integer programming*, J. Wiley, New York, 1972.

[15] D. GOLDFARB en W.Y. SIT, *Worst case behavior of the steepest edge simplex method*, Discrete Applied Math. 1 (1979) 277-285.

[16] D. GOLDFARB en M.J. TODD, *Modifications and implementation of the Shor-Khachian algorithm for linear programming*, Tech. Report 406, School of Oper. Res. and Industrial Engineering, Cornell University, Ithaca, N.Y., 1980.

[17] R.E. GOMORY, *Outline of an algorithm for integer solutions to linear programs*, Bull. Amer. Math. Soc. 64 (1958) 275-278.

[18] M. GRÖTSCHEL, L. LOVÁSZ en A. SCHRIJVER, *The ellipsoid method and its consequences in combinatorial optimization*, verschijnt in Combinatorica.

[19] R. JEROSLOW, *The simplex algorithm with the pivot rule of maximizing criterion improvement*, Discrete Math. 4 (1973) 367-377.

[20] D.B. JUDIN en A.S. NEMIROVSKII, *Informational complexity and effective methods of solution for convex extremal problems*, Ekonomika i Matematicheskie Metody 12 (1976) 357-369 (vertaald: Matekon: Translations of Russian and East European Math. Economics 13 (2) Spring '77, 25-45).

[21] D.B. JUDIN en A.S. NEMIROVSKII, *Optimization methods adapting to the "significant" dimension of the problem*, Automatika i Telemekhanika 38 (1977) No. 4, 75-87 (vertaald: Automation and Remote Control 38 (1977) No 4, 513-524).

[22] R.M. KARP en C.H. PAPADIMITRIOU, *On linear characterizations of combinatorial optimization problems*, Report MIT/LCS/TM-154, Mass. Inst. of Technology, Cambridge, Mass., 1980.

[23] L.G. KHACHIAN, *A polynomial algorithm in linear programming*, Doklady Akademiia Nauk SSSR 244:5 (1979) 1093-1096 (vertaald: Soviet Math. Doklady 20:1 (1979) 191-194).

[24] V. KLEE en G.L. MINTY, *How good is the simplex algorithm?*, in: "Inequalities, III" (O. Shisha, ed.), Academic Press, New York, 1972, pp. 159-175.

[25] G.B. KOLATA, *Mathematicians amazed by Russian discovery*, Science 206 (1979) 545-546.

[26] M.K. KOZLOV, S.P. TARASOV en L.G. KHACHIAN, *Polynomial solvability of convex quadratic programming*, Doklady Akademiia Nauk SSSR **248**:5 (1979) 1049-1050 (vertaald: Soviet Math. Doklady 20:5 (1979) 1108-1111).

[28] J. van LEEUWEN, *Computers en (on-)doenlijke problemen,* dit colloquium.

[29] L. LOVÁSZ, *Normal hypergraphs and the perfect graph conjecture*, Discrete Math. 2 (1972) 253-267.

[30] L. LOVÁSZ, *On the Shannon capacity of a graph*, IEEE Trans. on Information Theory 25 (1979) 1-7.

[31] E.H. McGALL, *A study of the Khachian algorithm for real-world linear programming*, rapport Univac, 1980.

[32] J. von NEUMANN, *On a maximization problem*, manuscript, Institute for Advanced Studies, Princeton, N.J., 1947.

[33] M.W. PADBERG en M.R. RAO, *The Russian method for linear inequalities*, rapport Graduate School of Business Administration, New York University, New York, 1979.

[34] M.W. PADBERG en M.R. RAO, *Minimum cut-sets and* b*-matchings*, rapport Graduate School of Business Administration, New York University, New York, 1979.

[35] M.W. PADBERG en M.R. RAO, *The Russian method and integer programming*, rapport Graduate School of Business Administration, New York University, New York, 1980.

[36] A. SCHRIJVER, *On cutting planes,* Annals of Discrete Math. 9 (1980) 291-296.

[37] N.Z. SHOR, *Generalized gradient methods of nondifferentiable function minimization and their application to problems of mathematical programming*, Ekonomika i Matematicheskie Metody 12 (1976) 337-356.

[38] P. WOLFE, *A bibliography for the ellipsoid method*, rapport IBM Research Center, Yorktown Heights, N.Y., 1980.