

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

ZW 1955-014

Het extreem van een lineaire functie onder
lineaire bijvoorwaarden

Voordracht in de serie
"Actualiteiten"

J. Kriens



1955

Het extreem van een lineaire functie onder lineaire bijvoorwaarden¹⁾

door

J. Kriens

Voordracht in de serie "Actualiteiten".

29 October 1955

In de praktijk stuit men vaak op problemen, waarin één van de extremen van een functie bepaald moet worden, terwijl de variabelen moeten voldoen aan een aantal vergelijkingen en slechts niet negatieve waarden mogen aannemen. De klassieke wijze van oplossen met b.v. Lagrange-multiplicatoren eist reeds bij problemen van matige omvang zeer veel tijd. Mede hierom heeft men gezocht naar methoden, waarbij de oplossing sneller bereikt kan worden.

Voor het geval de functie lineair is en ook de bijvoorwaarden lineair zijn, is hiervoor een methode ontwikkeld door G.B. DANTZIG [1], welke later is aangevuld en verbeterd, o.a. door A. CHARNES [2], [3]. Het principe waarop deze methode berust zullen wij hier beschrijven.

Probleemstelling

Gevraagd de extremen te bepalen van de functie

$$z_0 = f(\lambda) = \sum_{h=1}^n c_h \lambda_h \quad (1)$$

onder de bijvoorwaarden

$$\begin{array}{ccccccc} a_{11} \lambda_1 + \dots + a_{1n} \lambda_n & = & b_1 & & & & \\ \vdots & & \vdots & & & & \\ a_{m1} \lambda_1 + \dots + a_{mn} \lambda_n & = & b_m & & & & \end{array} \quad (2)$$

$$\lambda_h \geq 0 \quad (h=1, \dots, n) \quad (3)$$

1) Deze voordracht is gebaseerd op de lezingen III, IV en V uit [3].

The Mathematical Centre at Amsterdam, founded the 11th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for Pure Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

Aangezien maximum en minimum op analoge wijze verkregen worden, zullen wij ons hier beperken tot het geval, waarin het maximum bepaald moet worden.

Wij beschouwen de coëfficiënten van de variabelen en de rechterleden van (2) als vectoren in de m -dimensionale ruimte W ; (2) gaat dan over in

$$\sum_{h=1}^n \lambda_h p_h = p_0 \quad (4)$$

waardoor het probleem luidt:

Bepaal die λ -grepen met niet-negatieve coördinaten uit de n -dimensionale ruimte U , welke voldoen aan (4) en (1) maximaliseren.²⁾

Wij nemen aan dat de rang van de matrix $[p_1, \dots, p_n]$ maximaal is. Dit is geen belangrijke beperking, want in het geval dat deze rang lager is en ook de rang van de matrix $[p_0, p_1, \dots, p_n]$ lager is, zijn de vergelijkingen (2) lineair afhankelijk, terwijl in het andere geval de vergelijkingen (2) strijdig zijn.

2. Eigenschappen van de verzameling van toegelaten λ 's

In het algemeen zal er meer dan één punt λ uit U met niet-negatieve coördinaten voldoen aan (4), dus toegelaten zijn. Voor de verzameling Λ van toegelaten punten geldt de volgende stelling.

Stelling 1:

De verzameling Λ van punten met niet-negatieve coördinaten uit U , welke voldoen aan (4) is convex.

Bewijs: De lege verzameling en een verzameling bestaande uit één punt zijn convex.

Voldoen aan (4) de punten λ^1 en λ^2 , dan geldt voor

$$\lambda = \gamma \lambda^1 + (1-\gamma) \lambda^2 \quad (0 \leq \gamma \leq 1),$$

$$\begin{aligned} \sum_{h=1}^n \lambda_h p_h &= \sum_{h=1}^n (\gamma \lambda_h^1 + (1-\gamma) \lambda_h^2) p_h = \gamma \sum_{h=1}^n \lambda_h^1 p_h + (1-\gamma) \sum_{h=1}^n \lambda_h^2 p_h = \\ &= \gamma p_0 + (1-\gamma) p_0 = p_0, \end{aligned}$$

dus $\lambda \in \Lambda$.

