STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING ZUIVERE WISKUNDE

ZW 1966-004

Voordracht in de serie

"Elementaire onderwerpen vanuit hoger standpunt belicht"

door

Prof. J. Kriens

23 maart 1966

De stelling van KUHN en TUCKER als generalisatie van klassieke optimalisatiemethoden

Onder klassieke optimalisatiemethoden verstaan wij de uit de analyse bekende methoden om maxima en minima te berekenen. Wij brengen de volgende stellingen in herinnering.

Stelling 1. Zij $f(x)$ een reële n keer differentieerbare functie ($n \geq 2$),
$f^{(\nu)}(\bar{x}) = 0$ voor $1 \leq \nu \leq n-1$,
$f^{(n)}(\bar{x}) \neq 0$,
dan volgt uit
n even , $f^{(n)}(\bar{x}) > 0$: $f(x)$ heeft een relatief minimum in $x = \bar{x}$,
n even , $f^{(n)}(\bar{x}) < 0$: $f(x)$ heeft een relatief maximum in $x = \bar{x}$,
n oneven, $f^{(n)}(\bar{x}) > 0$: $f(x)$ stijgt in $x = \bar{x}$,
n oneven, $f^{(n)}(\bar{x}) < 0$: $f(x)$ daalt in $x = \bar{x}$.

Stelling 2.[1] Zijn $f(X)$ en $g_i(X)$ $(i = 1,\ldots,m)$ reële differentieerbare functies van n variabelen $x_1, \ldots, x_n$,
zijn $b_1, \ldots, b_m$ constanten,

$$\text{zij } F(X,Y) = f(X) + \sum_{i=1}^{m} y_i \left[b_i - g_i(X) \right], \qquad (1)$$

zij $\mathcal{V}$ de verzameling van vektoren X, waarvoor
$g_i(X) = b_i$ $(i = 1,\ldots,m)$,
zij $\mathcal{G}$ de matrix van de 1e partiële afgeleiden van de functies g_i naar de variabelen x_j, dus $\mathcal{G} = \left(\frac{\partial g_i}{\partial x_j} \right)$,
dan kan, wanneer rang $(\mathcal{G}) = m$ in $X = \overline{X}$, de functie $f(X)$ op $\mathcal{V}$ slechts dan een extreme waarde bezitten in $X = \overline{X}$, wanneer er een vektor $Y = \overline{Y}$ bestaat, zodanig dat

$$\frac{\partial F}{\partial x_j} (\overline{X},\overline{Y}) = 0 \qquad (j = 1,\ldots,n) \qquad (2)$$

$$\frac{\partial F}{\partial y_i} (\overline{X},\overline{Y}) = 0 \qquad (i = 1,\ldots,m). \qquad (3)$$

Stelling 1 kan bewezen worden met behulp van de formule van TAYLOR; er bestaan generalisaties van, die uitspraken doen omtrent extreme waarden van functies van n variabelen.
Stelling 2 (LAGRANGE) wordt bewezen door gebruik te maken van stellingen over impliciet gegeven functies. De stelling kan uitgebreid worden tot a) een stelling die ook geldt voor het geval waarin rang $(\mathcal{G})$ in $X = \overline{X}$ kleiner is dan m en b) een stelling die niet alleen nodige maar ook voldoende voorwaarden voor een extreme waarde in $X = \overline{X}$ geeft. Zie b.v. C. CARATHÉODORY [1] of G. HADLEY [2].
In stelling 1 en de generalisatie daarvan kunnen alle variabelen vrij variëren over de reële as. In stelling 2 is deze vrijheid beperkt doordat de functie $f(X)$ niet onderzocht wordt op de gehele ruimte E_n, doch

---

1) Skalairen worden aangegeven met x, y, a, b, ..., vektoren met X, Y, A, B, ..., matrices en verzamelingen met $\mathcal{A}, \mathcal{B}, \mathcal{V}, \ldots$.