2) Een punt uit de ruimte U wordt aangegeven door λ , eventueel met bovenindex, wanneer verschillende punten moeten worden onderscheiden; benedenindices representeren de coördinaten van de punten uit U .

Definitie 1:

Een convexe combinatie van k punten $\lambda_1, \dots, \lambda_k$ uit een verzameling is een combinatie van de vorm

$$\sum_{i=1}^k \gamma_i \lambda_i,$$

waarin

$$0 \leq \gamma_i < 1 \quad (i = 1, \dots, k)$$

en

$$\sum_{i=1}^k \gamma_i = 1.$$

Definitie 2:

Een extreem punt van een verzameling is een punt dat niet geschreven kan worden als convexe combinatie van andere punten uit de verzameling.

Stelling 2:

Het punt λ is dan en slechts dan een extreem punt van Λ , wanneer de $\lambda_j > 0$, coëfficiënten zijn van vectoren P_j , die een lineair onafhankelijk stelsel vormen.

Bewijs:

a) voldoende voorwaarde.

Gegeven: $\lambda_1, \dots, \lambda_k$ ($k \leq m$) zijn > 0 en $P_1, \dots, P_k$ vormen een lineair onafhankelijk stelsel vectoren. Indien $k < m$ is, vullen wij dit stelsel aan met $P_{k+1}, \dots, P_m$ zodat dan $P_1, \dots, P_m$ een basis vormen van W . Er geldt dus

$$\sum_{i=1}^m \lambda_i P_i = P_0$$

en wij bewijzen

$$\lambda = (\lambda_1, \dots, \lambda_m, 0, \dots, 0),$$

waarin

$$\lambda_i \geq 0 \quad (i = 1, \dots, m) \tag{5}$$

is extreem punt van Λ .

Stel er bestaan twee verschillende punten λ^1 en λ^2 zodat dan

$$\lambda = \gamma \lambda^1 + (1 - \gamma) \lambda^2 \quad (0 < \gamma < 1)$$

Uit (5), $\gamma > 0$ en $1 - \gamma > 0$ volgt, dat ook λ^1 en λ^2 zijn van de gedaante

$$\begin{aligned} \lambda_1^1, \dots, \lambda_m^1, 0, \dots, 0 \\ \lambda_1^2, \dots, \lambda_m^2, 0, \dots, 0 \end{aligned}$$

met $\lambda_i^\ell \geq 0 \quad (\ell=1,2 ; i=1,\dots,m).$

Dus is $\sum_{i=1}^m \lambda_i^\ell p_i = p_0 \quad (\ell=1,2)$

Omdat $p_1, \dots, p_m$ een basis vormen van W , geldt

$$\lambda_i^1 = \lambda_i^2 = \lambda_i \quad (i=1,\dots,m)$$

en dus

$$\lambda = \lambda^1 = \lambda^2,$$

wat in strijd is met de gemaakte onderstelling

$$\lambda^1 \neq \lambda^2;$$

m.a.w. λ is extreem punt van Λ .

b) noodzakelijke voorwaarde.

Gegeven: λ is een extreem punt van Λ ; de coëfficiënten > 0 zijn $\lambda_1, \dots, \lambda_p$, zodat voldaan is aan

$$\sum_{i=1}^p \lambda_i p_i = p_0 \quad (6)$$

Stel $p_1, \dots, p_p$ zijn lineair afhankelijk; dan bestaan getallen $d_1, \dots, d_p$, niet alle gelijk aan nul, zodat dan

$$\sum_{i=1}^p d_i p_i = \vec{0} \quad (7)$$

waarin $\vec{0}$ de m -dimensionale nulvector is.

Uit (6) en (7) volgt voor iedere positieve k

$$p_0 = \sum_{i=1}^p \lambda_i p_i \pm k \sum_{i=1}^p d_i p_i = \sum_{i=1}^p (\lambda_i \pm k d_i) p_i$$

Wegens $\lambda_i > 0$ kan k zo gekozen worden, dat alle getallen $\lambda_i \pm k d_i$ positief zijn, waaruit volgt dat

$$\lambda_1 = (\lambda_1 + k d_1, \dots, \lambda_p + k d_p, 0, \dots, 0)$$

en

$$\lambda_2 = (\lambda_1 - k d_1, \dots, \lambda_p - k d_p, 0, \dots, 0)$$

beide elementen zijn van Λ .

Bovendien is

$$\lambda = \frac{1}{2}(\lambda^1 + \lambda^2)$$

en dus is λ een convexe combinatie van andere punten uit Λ , wat in strijd is met het gegeven; $P_1, \dots, P_p$ vormen dus een lineair onafhankelijk stelsel vectoren.

Gevolgtrekking 1:

Ieder extreem punt van de verzameling Λ heeft hoogstens m coördinaten $\neq 0$.

Gevolgtrekking 2:

Het aantal extreme punten van Λ is eindig, omdat men uit $P_1, \dots, P_n$ slechts eindig veel verschillende stelsels van lineair onafhankelijke vectoren kan vormen. Bovendien is Λ afgesloten, omdat in alle voorwaarden $\leq$ staat en nergens $<$. Verder kan men bewijzen, dat een convexe afgesloten verzameling met eindig veel extreme punten een convexe polyeder is; de extreme punten zijn de hoekpunten.

Stelling 3:

Een lineaire functie $f(u)$, gedefinieerd op een convexe polyeder K , bereikt de extremen in extreme punten van K . Wanneer een extreem in meer dan één punt bereikt wordt, heeft $f(u)$ dezelfde waarde op de gehele convexe verzameling, voortgebracht door deze punten.

Bewijs:

a) Het maximum wordt in een hoekpunt bereikt.

Laat $f(u)$ het maximum bereiken voor $u=x$. Wanneer x een extreem punt is, is het gestelde bewezen.

Stel x is geen extreem punt. Wij kunnen x schrijven als convexe combinatie van extreme punten $y^1, \dots, y^r$ van K :

$$x = \sum_{i=1}^r \gamma_i y^i,$$

waarin $0 \leq \gamma_i < 1 \quad (i=1, \dots, r)$

en $\sum_{i=1}^r \gamma_i = 1$.

Omdat f lineair is, geldt

$$f(x) = f\left(\sum_{i=1}^r \gamma_i y^i\right) = \sum_{i=1}^r \gamma_i f(y^i).$$

Zij $f(y^k)$ gelijk aan het maximum van de waarden $f(y^i) \quad (i=1, \dots, r)$ dan is wegens $\gamma_i \geq 0 \quad (i=1, \dots, r)$

$$f(x) = \sum_{i=1}^r \gamma_i f(y^i) \leq f(y^k)$$

en omdat $f(x)$ het maximum is geldt

$$f(x) = f(y^k),$$

zodat het maximum ook in een extreem punt wordt bereikt.

b) Wanneer het maximum M bereikt wordt in de punten $x^1, \dots, x^\ell$ dan wordt het bereikt op de gehele convexe verzameling, voortgebracht door deze punten.

Ieder punt uit de door $x^1, \dots, x^\ell$ voortgebrachte convexe verzameling kan geschreven worden in de vorm

$$x = \sum_{i=1}^{\ell} \gamma_i x^i,$$

waarin $\gamma_i \geq 0 \quad (i = 1, \dots, \ell)$

$$\sum_{i=1}^{\ell} \gamma_i = 1$$

Dan is

$$f(x) = f\left(\sum_{i=1}^{\ell} \gamma_i x^i\right) = \sum_{i=1}^{\ell} \gamma_i f(x_i) = \sum_{i=1}^{\ell} \gamma_i \cdot M = M,$$

wat te bewijzen was.

Voor een minimum lopen de onder a) en b) gegeven bewijzen analoog.

Gevolgtrekking 3:

Het maximum van z_0 onder de voorwaarden (4) kan worden bepaald door de hoekpunten van de verzameling Λ af te tasten.

Wanneer men de waarde van z_0 kent in één hoekpunt, tracht men daarom een nieuw hoekpunt te vinden, waarin z_0 een grotere waarde heeft. Wij laten nu eerst zien hoe men vanuit één hoekpunt een ander kan bereiken.

3. Methode waarmee men de coördinaten van een nieuw hoekpunt van kan bepalen, wanneer de coördinaten van een ander gegeven zijn.