slechts op een deelverzameling $\mathcal{V}$ daarvan, bepaald door de vergelijkingen

$$g_i(X) = b_i \quad (i = 1,\ldots,m). \tag{4}$$

Men kan zich afvragen of het ook mogelijk is nodige en/of voldoende voorwaarden te geven voor het geval waarin de verzameling $\mathcal{V}$ van toegelaten vektoren X bepaald wordt door voorwaarden in de vorm van ongelijkheden

$$g_i(X) \leq b_i \quad (i = 1,\ldots,m). \tag{5}$$

De stelling van KUHN en TUCKER [3] geeft een antwoord op deze vraag wanneer

a) de functies $g_i(X)$ differentieerbaar zijn en convex,

b) de functie f(X) differentieerbaar is en
- b1) convex in het minimaliseerprobleem,
- b2) concaaf in het maximaliseerprobleem.

<u>Definitie 1</u>. Een <u>verzameling</u> $\mathcal{V}$ in E_n is <u>convex</u>, als voor ieder element $A \varepsilon \mathcal{V}$, ieder element $B \varepsilon \mathcal{V}$ en iedere λ, $0 < \lambda < 1$, geldt

$$(1 - \lambda)A + \lambda B \varepsilon \mathcal{V}. \tag{6}$$

<u>Definitie 2</u>. Een <u>functie</u> f(X), gedefiniëerd op de convexe verzameling $\mathcal{V}$ in E_n is <u>convex</u> als voor ieder element $A \varepsilon \mathcal{V}$, ieder element $B \varepsilon \mathcal{V}$, en iedere λ, $0 < \lambda < 1$, geldt

$$f((1 - \lambda)A + \lambda B) \leq (1 - \lambda)f(A) + \lambda f(B). \tag{7}$$

<u>Definitie 3</u>. Een <u>functie</u> f(X), gedefiniëerd op de convexe verzameling $\mathcal{V}$ in E_n is <u>concaaf</u> als voor ieder element $A \varepsilon \mathcal{V}$, ieder element $B \varepsilon \mathcal{V}$ en iedere λ, $0 < \lambda < 1$, geldt

$$f((1 - \lambda)A + \lambda B) \geq (1 - \lambda)f(A) + \lambda f(B). \tag{8}$$

<u>Stelling 3</u>. Zij f(X) een concave, continu differentieerbare functie op het inwendige van de convexe verzameling $\mathcal{V}$ in E_n,

zij $\nabla f(X)$ de gradient van $f(X)$,
dan geldt voor iedere A uit het inwendige van $\mathcal{V}$ en iedere $B \varepsilon \mathcal{V}$

$$f(B) \leq f(A) + (B - A)'\nabla f(A). \tag{9}$$

<u>Bewijs</u>. Uit (8) volgt voor $0 < \lambda < 1$

$$f(B) \leq f(A) + \frac{f((1-\lambda)A + \lambda B) - f(A)}{\lambda} . \tag{10}$$

Toepassing van de formule van TAYLOR op $f((1-\lambda)A + \lambda B) = f(A + \lambda(B-A))$ leidt tot

$$f(A + \lambda(B-A)) = f(A) + \lambda(B-A)'\nabla f(A+\theta\lambda(B-A)) \quad \text{met } 0 \leq \theta \leq 1.$$

Substitutie hiervan in (10) geeft:

$$f(B) \leq f(A) + (B-A)'\nabla f(A+\theta\lambda(B-A)).$$

Laten wij hierin λ naderen tot 0, dan wordt (9) gevonden.

Wij bekijken vervolgens het zogenaamde <u>concave programmeringsprobleem</u>:
maximaliseer $f(X)$
onder de voorwaarden $g_i(X) \leq b_i \quad (i = 1,\dots,m)$ (5)

$$X \geq 0, \tag{11}$$

waarin de functies $g_i(X)$ convex zijn en continu differentieerbaar en de functie $f(X)$ concaaf en continu differentieerbaar voor alle niet-negatieve X.

<u>Stelling 4</u>. Zij $f(X)$ concaaf en continu differentieerbaar voor $X \geq 0$,
zijn de functies $g_i(X)$ $(i = 1,\dots,m)$ convex en continu differentieerbaar voor $X \geq 0$,
zijn $b_1, \dots, b_m$ constanten,
zij $F(X,Y) = f(X) + \sum_{i=1}^{m} y_i \left[b_i - g_i(X)\right]$,
zij $\mathcal{V}$ de verzameling van vektoren X waarvoor $g_i(X) \leq b_i$ $(i = 1,\dots,m)$
en $X \geq 0$,

zij $\overline{Y}$ een niet-negatieve vektor en $\overline{X}\varepsilon\mathcal{V}$, zodanig dat

$$\nabla f(\overline{X}) - \sum_{i=1}^{m} \overline{y}_i \nabla g_i(\overline{X}) \leq 0 \tag{12}$$

$$\left[\nabla f(\overline{X}) - \sum_{i=1}^{m} \overline{y}_i \nabla g_i(\overline{X})\right]'\overline{X} = 0 \tag{13}$$

$$\sum_{i=1}^{m} \left[b_i - g_i(\overline{X})\right] \overline{y}_i = 0 \tag{14}$$

dan bereikt f(X) de maximale waarde op $\mathcal{V}$ in $X = \overline{X}$.

Bewijs. Uit $g_i(X)$ convex volgt $-g_i(X)$ concaaf. Toepassing van formule (9) op f(X) resp. $-g_i(X)$ geeft

$$f(X) \leq f(\overline{X}) + (X - \overline{X})'\nabla f(\overline{X}) \tag{15}$$

$$-g_i(X) \leq -g_i(\overline{X}) - (X - \overline{X})'\nabla g_i(\overline{X}). \tag{16}$$

Voor een willekeurige $X\varepsilon\mathcal{V}$ blijkt met behulp van (12) t/m (16):

$$f(X) \leq f(X) + \sum_{i=1}^{m} \overline{y}_i\left[b_i - g_i(X)\right] \leq$$

$$\leq f(\overline{X}) + (X-\overline{X})'\nabla f(\overline{X}) + \sum_{i=1}^{m} \overline{y}_i\left[b_i - g_i(\overline{X})\right] - \sum_{i=1}^{m} \overline{y}_i(X-\overline{X})'\nabla g_i(\overline{X}) =$$

$$= f(\overline{X}) + X'\left[\nabla f(\overline{X}) - \sum_{i=1}^{m} \overline{y}_i\nabla g_i(\overline{X})\right] - \overline{X}'\left[\nabla f(\overline{X}) - \sum_{i=1}^{m} \overline{y}_i\nabla g_i(\overline{X})\right] \leq f(\overline{X});$$

m.a.w. $f(\overline{X})$ is het absolute maximum van f(X) op $\mathcal{V}$.

Stelling 5 houdt dat deel van de stelling van KUHN en TUCKER in, dat zowel gemakkelijk is te formuleren als eenvoudig te bewijzen. Het tweede deel van de stelling leert dat de voorwaarden (12) t/m (14) ook noodzakelijk zijn voor een maximum mits de door (5) en (11) gedefiniëerde verzameling $\mathcal{V}$ voldoet aan bepaalde regelmatigheidseisen, de zogenaamde "constraint qualification". Deze regelmatigheidseisen kunnen zowel meetkundig worden geïnterpreteerd als in de vorm van voorwaarden, vergelijkbaar met de voorwaarde rang $(\mathcal{J}) = m$ uit stelling 2 en de existentie van bepaalde partiële differentiaalquotiënten.

Specialiseert men de stelling van KUHN en TUCKER, dan kunnen daaruit afgeleid worden:

1. nodige en voldoende voorwaarden voor een maximum in een lineair programmeringsprobleem,
2. nodige en voldoende voorwaarden voor een maximum in een kwadratisch programmeringsprobleem,
3. stelling 2 en de uitbreiding daarvan tot voldoende voorwaarden.

Literatuur

[1] C. CARATHEODORY, Variationsrechnung und Partielle Differentialgleichungen Erster Ordnung, Teubner (1935), hoofdstuk XI.

[2] G. HADLEY, Nonlinear and Dynamic Programming, Addison-Wesley (1964), hoofdstukken 3 en 6.

[3] H.W. KUHN and A.W. TUCKER, Non-Linear Programming, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, (1950), 481-492.