Stel λ is het hoekpunt waarvan de coördinaten bekend zijn; wij nemen aan, dat m coördinaten van λ een positieve waarde bezitten. Laten de coördinaten $\lambda_1, \dots, \lambda_m$ positief zijn, dan is voldaan aan de betrekking

$$\sum_{i=1}^m \lambda_i p_i = p_0 \quad (8)$$

Wij noemen λ een basisoplossing van (4), omdat $p_1, \dots, p_m$ een basis vormen van W .

De vectoren $p_1, \dots, p_n$ kunnen in deze basis worden uitgedrukt

$$p_j = \sum_{i=1}^m x_{ij} p_i \quad (j = 1, \dots, n) \quad (9)$$

Stel dat één of meer van de coëfficiënten x_{ik} ($k > m$) positief zijn. Wij zoeken dan naar een oplossing van (4), waarin de coëfficiënt van $p_k > 0$ is; opdat het nieuwe punt λ' een extreem punt van Λ zal zijn, moet in ieder geval één van de coëfficiënten van $p_1, \dots, p_m$ in de nieuwe oplossing van (4) nul zijn (zie gevolgtrekking 1).

Wij geven λ'_k voorlopig de nog nader te specificeren waarde θ en vervormen (8) als volgt:

$$\sum_{i=1}^m \lambda_i p_i - \theta p_k + \theta p_k = p_0,$$

of met (9)

$$\sum_{i=1}^m \lambda_i p_i - \theta \sum_{i=1}^m x_{ik} p_i + \theta p_k = p_0,$$

of

$$\sum_{i=1}^m (\lambda_i - \theta x_{ik}) p_i + \theta p_k = p_0. \quad (10)$$

De coëfficiënten van $p_1, \dots, p_m, p_k$ vormen alleen dan een oplossing indien ze alle ≥ 0 zijn, dus als geldt

$$\theta \geq 0 \quad (11)$$

en

$$\lambda_i - \theta x_{ik} \geq 0 \quad (i = 1, \dots, m) \quad (12)$$

Voor $\theta \geq 0$ en $x_{ik} \leq 0$ is steeds aan (12) voldaan; voor de $x_{ik} > 0$ moet gelden

$$\theta \leq \frac{\lambda_i}{x_{ik}}. \quad (13)$$

Laat het minimum van de quotiënten $\frac{\lambda_i}{x_{ik}}$ alleen bereikt worden voor $i = r$, dan kiezen wij

$$\theta = \frac{\lambda_r}{x_{rk}}$$

De coëfficiënt van p_r in (10) wordt dan nul.

De vectoren $P_1, \dots, P_{r-1}, P_{r+1}, \dots, P_m$ en P_k vormen weer een lineair onafhankelijk stelsel.

Was het tegendeel waar, dan was

$$P_k = \sum_{\substack{i=1 \\ i \neq r}}^m d_i P_i, \quad (14)$$

waarin niet alle $d_i = 0$ zijn.

Hieruit volgt met (9)

$$\sum_{i=1}^m x_{ik} P_i - \sum_{\substack{i=1 \\ i \neq r}}^m d_i P_i = x_{rk} P_r + \sum_{\substack{i=1 \\ i \neq r}}^m (x_{ik} - d_i) P_i = 0.$$

$P_1, \dots, P_m$ vormen echter een basis en dus moet $x_{rk} = 0$ zijn, wat in strijd is met de gemaakte onderstelling $x_{rk} > 0$.

De vectoren $P_1, \dots, P_{r-1}, P_{r+1}, \dots, P_m$ en P_k vormen dus een lineair onafhankelijk stelsel en de coëfficiënten van (10) met

$$\theta = \frac{\lambda_r}{x_{rk}}$$

zijn de coördinaten van een nieuw extreem punt λ' van Λ .

Op deze wijze is het dus mogelijk van één hoekpunt van Λ over te springen naar een ander. Er zijn echter twee beperkingen gemaakt: van de oorspronkelijke greep $(\lambda_1, \dots, \lambda_n)$ werd aangenomen dat m coëfficiënten > 0 waren en bovendien onderstellen wij dat het minimum van de quotiënten $\frac{\lambda_i}{x_{ik}}$ ($x_{ik} > 0$) bereikt werd voor precies één waarde van i . Dit betekent dat ook de nieuwe oplossing een basisoplossing wordt. Wanneer hieraan niet is voldaan houdt men minder dan m vectoren P_i over, in welk geval men het probleem ontaard noemt. Het is dan mogelijk P_0 uit te drukken in minder dan m lineair onafhankelijke vectoren uit $P_1, \dots, P_n$.

4. Constructie van een optimale oplossing.

Teneinde een extreem punt te vinden, waarvoor (1) het maximum bereikt, gaan wij na, onder welke omstandigheden z_0 toeneemt, wanneer men van één hoekpunt overgaat op een ander. Hiertoe vergelijken wij bij de hierboven beschreven constructie de waarden van $z_0 = f(\lambda)$ en $z'_0 = f(\lambda')$:

$$z'_0 = \sum_{i=1}^m (\lambda_i - \theta x_{ik}) c_i + \theta c_k = \sum_{i=1}^m \lambda_i c_i + \theta (c_k - \sum_{i=1}^m x_{ik} c_i)$$

of

$$z'_0 = z_0 + \theta (c_k - z_k), \quad (15)$$

waarin

$$z_k = \sum_{i=1}^m x_{ik} c_i . \quad (16)$$

Indien $c_k - z_k > 0$ is, leidt het overgaan van het oude naar het nieuwe hoekpunt, d.w.z. het vervangen van P_k in de basis door P_k , dus tot een grotere waarde van de functie $f(\lambda)$.

In ieder stadium van de berekening doet zich nu één van de volgende, elkaar uitsluitende gevallen voor.

Wij nemen meer aan, dat λ een basisoplossing is, dat de coördinaten van λ groter dan nul zijn: $\lambda_1, \dots, \lambda_m$ en dat dus geldt

$$P_o = \sum_{i=1}^m \lambda_i P_i . \quad (17)$$

I. Er bestaat een waarde van j , waarvoor $c_j - z_j > 0$ is³⁾, terwijl alle waarden x_{ij} uit (9) ($i=1, \dots, m$) ≤ 0 zijn.

Uit (12) volgt dat voor iedere positieve θ alle coëfficiënten in (10) > 0 zijn; θ kan dan willekeurig groot gekozen worden en het maximum van (1) onder de voorwaarden (3) en (4) is oneindig.

II. Voor iedere j is $c_j - z_j \leq 0$.

In dit geval is het maximum van z_o onder de voorwaarden (3) en (4) bereikt.

Bewijs:

Zij $\lambda^1 = (\lambda_1^1, \dots, \lambda_n^1)$ een ander punt uit Λ , dan is

$$P_o = \sum_{h=1}^n \lambda_h^1 P_h ; \quad (18)$$

de bijbehorende waarde van de functie is

$$z_o^* = \sum_{h=1}^n \lambda_h^1 c_h . \quad (19)$$

Met behulp van de ongelijkheden $c_j - z_j \leq 0$ en (16) volgt uit (19)

$$z_o^* = \sum_{h=1}^n \lambda_h^1 c_h \leq \sum_{h=1}^n \lambda_h^1 z_h = \sum_{h=1}^n \lambda_h^1 \left(\sum_{i=1}^m x_{ih} c_i \right) = \sum_{i=1}^m \left(\sum_{h=1}^n \lambda_h^1 x_{ih} \right) c_i . \quad (20)$$

Verder vinden wij met (9) en (18)

$$P_o = \sum_{h=1}^n \lambda_h^1 P_h = \sum_{h=1}^n \lambda_h^1 \left(\sum_{i=1}^m x_{ih} P_i \right) = \sum_{i=1}^m \left(\sum_{h=1}^n \lambda_h^1 x_{ih} \right) P_i . \quad (21)$$

3) Als een dergelijke waarde bestaat, is zeker $j > m$, want voor $j \leq m$ geldt: $c_j - z_j = c_j - \sum_{i=1}^m x_{ij} c_i = c_j - c_j = 0$.

Omdat $P_1, \dots, P_m$ een basis vormen van W volgt uit (17) en (21)

$$\lambda_i = \sum_{h=1}^n \lambda_h^1 x_{ih},$$

wat, gesubstitueerd in (20), oplevert

$$z_o^* \leq \sum_{i=1}^m c_i \lambda_i = z_o,$$

zodat voor ieder element λ^1 uit Λ geldt

$$f(\lambda^1) = z_o^* \leq z_o = f(\lambda).$$

III. Er bestaan één of meer waarden van j , waarvoor $c_j - z_j > 0$ is en voor al deze j geldt dat minstens één van de elementen $x_{ij} > 0$ is. In dit geval leidt het opnemen in de basis van een vector P_k , waarvoor $c_k - z_k > 0$ is tot een grotere eindige waarde van de functie.

De nieuwe vectoren vormen weer een basis van W en men kan na het berekenen van de formules (9) en (16) wederom nagaan in welk van de gevallen I, II of III men verkeert.

Er van uitgaande, dat men een basisoplossing kent, is dus een methode aangegeven, waarmee men het maximum kan bepalen in de onderstelling, dat men voortdurend m coëfficiënten $\lambda_i > 0$ behoudt. Aangezien de waarde van de functie bij iedere stap toeneemt, is het niet mogelijk, dat men terugkeert op een basisoplossing, die men reeds eerder heeft gevonden. Daar bovendien uit n vectoren P_i hoogstens $\binom{n}{m}$ verschillende stelsels van m lineair onafhankelijke vectoren gevormd kunnen worden, is het aantal basisoplossingen eindig en het proces dus in een eindig aantal stappen voltooid.

Opmerking

Bij iedere stap wordt de oplossing, het punt λ , bepaald door $n-m$ vergelijkingen van het type

$$\lambda_j = 0 \tag{22}$$

en m vergelijkingen van het type

$$\sum_{h=1}^n a_{ih} \lambda_h = b_i, \tag{23}$$

welke samen n lineair onafhankelijke hypervlakken vormen en dus, één en niet meer dan één punt gemeenschappelijk hebben.

Vanuit λ lopen rechten die behoren tot de begrenzing van Λ en die de snijlijnen vormen van $n-1$ van de vlakken (22) en (23).

Strekt Λ zich in de richting van één van deze rechten niet oneindig ver uit, dan leidt deze rechte tot een ander extreem punt λ' van Λ . Het punt λ' ligt in $n-1$ van de vlakken (22) en (23) en wordt een naburig hoekpunt van λ genoemd.

Bij het berekenen van een nieuw extreem punt wordt één van de vlakken (22) vervangen door een vlak $\lambda=0$, dat niet in (22) is opgenomen. Men beweegt zich dus voortdurend van één hoekpunt van Λ naar één van de naburige hoekpunten.

5. Het bepalen van de oplossing wanneer ontaarding optreedt.

Ontaarding van het probleem kan optreden, wanneer P_0 een lineaire combinatie is van minder dan m van de vectoren P_i ($i=1, \dots, n$), d.w.z. wanneer P_0 ligt in een ruimte die door minder dan m van de vectoren P_i wordt opgespannen.

De behandeling van dit geval komt er nu op neer, dat men het probleem vervanet door een ander, waarin de P_0 -vector willekeurig weinig verschoven is tot de vector $P_0(\epsilon)$ en wel zodanig, dat $P_0(\epsilon)$ niet geschreven kan worden als lineaire combinatie van minder dan m lineair onafhankelijke vectoren uit $P_1, \dots, P_n$. De oplossing van dit nieuwe probleem kan worden verkregen op de boven geschetste wijze. en het is dan niet moeilijk om in te zien, dat de zo verkregen oplossing ook de oplossing vormt van het oorspronkelijke probleem. Voor een uitvoerige bespreking van deze methoden verwijzen wij naar [2] en [3].

Literatuur

- [1] G.B. DANTZIG, Maximalization of a linear function of variables subject to linear inequalities, hoofdstuk XXI in Activity Analysis of production and allocation, onder redactie van Tj.C. Koopmans, John Wiley and Sons (1951).
- [2] A. CHARNES, Optimality and degeneracy in linear programming, Econometrica 20 (1952) p. 160-170.
- [3] A. CHARNES, W.W. COOPER and A. HENDERSON, An introduction to linear programming, John Wiley and Sons (1953).